Question 1. (a) Briefly explain the function of following proteins/terms :

(i) Abortive transcription

- Abortive transcription is a phenomenon that occurs during the initial stages of transcription, primarily in bacteria.
- RNA polymerase often synthesizes short RNA transcripts (typically less than 10 nucleotides long) that are released from the template before the polymerase enters the elongation phase.
- This process is considered a "frustrated" attempt at transcription, as the polymerase repeatedly initiates and releases short RNAs without stably moving away from the promoter.
- It is thought to play a role in promoter clearance, allowing the RNA polymerase to transition from a closed to an open complex and then into a stable elongation complex.

(ii) CAP

- CAP stands for Catabolite Activator Protein, also known as CRP (cAMP Receptor Protein).
- CAP is a homodimeric protein in *E. coli* that acts as a transcriptional activator.
- It binds to cyclic AMP (cAMP) when glucose levels are low. The cAMP-CAP complex then binds to specific DNA sequences located upstream of the promoter of certain operons, such as the *lac* operon.
- This binding enhances the affinity of RNA polymerase for the promoter, thereby positively regulating the transcription of genes involved in the metabolism of alternative carbon sources when glucose, the preferred carbon source, is scarce.

(iii) RNA Pol III

- RNA Pol III refers to RNA Polymerase III, one of the three main nuclear RNA polymerases in eukaryotes.

- Its primary function is to synthesize small, stable RNA molecules that are essential for various cellular processes.
- Specifically, RNA Pol III transcribes:
    - Transfer RNAs (tRNAs), which carry amino acids to the ribosome during protein synthesis.
    - 5S ribosomal RNA (5S rRNA), a component of the large ribosomal subunit.
    - A variety of other small nuclear RNAs (snRNAs) and small cytoplasmic RNAs (scRNAs).
- Its promoters are typically located within the genes it transcribes.

(iv) EF-Tu

- EF-Tu stands for Elongation Factor Tu.
- EF-Tu is a prokaryotic elongation factor that plays a crucial role in protein synthesis (translation).
- Its primary function is to bind to aminoacyl-tRNAs (tRNAs carrying their specific amino acids) and escort them to the A-site (aminoacyl site) of the ribosome.
- EF-Tu, in complex with GTP, ensures the correct delivery of the charged tRNA to the ribosome, where codon-anticodon pairing is checked. GTP hydrolysis by EF-Tu occurs only if the correct codon-anticodon match is made, providing a proofreading step that enhances the fidelity of translation.

(v) RISC

- RISC stands for RNA-Induced Silencing Complex.
- RISC is a multi-protein complex that plays a central role in RNA interference (RNAi), a gene silencing mechanism in eukaryotes.

- Its primary function is to bind to small interfering RNAs (siRNAs) or microRNAs (miRNAs) and then use these small RNAs as guides to locate and silence complementary messenger RNA (mRNA) molecules.
- Depending on the degree of complementarity between the small RNA and the target mRNA, RISC can either cleave the target mRNA (for perfect complementarity, typically with siRNA) or inhibit its translation and promote its decay (for imperfect complementarity, typically with miRNA).

(vi) Polycistronic mRNA

- Polycistronic mRNA is a single messenger RNA molecule that encodes for more than one polypeptide chain.
- It is characteristic of prokaryotic gene organization, where genes involved in a common metabolic pathway are often clustered together in an operon.
- A single promoter controls the transcription of the entire operon, producing one polycistronic mRNA molecule.
- This mRNA then contains multiple start and stop codons, allowing for the translation of several distinct proteins from a single transcript. This arrangement allows for coordinated expression of functionally related genes.

(vii) Enhancers

- Enhancers are regulatory DNA sequences that significantly increase the transcription of target genes.
- They can be located a considerable distance from the promoter, either upstream or downstream of the gene, or even within introns.
- Enhancers function by binding to specific transcription factors (activators). These DNA-bound activators then interact with the general transcription machinery at the promoter, often through DNA

looping and the involvement of co-activators and the Mediator complex, to stimulate transcription.

- Enhancers can act in an orientation-independent manner.

(viii) Insulators

- Insulators are DNA sequences that act as boundaries to prevent inappropriate interactions between genes and their regulatory elements.
- Their primary function is to define independent domains of gene expression.
- They can block the enhancing effects of an enhancer on a gene it is not intended to regulate (enhancer-blocking activity) or protect a gene from being silenced by spreading heterochromatin (barrier activity).
- Insulators achieve this by binding specific proteins that help to organize chromatin structure and regulate accessibility.

(ix) Mediator complex

- The Mediator complex is a multi-protein complex found in eukaryotes.
- Its primary function is to act as a crucial bridge or co-activator that links gene-specific transcription factors (activators bound to enhancers or upstream promoter elements) to the general transcription machinery at the core promoter.
- It facilitates communication between these regulatory elements and RNA Polymerase II, influencing the rate of transcription initiation.
- The Mediator complex can also interact with components of the chromatin remodeling machinery and general transcription factors, integrating signals from various regulatory proteins to fine-tune gene expression.

(x) Release factors

- Release factors are proteins that play a critical role in the termination phase of protein synthesis (translation).
- Their primary function is to recognize stop codons (UAA, UAG, UGA) in the mRNA A-site of the ribosome.
- In prokaryotes, RF1 recognizes UAA and UAG, while RF2 recognizes UAA and UGA. RF3 assists RF1/RF2.
- In eukaryotes, eRF1 recognizes all three stop codons, and eRF3 assists eRF1.
- Upon stop codon recognition, release factors promote the hydrolysis of the bond between the polypeptide chain and the tRNA in the P-site, leading to the release of the newly synthesized polypeptide from the ribosome. They also facilitate the dissociation of ribosomal subunits.

Question 1. (b) Briefly explain the following statements :

(i) RNA polymerase core enzyme is catalytically active but non-specific.

- The RNA polymerase core enzyme in prokaryotes consists of four subunits: two α subunits, one β subunit, and one β' subunit ($\alpha_2\beta\beta'$).
- **Catalytically Active:** The core enzyme contains the active site for RNA synthesis, meaning it can catalyze the formation of phosphodiester bonds between ribonucleotides, thereby synthesizing an RNA molecule from a DNA template. In vitro, it can perform this catalytic function.
- **Non-specific:** However, the core enzyme alone lacks the ability to specifically recognize and bind to promoter sequences on the DNA template. Without the ability to find a promoter, it will initiate transcription randomly at various sites on the DNA, producing heterogeneous and non-functional RNA transcripts.
- This non-specific binding and initiation highlight the need for an additional factor, the sigma (σ) factor, to confer promoter specificity.

(ii) Transcriptional regulation in eukaryotes is primarily positive.

- This statement refers to the predominant mode of gene regulation in eukaryotes compared to prokaryotes.
- **Positive Regulation:** In eukaryotic transcription, the default state of most genes is "off" or repressed. This is largely due to the condensed nature of chromatin (DNA tightly wound around histones) which makes promoters inaccessible to RNA polymerase and general transcription factors.
- Therefore, for a gene to be actively transcribed, there is a general requirement for the recruitment of activators. These are gene-specific transcription factors that bind to enhancer elements or promoter-proximal elements.
- These activators work by recruiting co-activators, chromatin remodeling complexes, histone modifying enzymes, and the Mediator complex to open up the chromatin, make the promoter accessible, and facilitate the assembly and activity of the RNA Polymerase II pre-initiation complex.
- While negative regulation (repression) also occurs, the primary mechanism for turning on genes and achieving specific expression patterns involves activators that positively stimulate transcription, overcoming the repressive chromatin environment. In contrast, prokaryotic regulation often involves repressors blocking an otherwise accessible promoter (negative regulation), though activators are also present.

Question 2. Differentiate between the following (any 5) :

(a) Group I and Group II introns

Feature	Group I Introns	Group II Introns
Catalytic Mechanism	Self-splicing; requires an external guanosine (GTP, GDP, or GMP) as a cofactor.	Self-splicing; uses an internal A residue to form a lariat intermediate.
Mechanism Steps	Two transesterification reactions: G attacks 5' splice site; 3'-OH of exon attacks 3' splice site.	Two transesterification reactions: internal A's 2'-OH attacks 5' splice site forming lariat; 3'-OH of exon attacks 3' splice site.
Intermediate	Linear intron with G added to 5' end.	Lariat structure (branched RNA).
Energy Source	Does not require ATP hydrolysis or protein factors for the splicing reaction itself.	Does not require ATP hydrolysis or protein factors for the splicing reaction itself.
Location	Found in nuclear pre-rRNA (e.g., Tetrahymena), mitochondrial, and chloroplast genes, some bacterial genes.	Found in mitochondrial and chloroplast genes, some bacterial genes.
Evolutionary Relation	Thought to be related to telomerase RNA.	Thought to be evolutionary ancestors of spliceosomal introns.

(b) Rho dependent and independent transcription termination.

Feature	Rho-Dependent Termination	Rho-Independent Termination
Protein Factor Involved	Requires the Rho protein (a hexameric helicase).	Does not require any protein factor.
Mechanism	Rho protein binds to a specific C-rich, G-poor sequence on nascent RNA (rut site), moves along RNA towards RNA polymerase. When it reaches a stalled RNA polymerase (often at a "pause site"), its helicase activity unwinds the RNA-DNA hybrid, releasing the RNA transcript.	Relies solely on specific RNA sequences and structures.
RNA Structure Required	No specific hairpin structure required, but a Rho utilization site (rut site) on RNA.	Formation of a G-C rich hairpin (stem-loop) structure followed by a run of 6-8 U residues in the nascent RNA.
Stalling of RNA Pol	Often involves a transcription pause site where RNA polymerase slows down or stalls, allowing Rho to catch up.	RNA polymerase stalls at the U-rich tract immediately after the hairpin.
Dissociation	ATP-dependent helicase activity of Rho protein unwinds RNA-DNA hybrid.	Weak A-U base pairing in the RNA-DNA hybrid facilitates dissociation after hairpin formation.

(c) RNA Pol I and RNA Pol II

Feature	RNA Polymerase I (Pol I)	RNA Polymerase II (Pol II)
Transcribes	Most ribosomal RNA (rRNA) genes (18S, 5.8S, 28S rRNAs).	All protein-coding genes (pre-mRNA) and some small RNAs (e.g., snRNAs, miRNAs).
Location	Nucleolus	Nucleoplasm
Sensitivity to α-Amanitin	Insensitive	Highly sensitive (inhibited by very low concentrations)
Promoters	Generally upstream sequence elements.	Highly complex and diverse promoters, including TATA box, initiator, DPE.
Transcript Processing	Extensive processing, cleavage, and modification (e.g., methylation) of pre-rRNA.	Capping at 5' end, splicing, polyadenylation at 3' end.
Complexity of Regulation	Relatively simple, primarily regulated by growth conditions.	Highly complex, regulated by numerous transcription factors and chromatin modifications.

(d) Transcription and Replication

Feature	Transcription	Replication
Purpose	Synthesis of RNA from a DNA template.	Synthesis of DNA from a DNA template (DNA duplication).
Enzyme	RNA Polymerase	DNA Polymerase

Feature	Transcription	Replication
Template	Single strand of DNA (template strand)	Both strands of DNA (leading and lagging strands)
Product	RNA (single-stranded)	DNA (double-stranded)
Specificity	Selective: only specific genes are transcribed at a given time.	Entire genome is replicated.
Base Pairing	A pairs with U, T pairs with A, G pairs with C, C pairs with G.	A pairs with T, T pairs with A, G pairs with C, C pairs with G.
Primer Requirement	Generally no primer required (de novo synthesis).	Requires an RNA primer to initiate synthesis.
Start/Stop Sites	Specific start (promoter) and stop (terminator) signals.	Specific origins of replication (start sites), ends at chromosome ends or termination sites.
Fidelity/Proofreading	Lower fidelity, some proofreading mechanisms.	Very high fidelity, extensive proofreading and repair mechanisms.

(e) siRNA and miRNA

Feature	siRNA (Small Interfering RNA)	miRNA (Micro RNA)
Origin	Typically exogenous (from viruses, experimental introduction) or endogenous dsRNA (e.g., transposons).	Endogenous (transcribed from specific genes in the genome).
Precursor	Long double-stranded RNA (dsRNA).	Primary miRNA (pri-miRNA) with hairpin structures, processed by Drosha/Pasha.
Processing	Processed by Dicer into ~20-25 bp double-stranded fragments.	Processed by Dicer (after Drosha) into ~20-25 bp double-stranded fragments.
Targeting	Usually perfect or near-perfect complementarity to target mRNA.	Typically imperfect complementarity to target mRNA.
Mechanism of Action	Primarily mRNA cleavage/degradation (slicing) by RISC.	Primarily translational repression and mRNA destabilization/decay by RISC.
Function	Defense against viruses/transposons, research tool for gene knockdown.	Gene regulation (development, differentiation, cell growth, apoptosis).
Specific Targets	Highly specific, often targets one or very few mRNAs.	Can target multiple mRNAs due to imperfect complementarity.

(f) Lac Repressor and Trp Repressor

Feature	Lac Repressor	Trp Repressor
Operon Regulated	*lac* operon (lactose metabolism).	*trp* operon (tryptophan biosynthesis).
Ligand/Corepressor	Allolactose (inducer)	Tryptophan (corepressor)
Binding State (to DNA)	Binds to DNA (operator) in the **absence** of allolactose (lactose).	Binds to DNA (operator) in the **presence** of tryptophan.
Regulatory Mode	**Negative Inducible:** Repressor is active and binds DNA, blocking transcription unless inactivated by an inducer.	**Negative Repressible:** Repressor is inactive unless activated by a corepressor, then binds DNA to block transcription.
Function	Prevents transcription of *lac* operon genes when lactose is absent.	Prevents transcription of *trp* operon genes when tryptophan is abundant.
Active Form	Tetramer, active in absence of allolactose.	Dimer, active only when bound to tryptophan.

Question 3. (a) What are the different consensus sequences in the promoter region of prokaryotes and eukaryotes? Explain with the help of a diagram.

Please note: As per your instructions, I cannot create schematic diagrams. I will explain the sequences verbally and describe their positions.

Prokaryotic Promoter Consensus Sequences: Prokaryotic promoters are relatively simple, typically consisting of two primary conserved sequences located upstream of the transcription start site (+1).

1. **-35 Sequence (Recognition Sequence):**

- **Location:** Approximately 35 base pairs upstream from the transcription start site.
- **Consensus Sequence:** TTGACA
- **Function:** This sequence is recognized by the sigma (σ) subunit of the RNA polymerase holoenzyme. It helps in the initial binding and recognition of the promoter by the RNA polymerase, orienting the enzyme correctly on the DNA.

2. **-10 Sequence (Pribnow Box):**
    - **Location:** Approximately 10 base pairs upstream from the transcription start site.
    - **Consensus Sequence:** TATAAT
    - **Function:** This sequence is rich in A-T base pairs, which are easier to unwind due to fewer hydrogen bonds. It is the site where the DNA helix is unwound (melted) to form the transcription bubble, allowing RNA polymerase to access the template strand and initiate transcription. It's crucial for facilitating the transition from a closed to an open promoter complex.

Diagrammatic Representation (Verbal Description): Imagine a linear DNA molecule.

- The transcription start site is marked as +1.
- Moving upstream (to the left), you would find the -10 sequence (Pribnow Box: TATAAT).
- Further upstream, around -35, you would find the -35 sequence (TTGACA).
- The spacing between the -35 and -10 sequences is also important, typically around 17-19 base pairs.
- RNA polymerase holoenzyme binds to both the -35 and -10 regions.

Eukaryotic Promoter Consensus Sequences (for RNA Polymerase II, which transcribes protein-coding genes): Eukaryotic promoters are much more complex and variable than prokaryotic ones. They typically involve multiple *cis*-acting elements that work together to recruit RNA Polymerase II and the general transcription factors.

3. **TATA Box:**
    - **Location:** Approximately 25-30 base pairs upstream from the transcription start site (hence often called the -25 or -30 element).
    - **Consensus Sequence:** TATAAA (or a variation).
    - **Function:** It is a binding site for the TATA-binding protein (TBP), a subunit of Transcription Factor IID (TFIID). TBP binding to the TATA box is crucial for the initial recruitment of TFIID and the subsequent assembly of the pre-initiation complex (PIC) at the promoter. It helps position RNA Polymerase II correctly for transcription initiation. Not all eukaryotic genes have a TATA box.
4. **Initiator Element (Inr):**
    - **Location:** Spans the transcription start site (+1), often from -2 to +4 or +5.
    - **Consensus Sequence:** YYAN(T/A)YY (where Y is pyrimidine, N is any nucleotide).
    - **Function:** It is recognized by TFIID and helps to define the transcription start site, particularly in genes that lack a TATA box.
5. **Downstream Promoter Element (DPE):**
    - **Location:** Approximately +28 to +32 relative to the transcription start site (i.e., downstream of +1).
    - **Consensus Sequence:** Often AGAC.

- **Function:** Works cooperatively with the Initiator element, especially in TATA-less promoters, to recruit TFIID and stabilize the pre-initiation complex.

6. **TFIIB Recognition Element (BRE):**
    - **Location:** Approximately -37 to -32 upstream of the transcription start site (immediately upstream of the TATA box).
    - **Consensus Sequence:** G/C G/C G/A C G C C
    - **Function:** Recognized by the general transcription factor TFIIB, which helps position RNA Polymerase II.

Diagrammatic Representation (Verbal Description): Imagine a linear DNA molecule, more complex than the prokaryotic one.

- The transcription start site is at +1.
- Immediately surrounding +1 is the Initiator (Inr) element.
- Downstream of +1, around +28 to +32, is the DPE.
- Upstream of +1, around -25 to -30, is the TATA box.
- Further upstream of the TATA box, around -37 to -32, is the BRE.
- Beyond these core promoter elements, further upstream and downstream, you would find more distant regulatory elements like enhancers and silencers, which are bound by gene-specific transcription factors.

Question 3. (b) You are provided with two proteins (X and Y), one of which can potentially bind onto the promoter sequence of gene Z and help RNA polymerase in regulating gene expression. Describe a technique with which you can confirm which one out of X and Y does bind to the promoter sequence.

The technique to confirm which protein (X or Y) binds to the promoter sequence of gene Z is **Electrophoretic Mobility Shift Assay (EMSA)**, also known as a Gel Shift Assay.

- **Principle:** EMSA is based on the principle that a protein-DNA complex migrates more slowly through a non-denaturing polyacrylamide gel than the free DNA probe. The "shift" in mobility indicates that a protein has bound to the DNA.
- **Steps of the Technique:**
    a. **Promoter DNA Probe Preparation:**
        - Obtain or synthesize a short, double-stranded DNA oligonucleotide that contains the specific promoter sequence of gene Z that you suspect binds the protein.
        - Label this DNA probe. Common labeling methods include radioactive labeling (e.g., with ^{32}P) or non-radioactive labeling (e.g., with biotin or a fluorescent tag).

    b. **Protein Preparation:**
        - Prepare purified samples of protein X and protein Y.
        - Prepare a control sample (e.g., buffer only or a non-binding protein) to ensure that any observed shifts are specific to X or Y.
    c. **Binding Reactions:**
        - Set up multiple reaction tubes. Each tube will contain:
            - The labeled promoter DNA probe for gene Z.
            - Binding buffer, which provides the appropriate ionic strength and pH for protein-DNA interaction.
            - Optionally, a non-specific competitor DNA (e.g., poly(dI-dC)) to absorb any proteins that bind non-

specifically to DNA, ensuring that only specific binding to the promoter sequence is detected.

- In separate tubes, add varying concentrations of protein X, protein Y, or the control protein.
- Incubate the reactions for a specific time (e.g., 20-30 minutes) at a suitable temperature (e.g., room temperature or 4°C) to allow protein-DNA binding to occur.

d. **Electrophoresis:**

- Load the entire binding reaction mixtures onto a non-denaturing polyacrylamide gel.
- Apply an electric current, causing the DNA and protein-DNA complexes to migrate through the gel.
- Because protein-DNA complexes are larger and heavier than free DNA probes, they will migrate more slowly, resulting in a "shift" or retardation in their electrophoretic mobility.

e. **Detection and Analysis:**

- After electrophoresis, the gel is analyzed to visualize the labeled DNA.
- If the probe is radioactive, the gel is exposed to X-ray film (autoradiography).
- If non-radioactive labels are used, the gel is scanned using a compatible detection system (e.g., chemiluminescence for biotin, fluorescence for fluorescent tags).
- **Interpretation:**

- **Free DNA Probe:** A band corresponding to the faster-migrating, unbound DNA probe will be observed in lanes where no protein binds or in the control lane.
- **Protein-DNA Complex (Shifted Band):** If protein X (or Y) binds to the promoter DNA, a slower-migrating band (the "shifted band") will appear in the corresponding lane, and the intensity of the free DNA band will decrease.
- By comparing the lanes containing protein X and protein Y, you can determine which protein causes a specific shift, indicating its binding to the promoter sequence of gene Z.

- **Additional Controls (Supershift and Competition):**
  - **Supershift Assay:** To confirm the identity of the binding protein, an antibody specific to protein X (or Y) can be added to the binding reaction. If the antibody binds to the protein-DNA complex, it will create an even larger, slower-migrating "supershifted" band, confirming that protein X (or Y) is indeed part of the complex.
  - **Cold Competition Assay:** Unlabeled (cold) DNA containing the promoter sequence is added in increasing amounts to the reaction along with the labeled probe and protein. If the protein binding is specific, the unlabeled competitor DNA will outcompete the labeled probe for protein binding, leading to a reduction in the shifted band intensity and an increase in the free labeled probe. A non-specific unlabeled DNA will not compete.

This comprehensive approach using EMSA would definitively confirm whether protein X or Y binds to the promoter sequence of gene Z.

Question 3. (c) Briefly explain how mRNA ends are processed with the help of a diagram.

Please note: As per your instructions, I cannot create schematic diagrams. I will explain the processing steps verbally and describe their order.

In eukaryotes, messenger RNA (mRNA) undergoes extensive processing at both its 5' and 3' ends before it can be exported from the nucleus and translated into protein. This processing is crucial for mRNA stability, export, translation efficiency, and protection against degradation.

1. 5' Capping:

- **Timing:** This process occurs very early, co-transcriptionally, when the nascent mRNA transcript is only about 20-40 nucleotides long.
- **Mechanism:**
    - **Removal of γ-phosphate:** A phosphatase enzyme removes the terminal γ-phosphate from the 5' end of the nascent mRNA.
    - **Addition of GMP:** Guanylyl transferase then adds a guanosine monophosphate (GMP) to the 5' end of the mRNA via an unusual 5'-to-5' triphosphate linkage (inverted linkage).
    - **Methylation:** Methyltransferase enzymes then methylate this added guanosine (at position N7) and often the 2'-O position of the first one or two nucleotides of the mRNA itself. The resulting structure is called a 7-methylguanosine cap (m^7G cap).
- **Functions of 5' Cap:**
    - **Protection from Degradation:** The 5'-to-5' linkage and methylation protect the mRNA from degradation by 5' exonucleases.
    - **Ribosome Binding:** It is essential for the efficient binding of the ribosome (specifically the eIF4E initiation factor) during translation initiation in eukaryotes.

- **Nuclear Export:** It plays a role in the efficient transport of mRNA from the nucleus to the cytoplasm.
- **Splicing Efficiency:** It can influence the efficiency of pre-mRNA splicing.

2. 3' Polyadenylation (Addition of Poly-A Tail):

- **Timing:** This process occurs at the 3' end of the mRNA after transcription termination.
- **Mechanism:**
  - **Cleavage Signal Recognition:** The RNA polymerase II transcribes beyond the actual coding sequence. Specific sequence elements in the 3' untranslated region (3' UTR) of the pre-mRNA signal for cleavage. The most common signal is the polyadenylation signal sequence, typically AAUAAA, located about 10-30 nucleotides upstream of the cleavage site. A G/U-rich sequence is often found downstream of the cleavage site.
  - **Cleavage:** A multi-protein complex (including CPSF - Cleavage and Polyadenylation Specificity Factor, and CstF - Cleavage Stimulation Factor) recognizes these signals, leading to the cleavage of the pre-mRNA at a site downstream of the AAUAAA sequence.
  - **Poly-A Polymerase (PAP) Activity:** After cleavage, an enzyme called poly-A polymerase (PAP) adds approximately 80-250 adenosine (A) nucleotides to the newly generated 3' hydroxyl end. This addition does not require a DNA template.
  - **Poly-A Binding Protein (PABP) Binding:** Once added, the poly-A tail is quickly bound by poly-A binding proteins (PABPs).
- **Functions of Poly-A Tail:**
  - **mRNA Stability:** The poly-A tail and associated PABPs protect the mRNA from degradation by 3' exonucleases.

- **Translation Efficiency:** It enhances the efficiency of translation initiation by interacting with the 5' cap and ribosomal components (forming a circularized mRNA for re-initiation).
- **Nuclear Export:** It aids in the efficient export of mRNA from the nucleus to the cytoplasm.

Diagrammatic Representation (Verbal Description): Imagine a pre-mRNA molecule being transcribed from left to right.

- At the extreme left (5' end), shortly after transcription begins, a "cap" structure (m^7G) is added, showing its inverted 5'-to-5' linkage and methyl groups.
- The mRNA continues, containing coding sequences (exons) and non-coding sequences (introns).
- At the extreme right (3' end), after the coding sequence, there's a recognition sequence (AAUAAA) followed by a cleavage site.
- A "scissor" symbol indicating cleavage.
- Beyond the cleavage site, a long stretch of "A" nucleotides (poly-A tail) is added, with PABPs bound along its length.

These two processing events, capping and polyadenylation, are essential for the maturation and function of eukaryotic mRNA.

Question 4. (a) Diagrammatically explain how one eukaryotic gene can give rise to two functional proteins.

Please note: As per your instructions, I cannot create schematic diagrams. I will explain the process verbally.

One eukaryotic gene can give rise to two or more different functional proteins primarily through a process called **alternative splicing**.

- **Background:** Eukaryotic genes contain both coding regions (exons) and non-coding regions (introns). After transcription, the primary RNA transcript (pre-mRNA) contains both exons and introns. Splicing is

the process that removes introns and ligates exons together to form a mature mRNA.

- **Mechanism of Alternative Splicing leading to two proteins:**

  f. **Gene Structure:**

  - Imagine a eukaryotic gene on a DNA strand. This gene consists of a promoter region, followed by several exons and introns in a specific order (e.g., Exon 1, Intron A, Exon 2, Intron B, Exon 3, Intron C, Exon 4, etc.).
  - The full genetic information to produce all possible protein variants is contained within this single gene.

  g. **Transcription:**

  - RNA Polymerase II transcribes the entire gene, including all exons and introns, to produce a long primary RNA transcript (pre-mRNA).
  - This pre-mRNA also undergoes 5' capping and 3' polyadenylation as explained previously.

  h. **Alternative Splicing (The Key Step):**

  - Instead of always removing all introns and joining all exons in a fixed linear order, the spliceosome (the molecular machinery responsible for splicing) can selectively include or exclude certain exons, or use alternative splice sites within exons or introns.
  - **Example for two proteins:** Consider a gene with Exon 1 - Exon 2 - Exon 3 - Exon 4.
    - **Splicing Pathway 1 (Protein 1):** In one cell type or under certain conditions, the spliceosome might remove Intron A, Intron B, and Intron C, joining Exon 1, Exon 2, Exon 3, and Exon 4 together. This forms **mRNA 1**.

- **Splicing Pathway 2 (Protein 2):** In another cell type or under different conditions, the spliceosome might skip Exon 3 entirely. It removes Intron A, then joins Exon 2 directly to Exon 4 (while also removing Intron B and Intron C). This forms **mRNA 2**.
- (Other common forms of alternative splicing include alternative 5' splice sites, alternative 3' splice sites, and mutually exclusive exons.)

i. **Translation:**

- **mRNA 1:** Is transported to the cytoplasm and translated into **Protein 1**. This protein will have a specific amino acid sequence and therefore a specific three-dimensional structure and function.
- **mRNA 2:** Is also transported to the cytoplasm and translated into **Protein 2**. Because it lacks Exon 3 (or has a different exon combination), Protein 2 will have a different amino acid sequence compared to Protein 1. This difference in sequence will lead to a different three-dimensional structure and, consequently, a different or modified biological function.

- **Diagrammatic Representation (Verbal Description):**
  - **Line 1 (DNA):** Show a gene with "Promoter - Exon1 - Intron A - Exon2 - Intron B - Exon3 - Intron C - Exon4 - Terminator".
  - **Line 2 (Pre-mRNA):** Show the full primary transcript after transcription, maintaining the same sequence of exons and introns, with a 5' cap and 3' poly-A tail.
  - **Branching Arrows from Pre-mRNA:**
    - **Arrow 1 (Splicing Pathway 1):** Show a path where all introns are removed, and Exon1, Exon2, Exon3, and

Exon4 are ligated, leading to "mRNA 1 (Exon1-Exon2-Exon3-Exon4)".

- **Arrow 2 (Splicing Pathway 2):** Show a different path where Intron A, Intron B, and Intron C are removed, but Exon3 is skipped, and Exon1, Exon2, and Exon4 are ligated, leading to "mRNA 2 (Exon1-Exon2-Exon4)".

- **Final Products:** From mRNA 1, an arrow points to "Protein 1". From mRNA 2, an arrow points to "Protein 2", emphasizing that these are distinct proteins.

This process of alternative splicing is a major contributor to the complexity and diversity of proteins produced from a relatively limited number of genes in the eukaryotic genome, allowing a single gene to encode multiple protein isoforms with specialized functions.

Question 4. (b) Explain the phenomena of attenuation with the example of the Trp operon in bacteria.

- **Attenuation** is a mechanism of transcriptional regulation, primarily found in prokaryotes, that controls gene expression after transcription initiation has occurred but before the full-length mRNA is produced. It functions by causing premature termination of transcription, thereby "attenuating" (reducing) the amount of full-length transcript. This mechanism is particularly common in operons encoding enzymes for amino acid biosynthesis.

- **Example: Trp Operon in Bacteria**

  - The *trp* operon in *E. coli* encodes enzymes necessary for the biosynthesis of the amino acid tryptophan. It is regulated by two mechanisms: negative repression by the Trp repressor protein (as mentioned in Q2.f) and attenuation. Attenuation provides a fine-tuning mechanism that responds to fluctuating levels of charged tRNAs for tryptophan ($tRNA^{Trp}$).

- **Structure of the *trp* Operon Leader Region:**

- The key to attenuation lies in a specific sequence located in the mRNA *leader region* (the *trpL* gene) that is transcribed immediately before the structural genes of the *trp* operon. This leader region is about 162 nucleotides long and contains:
    - A short open reading frame (ORF) that codes for a "leader peptide" of 14 amino acids. Importantly, this leader peptide contains two adjacent tryptophan codons (UGG-UGG).
    - Four complementary regions (labeled 1, 2, 3, and 4) within the mRNA leader sequence that can form different stem-loop (hairpin) structures.

- **Mechanism of Attenuation (How it works):**

j. **Coupled Transcription and Translation:** In prokaryotes, transcription and translation are coupled, meaning that ribosomes can begin translating an mRNA molecule even while it is still being transcribed by RNA polymerase. This coupling is essential for the attenuation mechanism.

k. **Scenario 1: High Tryptophan Levels (Excess Tryptophan)**

- When tryptophan is abundant in the cell, there is a sufficient supply of charged tryptophan tRNAs ($tRNA^{Trp}$).
- As RNA polymerase transcribes the *trp* operon leader region, the ribosome immediately begins translating the leader peptide sequence.
- Since $tRNA^{Trp}$ is abundant, the ribosome does not stall at the two tryptophan codons (UGG-UGG) in region 1 of the leader peptide ORF. It quickly translates through region 1 and proceeds to region 2.
- By the time the ribosome reaches region 2, region 3 has been transcribed by RNA polymerase. Because the

ribosome is occupying region 2, it prevents region 2 from pairing with region 3.

- Instead, region 3 pairs with region 4, forming a stable **3-4 hairpin structure**. This 3-4 hairpin acts as a **rho-independent transcription terminator**.
- The formation of this terminator hairpin, followed by a stretch of U residues, causes RNA polymerase to pause and then dissociate from the DNA template, prematurely terminating transcription just before the structural genes of the *trp* operon.
- Result: Transcription of tryptophan biosynthetic genes is attenuated (turned off or significantly reduced), conserving cellular resources.

l. **Scenario 2: Low Tryptophan Levels (Tryptophan Starvation)**

- When tryptophan is scarce, there is an insufficient supply of charged tryptophan tRNAs ($tRNA^{Trp}$).
- As the ribosome translates the leader peptide, it encounters the two tryptophan codons (UGG-UGG) in region 1.
- Because of the lack of charged $tRNA^{Trp}$, the ribosome **stalls** at these tryptophan codons in region 1.
- The stalling of the ribosome at region 1 prevents region 1 from pairing with region 2.
- Since region 2 is now free, it pairs with region 3 to form a stable **2-3 hairpin structure**.
- Crucially, the formation of the 2-3 hairpin **prevents the formation of the 3-4 terminator hairpin**.

- With no terminator hairpin formed, RNA polymerase continues transcribing past the leader region and into the structural genes of the *trp* operon.
- Result: Transcription of tryptophan biosynthetic genes continues (or is not attenuated), allowing the cell to synthesize its own tryptophan.

- **Summary:** Attenuation effectively senses the availability of tryptophan (via charged $tRNA^{Trp}$ levels) by monitoring ribosome movement through the leader peptide. This allows for rapid and precise control of gene expression in response to immediate cellular needs.

Question 4. (c) Discuss the importance of GTP in the translation process. Explain how tRNA synthetases play the most crucial roles in determining the fidelity of translation.

- **Importance of GTP in the Translation Process:** GTP (Guanosine Triphosphate) is a crucial energy source and regulatory molecule in protein synthesis, participating in several key steps:

m. **Initiation:**

- **Prokaryotic Initiation:** Initiation factors IF-2 (prokaryotes) bind GTP. GTP hydrolysis by IF-2 is required for the binding of the large ribosomal subunit to the small subunit, releasing the initiation factors, and allowing the ribosome to fully assemble at the start codon.
- **Eukaryotic Initiation:** Eukaryotic initiation factors (e.g., eIF2, eIF5B) also bind GTP. eIF2-GTP complex escorts the initiator tRNA to the small ribosomal subunit. GTP hydrolysis by eIF2 (stimulated by eIF5B) is essential for release of eIF2-GDP and assembly of the 80S ribosome.

n. **Elongation:**

- **Aminoacyl-tRNA Delivery:** Elongation Factor Tu (EF-Tu in prokaryotes) or eEF1A (eukaryotes) binds to aminoacyl-tRNAs in a GTP-bound state. This EF-Tu/eEF1A-GTP complex escorts the charged tRNA to the A-site of the ribosome.
- **Proofreading and Accuracy:** GTP hydrolysis by EF-Tu/eEF1A occurs only if there is a correct (or near-correct) codon-anticodon match at the A-site. This hydrolysis promotes a conformational change that commits the tRNA to the ribosome, providing a kinetic proofreading step that enhances translational accuracy. Without GTP hydrolysis, the tRNA would dissociate.
- **Translocation:** Elongation Factor G (EF-G in prokaryotes) or eEF2 (eukaryotes) binds to the ribosome in a GTP-bound state. GTP hydrolysis by EF-G/eEF2 powers the translocation of the ribosome along the mRNA, moving the peptidyl-tRNA from the A-site to the P-site and the deacylated tRNA from the P-site to the E-site. This movement sets up the A-site for the next incoming aminoacyl-tRNA.

o. **Termination:**

- **Release Factor Activity:** In both prokaryotes and eukaryotes, certain release factors (RF3 in prokaryotes, eRF3 in eukaryotes) are GTPases. They bind to the ribosome and facilitate the release of the newly synthesized polypeptide chain from the tRNA in the P-site upon recognition of a stop codon. GTP hydrolysis provides the energy for the conformational changes required for this release and for the dissociation of the ribosomal subunits.

In summary, GTP acts as a molecular switch, driving conformational changes in various translation factors and providing the energy for

several critical, irreversible steps, thereby ensuring the efficiency, accuracy, and unidirectional progression of protein synthesis.

- **How tRNA Synthetases Play the Most Crucial Roles in Determining the Fidelity of Translation:** While the ribosome and elongation factors contribute to translational fidelity through proofreading during codon-anticodon recognition, the **aminoacyl-tRNA synthetases (aaRSs)** are considered to play the *most crucial* role in determining the overall fidelity of translation. Their function ensures that the correct amino acid is attached to its cognate tRNA, a process known as **"charging" or "aminoacylation"** of the tRNA.
    - **The "Second Genetic Code":** The accuracy of aminoacyl-tRNA synthetases is often referred to as the "second genetic code" because they are responsible for correctly matching each of the 20 amino acids with its corresponding set of tRNAs. If an aaRS makes a mistake at this step, the wrong amino acid will be incorporated into the growing polypeptide chain regardless of the correct codon-anticodon pairing at the ribosome. This is analogous to a typo in the "dictionary" of the genetic code.
    - **Specificity and Double Sieve Mechanism:** Aminoacyl-tRNA synthetases achieve high fidelity through a two-step process:
        i. **Amino Acid Recognition and Activation:**
            - Each aaRS has a highly specific active site that recognizes and binds to its cognate amino acid.
            - The amino acid is activated by ATP to form an aminoacyl-AMP intermediate.
            - The activated amino acid is then transferred to the 3' end of its specific tRNA.
            - **Initial Discrimination:** The primary selection occurs at this active site, based on the size, shape, and chemical properties of the amino acid.

However, some amino acids are structurally very similar (e.g., isoleucine and valine), making initial discrimination challenging.

ii. **Proofreading (Editing) Site:**

- Many aaRSs possess a separate, distinct "editing" or "proofreading" site within the enzyme.
- This site is designed to hydrolyze (remove) amino acids that are *smaller* than the correct amino acid and have mistakenly been attached to the tRNA or activated.
- For example, isoleucyl-tRNA synthetase (IleRS) can initially activate valine (which is similar but smaller than isoleucine) due to its size. However, the IleRS editing site then specifically accommodates and hydrolyzes valyl-tRNA, but it is too small to accommodate isoleucyl-tRNA.
- This "double sieve" mechanism (first discrimination at the synthetic site, second at the editing site) significantly improves the accuracy of aminoacylation, reducing errors to about 1 in 10,000 to 100,000 events.

- **tRNA Recognition:**
  - In addition to recognizing the correct amino acid, each aaRS must also recognize its *cognate tRNA* uniquely. This involves recognizing specific structural features on the tRNA molecule, often called "identity elements," which can be located in the anticodon loop, acceptor stem, or other parts of the tRNA.

- This ensures that the correct amino acid is loaded onto the correct tRNA, so that the genetic code is accurately read by the ribosome.

If an aaRS makes a mistake and links the wrong amino acid to a tRNA, the ribosome has no mechanism to detect this error during codon-anticodon pairing, as the tRNA itself is correctly paired with the mRNA codon. Therefore, the fidelity of aminoacyl-tRNA synthetases is paramount for ensuring that the correct protein sequence is synthesized, directly impacting the accuracy and functionality of all proteins produced in the cell.

Question 5. (a) Discuss the salient features of the Genetic code.

The genetic code is the set of rules by which information encoded in genetic material (DNA or RNA sequences) is translated into proteins (amino acid sequences) by living cells. Its salient features are:

7. **Triplet Code:** Each amino acid is specified by a sequence of three consecutive nucleotides called a codon. For example, AUG codes for methionine.
8. **Non-overlapping:** The codons are read sequentially, one after another, without any overlap. Each nucleotide is part of only one codon.
9. **Comma-less (No Punctuation):** There are no intervening nucleotides or "commas" between codons. The coding sequence is read continuously from a fixed starting point.
10. **Degenerate (Redundant):** Most amino acids are specified by more than one codon. For example, both UUA and UUG code for leucine. This degeneracy provides a buffer against the potentially harmful effects of point mutations.
11. **Universal (Nearly Universal):** The genetic code is remarkably universal across almost all forms of life, from bacteria to humans, implying a common evolutionary origin. A given codon specifies the

same amino acid in nearly all organisms. There are a few minor exceptions, primarily in mitochondrial genomes and some single-celcelled organisms, where certain codons may have different meanings.

12. **Unambiguous:** Each codon specifies only one particular amino acid. For example, UGG always codes for tryptophan and never for any other amino acid.

13. **Start Codon:** The codon AUG serves two functions: it specifies the amino acid methionine (or N-formylmethionine in prokaryotes) and also acts as the primary "start" signal for protein synthesis. In some cases, GUG or UUG can also serve as start codons, typically for methionine.

14. **Stop (Termination) Codons:** Three codons, UAA, UAG, and UGA, do not specify any amino acid. Instead, they act as "stop" signals, indicating the termination of protein synthesis. They are also known as nonsense codons.

15. **Reading Frame:** The coding sequence of an mRNA molecule must be read in the correct reading frame (a series of non-overlapping triplets) to produce the correct protein. The start codon establishes this reading frame. A single nucleotide insertion or deletion (frameshift mutation) can drastically alter the protein sequence downstream of the mutation.

Question 5. (b) Explain four mechanisms by which eukaryotic repressors repress transcription.

Eukaryotic repressors are transcription factors that bind to specific DNA sequences (often called silencer elements or operator sites within promoters) and inhibit the transcription of target genes. Unlike prokaryotic repressors which often simply block RNA polymerase access, eukaryotic repressors employ diverse mechanisms due to the complexity of eukaryotic chromatin and the transcriptional machinery.

Here are four major mechanisms:

16. **Competition for Activator Binding Sites:**

- **Mechanism:** A repressor protein can bind to the same or an overlapping DNA sequence that an activator protein would normally bind to.
- **Effect:** By occupying the binding site, the repressor physically prevents the activator from binding to the DNA, thereby blocking the activator's ability to recruit co-activators or interact with the basal transcription machinery.
- **Example:** A repressor might bind to an enhancer element, physically preventing an activating transcription factor from binding to that enhancer.

17. **Masking of Activation Domain:**

- **Mechanism:** A repressor protein binds to a DNA sequence near an activator's binding site (or even directly to the activator itself) and then physically interacts with the activation domain of an activator protein.
- **Effect:** This interaction does not prevent the activator from binding to DNA, but it masks or sterically hinders the activator's functional domain that is necessary for recruiting general transcription factors or co-activators (like the Mediator complex) to the promoter. The activator is present but rendered non-functional.
- **Example:** A repressor might bind to an activator and prevent its interaction with TFIIB or a subunit of the Mediator complex.

18. **Direct Interaction with General Transcription Factors or RNA Polymerase II:**

- **Mechanism:** A repressor protein, once bound to DNA, can directly interact with one or more components of the general transcription machinery (e.g., TFIID, TFIIB, or RNA Polymerase II itself).

- **Effect:** This direct interaction can physically impede the assembly of the pre-initiation complex (PIC), destabilize it, or interfere with its activity (e.g., preventing RNA Pol II from initiating transcription or efficiently clearing the promoter).
- **Example:** Some repressors have domains that directly bind to and inhibit the function of TBP or a subunit of TFIIB.

19. **Recruitment of Chromatin-Modifying Enzymes and Chromatin Remodelers:**

- **Mechanism:** This is a very common and powerful mechanism in eukaryotes. Repressor proteins recruit other proteins that modify chromatin structure, making the DNA less accessible for transcription.
- **Types of recruited proteins:**
  - **Histone Deacetylases (HDACs):** Repressors can recruit HDACs, which remove acetyl groups from histone tails. Deacetylation increases the positive charge of histones, leading to tighter binding of DNA to histones and a more condensed, repressive chromatin structure (heterochromatin), thereby blocking access for transcription factors and RNA polymerase.
  - **Histone Methyltransferases (HMTs):** Some repressors recruit HMTs, which add methyl groups to specific lysine or arginine residues on histone tails. Depending on the specific methylation mark (e.g., H3K9me3), this can lead to gene silencing by recruiting heterochromatin-binding proteins (e.g., HP1) that further condense chromatin.
  - **Nucleosome Remodeling Complexes:** Repressors can recruit ATP-dependent chromatin remodeling complexes that reposition, evict, or alter nucleosomes to create a more compact and less accessible chromatin state at the promoter region.

- **Effect:** All these actions collectively lead to chromatin compaction, reducing the accessibility of the promoter to transcriptional activators, general transcription factors, and RNA Polymerase II, thereby repressing transcription.
- **Example:** The binding of a repressor to a silencer element might recruit an HDAC, leading to localized chromatin condensation and gene silencing.

These mechanisms allow eukaryotic cells to exert fine-tuned, robust, and often long-lasting control over gene expression, which is crucial for development, differentiation, and maintaining cellular identity.

Question 5. (c) Write the mechanism of action of the following :

(i) Rifampicin

- **Mechanism of Action:** Rifampicin is a bacterial antibiotic that specifically targets **bacterial RNA Polymerase**.
- It binds to the β subunit of the bacterial RNA polymerase holoenzyme, specifically preventing the enzyme from elongating the RNA chain beyond the initial few nucleotides.
- It primarily inhibits the initiation step of transcription, blocking the formation of the first phosphodiester bond or severely inhibiting promoter clearance, rather than affecting ongoing elongation.
- This selective inhibition of bacterial RNA polymerase makes it effective against bacterial infections, particularly mycobacterial infections like tuberculosis. It does not significantly affect eukaryotic RNA polymerases.

(ii) α-Amanitin

- **Mechanism of Action:** α-Amanitin is a potent toxin derived from *Amanita phalloides* (death cap mushroom) that specifically inhibits eukaryotic RNA polymerases.

- Its mechanism of action differs depending on the specific eukaryotic RNA polymerase it encounters and its concentration:
    - **RNA Polymerase II (Pol II):** It binds very tightly to RNA Polymerase II (which transcribes mRNA), severely inhibiting its elongation activity at very low concentrations. This is its most significant effect, leading to the blockage of mRNA synthesis and ultimately cell death.
    - **RNA Polymerase III (Pol III):** It inhibits RNA Polymerase III (which transcribes tRNAs and 5S rRNA) at higher concentrations.
    - **RNA Polymerase I (Pol I):** It has virtually no effect on RNA Polymerase I (which transcribes most rRNAs), even at very high concentrations.
- This differential sensitivity is a key tool for researchers to distinguish the activities of different eukaryotic RNA polymerases. Its potent inhibition of Pol II makes it extremely toxic to eukaryotic cells.

(iii) Puromycin

- **Mechanism of Action:** Puromycin is an antibiotic that inhibits protein synthesis (translation) in both prokaryotic and eukaryotic cells, though it is more commonly used in eukaryotic cell biology research.
- It mimics the structure of the 3' end of an aminoacyl-tRNA, particularly the aminoacyl portion of tyrosyl-tRNA.
- Puromycin enters the A-site (aminoacyl site) of the ribosome.
- The peptidyl transferase activity of the ribosome then forms a peptide bond between the growing polypeptide chain (on the P-site tRNA) and puromycin.
- Since puromycin lacks a further amino acid or the ability to be translocated, the growing polypeptide chain is prematurely released from the ribosome, resulting in the production of truncated, non-

functional polypeptide chains. This causes premature termination of translation.

(iv) Actinomycin D

- **Mechanism of Action:** Actinomycin D is an antibiotic (and an anti-cancer drug, although its use is limited by toxicity) that primarily inhibits **transcription** in both prokaryotic and eukaryotic cells.
- Its mechanism involves intercalating (inserting itself) into the double helix of DNA, specifically between adjacent G-C base pairs.
- This intercalation distorts the DNA structure and physically blocks the movement of RNA polymerase along the DNA template.
- By preventing RNA polymerase from transcribing DNA, Actinomycin D inhibits the synthesis of all types of RNA (mRNA, tRNA, rRNA), thereby blocking gene expression and ultimately inhibiting protein synthesis.

Question 6. Write short notes on the following :

(a) DNA binding domains

- **Description:** DNA binding domains (DBDs) are specific structural motifs within proteins that enable them to recognize and bind to particular DNA sequences. These domains are crucial for the function of transcription factors, nucleases, and other DNA-interacting proteins. The interaction is typically sequence-specific, allowing proteins to regulate particular genes or perform specific enzymatic functions on DNA.
- **Common Motifs:** Several well-characterized DBD motifs exist:
    - **Helix-Turn-Helix (HTH):** Found in many prokaryotic repressors (e.g., Lac repressor, Trp repressor) and some eukaryotic proteins. It consists of two α-helices separated by a short turn; one helix (the recognition helix) fits into the major groove of DNA.

- **Zinc Finger:** A highly versatile motif, common in eukaryotes, characterized by zinc ions coordinating with cysteine and/or histidine residues to stabilize small protein domains. These domains typically form α-helices that insert into the major groove and recognize specific base sequences. Multiple zinc fingers can be arranged in tandem to increase binding specificity.
- **Leucine Zipper (bZIP):** Characterized by a region of leucine residues spaced seven amino acids apart, forming an α-helix. Two such helices can "zip" together to form a coiled-coil dimer. Adjacent to this dimerization region is a basic region rich in lysine and arginine, which interacts with the major groove of DNA.
- **Helix-Loop-Helix (bHLH):** Similar to bZIP, it involves two α-helices connected by a loop. Like bZIP, it often forms dimers, and a basic region facilitates DNA binding.

- **Function:** DBDs enable proteins to selectively bind to promoters, enhancers, silencers, or other regulatory regions, thereby controlling gene expression, DNA replication, DNA repair, and recombination. The specificity of these interactions is fundamental to genome function.

(b) Spliceosome

- **Description:** The spliceosome is a large, dynamic, and highly complex molecular machine found in the nucleus of eukaryotic cells. It is responsible for the removal of introns (non-coding sequences) from pre-mRNA (primary RNA transcript) and the precise ligation of exons (coding sequences) to form mature mRNA. This process is called RNA splicing.
- **Composition:** The spliceosome is primarily composed of five small nuclear ribonucleoproteins (snRNPs), denoted as U1, U2, U4, U5, and U6. Each snRNP consists of one or two small nuclear RNAs

(snRNAs) and a set of associated proteins. In addition to snRNPs, numerous other proteins (non-snRNP proteins) are transiently associated with the spliceosome.

- **Mechanism (Brief):** The spliceosome assembles sequentially on the pre-mRNA, guided by base pairing between snRNAs and consensus sequences at the intron-exon junctions (5' splice site, 3' splice site, and branch point A within the intron). It catalyzes two sequential transesterification reactions that excise the intron in a lariat form and join the exons together.
- **Significance:** Splicing is essential for generating functional mRNA in eukaryotes. Furthermore, the spliceosome enables alternative splicing, a mechanism by which a single pre-mRNA can produce multiple distinct mRNA isoforms (and thus different proteins) by selectively including or excluding exons, thereby greatly increasing the coding capacity of the genome.

(c) Sigma Factor

- **Description:** The sigma (σ) factor is a detachable subunit of the bacterial RNA polymerase holoenzyme. It is a key protein that plays a crucial role in initiating transcription in prokaryotes.
- **Function:** The core RNA polymerase enzyme ($\alpha_2\beta\beta'$) is catalytically active but cannot specifically recognize promoter sequences. The sigma factor's primary function is to confer promoter specificity to the RNA polymerase.
    - **Promoter Recognition:** The sigma factor specifically recognizes and binds to the consensus sequences in the prokaryotic promoter, particularly the -35 and -10 regions (Pribnow box). This binding positions the RNA polymerase holoenzyme correctly at the transcription start site.
    - **DNA Unwinding:** It assists in the initial unwinding of the DNA double helix at the -10 region to form the open promoter complex.

- **Promoter Clearance:** Once transcription initiation has occurred and the RNA polymerase has synthesized a short RNA transcript (typically ~10 nucleotides), the sigma factor usually dissociates from the core enzyme. This dissociation allows the core enzyme to enter the elongation phase of transcription.

- **Diversity:** Bacteria have multiple sigma factors (e.g., σ^{70} for housekeeping genes, σ^{32} for heat shock genes, σ^{N} for nitrogen metabolism), each recognizing different sets of promoters. This allows for global regulation of gene expression in response to various environmental conditions.

(d) Regulation of Gal genes in Yeast

- **Description:** The *GAL* genes in the yeast *Saccharomyces cerevisiae* encode enzymes required for metabolizing galactose (e.g., Gal1, Gal7, Gal10). Their expression is tightly regulated, primarily by the availability of galactose and glucose. This system is a classic example of eukaryotic transcriptional regulation involving both positive and negative control.

- **Key Regulatory Proteins:**

  - **Gal4 (Activator):** A sequence-specific DNA-binding protein (transcription factor) that binds to upstream activating sequences (UAS_G) in the promoters of *GAL* genes. In the absence of galactose, Gal4's activation domain is inhibited.

  - **Gal80 (Repressor):** A protein that binds directly to the activation domain of Gal4, preventing it from activating transcription.

  - **Gal3 (Sensor/Inducer):** A cytoplasmic protein that acts as a galactose sensor. In the presence of galactose (and ATP), Gal3 undergoes a conformational change and binds to Gal80.

  - **Mig1 (Glucose Repressor):** A glucose-sensing protein that, in the presence of glucose, recruits the Tup1/Cyc8 repressor

complex, leading to chromatin condensation and repression of *GAL* genes.

- **Mechanism:**
    - **Absence of Galactose:** Gal80 binds to Gal4, blocking its activation domain. *GAL* genes are repressed.
    - **Presence of Galactose:** Galactose binds to Gal3. The Gal3-galactose complex then binds to Gal80, causing a conformational change in Gal80. This releases Gal4's activation domain, allowing Gal4 to activate *GAL* gene transcription (even if Gal80 is still bound to Gal4, it no longer represses).
    - **Presence of Glucose (Catabolite Repression):** Glucose levels negatively regulate *GAL* genes (even if galactose is present). High glucose activates Mig1, which binds to specific DNA sites and recruits the Tup1/Cyc8 repressor complex. This complex promotes histone deacetylation and chromatin compaction, actively repressing *GAL* gene expression, ensuring yeast preferentially uses glucose over galactose.

- **Significance:** This system demonstrates sophisticated eukaryotic regulation involving activator-repressor interactions, ligand-mediated conformational changes, and chromatin-level control, allowing yeast to adapt its metabolism to available carbon sources.

(e) Histone modifications

- **Description:** Histone modifications are reversible, post-translational chemical alterations that occur on the N-terminal tails (and sometimes globular domains) of core histone proteins (H2A, H2B, H3, H4) that form the nucleosome, the basic unit of chromatin. These modifications act as a "histone code" that influences chromatin structure and gene expression.
- **Types of Modifications:**

- **Acetylation:** Addition of acetyl groups to lysine residues, primarily by Histone Acetyltransferases (HATs). Acetylation neutralizes the positive charge of lysine, reducing histone-DNA interaction, leading to a more open, euchromatic state that is generally associated with transcriptional activation. Removed by Histone Deacetylases (HDACs).
- **Methylation:** Addition of methyl groups to lysine and arginine residues, by Histone Methyltransferases (HMTs). Methylation can be mono-, di-, or tri- on lysines. Depending on the specific residue and degree of methylation, it can be associated with either transcriptional activation (e.g., H3K4me3) or repression (e.g., H3K9me3, H3K27me3). Removed by Histone Demethylases (HDMs).
- **Phosphorylation:** Addition of phosphate groups to serine, threonine, or tyrosine residues, by kinases. Often involved in chromatin condensation during mitosis (e.g., H3S10ph) or DNA repair (e.g., H2AX phosphorylation to γH2AX).
- **Ubiquitination:** Addition of ubiquitin protein to lysine residues. Often associated with transcriptional elongation or repression (e.g., H2BK120ub1 linked to active transcription, H2AK119ub1 linked to repression).
- **Sumoylation, ADP-ribosylation:** Other less common modifications.

- **Function (Histone Code Hypothesis):** These modifications do not directly alter DNA sequence but significantly impact chromatin accessibility and function. Specific combinations of modifications on histone tails serve as binding sites for "reader" proteins (e.g., bromodomains for acetylated lysines, chromodomains for methylated lysines). These reader proteins then recruit other complexes (e.g., chromatin remodelers, transcription factors) that either promote or inhibit gene transcription, DNA replication, DNA repair, or

chromosome segregation. The dynamic interplay of these modifications is central to epigenetic regulation and cellular identity.