1 (a) Define the following (any five): (1x5=5)

- (i) Codon: A codon is a sequence of three successive nucleotides in a DNA or RNA molecule that codes for a specific amino acid or a stop signal during protein synthesis.
- (ii) Split gene: A split gene is a gene that contains non-coding intervening sequences (introns) interspersed between the coding sequences (exons). Both introns and exons are transcribed into a primary RNA transcript, but introns are subsequently removed by splicing.
- (iii) Inducer: An inducer is a molecule that triggers the expression of a gene or operon by inactivating a repressor protein or by activating an activator protein. For example, lactose acts as an inducer for the *lac* operon.
- (iv) Primosome: A primosome is a protein complex that forms on single-stranded DNA during DNA replication and is responsible for synthesizing RNA primers required for DNA polymerase to initiate DNA synthesis. It typically consists of a primase enzyme and helicase.
- (v) Okazaki fragment: Okazaki fragments are short, newly synthesized DNA strands that are formed on the lagging strand during DNA replication. They are synthesized discontinuously in the 5' to 3' direction and are later joined together by DNA ligase to form a continuous strand.
- (vi) Glycosylation: Glycosylation is a post-translational modification process in which carbohydrate molecules (glycans) are covalently attached to proteins or lipids. This modification plays crucial roles in protein folding, stability, cell-cell recognition, and immune response.

 (b) Expand the following (any five): (1x5=5)

- (i) ORF: Open Reading Frame
- (ii) PCNA: Proliferating Cell Nuclear Antigen

- (iii) miRNA: microRNA
- (iv) TBP: TATA-Binding Protein
- (v) Tm: Melting Temperature (of DNA)
- (vi) CRP: cAMP Receptor Protein

 (c) Answer the following in one word (any five): (1x5=5)

- (i) Enzyme encoded by the *lacZ* gene. **Beta-galactosidase**
- (ii) Sequence of 5' and 3' splice site. **GT-AG (GU-AG in RNA)**
- (iii) Enzyme involved in charging tRNA during protein synthesis. **Aminoacyl-tRNA synthetase**
- (iv) Give the percentage of adenine present in a DNA molecule having 30% guanine in its base composition. **20%** (According to Chargaff's rules, G = C. If G = 30%, then C = 30%. Total G+C = 60%. Therefore, A+T = 40%. Since A = T, A = 20%.)
- (v) RNA polymerase subunit required to initiate transcription in prokaryote. **Sigma (σ) subunit**
- (vi) Name the enzyme involved in removal of RNA primer during eukaryotic DNA replication process. **RNase H and FEN1 (Flap Endonuclease 1)**

2 Differentiate between the following (any five): (3x5=15)

- (a) A-DNA and B-DNA
    - **A-DNA:**
        - Right-handed double helix.
        - Shorter and wider helix.
        - Approximately 11 base pairs per turn.

- Major groove is narrow and deep, minor groove is wide and shallow.
- Favored under dehydrating conditions or in solutions with low water activity.

- **B-DNA:**
  - Right-handed double helix.
  - Longer and narrower helix.
  - Approximately 10.5 base pairs per turn.
  - Major groove is wide and deep, minor groove is narrow and shallow.
  - Most common form of DNA found in living cells under physiological conditions.

- (b) Topoisomerase-I and Topoisomerase-II
  - **Topoisomerase-I:**
    - Creates transient single-stranded breaks in DNA.
    - Relieves supercoiling by passing one strand through the break.
    - Does not require ATP hydrolysis for its activity.
    - Changes the linking number by ±1.
  - **Topoisomerase-II:**
    - Creates transient double-stranded breaks in DNA.
    - Relieves supercoiling by passing a segment of DNA helix through the break.
    - Requires ATP hydrolysis for its activity.
    - Changes the linking number by ±2.

- (c) Denaturation and Renaturation of DNA
    - **Denaturation of DNA (Melting):**
        - The process where the double-stranded DNA helix unwinds and separates into two single strands.
        - Caused by breaking of hydrogen bonds between complementary base pairs, usually by heating or extreme pH.
        - Results in an increase in UV absorbance (hyperchromic effect).
        - It is a reversible process.
    - **Renaturation of DNA (Annealing):**
        - The process where two complementary single strands of DNA re-associate to form a double helix.
        - Occurs when denatured DNA is slowly cooled, allowing hydrogen bonds to reform between complementary bases.
        - Results in a decrease in UV absorbance.
        - Used in techniques like hybridization.
- (d) Negative and Positive Gene Regulation in Lac Operon
    - **Negative Gene Regulation in *Lac* Operon:**
        - Involves a repressor protein that binds to the operator region and inhibits transcription.
        - Occurs when lactose is absent: The *lac* repressor binds to the operator, blocking RNA polymerase from transcribing the structural genes.

        - Lactose (or allolactose) acts as an inducer, binding to the repressor and causing it to release from the operator, thereby allowing transcription.

    - **Positive Gene Regulation in *Lac* Operon:**

        - Involves an activator protein that binds to DNA and promotes transcription.
        - Occurs when glucose levels are low: cAMP levels rise and bind to cAMP Receptor Protein (CRP, also known as CAP).
        - The cAMP-CRP complex binds to the promoter region, enhancing the binding of RNA polymerase and stimulating transcription of the *lac* operon. This ensures that lactose metabolism is preferred only when glucose is scarce.

- (e) Monocistronic and Polycistronic RNA
    - **Monocistronic RNA:**
        - An mRNA molecule that codes for only one polypeptide chain (one gene).
        - Characteristic of eukaryotic mRNA.
        - Each mRNA typically has one start codon and one stop codon.
    - **Polycistronic RNA:**
        - An mRNA molecule that codes for multiple polypeptide chains (multiple genes) from a single transcript.
        - Characteristic of prokaryotic mRNA (e.g., in operons like the *lac* operon).
        - Contains multiple start and stop codons, each corresponding to a different protein.

- (f) Prokaryotic and Eukaryotic Ribosome
    - **Prokaryotic Ribosome (70S):**
        - Smaller in size (70S, where S is Svedberg unit).
        - Composed of two subunits: a large 50S subunit (containing 23S and 5S rRNA) and a small 30S subunit (containing 16S rRNA).
        - Found free in the cytoplasm.
        - Initiation of translation occurs at a Shine-Dalgarno sequence.
    - **Eukaryotic Ribosome (80S):**
        - Larger in size (80S).
        - Composed of two subunits: a large 60S subunit (containing 28S, 5.8S, and 5S rRNA) and a small 40S subunit (containing 18S rRNA).
        - Found free in the cytoplasm and attached to the endoplasmic reticulum.
        - Initiation of translation typically involves scanning from the 5' cap to find the first AUG codon (Kozak sequence).

3 Write short notes on (any three): (5x3=15)

- (a) Mechanism of splicing
    - Splicing is a crucial step in eukaryotic gene expression where non-coding introns are removed from the primary RNA transcript (pre-mRNA), and the coding exons are joined together to form a mature mRNA.
    - **Key components:** Splicing is carried out by a large ribonucleoprotein complex called the **spliceosome**, composed

of small nuclear ribonucleoproteins (snRNPs, often U1, U2, U4, U5, and U6) and other proteins.

- **Mechanism (two transesterification reactions):**
  - **Step 1: Cleavage at 5' splice site:** The U1 snRNP recognizes the 5' splice site (GU sequence) and U2 snRNP recognizes the branch point A within the intron. A series of interactions brings U4/U5/U6 snRNPs to form the spliceosome. The 2'-OH group of the branch point adenine attacks the phosphodiester bond at the 5' splice site, creating a lariat intermediate (a loop structure) and releasing the 5' exon.
  - **Step 2: Cleavage at 3' splice site and exon ligation:** The free 3'-OH group of the first exon then attacks the phosphodiester bond at the 3' splice site (AG sequence). This releases the intron (still in lariat form) and ligates (joins) the two exons together.
  - The mature mRNA can then be translated. Splicing ensures that only the coding sequences are translated into functional proteins, and alternative splicing allows a single gene to produce multiple protein isoforms.

- (b) Attenuation in Trp operon
  - Attenuation is a mechanism of gene regulation, primarily found in prokaryotes, that controls the transcription of operons involved in amino acid biosynthesis (like the *trp* operon). It works by prematurely terminating transcription based on the availability of the amino acid.
  - The *trp* operon in *E. coli* regulates the synthesis of tryptophan. When tryptophan levels are high, the operon is "attenuated," meaning transcription is halted before the structural genes are fully transcribed.

- **Mechanism:** The *trp* operon has a leader sequence containing four regions (1, 2, 3, 4) and a short peptide coding for two tryptophan residues.
    - **High Tryptophan:** When tryptophan is abundant, the ribosome quickly translates the leader peptide. This causes the ribosome to cover region 2, allowing regions 3 and 4 to form a stable hairpin loop (terminator loop). This terminator loop signals RNA polymerase to detach from the DNA, prematurely terminating transcription.
    - **Low Tryptophan:** When tryptophan is scarce, the ribosome stalls at the tryptophan codons in region 1 (because there's a shortage of Trp-tRNA). This stalling leaves region 2 exposed, allowing it to pair with region 3, forming an anti-terminator hairpin loop (2-3 loop). This anti-terminator structure prevents the formation of the 3-4 terminator loop, allowing RNA polymerase to continue transcription through the structural genes, thus synthesizing enzymes for tryptophan production.
- Attenuation provides a fine-tuned control mechanism that responds rapidly to cellular tryptophan levels.

- (c) Theta mode of DNA replication
    - Theta replication is a common mode of DNA replication found in circular DNA molecules, characteristic of prokaryotic chromosomes (e.g., *E. coli*) and some plasmids. It is named "theta" because the replicating structure resembles the Greek letter theta (θ).
    - **Mechanism:**
        - **Origin of Replication (oriC):** Replication begins at a single, specific origin of replication (*oriC*) on the circular DNA molecule.

- **Bidirectional Replication:** Two replication forks are formed, moving in opposite directions around the circular chromosome.
- **Unwinding:** DNA helicases unwind the DNA helix at the origin, creating a replication bubble. Single-strand binding proteins (SSBs) stabilize the separated strands.
- **Primer Synthesis:** Primase synthesizes short RNA primers, providing 3'-OH groups for DNA polymerase.
- **Leading and Lagging Strands:** At each replication fork, one strand is synthesized continuously (leading strand) in the 5' to 3' direction, while the other strand is synthesized discontinuously (lagging strand) as Okazaki fragments.
- **Elongation:** DNA polymerase III (in prokaryotes) synthesizes new DNA strands.
- **Primer Removal and Ligation:** RNA primers are removed by DNA polymerase I (which also fills the gaps), and Okazaki fragments are joined by DNA ligase.
- **Supercoiling Relief:** Topoisomerases relieve supercoiling ahead of the replication forks.
- **Termination:** The two replication forks meet at a termination region, resulting in two interlinked (catenated) circular DNA molecules.
- **Decatenation:** DNA topoisomerase IV (a type II topoisomerase) separates the interlinked circles, yielding two complete, identical circular DNA molecules.

- (d) Central Dogma
    - The Central Dogma of Molecular Biology, proposed by Francis Crick in 1957, describes the fundamental flow of genetic

information within a biological system. It states that genetic information flows generally from DNA to RNA to protein.

- **Key processes:**
  - **Replication (DNA to DNA):** The process by which DNA makes identical copies of itself. This ensures that genetic information is faithfully passed from one generation to the next.
  - **Transcription (DNA to RNA):** The process by which the genetic information encoded in a DNA strand is transcribed into an RNA molecule. This step is catalyzed by RNA polymerase. There are different types of RNA (mRNA, tRNA, rRNA), each with specific functions.
  - **Translation (RNA to Protein):** The process by which the genetic information encoded in messenger RNA (mRNA) is decoded to synthesize proteins. This occurs on ribosomes, with the help of transfer RNA (tRNA) molecules that carry specific amino acids.
- **Exceptions/Modifications to the basic dogma:**
  - **Reverse Transcription (RNA to DNA):** Some viruses (retroviruses like HIV) use reverse transcriptase to synthesize DNA from an RNA template.
  - **RNA Replication (RNA to RNA):** Some RNA viruses replicate their RNA directly using RNA-dependent RNA polymerase.
  - **Direct protein synthesis from DNA (not common or widespread):** While theoretically possible, it's not a major biological pathway.
  - **Prions:** Misfolded proteins that can induce other proteins to misfold, representing information flow from protein to protein, though not involving nucleic acids directly.

- Despite these exceptions, the Central Dogma remains a foundational concept, illustrating the primary direction of genetic information flow essential for life.

- (e) Transcription termination in eukaryotes
  - Transcription termination in eukaryotes is more complex and less well-understood than in prokaryotes, and it does not typically involve specific terminator sequences recognized directly by RNA polymerase. It varies depending on the type of RNA polymerase (I, II, or III) involved.
  - **RNA Polymerase I (rRNA synthesis):**
    - Termination signals are specific DNA sequences recognized by a termination factor (e.g., TTF-1 in mammals).
    - The termination factor binds to this sequence downstream of the rRNA gene and causes RNA Pol I to dissociate.
  - **RNA Polymerase II (mRNA, snRNA, miRNA synthesis):**
    - Transcription by RNA Pol II often proceeds past the coding sequence of the gene.
    - Termination is coupled with **3'-end processing** of the mRNA, which involves cleavage and polyadenylation.
    - The primary transcript is cleaved at a specific site downstream of the polyadenylation signal (consensus sequence AAUAAA).
    - This cleavage leads to the dissociation of RNA Pol II from the DNA template.
    - Further termination can involve a "torpedo" model, where an exonuclease (e.g., Xrn2 in mammals) degrades the unprotected RNA downstream of the cleavage site,

eventually catching up to and dissociating the RNA polymerase.

- **RNA Polymerase III (tRNA, 5S rRNA, some snRNA synthesis):**
  - RNA Pol III terminates transcription after transcribing a series of U's (poly-U stretch) in the template DNA, similar to rho-independent termination in prokaryotes.
  - This poly-U stretch forms a weak RNA-DNA hybrid, allowing the polymerase to dissociate without the help of additional protein factors.

4 (a) With a well labeled diagram, discuss the mechanism of initiation in prokaryotic translation and compare it with that of eukaryote. (8)

- **Mechanism of Initiation in Prokaryotic Translation (with diagram):**
  - [**Diagram of Prokaryotic Translation Initiation**]
    - **Components:** 30S ribosomal subunit, mRNA, fMet-tRNAfMet (initiator tRNA carrying N-formylmethionine), Initiation Factors (IF1, IF2, IF3), GTP.
    - **Steps:**
      1. **Small Subunit Binding:** The 30S ribosomal subunit, associated with IF1 and IF3, binds to the mRNA. IF3 prevents the premature association of the 50S subunit.
      2. **Shine-Dalgarno Interaction:** The 16S rRNA within the 30S subunit recognizes and base-pairs with the **Shine-Dalgarno sequence** (a

purine-rich sequence 6-8 bases upstream of the AUG start codon) on the mRNA. This positions the AUG start codon correctly at the P-site of the ribosome.

3. **Initiator tRNA Binding:** IF2 (bound to GTP) facilitates the binding of the initiator tRNA, fMet-tRNAfMet, to the AUG start codon at the P-site. IF1 binds to the A-site, preventing tRNA from binding there prematurely.

4. **Large Subunit Association:** GTP is hydrolyzed by IF2, leading to the release of all initiation factors (IF1, IF2, IF3). The 50S ribosomal subunit then associates with the 30S subunit, forming the complete 70S initiation complex.

5. **P-site Occupancy:** The initiator fMet-tRNAfMet is now positioned in the P-site, and the A-site is ready to accept the next aminoacyl-tRNA.

- **Comparison of Prokaryotic and Eukaryotic Translation Initiation:**

Feature	Prokaryotic Translation Initiation	Eukaryotic Translation Initiation
Ribosome Size	70S (30S + 50S subunits)	80S (40S + 60S subunits)

Feature	Prokaryotic Translation Initiation	Eukaryotic Translation Initiation
Initiator Amino Acid	N-formylmethionine (fMet)	Methionine (Met)
Initiator tRNA	fMet-tRNAfMet	Met-tRNAiMet
mRNA Recognition	**Shine-Dalgarno sequence** (16S rRNA interaction)	5' cap binding (eIF4E), scanning from 5' end, **Kozak sequence**
Start Codon Position	AUG positioned directly at P-site by Shine-Dalgarno	Ribosome scans from 5' cap to find the first AUG (often in Kozak context)
Initiation Factors	Fewer (IF1, IF2, IF3)	Many (eIF1, eIF2, eIF3, eIF4F complex, eIF5, eIF6, etc.)
mRNA Structure	Often polycistronic (multiple ORFs)	Always monocistronic (one ORF per mRNA)
Cyclization of mRNA	No	Yes, via interactions between eIF4E (cap) and poly(A) binding protein (PABP)
Energy Source	GTP	GTP, ATP (for scanning)

- Export to Sheets

 (b) Explain the salient features of genetic code. (5)

- 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.

- **Salient Features:**
    - **1. Triplet Code:** Each amino acid is specified by a sequence of three nucleotides, called a codon. There are 43=64 possible codons.
    - **2. Degenerate (Redundant):** Most amino acids are specified by more than one codon. For example, Leucine is encoded by six different codons (UUA, UUG, CUU, CUC, CUA, CUG). This redundancy provides some protection against the effects of point mutations.
    - **3. Unambiguous (Specific):** Each codon specifies only one particular amino acid. A given codon will always code for the same amino acid.
    - **4. Non-overlapping:** The codons are read sequentially, one after another, without any overlap. Each nucleotide is part of only one codon.
    - **5. Commaless (No Punctuation):** There are no intervening nucleotides or "commas" between successive codons. The sequence is read continuously from a fixed starting point.
    - **6. Universal (Nearly Universal):** The genetic code is largely universal across all forms of life, from bacteria to humans, meaning the same codons specify the same amino acids. This universality strongly supports the idea of a common ancestor for all life. (However, there are a few minor exceptions, particularly in mitochondrial genomes and some single-celled organisms).
    - **7. Start Codon:** The codon AUG serves as the primary start signal for translation and also codes for Methionine (Met). In prokaryotes, it codes for N-formylmethionine (fMet).

- **8. Stop Codons (Nonsense Codons):** Three codons (UAA, UAG, UGA) do not code for any amino acid. They act as stop signals, terminating protein synthesis.

 (c) Name two unusual bases present in tRNA. (2)

- Two unusual (modified) bases present in tRNA are:
  - **Pseudouridine (Ψ)**
  - **Dihydrouridine (D)**
  - (Other examples include Inosine (I), Methylguanosine, Ribothymidine, etc.)

5 Attempt any two of the following: (7.5x2=15)

- (a) With the help of a well labeled diagram, explain the mechanism of RNA interference.
  - [**Diagram of RNA Interference (RNAi) Mechanism**]
    - **Key Components:** Double-stranded RNA (dsRNA), Dicer, Argonaute (AGO) protein, RNA-induced silencing complex (RISC), target mRNA.
    - **Overall Mechanism:** RNA interference (RNAi) is a biological process in which RNA molecules inhibit gene expression or translation, by neutralizing targeted mRNA molecules. It plays a crucial role in gene regulation, antiviral defense, and maintaining genome stability.
  - **Steps of RNA Interference:**
    - **1. dsRNA Processing (Dicer Cleavage):**
      - The process begins with the presence of long double-stranded RNA (dsRNA) molecules in the cell. These dsRNAs can be derived from viral replication, transposable elements, or intentionally introduced in experiments.

- An enzyme called **Dicer**, which is an RNase III-type endonuclease, recognizes and cleaves these long dsRNAs into shorter, double-stranded fragments of approximately 21-25 nucleotides. These small dsRNAs are called **small interfering RNAs (siRNAs)**.
- (Alternatively, in the case of **microRNAs (miRNAs)**, a primary miRNA transcript (pri-miRNA) is processed by Drosha and Pasha in the nucleus to form a pre-miRNA, which is then exported to the cytoplasm and further processed by Dicer into a mature miRNA duplex.)

- **2. RISC Loading:**
  - The siRNA (or miRNA) duplex is then loaded into a multiprotein complex called the **RNA-induced silencing complex (RISC)**.
  - During this loading, one of the two strands of the siRNA/miRNA duplex (the "passenger strand") is usually degraded and ejected from the RISC.
  - The remaining single strand, known as the "guide strand" (or mature siRNA/miRNA), remains associated with the RISC complex, specifically with the **Argonaute (AGO) protein**, which is the catalytic core of RISC.
- **3. Target mRNA Recognition and Silencing:**
  - The guide strand in the RISC complex acts as a template to search for and identify complementary messenger RNA (mRNA) molecules in the cytoplasm.

- **siRNA Pathway (Perfect Complementarity):** If there is perfect or nearly perfect complementarity between the guide siRNA and the target mRNA, the Argonaute protein (specifically, its slicer activity) cleaves the target mRNA, leading to its rapid degradation. This prevents the mRNA from being translated into protein, effectively silencing the gene.
- **miRNA Pathway (Imperfect Complementarity):** If there is imperfect complementarity (as is often the case with miRNAs), the RISC complex does not typically cleave the mRNA. Instead, it inhibits translation of the target mRNA or promotes its deadenylation and eventual degradation by other cellular machinery.

- **4. Gene Silencing:** The result is the reduction or complete inhibition of protein production from the targeted gene, leading to post-transcriptional gene silencing.

- (b) How did Fraenkel-Conrat prove that RNA is genetic material in some viruses?
  - The Hershey-Chase experiment (1952) had already established DNA as the genetic material in bacteriophages. However, the question remained for viruses that contain RNA but no DNA. The experiment by Heinz Fraenkel-Conrat and Beatrice Singer in 1956 provided crucial evidence that RNA can serve as the genetic material in some viruses, specifically using **Tobacco Mosaic Virus (TMV)**.
  - **Background on TMV:** TMV is an RNA virus composed of an RNA genome surrounded by a protein coat.
  - **Experimental Design:**

- **Step 1: Separation of RNA and Protein:** Fraenkel-Conrat and Singer developed methods to separate the RNA genome from the protein coat of TMV. They also had different strains of TMV (e.g., strain A and strain B) that produced distinct types of lesions on tobacco leaves and had differences in their protein composition.
- **Step 2: Reconstitution of Hybrid Viruses:**
  - They then mixed the RNA from one strain (e.g., strain A RNA) with the protein coat from another strain (e.g., strain B protein) to create **reconstituted hybrid TMV particles**.
  - They also created the reciprocal hybrid (strain B RNA with strain A protein).
  - As controls, they used the intact original viruses, and also RNA alone and protein alone.
- **Step 3: Infection of Tobacco Plants:** The reconstituted hybrid viruses, as well as the isolated RNA and protein, were used to infect tobacco plants.
- **Step 4: Observation of Progeny Viruses:** After infection, they observed the characteristics of the progeny viruses produced in the infected plants.

- **Results:**
  - Infection with protein alone or RNA alone (without being reconstituted) did not produce viable progeny viruses or caused very little infection. This showed that both components are generally needed for infection.
  - The crucial observation was that the progeny viruses produced by the **hybrid viruses always matched the strain from which the RNA was obtained, not the protein coat.**

        - For example, if a hybrid virus was made with RNA from strain A and protein from strain B, the progeny viruses produced in the infected plant were all characteristic of strain A (e.g., produced lesions typical of strain A, and had protein coats identical to strain A).

    - **Conclusion:** This experiment definitively demonstrated that the genetic information responsible for directing the synthesis of new viral components (both RNA and protein) resided in the RNA molecule, not the protein coat. Therefore, RNA is the genetic material in TMV.

- (c) What are consensus sequences? Explain them with reference to prokaryotic and eukaryotic promoter regions.

    - **Consensus Sequences:**

        - A consensus sequence is a calculated sequence of DNA, RNA, or protein that represents the most common nucleotide or amino acid at each position in a stretch of aligned homologous sequences.

        - It is not a perfectly identical sequence but rather a "statistical" sequence that reflects the most frequently occurring bases/amino acids, indicating functionally important regions.

        - Deviations from the consensus sequence often lead to reduced efficiency or altered function of the molecular process (e.g., transcription initiation or splicing).

    - **Consensus Sequences in Prokaryotic Promoter Regions:**

        - Prokaryotic gene expression is primarily regulated at the transcriptional level. RNA polymerase recognizes specific DNA sequences, known as promoter regions, to initiate transcription.

- There are two main consensus sequences upstream of the transcription start site (+1) in prokaryotic promoters:
    - **-35 region (Recognition Sequence):** This sequence is approximately 35 base pairs upstream of the transcription start site and has the consensus sequence **TTGACA**. It is recognized by the sigma (σ) subunit of the RNA polymerase holoenzyme, which helps in initial binding and recognition of the promoter.
    - **-10 region (Pribnow Box):** This sequence is approximately 10 base pairs upstream of the transcription start site and has the consensus sequence **TATAAT**. This region is AT-rich, making it easier to unwind the DNA strands for transcription initiation, as it is where the DNA helix begins to separate to form the open complex.
- The spacing between these two elements is also critical (typically 17-19 base pairs). The closer a promoter sequence matches these consensus sequences, the stronger its binding affinity for RNA polymerase and the higher the rate of transcription.

- **Consensus Sequences in Eukaryotic Promoter Regions:**
    - Eukaryotic transcription is more complex, involving three different RNA polymerases (Pol I, Pol II, Pol III), each with distinct promoter elements. Here, we focus on RNA Polymerase II promoters, which transcribe protein-coding genes.
    - Eukaryotic RNA Pol II promoters often have several core promoter elements, which are recognized by general transcription factors (GTFs) before RNA polymerase II can bind.

- **TATA Box:** This is the most common and well-studied core promoter element, typically located around -25 to -30 base pairs upstream of the transcription start site. Its consensus sequence is **TATA(A/T)A(A/T)**. It is recognized by the TATA-binding protein (TBP), a subunit of the TFIID complex, which is crucial for initiating the assembly of the pre-initiation complex.
- **Initiator Element (Inr):** Found at the transcription start site (around -2 to +4), the consensus for Inr is often **YYAN(T/A)YY** (where Y is pyrimidine, N is any nucleotide). It is important for accurate initiation in the absence or presence of a TATA box.
- **Downstream Promoter Element (DPE):** Located downstream of the transcription start site (around +28 to +34), this element has the consensus sequence **R(A/G)G(A/T)(C/T)G**. It is often found in TATA-less promoters and is recognized by TFIID.
- **TFIIB Recognition Element (BRE):** Found immediately upstream of the TATA box (around -37 to -32), with consensus **G/C G/C G/A C G/C C**. It is recognized by transcription factor TFIIB.

- The specific combination and spacing of these elements determine the strength and regulation of eukaryotic promoters.

6 (a) Discuss the mechanism of transcription termination in prokaryotes. (8)

- Transcription termination in prokaryotes is the process by which RNA polymerase stops synthesizing RNA and dissociates from the DNA template. There are two primary mechanisms of termination in

prokaryotes: Rho-independent (intrinsic) termination and Rho-dependent termination.

- **1. Rho-Independent (Intrinsic) Termination:**
    - This mechanism relies solely on specific sequences within the newly synthesized RNA molecule and does not require the Rho protein.
    - **Mechanism:**
        - **a. Hairpin Loop Formation:** The terminator sequence in the DNA template is transcribed into an RNA sequence that contains an inverted repeat. This inverted repeat allows the newly synthesized RNA to fold back on itself and form a stable **hairpin loop** (stem-loop structure) through complementary base pairing. This hairpin is typically rich in G-C base pairs, making it thermodynamically stable.
        - **b. Poly-Uracil Stretch:** Immediately downstream of the inverted repeat, the RNA sequence contains a stretch of approximately 6-8 uridine residues (poly-U tail). This poly-U tail corresponds to a poly-adenine stretch in the template DNA.
        - **c. Dissociation:** When RNA polymerase encounters the hairpin loop, its movement is momentarily paused. The weak hydrogen bonds between the poly-U tail in the RNA and the poly-A stretch in the DNA template are not strong enough to hold the RNA-DNA hybrid together during this pause, especially with the pull from the translocating polymerase. The RNA molecule then dissociates from the DNA template and the RNA polymerase detaches, ending transcription.
- **2. Rho-Dependent Termination:**

- This mechanism requires the involvement of a protein called **Rho (ρ) factor**, which is an ATP-dependent helicase.
- **Mechanism:**
  - **a. Rho Binding Site (Rut Site):** The DNA template contains a specific sequence called the **Rho utilization site (rut site)**, which is transcribed into an untranslated, C-rich, G-poor sequence in the nascent RNA molecule.
  - **b. Rho Loading and Translocation:** The Rho factor binds to this rut site on the nascent RNA. Rho is a hexameric protein that acts as an RNA helicase. It uses ATP hydrolysis to move along the RNA molecule in a 5' to 3' direction, following behind the RNA polymerase.
  - **c. RNA Polymerase Pause:** RNA polymerase pauses at a specific Rho-dependent terminator sequence on the DNA, which often lacks a strong hairpin structure. This pause is crucial as it allows Rho to catch up.
  - **d. Unwinding and Dissociation:** Once Rho reaches the paused RNA polymerase, its helicase activity unwinds the RNA-DNA hybrid within the transcription bubble. This unwinding disrupts the association between the RNA transcript and the DNA template, leading to the release of the RNA polymerase and the newly synthesized RNA molecule, thus terminating transcription.

 (b) With a well labeled diagram, discuss end replication in eukaryotes.

- **End Replication Problem in Eukaryotes:**
  - Eukaryotic chromosomes are linear, unlike circular prokaryotic chromosomes. This linearity presents a unique challenge during DNA replication known as the "end replication problem."

- During DNA replication, DNA polymerase synthesizes DNA in the 5' to 3' direction and requires an RNA primer to initiate synthesis.
- On the **leading strand**, replication proceeds continuously to the end of the chromosome.
- On the **lagging strand**, DNA is synthesized as Okazaki fragments. When the RNA primer at the very 3' end of the lagging strand is removed, there is no upstream DNA fragment to provide a 3'-OH group for DNA polymerase to fill the resulting gap.
- Consequently, a small portion of the DNA at the 3' end of the lagging strand template (i.e., the 5' end of the newly synthesized strand) cannot be replicated. This leads to progressive shortening of the chromosome with each round of replication, which could eventually lead to loss of vital genetic information.

- **Telomeres and Telomerase: The Solution:**
  - Eukaryotic chromosomes have specialized repetitive DNA sequences at their ends called **telomeres**. These sequences (e.g., TTAGGG in humans) are non-coding and protect the genes from being lost during replication.
  - The enzyme **telomerase** solves the end-replication problem by extending the telomeric DNA. Telomerase is a unique ribonucleoprotein complex, meaning it contains both protein and an RNA component.
  - **RNA Component:** The RNA component of telomerase serves as a template for synthesizing the telomeric DNA repeats.
- **Mechanism of End Replication (with telomerase) (with diagram):**
  - [**Diagram of Telomere Replication**]

- **Components:** Linear eukaryotic chromosome, unreplicated lagging strand end, RNA primer, Telomerase enzyme (with its RNA template), dNTPs (deoxynucleotide triphosphates).
- **Steps:**
    - 1. **Primer Removal and Gap:** After the last RNA primer on the lagging strand is removed by RNase H, a gap remains at the 5' end of the newly synthesized lagging strand, leaving a single-stranded 3' overhang on the parental template strand.
    - 2. **Telomerase Binding:** Telomerase binds to this 3' overhang of the parental template strand. The RNA component of telomerase (e.g., 3'-AAUCCC-5' in humans) base-pairs with the telomeric DNA sequence (e.g., 5'-TTAGGG-3') of the overhang.
    - 3. **Telomere Elongation (Reverse Transcription):** Using its internal RNA molecule as a template, the telomerase enzyme (which is a reverse transcriptase) synthesizes new DNA repeats directly onto the 3' end of the parental template strand. It extends this strand, creating a longer 3' overhang (e.g., 5'-TTAGGG TTAGGG-3'). Telomerase can move along the DNA, repeatedly adding these short repeats.

4. **Primer Synthesis and Lagging Strand Elongation:** Once the 3' overhang is sufficiently extended, primase can synthesize a new RNA primer on this extended template.

5. **DNA Polymerase Synthesis:** DNA polymerase (specifically DNA Pol δ in eukaryotes) then uses this new primer to synthesize the complementary DNA strand, filling in the gap and extending the newly synthesized lagging strand.

6. **Primer Removal:** The final RNA primer is removed. Although a short gap may still remain at the very end of the newly synthesized lagging strand, the overall chromosome length is maintained because the previously lost sequences have been replaced by the non-coding telomeric repeats.