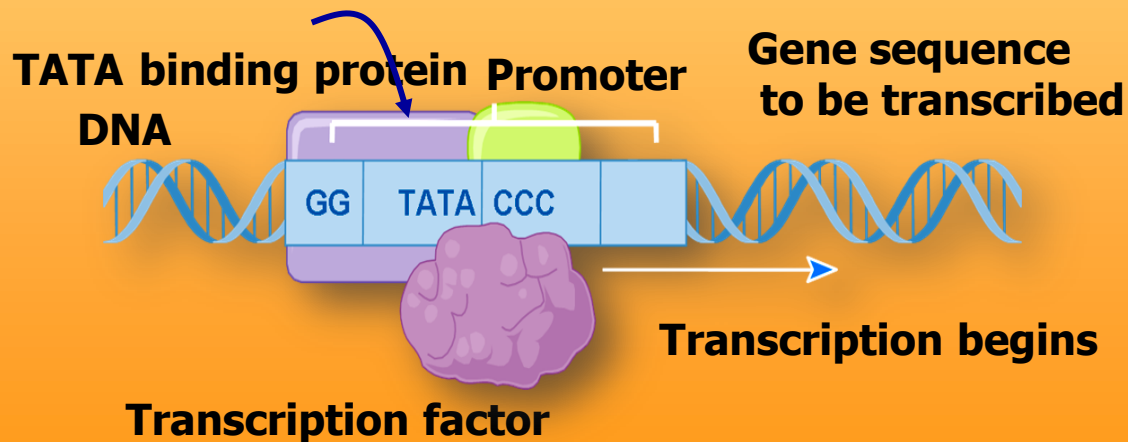
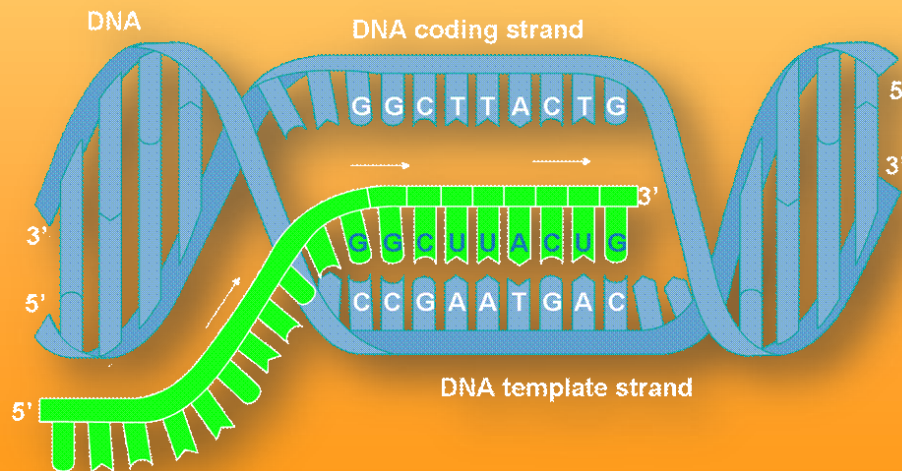


TRANSCRIPTION



Transcription

- Transfer of the information from DNA to RNA is called transcription, and the molecule that encodes one or more polypeptides (proteins) is called messenger RNA (mRNA).
- The principle of complementarity governs the process of transcription, except the adenosine now forms base pair with uracil instead of thymine.





Transcription

- However, unlike in the process of replication, which once set in, the total DNA of an organism gets duplicated, in transcription only a segment of DNA and only one of the strands is copied into RNA.
- This necessitates defining the boundaries that would demarcate the region and the strand of DNA that would be transcribed.
- It is a hetero-catalytic function of DNA, which involves transfer of coded information from DNA to RNA through the synthesis of RNA over the template of DNA.
- The mechanism of RNA synthesis was worked out in the late 1950's by the American investigators Jerard Hurwitz, Samuel B. Weiss, and Audrey Stevens by independent in vitro experiments.



Transcription contd...

- Both RNA and DNA are nucleic acids, which use base pairs of nucleotides as a complementary language that can be converted back and forth from DNA to RNA in the presence of the correct enzymes.
- During transcription, a DNA sequence is read by RNA polymerase, which produces a complementary, antiparallel RNA strand.
- As opposed to DNA replication, transcription results in an RNA complement that includes uracil (U) in all instances where thymine (T) would have occurred in a DNA complement.
- Transcription is the first step leading to gene expression. The stretch of DNA transcribed into an RNA molecule is called a transcription unit and encodes at least one gene.



Transcription contd...

- **Transcription involves rewriting the genetic message coded in DNA into an RNA molecule.**
- **Transcription occurs in the nucleus during the G1 and G2 phases of cell cycle.**
- **In eukaryotes, transcription occurs inside the nucleus and the transcription products move out into cytoplasm for translation.**
- **In prokaryotes, transcription occurs in contact with the cytoplasm as their DNA lies in the cytoplasm.**
- **As DNA contains all the hereditary information, it is called master copy of genetic information.**
- **It replicates its carbon copies when new master copies are required.**
- **Normally it forms working copies in the form of mRNA.**



Transcription contd...

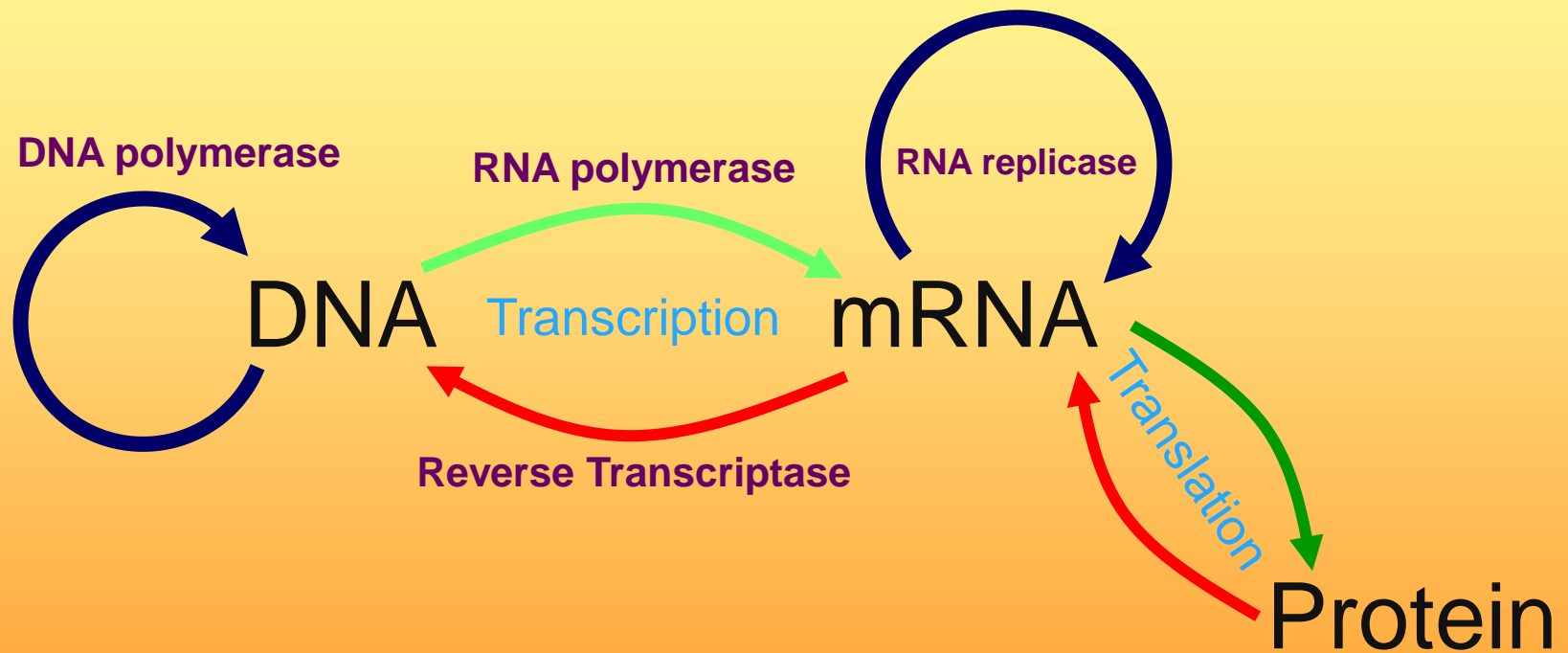
- Transcription requires a DNA template, activated precursors, a divalent metal ion, and RNA polymerase.
- Both the strands of DNA do not transcribe RNA but only one of them called mater strand or template or sense strand does it.
- The segment of DNA involved in transcription is called Cistron, in functional unit of gene.
- Transcribing segment i.e., Cistron has a promoter region where initiation of transcription begins, and a terminator region where transcription ends.
- Promoter has different parts for attachment to various transcription factors. In many cases, the promoter has an AT rich sequence called TATA box.



Transcription contd...

- The area has groove to which specific protein components can combine.
- TATA containing region is also called Pribnow box after the name of its discoverer.
- Enzyme required for transcription is RNA polymerase (Chamberlain and Berg, 1962; isolated from *E.coli*).
- Eukaryotic RNA polymerases also require transcription factors for initiation.
- The entire enzyme (holozyme) consists of a core enzyme and a sigma (σ) factors.
- The core enzyme consists of four polypeptide chains, β , β , and ω .

Central dogma





Transcription contd...

- The holozyme has a molecular weight 4.5×10^5 . RNA polymerase not only initiates but also extends the RNA (chain elongation) and functions always in 5' and 3' direction.
- Sigma (σ) factor recognizes the start signal or promoter region of DNA.
- If the gene transcribed encodes for a protein, the result of transcription is messenger RNA (mRNA), which will then be used to create that protein via the process of translation.
- Alternatively, the transcribed gene may encode for either ribosomal RNA (rRNA) or transfer RNA (tRNA), other components of the protein-assembly process, or other ribozymes.



Transcription contd...

- **Prokaryotes have only one type of RNA polymerase, which synthesizes all types of RNA.**
- **Eukaryotes have three types of RNA polymerases, I for rRNA, II for mRNA and III for tRNAs.**
- **A specific protein called rho factor (PR) is required for termination of transcription.**
- **In many bacteria, genes of related functions are grouped together in operons.**
- **An operon acts as single transcription unit and thus produces polycistronic mRNA. In eukaryotes, only monocistronic mRNAs are generally produced.**
- **Actinomycin D prevents transcription.**



Transcription contd...

- **First, if both strands act a template, they would code for RNA molecule with different sequences and in turn, if they code for proteins, the sequence of amino acids in the proteins would be different.**
- **One segment of the DNA would be coding for two different proteins, and this would complicate the genetic information transfer machinery.**
- **Second, the two RNA molecules if produced simultaneously would be complementary to each other, hence would form a double stranded RNA.**
- **Which prevents RNA from being translated into protein and the exercise of transcription would become a futile one.**



Transcription Unit

- **A transcription unit in DNA is defined primarily by the three regions in the DNA:**
 - **A Promoter**
 - **The Structural gene**
 - **A Terminator**
- **A DNA transcription unit encoding for a protein contains not only the sequence that will eventually be directly translated into the protein but also regulatory sequences that direct and regulate the synthesis of that protein.**
- **The regulatory sequence before (upstream from) the coding sequence is called the five prime untranslated region (5'UTR), and the sequence following (downstream from) the coding sequence is called the three prime untranslated region (3'UTR).**



Transcription Unit contd...

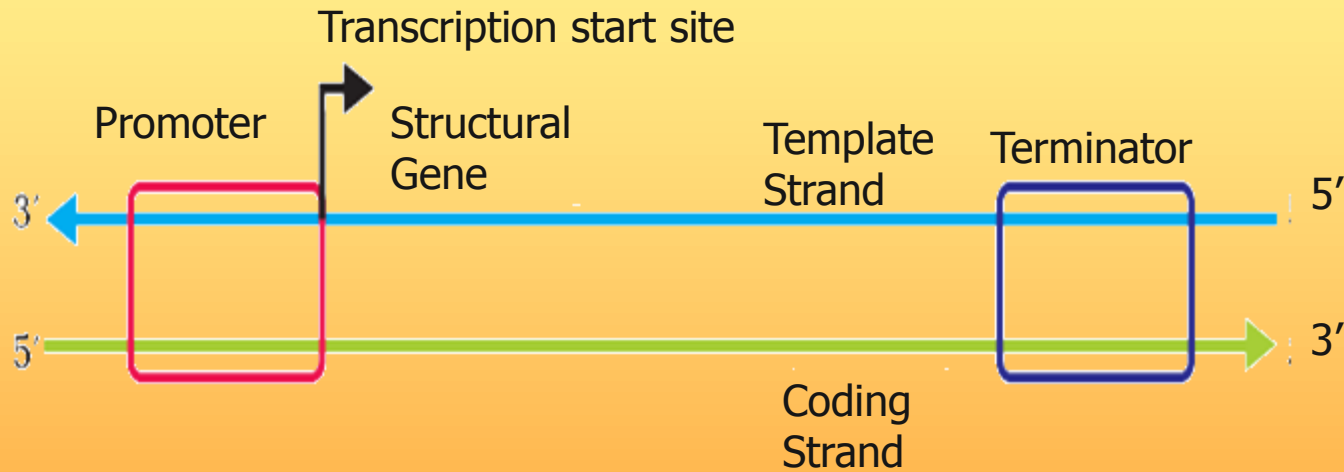
- **Transcription has some proofreading mechanisms, but they are fewer and less effective than the controls for copying DNA; therefore, transcription has a lower copying fidelity than DNA replication.**
- **As in DNA replication, DNA is read from 3' → 5' during transcription.**
- **Meanwhile, the complementary RNA is created from the 5' → 3' direction.**
- **Although DNA is arranged as two antiparallel strands in a double helix, only one of the two DNA strands, called the template strand, is used for transcription.**
- **This is because RNA is only single-stranded, as opposed to double-stranded DNA.**



Transcription Unit

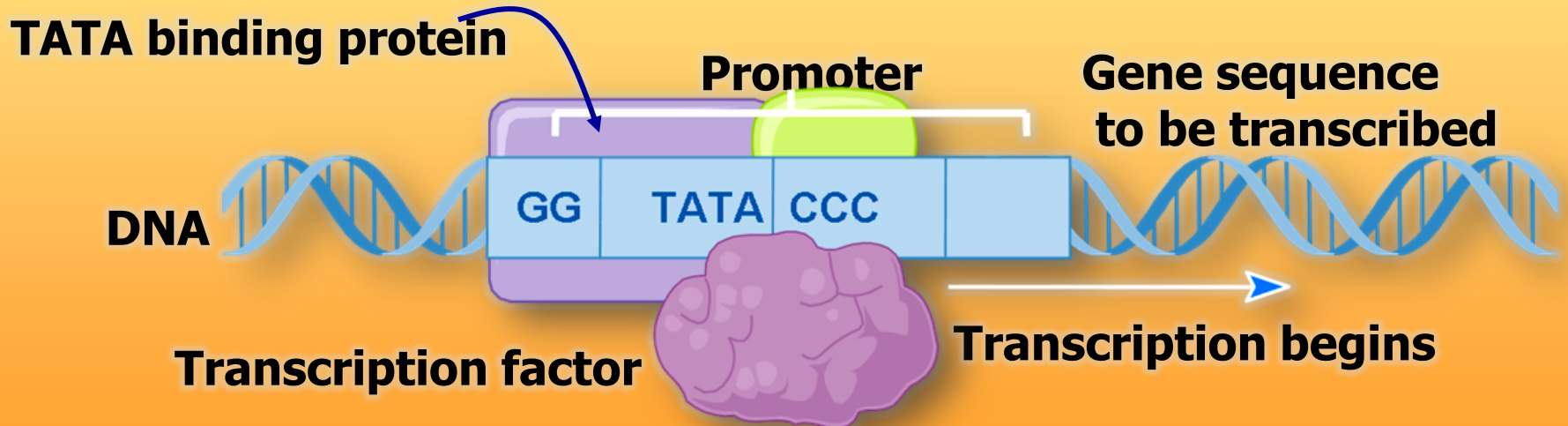
- The other DNA strand is called the coding strand, because its sequence is the same as the newly created RNA transcript.
- The use of only the 3' → 5' strand eliminates the need for the Okazaki fragments seen in DNA replication.
- Transcription is divided into 5 stages-
 - Pre-initiation,
 - initiation (at a promoter site) ,
 - promoter clearance,
 - elongation (3'→5' on template strand, $\approx 50 \text{ s}^{-1}$).
 - termination (at termination signal).

DNA Transcription

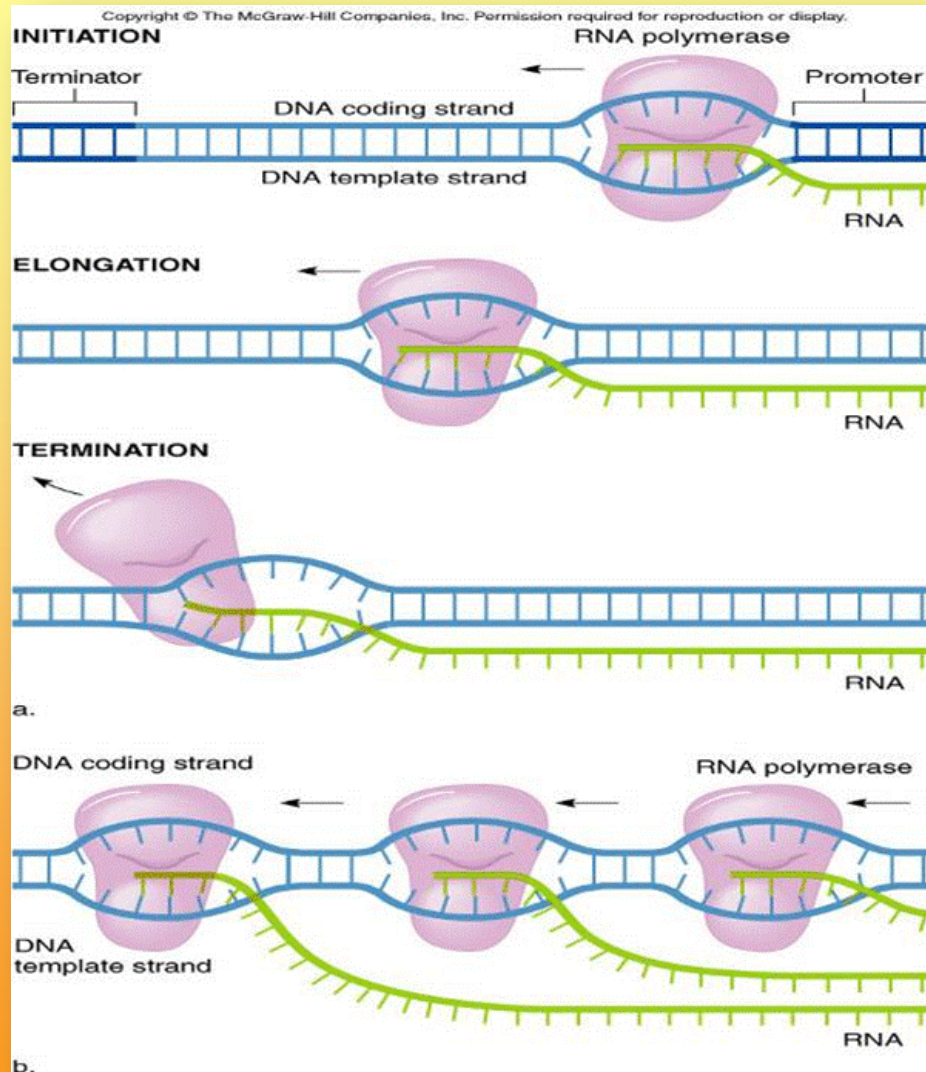


DNA transcription: initiation

- Transcription begins at the 3' end of the gene in a region called the promoter.
- The promoter recruits TATA protein, a DNA binding protein, which in turn recruits other proteins.



When a complete transcription complex is formed RNA polymerase binds and transcription begins.



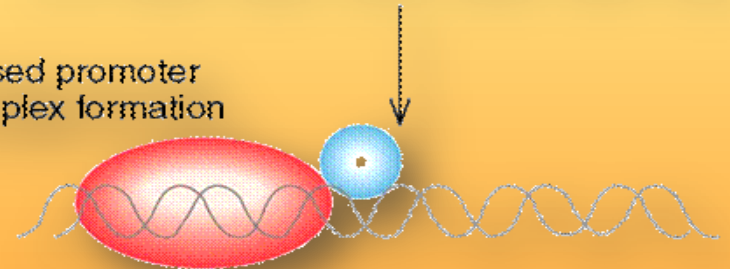
DNA transcription: initiation

- Sigma factor associates with -10 (TATA box) and -35 sequences.
- RNA polymerase binds.
- Bubble forms at -10→3.

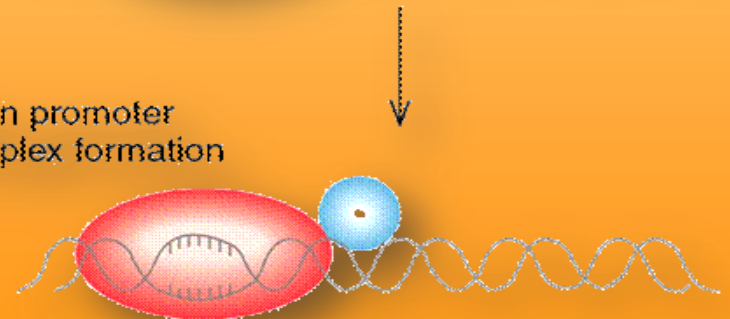
Promoter search



Closed promoter complex formation

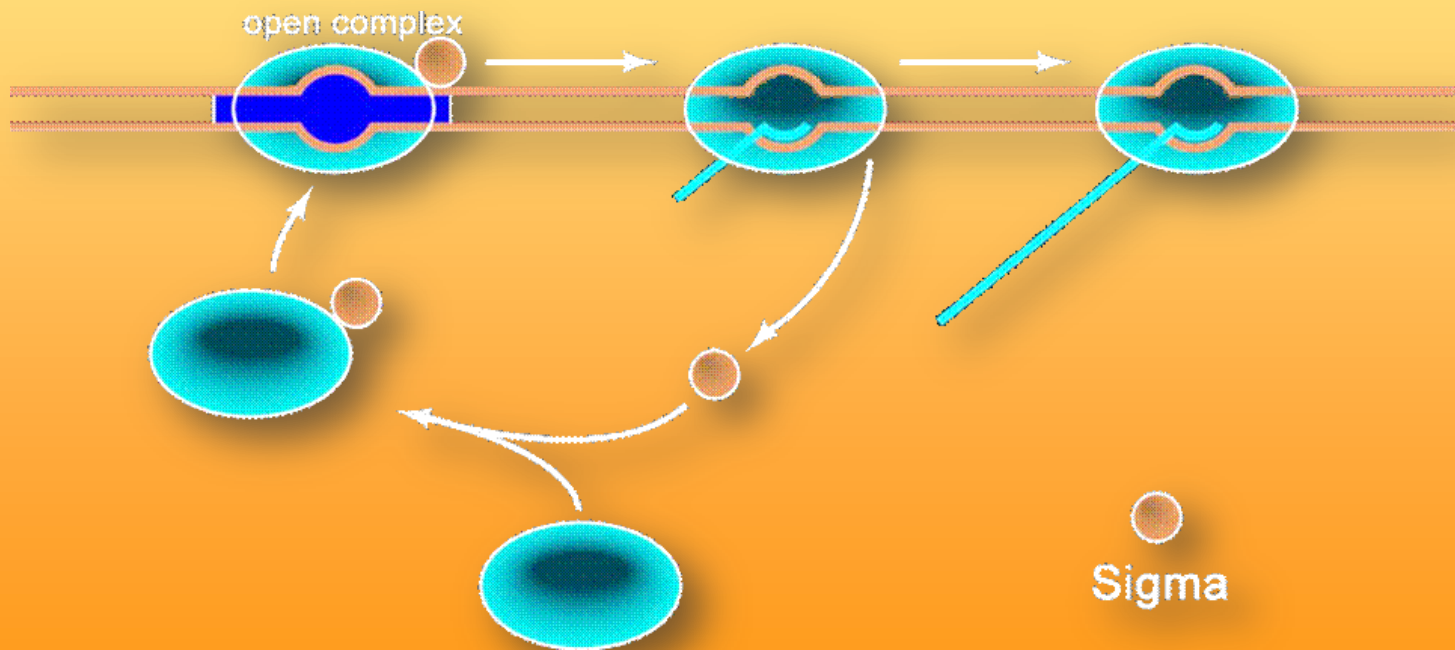


Open promoter complex formation



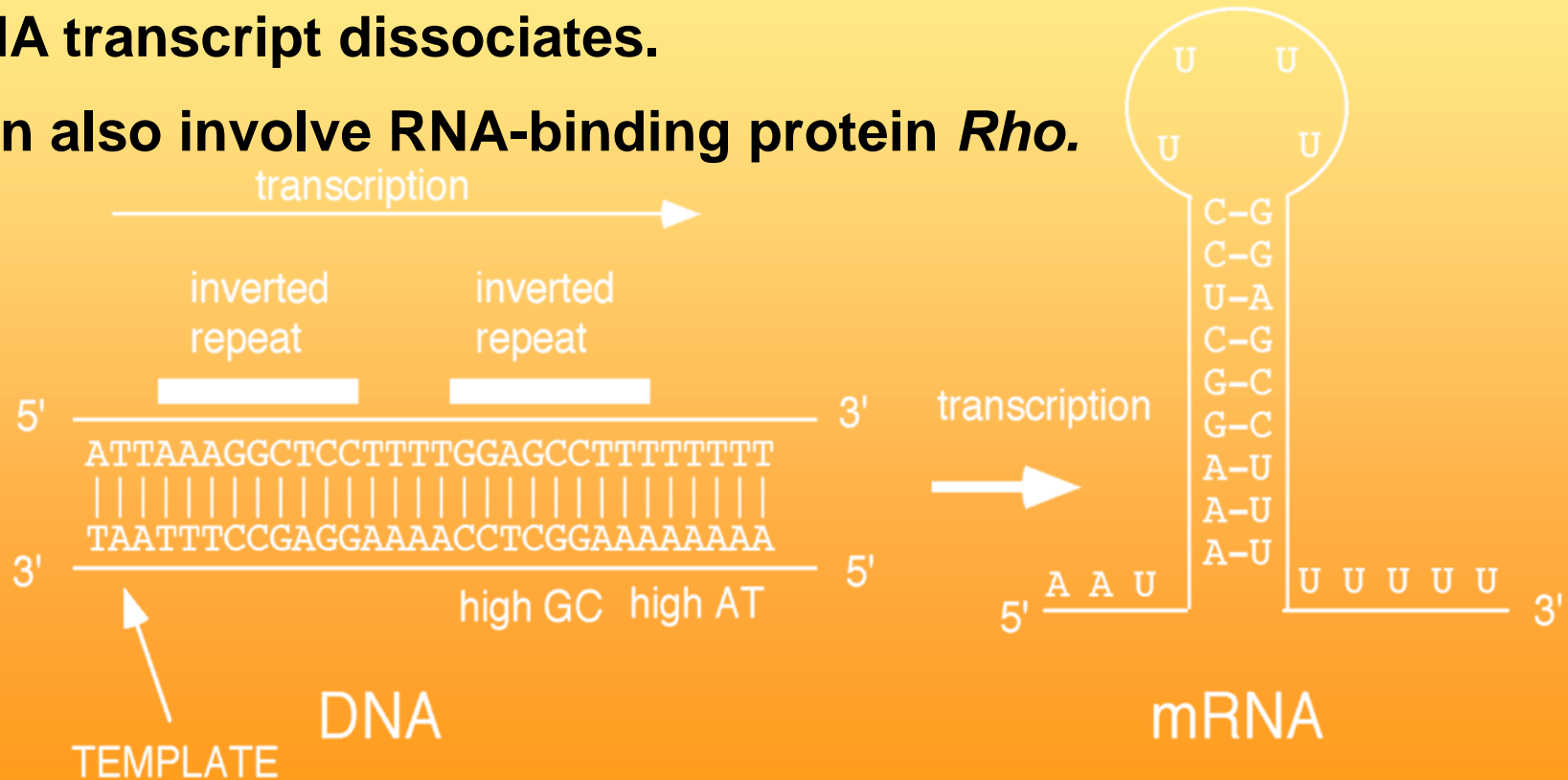
DNA transcription: elongation

- Form ≈ 10 bp RNA-DNA hybrid
- 5'-end of RNA dissociates
- σ factor dissociates and recycles

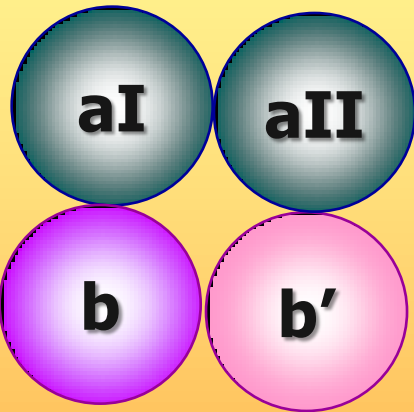


DNA transcription: termination

- Inverted repeat preceding A-rich region.
- Hairpin formation competes with RNA-DNA hybrid.
- RNA transcript dissociates.
- Can also involve RNA-binding protein *Rho*.



The assembly pathway of the core enzyme



CORE ENZYME
Sequence-independent,
nonspecific transcription
initiation

+



vegetative
(principal s)



heat shock
(for emergencies)

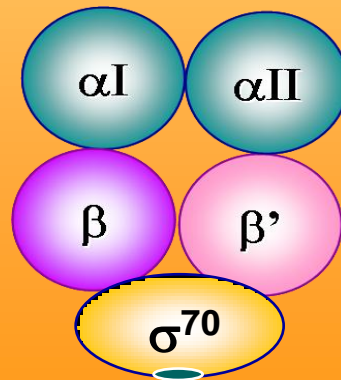


nitrogen starvation
(for emergencies)

SIGMA SUBUNIT
interchangeable,
promoter
recognition

In the Holoenzyme

- β' binds DNA.
- β binds NTPs.
- β and β' together make up the active site.
- α subunits appear to be *essential* for assembly and for activation of enzyme by regulatory proteins. They also bind DNA.
- σ recognizes promoter sequences on DNA.



RNAP HOLOENZYME - σ^{70}

Promoter-specific
transcription initiation



Binding of polymerase to Template DNA

- **Polymerase binds nonspecifically to DNA with low affinity and migrates, looking for promoter.**
- **Sigma subunit recognizes promoter sequence.**
- **RNA polymerase holoenzyme and promoter form "closed promoter complex" (DNA not unwound).**
- **Polymerase unwinds about 12 base pairs to form "open promoter complex".**



Chain Elongation Core polymerase - no sigma

- **Polymerase is pretty accurate - only about 1 error in 10,000 bases (not as accurate as DNAP III).**
- **Even this error rate is OK, since many transcripts are made from each gene.**
- **Elongation rate is 20-50 bases per second - slower in G/C-rich regions and faster elsewhere.**
- **Topoisomerases precede and follow polymerase to relieve supercoiling.**

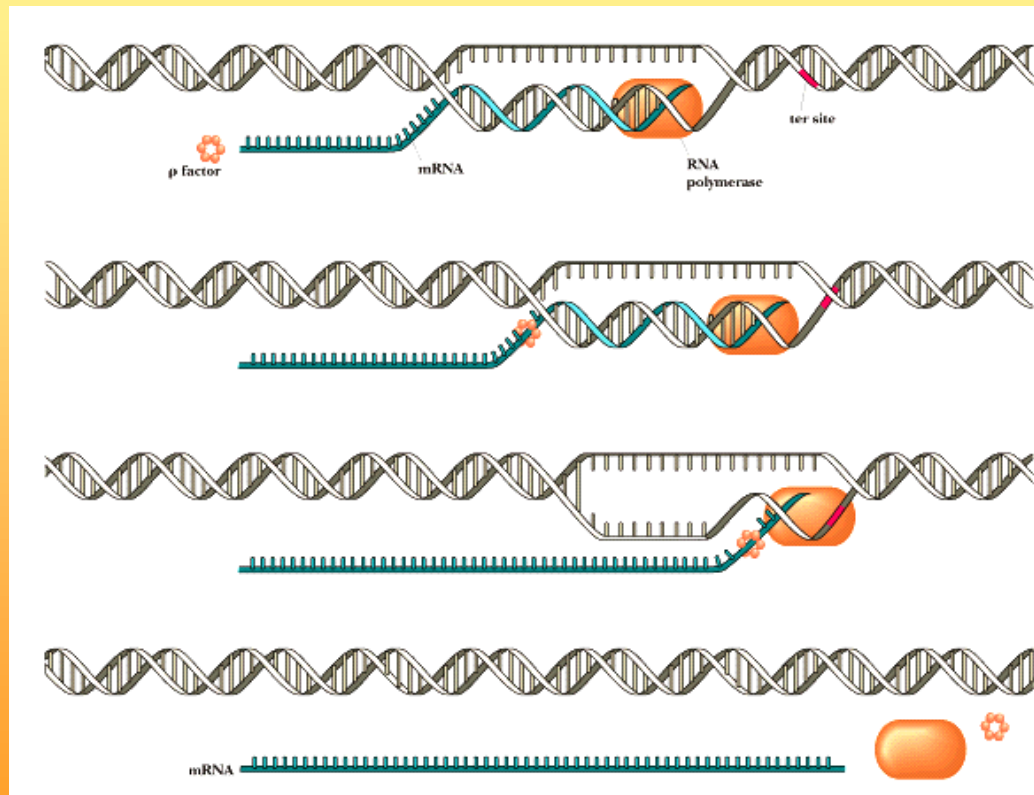


Chain Termination -Two mechanisms

1) Rho - the termination factor protein.

- ✓ **rho is an ATP-dependent helicase.**
- ✓ **it moves along the RNA transcript, finds the "bubble", unwinds it and releases the RNA chain.**

Rho-Dependent Transcription: Termination (depends on a protein and a *DNA* sequence)



Elongating complex is disrupted

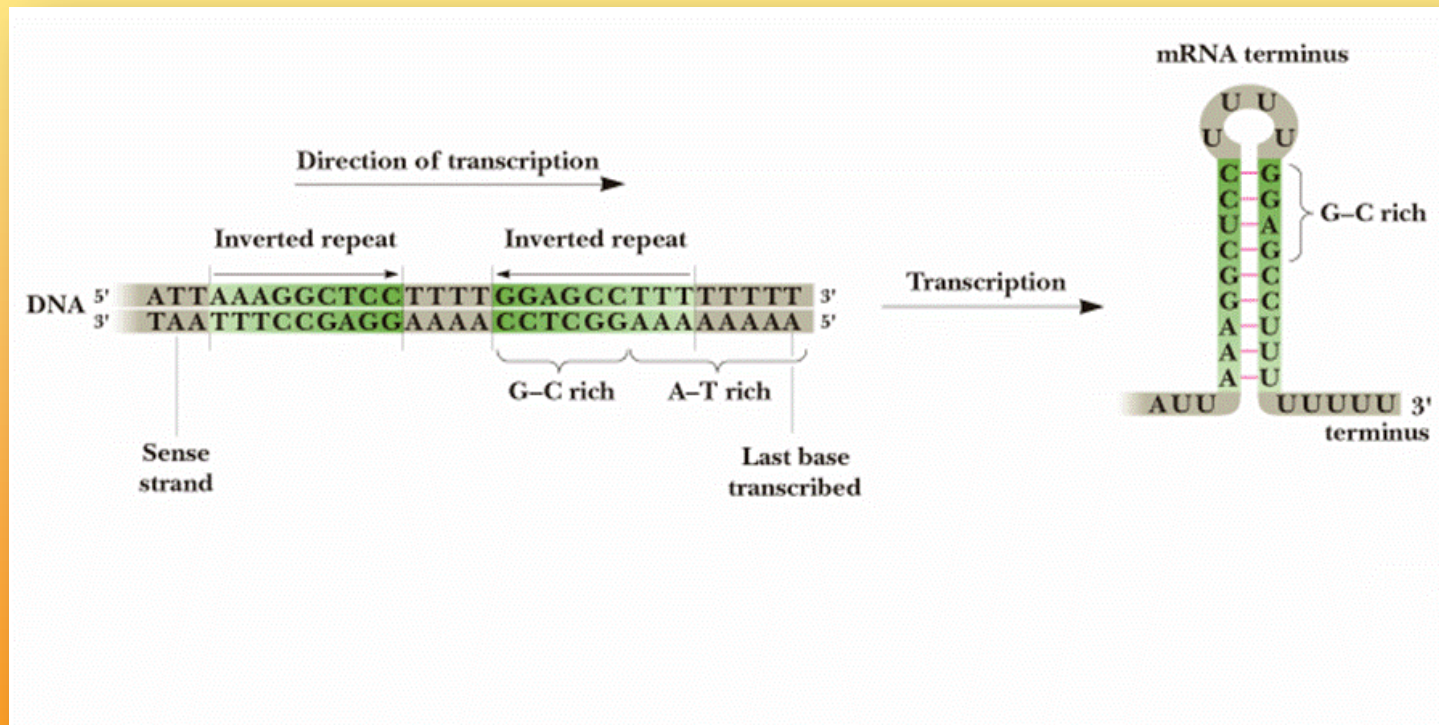


Chain Termination Two mechanisms

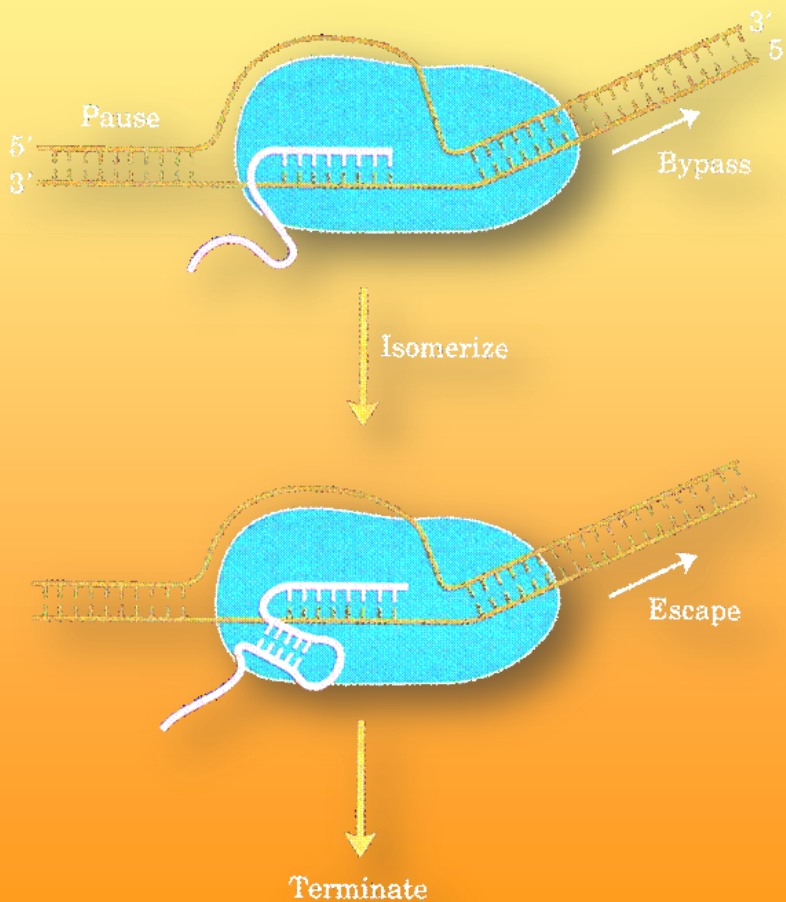
Rho-Independent

- **termination sites in DNA.**
- **Inverted repeat, rich in G:C, which forms a stem-loop in RNA transcript.**
- **6-8 A's in DNA coding for U's in transcript.**

Rho-Independent Transcription: Termination (depends on *DNA sequence* – *NOT* a protein factor)

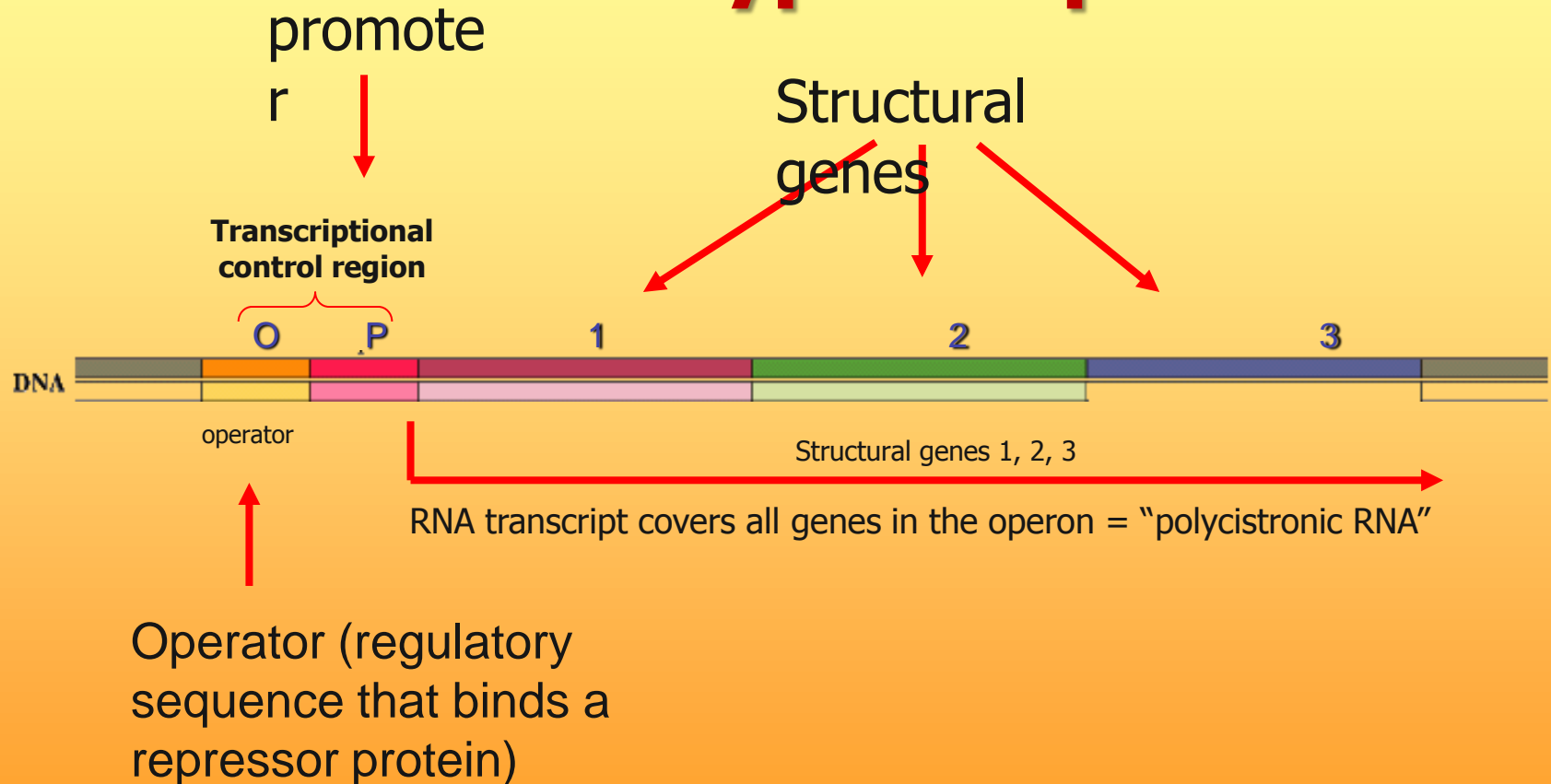


Rho-independent transcription termination



- RNAP pauses when it reaches a termination site.
- The pause may give the hairpin structure time to fold
- The fold disrupts important interactions between the RNAP and its RNA product
- The U-rich RNA can dissociate from the template
- The complex is now disrupted and elongation is terminated

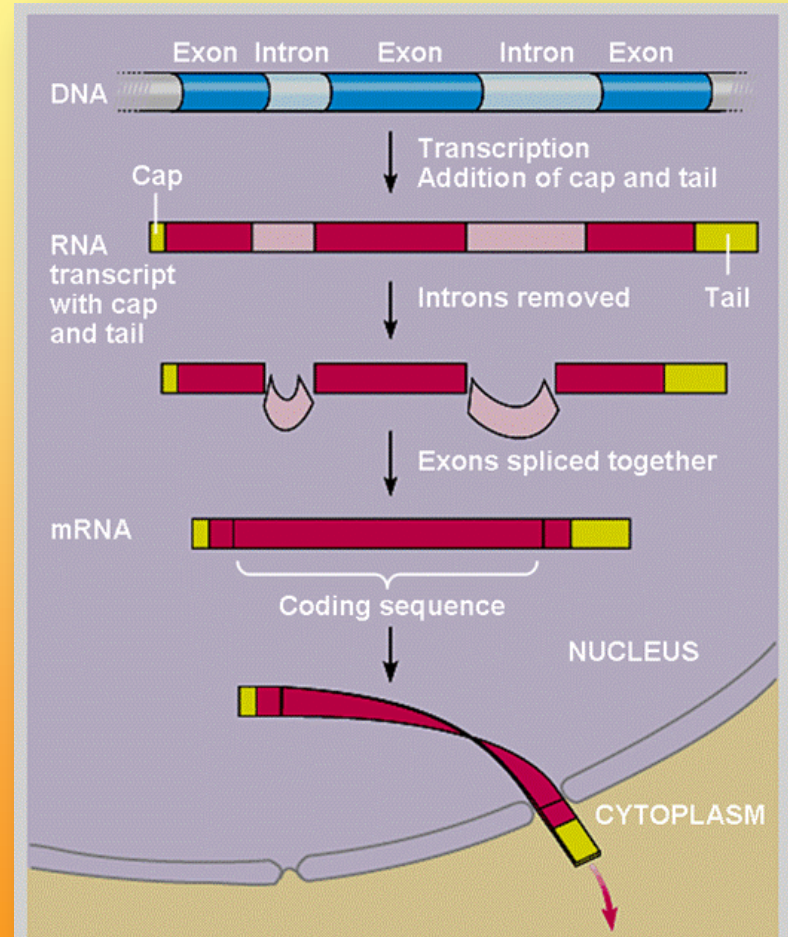
Architecture of a typical operon



By regulating a single promoter you can co-ordinate the expression of three genes (in this example)

Architecture of a typical operon

- Non-coding segments called introns are spliced out
- A cap and a tail are added to the ends to protect against degradation in the cytoplasm





Transcription Unit and the Gene

- **A gene is defined as the functional unit of inheritance.**
- **Though there is no ambiguity that the genes are located on the DNA, it is difficult to literally define a gene in terms of DNA sequence.**
- **The DNA sequence coding for tRNA or rRNA molecule also define a gene.**
- **However by defining a Cistron as a segment of DNA coding for a polypeptide, the structural gene in a transcription unit could be said as monocistronic (mostly in eukaryotes) or polycistronic (mostly in bacteria or prokaryotes).**



Transcription Unit and the Gene

- In eukaryotes, the monocistronic structural genes have interrupted coding sequences – the genes in eukaryotes are split.
- The coding sequences or expressed sequences are defined as exons.
- Exons are said to be those sequence that appear in mature or processed RNA.
- The exons are interrupted by introns. Introns or intervening sequences do not appear in mature or processed RNA.
- The split-gene arrangement further complicates the definition of a gene in terms of a DNA segment.
- Inheritance of a character is also affected by promoter and regulatory sequences of a structural gene.



Transcription Unit and the Gene

- In eukaryotes, the monocistronic structural genes have interrupted coding sequences – the genes in eukaryotes are split.
- The coding sequences or expressed sequences are defined as exons.
- Exons are said to be those sequence that appear in mature or processed RNA.
- The exons are interrupted by introns. Introns or intervening sequences do not appear in mature or processed RNA.
- The split-gene arrangement further complicates the definition of a gene in terms of a DNA segment.
- Inheritance of a character is also affected by promoter and regulatory sequences of a structural gene.
- Hence, sometime the regulatory sequences are loosely defined as regulatory genes, even though these sequences do not code for any RNA or protein.



Three Different Classes of RNA

1) rRNA (ribosomal)

- large (long) RNA molecules
- structural and functional components of ribosomes
- highly abundant

2) mRNA (messenger)

- typically small (short)
- encode proteins
- multiple types, not abundant

3) tRNA (transfer) and small ribosomal RNAs

- very small
- Important in translation

Not all genes encode proteins



Different Types of RNA Polymerase

In Bacteria (simple system)

- all three classes are transcribed by the same RNA polymerase (RNAP for short)

In Eukaryotes (complex system)

- each class is transcribed by a different RNA Polymerase:

- **RNAP I - rRNAs**
- **RNAP II - mRNAs**
- **RNAP III - tRNAs & small ribosomal RNAs**



Types of RNA and the process of Transcription

- In bacteria, there are three major types of RNAs: mRNA (messenger RNA), tRNA (transfer RNA), and rRNA (ribosomal RNA).
- All three RNAs are needed to synthesise a protein in a cell.
- The mRNA provides the template, tRNA brings amino acids and reads the genetic code, and rRNAs play structural and catalytic role during translation.
- There is single DNA dependent RNA polymerase that catalyses transcription of all types of RNA in bacteria.
- RNA polymerase binds to promoter and initiates transcription (Initiation).
- It uses nucleoside triphosphates as substrate and polymerises in a template depended fashion following the rule of complementarity.
- It somehow also facilitates opening of the helix and continues elongation.

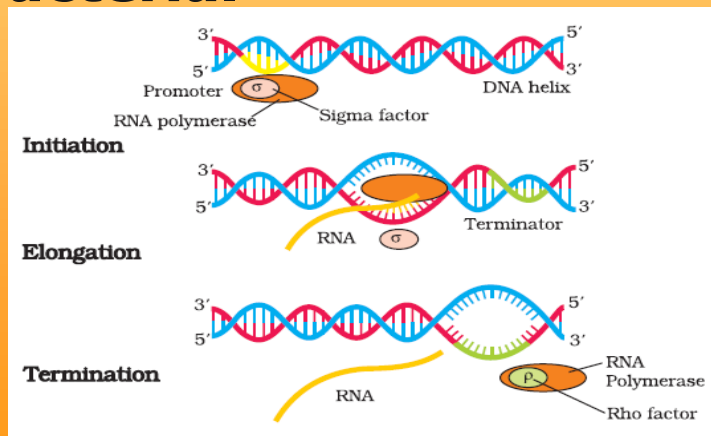


Types of RNA and the process of Transcription

- Only a short stretch of RNA remains bound to the enzyme.
- Once the polymerases reaches the terminator region, the nascent RNA falls off, so also the RNA polymerase.
- This results in termination of transcription.
- An intriguing question is that how is the RNA polymerases able to catalyse all the three steps, which are initiation, elongation and termination.
- The RNA polymerase is only capable of catalysing the process of elongation.
- It associates transiently with initiation-factor (σ) and termination-factor (ρ) to initiate and terminate the transcription, respectively.
- Association with these factors alter the specificity of the RNA polymerase to either initiate or terminate.

Types of RNA and the process of Transcription

- In bacteria, since the mRNA does not require any processing to become active, and also since transcription and translation take place in the same compartment (there is no separation of cytosol and nucleus in bacteria), many times the translation can begin much before the mRNA is fully transcribed.
- Consequently, the transcription and translation can be coupled in bacteria.





Types of RNA and the process of Transcription

- **In eukaryotes, there are two additional complexities –**
 - **There are at least three RNA polymerases in the nucleus (in addition to the RNA polymerase found in the organelles). There is a clear cut division of labour. The RNA polymerase I transcribes rRNAs(28S, 18S, and 5.8S), whereas the RNA polymerase III is responsible for transcription of tRNA, 5srRNA, and snRNAs (small nuclear RNAs). The RNA polymerase II transcribes precursor of mRNA, the heterogeneous nuclear RNA (hnRNA).**



Types of RNA and the process of Transcription

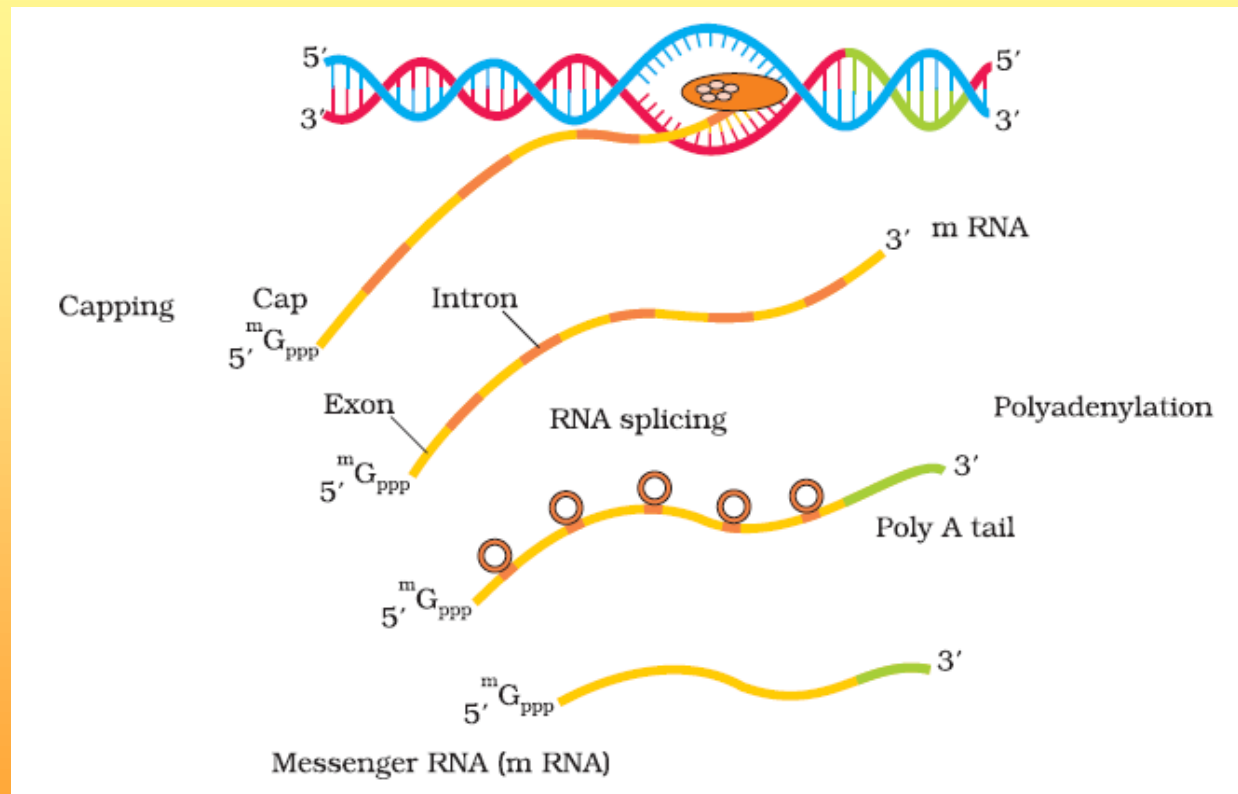
- **The second complexity is that the primary transcripts contain both the exons and the introns and are non-functional. Hence, it is subjected to a process called splicing where the introns are removed and exons are joined in a defined order. hnRNA undergo two additional processing called as capping and tailing. In capping an unusual nucleotide (methyl guanosine triphosphate) is added to the 5'-end of hnRNA. In tailing, adenylate residues (200-300) are added at 3'-end in a template independent manner. It is the fully processed hnRNA, now called mRNA, that is transported out of the nucleus for translation.**



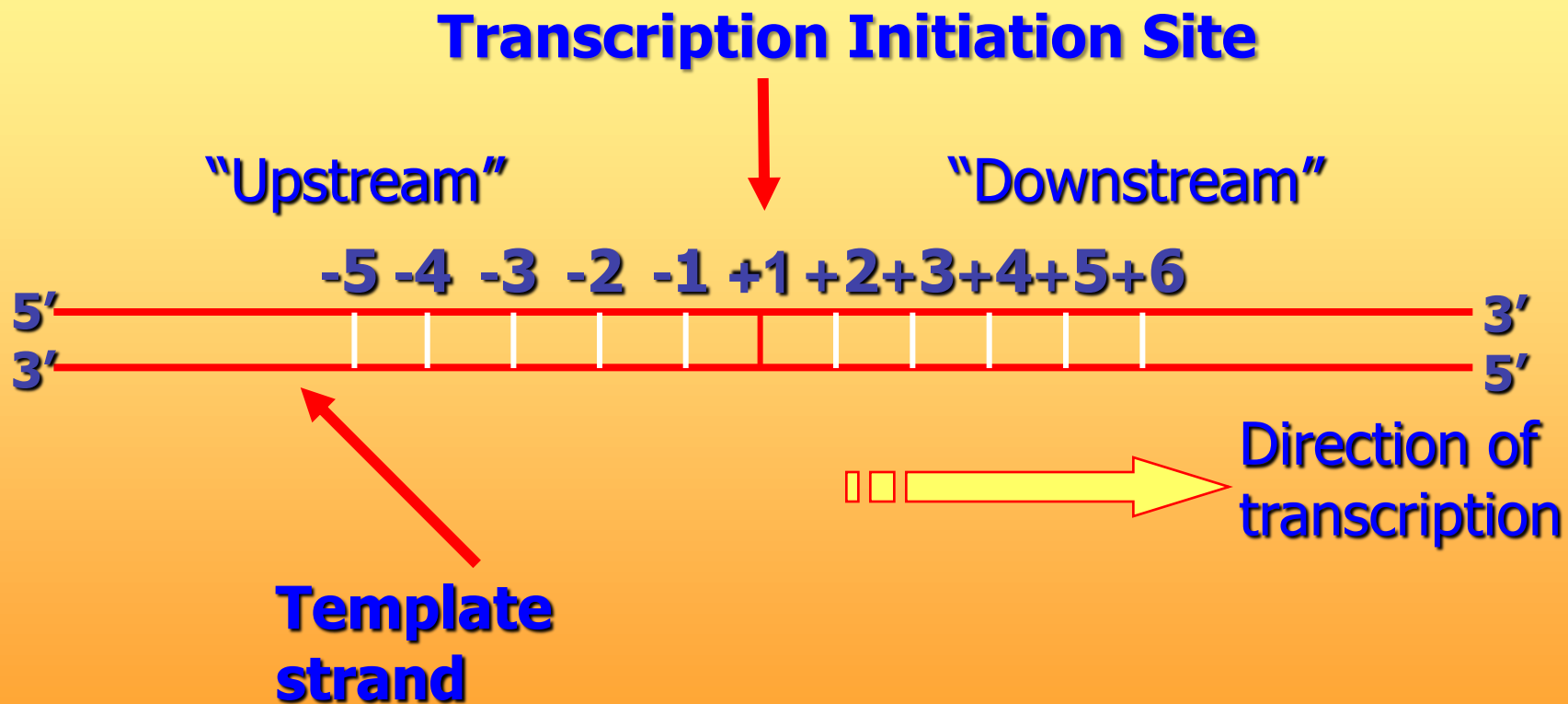
Types of RNA and the process of Transcription

- **The significance of such complexities is now beginning to be understood. The split-gene arrangements represent probably an ancient feature of the genome. The presence of introns is reminiscent of antiquity, and the process of splicing represents the dominance of RNA-world. In recent times, the understanding of RNA and RNA-dependent processes in the living system have assumed more importance.**

Types of RNA and the process of Transcription



Some nomenclature conventions



There is no "zero"



Transcription Factors (TAFs)

- **Activators: specific DNA-binding proteins that activate transcription.**
- **Repressors: specific DNA-binding proteins that repress transcription.**
- **Some regulatory proteins can work as both activators and repressors for different genes.**
- **TAF sites are more difficult to locate than genes.**
- **Nucleosome positioning influences gene transcription.**

RNAP binds a region of DNA from -40 to +20

The sequence of the non-template strand is shown

Gene	-35 region	Pribnow box (-10 region)	Initiation site (+1)
<i>araBAD</i>	GGATGCTACCTGACGCTTTT	TATCGGCAACTGTC	TACTGTTTGTGCAATAGCCGGTTTTT
<i>araC</i>	GGCGTGATTATAGACACTTTT	TGTTACGGCTTTT	TGTCATCGCCTTTGCTGCCGCTTTE
<i>blaA</i>	TTCCAAAAGCTGTTTTTTGTTG	TTAATTCCGGTG	TAGACTTGTAAAGCTAAATCTTTT
<i>blaB</i>	CATAATCGACTTGTAACCAAA	ATTGAAAAGATT	TAGGTTTACAAGTCTAGACCGGAAT
<i>galP2</i>	ATTTATTCCATGTCACACTTTT	TCCGATCTTTCT	TATGCTATGCTTATTTGATACCAT
<i>lac</i>	ACGCCAGGCTTTACACTTTT	ATGGTTCCGGCTCG	TATGTTGTGTGGAATTGTGAGCGGG
<i>lacI</i>	CCATCGAATGGCGCAAAACCT	TTCCGGGGTATCG	CATGATAGCCCCCGGAAGACAGTC
<i>rnaI</i>	AAAAATAAATGCTTGACTCTGT	AGCCGGGAAGCCG	TATTATCACACCCCGCCGCCGCTG
<i>rnaII</i>	CAAAAAAATACTTGTCAAAAA	ATTGGGATCCG	TATAATCCGGCTCCGTTGACAGCA
<i>rnaIII</i>	CAATTTTTCTATTGCGGCCTG	CGGCAGAACTCCG	TATAATCCGGCTCCATCGACAGCG
<i>rRNA^{16S}</i>	CAACGTAACTTTACAGCGG	CGCGTCATTTGAT	TATGATCCGCCCCCGCTTCCCGATA
<i>trp</i>	AAATGAGCTGTTGACAATTA	ATCATCGAAGTAG	TTAACTAGTACGCCAAGTTACGTA

	-35 region	Pribnow box	Initiation site
Consensus sequence:	T C T T G A C A T	T A T A A T	A
	42 38 82 84 79 64 53 45 41	79 95 44 59 51 96	51
		-10 region	C 55 T 48 G 42

TTGACA...16-19 bp... TATAAT
"-35" spacer "-10"



Important Promoter Features (tested by mutations)

- **The closer the match to the consensus the stronger the promoter (-10 and -35 boxes).**
- **The absolute sequence of the spacer region (between the -10 and -35 boxes) is not important.**
- **The length of the spacer sequence IS important- TTGACA - spacer (16 to 19 base pairs) – TATAAT**
- **Spacers that are longer or shorter than the consensus length make weak promoters.**



Properties of Promoters

- Promoters typically consist of a 40 bp region on the 5'-side of the transcription start site.
- Two consensus sequence elements:
 - The "-35 region", with consensus TTGACA
 - The Pribnow box near -10, with consensus TATAAT - this region is ideal for unwinding.



RNA Polymerase has many functions

- **Scan DNA and identify promoters**
- **Bind to promoters**
- **Initiate transcription**
- **Elongate the RNA chain**
- **Terminate transcription**
- **Be responsive to regulatory proteins (activators and repressors)**

Thus, RNAP is a multi-subunit enzyme



Points to Remember

- RNA is the genetic material in some viruses.
- The synthesis of RNA molecule from and complementary to DNA template is called transcription.
- It occurs during G1 and G2 phases of cell cycle in the nucleus.
- The different types of RNAs, namely, tRNA, mRNA and rRNA are transcribed from different regions of the DNA molecule catalysed by RNA polymerase I, II and III respectively.
- In RNA transcription, a specific region of DNA unwinds and splits into single to expose the bases.



Points to Remember

- Ribonucleotide units are joined by the enzyme RNA polymerase through hydrogen bonding to the appropriate bases of the sense DNA strand, forming a complementary RNA strand.
- The RNA strand separates from the DNA template; and the split region of the DNA molecule gets hydrogen bonded and spirally coiled to resume its original form.
- Formation of mRNA from DNA is called transcription.
- During transcription the DNA site at which RNA polymerase binds is called enhancer.
- The enzyme involved in transcription is RNA polymerase.



Points to Remember

- During transcription if nucleotide sequence of DNA strand that is being coded is ATACG then the nucleotide sequence in mRNA would be UAUGC.
- During transcription , RNA polymerase holoenzyme binds to a gene promoter and assumes a saddle like structure. Its DNA binding sequence will be TATA.
- Molecular basis of organ differentiation depends on the modulation in transcription by RNA polymerase.
- The coding segment of DNA is exon.
- Removal of Introns and joining the exons in a defined order in a transcription is called splicing.



Thanks...