

Transcription

Dr. Kalsoom Tariq

The Genetic code

- The genetic code is a dictionary that identifies the correspondence between a sequence of nucleotide bases and a sequence of amino acids
- Each individual word in the code is composed of three nucleotides bases
- These genetic words are called **codons**.

- Codons are presented in the mRNA language of adenine (A), Guanine (G), Cytosine (C) and Uracil (U)
- The four nucleotide bases produce 64 different combinations of three base codons
- The nucleotide sequence of codon on mRNA is written from 5' to 3' end
- 61 codons code for 20 amino acids found in protein.

Genetic Code- Table

		Second Letter				
		U	C	A	G	
1st letter	U	UUU Phe UUC UUA Leu UUG	UCU UCC Ser UCA UCG	UAU Tyr UAC UAA Stop UAG Stop	UGU Cys UGC UGA Stop UGG Trp	U C A G
	C	CUU CUC Leu CUA CUG	CCU CCC Pro CCA CCG	CAU His CAC CAA Gln CAG	CGU CGC Arg CGA CGG	U C A G
	A	AUU AUC Ile AUA AUG Met	ACU ACC Thr ACA ACG	AAU Asn AAC AAA Lys AAG	AGU Ser AGC AGA Arg AGG	U C A G
	G	GUU GUC Val GUA GUG	GCU GCC Ala GCA GCG	GAU Asp GAC GAA Glu GAG	GGU GGC Gly GGA GGG	U C A G

Non Sense Codon

- Three of the codons, UAG, UGA and UAA do not code for amino acids and are called **termination codon**
- When one of these codons appear in mRNA sequence, it signals synthesis of the protein coded for by that mRNA is completed
- AUG codes for methionine which is the first amino acid in the poly peptide ,known as the **start codon**.

Characteristics of Genetic Code

- Specificity (unambiguity)
- Universality
- Degeneracy
- Non overlapping
- Comma less ness

Specificity

- A particular codon always codes for the same amino acid
- For example: UGG is a code specified for Tryptophan

Universality

- The same codons are used to code for the same amino acids in all living organisms.
- Thus genetic code has been conserved during the course of evolution.
- Exception to universality is found in mitochondrial genome where 'AUA' codes for methionine instead of isoleucine & UGA for tryptophan instead of stop codon.

Degeneracy

- Degeneracy means having multiple domain elements corresponding to one element.
- Most of the amino acids have more than one codon
- The codon is degenerate since there is 61 codons for 20 amino acids
- The codons that designate the same amino acids are called **synonyms**.

Degeneracy

- Often the base of the third position is insignificant meaning the codons representing the same amino acid differ in the third base ,e.g; UCU,UCC, UCA, UCG all code for serine.
- The reduced specificity of the last position is known as “**third base degeneracy or wobbling phenomenon**”.
- This feature minimizes the effects of mutations.

Non-overlapping

- All codons are independent set of 3 bases, read from a fixed point as a continuous base sequence.
- There is non-overlapping, i.e. no base functions as a common member of two consecutive codons.
- For example: UUUCUUAGA is read as UUU/CUU/AGA/
- Addition or deletion of one or two bases will change the message sequence in mRNA

Comma less ness

- Codons are arranged in a continuous structure.
- There is not one or more nucleotides between consecutive codons.
- The last nucleotide of preceding codon is immediately followed by the first nucleotide of the succeeding nucleotide.

Consequences of Altering the Nucleotide sequence

Changing a single nucleotide base on the mRNA chain results in any one of the following mutation.

A. Point mutations

B. Frame shift mutations

Point mutations can either be

1. Transitions

2. Transversions ,
which can be further subdivided into

- Silent mutation
- Mis sense mutation
- Nonsense mutation

Silent Mutation

- means (no detectable effect).
- The codon containing the changed base may code for the same amino acid.
- For example: if serine codon UCA is given a different third base-U-to become UCU, it still codes for serine. Therefore it is termed as 'silent mutation'

Missense Mutation

- The codon containing the changed base may code for different amino acid
- For example: if serine codon UCA is given a different first base-C-to become CCA, it will code for different amino acid that is proline.
- The substitution of incorrect amino acid is called mis sense mutation

Nonsense Mutation

- The codon containing the changed base may become a termination codon.
- For example: if serine codon UCA is given a different second base-A-to become UAA, the new codon causes the termination of translation at that point.
- The creation of termination codon at an inappropriate place is called nonsense mutation.

Other Mutation

- These can alter the **amount** or **structure** of protein produced by translation
- Tri nucleotide repeat expansion
- Splice site mutation
- Frame-shift mutation

Trinucleotide repeat expansion

- Occasionally the sequence of three bases become amplified in number so that many copies of the triplet occur.
- If this occur in the coding region of a gene, the protein will contain many extra copies of one amino acid
- Eg; Amplification of CAG codon leads to the insertion of many extra glutamine residues in huntington protein causing the neuro-generative disorder Huntington disease

Splice Site Mutation

- Mutations at splice sites can change the way in which introns are removed from pre m RNA molecules, which can produce abnormal protein.

Frame-Shift Mutation

- Two types, **Deletion type** or **Insertion type**
- If one or two nucleotide is either deleted from or added to the interior of a message sequence, a frame-shift mutation occurs and the reading frame is altered
- The resulting amino acid sequence may become different from this point
- If three nucleotides are added or deleted from the strand, the reading frame is not affected

Mechanisms of Repair

Several methods are available. These are:

- **A. Excision repair**
- **B. Photo reactivation**
- **C. Recombinational repair**
- **D. Mismatch repair.**

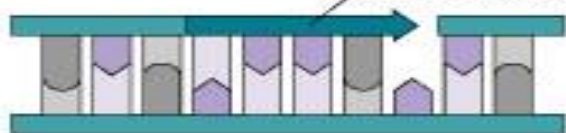
Ultraviolet light



Thymine dimer



New DNA



1 Exposure to ultraviolet light causes adjacent thymines to become cross-linked, forming a thymine dimer and disrupting their normal base pairing.

2 An enzyme cuts out and removes the damaged DNA.

3 DNA polymerase fills the gap by synthesizing new DNA, using the intact strand as a template.

4 DNA ligase seals the remaining gap by joining the old and new DNA.

- ***Photo reactivation: Also called light induced repair.***
- The enzyme ***photo reactivating (PR) enzyme brings about***
- an enzymatic cleavage of thymine dimers activated by
- the visible light (**300-600 mμ**) leading to a restoration

- ***Recombinational Repair***
- Also called as ***sister-strand exchange***.
- In this process, the un mutated single stranded segment from homologous DNA is excised from the “good” strand and inserted into the “gap” opposite the dimer.
- It occurs after the first round of DNA replication.
- ***Mismatch Repair of DNA***
- Mismatch repair corrects errors made when DNA is copied.
- ***Mechanism:*** This mechanism corrects a single mismatch base pair, e.g. C to A rather than T to A or a short region of unpaired DNA.
The defective region is recognized by an ***endonuclease***

Base Excision Repair (BER)

Variety of **DNA glycosylases**, for different types of damaged bases.

AP endonuclease recognizes sites with a missing base; cleaves sugar-phosphate backbone.

Deoxyribose phosphodiesterase removes the sugar-phosphate lacking the base.

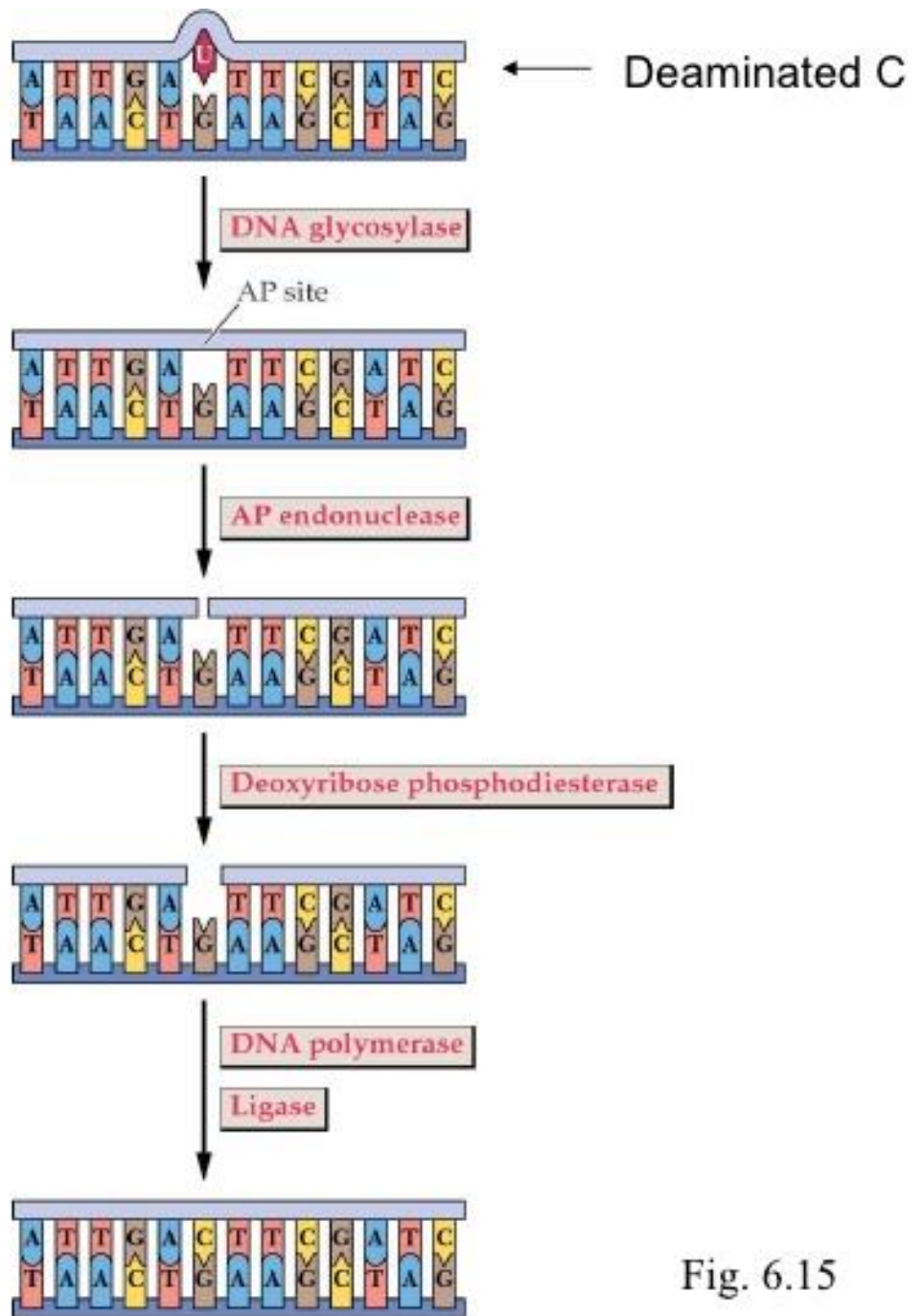


Fig. 6.15

Transcription

- Transcription is the process in which ribonucleic acid (RNA) is synthesized from DNA
- For this purpose, one of the two strands of DNA serves as a template to produce working copies of RNA molecules

Transcription is Selective

- The entire molecule of DNA is not expressed in transcription
- RNA are synthesized only for some selected regions of DNA
- The product formed in transcription is referred to as **primary transcript**
- The primary transcript is **inactive**. They undergo post transcriptional modification to produce functionally active RNA molecules

Transcription

- A single enzyme DNA dependent RNA polymerase or simply RNA polymerase is responsible for transcription
- RNA polymerase is a complex holoenzyme with five polypeptide units ___ 2α , 1β and $1\beta'$ and 1σ factor
- The enzyme with out a sigma factor is called a **core enzyme**
- And with a core enzyme is called **Holo enzyme**

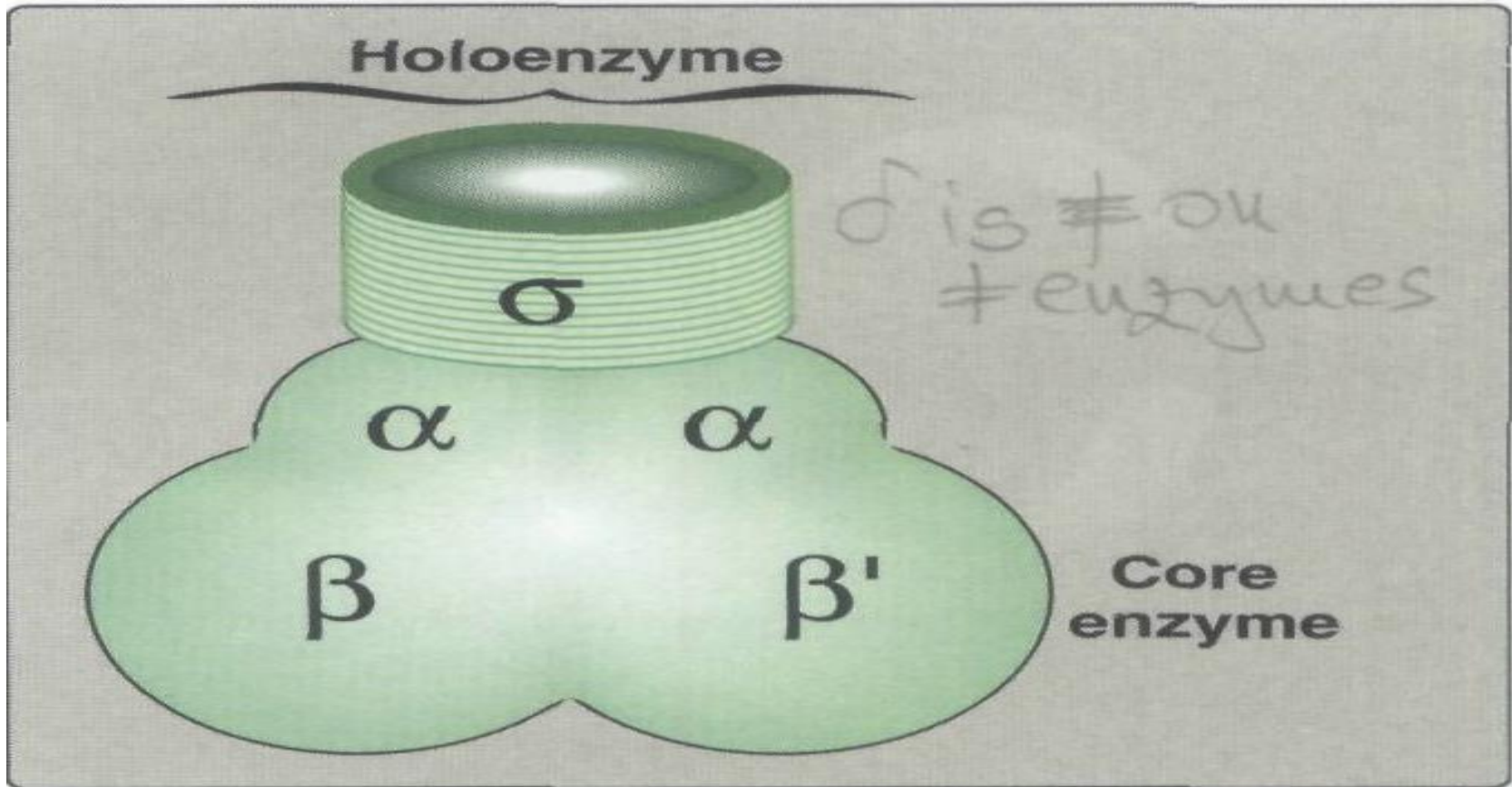


Figure 30.6
Prokaryotic *RNA polymerase*.

Steps of Transcription

Transcription involves three different stages:

- Formation of transcription complex
- Initiation
- Elongation
- Termination

Initiation

- The binding of the enzyme RNA polymerase is prerequisite for the transcription to start
- The specific region on the DNA where the enzyme binds is known as **promoter region**
- There are **two base sequences** on the coding DNA strand which the sigma factor of RNA polymerase can recognize for initiation of transcription

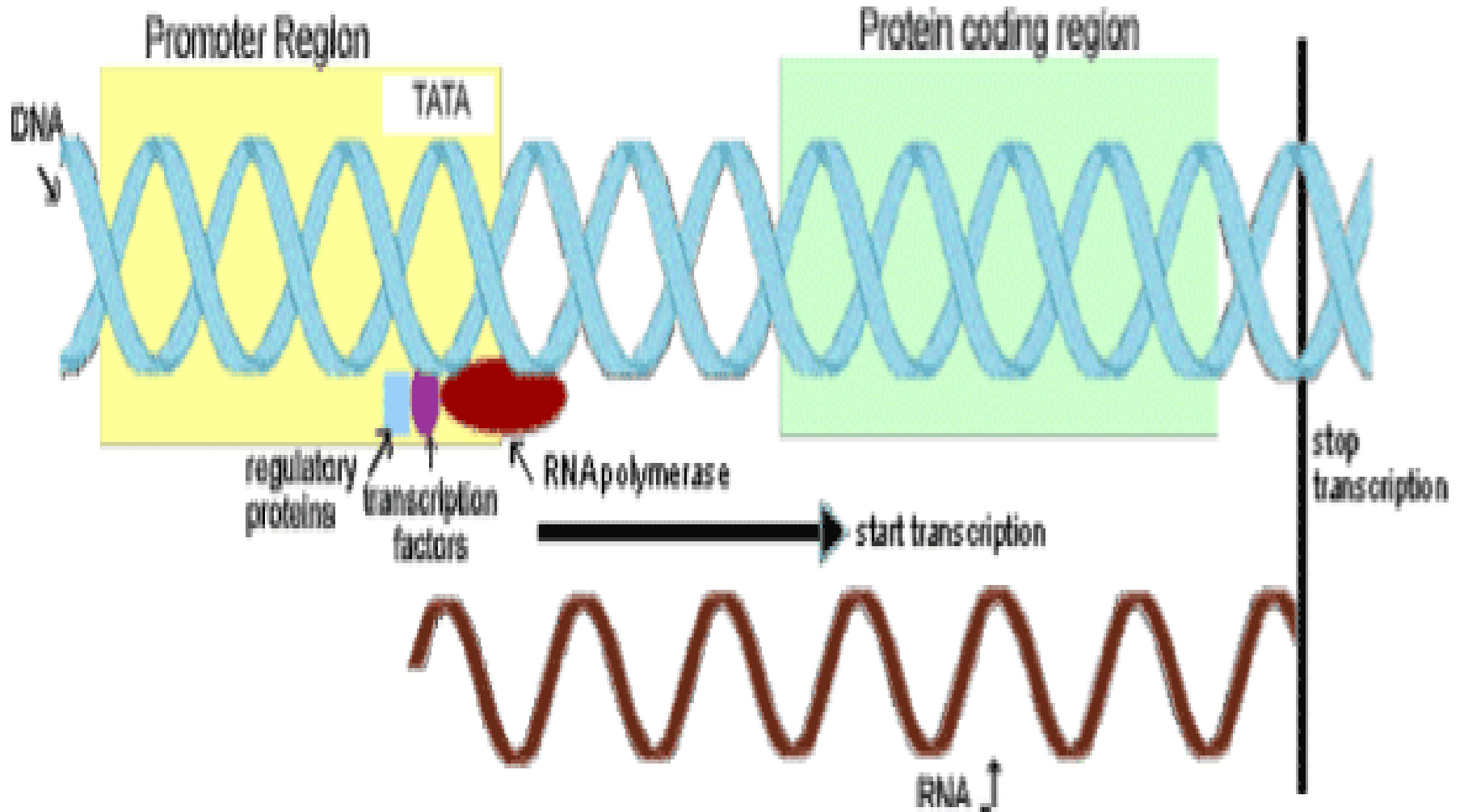
Pribnow Box

- This consist of 6 nucleotide bases (TATAAT) located on the left side about 10 bases away (upstream) from the starting point of transcription
- The T at position 6 is present in every promotor region of pribnow box & is known as **(conserved T)**.

The '---35' sequence

- This is the second recognition site in the promoter region of DNA
- It contains a base sequence **TTGACA**, which is located about 35 bases away on the left side from the site of transcription started.
- It is thought to be the initial site of sigma subunit binding, it contains up to nine bases.

Typical gene organization



Elongation

- As the holo enzyme, RNA polymerase recognizes the promoter region, the sigma factor is released and transcription proceeds
- RNA is synthesized from 5' to 3' direction, anti parallel to DNA template
- RNA polymerase utilizes ribonucleotide triphosphates (ATP, GTP, CTP and UTP) for the formation of RNA

Elongation

- The sequence of nucleotide bases in the mRNA is complementary to the template DNA strand. It is however, identical to the coding strand except that RNA contains U in place of T in DNA

Elongation

- RNA polymerase differs from DNA polymerase in **two** aspects.
- 1) No primer is required
- 2) RNA polymerase do not have endo and exonuclease activity and also do not have proof reading activity
- RNA polymerase has no ability to repair the mistakes in the RNA synthesized in contrast to DNA replication which is carried out with high fidelity

Elongation

- It is however fortunate that mistakes in RNA synthesis are less dangerous, since they are not transmitted to the daughter cells
- The double helical structure of DNA unwinds as the transcription goes on, resulting in super coils
- The problem of super coils is overcome by **topo isomerase**.
- The DNA helix recloses after RNA polymerase transcribes through it & the growing RNA chain dissociates from it.

Termination

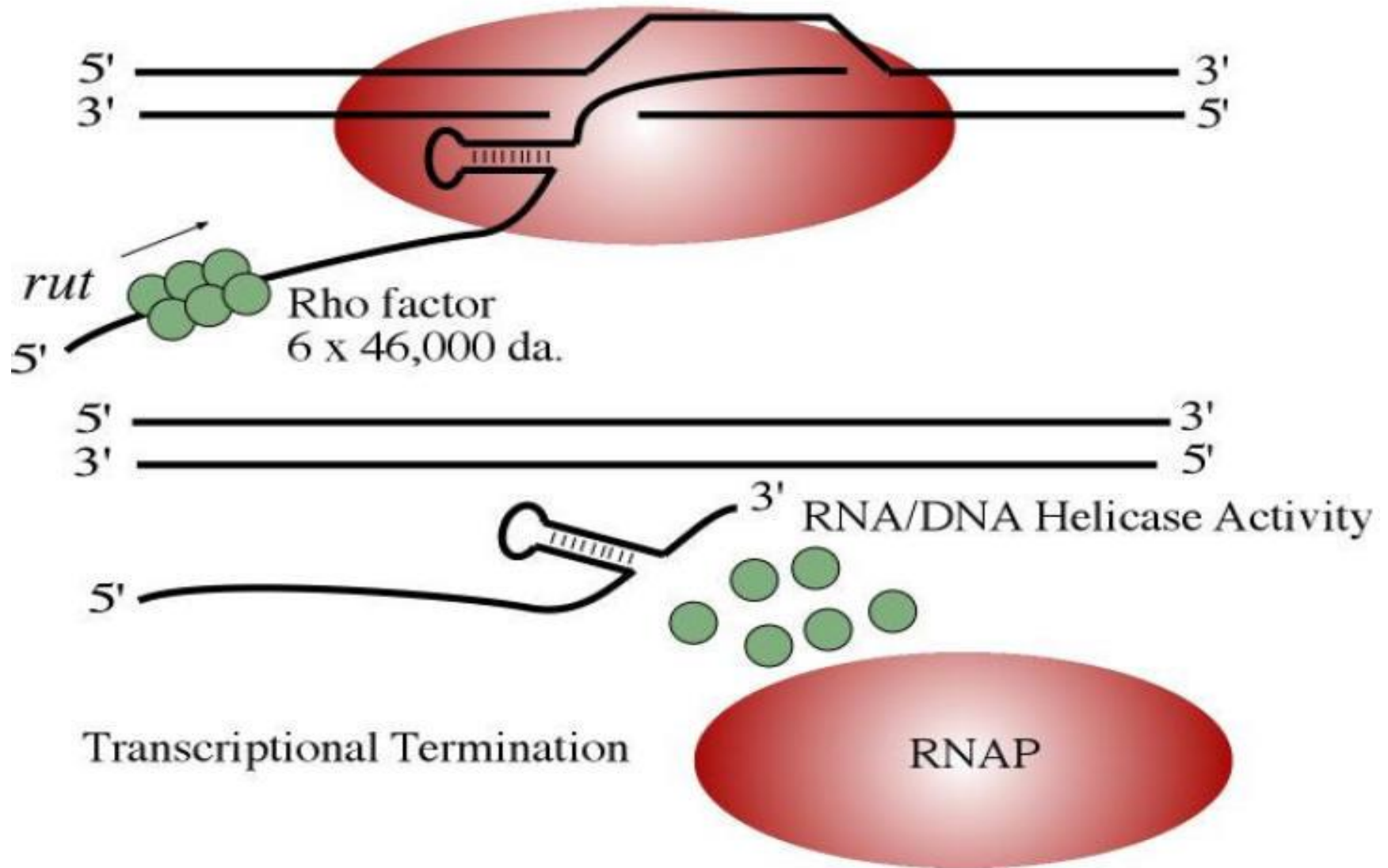
Two types of termination are identified

- Rho dependent termination
- Rho independent termination

Rho Dependent Termination

- A specific protein, named rho factor, binds to the growing RNA , and **terminates transcription and release RNA**
- The Rho factor is also responsible for the **dissociation of RNA polymerase from DNA**

Rho-Dependent Termination



Rho Independent Termination

- This requires that a sequence in template DNA generates a sequence in RNA which is self complementary
- This allows the RNA to fold back on itself, forming GC rich stem plus loop. This structure is known as **hairpin**
- Beyond the hairpin the RNA transcript contain a string of Uracils
- The bonding of Us with Adenines of DNA is weak and thus RNA strand splits out

T G C G	GGT C G A C T G	C C G A T	C A G T C G A C C	T T T T
A C G C	C C A G C T G A C	G G C T A	G T C A G C T G G	A A A A

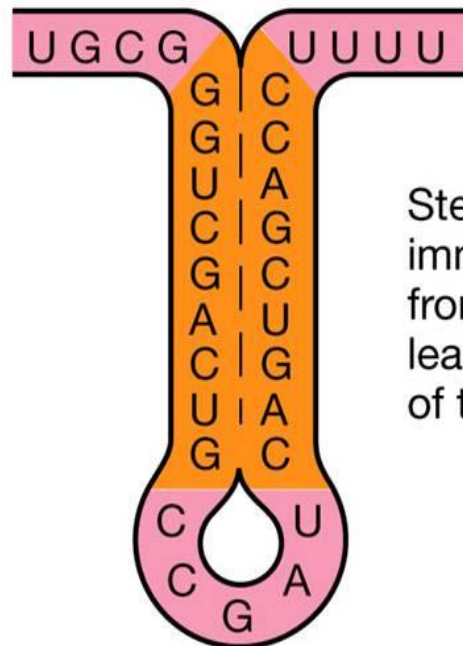
DNA with inverted repeats

↓ Transcription of one strand

U G C G	GGU C G A C U G	C C G A U	C A G U C G A C C	U U U U
---------	-----------------	-----------	-------------------	---------

RNA

↓ Fold to form secondary structure



Stem-loop in RNA immediately upstream from a run of uracils leads to termination of transcription

Transcription in Eukaryotes

- RNA synthesis in eukaryotes is a much more complicated process than the transcription described in prokaryotes
- The difference in the eukaryotic transcription occurs in
 - RNA polymerases
 - Promoter Sites
 - Initiation step

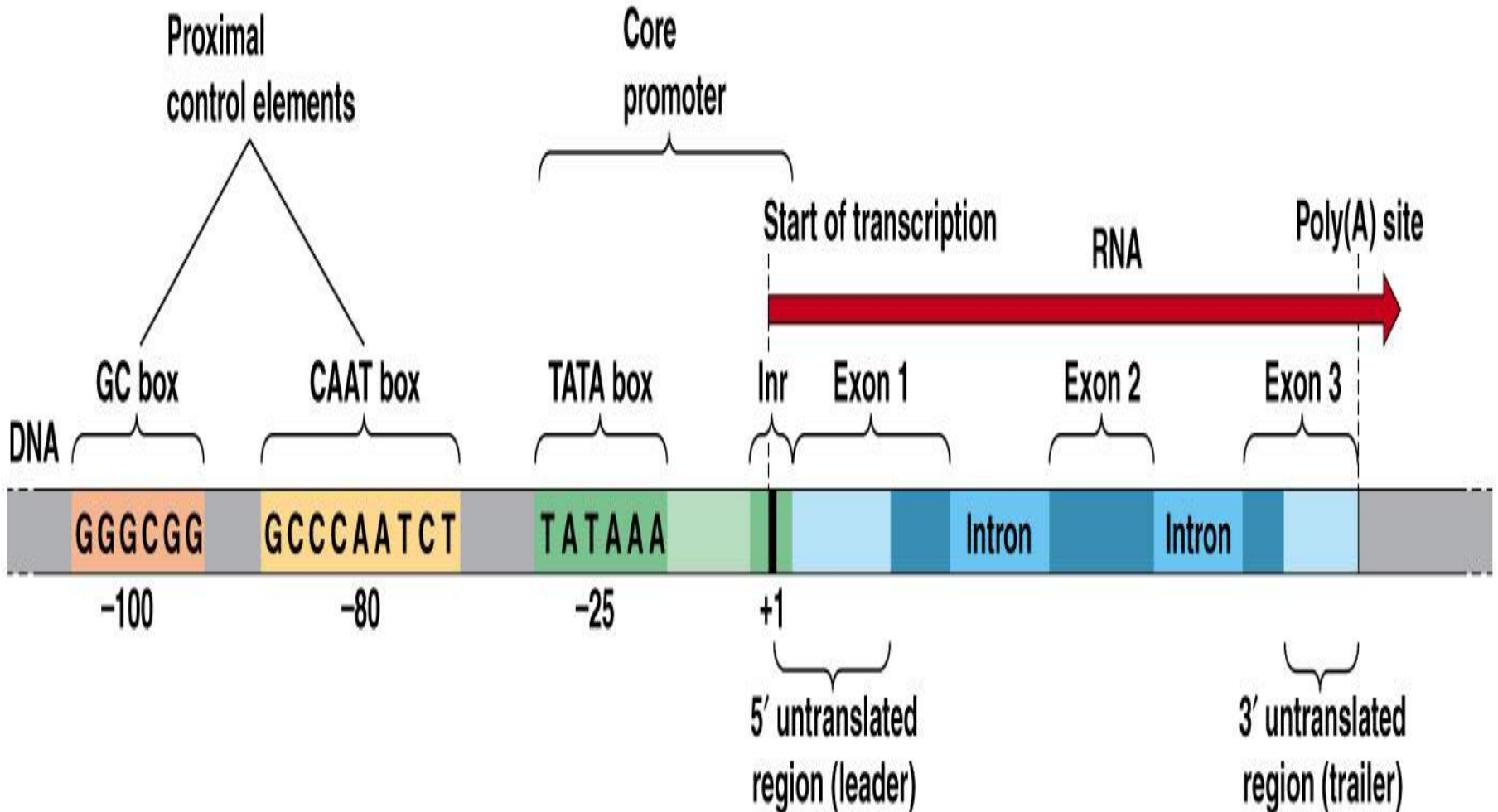
Eukaryotic RNA Polymerases

- The nuclei of eukaryotic cells possess three distinct RNA polymerases
- **RNA polymerase I:** responsible for the synthesis of ribosomal RNA
- **RNA polymerase II:** responsible for the synthesis of messenger RNA
- **RNA polymerase III:** responsible for the synthesis of transfer RNA

Promoter Sites

- In eukaryotes three different promoter regions are identified
- **Hogness box:** also called **TATA box**, located on the left about **25** nucleotide away from the starting site of mRNA synthesis
- **CAAT box:** it is located to the left, **80** bases away from the starting site of mRNA synthesis
- **GC region:** **100** bases towards the left upstream from the starting site of mRNA synthesis.

Promoter Sites

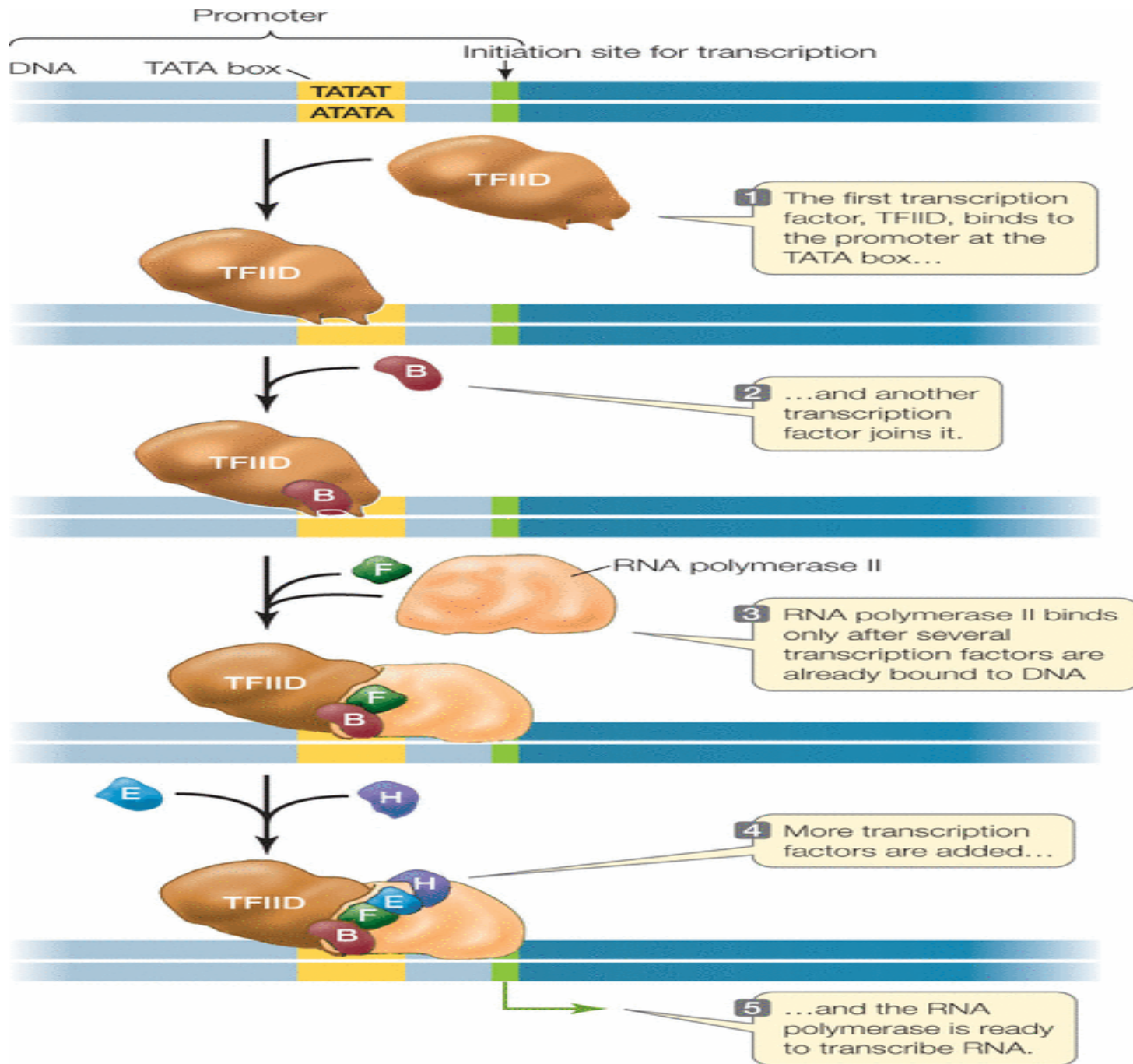


Initiation of Transcription

- The molecular events required for the initiation of transcription in eukaryotes are complex and broadly involve three stages
- **Attachment of sigma factor** of RNA polymerase II to the promoter region
- **Binding of transcription factor (TFs)** to the promoter region
- **Attachment of enhancer** to the promoter region and thus increasing the gene expression by 100 folds

Transcription factors

- TF are a very diverse family of proteins & generally function in multi subunit protein complexes. They may bind directly to the special “promotor “ region or directly to the RNA Polymerase molecule.
- Their function is to **initiate & regulate** the transcription of genes.



Heterogeneous Nuclear RNA (hnRNA)

- The primary mRNA transcript produced by RNA polymerase II in eukaryotes is often called **heterogeneous nuclear RNA (hnRNA)**
- It is the precursor of mRNA & is in inactive form.

Post Transcriptional Modification

- The RNAs produced during transcription are called **primary transcript**. They undergo many alterations, terminal base additions, base modifications and splicing.
- These alterations are collectively called **post-transcriptional modifications**
- This process is required to convert the RNAs into active form
- **Ribo nucleases** are responsible for post transcriptional modifications

Post Transcriptional Modification

- The primary transcript of mRNA is hn RNA in eukaryotes, which is subjected to many changes before functional mRNA is produced
- 5' capping
- Poly A tail
- Removal of Introns

The 5' Capping

- The 5' end of mRNA is capped with 7-methyl guanosine produced by S-Adenosyl-methionine + GTP.
- This group is required for translation, besides stabilizing the structure of mRNA

Poly A tail

- A large number of eukaryotic mRNAs possess an adenine nucleotide chain at the 3' end.
- This poly A tail, as such, is not produced during transcription.
- It is later added to stabilize mRNA

RNA capping and polyadenylation

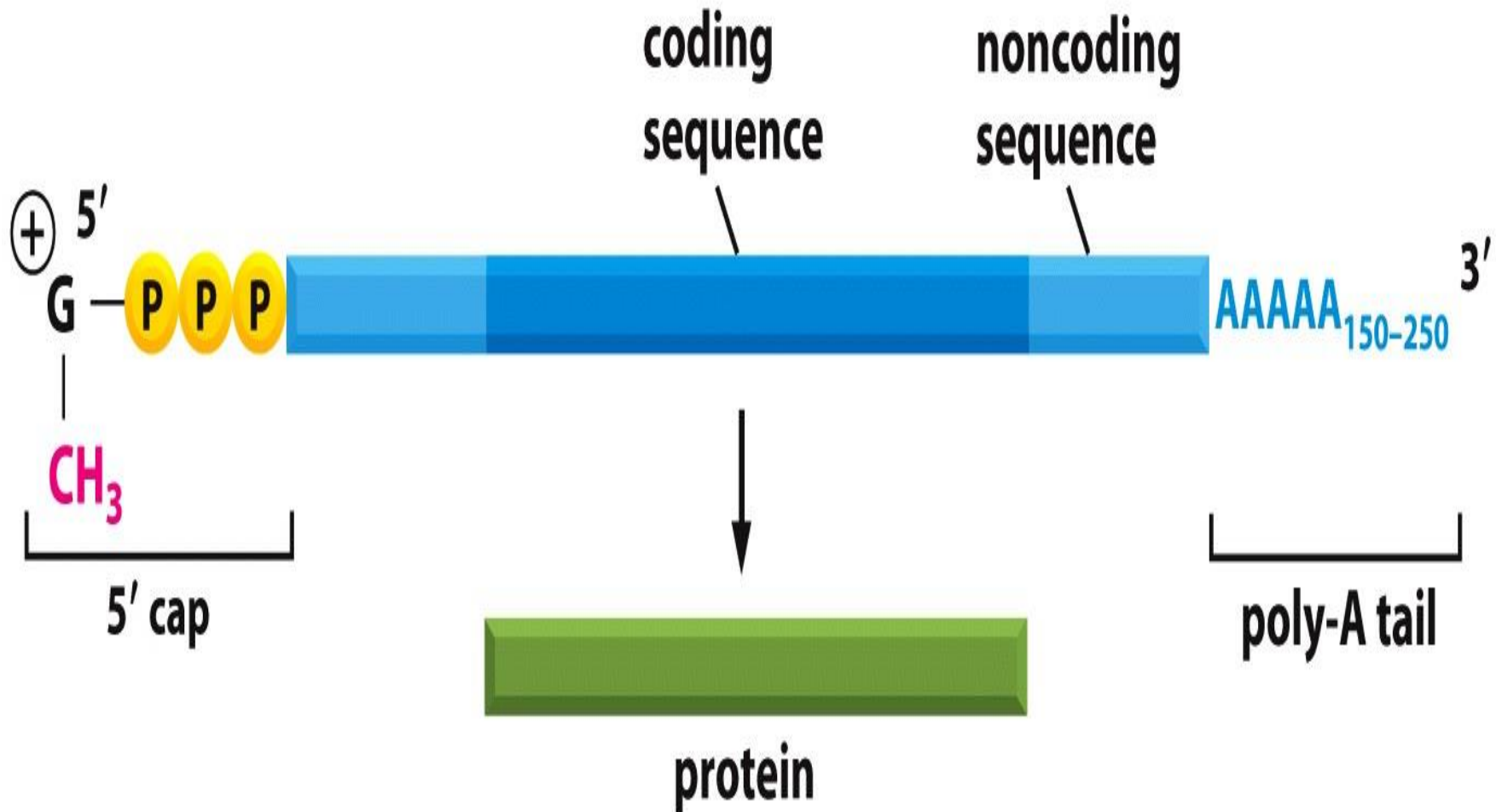
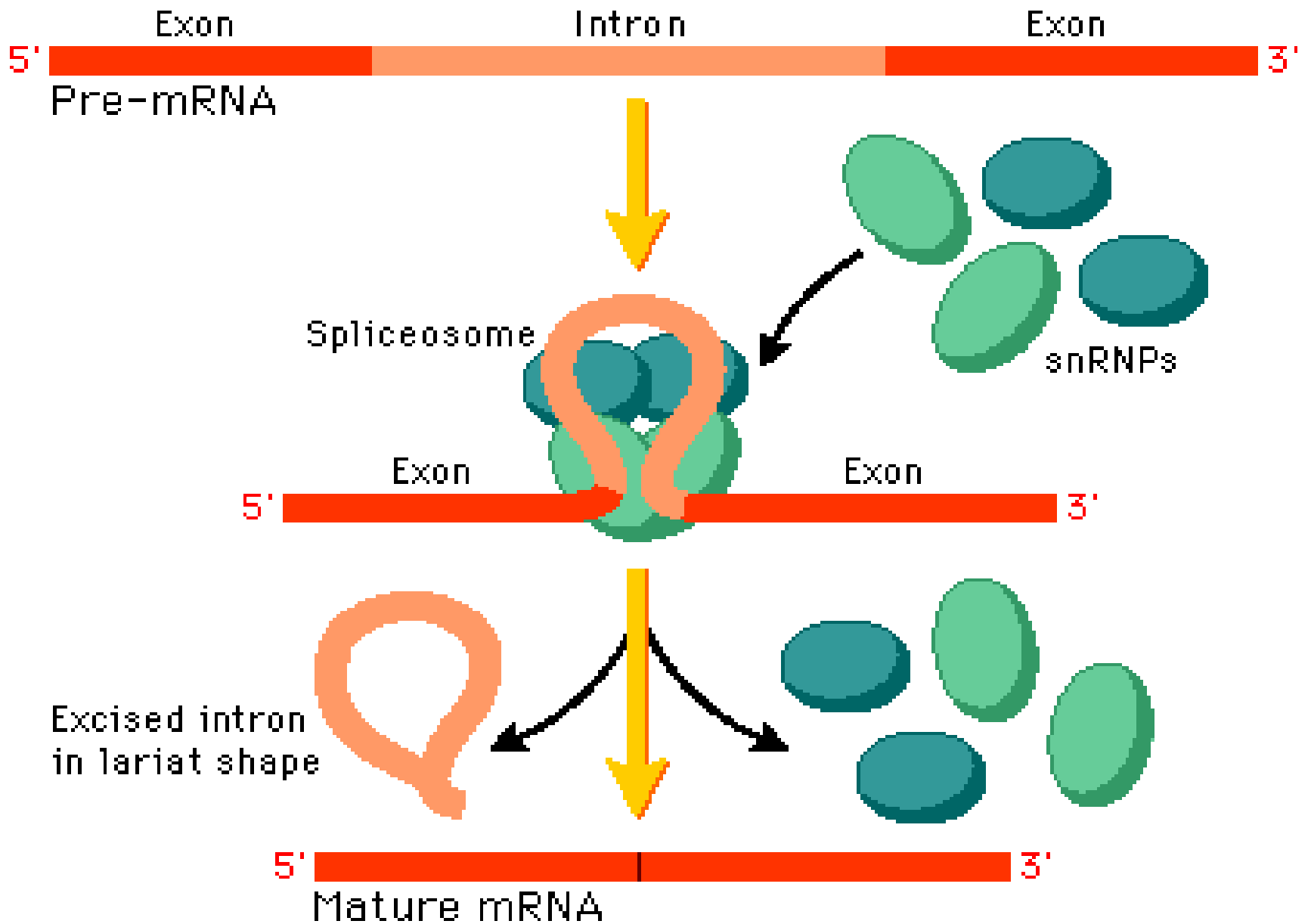


Figure 7-16a Essential Cell Biology 3/e (© Garland Science 2010)

Removal of Introns

- Introns are the intervening nucleotide sequence in mRNA which do not code for protein
- Exons of mRNA possess genetic code and are responsible for protein synthesis
- The removal of introns are promoted by **small nuclear ribo nuclear protein particles (SnRNPs)**



Inhibitors of Transcription

- **Actinomycin D:** It binds with DNA template strand and blocks the movement of RNA polymerase. It is widely used in the treatment of tumors
- **Rifamycin:** It binds to the beta subunit of RNA polymerase and inhibits its activity
- **Amantin:** It tightly binds to the RNA polymerase II in eukaryotes and stops the transcription process
- **Heparin:** It binds to the beta subunit & inhibits transcription in vitro.