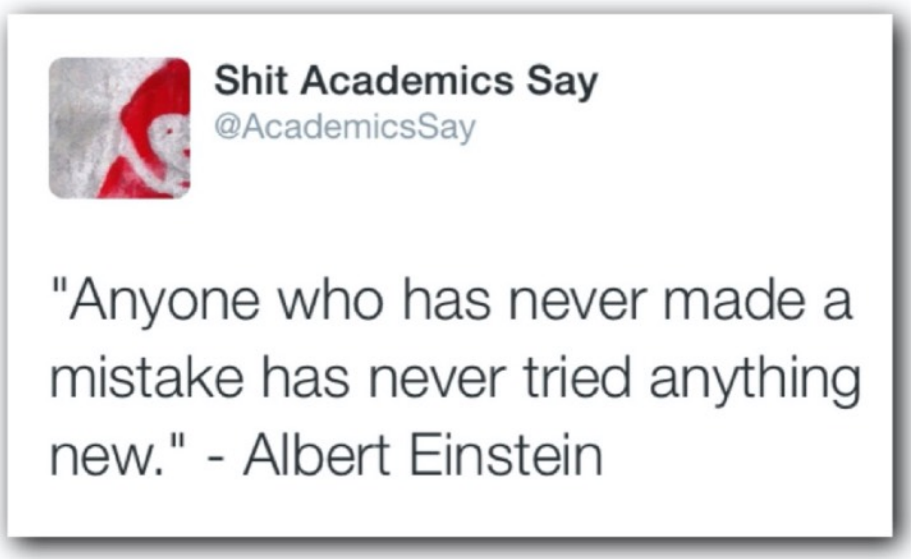
Lecture 13:

Transcription in eukaryotes

Course 281

Lessons for life



AIMS

- Understand the transcription process in eukaryotes.
- Understand the eukaryotic gene structure.
- Understand the sequence of events and the molecular machinery.
- Understand the molecular modifications of eukaryotic transcript and why they take place.
- Compare the transcription process of prokaryotes to that of eukaryotes.

Eukaryotes RNA polymerase

Unlike prokaryotes, eukaryotes have different kinds of RNA polymerases. Each RNA polymerase is responsible for transcription of specific RNA molecules.

Eukaryotes RNA polymerase

1. **RNA Pol I:** Transcribes RNA found in ribosomes
   - 28s ribosomal RNA
   - 18s rRNA
   - 5.8s rRNA

2. **RNA Pol II:** Transcribes
   - m-RNA
   - Some sn-RNA

3. **RNA Pol III:** Transcribes
   - t-RNA
   - 5s rRNA
   - Some sn-RNA

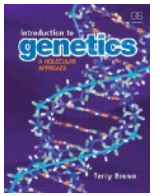
- Eukaryotes possess more complex RNA polymerases.

- It requires three different RNA polymerases namely RNA polymerase I, RNA polymerase II, RNA polymerase III.

TABLE 4.1 FUNCTIONS OF THE THREE EUKARYOTIC RNA POLYMERASES

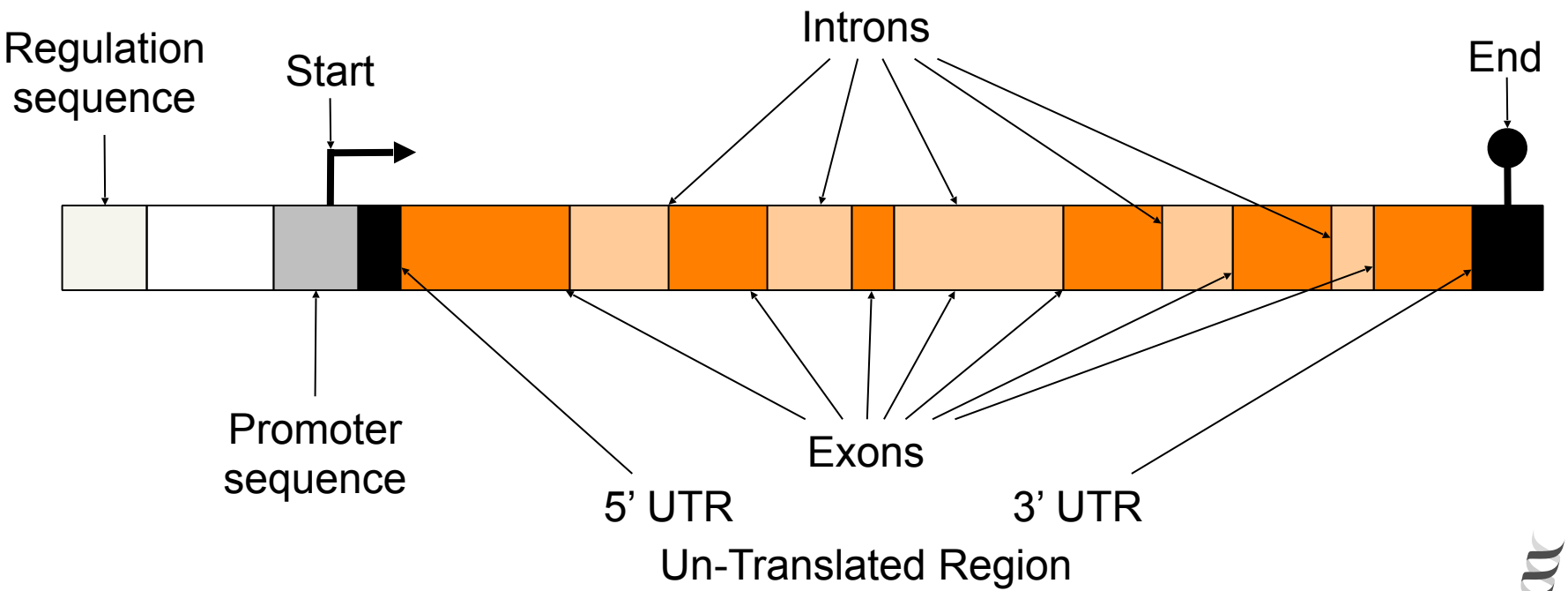
Polymerase	Genes transcribed
RNA polymerase I	28S, 5.8S, and 18S rRNA genes
RNA polymerase II	Protein-coding genes, most snRNA genes, miRNA genes
RNA polymerase III	Genes for 5S rRNA, tRNAs, snoRNAs, U6-snRNA

Table 4.1 Introduction to Genetics (© Garland Science 2012)



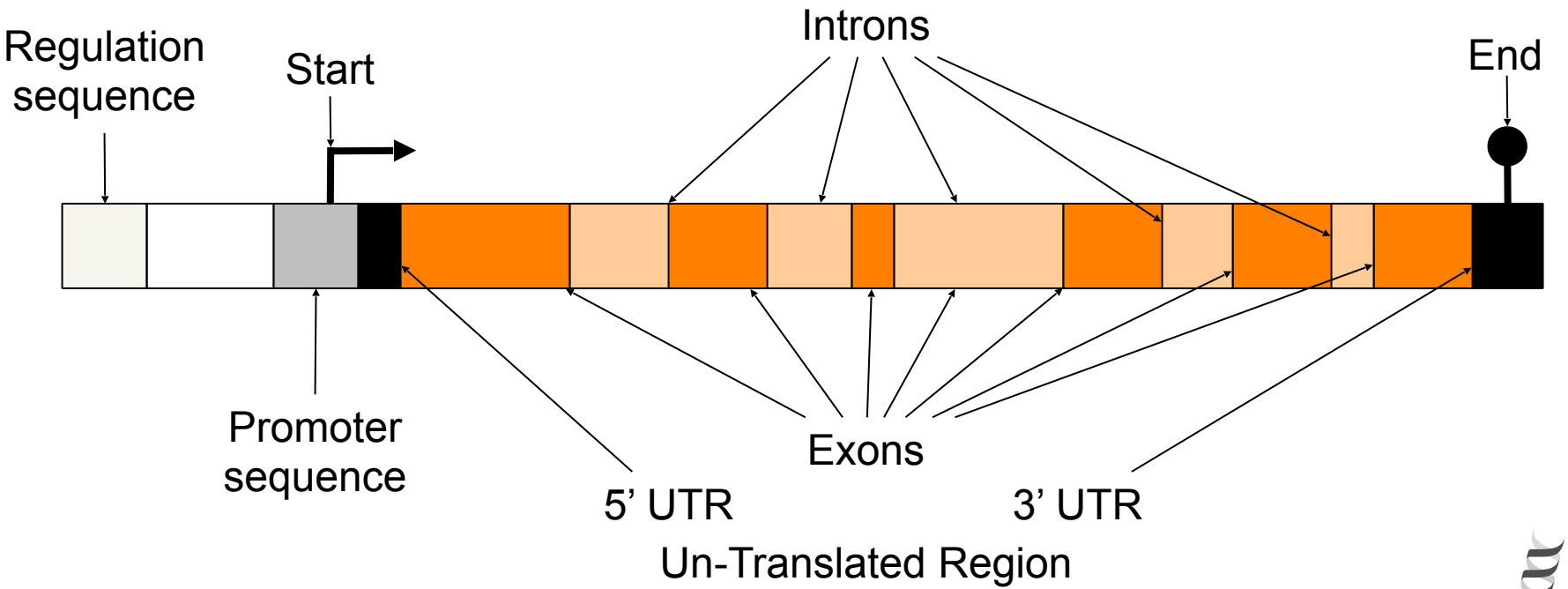
Eukaryotes Gene structure

- Promoter.
- Promoter proximal elements.
- 5’ Un-translated region (5’ UTR).



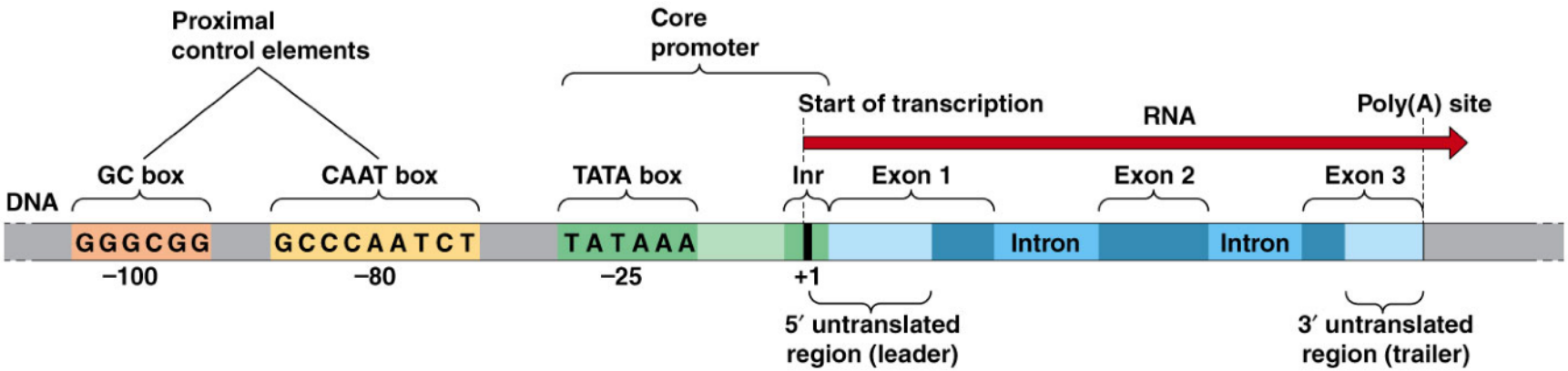
Eukaryotes Gene structure

- Protein coding sequence.
- 3’ Un-translated region (3’ UTR).
- Exons: amino acid coding sequence.
- Introns: no code for amino acid.



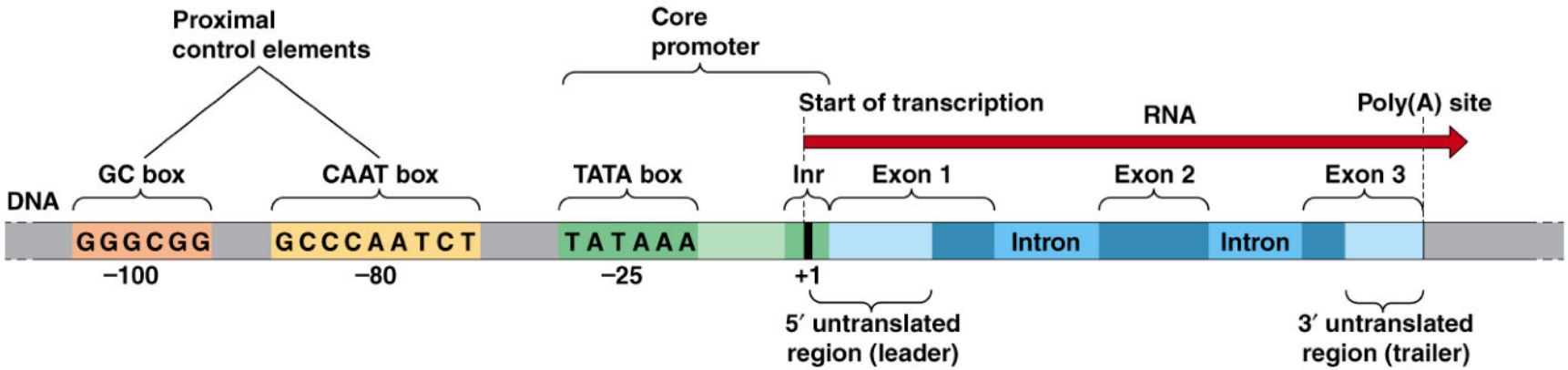
Eukaryotic promoter

- Short sequence called (**initiator**):
  - **Inr** located at (seq +1).
- **TATA box :**
  - Located at -30 sequence.



Eukaryotic promoter

- **Promoter proximal element:**
  - located upstream of the TATA box (-50 - -200)
  - Activators.
  - Enhancers.

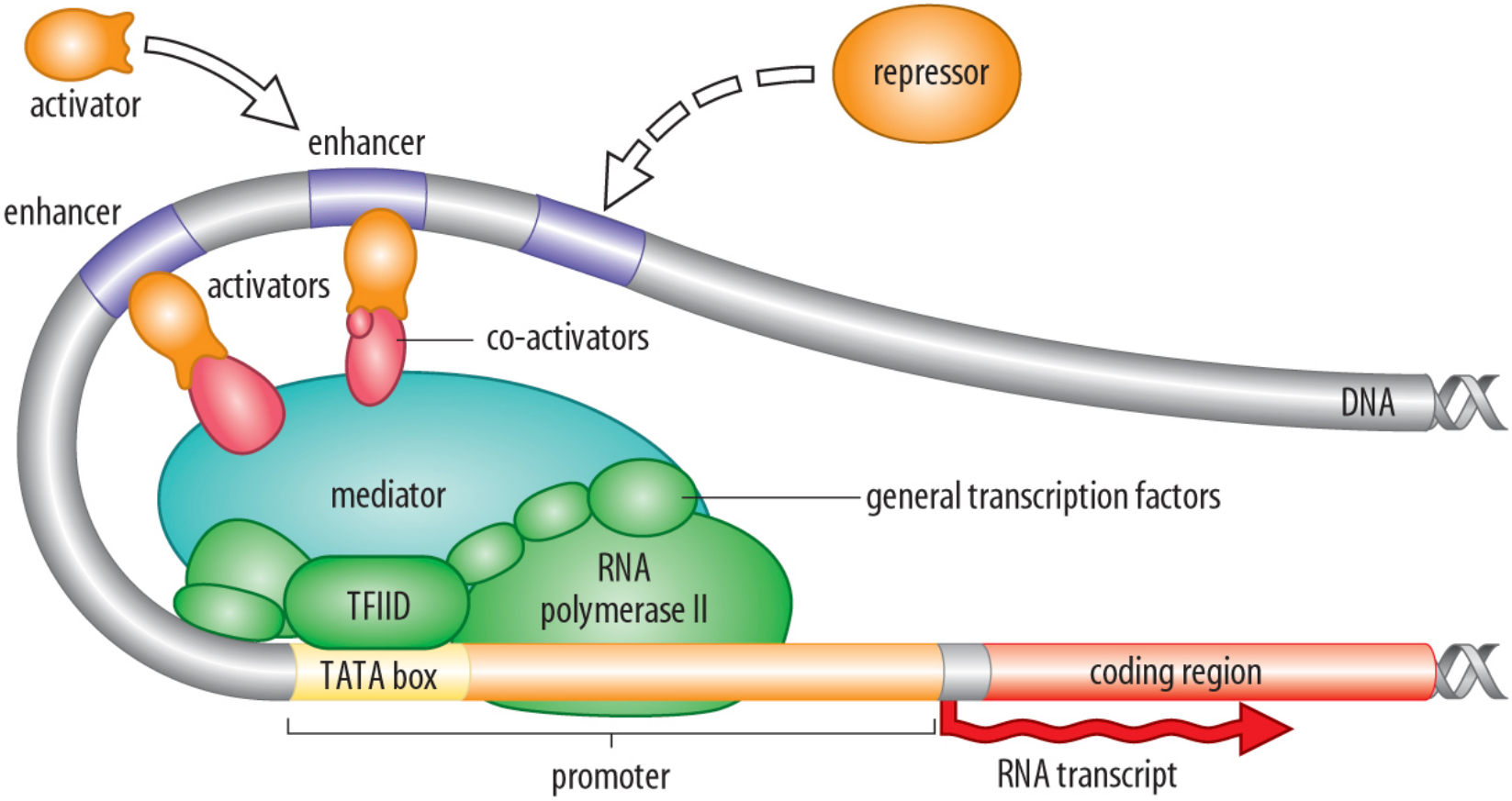


Eukaryotes RNA polymerase

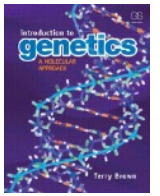
TATA Box

Promoter proximal element:

Activator, Enhancer, or Repressor



- Processing events are more extensive in eukaryotes.
- Like prokaryotes noncoding RNA has to be processed.
- Additionally precursor mRNAs also has to be processed to remove introns, some proteins are also processed.



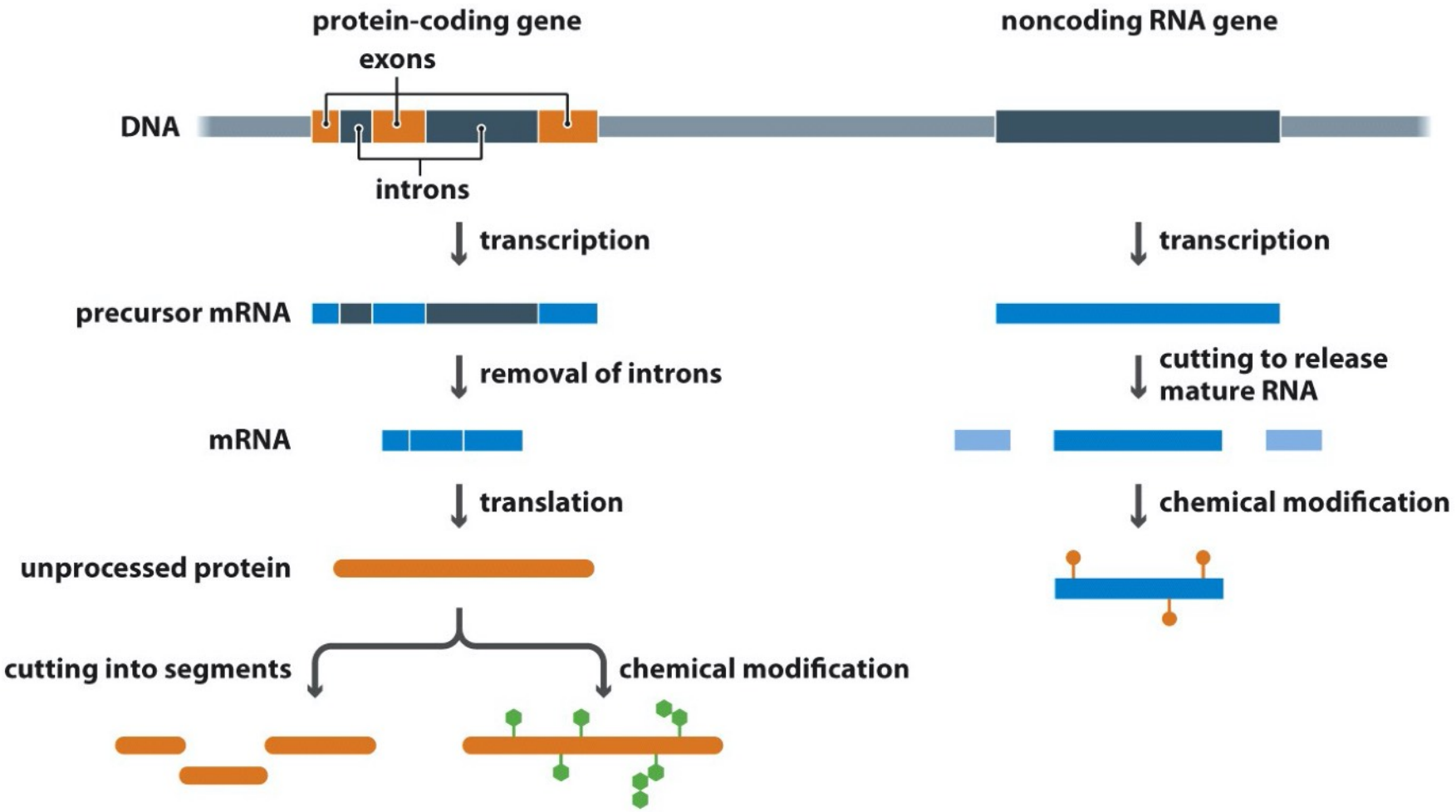
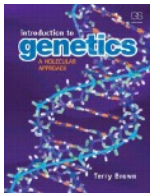


Figure 4.9 Introduction to Genetics (© Garland Science 2012)



• Eukaryotic promoters are more complex.

• For some genes include the core promoter, and upstream elements.

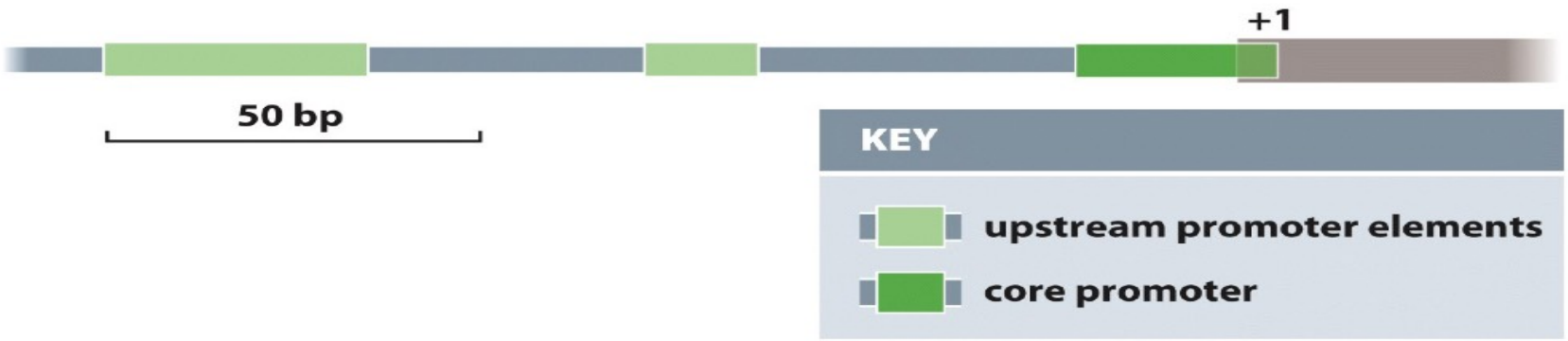
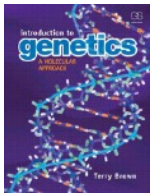


Figure 4.13 Introduction to Genetics (© Garland Science 2012)

• The difference between the promoters defines which genes are transcribed by which polymerase.



• The promoters for RNA polymerase II, stretch several kb upstream of the transcription start site.

• RNA polymerase II core promoter consists of two main segments.

• The first is the –25 or **TATA box, which has the consensus sequence 5'-TATAWAAR-3'.**

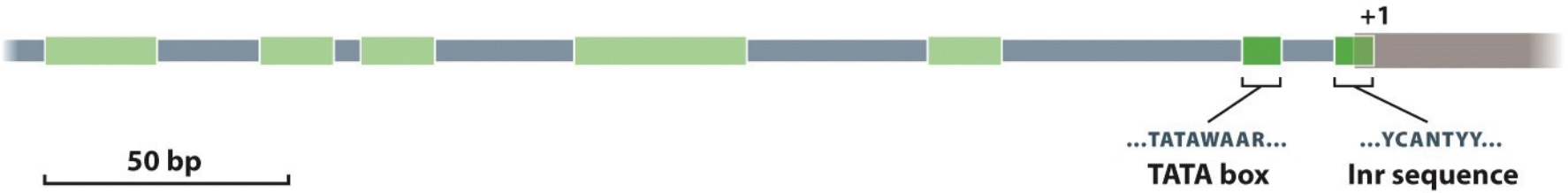
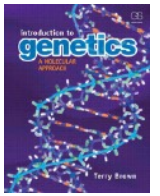


Figure 4.14 Introduction to Genetics (© Garland Science 2012)

Chapter4



• The second part of the core promoter is the initiator (Inr) sequence, located around +1.

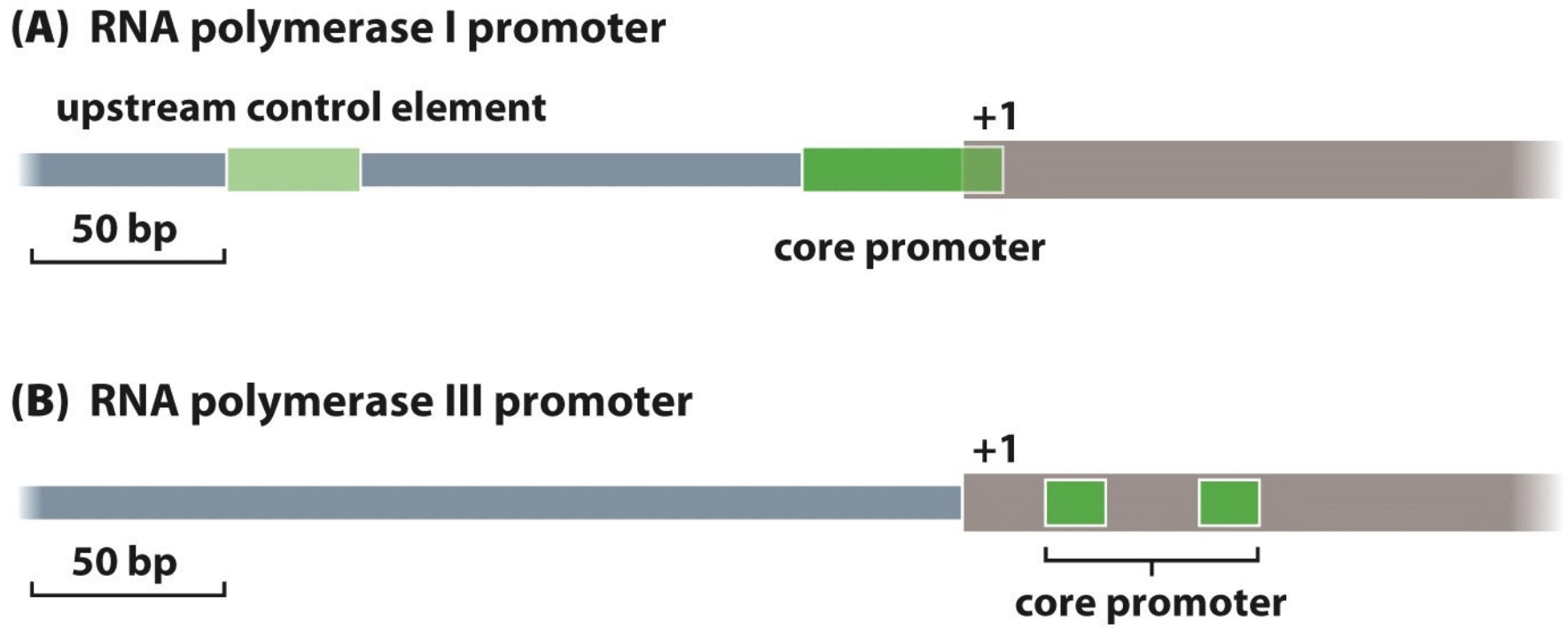
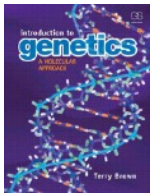


Figure 4.15 Introduction to Genetics (© Garland Science 2012)

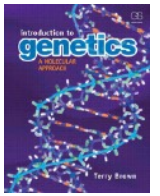


• RNA polymerase I promoters consist of a core promoter region in the transcription start point, **upstream control element** is in nucleotides –45 and +20.

• RNA polymerase III promoters are variable.

• Some are located within the genes whose transcription they promote.

• These sequences span 50 to 100 bp and comprise two segments separated by variable region.



Transcription process in Eukaryotes

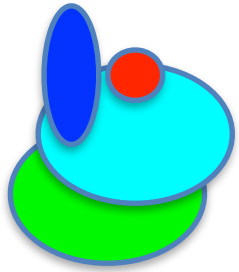
We will consider the transcription process to make m-RNA by RNA Pol II

Transcription process in Eukaryotes

1. **Initiation:** assembly of RNA Pol II with general transcription factors (GTFs) on promoter.

    1. TFIID binds to the TATA box.

    2. **TFIID is composed of two units:**



        - TATA boxes binding protein (TBP)
        - TBP associated factors (TAFs)

    3. TFIIA and TFIIB bind to TFIID.

Transcription process in Eukaryotes

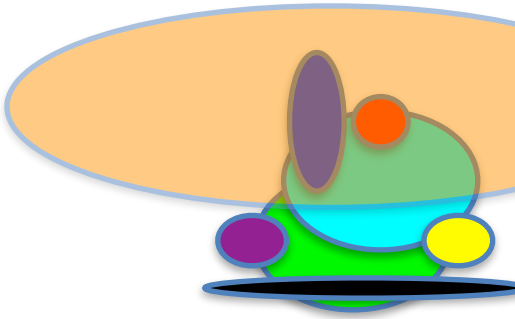
1. **Initiation:** assembly of RNA Pol II with general transcription factors (GTFs) on promoter.

    5. RNA Pol II and TFIIF bind to (TFIID, TFIIA, TFIIB)

    6. TFIIE binds to the complex.

    7. TFIIH (helicase) binds to the complex – will be unwinding the promoter.

    8. All elements make the transcription **Pre-initiation complex (PIC).**

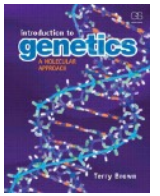


4.4 INITIATION OF TRANSCRIPTION IN BACTERIA AND EUKARYOTES

- The σ subunit recognizes the bacterial promoter
- Formation of the RNA polymerase II initiation complex
- Initiation of transcription by RNA polymerases I and III

• For RNA polymerase II, the initial contact is made by the protein called **transcription factor IID.**

• **It is a multisubunit** complex, made up of the **TATA-binding protein (TBP) and twelve TBP-associated factors, or TAFs.**



- Initiation complex must be activated before transcription.

- Activation involves addition of phosphate to **C-terminal** domain (CTD).

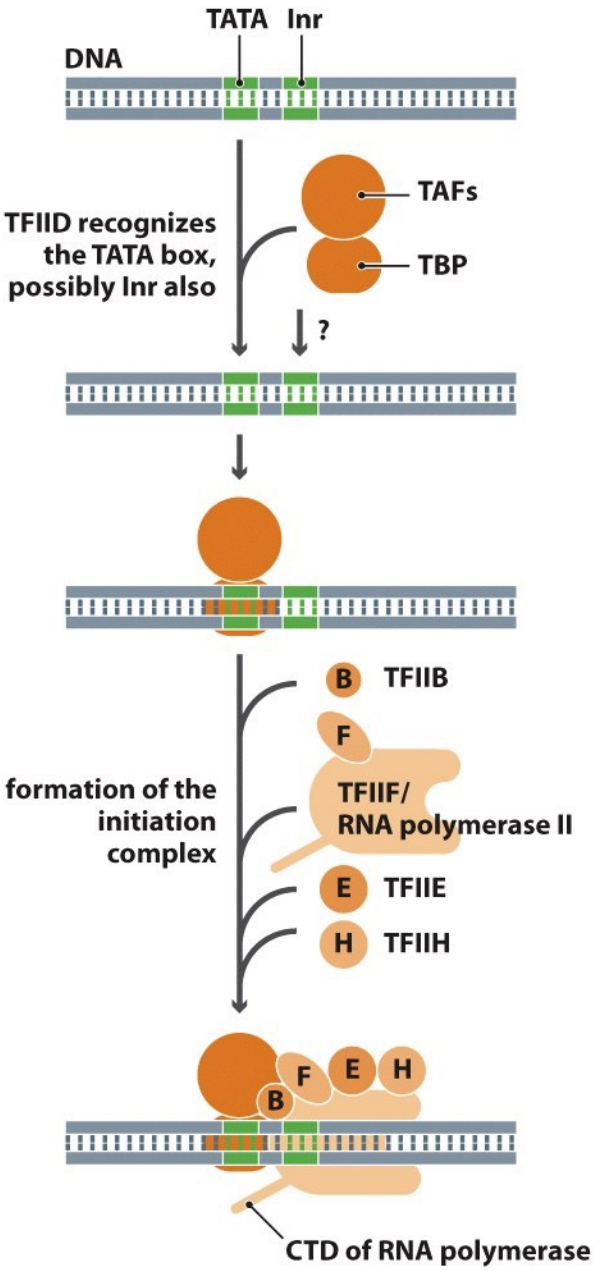
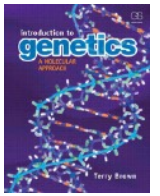


Figure 4.20 Introduction to Genetics (© Garland Science 2012)



• Once these phosphates have been attached, the polymerase is able to leave the initiation complex and begin synthesizing RNA.

• After departure of the polymerase, transcription factors except TFIID and TFIIH remain, enabling re-initiation to occur if needed.

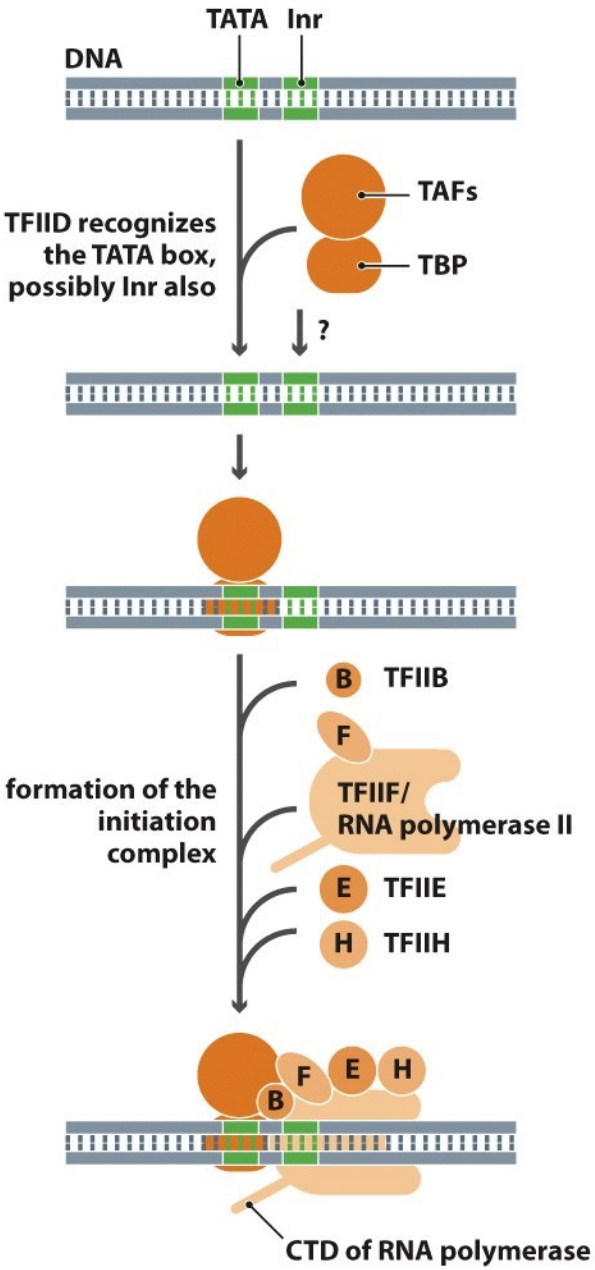
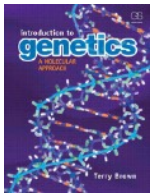


Figure 4.20 Introduction to Genetics (© Garland Science 2012)



- RNA polymerase I forms an initiation complex with four multi-subunit proteins, one contains TBP.

•This initiation complex then attaches to the RNA polymerase I core promoter.

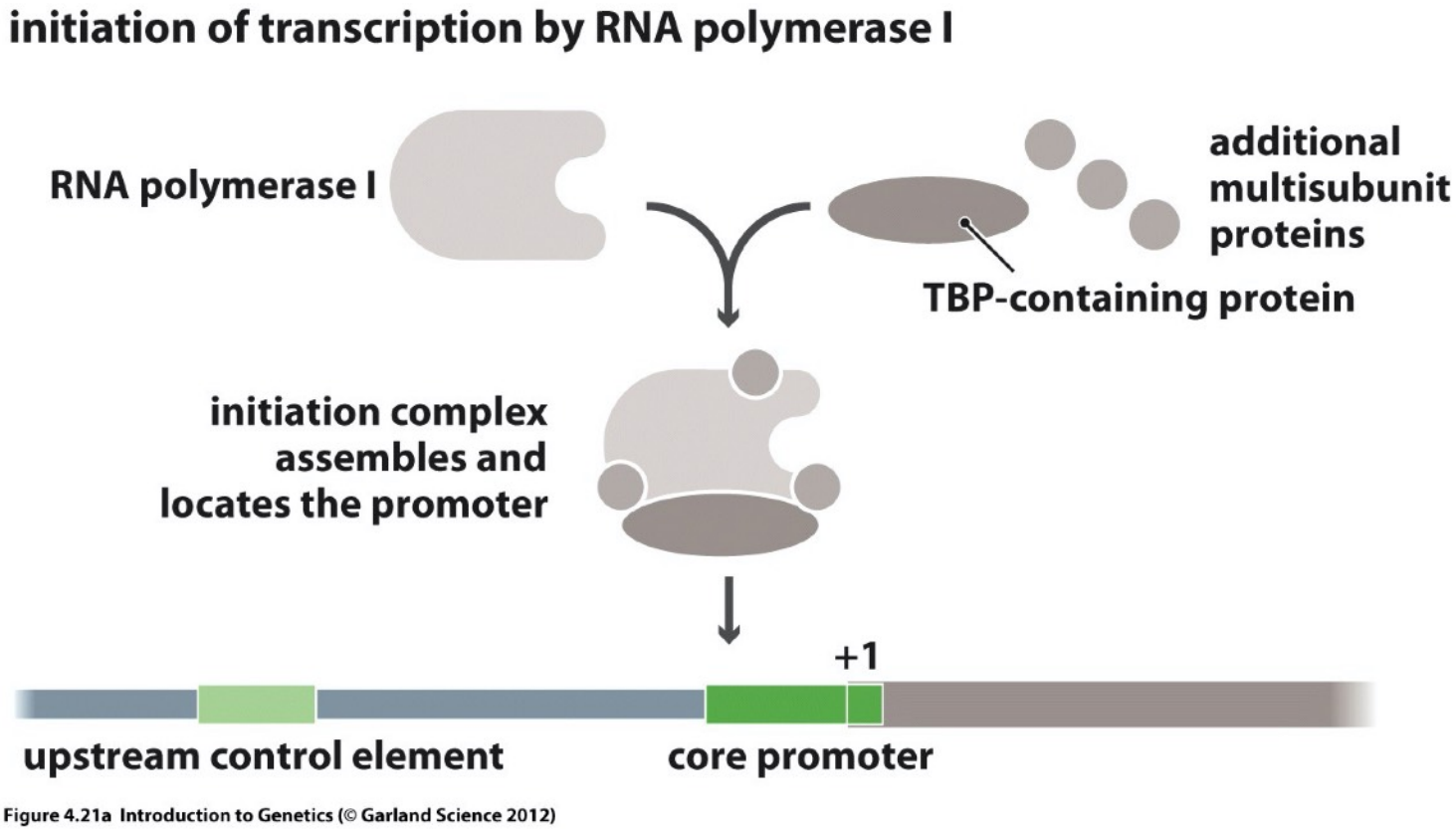
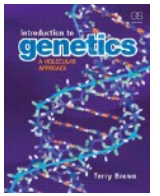


Figure 4.21a Introduction to Genetics (© Garland Science 2012)



- Initiation by RNA polymerase III begins with attachment of TBP of TFIIIB to core promoter.

- The RNA polymerase then attaches to the bound TBP.

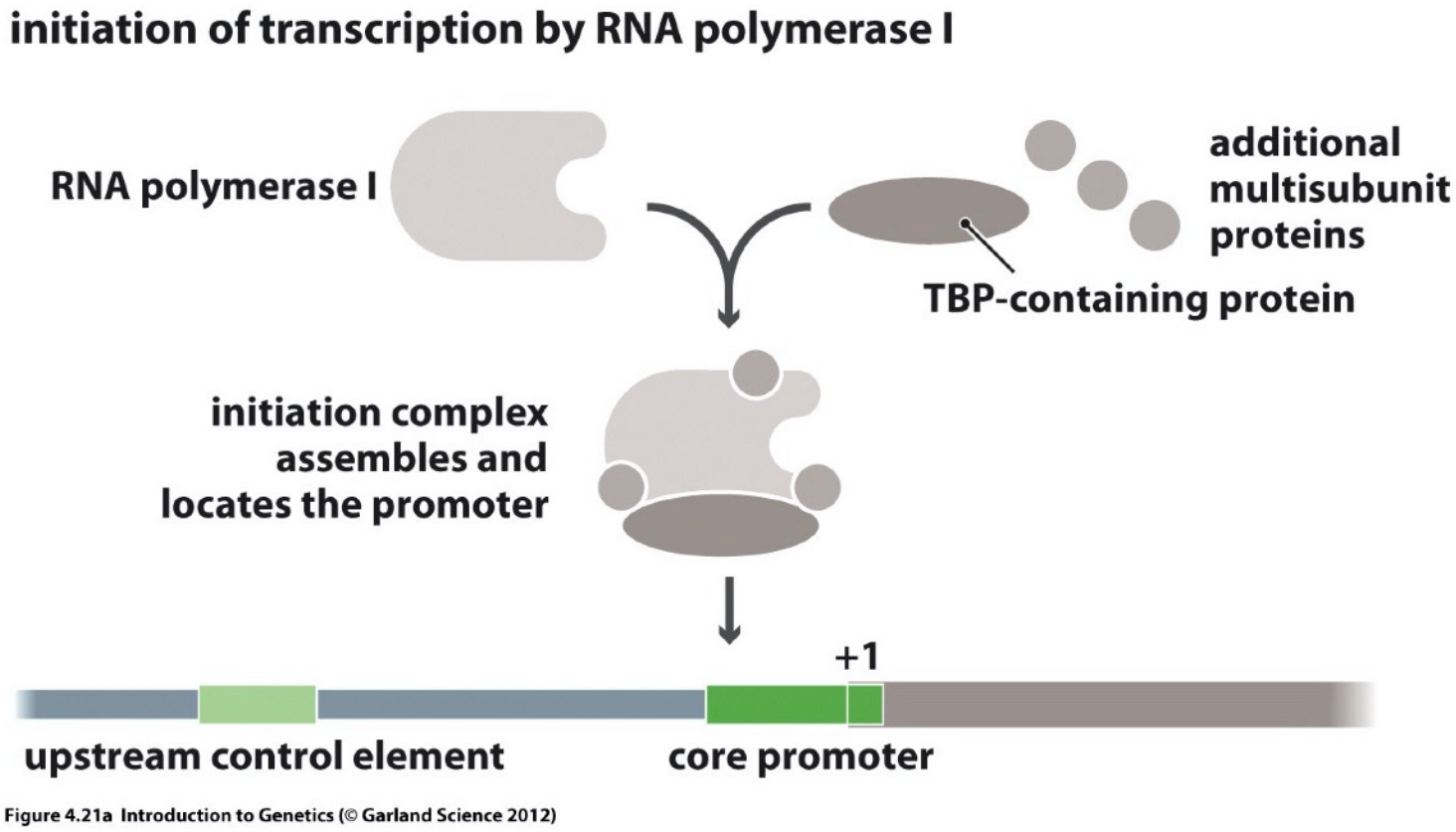
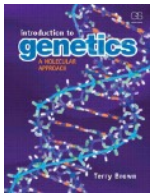


Figure 4.21a Introduction to Genetics (© Garland Science 2012)



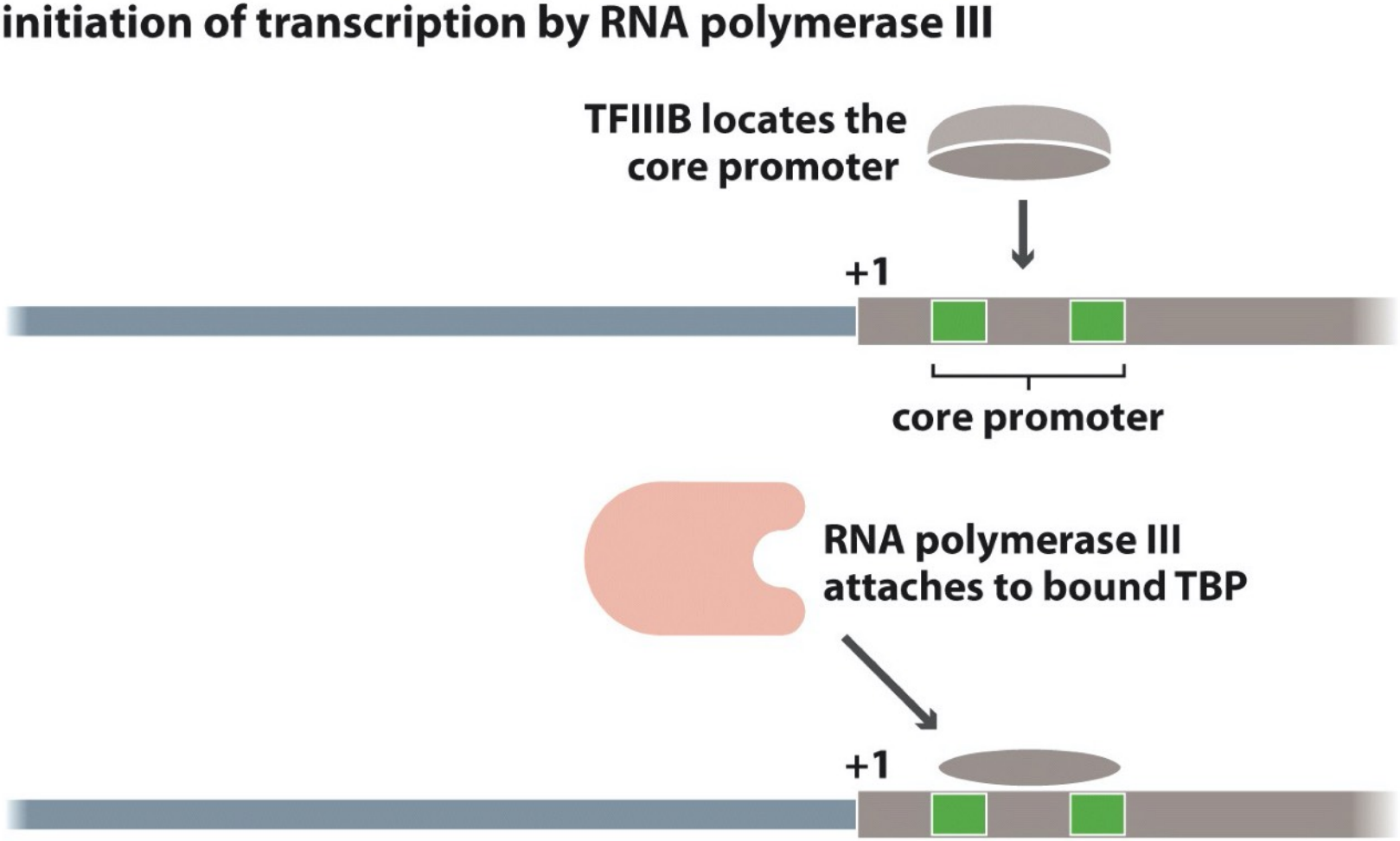
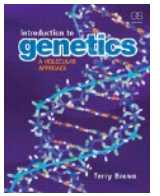
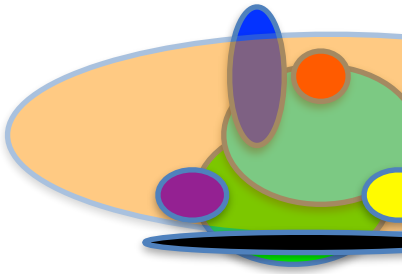


Figure 4.21b Introduction to Genetics (© Garland Science 2012)



Transcription process in Eukaryotes



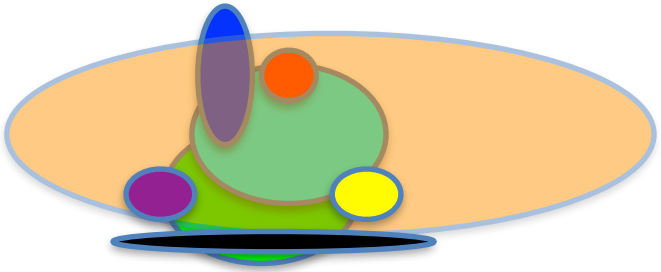
2. **Promoter melting by TFIIH:**

   - The helicase activity of TFIIH unwind the promoter sequence to use a template.

3. **Abortive initiation:**

   - The RNA polymerase synthesizes short sequences of RNA but can not escape the promoter.

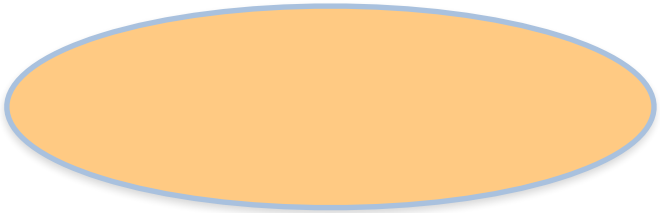
Transcription process in Eukaryotes



4. **Promoter escape:**

   - RNA Pol breaks the interactions with transcription factors and escape the promoter and start synthesizing RNA.

Transcription process in Eukaryotes



5. **Elongation:**

- RNA Pol adds the correct complementary NTPs to the template and continue. IF errors made proofreading capabilities fixes the problems.

Transcription process in Eukaryotes

6. **Termination:**

- When RNA Pol synthesize pass a poly(A) site in the transcript (5’ AAUAAA 3’), the RNA synthesized is cleaved by:
    - CPSF (cleavage and polyadenylation specificity factor) protein.
    - CstF (cleavage stimulating factor) protein.
    - CFI and CFII (cleavage factor proteins).

• The three eukaryotic RNA polymerases each use different mechanisms for transcript termination.

• RNA polymerase I: termination involves a DNA-binding protein that attaches DNA at a recognition sequence located 12 to 20 bp downstream of the point at which transcription terminates.

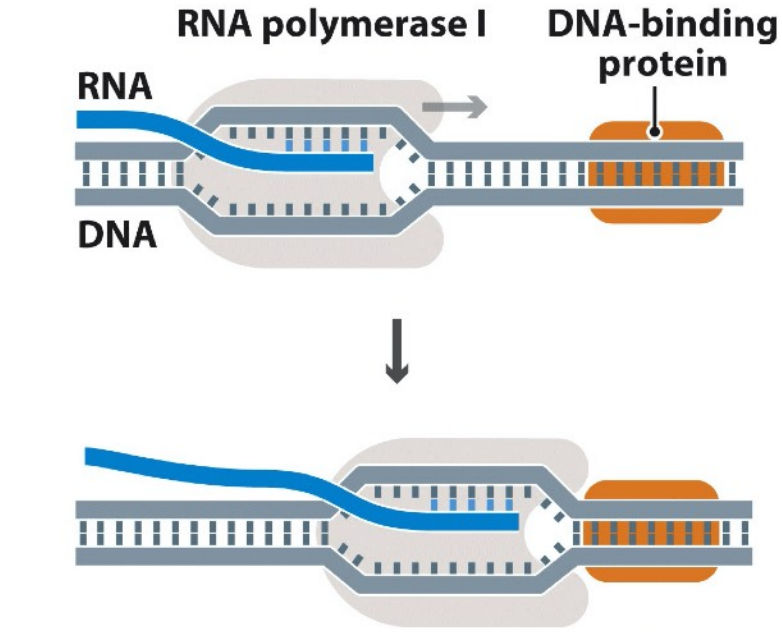
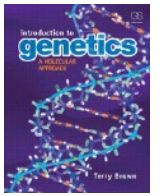


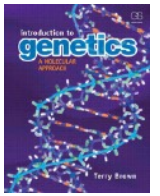
Figure 4.31 Introduction to Genetics (© Garland Science 2012)

Chapter4



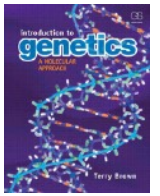
- RNA polymerase II: termination is accompanied by addition of a series of adenosine nucleotides, called the **poly(A) tail, to the 3' end of the mRNA** that is added by a DNA-*independent* RNA polymerase called **poly(A) polymerase.**

• This polymerase act at an internal site **polyadenylation site** which is cleaved to create a new 3' end to which the poly(A) tail is added.



• In mammals, **polyadenylation site is located between 10 and 30 nucleotides** downstream of a signal sequence, almost always 5'-AAUAAA-3'.

• The signal sequence is the binding site for a multisubunit protein **cleavage and polyadenylation specificity factor (CPSF).**



• **CPSF** attaches to the polymerase complex during initiation of transcription, and ride along the template with RNA polymerase II.

• It binds to the signal sequence as soon as it is transcribed, initiating the polyadenylation reaction.

• As a result, transcription stops soon after the poly(A) signal sequence has been transcribed.

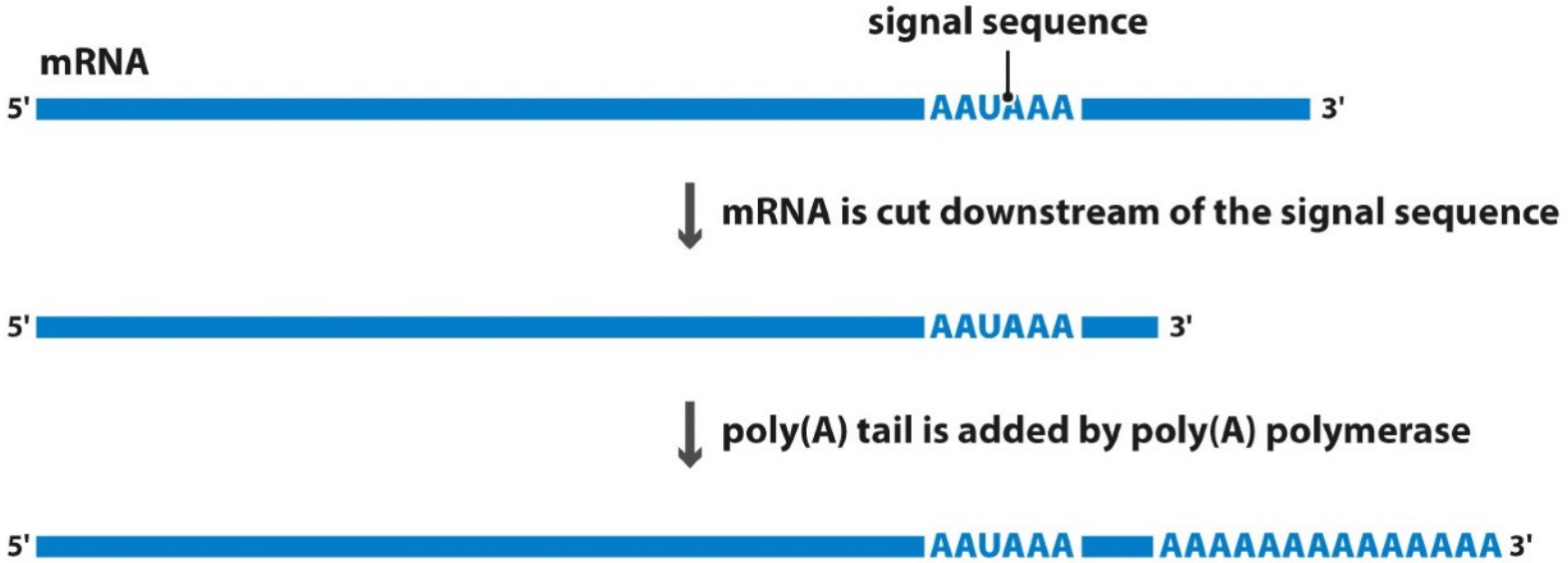
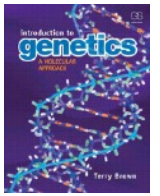


Figure 4.33 Introduction to Genetics (© Garland Science 2012)



pre-mRNA

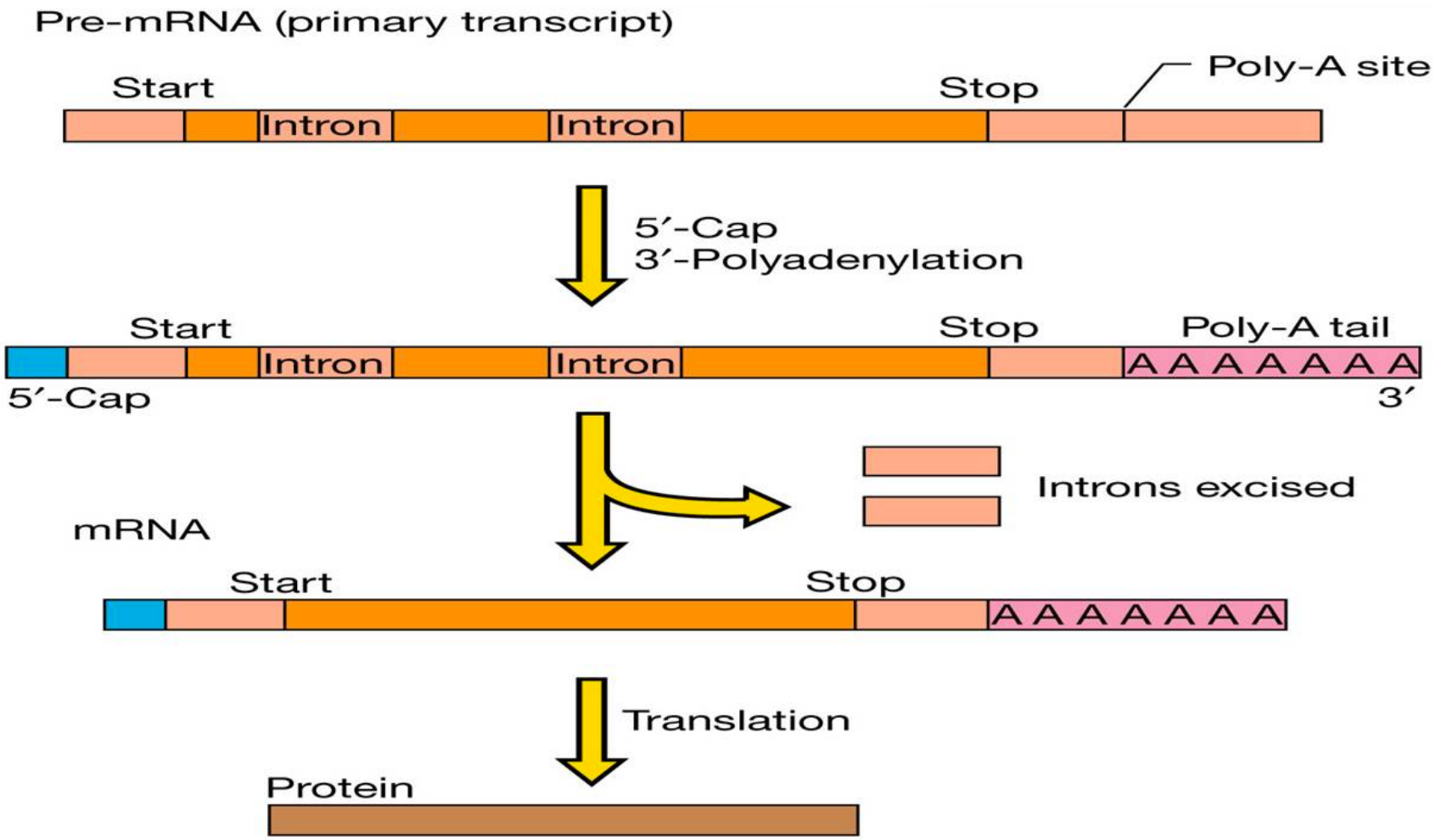
- The first product of transcription in eukaryotes is called pre-mRNA.
- Unlike the transcript of a prokaryote that get translated while being transcribed (**coupled transcription and translation**), the eukaryotic transcript needs some modification before leaving the nucleus to be translated.
- The pre-mRNA of eukaryotes gets processed to a mature mRNA.

pre-mRNA

What are the processes to produce mature mRNA?

(1) 5’ modification
(2) 3’ modification
(3) Intron splicing

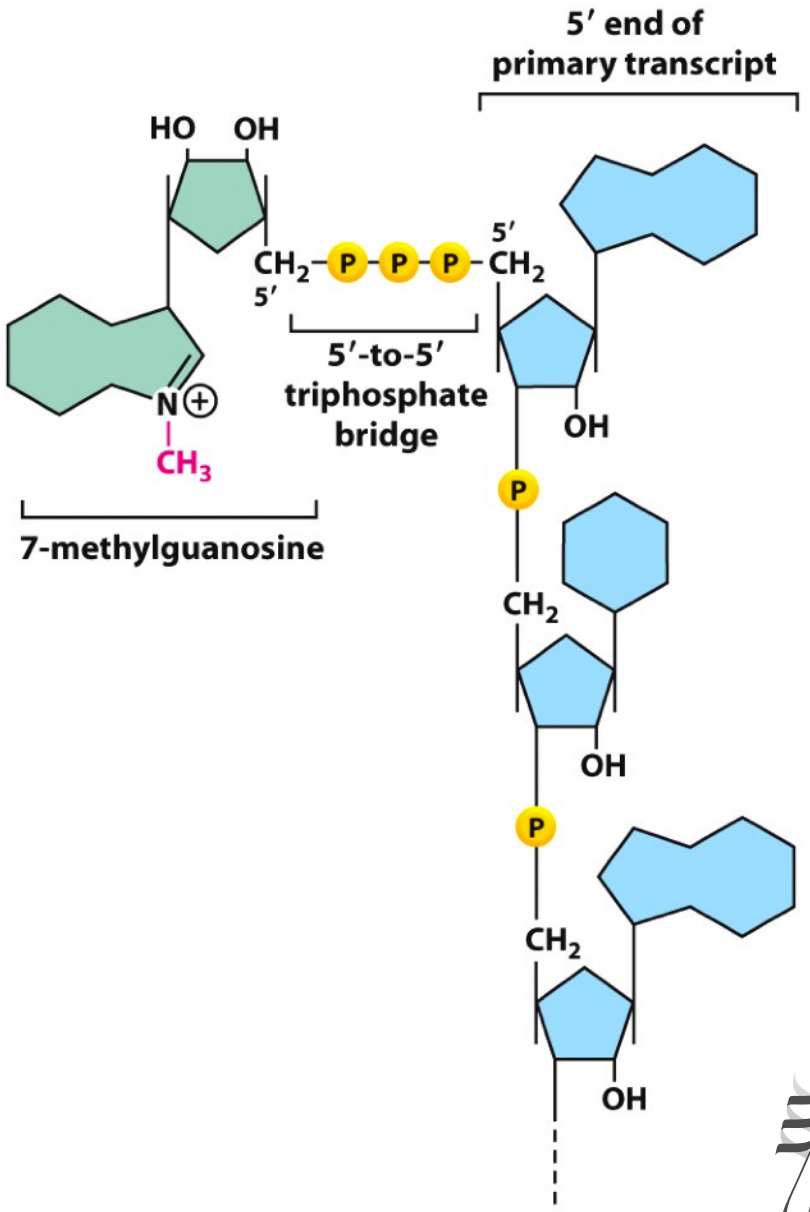
pre-mRNA maturation



5’ modification of eukaryotic mRNA

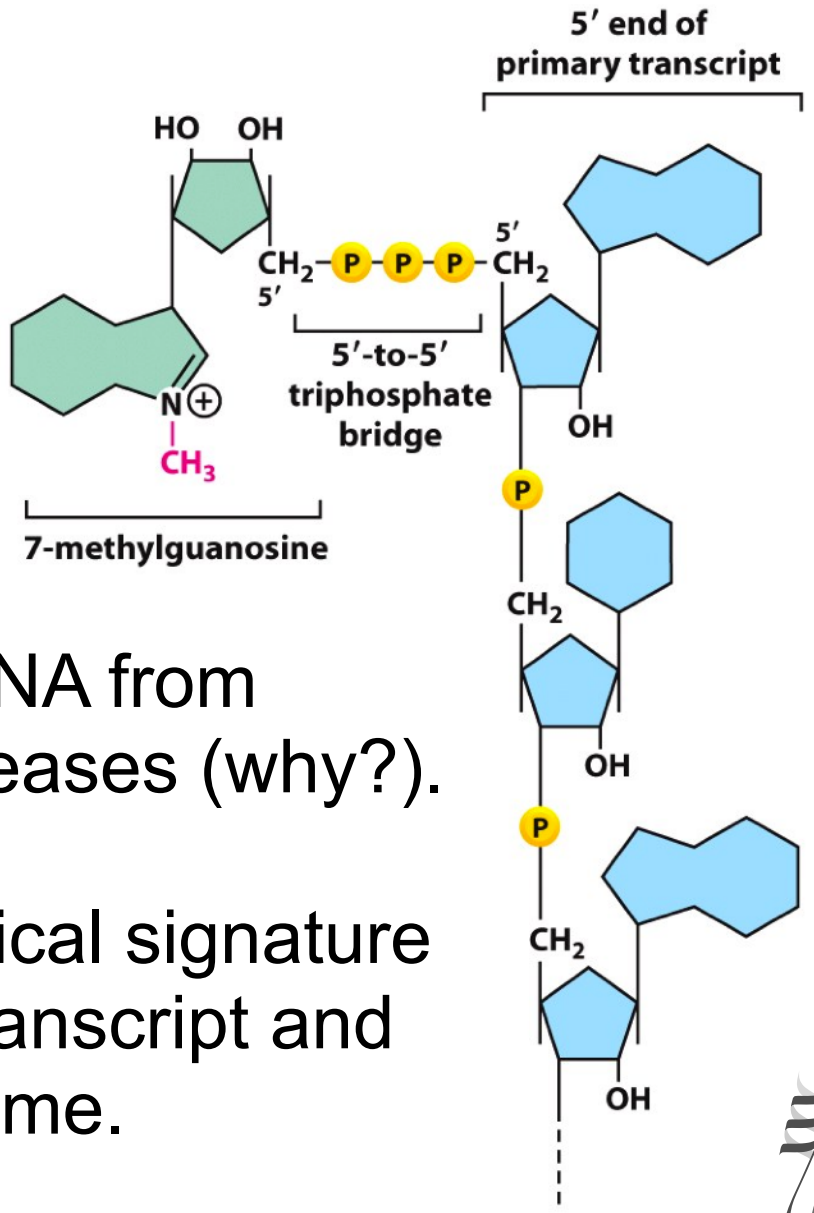
- It is referred to as the **Capping of the 5’ end.**

- Adding a 7-methyl Guanine (m^7G) to the 5’ end using a (5’ – 5’) linkage.



5’ modification of eukaryotic mRNA

This modification is important for:



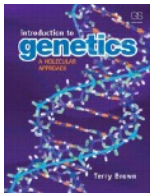
- Protecting the mature mRNA from degradation from exonucleases (why?).
- This is an important chemical signature for the translation of the transcript and initial binding to the ribosome.

• There is usually a triphosphate group at the 5’ end of a polynucleotide.

• The RNAs made by RNA polymerase II are exceptions to this rule.

• 5’ terminus has a more complex chemical structure, a 7-MeG is the nucleotide carrying the modified base 7-methylguanosine. This is referred to as the **cap structure.**

• It is put in place soon after the RNA polymerase leaves the promoter, before the mRNA reaches 30 nucleotides in length.



Synthesis of the type 0 cap.

- The nucleotide at the 5’ end “pppN” (terminal triphosphate).

- First GTP is attached to the terminal nucleotide, forming a 5’–5’ triphosphate one phosphate from the GTP and two from terminal nucleotide of the mRNA.

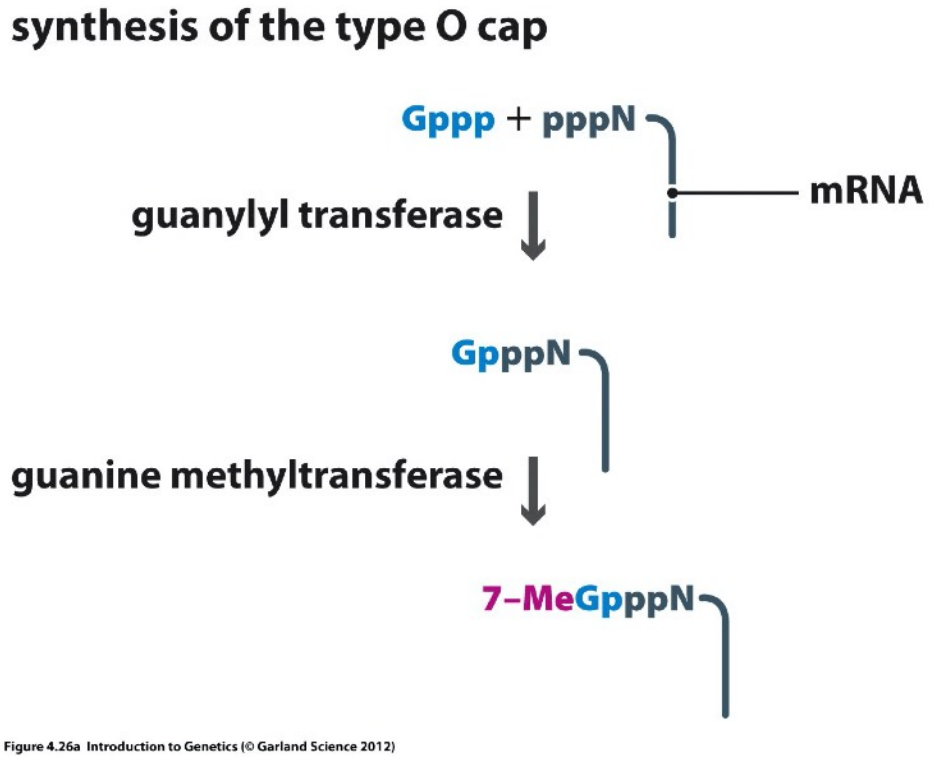
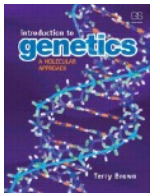


Figure 4.26a Introduction to Genetics (© Garland Science 2012)

Chapter4



• In second step a methyl group is added to position 7 of the guanine base.

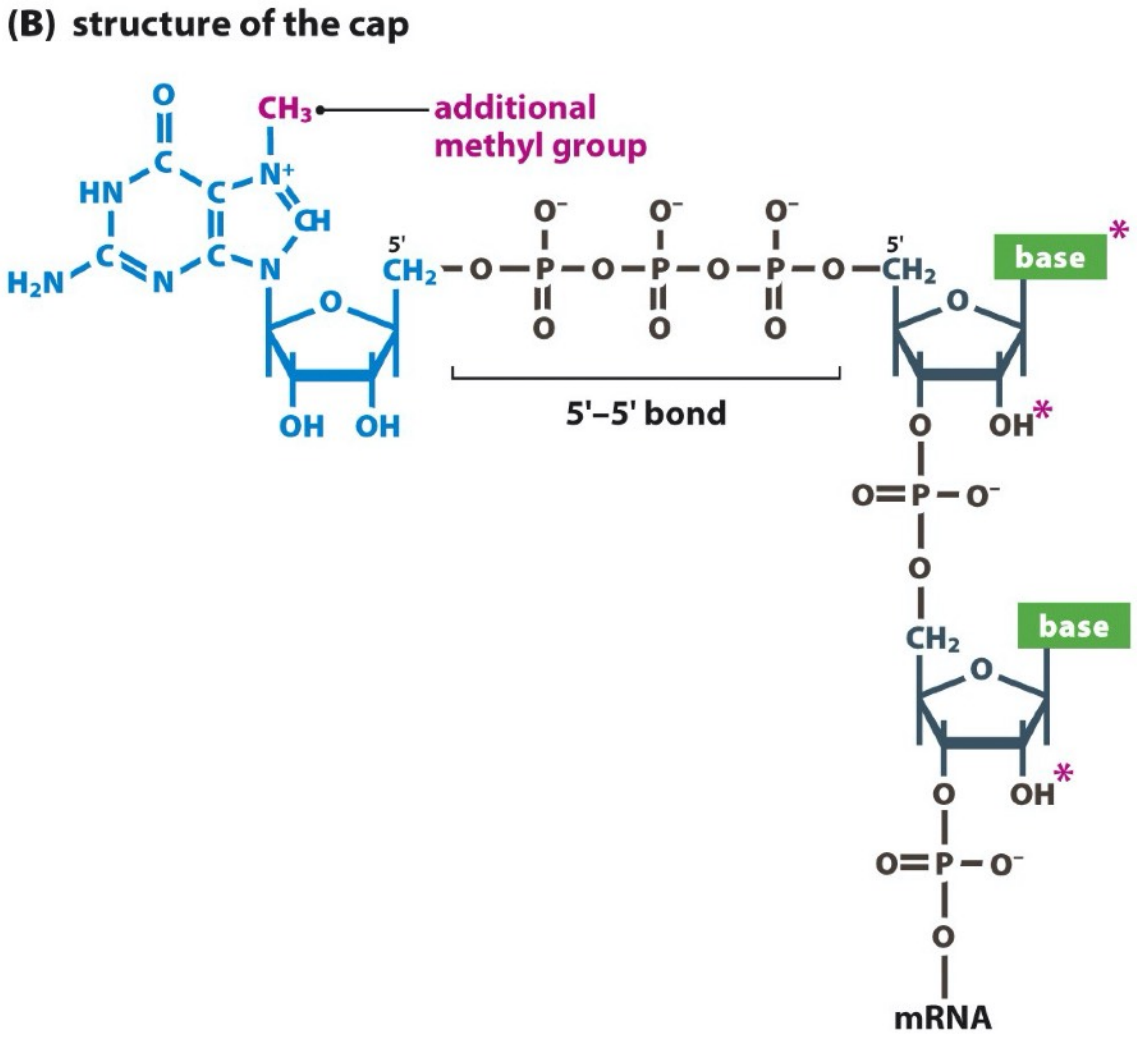
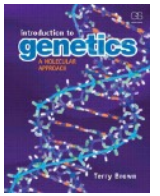


Figure 4.26b Introduction to Genetics (© Garland Science 2012)

Chapter4



3’ modification of eukaryotic mRNA

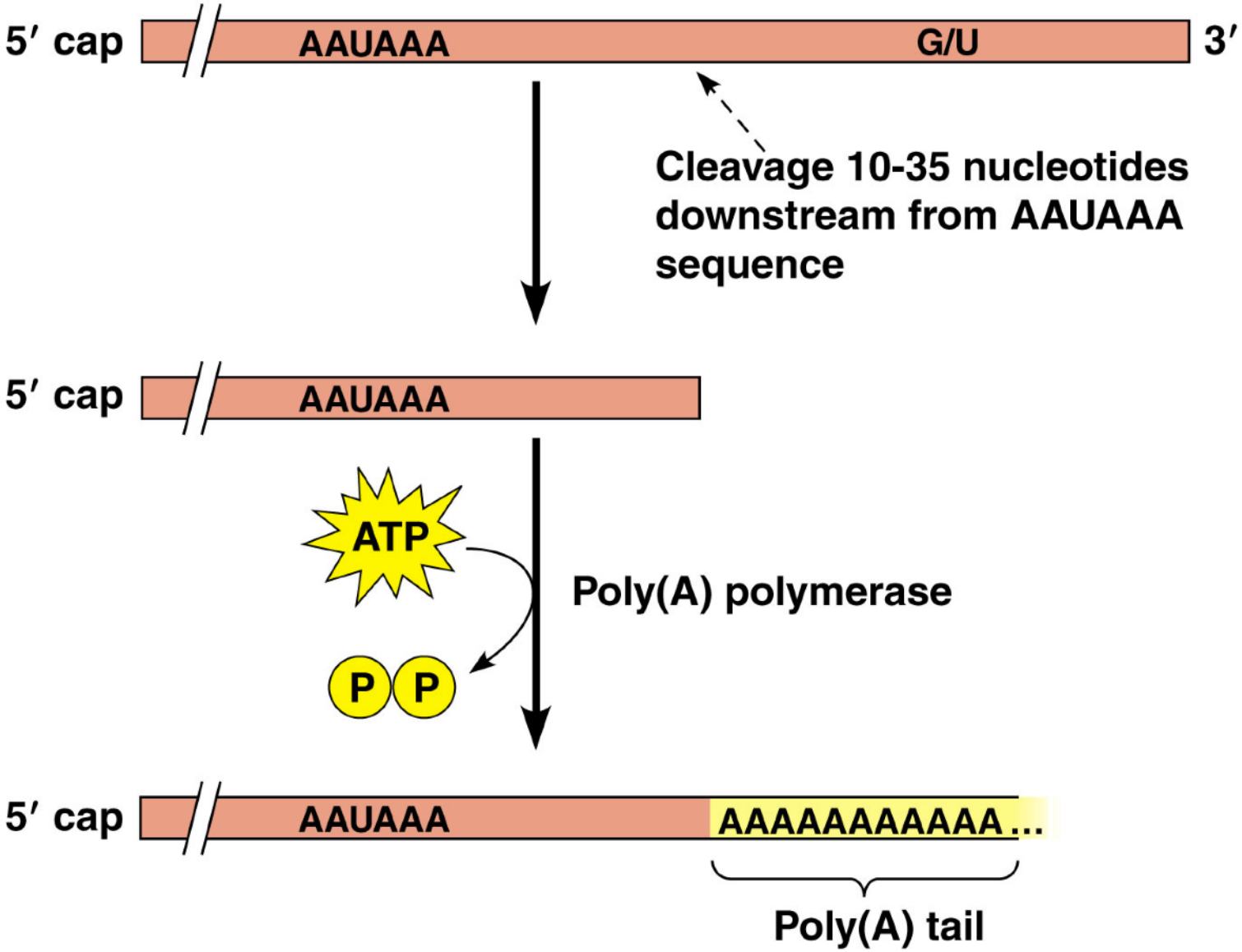
- This modification involve the addition of 50-250 adenine nucleotides at the 3’ end.
- This is called **poly(A) tails.**
- The poly(A) tails is synthesized without a template by an enzyme called **poly(A) polymerase (PAP).**

3’ modification of eukaryotic mRNA

Why poly(A) tails?

- Help to export the mRNA from the nucleus.
- Prevent the mRNA from degradation by exonucleases.
- Helps in the initiation of translation.
- Helps in the regulation of translation.
- Helps to terminate transcription.

3’ modification of eukaryotic mRNA



Intron splicing

- Introns are the non-coding sequence of the gene that get transcribed but has to be removed before translation.
- Introns usually start with **5’-GU**.
- Introns usually end with **3’-AG**.
- The removal of introns from the pre-mRNA is called **intron splicing.**

Intron splicing

- The molecular machinery that carries out splicing of the premature transcript is called **spliceosome.**

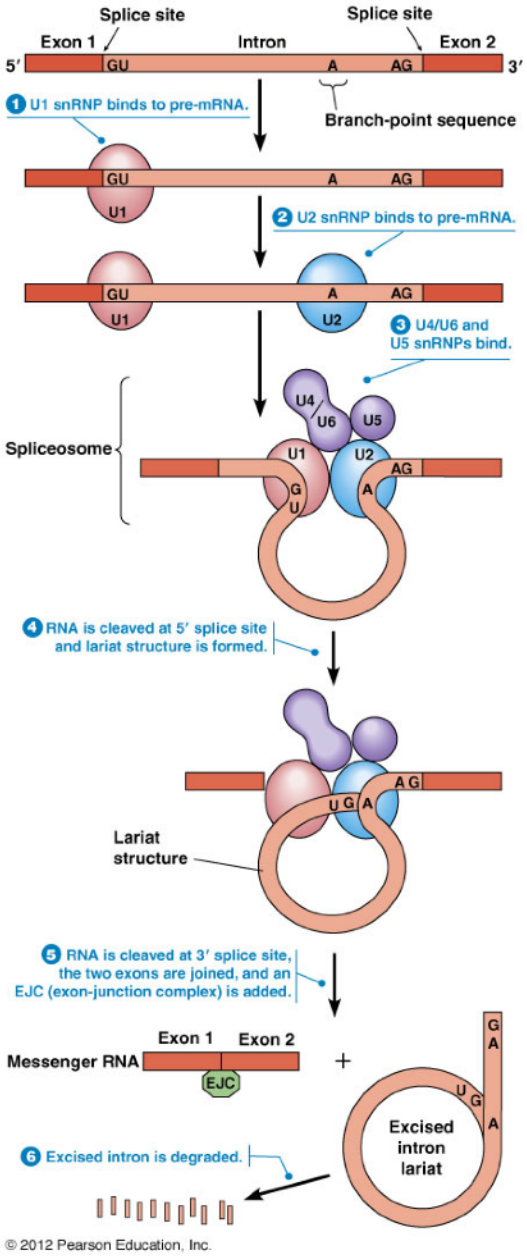
- The pre-mRNA binds to **small nuclear ribonucleoprotein particles (snRNPs).**

- **snRNPs** (U1, U2, U4, U5, U6) **are composed of:**
    - Small nuclear RNA (snRNA).
    - Associated proteins.

Intron splicing process

1. U1 snRNP binds to the introns 5’ bind site.

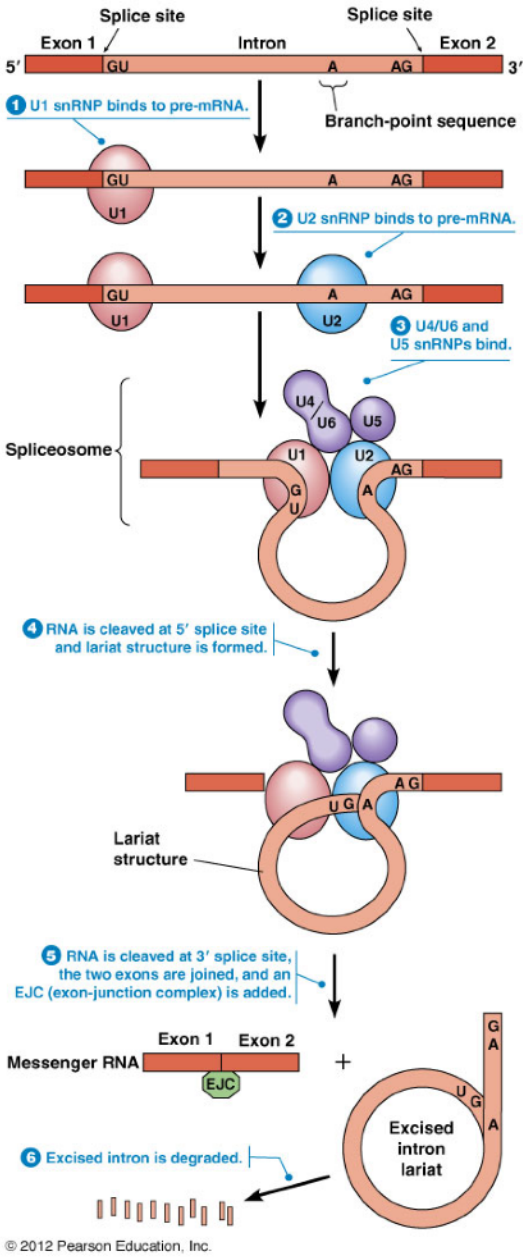
2. U2 snRNP binds to the **branch point sequence** in the middle of the intron (usually Adenine (A)).



Intron splicing process

3. U4/U6 snRNP and U5 snRNP interact forming a unit.

4. The U4,U5,U6 bind to U1 and U2 making the intron loop and bend.



Intron splicing process

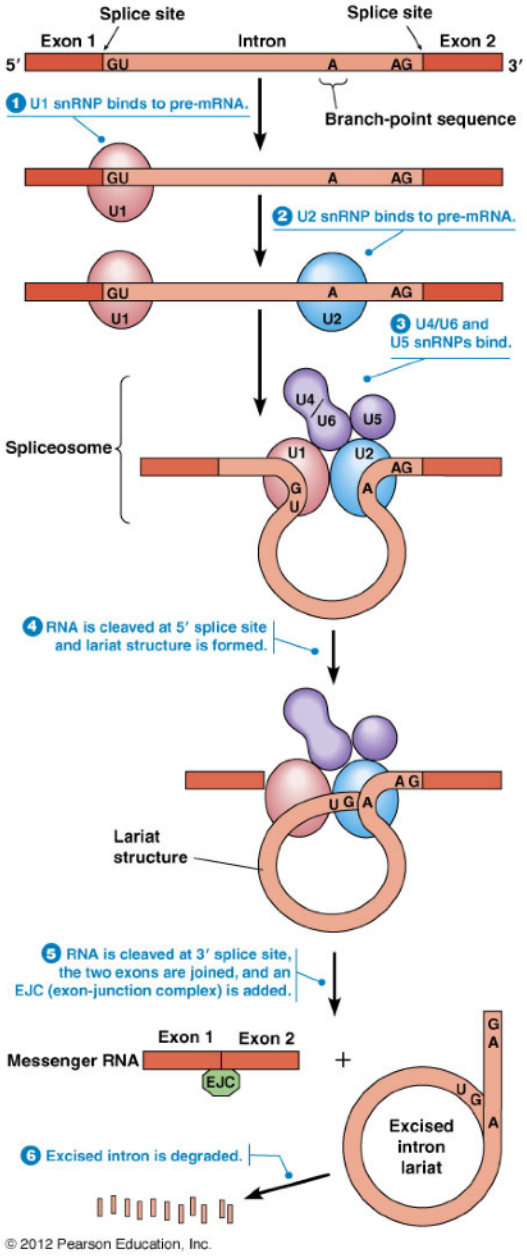
5. This brings the 5’ splice site close to the 3’ splice site.

6. U4 snRNP dissociates which activate the **spliceosome.**

7. The splice sites are excised and the intron is removed.

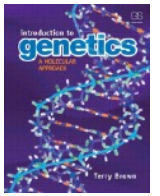
8. The two exons are ligated.

9. This takes place for every two exons in the gene.



• The nucleus can be divided into the **nucleolus, where ribosomal RNA genes** are transcribed, and the **nucleoplasm, where other genes, including** those for mRNA, are transcribed.

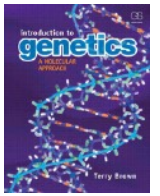
• The nucleoplasmic RNA fraction is called **heterogeneous nuclear RNA.**



• It is made up of a complex mixture of RNA molecules, many over 20 kb in length.

• The mRNA in the cytoplasm is also heterogeneous, but its average length is only 2 kb.

• If the mRNA in the cytoplasm is derived from hnRNA, then the primary transcripts are being shortened before the mRNA leaves the nucleus.



• There are consensus sequences for intron- exon splice site , which vary in different types of eukaryote.

• In vertebrates :

5’ splice site: 5’-AGGUAAGU-3’

3’ splice site: 5’-PyPyPyPyPyPyNCAG-3’

Py is one of the two pyrimidine nucleotides (U or C), N is any nucleotide, and the arrow indicates the exon–intron boundary.

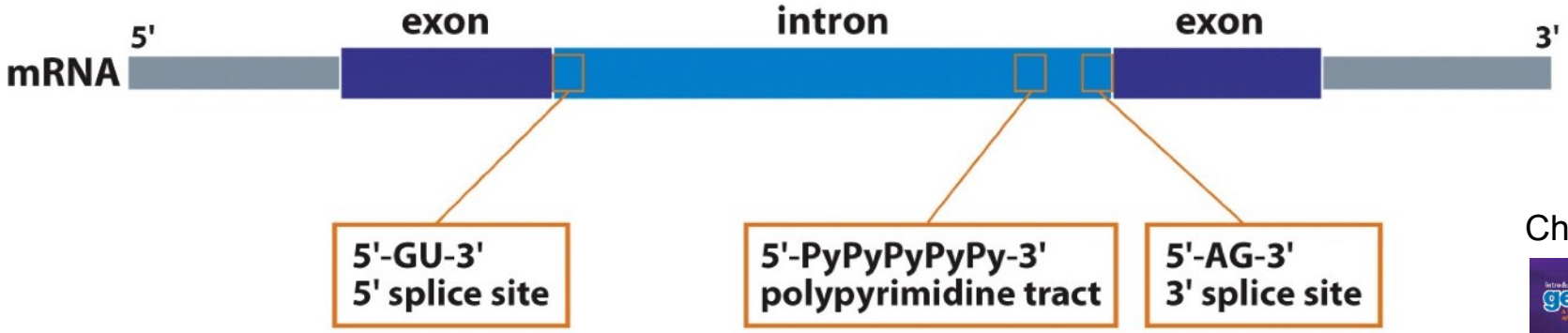
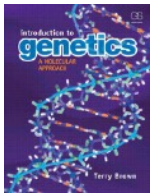


Figure 5.18 Introduction to Genetics (© Garland Science 2012)

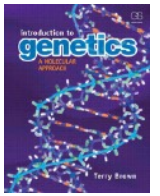


• The splicing pathway involves three steps.

• The first step is cleavage at the 5’ splice site, called the **donor site.**

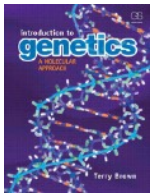
• The resulting free 5’ end is then attached to an internal branch site within the intron to form a lariat structure.

• The lariat is formed by creating a phosphodiester bond between the 5’ carbon of the first nucleotide of the intron (the G of the GU motif) and the 2’ carbon of the internal adenosine.



• The 3’ splice site—the **acceptor site**—is then cleaved and the two exons joined together.

• The intron is released as the lariat structure, which later converts back to linear RNA and degraded.



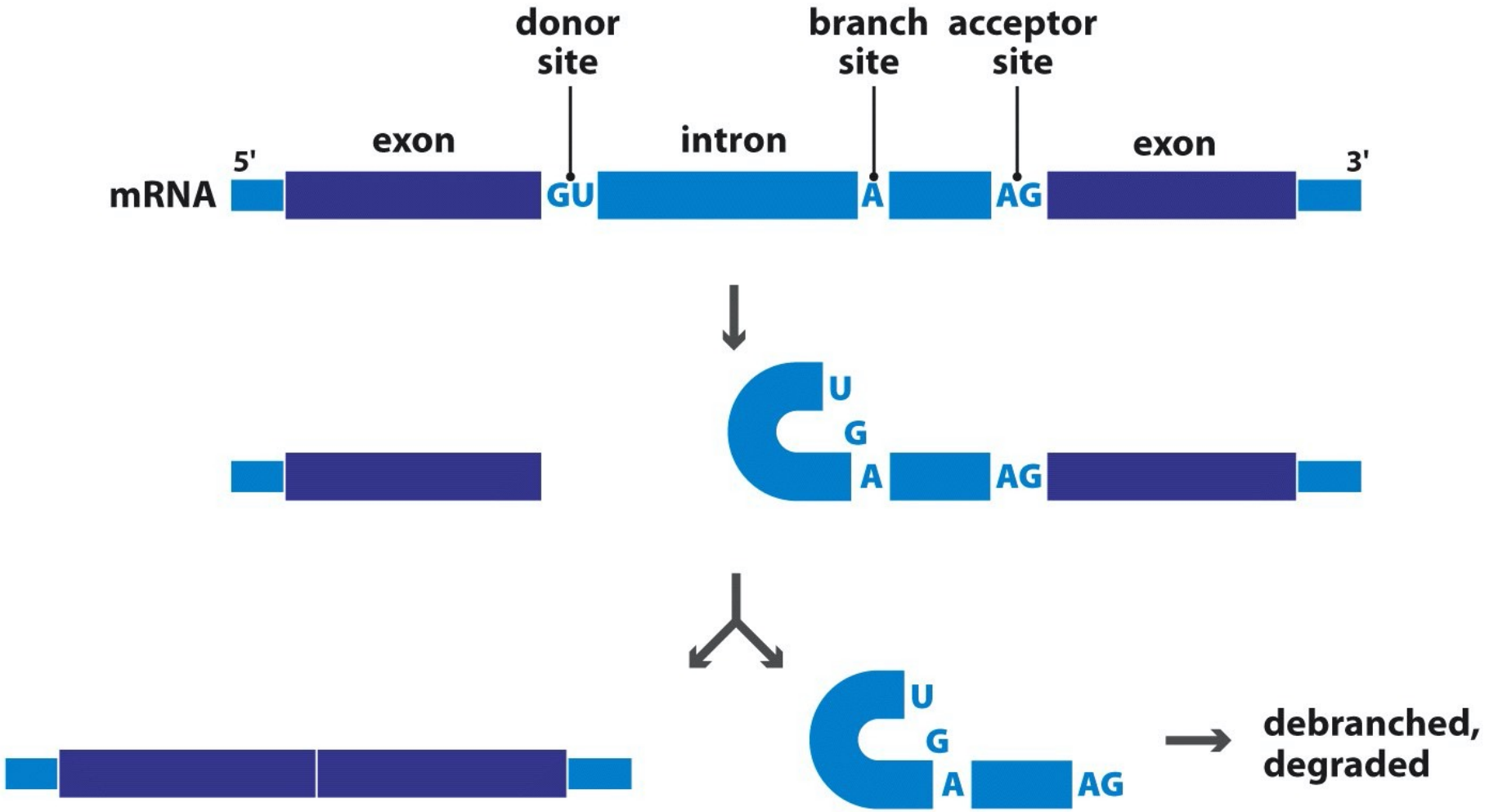
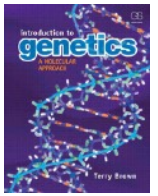


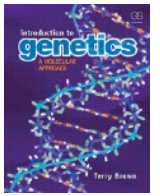
Figure 5.19 Introduction to Genetics (© Garland Science 2012)



• The splicing reaction is carried out by a set of RNA–protein complexes called **small nuclear ribonucleoproteins (snRNPs).**

• Each of these contains several proteins and one or two of the noncoding snRNAs.

• There are number of different snRNAs like U1, U2, U3, U4, U5, and U6 . The U stands for “uracil-rich.”



• Splicing process begins with U1–snRNP binding to a donor site.

• U2–snRNP then attaches to the branch site.

• U1– and U2–snRNPs move towards each other, bringing donor site close to the branch point.

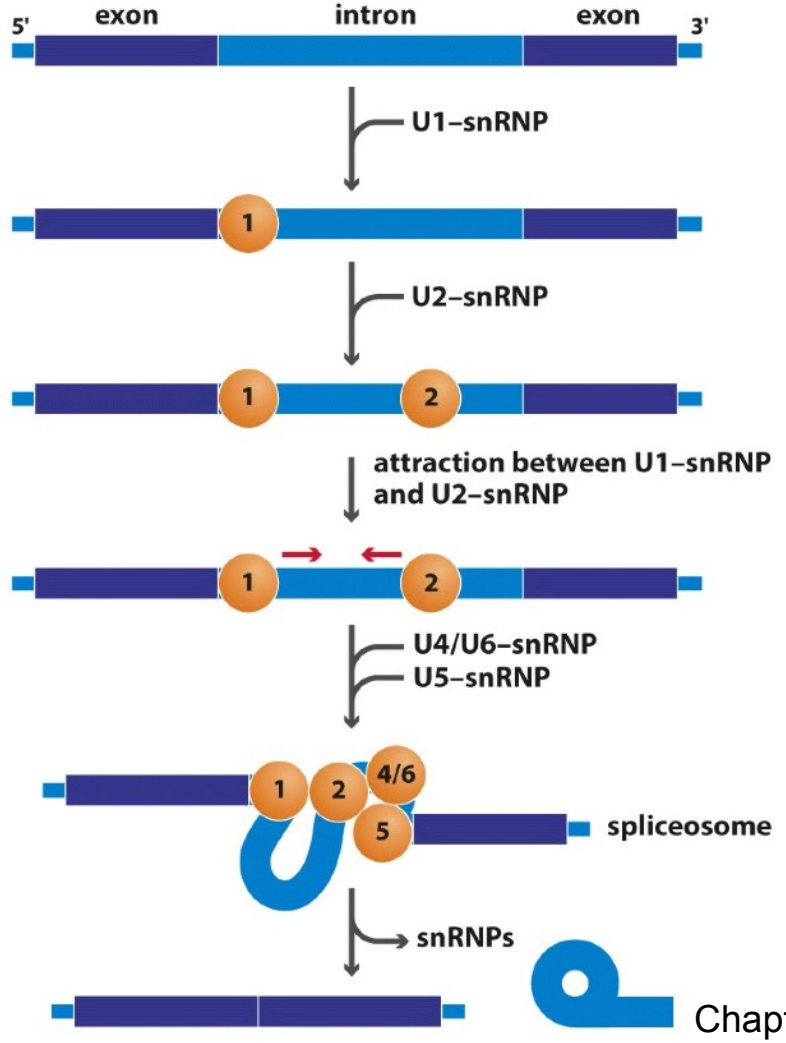
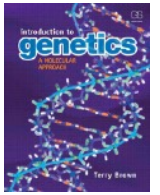


Figure 5.20 Introduction to Genetics (© Garland Science 2012)

Chapter5



• U5–snRNP and U4/U6–snRNP then attach, forming the spliceosome.

• The donor and acceptor sites now cleaves and two exons join together.

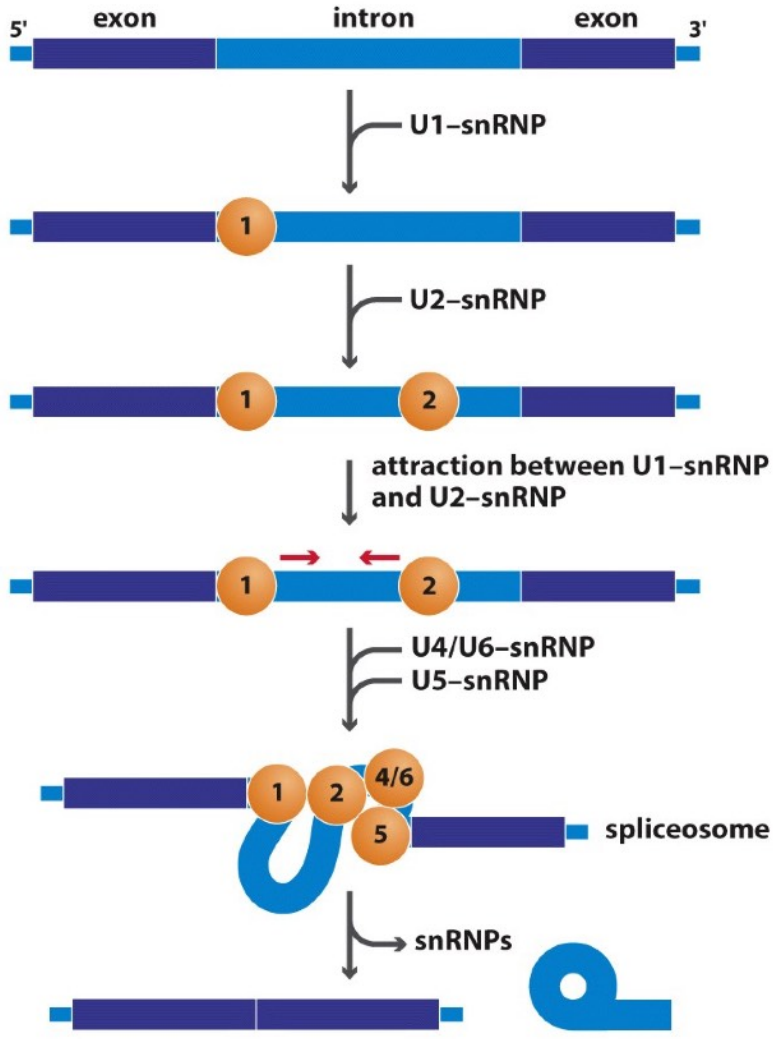
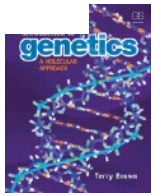
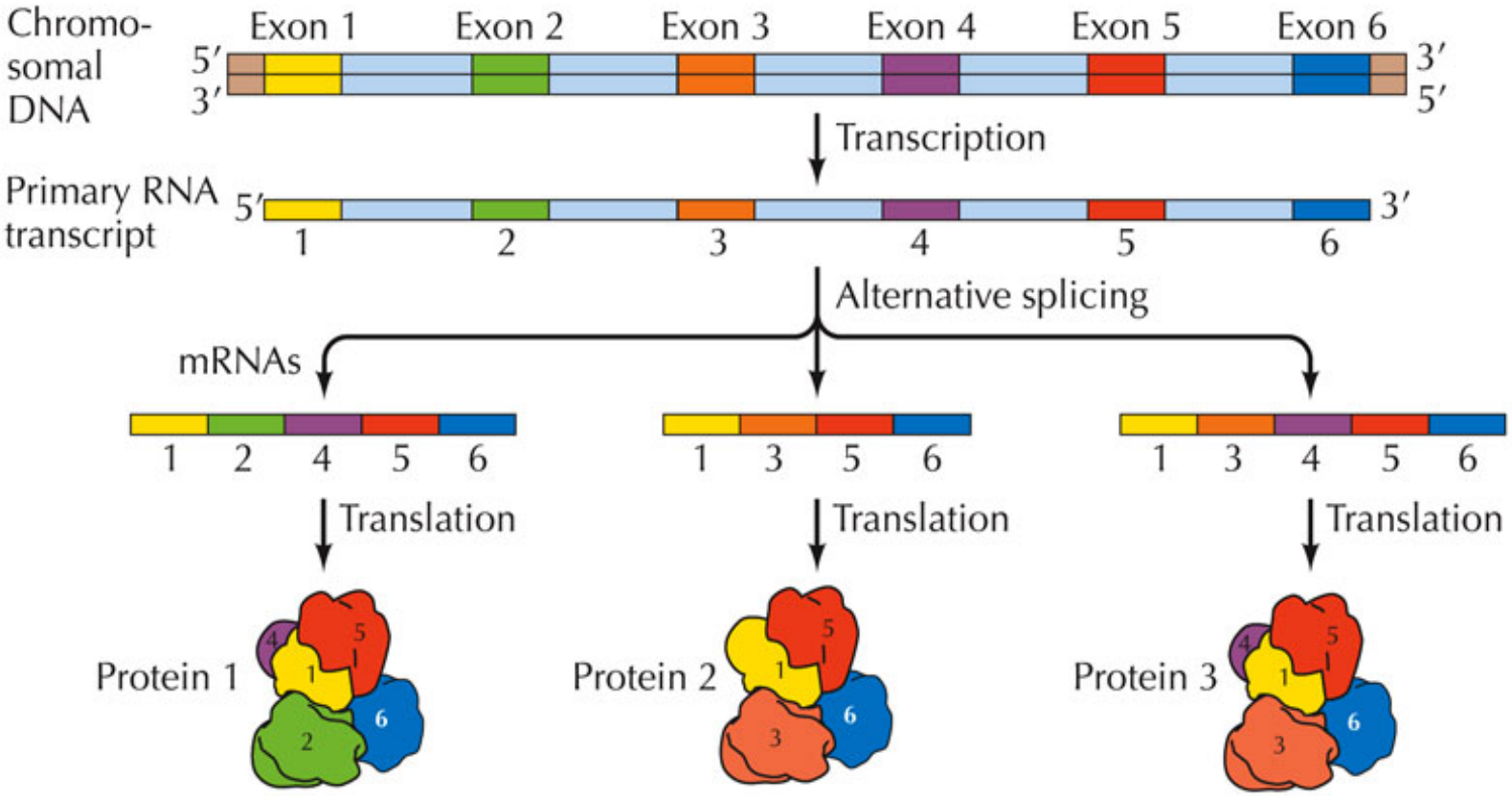


Figure 5.20 Introduction to Genetics (© Garland Science 2012)

ter5



Alternative splicing

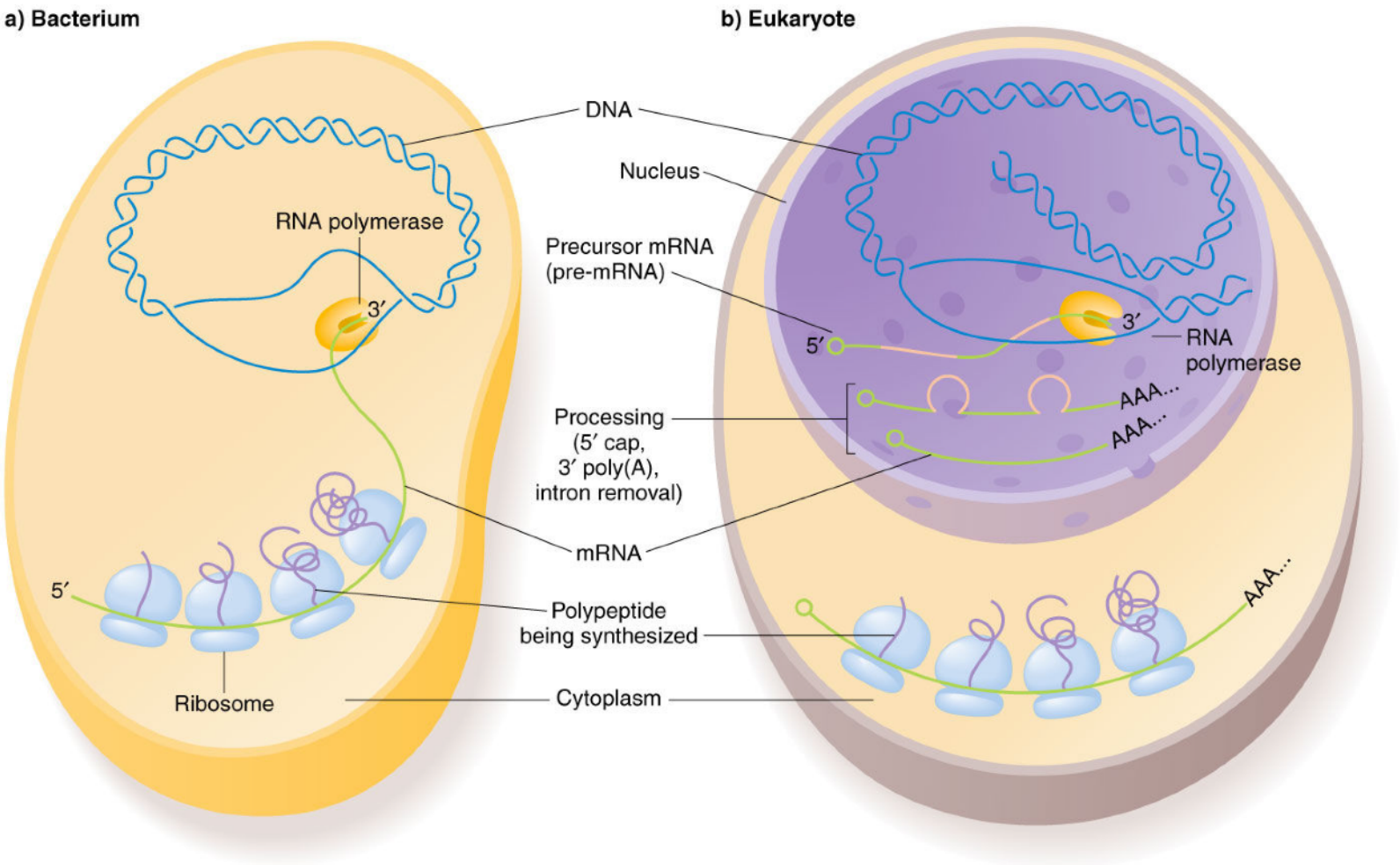


The splicing of introns allows for generating many different combination of exons and thus many different proteins.

Summary and review

Prokaryotes	Eukaryotes
No nucleus	**Nucleus**
DNA in cytoplasm	DNA in nucleus
Transcription in cytoplasm	**Transcription in nucleus**
Translation in cytoplasm	Translation in cytoplasm
Polcistronic transcripts (one transcript many genes)	**Monocistronic transcript (one transcript one gene)**
Coupled transcription and translation	Transcription and translation **NOT** coupled
mRNA not processed	**RNA processed**
One RNA polymerase	Many RNA polymerases

Summary and review



To know

polycistronic

Alternative splicing

5’ capping

TFIIB

snRNPs

3’UTR

5’UTR

TFIIA

PIC

RNA Pol I

U4

Spliceosome

AG

promoter

7-methy guanine

Pre-mRNA

Promoter proximal element

exon

RNA Pol III

intron

TFIID

Branch point sequence

Poly(A) tail

U1

U2

PAP

initiator

RNA Pol II

activator

U5

Pre-initiation complex

enhancer

TATA box

U6

monocistronic

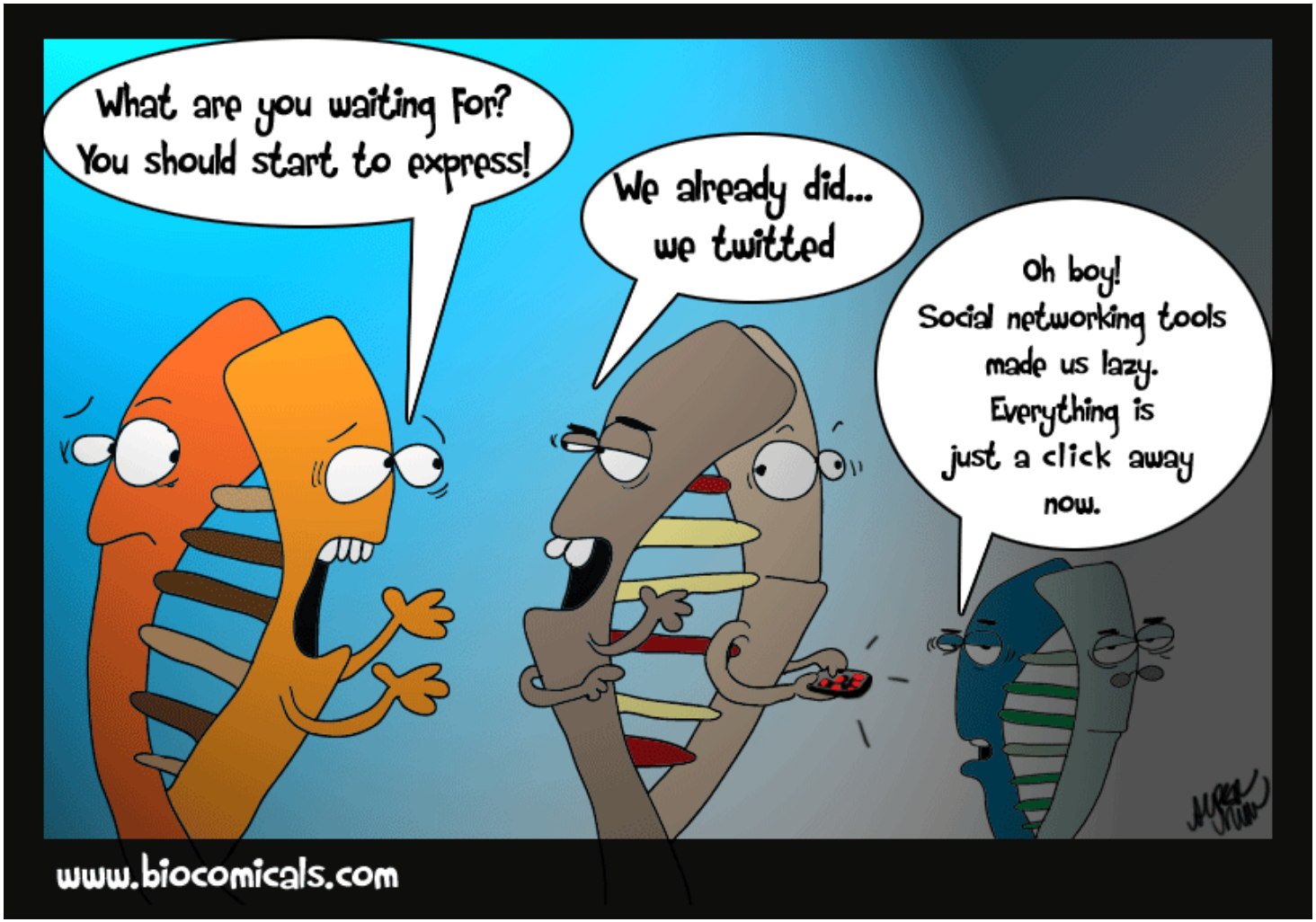
5’-5’ bond

GU

Expectations

- You know the transcription process in eukaryotes.
- You know what is the eukaryotic gene is composed of.
- You know and memorize the names of the molecular machinery involved in eukaryotic transcription.
- You know the modifications that take place to the mRNA in eukaryotes.
- You know the similarities and differences between the transcription process in prokaryotes to that of eukaryotes.

For a smile


What are you waiting for?
You should start to express!
We already did...
we twitted
Oh boy!
Social networking tools
made us lazy.
Everything is
just a click away
now.
www.biocomicals.com