MD637702 Medical Biochemistry and Molecular Biology
Academic Year 2024

Pre-mRNA splicing
&
RNA editing

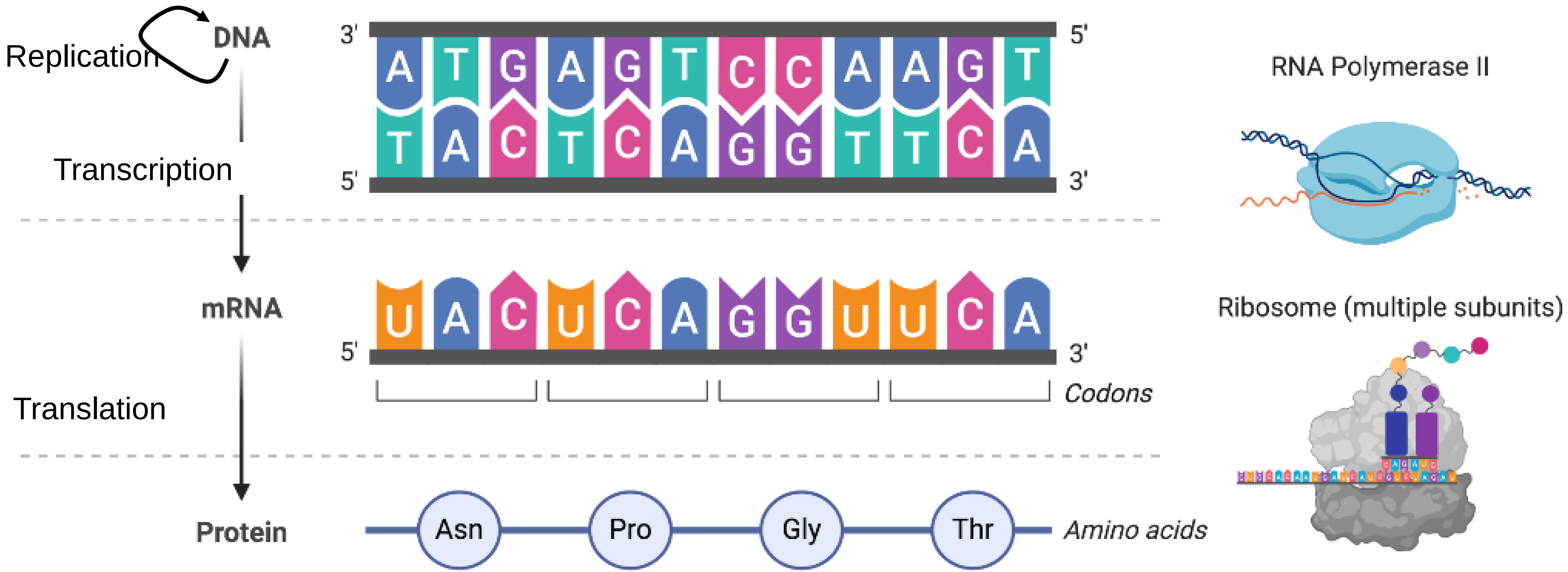
Raynoo Thanan
email: raynoo@kku.ac.th

Outline

- **Introduction**
  - Central dogma
  - Gene structures & Human genes
  - Types of RNA
  - RNA biosynthesis (Transcription)
- **Post-transcriptional modification**
  - Key RNA processing events that occurred on pre-mRNA to product mature RNA
  - Pre-mRNA splicing (key molecules & important process)
  - The importance of RNA splicing
  - RNA editing (key molecules & important process)
  - The importance of RNA editing

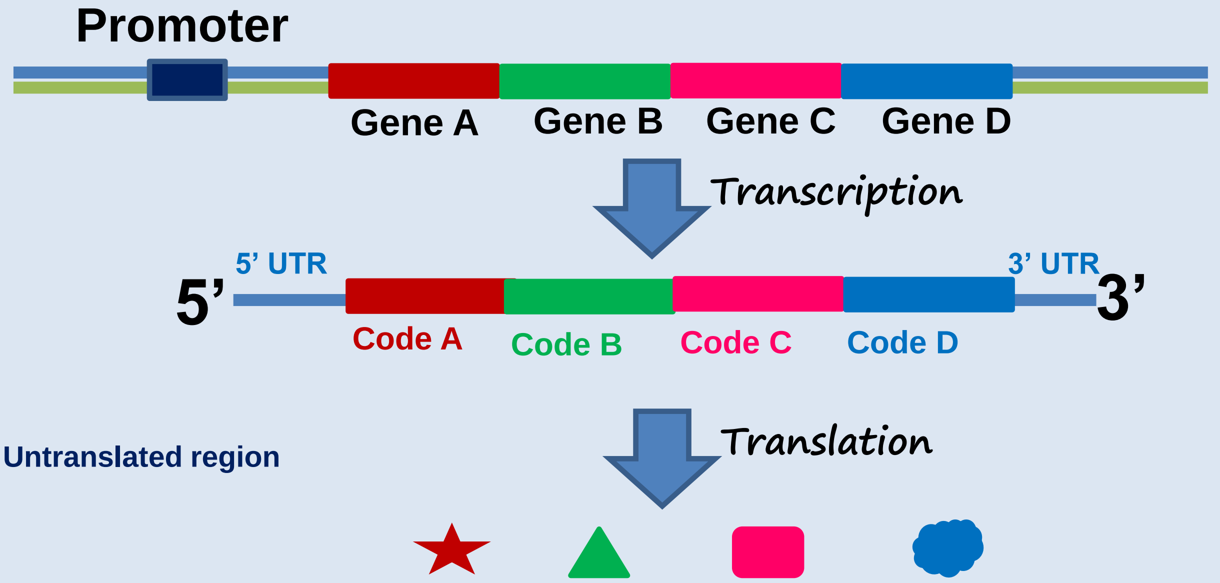
How cell read the genome: From DNA to Protein

Central dogma is an organizing principle of molecular biology : genetic information flows between nucleic acids and from nucleic acid to protein



https://elizabethmcd.github.io/

Gene structures



Promoter

Gene A Gene B Gene C Gene D

Prokaryotic gene (Operon)

Transcription

5' UTR 3' UTR

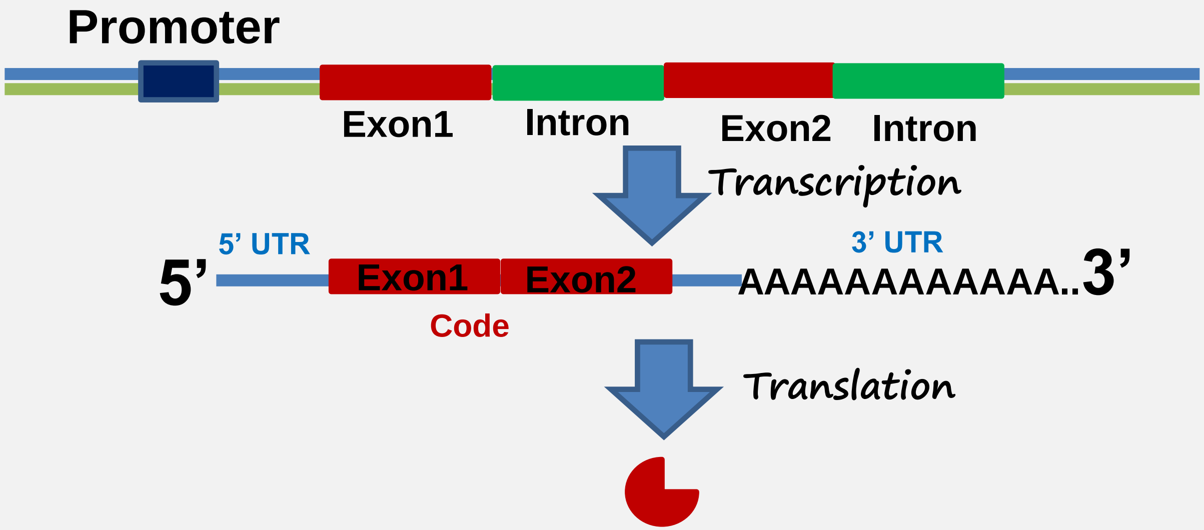
5' Code A Code B Code C Code D 3'

Polycistronic mRNA

*UTR= Untranslated region

Translation

Proteins



Promoter

Exon1 Intron Exon2 Intron

Eukaryotic gene

Transcription

5' UTR 3' UTR

5' Exon1 Exon2 AAAAAAAAAAA..3'

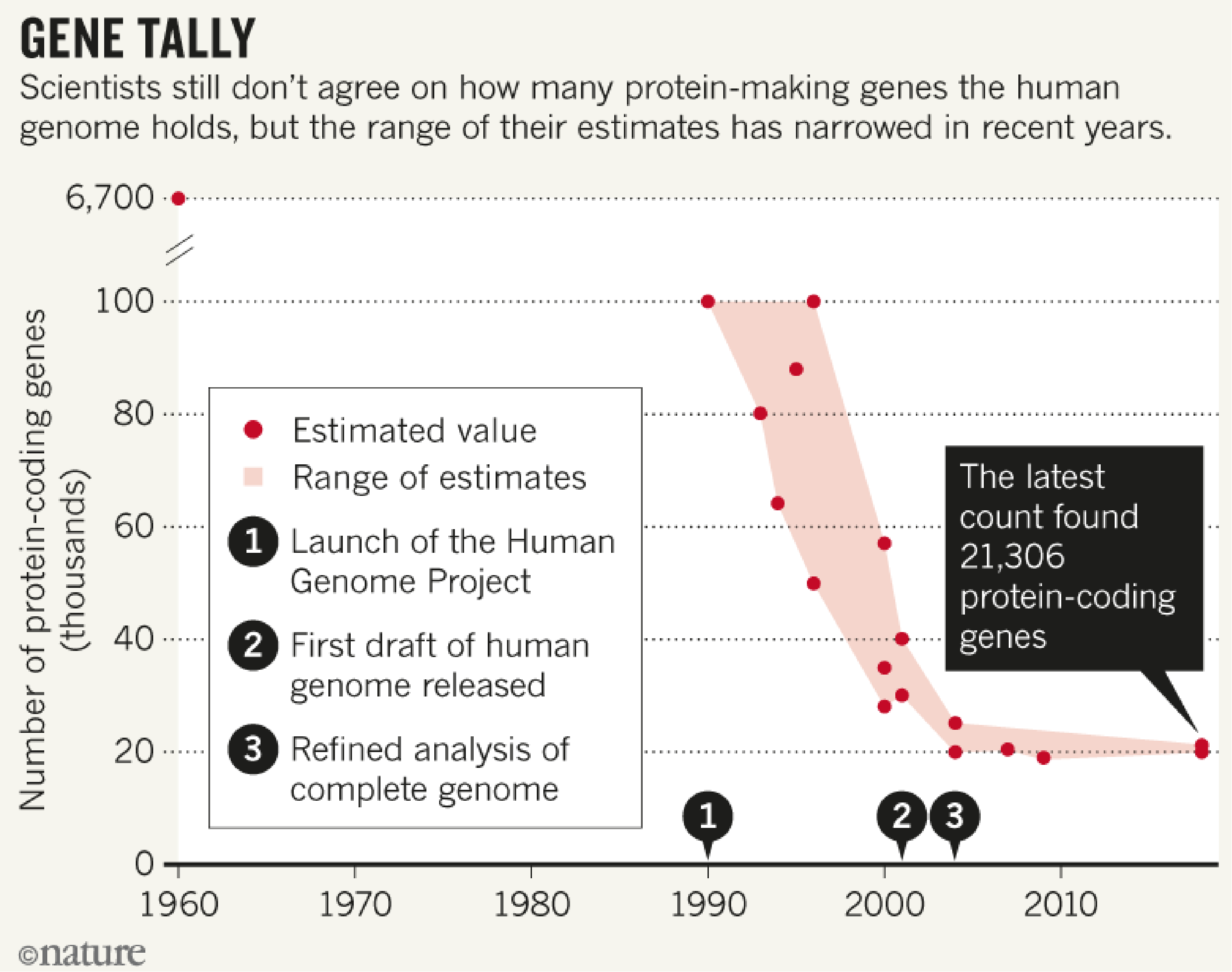
Code

Monocistronic mRNA

Translation

Protein

Human genes



GENE TALLY

Scientists still don't agree on how many protein-making genes the human genome holds, but the range of their estimates has narrowed in recent years.

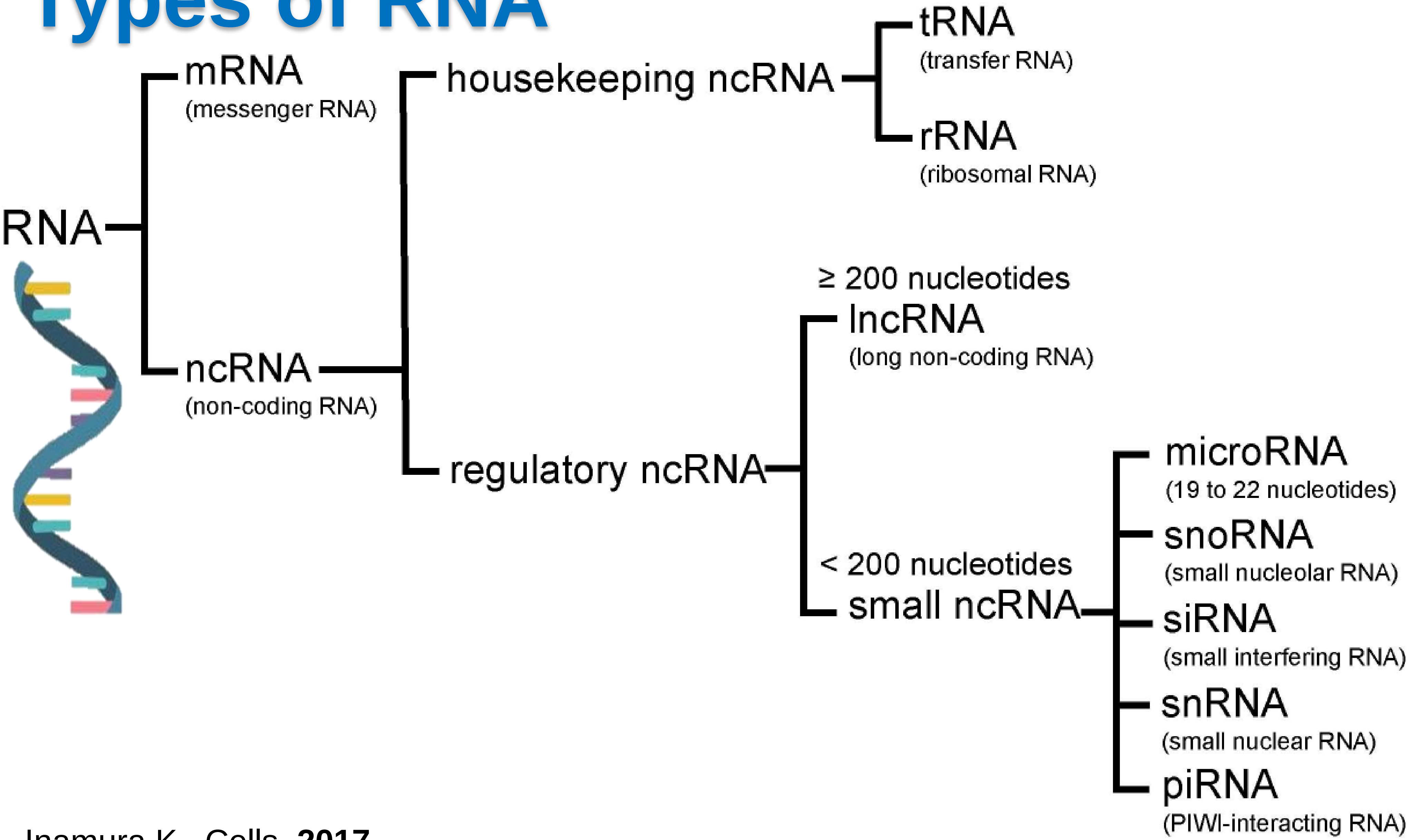
The latest count found 21,306 protein-coding genes

~20,000 different genes were identified in human genome.

>100,000 different proteins were identified in human cells.

How is this possible?

Types of RNA



- RNA
  - mRNA (messenger RNA)
  - ncRNA (non-coding RNA)
    - housekeeping ncRNA
      - tRNA (transfer RNA)
      - rRNA (ribosomal RNA)
    - regulatory ncRNA
      - ≥ 200 nucleotides lncRNA (long non-coding RNA)
      - < 200 nucleotides small ncRNA
        - microRNA (19 to 22 nucleotides)
        - snoRNA (small nucleolar RNA)
        - siRNA (small interfering RNA)
        - snRNA (small nuclear RNA)
        - piRNA (PIWI-interacting RNA)

Inamura K., Cells, **2017**

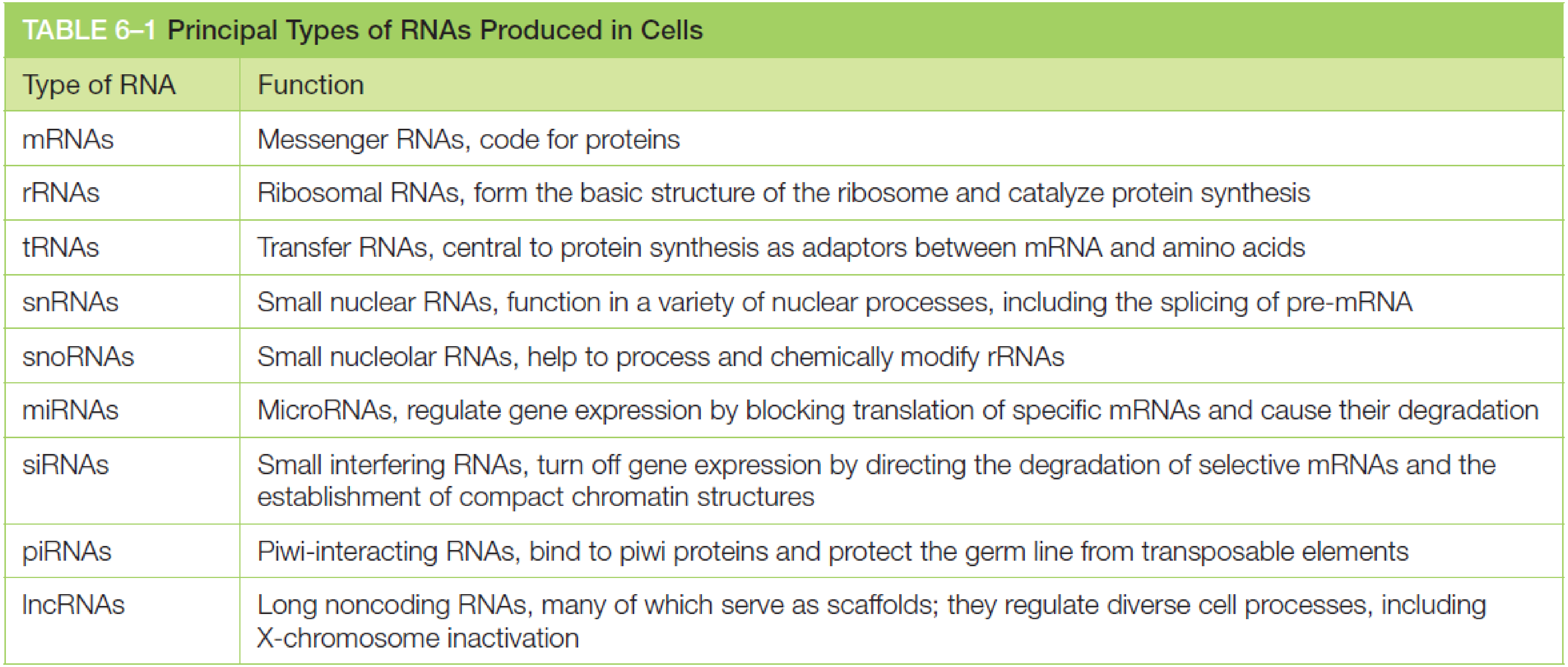


TABLE 6–1 Principal Types of RNAs Produced in Cells

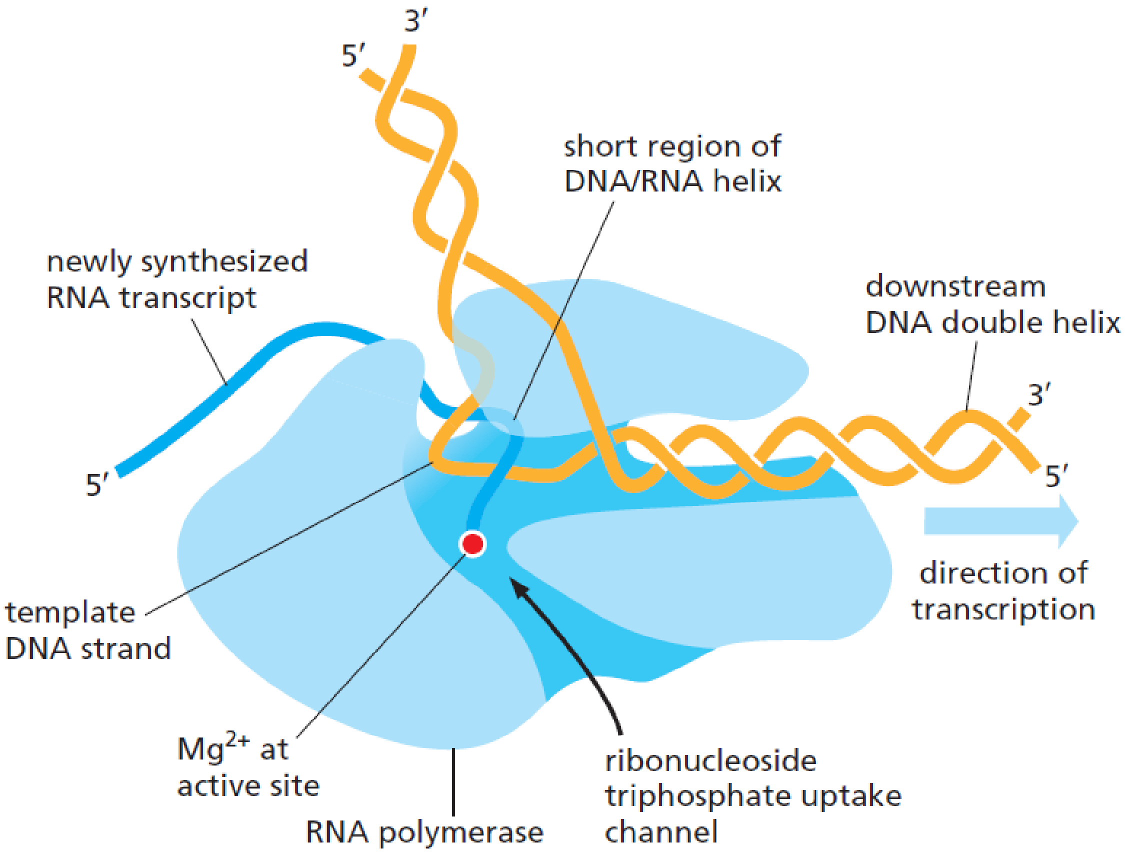
Type of RNA	Function
mRNAs	Messenger RNAs, code for proteins
rRNAs	Ribosomal RNAs, form the basic structure of the ribosome and catalyze protein synthesis
tRNAs	Transfer RNAs, central to protein synthesis as adaptors between mRNA and amino acids
snRNAs	Small nuclear RNAs, function in a variety of nuclear processes, including the splicing of pre-mRNA
snoRNAs	Small nucleolar RNAs, help to process and chemically modify rRNAs
miRNAs	MicroRNAs, regulate gene expression by blocking translation of specific mRNAs and cause their degradation
siRNAs	Small interfering RNAs, turn off gene expression by directing the degradation of selective mRNAs and the establishment of compact chromatin structures
piRNAs	Piwi-interacting RNAs, bind to piwi proteins and protect the germ line from transposable elements
lncRNAs	Long noncoding RNAs, many of which serve as scaffolds; they regulate diverse cell processes, including X-chromosome inactivation

RNA synthesis (Transcription)

Transcription: The process of transcribing DNA sequence information into RNA sequence information.

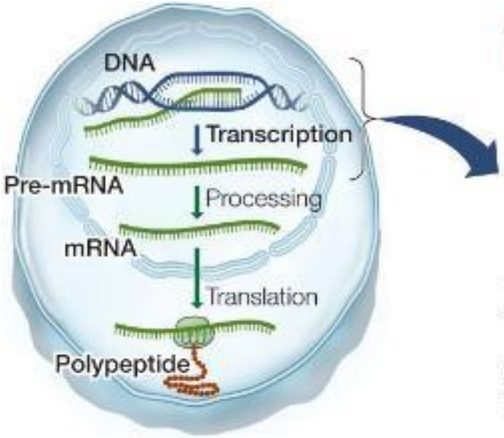
DNA and RNA use the complementary language, and the information is simply transcribed, or copied, from one molecule to the other

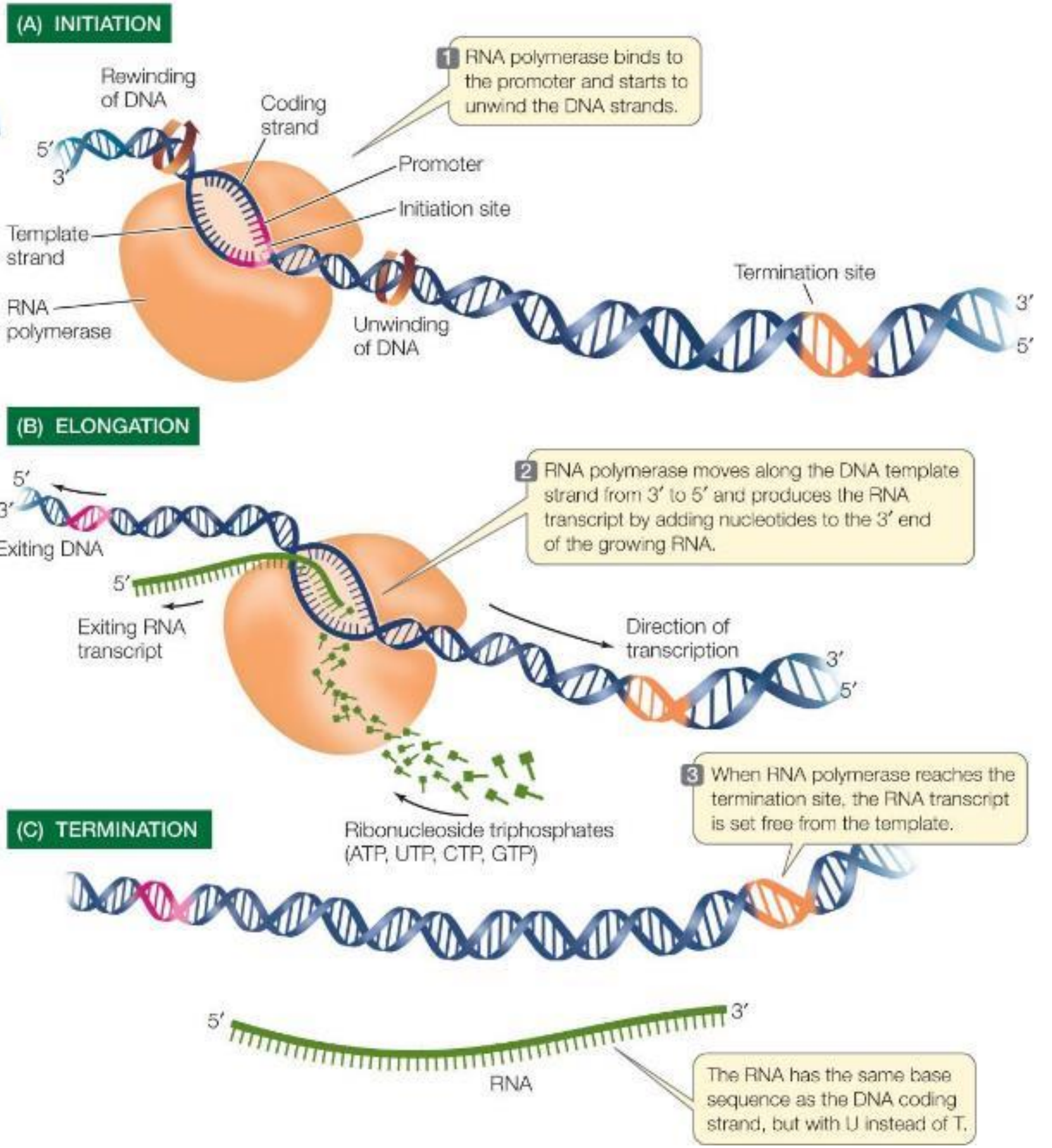
DNA is transcribed by the enzyme RNA polymerase.



-The RNA polymerase (pale blue) moves stepwise along the DNA, unwinding the DNA helix at its active site indicated by the Mg^{2+} (red), which is required for catalysis.

-The polymerase adds nucleotides one by one to the RNA chain at the polymerization site, using an exposed DNA strand as a template.





1 RNA polymerase binds to the promoter and starts to unwind the DNA strands.

2 RNA polymerase moves along the DNA template strand from 3′ to 5′ and produces the RNA transcript by adding nucleotides to the 3′ end of the growing RNA.

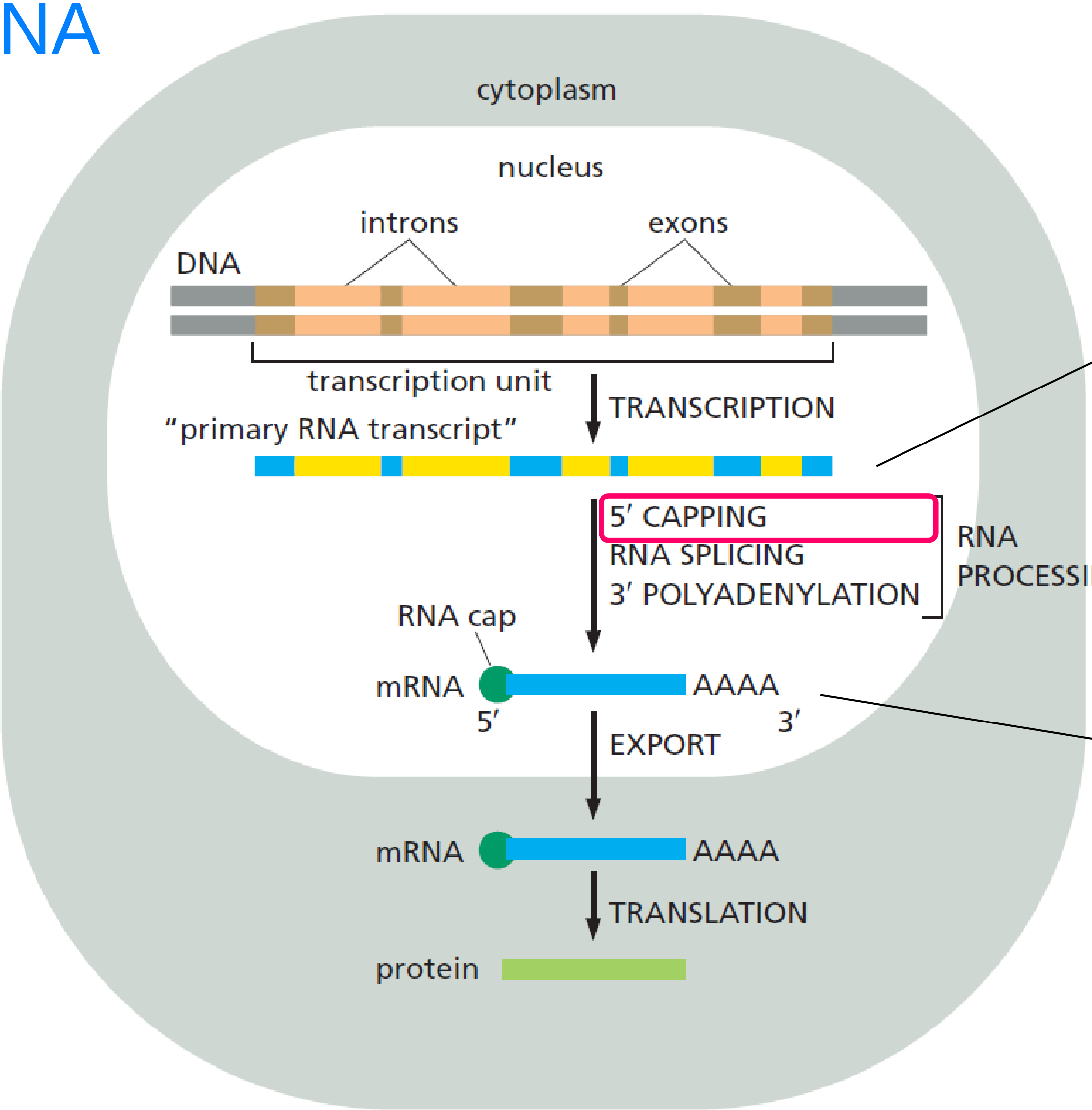
3 When RNA polymerase reaches the termination site, the RNA transcript is set free from the template.

The RNA has the same base sequence as the DNA coding strand, but with U instead of T.

Transcription steps

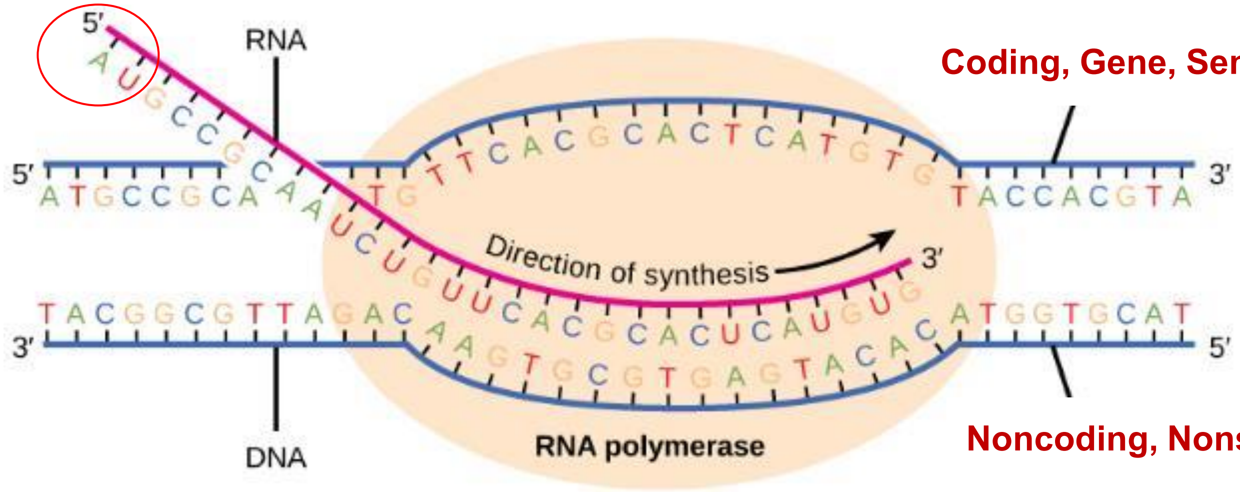
microbenotes.com

Transcription Elongation in Eukaryotes Is Tightly Coupled to RNA Processing



Primary mRNA, Pre-mRNA

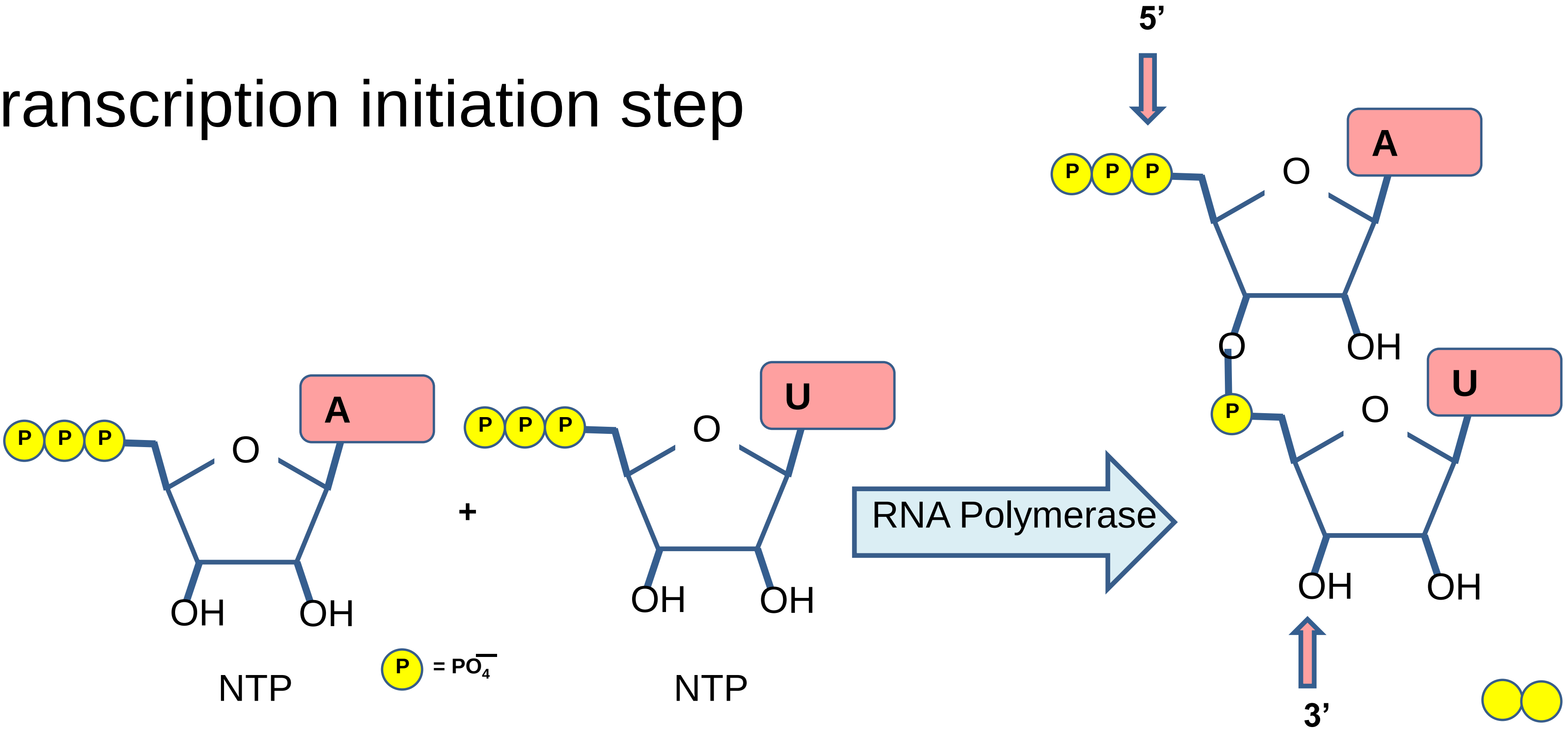
Mature RNA



Coding, Gene, Sense, Non-template strand

Noncoding, Nonsense, Template strand

Transcription initiation step



5’ RNA Capping

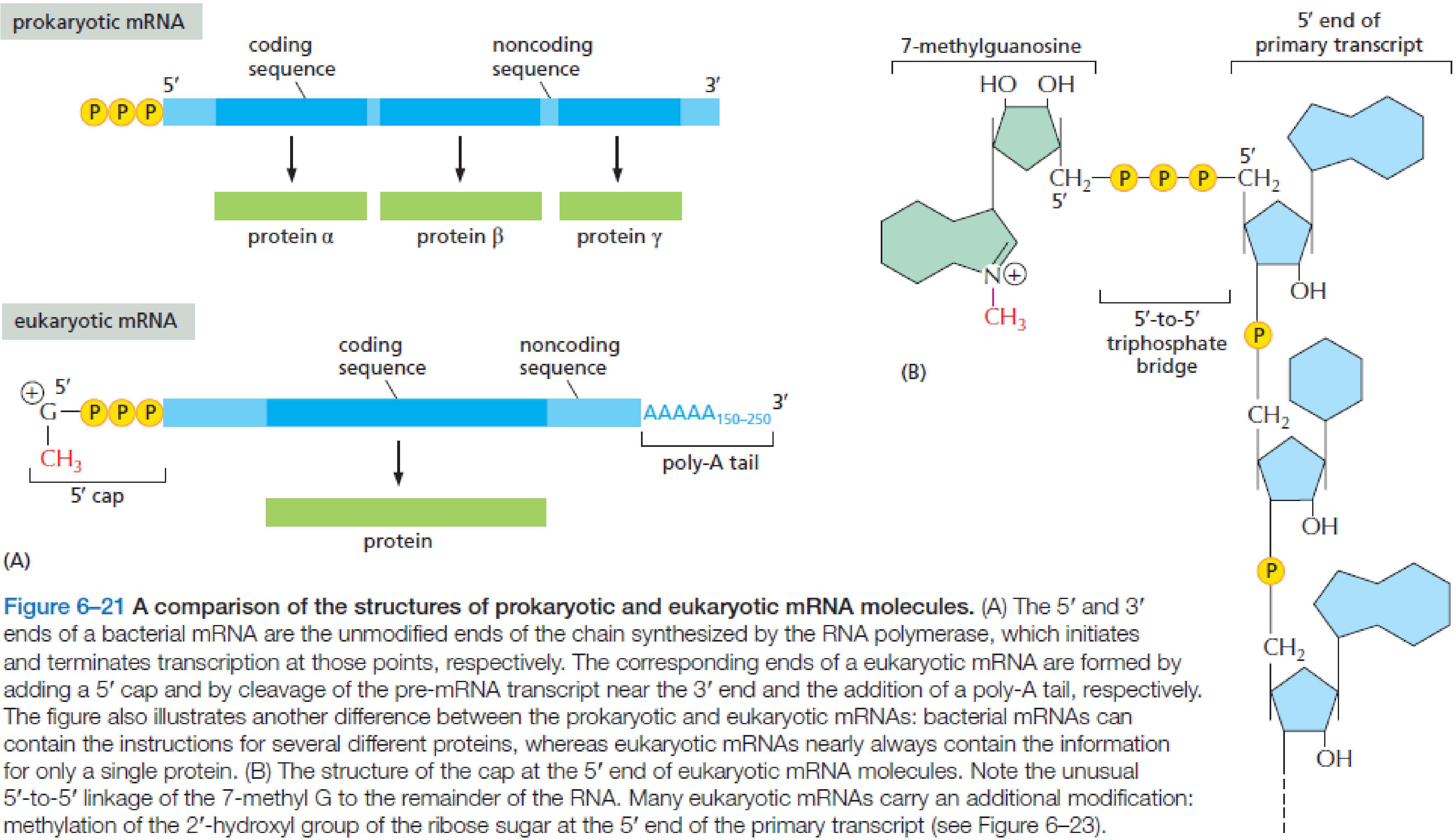
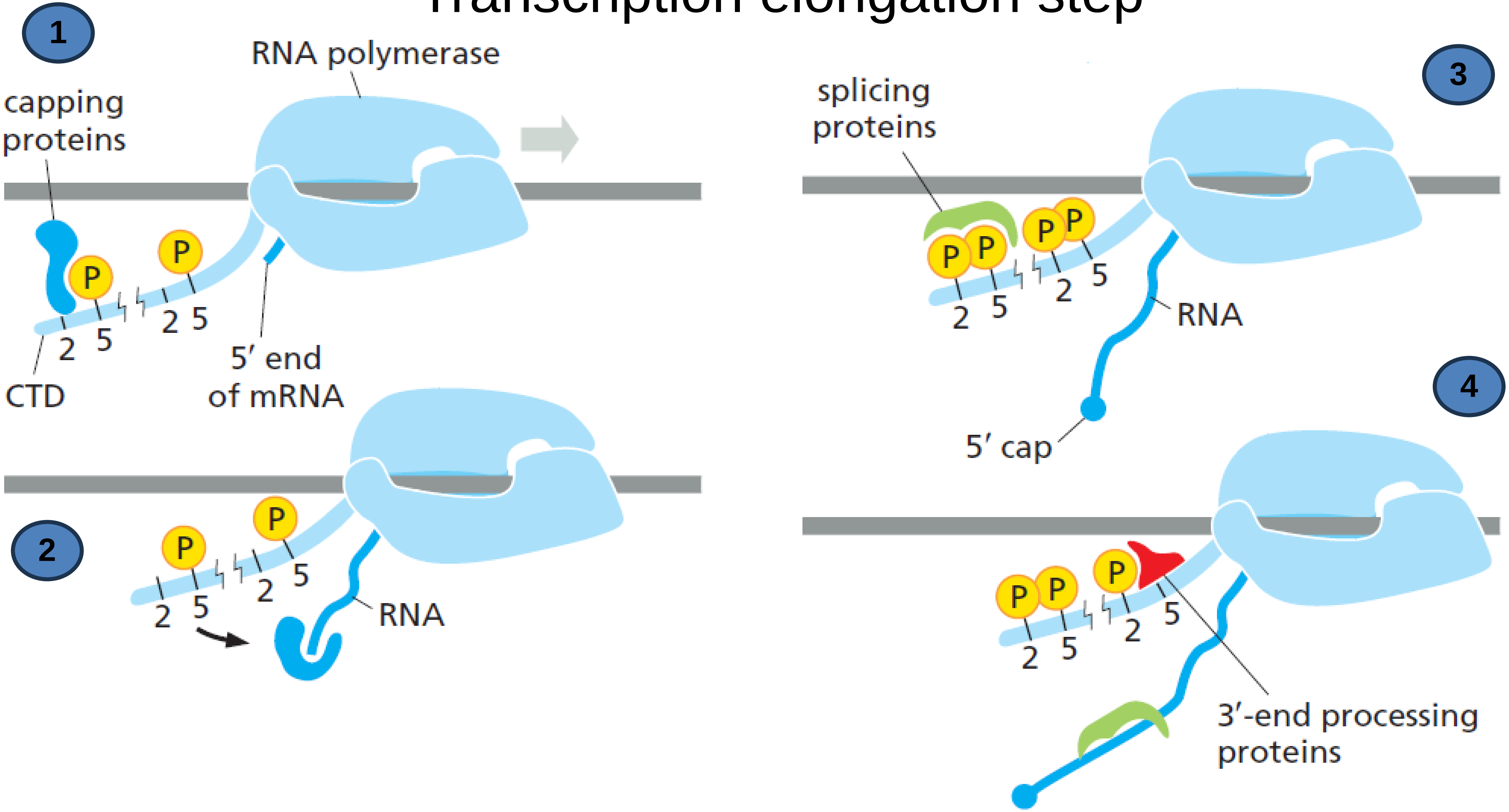


Figure 6–21 A comparison of the structures of prokaryotic and eukaryotic mRNA molecules. (A) The 5′ and 3′ ends of a bacterial mRNA are the unmodified ends of the chain synthesized by the RNA polymerase, which initiates and terminates transcription at those points, respectively. The corresponding ends of a eukaryotic mRNA are formed by adding a 5′ cap and by cleavage of the pre-mRNA transcript near the 3′ end and the addition of a poly-A tail, respectively. The figure also illustrates another difference between the prokaryotic and eukaryotic mRNAs: bacterial mRNAs can contain the instructions for several different proteins, whereas eukaryotic mRNAs nearly always contain the information for only a single protein. (B) The structure of the cap at the 5′ end of eukaryotic mRNA molecules. Note the unusual 5′-to-5′ linkage of the 7-methyl G to the remainder of the RNA. Many eukaryotic mRNAs carry an additional modification: methylation of the 2′-hydroxyl group of the ribose sugar at the 5′ end of the primary transcript (see Figure 6–23).

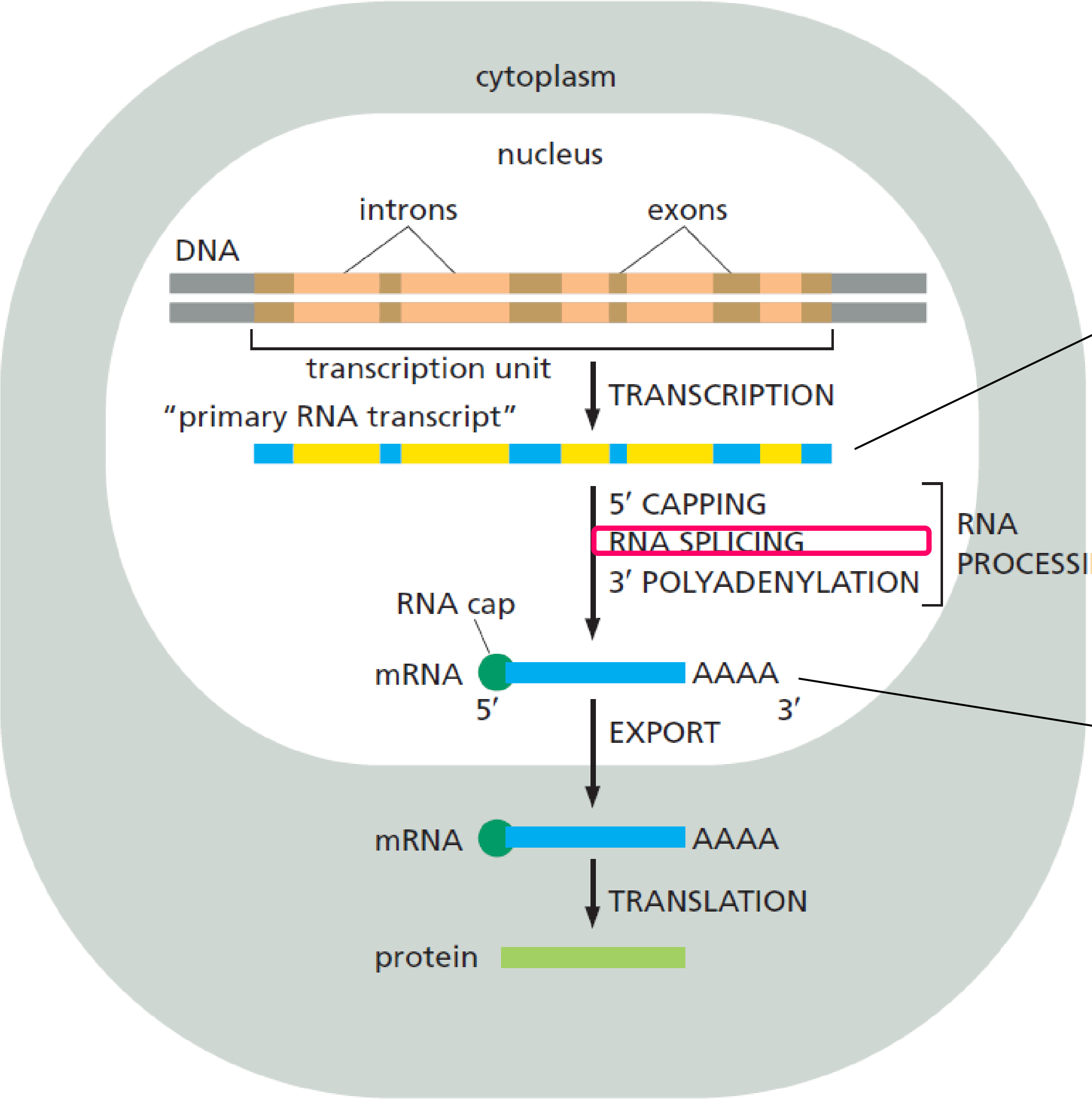
RNA Capping Is the First Modification of Eukaryotic Pre-mRNAs

Transcription elongation step



As the RNA polymerase transcribes DNA into RNA, it carries RNA-processing proteins on its tail that are transferred to the nascent RNA at the appropriate time.

Transcription Elongation in Eukaryotes Is Tightly Coupled to RNA Processing

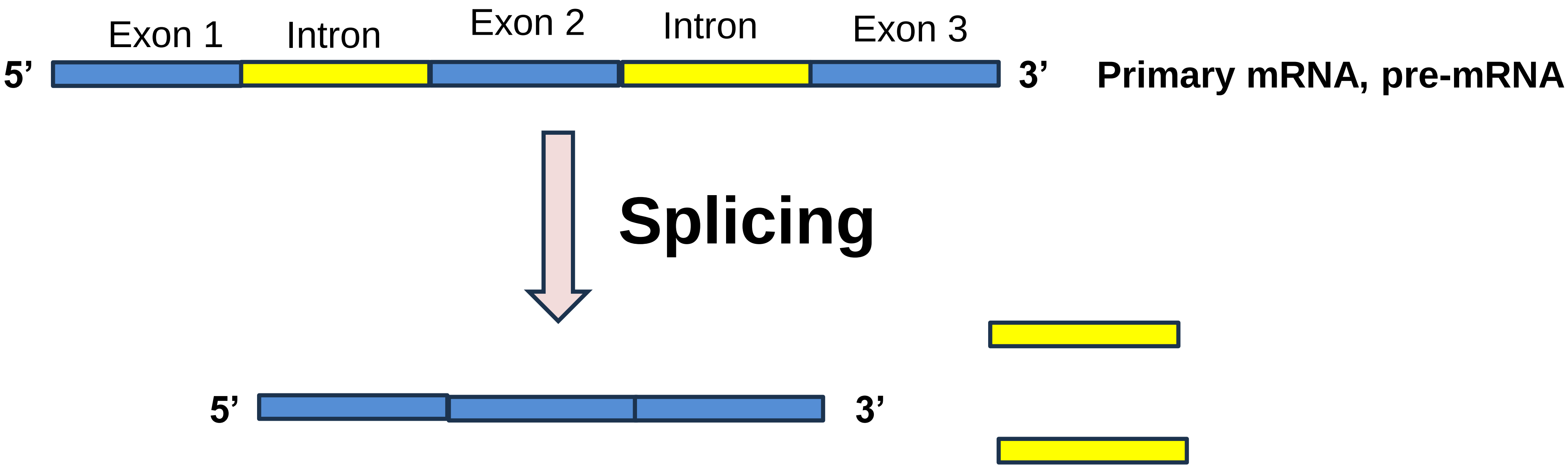


Primary mRNA, Pre-mRNA

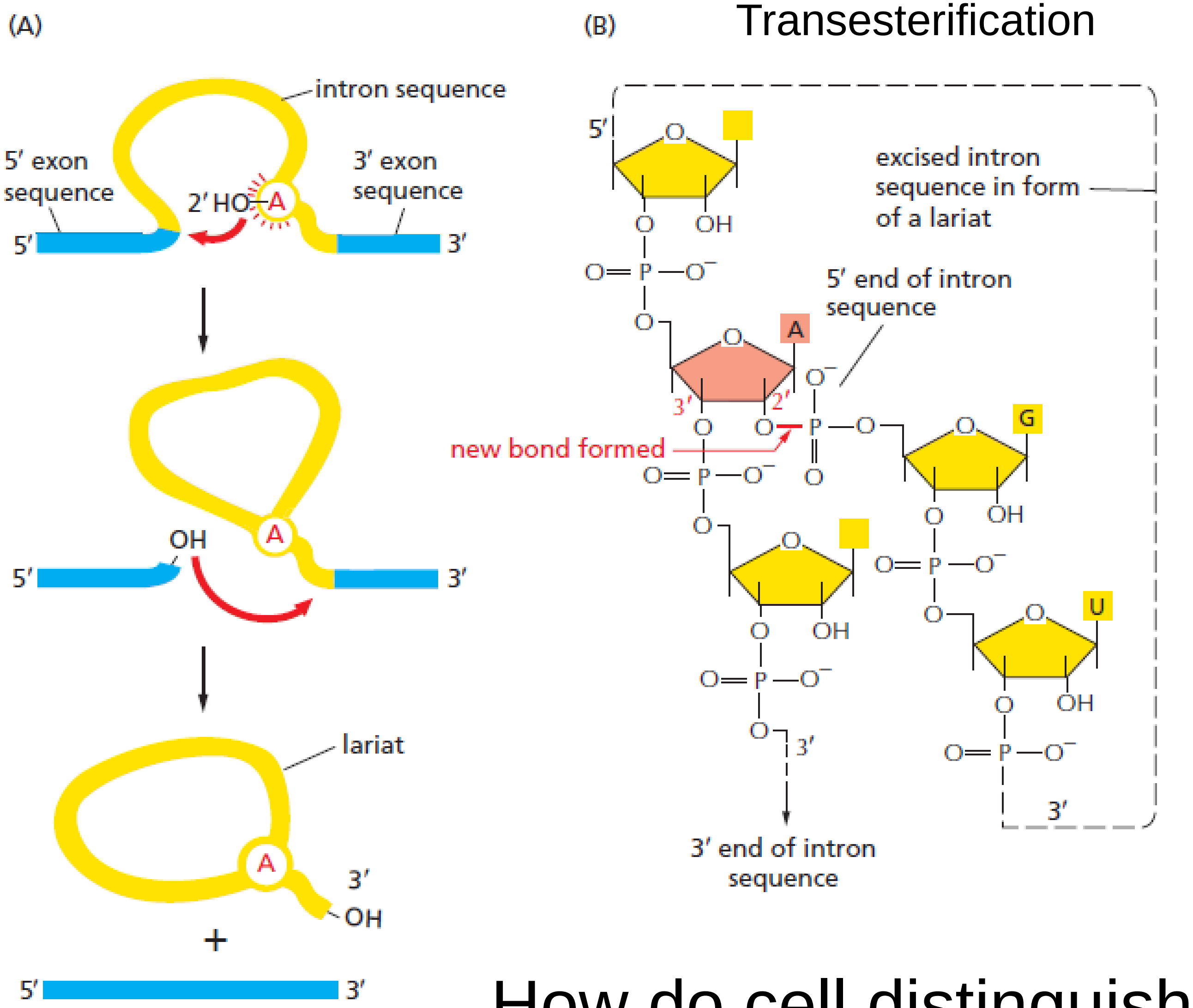
Mature RNA

Pre-mRNA splicing

Removal of introns and joining of exons in a primary transcript.
Also called simply splicing.



Pre-mRNA splicing

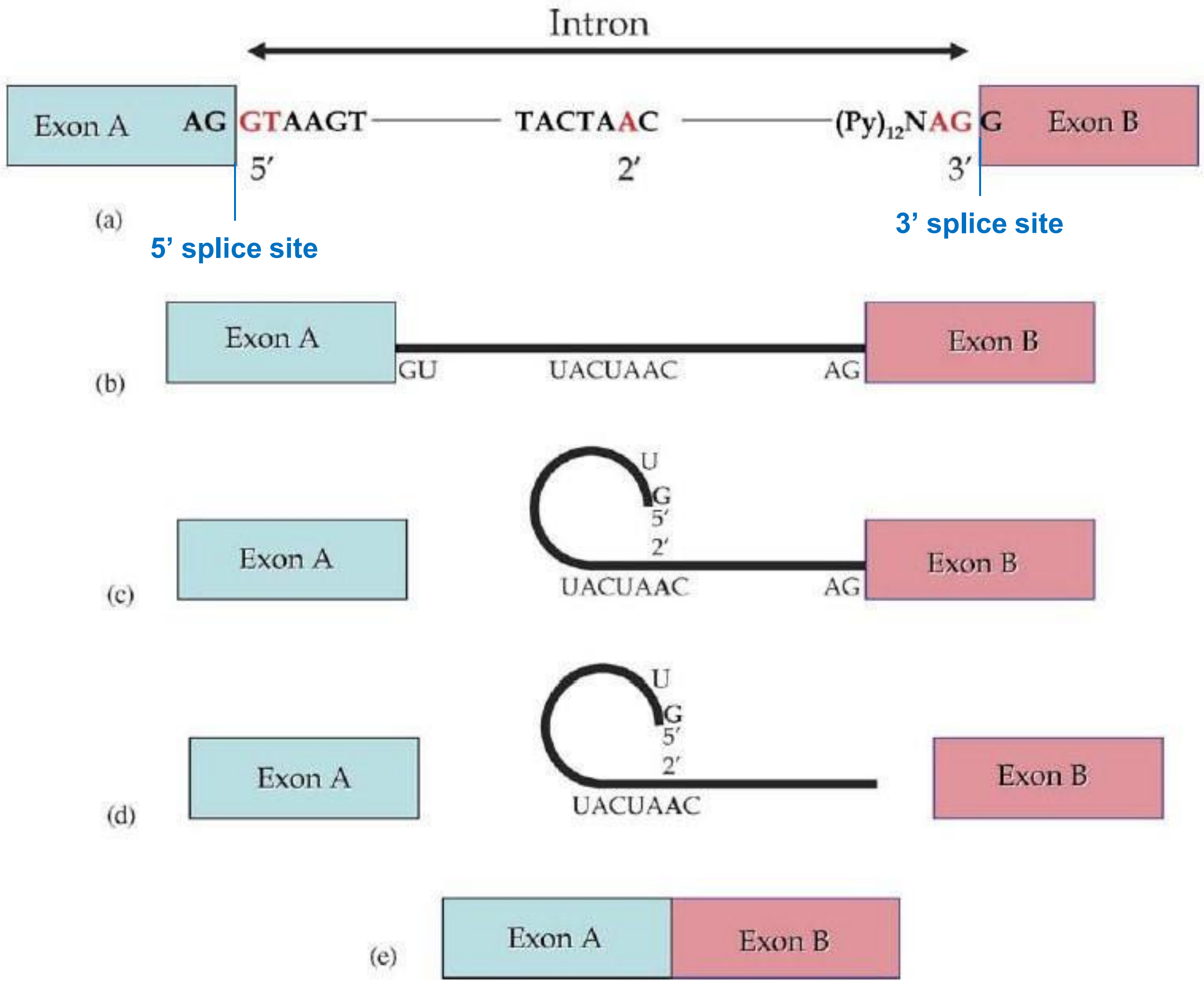


Transesterification

Figure 6–25 The pre-mRNA splicing reaction. (A) In the first step, a specific adenine nucleotide in the intron sequence (indicated in *red*) attacks the 5′ splice site and cuts the sugar-phosphate backbone of the RNA at this point. The cut 5′ end of the intron becomes covalently linked to the adenine nucleotide, as shown in detail in (B), thereby creating a loop in the RNA molecule. The released free 3′-OH end of the exon sequence then reacts with the start of the next exon sequence, joining the two exons together and releasing the intron sequence in the shape of a *lariat*. The two exon sequences thereby become joined into a continuous coding sequence. The released intron sequence is eventually broken down into single nucleotides, which are recycled.

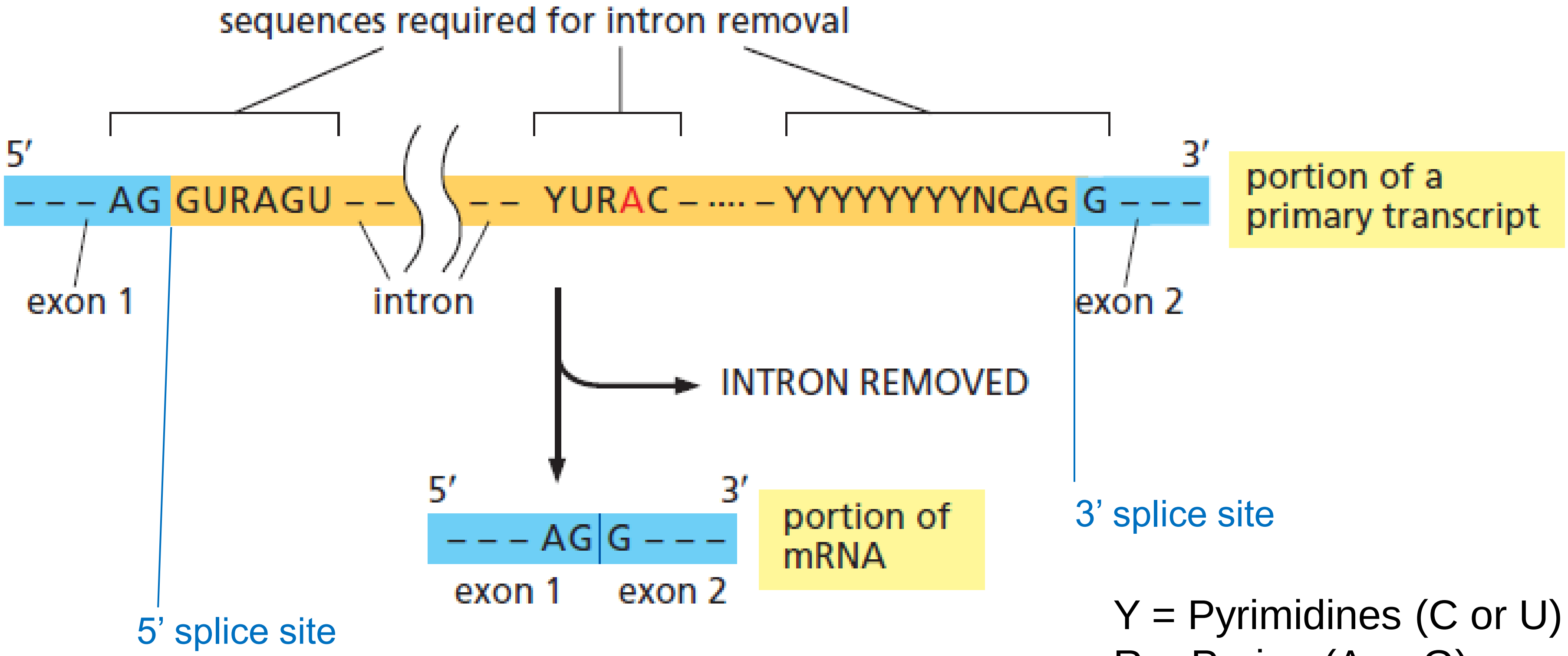
How do cell distinguish exon and intron?

Pre-mRNA splicing

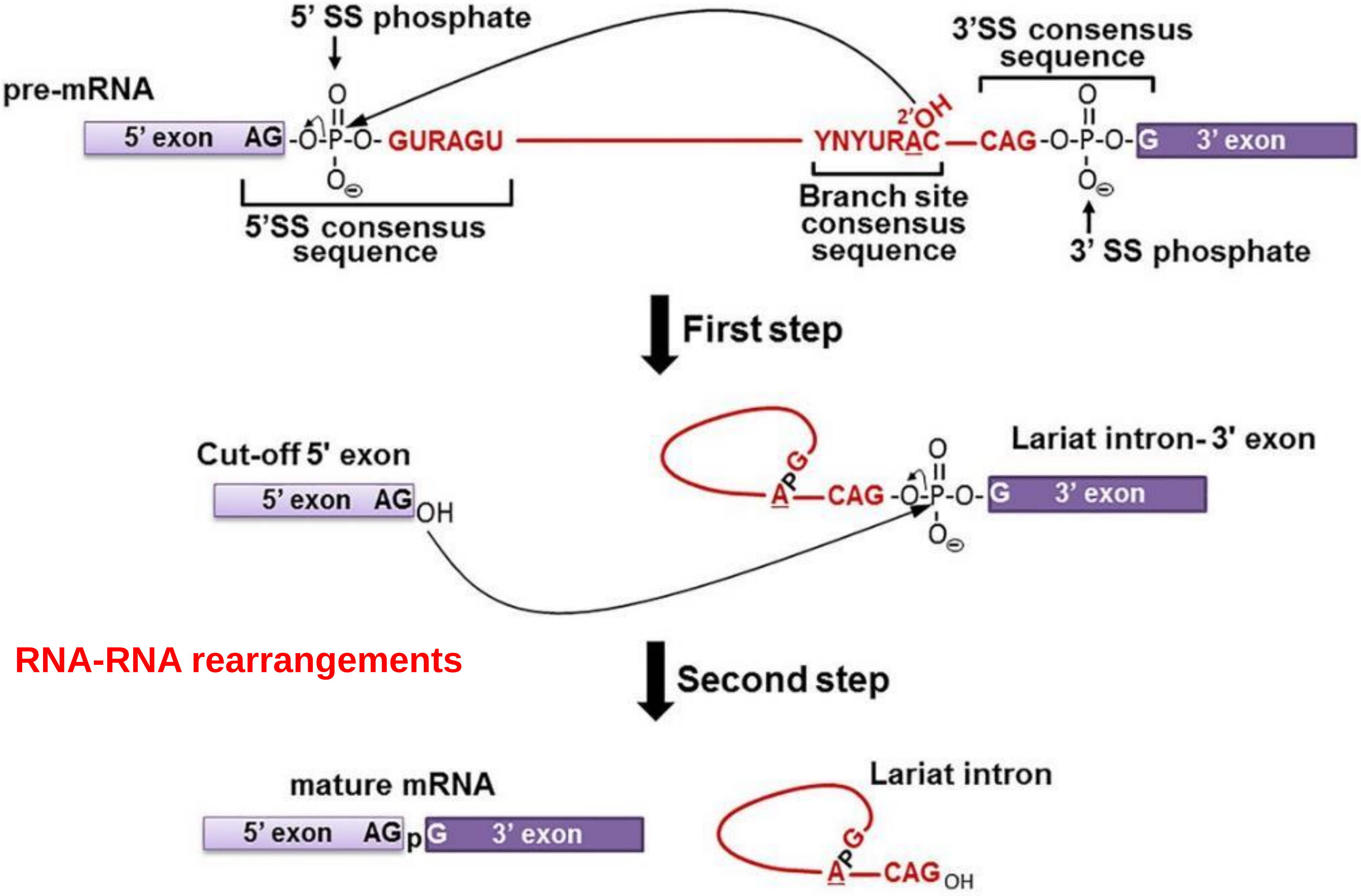


Nuclear splicing. (a) Schematic representation of an intron, flanked by two exons (A, light blue) and (B, pink). The essential splicing signals that define the exon boundaries are relatively short and poorly-conserved sequences. Only the GT (at the 5' end of the intron), AG (at the 3' end) and the branchpoint adenosine at the 2' position are always conserved (all shown in red). (b) Pre-mRNA, showing the two flanking exons, branchpoint sequence and the conserved nucleotides as in (a). (c) The spliceosome cuts the pre-mRNA at its 5' end, and then the intron forms a lariat by joining G at the 5' end with the branchpoint A at the 2' position. (d) The spliceosome cuts at the 3' end of the intron and the intron is released as a lariat. (e) Exons A and B are then joined and released from the spliceosome.

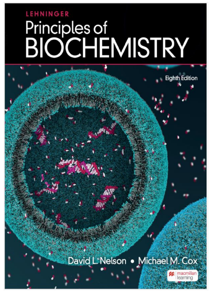
The consensus nucleotide sequences in an RNA molecule that signal the beginning and the end of most introns in humans.



Y = Pyrimidines (C or U)
R = Purine (A or G)
N = Pyrimidines or Purines



RNA-RNA rearrangements



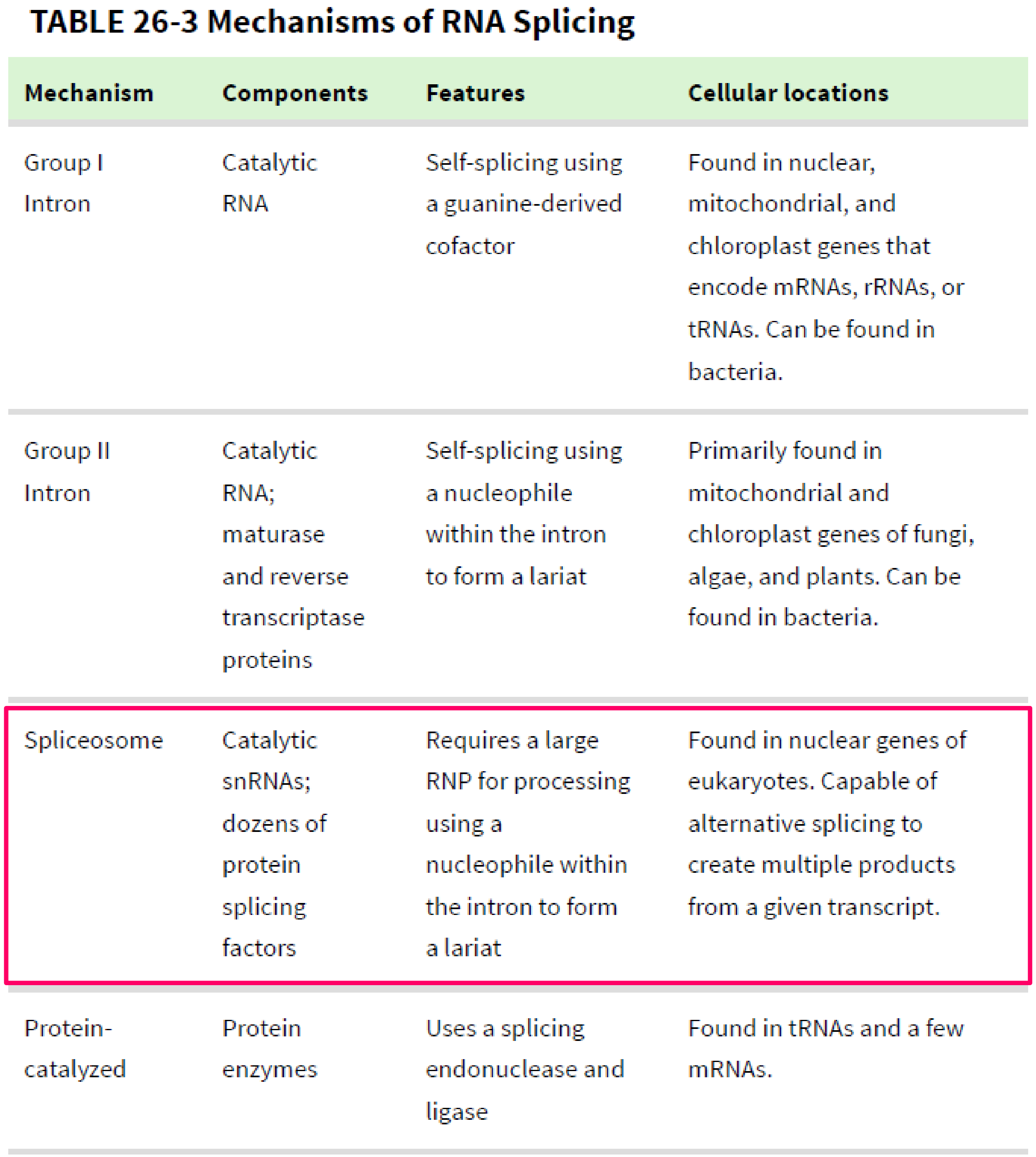
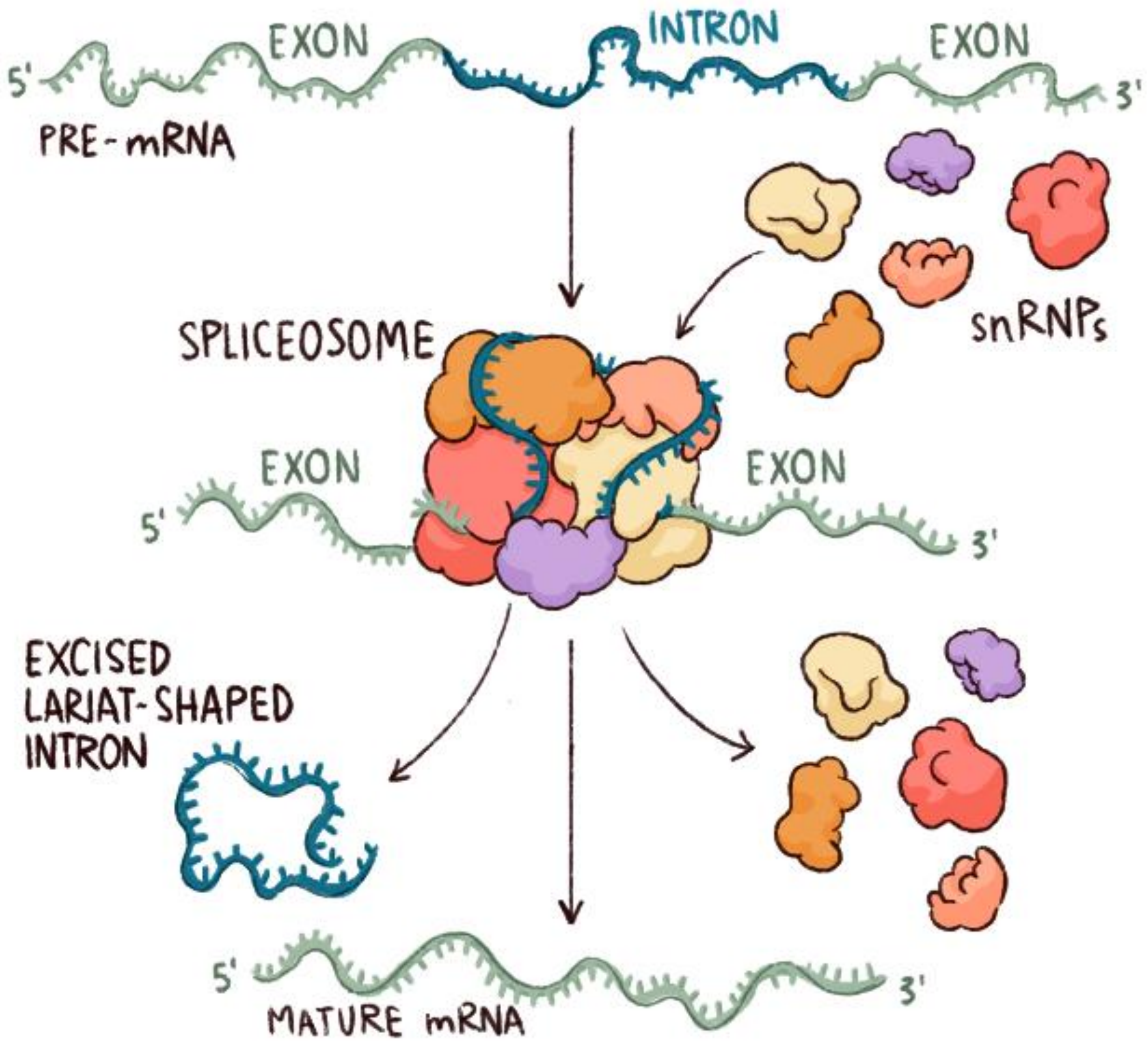


TABLE 26-3 Mechanisms of RNA Splicing

Mechanism	**Components**	**Features**	**Cellular locations**
Group I Intron	Catalytic RNA	Self-splicing using a guanine-derived cofactor	Found in nuclear, mitochondrial, and chloroplast genes that encode mRNAs, rRNAs, or tRNAs. Can be found in bacteria.
Group II Intron	Catalytic RNA; maturase and reverse transcriptase proteins	Self-splicing using a nucleophile within the intron to form a lariat	Primarily found in mitochondrial and chloroplast genes of fungi, algae, and plants. Can be found in bacteria.
Spliceosome	Catalytic snRNAs; dozens of protein splicing factors	Requires a large RNP for processing using a nucleophile within the intron to form a lariat	Found in nuclear genes of eukaryotes. Capable of alternative splicing to create multiple products from a given transcript.
Protein-catalyzed	Protein enzymes	Uses a splicing endonuclease and ligase	Found in tRNAs and a few mRNAs.

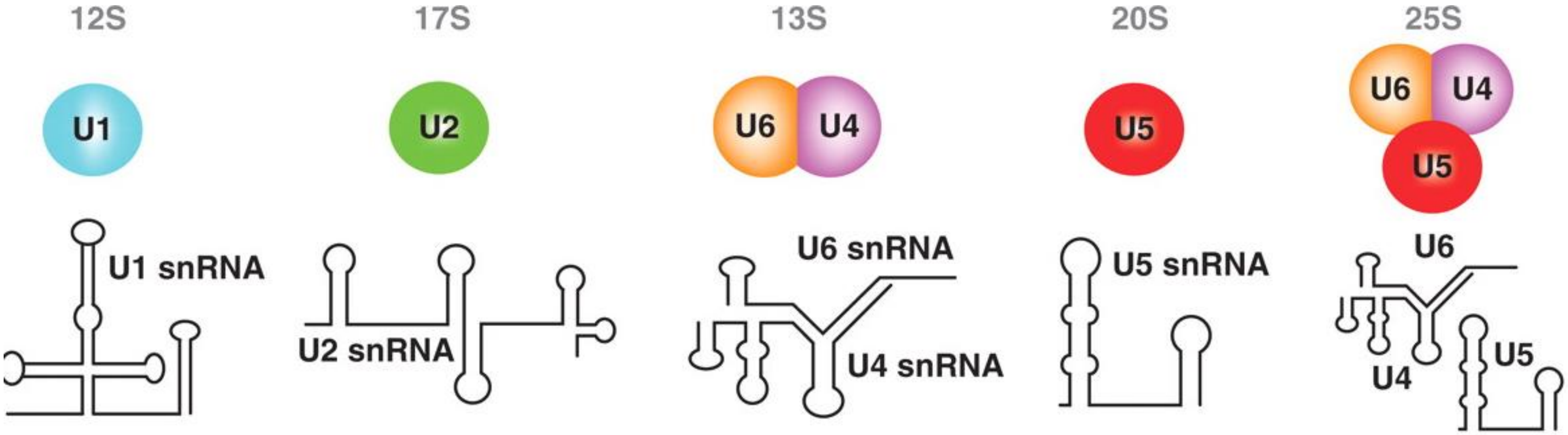
Most of RNA splicing in eukaryotes is performed by the spliceosome



snRNA + Proteins

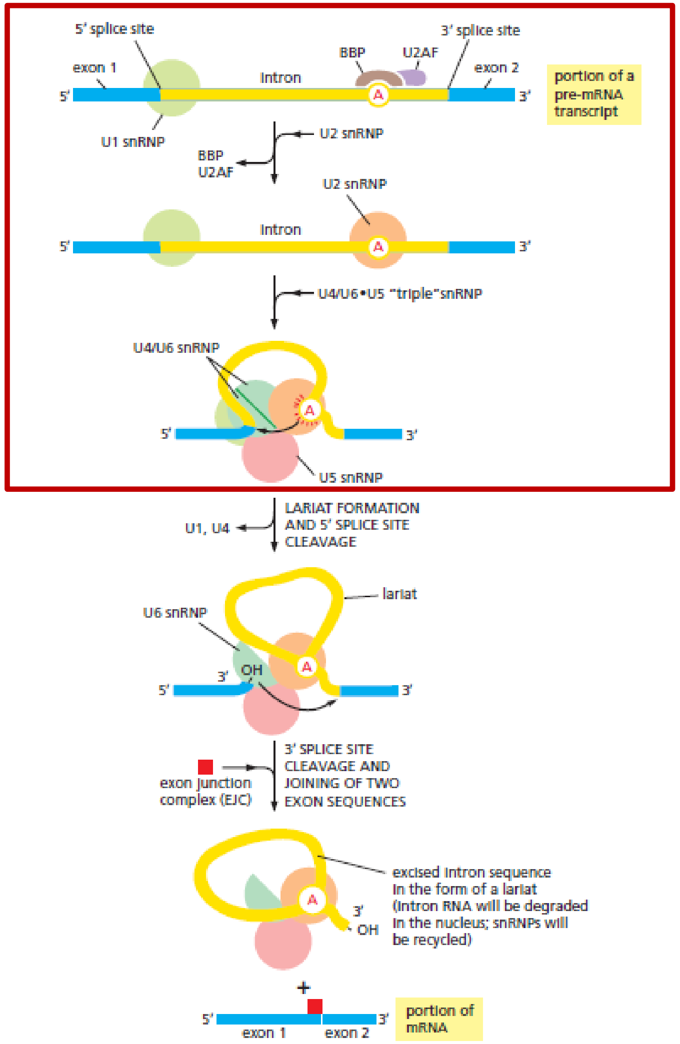
chemistryworld.com

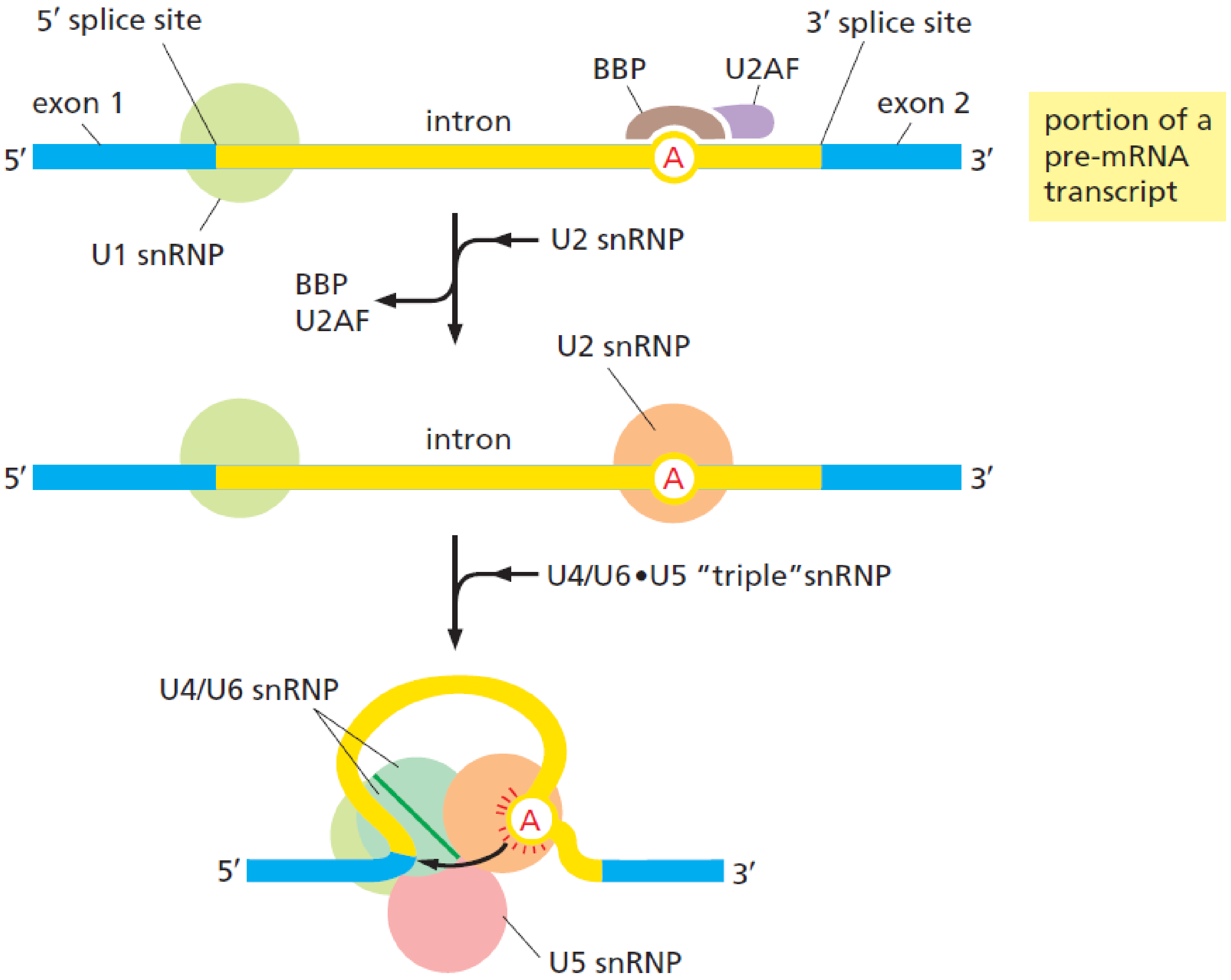
Spliceosome components



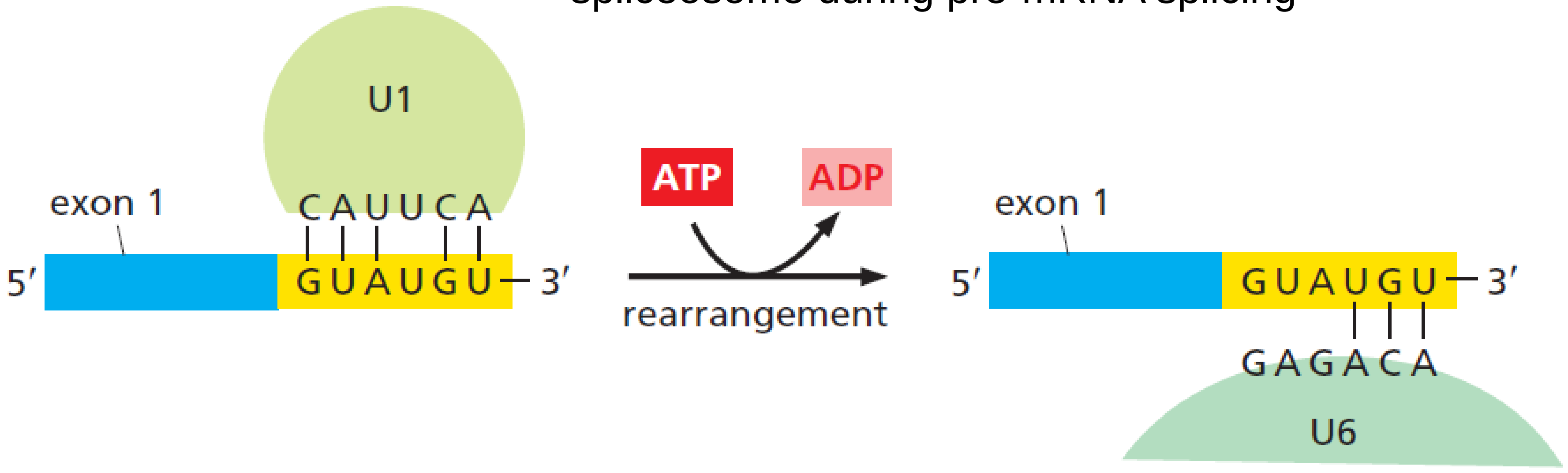
Most of RNA splicing in eukaryotes is performed by the spliceosome

Pre-mRNA splicing process



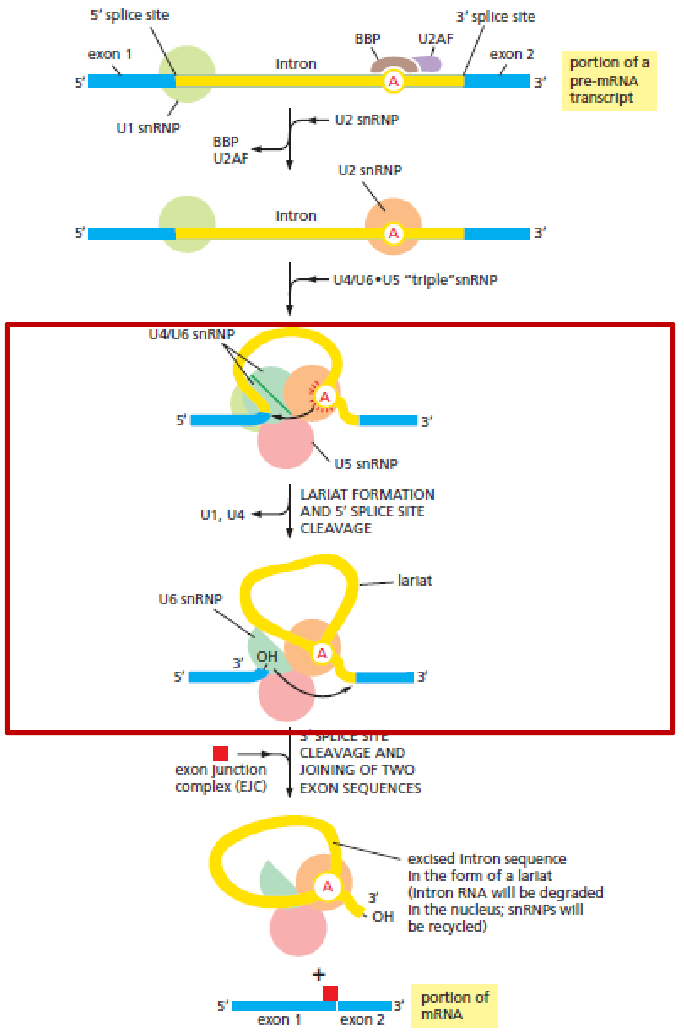


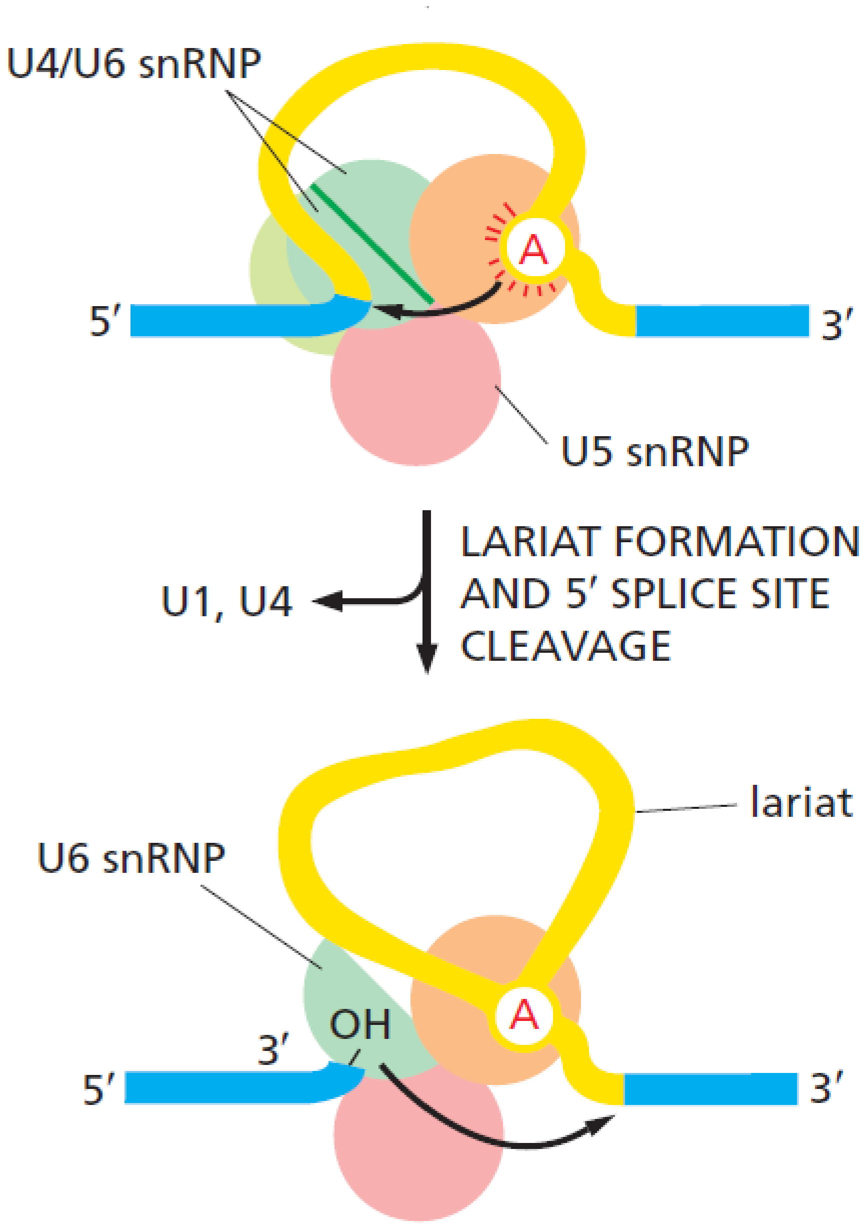
One of the many
rearrangements that take place in the
spliceosome during pre-mRNA splicing



Most of RNA splicing in eukaryotes is performed by the spliceosome

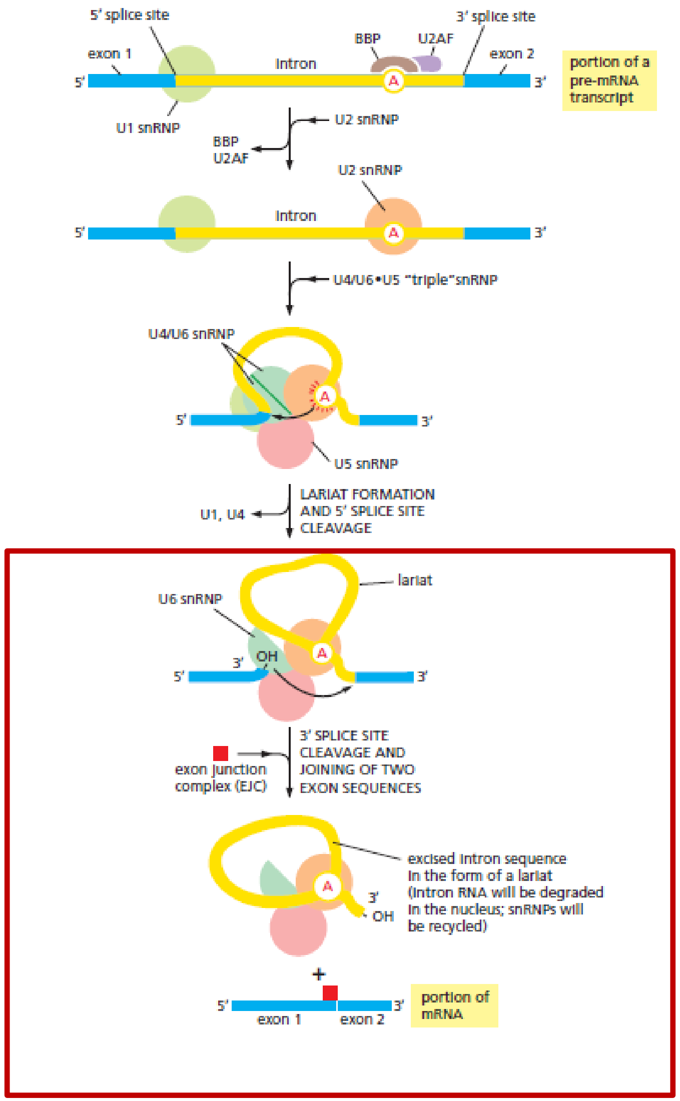
Pre-mRNA splicing process

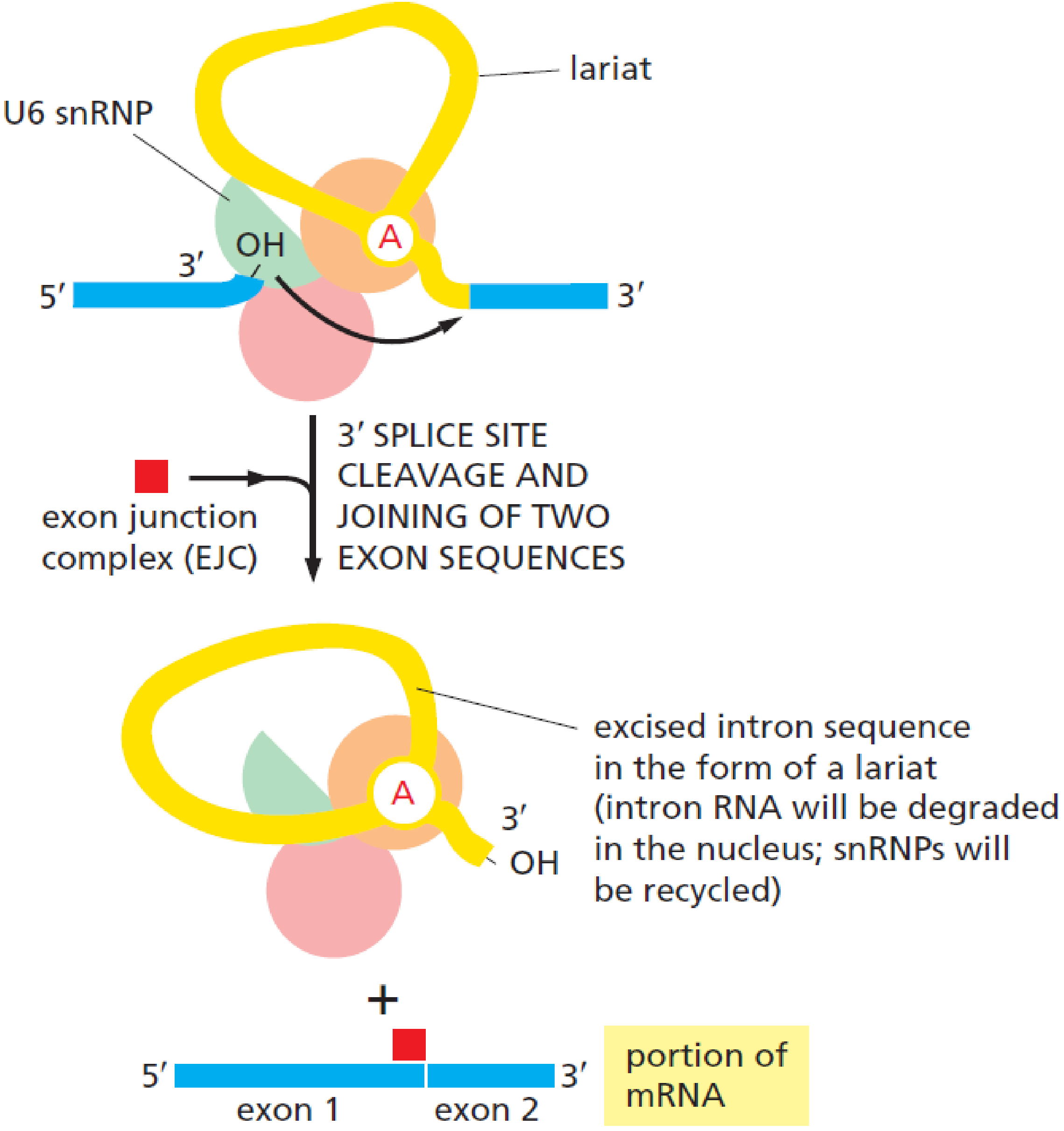




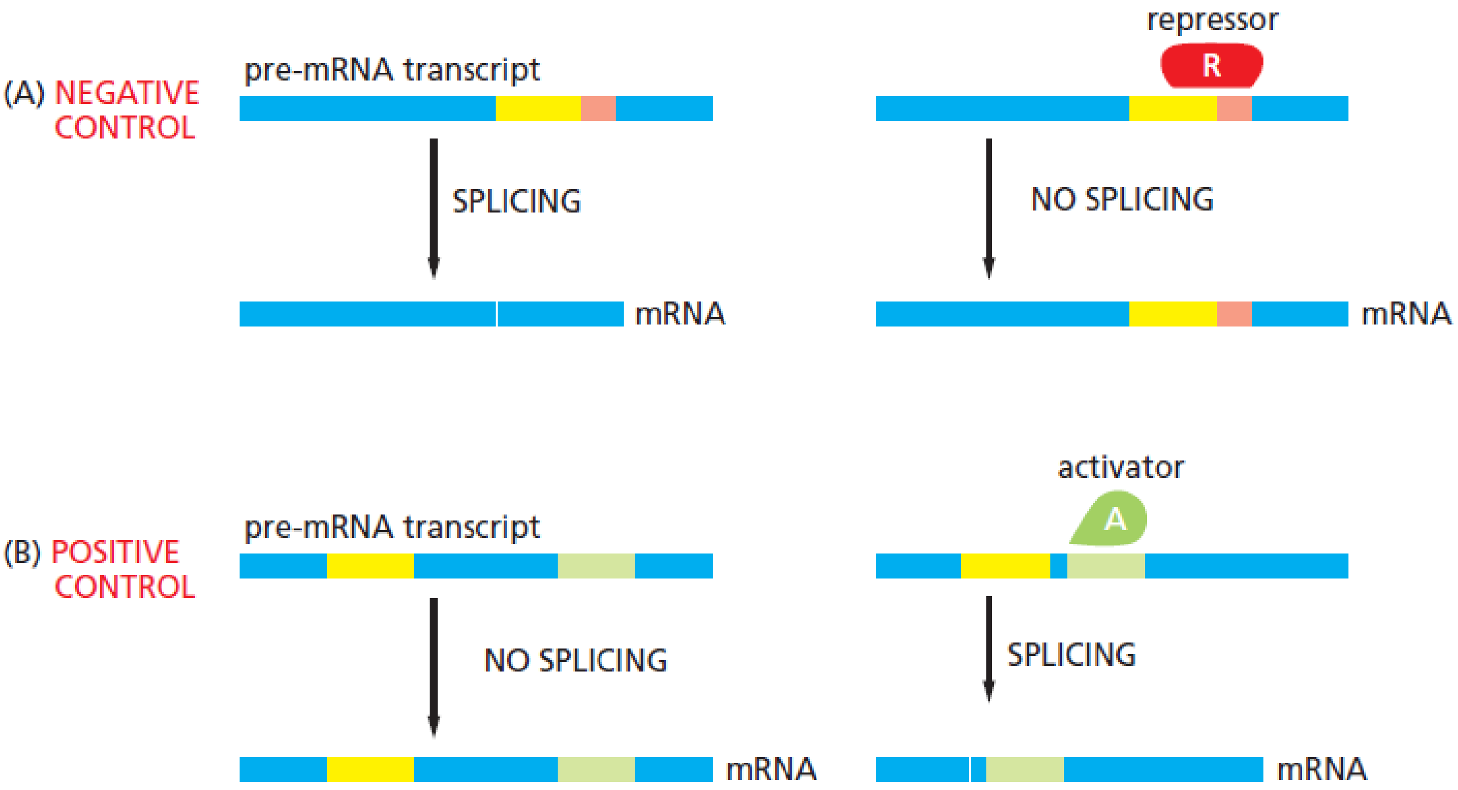
Most of RNA splicing in eukaryotes is performed by the spliceosome

Pre-mRNA splicing process





Control of RNA splicing



(A) NEGATIVE CONTROL

pre-mRNA transcript

SPLICING

mRNA

repressor

R

NO SPLICING

mRNA

(B) POSITIVE CONTROL

pre-mRNA transcript

NO SPLICING

mRNA

activator

A

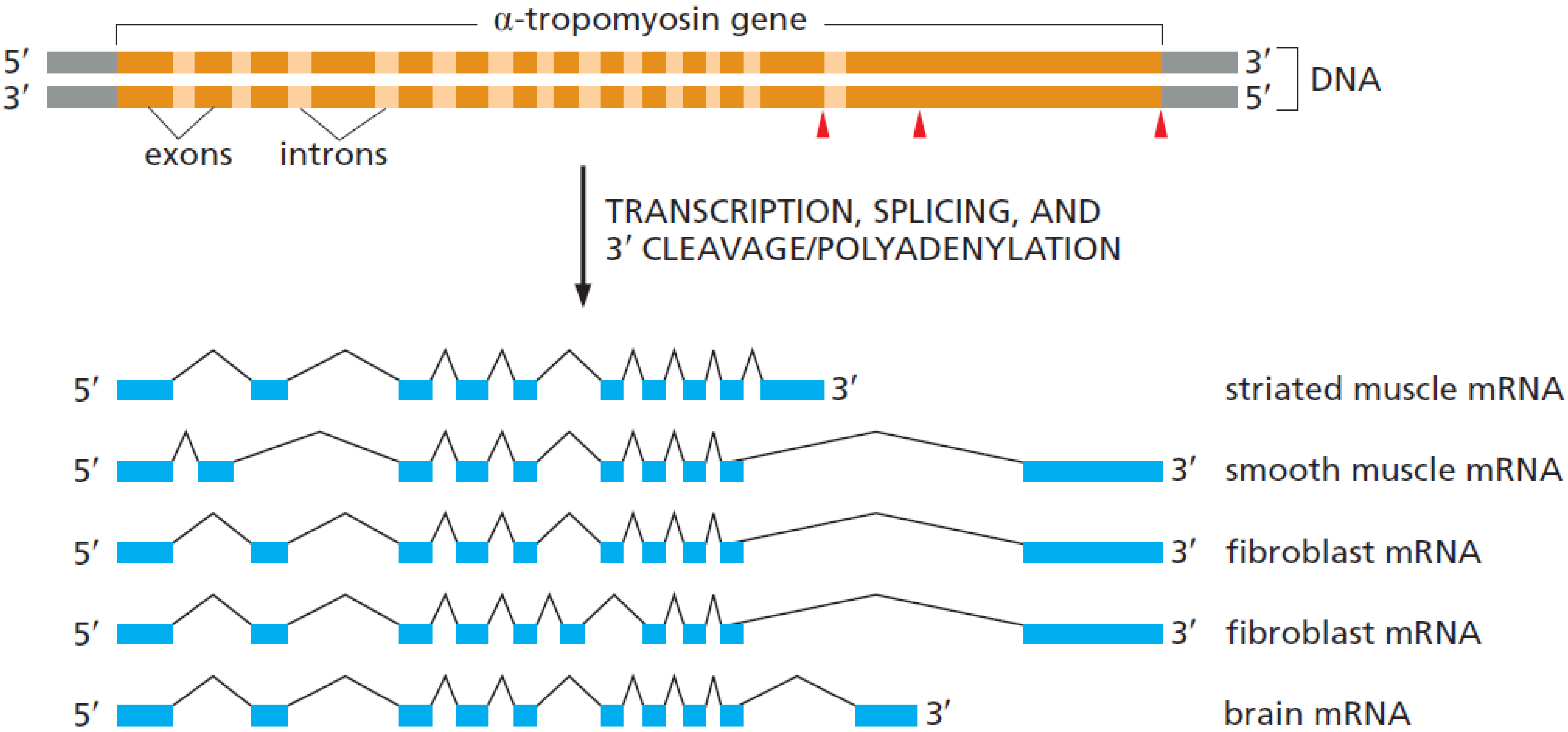
SPLICING

mRNA

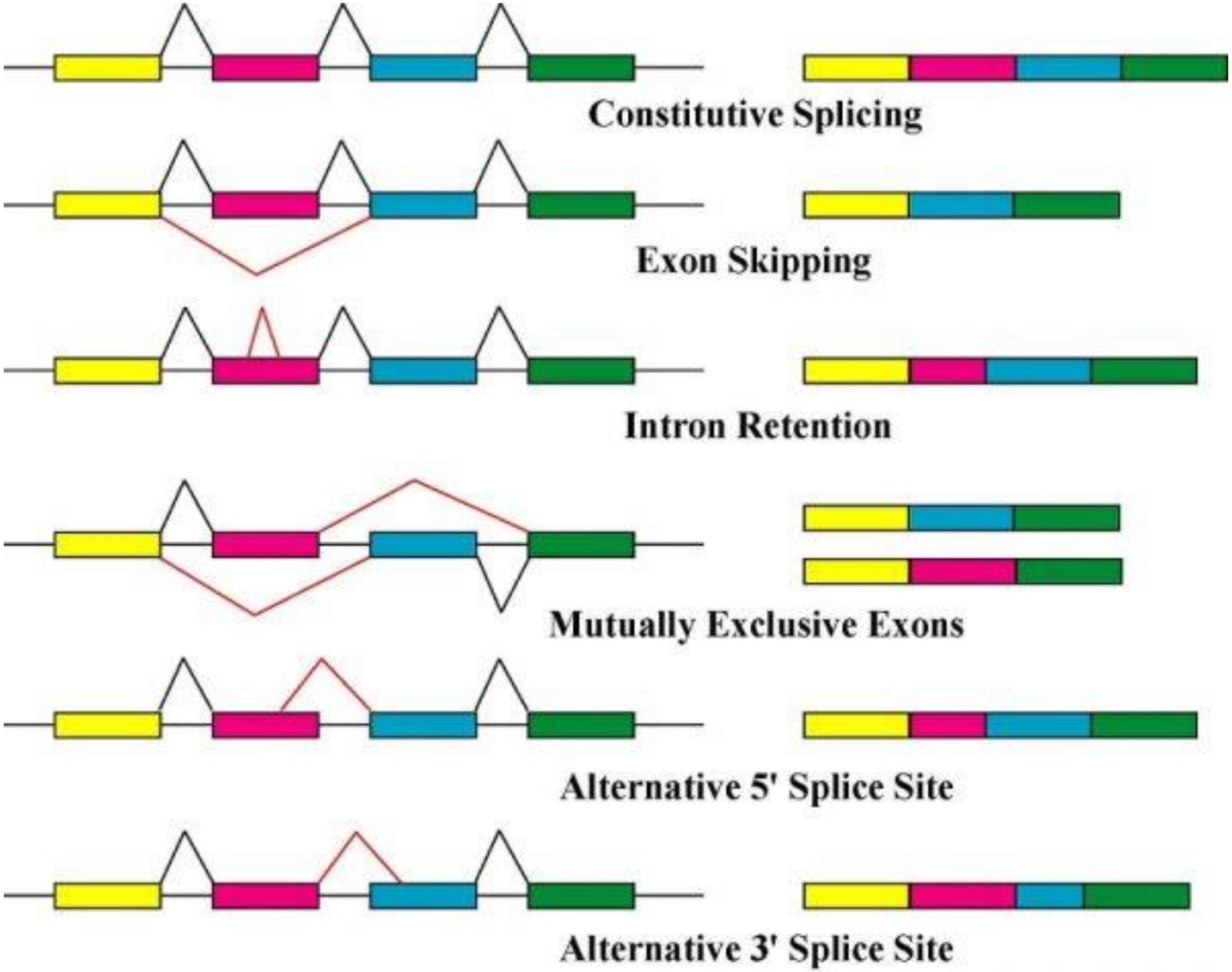
Alternative splicing

Alternative splicing is a process by which the exons of the RNA produced by transcription of a gene (a primary gene transcript or pre-mRNA) are reconnected in multiple ways during RNA splicing. The resulting different mRNAs may be translated into different protein isoforms; thus, a single gene may code for multiple proteins.

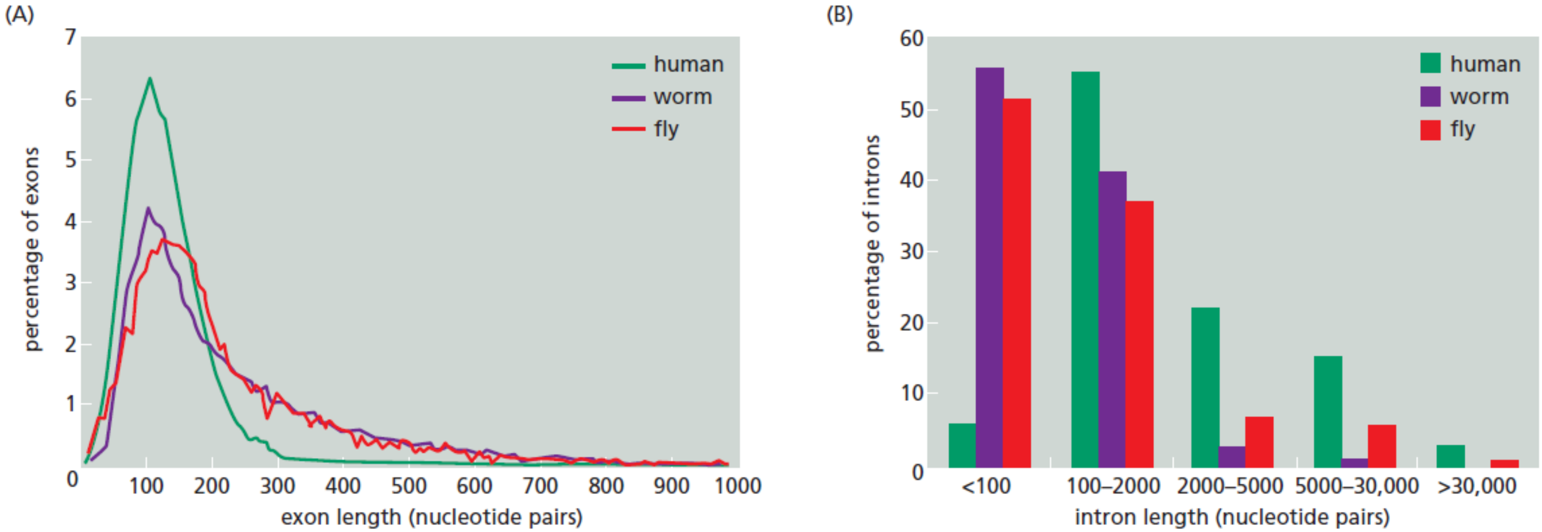
Example of alternative splicing

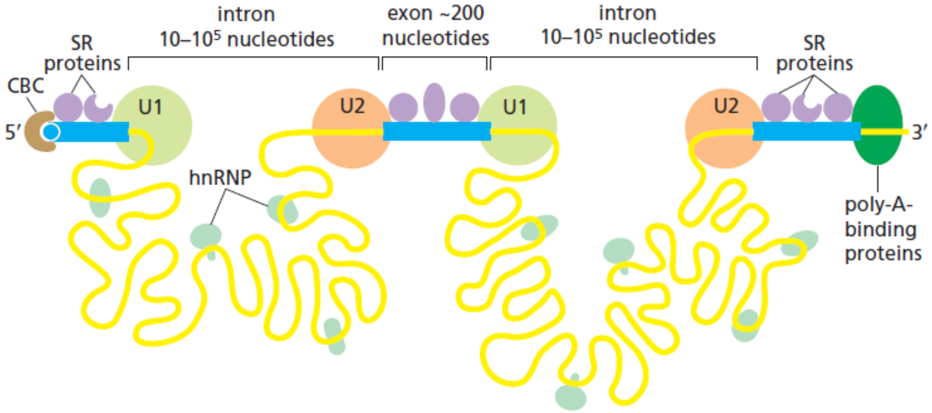


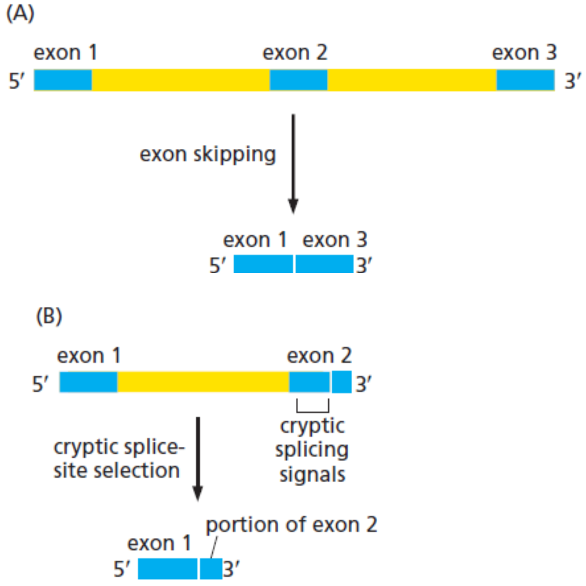
Mechanisms of alternative splicing



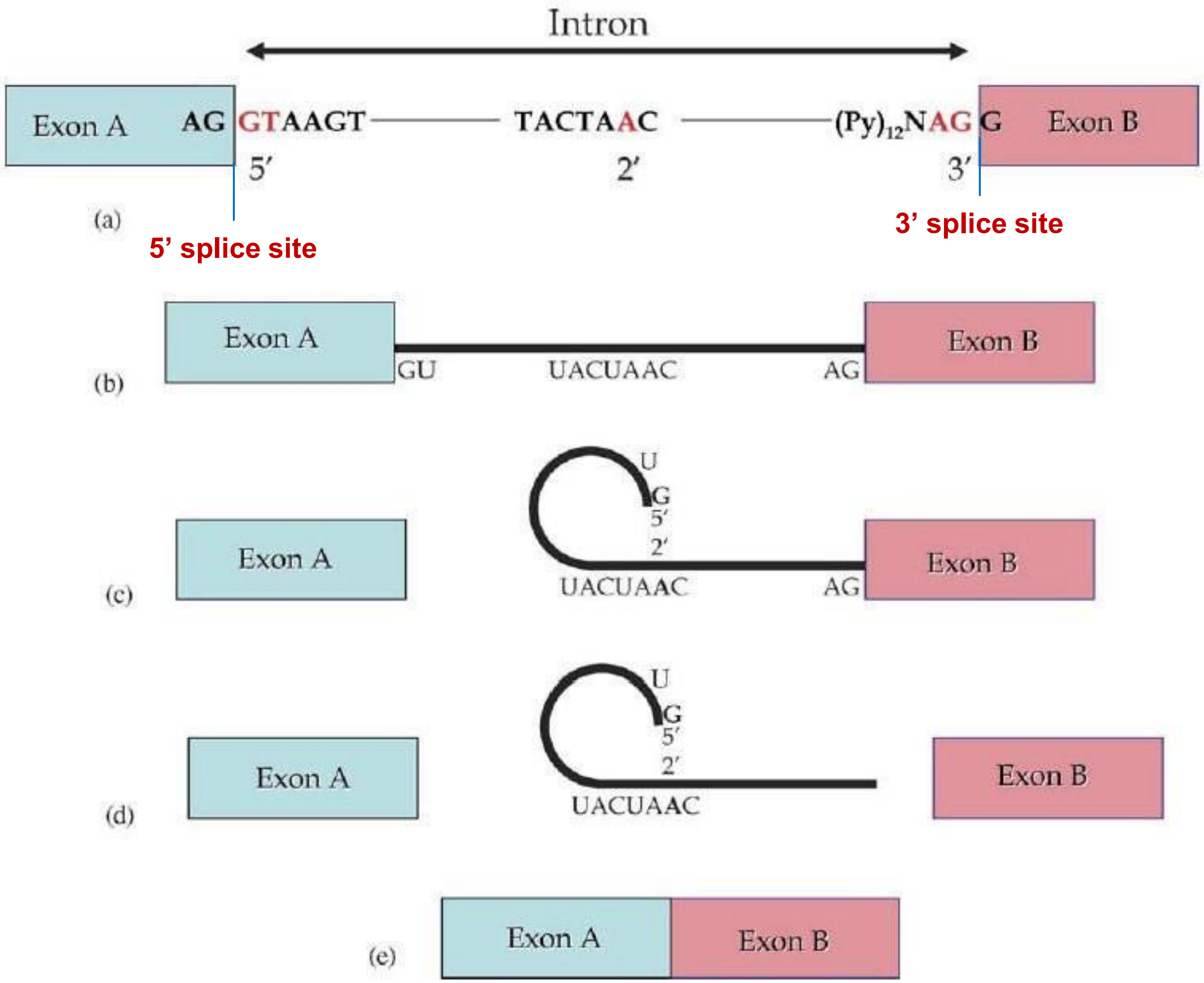
How cell regulates alternative splicing?







Controlled by the splicing machinery, transcription process, and chromatin structure



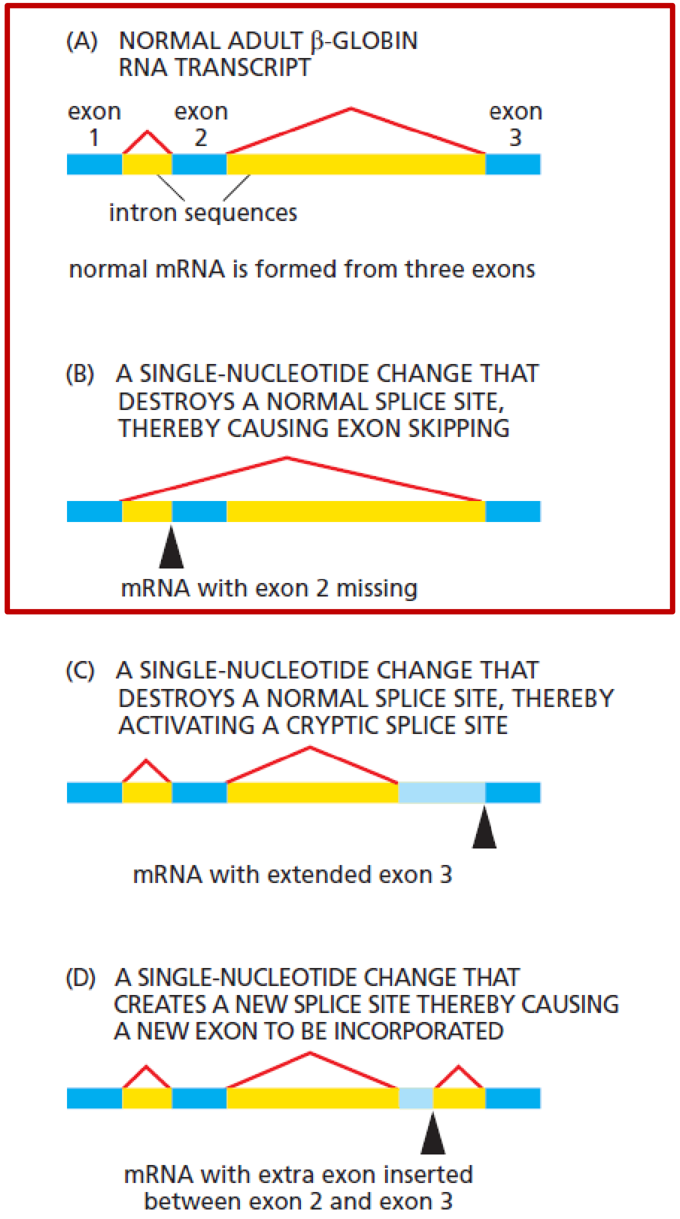
DNA mutation at the splice site

DNA mutation that make new splice site

↓

mRNA splicing error

Examples of mRNA splicing error



(A) NORMAL ADULT β-GLOBIN RNA TRANSCRIPT

normal mRNA is formed from three exons

(B) A SINGLE-NUCLEOTIDE CHANGE THAT DESTROYS A NORMAL SPLICE SITE, THEREBY CAUSING EXON SKIPPING

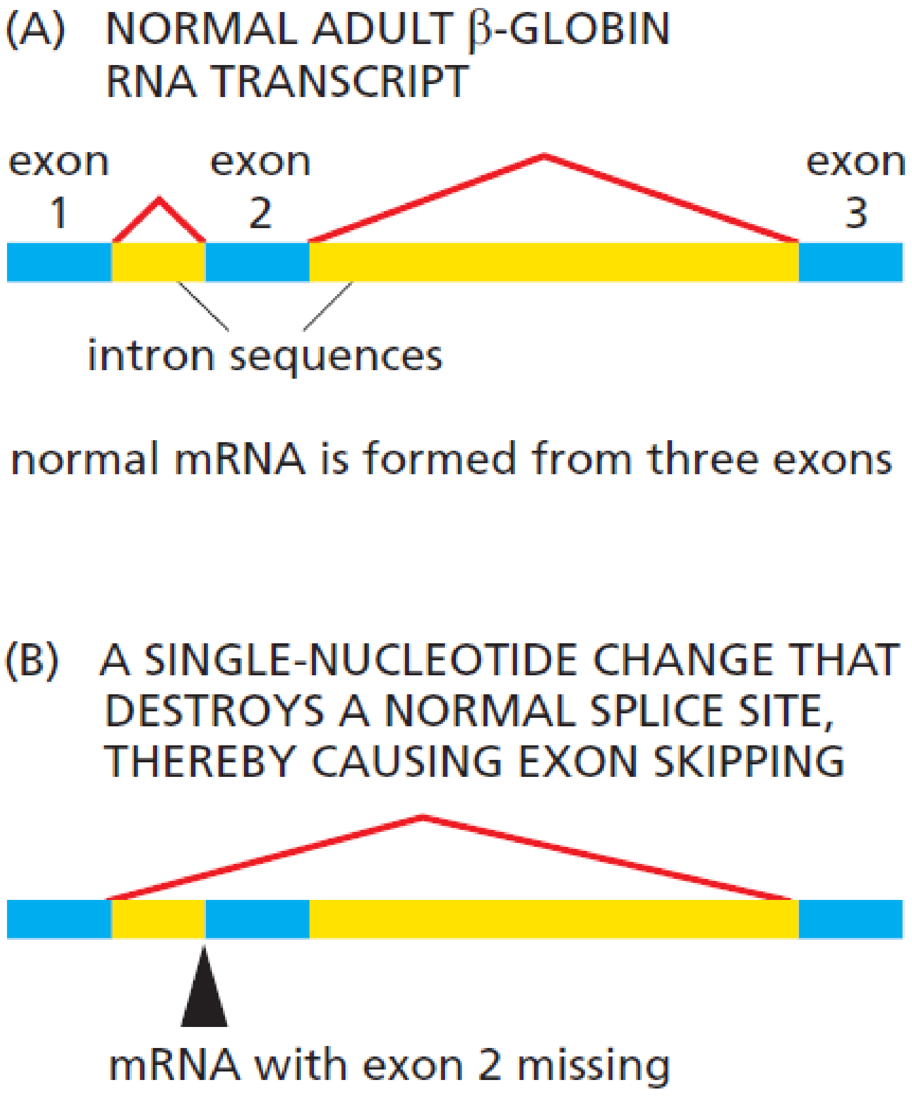
mRNA with exon 2 missing

(C) A SINGLE-NUCLEOTIDE CHANGE THAT DESTROYS A NORMAL SPLICE SITE, THEREBY ACTIVATING A CRYPTIC SPLICE SITE

mRNA with extended exon 3

(D) A SINGLE-NUCLEOTIDE CHANGE THAT CREATES A NEW SPLICE SITE THEREBY CAUSING A NEW EXON TO BE INCORPORATED

mRNA with extra exon inserted between exon 2 and exon 3



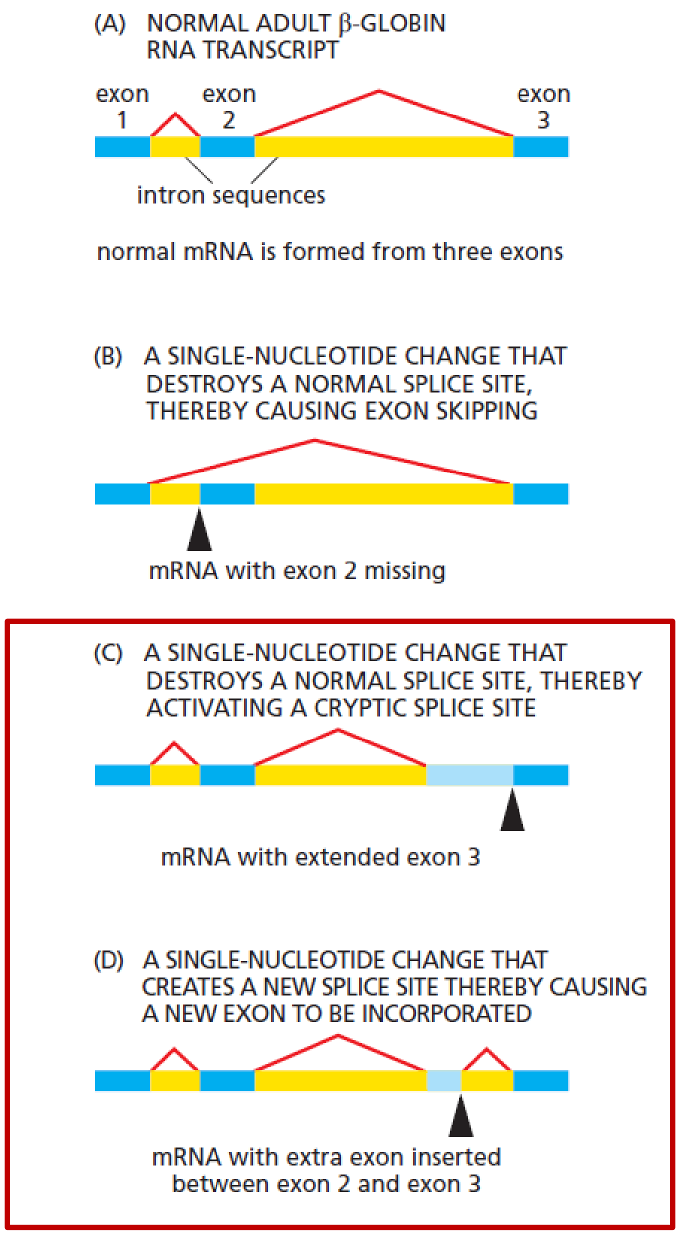
(A) NORMAL ADULT β-GLOBIN RNA TRANSCRIPT

normal mRNA is formed from three exons

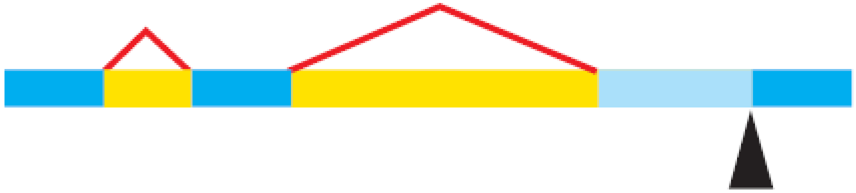
(B) A SINGLE-NUCLEOTIDE CHANGE THAT DESTROYS A NORMAL SPLICE SITE, THEREBY CAUSING EXON SKIPPING

mRNA with exon 2 missing

Examples of mRNA splicing error

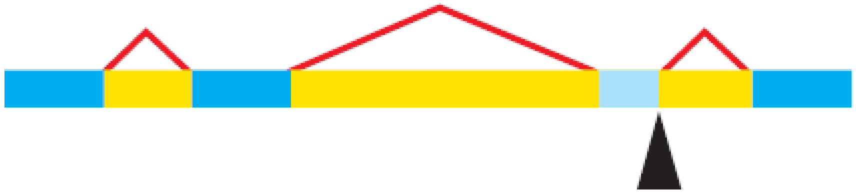


(C) A SINGLE-NUCLEOTIDE CHANGE THAT DESTROYS A NORMAL SPLICE SITE, THEREBY ACTIVATING A CRYPTIC SPLICE SITE



mRNA with extended exon 3

(D) A SINGLE-NUCLEOTIDE CHANGE THAT CREATES A NEW SPLICE SITE THEREBY CAUSING A NEW EXON TO BE INCORPORATED

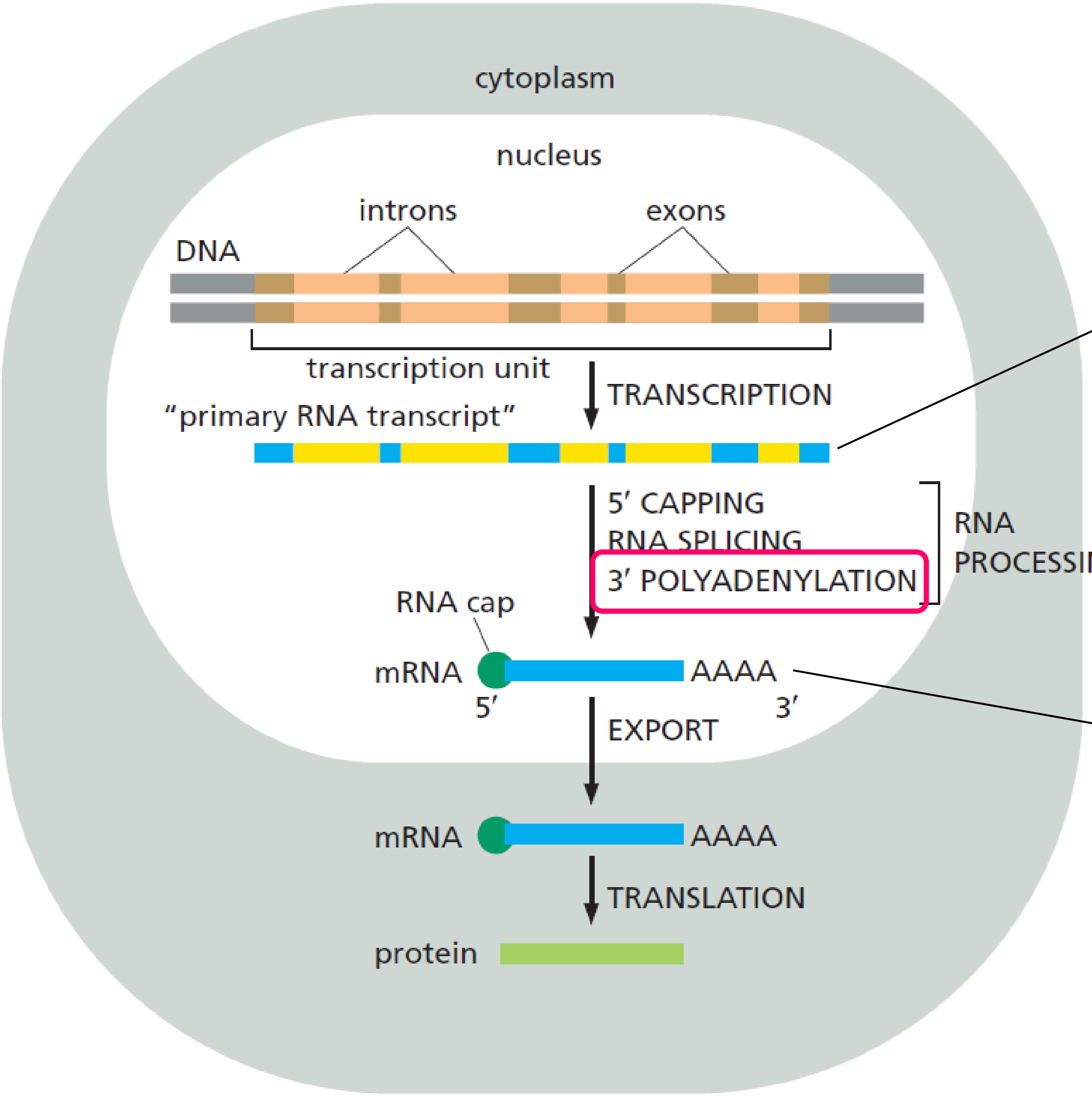


mRNA with extra exon inserted
between exon 2 and exon 3

the *light blue boxes* depict new nucleotide sequences included in the final mRNA molecule as a result of the mutation denoted by the *black arrowhead*.

What is the difference between alternative splicing & splicing error?

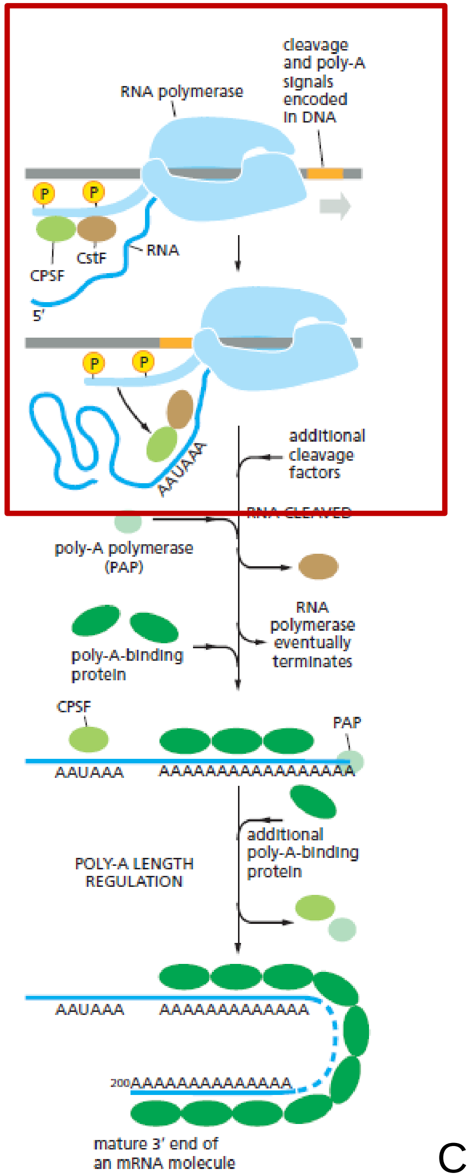
Transcription Elongation in Eukaryotes Is Tightly Coupled to RNA Processing

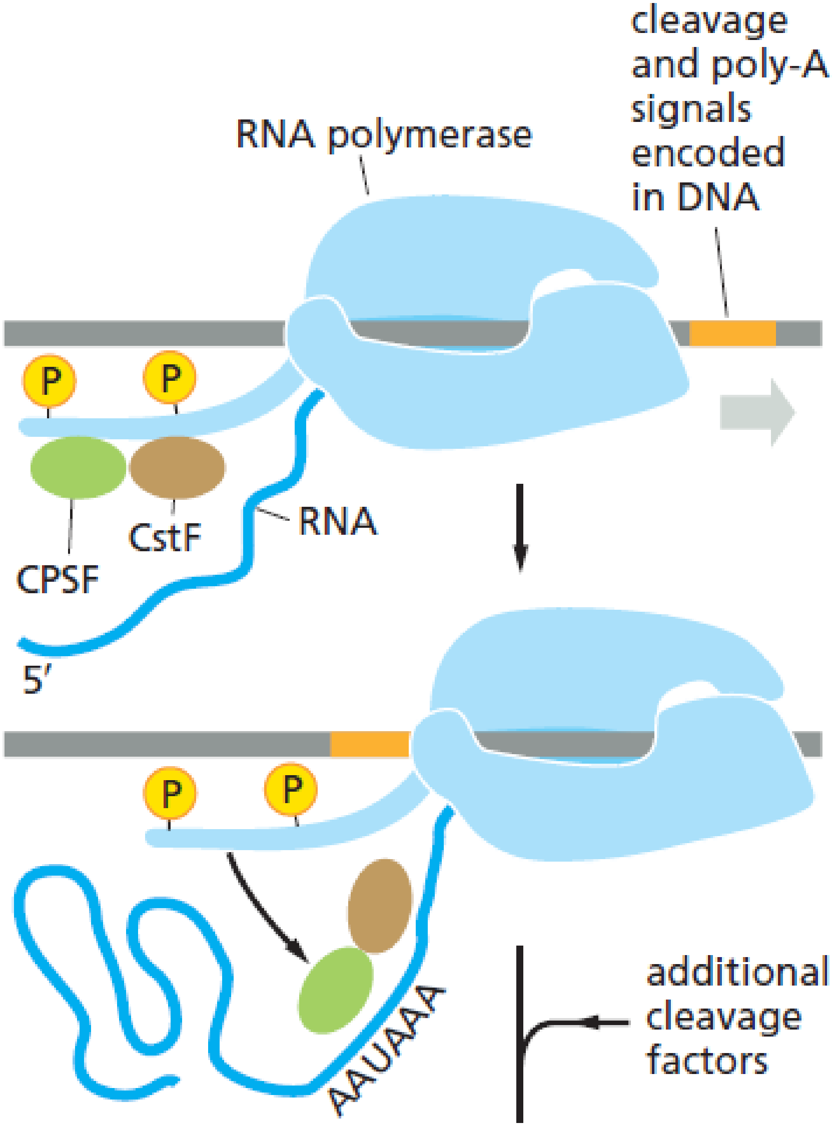


Primary mRNA, Pre-mRNA

Mature RNA

3’ Polyadenylation of mRNA

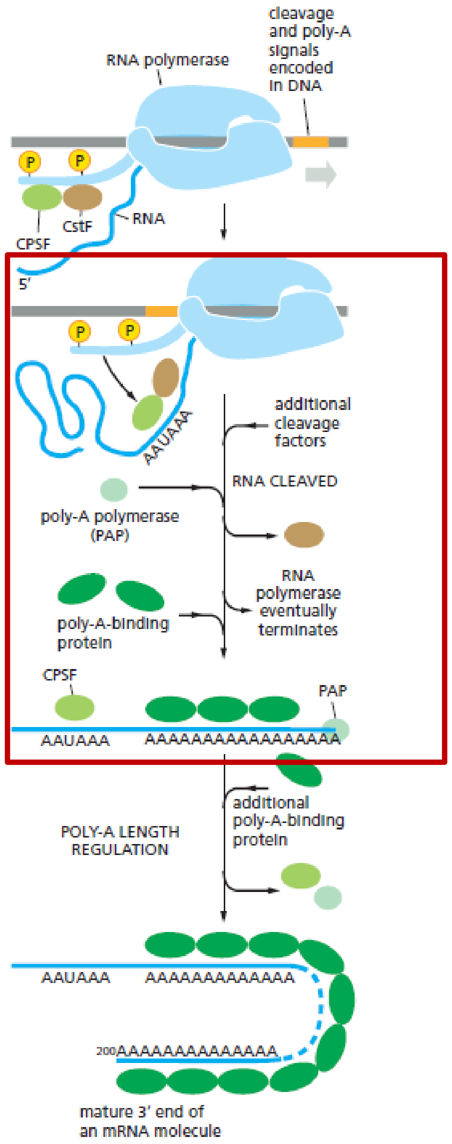


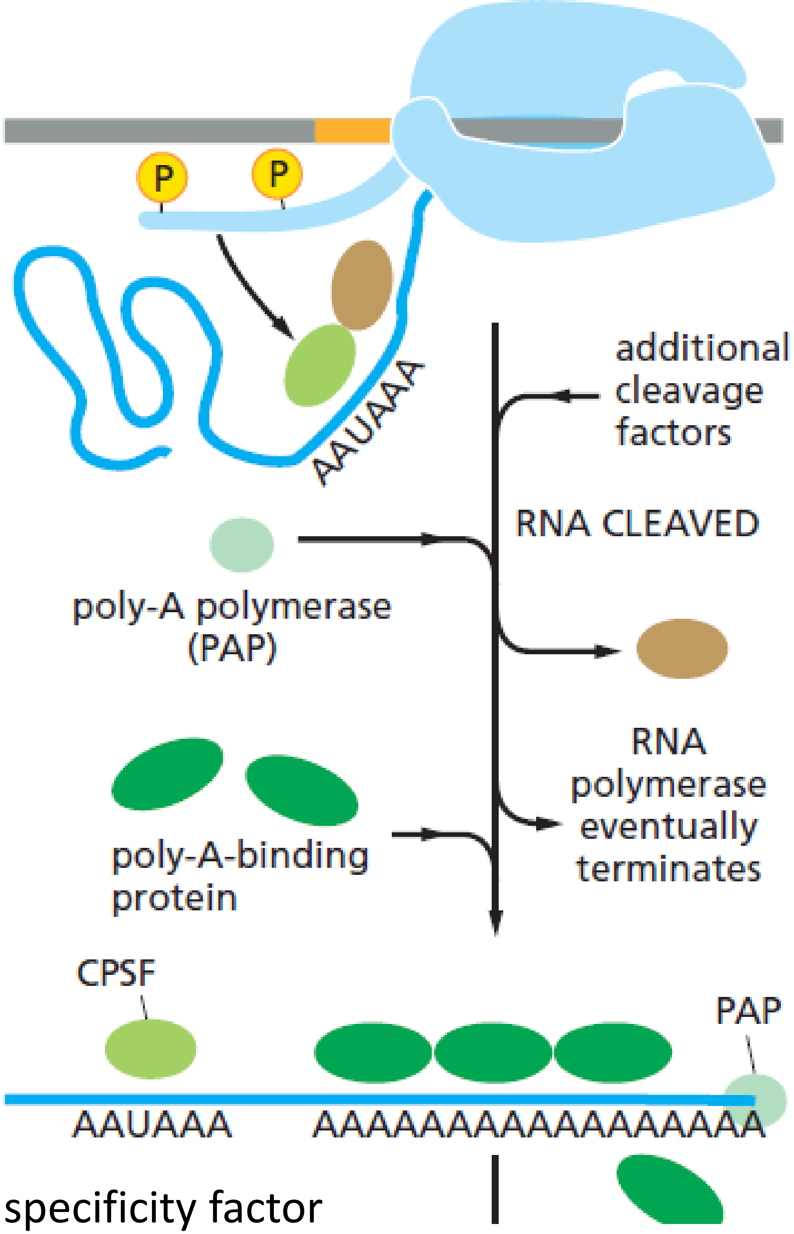


CstF = cleavage stimulation factor

CPSF = cleavage and polyadenylation specificity factor

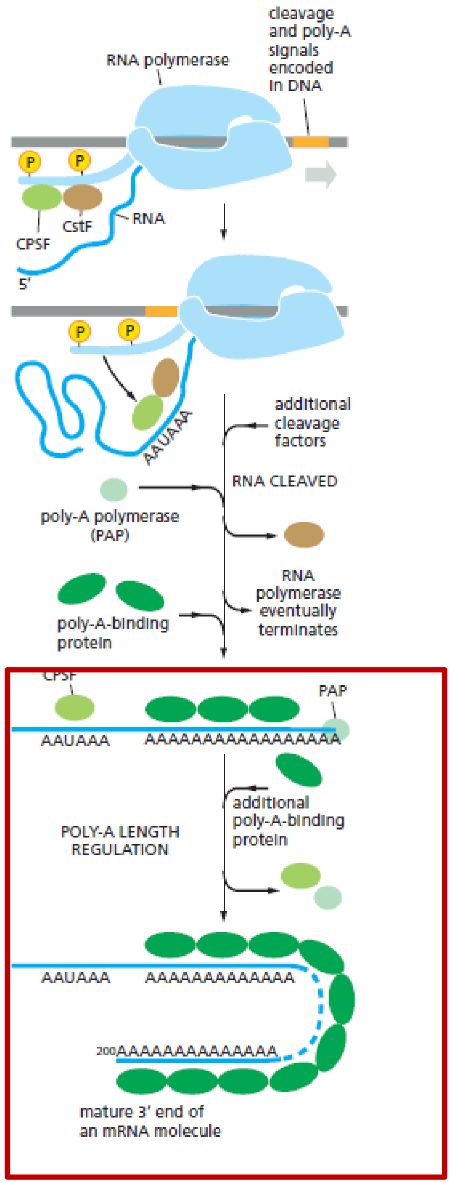
3’ Polyadenylation of mRNA

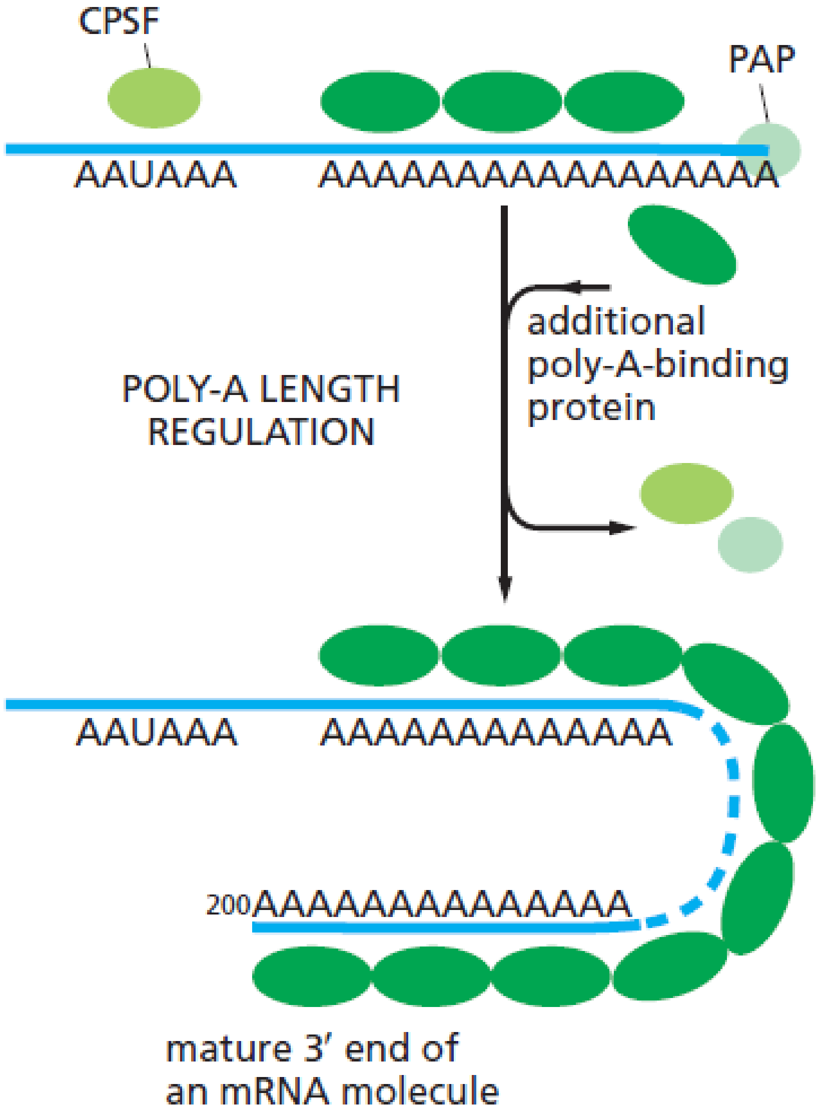




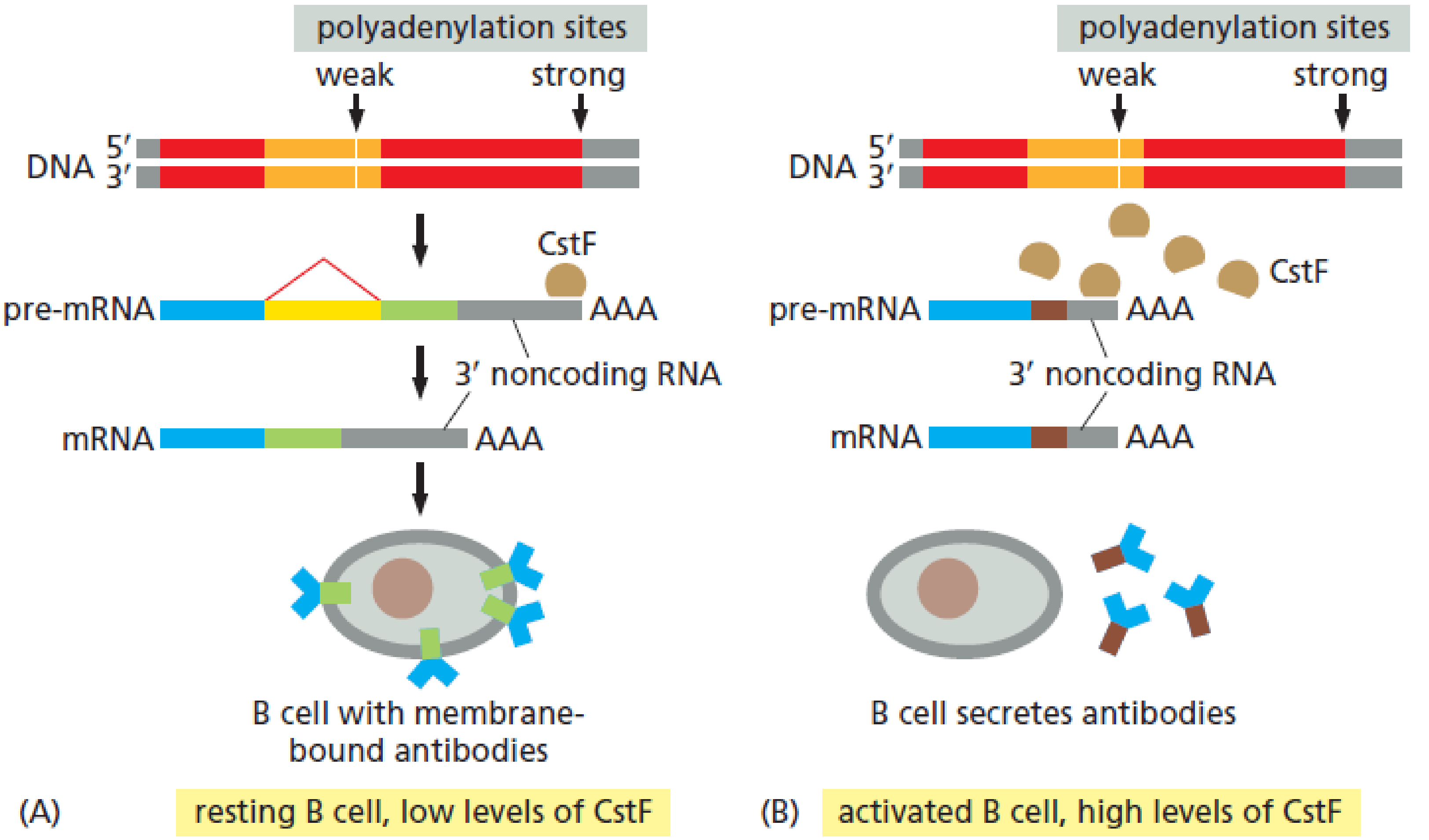
CPSF = cleavage and polyadenylation specificity factor

3' Polyadenylation of mRNA

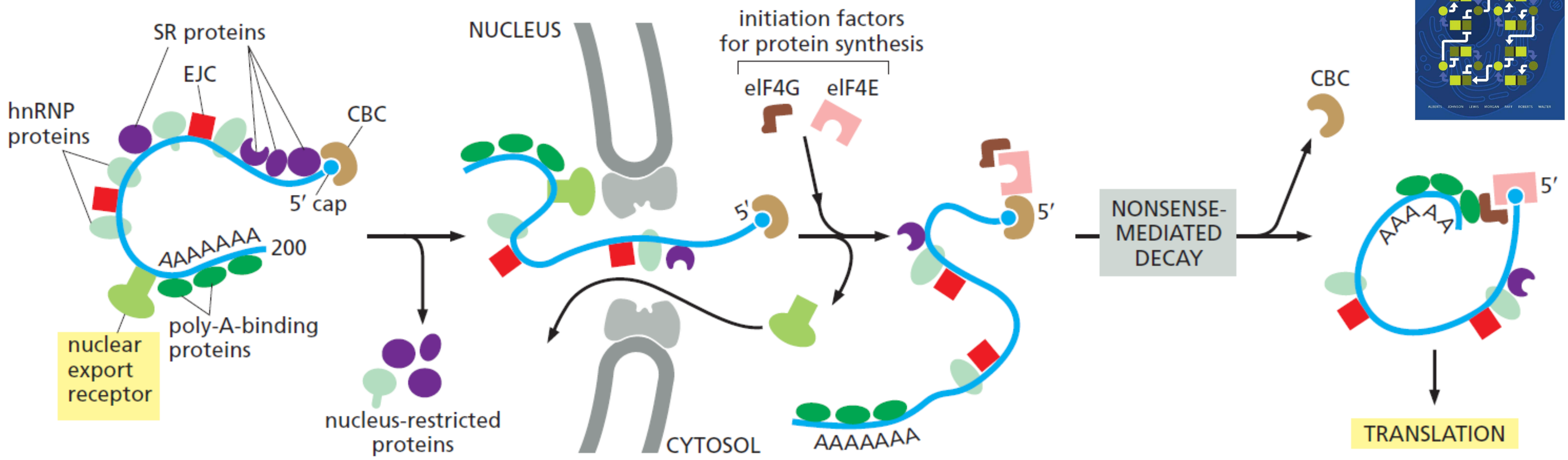


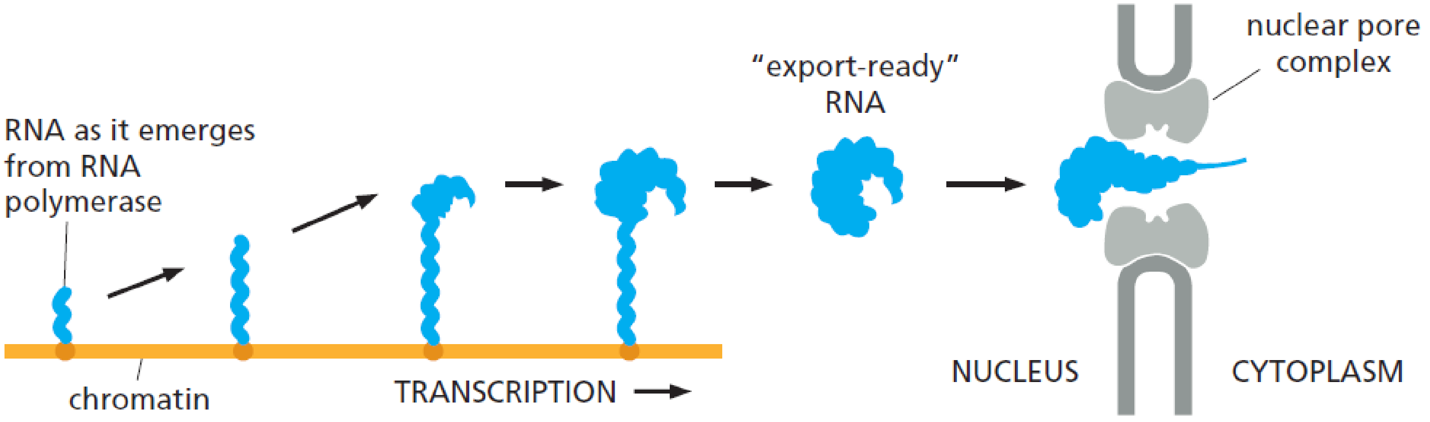


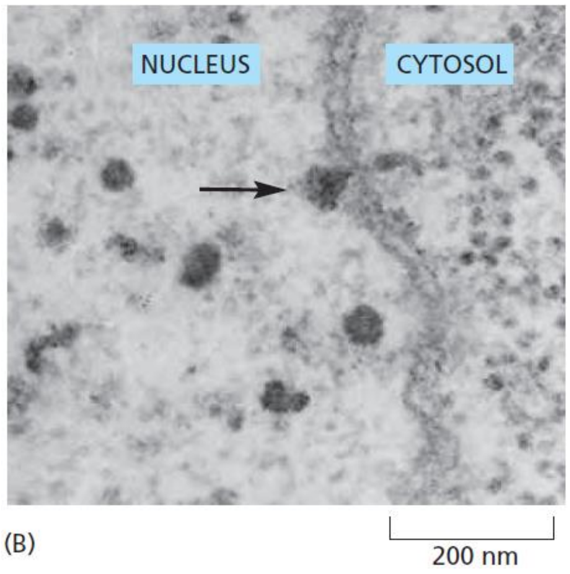
Regulation of the site of RNA cleavage and poly-A addition determines whether an antibody molecule is secreted or remains membrane-bound.



mRNA transport from nucleus to cytoplasm







RNA editing

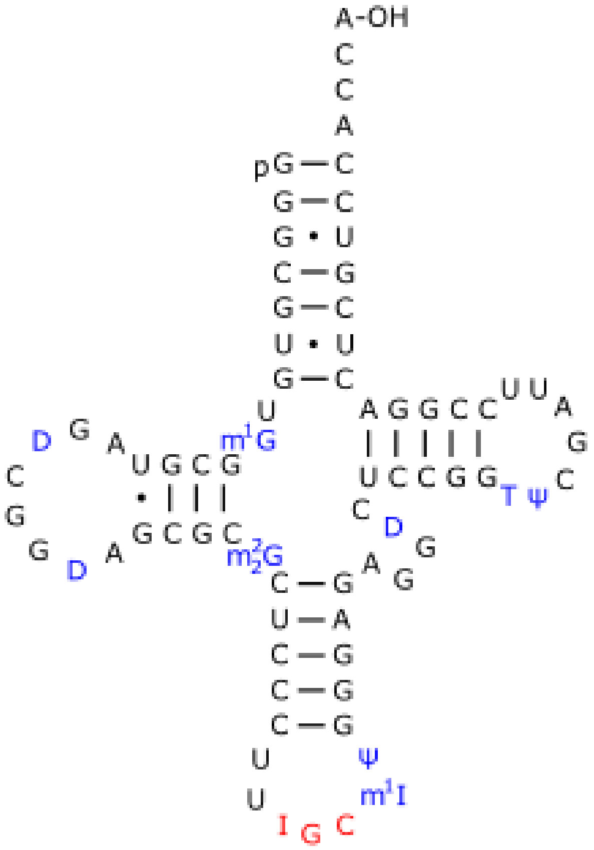
-The process by which RNA molecules are enzymatically modified post synthesis on specific nucleosides

-Can involve the addition, deletion, or alteration of nucleotides in the RNA in a manner that affects the meaning of the transcript when it is translated

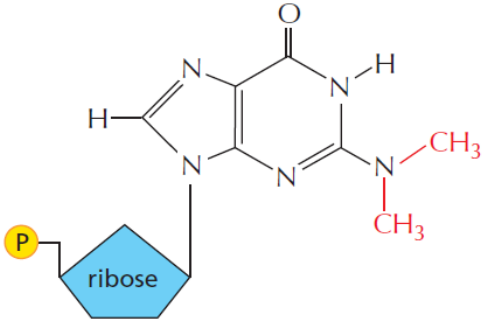
-One of the most prevalent and abundant forms of post-transcriptional RNA modification in normal physiological processes.

-The major types of RNA editing
    Pseudouridylation (the isomerization of uridine residues)
    U Insertion/deletion
    Deamination (removal of an amine group from nitrogenous bases (A or C)

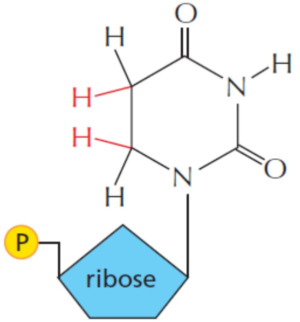
Examples of the unusual nucleotides found in tRNA molecules



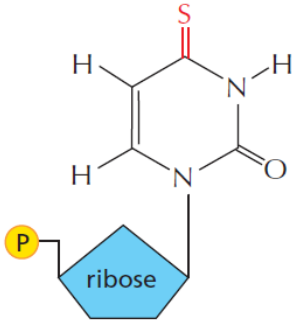
tRNA-Ala from yeast showing also modified bases in blue m1G: 1-methyl-guanosine D: 5,6-Dihydrouridine m22G: N2-dimethyl-guanosine I: Inosine m1I: 1-methyl-inosine Ψ: pseudouridine T: 5-Methyluridine (Ribothymidine)



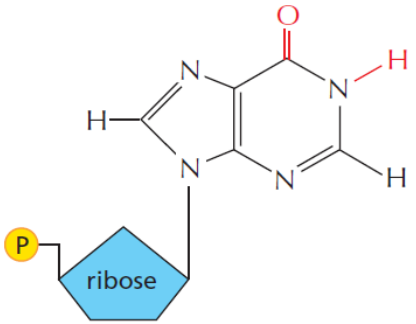
two methyl groups added to G
(*N,N*-dimethyl G)



two hydrogens added to U
(dihydro U)

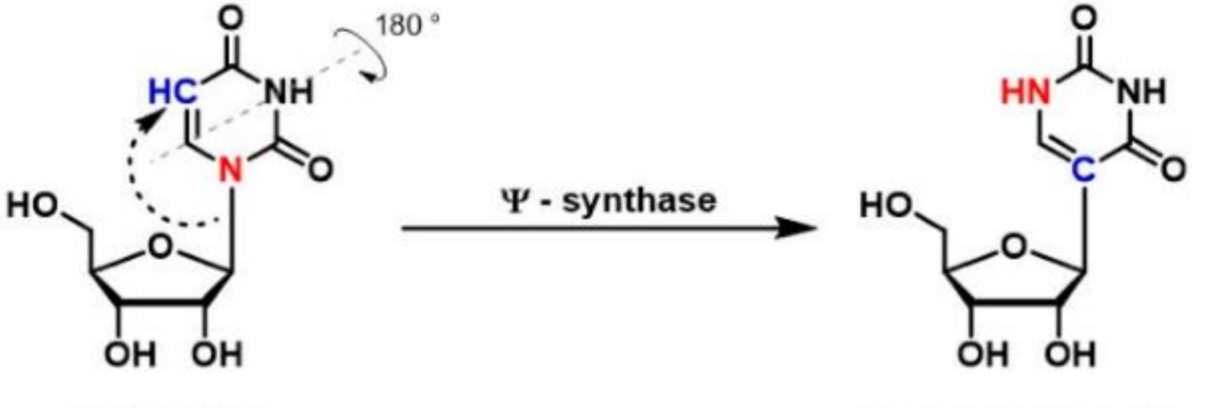


sulfur replaces oxygen in U
(4-thiouridine)



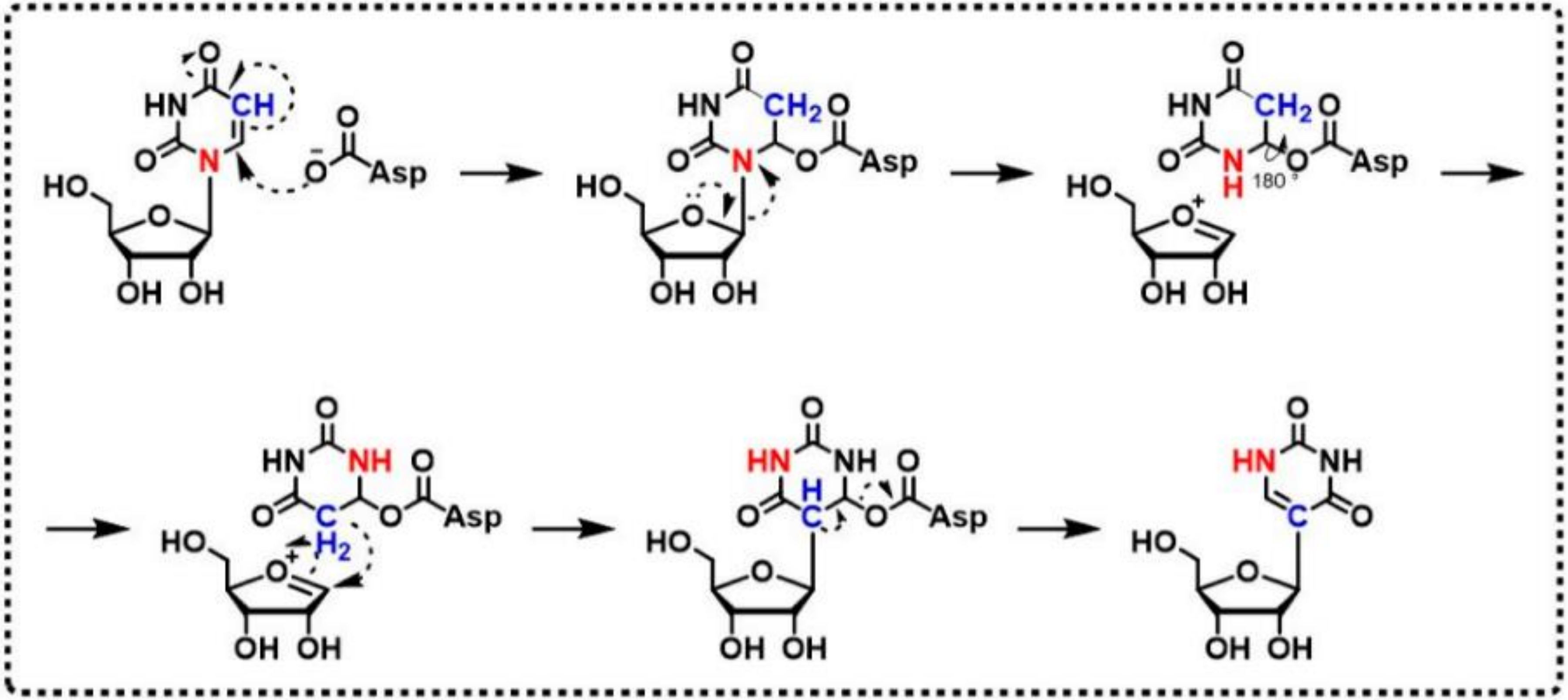
deamination of A
(inosine)

Pseudouridylation

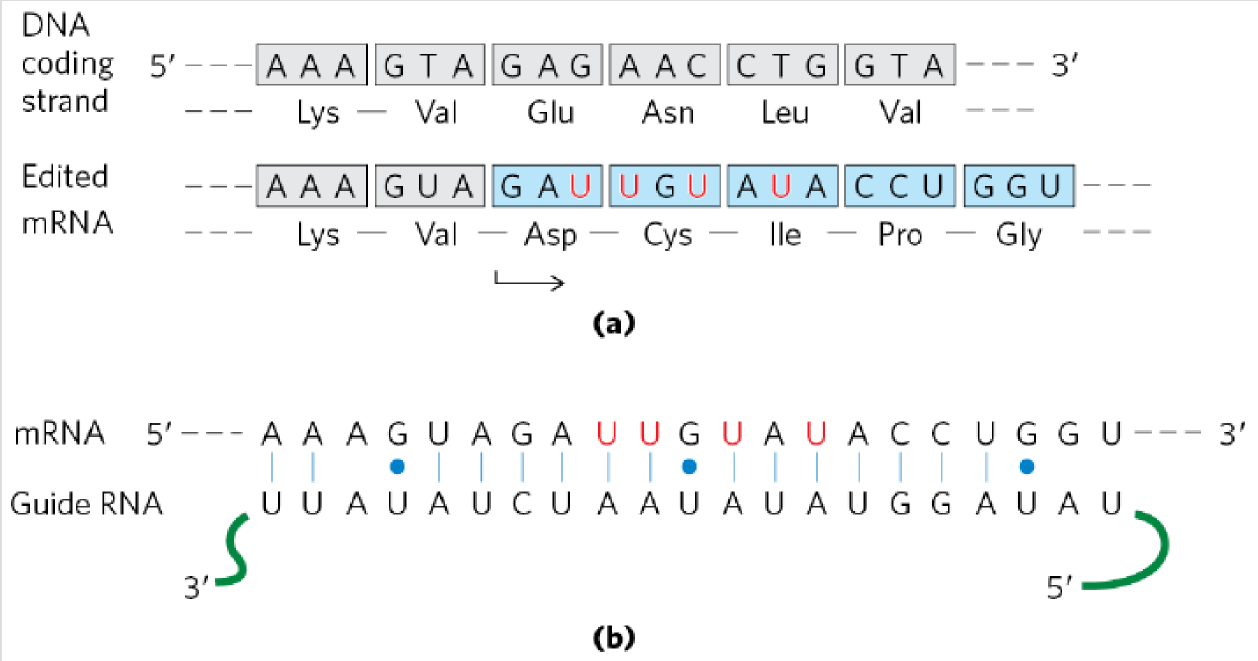


Uridine (U)

Pseudouridine (Ψ)



Uridine (U) insertion



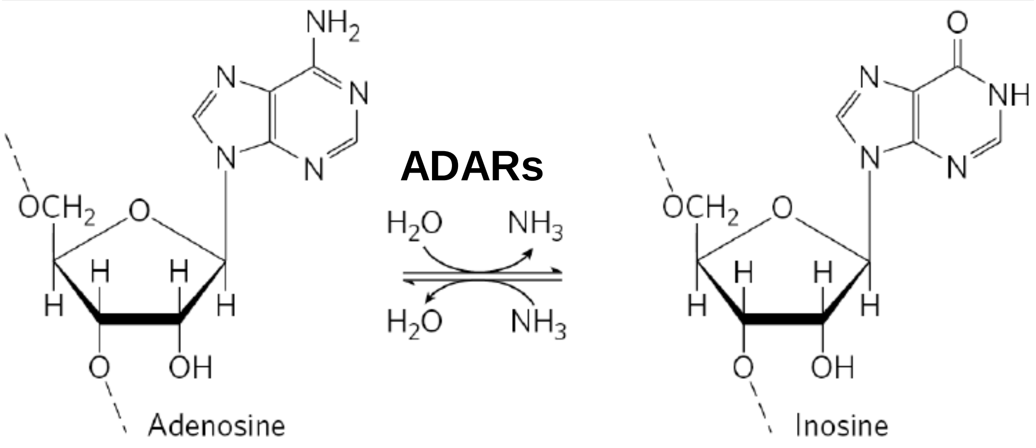
Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

FIGURE 27-10 RNA editing of the transcript of the cytochrome oxidase subunit II gene from *Trypanosoma brucei* mitochondria. (a) Insertion of four U residues (red) produces a revised reading frame. (b) A special class of guide RNAs, complementary to the edited product, acts as templates for the editing process. Note the presence of three G═U base pairs, signified by blue dots to indicate non-Watson-Crick pairing.

The RNA editing alters the sequences of mRNAs, it makes them different from the corresponding genomic template.

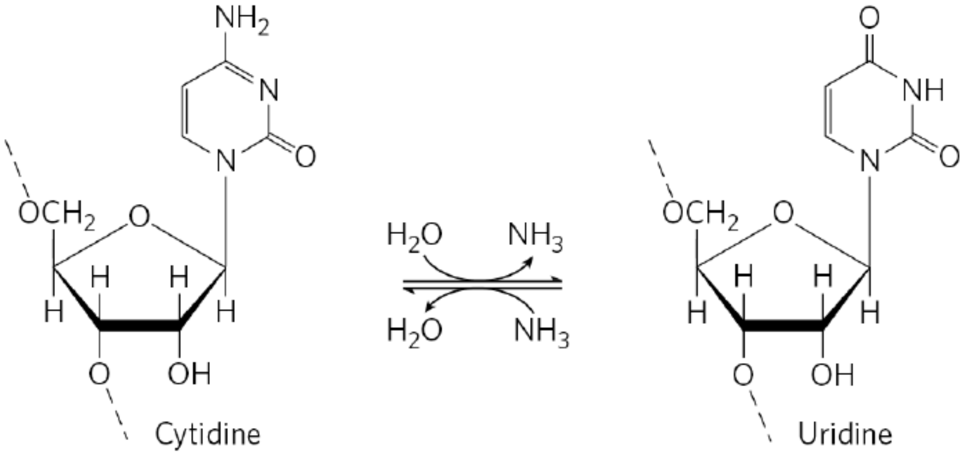
Deamination

A-to-I RNA editing



(a)

C-to-U RNA editing



(b)

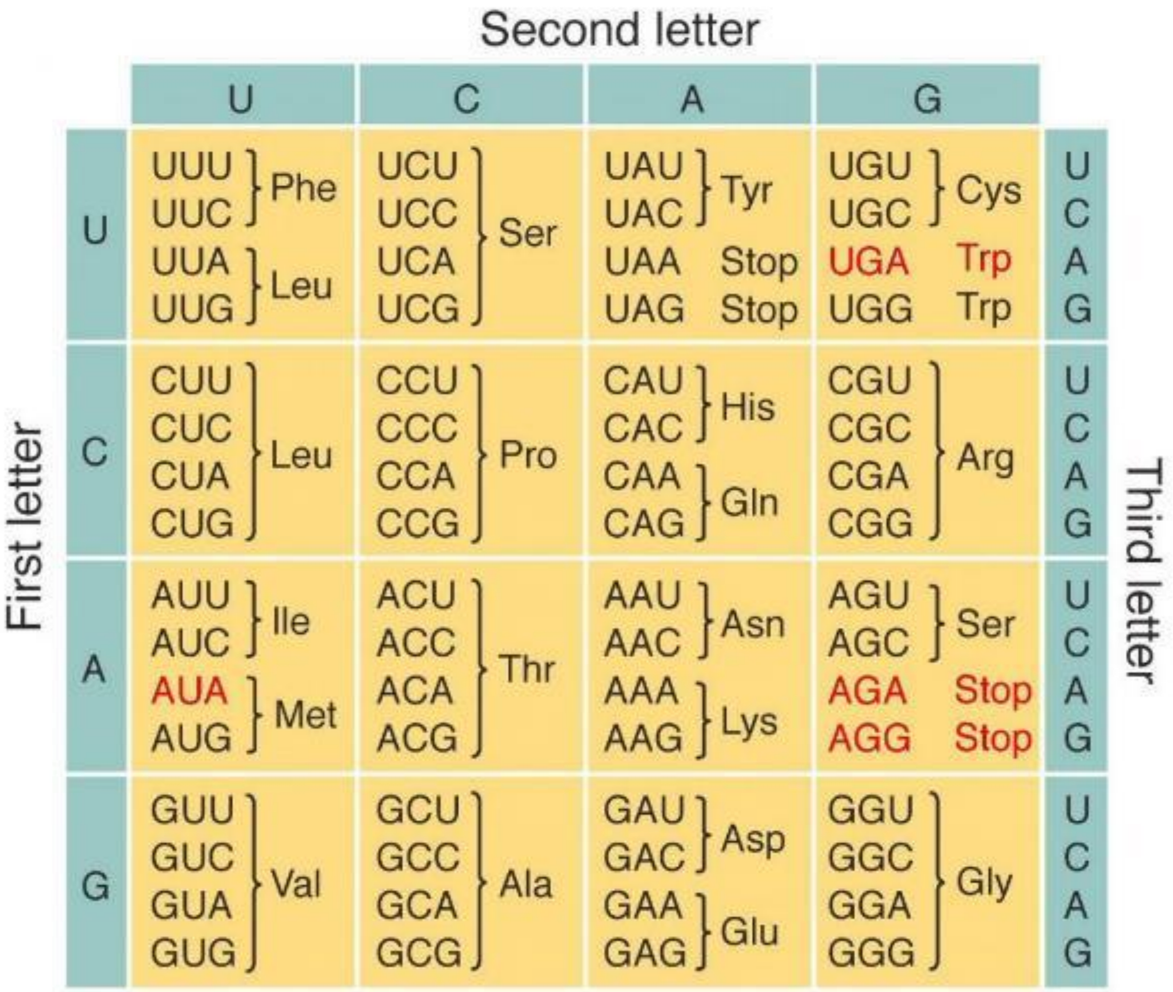
The A-to-I and C-to-U RNA editing can alter the sequences of mRNAs, it makes them different from the corresponding genomic template.

Translation

I pairs with C

U pairs with A

ADARs = Adenosine deaminases



Second letter

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

Translation

I pairs with C

U pairs with A

CAA → CIA

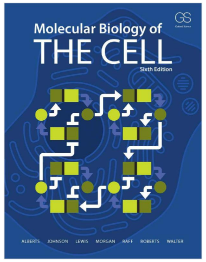
CAG → CIG

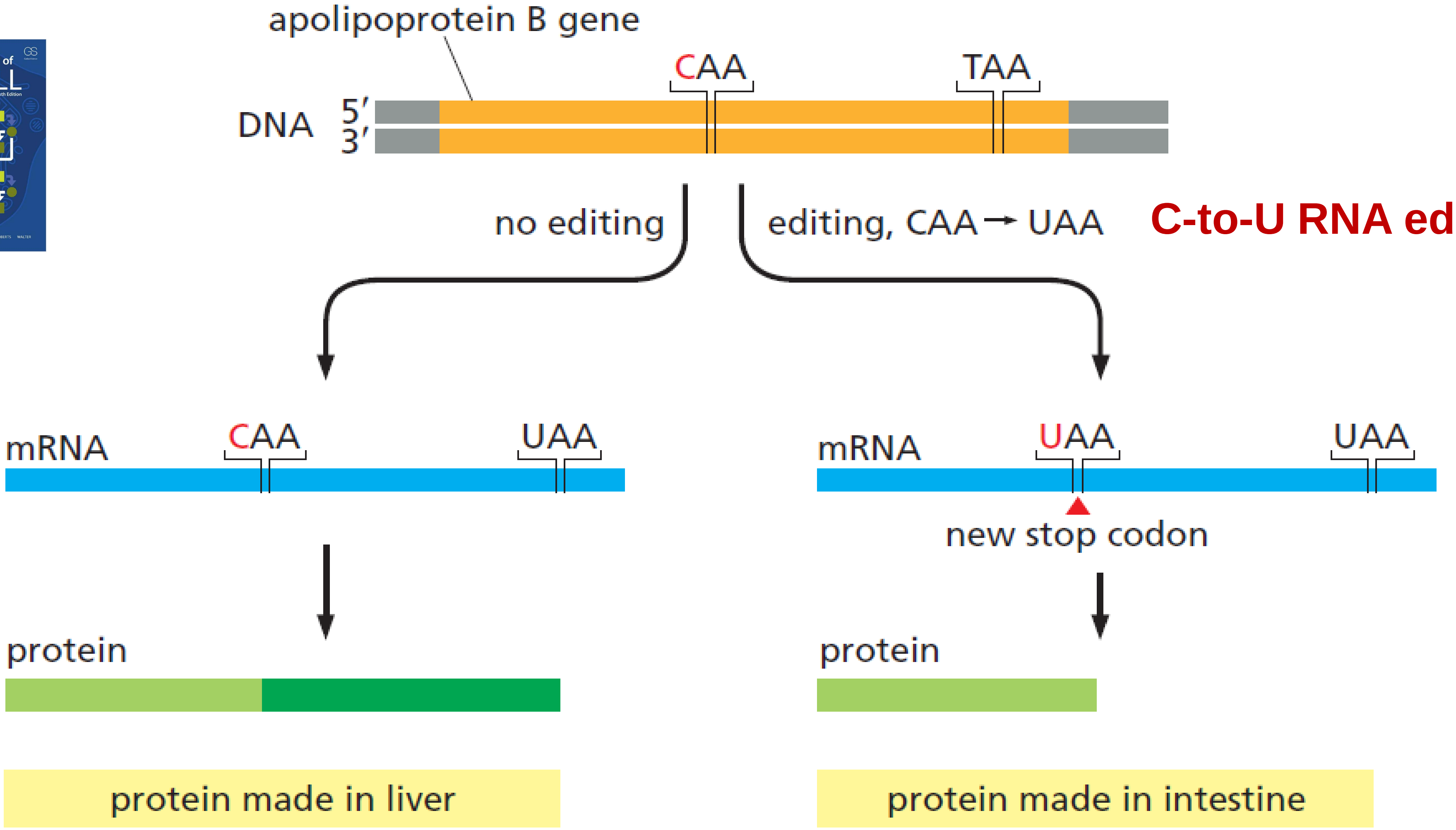
Gln → Arg

CAA → UAA

Gln → Stop

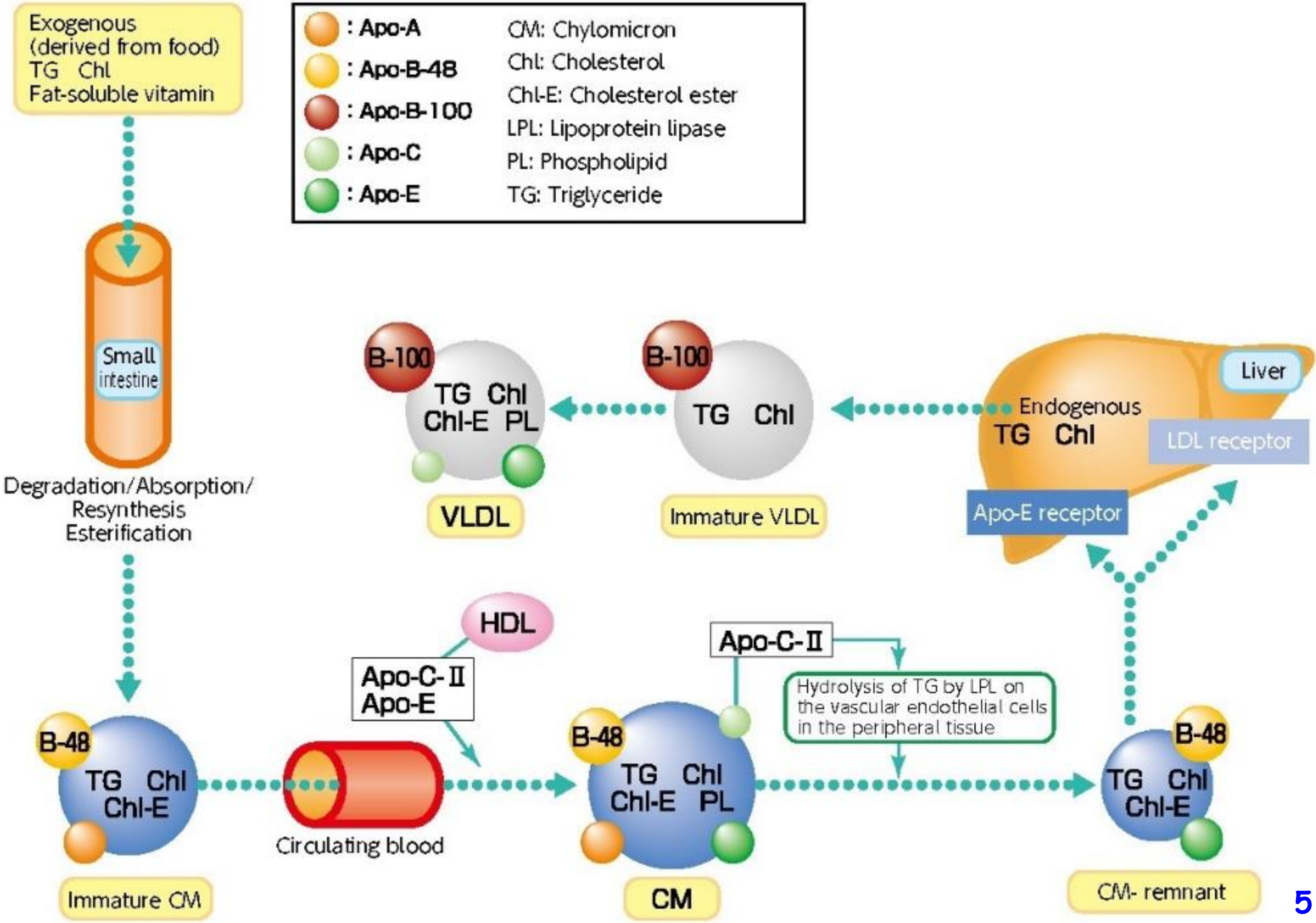
Example of C-to-U RNA editing in apolipoprotein B gene



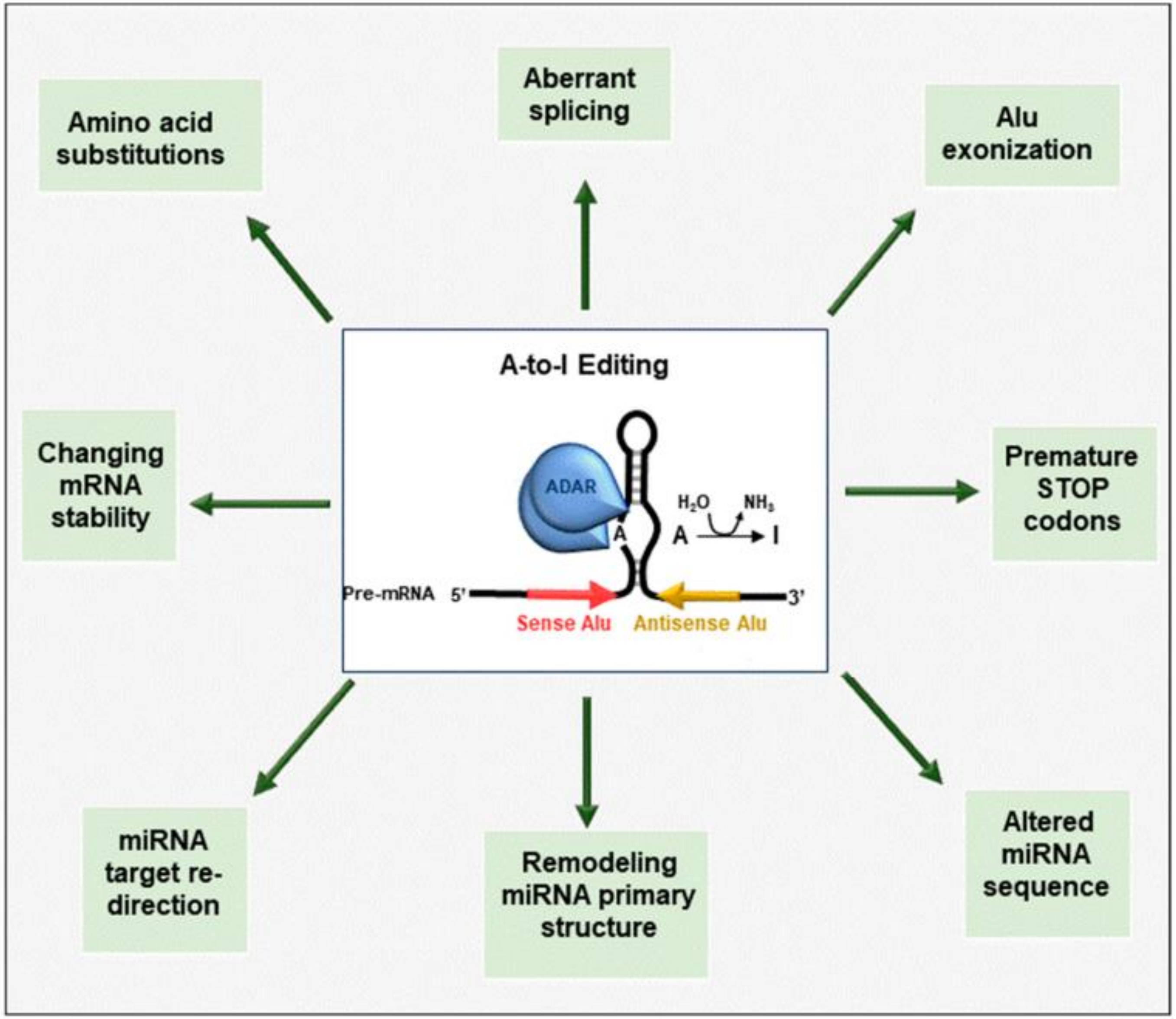


C-to-U RNA editing

“RNA Editing Can Change the Meaning of the RNA Message”

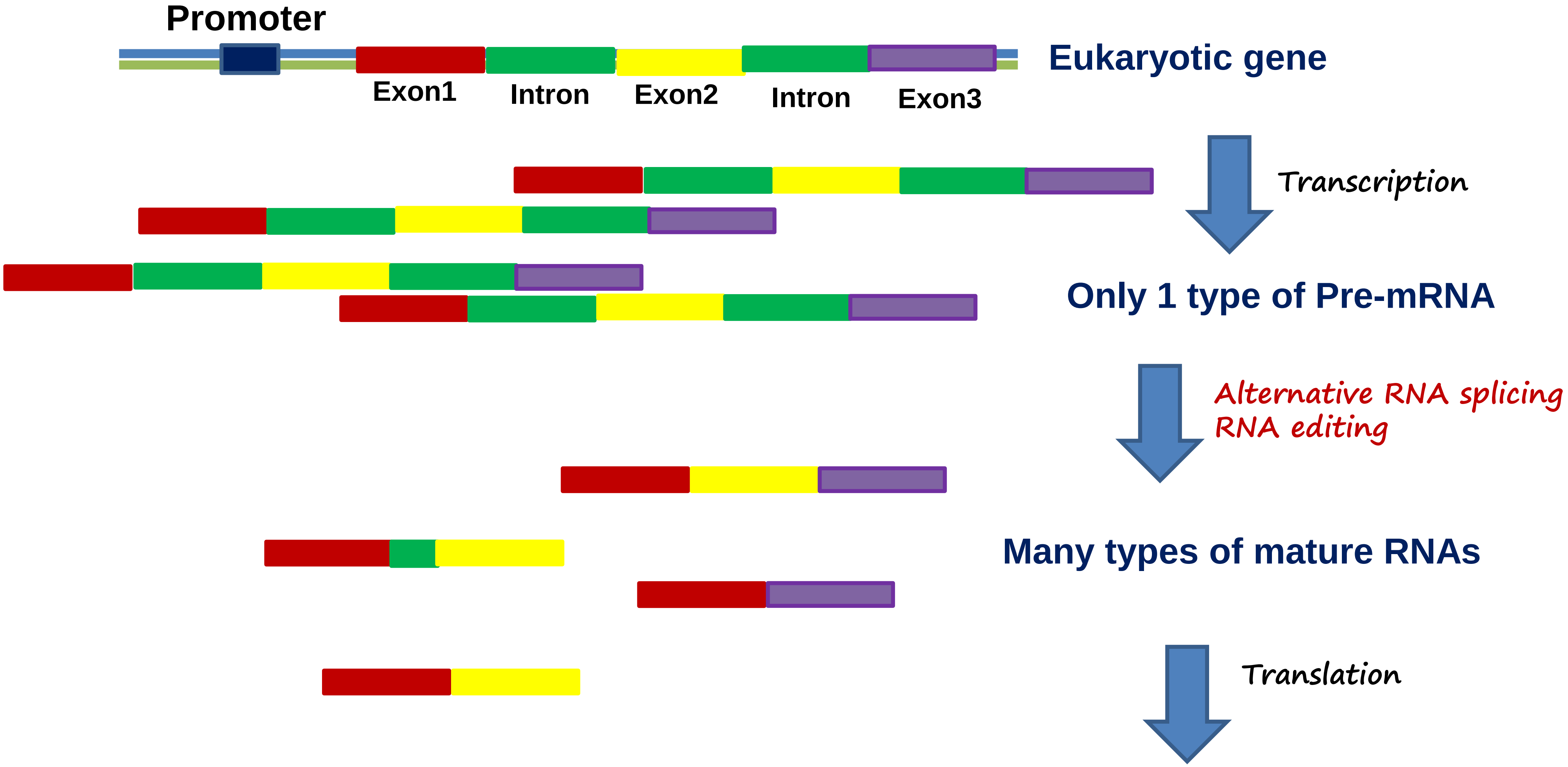

Exogenous
(derived from food)
TG Chl
Fat-soluble vitamin
● : Apo-A
● : Apo-B-48
● : Apo-B-100
● : Apo-C
● : Apo-E
CM: Chylomicron
Chl: Cholesterol
Chl-E: Cholesterol ester
LPL: Lipoprotein lipase
PL: Phospholipid
TG: Triglyceride
Small intestine
Degradation/Absorption/
Resynthesis
Esterification
B-100
TG Chl
Chl-E PL
VLDL
B-100
TG Chl
Immature VLDL
Liver
Endogenous
TG Chl
LDL receptor
Apo-E receptor
HDL
Apo-C-II
Apo-E
Apo-C-II
Hydrolysis of TG by LPL on the vascular endothelial cells in the peripheral tissue
B-48
TG Chl
Chl-E
Immature CM
Circulating blood
B-48
TG Chl
Chl-E PL
CM
B-48
TG Chl
Chl-E
CM- remnant

Effects of A-to-I RNA editing in cancer



Frezza V, et al, Cancers, 2023

Summary



Eukaryotic gene

Transcription

Only 1 type of Pre-mRNA

Alternative RNA splicing
RNA editing

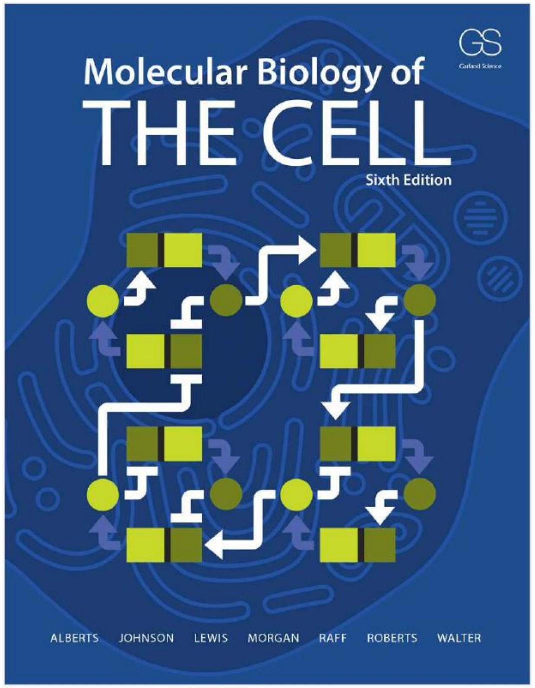
Many types of mature RNAs

Translation

Many types of proteins

State a hypothesis for why eukaryotes evolved complex RNA processing steps.

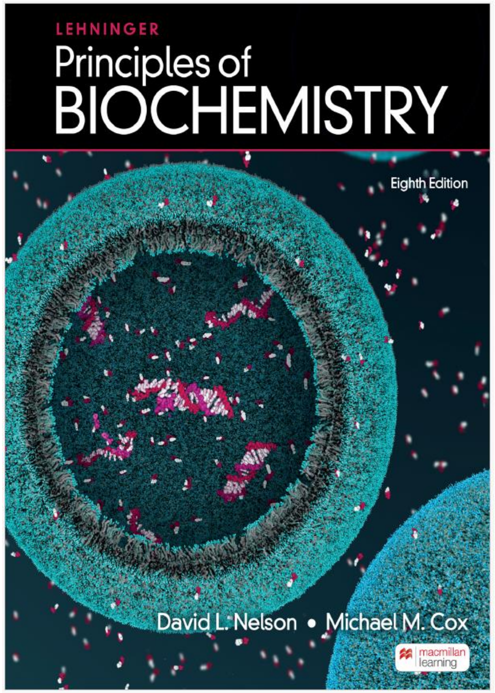
References



GS
Garland Science

Molecular Biology of
THE CELL
Sixth Edition

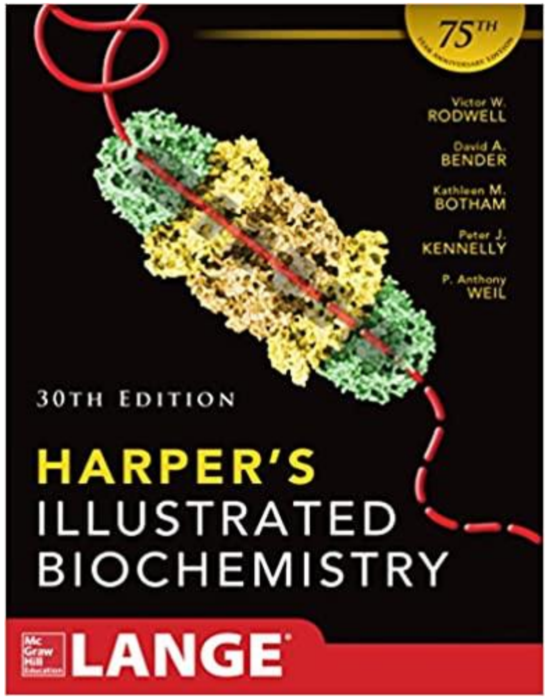
ALBERTS JOHNSON LEWIS MORGAN RAFF ROBERTS WALTER



LEHNINGER
Principles of
BIOCHEMISTRY
Eighth Edition

David L. Nelson • Michael M. Cox

macmillan learning



75TH

Victor W. RODWELL
David A. BENDER
Kathleen M. BOTHAM
Peter J. KENNELLY
P. Anthony WEIL

30TH EDITION
HARPER'S
ILLUSTRATED
BIOCHEMISTRY

Mc Graw Hill Education
LANGE