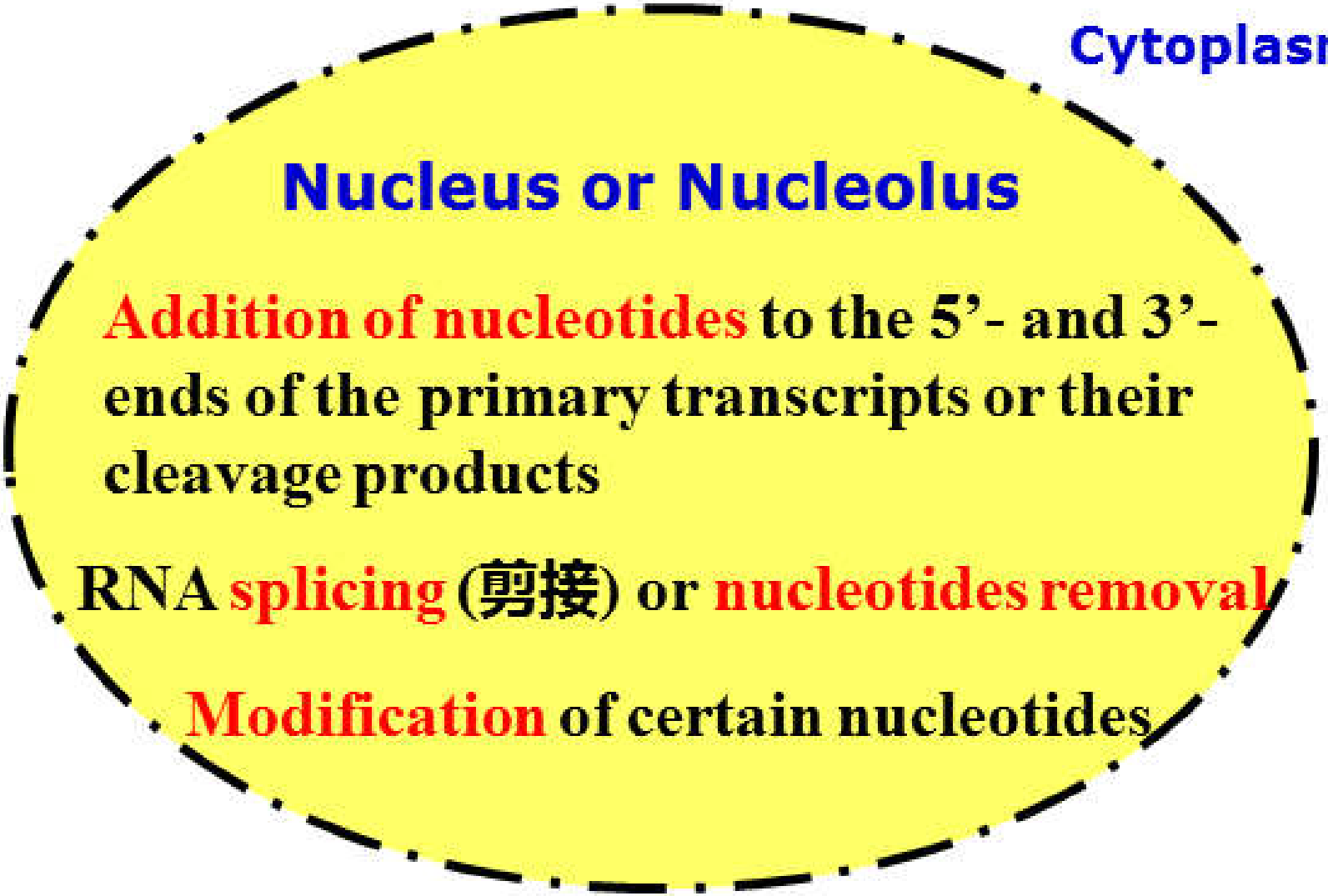
Chapter 7 RNA processing

- **RNA processing** is the collective term used to describe the molecular events allowing the primary transcripts to become the mature RNA.

RNA加工是初生转录物转变为成熟RNA过程的分子事件的总称。



加工类型

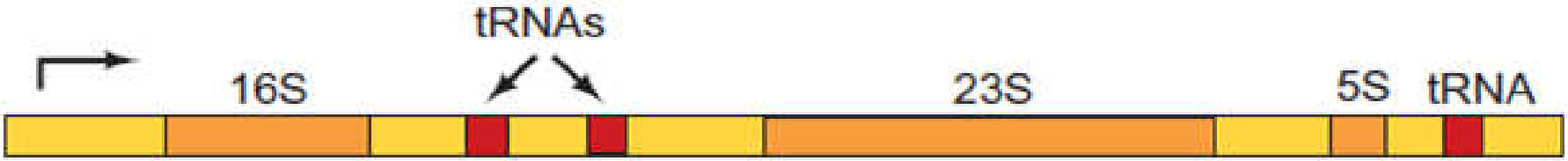
- RNA splicing is the process of cutting the pre-RNA to remove introns and joining together of the exons.

1. rRNA processing

1.1 rRNA processing in prokaryotes

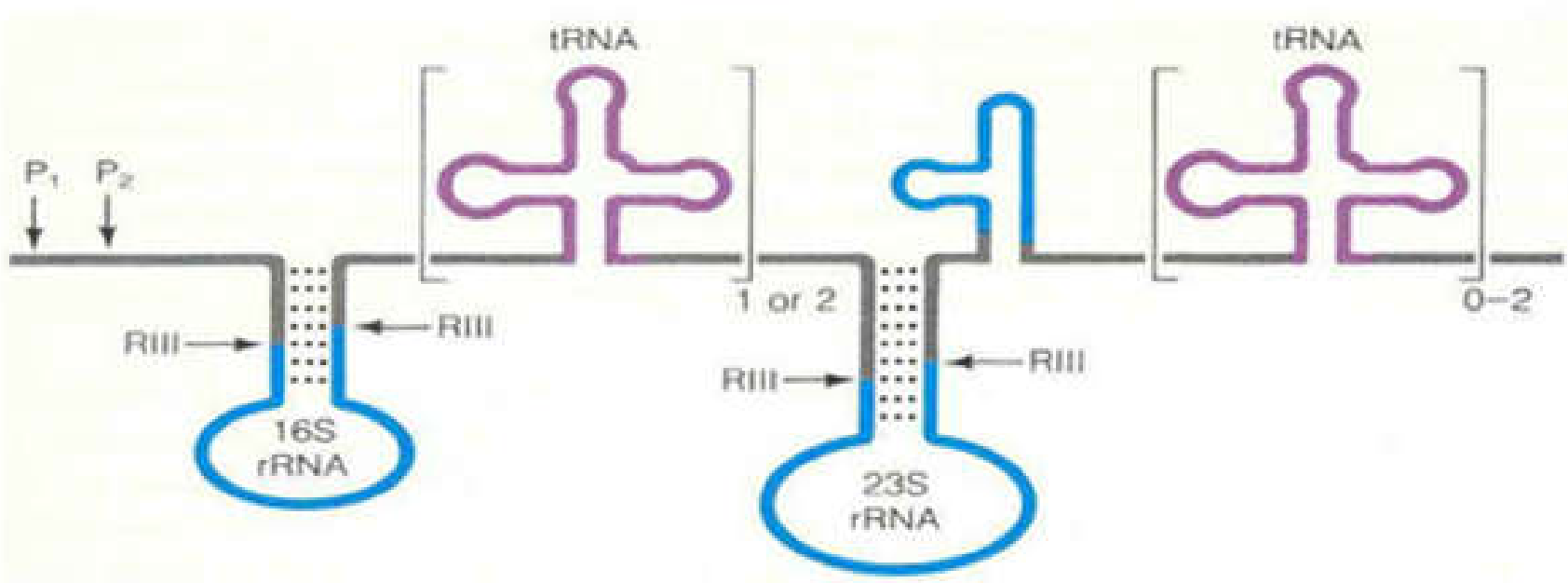
1.1.1 Bacterial rRNA gene

- There are 7 different operons (操纵子) for rRNA that are dispersed throughout the genome of *E.coli*.



- Each operon contains one copy of each of the 16S, the 23S and the 5S rRNA sequences.
- About 1~4 coding sequences for tRNA molecules are also present.

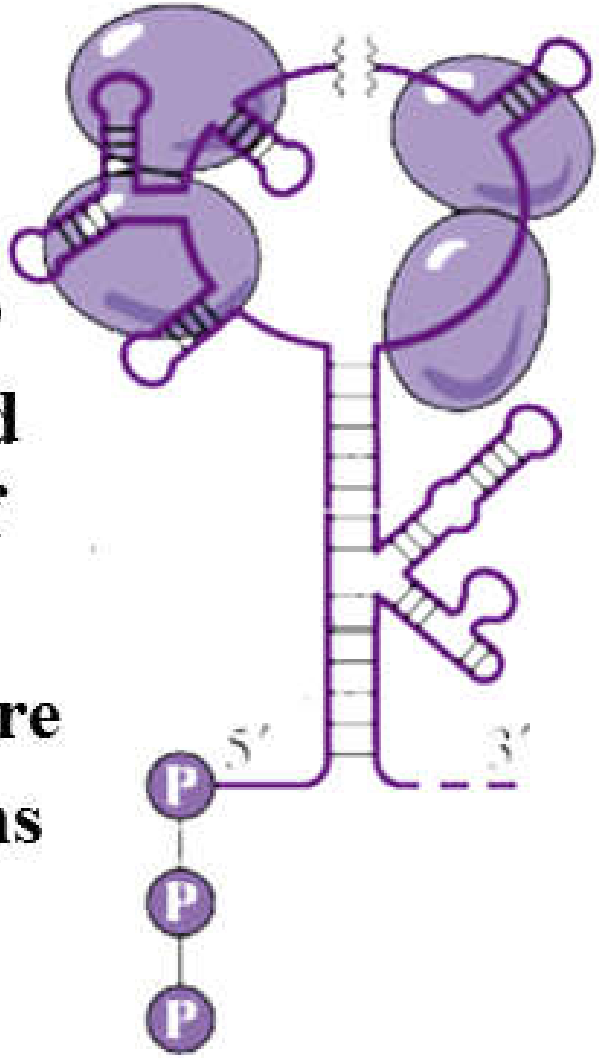
- The RNA folds up into a number of **stem-loop structures** by base pairing between complementary sequences.



1.1.2 rRNA processing steps in prokaryotes

(1) Form a RNP complex

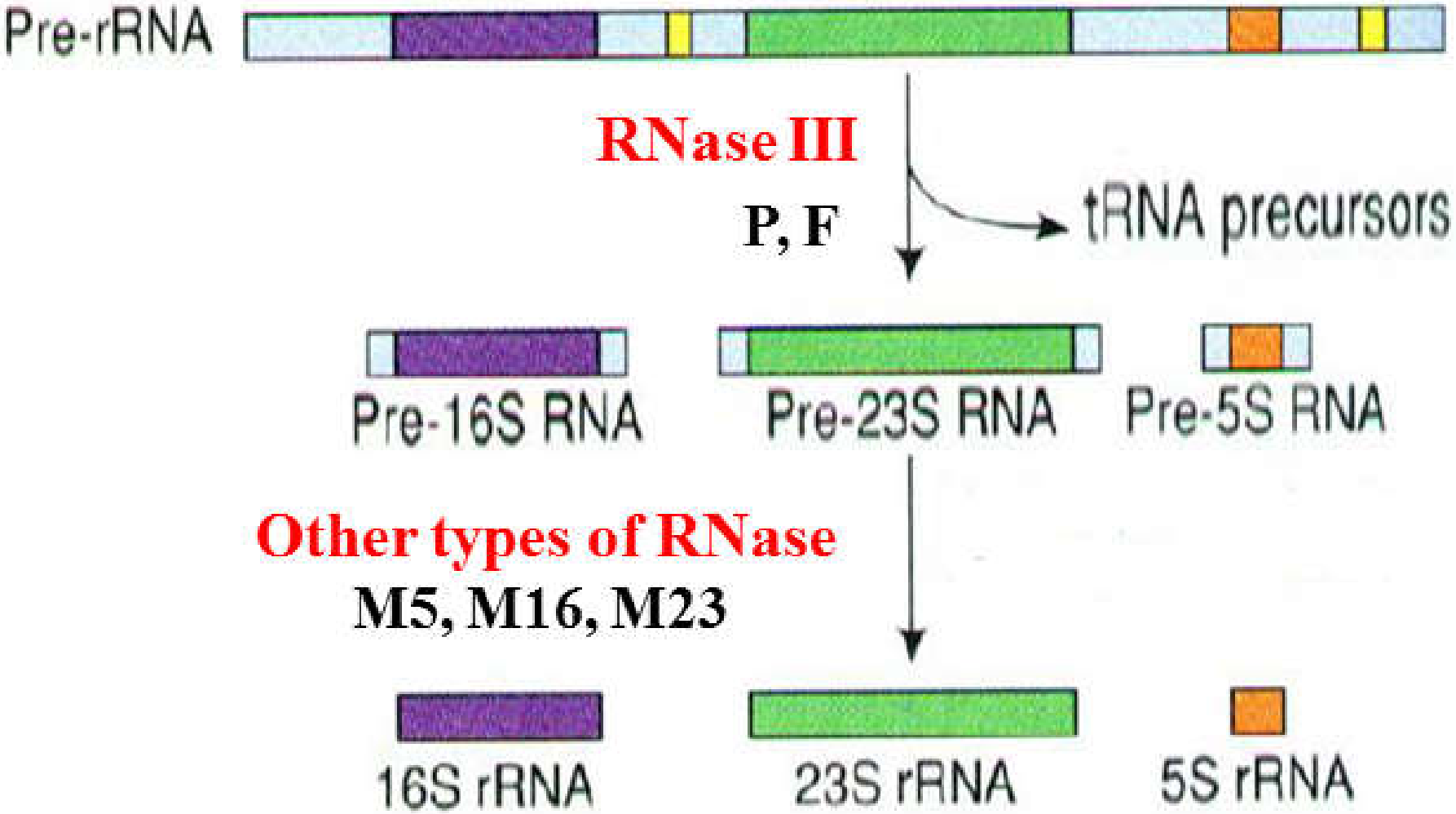
- The formation of stems and loops of rRNA allows some proteins to bind to form a RNP complex which remain attached to the RNA and become part of the ribosome.
  - ➢ Ribonucleoproteins (RNP) are complexes of specific proteins attach to specific RNAs.
    RNP = RNA + protein



(2) Nucleotide modifications

- **After the binding of proteins, nucleotide modifications take place.**
  - e.g. methylation by methylating agent S-adenosylmethonine (SAM) S-腺苷甲硫氨酸

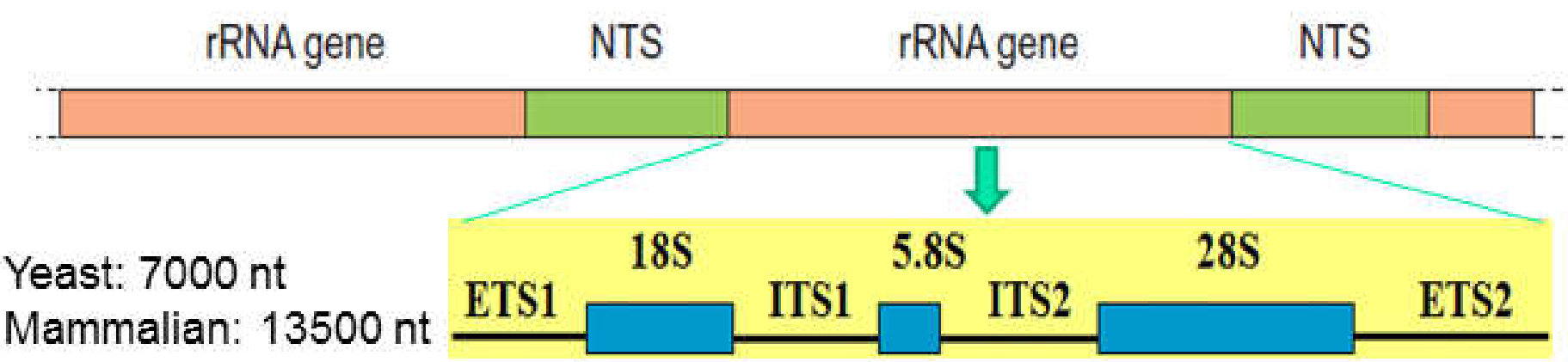
(3) RNA cleavage



1.2 rRNA processing in eukaryotes

1.2.1 Eukaryotic rRNA genes

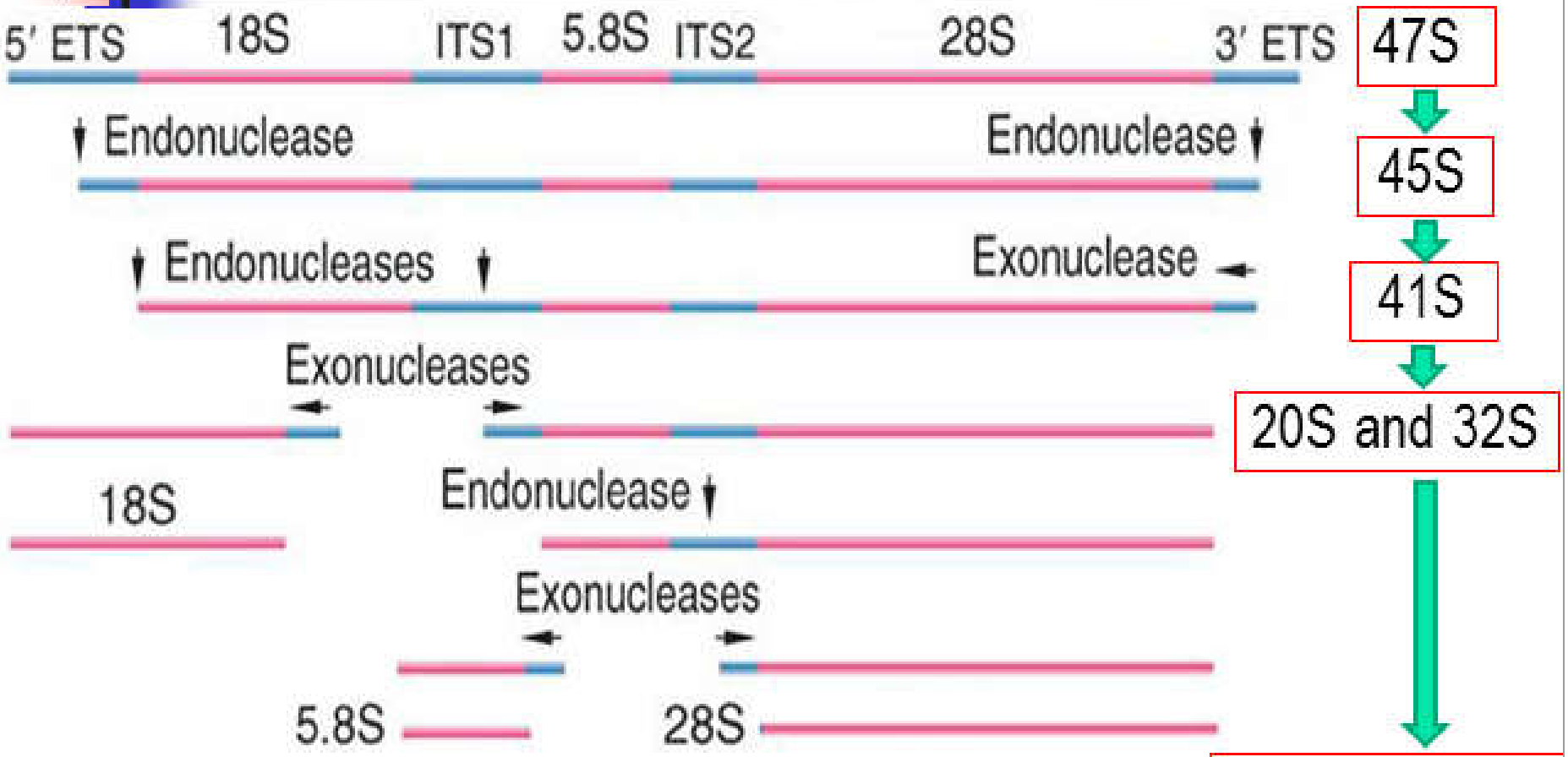
- The 5S rRNA is separately transcribed by RNA Pol III. The 5S rRNA transcript undergoes little or no processing.
- The 18S, 5.8S and 28S rRNA genes are present in a tandemly repeated cluster, and are transcribed in nucleolus by RNA Pol I.



External transcribed spacer (ETS, 外部转录间隔区)
Internal transcribed spacer (ITS, 内部转录间隔区)

1.2.2 rRNA cleavage in eukaryotes

Nucleolus



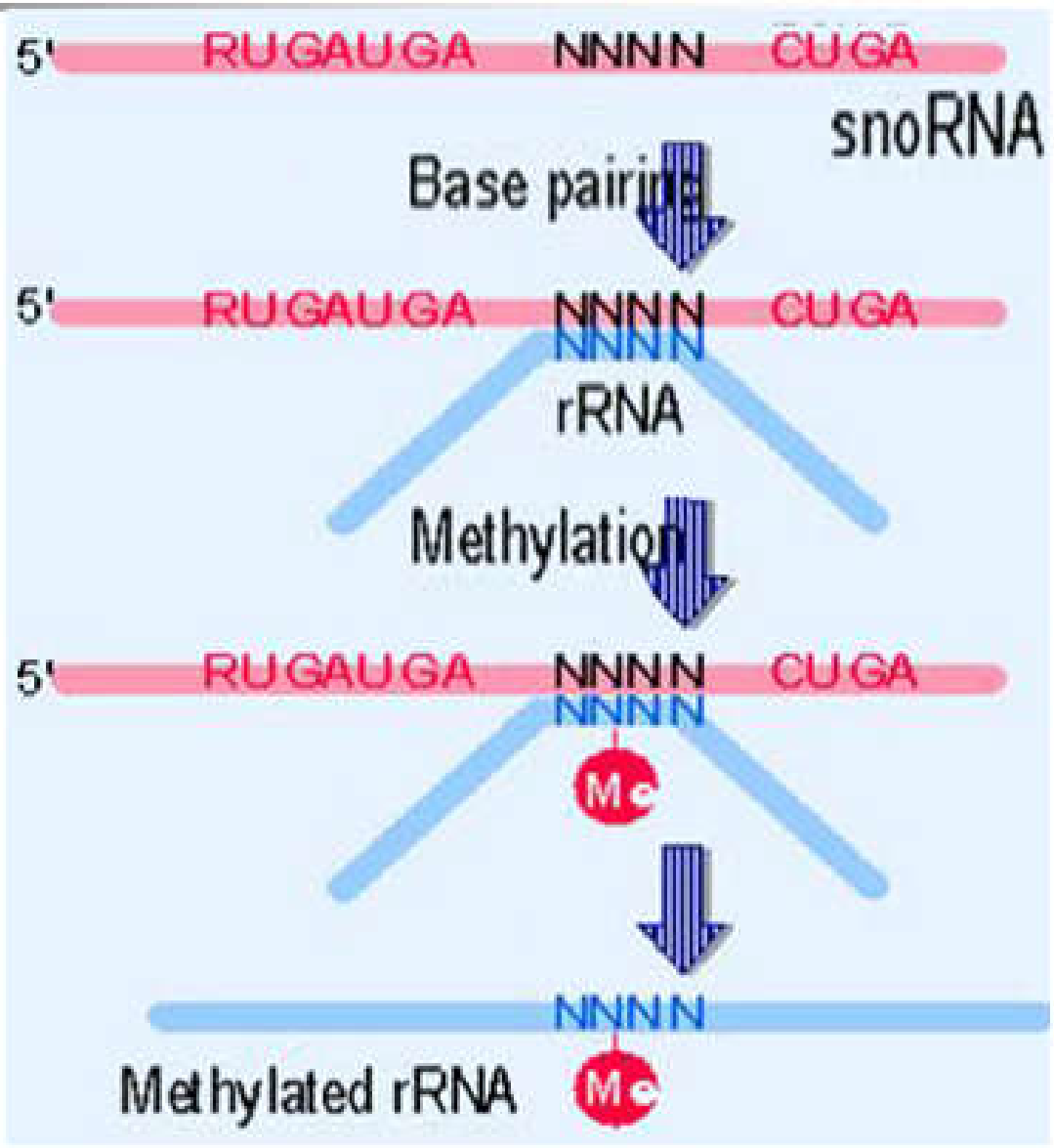
The 5.8S region must base-pair to the 28S rRNA before the mature molecules are produced.

1.2.3 Modification carried out by snoRNP

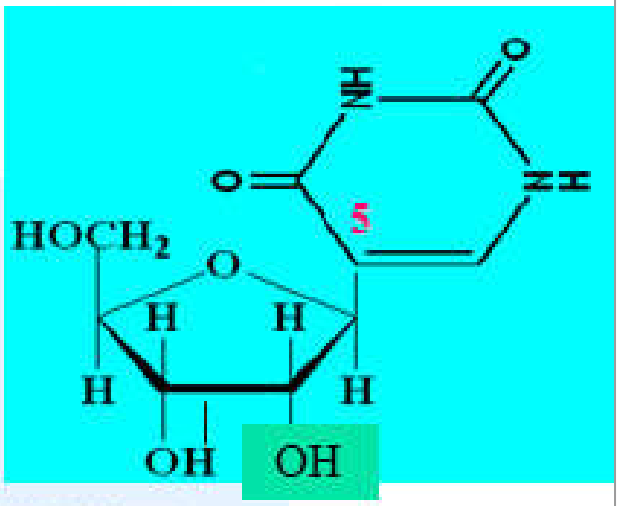
snoRNP (小核仁核糖核蛋白体)= snoRNA (小核仁RNA) + proteins

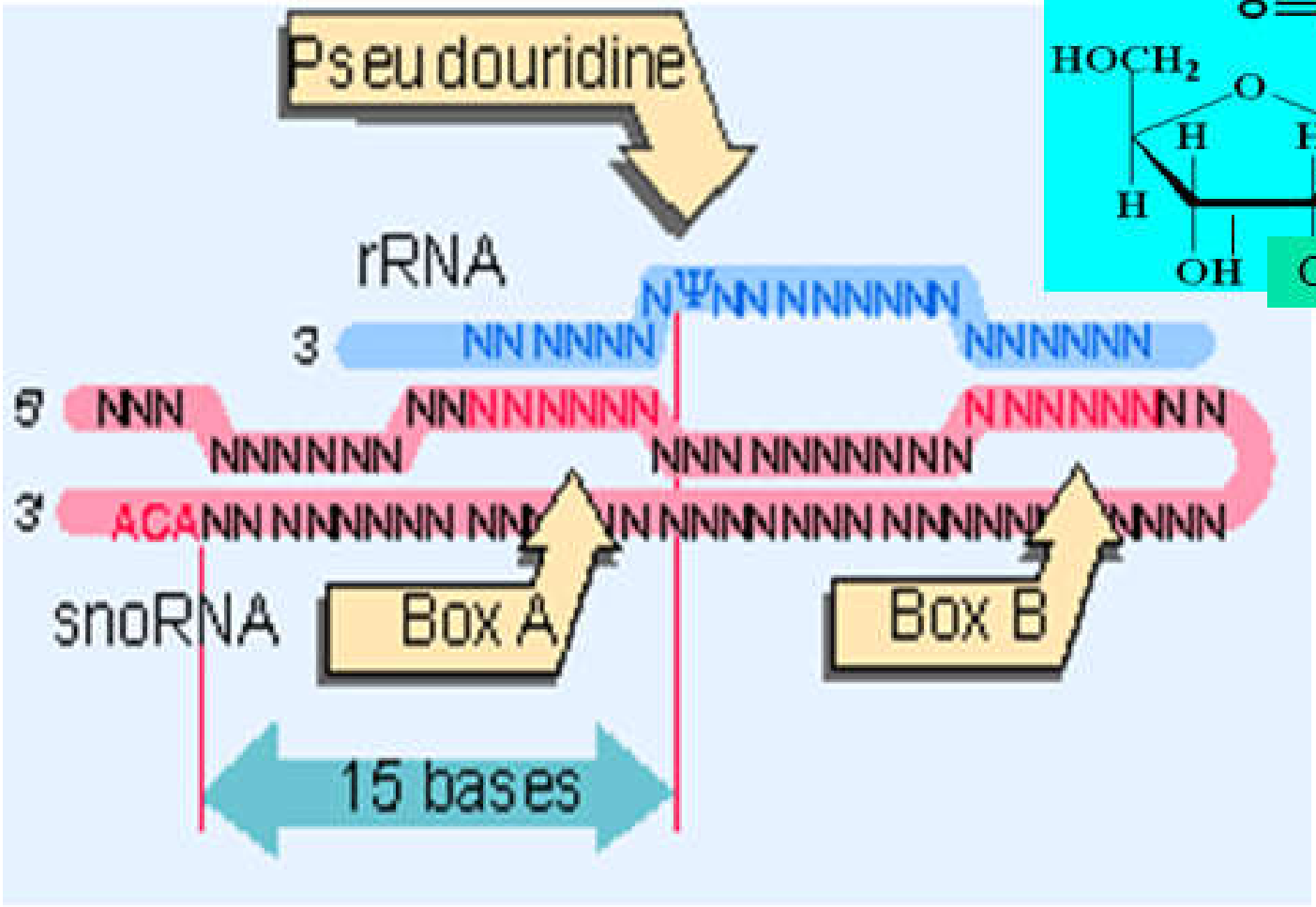
A snoRNA base pairs with a region of rRNA that is to be methylated — define methylation sites

2'-O-methylribose



Pseudouridine (假尿嘧啶) modification





2. tRNA processing

- Overly long precursors. Processed by removing nucleotides at both ends and base modifications.
- **Prokaryotes:** a precursor may contain one or more tRNAs, and sometimes a mixture of rRNAs and tRNAs.
- **Eukaryotes:** each tRNA precursor contains a single tRNA **with an intron.** No CCA at 3'-end.

2.1 tRNA processing in prokaryotes

RNase III

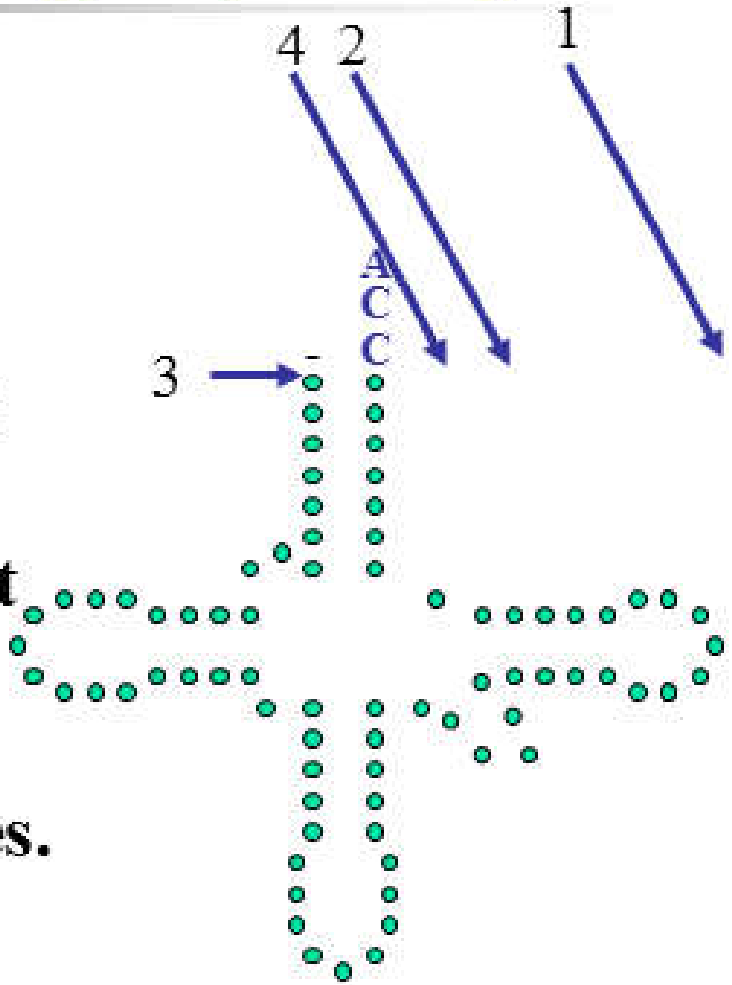
(1) Endo-RNase E or F cleave 3'-end;

(2) Exo-RNase D trims (修剪) the 3'-end to within 2 nt of mature length;

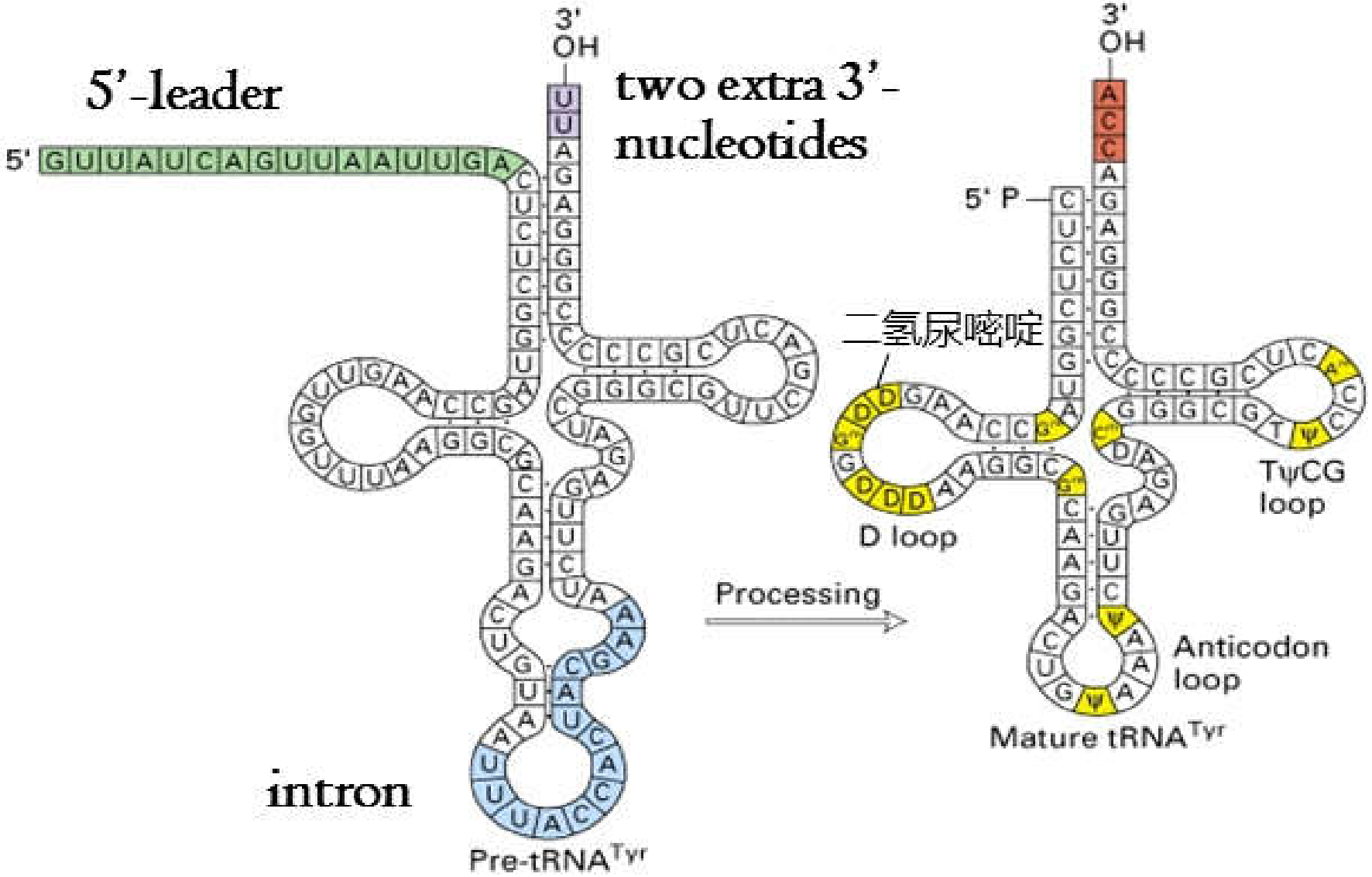
(3) Endo-RNase P can then cut to give the mature 5'-end;

(4) Exo-RNase D finally removes the two 3'-residues.

(5) Base modifications: are unique to each particular tRNA type.



2.2 tRNA processing in eukaryotes

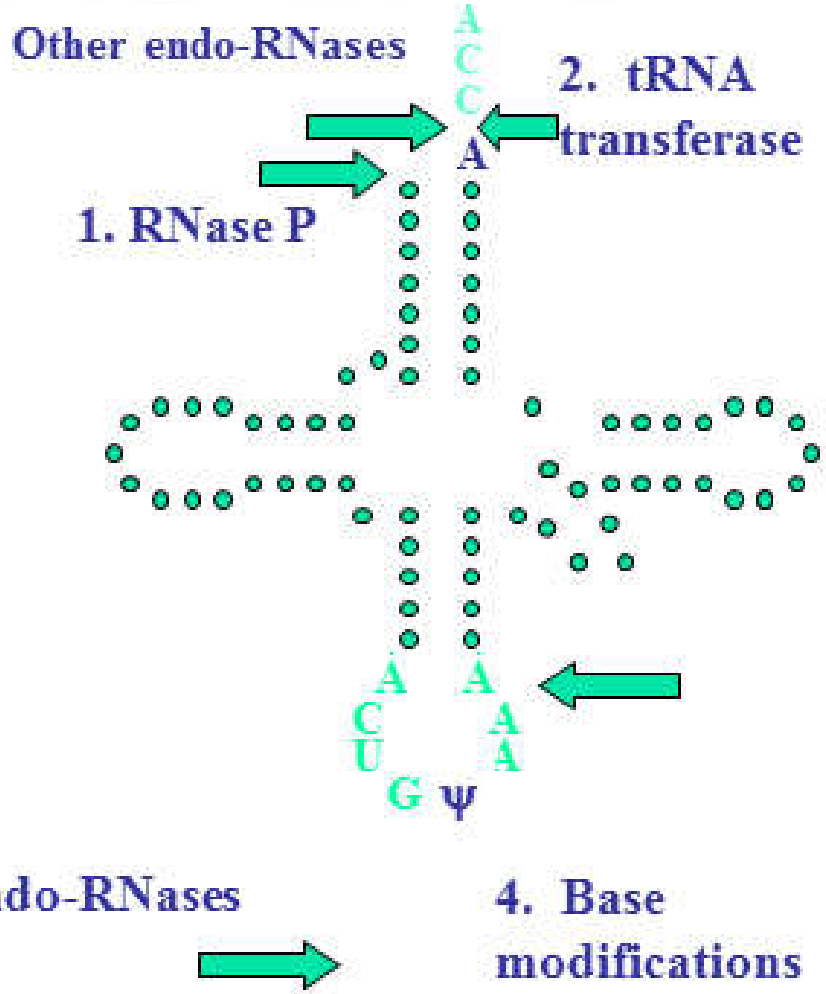


(1) Specific cleavages by RNase P for 5' leader and other endo-RNases for 3' extra nucleotide removal

(2) tRNA nucleotidyl transferase (核苷酸转移酶) adds CCA to the 3'-end to generate the mature 3'-end.

(3) Intron removal

(4) Base modifications



2.3 RNase P

- Ribonuclease P (RNase P) is an enzyme involved in tRNA processing that removes the 5' leader sequences from tRNA precursors.
- RNase P are found in both prokaryotes and eukaryotes, being located in the nucleus of the latter where they are small nuclear RNPs (snRNPs).

Sidney Altman
Yale University, USA

- **RNase P is an endonuclease composed of one RNA molecule (ribozyme 核酶, higher Mg^{2+} required in vitro) and one protein molecule (helps to catalyze).**

The 1989 Nobel Prize in chemistry for the discovery of catalytic properties of RNA.

Thomas Cech
University of Colorado, USA

?

3. mRNA processing

- There is essentially no processing of prokaryotic mRNA, it can start to be translated before it has finished being transcribed.
- Prokaryotic mRNA is degraded rapidly from the 5' end (short-lived).

3.1 hnRNP and snRNP

3.1.1 hnRNA and pre-mRNA

- **Heterogeneous nuclear RNA (hnRNA, 核内不均一RNA): RNA that comprises transcripts of nuclear genes made by RNA Pol II; it has a wide size distribution.**
- **Among hnRNA, those processed to give mature mRNAs are called pre-mRNAs. hnRNA≠pre-mRNA**
- **hnRNAs are not exported until processing is complete; thus they are found only in the nucleus.**

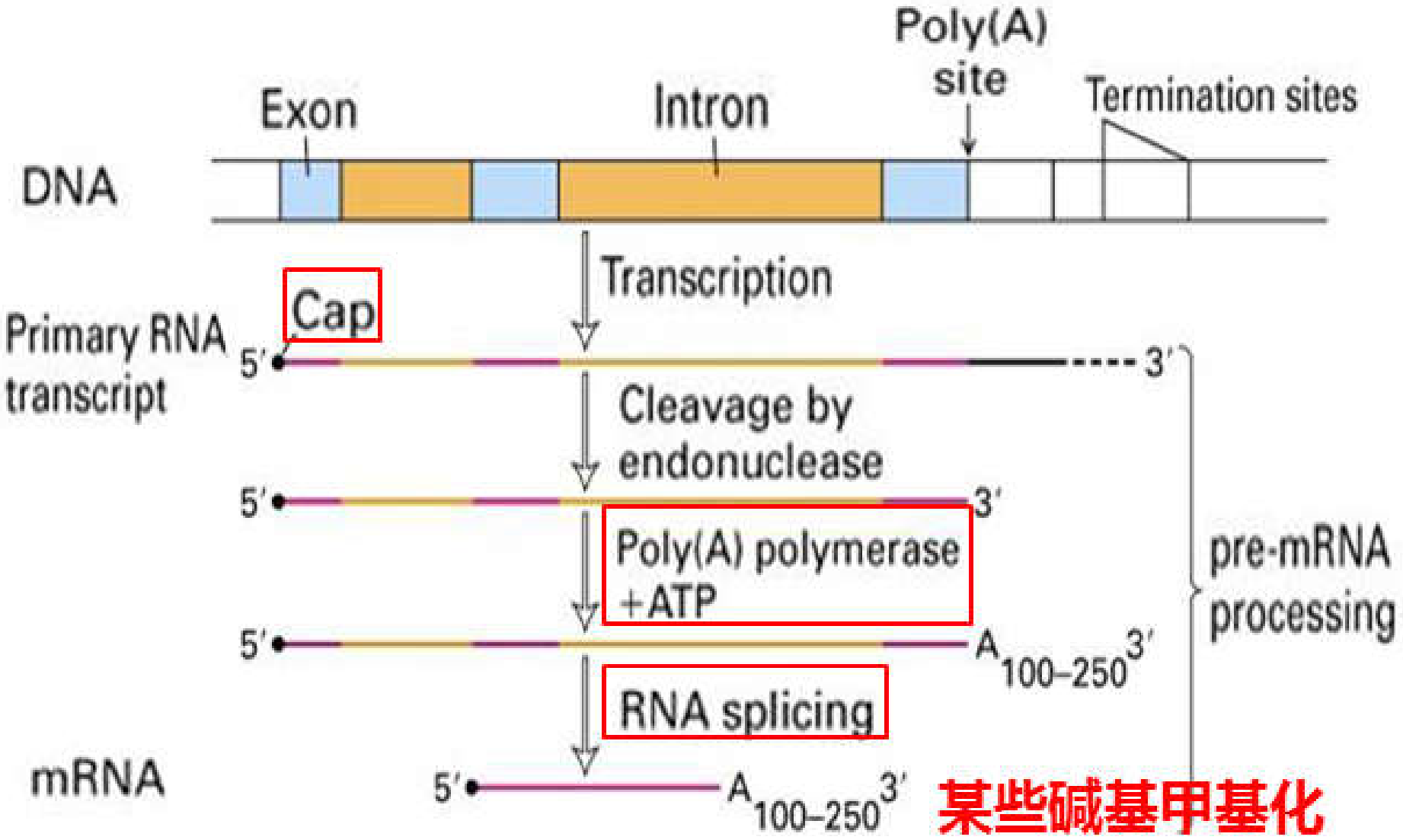
3.1.2 hnRNP

- **hnRNP = hnRNA + specific proteins.**
- **hnRNP proteins are thought to**
  - **Help keep the hnRNA in a single-stranded form**
  - **Assist in the various RNA processing reactions**

3.1.3 snRNP

- Small nuclear RNA (snRNA) are rich in the base uracil (尿嘧啶), which complex with specific proteins to form snRNPs.
- Several of snRNP are involved in splicing or other RNA processing reactions (define modification sites in pre-rRNA).
- The most abundant snRNP are involved in pre-mRNA splicing, U1,U2,U4,U5 and U6.
- snRNAs synthesized in the nucleus by RNA Pol II have a normal 5'-cap.

5’端加帽，3’端加poly (A)尾，剪接去掉内含子



3.2 5' Capping

3.2.1 Structure of the cap

7-methylguanosine

5', 5'-triphosphate linkage

7-甲基鸟苷三磷酸

Cap 0 (0型):

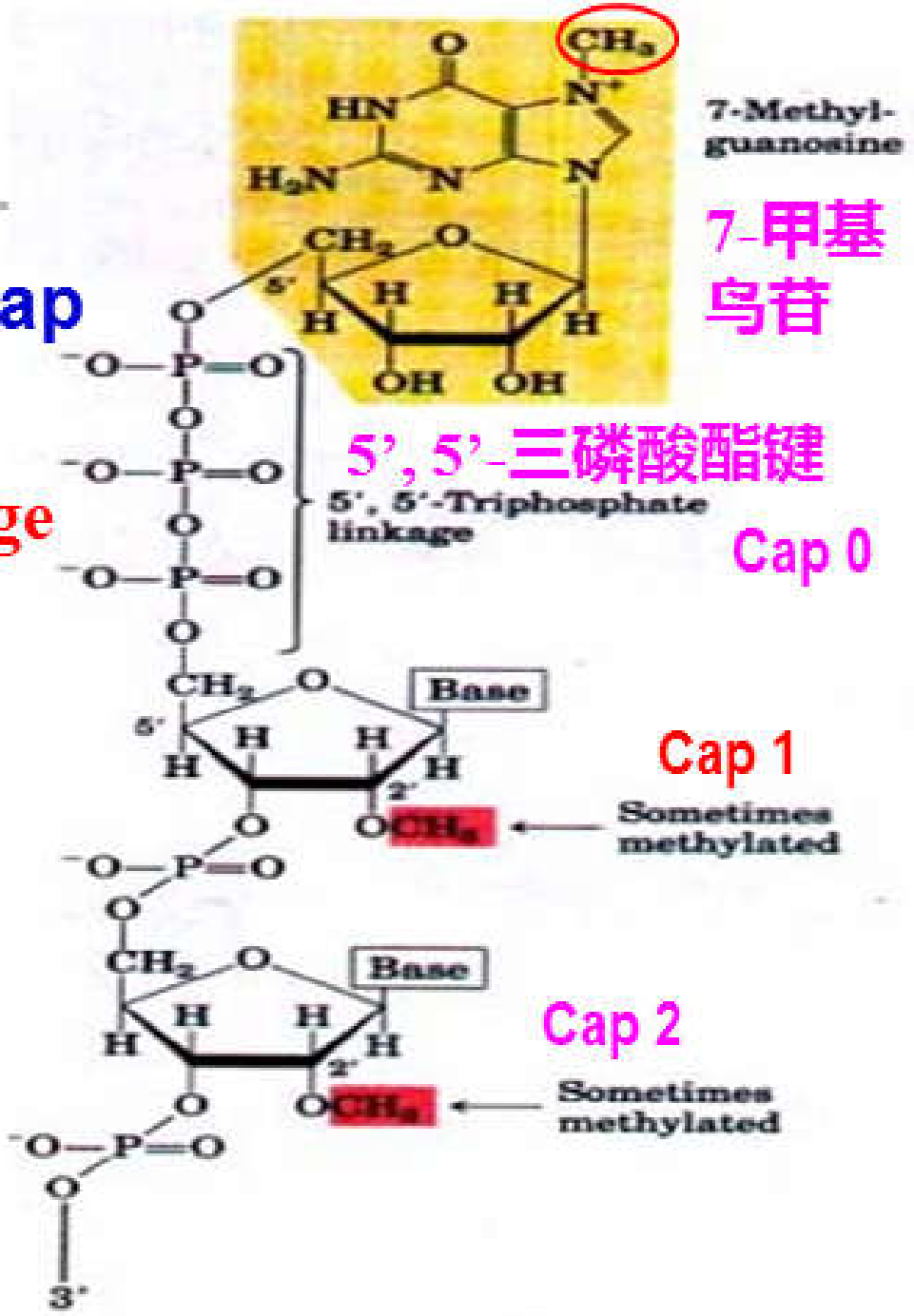
$m^7G^{5'}ppp^{5'}Xp$

Cap 1 (I型):

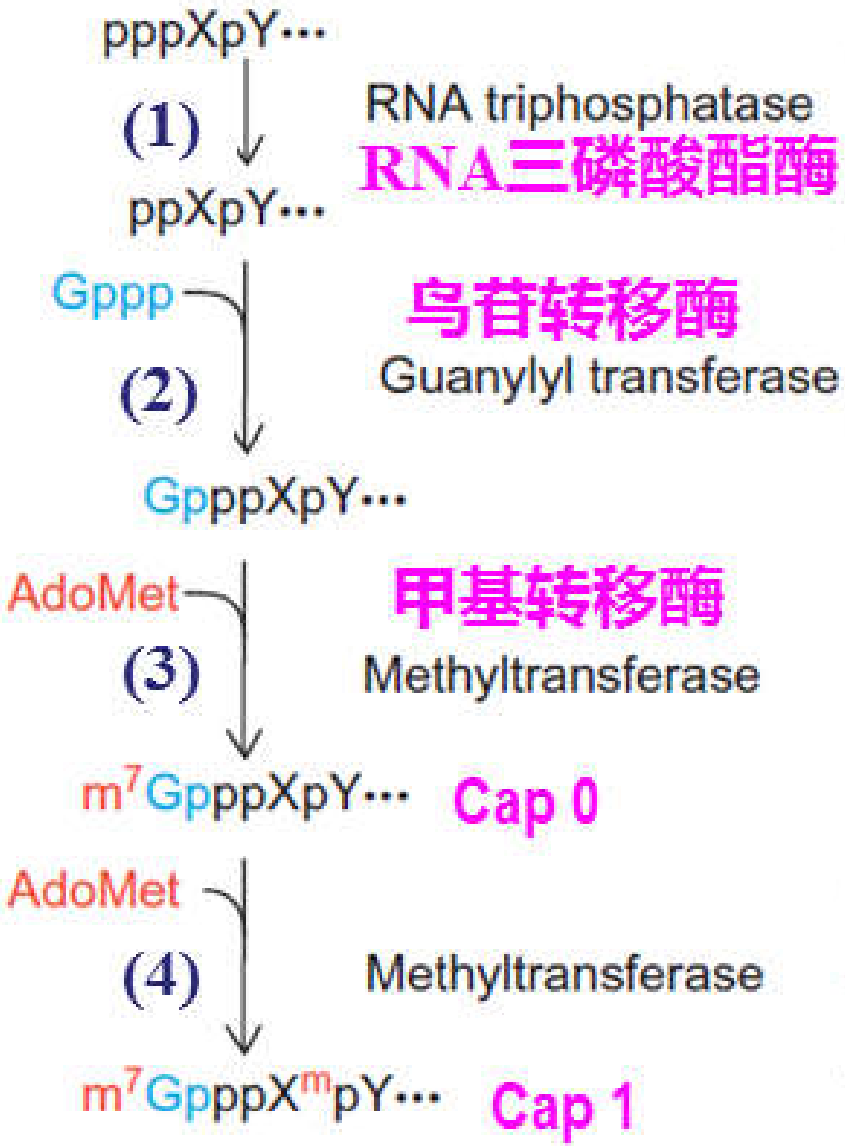
$m^7G^{5'}ppp^{5'}X^mpYp$

Cap 2 (II型):

$m^7G^{5'}ppp^{5'}X^mpY^mp$



3.2.2 Synthesis of the cap



(1) **RNA三磷酸酯酶**切去pre-mRNA 5'端γ-磷酸基;

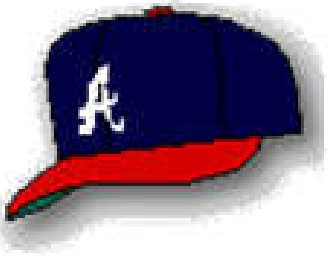
(2) **鸟苷转移酶**在末端加上GMP (来源于GTP);

(3) **甲基转移酶**从S-腺苷Met上转移甲基到G的7号N上;

(4) 另一个**甲基转移酶**将倒数第二(和三)个核苷酸2'-OH甲基化。

Capping occur early in the transcription process, before the chain length reaches 30nt.

3.2.3 Functions of the cap



(1) **Protection** of the mRNA from degradation;

(2) Enhancement of the mRNA's **translatability**;

(3) **Transport** of at least some RNAs out of the nucleus. e.g. mRNA, U1 snRNA made by RNA Pol II.

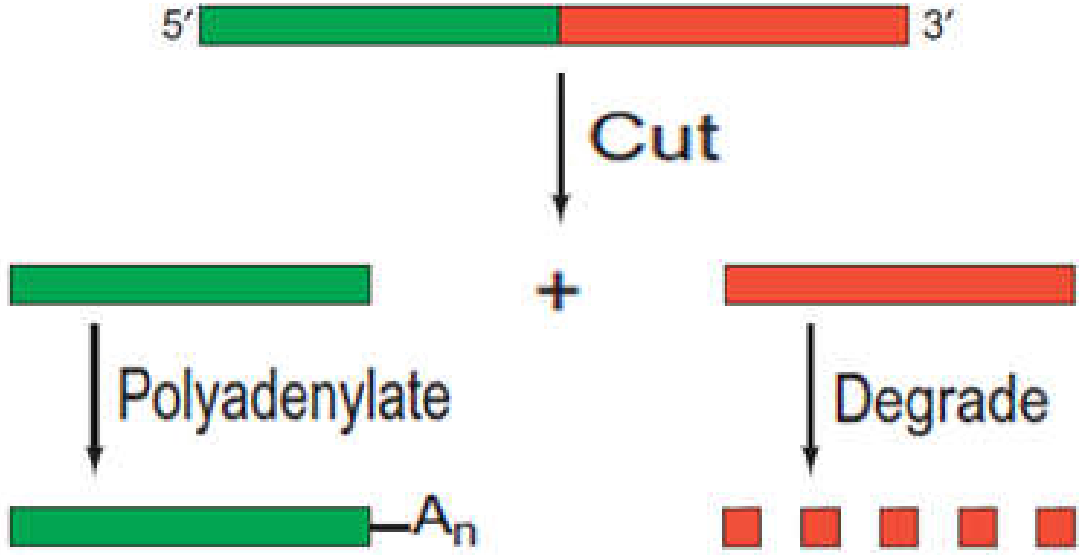
(4) **Assist splicing** of the pre-mRNA.

3.3 3' polyadenylation

- The process of adding poly(A) to RNA is called polyadenylation (聚腺苷酸化).
- Most eukaryotic mRNAs and their precursors have a chain of AMP residues about 250 nt long at their 3'-ends.
- Neither rRNA nor tRNA has a poly(A) tail.
- The poly(A) is added by poly(A) polymerase (PAP, 聚腺苷酸聚合酶).

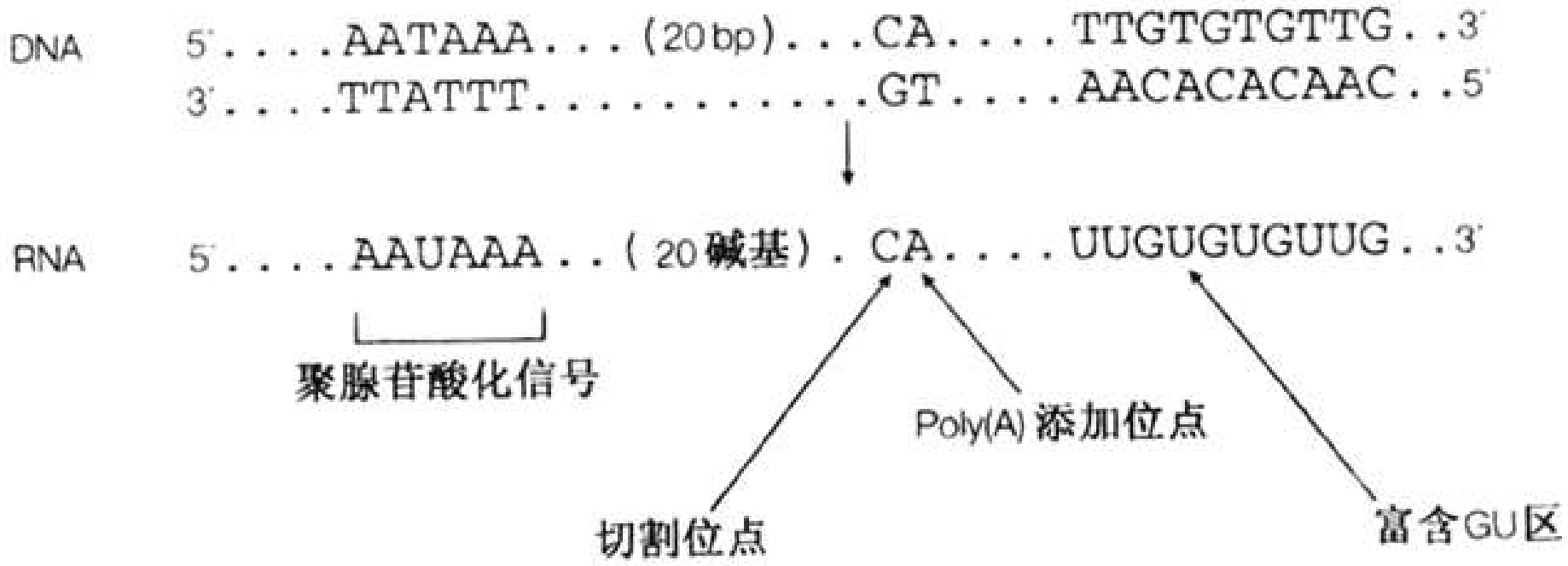
3.3.1 Basic mechanism of polyadenylation

- **Cleaving the mRNA transcript, which may actually still be in the process of being made.**
- **The PAP adds ~250 AMPs to the 3'-end of the mRNA.**
- **Degradation of the extra RNA**



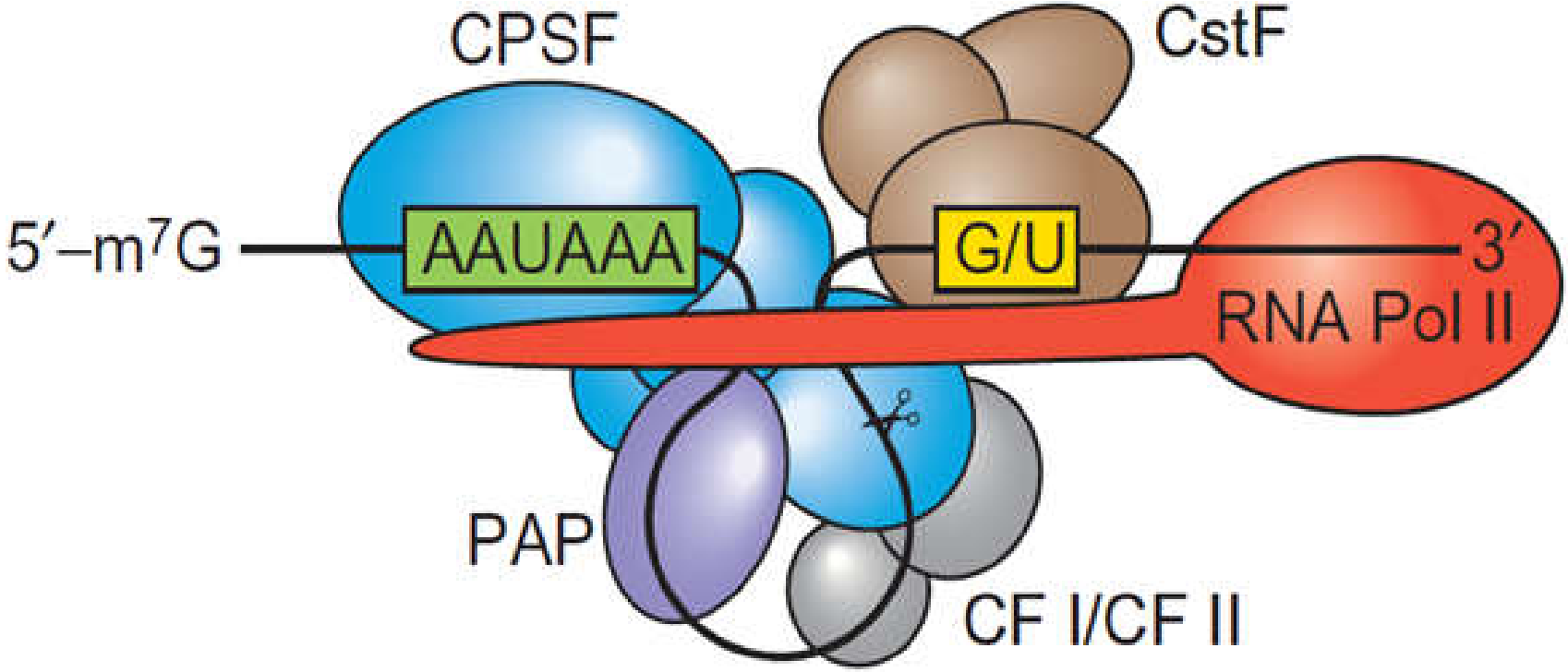
3.3.2 Signal of polyadenylation

- **AAUAAA motif about 20 nt upstream of a polyadenylation site**
- **GU-rich motif about 23 or 24 nt downstream of a polyadenylation site**



Cleavage and polyadenylation specificity factors
剪切和多腺苷酸化特异性因子

Cleavage stimulation factors
剪切激活因子



剪切复合体模型

Cleavage factors
剪切因子

3.3.3 Functions of Poly (A)

Most mRNAs in eukaryotes contain poly (A).
One noteworthy exception is the histone mRNAs.

(1) Protection of mRNA

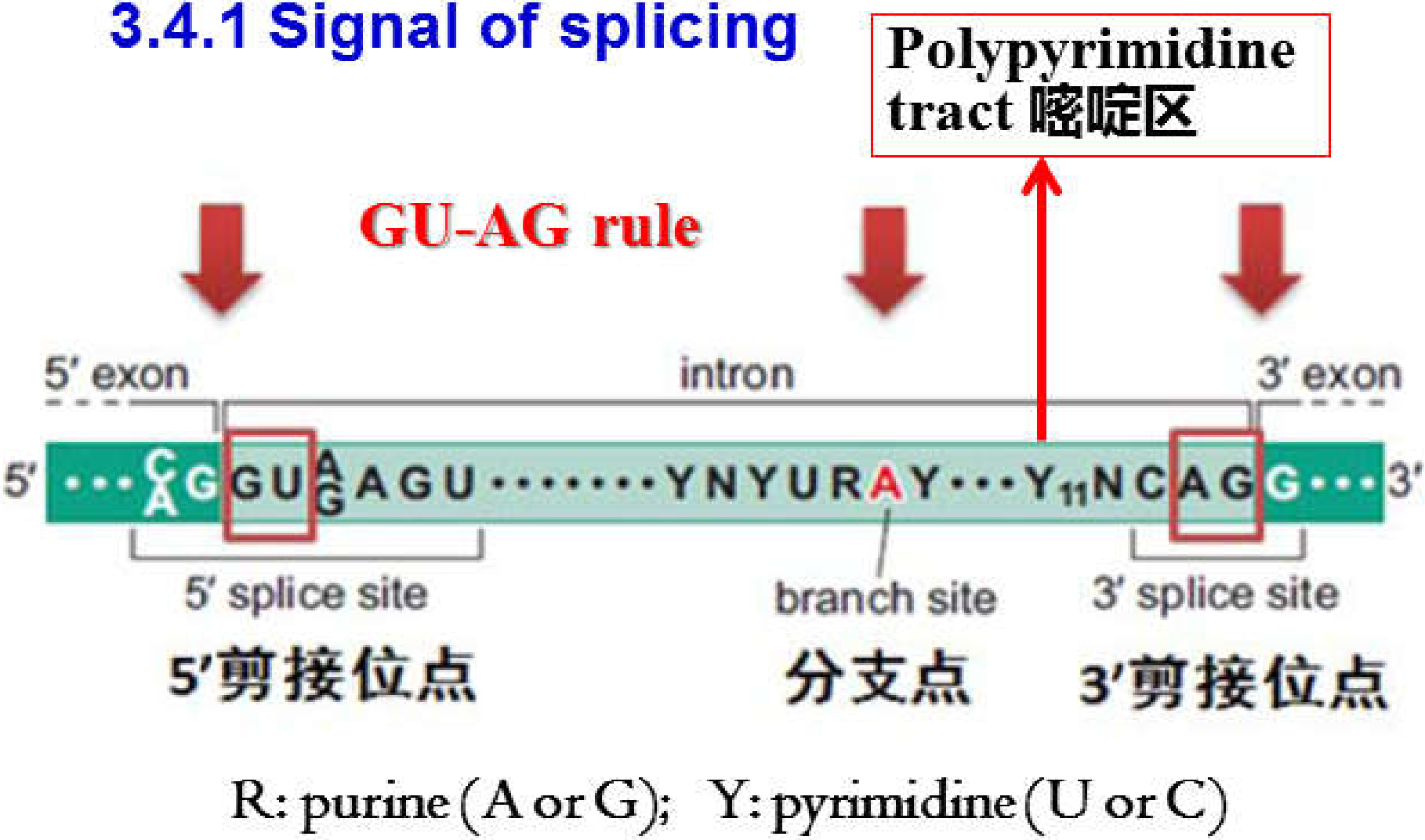
(2) Enhance translatability of mRNA

(3) Assist splicing of the pre-mRNA.

(4) Related to transcriptional termination

3.4 mRNA splicing

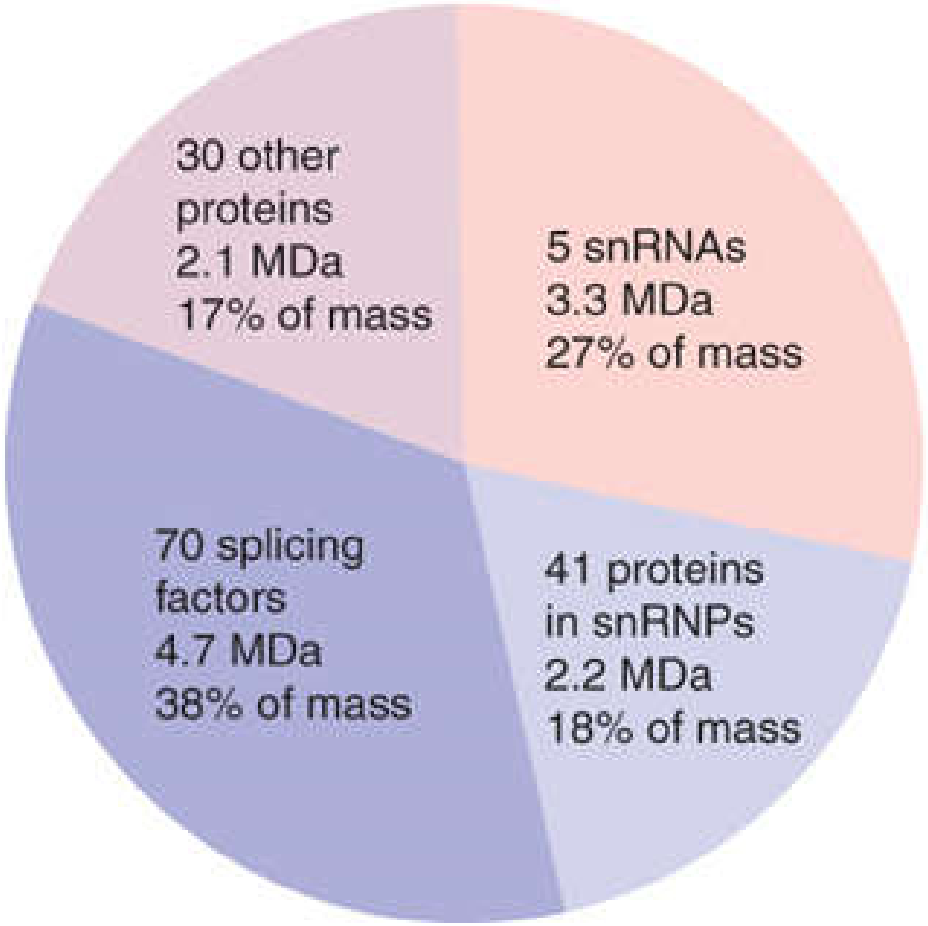
3.4.1 Signal of splicing



R: purine (A or G); Y: pyrimidine (U or C)

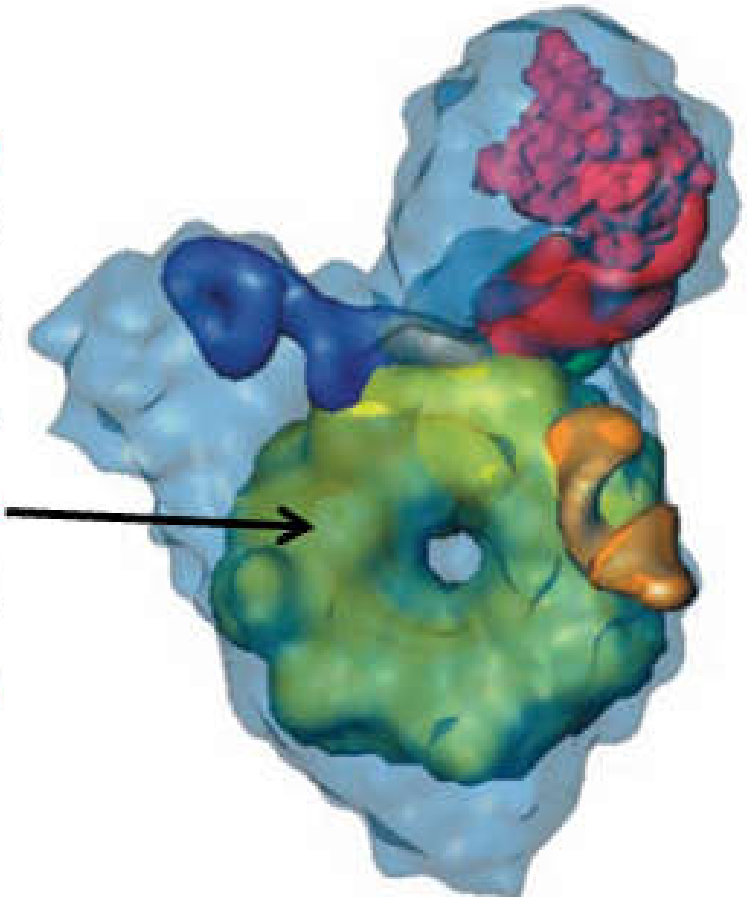
3.4.2 mechanism of pre-mRNA splicing

- Splicing takes place on a particle called a spliceosome (剪接体).
- Spliceosomes contain the pre-mRNA, as well as snRNPs and protein splicing factors that recognize key splicing signals.

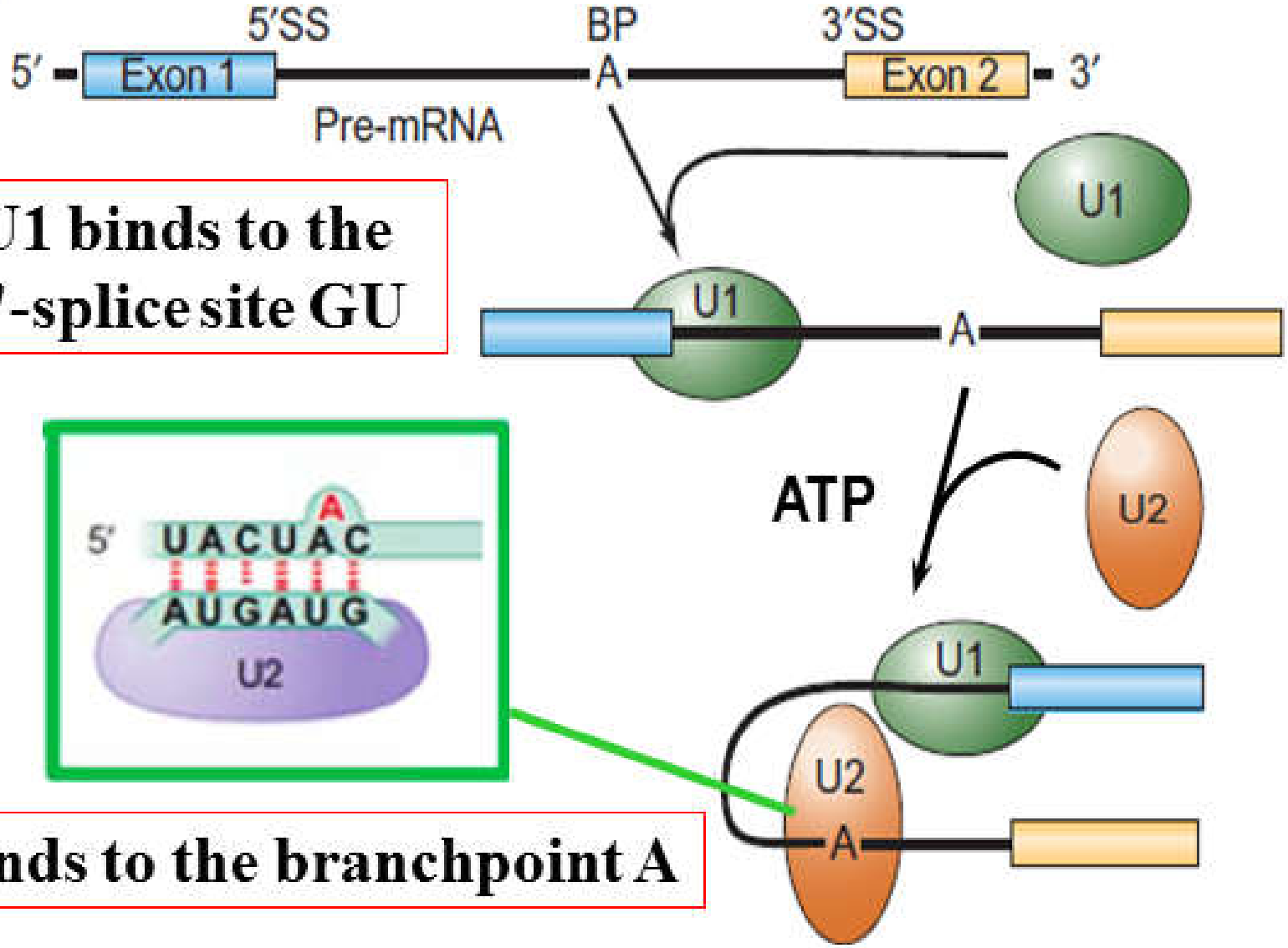


The spliceosome is a large particle.

- The five snRNPs (U1, U2, U4, U5, U6) that participate in splicing all contain a common set of seven Sm proteins and several specific proteins to each snRNP.



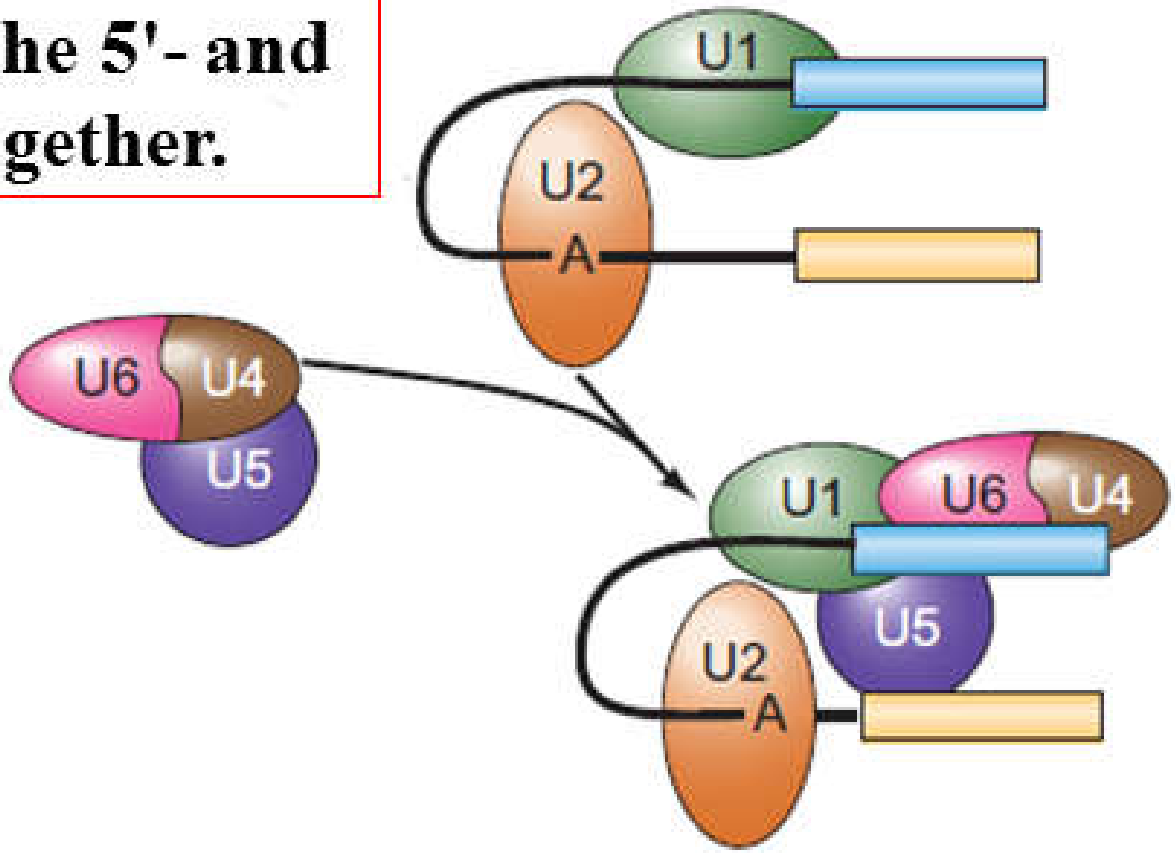
(1) Formation of spliceosome



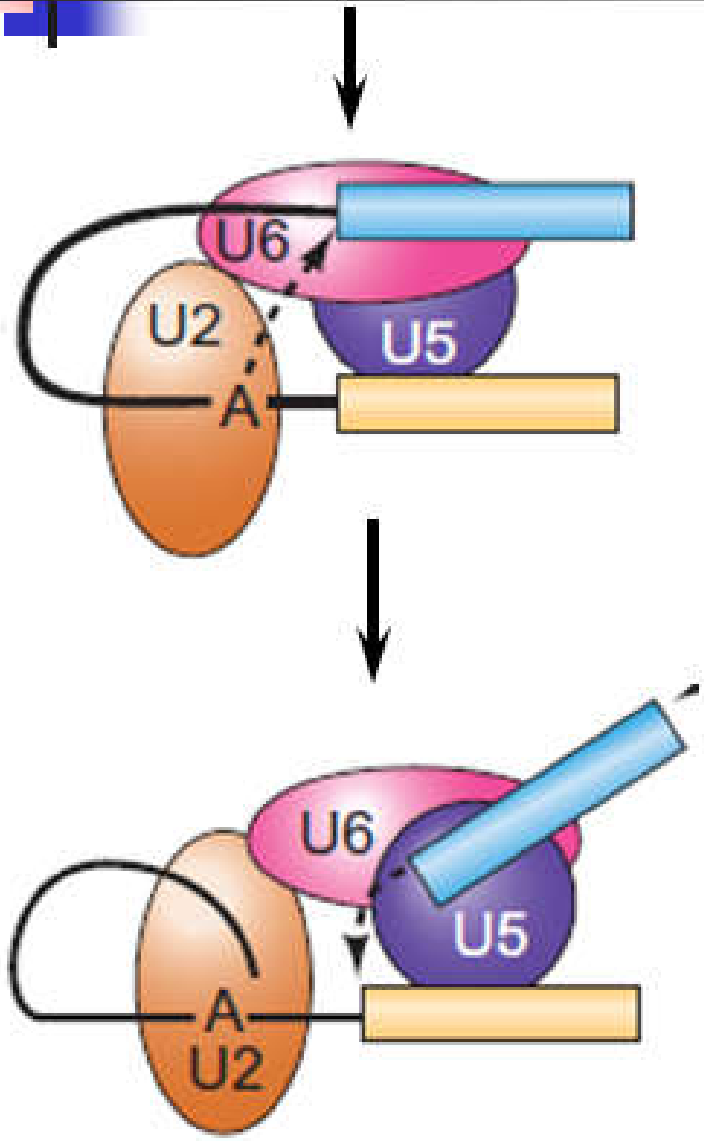
U1 binds to the 5'-splice site GU

U2 binds to the branchpoint A

U4, U5 and U6 can then bind, and the intron is looped out and the 5'- and 3'-exons close together.

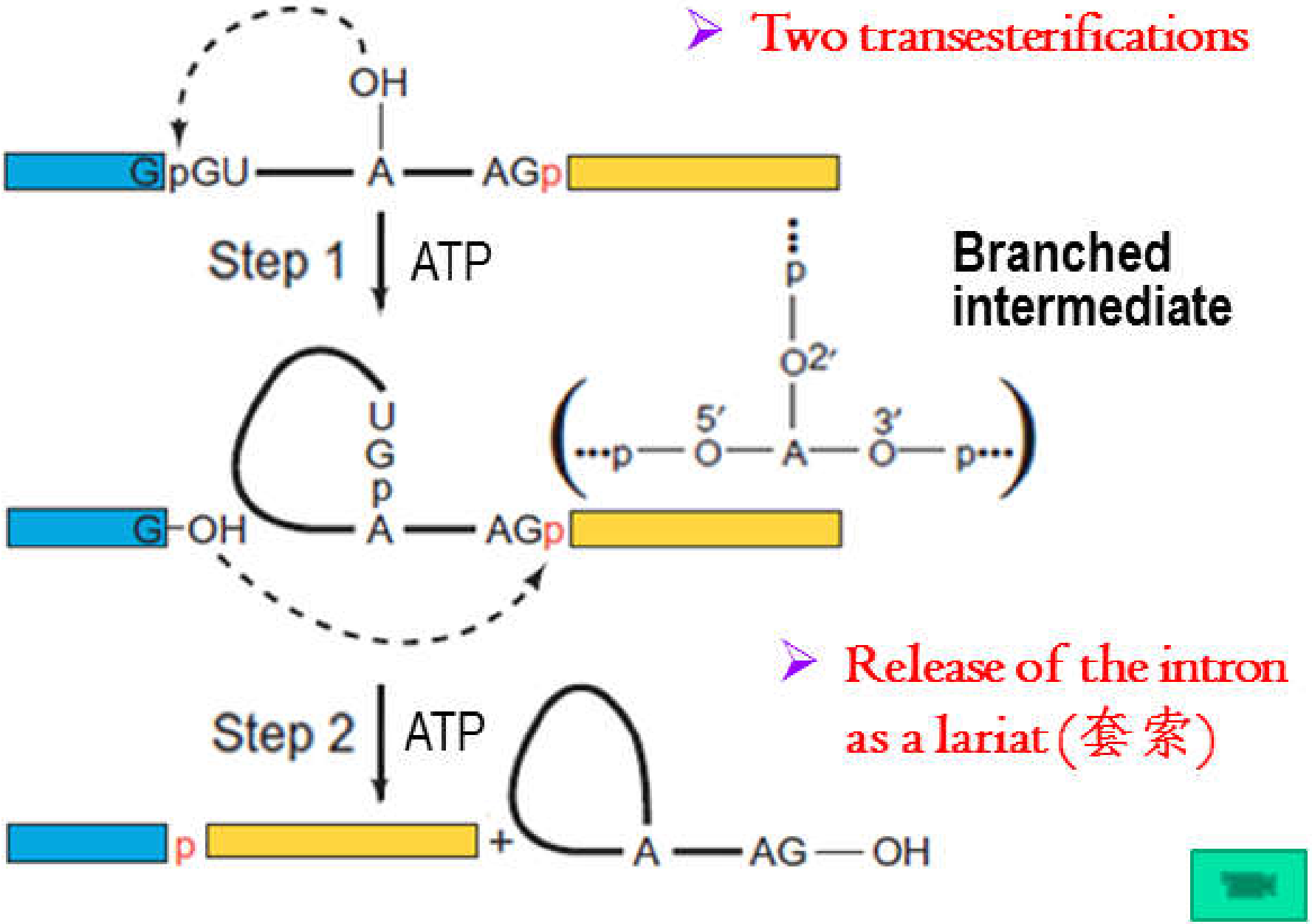


(2) Transesterification (转酯反应)



U1 and U4 leave the spliceosome. U6 binds to the 5'-splice site and interacts with U2. The first transesterification happens.

U5 assist the reaction of the second transesterification.

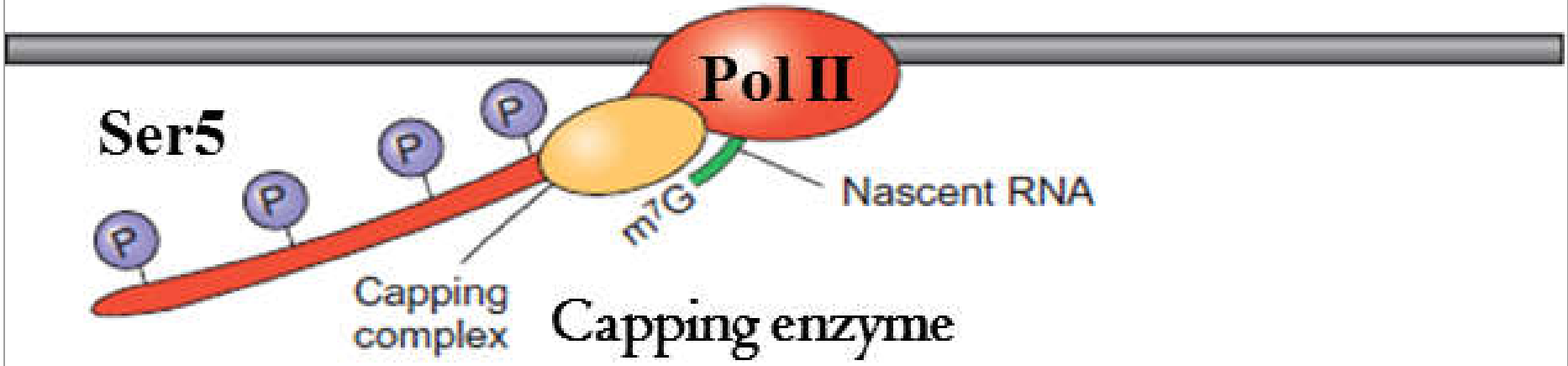

Two transesterifications
OH
GpGU
A
AGp
Step 1
ATP
Branched intermediate
p
O2′
5′
3′
U
G
p
···p
O
A
O
p···
G
OH
A
AGp
Release of the intron as a lariat (套索)
Step 2
ATP
p
+
A
AG
OH

3.5 CTD of RNA Pol II and mRNA processing

Capping, polyadenylation, and splicing proteins all associate with the CTD during transcription.

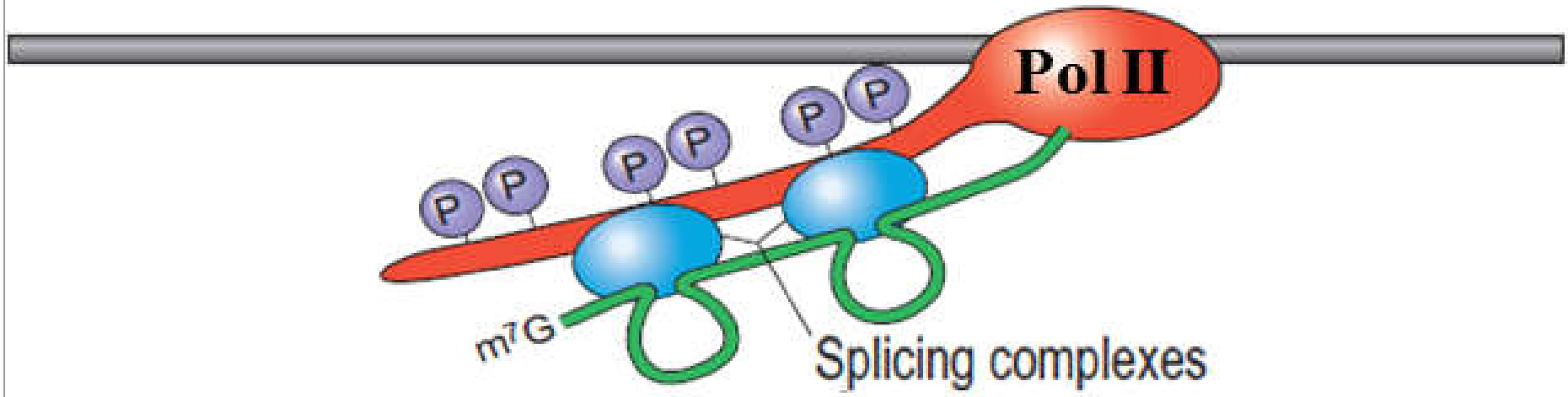
3.5.1 CTD and 5'capping

- Transcription complexes close to the promoter contain phosphorylated Ser5.
- Capping complex is recruited (募集) by **phosphorylation of Ser5** in Pol II CTD.



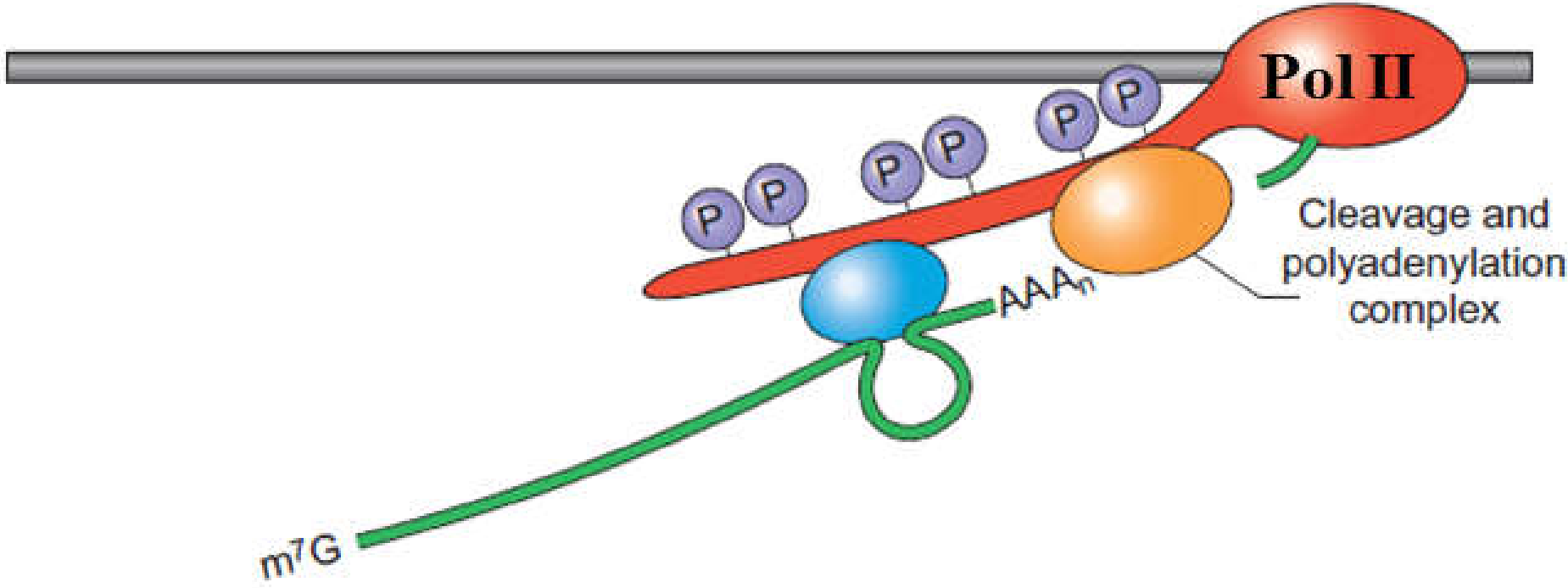
3.5.2 CTD and pre-mRNA splicing

- The CTD has become further phosphorylated (presumably including a shift from Ser5 to Ser2 phosphorylation) and has attracted the splicing complex.



3.5.3 CTD and 3' polyadenylation

- **The CTD is associated with the cleavage and polyadenylation complex.**



3.6 Pre-mRNA methylation

- A small percentage (0.1%) of A residues are methylated at the N6 position.
- 5'-RRACX-3' (R = purine, X is rarely G)

4. RNA self-splicing (自剪接)

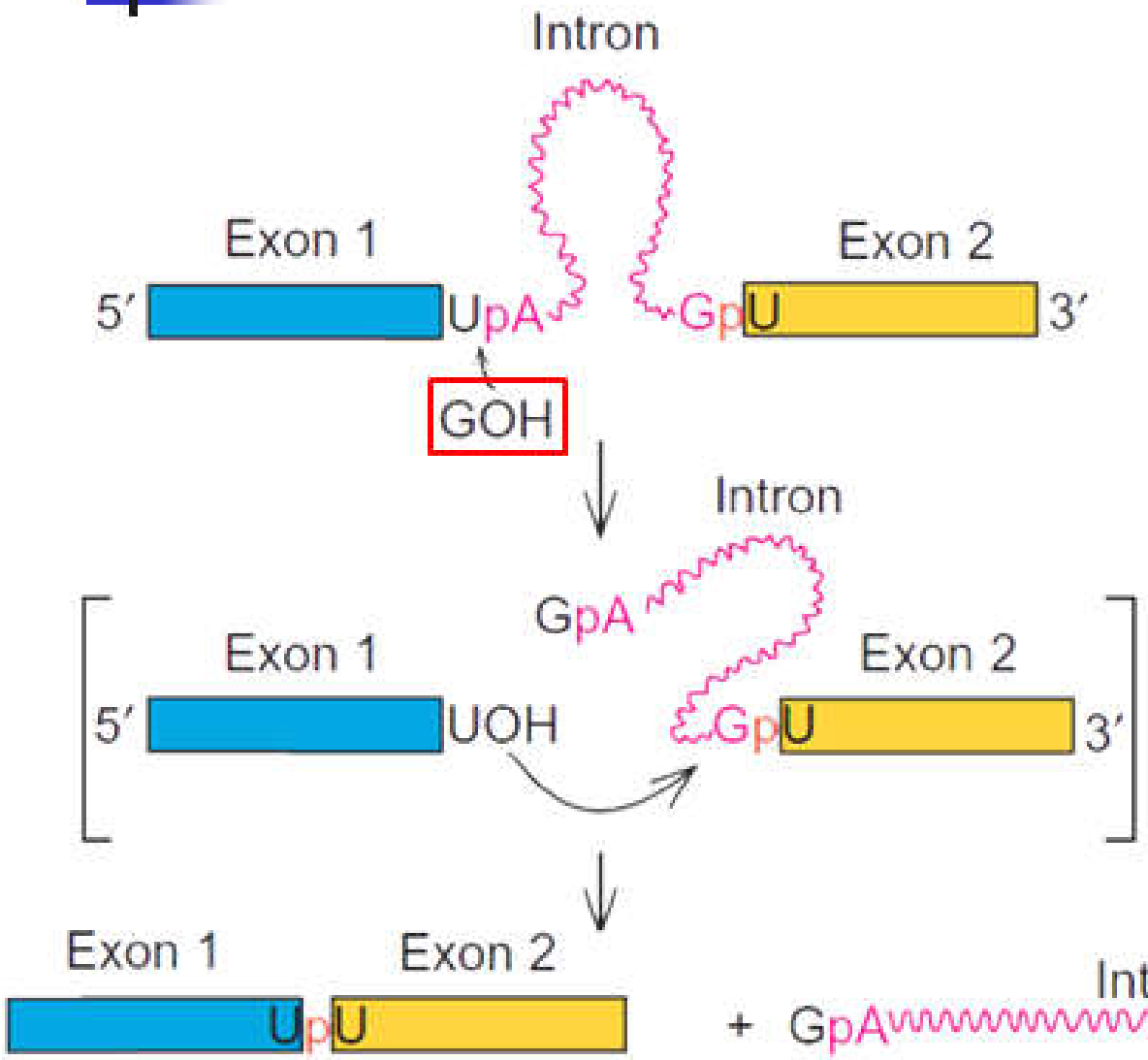
- **RNA self-splicing** describes the ability of an intron to excise itself from an RNA by a catalytic action that depends only the sequence of RNA in the intron. 内含子的RNA本身具有催化活性，能进行内含子的剪接。
- Thomas Cech and his coworkers made this discovery in their study of the 26S rRNA gene of the *Tetrahymena* (四膜虫).

4.1 Group I introns

I型内含子普遍存在于低等真核生物rRNA基因以及某些生物线粒体和叶绿体基因中，在体外无其他蛋白质参与的情况下能自剪接。

I型内含子内部无分支点腺苷酸。

I型内含子剪接机制

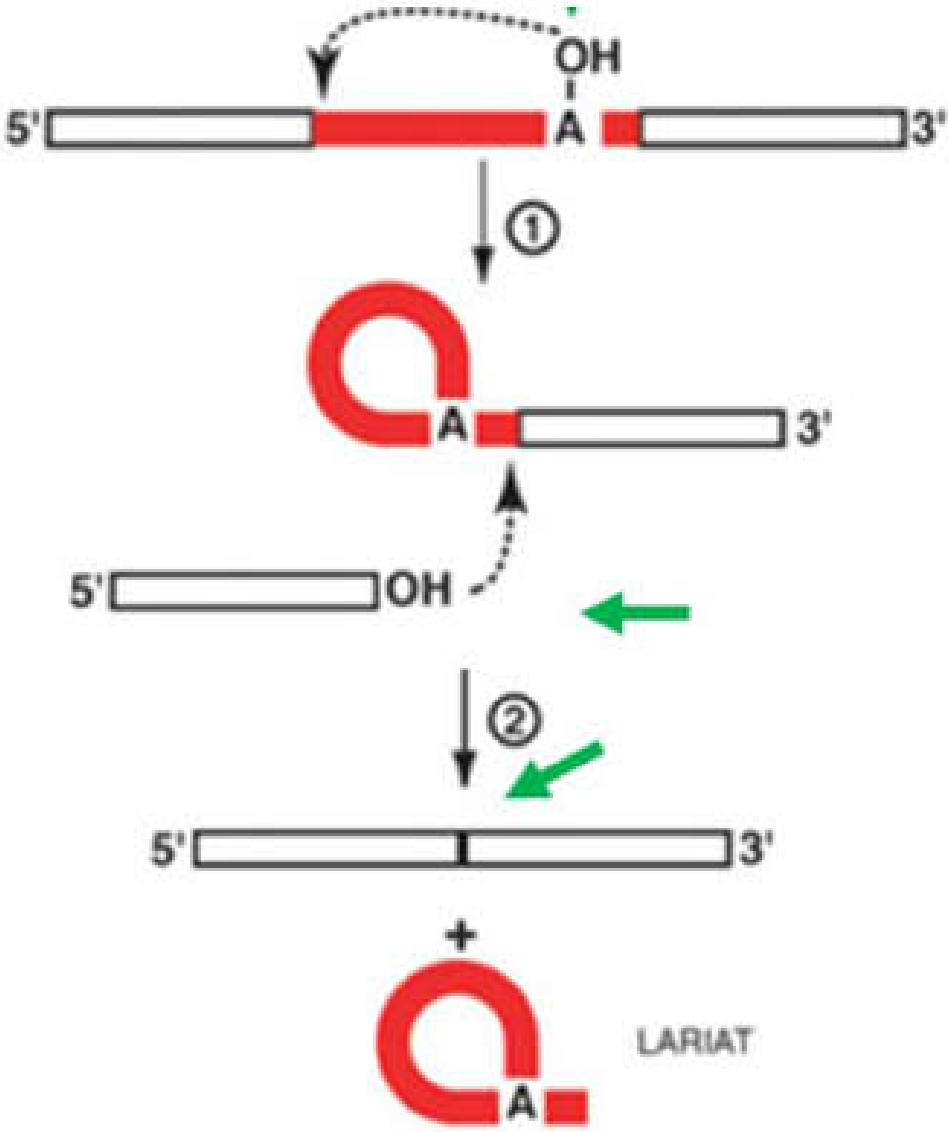


(1) **游离的鸟苷酸**攻击内含子的5'剪接位点，G被添加到内含子5'端，并释放第一个外显子。

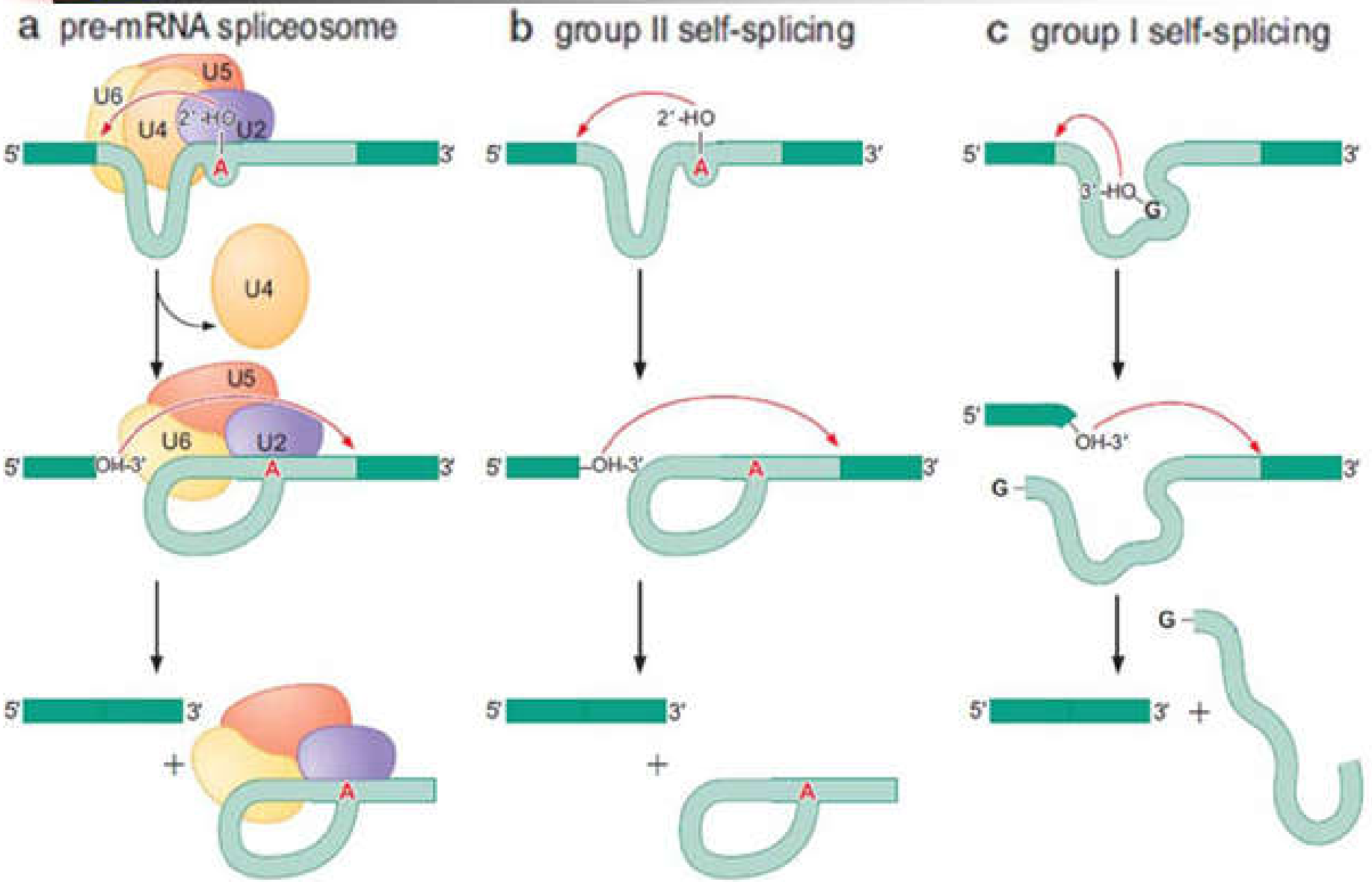
(2)第一个外显子攻击3'-剪接位点，两个外显子连接，释放线性内含子。

4.2 Group II introns

- II型内含子存在于原生生物、真菌、藻类以及植物的线粒体和叶绿体基因中，也能发生自剪接。
- II型内含子的自剪接**由内含子内部的腺苷酸启动**，并形成套索结构，类似于hnRNA剪接方式。



小结： Three classes of RNA splicing



5. Ribozyme (核酶)

Ribozymes are catalytic RNA molecules that catalyze particular biochemical reactions.

核酶是一种可以催化特定生化反应的RNA分子。

5.1 Types of ribozymes

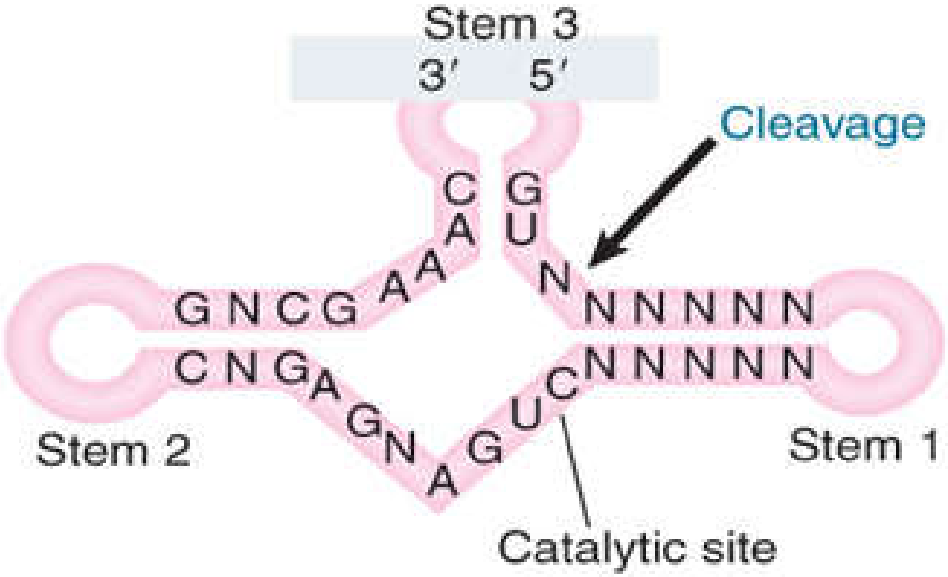
(1) Self-splicing introns: Group I introns and Group II introns

(2) RNase P is a common ribozyme that matures tRNA and acts as an endonuclease.

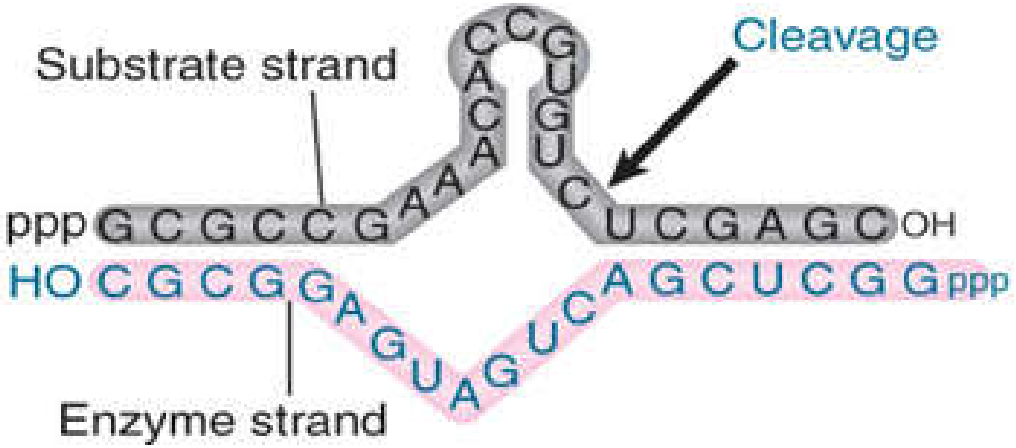
(3) Self-cleaving RNA

e.g. hairpin ribozyme, hammer head ribozyme

Consensus hammerheads have three stem loops and conserved bases



Hammerheads can be created by interaction between two complementary RNA molecules

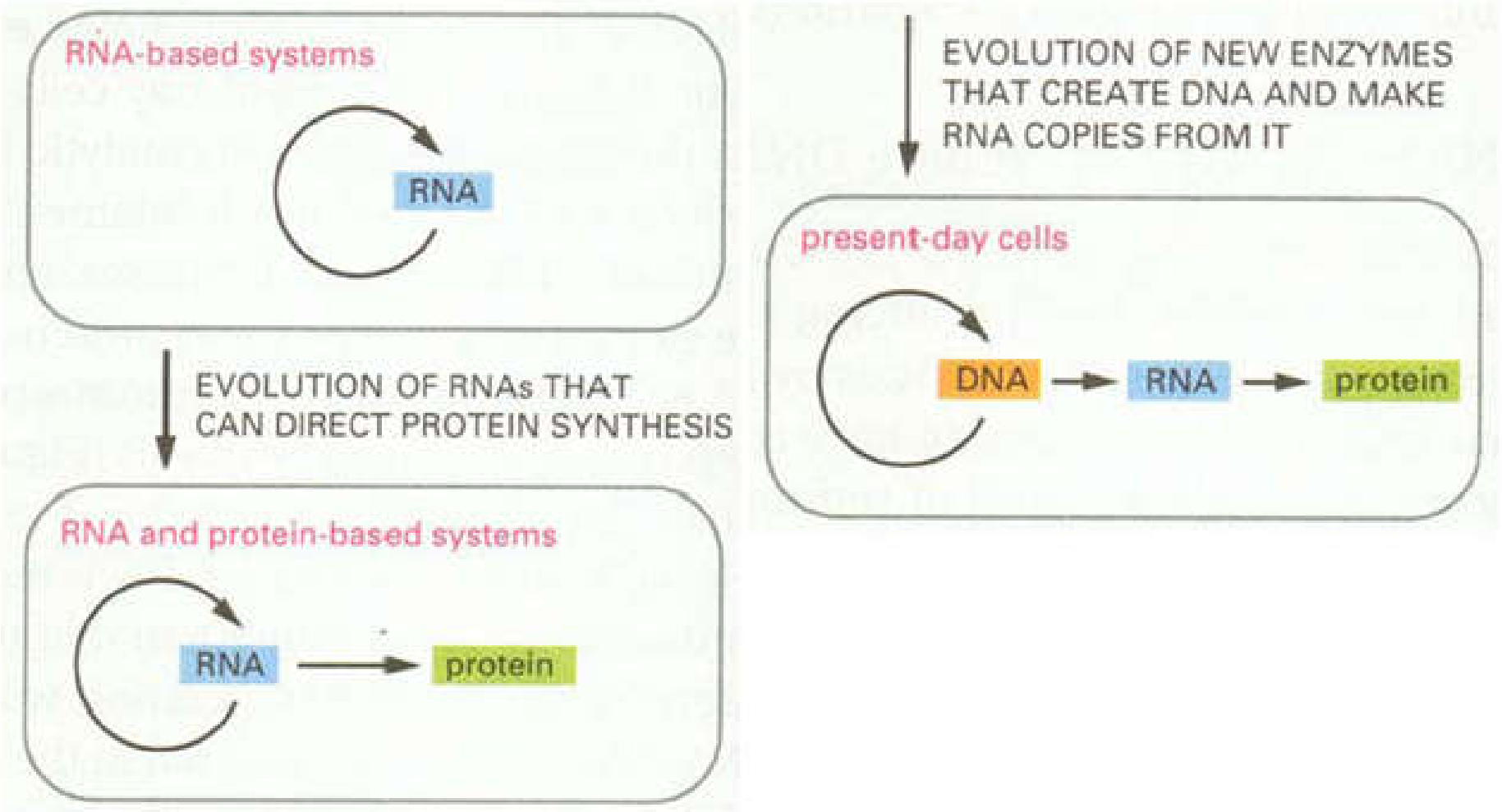


5.2 Applications of ribozymes

Ribozymes can be used in inhibiting unwanted gene expression by cleaving mRNA molecules.

- Kill cancer cells
- Prevent virus replication
- Discover the function of new genes

5.3 The position of RNA in the origin of life and its evolution

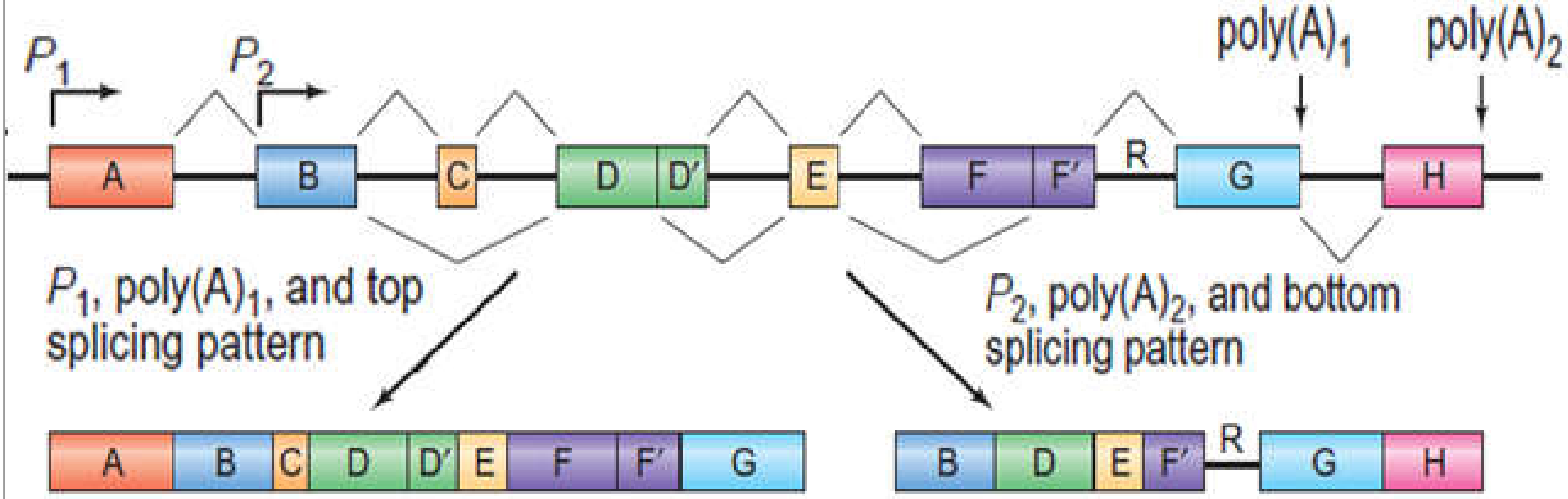


6. Alternative mRNA processing

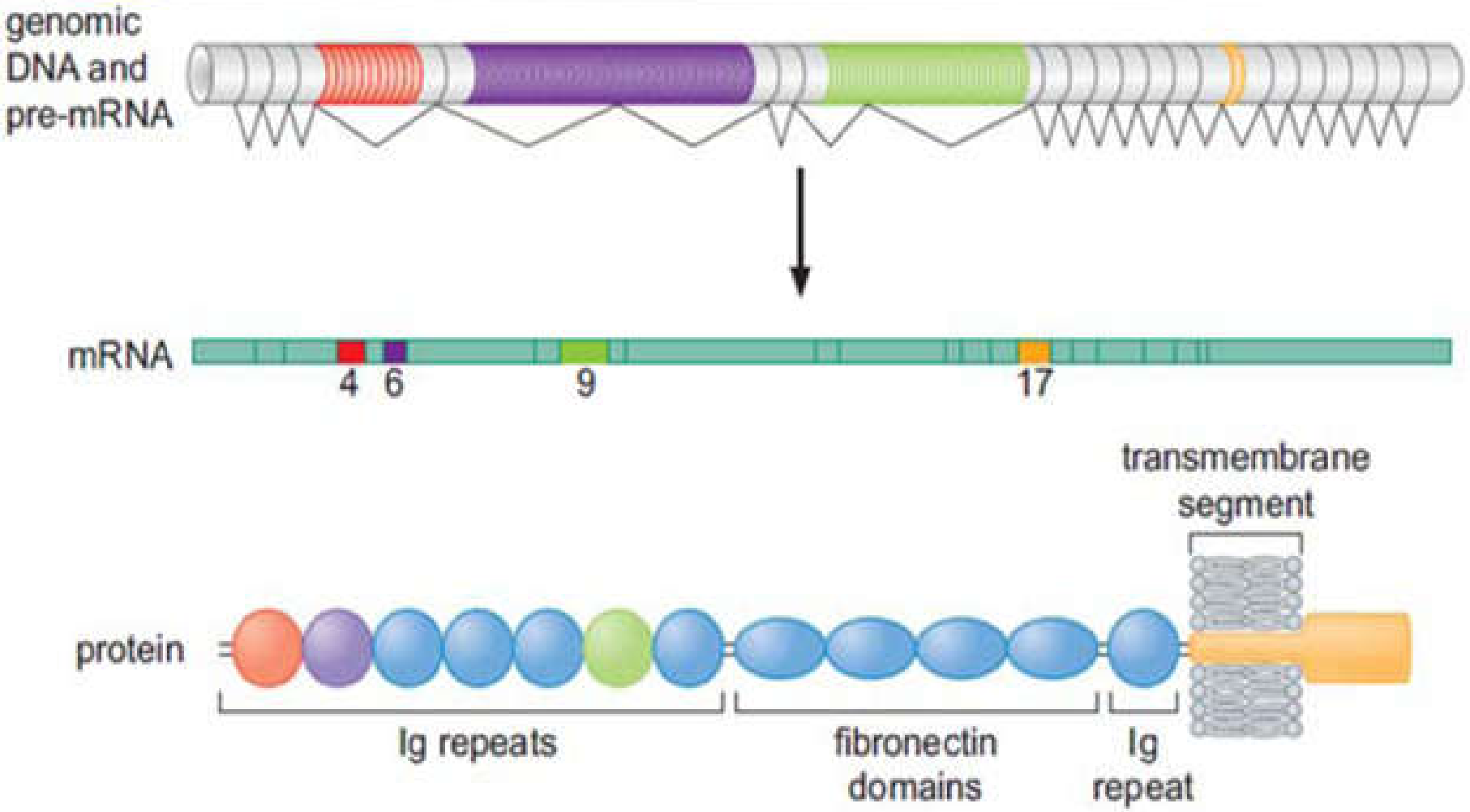
- **Alternative mRNA processing (splicing)** is the conversion of pre-mRNA species into more than one type of mature mRNA.
- The transcripts of many eukaryotic genes are subject to alternative processing.
- **Functions**: One gene encodes multiple proteins; regulate the expression of a given gene.

6.1 Types of alternative mRNA processing

- Using different promoters
- Retaining certain introns
- Retaining or removing certain exons
- Using different poly(A) sites



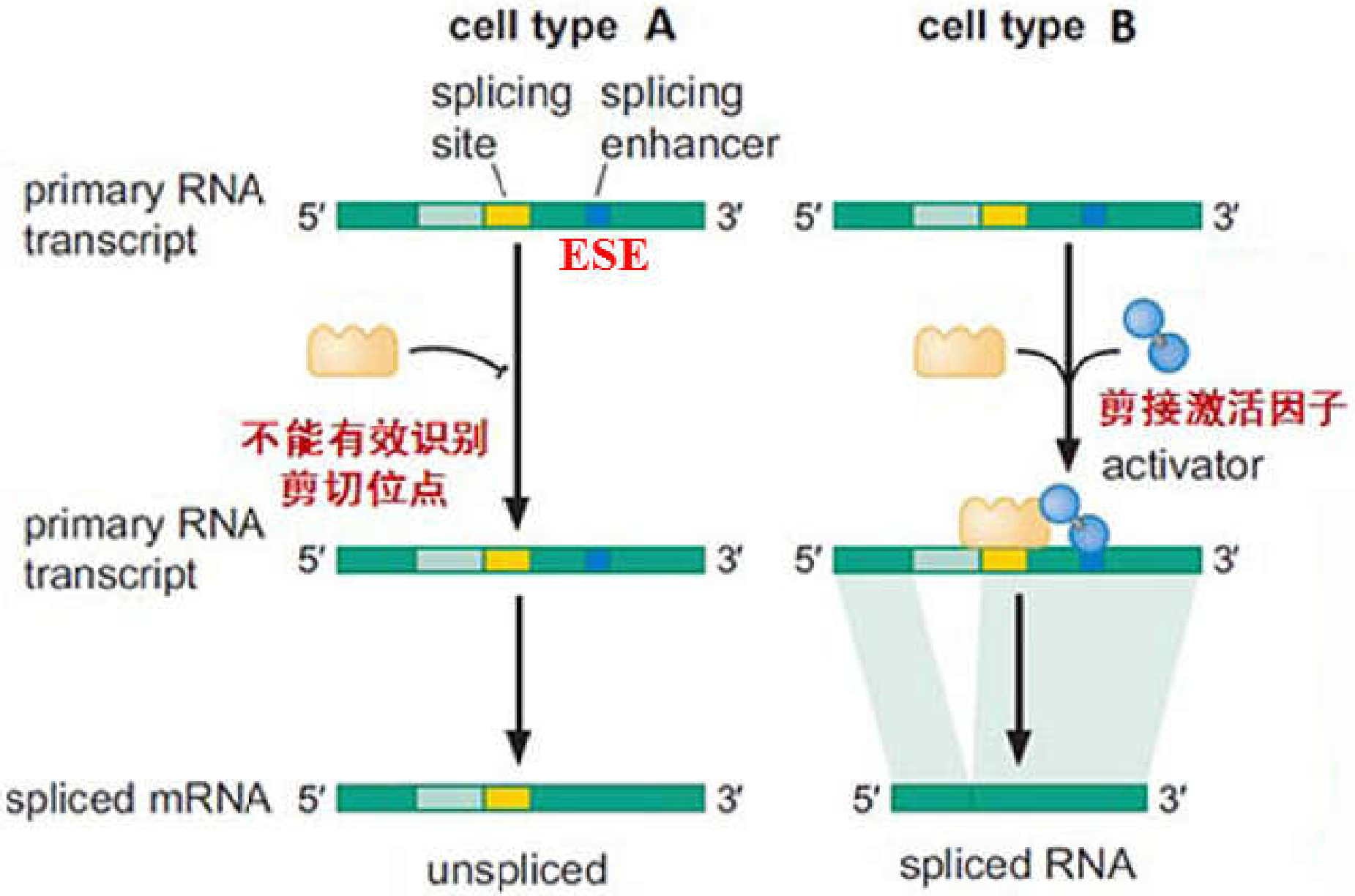
e.g. 果蝇*Dscam*基因



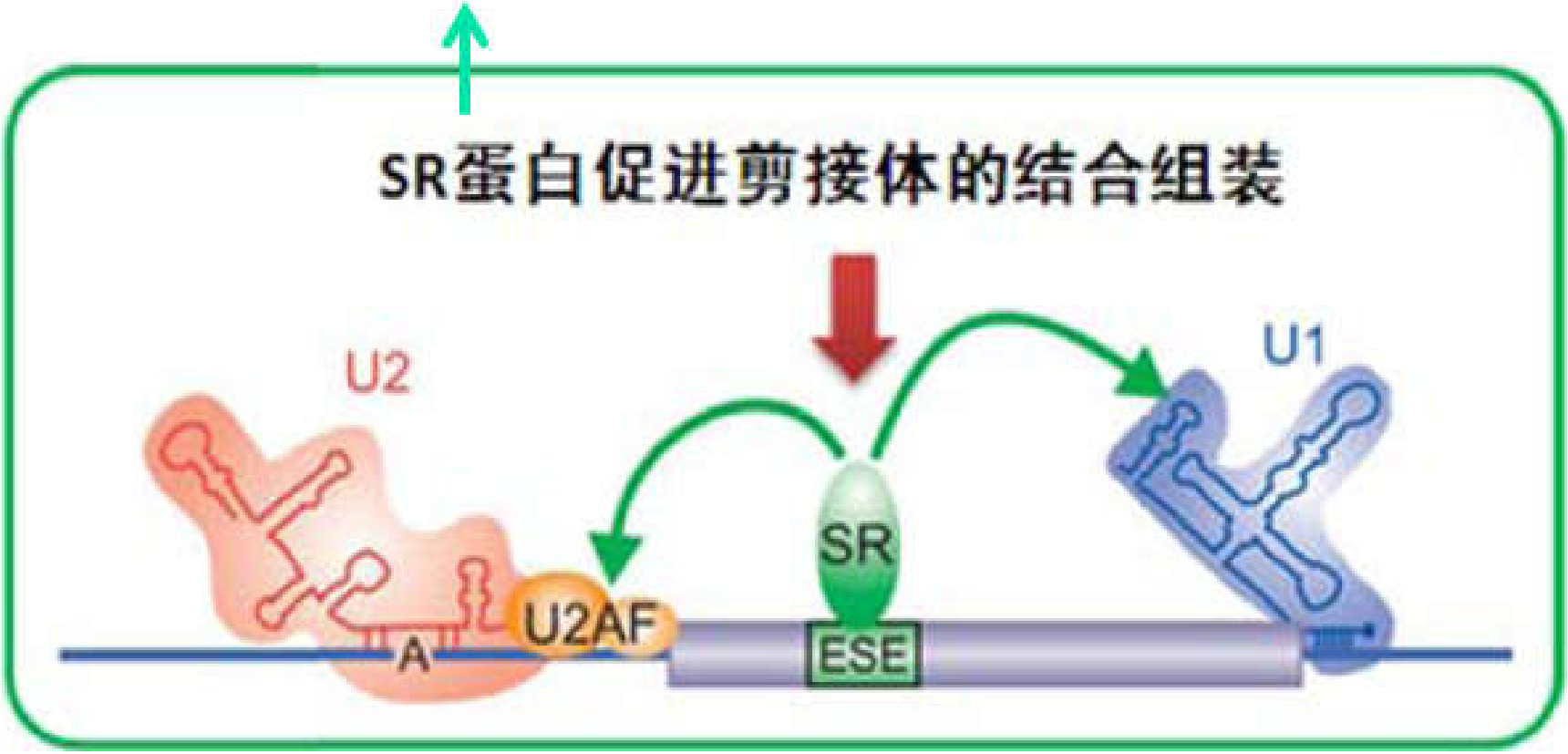
可能的可变剪切产物数量 12 X 48 X 33 X 2 = 38016

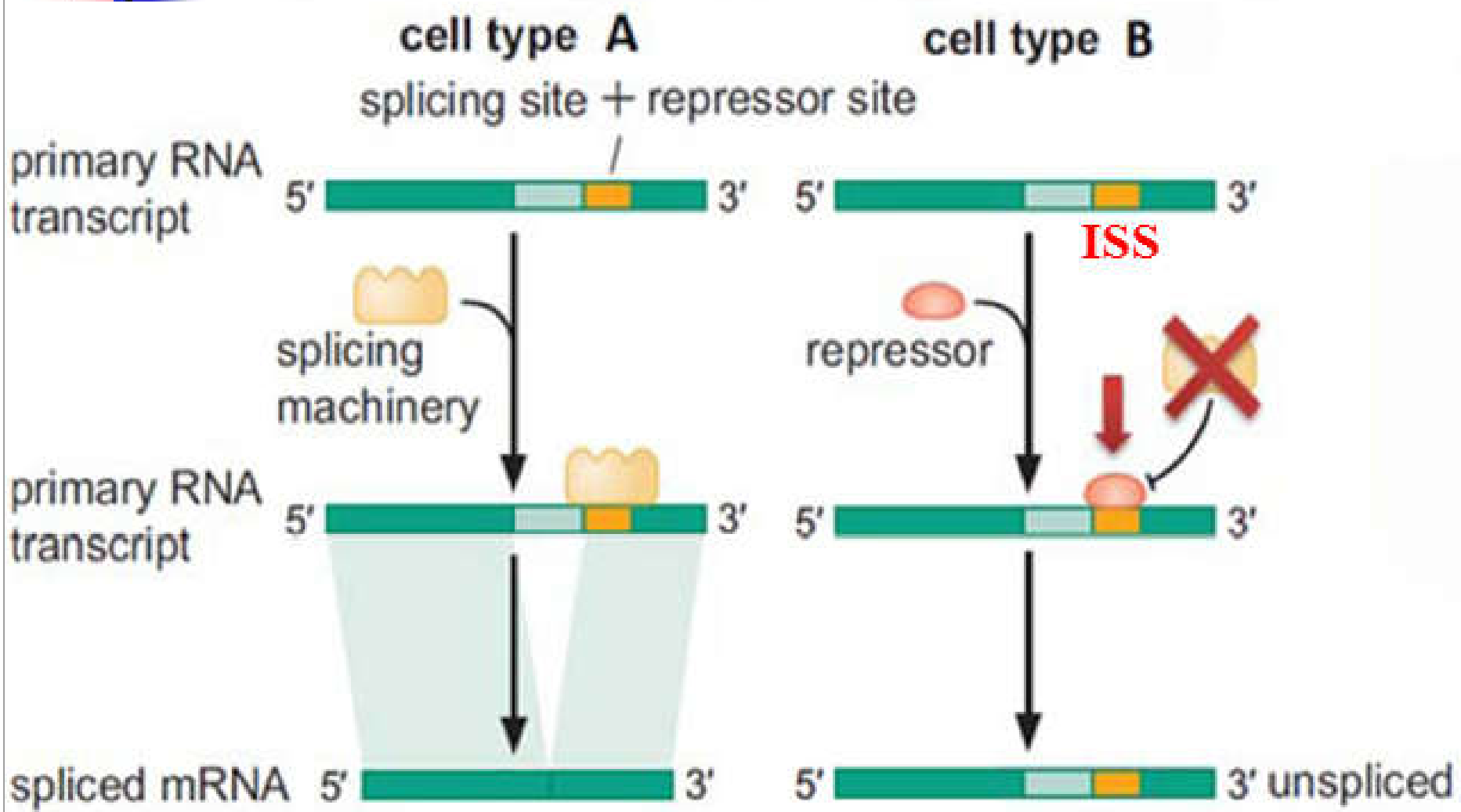
6.2 Regulation of the alternative mRNA splicing site

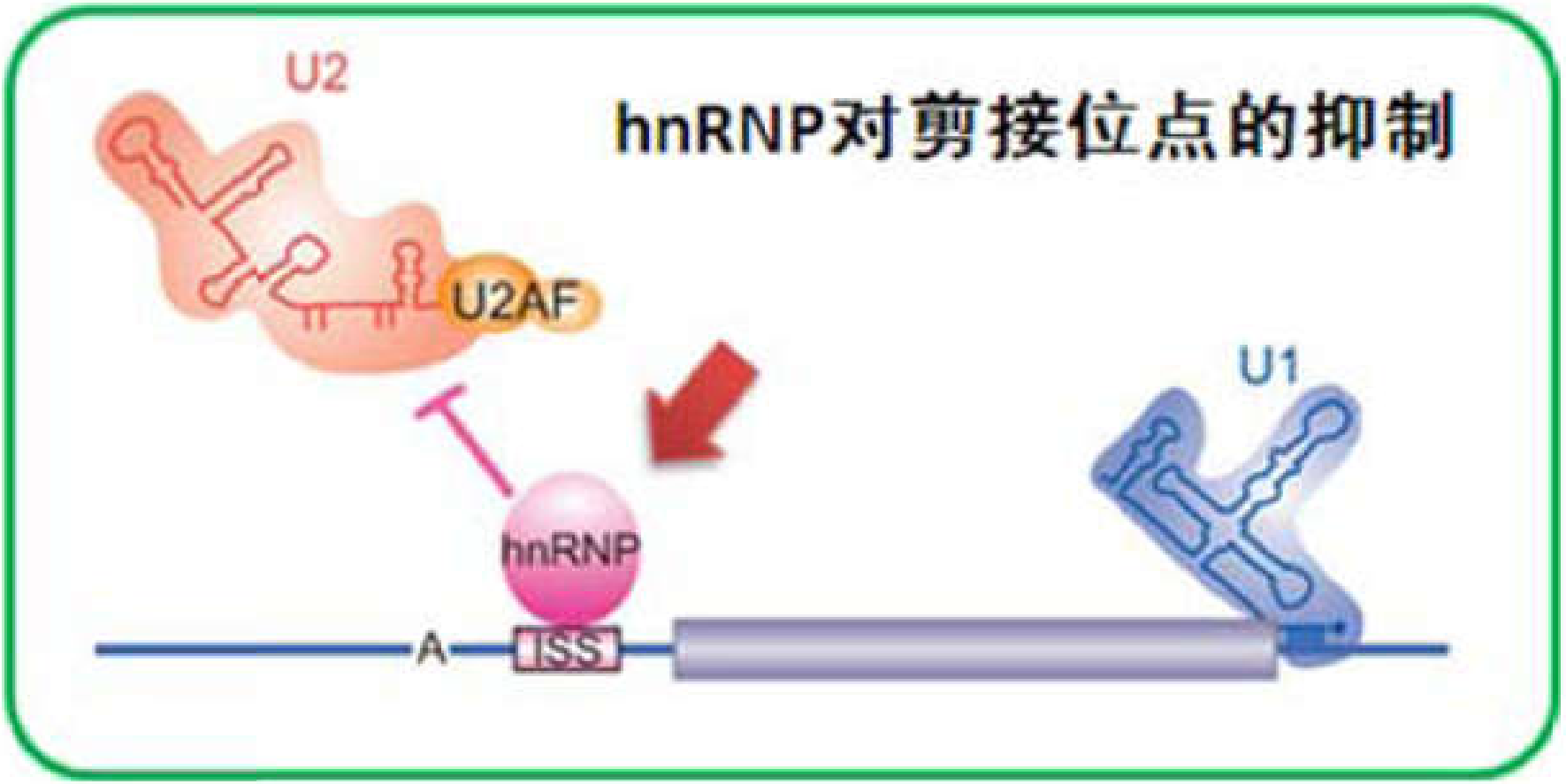
- **Splicing factors that stimulate commitment (定向) at certain splice sites.**
- **Exonic splicing enhancers (ESEs, 外显子剪接增强子) stimulate splicing, while exonic splicing silencers (ESSs, 外显子剪接沉默子), which inhibit splicing.**
- **Intronic splicing enhancers (ISE) and silencers (ISS) also exist.**


cell type A
cell type B
splicing site
splicing enhancer
primary RNA transcript
5′
3′
ESE
不能有效识别剪切位点
剪接激活因子
activator
primary RNA transcript
spliced mRNA
unspliced
spliced RNA

Ser and Arg-rich



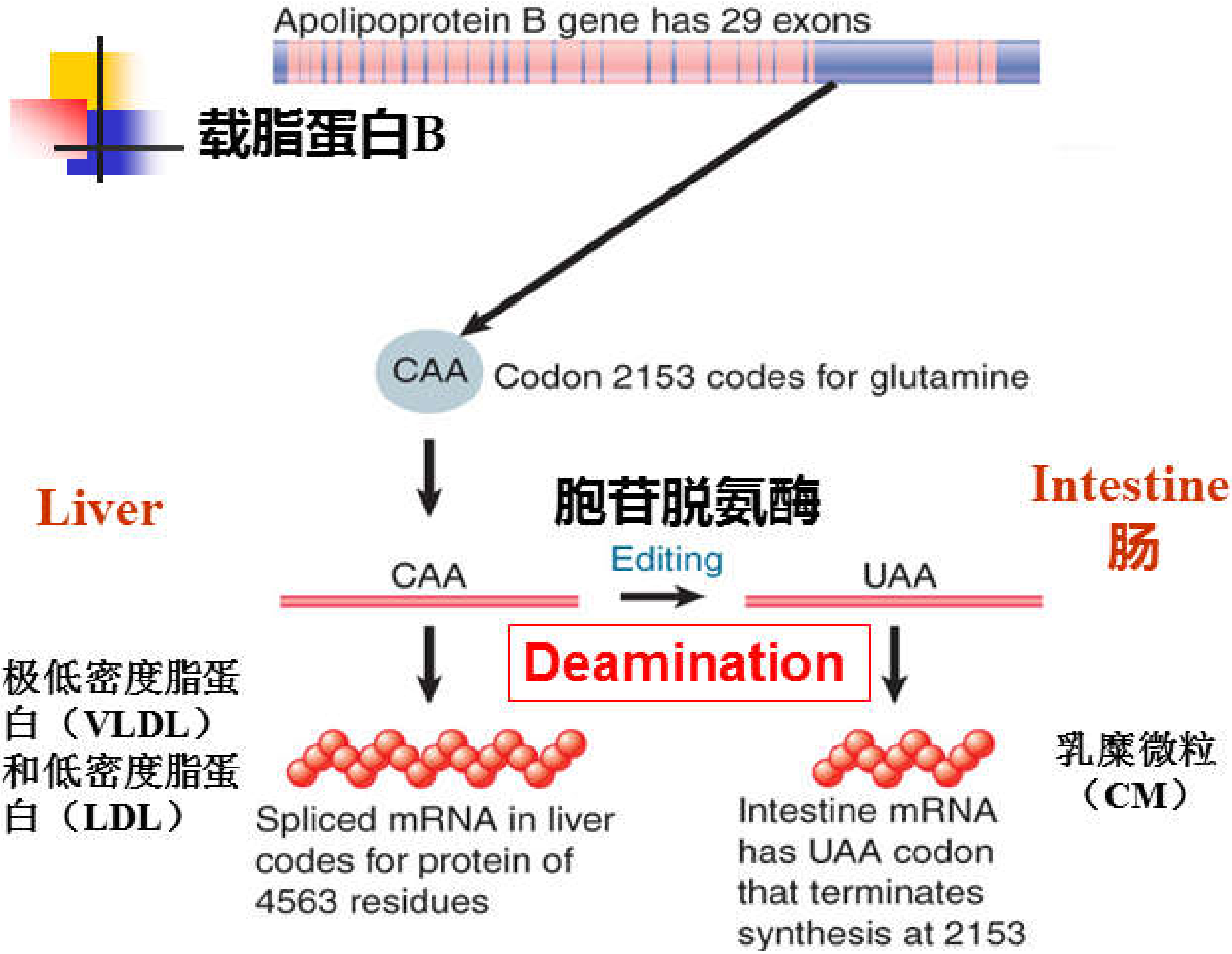

cell type A
cell type B
splicing site + repressor site
primary RNA transcript
5′
3′
ISS
splicing machinery
repressor
primary RNA transcript
spliced mRNA
unspliced


hnRNP对剪接位点的抑制
U2
U2AF
U1
hnRNP
A
ISS

7. RNA editing

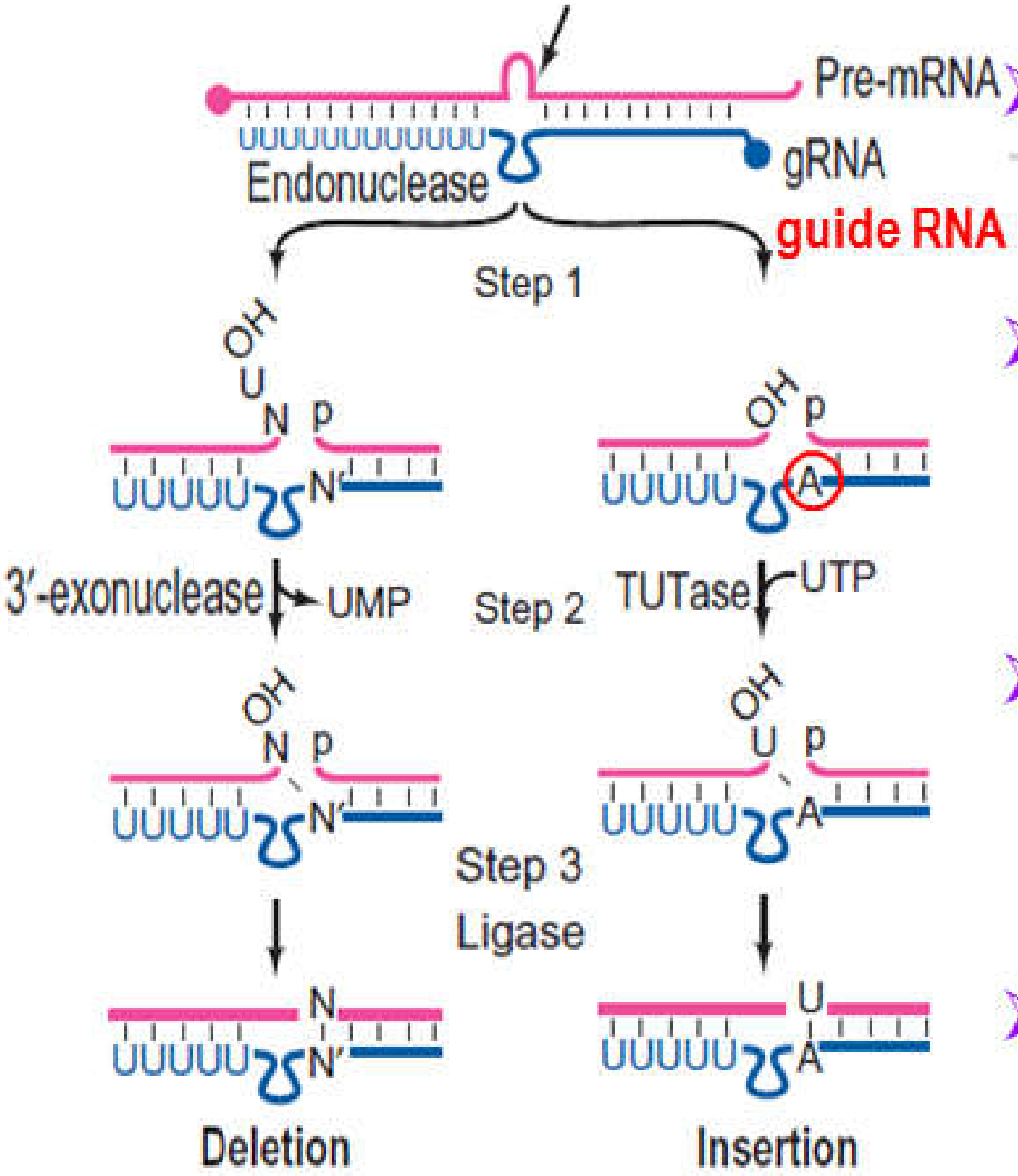
- **RNA editing** is an unusual form of RNA processing in which the nucleotide sequence of the primary transcript is altered by either changing, inserting or deleting residues at specific points along the molecule.

  RNA编辑是一种不同寻常的RNA加工模式，是通过改变、插入或删除初生转录物特定部位的碱基而改变其中的核苷酸序列。


Apolipoprotein B gene has 29 exons
载脂蛋白B
CAA
Codon 2153 codes for glutamine
Liver
胞苷脱氨酶
Intestine
肠
CAA
Editing
UAA
Deamination
极低密度脂蛋白（VLDL）和低密度脂蛋白（LDL）
Spliced mRNA in liver codes for protein of 4563 residues
Intestine mRNA has UAA codon that terminates synthesis at 2153
乳糜微粒（CM）

Insertion or deletion of U in protozoa 原生动物

Part of the mRNA sequence of *Trypanosome brucei* (布氏锥虫) *CoxIII* shows many uridines that are not coded in the DNA (shown in red) or that are removed from the RNA (shown as T).



- gRNA与pre-mRNA杂交，不匹配处被内切酶切开；
- gRNA提供A, G，可作为U掺入的模板。增加的U来源于UTP；
- 若gRNA无A或G与U配对，pre-mRNA中的U将被外切酶移除；
- 断开的pre-mRNA由RNA连接酶连接。

Summary

1. Types of RNA processing
2. Mechanisms of rRNA, tRNA and mRNA processing
3. Self-splicing
4. Ribozymes in RNA processing
5. Types of alternative mRNA processing and definition of RNA editing