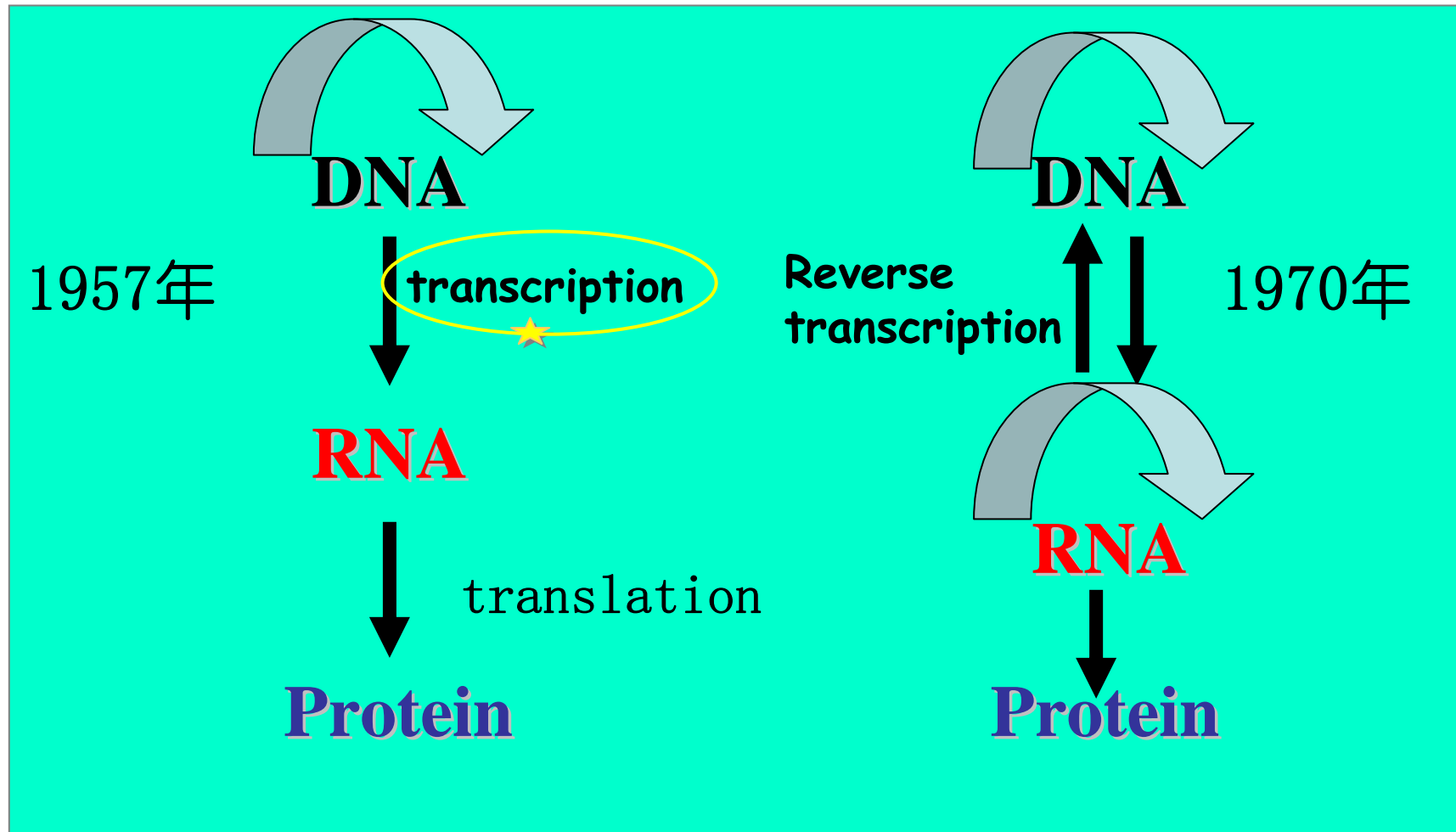


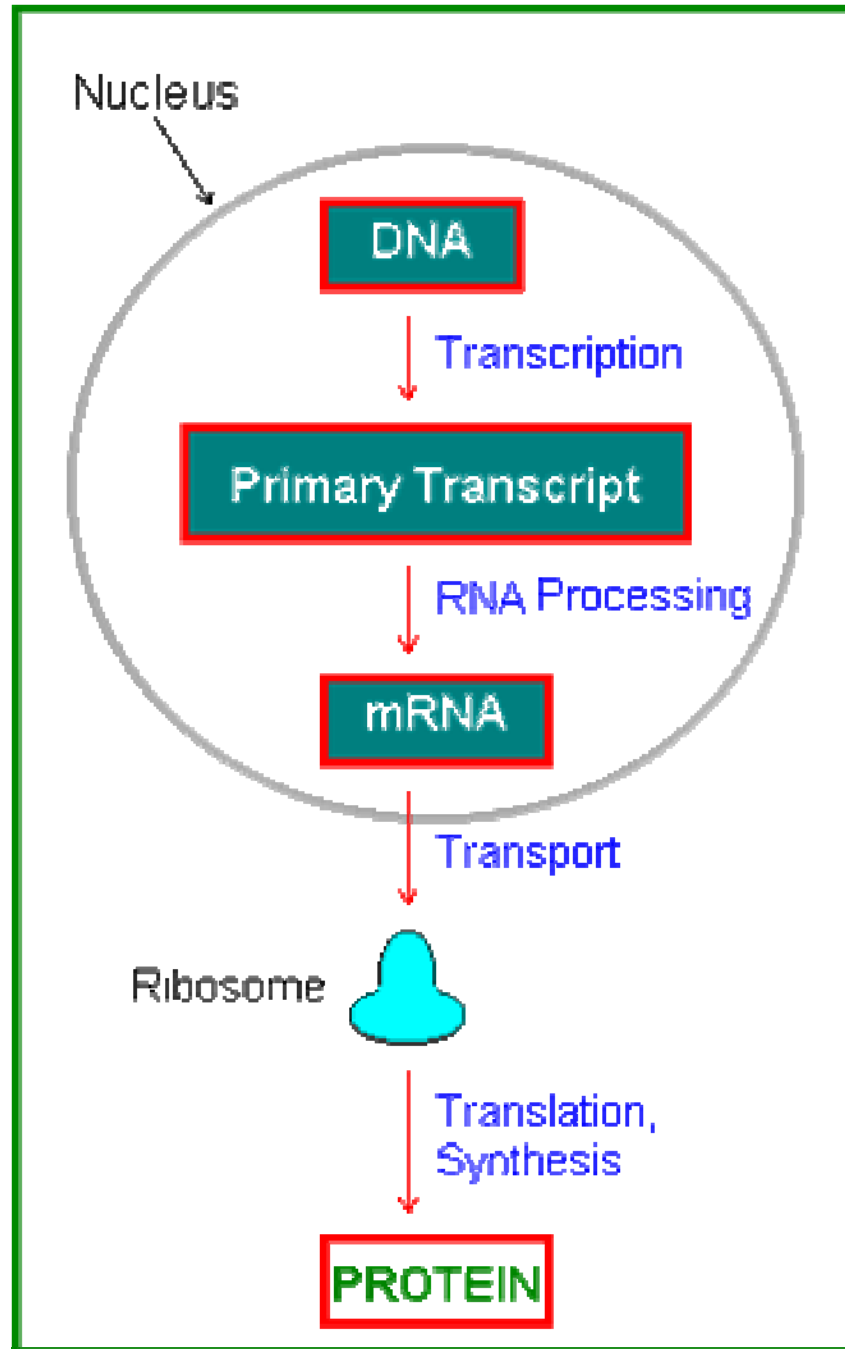
# 第三讲 生物信息的传递

## (上) 从DNA到RNA

# Crick的中心法则 (central dogma)



## From DNA to Protein



**DNA**序列是遗传信息的贮存者，它通过自主复制得到永存，并通过**转录**生成**信使RNA**，**翻译**生成蛋白质的过程来控制生命现象。

基因表达包括**转录**（transcription）和**翻译**（translation）两个阶段。

**转录**是指拷贝出一条与DNA链序列完全相同(除了T→U之外)的RNA单链的过程，是基因表达的核心步骤。

**翻译**是指以新生的mRNA为模板，把核苷酸三联遗传密码子翻译成氨基酸序列、合成多肽链的过程，是**基因表达的最终目的**。

- **Replication:** synthesis of two DNA strands using both parental DNA strands as templates. *Duplication of a DNA molecule*  
1 DNA molecule → 2 DNA molecules

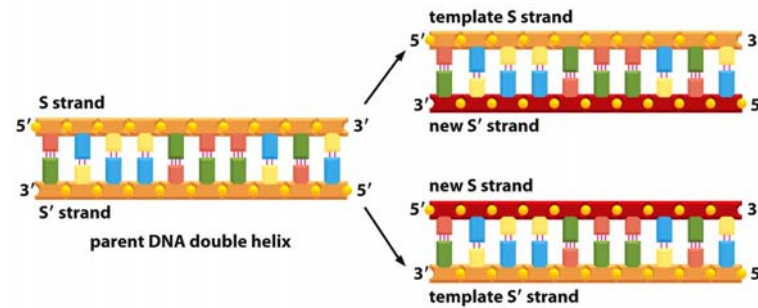


Figure 6-3 Essential Cell Biology 3/e (© Garland Science 2010)

- **Transcription:** synthesis of one RNA molecule using one of the two DNA strands as a template.

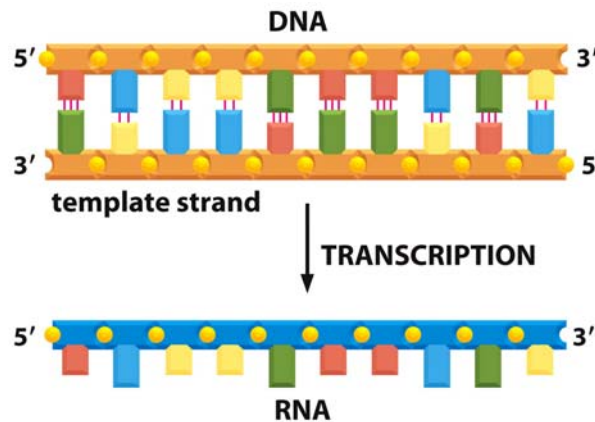


Figure 7-6 Essential Cell Biology 3/e (© Garland Science 2010)

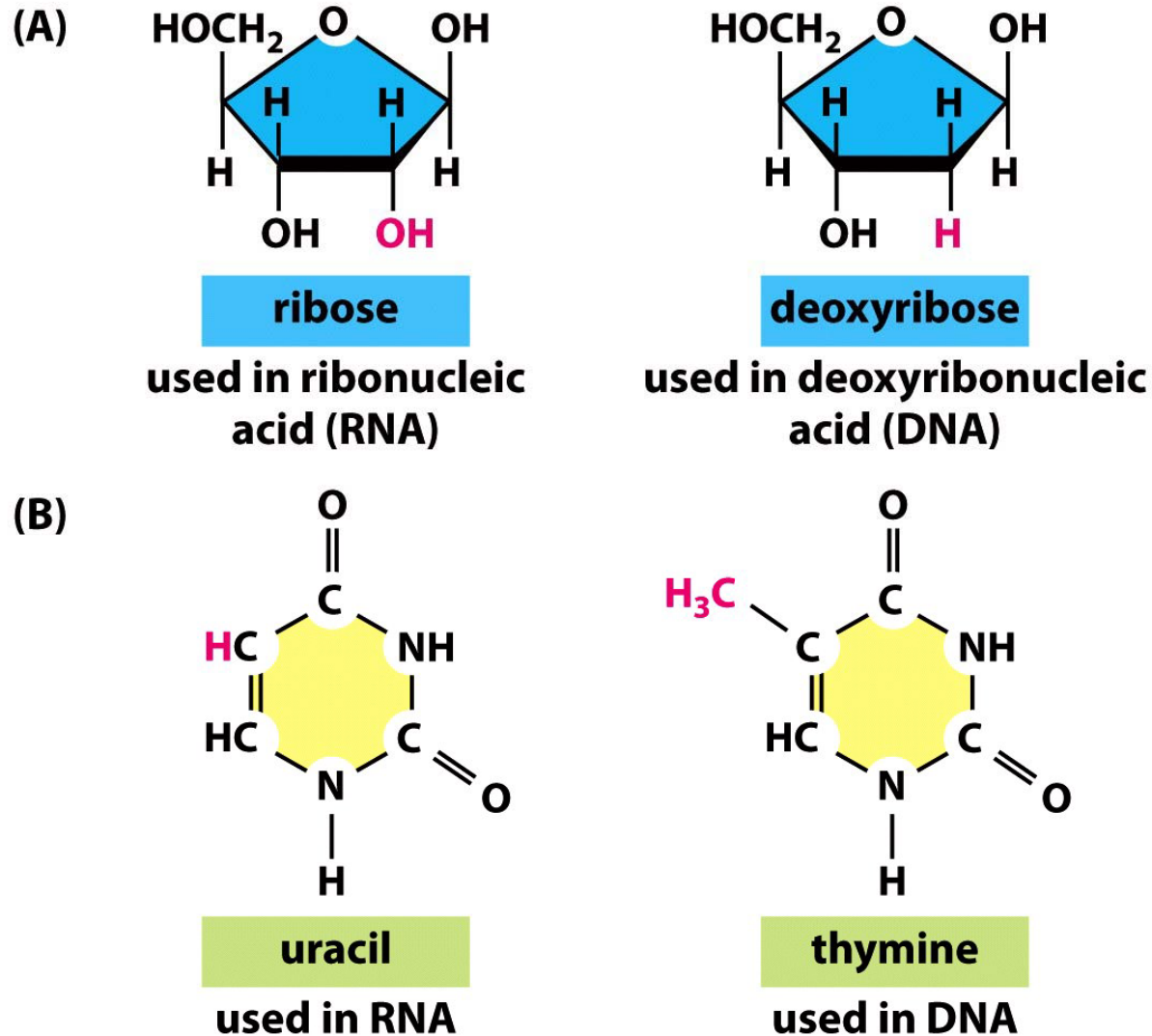
# Transcription vs. Replication

Transcription is **chemically** and **enzymatically** very similar to DNA replication.

## Some important differences:

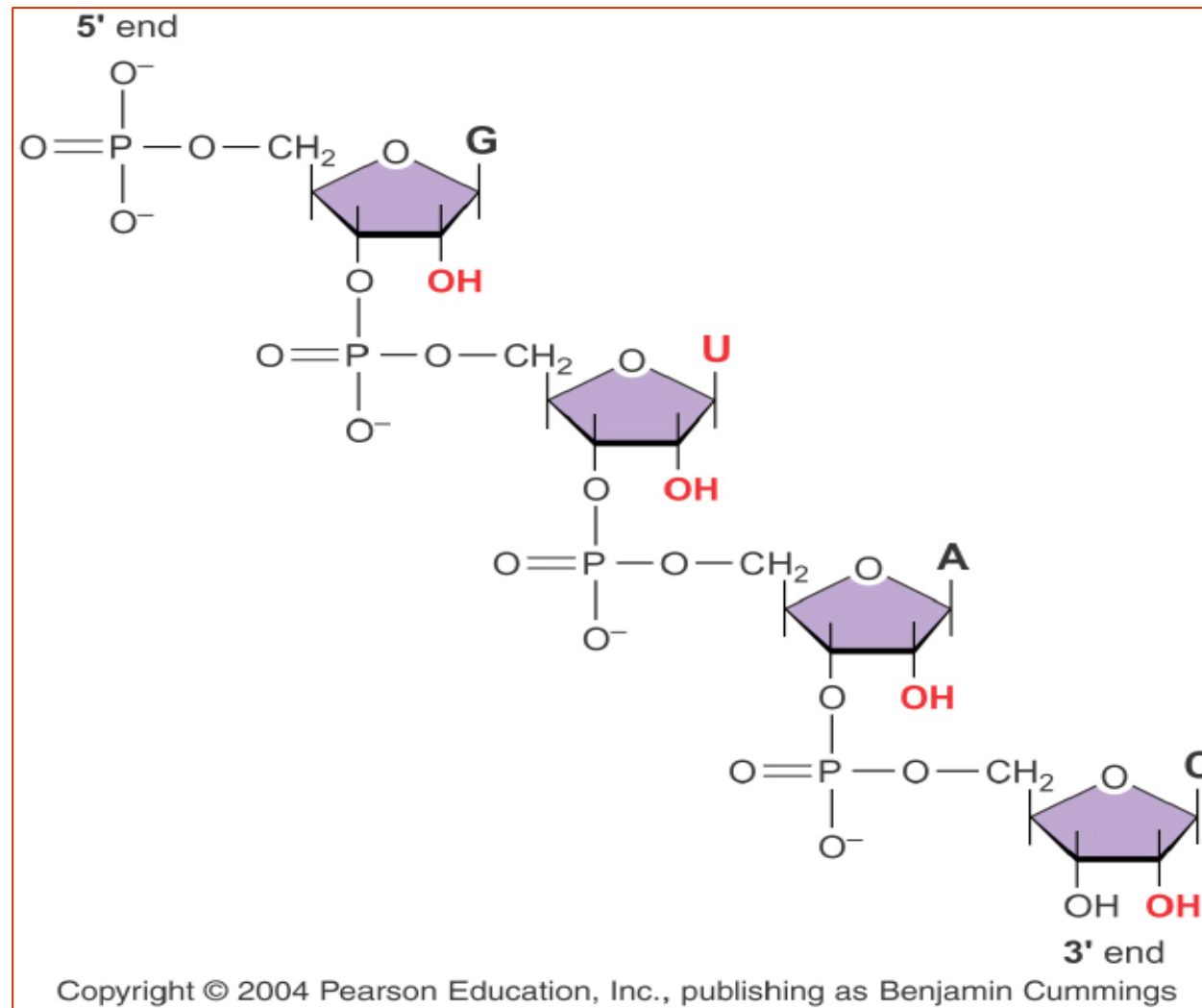
1. RNA is made of ribonucleotides.
2. RNA polymerase catalyzes the reaction, which does not require a primer (*de novo* synthesis).
3. The RNA product does not remain base-paired to the template DNA strand.
4. Less accurate (error rate:  $10^{-4}$ ).

# RNA的化学结构与DNA有略微差异





# RNA主要以单链形式存在于生物体内



# RNA分子能够形成分子内的碱基配对， 进而折叠成特殊的结构

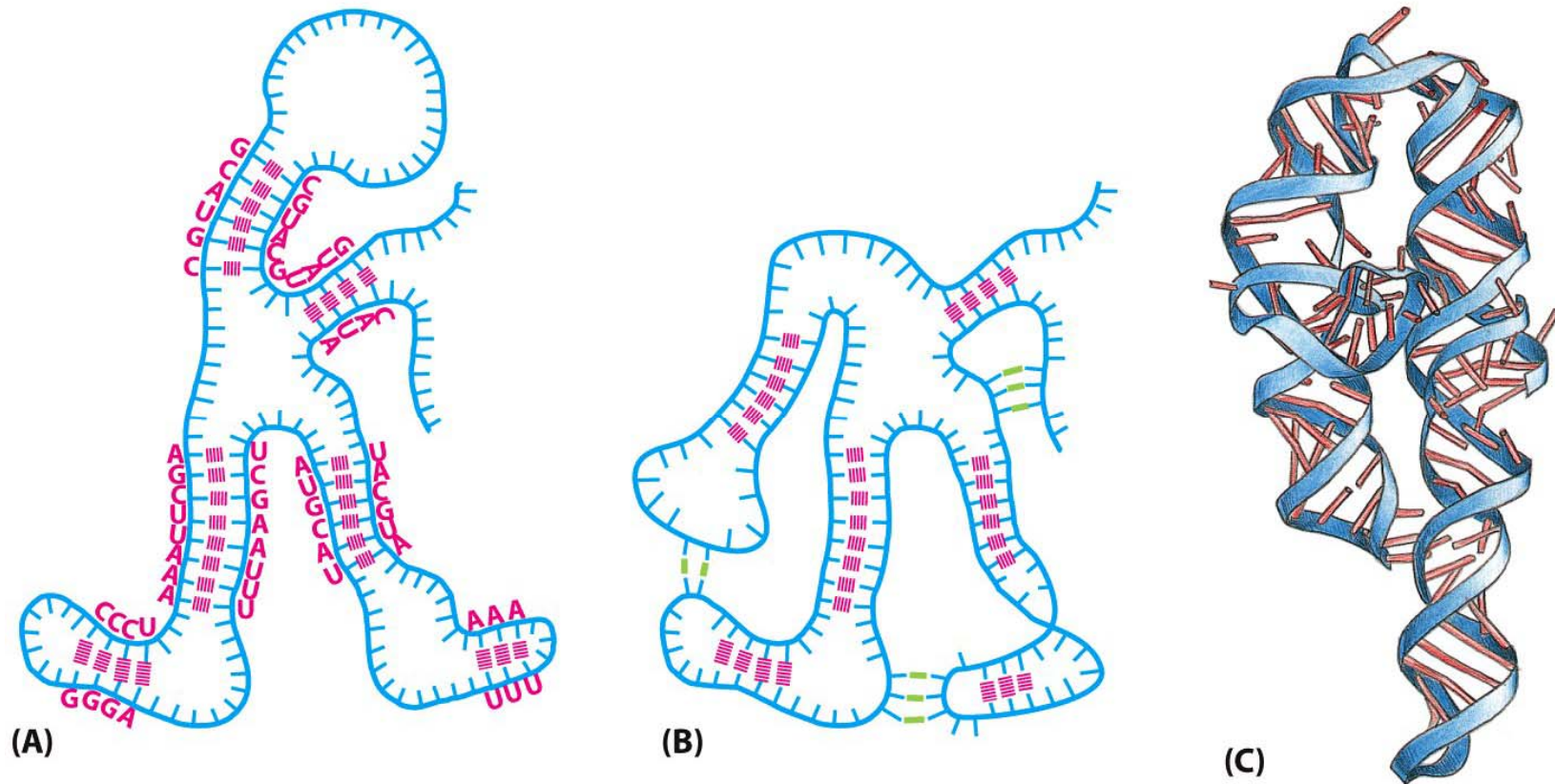
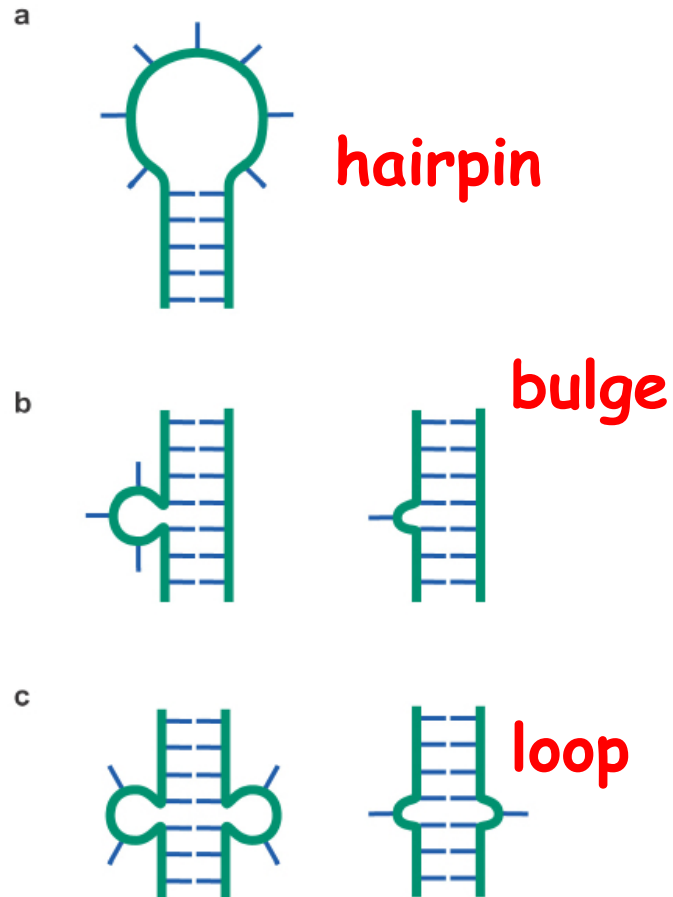


Figure 7-5 Essential Cell Biology 3/e (© Garland Science 2010)

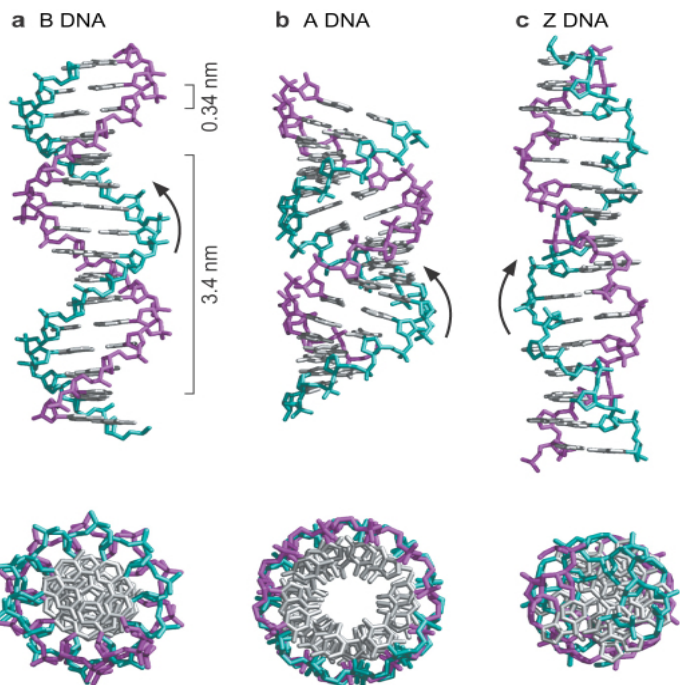
# Secondary structure of RNA



Copyright © 2004 Pearson Education, Inc., publishing as Benjamin Cummings

RNA chains fold back on themselves to form local regions of double helix.

# A-, B-, Z-型DNA三种结构比较

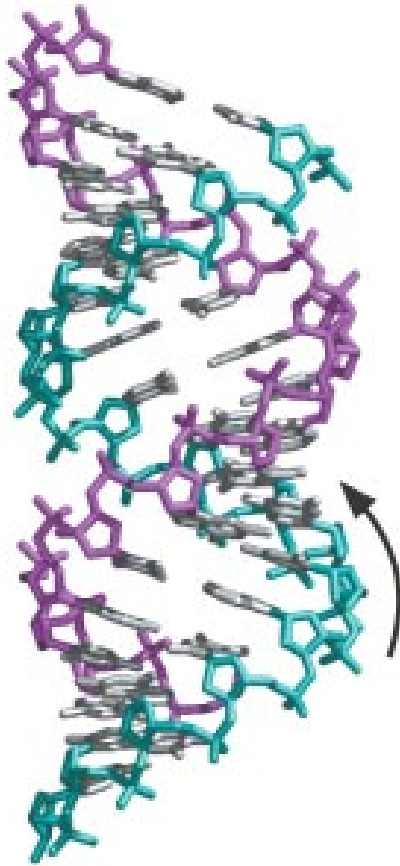


Copyright © 2004 Pearson Education, Inc., publishing as Benjamin Cummings

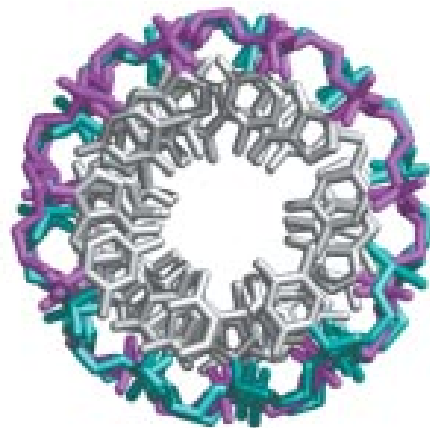
	<b>A</b>	<b>B</b>	<b>Z</b>
外型	粗短	适中	细长
螺旋方向	右手	右手	左手
螺旋直径	<b>2.55nm</b>	<b>2.37nm</b>	<b>1.84nm</b>
碱基直升	<b>0.23nm</b>	<b>0.34nm</b>	<b>0.38nm</b>
每圈碱基数	<b>11</b>	<b>10</b>	<b>12</b>
碱基倾角	<b>20°</b>	<b>0°</b>	<b>7°</b>
大沟	很窄很深	很宽较深	平坦
小沟	很宽、浅	窄、深	较窄很深

# The double helical structure of RNA resembles the A-form structure of DNA

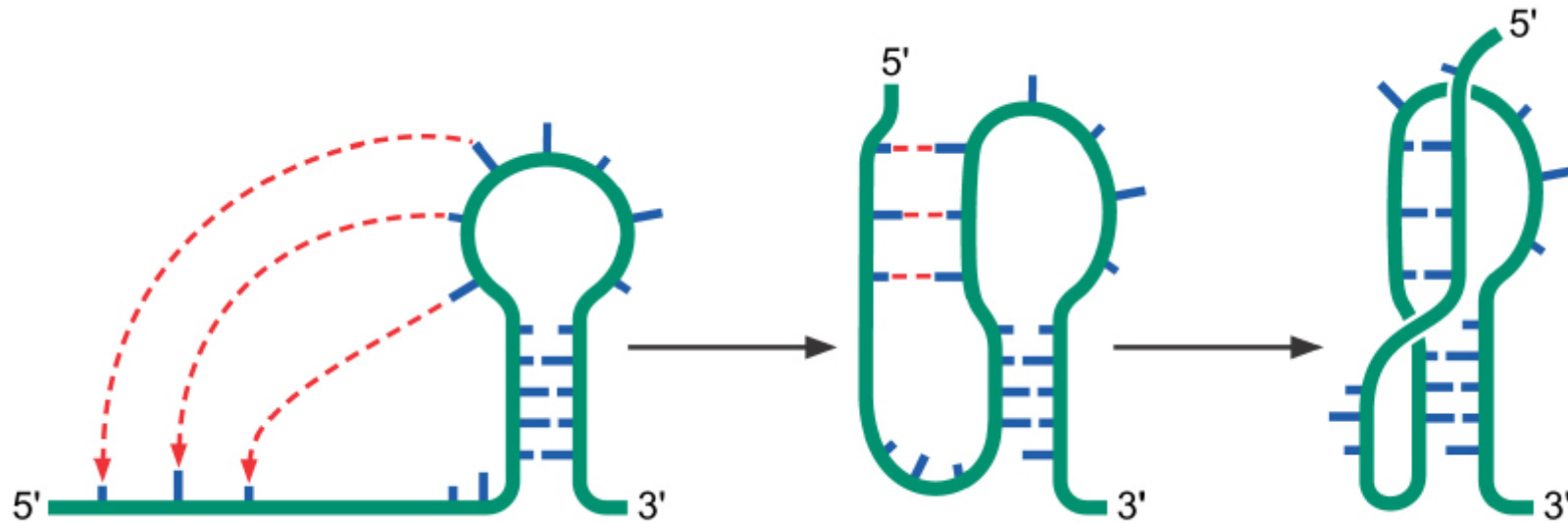
b A DNA



The **minor groove** is wide and shallow, but offers little sequence-specific information. The **major groove** is so narrow and deep that it is not very accessible to amino acid side chains from interacting proteins. Thus RNA structure is less well suited for sequence-specific interactions with proteins.



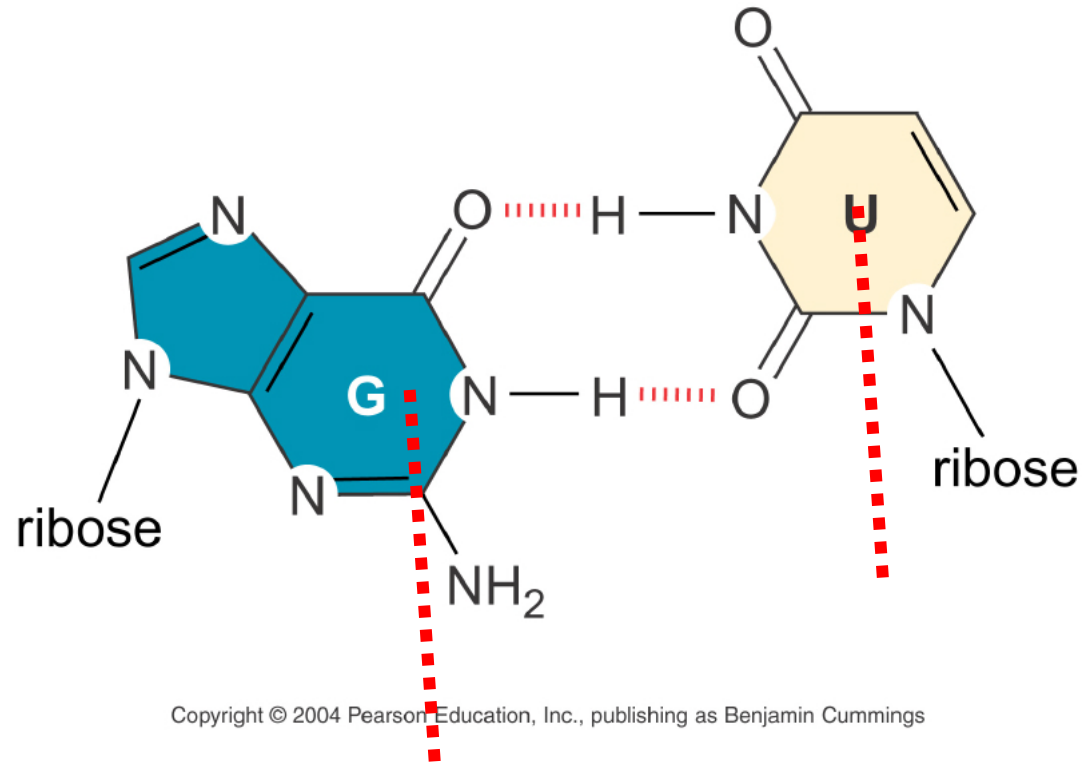
# Pseudoknot



Copyright © 2004 Pearson Education, Inc., publishing as Benjamin Cummings

**Pseudoknots are complex structure resulted from base pairing of discontinuous RNA segments**

# G:U base pair



Copyright © 2004 Pearson Education, Inc., publishing as Benjamin Cummings

**Non-Watson-Crick G:U base pairs represent additional regular base pairing in RNA, which enriched the capacity for self-complementarity**

# Schematic illustration of transcription

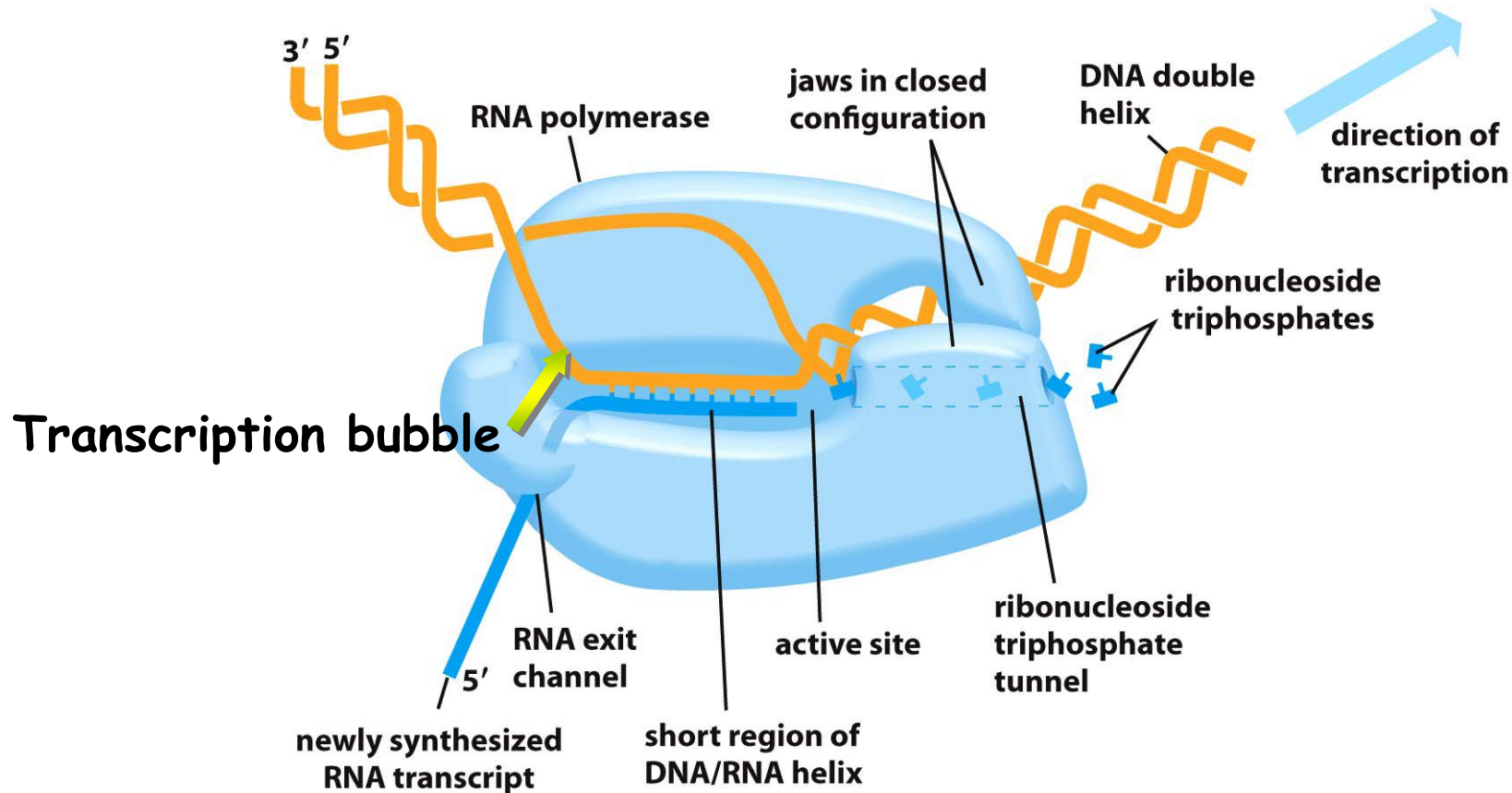


Figure 7-7 Essential Cell Biology 3/e (© Garland Science 2010)

The length of the bubble is ~12-14 bp, and the length of RNA-DNA hybrid within it is ~8-9 bp.



# 电镜下观察到的rRNA转录

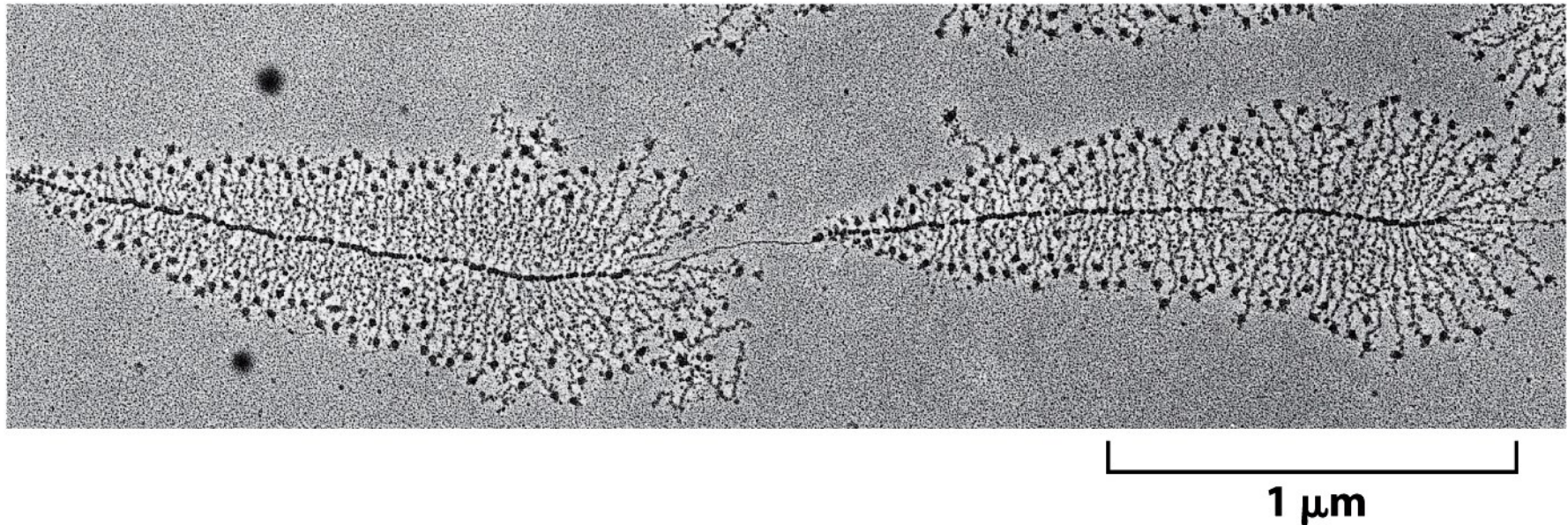
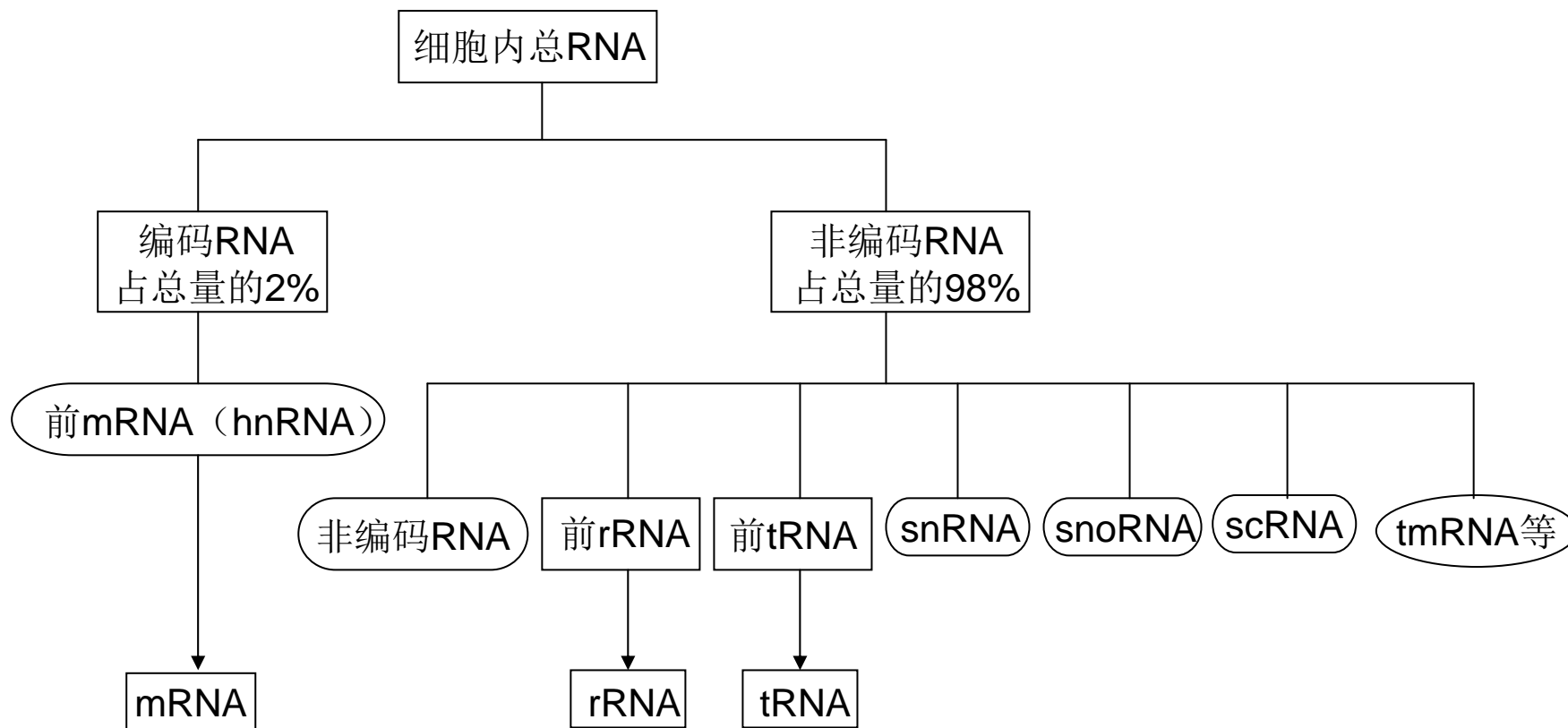


Figure 7-8 Essential Cell Biology 3/e (© Garland Science 2010)

许多RNA聚合酶分子同时转录两个相邻的基因.

# 细胞内产生的RNA种类



□ 所有生物

○ 仅真核生物

○ 仅细菌

# Functions of RNAs

- **Functions in protein synthesis**

- a. **mRNA**: as the intermediate between the gene and the protein-synthesizing machinery.

- b. **tRNA**: as an adaptor between the codons in the mRNA and amino acids.

- c. **rRNA**: play a structural role, as in the case of the RNA components of the ribosome.

- **As genetic material**

- Serving as a template for its own replication in certain viruses

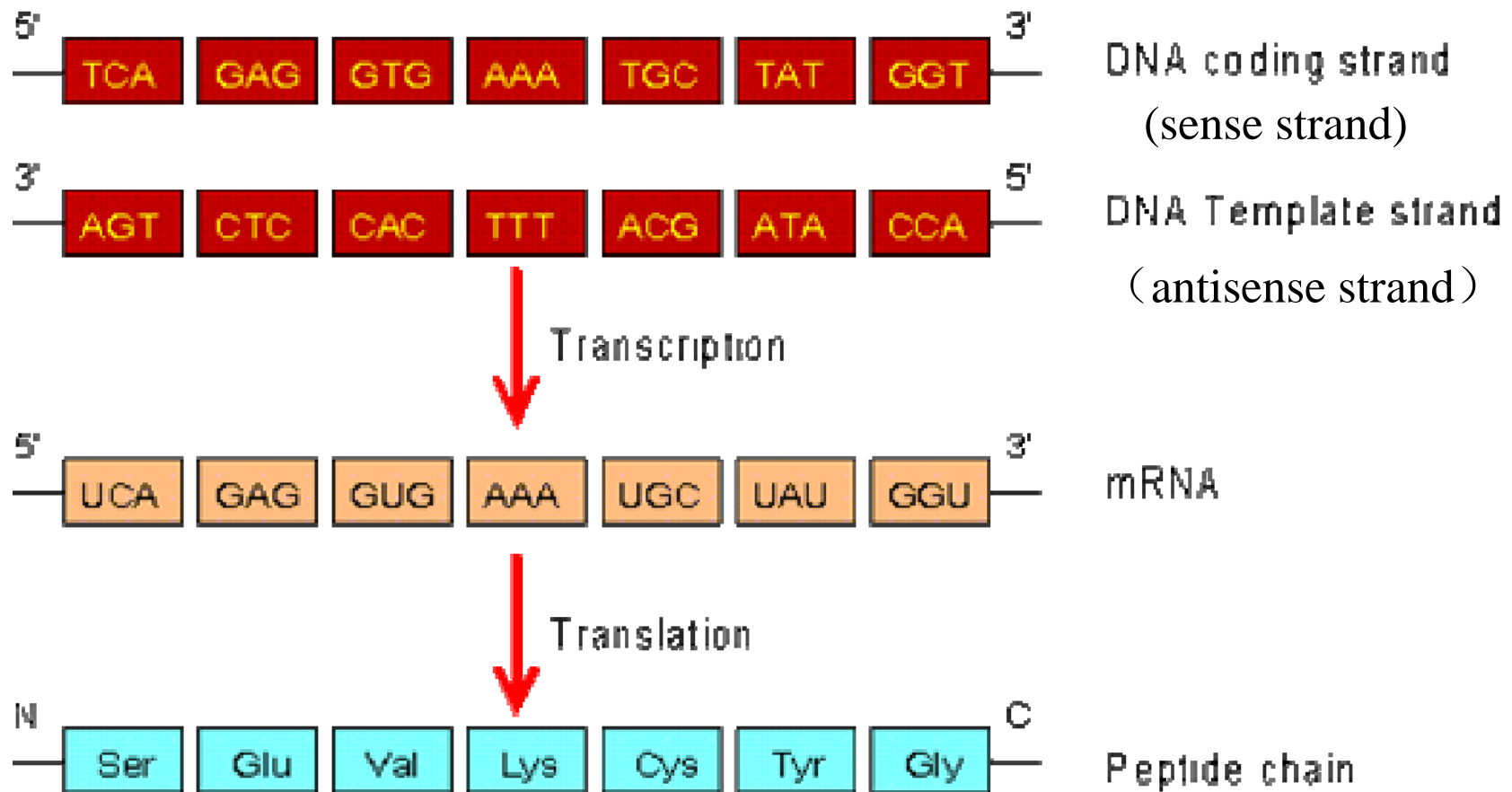
- **RNA as catalysts (ribozyme)**

- Some RNAs (including one of the structural RNAs of the ribosome) are enzymes that catalyze essential reactions in the cell.

- **RNA is a regulatory molecule**

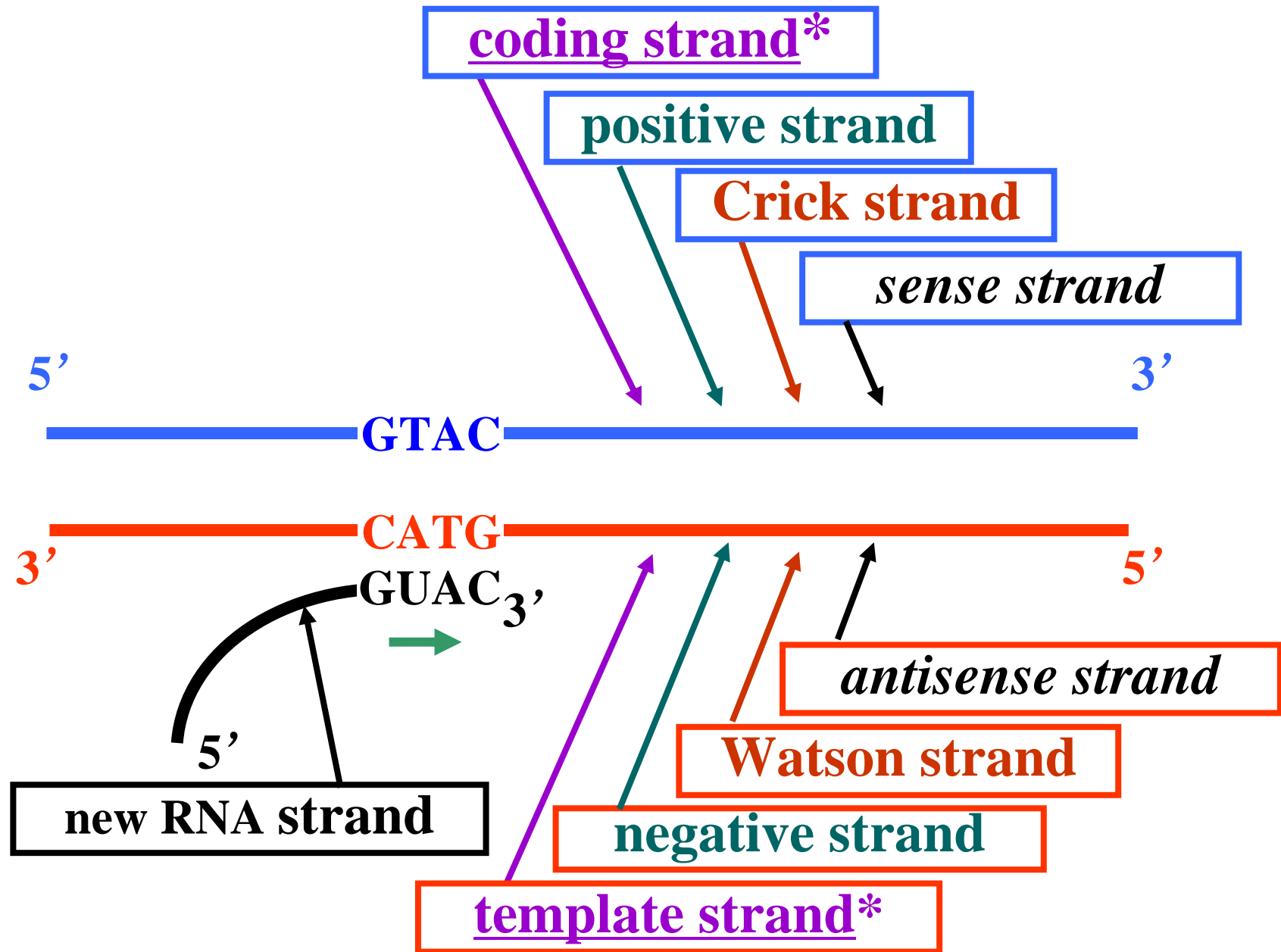
- Small non-coding RNA** which through sequence complementarity binds to, and interferes with the translation of certain mRNAs.

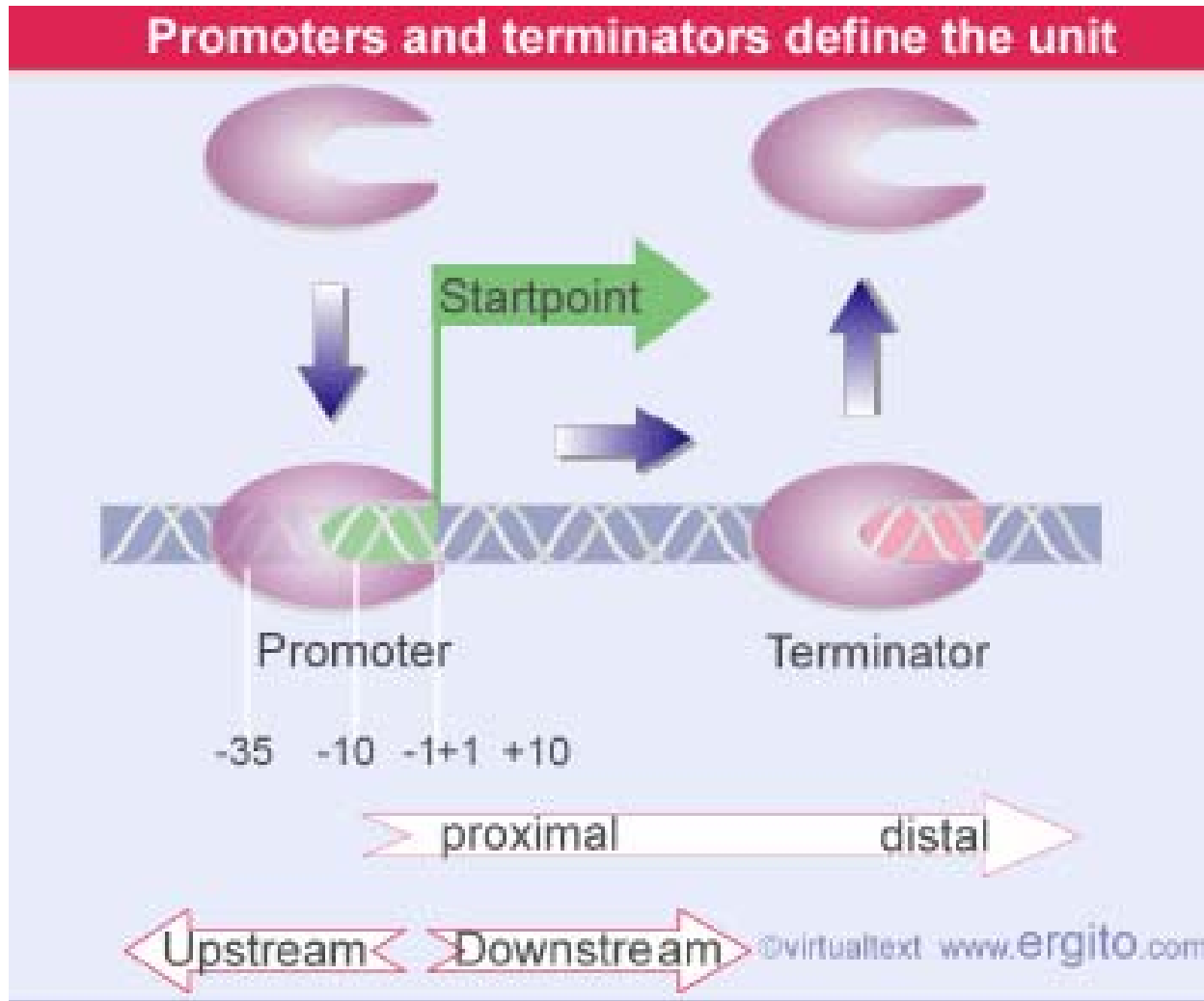
# DNA-mRNA-the encoded peptide



**编码链 (coding strand) :** 与mRNA序列相同的那条DNA链, 或称有意义链 (sense strand) 。

**模板链 (template strand) :** 根据碱基互补原则指导mRNA合成的DNA链, 或称反义链 (antisense strand) 。





A **transcription unit** is a sequence of DNA transcribed into a single RNA, starting at the **promoter** and ending at the **terminator**.

# Key terms related to transcription

**RNA polymerases** are enzymes that synthesize RNA using a DNA template (formally described as DNA-dependent RNA polymerases).

**Promoter** is a region of DNA where RNA polymerase binds to initiate transcription.

**Startpoint (Startsite)** refers to the position on DNA corresponding to the first base incorporated into RNA.

**Terminator** is a sequence of DNA that causes RNA polymerase to terminate transcription.

**Transcription unit** is the distance between sites of initiation and termination by RNA polymerase.

**Upstream** identifies sequences proceeding in the opposite direction from expression; for example, the bacterial promoter is upstream of the transcription unit, the initiation codon is upstream of the coding region.

**Downstream** identifies sequences proceeding farther in the direction of expression; for example, the coding region is downstream of the initiation codon.

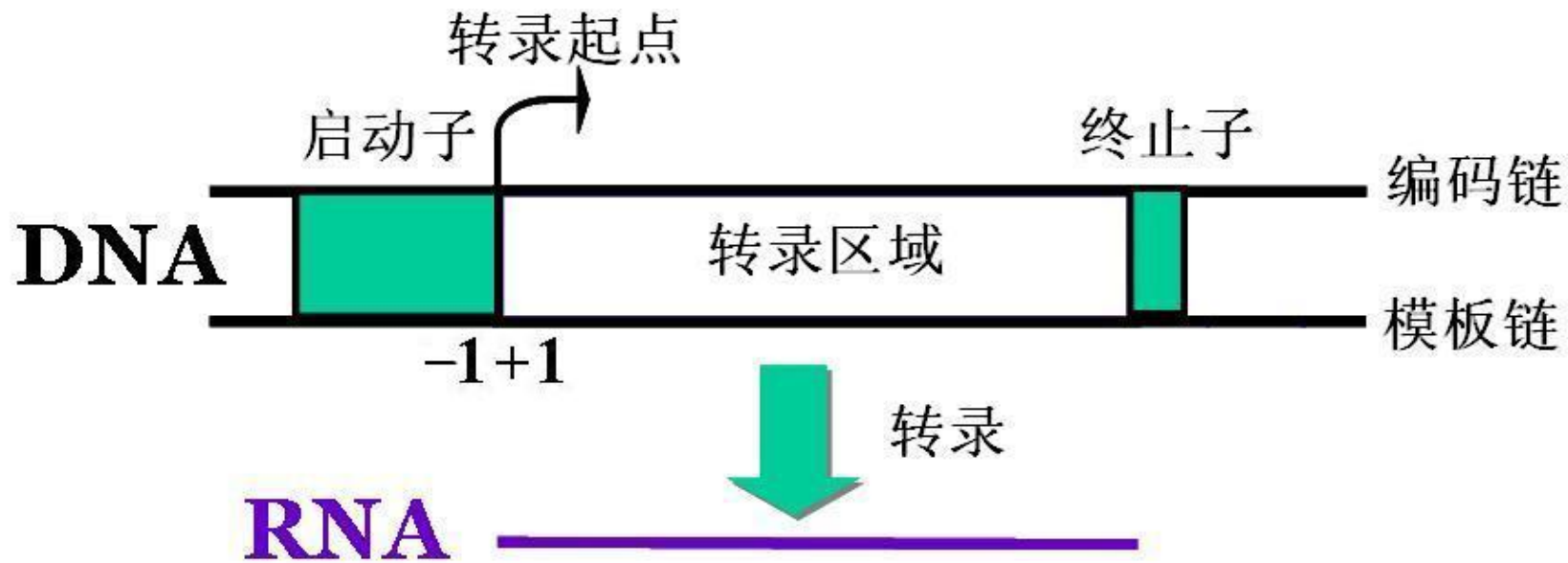
**Primary transcript** is the original unmodified RNA product corresponding to a transcription unit.



- How does RNA polymerase find promoters on DNA? ( how do proteins distinguish their specific binding sites in DNA from other sequences?)

- How do regulatory proteins interact with RNA polymerase to activate or to repress specific steps in the initiation, elongation, or termination of transcription?

# RNA的转录 (Transcription)



# RNA合成的特点

1. RNA 是以**5'→3'**方向合成的，它的序列是与DNA编码链（意义链）相同。
2. RNA 的合成是以**反义链(模板链)**为模板。
3. 同在DNA中一样，形成**磷酸二脂键 (Phosphodiester bonds)**。
4. 必需的成分：

**RNA polymerase, rNTPs, transcription factors, promoter & terminator/template**

# 转录的基本过程

- 1、模板识别 (**Template Recognition**)
- 2、转录起始 (**Initiation**)
- 3、转录的延伸 (**Elongation**)
- 4、转录的终止 (**Termination**)

# 基因序列上的信号指示细菌RNA聚合酶起始和终止转录

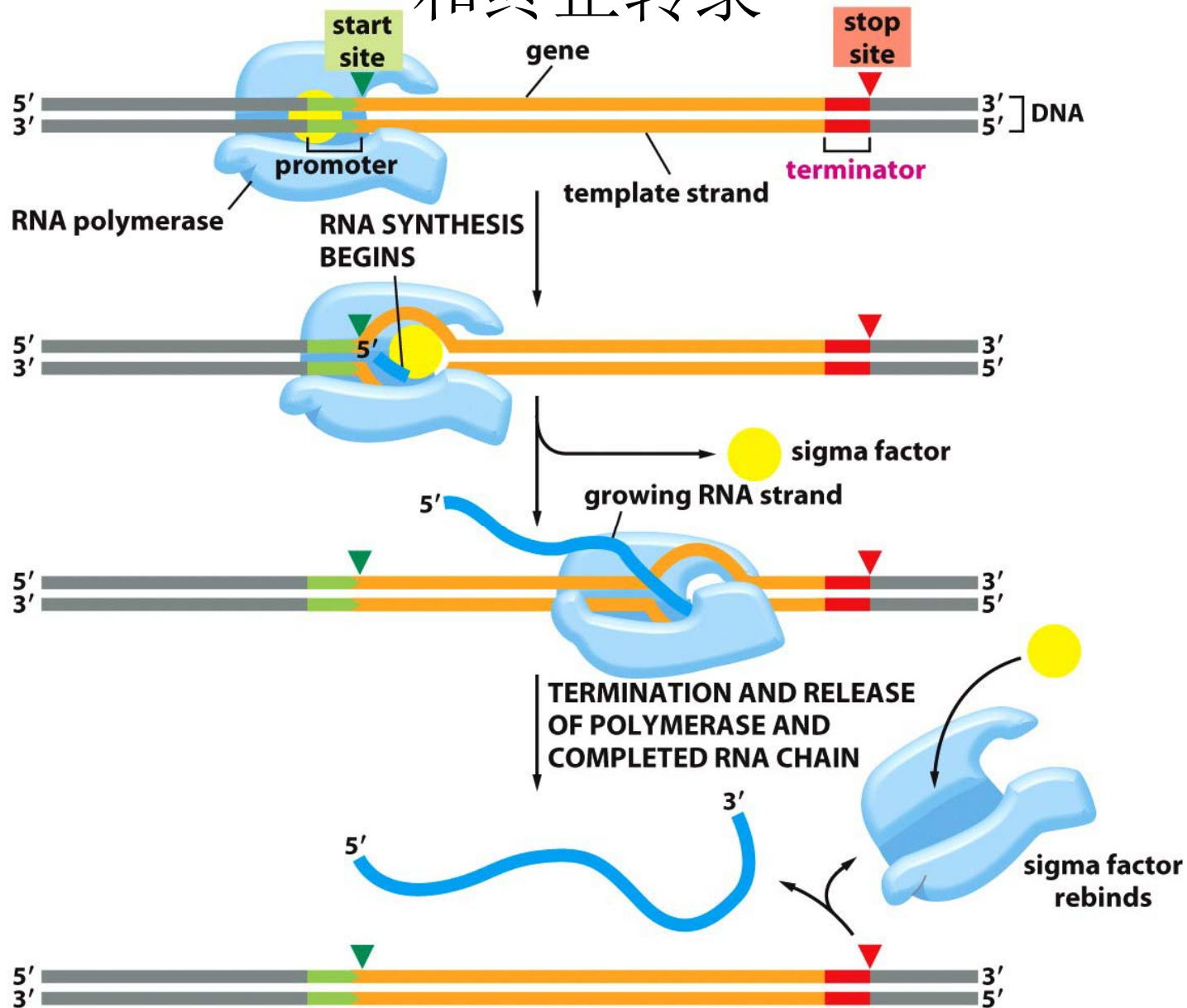
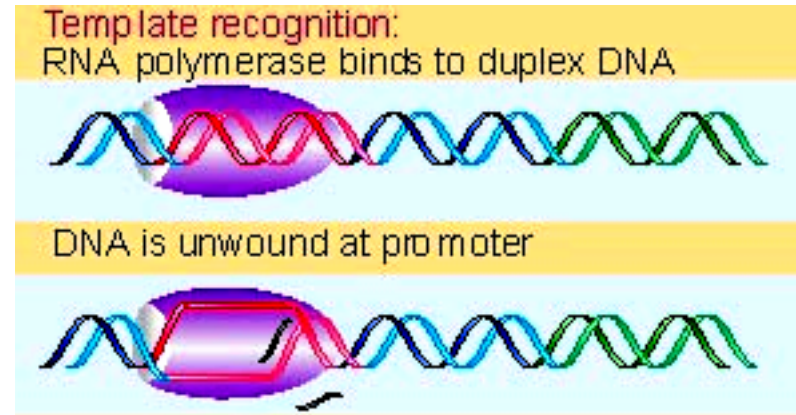


Figure 7-9 Essential Cell Biology 3/e (© Garland Science 2010)

# 模板识别



- 1、RNA聚合酶与启动子DNA双链相互作用并与之相结合的过程。
- 2、转录起始前，启动子附近的DNA双链分开形成转录泡以促使底物核糖核苷酸与模板DNA的碱基配对。

# 转录起始



转录起始就是**RNA**链上第一个核苷酸键的产生。

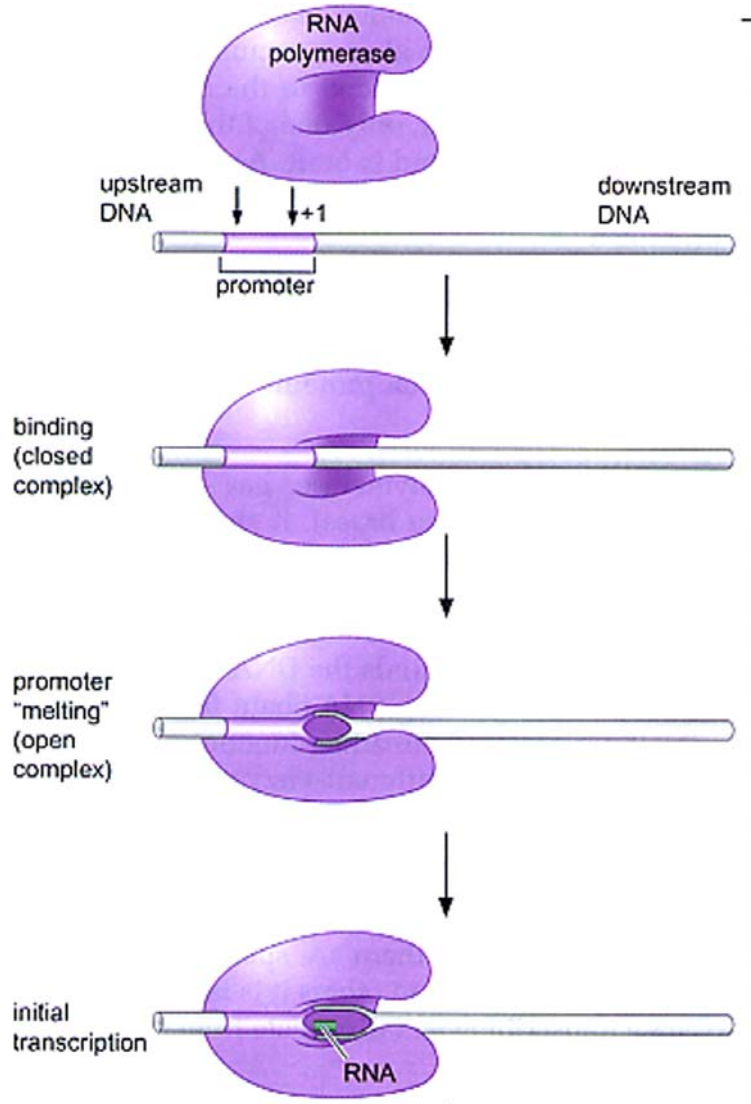
- 1.在起始位点合成**RNA**链：第一个核苷酸键的产生，该位点被称为**position +1**。
2. 转录起始后直到形成**9**个核苷酸短链是通过启动子阶段。
3. 通过启动子的时间代表一个启动子的强弱：**时间越短，转录起始的频率也越高。**

一般情况下，转录起始复合物可以进入两条不同的反应途径：

- 合成并释放2-9个核苷酸的短**RNA**转录物，即所谓的**流产式起始**；
- 一旦**RNA**聚合酶成功地合成9个以上核苷酸，尽快释放  $\sigma$  亚基，**RNA**聚合酶离开启动子区，转录就进入正常的延伸阶段。转录起始复合物通过上游启动子区并生成由核心酶、**DNA**和新生**RNA**所组成的**转录延伸复合物**。



# Transcription initiation involves three defined steps

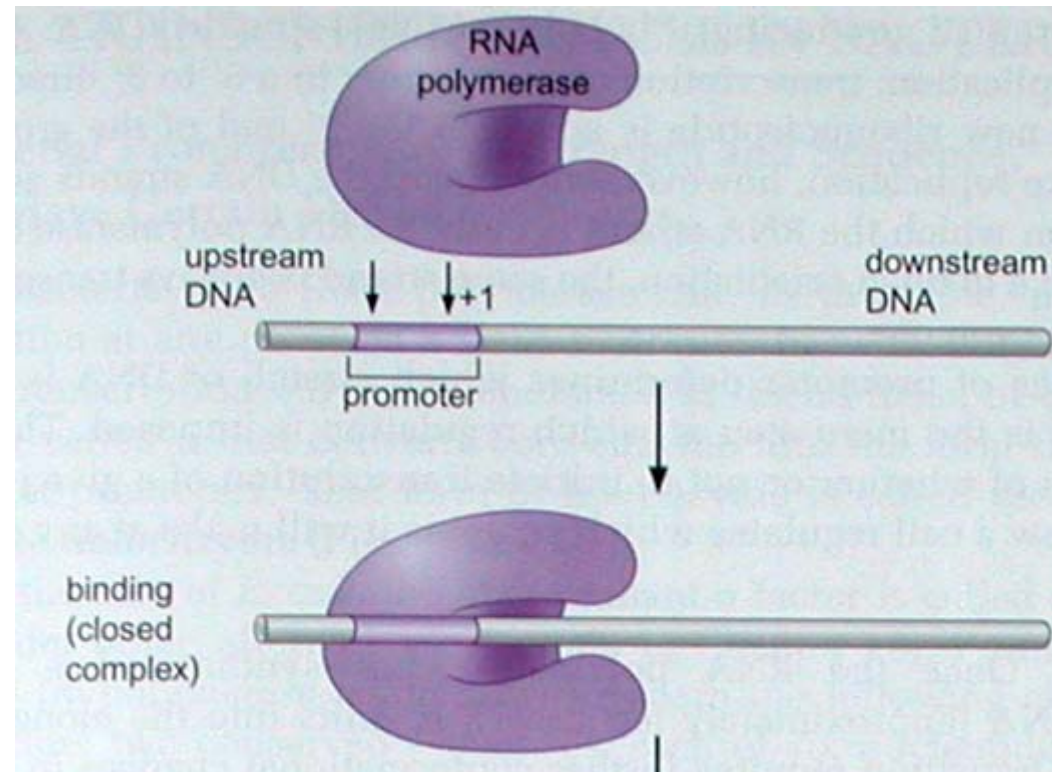


封闭复合物 (closed complex)

开放复合物 (open complex)

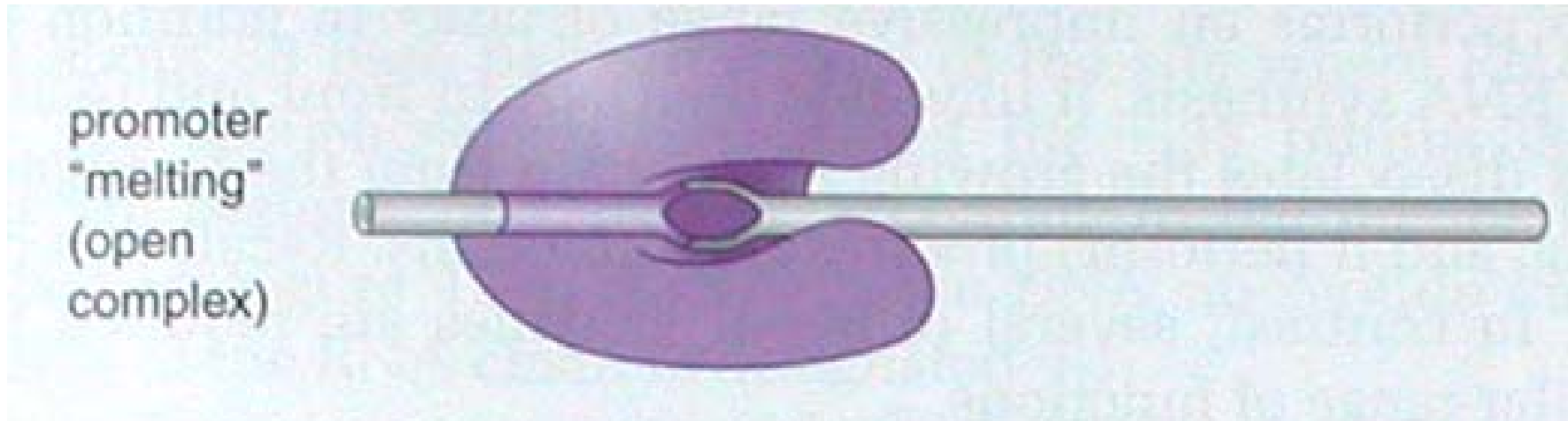
三元复合物 (ternary complex)

# 封闭复合物 (closed complex)



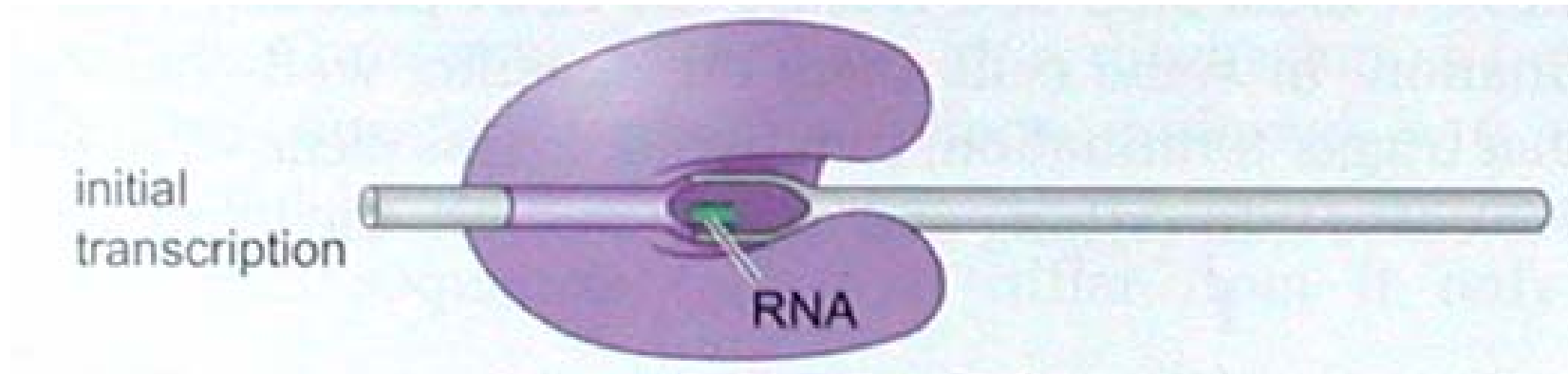
- The initial binding of polymerase to a promoter.
- DNA remains double stranded.
- The enzyme is bound to one face of the helix.

# 开放复合物 (open complex)



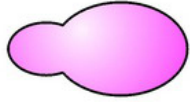




- The DNA strands separate over a distance of **~14 bp (-11 to +3)** around the **start site (+1 site)**
- Transcription bubble forms

# 三元复合物 (ternary complex)



- The enzyme **escapes** from the promoter
- The transition to the elongation phase
- **Stable ternary complex**  
= DNA + RNA + enzyme

# RNA合成的起始

反应成分	参数	状态
		全酶
聚合酶与DNA结合 	结合平衡常数 $K_B = 10^6 \sim 10^9$	二元闭合复合物
DNA开始解链 	反应速率 $k_2 = 10^{-3} \sim 10^{-1} \text{ base/s}$	二元开链复合物
转录起始 	反应速率 $k_i = 10^{-3} \text{ base/s}$	三元复合物
$\sigma$ 因子被释放 	聚合酶 离开启动区时间 $> 1 \sim 2 \text{ s}$	RNA合成开始

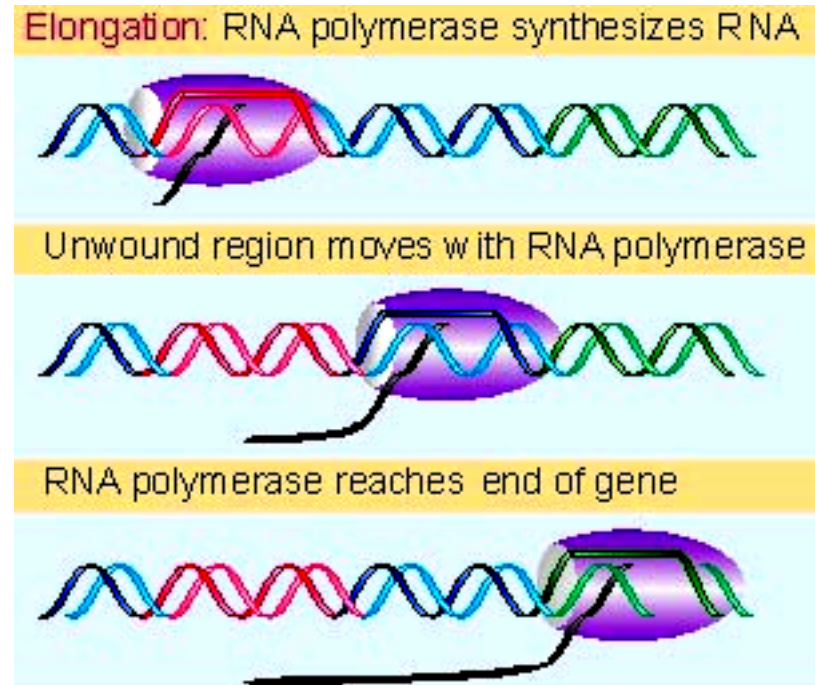
**RNA聚合酶全酶+启动子  
DNA处于双链状态**

**聚合酶全酶所结合的DNA序列中有一小段双链被解开**

**RNA聚合酶、DNA和新生RNA**

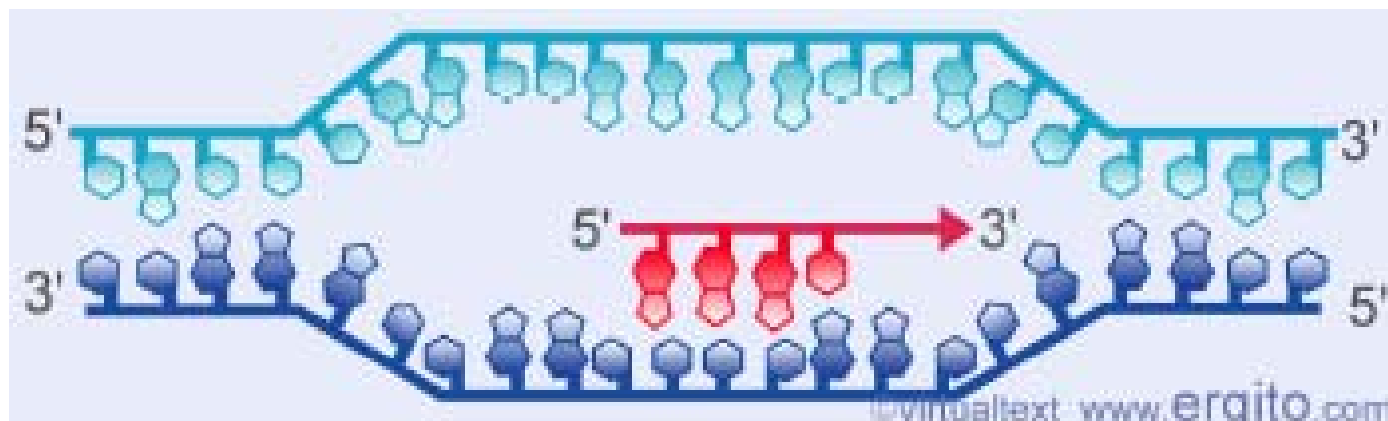
# 转录的延伸

**RNA聚合酶**离开启动子，沿**DNA链**移动并使新生**RNA链**不断伸长的过程就是**转录的延伸**。

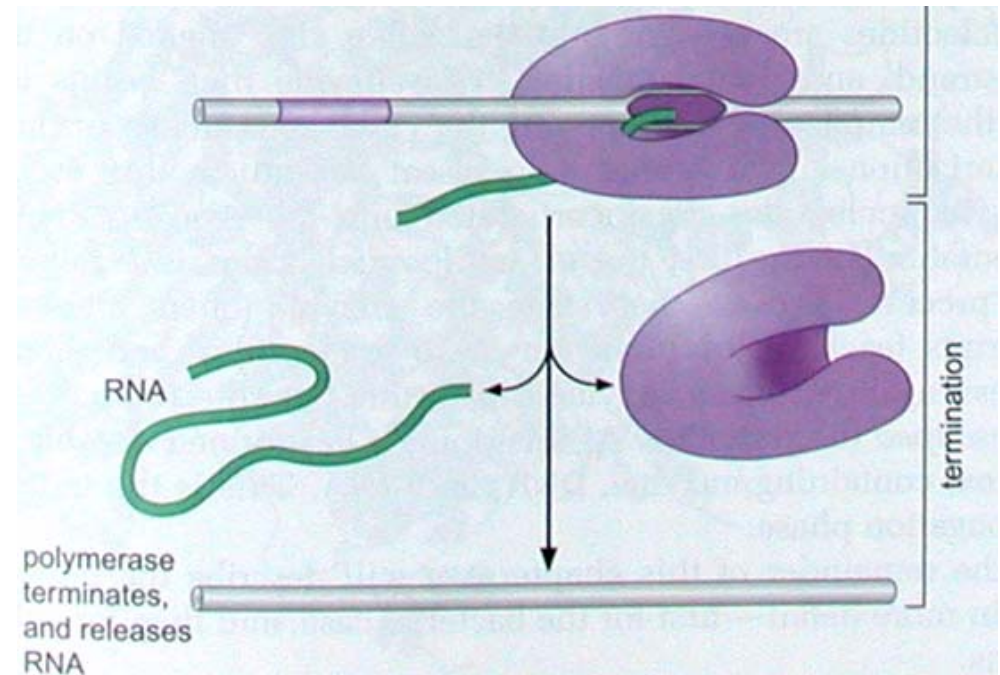


- 共价地向生长**RNA链**的**3'**端添加核糖核苷酸
- **RNA聚合酶**是以**5' → 3'**方向来延长**RNA链**
- **RNA聚合酶**本身沿着反义链以**3' → 5'**方向移动

- 大肠杆菌RNA聚合酶的活性一般为每秒50-90个核苷酸。
- 随着RNA聚合酶的移动，DNA双螺旋持续解开，暴露出新的单链DNA模板，新生RNA链的3'末端不断延伸，在解链区形成RNA-DNA杂合物。



## 转录的终止



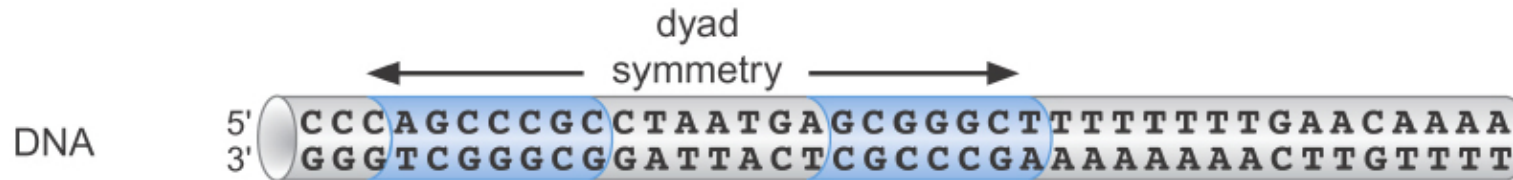
- 合成的终止：当RNA链延伸到转录终止位点时，RNA polymerase 和RNA链均从DNA模板上释放出来。
- **Terminator**: 通常含有自我互补区域(**self-complementary regions**)，在RNA产物中可以形成**stem-loop**或**hairpin**结构。



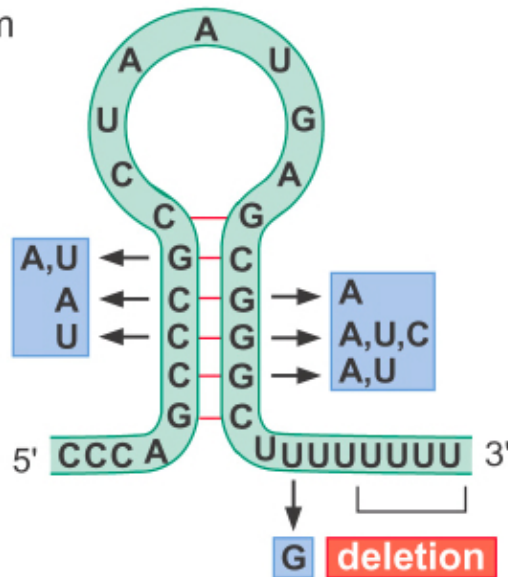
# Two types of terminators in *E. coli*

- 不依赖于  $\rho$  因子的终止  
intrinsic terminator (内在)
- 依赖于  $\rho$  因子的终止

# Rho非依赖型终止子的序列



transcript folded to form termination hairpin



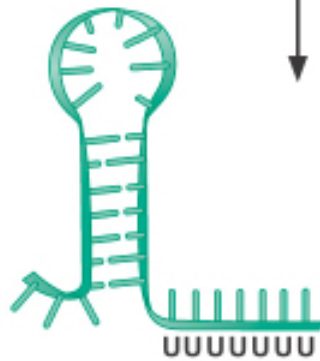
- 由两个序列原件组成：
- 一段短的反向重复序列（大约20个核苷酸）；
  - 其后是一段大约8个A:T碱基对的序列。

↓

由它转录出mRNA可形成茎环结构，可阻止RNA pol的前进。



不依赖于  $\rho$  因子的终止子 (内在终止子)



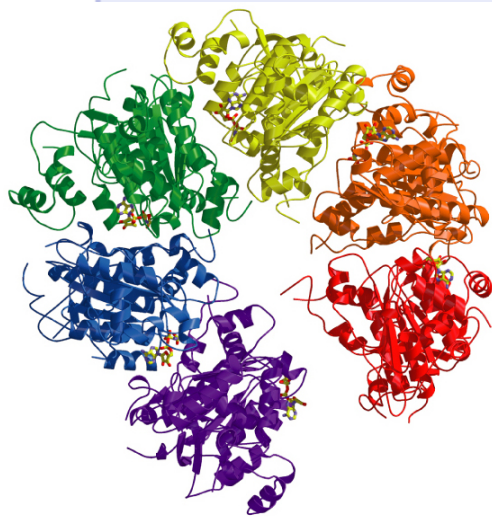
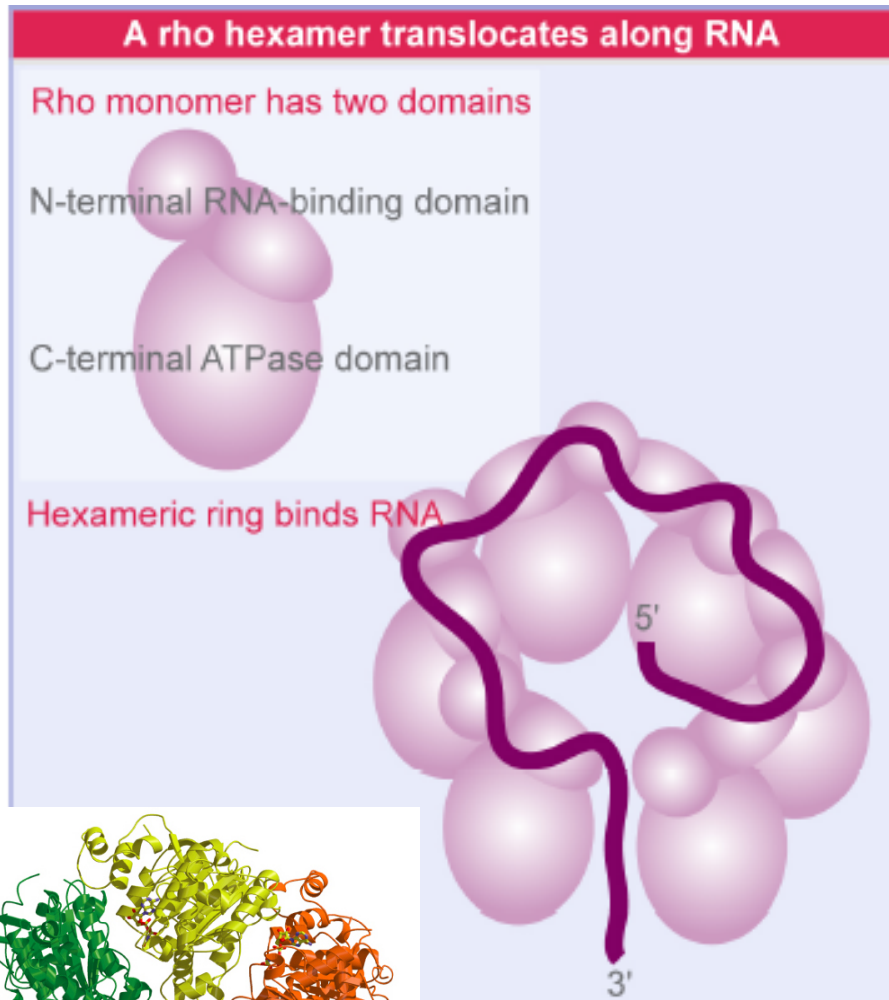
Weakest base pairing: A:U make the dissociation easier



# 依赖于 $\rho$ 因子的终止

提纯的RNA聚合酶并不能识别特异性的转录终止信号，而加入大肠杆菌  $\rho$  因子后该聚合酶就能在DNA模板上准确地终止转录。

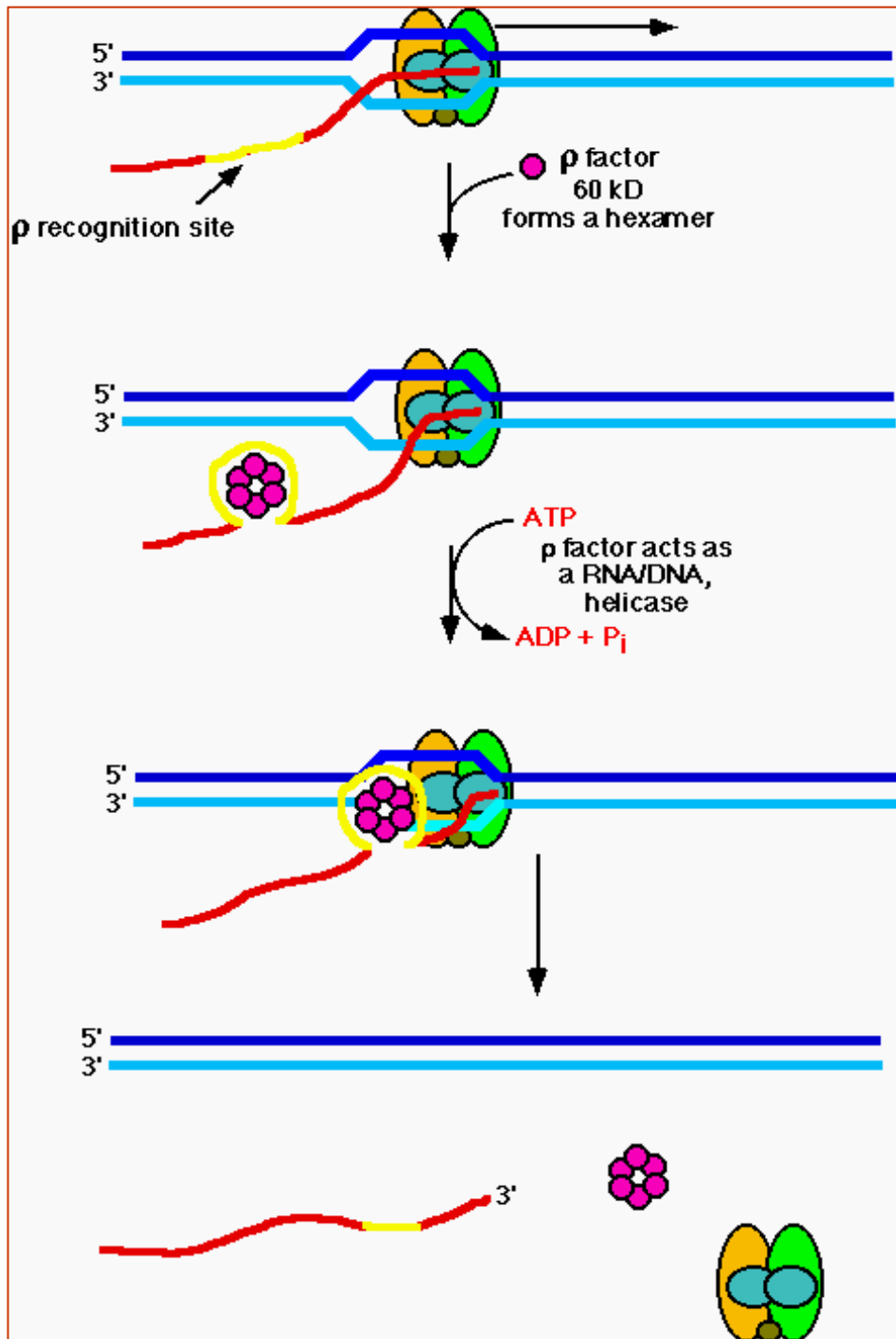
- 只含有自我互补区域，可形成茎环结构，但在茎中的G.C含量少，茎环易打开。
- 其终止需要  $\rho$  因子的参与。
- $\rho$  因子与ssRNA的特定位点结合(C丰富，G缺乏)。
- $\rho$  通过催化NTP的水解促使新生RNA链从三元转录复合物中解离。



## $\rho$ 因子

$\rho$  因子是一个相对分子质量为 $2.0 \times 10^5$ 的六聚体蛋白。

其功能是作为RNA pol的一种辅助因子，当其浓度为RNA pol浓度的10%时在体外可发挥最高的活性。



## $\rho$ 因子参与的RNA合成终止模式 “穷追” (hot pursuit) 模型

$\rho$  结合到RNA链终止子上游的某一点

$\rho$  因子结合以后延着RNA向3'端移动，跟踪聚合酶

$\rho$  追上在终止位点暂停的RNA聚合酶

终止---三元复合物解体

## 常识数据

• 转录生成mRNA的速度大约是每分钟2500个核苷酸(14个密码子/秒)，与翻译速度(15aa/秒)基本相等，但比DNA复制的速度要慢得多(800bp /秒)。

• 基因开始表达→mRNA的间隔约为2.5分钟，而再过半分钟就能在细胞内测到相应的蛋白质。

# RNA聚合酶 (RNA polymerase)

**RNA聚合酶**是转录过程中**最关键**的酶。

1. 主要以双链**DNA**为模板，以**4种核苷三磷酸**作为活性前体。
2. 需要**Mg<sup>2+</sup>/Mn<sup>2+</sup>**为辅助因子。
3. 它**不需要任何引物**。
4. 以**5' → 3'**方向合成**RNA**链。
5. 缺乏**3' → 5'**外切酶活性。
6. 是一个含有多个亚单位(**multi-subunit**)的酶。



# RNA聚合酶需执行的功能

- (1) 识别**DNA**双链上的起始子；
- (2) 使**DNA**变性在启动子处解旋成单链；
- (3) 通过阅读启动子序列，**RNA pol** 确定它自己的转录方向和模板链。
- (4) 最后当它达到终止子时，通过识别停止转录。

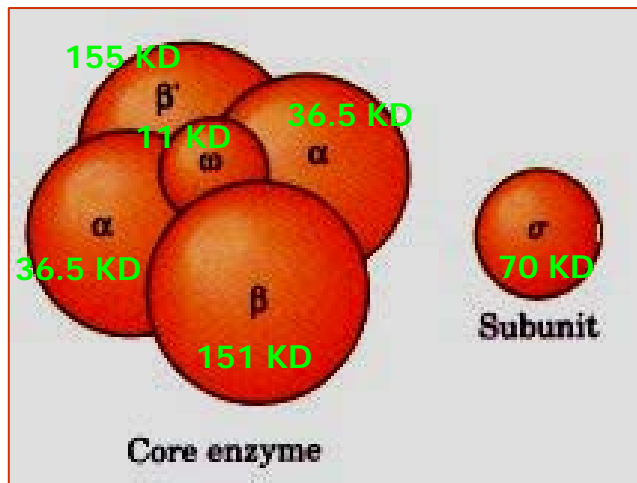
原核和真核生物的RNA聚合酶虽然都能催化RNA的合成，但在其分子组成、种类和生化特性上各有特色。

# 原核生物(*E.coli*)的RNA聚合酶

2 alpha ( $\alpha$ ) subunit,  
1 beta ( $\beta$ ) subunit,  
1 beta prime ( $\beta'$ ) subunit,  
1 omega ( $\omega$ ) subunit,  
1 sigma ( $\sigma$ ) subunit

Core  
enzyme

Holoenzyme  
聚合酶全酶



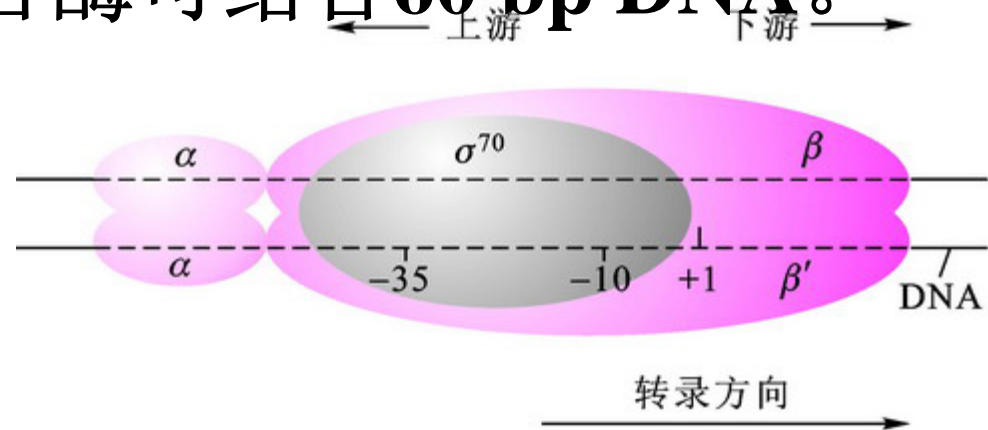
相对分子量:  $4.65 \times 10^5$

只与转录的  
起始有关

参与转录延伸

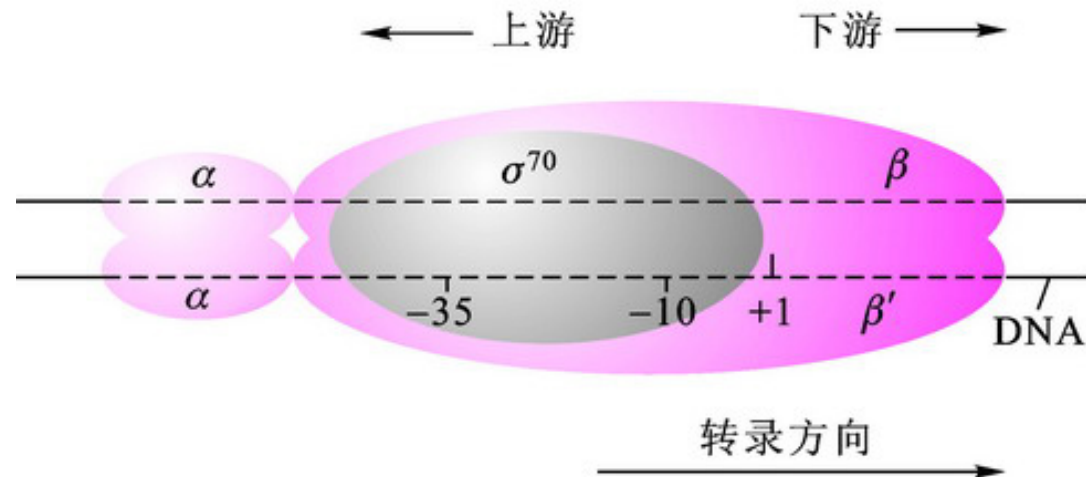
# 原核生物(*E.coli*)的RNA聚合酶

- *E. coli* 只有一个 DNA-directed RNA 聚合酶，来合成所有类型的RNA。
- 是细胞中最大的酶之一。
- 由5种subunits 组成 聚合酶全酶(holoenzyme), 包括 2  $\alpha$ , 1  $\beta$ , 1  $\beta'$ , 1  $\omega$  以及1  $\sigma$  subunits 。
- 形状象一个圆筒状通道，可以直接与16 bp DNA结合。整个聚合酶可结合60 bp DNA。
- RNA 合成速率：  
40 nt /秒, 37°C 。



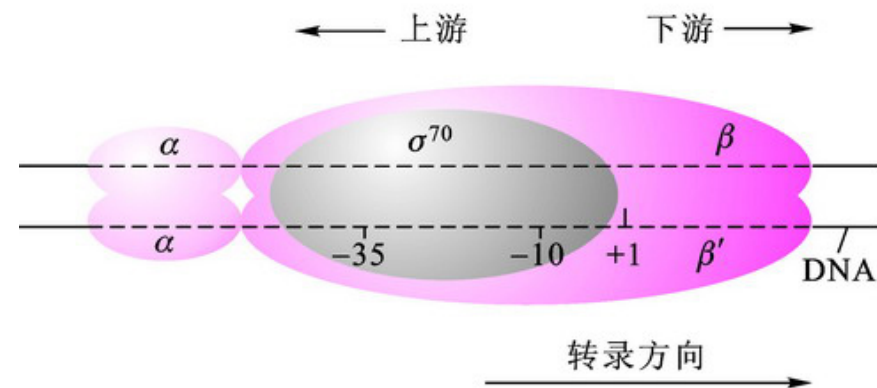
# *E. coli* RNA polymerase : $\alpha$ subunit

- 是核心酶中的两个相同的亚单位
- 由*rpoA* 基因编码
- 与核心酶的组装有关
- 参与RNA聚合酶和部分调节因子的相互作用

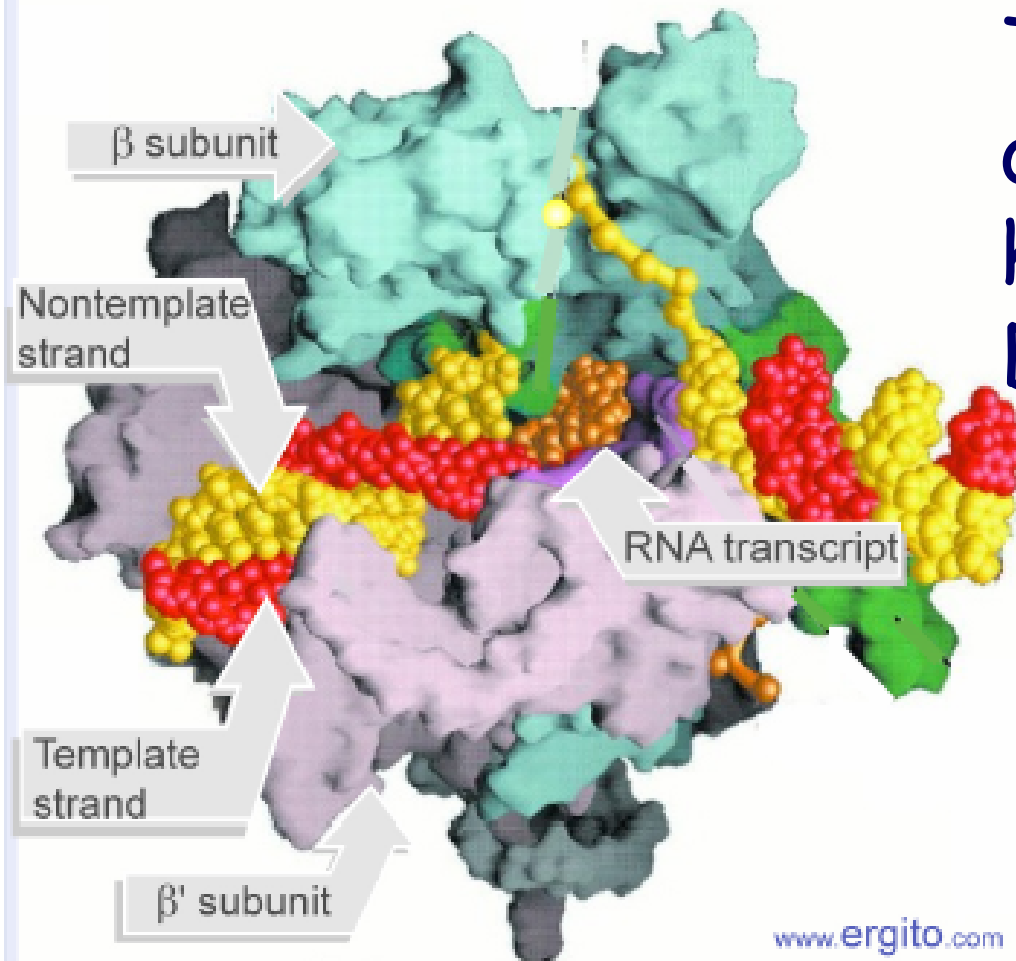


## *E. coli* RNA polymerase : $\beta$ & $\beta'$ subunit

- $\beta$  和  $\beta'$  分别由 *rpoB* 和 *rpoC* 基因编码。
- 由  $\beta$  和  $\beta'$  亚基组成了聚合酶的催化中心。
- 它们在序列上与真核生物RNA聚合酶的两个大亚基有同源性。
- $\beta$  亚基能与模板DNA、新生RNA链及核苷酸底物相结合。 *rpoB* 和 *rpoC* 基因的突变会影响转录所有的阶段。



RNA polymerase has a channel for DNA

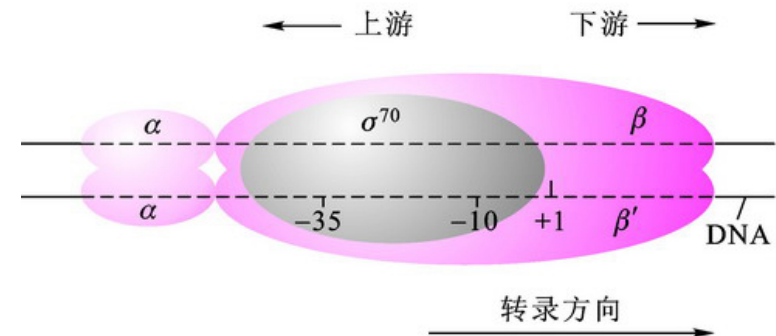


The  $\beta$  and  $\beta'$  subunit of RNA polymerase have a channel for the DNA template.

# *E. coli* RNA polymerase : $\sigma$ factor

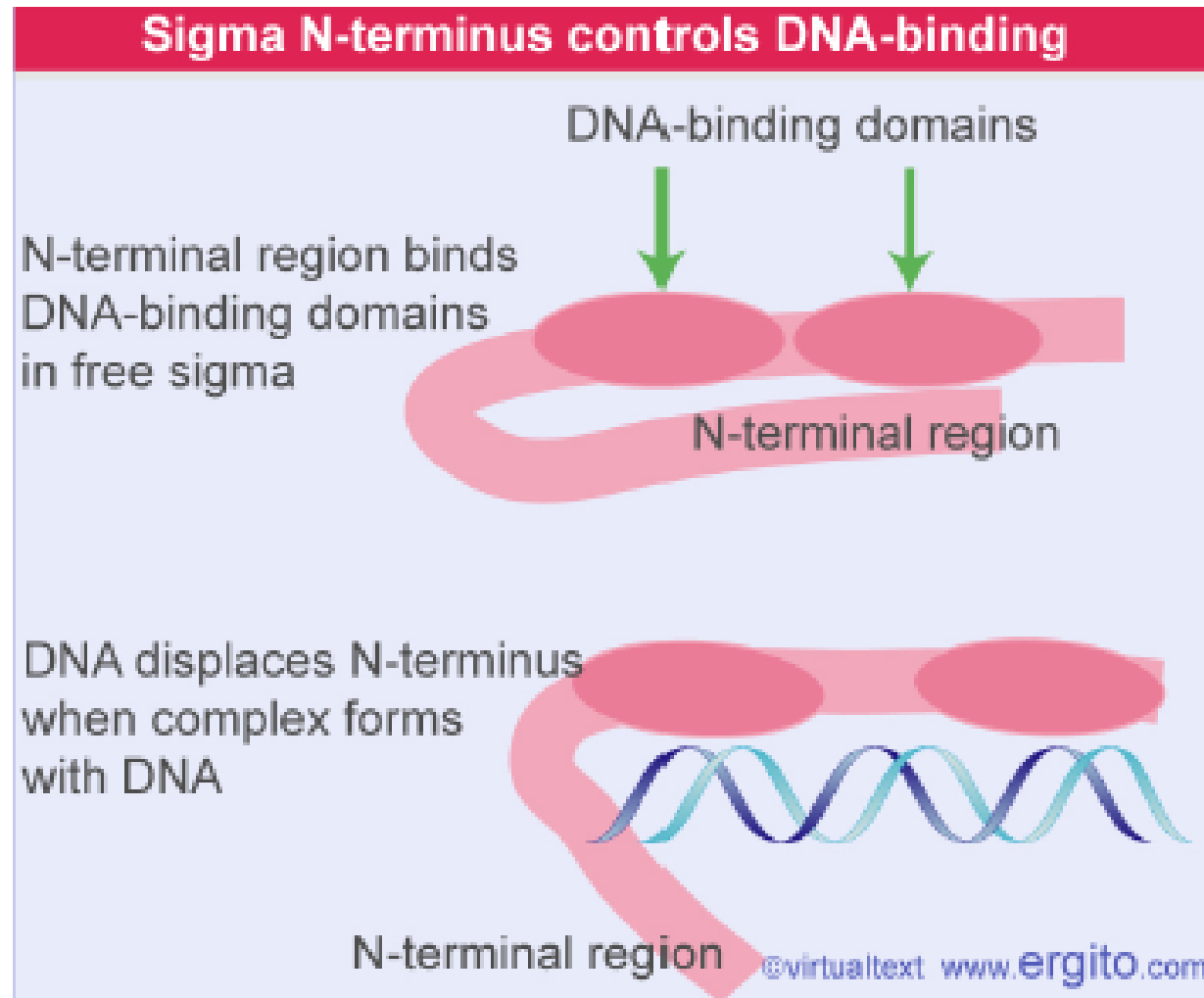
负责模板链的选择和转录的起始:

- 与  $\sigma$  因子的结合使RNA聚合酶从核心酶转变为聚合酶全酶。
- 是启动子识别的关键的酶，不仅增加聚合酶对启动子的亲和力(提高 $10^3$ 倍)，还可降低它对非专一位点的亲和力(降低 $10^4$ 倍)，使酶底复合物的半衰期小于1s。
- 在细胞中对  $\sigma$  因子量的需求少于聚合酶中其它亚单位。





# The conformational change in $\sigma$ at open complex formation



# 真核生物的RNA聚合酶

- 有**3类RNA聚合酶**；
- 结构比大肠杆菌**RNA聚合酶**复杂；
- 在细胞核中的位置不同；
- 负责转录的基因不同，对  $\alpha$ -鹅膏蕈碱的敏感性也不同。
- 真核生物**RNA聚合酶**一般有**8-14个亚基**所组成，相对分子质量超过 **$5 \times 10^5$** 。

# 真核细胞中三类RNA聚合酶特性比较

酶	细胞内定位	转录产物	相对活性	对 $\alpha$ -鹅膏蕈碱的敏感程度
RNA聚合酶I	核仁	rRNA (28S,18S, 5.8S)	50%- 70%	不敏感
RNA聚合酶II	核质	hnRNA*, snRNA, mRNA	20%- 40%	敏感
RNA聚合酶III	核质	tRNA 5SRNA 某些涉及 RNA加工的 snRNA	约 10%	存在物种 特异性

\*hnRNA: heterogeneous nuclear RNA,核内不均一RNA, RNA的前体

## 转录的抑制剂

抑制剂	靶酶	抑制作用
利福霉素	细菌全酶	和 $\beta$ 亚基结合，抑制起始
链霉溶菌素	细菌核心酶	和 $\beta$ 亚基结合，抑制起始
放射线素D	真核Po1 I	和DNA结合，阻止延伸
$\alpha$ -鹅膏蕈	真核Po1 II	和RNA Po1 II结合

**RNA合成抑制剂主要分两类：**

- 1. 模板结合抑制剂**
- 2. 聚合酶抑制剂**

- 在动、植物及昆虫的细胞中，RNA pol II 的活性可被低浓度的  $\alpha$ -鹅膏蕈碱所抑制。但却不抑制 pol I。
- Pol III 对  $\alpha$ -鹅膏蕈的反应，不同的生物有所差异。在动物细胞中高浓度的  $\alpha$ -鹅膏蕈可抑制转录，在昆虫中不受抑制。

# 真核生物RNA聚合酶的主要特征

- 聚合酶中有两个相对分子质量超过 $1 \times 10^5$ 的大亚基；

- 同种生物3类聚合酶有“共享”小亚基的倾向，即有几个小亚基是其中3类或2类聚合酶所共有的。

# 真核生物RNA聚合酶的亚基

RNA聚合酶I	RNA聚合酶II	RNA聚合酶III
RPA 1	RPB 1 ( $\beta'$ )	RPC 1
RPA 2	RPB 2 ( $\beta$ )	RPC 2
RPC 5	RPB 3 ( $\alpha'$ )	RPC 5
RPC 9	RPB 11 ( $\alpha''$ )	RPC 9
RPB 6	RP 6 ( $\omega$ )	RPB 6
其它9个亚基	其它7个亚基	其它11个亚基

注：亚基按照分子量由大到小的顺序排列。

真核生物RNA聚合酶一般有**8-16**个亚基所组成, 相对分子质量超过 **$5 \times 10^5$** 。

# 真核生物线粒体和叶绿体中存在不同的**RNA**聚合酶

## 线粒体中**RNA**聚合酶:

- 只有一条多肽链，相对分子量小于 **$7 \times 10^4$** ，是已知最小的**RNA**聚合酶之一；
- 与**T7**噬菌体**RNA**聚合酶有同源性。

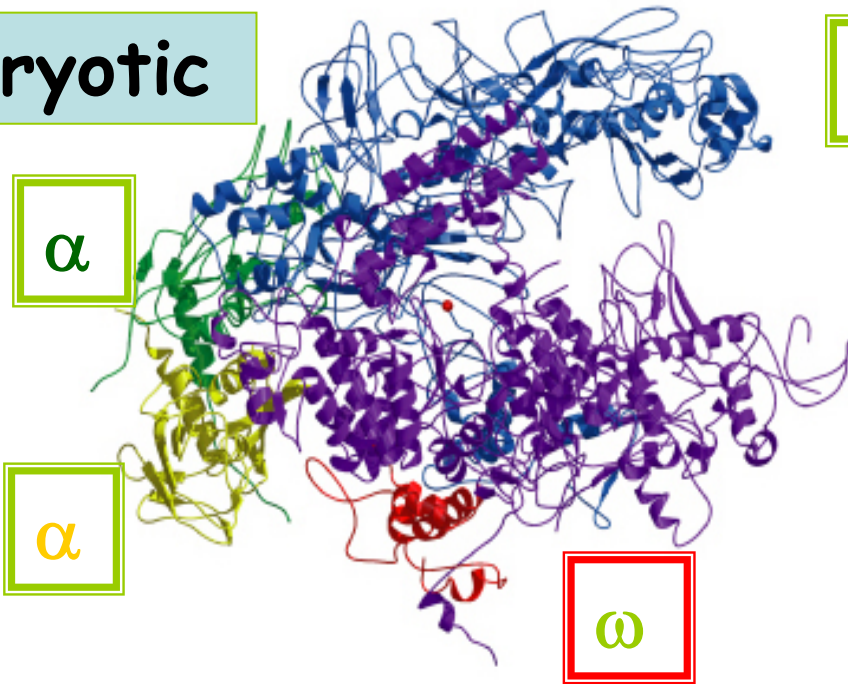
## 叶绿体中**RNA**聚合酶:

- 比较大；
- 结构上与细菌中的聚合酶相似，由多个亚基组成，部分亚基由叶绿体基因编码。

线粒体和叶绿体**RNA**聚合酶活性不受  $\alpha$ -鹅膏蕈碱所抑制。



prokaryotic



$\beta'$

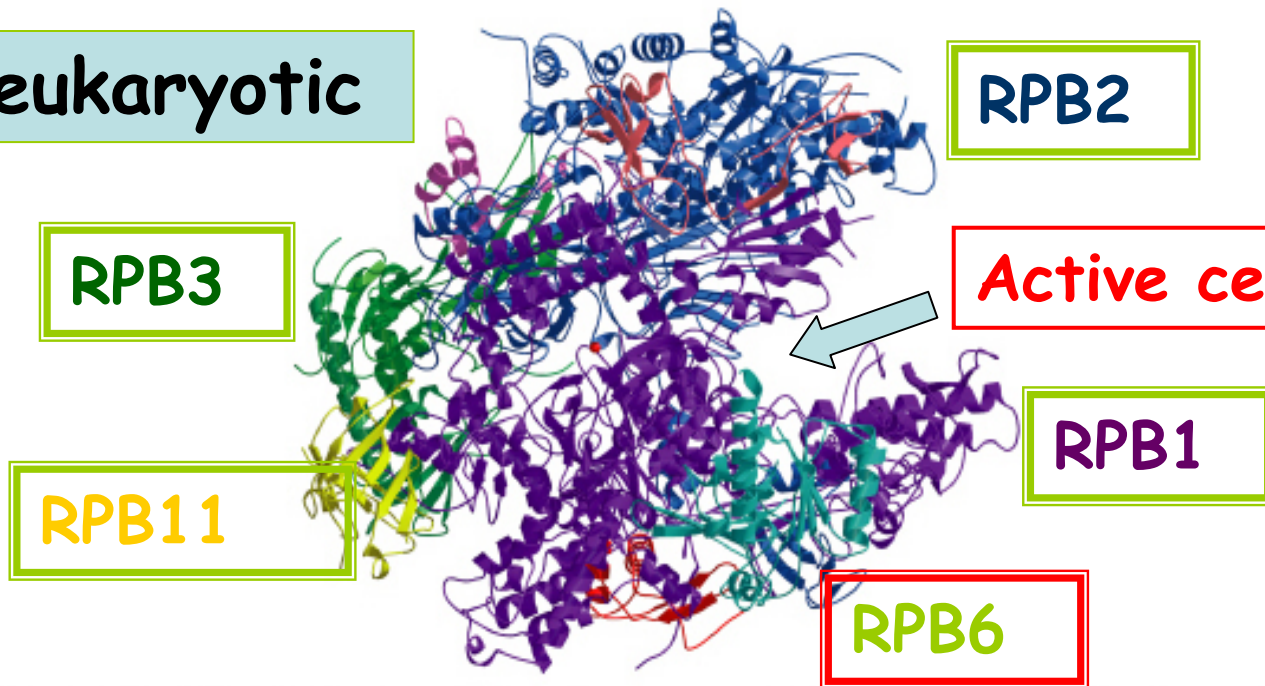
“Crab claw”  
shape of RNAP

$\beta$

The same color indicate  
the homologous of the  
two enzymes

$\omega$

eukaryotic



RPB2

Active center cleft

RPB1

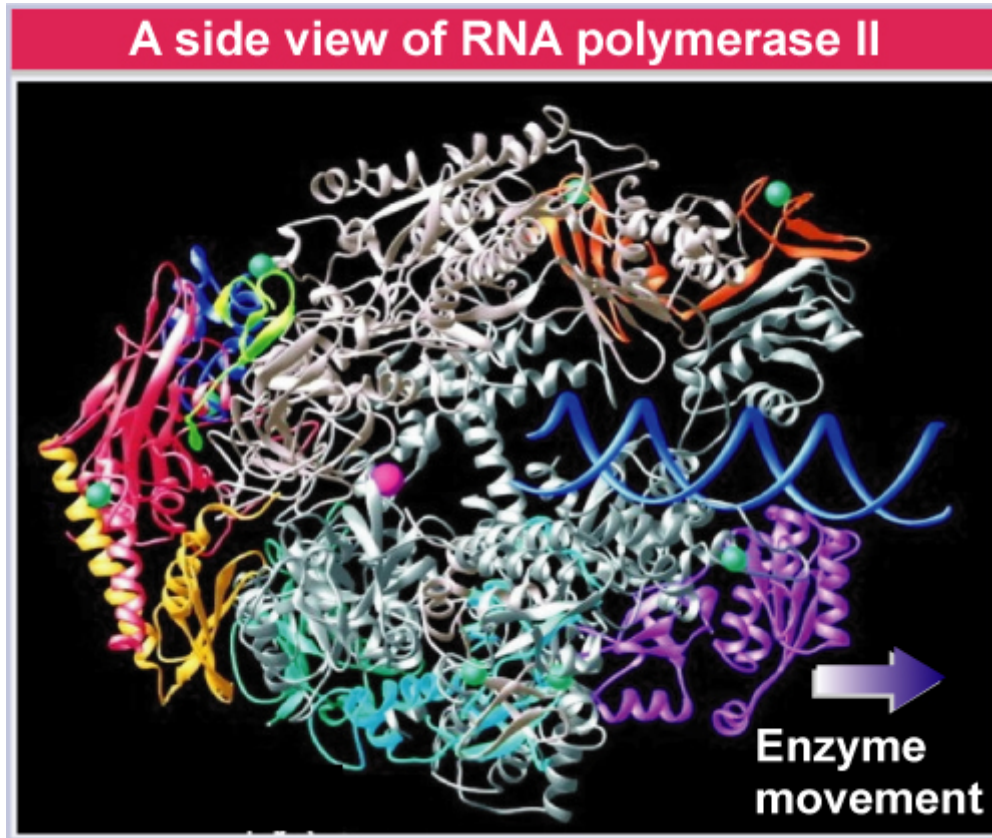
RPB6

RPB3

RPB11



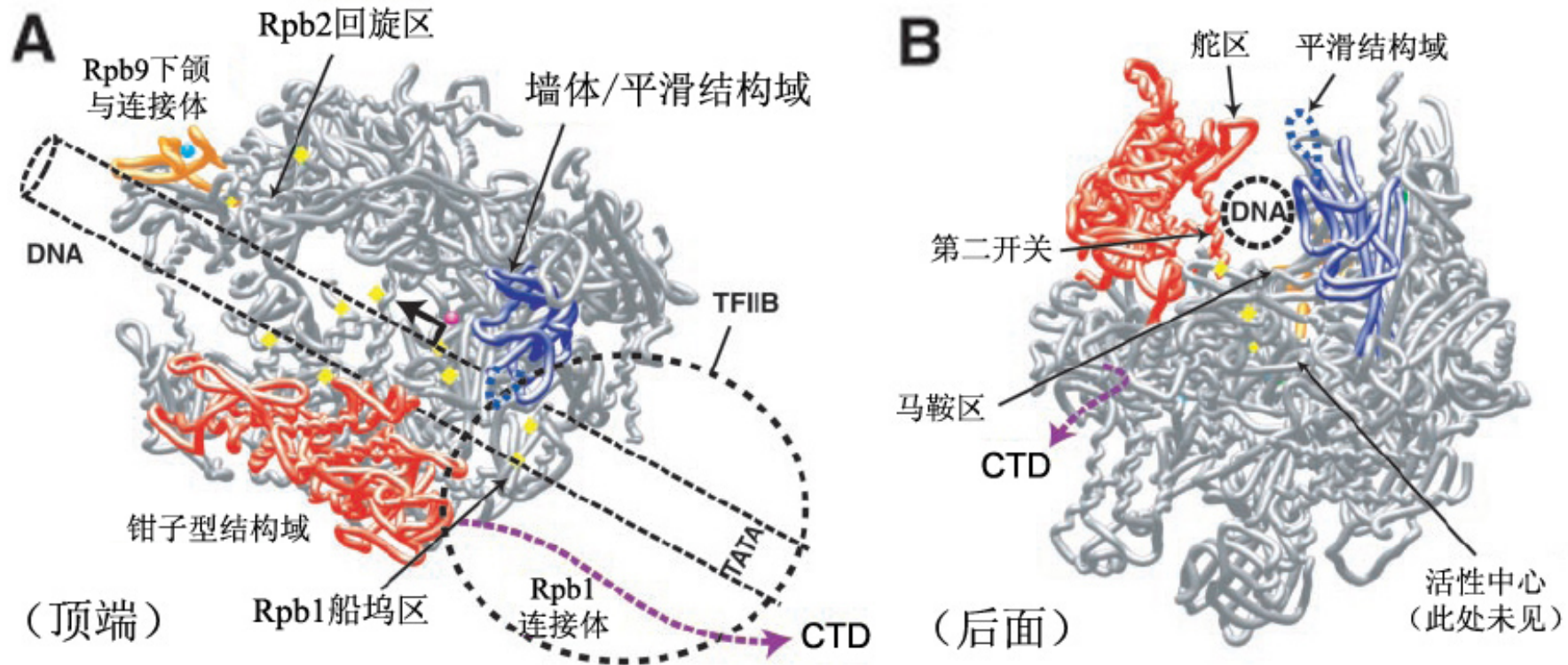
# 酵母RNA聚合酶II晶体结构



By Roger Kornberg



# 模板DNA进入转录复合体的路径



There are **various channels** allowing DNA, RNA and ribonucleotides (rNTPs) into and out of the enzyme's active center cleft

# RNA polymerase/transcription and DNA polymerase/replication

	RNA pol	DNA pol
Template	dsDNA	dsDNA
Require primer	No	Yes
Initiation	promoter	origin
elongation	40 nt/ sec	900 bp/sec
Exonuclease activity	No	Yes
terminator	Synthesized RNA	Template DNA

# 前起始复合物

(preinitiation transcription complex, PIC)

**真核生物**RNA聚合酶不能直接识别基因的启动子区，需要一些被称为**转录调控因子**的辅助蛋白质按特定顺序结合于启动子上，**RNA聚合酶**才能与之相结合并形成复杂的**前起始复合物(PIC)**以保证有效地起始转录。

# 真核生物RNA聚合酶II需形成转录起始复合物

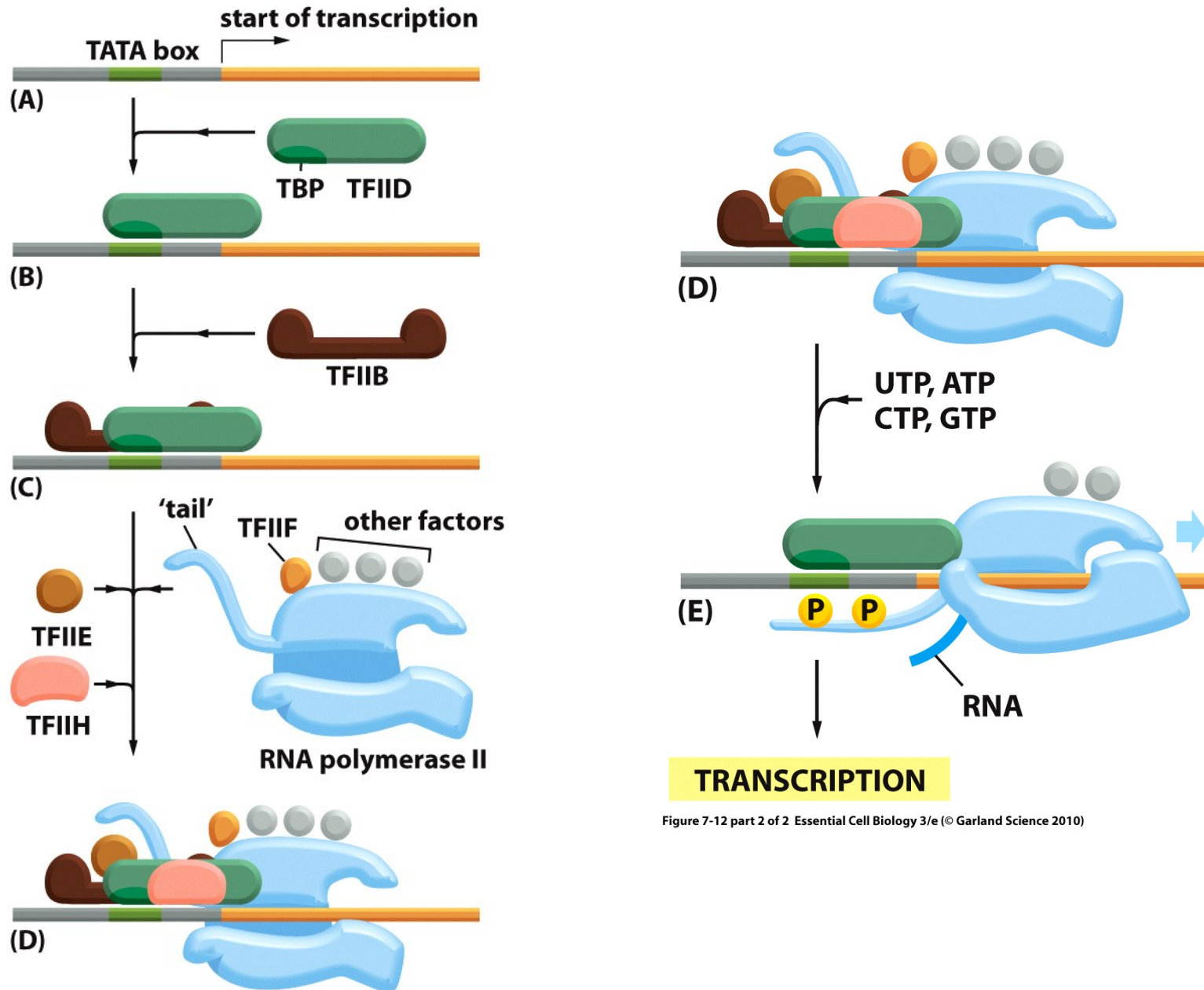
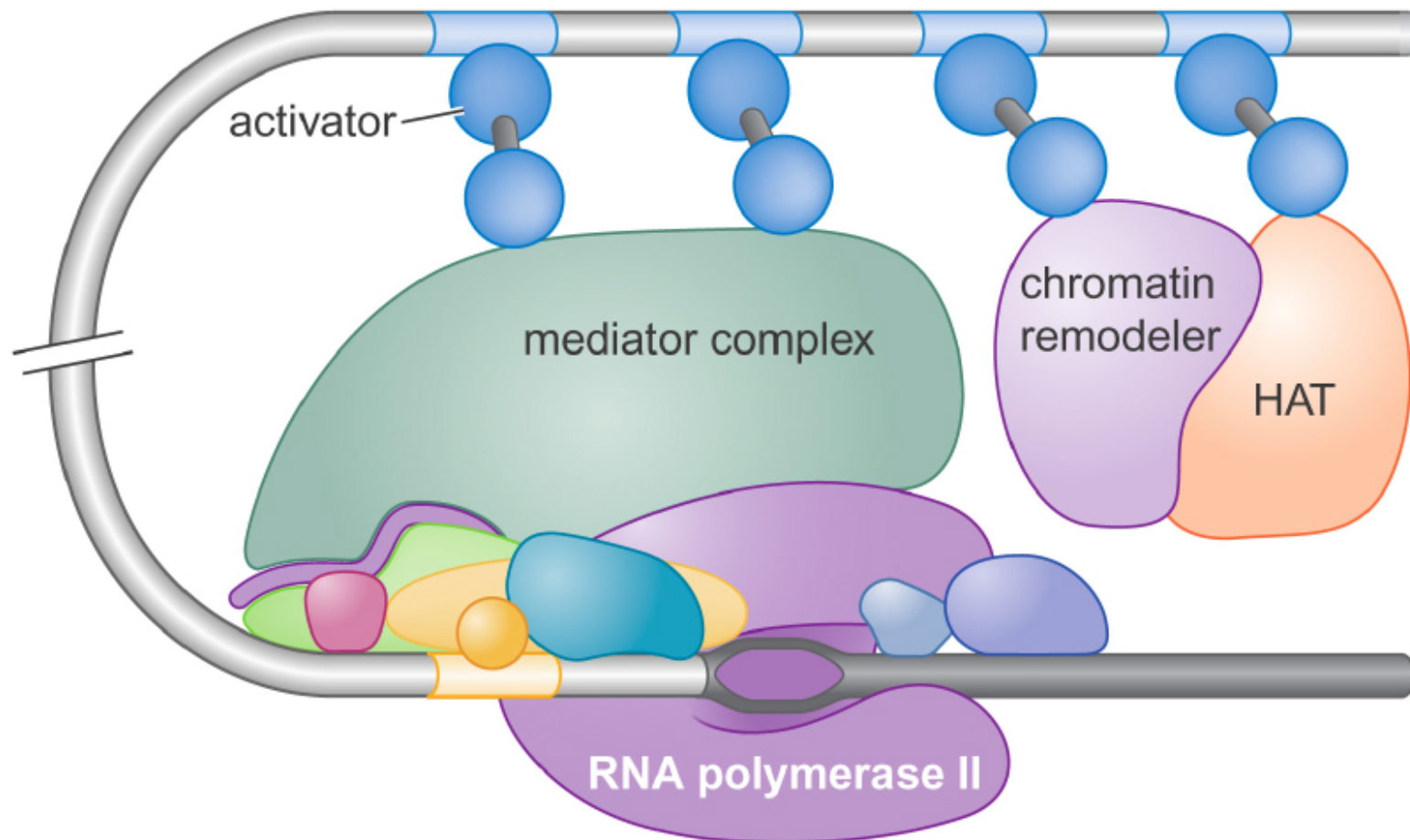
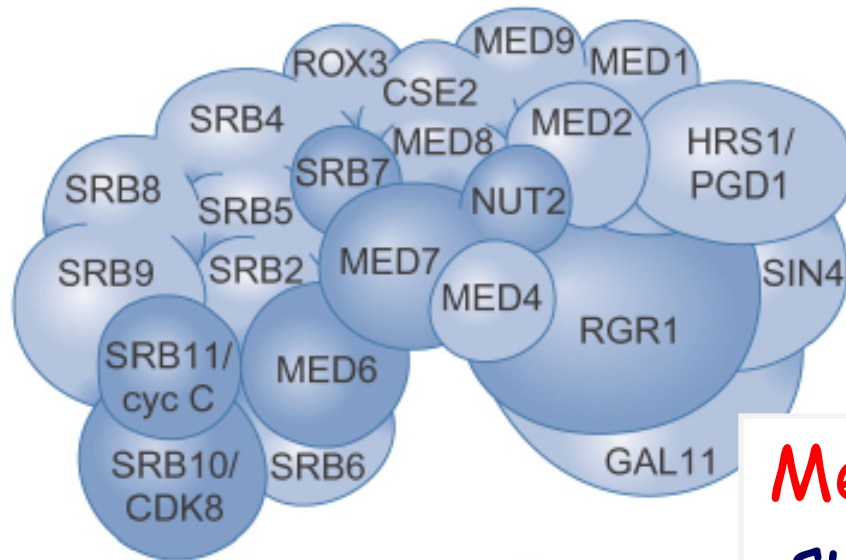


Figure 7-12 part 2 of 2 Essential Cell Biology 3/e (© Garland Science 2010)

Figure 7-12 part 1 of 2 Essential Cell Biology 3/e (© Garland Science 2010)

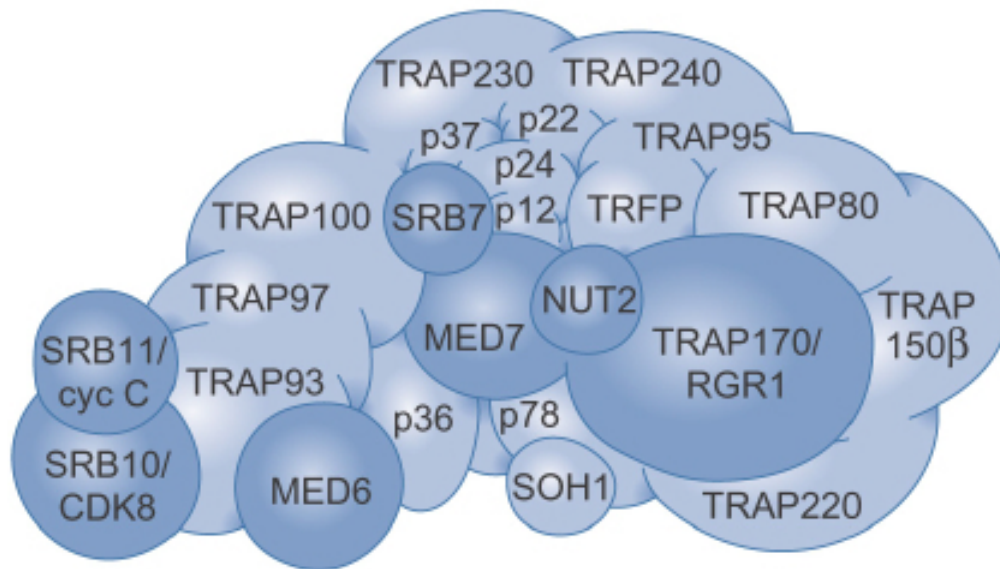
# Assembly of the pre-initiation complex in presence of mediator, nucleosome modifiers and remodelers, and transcriptional activators





yeast mediator

**Mediator** consists of many subunits, some conserved from yeast to human.



human mediator

Only subunit **Srb4** is essential for transcription of essentially all Pol II genes *in vivo*.



# The molecular basis of eukaryotic transcription

Roger D. Kornberg\*

Stanford University School of Medicine, Stanford, CA 94305-5400

I am deeply grateful for the honor bestowed on me by the Nobel Committee for Chemistry and the Royal Swedish Academy of Sciences. It is an honor I share with my collaborators. It is also recognition of the many who have contributed over the past quarter century to the study of transcription.

## The Nucleosome

My own involvement in studies of transcription began with the discovery of the nucleosome, the basic unit of DNA coiling in eukaryote chromosomes (1). X-ray studies and protein chemistry led me to propose the wrapping of DNA around a set of eight histone molecules in the nucleosome (Fig. 1). Some years later, Yuhli Lorch and I found that this wrapping of DNA prevents the initiation of transcription *in vitro* (2). Michael Grunstein and colleagues showed nucleosomes interfere with transcription *in vivo* (3). The nucleosome serves as a general gene repressor. It assures the inactivity of all of the many thousands of genes in eukaryotic cells except those whose transcription is brought about by specific positive regulatory mechanisms.

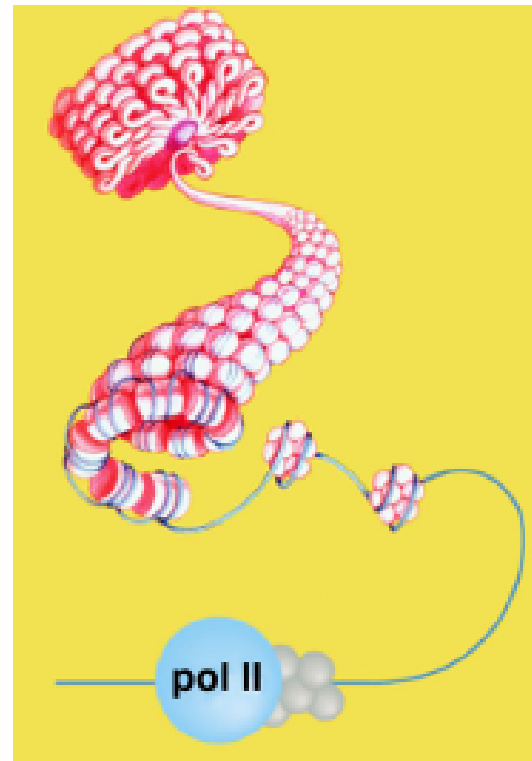
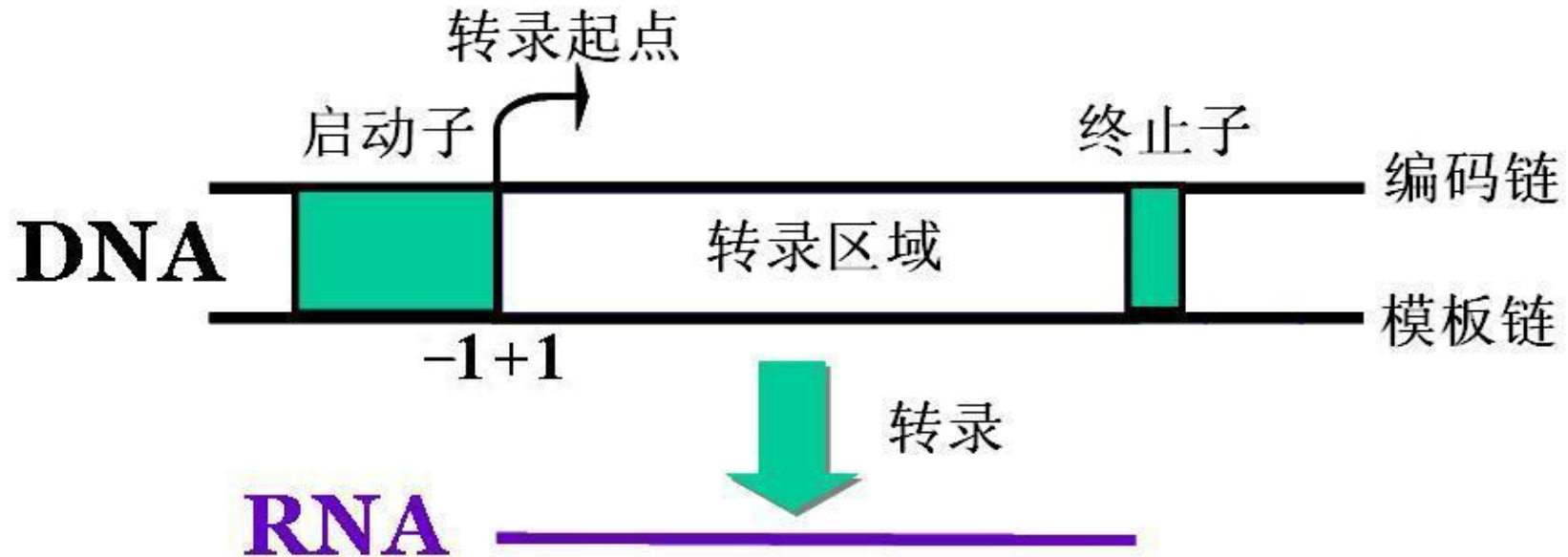


Fig. 1. The nucleosome, the fundamental particle of the eukaryote chromosome. Schematic shows the coiling of DNA around a set of eight histones in the nucleosome, the further coiling in condensed (transcriptionally inactive) chromatin, and uncoiling for interaction with the RNA polymerase II (pol II) transcription machinery.

scription system, that it would support not only accurately initiated but also appropriately regulated transcription. Communication from a regulatory protein to the transcription machinery at a promoter was believed to be direct. We found, however, that an additional crude fraction was required for regulation in the yeast system (Fig. 2). We referred to this activity as Mediator (9, 10), and in 1994, Stefan Bjorklund and Young-Joon Kim isolated the active protein as an assembly of more than 20 subunits, with a total mass in excess of a million daltons (11). Thirteen of the subunits were products of genes previously identified in screens for molecules involved in transcriptional regulation. These were disparate screens, done on different promoters in different laboratories at different times. With the isolation of Mediator, the products of all of the screens were united in a common biochemical entity. Still, the Mediator idea did not gain wide acceptance, as regulation in higher organisms was believed to be direct, through interaction of gene activator proteins with the so-called TAF subunits of the general transcrip-

# 转录单元 (transcription unit)

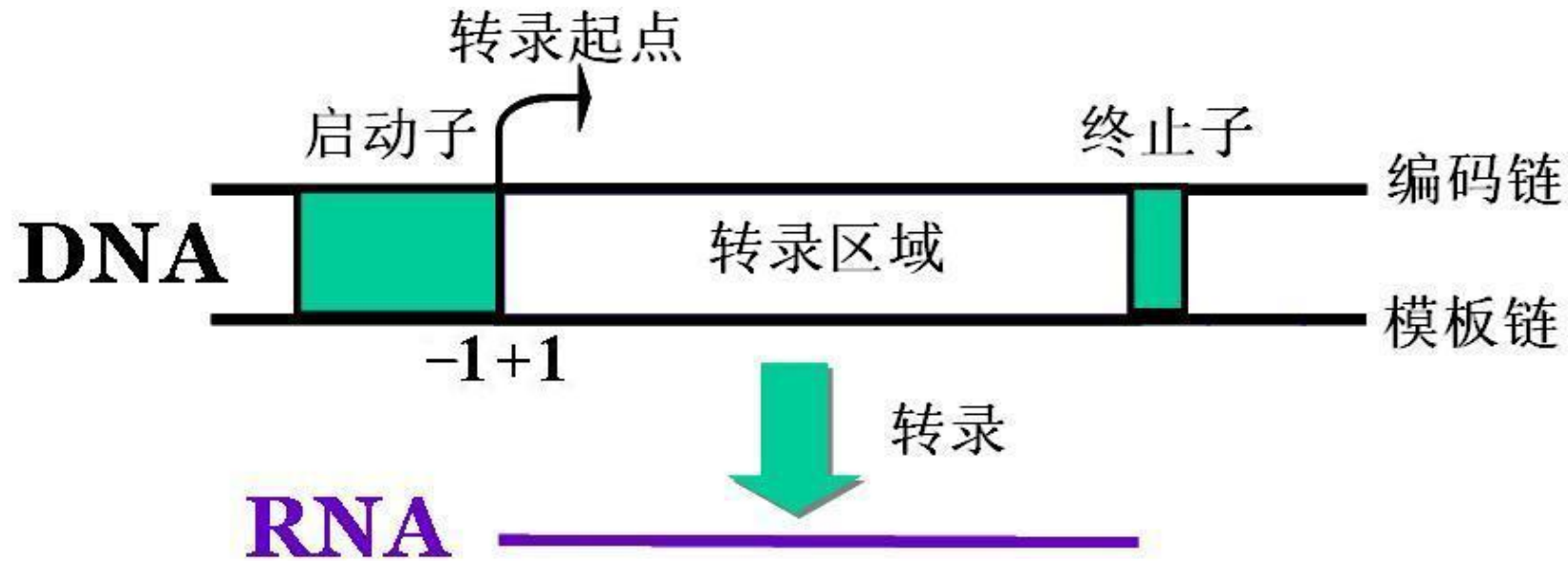


**转录起点**是指与新生**RNA**链第一个核苷酸相对应**DNA**链上的碱基，通常为**嘌呤**。

把起点5'末端的序列称为**上游** (upstream) ，

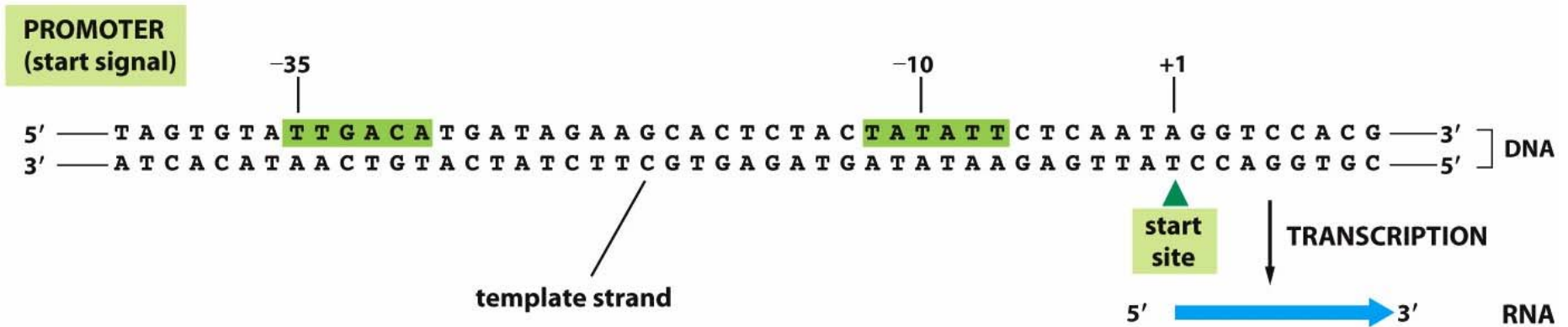
把其3'末端的序列称为**下游** (downstream) 。

# 转录起始位点 (Transcription start site)



- 在原核中90%为嘌呤，A或G；
- 位置固定；
- 通常在起始核苷酸的两侧为C和T  
(i.e. C**G**T or C**A**T)

# 启动子区的基本结构



启动子是一段位于结构基因**5'端上游区**的保守的**DNA**序列，能**活化RNA**聚合酶，使之与模板**DNA**准确地相结合并具有**转录起始**的特异性。

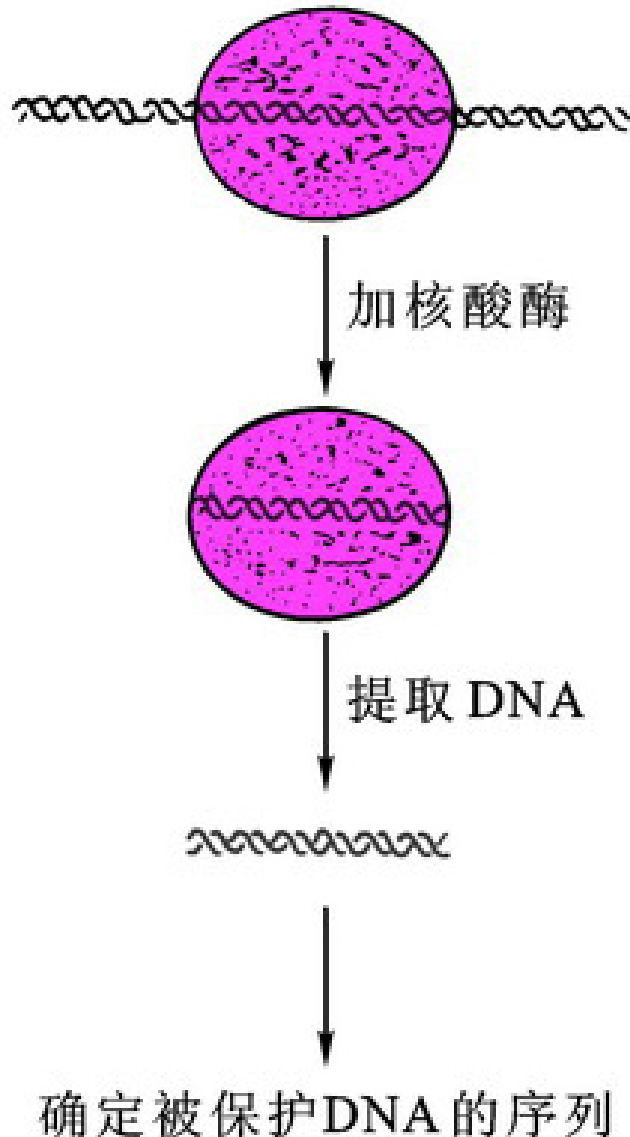
# 启动子与转录起始

大肠杆菌**RNA**聚合酶与启动子的相互作用主要包括：

- 启动子区的识别；
- 酶与启动子的结合；
- $\sigma$  因子的结合与解离。

# 原核生物的启动子

RNA 聚合酶结合到  
DNA 特定位点



## -10区(Pribnow区)的发现

1975年，Pribnow和Schaller将RNA聚合酶全酶与模板DNA结合后，用DNaseI降解DNA，得到41~44个核苷酸对的DNA片段。

序列分析发现，在被保护区内有一个由5个核苷酸组成的保守序列，是聚合酶结合位点，称为Pribnow区，其中央大约位于起点上游10bp处，所以又称为-10区。

# -10序列 (Pribnow框盒)

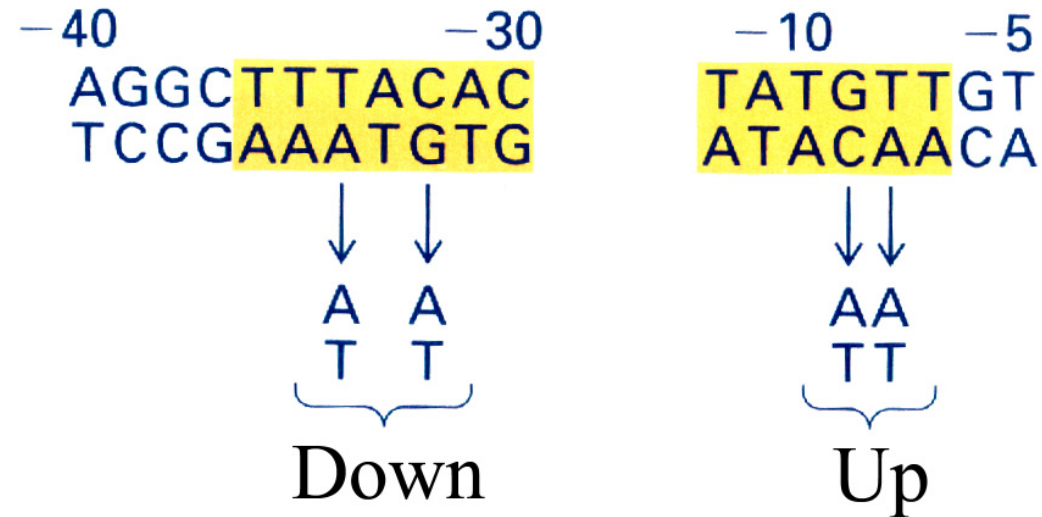
- 其保守序列为**TATAAT**，位于**-10bp**左右，其中**3'**端的“**T**”十分保守。
- **A.T**较丰富，易于解链。
- 它和转录起始位点**I**一般相距**5bp**。

## 功能:

- (1) **RNA pol**紧密结合；
- (2) 形成开放启动复合体；
- (3) 使**RNA pol**定向转录。

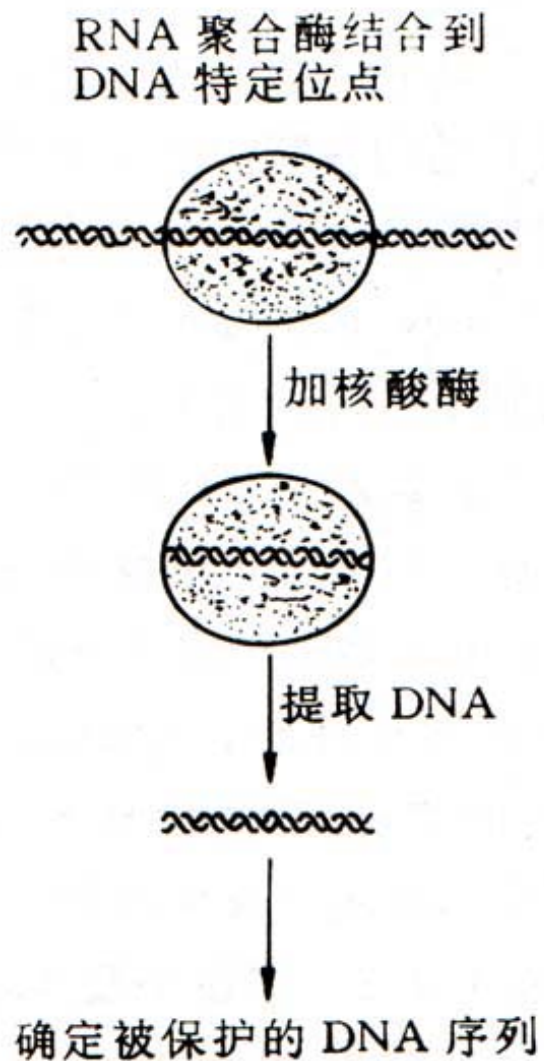


## (b) Promoter mutations



如果把Pribnow区从TATAAT变成AATAAT就会使该启动子发生下降突变(down mutation);

如果增加Pribnow区的共同序列, 将乳糖操纵子的启动子中的TATGTT变成TATATT, 就会提高启动子的效率, 称为上升突变(up mutation)。



提纯被保护的片段后却发现，**RNA**聚合酶并不能重新结合或并不能选择正确的起始点，表明在保护区外可能还存在与**RNA**聚合酶对启动子的识别有关的序列。



图 3-5 分离转录启动子区 DNA 序列程序图示

## -35区的发现



科学家又从噬菌体的左、右启动子 $P_L$ 及 $P_R$ 和SV40启动子的-35 bp附近找到了另一段共同序列：**TTGACA**。

## **-35序列 (Sextama盒)**

- 其保守序列为**TTGACA**,
- 与**-10**序列相隔**16-19bp**。

功能:

**(1)**为**RNA pol**的识别位点。

**(2)****RNA Pol**的核心酶只能起到和模板结合和催化的功能，并不能识别**-35**序列，只有 $\sigma$ 亚基才能识别**-35**序列，为转录选择模板链。

# 100个*E.coli*的不同启动子的序列测定结果

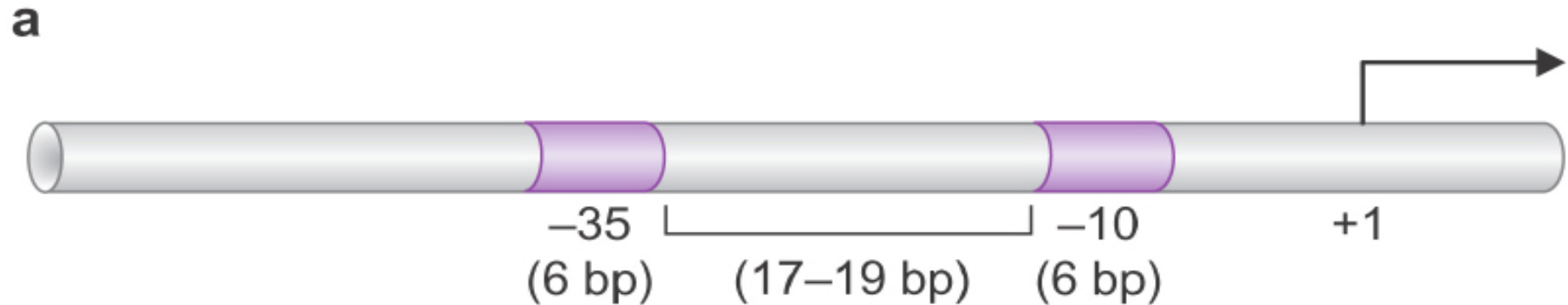
-35区

-10区

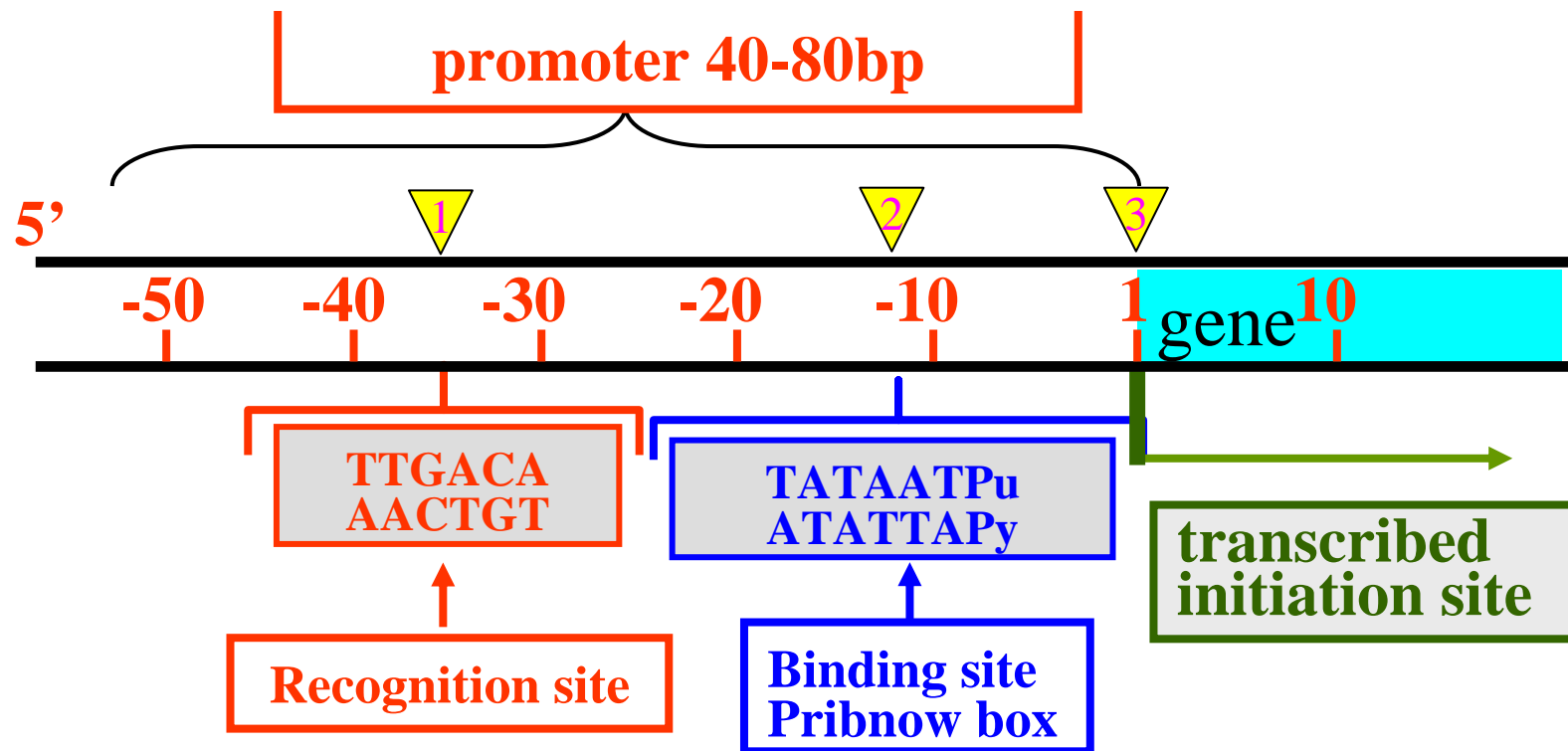
.....T<sub>85</sub>T<sub>83</sub>G<sub>81</sub>A<sub>61</sub>C<sub>69</sub>A<sub>52</sub>.....T<sub>89</sub>A<sub>89</sub>T<sub>50</sub>A<sub>65</sub>A<sub>100</sub>.....

- 大部分启动子都存在这两段共同序列，即位于 **-10 bp** 处的 **TATA** 区和 **-35 bp** 处的 **TTGACA** 区。
- 它们是 **RNA** 聚合酶与启动子的结合位点，能与  $\sigma$  因子相互识别而具有很高的亲和力。

## -10区和-35区的最佳距离



- 在原核生物中，-10区和-35区的最佳距离大约是**16~19 bp**。
- 过大或过小都会降低转录活性。  
这可能是因为**RNA Pol**本身的大小和空间结构有关。



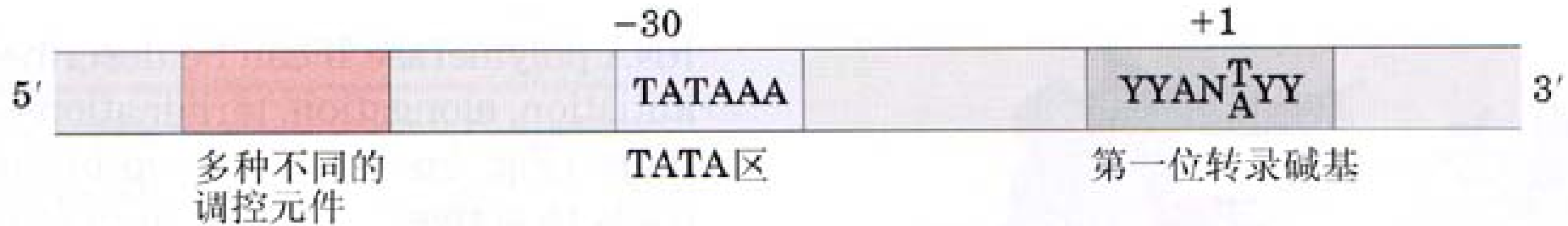
- (1) 结构典型，都含有识别(R)，结合(B)和起始(I)三个位点；
- (2) 序列保守，如-35序列，-10序列结构都十分保守；
- (3) 位置和距离都比较恒定；
- (4) 直接和多聚酶相结合；
- (5) 常和操纵子相邻；
- (6) 都在其控制基因的5'端；
- (7) 决定转录的启动和方向。

## 原核生物启动子的共同的特点

# 真核生物启动子的结构特点



# 真核生物RNA聚合酶II所识别的启动子区



Hogness等发现类似Pribnow区的**Hogness区**，在转录起始点上游**-25~-30 bp**处，保守序列为**TATAAA**，也称**TATA区**。

在起始位点上游**-70~-78 bp**处还有另一段共同序列**CCAAT**，称为**CAAT区**（**CAAT box**）。

# 真核基因的启动子

## 1. 核心元件

**TATA box:** -25~-30 bp区

**启始子(initiator, Inr):** 转录起始位点附近

## 2. 上游启动子元件 (UPE)

**CAAT box:** -70~-80区CCAAT序列

**GC box:** -80~-110含有GCCACACCC  
或GGGCGGG序列。

## 上游启动子元件

**(upstream promoter element, UPE):**

将**TATA**区上游的保守序列称为上游启动子元件或称**上游激活序列**（**upstream activating sequence, UAS**）。

# 真核生物启动子对转录的影响

- **TATA区--使转录精确地起始:**

如果除去**TATA**区或进行碱基突变，转录产物下降的相对值不如**CAAT**区或**GC**区突变后明显，但发现所获得的**RNA**产物起始点不固定。

- **CAAT区和GC区主要控制转录起始频率:**

基本不参与起始位点的确定。

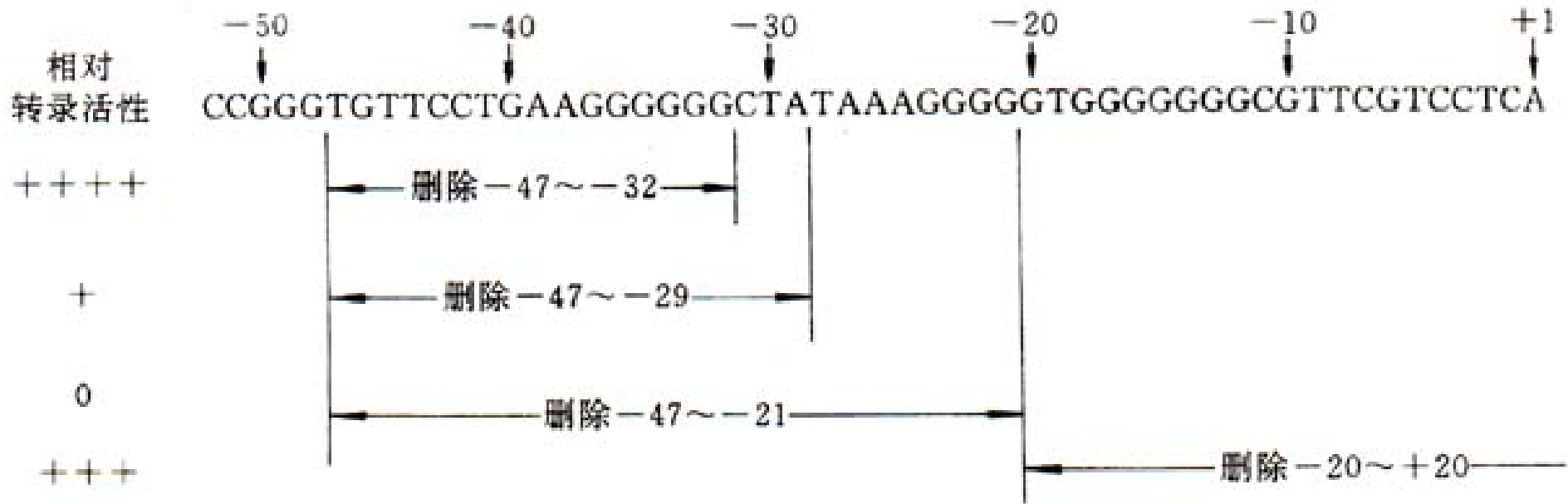


图3-11 SV40 基因启动子上 TATAAA 及邻近区域对基因转录活性的影响

研究SV40晚期基因启动子发现，上游激活区的存在与否，对该启动子的生物活性有着根本性的影响。  
若将该基因5'上游-21~-47核苷酸序列切除，基因完全不表达。

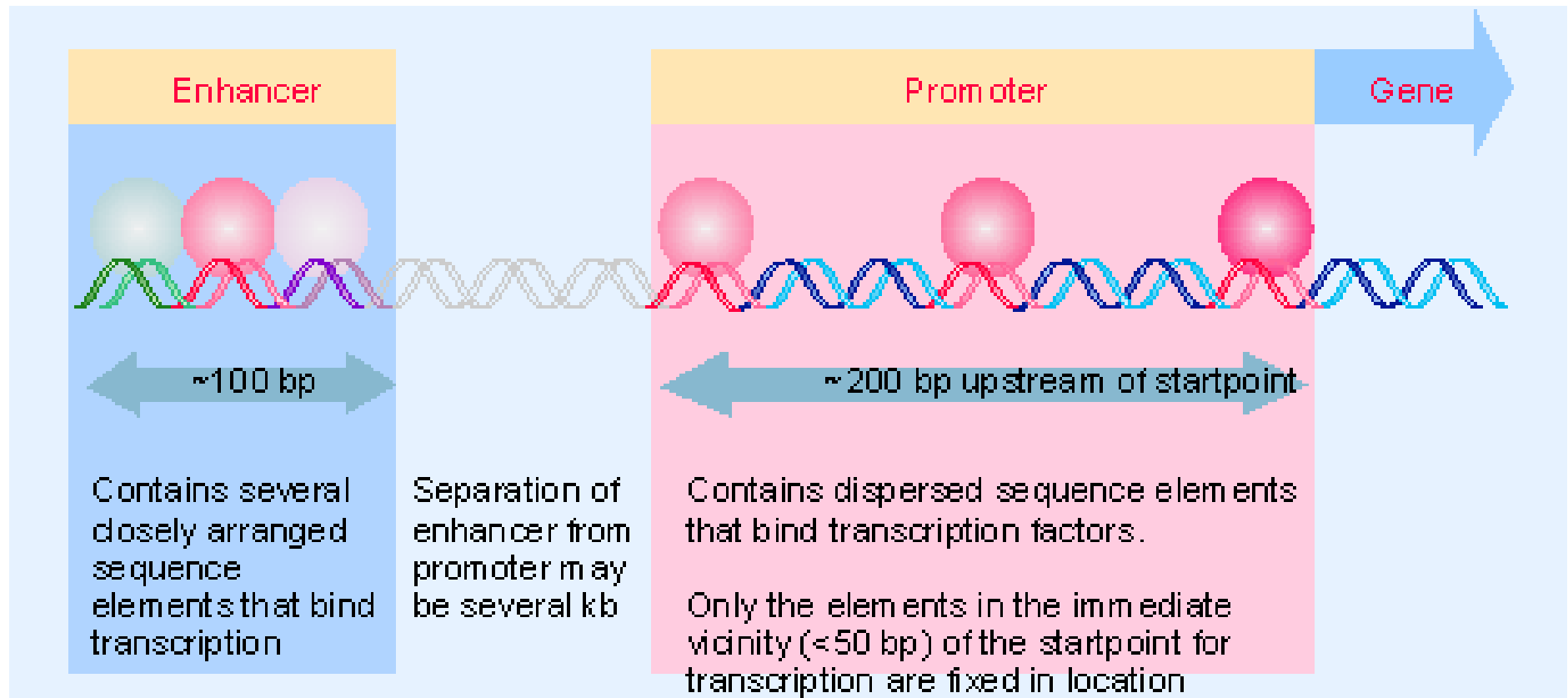
# 远端调控区

## 增强子(enhancer)

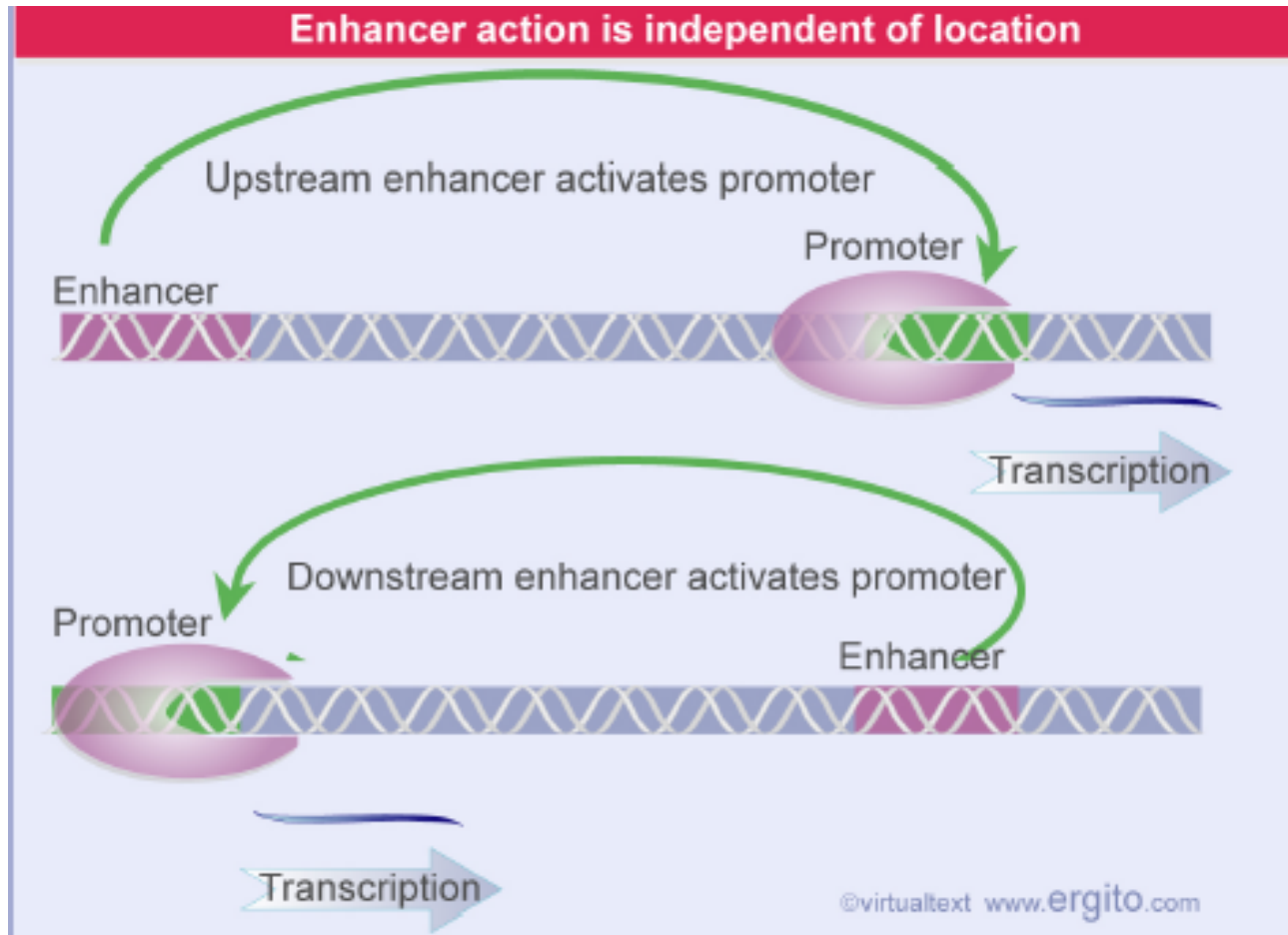
1981年由Benerji, Rusconi小组和 Chambom等发现的，又称远上游序列（far upstream sequence）。

已在SV40的转录单元上发现其转录起始位点上游约200bp处有两段72bp长的重复序列，它们不是启动子的一部分，但能增强或促进转录的起始，因此，称这种能强化转录起始的序列为增强子或强化子（enhancer）。

# 增强子 (Enhancer)



增强子很可能通过影响染色质DNA-蛋白质结构或改变超螺旋的密度而改变模板的整体结构，从而使得RNA聚合酶更容易与模板DNA相结合，起始基因转录。



**An enhancer can activate a promoter from upstream or downstream locations, and its sequence can be inverted relative to the promoter.**



# 增强子特点

## ① 具有远距离效应。

常在上游-**200bp**处，但可增强远处启动子的转录，即使相距十几**Kb**也能发挥其作用；

## ② 无方向性。

无论在靶基因的上游，下游或内部都可发挥增强转录的作用；

## ③ 顺式调节。

只调节位于同一染色体上的靶基因，而对其它染色体上的基因无作用；

## ④ 无物种和基因的特异性，

可以接到异源基因上发挥作用；

## ⑤ 具有组织的特异性。

**SV40**的增强子在**3T3**细胞中比多瘤病毒的增强子要弱，但在**HeLa**细胞中**SV40**的增强子比多瘤病毒的要强**5**倍。增强子的效应需特定的蛋白质因子参与。

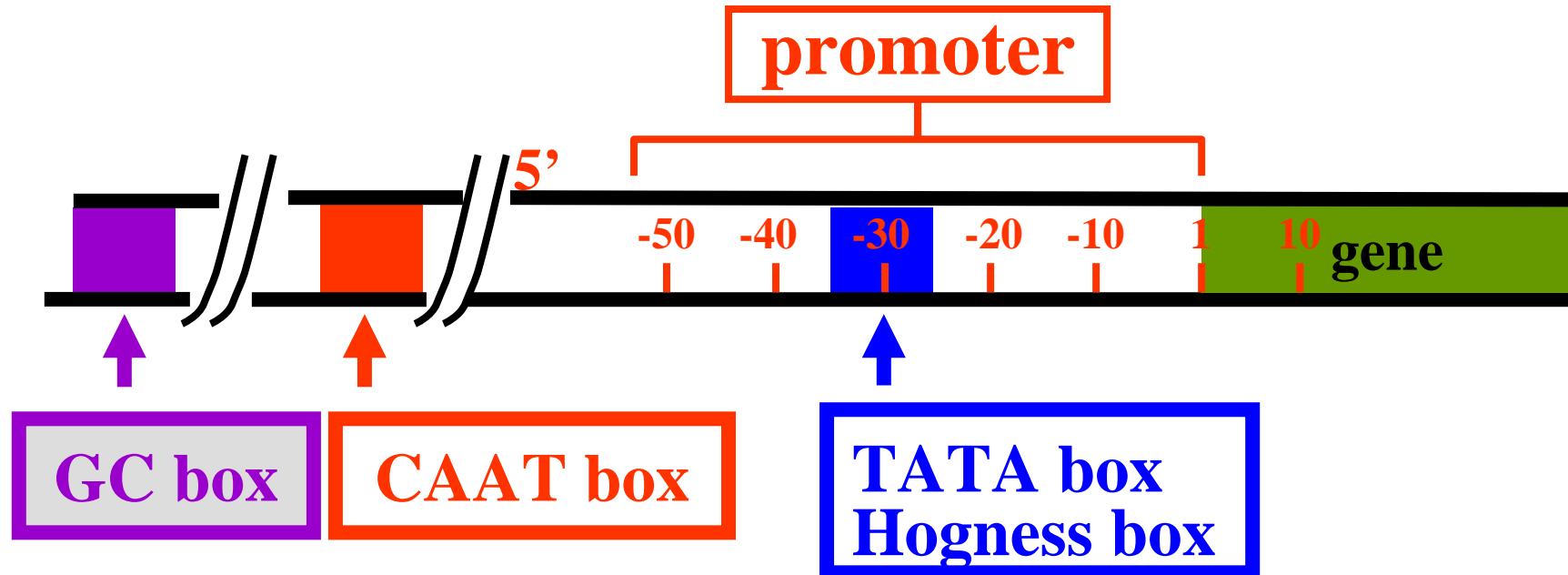
## ⑥ 有相位性。

其作用和**DNA**的构象有关。

## ⑦ 有的增强子可以对外部信号产生反应。

如热休克基因在高温下才表达。编码重金属蛋白的金属硫蛋白基因在镉和锌存在下才表达。某些增强子可以被固醇类激素所激活。

# 真核生物的启动子特点



- (1) 有多种元件：**TATA框**，**GC框**，**CATT框**，**OCT**等；
- (2) 结构不恒定。有的有多种框盒，如组蛋白**H2B**；有的只有**TATA框**和**GC框**，如**SV40**早期转录蛋白；
- (3) 它们的位置、序列、距离和方向都不完全相同；
- (4) 有的有远距离的调控元件存在，如增强子；
- (5) 这些元件常常起到控制转录效率和选择起始位点的作用；
- (6) 不直接和**RNA pol**结合。转录时先和其它转录激活因子相结合，再和聚合酶结合。



**谢谢!**