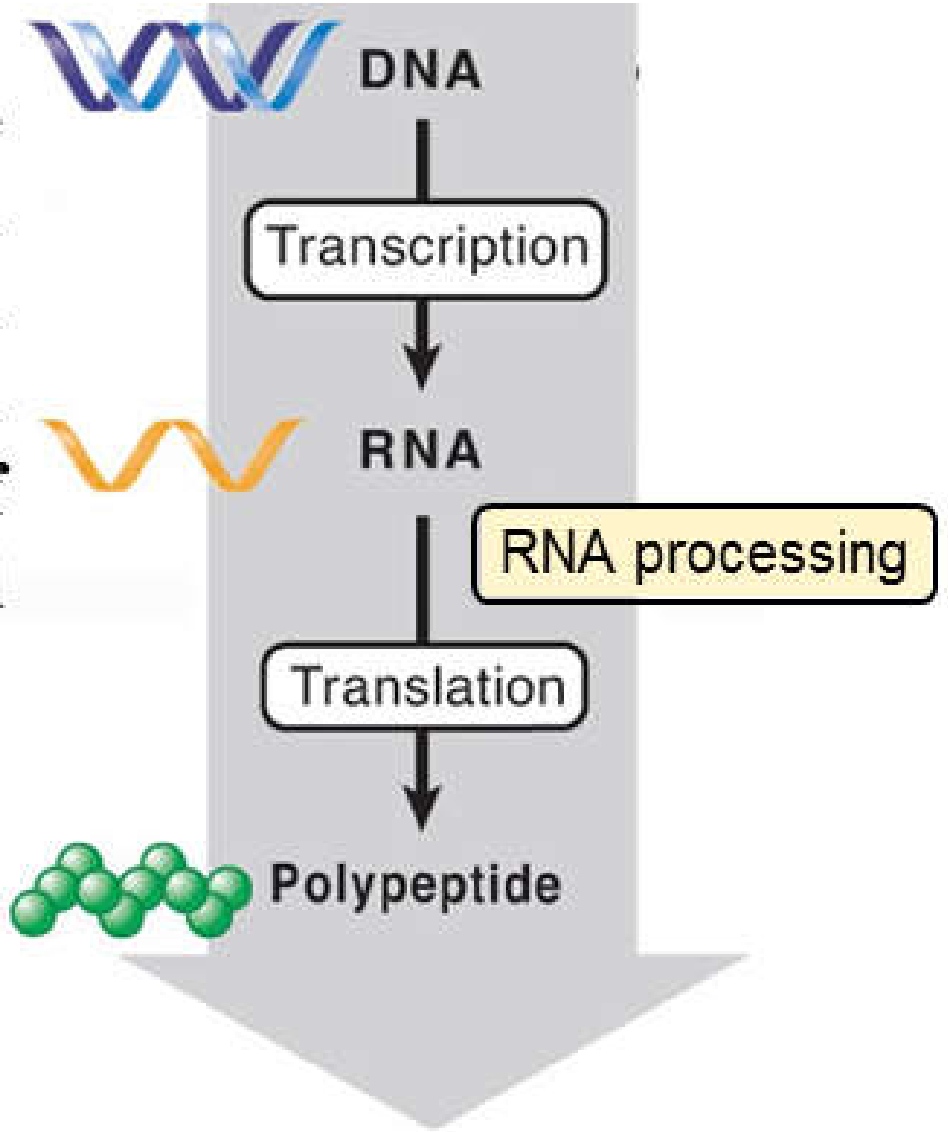
Part Ⅲ Gene expression

- **Gene expression is the process by which genetic information from a gene is used in the synthesis of proteins and functional RNA.**

基因表达是指遗传信息由DNA到蛋白质或功能RNA的过程。



Chapter 5 Transcription in prokaryotes

1. Transcription: an overview

- **Transcription** is the synthesis of a single-stranded RNA from one of the double-stranded DNA as a template.

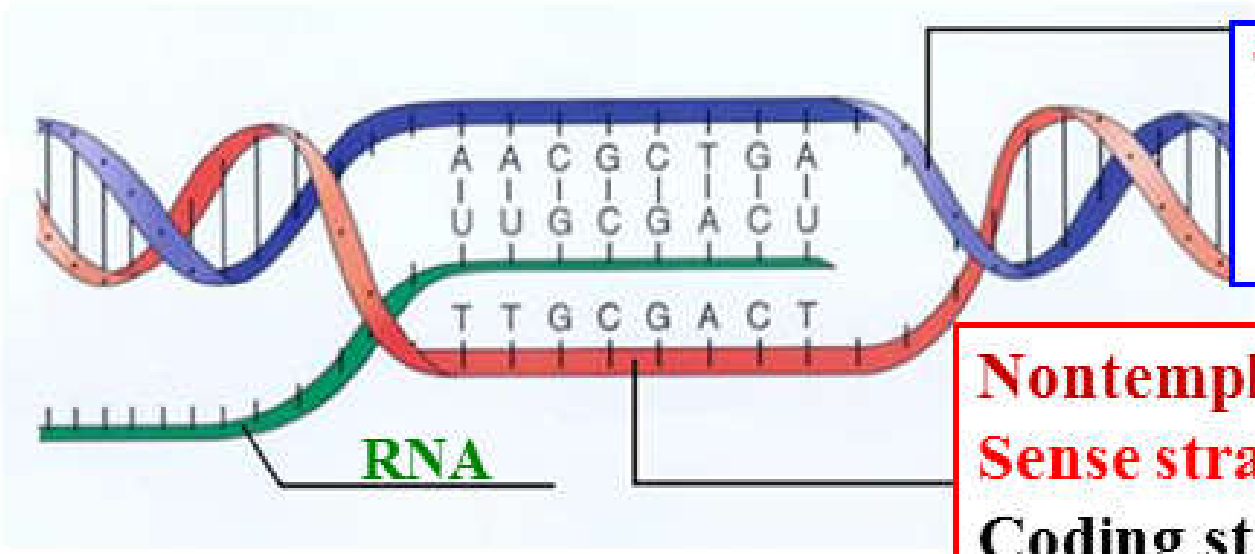
  转录是以双链DNA中的一条链为模板，合成单链RNA的过程（遗传信息由DNA到RNA的过程）。

1.1 The general characteristics of transcription

1.1.1 Asymmetric transcription (不对称转录)

Only one strand of the DNA molecule is used as template for the enzymatic formation of RNA.

转录时，双链DNA中只有一条链作为转录的模板，这种转录方式称作不对称转录。



Template strand
Antisense strand
模板链/反义链

Nontemplate strand
Sense strand 有义链/正义链
Coding strand 编码链

1.1.2 Partial transcription (部分转录)

RNA在转录时每次只转录DNA分子上的一部分遗传信息，这种转录方式称作**部分转录**。

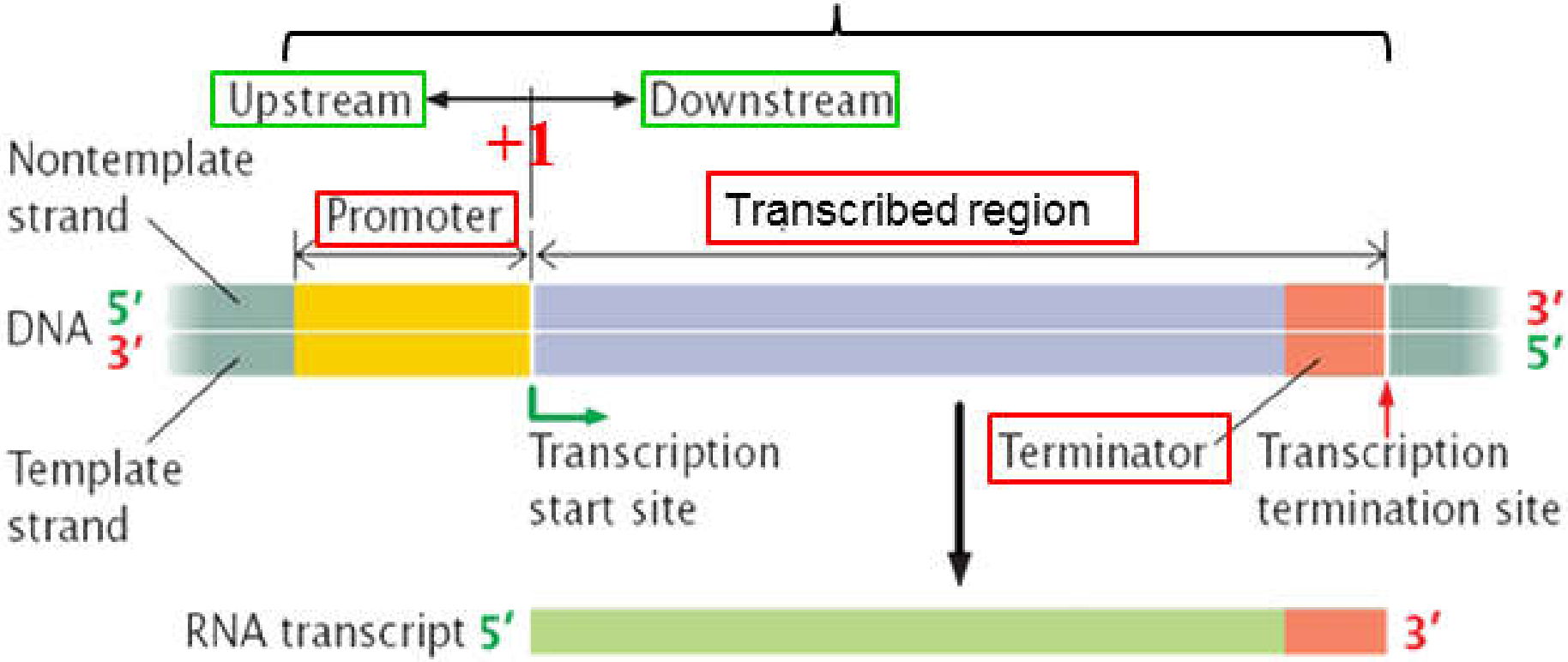
基因表达受时间、空间（环境）因素的影响，所有基因不同时转录。

http://www.nature.com/scitable/definition/transcription-unit-260

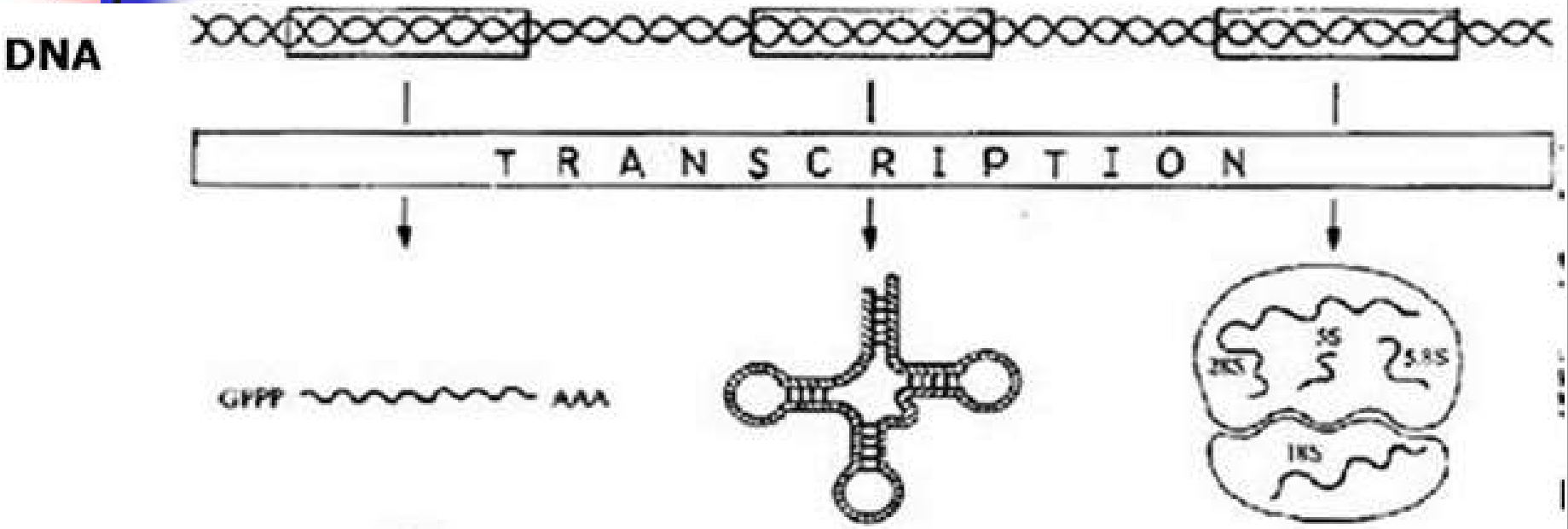
1.2 Transcription unit

Transcription unit is the DNA sequence from the promoter to the terminator.

转录单位是从启动子到终止子的一段DNA序列.



1.3 The types of transcription products



mRNA	tRNA	rRNA
	microRNA, snRNA, snoRNA, lncRNA…	

Coding RNA

Non-coding RNA are **functional RNA** that are **not translated to proteins** (非编码RNA是指不被翻译成蛋白质的功能RNA).

1.4 Properties of RNA polymerase

DDRP or transcriptase (转录酶)
RDRP or RNA replicase

DDRP

- **Templates:** ssDNA
- **Substrates:** NTP
- **Cofactor:** Mg^{2+}/Mn^{2+}, Zn^{2+}
- **Direction of RNA synthesis:** 5'→3'
- Requires **no primer**

2. *E. coli* RNA polymerase

One of the largest enzyme in *E. coil.* 1种

		组成	数目	分子量	主要功能
Holoenzyme 全酶	Core enzyme 核心酶	α	2	36.5KD	与启动子结合
		β	1	150KD	催化作用
		β'	1	160KD	与DNA模板结合
		ω	1	11KD	功能不清楚
		σ	1	70KD	识别结合启动子

RNA polymerase differs from organism to organism.

e.g. The RNA polymerases of phage T3 and T7 are smaller single polypeptide chains.

2.1 α subunit

- Two identical subunits in the core enzyme
- Encoded by the *rpoA* gene
- Required for assembly of the core enzyme
- Plays a role in promoter binding
- Plays a role in the interaction of RNA polymerase with some regulatory factors

2.2 β subunit

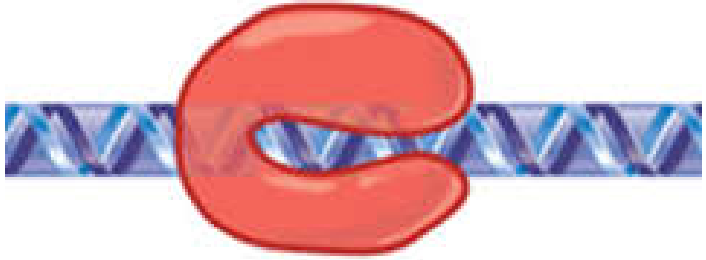
- Encoded by the *rpoB* gene
- Make up the catalytic center of the RNA pol.
- Some antibiotics inhibit the catalytic activity of β subunit.
  - Rifampicin (利福平) is an inhibitor of RNA polymerase that blocks initiation.
  - Streptolydigin (利迪链霉素) blocks RNA elongation.
  - β subunit may contain two domains responsible for transcription initiation and elongation.

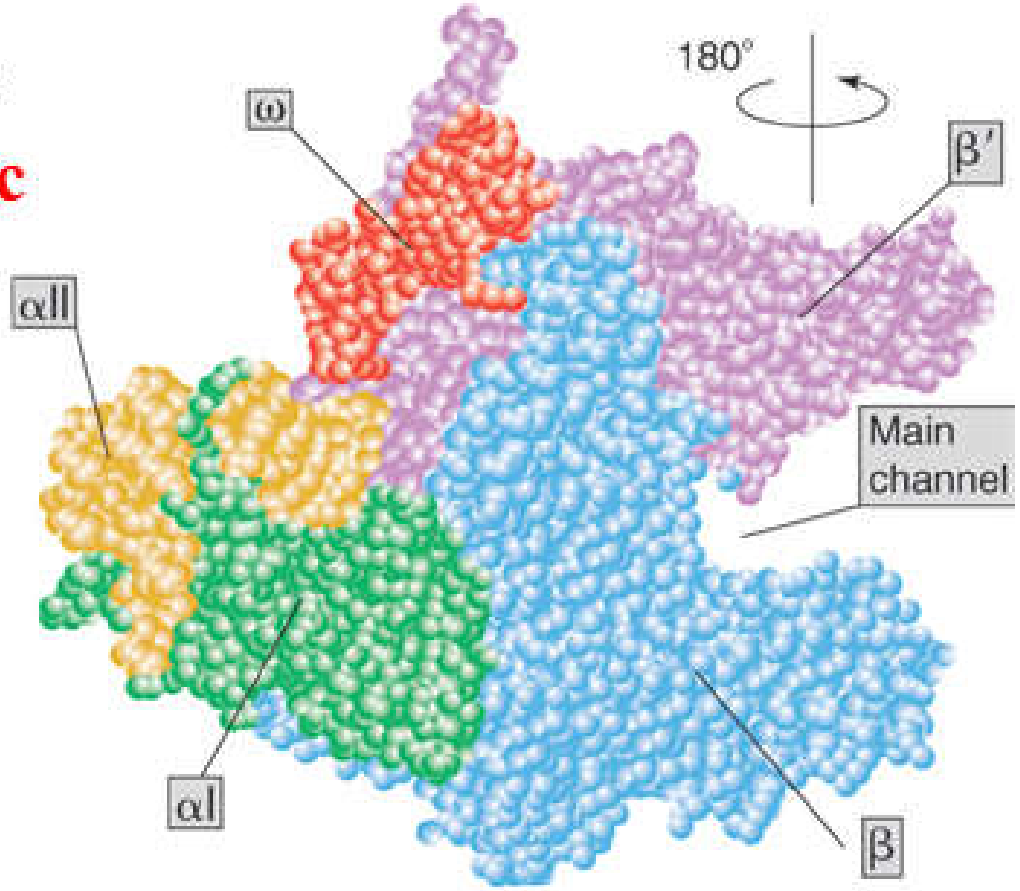
2.3 β' subunit

- Encoded by the *rpoC* gene
- Binds two Zn^{2+} and may participate in the catalytic function of the polymerase
- Responsible for binding to the template DNA.
  - A valine residue in the β' subunit inserts into the minor groove of the DNA, and prevent the DNA from slipping.
  - Hyparin (肝素) competes with DNA for binding to the β' subunit and inhibits transcription *in vitro* (体外).

Core enzyme: α2ββ'ω

The core enzyme has a general non-specific affinity for DNA, which is referred to as loose binding.





2.4 σ factor

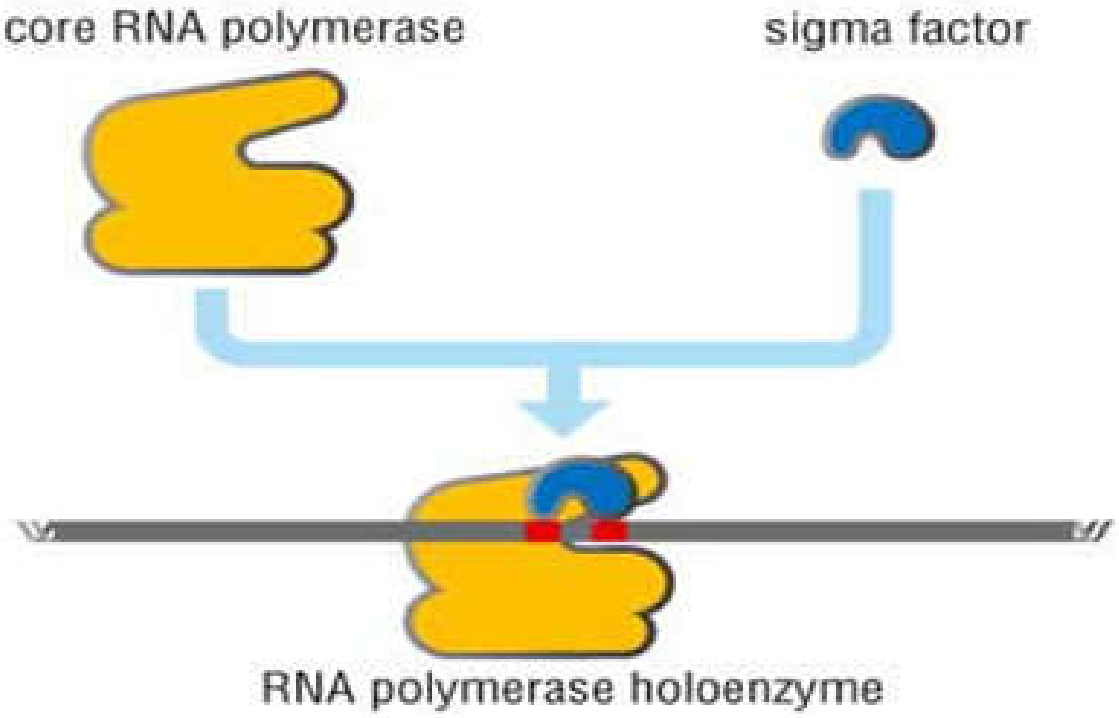
2.4.1 Several σ factors in *E. coli*

Gene	Factor	Use
rpoD	σ^{70}	most required functions
rpoS	σ^{S}	stationary phase/some stress responses
rpoH	σ^{32}	heat shock
rpoE	σ^{E}	periplasmic/extracellular proteins
rpoN	σ^{54}	nitrogen assimilation 同化
rpoF	σ^{F}	flagellar synthesis/chemotaxis 趋化性
fecI	σ^{fecI}	iron metabolism/transport

70KD

2.4.2 Function of σ factor

Binding of the σ factor converts (将...转换为) the core RNA polymerase into the holoenzyme which tightly binds to DNA.



- The σ factor has a critical role in promoter recognition.
  - σ factor binding dramatically increases the specificity of the holoenzyme for correct promoter-binding site.
- The σ factor by itself cannot bind to DNA, but interaction with core enzyme unmasks (暴露) a DNA-binding region of σ.
- Not required for transcription elongation.

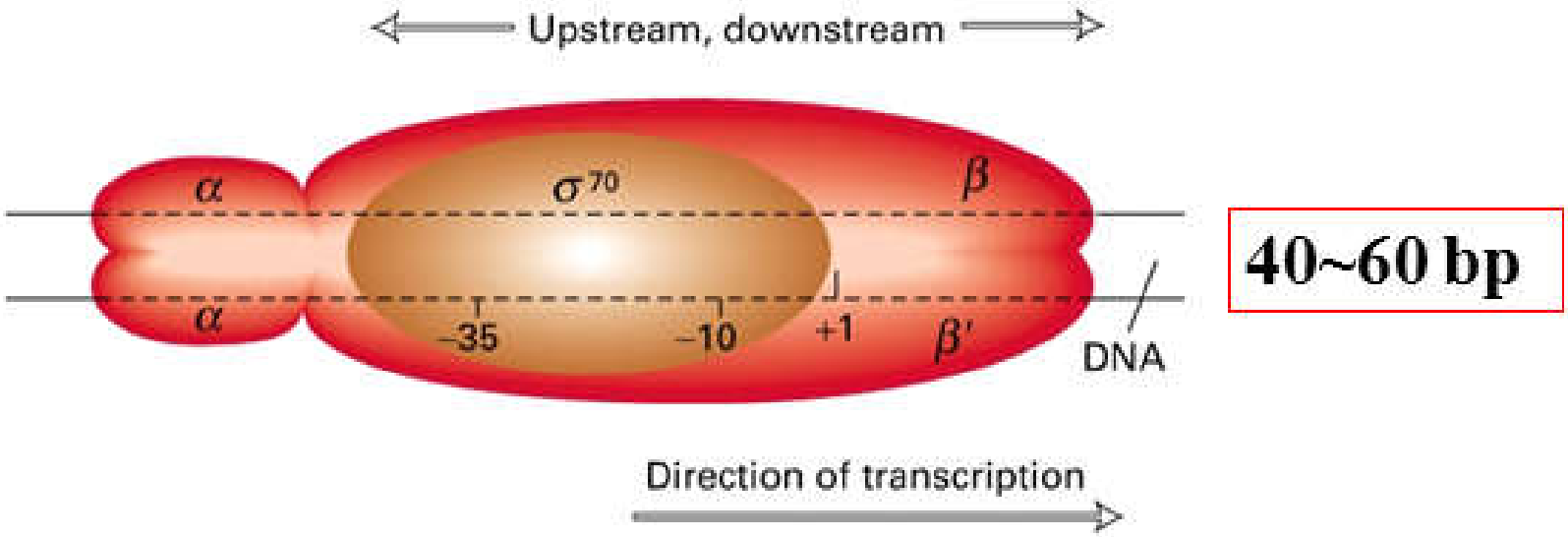
3. The *E. coli* σ^{70} promoter

Promoters are DNA sequences which are involved in binding of RNA polymerase to initiate transcription.

启动子是能与RNA聚合酶结合并起始转录的一段DNA序列。

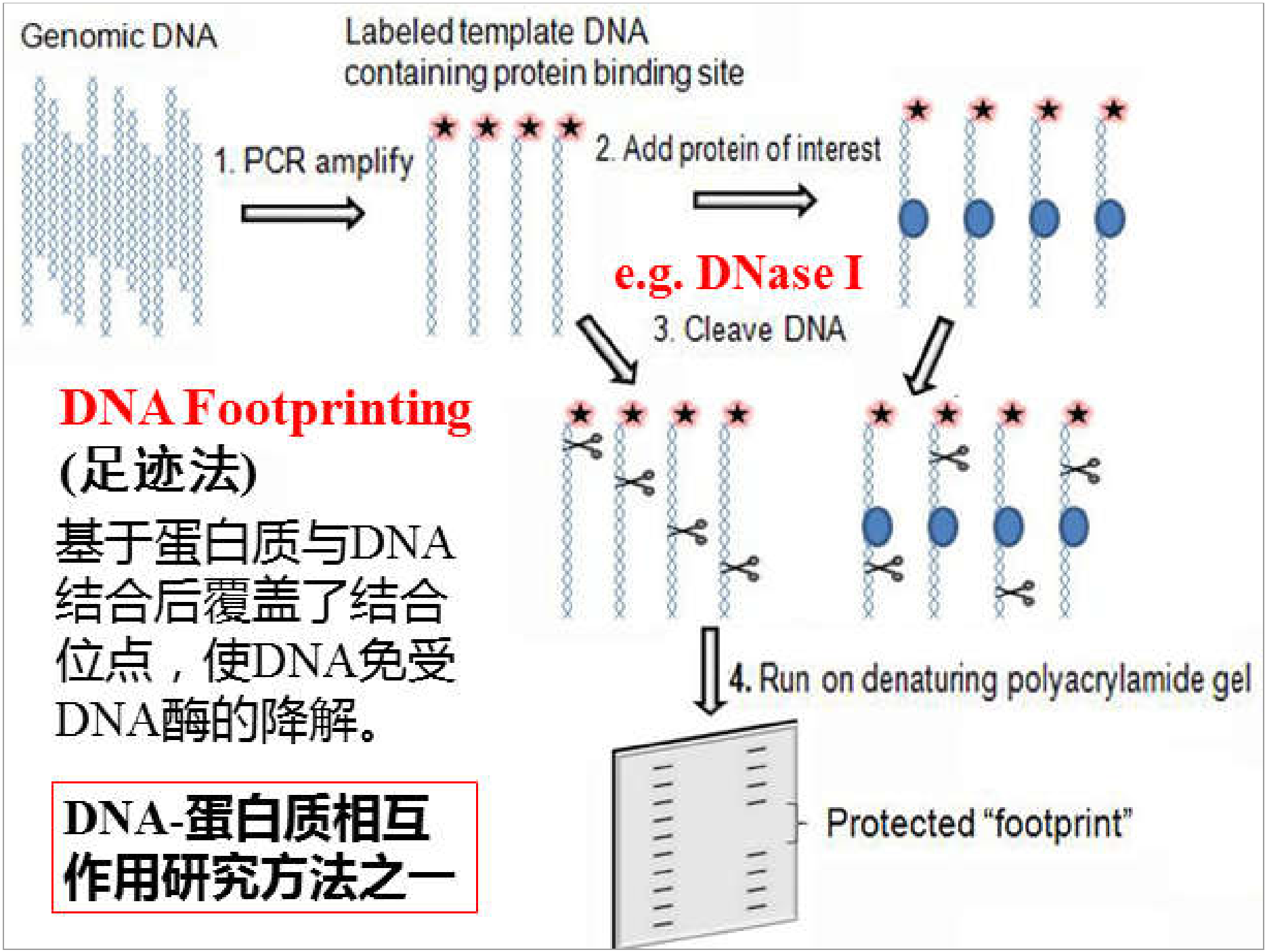
启动子有时也称为启动基因（本身不被转录）。通常位于转录区上游。

3.1 σ^{70} promoter size and sequences



40~60 bp

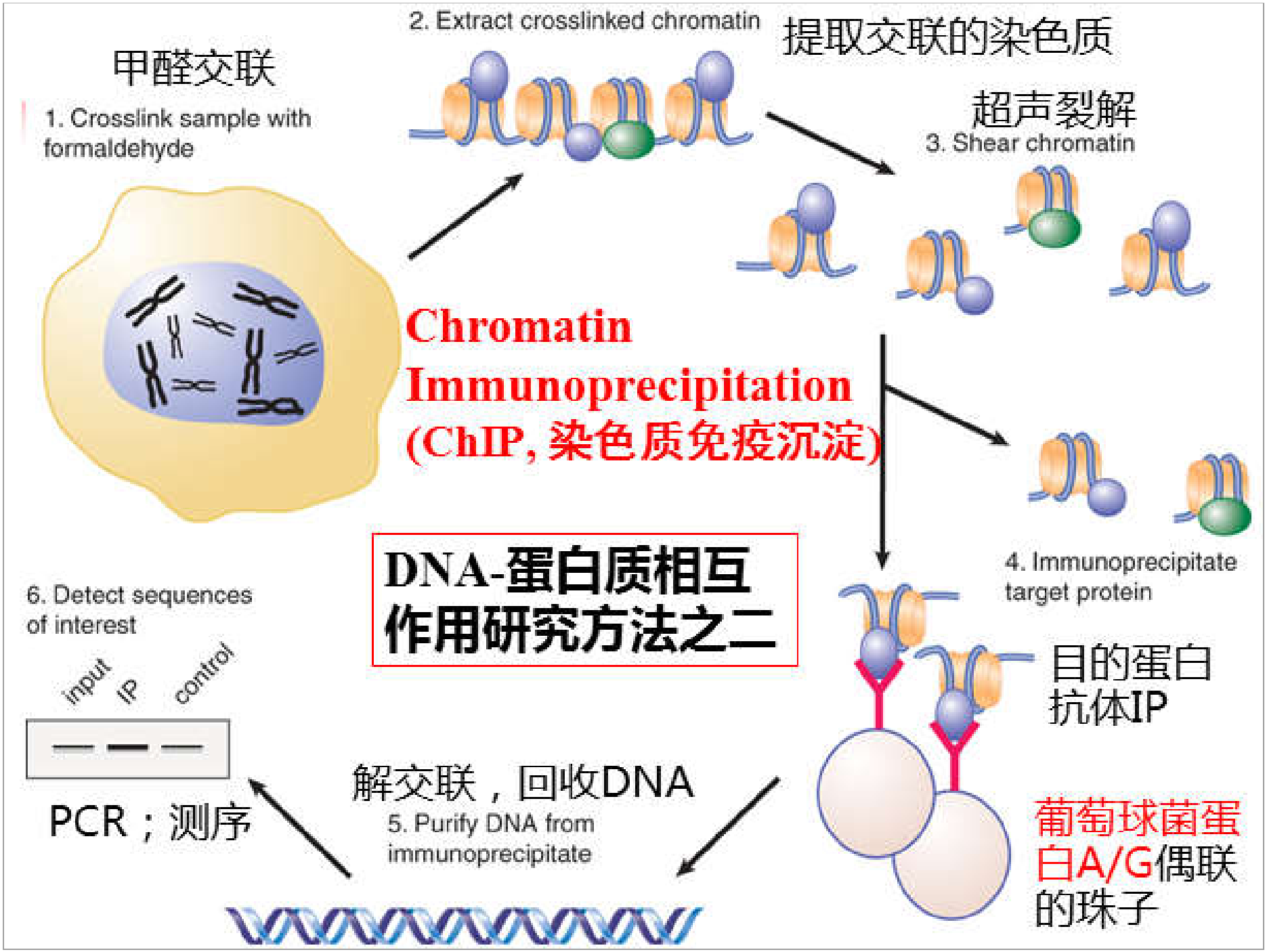
- -55 to +20: bound by the polymerase
- -20 to +20: tightly associated with the polymerase and protected from nuclease digestion by DNase I



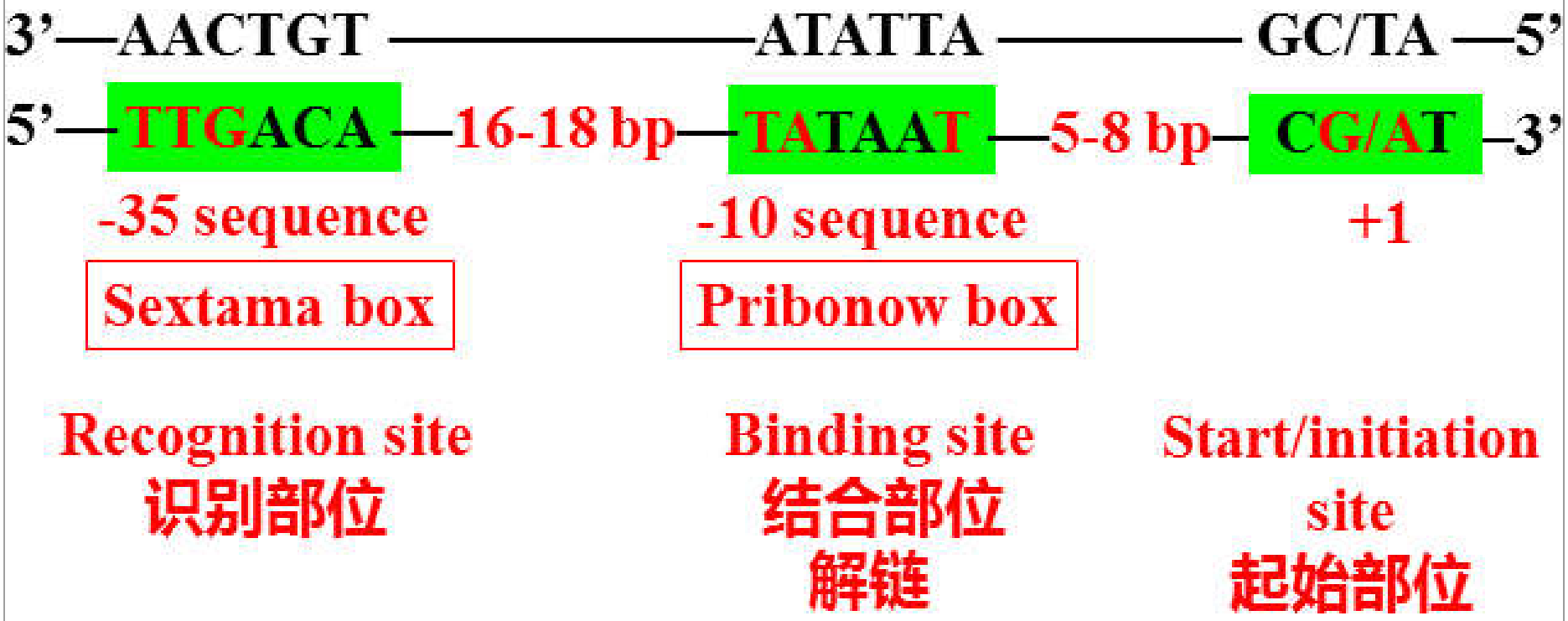
DNA Footprinting
(足迹法)

基于蛋白质与DNA结合后覆盖了结合位点，使DNA免受DNA酶的降解。

DNA-蛋白质相互作用研究方法之一


甲醛交联
1. Crosslink sample with formaldehyde
2. Extract crosslinked chromatin
提取交联的染色质
超声裂解
3. Shear chromatin
Chromatin Immunoprecipitation (ChIP, 染色质免疫沉淀)
DNA-蛋白质相互作用研究方法之二
4. Immunoprecipitate target protein
目的蛋白抗体IP
葡萄球菌蛋白A/G偶联的珠子
解交联，回收DNA
5. Purify DNA from immunoprecipitate
6. Detect sequences of interest
input
IP
control
PCR；测序

Core promoter sequence



3.1.1 -35 sequence

- -35 sequence is also named Sextama box or -35 box.
- Consensus sequence (共有序列): TTGACA. The first three positions (TTG) are the most conserved among *E. coli* promoters.
- Enhances recognition and interaction with σ factor.
- Separated by 16-18 bp from the -10 sequence in 90% of all promoters.

3.1.2 -10 sequence (Pribonow box)

Pribonow box was firstly recognized by David Pribnow in 1975.

- Consensus sequence (共有序列): TATAAT. The first two bases (TA) and the final T are most highly conserved (保守的).
- Important for RNA polymerase binding and DNA unwinding.
- This hexamer (六聚体) is separated by 5-8 bp from position +1, and the distance is critical.

3.1.3 Transcription start site

- Transcription start/initiation site is a purine in 90% of all gene.
- G is more common at position +1 than A
- The sequence around the start site influence initiation efficiency.
- There are usually a C and T on either side of the start nucleotide.

3.2 Promoter efficiency

- There is considerable variation in sequence between different promoters, and the transcription efficiency can vary by up to 1000-fold.

  - 并非所有启动子的-10序列和-35序列都完全一样。事实上很难找到与共有序列完全匹配的-10和-35 序列。**完全匹配序列往往出现在极强的启动子中**，这些启动子异常活跃地启动转录。 **弱启动子**常常与-35序列的共有序列很不相似。转录起始需要其他蛋白帮助。

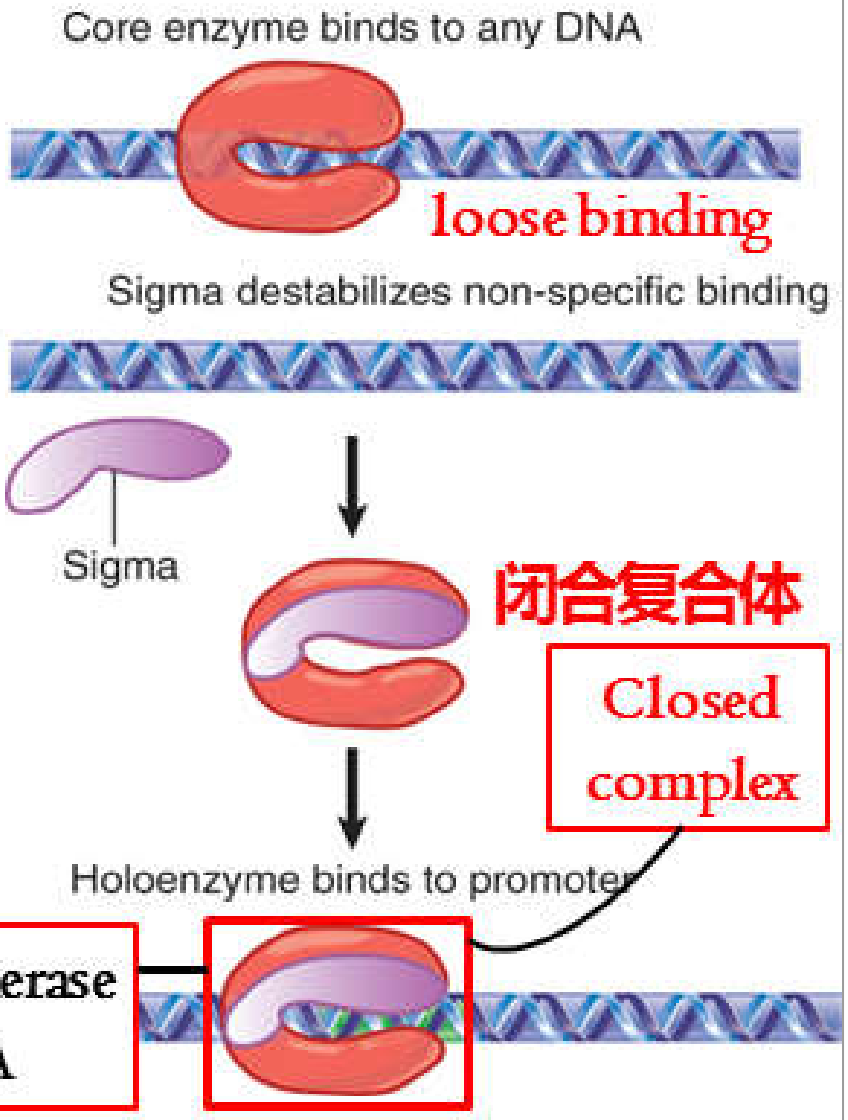
4. Transcription process in prokaryotes

4.1 Initiation

4.1.1 Promoter binding

(1) Core enzyme recognizes and binds to -35 sequence of the promotor with the participation of σ factor.

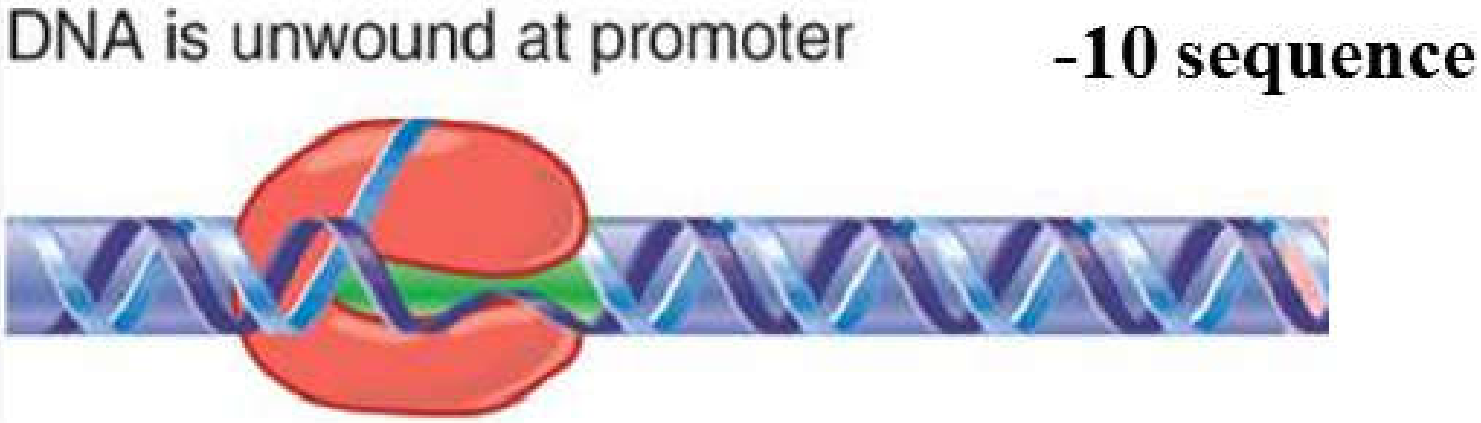
(2) Holoenzyme migrates to -10 sequence of the promotor to form closed complex.



The initial complex of RNA polymerase with the base-paired promoter DNA

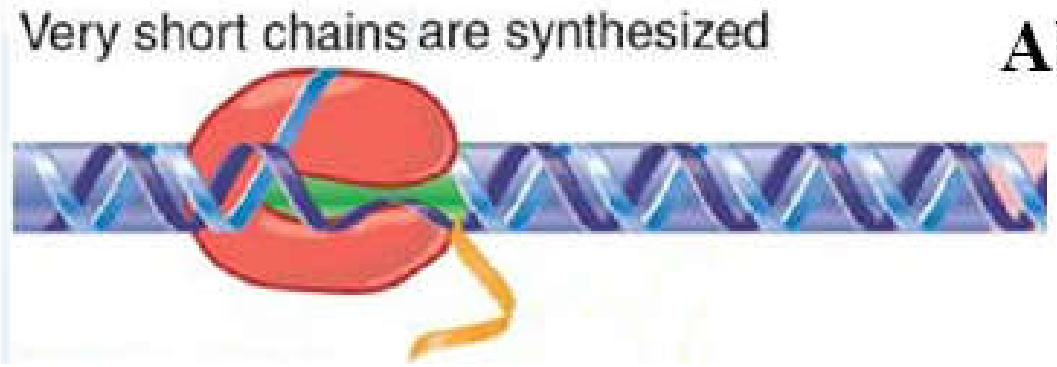
4.1.2 DNA unwinding

The initial unwinding of the DNA results in formation of an open complex with the polymerase, and this process is referred to as tight binding.



4.1.3 RNA chain initiation

- **No primer is needed.**
- **The first 9~10 nt are added without enzyme movement along the DNA.**
- **After each one of the first 9~10 nt is added to the chain, there is a significant probability that the chain will be aborted (中止) - Abortive initiation (无效起始).**

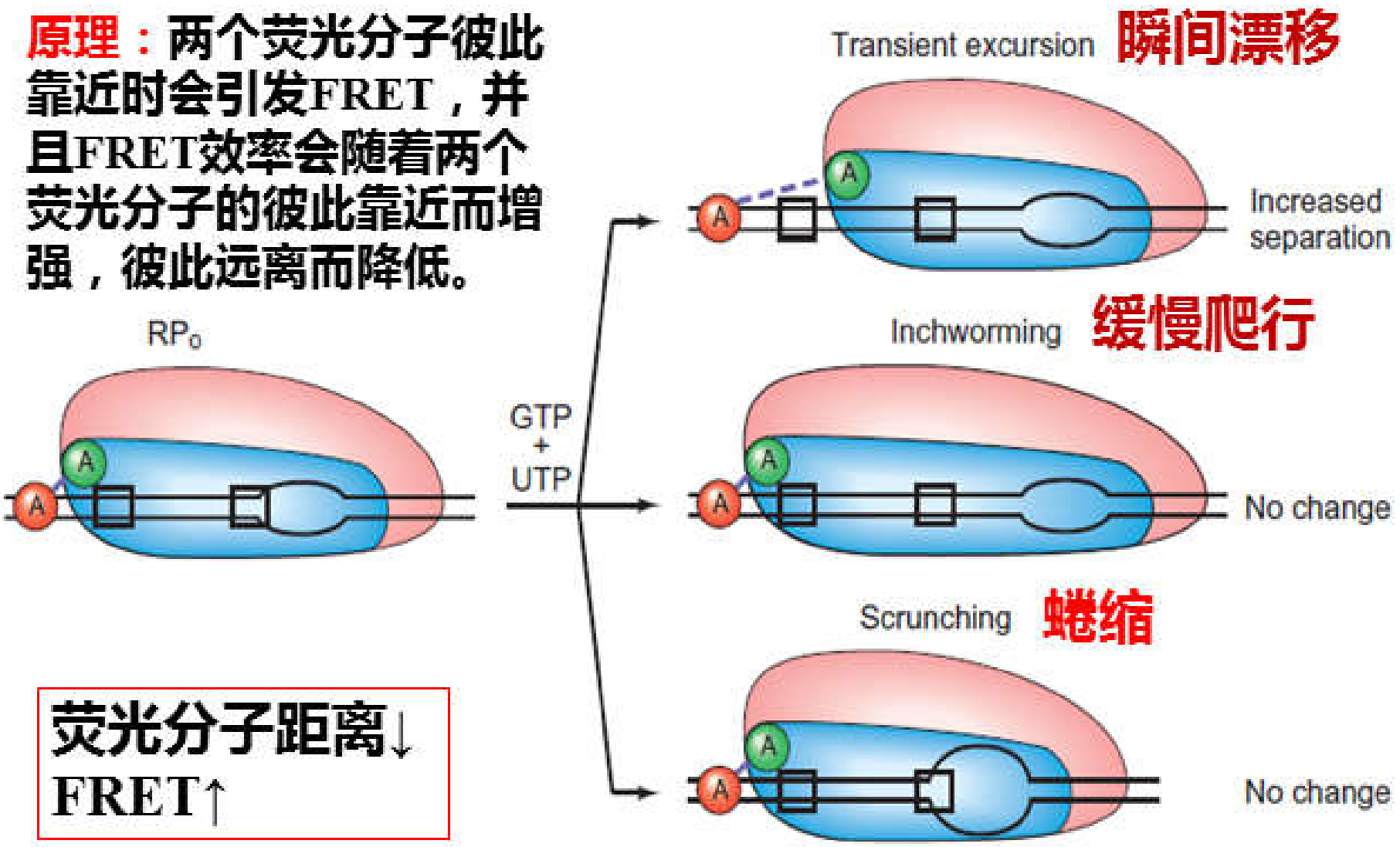


Abortive transcript
中断转录物
<10 nt

- **Promoter clearance (启动子清除):** when initiation succeeds, the polymerase leaves the promoter to enter the stage of elongation and allow the reinitiation of transcription.
  - The minimum time for promoter clearance is 1-2 s (a long event, the synthesis is 40 nt/s).

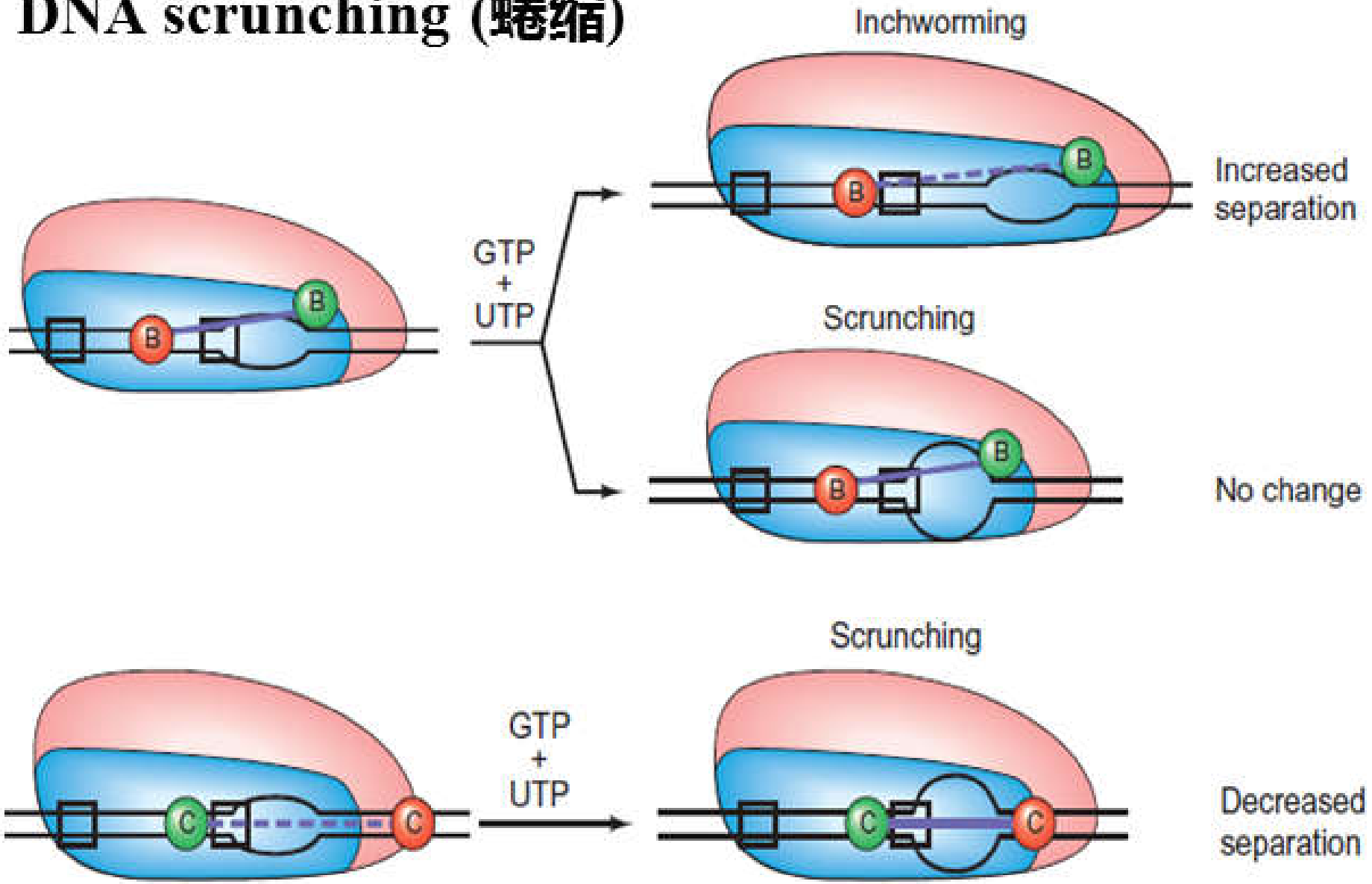
荧光共振能量转移 (fluorescence resonance energy transfer, FRET)

原理：两个荧光分子彼此靠近时会引发FRET，并且FRET效率会随着两个荧光分子的彼此靠近而增强，彼此远离而降低。



荧光分子距离↓ FRET↑

DNA scrunching (蠕缩)



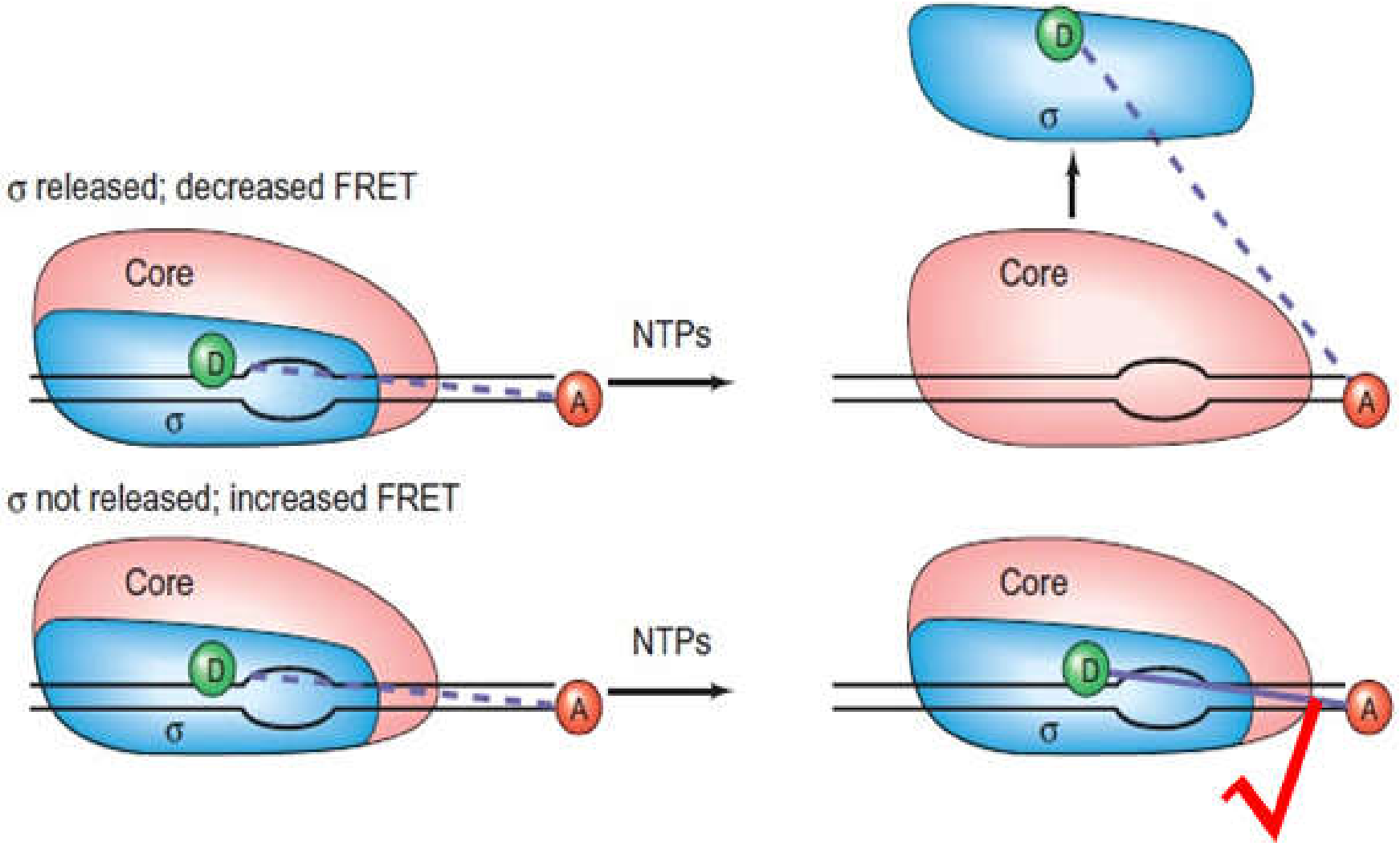
4.2 Elongation

4.2.1 σ factor release

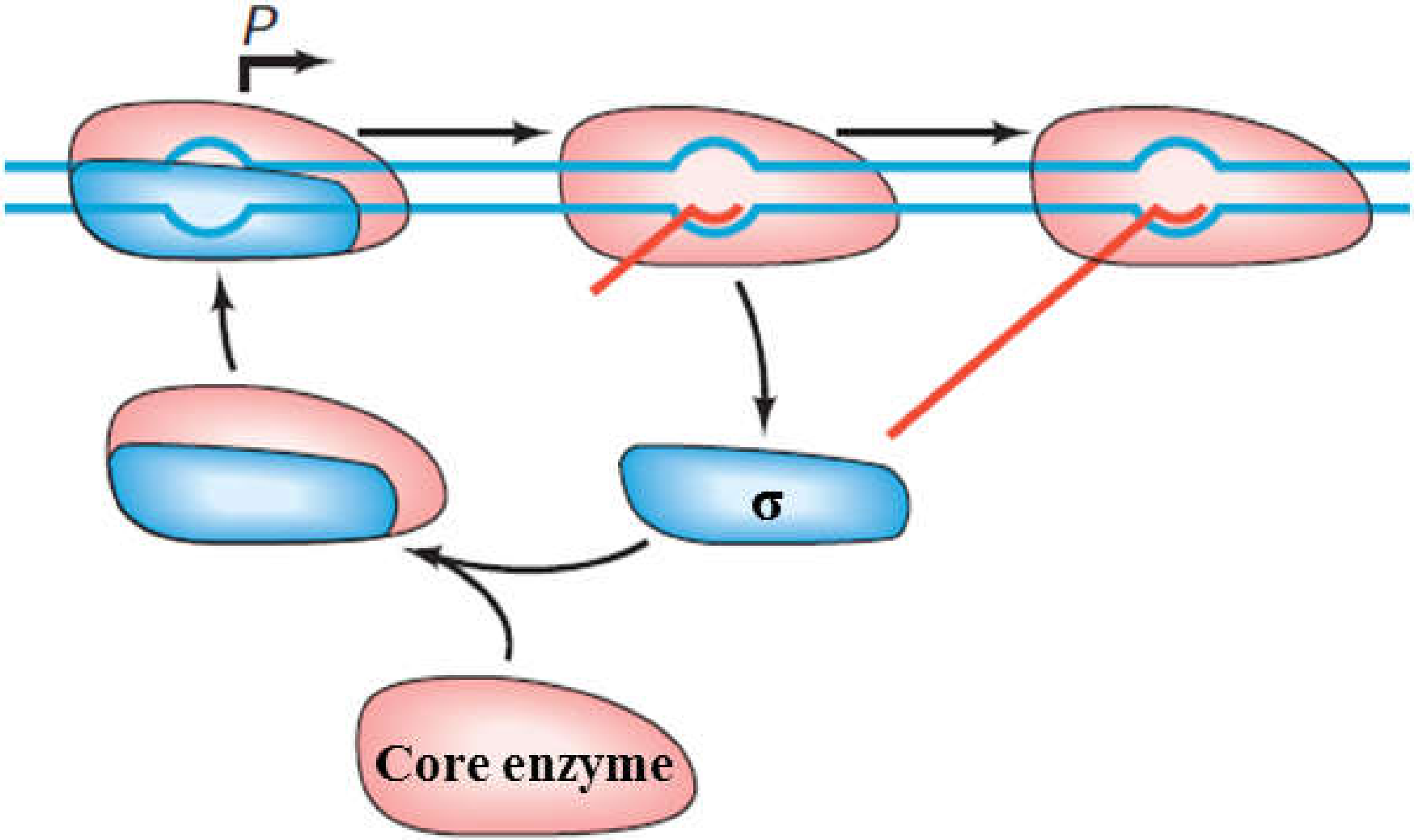
(1) Stochastic release model (随机释放模型)

- **Sigma factor is released** from core RNA polymerase, but there is **no discrete** (分离) **point** during transcription; rather, it is **released randomly**.
- **注意：**σ因子并不是在启动子清除后立即从核心酶释放，而是在延伸过程中随机释放的。但长转录物的延伸复合体更倾向于丢失σ因子。

将两个荧光分子分别标记到σ因子（供体）和其下游的DNA（受体）上。

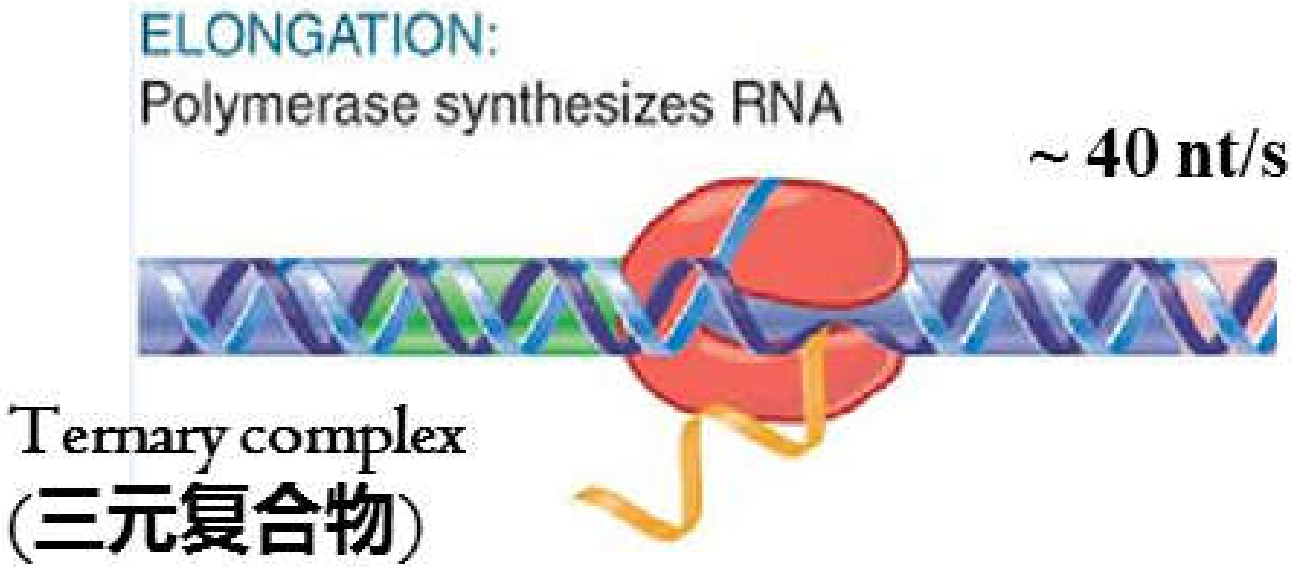


(2) The σ cycle



4.2.2 RNA chain elongation

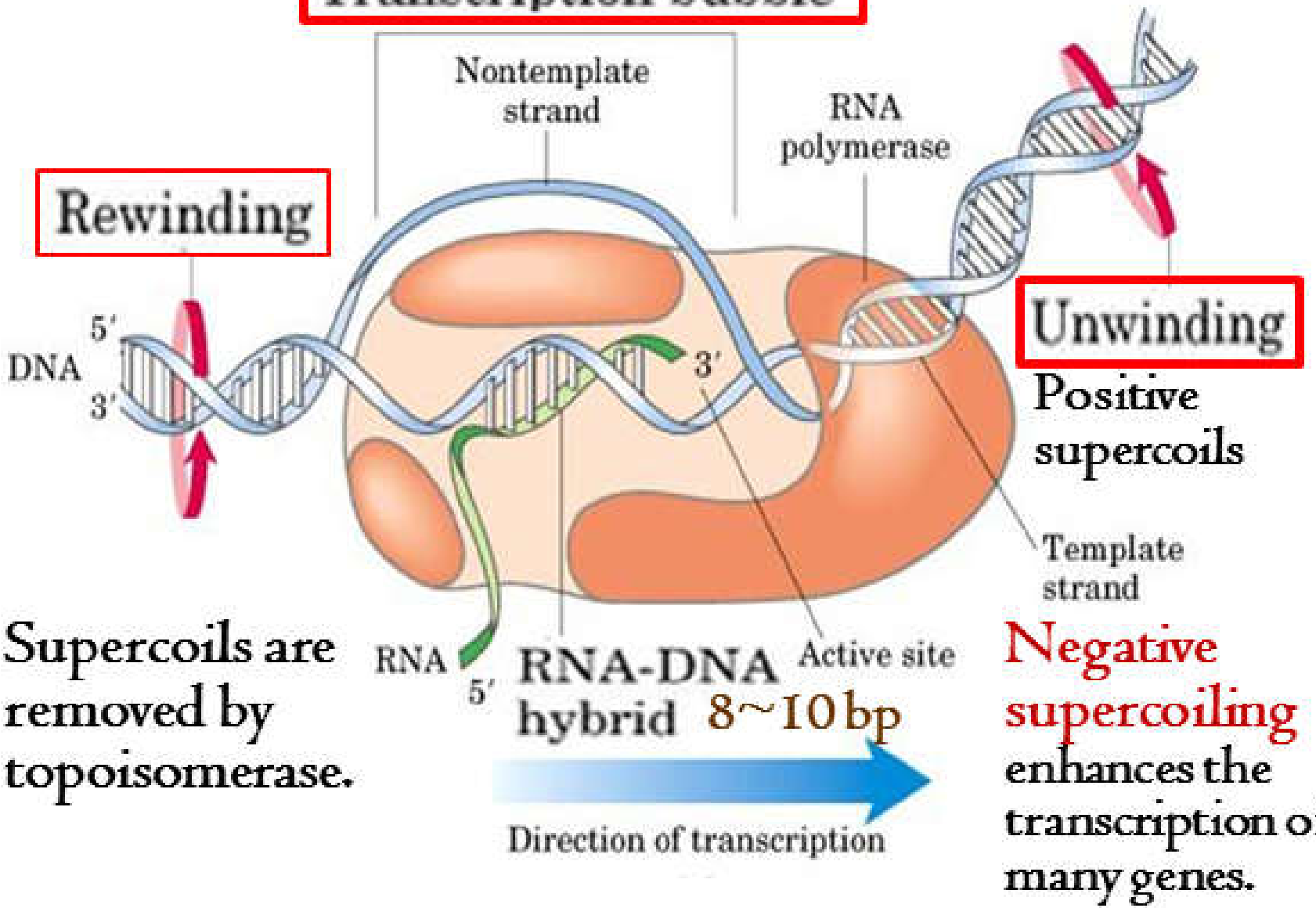
- The core enzyme progresses along the DNA.
- RNA chain is extended in the 5′→3′ direction by adding nucleotides to the 3′ end.



~ 17 bp of unwound DNA region

Helicase ?

Transcription bubble



4.3 Termination

When RNA polymerase reaches a terminator (终止子) at the end of a gene it falls off the template, releasing the RNA.

A terminator is a specific DNA sequence where the transcription complex dissociate and transcription stops.

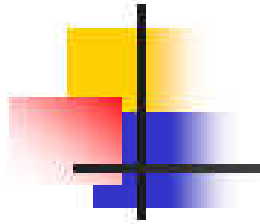
- Rho (ρ)-independent termination or intrinsic (内源性) termination — 强终止
- Rho-dependent termination — 弱终止

4.3.1 Rho (ρ)-independent termination

Structure of an intrinsic terminator:

(1) GC-rich inverted repeat can form hairpin (发夹) structure (self-complementary 自身互补).

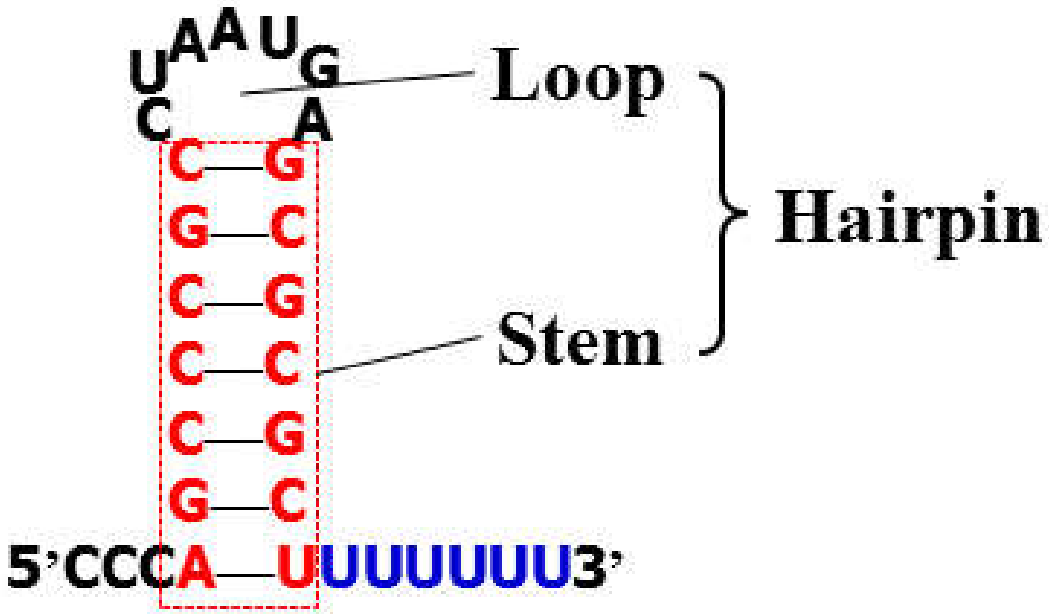
(2) A string of Ts (4~8) in the nontemplate DNA strand (Us in transcribed RNA)



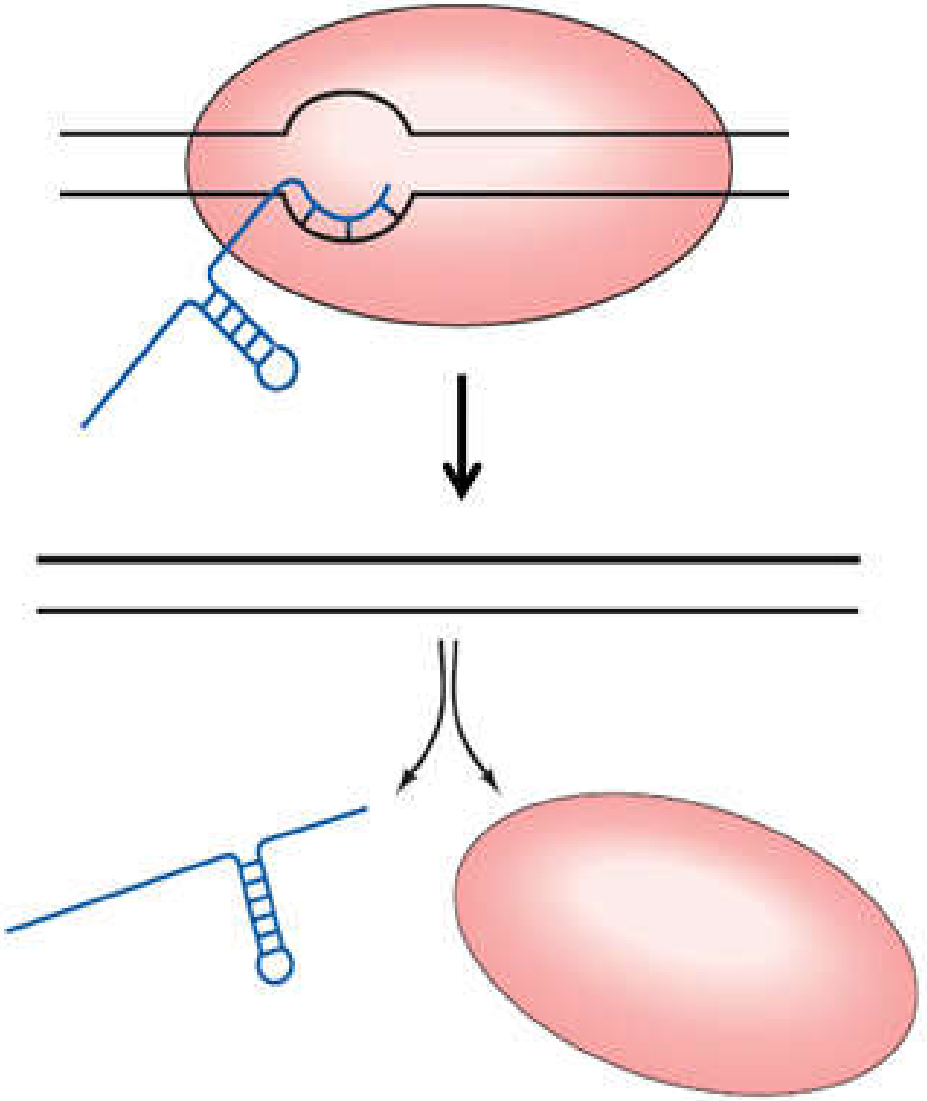
5’CCC AGCCCGC CTAATGA GCGGGCTTTTTTT 3’

模板DNA: 3’GGGTCGGGCG GATTACT CGCCCGA AAAAAA5’

RNA: 5’CCC AGCCCGC CUAAUGA GCGGGCUUUUUUU3’

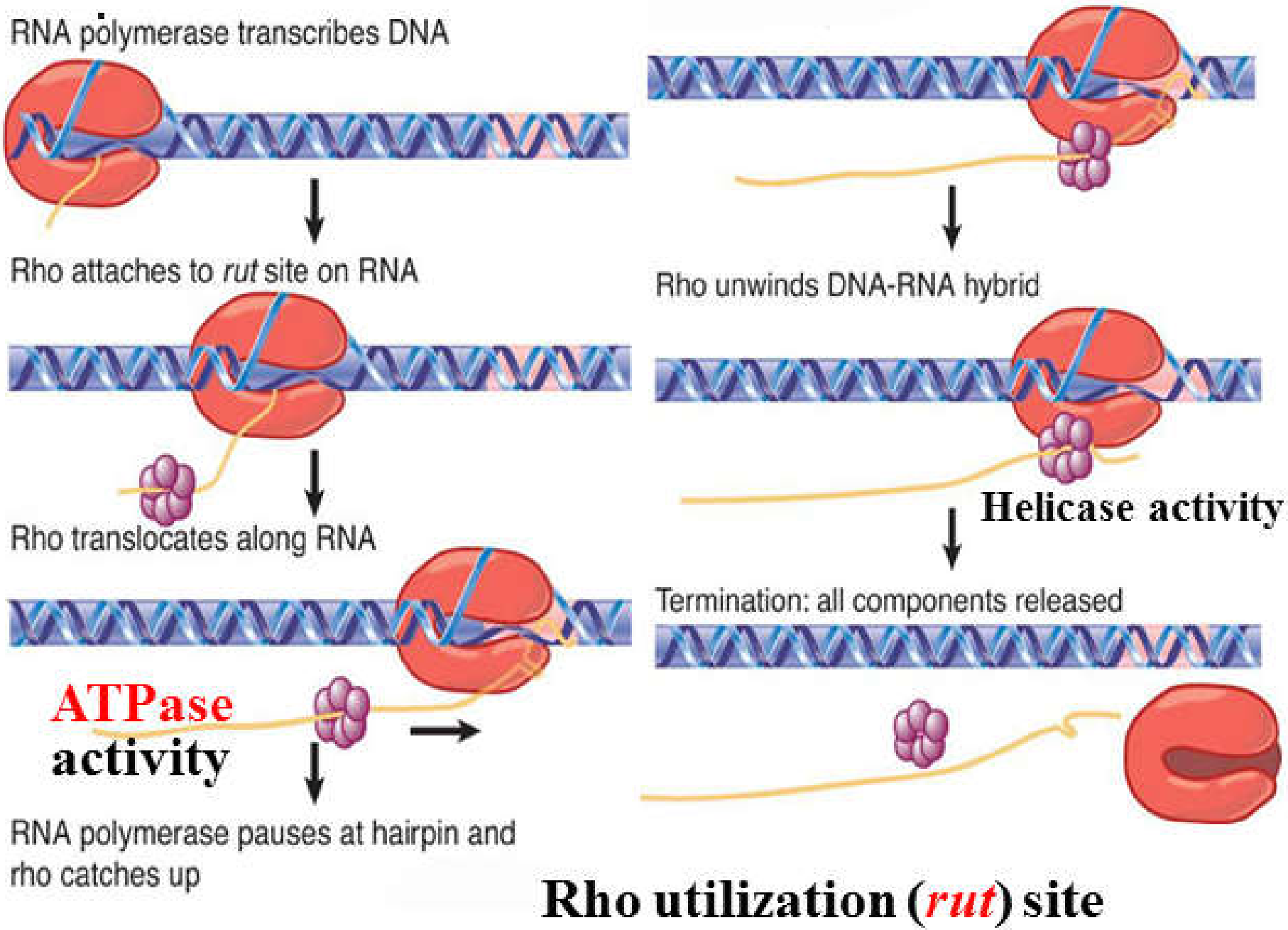


- As the hairpin forms, it further destabilizes the RNA–DNA hybrid.
- The A-U base-pairing is less stable that favors the dissociation.



4.3.2 Rho (ρ)-dependent termination

- Rho (ρ) is a termination factor (终止因子) which assist RNA polymerase in recognizing termination signals.
- Rho-dependent terminators consist of an inverted repeat (not GC-rich), but no string of Ts.


RNA polymerase transcribes DNA
Rho attaches to *rut* site on RNA
Rho translocates along RNA
ATPase activity
RNA polymerase pauses at hairpin and rho catches up
Rho unwinds DNA-RNA hybrid
Helicase activity
Termination: all components released
Rho utilization (*rut*) site

Vol 463|14 January 2010|doi:10.1038/nature08669 nature

His_6标记镍珠上的Rho，发现被镍珠挡住的延伸复合体(EC)中RNA产物长度仅11 nt，意味着Rho与EC结合涉及RNA Pol而不是RNA。

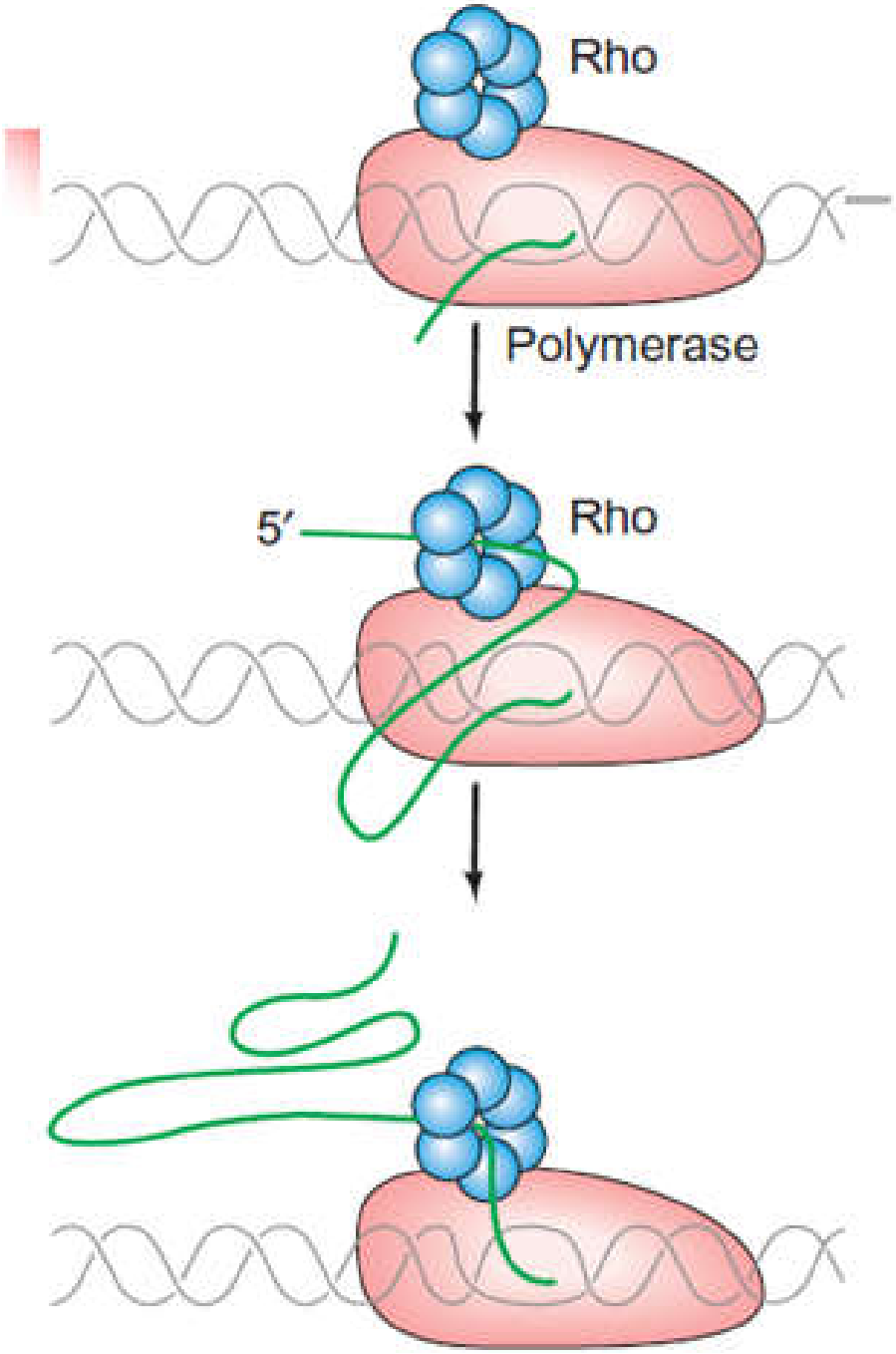
LETTERS

An allosteric mechanism of Rho-dependent transcription termination

Vitaly Epshtein[1]*, Dipak Dutta[1]*, Joseph Wade[2] & Evgeny Nudler[1]

Rho is the essential RNA helicase that sets the borders between transcription units and adjusts transcriptional yield to translational needs in bacteria[1–3]. Although Rho was the first termination factor to be discovered[4], the actual mechanism by which it reaches and disrupts the elongation complex (EC) is unknown. Here we show that the termination-committed Rho molecule associates with RNA polymerase (RNAP) throughout the transcription cycle; that is, it does not require the nascent transcript for initial binding. Moreover, the formation of the RNAP–Rho complex is crucial for termination. We show further that Rho-dependent termination is a two-step process that involves rapid EC inactivation (trap) and a relatively slow dissociation. Inactivation is the critical rate-limiting step that establishes the position of the termination site. The trap mechanism depends on the allosterically induced rearrangement of the RNAP catalytic centre by means of the evolutionarily conserved mobile trigger-loop domain, which is also required for EC dissociation. The key structural and functional similarities, which we found between Rho-dependent and intrinsic (Rho-independent) termination pathways, argue that the allosteric mechanism of termination is general and likely to be preserved for all cellular RNAPs throughout evolution.

determine the parameters of the EC that define the termination site, we selected two representative complexes on each template for detailed biochemical analyses: EC131 (termination-prone) and EC154 (termination-resistant) on the trpt' template (Fig. 1a), and EC158 (termination-prone) and EC169 (termination-resistant) on the aRut template (Fig. 1b). Each individual complex was prepared by walking of the EC followed by treatment with Rho (Supplementary Figs 2 and 3). We monitored inactivation (shown by the inability of the EC to extend the nascent RNA after chasing with a limited set of NTPs) and dissociation (shown by RNA release after washing the beads). Whereas the rate of dissociation was similar between EC131/EC158 and EC154/EC169, the rate of inactivation was significantly faster for EC131/EC158 (Fig. 1). The inactivation and dissociation rates were identical for termination-resistant EC154 and EC169. These rates were similar to the dissociation rates of EC131 and EC158 (Fig. 1, middle and right panels). Similar results were obtained for other complexes (Supplementary Figs 2 and 3). Thus, Rho termination can be viewed as a two-phase process: a fast inactivation (irreversible trap) followed by slow dissociation with inactivation as the rate-limiting step, which determines the position of natural termination sites.



(1) Rho binds to the RNA polymerase in an elongation complex.

(2) When the RNA transcript is long enough, Rho binds to it via *rut* site and feed it through Rho itself.

(3) Polymerase pauses at a terminator, which allows rho to tighten the RNA and trap the elongation complex.

(4) Terminating transcription

原核生物转录过程小结

起始
- 识别
- 结合
- 解链 (-10)
- 合成9~10 nt

延伸
- NTP为底物，延伸RNA链5’→3’
- 延伸过程σ随机解离

终止
- RNA pol滑动到终止子
- 不依赖/依赖rho的终止：RNA链脱落、RNA pol释放

Summary

1. Definition and general characteristics of transcription
2. Composition and function of *E. coli* RNA polymerase
3. Structures of *E. coli* σ^{70} promoter
4. Transcription process in prokaryotes
5. Structural features and termination modes of terminators in prokaryotes