

# 遗传学 (第3版)

## 第14章 基因表达调控

1. 大肠杆菌乳糖操纵子的结构与调控机制
2. 原核生物的其他类型操纵子及其调控机制
3. 原核生物基因表达的翻译调节
4. 真核生物基因转录水平的调节
5. 真核生物基因转录后水平的调节
6. 真核生物基因的翻译和翻译后水平的调节
7. 非编码RNA对基因表达的调控作用



原核生物是单细胞生物，以**操纵子（operon）**为单位的**调控系统是其主要的调控方式**。此外，也有许多翻译水平上的调控，如严紧反应与核糖体蛋白质合成的自身调节等。

真核生物有细胞质和细胞核，其基因的转录和翻译分别在细胞核和细胞质中进行，因而**处在一个非常纷繁复杂的控制系统中**，表达通路的每一步都受到严格的调节。为适应机体生长和发育分化，**各类细胞中都具有特异基因的表达和相应的调控机制**。本章在真核生物基因表达调控的讨论中我们将侧重于RNA聚合酶 II 所转录的这类基因的表达调控。



## 14.1 大肠杆菌乳糖操纵子的结构与调控机制

### 14.1.1 大肠杆菌乳糖操纵子的结构

1961年 F.Jacob和J.Monod 研究大肠杆菌乳糖代谢系统，提出操纵子学说（operon theory）

实验证明：结构基因 *lacZ*、*lacY*和*lacA*连接在一起，分别编码 $\beta$ -半乳糖苷酶（Z）、半乳糖苷透性酶（Y）和硫代半乳糖苷转乙酰基酶（A），形成一个转录单位。转录从启动子（promoter, *P*）开始，并受操纵基因（operator, *O*）和调节基因（regulator, *I*）的控制。启动子和操纵基因位于乳糖结构基因的上游，依次互相连接。这样依次排列的*P*、*O*、*Z*、*Y*、*A*序列片段便构成了一个共表达的遗传单位——乳糖操纵子（*lac operon*）（图14-1）。调节基因是一个独立的转录单位，它有自己的启动子。其表达产物为阻遏物（repressor），阻遏物既能阻止转录，又能识别小分子的诱导物。

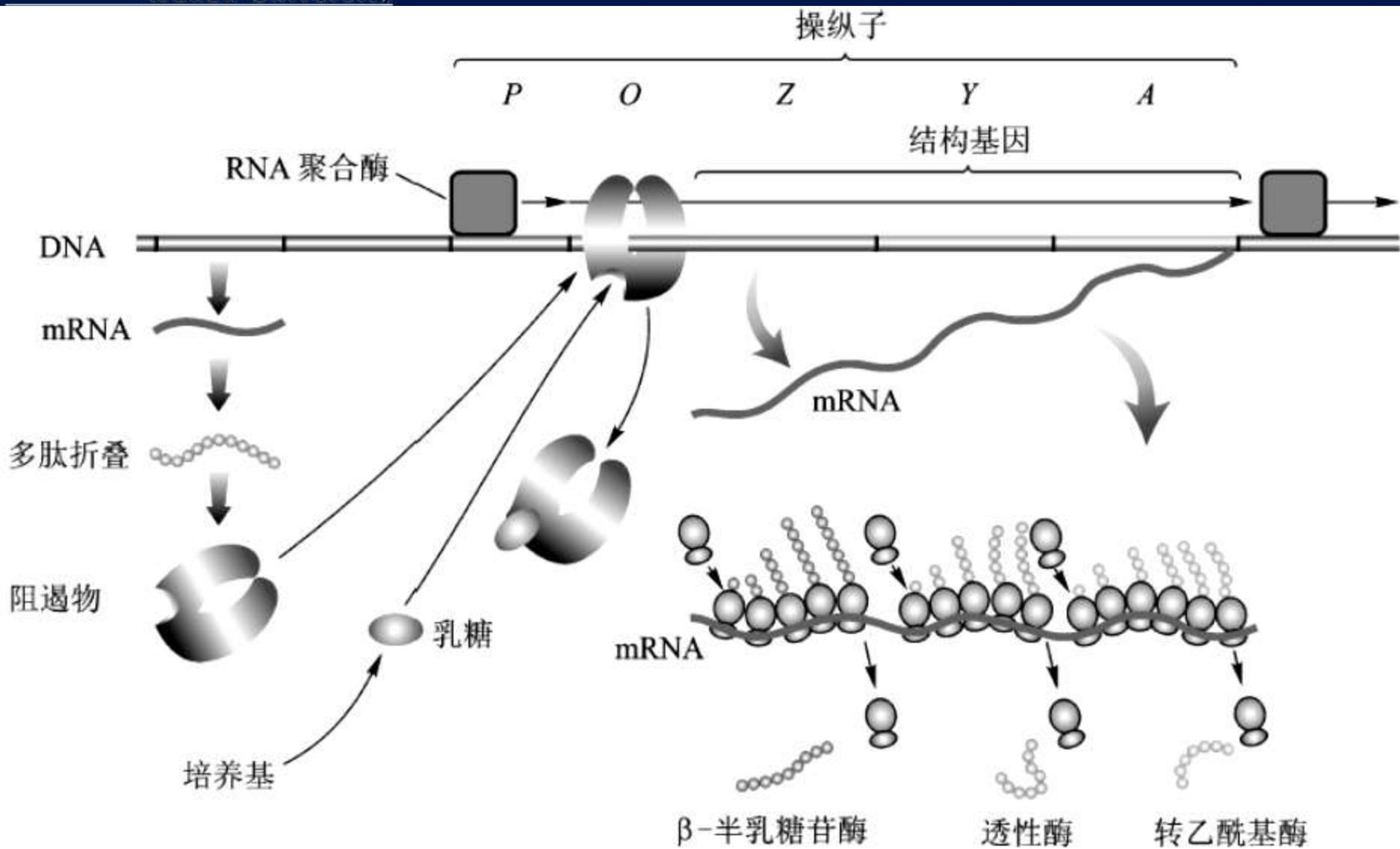


图 14 - 1 乳糖操纵子的结构(引自 Griffiths 等,2005)



## 14.1.2 乳糖操纵子的调控机制

### (1) 乳糖操纵子的负控制

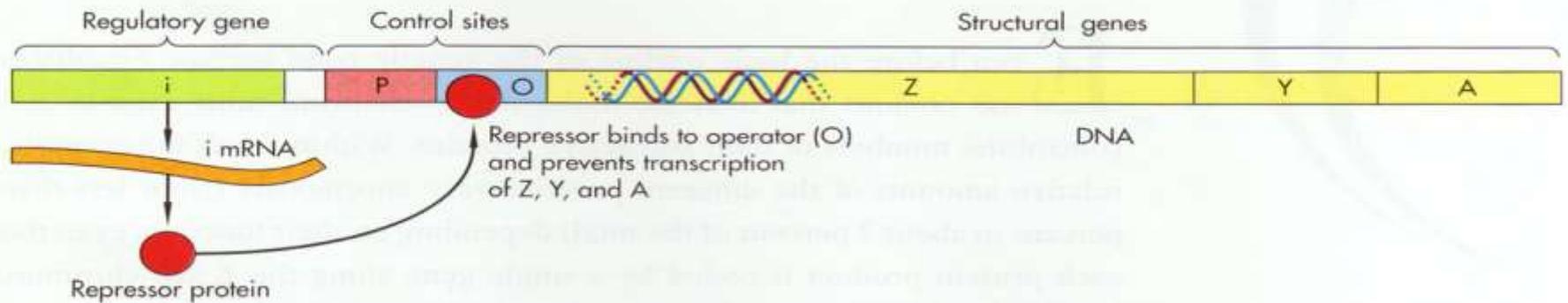
乳糖操纵子中结构基因 $Z$ 、 $Y$ 、 $A$ 转录为mRNA受到操纵基因的控制：  
阻遏物结合到操纵基因上 $\Rightarrow$ 阻断RNA聚合酶的转录活动 $\Rightarrow$ 结构基因转录关闭，未结合则使基因转录开启。

原因是P和O在序列上有一定重叠，O被阻遏物占据时，RNA聚合酶就不可能与P结合，因而不能催化转录。

当没有乳糖等诱导物时，阻遏物与操纵基因结合，使结构基因不能表达；它一旦与诱导物结合，便会发生构象变化、丧失活性，不再能结合操纵基因。在这种情况下，RNA聚合酶可顺利地通过操纵基因，启动结构基因的转录，3个基因所转录出的mRNA进而翻译成蛋白质（图14-1）。可见，调节基因表达的阻遏物的作用在于阻止转录，这种作用原理称为负控制（negative control）。



## REPRESSED STATE OF LACTOSE OPERON



## INDUCED STATE OF LACTOSE OPERON

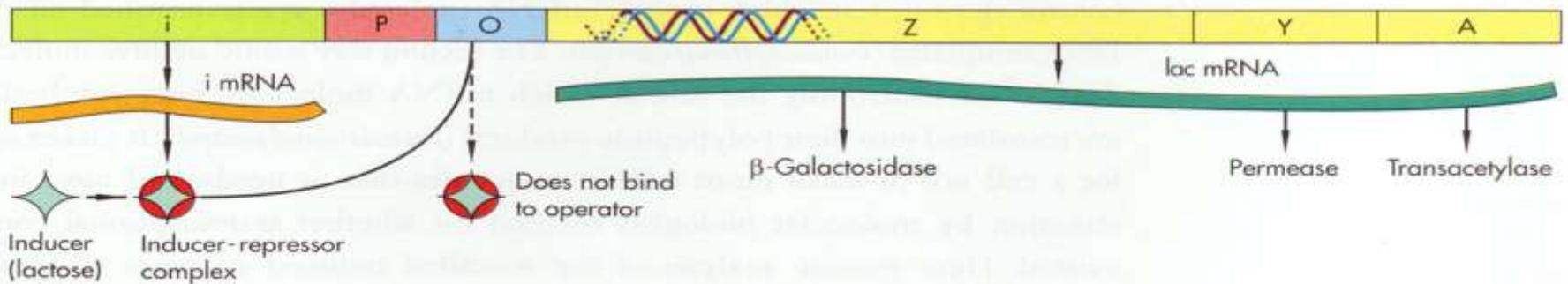
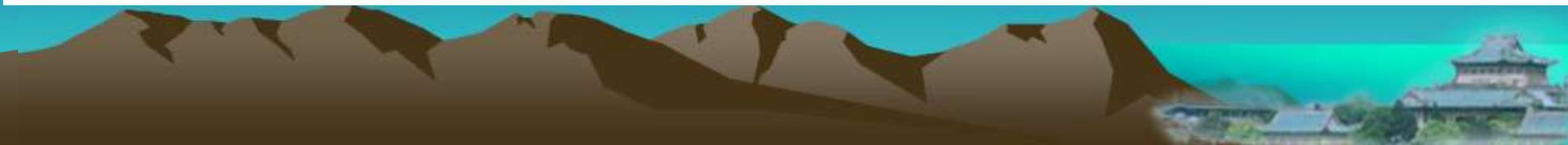


FIGURE 4-1

Repressors and inducers control the functioning of the genes belonging to the lactose (*lac*) operon. The regulatory gene (*i*) codes for the lactose repressor. The *P* segment of the DNA chain is the “promoter” and is discussed in the next figure.



Jacob和Monod的研究表明，若 $\beta$ -半乳糖苷酶基因  $lacZ$  由  $Z^+$  突变为  $Z^-$ ，则失去了合成 $\beta$ -半乳糖苷酶的能力

**组成型（或恒定型）突变体**（constitutive mutant）：  
原来只有诱导物存在时才能进行酶合成的诱导性菌株，变成没有诱导物时也能进行酶合成的突变型。**组成型突变多发生在  $I$  基因和  $O$  基因上，如野生型  $I^+$  突变为  $I^-$ ，而  $O^+$  则突变成  $O^c$ 。**

**超阻遏突变体**（superrepression mutant）：丧失了所有合成结构基因产物的能力，如  $I^s$ 。



## ① 调节基因 $I$ 的遗传分析

Jacob和Monod在对  $I^-$ 和  $I^+$ 的部分二倍体研究中，发现对结构基因  $lacZ$  的调控， $I^+$ 对  $I^-$ 呈显性（比较表14-1中的第1~3号菌株）。第3号和第4号菌株的结果说明  $I$  基因是反式作用，即跟其靶基因之间无论是顺式排列还是反式排列，都能够起到调控作用。

$I$  基因的另一突变型  $I^s$  则丧失了诱导所有结构基因表达的能力，在部分合子  $I^sZ^+Y^+/F' I^+Z^+Y^+$ 中  $I^s$ 呈显性，该菌株在诱导物存在时既不合成  $\beta$ -半乳糖苷酶，也不合成半乳糖苷透性酶（比较表14-1中的第5~6号菌株）。

## ② 操纵基因O的遗传分析

阻遏物产生功能失活⇨组成型突变，但在*I*基因未发生突变、能产生正常阻遏物的情况下，为什么也会出现菌株由诱导型转变为组成型呢？Jacob和Monod等设想，在邻近整套结构基因的上游有一段操纵基因序列，是阻遏物所识别并专一结合的部位。操纵基因序列的突变可导致阻遏物不能识别和结合到该部位上，从而造成乳糖操纵子的结构基因持续表达（图14-2）

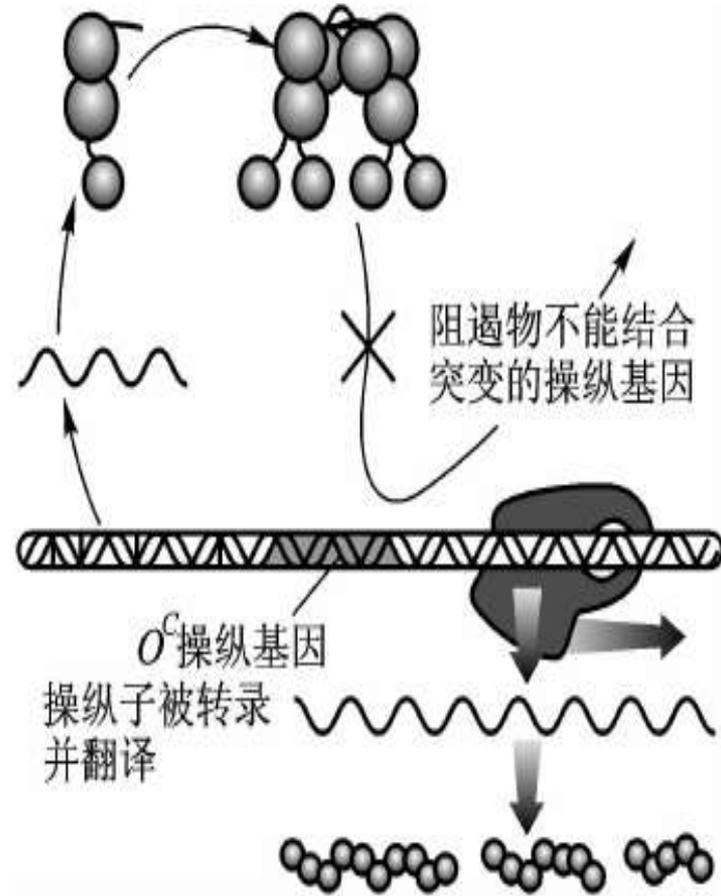


图 14 - 2 操纵基因突变导致组成型表达(引自 Lewin,2007)

## 操纵基因组成型 ( $O^c$ ) 和调节基因组成型突变体 ( $I^-$ ) 二者究竟有什么区别呢?

Monod等利用F' lac质粒构建了多种组合的部分二倍体, 以研究说明二者的关系 (图14-3):

(a) 表型为野生型: 野生型LacI蛋白是一种反式作用因子。 $lacI^+$ 基因编码的野生型阻遏蛋白能够在细胞内自由扩散, 它既能与质粒本身的操纵基因位点 $O^+$ 结合, 也能够跟细菌染色体上的操纵基因位点 $O^+$ 结合

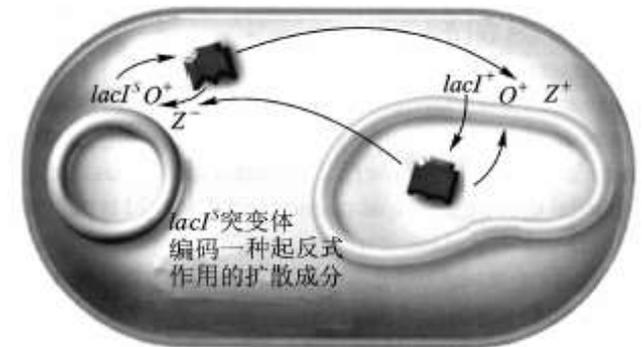
(b) 维持阻遏状态: LacI<sup>S</sup>蛋白也是一种反式作用因子。由 $lacI^S$ 编码的超阻遏蛋白即使在有诱导物存在时也依然能够结合在质粒和细菌染色体的操纵基因位点 $O^+$ 上

(c)  $lacO^c$ 是一种顺式作用位点。 $O^c$ 突变只能影响跟它位于同一条DNA分子上的结构基因 $lacZ^+$ , 使其表现为恒定的表达, 而质粒中 $O^+$ 的存在并不影响细菌染色体中 $lacZ^+$ 基因的表达

(a)  $lacI^+O^+Z^-$  菌中的 F' $lacI^+O^+Z^-$  质粒



(b)  $lacI^S Z^+$  菌中的 F' $lacI^S Z^+$  质粒



(c)  $lacI^+O^c lacZ^+$  菌中的 F' $lacI^+O^+ lacZ^-$  质粒

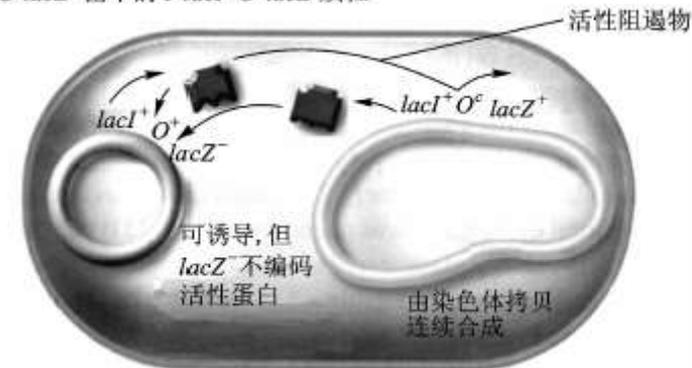


图 14-3 反式作用蛋白与顺式作用位点  
(引自 Hartwell 等, 2015)



## (2) 乳糖操纵子的正调控

在大肠杆菌中还发现有一种分子，其作用与阻遏物相反，它结合到操纵子的适当部位后可启动转录。这种调节机制属于正调控（positive regulation）。

1965年，B. Magasonik偶然发现，大肠杆菌中含有cAMP，而且胞内cAMP的浓度与培养基中的葡萄糖有关，葡萄糖浓度越高，胞内cAMP越少；反之，葡萄糖的浓度降低，则cAMP的浓度会相应提高[图14-4 (a)]

乳糖操纵子突变体的分析结果证明cAMP与分解物阻遏现象确实存在一定关系。

采用保护降解法已证实，在*lacP*区上游-72~-52 bp的部位有一个CAP结合位点；但是CAP需要与cAMP先形成复合物，使CAP构象发生变化后，才能与DNA上的这一特定序列结合，促进RNA聚合酶更有效地结合到*lac*启动子上，增强结构基因的转录[图14-4 (b)]。

研究表明，细菌只有在只有乳糖而没有葡萄糖时，才会产生乳糖代谢的一系列酶，而这种有效的调控是通过CAP-cAMP构成的正调控系统和诱导物-阻遏物构成的负调控系统这两条途径的相互配合实现的（图14-5）。

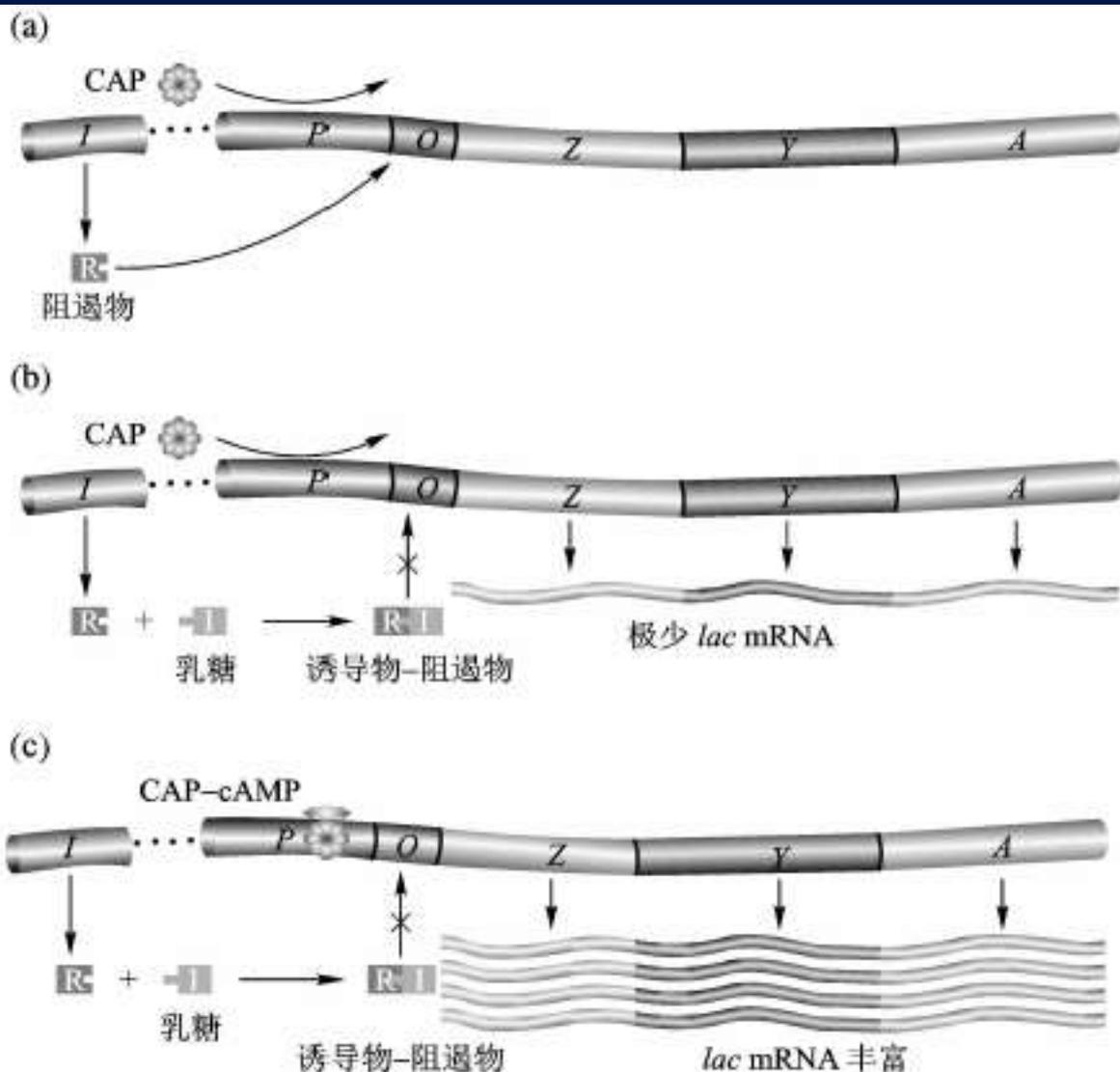


图 14-5 乳糖操纵子的正、负调控图解

(引自 Griffiths 等, 2015)

(a) 葡萄糖存在 (cAMP 低), 无乳糖, 无 *lac* mRNA (b) 葡萄糖存在 (cAMP 低), 乳糖存在, 产生少量 *lac* mRNA (c) 无葡萄糖 (cAMP 高), 乳糖存在, 大量产生 *lac* mRNA

## 14.2 原核生物的其他类型操纵子及其调控机制

### 14.2.1 半乳糖操纵子中的双重控制

半乳糖操纵子 (galactose operon) 由操纵基因 *galO*、启动子 *galP* 和紧密连锁的3个结构基因组成。

结构基因 *galk* 编码：半乳糖激酶 (galactokinase, *Galk*)

*galt* 编码：半乳糖转移酶 (galactose transferase, *GalT*)

*gale* 编码：半乳糖差向异构酶 (galactose epimerase, *GalE*)

*gal* 操纵子也有正、负两种调控方式。在 *gal* 操纵子中也发现有 *galI<sup>-</sup>* 和 *galO<sup>c</sup>* 两种组成型突变，表明 *galI* 阻遏物的作用机制是通过与操纵基因 *galO* 的结合而阻遏 *gal* mRNA 的合成。从转录水平上看，*gal* 操纵子是一个典型的负控制系统。该系统的诱导物是半乳糖。



分析*galP*的突变发现，*gal*操纵子有两个相互重叠的启动子：*galP*<sub>1</sub>和*galP*<sub>2</sub>。前者的起始依赖CAP-cAMP，转录起点为S1，位于+1；后者不依赖CAP-cAMP，转录起始点为S2，位于-5，两者相距4 bp。两个启动子各有自己的Pribnow框：分别位于-12~-6和-17~-11（[图14-6](#)）。测定不同条件下所得到的*gal* mRNA 5'端序列表明，转录时启动子的选择取决于CAP-cAMP存在与否。当CAP-cAMP存在而无葡萄糖时，RNA聚合酶与*galP*<sub>1</sub>结合，以S1为转录起始点，说明该过程可被CAP-cAMP激活。当CAP-cAMP不存在而有葡萄糖时，RNA聚合酶就与*galP*<sub>2</sub>结合，采用S2为转录起始点，该过程可被CAP-cAMP抑制，这就是*gal*操纵子基因表达的双重控制



## 14. 2. 2 阿拉伯糖操纵子的双向控制

在阿拉伯糖操纵子 (arabinose operon, *ara* operon) 中, 同一个 DNA 结合蛋白既可充当阻遏物, 又可作为激活因子, 提供了一个双向控制极好的例子。

催化阿拉伯糖代谢的酶分别由 *araA*、*araB*、*araD* 基因编码

阿拉伯糖操纵子的结构与组成见 [图14-7 \(a\)](#), 其双向控制的关键是:  
**AraC 蛋白能够以两种构象存在, 对应两种功能:**

(1) 在有阿拉伯糖时, AraC 蛋白与诱导物阿拉伯糖结合, 被激活成为一种诱导性蛋白, 跟 CAP-cAMP 复合物共同结合到 *araI* 区域, 促使 RNA 聚合酶结合于启动子上, 促进 *ara* 操纵子的转录, 发挥正调控作用 [[图14-7 \(b\)](#)];

(2) 在缺乏阿拉伯糖时, AraC 蛋白则会同时结合到 *araI* 和 *araO* 上, 形成一个环, 从而表现出阻遏物的性质, 抑制转录 [[图14-7 \(c\)](#)]。

## 14.2.3 色氨酸操纵子基因表达的衰减作用

### (1) 色氨酸操纵子的结构与调控因子

色氨酸操纵子 (tryptophan operon, *trp* operon) 5个编码酶的结构基因紧密连锁, 其中 *trpE* 和 *trpD* 的基因产物形成一个复合物, *trpB* 和 *trpA* 的基因产物也构成一个四聚体的复合物, 为色氨酸合成酶, 催化色氨酸合成的最后两步反应 (图14-8)。

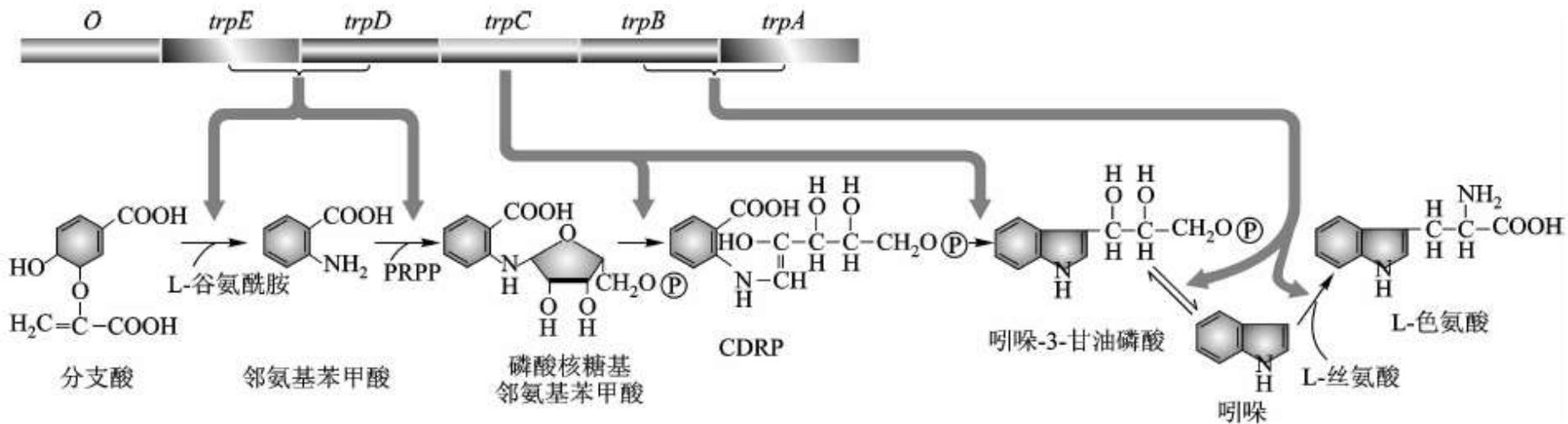


图 14 - 8 *trp* 操纵子的结构基因及其产物所催化的反应序列

(引自 Griffiths 等, 2015)



在 *trp* 操纵子的第一个结构基因 *E* 与操纵基因 *O* 之间有一段前导序列 (leader sequence, L), 在色氨酸合成中起着特殊的调控作用。此外, 编码阻遏物的基因 *trpR* 远离 *trp* 操纵子区。另外, 还有两个因子参与调控色氨酸操纵子—— $tRNA^{trp}$  (或称 Trp- $tRNA$ ) 和色氨酸- $tRNA$  合成酶; 它们的编码基因也在染色体上不同于色氨酸操纵子的位置。

### (3) 色氨酸操纵子的衰减作用

衰减调控仅出现在细菌中，它要求转录与翻译同步进行。此外，转录与翻译的速度必须大体一致，如果转录太快，或翻译太慢都无法协调衰减控制必需的几种因素之间的互作。

衰减作用主要涉及氨基酸如色氨酸合成有关的操纵子。*Trp*操纵子启动子下游有一段140 bp的引导顺序，终止信号位于+100~+140 bp之间，可形成发夹结构，但取决于RNA多聚酶与核糖体之间的相对位置。引导顺序+50~+60 bp有两个Trp密码子，当细胞中Trp水平很低时，核糖体会在这儿停留，拉开与RNA多聚酶间的距离，阻止终止信号形成发夹结构，转录正常进行。

当细胞中有足够的Trp存在时，核糖体紧跟RNA聚合酶，在转录到达近+140 bp位置时，终止信号形成发夹结构，RNA多聚酶脱离模板，转录停止。

其它一些氨基酸，如组氨酸，亮氨酸和苏氨酸的生物合成也采取衰减控制模式

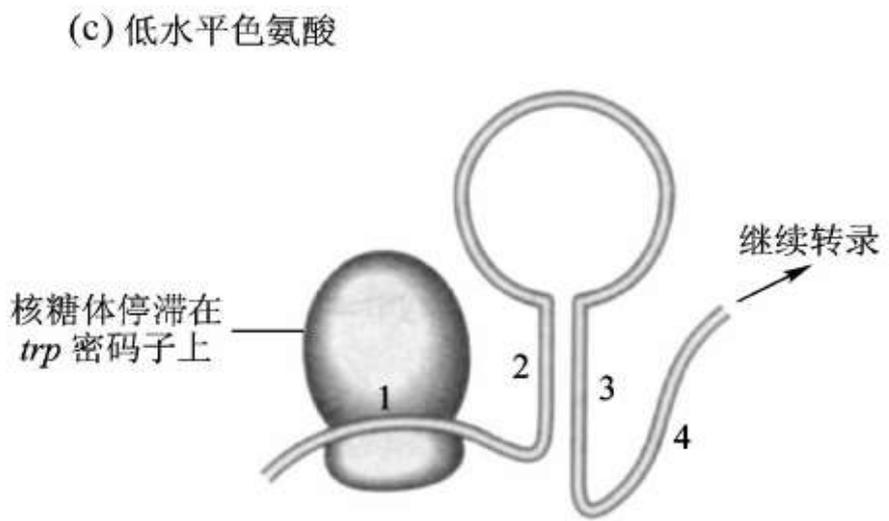
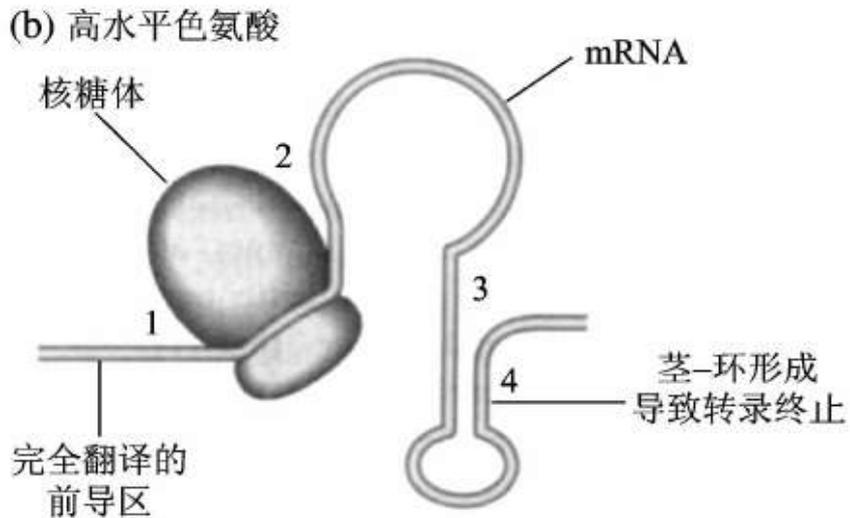
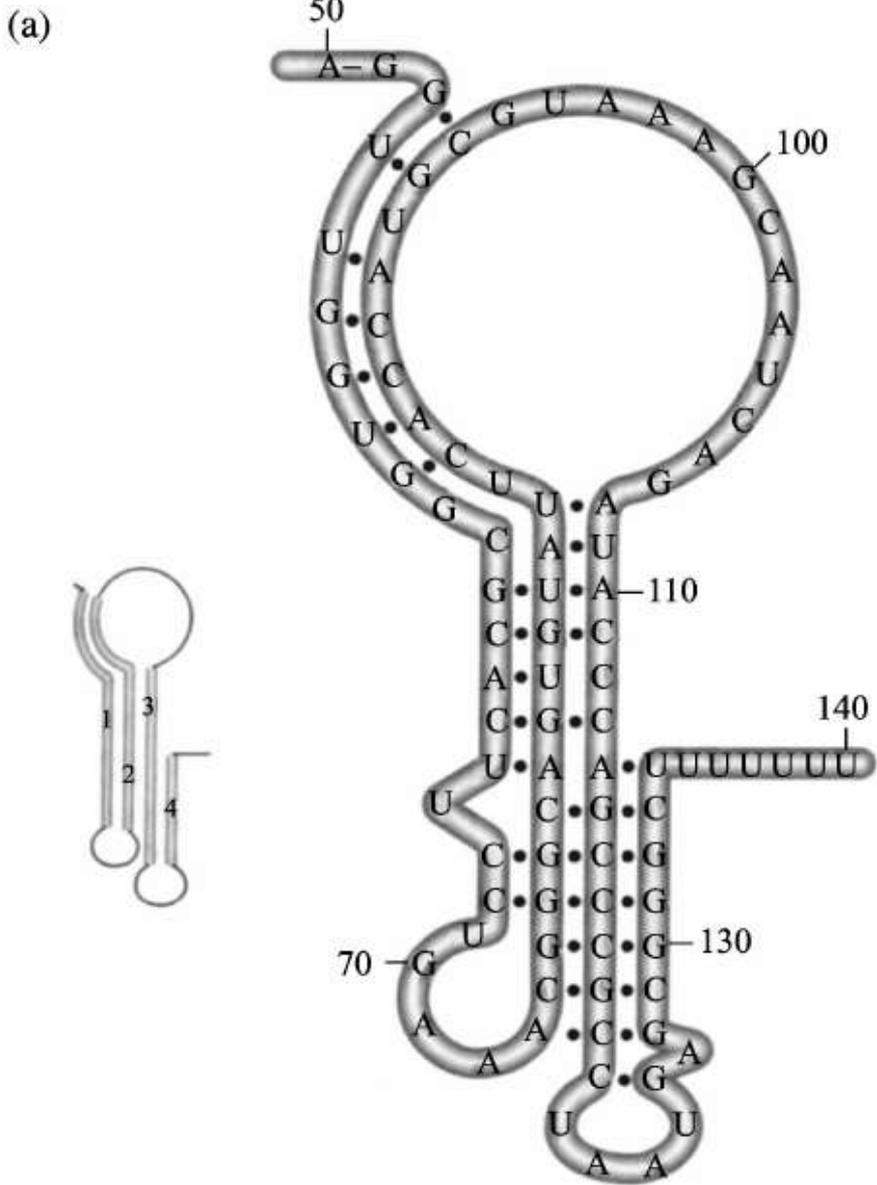


图 14 - 10 *trp* 操纵子衰减作用机制的模型(引自 Griffiths 等,2015)



## (4) 衰减作用的普遍性及其生物学意义

对衰减作用机制可以严格控制基因表达，依据细胞内某一氨基酸水平的高低进行表达调控，是一种灵活的多重调控方式。

一方面，当有活性阻遏物向无活性阻遏物转变的速率极低时，衰减系统能更加迅速地作出反应，使相应的氨基酸从较高浓度迅速下降；

另一方面，若外源氨基酸的浓度过低，细菌又没有其他相应的内源氨基酸的合成体系，那么细菌将难以支持自身的生长。有了衰减体系的调节，便可通过转录mRNA来增加相应的氨基酸合成酶的合成，提高该种内源性氨基酸的浓度。

衰减机制在控制基因产物的数量和种类的配比上起着快速而灵敏的调节作用，与阻遏物一起协同控制基因的表达，使之更为精密有效。因而，从生物进化角度来看，在细菌中像trp操纵子这样，除阻遏作用外，还演化出衰减子调控系统，具有重要的生物学意义。

## 14.3.1 原核生物蛋白质合成的严紧反应

在蛋白质合成中，核糖体直接或间接地控制着一系列酶的合成。当细胞饥饿时，蛋白质合成会骤然下降，细胞中的核糖体数目随之减少，rRNA和tRNA的合成大幅下降（可达10~20倍）。这种rRNA合成受控于氨基酸饥饿的现象称为**严紧反应（stringent response）**或**严紧控制（stringent control）**。

在大肠杆菌中发现野生型细胞在氨基酸饥饿时，细胞内很快增加一类异常的小分子化合物——鸟苷四磷酸和鸟苷五磷酸（ppGpp和pppGpp），其浓度可达500 $\mu\text{mol/L}$ ，同时还出现rRNA合成的突然停止。

更多研究表明，任何一种氨基酸的匮乏或者任何一种氨酰基tRNA合成酶失活的突变都会导致严紧反应，该反应的触发物是处于核糖体A位的无负载tRNA。当这种无负载tRNA进入A位后，由于氨基酸缺乏不能形成新肽键，而GTP不断消耗，于是出现空转反应（idling reaction），使ppGpp和pppGpp合成达到最高水平。

说明在细胞饥饿时，大量的GTP被用于合成ppGpp和pppGpp。结合在核糖体A位点的无负载tRNA被有负荷的tRNA替代后，这两种异常核苷酸的合成率便大大下降，并被spoT基因编码的酶催化而迅速降解，于是严紧反应得以逆转。

在氨基酸短缺的情况下，recA基因产物催化GDP转变为ppGpp；ppGpp能够直接与RNA聚合酶作用，从而改变RNA聚合酶的结构，影响其启动转录的能力，导致RNA的合成停止，rRNA的数量便急剧下降，最终使核糖体蛋白失去结合对象，核糖体的装配受阻。

由于这两种异常核苷酸是在细胞饥饿时合成，从而能够使细胞在整个调控网络中作出应急反应，如抑制核糖体和其他大分子的合成，抑制与氨基酸运转无关的转运系统，活化某些氨基酸操纵子的转录，表达蛋白水解酶等，以期节省或开发能源，帮助细胞渡过难关。

## 14.3.2 原核生物核糖体蛋白质合成的自身调节

在一些情况下，蛋白质的贮积常会限制它本身和其他一些基因产物的进一步合成。因此，当细胞内有游离的rRNA存在时，新合成的核糖体蛋白就首先与它结合，以启动核糖体的装配，使翻译继续进行；但是只要rRNA的合成减少或停止，游离的核糖体蛋白就开始积累，它们就会与自身的mRNA结合，阻断自身的翻译。同时也阻断同一顺反子mRNA中其他核糖体蛋白编码区的翻译（图14-11），使核糖体蛋白的合成及rRNA的合成几乎同时停止。不过，rRNA的合成是在转录层次上的调节，而核糖体蛋白的合成是在翻译层次上的调控。

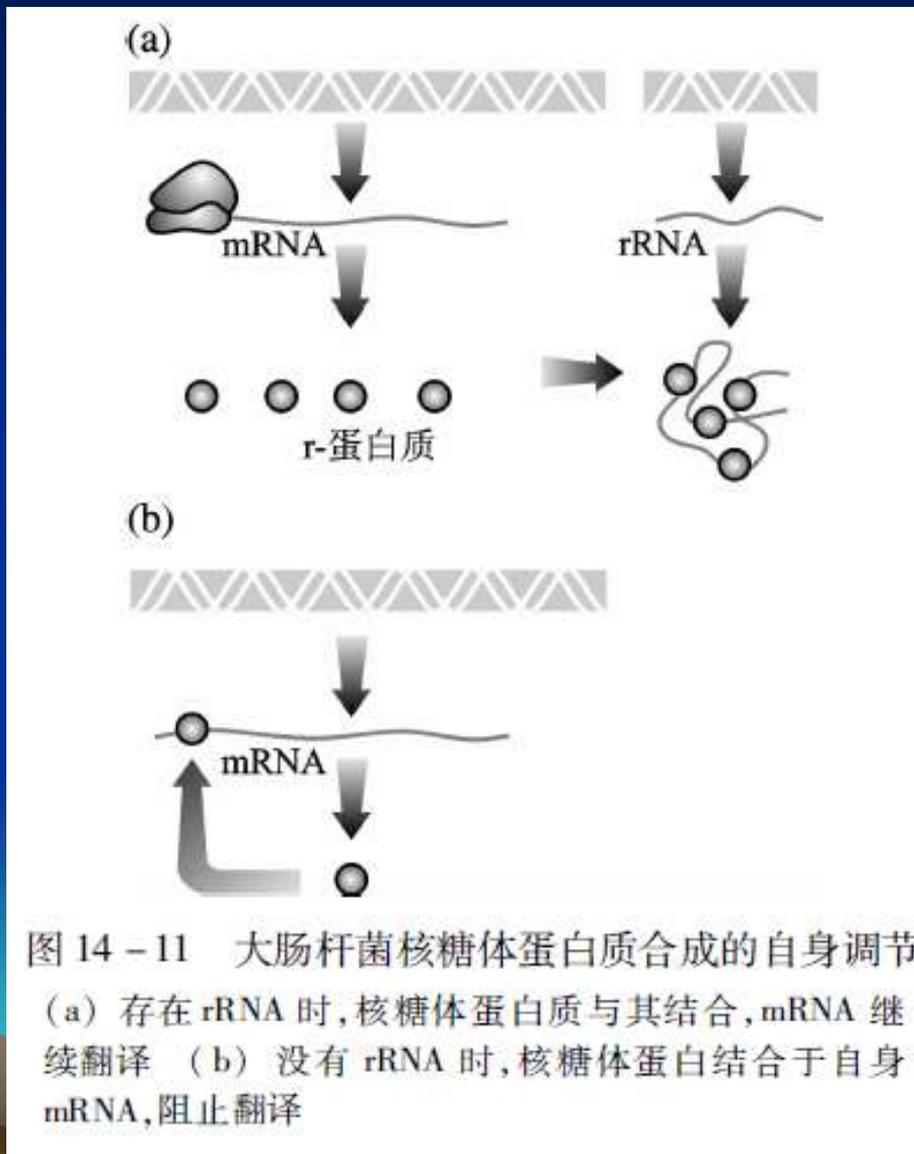


图 14 - 11 大肠杆菌核糖体蛋白质合成的自身调节  
(a) 存在 rRNA 时,核糖体蛋白质与其结合,mRNA 继续翻译 (b) 没有 rRNA 时,核糖体蛋白结合于自身 mRNA,阻止翻译



# 14.4 真核生物基因转录水平的调节

## 14.4.1 顺式调节元件与转录调节蛋白的结构与功能

### (1) 顺式调节元件

顺式调节元件 (cis-acting regulatory element)，又称顺式作用元件，是指同一DNA序列上一些对基因表达有调节活性的特定DNA调控序列 (regulatory sequence)。这种序列通常不编码蛋白质，多位于基因傍侧或内含子中。

不论是启动子、近启动子元件，还是远距离的增强子或沉默子都是顺式调节元件，都是不同反式作用的DNA结合蛋白的靶位点 (图14-12)。

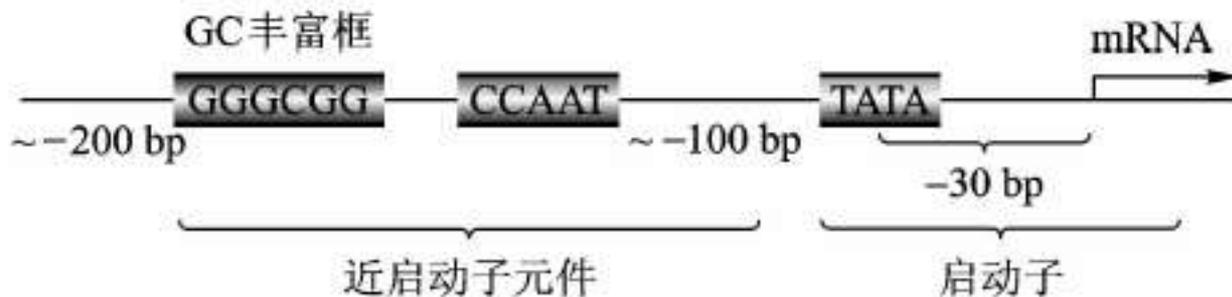


图 14 - 12 真核生物转录起始位点的上游序列

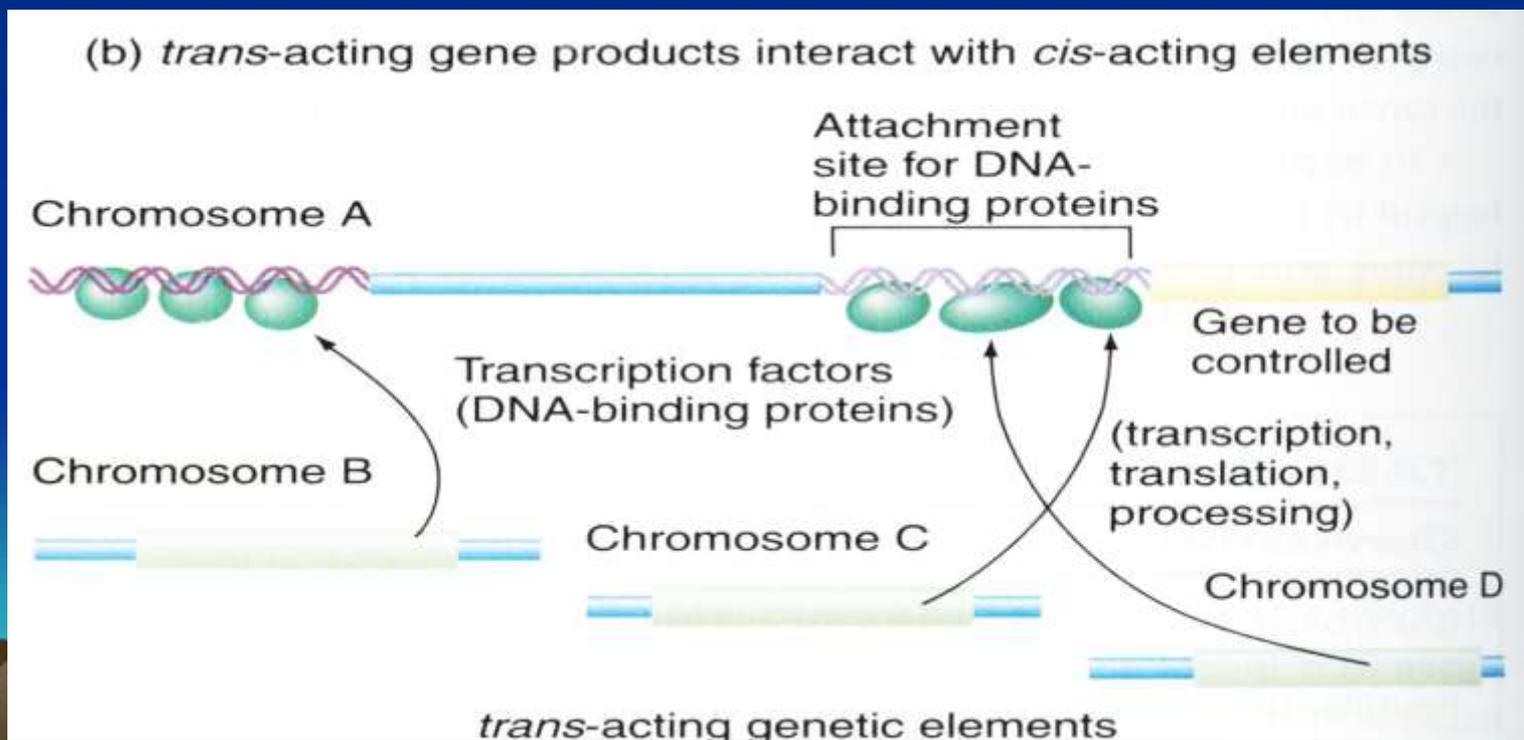
(引自 Griffiths 等, 2015)



## (2) 转录调节蛋白分类

反式作用：游离基因产物扩散至目标场所的过程。因此反式作用因子的编码基因与其识别或结合的靶核苷酸序列一般不在同一个DNA分子上。

反式作用因子（trans-acting factor）：通常是蛋白质或蛋白质亚基。





依据其功能反式作用因子大致可分为4类：

- 转录因子 (transcription factor)
- 激活因子 (activator)
- 辅激活因子 (coactivator)
- 阻遏物 (repressor)

① **通用转录因子** (general transcription factor, GTF)：在转录起始时首先与核心启动子结合并组成转录起始基本装置的调节蛋白，如TF II D等

② **激活因子**：那些能识别特定的共有元件，结合到启动子或增强子短序列上，以加强启动子上基本装置的效能，并提高转录频率的调节蛋白。

③ **辅激活因子**：是在激活因子和基本转录装置间提供一个“连接”，利用蛋白质与蛋白质之间的相互作用来辅助转录激活。

④ **阻遏物**：在真核细胞中发现的一类与转录起始相关的结合蛋白，它们结合于上游启动子元件或更远的沉默子位点，可抑制转录起始。



## (3) 转录调节蛋白的结构

### ① 转录调节蛋白的DNA结合结构域

真核生物转录调节蛋白的DNA结合功能域各具不同的特征，目前研究得较为清楚的有4种类型： $\alpha$ 螺旋-转角- $\alpha$ 螺旋结构域（helix-turn-helix motif）、锌指结构域（zinc finger motif）、亮氨酸拉链结构域（leucine zipper motif）、螺旋-袢环-螺旋结构域（helix-loop-helix motif）。

### ② 转录调节蛋白的转录激活结构域

激活域与DNA结合域不同，它没有特定的结构模式，而是以一种黏性表面与前起始复合物之间形成接触点来起作用。大致可分为3类：酸性功能域（acidic domain），富含酸性氨基酸残基；富含谷氨酰胺残基的功能域（glutamine-rich domain）；富含脯氨酸残基的功能域（proline-rich domain）。

### ③ 转录调节蛋白的模块式结构特征

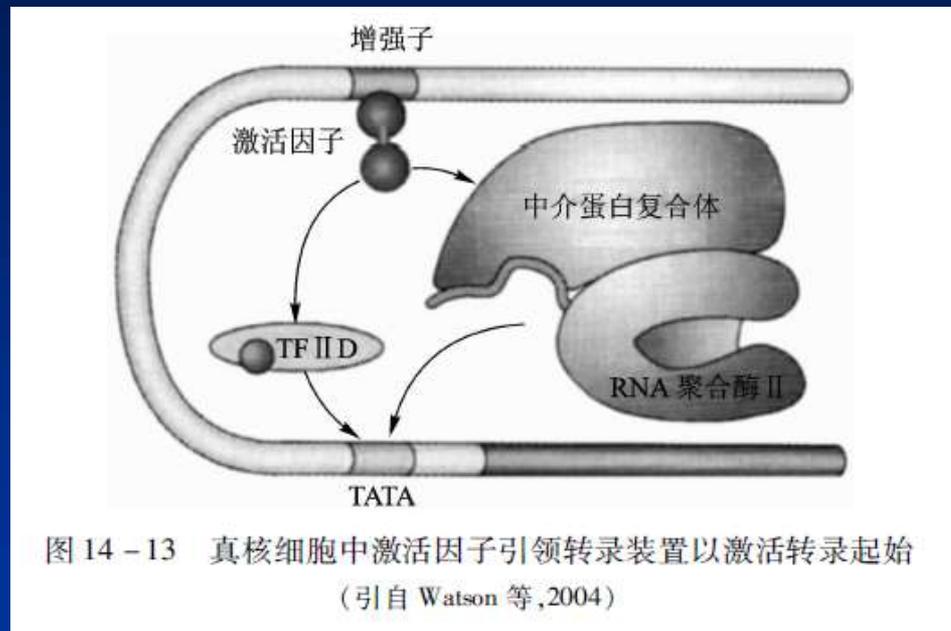
这些调节蛋白都是由结构上独立的DNA结合域和激活域共同组成的模块式结构。模块中含有一个或几个激活功能域，它们通过可弯曲的蛋白质区域连接到一个专一序列的DNA结合功能域上；有时还可对转录激活起一定的作用；激活功能域是通过与转录相关的其他蛋白质的结合来发挥功能。



## (4) 转录调节蛋白的主要功能

### ① 转录激活因子的功能

在真核生物基因转录的起始除需要一系列的通用转录因子外，还要有其他转录调节蛋白，如激活因子、中介蛋白复合体（mediator complex）等的协助，才能使RNA聚合酶



组合到转录装置上，并稳定地结合于启动子。除RNA聚合酶外，激活因子还可与转录装置中的其他复合物（如TF II D）相互作用，将它们引领到基因上（图14-13）。正是通过上述这些作用，激活因子促进完整的前起始复合物的形成。

又例如： $\beta$  干扰素的增强子中有4个控制元件，它们可在DNA双螺旋的同一面同时结合4种不同的转录调节蛋白，在增强子上产生一个多蛋白复合物，称为**增强体（enhanceosome）**（图14-14）。增强体一旦形成，其中的激活因子将同转录装置联系，并高水平地激活基因转录。增强子中各元件上分别结合的激活因子具有高度的协同性，保证了信号的紧密结合。

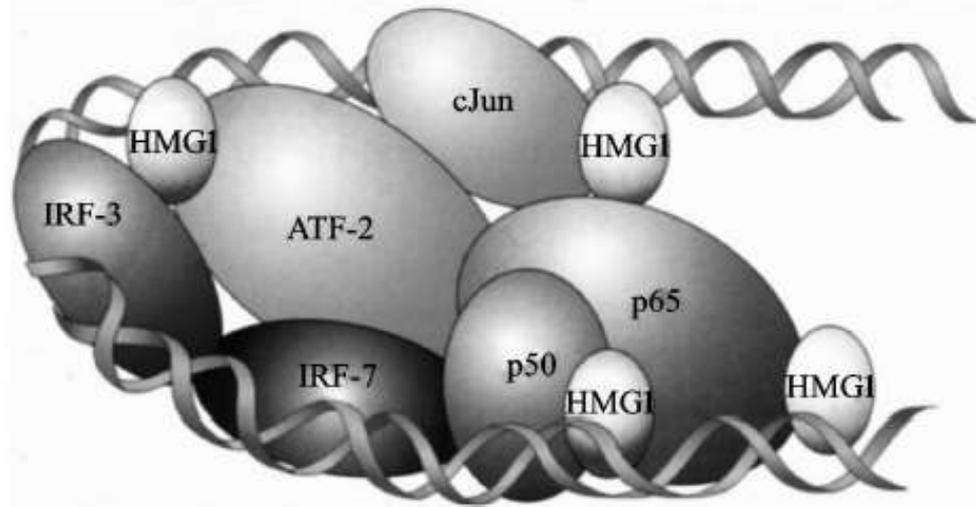


图 14 - 14 在人类  $\beta$  干扰素增强子上形成的增强体  
(引自 Lodish 等,2004)

IRF-3 和 IRF-7 是两个单体转录因子,cJun/ATF-2 和 p50/p65 (NF -  $\kappa$ B) 是两个异二聚体转录因子, 它们结合到增强子的 4 个不同的控制元件上,并在 HMG1 的促进下协同结合形成增强体

## ② 阻遏物对基因表达的影响

阻遏物的作用是它们结合于上游近启动子元件或更远的沉默子中的位点，**阻遏基因表达**。不同的阻遏物以不同的方式行使其抑制功能(图14-15)。有些阻遏物是通过与激活因子竞争同一增强子元件来阻挠激活因子，以致无法激活转录[图14-15(a)]。有的阻遏物则是采取直接抑制激活蛋白的方式，迫使后者无法发挥作用[图14-15(b)]。

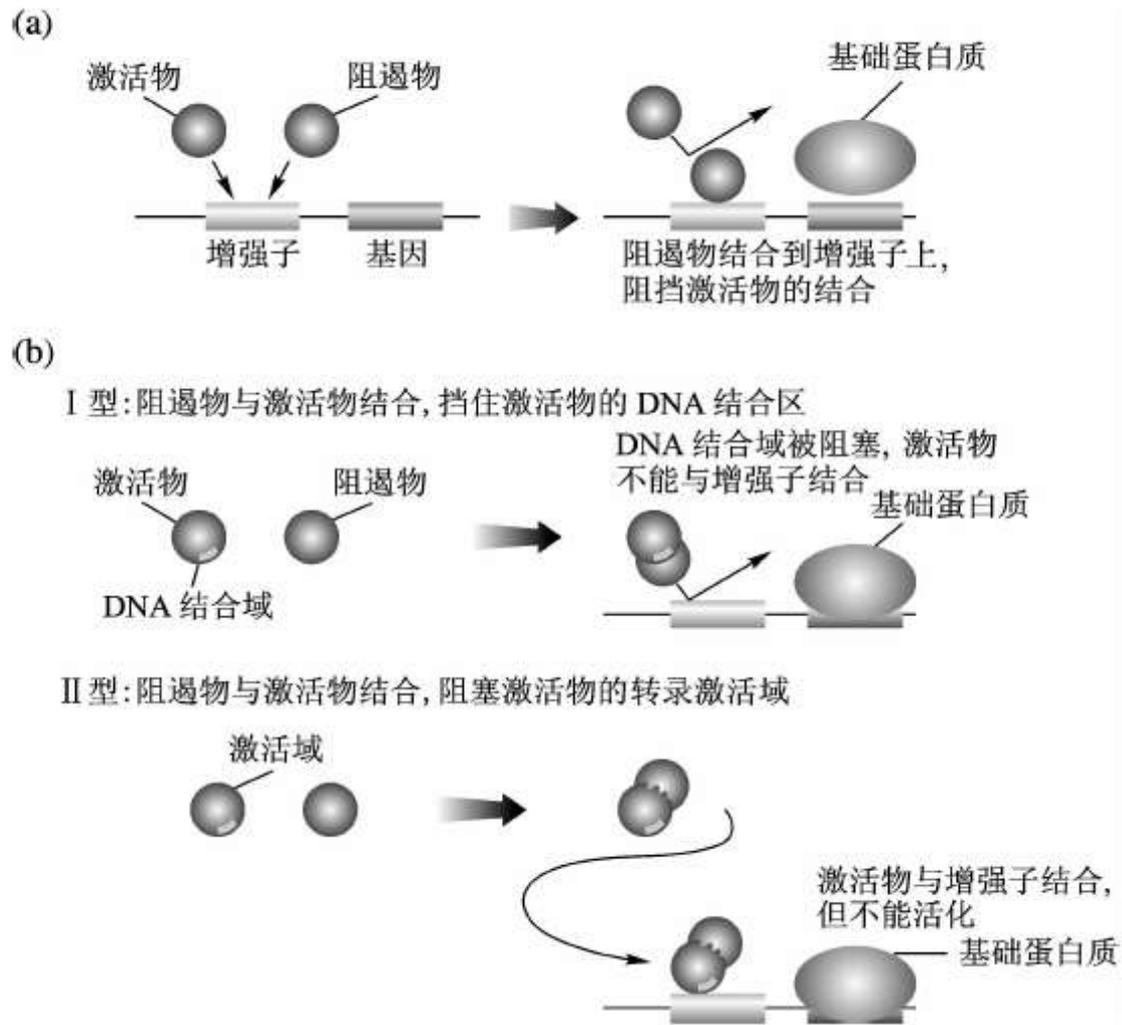


图 14-15 阻遏物抑制基因表达

(引自 Hartwell 等, 2011)

(a) 阻遏物与激活物竞争结合 (b) 遏制



## 14.4.2 基因表达的激素调节

多细胞真核生物的一些基因表达常受内、外激素的调控。许多甾类激素如蜕皮素、皮质素、雌激素、睾酮、甲状腺素、糖皮质激素和一些多肽激素（如胰岛素）等，都可以促进某些基因的转录。

甾类激素是一些较小的疏水性分子，可以穿过质膜进入靶细胞，与细胞质内或核内的相应受体形成复合物，**激素受体复合物可直接进入细胞核，调控基因转录**（图14-16）。



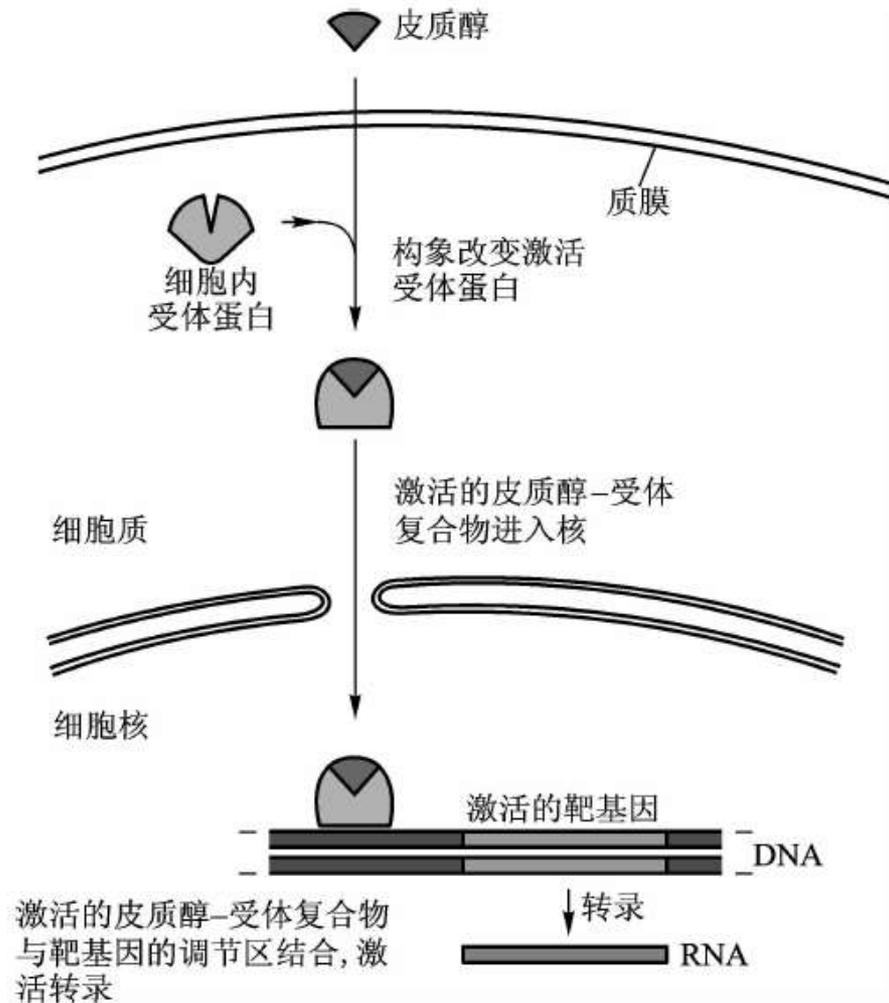


图 14-16 皮质激素激活基因调节蛋白以激发基因转录

(仿自 Alberts 等, 2004)

皮质醇扩散穿过质膜, 与细胞质中的受体结合。皮质醇-受体复合物经核孔进入核, 皮质醇所结合的受体被激活, 而后受体与 DNA 专一调节序列结合, 激活基因转录

## 激素是怎样调控基因转录的呢？

研究发现激素可以使与它相结合的受体蛋白发生某些变化，从而结合到染色质上以促进转录。如图14-17所示，一旦激素分子结合到受体的激素结合区，便导致抑制蛋白复合物解离，暴露出DNA结合区，受体被激活。

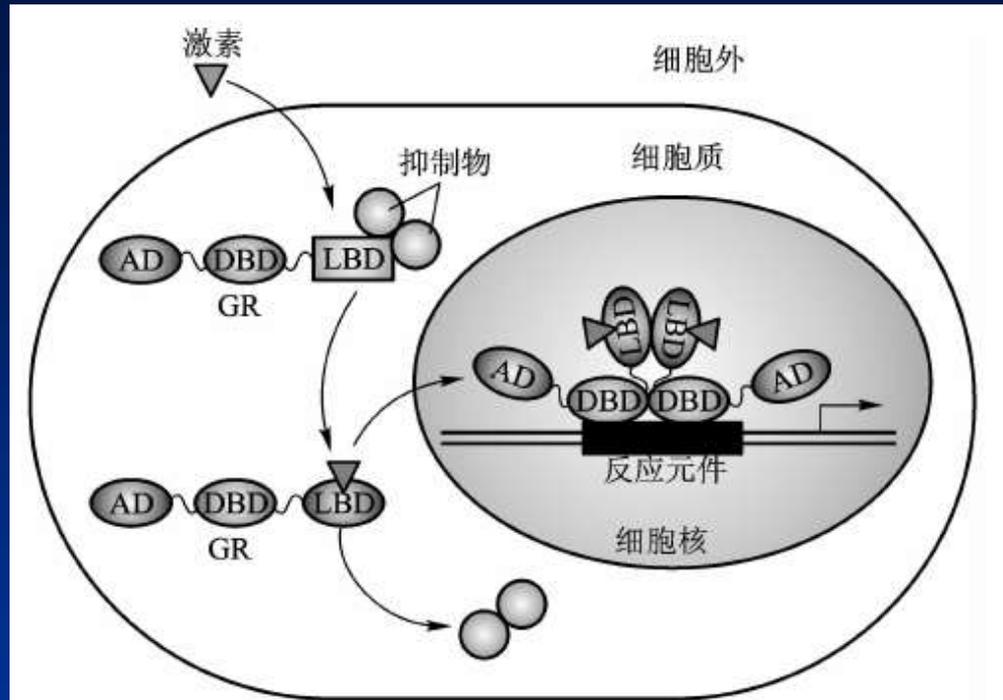


图14-17 激素激活细胞内受体的机制  
(引自 Lodish 等, 2013)

无激素时，受体的配体结合域（LBD）保持与抑制蛋白相互作用，处于失活状态。激素进入细胞后便结合到受体的配体结合域，使受体发生构象改变，与抑制物解离。随后结合着激素的受体转移入核内，其暴露出的DNA结合域（DBD）便结合到DNA上，使配体结合域和N端的另一个激活域（AD）刺激靶基因的转录。GR为糖皮质激素受体



## 14.5 真核生物基因转录后水平的调节

### 14.5.1 选择性剪接

一个基因的外显子和内含子共同转录在一条转录产物中，然后将内含子去除而把外显子连接起来形成成熟的RNA分子，这一过程称为**RNA剪接 (RNA splicing)**。

内含子的剪接一般都是发生在同一个基因内，切除内含子，相邻的外显子彼此连接，称为**顺式剪接 (cis-splicing)**。

**组成型剪接 (constructive splicing)**是指mRNA前体 (hnRNA) 只有一种剪接方式，剪接后仅产生一种成熟的mRNA分子。

**选择性剪接 (alternative splicing)** 又称可变剪接、变位剪接，是指同一种hnRNA可以采用几种不同的剪接方式，从而产生出不同的mRNA，有时甚至还会产生某种非编码的RNA分子。通过可变剪接产生的RNA分子又称为剪接变体 (splice variant)



例如，原肌球蛋白基因可以为脑、肝、骨骼肌、平滑肌和成纤维细胞编码  $\alpha$  原肌球蛋白，由这一基因转录出的 hnRNA 含有 11 个外显子，通过选择性剪接，可在不同细胞中产生不同的 mRNA，最终编码出不同的  $\alpha$  原肌球蛋白（图 14-18）。

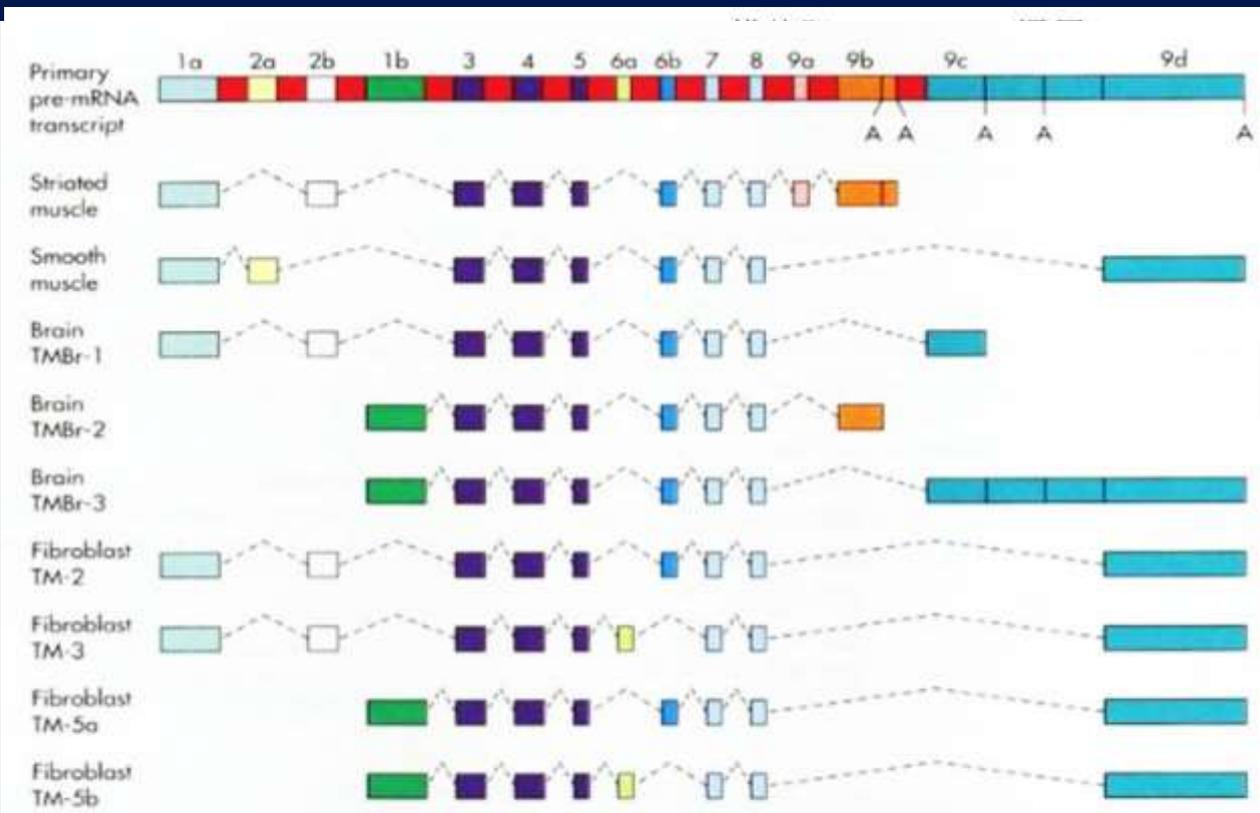


FIGURE 8-8

Complex patterns of eukaryotic mRNA splicing. The pre-mRNA transcript of the  $\alpha$ -tropomyosin gene is alternatively spliced in different cell types. The red boxes represent introns; the other colors represent exons. Polyadenylation signals are indicated by an A. Dashed lines in the mature mRNAs indicate regions that have been removed by splicing. TM, tropomyosin. (After J. P. Lees, et al., 1990.)

选择性剪接是一种重要的调节手段，使得一个基因所携带的遗传信息在转录后有所扩展。



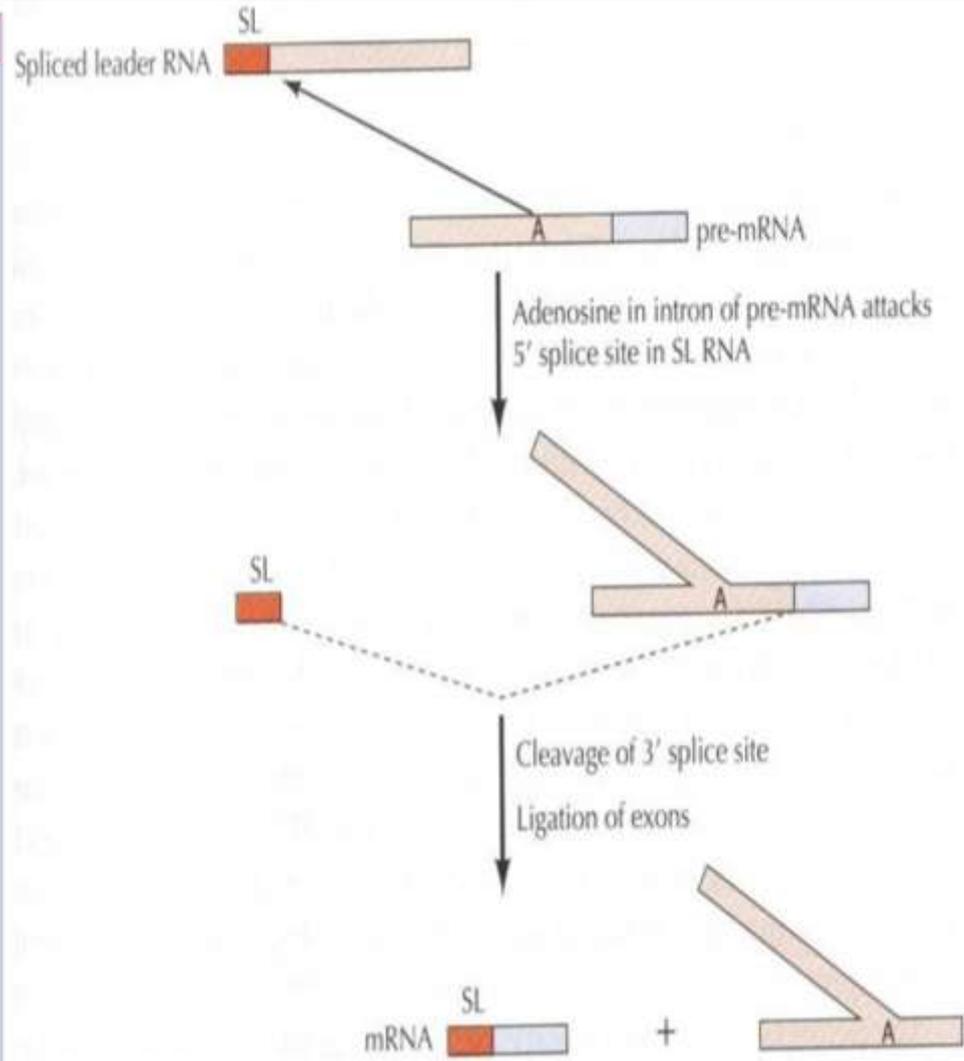
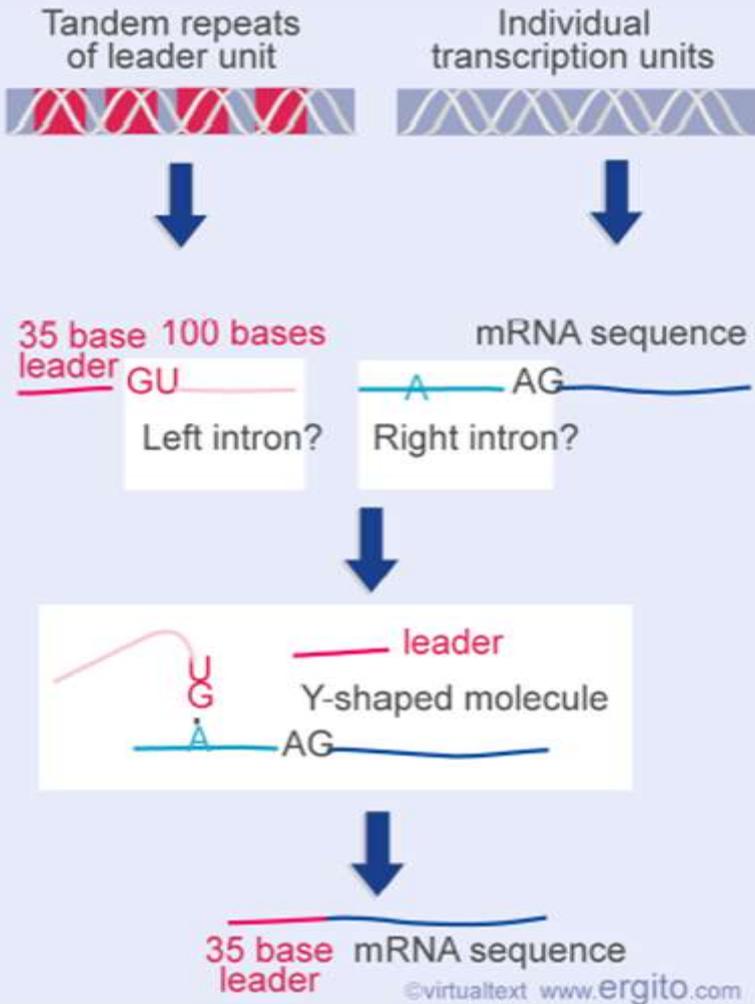
## 14.5.2 反式剪接

**反式剪接 (trans-splicing)** 是指将不同基因的外显子剪接后相互连接，成为一条成熟的mRNA分子。

通常经过这种剪接方式在mRNA上游非编码区的5'端拼接上一段**剪接前导序列 (splicing leader, SL)** 或称为小外显子 (mini-exon) 的RNA片段。这些片段原本不存在于相应的编码基因内，而是由其他DNA链转录而来。

现在已知，锥虫、线虫，以及植物叶绿体与线粒体中，反式剪接是RNA的主要剪接方式。

### SL RNA is *trans*-spliced



**Figure 24.25** The SL RNA provides an exon that is connected to the first exon of an mRNA by *trans*-splicing. The reaction involves the same interactions as nuclear *cis*-splicing, but generates a Y-shaped RNA instead of a lariat.





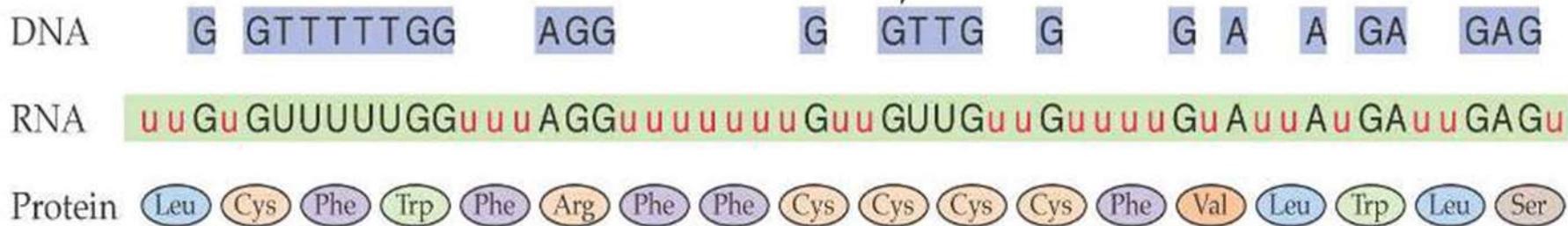
## 14.5.3 RNA编辑

**RNA编辑 (RNA editing)** 是指对前信使RNA (pre-messenger RNA, pre-mRNA) 的编码区进行碱基插入、删除或替换，以改变来源于DNA模板的遗传信息，翻译出不同于基因原编码的氨基酸序列的蛋白质。

或：基因转录产生的mRNA分子中，由于核苷酸的缺失、插入或置换，基因转录物的序列不与基因编码序列互补，使翻译生成的蛋白质的氨基酸组成，不同于基因序列中的编码信息，这种现象称为**RNA编辑**。

例如：四膜虫线粒体 *COIII* 基因转录区的 mRNA 的编辑

Region of *COIII* gene transcript:



例如图 14-19 哺乳动物载脂蛋白 B 前体 mRNA 的编辑

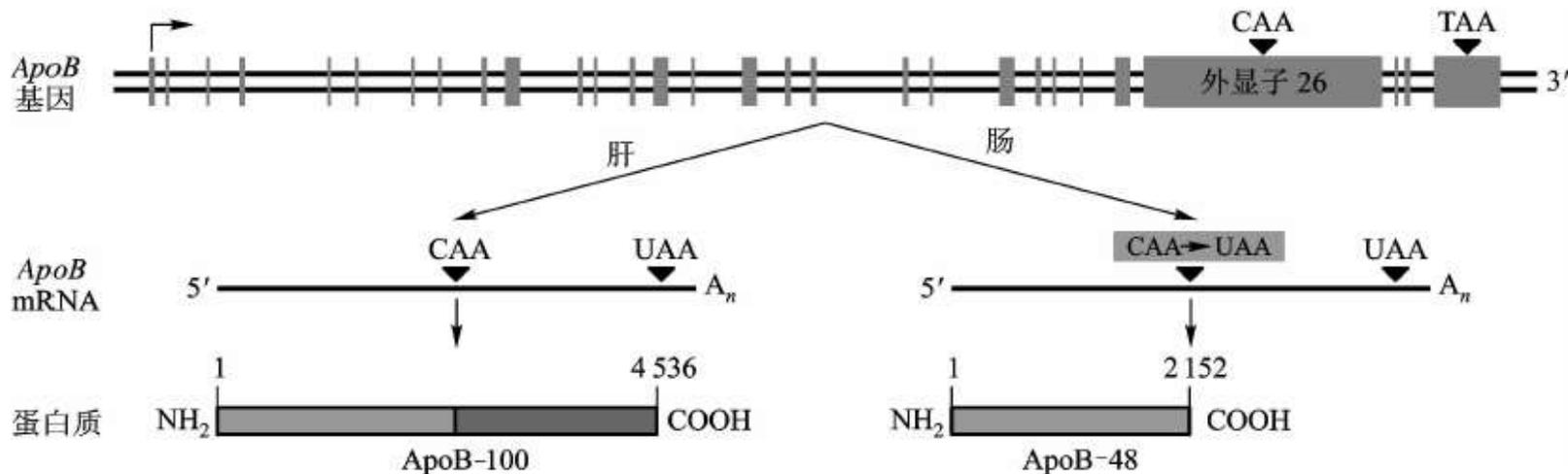


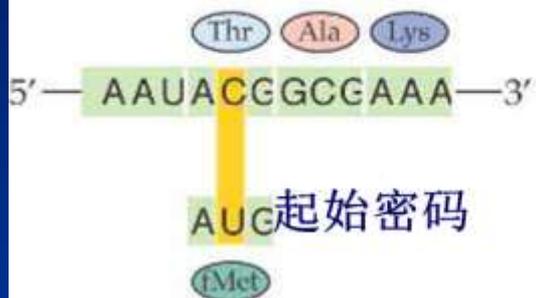
图 14-19 *ApoB* 前体 mRNA 的 RNA 编辑

(仿自 Lodish 等, 2013)

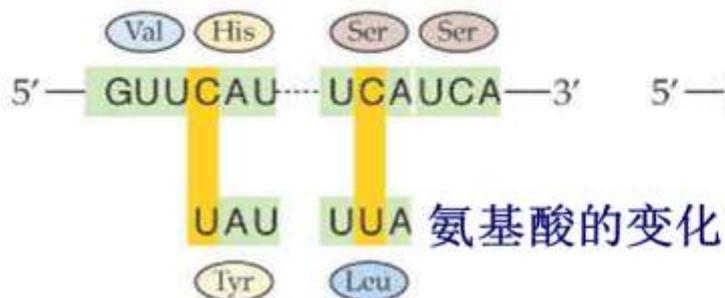
肝中产生的 *ApoB* mRNA 与初级转录物具有同样的序列, 其 mRNA 翻译成 ApoB-100。ApoB-100 含有两个功能域: 一个是与脂质结合的 N 端域, 另一个是与质膜上的 LDL 受体结合的 C 端域。小肠产生的 *ApoB* mRNA 中, 外显子 26 中的 CAA 密码子被编辑成 UAA 终止密码子, 结果小肠细胞产生 ApoB-48, 相当于 ApoB-100 的 N 端域

# mRNA编辑的几种类型

(A) Creation of initiation codon



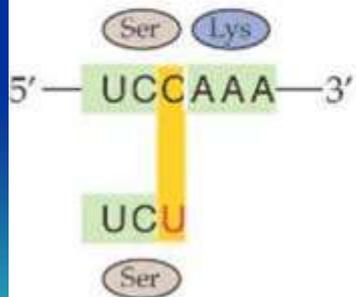
(B) Amino acid changes



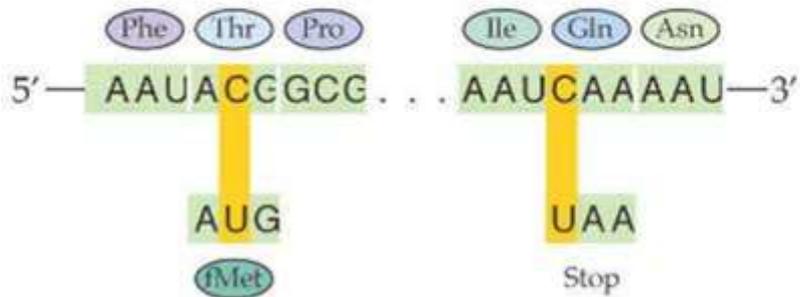
(C) Creation of a stop codon



(D) Silent editing 沉默编辑



(E) Creation of both initiation and stop codons



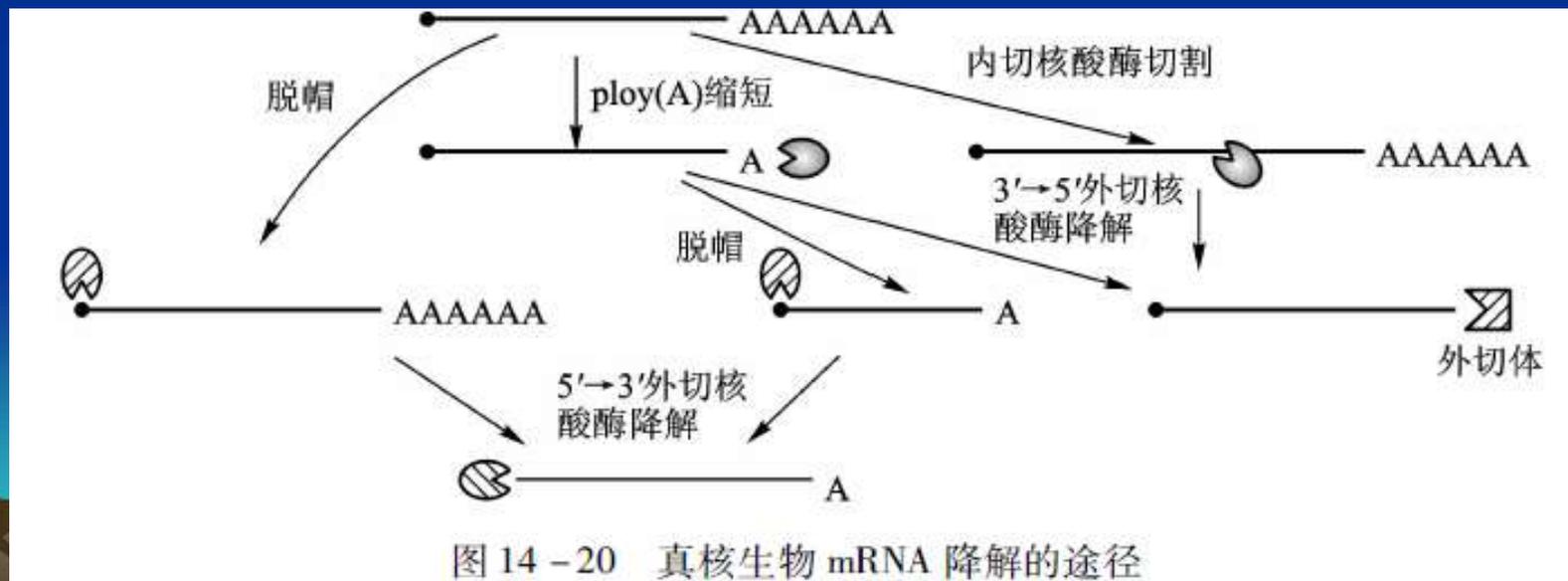


## 14.6 真核生物基因的翻译和翻译后水平的调节

### 14.6.1 翻译调节

#### (1) mRNA的稳定性

mRNA翻译水平的调控主要是在翻译过程的起始阶段。mRNA的正常翻译总是需要5'端与3'端之间的相互作用，终止翻译则需要破坏这两者间的正常作用。mRNA降解过程都起始于3'端。当3'端poly(A)发生变化时，也影响到5'端帽子结构，从而引起mRNA由5'端向3'端的降解（图14-20）。



## (2) mRNA非翻译区与翻译调控的关系

已知蛋白质的生物合成不仅与其mRNA的编码序列有关，而且还受到5'端和3'端非翻译区结构的调控。

① 5' 非翻译区与翻译调控 翻译起始时，起始因子对“帽子”的识别非常重要。研究表明，只有当此帽被甲基化形成m7G状态时，mRNA的翻译才更为有效。

② 3'非翻译区与翻译调控 真核mRNA的3'-UTR包括终止密码、poly (A) 尾以及前两者间的非编码序列，它们在翻译过程中同样具有重要的调控作用。mRNA 3'-UTR序列，包含着富含UA的保守序列。若除去这段序列，mRNA的稳定性明显提高。可见UA序列是抑制翻译作用的元件；其调控特点是随着它在3'-UTR中拷贝数的增加，对翻译的抑制效率也提高。

3'poly (A) 尾不仅对mRNA由核内向细胞质转运时具有保护作用，而且对mRNA的稳定性和翻译效率还具有调控作用。



## (3) 翻译起始因子的可逆磷酸化与翻译调控

翻译因子的磷酸化修饰直接关系到翻译作用的激活和抑制。**eIF-4E**是识别和结合mRNA 5'-m7G帽子结构的翻译起始因子。**eIF-4E在mRNA翻译中的重要调控作用是通过其亚基的可逆磷酸化来实现的。**eIF-4E的磷酸化作用可能有助于刺激eIF-4E的3个亚基形成复合物；或是促进eIF-4B、eIF-4A与eIF-3组装成更高级的复合物，因为这类复合物对翻译起始因子识别mRNA更为重要，可加快翻译的起始效率。

与eIF-4E磷酸化的激活效应不同，**eIF-2的磷酸化对翻译起始具有抑制作用。**eIF-2也是由3个相对分子质量不同的亚基组成。在有GTP存在时介导核糖体40S小亚基与甲硫氨基化的起始tRNA (Met-tRNA<sub>i</sub>) 特异结合，起始翻译。然而，eIF-2的磷酸化会抑制它的再利用，因为磷酸化的eIF-2对GDP和eIF-2B有很高的亲和力，因而阻抑了eIF-2参与前述复合物的形成而再循环。



## 14.6.2 翻译后水平的调节

### (1) 新生肽链的剪接

新生的多肽链经过一系列的加工修饰才能成为有活性的蛋白质。

加工包括：切除前蛋白的信号序列—信号肽；蛋白质原（或酶原）的加工；氨基酸的修饰；N端fMet或Met的切除以及二硫键的形成等

① **多肽链末端的切割和多蛋白肽链的水解加工** 多肽链末端的裂解加工常见于分泌型多肽，如胰岛素多肽链的翻译后剪接加工。胰岛素是由胰岛的 $\beta$ 细胞合成，最初在内质网膜上合成的肽链前胰岛素原（preproinsulin）。信号肽段在内质网中被信号肽酶切除，产生了胰岛素原（proinsulin）。胰岛素原进入高尔基体中，被PC3和PC2两种转变酶（convertase）从C片段两端切割，除去C片段。剩下的A和B片段通过二硫键结合，成为具有活性的胰岛素（insulin）（图14-21）。

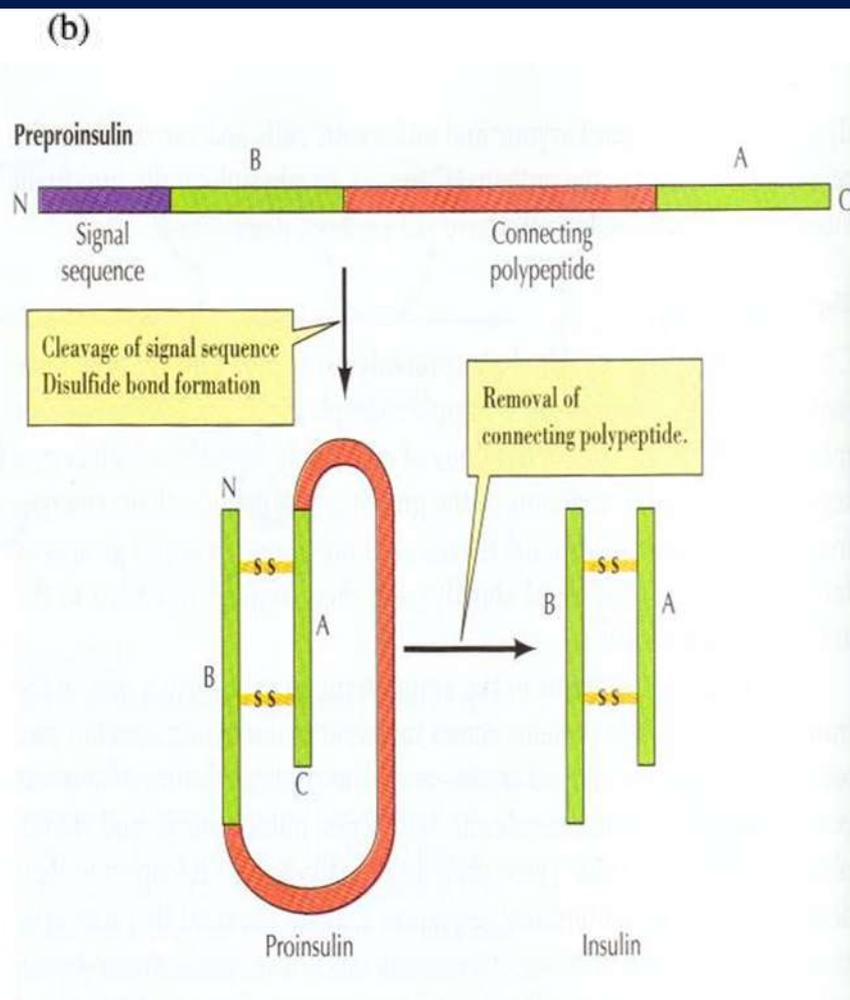
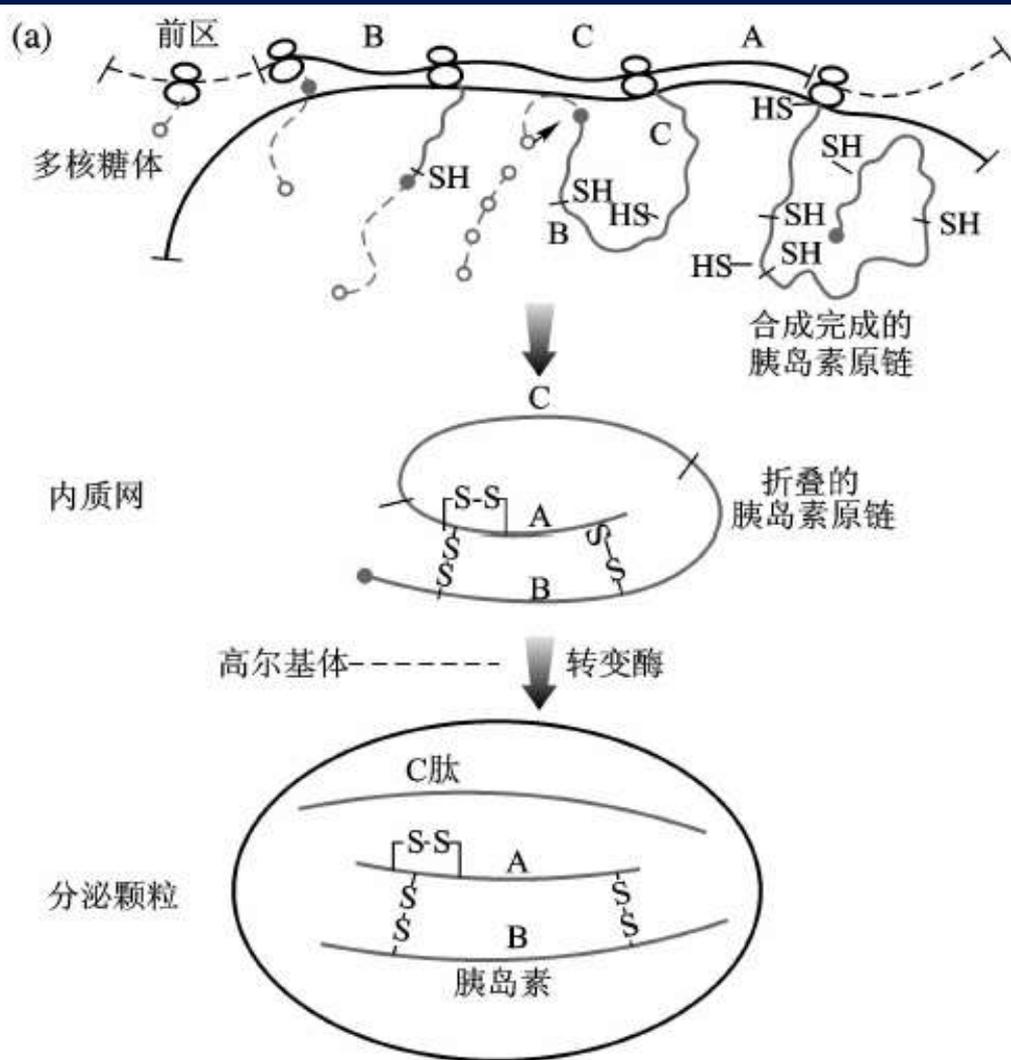
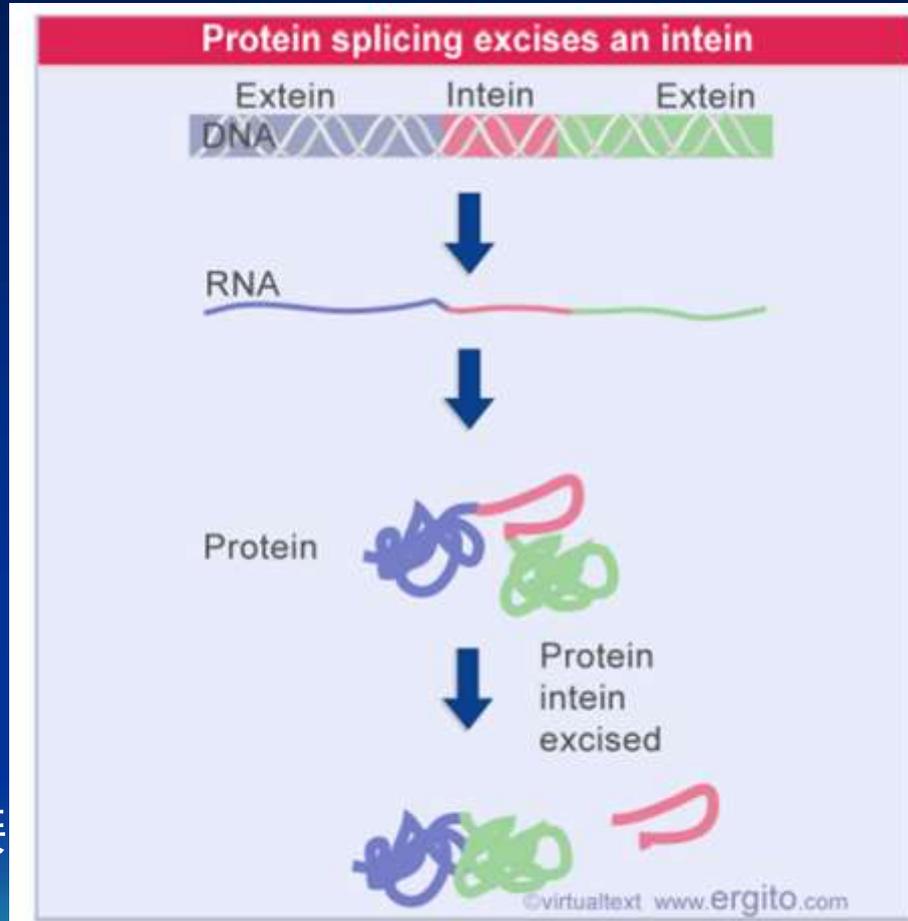


图 14-21 前胰岛素原加工成胰岛素示意图

## ② 内含肽剪接

蛋白内含子又称为**内含肽 (intein)** 是一种翻译后加工的产物。蛋白外显子又称为**外显肽 (extein)**。

内含肽的基因不是单独的开放阅读框 (ORF)，它是插入在外蛋白子的基因中，和内含子的区别在于它可以和外蛋白子的基因一道表达，而不是mRNA阶段被切除，它在产生前体蛋白以后再从前体中被切除掉，即**内含肽剪接 (intein splicing)**，余下的外蛋白子连接在一起成为成熟的蛋白，所以也被称为**蛋白质剪接 (protein splicing)**。



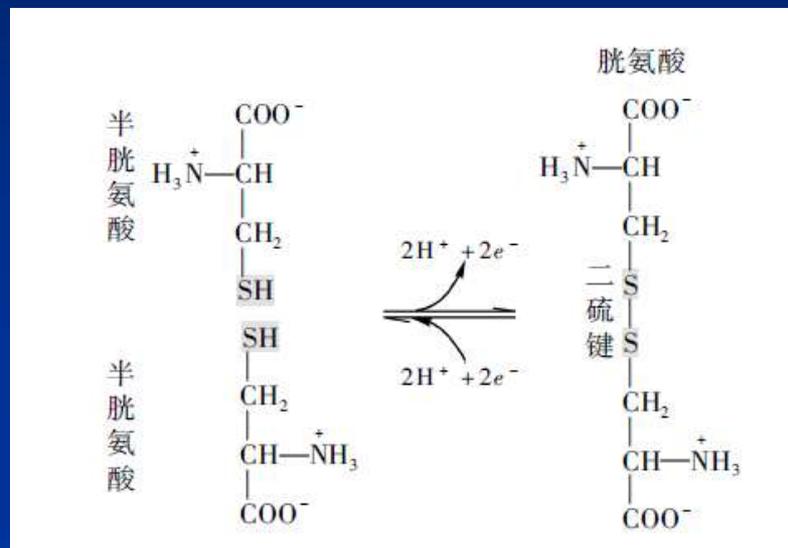
**Figure 26.24** In protein splicing the exteins are connected by removing the intein from the protein.



## (2) 新生肽链的化学修饰

白质的氨基酸组成一般有20种，但在成熟蛋白质中常存在非常规氨基酸，这些氨基酸是在翻译后经化学修饰形成的，从而增加了氨基酸的种类，以适应蛋白质功能的需要。例如：胱氨酸的形成

最简单的化学修饰类型是氨基酸的侧链或多肽链末端氨基酸的氨基或羧基加上小的化学基团，如乙酰基、甲基或磷酸基。修饰的方式有多种，如胶原中的脯氨酸和赖氨酸的羟基化。目前已发现



有150余种修饰氨基酸，每一种都有特定的修饰方式。这些修饰往往对蛋白质的精确活性起决定作用。如组蛋白乙酰化、甲基化对确定染色质的精细结构和表观遗传调控具有重要作用

### (3) 肽链的折叠

新生肽链经过上述几种方式修饰后，还必须进行正确的折叠，形成特定的三维构型和构象，才能成为有功能的蛋白质。无论是原核细胞还是真核细胞，都含有一类能使蛋白质肽链正确折叠的蛋白质，这类蛋白质称为**分子伴侣 (chaperone)**。在蛋白质折叠和组装过程中，分子伴侣能够防止多肽链的链内和链间错误折叠或聚集作用，并且还能破坏多肽链中已形成的错误结构，但自身不参加最终产物的组成。

### (4) 蛋白质更换

在真核生物中，一个细胞内各种蛋白质的数量，既取决于新生肽链的合成速率，又取决于它们存活的寿命，所以在一定时期内，细胞内的一些蛋白质会被降解，用以调节特定蛋白质的数量。除了可以通过溶酶体水解外，还具有蛋白酶解 (proteolysis) 的专门途径。这一专门的蛋白酶解活动与细胞的生理活动、细胞周期以及细胞分化等过程密切相关。

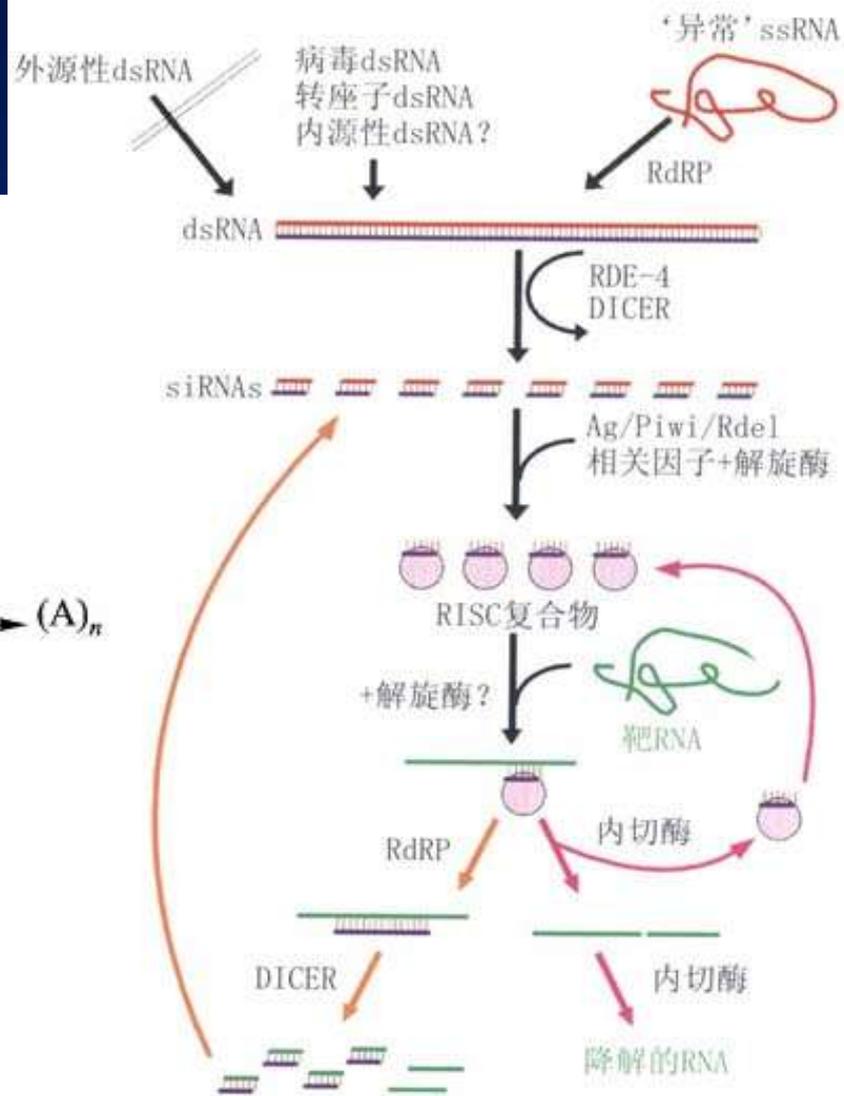


## 14.7 非编码RNA对基因表达的调控作用

### 14.7.1 RNA干扰

RNA干扰(RNA interference, RNAi)是正常生物体内一些小的双链RNA可以抑制特定基因表达的一种现象。当向细胞中导入与内源性mRNA编码区同源的小的双链RNA (double stranded RNA, dsRNA)时,可以通过促使该mRNA降解来高效、特异的阻断体内特定基因的表达,从而导致基因表达沉默的现象,因发生在转录后水平,又称为转录后基因沉默(post-transcriptional gene silencing, PTGS)。这些小的双链RNA称为siRNA (Small /short interfering RNA, siRNA)。

有关RNA干扰的分子机制已经研究得比较清楚(图14-22)。目前认为RNAi主要包括几个重要步骤:①首先, dsRNA被Dicer切割,产生siRNA;②RISC复合体(RNA-induced silencing complex)识别并结合siRNA,导致靶基因mRNA降解;③最后,在RNA聚合酶RdRP (RNA-dependent RNA polymerase, )作用下合成更多新的dsRNA,新合成的dsRNA再由Dicer切割产生大量的次级siRNA,从而使RNAi的作用进一步放大,最终将靶mRNA完全降解。



RNAi的作用模型

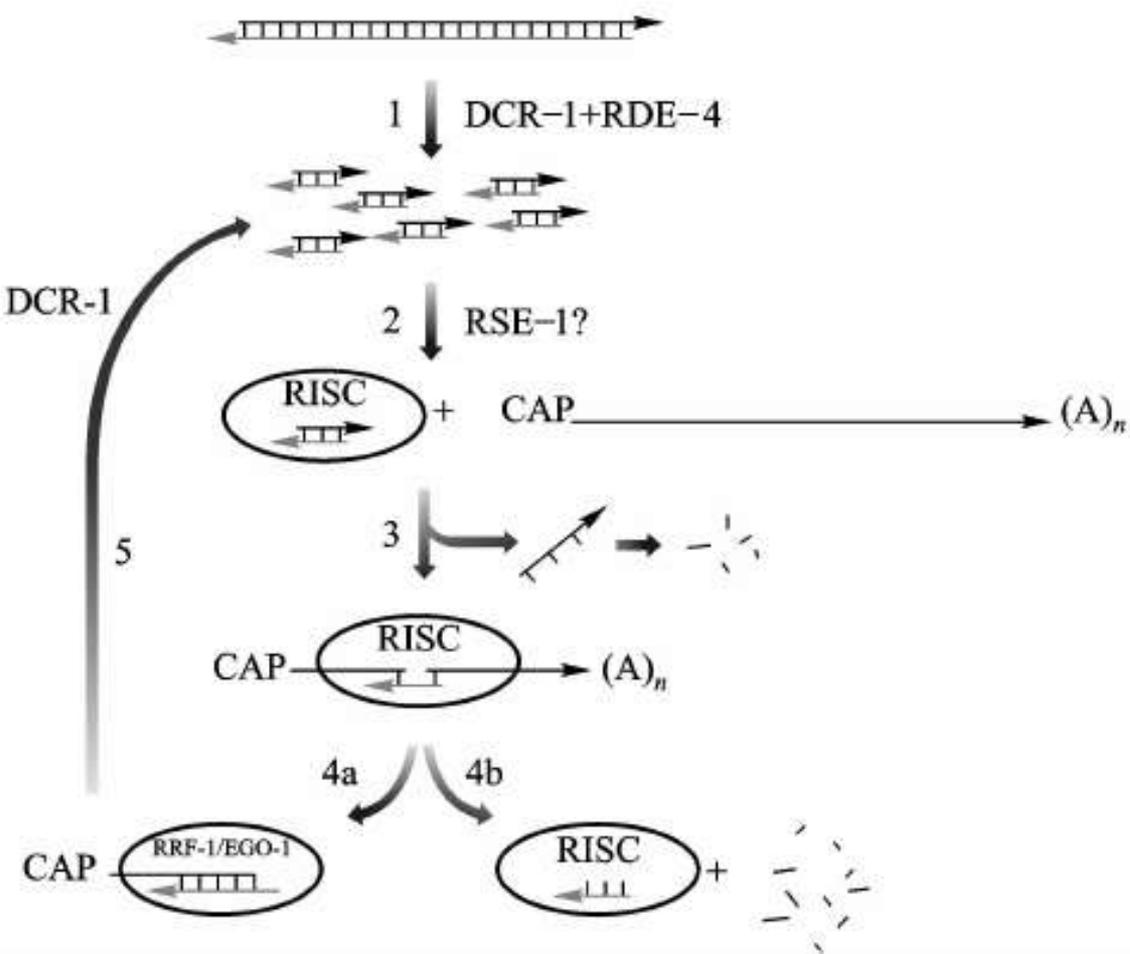
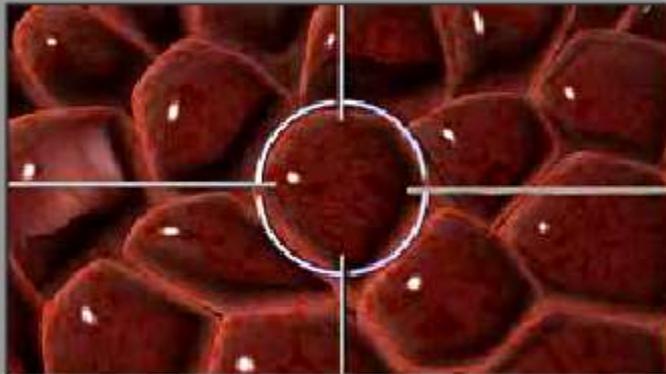


图 14-22 RNAi 的作用机制模型  
 (引自 Ketting 等, 2003)





RNA interference (RNAi) was first discovered in *C. elegans*. When long double-stranded (ds) RNAs were injected into a worm's gonad (a standard way of introducing transgenes into worms), they blocked the expression of endogenous genes in a sequence-specific manner.



Part 1: Gene expression



Glossary



Clear

Download

Copyright © 2003 Nature Publishing Group. Created by [Arkitek](#) for Nature Reviews Genetics

RNA干涉



## 14.7.2 miRNA对基因表达的调节

**miRNA (microRNA, 微RNA)** 是一类细胞内源表达的、非编码的、长度为19~23 nt的小RNA分子，它通过与靶基因mRNA中3'非翻译区的不完全匹配来识别并结合靶序列，诱导特异mRNA的降解或者翻译抑制，从而对靶基因的表达产生负调控作用。这属于跟RNAi类似的一种转录后基因沉默现象。*lin-4* (*lin*: abnormal cell LINEage) 和*let-7* (*let*: LETHal) 是最早发现的两个miRNA；它们都对线虫的发育在时间顺序上具有调节功能，被称为小时序RNA (*stRNA*, small temporal RNA)。

例如：*lin-4*和*lin-14*是调控线虫从一龄幼虫 (L1) 期向二龄幼虫 (L2) 期转换的两个关键基因；两者作用相反，并且*lin-14*基因对*lin-4*基因呈上位作用。*lin-14*基因编码一种核定位的蛋白，其表达呈现时间梯度：L1期基因产物较多，到了L2期蛋白含量急剧下降，但是其mRNA的表达水平则没有剧烈的变化。这种时间梯度变化受到其mRNA中3'-UTR序列的负调控。



1993年，V.Ambros等发现lin-4并不编码蛋白质，而是产生一对小RNA，一个长度约为22 nt，另一个为61 nt，前者表达量较高。这两个RNA产物与lin-14基因mRNA的3'-UTR的某些区域在序列上存在（不完全的）互补性，这种互补性导致lin-14基因mRNA的翻译受到抑制，从而造成LIN-14蛋白质水平降低，但并不影响其mRNA的表达（图14-23）

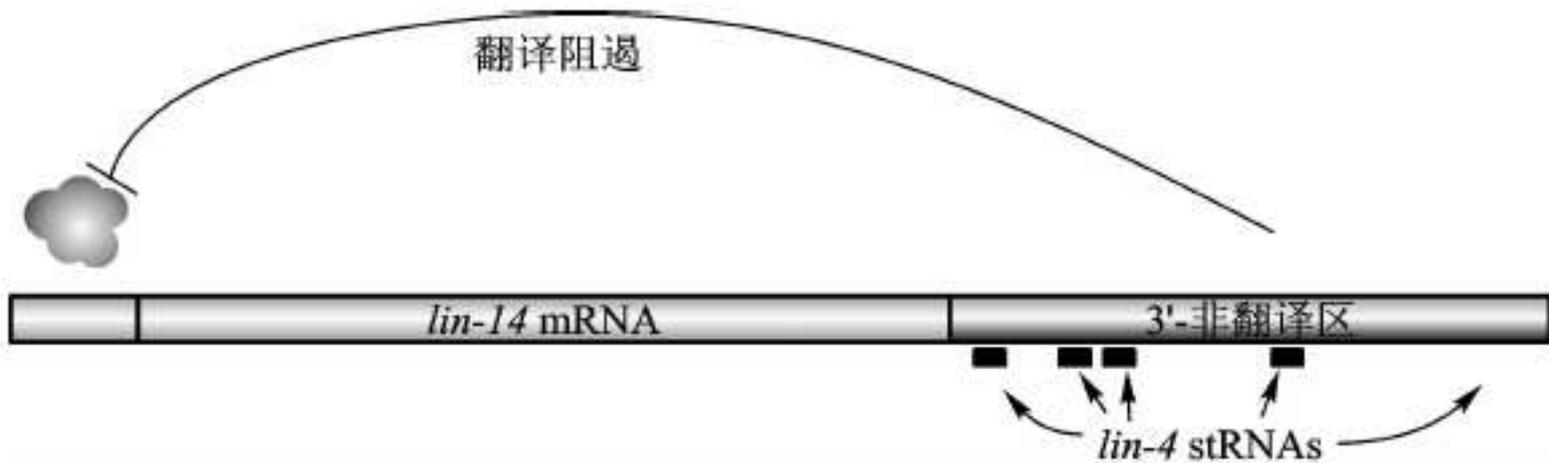


图 14 - 23 *lin - 4* miRNA 结合 *lin - 14* mRNA 3' - UTR 上的多个部位  
(仿自 Mamus 等,2003)

stRNA, small temporal RNA(小时序 RNA)

miRNA在植物、线虫、果蝇、斑马鱼和哺乳动物等多种物种中广泛存在，并且非常保守。迄今已鉴定与预测出几百种miRNA。每一个miRNA可以调控多个靶基因，每一个靶基因可同时受到多种miRNA的调控。

miRNA基因的表达包含了一系列复杂的步骤。

miRNA的功能呈现惊人的多样性，并且涵盖机体的发育过程、生理过程以及病理过程。

当前miRNA领域的研究热点主要集中在靶基因的寻找与鉴定以及miRNA的生物学功能与作用机制研究，以及LncRNA、miRNA、circRNA与mRNA的网络调控关系。



## 14.7.3 原核生物中小分子RNA在基因表达中的调控作用

### (1) 原核生物反义RNA在基因表达中的调控作用

**反义RNA (antisense RNA)** 是指能与所调控的RNA (或有意义的RNA) 互补配对, 抑制翻译的RNA序列。

大肠杆菌Tn10转座子中转座酶的合成调节是说明反义RNA调节基因表达的典型例证之一。Tn10两端各为IS 10R和IS 10L, 其中IS 10R有转座酶编码基因。该转座酶催化Tn10在大肠杆菌染色体上转座。**反义RNA来调节转座酶的表达**。启动子Pin和Pout转录的方向相反, 二者所转录的RNA有一段约36 bp的重叠区 [图14-24 (a)], 因此转录产物可以通过碱基配对阻止核糖体结合到由Pin启动转录的RNA上, 也就阻止了转座酶蛋白的合成 [图14-24 (b)]。通过这一机制, 细胞内的Tn10拷贝数越多, 反义RNA也会越多, 反过来却限制了转座酶基因的表达, 以致这个细胞株中的转座效率很低; 相反, 如果细胞中只有一个Tn10拷贝, 反义RNA水平就会很低, 转座酶便可有效地合成, 转座发生的频率就会大大增加 [图14-24 (c)]。



对原核生物的研究表明，反义RNA作用的基本原理是通过碱基配对与mRNA结合，形成二聚体，从而阻断后者的表达功能。

这种作用的可能途径：

第一是反义RNA与mRNA的SD序列或/和编码区互补结合，形成RNA-RNA二聚体，使mRNA不能与核糖体结合，而阻止了翻译过程；

第二是在复制水平上，反义RNA可与引物RNA互补结合，抑制DNA复制，从而控制着DNA（如质粒CoIE1）的复制频率；

第三是在转录水平上，反义RNA还可以与mRNA 5'端互补结合，阻止mRNA完整转录。

反义RNA的研究具有重要意义，由于反义RNA能高度特异性地与mRNA结合，抑制特定基因的表达，因此在基础研究中为基因分析提供了更好的手段，不需改变基因结构就可以分析特定基因在细胞内的功能



## (2) 细菌中的RNA调节物

近几年发现，细菌中有些小RNA也具有调节基因表达的作用。其作用机制分两个方面：一是与目标核酸序列形成双螺旋区段，直接阻止蛋白质结合到目标RNA单链上，阻断翻译的起始，导致转录终止；或是由于形成双链区而为内切核酸酶创造一个靶位点。二是在目标分子某一部分形成双链区，使得目标分子的另一部分发生构象变化，间接影响该区段的功能。

这些小RNA它们共同的特性都是通过目标序列在二级结构方面的变化来控制其活性。RNA调节物与阻遏操纵子的蛋白质不同，它没有变构的属性，不能用改变其识别目标的方法对其他小分子产生影响，它是用控制其基因的转录作用来发挥作用，或是由于降解RNA调节物产量酶的影响而停止活动。



目前在大肠杆菌中已鉴定出约50种小分子RNA，其中有一些是作用于许多目标基因的通用调节物。在通用控制系统中氧化应激（oxidative stress）是一个很好的例子，小分子RNA是这个通用控制系统中的一个调节物，当细菌处在活性氧的环境中，就会采用诱导抗氧化剂防御基因来应答。过氧化氢激活转录作用激活因子OxyR，以此控制几种可诱导基因的表达。这些基因中的*oxyS* 编码一个小分子RNA。

*oxyS* RNA是一条只含有109个核苷酸的短小分子，不编码蛋白质，它是一个反式作用调节物，在转录后水平上影响基因表达。它有10个以上的目标基因，在有些基因上它激活表达，在另一些基因上它阻遏表达，*oxyS*是通过与目标mRNA碱基配对机制来行使其阻遏表达的功能。在*oxyS* RNA的二级结构中有3个突出的茎-环结构，当与一个目标分子*flhA* mRNA相互作用时，它靠近3'端的茎-环正好与*flhA* mRNA起始密码子前方的序列碱基配对，使目标分子翻译受阻（图14-25）



图 14-25 *oxyS* RNA 抑制 *flhA* mRNA 翻译示意图  
(引自 Krebs 等, 2014)

所有上述几种小RNA都利用RNA分子伴侣Hfg来增强其效应。Hfg是一种多效蛋白，能与大肠杆菌的许多小RNA分子结合。例如，它与*oxyS* RNA结合，通过增强后者与其目标mRNA结合的能力来提高*oxyS*的效应。关于大肠杆菌中小RNA调节物的研究，近年已取得不少成果，为在其他生物中探询这类重要的RNA调节物的作用提供了范例。



武汉大学

Wuhan University

谢谢！

本章结束

