



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. RNA介导的原核生物基因表达调控
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



单细胞原核生物：以操纵子（operon）为单位的调控系统是其主要的调控方式。另有许多翻译水平上的调控，以及RNA介导的原核生物基因表达调控。

真核生物：基因转录在细胞核中，而翻译在细胞质中进行，因而处在一个非常纷繁复杂的控制系统中。各类细胞中都具有特异基因的表达和相应的调控机制。

近年来，表观遗传学的研究表明，通过影响染色质的结构与活性调控基因表达是真核生物基因表达调控的一种重要方式；而非编码RNA对基因的表达调控作用则揭示了基因表达调控的新机制，值得我们认真探究



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)

乳糖操纵子：结构基因*lacZ*、*lacY* 和*lacA*连接在一起，分别编码 β -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。



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）。



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

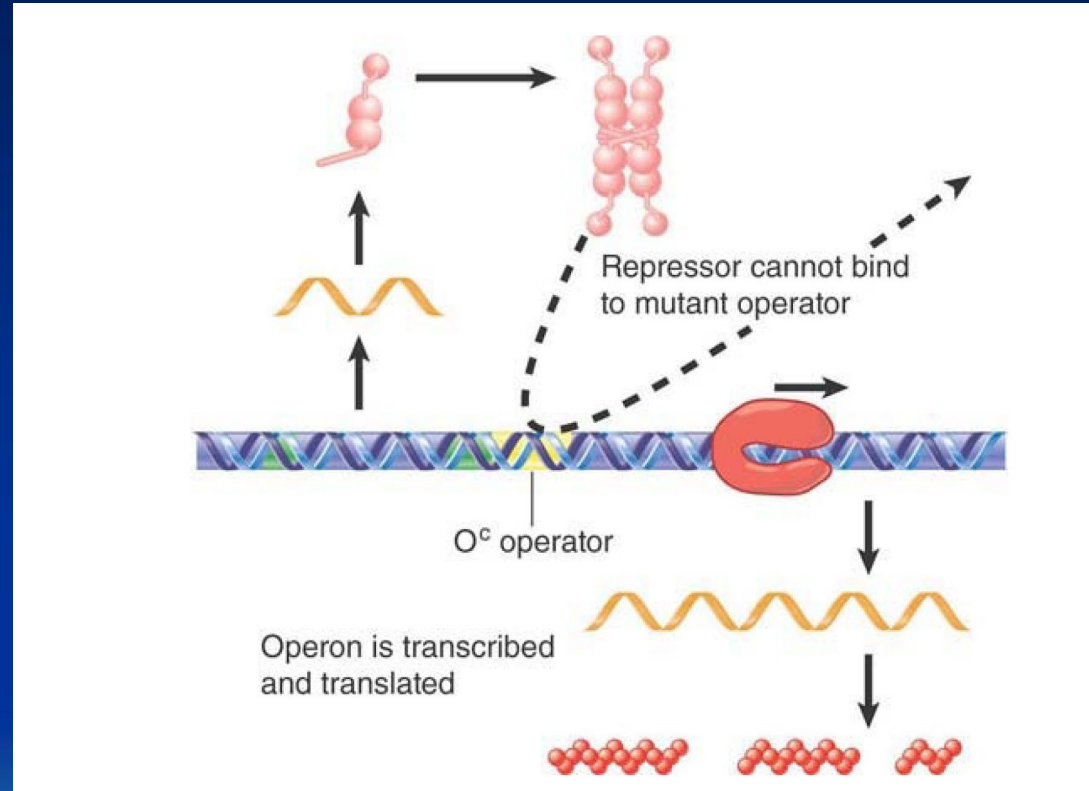


图14-2 操纵基因突变导致组成型表达
(引自Krebs, 2018)

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?

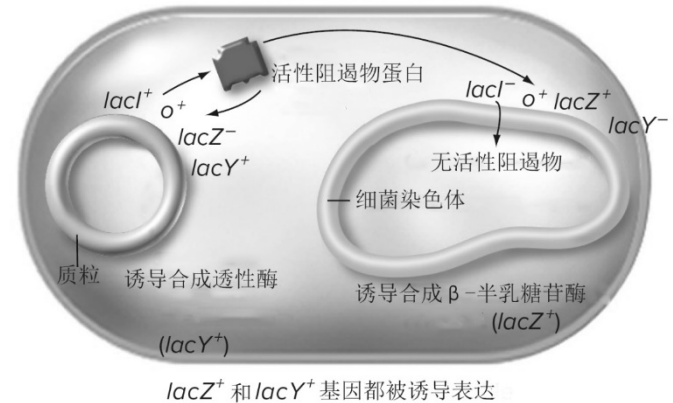
Monod等利用F' lac质粒构建了多种组合的部分二倍体，以研究说明二者的关系 (图14-3)：

(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

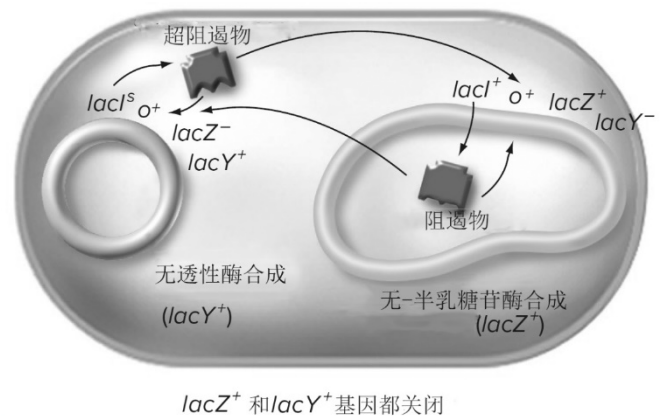
(b) 维持阻遏状态：LacI^S蛋白也是一种反式作用因子。由 $lacI^S$ 编码的超阻遏蛋白即使在有诱导物存在时也依然能够结合在质粒和细菌染色体的操纵基因位点 O^+ 上

(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

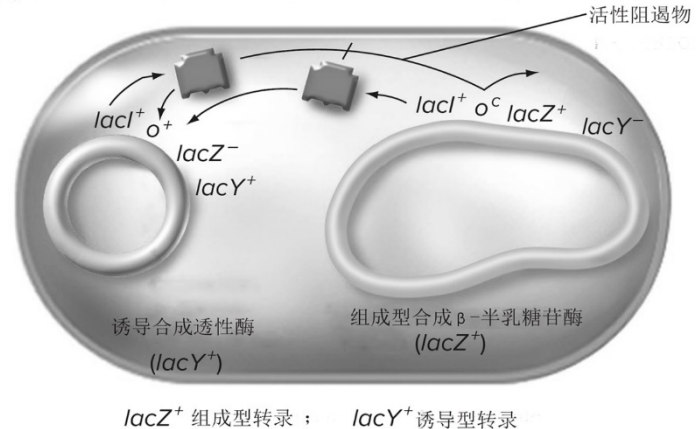
(a) 引入 $lacI^- o^+ lacZ^+ lacY^-$ -细菌中的F' ($lacI^+ o^+ lacZ^- lacY^+$) 质粒



(b) 引入 $lacI^S o^+ lacZ^+ lacY^-$ -细菌中的F' ($lacI^S o^+ lacZ^- lacY^+$) 质粒



(c) 引入 $lacI^+ o^c lacZ^+ lacY^-$ -细菌中的F' ($lacI^+ o^+ lacZ^- lacY^+$) 质粒





(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，而且细胞内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)]

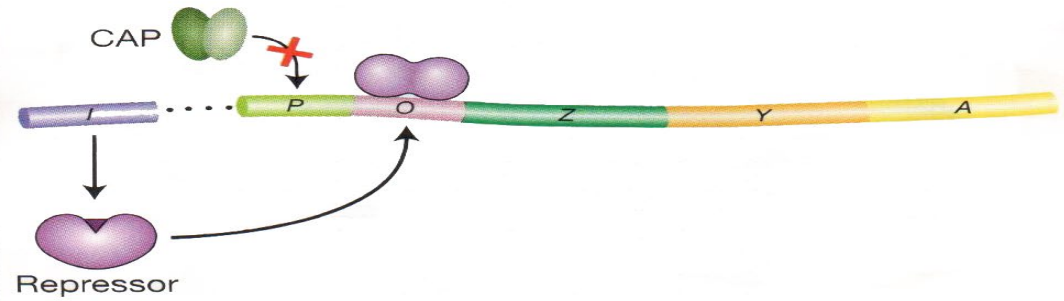
乳糖操纵子突变体的分析结果证明cAMP与分解物阻遏现象确实存在一定关系。

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)]。

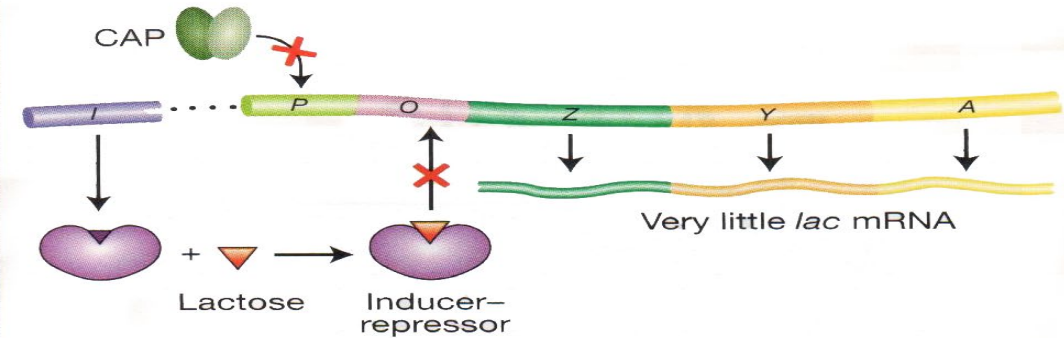
研究表明，细菌只有在只有乳糖而没有葡萄糖时，才会产生乳糖代谢的一系列酶，而这种有效的调控是通过CAP-cAMP构成的正调控系统和诱导物-阻遏物构成的负调控系统这两条途径的相互配合实现的（图14-5）。

Negative and positive regulation of the *lac* operon

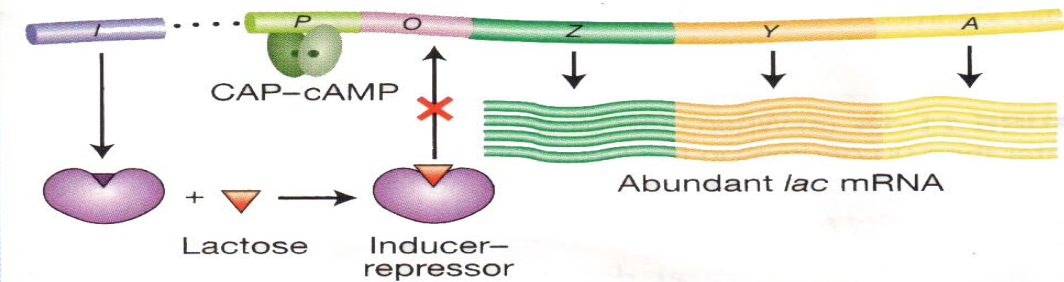
(a) Glucose present (cAMP low); no lactose; no *lac* mRNA



(b) Glucose present (cAMP low); lactose present



(c) No glucose present (cAMP high); lactose present



(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

半乳糖也是 *E. coli* 的一种碳源。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。

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 操纵子与其调节基因 *galR* 相距甚远 (图16-6a)。

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⁻* 和 *galO^c* 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：（1）*gal* 操纵子有两个相互重叠的启动子：*galP1* 和 *galP2*。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 14-6b)。（2）*gal* 操纵子有两个 *O* 区，一个为 O_E ，在 *P* 区上游 -67 ~ -73，另一个 O_I 在结构基因 *galE* 内部（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 *galP1* 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 *galP2* 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:**

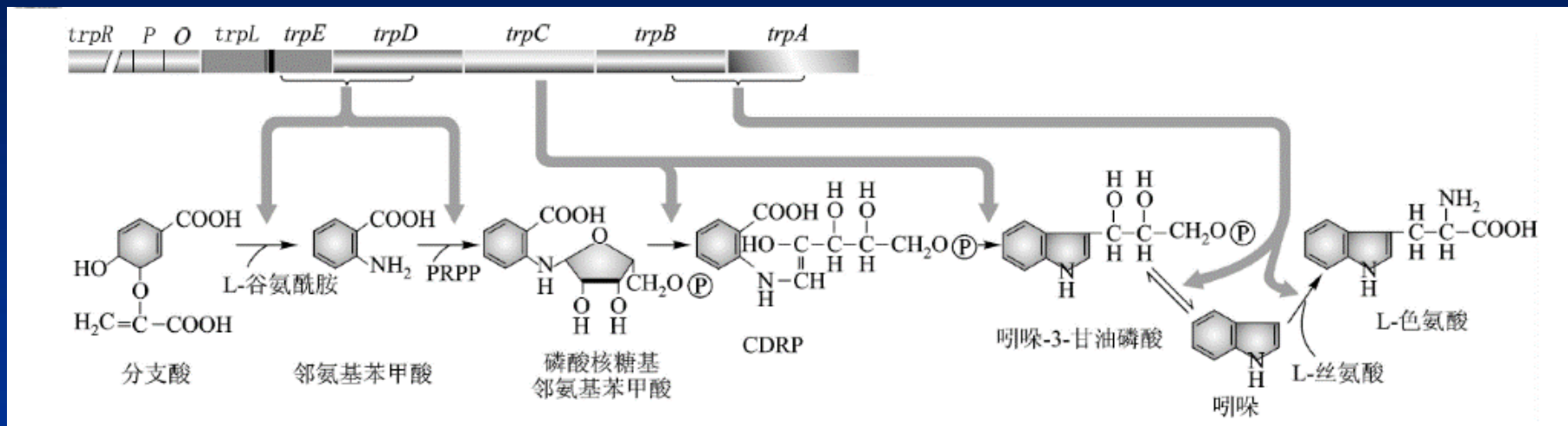
(1) 在有阿拉伯糖时, *AraC* 蛋白与诱导物阿拉伯糖结合, 被激活成为一种诱导性蛋白, 跟 CAP-cAMP 复合物共同结合到 *araI* 区域, 促使 RNA 聚合酶结合于启动子上, 促进 *ara* 操纵子的转录, 发挥正调控作用 [[图14-7 \(b\)](#)];

(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 (*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。

与*lac*操纵子的结构不同，在*trp*操纵子的第一个结构基因E与操纵基因O之间有一段由*trpL*基因编码的前导序列 (leader sequence, L) 和衰减子区 (黑方块)。***trpR*和*trpL*这两个基因在*trp*操纵子的调控中起重要作用**。*trpL*基因是*trp*操纵子的一部分，而编码阻遏蛋白的*trpR*基因远离*trp*操纵子，具有自己的启动子，不是*trp*操纵子的一部分。



(2) *trp*操纵子的阻遏调控

trp 操纵子是如何受编码 trp 阻遏蛋白的 *trpR* 基因调控的？当细胞内的色氨酸水平非常低时，色氨酸不与 trp 阻遏蛋白结合，无活性的阻遏蛋白不能与操作基因位点结合。因此，RNA聚合酶转录 *trp* 操纵子，细胞表达色氨酸合成所需的基因（[图14-9a](#)）。

当细胞内的色氨酸水平变高时，色氨酸会充当与 trp 阻遏物结合的辅阻遏物。这会导致 trp 阻遏物的构象变化，色氨酸- trp 阻遏复合物然后能够结合到 *trp* 操纵基因位点（[图14-9b](#)）。这抑制了RNA聚合酶转录操纵子的能力。因此，当细胞内存在大量色氨酸时——当细胞不需要产生更多色氨酸时——*trp* 操纵子将关闭。



(3) 色氨酸操纵子的衰减作用

1970年，Yanofsky等人发现缺少 trp 阻遏物的突变株。

经研究，Yanofsky在 trp 操纵子中发现了第二种调控机制——衰减（attenuation），它是由包括 $trpL$ 基因的区域介导的（[图14-9c](#)）。

将 trp 操纵子的mRNA分离出来进行序列分析，发现在 $trpE$ 基因5'端起始密码子与 $trpP$ 之间有一段长达160个碱基的序列——前导序列。在能够大量产生 trp mRNA的突变体中，这一前导序列内缺失了含有 trp 密码子的一段碱基序列。当培养基中含有色氨酸时，由于前导序列中这一区段的存在，RNA聚合酶不再前进，mRNA分子的合成便终止于这一区域，说明这是一个直接参与色氨酸操纵子调控的区段，称为衰减子或弱化子（attenuator）（[图14-10a](#)）。



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

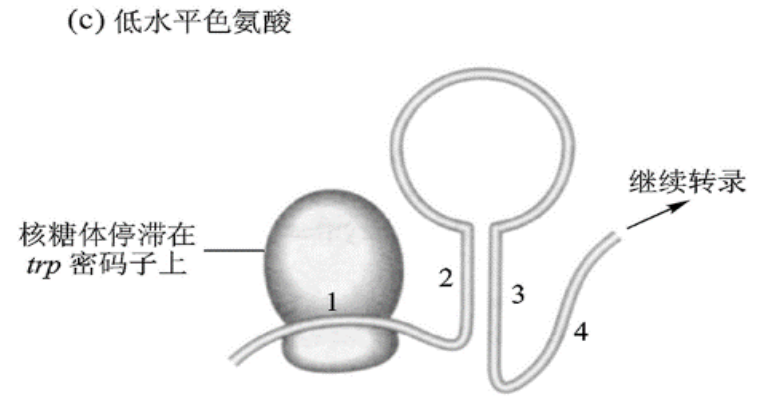
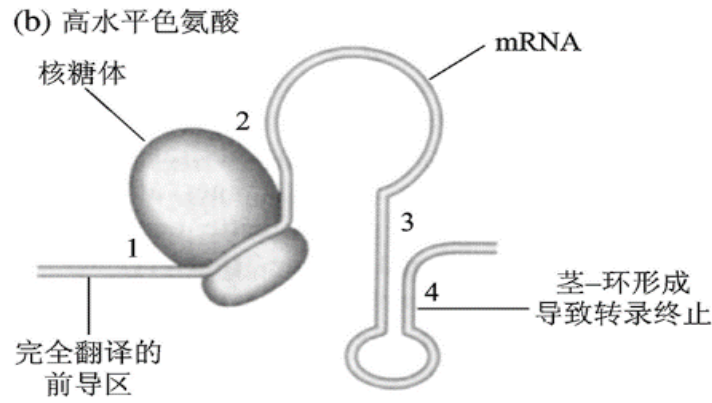
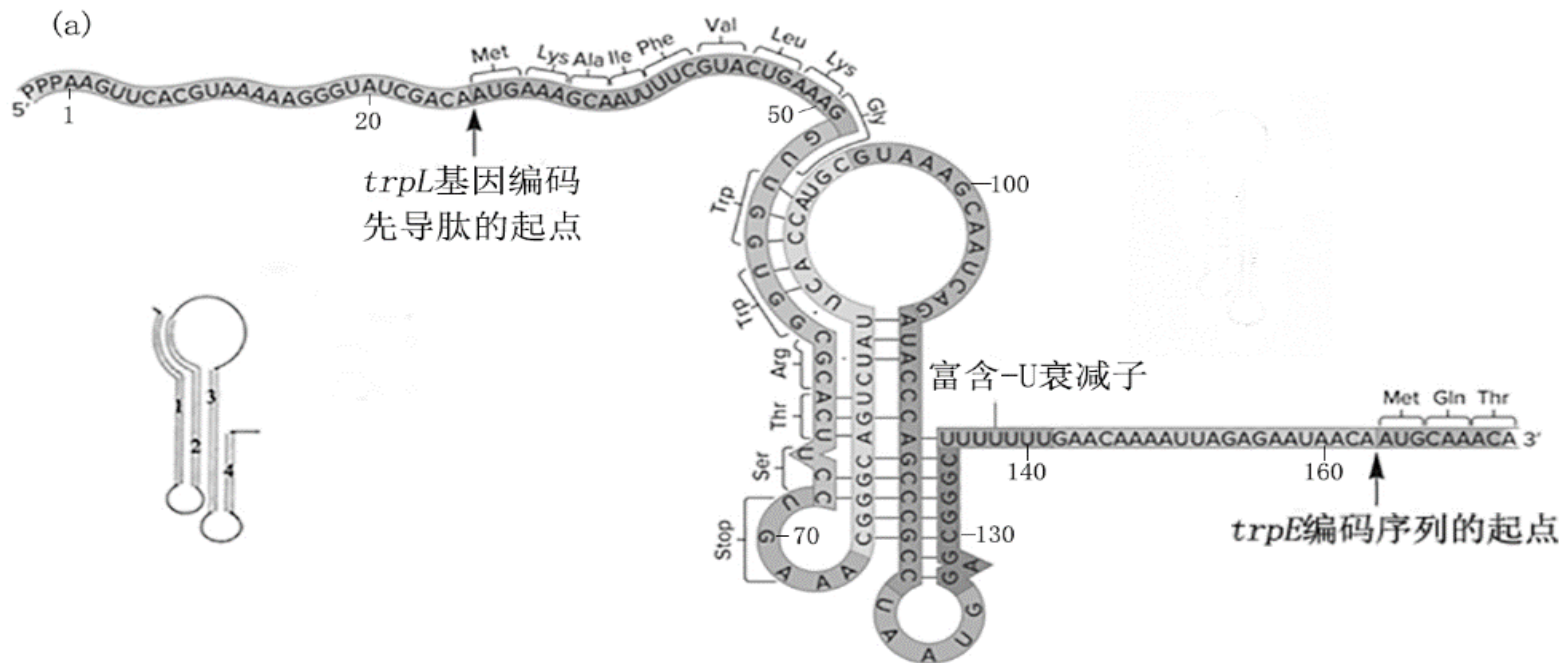


图14-10 *trp*操纵子衰减作用机制的模型 (改自Brooker, 2018和Griffiths等, 2020)





(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 RNA介导的原核生物基因表达调控

而近些年来的一些重要发现是，许多在细菌细胞中控制基因表达的机制依赖于表现出形状变构的RNA分子。

已知在转录中，所有细菌的mRNA均始于称为5' UTR或RNA前导序列的非翻译区。通过互补的碱基配对，许多RNA前导序列形成茎环（或发夹环）二级结构。这些茎环结构可以提前终止mRNA的其余部分的转录，或者可以通过阻止mRNA进入核糖体结合位点来阻止翻译。这些RNA前导分子是变构的，因为它们可以响应各种环境胁迫而改变其茎环结构，从而改变其功能。



(1) 衰减子

以上，我们介绍的大肠杆菌色氨酸操纵子中的衰减子是RNA前导装置顺式调节基因表达的典型例子之一。

(2) 核糖开关

(3) 调控小RNA反式调控mRNA的翻译

(4) 反义RNAs的调控





(2) 核糖开关

核糖开关 (Riboswitches) 是指mRNA一些非编码区的序列折叠成一定的构象, 这些构象的改变应答于体内的一些代谢分子, 从而通过这些构象的改变达到调节基因表达的目的。

典型的核糖开关由**适体** (aptamer) 和**表达平台** (expression platform) 两部分组成 (图14-11)

许多细菌基因和操纵子mRNA的前导序列是一个核糖开关。在一些核糖开关中, **表达平台控制转录的终止** (图14-11a); 在其他核糖开关中, **表达平台通过阻断或解除核糖体结合位点来调节翻译** (图14-11b)。在大肠杆菌等原核生物基因组中已经发现了20多种核糖开关适体, 它们都能广泛结合嘌呤、氨基酸、离子、辅酶及其衍生物等。

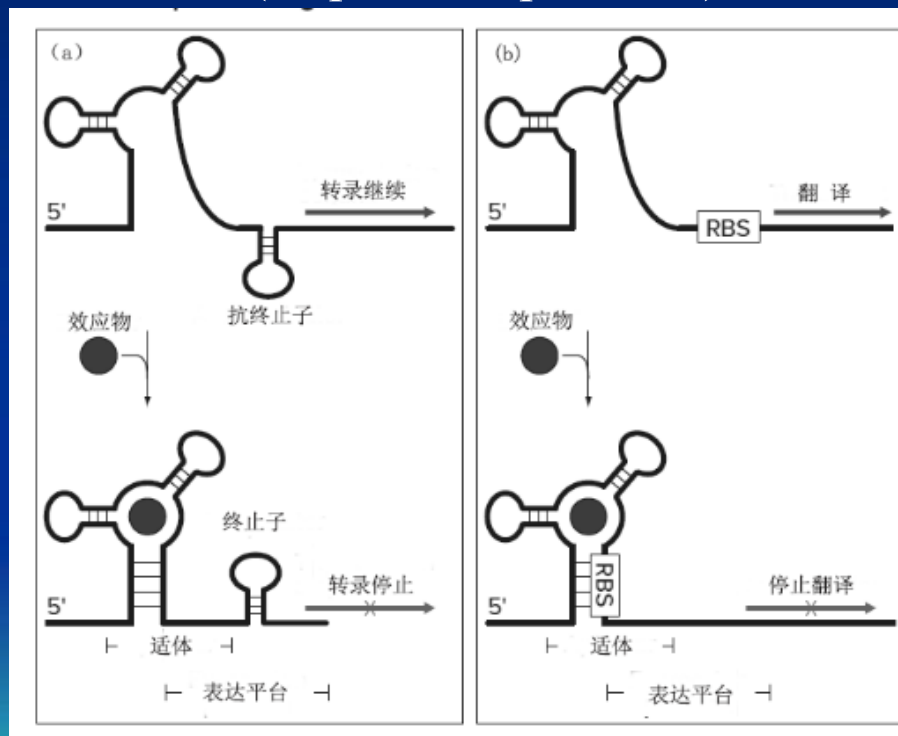


图14-11 核糖开关

(引自Hartwell等, 2018)

(3) 调控小RNA反式调控mRNA的翻译

细菌基因组编码许多小RNA分子，或sRNAs，它们通过与mRNAs的碱基配对来反式调节蛋白质翻译(图14-12)。调节sRNAs长50-400nt，含一个与几个不同mRNA靶点互补的区域。大多数sRNA是抑制性的，它们通过与核糖体结合位点的碱基配对来抑制其靶mRNA的翻译(图14-12a)。

然而，还有一些sRNAs通过破坏mRNA前导序列中的茎环结构的形成来激活其靶mRNAs的翻译，否则将阻断核糖体结合位点(图14-12b)。

另一种方式，某些sRNAs可以影响特定靶基因的表达，是通过促进mRNA的降解：双链sRNA与mRNA结合而产生的RNA区域导致mRNA被核糖核酸酶降解，即RNA干扰。

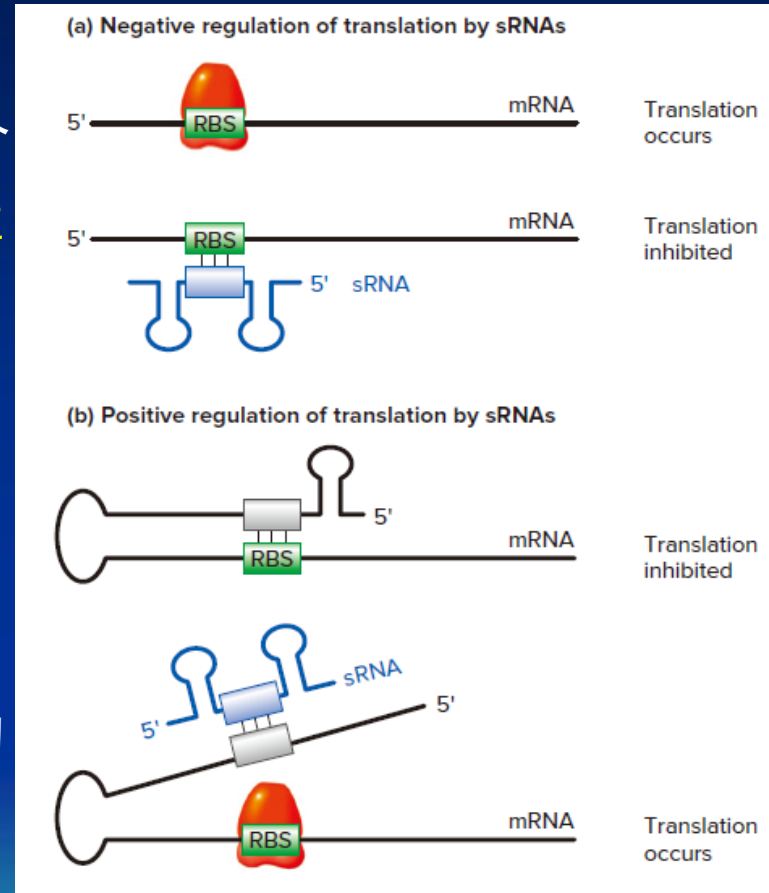


图14-12 sRNAs 的反式作用对翻译的调节



(4) 反义RNAs的调控

某些细菌基因受与mRNA序列互补的RNA调控，因为它们的转录模板是DNA的相反链。这些调节性RNA称为反义RNA（antisense RNAs）。它们调节的mRNA是有义RNA（sense RNAs）（图14-13）。反义RNA的大小范围为10-1000 nt，可能与相反的DNA链上编码的整个mRNA互补，或者它们可能仅重叠一部分。注意，用于产生有义mRNA和反义调节RNA的启动子位于编码区的相对侧（图14-13）。

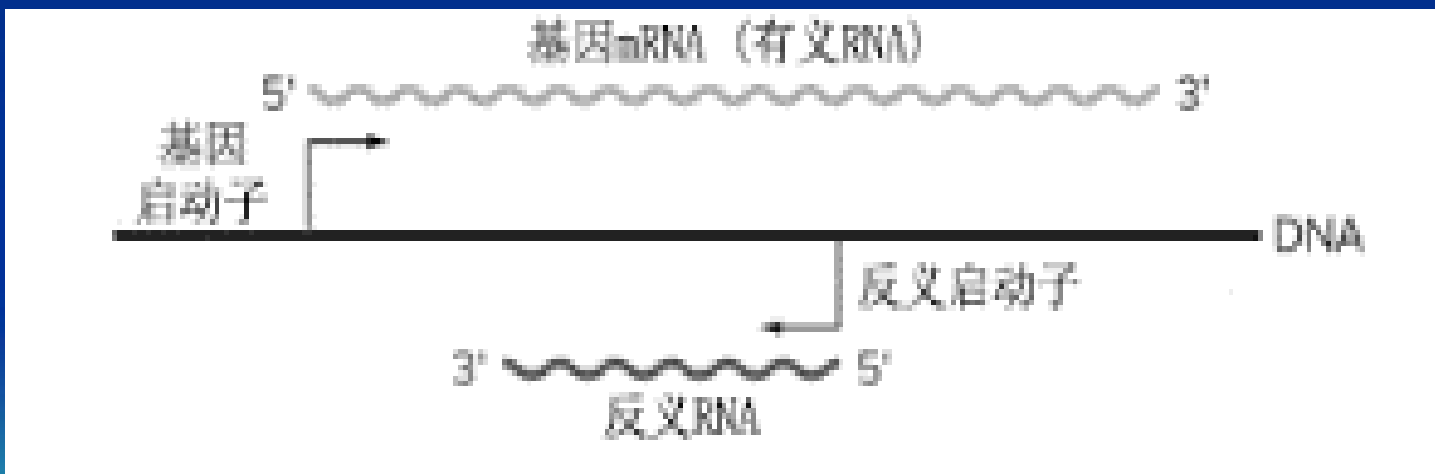


图14-13 反义RNAs抑制基因表达



14.4 真核生物基因转录水平的调节

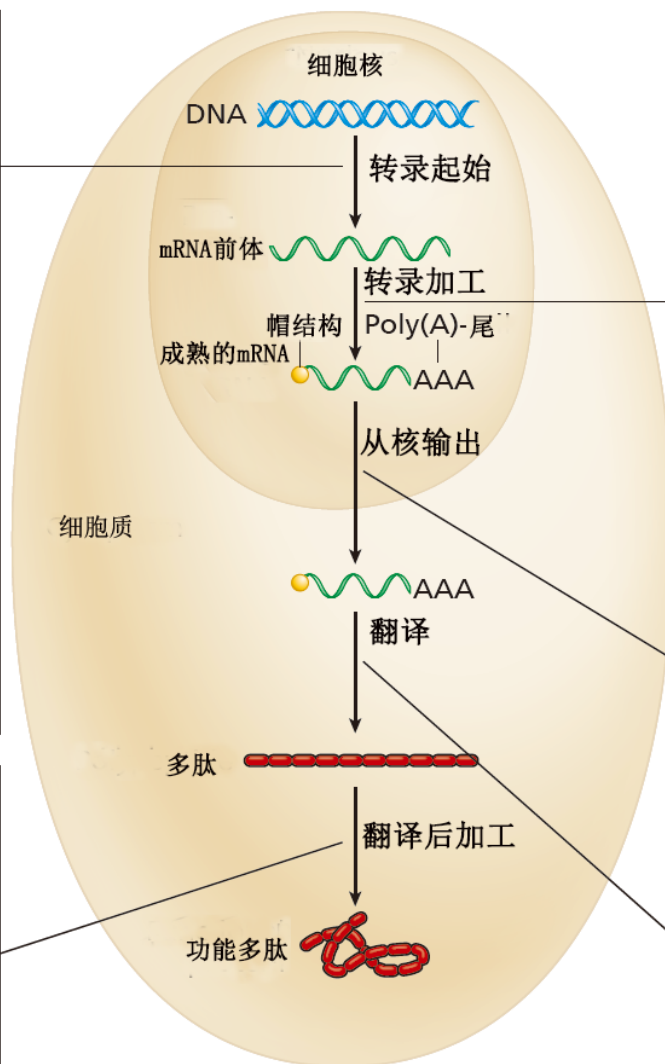
14.4.1 真核生物基因表达调控概述

(1) 转录调控

- a. 调节蛋白和转录因子与共有 DNA 序列（启动子区域）结合以促进转录。
- b. 其他调节性 DNA 序列（增强子和沉默子）结合调节蛋白，以促进每种细胞类型中特定基因的转录。
- c. 通过蛋白质作用形成有利于转录的开放染色质结构。
- d. 在不同的细胞类型中利用替代的启动子来产生不同的前体 mRNA 分子。
- e. DNA 的甲基化抑制转录。

(5) 翻译后调节

- a. 在转运出细胞之前，多肽先在高尔基体中加工和修饰。
- b. 调节分子与多肽结合以改变其功能。
- c. 蛋白质的稳定性受到调节。



(2) mRNA 加工

- a. 5' 末端的加帽，3' 末端的聚腺苷酸化和内含子剪接修饰了前体 mRNA。
- b. 选择加帽和聚腺苷酸化位点可被用于不同的细胞类型。
- c. 选择性剪接可从某些细胞类型中产生不同的成熟 mRNA 分子。
- d. RNA 编辑可修饰 mRNA 的碱基序列。

(3) 成熟 mRNA 的调节

- a. 翻译调节蛋白与成熟的 mRNA 结合以延迟翻译起始。
- b. 小 RNA 调节 mRNA 的稳定性或 mRNA 的翻译。
- c. 成熟的 mRNA 运输到细胞质受到调节。
- d. RNA 稳定性受到调节。

(4) 翻译

掩蔽 mRNA 会延迟或阻止翻译。



14.4.2 真核生物转录水平的调控

(1) 真核生物结构基因的一般结构组成

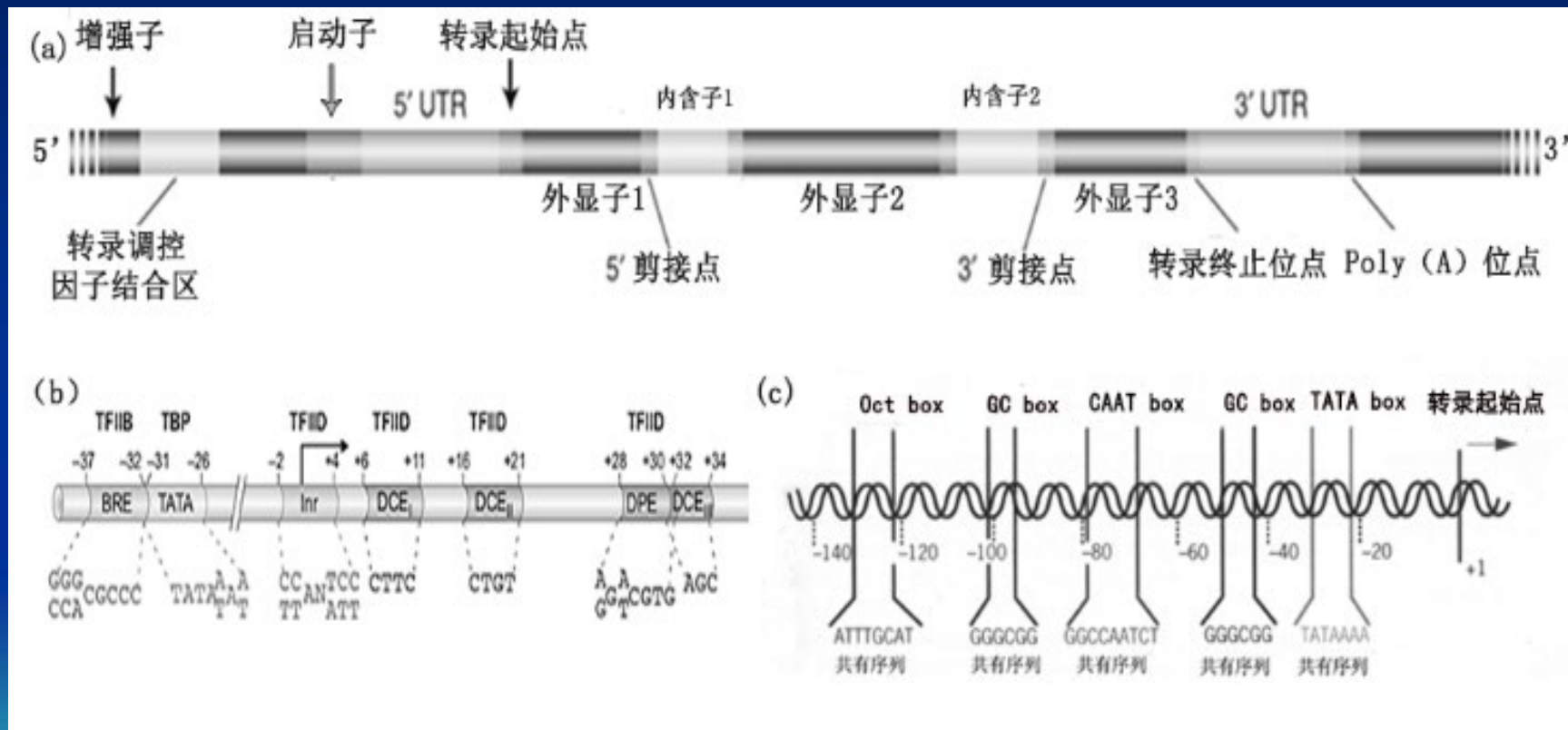


图14-15 一个真核蛋白质编码基因的一般结构模式

a. 一个蛋白质编码基因的一般结构组成 b. Pol II 核心启动子 c. 启动子的基本元件





(1) 顺式作用元件

顺式作用元件 (cis-acting element)，又称顺式调节元件，是指DNA序列上一些对基因表达有调节活性的特定调控序列 (regulatory sequence)。这种序列上分布着调节蛋白的结合位点 (regulator site)。真核生物的顺式作用元件主要包括启动子、增强子、绝缘子 (insulator) 和沉默子 (silencer) 等。

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)。

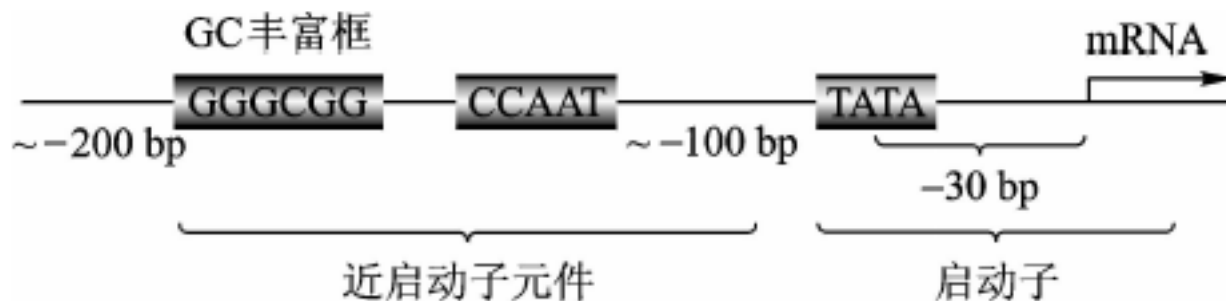


图 14 - 12 真核生物转录起始位点的上游序列

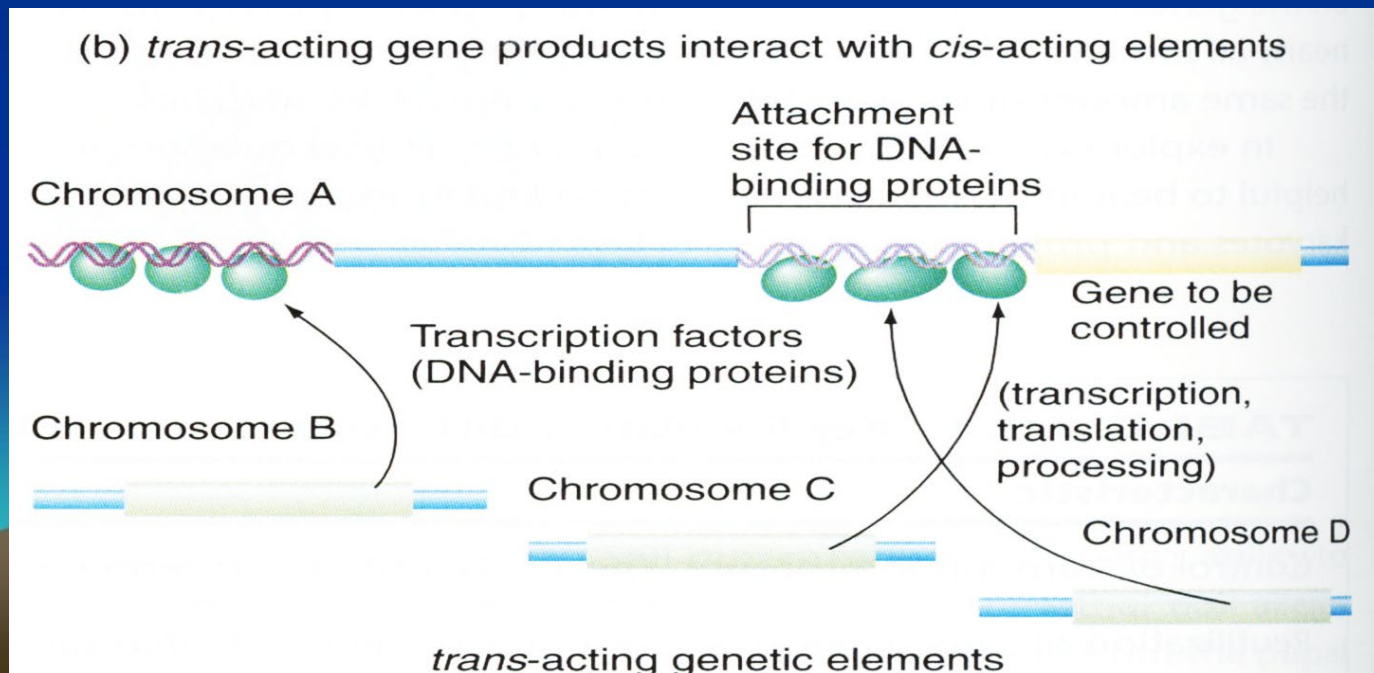
(引自 Griffiths 等, 2015)



(2) 反式作用因子

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

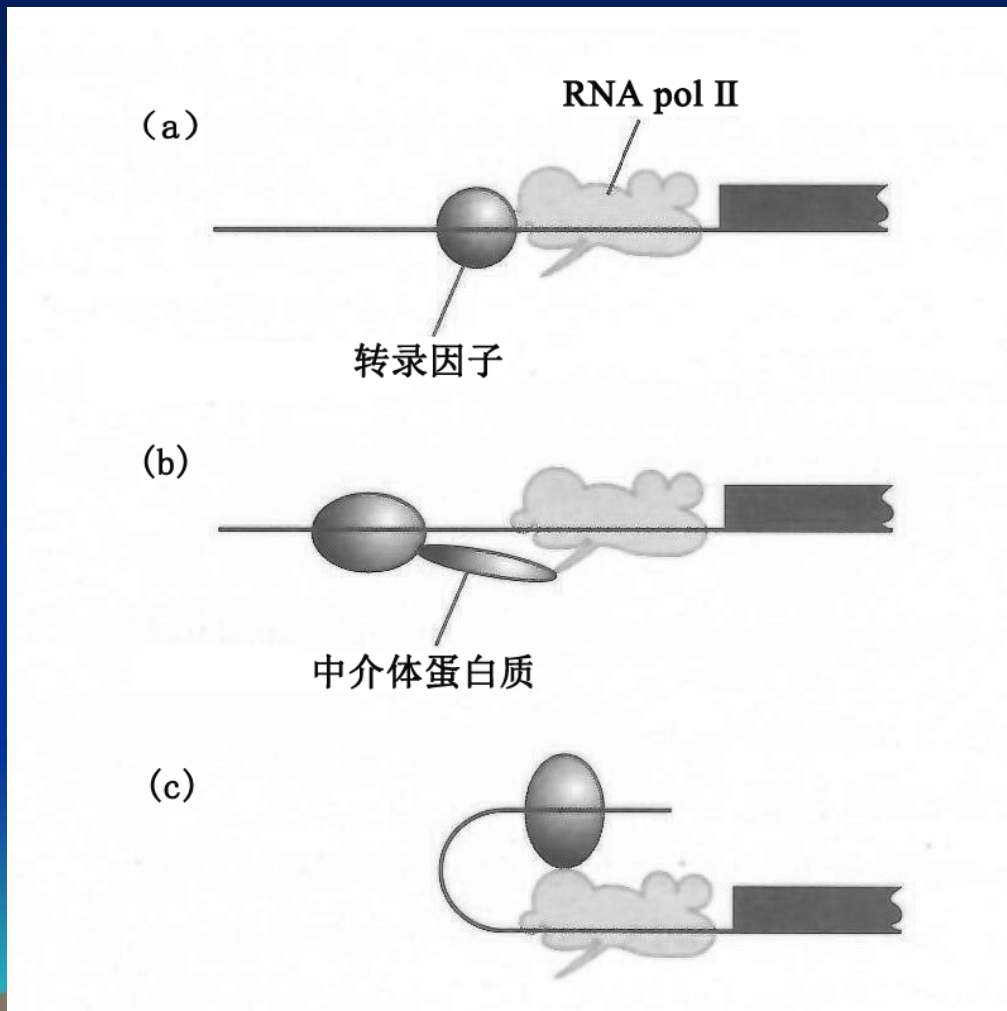
反式作用因子 (trans-acting factor)：一些直接或间接地识别或结合各种顺式作用元件的8~12bp核心序列，并参与调控靶基因转录效率的一组蛋白（少量是RNA）统称为调节基因转录的反式作用因子 (trans-acting factor)。反式作用因子大多数是DNA结合蛋白，参与转录调控，所以也称为转录因子 (transcription factor)。

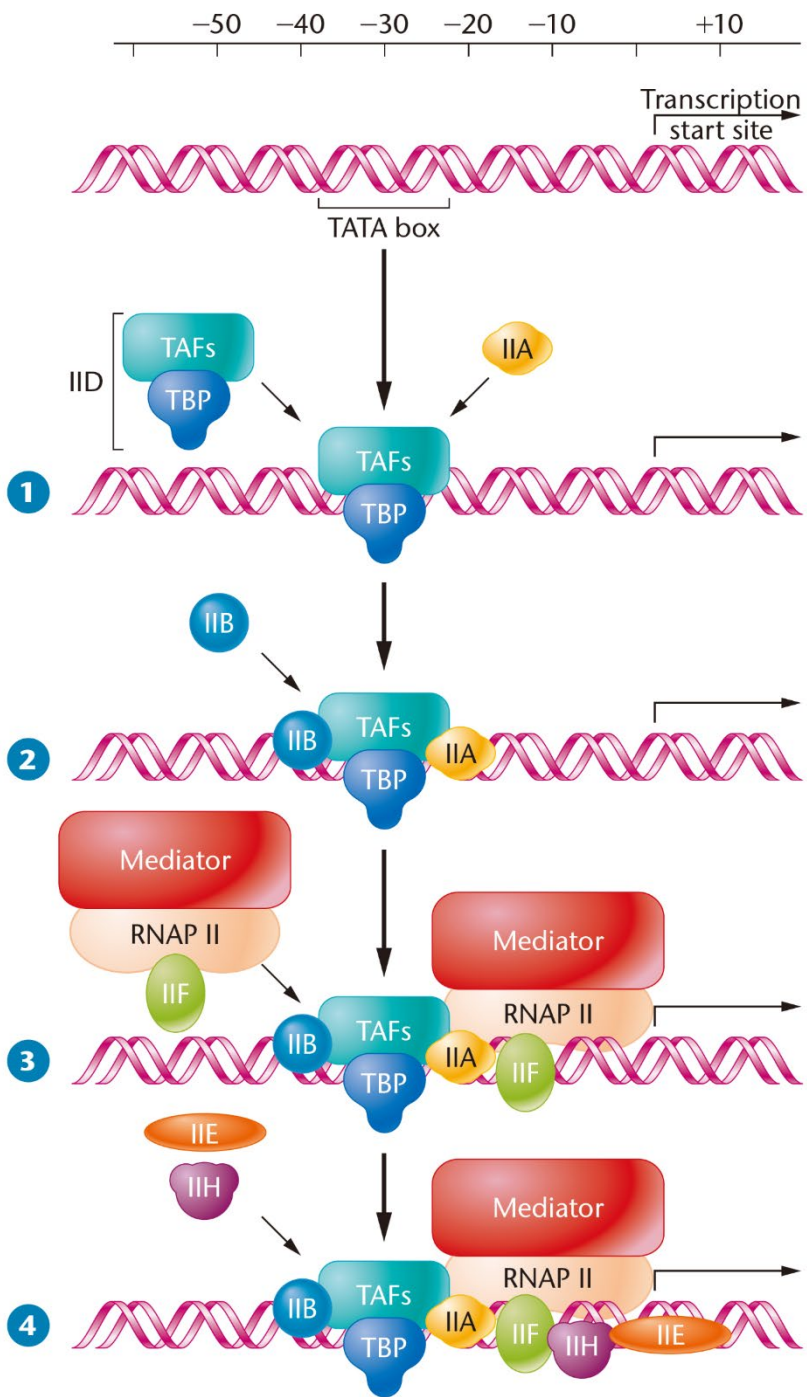




(3) 调控转录起始复合物形成的一般模型

研究表明，转录因子通过三种方式激活由RNA pol II进行的转录起始：（a）一些转录因子与靠近核心启动子的靶序列结合，然后直接与RNA pol II前起始复合物接触；（b）一些转录因子通过与辅激活因子或者中介体蛋白在转录因子与前起始复合物间形成蛋白质桥；（c）有些转录因子通过DNA弯曲而接触RNA pol II。

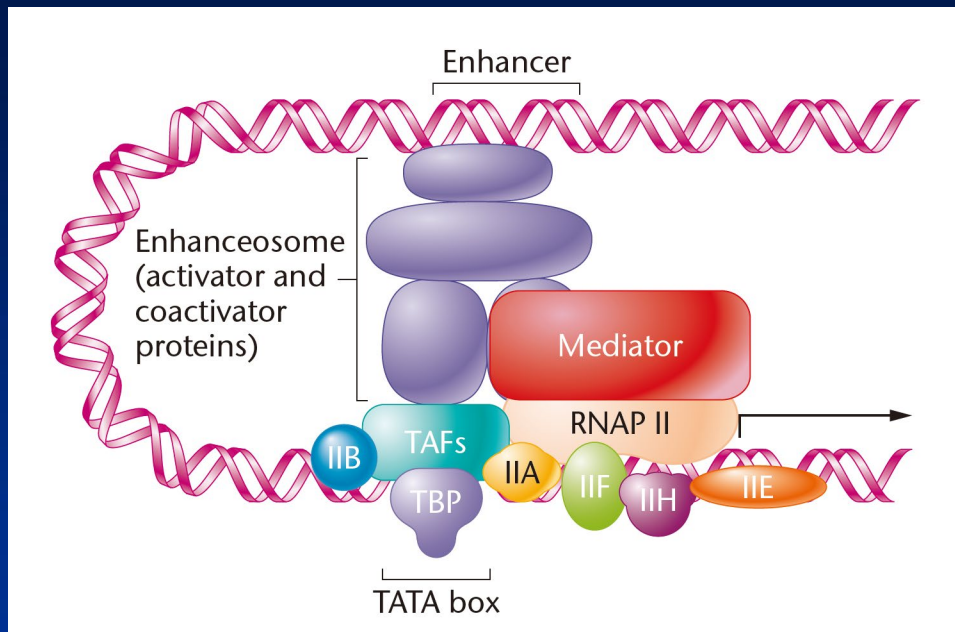




在真核生物转录水平的调控中，**转录起始是最关键的调控之一**。其关键的一个步骤是**形成预起始复合物**(pre-initiation complex, PIC)。第1步（图中的步骤1）是TFIID与TATA框的结合。TFIID是一种多亚基复合物，包含TATA结合蛋白（TATA-binding Protein, TBP）和大约13种称为TBP相关因子（TBP-associated factors, TAFs,）的蛋白；第2步TFIIB与TATA框附近的BRE (TFIIB recognition element)结合；最后，其他GTF（general transcription factor）(中介蛋白复合体、IIF、IIE和IIH)帮助招募RNA pol II到启动子上（步骤3和4）



而当转录起始位置DNA环形成时，使得在启动子远距离处的增强子结合因子与预起始复合物中的普通转录因子相互作用，形成**增强体（enhanceosome）**。这种由增强子、启动子以及由激活因子结合排列成的核蛋白结构称为增强体。基本转录机构（general transcription machinery）由Pol II、通用转录因子



TFIIA,TFIIB、TFIID、TFIIE、TFIIF、TFIIH以及被称为中介体蛋白的辅激活因子复合体组成。增强体与基本转录机构和核心启动子构成一个蛋白质—蛋白质和蛋白质—DNA作用的复杂体系，协同控制转录起始的频率，从而调节转录水平（图14-19）；或当沉默子存在时，阻遏蛋白结合到沉默子DNA序列上，会阻碍RNA聚合酶转录DNA序列，阻碍基因的表达。



14.4.3 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）。



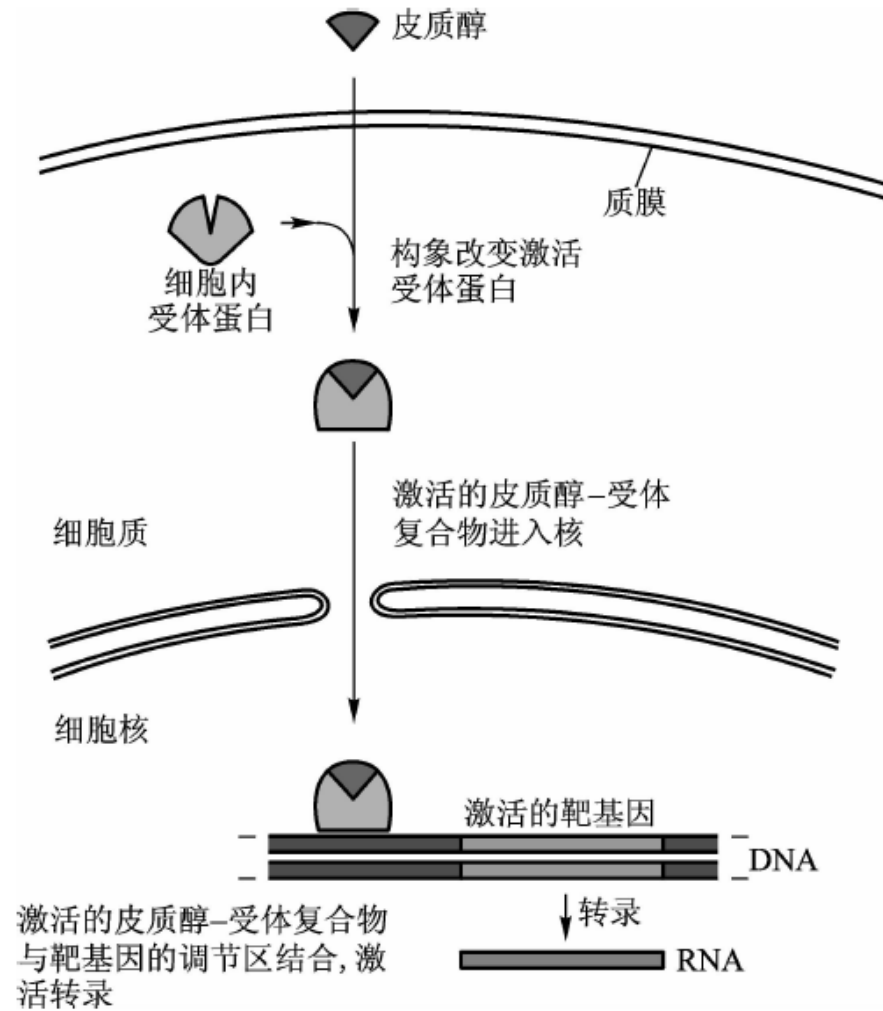


图 14-16 皮质激素激活基因调节蛋白以激发基因转录

(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。

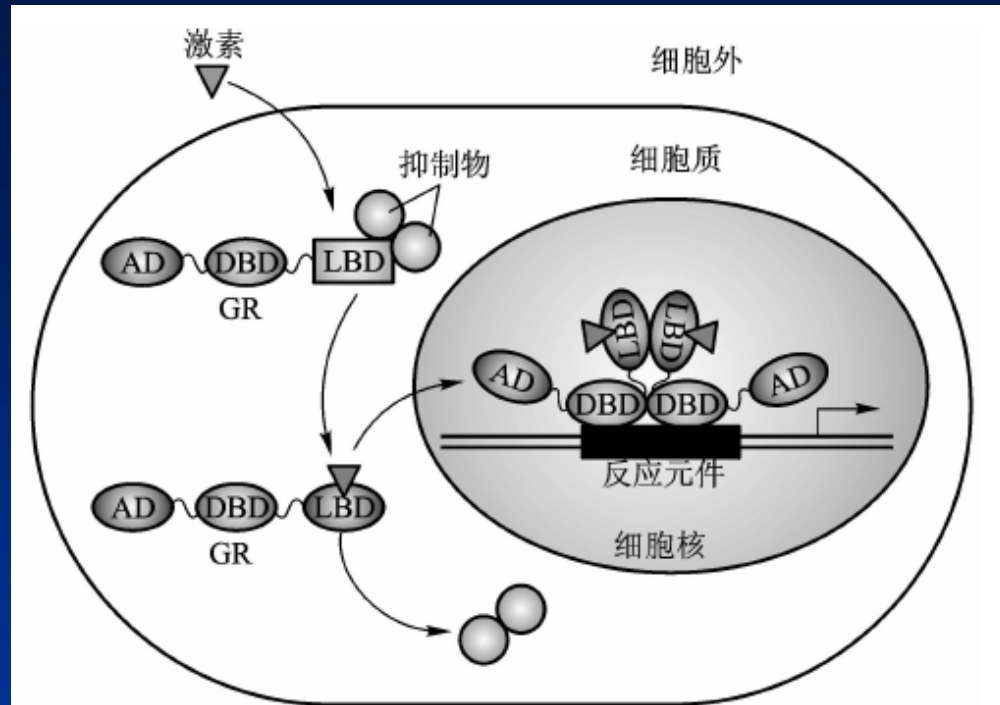


图14-17 激素激活细胞内受体的机制
(引自 Lodish 等, 2013)

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)**。

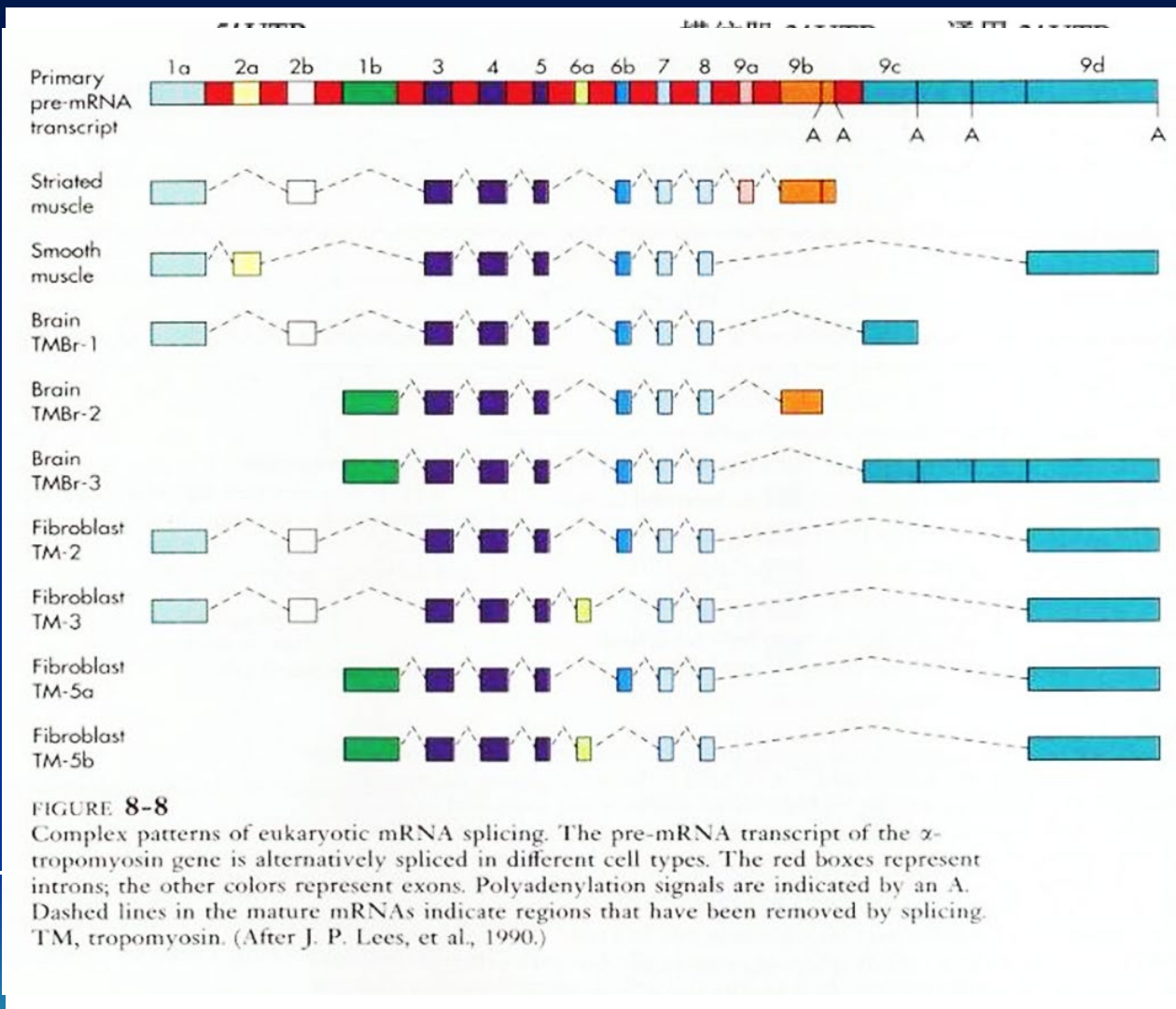
内含子的剪接一般都是发生在同一个基因内，切除内含子，相邻的外显子彼此连接，称为**顺式剪接 (cis-splicing)**。

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 α 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 α 原肌球蛋白（图 14-18）。



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



35 base leader GU

Left intron?

mRNA sequence A AG

Right intron?



35 base leader mRNA sequence

©virtualtext www.ergito.com

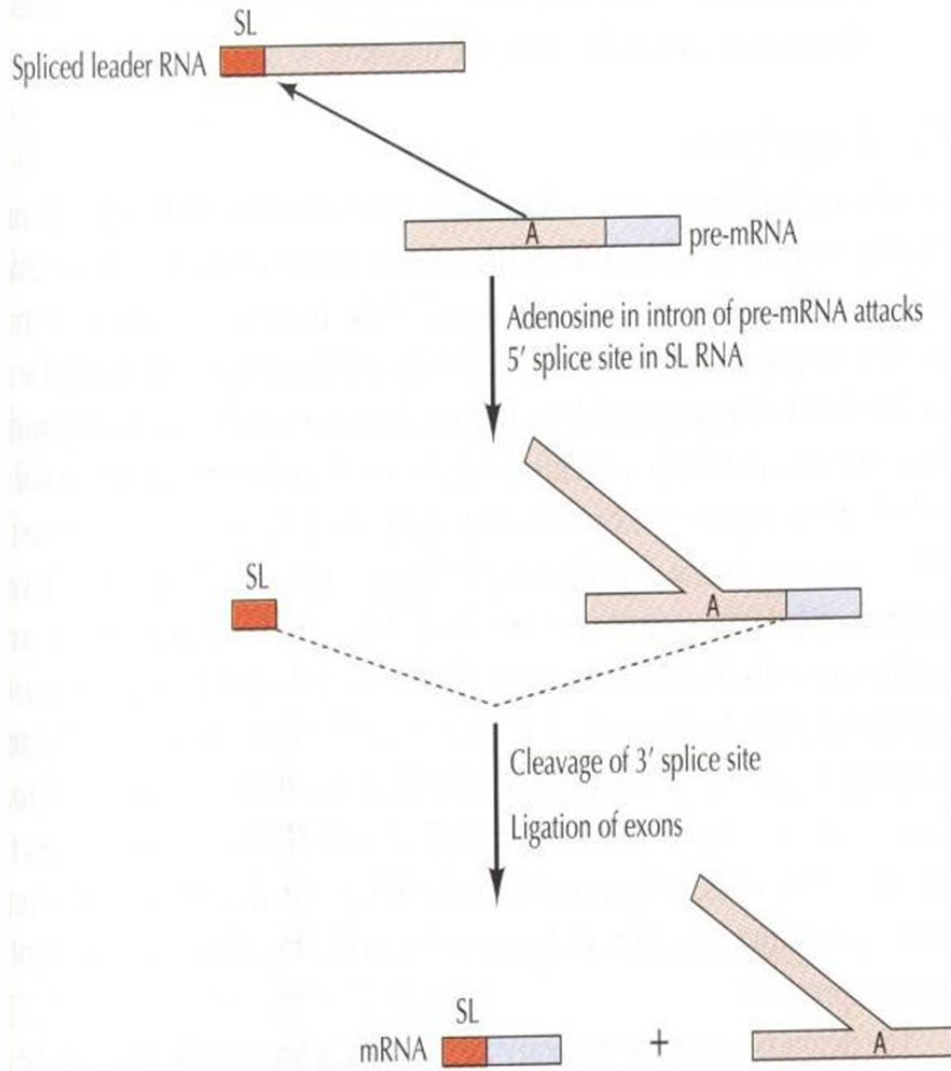


Figure 24.25 The SL RNA provides an exon that is connected to the first exon of an mRNA by *trans*-splicing. The reaction involves the same interactions as nuclear *cis*-splicing, but generates a Y-shaped RNA instead of a lariat.





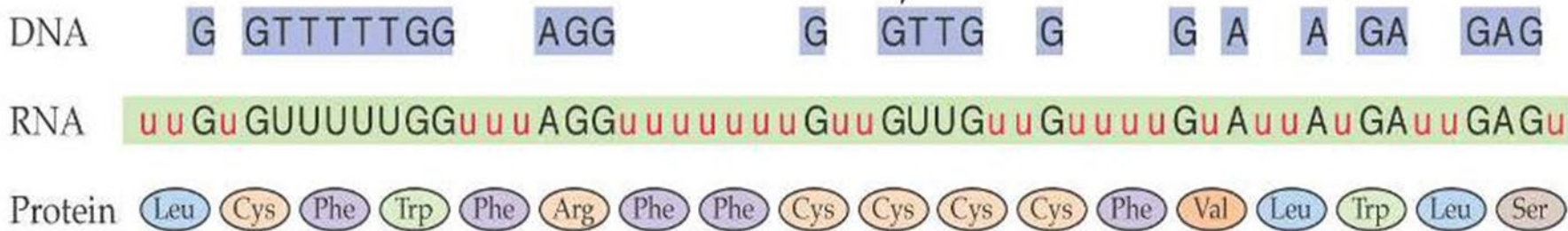
14.5.3 RNA编辑

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**。



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

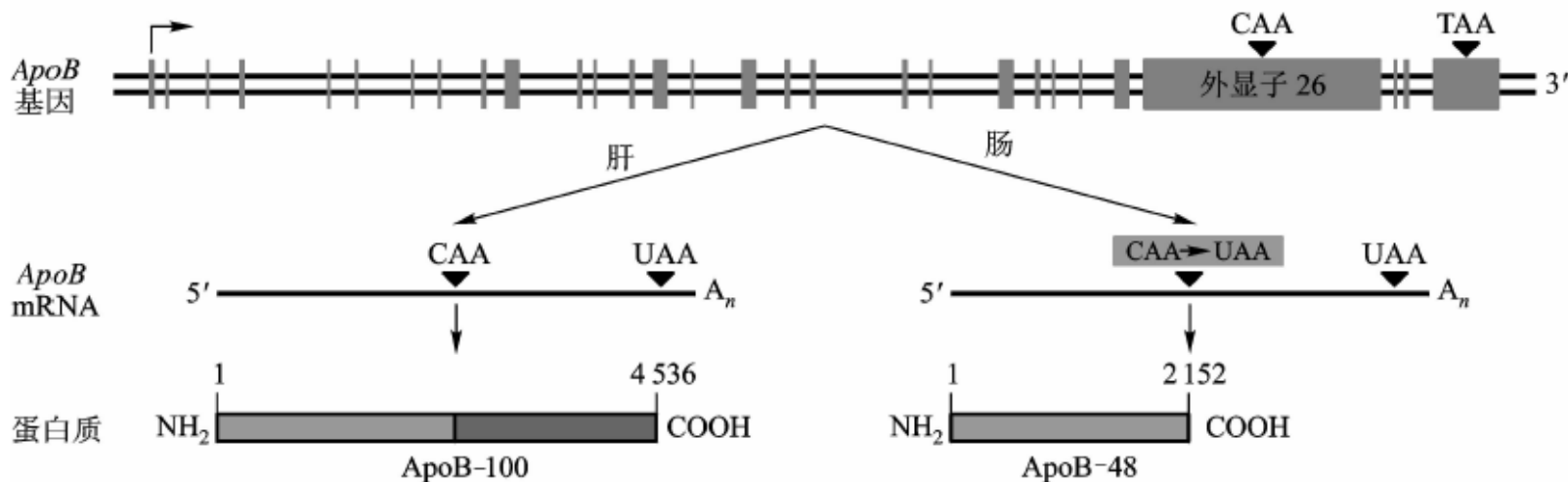


图 14 - 19 ApoB 前体 mRNA 的 RNA 编辑

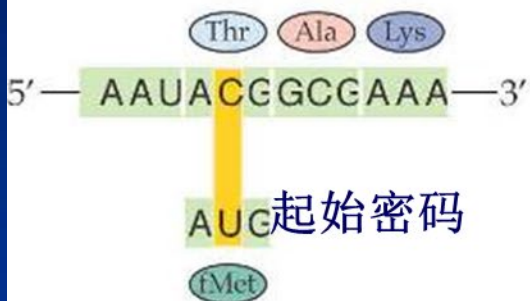
(仿自 Lodish 等, 2013)

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

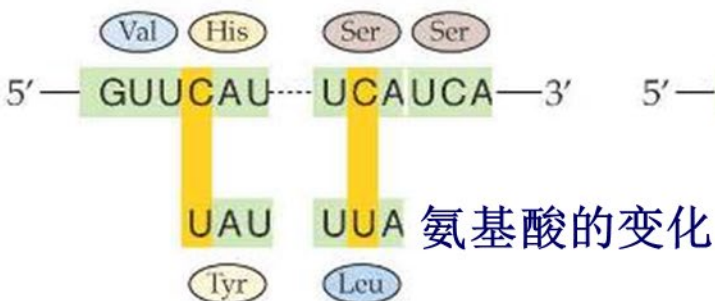


mRNA编辑的几种类型

(A) Creation of initiation codon



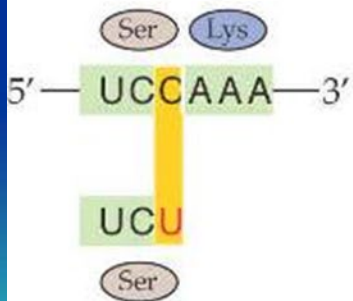
(B) Amino acid changes



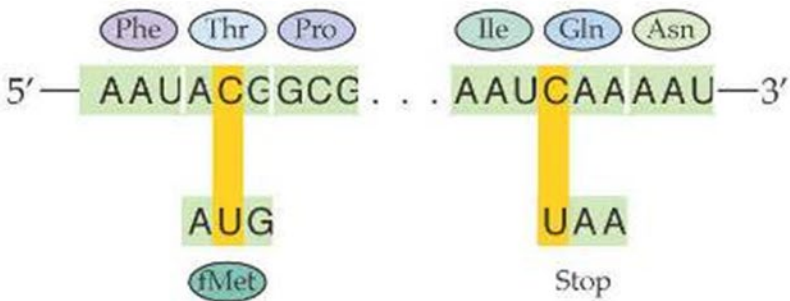
(C) Creation of a stop codon



(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

翻译调节是真核生物基因表达多步骤调节过程中的一个重要环节。

一般涉及 翻译起始复合体的装配调节

mRNA的稳定性及其5'端3'端结构（UTR）对翻译的影响

多种蛋白质因子间的相互作用等。

(1) 调节翻译起始复合体的装配

mRNA的5'帽结构是起始因子eIF-4F(由eIF-4A、eIF-4E和eIF-4G三个亚基组成)识别并结合于mRNA以及最终形成翻译起始复合物所必需的。

大多数真核生物mRNA的翻译起始活性依赖于5'端帽子结构的存在。



核糖体的小亚基不能单独识别mRNA的5'帽结构，而是可以识别帽周围更复杂的结构（图14-24）。由三个起始因子eIF-4A、eIF-4E和eIF-4G组成的复合体与5'帽结合，然后，该复合物中的eIF-4G蛋白在poly-A尾处与poly-A结合蛋白（Poly-A binding protein, PABP）相互作用。结果，mRNA环化，正是这种具有复杂起始因子结合的mRNA环化结构，从而启动了翻译。核糖体小亚基在RNA/蛋白质复合物的背景下识别mRNA 5'帽结构，这使得其能够通过控制复合物装配来调节翻译。

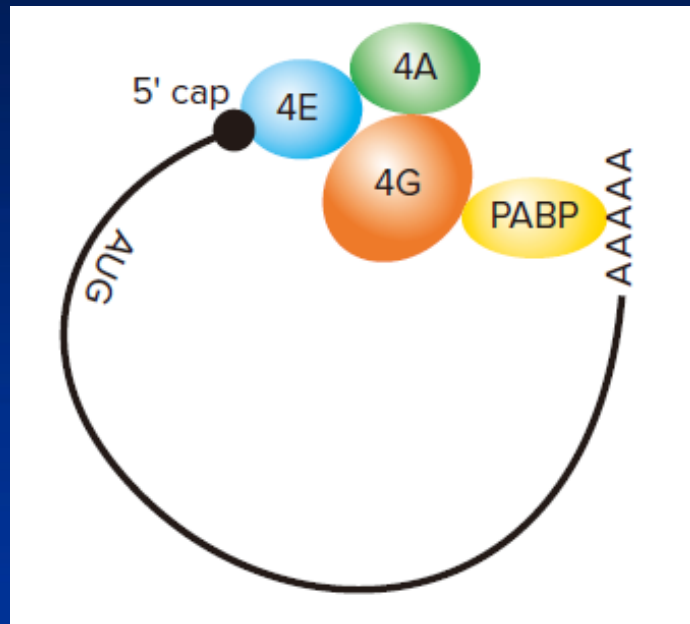


图14-24 真核生物翻译起始复合物



同时，细胞外环境中的营养水平也可调节mRNA的翻译起始水平。

对细胞外刺激做出适当反应的重要途径取决于称为eIF4E结合蛋白1（4E binding protein1, 4E-BP1）的蛋白。

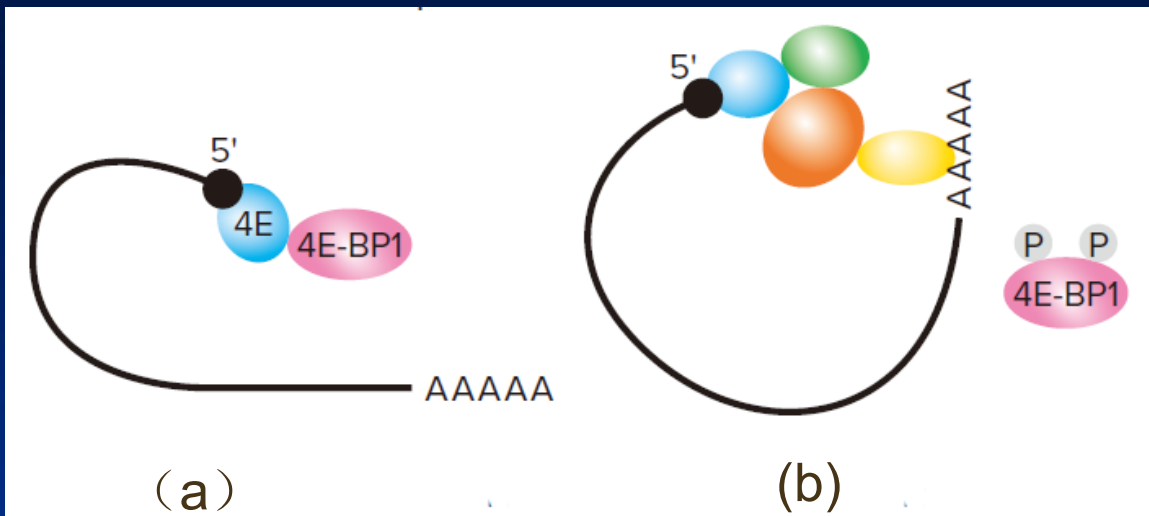


图14-25 受4E-BP1控制的翻译调控

4E-BP1可以与起始因子eIF-4E结合，从而阻止起始复合物的其余部分在5'帽上的组装（图14-25a）。在细胞外环境中营养物质和生长因子的存在会激活激酶将磷酸基团添加到4E-BP1上。当4E-BP1磷酸化时，它不再与eIF-4E结合，因此起始复合物可以在mRNA的5'端组装（图14-25b）。当细胞需要生长并产生大量蛋白质时，这种机制可确保mRNA翻译的整体频率显著增加。在这种情况下下的研究表明，营养物质缺乏时，出现低mRNA翻译，而营养物质增加时，表现mRNA翻译升高。



通过5'帽的起始因子和3'poly-A尾部的PABP的结合,使mRNA的环化是通过控制poly-A尾长度间接调节翻译起始机制的物理基础。较长的尾比较短的尾更有效地吸引PABP,因此Poly-A尾越长,翻译起始复合物形成的效率越高,翻译越多(图14-26)。序列特异的RNA结合蛋白结合mRNA上的位点,并募集可将As添加至尾部或删除它们的酶。

有趣的是,该机制部分负责一天中不同时间某些蛋白质的昼夜节律振荡。尽管这些蛋白质的mRNA量保持恒定,但poly-A尾巴的大小以及mRNA的翻译效率以及这些蛋白质的量在深夜和清晨达到峰值。

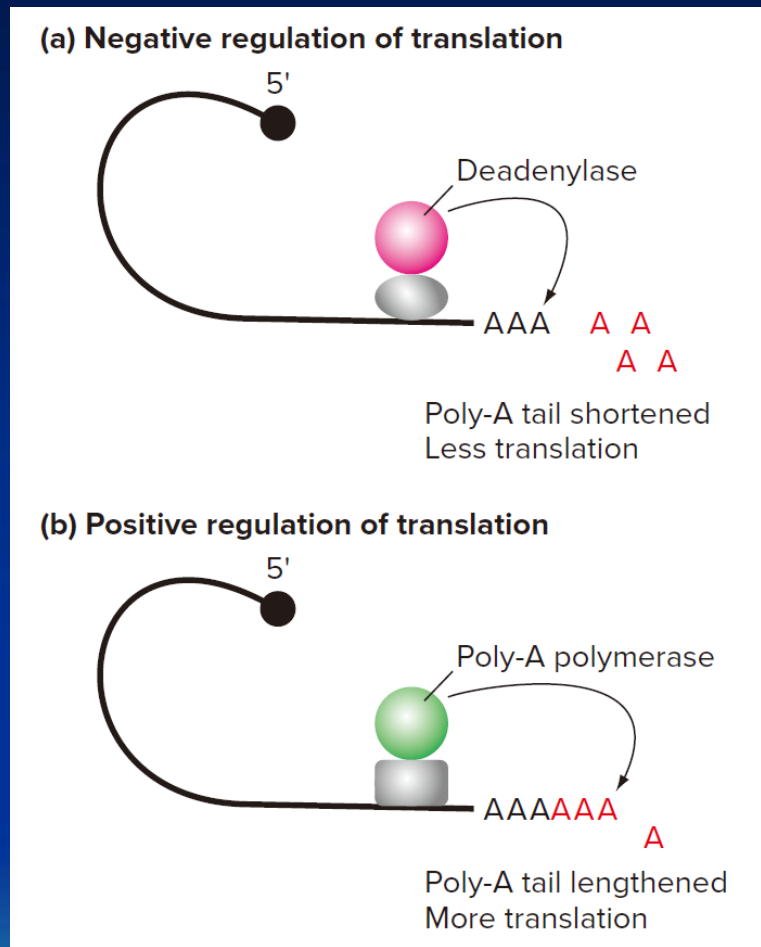
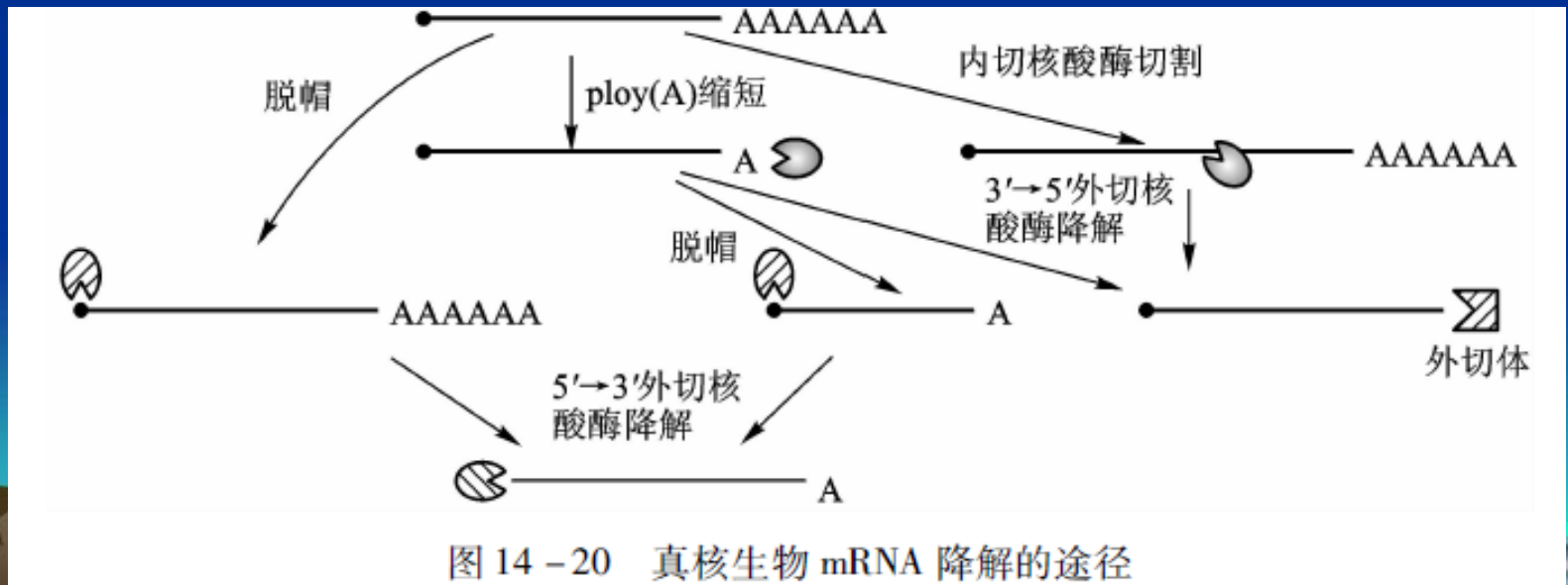


图14-26 通过poly-A尾部长度的翻译控制

(3) 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。根据对酵母以及哺乳动物细胞mRNA降解的种种分析表明,它们的降解过程都起始于3'端,或是去除尾部腺苷酸,或是在3'-UTR区发生专一性剪接,或是破坏并消除3'端茎-环结构。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。





有研究表明，真核细胞中mRNA的衰减（降解）途径主要有6种（表14-1）。其中绝大多数mRNA的降解是依赖于脱腺苷化作用的。

表14-1 在真核细胞中mRNA衰减（降解）途径的关键因素总结

Pathway 途径	Initiating Event 启动事件	Secondary Step(s) 第二步（多步）	Substrates 底物
依赖于脱腺苷化的 5'至3'消化	脱腺苷化寡聚（A）	通过Lsm复合物结合Oligo（A）脱帽以及XRN1的5'到3'核酸外切酶的消化	可能是大多数多腺苷化的mRNA
依赖于脱腺苷化的 3'至5'消化	脱腺苷化寡聚（A）	3'至5'核酸外切酶消化	可能是大多数多腺苷化的mRNA
不依赖于脱腺苷化的脱帽	5'端脱帽	5'至3'核酸外切酶消化	很少的特定mRNA
核酸内切酶途径	核酸内切酶切割	5'至3'和3'至5'核酸外切酶消化	很少的特定mRNA
组蛋白mRNA途径	寡聚尿苷	通过Lsm复合物结合Oligo（U）脱帽和XRN1 5'至3'外切酶消化以及外泌体3'至5'消化	哺乳动物中的组蛋白mRNA
miRNA途径	在RISC中与miRNA的碱基配对	核酸内切酶切割或翻译抑制	许多mRNA（程度未知）



通过去除poly (A) 可以启动两种不同的mRNA降解途径。

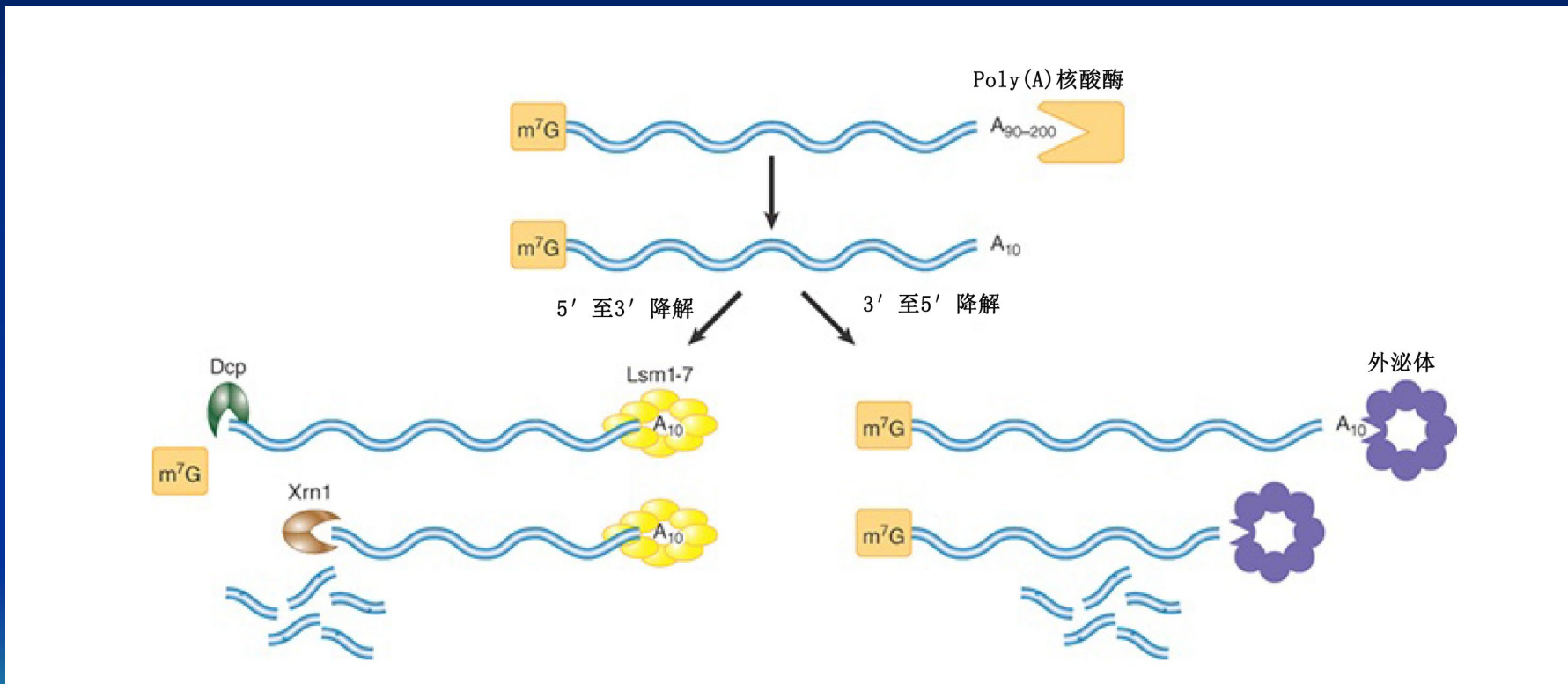


图14-27 真核生物中mRNA主要依赖于脱腺苷化的降解途径

(引自Krebs等, 2018)

, 它们适用于特定的mRNAs亚群, 通常涉及调节降解事件。

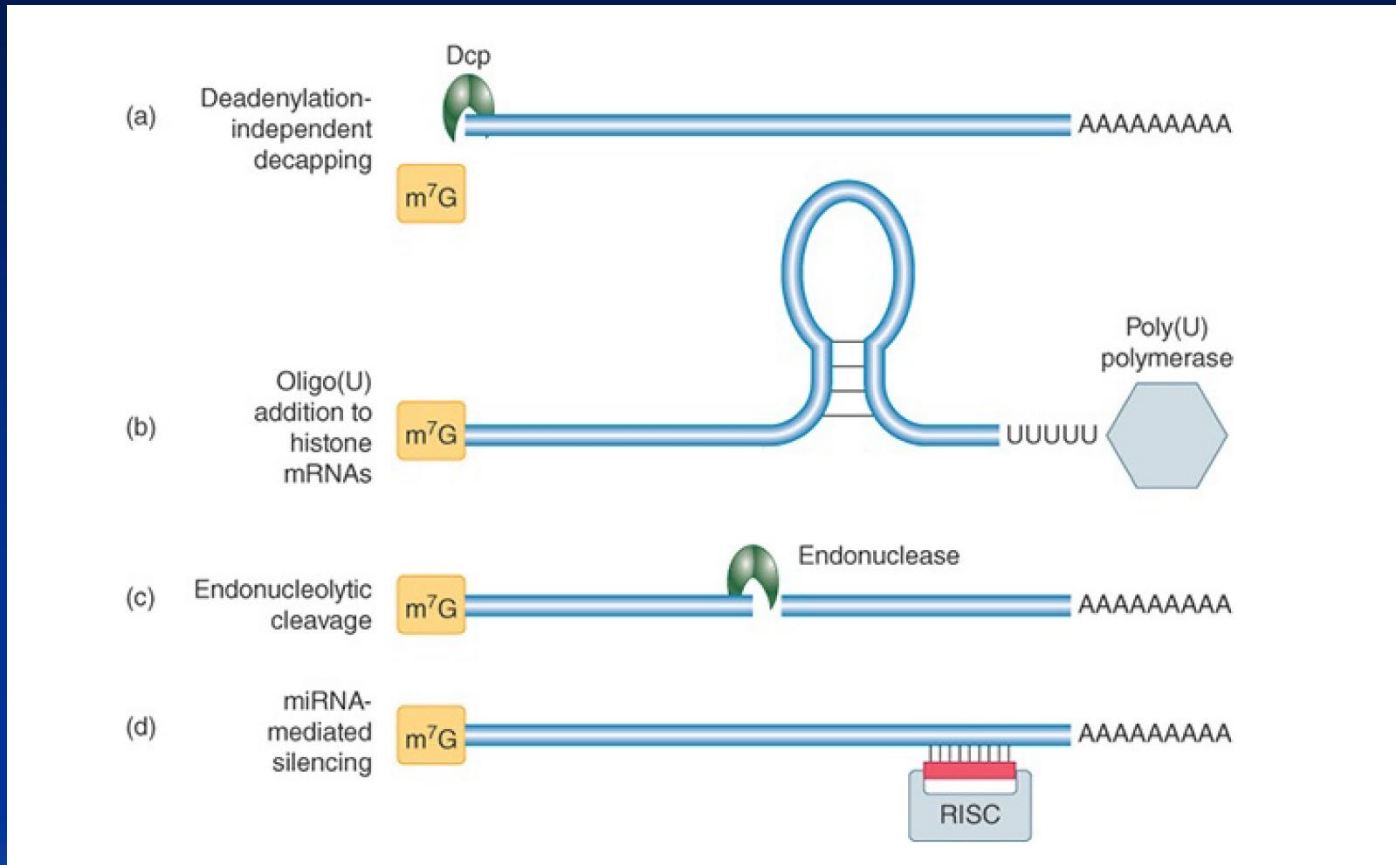


图14-28 在真核细胞中mRNA的其他4种降解途径

(a) 一些mRNA在脱腺苷化之前可能已脱帽。(b) 组蛋白的mRNA接收一条短的poly(U)尾巴, 成为衰变的底物。(c) 某些mRNA的降解可以通过序列特异性的内切核酸酶切来引发。(d) 一些mRNA可以通过互补的引导miRNA靶向降解或翻译沉默。



(4) 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。

某些上游ORFs可实现诱饵ORF对翻译的调节，某些mRNAs有一个或多个上游开放阅读框(upstream open reading frames, UORFs)，从诱饵AUG开始，编码没有功能的小肽。如果核糖体翻译uORF，则mRNA中主ORF的翻译受到抑制。因此，在翻译uORF和主ORF之间的选择是一个潜在的基因表达调控点



例如果蝇中，一个由uORFs翻译调节的例子是果蝇性别分化途径，一种“性致死 (Sex lethal, Sxl)”蛋白只在XX蝇中表达，它是XX蝇作为雌性发育所必需的。另一方面，如果雌性产生一种不同的、通常是雄性特有的蛋白质Ms1-2，则它们会死亡。Sxl蛋白可通过多种机制确保雌性体内不合成Ms1-2蛋白，其中之一是Sxl蛋白阻止了*msl-2* mRNA的翻译(图14-29)。Sxl蛋白在特定的结合位点与*msl-2* mRNA结合，这种结合促进uORF的翻译，从而阻止编码Ms1-2蛋白的主ORF的翻译，从而使这些XX动物得以生存(图14-29)

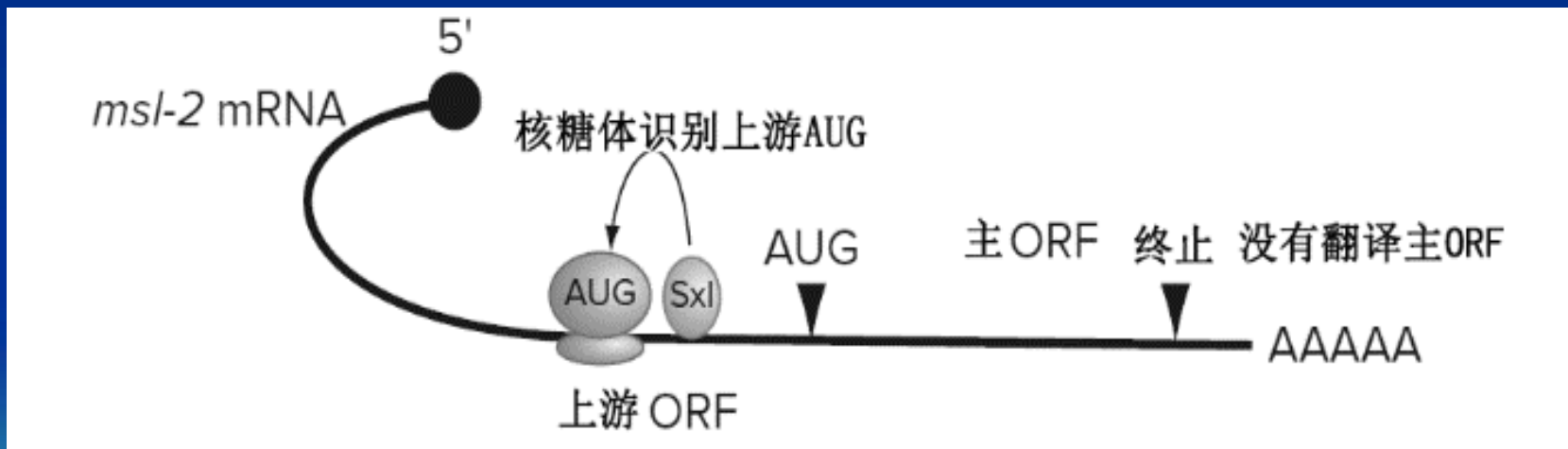


图14-29 一种诱饵ORF对翻译的调节

为了防止果蝇雌性中*msl-2* mRNA的翻译，Sxl蛋白与mRNA结合，导致核糖体翻译uORF。因此，Sxl蛋白阻止了编码Ms1-2蛋白的主ORF的翻译。



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。



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 β 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）。

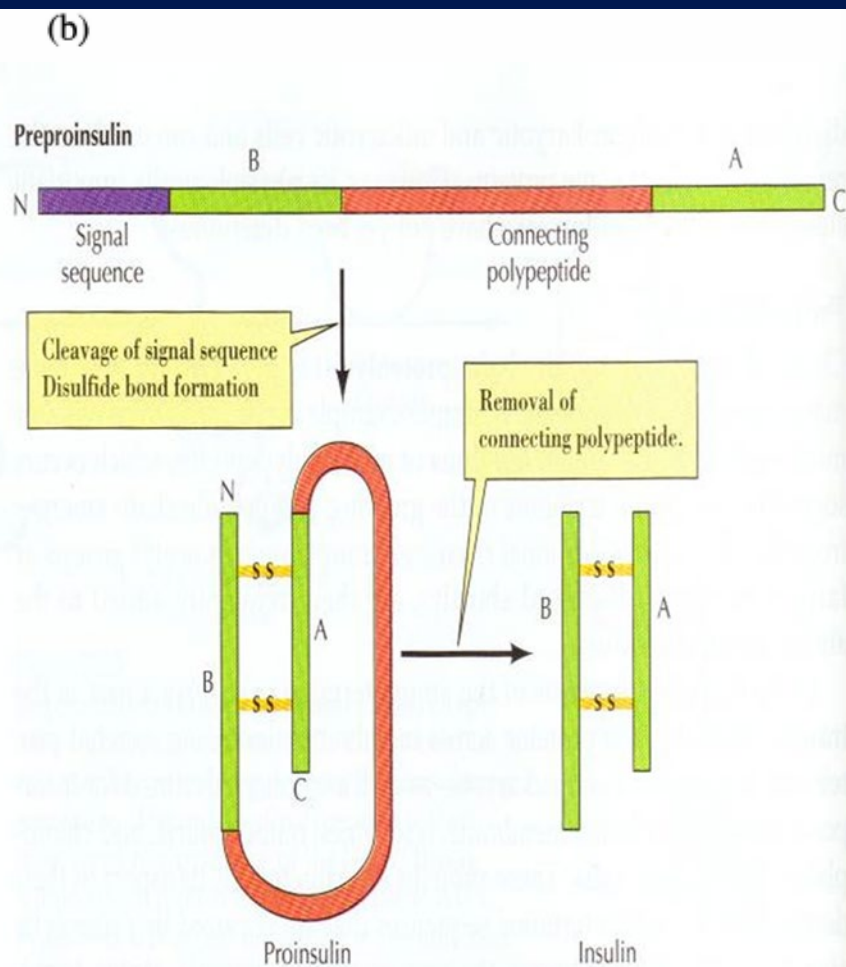
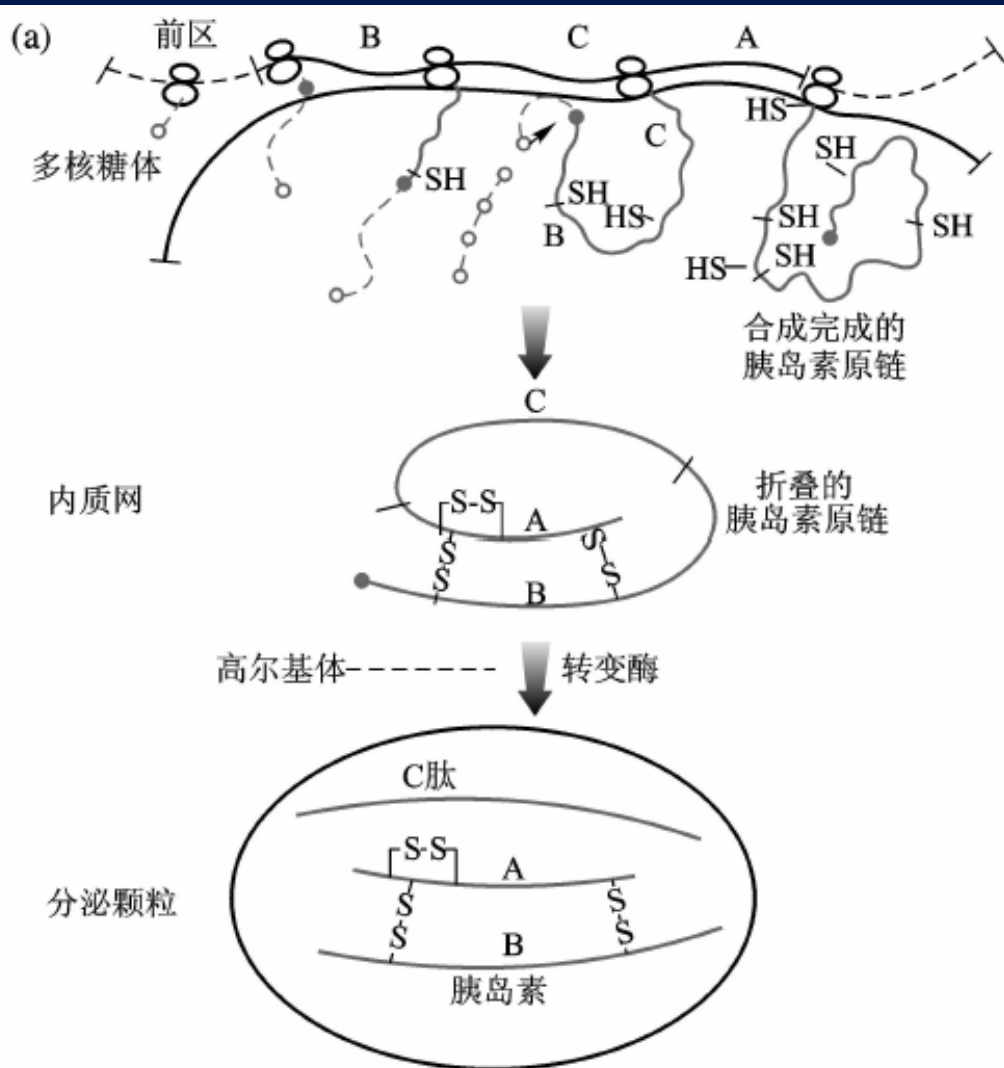


图 14 - 21 前胰岛素原加工成胰岛素示意图



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)**。

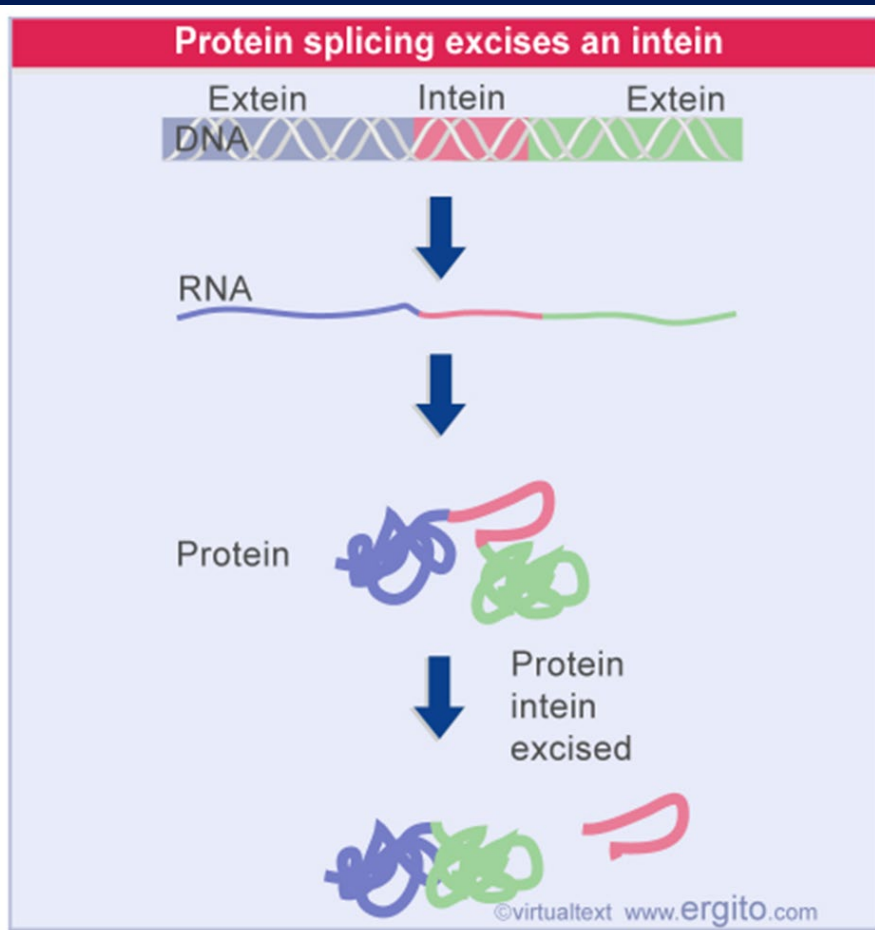


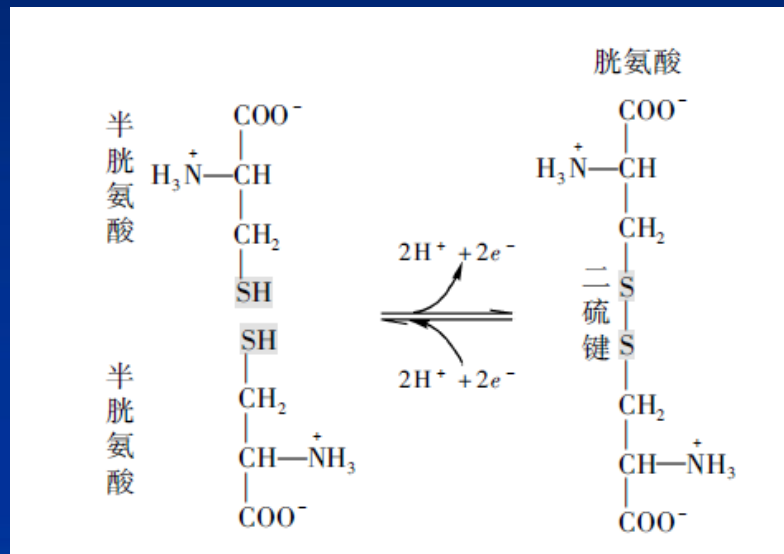
Figure 26.24 In protein splicing the exteins are connected by removing the intein from the protein.



(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。



14.7.1 RNA干扰与基因表达调控

(1) RNA干扰的发现

RNA干扰 (RNA interference, RNAi) 是指正常生物体内一些小的双链RNA有效地阻断靶基因表达的现象。当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, 这种现象发生在转录后水平, 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)

RNA干扰现象首先是在植物的转基因中发现的

1990年, R. Jorgensen小组 C. Napoli试图通过引入嵌合矮牵牛查尔酮合成酶 (chalcone synthase) 基因到有色矮牵牛花瓣中过表达, 以加深矮牵牛花的紫色。结果, 转基因的植株一些花瓣颜色并未加深, 反而呈杂色甚至白色



1997, R. Jorgensen小组的Que等继续进行该研究, 发现不但自身转入的酶基因不能正常表达, 而且使植物本身内源的查耳酮合成酶的表达也出现了异常, 从而形成全白到各种花色嵌合体的表型, 当时这种现象被他们称为“共抑制” (co-suppression)。

1995年, Su Guo等发现注射正义RNA和反义RNA均能有效并特异性地抑制秀丽新小杆线虫(*C. elegans*) *par-1*基因的表达, 该结果不能使用反义RNA技术的理论做出合理解释。

1998年, Fire等深入研究后发现植物中dsRNA能触发RNA沉默以及线虫中dsRNA可触发RNA干扰的现象, 可能存在一种共同的机制。他们发现Su Guo博士遇到的正义RNA抑制基因表达的现象, 以及过去的反义RNA技术对基因表达的阻断, 都是由于体外转录所得RNA中污染了微量双链RNA (dsRNA) 而引起。

而一旦混入dsRNA, 其阻抑效应比仅仅任一单链RNA强10倍以上, 纯化的dsRNA则可产生更强的基因阻抑效应。



- 实际上每个细胞只要很少几个分子的dsRNA已经足够完全阻断同源基因的表达。后来的实验表明在线虫消化道中注入双链RNA不单可以阻断整个线虫的同源基因表达，还会导致其第一代子代的同源基因沉默。他们将这种现象称为RNA干扰（RNA interference, 简称RNAi,）
- 当他们将体外转录得到的
单链RNA→纯化→注射线虫→基因抑制效应变得十分微弱
- 双链RNA→纯化→注射线虫→能够高效特异性阻断相应基因的表达，其抑制基因表达的效率比纯化后的反义RNA至少高2个数量级。

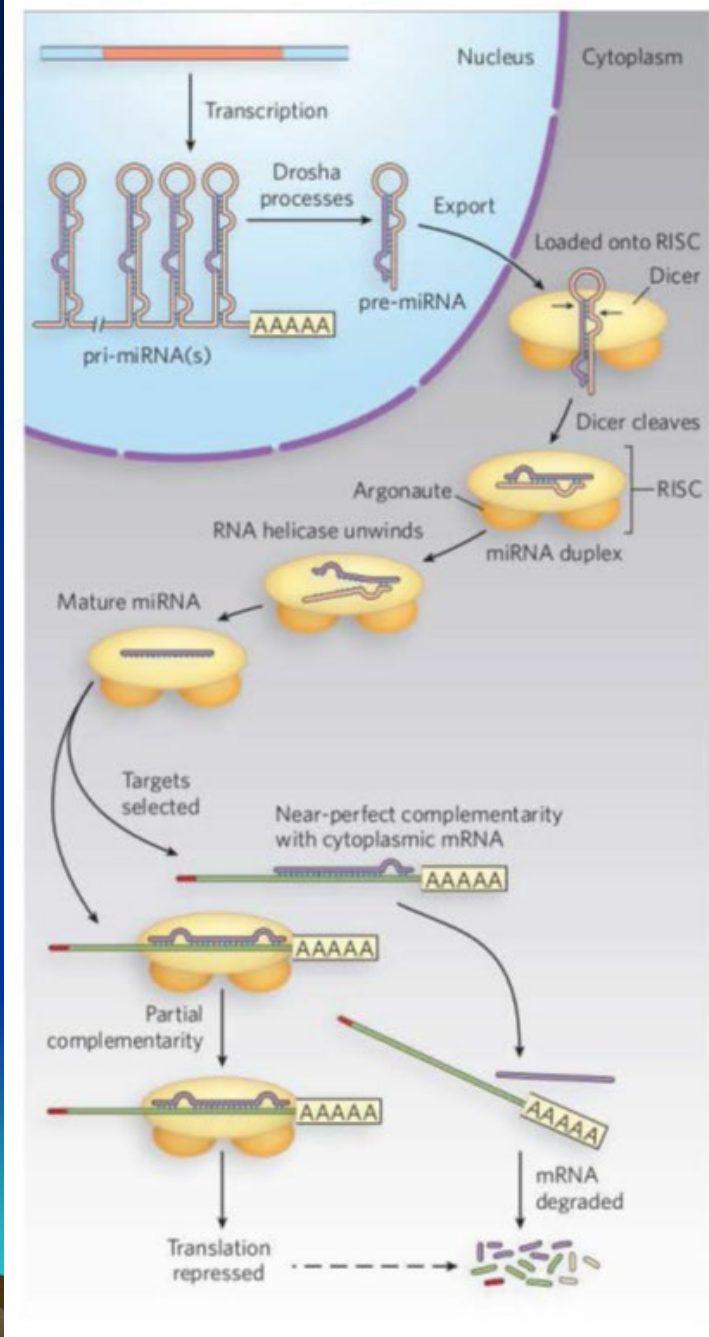
该小组将这一现象称为RNA干扰



(2) RNA干扰与基因沉默

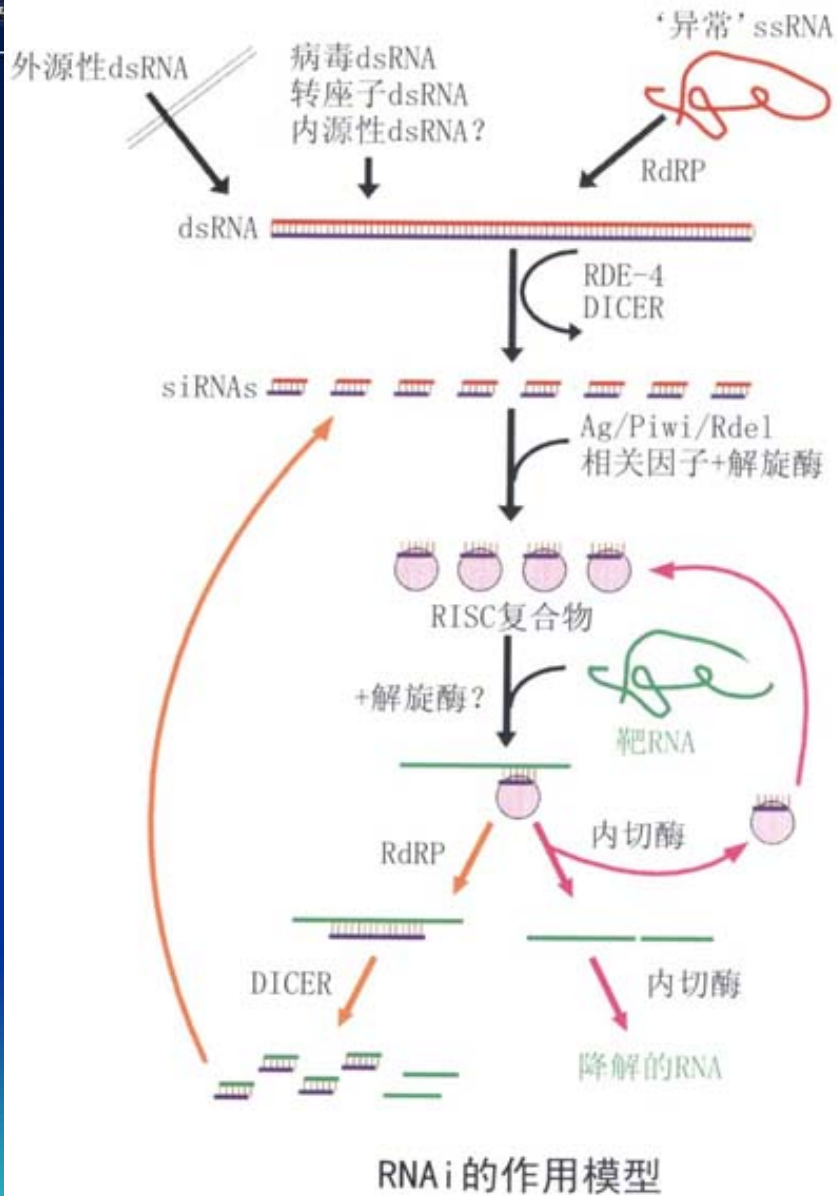
在动植物中，最丰富的一类小RNA是microRNA (miRNA)。

人类基因组有近1000个编码miRNA的基因。这些基因被RNA Pol II转录成为pri-miRNA的长初级转录本，其中大多数包含一个或多个以双链茎环形式存在的miRNA序列。pri-miRNA经过加工形成活性miRNA，它们是短链的单链。加工过程由Drosha和Dicer两种核糖核酸酶辅助进行（图14-7-1）。



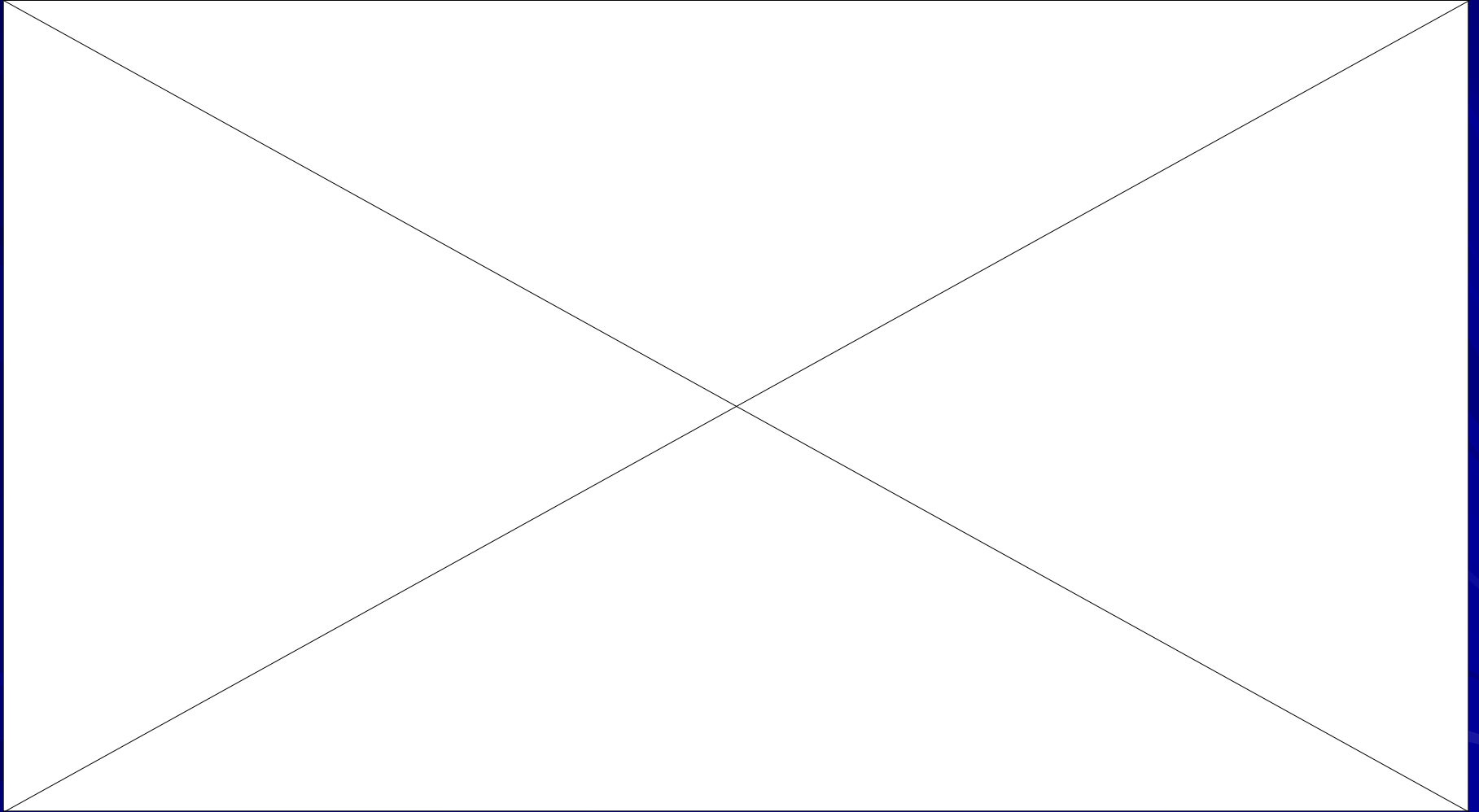


miRNA序列从细胞核转运到细胞质。接着，miRNA被整合并入到核糖核蛋白复合物中，形成miRNA诱导的沉默复合物（miRNA-induced silencing complex，miRISC）。每个miRISC都包含Argonaute蛋白家族的特定成员。miRISC可以介导各种功能，具体取决于它们拥有的特定Argonaute蛋白质，以及复合物中的miRNA与mRNA 3'UTR中的靶序列之间的序列互补程度。一个带引导miRNA的miRISC与目标mRNA完全互补时，可引起mRNA切割。而互补性较低时，其机制通常是抑制翻译（图14-7-1），但目前尚未确切了解miRISC是如何调节翻译活性的。



研究还发现siRNA有一个扩增的过程，即由siRNA导致mRNA切断后，含polyA的mRNA后半段会在核酸酶的作用下从5'到3'外切而降解，而含帽子的mRNA前半段在依赖于RNA的RNA聚合酶（RNA depended RNA polymerase, RdRP）作用下合成双链RNA。该双链RNA又可作为Dicer的底物，产生更多新的siRNA，这些新生的siRNA亚群也称二级siRNA，而这些二级siRNA又能继续作用于靶mRNA令其降解。显然，该扩增机制赋予了RNAi高效性和持久性，小剂量dsRNA就能诱发强烈的基因沉默效应（图17-4-4）。

RNA干涉





14.7.2 LncRNA

LncRNA (Long non-coding RNA) 为长链非编码RNA，**长度在200nt以上的不具有编码蛋白的调控非编码RNA**，无长的开放阅读框(ORF)，在某些特定条件下有些lncRNA也可以编码功能性短肽。lncRNA种类很多，可以通过多种机制调节基因表达，广泛参与了生物生长发育、生理、病理和抗逆等过程，主要涉及基因转录、染色质修饰、转录后调控、基因组修饰、mRNA降解、协助蛋白定位和调节蛋白活性等生命过程

根据其**与编码蛋白基因的相对位置**，一般可以将lncRNA分为5类：

正义 lncRNA (sense lncRNA)

反义 lncRNA (antisense lncRNA)

内含子lncRNA (intronic lncRNA)

双向lncRNA (bidirectional lncRNA)

基因间lncRNA (intergenic lncRNA)

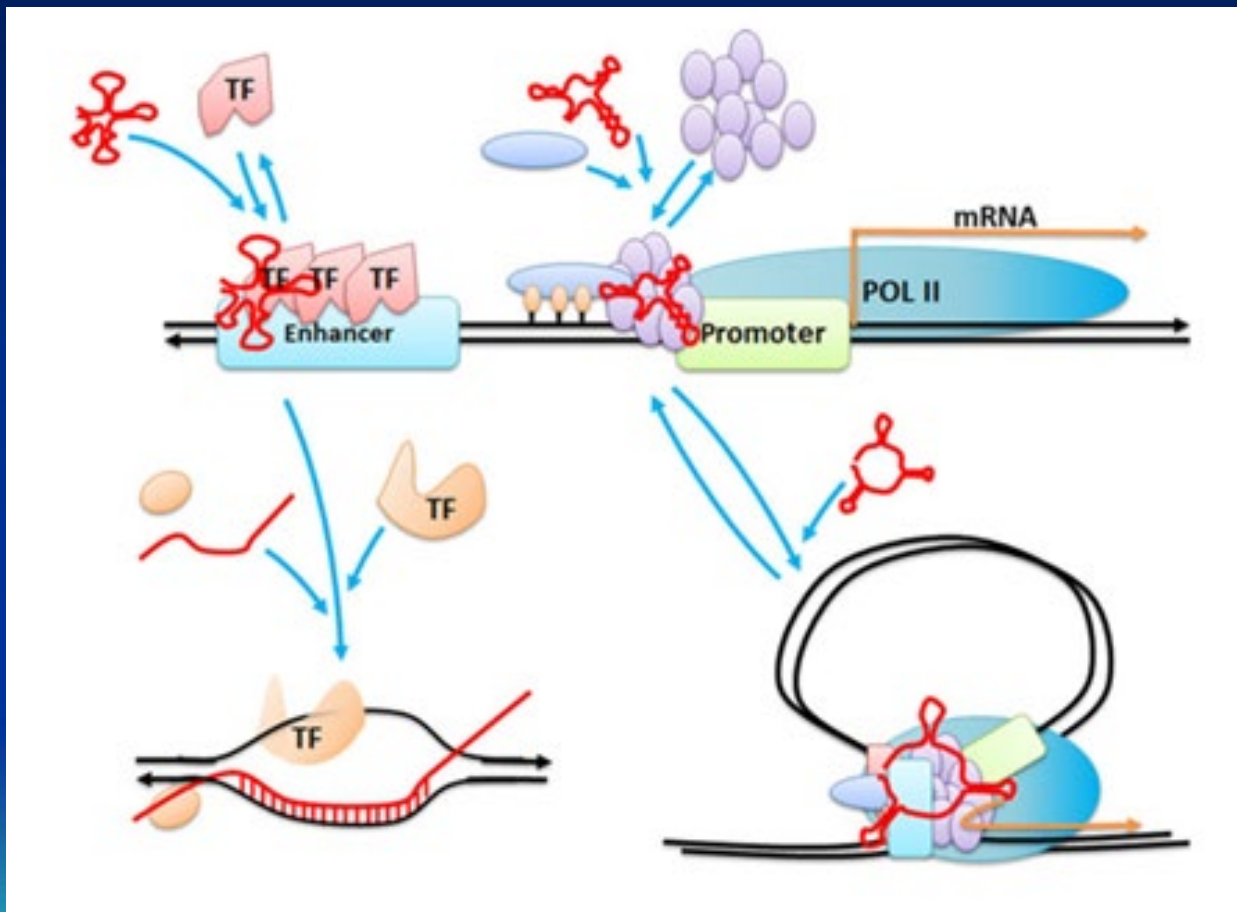


lncRNA参与调控的作用机制非常复杂，至今尚未完全清楚。

(1) 在转录起始过程中，lncRNA可以多种方式直接调节Pol II转录机构

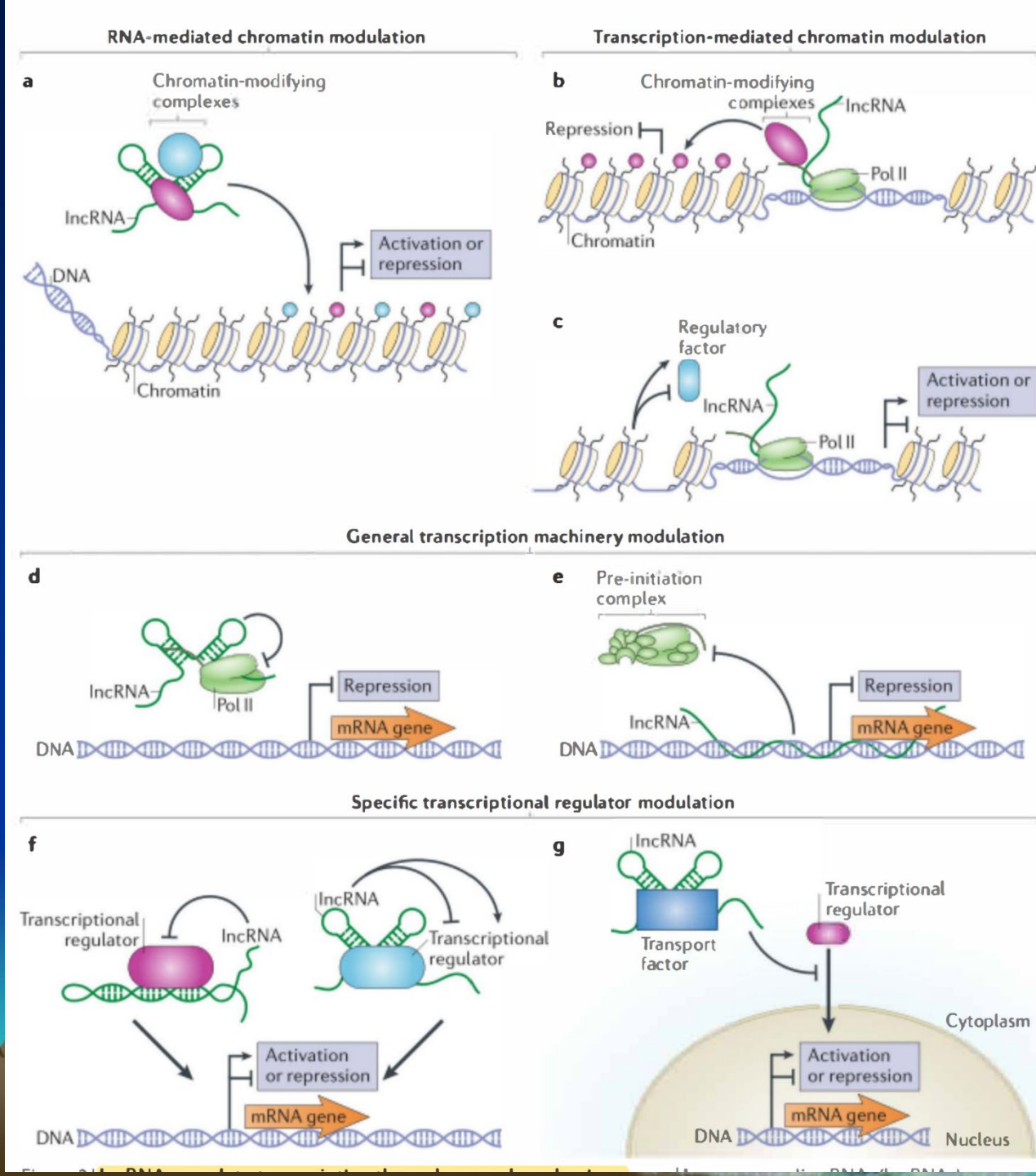
lncRNA（红色）可以在启动子区域调节转录机构和染色质修饰，以及在核内增强子区域调节转录因子结合亲和力。

lncRNA还可以调节RNA与DNA的杂交和染色质的拓扑结构



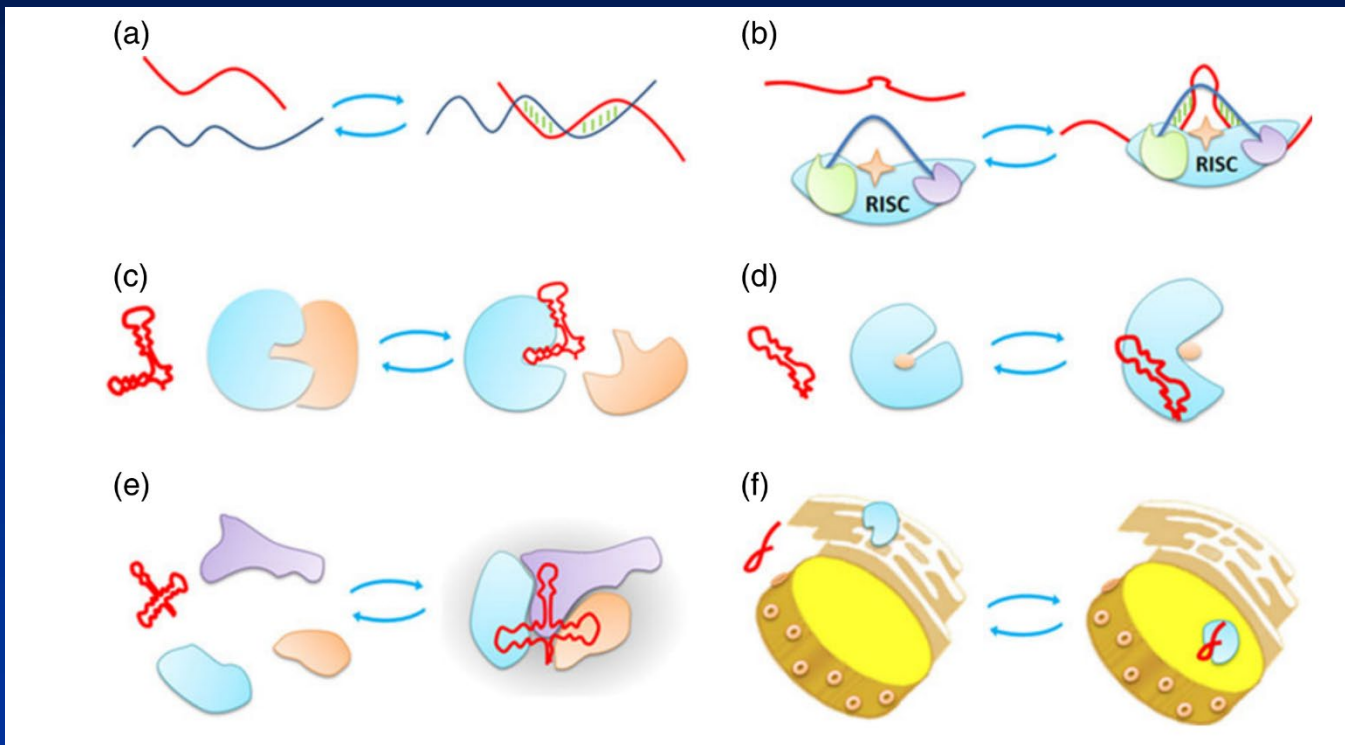
(2) lncRNA可以通过

多种机制调节转录：转录独立机制（a）和转录依赖机制（b和c）来调节染色质状态。在转录独立机制中，采用RNA介导的染色质调节；在转录依赖机制中，通过转录介导的染色质调节来调节染色质从而调节转录。





(3) 转录后和翻译后 lncRNA 也可以以多种机制参与调节



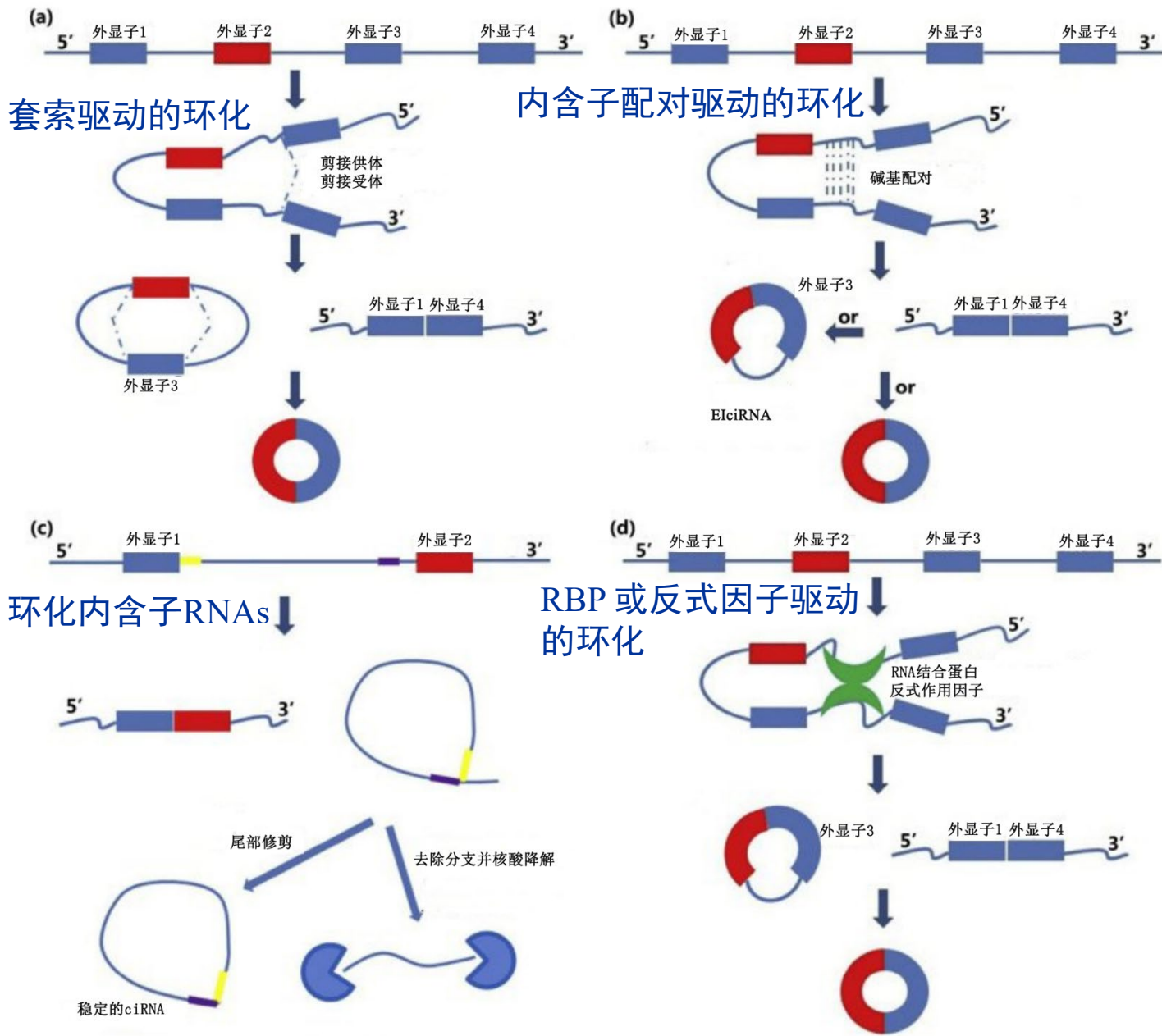
lncRNA 与其靶 RNA 互补序列形成双链 RNA 小片段 (图 a)；lncRNA 作为一个模拟靶标，与 RNA 诱导的沉默复合物 (RISC) 中的 RNA 结合下调 miRNA 的活性 (图 b)；lncRNA 也可调节蛋白质-蛋白质相互作用 (图 c)；lncRNA 可以改变蛋白质结构以暴露或保护特定氨基酸进行修饰 (图 d)；lncRNA 还可作为支架调节蛋白质复合物亚基的组装 (图 e)；lncRNA 也可以指导 RNA 结合蛋白在细胞中的重定位 (图 f)。



环状RNA (circular RNA, circRNA) 是一类不具有5' 末端帽子和3' 末端 poly(A)尾巴, 并以共价键形成的单链闭合的环状RNA分子, 主要由pre-mRNA通过可变剪切加工产生, 大量存在于真核生物转录组中。

CircRNA 的来源方式多种多样, 按照其在生物体内产生的模式, 主要分为

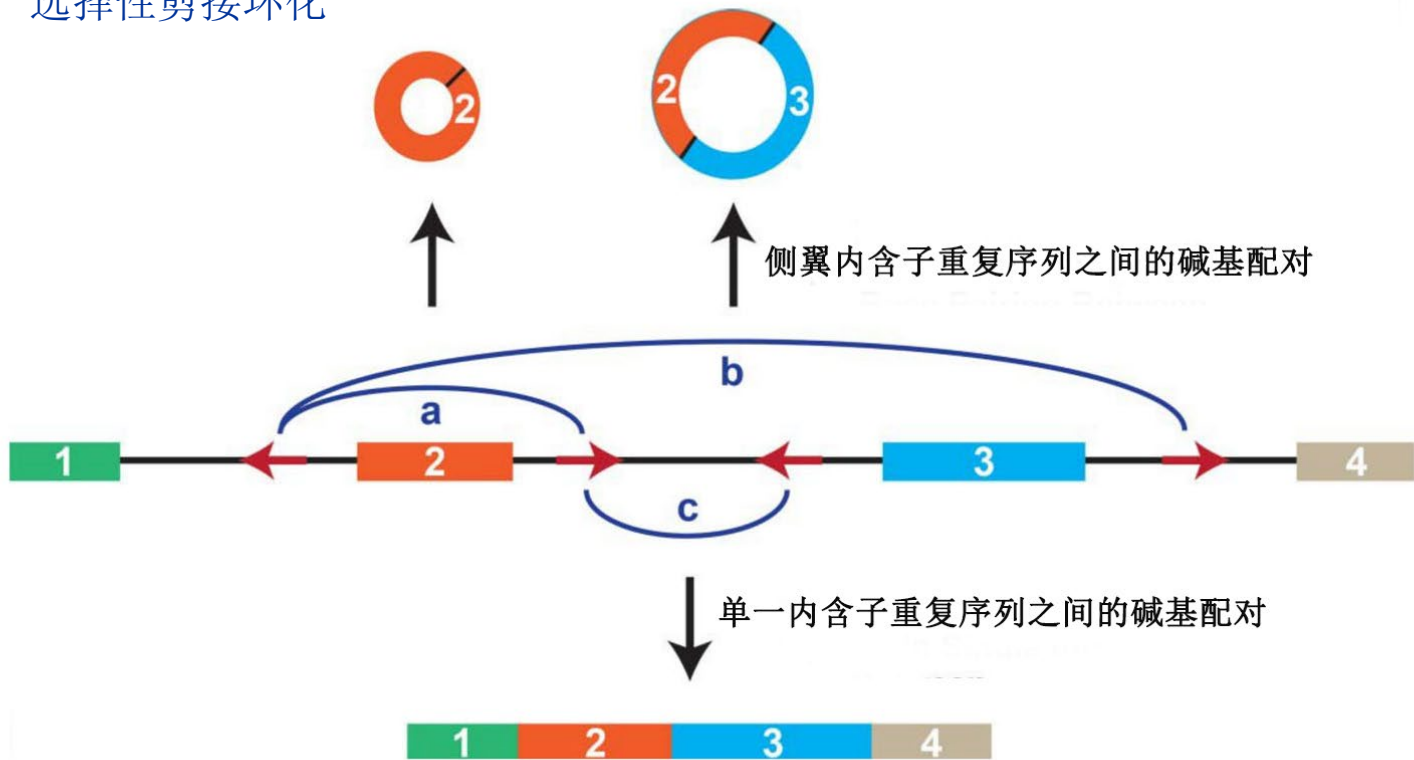
- a. 套索驱动环化 (ariat-driven circularization)、
- b. 内含子配对驱动环化 (intron- pairing-driven circularization)、
- c. 环化内含子RNAs (circular intronic RNAs)、
- d. RBP 或反式因子驱动环化 (RNA-binding proteins or trans-factor driven circularization)
- e. 选择性剪接环化 (circularization by alternative splicing) 等方式



下合成双



选择性剪接环化

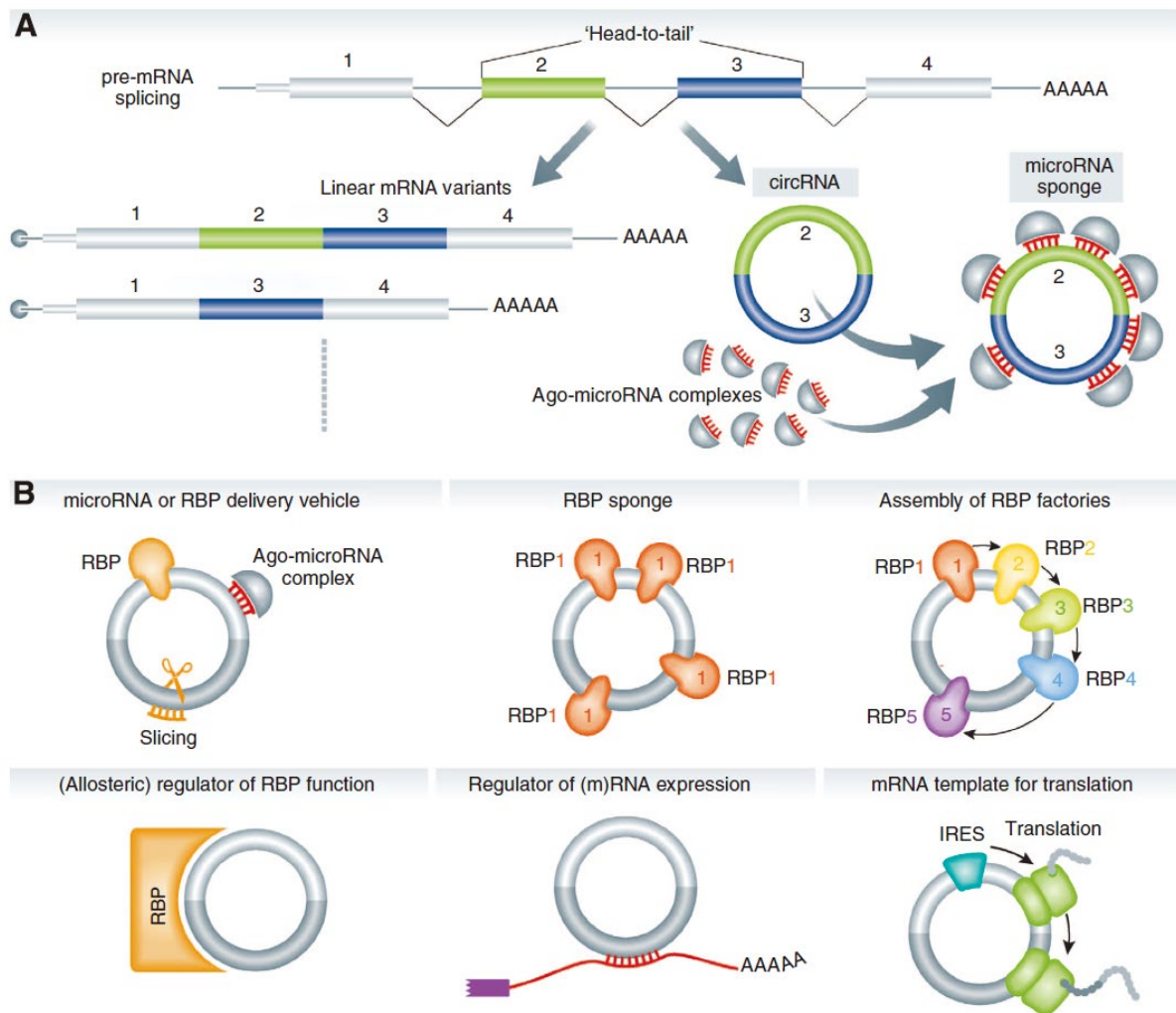




circRNAs在不同物种中具有保守性，并且存在组织表达差异，circRNA的环状结构对核酸酶不敏感，因此比线性RNA更为稳定。

部分circRNA分子含miRNA应答元件(miRNA response element, MRE)，可充当竞争性内源RNA (competing endogenous RNA, ceRNA)，与miRNA结合，在细胞中起到miRNA海绵 (miRNA sponge) 的作用，进而解除miRNA对其靶基因的抑制作用，上调靶基因mRNA的表达水平；有些还可以翻译成蛋白质，但大部分是非编码RNA。

Hentze 等 (2013) 提出 **circRNA** 的多种功能, 可以作为 miRNA 海绵或 RBP 的运载工具、RBP 海绵 (RBP sponge)、装配 RBP 的工厂、RBP 功能 (变构) 调节子、(m)RNA 表达的调节子以及作为 mRNA 模板被翻译。





14.7.4 ceRNA假说——

miRNA、circRNA、lincRNA以及mRNA的联系

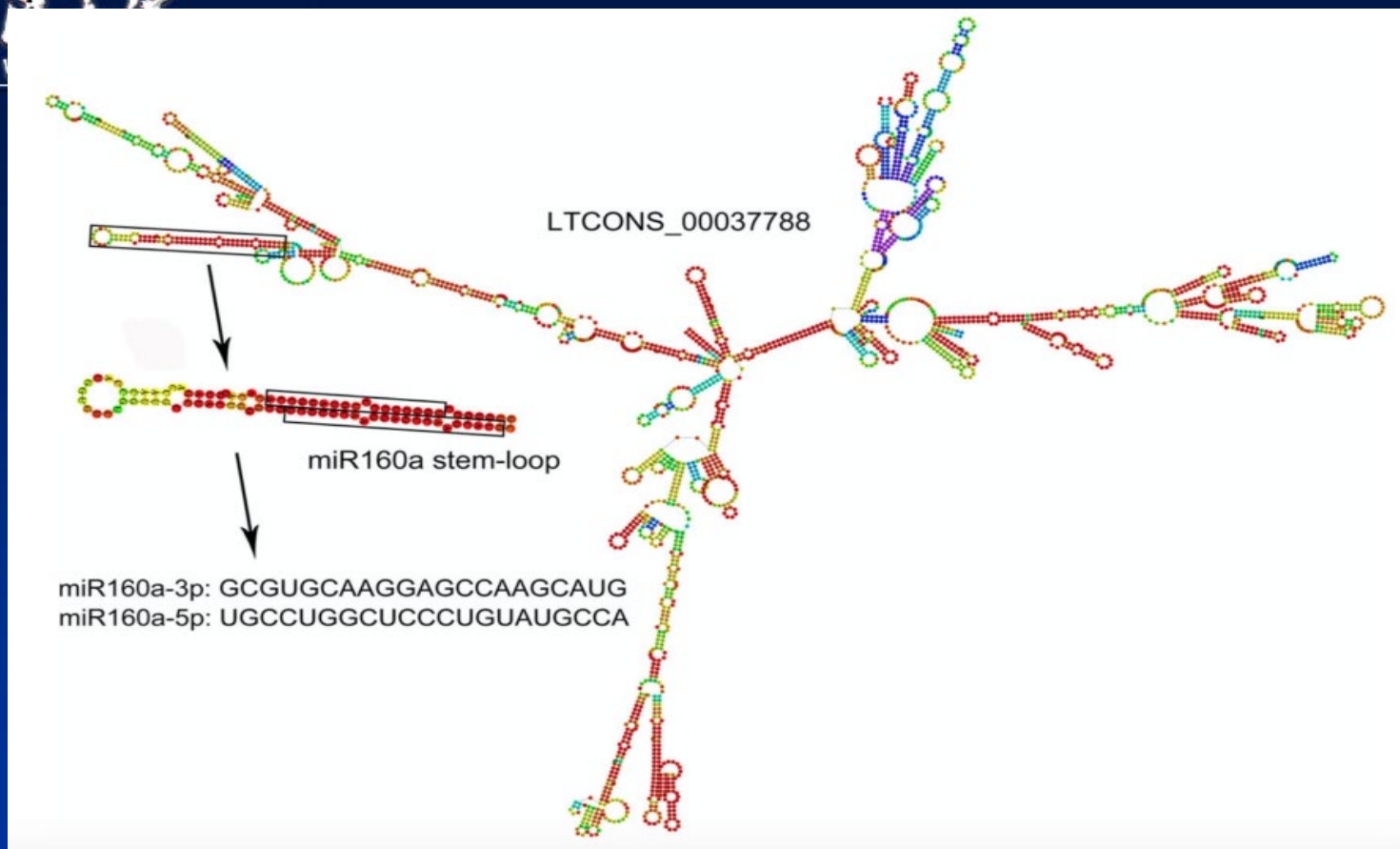
Salmena 和Pandolfi等人于2011年提出了“ceRNA假说（A ceRNA hypothesis）”，揭示出一种RNA间相互作用的新机制，认为细胞内存在ceRNA（包括mRNA、lincRNA、circRNA、miRNA、假基因等），RNA转录本通过ceRNA语言（碱基配对）进行通讯来相互交流，ceRNA分子能够通过miRNA应答元件竞争结合相应的miRNA以达到调节彼此的表达水平，ceRNA活性能形成一种大规模的转录调控网络。



根据“ceRNA假说”以及一系列的实验报道，证实miRNA、circRNA、lincRNA以及mRNA等RNA分子在细胞内通过复杂的网络联系起来，调节基因表达、转录和翻译等各种遗传信息的传递与表达，调控生物的生长发育等各种生命现象。

(1) lincRNA作为miRNA等微小RNA的前体

lincRNA作为蛋白编码基因表达的调控因子，可以作为微小RNA（例如miRNA和siRNA）的前体，通过加工剪切产生较短的微小RNA以发挥功能。



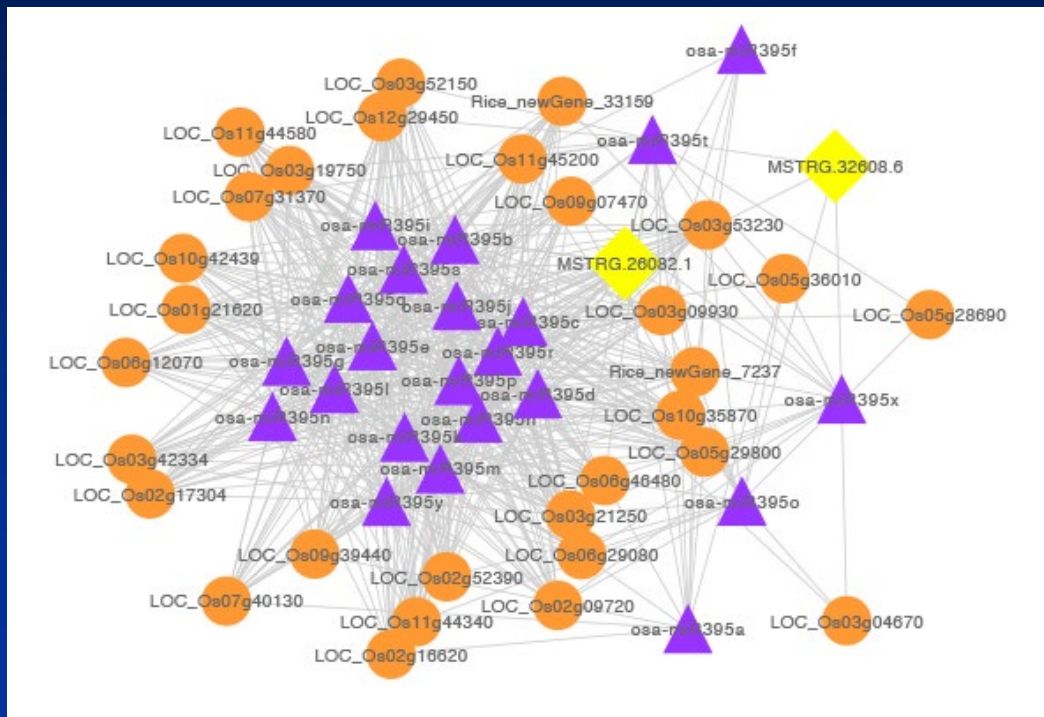
例如，水稻lncRNA LTCONS_00037788作为潜在的miRNA前体分子，在核酸内切酶的作用下剪切形成miR160a（图14-@15）。这些最初形成的miR160a均为非成熟的miRNA，需进一步加工生成成熟的miRNA才具有调控靶基因的功能。



(2) lncRNA作为miRNA的伪靶标调控mRNA基因表达

lncRNA也可以通过ceRNA网络竞争性地与miRNA结合，从而保护相应的mRNA靶基因免受miRNA介导的抑制或降解作用，这类lncRNA也被称作miRNA的内源性伪靶标 (endogenous target mimics, eTMs)。

图中所示为水稻武香S的两种lncRNA、22种miRNA和30种mRNA在育性转化过程中，所形成的一个小的lncRNA-miRNA-mRNA调控网络。

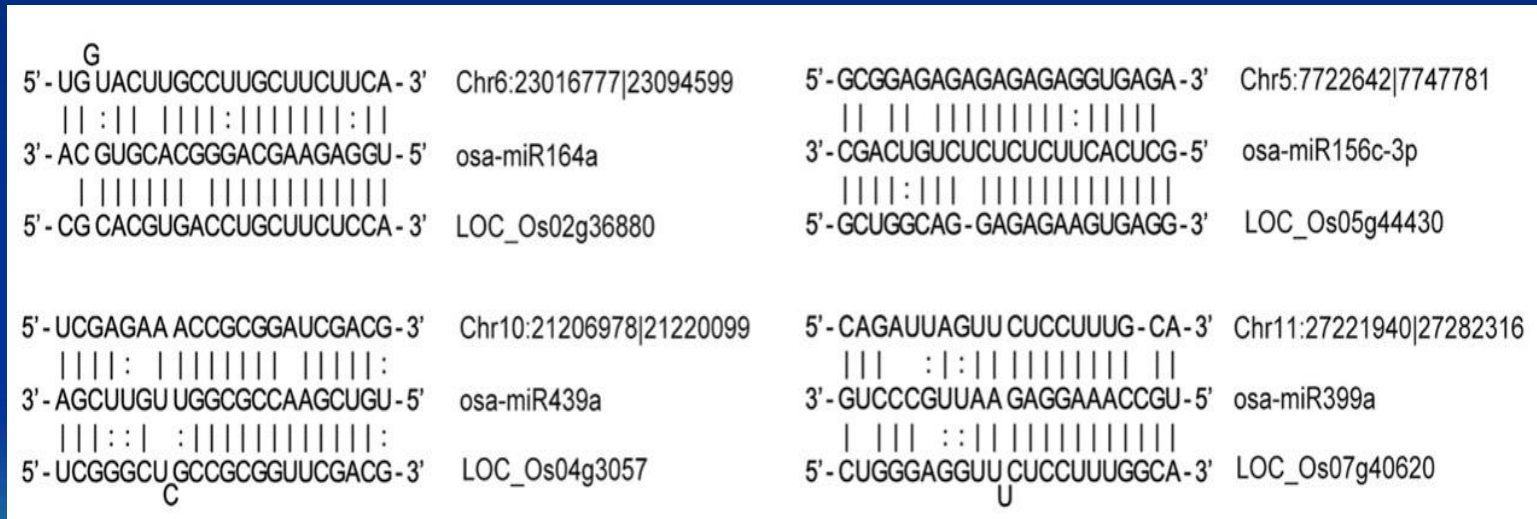


lncRNA-miRNA-mRNA互作网络关系图

(菱形：lncRNA；三角形：miRNA；圆形：mRNA)

(3) circRNA作为miRNA海绵调控mRNA基因表达

如前所述，充当miRNA海绵是circRNA较为普遍的功能，即circRNA可以作为ceRNA发挥海绵作用，通过“诱饵”的方式吸附一些特定的miRNA，以减少miRNA与其mRNA靶基因的结合，从而解除miRNA对mRNA表达的抑制作用，形成基于序列互补的circRNA-miRNA-mRNA互作调控网络。



水稻武香S在育性转换过程中，4种circRNA与miRNA和mRNA也能形成一种互作关系，如上图，以调节育性的转换。



武汉大学

Wuhan University

谢谢！

本章结束

