

基因表达调控及表观遗传



刘贇

yliu39@fudan.edu.cn

<https://liulab.fudan.edu.cn/>

基因表达调控为生物体生长、发育所必需

(一) 以适应环境、维持生长和增殖

生物体所处的内、外环境是在不断变化的。通过一定的程序调控基因的表达，可使生物体表达出合适的蛋白质分子，以便更好地适应环境，维持其生长和增殖。

基因表达调控为生物体生长、发育所必需

(二) 以维持细胞分化与个体发育

在多细胞个体生长、发育的不同阶段，或同一生长发育阶段，不同组织器官内蛋白质分子分布、种类和含量存在很大差异，这些差异是调节细胞表型的关键。

第一节

基因表达与基因表达调控

Gene Expression and Gene Expression Regulation

基因组 (genome)

来自一个生物体的一整套遗传物质。

基因表达 (gene expression)

遗传信息以各种生命现象表现出来的过程，包括基因转录及翻译。

基因表达调控 (regulation of gene expression)

细胞或生物体在接受内外环境信号刺激或适应环境变化的过程中，基因表达水平的改变方式及其过程。

一、基因表达及其特点

(一) 基因表达的时间特异性

- ▶ 按功能需要，某一特定基因的表达严格按特定的时间顺序发生，称之为基因表达的**时间特异性(temporal specificity)**。
- ▶ 多细胞生物基因表达的时间特异性又称**阶段特异性(stage specificity)**。

一、基因表达及其特点

(二) 基因表达的空间特异性

- ▶ 在个体生长全过程，某种基因产物在个体按不同组织空间顺序出现，称之为基因表达的**空间特异性(spatial specificity)**。
- ▶ 基因表达所表现出的这种分布差异，实际上是由细胞在器官的分布决定的，所以空间特异性又称**细胞或组织特异性(cell or tissue specificity)**。

一、基因表达及其特点

(三) 基因表达的持续性

➤ 某些基因在一个个体的几乎所有细胞中持续表达，不易受环境条件影响的基因，通常被称为**管家基因 (housekeeping gene)**。

➤ 管家基因较少受环境因素影响，在正常生理条件下基本不变或变化很小。区别于其他基因，这类基因表达被视为**组成性基因表达 (constitutive gene expression)**。

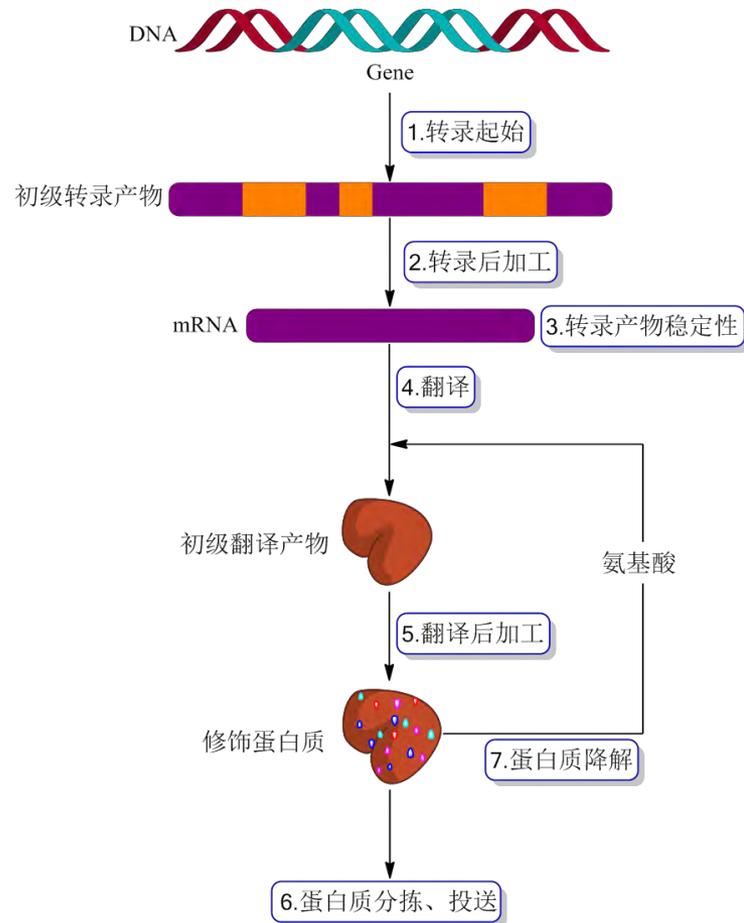
一、基因表达及其特点

(四) 基因表达的可诱导性

- 在特定环境信号刺激下，相应的基因被激活，基因表达产物增加，这种基因称为**可诱导基因(inducible gene)**。如果基因对环境信号应答是被抑制，这种基因是**可阻遏基因(repressible gene)**。
- 可诱导基因在特定环境中表达增强的过程，称为**诱导(induction)**。
- 可阻遏基因表达产物水平降低的过程称为**阻遏(repression)**。

二、基因表达调控

(一) 基因表达调控的多层次



二、基因表达调控

(一) 基因表达调控的多层次

基因激活

拷贝数

重排

DNA 甲基化

转录起始

转录后加工

mRNA降解

蛋白质翻译

翻译后加工修饰

蛋白质降解等

二、基因表达调控

(二) 基因表达调控的协调性

- 在一定机制控制下，功能上相关的一组基因，无论其为何种表达方式，均需协调一致、共同表达，即为**协同调节**(coordinate regulation)。

二、基因表达调控

(三) 基因表达调控的主要方式

顺式作用元件(cis-acting element)

存在于基因旁侧序列中能影响基因表达的序列。顺式作用元件包括启动子、增强子、调控序列和可诱导元件等，它们的作用是参与基因表达的调控。顺式作用元件本身不编码任何蛋白质，仅提供一个作用位点，要与反式作用因子相互作用而起作用。

反式作用因子(trans-acting factor)

反式作用因子是转录模板上游基因编码的一类蛋白调节因子，包括激活因子和阻遏因子等，它们与顺式作用元件中的上游激活序列特异性结合，对基因的转录分别起促进和阻遏作用。

分类：

- ①通用转录因子。普遍存在的转录因子，具有识别启动子元件功能的基本转录因子。
- ②组织特异性转录因子：与基因表达的组织特异性有很大关系，能识别增强子或沉默子的转录调节因子。
- ③诱导性反式作用因子：活性能被特异的诱导因子所诱导，不需要通过DNA-蛋白质相互作用就参与转录调控。

二、基因表达调控

➤有些反式作用因子具有特殊的DNA结合结构域(DNA-binding domain)能特异识别某些DNA序列并与顺式作用元件相互作用。如DNA双螺旋的大沟是最常见的调节蛋白与DNA序列发生相互作用的部位。

二、基因表达调控

(四) 基因表达调控的生理意义

➤ 适应环境

➤ 维持生长

➤ 细胞增殖

第二节

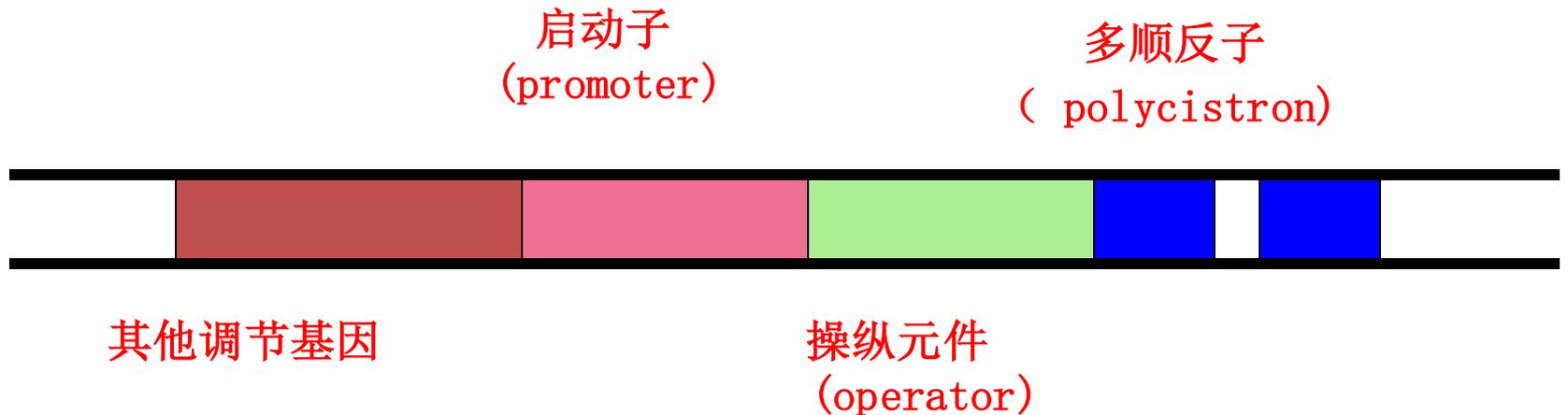
原核基因表达调控

Regulation of Gene Expression in Prokaryotes

一、原核基因转录调节特点

原核生物

—— 操纵子(operon) 机制



一、原核基因转录调节特点

1、启动子

是RNA pol和各种调控蛋白作用的部位，是决定基因表达效率的关键元件。

	-35区		-10区		RNA转录起始
trp	TTGACA	N17	TTAACT	N7	A
tRNATyr	TTTACA	N16	TATGAT	N7	A
lac	TTTACA	N17	TATGTT	N6	A
recA	TTGATA	N16	TATAAT	N7	A
Ara BAD	CTGACG	N16	TACTGT	N6	A
共有序列	TTGACA		TATAAT (pribnow box)		

图13-1 五种*E.coli*启动序列的共有序列

一、原核基因转录调节特点

- 共有序列 (consensus sequence) 决定启动序列的转录活性大小 (与保守序列之间的差异)。
- 某些特异因子 (蛋白质) 决定RNA聚合酶对一个或一套启动序列的特异性识别和结合能力。
- Pribnow框功能: (1) RNA pol (RNA 聚合酶) 紧密结合; (2) 形成开放启动复合体; (3) 使RNA pol定向转录。
- -35区 (TTGACA) 功能: 为RNA聚合酶对转录起始的识别序列

一、原核基因转录调节特点

2、操纵元件

——阻遏蛋白(repressor)的识别和结合位点

当操纵序列结合阻遏蛋白时，会阻碍RNA聚合酶与启动序列的结合，或是RNA聚合酶不能沿DNA向前移动，阻碍转录，介导负性调节(negative regulation)。



一、原核基因转录调节特点

- 特异DNA序列可结合激活蛋白(activator)：可结合启动序列邻近的DNA序列，促进RNA聚合酶与启动序列的结合，增强RNA聚合酶活性，介导正性调节(positive regulation)。
- 有些基因在没有激活蛋白存在时，RNA聚合酶很少或完全不能结合启动序列。

一、原核基因转录调节特点

调节基因(regulation gene)编码能够与操纵序列结合的调控蛋白，可以分为3类：

➤特异因子：决定RNA pol对一个或一套启动序列的特异性识别和结合能力。特异因子与RNA聚合酶形成全酶后，可以改变原来RNA聚合酶识别启动序列的特异性。RNA聚合酶是笔杆，那么特异因子就是笔芯，原核生物就是这样通过不断换笔芯来实现不同基因的表达。

➤阻遏蛋白：识别、结合特异操纵序列，抑制基因转录，介导负性调节。

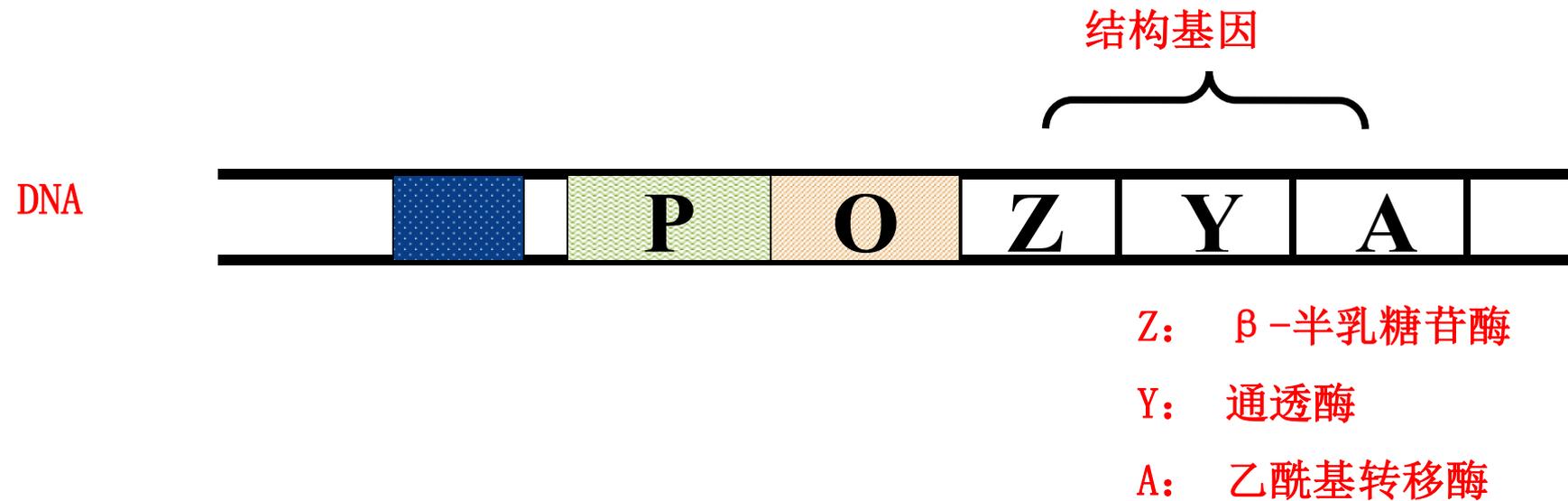
➤激活蛋白：结合启动序列邻近的DNA序列，提高RNA pol与启动序列的结合能力，从而增强RNA pol的转录活性，是一种正调控。分解（代谢）物基因激活蛋白(catabolite gene activator protein, CAP就是一种典型的激活蛋白。

一、原核基因转录调节特点

- 凡是能够诱导基因表达的分子称为诱导剂。
- 凡是能够阻遏基因表达的分子称为阻遏剂。
- 诱导和阻遏是原核生物转录调控的基本方式。

二、乳糖操纵子调节机制

(一) 乳糖操纵子 (lac operon)



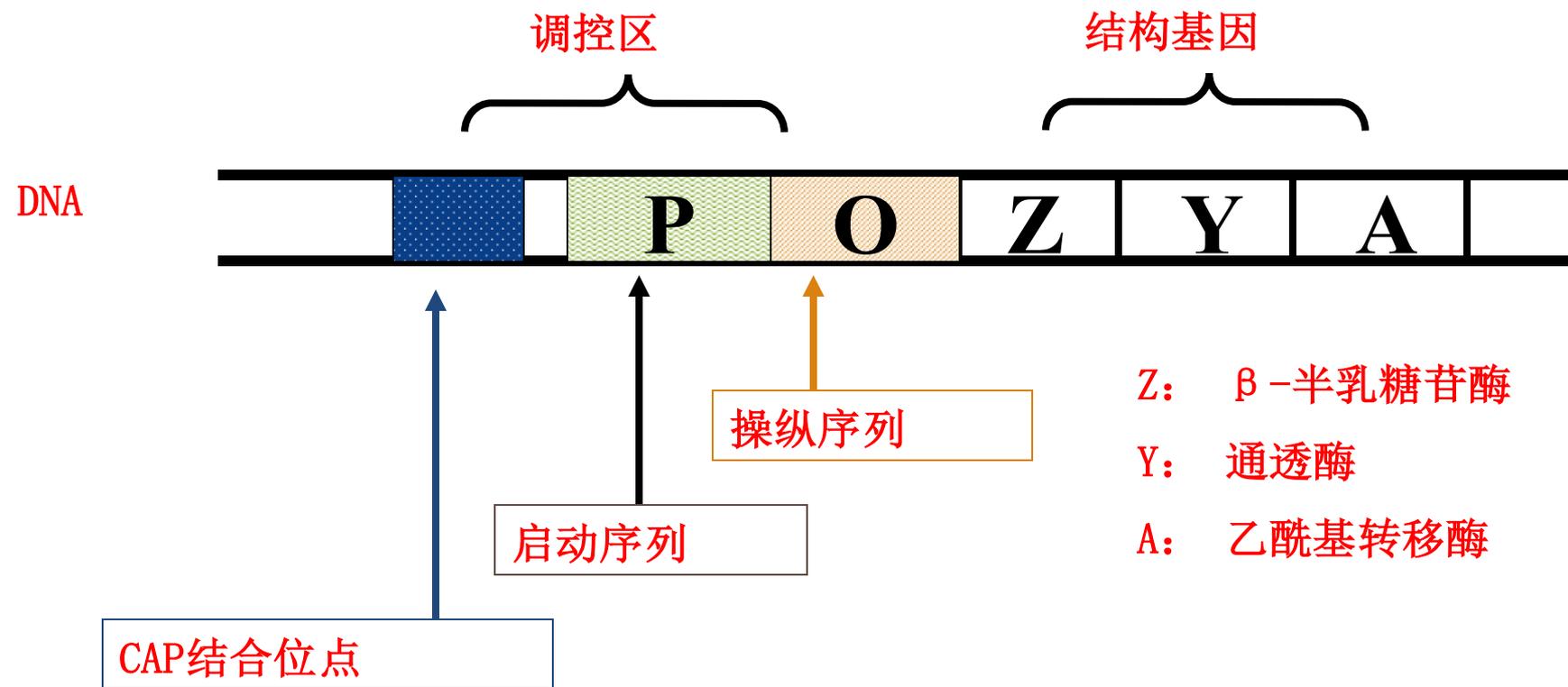
乳糖操纵子编码将乳糖运输到细胞中所需的蛋白质，使细胞能够在没有葡萄糖的情况下利用乳糖。

乳糖操纵子需要整合两个信号，只有在下列情况下操纵子才会高表达：

- 必须有乳糖
- 必须没有葡萄糖

二、乳糖操纵子调节机制

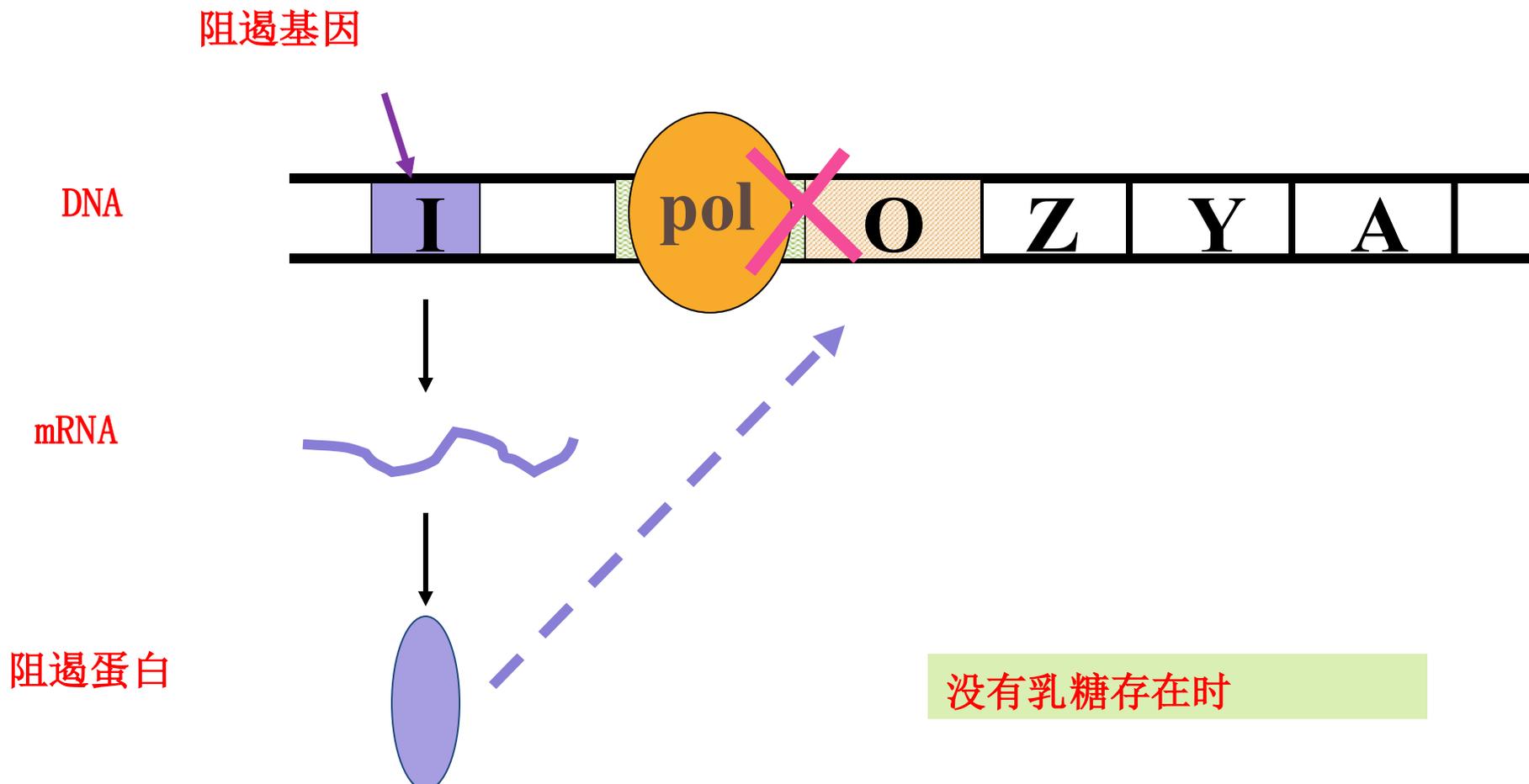
(一) 乳糖操纵子(lac operon)的结构



二、乳糖操纵子调节机制

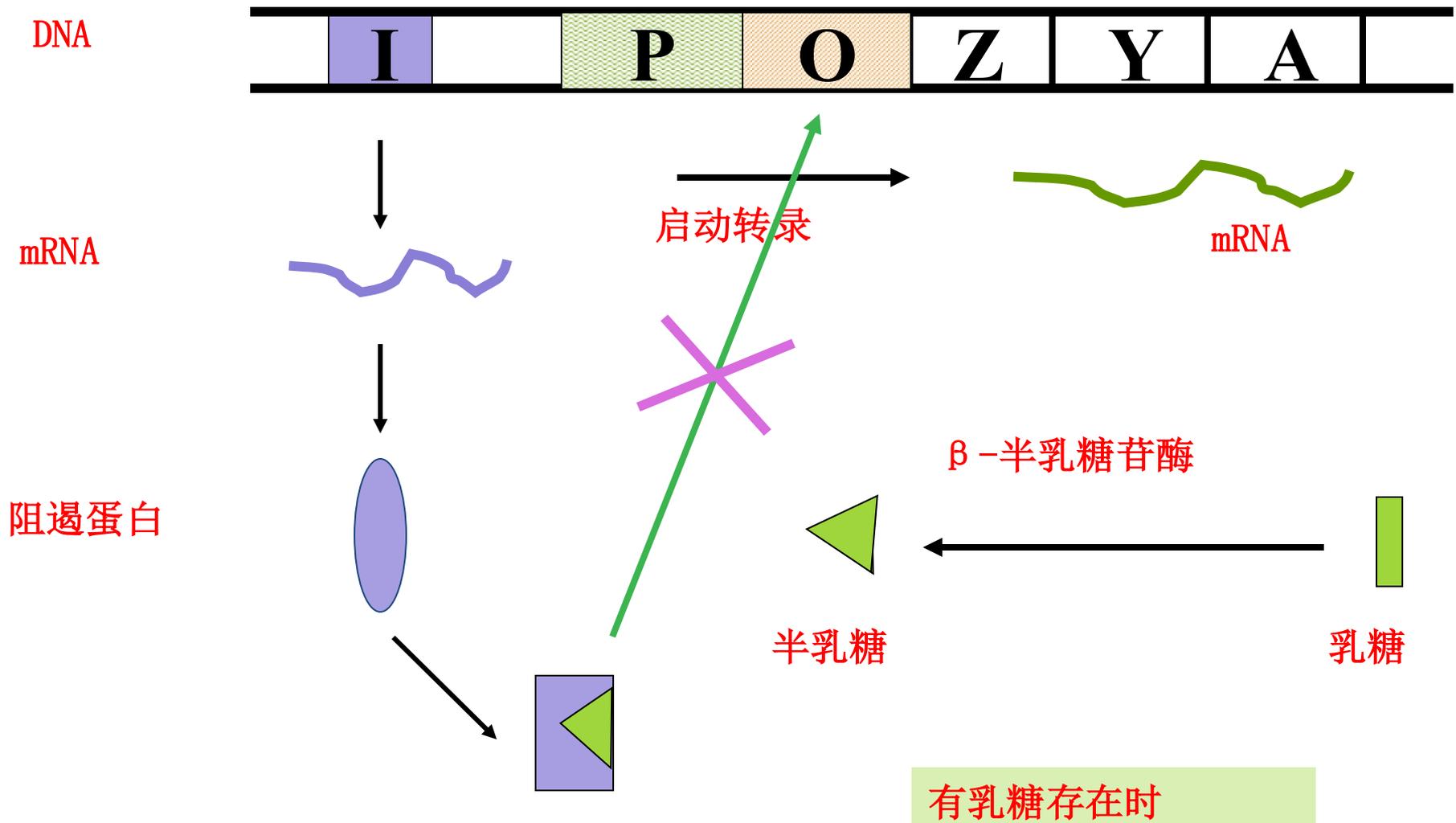
(一) 乳糖操纵子 (lac operon) 的结构

1、阻遏蛋白的负调节



二、乳糖操纵子调节机制

(一) 乳糖操纵子(lac operon)的结构

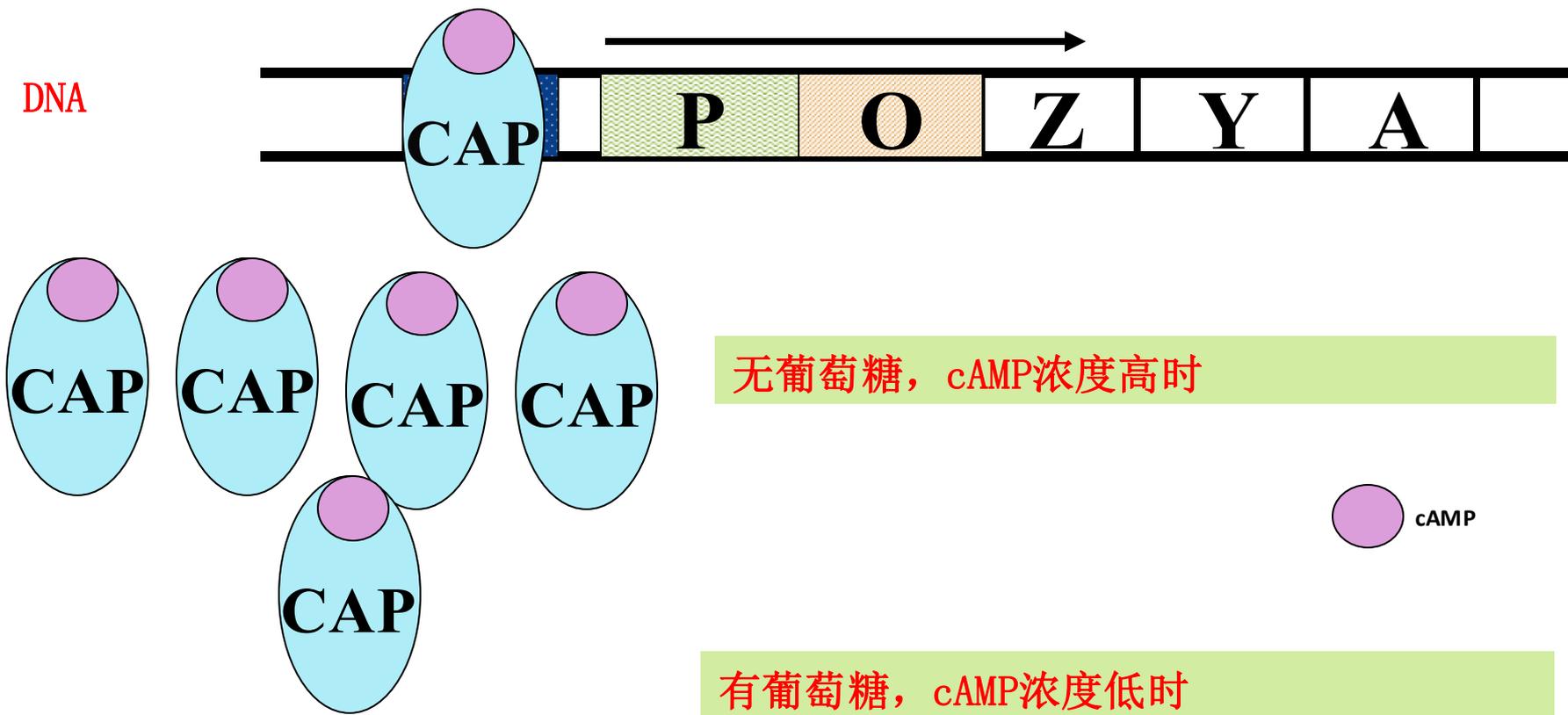


二、乳糖操纵子调节机制

(一) 乳糖操纵子 (lac operon) 的结构

2、CAP的正性调节

++++ 转录



二、乳糖操纵子调节机制

(一) 乳糖操纵子(lac operon)的结构

3、协同调节

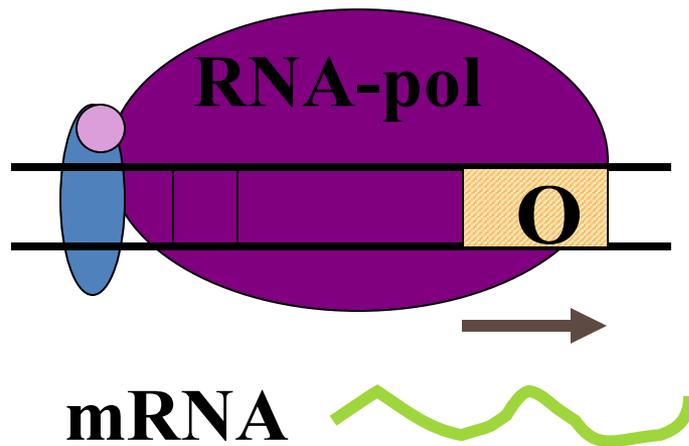
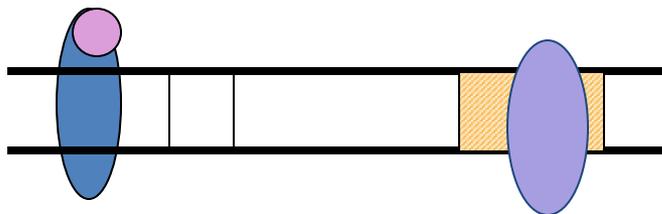
- 当阻遏蛋白封闭转录时，CAP对该系统不能发挥作用。
- 如无CAP存在，即使没有阻遏蛋白与操纵序列结合，操纵子仍无转录活性。
- 单纯乳糖存在时，细菌利用乳糖作碳源；若有葡萄糖或葡萄糖/乳糖共同存在时，细菌首先利用葡萄糖。葡萄糖对 lac 操纵子的阻遏作用称分解代谢阻遏(catabolic repression)。

二、乳糖操纵子调节机制

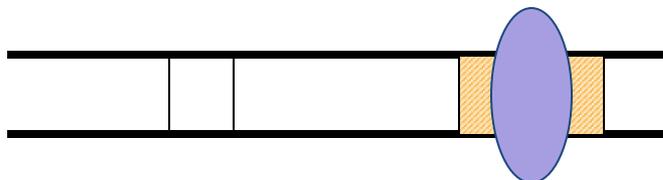
低半乳糖时

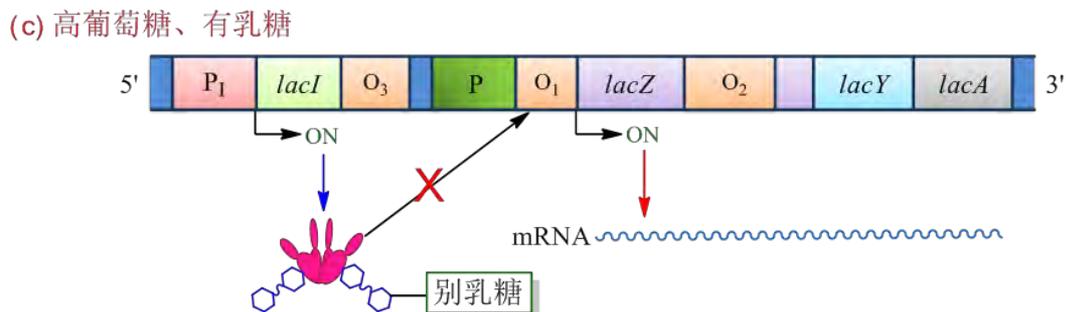
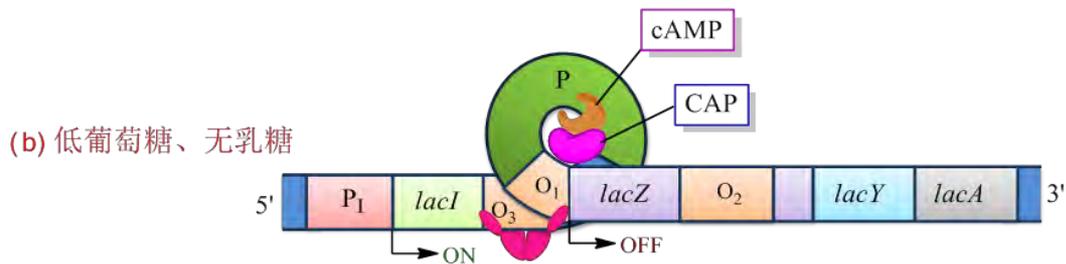
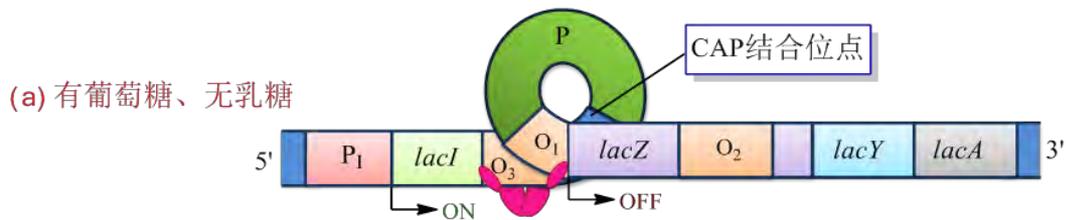
高半乳糖时

葡萄糖低
cAMP浓度高

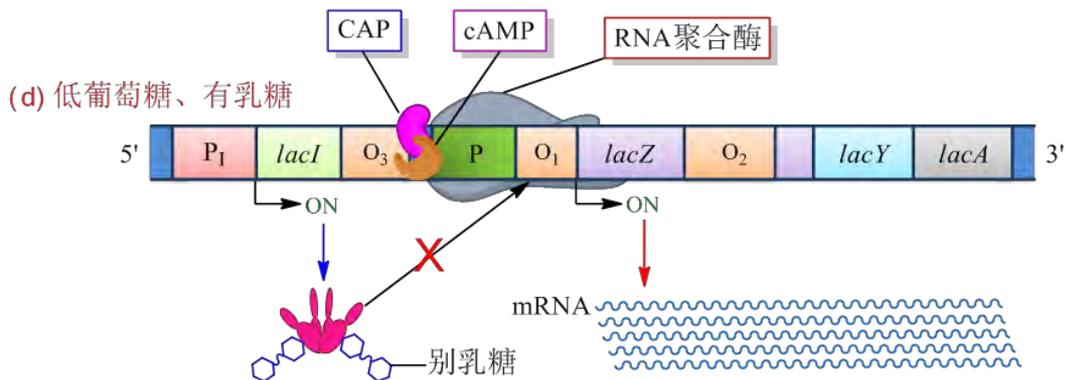


葡萄糖高
cAMP浓度低





乳糖操纵子的协同调节



二、乳糖操纵子调节机制

乳糖操纵子的协同调节

(a) 有葡萄糖、无乳糖，阻遏蛋白封闭转录，cAMP低浓度，CAP不能发挥作用，基因关闭； (b) 低葡萄糖、无乳糖，阻遏蛋白封闭转录，cAMP高浓度，CAP能发挥作用，基因依然关闭； (c) 高葡萄糖、有乳糖，去阻遏，但因有葡萄糖存在，cAMP低浓度，CAP不能发挥作用，基因低水平表达，催化乳糖转化为别乳糖，解除阻遏； (d) 无葡萄糖、有乳糖，既去阻遏，又因cAMP高浓度，CAP发挥作用，基因高水平表达

三、其他转录调节机制

大肠杆菌在转录终止阶段有两种调控终止的方式

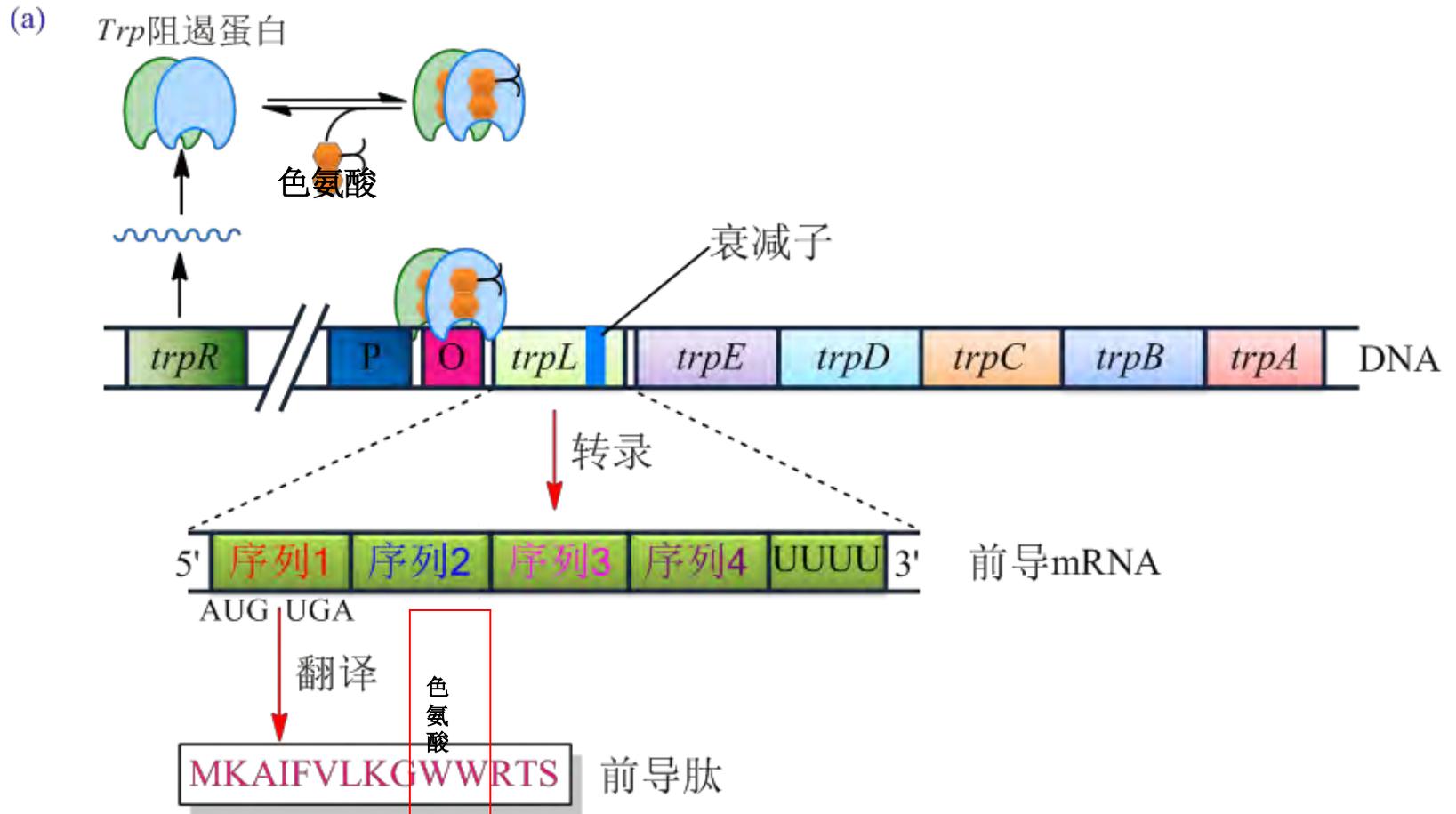
➤ **转录衰减**

RNA链在转录过程中翻译产生的特殊蛋白质与自身基因的调节序列结合而导致转录提前终止。

➤ **抗终止**

阻止前者的发生，使下游基因得以表达。

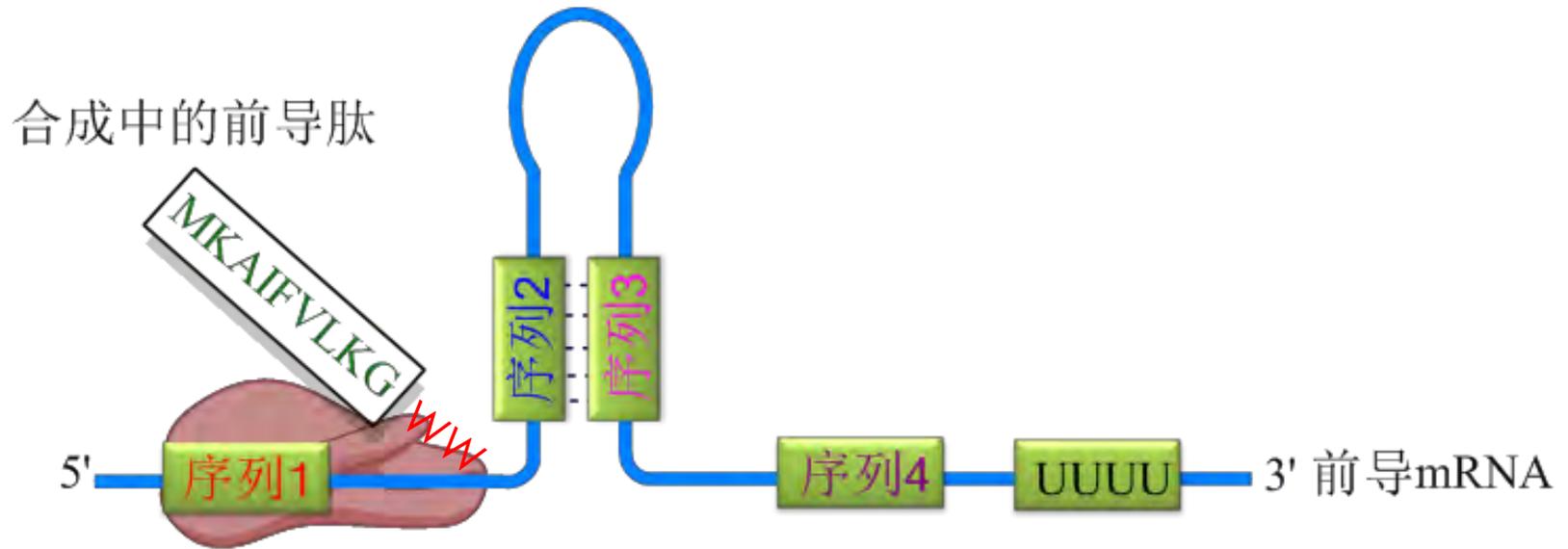
三、其他转录调节机制



色氨酸操纵子的基本结构及前导序列转录产物结构

三、其他转录调节机制

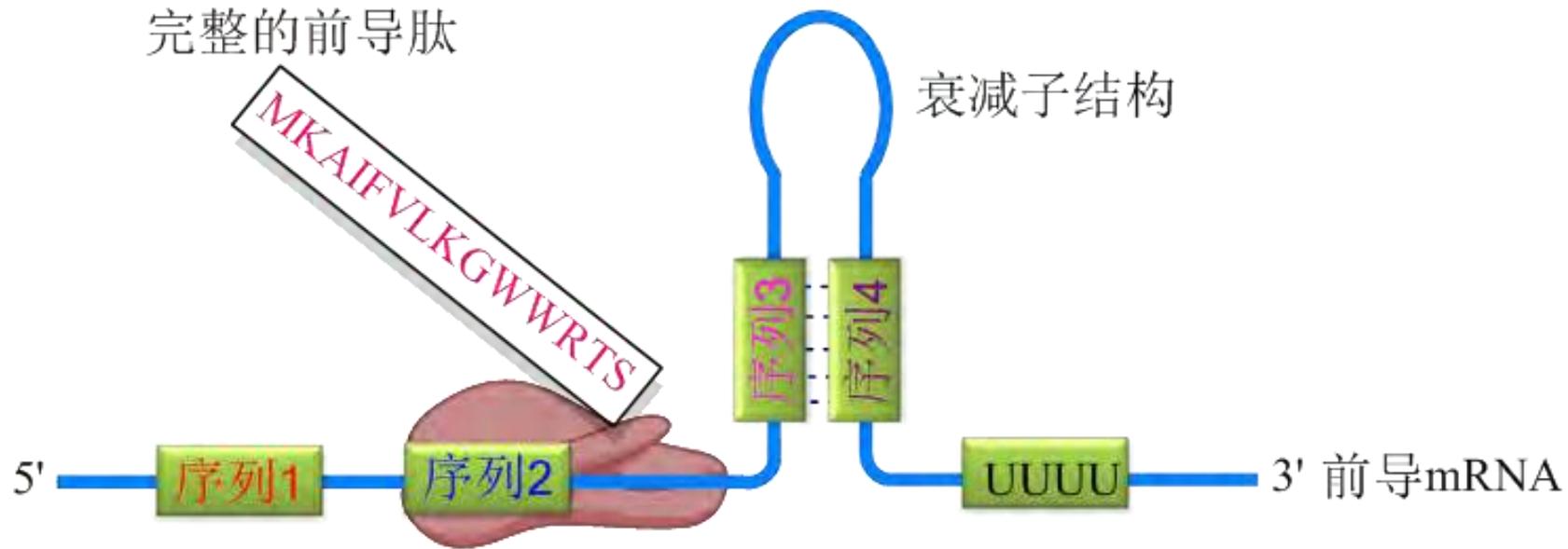
(b)



低Trp，前导肽的合成到WW(色氨酸)时就停滞了，核糖体停止在序列1上，形成2/3发夹，不影响转录

三、其他转录调节机制

(c)



高Trp，前导肽继续向前合成，完整的前导肽与3/4发夹结构作用，合并后续多聚U序列使得转录提前终止

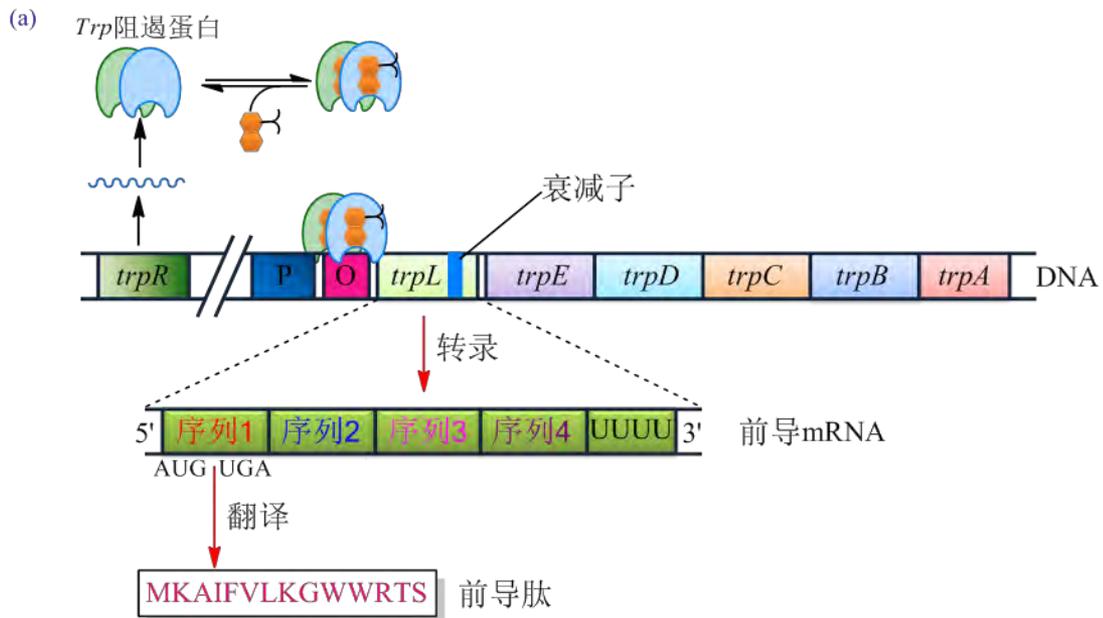
色氨酸操纵子的结构及其关闭机制

(b)

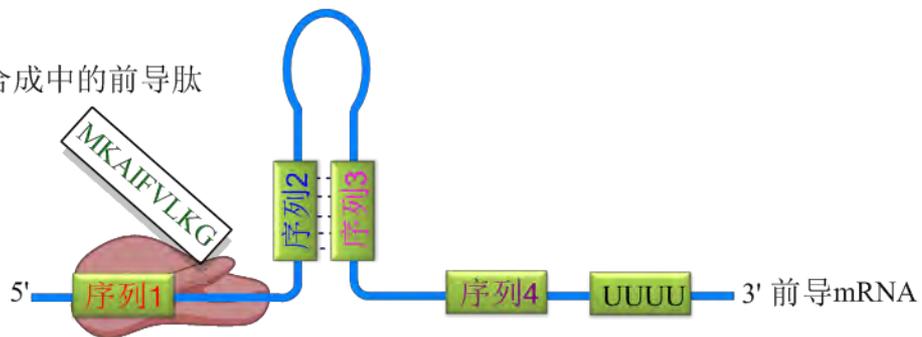
两个“开关”

- (1) 粗调开关：阻遏蛋白的调节
- (2) 细调开关：前导序列的调节

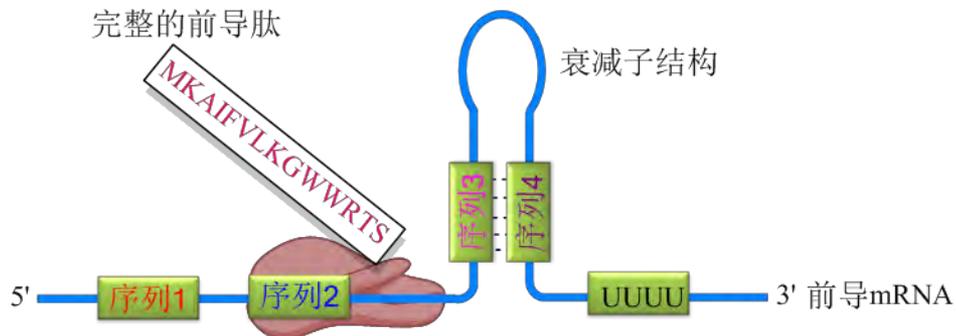
(c)



合成中的前导肽



完整的前导肽



第三节

真核基因表达调控

Regulation of Gene Expression in Eukaryotes

一、真核基因的结构特点

◆ 真核基因组结构庞大

- 哺乳类动物基因组结构庞大，DNA 约 3×10^9 碱基对。
- 人编码基因约4万个，编码序列仅占总长的1%。
- rDNA等重复基因约占5%~10%。

一、真核基因的结构特点

◆ 真核基因组含有大量的重复序列

多拷贝序列 { 高度重复序列 (10^6 次)
 { 中度重复序列 ($10^3 \sim 10^4$ 次)

单拷贝序列 (一次或数次)

◆ 真核基因转录产物为单顺反子

单顺反子(monocistron)

即一个编码基因转录生成一个mRNA分子，经翻译生成一条多肽链。

一、真核基因的结构特点

- ◆ 真核基因中存在非编码序列和间隔区，故：具有不连续性

真核结构基因两侧存在有不被转录的非编码序列，往往是基因表达的调控区。在编码基因内部尚有内含子（intron）、外显子（exon）之分，因此真核基因是不连续的。

二、转录起始的调控

真核的**每种RNA pol**都是巨大的蛋白（500kD以上），约由12个左右的亚基组成，其中有些是3种酶共有的。每个聚合酶最大的亚基彼此同源。真核生物没有 σ 因子，因此需要其它蛋白质帮助识别或结合启动子

RNA pol I 转录rRNA 基因产生45S rRNA 前体，加工成5.8S、18S 和28S rRNA

RNA pol III 的转录产物为5S rRNA、tRNA 和snRNA

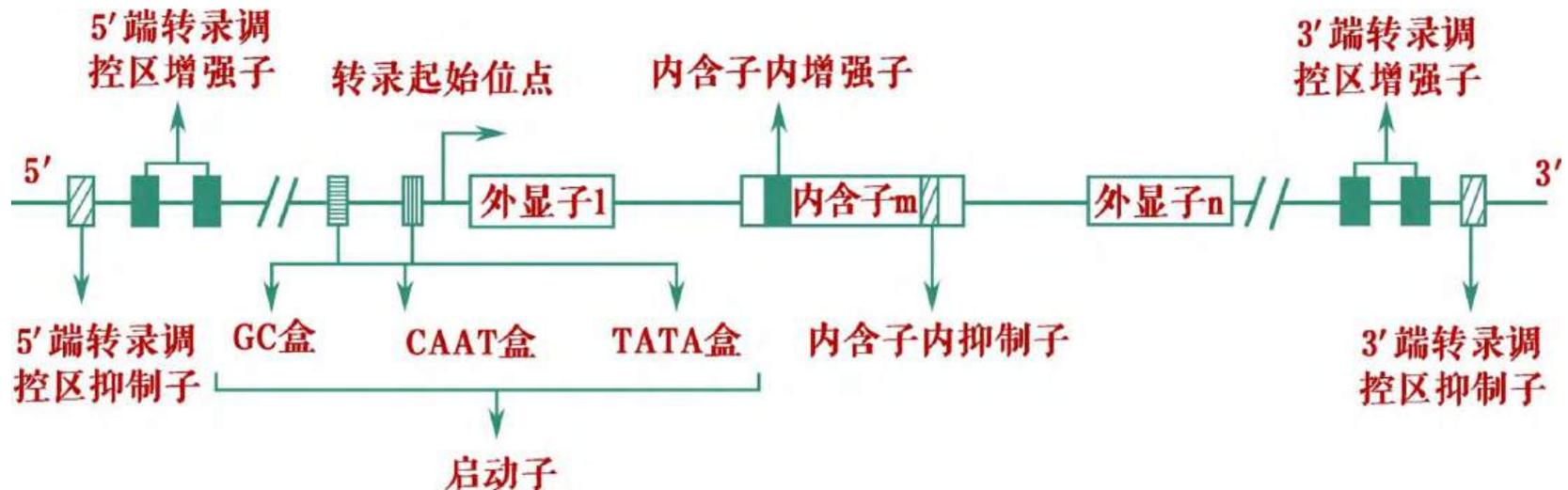
RNA pol II 转录mRNA 和一些snRNA 的基因

二、转录起始的调控

真核生物

(一) 顺式作用元件 (cis-acting element)

——可影响自身基因表达活性的DNA序列

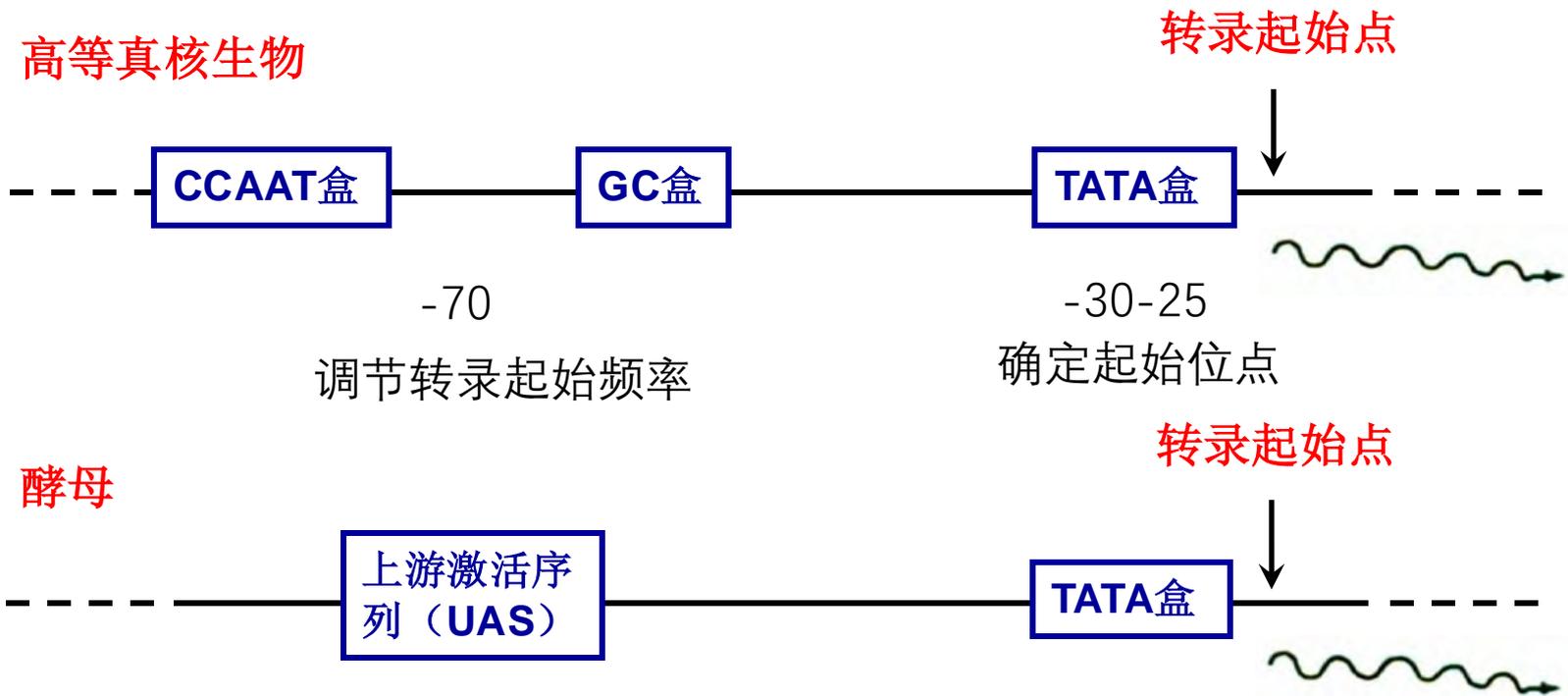


- 顺式作用元件

二、转录起始的调控

1. 启动子

真核基因启动子是RNA聚合酶结合位点周围的一组转录控制组件，至少包括一个转录起始点以及一个以上的功能组件。



真核基因启动子的典型结构

二、转录起始的调控

很多启动子不含TATA盒，这类启动子分为两类

➤一类为富含GC的启动子，最初发现于一些管家基因。这类启动子一般含数个分离的转录起始点，并有数个转录因子SP1结合位点，对基本转录活化有重要作用。

➤另一类启动子既不含TATA盒，也没有GC富含区，这类启动子可有一个或多个转录起始点，大多转录活性很低或根本没有转录活性，而是在胚胎发育、组织分化或再生过程中受调节。

二、转录起始的调控

2. 增强子(enhancer)

指远离转录起始点、决定基因的时间、空间特异性、增强启动子转录活性的DNA序列。其发挥作用的方式通常与方向、距离无关。其本身不具备启动子活性。

二、转录起始的调控

增强子有以下特点：

- ①增强子的增强转录效应十分明显，一般能使基因转录效率提高10~200倍，有的甚至可以增加上千倍。
- ②增强子发挥作用的方式，通常与其方向或与其所在部位与转录起始点的距离无关。
- ③增强子大多数为重复序列。
- ④增强子的增强转录效应有严格的组织细胞特异性。
- ⑤增强子无基因专一性，可以在不同基因的转录中发挥作用。
- ⑥增强子的活性与其在DNA双螺旋结构中的空间方向性有关。
- ⑦增强子如果受外部信号驱使发挥作用，该类增强子又被称为反应元件，如cAMP反应元件、激素反应元件、金属反应元件和血清反应元件等。

二、转录起始的调控

3. 沉默子(silencer)

某些基因的负性调节元件，当其结合特异蛋白因子时，对基因转录起阻遏作用。

*沉默子是到20世纪80年代末才被证实的一类**负性转录调控元件**。

*同一DNA元件时而表现增强子的活性，时而又表现沉默子的活性，这是由该元件的结合蛋白的性质决定的。这些负调控元件不受距离和方向的限制，并可对异源基因的表达发挥作用。

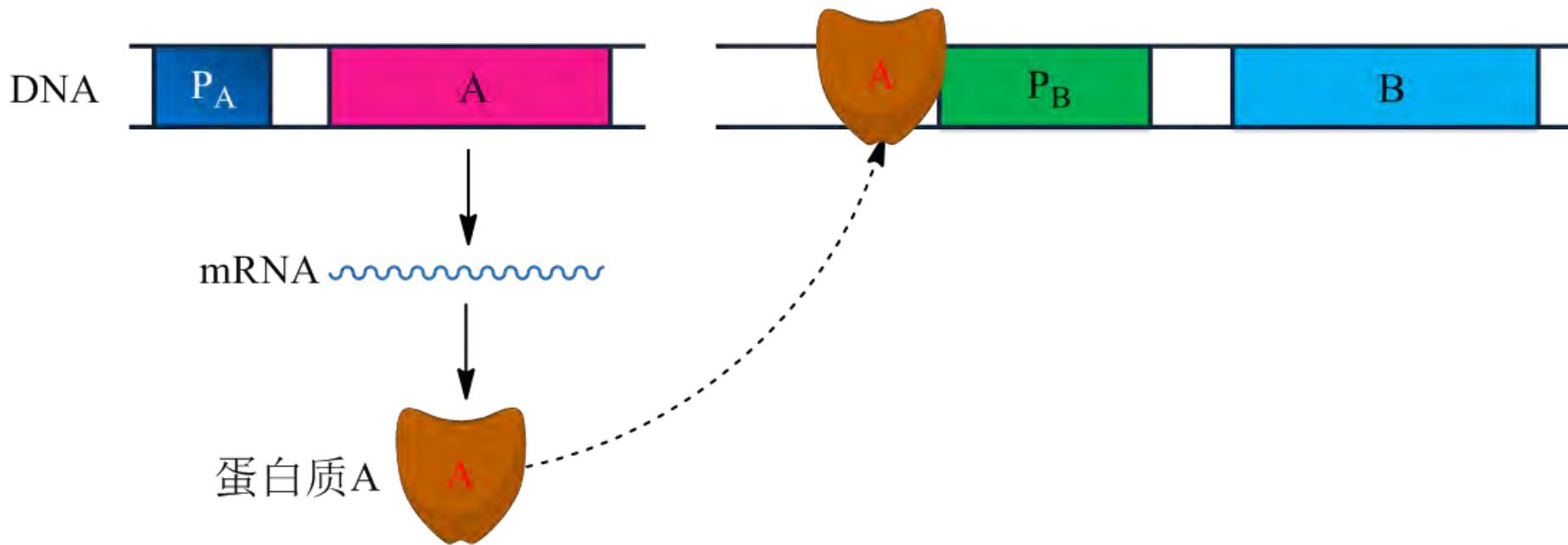
二、转录起始的调控

(二) 转录因子

- 转录因子是由特定基因表达产生的蛋白质因子，通过与特异的顺式作用元件相互作用，反式激活另一基因的转录，称为反式作用因子(trans-acting factor)。
- 还有蛋白质因子可特异识别、结合自身基因的调节序列，调节自身基因的表达，称顺式作用。具有这种调节方式的调节蛋白称为顺式作用因子。

二、转录起始的调控

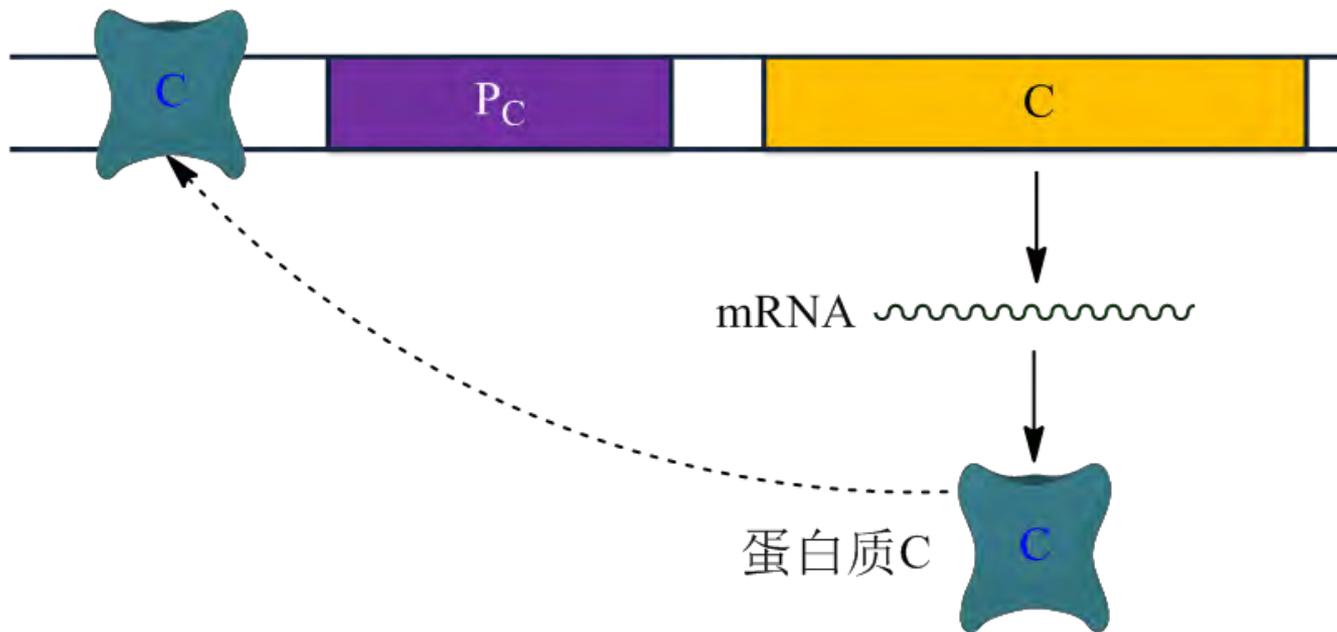
反式调节



反式调节：基因A的表达产物蛋白质A调节基因B的表达，蛋白质A为反式作用因子； P_A 和 P_B 分别为基因A和B的启动子。

二、转录起始的调控

顺式调节



顺式调节：基因C的表达产物蛋白质C调节自身基因的表达，蛋白质C为顺式作用因子； P_C ：基因C的启动子

二、转录起始的调控

■ 转录因子分类（按功能特性）

➤ 通用转录因子 (general transcription factors)

是RNA聚合酶结合启动子所必需的一组蛋白因子，决定三种RNA (mRNA、tRNA及rRNA) 转录的类别，也称基本转录因子。

二、转录起始的调控

➤ 特异转录因子 (special transcription factors)

为个别基因转录所必需，决定该基因的时间、空间特异性表达。

- 转录激活因子
- 转录抑制因子
- 转录起始过程还有上游因子 (upstream factor) 和可诱导因子 (inducible factor)。

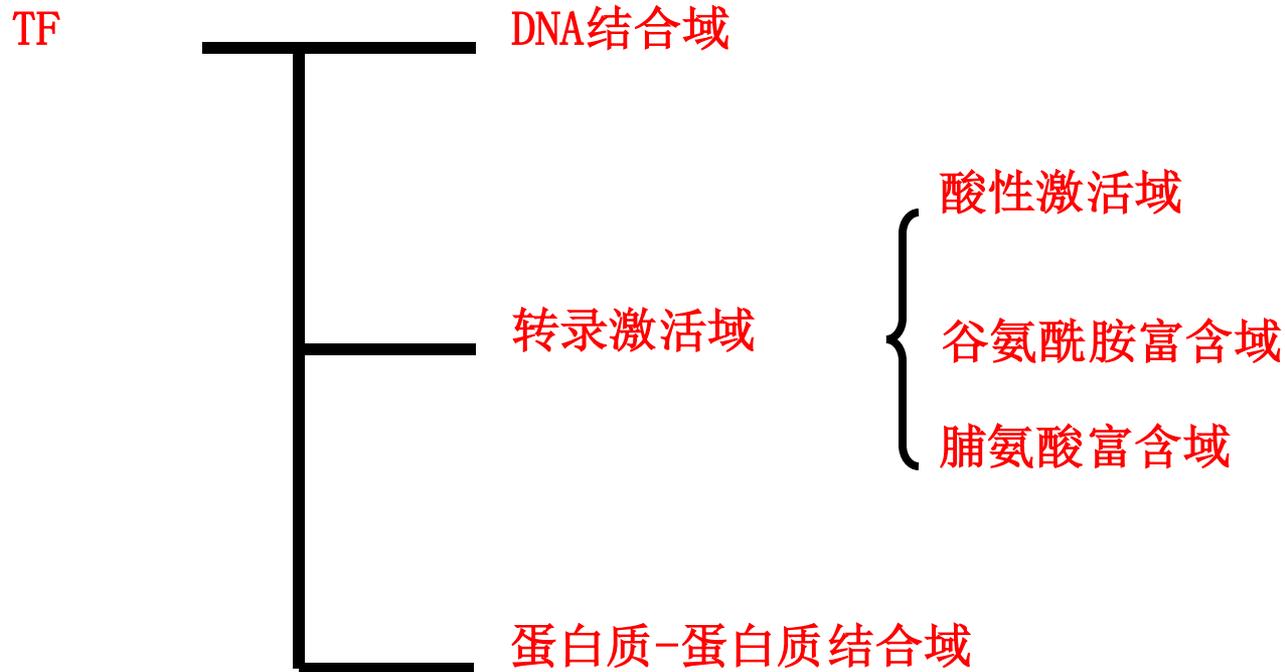
二、转录起始的调控

RNA聚合酶II启动转录需要多种蛋白质因子的协同作用。

- 可诱导因子或上游因子与增强子或启动子上游元件的结合；
- 通用转录因子在启动子处的组装；辅激活因子和（或）中介子在通用转录因子或RNA pol II复合物与可诱导因子、上游因子之间的辅助和中介作用。
- 因子与因子之间互相辨认、结合，以准确控制基因是否转录，何时转录。

二、转录起始的调控

转录因子结构



(二聚化结构域)

二、转录起始的调控

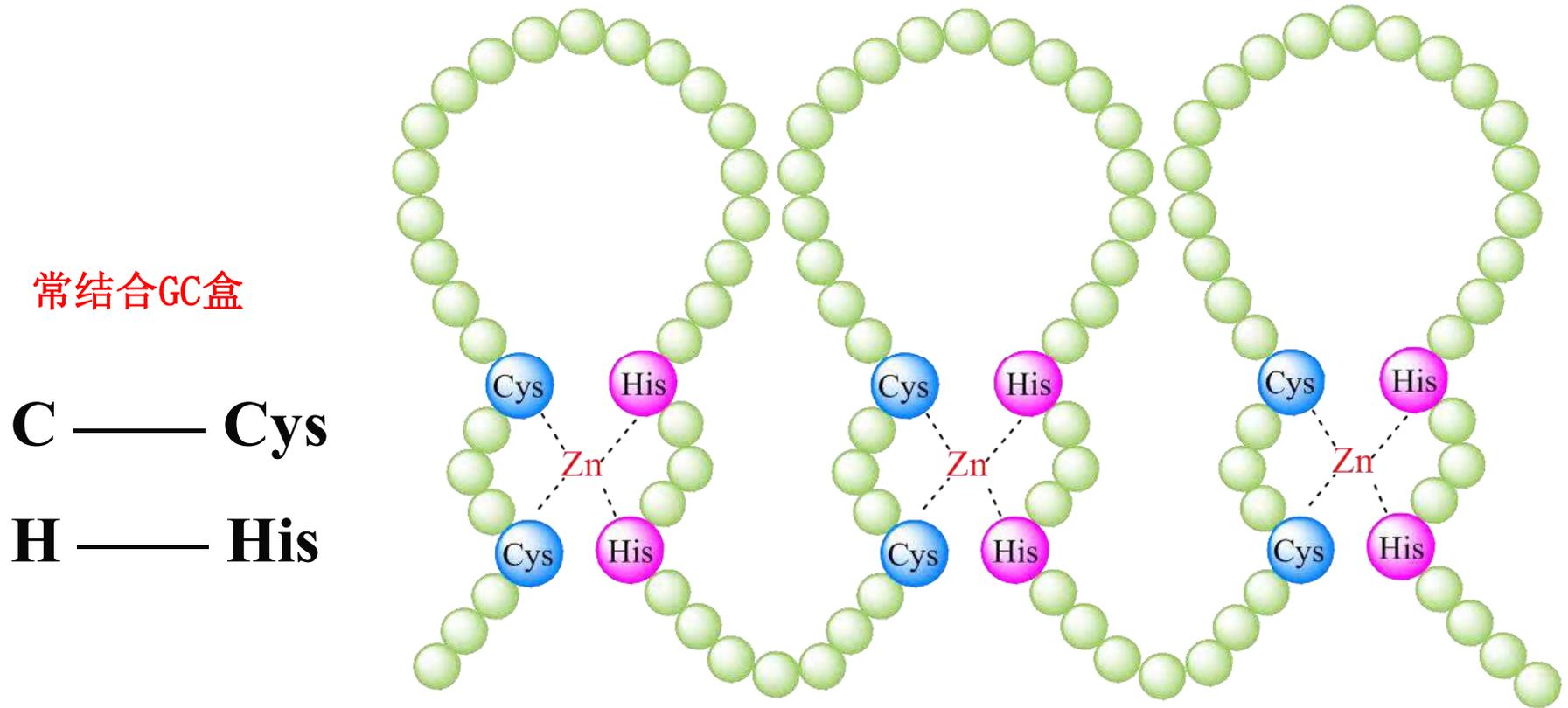
(三) 顺式作用元件与转录因子的结合

- 反式作用因子与顺式作用元件之间的特异结合主要依赖于转录因子DNA结合域中的特殊蛋白质模体。

二、转录起始的调控

最常见的DNA结合域：

1. 锌指(zinc finger)结构

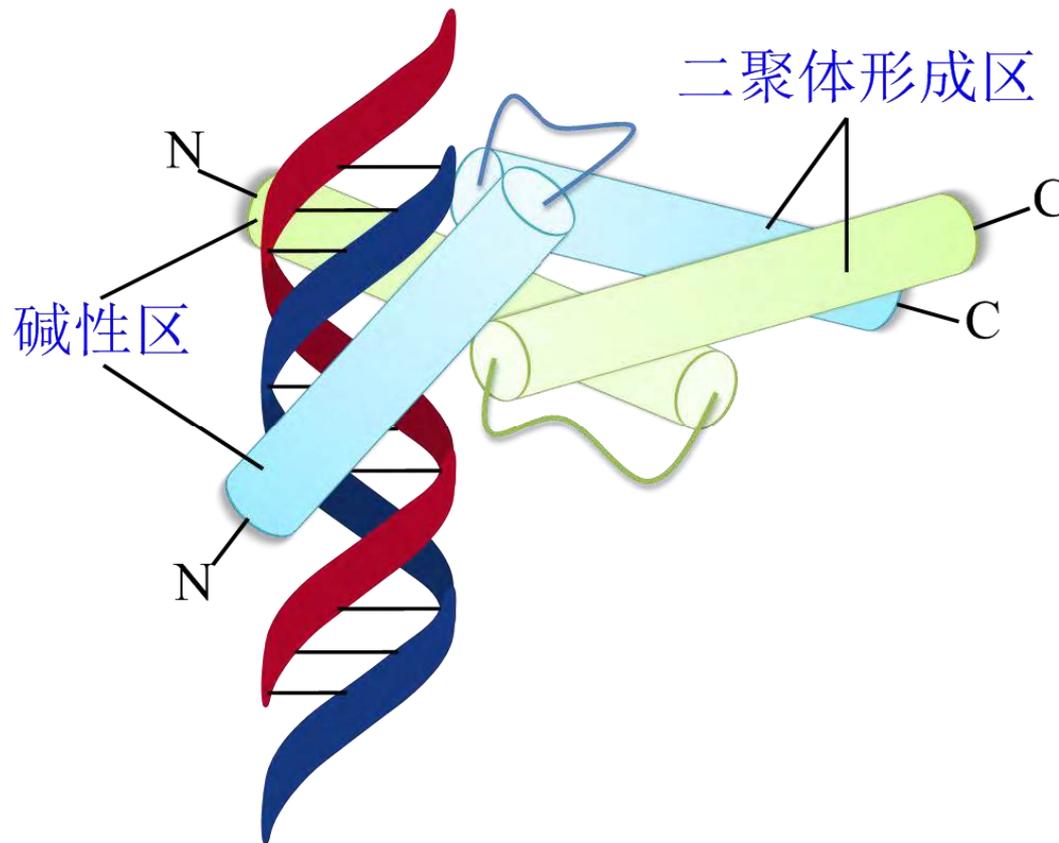


锌指 (zincfinger): 每个重复的指状结构约含**23**个氨基酸残基，锌以**4**个配价键与**4**个半胱氨酸或**2**个半胱氨酸和**2**个组氨酸相结合。整个蛋白质分子可有**2—9**个这样的锌指重复单位。每一个单位可以其指部伸入**DNA**双螺旋的深沟

二、转录起始的调控

最常见的DNA结合域：

2. 碱性螺旋-环-螺旋(basic helix-loop-helix, bHLH) 结构

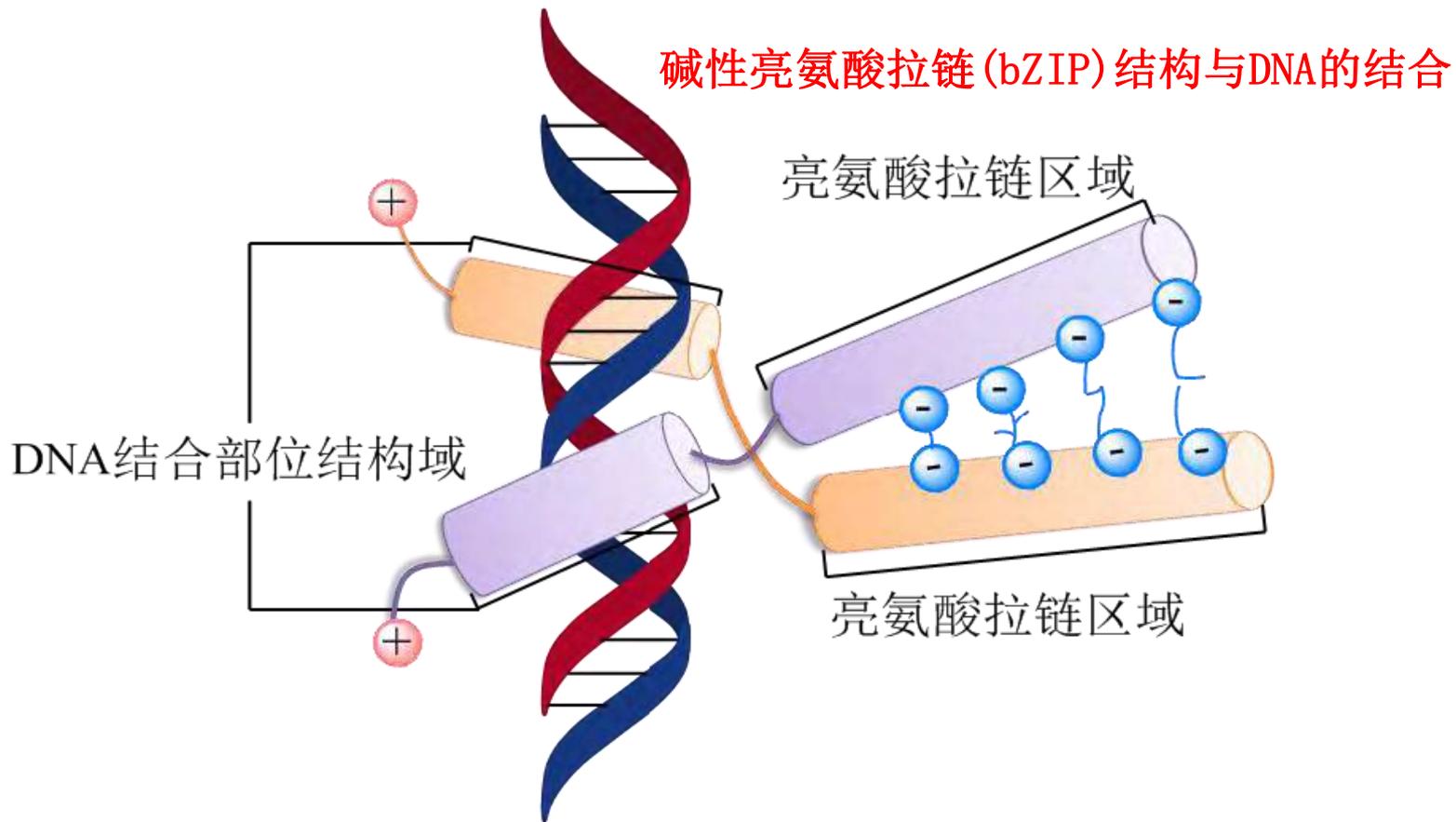


螺旋—环—螺旋(helix-loop-helix, HLH)：两个两性 α —螺旋通过一个肽段连结形成螺旋—环—螺旋结构，具备两种特性：具有一些与DNA结合的螺旋区，并能形成二聚体

二、转录起始的调控

最常见的DNA结合域：

3 碱性亮氨酸拉链(basic leucine zipper, bZIP) 结构

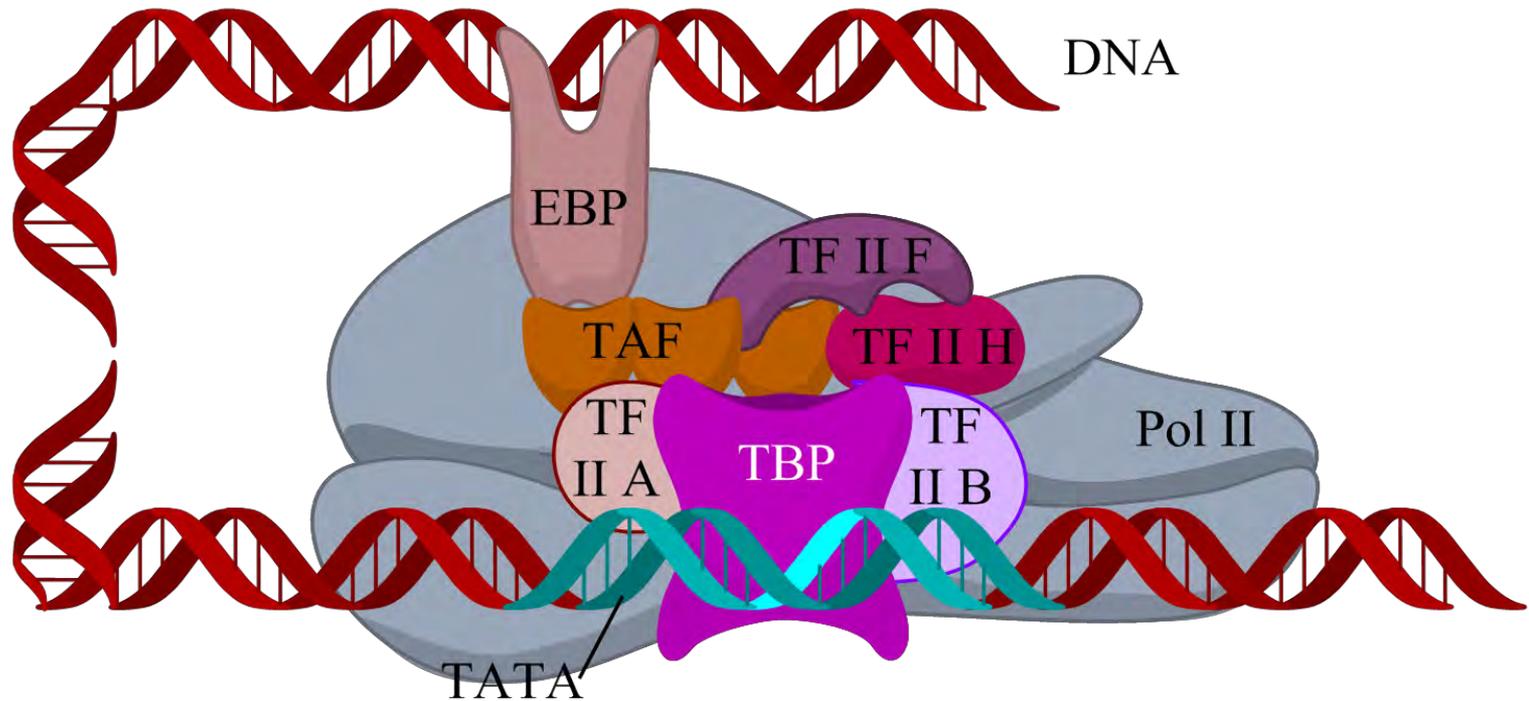


亮氨酸拉链：两条平行走向的肽链单体的 α 螺旋，通过亮氨酸残基，组成形似拉链的对称二聚体结构，**亮氨酸**总是有规律地每隔7个**氨基酸**就出现一次。

二、转录起始的调控

(四) 转录起始复合物的形成

真核RNA聚合酶II在转录因子帮助下，形成转录起始复合物（PIC）。



转录起始复合物, TAF (转录起始因子), EBP (增强子结合蛋白), TBP (TATA结合蛋白), TF (转录因子)

三、其他水平的调控

(一) 转录后的调控

1. mRNA的稳定性

- 稳定性受很多因素的影响，如加帽、加尾等。
- 组蛋白3' -端会形成一种发夹结构，使其免受核酸酶的攻击。
- 3' -UTR区富含AU序列(AU-rich element, ARE)结合蛋白可升高或降低mRNA的稳定性。

三、其他水平的调控

➤核蛋白体复合物(ribonucleoprotein, RNP)中的相关蛋白质含量可直接影响mRNA 的运输及在胞质内的稳定性。

➤IRE结合蛋白(IRE-binding protein, IRE-BP)与IRE (iron response element, IRE)结合可以延长TfR (transferrin receptor, TfR) 的半寿期。

2. mRNA前体的选择性剪接

三、其他水平的调控

(二) 翻译及翻译后的调控

- 蛋白质合成速率的快速变化在很大程度上取决于起始水平，通过磷酸化调节翻译起始因子 (eukaryotic initiation factor, eIF) 的活性对起始阶段有重要的控制作用。
- eIF2a的磷酸化导致蛋白质翻译受阻，但eIF-4E及其结合蛋白的磷酸化则激活翻译起始。

三、其他水平的调控

2. RNA结合蛋白的调节

- RNA结合蛋白 (RNA binding protein, RBP)，是指那些能够与RNA特异序列结合的蛋白质。
- 基因表达的许多调节环节都有RBP的参与，如前述转录终止、RNA剪接、RNA转运、RNA胞浆内稳定性控制以及翻译起始等。

三、其他水平的调控

3 . 翻译产物的调节

- 新合成蛋白质的半衰期长短是决定蛋白质生物学功能的重要影响因素。因此，通过对新生肽链的水解和运输，可以控制蛋白质的浓度在特定的部位或亚细胞器保持在合适的水平。
- 许多蛋白质需要在合成后经过特定的修饰才具有功能活性。通过对蛋白质的可逆的磷酸化、甲基化、酰基化修饰，可以达到调节蛋白质功能的作用，是基因表达的快速调节方式。

三、其他水平的调控

(三) 小分子RNA对基因表达的调控

➤某些小分子RNA也可调节真核基因表达，这些RNA均属于非编码RNA (non-coding RNA, ncRNA) 。

1. 微小RNA (microRNA, miRNA)

是一大家族小分子非编码单链RNA，长度约20~25个碱基，由一段具有发夹环结构，长度为70~90个碱基的单链RNA 前体 (pre-miRNA) 经Dicer酶剪切后形成。

三、其他水平的调控

■ miRNA的特点：

- 其长度一般为20~25个碱基；
- 在不同生物体中普遍存在；
- 其序列在不同生物中具有一定的保守性；
- 具有明显的表达阶段特异性和组织特异性；
- miRNA 基因以单拷贝、多拷贝或基因簇等多种形式存在于基因组中，大多位于基因间隔区。

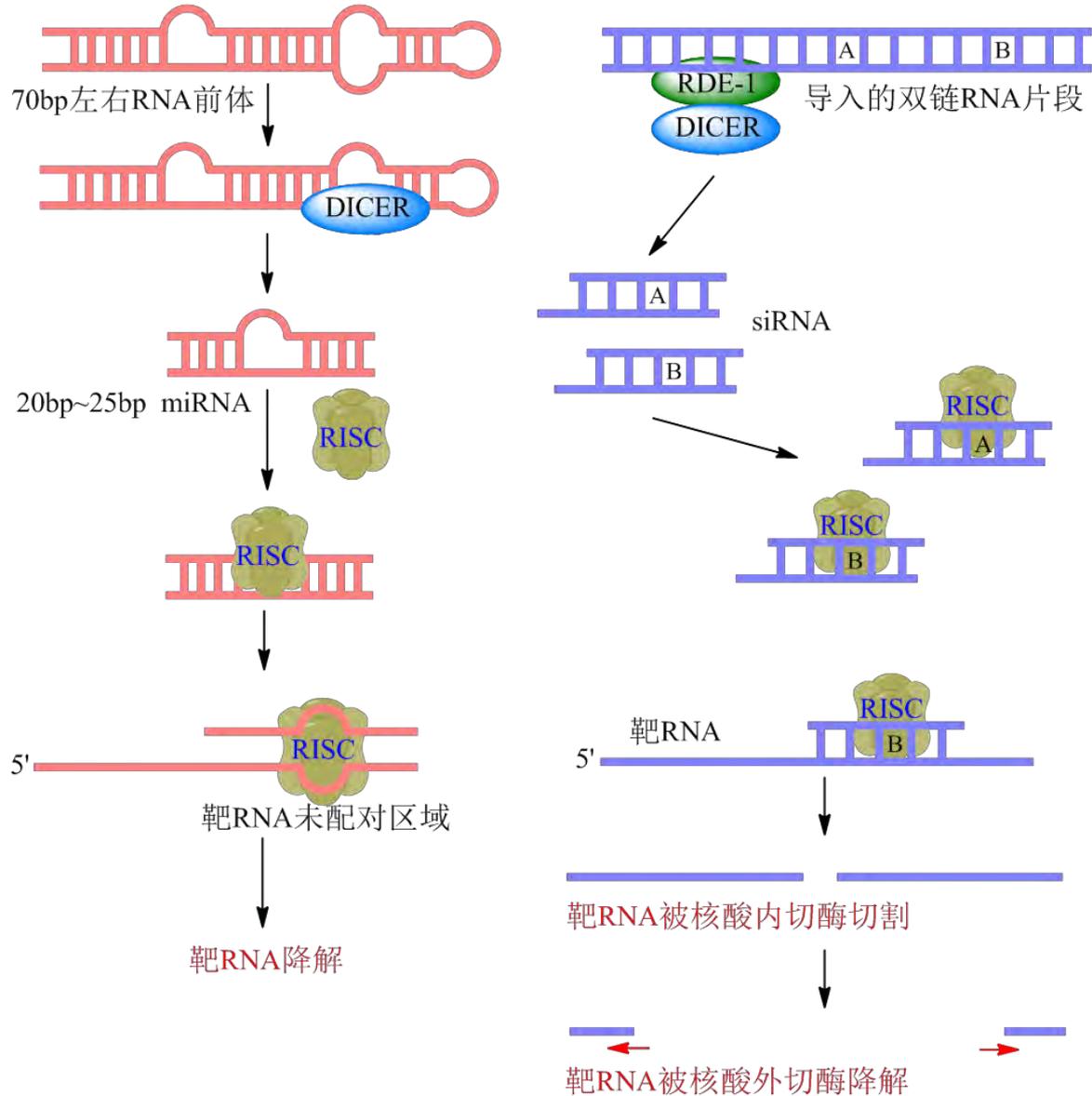
三、其他水平的调控

2. 干扰小RNA (small interfering RNA, siRNA)

- 是细胞内一类双链RNA (double-stranded RNA, dsRNA)在特定情况下通过一定酶切机制,转变为具有特定长度(21~23个碱基)和特定序列的小片段RNA。
- 双链siRNA参与RISC组成,与特异的靶mRNA完全互补结合,导致靶mRNA降解,阻断翻译过程。
- 由siRNA介导的基因表达抑制作用被称为RNA干涉 (RNA interference, RNAi)。

三、其他水平的调控

RNA干扰作用



三、其他水平的调控

(四) 基因重排

真核细胞在分化过程中，DNA分子里核苷酸序列的重新排列称为**基因重排**。

基因重排能形成新的基因，也可以调节基因的表达。

三、其他水平的调控

(五) . 基因拷贝数变化

DNA拷贝数变化 (copy number variant, CNV) 是一种介于1kb~3Mb的DNA片段的变异，其中包括缺失、重复、倒位和易位。

CNV通过改变基因剂量、调节基因活性来影响基因表达（癌基因激活与抑癌基因失活）、表型差异和表型适应，从而引起肿瘤发生以及其他遗传疾病

第四节

表观遗传学

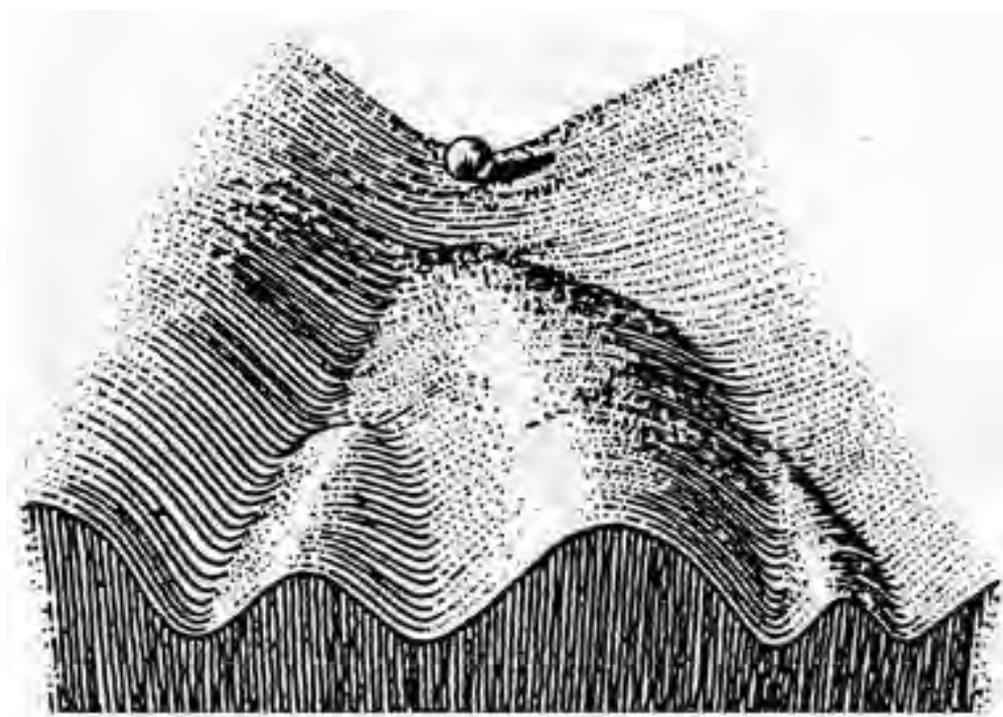
Epigenetics

表观遗传： Epigenetics

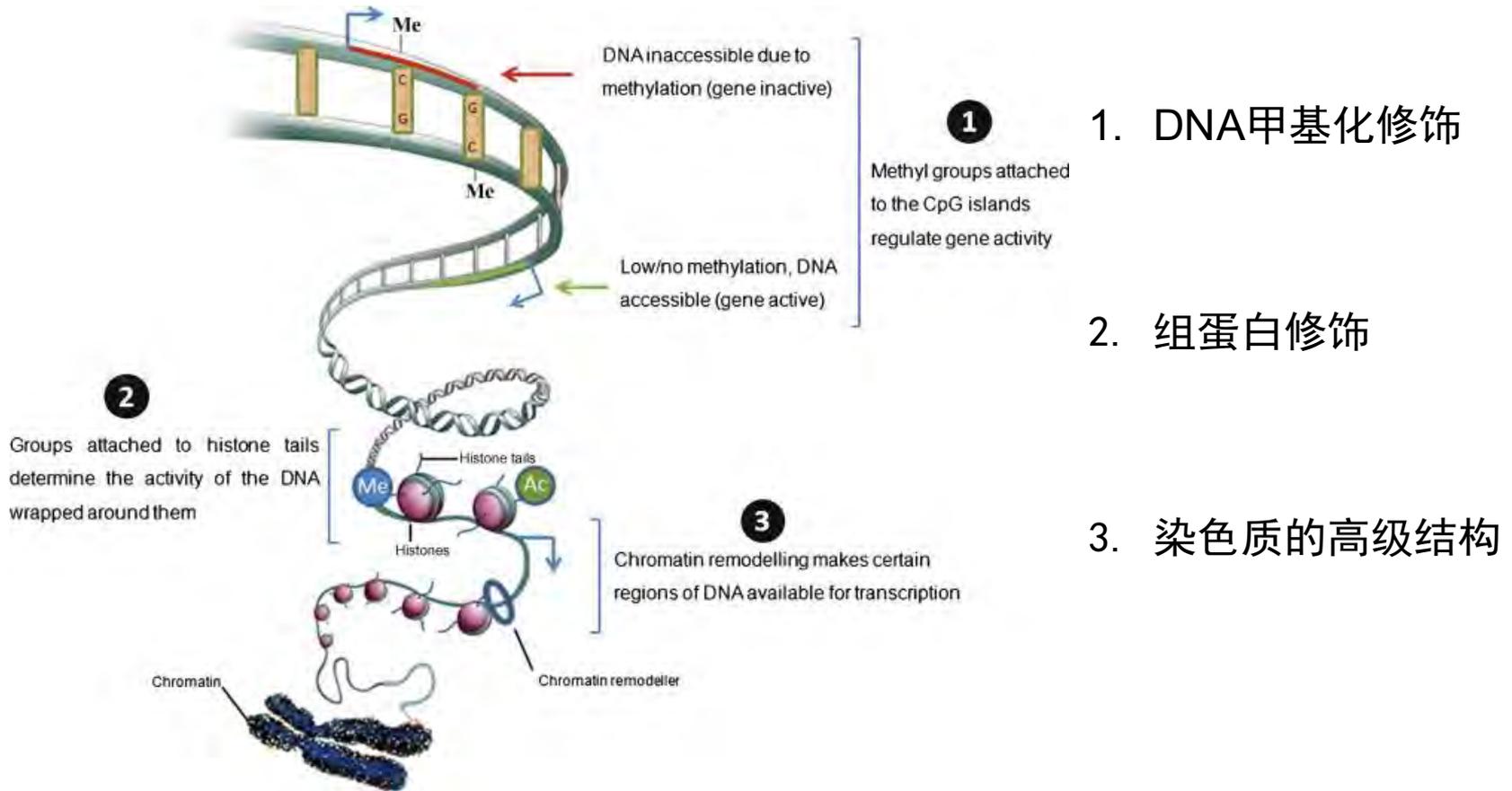
The branch of biology which studies the causal interactions between genes and their products which bring the phenotype into being.

Conrad Waddington, 1956

Waddington's epigenetic landscape:



不同的表观遗传修饰



遗传与表观遗传

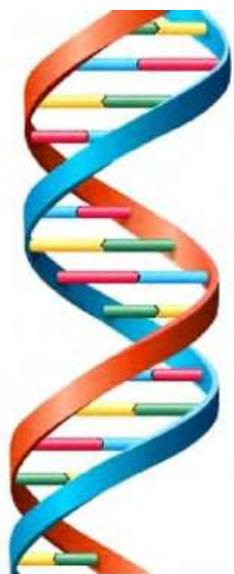
Epigenomics is essential for development. The cells making up the different organs of our body have identical genomes, but different epigenomics!



Genetics: the alphabet of life

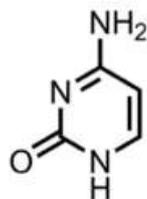
Epigenetics: the grammar of life

DNA 修饰



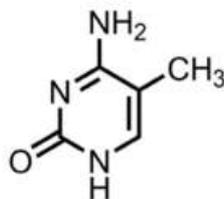
▶ A
▶ T
▶ C
▶ G

胞嘧啶



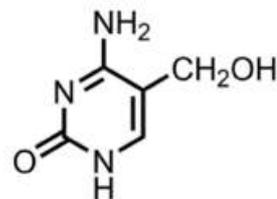
cytosine
(C)

5甲基胞嘧啶



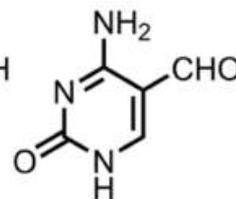
5-methylcytosine
(5mC)

5羟甲基胞嘧啶



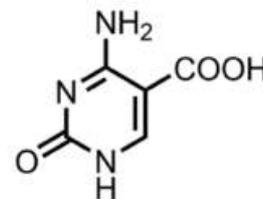
5-hydroxymethyl
cytosine
(5hmC)

5甲酰基胞嘧啶

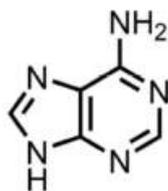


5-formyl
cytosine
(5fC)

5羧基胞嘧啶

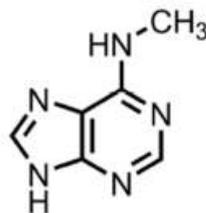


5-carboxylcytosine
(5caC)



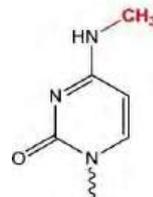
adenine
(A)

腺嘌呤



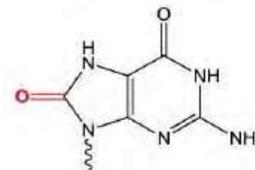
*N*⁶-methyladenine
(6mA)

*N*⁶甲基腺嘌呤



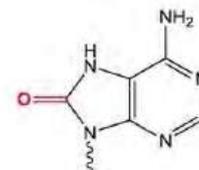
4-mC

4甲基胞嘧啶



8-oxoG

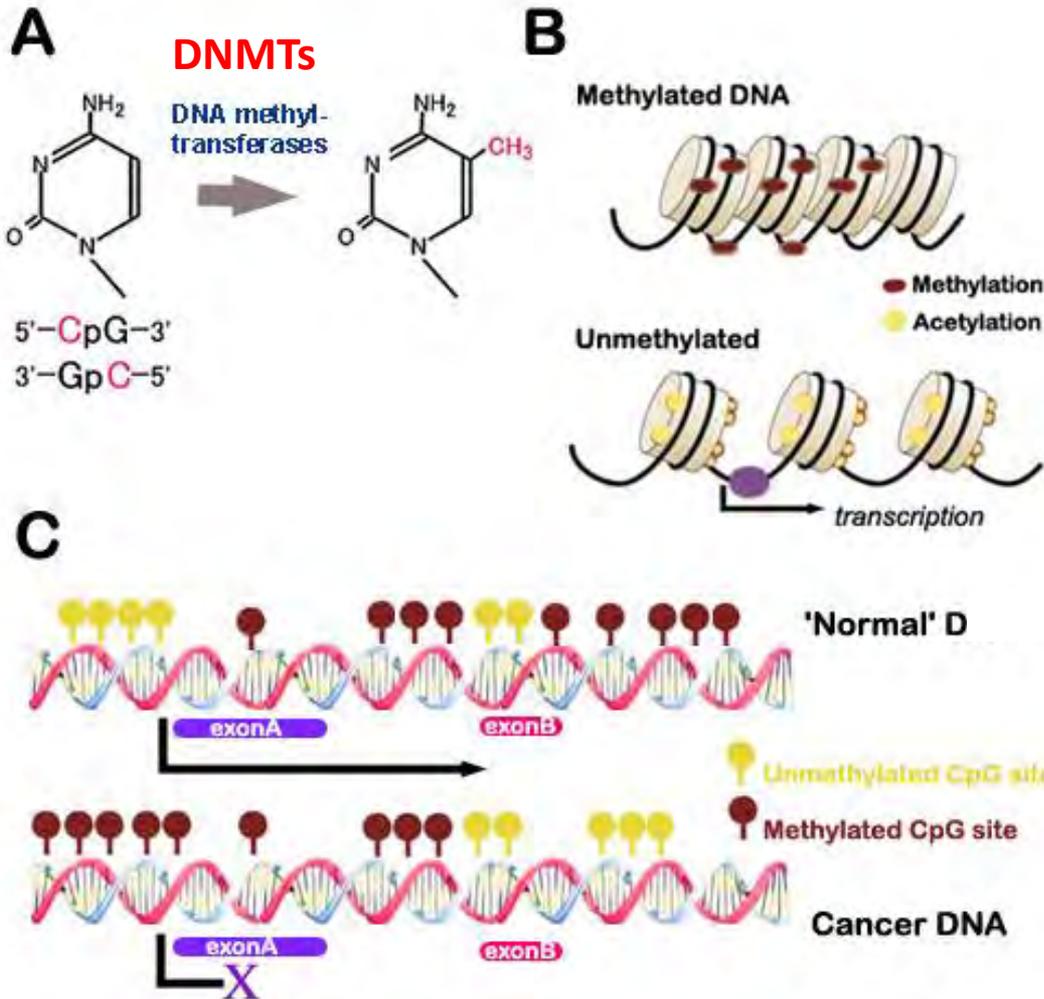
8氧代鸟嘌呤



8-oxoA

8氧代腺嘌呤

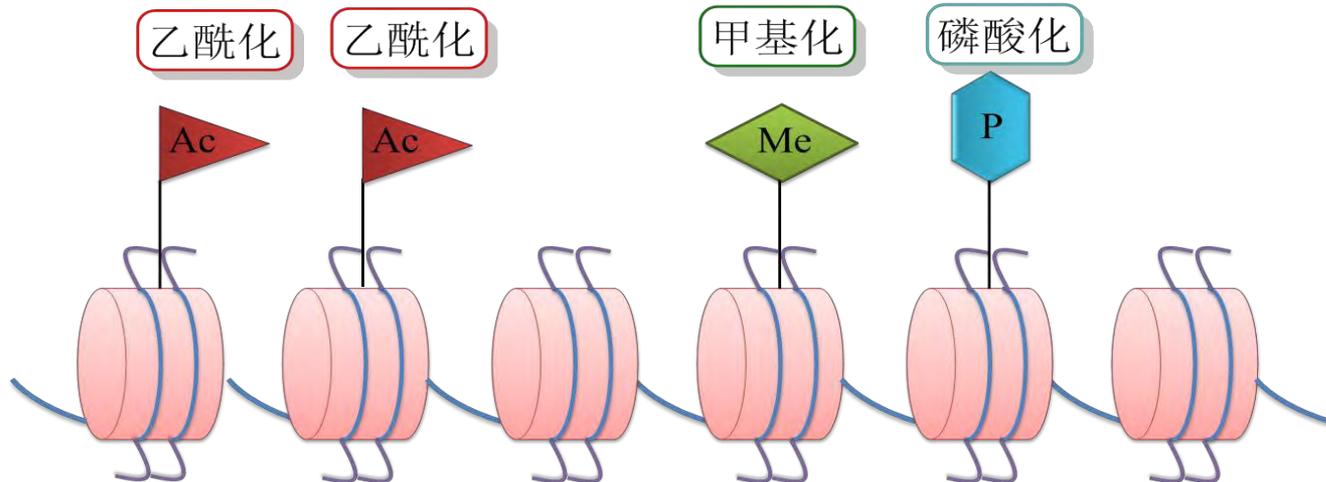
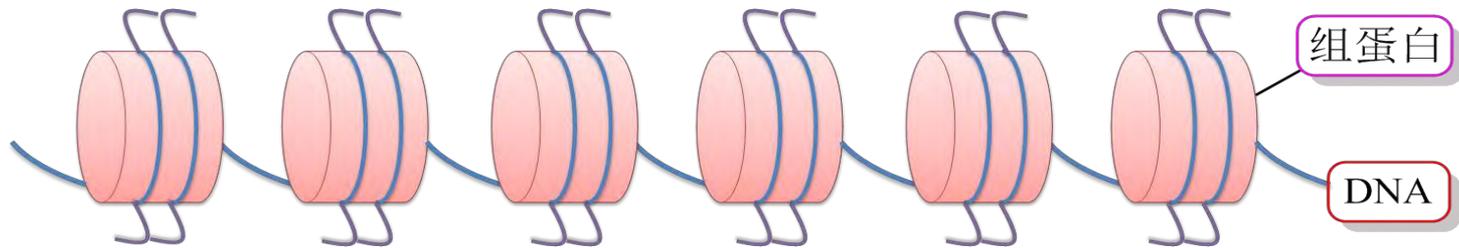
DNA 甲基化修饰



- Most occurs at CpG sites.
- ~30 million CpGs in human genome.
- CpGs often cluster in CpG dense regions (CpG islands - CGIs).
- 60-70% of gene promoters are covered by CGIs.
- Generally, CGI promoters are unmethylated.
- Inter-genic regions are generally methylated

组蛋白修饰

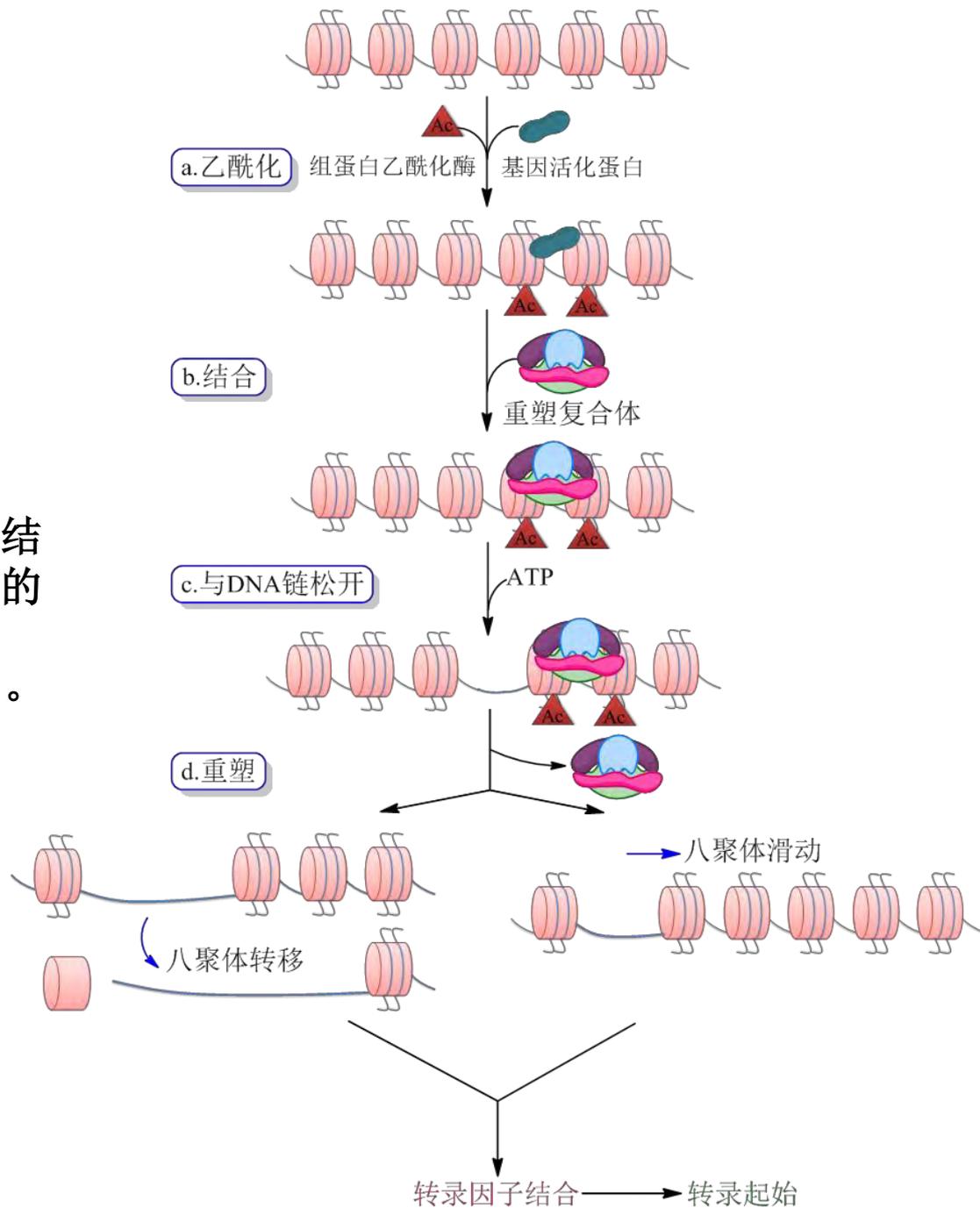
组蛋白结构及其化学修饰



组蛋白修饰

- 组蛋白的乙酰化修饰中和组蛋白尾巴上碱性氨基酸残基的正电荷，减弱组蛋白与带有负电荷的DNA之间的结合，选择性地使某些染色质区域的结构从紧密变得松散，有利于转录因子与DNA的结合，从而开放某些基因的转录，提高其表达水平。
- 组蛋白的甲基化增加其碱性度和疏水性，增强与DNA的亲合力。
- 特定定位点的乙酰化修饰和甲基化修饰互相抑制。
- 组蛋白的磷酸化修饰在细胞有丝分裂和减数分裂期间染色体浓缩以及基因转录激活过程中发挥着重要的调节作用。

组蛋白修饰引起局部染色质结构改变并进而影响转录活性的过程称为染色质重塑 (chromatin remodeling)。



组蛋白修饰

组蛋白修饰对染色质结构与功能的影响

组蛋白	氨基酸残基位点	修饰类型	功能
H3	Lys - 4	甲基化	激活
H3	Lys - 9	甲基化	染色质浓缩
H3	Lys - 9	甲基化	DNA 甲基化必需
H3	Lys - 9	乙酰化	激活
H3	Ser - 10	磷酸化	激活
H3	Lys - 14	乙酰化	防止 Lys - 9 的甲基化
H3	Lys - 79	甲基化	端粒沉默
H4	Arg - 3	甲基化	
H4	Lys - 5	乙酰化	核小体装配
H4	Lys - 12	乙酰化	核小体装配
H4	Lys - 16	乙酰化	核小体装配
H4	Lys - 16	乙酰化	Fly X 激活

Lys= 赖氨酸; Ser= 丝氨酸; Arg= 精氨酸。译自: Gene IX, by Benjamin Lewin

组蛋白修饰

➤ 组蛋白修饰对于基因表达影响的机制也包括两种相互包容的理论。即：组蛋白的修饰直接影响染色质或核小体的结构，以及化学修饰募集了其他调控基因转录的蛋白质，为其他功能分子与组蛋白结合搭建了一个平台。这些理论构成了“组蛋白密码 (histone code)” 的假说。

染色质的活化

真核生物细胞核中，常染色质占大部分，呈松散分布，具有很强的转录活性。

常染色质在转录前多已被解旋或松弛，可包括核小体结构的改变或DNA本身局部结构的变化（如双螺旋的局部解旋、右旋DNA变成左旋等），暴露结构基因，使RNA聚合酶得以结合，促进转录因子结合到启动区DNA，发生基因转录

染色质的活化

具有转录活性的染色质被称为活性染色质。

染色质活化后，出现对核酸酶高度敏感的位点，称为超敏位点，位于调节蛋白结合位点附近。

染色质的活化

染色质重塑 (chromatin remodeling)：通过调整核小体的相位，中和组蛋白尾巴碱性氨基酸残基（赖氨酸K、精氨酸R、组氨酸H等）的正电荷，减弱核小体中碱性氨基酸与DNA的结合，降低相邻核小体间的聚集使核小体滑动暴露本来被遮蔽的元件，或使核小体表面的元件瞬间暴露的动态变化过程

1. 依赖ATP的染色质物理重塑

ATP水解供能使核小体沿DNA滑动，或使核小体解离并重新装配。使DNA超螺旋旋矩和旋相发生变化，使转录因子更易接近并结合核小体DNA，从而调控基因的转录过程。

2. 染色质的化学修饰——多发生在组蛋白末端“尾巴”，包括乙酰化、磷酸化、甲基化和泛素化等。

➤ 富含Lys的H1组蛋白水平降低

➤ H2A • H2B二聚体不稳定性增加

➤ 核心组蛋白H3、H4发生乙酰化、甲基化或磷酸化修饰

原核与真核基因的基因表达调控

Chromatin Activation

Transcription

Post-transcriptional Modification

RNA Transport

RNA degradation

Translation

Post-translational Modifications

