

# 第十三章 基因表达调控

**Chapter 13**

**Gene expression and regulation**

# 本章要求

- 1.掌握基因表达及调控的基本含义和规律
- 2.掌握乳糖操纵子的结构特征和转录调控原理
- 3.掌握真核基因转录激活的调控
- 4.掌握下列概念：

**gene expression, regulation of gene expression, operon, promoter, enhancer, SD sequence, transcription unit, cis-acting element, trans-acting factor, housekeeping gene**

# 第一节 基因表达及其调控的概念及特点

## 一、基本概念

### 1. 基因表达 (**gene expression**)

细胞将储存在**DNA**中的遗传信息经过转录或转录-翻译过程转变为具有生物学活性分子（**RNA**或蛋白质）的过程。

### 2. 基因表达调控 (**regulation of gene expression**)

在基因表达的不同阶段控制基因表达速率和产量的过程。

## 二、基因表达的特点和方式

### （一）基因表达的特性

#### 1. 基因表达的时间特异性

生物体内某一特定基因的表达严格按照特定的时间顺序发生。

#### 2. 基因表达的空间特异性

特定基因按组织需求进行表达的特性  
又称作： 细胞特异性  
          组织特异性

## （二）基因表达的基本方式

### 1. 组成性表达（**constitutive expression**）

生物体内一些不受环境变化影响的基因表达。

以组成性方式表达的基因，称作：

管家基因（**housekeeping gene**）

### 2. 适应性表达（**adaptive expression**）

生物体内一些受环境因素影响的基因表达。

环境因素上调基因表达——诱导性表达

环境因素下调基因表达——阻遏性表达

# 三、基因表达调控的特点及意义

## （一）基因表达调控的特性

### 1. 转录起始的调控特点

转录起始是RNA聚合酶与DNA序列相互作用的结果。

原核生物RNA聚合酶能直接与DNA结合

真核生物RNA聚合酶需要其他蛋白质帮助

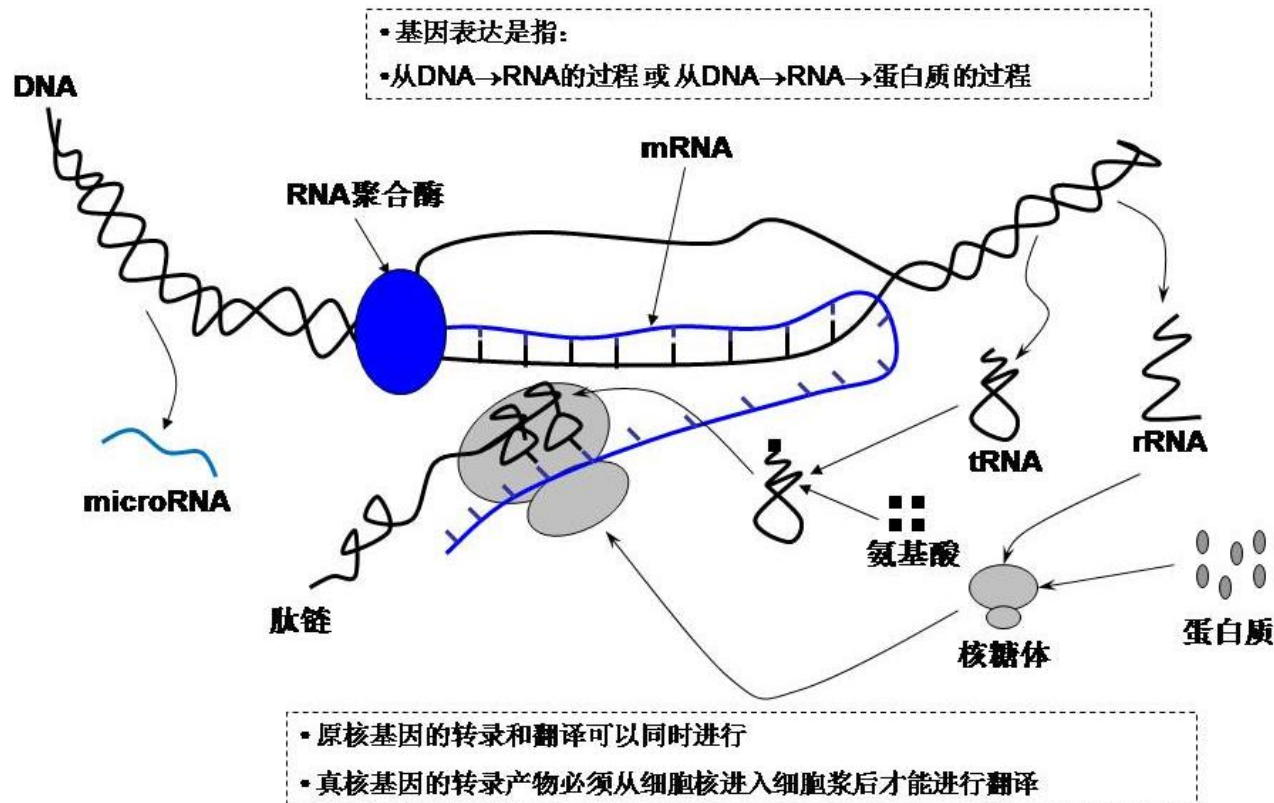
### 2. 翻译起始的调控特点

翻译起始是核糖体与mRNA相互作用的结果

### 3. 对转录和翻译产物的调控

## (二) 基因表达调控的生物学意义

1. 适应内外环境的变化
2. 维持个体生长、发育和分化



## 第二节 原核基因表达调控

### 一、转录水平的调控

#### (一) 转录调控的关键因素

##### 1. 转录调控序列

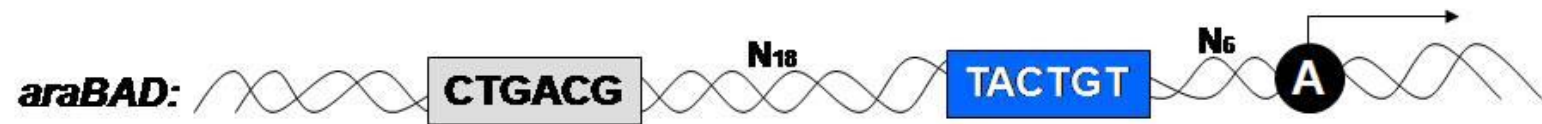
原核生物的多数基因以操纵子为转录单位

操纵子（operon）：调控区+结构基因

调控区：启动子、操作子

启动子：原核生物RNA聚合酶识别及结合的一段DNA序列

启动子:



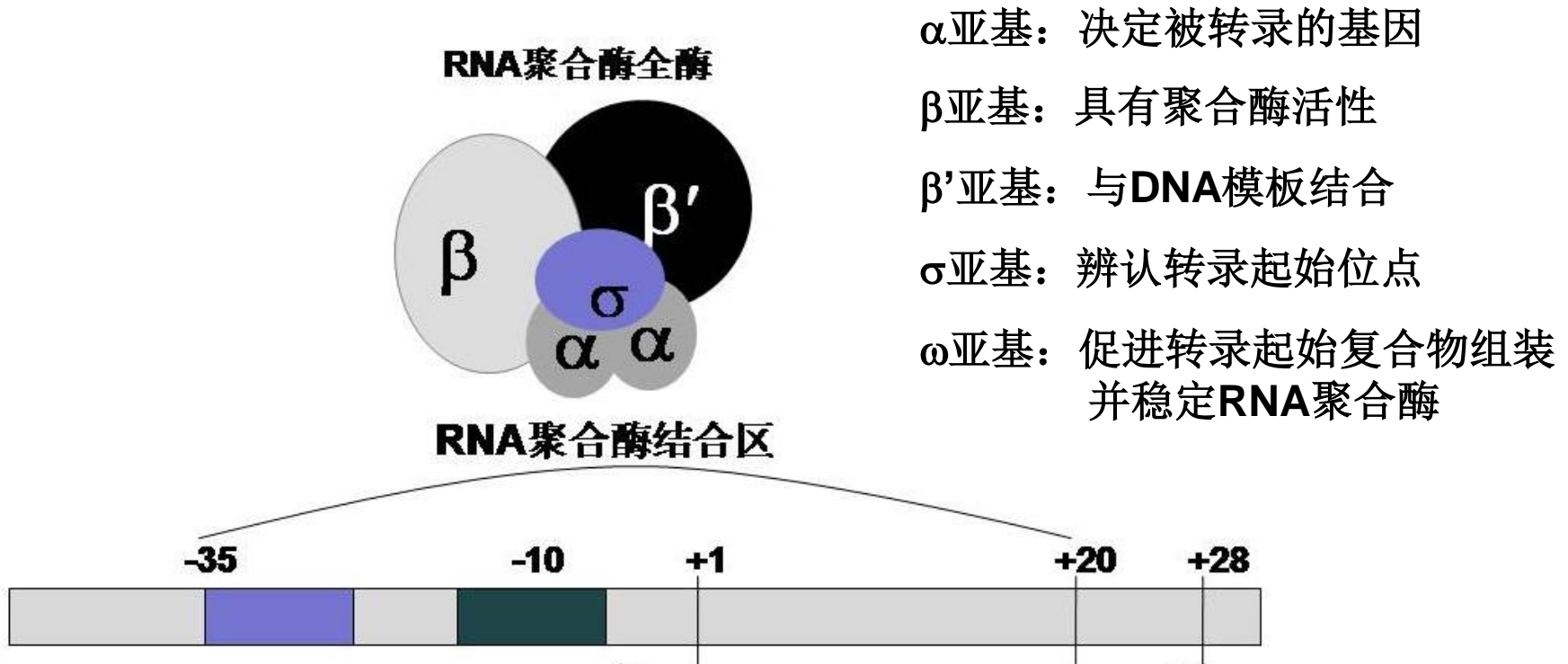
Consensus  
sequence:

TTGACA	(-35区域)
TATAAT	(-10区域)

## 2. RNA聚合酶 (RNA polymerase)

大肠埃希菌只有一种RNA聚合酶

特点：6个亚基组成RNA聚合酶全酶  
直接识别和结合启动子

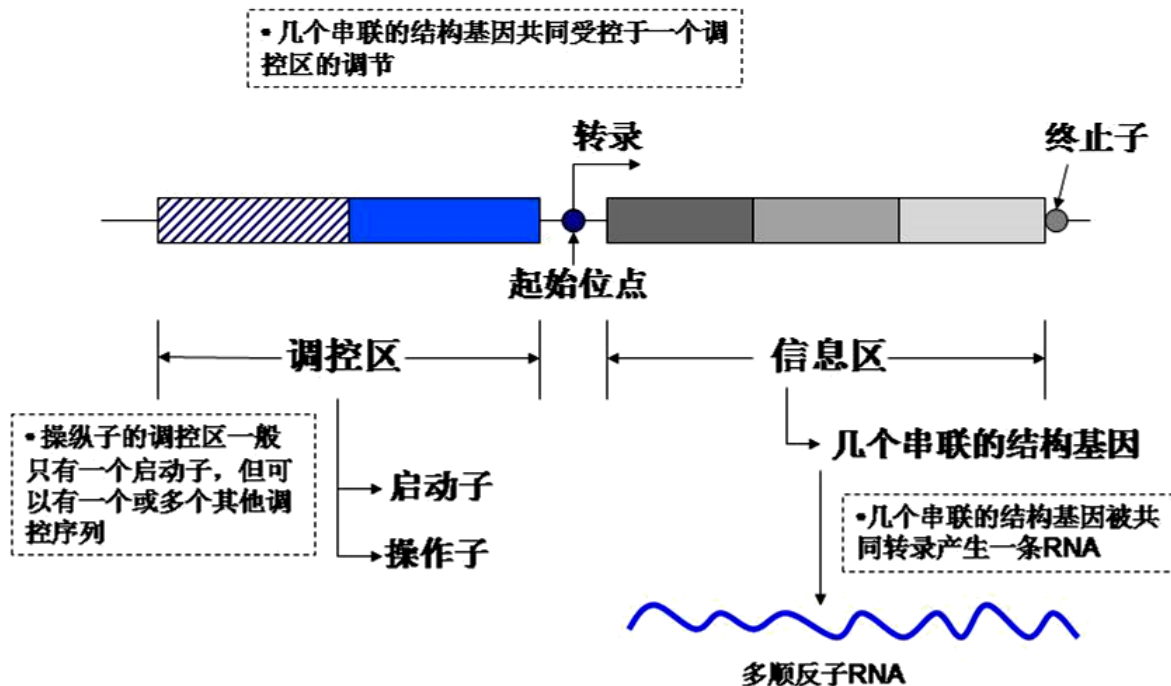
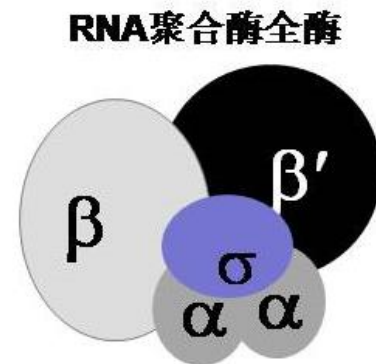


## (二) 转录调控的基本特点

### 1. $\sigma$ 因子决定基因的转录

即原核生物RNA聚合酶的 $\sigma$ 亚基

### 2. 操纵子是转录调控的基本单位

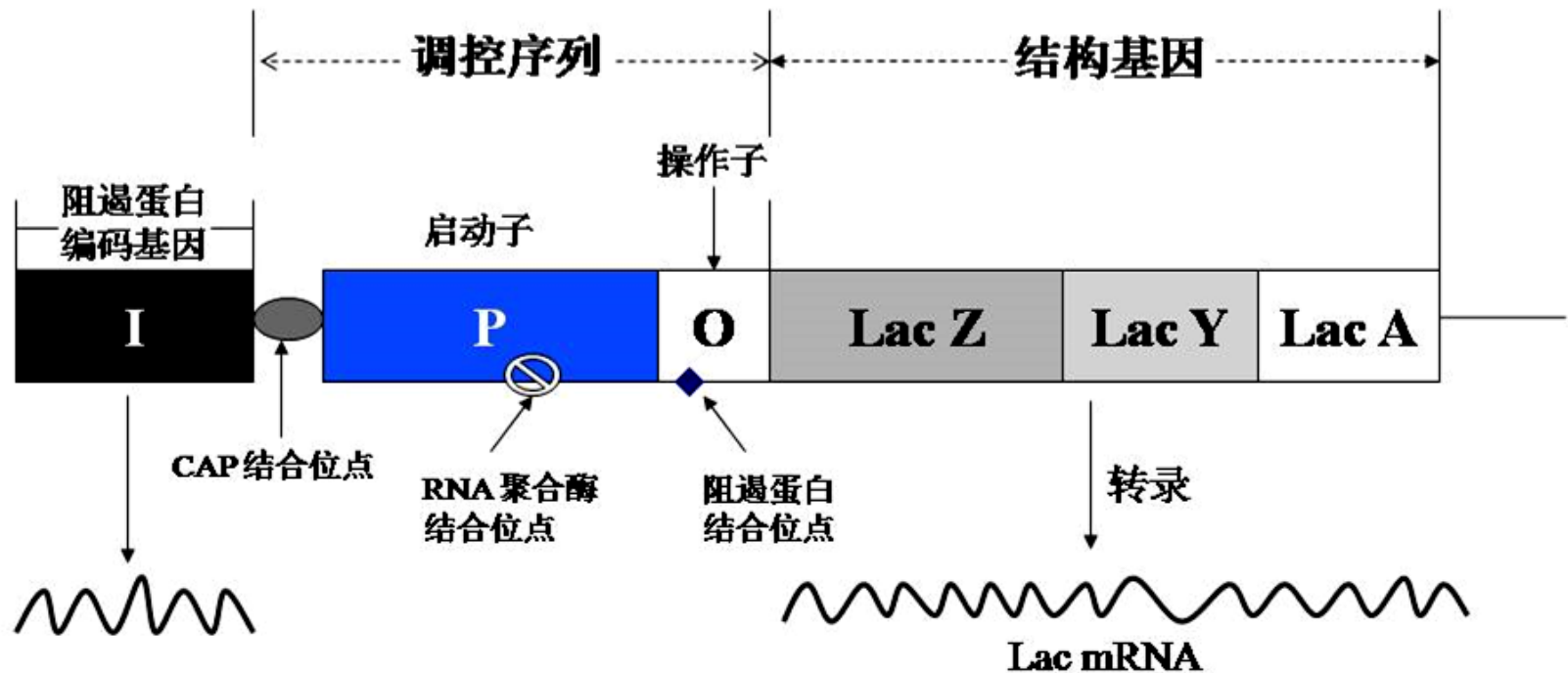


### 3. 阻遏调控是原核基因表达调控的基本原理

### (三) 操纵子的转录调控

#### 1. 乳糖操纵子的基本结构

三个结构基因串联在一起共同受一个调控区调控的一段DNA序列。



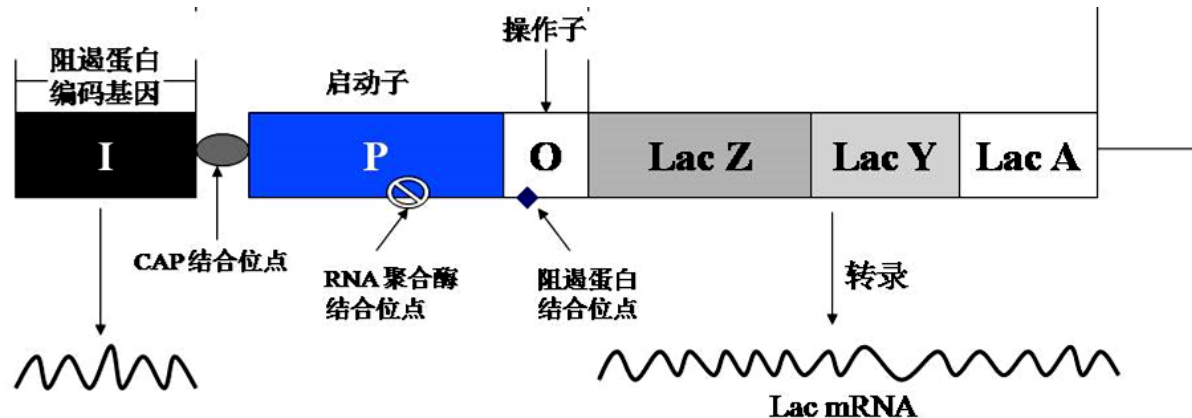
# (1) 乳糖操纵子的结构基因

三个结构基因：

- Z基因：编码 $\beta$ -半乳糖苷酶（Lac Z）
- Y基因：编码透性酶（Lac Y）
- A基因：编码乙酰基转移酶（Lac A）

ZYA结构基因：

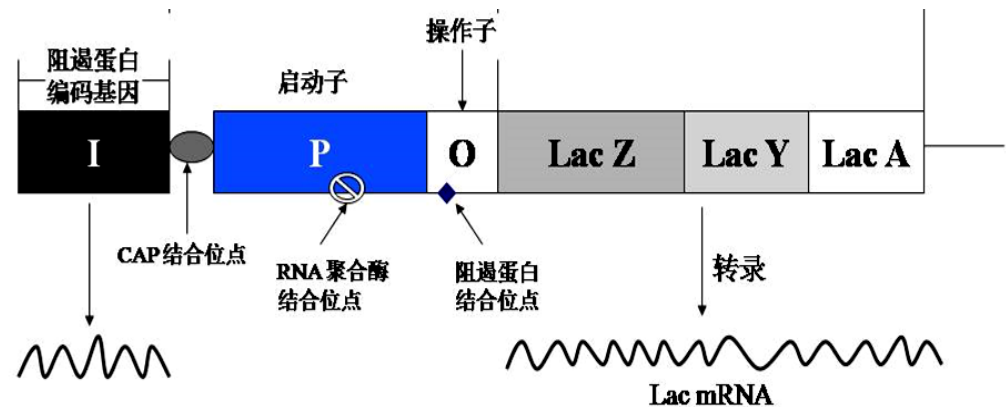
- 位于乳糖操纵子的下游
- 共同转录产生一条多顺反子mRNA
- 三个结构基因有各自独立的可读框（ORF）



## (2) 乳糖操纵子的调控序列

两种调控序列：

- 启动子：位于操作子的上游  
有RNA聚合酶识别和结合的位点
- 操作子：位于启动子的下游  
部分序列与启动子重叠  
有阻遏蛋白的结合位点



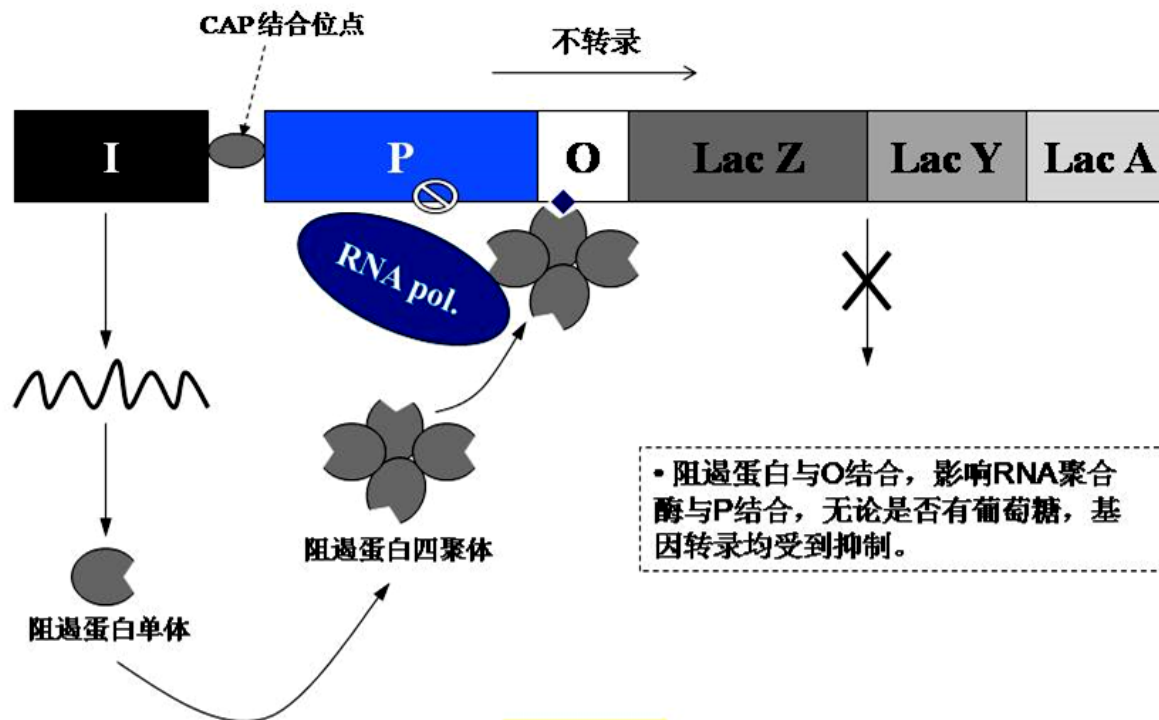
**CAP结合位点：**

- 位于启动子上游
- 能与分解代谢物基因激活蛋白（CAP）结合，促进RNA聚合酶的转录活性

### (3) 阻遏蛋白的编码基因

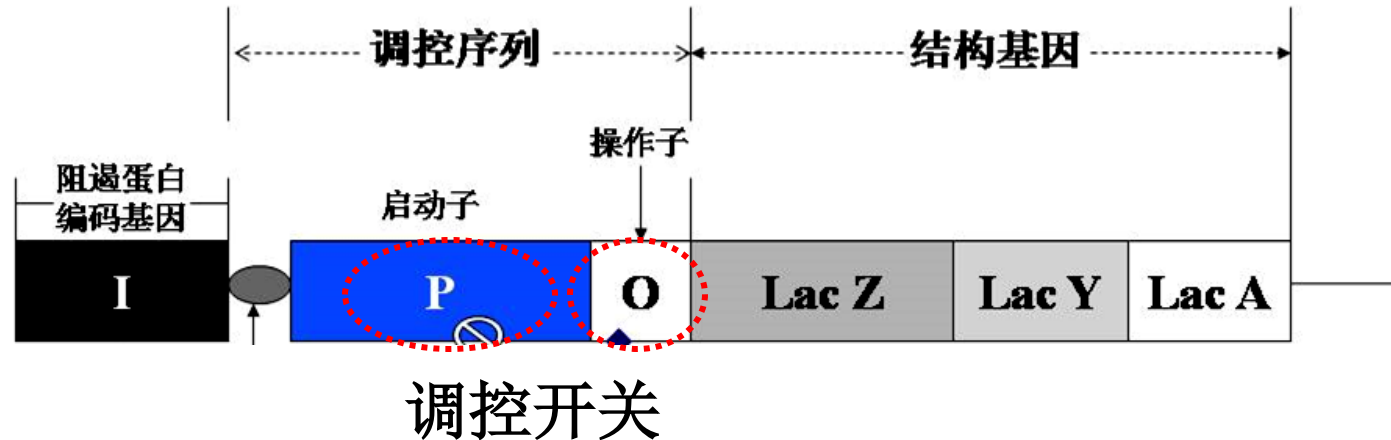
#### I基因:

- 编码阻遏蛋白，其编码产物参与乳糖操纵子的调控
- 不属于乳糖操纵子成员，其转录不受乳糖操纵子调控
- 是一种低效率持续转录的基因



## 2. 乳糖操纵子的转录调控

调控序列中的**P**、**O**是两个关键的调控点



基因表达顺序：

- RNA聚合酶与启动子（P）结合
- 经过操作子（O）
- 到达三个串联的结构基因
- 转录一条多顺反子Lac mRNA

# (1) 基本原理

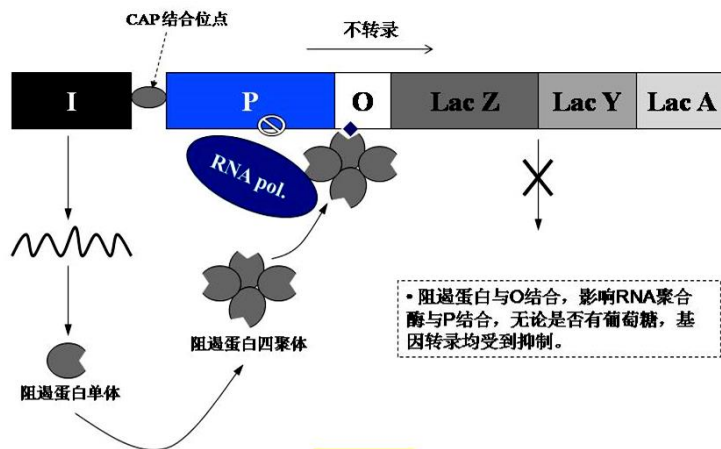
乳糖操纵子转录调控的正、负两种模式：

- 阻遏蛋白的负调控
- cAMP-CAP的正调控

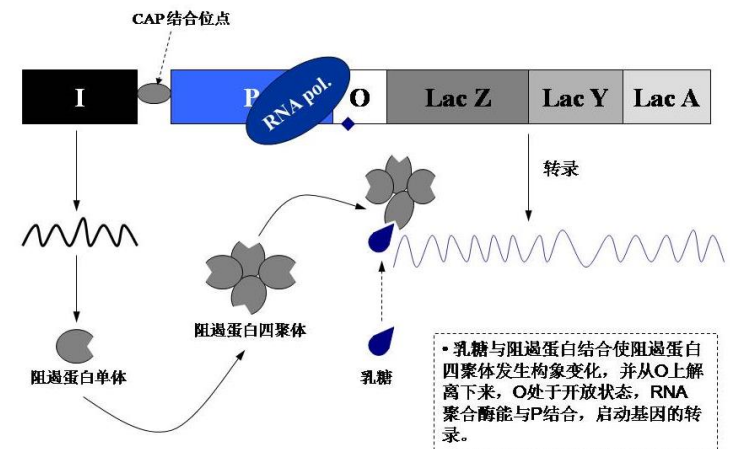
## 1) 阻遏蛋白的负调控

- 阻遏蛋白四聚体与O结合阻碍RNA聚合酶与P结合

没有乳糖时：



有乳糖时：

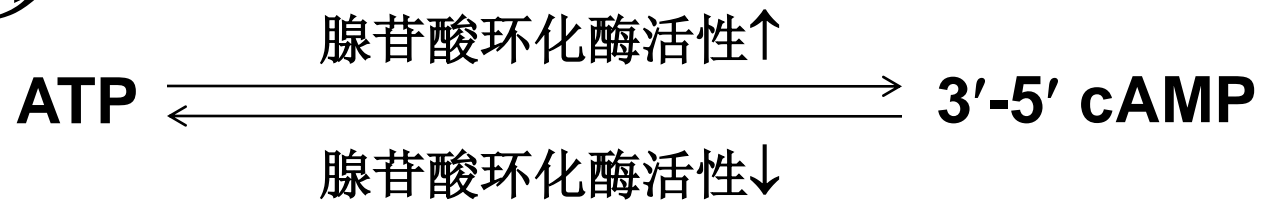


## 2) cAMP-CAP的正调控

- 乳糖操纵子中的CAP结合位点是cAMP-CAP复合物促进RNA聚合酶转录活性的作用靶点

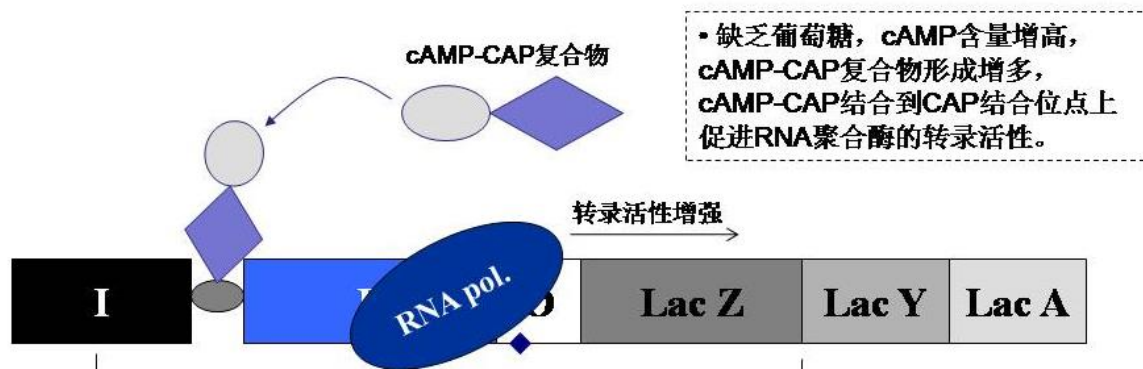
cAMP-CAP复合物的产生:

- 有葡萄糖时，葡萄糖代谢产物抑制腺苷酸环化酶的活性  
cAMP产生减少



- 没有葡萄糖时，cAMP↑

cAMP-CAP复合物↑ → 乳糖操纵子的CAP结合位点

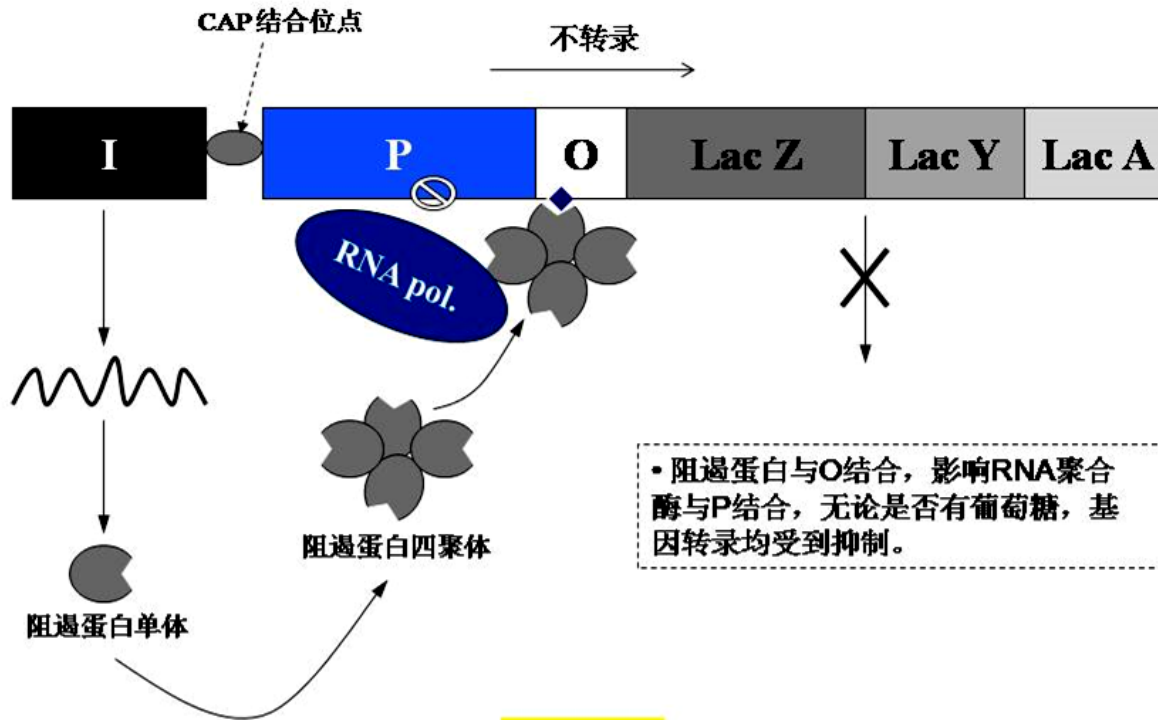


### 3) 正负调控系统的协同调节

- 阻遏蛋白的负调控
- cAMP-CAP复合物的正调控

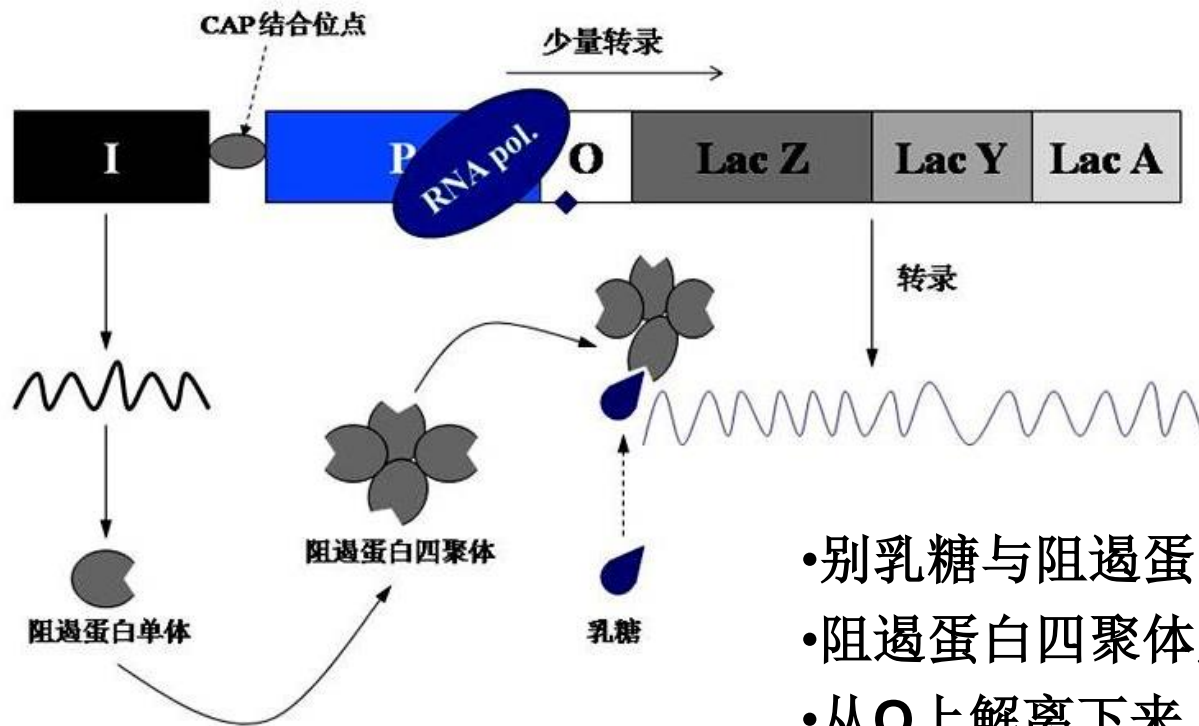
#### ① 没有乳糖时:

- 不管是否有葡萄糖都不能启动基因的转录，因为阻遏蛋白与O结合，O处于关闭状态



## ②乳糖和葡萄糖共同存在时，

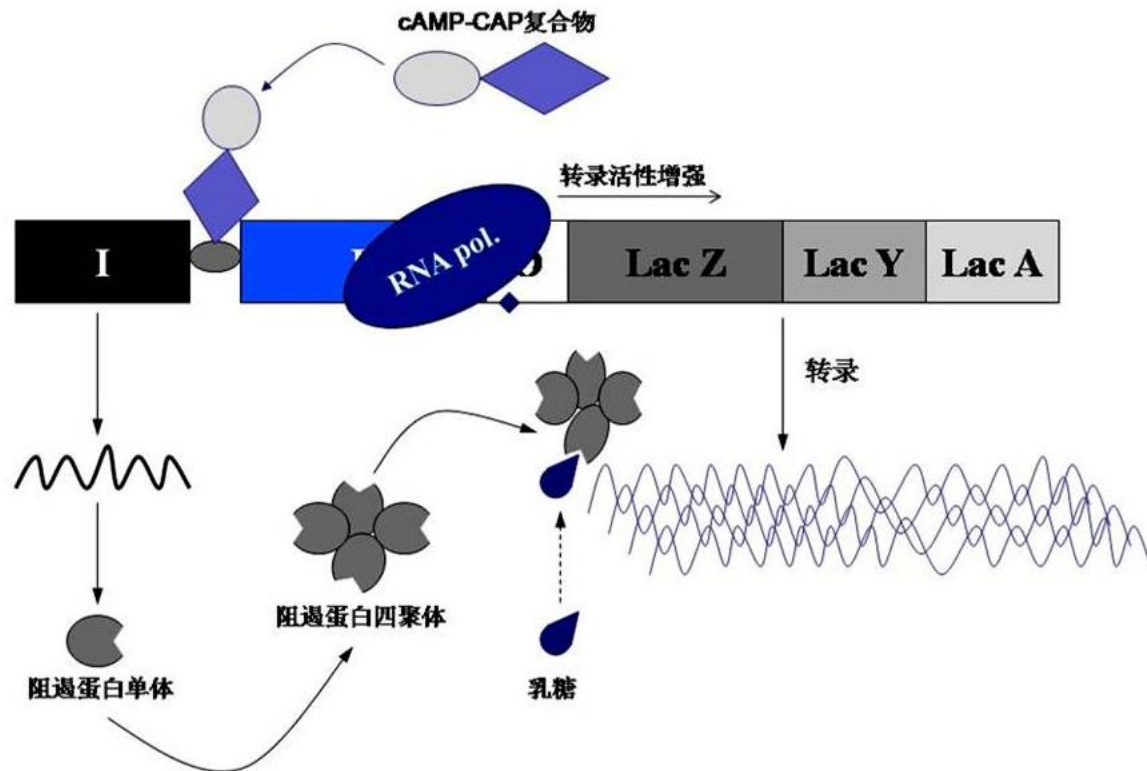
- 阻遏蛋白与别乳糖结合，O处于开放状态
- cAMP处于低水平，cAMP-CAP复合物形成受阻
- Lac mRNA低水平转录



- 别乳糖与阻遏蛋白结合
- 阻遏蛋白四聚体发生构象变化
- 从O上解离下来
- O处于开放状态

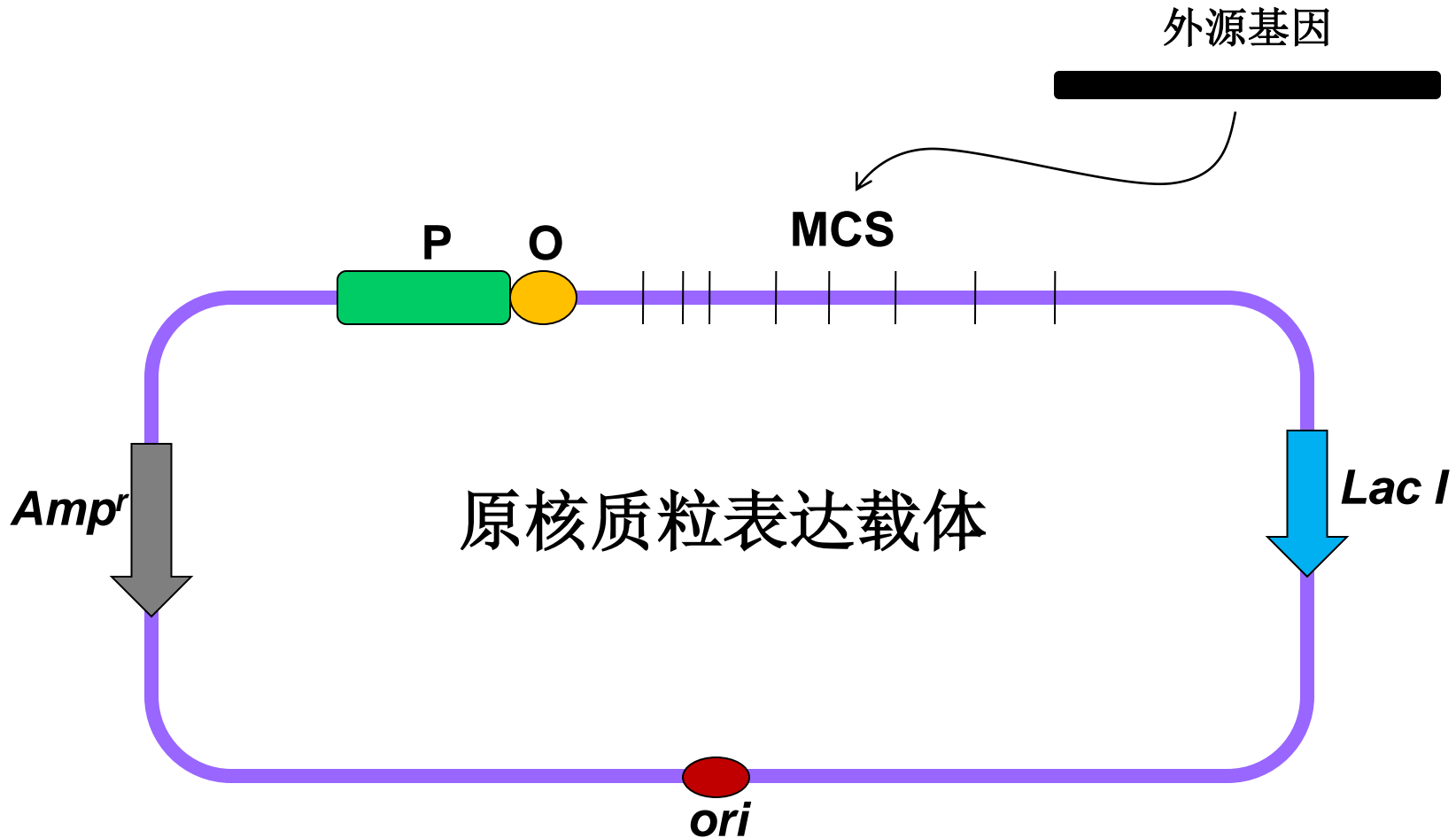
### ③有乳糖，没有葡萄糖，

- 阻遏蛋白与别乳糖结合，O处于开放状态
- cAMP含量 $\uparrow$ ，cAMP-CAP复合物形成 $\uparrow$ ，cAMP-CAP复合物结合到乳糖操纵子的CAP结合位点，促进RNA聚合酶与P结合
- Lac mRNA高水平转录



## (2) 乳糖操纵子转录调控原理的应用

- 广泛用于基因工程中原核表达载体的构建

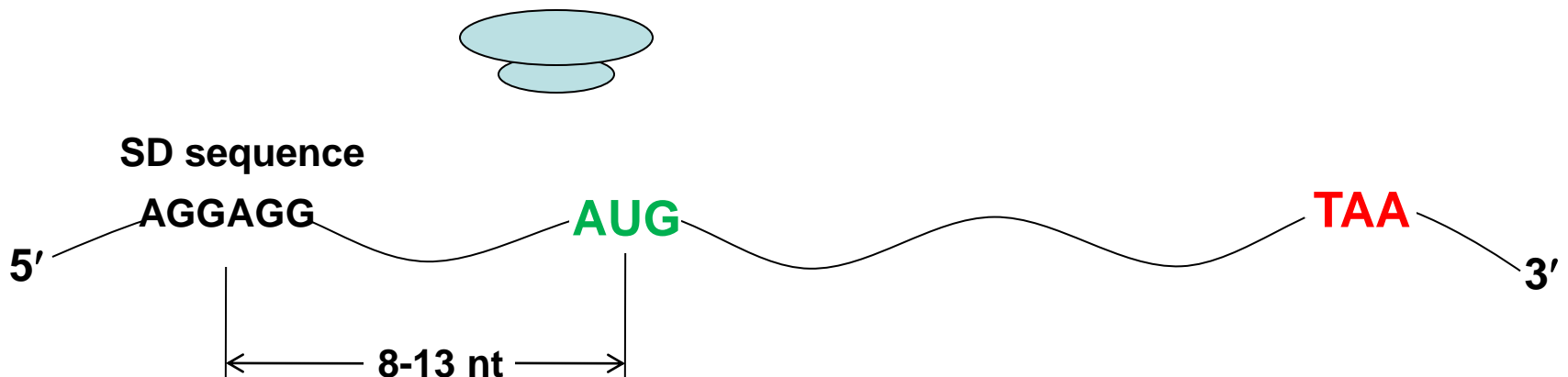


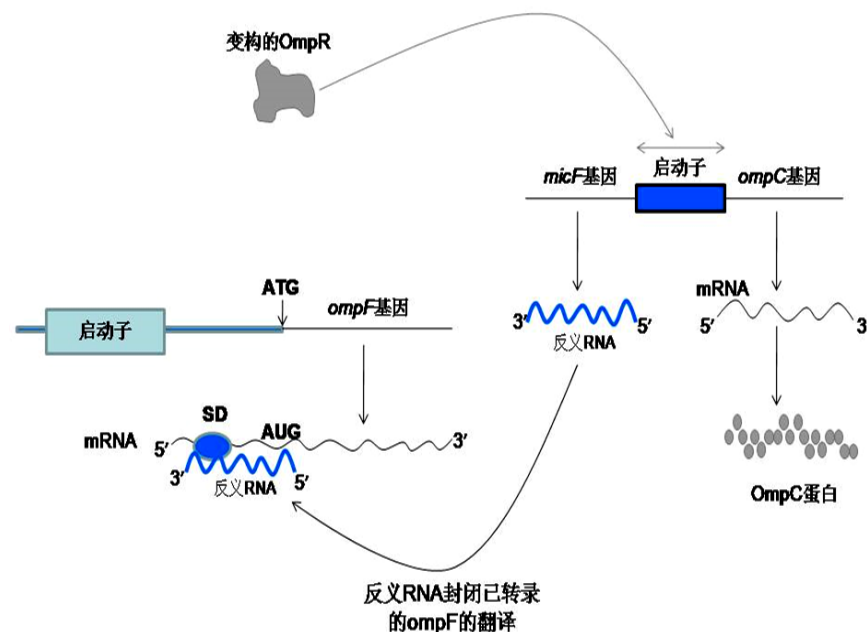
## 二. 翻译水平的调控

### (一) SD序列对翻译起始的影响

#### SD序列:

- 位于mRNA的起始密码子AUG上游8-13个核苷酸
- 长度4-9个核苷酸
- 共有序列是AGGAGG
- 可被核糖体小亚基中的16S rRNA通过碱基互补特异性识别和结合
- 被称作核糖体结合位点 (ribosomal binding site, RBS)





### **(三) mRNA的稳定性及密码子偏爱性 对翻译的影响**

#### **原核生物的mRNA:**

- 没有5'端帽结构和3'端尾结构
- 5'端和3'端可以形成发夹结构，抵抗外切核酸酶的水解
- 一些调节蛋白可结合mRNA的发夹结构，调节mRNA的稳定性

#### **密码子对翻译的影响:**

- 偏爱密码子 (preferred codon)，翻译速度快
- 稀有密码子 (rare codon)，翻译速度慢

# 第三节 真核基因表达调控

## 一、真核基因的结构和表达特点

### (一) 真核基因的结构特点

真核基因是断裂基因（split gene），包含：

- 编码序列（coding sequence）
- 非编码序列（non-coding sequence）

## 1. 真核基因的编码序列:

- 被非编码序列间隔开, 呈不连续方式排列
- 属于外显子 (**exon**)

### 外显子:

- 能体现在成熟**mRNA**中的序列
- 包括: 基因的编码序列
  - 5'-非翻译区 (5'-UTR)**
  - 3'-非翻译区 (3'-UTR)**

## 2. 真核基因的非编码序列:

包括: 内含子 (intron)

5'-非翻译区 (5'-UTR)

3'-非翻译区 (3'-UTR)

内含子:

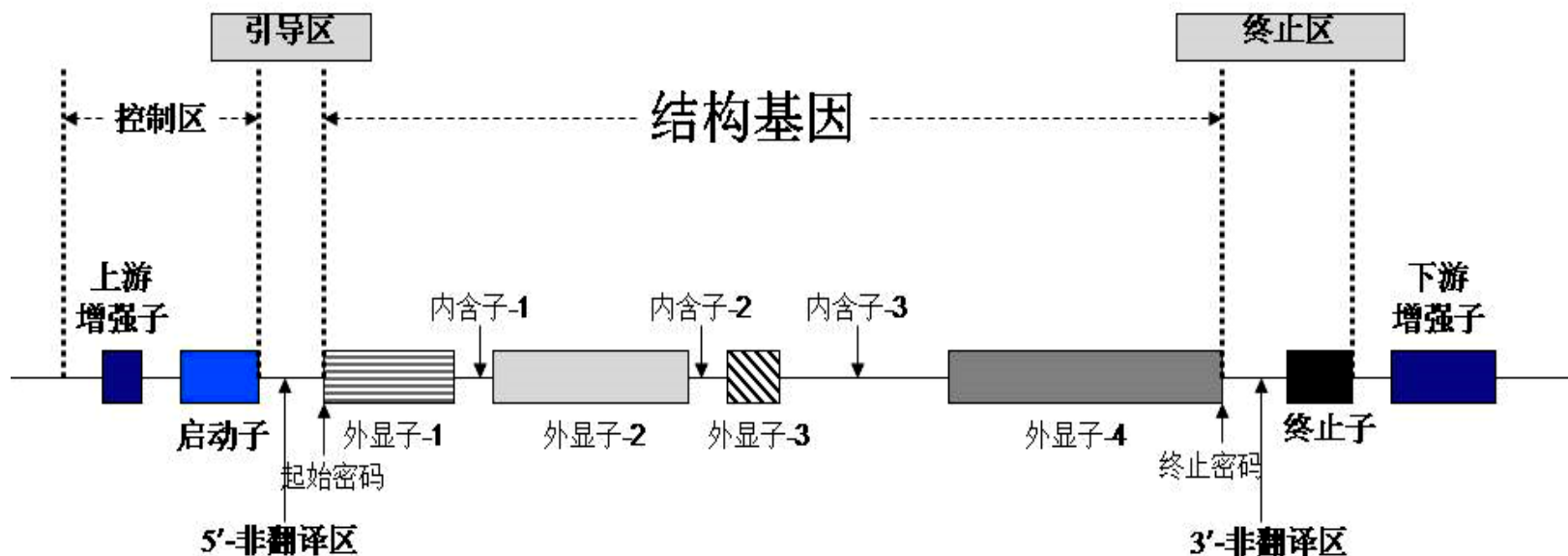
- 位于基因外显子之间的**DNA**序列
- 能体现在基因的初级转录产物中
- 在**mRNA**成熟过程中被剪切去除

### 3. 真核基因的转录单位：

- 一般由一个结构基因及其调控序列组成
- 调控序列：启动子（promoter）  
增强子（enhancer）

综上，

真核基因的结构特点：



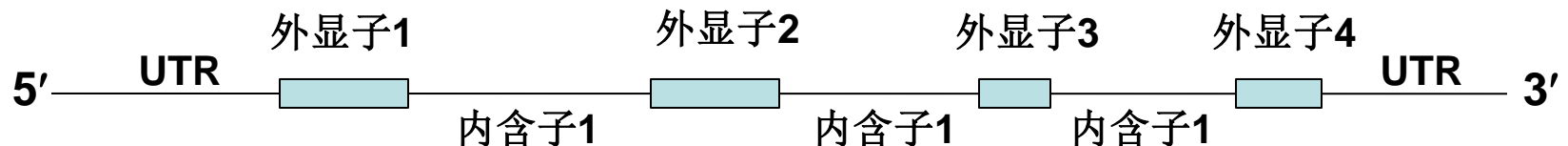
## (二) 真核基因的表达特点

真核生物有细胞核，转录和翻译在不同空间进行。

### 1. 转录在细胞核中

初级转录产物：

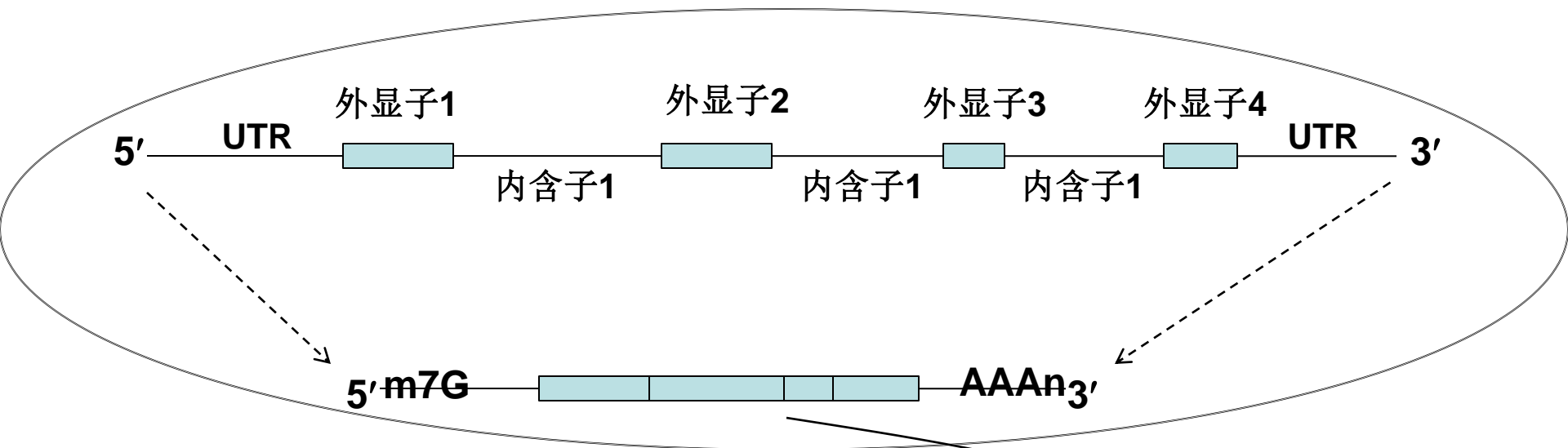
- 是从转录起始位点一直到终止位点之间的全部基因序列
- 可以写成 “5'-UTR-外显子-内含子交替排列-3'-UTR”



## 2. mRNA需要剪接和运输

### RNA剪接（RNA splicing）：

- 去掉内含子并重新连接起来的过程



### 真核生物mRNA：

- 5'端有帽结构
- 3'端有poly(A)尾
- 成熟mRNA需要从核里被运输到细胞质中



### 3. 翻译在细胞质中

- 成熟mRNA在核糖体作用下翻译成蛋白质
- 蛋白质经过加工、折叠、运输及定位变成具有生物学活性的蛋白质

总之，

真核基因的表达过程：

- DNA变构
- RNA合成
- RNA剪接及出核
- 蛋白质合成及加工修饰

## 二、真核基因表达的转录前调控

### (一) 基因重排和基因扩增对转录的影响

#### 1. 基因重排 (gene rearrangement)

- 是基因在转录前DNA序列被重新排列的一种调控方式

例如：

- 免疫球蛋白 (Ig) 的VDJ重排

#### 2. 基因扩增 (gene amplification)

- 是通过增加基因在基因组上的数目达到增加基因表达量的一种调控方式
- 是许多癌基因表达量增加的机制之一

## (二) 染色质变构对基因转录的影响

真核细胞的染色质两种状态：

- 活化状态
- 异染色质化状态

染色质活化时，

基因组DNA和组蛋白的结合变松散

- 利于转录因子接近
  - 利于双链DNA解链
- 促进基因的转录

异染色质化时，

染色质凝集成致密结构

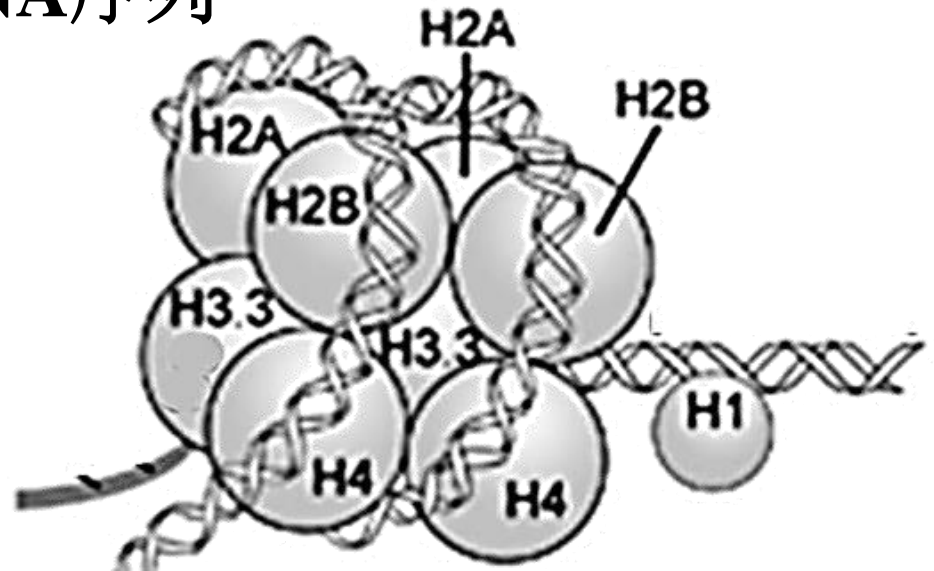
- 不利于转录因子靠近
  - 不利于双链DNA解链
- 抑制基因的转录

核小体（nucleosome）：

- 是组成真核生物染色体的基本单位

核小体的组成：

- 长度为200bp左右的DNA序列
- 四对组蛋白
- H1组蛋白



改变核小体结构：

- 是基因表达调控的一种特殊方式
- 例如：改变核小体在基因启动子区的排列，从而影响启动子的易接近性。

### (三) 表观遗传修饰对基因转录的影响

表观遗传 (epigenetic) :

- 在DNA序列不变情况下通过修饰改变基因功能的遗传现象

表观遗传修饰:

- DNA甲基化修饰
- 组蛋白的修饰
- 印记基因 (imprinted gene)

#### 1. DNA甲基化修饰

- 是最常见的表观修饰方式
- 甲基化程度越高, 基因的转录活性越低
- 绝大多数发生在CG序列中
- 哺乳细胞基因组DNA中2%~7%的胞嘧啶在5-C上甲基化

## 2. 组蛋白修饰

- 组蛋白的修饰可以直接影响核小体的结构，从而影响基因的转录。
- 最常见的是乙酰化修饰和甲基化修饰
- 组蛋白在N-端修饰的规律性被定义为组蛋白密码  
(histone code)

## 3. 印记基因

- 表观遗传修饰相当于在基因上打上印记
- 通过修饰被打上标记的基因称作印记基因 (imprinted gene)
- 在生物个体发育过程中扮演了重要角色
- 也可以解释环境因素影响基因表达的分子机制

# 三、真核基因表达的转录调控

两类：

- 发育调控
- 瞬时调控

**发育调控（developmental regulation）**

- 是指真核生物在自身生长、发育、分化等阶段对基因表达按“预定”和“有序”程序进行的调控。
- 不受环境因素影响
- 不可逆

**瞬时调控（transient regulation）**

- 是指真核生物在内外环境刺激下所进行的适应性转录调控，是可逆过程。

# 真核基因转录调控的三种因素：

- RNA聚合酶
- 顺式作用元件
- 反式作用因子

## 1. RNA聚合酶

### （1）特点

- ①能识别基因的启动子，不能与启动子结合
- ②借助蛋白质复合物间接结合启动子

### （2）类别

- RNA聚合酶I
- RNA聚合酶II → 负责转录能编码蛋白质的mRNA
- RNA聚合酶III

## 2. 顺式作用元件（cis-acting element）

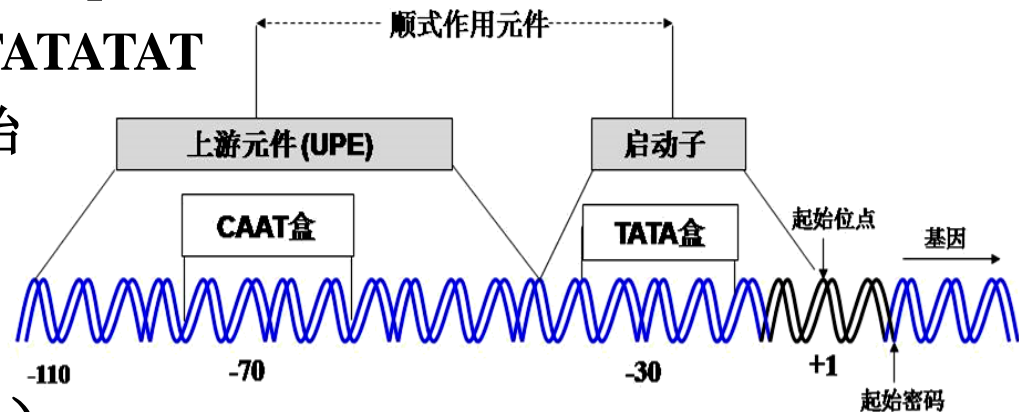
### （1）启动子（promoter）

#### 1) II类基因启动子的结构特点

将RNA聚合酶II转录的蛋白质基因称作II类基因

##### ①TATA盒

- 位于转录起始位点上游-25 ~ -30bp
- 富含AT序列：TATAAAA或TATATAT
- 负责控制基因转录的精确起始



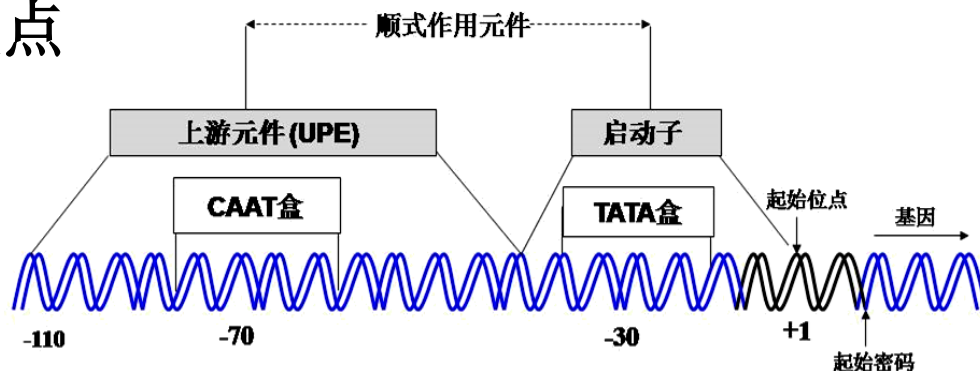
##### ②上游启动子元件（UPE）

- 在TATA盒上游-30 ~ -110bp
- 有GC盒（GGGCGG）和CAAT盒（GGCCAAT）
- 负责控制基因转录的频率和强度

## 2) II类基因启动子的类型

### ①有TATA盒的启动子

- 典型的结构包括TATA盒、CAAT盒/GC盒
- 一般只有一个转录起始位点
- 转录活性较高



### ②没有TATA盒的启动子

两类:

- 一类是不含TATA盒、富含GC盒的启动子  
一般有数个分离的转录起始位点
- 另一类是不含TATA盒、也不含GC盒的启动子  
一般有一个或多个转录起始位点，大多数转录活性很低/无  
只在胚胎发育、组织分化或再生过程中发挥转录活性

## (2) 增强子 (enhancer)

### 1) 增强子的特点

- ①位置灵活不固定
- ②无方向性但有相位性
- ③有细胞或组织特异性但无基因特异性
- ④本身没有转录活性

### 2) 增强子的作用机制

环化学说:

- 增强子与细胞内的增强子因子结合
- 引起DNA构象发生变化
- 变构DNA折叠成环
- 增强子从空间上靠近启动子, 从而增强启动子的转录活性

### 3. 反式作用因子（trans-acting factor）

#### （1）分类

#### 1) 通用转录因子（general transcription factor） 或基本转录因子

- 识别启动子TATA盒
- 是RNA聚合酶结合启动子所必需的蛋白质  
如：转录因子IIA（TFIIA）、TFIIB等

#### 2) 转录调控因子

- 识别上游启动子元件
- 能对基本转录因子其增效作用  
如：SP1与GC盒结合可提高10~25倍转录效率

## **(2) 结构特点**

- **DNA结合结构域 (DNA binding domain, DBD)**
- **转录激活结构域 (transcription activation domain, TAD)**

### **1) DNA结合结构域**

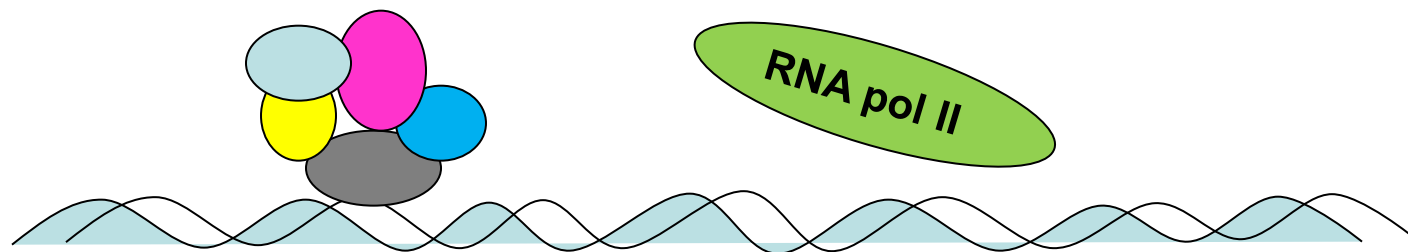
- **是反式作用因子与DNA结合的一段肽链**
- **一般由60~100aa组成**
- **有4种形式:**
  - ① **锌指结构 (zinc finger)**
  - ② **亮氨酸拉链 (leucine zipper)**
  - ③ **螺旋-转角-螺旋 (helix-turn-helix)**
  - ④ **螺旋-环-螺旋 (helix-loop-helix)**

## 2) 转录激活结构域

- 是蛋白质-蛋白质相互作用的结构基础
- 一般由30~100aa组成：
  - 酸性激活结构域
  - 富含谷氨酰胺结构域
  - 富含脯氨酸结构域

## 真核基因的转录

- 蛋白质复合物为RNA聚合酶提供“脚手架”



因此，反式作用因子：

- 不一定都有DNA结合结构域
- 应具备蛋白质相互作用结构域

### (3) 活化方式

#### 两类活化方式:

- 天然具有活性的活化方式
- 需要诱导的活化方式

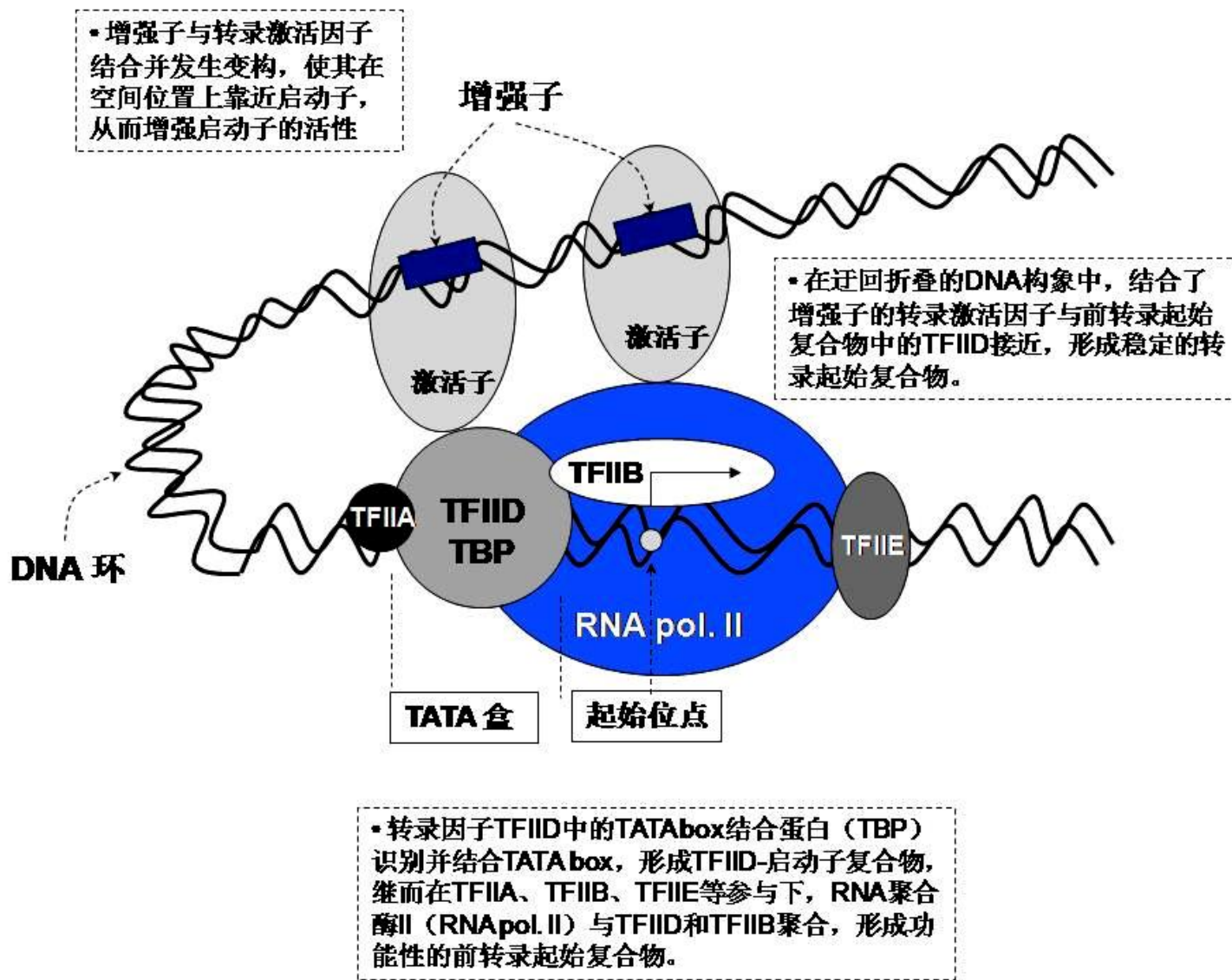
#### 绝大多数反式作用因子

- 通过蛋白质-蛋白质相互作用形成二聚体或多聚体之后才能与DNA序列结合

#### 二聚体 (dimer)

- 是反式作用因子结合DNA时最常见的分子结构
- 两个相同分子组成的二聚体为同源二聚体 (homodimer)
- 两个不同分子组成的二聚体为异源二聚体 (heterodimer)
- 异源二聚体结合DNA的能力更强

## 4. 转录调控中三种因素的相互作用



# 转录调控模式:

- 正调控模式
- 负调控模式

## (1) 正调控模式

- 1) DNA成环靠近RNA聚合酶结合位点促进转录
- 2) 反式作用因子使DNA变构促进转录
- 3) 反式作用因子沿着DNA滑动促进转录
- 4) 反式作用因子的连锁反应促进转录

## (2) 负调控模式

- 1) 抑制性反式作用因子与活化性反式作用因子的DNA结合位点有部分重叠，从而竞争抑制基因转录
- 2) 抑制性反式作用因子直接与转录因子结合，使转录因子与RNA聚合酶的结合位点被封闭，抑制基因转录

# 四、真核基因表达的转录后调控

## 1. RNA剪接的调控作用

两种剪接方式：

### 1) 组成型剪接（**constitutive splicing**）

指剪接后的外显子按照编码蛋白质的顺序规范拼接

一个基因产生一种成熟mRNA

产生一种肽链

### 2) 可变剪接（**alternative splicing**）

指剪接后的外显子以不同方式拼接

一个基因产生一种以上成熟mRNA

产生一种以上肽链

## 2. 5'端加帽 (capping)

5'端帽结构：7-甲基鸟苷三磷酸 ( $m^7G$ )

帽结构的作用：

- 保护mRNA免受外切核酸酶的降解
- 为蛋白质合成提供识别标志

3'端加多聚腺苷酸尾 (polyA tail)

加尾信号：AAUAAA

polyA尾的作用：

- 帮助mRNA从细胞核进入细胞质
- 防止外切核酸酶水解mRNA
- polyA尾的长短可影响mRNA半衰期，从而调节基因表达



# 五、真核基因表达的翻译和翻译后调控

## （一）翻译起始的调控

### 1. AUG旁侧序列对翻译起始的影响

**Kozak序列（Kozak sequence）：**

**(1) AUG旁侧的序列在不同位置有不同作用**

**+4G、-6G：对翻译起始非常重要**

**-1、-2位：序列不保守，对kozak序列的强度有影响**

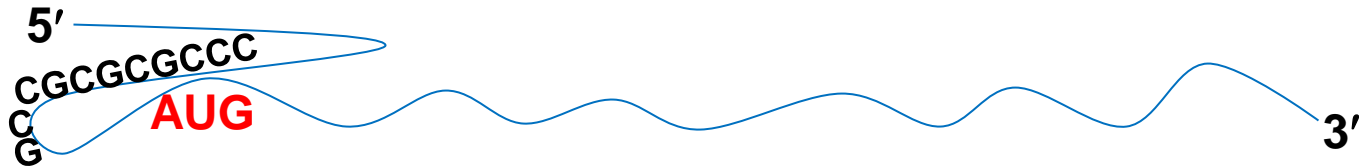
**(2) 不同真核生物的kozak序列有一定差异**

**共有序列：(gcc)gccRcc**AUGG**（-3R可以是A或G）**

## 2. mRNA 5'-非翻译区对翻译起始的影响

富含GC序列:

- 易形成环状结构
- 若AUG位于环中，翻译起始受抑



富含AT序列:

- 不易成环
- AUG暴露充分，有利于翻译起始



### **3. mRNA本身结构对翻译起始的影响**

**无义介导的mRNA降解（nonsense-mediated mRNA decay, NMD）**

- **检测mRNA的分子结构**
- **若发现终止密码提前出现（premature termination codon, PTC）**
  - 1) **启动NMD的降解机制**
  - 2) **诱导mRNA降解**
  - 3) **需外显子拼接复合体（exon-junction complex, EJC）**

**EJC:**

- **在RNA剪切和拼接中黏附在RNA分子上**
- **若基因突变，EJC会出现在RNA的错误位置上**
- **错误位置上的EJC诱导NMD，使mRNA降解**

## **（二）非编码RNA的调控作用**

### **非编码RNA（non-coding RNA, ncRNA）**

- 指不具有可读框或编码蛋白质功能的RNA分子
- 能在翻译水平调控基因的表达

### **非编码RNA的类型：**

- 非编码小RNA（small non-coding RNA, scnRNA）
- 微小RNA（microRNA, miRNA）
- 长链非编码RNA（long-chain non-coding RNA, lncRNA）

# 1. 非编码RNA的调控机制

- 诱导mRNA降解
- 位阻效应干扰蛋白质的翻译

例如：

- 靶向启动子调节基因的转录起始
- 靶向mRNA的3'-非翻译区抑制蛋白质的翻译

长链非编码RNA（lncRNA）：

- 参与表观遗传修饰，包括组蛋白修饰和DNA甲基化

例如：

一个与*p53*基因互补的lncRNA

- 通过改变组蛋白甲基化，抑制*p15*基因的表达

## 2. miRNA与转录因子的比较

	miRNA	转录因子
调控多个基因表达	靶向不同mRNA	靶向不同顺式作用元件
识别位点的亲和性	RNA结合蛋白对位点的识别	依赖核小体对位点的覆盖率
调节效应	一般是阻遏调控	在转录水平对基因进行正、负调控
自身表达	细胞特异性表达受转录因子的调控	本身的基因表达受某些miRNA的调控
靶序列长度	靶序列平均长度<1000nt，mRNA的3'UTR	靶序列一般可覆盖几个万碱基序列
调控对象及方式	调控发生在mRNA水平，调节速度快，具有可逆性	调控发生在DNA水平，一般是不可逆的

### （三）翻译后水平的调控

主要指

- 蛋白质本身的各种加工修饰及折叠剪切等

如：

- 氨基酸的糖基化、磷酸化、乙酰化等修饰
- 蛋白质的折叠
- 信号肽的剪切

## 思考题：

1. 解释结构基因和调节基因，叙述真核基因的结构特点。
2. 解释基因表达的概念，举例说明基因表达的组织特异性和阶段特异性。
3. 解释操纵子的基本结构特点，说明乳糖操纵子的工作原理。
4. 解释基因转录单位在原核和真核生物的异同。
5. 解释顺式作用元件和反式作用因子的概念，举例说明二者之间相互作用的基本特点。
6. 说明基因表达的多级调控特点。
7. 说明miRNA在基因表达调控中的作用。
8. 叙述表观遗传在基因表达调控中的作用。