

第三节

真核基因表达调控

Regulation of Gene Expression in
Eukaryotes

- 1970年，猿猴病毒40(simian virus 40, SV40) 的生命周期被确定为早、晚两个阶段，使它们成为基因表达调控的典型模型。
- 1975年D. Pribonow发现T7启动子序列。
- 1982年C. Benoist和P. Chambon揭示了SV40的早期启动子序列。
- 1983年，W.S.Dynan和R.Tjian分离、鉴定了真核转录因子AP1。
- 1986年，J.Clarke和Chambon两个实验室同时报告了SV40增强子的多功能元件组成。
- 1990年代，信号转导-基因转录偶联机制成为热点课题。

真核生物基因表达特点

(一) 真核基因组结构庞大

- 哺乳类动物基因组DNA 约 3×10^9 碱基对。
- 人类编码基因约2.3~2.5万个
- rDNA等重复基因约占5%~10%。

(二) 真核基因转录产物为单顺反子

(三) 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 ($10 \sim 10^4$ 次)

单拷贝序列 (一次或数次)

(四) 真核基因中存在非编码序列和间隔区，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。

(五) 真核细胞内含有多种RNA聚合酶

真核RNA聚合酶有三种，即RNA pol I、II及III，分别负责三种RNA转录。

(六) 在真核细胞中转录与翻译分隔进行

真核细胞有细胞核及胞浆等区间分布，转录与翻译在不同细胞部位进行，转录在细胞核，翻译在细胞浆。因此，转录与翻译产物的分布、定位等环节均可以被调控。

(七) 转录后修饰、加工更为复杂

主要内容

真核基因表达调控的特征

真核生物染色质水平的基因表达调控

真核生物转录水平的基因表达调控

真核生物翻译水平的基因表达调控

真核基因表达调控的特征

基因表达的多级调控

基因激活

拷贝数

重排

甲基化程度

转录起始

转录后加工

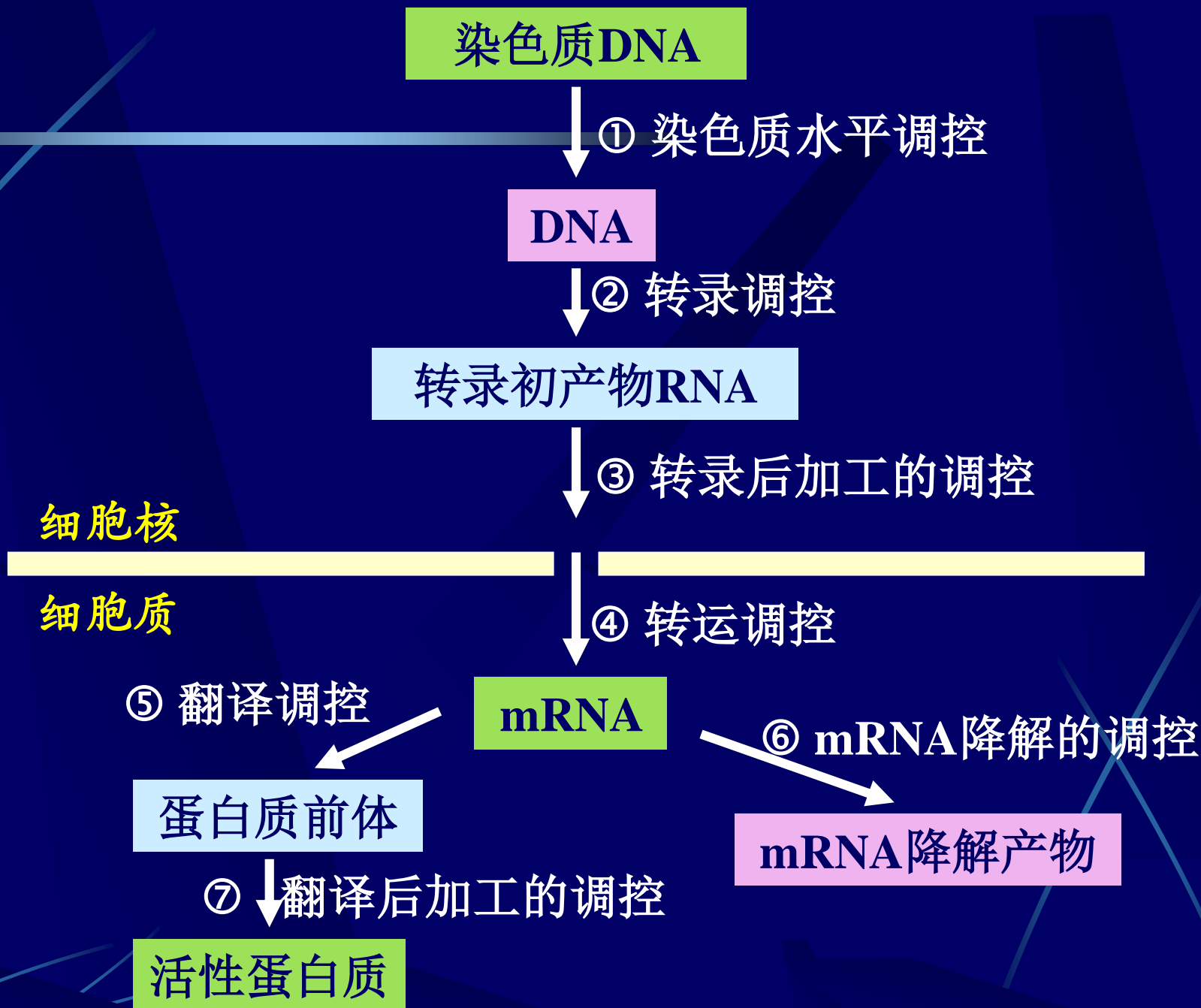
mRNA降解

蛋白质翻译

翻译后加工修饰

蛋白质降解等

真核基因表达的多级调控



在真核基因表达调控中以正性调节占主导

- 采用负性调节不经济
- 采用正性调节机制可提高特异性：一个负性调节元件的结合足可阻断RNA聚合酶的结合，因此同时采用几个负性调节元件一般不会改变特异性；相反，如果采用多种正性调节元件、正性调节蛋白可提高基因表达调节的特异性。
- 采用正性调节机制更精确：在正性调节中，如果基因不结合调节蛋白，则没有活性；只要细胞表达一组激活蛋白时，相关靶基因即可从多个层次依次进入转录激活状态。

染色质水平的 基因表达调控

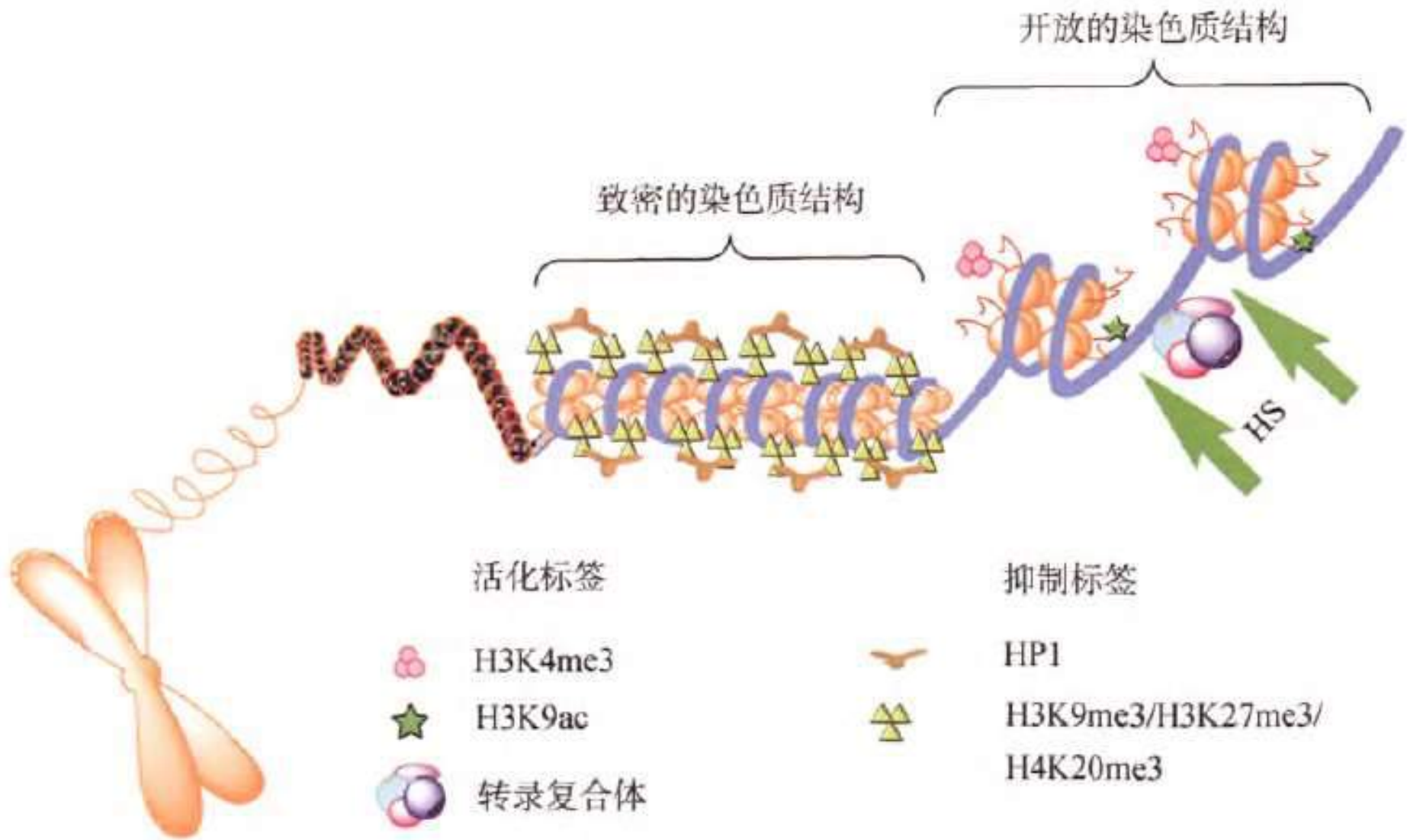
**Chromosomal Regulation of
Eukaryotic Gene Expression**

(一) 染色质结构与基因活性

转录活化染色质与非活化染色质在结构上有很大不同。

- **常染色质 (euchromatin)** 染色体的大部分在间期时松开分散在核内，松散的染色质中的基因可以转录
- **异染色质 (heterochromatin)** 染色体中的某些区段到分裂期后不像其他部分解旋松开，仍保持紧凑折叠的结构，在间期核中可以看到其浓集的斑块

- **常染色质**中的转录活性区域对核酸酶敏感，特别是转录基因的5'-侧翼区1000 nt以内是高敏感位点(HS, hypersensitive sites)。很多高敏感位点是调节蛋白质的结合序列，而这些区域核小体的相对缺少也使得这些蛋白质易于与之结合。
- **异染色质**是转录非活性的，然而最近研究表明异染色质中也会发生有限的转录。



HS: 核酸酶高敏感位点

二、染色质重塑 (chromatin remodeling)

基因活化蛋白质可以通过改变基因的启动子和调节序列区域的染色质结构来促进转录开始。这种改变局部染色质结构的过程被称为**染色质重塑**。

- **最主要的三种方式：**

组蛋白的共价修饰

核小体重塑 (nucleosome remodeling)

DNA的甲基化

(一) 组蛋白的共价修饰

- 在5种组蛋白中，H1的N端富含疏水氨基酸，C端富含碱性氨基酸，H2A、H2B、H3和H4都是N端富含碱性氨基酸(如精氨酸、赖氨酸)，C端富含疏水氨基酸(如缬氨酸、异亮氨酸)。
- 组蛋白N端的15-38个氨基酸残基是翻译后修饰的主要位点，可同其他调节蛋白和DNA相互作用，调节DNA的生物学功能。

转录活化染色质与非活化染色质的组蛋白共价修饰的方式不同。

- 核小体的核心组蛋白（H2A，H2B，H3，H4）中赖氨酸残基的非可逆性甲基化，丝氨酸与苏氨酸残基的磷酸化，乙酰化以及泛素化多是转录活性染色质的特点。

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甲基化所必需
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	装配
H4	Lys-12	乙酰化	装配
H4	Lys-16	乙酰化	核小体装配
H4	Lys-16	乙酰化	Fly X激活

●组蛋白乙酰化

时间：基因活化蛋白质结合在转录调节区域之后

酶：组蛋白乙酰基转移酶（histone acetyl transferases, HAT），也称组蛋白乙酰化酶（histone acetylase）

位点：核小体的核心组蛋白所富含的赖氨酸残基，大多在组蛋白**H3赖氨酸**的9、14、18、23和**H4赖氨酸**5、8、12、16等位点。组蛋白乙酰化是可逆的动态过程，HAT将乙酰辅酶A的乙酰基部分转移到核心组蛋白氨基末端特定赖氨酸残基的 ϵ -氨基基团上

●组蛋白乙酰化

作用：

1. 导致组蛋白表面正电荷减少，组蛋白与DNA结合能力下降，引起核小体解聚并阻止核小体装配，使得染色体处于松弛状态，从而使转录因子和RNA聚合酶顺利结合在基因的DNA上，促进基因转录
2. 组蛋白乙酰化是许多转录调控蛋白相互作用的一种“识别信号”，如H4组蛋白的乙酰化参与了指示和吸引TFIID到相应的启动子，促进转录前起始复合物的装配
3. 在细胞分裂期，组蛋白的乙酰化参与细胞周期和细胞分裂的调控

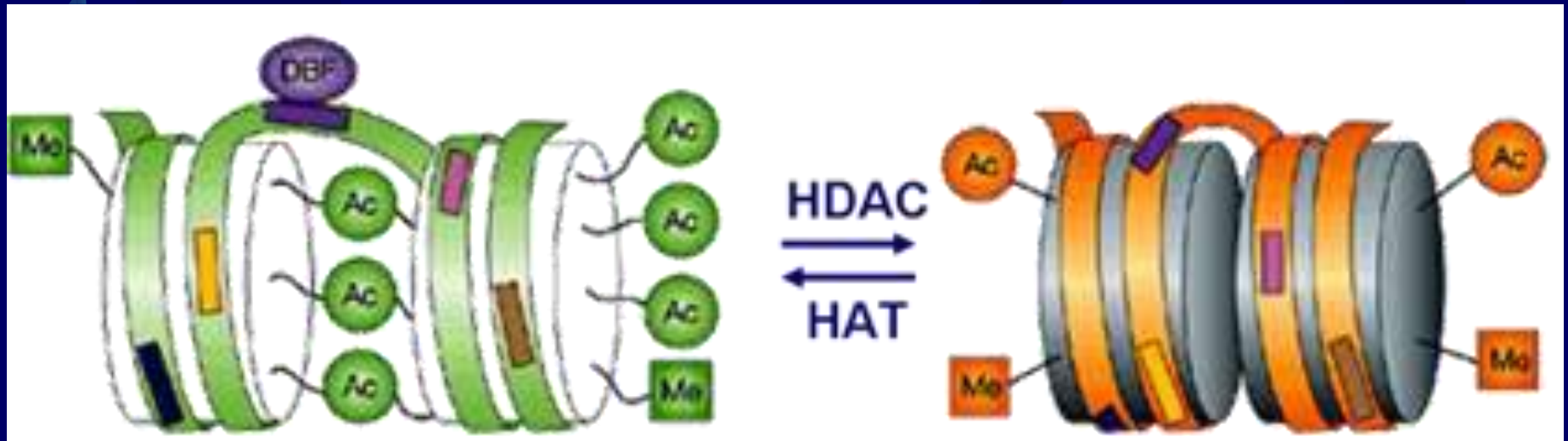
●组蛋白乙酰化

逆转：组蛋白去乙酰化酶(histone deacetylase, HDAC)减少核小体的乙酰化，使染色质恢复转录非活性状态。

组蛋白去乙酰化酶移去组蛋白赖氨酸残基上的乙酰基，恢复组蛋白的正电性，带正电荷的赖氨酸残基与DNA分子的电性相反，增加了DNA与组蛋白之间的吸引力，使启动子不易接近转录调控元件，从而抑制转录。

转录活性状态

转录非活性状态



组蛋白去乙酰化酶（histone deacetylase, **HDAC**）

组蛋白乙酰基转移酶（histone acetyl transferases, **HAT**）

- **意义：**组蛋白乙酰化/去乙酰化循环参与了真核基因组整体表达水平的调控，基因组具备较低的基础乙酰化水平有两个明显的好处：一是方便地开启或关闭基因转录，二是实现基因组水平上复制和**DNA**损伤修复在内的整体调控。

●组蛋白甲基化

- 主要位于核小体核心组蛋白H3和H4外周结构域氨基末端的赖氨酸和精氨酸残基。
- 由组蛋白甲基转移酶（histone methyltransferase, HMT）催化。
- 组蛋白甲基化可改变染色质结构，参与基因表达调控。

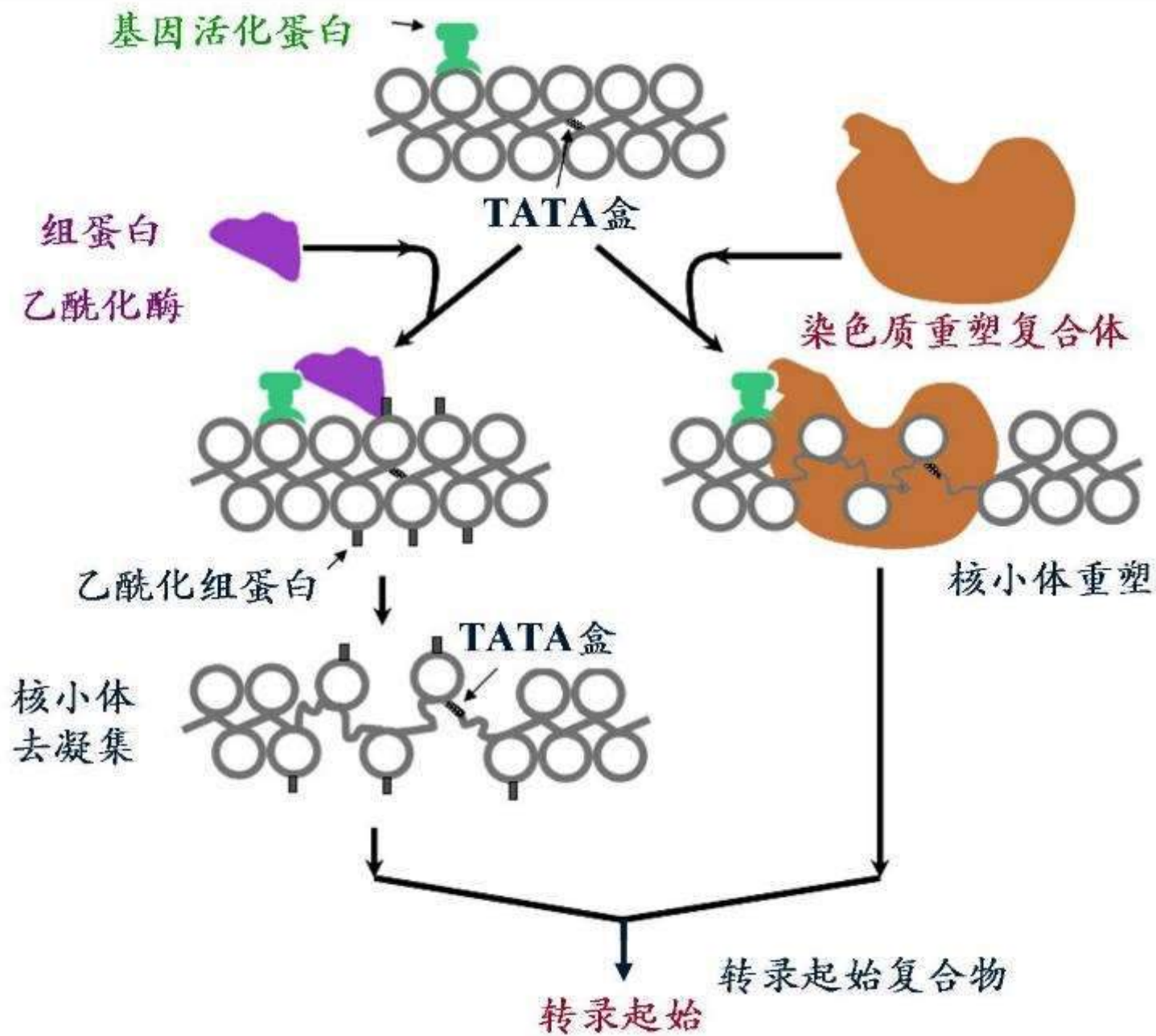
（二）核小体重塑

- 核小体在染色体上的精确定位有两方面的作用：
 - 1。提供一个框架，使转录因子之间的信息传递更有效
 - 2。染色体结构的不均一性，即它的某些区域核小体结构变化，转录因子易于接近染色质模板，启动转录过程。
- 虽然基因编码区的核小体对基因影响较小，但核心启动子区核小体结构是否改变将直接影响转录的起始。
- 研究表明，核小体占据的位置是基因序列中的特殊部位，比如启动子的部位，而增强子则往往处于核小体末端或核小体之间的**DNA** 区域。

核小体重塑

- 组蛋白H3和H4的修饰直接影响核小体的结构，并为其它蛋白提供了和DNA作用的结合位点。
- 染色质重塑因子复合物利用ATPase和解旋酶活性来改变核小体在DNA上的位置。ATP依赖的染色质重塑可以使与核小体结合的DNA暴露出来，使核小体沿着DNA滑动并重新分布，在改变单个核小体结构的同时改变染色质的高级结构，从而在DNA修复、重组、复制及转录过程中调节全基因组的柔顺性和可接近性。

染色质重塑



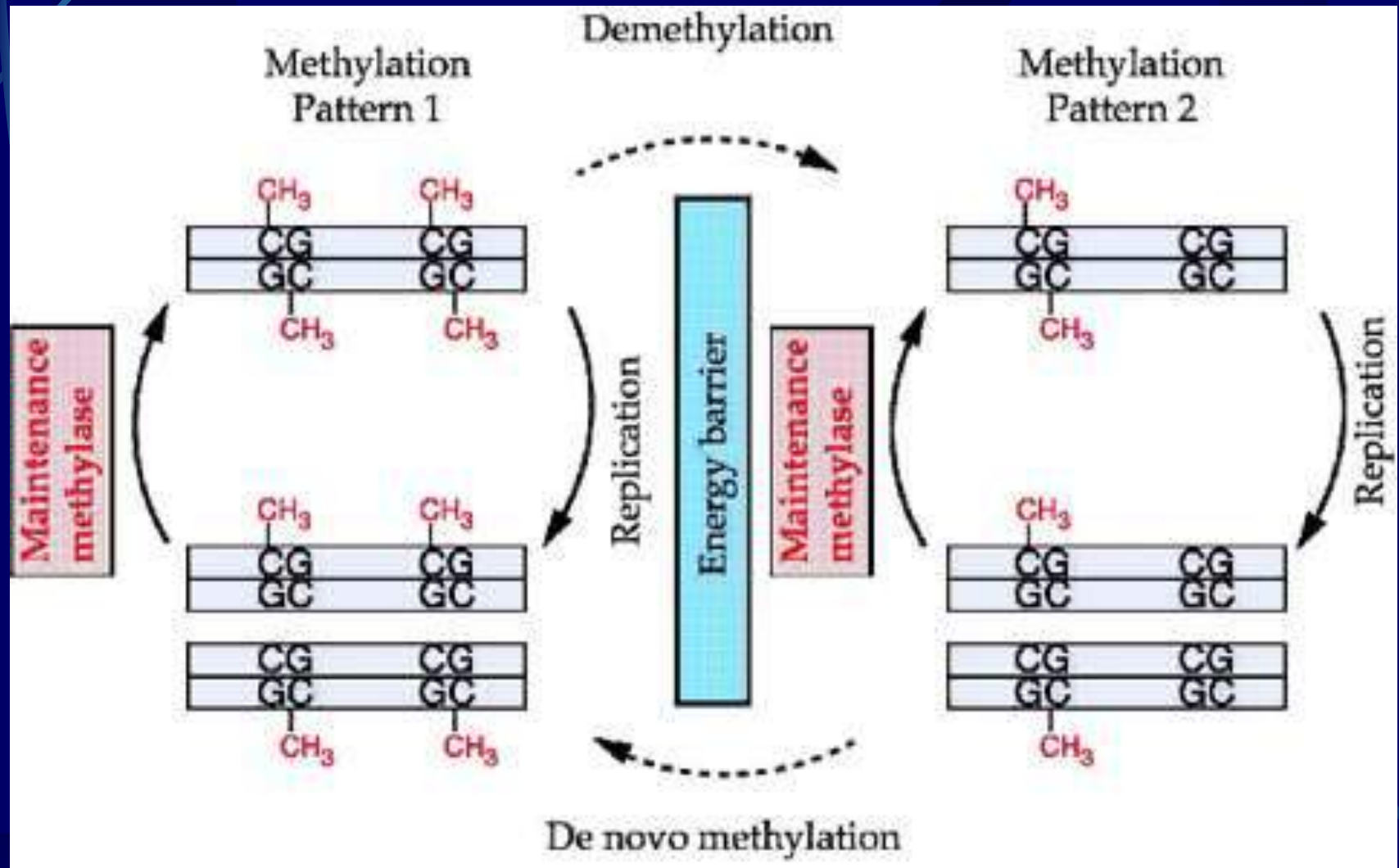
(三) DNA甲基化

- **位点：** CpG序列胞嘧啶的第5位碳原子
- **酶：** 在**DNA甲基转移酶 (DNA methyltransferases, DNMTs)**作用下，以**S-腺苷甲硫氨酸(SAM)**为甲基供体，甲基基团合成到胞嘧啶的第5位碳原子上，形成**5-甲基胞嘧啶(5-methylcytosine, 5-mC)**。



- DNA甲基转移酶分两种：一种是**维持甲基化酶**，另一种是**重新甲基化酶**。
- 在细胞分化的过程中，基因的甲基化状态将遗传给后代细胞。

DNA分子胞嘧啶的甲基化修饰是一个动态的过程



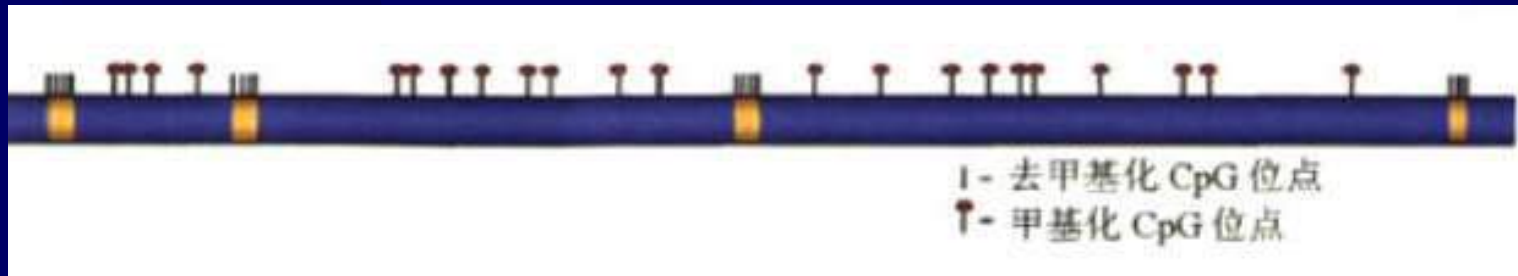
CpG岛

- 在DNA中，有些CpG序列呈局部聚集分布，形成GC含量较高、CpG双核苷酸相对集中的区域，称为**CpG岛**。
- CpG岛分为两类：
 - ① 转录起始点附近的CpG岛
 - ② 非转录起始点附近的CpG岛，多数位于高度重复序列附近

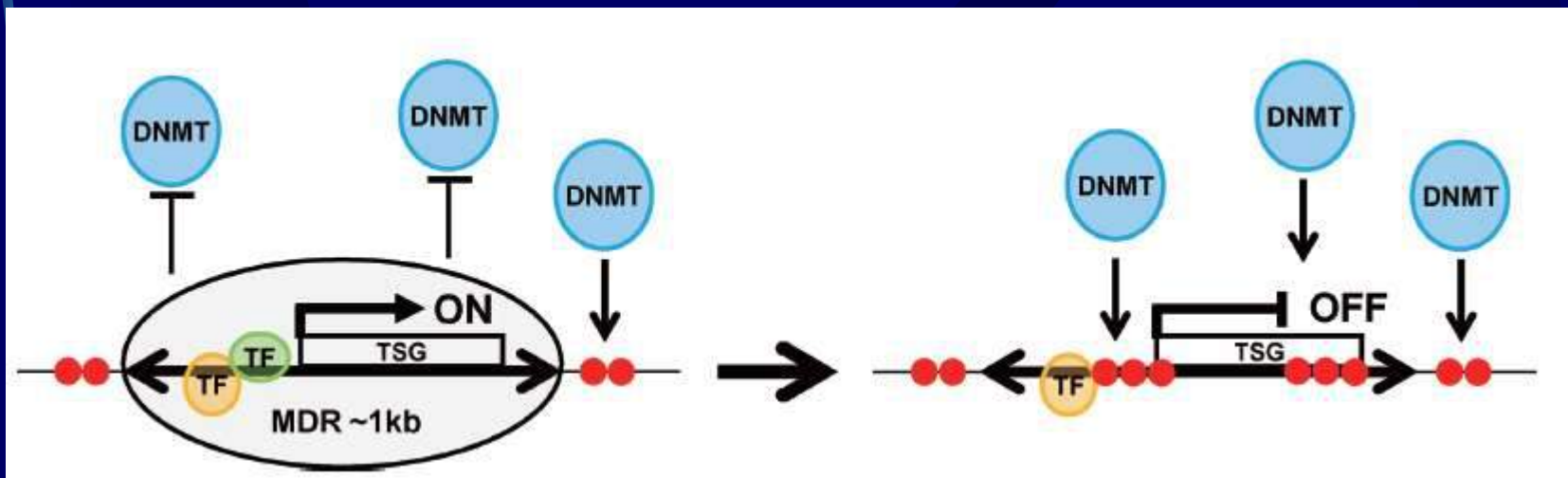
转录活性基因启动子区甲基化程度低

- 转录起始点附近的CpG岛通常位于基因的启动子区或第一个外显子区。
- 健康人基因组中，CpG岛中的CpG位点通常是处于非甲基化状态，约有5%胞嘧啶被甲基化为5-甲基胞嘧啶，而在CpG岛外的CpG位点则通常是甲基化的。
- 调控区CpG岛甲基化的范围、程度与基因表达水平呈反比，转录活性染色质区域胞嘧啶被甲基化的程度降低。

- DNA 甲基化调控基因表达的直接机制为 5-甲基胞嘧啶深入到 DNA 双螺旋的大沟中，阻碍转录因子结合到包含有 CpG 的识别位点。



抑癌基因调控区域的CpG岛高甲基化能导致其转录失活，从而促进肿瘤的发生、发展



- DNMT: DNA甲基转移酶
- TSG (tumor suppressor gene): 抑癌基因
- TF: 转录因子

非转录起始点附近的CpG岛

- 多数位于高度重复序列附近，在正常组织通常呈高度甲基化状态。
- 在肿瘤组织甲基化水平降低，降低程度与癌症的恶性程度相关。

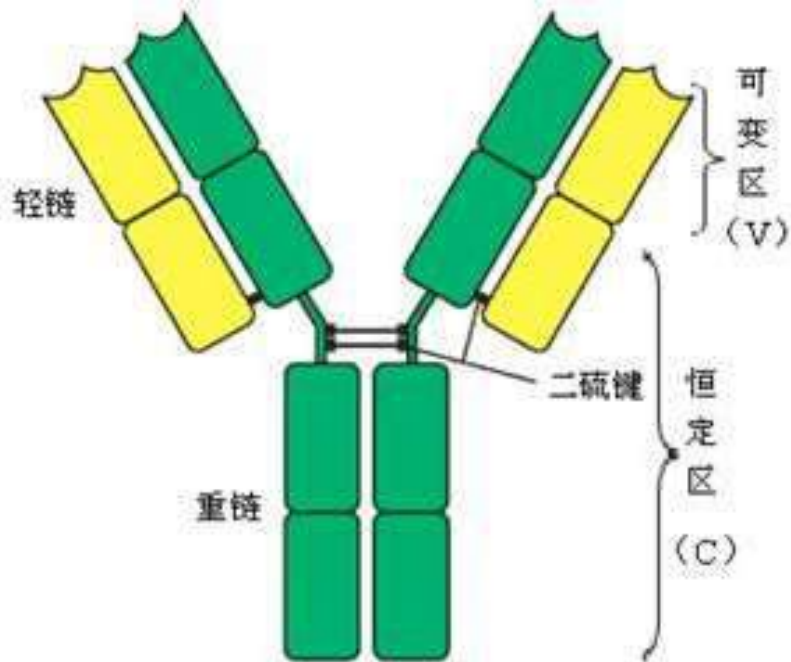
三、DNA重排与扩增

- **基因扩增**：细胞中特定基因的拷贝数专一性大量增加的现象，可增加基因表达产物。

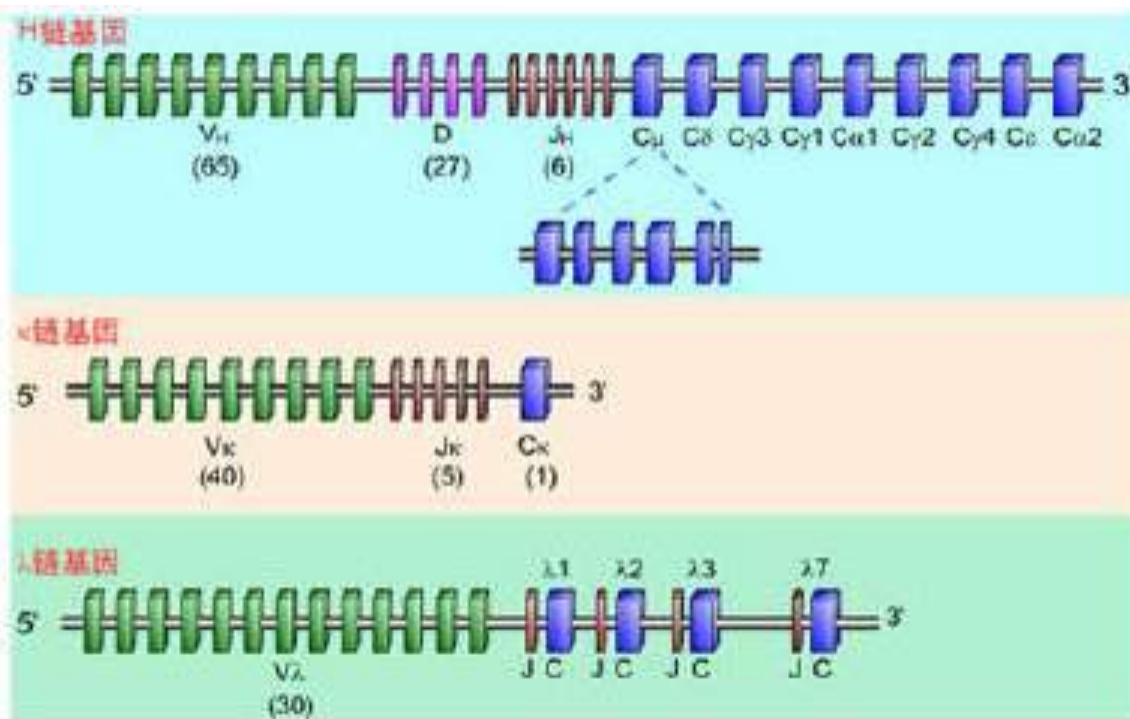
如非洲爪蟾在成熟过程中产生18S、28SrRNA的rDNA的选择性复制，使成熟卵细胞具有 2×10^6 个DNA拷贝，与二倍体细胞（含有900个拷贝）相比，扩增1000倍。

- **基因重排：**真核生物基因组中的DNA序列可发生改变而进行重新组合，重排成为一个完整的转录单位。这种重排是由特定基因的遗传信息所决定的，是某些基因调控的重要机制。

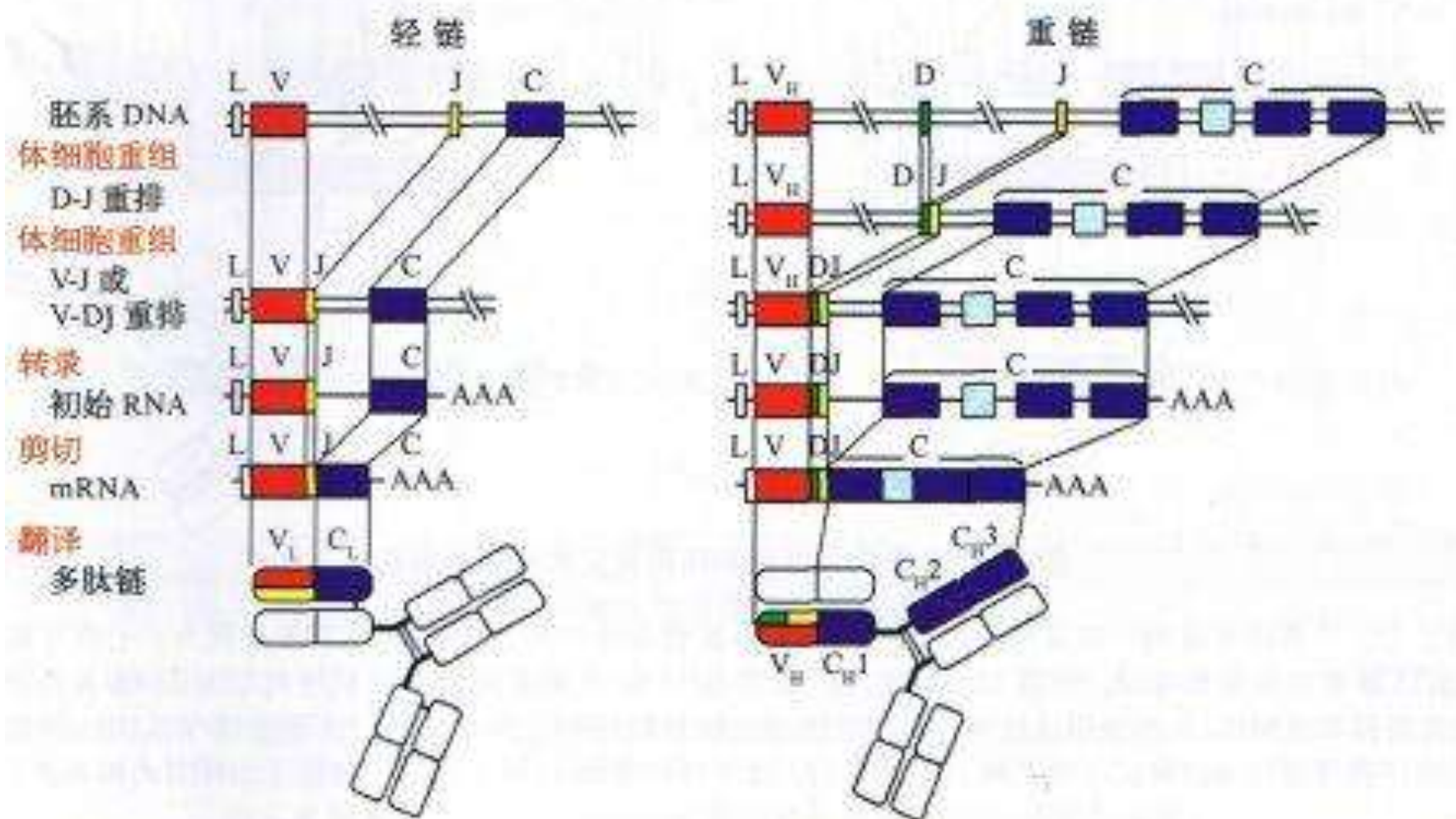
在人类基因组中，所有抗体的重链和轻链都不是由一个完整的抗体基因编码的，而是由不同基因片段经重排后形成的。



编码多肽链	基因符号(人)	基因染色体定位	
		人	小鼠
κ 轻链	Ig κ	2	6
λ 轻链	Ig λ	22	16
重链	IgH	14	12



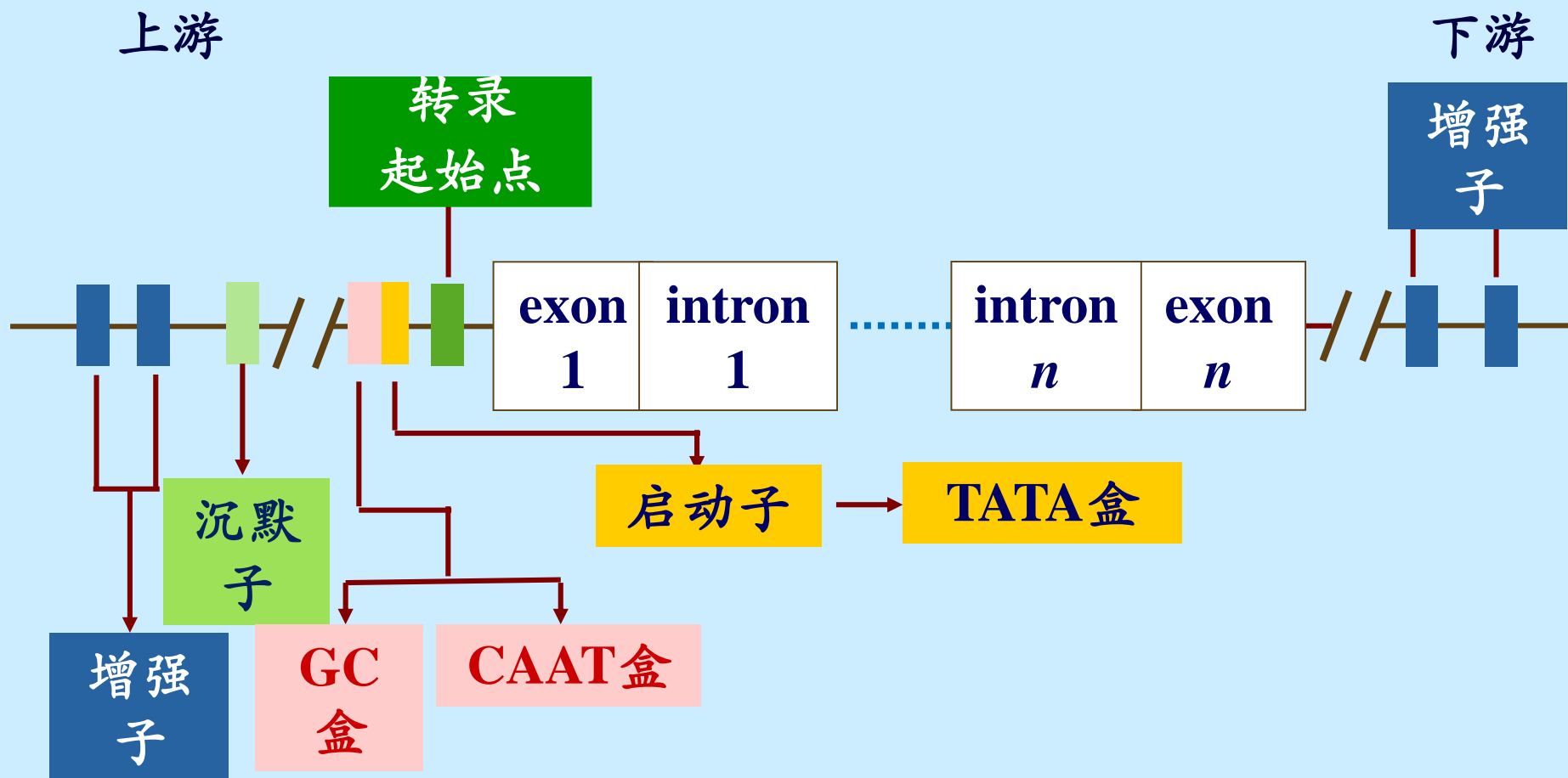
多肽链	基因片段数			V区基因重组方式	经重排的随机配对后* 推算的多样性数目	
	V		J			
H链	1000	12	4	V-D-J	4.8×10^4	4.8×10^7
κ 链	250	-	4	V-J	1.0×10^3	



转录水平的基因表达调控

**Transcriptional Regulation of
Eukaryotic Gene Expression**

一、转录起始的调控

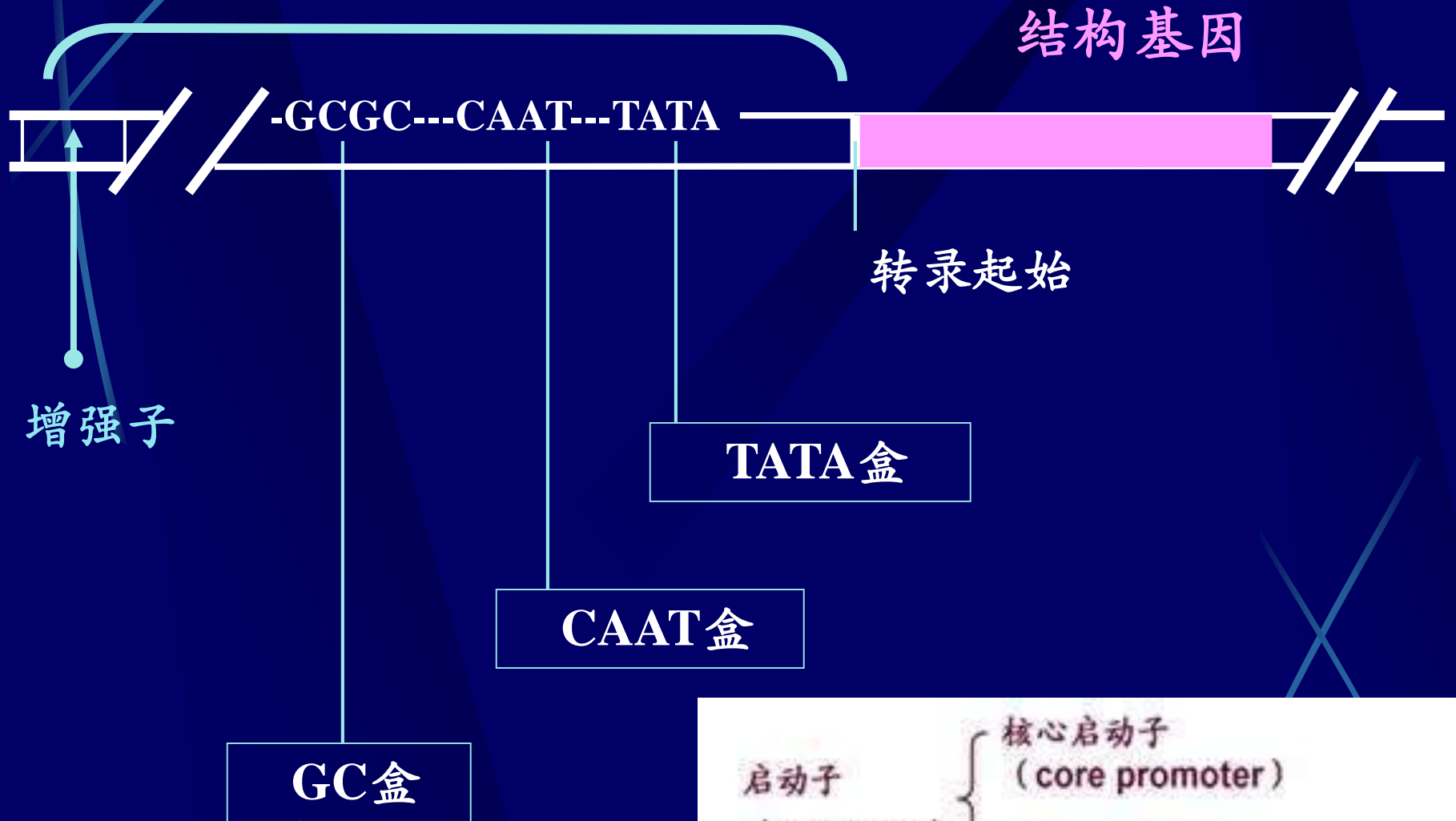


(一) 转录起始调控区

1. 启动子是RNA聚合酶结合并启动基因转录的特异DNA序列

真核基因的启动子指的是RNA聚合酶 (RNA polymerase, RNA Pol) 识别、结合的基因转录调控区中启动基因转录的一段特异DNA序列，包含一组转录调控功能组件，其中每一个功能组件的DNA序列约7~20 bp。

真核生物启动子保守序列



启动子 (promoter) { 核心启动子 (core promoter)
启动子近端元件 (promoter proximal element)

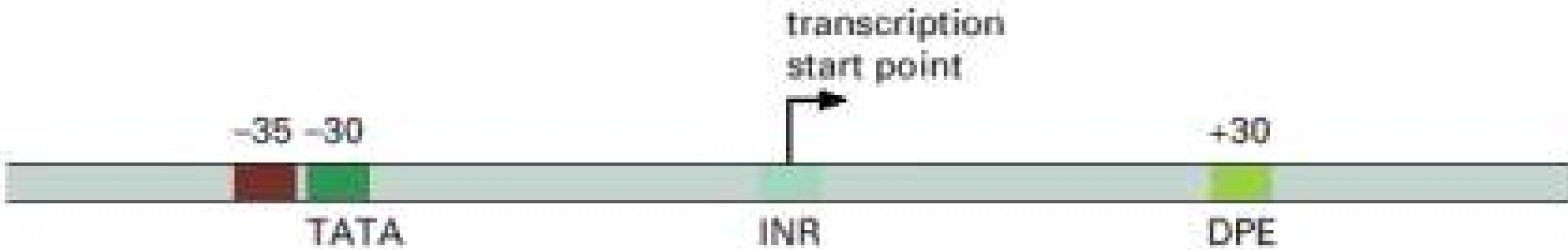
1) 近起始点的TATA盒/起始子及CpG岛是真核生物启动子的核心序列

- 真核细胞的启动子是包括转录起始点（**transcription initiation site**或**transcription start site**）在内的、有通用转录因子及RNA聚合酶组装、结合的一段DNA序列。
- 典型的启动子核心序列（**core sequences**）是在转录起始位点上游25~35 bp处，有一保守的TATA序列，被称为**TATA盒**（**TATA box**），真核细胞的TATA盒多为TATAAAA序列。TATA盒与原核细胞的启动子一样，对RNA聚合酶II的转录起始位点起定位作用。

❁ 一些真核细胞基因含有另一种启动子元件，称为起始子(initiator, Inr)，决定启动子的强度。

❁ 起始子共有序列为：(5') Y-Y-A⁺¹-N-T/A-Y-Y-Y (3')；Y代表胞嘧啶C或胸腺嘧啶T，N代表任意碱基，T/A代表T或A。

真核核心启动子共有序列



element	consensus sequence	general transcription factor
TATA	TATAA/TAA/T	TBP
INR	C/T C/T A N T/A C/T C/T	TFIID
DPE	A/G G A/T C G T G	TFIID

DPE (downstream promoter element)

- ◆有一些编码蛋白质基因的转录可有多个起始点，可产生含不同5'末端的mRNA。
- ◆这些基因大多为低转录基因，编码中间代谢酶的管家基因。
- ◆它们不含TATA盒或起始子，多在起始位点上游约100 bp内含有20~50个核苷酸的CG序列，被称做**CpG岛 (CpG island)**。

2) 启动子上游元件协助真核基因调节

- ◆ 启动子上游元件 (promoter-proximal elements, 或 upstream promoter elements, UPE) 是一些位于 TATA 盒上游的 DNA 序列, 与调节蛋白结合, 调节通用转录因子与 TATA 盒的结合、RNA 聚合酶与启动子的结合, 以及转录起始复合物的形成, 从而决定基因的转录效率与专一性。
- ◆ 常见的序列是:

{ CAAT 盒 (CAAT box)
GC 盒 (GC box)

● CAAT盒

在5'端转录起始点上游约70~100个核苷酸，碱基序列为GGCTCAATCT。CAAT盒的作用还不很肯定，一般认为它控制着转录的起始频率，而不影响转录的起始点。当这段序列被改变后，mRNA的形成量会明显减少。

● GC盒

有两个拷贝，位于CAAT盒的两侧，由GGGCGG组成，转录因子SP1能识别GC盒并且与之结合，有激活转录的作用。

真核基因启动子的典型结构

高等真核生物



酵母



2. 增强子决定基因的时空特异性表达

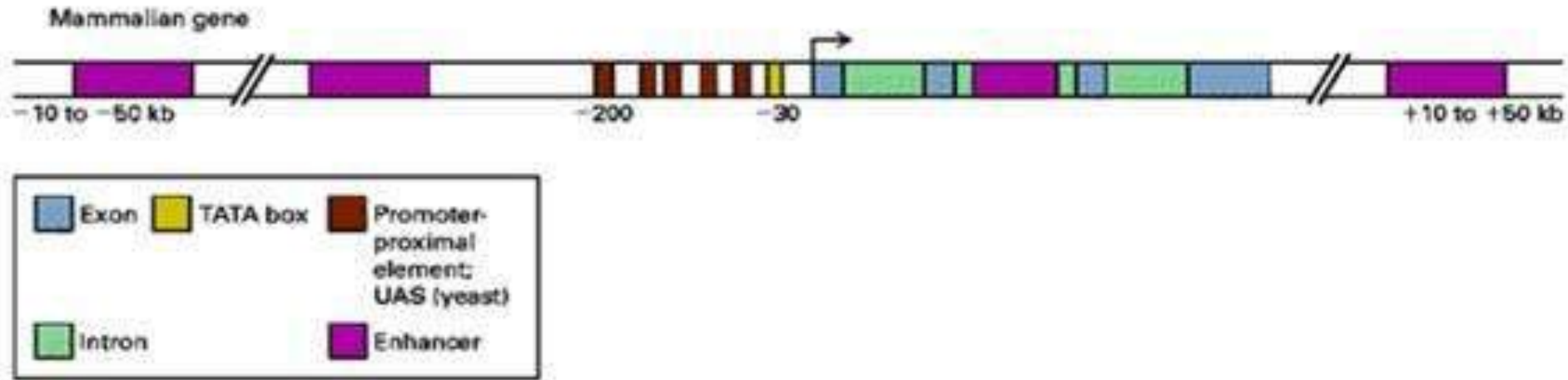
增强子 (enhancer)

是能够结合特异的基因调节蛋白，促进邻近或远隔特定基因表达的DNA序列；在酵母中，被称为上游活化序列 (upstream activator sequences, UASs)。

- 增强子的作用通常与其所处的位置和方向无关。

•增强子所处位置

- 在所调控基因的上游或下游，但主要位于上游。下游内含子当中，乃至下游最后外显子以外的序列也可含有增强子。
- 增强子距转录起始位点的距离变化很大，从100 nt到50,000 nt，甚至更大，但总是作用于最近的启动子。



很多哺乳动物基因受一个以上的增强子调控。

● 增强子的作用特点

- ① 在转录起始点5'或3'侧均能起作用
- ② 相对于启动子的任一指向均能起作用
- ③ 发挥作用与受控基因的远近距离相对无关
- ④ 对异源性启动子也能发挥作用
- ⑤ 通常具有一些短的重复顺序
- ⑥ 功能可以累加

● 增强子的种类

1. 细胞特异性增强子

能够在特定的细胞或特定的细胞发育阶段选择性调控基因转录表达的增强子。

人类胰岛素基因5'末端上游约250个核苷酸处有一组织特异性增强子，在胰岛 β 细胞中有一种特异性蛋白因子，可以作用于这个区域以增强胰岛素基因的转录。在其他组织细胞中没有这种蛋白因子，所以也就没有此作用。

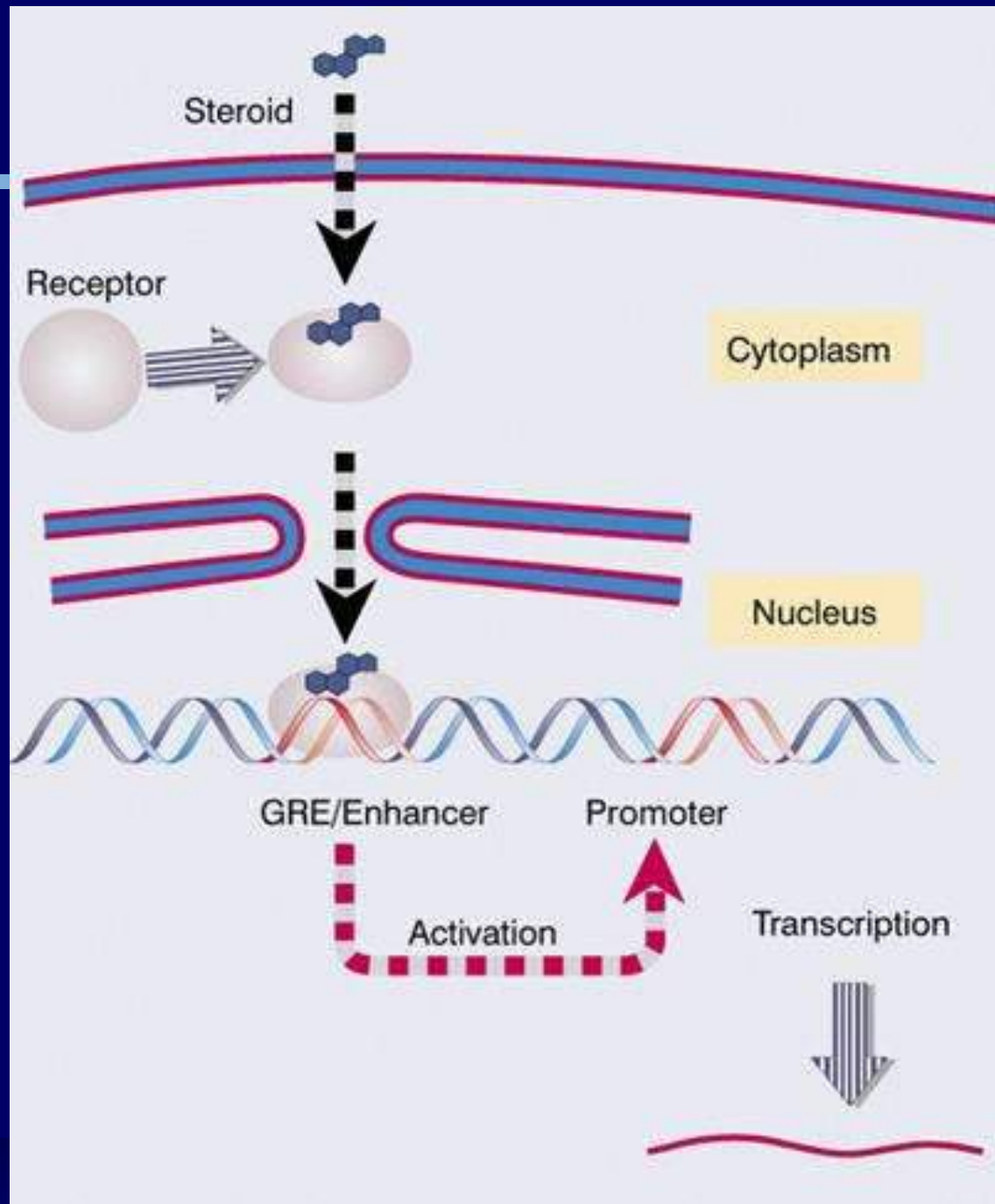
2. 诱导性增强子

在特定刺激因子的诱导下，才能发挥其增强基因转录活性的增强子。

例：激素应答元件（HRE）

在糖皮质激素进入细胞后，首先与胞液中的特异受体作用形成激素-受体复合物，后者进入细胞核，结合于糖皮质激素的增强子上，从而引发特定的一套基因表达，其结果表现出糖皮质激素的生理效应。因此糖皮质激素增强子仅在具有糖皮质激素受体的细胞中才能发挥作用。

类固醇激素受体与
糖皮质激素应答元
件（GRE）结合诱
导基因表达



3. 沉默子阻遏基因转录

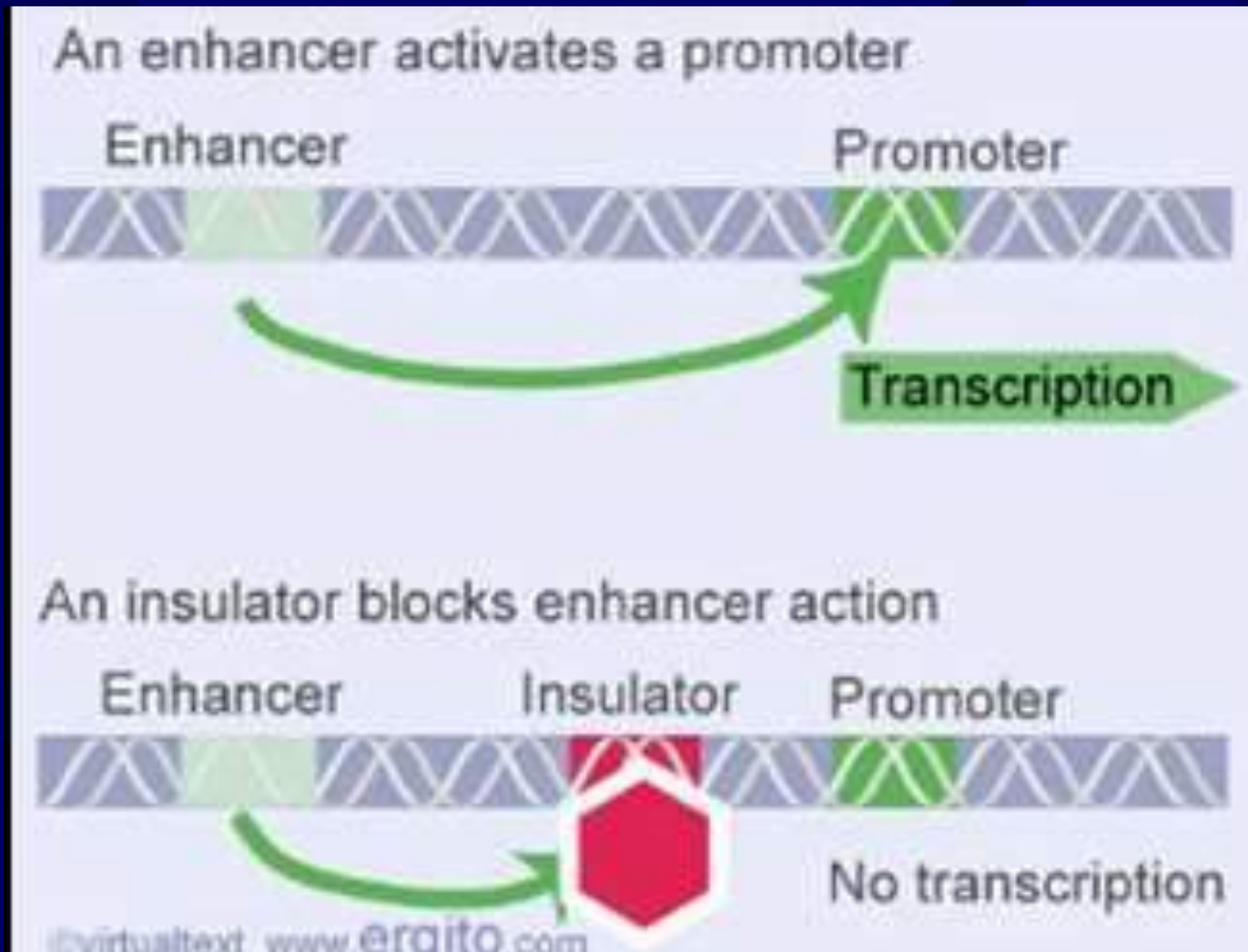
沉默子(silencer)是指某些真核基因转录调控区中抑制或阻遏基因转录的一段（数百bp）DNA序列。沉默序列促进局部DNA的染色质形成致密结构，从而阻止转录激活因子结合DNA，是基因转录的负性调节因素。

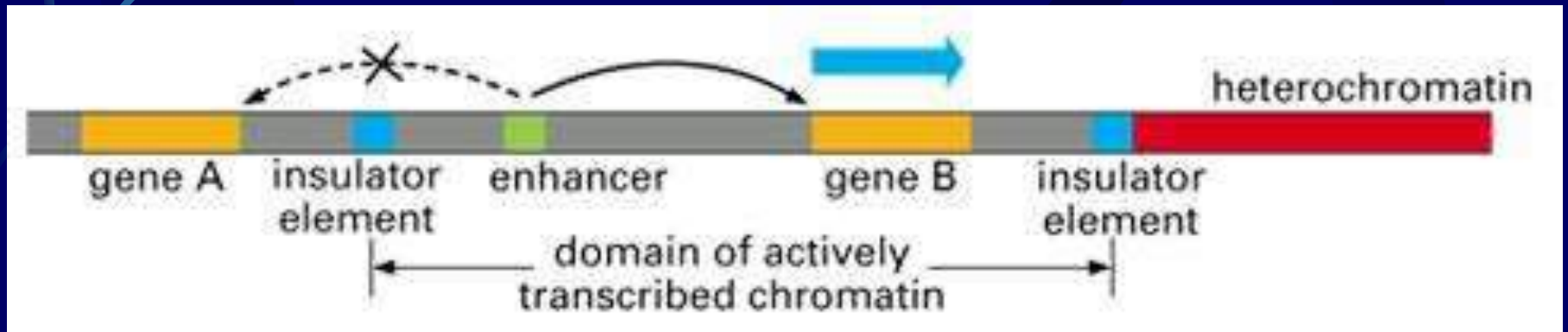
4. 绝缘子

绝缘子 (insulator) 是位于正调控元件(增强子)与启动子之间的小段DNA调控序列, 也可位于负调控元件 (异染色质或沉默子) 与启动子之间。

绝缘子本身对基因的表达既没有正效应, 也没有负效应, 其作用只是不让其他调控元件对基因的活化效应或失活效应发生作用。

绝缘子通过与特定蛋白的作用可阻断增强子对启动子的激活





绝缘子对增强子的抑制作用具有方向性，即只对处于其所在边界另一侧的增强子有抑制作用，而并不能抑制与其处于同一染色质结构域的增强子。

(二) 转录调控蛋白

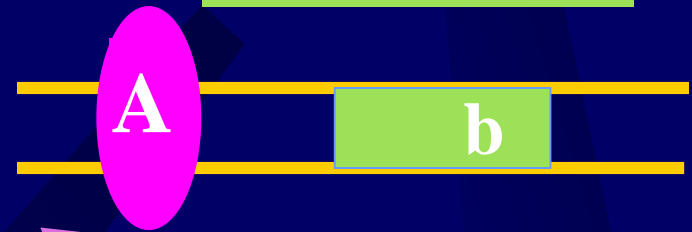
- 在真核细胞核中，有许多蛋白质能够帮助RNA聚合酶转录RNA。这类蛋白质统称转录因子（transcription factors, TF）或转录调节蛋白（transcription regulatory protein）。
- 以反式作用方式调节基因转录的转录因子称为反式作用因子（trans-acting factor），以顺式作用方式调节基因转录的转录因子称为顺式作用蛋白（cis-acting protein）。

DNA

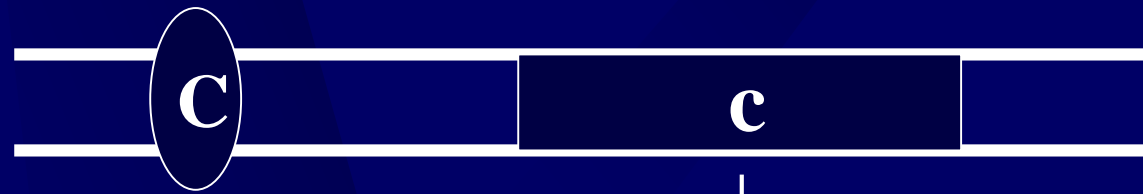
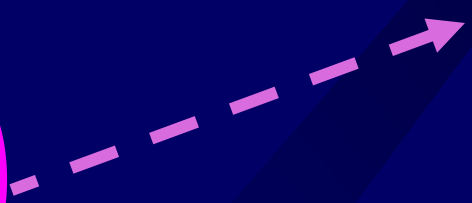


反式调节

mRNA



蛋白质A



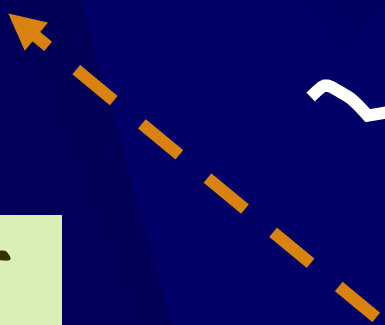
DNA

mRNA



蛋白质C

顺式调节



Pol II 转录需要通用转录因子和特异转录因子

* 基础/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。

* 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

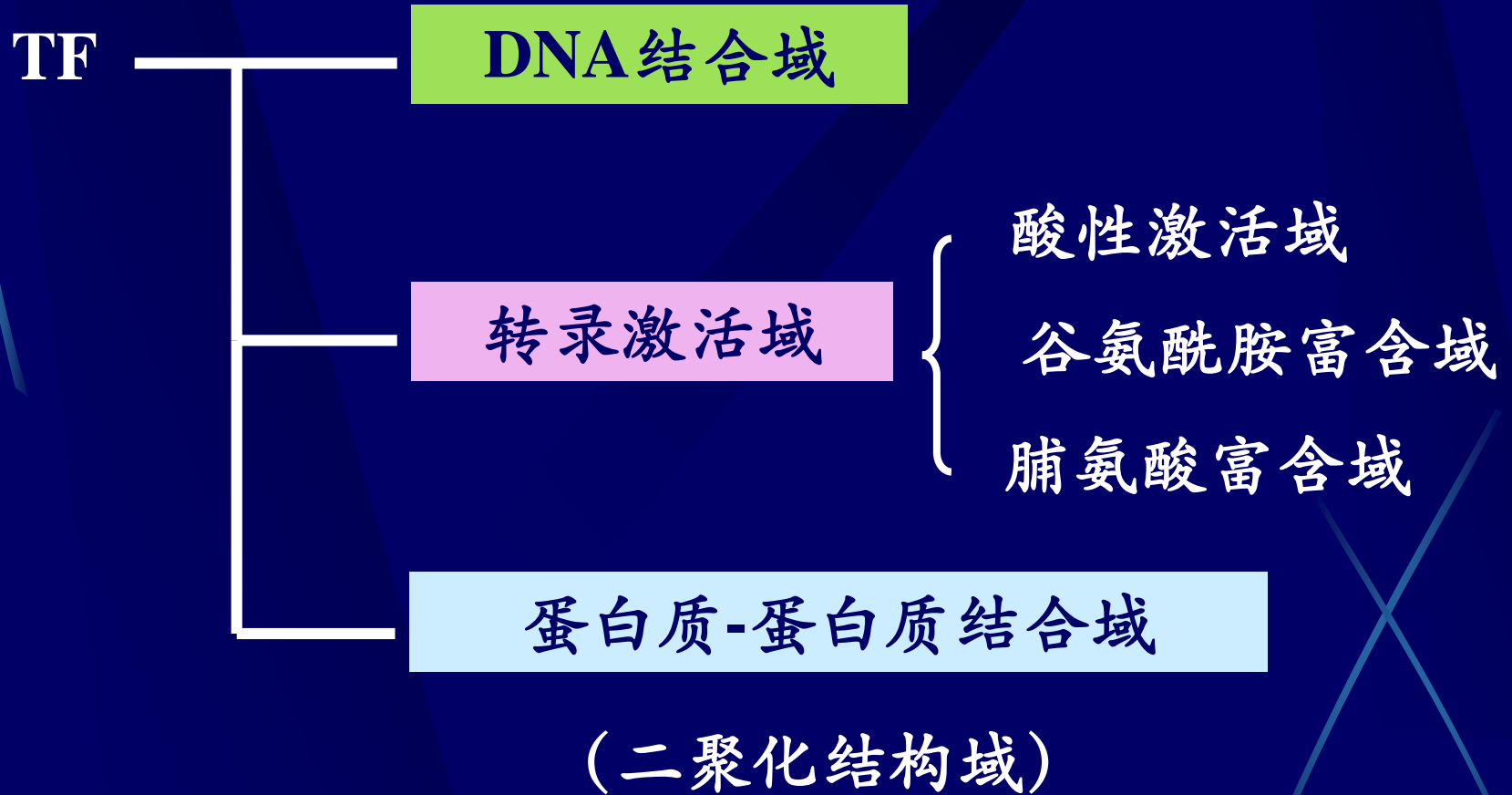
转录抑制因子

转录激活因子通常是增强子结合蛋白。

转录抑制因子有些是沉默子或增强子结合蛋白，有些是通过蛋白质-蛋白质相互作用抑制基因转录。

- * 不同的**DNA**元件组合可产生多种类型的转录调节方式；
- * 多种转录因子又可结合相同或不同的**DNA**元件；
- * 转录因子与**DNA**元件结合后，对转录激活过程所产生的效果各异，有正性调节或负性调节之分。

转录因子的基本结构

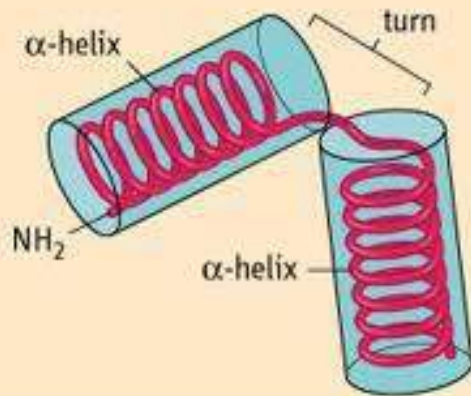


1. 转录因子含有不同的DNA结合域

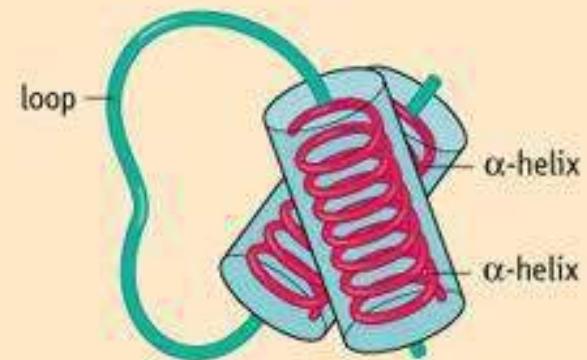
- **螺旋-转角-螺旋**是转录因子中常见的DNA结合域
- **锌指结构**是一类含锌离子转录因子的DNA结合域
- **亮氨酸拉链**既可介导结合DNA又可介导蛋白质二聚体化
- **碱性螺旋-环-螺旋**也可同时介导结合DNA和蛋白质二聚体化

DNA binding motifs commonly found in transcription factors

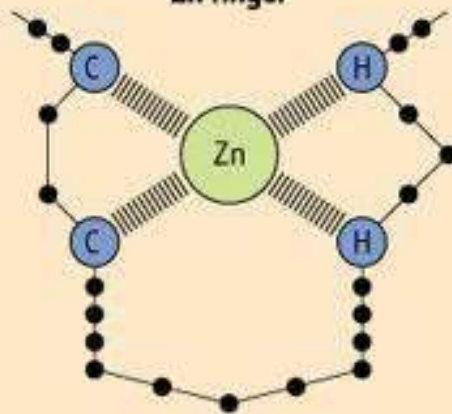
HTH



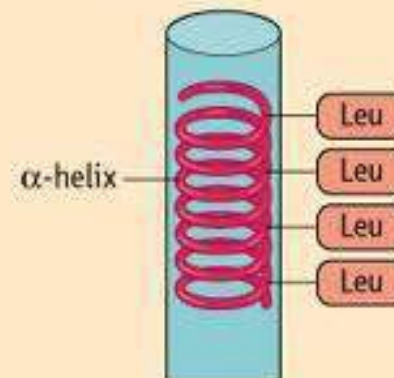
HLH



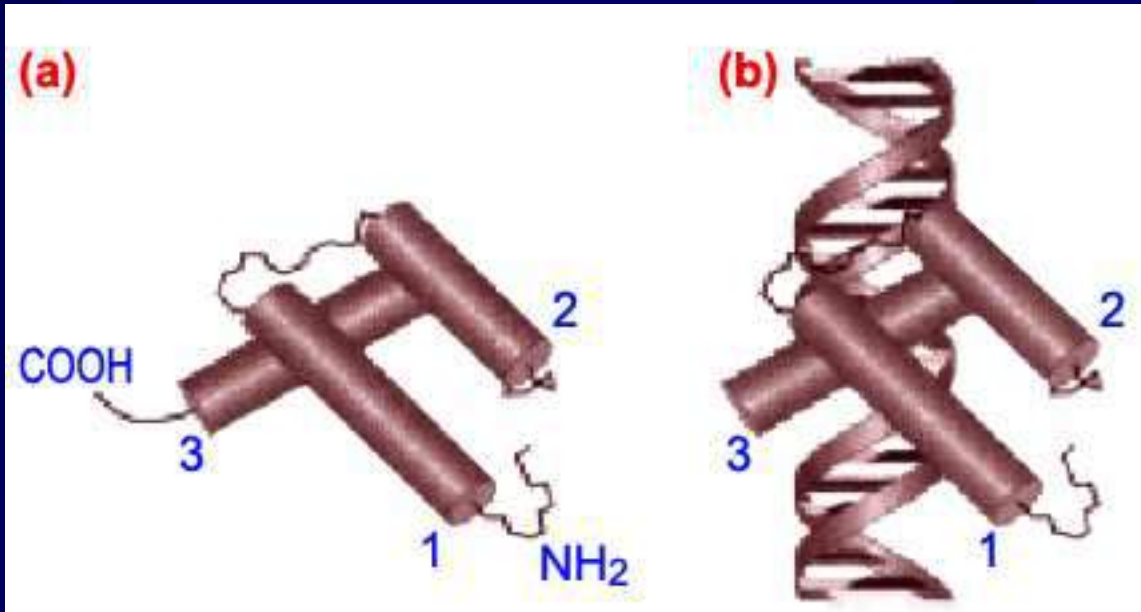
Zn finger



Leucine zipper

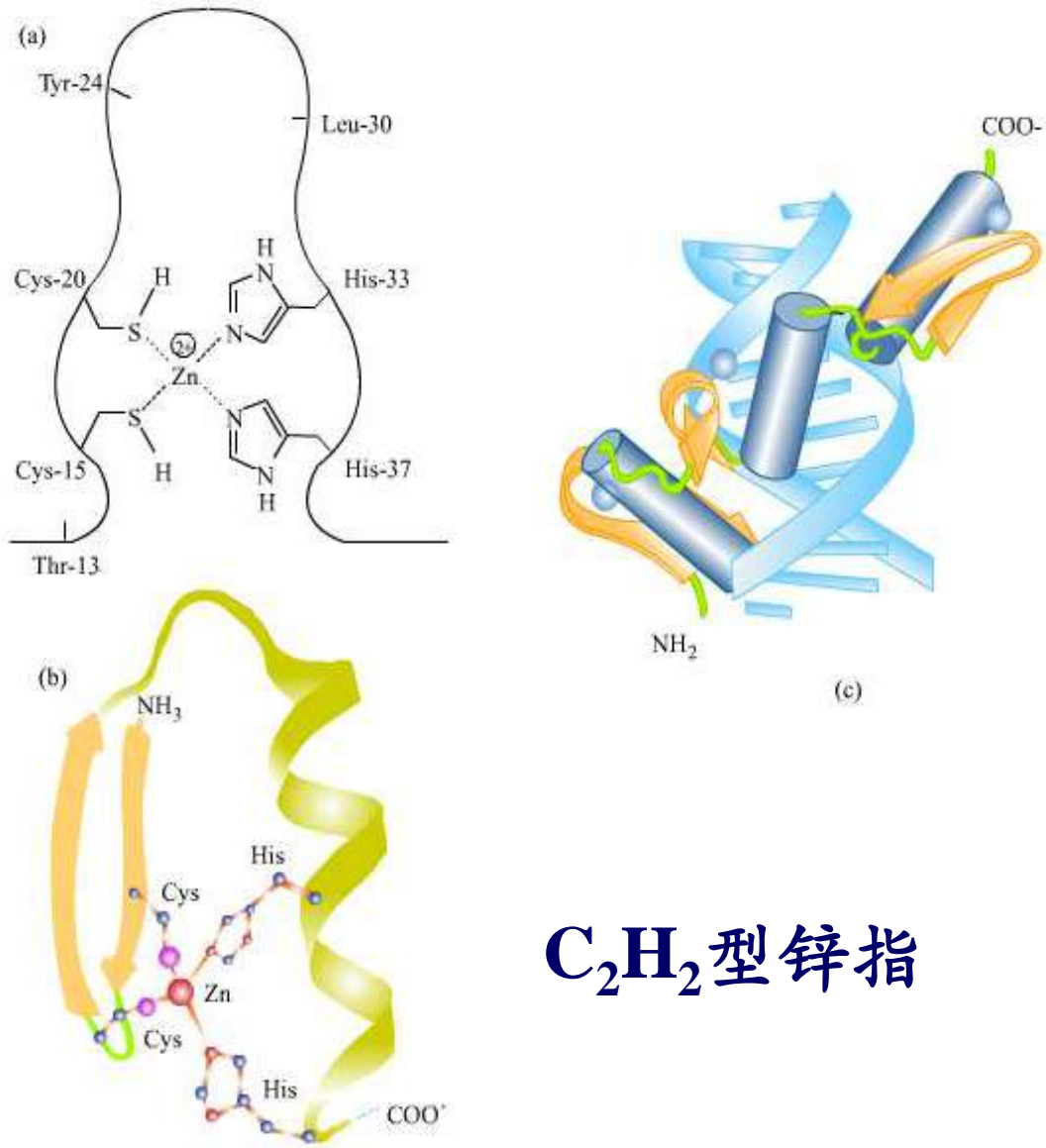


(1) 螺旋-转角-螺旋 (helix-turn-helix, HTH)



- HTH是由3个 α -螺旋（7~9氨基酸残基）经 β -转角连接而成。3个 α -螺旋通过侧链间的相互作用维持固定的角度。
- C末端的 α -螺旋（螺旋3）是识别螺旋，嵌入DNA大沟。

(2) 锌指结构 (Zinc finger)

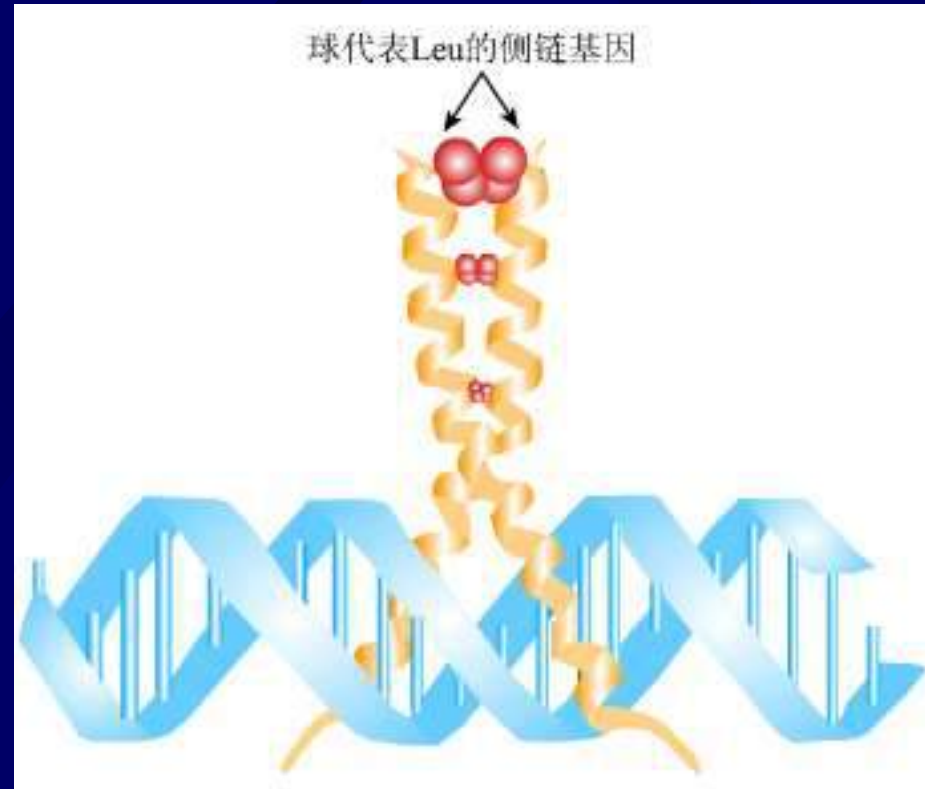


C₂H₂型锌指

- 锌指结构是一类含锌的DNA结合蛋白质模体，由一个含有大约30个氨基酸的环和一个与环上的4个Cys或2个Cys和2个His配位的Zn²⁺构成 (a)
- 单个锌指的三维结构是由一个α螺旋和两个反向平行的β折叠组成 (b)
- 含锌指结构的转录因子通过α螺旋与DNA大沟接触来影响转录 (c)

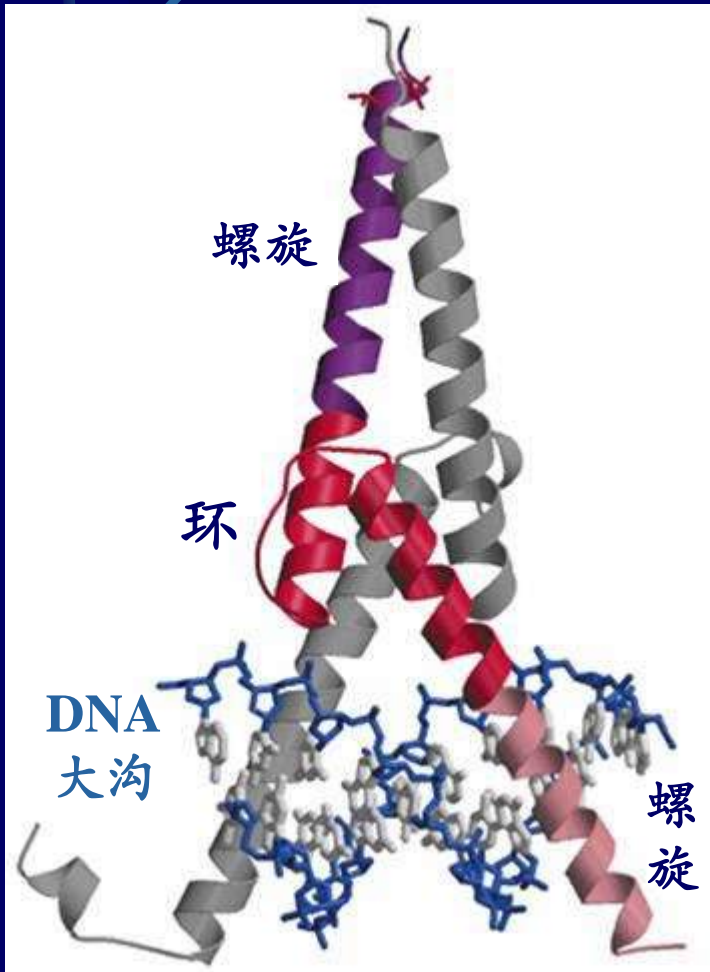
(3) 亮氨酸拉链 (leucine zipper)

- 含亮氨酸拉链的转录因子以二聚体的形式识别和结合DNA，象一个倒置的Y，包括疏水区和亲水区。
- 疏水区是两个相同 α -螺旋靠疏水相互作用彼此缠绕形成的二聚体区。螺旋中每隔6个氨基酸残基就出现一个亮氨酸残基，相邻 α -螺旋中的亮氨酸残基侧链都伸向对方，亮氨酸侧链象两手手指那样交叉锁住，提供了可将两个 α -螺旋结合在一起的疏水堆积相互作用。
- 亲水区为DNA结合域，是分开的两个 α -螺旋，螺旋中含有很多碱性氨基酸残基。



亮氨酸拉链二聚体
结合DNA大沟

(4) 碱性螺旋-环-螺旋 (basic helix-loop-helix, bHLH)



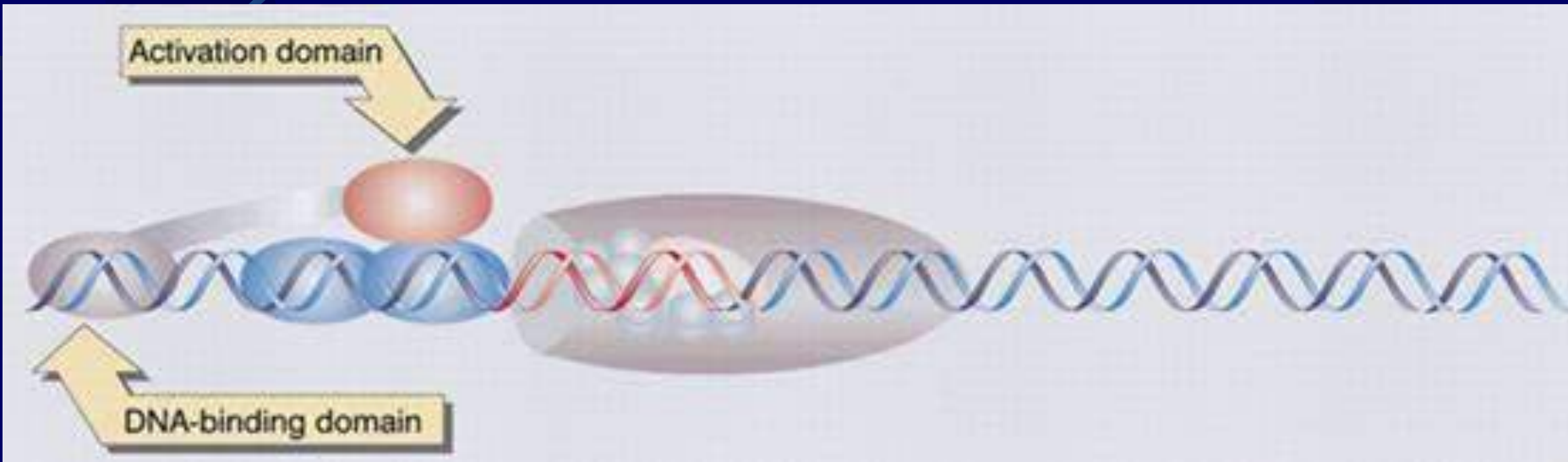
- bHLH结构也可调节DNA结合及蛋白质二聚体化。
- 一个bHLH结构是由一个短的 α -螺旋通过一个环与另一个长的 α -螺旋组成。

bHLH双体

2. 转录因子含有不同的转录激活域

- 富含谷氨酰胺的激活结构域
- 富含脯氨酸的激活结构域
- 富含酸性氨基酸的激活结构域

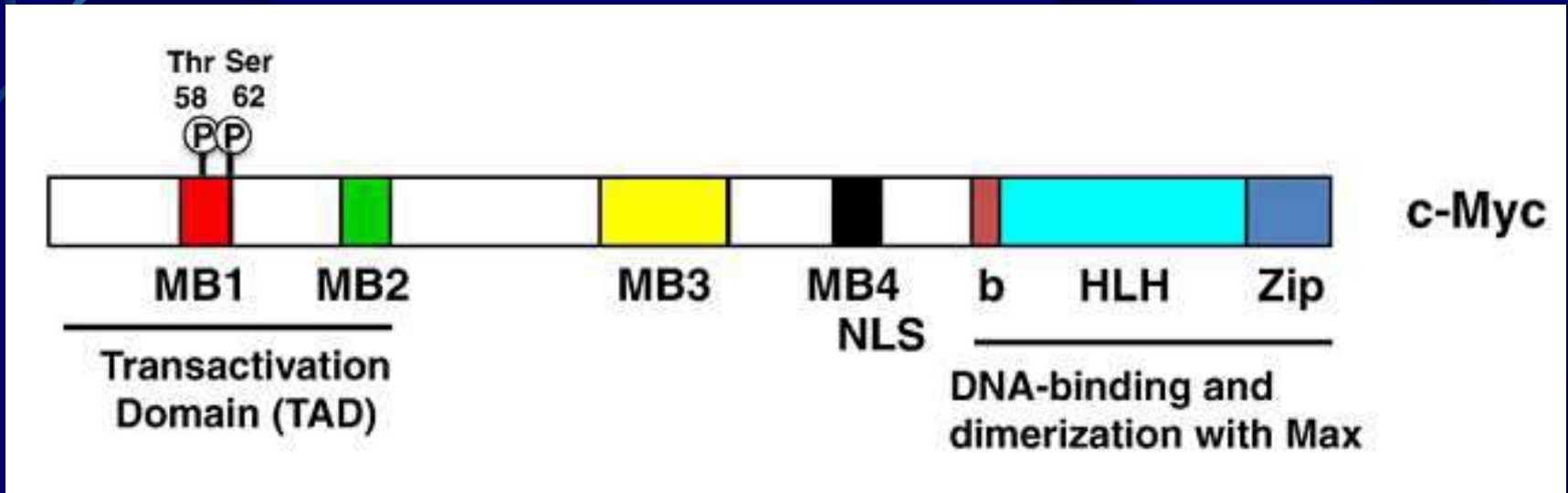




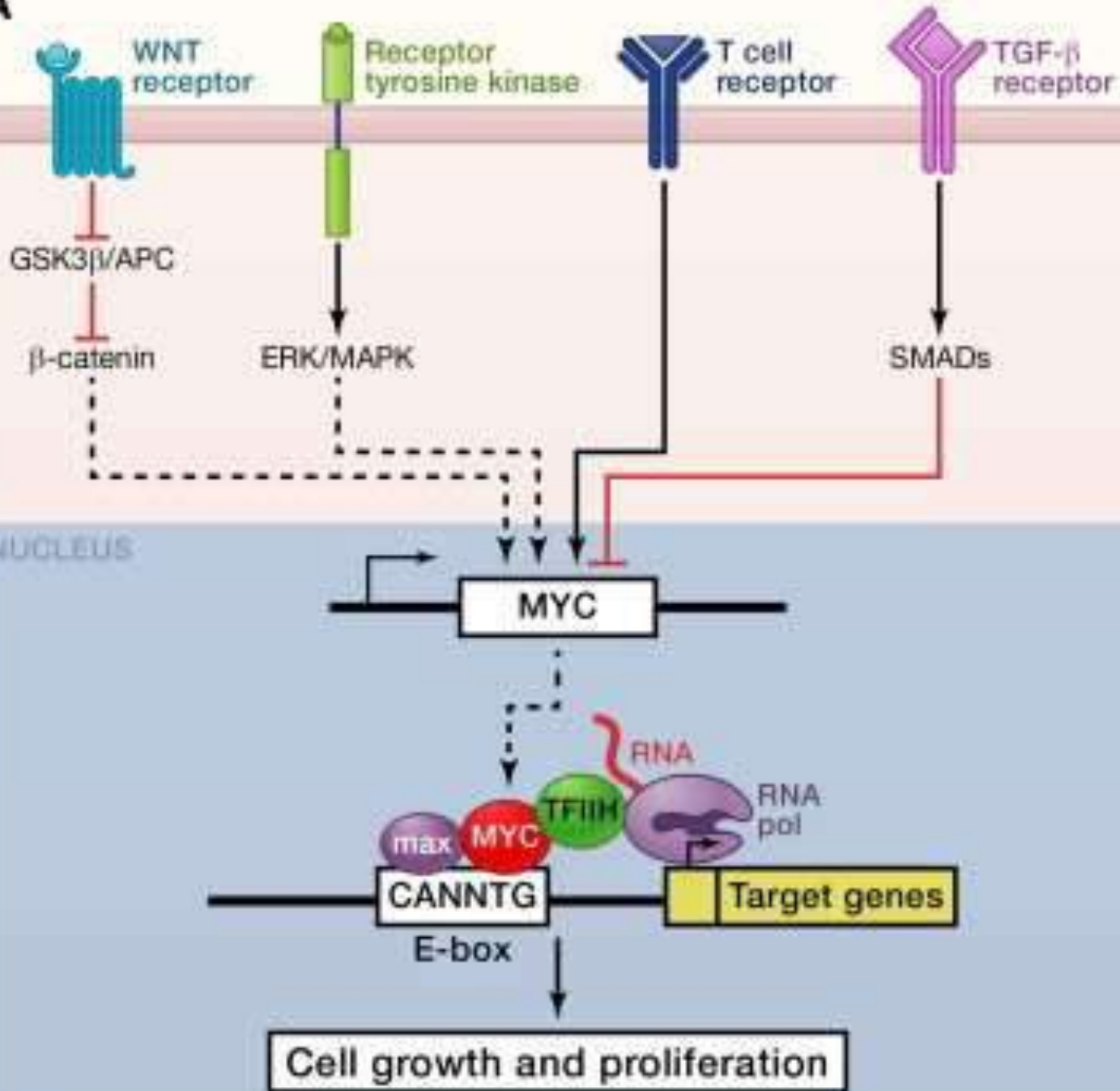
作用

- 1) 改变启动子周围的染色质结构；
- 2) 招募RNA聚合酶，使其快速、准确定位于启动子TATA盒；
- 3) 介导转录相关蛋白质的化学修饰与构象改变，促进转录起始。

转录因子的作用模式



- **C-myc**主要通过C端的碱性区域/螺旋-环-螺旋/亮氨酸拉链区(**b-HLH-Zip**)与同样含有b-HLH-Zip结构的max蛋白形成异二聚体，特异地识别其靶基因DNA序列中的**CACGTG**核心序列(E盒)并与之结合，使被调节的基因激活或转录增强。

A

基因转录激活调节的基本要素：

- ❖ 基因的结构、性质
- ❖ 生物个体或细胞所处的内、外环境
- ❖ 细胞内所存在的转录调节蛋白

真核细胞基因开关的分子机制比原核复杂

- **基本共同点：** 转录起始的调节是关键点
- **根本不同点：**

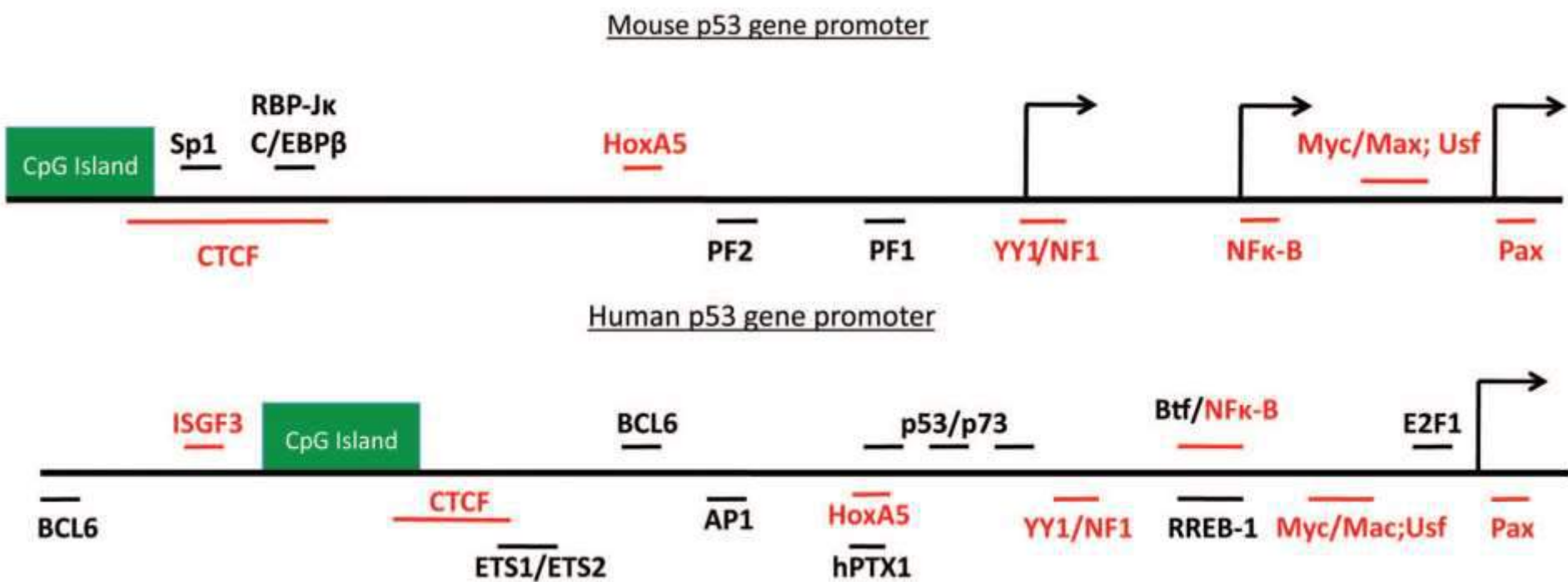
原核生物： 启动子在缺少转录因子情况下
就具有天然活性

真核生物： 强力启动子在缺少调节蛋白的
情况下往往没有活性

真核细胞对基因表达起始的调控至少在下面几个方面与原核细胞存在不同：

- * 真核细胞基因及其调控区域受到**染色质结构的限制**，转录的活化与转录调控区域、转录区域内染色质结构的诸多变化相关；
- * **正性调节**是主要形式。因此，在转录的基本状态受到制约的情况下，使得每个真核细胞基因需要活化才能被转录；
- * 真核细胞有更大、更为复杂、种类更多的**调节蛋白**。单一启动子可以被分散在DNA分子上、数量近乎无限的调节序列所控制。

● p53基因启动子区各种转录因子的识别位点



3. 转录起始的调控需要转录调控蛋白与 转录调控区DNA的相互作用

RNA聚合酶II与启动子的结合、启动转录需要诸多蛋白质因子的协同作用。这通常包括：

- ①基因活化蛋白质与增强子或UASs的结合；
- ②通用转录因子在启动子处的组装；
- ③共激活蛋白(coactivator)和/或中介子(mediator)在通用转录因子/RNA聚合酶II复合物与基因活化蛋白质之间的辅助和中介作用。

基因活化蛋白质与增强子结合后，通过与全酶复合体中的中介子相互反应，使全酶复合体在空间上更接近启动子并有效组装。

此外，多数全酶复合体中缺少一些通用转录因子，如TFIID与TFIIA，它们需要在启动子处分别组装，最后形成稳定的**转录起始复合物**（**transcription initiation complex**），启动转录。

基因活化蛋白与增强子结合后如何影响到远距离的RNA聚合酶结合位点，有以下几种模式：

① 扭曲 (twisting)

通过扭曲作用使DNA链发生构型变化，更适合于通用转录因子与RNA聚合酶结合，并通过直接接触或通过辅活化子/中介子而影响通用转录因子和RNA聚合酶的组装。

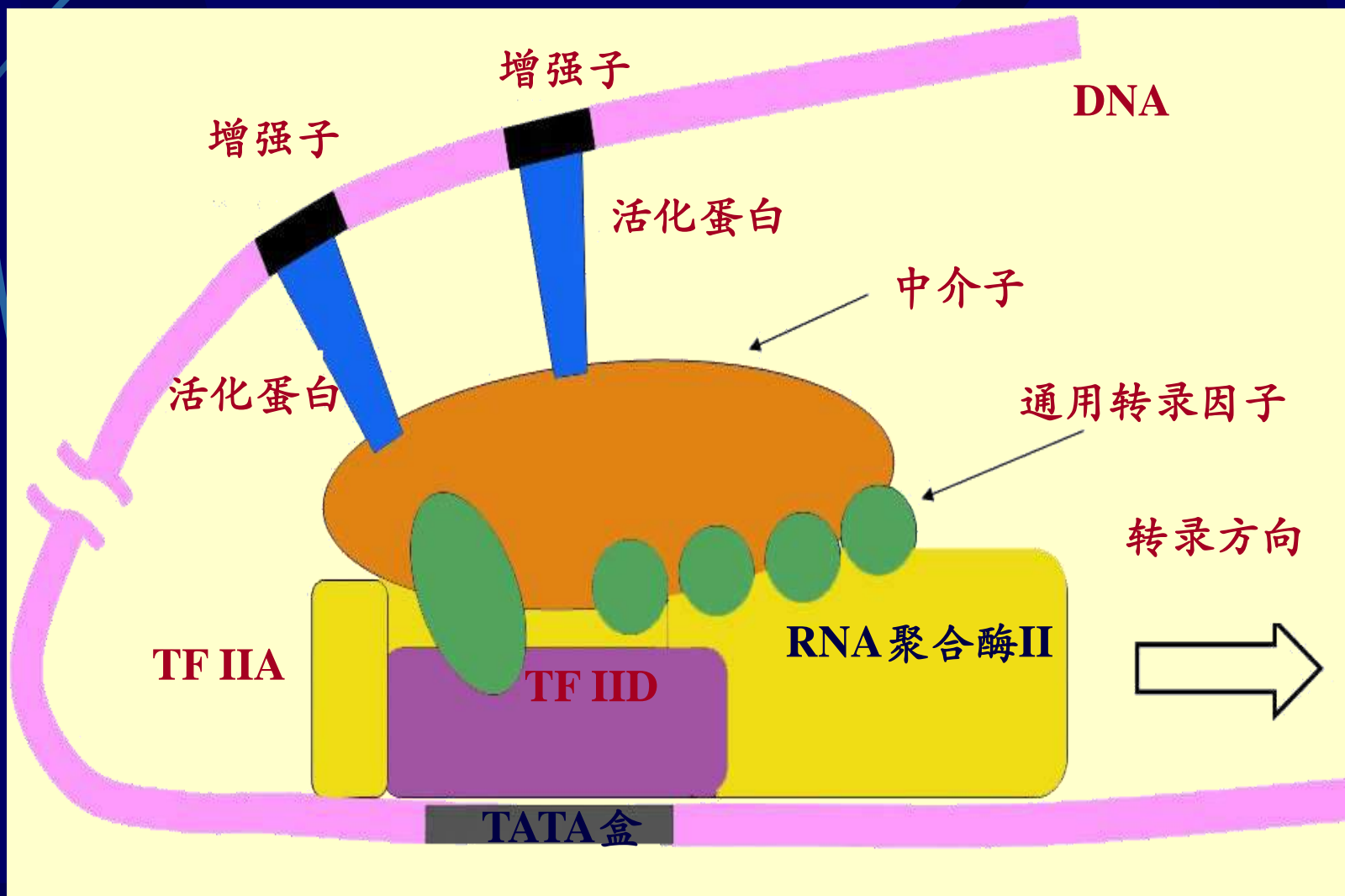
②滑动 (sliding)

沿DNA滑动，直到接触另一个特异DNA序列结合的转录因子，发挥作用。

③成环 (looping)

利用DNA分子的柔曲性弯曲成环，使增强子区域与RNA聚合酶结合位点靠近，直接接触或通过辅活化子/中介子而发挥作用。

真核基因转录起始复合物



二、转录后加工运输的调控

(一) mRNA的剪接加工调控

- 顺式剪接 (**cis-splicing**)

切除RNA初产物的内含子，连接相邻的外显子。

- 反式剪接 (**trans-splicing**)

切除RNA初产物的内含子后，不同基因的外显子相互连接。

- 可变/选择性剪接 (**alternative splicing**)

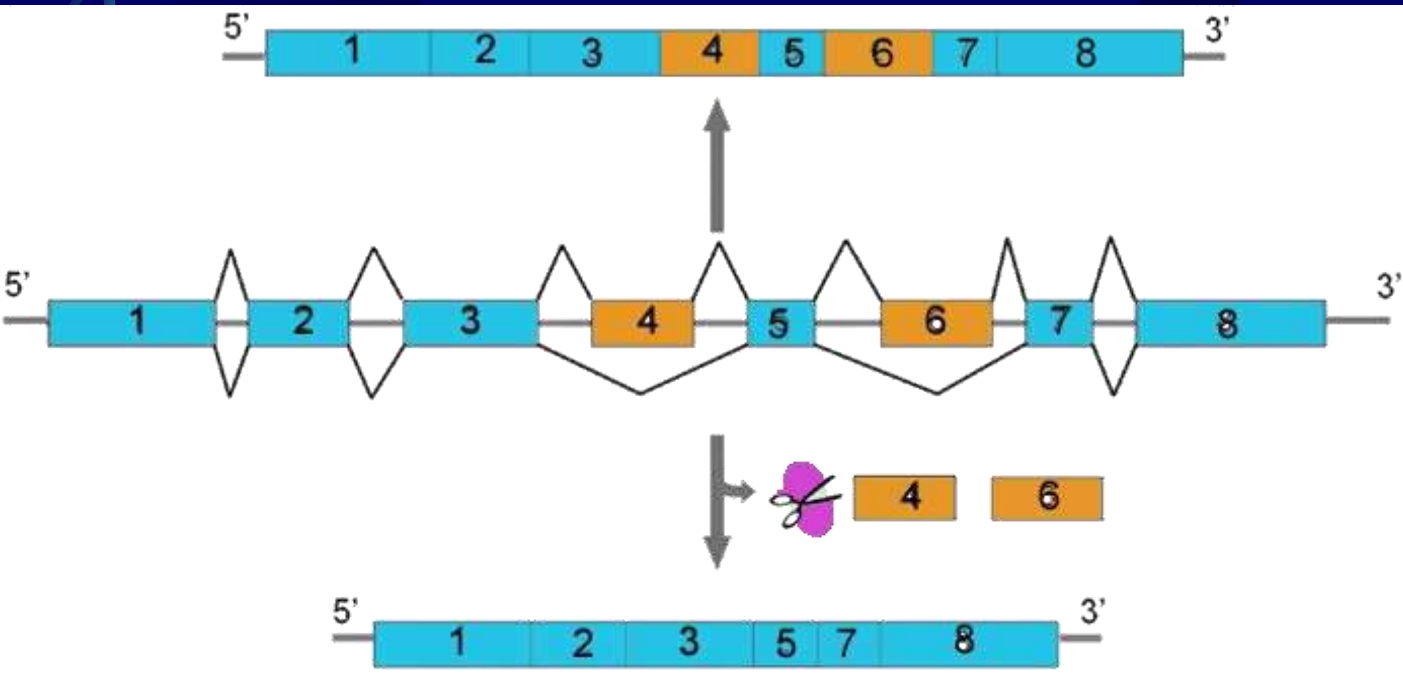
mRNA初产物可以通过一种以上的选择性剪接方式，去除不同的内含子而被加工形成不同的mRNAs，因而形成不同的多肽。

选择性剪接的不同方式

- 外显子缺失
- 内含子保留
- 多位点剪接

1/3以上的基因存在选择性剪接，有限的基因可产生数量繁多的蛋白质。

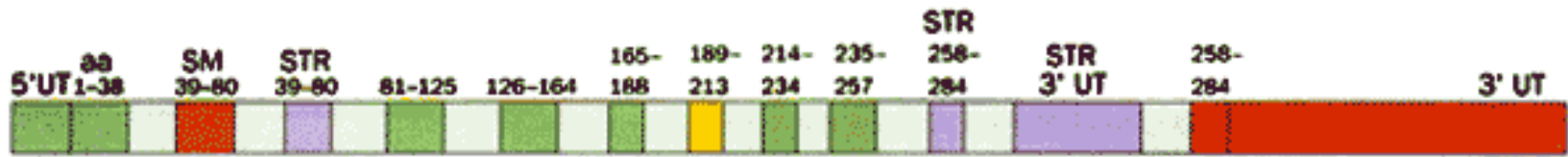
人视黄醛还原酶mRNA的可变剪接



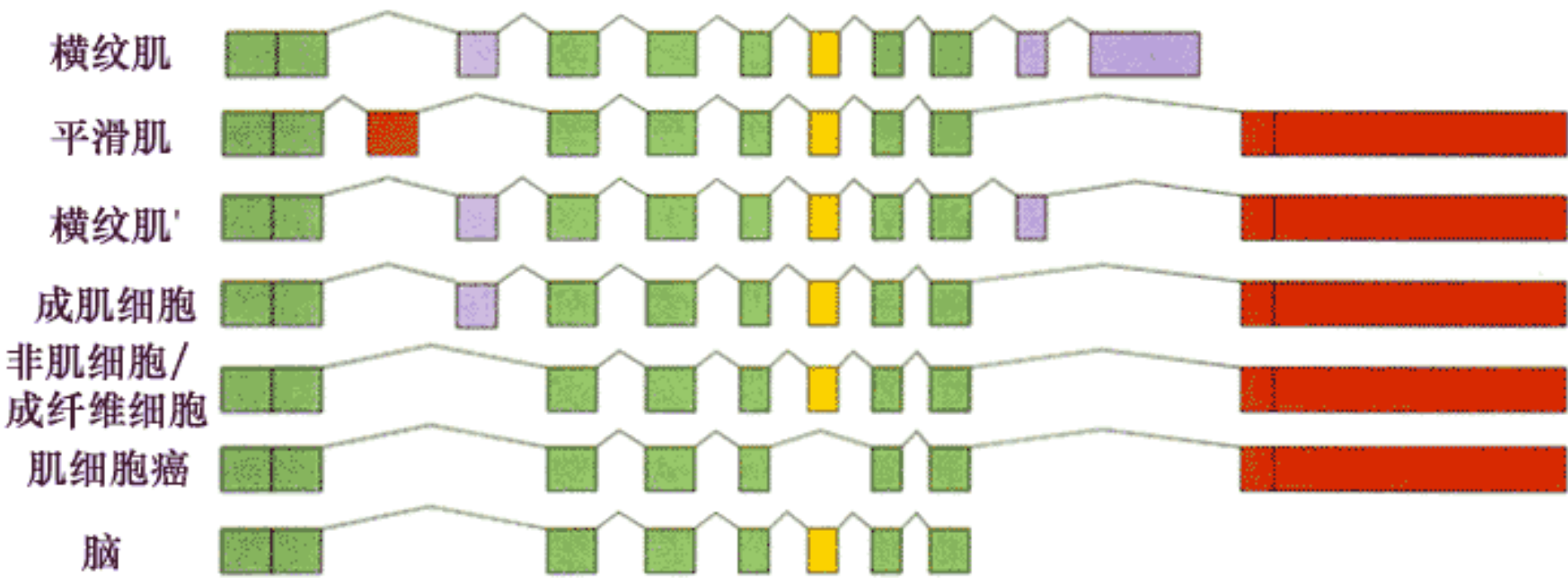
mRNA

Pre-mRNA

同种异型mRNA



mRNA转录物



大鼠的 α -原肌球蛋白 (α -tropomyosin) 基因可通过改变剪接位点, 以不同的剪接方式进行加工, 编码7种组织特异性的原肌球蛋白。

选择性剪接的调控

初始转录本含有所有选择性加工途径所需要的分子信号。

一种细胞偏好何种选择性加工途径取决于加工因子——**RNA结合蛋白**的特异性。

选择性剪接可以被正负调节分子调节：

负调节：抑制蛋白质可以通过与原始转录本的结合来防止剪接复合体切除内含子序列。

正调节：不能正常剪接的剪接复合体可以在活化蛋白的帮助下发挥剪接功能。

(二) mRNA的编辑加工调控

某些mRNAs在翻译前被编辑(editing)。

- mRNA的编辑和可变剪接一样，能使同一基因产生几种不同的蛋白质。
- 主要通过插入或缺失核苷酸，或使编码区内的碱基发生转换或颠换，从而改变遗传信息。

mRNA的核苷酸插入

结果：转录后编辑过程插入了4个U残基，从而改变了转录本的翻译读码框。

机制：尚不清楚。研究人员已经发现线粒体转录一类特殊的RNA分子，其3'端有一段 poly(U)，其5'端序列与mRNAs被编辑的区域互补，被称为**引导RNA (guide RNA)**。引导RNA可能作为编辑过程的模板，并将其3'端的U转移给被编辑的mRNAs。

Original RNA transcript



Guide RNA used as template for addition of U's



Functional mRNA



Inserted uridines

Guide RNA

(三) mRNA 的稳定性调控

降解途径保证mRNA不在细胞中累积并避免合成过多的蛋白质。

不同真核基因的mRNA的降解速率大不相同。

暂时需要的基因产物：半衰期可能仅为几分钟、甚至几秒钟。

经常需要的基因产物：其mRNA可在多代细胞中稳定存在。

除组蛋白外，所有真核细胞mRNA都有5'端的“帽”和3'端的Poly(A)尾结构。

5'端的“帽”和3'端的Poly(A)尾均有其特有的作用。

5'加帽的作用在于:

- ①有助于保护mRNA免于被核糖核酸酶降解;
- ②协助mRNA的剪接。在剪接第一个外显子时,剪接体的形成需要帽结合蛋白的参与;
- ③促进mRNA从细胞核运输到细胞浆;
- ④5'帽结合蛋白复合体参与mRNA和核糖体的结合来起始翻译。

poly(A)尾的作用：

poly(A)尾可结合一种或多种特殊蛋白，避免mRNA被酶降解，并在翻译过程中具有重要作用。许多原核mRNA也含有Poly(A)尾，但是此尾的功能是促进mRNA降解，而不是保护mRNA免于被降解。

真核细胞mRNA降解有两种途径，由每种mRNA分子的序列所决定

• Poly(A)的逐渐短缩：最常见的途径

- ❖ mRNA分子一旦进入细胞浆中，核酸外切酶会使Poly(A)末端逐步短缩，当剩下约30个A时，5'端发生脱帽，mRNA分子被迅速降解。
- ❖ 一些mRNA分子的3'UTR的特殊序列有助于特殊蛋白质的结合，增加或降低Poly(A)短缩的速度。

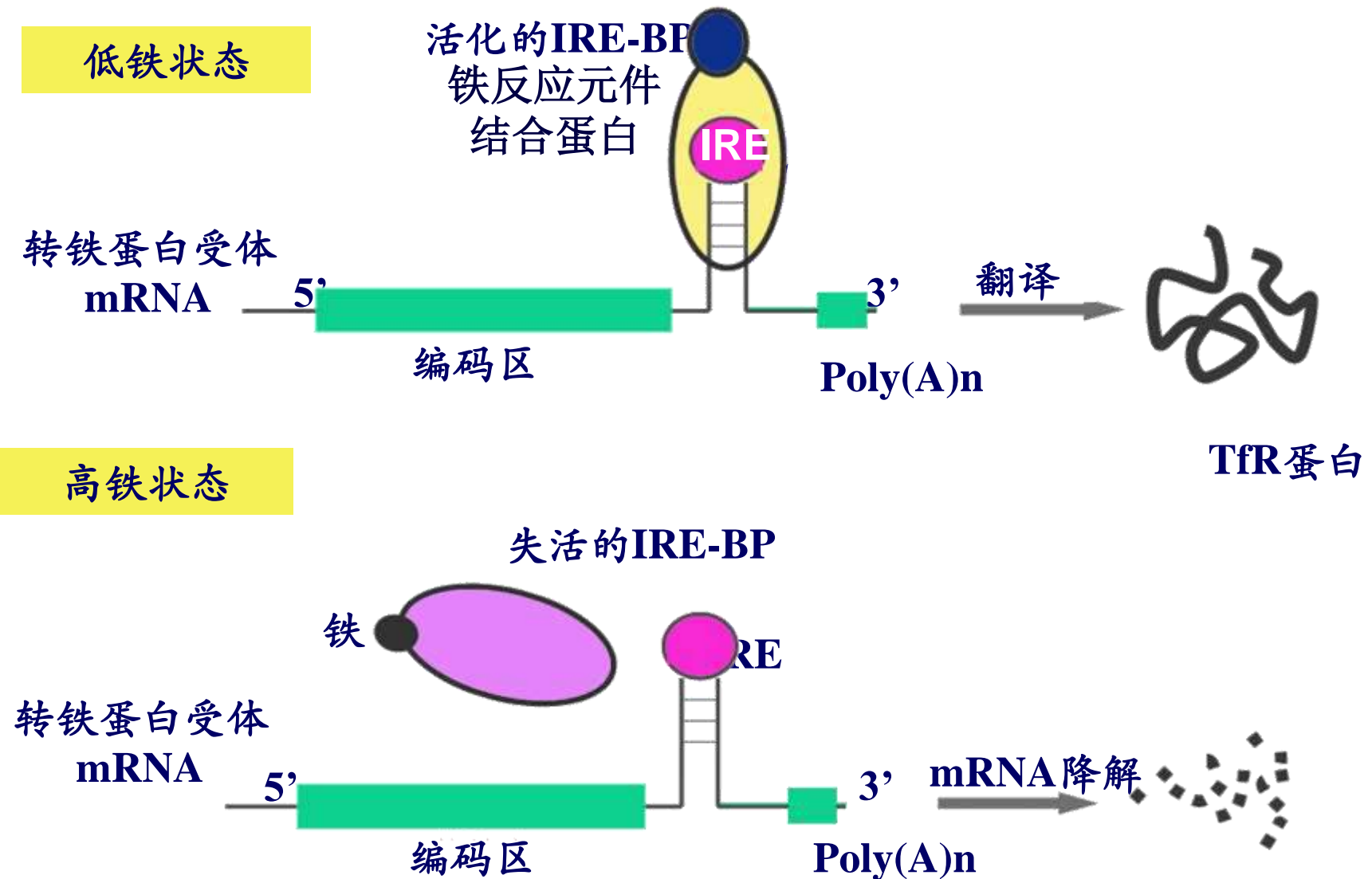
• 特异核酸内切酶切割途径

- ❖ 可将Poly(A)直接从mRNA分子上切除。这种切除也有赖于mRNA分子的3'UTR特殊序列可以被内切酶识别。

例如：

转铁蛋白受体mRNA分子3'UTR存在茎环（stem-loop）结构——铁反应元件（iron-response element, **IRE**）。IRE及其附近有内切酶识别位点。

IRE-BP对转铁蛋白受体mRNA稳定性的调节



(四) 成熟mRNA的核外转运和定位

• 核外转运

现象： 估计有1/5的核内成熟mRNAs能进入细胞浆。留在核内的mRNAs约在1小时内降解。mRNA通过核膜的过程是一个主动运输过程，常常需要借助于核输出受体(nuclear export receptors)才可穿过9nm的核孔通道。

机制： 调控RNA从核运输至细胞浆的机制还不很清楚。

• mRNA定位

现象：一些mRNA分子携带有信息，可以在翻译开始前自我导向定位于细胞内的特定位置。

作用：在细胞的特定部位集中产生所需要的大量蛋白质，避免产物的远距离运输。

机制： 导向信号存在于mRNA的3'端非翻译区（3'untranslated region, 3'UTR）。

• **第一种情况**

mRNA 被连接在附着于细胞骨架的动力蛋白（motor proteins）上，利用其水解ATP所提供的能量沿着骨架成分朝目的方向移动，最终在目的地被**锚蛋白**（anchor proteins）固定。

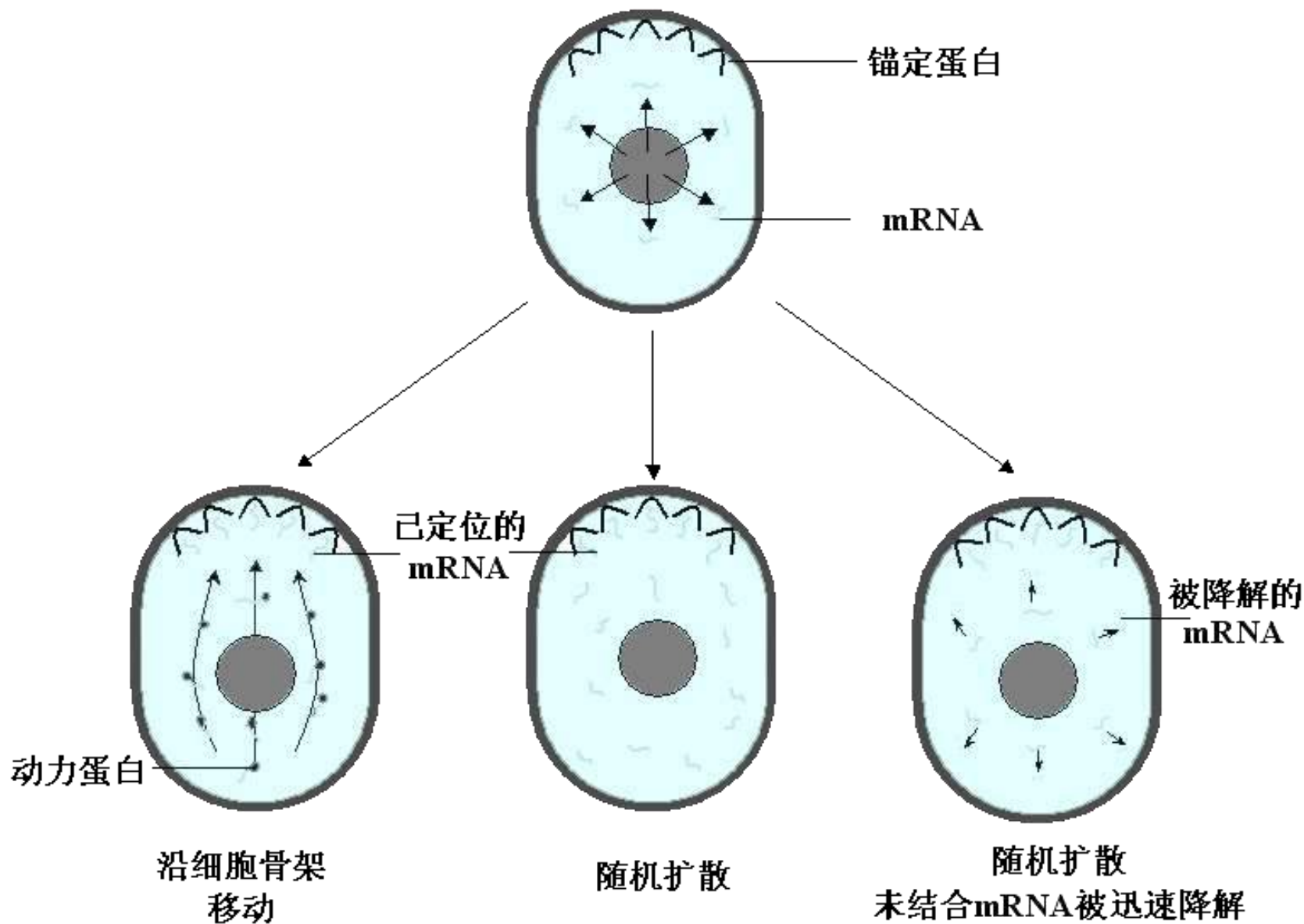
• **第二种情况**

mRNA通过在细胞浆中的随机扩散，在其定位处被锚蛋白捕捉、固定。

• **第三种情况**

mRNA在细胞浆中随机扩散并被不断地降解，只有碰上锚蛋白才能得到保护。

mRNA分子的3种不同定位过程



(五) 小分子RNA介导的转录后基因沉默

一些小的双链RNA可以高效、特异地阻断体内特定基因表达。这是由于此类小RNAs与mRNAs（经常是3'UTR）相互作用，导致mRNA降解或者翻译抑制，使mRNA及其相应基因无法表达而沉默（silence）。

意义：

控制某些生物体的适时发育。它也是一种保护生物体免受RNA病毒侵袭和控制转座子活性的机制。

参与转录后基因沉默的一些活性分子

- ✓ **dsRNA**: 双链RNA, 包含正义链和反义链
- ✓ **siRNA**: small interfering RNA, RNAi的关键效应分子, 21-23个nt大小的双链RNA, 多数是外源性的
- ✓ **miRNA**: microRNA, 21-23个nt大小的单链RNA, 由内源性基因编码
- ✓ **Dicer**: 属于RNaseIII 家族, 是dsRNA的特异性核酸内切酶
- ✓ **RISC**: RNA-inducing silencing complex, 具有核酸内切、外切以及解旋酶活性

RNA诱导的沉默复合物 (RNA-inducing silencing complex, RISC)

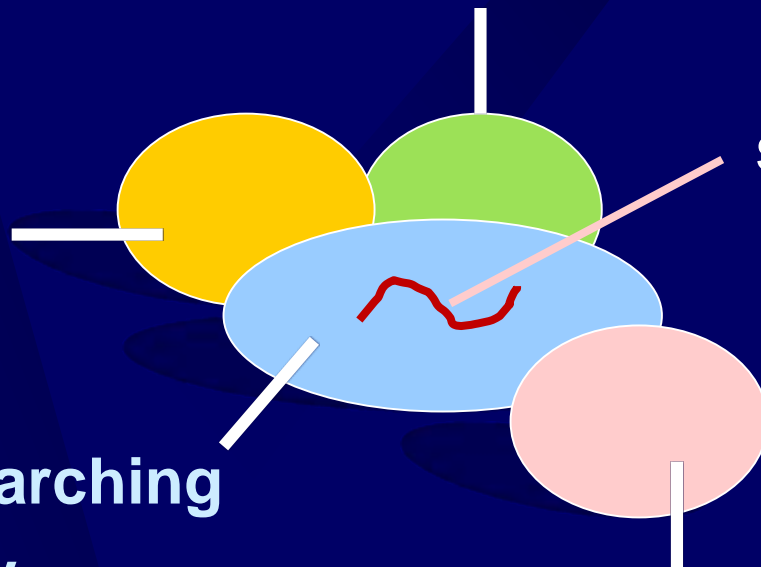
Exonucleolytic nuclease
核酸外切酶

siRNA

Helicase
解旋酶

Hemology-searching
activity

Endonucleolytic nuclease
核酸内切酶



● siRNA介导的转录后基因沉默

siRNA与RNA诱导的沉默复合物(RISC)结合，将RISC激活进而使双链siRNA解链，其中的一条链将去识别与其序列互补的特异性mRNA，然后RISC对特异性mRNA进行酶切，阻断相应蛋白的翻译，抑制相关基因的表达。

因为RNA诱导的RISC可以循环使用，能够继续酶切其他的靶mRNA，所以siRNA对靶mRNA的抑制作用可以被放大，特异性、高效性抑制基因表达。

● miRNA介导的转录后基因沉默

miRNA的特点

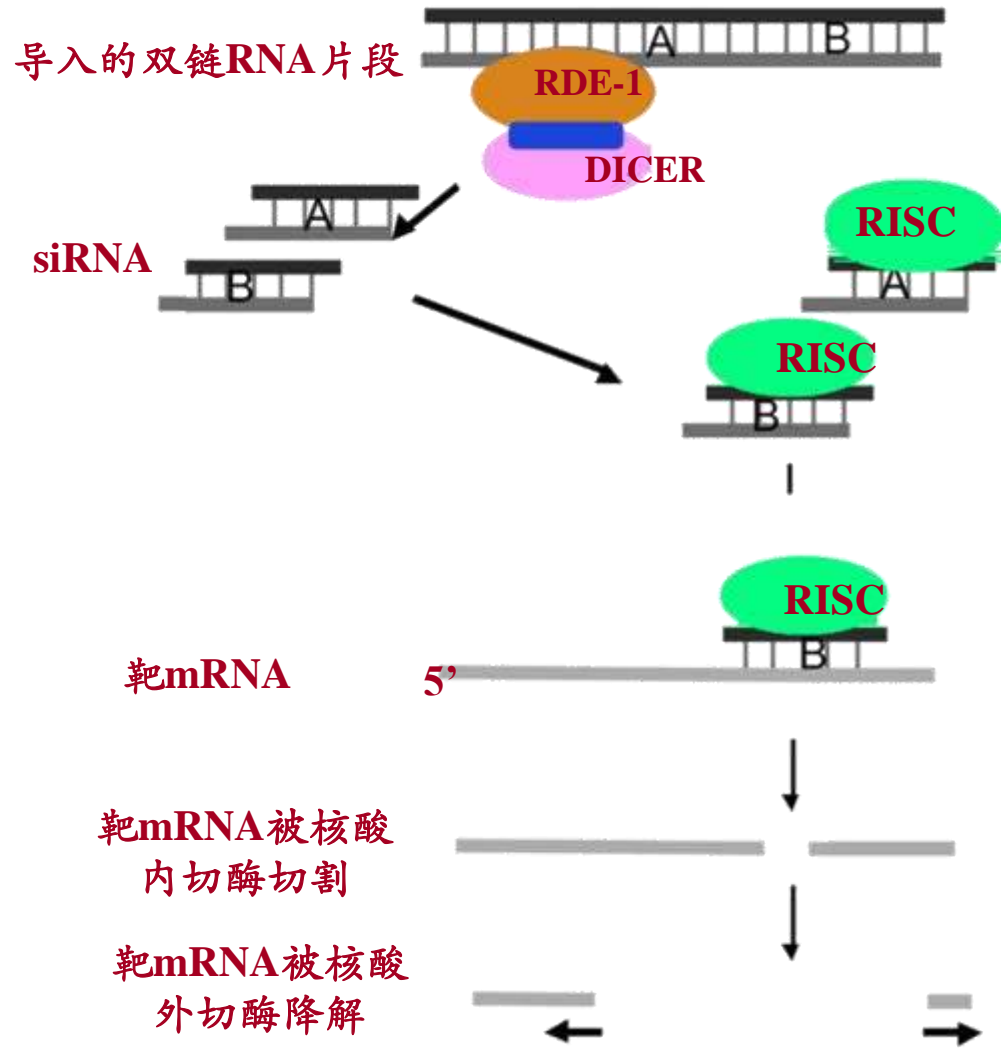
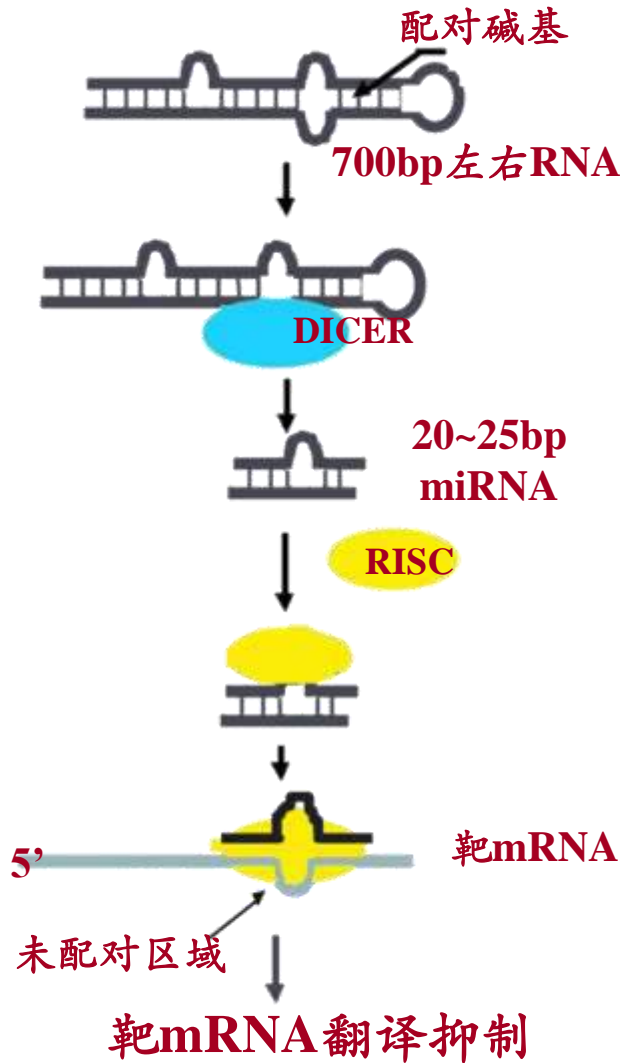
- 在真核细胞中广泛存在，可以稳定存在于体液中，包括唾液、尿液、母乳和血液
- 没有开放阅读框，是一类非编码RNA
- 普遍长度约为20~23 nt，在3'端有1~2 nt的变化，成熟的miRNA 5'端有磷酸基团，3'端有羟基
- 序列有高度保守性，表达水平具有时空特异性

- 生物信息学预测发现，每个miRNA都有众多的靶基因，而每个基因的mRNA又有可能受到多个miRNA的调控，由此构成了复杂的调控网络。
- 在哺乳动物基因组中，30%以上基因的mRNA都受到miRNA的调节。
- 此外，一些miRNA隶属于同一个家族，它们具有相同的种子区域——与目标mRNA结合的关键部位，即从5'端起2~7位具有相同的序列。

miRNA与siRNA的区别

- **miRNA**是内源性的；**siRNA**是外源性或内源性的
- **miRNA**合成起始于胞核，再被转运至胞质；**siRNA**一般不在胞核合成
- **miRNA**前体是不完整的发夹结构RNA；**siRNA**前体是完全互补的长双链RNA
- 成熟的**miRNA**是单链；成熟的**siRNA**是双链结构
- **miRNA**与靶mRNA并非严格互补；**siRNA**通常与靶mRNA完全互补
- **miRNA**参与正常机体生长发育调节；而**siRNA**通常不参与

miRNA 与 siRNA 使靶 mRNA 沉默机制



表观遗传学（epigenetics）

研究基因的核苷酸序列不发生改变的情况下，基因功能的可逆的、可遗传的改变。表观遗传的现象很多，已知的有DNA甲基化（DNA methylation），基因组印迹（genomic imprinting），母体效应（maternal effects），基因沉默（gene silencing），核仁显性，休眠转座子激活和RNA编辑（RNA editing）等。

- 表观遗传密码构成了基因(DNA序列) 和表型(由基因表达和环境因素所决定)间的关键信息界面，这使经典的遗传密码中所隐藏的信息产生了意义非凡的扩展。
- 表观遗传有3个密切相关的含义：
 - (1)可遗传的，即这类改变通过有丝分裂或减数分裂能在细胞或个体世代间遗传
 - (2)可逆性的基因表达调节
 - (3)没有DNA序列的变化，或不能用DNA序列变化来解释

- 同卵双生子(**monozygotic twins, MZ**)是由一个受精卵分裂发育而成的双胞胎，二者具有完全相同的基因组。从经典遗传学的角度，使用短串联重复序列等遗传标记均不能对其进行有效的个体甄别。
- **MZ**在**DNA**甲基化水平、**X**染色体失活以及组蛋白位点特异性乙酰化上存在明显差异，这种差异随着年龄的增长而显著增加。
- 同卵双生子在某些**CpG**岛区域的甲基化程度的差异甚至超过了无关个体间的差异。

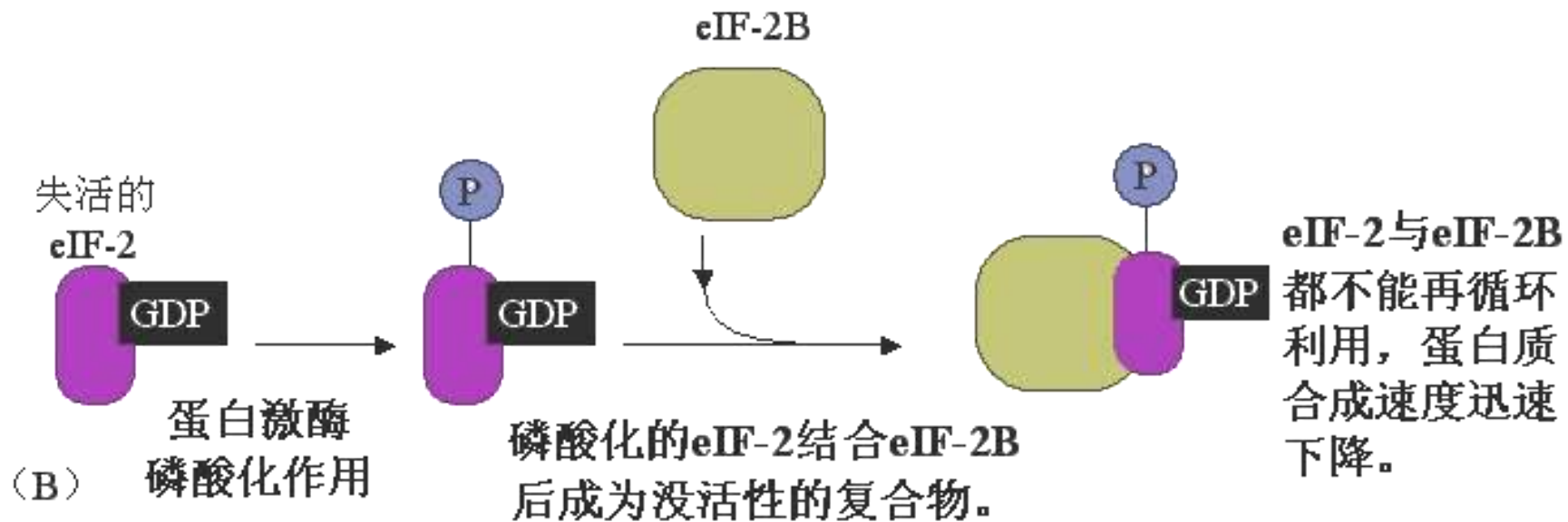
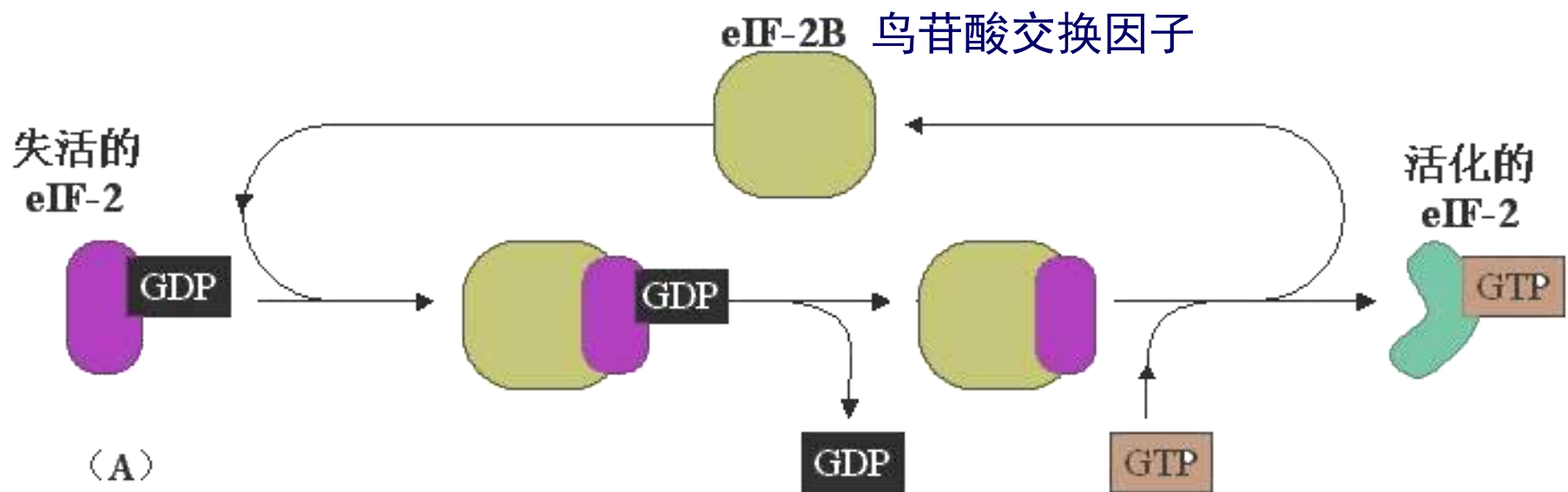
翻译水平的 基因表达调控

**Translational Regulation of
Eukaryotic Gene Expression**

一、翻译起始因子-2的磷酸化 调节蛋白质翻译

翻译起始因子eIF-2 (eukaryotic initiation factor, eIF-2) 与GTP形成的复合物介导Met-tRNA与核糖体小亚基的结合。

真核细胞内各种条件的变化可激活特殊的蛋白激酶，使翻译起始因子eIF-2磷酸化。



低血红素、亚砷酸盐中毒或热休克

氨基酸饥饿或紫外线损伤

双链RNA

内质网应激

HRI

GCN2

PKR (protein kinase RNA)

PERK

eIF-2磷酸化

整体翻译水平的降低和特定 mRNA 翻译的增强

二、mRNA非翻译区的调控

- **5'UTR (untranslated region)** 包括帽子结构、起始密码AUG和二者之间的DNA序列。
- 帽子结构和AUG的距离影响翻译起始的效率
- 上游开放阅读框架 (upstream open reading frames , uORFs) : 在有些mRNA分子中, 起始密码子AUG的上游 (5'UTR) 有一个或几个AUG, 这些上游开放阅读框架与正常的开放阅读框架多不一致, 翻译后很快会遇到终止密码子而释出无功能的多肽。

- ❖ **易遗漏扫描：**当不利于识别时，扫描的核糖体小亚基有时可以无视第一个AUG而滑向第二个，甚至第三个AUG，这种现象被称为易遗漏扫描（leaky scanning），可以使一个mRNA分子产生两个或更多仅氨基末端不同的相关的蛋白质。在一些情况下，细胞可以通过易遗漏扫描调节这些不同长度蛋白质的相对丰度。

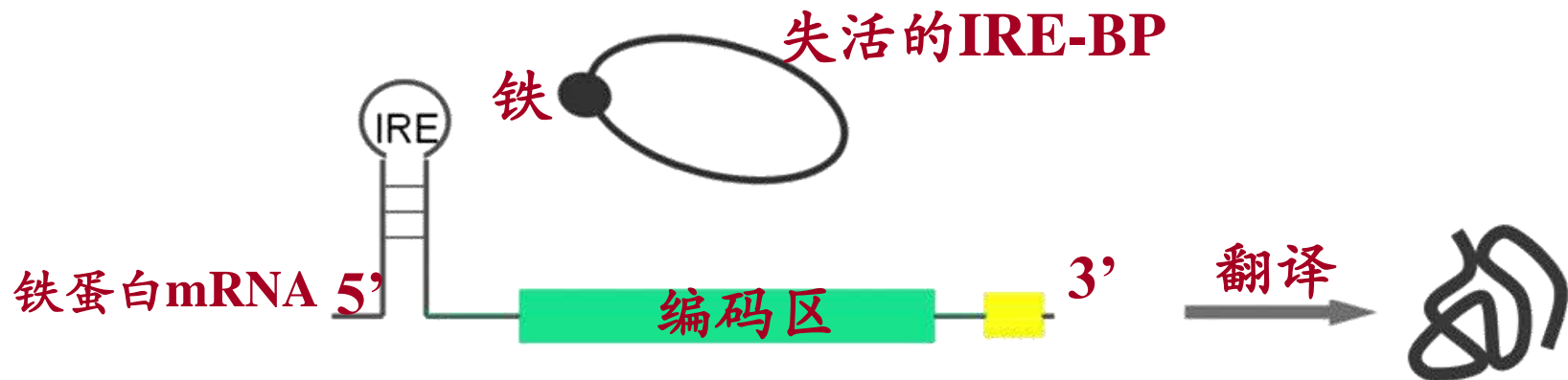
- **3'UTR (untranslated region)** 包括终止密码、PolyA和二者之间的DNA序列。
- 某些抑制蛋白质识别特殊mRNA分子的3'UTR，通过干扰3'Poly(A)尾与5'端帽的联络而减少翻译的启动。
- 终止密码的旁侧序列影响终止密码的作用，与翻译终止的调控密切相关。

三、mRNA特异结合蛋白的调控

低铁状态



高铁状态



IRE-BP对铁蛋白mRNA翻译的负性调节

Key points

- ❖ 顺式作用元件

 - 启动子

 - 增强子(enhancer)

 - 沉默子(silencer)

- 转录因子

 - DNA结合域

 - 转录激活域

- miRNA与siRNA

- 可变剪接

专题讨论

基因及基因表达与疾病发生的关系

- 时间：3月20日1-2节
- 每组3-5人，20min
- 分工协作：查阅资料、制作PPT、课堂陈述
- 选择一个主题，避免泛泛而谈
- PPT需上交作为存档(注明人员组成及分工)，
计入平时成绩