

## 第一部分：原核生物的基因调控的复习

原核生物通过操纵子 (Operon) 实现对基因表达的精细调控 (Fine Control)，主要涉及分解代谢 (catabolism) 和合成代谢 (anabolism) 两类基因的调控。

### 1. 乳糖操纵子 (Lac Operon)

- **结构:** 包括调控基因 (lacI)、启动子 (Promoter)、操纵序列 (Operator) 以及三个结构基因 (lacZ, lacY, lacA)。 lacI 编码阻遏蛋白 (Repressor protein)。
- **调控机制:**
  - **负调控 (Negative Control):**
    - **无乳糖时:** lacI 基因表达的阻遏蛋白单体形成四聚体 (Tetramer) 并与操纵序列结合，阻止 RNA 聚合酶转录结构基因，操纵子关闭。
    - **有乳糖时:** 乳糖的代谢产物异乳糖 (Allolactose) 作为诱导剂 (Inducer)，与阻遏蛋白结合，使其构象改变，从操纵序列上脱落，结构基因得以转录，操纵子开启 (解除阻遏, derepression)。
  - **正调控 (Positive Control) / 代谢物阻遏 (Catabolite Repression):**
    - **有葡萄糖时:** 葡萄糖水平高，导致细胞内 cAMP (环腺苷酸) 水平低。cAMP 不能与 CAP (分解代谢物激活蛋白, Catabolite Activator Protein) 结合形成 cAMP-CAP 复合物。因此，即使有乳糖存在，RNA 聚合酶与启动子的结合效率也较低，转录水平不高。这是因为腺苷酸环化酶 (Adenylate Cyclase) 的活性受到葡萄糖分解代谢产物的抑制。
    - **无葡萄糖时:** 葡萄糖水平低，导致细胞内 cAMP 水平升高。cAMP 与 CAP 结合形成 cAMP-CAP 复合物，该复合物结合到 Lac 启动

子上游的特定位点，促进 RNA 聚合酶与启动子的结合，从而高效启动转录。

- **双重调控总结:**

- 有葡萄糖，无乳糖: 操纵子被阻遏 (负调控关闭，正调控关闭)。
- 有葡萄糖，有乳糖: 操纵子解除阻遏，但正调控未激活，转录水平低 (刹车松开，油门未踩)。Lac 启动子本身是弱启动子。
- 无葡萄糖，无乳糖: 操纵子被阻遏。此时其他糖类如 galactose (半乳糖) 或 arabinose (阿拉伯糖) 的操纵子可能会开启 (如果这些糖源存在)。
- 无葡萄糖，有乳糖: 操纵子解除阻遏且正调控激活，转录水平高 (刹车松开，油门踩下)，形成开放型启动子复合物 (Open promoter complex)。

## 2. 色氨酸操纵子 (Trp Operon)

- **结构:** 包括阻遏基因 (trpR)、启动子、操纵序列、前导序列 (leading segment, L.s.) 包括衰减子 (Attenuator)，以及编码色氨酸合成途径中各种酶的结构基因 (trpE, D, C, B, A)。
- **调控机制:**
  - **阻遏调控:**
    - 色氨酸充足时: 色氨酸作为辅阻遏物 (co-repressor) 与 TrpR 蛋白结合，激活的阻遏蛋白结合到操纵序列，阻止转录。
    - 色氨酸缺乏时: TrpR 蛋白无活性，不能结合操纵序列，操纵子开启。
  - **衰减调控 (Attenuation):** 通过转录与翻译的偶联 (Coupling of transcription

and translation) 实现对转录的精细调控。前导序列编码一个含连续色氨酸密码子 (UGGUGG) 的短肽。

- **色氨酸缺乏时 (Tryptophan starvation):** 核糖体在翻译前导肽时会在色氨酸密码子处停顿, 导致 mRNA 形成一种特殊的二级结构 (抗终止子发卡结构, antiterminator hairpin / non-terminator hairpin), 使得 RNA 聚合酶可以继续转录下游的结构基因。衰减作用被解除。
- **色氨酸充足时 (Tryptophan abundance):** 核糖体顺利通过色氨酸密码子, 导致 mRNA 形成另一种二级结构 (终止子发卡结构, terminator hairpin), RNA 聚合酶在该结构后的 UUUUUU 序列处终止转录。

- **Lac Operon 与 Trp Operon 的不同:**

- **代谢类型:** Lac Operon 参与分解代谢 (分解乳糖), Trp Operon 参与合成代谢 (合成色氨酸)。
- **诱导物/辅阻遏物:** 乳糖是 Lac Operon 的诱导剂; 色氨酸是 Trp Operon 的辅阻遏物。
- **调控逻辑:** 无乳糖时 Lac Operon 关闭; 色氨酸缺乏时 Trp Operon 开启。

### 3. 核糖开关 (Riboswitch)

- 是 mRNA 5' 非翻译区 (UTR) 的一部分, 能直接结合小分子配体 (Ligand), 并根据配体的结合与否改变其二级结构, 从而调控基因表达, 通常影响转录终止或翻译起始。
- 包含适体域 (Aptamer) 用于结合配体, 和表达平台 (Expression platform) 用于调控基因表达。

## 第二部分：真核生物的 RNA 聚合酶及其启动子

### 1. 真核细胞中的 RNA 种类与比例

- 主要有三类 RNA: mRNA (信使 RNA), rRNA (核糖体 RNA), tRNA (转运 RNA)。
- 还有定义尚不完全清楚的 RNA: hnRNA (核内不均一 RNA, heterogenous nuclear RNAs), 大部分是 mRNA 的前体 (precursors); snRNA (核内小 RNA, small nuclear RNAs), 参与 hnRNA 向 mRNA 的成熟 (maturation) 过程。
- RNA 比例:
  - rRNA: **80-85%** (包括 28S, 18S, 5.8S, 5S rRNA)
  - tRNA 及小分子 RNA (包括 snRNA): 10-15%
  - mRNA (hnRNA): 1-5%
- 尽管 mRNA 含量最少, 但它是细胞中最重要的 RNA 类型, 通常重要分子的拷贝数较少。

### 2. 真核生物的 RNA 聚合酶 (RNA Polymerase, RNAP)

- 原核生物只有一种 RNA 聚合酶 (核心酶由  $\beta$ ,  $\beta'$ ,  $2\alpha$  亚基组成, 全酶还包括  $\sigma$  因子, 有时有  $\omega$  亚基), 通过更换  $\sigma$  因子来识别不同启动子, 负责合成所有种类的 RNA。
- 真核生物细胞核 (Nucleus) 中有**三种**主要的 RNA 聚合酶, 每种负责转录特定类型的基因, 并识别不同类型的启动子。
- 分离与鉴定: 通过 DEAE-Sephadex 离子交换层析法 (ion-exchange chromatography) 可将三种聚合酶分离出来, 按洗脱顺序命名为 RNA Pol I, RNA

Pol II, RNA Pol III。

- **细胞内定位 (Localization):**

- **RNA Pol I:** 主要位于核仁 (Nucleolus), 负责 **rRNA** 的合成。
- **RNA Pol II:** 位于核质 (Nucleoplasm)。
- **RNA Pol III:** 也位于核质 (Nucleoplasm)。
- 三种 RNA 聚合酶均位于细胞核内。

- **转录产物:**

- **RNA Pol I:** 合成大的 rRNA 前体 (45S rRNA), 该前体再加工成熟为 28S, 18S 和 5.8S rRNA (在脊椎动物中)。
- **RNA Pol II:** 合成 hnRNA (**mRNA** 前体) 和大部分 snRNA。
- **RNA Pol III:** 合成 5S rRNA 前体、tRNA 前体以及一些其他小 RNA (如 U6 snRNA, 7SL RNA, 7SK RNA)。

### 3. RNA 聚合酶 II (RNA Polymerase II) 结构与功能

- 研究最清楚的是面包酵母 (baker's yeast) 的 RNA Pol II。
- **亚基组成:** 包含 **12** 个亚基, 命名为 RPB1 至 RPB12 (**RPB** 代表 **RNA Polymerase B**)。
  - **鉴定方法:** 表位标签技术 (epitope tagging) 被用于鉴定和纯化这些亚基。
  - **亚基分类:**
    - **核心亚基 (Core subunits):** RPB**1**, RPB**2**, RPB**3**。对酶活性绝对必需, 与细菌 RNA 聚合酶的核心亚基同源。

- **RPB1**: 负责结合 DNA, 类似细菌的  $\beta'$  亚基。
- **RPB2**: 位于或接近核苷酸连接的活性位点, 类似细菌的  $\beta$  亚基。
- **RPB3**: 与细菌的  $\alpha$  亚基在大小、化学计量 (stoichiometry, 每个全酶有两个单体) 及部分序列上相似。
- **共同亚基 (Common subunits)**: **RPB5**, **RPB6**, **RPB8**, **RPB10**, **RPB12**。  
存在于酵母三种核 RNA 聚合酶中, 表明它们在转录中起基础性作用, 但具体功能尚不完全清楚。
- **非必需亚基 (Nonessential subunits)**: **RPB4**, **RPB9**。在正常温度下不是酶活性所必需的, 但这些亚基的突变体在高温或低温下可能无法存活。
  - **RPB4**: 可能参与锚定 (anchoring) **RPB7** 到酶上。**RPB4** 和 **RPB7** 相对容易从 RNA Pol II 上解离 (dissociate), 可能在不同聚合酶间穿梭 (shuttle)。
- **其他必需亚基**: **RPB7**, **RPB11**。对聚合酶活性绝对必需, 其突变体无活性 (viable)。
- **RPB1 亚基的异质性 (Heterogeneity)**:
  - 主要由其羧基末端结构域 (Carboxyl-terminal domain, **CTD**) 的磷酸化 (Phosphorylation) 程度不同引起。
  - **CTD 结构**: 由一个富含丝氨酸的**七肽重复序列** (Tyr-**Ser**-Pro-Thr-**Ser**-Pro-**Ser**) 多次重复构成。
  - **磷酸化位点**: 主要发生在七肽重复序列中的丝氨酸 (Ser) 和苏氨酸

(Thr), 有时酪氨酸 (Tyr) 也会被磷酸化。

- **RPB1 的不同形式:**

- **IIa 形式:** 未磷酸化的形式 (或低磷酸化)。
- **IIo 形式:** 高度磷酸化的形式。
- **IIb 形式:** CTD 被蛋白酶切除后的形式。
- 这些形式可以通过激酶 (Kinase) 磷酸化、磷酸酶 (Phosphatase) 去磷酸化以及蛋白酶 (Protease) 切割而相互转换。

- **不同形式的功能:**

- **IIa 形式:** 主要结合到启动子上, 参与转录起始。
- **IIo 形式:** 主要参与转录延伸 (transcript elongation)。
- CTD 的磷酸化状态变化决定了 mRNA 前体加工 (processing) 相关蛋白 (如剪接、加帽、多聚腺苷酸化所需蛋白) 的结合与解离。

#### 4. 真核生物的启动子 (Promoters)

- 启动子是 RNA 聚合酶特异性结合的一段 DNA 序列, 是基因的组成部分, 像开关一样控制基因转录的活性。
- 三种真核 RNA 聚合酶识别不同类型的启动子, 转录不同类别的基因。

##### a. II 类启动子 (Class II Promoters) - RNA Pol II 识别

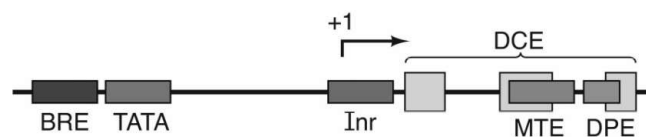
- **结构:** 通常包含两部分: 核心启动子 (Core promoter) 和近端启动子 (Proximal promoter)。

- **核心启动子功能**: 在基础水平 (basal level) 吸引通用转录因子 (General Transcription Factors, GTFs) 和 RNA Pol II, 并设定转录起始位点 (Transcription Start Site, TSS) 及转录方向。
- **近端启动子功能**: 帮助吸引 GTFs 和 RNA Pol II。
- **核心启动子元件 (Core Promoter Elements)**: 具有模块化 (modular) 特点, 可以包含以下元件的不同组合:
  - **TATA 盒 (TATA box)**: 通常位于 TSS 上游约 -25 至 -30 bp 处, 与原核生物的 -10 盒非常相似。
    - **功能**: **主要参与定位 (positioning) 转录起始位点**。在某些基因中, 移除 TATA 盒会导致转录在多个位点起始, 有时甚至会提高转录效率; 而在另一些启动子中, 移除 TATA 盒会严重影响启动子功能, 导致转录无法检测。
    - **分布**: 对于高度**特化基因** (highly specialized genes) 的启动子倾向于**含有 TATA 盒**。而管家基因 (Housekeeping genes, 组成型表达) 和发育调控基因 (developmentally regulated genes) 的启动子通常缺乏可识别的 TATA 盒 (TATA-less)。
  - **起始子 (Initiator, Inr)**: 围绕转录起始位点的保守序列, 对最佳转录 (optimal transcription) 至关重要。
  - **TFIIB 识别元件 (TFIIB Recognition Element, BRE)**: 位于 TATA 盒上游的 DNA 元件, **帮助转录因子 TFIIB 结合到 DNA 上, 最终形成预起始复合物 (preinitiation complex)**。
  - **下游启动子元件 (Downstream Promoter Element, DPE)**: 位于转录起始位



点下游约 +30 bp 处，可以**补偿启动子中 TATA 盒的缺失**。 TATA-less 启动子倾向于拥有 DPEs。

- **基序十元件 (Motif Ten Element, MTE)**: 参与转录起始的调控。
- **下游核心元件 (Downstream Core Element, DCE)**: 位于转录起始位点下游，对转录起始有重要影响。
- 大多数启动子至少缺失四个核心元件中的一个。
- **近端启动子元件 (Proximal Promoter Elements)**: 通常位于 II 类核心启动子的上游。
  - 常见的有 **GC 盒 (GC boxes)** (富含 GC 序列) 和 **CCAAT 盒 (CCAAT box)** (发音为 "**cat box**").
  - 它们与核心启动子的不同在于，它们结合的是相对基因特异性的转录因子。
  - 这些上游启动子元件可以是**方向独立**的 (orientation-independent)，但其位置相对依赖 (relatively position-dependent)，这与经典的增强子不同。



**图 10.18 一般性 II 类核心启动子。**核心启动子由 6 个元件组成，从 5' 端至端 3' 依次为：TFIIB 识别元件 (BRE, 紫色)；TATA 框 (红色)；起始子 (绿色)；由三部分构成的下游核心元件 (DCE, 黄色)；十基序元件 (MTE, 蓝色) 和下游启动子元件 (DPE, 橙色)。文中介绍了这些启动子元件的准确定位。

### b. I 类启动子 (Class I Promoters) - RNA Pol I 识别

- 其序列在不同物种间的变异性很大，比 II 类启动子更大。

- 但 I 类启动子确实拥有一个保守序列，即围绕转录起始位点的**富含 AT 的启动子** (AT-rich initiator, rINR)。
- 示例：人类 rRNA 基因启动子包含两个关键区域：
  - **核心元件 (Core element)**: 位于转录起始处，大约在 -45 到 +20 位置。
  - **上游启动子元件 (Upstream Promoter Element, UPE)**: 位于大约 -156 到 -107 位置。
- 这两个启动子元件之间的间距 (spacing) 对启动子活性非常重要。

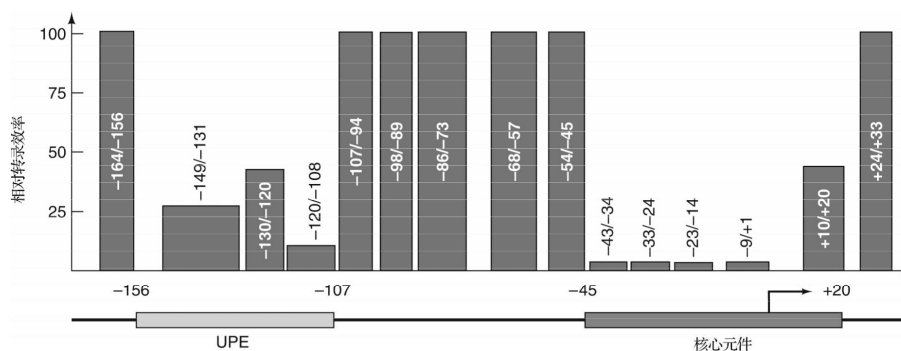


图 10.21 rRNA 启动子的两个元件。Tjian 及同事用接头扫描突变对人类 rRNA 基因的 5'侧翼区进行系统突变，然后利用体外转录体系检测各个区段突变后对启动子起始转录活性的影响。研究者用柱形图说明实验结果。结果分析显示，启动子有两个重要区域：UPE（上游启动子元件）和核心元件。UPE 是最佳转录所必需的，但本底水平的转录不一定需要 UPE 的参与，而核心元件是任何转录所必需的。（Source: Adapted from Learned, R. M., T. K. Learned, M. M. Haltiner, and R. T. Tjian, Human rRNA transcription is modulated by the coordinated binding of two factors to an upstream control element. *Cell* 45: 848, 1986.）

### c. III 类启动子 (Class III Promoters) - RNA Pol III 识别

- RNA Pol III 转录多种编码小 RNA 的基因 (称为 III 类基因)。
- "经典" III 类基因 (Classical class III genes): 如 5S rRNA 基因、tRNA 基因和腺病毒 VA RNA 基因。它们的启动子完全位于基因内部 (internal promoters)。
  - **类型 I (Type I)**: 例如 5S rRNA 基因，其内部启动子包含 Box A、中间元件 (Intermediate element) 和 Box C。通过对非洲爪蟾 (*Xenopus borealis*) 5S rRNA 基因进行 5' 端删除实验，证明了**其启动子位于基因内部**。
  - **类型 II (Type II)**: 例如 tRNA 基因或 VA RNA 基因，其内部启动子包含

Box A 和 Box B。

- **"非经典" III 类基因 (Nonclassical class III genes):** 如 U6 snRNA 基因、7SL RNA 基因、7SK RNA 基因和 Epstein-Barr 病毒 EBER2 基因。它们的启动子与 II 类基因的启动子相似，位于 5'-侧翼区 (5'-flanking region)，可能包含 TATA 盒。

- **类型 III (Type III):** 例如人类 U6 snRNA 基因，其启动子可能包含位于上游的 TATA 盒、近端序列元件 (Proximal Sequence Element, PSE) 和远端序列元件 (Distal Sequence Element, DSE)。

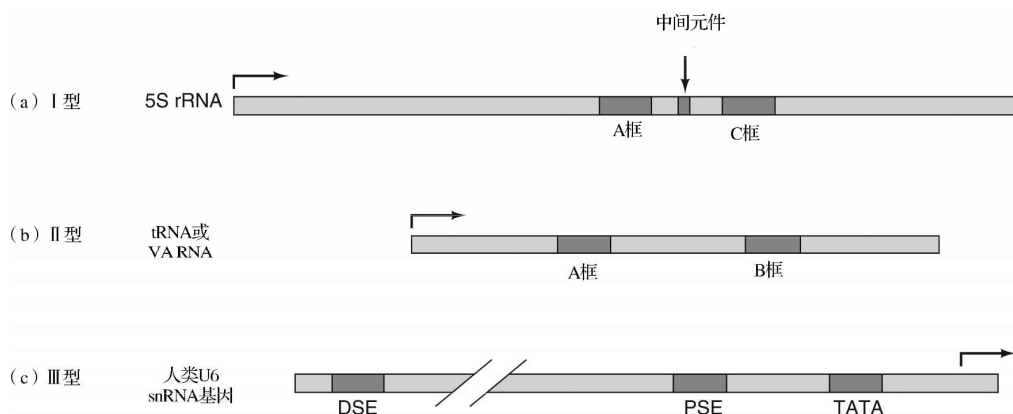


图 10.23 某些 III 类基因的启动子。图中显示了 5S rRNA 基因启动子、tRNA 基因启动子和 U6 RNA 基因启动子的结构，蓝色条框为基因的调控元件。

## 5. 增强子 (Enhancers) 和沉默子 (Silencers)

- 两者都是 DNA 序列，且为非启动子元件。
- **增强子 (Enhancers):**
  - 能够刺激 (stimulate) 转录。
  - 第一个发现的增强子位于 SV40 早期基因的 5'-侧翼区，包含一个 72 bp 的序列重复。
  - 特点: 具有**位置**和**方向不依赖性** (position- and orientation-independent)，即使将其反向或移动到距离启动子很远的位置 (如超过 2 kb)，仍能发挥作

用。

- 作用机制: 通过结合称为激活因子 (activators) 的蛋白质来发挥作用。这种相互作用促进预起始复合物的形成, 从而增强转录。

- **沉默子 (Silencers):**

- 能够抑制 (depress / inhibit) 转录。
- 作用机制: 可能通过使染色质**卷曲** (coil up) 成一种浓缩的 (condensed)、不可到达的 (inaccessible) 且因此失活的 (inactive) 形式, 从而阻止邻近基因的转录。

- **双重功能:** 有时同一个 DNA 元件既可以作为增强子也可以作为沉默子, 这取决于与其结合的蛋白质。

- 例如, 甲状腺激素应答元件 (thyroid hormone response element): 当甲状腺激素受体在没有配体 (甲状腺激素) 的情况下结合时, 它作为沉默子; 而当受体与甲状腺激素一同结合时, 它作为增强子。

## 总结与复习要点

- **原核调控:** 理解 Lac 和 Trp 操纵子的结构、正负调控、阻遏、诱导、衰减等机制, 以及 Riboswitch 的作用原理。
- **真核 RNA 聚合酶:** 掌握三种 RNA 聚合酶的种类、定位、转录产物, 重点关注 RNA Pol II 的亚基组成、RPB1 亚基 CTD 的磷酸化及其功能意义。
- **真核启动子:** 区分 I、II、III 类启动子的结构特征、核心元件 (特别是 TATA 盒、Inr、DPE、BRE 的功能和分布) 以及与相应 RNA 聚合酶的对对应关系。理解经典与非经典 III 类启动子的区别。
- **调控元件:** 明确增强子和沉默子的定义、作用特点 (位置与方向无关性) 以及大

致的作用机制。理解同一 DNA 元件可能具有双重调控功能的原因。