

# 真核生物 II 类基因的转录起始

尤忠胜 汪 垣

(中国科学院上海生物化学研究所, 上海 200031)

**摘要** 真核生物 II 类基因的转录起始是一个非常复杂并且受到严格调控的过程. 它涉及到染色体结构的改变、顺式作用元件和反式作用因子相互作用等多种过程. 关于 II 类基因基础转录起始复合物在启动子区的装配的研究, 目前存在两种模型, 即分步组装模型与 RNAP II 全酶模型, 它们分别是基于体外重建转录实验的结果以及酵母中 RNAP II 全酶巨酶体系的发现而提出的. 在转录激活的分子机制上目前认为存在两个相互联系的过程, 即解抑制作用和真正的激活作用.

**关键词** RNA 聚合酶 II, 基础转录起始复合物, 转录激活

真核生物 II 类基因 (即由 RNA polymerase II 转录的基因) 的转录起始是指起始复合物在启动子区的装配到起始复合物离开启动子区, 在核苷三磷酸存在下生成第一个磷酸二酯键的过程, 它的高度复杂性为基因转录的调控提供了多种可能. 近几年来, 随着酵母遗传学的发展以及分子生物学、生物化学的技术和方法的进步, 科学家们在 II 类基因转录起始过程的研究方面取得了长足进展. 本文将着重介绍 RNA 聚合酶 II (RNAP II) 转录起始复合物在启动子区的组装以及转录起始的调控机制.

## 1 II 类基因启动子区的结构

除 SnRNA 的启动子外, 迄今发现的所有 II 类基因的启动子区都至少包含有以下两类启动子核心元件的一种, 它们是 TATA 盒与起始子元件 Inr (initiator). 前者一般位于转录起始位点上游 - 30 bp 处, 后者在起始位点附近, 并且包含起始位点. 它们对转录起始过程是必不可少的. TF II D 与 TATA 盒、起始子结合蛋白 IBP (initiator binding proteins) 与 Inr 的结合对 RNAP II 在启动子上的精确定位至关重要<sup>[1~3]</sup>.

## 2 参与转录起始的因子

RNAP II 选择性地转录起始需要多种因子

的参与. 目前已发现至少有 20 种以上的蛋白质参与这一过程<sup>[4]</sup>. 这些因子除 RNAP II 以外, 根据其功能大致可以划分为以下几类.

通用转录因子 GTFs (general transcription factors). 目前发现的 GTFs 有 TBP (TATA-binding protein), TF II A, TF II B, TF II E, TF II F, TF II H, TF II J 等. 体外重建 (*in vitro* reconstitution) 转录实验表明, 它们是基础转录器 (basal transcription apparatus) 的组分. 而且, 从酵母到人, RNAP II 和大多数 GTFs 的肽链组成都是高度保守的<sup>[5]</sup>.

激活因子 (activators) 和抑制因子 (repressors). 在结构特征上, 这类因子基本上都有两个功能域, 即 DNA 结合结构域以及转录激活/抑制结构域. 它们通过与基础转录起始复合物中的组分相互作用, 或者直接干扰 TF II D 与 DNA 的结合等方式来实现对基因的转录进行激活或抑制.

辅助激活因子 (coactivators)、中介因子 (mediators) 和接头 (adaptors). 它们是激活因子和抑制因子转录调控作用的中介. 如在 HeLa 细胞发现的 Dr2, ACF 等. 另外, TF II D 中的 TBP 偶联因子 TAFs (TBP-associated factors), 在酵母中发现的 ADA2, ADA3 也属于此类<sup>[1]</sup>.

收稿日期: 1996-06-24, 修回日期: 1996-11-03

### 3 基础转录起始复合物的装配

#### 3.1 分步组装模型

通过 DNA 足迹、凝胶阻滞、以及动力学分析等实验手段, 人们发现, 体外条件下 GTFs 和 RNAP II 通过有序的分步结合途径在启动子区组装成基础转录起始复合物。藉此, 科学家们提出所谓分步组装模型 (stepwise assembly model)。该模型可概括如下。

对于含有 TATA 盒的启动子而言, 首先是 TF II D 与 TATA 盒区结合, 接着 TF II A 结合

到 TF II D 上, 稳定 TF II D 与 TATA 盒的相互作用。此后, TF II A 可能离开也可能留在形成中的复合物中。TF II B 可以识别 TF II D-DNA 复合物或 DA-DNA 复合物 (即 TF II D, TF II A, DNA 三者的复合物, 如下类同), 形成 DB 或 DAB 复合物。然后, TF II F 和 RNAP II 一起结合上去, 形成 DBpolF 或 DABpolF。最后, TF II E, TF II H, TF II J 结合进来, 形成完整的基础起始复合物 (图 1)。这时如果有核苷三磷酸存在, 就可以起始转录, 形成子代 RNA 中的第一个磷酸二酯键。

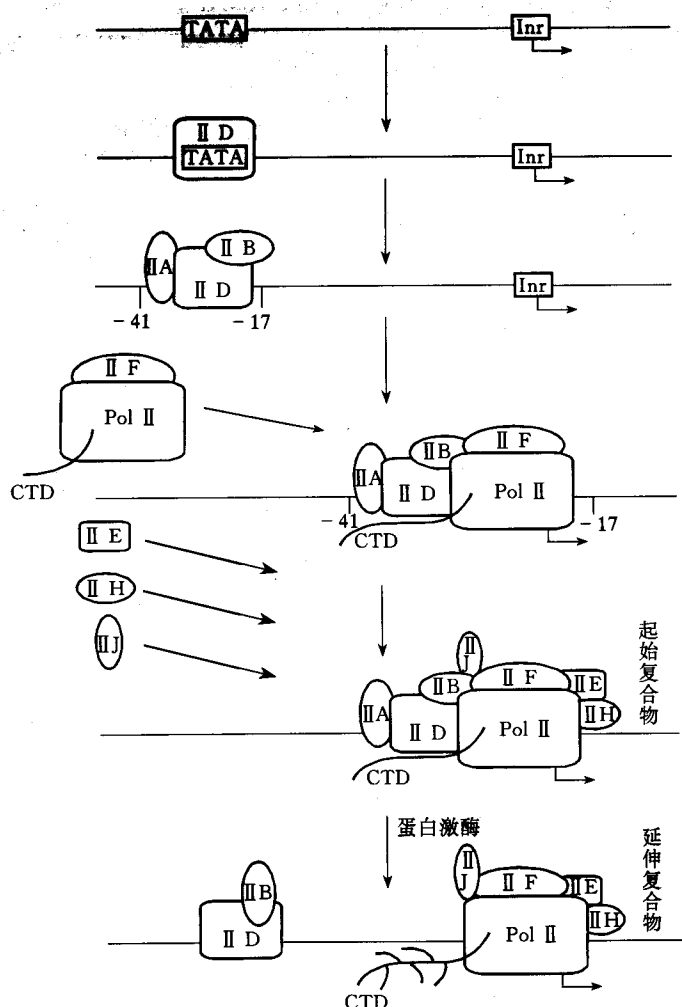


图 1 真核生物 II 类基因转录起始示意图 (分步组装模型)

对于不含 TATA 盒的启动子的转录起始, 目前已至少发现了三种模式: a. RNAP II 直接

识别结合 Inr 元件。含有保守的 CTCA 花样结构的启动子 (如 Ad-MLP, TdT, Ad-IVa2 启

动子) 的 Inr 元件可以被 RNAP II 识别. 在 TF II D, TF II B, TF II F 的存在下, RNAP II 可稳定结合在 Inr 元件上. 以此为核 心, 其他 GTFs 通过因子间的相互作用装配成基础转录起始复合物<sup>[1]</sup>. b. 通过某些 TAFs 识别并结合 Inr 元件. 比如, Tjian 等<sup>[6]</sup> 发现, 果蝇的 TAF150 能独立地特异性地识别包含 Inr 元件的 DNA. c. 通过 IBPs, 如 TF II-I 或 YY1 等与 TBP 或基础转录器中其他成分相互作用来引导转录起始复合物在启动子区的装配<sup>[1]</sup>.

3.2 RNAP II 全酶模型

Whitehead 生物医学研究所和 MIT 的两个研究小组通过筛选能够抑制酵母 RNAP II 最大亚基羧端结构域 CTD 突变的突变体, 得到 9 个 SRB (surpressors of RNA polymerase B) 基因, 其产物均参与了转录起始过程<sup>[5]</sup>. 并且,

他们从酵母中 分离纯化得到一种被称为 RNAP II 全酶 (RNAP II holoenzyme) 的巨酶体系, 它是包括 RNAP II、多种 SRB 在内的多因子复合物<sup>[7]</sup>. 研究发现, 它在无 DNA 模板时仍高度稳定, 并在 TBP 和 TF II E 的存在下就可以有效地进行转录起始. 另外, 激活因子 GAL4-VP16 可以刺激它引导的转录, 但纯化的酵母 RNAP II 和 GTFs 对 GAL4-VP16 没有反应. 藉此, 他们提出了不同的 RNAP II 转录起始的模型 (图 2). 该模型认为, 转录起始过程中, 首先是 TF II D 与 DNA 模板结合, 接着, RNAP II 全酶通过其组分分别与 TF II D-DNA 复合物、TF II E 和激活因子等相互作用, 从而形成稳定的有效的转录起始复合物<sup>[5]</sup>.

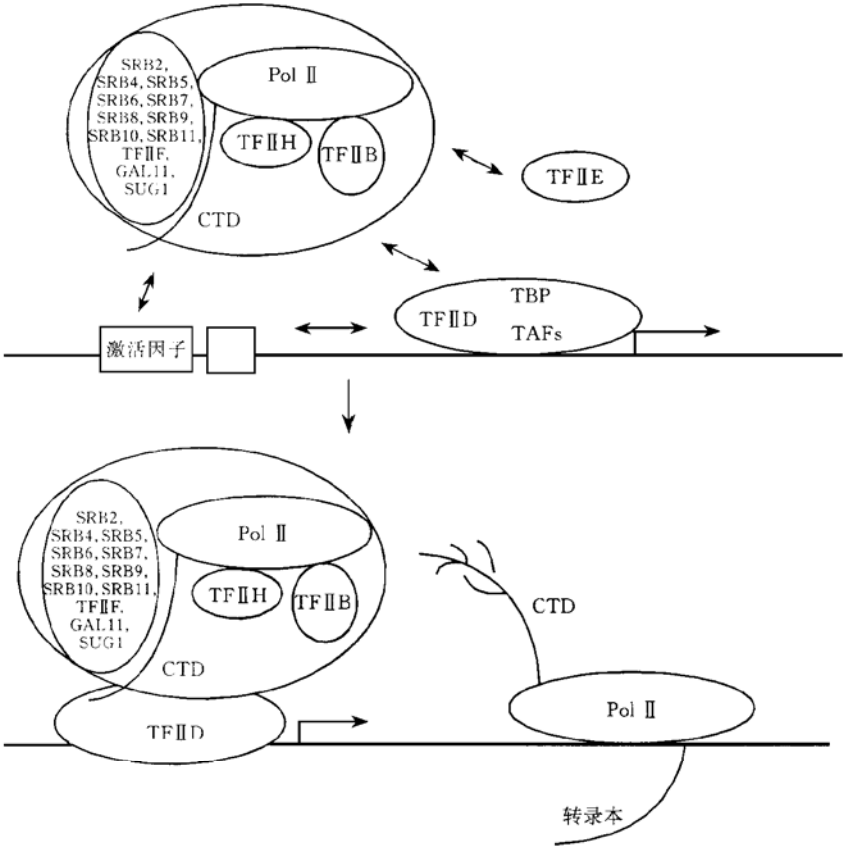


图 2 真核生物 II 类基因转录起始示意图 (RNAP II holoenzyme 模型)

Kim 等<sup>[8]</sup>在酵母中发现了另一种形式的 RNAP II 全酶, 它更象是前者的亚复合物, 包含有 RNAP II, TF II F 和多种 SRB 蛋白, 但没有 TF II B 和 TF II H, 可能后者是前者在纯化过程中的产物, 也可能这两种形式同时存在于细胞内。

另外一个有力的证据是, 在体外条件下, 哺乳动物的 RNAP II、TF II F、TF II H 和 TF II B 也可以不依赖于 DNA 模板而形成稳定的复合物<sup>[5]</sup>。

但是, 现在已发现的 RNAP II 全酶并未包括所有的参与转录起始的基本成分。另外, 尚不清楚这类复合物中的 GTFs 是否按固定比例存在。目前使用的分离纯化方法所得到的结果表明, 至多只有 50% 的 RNAP II 与 SRB 等蛋白以复合物的形式存在<sup>[5]</sup>。因此, 后一模型仍需要更多的实验证据。以上两种模型是根据不同的研究方法所得到的结果而提出的, 目前尚不能确定到底哪一个更能反映体内的真实情况, 也可能细胞内转录起始复合物的组装本身就存在多种方式。

## 4 转录起始调控的分子机制

在不考虑其他调控成分的情形下, 由核心启动子区 (TATA 盒, Inr 元件) 引导的转录, 称为基础转录。在细胞内, 基因的转录显然受到严格的调控。由于 DNA 被包裹在核小体中, 加上非组蛋白抑制因子 (包括序列特异性的静子结合蛋白等) 的作用, 大多数基因在体内一般处于沉默的状态。目前认为, 基因的转录激活可分为解抑制作用 (antirepression) 和真正的激活作用 (true activation) 两个过程<sup>[9]</sup>。

### 4.1 解抑制作用

近来研究表明, 某些启动子/增强子结合蛋白的一个重要功能就是解除染色体介导的抑制作用<sup>[9]</sup>。比如, 某些转录激活因子是通过改变核小体的结构从而有利于 GTFs 与模板的结合来激活基因转录的<sup>[1]</sup>。组蛋白的修饰作用以及核小体定位 (nucleosome positioning) 也有明显的解抑制作用。已证明组蛋白 H3 的酰化

作用有利于转录因子接触模板。一些启动子区的特定核小体结构的形成有利于转录因子识别并结合特异的序列 (如 MMTV 的启动子), 或介导远距离的 DNA 元件间的相互通讯 (如果蝇的 hsp60 和醇脱氢酶的启动子等)<sup>[10]</sup>。

### 4.2 真正的激活作用

真正的激活作用是指激活因子与基础转录器相互作用, 从而激活基因的转录起始。它可以分为两种方式。一种是通过与基础转录器中的组分直接作用来实现的。TF II A、TF II B、TF II D、TF II E、TF II F 以及 TF II H 等都是潜在的激活蛋白的靶因子<sup>[6]</sup>。RNAP II 本身也是转录调控因子的作用位点。它的最大亚基羧基末端结构域 CTD 中的重复序列含有丰富的羟基基团, 为这种作用提供了多种位点<sup>[5]</sup>。

激活蛋白与基础转录器相互作用的另一种情形是需要中介物的介导。体外重建转录实验表明, 重组的 TBP 可以识别 TATA 盒元件, 并介导其他基础转录组分的介入, 但 TBP 并不能象 TF II D 那样引导 SP1, MLTF, NTF-1 等调节蛋白介导的激活转录, 表明 TF II D 中的 TAFs 在功能上起着辅助激活因子 (coactivators) 的作用<sup>[2]</sup>。目前在果蝇中已经发现了 8 种以上的 TAFs (250、150、110、80、60、40、30 $\alpha$ 、30 $\beta$  等)<sup>[6]</sup>。研究证实, 果蝇的 TAF40 与 VP16, TAF150 与 NTF-1, TAF110 与 SP1 以及人的 TAF55 与 CTF 可以相互作用<sup>[1]</sup>。

在 HeLa 细胞的抽提物中也分离得到另外一些起中介作用的蛋白。如起正调控作用的有 PC1、PC2、PC3、PC4、ACF 等, 它们并不影响基础转录, 但能增强激活因子的效应; 起负调控作用的有 NC1、NC2/Dr1、Dr2/PC3/topoisomerase I、MOT/ADI、DT1-4、TUP1 等, 它们抑制基础转录, 但这种抑制可以为激活因子所解除。在酵母中也存在中介蛋白, 如 ADA2 在 GAL4-VP16 与 TF II B 的相互作用中起桥梁作用, 被称为接头 (adaptor)<sup>[11]</sup>。

近来, 人们还发现, 某些天然的增强子发挥作用需要形成立体结构高度专一的核蛋白复

合物. 这种复合物的形成需要一些序列特异性的因子的存在. 这些因子是一些结构性的成分, 是该复合物的形成所必需的. 如鼠 T 细胞受体  $\alpha$  基因 TCR $\alpha$  的增强子上结合的 LEF-1 蛋白<sup>[11]</sup>、人  $\beta$  干扰素基因 IFN- $\beta$  的增强子上结合的 HMG I(Y)<sup>[12]</sup> 等. 它们均介导了 DNA 的弯曲, 使结合在附近区域的因子间能相互作用, 形成有效的三维增强子复合物.

尽管科学家们在 RNAP II 转录起始研究方面取得了较大的进展, 但显然还有许多问题亟待解决. 首先仍须弄清包括 RNAP II 在内的众多的蛋白质因子究竟是以什么方式在启动子区形成有效的基础转录器的. 其次, 深入地研究各因子之间的相互作用及其效应将有助于理解基因转录正负调控的分子机制. 目前尚不清楚, 各顺式元件之间以及各因子之间在体内到底是如何发挥效应的. 有人认为, 当激活蛋白结合在转录起始位点附近区域时可能是增强转录的速率, 但是若结合在远距离位点时, 则可能是起着稳定转录的作用<sup>[13]</sup>. 再次, 染色质、DNA 拓扑结构以及其他调节因子对转录起始的影响也将是研究的焦点之一.

## 参 考 文 献

- 1 Zawel L, Reiberg D. Common themes in assembly and function of eukaryotic transcription complexes. *Annu Rev Biochem*, 1995, **64**: 533~ 561
- 2 Zawel L, Reinberg D. Advance in RNA polymerase II transcription. *Curr Opin Cell Biol*, 1992, **4** (4): 488~ 495
- 3 Kollmar R, Farnham P J. Site specific initiation of transcription by RNA polymerase II. *Proc Soc Exp Biol Med*, 1993, **203** (2): 127~ 139
- 4 Buratowski S. The basics of basal transcription by RNA polymerase II. *Cell*, 1994, **77** (1): 1~ 3
- 5 Koleske A J, Young R A. The RNA polymerase II holoenzyme and its implications for gene regulation. *TIBS*, 1995, **20** (3): 113~ 116
- 6 Tjian R, Maniatis T. Transcriptional activation: a complex puzzle with few easy pieces. *Cell*, 1994, **77** (1): 5~ 8
- 7 Koleske A J, Young R A. An RNA polymerase II holoenzyme responsive to activators. *Nature*, 1994, **368** (6470): 466~ 469

- 8 Kim Y J, Bjorklund S, Li Y *et al.* A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell*, 1994, **77** (4): 599~ 608
- 9 Paranjape S M, Kamakaka R T, Kadonaga J T. Role of chromatin structure in the regulation of transcription by RNA polymerase II. *Annu Rev Biochem*, 1994, **63**: 265~ 297
- 10 Wolffe A P. Transcription: in tune with the histones. *Cell*, 1994, **77** (1): 13~ 16
- 11 Giese K, Cox J, Grosschedl R. The HMG domain of lymphoid factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures. *Cell*, 1992, **69** (1): 185~ 195
- 12 Du W, Thanos D, Maniatis T. Mechanisms of transcription synergism between distinct virus inducible enhancer elements. *Cell*, 1993, **74** (5): 887~ 898
- 13 Walters M C, Magis W, Fiering S *et al.* Transcriptional enhancers act *in cis* to suppress position effect variegation. *Genes Dev*, 1996, **10** (2): 185~ 195

**Transcription Initiation of Eukaryotic Class II Genes.** YOU Zhongsheng, WANG Yuan (Shanghai Institute of Biochemistry, The Chinese Academy of Sciences, Shanghai 200031, China).

**Abstract** The transcription initiation of eukaryotic class II genes is very complicated and strictly regulated. Many events such as the alteration of chromatin structure, the interaction between *cis*-elements and *trans*-factors are involved in this process. Based on the results of *in vitro* reconstitution transcription and the discovery of RNAP II holoenzyme in yeast respectively, two models, step-wise assembly model and RNAP II holoenzyme model, were proposed on the assembly of basal transcription initiation complex at promoters. Two related processes, antirepression and true activation, were supposed to be the molecular mechanism of the transcriptional activation.

**Key words** RNA polymerase II, basal transcription initiation complex, transcriptional activation