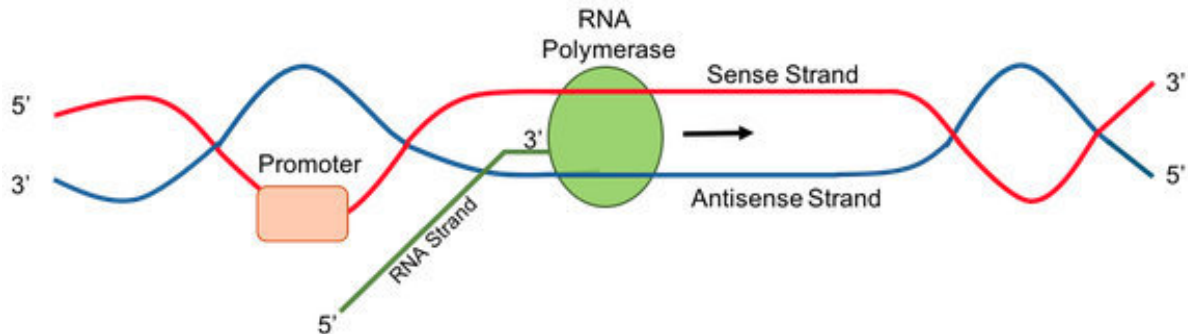


[Promoters][<https://www.addgene.org/mol-bio-reference/promoters/>]

启动子是DNA区域上基因转录起始的位置。对于表达至关重要，该位置控制RNA核糖体结合DNA。RNA聚合酶转录DNA为mRNA最终翻译为功能蛋白。



启动子大约100-1000bp长，典型地临近转录基因编码链的5'上游区域。编码链为编码密码子的DNA链，并且其序列对应为mRNA转录产物。反义链为模板链非编码链，是RNA聚合酶转录的链。

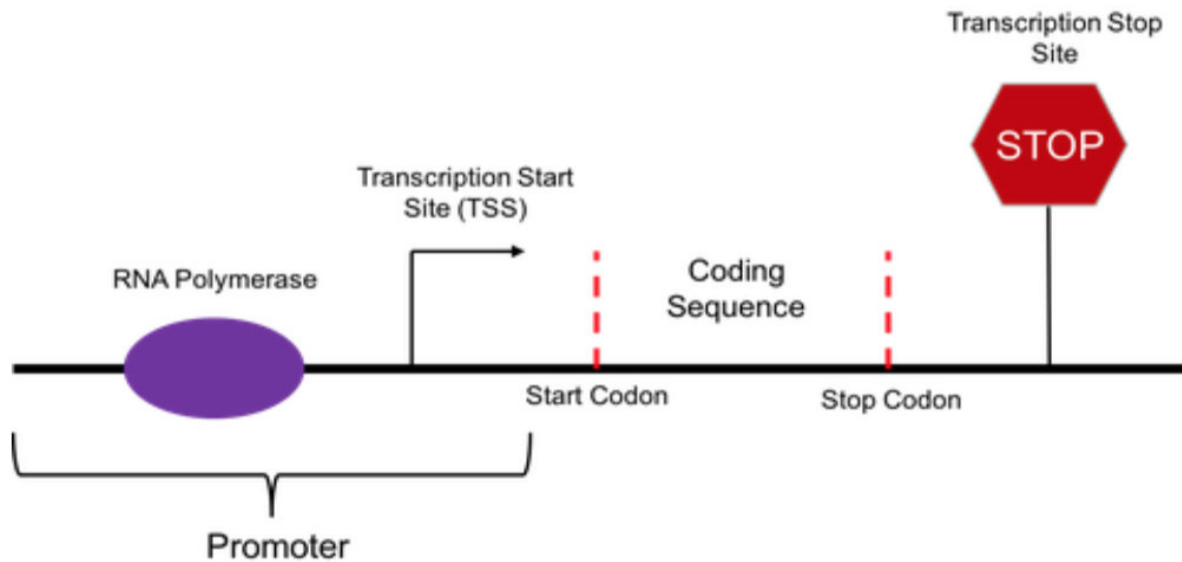
称为response elements的DNA序列位于启动子区域内，提供RNA聚合酶和转录因子稳定的结合位点。转录因子为蛋白，能够招募RNA聚合酶，控制和调节DNA转录为mRNA。

启动子位置的结合在细菌和真核生物中差异很大。在细菌中，core RNA polymerase需要相关的sigma因子用于启动子的识别和结合。然而，真核生物更加复杂。真核生物需要最少7个转录因子用于RNA聚合酶 II(一个真核生物特异性的RNA聚合酶)来结合到启动子区域。启动子被多种DNA调节序列所控制，包括enhancers, boundary elements, insulators和silencers。

promoter regions

启动子区域由3个主要蛋白所构成，core promoter, proximal promoter和dital promoter。

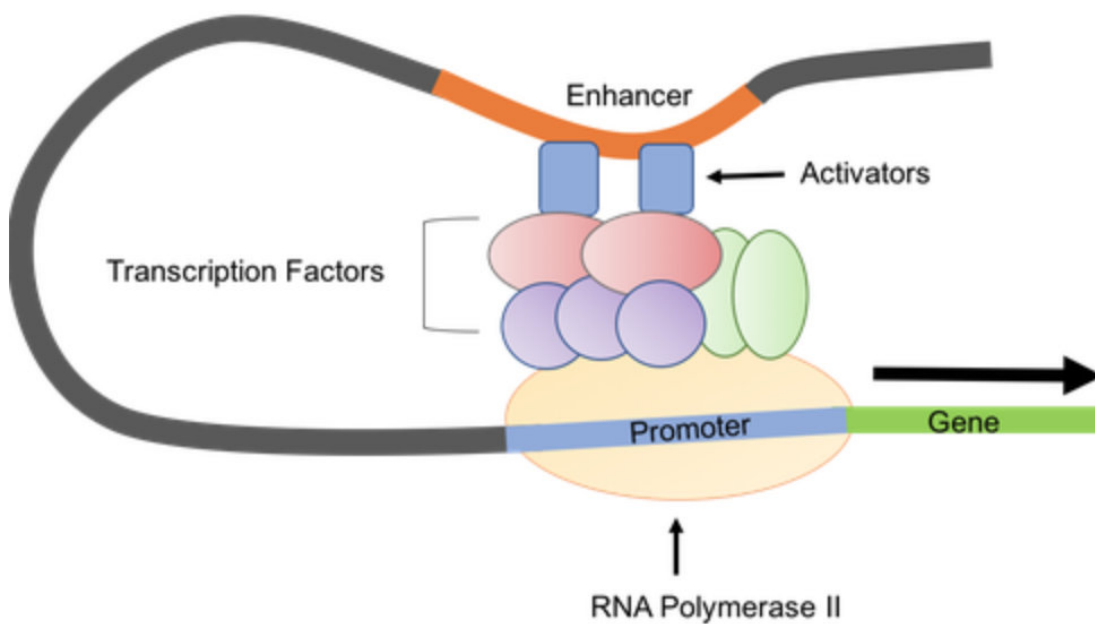
core promoter区域最邻近起始密码子，包含RNA聚合酶结合位点，TATA box，和转录起始位置(TSS)。RNA聚合酶稳定结合到该core promoter区域，开启模版链的转录。TATA box为位于core promoter区域的DNA序列(5' - TATAAA -3')，一般的转录因子和组蛋白可以结合。组蛋白是真核生物中发现的蛋白，将DNA包装成为核小体。组蛋白的结合阻止了转录起始，相反，转录因子促进转录的开始。核心启动子最3短部分(closest to the gene's start codon)，为TSS区域，为转录实际上开始的区域。只有在真核生物和古细菌中，含有TATA box。大多数原核生物包含一个功能上相似的序列， Pribnow box，常包含6个核苷酸，TATAAT。



proximal promoter, 位于core promoter的上游区域, 该区域包含很多主要的调节单元。proximal promoter位于TSS上游约250bp, 一般为转录因子结合位置。

distal promoter, 启动子区域的最后一部分为distal promoter, 位于proximal promoter上游。该区域同样包含转录因子结合位点, 但是大部分包含调节单元。

Eukaryotic Promoters



Eukaryotic Transcription

真核生物启动子比原核生物启动子要复杂的多。真核生物启动子长跨越很大范围的DNA序列。常含有多个调节单元, 例如距离TSS数千bp的增强子。真核生物结构上如此负责, DNA倾向与自身折叠, 有助于多个物理空间远距离DNA序列影响指定的基因转录。TATA-结合蛋白结合到TATA box, 帮助RNA聚合酶结合。一个转录复合体结构上包含RNA聚合酶和多个转录因子蛋白。

bacterial promoters

细菌的启动子包含两个短DNA序列位于-10位置(10bp 5'端)和-35bp转录起始位置(距离TSS)。Paribnow box(普里布诺盒 Pribnow box, TATA)位于-10位置, 是转录起始所必需的。-35位置, 典型地包含TTGACA序列, 该单元控制转录的速率。细菌细胞包含sigma因子, 帮助RNA聚合酶集合到启动子区域。每个sigma因子识别不同的core promoter序列。

operons

尽管原核生物的转录比真核生物的转录要简单, 但是细菌仍然拥有复杂的基因调节系统, 例如 operons。operons为一个不同的基因的cluster, 共同由一个promoter和operator所调控。operons一般存在于原核生物中, 尤其是细菌。operons包含一个启动子, 被RNA聚合酶所识别, 一个DNA片段, repressor或activator能够结合。

operons的调节可以是正负的。负调控可抑制operons, 一般由一个抑制蛋白结合阻止转录。当一个诱导分子结合到repressor时, 它改变其构象, 阻止抑制蛋白结合到operator上, 因此转录开始。Lac操纵子时细菌中负调控的operon。

正调控可抑制operons, 一般处于转录状态, 直到有reprssor/corepressor蛋白结合到operator区域, 阻止转录。涉及到色氨酸合成的trp operon就是正调控operon。

Common Bacterial Promoters used in Research

Promoter	Expression	Description
T7	Constitutive but requires T7 RNA polymerase	Promoter from T7 bacteriophage
Sp6	Constitutive but requires Sp6 RNA polymerase	Promoter from Sp6 bacteriophage
lac	Constitutive in the absense of lac repressor (lacI or lacIq). Can be induced by IPTG or lactose	Promoter from Lac operon
araBad	Inducible by arabinose	Promoter of the arabinose metabolic operon
trp	Repressible by tryptophan	Promoter from <i>E. coli</i> tryptophan operon
Ptac	Regulated like the lac promoter	Hybrid promoter of lac and trp

types of RNA polymerases

启动子控制RNA聚合酶结合到DNA, 开启基因的转录。共有三种的RNA 聚合酶转录不同的基因。

RNA polymerase I, 转录编码核糖体RNA(rRNA), 是细胞核糖体结构的主要成分。核糖体为蛋白合成的场所。

RNA polymerase II, 转录信使RNA(mRNA), 为蛋白的翻译提供稳定的转录模版。

RNA polymerase III, 转录基因编码转运RNA(tRNA), 该adaptor molecules用于转运氨基酸到核糖体, 用于蛋白的合成。同时RNA polymerase III也转录小RNAs, 例如shRNAs和gRNAs。

[Bacterial transcription]

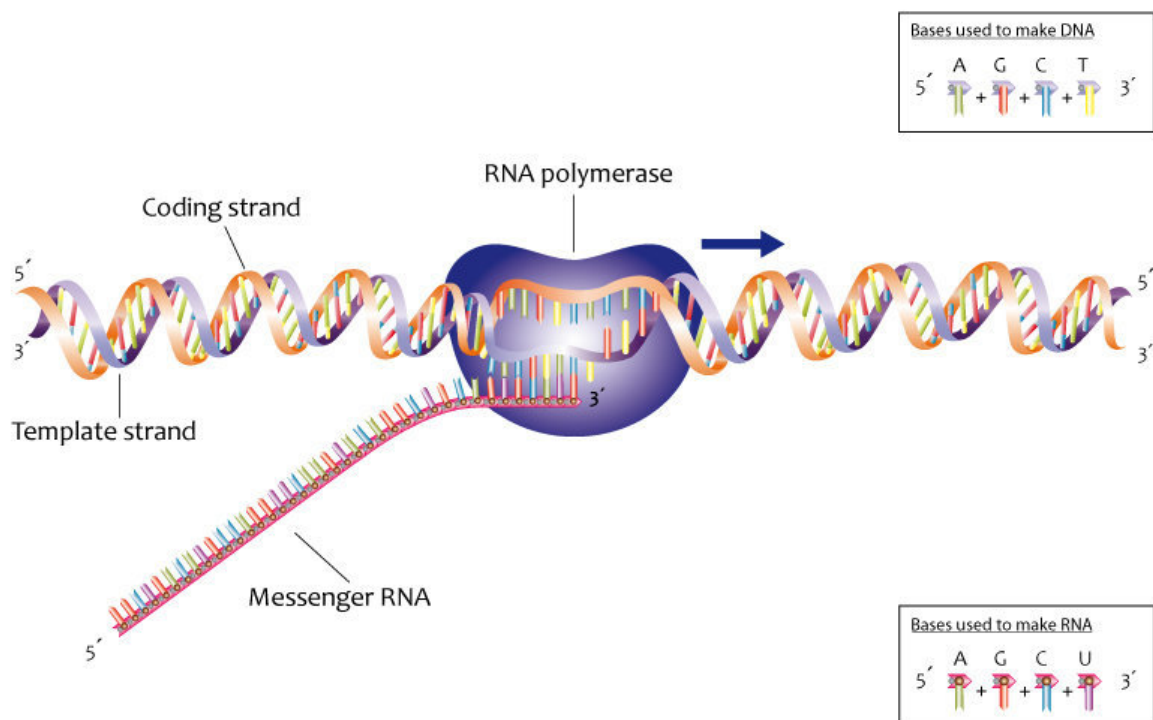
[file:///Users/carlos/Desktop/Bacterial%20transcription%20-%20Wikipedia.webarchive]

实际上,许多原核基因以操纵子的形式出现,一个操纵子就是一系列编码相同蛋白或基因产物的基因,同时被一个启动子所调控.原核RNA聚合酶由4个亚基构成,当第5个亚基结合后(σ 因子),该聚合酶复合体能够识别DNA上特殊的序列,称之为启动子. σ 因子结合到启动子是initiation的第一步.一旦 σ 因子从聚合酶上释放,便开始elongation.聚合酶沿着双链DNA开始合成新的mRNA,遇到终止位点合成终止.

转录过程是由RNA聚合酶所带来的,同时被序列特异性的DNA结合蛋白所控制,称为转录因子.转录因子根据细胞需要识别特殊的DNA序列,促进或者抑制其他的转录.

RNA polymerase

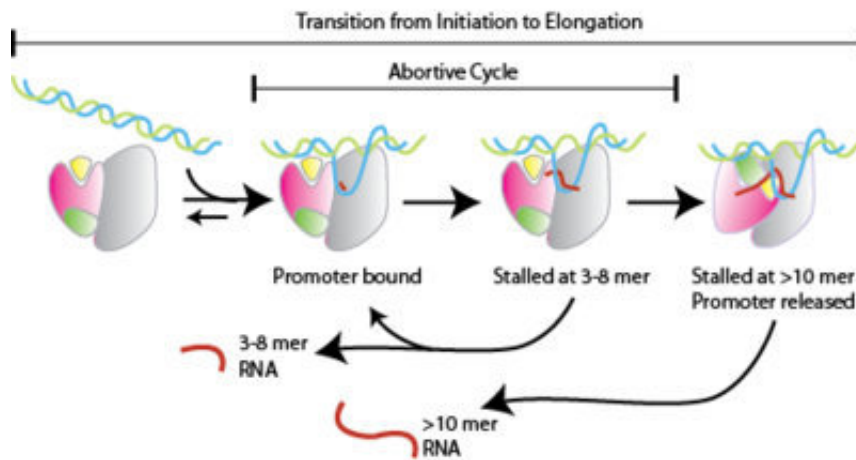
RNA聚合酶由core和holoenzyme(全酶)结构组成.core酶包含RNA聚合酶的催化性质,由 β β' α 2 ω 亚基构成.该序列在所有细菌中都保守存在.holoenzyme由特殊的组成构成, σ 因子. σ 因子有助于识别启动子,正确定位RNA聚合酶,同时在转录起点打开双链.



Initiation

转录的起始需要启动子区域,该特异性的核酸序列告诉RNA聚合酶上的 σ 因子去哪里结合DNA.启动子常位于距离基因15到19bp位置,常位于所调控基因的上游.RNA聚合酶由4个亚基构成,包含两个alpha,一个beta,一个beta prime(α , α , β , β').第5个亚基, sigma, 仅在initiation时期存在,同时在elongation前脱离. σ 因子在转录initiation时必须存在,当所有的 σ 因子存在时, RNA聚合酶表现为holoenzyme形式,当 σ 因子脱离, RNA聚合酶表现为core polymerase形式. σ 因子是被启动子区域位于-35到-10区域,同时转录起点为+1. -10位置序列为TATAAT, -35位置为TTGACA.

σ 因子结合到-35 启动子区域; 一旦 σ 因子结合, 聚合酶剩下的亚基开始结合, 高保守的adenine-thymine(腺嘌呤胸腺嘧啶)结合到-10位置促进DNA解双链; 转录开始于产生接近10bp的核酸序列, 该短序列没有功能, 产生后然后释放. 该序列有助于稳定RNA聚合酶; 接着, σ 因子和RNA聚合酶分离, 开始 elongation.



启动子区域的序列碱基对具有重要意义, 启动子序列和consensus序列越相似, RNA聚合酶结合的越紧密, 紧密结合有助于转录延长阶段的稳定, 发挥更有效功能. 此外, 在指定的细菌细胞内, 所提供的RNA聚合酶和 σ 因子的量有限. 因此, σ 因子结合到启动子区域受限于此, 所有启动子区域均包含non-consensus序列, 有助于 σ 因子在整个基因组的分布.

Elongation

RNA-DNA复合体的移动是RNA聚合酶的催化机制所必须的. 此外, RNA聚合酶通过连接RNA和DNA链增加了该过程的整体稳定性. RNA聚合酶以每秒约40bp的速度沿着DNA移动, 该快速移动使得在RNA聚合酶前的DNA双链解开, 一旦RNA聚合酶完成转录, 后续DNA双链结合.

启动子序列决定了对应基因的转录频率

Termination

内在终止(Rho-independent termination): 特殊的DNA核酸序列指导RNA聚合酶停止. 该序列一般为回文序列, 通过将链形成环拖拽RNA聚合酶(发卡环), 使得RNA聚合酶从模版链脱离, 转录停止

Rho-dependent termination: rho因子是一个辅助到RNA链上的终止蛋白, 同时在elongation过程中跟随聚合酶. 一旦聚合酶接近转录基因的末端, 聚合酶遇到一系列G核苷酸, 导致转录stall. 该stall使得rho因子捕获到RNA聚合酶. rho蛋白将RNA聚合酶从DNA模版上拉下, 释放新合成的mRNA, 转录终止.

转录机制

RNA聚合酶: 包含核心酶和 σ 因子

σ 因子: 在体外转录时, 如果没有 σ 因子, 则转录不具有专一性, RNA聚合酶转录所有的DNA双链, 而在加入 σ 因子后, 转录开始变得具有特异性.

启动子: 当我们用带有缺口的DNA转录时, RNA聚合酶可以正常转录, 但是明显不具有特异性, 然而在完整的T4DNA中是不具有缺口的, RNA聚合酶很难遇到这种天然的缺口, 因此其转录活性必然大大下降, 如果 σ 因子存在, 那么RNA聚合酶就可以正常识别起始位点并进行结合转录, 聚合酶结合的位点成为启动子, 因此 σ 因子的作用是指导聚合酶在特异的启动子处开始转录DNA.

RNA聚合酶与启动子的结合：

我们发现，RNA聚合酶全酶可以与DNA进行紧密结合，而核心酶只能松散的结合，这取决于是否含有 σ 因子，全酶结合半衰期为30-60h，而核心酶结合的半衰期只有1min，进一步实验发现，结合紧密的位点有8个，这与启动子数目相近，而结合松散的位置有1300个，并出现在任何位置，因此我们可以得出结论：核心聚合酶不可以特异性起始转录DNA，因为其实转录必须有启动子参与。继续研究发现，当我们提高温度时，可以增强结合的紧密程度，25°的解离速率明显高于37°C，因为温度高可以促进DNA解链，因此我们可以总结以下假说：RNA聚合酶先在DNA上进行松散的结合，直到发现启动子，全酶和启动子的松散结合复合体称为**封闭起始复合物**，然后，全酶使得一部分DNA解旋，这时的复合物称为**开放起始复合物**。

启动子结构：细菌需要怎样的DNA结构才能使RNA聚合酶与之结合呢？David Pribnow对E.coli和噬菌体多个启动子加以比较后，发现一个共有区域，其中心位于转录起始位点上方大概10bp处，长度大概6-7bp，我们现在称之为**-10序列**或**-10框**；其后，Mark Ptashne及其同事还发现一个短序列，其中心位于转录起点位置上游35bp，分析数千个启动子后发现，每个框都存在共有序列

```
-----TTGACa-----TAtAaT-----  
-----AACTGt-----ATaTtA-----
```

大写字母表示这个碱基在这个位置出现的频率比较高，这种概率性使得很难有与共同序列完全一样的-10序列和-35序列，然而，只要出现完全匹配的，其转录一定非常活跃，实际上，越不匹配的序列，启动子活性越低。除此之外，启动子元件之间的距离也十分重要，-10与-35远离或者靠近都会使得转录活性下降。

除了-10和-35序列（我们称之为核心启动子元件）有些极具活性的启动子还存在一个额外的元件，称为**UP元件**，比如我们E.coli细胞有7个编码rRNA的基因，当快速生长需要大量rRNA时，这7个基因自身可以引导转录的大量发生，我们认为UP元件是真正的启动子，因为RNA聚合酶可以识别UP元件，并且仅在UP元件存在的情况下，转录活性就可以被提升30倍。那么什么时候才能决定是否增强转录呢？实际上，该启动子还涉及-60~-150之间的三个Fis位点，他们是转录激活蛋白Fis的结合位点，他们不与RNA聚合酶结合，因此不叫启动子，而是一类被称为**增强子**的转录激活DNA元件。

E.coli RNA基因启动子还被一对小分子所调控，既起始NTP（iNTP）和预警素鸟苷-5'-二磷酸-3'-二磷酸（ppGpp），当大量iNTP存在时说明核苷酸的浓度很高。

转录起始：

σ 因子促进转录的开始，而不是延伸

σ 因子在促进转录完成后的某个阶段，显然与核心酶发生了解离，此后由核心酶继续完成转录的延伸， σ 因子则可以被不同的核心酶重新利用，。

σ 因子的释放并非是离开启动子的时候，而是在延伸过程中随机的方式释放。

在与启动子结合后，RNA聚合酶会引起临近转录起始位点的DNA解链，解链长度为10~17bp。

我们已经知道，RNA聚合酶可以识别启动子并与其紧密结合，但是要想进入延伸阶段，就必须破坏这种紧密结合，那么如何破坏这种紧密得结合呢？

经过复杂的实验验证，证明了RNA聚合酶脱离启动子的机质：ENApol通过蜷缩DNA，将下游DNA拉进自身内部，由于蜷缩的DNA具有的能量，使得聚合酶破坏了其与启动子的结合，使得下游转录开始。