

# TATA 结合蛋白

July 2012 Molecule of the Month by David Goodsell

译者：王权威（农科院特产所）

张昊（农科院生物所）

**关键词：** 转录调控，转录因子，DNA 弯曲，序列特异性的 DNA 结合活性

## 引言

RNA 聚合酶可以使双链 DNA 解旋并将遗传信息转录到一条 RNA 链上，但是我们并不知道它在基因上确切的结合位点。我们的细胞含有 30000 个编码基因，有数十亿计的核苷酸。对于这些编码基因来说，细胞必须能够在准确的时间和位置上开始转录。

## 转录起始

靠近编码基因的位置有一段特殊的启动子序列，它决定转录起始的准确位置和方向。不同的生物体中启动子的序列和位置是不同的。在细菌中，典型的启动子包含两个与 RNA 聚合酶 $\sigma$ 亚基相互作用的区域。 $\sigma$ 亚基与这些 DNA 序列结合后，促进转录起始，然后脱离聚合酶继续进行基因转录。

## TATA 盒

我们的蛋白质编码基因中有一段特有的被称为 TATA 盒的核苷酸序列，位于转录起始位点前。这个典型的序列是一段类似 T-A-T-A-a/t-A-a/t 一样的序列，其中 a/t 所在的位置可以是一个 A 或 T。令人惊奇的是，该序列的许多突变体也能发挥作用，研究转录的热点之一就是为什么一些突变体能够发挥作用而其他突变体却不能。TATA 结合蛋白（简称 TBP）能够识别结合该 TATA 序列，并构建了转录起始位点。TATA 结合蛋白的第一个结构被测定出来后，研究人员发现，TATA 结合蛋白结合到 DNA 上时并不稳定。相反，TATA 结合蛋白结合到 TATA 序列并使 TATA 序列极度弯曲，在 PDB 中的 1ytb、1tgh 和 1cdw 可以看到。

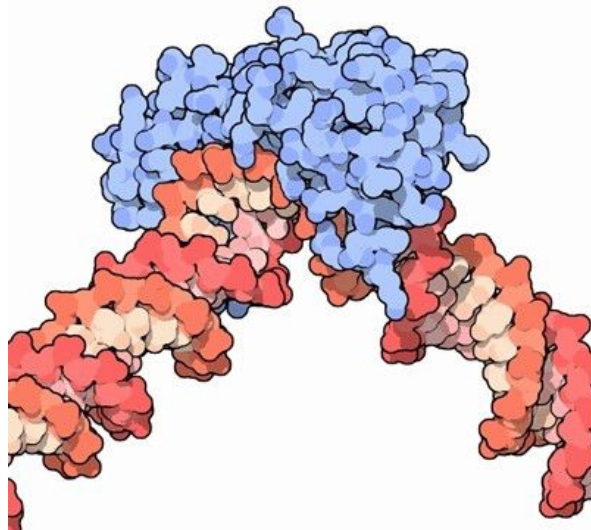


图 1 TATA-结合蛋白

## 其他辅助因子

TFIID 是一类大的转录因子，用于启动转录，TATA 结合蛋白是其中的一部分。它结合到启动子上后，又结合其他的转录因子。

如图 2 所示，然后 TFIIB 也结合到启动子上。接着还有一系列其他的转录因子结合，构建一个大的蛋白复合体，这个蛋白复合体决定是否启动转录。这些蛋白复合体可以包括转录增强因子，如图 3 所示的 TFIIA，它能够促进转录的起始。

另有他抑制转录起始的因子，如图 4 所示的转录调节因子 NC2（负辅因子 2）。

所有图片中，蓝色为 TATA 结合蛋白，红色为一小段 DNA 片段，绿色为转录因子。

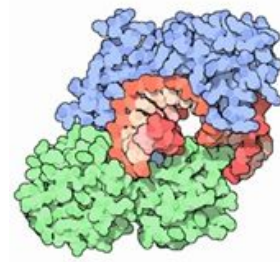


图 2

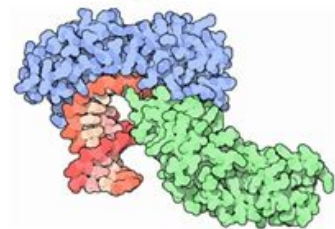


图 3

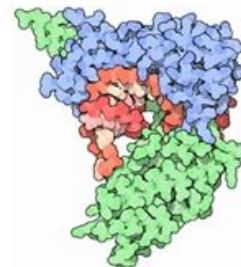


图 4

## 结构的探索

如图 5 所示，TATA 结合蛋白通过两种类型的相互作用去识别结合 TATA 序列。第一种，我们可以在它的顶端看到，有一串与 DNA 的磷酸基团（明亮的黄色和红色）相互作用的赖氨酸和精氨酸（深蓝色），这种作用可以使蛋白质和 DNA 结合在一起。第二种，蛋白质使用特定位置的氨基酸与 DNA 碱基相互作用。如图 6 所示，四个苯丙氨酸插入 DNA 小沟形成卷曲从而折叠 DNA。在正中央的位置的两个对称的天冬酰胺氨基酸形成了氢键。这些具有不寻常灵活性的 DNA 序列结合体和特定的氢键使 TATA 结合蛋白识别正确的序列。当你看到 TATA 蛋白的这些结构时，你可能会注意到，尽管它是由一条多肽链组成的，但是却形成了两个相互对称的部分。图 6 中，我们可以清晰的看到，两对苯丙氨酸和精氨酸是相互对称的。一般认为是一种原始基因复制的两个副本相结合产生了这种蛋白质。

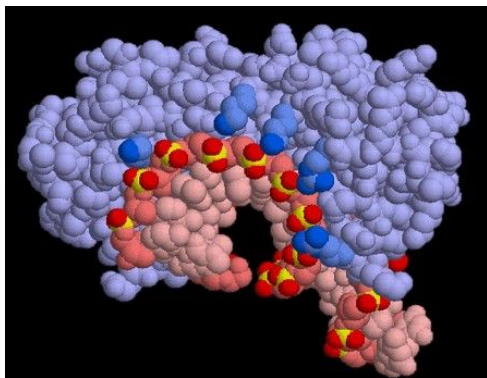


图 5

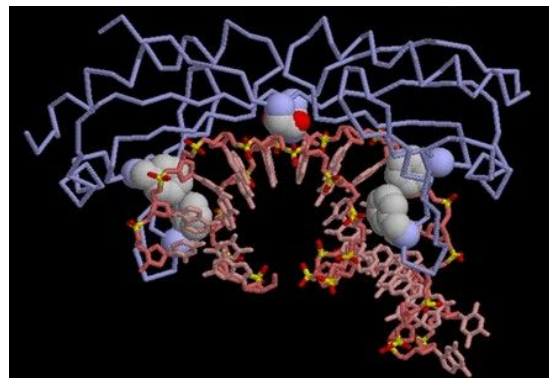


图 6

想从基因组学角度了解更多关于 TATA 结合蛋白的信息，请访问 [Protein of the Month at the European Bioinformatics Institute](#)。

这些图片是由 RasMol 制作的。你可以通过点击这里的加入代码并选择视图结构中的一个选项制作类似的图片。上面显示的苯丙氨酸数字是 99, 116, 190, 和 207, 天冬酰胺数字是 69 和 159。

## 拓展阅读

R. G. Roeder (1996) The role of general initiation factors in transcription by RNA polymerase II. *Trends in Biochemical Sciences* 21, 327-335.

Z. S. Juo, T. K. Chiu, P. M. Leiberman, I. Baikalov, A. J. Berk and R. E. Dickerson (1996) How proteins recognize the TATA box. *Journal of Molecular Biology* 261, 239- 254.