

酿酒酵母 GAL 基因的表达调控

刘巍峰 高东

(山东大学生命科学院微生物系 济南 250100)

酵母菌一直是人们进行遗传分析的理想材料,对许多酵母基因的表达调节途径人们也已有了较深入的了解。酵母中研究最多的基因表达调节途径是 GAL 基因表达调节过程,它不仅为比较研究原核和真核生物的协同表达调节机制提供了可能性,而且由于酵母 GAL 基因表达可以被生长条件所控制,所以 GAL 基因区域对外源基因在酵母中的克隆表达也具有潜在的应用价值^[1,2]。现在就酵母 GAL 基因的表达调节做一概述。

1 GAL 基因组成及转录调节模型

半乳糖是通过转化为 6-磷酸葡萄糖后进入糖解途径而被酵母利用的。其中至少有五种基因产物参与从半乳糖的运输到 6-磷酸葡萄糖的转化,这五种酶分别由 GAL2(透性酶)、GAL1(激酶)、GAL7(转移酶)、GAL10(差向异构酶)和 GAL5(变位酶)基因编码^[3]。体外翻译及 Northern blot 分析发现 GAL1、GAL10 和 GAL7 基因是从各自的启动子分别转录的^[4]。除 GAL5 外, GAL 基因的转录是被紧密调控的, GAL1、GAL7、GAL10 和 GAL2 基因在没有半乳糖的条件下不表达,而在半乳糖存在条件下表达水平提高约 1000 倍;诱导的 GAL1、GAL7 及 GAL10 基因编码产物占细胞总蛋白含量的 0.5-1.5%^[5];而其转录产物占总聚腺苷酸化 RNA 的 0.25~1%。

早期的遗传工作及后来更直接的生化分析发现主要有两种基因参与了 GAL 基因系统的这种表达调节,并由此建立了 GAL 基因表达调节基本模型。在此模型中位于染色体 XVI 上的 GAL4 基因编码一种能结合上述五种基因的上游序列位点并激活这些基因转录的 GAL4 蛋白;而位于染色体 XIII 上的 GAL80 基因编码的 GAL80 蛋白则能和 GAL4 直接结合,从而抑制 GAL4 的转录激活功能。至于半乳糖产生诱导信号的作用机制还不清楚,但其发挥作用,至少在产生迅速诱导方面需要 GAL3 基因产物^[6]。

2 GAL 基因参与转录调控的顺式序列元件

如模型中所述, GAL4 蛋白参与 GAL 基因的转录激活。它是通过结合 GAL 基因的上游 DNA 序列而发挥作用的,此种目的序列在酵母中通称为上游激活系列(upstream activating sequence, UAS)。研究发现,如果以位于 GAL1 和 GAL10 基因上游的一段 365bp DNA 片段取代 CYC1 或 HIS3 基因的上游区域,发现 CYC1 和 HIS3 基因的表达依赖于 GAL4 蛋白。从而首先定位了 UAS_{GAL} ^[7]。 GAL4 识别位点可进一步定位于与 GAL1 和 GAL10 转录起始位点等距离的一段富含 GC 的 108bp 区域^[8]。对此 DNA 片段精细的缺失突变分析发现了一含四个部分二分对称短序列的小区域,其中的二个短序列是 GAL1 和 GAL10 表达所需要的^[9]。 GAL7 基因的表达也需要两个部分二分对称的短序列^[10],这六个序列之间不仅具有同源性,而且与其它受半乳糖调控基因的上游序列具有同源性。通过对 20 个此种序列的比较得出了 GAL4 蛋白结合位点一致序列如下: CGGA(GC)GACATCAGGCAGGC。它们在各调控基因中与转录起始位点的距离从 100bp(GAL80)至 385bp(GAL1)不等^[11],此种序列介导 GAL4 转录激活功能与周围序列无关,当与之相似的 17bp 或 21bp 人工合成寡核苷酸置于某些其它基因的上游序列中时都可以引起依赖于 GAL4 的高水平基因表达^[12]。 GAL4 结合位点中最保守的碱基是外端的 CGG,中间的 G 在 20 个结合位点中保持一致。 GAL4 蛋白结合位点的几个特征使之与较高等真核生物的增强子有可比之处。第一,它是能增加任何基因转录的模式元件,并仅在一定条件下(如存在 GAL4)具有活性;第二,它可以两种方向发挥功能;第三, $GAL4/UAS_{GAL}$ 系统能在哺乳动物细胞中激活报道基因的转录^[13],并且 UAS_G 在 GAL4 存在的条

1997-01-22收稿

件下也可以在基因启动子的下游以增强子的形式发挥作用^[14]。第四, UAS_G可以在距转录起始点不同距离处发生作用, 但此 UAS 的功能伸缩性仅为几百个碱基对, 而许多哺乳动物基因增强子可以在几千个碱基对之外仍具有功能, 这可能也反映了 GAL4 蛋白功能的一种局限性。但在真核基因表达中对于研究结合于增强子元件的蛋白因子如何在相当大的距离处激活转录, GAL4 蛋白无疑是一个很好的模型。

3 GAL4 蛋白的结构及功能

GAL4 基因已被克隆并测定了其序列, 它编码一个含 881 个氨基酸的较大蛋白质, 在酵母中低水平存在^[15]。研究发现 GAL4 蛋白至少有以下六种功能: (i) DNA 结合; (ii) 转录激活; (iii) 专一性核定位功能; (iv) 形成多聚体; (v) 与 GAL80 相互作用 (vi) 可能直接参与代谢物阻遏。在确定该蛋白的功能域方面已取得了很大进展, 发现 GAL4 蛋白的 DNA 结合和转录激活功能并不相互依存^[16]。GAL4 蛋白的 DNA 结合区域包括 N-末端的 74 个氨基酸, 其中的两对半胱氨酸形成锌指结构, 锌指结构中有较高比例的能和 DNA 磷酸骨架结合的赖氨酸和精氨酸^[17]。GAL4 蛋白中紧邻锌指结构 C-端的 20 个氨基酸也参与 DNA 的结合并保证结合的专一性^[18]。

GAL4 蛋白中负责转录激活的部分位于蛋白 C-末端 90% 区域, 该区域可以独立于 DNA 结合区域发挥作用, 因为如果 GAL4 蛋白的 DNA 结合域被细菌抑制子 (由 *E. coli* *lexA* 编码) 的 DNA 结合域所取代, 产生的杂合蛋白仍能激活启动子中含 *lexA* 结合位点的酵母基因^[19], 由此可见 GAL4 蛋白 DNA 结合域的唯一功能就是把该蛋白的转录激活域在启动子上适当定位^[20]。通过研究不同的 GAL4 缺失突变体已经定位了两个具有转录激活功能的 GAL4 区域, 即邻近 DNA 结合结构域的 148~196 区和羧基端的 768~881 区。激活结构域的负净电荷和转录激活效率有很强的相关性, 这与许多其它转录激活因子相似, 因此可能暗示了酵母乃至整个真核生物调节蛋白激活转录的一种共有机制。GAL4 蛋白可能是通过它的转录激活结构域与较直接参与转录的其它蛋白因子间的相互作用 (如 TATA 结合蛋白或 RNA 聚合酶 II) 或通过改变基因上游的染色质结构而促进基因表达的^[21,22]。

研究发现, GAL4 蛋白激活转录的功能可以被另一种基因产物——GAL80 所抑制。体内和体外实验发现

GAL80 蛋白是直接抑制转译后 GAL4 蛋白的活性, GAL4 蛋白中被 GAL80 蛋白识别并与之相互作用的区域位于其 C-末端的转录激活区域^[23]。另外最近发现 GAL80 在被调控基因上游对染色质结构的改变具有重要的作用。由于 GAL4 蛋白和 GAL80 蛋白直接相互作用, 因此这两种蛋白细胞内水平的精细平衡对 GAL 基因表达调节是必需的, 任一蛋白的过多表达都将打破 GAL 基因的诱导表达模式。研究发现 GAL80 基因的转录受 GAL4 蛋白的调节^[24], 诱导效应物这种既降低 GAL80 蛋白活性同时又增加 GAL80 蛋白表达水平的机制可能是酵母 GAL 基因表达中保持各作用因子稳态的一种途径。

4 GAL4 基因表达的代谢物阻遏

对酵母代谢物阻遏机制的了解要比大肠杆菌少得多, 而且葡萄糖引起的抑制程度强烈地依赖于目的基因和使用的菌株。阻遏机制的实现至少需要三种元素: 由一种或多种活性蛋白催化葡萄糖产生阻遏信号; 反式作用调节蛋白级联传导此种信号; 被调控基因启动子中的顺式作用元件。到目前为止, 关于葡萄糖产生的阻遏信号的性质还不清楚, 但已经分离到许多参与代谢物阻遏过程的中间元素突变体, 这些中间元素就是反式调控基因的产物^[25]。葡萄糖可能在干扰诱导信号的形成, 促进负调节蛋白对启动子的结合或干扰 GAL4 蛋白的转录激活功能等几个水平上来阻止转录。研究发现在 *gal80* 菌株中由 2% 葡萄糖引起的 GAL1 抑制达 15~170 倍, 而野生型菌株在葡萄糖和半乳糖混合培养物中 GAL1 基因表达抑制 80~600 倍^[26], 因此很明显葡萄糖能干扰半乳糖的诱导。但最近结果显示调节 GAL4 的合成是葡萄糖介导阻遏的主要方式, 研究发现在葡萄糖存在情况下 GAL4 基因启动子序列能结合负调控蛋白 MIG1, 从而引起 GAL4 基因转录的抑制 (4~7) 倍, 就是这一弱表达调节基因的这种轻度调节而对环境变化产生极其扩大的反应^[27]。总之, GAL 基因表达的代谢物阻遏调控可能远非如此简单, 因为还有许多反式调控基因产物, 如 HXK2, HEX2, GAL82, GAL83, CYC8 等参与此调控, 但这些基因间的相互作用模式及在整个途径中的位置还不清楚。

5 与几种其它调节途径的比较

许多酵母基因调节途径和 GAL 基因调节机制相似, 其特征一般包括 (1) 由一种转录激活因子 (如 GAL4 蛋白) 介导的基因表达正控制。 (2) 由一种能使转录激

活蛋白失去活性而抑制转录的蛋白在转录后或翻译后水平负控制基因表达。其中了解得比较清楚的例子包括酵母菌的 DAL1-4, ADH2 及粗糙脉胞菌的 QA-2, -3, -4 和 PRT 等基因调节途径。这些调节途径中的转录激活因子(如 DAL1, ADR1, QAF1 及 CYS3 等)的结构特征以及与抑制因子间的相互作用方式和 GAL4 及 GAL80 相似。酵母菌的 CYC1 和 CYC7 基因的表达调节虽然受转录因子 HAP1 / HAP2 / HAP3 的正调控,但 HAP1 / HAP2 / HAP3 的活性受效应因子血红蛋白的直接调节。另外两个较特殊的情况是酵母 HIS4 基因和 PHO5 基因。HIS4 基因的抑制因子 GCD1 不是和激活因子 GCN4 直接作用,而是负调控 GCN4 的翻译^[28];在酵母和丝状真菌的 PHO 基因表达调节中,存在次级激活蛋白 PHO81。PHO 基因表达的效应因子(无机磷)阻遏了次级激活蛋白,使抑制因子 PHO80 + PHO85 失活初级激活蛋白 PHO4 + PHO2 从而抑制转录,当效应因子耗尽时次级激活蛋白得以释放来阻遏抑制因子 PHO80 + PHO85 对初级激活蛋白的作用,最终释放初级激活蛋白来激活转录,此种级式调节更可能是真菌基因表达调节途径的一致特征。研究还发现哺乳动物细胞中某些受类固醇激素诱导的基因通过类固醇激素与胞内受体蛋白的结合来激活自身的转录。该家族受体蛋白在结构功能特征及诱导物激活基因表达的方式上与 GAL4 蛋白颇为相似,不同之处是受体蛋白中有一个相当于 GAL80 蛋白功能的结构域^[29]。另外还发现哺乳动物细胞中的转录激活因子可以与经过改造的 GAL4 蛋白协同激活哺乳动物细胞基因的转录^[30]。这些都暗示了整个真核生物基因表达调节机制方面的某些相似性。

总之,上述研究表明 GAL 基因表达调节极其复杂,其表达调控的实现依赖于众多正负控制因子间功能的平衡。由于 GAL 基因表达调节途径可能是酵母菌中了解得最为深入的一种调节途径,并且其许多特征也是许多其它酵母基因表达调节途径所共有的,因此对此调节机制的详细了解将有助于我们对酵母基因表达乃至整个真核生物基因表达的调控有更广泛的认识。

参 考 文 献

- [1] Broach JR. *Methods Enzymol*, 1983, 101C:307~324.
- [2] Romanos MA, Scorer CA, Clare JJ. *Yeast*, 1992, 8: 423~488.
- [3] Douglas HC. *J Bacteriol*, 1954, 68:662~670.
- [4] John St, Davis RW. *J Mol Biol*, 1981, 152:285~315.
- [5] Fukasawa TK, Obonai TS, Nogi Y. *J Biol Chem*, 1980, 255:2705~2707.
- [6] Bhat PJ, Hopper JE. *Mol Cell Biol*, 1992, 11:5454~5461.
- [7] Guarente L, Yocum RR, Gifford P. *Proc Natl Acad Sci USA*, 1982, 79:7410~7416.
- [8] Yocum RR, Hanley S, West M. *Mol Cell Biol*, 1984, 4:1985~1988.
- [9] West RW, Yocum RR, Ptashne M. *Mol Cell Biol*, 1984, 4:2467~2478.
- [10] Tajima M, Nogi Y, Fukasawa T. *Mol Cell Biol*, 1986, 6:246~256.
- [11] Johnston M. *Microbiol Rev*, 1987, 57:458~476.
- [12] Lorch Y, Kornberg RD. *J Mol Biol*, 1985, 186:821~824.
- [13] Kakidani H, Ptashne M. *Cell*, 1988, 52:161~167.
- [14] Webster N, Jim JR, Green S. *Cell*, 1988, 52:169~178.
- [15] Laughon A, Gesteland RF. *Mol Cell Biol*, 1985, 4: 260~267.
- [16] Keegan L, Gill G, Ptashne M. *Science*, 1986, 231: 699~704.
- [17] Johnston M, Dover J. *Proc Natl Acad Sci USA*, 1987, 84:2401~2405.
- [18] Saimeron JM, Johnston SA. *Nucleic Acids Res*, 1986, 14:7767~7781.
- [19] Brebt R, Ptashne M. *Cell*, 1985, 43:729~736.
- [20] Struhl K. *Cell*, 1987, 49:295~297.
- [21] Lohr D, Lopez J. *J Biol Chem*, 1995, 270:27671~27678.
- [22] Grace G, Ptashne M. *Nature*, 1988, 334:721~724.
- [23] Lohr D, Venkov D, Zlatanova J. *FASEB J*, 1995, 9: 777~787.
- [24] Shimada H, Fukasawa T. *Gene*, 1985, 39:1~9.
- [25] Gancedo JM, Gancedo C. *FEMS Microbiol Rev*, 32: 179~187.
- [26] Nehlin JO, Carlberg M, Ronne H. *EMBO J*, 1991, 10:3373~3378.
- [27] Griggs DW, Johnston M. *Proc Natl Acad Sci USA*, 1991, 88:8597~8601.
- [28] Hinnebusch AG. *TIBS*, 1990, 15:148~162.
- [29] Godowski PJ, Fusconis S, Misefield R, et al. *Nature*, 1987, 325:365~368.
- [30] Lin YS, Carey M, Ptashne M, et al. *Nature*, 1990, 345:359~361.