



## 结瘤基因的表达调控

刘嵩涛 洪国藩

(中国科学院上海生物化学研究所 上海 200031)

**关键词** 根瘤菌, 结瘤基因, 作用机理

**分类号** Q939.11

根瘤菌和豆科植物的共生体系一方面因为它的固氮效应在农业上有重要意义, 另一方面因为它可以作为发育生物学的良好模型, 所以得到了广泛的研究。通常情况下, 一种根瘤菌只能在一种或几种豆科植物上结瘤; 而一种豆科植物也只能接纳一种或几种根瘤菌。在根瘤菌方面, 结瘤基因及其产物蛋白的活动是决定这种宿主特异性的主要因素。对于结瘤基因表达调控的深入认识将有助于实现人们梦寐以求的扩大根瘤菌宿主范围的愿望。

通常所说的根瘤菌实际上包括四个属: (快生型) 根瘤菌属 (*Rhizobium*)、慢生型根瘤菌属 (*Bradyrhizobium*)、固氮根瘤菌属 (*Azorhizobium*) 和中华根瘤菌属 (*Sinorhizobium*), 属革兰氏阴性菌。八十年代以来通过综合使用转座子突变、互补分析、DNA 序列分析、异源探针杂交、同源性比较等方法已经在各种根瘤菌中鉴定了大量的结瘤基因, 即 *nod*、*nol* 和 *noe* 基因<sup>[1,2]</sup>。结瘤基因在快生型根瘤菌中定位于一个称作 *sym* (共生) 的巨大质粒 (*R. leguminosarum*: 200~300kb; *R. meliloti*: 1, 200~1, 500kb) 上, 在慢生型根瘤菌中则位于细菌染色体上, 它们大都与结瘤因子的合成及外运有关。结瘤因子是三到六聚的 N-乙酰葡萄糖胺组成的寡糖, 其还原和还原端常受到多种修饰如长链脂肪酰基化、磺基化等。不同根瘤菌的结瘤因子各有一些独特的结构特征, 只能诱导相应的宿主植物作出回应, 因此被认为是根瘤菌与宿主特异识别的主要信号。负责合成寡糖骨架的 *nodA*, *B*, *C* 被称作通用结瘤基因, 它们存在于所有已知根瘤菌中; 负责添加修饰基团的其它基因被称为宿主特异基因, 它们只存在于某些或某种根瘤菌株中<sup>[2]</sup>。

但是, 绝大多数的结瘤基因在培养细菌中并不表

达。它们的启动受到了严格的调控。结瘤基因的表达调控目前主要在转录水平研究, 已鉴定了顺式元件 *nod box* 和 *nodD*, *syrM*, *nolR* 等调控基因, 并探讨了它们的作用机理; 其它遗传因子和生理因子对结瘤基因表达的影响也有一些报道。我们下面分别介绍。

### 1 *nodD*

*nodD* 在植物诱导分子存在时激活其它结瘤基因的表达, 它是目前了解最清楚的调控基因。

存在 *nodD* 基因以不同的拷贝数存在于各根瘤菌中。 *R. leguminosarum* bv. *viciae* 和 *R. leguminosarum* bv. *trifolii* 各有一个拷贝; 三种广谱根瘤菌 *R. fredii*, *Rhizobium* sp. NGR234 和 *Bradyrhizobium japonicum* 都有两个不连锁的拷贝; *R. meliloti* 和 *R. leguminosarum* bv. *phaseoli* 有三个拷贝; 而 *R. loti* 和 *R. tropici* 则分别有四和五个拷贝<sup>[1]</sup>。 *R. leguminosarum* bv. *viciae* 中, *NodD* 是一种两亲性的细胞膜蛋白, 可能插入细胞质侧的膜单层中。但在 *R. meliloti* 中, *NodD1* 和 *NodD3* 在生化分离时基本上都存在于可溶部分<sup>[3]</sup>。

**LysR 家族成员** 根据同源性分析, *NodD* 被列入原核 LysR 转录调控蛋白家族<sup>[4]</sup>。

**DNA 结合活性** 早在 1987 年就利用凝胶阻滞技术观察到了 *NodD* 在体外能够结合 DNA<sup>[5]</sup>。在 *R. meliloti* 和 *R. leguminosarum* bv. *viciae* 中, 已经用 DNaseI, o-Phen-Cu 等试剂进行足迹保护实验确认了 *NodD* 在结瘤基因启动子上的保护区<sup>[6,7,8]</sup>, 这些保护区都包括了 *nod box*。

本实验室工作得到攀登计划的资助

1997-04-07 收稿

nod box 最初是通过顺序比较在结瘤基因上游发现的保守顺序。在快生型根瘤菌中, nod box 由三个高度保守的部分: 7bp, 5bp, 25bp 组成<sup>[9]</sup>; 慢生型大豆根瘤菌的 nod box 则被分为四个 9bp 的直接重复块<sup>[10]</sup>; 但所有的已知 nod box 都存在两个具有 ATC-N<sub>9</sub>-GAT 顺序的反转重复结构<sup>[11]</sup>。早期的实验指出 NodD 可能以二体或四体结合 nod box。上面提到的 nod box 保守结构中后面两种都支持四体结合, LysR 家族中的其它两个成员 CysB 和 NahR 也是以四体形式结合 DNA 的; 但是 Fisher 和 Long 通过对 nod box 的缺失和插入突变分析坚持认为 NodD 是以二体形式结合到 nod box DNA 上同一侧面的两个位点<sup>[12]</sup>。我们注意到所有这些分析的依据都是顺式元件的结构而不是蛋白-DNA 复合物本身, 所以他们的结论尚需进一步验证。

Fisher 和 Long 发现 *R. meliloti* 的 NodD3 与 DNA 结合时诱导 DNA 发生弯折<sup>[12]</sup>。我们用 *R. leguminosarum* bv. *viciae* 的 NodD 观察到类似现象。已经知道转录调控蛋白与 DNA 结合时常引起 DNA 分子的结构变化如弯折、成环、解旋、扭曲、缠绕等, 这些变化一般是调控蛋白发挥功能所必需的。NodD 诱导弯折的能力可能也与其调控功能相关。

Long 实验室发现 Tn5 插入突变 *R. meliloti* 染色体上的一个 groEL 基因座导致 NodD 结合 DNA 及激活转录的活性均丧失<sup>[13]</sup>; 无论用免疫亲和纯化还是 DNA 亲和纯化的 NodD 制备物中都曾发现含有 GroEL 同源蛋白, 说明两者至少在细胞生长的某个时期存在生化上的联系。但 GroEL 是直接影响 NodD 蛋白的折叠还是影响 NodD 与其它蛋白形成复合物时的组装, 作为细胞基本分子它又怎么特异影响与结瘤有关的 NodD 蛋白的性质, 其细节还不得而知。

**与诱导分子的结合** 如前所述, 结瘤结构基因的表达主要依赖于调控基因 nodD 在植物诱导分子存在下的激活。已经发现这些诱导分子除少数是酚类化合物外, 绝大多数是黄酮或二氢黄酮<sup>[14]</sup>。它们是一些三环芳香化合物, 在植物中由苯丙酮 (phenylpropanoid) 代谢产生。不同植物或同一种植物在发育的不同时期能分泌不同的类黄酮诱导分子。在苜蓿和三叶草中, 最强的诱导物是黄酮, 如 luteolin (3', 4', 5, 7-四羟基黄酮) 和 DHF (7, 4'-二羟基黄酮); 在大豆中, 天然的诱导物包括异黄酮如木黄酮 (genistein) 和 daidzein, 但后者在苜

蓿系统中是诱导抑制物<sup>[15]</sup>。

植物诱导分子除了黄酮类物质外, 还有另一类分子: 甜菜碱。在 *R. meliloti* 中, NodD2 对葫芦巴碱 (trigonelline) 和水苏碱 (stachydrine) 都有反应。前者是一种取代的吡啶, 后者是一种取代的吡咯烷<sup>[16]</sup>。这类分子和类黄酮的结构不同, 物理化学性质不同, 合成的代谢途径不同, 却都能激活 NodD 蛋白, 这是很有趣的现象。

NodD 直接结合诱导物的生化证据还不充分。不过许多遗传实验都提示 NodD 可以充当类黄酮的特异受体, 并且已经把根瘤菌对特定化合物的响应与特定的 nodD 等位基因联系起来。一些早期的研究认为 NodD 的 C 端负责与类黄酮结合, 但现在认为 NodD 整体的三维结构都很重要。在体内, NodD 和类黄酮间的相互作用可能发生在细胞膜上, 因为二者均定位于此间<sup>[3]</sup>。

Gyocergypal 等提出 NodD 蛋白和哺乳动物甾体激素 / 甲状腺激素受体间有相似性。他们发现这类激素或其类似物如雌二醇 (estradiol), auronos, coumestrol 在特定 NodD 蛋白存在下能够诱导 nodC-lacZ 融合基因的表达; 而激素受体与 NodD 蛋白的顺序上有两个同源的模块 module 1 和 module 2<sup>[17]</sup>。这种比较对人们了解 NodD 的作用机理是有益的。

很多作者都认为有无诱导分子存在时 NodD 和靶 DNA 的结合能力是一样的, 但在 *R. meliloti* AK41 和 *A. caulinodans* 中, 有报道称诱导分子的存在使 NodD 与 nod box 的亲和力升高<sup>[7, 11]</sup>; 而我们实验室则观察到 *R. leguminosarum* bv. *viciae* 的 NodD 在其诱导物 Naringenin 的浓度上升到一定程度时, 不能与含 nod box 的 DNA 片段保持结合<sup>[18]</sup>, 其生理意义尚待研究。

**自身的表达调控** 早期曾认为 nodD 基因在所有根瘤菌中都是组成型表达的, 但现在发现它的表达调控在不同的菌种、不同的菌株乃至同一菌株的不同 nodD 拷贝间都会有差异, 这反映了结瘤调控的复杂性<sup>[1-3]</sup>。*R. leguminosarum* bv. *viciae* 和 bv. *trifolii* 的唯一 nodD 自身负反馈调控; *R. leguminosarum* bv. *phaseoli* 和 *B. japonicum* 的 nodD1 则在某些类黄酮存在下自身激活; *R. meliloti* 的 nodD3 与 syrM 构成正反馈回路; 该菌种的某些菌株中存在 nolR, 抑制 nodD1 和 nodD2 的表达。结合态的氮和根瘤菌的生长时期也影响 nodD 基因的表达 (见下)。

**作用机理** 遗传学证据表明激活结瘤基因的表达

达大多数情况下需要同时存在三个条件:有活性的NodD蛋白;植物诱导分子;完整的顺式元件(nod box)。NodD的激活已表明作用于转录一步,那么NodD究竟是怎样激活其它结瘤基因的转录的呢?它与RNA聚合酶的相互作用如何?它与诱导物间的作用又是怎样的?它的DNA结合活性与激活功能间有无必然联系?当存在自身负反馈时,负反馈与激活功能间是如何协调的?两种功能需要RNA聚合酶作出相应的变化或者需要不同的额外因子吗?所有这些问题对于理解结瘤基因的调控过程都是至关重要的,但遗憾的是我们还没有一个清楚的图景。不过有一点可以肯定:激活结瘤基因的表达,仅有NodD蛋白是不够的<sup>[19]</sup>。已经发现一些sym质粒以外的基因如groEL<sup>[13]</sup>和dctB<sup>[20]</sup>的突变会降低结瘤基因的表达,我们实验室正在研究其它反式因子和RNA聚合酶不同形式的 $\sigma$ 亚基在这个过程中的作用。

## 2 syrM

1989年Long实验室在*R. meliloti*中检测到一个与nodD具有同源性的基因座syrM<sup>[21]</sup>。这个基因也存在于sym质粒上,并与nodD3形成正激活回路。二者都以多拷贝存在时,会引起结瘤基因不依赖诱导物的表达,宿主范围也扩大到一般情况下的非宿主*Macropitium*。SyrM蛋白属于LysR家族的正调控蛋白,它在根瘤中的表达水平比根瘤菌自由生长时要高。除影响结瘤基因表达外,它还控制胞外多糖的产生。在*Rhizobium* sp. NGR234和*R. leguminosarum* bv. *phaseoli*中也已发现了syrM的存在。

## 3 nolR

NolR是一分子量为13349的DNA结合蛋白,抑制结瘤基因的表达,首先由Kondorosi实验室在*R. meliloti*中发现<sup>[7,22,23]</sup>。它包含一个与NodD, SyrM, NahR等LysR家族成员同源的螺旋-转折-螺旋结构,可能以二体结合到nodD1和nodD2的启动子区,对这两个调控基因及通用结瘤基因nodABC的表达进行负调,进而降低结瘤因子的产生量并改变其组成,但它不直接影响宿主特异的结瘤基因。NolR的负调可能为根瘤菌保持最佳结瘤状态所需要,因为nolR<sup>-</sup>的菌株比nolR<sup>+</sup>的结瘤效率要低。

nolR在自由生长的细菌和类菌体中均有较高表达,但被自身产物和结瘤基因诱导物luteolin下调。杂

交实验在*B. japonicum*中发现了nolR的一个同源基因座,突变该座引起结瘤基因诱导水平升高;在*R. leguminosarum* bv. *viciae*中杂交没有发现nolR同源序列。

## 4 影响结瘤基因表达的其它因子

有人认为是结瘤调控基因<sup>[24]</sup>的还有:慢生型根瘤菌中的nodV, nodW, nwsA, nwsB, nolA, nodZ, *R. fredii*中的nolC, nolBTUV, nolW, nolX, nolJ, 其中有些基因或基因群的研究是有意义的,如NodV-NodW可能组成原核生物中常见的二组份调控系统,将结瘤基因的调控链作了纵向的延伸。但作者有些混淆了调控基因和影响结瘤性状的基因间的区别,如nodZ现在认为与结瘤因子上的岩藻糖残基修饰有关,实际上应视作结构基因。

从蛋白方面来说,在*A. caulinodans*中,已经发现三种分子量小于NodD的蛋白可以与nod box DNA结合<sup>[11]</sup>。它们是否真的是结瘤基因表达的调控因子,还需要体内实验的支持。

在*R. meliloti*和*B. japonicum*中,高水平的氨降低结瘤基因的表达。已经发现这一控制分别是由NodD3和NodD1介导的。NodD3在此情况下受总氮调控系统(global nitrogen regulation system)中一个新成分的负调。突变该基因能导致逃脱总氮控制<sup>[25]</sup>。

Djordjevic等发现结瘤基因受类黄酮诱导而引起的表达,其水平与细菌的生长时期有关<sup>[26]</sup>;Schlaman则发现当根瘤菌侵入植物并分化成可以固氮的类菌体后,nodD的表达水平大幅下降,而其它结瘤基因则完全停止表达<sup>[27]</sup>。

## 参 考 文 献

- [1] Pueppke S G. Criti Rev Biotech, 1996, 16:1~51.
- [2] Denarie J. Annu Rev Biochem, 1996, 65:503~535.
- [3] Schlaman H R M. J Bacteriol, 1992, 174:5177~5182.
- [4] Henikoff S. Proc Natl Acad Sci USA, 1988, 85:6602~6606.
- [5] Hong G F. Nucleic Acids Res, 1987, 15:9677~9690.
- [6] Fisher R F, S R Long. J Bacteriol, 1989, 171:5492~5502.
- [7] Kondorosi E M. EMBO J, 1989, 8:1331~1341.
- [8] 吕幼仪,徐有成,洪国藩. 科学通报, 1993, 38:1998~2000.

- [9] Rostas K. Proc Natl Acad Sci USA, 1986, 83:1757~1761.
- [10] Wang S P, G Stacey. J Bacteriol, 1991, 173:3356~3365.
- [11] Goethals K M. Proc Natl Acad Sci USA, 1992, 89:1646~1650.
- [12] Fisher R F, S R Long. J Mol Biol, 1993, 233:336~348.
- [13] Ogawa J, S R Long. Genes Dev, 1995, 9:714~729.
- [14] Peters N K. Science, 1986, 233:977~980.
- [15] Long S R. Cell, 1989, 56:203~214.
- [16] Philips D A. Plant Physiol, 1992, 99:1526~1531.
- [17] Gyoergypal Z, A Kondorosi. Mol Gen Genet, 1991, 226:337~340.
- [18] 洪国藩, 曹慧敏. 生物工程学报, 1993, 9:113~116.
- [19] Yelton M M. J Bacteriol, 1987, 169:3094~3098.
- [20] Mavridou A. Microbiology, 1995, 141:103~111.
- [21] Mulligan J T, S R Long. Genetics, 1989, 122:7~18.
- [22] Kondorosi E. J Mol Biol, 1991, 222:885~896.
- [23] Cren M. Mol Microbiol, 1995, 15:733~747.
- [24] van Rhijn P, Vanderleyden J. Microbiol Rev, 1995, 59:124~142.
- [25] Dusha I. Mol Gen Genet, 1989, 219:89~96.
- [26] Djordjevic M A. EMBO J, 1987, 6:1173~1179.
- [27] Schlaman H R M. J Bacteriol, 1991, 173:4277~4287.