

巴西固氮螺菌 Yu62 draTG 基因启动子区域的核苷酸序列及其功能分析

马旅雁 李季伦

(北京农业大学农业生物技术国家重点实验室 北京 100094)

摘要 对巴西固氮螺菌 draTG 上游区域进行了全序列分析, 结果表明该区域除了编码部分 nifH 基因外(nifH 与 draTG 转录方向相反), 不编码任何其它已知的基因。但在该区域发现了一些可能的调控序列, 它们包括上游激活序列(UAS)、下游启动子组份(DPE)和富 A + T 区。这说明: nifH 与 draT 间的区域可能主要起调控功能而非编码功能; dra 操纵元(operon)的启动子很可能是 RpoN - 依赖型。用 pAF300 做载体, 构建了 draT::cam 转录融合质粒 pAT1, 并通过检测 Cm^r 以检测 draT 在大肠杆菌和巴西固氮螺菌中的表达, 结果表明 draT 在 LD 丰富培养基上, 好氧条件下, 只在巴西固氮螺菌中才表达。这说明, draT 的转录需要某种大肠杆菌中没有的因子, 同时也表明 draTG 上游区域有启动子功能。利用启动子探针质粒载体 pCB182, 构建了 draT::lacZ 转录融合质粒 pCT1。在大肠杆菌中测定肺炎克氏杆菌 NifA 对 draT::lacZ 的转录激活作用。结果表明 nifA 并不参与 draT 的转录调控。

关键词 巴西固氮螺菌, draTG 启动子, 核苷酸序列, 转录融合质粒

固氮螺菌(*Azospirillum*)是一类可与许多禾本科作物及牧草根际联合固氮的微生物^[1], 由于其具有分布广、数量多、适应性强等优点, 成为最受重视的联合固氮菌。对其固氮遗传学研究的主要对象是巴西固氮螺菌(*A. brasilense*), 该菌与深红红螺菌(*Rhodospirillum rubrum*)一样具有固氮酶翻译后的活性调控机制^[2]。该调控是由 DRAT (Dinitrogenase reductase ADP-ribosyltransferase)和 DRAG(Dinitrogenase reductase activating glycohydrolase)催化的可逆 ADP - 核糖修饰系统来完成^[3]。分别编码这两个酶的 draT 和 draG 基因的核苷酸序列在 *A. brasilense* Sp7、*A. lipoferum*、*R. rubrum* 和 *Rhodobacter capsulatus* 中都已报道^[4~7]。序列分析的结果表明 draTG 可能是以一个操纵元的方式转录, 但至今 draTG 的启动子类型及其转录调控尚不清楚。在上述的前三个菌中 draTG 的上游均与 nifH 相邻, 但转录方向相反。*A. brasilense* Yu62 菌株是 1984 年杨洁彬等^[8]从北京郊区玉米根际分离得到的。该菌的 draTG 同源片段已得到克隆, 其 draTG 基因及其下游约 1kb 区域已进行了全序列分析^[9]。本文报道的是 draTG 上游区域的核苷酸序列及 draTG 转录的分析结果。

863 计划资助项目。

本文部分内容曾在 1996 年“第二届中韩生物技术研讨会”上报告。

本文于 1997 年 2 月 28 日收到。

1 材料和方法

1.1 菌株与质粒(表 1)

表 1 菌株与质粒

Table 1 Bacterial strains and plasmids

| Strains and plasmids | Phenotype and/or genotype | Reference or source |
|-----------------------|---|---------------------|
| Strain | | |
| <i>E. coli</i> | | |
| CB454 | F ⁻ Δ <i>lacZ</i> ⁻ <i>lacY</i> ⁺ <i>galK</i> <i>rspL</i> <i>thi</i> <i>recA56</i> | |
| JM101 | <i>supE</i> <i>thi</i> Δ(<i>lac</i> - <i>proAB</i>) F' ⁺ [<i>traD36</i> <i>proAB</i> ⁺ <i>lac</i> ^g <i>lacZAM15</i>] | [10] |
| <i>A. brasiliense</i> | | |
| Yu62 | Ap ^r Nx ^r , wild type | [8] |
| Yu62(pAT1) | Ap ^r Nx ^r Tc ^r (harboring pAT1) | This report |
| Yu62(pAF300) | Ap ^r Nx ^r Tc ^r (harboring pAF300) | This report |
| Plasmid | | |
| pBSKS(-) | Ap ^r | Stratagene Corp. |
| pCB182 | Ap ^r <i>lacZ</i> | [11] |
| pST1021 | Cm ^r , <i>Klebsiella pneumoniae</i> nif ^r BAL ^r | [12] |
| pRK2013 | Km ^r Tra ⁺ | [13] |
| pAF300 | Tc ^r Cm ^r mob ⁺ | [14] |
| pLYM106 | Ap ^r , <i>A. brasiliense</i> Yu62 <i>draTG</i> (8.0kb <i>Sal</i> I fragment) in pUC19 | This report |
| pBE(+) | Ap ^r , <i>A. brasiliense</i> Yu62 <i>draTG</i> upstream region (1.9kb <i>EcoR</i> I Fragment) in pBSKS(-) | This report |
| pAT1 | Tc ^r , mob ⁺ , 2.5kb fragment in pAF300 (<i>draT</i> :: <i>cam</i>) | This report |

1.2 培养方法

巴西固氮螺菌用 LD 培养基^[4]30℃培养,大肠杆菌用 LB 培养基 37℃培养。

1.3 DNA 的制备

质粒和单链 M13 噬菌体 DNA 的提取按文献[15]进行。

1.4 DNA 序列测定及计算机分析

采用双脱氧末端终止法^[16],以单链 M13 噬菌体 DNA 为模板,按 Perkin Elmer Cetus 公司 Ampli Taq DNA 聚合酶测序试剂盒说明书进行。用 Beckman 公司的 MicroGenie 序列分析软件进行序列拼接、分析等。

1.5 质粒的接合转移

按文献[17]进行。

1.6 酶及试剂

核酸限制酶、DNA 聚合酶 Klenow 片段购自 Promega 公司和华美公司, T4DNA 连接酶购自 New England Biolabs 公司, Ampli Taq DNA 聚合酶测序试剂盒购自 Perkin Elmer

Cetus 公司, α - 32 P-dATP 购自生辉公司, Erase-a-Base-System 购自 Promega 公司。

2 结果与讨论

2.1 draTG 基因上游区域的 DNA 序列分析

从 pLYM106 (带有巴西固氮螺菌 Yu62 8.0kb draTG 同源片段) 上回收 1.9kb 的 EcoRI 片段, 亚克隆至 pBKS 上, 获得质粒 pBE (+)。用 ExoIII 缺失法分别从插入片段的两端进行缺失, 从而获得两个方向的不同缺失程度的一系列亚克隆。用这些亚克隆的质粒 DNA 做模板直接测序, 结果只有少数的亚克隆能获得较满意的结果, 绝大多数的反应几乎无法延伸, 原因可能是 *A. brasilense* 的 (G+C)% 较高 (64%~71%), 使双链 DNA 模板不易变性。因此, 又将各亚克隆中的插入片段回收后分别亚克隆到 M13mp 18/19 上, 提取单链模板测序, 最终获得了该 1900bp 片段的全核苷酸序列 (有约 50bp 片段是来自 pUC19 载体多克隆位点和 EMBL3 载体), 其中 90% 为双向测序结果。

序列分析结果表明该片段含有 94bp 的 nifH 基因编码区, 其转录方向与 draTG 相反, 这与 *A. brasilense* Sp7 和 *R. rubrum* 中的情况一致。draT 与 nifH 相距 1866bp。而在 *R. rubrum* 中, draT 与 nifH 仅隔 400bp。同源比较表明该部分 nifH 基因及其上游启动子区与阎大来^[18]用 PCR 法从 *A. brasilense* Yu62 中克隆出的 nifH 启动子区完全一致。根据 *A. brasilense* 基因的密码子利用率^[19], 分析该 1.9kb 片段中可能的阅读框架 (ORF), 发现有三个 ORF 的密码子利用符合 *Azospirillum* 基因的规律, 即基因序列内密码子的第三位为 G 或 C 的比例均在 90% 以上。其中有两个 ORF 与 draT 转录方向相同, 另一个与 draT 转录方向相反。但这三个 ORF 与基因库中的已知基因都没有同源性。在 *A. lipoferum* 中, draT 与 nifH 之间的区域通过杂交发现与 *K. pneumoniae* nifJ 有较弱的同源性。对 *A. brasilense* Yu62 该区域的序列分析表明该区域没有 nifJ 基因。

在 nifH 与 draT 之间的区域内, 有一些可能的调控序列, 它们分别是: (1) σ^{54} 因子结合序列: GG-N₁₀-GC, 该序列是 RpoN-依赖型 (或 σ^{54} -依赖型) nif 启动子的特征序列, 它也被称为下游启动子组份 (DPE: downstream promoter element)。在该区域有三个可能的 DPE。其中之一在 nifH 上游 (在图 1 中以虚下划线标注), 该 DPE 已证实是 nifH 的启动子^[20]。另外两个都在 draT 上游, 分别被命名为 DPE1 和 DPE2 (在图 1 中以下线标注)。DPE1 在 draT 起始密码子上游 -1483 至 -1498 之间, 与 nifH 起始密码子相距 380bp, 与 DPE1 相距 282bp 处正好是 nifH 启动子的 UAS 区 (NifA 结合位点)。从与 nifH 相距的位置, 及同源比较推测该 DPE 可能就是 de Zamaroczy 等^[20]在 nifH 上游互补链上发现的 -380 至 -396 区的 RpoN-依赖型启动子 (CTGGC-N₉-GCA)。Zhang 等^[4]曾推测 draT 可能从该启动子转录。而 DPE2 位置恰在 draT 起始密码子前, 仅隔一个碱基 (如图 1)。(2) 上游激活序列: TGT-N₁₀-ACA (UAS: upstream activator sequence), 该序列是 NifA 的结合位点。在该区域只有一个 UAS, 该 UAS 就是 nifH 启动子的 UAS 序列, 它具有两个相互重叠的 UAS (在图 1 中以上线标注)。该 UAS 也有可能参与 dra 操纵元的转录。(3) 富 A+T 区: 在该区找到了三个富 A+T 区, 一个在 nifH 的 DPE 与其 UAS 之间, 第二个在 nifH 启动子的 UAS 与 DPE1 之间, 第三个在 draT 上游约 -160bp 处 (如图 1)。 *A. brasilense* 的 DNA (G+C)% 通常在 64%~71% 之间, 而这三个富 A+T 区, (G+C)% 只

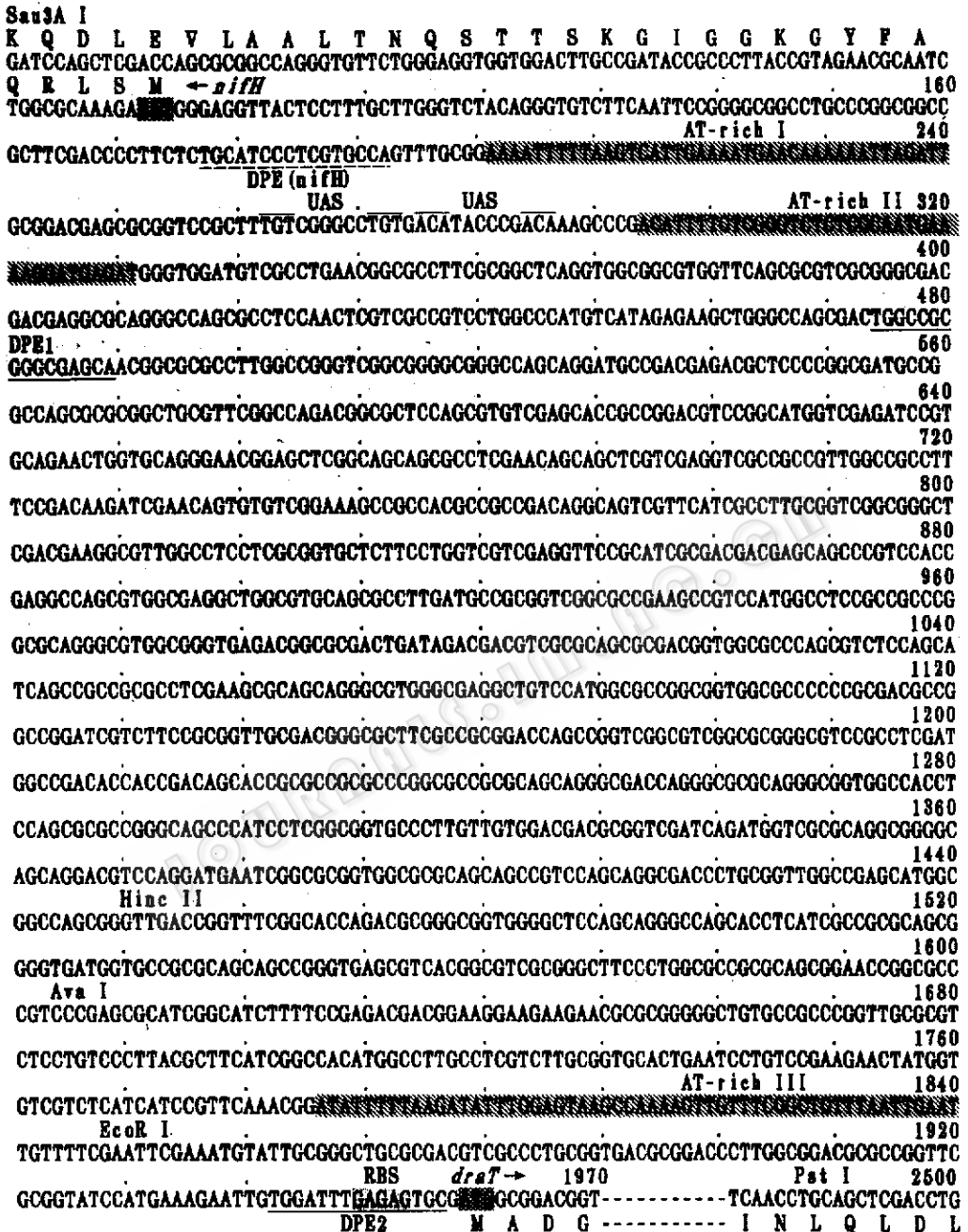


图 1 巴西固氮螺菌 Yu62 *draTG* 基因上游区域的核苷酸序列

Fig. 1 Nucleotide sequence of upstream region of *draTG* genes from *A. brasilense* Yu62

For the genes orientated in 5'→3' direction, the amino acid sequences are given in one-letter code below the nucleotide sequences; for gene in opposite direction, above. Putative RpoN-dependent promoter sequences are marked DPE and shown by underlining (in 5'→3' direction) and dashed underlining (in opposite direction); upstream activator sequences are marked UAS and overlined; putative ribosome binding sites are marked RBS and shaded; and A + T-rich regions are marked AT-rich and shaded. The potential start condons are shaded and stop codons are shown by asterisks. The arrows indicate the orientation of transcription. The short lines indicate the omitted sequences.

有6%~26%。在 *Klebsiella pneumoniae* 和 *Azotobacter vinelandii* 的一些 nif 基因(如: nifH, nifE, nifU, nifB)启动子的 DPE 与 UAS 之间也发现有同样的富 A+T 区^[21,22]。Baynon^[21]认为,这些富 A+T 区可能是普通的 RNA 聚合酶识别位点,富 A+T 的比例及它们在 nif 特异性启动子中的位置可能影响转录起始的频率。此外,不同根瘤菌 nifA 基因的序列比较表明,-30 至 -40 之间的富 A+T 区可能是 nifA 的启动子成分^[23]。

由以上的分析可看出:nifH 与 draT 之间的区域可能主要起调控功能而非编码功能, dra 操纵元的启动子很可能是 RpoN-依赖型。在 *R. rubrum* 中,从 DRAT 的 N-端分析表明 draT 以 TTG 为起始密码子^[6],从该起始密码子开始的翻译是很差的,它实质上控制了细胞内 DRAT 和 DRAG 的量。在 *A. brasilense* 中 DRAT 酶未得到纯化。但从 draT 的序列分析推测可能以 ATG 作为起始密码子,如果这种推测正确的话, *A. brasilense* 中,细胞内 DRAT 和 DRAG 的量主要通过 draTG 基因转录水平来控制。与 *A. brasilense* Yu62 draT 基因起始密码子仅隔一个碱基处以及相距约 1.5kb 处都发现了一个类 DPE 序列(DPE1 和 DPE2)。DPE 的这种特殊位置是否即是控制 draTG 操纵元转录的一种方式? 这需要进一步研究。

2.2 draTG 上游区域的启动子功能鉴定

为了证实 draTG 上游区域的启动子功能,我们从 pLYM106 质粒上回收了 2.5kb 的 Pst I 片段,该片段含有 draT 上游区域和 600bp 的 draT 基因编码区。将该片段亚克隆到启动子检测载体 pAF300 Cm 抗性基因上游的 Pst I 位点,从而获得重组质粒 pAT1 和 pAT2。对这两个重组质粒进行酶切分析,结果表明 pAT1 是所需的 draT::cam 融合质粒(如图 2)。由于 pAF300 中 cam 基因没有启动子,它只能通过 draT 的启动子来转录表达,因而通过检测该转录融合质粒的 cam 基因是否表达(检测 Cm 抗性或细胞粗体液的 CAT 活性)即可说明 draT 是否转录。用三亲本杂交法分别将 pAT1 及 pAF300 引入 *A. brasilense* Yu62 菌株,转交子经纯化后提取质粒进行鉴定。将 Yu62(pAT1)转交子和 Yu62(pAF300)转交子接种到 Cm 抗性(25μg/ml)LD 平板上,两天后观察这两种菌的生长情况,同时以 *E. coli* (pAT1)、*E. coli* (pAF300)和 *A. brasilense* Yu62 菌株做对照,并将 Yu62(pAT1)接种到 Ap^rNx^rTc^r 的 LD 平板上,作为菌生长的正对照。结果表明 Yu62(pAT1)可在 Cm^r 平板上生长, Yu62(pAF300)、*E. coli* (pAT1)、*E. coli* (pAF300)只有很微弱地生长而 Yu62 不生长。这结果说明 draT 可以在好氧的条件下转录,而且只能在 *A. brasilense* 中才能转录。因而显然 draT 的转录需要 *A. brasilense* 中的某一特殊因子,而该因子在 *E. coli* 中是

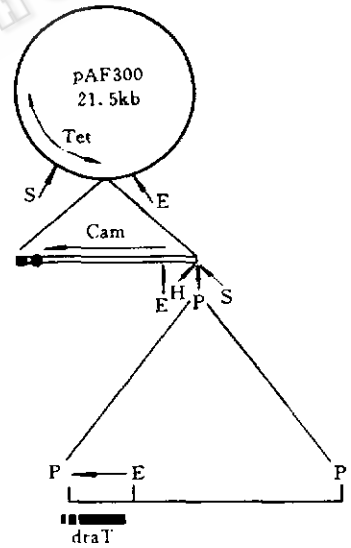


图2 draT::cam 转录融合质粒 pAT1 的构建

Fig.2 Construction of draT::cam fusion plasmid pAT1

Light arrows indicate the direction of transcription. ■ and ● indicate T1 and T2 transcription terminators respectively. Restriction sites: E-EcoRI, H-Hind III, P-PstI, S-SalI.

不存在的。已知 *A. brasilense* 的 NifA 在好氧或高 NH_4^+ 时都没有活性,因而在这种条件下 draT 的转录并非由 NifA 激活,但不能排除在微好氧限铵条件下, NifA 激活 dra 操纵子转录的可能性。同时,上述结果还说明 draTG 上游区域具有启动子功能。

2.3 NifA 是否参与 draTG 转录调控

为了确定 NifA 是否参与 draTG 的转录调控,我们从 pAT2 上再次回收了上述 2.5kb 片段并定向克隆到启动子探针载体 pCB182 上(该质粒的图谱参见文献[11]),从而获得 draT::lacZ 转录融合质粒 pCT1。将该质粒分别引入大肠杆菌 CB454 及 CB454 (pST1021)中,测定各转化子的 β -半乳糖苷酶活性。同时测定大肠杆菌 CB454(pCT1)和大肠杆菌 CB454(pST1021)的 β -半乳糖苷酶活性,以作参照。结果各菌的 β -半乳糖苷酶活性单位均小于 1。由于阎大来等^[18]已证明肺炎克氏杆菌的 NifA 可有效启动 nifH 的表达,即 nifH 上游的 UAS 可与肺炎克氏杆菌的 NifA 结合。而 pST1021 含有肺炎克氏杆菌的 nifA,该 nifA 在四环素抗性基因启动子下组成型表达。因而,上述的结果说明 NifA 并不参与 draTG 的转录调控。同时也说明 draT 上游的两个重叠的 UAS 并非 dra 操纵元的启动子组份。

3 讨 论

Inoue A. 等^[5]用 draT::lacZ 融合子研究 *A. lipoferum* draT 基因的表达。他们的结果表明:draT 只在微好氧条件(30%空气)下表达,好氧(100%空气)条件下没有或只有微量表达。而且 draT 的表达与铵浓度无关。我们对 *A. brasilense* draT::cam 融合子的研究则表明 draT 可在好氧条件下表达。但由于我们没有检测融合子的 CAT 活性,无定量分析的数据。因而在 Cm^r 平板上 draT::cam 融合子的生长也有可能是 draT 微量表达的结果。至于融合子在丰富培养基上生长这一结果则与 Inoue A. 等认为的 draT 的表达与铵浓度无关是一致的。*A. lipoferum* 和 *A. brasilense* dra 操纵元的表达调控是否真有什么区别,尚需进一步证实。此外,我们的实验结果还表明,dra 操纵元的表达需要某种 *E. coli* 中不存在的特殊因子,而且该因子并非 NifA。在共生固氮菌中,nifA、fixK 等基因的氧调控是通过 fixLJ 基因来完成的^[24]。Inoue A. 等认为 dra 操纵元的氧调控可能类似于这些基因。在 *A. brasilense* 中曾发现一些可能的 fix 位点^[25],但尚未有人证实该菌中有类 fixLJ 的基因。

参 考 文 献

- [1]Pedrosa F O. CRC Crit Rev Plant Sci., 1985, 6:345~384.
- [2]Hartmann A, Fu H-A, Burris R H. J Bacteriol, 1986, 165:864~870.
- [3]Ludden P W, Roberts G P. Curr Top Cell Regul, 1989, 30:23~55.
- [4]Zhang Y P, Burris R H, Roberts G. J Bacteriol, 1992, 174:3364~3369.
- [5]Inoue A, Shigematsu T, Hidaka M *et al.* Gene, 1996, 170:101~106.
- [6]Fitzmaurice W P, Saari L L, Lowery R G, *et al.* Mol Gen Genet, 1989, 218:340~347.
- [7]Masepohl B, Krey R, Klipp W. J Gen Microbiol, 1993, 139:2667~2675.
- [8]杨洁彬,曹增良,李季伦.北京农业大学学报,1984,10(3):321~329.
- [9]马旅雁,李季伦.生物工程学报,1997,13(3):227~235.

- [10] Messing J. *Recom DNA Tech. Bull.*, 1979, 2(2):43.
- [11] Schneider K, Beck C F. *Gene*, 1986, 42:37~48.
- [12] 朱家璧, 俞冠翊, 江益群. *中国科学(B辑)*, 1983, 8:688~696.
- [13] Figurski D & Helinski D R. *Proc Natl Acad Sci UAS*, 1979, 76:1648~1652.
- [14] Fani R, Bazzicalupo M, Ricci F *et al.* *FEMS Microbiol. Lett*, 1988, 5:271~276.
- [15] Sambrook J, Maniatis T, Fritsch EF. *Molecular Cloning: A Laboratory Manual*, 2nd ed, Cold Spring Harbor Laboratory Press, 1989.
- [16] Sanger F. *Proc Natl Acad Sci USA*, 1977, 74:5463~5467.
- [17] Pedrosa F O, Yate M G. *FEMS Microbiol Lett*, 1985, 23:95~101.
- [18] 阎大来, 何路红, 马旅雁等. *生物工程学报*, 1995, 11(3):205~210.
- [19] Fani R, Allotta G, Bazzicalupo M. *Mol Gen Genet*, 1989, 220:81~87.
- [20] De Zamaroczy M, Delorme F, Elmerich C. *Mol Gen Genet*, 1989, 220:88~94.
- [21] Beynon J, Cannon M, Buchanan-Wollaston V *et al.* *Cell*, 1983, 34:665~671.
- [22] Jacobson M R, Brigle K, Bennett L T *et al.* *J Bacteriol*, 1989, 171:1017~1027.
- [23] Iismaa S E, Watson J M. *Mol Microbiol*, 1989, 3:943~955.
- [24] Fisher H M. *Microbiol Rev*, 1994, 58:352~386.
- [25] Galimand M, Perroud B, Delorme F *et al.* *J Gen Microbiol*, 1989, 135:1047~1059.

Sequencing and Analysis of Function of the Promoter Region of draTG Genes from *Azospirillum brasilense* Yu62

Ma Lüyan Li Jilun

(National Laboratories for Agrobiotechnology, Beijing Agricultural University, Beijing 100094)

Abstract From *Azospirillum brasilense* Yu62, we cloned and sequenced the upstream region of draTG. No identified genes were found in this region except for portion of nifH, which is transcribed divergently from draTG. However, some potential regulatory elements were found, which include downstream promoter elements (DPE), upstream activator sequences (UAS), and A + T-rich regions. It suggests that the region between draT and nifH might be a regulatory region rather than an encoding region, and the promoter of dra operon would probably be a RpoN-dependent promoter. A draT::cam transcriptional fusion plasmid pAT1 was constructed in pAF300. The expression of draT was studied in *Escherichia coli* and *A. brasilense* by Cm^r assay. The result showed that draT could be transcribed in *A. brasilense* but not in *E. coli* while grown aerobically on LD plate. It suggests that the transcription of draT needs some factors which are absent in *E. coli*, and the draT upstream region has the promoter function. Using promoter probe vector pCB182, a draT::lacZ transcriptional fusion plasmid pCT1 was constructed. β -galactosidase activity was determined *in vivo* in *E. coli*, in the presence or absence of NifA of *Klebsiella pneumoniae*, respectively. The results demonstrated that NifA wasn't involved in the transcription of draTG.

Key words *Azospirillum brasilense*, draTG promoter, nucleic acid sequence, transcriptional fusion plasmid