

5'端非翻译区 对 LT-B 基因表达的影响

叶棋浓 陈添弥

(军事医学科学院生物工程研究所, 北京 100850)

摘要 mRNA 5'端非翻译区的不同结构可影响基因表达, 为了改善编码人毒素源性大肠杆菌热敏感肠毒素 B 亚单位的 LT-B 基因的表达水平, 我们把该基因置于 pBV220 载体的 P_RPL 串联启动子下游, 构建了带有不同核苷酸组成的 5'端非翻译区的重组体。这些重组体分别在大肠杆菌 HB101 和 DH5 α 中表达。结果表明, 起始密码前有两个连续串联 SD 序列的 LT-B 基因的表达水平低于只有单个 SD 序列下的表达水平, 而翻译偶联可使表达改善; 用不同的 SD 序列 LT-B 基因的表达水平也有所不同, 用基因本身 SD 序列可能要比用 pBV220 P_L 启动子下游的 SD 序列好; 在只含单个 LT-B 基因 SD 序列的重组体中, 5'端非翻译区序列的长短对 LT-B 基因表达没有什么影响, 重组体在 HB101 中的表达水平高于在 DH5 α 中的表达水平。

关键词 热敏感肠毒素, 5'端非翻译区, SD 序列, 基因克隆与表达

大肠杆菌可引起急性腹泻, 据世界卫生组织统计, 急性腹泻中 25% 是由大肠杆菌引起的, 其中毒素源性大肠杆菌 (ETEC) 是很多发展中国家和一些发达国家引起腹泻的主要病原菌^[1]。在人源 ETEC 的研究领域里, 国内外都已成功地克隆了 LT-B 基因, 并进行了免疫和保护试验^[2-4], 但对其表达调控研究甚少。本实验室曾把 LT-B 基因分别置于 lac, tac, cat, lpp 启动子下游进行表达^[5], 结果在启动效率较高的 tac, lpp 启动子下的表达水平也不高, 说明基因表达不但受转录水平的调控, 而且还受翻译水平的调控。本文从 SD 序列数目对 LT-B 基因表达的影响入手, 构建了带有不同 5'端非翻译区的重组质粒, 从翻译水平上研究了 SD 序列数目、不同的 SD 序列及 5'端非翻译区长度等对 LT-B 基因表达的影响。

1 材料与方法

1.1 菌株和质粒

E. coli H10407 为产 LT 的野生菌株, 国际上常用作标准人 ETEC 株, pCHP23 为本实验室构建的带有人 LT-B 基因的重组质粒, pBV220 为质粒载体, 由中国预防医学科学院提供, *E. coli* HB101 和 DH5 α 为受体菌。

1.2 培养基

普通 LB 培养基和 CAYE-2 产毒培养基按文献〔6〕配制。

1.3 试剂

限制酶、T4 连接酶和 Bal31 外切酶等为 Biolab, BRL, Promega 和 Pharmacia 等公司的产品。16 碱基寡核苷酸引物 5'CCGAACATAGTTCTGT3' 由本室合成, 作序列分析

用。T7 sequencing kit 为 Pharmacia 公司产品。SPA 标记的 LT 抗体, LT 单克隆抗体和辣根过氧化物酶标记的羊抗鼠 IgG 由 302 医院魏燕玲同志提供。

1.4 DNA 体外重组方法

质粒 DNA 提取、酶切、回收、连接、转化和菌落原位杂交按参考文献〔7〕进行。

1.5 SPA 协同凝集反应——用作菌株产 LT 的定性试验

细菌在产毒培养基中 150 r/min 振荡培养至 OD₆₀₀ 值为 0.3—0.5, 42℃ 诱导 4—6 小时。取 1.5ml 菌液, 于 Eppendorff 管中离心 5 分钟收集菌体, 菌体悬浮于 100μl 0.25 mg/ml 的多粘菌素 B 中, 37℃ 1 小时, 离心收集上清。在洁净玻片上, 取上述上清 1.5μl, 加 1.5μl SPA 标记的 LT 抗体, 混匀, 若出现均匀颗粒状凝集即为阳性反应, 反之为阴性反应。

1.6 GM1-ELISA⁽⁸⁾——LT 的定量检测

用上述同样条件下培养及诱导后, 取菌液用 LB 培养液调 OD₆₀₀ 值到 0.5, 取该调整好的菌液 1.5ml, 离心收集菌体, 再悬浮于 200μl 0.25mg/ml 的多粘菌素 B 溶液中, 37℃ 1 小时, 离心收集上清。取稀释后的 100μl 上清进行 GM1-ELISA 试验。此方法与传统方法的不同之处在于它用特异性的受体 GM1 代替纯化抗体包被酶联板。

1.7 DNA 序列测定

根据 Sanger 双脱氧末端终止法原理测定 DNA 序列, 按 Pharmacia 公司的 T7 sequencing kit 说明进行。

1.8 SDS 聚丙烯酰胺凝胶电泳和 Western blot

SDS 聚丙烯酰胺凝胶电泳按参考文献〔7〕进行, 电泳毕在蛋白转移装置上进行电转移, 蛋白条带即吸附于孔径为 0.45μm 的硝酸纤维素膜上, 然后进行免疫杂交反应。用 3% BSA 封闭膜 2 小时, 与 LT 单克隆抗体 (1:300 稀释) 反应, 室温振摇 4 小时, 洗膜后再与辣根过氧化物酶标记的羊抗鼠 IgG (1:300 稀释) 反应 1 小时, 洗膜后加 DAB 进行显色反应。

2 结 果

2.1 重组体的构建及鉴定

2.1.1 带有载体 pBV220 P_L 启动子下游的 SD 序列和 LT-B 基因本身 SD 序列重组子的构建: 重组质粒构建示意图如图 1。用 EcoRI 和 Hind III 双酶切质粒 pCHP23, 回收 0.6kb LT-B 基因片段, 补平后插入到经 EcoRI 消化并补平的 pBV220 中, 转化大肠杆菌 HB101, 转化子经用 LT-B 基因作探针进行原位杂交筛选, DNA 酶切鉴定后得到预期的重组子命名为 pMG208。

2.1.2 去掉载体 pBV220 P_L 启动子下游的 SD 序列但带有 LT-B 基因本身 SD 序列重组子的构建: 载体 pBV220 经 EcoRI 酶切后, 用 Bal31 外切酶控制不同时间 (5、10、30、60、120 分钟) 对相同浓度的样品进行酶切, 补平后与含 LT-B 基因的片段连接, 转化大肠杆菌 HB101。由于 Bal31 酶切情况不同, 可能去掉了载体中的 SD 序列并带有不同长度的 5' 端非翻译区序列。转化子经原位杂交筛选、酶切鉴定后得到预期的重组子命名为 pMG209-217。图版 I-A 是其中 2 个重组质粒即 pMG213 和 pMG214 的酶切图谱 (其余重

组子图略), 所得结果与预期的相同。

2.1.3 缩短 LT-B 基因 5'端非翻译区可能去掉了 LT-B 基因本身 SD 序列但带有载体 pBV220 P_L 启动子下游的 SD 序列的重组子的构建: 含有 LT-B 基因的 pCHP23 质粒经 EcoRI 酶切后, 用 Bal31 酶切不同时间 (10、30、60、180、600 秒), 再用 Hind III 酶切, 回收含 LT-B 基因的片段, 补平后插入到经 EcoRI 消化并补平的 pBV220 中, 转化 HB101。转化子经菌落原位杂交筛选, 酶切鉴定后得到预期的重组子命名为 pMG218-220。

以上 3 种构建由于 LT-B 基因片段与载体之间都是平端连接, 所以 LT-B 基因插入载体有两种可能的方向。我们用 SPA 协同凝集反应鉴定其插入的正反方向, 当正向插入时, LT-B 基因可在载体 P_RP_L 启动子控制下在大肠杆菌中表达, SPA 协同凝集反应为阳性, 反之为阴性。还曾用酶切方法鉴定, 分别选择目的基因和载体上的 Cla I 和 Bgl II 进行双酶切, LT-B 基因正向与反向插入可得到不同大小的片段, 经电泳后可区分开 (图略)。我们又把经 SPA 协同凝集反应鉴定为阳性的重组质粒转化到 DH5 α 中, 以便比较重组质粒在不同受体菌中的表达情况。

2.2 LT-B 基因在大肠杆菌中的表达

用 GM1-ELISA 方法测定胞内的 LT-B 基因的表达水平 (结合序列分析见表 1), 从表中我们可以看到: LT-B 基因起始密码前有两个连续串联 SD 序列的表达水平低于只有单个 SD 序列下的表达水平, 但前者不同重组体的表达水平也存在着差异; 用基因本身 SD 序列表达 LT-B 基因可能要比用载体 pBV220 P_L 启动子下游的 SD 序列好; 除 pMG212 外, 其它只含单个 LT-B 基因 SD 序列的重组体中, 5'端非翻译区序列的长短对 LT-B 基因表达没有什么影响, pMG212 表达水平偏低可能是由缺失 P_L 启动子中 -10 区所致; 重组体在 HB101 中的表达水平高于在 DH5 α 中的表达水平, 并且在 HB101 和在 DH5 α 中有上述相似的规律。

重组质粒转化到大肠杆菌 HB101 中进行表达, 细菌总蛋白作 SDS 聚丙烯酰胺凝胶电泳, 由于蛋白条带较多, LT-B 基因产物在凝胶中看不太清楚。鉴于这种情况, 我们进行了 Western 免疫印迹杂交, 从图版 I-B 可以看出, 重组子确实表达了 LT-B 基因产

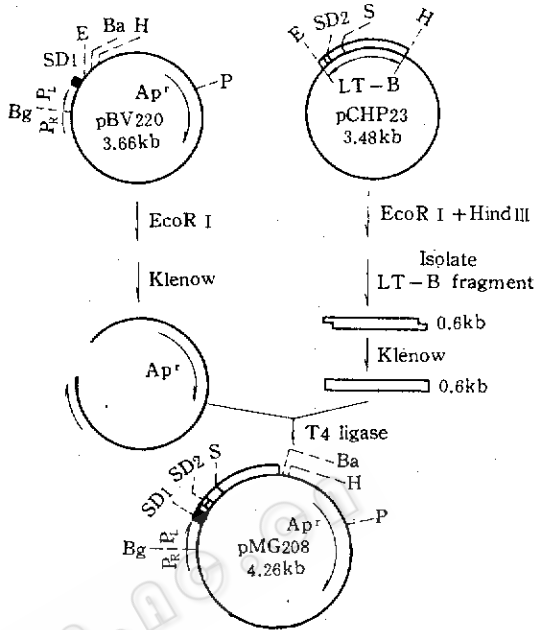


图 1 带有载体 pBV220 P_L 启动子下游的 SD 序列和 LT-B 基因本身 SD 序列重组子的构建

Fig. 1 Construction of the recombinant plasmid containing a SD sequence downstream from P_L promoter of the vector pBV220 and a SD sequence of LT-B gene itself

SD1: SD sequence downstream from P_L promoter of the vector pBV220, SD2: SD sequence of LT-B gene itself, Ba: BamHI, Bg: Bgl II, E: EcoRI, H: Hind III, P: PvuI, S: SacI

物。

表 1 用 GM1-ELISA 方法测定 LT-B 基因在不同构建中的表达

Table 1 Expression of LT-B gene in various constructions detected by GM1-ELISA

	Nucleotide sequence upstream from the ATG codon	OD ₄₉₀	OD ₄₉₀
		(HB101)	(DH5 α)
Recombinant:			
pMG208	<u>TTAAGGAGGAATTAATTCGGAATGAATTATG</u>	0.60	0.27
pMG209	<u>ACGAAACGAAGCAAATTCGGAATGAATTATG</u>	1.84	0.91
pMG210	<u>TGATACTGAGCACAATTCGGAATGAATTATG</u>	1.94	0.94
pMG211	<u>AGCATTCAAAGCAAATTCGGAATGAATTATG</u>	1.87	0.91
pMG212	<u>GACATAAATACCAAATTCGGAATGAATTATG</u>	1.07	0.52
pMG213	<u>AGGGCAGCATTCAAATTCGGAATGAATTATG</u>	1.89	0.94
pMG214	<u>AAAATTAAGCCCTAATTCGGAATGAATTATG</u>	1.83	0.97
pMG215	<u>GAAGGTGACGCTCAATTCGGAATGAATTATG</u>	1.82	0.96
pMG216	<u>ATTAAGCCCTGAAAATTCGGAATGAATTATG</u>	1.81	0.91
pMG217	<u>TGAAGAAGGGCAGAATTCGGAATGAATTATG</u>	1.82	0.95
pMG218	<u>AAATTAAGGAGGAATTCGGAATGAATTATG</u>	1.13	0.51
pMG219	<u>AAAATTAAGGAGGAATTCGGAATGAATTATG</u>	1.12	0.51
pMG220	<u>GCATTGGTTAAAAATTAAGGAGGAATTTATG</u>	0.51	0.27
Control:			
H10407			0.54
pCHP23 (JM103)			0.60
pBV220		0.05	
pBV220			0.05

TAAGGAGG; SD sequence downstream from P_L promoter of the vector pBV220, GGAA: SD sequence of LT-B gene itself

2.3 重组质粒的序列分析

我们用 GM1-ELISA 方法测定重组子的表达水平后,根据表达水平高低从3种构建中挑选了一些重组子进行序列分析,即从第一种构建的重组子中挑1个(命名为 pMG208),从第二种中挑9个(命名为 pMG209-217),从第三种中挑3个(命名为 pMG218-220)作序列分析,为了直接用双链质粒作模板测定重组质粒5'端非翻译区的碱基序列,我们在目的基因 LT-B 上设计并合成了一个长16碱基的引物,即5'CCGAACATAGTTCTGT3',该引物距离 ATG 起始密码75个碱基,与 LT-B 基因的一条正链互补,经计算机检索表明,该引物本身不能形成发夹或茎环结构,不存在与载体 DNA 和 LT-B 基因 DNA 8个碱基以上的互补序列。载体 pBV220中从 P_L 启动子开始到多克隆位点第一个 EcoRI 位点为止及目的基因 LT-B 中从5'端 EcoRI 位点开始到引物为止的碱基组成如图2所示,重组质粒3种不同构建的部分重组子的5'端序列分析结果如图版 I-C 所示。

3 讨论

由 mRNA 翻译成蛋白质的限速阶段发生在翻译起始中, mRNA 5'端翻译起始区(TIR)的长短和组成可影响基因的表达水平^[9]。我们将 LT-B 基因置于 P_RP_L 强启动子下游,比较了 LT-B 基因 ATG 起始密码前的 SD 序列数目对其表达的影响,结果和国外大多数报道的结果一致^[10-12]。

```

ACAGATAACC ATCTGCGGTG ATAAATTATC TCTGGCGGTG TTGACATAAA
      -35 like sequence           -35 region
TACCA*CTGGC GGTGATACTG AGCAC*ATCAG CAGGACGCAC TGACCACCAT
      -10 region           mRNA
GAAGGTGACG CTC*TTAAAAA TTAAGCCCT*G AA*GAAGGGCA G*CATTCA* AAG

CA*GAAGGCTT TGGGGTGTGT GATACGAAAC GAAGCA*TTGG TTA AAAATTA

AGGAGGAATTC
SD1 EcoRI

      GAAT*TC GGAATGAAT*T ATGAATAAAG TAAATGTTA TGTTTTATTT
      EcoRI SD2 Initiation codon
ACGGCGTTAC TATCCTCTCT ATGTGCATAC GGAGCTCCCC AGTCTATTAC

AGAACTATGT TCGGAATATC
Primer binding region

```

图2 载体 pBV220 中从 P_L 启动子开始到多克隆位点第一个 EcoRI 位点为止及目的基因 LT-B 中从 5' 端 EcoRI 位点开始到引物为止的碱基组成

Fig. 2 Nucleotide sequence of the vector pBV220 from P_L promoter to EcoRI site and of LT-B gene from 5' EcoRI site to primer binding region

* * indicates the location of deleted 5' non-coding region of different recombinants

不同的 SD 序列及其与 ATG 之间的距离对基因表达也有影响^[13,14]。我们的实验结果表明, LT-B 基因在本身的 SD 序列控制下的表达水平是比较高的, pMG220 表达水平偏低一方面可能是由于缺失了 LT-B 基因本身 SD 序列而带有载体 SD 序列, 另一方面, pMG220 中的 SD 序列与 ATG 间的距离为 5 个碱基, 在较高水平表达的只含单个 LT-B 基因 SD 序列的重组子中, 距离为 6 个碱基, 推测后者可能更合适。在我们构建的只含一个 LT-B 基因的 SD 序列的重组子中, 5' 端非翻译区长度分别为 15, 53, 69, 72, 81, 87, 92, 126 个碱基, 结果表明, 其长度对 LT-B 基因的表达水平没有什么影响。我们用计算机辅助的检索表明, 在 SD 序列及其附近不存在阻碍核糖体结合的二级结构, 因而这些重组子可能有相同效率的翻译起始^[15], 而使它们的表达水平大体相同。

参 考 文 献

- [1] Qualk C *et al.* J Clin Microbiol, 1987, 25:1572.
- [2] Klipstein F A *et al.* Infect. Immun, 1982, 37: 550.
- [3] Maskell D J *et al.* Microb Pathogenesis, 1987, 2:211.
- [4] Clement J D *et al.*, Infect Immun, 1990, 58:1159.
- [5] 程新波等. 第二届基因工程学术讨论会摘要汇编, 1989, p64.
- [6] 李丰生等. 生物工程学报, 1987, 3:102.
- [7] Sambrook J *et al.* Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, 1989.
- [8] 刘传喧等. 中华微生物学与免疫学杂志, 1989, 9:62.
- [9] Jacques N *et al.* Mol Micro, 1990, 4:1063.
- [10] Dalboge H *et al.* DNA, 1988, 7:339.
- [11] Shigesada K *et al.* Gene, 1987, 53:163
- [12] Goeddel D V *et al.* Methods in Enzymology, 1990, 185:94.
- [13] Curry K A *et al.* DNA, 1988, 7:173.
- [14] Spanjaard R A *et al.* Gene, 1989, 80:345
- [15] Petersen C *et al.* J Mol Biol, 1989, 206:323.

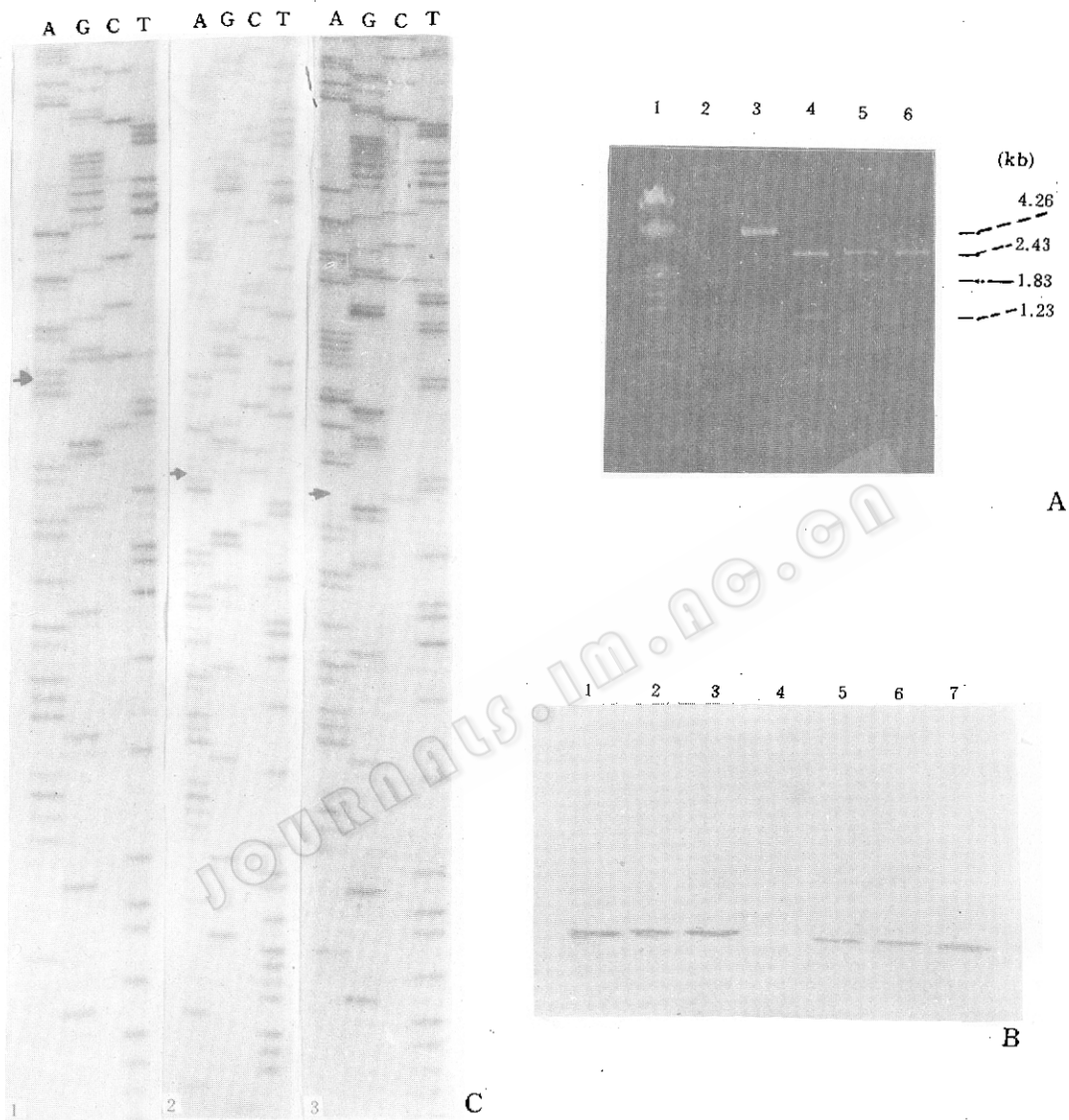
Effect of 5' non-coding Region on Expression of LT-B Gene

Ye Qinong Chen Tianmi

(Molecular Genetics Center, Institute of Biotechnology, Academy of Military Medical Science, Beijing 100850)

Abstract The different construction of mRNA 5' non-coding region may affect gene expression. In order to improve expression level, we constructed recombinant plasmids with different nucleotide composition of the 5' non-coding region which direct the synthesis of human toxicogenic *E. coli* heat-labile enterotoxin B subunit coded by LT-B gene under the control of the P_{RP1} , tandem promoter of vector pBV220. These recombinants were expressed in *E. coli* HB101 and DH5 α , respectively. The results show that the expression levels of LT-B gene with two tandem SD sequences upstream from the initiation codon are lower than those with only one SD sequence and translation coupling can improve the expression levels; The different SD sequences can affect the expression of LT-B gene a little. The SD sequence of LT-B gene itself may be better for expression than that of the vector pBV220. The length of 5' non-coding region in the recombinants harboring only one SD sequence has no effect on the expression of LT-B gene. The expression of LT-B gene in HB101 is higher than that in DH5 α .

Key words Heat-labile enterotoxin, 5' non-coding region, SD sequence, gene expression



A. Restriction analysis of recombinant plasmids

1. λ DNA + EcoRI + Hind III, 2. pMG213, 3. pMG213 + Bgl I, 4. pBV220 + Bgl I + Pvu I 5. pMG213 + Bgl I + Pvu I, 6. pMG214 + Bgl I + Pvu I

B. Western blot analysis of LT-B expressed in *E. coli*

1. pMG209, 2. pMG212, 3. pMG215, 4. pMG220, 5. H10407, 6. pMG208, 7. pMG219

C. The result of DNA sequencing of the partial recombinants

Over the arrows are the sequences of vector; Under the arrows are the target sequences.

1. pMG209, 2. pMG210, 3. pMG219