

棒状杆菌 2,5-二酮基-D-葡萄糖酸 还原酶 I 基因在大肠杆菌中的克隆与表达*

陈策实 尹光琳

(中国科学院上海生物工程研究中心 上海 200233)

摘 要 从棒状杆菌 (*Corynebacterium* sp. SCB3058) 初步纯化得到两个 2,5-二酮基-D-葡萄糖酸 (2,5-DKG) 还原酶, 在此基础上利用 PCR 技术, 以基因组 DNA 为模板, 扩增得到含有 2,5-DKG 还原酶 I 基因的片段, 定向连接到 PGEM3Zf(+) 并转化大肠杆菌 DH5 α , 筛选得到阳性克隆 pGEM813。序列分析表明, 片段全长 1107bp, 含有一个 834bp 的开放阅读框架, 编码产生由 278 个氨基酸组成的分子量为 34kD 的蛋白。将目的基因的调控序列进行缺失突变后, 克隆到原核表达载体 pBL 上获得表达质粒 pBL4, 通过温度诱导, 经 SDS-PAGE 分析有明显的表达条带, 约占菌体总蛋白的 20%, 并且表达的蛋白具有较高的酶活力。构建的优良基因工程菌为最终实现从葡萄糖一步发酵产生维生素 C 前体 2-酮基-L-古龙酸打下了基础。

关键词 2,5-DKG 还原酶 I, 棒状杆菌, PCR, 克隆, 表达

分类号 Q554

2,5-二酮基-D-葡萄糖酸 (简称 2,5-DKG) 还原酶能够选择性和立体特异性地还原 2,5-DKG 的 C-5 位酮基, 生成维生素 C 重要前体 2-酮基-L-古龙酸 (简称 2-KLG)。它广泛分布于 17 个不同的细菌属, 其中酶活力最高的菌株是在棒状杆菌属中 (*Corynebacterium* sp.)^[1]。在某些棒状杆菌中能同时纯化得到两种 2,5-DKG 还原酶, 它们在分子量和催化能力方面有比较明显的差别^[2]。

棒状杆菌 SCB3058 是我们实验室经诱变筛选得到的一株酶活力较高的菌株, 已经成功地应用于从葡萄糖串联发酵产生 2-KLG 的研究^[3], 葡萄糖经欧文氏菌 (*Erwinia* sp.) SCB125 发酵生成重要的中间体 2,5-DKG, 再由棒状杆菌把 2,5-DKG 还原为 2-KLG。利用重组 DNA 技术, 把棒状杆菌中的 2,5-DKG 还原酶基因克隆表达达到欧文氏菌中, 构建一个基因工程菌, 将实现葡萄糖一步发酵产生 2-KLG。

Anderson 等人首先采用筛选基因文库的方法成功地从棒状杆菌克隆到了一个 2,5-DKG 还原酶基因, 并且构建了利用葡萄糖直接发酵产生 2-KLG 的基因工程欧文氏菌^[4]。此后另一个 2,5-DKG 还原酶基因也被 Hardy 等人用类似的方法从棒状杆菌克隆得到^[5]。而采用 PCR 方法克隆 2,5-DKG 还原酶基因尚未见报道, 关于构建利用葡萄糖直接发酵产生 2-KLG 的基因工程菌这方面的研究至今在国内尚未见报道。

* 本研究得到法国贡比涅科技大学科研合作经费资助; 中国自然科学基金资助项目。

收稿日期: 1997-07-14

1 材料和方法

1.1 菌株和质粒

棒状杆菌 (*Corynebacterium* sp.) SCB3058 和欧文氏菌 (*Erwinia* sp.) SCB125 由本实验室诱变筛选得到。E. coli DH5 α 和 pBL 表达载体由本实验室保存。pGEM3Zf(+) 为 Promega 公司产品。

1.2 培养基

BPY 培养基^[6]用于棒状杆菌培养, LB 培养基^[6]为大肠杆菌完全培养基, 用前根据需要加入 100 μ g / ml 氨苄青霉素 (Amp)。

1.3 酶与试剂

PCR 引物由中国科学院植物生理研究所合成, Taq DNA Polymerase、所有内切酶及连接酶、PCR 产物纯化试剂盒为 Promega 公司产品。蛋白酶 K 为 Merck 公司产品。RNase A 为华美生物工程公司产品。溶菌酶为 Serva 公司产品。123bp Ladder Marker 为 Life Technologies 公司产品。NADPH 购自 Boehringer 公司, 2, 5-DKG 从葡萄糖发酵制备^[3]。所用的化学试剂均为分析纯。

1.4 棒状杆菌 2, 5-DKG 还原酶提取和纯化

参照文献 [2] 进行。

1.5 棒状杆菌基因组 DNA 的提取方法

参考文献 [7] 并略加改进。

1.6 引物合成与 PCR 扩增

根据文献 [4] 报道的 2, 5-DKG 还原酶基因全序列, 用 PC / GENE 软件中 PCR 引物设计程序, 设计了 PCR 反应的引物, 合成时在 5' 端加上酶切位点和保护碱基可以方便地进行克隆:

5' 端引物: 5' TCTGAATTCGCGCCTACCCTGGAAGACATGACAG 3'
EcoRI

3' 端引物: 5' TCTAAGCTTATCCTCGAAGCTCTCCCGTGCCATC 3'
HindIII

5' 端突变引物: 5' CGCCTACCCTGGAAATTCATGACAG 3'
EcoRI

以棒状杆菌基因组 DNA 为模板, 在 50 μ l PCR 反应混合液含 10 \times Buffer(Mg²⁺ free) 5 μ l, 5mol / L dNTP 2 μ l, 20pmol / μ l 引物各 2 μ l, 模板 1 μ l, Taq DNA Polymerase 0.5 μ l, 根据需要加入 25mmol / L MgCl₂ 和甘油等若干, 用无菌水调整至 50 μ l。循环参数: 95 $^{\circ}$ C 变性 1 min; 50 $^{\circ}$ C 复性 1 min; 72 $^{\circ}$ C 延伸 3 min; 35 个循环, 72 $^{\circ}$ C 10 min 使产物延伸完全。混合物取 5 μ l 用琼脂糖凝胶电泳检测。扩增产物通过 Wizard PCR 回收试剂盒纯化。

1.7 PCR 扩增产物的克隆与鉴定^[7]

PCR 扩增产物经纯化后, 用 EcoRI 和 HindIII 酶切, 定向克隆到 pGEM3Zf(+) 中。通过蓝白斑变化和小规模抽质粒鉴定筛选阳性菌落。所得质粒 pGEM813 通过酶切图谱鉴定后再进行核酸序列分析。

1.8 序列分析

采用 Sanger 双脱氧末端终止法,测序采用荧光掺入自动测序。

1.9 大肠杆菌表达分析

1.9.1 表达载体构建: 用 EcoRI 和 HindIII 从阳性重组质粒 pGEM813 上切下 2,5-DKG 还原酶 I 基因片段,插入到原核表达载体 pBL 上的 EcoRI 和 HindIII 位点间,将连接产物转化大肠杆菌 DH5 α ,在 AmpLB 平板上筛选得到重组表达载体,酶切鉴定,命名为 pBL4。

1.9.2 2,5-DKG 还原酶 I 蛋白表达: 挑取含 pBL4 的 DH5 α 菌接种于 AmpLB 液体培养基中,30℃ 培养过夜,次日以 1% 接种新鲜 AmpLB 培养基中,30℃ 培养 3~4 h (A_{600} 约 0.4),升温至 42℃ 诱导继续培养 4 h。

1.9.3 SDS-PAGE 分析^[7]: 取 1ml 菌液离心收集菌体,悬浮于 50 μ l 无菌水,加等体积 2 \times 上样缓冲液,煮沸 10 min,离心后取上清进行 SDS-PAGE (分离胶 10%) 分析。用考马斯亮蓝 R250 染色,灰度扫描测定表达产物含量。

1.9.4 2,5-DKG 还原酶 I 比活测定方法^[2]: 离心收集表达的菌体,用 Buf A 0.1mol/L Tris-Cl (pH7.0) 洗涤,重悬浮于一定体积的 Buf A 中,用超声波破碎菌体,离心取上清作为测定用的酶液。在 30℃,100 μ l 酶液加到 2.7ml Buf A 中,再加 100 μ l 浓度为 3~5 mmol/L 的 NADPH 溶液及 100 μ l 过量底物 2,5-DKG,测定 A_{340} 的变化,根据标准曲线换算成 NADPH 浓度变化,计算酶活力。酶活力单位 (U) 定义为:每分钟氧化 1 μ mol NADPH 所需的酶量。蛋白浓度采用考马斯亮蓝染色法以牛血清白蛋白为标准测定^[8],计算酶的比活力。

2 结果

2.1 2,5-DKG 还原酶提取和纯化

棒状杆菌经添加 2,5-DKG 诱导培养后,离心收集菌体并破碎,经硫酸铵沉淀和透析纯化得到粗酶液,测定具有 2,5-DKG 还原酶活力,酶蛋白经 SDS-PAGE 分析,有两个主要条带分子量约为 34kD 和 27kD。

2.2 模板制备和 PCR 扩增

本文采用的染色体 DNA 提取方法,是由哺乳动物细胞染色体 DNA 提取方法改进而来。对提取到的棒状杆菌 DNA 电泳检测 (0.4% Agarose),所抽提到的 DNA 分子量为 23kb 左右。

以棒状杆菌 SCB3058 全基因组 DNA 为模板,按所设计的优化反应条件进行 PCR 扩增。扩增产物经 1.5% 琼脂糖凝胶电泳检查为单一条带,长约 1.1kb,大小与预期值完全一致。为了鉴定 PCR 反应的产物是否正确,从产物的酶切结果方面进行了初步判别, BamHI、XhoI 等内切酶消化位点与理论值符合。

2.3 扩增产物的克隆和序列分析

纯化后的 PCR 扩增产物用两端设计的内切酶消化成粘性末端,利用质粒 pGEM3Zf(+) 多克隆位点上酶切位点 EcoRI 和 HindIII,将产物定向克隆到该载体上。筛选得到多个重组质粒。增大的重组克隆用 EcoRI、HindIII 和基因内部酶切位点进行酶切初步鉴定。

直接以重组质粒 pGEM813 双链 DNA 为模板,采用 M13 正向通用引物和中间合成的

引物测出两端的序列,中间部分采用 SalI-PstI 亚克隆后测序,结果全序列与发表的 2,5-DKG还原酶 I 基因顺序进行比较基本相同。从而证明 PCR 方法扩增基因的有效性。但是在基因编码区 434 位是 G 而不是 A,734 位是 C 而不是 T。从而导致了两个氨基酸差异。

2.4 表达载体的构建和在大肠杆菌中的表达分析

将 2,5-DKG还原酶 I 基因直接从 EcoRI 和 HindIII 位点克隆到原核表达载体 pBL 中,利用棒状杆菌的翻译信号进行翻译,构建成大肠杆菌表达载体 pBL1028,经过 SDS-PAGE 分析未能实现高效表达。推测为 pBL 上的 SD 序列与起始密码子 ATG 之间距离太长(26bp),大大影响了翻译效率。因而合成了新的 5' 端突变引物,在 ATG 前面突变两个碱基,构建出一个新的 EcoRI 酶切位点,序列分析证实了产生的新片段缺失掉了棒状杆菌 2,5-DKG还原酶 I 基因原来的 5' 端调控序列,全部用表达载体 pBL 上的 P1 启动子、SD 序列等调控基因表达,SD 序列与 ATG 起始密码子之间间距变为 8bp,接近文献报道的最佳距离^[9]。构建出大肠杆菌高效报道载体 pBL4(图 1)。

含有重组表达载体 pBL4 的大肠杆菌 DH5 α 经过 42℃ 热诱导,产生了对照(含有质粒 pBLpho4*)所没有的明显表达产物,在 SDS—PAGE 凝胶上可见明显的表达产物条带,测得分子量约为 34kD,与理论值完全一致,灰度扫描结果表明,表达的重组蛋白约占菌体总蛋白的 20%(图 2)。

按酶活力测定方法测定表达在大肠杆菌中的 2,5-DKG 还原酶 I 活力,结果表明表达的 2,5-DKG还原酶 I 具有很高的生物学活性,达到了 1660U / mg,比直接从棒状杆菌纯化

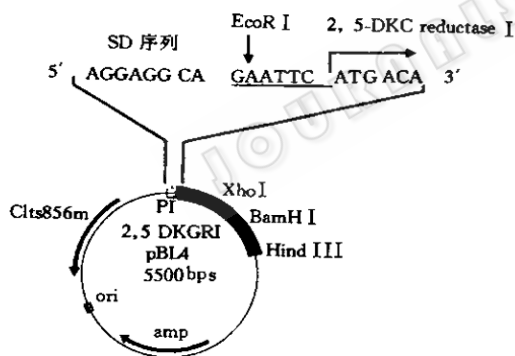


图 1 pBL4 物理图和 2,5-DKG 还原酶 I 基因 N-端部分核酸序列

Fig.1 Physical map of pBL4 and N-terminal partial nucleic acid sequence of the 2,5-DKG reductase I gene

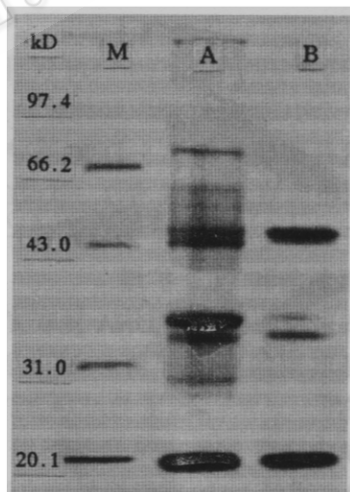


图 2 2,5-DKG 还原酶 I 基因编码的 34kD 蛋白在大肠杆菌中表达的 SDS-PAGE 分析

Fig.2 SDS-PAGE of the 2,5-DKG reductase I gene coded 34kD protein expression in *E. coli*
M: Molecular weight standard

A. DH5 α (pBL4) induced at 42℃ for 4 h;
B. DH5 α (pBLpho4) induced at 42℃ for 4 h.

* 因为没有单纯的 pBL 载体,选择克隆表达了酸性磷酸酯酶的 pBLpho4 作为对照。

了28倍后的比活力 $3\text{U}/\text{mg}^{[2]}$ 提高了近550倍。通过不同的温度诱导试验,发现 42°C 诱导能够得到高表达量,但表达的酶蛋白主要存在于沉淀中,上清液中酶的比活力和 37°C 诱导保持基本相同的水平。

3 讨论

在棒状杆菌SCB3058成功应用于葡萄糖串联发酵以及从中提取和纯化得到2,5-DKG还原酶的基础上,我们首先克隆2,5-DKG还原酶I基因,它的长度比较适中,利用PCR技术扩增简洁、高效,用PCR扩增克隆棒状杆菌基因很少有报道,棒状杆菌SCB3058基因组DNA的G+C%含量高达71%,复杂的二级结构使其产生较大的碱基堆积力,这给PCR扩增和序列分析带来了一些困难。我们从反应的温度、时间、镁离子浓度及反应促进剂等多方面摸索到PCR的最佳反应条件^[10],成功地扩增了2,5-DKG还原酶I基因。同时在设计PCR引物时引入合适的酶切位点,使扩增产物定向克隆十分方便。

2,5-DKG还原酶I是一个由278个氨基酸组成的蛋白,其基因编码区长834bp,在基因编码区碱基变化导致了蛋白第145位和第245位两个氨基酸差异,这可能是由菌种传代中的自发突变引起,但也可能由于是Taq DNA聚合酶缺乏3'至5'校正活性,在PCR扩增中错误掺入的结果。所表达的蛋白具有较高的酶活力表明了这两个氨基酸变化没有对酶的活性中心产生明显的影响。

外源基因能否在大肠杆菌中正确、高效地表达为有活性的蛋白,关键是要有合适的表达系统使外源基因的转录和翻译能顺利进行,表达的蛋白能正确折叠^[9]。表达载体pBL包括 λ 噬菌体阻遏蛋白温度敏感性CI_{ts}857m基因、P1启动子、人工合成的SD序列和多克隆位点,pBL4的SD序列到起始密码子ATG之间距离为8bp。本文的实验结果说明选择合适的表达载体和相应的宿主,启动子、SD序列、SD序列与起始密码子之间的距离等因素的重要性,尤其是SD序列与起始密码子之间的距离特别明显影响基因的表达量。由于大肠杆菌缺乏翻译后修饰,很多蛋白表达后形成没有活性的包含体,2,5-DKG还原酶I来源于原核生物,有活性的酶蛋白是由单体组成,分子内又没有二硫键,这对高效表达的蛋白正确折叠成为有活性的分子比较有利,但本文的实验结果表明: 42°C 诱导的大量表达产物仍旧形成了包含体,胞质中的可溶性蛋白可能已经达到饱和。 30°C 生长也能达到相同酶活水平,说明突变的CI_{ts}857蛋白抑制P1启动子转录不是十分严谨,这对我们将该基因表达在欧文氏菌中十分有利,因为欧文氏菌SCB125正常发酵温度接近于 30°C ,如果2,5-DKG还原酶在欧文氏菌SCB125中组成型表达对菌生长、发酵产酸没有不利影响,这将是比较合适的载体。

大肠杆菌是最为成熟的原核表达系统,有很多载体和宿主细胞可供选择,既可以高效地表达外源蛋白,还可以有效地加以调控。2,5-DKG还原酶I基因预先在大肠杆菌DH5 α 中实现高效表达,了解其表达规律对下一步在欧文氏菌中成功表达至关重要,目前对棒状杆菌SCB3058中2,5-DKG还原酶II基因的克隆与表达也在研究中。

致谢 本工作承杨胜利研究员提出宝贵意见和给予很大支持,董文玲、邓貽默和徐炜等同志参加早期部分工作,叶晴、宋祺同志提供实验用菌种及进行2,5-DKG还原酶的纯化,序

列分析由中国科学院植物生理研究所 Beckman 生物技术示范实验室完成,在此一并致谢。

参 考 文 献

- [1] Sonoyama T, Kageyama B, Yage S. *Agric Biol Chem*, 1987, **51**(7):2003~2004.
- [2] Sonoyama T, Kobayashi K. *Ferment Technol*, 1987, **65**(3):311~317.
- [3] 尹光琳,马志方,董文玲等. *微生物学报*, 1991, **31**(3):198~205.
- [4] Anderson S, Marks C, Lazarus R *et al*. *Science*, 1985, **230**:144~149.
- [5] Grindley J F, Payton M A, Hardy K G. *Appl Environ Microbiol*, 1988, **54**:1770~1775.
- [6] 中国微生物菌种保藏管理委员会编. *中国菌种目录*. 北京:轻工业出版社,1983. 408.
- [7] Sambrook J, Fritsch E F, Maniatis T. *Molecular Cloning. A Laboratory Manual*. 2nd ed. New York: Cold Spring Harbor Laboratory Press, 1989.
- [8] 李建武,陈丽蓉,陈雅惠等. *生物化学实验原理和方法*. 北京:北京大学出版社,1994. 174~176.
- [9] David V L. *Methods in Enzymology*, 1990, **185**:3~195.
- [10] 林万明,杨瑞霞,黄尚志等. *PCR 技术操作和应用指南*. 北京:人民军医出版社,1993. 1~62.

CLONING AND EXPRESSION IN *E. COLI* OF 2,5-DKG REDUCTASE I FROM *CORYNEBACTERIUM*

Chen Ceshi Yin Guanglin

(Shanghai Research Centre of Biotechnology, Chinese Academy of Sciences, Shanghai 200233)

Abstract Two different 2,5-DKG reductases was purified from *Corynebacterium* sp. SCB3058, then its genomic DNA was used as template, a segment containing 2,5-DKG reductase I was amplified by PCR and cloned into pGEM3zf(+) to obtain the recombinant plasmid pGEM813. The sequence analysis showed that the cloned segment was 1107bp in length, contained a single open reading frame of 834 nucleotides, which encoded a 34kD protein consisted of 278 amino acids. After the primary control sequence was deleted, the expression vector pBL4 was constructed by plasmid pBL. With induction of temperature, a 34kD protein was specifically expressed in *E. coli* DH5 α in high level. The expressed recombinant protein accounted for 20% of the total cell protein and had high specific enzyme activity. In spite of 30°C, 37°C or 42°C induction, the specific activity of enzyme was almost the same level. These results indicate that most of the recombinant protein induced at 42°C was existed with inclusion bodies. This work was the bases of constructing a recombinant *Erwinia* which can produce 2-KLG directly from D-glucose by one-step fermentation.

Key words 2,5-DKG reductase I, *Corynebacterium*, PCR, Cloning, Expression