

肌苷和鸟苷生产菌中嘌呤核苷合成途径三段基因 序列的分析*

钱江潮 蔡显鹏 储 炬 庄英萍 张嗣良**

(华东理工大学生物反应器工程国家重点实验室 上海 200237)

姚泉洪 彭日荷

(上海市农业科学院生物中心 上海 201106)

摘 要: 为了研究肌苷和鸟苷生产菌中与产苷有关的嘌呤核苷合成途径的遗传背景, 选择了 *pur* 操纵子的启动子序列、编码 SAMP 合成酶的 *purA* 基因和编码 GMP 合成酶的 *guaA* 基因, 设计合适的引物, 分别从野生菌、一株肌苷低产菌和肌苷鸟苷高产菌中扩增出相应片段, 经克隆和测序后, 对它们进行比较和分析。分析结果表明两株生产菌的 *purA* 基因发生了 1 个碱基缺失, 导致阅读框发生移码突变, 而鸟苷高产菌在 *pur* 操纵子的启动子部分和操纵子抑制蛋白结合区域发生了近 10% 的突变, 可能影响整个操纵子的表达调控。

关键词: 肌苷 鸟苷 生物合成 基因序列分析

中图分类号: Q933 文献标识码: A 文章编号: 1001-6209(2003)02-0200-06

肌苷和鸟苷是两种多用途的嘌呤核苷, 它们分别广泛应用于合成食品增鲜剂 IMP 和 GMP。此外, 肌苷还可用于心脏病、肝病、白血球减少症、血小板减少症、视神经萎缩和中心性视网膜炎等疾病的治疗, 能预防及解除由血防药物引起的对心脏或肝脏的副作用, 鸟苷则还可作为三氮唑核苷、无环鸟苷等许多抗病毒药物的合成原料。因此, 有关它们的生产从六十年代以来就开展了大量的研究工作^[1, 2]。

在生产菌的选育和改造方面, 日本味之素公司和武田制药厂都利用磷酸单酯酶活力很强的枯草芽孢杆菌或短小芽孢杆菌作为诱变出发菌株, 采用物理或化学诱变方法, 选育出产苷水平较高的生产菌株。六十年代末至八十年代初, 针对细菌体内核苷酸合成途径和调节机理开展了大量研究, 并完成了嘌呤核苷生物合成途径操纵子的基因克隆和全序列测定^[3], 研究的重点随之转移到利用基因工程的方法构建重组菌^[4, 5]。在发酵工艺研究方面, 通过发酵过程中多参数的相关分析, 实时分析和调控发酵过程的代谢流迁移, 在鸟苷发酵中取得了突破性进展, 将产苷水平稳定在 30g/L^[6]。为了进一步提高生产水平, 并在基因水平了解生产菌株的特性, 对生产菌中嘌呤核苷合成途径的遗传背景进行研究, 为基因改造提供理性指导。

为此, 我们采用野生型和一株肌苷低产菌作为对照, 选择 *pur* 操纵子(编码嘌呤生物合成途径中的 12 个蛋白)的启动子序列, *purA* 基因(编码 SAMP 合成酶), *guaA* 基因(编码 GMP 合成酶), 分别进行比较。结果表明, 肌苷和鸟苷生产菌的 *purA* 基因发生了碱基缺

* 上海市曙光计划资助项目(01SG28)

** 通讯作者。E-mail: siliangz@163.net

作者简介: 钱江潮(1971-), 女, 浙江人, 博士, 讲师, 主要研究生物化学和分子生物学。E-mail: jicqian@yahoo.com

收稿日期: 2002-06-27, 修回日期: 2002-10-22

失,它们所编码的 SAMP 合成酶应已失活。与其它菌株相比,鸟苷生产菌的 *pur* 操纵子在启动子部分发生了多处突变,可能造成表达调控方式的改变。研究从分子水平揭示了生产菌的部分遗传特性,有助于在基因水平理解产物在生产菌中的积累。

1 材料和方法

1.1 菌种质粒和培养基

枯草芽孢杆菌野生型菌株 160 购自上海工业微生物研究所,肌苷低产型菌株 161 由野生菌株诱变得来,肌苷高产菌株 162 由肌苷低产菌株诱变所得,鸟苷高产菌株 163 由肌苷高产菌株诱变筛选得到。枯草杆菌培养基含葡萄糖、酵母粉、 $(\text{NH}_4)_2\text{SO}_4$ 、 KH_2PO_4 、KCl 等。大肠杆菌 DH5 α 质粒 pBluscriptSK。

1.2 工具酶和化学试剂

Ex-Taq DNA 聚合酶,限制性内切酶和其他基因工程工具酶为日本 TaKaRa 公司产品,其他试剂为国产 AR 级试剂。

1.3 PCR 扩增

根据 GenBank 公布的枯草杆菌基因组序列,分别设计 3 对引物:*pur* 操纵子启动子区域,以第一个基因 *purE* 的翻译起始密码子为 1,上游引物 N291(-438 ~ -414),GATATCT-GCCTGTAAGGCCAATGGA,下游引物 N292(58 ~ 82)GTTCGTCAGATATGTCGCATGCCGTG,扩增片段全长 520bp;*purA* 基因,以翻译起始密码子为 1,上游引物 N194(-66 ~ -40),GAGATCTTCACTGATAAACTTGATTTGTTTG,下游引物 N195(1290 ~ 1302),GCATATGCTG-CAGACGATATTCTATTTAG,扩增片段全长 1388bp;*guaA* 基因,以翻译起始密码子为 1,上游引物 N293(-26 ~ -1),AACTTAGAGGAGATTAGGTGACAACC,下游引物 N294(1547 ~ 1574),TTGCCAAAAAGATGGTTTCCATTAATTG,扩增片段全长 1600bp。在 50 μL 的反应体系中,加入 1 μL 抽提的枯草杆菌 DNA 模板(约 100ng),1 μL 上游引物和 1 μL 下游引物(各 100ng),10 \times 缓冲液 5 μL ,dNTP4 μL (各 200nmol/L),*Taq* 酶 2.5U。PCR 反应条件:*pur* 操纵子启动子区域和 *guaA* 基因均是 94 $^{\circ}\text{C}$ 变性 30s,64 $^{\circ}\text{C}$ 退火 30s,72 $^{\circ}\text{C}$ 反应 2min,进行 30 个循环;*purA* 基因的扩增条件是 94 $^{\circ}\text{C}$ 变性 30s,56 $^{\circ}\text{C}$ 退火 30s,72 $^{\circ}\text{C}$ 反应 2min,进行 30 个循环。PCR 反应结束后,产物经电泳后回收。

1.4 DNA 操作

枯草杆菌总 DNA 抽提按文献 [7] 方法进行,其他的 DNA 操作按标准的分子克隆手册进行,PCR 产物经 T/A 克隆后,制备的双链质粒 DNA 在 ABI377 型自动测序仪上进行序列测定。每个菌株挑选两个转化子,制备测序模板。对于 *pur* 操纵子的启动子序列(520bp),分别从正、反两向测序。*purA* 基因(1388bp)分 3 段测序,前两段正向测,最后一段反向测,3 段序列均有重叠部分,*guaA* 基因较长,有 1600bp,分 4 段测定,前 3 段正向测,最后一段反向测,4 个片段均有重叠部分,经拼接后得到最后测定结果。

2 结果

2.1 *pur* 操纵子 *cis* 序列测定和分析

比较野生型菌株(160)、肌苷低产菌株(161)、肌苷高产菌株(162)和鸟苷高产菌株

(163)的 *pur* 操纵子启动子序列,163 菌株在该区域发生了较多突变,与 160 相比,有 51 个突变位点(以 *purE* 基因翻译起始密码子为 1 计):-407,-404,-399,-398,-392,-389,-380,-371,-370,-362,-354,-321,-309,-308,-303,-294,-279,-262,-181,-180,-172,-171,-147,-146,-144,-142,-141,-140,-127,-126,-106,-99,-98,-96,-95,-65,-57,-55,-54,-53,-52,-47,-46,-40,-39,-38,-37,-27,-26,-18,-7)。161 和 162 分别有一个核苷酸发生了突变,161 在 -78 处,162 在 -272 位。与文献报道的枯草杆菌 168 菌株相比较,野生型 160 与 168 的序列一致。

2.2 *purA* 基因

以起始密码子为 1 计,比较 4 个菌株在基因编码区发生的核苷酸序列变化。与野生菌(160)相比,肌苷高产菌(162)和鸟苷高产菌(163)在第 55 位发生了一个碱基(A)的缺失,这将导致后面所有的密码子发生移码突变,可能使它所编码的 SAMP 合成酶失活,进而导致肌苷和鸟苷积累,产量提高。而肌苷低产菌 161 只有一个位点的核苷酸发生了突变,1003 位的核苷酸由野生型的 G 变成了 A,导致密码子由 GCA 变成了 ACA,所编码的氨基酸由丙氨酸变成了苏氨酸。

与文献报道的 168 菌株 *purA* 基因序列相比,野生型菌株 160 在 6 个位点发生了不引起氨基酸改变的碱基变化,但在 734~772 位点间,所测 4 个菌株与 168 相比有较大差异(暂时不统计 162 和 163 发生的移码突变),在 734 处增加了一个碱基(G),而在 772 处又缺失了一个碱基(A),中间发生了两个碱基的互置,核苷酸序列由文献报道的 G-CGGTGTCACAGTCGGTTCTGGTGTGCGCCCGACCAAAAATC 变成了所测的 GCGGTGTCACGATCGGT-TCTGGTGTGCGCCCGACCAAAA-TC,所编码的氨基酸由 Ala Val Ser Gln Ser Val Leu Val Ser Ala Arg Pro Lys Ile 变成了 Gly Gly Val Thr Ile Gly Ser Gly Val Gly Pro Thr Lys Ile,因此导致这一区域所编码的氨基酸有 12 个不同。此外,在 911~912 和 1033~1034 位点,4 个菌株也与 168 有差异,都是由 168 的 CG 变成了 GC,各导致一个氨基酸的变化,脯氨酸变为精氨酸,精氨酸变为丙氨酸。这两处差异可能是由于菌株的不同所引起,也可能是 168 菌株测序误差造成的。

2.3 *guaA* 基因

对 4 株菌的 *guaA* 基因序列进行分析,与野生菌 160 相比,肌苷低产菌在 739 位(以起始密码子为 1 计)发生了不引起氨基酸改变的碱基变化,鸟苷高产菌也在 105 位和 726 位发生了没有氨基酸变化的核苷酸变化,而肌苷高产菌 162 则与 160 序列一致。

与 *purA* 基因相似,所测的 4 个枯草杆菌菌株中,共有 23 个碱基与文献报道的菌株 168 的 *guaA* 序列不同,涉及的氨基酸变化有 12 处,详见表 1。但所有的氨基酸变化均同时发生在高产菌和低产及野生菌中,说明生产菌的 *guaA* 基因没有发生变异,GMP 合成酶的活性在 4 个菌株中应该是一致的,它应该不是导致生产菌高产苷的直接因素或低产的限制因素。

3 讨 论

在枯草杆菌中,嘌呤核苷酸的合成,从 PRPP(5-磷酸核糖焦磷酸)起,经 11 个酶催化 10 步反应,生成 IMP(次黄苷酸),再经两个支路,分别生成鸟苷酸和腺苷酸(图 1)。

表 1 *guaA* 基因突变统计结果(与文献报道的 168 菌株序列进行比较)

Table 1 List of mutations in *guaA* gene of *B. subtilis* strain 160 compared with that of strain 168

Position	Strain 168	Strain160	Amino Acid Change
308 309)	CG(ACG)	GC(AGC)	Thr→Ser
448 9 450)	GC(GCG)	-	Ala deletion
488 489)	GC(GGC)	CG(GCG)	Gly→Ala
502)	C(CAC)	T(TAT)	His→Tyr
504)	C(CAC)	T(TAT)	
587 588)	CG(TCG)	GC(TGC)	Ser→Cys
646 647)	CA(CAG)	AC(ACG)	Gln→Thr
674 675)	GC(GGC)	CG(GCG)	Gly→Ala
784 785)	CG(CGT)	GC(GCT)	Arg→Ala
1378)	-	G	
1382)	-	C	
1389)	-	T	
1378 - 1389	-GAAT-CGCGCCGT-	GGAATCCGCGCCGTT	Glu Ser Arg Arg→Gly Ile Arg Ala Val

在这一合成途径中,编码 PRPP 生成 IMP 所需的 11 个酶都位于包含了 12 个基因的多顺反子 *pur* 操纵子上^[3]。该操纵子的转录起始于第一个结构基因 *purE* 上游 242nt 处,启动子为 σ 依赖型,腺苷和鸟苷分别调控转录的起始和提前终止^[3,8]。去除启动子的 -193 ~ -37 (以转录起始为 1 计)区域可导致腺嘌呤抑制作用的解除,而 -145 ~ -29 区域则是 *pur* 操纵子抑制蛋白 (PurR, *pur* Operon Repressor) 的结合位点^[9]。对抑制蛋白 PurR 的研究发现,它的分子量为 62kD,具有与 DNA 结合的 N 端区和一个 PRPP 结合位点^[10],它与 DNA 的结合不受腺嘌呤、腺苷或腺苷酸的影响^[9],但 PRPP 可抑制它们的结合^[10]。Weng 等提出了一个调控模型^[10],在过量腺嘌呤存在时,它可转化为腺苷酸,消耗 PRPP。同时,当 ATP 浓度升高时,又将加强 ADP 对 PRPP 合成酶的抑制作用。在这两种过程的作用下,PRPP 浓度降低,使得 PurR 与操纵子调控区域的结合增强,最终导致转录下降。进一步对 PurR 识别位点的研究^[11],发现它在体外还能识别 *purA*、*purR* (自身)和 *pyr* 操纵子,但并未在这些 DNA 片段中发现特定的蛋白识别位点,只在 *pur* 操纵子中发现了一段 GAAC-N₂₄-GTTC 的必需序列。

根据以上调控模型,再对所测菌株 *pur* 操纵子启动子中的 PurR 结合位点进行分析 (图 2)。

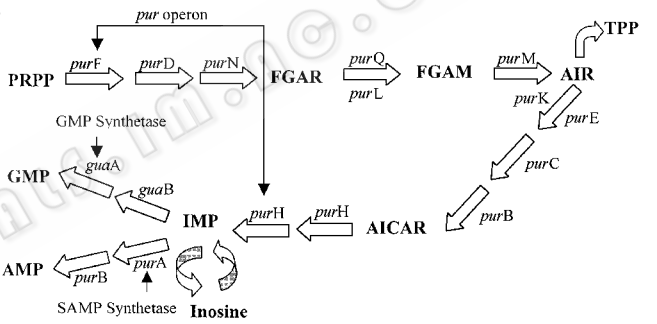


图 1 嘌呤核苷生物合成示意图

Fig. 1 Biosynthesis of purine nucleotides in *B. subtilis*

PRPP. Phosphoribosylpyrophosphate;FGAR. Formylglycinamide ribonucleotide;FGAM. Formylglycinamide ribonucleotide;AIR. Aminoimidazole ribonucleotide;AICAR. Aminoimidazole-carboxamide ribonucleotide;IMP. Inosine monophosphate;GMP. Guanosine monophosphate;AMP. Adenosine monophosphate.

- [2] Matsui H , Sato K , Enei H , *et al.* Production of guanosine by sulfaguanidine resistant mutants of *Bacillus subtilis*. *Agr Biol Chem* , 1978 , **42** (3) : 637 ~ 641 .
- [3] Ebbole D J , Zalkin H. Cloning and characterization of a 12-gene cluster from *Bacillus subtilis* encoding nine enzymes for *de Novo* purine nucleotide synthesis. *J Biol Chem* , 1987 , **262** (17) : 8274 ~ 8287 .
- [4] Miyagawa K , Kimura H , Nakahama K , *et al.* Cloning of the *Bacillus subtilis* IMP dehydrogenase gene and its application to increased production. *Bio/Technology* , 1986 , **4** : 225 ~ 228 .
- [5] Miyagawa K , Doi M , Akiyana S. Method of producing inosine and/or guanosine US Patent 4 , 701 413 , 1987 .
- [6] 蔡显鹏 , 陈双喜 , 张嗣良 , 等 . 基于过程参数相关分析的鸟苷发酵过程优化 . 微生物学报 , 2002 , **42** (2) : 232 ~ 235 .
- [7] Chen W P , Kuo T T. A simple and rapid method for the preparation of gram-negative bacterial genomic DNA. *Nucl Acids Res* , 1993 , **21** (9) : 2260 .
- [8] Ebbole D J , Zalkin H. *Bacillus subtilis* pur operon expression and regulation. *J Bacteriol* , 1989 , **171** (4) : 2136 ~ 2141 .
- [9] Ebbole D J , Zalkin H. Interaction of putative repressor protein with an extended control region of the *Bacillus subtilis* pur operon. *J Biol Chem* , 1989 , **264** (6) : 3553 ~ 3561 .
- [10] Weng M , Nagy P L , Zalkin H. Identification of the *Bacillus subtilis* pur operon repressor. *Proc Natl Acad Sci USA* , 1995 , **92** (16) : 7455 ~ 7459 .
- [11] Shin B S , Stein A , Zalkin H. Interaction of *Bacillus subtilis* purine repressor with DNA. *J Bacteriol* , 1997 , **179** (23) : 7394 ~ 7402 .

Analysis of Three Nucleotide Sequences Involved in the Purine Nucleotides Biosynthesis in Inosine and Guanosine-producing *Bacillus subtilis* *

Qian Jiangchao Cai Xianpeng Chu Ju Zhuang Yingping Zhang Siliang**

(State Key Laboratory of Bioreactor Engineering , ECUST , Shanghai 200237 , China)

Yao Quanhong Peng Rihe

(Agro-biotechnology Research Center , Shanghai Academy of Agricultural Sciences , Shanghai 201106 , China)

Abstract : In order to study the genetic background involved in the purine nucleotides biosynthesis in the inosine and guanosine-producing *Bacillus subtilis* , the promoter domain of the pur operon , *purA* gene encoding SAMP synthetase , and *guaA* gene encoding GMP synthetase , were amplified by PCR from a wild-type strain , a low-yield inosine-producing strain , a inosine-producing strain and a guanosine-producing strain . After cloning and sequencing of PCR products , the nucleotide sequences from four strains were aligned with the reported corresponding sequences . A one-base deletion is discovered in *purA* gene from inosine-producing strain and guanosine-producing strain , which will cause the open-reading-frame-shift mutation . And nearly 10% base mutations in the entire and pur operon repressor binding domain of pur operon are observed in guanosine-producing strain , which may affect the expression regulation mode of the entire operon .

Key words : Inosine , Guanosine , Biosynthesis , Nucleotide sequence analysis

* Shanghai Shuguang Plan (01SG28)

** Corresponding author . E-mail : siliangz@163.net