

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200580044130.4

[51] Int. Cl.

C12N 15/52 (2006.01)

C07K 14/34 (2006.01)

C12R 1/15 (2006.01)

[43] 公开日 2007 年 12 月 5 日

[11] 公开号 CN 101084312A

[22] 申请日 2005.12.13

[21] 申请号 200580044130.4

[30] 优先权

[32] 2004.12.22 [33] DE [31] 102004061696.5

[86] 国际申请 PCT/EP2005/013372 2005.12.13

[87] 国际公布 WO2006/066758 英 2006.6.29

[85] 进入国家阶段日期 2007.6.21

[71] 申请人 德古萨有限责任公司

地址 德国杜塞尔多夫市

[72] 发明人 S·汉斯 B·巴特 G·蒂尔巴赫

[74] 专利代理机构 永新专利商标代理有限公司

代理人 林晓红

权利要求书 1 页 说明书 20 页 序列表 11 页

[54] 发明名称

棒状细菌的 proB 基因的突变体

[57] 摘要

本发明涉及编码 γ -谷氨酰激酶的棒状细菌 proB 基因的突变体, 及涉及使用含有这种突变的细菌经发酵生产 L-脯氨酸的方法。

1.在第 149 位或类似氨基酸位置具有除了甘氨酸之外的蛋白质氨基酸的 γ -谷氨酰激酶。

2.权利要求 1 的 γ -谷氨酰激酶，特征在于在权利要求 1 所述的氨基酸位置或者类似位置存在选自如下一组的一个氨基酸：Lys、Asn、Arg、Ser、Thr、Ile、Met、Glu、Asp、Ala、Val、Gln、His、Pro、Leu、Tyr、Trp、Cys 或 Phe。

3. 编码权利要求 1 或 2 的 γ -谷氨酰激酶的核酸序列。

4.权利要求 3 的核酸序列，特征在于其：

a)与 SEQ. ID No:1 至少 70%相同，或者

b)编码权利要求 1 或 2 的 γ -谷氨酰激酶，其长度为 369 ± 40 个氨基酸，或者

c) 编码权利要求 1 或 2 的 γ -谷氨酰激酶，其具有 SEQ. ID No:2 的至少第 145—154 位或类似位置的氨基酸序列，或者

d) 编码权利要求 1 或 2 的 γ -谷氨酰激酶，其具有 SEQ. ID No:2 的至少第 145—154 位或类似位置的氨基酸序列，及在严格实验条件下与 SEQ. ID NO:1 或 SEQ. ID NO:3 互补的核苷酸序列杂交。

5.具有权利要求 3 或 4 的核苷酸序列的重组载体。

6.通过如下步骤生产 L-脯氨酸的方法：

a) 发酵表达或过表达权利要求 3 或 4 的核苷酸序列的宿主生物体，及

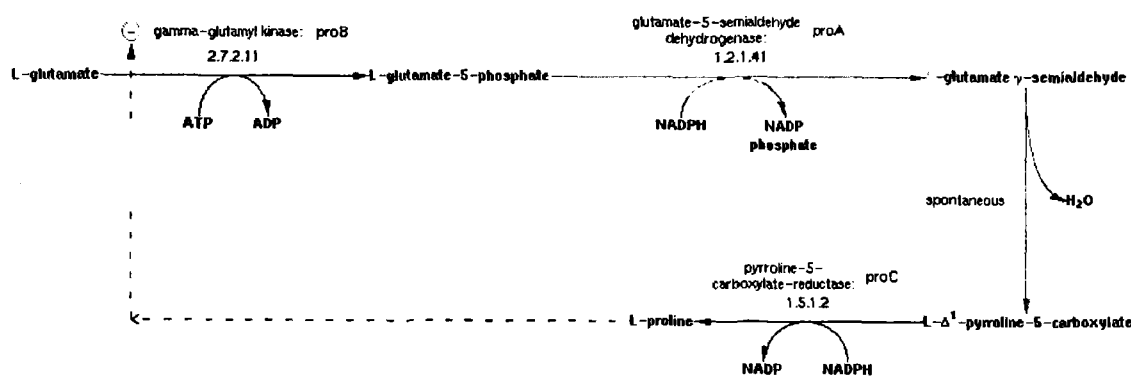
b) 分离或收集 L-脯氨酸，其中当合适时具有发酵肉汤的成分和/或生物量。

棒状细菌的 proB 基因的突变体

本发明涉及一种具有 γ -谷氨酰激酶活性的酶。更特别地，本发明鉴定了在氨基酸序列的第 149 位或类似氨基酸位置（a comparable position）具有除甘氨酸之外的蛋白质氨基酸（proteinogenic amino acid）的那些酶。

具有 γ -谷氨酰激酶活性的酶优选用于发酵生产 L-脯氨酸。从谷氨酸起始的 L-脯氨酸生物合成途径示于方案 1。

方案 1



已知可以通过发酵例如棒状细菌（coryneform bacteria）菌株、特别是谷氨酸棒杆菌菌株而生产氨基酸。由于其极其重要性，已持续进行改良生产方法的尝试。方法的改良可涉及发酵措施，如搅拌和供氧，或营养培养基的组分如发酵期间的糖浓度，或产物形式的加工，例如通过离子交换层析，或微生物本身的固有性质。

为改良这些微生物的性质，可使用诱变，选择及突变体选择等方法，以此方法可获得对抗代谢物有抗性或调节重要性代谢物是营养缺陷的并产生氨基酸的菌株。一种已知的抗代谢物是脯氨酸类似物 3,4-脱氢-DL-脯氨酸(DHP)。

一段时间以来,重组 DNA 方法也已经用于改良生产 L-氨基酸的棒杆菌菌株的改良,其通过扩增各个氨基酸生物合成基因,并研究对氨基酸生产的作用而进行改良。

谷氨酸棒杆菌基因组的核苷酸序列在 EP-A-1108790 中描述,并且储存在国立医学图书馆(Bethesda, MD, USA)的国立生物技术信息中心(NCBI)数据库中,登记号为 NC_003450.2 和 BX927148.1 至 BX927157.1。

Sleator 等报道了通过突变 *Listeria monocytogenes* proB-Gen 而引起脯氨酸生物合成过量生产的可能性(Appl. Environ. Microbiol. 2001, 67, 4560-5)。所述突变在该文中描述并称作编码 γ -谷氨酰激酶的 proB 基因,其具有如下突变: V121I、A144V 和 E146K。另外,需要提及的是发生这些突变的区域良好对应于也在其它生物体中作为有益突变的靶区域而被鉴别的区域。

因此本发明的目的是产生 γ -谷氨酰激酶的进一步突变的以及当合适时改良的蛋白质变体,它们可有利地用于 L-脯氨酸的发酵生产的技术方法中。

这一目的及其它未详细说明但从本领域可明显得出的目的通过权利要求 1 所述的 γ -谷氨酰激酶即可实现。权利要求 2 描述优选的酶。权利要求 3 和 4 分别涉及编码这些酶的核苷酸序列,而权利要求 5 描述重组产生的具有刚刚提及的核苷酸序列的载体。最后,权利要求 5 涉及本发明的借助于所述酶生产 L-脯氨酸的方法。

通过提供在第 149 位或类似氨基酸位置具有除甘氨酸之外的蛋白质氨基酸的 γ -谷氨酰激酶 (proB 基因的产物),令人惊奇地且有利地实现了所述目的。与野生型酶相对比,具有合适突变的 γ -谷氨酰激酶有助于在发酵生产过程中以改良的方式生产 L-脯氨酸。使用本发明的方法可以使得宿主生物体或者发酵方法的选自如下一组的一或多个参数的性能基于起始菌株或亲代菌株或者使用所述酶的发酵方法改良至少 0.5%、至少 1%、至少 1.5%或者至少 2%,所述的组为:产物

浓度(每体积的产物)、产物产量(形成的产物/消耗的碳源)及产物形成(每体积和时间形成的产物), 或者其它方法参数及其组合。

优选提供 γ -谷氨酰激酶, 其中在本发明定义的氨基酸位置或者类似位置存在选自如下的氨基酸、优选 L-氨基酸: Lys、Asn、Arg、Ser、Thr、Ile、Met、Glu、Asp、Ala、Val、Gln、His、Pro、Leu、Tyr、Trp、Cys 或 Phe。(所述氨基酸、包括甘氨酸, 在本领域也称作蛋白质氨基酸)。非常优选的是在所述酶的第 149 位用 L-天冬氨酸取代甘氨酸(G149D)。最优选的是如上所述的本发明的 γ -谷氨酰激酶, 其长度为 369 ± 40 、优选 ± 20 、更优选 ± 10 及特别优选 ± 5 个氨基酸或者 ± 3 个氨基酸。已知称作氨基肽酶的宿主固有的酶能从形成的蛋白质中切割下 N 末端甲硫氨酸。进一步已知从蛋白质的 C 末端切割下 1 或 2 个及不超过 3 个氨基酸即使削弱所述酶活性, 最多也仅是可忽略程度地削弱。然而, 所述酶由于附加某些融合蛋白而可以增加长度(见下文描述)。

本发明还涉及 SEQ ID NO:2、优选 SEQ ID NO:4 所示氨基酸序列或者与其至少 90% 相同的序列, 所述序列在第 149 位或类似位置含有甘氨酸由另一氨基酸的取代, 特别是由所述的任何优选的氨基酸的取代。本发明因此还涉及在氨基酸水平与 SEQ ID NO:2、优选 SEQ ID NO:4 具有上述程度相同性的那些酶。这些酶可同样源自天然来源。或者, 它们可以通过重组 DNA 技术修饰, 由此技术人员可预测所述酶活性被保留或基本保留(参见例如 Sambrook et al, "Molecular Cloning, A Laboratory Handbook", 2nd edition 1989, CSH Press, Cold Spring Harbor, Ausubel et al. "Current Protocols in Molecular Biology", John Wiley & Sons, NY 2001 所述)。因此, 不存在于活性位点并且其由“相同种类的”氨基酸的取代预期不导致三维结构实质改变的氨基酸可以由“相同种类的”氨基酸取代。可预期例如具有非极性侧链的某些氨基酸(相同种类的氨基酸)可以被取代, 例如异亮氨酸由缬氨酸取代, 而这对本发明的酶的生物学功能或酶功能或者对酶活性

无（实质）影响。技术人员基于这种知识，可相应总结出其它类型氨基酸的取代(例如用其它碱性氨基酸置换碱性氨基酸，或者用其它不带电极性侧链氨基酸置换这一组的氨基酸)。

同样优选的是具有至少相应于 SEQ. ID NO:2、优选 SEQ. ID NO:4 的第 145—154 位、更优选第 130—169 位、特别优选第 110—189 位或类似位置的氨基酸序列的 γ -谷氨酰激酶， γ -谷氨酰激酶氨基酸序列的长度可相应于上述数值。

在另一优选的实施方案中，本发明的酶另外包含至少一个异源氨基酸节段，由于这个节段这些多肽的特征为融合蛋白。本发明的融合蛋白的异源成分的例子可以是标记(例如 His 标记或 Flag 标记)，其可用于本发明的融合蛋白的纯化中。在其它实施方案中，所述异源成分可具有单独的酶活性。在这种情况下，所述两种酶成分优选通过接头连接在一起，所述接头如长度为 6—10 个氨基酸的柔性甘氨酸或甘氨酸-丝氨酸接头，以保证所述成分的功能性。如本文所用，术语“异源”一方面可意味着所述融合蛋白的成分不是以共价键连接形式一起天然发生的，另一方面是指所述成分衍生自不同物种。融合蛋白通常是通过重组 DNA 技术制备(见 Sambrook et al., 上述引文)。

在另一个实施方案中，本发明涉及编码本发明 γ -谷氨酰激酶的核苷酸序列（核酸序列）。因此，本发明还涉及编码 γ -谷氨酰激酶的可复制的核苷酸序列，由这些核苷酸序列编码的相应氨基酸序列可以在第 149 位或类似位置具有除了甘氨酸之外的任何蛋白质氨基酸。

本发明的核苷酸序列优选编码 γ -谷氨酰激酶，相应的被编码的氨基酸序列在第 149 位或类似位置包含选自如下一组的一种氨基酸：Lys、Asn、Arg、Ser、Thr、Ile、Met、Glu、Asp、Ala、Val、Gln、His、Pro、Leu、Tyr、Trp、Cys 或 Phe。特别优选的是编码在第 149 位或类似位置具有 L-天冬氨酸的 γ -谷氨酰激酶的核苷酸序列。

本发明还涉及编码在第 149 位或类似位置具有氨基酸取代的 γ -谷氨酰激酶的可复制的核酸序列，这些序列：

- a) 与 SEQ. ID NO: 1、优选 SEQ. ID NO:3 至少 70% 相同，或者
- b) 编码本发明的 γ -谷氨酰激酶，其长度为 369 ± 40 个氨基酸，或者
- c) 编码本发明的 γ -谷氨酰激酶，其具有 SEQ. ID NO:2、优选 SEQ. ID NO:4 的氨基酸序列，至少在第 145—154 位，或者
- d) 编码本发明的 γ -谷氨酰激酶，其具有 SEQ. ID NO:2、优选 SEQ. ID NO:4 的氨基酸序列，至少在第 145—154 位，并且在严格实验条件下与互补于 SEQ. ID NO:1 或 SEQ. ID NO:3 的核苷酸序列杂交，或者
- f) 编码本发明 γ -谷氨酰激酶的可复制的核苷酸序列，其碱基序列在第 446 位是腺嘌呤，如 SEQ ID NO:3 所示。

权利要求书的范围优选地同样包含编码本发明 γ -谷氨酰激酶的可复制的核酸序列，其长度为 369 ± 20 、更优选 ± 10 及最优选 ± 5 或 ± 3 个氨基酸残基。

本发明还涉及如 SEQ. ID NO:1、优选 SEQ. ID NO:3 所示的核苷酸序列。正如已经提及的，本发明还包含在核苷酸水平与所述序列至少 70% 相同的那些序列。与 SEQ ID NO:3 至少 70% 相同的核苷酸序列例如 SEQ ID NO:5 所示。SEQ ID NO:6 示出由 SEQ ID NO:5 编码的 γ -谷氨酰激酶的氨基酸序列。

本发明还涉及编码 γ -谷氨酰激酶的可复制的核酸序列，所述的 γ -谷氨酰激酶优选包括至少相应于 SEQ ID NO: 2、优选 SEQ. ID NO:4 的第 130—169 位或类似位置、特别优选第 110—189 位的氨基酸序列，并且所述核酸序列在严格实验条件下与互补于 SEQ ID NO: 1 或 SEQ ID NO:3 的核苷酸序列杂交。

术语“互补”根据本发明是指互补性无缺口地延伸穿过本发明的核酸分子的全部区域。换句话说，优选本发明序列的全部区域、

即从 5'末端至 3'末端、特别是编码区(cds)的 100%互补性。在另外优选的实施方案中,互补性延伸穿过优选地不编码酶活性的活性位点的至少 19 个、优选至少 21 个连续核苷酸的区域。

优选所述核酸序列衍生自棒状细菌,优选衍生自谷氨酸棒杆菌。存在于一个物种种群中的基因或等位基因的核酸序列在本领域也称作内源基因或内源等位基因。

本发明进一步涉及具有本发明的核苷酸序列的重组(rec)运载体。合适的运载体可以是技术人员认为适合该目的的任何运载体,特别是载体和宿主生物体。

在这方面可以提及的宿主生物体的例子是酵母如 *Hansenula polymorpha*、毕赤氏酵母 (*Pichia sp.*)、酿酒酵母 (*Saccharomyces cerevisiae*), 原核生物如大肠杆菌、枯草芽孢杆菌 (*Bacillus subtilis*), 棒状细菌如谷氨酸棒杆菌, 或者真核细胞如哺乳动物细胞、昆虫细胞或植物细胞。当合适时在本发明之前这些生物体已经在其细胞中或其周围的发酵培养基中积累 L-脯氨酸。因此,在这个优选的实施方案中,本发明的宿主是已经用本发明的核酸序列或本发明的载体(见下文描述)转化或转染或通过接合方式被提供了本发明的核酸序列或载体的重组细胞(术语“转化”、“转染”和“接合”根据本发明可同义使用)。转化和转染可分别根据已知方法进行,例如通过磷酸钙共沉淀、脂染、电穿孔、粒子轰击或病毒感染等方法进行。本发明的细胞可含有染色体外的或染色体整合形式的重组核酸。换句话说,所述转染/转化可以是稳定或瞬时转染/转化。克隆方法为技术人员所熟知(Sambrook, J.; Fritsch, E. F. and Maniatis, T. (1989), *Molecular cloning: a laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory Press, New York)。大肠杆菌菌株可用于重组制备和诱变方法中,所述大肠杆菌菌株特别包括: *E. coli* XL1 Blue、NM 522、JM101、JM109、JM105、RR1、DH5 α 、TOP10、HB101、BL21 codon plus、BL21 (DE3) codon plus、BL21、BL21 (DE3)、

MM294。用于将具有本发明的核酸的基因构建体克隆进宿主生物体中的质粒同样为技术人员所已知(见 PCT/EP03/07148 所述; 见下文所述)。

可特别提及的棒杆菌属的宿主是本领域已知的谷氨酸棒杆菌物种。已知的棒杆菌属的野生型菌株的例子是:

谷氨酸棒杆菌 ATCC13032

乙酰谷氨酸棒杆菌 (*Corynebacterium acetoglutamicum*)
ATCC15806

乙酰谷氨酸棒杆菌 ATCC13870

Corynebacterium efficiens DSM 44549

Corynebacterium melassecola ATCC17965

Corynebacterium thermoaminogenes FERM BP-1539

黄色短杆菌 (*Brevibacterium flavum*) ATCC14067

乳发酵短杆菌 (*Brevibacterium lactofermentum*) ATCC13869 及

分支短杆菌 (*Brevibacterium divaricatum*) ATCC14020。

关于这组细菌菌株的分类信息可见于 Kämpfer and Kroppenstedt (Canadian Journal of Microbiology 42, 989-1005 (1996)) 及 US-A-5,250,434 所述。近年来(Liebl et al., International Journal of Systematic Bacteriology 41(2), 255-260 (1991)), 种名称作“黄色短杆菌 (*Brevibacterium flavum*)”、“乳发酵短杆菌 (*Brevibacterium lactofermentum*)”和“分支短杆菌 (*Brevibacterium divaricatum*)”的棒状细菌已经归划在谷氨酸棒杆菌物种中。种名为“*Corynebacterium melassecola*”的棒状细菌同样属于谷氨酸棒杆菌物种。

产生 L-脯氨酸的棒状细菌菌株的例子是:

乳发酵短杆菌 NRRL B-11421

黄色短杆菌 NRRL B-11422

谷氨酸棒杆菌 NRRL B-11423

Microbacterium ammoniophilum NRRL B-11424

谷氨酸棒杆菌 ATCC 21157

谷氨酸棒杆菌 ATCC 21158

谷氨酸棒杆菌 ATCC 21159

谷氨酸棒杆菌 ATCC 21355

Corynebacterium acetophilum FERM-P 4045

Corynebacterium acetoacidophilum FERM-P 4962

Arthrobacter citreus FERM-P 4963

Microbacterium ammoniophilum FERM-P 4964

所有这些菌株均在 US 4,224,409 和 US 4,444,885 中描述。

在本发明的载体中，本发明的核酸序列优选与一个表达控制序列可操纵地连接，以便能被转录及当合适时在合适的宿主细胞中翻译。表达控制序列通常包含一个启动子，当合适时还包含调节序列如操纵子或增强子。其中还可以存在翻译起始序列。对于原核或真核宿主细胞而言合适的表达控制序列为技术人员所已知(见例如 Sambrook et al., 见上述引文)。本发明的重组载体可进一步包括通常的元件如复制起点和选择标记基因。合适的重组载体的例子是质粒、粘粒、噬菌体或病毒(见例如 Sambrook et al., 如前述)。

合适的质粒原则上可以是技术人员为此目的可获得的任何质粒。这类质粒可见例如 Studier 及其同事的文章所述(Studier, W. F.; Rosenberg A. H.; Dunn J. J.; Dubendorff J. W.; (1990), Use of the T7 RNA polymerase to direct expression of cloned genes, *Methods Enzymol.* 185, 61-89), 或者 Novagen, Promega, New England Biolabs, Clontech 或 Gibco BRL 的手册所述。另外优选的质粒和载体可见于 Glover, D. M. (1985), *DNA cloning: a practical approach*, Vol. I-III, IRL Press Ltd., Oxford; Rodriguez, R.L. and Denhardt, D. T (eds) (1988), *Vectors: a survey of molecular cloning vectors and their uses*, 179-204, Butterworth, Stoneham; Goeddel, D. V. (1990), *Systems for heterologous gene*

expression, Methods Enzymol. 185, 3-7; Sambrook, J.; Fritsch, E. F. and Maniatis, T. (1989), Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York 所述。

可用于以特别优选的方式将具有预定核酸序列的基因构建体克隆进宿主生物体中的质粒是或基于：pUC18/19 (Roche Biochemicals)、pKK-177-3H (Roche Biochemicals)、pBTac2 (Roche Biochemicals)、pCR (Invitrogen)、pKK223-3 (Amersham Pharmacia Biotech)、pKK-233-3 (Stratagene)或pET (Novagen)。其它优选的质粒是 pBR322 (DSM3879)、pACYC184 (DSM4439) 和 pSC101 (DSM6202), 所有这些质粒均可得自 DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Brunswick, Germany。优选的启动子例如是 T7 启动子、lac 启动子、tac 启动子、trp 启动子、rha 启动子和 ara 启动子。

优选地, 本发明的 γ -谷氨酰激酶或编码其的核酸在棒状细菌、优选在棒杆菌属、特别优选在谷氨酸棒杆菌菌种中过表达。

过表达意味着本发明的 γ -谷氨酰激酶的胞内浓度或活性增加。

所述过表达措施使得相应蛋白质的活性或浓度基于起始微生物中蛋白质的活性或浓度增加至少 10%、25%、50%、75%、100%、150%、200%、300%、400%或 500%, 直至最大 1000%或 2000%。

过表达可以通过使本发明的基因或等位基因的拷贝数增加至少一个拷贝或者通过突变启动子区域和调节区或者位于结构基因上游的核糖体结合位点而实现。掺入结构基因上游的表达盒以相同方式起作用。另外, 诱导型启动子可以增加发酵生产 L-脯氨酸期间的表达。延长 mRNA 寿命的措施也改良表达。酶活性通过阻止酶的降解而进一步增强。基因或基因构建体可以不同拷贝数存在于质粒中或者以整合及扩增的方式存在于染色体中。或者, 感兴趣的基因可通过改变培养基成分和培养方法而进一步过表达。

在棒状细菌中复制的质粒可用于增加本发明的 *proB* 等位基因的拷贝数。许多已知的质粒载体例如 pZ1 (Menkel et al., *Applied and Environmental Microbiology* (1989) 64: 549-554)、pEKEx1 (Eikmanns et al., *Gene* 102:93-98 (1991))或者 pHS2-1 (Sonnen et al., *Gene* 107:69-74 (1991))是基于隐蔽性质粒 pHM1519、pBL1 或 pGA1。其它质粒载体如基于 pCG4 (US-A 4,489,160)或 pNG2 (Serwold-Davis et al., *FEMS Microbiology Letters* 66, 119-124 (1990))或 pAG1 (US-A 5,158,891)的那些载体可以相同方式应用。关于谷氨酸棒杆菌质粒载体的综述可见于 Tauch et al. (*Journal of Biotechnology* 104(1-3), 27-40 (2003))所述。

也可以通过应用染色体基因扩增方法增加拷贝数, 所述方法例如 Reinscheid et al. (*Applied and Environmental Microbiology* 60, 126-132 (1994))在描述 *hom-thrB*-operon 的复制和扩增中所描述。这种方法包括将完整的基因或等位基因克隆进可以在宿主(典型为大肠杆菌)中复制但在谷氨酸棒杆菌中不复制的质粒载体中。合适的载体例如是 pSUP301 (Simon et al., *Bio/Technology* 1, 784-791 (1983))、pK18mob 或 pK19mob (Schäfer et al., *Gene* 145, 69-73 (1994))、pGEM-T (Promega Corporation, Madison, WI, USA)、pCR2.1-TOPO (Shuman, *Journal of Biological Chemistry* 269:32678-84 (1994) ; US-A 5,487,993) 、 pCR®Blunt (Firma Invitrogen, Groningen, the Netherlands; Bernard et al., *Journal of Molecular Biology*, 234: 534-541 (1993))、pEM1 (Schrumpf et al., *Journal of Bacteriology* 173:4510-4516 (1991))或者 pBGS8 (Spratt et al., *Gene* 41: 337-342 (1986))。随后将含有被扩增的基因或等位基因的质粒载体通过接合或转化转移进希望的谷氨酸棒杆菌菌株中。接合方法例如在 Schäfer et al. (*Applied and Environmental Microbiology* 60, 756-759 (1994))中描述。转化方法例如由 Thierbach et al. (*Applied Microbiology and Biotechnology* 29, 356-362 (1988)), Dunican and Shivnan (*Bio/Technology* 7, 1067-1070 (1989))

及 Tauch et al. (FEMS Microbiological Letters 123, 343-347 (1994))描述。在通过“交换”同源重组后,所得菌株含有所述基因或等位基因的至少两个拷贝。为了使拷贝数增加至少 1、2 或 3 个,也可以使用如 WO 03/014330 所述的串联扩增方法,或者使用如 WO 03/040373 所述的通过在希望的位置整合的扩增方法。

使用现有技术描述的诱变方法在 γ -谷氨酰激酶中产生本发明的氨基酸取代(例如 SEQ ID NO: 2)及本发明的其它 proB 突变,其特征在于第 149 位的氨基酸取代。合适的诱变方法是技术人员据此目的可利用的任何方法。

使用诱变物质例如 N-甲基-N'-硝基-N-亚硝基胍(MNNG)或者紫外光的传统体内诱变方法可用于进行诱变。在适当情况中,被诱变的细胞随后被加至含有浓度为约 0.5 – 1 g/l、约 1 – 2 g/l 或者约 2 – 3 g/l 的 3,4-脱氢-DL-脯氨酸的基本琼脂中。在克隆之后当合适时,分离各个突变体,确定 proB 基因或等位基因的核苷酸序列。

也可以使用体外诱变方法,例如用羟胺处理(Miller, J. H.: A Short Course in Bacterial Genetics. A Laboratory Manual and Handbook for Escherichia coli and Related Bacteria, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1992)或者用诱变寡核苷酸处理(T. A. Brown: Gentechnologie für Einsteiger [genetic engineering for beginners], Spektrum Akademischer Verlag, Heidelberg, 1993; textbook by Knippers („Molekulare Genetik“ [molecular genetics], 6th edition, Georg Thieme Verlag, Stuttgart, Germany, 1995); Winnacker („Gene und Klone“ [genes and clones], VCH Verlagsgesellschaft, Weinheim, Germany, 1990); Hagemann („Allgemeine Genetik“ [general genetics], Gustav Fischer Verlag, Stuttgart, Germany, 1986)), 或者如 Newton and Graham 的手册 (PCR, Spektrum Akademischer Verlag, Heidelberg, Germany, 1994)所述的聚合酶链反应。体外诱变方法特别包括饱和诱变、随机诱变、体外重组方法及定向诱变(Eigen, M. and Gardiner, W.

(1984), Evolutionary molecular engineering based on RNA replication, *Pure Appl. Chem.* 56, 967-978; Chen, K. and Arnold, F. (1991), Enzyme engineering for nonaqueous solvents: random mutagenesis to enhance activity of subtilisin E in polar organic media. *Bio/Technology* 9, 1073-1077; Horwitz, M. and Loeb, L. (1986), Promoters Selected From Random DNA-Sequences, *Proc Natl Acad Sci USA* 83, 7405-7409; Dube, D. and L. Loeb (1989), Mutants Generated By The Insertion Of Random Oligonucleotides Into The Active-Site Of The Beta-Lactamase Gene, *Biochemistry* 28, 5703-5707; Stemmer, P.C. (1994), Rapid evolution of a protein in vitro by DNA shuffling, *Nature* 370, 389-391 及 Stemmer, P.C. (1994), DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution. *Proc Natl Acad Sci USA* 91, 10747-10751)。

体外方法的使用包括从分离自野生型菌株的总 DNA 开始, 借助于聚合酶链反应如现有技术所述扩增 *proB* 基因, 当合适时, 将所述基因克隆进合适的质粒载体中, 然后将所述 DNA 进行诱变。借助于聚合酶链反应 (PCR) 扩增 DNA 序列的指导可见于 Gait: *Oligonucleotide Synthesis: A Practical Approach* (IRL Press, Oxford, UK, 1984) 及 Newton and Graham: *PCR* (Spektrum Akademischer Verlag, Heidelberg, Germany, 1994) 所述。同样也可以使用例如熟知的 Sambrook et al. (*Molecular Cloning, A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA, 1989) 所述的体外诱变方法。相应方法也可以以“试剂盒”形式商购, 例如 Stratagene (La Jolla, USA) 的 QuikChange 定向诱变试剂盒, 如 Papworth et al. (*Strategies* 9(3), 3-4 (1996)) 所述。随后选择合适的 *proB* 突变体并通过上述方法研究。

为了生产 L-脯氨酸, 除了使用已经被过表达或扩增的或未被过表达或扩增的本发明的 γ -谷氨酰激酶之外, 还同时扩增特别是过表达

除了所述激酶之外的脯氨酸生物合成的一或多种酶是有益的。优选使用内源基因。

“内源基因”或“内源核苷酸序列”是指存在于一个物种种群中的基因或核苷酸序列和等位基因。

文中术语“扩增”描述了微生物中由相应 DNA 编码的一或多种酶或蛋白质的胞内活性或浓度的增加，例如通过增加基因的拷贝数、使用强启动子或使用编码高活性相应酶或蛋白质的基因或等位基因，及组合这些措施。

扩增、特别是过表达措施使得相应酶或蛋白质的活性或浓度基于野生型蛋白质或起始微生物中所述蛋白质的活性或浓度增加至少 10%、25%、50%、75%、100%、150%、200%、300%、400% 或 500%，直至最大 1000% 或 2000%。

为了生产 L-脯氨酸，除了使用本发明的 proB 基因的变体之外，也可以扩增、特别是过表达选自如下的一或多个基因：

- 编码谷氨酸脱氢酶(EC 1.4.1.4)的 gdh 基因，
- 编码γ-谷氨酰-磷酸还原酶(EC 1.2.1.41)的 proA 基因，
- 编码吡咯啉-5-羧酸还原酶(EC 1.5.1.2)的 proC 基因，
- 编码的鸟氨酸环化脱氨酶(EC 4.3.1.12)的 ocd 基因。

为了生产 L-脯氨酸，除了使用本发明的 proB 基因突变体之外同时弱化、特别是降低选自如下的一或多个内源基因的表达也是有益的：

- 编码苏氨酸脱氨酶(EC 4.2.1.16)的 ilvA 基因，
- 编码脯氨酸脱氢酶/吡咯啉-5-羧酸脱氢酶(EC 1.5.99.8)的 putA 基因，
- 编码 2-酮戊二酸脱氢酶(EC 1.2.4.2)的 sucA 基因，
- 编码二氢硫辛酰胺琥珀酰转移酶(EC 2.3.1.61)的 sucB 基因，及
- 编码乙酰鸟氨酸氨基转移酶(EC 2.6.1.11)的 argD 基因。

文中术语“弱化”描述了微生物中由相应 DNA 编码的一或多种酶或蛋白质的胞内活性或浓度的降低或消除，例如通过使用弱启动子或者使用编码低活性相应酶的基因或等位基因或者使相应基因或酶或蛋白质失活及在适当情况中组合这些措施而实现。

弱化可以通过降低或消除所述基因的表达或者所述酶蛋白质的催化或调节性质而实现。在适当情况中可以组合这两种措施。

基因表达可以通过合适的培养方法或通过对基因表达的信号结构进行遗传修饰（突变）而降低。基因表达的信号结构的例子是抑制物基因、激活物基因、操纵基因、启动子、弱化子、核糖体结合位点、起始密码子及终止子。关于这方面的信息可见于例如专利申请 WO 96/15246、Boyd and Murphy (Journal of Bacteriology 170: 5949-5952 (1988))、Voskuil and Chambliss (Nucleic Acids Research 26: 3584-3590 (1998))、Pátek et al. (Microbiology 142: 1297-309 (1996))及 Journal of Biotechnology 104: 311-323 (2003))及已知的遗传和分子生物学教科书所述，如 Knippers (Molekulare Genetik, 6th edition, Auflage, Georg Thieme Verlag, Stuttgart, Germany, 1995)或者 Winnacker (Gene und Klone, VCH Verlagsgesellschaft, Weinheim, Germany, 1990)所著教科书中描述。

对基因表达进行特异性调节的例子是克隆被弱化的基因，使其在可通过添加定量的 IPTG (异丙基- β -D-硫代半乳糖吡喃糖苷)而诱导的启动子的控制下，例如 trc 启动子或 tac 启动子。本发明可用的载体例如是大肠杆菌表达载体 pXK99E (WO0226787; 根据布达佩斯条约于 2001 年 7 月 31 日保藏在 Deutsche Sammlung für Mikroorganismen und Zellkulturen (DSMZ, Brunswick, Germany) 的 DH5alpha/pXK99E 中，保藏号为 DSM14440) 或者 pVWEx2 (Wendisch, Ph. D. thesis, Berichte des Forschungszentrums Jülich, Jül-3397, ISSN 0994-2952, Jülich, Germany (1997))，其可以使克隆的基因以 IPTG 依赖性方式在谷氨酸棒杆菌中表达。

这种方法在例如专利 WO02/26787 中用于调节 *deaD* 基因的表达，通过将载体 pXK99EdeaD 整合在谷氨酸棒杆菌基因组中而进行，或者由 Simic et al. (*Applied and Environmental Microbiology* 68: 3321-3327 (2002)) 用于调节 *glyA* 基因的表达，通过将载体 pK18mobglyA' 整合进谷氨酸棒杆菌中而进行。

特异性降低基因表达的另一种方法是反义技术，包括向靶细胞中导入短的寡核糖核苷酸或寡脱氧核糖核苷酸或合成较长的反义 RNA 的载体。在此所述反义 RNA 可以结合特异性 mRNA 的互补片段并降低其稳定性或阻断可翻译性。这种例子可见于 Srivastava et al. (*Applied Environmental Microbiology* 2000 Oct; 66 (10): 4366 – 4371) 所描述。

导致酶蛋白质的催化性质改变或降低的突变为本领域技术人员所已知，可以提及的例子可见于 Qiu and Goodman (*Journal of Biological Chemistry* 272: 8611-8617 (1997))、Sugimoto et al. (*Bioscience Biotechnology and Biochemistry* 61: 1760-1762 (1997)) 和 Möckel (Ph. D. thesis, Berichte des Forschungszentrums Jülich, Jül-2906, ISSN09442952, Jülich, Germany (1994)) 所进行的研究工作。这方面的综述可见于已知的遗传和分子生物学教科书描述，例如 Hagemann (*Allgemeine Genetik*, Gustav Fischer Verlag, Stuttgart, Germany, 1986) 所述。

可提及的突变是转换、颠换、插入和缺失。根据氨基酸取代对酶活性的作用，可提及的是错义取代或无义突变。无义突变导致至少一个终止密码子位于基因的编码区中，结果翻译被过早地终止。在基因中插入或缺失至少一个碱基对导致移码突变，结果掺入不正确的氨基酸或者翻译被过早终止。一或多个密码子的缺失典型导致酶活性全部丧失。产生这种突变的指导可见于已知的遗传和分子生物学教科书所述，例如 Knippers (*"Molekulare Genetik"*, 6th edition, Georg Thieme Verlag, Stuttgart, Germany, 1995)、Winnacker (*"Gene*

und Klone”, VCH Verlagsgesellschaft, Weinheim, Germany, 1990)或者 Hagemann (“Allgemeine Genetik”, Gustav Fischer Verlag, Stuttgart, 1986)所著教科书中描述。

使用实现弱化的措施的结果是相应蛋白质的活性或浓度通常降低至野生型蛋白质的活性或浓度或起始微生物中所述蛋白质的活性或浓度的 0—75%、0—50%、0—25%、0—10%、0—5%或者 0—1%。

根据本发明制备的微生物（也是本发明的主题）可以分批或补料分批或者重复补料分批方法连续或非连续培养，以生产 L-脯氨酸。已知的培养方法综述在 Chmiel (Bioprozesstechnik 1. Einführung in die Bioverfahrenstechnik [Bioprocess Technology 1. Introduction to Bioprocess Technology] (Gustav Fischer Verlag, Stuttgart, 1991))或者 Storhas (Bioreaktoren und periphere Einrichtungen [Bioreactors and Peripheral Equipment] (Vieweg Verlag, Brunswick/Wiesbaden, 1994))所著教科书中描述。

使用的培养基必须符合特定菌株的需要。关于培养各种微生物的培养基的描述见 American Society for Bacteriology(Washington D.C., USA, 1981)出版的 Manual of Methods for General Bacteriology 所描述。

可使用的碳源包括糖及碳水化合物，例如葡萄糖，蔗糖，乳糖，果糖，麦芽糖，糖蜜，淀粉和纤维素，油和脂肪如豆油，葵花油，落花生油和椰子油，脂肪酸如棕榈酸，硬脂酸和亚油酸，醇如甘油和乙醇，糖醇如核糖醇和甘露醇，及有机酸如乙酸，这些物质可单独或混合使用。

可使用的氮源包括含氮的有机化合物如胨，酵母提取物，肉膏，麦芽提取物，玉米浸液、大豆粉和尿素，或无机化合物如硫酸铵，氯化铵，磷酸铵，碳酸铵和硝酸铵。这些氮源可单独或混合使用。

可使用的磷源包括磷酸，磷酸二氢钾或磷酸氢二钾，或相应钠盐。另外培养基还必须含有为生长所需的金属盐如硫酸镁或硫酸铁。最后，除了上述物质之外，可使用生长必需物质如氨基酸和维生素。此外，可将适当前体加入培养基中。上述加入的物质可以分批形式或在培养期间以适当方式补加入培养物中。

可以适当方式加入碱性化合物如氢氧化钠，氢氧化钾，氨或氨水，或酸性化合物如磷酸或硫酸，以调节培养物的 pH 值。抗泡沫剂例如脂肪酸聚乙二醇酯可用于控制泡沫产生。适当的选择性作用物质例如抗生素，可加入培养基中以保持质粒的稳定性。氧气或含氧混合气，例如空气，可充入培养物中以保持有氧条件。培养温度通常在 20℃—45℃、优选 25℃—40℃。持续培养直至形成最大量的 L-脯氨酸或者直至产量或生产力已经达到希望的最佳水平。此目的通常在 10~160 小时范围达到。

确定 L-脯氨酸的方法是现有技术中已知的。所述分析方法可如 Speckman et al. (Analytical Chemistry, 30 (1958), 1190)所述通过阴离子交换层析，随后经茚三酮衍生化作用及在适当波长进行检测。

因此，本发明的另一实施方案包括通过如下步骤生产 L-脯氨酸的方法：

- a) 发酵表达或过表达本发明的至少一个核苷酸序列的宿主生物体，
- b) 分离或收集 L-脯氨酸，当合适时伴有发酵肉汤和/或生物量成分。

优选应用生产 L-脯氨酸的方法，所述方法包括如下步骤：

- a) 发酵表达或过表达本发明的至少一个核苷酸序列的棒状细菌，
- b) 浓缩发酵肉汤或棒状细菌细胞中的 L-脯氨酸，
- c) 从发酵肉汤中分离或收集 L-脯氨酸，适当时

d) 具有发酵肉汤和/或生物量成分(具有从 > 0 至 $<100\%$ 重量的生物量, 优选从 $10-80\%$ 、更优选 $20-60\%$ 重量的生物量)。

技术人员可以收集和分离以这种方式产生的 L-脯氨酸并在适当时进行纯化。

本发明的方法可用于经发酵生产 L-脯氨酸。

本发明所用术语“类似位置”是指这样一个位置, 其通过应用序列对比程序(BLAST, Altschul et al. J. Mol. Biol. 1990, 215, 403-10)在起始序列预定位置对比起始序列与相应序列, 提供在相应序列中与对比的所述位置相差不超过 ± 5 、更优选 ± 4 、更优选 ± 3 、更优选 ± 2 、最优选 ± 1 甚至为 0 的氨基酸位置。

关于杂交的指导可见于 Boehringer Mannheim GmbH (Mannheim, Germany, 1993) 的 The DIG System Users Guide for Filter Hybridization 及 Liebl et al. (International Journal of Systematic Bacteriology 41: 255-260 (1991))所著手册描述。所述杂交是在严格条件下进行, 即仅形成其中探针例如与 SEQ ID NO: 3 互补的核苷酸序列与靶序列(即用所述探针处理的多核苷酸)至少 70%相同的杂交体。杂交、包括洗涤步骤的严格性已知通过改变缓冲液成分、温度和盐浓度而受影响或决定。杂交反应通常在与洗涤步骤相比相对低的严格性下进行 (Hybaid Hybridisation Guide, Hybaid Limited, Teddington, UK, 1996)。

例如。可以使用相应于 $5\times\text{SSC}$ 缓冲液的缓冲液及在 $50^{\circ}\text{C}-68^{\circ}\text{C}$ 的条件进行杂交反应。在这种情况下, 探针也可以和与探针序列低于 70%相同的多核苷酸杂交。这种杂交体稳定性差, 通过在严格条件下洗涤除去。这可以例如通过将盐浓度降低至 $2\times\text{SSC}$ 及适当的随后至 $0.5\times\text{SSC}$ (The DIG System User's Guide for Filter Hybridisation, Boehringer Mannheim, Mannheim, Germany, 1995), 温度设定在大约 $50^{\circ}\text{C}-68^{\circ}\text{C}$ 、 $52^{\circ}\text{C}-68^{\circ}\text{C}$ 、 $54^{\circ}\text{C}-68^{\circ}\text{C}$ 、 $56^{\circ}\text{C}-68^{\circ}\text{C}$ 、 $58^{\circ}\text{C}-68^{\circ}\text{C}$ 、 $60^{\circ}\text{C}-68^{\circ}\text{C}$ 、 $62^{\circ}\text{C}-68^{\circ}\text{C}$ 、 $64^{\circ}\text{C}-68^{\circ}\text{C}$ 、 $66^{\circ}\text{C}-68^{\circ}\text{C}$ 而实

现。优选在大约 62°C – 68°C、特别优选在大约 64°C – 68°C 或 66°C – 68°C 温度进行洗涤步骤。在适当情况中，可以将盐浓度降低至相应于 0.2× SSC 或者 0.1× SSC 浓度。通过以大约 1–2°C 梯度将杂交温度从 50°C 逐步升高至 68°C，可以分离与应用的探针序列至少 70%、或至少 80%、或者至少 90%–95%、或者至少 96%–98%或者至少 99% 相同的多核苷酸片段。关于杂交的进一步指导可以试剂盒形式商购(例如 DIG Easy Hyb from Roche Diagnostics GmbH, Mannheim, Germany, Catalogue No. 1603558)。

根据本发明，请求保护的多肽(氨基酸序列)及核酸序列也包含与任何这些序列分别具有 70% 以上、优选 80%、更优选 85% 以上(关于核酸序列)或 90% (也关于多肽)、优选 91%、92%、93% 或 94%、更优选 95% 或 96% 及特别优选 97%、98% 或 99% 以上(关于这两种序列)的同源性(在氨基酸水平)及相同性(在核酸水平，天然简并除外)的那些序列，只要这个序列的作用或目的保留即可。如本文所用，术语“同源性”(或相同性)可以通过等式 $H(\%) = [1 - V/X] \times 100$ 表明，其中 H 是同源性，X 是对比序列的核碱基/氨基酸总数，V 是基于对比序列在研究序列中不同的核碱基/氨基酸数目。在任何情况中，术语编码多肽的核酸序列均包括根据遗传密码简并性似乎可能的任何序列。

与在本说明书中以 SEQ ID 编号示出的氨基酸序列的相同性百分比可易于由技术人员使用本领域已知的方法确定。根据本发明可以应用的一个合适程序是 BLASTP (Altschul et al., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25(17):3389-3402.)。本发明的核酸序列可以是 DNA 分子或者 RNA 分子。优选所述核酸序列是 DNA 分子或 mRNA 分子。根据本发明，所述 DNA 分子可进一步是基因组或分离的 DNA 分子。本发明进一步包含其中所述 DNA 分子是 PNA 分子或者 DNA 分子的另一种衍生物的实施实施方案。

本申请中以 DSMZ 编号提及的微生物可得自 Deutsche Sammlung für Mikroorganismen und Zellkulturen, Mascheroder Weg 4, Brunswick (Germany)。

```

<110> 德古萨有限责任公司
<120> 棒状细菌的proB基因的突变体
<130> 040071 AM
<160> 6
<170> PatentIn version 3.3

<210> 1
<211> 1110
<212> DNA
<213> coryneform bacterium

<220>
<221> CDS
<222> (1)..(1107)

<220>
<221> misc_feature
<222> (445)..(447)
<223> Xaa all proteinogenic amino acids other than glycine

<400> 1
atg cgt gag cgc atc tcc aac gct aag cga gtg gtg gtg aaa att ggt      48
Met Arg Glu Arg Ile Ser Asn Ala Lys Arg Val Val Val Lys Ile Gly
1          5          10          15

tcg tcc tca ttg act aac gat gag gac gga cac acc gtc gat ccc aac      96
Ser Ser Ser Leu Thr Asn Asp Glu Asp Gly His Thr Val Asp Pro Asn
20          25          30

cgc atc aac act att gtc aat gcc ttg caa gca cgc atg gaa gct ggc      144
Arg Ile Asn Thr Ile Val Asn Ala Leu Gln Ala Arg Met Glu Ala Gly
35          40          45

tcg gac ctc atc gtt gtg tcc tct ggc gca gtg gcc gcg gga atg gcc      192
Ser Asp Leu Ile Val Val Ser Ser Gly Ala Val Ala Ala Gly Met Ala
50          55          60

ccg ctt gga ttg agc acc cgg ccc acg gaa ttg gca gtc aag cag gct      240
Pro Leu Gly Leu Ser Thr Arg Pro Thr Glu Leu Ala Val Lys Gln Ala
65          70          75          80

gca gca gca gtg ggg caa gtt cac ctc atg cac cag tgg gga cgt tct      288
Ala Ala Ala Val Gly Gln Val His Leu Met His Gln Trp Gly Arg Ser
85          90          95

ttt gcc cgg tat ggt cgc ccc atc ggc cag gtg ctt ctt acc gca gct      336
Phe Ala Arg Tyr Gly Arg Pro Ile Gly Gln Val Leu Leu Thr Ala Ala
100          105          110

gat gca gga aag cgt gat cgt gcg agg aat gcg cag cgt acc atc gac      384
Asp Ala Gly Lys Arg Asp Arg Ala Arg Asn Ala Gln Arg Thr Ile Asp
115          120          125

aag ctg cgc att ttg ggc gcg gtt cct atc gtc aat gaa aat gac acc      432

```

Lys 130	Leu	Arg	Ile	Leu	Gly 135	Ala	Val	Pro	Ile	Val	Asn 140	Glu	Asn	Asp	Thr	
gtg	gca	acc	acc	nnn	gtg	aat	ttt	ggg	gac	aac	gac	cga	ctt	gct	gca	480
Val	Ala	Thr	Thr	Xaa	Val	Asn	Phe	Gly	Asp	Asn	Asp	Arg	Leu	Ala	Ala	
145					150					155					160	
att	gtg	gcg	cac	ctg	gtg	tcg	gct	gat	gct	ttg	gtg	ctg	ctc	agt	gac	528
Ile	Val	Ala	His	Leu	Val	Ser	Ala	Asp	Ala	Leu	Val	Leu	Leu	Ser	Asp	
				165					170					175		
gtg	gat	gga	ctt	ttt	gat	aaa	aac	cct	act	gat	ccc	acc	gcg	aag	ttt	576
Val	Asp	Gly	Leu	Phe	Asp	Lys	Asn	Pro	Thr	Asp	Pro	Thr	Ala	Lys	Phe	
			180					185					190			
att	tcc	gag	gtt	cgt	gac	ggc	aat	gat	ttg	aaa	ggg	gtc	att	gcc	ggc	624
Ile	Ser	Glu	Val	Arg	Asp	Gly	Asn	Asp	Leu	Lys	Gly	Val	Ile	Ala	Gly	
		195					200					205				
gac	ggc	gga	aaa	gtg	ggc	acc	ggg	ggc	atg	gca	tca	aag	gtg	tct	gct	672
Asp	Gly	Gly	Lys	Val	Gly	Thr	Gly	Gly	Met	Ala	Ser	Lys	Val	Ser	Ala	
	210					215					220					
gca	cgt	ttg	gct	tcc	cga	agt	ggc	gtg	cct	gtg	ctg	ttg	acc	tct	gcg	720
Ala	Arg	Leu	Ala	Ser	Arg	Ser	Gly	Val	Pro	Val	Leu	Leu	Thr	Ser	Ala	
225					230					235					240	
gca	aac	att	ggc	cca	gca	ctg	gaa	gac	gcc	cag	gtg	ggc	act	gta	ttc	768
Ala	Asn	Ile	Gly	Pro	Ala	Leu	Glu	Asp	Ala	Gln	Val	Gly	Thr	Val	Phe	
				245					250					255		
cac	ccc	aag	gac	aac	cgc	ctc	tcc	gcg	tgg	aag	ttc	tgg	gct	ttg	tat	816
His	Pro	Lys	Asp	Asn	Arg	Leu	Ser	Ala	Trp	Lys	Phe	Trp	Ala	Leu	Tyr	
			260					265					270			
gcc	gca	gat	act	gca	gga	aag	atc	cga	ctc	gat	gac	ggc	gcg	gtg	gaa	864
Ala	Ala	Asp	Thr	Ala	Gly	Lys	Ile	Arg	Leu	Asp	Asp	Gly	Ala	Val	Glu	
		275					280					285				
gca	gtg	acc	tcc	ggg	ggg	aaa	tct	ttg	ctg	gct	gtg	ggc	att	act	gaa	912
Ala	Val	Thr	Ser	Gly	Gly	Lys	Ser	Leu	Leu	Ala	Val	Gly	Ile	Thr	Glu	
	290					295					300					
atc	att	ggg	gat	ttc	cag	cag	ggg	gag	atc	gtg	gag	atc	ttg	gga	cct	960
Ile	Ile	Gly	Asp	Phe	Gln	Gln	Gly	Glu	Ile	Val	Glu	Ile	Leu	Gly	Pro	
305					310					315					320	
gcc	ggc	caa	atc	atc	ggg	cga	ggc	gag	gtg	tcc	tac	gat	tct	gat	acc	1008
Ala	Gly	Gln	Ile	Ile	Gly	Arg	Gly	Glu	Val	Ser	Tyr	Asp	Ser	Asp	Thr	
				325					330					335		
ttg	caa	tca	atg	gtt	ggg	atg	caa	acg	cag	gac	ctt	cca	gat	ggc	atg	1056
Leu	Gln	Ser	Met	Val	Gly	Met	Gln	Thr	Gln	Asp	Leu	Pro	Asp	Gly	Met	
			340					345					350			
cag	cgc	ccg	gta	gtg	cat	gca	gat	tat	ctg	tcc	aac	tac	gcc	agc	cgc	1104
Gln	Arg	Pro	Val	Val	His	Ala	Asp	Tyr	Leu	Ser	Asn	Tyr	Ala	Ser	Arg	
		355					360					365				
gcg	taa															1110

Ala

<210> 2
 <211> 369
 <212> PRT
 <213> coryneform bacterium

 <220>
 <221> misc_feature
 <222> (149)..(149)
 <223> The 'Xaa' at location 149 stands for Lys, Asn, Arg, Ser, Thr,
 Ile, Met, Glu, Asp, Ala, Val, Gln, His, Pro, Leu, Tyr, Trp,
 Cys, or Phe.

<400> 2

Met Arg Glu Arg Ile Ser Asn Ala Lys Arg Val Val Val Lys Ile Gly
 1 5 10 15

Ser Ser Ser Leu Thr Asn Asp Glu Asp Gly His Thr Val Asp Pro Asn
 20 25 30

Arg Ile Asn Thr Ile Val Asn Ala Leu Gln Ala Arg Met Glu Ala Gly
 35 40 45

Ser Asp Leu Ile Val Val Ser Ser Gly Ala Val Ala Ala Gly Met Ala
 50 55 60

Pro Leu Gly Leu Ser Thr Arg Pro Thr Glu Leu Ala Val Lys Gln Ala
 65 70 75 80

Ala Ala Ala Val Gly Gln Val His Leu Met His Gln Trp Gly Arg Ser
 85 90 95

Phe Ala Arg Tyr Gly Arg Pro Ile Gly Gln Val Leu Leu Thr Ala Ala
 100 105 110

Asp Ala Gly Lys Arg Asp Arg Ala Arg Asn Ala Gln Arg Thr Ile Asp
 115 120 125

Lys Leu Arg Ile Leu Gly Ala Val Pro Ile Val Asn Glu Asn Asp Thr
 130 135 140

Val Ala Thr Thr Xaa Val Asn Phe Gly Asp Asn Asp Arg Leu Ala Ala
 145 150 155 160

Ile Val Ala His Leu Val Ser Ala Asp Ala Leu Val Leu Leu Ser Asp
 165 170 175

Val Asp Gly Leu Phe Asp Lys Asn Pro Thr Asp Pro Thr Ala Lys Phe
180 185 190

Ile Ser Glu Val Arg Asp Gly Asn Asp Leu Lys Gly Val Ile Ala Gly
195 200 205

Asp Gly Gly Lys Val Gly Thr Gly Gly Met Ala Ser Lys Val Ser Ala
210 215 220

Ala Arg Leu Ala Ser Arg Ser Gly Val Pro Val Leu Leu Thr Ser Ala
225 230 235 240

Ala Asn Ile Gly Pro Ala Leu Glu Asp Ala Gln Val Gly Thr Val Phe
245 250 255

His Pro Lys Asp Asn Arg Leu Ser Ala Trp Lys Phe Trp Ala Leu Tyr
260 265 270

Ala Ala Asp Thr Ala Gly Lys Ile Arg Leu Asp Asp Gly Ala Val Glu
275 280 285

Ala Val Thr Ser Gly Gly Lys Ser Leu Leu Ala Val Gly Ile Thr Glu
290 295 300

Ile Ile Gly Asp Phe Gln Gln Gly Glu Ile Val Glu Ile Leu Gly Pro
305 310 315 320

Ala Gly Gln Ile Ile Gly Arg Gly Glu Val Ser Tyr Asp Ser Asp Thr
325 330 335

Leu Gln Ser Met Val Gly Met Gln Thr Gln Asp Leu Pro Asp Gly Met
340 345 350

Gln Arg Pro Val Val His Ala Asp Tyr Leu Ser Asn Tyr Ala Ser Arg
355 360 365

Ala

<210> 3
<211> 1110
<212> DNA
<213> coryneform bacterium

<220>
<221> CDS
<222> (1)..(1107)

```

<220>
<221> mutation
<222> (446)..(446)
<223> adenine

<400> 3
atg cgt gag cgc atc tcc aac gct aag cga gtg gtg gtg aaa att ggt      48
Met Arg Glu Arg Ile Ser Asn Ala Lys Arg Val Val Val Lys Ile Gly
1          5          10          15

tcg tcc tca ttg act aac gat gag gac gga cac acc gtc gat ccc aac      96
Ser Ser Ser Leu Thr Asn Asp Glu Asp Gly His Thr Val Asp Pro Asn
          20          25          30

cgc atc aac act att gtc aat gcc ttg caa gca cgc atg gaa gct ggc      144
Arg Ile Asn Thr Ile Val Asn Ala Leu Gln Ala Arg Met Glu Ala Gly
          35          40          45

tcg gac ctc atc gtt gtg tcc tct ggc gca gtg gcc gcg gga atg gcc      192
Ser Asp Leu Ile Val Val Ser Ser Gly Ala Val Ala Ala Gly Met Ala
          50          55          60

ccg ctt gga ttg agc acc cgg ccc acg gaa ttg gca gtc aag cag gct      240
Pro Leu Gly Leu Ser Thr Arg Pro Thr Glu Leu Ala Val Lys Gln Ala
65          70          75          80

gca gca gca gtg ggg caa gtt cac ctc atg cac cag tgg gga cgt tct      288
Ala Ala Ala Val Gly Gln Val His Leu Met His Gln Trp Gly Arg Ser
          85          90          95

ttt gcc cgg tat ggt cgc ccc atc ggc cag gtg ctt ctt acc gca gct      336
Phe Ala Arg Tyr Gly Arg Pro Ile Gly Gln Val Leu Leu Thr Ala Ala
          100          105          110

gat gca gga aag cgt gat cgt gcg agg aat gcg cag cgt acc atc gac      384
Asp Ala Gly Lys Arg Asp Arg Ala Arg Asn Ala Gln Arg Thr Ile Asp
          115          120          125

aag ctg cgc att ttg ggc gcg gtt cct atc gtc aat gaa aat gac acc      432
Lys Leu Arg Ile Leu Gly Ala Val Pro Ile Val Asn Glu Asn Asp Thr
          130          135          140

gtg gca acc acc gat gtg aat ttt ggt gac aac gac cga ctt gct gca      480
Val Ala Thr Thr Asp Val Asn Phe Gly Asp Asn Asp Arg Leu Ala Ala
145          150          155          160

att gtg gcg cac ctg gtg tcg gct gat gct ttg gtg ctg ctc agt gac      528
Ile Val Ala His Leu Val Ser Ala Asp Ala Leu Val Leu Leu Ser Asp
          165          170          175

gtg gat gga ctt ttt gat aaa aac cct act gat ccc acc gcg aag ttt      576
Val Asp Gly Leu Phe Asp Lys Asn Pro Thr Asp Pro Thr Ala Lys Phe
          180          185          190

att tcc gag gtt cgt gac ggc aat gat ttg aaa ggt gtc att gcc ggc      624
Ile Ser Glu Val Arg Asp Gly Asn Asp Leu Lys Gly Val Ile Ala Gly
          195          200          205

gac ggc gga aaa gtg ggc acc ggt ggc atg gca tca aag gtg tct gct      672

```

Asp	Gly	Gly	Lys	Val	Gly	Thr	Gly	Gly	Met	Ala	Ser	Lys	Val	Ser	Ala	
	210					215					220					
gca	cgt	ttg	gct	tcc	cga	agt	ggc	gtg	cct	gtg	ctg	ttg	acc	tct	gcg	720
Ala	Arg	Leu	Ala	Ser	Arg	Ser	Gly	Val	Pro	Val	Leu	Leu	Thr	Ser	Ala	
225					230					235					240	
gca	aac	att	ggc	cca	gca	ctg	gaa	gac	gcc	cag	gtg	ggc	act	gta	ttc	768
Ala	Asn	Ile	Gly	Pro	Ala	Leu	Glu	Asp	Ala	Gln	Val	Gly	Thr	Val	Phe	
				245					250					255		
cac	ccc	aag	gac	aac	cgc	ctc	tcc	gcg	tgg	aag	ttc	tgg	gct	ttg	tat	816
His	Pro	Lys	Asp	Asn	Arg	Leu	Ser	Ala	Trp	Lys	Phe	Trp	Ala	Leu	Tyr	
			260					265					270			
gcc	gca	gat	act	gca	gga	aag	atc	cga	ctc	gat	gac	ggc	gcg	gtg	gaa	864
Ala	Ala	Asp	Thr	Ala	Gly	Lys	Ile	Arg	Leu	Asp	Asp	Gly	Ala	Val	Glu	
		275					280					285				
gca	gtg	acc	tcc	ggt	ggt	aaa	tct	ttg	ctg	gct	gtg	ggc	att	act	gaa	912
Ala	Val	Thr	Ser	Gly	Gly	Lys	Ser	Leu	Leu	Ala	Val	Gly	Ile	Thr	Glu	
	290					295					300					
atc	att	ggt	gat	ttc	cag	cag	ggt	gag	atc	gtg	gag	atc	ttg	gga	cct	960
Ile	Ile	Gly	Asp	Phe	Gln	Gln	Gly	Glu	Ile	Val	Glu	Ile	Leu	Gly	Pro	
305					310					315					320	
gcc	ggc	caa	atc	atc	ggg	cga	ggc	gag	gtg	tcc	tac	gat	tct	gat	acc	1008
Ala	Gly	Gln	Ile	Ile	Gly	Arg	Gly	Glu	Val	Ser	Tyr	Asp	Ser	Asp	Thr	
				325					330					335		
ttg	caa	tca	atg	ggt	ggt	atg	caa	acg	cag	gac	ctt	cca	gat	ggc	atg	1056
Leu	Gln	Ser	Met	Val	Gly	Met	Gln	Thr	Gln	Asp	Leu	Pro	Asp	Gly	Met	
			340					345					350			
cag	cgc	ccg	gta	gtg	cat	gca	gat	tat	ctg	tcc	aac	tac	gcc	agc	cgc	1104
Gln	Arg	Pro	Val	Val	His	Ala	Asp	Tyr	Leu	Ser	Asn	Tyr	Ala	Ser	Arg	
		355					360					365				
gcg	taa															1110
Ala																
<210>	4															
<211>	369															
<212>	PRT															
<213>	coryneform bacterium															
<400>	4															
Met	Arg	Glu	Arg	Ile	Ser	Asn	Ala	Lys	Arg	Val	Val	Val	Lys	Ile	Gly	
1				5					10					15		
Ser	Ser	Ser	Leu	Thr	Asn	Asp	Glu	Asp	Gly	His	Thr	Val	Asp	Pro	Asn	
			20					25					30			
Arg	Ile	Asn	Thr	Ile	Val	Asn	Ala	Leu	Gln	Ala	Arg	Met	Glu	Ala	Gly	
		35					40					45				

Ser Asp Leu Ile Val Val Ser Ser Gly Ala Val Ala Ala Gly Met Ala
50 55 60

Pro Leu Gly Leu Ser Thr Arg Pro Thr Glu Leu Ala Val Lys Gln Ala
65 70 75 80

Ala Ala Ala Val Gly Gln Val His Leu Met His Gln Trp Gly Arg Ser
85 90 95

Phe Ala Arg Tyr Gly Arg Pro Ile Gly Gln Val Leu Leu Thr Ala Ala
100 105 110

Asp Ala Gly Lys Arg Asp Arg Ala Arg Asn Ala Gln Arg Thr Ile Asp
115 120 125

Lys Leu Arg Ile Leu Gly Ala Val Pro Ile Val Asn Glu Asn Asp Thr
130 135 140

Val Ala Thr Thr Asp Val Asn Phe Gly Asp Asn Asp Arg Leu Ala Ala
145 150 155 160

Ile Val Ala His Leu Val Ser Ala Asp Ala Leu Val Leu Leu Ser Asp
165 170 175

Val Asp Gly Leu Phe Asp Lys Asn Pro Thr Asp Pro Thr Ala Lys Phe
180 185 190

Ile Ser Glu Val Arg Asp Gly Asn Asp Leu Lys Gly Val Ile Ala Gly
195 200 205

Asp Gly Gly Lys Val Gly Thr Gly Gly Met Ala Ser Lys Val Ser Ala
210 215 220

Ala Arg Leu Ala Ser Arg Ser Gly Val Pro Val Leu Leu Thr Ser Ala
225 230 235 240

Ala Asn Ile Gly Pro Ala Leu Glu Asp Ala Gln Val Gly Thr Val Phe
245 250 255

His Pro Lys Asp Asn Arg Leu Ser Ala Trp Lys Phe Trp Ala Leu Tyr
260 265 270

Ala Ala Asp Thr Ala Gly Lys Ile Arg Leu Asp Asp Gly Ala Val Glu
275 280 285

Ala Val Thr Ser Gly Gly Lys Ser Leu Leu Ala Val Gly Ile Thr Glu
290 295 300

Ile Ile Gly Asp Phe Gln Gln Gly Glu Ile Val Glu Ile Leu Gly Pro
305 310 315 320

Ala Gly Gln Ile Ile Gly Arg Gly Glu Val Ser Tyr Asp Ser Asp Thr
325 330 335

Leu Gln Ser Met Val Gly Met Gln Thr Gln Asp Leu Pro Asp Gly Met
340 345 350

Gln Arg Pro Val Val His Ala Asp Tyr Leu Ser Asn Tyr Ala Ser Arg
355 360 365

Ala

<210> 5
<211> 1110
<212> DNA
<213> coryneform bacterium

<220>
<221> CDS
<222> (1)..(1107)

<220>
<221> mutation
<222> (446)..(446)
<223> adenine

<400> 5
atg cgt gaa cgc atc tcc aac gct aag cga gtg gtg gtg aaa att ggt 48
Met Arg Glu Arg Ile Ser Asn Ala Lys Arg Val Val Val Lys Ile Gly
1 5 10 15

tcg tcc tca ttg act aac gat gag gac gga cac acc gtc gat ccc aac 96
Ser Ser Ser Leu Thr Asn Asp Glu Asp Gly His Thr Val Asp Pro Asn
20 25 30

cgc atc aac act att gtc aat gcc ttg caa gca cgc atg gaa gct ggc 144
Arg Ile Asn Thr Ile Val Asn Ala Leu Gln Ala Arg Met Glu Ala Gly
35 40 45

tcg gac ctc atc gtt gtg tcc tct ggc gca gtg gcc gcg gga atg gcc 192
Ser Asp Leu Ile Val Val Ser Ser Gly Ala Val Ala Ala Gly Met Ala
50 55 60

ccg ctt gga ttg agc acc cgg ccc acg gaa ttg gca gtc aag cag gct 240
Pro Leu Gly Leu Ser Thr Arg Pro Thr Glu Leu Ala Val Lys Gln Ala
65 70 75 80

gca gca gca gtg ggg caa gtt cac ctc atg cac cag tgg gga cgt tct 288

Ala	Ala	Ala	Val	Gly	Gln	Val	His	Leu	Met	His	Gln	Trp	Gly	Arg	Ser	
				85					90					95		
ttt	gcc	cgg	tat	ggt	cgc	ccc	atc	ggc	cag	gtg	ctt	ctt	acc	gca	gct	336
Phe	Ala	Arg	Tyr	Gly	Arg	Pro	Ile	Gly	Gln	Val	Leu	Leu	Thr	Ala	Ala	
			100					105					110			
gat	gca	gga	aag	cgt	gat	cgt	gcg	agg	aat	gcg	cag	cgt	acc	atc	gac	384
Asp	Ala	Gly	Lys	Arg	Asp	Arg	Ala	Arg	Asn	Ala	Gln	Arg	Thr	Ile	Asp	
		115					120					125				
aag	ctg	cgc	att	ttg	ggc	gcg	gtt	cct	atc	gtc	aat	gaa	aat	gac	acc	432
Lys	Leu	Arg	Ile	Leu	Gly	Ala	Val	Pro	Ile	Val	Asn	Glu	Asn	Asp	Thr	
	130					135					140					
gtg	gca	acc	acc	gat	gtg	aat	ttt	ggc	gac	aac	gac	cga	ctt	gct	gca	480
Val	Ala	Thr	Thr	Asp	Val	Asn	Phe	Gly	Asp	Asn	Asp	Arg	Leu	Ala	Ala	
	145				150					155					160	
att	gtg	gcg	cac	ctg	gtg	tcg	gct	gac	gct	ttg	gtg	ctg	ctc	agt	gac	528
Ile	Val	Ala	His	Leu	Val	Ser	Ala	Asp	Ala	Leu	Val	Leu	Leu	Ser	Asp	
			165						170					175		
gtg	gat	gga	ctt	ttt	gat	aag	aac	cct	act	gat	ccc	acc	gcg	aag	ttt	576
Val	Asp	Gly	Leu	Phe	Asp	Lys	Asn	Pro	Thr	Asp	Pro	Thr	Ala	Lys	Phe	
			180					185					190			
att	tcc	gag	gtt	cgt	gac	ggc	aat	gat	ttg	aaa	ggc	gtc	att	gcc	ggc	624
Ile	Ser	Glu	Val	Arg	Asp	Gly	Asn	Asp	Leu	Lys	Gly	Val	Ile	Ala	Gly	
		195					200					205				
gac	ggc	gga	aaa	gtg	ggc	acc	ggc	ggc	atg	gca	tca	aag	gtg	tct	gct	672
Asp	Gly	Gly	Lys	Val	Gly	Thr	Gly	Gly	Met	Ala	Ser	Lys	Val	Ser	Ala	
	210					215					220					
gca	cgt	ttg	gct	tcc	cga	agt	ggc	gtg	cct	gtg	ctg	ttg	acc	tct	gcg	720
Ala	Arg	Leu	Ala	Ser	Arg	Ser	Gly	Val	Pro	Val	Leu	Leu	Thr	Ser	Ala	
	225				230					235					240	
gca	aac	att	ggc	cca	gca	ctg	gaa	gac	gcc	cag	gtg	ggc	act	gta	ttc	768
Ala	Asn	Ile	Gly	Pro	Ala	Leu	Glu	Asp	Ala	Gln	Val	Gly	Thr	Val	Phe	
			245						250					255		
cac	ccc	aag	gac	aac	cgc	ctc	tcc	gcg	tgg	aag	ttc	tgg	gct	ttg	tat	816
His	Pro	Lys	Asp	Asn	Arg	Leu	Ser	Ala	Trp	Lys	Phe	Trp	Ala	Leu	Tyr	
			260					265					270			
gcc	gca	gat	act	gca	gga	aag	atc	cga	ctt	gat	gat	ggc	gcg	gtg	gaa	864
Ala	Ala	Asp	Thr	Ala	Gly	Lys	Ile	Arg	Leu	Asp	Asp	Gly	Ala	Val	Glu	
		275					280					285				
gca	gtg	acc	tcc	ggt	ggt	aaa	tct	ttg	ctg	gct	gtg	ggc	att	act	gag	912
Ala	Val	Thr	Ser	Gly	Gly	Lys	Ser	Leu	Leu	Ala	Val	Gly	Ile	Thr	Glu	
	290					295					300					
atc	att	ggt	gat	ttc	caa	cag	ggt	gag	atc	gtg	gag	atc	ttg	gga	cct	960
Ile	Ile	Gly	Asp	Phe	Gln	Gln	Gly	Glu	Ile	Val	Glu	Ile	Leu	Gly	Pro	
	305				310					315					320	
gcc	ggc	caa	atc	atc	ggg	cga	ggc	gag	gtg	tcc	tac	gat	tct	gat	acc	1008

32

Ile Val Ala His Leu Val Ser Ala Asp Ala Leu Val Leu Leu Ser Asp
165 170 175

Val Asp Gly Leu Phe Asp Lys Asn Pro Thr Asp Pro Thr Ala Lys Phe
180 185 190

Ile Ser Glu Val Arg Asp Gly Asn Asp Leu Lys Gly Val Ile Ala Gly
195 200 205

Asp Gly Gly Lys Val Gly Thr Gly Gly Met Ala Ser Lys Val Ser Ala
210 215 220

Ala Arg Leu Ala Ser Arg Ser Gly Val Pro Val Leu Leu Thr Ser Ala
225 230 235 240

Ala Asn Ile Gly Pro Ala Leu Glu Asp Ala Gln Val Gly Thr Val Phe
245 250 255

His Pro Lys Asp Asn Arg Leu Ser Ala Trp Lys Phe Trp Ala Leu Tyr
260 265 270

Ala Ala Asp Thr Ala Gly Lys Ile Arg Leu Asp Asp Gly Ala Val Glu
275 280 285

Ala Val Thr Ser Gly Gly Lys Ser Leu Leu Ala Val Gly Ile Thr Glu
290 295 300

Ile Ile Gly Asp Phe Gln Gln Gly Glu Ile Val Glu Ile Leu Gly Pro
305 310 315 320

Ala Gly Gln Ile Ile Gly Arg Gly Glu Val Ser Tyr Asp Ser Asp Thr
325 330 335

Leu Gln Ser Met Val Gly Met Gln Thr Gln Asp Leu Pro Asp Gly Met
340 345 350

Gln Arg Pro Val Val His Ala Asp Tyr Leu Ser Asn Tyr Ala Ser Arg
355 360 365

Ala