

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

C12N 9/28 (2006.01)

C11D 3/386 (2006.01)



# [12] 发明专利说明书

专利号 ZL 00805949.7

[45] 授权公告日 2006 年 4 月 19 日

[11] 授权公告号 CN 1252256C

[22] 申请日 2000.3.28 [21] 申请号 00805949.7  
[30] 优先权  
[32] 1999.3.30 [33] DK [31] PA199900437  
[86] 国际申请 PCT/DK2000/000148 2000.3.28  
[87] 国际公布 WO2000/060059 英 2000.10.12  
[85] 进入国家阶段日期 2001.9.30  
[71] 专利权人 诺维信公司  
地址 丹麦鲍斯韦  
[72] 发明人 卡斯滕·安德森  
克里斯特尔·T·乔根森  
亨里克·比斯加德-弗兰岑  
艾伦·斯文德森 索伦·杰拉尔夫  
审查员 徐益君

[74] 专利代理机构 北京市柳沈律师事务所  
代理人 巫肖南

权利要求书 3 页 说明书 68 页

[54] 发明名称

$\alpha$  - 淀粉酶变体

[57] 摘要

本发明涉及亲本 Termamyl - 样  $\alpha$  淀粉酶的变体, 此变体表现了改变的特性, 尤其降低了靠近分支点酶解底物的能力, 而且改进了底物特异性和/或改进了相对于亲本  $\alpha$  - 淀粉酶的比活性。

1. 一种亲本 Termamyl-样 $\alpha$ -淀粉酶的变体,其在选自下组的一或多个位置处包含改变: T49, Q51 和 G107, 其中
- 5 (a)所述每一改变是
- (i)在占据此位置的氨基酸下游插入氨基酸,
- (ii)占据此位置的氨基酸的缺失, 或
- (iii)占据此位置的氨基酸用另一氨基酸取代,
- (b)每一位置对应于具有 SEQ ID NO:4 之氨基酸序列的亲本 Termamyl-样 $\alpha$ -淀粉酶氨基酸序列的位置,
- 10 (c)所述变体具有 $\alpha$ -淀粉酶活性, 且
- (d)所述变体与亲本 Termamyl-样 $\alpha$ -淀粉酶相比, 具有降低了的在靠近分支点酶解寡糖底物的能力, 改进了的底物特异性, 改进了的比活性, 或三种特性的任意组合。
- 15 2. 权利要求 1 的变体, 其包含对应于 SEQ ID NO:4 所示氨基酸序列的至少一个下列突变:
- A52S+V54N, T49L, A52S+V54N+T49L, G107A, Q51R+A52S, T49V+G107A, T49D+G107A, T49Y+G107A, T49N+G107A, T49I+G107A, T49L+A52S+G107A, T49L+A52T+G107A, T49L+A52F+G107A,
- 20 T49L+A52L+G107A, T49L+A52I+G107A, T49L+A52V+G107A, T49F, T49L+G107A, A52S+V54N+T49L+G107A。
3. 权利要求 1-2 中任一项的变体, 其中亲本 Termamyl-样 $\alpha$ -淀粉酶是 SEQ ID NO:4 和 SEQ ID NO:6 的杂合型 $\alpha$ -淀粉酶。
4. 权利要求 1-2 中任一项的变体, 其中杂合型亲本 Termamyl-样 $\alpha$ -淀粉酶是含有 SEQ ID NO:4 中所示地衣芽孢杆菌 $\alpha$ -淀粉酶的 445 个 C-末端氨基酸残基和 SEQ ID NO:6 中所示解淀粉芽孢杆菌 $\alpha$ -淀粉酶的 37 个 N-末端氨基酸残基的杂合型 $\alpha$ -淀粉酶。
5. 权利要求 1-2 中任一项的变体, 其中杂合型亲本 Termamyl-样 $\alpha$ -淀粉酶还含有下列突变: 根据 SEQ ID NO:4 编号的
- 30 H156Y+A181T+N190F+A209V+Q264S 或 LE174。

6. 权利要求 1-2 中任一项的变体, 其中杂合型亲本 Termamyl-样 $\alpha$ -淀粉酶还含有下列突变: 根据 SEQ ID NO:4 编号的 H156Y+A181T+N190F+A209V+Q264S+I201F 或 LE429。

5 7. DNA 构建体, 其含有编码权利要求 1-6 中任一项之 $\alpha$ -淀粉酶变体的 DNA 序列。

8. 重组表达载体, 其携有权利要求 7 的 DNA 构建体。

9. 用权利要求 8 的 DNA 构建体或权利要求 8 的载体转化的细胞。

10. 权利要求 9 的细胞, 其作为一种微生物。

11. 权利要求 10 的细胞, 其中所述微生物是细菌或真菌。

10 12. 权利要求 11 的细胞, 其中所述细菌为革兰氏阳性细菌。

13. 权利要求 12 的细胞, 其中所述革兰氏阳性细菌为枯草芽孢杆菌, 地衣芽孢杆菌, 迟缓芽孢杆菌, 短芽孢杆菌, 嗜热脂肪芽孢杆菌, 嗜碱芽孢杆菌, 解淀粉芽孢杆菌, 凝结芽孢杆菌, 环状芽孢杆菌, 灿烂芽孢杆菌或苏云金芽孢杆菌。

15 14. 一种组合物, 其包括:

衍生自作为亲本 Termamyl-样 $\alpha$ -淀粉酶的杂合型 $\alpha$ -淀粉酶的一或多个如权利要求 1-6 所述变体的混合物, 其中所述杂合型 $\alpha$ -淀粉酶是由至少一种 SEQ ID NO:6 中所示解淀粉芽孢杆菌 $\alpha$ -淀粉酶和至少一种 SEQ ID NO:4 中所示地衣芽孢杆菌 $\alpha$ -淀粉酶杂交而成, 并且该杂交而成的杂合型 $\alpha$ -淀粉酶是

20 Termamyl-样 $\alpha$ -淀粉酶。

15. 权利要求 14 的组合物, 其中杂合型 $\alpha$ -淀粉酶是包含 SEQ ID NO:4 中所示地衣芽孢杆菌 $\alpha$ -淀粉酶的 445 个 C-末端氨基酸残基和 SEQ ID NO:6 中所示解淀粉芽孢杆菌 $\alpha$ -淀粉酶的 37 个 N-末端氨基酸残基的杂合型 $\alpha$ -淀粉酶。

25 16. 权利要求 15 的组合物, 其中杂合型 $\alpha$ -淀粉酶进一步含有下列突变: 根据 SEQ ID NO:4 编号的 H156Y+A181T+N190F+A209V+Q264S 或 LE174。

17. 权利要求 15 的组合物, 其中杂合型 $\alpha$ -淀粉酶进一步含有下列突变: 如 SEQ ID NO:2 所示的 H156Y+A181T+N190F+A209V+Q264S+I201F 或 LE429。

30 18. 权利要求 1-6 中任一项的 $\alpha$ -淀粉酶变体或权利要求 16-17 中任一项的组合物的应用, 所述应用为用于淀粉液化; 用于洗涤剂组合物或洗衣、

---

碗碟清洗和硬表面清洗组合物；用于乙醇生产或燃料、酒类和工业用乙醇的生产；用于织物，纤维或衣物的退浆。

## $\alpha$ -淀粉酶变体

5

### 发明领域

本发明特别涉及亲本 Termamyl-样 $\alpha$ -淀粉酶的新变体，尤其是显示改变的特性的变体，具体是改变了酶解模式(相对于亲本)，这在此类变体的应用中，特别是在工业淀粉处理中(例如，淀粉液化或糖化)是有利的。

10

### 发明背景

$\alpha$ -淀粉酶( $\alpha$ -1,4-葡聚糖-4-葡聚糖水解酶，EC 3.2.1.1)组成催化淀粉和其它直链及支链 1,4-葡糖苷寡糖和多糖水解的一组酶。

有大量的专利和科学文献涉及此类在工业上很重要的酶。许多 $\alpha$ -淀粉酶，例如 Termamyl-样 $\alpha$ -淀粉酶可从，例如 WO 90/11352，WO 95/10603，  
15 WO 95/26397，WO 96/23873，WO 96/23874 和 WO 97/41213 中了解。

在最近涉及 $\alpha$ -淀粉酶的公开内容中，WO 96/23874 提供了名为 BA2 的 Termamyl-样 $\alpha$ -淀粉酶的三维，X-射线晶体结构数据，其由解淀粉芽孢杆菌 $\alpha$ -淀粉酶(包含本文 SEQ ID NO:6 中所示氨基酸序列)的 300 个 N-末端氨基酸残基，以及地衣芽孢杆菌 $\alpha$ -淀粉酶(包含本文 SEQ ID NO:4 中所示氨基酸序列)  
20 列)C-末端的氨基酸 301-483(后者可以以商品名 Termamyl™买到)组成，它因此与工业上很重要的芽孢杆菌 $\alpha$ -淀粉酶(在本文中包括在术语“Termamyl-样 $\alpha$ -淀粉酶”的意义中，并且尤其包括地衣芽孢杆菌，解淀粉芽孢杆菌，和嗜热脂肪芽孢杆菌 $\alpha$ -淀粉酶)密切相关。WO 96/23874 进一步描述了根据对亲本 Termamyl-样 $\alpha$ -淀粉酶的结构分析来设计亲本 Termamyl-样 $\alpha$ -淀粉酶的表现  
25 出特性改变的变体的方法。

WO 96/23874 和 WO 97/41213(Novo Nordisk)公开了 Termamyl-样 $\alpha$ -淀粉酶的一种变体，其酶解模式已发生改变、并包含在本文示于 SEQ ID NO:4 的序列中氨基酸残基 V54，D53，Y56，Q333，G57 和 A52 的突变。

30

### 发明简述

本发明涉及 Termamyl-样 $\alpha$ -淀粉酶的新的 $\alpha$ -淀粉分解变体(突变体), 尤其显示出改变的酶解模式(对照于亲本)的变体, 所述改变在淀粉工业处理方面(淀粉的液化, 糖化及类似方面)是有利的。

5 本发明人意外发现一些具有改变的特性的变体, 它们与 WO 96/23874 和 WO 97/41213(Novo Nordisk)中披露的 Termamyl-样 $\alpha$ -淀粉酶变体(其酶解模式发生了改变, 并包含 SEQ ID NO:4 中所示氨基酸序列中残基 V54, D53, Y45, Q333, G57 和 A52 的突变)相比, 特性已改变, 尤其改变了酶解模式—降低了靠近分支点酶解底物的能力, 而且进一步改进了底物特异性和/或改进了比活性。

10 本发明进一步涉及编码本发明变体的 DNA 构建体, 包含本发明变体的组合物, 制备本发明变体的方法, 以及本发明变体和组合物单独或其它 $\alpha$ -淀粉分解酶联合在各种工业处理过程中的应用, 例如, 淀粉液化, 在洗涤剂组合物中, 例如洗衣, 清洗碗碟和硬表面清洗的组合物; 乙醇的生产, 例如燃料, 酒类和工业乙醇的生产; 织物, 纤维或衣物的退浆等。

15

### 命名法

在本说明和权利要求中, 使用传统的一个和三个字母表示氨基酸残基。为便于引用, 本发明的 $\alpha$ -淀粉酶变体采用下述的命名法:

原氨基酸: 位置: 用于取代的氨基酸

20 根据该命名法, 用天冬酰胺取代第 30 位的丙氨酸被示于: Ala30Asn 或 A30N, 在相同位置处缺失丙氨酸被示于: Ala30\*或 A30\*, 而插入另一个氨基酸残基, 如赖氨酸被示于: Ala30AlaLys 或 A30AK。缺失连续的一段氨基酸残基, 如氨基酸残基 30-33 被示于(30-33)\*或 $\Delta$ (A30-N33)。

25 当特定的 $\alpha$ -淀粉酶相对于其它 $\alpha$ -淀粉酶而言含有“缺失”, 并在该位置插入某个氨基酸, 例如在第 36 位插入天冬氨酸时, 被示于\*36Asp 或\*36D。

多个突变由加号隔开, 即: Ala30Asp+Glu34Ser 或 A30N+E34S 分别表示在第 30 和 34 位, 丙氨酸和谷氨酸分别用天冬酰胺和丝氨酸取代。多个突变也可以由下列与加号意义相同的符号隔开: Ala30Asp /Glu34Ser 或 A30N/E34S。

30 当在给定的位置插入一个或多个其它氨基酸残基时, 被示于 A30N, E 或 A30N 或 A30E。

另外，当本文鉴定出适于修饰的位置，而没有暗示任何具体的修饰时，应理解为可用任一氨基酸残基取代该位置存在的氨基酸残基。因此，例如，当提到修饰第 30 位的丙氨酸，但未具体说明或示于“X”时，应理解为丙氨酸可缺失，或用任一其它氨基酸，即下列任一氨基酸所取代：R, N, D, C, Q, E, G, H, I, L, K, M, F, P, S, T, W, Y, V。

### 发明详述

#### Termamyl-样 $\alpha$ -淀粉酶

众所周知，通过芽孢杆菌种产生的多种 $\alpha$ -淀粉酶在氨基酸水平上高度同源。例如，已发现含有 SEQ ID NO:4 中所示氨基酸序列的地衣芽孢杆菌 $\alpha$ -淀粉酶(商品为 Termamyl™)与含有 SEQ ID NO:5 所示氨基酸序列的解淀粉芽孢杆菌 $\alpha$ -淀粉酶约 89%同源，与含有 SEQ ID NO:3 所示氨基酸序列的嗜热脂肪芽孢杆菌 $\alpha$ -淀粉酶约 79%同源。其它同源的 $\alpha$ -淀粉酶包括衍生自芽孢杆菌菌种 NCIB 12289, NCIB 12512, NCIB 12513 或 DSM 9375 的 $\alpha$ -淀粉酶(皆详细描述于 WO 95/26397)，和描述于 Tsukamoto 等，生物化学和生物物理研究通讯, 151(1988), p25-31 中的 $\alpha$ -淀粉酶。

其它同源的 $\alpha$ -淀粉酶包括由 EP 0252666 所述的地衣芽孢杆菌菌株(ATCC 27811)所产生的 $\alpha$ -淀粉酶，和 WO 91/00353 和 WO 94/18314 中鉴定的 $\alpha$ -淀粉酶。其它可商购的 Termamyl-样地衣芽孢杆菌 $\alpha$ -淀粉酶是 Optitherm™和 Takatherm™(可衍生自 Solvay)，Maxamyl™(可衍生自 Gist-brocades/Genencor)，Spezym AA™和 Spezyme Delta AA™(可衍生自 Genencor)和 Keistase™(可衍生自 Daiwa)。

由于这些 $\alpha$ -淀粉酶基本上同源，因此可以认为它们属于同一类 $\alpha$ -淀粉酶，即“Termamyl-样 $\alpha$ -淀粉酶”。

因此，在本发明的上下文中，术语“Termamyl-样 $\alpha$ -淀粉酶”欲指在氨基酸水平上与 Termamyl™(即具有本文 SEQ ID NO:4 中所示氨基酸序列的地衣芽孢杆菌 $\alpha$ -淀粉酶)基本上同源的 $\alpha$ -淀粉酶。换句话说，Termamyl-样 $\alpha$ -淀粉酶是具有本文 SEQ ID NO:2, 4, 6 或 8 中所示氨基酸序列，和 WO 95/26397 或 Tsukamoto 等, 1988 的 SEQ ID NO:1 或 2 所示氨基酸序列的 $\alpha$ -淀粉酶，或者，Termamyl-样 $\alpha$ -淀粉酶是：i)与至少一个上述氨基酸序列的同源性至少为 60%，优选至少为 70%，更优选至少为 75%，甚至更优选至少为 80%，尤其

是至少为 85%，尤其优选至少为 90%，尤其是至少为 95%，甚至尤其更优选至少为 97%，尤其是至少为 99%的 $\alpha$ -淀粉酶，和/或 ii)能与针对一种或多种所述 $\alpha$ -淀粉酶产生的抗体发生免疫交叉反应的 $\alpha$ -淀粉酶，和/或 iii)由能与编码上述-特定 $\alpha$ -淀粉酶的 DNA 序列杂交的 DNA 序列编码的 $\alpha$ -淀粉酶，所述  
5 特定 $\alpha$ -淀粉酶的 DNA 序列分别示于本申请的 SEQ ID NO:1, 3, 5 和 7，以及 WO 95/26397 的 SEQ ID NO:4。

关于特性 i)，可利用任何常规的算法测定“同源性”(同一性)，优选使用 GCG 程序包第 7.3 版(1993 年 6 月)的缺口程序，该程序使用缺口得分的缺省值，即缺口产生得分为 3.0，缺口延伸得分为 0.1(遗传学计算机组(1991)GCG  
10 软件包程序手册，第 7 版，575 Science Drive, Madison, Wisconsin, 美国 53711)。

可使用 Termamyl 和 Termamyl-样 $\alpha$ -淀粉酶之间的结构对比来鉴定其它 Termamyl-样 $\alpha$ -淀粉酶中等同的/相应的位置。得到所述结构对比的一个方法是使用 GCG 程序包中的 Pile Up 程序，该程序使用缺省的缺口得分值，即缺口产生得分为 3.0，缺口延伸得分为 0.1。其它结构对比方法包括疏水簇分析  
15 (Gaboriaud 等, (1987), FEBS LETTERS 224, p.149-155)和反向穿梭(threading)法(Huber, T; Torda, AE, 蛋白质科学, 卷 7, 1 期, 142-149(1998))。 $\alpha$ -淀粉酶的特性 ii)，即免疫交叉反应性可利用针对或与相关 Termamyl-样 $\alpha$ -淀粉酶的至少一个表位反应的抗体进行分析。所述抗体(为单克隆或多克隆形式)可用本领域已知方法得到，如 Hudson 等，免疫学实践(Practical Immunology),  
20 第三版(1989)，Blackwell Scientific Publications 所述。免疫交叉反应性可利用本领域已知试验来测定，如 Western 印迹或放射性免疫扩散试验，如 Hudson 等，1989 所述。在这方面，在分别含有氨基酸序列 SEQ ID NOS: 2, 4, 6, 或 8 的 $\alpha$ -淀粉酶间发现了免疫交叉反应性。

可以在所述 $\alpha$ -淀粉酶的完整或部分核苷酸或氨基酸序列的基础上适当  
25 制备寡核苷酸探针，然后使用所述探针，根据上述特性 iii)鉴定 Termamyl-样 $\alpha$ -淀粉酶。

检测杂交的适当条件包括：在 5×SSC 中预浸泡，于~40℃，在 20%甲酰胺，5×Denhardt's 溶液，50mM 磷酸钠，pH6.8 和 50mg 变性的经超声处理的小牛胸腺 DNA 中预杂交 1 小时，接着于~40℃，在添加有 100mM ATP 的相  
30 同溶液中杂交 18 小时，接着于 40℃，用 2×SSC, 0.2% SDS 将滤膜洗涤 3 次，每次 30 分钟(低严紧性)，优选在 50℃(中度严紧性)，更优选在 65℃(高严紧



性), 甚至更优选在~75℃(很高的严紧性)进行洗涤。有关杂交方法的更多细节描述于 Sambrook 等, 分子克隆: 实验室手册, 第2版, 冷泉港, 1989。

在本发明的上下文中, “衍生自” 不仅仅表示由所述生物体产生或可由所述生物体产生的 $\alpha$ -淀粉酶, 还指由分离自所述菌株的 DNA 序列编码的 $\alpha$ -淀粉酶, 和由所述 DNA 序列转化的宿主生物体产生的 $\alpha$ -淀粉酶。最后, 该术语还表示由合成的和/或 cDNA 来源的 DNA 序列编码的, 并具有所述 $\alpha$ -淀粉酶的鉴定特征的 $\alpha$ -淀粉酶。该术语还表示亲本 $\alpha$ -淀粉酶可以是天然 $\alpha$ -淀粉酶的变体, 即修饰(插入, 取代, 缺失)天然 $\alpha$ -淀粉酶的一个或多个氨基酸残基而得到的变体。

#### 10 亲本杂合型 $\alpha$ -淀粉酶

亲本 $\alpha$ -淀粉酶可以是杂合型 $\alpha$ -淀粉酶, 即含有衍生自至少两种 $\alpha$ -淀粉酶的部分氨基酸序列组合的 $\alpha$ -淀粉酶。

亲本杂合型 $\alpha$ -淀粉酶可以是在氨基酸同源性和/或免疫交叉-反应性和/或 DNA 杂交(如上所述)的基础上被确定为属于 Termamyl-样 $\alpha$ -淀粉酶家族的杂合型 $\alpha$ -淀粉酶。此时, 杂合型 $\alpha$ -淀粉酶一般由 Termamyl-样 $\alpha$ -淀粉酶的至少一个部分和一种或多种其它 $\alpha$ -淀粉酶的部分组成, 所述其它 $\alpha$ -淀粉酶选自微生物(细菌或真菌)和/或哺乳动物来源的 Termamyl-样 $\alpha$ -淀粉酶或非-Termamyl-样 $\alpha$ -淀粉酶。

因此, 亲本杂合型 $\alpha$ -淀粉酶可含有衍生自至少两种 Termamyl-样 $\alpha$ -淀粉酶, 或衍生自至少一种 Termamyl-样和至少一种非-Termamyl-样细菌 $\alpha$ -淀粉酶, 或衍生自至少一种 Termamyl-样和至少一种真菌 $\alpha$ -淀粉酶的部分氨基酸序列组合。衍生得到部分氨基酸序列的 Termamyl-样 $\alpha$ -淀粉酶可以是例如本文所述的任何特定的 Termamyl-样 $\alpha$ -淀粉酶。

例如, 亲本 $\alpha$ -淀粉酶可含有衍生自地衣芽孢杆菌菌株的 $\alpha$ -淀粉酶的 C-末端部分, 和衍生自解淀粉芽孢杆菌菌株或嗜热脂肪芽孢杆菌菌株的 $\alpha$ -淀粉酶的 N-末端部分。例如, 亲本 $\alpha$ -淀粉酶可含有地衣芽孢杆菌 $\alpha$ -淀粉酶 C-末端部分的至少 430 个氨基酸残基, 也可含有如 SEQ ID NO:6 中所示解淀粉芽孢杆菌 $\alpha$ -淀粉酶的 37 个 N 端氨基酸残基的对应氨基酸片段及如 SEQ ID NO:4 中所示地衣芽孢杆菌 $\alpha$ -淀粉酶的 445 个 C 端氨基酸残基的对应氨基酸片段, 或 b) 对应于嗜热脂肪芽孢杆菌 $\alpha$ -淀粉酶(具有 SEQ ID NO:8 中所示氨基酸序列)的 68 个 N-末端氨基酸残基的氨基酸区段, 和对应于地衣芽孢杆菌

$\alpha$ -淀粉酶(具有 SEQ ID NO:4 中所示氨基酸序列)的 415 个 C-末端氨基酸残基的氨基酸区段。

在一优选实施方案中, 亲本 Termamyl-样 $\alpha$ -淀粉酶是与 SEQ ID NO:4 中所示地衣芽孢杆菌 $\alpha$ -淀粉酶相同的杂合型 Termamyl-样 $\alpha$ -淀粉酶, 不同在于: (成熟蛋白质)N-末端的 35 个氨基酸残基被 SEQ ID NO:6 中所示解淀粉芽孢杆菌 $\alpha$ -淀粉酶(BAN)的成熟蛋白 N-末端的 33 个氨基酸残基所取代。该杂合型可以进一步含有下列突变: H156Y+A181T+N190F+A209V+Q264S (使用 SEQ ID NO:4 的编号), 指 LE174。

另一优选的亲本杂合型 $\alpha$ -淀粉酶是 SEQ ID NO:6 中所示的 LE429。

10 非-Termamyl-样 $\alpha$ -淀粉酶可以是例如真菌 $\alpha$ -淀粉酶, 哺乳动物或植物 $\alpha$ -淀粉酶或细菌 $\alpha$ -淀粉酶(不同于 Termamyl-样 $\alpha$ -淀粉酶)。这种 $\alpha$ -淀粉酶的具体例子包括米曲霉 TAKA  $\alpha$ -淀粉酶, 黑曲霉酸性 $\alpha$ -淀粉酶, 枯草芽孢杆菌 $\alpha$ -淀粉酶, 猪胰腺 $\alpha$ -淀粉酶和大麦 $\alpha$ -淀粉酶。所有这些 $\alpha$ -淀粉酶都具有已阐明的, 与本文所述典型 Termamyl-样 $\alpha$ -淀粉酶的结构显著不同的结构。

15 上述真菌 $\alpha$ -淀粉酶, 即得自黑曲霉和米曲霉的 $\alpha$ -淀粉酶在氨基酸水平上高度同源, 一般认为它们属于相同的 $\alpha$ -淀粉酶家族。得自米曲霉的真菌 $\alpha$ -淀粉酶可以商购, 其商品名为 Fungamyl™。

另外, 当提及 Termamyl-样 $\alpha$ -淀粉酶的特定变体(本发明的变体)时, 提到该变体是通过常规方法, 修饰(例如缺失或取代)特定 Termamyl-样 $\alpha$ -淀粉酶氨基酸序列中的特定氨基酸残基而获得的, 应理解在另一个 Termamyl-样 $\alpha$ -淀粉酶的等同位置(由各个氨基酸序列之间最有可能的氨基酸序列对比测定)修饰得到的变体也包括在本发明的范围内。

20 本发明变体的优选实施方案是得自地衣芽孢杆菌 $\alpha$ -淀粉酶(作为亲代 Termamyl-样 $\alpha$ -淀粉酶)的变体, 例如, 上述 $\alpha$ -淀粉酶中的一个, 如具有 SEQ ID NO:4 中所示氨基酸序列的地衣芽孢杆菌 $\alpha$ -淀粉酶的变体。

#### 本发明变体的构建

可通过在适于生产变体的条件下, 培养含有编码所述变体之 DNA 序列的微生物来实现目标变体的构建。所述变体可随后从得到的培养液中回收。这进一步详述于下。

#### 30 改变的性质

下面讨论突变(本发明变体中出现的)及其导致的所需性质改变(相对于亲本 Termamyl-样 $\alpha$ -淀粉酶)的关系。

本发明第一方面涉及亲本 Termamyl-样 $\alpha$ -淀粉酶变体, 其在选自下组的一或多个位置上包含改变:

- 5           W13, G48, T49, S50, Q51, A52, D53, V54, G57, G107, G108, A111, S168, M197, 其中(a)所述每一改变是
- (i)在占据此位置的氨基酸下游插入氨基酸,
- (ii)占据此位置的氨基酸的缺失, 或
- (iii)占据此位置的氨基酸用另一氨基酸取代,
- 10          (b)所述变体具有 $\alpha$ -淀粉酶活性且(c)每一位置对应于具有 SEQ ID NO:4 之氨基酸序列的亲本 Termamyl-样 $\alpha$ -淀粉酶氨基酸序列的位置。

在优选实施方案中本发明上述变体包含一个突变, 其位置对应于 SEQ ID NO:4 中所示氨基酸序列中的至少一个下列突变:

- V54N, A52S, A52S+V54N, T49L, T49+G107A, A52S+V54N+T49L+G107A,
- 15   A52S+V54N+T49L, G107A, Q51R, Q51R+A52S, A52N; 或
- T49F+G107A, T49V+G107A, T49D+G107A, T49Y+G107A, T49S+G107A, T49N+G107A, T49I+G107A, T49L+A52S+G107A, T49L+A52T+G107A, T49L+A52F+G107A, T49L+A52L+G107A, T49L+A52I+G107A, T49L+A52V+G107A; 或
- 20   T49V, T49I, T49D, T49N, T49S, T49Y, T49F, T49W, T49M, T49E, T49Q, T49K, T49R, A52T, A52L, A52I, A52V, A52M, A52F, A52Y, A52W, V54M, G107V, G107I, G107L, G107C。

在优选实施方案中, 本发明变体在对应于 SEQ ID NO:4 中所示氨基酸序列中的下列突变的位置包含至少一个突变:

- 25           W13F, L, I, V, Y, A;
- G48A, V, S, T, I, L;
- \*48aD 或\*48aY(即 D 或 Y 的插入);
- T49X;
- \*49aX(即任一可能的氨基酸残基的插入)
- 30           S50X, 尤其 D, Y, L, T, V, I;
- Q51R,K;

- A52X, 尤其 A52S, N, T, F, L, I, V;  
D53E, Q, Y, I, N, S, T, V, L;  
V54X, 尤其 V54I, N, W, Y, F, L;  
G57S, A, V, L, I, F, Y, T;  
5 G107X, 尤其 G107A, V, S, T, I, L, C;  
G108X, 尤其 G108A, V, S, T, I, L;  
A111V, I, L;  
S168Y;  
M197X, 尤其 Y, F, L, I, T, A, G。
- 10 在优选实施方案中, 本发明变体包含 SEQ ID NO:4 所示氨基酸序列中下列位置的突变: T49X+A52X+V54N/I/L/Y/F/W+G107A, 并且可以进一步包含 G108A。
- 在优选实施方案中, 本发明变体在 SEQ ID NO:4 所示氨基酸序列中的下列位置包含至少一个突变:
- 15 T49L+G107A;  
T49I+G107A;  
T49L+G107A+V54I;  
T49I+G107A+V54I;  
A52S+V54N+T49L+G107A;  
20 A52S+V54I+T49L+G107A;  
A52S+ T49L+G107A;  
A52T+T49L+G107A;  
A52S+V54N+T49I+G107A;  
A52S+V54I+T49I+G107A;  
25 A52S+ T49I+G107A;  
T49L+G108A;  
T49I+G108A;  
T49L+G108A+V54I;  
T49I+G108A+V54I;  
30 所有上述本发明变体相对于亲本 $\alpha$ -淀粉酶具有改变了的特性(意指提高或降低的性质), 具体为至少一个下列性质: 降低了接近分支点酶解底物的能

力, 提高了底物特异性和/或提高了比活性, 改变了底物结合, 改变了热稳定性, 改变了 pH/活性关系(profile), 改变了 pH/稳定性关系, 改变了对氧化作用的稳定性, 改变了  $\text{Ca}^{2+}$  依赖性。

#### 稳定性

- 5        在本发明内容中, 在极低 pH (即 pH4-6) 下获得改变的稳定性, 具体如改进的稳定性(即更高或更低), 的重要突变(包括氨基酸取代和/或缺失), 包括列于上面“改变的特性”部分中的任一突变和下面提到的变体。

以下变体: Q360A, K; N102A, N326A, L, N190G, N190K; Y262A, K, E(采用 BAN, 即, SEQ ID N: 6, 编号)也进行 pH 稳定性测试。优选的  
10    亲本 $\alpha$ -淀粉酶可以是上述的 BA2。按照“材料与方法”部分所述测定 pH 稳定性。

#### $\text{Ca}^{2+}$ 稳定性

- 改变的  $\text{Ca}^{2+}$  稳定性意味着在  $\text{Ca}^{2+}$  耗尽(depletion)的情况下改善了酶稳定性, 即提高或降低了稳定性。在本发明内容中, 在极低 pH (即 pH4-6) 下获得  
15    改变的稳定性, 具体如改进的稳定性(即更高或更低), 的重要突变(包括氨基酸取代), 包括列于上面“改变的特性”部分中的任一突变。

#### 比活性

- 在本发明另一方面, 能获得比活性已改变的变体的重要突变, 尤其在 60-100 $^{\circ}\text{C}$ , 优选 70-95 $^{\circ}\text{C}$ , 特别是 80-90 $^{\circ}\text{C}$  提高或降低了比活性的突变, 包括  
20    列于上文“改变的特性”部分中的任一突变。

用“材料和方法”部分中所述 Phadebas<sup>®</sup>方法鉴定 LE174 和 LE429 的比活性为 16,000 NU/mg。

#### 改变的酶解模式

- 在淀粉液化过程中优选使用一种能将淀粉分子为降解长型带支链的寡糖的 $\alpha$ -淀粉酶, 而不优选使用产生较短的带支链寡糖的 $\alpha$ -淀粉酶(如传统的 Termamyl-样 $\alpha$ -淀粉酶)。短型带支链的寡糖(潘糖前体)不能被液化过程中 $\alpha$ -  
25    淀粉酶处理后使用, 或与糖化淀粉葡萄糖苷酶(葡糖淀粉酶)同时使用, 或在添加糖化淀粉葡萄糖苷酶(葡糖淀粉酶)之前使用的支链淀粉酶充分水解。故在存在潘糖前体的情况下, 葡糖淀粉酶处理后的产物混合物中含有很大比例  
30    的短型带支链的所谓极限糊精, 即三糖潘糖。潘糖的存在极大地降低了糖化产率, 这是不希望得到的。

以前有报道(US 专利 5,234,823), 当用葡糖淀粉酶和支链淀粉酶进行糖化时, 如果不在糖化前使来自于液化步骤的残余 $\alpha$ -淀粉酶失活, 其活性的存在能导致葡萄糖产量降低。典型的失活方法可以是在降低温度到 60℃ 以进行糖化之前, 于 95℃ 调 pH 到 4.7 以下。

5        这种处理对葡萄糖生产之负影响的原因还不完全清楚, 但是设想液化所用 $\alpha$ -淀粉酶(例如地衣芽孢杆菌的 Termamyl 120 L)通过水解支链淀粉中靠近和位于分支点两侧的 1,4- $\alpha$ -糖苷键而产生“极限糊精”(它们是支链淀粉酶较难处理的底物)。葡糖淀粉酶对这些极限糊精的水解导致三糖潘糖的积累, 而此糖只能被葡糖淀粉酶缓慢水解。

10        能避免该缺陷的热稳定性 $\alpha$ -淀粉酶的开发是一大进步, 因为无需单独进行失活步骤。

因此, 本发明的目标就是得到突变的 $\alpha$ -淀粉酶, 它具有适当修饰的淀粉降解特性但保留了亲本 Termamyl-样 $\alpha$ -淀粉酶的热稳定性。

相应地, 本发明涉及 Termamyl-样 $\alpha$ -淀粉酶变体, 其降低了靠近分支点  
15        酶解底物的能力, 并且进一步改进了底物特异性和/或比活性。

特定的目的是一种变体, 它酶解支链淀粉底物, 是从缩短的末端进行, 距分支点一个以上葡萄糖单位, 优选两或三个以上葡萄糖单位, 即比由野生型地衣芽孢杆菌 $\alpha$ -淀粉酶得到的距分支点更远的位置。

在此可以提及, 依照 WO 96/23874, 预计包含至少一个下列突变的变体  
20        能阻止靠近分支点的酶解:

V54L, I, F, Y, W, R, K, H, E, Q;

D53L, I, F, Y, W;

Y56W;

Q333W;

25        G57, 所有可能的氨基酸残基;

A52, 大于 A 的氨基酸残基, A52W, Y, L, F, I。

与获得本发明变体有关的目标突变包括 SEQ ID NO:4 所示地衣芽孢杆菌 $\alpha$ -淀粉酶中下列位置上的突变, 其中所述变体降低了靠近分支点酶解底物的能力, 并进一步改进了底物特异性和/或比活性: H156, A181, N190, A209,  
30        Q264 和 I201

应该强调的是，不但可以采用下面特别提及的 Termamyl-样 $\alpha$ -淀粉酶，而且也可以采用其它市售的 Termamyl-样 $\alpha$ -淀粉酶。下面所列并非这样的 $\alpha$ -淀粉酶的全部：产自 EP 0252666 中所述地衣芽孢杆菌(ATCC 27811)的 $\alpha$ -淀粉酶，在 WO 91/00353 和 WO 94/18314 中鉴定的 $\alpha$ -淀粉酶。其它市售的地衣芽孢杆菌 Termamyl-样 $\alpha$ -淀粉酶是 Optitherm™和 Takatherm™ (从 Solvay 可得)，Maxamyl(从 Gistbrocades/Genencor 可得)，Spezym AA™ Spezyme Delta AA™ (从 Genencor 可得)，和 Keistase™ (从 Daiwa 可得)。

所有 Termamyl-样 $\alpha$ -淀粉酶都适合用作制备本发明变体的骨架。

在本发明优选的实施方案中，亲本 Termamyl-样 $\alpha$ -淀粉酶是 SEQ ID NO:4 和 SEQ ID NO:6 的杂合型 $\alpha$ -淀粉酶。特别地，亲本杂合型 Termamyl-样 $\alpha$ -淀粉酶可以是这样的杂合型 $\alpha$ -淀粉酶，其包含示于 SEQ ID NO:4 的地衣芽孢杆菌 $\alpha$ -淀粉酶的 445 个 C-末端氨基酸残基和衍生自解淀粉芽孢杆菌示于 SEQ ID NO:6 的成熟 $\alpha$ -淀粉酶的 37 个 N-末端残基，它可以进一步有下列突变：H156Y+A181T+N190F+A209V+Q264S(SEQ ID NO:4 的编号)。此杂合型称为 LE174。LE174 杂合型可与另一突变 I201F 组合而形成具有如下突变 H156Y+A181T+N190F+A209V+Q264S +I201F(SEQ ID NO:4 的编号)的亲本 Termamyl-样 $\alpha$ -淀粉酶。此杂合型变体示于 SEQ ID NO:2 且用在下面的例子中，称作 LE429。

同样，LE174 或 LE429(SEQ ID NO:2)或 SEQ ID NO:4 中所示的包含一个或多个下列突变的地衣芽孢杆菌 $\alpha$ -淀粉酶可以用作骨架(采用 SEQ ID NO:4 的编号给突变编号)：

- E119C;
- S130C;
- D124C;
- R127C;
- A52 所有可能的氨基酸残基；
- S85 所有可能的氨基酸残基；
- N96 所有可能的氨基酸残基；
- V129 所有可能的氨基酸残基；
- A269 所有可能的氨基酸残基；
- A378 所有可能的氨基酸残基；

- S148 所有可能的氨基酸残基, 尤其 S148N;  
E211 所有可能的氨基酸残基, 尤其 E211Q;  
N188 所有可能的氨基酸残基, 尤其 N188S, N188P;  
M197 所有可能的氨基酸残基, 尤其 M197T, M197A, M197G, M197I,
- 5 M197L, M197Y, M197F, M197I;  
W138 所有可能的氨基酸残基, 尤其 W138Y;  
D207 所有可能的氨基酸残基, 尤其 D207Y;  
H133 所有可能的氨基酸残基, 尤其 H133Y;  
H205 所有可能的氨基酸残基, 尤其 H205H, H205C, H205R;
- 10 S187 所有可能的氨基酸残基, 尤其 S187D;  
A210 所有可能的氨基酸残基, 尤其 A210S, A210T;  
H405 所有可能的氨基酸残基, 尤其 H405D;  
K176 所有可能的氨基酸残基, 尤其 K176R;  
F279 所有可能的氨基酸残基, 尤其 F279Y;
- 15 Q298 所有可能的氨基酸残基, 尤其 Q298H;  
G299 所有可能的氨基酸残基, 尤其 G299R;  
L308 所有可能的氨基酸残基, 尤其 L308F;  
T412 所有可能的氨基酸残基, 尤其 T412A;
- 此外, SEQ ID NO:4 中所示、含至少一个下列突变的地衣芽孢杆菌 $\alpha$ -
- 20 淀粉酶可用作骨架:  
M15 所有可能的氨基酸残基;  
A33 所有可能的氨基酸残基;
- 当通过组合 LE174 与突变 I201F(SEQ ID NO:4 编号)而利用 LE429(SEQ ID NO:2 中所示)作骨架(即, 作为亲本 Termamyl-样 $\alpha$ -淀粉酶), 可根据本发
- 25 明在一或多个下列位置产生, 突变/改变, 尤其取代, 缺失和插入, 以降低靠近分支点酶解底物的能力, 改进底物特异性和/或比活性: W13, G48, T49, S50, Q51, A52, D53, V54, G57, G107, G108, A111, S168, M197(采用 SEQ ID NO:4 编号)
- 其中(a)所述每一改变是
- 30 (i)在占据此位置的氨基酸下游插入氨基酸,  
(ii)占据此位置的氨基酸的缺失, 或



(iii)占据此位置的氨基酸用另一氨基酸取代,

(b)所述变体具有 $\alpha$ -淀粉酶活性且(c)每一位置对应于具有 SEQ ID NO:4 之氨基酸序列的亲本 Termamyl-样 $\alpha$ -淀粉酶氨基酸序列的位置。

在优选实施方案中本发明的变体包含至少一个突变,其位置对应于下列

5 突变(在示于 SEQ ID NO:4 的氨基酸序列中的):

V54N, A52S, A52S+V54N, T49L, T49+G107A, A52S+V54N+T49L+G107A, A52S+V54N+T49L, G107A, Q51R, Q51R+A52S, A52N; 或

T49F+G107A, T49V+G107A, T49D+G107A, T49Y+G107A, T49S+G107A, T49N+G107A, T49I+G107A, T49L+A52S+G107A, T49L+A52T+G107A,

10 T49L+A52F+G107A, T49L+A52L+G107A, T49L+A52I+G107A, T49L+A52V+G107A; 或

T49V, T49I, T49D, T49N, T49S, T49Y, T49F, T49W, T49M, T49E, T49Q, T49K, T49R, A52T, A52L, A52I, A52V, A52M, A52F, A52Y, A52W, V54M, G107V, G07I, G107L, G107C。

15 在优选实施方案中,本发明变体在对应于 SEQ ID NO:4 中所示氨基酸序列的下列突变的位置包含至少一个突变:

W13F, L, I, V, Y, A;

G48A, V, S, T, I, L;

\*48aD 或\*48aY(即 D 或 Y 的插入);

20 T49X;

\*49aX(即任一氨基酸残基的插入)

S50X, 尤其 D, Y, L, T, V, I;

Q51R,K;

A52X, 尤其 A52S, N, T, F, L, I, V;

25 D53E, Q, Y, I, N, S, T, V, L;

V54X, 尤其 V54I, N, W, Y, F, L;

G57S, A, V, L, I, F, Y, T;

G107X, 尤其 G107A, V, S, T, I, L, C;

G108X, 尤其 G108A, V, S, T, I, L;

30 A111V, I, L;

S168Y;

M197X, 尤其 Y, F, L, I, T, A, G。

在优选实施方案中, 本发明变体包含至少一个下列突变, 其位置对应于 SEQ ID NO:4 中所示氨基酸序列中的下列突变: T49X+A52X+V54N/I/L/Y/F/W+G107A, 并且可以进一步包含 G108A。

- 5 在优选实施方案中, 本发明变体包含至少一个突变, 其位置对应于下列突变(在示于 SEQ ID NO:4 的氨基酸序列中):

- T49L+G107A;  
T49I+G107A;  
T49L+G107A+V54I;  
10 T49I+G107A+V54I;  
A52S+V54N+T49L+G107A;  
A52S+V54I+T49L+G107A;  
A52S+ T49L+G107A;  
A52T+T49L+G107A;  
15 A52S+V54N+T49I+G107A;  
A52S+V54I+T49I+G107A;  
A52S+ T49I+G107A;  
T49L+G108A;  
T49I+G108A;  
20 T49L+G108A+V54I;  
T49I+G108A+V54I;

本发明变体中的一般突变

- 优选本发明的变体除了含有上述变化外, 还含有一个或多个修饰。因此, 经修饰的 $\alpha$ -淀粉酶变体中一个或多个脯氨酸残基被非脯氨酸残基取代较为有利, 所述非脯氨酸残基可以是任何可能的天然非脯氨酸残基, 优选为丙氨酸, 甘氨酸, 丝氨酸, 苏氨酸, 缬氨酸或亮氨酸。
- 25

类似地, 优选亲代 $\alpha$ -淀粉酶被修饰的氨基酸残基中存在的一个或多个半胱氨酸残基被非半胱氨酸残基取代, 所述非半胱氨酸残基可以是例如丝氨酸, 丙氨酸, 苏氨酸, 甘氨酸, 缬氨酸或亮氨酸。

- 30 另外, 本发明的变体可以(作为唯一的修饰或与上述任何修饰联合)被修饰, 以使对应于 SEQ ID NO:4 中 185-209 的氨基酸片段中存在的一个或多个

Asp 和/或 Glu 分别被 Asn 和/或 Gln 取代。另外，用 Arg 取代对应于 SEQ ID NO: 4 中 185-209 的氨基酸片段中存在的一个或多个 Lys 残基也很有意义。

应理解本发明包含掺入了两个或多个上述修饰的变体。

另外，在本文所述的任何变体中导入点突变较为有利。

## 5      制备 $\alpha$ -淀粉酶的方法

将突变引入基因的多种方法为本领域所知。在对编码 $\alpha$ -淀粉酶的 DNA 序列的克隆进行简要讨论后，以下讨论在 $\alpha$ -淀粉酶编码序列内的特定定位点产生突变的方法。

### 克隆编码本发明 $\alpha$ -淀粉酶的 DNA 序列

10      可使用本领域众所周知的多种方法，从产生所述 $\alpha$ -淀粉酶的任何细胞或微生物中分离出编码亲代 $\alpha$ -淀粉酶的 DNA 序列。首先，应使用源自产生待研究之 $\alpha$ -淀粉酶的生物的染色体 DNA 或信使 RNA 构建基因组 DNA 和/或 cDNA 文库。然后，如果 $\alpha$ -淀粉酶的氨基酸序列是已知的，可合成同源的经标记的寡核苷酸探针，使用该探针从生产自所述生物的基因组文库中鉴定编  
15      码 $\alpha$ -淀粉酶的克隆。或者，将含有与已知 $\alpha$ -淀粉酶基因同源之序列的经标记的寡核苷酸探针用作探针，使用较低严紧度的杂交和洗涤条件鉴定编码 $\alpha$ -淀粉酶的克隆。

另一种鉴定编码 $\alpha$ -淀粉酶的克隆的方法包括：将基因组 DNA 片段插入表达载体，如质粒，用所得基因组 DNA 文库转化 $\alpha$ -淀粉酶-阴性细菌，然后  
20      将经转化的细菌铺于含有 $\alpha$ -淀粉酶底物的琼脂上，从而鉴定出表达 $\alpha$ -淀粉酶的克隆。

或者，可通过已建立的标准方法合成生产编码酶的 DNA 序列，所述方法例如 S. L. Beaucage 和 M. H. Caruthers(1981)所述的磷酰胺法或 Matthes 等(1984)所述的方法。在磷酰胺法中，可在例如自动化的 DNA 合成仪中合成寡  
25      核苷酸，对其进行纯化，退火，再连接并克隆至适当的载体。

最终，DNA 序列可以是根据标准技术，通过连接合成的，基因组的或 cDNA 来源的片段(适当时是对应于完整 DNA 序列的多个部分的片段)而生产的混合的基因组和合成来源，混合的合成和 cDNA 来源或混合的基因组和 cDNA 来源的 DNA 序列。也可以使用特定的引物，如 US 4,683,202 或 R. K.  
30      Saiki 等(1988)所述的引物经聚合酶链反应(PCR)生产 DNA 序列。

### 定点诱变

一旦分离出编码 $\alpha$ -淀粉酶的 DNA 序列，并鉴定出合乎需要的突变位点，可使用合成的寡核苷酸导入突变。这些寡核苷酸含有侧翼于所需突变位点的核苷酸序列；在合成寡核苷酸的过程中插入突变的核苷酸。在特定的方法中，在携有 $\alpha$ -淀粉酶基因的载体中产生桥连 $\alpha$ -淀粉酶-编码序列的单链 DNA 缺口。然后，使携有所需突变的合成核苷酸与该单链 DNA 的同源部分退火。然后用 DNA 聚合酶 I (Klenow 片段)补平其余缺口，使用 T4 连接酶连接构建体。该方法的特例描述于 Morinaga 等(1984)。US 4,760,025 公开了通过对盒进行微小的改变而导入编码多个突变的寡核苷酸。然而，由于 Morinaga 法可以导入多种长度的多个寡核苷酸，因此可在任何一次导入甚至更多个突变。

Nelson 和 Long(1989)描述了另一种将突变导入编码 $\alpha$ -淀粉酶的 DNA 序列的方法。所述方法包括以 3 个步骤产生含有所需突变的 PCR 片段，所述突变是通过使用化学合成的 DNA 链作为 PCR 反应的一个引物而导入的。通过用限制性内切核酸酶裂解，可从 PCR-产生的片段中分离出携有突变的 DNA 片段，并将该片段重新插入表达质粒。

#### 随机诱变

可在翻译成本文所示氨基酸序列的基因的至少 3 个部分中，或在整个基因内适当地进行随机诱变，所述诱变可以是定域的或区域-特异性的随机诱变。

利用本领域已知的任何方法，可以方便地对编码亲代 $\alpha$ -淀粉酶的 DNA 序列进行随机诱变。

关于上文，本发明的另一方面涉及产生亲代 $\alpha$ -淀粉酶的变体的方法，例如，其中变体相对于亲代而言，表现出经改变的或增加的热稳定性，所述方法包括：

- (a)对编码亲代 $\alpha$ -淀粉酶的 DNA 序列进行随机诱变，
- (b)在宿主细胞中表达步骤(a)中得到的突变 DNA 序列，和
- (c)筛选表达 $\alpha$ -淀粉酶变体的宿主细胞，所述变体相对于亲代 $\alpha$ -淀粉酶而言具有经改变的特性(如热稳定性)。

优选使用添加引物进行本发明上述方法中的步骤(a)。

例如，可使用适当的物理或化学诱变剂，使用适当的寡核苷酸，或通过 DNA 序列进行 PCR 以产生诱变来进行随机诱变。另外，可使用这些诱变

剂的任何组合进行随机诱变。诱变剂可以是例如诱导碱基转换，颠换，倒位，倒频，缺失和/或插入的试剂。

适用于本发明目的的物理或化学诱变剂的例子包括紫外线(UV)照射，羟胺，N-甲基-N'-硝基-N-亚硝基胍(MNNG)，邻甲基羟胺，亚硝酸，乙基甲磺酸(EMS)，亚硫酸氢钠，甲酸和核苷酸类似物。当使用这些诱变剂时，一般通过在适于发生诱变的条件下，在选定诱变剂的存在下保温待诱变的编码亲代酶的 DNA 序列来进行诱变，然后筛选具有所需特性的突变 DNA。

当利用寡核苷酸进行诱变时，在合成待改变位置处的寡核苷酸的过程中，可同时添加寡核苷酸和 3 个非-亲代核苷酸。添加的目的是避免不必要氨基酸的密码子。可通过任何公开的技术如使用 PCR，LCR 或适当时使用任何 DNA 聚合酶和连接酶，将添加的寡核苷酸掺入编码 $\alpha$ -淀粉酶的 DNA。

优选使用“恒随机的添加”进行添加，其中各个位置的野生型和突变的百分比是预先确定好的。另外，添加可以导致优先导入某些核苷酸，从而优先导入一个或多个特定的氨基酸残基。例如，进行添加后可在每个位置导入 90%野生型和 10%突变。选择添加方案的其它考虑基于遗传学以及蛋白质-结构的限制。可使用 DOPE 程序(见“材料和方法”部分)制订添加方案，所述程序可特别地确保不导入终止密码子。

当使用 PCR-产生的诱变时，可在增加核苷酸错误掺入的条件下，对经化学处理或未经处理的编码亲代 $\alpha$ -淀粉酶的基因进行 PCR(Deshler 1992; Leung 等, 技术, Vol.1, 1989, pp.11-15)。

可使用大肠杆菌(Fowler 等, Molec. Gen. Genet., 133, 1974, pp.179-191)，酿酒酵母或任何其它微生物的增变株对编码 $\alpha$ -淀粉酶的 DNA 进行随机诱变，诱变方法包括例如：将含有亲代糖基酶的质粒转化至增变株，培养含有质粒的增变株，从增变株中分离突变的质粒。随后，将突变的质粒转化至表达生物。

待诱变的 DNA 序列可以方便地存在于基因组或 cDNA 文库中，所述文库生产自表达亲代 $\alpha$ -淀粉酶的生物。或者，DNA 序列可以存在于适当的载体，如质粒或噬菌体上，再与诱变剂一起保温或要不然暴露于诱变剂中。待诱变的 DNA 也可存在于宿主细胞中，或者整合于所述细胞的基因组中，或者存在于细胞所携带的载体上。最后，待诱变的 DNA 也可以是分离的形式。应理解接受随机诱变的 DNA 序列优选为 cDNA 或基因组 DNA 序列。

在某些情况下可于实施表达步骤 b)或筛选步骤 c)前方便地扩增突变的 DNA 序列。可根据本领域已知的方法进行扩增，本发明优选的方法是：使用根据亲代酶的 DNA 或氨基酸序列生产的寡核苷酸引物进行 PCR 扩增。

与诱变剂保温或暴露于诱变剂之后，通过在允许发生表达的条件下培养  
5 携有突变 DNA 序列的适当宿主细胞来表达突变的 DNA。用于此目的的宿主细胞可以是被突变的 DNA 序列(任选其存在于载体上)转化的宿主细胞，或者是在诱变处理的过程中携有编码亲代酶的 DNA 序列的宿主细胞。适当宿主细胞的例子如下：革兰氏阳性细菌，如枯草芽孢杆菌，地衣芽孢杆菌，迟缓芽孢杆菌，短芽孢杆菌，嗜热脂肪芽孢杆菌，嗜碱芽孢杆菌，解淀粉芽孢杆菌，  
10 凝结芽孢杆菌，环状芽孢杆菌，灿烂芽孢杆菌，巨大芽孢杆菌，苏云金芽孢杆菌，浅青紫链霉菌或鼠灰链霉菌；和革兰氏阴性细菌，如大肠杆菌。

突变的 DNA 序列可进一步含有能使突变的 DNA 序列表达的 DNA 序列。

#### 定域随机诱变

15 随机诱变可有利地定域于所述亲代 $\alpha$ -淀粉酶的一部分。例如，当酶的某些区域被鉴定为对酶的给定特性特别重要时，以及当预期修饰能导致具有改良特性的变体时，定域随机诱变较为有利。当亲代酶的三级结构已被阐明，且所述结构与酶的功能相关时，一般可鉴定出所述区域。

通过使用上述的 PCR 产生的诱变技术或本领域已知的任何其它适当的技术，可以方便地进行定域的或区域-特异性的随机诱变。或者，通过例如插入  
20 适当的载体来分离编码待修饰的 DNA 序列部分的 DNA 序列，随后可使用上述任何诱变方法对所述部分进行诱变。

#### 提供 $\alpha$ -淀粉酶变体的其它方法

提供 $\alpha$ -淀粉酶变体的其它方法包括本领域已知的基因改组方法，所述方法  
25 包括例如 WO 95/22625(Affymax Technologies N. V.)和 WO 96/00343(Novo Nordisk A/S)所述的方法。

#### 本发明 $\alpha$ -淀粉酶变体的表达

根据本发明，可使用表达载体，以酶的形式表达通过上述方法，或通过本领域已知的任何其它方法产生的编码变体的 DNA 序列，所述表达载体一  
30 般包括编码启动子，操纵子，核糖体结合位点，翻译起始信号的控制序列，并任选包括阻抑基因或多种激活基因。

携有编码本发明 $\alpha$ -淀粉酶变体的 DNA 序列的重组表达载体可以是能对其方便地进行重组 DNA 操作的任何载体，载体的选择经常取决于导入该载体的宿主细胞。因此，载体可以是自我复制的载体，即作为染色体外实体存在的载体，所述载体的复制独立于染色体的复制，所述载体包括例如质粒，噬菌体或染色体外元件，微型染色体或人工染色体。或者，载体可以是当导入宿主细胞时可以整合至宿主细胞基因组，并与整合了该载体的染色体一起复制的载体。

在载体中，DNA 序列应该与适当的启动子序列可操作相连。启动子可以是在选定宿主细胞中显示出转录活性的任何 DNA 序列，启动子可以得自编码与宿主细胞同源或异源之蛋白质的基因。介导编码本发明 $\alpha$ -淀粉酶变体之 DNA 序列转录，尤其是在细菌宿主中的转录的适当启动子的例子是：大肠杆菌 *lac* 操纵子的启动子，天蓝色链霉菌琼脂酶基因 *dagA* 启动子，地衣芽孢杆菌 $\alpha$ -淀粉酶基因(*amyL*)启动子，嗜热脂肪芽孢杆菌产麦芽糖淀粉酶基因(*amyM*)启动子，解淀粉芽孢杆菌 $\alpha$ -淀粉酶(*amyQ*)启动子，枯草芽孢杆菌 *xylA* 和 *xylB* 基因的启动子等。为了在真菌宿主中转录，有用的启动子的例子是得自编码下列酶的基因的启动子：米曲霉 TAKA 淀粉酶，米赫根毛霉 (*Rhizomucor miehei*)天冬氨酸蛋白酶，黑曲霉中性 $\alpha$ -淀粉酶，黑曲霉酸稳定的 $\alpha$ -淀粉酶，黑曲霉葡糖淀粉酶，米赫根毛霉脂肪酶，米曲霉碱性蛋白酶，米曲霉丙糖磷酸异构酶或构巢曲霉乙酰胺酶。

本发明的表达载体也含有适当的转录终止子，在真核生物中，还含有与编码本发明 $\alpha$ -淀粉酶变体的 DNA 序列可操作相连的聚-腺苷酸化序列。终止和聚腺苷酸化序列可适当地得自与启动子相同的来源。

载体可进一步含有能使载体在所述宿主细胞中复制的 DNA 序列。所述序列的例子是质粒 pUC19, pACYC177, pUB110, pE194, pAMB1 和 pIJ702 的复制起点。

载体也可含有选择标记，例如其产物可以补偿宿主细胞缺陷的基因，如枯草芽孢杆菌或地衣芽孢杆菌的 *dal* 基因，或能赋予抗生素抗性的基因，所述抗生素抗性如氨苄青霉素，卡那霉素，氯霉素或四环素抗性。另外，载体可含有曲霉属选择标记，如 *amdS*, *argB*, *niaD* 和 *sC*，产生潮霉素抗性的标记，或者可通过 WO 91/17243 中所述的共-转化来完成选择。

尽管细胞内表达在某些方面(例如当使用某些细菌作为宿主细胞时)有利,但一般优选在细胞外进行表达。通常,本文所述的芽孢杆菌 $\alpha$ -淀粉酶含有允许表达的蛋白酶分泌至培养基的前区。必要时,该前区可被不同的前区或信号序列取代,通过取代编码各个前区的 DNA 序列可以方便地实现此目的。

分别连接编码 $\alpha$ -淀粉酶变体的本发明 DNA 构建体,启动子,终止子和其它元件,并将它们插入含有复制所需信息的适当载体的方法是本领域技术人员众所周知的(参见例如 Sambrook 等,分子克隆:实验室手册,第2版,冷泉港,1989)。

含有上文所定义的本发明 DNA 构建体或表达载体的本发明细胞可以有利地用作宿主细胞,以重组产生本发明的 $\alpha$ -淀粉酶变体。可以方便地通过将编码变体的本发明 DNA 构建体(一个或多个拷贝)整合至宿主染色体,用所述 DNA 构建体转化细胞。一般认为整合是有利的,因为整合后 DNA 序列可以在细胞中更加稳定地维持。根据常规方法,例如通过同源或异源重组可以将 DNA 构建体整合至宿主染色体。或者,可用与不同类型的宿主细胞有关的上述表达载体转化细胞。

本发明的细胞可以是高等生物,如哺乳动物或昆虫的细胞,但优选其为微生物细胞,例如细菌或真菌(包括酵母)细胞。

适合的细菌为革兰氏阳性菌,如枯草芽孢杆菌,地衣芽孢杆菌,迟缓芽孢杆菌,短芽孢杆菌,嗜热脂肪芽孢杆菌,嗜碱芽孢杆菌(*B.alkalophilus*),解淀粉芽孢杆菌,凝结芽孢杆菌,环状芽孢杆菌,灿烂芽孢杆菌,巨大芽孢杆菌,苏云金芽孢杆菌,浅青紫链霉菌或鼠灰链霉菌,或革兰氏阴性菌如大肠杆菌。细菌的转化可以被,例如原生质转化或以感受态细胞本身已知的方式使用该细胞而影响。

酵母生物可优选糖酵母属或裂殖酵母属,例如酿酒酵母。丝状真菌优选属于曲霉属,例如,米曲霉或黑曲霉。真菌细胞可用一种方法转化,该方法涉及原生质形成、原生质转化、然后细胞壁以其本身已知的方式再生。曲霉属宿主细胞转化的适合方法如 EP 238 023 所述。

在另一方面,本发明涉及生产本发明 $\alpha$ -淀粉酶变体的方法,该方法包含在有利于变体生产的条件下培养如上所述宿主细胞以及从细胞和/或培养液中回收变体。



用于培养细胞的培养基可以是适于培养所述宿主细胞，并表达本发明 $\alpha$ -淀粉酶变体的任何常规培养基。适当的培养基可以商购，或者可根据公开的配方(例如美国典型培养物保藏中心目录中所述的配方)生产。

5 通过众所周知的方法，包括通过离心或过滤将培养基与细胞分开，利用诸如硫酸铵等盐沉淀培养基中的蛋白质成分，接着利用层析法，如离子交换层析，亲和层析等，可以从培养基中方便地回收宿主细胞分泌的 $\alpha$ -淀粉酶变体。

#### 工业化应用

10 本发明的 $\alpha$ -淀粉酶变体具有能进行多种工业应用的有价值的特性。具体地，本发明酶变体适于作为洗衣，碗碟清洗，硬表面清洗的洗涤剂组合物的成分来应用。多种变体可特别有效用于从淀粉生产增甜剂和乙醇中，如燃料，酒类或工业用乙醇，和/或用于织物退浆。传统淀粉转化工艺包括淀粉液化和/或糖化加工的条件描述于，例如 US 3, 912, 590 和 EP 专利出版物 252 730 和 63 909。

#### 15 从淀粉生产增甜剂

将淀粉转化为果糖浆的“传统”过程通常由三个连续的酶处理步骤构成，即液化步骤，然后是糖化步骤和异构化步骤。在液化期间，淀粉在 pH 5.5-6.2, 95-160°C 下用 $\alpha$ -淀粉酶(例如 Termamyl™)降解约 2 小时而成为糊精。为确保在这些条件下优化酶稳定性，添加了 1mM 钙(40ppm 游离钙离子)。

20 液化步骤之后，通过添加葡糖淀粉酶(例如，AMG™)和脱支酶如异淀粉酶或支链淀粉酶(例如，Promozyme™)可使糊精转化为葡萄糖。在此步骤之前，将 pH 调到 4.5 以下，保持较高温度(95°C 以上)，使液化 $\alpha$ -淀粉酶失活。温度降低到 60°C，加入葡糖淀粉酶和脱支酶。糖化步骤持续 24-72 小时。

25 糖化步骤后，pH 升高到 6-8，优选 pH7.5，钙通过离子交换除去。接着使用例如固定化葡糖异构酶(如 Sweetzyme™)将葡萄糖浆转化为高果糖浆。

此步骤可至少有一个酶性改进：降低液化 $\alpha$ -淀粉酶的钙依赖性。需要加入游离钙以确保 $\alpha$ -淀粉酶的高度稳定性，但是游离钙会强烈限制葡糖异构酶的活性，且需要通过一种昂贵的单元操作除去到一定程度(游离钙水平在 3-5ppm 以下)。如果能避免使用这类操作就可以节约费用，液化步骤可以在未  
30 加入游离钙离子的情况下进行。

为达到这一点,需要一种在低游离钙浓度(<40ppm)下仍保持稳定和高活性的钙依赖性较小的 Termamyl-样 $\alpha$ -淀粉酶。这类 Termamyl-样 $\alpha$ -淀粉酶的最佳 pH 范围应为 4.5-6.5, 优选为 4.5-5.5。

本发明还涉及一种组合物, 其包括, 衍生自具有 SEQ ID NO:8 中所示序列之嗜热脂肪芽孢杆菌 $\alpha$ -淀粉酶(作为亲本 Termamyl-样 $\alpha$ -淀粉酶)的一或多个本发明变体与衍生自具有 SEQ ID NO:4 中所示序列的地衣芽孢杆菌 Termamyl-样 $\alpha$ -淀粉酶的混合物。

此外, 本发明还涉及一种组合物, 其包括, 衍生自具有 SEQ ID NO:8 中所示序列之嗜热脂肪芽孢杆菌 $\alpha$ -淀粉酶(作为亲本 Termamyl-样 $\alpha$ -淀粉酶)的一或多种本发明变体, 以及包含 SEQ ID NO:6 中所示解淀粉芽孢杆菌 $\alpha$ -淀粉酶的一部分和 SEQ ID NO:4 中所示地衣芽孢杆菌 $\alpha$ -淀粉酶的一部分的杂合型 $\alpha$ -淀粉酶。后一杂合型 Termamyl-样 $\alpha$ -淀粉酶包含 SEQ ID NO:4 中所示地衣芽孢杆菌 $\alpha$ -淀粉酶的 445 个 C-末端氨基酸残基和 SEQ ID NO:6 中所示解淀粉芽孢杆菌 $\alpha$ -淀粉酶的 37 个 N-末端氨基酸残基。该后一杂合型 $\alpha$ -淀粉酶优选含有下列突变: H156Y+A181T+N190F+A209V+Q264S(采用 SEQ ID NO:4 编号)。优选地, 这后一杂合型 $\alpha$ -淀粉酶最好含有下列突变: H156Y+A181T+N190F+A209V+Q264S+I201F(采用 SEQ ID NO:4 编号)。在以下实例中, 被称为 LE429(示于 SEQ ID NO:2)的该最后提及的亲本杂合型 Termamyl-样 $\alpha$ -淀粉酶用于制备本发明的变体, 这些变体可用于本发明的组合物中。

本发明的 $\alpha$ -淀粉酶变体或本发明的组合物在本发明的一个方面中可用于淀粉液化, 在洗涤剂组合物如洗衣、清洗碗碟和硬表面清洗的洗涤剂组合物中使用, 用于乙醇生产, 如燃料, 酒类和工业用乙醇的生产, 用于织物, 纤维和衣物的退浆。

## 25 材料和方法

### 酶:

LE174: 杂合型 $\alpha$ -淀粉酶变体:

LE174 是与 Termamyl 序列即, 示于 SEQ ID NO:4 的地衣芽孢杆菌 $\alpha$ -淀粉酶相同的杂合型 Termamyl-样 $\alpha$ -淀粉酶, 不同处在于其(成熟蛋白)N-末端 35 个氨基酸残基被 BAN, 即示于 SEQ ID NO:6 的解淀粉芽孢杆菌 $\alpha$ -淀粉酶(成熟蛋白)的 N-末端 33 个氨基酸残基取代。该 LE174 进一步具有

下列突变：H156Y+A181T+N190F+A209V+Q264S (SEQ ID NO:4)。

LE429：杂合型 $\alpha$ -淀粉酶变体：

与 Termamyl 序列即，示于 SEQ ID NO:4 的地衣芽孢杆菌 $\alpha$ -淀粉酶相同的杂合型 Termamyl-样 $\alpha$ -淀粉酶，不同处在于其(成熟蛋白)N-末端 35 个氨基酸残基被 BAN，即示于 SEQ ID NO:6 的解淀粉芽孢杆菌 $\alpha$ -淀粉酶(成熟蛋白)的 N-末端 33 个氨基酸残基取代。该 LE429 进一步具有下列突变：H156Y+A181T+N190F+A209V+Q264S+I201F (SEQ ID NO:4)。LE429 示于 SEQ ID NO:2，其通过 SOE-PCR (Higuchi 等,1988,核酸研究 16:7351)构建。

Dextrozyme™ E：衍生自黑曲霉和 *Bacillus deramificans*(可衍生自 Novo Nordisk A/S)所选菌株的葡糖淀粉酶(AMG)和支链淀粉酶的平衡混合物。

#### $\alpha$ -淀粉酶变体的发酵和纯化

携有相关表达质粒的一株枯草芽孢杆菌菌株在含有-80°C 贮存的 10 微克/毫升卡那霉素的 LB-琼脂平板上划线接种，并在 37°C 下过夜培养。将菌落转移到装在 500 ml 摇瓶中的添加有 10 微克/毫升卡那霉素的 100 ml BPX 培养基中。

BPX 培养基的组成：

|    |                                                       |         |
|----|-------------------------------------------------------|---------|
|    | 马铃薯淀粉                                                 | 100g/l  |
|    | 大麦粉                                                   | 50g/l   |
|    | BAN 5000 SKB                                          | 0.1g/l  |
| 20 | 酪蛋白酸钠                                                 | 10 g/l  |
|    | 大豆粉(Soy Bean Meal)                                    | 100g/l  |
|    | Na <sub>2</sub> HPO <sub>4</sub> · 12H <sub>2</sub> O | 9 g/l   |
|    | Pluronic™                                             | 0.1 g/l |

培养物在 37°C 下以 270rpm 振荡培养 5 天。

通过 4500rpm 离心 20-25 分钟除去发酵液中的细胞和细胞碎片。然后将上清液过滤来获得完全澄清的溶液。浓缩滤液并在 UF 过滤器上进行洗涤(10000 截留膜)，再将缓冲液换成 20mM 醋酸(Acetate) pH5.5。将 UF 滤液上样到 S-Sepharose F.F 上并且在同样的缓冲液中通过采用 0.2MNaCl 的逐步洗脱法进行洗脱。洗脱液对 10 mM Tris pH9.0 透析，上样到 Q-Sepharose F.F 上，利用 0-0.3M NaCl 的线性梯度洗脱 6 个柱体积。收集具有活性(通过 Phadebas 分析

法测定)的馏分,调节 pH 到 pH7.5 并通过采用 0.5%W/vol.的活性碳处理 5 分钟来除去残留色素。

#### 活性测定—(KNU)

5 一千个 $\alpha$ -淀粉酶单位(1 KNU)是指,用 Novo Nordisk 的在下列条件下测定 $\alpha$ -淀粉酶的标准方法所测,每小时分解 5.26g 淀粉(Merck, Amylum Solubile, Erg. B 6, 批号 9947275)所需的酶量:

|        |         |
|--------|---------|
| 底物     | 可溶性淀粉   |
| 溶剂中钙含量 | 0.0043M |
| 反应时间   | 7-20 分钟 |
| 10 温度  | 37℃     |
| pH     | 5.6     |

只要提出要求就可得到 Novo Nordisk 分析方法(AF9)的详细描述。

#### $\alpha$ -淀粉酶活性的测定

15 通过以 Phadebas<sup>®</sup>片为底物的方法测定 $\alpha$ -淀粉酶活性。Phadebas 片(Phadebas<sup>®</sup>淀粉酶试验,由 Pharmacia Diagnostic 提供)含有交联的不可溶蓝色淀粉聚合物,其与牛血清白蛋白和一种缓冲物质相混合并被压片。

20 每次测定时,将一个片剂悬浮在含有 5ml 50mM Britton-Robinson 缓冲液(50mM 醋酸、50mM 磷酸、50mM 硼酸、0.1mM  $\text{CaCl}_2$ ,用 NaOH 将 pH 调到所需值)的试管中。试验在所需温度的水浴中进行。要测定的 $\alpha$ -淀粉酶用 x ml 的 50mM Britton-Robinson 缓冲液稀释。1 ml 的该 $\alpha$ -淀粉酶溶液被加到 5ml 50mM Britton-Robinsond 缓冲液中。淀粉被该 $\alpha$ -淀粉酶水解并产生可溶的蓝色片段。在 620nm 下分光光度法测定蓝色溶液的吸光度,其是 $\alpha$ -淀粉酶活性的函数。

25 重要的是经 10 或 15 分钟温育(试验时间)后在 620nm 下测定的吸光度是在 0.2 至 2.0 个 620nm 光吸收单位的范围内。在这一光吸收范围内活性与光吸收值之间成线性关系(Lambert-Beer 定律)。因此必须调节酶的稀释度以适合这一标准。在一系列特定的条件(温度、pH、反应时间、缓冲液)下,1 mg 的给定 $\alpha$ -淀粉酶将水解一定量的底物并产生蓝色。在 620nm 下测定颜色的强度。测定的吸光值直接与在所述一系列特定条件下测得的目标 $\alpha$ -淀粉酶的比  
30 活性成比例(活性/mg 纯 $\alpha$ -淀粉酶蛋白)。

#### 比活性的测定

采用 Phadebas 分析方法(Pharmacia)将比活性测定为活性/mg 酶。

#### 测定 pH 活性图谱(pH 稳定性)

将变体保存在 20mM TRIS pH7.5, 0.1mM, CaCl<sub>2</sub> 中, 检测在 30℃, 50mM Britton-Robinson, 0.1mM CaCl<sub>2</sub> 中进行。使用上述 Phadebas 试验在 pH4.0, 5 4.5, 5.0, 5.5, 6.0, 7.0, 8.0, 9.5, 9.5, 10, 和 10.5 下测定 pH 活性。

#### 测定 AGU 活性并表示为 AGU/mg

一个 Novo 淀粉葡萄糖苷酶单位(AGU)定义为 37℃, pH4.3 下每分钟水解 1 微摩尔麦芽糖的酶量。此分析方法的详述(AEL-SM-0131)可要求 Novo Nordisk 提供。

10 用 Boehringer Mannheim, 124036 的葡萄糖 GOD-Perid 试剂盒以修饰的 (AEL-SM-0131)方法将活性为测定 AGU/ml。标准品: AMG-标准品, 批次 7-1195, 195 AGU/ml。

375μL 底物(50mM 乙酸钠中加入 1%麦芽糖, pH4.3)37℃温育 5 分钟。添加 25μL 的酶的乙酸钠稀释液。10 分钟后, 加入 100μL 0.25M NaOH 终止 15 反应。20μL 转到一 96 孔微滴板上, 加入 200μL GOD-Perid 溶液。室温 30 分钟后, 测定 650nm 吸光度, 参照 AMG 标准品以 AGU/mg 计算活性。

活性(AGU/ml)除以蛋白浓度(mg/ml)可算得以 AGU/mg 表示的比活性。

#### 实施例

##### 实施例 1

20 本发明 Termamyl 变体的构建

Termamyl(地衣芽孢杆菌α-淀粉酶 SEQ ID NO:4)由枯草芽孢杆菌中的质粒 pDN1528 表达。此质粒包含了编码了 Termamyl 的完整基因, *amyl*, 其表达其自身启动子指导。此外, 该质粒包含质粒 pUB110 的复制起点 *ori* 和来自质粒 pC194 的赋予氯霉素抗性的 *cat* 基因。pDN1528 示于 WO 96/23874 25 的图 9 中。制备含有 SEQ ID NO:3 编码区之主要部分的特异性诱变载体。该载体 pJeEN1 的重要特征包括来自 pUC 质粒的复制起点, 赋予对氯霉素抗性的 *cat* 基因, *bla* 基因的移码型, 所述 *bla* 基因的野生型(*amp<sup>R</sup>*表型)通常赋予氨苄青霉素抗性。此突变型产生 *amp<sup>S</sup>*表型。质粒 pJeEN1 示于 WO 96/23874 的图 10 中, 大肠杆菌复制起点 *ori*, *bla*, *cat*, Termamyl 淀粉酶基因的 5' - 30 截短型, 以及所选限制性位点都指示在该质粒上。

按照 Deng 和 Nickoloff (1992, Anal. Biochem. 200, 81-88 页)所述方法将突变引入 *amyl*, 不同处在于, 带有“选择引物”(引物 #6616; 见下面)的质粒根据转化的大肠杆菌细胞(该细胞携有含修复的 *bla* 基因的质粒)的 *amp<sup>R</sup>* 表型来选择, 而不采用 Deng 和 Nickoloff 所述的利用限制性酶消化的选择方法。

- 5 用于诱变的化学品和酶由 Stratagene 的 Chameleon<sup>®</sup> 诱变试剂盒(目录号 200509)得到。

在确认变异质粒中的 DNA 序列后, 含有所需改变的截短型基因被作为 *PstI-EcoRI* 片段亚克隆至 pDN1528, 并转化至蛋白酶和淀粉酶已耗尽的枯草芽孢杆菌 SHA273 株(述于 WO92/11357 和 WO95/10603 中)以表达这种酶变体。

10

通过使用下列诱变引物(5' 到 3' 从左到右书写), 构建 Termamyl 变体 V54W:

PG GTC GTA GGC ACC GTA GCC CCA ATC CGC TTG (SEQ ID NO:9)

- 15 通过使用下列诱变引物(5' 到 3' 从左到右书写), 构建 Termamyl 变体 A52W+V54W:

PG GTC GTA GGC ACC GTA GCC CCA ATC CCA TTG GCT CG (SEQ ID NO:10)

引物#6616(5' 到 3' 从左到右书写; P 表示 5' 磷酸):

P CTG TGA CTG GTG AGT ACT CAA CCA AGT C (SEQ ID NO:11)

- 20 通过使用下列诱变引物(5' 到 3' 从左到右书写), 构建 Termamyl 变体 V54E:

PGG TCG TAG GCA CCG TAG CCC TCA TCC GCT TG (SEQ ID NO:12)

- 25 通过使用下列诱变引物(5' 到 3' 从左到右书写), 构建 Termamyl 变体 V54M:

PGG TCG TAG GCA CCG TAG CCC ATA TCC GCT TG (SEQ ID NO:13)

通过使用下列诱变引物(5' 到 3' 从左到右书写), 构建 Termamyl 变体 V54I:

- 30 PGG TCG TAG GCA CCG TAG CCA ATA TCC GCT TG (SEQ ID NO:14)

通过使用下列诱变引物(5' 到 3' 从左到右书写), 构建 Termamyl 变体 Y290E 和 Y290K:

PGC AGC ATG GAA CTG CTY ATG AAG AGG CAC GTC AAA  
C(SEQ ID NO:15)

- 5 Y 代表 C 和 T 的等量混合物。在位置 290 出现的编码谷氨酸或赖氨酸的密码子由 DNA 测序证实。

通过使用下列诱变引物(5' 到 3' 从左到右书写)构建 Termamyl 变体 N190F:

PCA TAG TTG CCG AAT TCA TTG GAA ACT TCC C(SEQ ID NO:16)

- 10 通过使用下列诱变引物(5' 到 3' 从左到右书写)构建 Termamyl 变体 N188P+N190F:

PCA TAG TTG CCG AAT TCA GGG GAA ACT TCC CAA TC(SEQ ID NO:17)

- 15 通过使用下列诱变引物(5' 到 3' 从左到右书写)构建 Termamyl 变体 H140K+H142D:

PCC GCG CCC CGG GAA ATC AAA TTT TGT CCA GGC TTT AAT  
TAG (SEQ ID NO:18)

通过使用下列诱变引物(5' 到 3' 从左到右书写), 构建 Termamyl 变体 H156Y:

- 20 PCA AAA TGG TAC CAA TAC CAC TTA AAA TCG CTG(SEQ ID NO:19)

通过使用下列诱变引物(5' 到 3' 从左到右书写), 构建 Termamyl 变体 A181T:

PCT TCC CAA TCC CAA GTC TTC CCT TGA AAC(SEQ ID NO:20)

- 25 通过使用下列诱变引物(5' 到 3' 从左到右书写), 构建 Termamyl 变体 A209V:

PCTT AAT TTC TGC TAC GAC GTC AGG ATG GTC ATA ATC(SEQ ID NO:21)

- 30 通过使用下列诱变引物(5' 到 3' 从左到右书写), 构建 Termamyl 变体 Q264S:

PCG CCC AAG TCA TTC GAC CAG TAC TCA GCT ACC GTA  
AAC(SEQ ID NO:22)

通过使用下列诱变引物(5' 到 3' 从左到右书写), 构建 Termamyl 变体 S187D:

5 PGC CGT TTT CAT TGT CGA CTT CCC AAT CCC(SEQ ID NO:23)

通过使用下列诱变引物(5' 到 3' 从左到右书写), 构建 Termamyl 变体  $\delta$  (K370-G371-D372)(即 370, 371, 372 号氨基酸缺失):

PGG AAT TTC GCG CTG ACT AGT CCC GTA CAT ATC CCC(SEQ ID NO:24)

10 通过使用下列诱变引物(5' 到 3' 从左到右书写), 构建 Termamyl 变体  $\delta$  (D372-S373-Q374):

PGG CAG GAA TTT CGC GAC CTT TCG TCC CGT ACA TAT C(SEQ ID NO:25)

15 用能切割 pDN1528 质粒两次, 产生 1116bp 的片段和 3850bp 的载体部分(即包含该质粒的复制起点)的限制酶 *Cla*I, 消化含有 A181T 的 pDN1528 质粒样质粒(即在 *amyl* 内含有导致 A181T 改变之突变的 pDN1528)和含有 A209V 的 pDN1528 质粒样质粒(即在 *amyl* 内含有导致 A209V 改变之突变的 pDN1528), 可将 Termamyl 变体 A181T 和 A209V 组合为 A181T+A209V。包含 A209V 突变的片段和包含 A181T 突变的载体部分在琼脂糖胶上分离  
20 后, 用 QIAquick 凝胶提取试剂盒(购买自 QIAGEN)纯化。将所述片段和载体连接并转化至上述蛋白酶和淀粉酶耗尽型枯草芽孢杆菌菌株。分析 *amy*+(含淀粉琼脂平板上的透明区)及氯霉素抗性转化体的质粒上两个突变是否并存。

25 以上述相似的方式, 利用限制性内切酶 *Acc*65I 和 *Eco*RI 将 H156Y 和 A209V 相结合, 得到 H156Y+A209V。

利用限制性内切酶 *Acc*65I 和 *Hind*III 将 H156Y+A209V 和 A181T+A209V 结合为 H156Y+ A181T+A209V。

30 通过 SOE-PCR 法(Higuchi 等 1988, 核酸研究 16: 7351)将 Termamyl 变体 H156Y+ A181T+A209V 的成熟部分的 35 个 N-末端残基用解淀粉芽孢杆菌 $\alpha$ -淀粉酶(SEQ ID NO:4)(在此上下文中称为 BAN)33 个 N-末端残基取代, 如下所示:



引物 19364(5' -3' 排列): CCT CAT TCT GCA GCA GCA GCC GTA  
AAT GGC ACG CTG (SEQ ID NO:26)

引物 19362: CCA GAC GGC AGT AAT ACC GAT ATC CGA TAA  
ATG TTC CG (SEQ ID NO:27)

5 引物 19363: CGG ATA TCG GTA TTA CTG CCG TCT GGA TTC (SEQ  
ID NO:28)

引物 1C: CTC GTC CCA ATC GGT TCC GTC (SEQ ID NO:29)

采用来自 Boehringer Mannheim 的 Pwo 热稳定聚合酶按照生产商的指  
导进行标准的 PCR, 聚合酶链式反应, 温度循环为: 94℃ 5 分钟, (94℃ 30  
10 秒, 50℃ 45 秒, 72℃ 1 分钟)的 25 个循环, 72℃ 10 分钟。

在首轮 PCR (PCR1)中, 用引物 19364 和 19362 在含有编码解淀粉芽孢  
杆菌 $\alpha$ -淀粉酶基因的 DNA 片段上对约 130bp 的片段进行扩增。

在另一轮 PCR (PCR2)中, 用引物 19363 和 1C 在模板 pDN1528 上对约  
400bp 的片段进行扩增。

15 PCR1 和 PCR2 用琼脂糖凝胶纯化, 并在 PCR3 中用作模板, 以引物 19364  
和 1C 获得约 520bp 的片段。此片段因此包含编码 BAN N-末端的一段 DNA,  
所述 BAN 与编码 Termamyl 的 DNA 的一部分从第 35 个氨基酸融合。

通过限制性内切酶 *Pst*I 和 *Sac*II 消化, 如上所述进行枯草芽孢杆菌的连  
接以及转化, 将该 520bp 片段亚克隆至 pDN1528 样质粒(包含编码 Termamyl  
20 变体 H156Y+ A181T+A209V 的基因)。在来自 amy+及氯霉素抗性转化体的  
提纯质粒中进行 DNA 测序, 以确定在限制性位点 *Pst*I 和 *Sac*II 间的 DNA 序  
列。

最终的构建体包含来自 BAN 的正确的 N-末端和 H156Y+A181T+  
A209V, 将其表示为 BAN(1-35)+H156Y+A181T+A209V。

25 通过进行上述诱变, 但用编码 Termamyl 变体 BAN(1-35)+H156Y+  
A181T+A209V 的 DNA 序列取代 pJeEN1 中的 amyL 序列, 而将 N190F 与  
BAN(1-35)+H156Y+A181T+A209V 组合为 BAN(1-35)+H156Y+A181T+  
N190F+A209V。

通过进行上述诱变, 但将 pJeEN 中的 amyL 序列变为编码 Termamyl 变  
30 体 BAN(1-35)+H156Y+A181T+A209V 的 DNA 序列, 而将 Q264S 与

BAN(1-35)+H156Y+A181T+A209V 组合为 BAN(1-35)+H156Y+A181T+A209V+Q264S。

利用限制性内切酶 *Bsa*HI(将 *Bsa*HI 位点引至 A209V 突变附近)和 *Pst*I, 将 BAN(1-35)+H156Y+A181T+A209V+Q264S 和 BAN(1-35)+H156Y  
5 +A181T+N190F+A209V 组合为 BAN(1-35)+H156Y+A181T+N190F+A209V+Q264S。

通过进行上述诱变使 I201F 与 BAN(1-35)+H156Y+A181T+N190F+A209V+Q264S 组合为 BAN(1-35)+H156Y+A181T+N190F+A209V+Q264S+I201F(SEQ ID NO:2)。使用诱变引物 AM100, 引入 I201F 取代, 同时除去  
10 *Cla*I 限制性位点, 其有利于突变体的简便的稳定定位(pin-point)。

引物 AM100:

5' GATGTATGCCGACTTCGATTATGACC 3' (SEQ ID NO:30)

## 实施例 2

本发明具有改变的酶解模式的 Termamyl-样 $\alpha$ -淀粉酶变体的构建

15 热稳定的地衣芽孢杆菌 $\alpha$ -淀粉酶变体含有示于 SEQ ID NO:4 的地衣芽孢杆菌 $\alpha$ -淀粉酶 445 个 C-末端氨基酸残基和示于 SEQ ID NO:6 的解淀粉芽孢杆菌 $\alpha$ -淀粉酶 37 个 N-末端氨基酸残基, 并且进一步含有下列突变: H156Y+A181T+N190F+A209V+Q264S+ I201F(此变体的构建如实施例 1 所述, 氨基酸序列示于 SEQ ID NO:2), 此变体在接近分支点处酶解底物的能力  
20 减小。

在进一步改进该 $\alpha$ -淀粉酶变体在接近分支点处减小的酶解底物能力的尝试中, 采用 Sarkar 和 Sommer,1990(生物技术(BioTechniques) 8: 404-407)所述的大范围引物(mega-primer)法进行定点诱变。

25 LE313 的构建: BAN/Termamyl 杂合型 + H156Y+A181T+N190F+A209V+Q264S+V54N:

用基因特异性引物 27274 和诱变引物 AM115 通过 PCR 扩增 pDN1528 样质粒(在编码 SEQ ID NO:4 的淀粉酶的基因中携有突变 BAN(1-35)+H156Y+A181T+N190F+I201F+A209V+Q264S)中约 440bp 的 DNA 片段。

30 此 440bp 片段经琼脂糖凝胶纯化, 并作为大范围引物与引物 113711 一起以相同模板进行第二次 PCR 中。

所得约 630bp 的片段经限制性内切酶 EcoR V 和 Acc65 I 消化, 所得约 370bp 的 DNA 片段经纯化并连接至以相同酶消化的 pDN1528 样质粒。枯草芽孢杆菌 SHA273(低淀粉酶和蛋白酶)的感受态细胞用该连接物转化, 通过 DNA 测序检测氯霉素抗性转化体, 以证实质粒上存在正确的突变。

5 引物 27274:

5' CATAGTTGCCGAATTCATTGGAACTTCCC 3' (SEQ ID NO:31)

引物 1B:

5' CCGATTGCTGACGCTGTTATTTGC 3' (SEQ ID NO:32)

引物 AM115:

10 5' GCCAAGCGGATAACGGCTACGGTGC 3' (SEQ ID NO:33)

LE314 的构建: BAN/Termamyl 杂合型+ H156Y+A181T+N190F+A209V +Q264S+A52S 以相似的方式进行, 但使用的诱变引物是 AM116。

AM116:

5' GAACGAGCCAATCGGACGTGGGCTACGG 3' (SEQ ID NO:34)

15 LE315 的构建: BAN/Termamyl 杂合型+ H156Y+A181T+N190F+A209V +Q264S+A52S+V54N 以相似的方式进行, 但使用的诱变引物是 AM117。

AM117:

5' GGAACGAGCCAATCGGATAACGGCTACGGTGC 3' (SEQ ID NO:35)

20 LE316 的构建: BAN/Termamyl 杂合型+ H156Y+A181T+N190F+A209V +Q264S+T49L 以相似的方式进行, 但使用的诱变引物是 AM118。

AM118:

5' GCATATAAGGGACTGAGCCAAGCGG 3' (SEQ ID NO:36)

LE317 的构建: BAN/Termamyl 杂合型+ H156Y+A181T+N190F+A209V +Q264S+T49L+G107A 以相似的方式进行, 但同时使用诱变引物 AM118 和 AM119。

AM119:

5' CAACCACAAAGCCGGCGCTGATGCG 3' (SEQ ID NO:37)

30 LE318 的构建: BAN/Termamyl 杂合型+ H156Y+A181T+N190F+A209V +Q264S+A52S+V54N+T49L+G107A 以相似的方式进行, 但同时使用诱变引物 AM120 和 AM119。

AM120:

5' GCATATAAGGGACTGAGCCAATCGGATAACGGCTACGGTGC  
3' (SEQ ID NO:38)

LE319 的构建: BAN/Termamyl 杂合型+ H156Y+A181T+N190F+A209V  
5 +Q264S+A52S+V54N+T49L 以相似的方式进行, 但使用诱变引物 AM120。

LE320 的构建: BAN/Termamyl 杂合型+ H156Y+A181T+N190F+A209V  
+Q264S+G107A 以相似的方式进行, 但使用诱变引物 AM119。

LE322 的构建: BAN/Termamyl 杂合型+ H156Y+A181T+N190F+A209V  
+Q264S+Q51R+A52S 以相似的方式进行, 但使用诱变引物 AM121。

10 AM121:

5' GAACGAGCCGATCGGACGTGGGCTACGG 3' (SEQ ID NO:39)

LE323 的构建: BAN/Termamyl 杂合型+ H156Y+A181T+N190F+A209V  
+Q264S+A52N 以相似的方式进行, 但使用诱变引物 AM122。

AM122:

15 5' GAACGAGCCAAAACGACGTGGGCTACGG 3' (SEQ ID NO:40)

### 实施例 3

#### LE429 变体的检验(糖化):

标准反应条件:

|    |            |               |
|----|------------|---------------|
|    | 底物浓度       | 30% w/w       |
| 20 | 温度         | 60℃           |
|    | 初始 pH(60℃) | 5.5           |
|    | 酶剂量        |               |
|    | 葡糖淀粉酶      | 0.18 AGU/g DS |
|    | 支链淀粉酶      | 0.06 PUN/g DS |
| 25 | α-淀粉酶      | 10 μg 酶/g DS  |

Dextrozyme™ E 用于提供葡糖淀粉酶和支链淀粉酶活性。

糖化底物的制备如下, 在去离子水中溶解普通玉米淀粉, 调整至干物质含量约 30% w/w。pH 调到 5.5(在 60℃测定), 将相当于 10g 干重的底物等分样移入蓝盖玻璃摇瓶中。

30 然后将摇瓶置于 60℃平衡的振动水浴中, 加入酶。如果需要, 再次将 pH 调整至 5.5。糖化 48 小时后取样; 将 pH 调整至 3.0, 然后沸水浴 15 分钟

使酶失活。冷却后，样品用 0.1g 混合床离子交换树脂(BIO-RAD 501 X8 (D)) 在旋转混合器上处理 30 分钟，除去盐和可溶性 N。过滤后，经 HPLC 测定糖组成。得到下面结果：

此变体的亲本 $\alpha$ -淀粉酶是 LE 429。

| 加入的 $\alpha$ -淀粉酶变体  | DP1       | DP2       | DP3       | 比活性<br>(NU/mg) |
|----------------------|-----------|-----------|-----------|----------------|
| V54N                 | 96.1      | 1.75      | 1.18      | 8200           |
| A52S                 | 95.9      | 1.80      | 1.11      | 18800          |
| A52S+V54N            | 96.3      | 1.84      | 1.08      | 10000          |
| T49L                 | 96.3      | 1.77      | 1.11      | 12300          |
| T49L+G107A           | 96.4      | 1.87      | 0.72      | 13600          |
| A52S+V54N+T49L+G107A | 80.5      | 2.55      | 0.43      | 10000          |
| A52S+V54N+T49L       | 95.8      | 1.76      | 0.84      | 8400           |
| G107A                | 94.4      | 1.89      | 1.04      | 19600          |
| Q51R+A52S            | 95.9      | 1.77      | 1.27      | 16500          |
| A52N                 | 95.5      | 1.89      | 1.56      | 17600          |
| LE174 (对照)           | 95.9/95.8 | 1.87/1.83 | 1.17/1.35 | 16000          |

5           与对照相比，液化过程中，本发明活性 $\alpha$ -淀粉酶变体的存在导致了潘糖水平的降低(DP3)。

        尤其,LE429 的 T49L+G107A 变体和 LE429 的 A52S+V54N+T49L 变体，分别导致潘糖水平的显著降低(DP3)。如果这些 $\alpha$ -淀粉酶变体用于淀粉液化，将不需要在糖化开始之前使酶失活。

10           **实施例 4**

        LE429 变体的液化和糖化

        在相同条件下对其它多个 LE429 变体重复实施例 3 的实验。

        结果如下：

| 变体/蔗糖图谱             | DP1   | DP2   | DP3   | DP4+  |
|---------------------|-------|-------|-------|-------|
| 15       T49V+G107A | 95.9% | 1.72% | 1.27% | 1.11% |
| T49Y+G107A          | 95.3% | 1.73% | 1.29% | 1.65% |
| T49N+G107A          | 95.7% | 1.64% | 1.51% | 1.18% |

|   |                       |       |       |       |       |
|---|-----------------------|-------|-------|-------|-------|
| 5 | T49L+A52S+G107A       | 95.7% | 1.73% | 0.95% | 1.67% |
|   | T49L+A52T+G107A       | 95.8% | 1.66% | 1.03% | 1.48% |
|   | T49L+A52F+G107A       | 95.7% | 1.69% | 1.16% | 1.42% |
|   | T49L+A52L+G107A       | 95.5% | 1.70% | 1.40% | 1.38% |
|   | T49L+A52I+G107A       | 95.9% | 1.72% | 1.31% | 1.07% |
|   | T49L+A52V+G107A       | 94.7% | 1.69% | 1.16% | 2.44% |
|   | T49L+A52V+G107A+A111V | 94.5% | 1.75% | 0.72% | 2.99% |
|   | LE429                 | 94.9% | 1.71% | 1.85% | 1.51% |

### 实施例 5

- 10 对多个 LE429 变体重复实施例 3 的实验，但液化在 95℃，pH6.0 进行，糖化在 60℃，pH4.5，40ppm CaCl<sub>2</sub> 下进行，然后使酶失活。下述变体是 LE 429 变体。实验结果如下：

|    | 变体/蔗糖图谱         | DP4+ | DP3  | DP2  | DP1   |
|----|-----------------|------|------|------|-------|
| 15 | T49F            | 1.15 | 0.92 | 1.83 | 96.12 |
|    | T49D+G107A      | 0.84 | 1.03 | 1.82 | 96.3  |
|    | T49I+G107A      | 0.97 | 0.64 | 1.84 | 96.55 |
|    | T49L+G107A      | 0.96 | 0.81 | 1.82 | 96.42 |
|    | T49L+A52S+G107A | 1.37 | 0.75 | 1.88 | 96.01 |
|    | T49L+A52T+G107A | 0.87 | 0.81 | 1.8  | 96.52 |
| 20 | T49L+A52F+G107A | 0.98 | 0.83 | 1.87 | 96.31 |
|    | T49V+G107A      | 0.65 | 0.8  | 2.13 | 96.43 |
|    | T49Y+G107A      | 0.83 | 0.94 | 1.89 | 96.35 |
|    | LE429           | 1.16 | 1.21 | 1.77 | 95.87 |

25

### 参考文献

- Klein, C., 等, 生物化学 1992, 31, 8740-8746,
- 30 Mizuno, H., 等, 分子生物学杂志(J. Mol. Biol.) (1993)234, 1282-1283,
- Chang, C., 等, 分子生物学杂志(J. Mol. Biol.) (1993)229, 235-238,

- Larson, S.B., 分子生物学杂志(J. Mol. Biol.)(1994)235, 1560-1584,  
 LawSon, C.L., 分子生物学杂志(J. Mol. Biol.)(1994)236, 590-600,  
 Qian, M., 等, 分子生物学杂志(J. Mol. Biol.)(1993)231, 785-799,  
 Brady, R.L.等, 结晶学报 B 辑(Acta Crystallogr. sect. B), 47, 527-535,  
 5 Swift, H.J., 等, 结晶学报 B 辑, 47, 535-544  
 A. Kadziola, 博士论文: “来自大麦 $\alpha$ -淀粉酶及其与底物类似物抑制物的结晶学 X-射线研究”, 哥本哈根大学化学系, 1993,  
 MacGregor, E.A., 食物胶体学(Food Hydrocolloids), 1987, 卷 1, 5-6 期, p.  
 B . Diderichsen and L . Christiansen ,嗜热脂肪芽孢杆菌麦芽糖基因淀粉  
 10 酶的克隆, FEMS Microbiol. letters: 56: pp. 53-60 (1988)  
 Hudson 等,免疫学实践(Practical Immunology), 3 版( i 9 8 9 ), Blackwell  
 科学出版社,  
 Sambrook 等, 分子克隆: 实验室手册(Molecular Cloning: A Laboratory  
 Manual), 2 版, 冷泉港, 1989  
 15 S.L. Beaucage and M.H. Caruthers, Tetrahedron Letters 22, 1981,pp. 1859-  
 1869  
 Matthes 等, The EMBO J. 3, 1984, pp. 801-805.  
 R.K.Saiki 等, 科学 239 1988 pp. 487-491.  
 Morinaga 等, (1984, 生物技术 2:646-639)  
 20 Nelson and Long, 分析生物化学(Analytical Biochemistry)180, 1989, pp  
 147-151  
 Hunkapiller 等, 1984, 自然 310:105-111  
 R.Higuchi, B.Krummel, 和 R.K.Saiki (1988)— DNA 片段的胞外制备和特  
 异诱变: 蛋白质和 DNA 相互作用研究. 核酸研究(Nucl. Acids Res.)16:7351 -  
 25 7367 .  
 Dubnau 等, 1971, 分子生物学杂志. 56 pp. 209-221.  
 Gryczan 等, 1978, 细菌学杂志(J.Bacteriol.)134,pp.318-329  
 S.D. Erlich, 1977, Proc.Natl.Acad. Sci. 74 pp. 1680-1682. ,  
 Boel 等, 1990, 生物化学 29 pp. 6244-6249.  
 30 Sarkar 和 Sommer, 1990, 生物技术 8, pp.404-407

## &lt;110&gt; Novo Nordisk A/S

 $\langle 120 \rangle$ 

<130>

<160> 40

&lt;170&gt; PatentIn Ver. 2.1

<210> 1

<211> 1443

&lt;212&gt; DNA

### 〈213〉 解淀粉芽孢杆菌

 $\langle 220 \rangle$ 

<221> CDS

<222> (1).. (1443)

<400> 1

gta aat ggc acg ctg atg cag tat ttt gaa tgg tat acg ccg aac gac 48  
Val Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Thr Pro Asn Asp  
1 5 10 15

ggc cag cat tgg aaa cga ttg cag aat gat gcg gaa cat tta tcg gat 96  
Gly Gln His Trp Lys Arg Leu Gln Asn Asp Ala Glu His Leu Ser Asp  
20 25 30

atc ggt att act gcc gtc tgg att ccc ccg gca tat aag gga acg agc 144  
Ile Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly Thr Ser  
35 40 45

caa gcg gat gtg ggc tac ggt gct tac gac ctt tat gat tta ggg gag. 192  
Gln Ala Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu Gly Glu  
50 55 60

ttt cat caa aaa ggg acg gtt cgg aca aag tac ggc aca aaa gga gag 240  
Phe His Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Gly Glu  
65 70 75 80



---

|                                                                 |     |
|-----------------------------------------------------------------|-----|
| ctg caa tct gcg atc aaa agt ctt cat tcc cgc gac att aac gtt tac | 288 |
| Leu Gln Ser Ala Ile Lys Ser Leu His Ser Arg Asp Ile Asn Val Tyr |     |
| 85 90 95                                                        |     |
| ggg gat gtg gtc atc aac cac aaa ggc ggc gct gat gcg acc gaa gat | 336 |
| Gly Asp Val Val Ile Asn His Lys Gly Gly Ala Asp Ala Thr Glu Asp |     |
| 100 105 110                                                     |     |
| gta acc gcg gtt gaa gtc gat ccc gct gac cgc aac cgc gta att tca | 384 |
| Val Thr Ala Val Glu Val Asp Pro Ala Asp Arg Asn Arg Val Ile Ser |     |
| 115 120 125                                                     |     |
| gga gaa cac cta att aaa gcc tgg aca cat ttt cat ttt ccg ggg cgc | 432 |
| Gly Glu His Leu Ile Lys Ala Trp Thr His Phe His Phe Pro Gly Arg |     |
| 130 135 140                                                     |     |
| ggc agc aca tac agc gat ttt aag tgg tat tgg tac cat ttt gac gga | 480 |
| Gly Ser Thr Tyr Ser Asp Phe Lys Trp Tyr Trp Tyr His Phe Asp Gly |     |
| 145 150 155 160                                                 |     |
| acc gat tgg gac gag tcc cga aag ctg aac cgc atc tat aag ttt caa | 528 |
| Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys Phe Gln |     |
| 165 170 175                                                     |     |
| ggg aag act tgg gat tgg gaa gtt tcc aat gaa ttc ggc aac tat gat | 576 |
| Gly Lys Thr Trp Asp Trp Glu Val Ser Asn Glu Phe Gly Asn Tyr Asp |     |
| 180 185 190                                                     |     |
| tat ttg atg tat gcc gac ttt gat tat gac cat cct gat gtc gta gca | 624 |
| Tyr Leu Met Tyr Ala Asp Phe Asp Tyr Asp His Pro Asp Val Val Ala |     |
| 195 200 205                                                     |     |
| gag att aag aga tgg ggc act tgg tat gcc aat gaa ctg caa ttg gac | 672 |
| Glu Ile Lys Arg Trp Gly Thr Trp Tyr Ala Asn Glu Leu Gln Leu Asp |     |
| 210 215 220                                                     |     |
| ggt ttc cgt ctt gat gct gtc aaa cac att aaa ttt tct ttt ttg cgg | 720 |
| Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Phe Ser Phe Leu Arg |     |
| 225 230 235 240                                                 |     |
| gat tgg gtt aat cat gtc agg gaa aaa acg ggg aag gaa atg ttt acg | 768 |
| Asp Trp Val Asn His Val Arg Glu Lys Thr Gly Lys Glu Met Phe Thr |     |
| 245 250 255                                                     |     |

|                                                                 |      |
|-----------------------------------------------------------------|------|
| gta gct gag tac tgg tcg aat gac ttg ggc gcg ctg gaa aac tat ttg | 816  |
| Val Ala Glu Tyr Trp Ser Asn Asp Leu Gly Ala Leu Glu Asn Tyr Leu |      |
| 260 265 270                                                     |      |
| aac aaa aca aat ttt aat cat tca gtg ttt gac gtg ccg ctt cat tat | 864  |
| Asn Lys Thr Asn Phe Asn His Ser Val Phe Asp Val Pro Leu His Tyr |      |
| 275 280 285                                                     |      |
| cag ttc cat gct gca tcg aca cag gga ggc ggc tat gat atg agg aaa | 912  |
| Gln Phe His Ala Ala Ser Thr Gln Gly Gly Gly Tyr Asp Met Arg Lys |      |
| 290 295 300                                                     |      |
| ttg ctg aac ggt acg gtc gtt tcc aag cat ccg ttg aaa tcg gtt aca | 960  |
| Leu Leu Asn Gly Thr Val Val Ser Lys His Pro Leu Lys Ser Val Thr |      |
| 305 310 315 320                                                 |      |
| ttt gtc gat aac cat gat aca cag ccg ggg caa tcg ctt gag tcg act | 1008 |
| Phe Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu Ser Thr |      |
| 325 330 335                                                     |      |
| gtc caa aca tgg ttt aag ccg ctt gct tac gct ttt att ctc aca agg | 1056 |
| Val Gln Thr Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg |      |
| 340 345 350                                                     |      |
| gaa tct gga tac cct cag gtt ttc tac ggg gat atg tac ggg acg aaa | 1104 |
| Glu Ser Gly Tyr Pro Gln Val Phe Tyr Gly Asp Met Tyr Gly Thr Lys |      |
| 355 360 365                                                     |      |
| gga gac tcc cag cgc gaa att cct gcc ttg aaa cac aaa att gaa ccg | 1152 |
| Gly Asp Ser Gln Arg Glu Ile Pro Ala Leu Lys His Lys Ile Glu Pro |      |
| 370 375 380                                                     |      |
| atc tta aaa gcg aga aaa cag tat gcg tac gga gca cag cat gat tat | 1200 |
| Ile Leu Lys Ala Arg Lys Gln Tyr Ala Tyr Gly Ala Gln His Asp Tyr |      |
| 385 390 395 400                                                 |      |
| ttc gac cac cat gac att gtc ggc tgg aca agg gaa ggc gac agc tcg | 1248 |
| Phe Asp His His Asp Ile Val Gly Trp Thr Arg Glu Gly Asp Ser Ser |      |
| 405 410 415                                                     |      |
| gtt gca aat tca ggt ttg gcg gca tta ata aca gac gga ccc ggt ggg | 1296 |
| Val Ala Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly Gly |      |
| 420 425 430                                                     |      |

---

gca aag cga atg tat gtc ggc cgg caa aac gcc ggt gag aca tgg cat 1344  
 Ala Lys Arg Met Tyr Val Gly Arg Gln Asn Ala Gly Glu Thr Trp His  
           435                    440                    445

gac att acc gga aac cgt tcg gag ccg gtt gtc atc aat tcg gaa ggc 1392  
 Asp Ile Thr Gly Asn Arg Ser Glu Pro Val Val Ile Asn Ser Glu Gly  
           450                    455                    460

tgg gga gag ttt cac gta aac ggc ggg tcg gtt tca att tat gtt caa 1440  
 Trp Gly Glu Phe His Val Asn Gly Gly Ser Val Ser Ile Tyr Val Gln  
 465                    470                    475                    480

aga 1443  
 Arg

<210> 2

<211> 481

<212> PRT

<213> 解淀粉芽孢杆菌

<400> 2

Val Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Thr Pro Asn Asp  
       1                    5                    10                    15

Gly Gln His Trp Lys Arg Leu Gln Asn Asp Ala Glu His Leu Ser Asp  
           20                    25                    30

Ile Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly Thr Ser  
           35                    40                    45

Gln Ala Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu Gly Glu  
       50                    55                    60

Phe His Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Gly Glu  
       65                    70                    75                    80

Leu Gln Ser Ala Ile Lys Ser Leu His Ser Arg Asp Ile Asn Val Tyr  
           85                    90                    95

Gly Asp Val Val Ile Asn His Lys Gly Gly Ala Asp Ala Thr Glu Asp  
           100                    105                    110

Val Thr Ala Val Glu Val Asp Pro Ala Asp Arg Asn Arg Val Ile Ser

| 115                                                             | 120 | 125         |
|-----------------------------------------------------------------|-----|-------------|
| Gly Glu His Leu Ile Lys Ala Trp Thr His Phe His Phe Pro Gly Arg |     |             |
| 130                                                             | 135 | 140         |
| Gly Ser Thr Tyr Ser Asp Phe Lys Trp Tyr Trp Tyr His Phe Asp Gly |     |             |
| 145                                                             | 150 | 155 160     |
| Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys Phe Gln |     |             |
|                                                                 | 165 | 170 175     |
| Gly Lys Thr Trp Asp Trp Glu Val Ser Asn Glu Phe Gly Asn Tyr Asp |     |             |
|                                                                 | 180 | 185 190     |
| Tyr Leu Met Tyr Ala Asp Phe Asp Tyr Asp His Pro Asp Val Val Ala |     |             |
|                                                                 | 195 | 200 205     |
| Glu Ile Lys Arg Trp Gly Thr Trp Tyr Ala Asn Glu Leu Gln Leu Asp |     |             |
|                                                                 | 210 | 215 220     |
| Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Phe Ser Phe Leu Arg |     |             |
|                                                                 | 225 | 230 235 240 |
| Asp Trp Val Asn His Val Arg Glu Lys Thr Gly Lys Glu Met Phe Thr |     |             |
|                                                                 | 245 | 250 255     |
| Val Ala Glu Tyr Trp Ser Asn Asp Leu Gly Ala Leu Glu Asn Tyr Leu |     |             |
|                                                                 | 260 | 265 270     |
| Asn Lys Thr Asn Phe Asn His Ser Val Phe Asp Val Pro Leu His Tyr |     |             |
|                                                                 | 275 | 280 285     |
| Gln Phe His Ala Ala Ser Thr Gln Gly Gly Gly Tyr Asp Met Arg Lys |     |             |
|                                                                 | 290 | 295 300     |
| Leu Leu Asn Gly Thr Val Val Ser Lys His Pro Leu Lys Ser Val Thr |     |             |
|                                                                 | 305 | 310 315 320 |
| Phe Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu Ser Thr |     |             |
|                                                                 | 325 | 330 335     |
| Val Gln Thr Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg |     |             |
|                                                                 | 340 | 345 350     |

Glu Ser Gly Tyr Pro Gln Val Phe Tyr Gly Asp Met Tyr Gly Thr Lys  
 355 360 365  
 Gly Asp Ser Gln Arg Glu Ile Pro Ala Leu Lys His Lys Ile Glu Pro  
 370 375 380  
 Ile Leu Lys Ala Arg Lys Gln Tyr Ala Tyr Gly Ala Gln His Asp Tyr  
 385 390 395 400  
 Phe Asp His His Asp Ile Val Gly Trp Thr Arg Glu Gly Asp Ser Ser  
 405 410 415  
 Val Ala Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly Gly  
 420 425 430  
 Ala Lys Arg Met Tyr Val Gly Arg Gln Asn Ala Gly Glu Thr Trp His  
 435 440 445  
 Asp Ile Thr Gly Asn Arg Ser Glu Pro Val Val Ile Asn Ser Glu Gly  
 450 455 460  
 Trp Gly Glu Phe His Val Asn Gly Gly Ser Val Ser Ile Tyr Val Gln  
 465 470 475 480

Arg

<210> 3

<211> 1920

<212> DNA

<213> 地衣芽孢杆菌

<220>

<221> CDS

<222> (421)..(1872)

<400> 3

cggaagattg gaagtacaaa aataagcaaa agattgtcaa tcatgtcatg agccatgcgg 60

gagacggaaa aatcgtctta atgcacgata ttatgcaac gtgcgagat gctgctgaag 120

agattattaa aaagctgaaa gcaaaaggct atcaattggc aactgtatct cagcttgaag 180

aagtgaagaa gcagagaggc tattgaataa atgagtagaa gcgcatatc ggcgcttttc 240

ttttggaaga aaatataggg aaaatggtac ttgttaaaaa ttcggaatat ttatacaaca 300

tcatatgttt cacattgaaa ggggaggaga atcatgaaac aacaaaaacg gctttacgcc 360

cgattgctga cgctgttatt tgcgctcatc ttcttgctgc ctcatcttgc agcagcggcg 420

gca aat ctt aat ggg acg ctg atg cag tat ttt gaa tgg tac atg ccc 468

Ala Asn Leu Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Met Pro

1 5 10 15

aat gac ggc caa cat tgg agg cgt ttg caa aac gac tcg gca tat ttg 516

Asn Asp Gly Gln His Trp Arg Arg Leu Gln Asn Asp Ser Ala Tyr Leu

20 25 30

gct gaa cac ggt att act gcc gtc tgg att ccc ccg gca tat aag gga 564

Ala Glu His Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly

35 40 45

acg agc caa gcg gat gtg ggc tac ggt gct tac gac ctt tat gat tta 612

Thr Ser Gln Ala Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu

50 55 60

ggg gag ttt cat caa aaa ggg acg gtt cgg aca aag tac ggc aca aaa 660

Gly Glu Phe His Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys

65 70 75 80

gga gag ctg caa tct gcg atc aaa agt ctt cat tcc cgc gac att aac 708

Gly Glu Leu Gln Ser Ala Ile Lys Ser Leu His Ser Arg Asp Ile Asn

85 90 95

gtt tac ggg gat gtg gtc atc aac cac aaa ggc ggc gct gat gcg acc 756

Val Tyr Gly Asp Val Val Ile Asn His Lys Gly Gly Ala Asp Ala Thr

100 105 110

gaa gat gta acc gcg gtt gaa gtc gat ccc gct gac cgc aac cgc gta 804

Glu Asp Val Thr Ala Val Glu Val Asp Pro Ala Asp Arg Asn Arg Val

115 120 125

att tca gga gaa cac cta att aaa gcc tgg aca cat ttt cat ttt ccg 852

Ile Ser Gly Glu His Leu Ile Lys Ala Trp Thr His Phe His Phe Pro

130 135 140

ggg cgc ggc agc aca tac agc gat ttt aaa tgg cat tgg tac cat ttt 900

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|--|
| Gly | Arg | Gly | Ser | Thr | Tyr | Ser | Asp | Phe | Lys | Trp | His | Trp | Tyr | His | Phe |      |  |
| 145 |     |     |     |     | 150 |     |     |     | 155 |     |     |     |     |     | 160 |      |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |  |
| gac | gga | acc | gat | tgg | gac | gag | tcc | cga | aag | ctg | aac | cgc | atc | tat | aag | 948  |  |
| Asp | Gly | Thr | Asp | Trp | Asp | Glu | Ser | Arg | Lys | Leu | Asn | Arg | Ile | Tyr | Lys |      |  |
|     |     |     | 165 |     |     |     |     |     | 170 |     |     |     |     | 175 |     |      |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |  |
| ttt | caa | gga | aag | gct | tgg | gat | tgg | gaa | gtt | tcc | aat | gaa | aac | ggc | aac | 996  |  |
| Phe | Gln | Gly | Lys | Ala | Trp | Asp | Trp | Glu | Val | Ser | Asn | Glu | Asn | Gly | Asn |      |  |
|     |     |     | 180 |     |     |     |     | 185 |     |     |     |     |     | 190 |     |      |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |  |
| tat | gat | tat | ttg | atg | tat | gcc | gac | atc | gat | tat | gac | cat | cct | gat | gtc | 1044 |  |
| Tyr | Asp | Tyr | Leu | Met | Tyr | Ala | Asp | Ile | Asp | Tyr | Asp | His | Pro | Asp | Val |      |  |
|     |     |     | 195 |     |     |     |     | 200 |     |     |     |     | 205 |     |     |      |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |  |
| gca | gca | gaa | att | aag | aga | tgg | ggc | act | tgg | tat | gcc | aat | gaa | ctg | caa | 1092 |  |
| Ala | Ala | Glu | Ile | Lys | Arg | Trp | Gly | Thr | Trp | Tyr | Ala | Asn | Glu | Leu | Gln |      |  |
|     |     | 210 |     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |      |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |  |
| ttg | gac | ggt | ttc | cgt | ctt | gat | gct | gtc | aaa | cac | att | aaa | ttt | tct | ttt | 1140 |  |
| Leu | Asp | Gly | Phe | Arg | Leu | Asp | Ala | Val | Lys | His | Ile | Lys | Phe | Ser | Phe |      |  |
| 225 |     |     |     |     | 230 |     |     |     |     | 235 |     |     |     |     | 240 |      |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |  |
| ttg | cgg | gat | tgg | gtt | aat | cat | gtc | agg | gaa | aaa | acg | ggg | aag | gaa | atg | 1188 |  |
| Leu | Arg | Asp | Trp | Val | Asn | His | Val | Arg | Glu | Lys | Thr | Gly | Lys | Glu | Met |      |  |
|     |     |     |     | 245 |     |     |     |     |     | 250 |     |     |     | 255 |     |      |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |  |
| ttt | acg | gta | gct | gaa | tat | tgg | cag | aat | gac | ttg | ggc | gcg | ctg | gaa | aac | 1236 |  |
| Phe | Thr | Val | Ala | Glu | Tyr | Trp | Gln | Asn | Asp | Leu | Gly | Ala | Leu | Glu | Asn |      |  |
|     |     |     | 260 |     |     |     |     |     |     | 265 |     |     |     | 270 |     |      |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |  |
| tat | ttg | aac | aaa | aca | aat | ttt | aat | cat | tca | gtg | ttt | gac | gtg | ccg | ctt | 1284 |  |
| Tyr | Leu | Asn | Lys | Thr | Asn | Phe | Asn | His | Ser | Val | Phe | Asp | Val | Pro | Leu |      |  |
|     |     |     | 275 |     |     |     |     | 280 |     |     |     |     | 285 |     |     |      |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |  |
| cat | tat | cag | ttc | cat | gct | gca | tcg | aca | cag | gga | ggc | ggc | tat | gat | atg | 1332 |  |
| His | Tyr | Gln | Phe | His | Ala | Ala | Ser | Thr | Gln | Gly | Gly | Gly | Tyr | Asp | Met |      |  |
|     |     | 290 |     |     |     | 295 |     |     |     |     | 300 |     |     |     |     |      |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |  |
| agg | aaa | ttg | ctg | aac | ggt | acg | gtc | gtt | tcc | aag | cat | ccg | ttg | aaa | tcg | 1380 |  |
| Arg | Lys | Leu | Leu | Asn | Gly | Thr | Val | Val | Ser | Lys | His | Pro | Leu | Lys | Ser |      |  |
| 305 |     |     |     |     | 310 |     |     |     |     | 315 |     |     |     |     | 320 |      |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |  |
| gtt | aca | ttt | gtc | gat | aac | cat | gat | aca | cag | ccg | ggg | caa | tcg | ctt | gag | 1428 |  |

|                                                                 |      |
|-----------------------------------------------------------------|------|
| Val Thr Phe Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu |      |
| 325 330 335                                                     |      |
| tcg act gtc caa aca tgg ttt aag ccg ctt gct tac gct ttt att ctc | 1476 |
| Ser Thr Val Gln Thr Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu |      |
| 340 345 350                                                     |      |
| aca agg gaa tct gga tac cct cag gtt ttc tac ggg gat atg tac ggg | 1524 |
| Thr Arg Glu Ser Gly Tyr Pro Gln Val Phe Tyr Gly Asp Met Tyr Gly |      |
| 355 360 365                                                     |      |
| acg aaa gga gac tcc cag cgc gaa att cct gcc ttg aaa cac aaa att | 1572 |
| Thr Lys Gly Asp Ser Gln Arg Glu Ile Pro Ala Leu Lys His Lys Ile |      |
| 370 375 380                                                     |      |
| gaa ccg atc tta aaa gcg aga aaa cag tat gcg tac gga gca cag cat | 1620 |
| Glu Pro Ile Leu Lys Ala Arg Lys Gln Tyr Ala Tyr Gly Ala Gln His |      |
| 385 390 395 400                                                 |      |
| gat tat ttc gac cac cat gac att gtc ggc tgg aca agg gaa ggc gac | 1668 |
| Asp Tyr Phe Asp His His Asp Ile Val Gly Trp Thr Arg Glu Gly Asp |      |
| 405 410 415                                                     |      |
| agc tcg gtt gca aat tca ggt ttg gcg gca tta ata aca gac gga ccc | 1716 |
| Ser Ser Val Ala Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro |      |
| 420 425 430                                                     |      |
| ggt ggg gca aag cga atg tat gtc ggc cgg caa aac gcc ggt gag aca | 1764 |
| Gly Gly Ala Lys Arg Met Tyr Val Gly Arg Gln Asn Ala Gly Glu Thr |      |
| 435 440 445                                                     |      |
| tgg cat gac att acc gga aac cgt tcg gag ccg gtt gtc atc aat tcg | 1812 |
| Trp His Asp Ile Thr Gly Asn Arg Ser Glu Pro Val Val Ile Asn Ser |      |
| 450 455 460                                                     |      |
| gaa ggc tgg gga gag ttt cac gta aac ggc ggg tcg gtt tca att tat | 1860 |
| Glu Gly Trp Gly Glu Phe His Val Asn Gly Gly Ser Val Ser Ile Tyr |      |
| 465 470 475 480                                                 |      |
| ggt caa aga tag aagagcagag aggacggatt tcctgaagga aatccgtttt     | 1912 |
| Val Gln Arg                                                     |      |
| tttatttt                                                        | 1920 |



&lt;210&gt; 4

&lt;211&gt; 483

&lt;212&gt; PRT

&lt;213&gt; 地衣芽孢杆菌

&lt;400&gt; 4

Ala Asn Leu Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Met Pro  
 1 5 10 15

Asn Asp Gly Gln His Trp Arg Arg Leu Gln Asn Asp Ser Ala Tyr Leu  
 20 25 30

Ala Glu His Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly  
 35 40 45

Thr Ser Gln Ala Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu  
 50 55 60

Gly Glu Phe His Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys  
 65 70 75 80

Gly Glu Leu Gln Ser Ala Ile Lys Ser Leu His Ser Arg Asp Ile Asn  
 85 90 95

Val Tyr Gly Asp Val Val Ile Asn His Lys Gly Gly Ala Asp Ala Thr  
 100 105 110

Glu Asp Val Thr Ala Val Glu Val Asp Pro Ala Asp Arg Asn Arg Val  
 115 120 125

Ile Ser Gly Glu His Leu Ile Lys Ala Trp Thr His Phe His Phe Pro  
 130 135 140

Gly Arg Gly Ser Thr Tyr Ser Asp Phe Lys Trp His Trp Tyr His Phe  
 145 150 155 160

Asp Gly Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys  
 165 170 175

Phe Gln Gly Lys Ala Trp Asp Trp Glu Val Ser Asn Glu Asn Gly Asn  
 180 185 190

Tyr Asp Tyr Leu Met Tyr Ala Asp Ile Asp Tyr Asp His Pro Asp Val

| 195                                                             | 200 | 205     |
|-----------------------------------------------------------------|-----|---------|
| Ala Ala Glu Ile Lys Arg Trp Gly Thr Trp Tyr Ala Asn Glu Leu Gln |     |         |
| 210                                                             | 215 | 220     |
| Leu Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Phe Ser Phe |     |         |
| 225                                                             | 230 | 235 240 |
| Leu Arg Asp Trp Val Asn His Val Arg Glu Lys Thr Gly Lys Glu Met |     |         |
|                                                                 | 245 | 250 255 |
| Phe Thr Val Ala Glu Tyr Trp Gln Asn Asp Leu Gly Ala Leu Glu Asn |     |         |
|                                                                 | 260 | 265 270 |
| Tyr Leu Asn Lys Thr Asn Phe Asn His Ser Val Phe Asp Val Pro Leu |     |         |
|                                                                 | 275 | 280 285 |
| His Tyr Gln Phe His Ala Ala Ser Thr Gln Gly Gly Gly Tyr Asp Met |     |         |
|                                                                 | 290 | 295 300 |
| Arg Lys Leu Leu Asn Gly Thr Val Val Ser Lys His Pro Leu Lys Ser |     |         |
| 305                                                             | 310 | 315 320 |
| Val Thr Phe Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu |     |         |
|                                                                 | 325 | 330 335 |
| Ser Thr Val Gln Thr Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu |     |         |
|                                                                 | 340 | 345 350 |
| Thr Arg Glu Ser Gly Tyr Pro Gln Val Phe Tyr Gly Asp Met Tyr Gly |     |         |
|                                                                 | 355 | 360 365 |
| Thr Lys Gly Asp Ser Gln Arg Glu Ile Pro Ala Leu Lys His Lys Ile |     |         |
|                                                                 | 370 | 375 380 |
| Glu Pro Ile Leu Lys Ala Arg Lys Gln Tyr Ala Tyr Gly Ala Gln His |     |         |
| 385                                                             | 390 | 395 400 |
| Asp Tyr Phe Asp His His Asp Ile Val Gly Trp Thr Arg Glu Gly Asp |     |         |
|                                                                 | 405 | 410 415 |
| Ser Ser Val Ala Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro |     |         |
|                                                                 | 420 | 425 430 |

---

Gly Gly Ala Lys Arg Met Tyr Val Gly Arg Gln Asn Ala Gly Glu Thr  
                   435                                  440                                  445

Trp His Asp Ile Thr Gly Asn Arg Ser Glu Pro Val Val Ile Asn Ser  
           450                                  455                                  460

Glu Gly Trp Gly Glu Phe His Val Asn Gly Gly Ser Val Ser Ile Tyr  
 465                                  470                                  475                                  480

Val Gln Arg

<210> 5

<211> 2604

<212> DNA

<213> 解淀粉芽孢杆菌

<220>

<221> -10\_信号

<222> (707).. (712)

<220>

<221> -35\_信号

<222> (729).. (734)

<220>

<221> RBS

<222> (759).. (762)

<220>

<221> 信号\_肽

<222> (770).. (862)

<220>

<221> 成熟\_肽

<222> (863).. (2314)

<220>

<221> 终止子

<222> (2321).. (2376)

<220>

<221> CDS

<222> (863).. (2314)

&lt;400&gt; 5

aagcttcaag cggtcaatcg gaatgtgcat ctgcttcat acttaggttt tcacccgcat 60

attaagcagg cgtttttgaa ccgtgtgaca gaagctgttc gaaaccccg cgggcggttt 120

gattttaagg ggggacagta tgctgcctct tcacattaat ctgagcgga aaagaatcat 180

cattgctggc gggggcaatg ttgcattaag aaggctgaaa cgggtgtttcc ggaaggcgt 240

gatattaccg tgatcagtct gagcctgcct gaaattaaaa agctggcgga tgaaggacgc 300

atccgctgga tcccccgag aattgaaatg aaagatctca agcccgcttt tttcattatt 360

gccgcgacaa atgaccgagg cgtgaatcag gagatagccg caaacgcttc tgaaacgcag 420

ctggctcaact gtgtaagcaa ggctgaacaa ggcagcgtat atatgccgaa gatcatccgc 480

aaagggcgca ttcaagtatc agtatcaaca agcggggcaa gccccgcaca tacgaaaaga 540

ctggctgaaa acattgagcc tttgatgact gatgatttgg ctgaagaagt ggatcgattg 600

tttgagaaaa gaagaagacc ataaaaatac cttgtctgtc atcagacagg gtatttttta 660

tgctgtccag actgtccgct gtgtaaaaaa taggaataaa ggggggttgt tattatttta 720

ctgatatgta aaatataatt tgtataagaa aatgagaggg agaggaaaca tgattcaaaa 780

acgaaagcgg acagtttctg tcagacttgt gcttatgtgc acgctgttat ttgtcagttt 840

gccgattaca aaaacatcag cc gta aat ggc acg ctg atg cag tat ttt gaa 892

Val Asn Gly Thr Leu Met Gln Tyr Phe Glu

1

5

10

tgg tat acg ccg aac gac ggc cag cat tgg aaa cga ttg cag aat gat 940

Trp Tyr Thr Pro Asn Asp Gly Gln His Trp Lys Arg Leu Gln Asn Asp

15

20

25

gcg gaa cat tta tcg gat atc gga atc act gcc gtc tgg att cct ccc 988

Ala Glu His Leu Ser Asp Ile Gly Ile Thr Ala Val Trp Ile Pro Pro

30

35

40

gca tac aaa gga ttg agc caa tcc gat aac gga tac gga cct tat gat 1036

Ala Tyr Lys Gly Leu Ser Gln Ser Asp Asn Gly Tyr Gly Pro Tyr Asp

| 45                                                              | 50   | 55  |     |
|-----------------------------------------------------------------|------|-----|-----|
| ttg tat gat tta gga gaa ttc cag caa aaa ggg acg gtc aga acg aaa | 1084 |     |     |
| Leu Tyr Asp Leu Gly Glu Phe Gln Gln Lys Gly Thr Val Arg Thr Lys |      |     |     |
| 60                                                              | 65   | 70  |     |
| tac ggc aca aaa tca gag ctt caa gat gcg atc ggc tca ctg cat tcc | 1132 |     |     |
| Tyr Gly Thr Lys Ser Glu Leu Gln Asp Ala Ile Gly Ser Leu His Ser |      |     |     |
| 75                                                              | 80   | 85  | 90  |
| cgg aac gtc caa gta tac gga gat gtg gtt ttg aat cat aag gct ggt | 1180 |     |     |
| Arg Asn Val Gln Val Tyr Gly Asp Val Val Leu Asn His Lys Ala Gly |      |     |     |
| 95                                                              | 100  | 105 |     |
| gct gat gca aca gaa gat gta act gcc gtc gaa gtc aat ccg gcc aat | 1228 |     |     |
| Ala Asp Ala Thr Glu Asp Val Thr Ala Val Glu Val Asn Pro Ala Asn |      |     |     |
| 110                                                             | 115  | 120 |     |
| aga aat cag gaa act tcg gag gaa tat caa atc aaa gcg tgg acg gat | 1276 |     |     |
| Arg Asn Gln Glu Thr Ser Glu Glu Tyr Gln Ile Lys Ala Trp Thr Asp |      |     |     |
| 125                                                             | 130  | 135 |     |
| ttt cgt ttt ccg ggc cgt gga aac acg tac agt gat ttt aaa tgg cat | 1324 |     |     |
| Phe Arg Phe Pro Gly Arg Gly Asn Thr Tyr Ser Asp Phe Lys Trp His |      |     |     |
| 140                                                             | 145  | 150 |     |
| tgg tat cat ttc gac gga gcg gac tgg gat gaa tcc cgg aag atc agc | 1372 |     |     |
| Trp Tyr His Phe Asp Gly Ala Asp Trp Asp Glu Ser Arg Lys Ile Ser |      |     |     |
| 155                                                             | 160  | 165 | 170 |
| cgc atc ttt aag ttt cgt ggg gaa gga aaa gcg tgg gat tgg gaa gta | 1420 |     |     |
| Arg Ile Phe Lys Phe Arg Gly Glu Gly Lys Ala Trp Asp Trp Glu Val |      |     |     |
| 175                                                             | 180  | 185 |     |
| tca agt gaa aac ggc aac tat gac tat tta atg tat gct gat gtt gac | 1468 |     |     |
| Ser Ser Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp |      |     |     |
| 190                                                             | 195  | 200 |     |
| tac gac cac cct gat gtc gtg gca gag aca aaa aaa tgg ggt atc tgg | 1516 |     |     |
| Tyr Asp His Pro Asp Val Val Ala Glu Thr Lys Lys Trp Gly Ile Trp |      |     |     |
| 205                                                             | 210  | 215 |     |
| tat gcg aat gaa ctg tca tta gac ggc ttc cgt att gat gcc gcc aaa | 1564 |     |     |
| Tyr Ala Asn Glu Leu Ser Leu Asp Gly Phe Arg Ile Asp Ala Ala Lys |      |     |     |

| 220                                                             | 225 | 230 |      |
|-----------------------------------------------------------------|-----|-----|------|
| cat att aaa ttt tca ttt ctg cgt gat tgg gtt cag gcg gtc aga cag |     |     | 1612 |
| His Ile Lys Phe Ser Phe Leu Arg Asp Trp Val Gln Ala Val Arg Gln |     |     |      |
| 235                                                             | 240 | 245 | 250  |
| gcg acg gga aaa gaa atg ttt acg gtt gcg gag tat tgg cag aat aat |     |     | 1660 |
| Ala Thr Gly Lys Glu Met Phe Thr Val Ala Glu Tyr Trp Gln Asn Asn |     |     |      |
| 255                                                             | 260 | 265 |      |
| gcc ggg aaa ctc gaa aac tac ttg aat aaa aca agc ttt aat caa tcc |     |     | 1708 |
| Ala Gly Lys Leu Glu Asn Tyr Leu Asn Lys Thr Ser Phe Asn Gln Ser |     |     |      |
| 270                                                             | 275 | 280 |      |
| gtg ttt gat gtt ccg ctt cat ttc aat tta cag gcg gct tcc tca caa |     |     | 1756 |
| Val Phe Asp Val Pro Leu His Phe Asn Leu Gln Ala Ala Ser Ser Gln |     |     |      |
| 285                                                             | 290 | 295 |      |
| gga ggc gga tat gat atg agg cgt ttg ctg gac ggt acc gtt gtg tcc |     |     | 1804 |
| Gly Gly Gly Tyr Asp Met Arg Arg Leu Leu Asp Gly Thr Val Val Ser |     |     |      |
| 300                                                             | 305 | 310 |      |
| agg cat ccg gaa aag gcg gtt aca ttt gtt gaa aat cat gac aca cag |     |     | 1852 |
| Arg His Pro Glu Lys Ala Val Thr Phe Val Glu Asn His Asp Thr Gln |     |     |      |
| 315                                                             | 320 | 325 | 330  |
| ccg gga cag tca ttg gaa tcg aca gtc caa act tgg ttt aaa ccg ctt |     |     | 1900 |
| Pro Gly Gln Ser Leu Glu Ser Thr Val Gln Thr Trp Phe Lys Pro Leu |     |     |      |
| 335                                                             | 340 | 345 |      |
| gca tac gcc ttt att ttg aca aga gaa tcc ggt tat cct cag gtg ttc |     |     | 1948 |
| Ala Tyr Ala Phe Ile Leu Thr Arg Glu Ser Gly Tyr Pro Gln Val Phe |     |     |      |
| 350                                                             | 355 | 360 |      |
| tat ggg gat atg tac ggg aca aaa ggg aca tcg cca aag gaa att ccc |     |     | 1996 |
| Tyr Gly Asp Met Tyr Gly Thr Lys Gly Thr Ser Pro Lys Glu Ile Pro |     |     |      |
| 365                                                             | 370 | 375 |      |
| tca ctg aaa gat aat ata gag ccg att tta aaa gcg cgt aag gag tac |     |     | 2044 |
| Ser Leu Lys Asp Asn Ile Glu Pro Ile Leu Lys Ala Arg Lys Glu Tyr |     |     |      |
| 380                                                             | 385 | 390 |      |
| gca tac ggg ccc cag cac gat tat att gac cac ccg gat gtg atc gga |     |     | 2092 |
| Ala Tyr Gly Pro Gln His Asp Tyr Ile Asp His Pro Asp Val Ile Gly |     |     |      |

| 395                                                                | 400 | 405 | 410 |      |
|--------------------------------------------------------------------|-----|-----|-----|------|
| tgg acg agg gaa ggt gac agc tcc gcc gcc aaa tca ggt ttg gcc gct    |     |     |     | 2140 |
| Trp Thr Arg Glu Gly Asp Ser Ser Ala Ala Lys Ser Gly Leu Ala Ala    |     |     |     |      |
|                                                                    | 415 | 420 | 425 |      |
| tta atc acg gac gga ccc ggc gga tca aag cgg atg tat gcc ggc ctg    |     |     |     | 2188 |
| Leu Ile Thr Asp Gly Pro Gly Gly Ser Lys Arg Met Tyr Ala Gly Leu    |     |     |     |      |
|                                                                    | 430 | 435 | 440 |      |
| aaa aat gcc ggc gag aca tgg tat gac ata acg ggc aac cgt tca gat    |     |     |     | 2236 |
| Lys Asn Ala Gly Glu Thr Trp Tyr Asp Ile Thr Gly Asn Arg Ser Asp    |     |     |     |      |
|                                                                    | 445 | 450 | 455 |      |
| act gta aaa atc gga tct gac ggc tgg gga gag ttt cat gta aac gat    |     |     |     | 2284 |
| Thr Val Lys Ile Gly Ser Asp Gly Trp Gly Glu Phe His Val Asn Asp    |     |     |     |      |
|                                                                    | 460 | 465 | 470 |      |
| ggg tcc gtc tcc att tat gtt cag aaa taa ggtaataaaa aaacacctcc      |     |     |     | 2334 |
| Gly Ser Val Ser Ile Tyr Val Gln Lys                                |     |     |     |      |
| 475                                                                | 480 |     |     |      |
| aagctgagtg cgggtatcag ctiggagggtg cgttttat ttcagccgta tgacaaggtc   |     |     |     | 2394 |
| ggcatcaggt gtgacaaata cggatatgctg gctgtcatag gtgacaaatc cgggttttgc |     |     |     | 2454 |
| gccgtttggc tttttcacat gtctgatttt tgtataatca acaggcacgg agccggaatc  |     |     |     | 2514 |
| tttcgccttg gaaaaataag cggcgatcgt agctgcttcc aatatggatt gtcatcggg   |     |     |     | 2574 |
| atcgtctgctt ttaatcaciaa cgtgggatcc                                 |     |     |     | 2604 |
|                                                                    |     |     |     |      |
| <210> 6                                                            |     |     |     |      |
| <211> 483                                                          |     |     |     |      |
| <212> PRT                                                          |     |     |     |      |
| <213> 解淀粉芽孢杆菌                                                      |     |     |     |      |
|                                                                    |     |     |     |      |
| <400> 6                                                            |     |     |     |      |
| Val Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Thr Pro Asn Asp    |     |     |     |      |
| 1                                                                  | 5   | 10  | 15  |      |
| Gly Gln His Trp Lys Arg Leu Gln Asn Asp Ala Glu His Leu Ser Asp    |     |     |     |      |
| 20                                                                 | 25  | 30  |     |      |

Ile Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly Leu Ser  
                   35                                  40                                  45

Gln Ser Asp Asn Gly Tyr Gly Pro Tyr Asp Leu Tyr Asp Leu Gly Glu  
           50                                  55                                  60

Phe Gln Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Ser Glu  
   65                                  70                                  75                                  80

Leu Gln Asp Ala Ile Gly Ser Leu His Ser Arg Asn Val Gln Val Tyr  
                                   85                                  90                                  95

Gly Asp Val Val Leu Asn His Lys Ala Gly Ala Asp Ala Thr Glu Asp  
                   100                                  105                                  110

Val Thr Ala Val Glu Val Asn Pro Ala Asn Arg Asn Gln Glu Thr Ser  
           115                                  120                                  125

Glu Glu Tyr Gln Ile Lys Ala Trp Thr Asp Phe Arg Phe Pro Gly Arg  
   130                                  135                                  140

Gly Asn Thr Tyr Ser Asp Phe Lys Trp His Trp Tyr His Phe Asp Gly  
  145                                  150                                  155                                  160

Ala Asp Trp Asp Glu Ser Arg Lys Ile Ser Arg Ile Phe Lys Phe Arg  
                                   165                                  170                                  175

Gly Glu Gly Lys Ala Trp Asp Trp Glu Val Ser Ser Glu Asn Gly Asn  
                   180                                  185                                  190

Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp Tyr Asp His Pro Asp Val  
           195                                  200                                  205

Val Ala Glu Thr Lys Lys Trp Gly Ile Trp Tyr Ala Asn Glu Leu Ser  
   210                                  215                                  220

Leu Asp Gly Phe Arg Ile Asp Ala Ala Lys His Ile Lys Phe Ser Phe  
  225                                  230                                  235                                  240

Leu Arg Asp Trp Val Gln Ala Val Arg Gln Ala Thr Gly Lys Glu Met  
                                   245                                  250                                  255

Phe Thr Val Ala Glu Tyr Trp Gln Asn Asn Ala Gly Lys Leu Glu Asn



|                                                                 |     |     |
|-----------------------------------------------------------------|-----|-----|
| 260                                                             | 265 | 270 |
| Tyr Leu Asn Lys Thr Ser Phe Asn Gln Ser Val Phe Asp Val Pro Leu |     |     |
| 275                                                             | 280 | 285 |
| His Phe Asn Leu Gln Ala Ala Ser Ser Gln Gly Gly Gly Tyr Asp Met |     |     |
| 290                                                             | 295 | 300 |
| Arg Arg Leu Leu Asp Gly Thr Val Val Ser Arg His Pro Glu Lys Ala |     |     |
| 305                                                             | 310 | 315 |
| Val Thr Phe Val Glu Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu |     |     |
| 325                                                             | 330 | 335 |
| Ser Thr Val Gln Thr Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu |     |     |
| 340                                                             | 345 | 350 |
| Thr Arg Glu Ser Gly Tyr Pro Gln Val Phe Tyr Gly Asp Met Tyr Gly |     |     |
| 355                                                             | 360 | 365 |
| Thr Lys Gly Thr Ser Pro Lys Glu Ile Pro Ser Leu Lys Asp Asn Ile |     |     |
| 370                                                             | 375 | 380 |
| Glu Pro Ile Leu Lys Ala Arg Lys Glu Tyr Ala Tyr Gly Pro Gln His |     |     |
| 385                                                             | 390 | 395 |
| Asp Tyr Ile Asp His Pro Asp Val Ile Gly Trp Thr Arg Glu Gly Asp |     |     |
| 405                                                             | 410 | 415 |
| Ser Ser Ala Ala Lys Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro |     |     |
| 420                                                             | 425 | 430 |
| Gly Gly Ser Lys Arg Met Tyr Ala Gly Leu Lys Asn Ala Gly Glu Thr |     |     |
| 435                                                             | 440 | 445 |
| Trp Tyr Asp Ile Thr Gly Asn Arg Ser Asp Thr Val Lys Ile Gly Ser |     |     |
| 450                                                             | 455 | 460 |
| Asp Gly Trp Gly Glu Phe His Val Asn Asp Gly Ser Val Ser Ile Tyr |     |     |
| 465                                                             | 470 | 475 |
| Val Gln Lys                                                     |     |     |

&lt;210&gt; 7

&lt;211&gt; 1548

&lt;212&gt; DNA

&lt;213&gt; 嗜热脂肪芽孢杆菌

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1548)

&lt;400&gt; 7

```

gcc gca ccg ttt aac ggc acc atg atg cag tat ttt gaa tgg tac ttg 48
Ala Ala Pro Phe Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr Leu
  1           5           10           15

```

```

ccg gat gat ggc acg tta tgg acc aaa gtg gcc aat gaa gcc aac aac 96
Pro Asp Asp Gly Thr Leu Trp Thr Lys Val Ala Asn Glu Ala Asn Asn
          20           25           30

```

```

tta tcc agc ctt ggc atc acc gct ctt tgg ctg ccg ccc gct tac aaa 144
Leu Ser Ser Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys
          35           40           45

```

```

gga aca agc cgc agc gac gta ggg tac gga gta tac gac ttg tat gac 192
Gly Thr Ser Arg Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu Tyr Asp
          50           55           60

```

```

ctc ggc gaa ttc aat caa aaa ggg acc gtc cgc aca aaa tac gga aca 240
Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr
          65           70           75           80

```

```

aaa gct caa tat ctt caa gcc att caa gcc gcc cac gcc gct gga atg 288
Lys Ala Gln Tyr Leu Gln Ala Ile Gln Ala Ala His Ala Ala Gly Met
          85           90           95

```

```

caa gtg tac gcc gat gtc gtg ttc gac cat aaa ggc ggc gct gac ggc 336
Gln Val Tyr Ala Asp Val Val Phe Asp His Lys Gly Gly Ala Asp Gly
          100           105           110

```

```

acg gaa tgg gtg gac gcc gtc gaa gtc aat ccg tcc gac cgc aac caa 384
Thr Glu Trp Val Asp Ala Val Glu Val Asn Pro Ser Asp Arg Asn Gln
          115           120           125

```

```

gaa atc tcg ggc acc tat caa atc caa gca tgg acg aaa ttt gat ttt 432
Glu Ile Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe Asp Phe

```

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

| 305                                                             | 310 | 315 | 320 |      |
|-----------------------------------------------------------------|-----|-----|-----|------|
| aca ttg gcc gtc acc ttc gtt gat aat cat gac acc gaa ccc ggc caa |     |     |     | 1008 |
| Thr Leu Ala Val Thr Phe Val Asp Asn His Asp Thr Glu Pro Gly Gln |     |     |     |      |
| 325                                                             | 330 | 335 |     |      |
| gcg ctg cag tca tgg gtc gac cca tgg ttc aaa ccg ttg gct tac gcc |     |     |     | 1056 |
| Ala Leu Gln Ser Trp Val Asp Pro Trp Phe Lys Pro Leu Ala Tyr Ala |     |     |     |      |
| 340                                                             | 345 | 350 |     |      |
| ttt att cta act cgg cag gaa gga tac ccg tgc gtc ttt tat ggt gac |     |     |     | 1104 |
| Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr Gly Asp |     |     |     |      |
| 355                                                             | 360 | 365 |     |      |
| tat tat ggc att cca caa tat aac att cct tcg ctg aaa agc aaa atc |     |     |     | 1152 |
| Tyr Tyr Gly Ile Pro Gln Tyr Asn Ile Pro Ser Leu Lys Ser Lys Ile |     |     |     |      |
| 370                                                             | 375 | 380 |     |      |
| gat ccg ctc ctc atc gcg cgc agg gat tat gct tac gga acg caa cat |     |     |     | 1200 |
| Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr Gln His |     |     |     |      |
| 385                                                             | 390 | 395 | 400 |      |
| gat tat ctt gat cac tcc gac atc atc ggg tgg aca agg gaa ggg ggc |     |     |     | 1248 |
| Asp Tyr Leu Asp His Ser Asp Ile Ile Gly Trp Thr Arg Glu Gly Gly |     |     |     |      |
| 405                                                             | 410 | 415 |     |      |
| act gaa aaa cca gga tcc gga ctg gcc gca ctg atc acc gat ggg ccg |     |     |     | 1296 |
| Thr Glu Lys Pro Gly Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro |     |     |     |      |
| 420                                                             | 425 | 430 |     |      |
| gga gga agc aaa tgg atg tac gtt ggc aaa caa cac gct gga aaa gtg |     |     |     | 1344 |
| Gly Gly Ser Lys Trp Met Tyr Val Gly Lys Gln His Ala Gly Lys Val |     |     |     |      |
| 435                                                             | 440 | 445 |     |      |
| ttc tat gac ctt acc ggc aac cgg agt gac acc gtc acc atc aac agt |     |     |     | 1392 |
| Phe Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile Asn Ser |     |     |     |      |
| 450                                                             | 455 | 460 |     |      |
| gat gga tgg ggg gaa ttc aaa gtc aat ggc ggt tcg gtt tcg gtt tgg |     |     |     | 1440 |
| Asp Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser Val Trp |     |     |     |      |
| 465                                                             | 470 | 475 | 480 |      |
| gtt cct aga aaa acg acc gtt tct acc atc gct cgg ccg atc aca acc |     |     |     | 1488 |
| Val Pro Arg Lys Thr Thr Val Ser Thr Ile Ala Arg Pro Ile Thr Thr |     |     |     |      |

| 485                                                             | 490                 | 495         |      |
|-----------------------------------------------------------------|---------------------|-------------|------|
| cga ccg tgg act ggt gaa ttc gtc cgt tgg acc gaa cca cgg ttg gtg |                     |             | 1536 |
| Arg Pro Trp Thr Gly Glu Phe Val                                 | Arg Trp Thr Glu Pro | Arg Leu Val |      |
| 500                                                             | 505                 | 510         |      |
| gca tgg cct tga                                                 |                     |             | 1548 |
| Ala Trp Pro                                                     |                     |             |      |
| 515                                                             |                     |             |      |
| <210> 8                                                         |                     |             |      |
| <211> 515                                                       |                     |             |      |
| <212> PRT                                                       |                     |             |      |
| <213> 嗜热脂肪芽孢杆菌                                                  |                     |             |      |
| <400> 8                                                         |                     |             |      |
| Ala Ala Pro Phe Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr Leu |                     |             |      |
| 1                                                               | 5                   | 10          | 15   |
| Pro Asp Asp Gly Thr Leu Trp Thr Lys Val Ala Asn Glu Ala Asn Asn |                     |             |      |
| 20                                                              | 25                  | 30          |      |
| Leu Ser Ser Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys |                     |             |      |
| 35                                                              | 40                  | 45          |      |
| Gly Thr Ser Arg Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu Tyr Asp |                     |             |      |
| 50                                                              | 55                  | 60          |      |
| Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr |                     |             |      |
| 65                                                              | 70                  | 75          | 80   |
| Lys Ala Gln Tyr Leu Gln Ala Ile Gln Ala Ala His Ala Ala Gly Met |                     |             |      |
| 85                                                              | 90                  | 95          |      |
| Gln Val Tyr Ala Asp Val Val Phe Asp His Lys Gly Gly Ala Asp Gly |                     |             |      |
| 100                                                             | 105                 | 110         |      |
| Thr Glu Trp Val Asp Ala Val Glu Val Asn Pro Ser Asp Arg Asn Gln |                     |             |      |
| 115                                                             | 120                 | 125         |      |
| Glu Ile Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe Asp Phe |                     |             |      |
| 130                                                             | 135                 | 140         |      |

Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp Tyr His  
145 150 155 160

Phe Asp Gly Val Asp Trp Asp Glu Ser Arg Lys Leu Ser Arg Ile Tyr  
165 170 175

Lys Phe Arg Gly Ile Gly Lys Ala Trp Asp Trp Glu Val Asp Thr Glu  
180 185 190

Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met Asp His  
195 200 205

Pro Glu Val Val Thr Glu Leu Lys Asn Trp Gly Lys Trp Tyr Val Asn  
210 215 220

Thr Thr Asn Ile Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys  
225 230 235 240

Phe Ser Phe Phe Pro Asp Trp Leu Ser Tyr Val Arg Ser Gln Thr Gly  
245 250 255

Lys Pro Leu Phe Thr Val Gly Glu Tyr Trp Ser Tyr Asp Ile Asn Lys  
260 265 270

Leu His Asn Tyr Ile Thr Lys Thr Asp Gly Thr Met Ser Leu Phe Asp  
275 280 285

Ala Pro Leu His Asn Lys Phe Tyr Thr Ala Ser Lys Ser Gly Gly Ala  
290 295 300

Phe Asp Met Arg Thr Leu Met Thr Asn Thr Leu Met Lys Asp Gln Pro  
305 310 315 320

Thr Leu Ala Val Thr Phe Val Asp Asn His Asp Thr Glu Pro Gly Gln  
325 330 335

Ala Leu Gln Ser Trp Val Asp Pro Trp Phe Lys Pro Leu Ala Tyr Ala  
340 345 350

Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr Gly Asp  
355 360 365

Tyr Tyr Gly Ile Pro Gln Tyr Asn Ile Pro Ser Leu Lys Ser Lys Ile  
370 375 380

Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr Gln His  
385 390 395 400

Asp Tyr Leu Asp His Ser Asp Ile Ile Gly Trp Thr Arg Glu Gly Gly  
405 410 415

Thr Glu Lys Pro Gly Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro  
420 425 430

Gly Gly Ser Lys Trp Met Tyr Val Gly Lys Gln His Ala Gly Lys Val  
435 440 445

Phe Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile Asn Ser  
450 455 460

Asp Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser Val Trp  
465 470 475 480

Val Pro Arg Lys Thr Thr Val Ser Thr Ile Ala Arg Pro Ile Thr Thr  
485 490 495

Arg Pro Trp Thr Gly Glu Phe Val Arg Trp Thr Glu Pro Arg Leu Val  
500 505 510

Ala Trp Pro  
515

<210> 9

<211> 31

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述: 引物

<400> 9

ggtcgtaggc accgtagccc caatccgctt g

31

<210> 10

<211> 36

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 10

ggtcgtaggc accgtagccc caatcccatt ggctcg

36

<210> 11

<211> 28

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 11

ctgtgactgg tgagtactca accaagtc

28

<210> 12

<211> 31

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 12

ggtcgtaggc accgtagccc tcattccgctt g

31

<210> 13

<211> 31

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 13

ggtcgtaggc accgtagccc atatccgctt g

31



<210> 14

<211> 31

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 14

ggtcgtaggc accgtagcca atatccgctt g

31

<210> 15

<211> 36

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 15

gcagcatgga actgctyatg aagaggcacg tcaaac

36

<210> 16

<211> 30

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 16

catagttgcc gaattcattg gaaacttccc

30

<210> 17

<211> 34

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 17

catagttgcc gaattcaggg gaaacttccc aatc

34

<210> 18

<211> 41

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 18

ccgcgccccg ggaaatcaaa ttttgtccag gctttaatta g

41

<210> 19

<211> 32

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 19

caaaatggta ccaataccac ttaaaatcgc tg

32

<210> 20

<211> 29

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 20

cttccaatc ccaagtcttc ccttgaaac

29

<210> 21

<211> 36

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 21

cttaatttct gctacgacgt caggatggtc ataatc

36

<210> 22

<211> 38

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 22

cgcccaagtc attcgaccag tactcagcta ccgtaaac

38

<210> 23

<211> 29

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 23

gccgttttca ttgtcgactt cccaatccc

29

<210> 24

<211> 35

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 24

ggaatttcgc gctgactagt cccgtacata tcccc

35

<210> 25

<211> 36

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 25

ggcaggaatt tcgcgacctt tcgtcccgta catatc

36

<210> 26

<211> 36

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 26

cctcattctg cagcagcagc cgtaaattggc acgctg

36

<210> 27

<211> 38

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 27

ccagacggca gtaataccga tatccgataa atgttccg

38

<210> 28

<211> 30

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 28

cggatatcgg tattactgcc gtctggattc

30

<210> 29

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 29

ctcgtcccaa tcggttccgt c

21

<210> 30

<211> 26

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 30

gatgtatgcc gacttcgatt atgacc

26

<210> 31

<211> 30

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 31

catagttgcc gaattcattg gaaacttccc

30

<210> 32

<211> 24

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 32

ccgattgctg acgctgttat ttgc

24

<210> 33

<211> 25

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 33

gccaagcgga taacggctac ggtgc

25

<210> 34

<211> 28

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 34

gaacgagcca atcggacgtg ggctacgg

28

<210> 35

<211> 32

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 35

ggaacgagcc aatcggataa cggctacggt gc

32

<210> 36

<211> 25

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 36

gcatataagg, gactgagcca agcgg

25

<210> 37

<211> 25

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 37

caaccacaaa gccggcgctg atgcg

25

<210> 38

<211> 41

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 38

gcatataagg gactgagcca atcggataac ggctacggtg c

41

<210> 39

<211> 28

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 39  
gaacgagccg atcggacgtg ggctacgg 28

<210> 40  
<211> 28  
<212> DNA  
<213> 人工序列

<220>  
<223> 人工序列的描述：引物

<400> 40  
gaacgagcca aaacgacgtg ggctacgg 28