

[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 97194297.8

[45] 授权公告日 2006年3月22日

[11] 授权公告号 CN 1246455C

[22] 申请日 1997.4.30 [21] 申请号 97194297.8

[30] 优先权

[32] 1996.4.30 [33] DK [31] 0515/96

[32] 1996.6.28 [33] DK [31] 0712/96

[32] 1996.7.11 [33] DK [31] 0775/96

[32] 1996.11.8 [33] DK [31] 1263/96

[86] 国际申请 PCT/DK1997/000197 1997.4.30

[87] 国际公布 WO1997/041213 英 1997.11.6

[85] 进入国家阶段日期 1998.10.30

[71] 专利权人 诺沃奇梅兹有限公司

地址 丹麦鲍斯韦

[72] 发明人 A·斯文德森 T·V·博切特

H·比斯加德 - 弗兰特恩

审查员 张秀丽

[74] 专利代理机构 中国国际贸易促进委员会专利  
商标事务所

代理人 李 瑛

权利要求书3页 说明书67页 附图4页

[54] 发明名称

$\alpha$  - 淀粉酶变体

[57] 摘要

本发明涉及亲本类 - Termamyl  $\alpha$  - 淀粉酶的变体, 所说的变体具有  $\alpha$  - 淀粉酶活性, 并且显示出与所说亲本  $\alpha$  - 淀粉酶相关的下列性质中的至少一种改变: 底物特异性, 底物结合, 底物切割模式, 热稳定性, pH/活性轮廓, pH/稳定性轮廓, 氧化稳定性,  $\text{Ca}^{2+}$  依赖性 & 比活性。

1. 一种亲本类 Termamyl  $\alpha$ -淀粉酶之变体, 所述亲本具有与 SEQ ID NO: 2 氨基酸序列至少 80% 同源性的氨基酸序列, 该变体具有  $\alpha$ -淀粉酶活性, 并且显示出与所说亲本  $\alpha$ -淀粉酶相比的在 pH<7 的酸性 pH 和/或在低  $\text{Ca}^{2+}$  浓度下增加的热稳定性, 该变体含有对应于在 SEQ ID No. 2 中所示的氨基酸序列中的一种或多种下列突变:

在下列位置发生的取代: H156, A181, N188, N190, H205, D207, A209, A210, E211, Q264 或 N265, 包括 H156Y, D; A181T, E, D, Q, N, V; N188P; N190L, F; H205C; D207Y; A209L, T, V; A210S; E211Q; Q264A, E, L, K, S, T; N265A, S, T, Y, K, R, D; 或

取代: N188P + N190F;

H156Y+A181T+A209V;

H156Y+A181T+N190F+A209V+Q264S;

A1\*\*N2\*\*L3V+M15T+R23K+S29A+A30E+Y31H+A33S+E34D+H35I+H156Y+A181T+A209V;

A1\*\*N2\*\*L3V+M15T+R23K+S29A+A30E+Y31H+A33S+E34D+H35I+H156Y+A181T+A209V+Q264S;

A1\*\*N2\*\*L3V+M15T+R23K+S29A+A30E+Y31H+A33S+E34D+H35I+H156Y+A181T+N190F+A209V;

A1\*\*N2\*\*L3V+M15T+R23K+S29A+A30E+Y31H+A33S+E34D+H35I+H156Y+A181T+N190F+A209V+Q264S;

A1\*\*N2\*\*L3V+M15T+R23K+S29A+A30E+Y31H+A33S+E34D+H35I+H156Y+A181T+N190F+A209V+Q264S+E211Q; 或

A1\*\*N2\*\*L3V+M15T+R23K+S29A+A30E+Y31H+A33S+E34D+H35I+H156Y+A181T+N190F+H205C+D207Y+A209V+A210S+Q264S.

2. 按照权利要求 1 的变体, 其中所说的亲本类 Termamyl  $\alpha$ -淀粉酶选自:

地衣芽孢杆菌  $\alpha$ -淀粉酶, 其具有在 SEQ ID No. 2 所示的序列,

解淀粉芽孢杆菌 $\alpha$ -淀粉酶，其具有在 SEQ ID No. 4 所示的序列，嗜热脂肪芽孢杆菌 $\alpha$ -淀粉酶，其具有在 SEQ ID No. 6 所示的序列，芽孢杆菌菌株 NCIB 12512  $\alpha$ -淀粉酶，其具有在图 1 和图 2 所示的序列，芽孢杆菌菌株 NCIB 12513  $\alpha$ -淀粉酶，其具有在图 2 所示的序列，和 *Bacillus* sp. #707  $\alpha$ -淀粉酶，其具有在图 2 所示的序列。

3. 一种 DNA 构建体，其包括编码权利要求 1 的 $\alpha$ -淀粉酶变体的 DNA 序列。

4. 一种重组表达载体，其携带按照权利要求 3 的 DNA 构建体。

5. 一种细胞，其是由按照权利要求 3 的 DNA 构建体或权利要求 4 的载体转化的。

6. 按照权利要求 5 的细胞，其是微生物。

7. 按照权利要求 6 的细胞，其是细菌或真菌。

8. 按照权利要求 7 的细胞，其是革兰氏阳性菌，选自枯草芽孢杆菌，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，灿烂芽孢杆菌或者苏云金芽孢杆菌。

9. 按照权利要求 1 的 $\alpha$ -淀粉酶变体在洗涤中的用途。

10. 按照权利要求 9 的用途，其中所述洗涤为洗碟。

11. 按照权利要求 1 的 $\alpha$ -淀粉酶变体在纤维脱浆工艺中的用途。

12. 按照权利要求 1 的 $\alpha$ -淀粉酶变体在淀粉液化中的用途。

13. 一种去污添加剂，其包含按照权利要求 1 的 $\alpha$ -淀粉酶变体，其任选的可以是非尘粒形式，稳定化的液体形式或保护酶形式。

14. 按照权利要求 13 的去污添加剂，每克该添加剂含有 0.02-200 毫克的酶蛋白。

15. 按照权利要求 13 或 14 的去污添加剂，其还包含其它酶，选自蛋白酶，脂肪酶，过氧化物酶，另一种淀粉分解酶和/或纤维素酶。

16. 一种去污剂组合物，该组合物包含按照权利要求 1 的 $\alpha$ -淀粉酶变体。

17. 按照权利要求 16 的去污剂组合物, 其还包含其它酶, 选自蛋白酶, 脂肪酶, 过氧化物酶, 另一种淀粉分解酶和/或纤维素酶。

18. 一种手工或自动洗碟去污剂组合物, 该组合物包含按照权利要求 1 的 $\alpha$ -淀粉酶变体。

19. 按照权利要求 18 的手工或自动洗碟去污剂组合物, 其还包含其它酶, 选自蛋白酶, 脂肪酶, 过氧化物酶, 另一种淀粉分解酶和/或纤维素酶。

20. 一种手工或自动洗衣房洗涤组合物, 该组合物包含按照权利要求 1 的 $\alpha$ -淀粉酶变体。

21. 按照权利要求 20 的手工或自动洗衣房洗涤组合物, 其还包含其它酶, 选自蛋白酶, 脂肪酶, 过氧化物酶, 淀粉分解酶和/或纤维素酶。

22. 一种组合物, 该组合物包含:

(i) 得自地衣芽孢杆菌的具有 SEQ ID No. 2 所示序列的 $\alpha$ -淀粉酶与一种或多种按照权利要求 1 的变体的混合物, 所说的变体来源于作为亲本类 Termamyl  $\alpha$ -淀粉酶的具有 SEQ ID No. 6 所示序列的嗜热脂肪芽孢杆菌 $\alpha$ -淀粉酶; 或

(ii) 得自嗜热脂肪芽孢杆菌的具有 SEQ ID No. 6 所示序列的 $\alpha$ -淀粉酶与一种或多种按照权利要求 1 的变体的混合物, 所说的变体来源于一种或多种亲本类 Termamyl  $\alpha$ -淀粉酶; 或

(iii) 一种或多种来源于作为亲本类 Termamyl  $\alpha$ -淀粉酶的具有 SEQ ID No. 6 所示序列的嗜热脂肪芽孢杆菌 $\alpha$ -淀粉酶的按照权利要求 1 的变体, 与一种或多种来源于一种或多种亲本类 Termamyl  $\alpha$ -淀粉酶的按照权利要求 1 的变体的混合物。

## $\alpha$ -淀粉酶变体

### 发明领域

发明特别涉及亲本类-Termamyl  $\alpha$ -淀粉酶的新变体(突变体),尤其是相对于亲本而言显示出一种或多种性质上的改变的变体,所述改变对所说的变体在工业淀粉加工中(例如淀粉液化或者糖化)应用而言是有利的。

### 发明背景

$\alpha$ -淀粉酶( $\alpha$ -1,4-葡聚糖-4-葡聚糖氢化酶, EC 3.2.1.1)构成一组酶,其催化淀粉和其它线型和分支1,4-糖苷单糖和多糖的水解,并且有大量的与这类工业上十分重要的酶有关的专利和科学文献。

在最近的揭示 $\alpha$ -淀粉酶的有关文献中, WO 96/23874 提供了类-Termamyl  $\alpha$ -淀粉酶的晶体的三维, X-射线的结构数据,该淀粉酶由解淀粉芽孢杆菌的 $\alpha$ -淀粉酶(其具有 SEQ ID. No.4 中具有的氨基酸序列)的300个N端氨基酸和地衣芽孢杆菌 $\alpha$ -淀粉酶(其具有 SEQ ID.2 中具有的氨基酸序列)C端的端部301-483氨基酸而组成(后者以商业名 Termamyl<sup>TM</sup> 市售),并且该淀粉酶与工业上重要杆菌的 $\alpha$ -淀粉酶密切相关(其在本文中具有术语"类-Termamyl  $\alpha$ -淀粉酶"意义内的含义,并且包括,特别是,地衣芽孢杆菌,解淀粉芽孢杆菌和嗜热脂肪芽孢杆菌的 $\alpha$ -淀粉酶)。WO 96/23874 进一步描述了,在亲本类-Termamyl  $\alpha$ -淀粉酶亲本和类-Termamyl  $\alpha$ -淀粉酶变体(其具有与亲本相比性质上的改变)的结构分析基础上进行设计的方法。

### 发明简述

如上所述,本发明特别涉及亲本类-Termamyl  $\alpha$ -淀粉酶的新变体(突变体),尤其是显示出一些与淀粉的工业加工(淀粉液化作用,糖化等)相联系的优点的改变的性质的变体。

在本发明突变中得到改变的性质是在例如,底物特异性,底物结合,底物切割模式,热稳定性, pH/活性轮廓, pH/稳定性轮廓[例如在 pH 值

低(如 pH<6, 尤其是 pH<5)或高(如 pH>9)时增加的稳定性], 氧化稳定性, Ca<sup>2+</sup>依赖性及其活性和其他有利的性质等方面的改变。例如, 改变能导致产生, 与亲本类-Termamyl α-淀粉酶相比具有下降的 Ca<sup>2+</sup>依赖性和/或 pH/活性轮廓的变体。

本发明还特别涉及编码本发明的变体的 DNA 构建体, 制备本发明的变体的方法, 及本发明的变体在各种工业过程(例如淀粉液化作用)中单独或与其它α-淀粉酶结合起来的用途。

## 发明详述

### 类-Termamyl α-淀粉酶

众所周知, 许多由 *Bacillus spp.* 产生的α-淀粉酶在氨基酸水平上高度同源。例如, 人们发现, 具有 SEQ ID No. 2 中所显示的序列的地衣芽孢杆菌α-淀粉酶(可市售如 Termamyl™)与具有 SEQ ID No. 4 中所显示的氨基酸序列的解淀粉芽孢杆菌的α-淀粉酶有约 89 % 的同源性, 与具有 SEQ ID No. 6 中所显示的氨基酸序列的嗜热脂肪芽孢杆菌的α-淀粉酶有约 79 % 的同源性。进一步同源的α-淀粉酶包括来源于 *Bacillus sp.* 菌株 NCIB 12289, NCIB 12512, NCIB 12513 或 DSM 9375 的α-淀粉酶, 所有菌株在 WO 95/26397 得以详述, α-淀粉酶由 Tsukamoto 等在生化及生理研究通讯 151(1988), pp.25-31 所描述。进一步同源的α-淀粉酶包括由地衣芽孢杆菌菌株产生的α-淀粉酶(由 EP 0252666(ATCC 27811)所述), 和 WO 91/00353 与 WO 94/18314 中所测定的α-淀粉酶。其它市售的类-Termamyl 地衣芽孢杆菌α-淀粉酶是 Optitherm™ 和 Takatherm™(由 Solvay 提供), Maxamyl™(由 Gist-brocades/Genencor 提供), Spezym AA™(由 Genencor 提供), 及 Keistase™(由 Daiwa 提供)。

由于发现这些α-淀粉酶之间的大量同源性, 所以认为它们属于相同类型的α-淀粉酶, 即称为"类-Termamyl α-淀粉酶"类型。

因此在上下文中, 术语"类-Termamyl α-淀粉酶"旨在表明一类在氨基酸水平与 Termamyl™ 具有大量同源性的α-淀粉酶, 如具有 SEQ ID No. 2 中所显示氨基酸序列的地衣芽孢杆菌的α-淀粉酶。换句话说, 类-

Termamyl  $\alpha$ -淀粉酶是一类具有下列氨基酸序列的 $\alpha$ -淀粉酶: SEQ ID No. 2, No. 4 或 No. 6 所显示氨基酸序列, 或者 WO 95/26397 中 SEQ ID No. 1(其氨基酸序列见图 1 和图 2)或 WO 95/26397 中 SEQ ID No. 2(其氨基酸序列见图 2)中所显示的氨基酸序列, 或者在 Tsukamoto 等 1988(其氨基酸序列见图 2)中的氨基酸序列, 或者 i)与至少一种所说的氨基酸具有至少 60%, 如至少 70%, 如至少 75%, 或至少 80%, 如至少 85%, 至少 90%或至少 95%同源性的氨基酸序列, 和/或 ii)与至少抗一种所说的 $\alpha$ -淀粉酶产生的抗体发生免疫学交叉-反应的氨基酸序列, 和/或 iii)由与编码以上限定的 $\alpha$ -淀粉酶的 DNA 序列杂交的 DNA 序列编码的氨基酸序列, 这些序列分别貌似本申请中的 SEQ ID No. 1, 3 和 5(该编码序列分别编码本文 SEQ ID No. 2, 4 和 6 中所示的氨基酸序列), WO 95/26397 中的 SEQ ID No. 4(该 DNA 序列及终止密码子 TAA 示于本文图 1 中, 并且编码如本文图 1 所示的氨基酸序列), 及 WO 95/26397 中的 SEQ ID No. 5 的序列。

关于性质 i), "同源性"可用常规算法来检测, 优选采用 GCG 包版本 7.3(June 1993)利用 GAP 缺失的默认值[遗传计算机组(1991), GCG 包版本 7 的程序手册, 575 Science, Drive, Madison, Wisconsin, 美国 53711]。

$\alpha$ -淀粉酶的性质 ii)即免疫学交叉-反应性, 可利用至少抗类-Termamyl  $\alpha$ -淀粉酶一种表位或与之反应的抗体进行分析。该抗体(既能单克隆又能多克隆)可由本领域所熟知的方法制备, 如 1989 Hudson 等所描述的。免疫学交叉-反应性可用本领域所熟知的检测方法测定, 其例子是如 1989 Hudson 等所描述的 Western 杂交或径向免疫扩散分析。在这一方面, 分别在具有 SEQ ID No. 2, 4 和 6 氨基酸序列的 $\alpha$ -淀粉酶之间发现免疫学交叉-反应性。

用于类-Termamyl  $\alpha$ -淀粉酶鉴定(按照前面 iii)的性质)的寡核苷酸探针, 可根据所谈论的 $\alpha$ -淀粉酶的部分核苷酸或氨基酸适宜地制备。用以检验杂交的合适条件包括: 用 5 x SSC 先浸泡, 并在 20%甲酰胺、5 x Denhardt's 溶液、50mM 磷酸钠, PH 6.8, 和 50 $\mu$ g 变性 sonicated 小

牛胸腺 DNA 组成的溶液约 40 °C 预杂交一小时,接着在补充了 100 $\mu$ M ATP 的相同溶液约 40 °C 杂交 18 小时,或由 Sambrook 等 1989 所描述的其他方法。

在上下文中,“得自”不仅旨在表明由所谈论的有机体菌株产生或可产生的 $\alpha$ -淀粉酶,而且也表明由这些菌株中所分离出的 DNA 序列(且该序列由以所说的 DNA 序列转染的宿主而产生)编码的 $\alpha$ -淀粉酶。最后,术语旨在表明由合成和/或 cDNA 来源的 DNA 序列编码的 $\alpha$ -淀粉酶,且其具有所谈论 $\alpha$ -淀粉酶的特有特征。术语也旨在表明亲本 $\alpha$ -淀粉酶可能是天然存在的 $\alpha$ -淀粉酶的变体,即其为天然存在的 $\alpha$ -淀粉酶的一种或多种氨基酸基修饰(插入、取代、缺失)的结果。

#### 亲本杂种 $\alpha$ -淀粉酶

亲本 $\alpha$ -淀粉酶可能是杂种 $\alpha$ -淀粉酶,即具有得自至少两类 $\alpha$ -淀粉酶的部分氨基酸序列的组合物。 $\alpha$ -淀粉酶。

亲本杂种 $\alpha$ -淀粉酶是一种按照氨基酸同源性和/或免疫学交叉-反应性和/或 DNA 杂交(如上述定义)被确定为属于类-Termamyl  $\alpha$ -淀粉酶家族的 $\alpha$ -淀粉酶。在这种情况下,杂种 $\alpha$ -淀粉酶的典型组成是,至少一部分类-Termamyl  $\alpha$ -淀粉酶和一或更多部分得自微生物(细菌或真菌)和/或得自哺乳动物的 $\alpha$ -淀粉酶类-Termamyl  $\alpha$ -淀粉酶或者非类-Termamyl  $\alpha$ -淀粉酶。

这样,亲本杂种 $\alpha$ -淀粉酶包括几部分氨基酸序列的组合物,其来源于至少两种类-Termamyl  $\alpha$ -淀粉酶或者来源于至少一种类-Termamyl 和至少一种非类-Termamyl 细菌 $\alpha$ -淀粉酶,或来源于至少一种类-Termamyl 和至少一种真菌 $\alpha$ -淀粉酶。所说的类-Termamyl  $\alpha$ -淀粉酶(部分氨基酸序列来源于它)可能是例如,任何这里涉及的那些所明确的类-Termamyl  $\alpha$ -淀粉酶。

例如,亲本 $\alpha$ -淀粉酶可能具有来源于地衣芽孢杆菌菌株的 $\alpha$ -淀粉酶的 C-端部分,及来源于解淀粉芽孢杆菌的 $\alpha$ -淀粉酶的或来源于菌株嗜热脂肪芽孢杆菌菌株的 N-端部分。例如,亲本 $\alpha$ -淀粉酶可能具有地衣芽孢杆菌 $\alpha$ -淀粉酶 C-端部分至少 430 个氨基酸基,并且可能具有例如 a) 即一种对



应于解淀粉芽孢杆菌 $\alpha$ -淀粉酶 N-端部分的 37 个氨基酸基的氨基酸片段(其氨基酸序列如 SEQ ID NO. 2 中所示), 或 b)即一种对应于嗜热脂肪芽孢杆菌 $\alpha$ -淀粉酶 N-端部分的 68 个氨基酸基的氨基酸片段(其氨基酸序列如 SEQ ID NO. 6 中所示)及另一种对应于地衣芽孢杆菌 $\alpha$ -淀粉酶 C-端的 415 个氨基酸基的氨基酸片段(其氨基酸序列如 SEQ ID NO. 2 中所示)。

非类-Termamyl  $\alpha$ -淀粉酶可能是, 例如真菌 $\alpha$ -淀粉酶, 哺乳动物或植物 $\alpha$ -淀粉酶或者细菌 $\alpha$ -淀粉酶(其不同于类-Termamyl  $\alpha$ -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 $\alpha$ -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 $\alpha$ -淀粉酶的所阐明的结构明显不同于此处涉及的典型的类-Termamyl  $\alpha$ -淀粉酶结构。

上述真菌 $\alpha$ -淀粉酶(即来源于黑曲霉和米曲霉的 $\alpha$ -淀粉酶)在氨基酸水平上高度同源, 并且一般认为属于 $\alpha$ -淀粉酶的同一家族。来源于米曲霉的真菌 $\alpha$ -淀粉酶以商品名 Fungamyl<sup>TM</sup> 市售。

此外, 若以通常方式考虑特定类-Termamyl  $\alpha$ -淀粉酶的氨基酸序列中的氨基酸残基的修饰(例如缺失或取代)而考虑类-Termamyl  $\alpha$ -淀粉酶的特定变体(本发明的变体)时, 则认为包括另一种在相同位置修饰的类-Termamyl  $\alpha$ -淀粉酶变体(由各自的氨基酸序列之间最佳的氨基酸序列组合所确定)。

本发明的变体优选的实施方案是来源于地衣芽孢杆菌 $\alpha$ -淀粉酶(作为亲本类-Termamyl $\alpha$ -淀粉酶)的一种, 例如上述中的一种, 如具有在 SEQ ID NO. 2 中所示的氨基酸序列的地衣芽孢杆菌 $\alpha$ -淀粉酶。

#### 本发明的变体的构建

在产生突变的培养条件下, 通过培养具有编码变体的 DNA 序列的微生物从而完成所需变体的构建。然后从所形成的培养肉汤中回收变体。以下进一步详述这点。

#### 本发明的变体的改变的性质

下面讨论本发明的变体呈现的改变之间的关系, 及由此所形成性质中符合要求的改变(相对亲本类-Termamyl  $\alpha$ -淀粉酶)。

### 下降的 $\text{Ca}^{2+}$ 依赖性

降低类-Termamyl  $\alpha$ -淀粉酶的  $\text{Ca}^{2+}$  依赖性是非常值得做的。因此，本发明的一个方面涉及亲本类-Termamyl  $\alpha$ -淀粉酶变体，该变体具有  $\alpha$ -淀粉酶活性，并且具有与亲本  $\alpha$ -淀粉酶相比下降的  $\text{Ca}^{2+}$  依赖性。一般下降的  $\text{Ca}^{2+}$  依赖性将导致功能上的结果，即变体在钙离子浓度较低的体外培养基的情况下具有令人满意淀粉分解活性，而钙离子浓度对亲本酶来说是必要的。其常进一步产生这种结果即变体与亲本相比对钙离子-减少情况敏感性较低，如那些在含钙-络和剂的培养基(如某种去污剂)中所获得的。

例如，通过增加  $\text{Ca}^{2+}$  与亲本类-Termamyl  $\alpha$ -淀粉酶的结合亲和性，可有利地完成本发明的变体的下降  $\text{Ca}^{2+}$  依赖性；换句话说  $\text{Ca}^{2+}$  在酶中结合越强， $\text{Ca}^{2+}$  依赖性越低。

这里可能提到 WO 96/23874 阐述了钠或钙离子的定位在  $10\text{\AA}$  之内的氨基酸基被认为参与了酶的  $\text{Ca}^{2+}$  结合，或对其很重要，并且在这一联系中变体 N104D [具有 SEQ ID NO. 2 中所示的氨基酸序列的地衣芽孢杆菌  $\alpha$ -淀粉酶的变体或者另一种在相同位置的类-Termamyl  $\alpha$ -淀粉酶的等价的变体]被认为在类-Termamyl  $\alpha$ -淀粉酶的下降  $\text{Ca}^{2+}$  依赖性方面具有特殊的关系。

WO 96/23874 中所论及的与  $\text{Ca}^{2+}$  依赖性相联系的可能很重要的其它变体，包括此处通过化合形成例如半胱氨酸桥或盐桥而使类-Termamyl  $\alpha$ -淀粉酶三维结构的 C-结构域稳定化(如 WO 96/23874 所定义)，从而期待获得增加的钙结合性的变体。这样，WO 96/23874 揭示了地衣芽孢杆菌  $\alpha$ -淀粉酶的 C-结构域(其具有 SEQ ID NO. 2 中所示的氨基酸序列)的稳定化，其通过引入结构域 A 和 C 之间的半胱氨酸桥(如 WO 96/23874 所定义)及通过引入下列突变：

A349C+I479C 和/或 L346C+I430C。

WO 96/23874 同样揭示了为获得盐桥，可通过引入一个或多个下列突变到地衣芽孢杆菌  $\alpha$ -淀粉酶(其具有 SEQ ID NO. 2 中所示的氨基酸序列)中：

N457D, E

**N457D, E +K385R**

**F350D, E +I430R, K**

**F350D, E +I411R, K**

用任何其它氨基酸残基取代 H408 和/或 G303 氨基酸基从而可以稳定 C 结构域的钙位点, 尤其是引入下列取代基:

**H408Q, E, N, D 和/或 G303N, D, Q, E**

其被期待能提供更好的钙结合或保护以免受钙耗竭。

(此处包括在其它类-Termamyl  $\alpha$ -淀粉酶的等同位置上的类似变体)

WO 96/23874 中所揭示的有明显重要性(尤其在降低钙依赖性的情况下)的其它取代变体(相对于地衣芽孢杆菌 $\alpha$ -淀粉酶, SEQ ID NO. 2)包括下列: R23K, H156Y, A181T, A209V, R214, G310D 和 P345(或在另一类-Termamyl  $\alpha$ -淀粉酶的等同位置上的等同变体)。

在本发明的内容中, 其它取代突变(其特别是在关于钙依赖性的减弱上具有重要性)包括结构域 B 的下列突变(如 WO 96/23874 中所定义的):

A181E, D, Q, N, V(其导致在结构域 A 和结构域 B 之间的连接区中远离中心的  $\text{Ca}^{2+}$  结合位点在某种程度上的遮蔽);

I20I(大氨基酸), 如 I201W, F, L,(其导致在结构域 A 和结构域 B 之间的连接区中的紧挨着  $\text{Ca}^{2+}$ - $\text{Na}^{+}$ - $\text{Ca}^{2+}$  结合位点邻近区的几何型方面轻微的改变, 及洞/穴邻近的几何型方面和/或大小的轻微的改变); 及

Y203E, Q(认为其导致最外部  $\text{Ca}^{2+}$  离子更强的结合在位于结构域 A 和结构域 B 之间的连接区中的其结合位点上);

(或者另一种类-Termamyl  $\alpha$ -淀粉酶在等同的位置上的等同变体)

改变的最适 pH(改变的 pH/活性轮廓)

WO 96/23874 揭示了: 通过改变活性位点残基的 pKa 而改变类-Termamyl  $\alpha$ -淀粉酶的最适 pH 或改变在给定 pH 下的其酶促活性这一操作是有可能的, 并且这一操作还可以通过例如改变将被修饰的氨基酸和紧围绕着它的氨基酸侧链的官能团之间的静电作用或疏水相互作用而完成。

在本发明文中, 按照关于类-Termamyl  $\alpha$ -淀粉酶三维结构(WO

96/23874 中所揭示的)的静电因素[见, 例如, M.K. Gilson, 结构生物学的当前观点 5(1995) pp.216-223; 及 B.Honig 和 A. Nicholls, 科学 268(1995)pp. 1144-1149; 和那里所给出的参考文献]和疏水因素, 人们认为有关变体, 特别对于改变(增加或者减少)类-Termamyl  $\alpha$ -淀粉酶的最适 pH 的变体, 包括下列变体或其等同变体[此处参考地衣芽孢杆菌 $\alpha$ -淀粉酶的序列(SEQ ID NO. 2)]:

Q9K, L, E; F11R, K, E; E12Q; D100N, L; V101H, R, K, D, E, F; V102A, T; I103H, K; N104R, K, D; H105R, K, D, E, W, F; L196R, K, D, E, F, Y; I212R, K, D, E; L230H, K, I; A232G, H, F, S, V; V233D; K234L, E; I236R, K, N, H, D, E; L241R, K, D, E, F; A260S; W263H; 6Q264R, D, K, E; N265K, R, D; A269R, K, D, E; L270R, K, H, D, E; V283H, D; 8F284H; D285N, L; V286R, K, H, D, E; Y290R, E; V312R, K, D, E; F323H; D325N; N326K, H, D, L; H327Q, N, E, D, F; Q330L, E; G332D; Q333R, K, H, E, L; S334A, V, T, L, I, D; L335G, A, S, T, N; E336R, R375E; T337D, K; T338D, E; T339D; Q360K, R, E; D365N; G371D, R;

#### 低(酸性)pH 增加的稳定性

在本发明文中, 关于在低 pH 时达到增加稳定性的重要突变(氨基酸取代)包括对应于地衣芽孢杆菌 $\alpha$ -淀粉酶中的具有 SEQ ID NO. 2 中所示的氨基酸序列的下列突变:

在 H68, H91, H247, R305, K306, H382, K389, H405, H406, H450 或 R483 位置的突变;

突变:

H140Y; H142Y; H156Y; H159Y; H140D+H142R; H140K+H142D; 或 H142Y+H156Y

及任何两个或者更多的这些变体的组合。

增加热稳定性和/或改变的最适温度(改变的温度/活性轮廓)

本发明进一步涉及的方面是亲本类-Termamyl  $\alpha$ -淀粉酶的变体, 这种变体是在亲本 $\alpha$ -淀粉酶中缺失取代或添加一个或多个氨基酸残基, 以达

到增加变体的热稳定性的结果。

可能提到在关于达到增加的热稳定性中，WO 96/23874 揭示了类-Termamyl  $\alpha$ -淀粉酶的极为有意义的变体包括对应于下列突变之一的突变(利用 SEQ ID NO. 2 中的地衣芽孢杆菌 $\alpha$ -淀粉酶氨基酸序列的编号):

L61W, V, F;

Y62w;

F67W;

K106R, F, W;

G145F, W

I212F, L, W, Y, R, K;

S151 以任何其它氨基酸残基予以取代，特别是以 F, W, I 或 L;

R214W;

Y150R, K;

F143W; 和/或

R146W。

在这一关连中 WO 96/23874 进一步揭示了：对应于地衣芽孢杆菌 $\alpha$ -淀粉酶的具有 SEQ ID NO. 2 中所示氨基酸序列的下列变体的一种或多种变体，其在关于完成与亲本 $\alpha$ -淀粉酶相比的增加热稳定性中具有意义：

L241I, F, Y, W; 和/或

I236L, F, Y, W

L7F, I, W

V259F, I, L

F284W

F350W

F343W

L427F, L, W

V481, F, I, L, W

在本发明的内容中，从来源于各种杆菌种的 $\alpha$ -淀粉酶的氨基酸序列的线形排列，可看到地衣芽孢杆菌 $\alpha$ -淀粉酶和解淀粉芽孢杆菌 $\alpha$ -淀粉酶都含

有与如嗜热脂肪芽孢杆菌 $\alpha$ -淀粉酶有关的三个氨基酸的"插入"。

从按照 WO 96/23784 中所揭示的类-Termamyl  $\alpha$ -淀粉酶的三维结构(参考前面)并考虑到地衣芽孢杆菌 $\alpha$ -淀粉酶与被谈论的类-Termamyl  $\alpha$ -淀粉酶的同源性的基础上所建造的地衣芽孢杆菌 $\alpha$ -淀粉酶的结构模型中,可看到在 WO 96/23784 中上述"插入"位于称为"环 8"结构的一部分内,使得地衣芽孢杆菌 $\alpha$ -淀粉酶的该环比类-Termamyl  $\alpha$ -淀粉酶的大,并且引起环从结构中突出来,因此有可能使结构不稳定。因此人们认为所谈论的在地衣芽孢杆菌或解淀粉芽孢杆菌的 $\alpha$ -淀粉酶中的一个或多个氨基酸缺失将提高这些 $\alpha$ -淀粉酶的热稳定性。

这一关连中尤其有意义的是,从 T369 至 I377 的部分序列(参考如 SEQ ID NO. 2 中所示的地衣芽孢杆菌 $\alpha$ -淀粉酶的氨基酸序列)内的三个氨基酸的缺失,即所说的部分序列: T369-K370-G371-D372-S373-Q374-R375-E376-I377(或在解淀粉芽孢杆菌 $\alpha$ -淀粉酶中所对应的部分序列)。除这样的缺失外,后者部分序列之内的一个或多个未缺失氨基酸的取代也有利。

从 T369 至 I377 部分序列(在地衣芽孢杆菌 $\alpha$ -淀粉酶中)中三个氨基酸的优选缺失是 K370+G371+D372(即 K370\*+G371\*+D372\*)或 D372+S373+Q374(即 D372\*+S373\*+Q374\*)(或在解淀粉芽孢杆菌 $\alpha$ -淀粉酶中所对应的部分序列中的等同缺失)。

另一种对提高这些 $\alpha$ -淀粉酶的热稳定性有价值的突变类型是,从 T369 至 I377 整个该部分氨基酸序列(参考地衣芽孢杆菌 $\alpha$ -淀粉酶的氨基酸序列)被下列六个氨基酸的部分序列之一(编号从左至右增加): I-P-T-H-S-V; I-P-T-H-G-V; 和 I-P-Q-Y-N-I(或解淀粉芽孢杆菌 $\alpha$ -淀粉酶所对应的部分序列的相同取代的一种)所取代(替换)。

其它突变(其对于达到增加的热稳定性具有一定重要性)包括在下列位置的氨基酸取代(参考 SEQ ID NO. 2):

R169(例如 R169I, L, F, T);

R173(尤其是 R173I, L, F, T);

I201F;

**I212F;**

**A209L, T; 或**

**V208I**

及这些变体的任何两种或者多种的组合。

酸性 pH 和/或低浓度  $\text{Ca}^{2+}$  下增加的热稳定性

在本发明的内容中, 对于按照本发明获得具有酸性 pH (pH 小于 7) 和/或低浓度  $\text{Ca}^{2+}$  下增加的热稳定性的变体呈现特殊相关的突变包括在下列位置的突变(与地衣芽孢杆菌  $\alpha$ -淀粉酶相关, SEQ ID NO. 2):

**H156, N172, A181, N188, N190, H205, D207, A209, A210, E211, Q264, N265**

这里可能提到, 分别对应于 SEQ ID NO. 2 中的 N109 和 E211 位置的 N 和 E 氨基酸残基, 构成在许多类-Termamyl  $\alpha$ -淀粉酶中的保守的残基。这样, 例如, 这些已经提及的许多类-Termamyl  $\alpha$ -淀粉酶的氨基酸序列中对应位置的残基(参考前面)如下:

类-Termamyl $\alpha$ -淀粉酶	N 位置	E 位置
地衣芽孢杆菌(SEQ ID No. 2)	N190	E211
解淀粉芽孢杆菌(SEQ ID No. 4)	N190	E211
嗜热脂肪芽孢杆菌(SEQ ID No. 6)	N193	E210
芽孢杆菌 NCIB 12512(WO 95/26397)	N195	E212
芽孢杆菌 NCIB 12513(WO 95/26397)	N195	E212
" <i>Bacillus</i> sp. #707"(Tsukamoto 等)	N195	E212

这些保守氨基酸基的突变对于在酸性 pH 和/或低钙浓度下提高热稳定性十分重要, 并且下列突变对这一点尤其有意义(参考在 SEQ ID NO. 2 中所示的地衣芽孢杆菌氨基酸序列的编号):

**H156Y, D**

**N172R, H, K**

**A181T**

**N188P**

**N190L, F**

H205C

D207Y

A209L, T, V

A210S

E211Q

Q264A, E, L, K, S, T

N265A, S, T, Y

以及这些变体的两种或多种的任何组合。

在这一点上一种特别有意义的双突变体的例子是 Q264S +N265Y。

#### 改变的切割模式

在淀粉液化方法中，优选采用能将淀粉分子降解成长链、分支寡糖的 $\alpha$ -淀粉酶而非使其形成短链、分支寡糖的 $\alpha$ -淀粉酶(如通常的类-Termamyl  $\alpha$ -淀粉酶)。短链、分支寡糖(6- $\alpha$ -葡糖基麦芽糖前体)用支链淀粉酶(其在液化方法中淀粉处理之后，但在加入糖化淀粉葡糖苷酶(葡糖淀粉酶)之前加以利用)不能满意地将其水解。这样，在6- $\alpha$ -葡糖基麦芽糖前体的存在下，在葡糖淀粉酶处理后的当前产物混合物含有重要比例的短链、分支，所谓极限糊精，即是三糖6- $\alpha$ -葡糖基麦芽糖。6- $\alpha$ -葡糖基麦芽糖的存在明显降低了糖化产率，所以不合适。

这样，本发明的目的是得到一种具有适当地修饰降解-淀粉特性且保持亲本类-Termamyl  $\alpha$ -淀粉酶热稳定性的 $\alpha$ -淀粉酶变体。

这里可能提到，按照 WO 96/23874 具有至少下列突变中一种的突变被期待阻止接近分支位点的切割：

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

DS3L, I, F, Y, W

Y56W

Q333W

G57 所有可能的氨基酸基

大于 A 的 A52 氨基酸基，例如 A52W, Y, L, F, I。

增加的比活性



本发明的另一方面，关于获得具有增加的比活性的变体的重要突变包括对应于地衣芽孢杆菌 $\alpha$ -淀粉酶的具有 SEQ ID NO. 2 中所示氨基酸序列的下列突变：

在位置 S187(尤其是 S187D)或者 Q264(如 Q264R, K, S)的突变(氨基酸取代)；

在位置 Y290(尤其是 Y290E, K)的突变(取代)；

V54I 突变；

及后两种突变的两个或多个任意组合，或者后两种突变的一个，两个或多个突变与下列多突变的组合：

A1\*+N2\*+L3V+M15T+R23K+S29A+A30E+Y31H+A33S+E34D+H35I

#### 本发明的变体的一般突变

优选的为，本发明的变体含有除以上所列出的那些外还含有一种或多种修饰。这样，有利的是，位于被修饰 $\alpha$ -淀粉酶变体的部位中的一个或多个脯氨酸残基可被任何可能的，天然存在的非-脯氨酸基予以取代，并且优选的是丙氨酸，甘氨酸，丝氨酸，苏氨酸，缬氨酸或亮氨酸。

同样，可能优选的是，位于用来进行亲本 $\alpha$ -淀粉酶修饰的氨基酸中的一个或多个半胱氨酸残基可被非-半胱氨酸基诸如丝氨酸，丙氨酸，苏氨酸，甘氨酸，缬氨酸或亮氨酸所取代。

此外，本发明的变体(既是一种修饰也可是任何上述所列出的修饰的组合)可被修饰，那么位于对应于 SEQ ID NO. 2 185-209 氨基酸片段中的一个或多个 Asp 和/或 Glu 分别被 Asn 和/或 Gln 所取代。在类 Termamyl  $\alpha$ -淀粉酶中，精氨酸取代位于对应于 SEQ ID NO. 2 185-209 氨基酸片段中的一个或多个 Lys 基也是有意义的。

可认为本发明包括上述所列出修饰的两个或者多个结合的变体。

此外，在这里所描述的任何变体中引入点突变是有利的。

#### 制备 $\alpha$ -淀粉酶变体的方法

本领域内熟知用于将突变导入基因的几种方法。简述克隆 $\alpha$ -淀粉酶-编码 DNA 序列后，讨论用于产生在 $\alpha$ -淀粉酶-编码序列内特异位点处突变的方法。

### 克隆编码 $\alpha$ -淀粉酶的 DNA 序列

利用本领域熟知的各种方法，将编码亲本 $\alpha$ -淀粉酶的 DNA 序列从任何产生所谈论 $\alpha$ -淀粉酶的细胞或微生物中分离出来。首先，用来源于能产生所研究 $\alpha$ -淀粉酶的有机体的染色体 DNA 或信使 RNA 构建基因组 DNA 和/或 cDNA 文库。然后，若 $\alpha$ -淀粉酶的氨基酸序列是已知的，则合成同源的标记寡核苷酸探针，并且用来从由所说的有机体制备的基因组文库中识别 $\alpha$ -淀粉酶-编码的克隆。此外，利用杂交和严格程度较低的洗脱条件，可将具有与已知 $\alpha$ -淀粉酶基因同源序列的标记寡核苷酸探针作为识别 $\alpha$ -淀粉酶-编码克隆的探针。

目前，另一种用于识别 $\alpha$ -淀粉酶-编码克隆的方法包括将基因组 DNA 片段插入到表达载体如质粒，并用所产生的 cDNA 文库转染 $\alpha$ -淀粉酶-阴性细菌，然后将转染的细菌平板接种在含 $\alpha$ -淀粉酶底物的琼脂上，由此识别表达 $\alpha$ -淀粉酶的克隆。

此外，编码酶的 DNA 序列可用所建立的标准方法而合成制备，这些方法如 S. L. Beaucage 和 M. H. Caruthers(1981)所描述的亚磷酰胺方法或 Matthes 等(1984)所描述的方法。在亚磷酰胺方法中，寡核苷酸是在，如自动 DNA 合成仪中纯化，退火，连接并在适当载体中克隆。

最后，按照标准技术，通过连接合成的、基因组的或 cDNA 来源的片段(如对应于整个 DNA 序列的各种部分的适当片段)，则该 DNA 序列可能是基因组和合成混合来源的，合成和 cDNA 混合来源的或者基因组和 cDNA 混合来源的。用特异性引物通过聚合酶连锁反应(PCR)也可制备 DNA 序列，例如在 US 4,683,202 或 R.K. Saiki 等(1988)中所描述的。

### 定点诱变

一旦分离出 $\alpha$ -淀粉酶-编码 DNA 序列，并测定了突变的理想位点，则可利用合成的寡核苷酸诱发突变。这些寡核苷酸具有位于所需的变体位点两侧的核苷酸序列；且变体核苷酸在寡核苷酸合成期间插入。在具体的方法中，在携带 $\alpha$ -淀粉酶基因的载体中可产生桥接 $\alpha$ -淀粉酶-编码序列的 DNA 单链的缺口。然后，将携带所需突变的合成核苷酸退火到单链 DNA 的同源部分上。接着用 DNA 聚合酶 I(克列诺片段)填充余下的缺口，并用

T4 连接酶连接构建体。Morinaga 等(1984)描述了这一方法的具体例子。US 4,760,025 揭示了通过操作该盒的小改变而导入编码若干变体的寡核苷酸。然而,更大的某种变体用 Morinaga 方法可在任何时间导入,因为其能导入各种长度的大量寡核苷酸。

在 Nelson 和 Long(1989)中描述了另一种将突变导入 $\alpha$ -淀粉酶-编码 DNA 序列的方法。它包括 PCR 片段的 3 步产物,其具有通过利用化学合成的 DNA 作为 PCR 反应中引物之一所诱发的理想突变。通过用限制核酸内切酶切割,将携带突变的 DNA 片段从 PCR-产生的片段中分离出来,并再插入表达质粒。

### 随机诱变

至少在基因(其编码所说的氨基酸序列)三个部分或整个基因之内,可适宜地进行随机诱变如定位的或特定-区的随机诱变。

WO 96/23874 揭示了:与通过类-Termamyl  $\alpha$ -淀粉酶变体(其被修饰成(如更高)的底物特异性和/或修饰成(如更高)的底物切割(水解)特异性)而完成提高底物结合力(即提高碳水化合物物种的结合力,如直链淀粉或支链淀粉)过程有关的下列密码子位置(其编码的氨基酸序列如 SEQ ID NO. 2 所示)(或者本发明的内容中另一亲本 Termamyl  $\alpha$ -淀粉酶的等同的密码子位置)特别适于作靶位点:

13-18 50-56 70-76 102-109 163-172 189-199 229-235 360-364  
327-335

### 低 pH 和低浓度钙离子下提高液化作用效能

对于用于淀粉液化作用过程的 $\alpha$ -淀粉酶,具有热稳定性并能在低 pH 和低浓度钙下起作用是格外有意义的。为了提高亲本类-Termamyl  $\alpha$ -淀粉酶(特别是地衣芽孢杆菌 $\alpha$ -淀粉酶或其变体或其杂种细胞的这些性质),可进行随机诱变(优选利用添加的或掺入的寡核苷酸引物)接着对所形成的变体酶作适当的选择。随机区域的选择方向和添加的选择方向主要取决于已存在的钙离子的稳定化和低 pH 下的残基/残基或结构域/结构域之间的静电作用中的提高。此外,可选择已显示出包括对达到好的淀粉液化作用效能很重要的位置的区域。

为制备具有上述性质的亲本类-Termamyl  $\alpha$ -淀粉酶变体，可以有利地对下列区域中至少一种进行随机诱变(氨基酸基的编号如 SEQ ID NO. 2):

区	残基	描述
I:	153-163	结构域 A 和 B 之间的钙区，也含有 H156
II:	178-192	结构域 A 和 B 之间的钙区
III:	193-214	结构域 A 和 B 之间的钙区，也含有 A209
IV:	232-237	结构域 A 和 B 之间的钙区
V:	297-308	结构域 A 和 C 之间的钙区
VI:	403-409	结构域 A 和 C 之间的钙区
VII:	428-435	结构域 A 和 C 之间的钙区
VIII:	131-136	含有 H133 的区
IX:	164-175	与 H133 区联系的区
X:	262-278	含有 4264 的区

优选地是，在本发明的新 $\alpha$ -淀粉酶变体的构建中，对上述区的两个，三个或四个进行随机诱变。例如，下列区的组合非常适于进行随机诱变：

VIII+IX

VIII +IX +II

II+III +IV

IV +I

此外，更为优选的是，通过利用添加或掺入的寡核苷酸进行诱变。优选添加是为了导入有助于提高所形成的 $\alpha$ -淀粉酶变体在低 pH 时的稳定性和降低钙依赖性的氨基酸。此外，当选择添加方案时，由于 Asn 和 Gln 残基与低 pH 时的不稳定性有关，一般应避免导入 Asn 和 Gln 残基的可能性。优选地，当插入 Pro 残基具有潜在利益时(例如从脯氨酸结构考虑因素作估计)，进行添加方案包括优选导入 Pro 残基。

按照上述原则要进行随机诱变亲本类-Termamyl  $\alpha$ -淀粉酶可以是任何具有一种或多种突变的野生型 $\alpha$ -淀粉酶或者其变体。亲本可能是这里进一步详述的 $\alpha$ -淀粉酶中的至少两个之间的杂种。优选地，亲本 $\alpha$ -淀粉酶是

地衣芽孢杆菌 $\alpha$ -淀粉酶的变体，其序列如 SEQ ID NO. 2 中所示且具有至少一种变体，优选为多变体。亲本 $\alpha$ -淀粉酶另一选择是杂种 $\alpha$ -淀粉酶，至少含有地衣芽孢杆菌(SEQ ID NO. 2) $\alpha$ -淀粉酶的一部分。按照上述原则适宜于诱变的亲本 $\alpha$ -淀粉酶的具体例子包括：含有突变 H156Y， A181T， N190F， A209V 以及 Q264S 至少一种，即一，二，三，四或所有五个的地衣芽孢杆菌 $\alpha$ -淀粉酶(SEQ ID NO. 2)的变体；含有地衣芽孢杆菌 $\alpha$ -淀粉酶(SEQ ID NO. 2)一部分的杂种 $\alpha$ -淀粉酶，其优选为 C 端部分，如它的 35-483 个氨基酸，及另一种类-Termamyl  $\alpha$ -淀粉酶的一部分如解淀粉芽孢杆菌 $\alpha$ -淀粉酶(SEQ ID NO. 4)，其优选为 N 端部分如它的前 38 个氨基酸残基。

关于上述，本发明的另一方面涉及用于产生亲本类-Termamyl  $\alpha$ -淀粉酶变体的方法，所说的变体在低 pH 和低钙浓度时具有与亲本相比增加的稳定性，该方法包括：

- (a) 使编码亲本类-Termamyl  $\alpha$ -淀粉酶的 DNA 序列发生随机诱变，
- (b) 在宿主中表达步骤(a)中获得的变体 DNA 序列，并且
- (c) 筛选表达变体 $\alpha$ -淀粉酶的宿主细胞，其中所说的变体 $\alpha$ -淀粉酶具有在低 pH 和低钙浓度时与亲本相比增加的稳定性。

本发明后一种方法的步骤(a)优选利用掺加 ( doped ) 引物而进行，如本文实施例中所述的(参考前面)。

#### 产生随机诱变的方法

按照本发明上述方法的步骤(a)将能产生的编码亲本 $\alpha$ -淀粉酶的 DNA 序列的随机诱变，用任何本领域所熟知的方法即可便利地产生。

例如，通过利用合适的物理或化学诱变剂，通过利用合适的寡核苷酸，或使 DNA 序列发生 PCR 产生的诱变均可产生随机诱变。此外，利用这些诱变剂的任何组合可产生随机诱变。

诱变剂可能是，例如引起转变，横断，倒位，重叠，缺失，和/或插入的一种物质。

适于目前目标的物理或化学诱变剂的例子包括紫外(UV)扩散，羟胺，N-甲基-N'-硝基-N-亚硝基胍(MNNG)，O-甲基羟胺，亚硝酸，乙基甲烷

磺酰, 钠硫酸氢盐, 甲酸, 及核苷酸类似物。

当利用这些试剂时, 经过诱发在适于诱变发生的条件和所选择的诱变剂的存在下将会发生诱变的 DNA 序列(其编码亲本酶)进行诱变的典型操作, 并选择具有所需的性质的突变 DNA。

当利用寡核苷酸完成诱变时, 寡核苷酸合成期间在将要改变的位置处, 可添加或用三个非-亲本核苷酸掺入。做添加或掺入是为了避免得到不要的氨基酸的密码子。采用任何所出版的技术, 利用如 PCR, LCR 或任何 DNA 聚合酶和连接酶将添加或掺入寡核苷酸导入编码淀粉酶的 DNA 内。

优选地, 利用"经常随机掺加"进行掺加 (doping), 其中预先规定每一种位置的野生型和变体的百分比。此外, 指定的掺加倾向于一定核苷酸的导入, 因此, 优先选择一种或多种特定氨基酸基的导入。例如, 进行掺加是为了使在每一种位置导入 90%野生型和 10%变体。掺加方案选择中的额外因素是基因和蛋白质结构的限制。添加方案可利用 DOPE 方法(参见此处的实施例)而实施, 该方法特别保证避免终止密码子的导入。

当利用 PCR-产生的诱变时, 在增加核苷酸错配的条件下, 将经化学处理或未处理的编码亲本 $\alpha$ -淀粉酶的基因进行 PCR(Deshler 1992; Leung 等, 技术, 卷1, 1989, pp. 11-15)。

大肠杆菌的变体菌株(Fowler 等, Molec. Gen. Genet., 133, 1974, pp. 179-191), 啤酒糖酵母或任何其它微生物有机体, 通过如转化具有亲本酶的质粒, 培养带有质粒的变体菌株, 并把突变质粒从变体菌株中分离出, 可用于编码淀粉酶 DNA 的随机诱变。其后将突变质粒转染进表达有机体内。

发生诱变的 DNA 序列可能存在于从表达亲本淀粉酶的有机体所制备的基因或 cDNA 文库中。另一可取的是, 该 DNA 序列可能存在于合适的载体如质粒或细菌噬菌体中, 这些如同那些用诱变剂培育或以其它方式暴露在诱变剂下。通过整合到所说细胞的基因组中或存在于细胞内的载体中, 发生诱变的 DNA 也能存在于宿主细胞中。最后, 诱变的 DNA 可以是分离的形式。认为, 易发生随机诱变的 DNA 序列优选为 cDNA 或基因

组 DNA 序列。

在某些情况下，在表达步骤(b)或筛选步骤(c)的进行之前，能便利地扩增变体 DNA 序列。这样依据本领域所熟知的方法完成扩增，目前优选的方法，即在 DNA 或亲本酶的氨基酸序列的基础上，利用寡核苷酸引物而进行的 PCR-产生扩增。

在用诱变剂温育或者暴露其中后，通过在能使表达进行的条件下培养携带 DNA 序列的合适的宿主细胞而表达突变 DNA。用于这一目的宿主细胞可以用突变 DNA 序列转染的细胞，所说的突变序列可以存在于载体中或为在诱变处理期间携带编码亲本酶的 DNA 序列的序列。下列是合适的宿主细胞的例子：革兰氏阳性细菌，如枯草杆菌，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，灿烂芽孢杆菌或者苏云金芽孢杆菌，浅青紫链霉菌或鼠灰链霉菌；和革兰氏阳性细菌细菌如大肠杆菌。

变体 DNA 序列进一步包括编码功能(其允许突变 DNA 序列表达)的 DNA 序列。

定位随机诱变：随机诱变可有利地定位到所谈论的亲本 $\alpha$ -淀粉酶的某部分。当测定了对于酶所给定性质具有格外重要的酶的一定区域时，并且修饰能引起具有改进性质的变体时，这可能是有利的。当阐明了亲本酶的三级结构，通常能测定这些区，并将其与酶的功能相联系。

通过利用上述的 PCR-产生诱变技术，或本领域任何所熟知的其它合适的技术，可便利地完成定位随机诱变。

此外，通过例如插入到合适的载体中而分离出编码被修饰的 DNA 序列的一部分的 DNA 序列，利用上述的任何诱变方法可使所说的部分序列发生诱变。

关于本发明上述方法中的筛选步骤，利用与本文实施例 2 所描述的相关的分析可便利地完成这一点。一般地关于筛选，基于下面的滤膜测定通常是适用的：

将能表达有义变体的淀粉酶的微生物在合适的培养基上及适合于产生酶的条件下培养，培养基配有双层滤膜，第一层为蛋白质-结合滤膜和

顶端具有较低的蛋白质结合力的第二层滤膜。微生物位于第二层滤膜。温育后,将含有微生物分泌的酶的第一层滤膜与有微生物的第二层滤膜相分离。对第一层滤膜进行关于所需酶促活性的筛选,对应位于第二层滤膜上的微生物克隆则被测定。

用于结合酶促活性的滤膜可以是任何蛋白质结合滤膜例如尼龙或硝酸纤维素。携带表达有机体克隆的顶层-滤膜可以是任何没有或低结合蛋白亲和性的滤膜例如纤维素醋酸盐或 Durapore™。可在任何用于筛选的条件下预处理滤膜或者在酶促活性检测期间进行处理。

用染色,荧光,沉淀, pH 指示剂, IR-吸附剂或酶促活性检测中的任何其它已知技术来检测酶促活性。

用任何结合剂例如琼脂糖,琼脂,明胶,聚丙烯酰胺,淀粉,滤纸,布;或这些结合剂的任何组合将所检测的化合物结合。

由 Cibacron 红标记的支链淀粉检测 $\alpha$ -淀粉酶活性,其结合在琼脂糖上。对于筛选提高的热稳定性和高-pH 稳定性的变体,将结合了 $\alpha$ -淀粉酶变体的滤膜在 pH10.5 的缓冲液中 60 °C 或 65 °C 下温育至规定的时间,在去离子水中简单冲洗,并置于支链淀粉-琼脂糖基质上以检测活性。通过支链淀粉降解当红色 Cibacron 裂解时可观察到剩余活性。选择这些条件是为了 $\alpha$ -淀粉酶(其具有 SEQ ID NO. 2 中所示的氨基酸序列)的活性可以毫无遮蔽地检测到。在相同的条件下,因为红色 Cibacron 增强的释放,稳定化的变体具有增强的颜色强度。

对于筛选在更低温度下和/或更宽的温度范围内的具有最适活性的变体,将结合变体的滤膜直接置于支链淀粉-Cibacron 红底物平板上,并在所需温度(例如 4 °C, 10 °C 或 30 °C)下温育至规定的时间。在这之后, $\alpha$ -淀粉酶(其具有 SEQ ID NO. 2 中所示的氨基酸序列)的活性可以毫无遮蔽地检测到,其中在更低温度下具有最适活性的变体显示出增强的支链淀粉裂解。在支链淀粉基质上温育之前,在所有所需的培养基中(例如含  $\text{Ca}^{2+}$  的溶液,去污剂, EDTA 或其它有关添加剂)进行温育,以便用添加剂筛选出改变的依赖性 or 所谈论的变体的反应。



## 本发明的变体的检验

通过检测变体的降解-淀粉活性可适宜地对本发明的变体进行检验,例如通过在含淀粉的琼脂糖平板上培养被编码变体的 DNA 序列转染的宿主细胞并测定降解-淀粉的宿主细胞。按照本领域所熟知的方法可进一步检验改变的性质(包括比活性,底物特异性,切割模式,热活性,最适 pH, pH 依赖性,最适温度最适,及任何其它参数)。

## $\alpha$ -淀粉酶变体的表达

按照本发明,经过利用包括典型控制序列(其编码启动子,操纵基因,核糖体结合位点,翻译起始信号,和选择性的,阻抑物基因或各种激活物基因)的表达载体可将 DNA 序列(其编码用上述方法或任何本领域所熟知其它方法所产生变体)以酶形式进行表达。

携带编码本发明 $\alpha$ -淀粉酶变体的 DNA 序列的重组表达载体,可以是任何便于进行 DNA 重组过程的载体,对载体的选择经常取决于它将要导入的宿主细胞。这样,载体是自发复制的载体,即具有染色体外质体的载体,其复制不依赖于染色体复制,例如质粒,细菌噬菌体或染色体外因子,微型染色体或人工染色体。另一可取的是,当导入进宿主细胞内时,载体可以是整合到宿主细胞基因组上,并与其整合的染色体一起复制。

在载体中,应该将 DNA 序列与合适的启动子序列相连接。启动子可以是任何在选择的宿主细胞中具有转录活性的 DNA 序列,且其可能来源于编码与宿主细胞同源的或异源的蛋白质的基因。用于引起编码本发明 $\alpha$ -淀粉酶变体的 DNA 序列转录的合适启动子的例子,尤其在细菌宿主中,是大肠杆菌乳糖操纵子的启动子,天蓝色链霉菌琼脂糖酶基因 *dagA* 启动子,杆菌地衣淀粉菌 $\alpha$ -淀粉酶基因(*amyL*)的启动子,嗜热脂肪芽孢杆菌麦芽淀粉酶基因(*amyM*)的启动子,杆菌解凝淀粉菌 $\alpha$ -淀粉酶(*amyO*)的启动子,枯草芽孢杆菌 *xylA* 和 *xylB* 基因的启动子等等。对于在真菌宿主中的转录,有用的启动子例子是那些来源于编码下列物质的基因,米曲霉 TAKA 淀粉酶, *Rhizomucor miehei* 天冬氨酸蛋白酶,黑曲霉中性 $\alpha$ -淀粉酶,黑曲霉酸稳定 $\alpha$ -淀粉酶,黑曲霉葡糖淀粉酶, *Rhizomucor miehei* 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可包括合适的转录终止子及，在真核生物中，多腺苷酸序列，其与编码本发明 $\alpha$ -淀粉酶变体的 DNA 序列有效地连接。终止序列和聚腺苷酸化序列宜来源于与启动子相同的来源。

载体可进一步包含使载体在正谈论的宿主细胞内复制的 DNA 序列。这样序列的例子是质粒 pUC19，pACYC177，pUB110，pE194，pAMB1 和 pIJ702 的复制起点。

载体也可以包含选择性标记，例如其产物弥补了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，或者产生抗性如氨基青霉素，卡那霉素，氯霉素或者四环素的抗性。此外，载体可以包含黑曲霉选择标记如 *amdS*，*argB*，*niaD* 和 *sC*，引起潮霉素抗性的标记，或通过共转变完成的选择标记，等如 WO 91/17243 中的所描述的。

尽管在某些方面在细胞内可便利地进行表达，例如当用某细菌作为宿主细胞时，一般地还是优选细胞外表达。一般来说，这里所提到的芽孢杆菌  $\alpha$ -淀粉酶，其具有允许所表达蛋白酶分泌进培养基的前区（*preregion*）。如果需要，这一前区可被不同的前区或信号序列所取代，这可通过取代编码各自前区的 DNA 序列而便利地完成。

用于分别连接本发明编码 $\alpha$ -淀粉酶变体的 DNA 构建体，启动子，终止子以及其它元件，并将它们插入到包含复制必要信息的合适的载体中的方法，对本领域的技术人员是很熟悉的(cf., 例如, Sambrook 等(1989))。

本发明的细胞(其包含如上所述的本发明的 DNA 构建体或包含本发明的表达载体)在本发明 $\alpha$ -淀粉酶变体的重组体制备中，利于用作宿主细胞。通过将 DNA 构建体(在一种或多种拷贝上)便利地整合到宿主染色体中，则可使本发明编码变体的 DNA 构建体转染细胞。当 DNA 序列在细胞中稳定保持可能性更高是，一般认为这一整合有效。按照常规方法，如通过同源的或异源重组可使 DNA 构建体整合到宿主染色体中。此外，如上所述，可用与不同类型宿主细胞相关的表达载体转染细胞。

本发明的细胞可以是高等有机体细胞，诸如哺乳类或昆虫的细胞，但优选地是微生物细胞，例如细菌或真菌(包括酵母)细胞。

合适细菌的例子是革兰氏阳性菌，如枯草杆菌，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，灿烂芽孢杆菌或者苏云金芽孢杆菌，浅青紫链霉菌或鼠灰链霉菌；和革兰氏阳性细菌细菌如大肠杆菌。可通过熟知其本质的方式如，原生质体转化或用感受态细胞来实现细菌的转化。

酵母有机体可从酿酒酵母属或裂殖酿酒酵母属中顺利地选择到，例如酿酒酵母。丝状真菌优选为曲霉属的某个种，如米曲霉或黑曲霉。通过包括所熟知其本质方式的原生质体形成和原生质体转化及随后的细胞壁再生的方法来转染真菌细胞。EP 023 238 中描述了一种适于转化曲霉属宿主细胞的方法。

目前在另一方面，本发明涉及产生本发明 $\alpha$ -淀粉酶变体的方法，该方法包括在利于变体产生的条件下培养上述宿主细胞及从细胞和/或培养基中回收变体。

用于培养细胞的培养基是任何适于所说的宿主细胞生长和获得本发明 $\alpha$ -淀粉酶变体表达的常规培养基。由供给商或按照所出版的配方(例如美国典型培养物保藏中心的目录中所叙述的)可有效地获得或制备合适的培养基。

宿主细胞所分泌的 $\alpha$ -淀粉酶变体可由所熟知的方法从培养基中便利地回收，该方法包括通过离心或过滤从培养基中分离出细胞，并用盐如硫酸铵盐沉淀蛋白质组分，接着利用层析方法如离子交换层析，亲和性层析或类似的方法。

### 工业应用

本发明 $\alpha$ -淀粉酶变体具有为各种工业应用有价值的性质。尤其是，本发明的酶变体可用作为洗脱剂，厨具清洁剂和固体表面清洁去污剂中的组分。许多变体在由淀粉加工的甜料和乙醇产生中，和/或纺织脱浆工艺中特别有用。US 3,912,590 和 EP 专利出版物第 2,730 和 63,909 中描述了常规淀粉-转化方法(包括淀粉液化作用和/或糖化过程)的条件。

淀粉加工甜料的产生：将淀粉转化为果糖糖浆的"传统的"方法通常由三个连续酶促方法组成，即液化方法，接着为糖化方法和异构化方法。在液

化方法中，在 pH 5.5 和 pH 6.2 之间，温度为 95-160 °C 下，时间约 2 小时  $\alpha$ -淀粉酶(如 Termamyl<sup>TM</sup>)将淀粉降解成糊精。为了确保在这些条件下酶的最适稳定性，加入 1 mM 的钙(40 ppm 的游离钙离子)。

在液化作用处理后，通过加入葡糖淀粉酶(如 AMG<sup>TM</sup>)及脱支酶如异淀粉酶和支链淀粉酶(如蛋白酶<sup>TM</sup>) 将糊精转化成葡萄糖。在这一步骤前将 pH 值降至低于 4.5，保持高温(95 °C 以上)，并使液化  $\alpha$ -淀粉酶活性变性。温度降至 60 °C，并加入葡糖淀粉酶和脱支酶。糖化方法进行 24-72 个小时。

在糖化处理后，将 pH 升至 6-8 范围内的值，优选为 pH 7.5，并用离子交换将钙离子移走。然后将葡萄糖糖浆转化成高果糖糖浆利用，例如，结合的葡糖异构酶(如甜化酶<sup>TM</sup>)

本方法至少 3 个酶促促进效应可引起重视。三个促进效应均能被分别观察到其效应，但可使用它们的任何组合(例如 1+2，1+3，2+3 或 1+2+3):

#### 促进效应 1 液化 $\alpha$ -淀粉酶钙依赖性的减弱。

加入游离钙离子是为了充分确保  $\alpha$ -淀粉酶的高稳定性，但是游离钙严重地抑制了葡糖异构酶的活性并且需要移走，借助于一种昂贵的装置操作，在一定程度将游离钙离子水平降至 3-5 ppm 以下。若能避免这种操作则能节约成本，并且无须加游离钙离子即可完成液化操作。

为了做到那点，则需要更低的钙依赖性的类-Termamyl  $\alpha$ -淀粉酶(其在低浓度游离钙(小于 40 ppm)下保持稳定并具有高活性。这样的类-Termamyl  $\alpha$ -淀粉酶应该具有在 pH 4.5-6.5 范围内的最适 pH，优选为 pH 范围为 4.5-5.5。

#### 促进效应 2 不需要的 Maillard 产品形成的减弱

在液化过程期间，不需要的 Maillard 产品的形成程度依赖于 pH。低 pH 有助于降低 Maillard 产品的形成。这样则要求降低 pH 从约 pH 6.0 至约 pH 4.5；遗憾地是，如常规所知道的，热稳定性的类-Termamyl  $\alpha$ -淀粉酶在低 pH(即 pH 小于 6.0)不是十分稳定且比活性一般地较低。

达到上述目的就需要一种类-Termamyl  $\alpha$ -淀粉酶，它在低 pH 值

4.5-5.5 范围内是稳定的并且在游离钙浓度为 0-40 ppm 范围内保持高比活性。

### 促进效应 3

据以前报导(美国 5, 234, 823 专利), 当用黑曲霉葡糖淀粉酶和 *B. acidopullulyticus* 支链淀粉酶进行糖化时, 若在糖化阶段前没有灭活 $\alpha$ -淀粉酶, 液化作用剩余 $\alpha$ -淀粉酶活性的存在可导致葡萄糖的低产量。代表性灭活方法是在降低糖化温度至 60 °C 前, 在 95 °C 调整 pH 到 4.3 以下。

对葡萄糖产率的负面效应的原因尚不充分了解, 但推测液化的 $\alpha$ -淀粉酶(例如地衣芽孢杆菌的 Termamyl™ 120 L) 通过水解靠近支链淀粉分支点且在其两侧的 1,4- $\alpha$ -糖苷联接而产生"有限的糊精"(其对 *B. acidopullulyticus* 支链淀粉酶来说是不适底物)。葡糖淀粉酶对这些有限的糊精的水解可导致三糖葡糖基麦芽糖的构建, 其由葡糖淀粉酶的水解只能很慢地进行。

不受这一不利因素影响的热稳定 $\alpha$ -淀粉酶的制备将是一种非常重要的促进方法, 则不要求任何单独的灭活步骤。

如果制备了一种类 Termamyl 低-pH-稳定性的 $\alpha$ -淀粉酶, 特异性的改变会在与低 pH 时增加的稳定性的组合中是有利的。

本发明的方法与原则使得可以按照本发明设计和产生具有如上面列出的所需性质的变体。在这一点上, 特别有意义的变体是在类-Termamyl  $\alpha$ -淀粉酶中的变体[例如 Termamyl™ 本身(地衣芽孢杆菌 $\alpha$ -淀粉酶; SEQ ID NO. 2); 或具有某 N-端氨基酸序列(即至对应于 Termamyl™ 的 35 位置处的氨基酸位置的部分序列)的类-Termamyl $\alpha$ -淀粉酶, 且该序列等同于位于解淀粉芽孢杆菌 $\alpha$ -淀粉酶(SEQ ID NO. 4)的该序列, 即类-Termamyl  $\alpha$ -淀粉酶, 其具有与 Termamyl™ 氨基酸序列相关的下列 N 端序列:

A1\*+N2\*+L3V+M15T+R23K+S29A+A30E+Y31H+A33+H35I, 其中星号(\*)表明所论及氨基酸残基的缺失, 其所在位置对应于 Termamyl™ 中的任何下列位置:

H133 H156 A181 A209 G310 H450 V128 N104 V54 S187

**H293 A294** (在每一种末端氨基酸残基位置处, 可被任何其它氨基酸残基所取代, 即 A, R, N, D, C, Q, E, G, H, I, L, K, M, F, P, S, T, W, Y 和 V 中选择的任何其它氨基酸), 以及下列的三重缺失:

**K370\*+G371\*+D372\***

**D372\*+S373\*+Q374\***

在上面所指明位置更为优选的取代是下列:

**H133I H156Y A181T A209V G310D H450Y V128E N104D  
V54W, Y, F, I, L S187D H293Y A294V.**

在本文中, 用正谈论的类-Termamyl  $\alpha$ -淀粉酶可适当地实现任何一种或多种(即一种, 两个, 三个, 四个.....等)上述所表明突变的组合, 以及在淀粉液化作用行为方面能引起一种或多种上述促进效应的本发明的特别有意义的变体, 包含具有对应于下列 Termamyl<sup>TM</sup>(SEQ ID NO2)自身中的多突变组合的若干突变组合的变体:

H133I+H156Y+A181T+A209V+G310D+H450Y+V128E+N104D+V54W+S187D+H293Y+A294V+K370\*+G371\*+D372\*;

H133I+H156Y+A181T+A209V+G310D+H450Y+V128E+N104D+V54W+S187D+H293Y+A294V+D372\*+S373\*+Q374\*;

H133I+H156Y+A181T+A209V+G310D+H450Y+V128E+N104D+V54Y+S187D+H293Y+A294V+K370\*+G371\*+D372\*;

H133I+H156Y+A181T+A209V+G310D+H450Y+V128E+N104D+V54Y+S187D+H293Y+A294V+D372\*+S373\*+Q374\*;

H133I+H156Y+A181T+A209V+G310D+H450Y+V128E+N104D+V54F+S187D+H293Y+A294V+K370\*+G371\*+D372\*;

H133I+H156Y+A181T+A209V+G310D+H450Y+V128E+N104D+V54F+S187D+H293Y+A294V+D372\*+S373\*+Q374\*;

H133I+H156Y+A181T+A209V+G310D+H450Y+V128E+N104D+V54I+S187D+H293Y+A294V+K370\*+G371\*+D372\*;

H133I+H156Y+A181T+A209V+G310D+H450Y+V128E+N104D+V54I+S187D+H293Y+A294V+D372\*+S373\*+Q374\*;

H133I+H156Y+A181T+A209V+G310D+H450Y+V128E+N104D+V54L+S187D+H293Y+A294V+K370\*+G371\*+D372\*;

H133I+H156Y+A181T+A209V+G310D+H450Y+V128E+N104D+V54L+S187D+H293Y+A294V+D372\*+S373\*+Q374\*;

在本文中本发明的另一种有意义的变体包括具有对应于下列 Termamyl<sup>TM</sup> 自身中的单突变或多突变的单或多突变的变体:

突变(氨基酸取代), 其在位置 N172(例如 N172R, K), S187(例如 S187D), N188(例如 N188P), N190(例如 N190L, F), H205(例如 H205C), D207(例如 D207Y), A210(例如 A210S), Q264(例如 Q264S) 或 N265(例如 N265Y);

下列多突变:

H156Y+A181T+A209V;

H156Y+A181T+N190F+A209V+Q264S

A1\*+N2\*+L3V+M15T+R23K+S29A+A30E+Y31H+A33S+E34D+H35I+H156Y+A181T+A209V;

A1\*+N2\*+L3V+M15T+R23K+S29A+A30E+Y31H+A33S+E34D+H35I+H156Y+A181T+N190F+A209V; 或

A1\*+N2\*+L3V+M15T+R23K+S29A+A30E+Y31H+A33S+E34D+H35I+H156Y+A181T+N190F+A209V+Q264S

以及任何两个或多个的单突变或多突变的末端组合。

如已表明的, 按照本发明的许多变体特别适宜于应用于淀粉转化, 例如淀粉液化作用。在这一点上, 本发明的另一方面涉及包含以下混合物的组合物:

(i) 地衣芽孢杆菌的 $\alpha$ -淀粉酶(它具有 SEQ ID NO. 2 中所示的序列)与

按照本发明的一种或多种变体(突变体 $\alpha$ -淀粉酶)的混合物,所说的变体来源于(作为亲本类-Termamyl  $\alpha$ -淀粉酶)具有 SEQ ID NO. 6 中所示序列的嗜热脂肪芽孢杆菌 $\alpha$ -淀粉酶; 或者

(ii) 嗜热脂肪芽孢杆菌 $\alpha$ -淀粉酶(它具有 SEQ ID NO. 6 中所示的序列)与按照本发明的一种或多种变体(突变体 $\alpha$ -淀粉酶)的混合物,所说的变体来源于一种或多种其它亲本类-Termamyl  $\alpha$ -淀粉酶(如来源于具有 SEQ ID NO. 2 中所示序列的地衣芽孢杆菌 $\alpha$ -淀粉酶,或来源于本文提及的其它亲本类-Termamyl  $\alpha$ -淀粉酶); 或者

(iii) 按照本发明的一种或多种变体(突变体 $\alpha$ -淀粉酶)[其来源于(作为亲本类-Termamyl  $\alpha$ -淀粉酶)具有 SEQ ID NO. 6 中所示序列的嗜热脂肪芽孢杆菌 $\alpha$ -淀粉酶]与按照本发明的一种或多种变体(突变体 $\alpha$ -淀粉酶)[其来源于一种或多种其他亲本类-Termamyl  $\alpha$ -淀粉酶,如来源于具有 SEQ ID NO. 2 中所示序列的地衣芽孢杆菌 $\alpha$ -淀粉酶,或来源于本文提及的其它亲本类-Termamyl  $\alpha$ -淀粉酶]的混合物。

掺入到这样混合物内的嗜热脂肪芽孢杆菌 $\alpha$ -淀粉酶变体的优选变体包括在 N193 和/或 E210 的取代,和/或 R179\*+G180\*+或 I181\*+G182\*的双缺失(利用这一特定 $\alpha$ -淀粉酶的氨基酸编号)。

后一种类型的组合物(包括嗜热脂肪芽孢杆菌 $\alpha$ -淀粉酶或其按照本发明的变体)在淀粉液化作用中有很大的应用潜力。所给定混合物中的每种单个 $\alpha$ -淀粉酶组分之间的比率(如以术语“每升液体培养基中的活性淀粉酶蛋白的毫克数”表示)取决于每一种组分的确切种类和性质。

### 去污剂化合物

如前所述,本发明的变体适宜添加到去污剂组合物中。已经出版了有关参考文献,例如,WO 96/23874 和 WO 97/07202 进一步详述了去污剂组合物(如衣物或餐具去污剂)的有关组成,及这些去污剂组合物中制备变体的方法,以及如去污剂组合物的有关类型。

包含本发明的变体的去污剂组合物还另外包含一种或多种其它酶,如脂肪酶,角质酶,蛋白酶,纤维素酶,过氧化物酶或乳糖酶,和/或其它的 $\alpha$ -淀粉酶。

本发明的一个方面涉及一种去污添加剂,其包含本发明所述的 $\alpha$ -淀粉酶变体,其任选的可以是非尘粒形式,稳定化的液体形式或保护酶形式。在本发明的一个具体实施方案中,所述去污添加剂中每克该添加剂含有 0.02-200 毫克的酶蛋白。

本发明的又一个方面涉及一种手工或自动机洗去污剂组合物,该组合物包含本发明所述的 $\alpha$ -淀粉酶变体。任选的,其还包含其它酶,选自蛋



白酶, 脂肪酶, 过氧化物酶, 另一种淀粉分解酶和/或纤维素酶。

本发明的又一个方面涉及一种手工或自动洗衣房洗涤组合物, 该组合物包含本发明所述的 $\alpha$ -淀粉酶变体。任选的, 还包含其它酶, 选自蛋白酶, 脂肪酶, 过氧化物酶, 淀粉分解酶和/或纤维素酶。

用常规的浓度可将本发明的 $\alpha$ -淀粉酶变体掺入到去污剂中。目前通常认为, 在常规剂量水平, 用每升洗涤/餐具洗涤液 0.00001-1 mg $\alpha$ -淀粉酶(纯净物的计算, 活性酶蛋白)的量将本发明的变体掺入到去污剂中。

本发明又一方面还涉及本发明所述的 $\alpha$ -淀粉酶变体在纤维脱浆工艺中的用途。

参照附图进一步描述本发明, 其中:

图.1 显示 DNA 序列及终止密码子 TAA, 其编码 WO 95/26397 中所述的芽孢杆菌菌株 NCIB 12512 $\alpha$ -淀粉酶, 还显示了编码的 $\alpha$ -淀粉酶的氨基酸序列(比较 图 2)

图.2 是本发明中四个亲本类-Termamyl  $\alpha$ -淀粉酶的氨基酸序列的序列对比。最左端的编号代表如下的各种氨基酸序列:

1: WO 95/26397 所描述的芽孢杆菌菌株 NCIB 12512 $\alpha$ -淀粉酶的氨基酸序列;

2: WO 95/26397 所描述的芽孢杆菌菌株 NCIB 12513 $\alpha$ -淀粉酶的氨基酸序列;

3: 这里 SEQ ID NO. 6 所示的嗜热脂肪芽孢杆菌 $\alpha$ -淀粉酶的氨基酸序列;

4: *Bacillus* sp. # 707 $\alpha$ -淀粉酶的氨基酸序列(Tsukamoto 等在生物化学生物物理研究通讯 151(1988), pp. 25-31 中所描述的)。

最右侧图的数字阐明了在本文范围内每个氨基酸序列的连续总编号。注意编号为3的序列(与SEQ ID No. 6中的序列对应), 序列对比在分别对应于编号1, 2和4的序列的No. 1和No. 175的氨基酸的位置形成“缺口”。

图3说明了在实施例2中运用的PCR策略(参看下文)

### 材料和方法

#### pSNK101的构建

该大肠杆菌/芽孢杆菌穿梭载体用来引入在大肠杆菌中没有 $\alpha$ -淀粉酶表达的突变体, 并且经过在此方面的修饰使 $\alpha$ -淀粉酶在芽孢杆菌中活性表达。载体的构建如下: 通过将含有与大肠杆菌同源片段的 1.2 kb 的片段插入 $\alpha$ -淀粉酶基因 5'编码区的 PstI 位点而中断灭活 pX 载体中的 $\alpha$ -淀粉酶基因(在 amyL 之内的具有下列改变的 pDN1528: BAN(1-33), H156Y, A181T, N190F, A209V, Q264S, 质粒 pDN1528 在实施例 1 中进一步描述)。运用引物

从pUC19(基因库登记号: #X02514)扩增该1.2kb片段, 正向引物 5-gacctgcagtcaggcaacta-3'和反向引物 5'-tagagtcgacctgcaggcat-3'。用PstI在37℃ 酶切PCR 扩增子和具有 $\alpha$ -淀粉酶基因的pX质粒约2小时。在室温下连接pX载体片段和大肠杆菌来源的扩增子被连接约1小时, 并通过电转化法转化到大肠杆菌中。所得到的载体指定为pSnK101。

#### $\alpha$ -淀粉酶变体的发酵和纯化

将含有相关表达质粒的A.B.枯草芽孢杆菌菌株划线接种到具有得自-80℃保存的15 $\mu$ g/ml 氯霉素LB琼脂糖平板上, 并于37℃培养过夜。将克隆转到500ml摇瓶中, 该摇瓶含有100ml 补充了150 $\mu$ g/ml 氯霉素的BPX培养基。

#### BPX培养基的组分:

土豆粉	100 g/l
大麦面粉	50 g/l
BAN 5000 SKB	0.1 g/l
酪蛋白酸钠	10 g/l
大豆粉	20g/l
Na <sub>2</sub> HPO <sub>4</sub> ·12H <sub>2</sub> O	9g/l
Pluronic <sup>TM</sup>	0.1 g/l

培养物于37℃以270rpm振荡培养5天。

从发酵肉汤中移出细胞和细胞碎片并以4500rpm离心20-25分钟。然后, 过滤上清液以完全获得澄清溶液。浓缩滤过物并在UF-过滤器上洗脱(1000切口膜), 将缓冲液改为pH5.5的20mM 乙酸盐。UF-过滤器运用于SS琼脂F.F.上, 同时按步骤洗脱(洗脱在含有0.2M NaCl的相同缓冲液中进行)。在10mM的Tris(pH9.0)中透析洗脱物并应用在Q-琼脂糖F.F.上, 用6倍柱体积的含有从0-0.3M NaCl线型梯度进行洗脱。集中含有活性的组分(由Phadebas 测定测量), 调整pH到7.5, 同时去除剩余的颜色(通过用0.5%W/vol.的活性炭作用5分钟)。

#### 测定 $\alpha$ -淀粉酶活性

$\alpha$ -淀粉酶活性由使用Phadebas®片作为底物的方法决定。Phadebas片

(Phadebas® 淀粉酶检验, 由Pharmacia Diagnostic提供)含有交联的不溶性兰色淀粉多聚体, 其混合了牛血清白蛋白和缓冲液物质并压成片。

对于所有单一测量, 需将药片悬浮在含有5ml 50mM Britton-Robinson缓冲液的试管中(该缓冲液含50mM 乙酸, 50mM 磷酸, 50mM 硼酸, 0.1mM CaCl<sub>2</sub>, 并用NaOH调整pH到所要求的值)。检测在所需温度下的水浴中进行。将需检验的 $\alpha$ -淀粉酶溶解在xml 50mM的Britton-Robinson缓冲液中。再将该 $\alpha$ -淀粉酶溶液的1ml加入到5ml 50mM Britton-Robinson缓冲液中。用该 $\alpha$ -淀粉酶水解淀粉而得到可溶性蓝色片段。所得的蓝色溶液的吸收值是 $\alpha$ -淀粉酶活性的函数(用分光光度仪在620nm处测量)。

温育(检验时间)10或15分钟后测量的620nm的吸收值是重要的, 其在620 nm的范围是0.2 至2.0个吸收单位。在该吸收范围中, 有活性和吸收(Lambert-Beer律)之间的线性关系。因此, 必须调整酶的稀度以便适合于这一标准。在特定系列条件(温度, pH, 反应时间, 缓冲条件)下, 1mg所给定的 $\alpha$ -淀粉酶将水解一定量的底物并显出蓝色。在620 nm测量颜色强度。所测量的吸收对在所给定的系列条件下 $\alpha$ -淀粉酶的比活性( $\alpha$ -淀粉酶蛋白质的活性/毫克)直接成适当比例。

运用DOPE程序进行随机诱变的一般方法

用下列步骤可进行随机诱变:

- 1.在亲本酶中选择与修饰有关的区。
- 2.在所选择的区中测定突变位点和非突变位点。
- 3.确定所施行的突变种类。例如考虑所要构建的变体, 其所要求的稳定性和/或效能
- 4.选择在结构上合理的突变。
- 5.按照步骤4经步骤3调整选择的残基。
- 6.运用合适的dope方法分析核苷酸分布。
- 7.如果必要, 调整所要的残基以实现基因编码(例如考虑得自基因编码的约束因素(例如力求避免引入终止子))(技术人员必须清楚一些密码子组合不能在实际中运用而需要作相应调整)。

8. 制做引物。

9. 运用引物进行随机诱变。

10. 通过所要改进的性质筛选所要的 $\alpha$ -淀粉酶变体。

在步骤6中使用的合适的dope方法(algorithm)在本领域是众所周知的。Tomandl D.等(计算机辅助分子设计期刊, 11(1997), pp. 29-38)描述了其中一种方法; 下述将描述另一种方法, DOPE:

dope 程序

"DOPE"程序是一种计算机算法, 它用来优化核苷酸组合物的三联体密码(优化的方法使其编码氨基酸的分布能集中所要的最大多数氨基酸的分布)。为了评估那种可能的分布是最类似于所要的氨基酸分布, 我们需要一种记分函数。在"Dope"程序中下列函数是合适的:

$$S \equiv \prod_{i=1}^N \left( \frac{x_i^{y_i} (1-x_i)^{1-y_i}}{y_i^{y_i} (1-y_i)^{1-y_i}} \right)^{w_i}$$

其中,  $x_i$ '表示由程序计算的所得的氨基酸和氨基酸组的量,  $y_i$ '表示由程序用户定义的所需的氨基酸和氨基酸组的量(例如规定20个氨基酸或终止子需要引入, 例如需要一定的百分比(例如90 %Ala, 3%Ile, 7%Val)), 同时 $w_i$ '表示设计好的由程序用户定义的重量因子(例如取决于插入特定氨基酸残基到所论及的位置的重要性)。N为21加上程序用户定义的氨基酸组的数量。为了某一目的将 $0^0$  定义为1。

Monte-Carlo公式(Valleau, J.P. & Whittington, S.G.(1977) Mont Carlo统计机械学指南: 1 Highways。在"统计机械学, 部分A"中的平衡技术即 B.J. Berne, 纽约: 会议)可用于得到该函数的最大值。在每次重复中, 均需完成下列步骤:

1. 选择任何碱基的新的随机核苷酸组合物, 其中当前组合物和新组合物之间的绝对差别小于或等于d, 对密码子的所有三个位置中的四种核苷酸G, A, T, C的任意一种(参见下面的定义)而言。

2. 通过上述函数的运用比较新组合物和当前组合物的得分。如果新得分大于或等于当前得分, 则替换当前的而保存新组合物。如果新得分较

小, 则保持新组合物的可能性是 $\exp(1000(\text{新得分} - \text{当前得分}))$ 。

循环通常包括1000次如前述的重复, 其中d值从1至0线性递减。在优化过程中完成一百或更多次循环。最终获得最高得分的核苷酸组合物。

### 实施例1

#### 构建依据本发明的Termamyl<sup>TM</sup>变体

从质粒pDN1528 (地衣芽孢杆菌 $\alpha$ -淀粉酶SEQ ID No. 2)在枯草芽孢杆菌中表达Termamyl。该质粒含有编码Termamyl和淀粉的全部基因。该表达直接由它本身的启动子启动。此外, 该质粒含有得自质粒pUB110的复制起点(ori)和得自质粒pC194(其具有氯霉素抗性)的CAT基因。pDN1528在WO 96/23874的图9中显示。

制备含有SEQ ID No. 1编码区主要部分的特异性诱变载体。该载体的重要特性(指示pJeEN1)包括从pUC质粒衍生的复制起点, 具有氯霉素抗性的CAT基因, 以及bla基因的包含读框移位的形式, 其野生型通常具有氨苄青霉素抗性(amp<sup>R</sup>表型)。该变体形式形成amp<sup>R</sup>表型。质粒pJeEN1在WO 96/23874的图10中显示, 并且大肠杆菌的复制起点, ori, bla, CAT, Termamyl淀粉酶基因的5'截短形式, 所选限制酶切位点在质粒上表明。

将突变引入amyL中, 所用的是Deng和Nickoloff描述的方法(1992, 生物化学年报 200, pp. 81-88)。除了选择掺入的具有“选择引物”(引物#6616: 参见下面)的质粒是基于转化的大肠杆菌细胞(具有bla修复基因)的amp<sup>R</sup>表型外, 其替代用限制性内切酶酶切的选择方法(参见Deng和Nickoloff的概述)。化学的或酶学的诱变可通过Stratagene生产的Chameleono诱变试剂盒进行(目录号200509)。

证实变体质粒中突变的DNA序列后, 亚克隆含有所需改变的截短基因到pDN1528中作为PstI-EcoRI片段, 并转化蛋白酶缺失或淀粉酶缺失的枯草芽孢杆菌菌株 SHA273(在W092/11357与W095/10603中描述)以获得变体酶的表达。

通过运用下列诱变引物构建Termamyl 变体V54W(从5'到3', 从左到右写): PG GTC GTA GGC ACC GTA GCC CCA ATC CGC TTG

通过运用下列诱变引物构建Termamyl 变体A52W+V54W(从5'到3', 从左到右写): PG GTC GTA GGC ACC GTA GCC CCA ATC CCA TTG GCT CG

引物#6616(从5'到3', 从左到右写, P 指5'磷酸): P CTG TGA CTG GTG AGT ACT CAA CCA AGT C

通过运用下列诱变引物构建Termamyl 变体V54E(从5'到3', 从左到右写): PGG TCG TAGGCA CCG TAG CCC TCA TCC GCT TG

通过运用下列诱变引物构建Termamyl 变体V54M(从5'到3', 从左到右写): PG G TCG TAG GCA CCG TAG CCC ATA TCC GCT TG

通过运用下列诱变引物构建Termamyl 变体V54I(从5'到3', 从左到右写): PGG TCG TAG GCA CCG TAG CCA ATA TCC GCT TG

通过运用下列诱变引物构建Termamyl 变体Y290E与Y290K(从5'到3', 从左到右写): PGC AGC ATG GAA CTG CTY ATG AAG AGG CAC GTC AAA C.

Y表示C和T的等量混合物。编码谷氨酸或赖氨酸的位于290位的密码子的存在已通过DNA测序检测到。

通过运用下列诱变引物构建Termamyl 变体N190F(从5'到3', 从左到右写): PCA TAG TTG CCG AAT TCA TTG GAA ACT TCC C

通过运用下列诱变引物构建Termamyl 变体N188P+N190F(从5'到3', 从左到右写): PCA TAG TTG CCG AAT TCA GGG GAA ACT TCC CAA TC

通过运用下列诱变引物构建Termamyl 变体H140K+H142D(从5'到3', 从左到右写): PCC GCG CCC CGG GAA ATC AAA TTT TGT CCA GGC TTT AAT TAG

通过运用下列诱变引物构建Termamyl 变体H156Y(从5'到3', 从左到右写): PCA AAA TGG TAC CAA TAC CAC TTA AAA TCG CTG

通过运用下列诱变引物构建Termamyl 变体A181T(从5'到3', 从左到右写): PCT TCC CAA TCC CAA GTC TTC CCT TGA AAC

通过运用下列诱变引物构建Termamyl 变体A209V(从5'到3', 从左

到右写): PCTT AAT TTC TGC TAC GAC GTC AGG ATG GTC ATA  
ATC

通过运用下列诱变引物构建Termamyl 变体Q264S(从5'到3', 从左到  
右写): PCG CCC AAG TCA TTC GAC CAG TAC TCA GCTACC GTA  
AAC

通过运用下列诱变引物构建Termamyl 变体S187D(从5'到3', 从左到  
右写): PGC CGT TTT CAT TGT CGA CTT CCC AAT CCC

通过运用下列诱变引物构建Termamyl 变体DELTA(K370-G371-  
D372)(即缺失氨基酸残基370, 371和372)(从5'到3', 从左到右写): PGG  
AAT TTC GCG CTG ACT AGT CCC GTA CAT ATC CCC

通过运用下列诱变引物构建Termamyl 变体DELTA(D372-S373-  
Q374)(从5'到3', 从左到右写): PGG CAG GAA TTT CGC GAC CTT  
TCG TCC CGT ACA TAT C

合成Termamyl 变体A181T与A209V是通过用限制性酶ClaI消化具  
有类pDN1528质粒的A181T(即pDN1528具有amyl变体导致的A181T改变)  
和具有类pDN1528质粒的A209V(即pDN1528具有amyl变体导致的  
A290V改变), 该限制性酶切割类pDN1528质粒两次而得到一种1116bp的  
片段和3850bp的载体部分(即含有复制起点)。在琼脂糖凝胶上分离后, 通  
过QIA快速凝胶抽提试剂盒(从QIAGEN购买)纯化具有A209V突变的片  
段和具有A181T突变的载体部分的片段。连接该片段和载体并转化缺失蛋  
白酶和淀粉酶的枯草芽孢杆菌菌株(参见前述)。分析amy<sup>+</sup>质粒(在琼脂板  
上去除了淀粉的区域)和氯霉素抗性转化体, 分析质粒上两种突变的存  
在。

与上述类似的方式, 用限制性内切酶Acc65I和EcoRI组合H156Y和  
A209V 而得到H156Y+A209V。

用限制性内切酶 Acc65I 和 HindIII 组合 H156Y+A209V 和  
A181T+A209V而得到H156Y+A181T+A209V。

用解淀粉芽孢杆菌 $\alpha$ -淀粉酶(SEQ ID No. 4)(在本文中称为BAN)的33N  
端残基取代Termamyl 变体H156Y+A181T+A209V成熟部分的35N端残

基,其方法是SOE-PCR(Higuchi等,1988,核酸研究16:7351),具体如下:

引物19364(序列由5'到3'):

CCT CAT TCT GCA GCA GCA GCC GTA AAT GGC ACG CTG

引物19362:

CCA GAC GGC AGT AAT ACC GAT ATC CGA TAA ATG TTC CG

引物19363:

CGG ATA TCG GTA TTA CTG CCG TCT GGA TTC

引物1C:

CTC GTC CCAATCGGT TCC GTC

用Boehringer Mannheim生产的Pwo热稳定聚合酶进行标准PCR(即聚合酶链式反应),反应条件按照制造商说明并设定温度循环为:94℃5分钟,25个循环(94℃30秒,50℃45秒,72℃1分钟),72℃10分钟。

在定义为PCR1的第一次PCR中利用引物19364和19362扩增了大约130 bp的片段,所说的片段在含有编码解淀粉芽孢杆菌 $\alpha$ -淀粉酶的DNA片段上。

在定义为PCR2的另一次PCR中利用引物19363和1C在模板pDN1528上扩增了大约400 bp片段。

在琼脂糖凝胶上纯化PCR1和PCR2,并将其作为PCR3的模板,通过引物19364和1C获得大约520 bp的片段。因此,这一片段含有编码BAN之N端DNA的一部分,该部分融合到从第35个氨基酸编码Termamyl的DNA的一部分上。

通过限制性内切酶PstI和SacII进行酶切,然后将前述连接并转化进枯草芽孢杆菌菌株,将520 bp片段亚克隆到类似pDN1528的质粒中(含有编码Termamyl 变体H156Y+A181T+A209V的基因)。通过DNA测序(从amyl+和氯霉素抗性转化体中提取质粒)证实了限制酶切位点PstI和SacII之间的DNA序列。

将具有BAN和H156Y+A181T+A209V的正确N-端的最终构建体指定为BAN(1-35)H156Y+A181T+A209V。

按前述进行诱变,通过将N190F与BAN(1-35)+H156Y+A181T+A209V



组合而得到BAN(1-35)+H156Y+A181T+N190F+A209V, 但其中pJeEN1中的 amyL 序列由编码 Termamyl 变体 BAN(1-35)+H156Y+A181T+A209V的DNA序列替换。

按前述进行诱变, 通过将Q264S与BAN(1-35)+H156Y+A181T+A209V组合而得到BAN(1-35)+H156Y+A181T+A209V+Q264S, 但其中pJeEN1中的 amyL 序列由编码 Termamyl 变体 BAN(1-35)+H156Y+A181T+A209V的DNA序列替换。利用限制性内切酶 BsaHI(引入的BsaHI 酶切位点靠近A209V变体)和PstI组合H156Y+A181T+A209V+Q264S和BAN(1-35)+H156Y+A181T+N190F+A209V而得到BAN(1-35)+H156Y+A181T+N190F+A209V+Q264S。

### 实施例2

通过随机定位, doped诱变构建类Termamyl  $\alpha$ -淀粉酶变体, 与亲本酶比较, 其在低pH下具有改进的稳定性, 稳定性对钙离子的依赖性降低

$\alpha$ -淀粉酶对工业淀粉液化过程来说十分重要。热稳定的地衣芽孢杆菌之 $\alpha$ -淀粉酶变体含有由解淀粉芽孢杆菌淀粉酶(SEQ ID No. 4)1-33的氨基酸与地衣芽孢杆菌淀粉酶(SEQ ID No. 2)的36-483氨基酸的融合体, 并进一步具有下列突变: Y156, T181, F190, V209 以及S264(这些变体的构建已在实施例1中描述)在低pH和低钙浓度下有十分令人满意的稳定性。要进一步改进该 $\alpha$ -淀粉酶随机诱变变体在低pH和低钙浓度下的稳定性可在预选的区域中完成。

这些区域是:

区:	残基:
I:	Phe153-Thr163
II:	Gln178-Asn192
III:	His205-Arg214
IV:	Ala232-Asp237 和
VIII:	Gly131-Lys136
IX:	Asp164-Tyr175
X:	Tyr 262-Thr278

改变的区域	总量%	平均数%	残基数量
I:	35	88	11中的8个
II	20	86	15中的11个
III	27	88	10中的10个
IV	34	91	12中的11个
VIII	39	86	6中的6个
IX	46	93	12中的12个
X	27	90	13中的12
VIII+IX	18		
VIII+IX+II	4		
II+III+IV	2		
IV+I	12		

总量%之下的数字表示掺入后在所给定的区域中所需的野生型(wt)氨基酸总量。由突变位点数量(例如区域I为8)乘以它们各自的野生型获得该数字。对于区域I所需的总量%是 $80 \times 80 \times 90 \times 90 \times 90 \times 90 \times 95 \times 90 / 100 = 35\%$ 。

平均数%表示本范围内区域位点总量的平均掺入水平(例如有关区域I的11个位点)。对于区域I, 平均数%计算如下:  
 $80 + 80 + 90 + 90 + 90 + 90 + 95 + 90 = 705$ , 除以11 = 88%

DOPE软件(参见材料和方法)可用来测定掺入密码以引起在七个区域中每一种建议的变化而最大限度地减少终止子的量。核苷酸的确切分布在密码的三个位点中计算而得到建议的氨基酸群体的变化。在指定位置中向掺入区域特定位点掺入以高几率地获得所需残基, 但是, 任然允许其他可能性的存在。

例如, 在野生型序列中起始H156突变为Y(代表一种新的密码), 接着添加10%的其他残基。只有该DNA序列具有Y的密码子而不是H的密码子。在位置156处90%的设计是Tyr同时其他残基也允许自由地占领此位点。在有些位点不可能按照预先的建议改变氨基酸残基群体, 因为基因密码在结构和功能上限制了所需的残基。表I-7中列出了所得的七个掺入寡

核苷酸，野生型核苷酸和氨基酸以及每个掺加位点的核苷酸分布。所有引物文库作为有义链合成。

表1: 文库DASI(Phe153-Thr163)

153 154 155 156 157 158 159 160 161 162 163

Phe Lys Trp Tyr Trp Tyr His Phe Asp Gly Thr

引物: 5'-CGC GGC AGC ACA TAC AGC GAT T1T 2A3 TGG 45T  
TGG 67T 8AT TTT GAC GGA A9C GAT TGG GAC GAG TCC CGA  
AAG3

每一掺加位点的核苷酸分布:

- 1: 80 %T, 20%A
- 2: 96 %A, 2%G, 2%C
- 3: 98 %A, 2%T.
- 4: 93 %T, 4%G, 3%A.
- 5: 97 %A, 3%G.
- 6: 98 %T, 2%A.
- 7: 97 %A, 3%C.
- 8: 90 %C, 10% T.
- 9: 95 %C, 5%A.

表2: 文库DASII(Gln178-Asn192)

178 179 180 181 182 183 184 185 186 187 188 189 190 191 192

Gln Gly Lys Thr Trp Asp Trp Glu Val Ser Asn Glu Phe Gly Asn

引物: 5'CTG AAC CGC ATC TAT AAG TTT 1A2 34T AAG 567 TGG  
GAT 89G GA10 GTT A11T 1213T GAA T1415 161718 AAC TAT GAT  
TAT TTG ATG TAT3'

每一掺加位点的核苷酸分布:

- 1: 93 %C, 7%A.
- 2: 84 %G, 16%T.
- 3: 95 %G, 5%A.
- 4: 95 %G, 5%C.

- 5: 94 %A, 6%G.  
 6: 95 %C, 5%A.  
 7: 62 %T, 38 %G.  
 8: 87 %T, 13%A.  
 9: 91 %G, 9%C.  
 10: 92 %G, 8%T.  
 11: 90 %G, 5%A, 5%C.  
 12: 88 %A, 12%C.  
 13: 88 %A, 12%C.  
 14: 93 %T, 5%A, 2%C.  
 15: 97 %T, 3%G.  
 16: 86 %G, 14%A.  
 17: 89 %G, 11%C.  
 18: 60 %G, 40 %T.

表3: 文库DASIII(His205-Arg214)

205 206 207 208 209 210 211 212 213 214

His Pro Asp Val Val Ala Glu Ile Lys Arg

引物: 5'TAT GCC GAC ATC GAT TAT GAC 12T 3CT 456 7TT 8910  
 1112T 13A14 15T16 A17A 1819A TGG GGC ACT 3TGG TAT GCC  
 AAT 3'

每一掺加位点的核苷酸分布:

- 1: 89 %C, 11%T.  
 2: 89 %A, 11%G.  
 3: 95 %C, 2.5 %T, 2.5 %A.  
 4: 96 %G, 1%A, 3%T.  
 5: 96 %A, 4%C.  
 6: 98% T, 2%A.  
 7: 95 %G, 2.5 %A, 2.5 %C.  
 8: 93 %G, 7%A.

- 9: 96 %T, 4%A.  
 10: 84 %A, 16%G.  
 11: 81 %G, 7%A, 7%T, 5%C.  
 12: 98 %C, 2%A.  
 13: 96 %G, 4%C.  
 14: 94 %G, 6%T.  
 15: 82 %, A18% T.  
 16: 50 %A, 50 %T.  
 17: 90 %A, 10%G.  
 18: 70 %A, 30%C.  
 19: 86 %G, 14%A

表4: 文库DASIV(Ala232-Asp243)

232 233 234 235 236 237 238 239 240 241 242 243

Ala Val Lys His Ile Lys Phe Ser Phe Leu Arg Asp

引物: 5'TTG GAC GGT TTC CGT CTT GAT 12T G3T AAA 456 7TT  
 A8G T9T 1011T T12T 13T14 1516G GA17 TGG GTT AAT CAT GTC  
 AGG GAA 3'

每一掺加位点的核苷酸分布:

- 1: 93 %G, 3.5 %A, 3.5 %T.  
 2: 94 %C, 4%T.  
 3: 94 %T, 6%C.  
 4: 93 %C, 2%T, 2%A, 3%G.  
 5: 98 %A, 2%T.  
 6: 98 %T, 2%A.  
 7: 95 %A, 5%C.  
 8: 94 %A, 6%G.  
 9: 90 %T, 10%A.  
 10: 89 %T, 11%A.  
 11: 89 %C, 11%A.

12: 95%T, 5%A.

13: 64 %C, 33 %T, 3%A.

14: 93 %A, 7%T.

15: 90 %A, 10%C.

16: 90 %G, 5%A, 5%C

17: 90 %T, 10%A

表5: 文库DASVIII(Gly131-Lys136)

131 132 133 134 135 136

Gly Glu His Leu Ile lys

引物: 5'GCT GAC CGC AAC CGC GTA ATT TCA 123 GA4 56T 78A

9TA A10G GCC TGG ACA CTTTTT CAT TTT 3'

每一掺加位点的核苷酸分布:

1: 91 %G, 9%A.

2: 87 %G, 13%C.

3: 90% T, 10%G.

4: 90 %G, 10%T.

5: 85 %C, 8%T, 7%A.

6: 89 %A, 9%T, 2%C.

7: 88 %T, 12%A

8: 88 %T, 11%C, 1%G

9: 92 %A, 8%T

10: 93 %A, 7%G

表6: 文库DASIX(Asp164-Tyr175)

164 165 166 167 168 169 170 171 172 173 174 175

Asp Trp Asp Glu Ser Arg lys Leu Asn Arg Ile Tyr

引物: 5'9TGG TAC CAT TTT GAC GGA ACC GAT TGG 1A2 GAG

3CG CGA A4G 56A A7T A8G 9 1011 T12T AAG TTT CAA GGA AAG

GCT TGG3'

每一掺加位点的核苷酸分布:

- 1: 94 %G, 6%A.  
 2: 96 %T, 4%G.  
 3: 92 %T, 4%A, 4%G.  
 4: 95 %A, 5%G.  
 5: 93 %C, 7%A.  
 6: 92 %T, 8%A.  
 7: 90 %A, 5%G, 5%C.  
 8: 90 %G, 10%A.  
 9: 92 %A, 6%G, 2%T.  
 10: 92 %T, 8%A.  
 11: 50 %T, 50 %C.  
 12: 96 %A, 4%T.

表7: 文库DASX(Tyr262-Asn278)

262 263 264 265 266 267 268 269 270 271 272 273 374 275 276 277 278  
 Tyr Trp SAsn Asp Leu Gly Ala Leu Glu Asn Tyr Leu Asn Lys Thr Asn  
 引物: 5'GAA ATG TTT ACG GTA GCT GAA T1T TGG 234 56T 7A8  
 91011 1213T 1415'T 16T17 GA18 A19T T20T 21T22 A23C A24G ACA  
 25AT TTT AAT CAT TCA GTG TTT GAC3'

每一掺加位点的核苷酸分布:

- 1: 95 %A, 5%T.  
 2: 97 %A, 3%G.  
 3: 95 %G, 2.5 %A, 2.5 %C.  
 4: 94 %T, 6.2 %G.  
 5: 97 %A, 3%T.  
 6: 94 %A, 3%G, 3%C.  
 7: 95%G, 5%A.  
 8: 95 %T, 5%A.  
 9: 52 %T, 45%C, 3%A.  
 10: 96 %T, 4%C.

- 11: 60 %A, 40 %G.
- 12: 90 %G, 10%A.
- 13: 94 %G, 6%C.
- 14: 81 %G, 8%A, 8%T, 3% C.
- 15: 98 %C, 2%T.
- 16: 90 %C, 10%.
- 17: 50 %G, 50 %T.
- 18: 90 %, 10%T.
- 19: 90 %A, 5%G, 5%C.
- 20: 95 %A, 5%T.
- 21: 91%T, 9%A.
- 22: 92 %A, 8%G.
- 23: 94 %A, 3%G, 3%C.
- 24: 93 %G, 7%A.
- 25: 90 %A, 10%G.

### 随机诱变

表1-7中明显掺入的寡核苷酸(在图3中指定为普通术语FDAS)和每一区域的反向引物RDAS 以及具有SacII和Sall位点的特异性地衣芽孢杆菌引物可用来产生PCR文库片段, 其方法为重叠延伸方法(Horton等, 基因, 77(1988), pp. 61-68), 该方法通过一种21bp的重叠片段进行。图3 指出了PCR 策略。将PCR片段克隆到大肠杆菌/芽孢杆菌穿梭载体pSNK101内(参见材料和方法), 其能使诱变在大肠杆菌中产生并使其能立即在枯草芽孢杆菌中表达而抑制淀粉酶在大肠杆菌中的致死性积累。克隆的 PCR片段在大肠杆菌中形成后。在质粒外消化已修饰过的pUC19片段以及将突变的Termamyl 基因和启动子进行物理连接, 然后, 可在芽孢杆菌中表达。筛选

七个文库需按下述在低pH和低钙滤膜测定中进行筛选。

### 低pH值滤膜测定

将芽孢杆菌文库平板接种在纤维素乙酸盐(O.E 67, Schleicher &



Schuell, Dassel, 德国)和硝酸纤维素滤膜(Protran-Ba 85, Schleicher Schuell, Dassel, 德国)的夹层中, 并放入含有10 $\mu$ g/ml 氯霉素的TY 琼脂平板上于37 $^{\circ}$ C培育至少21小时。其中, 纤维素乙酸盐层位于TY 琼脂平板上。

平板接种后但在温育之前用针特异性地标记每个滤膜夹层, 以便能定位滤膜上的阳性变体, 将结合了变体的硝酸纤维素滤膜转到具有柠檬酸盐缓冲液, pH 4.5的容器中, 并在80 $^{\circ}$ C下温育15分钟, 将带有克隆的纤维素乙酸盐滤膜存储在TY平板上直到使用。经过温育后, 在含有1%琼脂糖, 0.2%淀粉, pH6.0和1%柠檬酸缓冲液的平板上检测剩余活性。用标记夹层硝酸纤维素滤膜测定平板, 然后在50 $^{\circ}$ C温育1小时。在移去滤膜后用10%鲁戈氏溶液将测定平板染色。在深蓝色的背景上的白斑即为淀粉降解变体, 接着, 将其在存储平板上进行鉴定。阳性变体在与第一次筛选相同的条件下筛选两次。

#### 低钙滤膜测定

测定用与低pH 滤膜相同方法进行, 具有下列改变: 结合了蛋白质的滤膜在95 $^{\circ}$ C, 用不同浓度的EDTA(0.001 mM-100 mM), 在pH6.0下温育1小时。

按上述方法获得下列栏中列出的变体(BAN指定为解淀粉芽孢杆菌 $\alpha$ -淀粉酶):

\* BAN/Termamyl杂种\*

+H156Y+A181T+N190F+A209V+Q264S+E211Q.

BAN/Termamyl杂种\*

+H156Y+A181T+N190F+A209V+Q264S+H205C+D207Y+A210S

通过随机诱变方法将栏中列出的突变体引入。有关这些变体的稳定性数据在实施例3中的表11列出。

按上述类似方式, 上述的随机诱变测定的七个区域在亲本地衣芽孢杆菌 $\alpha$ -淀粉酶(SEQ ID No. 2)中完成。掺加方案也以上述类似方式测定。

### 实施例3

钙和pH依赖稳定性的测量

通常，将pH6.0-6.2作为液化作用之pH进行工业液化过程，并加入40ppm的游离钙以在95℃-105℃条件下提高稳定性。一些在此处提出的取代使得稳定性提高在下列条件下成为可能：

1. pH低于6.2和/或
2. 游离钙水平低于40ppm游离钙。

有三种不同的方法可用来测量在Termamyl 中由不同的取代而获得的稳定性的提高：

1. 一种测量是在pH值微量减少(pH5.5)而40ppm的游离钙存在时测定该稳定性(由此，低pH条件下稳定性的提高则可以测量)。在下列条件下温育10 $\mu$ g变体：0.1M 乙酸盐溶液，调节pH到pH5.5，含有40ppm游离钙以及5%w/w的普通玉米粉(不含钙)。将其在水浴中95℃下温育30分钟。

2. 另一种测定则在缺少游离钙而pH值仍保持6.2的情况下测定稳定性。该测定在pH6.2被保持时测量在钙灵敏度上的减少。这一测定测量在钙灵敏度上减少：在下列条件下温育10 $\mu$ g变体：0.1M 乙酸盐溶液，调节pH到pH5.5，还含有5%w/w的普通玉米粉(不含钙)。将其在水浴中95℃下温育30分钟。

3. 第三种测定方法则结合了第一和第二两种方法的条件。在缺少钙和低pH值(pH5.5)的情况下进行稳定性测量。

4. 第四种测定方法类似于第三种方法。只是将pH值进一步降低至pH5.0。

#### 稳定性测定

所有1, 2, 3和4的稳定性实验按相同方式进行。方法是：

在相关条件(1-4)下温育该酶。各种样品分别温育0, 5, 10, 15和30分钟并在测定缓冲液(0.1M 50mM Britton 缓冲液, pH7.3)中稀释25次(所有样品都以相同方式进行稀释)，用 Phadebas 测定法(Pharmacia)在pH7.3, 37℃的标准条件下对活性进行测量。

在温育(0分钟)之前所测量的活性用作对照(100%)。将计算的百分值作为温育时间的函数。表中列出了在温育30分钟之后的剩余活性。

比活性测定

通过 Phadebas 测定法(Pharmacia)以活性/mg 酶测定比活性。通过本文材料和方法中所述的 $\alpha$ -淀粉酶测定法来测定活性。

结果:

表 8: 稳定性方法(第一种)/低 pH 值下的稳定性提高

变体	温育 30 分钟后的剩余活性	比活性,37 °C,pH7.3 . 相对 Termamyl 的百分比
Termamyl 野生型	5%	100%
H156Y	15%	100%
A209V	18%	100%
Q264S	30%	100%
H156Y+A181T+ A209V	39%	100%
H133Y+H156Y+ A181T+A209V+ H293Y+A294V+ G310D+H450Y	58%	85%
BAN/Termamyl 杂种* +H156Y+A181T+ A209V	48%	115%

表 9: 稳定性方法(第二种)/降低的钙灵敏度

变体	温育 30 分钟后的剩余活性	比活性,37 °C,pH7.3 . 相对 Termamyl 的百分比
Termamyl 野生型	52%	---
H156Y	70%	---
A209V	75%	---
A181T	57%	115%
N190F	34%	140%
N190F+N188P	57%	190%

表 10: 稳定性方法(第二种)/低 pH 值下的稳定性提高+降低的钙灵敏度

变体	温育 30 分钟后的剩余活性	比活性,37 °C,pH7.3 . 相对 Termamyl 的百分比
Termamyl 野生型	3%	---
BAN/Termamyl 杂种*+H156Y +A181T+A209V	20%	---
Q264S	5%	---
H140K+H142D	5%	115%
V128E	50%**	115%
BAN/Termamyl 杂种* +H156Y+A181T+ N190F+A209V	39%	170%
BAN/Termamyl 杂种* +H156Y+A181T+ A209V+Q264S	29%	175%
BAN/Termamyl 杂种* +H156Y+A181T+N1 90F+A209V+Q264S	57%	210%

\*BAN(解淀粉芽孢杆菌 $\alpha$ -淀粉酶(SEO ID No. 4)/Termamyl (地衣芽孢杆菌 $\alpha$ -淀粉酶(SEO ID No. 2)PCR 杂种。N 端的前 33 个氨基酸是 BAN 而其余的 36-483 是 Termamyl (变体的构建已经在实施例 1 中描述)。

\*\*温育 5 分钟后测量。与 Termamyl 野生型(其在相同条件下显示 36 % 的剩余活性)。

表 11: 稳定性方法(第四种)/低 pH(pH5.0)稳定性提高+降低的钙灵敏度

变体	温育 30 分钟后的剩余活性	比活性,37 °C,pH7.3. 相对 Termamyl 的百分比
BAN/Termamyl 杂种* +H156Y+A181T+ N190F+A209V+Q264S	9%	210%
BAN/Termamyl 杂种* +H156Y+A181T+N190F +A209V+Q264S +E211Q	28%	160%
BAN/Termamyl 杂种* +H156Y+A181T+ N190F+A209V+Q264S+ H205C+D207Y+ A210S	33%	130%

\*如表 10 相关部分所示

上述表 11 的变体构建需借助于在实施例 2 中所描述的随机诱变。

#### 实施例 4

在低 pH 和高温条件下的 $\alpha$ -淀粉酶稳定性

该实施例总结了变体的稳定性结果,其特性在两个不同的条件下通过荧光测定法于 70 °C 测定,该条件是(1)pH4.5 和 1mM CaCl<sub>2</sub> 和(2)pH6.2 和 10 $\mu$ M CaCl<sub>2</sub>。

方法简述

所有荧光实验在 Perkin-Elmer LS-50 发光分光计上通过 4 小池固定器完成。通过循环水浴控制温度并直接在小池中通过 Noronix 的数字温度计进行测量(模型 NTD 100)。在测量进行期间,通过磁性搅拌器高速搅拌可确保将试剂在小池中混合均匀。小池上加盖特氟隆小塞以最大限度地减

少蒸发。内在的蛋白质荧光(因为 Trp 侧链)在 350 nm 由激发监控而在 280 nm 发射。狭缝宽度是 5nm。

在动力学测量进行期间平行监测 4 个反应。将数据收集到波长程序目录中，其使得可以在延迟时期内自动收集数据。

在 70 °C 进行解折叠。解折叠条件是

(1) 50mM NaOAc pH4.5 和 1mM CaCl<sub>2</sub>

(2) 50mM NaOAc pH6.2 和 10μM CaCl<sub>2</sub>。

蛋白质浓度是 5μg/ml 而甘油为 3.5w/v(自蛋白质贮存液)

注意：因为微量的温度变化使解折叠进行一半的绝对值在每天都产生一些变化(例如水浴中不同水量的情况)。然而，Termamyl 总是作为在每种实验中所分析的四酶之一，实际上这已经是一种内部标准。与该内部标准相关的解折叠率可令人满意地重现(完成三次)。用 GraphPad Prism 软件进行。在 pH4.5 时，解折叠数据可令人满意地能随漂变而呈单指数衰变：

$$F(t) = A * \exp(-\ln(2) * t / t_{1/2}) + \text{漂变} * t + \text{补偿} \quad (1)$$

其中，F 指所测量的荧光，A 指解折叠的振幅 (amplitude)，t 指时间，而 t<sub>1/2</sub> 指解折叠的一半时间。

在 pH6.2 时，解折叠更为复杂(包括一种初始延滞期)，并且数据不能适用于公式 1。实际上，已衰变成为初始信号 50 % 的荧光信号用作表观 t<sub>1/2</sub>。

从这些一半时间，与 Termamyl 相关的解折叠自由能方面的变化可按式计算：

$$DDG = R * T * \ln(t_{1/2}^{\text{突变体}} / t_{1/2}^{\text{Termamyl}}) \quad (2)$$

其中 R 指大气常数而 T 指温度(R\*T 值为 0.5919，所得的 DDG 值单位为 kcal/mol)。

通过把数据转化为 DDG 值，不同变体的非稳定性/稳定性作用可直接进行比较，并检测加和性 (DDG<sub>1+2</sub>=DDG<sub>1</sub>+DDG<sub>2</sub>) 和协同性 (DDG<sub>1+2</sub>>DDG<sub>1</sub>+DDG<sub>2</sub>)，此处 DDG<sub>1+2</sub> 指引入突变体 1 和 2 的能量影响。

## 结果

淀粉酶在低 pH 和高温条件下的解折叠可以接着 Trp 荧光的衰减。在 pH4.5 和 1mM CaCl<sub>2</sub> 时，所有淀粉酶的折叠进行得相当迅速。

解折叠数据在 pH4.5 时更适宜于二指数方程式(与单指数方程式比较而言)。然而，由于第二期十分缓慢，其近似一种线性递减(方程式 1)。70 °C 下 pH6.2 和 10μM CaCl<sub>2</sub> 时的解折叠远不如在 pH4.5 时(尽管低 [Ca<sup>2+</sup>])。在 1 小时内解折叠远没有完成，数据也不能满足单指数方程式。实际上，已衰变为初始信号 50 % 的荧光信号用作表观 t<sub>1/2</sub>。荧光测定的结果已在表 12 中列出。

表 12

Termamyl 变体在 pH4.5 和 pH6.2/(70 °C 下)时的解折叠数据。

pH4.5 时,  $t_{1/2}^{\text{Termamyl}} = 200\text{s}$ ; pH6.2 时,  $t_{1/2}^{\text{Termamyl}} = 2800\text{s}$ 。

$$\text{DDG} = -RT * \ln(t_{1/2}^{\text{Termamyl}} / t_{1/2}^{\text{突变体}})$$

突变	pH 4.5, 1 mM CaCl <sub>2</sub>		pH 6.2, 10 μM CaCl <sub>2</sub>	
	DDG (kcal/mol)	$t_{1/2}/t_{1/2}^{\text{Termamyl}}$	DDG (kcal/mol)	$t_{1/2}/t_{1/2}^{\text{Termamyl}}$
野生型	0	1.0	0	1.0
A209V	-0.36	1.85	-0.72	3.39
H133Y+A209V	-0.77	3.67	-0.61	2.78
H156Y	0.06	0.90	-0.10	1.18
A181T	-0.06	1.10	-0.28	1.60
A181T+A209V	-0.44	2.09	< -1 (73%) <sup>a</sup>	> 5
S187D	0.37	0.54	0.19	0.72
H450Y	-0.49	2.29	0.15	0.78
L270D	-0.35	1.8	-0.10	1.2
A181T+H156Y	-0.17	1.34	-0.62	2.84
H133I	-0.33	1.75	-0.42	2.02
H133Y+H156Y+A181T+ A209V+H293Y+A294V+ G310D+H450Y	-0.96	5.10	< -1 (58%) <sup>a</sup>	>5
V128E	-0.10	1.2	-0.25	1.5
H156Y+A181T+A209V	-0.32	1.71	-0.49	2.30
H156Y+A181T+A209V+ H450Y	-0.42	2.05	-0.63	2.92
H156Y+A181T+A209V+ H450Y+H133I	-0.81	3.9	< -1 (65%) <sup>a</sup>	> 5
H156Y+A181T+A209V+ H133I	-0.70	3.3	< -1 (77%) <sup>a</sup>	> 5
Q264S	-0.26	1.6	-0.14	1.3
H156Y+A181T+A209V+ Delta(1,2)+	-0.43	2.1	-0.82	4.0



L3V+M15T+ R23K+S31A+A32E+Y33H+A3 5S+E36D+H37I	-			
Q264S+N265Y	-0.33	1.8	- 0.07	1.1
Q264S+N265Y+N190F	-1.07	6.1	-0.67	3.1
Q264S+N265Y+N190F+ H133I+A209V	-1.66	16.5	<-1(82%) <sup>a</sup>	> 5
H156Y+A181T+A209V+ Delta(1,2)+ L3V+M15T+ R23K+S31A+A32E+Y33H+ A35S+E36D+H37I+N190F	-0.30	1.7	<-1(66%) <sup>a</sup>	> 5
H156Y+A181T+A209V+Delt a(1,2) +L3V+M15T+ R23K+S31A+A32E+Y33H+ A35S+E36D+H37I+Q264S	-0.43	2.1	-0.86	4.3
H156Y+A181T+A209V+Delt a(1,2)+ L3V+M15T+ R23K+S31A+A32E+Y33H+ A35S+E36D+H37I Q264S+N190F	-0.36	1.8	<-1(76%) <sup>a</sup>	> 5
H156Y+A181T+A209V+N190 F+Q264S	-1.3	8.6	< -1 (66%) <sup>a</sup>	> 5

<sup>a</sup> 百分比表明在 70 °C 3 小时的期限内起始荧光水平已经下降的水平。缓慢降低是高稳定性的指示。

### 实施例 5

#### 具有增加的比活性的 $\alpha$ -淀粉酶变体

这一实施例总结了与野生型 Termamyl 相比具有增加的比活性为特征的变体的结果。这些取代相互组合或者作为单独的取代添加到稳定化取代中而存在时增加所形成的变体的比活性。采用在材料和方法中描述的 $\alpha$ -淀粉酶(Phadebas)测定法测定比活性, 其中测定活性/mg 酶。采用以下

的描述测定所说的活性，其中 pH 为 7.3，温度为 37 °C，试验时间为 15 分钟，缓冲液如所限定的。

突变	比活性(PHADEBAS 检测),相对 Termamyl 野生型的指数
S187D	260%
V54I	160%
BAN/Termamyl 杂种: ( $\Delta(1,2)$ +L3V+M15T+R23K+S31A+A3 2E+Y33H+A35S+E36D+H37I )	140%
$\Delta(D372+S373+Q374)$	125%
$\Delta(K370+G371+D372)$	125%
BAN/Termamyl 杂种: ( $\Delta(1,2)$ +L3V+M15T+R23K+S31A+A3 2E+Y33H+A35S+E36D+H37I ) + $\Delta(D372+S373+Q374)$ + V54I+S187D	360%
Q264S	130%
Y290E	155%
Y290K	140%
N190F	150%

## 实施例 6

### 检验特异性变体(糖化)

以前曾报导过(美国专利 5, 234, 823), 当用葡糖淀粉酶和支链淀粉酶进行糖化时, 剩余 $\alpha$ -淀粉酶活性会引起液化作用。因此, 如果 $\alpha$ -淀粉酶不能在糖化阶段前灭活的话, 则可以导致葡萄糖的低产率。典型灭活是将 pH 调整到低于 4.3, 并在 95 °C 进行处理(在糖化阶段前降低温度至 60 °C)。

对葡萄糖产率的负面影响产生的原因并不完全清楚, 但是我们假定通过水解位于支链淀粉的分支点两端或其附近的 1, 4- $\alpha$ -糖苷键, 液化的 $\alpha$ -淀粉酶(例如得自地衣芽孢杆菌 *termamyl* 120L)可产生“有限的糊精”(其是支链淀粉酶的弱底物)。通过葡糖淀粉酶水解这些有限的糊精可形成三糖潘糖(其只能被葡糖淀粉酶缓慢水解)。

热稳定 $\alpha$ -淀粉酶的开发则不要求单独的灭活步骤, 这种不利因素的去除是一种重要的改进。

在变体淀粉酶具有活性的条件下, 用黑曲霉葡糖淀粉酶和 *B. acidopullulyticus* 支链淀粉酶通过糖化 DE10 麦芽糖糊精来评估一定数量的地衣芽孢杆菌 $\alpha$ -淀粉酶变体(具有改变的特异性)。

通过 24 小时间隔并用 HPLC 分析的方法监控糖化反应。标准反应条件如下:

底物浓度 28.2% w/w

温度 60 摄氏度

初始 pH(60 摄氏度) 4.7

酶剂量

葡糖淀粉酶 0.18 AG/g DS

支链淀粉酶 0.06 PUN/g DS

$\alpha$ -淀粉酶 60 NU/g DS

运用了下列酶:

葡糖淀粉酶: AMG(Novo Nordisk) 153 AG/g

支链淀粉酶: Promozyme (Novo Nordisk) 295 PUN/g

**$\alpha$ -淀粉酶:**

Termamyl (Novo Nordisk)	135 KNU/g
V54Y	313 KUN/g
A52W	5116 NU/ml
D53E	3280 NU/ml
D53W	599 NU/ml
A52W+V54Y	134 NU/ml

在上述 $\alpha$ -淀粉酶表中列出的突变可用来表明地衣芽孢杆菌 $\alpha$ -淀粉酶的变体(SEQ ID NO. 2)(Termamyl), 其已被列出的突变修饰。

通过溶解 230g DE 10 喷干麦芽糖糊精制备糖化底物, 用普通玉米粉制备并在 460ml 煮沸的去离子水中调节干物质到约 30% w/w。将其 pH 值调到 4.7(在 60 摄氏度下测量), 并将与 15g 干重对应的底物等分测量, 然后将其转入 50ml 蓝帽玻璃瓶中。

接着, 将该瓶放置在振荡水浴中于 60 摄氏度中平衡并加入酶。必要时将 pH 调至 4.7。周期性地取出 2ml 样品, 将其 pH 调到 3.0, 再在沸水浴中加热 15 分钟以灭活酶活性。冷却后, 用大约 0.1g 的混合床离子交换树脂(BIO-Rad501 X8(D))在旋转混合器上处理 30 分钟以除去盐和可溶性 N。过滤后, 用 HPLC 测定碳氢化合物。72 小时之后, 可获得下列结果:

增加的 $\alpha$ -淀粉酶	%DP1	%DP2	%DP3	%DP4
空白(对照)	96.59	2.2	0.3	1.0
V54Y	96.5	2.2	0.4	0.9
A52W+V54Y	96.4	2.2	0.5	0.9
Termamyl	96.3	2.1	0.8	0.8

与对照比较(在液化过程中存在无活性的 $\alpha$ -淀粉酶), 活性淀粉酶变体 V54Y 和 A52W+V54Y 的存在不会引起潘糖水平的提高(DP3)。

如果这些 $\alpha$ -淀粉酶用于淀粉液化作用, 就没有必要在糖化开始前灭活酶。

**实施例 7**

### 在模拟液化条件下评估地衣芽孢杆菌变体

工业液化过程的标准过程包括两步，即通常所称的初始和第二步液化作用。在第一阶段，于 pH5.5-6.0 条件下匀浆 30-40% 的淀粉(已加入了地衣芽孢杆菌和嗜热脂肪芽孢杆菌的热稳定 $\alpha$ -淀粉酶)并在 105-110 摄氏度下于喷气炊具中加热，其中活性蒸气可注入淀粉流中。在此温度和一定压力下持续 5-10 分钟后，将液化的淀粉迅速冷却到 95 摄氏度并在此温度下持续 60-120 分钟。

在实验室规模上为了评估酶的小数量可运用下列方法：

溶在去离子水(约 30%w/w)中的普通玉米淀粉的 10g 等份悬浮液(Cerestar GL 3406)可于圆锥瓶(Schott GL 125)中称量，该瓶与紧螺帽配件匹配。溶液中的 pH，钙水平以及酶剂量都可以变化。

每一不同系列的实验需用 4 个瓶。这些瓶放置在振荡油浴(Heto VS 01)中 105 摄氏度下保持。约 7 分钟后，倒入凉油将油浴温度降到 95 摄氏度。再每个实验中，分别在 20，40，60 和 90 分后移去瓶并立即在流水中冷却。每瓶中加入一滴 1N 盐酸以灭活酶。通过 Neocuproine 方法测量 DE(降低以葡萄糖形式表示的糖量)来监控反应。

该方法的细节可在“测定糖量降低的精确改进。Dygart, Li, Florida 和 Thomas, 生物化学年报, 13, 368(1965)”中找到。

在 90 分钟后，记录了下列 DEs。

pH 6.0  
每 10.9 毫克酶蛋白/克淀粉,  
加入 5 ppm 钙

I.D 编号	Ter mam yl	杂种+	杂种+	H156 Y+	杂种+	杂种+	Q26 4S	Q264 S+N2 65Y	Q264 S+N2 65Y+ N190 F	H156Y + A181T + A209V
045 -96		16,0	16.9	13.2						
038 -96	6.5	13.9								
035 -96			15.2		12.9	9.9				
033 -96							6.7	7.2	12.1	
031 -96	4.5						7.0	8.8	12.5	
029 -96	4.0					8.7	5.2			7.7
039 -96		14.9	16.3		14.4					

• 杂种 = 按照实施例 3 所述的 BAN/Termamyl PCR 杂种

pH 6.0  
 每 10.9 毫克酶蛋白/克淀粉,  
 加入 40ppm 钙

I.D. 编号	Termamyl
045-96	12.6
007-97	12.1

pH 5.5  
 每 10.9 毫克酶蛋白/克淀粉,  
 加入 5ppm 钙

I.D. no.	杂种+ H156Y+ A181T+ A209V+ N190F+Q264S	杂种+ H156Y+ A181T+ A209V+ N190F	H156Y+ A181T+ A209V+N190F+Q2 64S
001-97	14.8	15.2	12.6

**参考文献:**

Klein, C.等, 生物化学 1992, 31, 8740-8746 ;

Mizuno, J.等, 分子生物学杂志(1993) 234, 1282-1283 ;

Chang, C.等, 分子生物学杂志(1993) 229, 235-238 ;

Larson, S.B.等, 分子生物学杂志(1994)235, 1560-1584 ;

Lawson, C.L.等, 分子生物学杂志(1994)236, 590-600 ;

Qian, M.等, 分子生物学杂志(1993)231, 785-799 ;

Brady, R.L.等 , Acta Crystallogr. Sect.B, 47, 527-535 ; .

Swift, H.J.等, Acta Crystallogr. Sect.B, 47, 535-544 ;

A. Kadziola, 博士论文: “大麦 $\alpha$ -淀粉酶及其底物类似物抑制剂复合物的X-射线晶体研究”, 哥本哈根大学化学系, 1993

MacGregor, E.A., Food Hydrocolloids, 1987, Vol.1, No. 5-6, P.

B. Diderichsen 和 L. Christiansen, 嗜热脂肪芽孢杆菌的 $\alpha$ -淀粉酶克隆, FEMS 微生物通讯: 56: PP.53-60(1988)

Hudson 等, 免疫学实践, 第三版(1989), Blackwell 科学出版社

Sambrook 等, 分子克隆: 实验手册, 第二版, 冷泉港, 1989

S.L. Beaucage 和 M.H. Caruthers, 四面体通讯, 22, 1981, pp.1859-1869

Mattes 等, The EMBO J., 3, 1984, pp.801-805

R.K. Saiki 等, 科学 239, 1977, pp.487-491

Morinaga 等, (1984, 生物技术 2: 646-639)

Nelson 和 Long, 分析生物化学 180, 1989, pp.147-151

Hunkapiller 等, 1984, 自然 310: 105-111

R. Higuchi, B.Krummel, R.K. Saiki(1988). DNA 片段的特异变体和体外制备的一般方法: 蛋白质和 DNA 相互作用的研究。核酸研究 16: 7351-7367

Dubnau 等, 1971, 分子生物学杂志, 56, pp.209-221

Gryczan 等, 1978, 细菌学杂志 134, pp.318-329

S.D. Erlich, 1977, 美国科学院院报 74, pp.1680-1682



---

Boel 等, 1990, 生物化学 29, pp.6244-6249 .

### 序列表

在下列 SEQ ID No.1, 3, 5 中列出了 $\alpha$ -淀粉酶相关基因的 5'和 3'编码序列。5'序列是用小写字母写出的序列的第一分开的部分, 编码序列是序列的中间部分, 此处信号序列用小写字母写出而编码成熟 $\alpha$ -淀粉酶的序列则用大写字母写出。3'序列是序列中用小写字母写出的序列的第三分开的部分。

SEQ ID No. 1

cggaagattggaagtacaaaaataagcaaaagattgtcaatcatgtcatgagccatgcgg-  
gagacggaaaaatcgtctta atgcacgatatttatgcaacgttcgcagatgctgctgaa-  
5 gagattattaaagctgaaagcaaaaggctatcaattggt aactgtatctcagcttga-  
agaagtgaagaagcagagaggctattgaataaatgagtagaagcgccatatacggcgcttttc  
ttttggaagaaaatatagggaaaatggtacttgttaaaaattcgggaatattttatacaacatc-  
atatgtttcacattgaaa ggggaggagaatc

10 atgaacaacaaaaacggctttacgcccgattgctgacgctggtatttgcgctcatcttctt-  
gctgc ctcatctcgcagcagcgggcgGCAAATCTTAATGGGACGCTGATGCAGTATTTT-  
GAATGGTACATGCCAATGACGGCCAA CATTGGAGGCGTTTGCAAAACGACTCGGCATAT-  
TTGGCTGAACACGGTATTACTGCCGTCTGGATTCCCCGGCATATAA GGAACGAGC-  
CAAGCGGATGTGGGCTACGGTGTACGACCTTTATGATTTAGGGGAGTTTCATCAAAAAG-  
15 GGACGGTTC GGACAAAGTACGGCACAAAAGGAGAGCTGCAATCTGCGATCAAAGTCTTC-  
ATTCCC CGACATTAACGTTTACGGGGAT GTGGTCATCAACCACAAAGGCGGCGCTGA-  
TGCGACCGAAGATGTAACCGCGGTTGAAGTCGATCCCGCTGACCGCAACCG CGTAATTT-  
CAGGAGAACACCTAATTAAAGCCTGGACACATTTTCATTTTCCGGGGCGCGGCAGCACATA-  
CAGCGATTTTA AATGGCATTGGTACCATTTTGACGGAACCGATTGGGACGAGTCCCGAAA-  
20 GCTGAACCGCATCTATAAGTTTCAAGGAAAG GCTTGGGATTGGGAAGTTTCCAATGAA-  
AACGGCAACTATGATTATTTGATGTATGCCGACATCGATTATGACCATCCTGA TGTCGCAG-  
CAGAAATTAAGAGATGGGGCACTTGGTATGCCAATGAACTGCAATTGGACGGTTTCCGTCTT-  
GATGCTGTCA AACACATTAAATTTTCTTTTTTTCGGGATTGGGTTAATCATGTCAGGGA-  
AAAACGGGGGAAGGAAATGTTTACGGTAGCT GAATATTGGCAGAAT-  
25 GACTTGGGCGCGCTGAAAACCTATTTGAACAAAACAAATTTTAATCATTAGTGTGTTGACGTGCC  
GCTTCATTATCAGTTCCATGCTGCATCGACACAGGGAGGCGGCTATGATATGAG-  
GAAATTGCTGAACGGTACGGTCGTTT CCAAGCATCCGTTGAAATCGGTTACATTTGTCTG-  
ATAACCATGATACACAGCCGGGGCAATCGCTTGAGTCGACTGTCCAA ACATGGTTTAAG-  
CCGCTTGCTTACGCTTTTATTCTCACAAGGGAATCTGGATACCCTCAGGTTTTCTACGGG-  
30 GATATGTA CGGGACGAAAGGAGACTCCCAGCGCGAAATTCCTGCCTTGAAACACAAAAT-  
TGAACCGATCTTAAAAGCGAGAAAACAGT ATGCGTACGGAGCACAGCATGATTATTTTCGAC-  
CACCATGACATTGTCTGGCTGGACAAGGGAAGGCGACAGCTCGGTTGCA AATTCAGGTTTGG-  
CGGCATTAATAACAGACGGACCCGGTGGGGCAAAGCGAATGTATGTCTGGCCGGCA-  
AAACGCCGGTGA GACATGGCATGACATTACCGGAAACCGTTCGGAGCCGGTTGTCATCA-  
35 ATTCGGAAGGCTGGGGAGAGTTTACGTAAACG GCGGGTCGGTTTCAATTTATGTTCAAA-  
GATAG

aagagcagagaggacggatctcctgaaggaaatccgtttttttatttt

## SEQ ID No. 2

ANLNGTLMQYFEWYMPNDGQHWRRLQND SAYLAEHGITAV  
 WIPPAYKGT SQADVGYGAYDLYDLGEFHQKGTVRTKYGTK  
 5 GELQSAIKSLHSRDIN VYGDV VINHKGGADATEDVTAVEV  
 DPADRNRVISGEHLIKAWTHFHFPGRGSTYSDFKWHWYHF  
 DGTDWDESRLNRIYKFQ GKAWDWEVSNENGN YDYL MYAD  
 IDYDHPDVAAEIKRWGTWYANELQLDGFRLDAVKHIKFSF  
 LRDWVNHVREKTGKEMFTVAEYWQNDLGALENYLNKTNFN  
 10 HSVFDVPLHYQFHAASTQGGGYDMRKLLNGTVVSKHPLKS  
 VTFVDNHDTQPGQSLESTVQTFWKPLAYAFILTRESGYPQ  
 VFYGD MYGTKGDSQREIPALKHKIEPILKARKQYAYGAQH  
 DYFDHHDIVGWTREGDSSVANSGLAALITDGP GGAKRMYV  
 GRQNAGETWHDITGNRSEPVVINSEGWGEFHVNGGSVSIY  
 15 VQR

## SEQ ID No. 3

gccccgcacatacgaaaagactggctgaaaacattgagcctttgatgactgatgatttgg-  
 20 ctgaagaagtggatcgattg tttgagaaaagaagaagaccataaaaataccttgtctgt-  
 catcagacagggtatTTTTTatgctgtccagactgtccgct gtgtaaaaataaggaata-  
 aaggggggttgttattTTTTtactgatatgtaaaatataatttgtataagaaaatgagaggg  
 agaggaaac

atgattcaaaaacgaaagcggacagtttcgttcagacttggtgcttatgtgcacgctggtt-  
 atttgtcagttt . gccgattacaaaaacatcagccGTAAATGGCACGCTGATGCAGTATT-  
 TTGAATGGTATACGCCGAACGACGGCCAGCATT GGAAACGATTGCAGAATGATGCGGAA-  
 CATTATCGGATATCGGAATCACTGCCGTCTGGATTCTCCCGCATACAAAGGA TTGAG-  
 5 CCAATCCGATAACGGATACGGACCTTATGATTTGTATGATTTAGGAGAATTCCAGCAAAA-  
 AGGGACGGTCAGAAC GAAATACGGCACAAAATCAGAGCTTCAAGATGCGATCGGCTCAC-  
 TGCATTCCCGGAACGTCCAAGTATACGGAGATGTGG TTTTGAATCATAAGGCTGGTGCT-  
 GATGCAACAGAAGATGTAAGTCCGTCGAAGTCAATCCGGCCAATAGAAATCAGGAA ACTTCG-  
 GAGGAATATCAAATCAAAGCGTGGACGGATTTTCGTTTTCCGGGCCGTGGAAACACGTACAG-  
 10 TGATTTTTAAATG GCATTGGTATCATTTTCGACGGAGCGGACTGGGATGAATCCC-  
 GGAAGATCAGCCGCATCTTTAAGTTTCGTGGGGAAGGAA AAGCGTGGGATTGGGAAGTAT-  
 CAAGTGAAAACGGCAACTATGACTATTTAATGTATGCTGATGTTGACTACGACCACCCT.  
 GATGTCGTGGCAGAGACAAAAAATGGGGTATCTGGTATGCGAATGAACTGTCATTAGACGG-  
 CTTCCGTATTGATGCCGC CAAACATATTAAATTTTCATTTCTGCGTGATTGGGTTCAGG-  
 15 CGGTGAGACAGGCGACGGGAAAAGAAATGTTTACGGTTG CGGAGTATTGGCAG-  
 AATAATGCCGGGAAACTCGAAAACACTTGAATAAAACAAGCTTTAATCAATCCGTGTTT-  
 GATGTT CCGCTTCATTTCAATTTACAGGCGGCTTCCTCACAAGGAGGCGGATATGATAT-  
 GAGGCGTTTGCTGGACGGTACCGTTGT GTCCAGGCATCCGGAAAAGGCGGTTACATTTGT-  
 TGAAAATCATGACACACAGCCGGGACAGTCATTGGAATCGACAGTCC AAACCTGGTTTTAA-  
 20 ACCGCTTGATACGCCTTTATTTTGACAAGAGAATCCGGTTATCCTCAGGTGTTCTATGGG-  
 GATATG TACGGGACAAAAGGGACATCGCCAAAGGAAATCCCTCACTGAAAGATAATATA-  
 GAGCCGATTTTAAAAGCGCGTAAGGA GTACGCATACGGGCCCCAGCAGATTATATTGAC-  
 CACCCGGATGTGATCGGATGGACGAGGGAAGGTGACAGCTCCGCCG CCAA-  
 ATCAGGTTTGGCCGCTTTAATCACGGACGGACCCGGCGGATCAAAGCGGATGTATGCCGG-  
 25 CCTGAAAATGCCGGC GAGACATGGTATGACATAACGGGCAACCGTTCAGATACTGTAA-  
 AAATCGGATCTGACGGCTGGGGAGAGTTTCATGTAAA CGATGGGTCCGTCTCCATTTAT-  
 GTTCAGAAATAA

ggtaataaaaaaacacctccaagctgagtgcggtatcagcttggg ggtgctttatatt-  
 30 tttcagccgtatgacaaggtcggcatcaggtgtgacaaatcgggtatgctggctgtcata-  
 ggtgaca aatccgggttttgcgctgttggctttttcacatgtctgattttgtataat-  
 caacaggcacggagccggaatctttcgc cttggaaaaataagcggcgatcgtagctgct-  
 tccaatatggattgttcacatcgggatcgctgcttttaatacacaacgtggg atcc

SEQ ID No. 4

VNGTLMQYFEWYTPNDGQHWKRLQNDAEHLSDIGITAVWI  
PPAYKGLSQSDNGYGPYDLYDLGEFQQKGTVRTKYGTKSE  
5 LQDAIGSLHSRNVQYGDVVLNHNKAGADATEDVTAVEVNP  
ANRNQETSEEYQIKAWTDERFPGRGNTYSDFKWHWHYHFDG  
ADWDESRKISRIFKFRGEGKAWDWEVSSSENGNYDILMYAD  
VDYDHPDVVAETKKWGIWYANELSLDGFRIIDAAKHIFSF  
LRDWVQAVRQATGKEMFTVAEYWQNNAGKLENYLNKTSFN  
10 QSVFDVPLHFENLQAASSQGGYDMRRLLDGTVVSRHPEKA  
VTFVENHDTQPGQSLESTVQTFWKPLAYAFILTRESGYPQ  
VfyGDMYGTkGTSPKEIPSLKDNIEPILKARKEYAYGPQH  
DYIDHPDVIGWTREGDSSAAKSGLAALITDGPGGSKRMYA  
GLKNAGETWYDITGNRSDTVKIGSDGWGEFHVNDGSVSIY  
15

SEQ ID No. 5

aaattcgatattgaaaacgattacaaataaaaattataatagacgtaaacgttcgagggg-  
 ttgctccctttttactcttt            ttatgcaatcgtttcccttaattttttggaagccaaacc-  
 5 gtcgaatgtaacatttgattaagggggaagggcatt

gtgct                    aacgtttcaccgcatcattcgaaaaggatggatgttcctgctcgcgtt-  
 tttgctcaactgtctcgtgtttctgcccacag            gacagcccgccaaggctGCCGCACCGT-  
 TTAACGGCACCATGATGCAGTATTTTGAATGGTACTTGCCGGATGATGGCAGG            TTATGG-  
 10 ACCAAAGTGGCCAATGAAGCCAACA ACTTATCCAGCCTTGGCATCACCGCTCTTTGGCTG-  
 CCGCCCCGCTTACAA            AGGAACAAGCCGCAGCGACGTAGGGTACGGAGTATACGACTTGTA-  
 TGACCTCGGCGAATTCAATCAAAAAGGGACCGTCC            GCACAAAATACGGAACAAAAGCTC-  
 AATATCTTCAAGCCATTCAAGCCGCCACGCCGCTGGAATGCAAGTGTACGCCGAT            GTC-  
 GTGTTTCGACCATAAAGGCGGCGCTGACGGCAGGAATGGGTGGACGCCGTCGAAGTCAATCCG-  
 15 TCCGACCGCAACCA            AGAAATCTCGGGCACCTATCAAATCCAAGCATGGACGAAATTTGA-  
 TTTTCCCGGGCGGGGCAACACCTACTCCAGCTTTA            AGTGGCGCTGGTACCATTTTG-  
 ACGGCGTTGATTGGGACGAAAGCCGAAAATTGAGCCGATTTACAAATTCGCGGCATC  
 GGCAAAGCGTGGGATTGGGAAGTAGACACGGAAAACGGAACTATGACTACTTAATGTAT-  
 GCCGACCTTGATATGGATCA            TCCCGAAGTCGTGACCGAGCTGAAAACTGGGGGAAATG-  
 20 GTATGTCAACACAACGAACATTGATGGGTTCCGGCTTGATG            CCGTCAAGCATATTAAGT-  
 TCAGTTTTTTTTCTGATTGGTTGTCGTATGTGCGTTCTCAGACTGGCAAGCCGCTATTTACC  
 GTCGGGGAATATTGGAGCTATGACATCAACAAGTTGCACAATTACATTACGAAAACAGAC-  
 GGAACGATGTCTTTGTTTGA            TGCCCCGTTACACAACAAATTTTATACCGCTTCCAAATCAGG-  
 GGGCGCATTTGATATGCGCACGTTAATGACCAATACTC            TCATGAAAGATCAAC-  
 25 CGACATTGGCCGTCACCTTCGTTGATAATCATGACACCGAACCCGGCCAAGCGCTGCAGTCATGG  
 GTCGACCCATGGTTCAAACCGTTGGCTTACGCCTTTATTCTAACTCGG-  
 CAGGAAGGATACCCGTGCGTCTTTTATGGTGA            CTATTATGGCATTCCACAATATAACAT-  
 TCCTTCGCTGAAAAGCAAATCGATCCGCTCCTCATCGCGCGCAGGGATTATG            CTTACG-  
 GAACGCAACATGATTATCTTGATCACTCCGACATCATCGGGTGGACAAGGGAAGGGGGCA-  
 30 CTGAAAACCAGGA            TCCGGACTGGCCGCACTGATCACCGATGGGCCGGGAGGAAGCAA-  
 TGGATGTACGTTGGCAAACAACACGCTGGAAAAGT            GTTCTATGACCTTACGGCAACCGGAG-  
 TGACACCGTCACCATCAACAGTGATGGATGGGGGGAATTCAAAGTCAATGGCG            GTT-  
 CGGTTTCGGTTTGGGTTCTAGAAAACGACCGTTTCTACCATCGCTCGGCCGATCACAA-  
 CCCGACCGTGGACTGGT GAATTCGTCCGTTGGACCGAACCACGGTTGGTGGCATGGCCTTGA  
 - 35

tgectgcga

## SEQ ID No. 6

AAPFNGTMMQYFEWYLPDDGTLWTKVANEANNLSSLGITA  
 LWLPPAYKGTSRSDVGYGVYDLYDLGEFNQKGTVRTKYGT  
 5 KAQYLQAIQAAHAAGMQVYADVVDHKGGADGTEWVDAVE  
 VNPSDRNQEISGTYQIQAWTKFDFPGRGNTYSSFKWRWYH  
 FDGVDWDESRKLSRIYKFRGIGKAWDWEVDTENGNYDYLM  
 YADLDMDHPEVVTELKNWGKWYVNTTNIDGFRLDAVKHIK  
 FSFFPDWLSYVRSQTGKPLEFTVGEYWSYDINKLHNYITKT  
 10 DGTMSLEFDAPLHNKFYTASKSGGAFDMRTLMTNTLMKDQP  
 TLAVTFVDNHDTEPGQALQSWVDPWEKPLAYAFILTRQEG  
 YPCVFYGDYYGIPQYNIPSLKSKIDPLLIARRDYAYGTQH  
 DYLDHSDIIGWTREGGTEKPGSGLAALITDGPGGSKWMYV  
 GKQHAGKVFYDLTGNRSDTVTINSDGWGEFKVNGGSVSVW  
 15 VPRKTTVSTIARPITTRPWTGEFVRWTEPRLVAW

## SEQ ID No. 10

1 ATPADWRSQS IYFLLTDREFA RTDGSTTATC  
 20 31 NTADQKYCGG TWQGIIDKLD YIQGMGFTAI  
 61 WITPVTAQLP QTTAYGDAYH GYWQQDIYSL  
 91 NENYGTADDL KALSSALHER GMYLMVDVVA  
 121 NHMGYDGAGS SVDYSVFKPF SSQDYFHPFC  
 151 FIQNYEDQTQ VEDCWLGDNT VSLPDLDTTK  
 25 181 DVVKNEWYDW VGSLVSNYSI DGLRIDTVKH  
 211 VQKDFWPGYN KAAGVYCIGE VLDGDPAYTC  
 241 PYQNVMDGVL NYPIYYPLLN AFKSTSGSMD  
 271 DLYNMINTVK SDCPDSTLLG TFVENHDNPR  
 301 FASYTNDIAL AKNVAAFIIIL NDGIPIIYAG  
 30 331 QEQHYAGGND PANREATWLS GYPTDSELYK  
 361 LIASANAIRN YAISKDTGFV TYKNWPIYKD  
 391 DITIAMRKGT DGSQIVTILS NKGASGDSYT  
 421 LSLSGAGYTA GQQLTEVIGC TTVTGVSDGN  
 451 VPVPMAGGLP RVLYPTEKLA GSKICSSS

1  
CAT CAT AAT GGA ACA AAT GGT ACT ATG ATG CAA TAT TTC GAA TGG TAT TTG CCA AAT GAC  
H H N G T N G T M H Q Y F E W Y L P N D

21  
GGG AAT CAT TGG AAC AGG TTG AGG GAT GAC GCA GCT AAC TTA AAG AGT AAA GGG ATA ACA  
G N H W N R L R D D A A N L K S K G I T

41  
GCT GTA TGG ATC CCA CCT GCA TGG AAG GGG ACT TCC CAG AAT GAT GTA GGT TAT GGA GCC  
A V W I P P A W K G T S Q N D V G Y G A

61  
TAT GAT TTA TAT GAT CTT GGA GAG TTT AAC CAG AAG GGG ACG GTT CGT ACA AAA TAT GGA  
Y D L Y D L G E F N Q K G T V R T K Y G

81  
ACA CGC AAC CAG CTA CAG GCT GCG GTG ACC TCT TTA AAA AAT AAC GGC ATT CAG GTA TAT  
T R N Q L Q A A V T S L K N N G I Q V Y

101  
GGT GAT GTC GTC ATG AAT CAT AAA GGT GGA GCA GAT GGT ACG GAA ATT GTA AAT GCG GTA  
G D V V M N H K G G A D G T E I V N A V

121  
GAA GTG AAT CCG AGC AAC CGA AAC CAG GAA ACC TCA GGA GAG TAT GCA ATA GAA GCG TGG  
E V N R S N R N Q E T S G E Y A I E A W

141  
ACA AAG TTT GAT TTT CCT GGA AGA GGA AAT AAC CAT TCC AGC TTT AAG TGG CGC TGG TAT  
T K F D F P G R G N N H S S F K W R W Y

161  
CAT TTT GAT GGG ACA GAT TGG GAT CAG TCA CGC CAG CTT CAA AAC AAA ATA TAT AAA TTC  
H F D G T D W D Q S R Q L Q N K I Y K F

181  
AGG GGA ACA GGC AAG GCC TGG GAC TGG GAA GTC GAT ACA GAG AAT GGC AAC TAT GAC TAT  
R G T G K A W D W E V D T E N G N Y D Y

201  
CTT ATG TAT CCA GAC GTG GAT ATG GAT CAC CCA GAA GTA ATA CAT GAA CTT AGA AAC TGG  
L M Y A D V D M D H P E V I H E L R N W

221  
GGA GTG TGG TAT ACG AAT ACA CTG AAC CTT GAT GGA TTT AGA ATA GAT GCA GTG AAA CAT  
G V W Y T N T L N L D G F R I D A V X H

241  
ATA AAA TAT AGC TTT ACG AGA GAT TGG CTT ACA CAT GTC CGT AAC ACC ACA GGT AAA CCA  
I X Y S F T R D W L T H V R N T T G K P

261  
ATG TTT GCA GTG GCT GAG TTT TGG AAA AAT GAC CTT GGT GCA ATT GAA AAC TAT TTG AAT  
M F A V A E F W K N D L G A I E N Y L N

281  
AAA ACA AGT TGG AAT CAC TCG GTG TTT GAT GTT CCT CTC CAC TAT AAT TTG TAC AAT GCA  
K T S W N H S V F D V P L H Y N L Y N A

图 1



301  
TCT AAT AGC GGT GGT TAT TAT GAT ATG AGA AAT ATT TTA AAT GGT TCT GTG GTG CAA AAA  
S N S G G Y Y D M R N I L N G S V V Q K

321  
CAT CCA ACA CAT GCC GTT ACT TTT GTT GAT AAC CAT GAT TCT CAG CCC GGG GAA GCA TTG  
H P T H A V T F V D N H D S Q P G E A L

341  
GAA TCC TTT GTT CAA CAA TGG TTT AAA CCA CTT GCA TAT GCA TTG GTT CTG ACA AGG GAA  
E S F V Q Q W F K P L A Y A L V L T R E

361  
CAA GGT TAT CCT TCC GTA TTT TAT GGG GAT TAC TAC GGT ATC CCA ACC CAT GGT GTT CCG  
Q G Y P S V F Y G D Y Y G I P T H G V P

381  
GCT ATG AAA TCT AAA ATA GAC CCT CTT CTG CAG GCA CGT CAA ACT TTT GCC TAT GGT ACG  
A M K S K I D P L L Q A R Q T F A Y G T

401  
CAG CAT GAT TAC TTT GAT CAT CAT GAT ATT ATC GGT TGG ACA AGA GAG GGA AAT AGC TCC  
Q H D Y F D H H D I I G W T R E G N S S

421  
CAT CCA AAT TCA GGC CTT GCC ACC ATT ATG TCA GAT GGT CCA GGT GGT AAC AAA TGG ATG  
H P N S G L A T I M S D G P G G N K W M

441  
TAT GTG GGG AAA AAT AAA GCG GGA CAA GTT TGG AGA GAT ATT ACC GGA AAT AGG ACA GGC  
Y V G K N K A G Q V W R D I T G N R T G

261  
ACC GTC ACA ATT AAT GCA GAC GGA TGG GGT AAT TTC TCT GTT AAT GGA GGG TCC GTT TCG  
T V T I N A D G W G N F S V N G G S V S

481  
GTT TGG GTG AAG CAA TAA  
V W V K Q \*

图 1 (续)

	10	20	30	40	50	60	
1	HHNGTNGTMMQYFEWYLPNDGNHWNRLRDDAANLKS KGITAVWI PPAWKGTSQNDVGYGA						60
3	-AAPFNGTMMQYFEWYLPDDGTLWTKVANEANNLSSLGITALWLP PAYKGTSSRSDVGYGV						59
2	HHNGTNGTMMQYFEWHL PNDGNHWNRLRDDASNLNRGITA IWI PPAWKGTSQNDVGYGA						60
4	HHNGTNGTMMQYFEWYLPNDGNHWNRLNSDASNLSKSGITAVWI PPAWKGASQNDVGYGA						60
	70	80	90	100	110	120	
1	YDLYDLGEFNQKGTVRTKYGTRNQLQA AVTSLKNNGIQVYGDVVMNHKGGADGTEIVNAV						120
3	YDLYDLGEFNQKGTVRTKYGTKAQYLQAIQA AHAAGMQVYADV VFDHKGGADGTEWVDAV						119
2	YDLYDLGEFNQKGTVRTKYGTRSQLES AIHALKNNGVQVYGDVVMNHKGGADATENVLAV						120
4	YDLYDLGEFNQKGTVRTKYGTRSQQA AVTSLKNNGIQVYGDVVMNHKGGADATEMVRVAV						120
	130	140	150	160	170	180	
1	EVNRSNRNQETSGEYAI EAWTKFDFPGRGNHSSFKWRWYHFDGTDWDQSRQLQNKIYKF						180
3	EVNPSDRNQEISGTYQIQAWTKFDFPGRGNTYSSFKWRWYHFDGVDWDESRKLS -RIYKF						178
2	EVNPNNRNQEISGDYTI EAWTKFDFPGRGNTYSDFKWRWYHFDGVDWQSRQFNRIYKF						180
4	EVNPNNRNQEV TGEYTI EAWTRFDFPGRGNTHSSFKWRWYHFDGVDWQSRRLNRIYKF						180
	190	200	210	220	230	240	
1	RGTGKAWDWEVDTE NGNYDYL MYADVDM DHPEVIHEL RNWGVWY TNLNLDGFRIDAVKH						240
3	RGIGKAWDWEVDTE NGNYDYL MYADLMDHPEV VTE LKNWGK WYVNTNIDGFR LDAVKH						238
2	RGDGKAWDWEVDSE NGNYDYL MYADVDM DHPEVVNELRRWGEWY TNLNLDGFRIDAVKH						240
4	RGGKAWDWEVDTE NGNYDYL MYADIDMDHPEV VNELRNWGVWY TNLGLDGFRI DAVKH						240
	250	260	270	280	290	300	
1	IKYSFTRDWLTHVRNTTGKPMFAVAEFWKNDLGA IENYLNKTSWNHSAFDVPLHYNL YNA						300
3	IKFSFFPDWLSYVRSQTGKPLFTVGEYWSYDINKLHNYITKTDG TMSLFDAPLHNKFYTA						298
2	IKYSFTRDWLTHVRNATGKEMFAVAEFWKNDLGA IENYLNKTNWNHSAFDVPLHYNL YNA						300
4	IKYSFTRDWINHVR SATGKNMFAVAEFWKNDLGA IENYLQKTNWNHSAFDVPLHYNL YNA						300
	310	320	330	340	350	360	
1	SNSGGYYDMRNILNGSVVQKHPH AVTFVDNHDSQPGEALESFVQQWFKPLAYALV LTRI						360
3	SKSGGAFDMRTLMTNTLMKDQPTLAVTFVDNHDTEPGQALQSWVDPWFKPLAYAFILTRQ						358
2	SNSGGNYDMAKLLNGT VVQKHPH AVTFVDNHDSQPGEALESFVQEWFKPLAYALIL TRE						360
4	SKSGGNYDMRNI FNGT VVQRHP SHAVTFVDNHDSQPPEALESFVEEWFKPLAYALTL TRE						360
	370	380	390	400	410	420	
1	QGYPSVFGDY YGIP THGVPAMKSKIDP LLQARQTFAYGTQH DYFDHDI IGW TREGNSS						420
3	EGYPCVFGDY YGIPQYNI PSLKSKIDP LLIARRDYAYGTQH DYLDHSDI IGW TREGGTE						418
2	QGYPSVFGDY YGIP THSV PAMKAKIDP ILEARQNFAYGTQH DYFDHNI IGW TREGNTT						420
4	QGYPSVFGDY YGIP THGVPAMRSKIDP ILEARQKYAYGKQNDYLDH HNI IGW TREGNTA						420
	430	440	450	460	470	480	
1	HPNSGLATIMSDGPGGNKWMYV GKNKAGQVWRDITGNRTGT V TINADGWGNFSVNGGSVS						480
3	KPGSGLAALITDGP GSKWMYV GKHAGKVFYDLTGNRSDT V TINSDGWGEFKNVNGGSVS						478
2	HPNSGLATIMSDGPGGEKWMYV GQNKAGQVWHDITGNKPGT V TINADGWANFSVNGGSVS						480
4	HPNSGLATIMSDGAGGSKWMFVGRNKAGQVWSDITGNRTGT V TINADGWGNFSVNGGSVS						480
	490	500	510	520	530	540	
1	VWVKQ						485
3	VWVPRKTTVSTIARPI TTRPWTGEFVRWTEPR LVAW						514
2	IWVKR						485
4	IWVNK						485

图 2

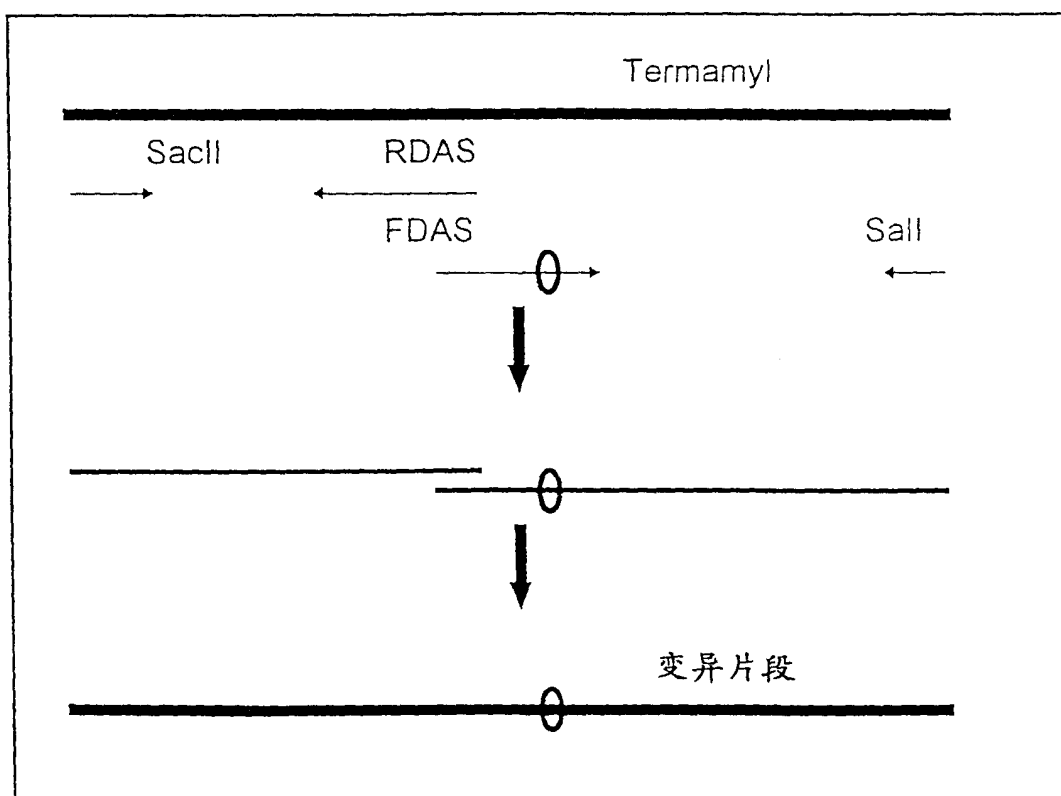


图 3