

[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 200410061736.6

[45] 授权公告日 2010 年 3 月 3 日

[11] 授权公告号 CN 100593034C

[22] 申请日 1998.10.30

[21] 申请号 200410061736.6

分案原申请号 98810871.2

[30] 优先权

[32] 1997.10.30 [33] DK [31] 1240/97

[32] 1998.7.14 [33] DK [31] PA199800936

[73] 专利权人 诺维信公司

地址 丹麦鲍斯韦

[72] 发明人 托宾·V·博彻特 阿伦·斯文森

卡斯滕·安德森

比贾尼·R·尼尔森

托宾·L·尼森 索伦·克贾鲁尔夫

[56] 参考文献

WO9510603 1995.4.20

CN1134725 1996.10.29

CN1118172 1996.3.6

WO9623873 1996.8.8

WO9623874 1996.8.8

审查员 滕 蕾

[74] 专利代理机构 北京市柳沈律师事务所

代理人 封新琴 巫肖南

权利要求书 3 页 说明书 90 页 附图 9 页

[54] 发明名称

α - 淀粉酶突变体

[57] 摘要

本发明涉及亲本类透阿米尔 α - 淀粉酶的变体，它相对所述亲本 α 淀粉酶，至少以下性能之一发生改变：i) 在 pH8 到 10.5 的 pH 稳定性提高；和/或 ii) 在 pH8 到 10.5 对 Ca^{2+} 的稳定性提高，和/或 iii) 在 10℃ 到 60℃ 的温度范围，特异活性增强。

1. 亲本类透阿米尔 α -淀粉酶的变体, 该变体具有 α -淀粉酶活性, 并且亲本 α -淀粉酶序列为 SEQ ID NO: 2, 其中至少 1 个突变选自下组: Δ (D183-G184); Δ Q174; M323L; R181S; K185R; A186T, R, I; K311R; N195F; T461P; E346Q; K385R 和 K269S。

2. 权利要求 1 的变体, 其中所述变体包含(i) Δ (D183-G184)缺失, 和(ii)至少 1 个选自下列的突变: Δ Q174; M323L; R181S; K185R; A186T, R, I; K311R; N195F; T461P; E346Q; K385R 和 K269S。

3. 权利要求 1 或 2 的变体, 其在 pH8 到 10.5 的 pH 稳定性提高; 并且具有 1 或多个对应 SEQ ID NO: 2 编号的下列位点的突变: R181S, A186T, 和 K311R。

4. 权利要求 1 或 2 的变体, 其在 pH8 到 10.5 对 Ca^{2+} 的稳定性提高, 具有 1 或多个对应 SEQ ID NO: 2 编号的下列位点的突变: K185R; A186T, I, R; N195F; E346Q; K385R。

5. 权利要求 1 或 2 的变体, 其中所述变体在温度为 10 到 60°C 具有增强的比活性, 并具有 1 个对应 SEQ ID NO: 2 编号的下列位点的突变: K269S。

6. 权利要求 5 的变体, 其中所述变体在温度为 20-50°C 具有增强的比活性, 并具有 1 个对应 SEQ ID NO: 2 编号的下列位点的突变: K269S。

7. 权利要求 5 的变体, 其中所述变体在温度为 30-40°C 具有增强的比活性, 并具有 1 个对应 SEQ ID NO: 2 编号的下列位点的突变: K269S。

8. 一种 DNA 构建体, 它含有编码权利要求 1-7 任何一项所述 α -淀粉酶变体的 DNA 序列。

9. 一种携带权利要求 8 所述 DNA 构建体的重组表达载体。

10. 一种用权利要求 8 所述 DNA 构建体或权利要求 9 所述载体进行转化的细胞。

11. 权利要求 10 的细胞, 它是一种微生物细胞。

12. 权利要求 11 的细胞, 它是细菌或真菌。

13. 权利要求 12 的细胞, 它是革兰氏阳性细菌。

14. 权利要求 13 的细胞, 其中所述革兰氏阳性细菌是枯草芽孢杆菌, 地

衣形芽孢杆菌,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. 权利要求 1-7 任何一项所述 α -淀粉酶变体用于洗涤和/或清洗餐具的用途。

16. 含有权利要求 1-7 任何一项所述 α -淀粉酶变体的洗涤剂添加剂,任选该添加剂是非 dusting 颗粒,稳定化的液体或被保护的酶。

17. 权利要求 16 的洗涤剂添加剂,它含有 0.02-200mg 酶蛋白/g 添加剂。

18. 权利要求 16 或 17 的洗涤剂添加剂,它还含有选自蛋白酶、脂酶和过氧化物酶的另一种酶,另一种淀粉水解酶和/或纤维素酶。

19. 一种包含权利要求 1-7 任何一项所述 α -淀粉酶变体的洗涤剂组合物。

20. 权利要求 19 的洗涤剂组合物,它另外包含选自蛋白酶、脂酶和过氧化物酶的另一种酶,另一种淀粉水解酶和/或纤维素酶。

21. 一种手洗或自动餐具洗涤剂组合物,该组合物包含权利要求 1-7 任何一项所述的 α -淀粉酶变体。

22. 权利要求 21 的清洗餐具的洗涤剂组合物,该组合物另外包含选自蛋白酶、脂酶和过氧化物酶的另一种酶,另一种淀粉水解酶和/或纤维素酶。

23. 一种手洗或自动的衣物洗涤组合物,该组合物包含权利要求 1-7 任何一项所述的 α -淀粉酶变体。

24. 权利要求 23 的衣物洗涤组合物,该组合物另外包含选自蛋白酶、脂酶和过氧化物酶的另一种酶,淀粉水解酶和/或纤维素酶。

25. 制备权利要求 1 的 α -淀粉酶的方法,

所述方法包括以下步骤:

i) 将两个或多个 pH, 温度和/或稳定性曲线迥异的 α -淀粉酶的 3D 结构之分子动力学进行比较,从而确定(a) α -淀粉酶突变的目标位点和/或区域;

ii) 在所确定的位点和/或区域进行 1 或多个氨基酸的取代,添加和/或缺失。

26. 权利要求 25 的方法,其中所述 α -淀粉酶至少 70%同源。

27. 权利要求 26 的方法,其中所述 α -淀粉酶至少 80%同源。

28. 权利要求 27 的方法,其中所述 α -淀粉酶至少 90%同源。

-
29. 权利要求 28 的方法，其中所述 α -淀粉酶至少 95%同源。
30. 权利要求 26 的方法，其中所比较的 α -淀粉酶是类透阿米尔 α -淀粉酶。
31. 权利要求 30 的方法，其中所比较的 α -淀粉酶是 SEQ ID NO: 1 到 SEQ ID NO: 8 所示的任何一个 α -淀粉酶。

α -淀粉酶突变体

本申请是题为“ α -淀粉酶突变体”的中国专利申请 98810871.2 的分案申请。该母案申请是国际申请 PCT/DK98/00471, 其公开号为 WO 99/23211。

技术领域

本发明涉及亲本类透阿米尔(Termamyl-like) α -淀粉酶的变体(突变体), 其中温和/或高 pH 下的活性更高。

背景技术

α -淀粉酶(α -1, 4-葡聚糖-4-葡聚糖水解酶, EC3.2.1.1)构成这样一组酶, 它们催化淀粉和其他线性及分枝 1, 4-葡萄糖苷寡-糖和多糖的水解。

有大量专利和科学文献涉及这类工业上非常重要的酶。从例如 WO90/11352, WO95/10603, WO95/26397, WO96/23873 和 WO96/23874 可以知道许多 α -淀粉酶, 比如类透阿米尔 α -淀粉酶变体。

在更近期的涉及 α -淀粉酶的公开文献中, WO96/23874 提供了一种类透阿米尔 α -淀粉酶的三维, X-射线晶体结构数据, 该酶含有解淀粉芽孢杆菌 α -淀粉酶(BANTM)的 300 个氨基端氨基酸残基以及含有所述氨基酸序列的地衣形芽孢杆菌 α -淀粉酶的羧基端 301-483 位氨基酸(后者以透阿米尔TM的商品名出售), 因此该酶与工业上很重要的芽孢杆菌 α -淀粉酶(其在本文中包含在术语“类透阿米尔 α -淀粉酶”中, 该术语特别包括地衣形芽孢杆菌, 解淀粉芽孢杆菌(BANTM)和嗜热脂肪芽孢杆菌(BSGTM) α -淀粉酶)有密切的关系。WO96/23874 还描述了在分析亲本类透阿米尔 α -淀粉酶之结构的基础上, 设计相对亲本改变了性质的亲本类透阿米尔 α -淀粉酶之变体的方法。

发明简述

本发明涉及新的类透阿米尔 α -淀粉酶的 α -淀粉分解酶变体(突变体), 该变体在高 pH 和中温的洗涤能力(相对亲本 α -淀粉酶)提高。

在本发明文中术语“中温”意指 10°C 到 60°C, 优选 20°C 到 50°C, 特别是 30°C 到 40°C。

术语“高 pH”意指目前用于洗涤的碱性 pH, 更具体说是大约 pH8 到 10.5。

在本发明文中, “低温 α -淀粉酶”意指一种在 0-30°C 的温度范围内具有

相对最佳活性的 α -淀粉酶。

在本发明文中,“中温 α -淀粉酶”意指一种在30℃-60℃的温度范围内具有最佳活性的 α -淀粉酶。比如,SP690和SP722 α -淀粉酶分别是“中温 α -淀粉酶”。

在本发明文中,“高温 α -淀粉酶”意指一种在60℃-110℃的温度范围内具有最佳活性的 α -淀粉酶。比如,透阿米尔是一种“高温 α -淀粉酶”。

本发明的变体可以实现这些特性的改变:

类透阿米尔 α -淀粉酶在pH8到10.5的稳定性;和/或在pH8到10.5对 Ca^{2+} 的稳定性;和/或在10℃-60℃,优选20℃-50℃,特别是30℃-40℃的特异活性。

应当注意的是,相对最适温度通常取决于所用的特定pH。换句话说,在例如pH8确定的相对最适温度可能与例如pH10确定的相对最适温度非常不同。

温度对酶活性的影响

活性部位及其周围的动力学取决于温度和氨基酸组成,并且对一种酶的相对最适温度有重要意义。通过比较中温和高温 α -淀粉酶的动力学,可以确定出高温 α -淀粉酶在中等温度下的功能关键区域。图2显示了SP722 α -淀粉酶(SEQ ID NO: 2)和地衣形芽孢杆菌 α -淀粉酶(以Termanyl®购自Novo Nodisk)(SEQ ID NO: 4)。

图中显示,在中温范围(30-60℃),SP722绝对活性的相对最适温度比同源的地衣形芽孢杆菌 α -淀粉酶高,后者在60-100℃左右具有最佳活性。这些分布图主要取决于活性部位残基及其周围的温度稳定性和动力学。另外,活性分布图取决于所用pH和活性部位残基的pKa。

本发明更具体地涉及:

1. 一种亲本类透阿米尔 α -淀粉酶的变体,该变体具有 α -淀粉酶活性,所述变体包含1或多个突变,其对应于下列SEQ ID NO: 2所示氨基酸序列中的突变:

T141, K142, F143, D144, F145, P146, G147, R148, G149, Q174, R181, G182, D183, G184, K185, A186, W189, S193, N195, H107, K108, G109, D166, W167, D168, Q169, S170, R171, Q172, F173, F267, W268, K269, N270, D271, L272, G273, A274, L275, K311, E346, K385, G456, N457, K458, P459, G460, T461, V462, T463。

2. 项1的变体,所述变体有1或多个下列取代或缺失:

T141A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, W, Y, V;
K142A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
F143A, D, R, N, C, E, Q, G, H, I, L, K, M, P, S, T, W, Y, V;
D144A, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
F145A, D, R, N, C, E, Q, G, H, I, L, K, M, P, S, T, W, Y, V;
P146A, D, R, N, C, E, Q, G, H, I, L, K, M, F, S, T, W, Y, V;
G147A, D, R, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;
R148A, D, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
G149A, D, R, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;
R181*, A, D, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
G182*, A, D, R, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;
D183*, A, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
G184*, A, R, D, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;
K185A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
A186D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
W189A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, Y, V;
S193A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, T, W, Y, V;
N195A, D, R, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
H107A, D, R, N, C, E, Q, G, I, L, K, M, F, P, S, T, W, Y, V;
K108A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
G109A, D, R, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;
D166A, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
W167A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, Y, V;
D168A, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
Q169A, D, R, N, C, E, G, H, I, L, K, M, F, P, S, T, W, Y, V;
S170A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, T, W, Y, V;
R171A, D, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
Q172A, D, R, N, C, E, G, H, I, L, K, M, F, P, S, T, W, Y, V;
F173A, D, R, N, C, E, Q, G, H, I, L, K, M, P, S, T, W, Y, V;
Q174*, A, D, R, N, C, E, G, H, I, L, K, M, F, P, S, T, W, Y, V;
F267A, D, R, N, C, E, Q, G, H, I, L, K, M, P, S, T, W, Y, V;
W268A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, Y, V;
K269A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
N270A, D, R, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
D271A, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
L272A, D, R, N, C, E, Q, G, H, I, K, M, F, P, S, T, W, Y, V;
G273A, D, R, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;
A274D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;

L275A, D, R, N, C, E, Q, G, H, I, K, M, F, P, S, T, W, Y, V;
 K311A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
 E346A, D, R, N, C, Q, G, H, I, K, L, M, F, P, S, T, W, Y, V;
 K385A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
 G456A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
 N457A, D, R, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
 K458A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
 P459A, D, R, N, C, E, Q, G, H, I, L, K, M, F, S, T, W, Y, V;
 G460A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
 T461A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, W, Y, V;
 V462A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y;
 T463A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, W, Y, V.

3. 项2的变体, 其中所述变体有1或多个下列取代或缺失:

K142R; S193P; N195F; K269R, Q; N270Y, R, D; K311R; E346Q; K385R;
 K458R; P459T; T461P; Q174*; R181Q, N, S; G182T, S, N; D183*; G184*;
 K185A, R, D, C, E, Q, G, H, I, L, M, N, F, P, S, T, W, Y, V; A186T, S, N, I, V, R;
 W189T, S, N, Q.

4. 项1-3的变体, 其中所述变体在D183+G184位点缺失, 还有1或多个下列取代或缺失: K142R; S193P; N195F; K269R, Q; N270Y, R, D; K311R;
 E346Q; K385R; K458R; P459T; T461P; Q174*; R181Q, N, S; G182T, S,
 N; D183*; G184*; K185A, R, D, C, E, Q, G, H, I, L, M, N, F, P, S, T, W, Y, V;
 A186T, S, N, I, V, R; W189T, S, N, Q.

5. 项1-4任何一项的变体, 其中相对亲本 α -淀粉酶所述变体至少下列性能之一发生改变:

- i) 在pH8到10.5的pH稳定性提高; 和/或
- ii) 在pH8到10.5的 Ca^{2+} 稳定性提高, 和/或
- iii) 在温度为10到60℃, 优选20-50℃, 特别是30-40℃特异活性增强。

6. 项1-5任何一项的变体, 在pH8到10.5稳定性提高, 具有1或多个对应下列位点(使用SEQ ID NO: 2的编号)的位点的突变: T141, K142, F143, D144, F145, P146, G147, R148, G149, R181, A186, S193, N195, K269, N270, K311, K458, P459, T461。

7. 项6的变体, 该变体具有1或多个下列取代:

T141A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, W, Y, V;
 K142A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
 F143A, D, R, N, C, E, Q, G, H, I, L, K, M, P, S, T, W, Y, V;
 D144A, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;

F145A, D, R, N, C, E, Q, G, H, I, L, K, M, P, S, T, W, Y, V;
 P146A, D, R, N, C, E, Q, G, H, I, L, K, M, F, S, T, W, Y, V;
 G147A, D, R, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;
 R148A, D, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
 G149A, D, R, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;
 K181A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
 A186D, R, N, C, E, Q, G, H, I, L, P, K, M, F, S, T, W, Y, V;
 S193A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, T, W, Y, V;
 N195A, D, R, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
 K269A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
 N270A, D, R, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
 K311A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
 K458A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
 P459A, D, R, N, C, E, Q, G, H, I, L, K, M, F, S, T, W, Y, V;
 T461A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, W, Y, V.

8. 项7的变体, 该变体具有1或多个下列取代: K142R, R181S, A186T, S193P, N195F, K269R, N270Y, K311R, K458R, P459T 和 T461P.

9. 项1-5的变体, 在 pH8 到 10.5 对 Ca^{2+} 的稳定性提高, 具有1或多个对应下列位点(使用 SEQ ID NO: 2 的编号)的突变: R181, G182, D183, G184, K185, A186, W189, N195, N270, E346, K385, K458, P459.

10. 项9的变体, 该变体具有1或多个下列取代或缺失:

R181*, A, D, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
 G182*, A, D, R, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;
 D183*, A, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
 G184*, A, R, D, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;
 K185A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
 A186D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
 W189A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, Y, V;
 N195A, D, R, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
 N270A, R, D, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;
 E346A, R, D, N, C, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
 K385A, R, D, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
 K458A, R, D, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
 P459A, R, D, N, C, E, Q, G, H, I, L, K, M, F, S, T, W, Y, V.

11. 项10的变体, 其中所述变体具有1或多个下列取代或缺失:

R181Q, N; G182T, S, N; D183*; G184*; K185A, R, D, C, E, Q, G, H, I, L, M, N, F, P, S, T, W, Y, V; A186T, S, N, I, V; W189T, S, N, Q; N195F; N270R, D; E346Q; K385R; K458R; P459T.

12. 项 1-11 的变体, 其中所述亲本类透阿米尔 α -淀粉酶选自:

具有 SEQ ID NO: 1 所示序列的芽孢杆菌菌株 NCIB12512 α -淀粉酶;

具有 SEQ ID NO: 5 所示序列的解淀粉芽孢杆菌 α -淀粉酶;

具有 SEQ ID NO: 4 所示序列的地衣形芽孢杆菌 α -淀粉酶。

13. 项 1-5 的变体, 在 10-60℃, 优选 20-50℃, 特别是 30-40℃ 的温度范围内特异活性提高, 在 1 或多个下列位点(使用 SEQ ID NO: 2 的编号)具有突变: H107, K108, G109, D166, W167, D168, Q169, S170, R171, Q172, F173, Q174, D183, G184, N195, F267, W268, K269, N270, D271, L272, G273, A274, L275, G456, N457, K458, P459, G460, T461, V462, T463。

14. 项 13 的变体, 该变体含有 1 或多个下列取代:

H107A, D, R, N, C, E, Q, G, I, L, K, M, F, P, S, T, W, Y, V;

K108A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;

G109A, D, R, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;

D166A, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;

W167A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, Y, V;

D168A, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;

Q169A, D, R, N, C, E, G, H, I, L, K, M, F, P, S, T, W, Y, V;

S170A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, T, W, Y, V;

R171A, D, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;

Q172A, D, R, N, C, E, G, H, I, L, K, M, F, P, S, T, W, Y, V;

F173A, D, R, N, C, E, Q, G, H, I, L, K, M, P, S, T, W, Y, V;

Q174*, A, D, R, N, C, E, G, H, I, L, K, M, F, P, S, T, W, Y, V;

D183*, A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, W, Y, V;

G184*, A, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;

N195A, D, R, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;

F267A, D, R, N, C, E, Q, G, H, I, L, K, M, P, S, T, W, Y, V;

W268A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, Y, V;

K269A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;

N270A, D, R, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;

D271A, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;

L272A, D, R, N, C, E, Q, G, H, I, K, M, F, P, S, T, W, Y, V;

G273A, D, R, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;

A274D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;

L275A, D, R, N, C, E, Q, G, H, I, K, M, F, P, S, T, W, Y, V;

G456A, D, R, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;

N457A, D, R, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;

K458A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;

P459A, D, R, N, C, E, Q, G, H, I, L, K, M, F, S, T, W, Y, V;

G460A, D, R, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;

T461A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, W, Y, V;

V462A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y;

T463A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, W, Y, V.

15. 项 14 的变体, 其中所述变体含有 1 或多个下列取代或缺失:

Q174*, D183*, G184*, N195F, K269S.

16. 项 13-15 的变体, 其中所述亲本类透阿米尔 α -淀粉酶是具有 SEQ ID NO: 4 所示序列的地衣形芽孢杆菌 α -淀粉酶。

17. 一种 DNA 构建体, 它含有编码项 1-16 任何一项所述 α -淀粉酶变体的 DNA 序列。

18. 一种携带项 17 所述 DNA 构建体的重组表达载体。

19. 一种用项 17 所述 DNA 构建体或项 18 所述载体进行转化的细胞。

20. 项 19 的细胞, 它是一种微生物细胞。

21. 项 20 的细胞, 它是细菌或真菌。

22. 项 21 的细胞, 它是革兰氏阳性细菌, 比如枯草芽孢杆菌, 地衣形芽孢杆菌, 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。

23. 项 1-16 任何一项所述 α -淀粉酶变体用于洗涤和/或清洗餐具的用途。

24. 含有项 1-16 任何一项所述 α -淀粉酶变体的洗涤剂添加剂, 任选该添加剂是非 dusting 颗粒, 稳定化的液体或被保护的酶。

25. 项 24 的洗涤剂添加剂, 它含有 0.02-200mg 酶蛋白/g 添加剂。

26. 项 24 或 25 的洗涤剂添加剂, 它还含有另一种酶, 比如蛋白酶, 脂酶, 过氧化物酶, 另一种淀粉水解酶和/或纤维素酶。

27. 一种包含项 1-16 任何一项所述 α -淀粉酶变体的洗涤剂组合物。

28. 项 27 的洗涤剂组合物, 它另外包含另一种酶, 比如蛋白酶, 脂酶, 过氧化物酶, 另一种淀粉水解酶和/或纤维素酶。

29. 一种手洗或自动餐具洗涤剂组合物, 该组合物包含项 1-16 任何一项所述的 α -淀粉酶变体。

30. 项 29 的清洗餐具的洗涤剂组合物, 该组合物另外包含另一种酶, 比如蛋白酶, 脂酶, 过氧化物酶, 另一种淀粉水解酶和/或纤维素酶。

31. 一种手洗或自动的衣物洗涤组合物, 该组合物包含项 1-16 任何一

项所述的 α -淀粉酶变体。

32. 项 31 的衣物洗涤组合物, 该组合物另外包含另一种酶, 比如蛋白酶, 脂酶, 过氧化物酶, 淀粉水解酶和/或纤维素酶。

33. 制备 α -淀粉酶的方法, 该 α -淀粉酶具有

- 1) 改变了的 pH 最佳值, 和/或
- 2) 改变了的温度最佳值, 和/或
- 3) 稳定性提高,

所述方法包括以下步骤:

- i) 将两个或多个 pH, 温度和/或稳定性曲线迥异的 α -淀粉酶的 3D 结构之分子动力学进行比较, 从而确定(a) α -淀粉酶突变的目标位点和/或区域;
- ii) 在所确定的位点和/或区域进行 1 或多个氨基酸的取代, 添加和/或缺失。

34. 项 33 的方法, 其中将一种中温 α -淀粉酶与一种高温 α -淀粉酶进行比较。

35. 项 33 的方法, 其中将一种低温 α -淀粉酶与一种中温或高温 α -淀粉酶进行比较。

36. 项 33-35 的方法, 其中所述 α -淀粉酶至少 70%, 优选 80%, 达到 90%, 比如达到 95%, 尤其是 95%同源。

37. 项 36 的方法, 其中所比较的 α -淀粉酶是类透阿米尔 α -淀粉酶。

38. 项 28 的方法, 其中所比较的 α -淀粉酶是 SEQ ID NO: 1 到 SEQ ID NO: 8 所示的任何一个 α -淀粉酶。

本发明更进一步涉及编码本发明所述变体的 DNA 构建体; 制备本发明所述变体的方法; 以及本发明所述变体单独地或与其他酶结合, 在各种工业产品或方法中的用途, 例如, 用在洗涤剂中或用于淀粉液化。

本发明最后一个方面涉及生产 α -淀粉酶的方法, 该酶的最适 pH, 和/或最适温度发生改变, 和/或稳定性提高。

命名法则

在本说明书和项书中, 使用了氨基酸残基的常规单字母和三字母编码。为了便于参考, 采用以下命名法则来描述本发明的 α -淀粉酶变体:

原氨基酸: 位点: 取代氨基酸

根据这一命名原则, 例如在第 30 位丙氨酸到天冬酰胺的取代应表示为:

Ala30Asn 或 A30N

在同一位点的丙氨酸缺失表示为:

Ala30* 或 A30*

添加的氨基酸残基, 比如赖氨酸插入表示为:

Ala30 AlaLys 或 A30AK

连续的氨基酸残基片段, 比如 30-33 位氨基酸残基缺失表示为(30-33)* 或 Δ (A30-N33)。

当特定 α -淀粉酶与其他 α -淀粉酶相比含有一个“缺失”, 并且在该位点做了一个插入, 这种情况表示为(对于 36 位插入了一个天冬氨酸):

36Asp 或 36D

多重突变用加号隔开, 即:

Ala30Asp+Glu34Ser 或 A30N+E34S

代表在 30 和 34 位分别将丙氨酸和谷氨酸取代为天冬氨酸和丝氨酸。

当在一个给定位点可以插入 1 或多个可选择的氨基酸残基时, 将其表示为:

A30N, E 或

A30N 或 A30E

另外, 当本文中确定了一个适合进行修饰的位置而没有建议具体的修饰方式, 应当理解为可以用任何氨基酸残基来取代该位点的氨基酸残基。因此, 例如, 如果提及 30 位的丙氨酸修饰, 但没有具体化, 应当理解可以将该丙氨酸缺失或用任何其他氨基酸来取代它, 即下列之一的氨基酸: R, N, D, A, C, Q, E, G, H, I, L, K, M, F, P, S, T, W, Y, V。

附图说明

图 1 是 6 个亲本类透阿米尔 α -淀粉酶的氨基酸序列对比。最左端的数字代表下列各氨基酸序列:

1: SEQ ID NO: 2

2: Kaoamyl

3: SEQ ID NO: 1

4: SEQ ID NO: 5

5: SEQ ID NO: 4

6: SEQ ID NO: 3

图 2 显示 SP722(SEQ ID NO: 2)(pH9)和地衣形芽孢杆菌 α -淀粉酶(SEQ ID NO: 4)(pH7.3)的温度活性关系图。

图3显示 SP690 (SEQ ID NO: 1), SP722(SEQ ID NO: 2), 地衣形芽孢杆菌 α -淀粉酶(SEQ ID NO: 4)在 pH10 的温度分布图。

图4是5个 α -淀粉酶的氨基酸序列对比。最左端的数字代表下列各氨基酸序列:

- 1: amyp-小鼠
- 2: amyp-大鼠
- 3: amyp-猪 猪胰腺 α -淀粉酶(PPA)
- 4: amyp-人
- 5: amy-altha A.haloplanctis α -淀粉酶(AHA)

发明详述

类透阿米尔 α -淀粉酶

众所周知,芽孢杆菌亚种所产生的许多 α -淀粉酶在氨基酸水平上同源性极高。例如,发现包含 SEQ ID NO: 4 所示氨基酸序列的地衣形芽孢杆菌 α -淀粉酶(商品名为透阿米尔™)与包含 SEQ ID NO: 5 所示氨基酸序列的解淀粉芽孢杆菌 α -淀粉酶大约 89%同源,与包含 SEQ ID NO: 3 所示氨基酸序列的嗜热脂肪芽孢杆菌 α -淀粉酶大约 79%同源。其他的同源 α -淀粉酶包括来源于芽孢杆菌菌株 NCIB12289, NCIB12512, NCIB12513 或 DSM9375 的 α -淀粉酶(这些在 WO95/26397 中都有详细的描述)以及 Tsukamoto 等描述过的 α -淀粉酶(生物化学和生物物理研究通讯, 151(1988), 25-31 页, 参见 SEQ ID NO: 6)。

其他同源 α -淀粉酶包括 EP0252666 中描述的地衣形芽孢杆菌(ATCC27811)产生的 α -淀粉酶,以及 WO91/00353 和 WO94/18314 中鉴定到的 α -淀粉酶。在产品 Optitherm™和 Takatherm™(Solvay 有售), Maxamyl™(Gist-brocades/Genencor 有售), Spezym AA™和 Spezym Delta AA™(Genencor 有售)以及 Keistase™ (Daiwa 有售)中含有其他一些商品类透阿米尔地衣形芽孢杆菌 α -淀粉酶。

由于发现这些 α -淀粉酶之间有极大的同源性,可以认为它们属于同类 α -淀粉酶,即“类透阿米尔 α -淀粉酶”类。

因此,在本文中,术语“类透阿米尔 α -淀粉酶”意指这类 α -淀粉酶,它在氨基酸水平上与透阿米尔™(即具有文中 SEQ ID NO: 4 所示氨基酸序列的地衣形芽孢杆菌 α -淀粉酶)有极大的同源性。换句话说,以下所有具有

SEQ ID NO: 1, 2, 3, 4, 5, 6, 7 或 8 所示氨基酸序列, 或者 WO95/26397 中的 SEQ ID NO: 1 所示氨基酸序列(与此处 SEQ ID NO: 7 所示氨基酸序列相同)之 α -淀粉酶或者 WO95/26397 中的 SEQ ID NO: 2 所示氨基酸序列(与此处 SEQ ID NO: 8 所示氨基酸序列相同)或者 Tsukamoto 等文中(1988, 该氨基酸序列在文中见 SEQ ID NO: 6)之 α -淀粉酶都被认为是“类透阿米尔 α -淀粉酶”。其他的类透阿米尔 α -淀粉酶是这样一些 α -淀粉酶: i) 与 SEQ ID NO: 1-8 所示氨基酸序列中的至少一个有至少 60%, 比如至少 70%(例如至少 75%), 或至少 80%(例如至少 85%), 至少 90%或至少 95%同源, 和/或 ii) 与至少一个所述 α -淀粉酶的抗体有免疫交叉反应, 和/或 iii) 由这样一种 DNA 序列编码, 这些序列能与编码以上特定 α -淀粉酶的 DNA 序列进行杂交, 根据本申请的 SEQ ID NO: 9, 10, 11 或 12(这些编码序列分别编码此处 SEQ ID NO: 1, 2, 3, 4 和 5 所示的氨基酸序列), WO95/26397 中的 SEQ ID NO: 4(文中 SEQ ID NO: 13 显示了该 DNA 序列和终止密码子 TAA, 并编码文中 SEQ ID NO: 8 所示的氨基酸序列)以及 WO95/26397 中的 SEQ ID NO: 5(文中 SEQ ID NO: 14 所示)很容易得到这些序列。

对于特性 i), 可以借助任何常规算法来确定“同源性”, 优选利用 GCG 程序包 7.3 版本(1993 年 6 月)中的 GAP 程序, 该程序中 GAP 罚分采用缺省值, 即 GAP 生成罚分为 3.0, GAP 延伸罚分为 0.1(Genetic computer group(1991)GCG 程序包程序手册, 版本 7, 575 Science Drive, Madison, Wisconsin, USA53711)。

可以利用透阿米尔(SEQ ID NO: 4)和类透阿米尔 α -淀粉酶之间的结构对比来鉴定其他类透阿米尔 α -淀粉酶中的等同/相应位点。获得所述结构对比的一个方法是利用 GCG 程序包中的 Pile Up 程序, 该程序中 GAP 罚分采用缺省值, 即 GAP 生成罚分为 3.0, GAP 延伸罚分为 0.1。其他结构对比方法包括疏水簇分析(Gaboriaud 等, (1987), FEBS 快报 224, 149-155 页)和反相成丝技术(reverse threading)(Huber T, Torda, AE, 蛋白质科学, 7 卷, 1 期: 142-149(1998))。

可以利用抗相关类透阿米尔 α -淀粉酶的至少一个表位, 或者与之反应的抗体来检测 α -淀粉酶的特性 ii), 即免疫交叉反应性。可以通过本领域的已知方法, 例如 Hudson 等(实用免疫学, 第 3 版(1989), Blackwell Scientific Publications)描述的方法来制备单克隆或多克隆抗体。可以利用本领域已知的检测方法(这类例子有 Western 印迹或者 Hudson 等(1989)描述的径向免疫

扩散法)来确定免疫交叉反应性。就这方面而言,发现具有 SEQ ID NO: 1, 2, 3, 4, 5, 6, 7 或 8 所示氨基酸序列的 α -淀粉酶之间分别有免疫交叉反应。

可以在所研究的 α -淀粉酶的全长或部分核苷酸或氨基酸序列的基础上适当地制备到用于根据特性 iii) 来鉴定类透阿米尔 α -淀粉酶的寡核苷酸探针。

检测杂交的适宜条件包括在 $5 \times \text{SSC}$ 中预浸泡, 于 $\sim 40^\circ\text{C}$ 在含有 20% 甲酰胺, $5 \times \text{Denhardt's}$ 溶液, 50mM 磷酸钠(pH6.8)以及 50mg 超声变性小牛胸腺 DNA 的溶液中预杂交 1 小时, 然后于 $\sim 40^\circ\text{C}$ 在补充有 100mM ATP 的同一溶液中杂交 18 小时, 随之将滤膜洗 3 次, 每次在 $2 \times \text{SSC}$, 0.2% SDS 中于 40°C (低严格性), 优选 50°C (中等严格性), 更优选 65°C (高严格性), 还要优选的是 $\sim 75^\circ\text{C}$ (极高严格性) 洗 30 分钟。有关杂交方法的其他细节可见 Sambrook 等, 分子克隆: 实验手册, 第 2 版, Cold Spring Harbor, 1989。

在本文中, “来源于” 不仅是指由所研究的微生物菌株产生或能产生的 α -淀粉酶, 也指分离自该菌株的 DNA 序列所编码的, 并在用所述 DNA 序列转化的宿主微生物中产生的 α -淀粉酶。最后, 该术语意指这样的 α -淀粉酶, 它由合成的和/或 cDNA 来源的 DNA 序列编码并且具备所述 α -淀粉酶的鉴定特征。该术语还用来表示亲本 α -淀粉酶可以是天然存在的 α -淀粉酶的变异体, 即由天然存在的 α -淀粉酶的 1 或多个氨基酸残基经过修饰(插入, 取代, 缺失)所产生的变体。

亲本杂交 α -淀粉酶

亲本 α -淀粉酶(即主链 α -淀粉酶)可以是一种杂交 α -淀粉酶, 即包含组合在一起的来源于至少两个 α -淀粉酶的部分氨基酸序列。

亲本杂交 α -淀粉酶可以是这样的酶, 在氨基酸同源性和/或免疫杂交反应性和/或 DNA 杂交(如上所述)的基础上可以确定它属于类透阿米尔 α -淀粉酶家族。在这种情况下, 杂交 α -淀粉酶通常包含类透阿米尔 α -淀粉酶的至少一部分和 1 或多个其他 α -淀粉酶的 1 个或多个部分, 后者选自微生物(细菌或真菌)和/或哺乳动物来源的类透阿米尔 α -淀粉酶或非类透阿米尔 α -淀粉酶。

因此, 亲本杂交 α -淀粉酶可以包括来源于至少两个类透阿米尔 α -淀粉酶, 或者来源于至少一个类透阿米尔 α -淀粉酶和至少一个非类透阿米尔细菌 α -淀粉酶, 或者来源于至少一个类透阿米尔和至少一个真菌 α -淀粉酶的部分氨基酸序列的组合。来源于部分氨基酸序列的类透阿米尔 α -淀粉酶可

以是,例如文中提到的那些特定类透阿米尔 α -淀粉酶中的任何一个。

例如,亲本 α -淀粉酶可以包括来源于地衣形芽孢杆菌菌株的 α -淀粉酶的羧基端部分和来源于解淀粉芽孢杆菌菌株或嗜热脂肪芽孢杆菌菌株的 α -淀粉酶的氨基端部分。例如,亲本 α -淀粉酶可以包括地衣形芽孢杆菌 α -淀粉酶的羧基端部分的至少430个氨基酸残基,还可以,例如,包括a)与具有SEQ ID NO: 5所示氨基酸序列的解淀粉芽孢杆菌 α -淀粉酶的37个氨基端氨基酸残基相对应的氨基酸片段和与具有SEQ ID NO: 4所示氨基酸序列的地衣形芽孢杆菌 α -淀粉酶的445个羧基端氨基酸残基相对应的氨基酸片段,或者杂交类透阿米尔 α -淀粉酶与透阿米尔序列(即SEQ ID NO: 4所示的地衣形芽孢杆菌 α -淀粉酶)相同,除了氨基端的35个氨基酸残基被BAN(成熟蛋白质)(即SEQ ID NO: 5所示的解淀粉芽孢杆菌 α -淀粉酶)的氨基端33个残基(成熟蛋白质)取代;或者包括b)与具有SEQ ID NO: 3所示氨基酸序列的嗜热脂肪芽孢杆菌 α -淀粉酶的68个氨基端氨基酸残基相对应的氨基酸片段和与具有SEQ ID NO: 4所示氨基酸序列的地衣形芽孢杆菌 α -淀粉酶的415个羧基端氨基酸残基相对应的氨基酸片段。

另一合适的亲本杂交 α -淀粉酶是先前在WO96/23874(Novo Nordisk 申请)中描述的酶,该酶由BAN(解淀粉芽孢杆菌 α -淀粉酶)的氨基端(成熟蛋白质的1-300位氨基酸)和透阿米尔的羧基端(成熟蛋白质的301-483位氨基酸)构成。通过将上述杂交 α -淀粉酶(BAN: 1-300/透阿米尔: 301-483)的1或多个以下位点取代使其活性提高: Q360, F290 和 N102。尤其有意义的取代是1或多个以下取代: Q360E, D; F290A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T; N102D, E。

SEQ ID NO: 2 所示 SP722 α -淀粉酶中的相应位点是 S365, Y295, N106 中的1或多个。在SEQ ID NO: 2 所示 α -淀粉酶中尤其有意义的相应取代是 S365D, E; Y295A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T; 以及 N106D, E 中的1或多个。

SEQ ID NO: 1 所示 SP690 α -淀粉酶中的相应位点是 S365, Y295, N106 中的1或多个。尤其有意义的相应取代是 S365D, E; Y295A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T; 以及 N106D, E 中的1或多个。

上面提到的非类透阿米尔 α -淀粉酶可以,例如,是(与类透阿米尔 α -淀粉酶不同的)真菌 α -淀粉酶,哺乳动物或植物 α -淀粉酶或者细菌 α -淀粉酶。这类 α -淀粉酶的具体例子包括米曲霉 TAKA α -淀粉酶,黑曲霉酸 α -

淀粉酶, 枯草芽孢杆菌 α -淀粉酶, 猪胰腺 α -淀粉酶和大麦 α -淀粉酶。所有这些 α -淀粉酶的结构与文中所述的典型类透阿米尔 α -淀粉酶的结构显著不同。

上面提到的真菌 α -淀粉酶, 即来源于黑曲霉和米曲霉, 在氨基酸水平上同源性极高, 通常认为属于同族 α -淀粉酶。来源于米曲霉的真菌 α -淀粉酶以商品名 Fungamyl™ 销售。

另外, 如果类透阿米尔 α -淀粉酶的特定变体(本发明的变体)是指—以常规方式—特定类透阿米尔 α -淀粉酶的氨基酸序列中的特异氨基酸残基发生的修饰(例如缺失或取代), 那么应将其理解为涵盖了在等同位点(由各氨基酸序列之间的最可能氨基酸对比所确定的)进行修饰的另一个类透阿米尔 α -淀粉酶的变体。

在本发明的一个优选实施方案中, α -淀粉酶主链来源于地衣形芽孢杆菌(与亲本类透阿米尔 α -淀粉酶一样), 例如前面提到过的, 比如具有 SEQ ID NO:4 所示氨基酸序列的地衣形芽孢杆菌 α -淀粉酶。

本发明变体的改变特性

以下讨论存在于本发明的变体中的突变与可能由该突变导致的预期性状变化(相对亲本类透阿米尔 α -淀粉酶的这些性状)之间的关系。

在 pH8-10.5 的稳定性提高

在本发明文中, 对于实现在高 pH 下稳定性提高有重要意义的突变(包括氨基酸取代)包括与 SP722 α -淀粉酶(具有 SEQ ID NO:2 所示氨基酸序列)中的 1 或多个下列位点所发生的突变相对应的突变: T141, K142, F143, D144, F145, P146, G147, R148, G149, R181, A186, S193, N195, K269, N270, K311, K458, P459, T461。

本发明的变体具有 1 或多个下列取代(使用 SEQ ID NO: 2 的编号):

T141A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, W, Y, V;

K142A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;

F143A, D, R, N, C, E, Q, G, H, I, L, K, M, P, S, T, W, Y, V;

D144A, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;

F145A, D, R, N, C, E, Q, G, H, I, L, K, M, P, S, T, W, Y, V;

P146A, D, R, N, C, E, Q, G, H, I, L, K, M, F, S, T, W, Y, V;

G147A, D, R, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;

R148A, D, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;

G149A, D, R, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;

K181A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
A186D, R, N, C, E, Q, G, H, I, L, P, K, M, F, S, T, W, Y, V;
S193A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, T, W, Y, V;
N195A, D, R, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
K269A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
N270A, D, R, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
K311A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
K458A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
P459A, D, R, N, C, E, Q, G, H, I, L, K, M, F, S, T, W, Y, V;
T461A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, W, Y, V.

优选的高 pH 稳定性变体包括 SP722 α -淀粉酶(具有 SEQ ID NO: 2 所示氨基酸序列)中的 1 或多个下列取代:

K142R, R181S, A186T, S193P, N195F, K269R, N270Y, K311R, K458R, P459T 和 T461P.

在具体实施方案中, 用具有 SEQ ID NO: 1 所示序列的芽孢杆菌菌株 NCIB12512 α -淀粉酶, 或者具有 SEQ ID NO: 3 所示序列的嗜热脂肪芽孢杆菌 α -淀粉酶, 或者具有 SEQ ID NO: 4 所示序列的地衣形芽孢杆菌 α -淀粉酶, 或者具有 SEQ ID NO: 5 所示序列的解淀粉芽孢杆菌 α -淀粉酶作为主链, 即亲本类透阿米尔 α -淀粉酶来进行这些突变。

从图 1 的对比结果可以看到, 嗜热脂肪芽孢杆菌 α -淀粉酶在与 SP722 中的 N270 对应的位点已有一个酪氨酸。另外, 芽孢杆菌菌株 NCIB12512 α -淀粉酶, 嗜热脂肪芽孢杆菌 α -淀粉酶, 地衣形芽孢杆菌 α -淀粉酶和解淀粉芽孢杆菌 α -淀粉酶在与 SP722 中的 K458 对应的位点早已有精氨酸。另外, 地衣形芽孢杆菌 α -淀粉酶在与 SP722 中的 T461 对应的位点已有一个脯氨酸。因此, 所述取代与这些 α -淀粉酶无关。

可以利用实施例 2 中提及的分子动力学模拟, 通过在所发现的区域进行取代来构建在高 pH 下稳定性提高的 α -淀粉酶变体。模拟描述了那些在高 pH(即 pH8-10.5)比在中等 pH 有更高柔性或可变性的区域。

利用任何与类透阿米尔 α -淀粉酶(BA2, 在 Novo Nodisk 提出的 WO96/23874 的附录 1 中公开了该酶的 3D 结构)同源(定义如下)的细菌 α -淀粉酶的结构, 即有可能建立这些 α -淀粉酶结构的模型并可以对它进行分子动力学模拟。用 GCG 程序包 7.3 版(1993 年 6 月)中的 UWGCG GAP 程序测量到所述细菌 α -淀粉酶的同源性可以为至少 60%, 优选 70% 以上, 更优选 80% 以上, 最优选 90% 以上同源于前面提及的类透阿米尔 α -淀粉酶(BA2), 测量采用缺省值作为 GAP 罚分(Genetic computer group(1991)GCG

程序包程序手册, 版本 7, 575 Science Drive, Madison, Wisconsin, USA53711)。
将不利的残基取代为其他残基是可行的。

在 pH8-10.5 的 Ca^{2+} 稳定性提高

Ca^{2+} 稳定性提高意味着酶在 Ca^{2+} 耗尽的情况下, 稳定性已有提高。在本发明文中, 对于获得在高 pH 下 Ca^{2+} 稳定性提高有重要意义的突变(包括氨基酸取代)包括与 SP722 α -淀粉酶(具有 SEQ ID NO: 2 所示氨基酸序列)中的下列位点相对应的 1 或多个位点突变或缺失: R181, G182, D183, G184, K185, A186, W189, N195, N270, E346, K385, K458, P459。

本发明的变体具有 1 或多个下列取代或缺失:

R181*, A, D, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
G182*, A, D, R, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;
D183*, A, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
G184*, A, R, D, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;
K185A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
A186D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
W189A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, Y, V;
N195A, D, R, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
N270A, R, D, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
E346A, R, D, N, C, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
K385A, R, D, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
K458A, R, D, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
P459A, R, D, N, C, E, Q, G, H, I, L, K, M, F, S, T, W, Y, V.

优选的是具有 1 或多个下列取代或缺失的变体:

R181Q, N; G182T, S, N; D183*; G184*;
K185A, R, D, C, E, Q, G, H, I, L, M, N, F, P, S, T, W, Y, V; A186T, S, N, I, V;
W189T, S, N, Q; N195F, N270R, D; E346Q; K385R; K458R; P459T.

在具体实施方案中, 用具有 SEQ ID NO: 1 所示序列的芽孢杆菌菌株 NCIB12512 α -淀粉酶, 或者具有 SEQ ID NO: 5 所示序列的解淀粉芽孢杆菌 α -淀粉酶, 或者具有 SEQ ID NO: 4 所示序列的地衣形芽孢杆菌 α -淀粉酶作为进行这些突变的主链。

从图 1 的对比结果可看到, 地衣形芽孢杆菌 α -淀粉酶没有与 SP722 中的 D183 和 G184 对应的位点。因此, 所述缺失与该 α -淀粉酶无关。

在一个优选实施方案中, 变体是这样的芽孢杆菌菌株 NCIB12512 α -淀粉酶, 它在 D183 和 G184 缺失, 并有下列取代之一: R181Q, N 和/或 G182T, S, N 和/或 D183*; G184*和/或 K185A, R, D, C, E, Q, G, H, I, L, M, N, F, P, S, T, W, Y, V 和/或 A186T, S, N, I, V 和/或 W189T, S, N, Q 和/或 N195F 和/或 N270R,

D 和/或 E346Q 和/或 K385R 和/或 K458R 和/或 P459T。

在中温的特异活性增强

本发明另外一个方面,对于获得在 10-60℃,优选 20-50℃,尤其是 30-40℃特异活性增强的变体来说,重要的突变包括与 SP722 α -淀粉酶(具有 SEQ ID NO: 2 所示氨基酸序列)中的 1 或多个下列位点相对应的突变:

H107, K108, G109, D166, W167, D168, Q169, S170, R171, Q172, F173, Q174, D183, G184, N195, F267, W268, K269, N270, D271, L272, G273, A274, L275, G456, N457, K458, P459, G460, T461, V462, T463.

本发明的变体具有 1 或多个下列取代:

H107A, D, R, N, C, E, Q, G, I, L, K, M, F, P, S, T, W, Y, V;
K108A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
G109A, D, R, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;
D166A, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
W167A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, Y, V;
D168A, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
Q169A, D, R, N, C, E, G, H, I, L, K, M, F, P, S, T, W, Y, V;
S170A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, T, W, Y, V;
R171A, D, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
Q172A, D, R, N, C, E, G, H, I, L, K, M, F, P, S, T, W, Y, V;
F173A, D, R, N, C, E, Q, G, H, I, L, K, M, P, S, T, W, Y, V;
Q174*, A, D, R, N, C, E, G, H, I, L, K, M, F, P, S, T, W, Y, V;
D183*, A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, W, Y, V;
G184*, A, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
N195A, D, R, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
F267A, D, R, N, C, E, Q, G, H, I, L, K, M, P, S, T, W, Y, V;
W268A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, Y, V;
K269A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;
N270A, D, R, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
D271A, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
L272A, D, R, N, C, E, Q, G, H, I, K, M, F, P, S, T, W, Y, V;
G273A, D, R, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;
A274D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
L275A, D, R, N, C, E, Q, G, H, I, K, M, F, P, S, T, W, Y, V;
G456A, D, R, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;
N457A, D, R, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V;
K458A, D, R, N, C, E, Q, G, H, I, L, M, F, P, S, T, W, Y, V;

P459A, D, R, N, C, E, Q, G, H, I, L, K, M, F, S, T, W, Y, V;
G460A, D, R, N, C, E, Q, H, I, L, K, M, F, P, S, T, W, Y, V;
T461A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, W, Y, V;
V462A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y;
T463A, D, R, N, C, E, Q, G, H, I, L, K, M, F, P, S, W, Y, V.

优选的变体具有 1 或多个下列取代或缺失: Q174*, D183*, G184*, K269S.

在具体实施方案中, 用具有 SEQ ID NO: 4 所示序列的地衣形芽孢杆菌 α -淀粉酶作为进行这些突变的主链。

本发明变体的通用突变: 在中温特异活性增强

特别有意义的氨基酸取代是那些能提高酶活性部位周围的可变性的取代。这可以通过那些能破坏活性部位附近(即优选与构成活性部位的任何残基在 10A 或 8A 或 6A 或 4A 之内)的稳定性相互作用的改变来实现。

这类例子是那些能降低侧链尺寸的突变, 比如

Ala 到 Gly

Val 到 Ala 或 Gly

Ile 或 Leu 到 Val, Ala 或 Gly

Thr 到 Ser

我们希望, 通过导入空腔或者通过能堵塞突变造成的空间的结构重排, 这些突变导致活性部位区域的柔性提高。

可以优选本发明的变体还包含 1 或多个除上面概括过的以外的修饰。因此, 将被修饰的 α -淀粉酶变体某部分存在的 1 或多个脯氨酸取代为非脯氨酸残基可能是有益的, 所述非脯氨酸可以是任何可能的, 天然非脯氨酸残基, 优选是 Ala, Gly, Ser, Thr, Val 或 Leu。

类似地, 可以优选将亲本 α -淀粉酶被修饰的那些氨基酸残基中存在的 1 或多个 Cys 残基取代为非半胱氨酸残基, 比如 Ser, Ala, Thr, Gly, Val 或 Leu。

另外, 可以将本发明的变体进行修饰—作为唯一的修饰或者结合上面概括的任何修饰—以便将与 SEQ ID NO: 4 中的 185-209 位氨基酸片段所对应的氨基酸片段中存在的 1 或多个 Asp 和/或 Glu 分别取代为 Asn 或 Gln。同样有意义的是在类透阿米尔 α -淀粉酶中, 将与 SEQ ID NO: 4 中的 185-209 位氨基酸片段所对应的氨基酸片段中存在的 1 或多个 Lys 残基取代为 Arg。

应当明白, 本发明涵盖了含有两个或多个上述修饰的变体。

另外, 向此处描述的任何变体导入点突变可能是有益的。

活性部位周围可变性提高的 α -淀粉酶变体

可以通过替换靠近底物部位的 1 或多个位点的 1 或多个氨基酸残基来提高本发明 α -淀粉酶变体的可变性。这些位点是(使用 SP722 α -淀粉酶(SEQ ID NO: 2)的编号): V56, K108, D168, Q169, Q172, L201, K269, L272, L275, K446, P459。

因此, 本发明的一个方面涉及在 1 或多个上述位点发生突变的变体。

优选的取代是下列中的 1 或多个:

V56A, G, S, T;

K108A, D, E, Q, G, H, I, L, M, N, S, T, V;

D168A, G, I, V, N, S, T;

Q169A, D, G, H, I, L, M, N, S, T, V;

Q172A, D, G, H, I, L, M, N, S, T, V;

L201A, G, I, V, S, T;

K269A, D, E, Q, G, H, I, L, M, N, S, T, V;

L272A, G, I, V, S, T;

L275A, G, I, V, S, T;

Y295A, D, E, Q, G, H, I, L, M, N, F, S, T, V;

K446A, D, E, Q, G, H, I, L, M, N, S, T, V;

P459A, G, I, L, S, T, V.

在本发明的具体实施方案中, 用具有 SEQ ID NO: 1 所示序列的芽孢杆菌菌株 NCIB12512 α -淀粉酶, 或者具有 SEQ ID NO: 3 所示序列的嗜热脂肪芽孢杆菌 α -淀粉酶, 或者具有 SEQ ID NO: 4 所示序列的地衣形芽孢杆菌 α -淀粉酶, 或者具有 SEQ ID NO: 5 所示序列的解淀粉芽孢杆菌 α -淀粉酶作为进行这些突变的主链。

从图 1 的对比可以看到, 地衣形芽孢杆菌 α -淀粉酶和解淀粉芽孢杆菌 α -淀粉酶在与 SP722 中的 K269 相对应的位点有一个 Glu。另外, 嗜热脂肪芽孢杆菌 α -淀粉酶在与 SP722 中的 K269 相对应的位点有一个 Ser。因此, 所述取代与这些 α -淀粉酶无关。

另外, 从图 1 的对比可以看到, 解淀粉芽孢杆菌 α -淀粉酶在与 SP722 中的 L272 相对应的位点有一个 Ala, 嗜热脂肪芽孢杆菌 α -淀粉酶在与 SP722 中的 L272 相对应的位点有一个 Ile。因此, 所述取代与这些 α -淀粉酶无关。

从图 1 的对比可以看到, 芽孢杆菌菌株 12512 α -淀粉酶在与 SP722 中的 L275 相对应的位点有一个 Ile。因此所述取代与该 α -淀粉酶无关。

从图 1 的对比可以看到, 解淀粉芽孢杆菌 α -淀粉酶在与 SP722 中的

Y295 相对应的位点有一个 Phe。另外，嗜热脂肪芽孢杆菌 α -淀粉酶在与 SP722 中的 Y295 相对应的位点有一个 Asn。因此，所述取代与这些 α -淀粉酶无关。

从图 1 的对比可以看到，地衣形芽孢杆菌 α -淀粉酶和解淀粉芽孢杆菌 α -淀粉酶在与 SP722 中的 K446 相对应的位点有一个 Asn。另外，嗜热脂肪芽孢杆菌 α -淀粉酶在与 SP722 中的 K446 相对应的位点有一个 His。因此，所述取代与这些 α -淀粉酶无关。

从图 1 的对比可以看到，地衣形芽孢杆菌 α -淀粉酶，解淀粉芽孢杆菌 α -淀粉酶以及嗜热脂肪芽孢杆菌 α -淀粉酶在与 SP722 中的 P459 相对应的位点有一个 Ser。另外，芽孢杆菌菌株 12512 α -淀粉酶在与 SP722 中的 P459 相对应的位点有一个 Thr。因此，所述取代与这些 α -淀粉酶无关。

在中温有较高活性的酶的稳定化

本发明另一个实施方案涉及提高低温 α -淀粉酶(例如河豚毒素交替单胞菌(Feller 等, (1994), 欧洲生物化学杂志 222: 441-447))和中温 α -淀粉酶(例如 SP722 和 SP690)的稳定性, 这些酶具有中温活性, 即通常称为嗜冷酶和嗜温酶。对于这一特定酶类, 稳定性应理解为热稳定性或在钙耗竭条件下的稳定性。

通常, 在中温显示高活性的酶在那些对酶有抑制作用的条件(比如温度或钙耗竭)下也会有严重问题。

从而, 本发明的目的是提供在中温显示所需高活性, 同时在轻微抑制条件下不丧失活性的酶。

应当优选稳定化变体于中温测定的活性是初始活性的 100%或以上与 50%之间, 更优选在 100%或以上与 70%之间, 最优选在 100%或以上与 85%之间, 所述初始活性是酶在稳定化前在该特定温度的活性, 并且得到的酶应能在抑制条件下比野生型的酶忍受更长时间的温育。

可以考虑的酶包括例如细菌或真菌来源的 α -淀粉酶。

一例这类低温 α -淀粉酶是分离自河豚毒素交替单胞菌(Feller 等, (1994), 欧洲生物化学杂志 222: 441-447)的酶。已得到了该 α -淀粉酶的晶体结构(Aghajari 等, (1998), 蛋白质科学 7: 564-572)。

河豚毒素交替单胞菌 α -淀粉酶(图 4 所示对比结果的 5)与猪胰腺 α -淀粉酶(PPA)(图 4 所示对比结果的 3)大约 66%同源。PPA 3D 结构是已知的, 并且可从 Brookhaven 数据库中以 1OSE 或 1DHK 的名字获得。根据与其他

更稳定的 α -淀粉酶的同源性,可以使得来自河豚毒素交替单胞菌 α -淀粉酶实现“低温高活性酶”的稳定化,并同时维持所需的中温高活性。

图4显示5个 α -淀粉酶,包括AHA和PPA α -淀粉酶的多重序列对比。使河豚毒素交替单胞菌 α -淀粉酶的稳定性提高的特异突变是:T66P, Q69P, R155P, Q177R, A205P, A232P, L243R, V295P, S315R。

制备 α -淀粉酶变体的方法

几种将突变导入基因的方法是本领域已知的。简要讨论 α -淀粉酶-编码DNA序列之后,将讨论在 α -淀粉酶-编码序列中的特定位点产生突变的方法。

克隆编码 α -淀粉酶的DNA序列

可以利用多种本领域的公知方法,从任何能产生目的 α -淀粉酶的细胞或微生物中分离编码亲本 α -淀粉酶的DNA序列。首先,应当利用得自能产生目的 α -淀粉酶的有机体的染色体DNA或信使RNA来构建基因组DNA和/或cDNA文库。然后,如果 α -淀粉酶的氨基酸序列是已知的,可以合成并合成并用同源的,标记的寡核苷酸来由从目的有机体制备的基因组文库鉴定 α -淀粉酶-编码克隆。另外,可以使用标记的寡核苷酸探针(该探针含有与已知 α -淀粉酶基因同源的序列)作为探针,采用低严谨性杂交和洗涤条件来鉴定 α -淀粉酶-编码克隆。

另一种鉴定 α -淀粉酶-编码克隆的方法包括将基因组DNA片段插入表达载体(比如质粒),用所得基因组DNA文库转化 α -淀粉酶阴性的细菌,以及随后将转化细菌铺到含有 α -淀粉酶底物的琼脂上,从而能鉴定表达 α -淀粉酶的克隆。

可选择地,可以通过已确立的标准方法来合成制备编码酶的DNA序列,例如S.L.Beaucage和M.H.Caruthers(1981)描述的磷酸脒方法或Matthes等(1984)描述的方法。在磷酸脒方法中,在例如自动DNA合成仪中合成寡核苷酸,将其纯化,退火,连接和克隆到合适的载体中。

最后,DNA序列可以是基因组和合成来源混合的,合成和cDNA来源混合的,或者基因组和cDNA来源混合的,其是依照标准技术,将合成的,基因组或cDNA来源的片段连接在一起而制备的(在合适时,对应完整DNA序列各部分的片段)。还可以如US4683202或R.K.Saiki等(1988)描述的,使用特异引物通过聚合酶链反应(PCR)来制备DNA序列。

表达 α -淀粉酶变体

根据本发明,可以利用表达载体将通过上述方法,或者通过本领域已知的任何替代方法制备的编码变体的DNA序列表达为酶的形式,载体通常包括调控序列,该序列编码启动子,操纵子,核糖体结合位点,翻译起始信号以及任选地,阻遏子基因或各种激活子基因。

携带编码本发明所述 α -淀粉酶变体之DNA序列的重组表达载体可以是任何能方便地进行重组DNA操作的载体,且载体的选择通常取决于它将要导入的宿主细胞。因此,载体可以是自主复制的载体,即以染色体外个体存在的载体,其复制独立于染色体的复制,例如质粒,噬菌体或染色体外元件,微小染色体或人工染色体。可选择地,载体可以是这样的,当被导入宿主细胞时,它会整合到宿主基因组中,并与所整合到其中的染色体一起复制。

在所述载体中,DNA序列应当可操纵地连接到合适的启动子序列。该启动子可以是任何在所选宿主细胞中显示转录活性的DNA序列,并可以来源于编码宿主细胞的同源或异源蛋白质的基因。适于引导编码本发明所述 α -淀粉酶变体的DNA序列的转录的合适的启动子序列是在所选宿主细胞中显示转录活性的序列,并可以来源于编码宿主细胞的同源或异源蛋白质的基因。适于引导编码本发明所述 α -淀粉酶变体的DNA序列进行转录(特别是在细菌宿主中)的启动子的例子是大肠杆菌 lac 操纵子的启动子,天蓝色链霉菌琼脂糖酶基因 dagA 的启动子,地衣形芽孢杆菌 α -淀粉酶(amyL)的启动子,嗜热脂肪芽孢杆菌产麦芽淀粉酶基因(amyM)的启动子,解淀粉芽孢杆菌 α -淀粉酶(amyQ)的启动子,枯草芽孢杆菌 xylA 和 xylB 基因等的启动子。在真菌宿主中进行转录时,有用的启动子的例子是来源于这样一些基因的启动子,所述基因编码米曲霉 TAKA 淀粉酶,米赫氏根霉天冬氨酸蛋白酶,黑曲霉中性 α -淀粉酶,黑曲霉酸稳定 α -淀粉酶,黑曲霉葡萄糖淀粉酶,米赫氏根霉脂酶,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。

本发明的表达载体还可以包含合适的转录终止子以及,在真核细胞时的多腺苷酸化序列,它们与编码发明所述 α -淀粉酶变体的DNA序列可操纵地连接在一起。终止和多腺苷酸化序列可以适当地来自与启动子相同的来源。

载体还可以包含能使载体在目的宿主细胞中进行复制的DNA序列。这类序列的例子是质粒 pUC19, pACYC177, pUB110, pE194, pAMB1 和 pIJ702

的复制原点。

载体还可以包含选择标记,例如一个其产物能补偿宿主细胞缺陷的基因,比如来自枯草芽孢杆菌或地衣形芽孢杆菌的 *dal* 基因;或者赋予抗生素抗性(比如氨苄青霉素,卡那霉素,氯霉素或四环素抗性)的基因。另外,载体可以包含曲霉选择标记,比如 *amdS*, *argB*, *niaD* 和 *sC*, 产生潮霉素抗性的标记,或者可以通过共转化(例如 WO91/17243 中描述的)来实现选择。

胞内表达在某些方面可能是有益的,例如用某种细菌作为宿主细胞时,但通常优选表达是胞外的。总之,此处提及的芽孢杆菌 α -淀粉酶包含一个允许被表达的蛋白酶分泌到培养基中的前导区。如果需要,可以方便地通过取代编码该前区的 DNA 序列来将该前区替换为不同的前导区或信号序列。

用于分别将编码 α -淀粉酶变体,启动子,终止子和其他元件的本发明所述 DNA 构建体连接在一起,并将其插入合适的包含复制所用的必要信息之载体的步骤是本领域技术人员所熟知的(参考,例如, Sambrook 等, 分子克隆: 实验指南, 第 2 版, Cold Spring Harbor, 1989)。

本发明的细胞,它包含如上所述的本发明 DNA 构建体或表达载体,可以有效地作为重组制备本发明 α -淀粉酶变体的宿主细胞。可以用本发明编码变体的 DNA 构建体,方便地通过将该构建体(以 1 或多拷贝)整合到宿主染色体中来转化所述细胞。通常认为整合是有益的,因为这样 DNA 序列更可能稳定地保持在细胞中。可以依照常规方法,例如通过同源或异源重组将 DNA 构建体整合到宿主染色体中。可选择地,可以根据宿主细胞的不同类型用上面描述过的表达载体来转化细胞。

本发明的细胞可以是更高等的生物,比如哺乳动物或昆虫的细胞,但优选是微生物细胞,比如细菌或真菌(包括酵母)细胞。

合适的细菌的例子是革兰氏阳性细菌,比如枯草芽孢杆菌,地衣形芽孢杆菌,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,巨大芽孢杆菌,苏云金芽孢杆菌,或者浅青紫链霉菌或鼠灰链霉菌,或者是革兰氏阴性细菌,比如大肠杆菌。可以通过例如原生质体转化或利用竞争细胞以已知方式来实现细菌的转化。

可以有利地从酵母属或裂殖酵母属中选择酵母微生物,例如酿酒酵母。丝状真菌最好是属于曲霉属的种,例如米曲霉或黑曲霉。可以通过一个包

括原生质体形成和转化，以及随后以已知方式再生细胞壁的方法来转化真菌细胞。在 EP238023 中描述了转化曲霉宿主细胞的适用步骤。

再一方面，本发明涉及制备发明所述 α -淀粉酶变体的方法，该方法包括在有助于变体生产的条件下培养以上描述的宿主细胞，以及从细胞和/或培养基中回收变体。

用于培养细胞的培养基可以是任何适合目的宿主细胞生长以及本发明所述 α -淀粉酶变体进行表达的常规培养基。可以从供应商那里获得合适的培养基或者可以根据公开的配方(例如，美国典型培养物保藏中心的目录中所描述的)来制备。

可以通过公知方法从培养基中方便地回收宿主细胞所分泌的 α -淀粉酶变体，这些方法包括通过离心或过滤从培养基中分离细胞，以及借助盐(比如硫酸铵)来沉淀培养基中的蛋白类成分，然后利用层析操作，比如离子交换层析，亲和层析等。

工业应用

本发明的 α -淀粉酶变体具备适合多种工业应用的有价值的特性。具体来说，本发明的酶变体可以作为洗涤，餐具清洗和硬金属表面清洁的洗涤剂组合物的成分。

许多变体在由淀粉生产甜味剂和乙醇，和/或纺织品退浆时尤其有用。在例如 US3912590 和欧洲专利公开号 252730 和 63909 中描述了在包括淀粉液化和/或糖化步骤的常规淀粉转化过程中使用的条件。

洗涤剂组合物

如上所述，可以适当地将本发明的变体加入洗涤剂组合物中。涉及洗涤剂组合物相关成分的进一步的细节，在这类洗涤剂组合物中配制变体的合适方法以及洗涤剂组合物的有关类型，可以参考例如，WO96/23874 和 WO97/07202。

包含本发明变体的洗涤剂组合物可以另外含有 1 或多种其他酶，比如脂酶，角质酶，蛋白酶，纤维素酶，过氧化物酶或漆酶，和/或另一种 α 淀粉酶。

可以将本发明的 α -淀粉酶变体以常规使用的浓度加入洗涤剂中。目前考虑可以将本发明的变体以相当于每升使用常规剂量水平洗涤剂的洗涤/餐具清洗液中含 0.00001-1mg α -淀粉酶的用量加入。

本发明还涉及提供具有 1)改变了的最适 pH，和/或 2)改变了的最适温

度, 和/或 3) 稳定性提高的 α -淀粉酶的方法, 该方法包括以下步骤:

i) 通过将两个或多个实质不同的 pH, 温度和/或稳定性曲线的 α -淀粉酶 3D 结构的分子动力学进行比较, 确定(a)进行 α -淀粉酶突变的位点和/或区域, ii) 在所确定的位点和/或区域取代, 添加和/或缺失 1 或多个氨基酸。

在本发明的实施方案中, 将一个中温 α -淀粉酶与一个高温 α -淀粉酶进行比较。在另一个实施方案中, 将低温 α -淀粉酶与中温或高温 α -淀粉酶进行比较。

进行比较的 α -淀粉酶应当优选相互之间至少 70%, 优选 80%, 到 90%, 比如达到 95%, 特别是 95% 同源。

进行比较的 α -淀粉酶可以是上面定义的一类透阿米尔 α -淀粉酶。在具体实施方案中, 进行比较的 α -淀粉酶是 SEQ ID NO: 1 到 SEQ ID NO: 8 所示的 α 淀粉酶。

在另一个实施方案中, 所比较的目的 α -淀粉酶的稳定性曲线是 Ca^{2+} 依赖曲线。

材料和方法

酶

SP722(SEQ ID NO: 2, Novo Nodisk 有售)

透阿米尔TM(SEQ ID NO: 4, Novo Nodisk 有售)

SP690(SEQ ID NO: 1, Novo Nodisk 有售)

枯草芽孢杆菌 SHA273: 见 WO95/10603

质粒

pJE1 包含编码 SP722 α -淀粉酶变体的基因(SEQ ID NO: 2), 即该基因缺失了成熟蛋白质中氨基酸 D183-G184 所对应的 6 个核苷酸。JE1 基因的转录由 amyL 启动子引导。该质粒还包含复制原点和来自质粒 pUB110(Gryczan, TJ 等, (1978), 细菌杂志 134: 318-329)的 cat-基因, 后者赋予对卡那霉素的抗性。

方法

构建文库载体 pDorK101

可以用大肠杆菌/芽孢杆菌穿梭载体 pDorK101(以下有描述)在大肠杆菌中导入突变, 而不表达 α -淀粉酶, 然后以在芽孢杆菌中 α -淀粉酶具有活性的方式进行修饰。如下构建载体: 在 SEQ ID NO: 2: SP722 的 5' 编码区中

的 PstI 位点, 用一个含有大肠杆菌复制原点的 1.2kb 片段通过基因打断使 pJE1 中的 JE1 编码基因(缺失 D183-G184 的 SP722)失活。所述片段是从 pUC19(GenBank 入册编号#: X02514)经 PCR 扩增到的, 扩增使用正向引物: 5'-gacctgcagtcaggcaacta-3'和反向引物: 5'-tagagtcgacctgcaggcat-3'。用 PstI 于 37℃ 将 PCR 扩增产物和 pJE1 载体消化 2 小时。于室温连接 pJE1 载体片段和 PCR 片段 1 小时, 并通过电转化转化到大肠杆菌中。所得载体命名为 pDorK101。

滤膜筛选检测

可以用该检测方法, 依靠筛选温度设置来筛选较之亲本酶在高 pH 的稳定性提高的类透阿米尔 α -淀粉酶变体, 以及较之亲本酶在高 pH 和中温的稳定性提高的类透阿米尔 α -淀粉酶变体。

高 pH 滤膜检测

将芽孢杆菌文库铺在含有 10 μ g/ml 卡那霉素的 TY 琼脂板上的醋酸纤维素 (OE67, Schleicher&Schuell,Dassel,Germany) 和硝酸纤维素滤膜 (Protran-Ba85, Schleicher&Schuell,Dassel,Germany) 夹心上, 于 37℃ 保持至少 21 小时。将醋酸纤维素层放在 TY 琼脂板上。

铺板之后, 但要在保温之前用针将每个滤膜夹心特异地标记一下, 以便能够确定阳性变体在滤膜上的位置, 并将结合了变体的硝酸纤维素滤膜转移到盛有甘氨酸-NaOH 缓冲液(pH8.6-10.6)的容器中, 于室温(可以在 10-60℃ 变化)温育 15 分钟。将带有菌落的醋酸纤维素滤膜于室温保存在 TY 平板上待用。温育后, 在含有 1% 琼脂糖, 0.2% 淀粉的甘氨酸-氢氧化钠缓冲液(pH8.6-10.6)的平板上检测残存的活性。用与滤膜夹心相同的方式标记带有硝酸纤维素滤膜的检测平板, 并于室温培养 2 小时。移去滤膜后, 用 10% Lugol 溶液将检测平板染色。降解淀粉的变体检测为深蓝色背景上的白色斑点, 然后在保存平板上鉴定。在与第一次筛选相同的条件下将阳性变体重筛两次。

低钙滤膜检测

将芽孢杆菌文库铺在含有关抗生素(例如卡那霉素或氯霉素)的 TY 琼脂板上的醋酸纤维素 (OE67, Schleicher&Schuell,Dassel,Germany) 和硝酸纤维素膜 (Protran-Ba85, Schleicher&Schuell,Dassel,Germany) 夹心上, 于 37℃ 保持至少 21 小时。将醋酸纤维素层放在 TY 琼脂板上。

铺板之后, 但要在培养之前用针将每个滤膜夹心特异地标记一下, 以

便能够将阳性变体固定在滤膜上,将结合了变体的硝酸纤维素滤膜转移到盛有碳酸-碳酸氢盐缓冲液(pH8.5-10)和不同浓度(0.001mM-100mM)EDTA的容器中,将膜于室温温育 1 小时。将带有菌落的醋酸纤维素滤膜于室温保存在 TY 平板上待用。温育后,在含有 1%琼脂糖,0.2%淀粉的碳酸盐-碳酸氢盐缓冲液(pH8.5-10)的平板上检测残存活性。以与滤膜夹心相同的方式标记带有硝酸纤维素滤膜的检测平板,并于室温培养 2 小时。移去滤膜后,用 10%Lugol 溶液将检测平板染色。降解淀粉的变体检测为深蓝色背景上的白色斑点,然后在保存平板上鉴定。在与第一次筛选相同的条件下将阳性变体重筛两次。

获得目的区域的方法

有三个已知的细菌 α -淀粉酶 3D 结构。两个是地衣形芽孢杆菌 α -淀粉酶的, Brookhaven 数据库 1BPL(Machius 等, (1995), 分子生物学杂志 246: 545-559)和 1VJS(Song 等, (1996), 用于糖类 163 工程的酶(生物技术进展, 12 卷)。这两个结构在两个钙离子和一个钠离子的结合位点周围所谓 B-结构域的地方缺少一段重要的结构。因此我们使用了 α -淀粉酶 BA2(WO96/23874, 它是 BANTM(SEQ ID NO: 5)和地衣形芽孢杆菌 α -淀粉酶(SEQ ID NO: 4)的杂交体)的 3D 结构。在该结构的基础上,建立了地衣形芽孢杆菌 α -淀粉酶和 SP722 α -淀粉酶的模式。

α -淀粉酶变体的发酵和纯化

可以利用本领域公知的方法进行发酵和纯化。

稳定性测定

全部稳定性实验使用相同的设置进行。方法是:

在相应条件(1-4)下温育酶。于不同时刻取样品,例如 0, 5, 10, 15 和 30 分钟后,将样品在检测缓冲液(0.1M 50mM Britton 缓冲液 pH7.3)中稀释 25 倍(所有取样稀释度相同),并采用 Phadebas 检测法(Pharmacia)在标准条件(pH7.3, 37℃)下测量活性。

以温育前(0 分钟)测量的活性作为对照(100%)。将下降的百分比作为温育时间的函数来计算活性。表中显示在温育例如 30 分钟后的残存活性。

特异活性测定

采用 Phadebas 检测法(Pharmacia)将特异活性确定为活性/mg 酶。测定依照产品说明书进行(还可见以下“ α -淀粉酶活性检测”)。

α -淀粉酶活性检测法

1. Phadebas 检测法

通过一种采用 Phadebas®片剂作为底物的方法来测定 α -淀粉酶的活性。Phadebas 片剂(Phadebas® Amylase Test,由 Pharmacia Diagnostic 提供)含有一种交联的不溶性兰色淀粉聚合物,该聚合物与牛血清白蛋白和缓冲物质混合在一起并做成片剂。

进行每次测量时,将一个药片悬浮于含有 5ml 50mM Britton-Robinson 缓冲液(50mM 醋酸, 50mM 磷酸, 50mM 硼酸, 0.1mM CaCl_2 , 用 NaOH 调至所需 pH)的试管中。于所需温度在水浴中进行实验。将待测 α -淀粉酶在 xml 50mM Britton-Robinson 缓冲液中进行稀释。将 1ml 该 α -淀粉酶溶液加入 5ml 50mM Britton-Robinson 缓冲液中。 α -淀粉酶水解淀粉产生水溶性兰色碎片。在 620nm 比色测定到的所得兰色溶液的吸光度是 α -淀粉酶活性的函数。

重要的是温育 10 或 15 分钟后测量的 620nm 吸光度应在 0.2 到 2.0 吸收单位内。在这样的吸光度范围内,活性和吸光度是线性关系(Lambert-Beer 定律)。因此,必须调节酶的稀释度以便符合该法则。在特别设置的条件下(温度, pH, 反应时间, 缓冲液条件)下, 1mg 给定 α -淀粉酶将水解一定量的底物并产生兰色。在 620nm 测量兰色的强度。在给定条件下,所测得的吸光度与受测 α -淀粉酶的特异活性直接成比例(活性/毫克纯 α -淀粉酶蛋白质)。

2. 替代方法

通过一种采用 PNP-G7 作为底物的方法来确定 α -淀粉酶的活性。PNP-G7(邻-硝基苯- α , D-麦芽庚糖苷的缩写)是一种被封闭的寡糖,可以被 α -淀粉酶切割。切割之后,试剂盒中包含的 α -葡萄糖苷酶消化底物释放出游离的 PNP 分子,它呈现黄色,因此可以在 $\lambda=405\text{nm}$ (400-420nm)的可见光处比色测量。包含 PNP-G7 底物和 α -葡萄糖苷酶的试剂盒由 Boehringer-Mannheim(cat.No.1054635)制造。

将一瓶底物加入 5ml 缓冲液(BM1442309)来制备底物。将一瓶 α -葡萄糖苷酶(BM1462309)加入 45ml 缓冲液(BM1442309)来制备 α -葡萄糖苷酶。将 5ml α -葡萄糖苷酶与 0.5ml 底物混匀制得工作溶液。

将 20 μl 酶溶液转移到一个 96 孔微量滴定板上并于 25°C 温育来进行检测。加入 200 μl 工作溶液(25°C)。将溶液混匀,预保温 1 分钟,并在 3 分钟内每 15 秒测量一次 OD405nm 处的吸光度。

在给定条件设置下,随时间变化的吸光度曲线的斜率与受测 α -淀粉酶特异活性(活性/毫克酶)直接成比例。

使用 DOPE 程序进行随机诱变的一般方法

可以通过以下步骤进行随机诱变:

1. 在亲本酶中选择目的修饰区域
2. 在所选区域内决定突变位点和非突变位点
3. 决定进行何种突变, 例如相对要构建的变体的预期稳定性和/或性能
4. 选择结构上可行的突变
5. 相对步骤 4 来调整步骤 3 所选的残基
6. 利用合适的预测算法来分析核苷酸分布状态
7. 如果必要, 根据遗传密码的可行性调节所需残基(例如, 考虑遗传密码带来的限制(例如, 为了避免导入终止密码子))(本领域技术人员应意识到某些密码子组合实际上不能用, 需要进行调整)。
8. 制备引物
9. 利用引物进行随机诱变
10. 通过筛选选择具有预期的改善特性的所得 α -淀粉酶变体。

适用于步骤 6 的预测算法是本领域所公知的。Tomandl, D.等(计算机辅助的分子设计杂志, 11(1997), 29-38)描述了一种算法。以下描述另一种 DOPE 算法:

预测算法

“DOPE”程序是一种可用来以一定方式优化三联密码子的核苷酸组成的算法, 从而, 三联密码子能编码与所需氨基酸分布最相似的氨基酸分布。为了评价哪一个可能的分布状态与所需氨基酸分布最相似, 需要一个评分公式。在“DOPE”程序中, 发现以下函数是合适的:

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

其中 X_i 是通过该程序计算得到的氨基酸和氨基酸组的数量, y_i 是程序使用者定义的氨基酸和氨基酸组的希望量(例如, 明确需要导入 20 个氨基酸或终止密码子中的哪一个, 比如体现为一定的百分比(例如, 90%Ala, 3%Ile, 7%Val)), w_i 是由程序使用者定义的重要因素(例如, 取决于将某个特定氨基酸残基插入所需位点的重要性)。N 是 21 加上由程序使用者定义的氨基酸组的数量。用于该函数, 规定 0^0 为 1。

利用 Monte-Carlo 算法(Valleau, J.P. & Whittington, S.G. (1977) 描述过一个例子: A guide to Mont Carlo for statictical mechanics: 1 highways. 《统计力学,

Part A:平衡技术》B.J.Berne 编,New York:Plenum)来找出该函数的最大值。在每次迭代中进行以下步骤:

1. 给每个碱基选择一个新的随机核苷酸组成,其中对于密码子所有三个位置的4种核苷酸 G, A, T, C 之一,目前和新组成之间的绝对差小于或等于 d (参见以下 d 的定义)。

2. 利用以上描述的函数 s 来比较新组成和目前组成的评分。如果新评分高于或等于目前组成的评分,保留新组成,将目前组成改为这个新的。如果新评分小,保留该组成的概率为 $\exp(1000(\text{新评分}-\text{目前评分}))$ 。

一个循环通常包括 1000 次上述迭代,其中的 d 由 1 线性降低至 0。在一个优化过程中,进行 100 或更多次循环。最终得到产生最高评分的核苷酸组成。

实施例

实施例 1

同源构建透阿米尔™的实施例

地衣形芽孢杆菌 α -淀粉酶与其他类透阿米尔 α -淀粉酶的总的同源性很高,相似百分比极高。用 Wisconsin 大学 Genetics Computer Group 的程序 GCG 计算的透阿米尔™与 BSG(具有 SEQ ID NO: 3 的嗜热脂肪芽孢杆菌 α -淀粉酶)和 BAN(具有 SEQ ID NO: 5 的解淀粉芽孢杆菌 α -淀粉酶)的相似率分别为 89%和 78%。与 BAN™和 BSG 相比,TERM 缺失残基 G180 和 K181 这两个残基。与 BAN 和透阿米尔™相比,BSG 在 G371 和 I372 之间缺失 3 个残基。另外与 BAN™和透阿米尔™相比,BSG 在羧基端多出 20 个残基。与 BSG™相比,在氨基端,BAN™少 2 个残基,透阿米尔™少 1 个残基。

在 WO96/23974 附录 1 所公开的结构的基础上,建立地衣形芽孢杆菌(透阿米尔™)和解淀粉芽孢杆菌 α -淀粉酶(BAN™)的结构模型。可以类似地建立其他类透阿米尔 α -淀粉酶(例如文中公开的)的结构。

与用来描述当前结构的 α -淀粉酶相比,透阿米尔™的区别在于它在 178-182 位附近缺少两个残基。为了在模型结构中补偿这一点,用 BIOSYM 的 HOMOLOGY 程序来取代该结构(不仅是结构保守的区域)中等同位点处的残基,缺失位点除外。使透阿米尔™(BAN™)中的 G179(G177)和 K180(K180)之间形成肽键。在模型中只发现有很少几个原子过于接近,这表明求解结构和模型结构之间结构关系密切(因此,后者是有效的)。

然后利用 INSIGHT 程序给这个非常粗略的透阿米尔™结构在与求解结

构相同的坐标处加上所述求解结构中的水(605)和离子(4个钙和1个钠)。这可以在只有很少的重叠下做到,换句话说,即吻合得很好。然后利用200步Steepest下降和600步共轭梯度(Conjugated gradient)对该模型结构进行最小化(参见Brook等,1983,计算化学杂志4:187-217)。然后对最小化了的结构进行分子动力学处理,加热5皮秒,随之平衡至多200皮秒但要在35皮秒以上。用Verlet算法进行动力学,利用偶联至水浴的Behrendsen(Behrendsen等,1984,化学物理学杂志81:3684-3690)保持300K的平衡温度。每皮秒去除旋转和翻译。

实施例2

提取鉴定 pH 稳定性提高和温度活性改变的 α -淀粉酶变体的重要区域的方法

对目的酶(此处是SP722和透阿米尔™)的X-射线结构和/或建立的模型结构进行分子动力学模拟。利用CHARMM程序(得自分子模拟(MSI)程序或其他合适程序,例如DISCOVER(来自MSI)做分子动力学模拟。在真空中进行分子动力学分析,更优选的是包含结晶水,或用包埋在水中(例如水球或水箱)的酶来进行。模拟进行300皮秒或更长时间,例如300-1200皮秒。提取结构中CA碳的各向同性波动,并在结构间进行比较。如果序列有缺失和/或插入,插入来自其他结构的各向同性波动从而使各向同性波动的差异为0。对各向同性波动的解释见CHARMM手册(可从MSI获得)。

可以利用标准电荷在可带电的氨基酸上进行分子动力学模拟。即令Asp和Glu带负电,Lys和Arg带正电。该条件类似大约7的中性pH。为了分析更高或更低pH,可以滴定分子以便使标准的可滴定残基的pKa(通常在pH2-10)发生改变,可滴定残基是Lys,Arg,Asp,Glu,Tyr和His,还有Ser,Thr和Cys也是可滴定的,但此处不考虑在内。此处将由pH导致的电荷改变描述为Asp和Glu在高pH带负电,Arg和Lys不带电。这是模拟10到11左右的pH,此时开始滴定Lys和Arg,因为这些残基正常pKa在9-11左右。

1. 提取鉴定高pH稳定性之 α -淀粉酶变体的重要区域的方法:

对于构建高pH稳定性的变体来说,重要的区域是那些在极端pH显示出最高可变性的区域,即有最高各向同性波动的区域。

通过进行两次分子动力学模拟来鉴定这种区域: i)一次高pH运行,这时将碱性氨基酸Lys和Arg看作是中性的(即非质子化的),酸性氨基酸Asp和Glu带(-1)电荷和 ii) 一次中性pH运行,碱性氨基酸Lys和Arg带有(+1)

净电荷, 酸性氨基酸带有(-1)电荷。

比较这两次运行, 鉴定出在高 pH 比在中性 pH 分析下可变性更高的区域。

导入这样一些残基, 这些残基能提高总的稳定性(例如氢键), 使该区域更刚性(通过突变, 比如用脯氨酸取代或替换甘氨酸残基), 或者提高电荷或它们的相互作用均能够提高酶的高 pH 稳定性。

2. 提取鉴定中温活性提高之 α -淀粉酶变体的区域的方法:

对于构建中温活性提高之变体来说, 重要的区域反映在 SP722 和透阿米尔各向同性波动的差值, 即 SP722 减去透阿米尔 CA 的各向同性波动。选择各向同性波动中具有最高可变性的区域。预计这些区域和残基能提高酶在中温的活性。只有存在可变性正确的特定残基, 才能表达 α -淀粉酶的活性。如果残基的可变性太低, 活性将下降或丧失。

实施例 3

通过定域随机, 预测的诱变来构建在中温 Ca^{2+} 稳定性比亲本酶提高的类透阿米尔 α -淀粉酶变体

为了提高 α -淀粉酶在低钙浓度下的稳定性, 在预选区域进行随机诱变。

区域: 残基:

SAI: R181-W189

利用 DOPE 软件(参见材料与方法)确定出 SAI 区中每个建议变化的峰值密码子, 这些变化能减少终止密码子的数量(见表 1)。计算密码子三个位置处的确切核苷酸分布状态以便给出氨基酸变化的建议数目。针对预测区域中的指定位点进行具体预测从而更可能得到所需残基, 但仍允许其他可能性。

表 1:

每个位点氨基酸残基的分布状态

R181: 72%R, 2%N, 7%Q, 4%H, 4%K, 11%S

G182: 73%G, 13%A, 12%S, 2%T

K185: 95%K, 5%R

A186: 50%A, 4%N, 6%D, 1%E, 1%G, 1%K, 5%S, 31%T

W187: 100%W

D188: 100%D

W189: 92%W, 8%S

将所得预测的寡核苷酸链作为有义链在表 2 中显示: 野生型核苷酸和氨基酸序列以及每个预测位点的核苷酸分布状态。

表 2

位点	181	182	185	186	187	188	189
氨基酸序列	Arg	Gly	Lys	Ala	Thr	Asp	Thr
野生型核苷酸序列	cga	ggt	aaa	gct	tgg	gat	tgg

正向引物(SEQ ID NO: 15)

FSA: 5'-caa aat cgt atc tac aaa ttc 123 456 a7g 8910 tgg gat t11g gaa gta gat
tcg gaa aat-3'

每个预测位点的核苷酸分布状态

- 1: 35%A, 65%C
- 2: 83%G, 17%A
- 3: 63%G, 37%T
- 4: 86%G, 14%A
- 5: 85%G, 15%C
- 6: 50%T, 50%C
- 7: 95%A, 5%G
- 8: 58%G, 37%A, 5%T
- 9: 86%C, 13%A, 1%G
- 10: 83%T, 17%G
- 11: 92%G, 8%C

反向引物(SEQ ID NO: 6)

RSA: 5'-gaa ttt gta gat acg att ttg-3'

随机诱变

利用表2中明显的峰值寡核苷酸(用一个常用术语称为FSA)和用于SA1区域的反向引物RSA以及跨越SacII和DraIII位点的SEQ ID NO: 2: SP722特异引物,经重叠延伸法(Horton等,基因77(1989): 61-68)来制备具有21个碱基对重叠的PCR-文库片段。质粒pJE1作为聚合酶链反应的模板。将PCR片段克隆到大肠杆菌/芽孢杆菌穿梭载体pDork101(见材料与方法)中,该载体能在大肠杆菌中进行诱变,在枯草芽孢杆菌中瞬时表达从而防止大肠杆菌中淀粉酶致死性的累积。将克隆的PCR片段导入大肠杆菌中后,从质粒中切下修饰过的pUC19片段,将启动子和突变了透阿米尔基因物理地连接在一起,在芽孢杆菌中进行表达。

筛选

可以用前面“材料与方法”部分描述的低钙滤膜检测法来筛选文库。

实施例4

构建淀粉酶SEQ ID NO: 1(SP690)的变体

将编码SEQ ID NO: 1之淀粉酶的基因放入WO96/23873中描述的质粒pTVB106中。从该构建体中的amyL启动子开始在枯草芽孢杆菌中表达淀

粉酶。

蛋白质的一个变体是 Δ (T183-G184)+Y243F+Q391E+K444Q。WO96/23873 中描述了该变体的构建过程。

经 Sarkar 和 Sommer(1990, 生物技术 8: 404-407)所描述的 mega-引物法来构建 Δ (T183-G184)+N195F。

利用基因特异引物 B1(SEQ ID NO: 17)和诱变引物 101458(SEQ ID NO: 19)经 PCR 来扩增一个接近 645bp 的 DNA 片段, 该片段来自类 pTVB106 质粒(在编码 SEQ ID NO: 1 之淀粉酶的基因中含有 Δ (T183-G184)突变)。

从琼脂糖凝胶中纯化该 645bp 片段, 并与引物 Y2(SEQ ID NO: 18)一起作为在同一模板上进行的第 2 次 PCR 中的 mega-引物。

用限制酶 BstEII 和 AflIII 消化所得近 1080bp 的片段, 纯化产生的约 510bp 的 DNA 片段, 并将其与用相同酶消化过的类 pTVB106 质粒(在编码 SEQ ID NO: 1 的淀粉酶的基因中含有 Δ (T183-G184)突变)进行连接。用连接产物和氯霉素抗性转化子来转化感受态枯草芽孢杆菌 SHA273(淀粉酶和蛋白酶低)细胞, 并经 DNA 测序检测来证明质粒中正确的突变存在。

引物 B1: (SEQ ID NO: 17)

5' CGA TTG CTG ACG CTG TTA TTT GCG 3'

引物 Y2: (SEQ ID NO: 18)

5' CTT GTT CCC TTG TCA GAA CCA ATG 3'

引物 101458: (SEQ ID NO: 19)

5' GT CAT AGT TGC CGA AAT CTG TAT CGA CTT C 3'

以相似的方法构建变体: Δ (T183-G184)+K185R+A186T, 除了所用诱变引物为 101638 外。

引物 101638: (SEQ ID NO: 20)

5' CC CAG TCC CAC GTA CGT CCC CTG AAT TTA TAT ATT TTG 3'

以相似的方法来构建变体: Δ (T183-G184)+A186T, Δ (T183-G184)+A186I, Δ (T183-G184)+A186S, Δ (T183-G184)+A186N, 除了所用模板为类 pTVB106 质粒(携有变体 Δ (T183-G184)+K185R+A186T), 并用它作为克隆载体。

诱变寡核苷酸(Oligo1)是:

5' CC CAG TCC CAG NTCTTT CCC CTG AAT TTA TAT ATT TTG 3'(SEQ ID NO:21)

N 代表用于合成诱变寡核苷酸(mutagenic oligonucleotide)的 4 种碱基: A, C, G 和 T 的混合物。

测定转化子的序列来确定成熟淀粉酶中 186 位氨基酸的正确密码子。

如下构建变体: Δ (T183-G184)+K185R+A186T+N195F:

用引物 X2 和引物 101458 在类 pTVB106 质粒(携有突变 Δ (T183-G184)+K185R+A186T)上进行 PCR。将所得 DNA 片段与引物 Y2 一起作为在类 pTVB106 质粒(携有突变 Δ (T183-G184)+N195F)上进行的 PCR 中的 mega-引物。用限制性内切酶 Acc65I 和 AflIII 消化第 2 次 PCR 的产物, 并将其克隆至用相同的酶消化过的类 pTVB106 质粒 (Δ (T183-G184)+N195F)。

引物 X2: (SEQ ID NO: 22)

5' GCG TGG ACA AAG TTT GAT TTT CCT G 3'

如下构建变体: Δ (T183-G184)+K185R +A186T+ N195F+ Y243F+ Q391E+K444Q:

用引物 X2(SEQ ID NO: 22)和引物 10+K185R1458(SEQ ID NO: 19)在类 pTVB106 质粒(携有突变 Δ (T183-G184)+K185R+A186T)上做 PCR。将所得 DNA 片段与引物 Y2(SEQ ID NO: 18)一起作为在类 pTVB106 质粒(携有突变 Δ (T183-G184)+Y243F+ Q391E+K444Q)上进行的 PCR 中的 mega-引物。用限制性内切酶 Acc65I 和 AflIII 消化第 2 次 PCR 的产物, 并将其克隆至用相同的酶消化过的类 pTVB106 质粒 (Δ (T183-G184)+Y243F+ Q391E+K444Q)。

实施例 5

构建亲本 SP722 α -淀粉酶(SEQ ID NO: 2)的定点 α -淀粉酶变体

如下所述构建淀粉酶 SEQ ID NO: 2(SP722)的变体。

将编码 SEQ ID NO: 2 之淀粉酶的基因转入 WO96/23873 中描述的质粒 pTVB112 中。从该构建体中的 amyL 启动子开始在枯草芽孢杆菌中表达淀粉酶。

经 Sarkar 和 Sommer(1990, 生物技术 8: 404-407)所描述的 mega-引物法来构建 Δ (D183-G184)+V56I。

利用基因特异引物 DA03 和诱变引物 DA07 经 PCR 来扩增一个接近 820bp 的 DNA 片段, 该片段来自类 pTVB112 质粒(在编码 SEQ ID NO: 2 之 α -淀粉酶的基因中含有 Δ (D183-G184)突变)。

从琼脂糖凝胶中纯化该 820bp 片段, 并与引物 DA01 一起作为在同一模板上进行的第 2 次 PCR 中的 mega-引物。

用限制酶 NgoM I 和 Aat II 消化所得近 920bp 的片段, 纯化产生的约 170bp 的 DNA 片段, 并将其与用相同酶消化过的类 pTVB112 质粒(在编码 SEQ ID NO: 2 所示淀粉酶的基因中含有 Δ (D183-G184)突变)进行连接。用连接产物来转化感受态枯草芽孢杆菌 SHA273(淀粉酶和蛋白酶低)细胞, 并经 DNA 测序检测氯霉素抗性转化子来证明质粒上含有正确的突变。

引物 DA01: (SEQ ID NO: 23)

5' CCTAATGATGGGAATCACTGG 3'

引物 DA03: (SEQ ID NO: 24)

5' GCATTGGATGCTTTTGAACAACCG 3'

引物 DA07: (SEQ ID NO: 25)

5' CGCAAAATGATATCGGGTATGGAGCC 3'

经 Sarkar 和 Sommer(1990, 生物技术 8: 404-407)描述的 mega-引物法来构建变体: Δ (D183-G184)+K108L, Δ (D183-G184)+K108Q, Δ (D183-G184)+K108E, Δ (D183-G184)+K108V。

用引物 DA03 和诱变引物 DA20 在类 pTVB112 质粒(携有突变 Δ (D183-G184))上做 PCR。将所得 DNA 片段与引物 DA01 一起作为在类 pTVB112 质粒(携有突变 Δ (D183-G184))上进行的 PCR 中的 mega-引物。用限制性内切酶 Aat II 和 Mlu I 消化约 920bp 的第 2 次 PCR 的产物, 并将其克隆至用相同的酶消化过的类 pTVB112 质粒(Δ (D183-G184))。

引物 DA20(SEQ ID NO: 26):

5' GTGATGAACCACSWAGGTGGAGCTGATGC 3'

S 代表用于合成诱变寡核苷酸的 2 种碱基: C 和 G 的混合物, W 代表用于合成诱变寡核苷酸的 2 种碱基: A 和 T 的混合物。

测定转化子的序列来确定成熟淀粉酶中 108 位氨基酸的正确密码子。

以相似的方法构建变体: Δ (D183-G184)+D168A, Δ (D183-G184)+D168I, Δ (D183-G184)+D168V, Δ (D183-G184)+D168T, 除了所用诱变引物为 DA14。

引物 DA14(SEQ ID NO: 27):

5' GATGGTGTATGGRYCAATCACGACAATTCC 3'

R 代表用于合成诱变寡核苷酸的 2 种碱基: A 和 G 的混合物, Y 代表用

于合成诱变寡核苷酸的 2 种碱基: C 和 T 的混合物。

测定转化子的序列来确定成熟淀粉酶中 168 位氨基酸的正确密码子。

以相似的方法构建变体: Δ (D183-G184)+Q169N,除了所用诱变引物为 DA15。

引物 DA15(SEQ ID NO: 28):

5' GGTGTATGGGATAACTCACGACAATTCC 3'

以相似的方法构建变体: Δ (D183-G184)+Q169L,除了所用诱变引物为 DA16。

引物 DA16(SEQ ID NO: 29):

5' GGTGTATGGGATCTCTCACGACAATTCC 3'

以相似的方法构建变体: Δ (D183-G184)+Q172N,除了所用诱变引物为 DA17。

引物 DA17(SEQ ID NO: 30):

5' GGGATCAATCACGAAATTTCCAAAATCGTATC 3'

以相似的方法构建变体: Δ (D183-G184)+Q172L,除了所用诱变引物为 DA18。

引物 DA18(SEQ ID NO: 31):

5' GGGATCAATCACGACTCTTCCAAAATCGTATC 3'

以相似的方法构建变体: Δ (D183-G184)+L201I,除了所用诱变引物为 DA06。

引物 DA06(SEQ ID NO: 32):

5' GGAAATTATGATTATATCATGTATGCAGATGTAG 3'

以相似的方法构建变体: Δ (D183-G184)+K269S,除了所用诱变引物为 DA09。

引物 DA09(SEQ ID NO: 33):

5' GCTGAATTTTGGTCGAATGATTAGGTGCC 3'

以相似的方法构建变体: Δ (D183-G184)+K269Q,除了所用诱变引物为 DA11。

引物 DA11(SEQ ID NO: 34):

5' GCTGAATTTTGGTCGAATGATTAGGTGCC 3'

以相似的方法构建变体: Δ (D183-G184)+N270Y,除了所用诱变引物为 DA21。

引物 DA21(SEQ ID NO: 35):

5' GAATTTTGAAGTACGATTAGGTCGG 3'

以相似的方法构建变体: Δ (D183-G184)+L272A, Δ (D183-G184)+L272I, Δ (D183-G184)+L272V, Δ (D183-G184)+L272T, 除了所用诱变引物为 DA12。

引物 DA12(SEQ ID NO: 36):

5' GGAAAACGATRYCGGTGCCTTGGAGAAC 3'

R 代表用于合成诱变寡核苷酸的 2 种碱基: A 和 G 的混合物, Y 代表用于合成诱变寡核苷酸的 2 种碱基: C 和 T 的混合物。

测定转化子的序列来确定成熟淀粉酶中 272 位氨基酸的正确密码子。

以相似的方法构建变体: Δ (D183-G184)+L275A, Δ (D183-G184)+L275I, Δ (D183-G184)+L275V, Δ (D183-G184)+L275T, 除了所用诱变引物为 DA13。

引物 DA13(SEQ ID NO: 37):

5' GATTTAGGTGCCTRYCAGAACTATTTA 3'

R 代表用于合成诱变寡核苷酸的 2 种碱基: A 和 G 的混合物, Y 代表用于合成诱变寡核苷酸的 2 种碱基: C 和 T 的混合物。

测定转化子的序列来确定成熟淀粉酶中 275 位氨基酸的正确密码子。

以相似的方法构建变体: Δ (D183-G184)+Y295E, 除了所用诱变引物为 DA08。

引物 DA08(SEQ ID NO: 38):

5' CCCCCTTCATGAGAATCTTTATAACG 3'

以 Sarkar 和 Sommer(1990, 生物技术 8: 404-407)描述的 mega-引物法来构建变体 Δ (D183-G184)+K446Q。

利用基因特异引物 DA04(该引物在终止密码子下游 214-231bp 退火)和诱变引物 DA10 经 PCR 来扩增一个接近 350bp 的 DNA 片段, 该片段来自类 pTVB112 质粒(在编码 SEQ ID NO: 2 的淀粉酶的基因中含有 Δ (D183-G184)突变)。

将所得 DNA 片段与引物 DA05 一起作为在类 pTVB112 质粒(携有突变 Δ (D183-G184))上进行的 PCR 中的 mega-引物。用限制性内切酶 SnaB I 和 Not I 消化约 460bp 的第 2 次 PCR 产物, 并将其克隆至用相同的酶消化过的类 pTVB106 质粒(携有 Δ (D183-G184)突变)。

引物 DA04(SEQ ID NO: 39):

5' GAATCCGAACCTCATTACACATTCG 3'

引物 DA05(SEQ ID NO: 40):

5' CGGATGGACTCGAGAAGGAAATACCACG 3'

引物 DA10(SEQ ID NO: 41):

5' CGTAGGGCAAAATCAGGCCGGTCAAGTTTGG 3'

以相似的方法构建变体: Δ (D183-G184)+K458R,除了所用诱变引物为 DA22。

引物 DA22(SEQ ID NO: 42):

5' CATAACTGGAAATCGCCCGGGAACAGTTACG 3'

以相似的方法构建变体: Δ (D183-G184)+P459S 和 Δ (D183-G184)+P459T, 除了所用诱变引物为 DA19。

引物 DA19(SEQ ID NO: 43):

5' CTGGAAATAAAWCCGGAACAGTTACG 3'

W 代表用于合成诱变寡核苷酸的 2 种碱基: A 和 T 的混合物。

测定转化子的序列来确定成熟淀粉酶中 495 位氨基酸的正确密码子。

以相似的方法构建变体: Δ (D183-G184)+T461P,除了所用诱变引物为 DA23。

引物 DA23(SEQ ID NO: 44):

5' GGAAATAAACCAGGACCCGTTACGATCAATGC 3'

以相似方法构建变体: Δ (D183-G184)+K142R,除了所用诱变引物为 DA32。

引物 DA32(SEQ ID NO: 45):

5' GAGGCTTGGACTAGGTTTGATTTTCCAG 3'

以相似方法构建变体: Δ (D183-G184)+K269R,除了所用诱变引物为 DA31。

引物 DA31(SEQ ID NO: 46):

5' GCTGAATTTTGGCGCAATGATTAGGTGCC 3'

实施例 6

构建亲本透阿米尔 α -淀粉酶(SEQ ID NO: 4)的定点 α -淀粉酶变体

将编码透阿米尔 α -淀粉酶的 amyL 基因转入 WO96/10603(Novo Nordisk) 中描述的质粒 pDN1528 中。通过 WO97/41213 或上面描述的“mega 引物法”

来构建分别携带所述亲本 α -淀粉酶中 N265R 和 N265D 取代的变体。

诱变寡核苷酸是：

用于 N265R 取代的引物 b11：

5' PCC AGC GCG CCT AGG TCA CGC TGC CAA TAT TCA G (SEQ ID NO:56)

用于 N265D 取代的引物 b12：

5' PCC AGC GCG CCT AGG TCA TCC TGC CAA TAT TCA G (SEQ ID NO:57)

P 代表磷酸根基团。

实施例 7

确定具有 SEQ ID NO: 2 所示氨基酸序列的亲本 α -淀粉酶之变体在碱性 pH 的 pH 稳定性

在这组分析中，利用了纯化的酶样品。用在 100mM CAPS 缓冲液(调至 pH10.5)中的各变体溶液进行测量。将溶液在 75℃温育。

温育 20 和 30 分钟后，用 PNP-G7 检测法(在“材料与方法”部分描述过)测量残存活性。用 Britton Robinson 缓冲液(pH7.3)测量样品中的残存活性。测量残存活性相对 0 分钟同一酶的相应对照溶液的下降值，所述对照溶液未在高 pH 和 75℃温育。

在下表中将初始活性的百分比作为亲本酶(SEQ ID NO: 2)和受测变体的一个函数。

变体	20 分钟后的残存活性	30 分钟后的残存活性
Δ (D183-G184)+M323L	56%	44%
Δ (D183-G184)+M323L+R181S	67%	55%
Δ (D183-G184)+M323L+A186T	62%	50%

在另一组分析中，使用了培养物上清液。用溶在 100mM CAPS 缓冲液(调至 pH10.5)中的各变体溶液进行测量。将溶液在 80℃温育。

温育 30 分钟后，用 Phadebas 检测法(在上述“材料与方法”部分描述过)测量残存活性。用 Britton Robinson 缓冲液(pH7.3)测量样品中的残存活性。测量残存活性相对 0 分钟同一酶的相应对照溶液的下降值，所述对照溶液未在高 pH 和 80℃温育。

在下表中将初始活性的百分比作为亲本酶(SEQ ID NO: 2)和受测变体的一个函数。

变体	30 分钟后的残存活性
Δ (D183-G184)	4%
Δ (D183-G184)+P459T	25%
Δ (D183-G184)+ K458R	31%
Δ (D183-G184)+K311R	10%

实施例 8

确定氨基酸序列如 SEQ ID NO: 1, SEQ ID NO: 2 和 SEQ ID NO: 4 所示的亲本 α -淀粉酶之变体在碱性 pH 的钙稳定性

A: SEQ ID NO: 1 序列之变体的钙稳定性

用溶在 100mM CAPS 缓冲液(调至 pH10.5)中的各变体溶液进行测量,并在(时间 $t = 0$ 时)向其中加入聚磷酸盐至终浓度为 2400ppm。将溶液于 50℃温育。

温育 20 和 30 分钟后,用 PNP-G7 检测法(上面描述过)测量残存活性。用 Britton Robinson 缓冲液(pH7.3)测量样品中的残存活性。测量残存活性相对 0 分钟时同一酶的相应对照溶液的下降值,所述对照溶液未在高 pH 和 50℃温育。

在下表中将初始活性的百分比作为亲本酶(SEQ ID NO: 1)和受测变体的一个函数。

变体	20 分钟后的残存活性	30 分钟后的残存活性
Δ (T183-G184)	32%	19%
Δ (T183-G184)+A186T	36%	23%
Δ (T183-G184)+K185R+A186I	45%	29%
Δ (T183-G184)+A186I	35%	20%
Δ (T183-G184+N195F	44%	n.d.

n.d.=未测

B. SEQ ID NO: 2 序列之变体的钙稳定性

在这组分析中使用了纯化的酶样品。用溶在 100mM CAPS 缓冲液(调至 pH10.5)中的各变体溶液进行测量,并在(时间 $t = 0$ 时)向其中加入聚磷酸盐至终浓度为 2400ppm。将溶液于 50℃温育。

温育 20 和 30 分钟后,用 PNP-G7 检测法(上面描述过)测量残存活性。用 Britton Robinson 缓冲液(pH7.3)测量样品中的残存活性。测量残存活性相

对 0 分钟时同一酶的相应对照溶液的下降值, 所述对照溶液未在高 pH 和 50℃温育。

在下表中将初始活性的百分比作为亲本酶(SEQ ID NO: 2)和受测变体的一个函数。

变体	20 分钟后的残存活性	30 分钟后的残存活性
Δ (D183-G184)+M323L	21%	13%
Δ (D183-G184)+M323L+R181S	32%	19%
Δ (D183-G184)+M323L+A186T	28%	17%
Δ (D183-G184)+M323L+A186R	30%	18%

变体	20 分钟后的残存活性	30 分钟后的残存活性
Δ (D183-G184)	30%	20%
Δ (D183-G184)+N195F	55%	44%

在这组分析中使用了培养物上清液。用溶在 100mM CAPS 缓冲液(调至 pH10.5)中的各变体溶液进行测量, 并在(时间 $t = 0$ 时)向其中加入聚磷酸盐至终浓度为 2400ppm。将溶液于 50℃温育。

温育 30 分钟后, 用 Phadebas 检测法(上面描述过)测量残存活性。用 Britton Robinson 缓冲液(pH7.3)测量样品中的残存活性。测量残存活性相对 0 分钟时同一酶的相应对照溶液的下降值, 所述对照溶液未在高 pH 和 50℃温育。

在下表中将初始活性的百分比作为亲本酶(SEQ ID NO: 2)和受测变体的一个函数。

变体	30 分钟后的残存活性
Δ (D183-G184)	0%
Δ (D183-G184)+P459T	19%
Δ (D183-G184)+K458R	18%
Δ (D183-G184)+T461P	13%
Δ (D183-G184)+E346Q+K385R	4%

C. SEQ ID NO: 4 中序列之变体的钙稳定性

用溶在 100mM CAPS 缓冲液(调至 pH10.5)中的各变体溶液进行测量, 并加入聚磷酸盐(在时间 $t=0$)至终浓度为 2400ppm。将溶液于 60℃温育 20

分钟。

温育 20 分钟后, 用 PNP-G7 检测法(上面描述过)测量残存活性。用 Britton Robinson 缓冲液(pH7.3)测量样品中的残存活性。测量残存活性相对 0 分钟时同一酶的相应对照溶液的下降值, 所述对照溶液未在高 pH 和 60 °C 温育。

在下表中将初始活性的百分比作为亲本酶(SEQ ID NO: 4)和受测变体的一个函数。

变体	20 分钟后的残存活性
透阿米尔(SEQ ID NO: 4)	17%
N265R	28%
N265D	25%

实施例 9

在中温测量氨基酸序列如 SEQ ID NO: 1 所示之 α -淀粉酶的活性

A: SEQ ID NO: 1 中序列的变体的 α -淀粉酶活性

用溶在 50mM Britton Robinson 缓冲液(调节至 pH7.3)中的各变体溶液, 利用 Phadebas 检测法(上面描述过)进行测量。在 37 °C 用 50mM Britton Robinson 缓冲液(pH7.3), 并在 25 °C 用 50mM CAPS 缓冲液(pH10.5)测量样品中的活性。

下表显示亲本酶(SEQ ID NO: 1)和受测变体的温度依赖活性和在 25 °C 的活性相对 37 °C 活性的百分比。

变体	NU/mg 25 °C	NU/mg 37 °C	NU(25 °C)/NU (37 °C)
SP690	1440	35000	4.1%
Δ (T183-G184)	2900	40000	7.3%
Δ (T183-G184)+K269S	1860	12000	15.5%
Δ (Q174)	3830	38000	7.9%

用溶在 50mM Britton Robinson 缓冲液(调至 pH7.3)中的各变体溶液, 采用 Phadebas 检测法(上面描述过)进行另一项测量。于 37 °C 和 50 °C 用 50mM Britton Robinson 缓冲液(pH7.3)测量样品中的活性。

下表显示亲本酶(SEQ ID NO: 1)和受测变体的温度依赖活性以及在 37 °C 的活性相对 50 °C 的活性的百分比。

变体	NU/mg 37 °C	NU/mg 50 °C	NU(37 °C)/NU(50 °C)
SP690(SEQ ID NO: 1)	13090	21669	60%
K269Q	7804	10063	78%

B: SEQ ID NO: 2 中的序列的变体的 α -淀粉酶活性

用溶在 50mM Britton Robinson 缓冲液(调至 pH7.3)中的各变体溶液, 采用 Phadebas 检测法(上面描述过)进行测量。于 25℃ 和 37℃ 用 50mM Britton Robinson 缓冲液(pH7.3)测量样品中的酶活性。

下表显示亲本酶(SEQ ID NO: 2)和受测变体的温度依赖活性以及在 25℃ 活性相对 37℃ 活性的百分比。

变体	NU/mg 25℃	NU/mg 37℃	NU(25℃) /NU(37℃)
Δ (D183-G184)+M323L	3049	10202	30%
Δ (D183-G184)+M323L+R181S	18695	36436	51%

C: SEQ ID NO: 4 中的序列的变体的 α -淀粉酶活性

用溶在 50mM Britton Robinson 缓冲液(调至 pH7.3)中的各变体溶液, 采用 Phadebas 检测法(上述)测量。在 37℃ 用 50mM Britton Robinson 缓冲液(pH7.3), 并在 60℃ 用 50mM CAPS 缓冲液(pH10.5)测量样品中的活性。

下表显示亲本酶(SEQ ID NO: 4)和受测变体的温度依赖活性以及在 37℃ 活性相对 60℃ 活性的百分比。

变体	NU/mg 37℃	NU/mg 60℃	NU(37℃)/NU(60℃)
透阿米尔	7400	4350	170%
Q264S	10000	4650	215%

实施例 10构建亲本杂交体 BAN:1-300/透阿米尔:301-483 α -淀粉酶之变体

质粒 pTVB191 包含编码杂交体 BAN:1-300/透阿米尔:301-483 的基因以及可在枯草芽孢杆菌中使用的复制原点和赋予氯霉素抗性的 cat 基因。

采用 mega-引物法(Sarkar 和 Sommer, 1990), 以质粒 pTVB191 为模板来构建变体 BM4(F290E)。

用引物 p1(SEQ ID NO: 52)和诱变寡核苷酸 bm4(SEQ ID NO: 47)经聚合酶链反应(PCR)在标准条件下扩增一个 444bp 的片段。

从琼脂糖凝胶上纯化该片段, 并与引物 p2(SEQ ID NO: 53)一起作为第 2 次 PCR 的“mega-引物”, 产生一个 531bp 的片段。用限制内切酶 HinD III 和 Tth111T 消化该片段。将这样产生的 389 bp 片段连接到用相同的两种酶切割过的质粒 pTVB191 中。将所得质粒转化至枯草芽孢杆菌 SHA273 中。

通过在含有氯霉素和不溶性淀粉的平板上培养转化子来挑选出氯霉素抗性克隆。用碘蒸汽将平板染色后,选出形成晕圈的克隆,即分离到了表达活性 α -淀粉酶的克隆。经DNA测序来证实所导入的突变。

以类似的方法构建变体 BM5(F290K), BM6(F290A), BM8(Q360E)和 BM11(N102D)。以下给出构建这些变体的细节。

变体: BM5 (F290K)

诱变寡核苷酸: bm5(SEQ ID NO: 48)

引物(第一次 PCR): p1(SEQ ID NO: 52)

所得片段的大小: 444bp

引物(第二次 PCR): p2(SEQ ID NO: 53)

限制内切酶: *HinD* III, *Tth*111I

切割片段的大小: 389bp

变体: BM6 (F290A)

诱变寡核苷酸: bm6(SEQ ID NO: 49)

引物(第一次 PCR): p1(SEQ ID NO: 52)

所得片段的大小: 444bp

引物(第二次 PCR): p2(SEQ ID NO: 53)

限制内切酶: *HinD* III, *Tth*111I

切割片段的大小: 389bp

变体: BM8 (Q360E)

诱变寡核苷酸: bm8(SEQ ID NO: 50)

引物(第一次 PCR): p1(SEQ ID NO: 52)

所得片段的大小: 230bp

引物(第二次 PCR): p2(SEQ ID NO: 53)

限制内切酶: *HinD* III, *Tth*111I

切割片段的大小: 389bp

变体: BM11 (N102D)

诱变寡核苷酸: bm11(SEQ ID NO: 51)

引物(第一次 PCR): p3(SEQ ID NO: 54)

所得片段的大小: 577bp

引物(第二次 PCR): p4(SEQ ID NO: 55)

限制内切酶: *HinD* III, *Pvu* I

切割片段的大小: 576bp

诱变寡核苷酸:

bm4(SEQ ID NO: 47): F290E

引物 5' GTG TTT GAC GTC CCG CTT CAT GAG AAT TTA CAG G

bm5(SEQ ID NO: 48): F290K

引物 5' GTG TTT GAC GTC CCG CTT CAT AAG AAT TTA CAG G

bm6(SEQ ID NO: 49): F290A

引物 5' GTG TTT GAC GTC CCG CTT CAT GCC AAT TTA CAG G

bm8(SEQ ID NO: 50): Q360E

引物 5' AGG GAA TCC GGA TAC CCT GAG GTT TTC TAC GG

bm11(SEQ ID NO: 51): N102D

引物 5' GAT GTG GTT TTG GAT CAT AAG GCC GGC GCT GAT G

其他引物

p1: 5' CTG TTA TTA ATG CCG CCA AAC C(SEQ ID NO:52)

p2: 5' G GAA AAG AAA TGT TTA CGG TTG CG(SEQ ID NO:53)

p3: 5' G AAA TGA AGC GGA ACA TCA AAC ACG(SEQ ID NO:54)

p4: 5' GTA TGA TTT AGG AGA ATT CC(SEQ ID NO:55)

实施例 11

亲本杂交体 BAN:1-300/透阿米尔:301-483 α -淀粉酶变体在碱性 pH 下的 α -淀粉酶活性

用各酶的溶液, 采用 Phadebas 检测法(上面描述过)进行测量。在 50mM Britton Robinson 缓冲液(用 NaOH 调至所需 pH)中于 30℃温育 15 分钟后, 测量活性。

NU/mg 酶						
pH	wt	Q360E	F290A	F290K	F290E	N102D
8.0	5300	7800	8300	4200	6600	6200
9.0	1600	2700	3400	2100	1900	1900

参考文献

- Klein, C.等,生物化学 1992, 31: 8740-8746,
- Mizuno, H.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., 分子生物学杂志(1994)235: 1560-1584
- Lawson, C.L., 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 α -淀粉酶及采用 X-射线晶体学研究的其与底物类似物抑制剂形成的复合体”, 哥本哈根大学化学系 1993
- MacGregor, E.A.,Food Hydrocolloids,1987, 1 卷, 5-6, P.B.Diderichsen 和 L.Christiansen,来自嗜热脂肪芽孢杆菌的产麦芽糖的 α -淀粉酶的克隆, FEMS 微生物快报: 56: 53-60(1988)
- Hudson 等, 实用免疫学, 第3版(1989), Blackwell Scientific Publications,
- Sambrook 等, 分子克隆: 实验指南, 第2版, Cold Spring Harbor,1989
- S.L.Beaucage 和 M.H.Caruthers,四面体快报 22, 1981: 1859-1869
- Matthes 等, EMBO 杂志 3, 1984: 801-805。
- R.K.Saiki 等, 科学 239, 1988: 487-491。
- Morinaga 等,1984, 生物技术 2: 646-639
- Nelson 和 Long, 分析生物化学 180, 1989: 147-151
- Hunkapiller 等, 1984, 自然 310: 105-111
- R. Higuchi, B. Krummel 和 R.K.Saiki(1988)。DNA 片段体外制备和特异突变的一般方法: 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: 209-221。
- Gryczan 等, 1978, 细菌杂志 134: 318-329。
- S.D.Erlich, 1977, Proc. Natl. Acad. Sci. 74: 1680-1682。
- Boel 等, 1990, 生物化学 29: 6244-6249。

序列表

(1) 一般资料:

(i) 申请人:

- (A) 姓名: Novo Nordisk A/S
 (B) 街道: Novo Alle
 (C) 城市: DK-2880 Bagsvaerd
 (E) 国家: 丹麦
 (F) 邮政编码(ZIP): DK-2880
 (G) 电话: +45 44 44 88 88
 (H) 传真: +45 44 49 32 56

(ii) 发明题目: α -淀粉酶变体

(iii) 序列数: 46

(iv) 计算机可读形式:

- (A) 介质类型: 软盘
 (B) 计算机: IBM PC 兼容机
 (C) 操作系统: PC-DOS/MS-DOS
 (D) 软件: Patent In Release #1.0, Version #1.25(EPO)

(2) SEQ ID NO:1 的资料:

(i) 序列特征:

- (A) 长度: 485 个氨基酸
 (B) 类型: 氨基酸
 (C) 链型: 单链
 (D) 拓扑结构: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO:1 :

His	His	Asn	Gly	Thr	Asn	Gly	Thr	Met	Met	Gln	Tyr	Phe	Glu	Trp	Tyr
1				5				10						15	
Leu	Pro	Asn	Asp	Gly	Asn	His	Trp	Asn	Arg	Leu	Arg	Asp	Asp	Ala	Ala
			20					25						30	
Asn	Leu	Lys	Ser	Lys	Gly	Ile	Thr	Ala	Val	Trp	Ile	Pro	Pro	Ala	Trp
		35					40						45		

Lys Gly Thr Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr
 50 55 60
 Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly
 65 70 75 80
 Thr Arg Asn Gln Leu Gln Ala Ala Val Thr Ser Leu Lys Asn Asn Gly
 85 90 95
 Ile Gln Val Tyr Gly Asp Val Val Met Asn His Lys Gly Gly Ala Asp
 100 105 110
 Gly Thr Glu Ile Val Asn Ala Val Glu Val Asn Arg Ser Asn Arg Asn
 115 120 125
 Gln Glu Thr Ser Gly Glu Tyr Ala Ile Glu Ala Trp Thr Lys Phe Asp
 130 135 140
 Phe Pro Gly Arg Gly Asn Asn His Ser Ser Phe Lys Trp Arg Trp Tyr
 145 150 155 160
 His Phe Asp Gly Thr Asp Trp Asp Gln Ser Arg Gln Leu Gln Asn Lys
 165 170 175
 Ile Tyr Lys Phe Arg Gly Thr Gly Lys Ala Trp Asp Trp Glu Val Asp
 180 185 190
 Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp Met
 195 200 205
 Asp His Pro Glu Val Ile His Glu Leu Arg Asn Trp Gly Val Trp Tyr
 210 215 220
 Thr Asn Thr Leu Asn Leu Asp Gly Phe Arg Ile Asp Ala Val Lys His
 225 230 235 240
 Ile Lys Tyr Ser Phe Thr Arg Asp Trp Leu Thr His Val Arg Asn Thr
 245 250 255
 Thr Gly Lys Pro Met Phe Ala Val Ala Glu Phe Trp Lys Asn Asp Leu
 260 265 270
 Gly Ala Ile Glu Asn Tyr Leu Asn Lys Thr Ser Trp Asn His Ser Val
 275 280 285
 Phe Asp Val Pro Leu His Tyr Asn Leu Tyr Asn Ala Ser Asn Ser Gly
 290 295 300
 Gly Tyr Tyr Asp Met Arg Asn Ile Leu Asn Gly Ser Val Val Gln Lys
 305 310 315 320
 His Pro Thr His Ala Val Thr Phe Val Asp Asn His Asp Ser Gln Pro
 325 330 335
 Gly Glu Ala Leu Glu Ser Phe Val Gln Gln Trp Phe Lys Pro Leu Ala
 340 345 350

```

Tyr Ala Leu Val Leu Thr Arg Glu Gln Gly Tyr Pro Ser Val Phe Tyr
    355                      360                      365

Gly Asp Tyr Tyr Gly Ile Pro Thr His Gly Val Pro Ala Met Lys Ser
    370                      375                      380

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

Gln His Asp Tyr Phe Asp His His Asp Ile Ile Gly Trp Thr Arg Glu
                      405                      410                      415

Gly Asn Ser Ser His Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asp
                      420                      425                      430

Gly Pro Gly Gly Asn Lys Trp Met Tyr Val Gly Lys Asn Lys Ala Gly
                      435                      440                      445

Gln Val Trp Arg Asp Ile Thr Gly Asn Arg Thr Gly Thr Val Thr Ile
    450                      455                      460

Asn Ala Asp Gly Trp Gly Asn Phe Ser Val Asn Gly Gly Ser Val Ser
    465                      470                      475                      480

Val Trp Val Lys Gln
                      485

```

(2) SEQ ID NO:2 的资料:

(i) 序列特征:

(A) 长度: 485 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO:2:

```

His His Asn Gly Thr Asn Gly Thr Met Met Gln Tyr Phe Glu Trp His
1          5          10          15

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

Asn Leu Arg Asn Arg Gly Ile Thr Ala Ile Trp Ile Pro Pro Ala Trp
          35          40          45

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

Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly
          65          70          75          80

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

```

Val Gln Val Tyr Gly Asp Val Val Met Asn His Lys Gly Gly Ala Asp
 100 105 110
 Ala Thr Glu Asn Val Leu Ala Val Glu Val Asn Pro Asn Asn Arg Asn
 115 120 125
 Gln Glu Ile Ser Gly Asp Tyr Thr Ile Glu Ala Trp Thr Lys Phe Asp
 130 135 140
 Phe Pro Gly Arg Gly Asn Thr Tyr Ser Asp Phe Lys Trp Arg Trp Tyr
 145 150 155 160
 His Phe Asp Gly Val Asp Trp Asp Gln Ser Arg Gln Phe Gln Asn Arg
 165 170 175
 Ile Tyr Lys Phe Arg Gly Asp Gly Lys Ala Trp Asp Trp Glu Val Asp
 180 185 190
 Ser Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp Met
 195 200 205
 Asp His Pro Glu Val Val Asn Glu Leu Arg Arg Trp Gly Glu Trp Tyr
 210 215 220
 Thr Asn Thr Leu Asn Leu Asp Gly Phe Arg Ile Asp Ala Val Lys His
 225 230 235 240
 Ile Lys Tyr Ser Phe Thr Arg Asp Trp Leu Thr His Val Arg Asn Ala
 245 250 255
 Thr Gly Lys Glu Met Phe Ala Val Ala Glu Phe Trp Lys Asn Asp Leu
 260 265 270
 Gly Ala Leu Glu Asn Tyr Leu Asn Lys Thr Asn Trp Asn His Ser Val
 275 280 285
 Phe Asp Val Pro Leu His Tyr Asn Leu Tyr Asn Ala Ser Asn Ser Gly
 290 295 300
 Gly Asn Tyr Asp Met Ala Lys Leu Leu Asn Gly Thr Val Val Gln Lys
 305 310 315 320
 His Pro Met His Ala Val Thr Phe Val Asp Asn His Asp Ser Gln Pro
 325 330 335
 Gly Glu Ser Leu Glu Ser Phe Val Gln Glu Trp Phe Lys Pro Leu Ala
 340 345 350
 Tyr Ala Leu Ile Leu Thr Arg Glu Gln Gly Tyr Pro Ser Val Phe Tyr
 355 360 365
 Gly Asp Tyr Tyr Gly Ile Pro Thr His Ser Val Pro Ala Met Lys Ala
 370 375 380
 Lys Ile Asp Pro Ile Leu Glu Ala Arg Gln Asn Phe Ala Tyr Gly Thr
 385 390 395 400

Gln His Asp Tyr Phe Asp His His Asn Ile Ile Gly Trp Thr Arg Glu
 405 410 415

Gly Asn Thr Thr His Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asp
 420 425 430

Gly Pro Gly Gly Glu Lys Trp Met Tyr Val Gly Gln Asn Lys Ala Gly
 435 440 445

Gln Val Trp His Asp Ile Thr Gly Asn Lys Pro Gly Thr Val Thr Ile
 450 455 460

Asn Ala Asp Gly Trp Ala Asn Phe Ser Val Asn Gly Gly Ser Val Ser
 465 470 475 480

Ile Trp Val Lys Arg
 485

(2) SEQ ID NO:3 的资料:

(i) 序列特征:

(A) 长度: 514 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO:3 :

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 Ala 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 Ser 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 Asp 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 Met Lys Thr Asn 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 Thr
 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 Val
 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 Trp Ser Ile Thr Thr
485 490 495

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

Ala Trp

(2) SEQ ID NO:4 的资料:

(i) 序列特征:

(A) 长度: 483 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO: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

(2) SEQ ID NO:5 的资料:

(i) 序列特征:

(A) 长度: 480 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO:5:

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
 305 310 315 320
 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
 385 390 395 400
 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
 465 470 475 480

(2) SEQ ID NO:6 的资料:

(i) 序列特征:

(A) 长度: 485 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO:6:

```

His His Asn Gly Thr Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr
1          5          10          15

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

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

Lys Gly Ala Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr
50          55          60

Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly
65          70          75          80

Thr Arg Ser Gln Leu Gln Ala Ala Val Thr Ser Leu Lys Asn Asn Gly
85          90          95

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

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

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

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

His Phe Asp Gly Val Asp Trp Asp Gln Ser Arg Arg Leu Asn Asn Arg
165         170         175

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

Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Ile Asp Met
195         200         205

Asp His Pro Glu Val Val Asn Glu Leu Arg Asn Trp Gly Val Trp Tyr
210         215         220

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

```

```

Ile Lys Tyr Ser Phe Thr Arg Asp Trp Ile Asn His Val Arg Ser Ala
      245                      250                      255

Thr Gly Lys Asn Met Phe Ala Val Ala Glu Phe Trp Lys Asn Asp Leu
      260                      265                      270

Gly Ala Ile Glu Asn Tyr Leu Gln Lys Thr Asn Trp Asn His Ser Val
      275                      280                      285

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

Gly Asn Tyr Asp Met Arg Asn Ile Phe Asn Gly Thr Val Val Gln Arg
      305                      310                      315                      320

His Pro Ser His Ala Val Thr Phe Val Asp Asn His Asp Ser Gln Pro
      325                      330                      335

Glu Glu Ala Leu Glu Ser Phe Val Glu Glu Trp Phe Lys Pro Leu Ala
      340                      345                      350

Tyr Ala Leu Thr Leu Thr Arg Glu Gln Gly Tyr Pro Ser Val Phe Tyr
      355                      360                      365

Gly Asp Tyr Tyr Gly Ile Pro Thr His Gly Val Pro Ala Met Arg Ser
      370                      375                      380

Lys Ile Asp Pro Ile Leu Glu Ala Arg Gln Lys Tyr Ala Tyr Gly Lys
      385                      390                      395                      400

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

Gly Asn Thr Ala His Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asp
      420                      425                      430

Gly Ala Gly Gly Ser Lys Trp Met Phe Val Gly Arg Asn Lys Ala Gly
      435                      440                      445

Gln Val Trp Ser Asp Ile Thr Gly Asn Arg Thr Gly Thr Val Thr Ile
      450                      455                      460

Asn Ala Asp Gly Trp Gly Asn Phe Ser Val Asn Gly Gly Ser Val Ser
      465                      470                      475                      480

Ile Trp Val Asn Lys
      485

```

(2) SEQ ID NO:7 的资料:

(i) 序列特征:

- (A) 长度: 485 个氨基酸
- (B) 类型: 氨基酸
- (C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO:7:

```

His His Asn Gly Thr Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr
1          5          10          15

Leu Pro Asn Asp Gly Asn His Trp Asn Arg Leu Arg Asp Asp Ala Ala
20          25          30

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

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

Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly
65          70          75          80

Thr Arg Asn Gln Leu Gln Ala Ala Val Thr Ser Leu Lys Asn Asn Gly
85          90          95

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

Gly Thr Glu Ile Val Asn Ala Val Glu Val Asn Arg Ser Asn Arg Asn
115         120         125

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

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

His Phe Asp Gly Thr Asp Trp Asp Gln Ser Arg Gln Leu Gln Asn Lys
165         170         175

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

Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp Met
195         200         205

Asp His Pro Glu Val Ile His Glu Leu Arg Asn Trp Gly Val Trp Tyr
210         215         220

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

Ile Lys Tyr Ser Phe Thr Arg Asp Trp Leu Thr His Val Arg Asn Thr
245         250         255

Thr Gly Lys Pro Met Phe Ala Val Ala Glu Phe Trp Lys Asn Asp Leu
260         265         270

Gly Ala Ile Glu Asn Tyr Leu Asn Lys Thr Ser Trp Asn His Ser Val
275         280         285

```



```

Phe Asp Val Pro Leu His Tyr Asn Leu Tyr Asn Ala Ser Asn Ser Gly
 290                      295                      300

Gly Tyr Tyr Asp Met Arg Asn Ile Leu Asn Gly Ser Val Val Gln Lys
305                      310                      315                      320

His Pro Thr His Ala Val Thr Phe Val Asp Asn His Asp Ser Gln Pro
                      325                      330                      335

Gly Glu Ala Leu Glu Ser Phe Val Gln Gln Trp Phe Lys Pro Leu Ala
                      340                      345                      350

Tyr Ala Leu Val Leu Thr Arg Glu Gln Gly Tyr Pro Ser Val Phe Tyr
                      355                      360                      365

Gly Asp Tyr Tyr Gly Ile Pro Thr His Gly Val Pro Ala Met Lys Ser
                      370                      375                      380

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

Gln His Asp Tyr Phe Asp His His Asp Ile Ile Gly Trp Thr Arg Glu
                      405                      410                      415

Gly Asn Ser Ser His Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asp
                      420                      425                      430

Gly Pro Gly Gly Asn Lys Trp Met Tyr Val Gly Lys Asn Lys Ala Gly
                      435                      440                      445

Gln Val Trp Arg Asp Ile Thr Gly Asn Arg Thr Gly Thr Val Thr Ile
                      450                      455                      460

Asn Ala Asp Gly Trp Gly Asn Phe Ser Val Asn Gly Gly Ser Val Ser
465                      470                      475                      480

Val Trp Val Lys Gln
                      485

```

(2) SEQ ID NO:8 的资料:

(i) 序列特征:

- (A) 长度: 485 个氨基酸
- (B) 类型: 氨基酸
- (C) 链型: 单链
- (D) 拓扑结构: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO:8:

```

His His Asn Gly Thr Asn Gly Thr Met Met Gln Tyr Phe Glu Trp His
1                      5                      10                      15

```

Leu Pro Asn Asp Gly Asn His Trp Asn Arg Leu Arg Asp Asp Ala Ser
 20 25 30
 Asn Leu Arg Asn Arg Gly Ile Thr Ala Ile Trp Ile Pro Pro Ala Trp
 35 40 45
 Lys Gly Thr Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr
 50 55 60
 Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly
 65 70 75 80
 Thr Arg Ser Gln Leu Glu Ser Ala Ile His Ala Leu Lys Asn Asn Gly
 85 90 95
 Val Gln Val Tyr Gly Asp Val Val Met Asn His Lys Gly Gly Ala Asp
 100 105 110
 Ala Thr Glu Asn Val Leu Ala Val Glu Val Asn Pro Asn Asn Arg Asn
 115 120 125
 Gln Glu Ile Ser Gly Asp Tyr Thr Ile Glu Ala Trp Thr Lys Phe Asp
 130 135 140
 Phe Pro Gly Arg Gly Asn Thr Tyr Ser Asp Phe Lys Trp Arg Trp Tyr
 145 150 155 160
 His Phe Asp Gly Val Asp Trp Asp Gln Ser Arg Gln Phe Gln Asn Arg
 165 170 175
 Ile Tyr Lys Phe Arg Gly Asp Gly Lys Ala Trp Asp Trp Glu Val Asp
 180 185 190
 Ser Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp Met
 195 200 205
 Asp His Pro Glu Val Val Asn Glu Leu Arg Arg Trp Gly Glu Trp Tyr
 210 215 220
 Thr Asn Thr Leu Asn Leu Asp Gly Phe Arg Ile Asp Ala Val Lys His
 225 230 235 240
 Ile Lys Tyr Ser Phe Thr Arg Asp Trp Leu Thr His Val Arg Asn Ala
 245 250 255
 Thr Gly Lys Glu Met Phe Ala Val Ala Glu Phe Trp Lys Asn Asp Leu
 260 265 270
 Gly Ala Leu Glu Asn Tyr Leu Asn Lys Thr Asn Trp Asn His Ser Val
 275 280 285
 Phe Asp Val Pro Leu His Tyr Asn Leu Tyr Asn Ala Ser Asn Ser Gly
 290 295 300
 Gly Asn Tyr Asp Met Ala Lys Leu Leu Asn Gly Thr Val Val Gln Lys
 305 310 315 320

His	Pro	Met	His	Ala	Val	Thr	Phe	Val	Asp	Asn	His	Asp	Ser	Gln	Pro		
325						330						335					
Gly	Glu	Ser	Leu	Glu	Ser	Phe	Val	Gln	Glu	Trp	Phe	Lys	Pro	Leu	Ala		
340							345					350					
Tyr	Ala	Leu	Ile	Leu	Thr	Arg	Glu	Gln	Gly	Tyr	Pro	Ser	Val	Phe	Tyr		
355								360				365					
Gly	Asp	Tyr	Tyr	Gly	Ile	Pro	Thr	His	Ser	Val	Pro	Ala	Met	Lys	Ala		
370						375					380						
Lys	Ile	Asp	Pro	Ile	Leu	Glu	Ala	Arg	Gln	Asn	Phe	Ala	Tyr	Gly	Thr		
385					390					395					400		
Gln	His	Asp	Tyr	Phe	Asp	His	His	Asn	Ile	Ile	Gly	Trp	Thr	Arg	Glu		
405								410					415				
Gly	Asn	Thr	Thr	His	Pro	Asn	Ser	Gly	Leu	Ala	Thr	Ile	Met	Ser	Asp		
420								425					430				
Gly	Pro	Gly	Gly	Glu	Lys	Trp	Met	Tyr	Val	Gly	Gln	Asn	Lys	Ala	Gly		
435							440					445					
Gln	Val	Trp	His	Asp	Ile	Thr	Gly	Asn	Lys	Pro	Gly	Thr	Val	Thr	Ile		
450						455					460						
Asn	Ala	Asp	Gly	Trp	Ala	Asn	Phe	Ser	Val	Asn	Gly	Gly	Ser	Val	Ser		
465						470					475					480	
Ile Trp Val Lys Arg																	
485																	

(2) SEQ ID NO: 9 的资料:

(i) 序列特征:

(A) 长度: 1455 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: DNA(基因组)

(xi) 序列描述: SEQ ID NO: 9:

CATCATAATG	GAACAAATGG	TACTATGATG	CAATATTTCTG	AATGGTATTT	GCCAAATGAC	60
GGGAATCATT	GGAACAGGTT	GAGGGATGAC	GCAGCTAACT	TAAAGAGTAA	AGGGATAACA	120
GCTGTATGGA	TCCCACCTGC	ATGGAAGGGG	ACTTCCCAGA	ATGATGTAGG	TTATGGAGCC	180
TATGATTTAT	ATGATCTTGG	AGAGTTTAAC	CAGAAGGGGA	CGGTTCGTAC	AAAATATGGA	240
ACACGCAACC	AGCTACAGGC	TGCGGTGACC	TCTTTAAAAA	ATAACGGCAT	TCAGGTATAT	300
GGTGATGTCG	TCATGAATCA	TAAAGGTGGA	GCAGATGGTA	CGGAAATTGT	AAATGCGGTA	360

GAAGTGAATC GGAGCAACCG AAACCAGGAA ACCTCAGGAG AGTATGCAAT AGAAGCGTGG	420
ACAAAGTTTG ATTTTCCTGG AAGAGGAAAT AACCATTCCA GCTTTAAGTG GCGCTGGTAT	480
CATTTTGATG GGACAGATTG GGATCAGTCA CGCCAGCTTC AAAACAAAAT ATATAAATTC	540
AGGGGAACAG GCAAGGCCTG GGACTGGGAA GTCGATACAG AGAATGGCAA CTATGACTAT	600
CTTATGTATG CAGACGTGGA TATGGATCAC CCAGAAGTAA TACATGAACT TAGAACTGG	660
GGAGTGTGGT ATACGAATAC ACTGAACCTT GATGGATTGA GAATAGATGC AGTGAAACAT	720
ATAAAATATA GCTTTACGAG AGATTGGCTT ACACATGTGC GTAACACCAC AGGTAAACCA	780
ATGTTTGCAG TGGCTGAGTT TTGGAAAAAT GACCTTGGTG CAATTGAAAA CTATTTGAAT	840
AAAACAAGTT GGAATCACTC GGTGTTTGAT GTTCCTCTCC ACTATAATTT GTACAATGCA	900
TCTAATAGCG GTGGTTATTA TGATATGAGA AATATTTTAA ATGGTTCTGT GGTGCAAAAA	960
CATCCAACAC ATGCCGTTAC TTTTGTGAT AACCATGATT CTCAGCCCGG GGAAGCATTG	1020
GAATCCTTTG TTCAACAATG GTTTAAACCA CTTGCATATG CATTGGTTCT GACAAGGGAA	1080
CAAGGTTATC CTTCCGTATT TTATGGGGAT TACTACGGTA TCCCAACCCA TGGTGTTCGG	1140
GCTATGAAAT CTAAAATAGA CCCTCTTCTG CAGGCACGTC AACTTTTGC CTATGGTACG	1200
CAGCATGATT ACTTTGATCA TCATGATATT ATCGGTTGGA CAAGAGAGGG AAATAGCTCC	1260
CATCCAAATT CAGGCCTTGC CACCATTATG TCAGATGGTC CAGGTGGTAA CAAATGGATG	1320
TATGTGGGGA AAAATAAAGC GGGACAAGTT TGGAGAGATA TTACCGGAAA TAGGACAGGC	1380
ACCGTCACAA TTAATGCAGA CGGATGGGGT AATTTCTCTG TTAATGGAGG GTCCGTTTCG	1440
GTTTGGGTGA AGCAA	1455

(2) SEQ ID NO: 10 的资料:

(i) 序列特征:

(A) 长度: 1455 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: DNA(基因组)

(xi) 序列描述: SEQ ID NO: 10:

CATCATAATG GGACAAATGG GACGATGATG CAATACTTTG AATGGCACTT GCCTAATGAT	60
GGGAATCACT GGAATAGATT AAGAGATGAT GCTAGTAATC TAAGAAATAG AGGTATAACC	120
GCTATTTTGA TTCCGCCTGC CTGGAAAGGG ACTTCGCAA ATGATGTGGG GTATGGAGCC	180

TATGATCTTT ATGATTTAGG GGAATTTAAT CAAAAGGGGA CGGTTCGTAC TAAGTATGGG	240
ACACGTAGTC AATTGGAGTC TGCCATCCAT GCTTTAAAGA ATAATGGCGT TCAAGTTTAT	300
GGGGATGTAG TGATGAACCA TAAAGGAGGA GCTGATGCTA CAGAAAACGT TCTTGCTGTC	360
GAGGTGAATC CAAATAACCG GAATCAAGAA ATATCTGGGG ACTACACAAT TGAGGCTTGG	420
ACTAAGTTTG ATTTTCCAGG GAGGGGTAAT ACATACTCAG ACTTTAAATG GCGTTGGTAT	480
CATTTTCGATG GTGTAGATTG GGATCAATCA CGACAATTCC AAAATCGTAT CTACAAATTC	540
CGAGGTGATG GTAAGGCATG GGATTGGGAA GTAGATTCGG AAAATGGAAA TTATGATTAT	600
TTAATGTATG CAGATGTAGA TATGGATCAT CCGGAGGTAG TAAATGAGCT TAGAAGATGG	660
GGAGAATGGT ATACAAATAC ATTAAATCTT GATGGATTTA GGATCGATGC GGTGAAGCAT	720
ATTAAATATA GCTTTACACG TGATTGGTTG ACCCATGTAA GAAACGCAAC GGGAAAAGAA	780
ATGTTTGCTG TTGCTGAATT TTGGAAAAAT GATTTAGGTG CCTTGGAGAA CTATTTAAAT	840
AAAACAACT GGAATCATTG TGTCTTTGAT GTCCCCCTTC ATTATAATCT TTATAACGCG	900
TCAAATAGTG GAGGCAACTA TGACATGGCA AAACCTCTTA ATGGAACGGT TGTTCAAAAG	960
CATCCAATGC ATGCCGTAAC TTTTGTGGAT AATCACGATT CTCAACCTGG GGAATCATTG	1020
GAATCATTTG TACAAGAATG GTTTAAGCCA CTTGCTTATG CGCTTATTTT AACAAAGAGAA	1080
CAAGGCTATC CCTCTGTCTT CTATGGTGAC TACTATGGAA TTCCAACACA TAGTGTCCCA	1140
GCAATGAAAG CCAAGATTGA TCCAATCTTA GAGGCGCGTC AAAATTTTGC ATATGGAACA	1200
CAACATGATT ATTTTGACCA TCATAATATA ATCGGATGGA CACGTGAAGG AAATACCACG	1260
CATCCCAATT CAGGACTTGC GACTATCATG TCGGATGGGC CAGGGGGAGA GAAATGGATG	1320
TACGTAGGGC AAAATAAAGC AGGTCAAGTT TGGCATGACA TAACTGGAAA TAAACCAGGA	1380
ACAGTTACGA TCAATGCAGA TGGATGGGCT AATTTTTCAG TAAATGGAGG ATCTGTTTCC	1440
ATTTGGGTGA AACGA	1455

(2) SEQ ID NO: 11 的资料:

(i) 序列特征:

- (A) 长度: 1548 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线性

(ii) 分子类型: DNA(基因组)

(xi) 序列描述: SEQ ID NO: 11:

GCCGCACCGT TTAACGGCAC CATGATGCAG TATTTTGAAT GGTACTTGCC GGATGATGGC	60
ACGTTATGGA CCAAAGTGGC CAATGAAGCC AACAACTTAT CCAGCCTTGG CATCACCGCT	120
CTTTGGCTGC CGCCCCGCTTA CAAAGGAACA AGCCGCAGCG ACGTAGGGTA CGGAGTATAC	180
GACTTGATG ACCTCGGCGA ATTCAATCAA AAAGGGACCG TCCGCACAAA ATACGGAACA	240
AAAGCTCAAT ATCTTCAAGC CATTCAAGCC GCCCAGCCG CTGGAATGCA AGTGTACGCC	300
GATGTCGTGT TCGACCATAA AGGCGGCGCT GACGGCACGG AATGGGTGGA CGCCGTCGAA	360
GTCAATCCGT CCGACCGCAA CCAAGAAATC TCGGGCACCT ATCAAATCCA AGCATGGACG	420
AAATTTGATT TTCCCGGGCG GGGCAACACC TACTCCAGCT TTAAGTGGCG CTGGTACCAT	480
TTTGACGGCG TTGATTGGGA CGAAAGCCGA AAATTGAGCC GCATTTACAA ATTCCGCGGC	540
ATCGGCAAAG CGTGGGATTG GGAAGTAGAC ACGGAAAACG GAAACTATGA CTACTTAATG	600
TATGCCGACC TTGATATGGA TCATCCCGAA GTCGTGACCG AGCTGAAAAA CTGGGGGAAA	660
TGGTATGTCA ACACAACGAA CATTGATGGG TTCCGGCTTG ATGCCGTCAA GCATATTAAG	720
TTCAGTTTTT TTCCTGATTG GTTGTCTGAT GTGCGTTCTC AGACTGGCAA GCCGCTATTT	780
ACCGTCGGGG AATATTGGAG CTATGACATC AACAGTTGC ACAATTACAT TACGAAAACA	840
GACGGAACGA TGTCTTTGTT TGATGCCCCG TTACACAACA AATTTTATAC CGCTTCCAAA	900
TCAGGGGGCG CATTTGATAT GCGCACGTTA ATGACCAATA CTCTCATGAA AGATCAACCG	960
ACATTGGCCG TCACCTTCGT TGATAATCAT GACACCGAAC CCGGCCAAGC GCTGCAGTCA	1020
TGGGTCGACC CATGGTTCAA ACCGTTGGCT TACGCCTTTA TTCTAACTCG GCAGGAAGGA	1080
TACCCGTGCG TCTTTTATGG TGA CTATTAT GGCATTCCAC AATATAACAT TCCTTCGCTG	1140
AAAAGCAAAA TCGATCCGCT CCTCATCGCG CGCAGGGATT ATGCTTACGG AACGCAACAT	1200
GATTATCTTG ATCACTCCGA CATCATCGGG TGGACAAGGG AAGGGGGCAC TGAAAAACCA	1260
GGATCCGGAC TGGCCGCACT GATCACCGAT GGGCCGGGAG GAAGCAAATG GATGTACGTT	1320
GGCAAACAAC ACGCTGGAAA AGTGTTCTAT GACCTTACCG GCAACCGGAG TGACACCGTC	1380
ACCATCAACA GTGATGGATG GGGGGAATTC AAAGTCAATG GCGGTTCGGT TTCGGTTTGG	1440
GTTCTAGAA AAACGACCGT TTCTACCATC GTCGGCCGA TCACAACCCG ACCGTGGACT	1500
GGTGAATTCG TCCGTTGGAC CGAACCACGG TTGGTGGCAT GGCCTTGA	1548

(2) SEQ ID NO: 12 的资料:

(i) 序列特征:

(A) 长度: 1920 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: DNA(基因组)

(ix)特性:

(A)名称/关键词: CDS

(B)位置: 421...1872

(xi) 序列描述: SEQ ID NO: 12:

CGGAAGATTG GAAGTACAAA AATAAGCAAA AGATTGTCAA TCATGTCATG AGCCATGCGG	60
GAGACGGAAA AATCGTCTTA ATGCACGATA TTTATGCAAC GTTCGCAGAT GCTGCTGAAG	120
AGATTATTAA AAAGCTGAAA GCAAAAGGCT ATCAATTGGT AACTGTATCT CAGCTTGAAG	180
AAGTGAAGAA GCAGAGAGGC TATTGAATAA ATGAGTAGAA GCGCCATATC GGCGCTTTTC	240
TTTTGGAAGA AAATATAGGG AAAATGGTAC TTGTTAAAAA TTCGGAATAT TTATACAACA	300
TCATATGTTT CACATTGAAA GGGGAGGAGA ATCATGAAAC AACAAAAACG GCTTTACGCC	360
CGATTGCTGA CGCTGTTATT TGCCTCATC TTCTTGCTGC CTCATTCTGC AGCAGCGGCG	420
GCA AAT CTT AAT GGG ACG CTG ATG CAG TAT TTT GAA TGG TAC ATG CCC	468
AAT GAC GGC CAA CAT TGG AGG CGT TTG CAA AAC GAC TCG GCA TAT TTG	516
GCT GAA CAC GGT ATT ACT GCC GTC TGG ATT CCC CCG GCA TAT AAG GGA	564
ACG AGC CAA GCG GAT GTG GGC TAC GGT GCT TAC GAC CTT TAT GAT TTA	612
GGG GAG TTT CAT CAA AAA GGG ACG GTT CGG ACA AAG TAC GGC ACA AAA	660
GGA GAG CTG CAA TCT GCG ATC AAA AGT CTT CAT TCC CGC GAC ATT AAC	708
GTT TAC GGG GAT GTG GTC ATC AAC CAC AAA GGC GGC GCT GAT GCG ACC	756
GAA GAT GTA ACC GCG GTT GAA GTC GAT CCC GCT GAC CGC AAC CGC GTA	804
ATT TCA GGA GAA CAC CTA ATT AAA GCC TGG ACA CAT TTT CAT TTT CCG	852
GGG CGC GGC AGC ACA TAC AGC GAT TTT AAA TGG CAT TGG TAC CAT TTT	900
GAC GGA ACC GAT TGG GAC GAG TCC CGA AAG CTG AAC CGC ATC TAT AAG	948
TTT CAA GGA AAG GCT TGG GAT TGG GAA GTT TCC AAT GAA AAC GGC AAC	996
TAT GAT TAT TTG ATG TAT GCC GAC ATC GAT TAT GAC CAT CCT GAT GTC	1044
GCA GCA GAA ATT AAG AGA TGG GGC ACT TGG TAT GCC AAT GAA CTG CAA	1092
TTG GAC GGT TTC CGT CTT GAT GCT GTC AAA CAC ATT AAA TTT TCT TTT	1140
TTG CGG GAT TGG GTT AAT CAT GTC AGG GAA AAA ACG GGG AAG GAA ATG	1188

TTT ACG GTA GCT GAA TAT TGG CAG AAT GAC TTG GGC GCG CTG GAA AAC	1236
TAT TTG AAC AAA ACA AAT TTT AAT CAT TCA GTG TTT GAC GTG CCG CTT	1284
CAT TAT CAG TTC CAT GCT GCA TCG ACA CAG GGA GGC GGC TAT GAT ATG	1332
AGG AAA TTG CTG AAC GGT ACG GTC GTT TCC AAG CAT CCG TTG AAA TCG	1380
GTT ACA TTT GTC GAT AAC CAT GAT ACA CAG CCG GGG CAA TCG CTT GAG	1428
TCG ACT GTC CAA ACA TGG TTT AAG CCG CTT GCT TAC GCT TTT ATT CTC	1476
ACA AGG GAA TCT GGA TAC CCT CAG GTT TTC TAC GGG GAT ATG TAC GGG	1524
ACG AAA GGA GAC TCC CAG CGC GAA ATT CCT GCC TTG AAA CAC AAA ATT	1572
GAA CCG ATC TTA AAA GCG AGA AAA CAG TAT GCG TAC GGA GCA CAG CAT	1620
GAT TAT TTC GAC CAC CAT GAC ATT GTC GGC TGG ACA AGG GAA GGC GAC	1668
AGC TCG GTT GCA AAT TCA GGT TTG GCG GCA TTA ATA ACA GAC GGA CCC	1716
GGT GGG GCA AAG CGA ATG TAT GTC GGC CGG CAA AAC GCC GGT GAG ACA	1764
TGG CAT GAC ATT ACC GGA AAC CGT TCG GAG CCG GTT GTC ATC AAT TCG	1812
GAA GGC TGG GGA GAG TTT CAC GTA AAC GGC GGG TCG GTT TCA ATT TAT	1860
GTT CAA AGA TAG AAGAGCAGAG AGGACGGATT TCCTGAAGGA AATCCGTTTT	1912
TTTATTTT	1920

(2) SEQ ID NO: 12 的资料:

(i) 序列特征:

(A) 长度: 2084 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: DNA(基因组)

(ix) 特性:

(A) 名称/关键词: CDS

(B) 位置: 343...1794

(xi) 序列描述: SEQ ID NO: 12:

GCCCCGCACA TACGAAAAGA CTGGCTGAAA ACATTGAGCC TTTGATGACT GATGATTGG	60
CTGAAGAAGT GGATCGATTG TTTGAGAAAA GAAGAAGACC ATAAAAATAC CTTGTCTGTC	120
ATCAGACAGG GTATTTTTTA TGCTGTCCAG ACTGTCCGCT GTGTAAAAAT AAGGAATAAA	180

GGGGGGTTGT TATTATTTTA CTGATATGTA AAATATAATT TGTATAAGAA AATGAGAGGG	240
AGAGGAAACA TGATTCAAAA ACGAAAGCGG ACAGTTTCGT TCAGACTTGT GCTTATGTGC	300
ACGCTGTTAT TTGTCAGTTT GCCGATTACA AAAACATCAG CC GTA AAT GGC ACG	354
CTG ATG CAG TAT TTT GAA TGG TAT ACG CCG AAC GAC GGC CAG CAT TGG	402
AAA CGA TTG CAG AAT GAT GCG GAA CAT TTA TCG GAT ATC GGA ATC ACT	450
GCC GTC TGG ATT CCT CCC GCA TAC AAA GGA TTG AGC CAA TCC GAT AAC	498
GGA TAC GGA CCT TAT GAT TTG TAT GAT TTA GGA GAA TTC CAG CAA AAA	546
GGG ACG GTC AGA ACG AAA TAC GGC ACA AAA TCA GAG CTT CAA GAT GCG	594
ATC GGC TCA CTG CAT TCC CGG AAC GTC CAA GTA TAC GGA GAT GTG GTT	642
TTG AAT CAT AAG GCT GGT GCT GAT GCA ACA GAA GAT GTA ACT GCC GTC	690
GAA GTC AAT CCG GCC AAT AGA AAT CAG GAA ACT TCG GAG GAA TAT CAA	738
ATC AAA GCG TGG ACG GAT TTT CGT TTT CCG GGC CGT GGA AAC ACG TAC	786
AGT GAT TTT AAA TGG CAT TGG TAT CAT TTC GAC GGA GCG GAC TGG GAT	834
GAA TCC CGG AAG ATC AGC CGC ATC TTT AAG TTT CGT GGG GAA GGA AAA	882
GCG TGG GAT TGG GAA GTA TCA AGT GAA AAC GGC AAC TAT GAC TAT TTA	930
ATG TAT GCT GAT GTT GAC TAC GAC CAC CCT GAT GTC GTG GCA GAG ACA	978
AAA AAA TGG GGT ATC TGG TAT GCG AAT GAA CTG TCA TTA GAC GGC TTC	1026
CGT ATT GAT GCC GCC AAA CAT ATT AAA TTT TCA TTT CTG CGT GAT TGG	1074
GTT CAG GCG GTC AGA CAG GCG ACG GGA AAA GAA ATG TTT ACG GTT GCG	1122
GAG TAT TGG CAG AAT AAT GCC GGG AAA CTC GAA AAC TAC TTG AAT AAA	1170
ACA AGC TTT AAT CAA TCC GTG TTT GAT GTT CCG CTT CAT TTC AAT TTA	1218
CAG GCG GCT TCC TCA CAA GGA GGC GGA TAT GAT ATG AGG CGT TTG CTG	1266
GAC GGT ACC GTT GTG TCC AGG CAT CCG GAA AAG GCG GTT ACA TTT GTT	1314
GAA AAT CAT GAC ACA CAG CCG GGA CAG TCA TTG GAA TCG ACA GTC CAA	1362
ACT TGG TTT AAA CCG CTT GCA TAC GCC TTT ATT TTG ACA AGA GAA TCC	1410
GGT TAT CCT CAG GTG TTC TAT GGG GAT ATG TAC GGG ACA AAA GGG ACA	1458
TCG CCA AAG GAA ATT CCC TCA CTG AAA GAT AAT ATA GAG CCG ATT TTA	1506
AAA GCG CGT AAG GAG TAC GCA TAC GGG CCC CAG CAC GAT TAT ATT GAC	1554

CAC CCG GAT GTG ATC GGA TGG ACG AGG GAA GGT GAC AGC TCC GCC GCC	1602
AAA TCA GGT TTG GCC GCT TTA ATC ACG GAC GGA CCC GGC GGA TCA AAG	1650
CGG ATG TAT GCC GGC CTG AAA AAT GCC GGC GAG ACA TGG TAT GAC ATA	1698
ACG GGC AAC CGT TCA GAT ACT GTA AAA ATC GGA TCT GAC GGC TGG GGA	1746
GAG TTT CAT GTA AAC GAT GGG TCC GTC TCC ATT TAT GTT CAG AAA TAA	1794
GGTAATAAAA AAACACCTCC AAGCTGAGTG CGGGTATCAG CTTGGAGGTG CGTTTATTTT	1854
TTCAGCCGTA TGACAAGGTC GGCATCAGGT GTGACAAATA CGGTATGCTG GCTGTCATAG	1914
GTGACAAATC CGGGTTTTGC GCCGTTTGGC TTTTTCACAT GTCTGATTTT TGTATAATCA	1974
ACAGGCACGG AGCCGGAATC TTTCGCCTTG GAAAAATAAG CGGCGATCGT AGCTGCTTCC	2034
AATATGGATT GTTCATCGGG ATCGCTGCTT TTAATCACAA CGTGGGATCC	2084

(2) SEQ ID NO: 13 的资料:

(i) 序列特征:

(A) 长度: 1455 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: DNA(基因组)

(xi) 序列描述: SEQ ID NO: 13:

CATCATAATG GAACAAATGG TACTATGATG CAATATTTTCG AATGGTATTT GCCAAATGAC	60
GGGAATCATT GGAACAGGTT GAGGGATGAC GCAGCTAACT TAAAGAGTAA AGGGATAACA	120
GCTGTATGGA TCCCACCTGC ATGGAAGGGG ACTTCCAGAG ATGATGTAGG TTATGGAGCC	180
TATGATTTAT ATGATCTTGG AGAGTTTAAC CAGAAGGGGA CGGTTCGTAC AAAATATGGA	240
ACACGCAACC AGCTACAGGC TGCAGTGACC TCTTTAAAAA ATAACGGCAT TCAGGTATAT	300
GGTGATGTCG TCATGAATCA TAAAGGTGGA GCAGATGGTA CGGAAATTGT AAATGCGGTA	360
GAAGTGAATC GGAGCAACCG AAACCAGGAA ACCTCAGGAG AGTATGCAAT AGAAGCGTGG	420
ACAAAGTTTG ATTTTCCTGG AAGAGGAAAT AACCATTCCA GCTTTAAGTG GCGCTGGTAT	480
CATTTTGATG GGACAGATTG GGATCAGTCA CGCCAGCTTC AAAACAAAAT ATATAAATTC	540
AGGGGAACAG GCAAGGCCTG GGAAGGGGAA GTCGATACAG AGAATGGCAA CTATGACTAT	600
CTTATGTATG CAGACGTGGA TATGGATCAC CCAGAAGTAA TACATGAACT TAGAACTGG	660
GGAGTGTGGT ATACGAATAC ACTGAACCTT GATGGATTTA GAATAGATGC AGTGAAACAT	720
ATAAAATATA GCTTTACGAG AGATTGGCTT ACACATGTGC GTAACACCAC AGGTAAACCA	780

ATGTTTGCAG TGGCTGAGTT TTGGAAAAAT GACCTTGGTG CAATTGAAAA CTATTTGAAT	840
AAAACAAGTT GGAATCACTC GGTGTTTGAT GTTCCTCTCC ACTATAATTT GTACAATGCA	900
TCTAATAGCG GTGGTTATTA TGATATGAGA AATATTTTAA ATGGTTCTGT GGTGCAAAAA	960
CATCCAACAC ATGCCGTTAC TTTTGTGAT AACCATGATT CTCAGCCCCG GGAAGCATTG	1020
GAATCCTTTG TTCAACAATG GTTTAAACCA CTTGCATATG CATTGGTTCT GACAAGGGAA	1080
CAAGGTTATC CTTCCGTATT TTATGGGGAT TACTACGGTA TCCCAACCCA TGGTGTCCG	1140
GCTATGAAAT CTAAATAGA CCCTCTTCTG CAGGCACGTC AACTTTTGC CTATGGTACG	1200
CAGCATGATT ACTTTGATCA TCATGATATT ATCGGTTGGA CAAGAGAGGG AAATAGCTCC	1260
CATCCAAATT CAGGCCTTGC CACCATTATG TCAGATGGTC CAGGTGGTAA CAAATGGATG	1320
TATGTGGGGA AAAATAAAGC GGGACAAGTT TGGAGAGATA TTACCGGAAA TAGGACAGGC	1380
ACCGTCACAA TTAATGCAGA CGGATGGGGT AATTTCTCTG TTAATGGAGG GTCCGTTTCG	1440
GTTTGGGTGA AGCAA	1455

(2) SEQ ID NO: 14 的资料:

(i) 序列特征:

(A) 长度: 1455 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: DNA(基因组)

(xi) 序列描述: SEQ ID NO: 14:

CATCATAATG GGACAAATGG GACGATGATG CAATACTTTG AATGGCACTT GCCTAATGAT	60
GGGAATCACT GGAATAGATT AAGAGATGAT GCTAGTAATC TAAGAAATAG AGGTATAACC	120
GCTATTTGGA TTCCGCCTGC CTGGAAAGGG ACTTCGCAA ATGATGTGGG GTATGGAGCC	180
TATGATCTTT ATGATTTAGG GGAATTTAAT CAAAGGGGA CGGTTCGTAC TAAGTATGGG	240
ACACGTAGTC AATTGGAGTC TGCCATCCAT GCTTTAAAGA ATAATGGCGT TCAAGTTTAT	300
GGGGATGTAG TGATGAACCA TAAAGGAGGA GCTGATGCTA CAGAAAACGT TCTTGCTGTC	360
GAGGTGAATC CAAATAACCG GAATCAAGAA ATATCTGGGG ACTACACAAT TGAGGCTTGG	420
ACTAAGTTTG ATTTTCCAGG GAGGGGTAAT ACATACTCAG ACTTTAAATG GCGTTGGTAT	480
CATTCGATG GTGTAGATTG GGATCAATCA CGACAATTCC AAAATCGTAT CTACAAATTC	540
CGAGGTGATG GTAAGGCATG GGATTGGGAA GTAGATTCGG AAAATGGAAA TTATGATTAT	600

TTAATGTATG CAGATGTAGA TATGGATCAT CCGGAGGTAG TAAATGAGCT TAGAAGATGG	660
GGAGAATGGT ATACAAATAC ATTAAATCTT GATGGATTTA GGATCGATGC GGTGAAGCAT	720
ATTAAATATA GCTTTACACG TGATTGGTTG ACCCATGTAA GAAACGCAAC GGGAAAAGAA	780
ATGTTTGCTG TTGCTGAATT TTGGAAAAAT GATTTAGGTG CCTTGGAGAA CTATTTAAAT	840
AAAACAACT GGAATCATTC TGTCTTTGAT GTCCCCCTTC ATTATAATCT TTATAACGCG	900
TCAAATAGTG GAGGCAACTA TGACATGGCA AACTTTCTTA ATGGAACGGT TGTTCAAAAG	960
CATCCAATGC ATGCCGTAAC TTTTGTGGAT AATCACGATT CTCAACCTGG GGAATCATT	1020
GAATCATTTG TACAAGAATG GTTTAAGCCA CTTGCTTATG CGCTTATTTT AACAAAGAGAA	1080
CAAGGCTATC CCTCTGTCTT CTATGGTGAC TACTATGGAA TTCCAACACA TAGTGTCCTCA	1140
GCAATGAAAG CCAAGATTGA TCCAATCTTA GAGGCGCGTC AAAATTTTGC ATATGGAACA	1200
CAACATGATT ATTTTGACCA TCATAATATA ATCGGATGGA CACGTGAAGG AAATACCACG	1260
CATCCCAATT CAGGACTTGC GACTATCATG TCGGATGGGC CAGGGGGAGA GAAATGGATG	1320
TACGTAGGGC AAAATAAAGC AGGTCAAGTT TGGCATGACA TAACTGGAAA TAAACCAGGA	1380
ACAGTTACGA TCAATGCAGA TGGATGGGCT AATTTTTCAG TAAATGGAGG ATCTGTTTCC	1440
ATTTGGGTGA AACGA	1455

(2) SEQ ID NO: 15 的资料:

(i) 序列特征:

(A) 长度: 60 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=正向引物 FSA

(ix)特性:

(A)名称/关键词: misc-feature

(B)位置:22-27、29、31-33、41

(C)其他资料: /Note=1:35%A,65%C

2:83%G,17%A

3:63%G,37%T
 4:86%G,14%A
 5:85%G,15%C
 6:50%T,50%C
 7:95%A,5%G
 8:58%G,37%A,5%T
 9:86%C,13%A,1%G
 10:83%T,17%G
 11:92%G,8%C

(xi) 序列描述: SEQ ID NO: 15:

caaaatcgta tctacaaatt cl23456a7g 8910tgggatt
 11ggaagtaga ttcggaaaat
 60

(2) SEQ ID NO: 16 的资料:

(i) 序列特征:

(A) 长度: 21 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=反向引物 RSA

(xi) 序列描述: SEQ ID NO: 16:

gaatttgtag atacgatttt g
 21

(2) SEQ ID NO: 17 的资料:

(i) 序列特征:

(A) 长度: 24 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 B1

(xi) 序列描述: SEQ ID NO: 17:

CGATTGCTGA CGCTGTTATT TGCG

24

(2) SEQ ID NO: 18 的资料:

(i) 序列特征:

(A) 长度: 24 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 Y2

(xi) 序列描述: SEQ ID NO: 18:

CTTGTTCCCT TGTCAGAACC AATG

24

(2) SEQ ID NO: 19 的资料:

(i) 序列特征:

(A) 长度: 30 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 101458

(xi) 序列描述: SEQ ID NO: 19:

GTCATAGTTG CCGAAATCTG TATCGACTTC

30

(2) SEQ ID NO: 20 的资料:

(i) 序列特征:

(A) 长度: 35 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 101638

(xi) 序列描述: SEQ ID NO: 20:

CCCAGTCCCA CGTACGTCCC CTGAATTTATATA TTTTG

35

(2) SEQ ID NO: 21 的资料:

(i) 序列特征:

(A) 长度: 21 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=Oligo 1

(A)名称/关键词: misc-feature

(B)位置: 12

(D)其他资料: /Note=N=25%A、25%C、25%G、25%T

(xi) 序列描述: SEQ ID NO: 21:

CCCAGTCCCA GNTCTTTCCC CTGAATTTAT ATATTTTG

38

(2) SEQ ID NO: 22 的资料:

(i) 序列特征:

(A) 长度: 25 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 X2

(xi) 序列描述: SEQ ID NO: 22:

GCGTGGACAA AGTTTGATTT TCCTG

25

(2) SEQ ID NO: 23 的资料:

(i) 序列特征:

(A) 长度: 21 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA01

(xi) 序列描述: SEQ ID NO: 23:

(2) SEQ ID NO: 24 的资料:

(i) 序列特征:

(A) 长度: 24 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA03

(xi) 序列描述: SEQ ID NO: 24:

GCATTGGATG CTTTGAACA ACCG

24

(2) SEQ ID NO: 25 的资料:

(i) 序列特征:

(A) 长度: 26 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA07

(xi) 序列描述: SEQ ID NO: 25:

CGCAAAATGA TATCGGGTAT GGAGCC

26

(2) SEQ ID NO: 26 的资料:

(i) 序列特征:

(A) 长度: 29 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA20

(ix)特性:

(A)名称/关键词: misc-feature

(B)位置: 13、14

(D)其他资料: /Note:S=C 和 G 的混合物
W=A 和 T 的混合物

(xi) 序列描述: SEQ ID NO: 26:

GTGATGAACC ACSWAGGTGG AGCTGATGC

29

(2) SEQ ID NO: 27 的资料:

(i) 序列特征:

(A) 长度: 30 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA14

(ix)特性:

(A)名称/关键词: misc-feature

(B)位置: 13、14

(D)其他资料: /Note:R= A 和 G 的混合物
Y= C 和 T 的混合物

(xi) 序列描述: SEQ ID NO: 27:

GATGGTGTAT GGRYCAATCA CGACAATTCC

30

(2) SEQ ID NO: 28 的资料:

(i) 序列特征:

(A) 长度: 28 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA15

(xi) 序列描述: SEQ ID NO: 28:

GGTGTATGGG ATAACCTCACG ACAATTCC

28

(2) SEQ ID NO: 29 的资料:

(i) 序列特征:

(A) 长度: 28 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA16

(xi) 序列描述: SEQ ID NO: 29:

GGTGTATGGG ATCTCTCACG ACAATTCC

28

(2) SEQ ID NO: 30 的资料:

(i) 序列特征:

(A) 长度: 32 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA17

(xi) 序列描述: SEQ ID NO: 30:

GGGATCAATC ACGAAATTTTC CAAAATCGTA TC

32

(2) SEQ ID NO: 31 的资料:

(i) 序列特征:

(A) 长度: 32 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA18

(xi) 序列描述: SEQ ID NO: 31:

GGGATCAATC ACGACTCTTC CAAAATCGTA TC

32

(2) SEQ ID NO: 32 的资料:

(i) 序列特征:

(A) 长度: 34 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA06

(xi) 序列描述: SEQ ID NO: 32:

GGAAATTATG ATTATATCAT GTATGCAGAT GTAG

34

(2) SEQ ID NO: 33 的资料:

(i) 序列特征:

(A) 长度: 30 个碱基对

- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线性
- (ii) 分子类型: 其他核酸
- (ix)特性:
 - (A)名称/关键词: misc-feature
 - (B)其他资料: /desc=引物 DA09
- (xi) 序列描述: SEQ ID NO: 33:
GCTGAATTTT GGTCGAATGA TTTAGGTGCC

30

(2) SEQ ID NO: 34 的资料:

- (i) 序列特征:
 - (A) 长度: 30 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑结构: 线性
- (ii) 分子类型: 其他核酸
- (ix)特性:
 - (A)名称/关键词: misc-feature
 - (B)其他资料: /desc=引物 DA11
- (xi) 序列描述: SEQ ID NO: 34:
GCTGAATTTT GGTCGAATGA TTTAGGTGCC

30

(2) SEQ ID NO: 35 的资料:

- (i) 序列特征:
 - (A) 长度: 27 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑结构: 线性
- (ii) 分子类型: 其他核酸
- (ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA21

(xi) 序列描述: SEQ ID NO: 35:

GAATTTTGGG AGTACGATTT AGGTCGG

27

(2) SEQ ID NO: 36 的资料:

(i) 序列特征:

(A) 长度: 29 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA12

(ix)特性:

(A)名称/关键词: misc-feature

(B) 位置: 12、13

(D) 其他资料: /Note:R=A 和 G 的混合物
Y=C 和 T 的混合物

(xi) 序列描述: SEQ ID NO: 36:

GGAAAAACGA TRYCGGTGCC TTGGAGAAC

29

(2) SEQ ID NO: 37 的资料:

(i) 序列特征:

(A) 长度: 27 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA13

(ix)特性:

(A)名称/关键词: misc-feature

(B) 位置: 14, 15

(D) 其他资料: /Note: R=A 和 G 的混合物
Y=C 和 T 的混合物

(xi) 序列描述: SEQ ID NO: 37:

GATTAGGTG CCTRYCAGAA CTATTTA

27

(2) SEQ ID NO: 38 的资料:

(i) 序列特征:

(A) 长度: 26 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA08

(xi) 序列描述: SEQ ID NO: 38:

CCCCCTTCAT GAGAATCTTT ATAACG

26

(2) SEQ ID NO: 39 的资料:

(i) 序列特征:

(A) 长度: 25 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA04

(xi) 序列描述: SEQ ID NO: 39:

GAATCCGAAC CTCATTACAC ATTCG

25

(2) SEQ ID NO: 40 的资料:

(i) 序列特征:

(A) 长度: 38 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA05

(xi) 序列描述: SEQ ID NO: 40:

CGGATGGACT CGAGAAGGAA ATACCACG

38

(2) SEQ ID NO: 41 的资料:

(i) 序列特征:

(A) 长度: 31 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA10

(xi) 序列描述: SEQ ID NO: 41:

CGTAGGGCAA AATCAGGCCG GTCAAGTTTG G

31

(2) SEQ ID NO: 42 的资料:

(i) 序列特征:

(A) 长度: 31 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA22

(xi) 序列描述: SEQ ID NO: 42:

CATAACTGGA AATCGCCCGG GAACAGTTAC G

31

(2) SEQ ID NO: 43 的资料:

(i) 序列特征:

(A) 长度: 29 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA19

(ix)特性:

(A)名称/关键词: misc-feature

(B)位置: 12

(D)其他资料:/Note:W=A 和 T 的混合物

(xi) 序列描述: SEQ ID NO: 43:

CTGGAAATAA AWCCGGAACA GTTACG

36

(2) SEQ ID NO: 44 的资料:

(i) 序列特征:

(A) 长度: 32 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA23

(xi) 序列描述: SEQ ID NO: 44:

GGAAATAAAC CAGGACCCGT TACGATCAAT GC

32

(2) SEQ ID NO: 45 的资料:

(i) 序列特征:

(A) 长度: 28 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA32

(xi) 序列描述: SEQ ID NO: 45:

GAGGCTTGGA CTAGGTTTGA TTTTCCAG

28

(2) SEQ ID NO: 46 的资料:

(i) 序列特征:

(A) 长度: 30 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 DA31

(xi) 序列描述: SEQ ID NO: 46:

GCTGAATTTT GGCGCAATGA TTTAGGTGCC

30

(2) SEQ ID NO: 47 的资料:

(i) 序列特征:

(A) 长度: 34 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 bm4

(xi) 序列描述: SEQ ID NO: 47:

GTGTTTGACG TCCCGCTTCA TGAGAATTTA CAGG

34

(2) SEQ ID NO: 48 的资料:

(i) 序列特征:

(A) 长度: 34 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 bm5

(xi) 序列描述: SEQ ID NO: 48:

GTGTTTGACG TCCCGCTTCA TAAGAATTTA CAGG

34

(2) SEQ ID NO: 49 的资料:

(i) 序列特征:

(A) 长度: 34 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 bm6

(xi) 序列描述: SEQ ID NO: 49:

GTGTTTGACG TCCCGCTTCA TGCCAATTTA CAGG

34

(2) SEQ ID NO: 50 的资料:

(i) 序列特征:

(A) 长度: 32 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 bm8

(xi) 序列描述: SEQ ID NO: 50:

AGGGAATCCG GATACCCTGA GGTTTTCTAC GG

32

(2) SEQ ID NO: 51 的资料:

(i) 序列特征:

(A) 长度: 34 个碱基对

- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线性
- (ii) 分子类型: 其他核酸
- (ix)特性:
 - (A)名称/关键词: misc-feature
 - (B)其他资料: /desc=引物 bm11
- (xi) 序列描述: SEQ ID NO: 51:
GATGTGTTT TGGATCATAA GGCCGGCGGT GATG

34

(2) SEQ ID NO: 52 的资料:

- (i) 序列特征:
 - (A) 长度: 22 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑结构: 线性
- (ii) 分子类型: 其他核酸
- (ix)特性:
 - (A)名称/关键词: misc-feature
 - (B)其他资料: /desc=引物 p1
- (xi) 序列描述: SEQ ID NO: 52:

CTGTTATTAA TGCCGCCAAA CC

22

(2) SEQ ID NO: 53 的资料:

- (i) 序列特征:
 - (A) 长度: 24 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑结构: 线性
- (ii) 分子类型: 其他核酸
- (ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 p2

(xi) 序列描述: SEQ ID NO: 53:

GGAAAAGAAA TGTTTACGGT TGCG

24

(2) SEQ ID NO: 54 的资料:

(i) 序列特征:

(A) 长度: 25 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 p3

(xi) 序列描述: SEQ ID NO: 54:

GAAATGAAGC GGAACATCAA ACACG

25

(2) SEQ ID NO: 55 的资料:

(i) 序列特征:

(A) 长度: 30 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其他核酸

(ix)特性:

(A)名称/关键词: misc-feature

(B)其他资料: /desc=引物 p4

(xi) 序列描述: SEQ ID NO: 55:

GTATGATTTA GGAGAATTCC

20

1	HHNGTNGTMM	QYFEWHLPLND	GNHWNRLRDD	ASNLRNRGIT	AIWIPPAWK	50
1	..NGTNGTMM	QYFEWYLPND	GNHWNRLRSD	ASNLDKKGIS	AVWIPPAWK	
2	HHNGTNGTMM	QYFEWYLPND	GNHWNRLRDD	AANLKSKGIT	AVWIPPAWK	
3	...VNGTLM	QYFEWYTPND	GQHWKRLQND	AEHLSDIGIT	AVWIPPAYK	
4	..ANLNGTLM	QYFEWYMPND	GQHWRRLQND	SAYLAEHGIT	AVWIPPAYK	
5	.AAPFNGTMM	QYFEWYLPDD	GTLWTKVANE	ANLSSLGIT	ALWLPPAYK	
6						
51	TSQNDVGCGA	YDLYDLGEFN	QKGTVRTKYG	TRSQLESaih	ALKNNGVQVY	100
1	ASQNDVGCGA	YDLYDLGEFN	QKGTIRTKYG	TRNQLQAaVN	ALKSNGIQVY	
2	TSQNDVGCGA	YDLYDLGEFN	QKGTVRTKYG	TRNQLQAAVT	SLKNNGIQVY	
3	LSQSDNGYGP	YDLYDLGEFQ	QKGTVRTKYG	TKSELQDAIG	SLHSRNVQVY	
4	TSQADVGCGA	YDLYDLGEFH	QKGTVRTKYG	TKGELQSAIK	SLHSRDINVY	
5	TSRSDVGCGV	YDLYDLGEFN	QKGTVRTKYG	TKAQYLQAIQ	AAHAAGMQVY	
6						
101	GDVVMNHKGG	ADATENVLAV	EVNPNNRNQE	ISGDYTiEAW	TKFDFPGRGN	150
1	GDVVMNHKGG	ADATEMVRaV	EVNPNNRNQE	VSGEYTiEAW	TKFDFPGRGN	
2	GDVVMNHKGG	ADGTEIVNAV	EVNRSNRNQE	TSGEYAIeAW	TKFDFPGRGN	
3	GDVVLNHKAG	ADATEDVTaV	EVNPANRNQE	TSEeYQIKaW	TDERFPGRGN	
4	GDVVINHKGG	ADATEDVTaV	EVDPADRNRV	ISGEHLIKaW	THFHFPGRGS	
5	ADVVFdHKGG	ADGTEWVDaV	EVNPSDRNQE	ISGTyQIQaW	TKFDFPGRGN	
6						

图 1

151						200
1	TYSDFKWRWY	HFDGVDWDQS	RQFQNRIYKF	RGDGKAWDWE	VDSENGNYDY	
2	THSNFKWRWY	HFDGVDWDQS	RKLNRIYKF	RGDGKAWDWE	VDTEGNYDY	
3	NHSSFKWRWY	HFDGTDWDQS	RQLQNKIYKF	RGTGKAWDWE	VDTEGNYDY	
4	TYSDFKWHWY	HFDGADWDES	RKI.SRIYKF	RGEKAWDWE	VSENGNYDY	
5	TYSDFKWHWY	HFDGTDWDES	RKL.NRIYKF	.QGKAWDWE	VSENGNYDY	
6	TYSSFKWRWY	HFDGVDWDES	RKL.SRIYKF	RGIGKAWDWE	VDTEGNYDY	
201						250
1	LMYADVMDH	PEVVNELRRW	GEWYTNLTNL	DGFRIDAVKH	IKYSFTRDWL	
2	LMYADIDMDH	PEVVNELRNW	GVWYTNLTGL	DGFRIDAVKH	IKYSFTRDWS	
3	LMYADVMDH	PEVIHELNRW	GVWYTNLTNL	DGFRIDAVKH	IKYSFTRDWL	
4	LMYADVVDYDH	PDVVAETKKW	GIWYANELSL	DGFRIDAACH	IKFSFLRDWV	
5	LMYADIDYDH	PDVAAEIKRW	GTWYANELQL	DGFRLDVAVKH	IKFSFLRDWV	
6	LMYADLDMDH	PEVVTELKNW	GKWYVNTTNI	DGFRLDVAVKH	IKFSFFPDWL	
251						300
1	THVRNATGKE	MFAVAEFWKN	DLGALENYLN	KTNWNHSHVED	VPLHYNLYNA	
2	IHVRSATGKN	MFAVAEFWKN	DLGALENYLN	KTNWNHSHVED	VPLHYNFYNA	
3	THVRNTTGKP	MFAVAEFWKN	DLGALENYLN	KTSWNHSAFD	VPLHYNLYNA	
4	QAVRQATGKE	MFTVAEYQON	NAGKLENYLN	KTSFNQSVFD	VPLHFNLQAA	
5	NHVREKTGKE	MFTVAEYQON	DLGALENYLN	KTNFNHSHVED	VPLHYQFHAA	
6	SYVRSQTGKP	LFTVGEYWSY	DINKLHNYIT	KTDGTMSLED	APLHNKFYTA	

图 1 (续)

[illegible]

图 1 (续)

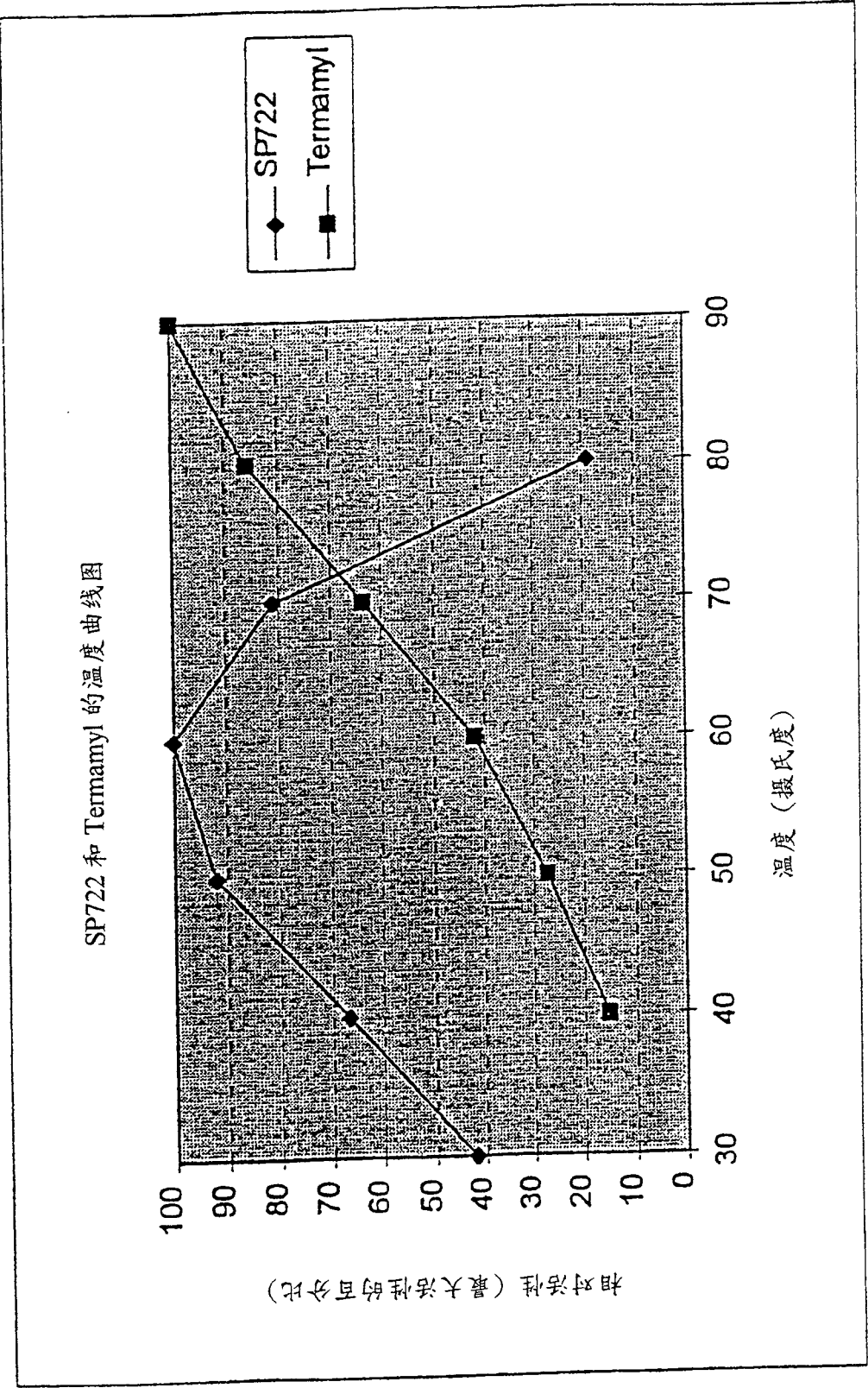


图 2

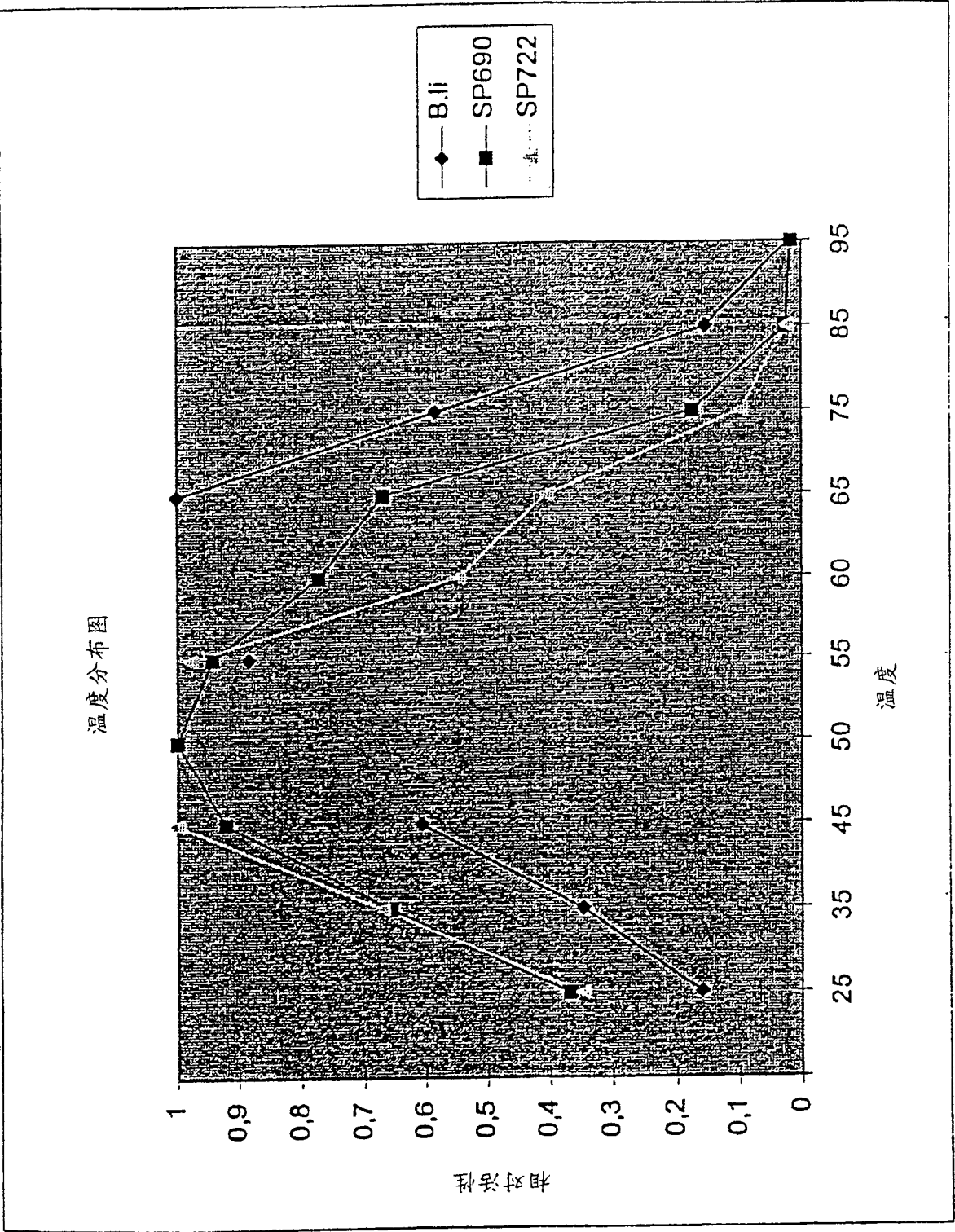


图 3

1
1 MKFVLLLSLI GFCWAQYDPH TSDG.RTAIV HLFWRWVDI AKECERYLAP 50
2 LLSLI GFCWAQYDPH TADG.RTAIV HLFWRWADI AKECERYLAP
3 QYAPQ TQSG.RTDIV HLFWRWVDI ALECERYLGP
4 MKFFLLLLFTI GFCWAQYSPN TQQG.RTSIV HLFWRWVDI ALECERYLAP
5 MKLNKIIITTA GLSLGLLPS IATATPTTFV HLFERNWQDV AQECEQYLG

51
1 KGFGGVQVSP PNENVVHNP SRPWERYQP ISYKICTRSG NEDEFKDMVT 100
2 KGFGGVQVSP PNENIIINP SRPWERYQP ISYKICSRSG NENEFKDMVT
3 KGFGGVQVSP PNENVVTNP SRPWERYQP VSYKLCTRS NENEFKDMVT
4 KGFGGVQVSP PNENVAIYNP FRPWERYQP VSYKLCTRS NEDEFKDMVT
5 KGAAAVQVSP PNEHI T GSQWTRYQP VSYELQSRGG NRAQFIDMVN

101
1 RCNNVGVRIY VDAVINHMC AGNPAGTSST CGSYLNPNNR EFPAVPYSAW 150
2 RCNNVGVRIY VDAVINHMC SGSAGTHST CGSYFNPNNR EFSAVPYSAW
3 RCNNVGVRIY VDAVINHMC SGAAAGTGT CGSYCNPCNR EFPAVPYSAW
4 RCNNVGVRIY VDAVINHMC NAVSAGTSST CGSYFNPGR DFPAVPYSGW
5 RCSAAGVDIY VDTLINHM AAGSGTGA GNSF GNK SFPI YSPQ

151
1 DFNDNKC GEIDNYDA YQVRNCRLTG LLDLALDKDY VRTKVADYMN 200
2 YFNDNKC GEINNYDA NQVRNCRLSG LLDLALDKDY VRTKVADYMN
3 DFNDGKCKTA SGGIESYNDP YQVRDCQLVG LLDLALDKDY VRSMIADYLN
4 DFNDGKCKTG SGDIENYDA TVVRDCRLTG LLDLALDKDY VRSKIAEYMN
5 DFHES.CTIN NSDYG NDR YRVQNCCELVG LADLDTASNY VQNTIAAYIN

图 4

201					250
1	HLIDIGVAGF	RLDAAKHMWP	RDIAVLDKL	HNLNTKWFSQ	GSRPFIFQEV
2	NLIDIGVAGF	RLDAAKHMWP	GDIKAVLDKL	HNLNTKWFSQ	GSRPFIFQEV
3	KLIDIGVAGF	RLDASKHMWP	GDIKAVLDKL	HNLNTNWFFA	GSRPFIFQEV
4	HLIDIGVAGF	RLDASKHMWP	GDIKAILDKL	HNLNSNWFFA	GSKPFIYQEV
5	DLQAIGVKGF	RFDASKHVAA	SDIQSLMAKV	N.....	.GSPVVFQEV
251					300
1	IDLGGEAIKG	SEYFGNGRVT	EFKYGAKLGT	VIRKWNGEKM	SYLKNWGEGW
2	IDLGGEAIKG	SEYFGNGRVT	EFKYGAKLGT	VIRKWNGEKM	SYLKNWGEGW
3	IDLGGEAIKS	GEYFSNGRVT	EFKYGAKLGT	VVRKWSGEKM	SYLKNWGEGW
4	IDLGGEPIKS	SDYFGNGRVT	EFKYGAKLGT	VIRKWNGEKM	SYLKNWGEGW
5	IDQGGEAVGA	SEYLSLGLVT	EFKYSTELGN	TFR...NGSL	AWLSNFGEGW
301					350
1	GLVPSDRALV	FVDNHDNQRG	HGAGGSSILT	FWDARMYKMA	VGFMLAHPYG
2	GFVPTDRALV	FVDNHDNQRG	HGAGGASILT	FWDARMYKMA	VGFMLAHPYG
3	GFMPSDRALV	FVDNHDNQRG	HGAGGSSILT	FWDAYRKLVA	VGFMLAHPYG
4	GFVPSDRALV	FVDNHDNQRG	HGAGGASILT	FWDARLYKMA	VGFMLAHPYG
5	GFMPSSSAVV	FVDNHDNQRG	HGGAG.NVIT	FEDGRLYDLA	NVFMLAYPYG

图 4 (续)

351
1 FTRVMSSYRW NRNFQNGKDQ NDWIGPPNNN GVTKEVTINA DTTCGNDWVC 400
2 FTRVMSSYRR TRNFQNGKDV NDWIGPPNNN GVTKEVTINP DTTCGNDWVC
3 FTRVMSSYRW ARNFVNGEDV NDWIGPPNNN GVIKEVTINA DTTCGNDWVC
4 FTRVMSSYRW PRQFQNGNDV NDWVGPPNNN GVIKEVTINP DTTCGNDWVC
5 YPKVMSSY...DFHGDTDA GGPVVPVHNN GNLE.....CFASNWKC

401
1 EHRWRQIRNM VAFRNVVNGQ .PFSNWWDNN SNQVAFSRGN RGFIVENNDD 450
2 EHRWRQIRNM VAFRNVVNGQ .PFANWWDNG SNQVAFSRGN RGFIVENNDD
3 EHRWREIRNM VWERNVVDGE .PFANWWDNG SNQVAFGRGN RGFIVENNDD
4 EHRWRQIRNM VIFRNVVDGQ .PFTNWDYDNG SNQVAFGRGN RGFIVENNDD
5 EHRWSYIAGG VDFRNNNTADN WAVTNWWDNT NNQISFGRGS SGHMAINKE

451
1 WALSATLQTG LPAGTYCDVI SGD KVDG..N CTGLRVNVGS DGKAHFSISN 500
2 WALSSLTQTG LPAGTYCDVI SGD KVG..N CTGLKVNVS DGKAHFSISN
3 WQLSSLTQTG LPAGTYCDVI SGD KVG..S CTGIKVVVSS DGKAQFSISN
4 WSFSLTQTG LPAGTYCDVI SGD KING..N CTGIKIYVSD DGKAHFSISN
5 STLTATVQTD MASGYCNVL KGELSADAKS CSGEVITVNS DGTINLNIGA

501
521
1 SAEDPFIH ADSKL.....
2 SAEDPFIH ADSKL.....
3 SAEDPFIH AESKL.....
4 SAEDPFIH AESKL.....
5 WDA...MAIH KNAKLNTSSA S

图 4 (续)