

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公开说明书

[21] 申请号 200480016564.9

[51] Int. Cl.

C12N 15/56 (2006.01)

C12N 9/34 (2006.01)

C12N 1/15 (2006.01)

C12P 19/20 (2006.01)

A61K 38/47 (2006.01)

[43] 公开日 2006年8月9日

[11] 公开号 CN 1816627A

[22] 申请日 2004.6.11

[21] 申请号 200480016564.9

[30] 优先权

[32] 2003.6.13 [33] DK [31] PA200300884

[86] 国际申请 PCT/DK2004/000406 2004.6.11

[87] 国际公布 WO2004/111218 英 2004.12.23

[85] 进入国家阶段日期 2005.12.13

[71] 申请人 诺和酶股份有限公司

地址 丹麦拜格斯威尔德

[72] 发明人 A·维克索-尼尔森 B·E·诺曼
S·兰德维克

[74] 专利代理机构 北京市中咨律师事务所

代理人 黄革生 林柏楠

权利要求书 5 页 说明书 40 页 序列表 8 页

[54] 发明名称

生产葡糖淀粉酶的方法及其用途

[57] 摘要

本发明涉及生产具有葡糖淀粉酶活性的分离的多肽、包含编码所述多肽的核酸序列的丝状真菌宿主细胞的方法，以及使用所述多肽的方法。

1、一种重组生产葡糖淀粉酶的方法，所述方法包括在丝状真菌宿主细胞中表达编码具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)的多核苷酸的步骤，其中所述多肽包含与 SEQ ID NO: 2 的 19 至 471 位（包括二者）所示序列具有至少 70% 同一性的氨基酸序列。

2、如权利要求 1 所述的方法，其中所述多肽还包含淀粉结合结构域(SBD)。

3、如权利要求 2 所述的方法，其中所述淀粉结合结构域包含与 SEQ ID NO: 2 的 483 至 579 位（包括二者）所示序列具有至少 80% 同一性的氨基酸序列。

4、如权利要求 1-3 任一项所述的方法，其中所述多肽包含与 SEQ ID NO: 2 的 19 至 579 位（包括二者）所示序列具有至少 80% 同一性的氨基酸序列。

5、如权利要求 1-4 任一项所述的方法，其中所述多肽包含位于淀粉结合结构域和其余多肽之间的至少为 2 个氨基酸的接头。

6、如权利要求 1-5 任一项所述的方法，其中所述多肽包含信号肽。

7、如权利要求 6 所述的方法，其中所述信号肽包含与 SEQ ID NO: 2 的 1 至 18 位（包括二者）所示序列具有至少 95% 同一性的氨基酸序列。

8、如权利要求 1-7 任一项所述的方法，其中所述多肽包含与 SEQ ID NO: 2 所示序列具有至少 70% 同一性的氨基酸序列。

9、如权利要求 1-8 任一项所述的方法，其中所述丝状真菌宿主细胞属于曲霉属。

10、如权利要求 9 所述的方法，其中所述曲霉属宿主细胞是泡盛曲霉、米曲霉或黑曲霉细胞。

11、如权利要求 1-10 任一项所述的方法，其中进行回收和/或纯化葡糖淀粉酶的后续步骤。

12、一种糖化液化淀粉的方法，包括用具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)处理液化淀粉，其中所述多肽包含与 SEQ ID NO: 2 的 19 至 471 位（包括二者）所示序列具有至少 70% 同一性的氨基酸序列，由此在 60℃ 30% w/w (g/100g 干物质)底物浓度下得到至少为 96%的葡萄糖(DX) %值，所述 DX 值如文中实施例 7 中所定义的那样测定。

13、如权利要求 12 所述的方法，其中所述多肽还包括淀粉结合结构域(SBD)。

14、如权利要求 13 所述的方法，其中所述淀粉结合结构域包含与 SEQ ID NO: 2 的 483 至 579 位（包括二者）所示序列具有至少 80% 同一性的氨基酸序列。

15、如权利要求 12 - 14 任一项所述的方法，其中所述多肽包含与 SEQ ID NO: 2 的 19 至 579 位（包括二者）所示序列具有至少 80% 同一性的氨基酸序列。

16、如权利要求 12 - 15 任一项所述的方法，其中所述多肽包含位于淀粉结合结构域和其余多肽之间的至少为 2 个氨基酸的接头。

17、如权利要求 12 - 16 任一项所述的方法，其中所述多肽包含信号肽。

18、如权利要求 17 所述的方法，其中所述信号肽包含与 SEQ ID NO: 2 的 1 至 18 位（包括二者）所示序列具有至少 95% 同一性的氨基酸序列。

19、如权利要求 12 - 18 任一项所述的方法，其中所述多肽包含与 SEQ ID NO: 2 的 1 至 579 位（包括二者）所示序列具有至少 70% 同一性的氨基酸序列。

20、丝状真菌宿主细胞，其包含至少一个拷贝的编码具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)的多核苷酸，所述多肽包含与 SEQ ID NO: 2 的 19 至 471 位（包括二者）所示序列具有至少 70% 同一性的氨基酸序列。

21、如权利要求 20 所述的宿主细胞，其中所述多肽包含淀粉结合结构域(SBD)。

22、如权利要求 21 所述的宿主细胞，其中所述淀粉结合结构域包含与 SEQ ID NO: 2 的 483 至 579 位（包括二者）所示序列具有至少 80% 同一性的氨基酸序列。

23、如权利要求 20 - 22 任一项所述的宿主细胞，其中所述多肽包含与 SEQ ID NO: 2 的 19 至 579 位（包括二者）所示序列具有至少 80% 同一性的氨基酸序列。

24、如权利要求 20 - 23 任一项所述的宿主细胞，其中所述多肽包含位于淀粉结合结构域和其余多肽之间的至少为 2 个氨基酸的接头。

25、如权利要求 20 - 24 任一项所述的宿主细胞，其中所述多肽包含信号肽。

26、如权利要求 25 所述的宿主细胞，其中所述信号肽包含与 SEQ ID NO: 2 的 1 至 18 位（包括二者）所示序列具有至少 95% 同一性的氨基酸序列。

27、如权利要求 20 - 26 任一项所述的宿主细胞，其中所述多肽包含与 SEQ ID NO: 2 的 1 至 579 位（包括二者）所示序列具有至少 70% 同一性的氨基酸序列。

28、如权利要求 20 - 27 任一项所述的宿主细胞，其属于曲霉属。

29、如权利要求 28 所述的宿主细胞，其为泡盛曲霉、米曲霉或黑曲霉细胞。

30、具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)在淀粉转化方法中的用途，其中所述多肽包含与 SEQ ID NO: 2 的 19 至 471 位（包括二者）所示序列具有至少 70% 同一性的氨基酸序列。

31、具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)在连续淀粉转化方法中的用途，其中所述多肽包含与 SEQ ID NO: 2 的 19 至 471 位（包括二者）所示序列具有至少 70% 同一性的氨基酸序列。

32、具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)在生产寡糖的方法中的用途，其中所述多肽包含与 SEQ ID NO: 2 的 19 至 471 位（包括二者）所示序列具有至少 70% 同一性的氨基酸序列。

33、具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)在生产专用糖浆的方法中的用途，其中所述多肽包含与 SEQ ID NO: 2 的 19 至 471 位（包括二者）所示序列具有至少 70% 同一性的氨基酸序列。

34、具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)在生产燃料酒精或饮用酒精（便携式酒精）方法中的用途，其中所述多肽包含与 SEQ ID NO: 2 的 19 至 471 位（包括二者）所示序列具有至少 70% 同一性的氨基酸序列。

35、具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)在生产饮料的方法中的用途，其中所述多肽包含与 SEQ ID NO: 2 的 19 至 471 位（包括二者）所示序列具有至少 70% 同一性的氨基酸序列。

36 具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)在生产有机化合物如柠檬酸、抗坏血酸、赖氨酸、谷氨酸的发酵方法中的用途，其中所述多肽包含与 SEQ ID NO: 2 的 19 至 471 位（包括二者）所示序列具有至少 70% 同一性的氨基酸序列。

37、具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)在洗涤剂中的用途，其中所述多肽包含与 SEQ ID NO: 2 的 19 至 471 位（包括二者）所示序列具有至少 70% 同一性的氨基酸序列。

38、具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)在淀粉挤压预处理中的用途，其中所述多肽包含与 SEQ ID NO: 2 的 19 至 471 位（包括二者）所示序列具有至少 70% 同一性的氨基酸序列。

39、具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)在混合肥料和生物废料处理中的用途，其中所述多肽包含与 SEQ ID NO: 2 的 19 至 471 位（包括二者）所示序列具有至少 70% 同一性的氨基酸序列。

40、具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)在纯化用于食品添加剂的植物提取物中的用途，其中所述多肽包含与 SEQ ID NO:

2 的 19 至 471 位（包括二者）所示序列具有至少 70% 同一性的氨基酸序列。

41、具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)在化妆品和药品中的用途，其中所述多肽包含与 SEQ ID NO: 2 的 19 至 471 位（包括二者）所示序列具有至少 70% 同一性的氨基酸序列。

42、具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)在烘焙工业中的用途，其中所述多肽包含与 SEQ ID NO: 2 的 19 至 471 位（包括二者）所示序列具有至少 70% 同一性的氨基酸序列。

43、具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)在宠物食品的生产中的用途，其中所述多肽包含与 SEQ ID NO: 2 的 19 至 471 位（包括二者）所示序列具有至少 70% 同一性的氨基酸序列。

生产葡糖淀粉酶的方法及其用途

发明领域

本发明涉及生产具有葡糖淀粉酶活性的分离的多肽、包含编码所述多肽的核酸序列的丝状真菌宿主细胞的方法，以及使用所述多肽的方法。

背景

高葡萄糖和果糖糖浆是通过酶促糖化液化淀粉制备的。所述糖化是通过形成葡萄糖的称为外切-1,4- α -D-葡萄糖苷酶（葡糖淀粉酶或淀粉葡萄糖苷酶）的外切淀粉酶实现的。该酶水解淀粉中的 1,4-以及 1,6- α -键。水解期间，葡萄糖单元被以逐步的方式从底物分子的非还原端除去，而见于分支糊精的 1,6- α -键被相对慢地解离。麦芽三糖和麦芽糖被以比高级寡糖更低的速率水解。葡糖淀粉酶也被用来降低啤酒的碳水化合物含量。

US 4,727,026 中描述了一种使用酶直接糖化未液化生淀粉的方法，特别是由担子菌株罗氏阿太菌 (*Athelia rolfsii*) (从前称之为罗尔伏革菌 (*Corticium rolfsii*)) 产生的淀粉酶。葡糖淀粉酶由同样的罗氏阿太菌 (罗尔伏革菌) 菌株分离到并被部分地表征 (Nagasaka, Y. 等人 (1998) “来自罗尔伏革菌的生淀粉降解葡糖淀粉酶的纯化和性能” (Purification and properties of the raw-starch-degrading glucoamylases from *Corticium rolfsii*), 《实用微生物学生物技术》(Appl Microbiol Biotechnol) 50: 323-330)。编码此葡糖淀粉酶 G2 型的基因被克隆并表达在面包酵母宿主细胞中，但是所述葡糖淀粉酶 G2 只能以极低的产量生产 (Nagasaka, Y. 等人 (1995) “罗尔伏革菌葡糖淀粉酶 cDNA 的克隆及其在酿酒酵母中的表达” (Cloning of *Corticium rolfsii* glucoamylase cDNA and its expression in *Saccharomyces cerevisiae*), 《实用微生物学生物技术》(Appl Microbiol Biotechnol) 44: 451-458)。

本领域需要新来源的具有改良性能的葡糖淀粉酶。

发明概述

本发明所要解决的一个问题是如何以商业意义的产量提供重组生产的葡糖淀粉酶。本发明人令人惊讶地发现本发明的葡糖淀粉酶能够在丝状真菌宿主细胞中表达，产生极高产量的葡糖淀粉酶。他们还证明，应用此葡糖淀粉酶处理液化淀粉时可以得到迄今未曾见过的高葡萄糖当量（或葡萄糖%值），这样提供了一种改良的和更经济的糖化液化淀粉的方法。

因此，本发明第一个方面涉及重组生产葡糖淀粉酶的方法，所述方法包括在丝状真菌宿主细胞中表达编码具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)的多核苷酸的步骤，其中所述多肽包含这样的氨基酸序列，所述氨基酸序列与 SEQ ID NO: 2 的 19 至 471 位所示序列具有至少 70% 同一性、优选至少 75%、或至少 80%、或至少 85%、或 90%、或至少 95%、或者甚至至少 98% 同一性。

一种罗氏阿太菌葡糖淀粉酶的氨基酸序列可以 SPTREMBL: Q12596 获得，除 SEQ ID NO: 2 115 位上的一个氨基酸残基外，其几乎与 SEQ ID NO: 2 所示的葡糖淀粉酶完全相同，该残基在数据库序列中是丝氨酸，而在 SEQ ID NO: 2 中是脯氨酸。数据库序列的注解确定 1-18 位氨基酸残基为信号肽，而 19-579 位残基是成熟的葡糖淀粉酶，其中残基 472-482 作为位于葡糖淀粉酶结构域和包含在残基 483-579 中的淀粉结构域之间的接头。

本发明还涉及重组生产葡糖淀粉酶的方法，所述方法包括在丝状真菌宿主细胞中表达编码具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)的多核苷酸的步骤，其中所述多肽选自：(a)在高严谨条件下与(i) SEQ ID NO: 1 的 1 至 2325 位所示多核苷酸、(ii) 介于 SEQ ID NO: 1 核苷酸 1 至 2325 之间的 cDNA 或无内含子的多核苷酸、(iii) 至少 100 个核苷酸的(i)或(ii) 的亚序列、或者(iv) (i)、(ii)或(iii)的互补链杂交的多核苷酸编码的多肽；(b) 具有 SEQ ID NO: 2 的氨基酸序列的多肽变体，所述变体包含一个或多个氨基酸的取代、缺失和/或插入；和(c) (a)或(b)中具有葡糖淀粉酶活性的片段。

第二个方面，本发明涉及糖化液化淀粉的方法，包括使用具有葡糖淀

粉酶活性的多肽(E.C. 3.2.1.3)处理液化淀粉, 其中所述多肽包含这样的氨基酸序列, 所述氨基酸序列与 SEQ ID NO: 2 的 19 至 471 位(含二者)所示序列具有至少 70% 同一性, 优选至少 75%, 或至少 80%, 或至少 85%, 或 90%, 或至少 95%, 或者甚至至少 98% 同一性, 藉此在 60°C 30% w/w (g/100g 干物质)底物浓度下得到至少为 96% 的葡萄糖(DX) %值, 所述 DX 值如文中实施例 7 所定义的那样测定。

本发明还涉及糖化液化淀粉的方法, 包括用具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)处理液化淀粉, 其中多肽选自: (a) 在高严谨条件下与(i) SEQ ID NO: 1 的 1 至 2325 位所示多核苷酸、(ii) 介于 SEQ ID NO: 1 的 1 至 2325 位核苷酸之间的 cDNA 或无内含子的多核苷酸、(iii) 至少 100 个核苷酸的(i)或(ii)的亚序列、或者(iv) (i)、(ii)或(iii)的互补链杂交的多核苷酸编码的多肽; (b) 具有 SEQ ID NO: 2 的氨基酸序列的多肽变体, 所说的变体包含一个或多个氨基酸的取代、缺失和/或插入; 和(c) (a)或(b)中具有葡糖淀粉酶活性的片段; 藉此在 60°C 30% w/w (g/100g 干物质)底物浓度下得到至少为 96% 的葡萄糖(DX) %值, 所述 DX 值如文中实施例 7 所定义的那样测定。

第三个方面, 本发明涉及丝状真菌宿主细胞, 其包含至少一个拷贝的编码具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)的多核苷酸, 其中所述多肽包含这样的氨基酸序列, 所述氨基酸序列与 SEQ ID NO: 2 的 19 至 471 位(含二者)所示序列具有至少 70% 同一性, 优选至少 75%, 或至少 80%, 或至少 85%, 或 90%, 或至少 95%, 或者甚至至少 98% 同一性。

本发明还涉及丝状真菌宿主细胞, 其包含至少一个拷贝的编码具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)的多核苷酸, 其中所述多肽选自: (a) 在高严谨条件下与(i) SEQ ID NO: 1 的 1 至 2325 位所示多核苷酸、(ii) 介于 SEQ ID NO: 1 的 1 至 2325 位核苷酸之间的 cDNA 或无内含子的多核苷酸、(iii) 至少 100 个核苷酸的(i)或(ii)的亚序列或者(iv) (i)、(ii)或(iii)的互补链杂交的多核苷酸编码的多肽; (b) 具有 SEQ ID NO: 2 的氨基酸序列的多肽变体, 所说的变体包含一个或多个氨基酸的替换、缺失和/或插入; 和(c) (a)或(b)中具有葡糖淀粉酶活性的片段。

第四方面，本发明涉及具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)在淀粉转化方法、在持续的淀粉转化方法、在生产寡糖的方法、在生产燃料酒精或饮用酒精（便携式酒精）的方法、在生产饮料的方法、在生产有机化合物如柠檬酸、抗坏血酸、赖氨酸、谷氨酸的发酵方法、或在清洁剂、在生产专用糖浆、在淀粉的挤压预处理、在混合肥料和生物废物处理、在用于食品添加剂的植物提取物的纯化、在化妆品和药品、在烘焙工业、在宠物食品的生产中的用途，其中所述多肽包含这样的氨基酸序列，所述氨基酸序列与 SEQ ID NO: 2 的 19 至 471 位（含二者）所示序列具有至少 70% 同一性，优选至少 75%，或至少 80%，或至少 85%，或 90%，或至少 95%，或者甚至至少 98% 同一性。

定义

术语“葡糖淀粉酶活性”在文中定义为从（糖）链非还原端连续水解末端 1,4- 交联的 α -D- 葡萄糖残基释放出 β -D- 葡萄糖的属于酶分类 EC 3.2.1.3 的葡聚糖 1,4- α - 葡萄糖苷酶。对本发明来说，葡糖淀粉酶活性按照 Fagershom 和 Kalkkinen, 1995, 《生物技术实用生物化学》（*Biotechnol. Appl. Biochem.*）21: 223-231 中描述的操作进行测定，在 pH 4、25°C 下使用 GO 葡萄糖氧化酶分析试剂盒 (Sigma Chemical Co., St. Louis, MO) 检测葡糖淀粉酶由 0.1M 麦芽三糖产生的葡萄糖。1 单位葡糖淀粉酶活性定义为在 pH 4、25°C 下每分钟产生的 1.0 μ mol 葡萄糖。

发明内容

第一个实施方案中，本发明涉及重组生产葡糖淀粉酶的方法，所述方法包括在丝状真菌宿主细胞中表达多核苷酸的步骤，该多核苷酸编码具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)，其中所述多肽包含这样的氨基酸序列，所述氨基酸序列与 SEQ ID NO: 2 的 19 至 471 位（包括二者）所示序列（即成熟多肽）具有至少 70% 同一性，优选至少 75%，或至少 80%，或至少 85%，或 90%，或至少 95%，或者甚至至少 98% 同一性；与 SEQ ID NO: 2 的 19 至 471 位（包括二者）所示序列具有至少大约 75%，优选

至少大约 80%，更优选至少大约 85%，甚至更优选至少大约 90%，最优选至少大约 95%，甚至最优选至少大约 97% 同一性。所述多肽在下文称之为“同源多肽”。在优选的实施方案中，同源多肽有 5 个氨基酸，优选 4 个氨基酸，更优选 3 个氨基酸，甚至更优选 2 个氨基酸，并且最优选 1 个氨基酸与 SEQ ID NO: 2 的 19 至 471 位所示氨基酸不同。对本发明来说，两个氨基酸序列之间的同一性程度由 ClustalW 法(Higgins, 1989, *CABIOS* 5: 151-153; Thompson 等人, 1994, 《核酸研究》(*Nucleic Acids Research*) 22: 4673-4680; Thompson 等人, 1997, 《核酸研究》(*Nucleic Acids Research*) 25: 4876-4882)使用 LASERGENE™ MEGALIGN™ 软件(DNASTAR, Inc., Madison, WI)测定，采用 blosum 权重矩阵和 0.03 到 0.05 的缺口延伸罚分以及如下的多重比对参数: 缺口罚分为 10 且缺口长度罚分为 10。成对比对参数为 Ktuple=1, 缺口罚分=3, 窗口=5, 而对角线=5。

优选的，本发明的多肽包含 SEQ ID NO: 2 氨基酸序列或其等位变体或其具有葡糖淀粉酶活性的片段。在更优选的实施方案中，本发明的多肽包含 SEQ ID NO: 2 氨基酸序列。在另一个优选的实施方案中，本发明的多肽包含 SEQ ID NO: 2 的 19 至 471 位氨基酸序列或其等位变体或其具有葡糖淀粉酶活性的片段。在另一个优选的实施方案中，本发明的多肽由 SEQ ID NO: 2 的 19 至 471 位氨基酸序列或其等位变体或其具有葡糖淀粉酶活性的片段组成。在另一个优选的实施方案中，本发明的多肽由 SEQ ID NO: 2 的氨基酸序列或其等位变异体或其具有葡糖淀粉酶活性的片段组成。

SEQ ID NO: 2 的“片段”是指从该氨基酸序列的氨基和/或羧基端缺失了一个或多个氨基酸的多肽。优选的，片段含有至少 400 个氨基酸残基，更优选至少 440 个氨基酸残基，并且最优选至少 450 个氨基酸残基。

“等位变体”是指两种或两种以上替代形式的基因占据相同的染色体基因座，等位变异通常由突变天然发生，并可导致群内的多样性。基因突变可以是沉默的（编码的多肽没有变化）或可能编码具有改变了的氨基酸序列的多肽。多肽的等位变体是由基因的等位变体编码的多肽。

在另一个实施方案中，本发明涉及重组生产葡糖淀粉酶的方法，所述方法包括在丝状真菌宿主细胞中表达可编码具有葡糖淀粉酶活性的多肽

(E.C. 3.2.1.3)的多核苷酸的步骤,其中多肽选自:(a)在极低严谨条件、优选低严谨条件、更优选中等严谨条件、更优选中-高严谨条件、甚至更优选高严谨条件、并且最优选极高严谨条件下与(i) SEQ ID NO: 1的1至2325位所示多核苷酸、(ii)介于SEQ ID NO: 1的1至2325位核苷酸之间的cDNA或无内含子的多核苷酸、(iii)至少100个核苷酸的(i)或(ii)的亚序列、或者(iv) (i)、(ii)或(iii)的互补链杂交的多核苷酸编码的多肽;(b)具有SEQ ID NO: 2氨基酸序列的多肽的变体,所说的变体包含一个或多个氨基酸的取代、缺失和/或插入;和(c) (a)或(b)中具有葡糖淀粉酶活性的片段(J. Sambrook, E. F. Fritsch 和 T. Maniatus, 1989,《分子克隆:实验室手册》(*Molecular Cloning, A Laboratory Manual*),第二版,冷泉港,纽约)。SEQ ID NO: 1亚序列可以是至少100个核苷酸或优选至少200个核苷酸。而且,该亚序列可以编码具有葡糖淀粉酶活性的多肽片段。所述多肽还可以是具有葡糖淀粉酶活性的多肽的等位变体或片段。

依据本领域熟知的方法,SEQ ID NO: 1核酸序列或其亚序列以及SEQ ID NO: 2的氨基酸序列或其片段,可用于设计核酸探针,以从不同的属或种中鉴定和克隆编码具有葡糖淀粉酶活性的多肽的DNA。具体地,可用此类探针与目的属或种的基因组或cDNA杂交,按照标准的Southern印迹操作,以鉴定和分离其中相应的基因。这些探针的长度可在相当大程度上短于全长序列,但应当至少15个、优选至少25个、且更优选至少35个核苷酸。也可以使用更长的探针。DNA和RNA探针都可以使用。典型地将所述探针标记以检测相应的基因(例如,用 ^{32}p 、 ^3H 、 ^{35}S 、生物素或抗生物素蛋白)。本发明涵盖此类探针。

这样,可从所述其它生物体制备的基因组DNA或cDNA文库中筛选能与上述探针杂交并编码具有葡糖淀粉酶活性的多肽的DNA。来自其它生物体的基因组或其它DNA可通过琼脂糖或聚丙烯酰胺电泳或其它分离技术分离。可将文库DNA或分离的DNA转移固定到硝酸纤维素或其它适合的载体材料上。为鉴定与SEQ ID NO: 1或其亚序列同源的克隆或DNA,在Southern印迹中使用所述载体。就本发明的目的而言,杂交是指所述核酸序列在从极低到极高严谨条件下,与对应于SEQ ID NO: 1所示的核酸

序列、其互补链、或其亚序列的标记核酸探针杂交。使用 X 射线胶片检测在这些条件下所述核酸探针与之杂交的分子。

在优选的实施方案中，所述核酸探针是编码 SEQ ID NO:1 多肽或其亚序列的核酸序列。在另一个优选的实施方案中，所述核酸探针是 SEQ ID NO:1 中的成熟多肽编码区。

对于长度是至少 100 个核苷酸的长探针来说，极低到极高严谨条件被定义为在 42°C、5X SSPE、0.3% SDS、200 µg/ml 剪切并变性的鲑精 DNA，以及对极低和低严谨条件 25% 甲酰胺、中和中 - 高严谨条件 35% 甲酰胺或高和极高严谨条件 50% 甲酰胺下，按照标准 Southern 印迹操作进行的预杂交和杂交。

对于长度至少 100 个核苷酸的长探针来说，载体材料最终使用 2 x SSC、0.2% SDS 优选在至少 45°C（极低严谨条件）、更优选在至少 50°C（低严谨条件）、更优选在至少 55°C（中严谨条件）、更优选在至少 60°C（中 - 高严谨条件）、甚至更优选在至少 65°C（高严谨条件）并且最优选在至少 70°C（极高严谨条件）下洗涤 3 次，每次 15 分钟。

对于长度约 15 个核苷酸到 70 个核苷酸的短探针，严谨条件定义为在低于按照 Bolton 和 McCarthy (1962, 《美国国家科学院进展》(Proceedings of the National Academy of Sciences USA)48: 1390)的计算法计算出的 T_m 值约 5°C - 约 10°C 下，在 0.9 M NaCl, 0.09 M Tris-HCl pH 7.6, 6 mM EDTA, 0.5% NP-40, 1X Denhardt's 溶液, 1 mM 焦磷酸钠, 1 mM 磷酸氢二钠, 0.1 mM ATP, 和每 ml 0.2 mg 酵母 RNA 中，根据标准 DNA 印迹操作进行预杂交、杂交和杂交后的洗涤。

对于长度约 15 个核苷酸到约 70 个核苷酸的短探针，载体材料在低于计算出的 T_m 值 5°C 到 10°C 下使用 6 x SCC 加 0.1% SDS 洗涤 15 分钟一次，并用 6 x SCC 洗涤两次各 15 分钟。

在优选的实施方案中，本发明具有葡糖淀粉酶活性的多肽包含淀粉结合结构域 (SBD)，并且优选所述淀粉结合结构域包含与 SEQ ID NO:2 的 483 到 579 位（包括二者）所示序列具有至少 80% 同一性的氨基酸序列。

另一个优选的实施方案涉及本发明的多肽，其中所述多肽包含位于淀

粉结合结构域和其余多肽之间的至少为 2 个氨基酸的接头。也可以优选所述多肽包含信号肽,例如信号肽优选包含与 SEQ ID NO: 2 的 1 到 18 位(包括二者)所示序列具有至少 95% 同一性的氨基酸序列。

在优选的实施方案中,本发明涉及的多肽变体具有这样的氨基酸序列,即包含了一个或多个氨基酸取代、缺失和/或插入的 SEQ ID NO: 2 的氨基酸序列中。

变体多肽的氨基酸序列可通过插入或缺失一个或多个氨基酸残基和/或由不同的氨基酸残基取代一个或多个氨基酸残基而与 SEQ ID NO: 2 或成熟多肽的氨基酸序列不同。优选的,氨基酸改变的是次要的特性,即不太影响蛋白的折叠和活性的保守氨基酸取代;小缺失,典型地 1 到大约 30 个氨基酸;小的氨基或羧基末端延伸;如氨基末端甲硫氨酸残基;长达约 20-25 个残基的小接头肽;或通过改变电荷或另一功能而便于纯化的小延伸;例如多聚组氨酸域,抗原表位或结合结构域。

保守取代的实例有碱性氨基酸(精氨酸、赖氨酸和组氨酸)、酸性氨基酸(谷氨酸和天门冬氨酸)、两性氨基酸(谷氨酰胺和天门冬酰胺)、疏水氨基酸(亮氨酸、异亮氨酸和缬氨酸)、芳族氨基酸(苯丙氨酸、色氨酸和酪氨酸)和小分子氨基酸(甘氨酸、丙氨酸、丝氨酸、苏氨酸和甲硫氨酸)组内的取代。一般不会改变比活的氨基酸取代是本领域公知的,并如 H. Neurath 和 R. L. Hill, 1979, 在《蛋白质》(*The Proteins*), Academic Press, New York 中所述。最常发生的调换是 Ala/Ser、Val/Ile、Asp/Glu、Thr/Ser、Ala/Gly、Ala/Thr、Ser/Asn、Ala/Val、Ser/Gly、Tyr/Phe、Ala/Pro、Lys/Arg、Asp/Asn、Leu/Ile、Leu/Val、Ala/Glu 和 Asp/Gly, 以及颠倒过来的这些调换。

在另一个实施方案中,本发明涉及分离的与具有 SEQ ID NO: 2 或其成熟多肽的氨基酸序列的多肽具有免疫化学相同性或部分免疫化学相同性的多肽。免疫化学特性通过公知的 Ouchterlony 双向免疫扩散操作经免疫交叉反应相同性实验进行测定。具体地,根据 Harboe 和 Ingild 在 N. H. Axelsen、J. Krøll 和 B. Weeks 编辑,《定量免疫电泳手册》(*A Manual of Quantitative Immunoelectrophoresis*), Blackwell Scientific Publications,

1973, 第 23 章中或 Johnstone 和 Thorpe 在《免疫化学实践》(*Immunochemistry in Practice*), Blackwell Scientific Publications, 1982 (更具体地是第 27-31 页)中所述的方法免疫接种兔子(或其它啮齿动物), 制备含有多克隆抗体的抗血清, 所述多克隆抗体与具有 SEQ ID NO: 2 或其成熟多肽的氨基酸序列的多肽的表位免疫反应或与之结合。具有免疫化学相同性的多肽是指以相同的方式与抗血清反应的多肽, 如沉淀物的完全融合, 相同的沉淀物形态, 和/或使用特定免疫化学技术的相同电泳迁移率。关于免疫化学相同性的进一步说明由 Axelsen、Bock 和 Kroll 在 N. H. Axelsen、J. Krøll 和 B. Weeks 编辑的《定量免疫电泳手册》, Blackwell Scientific Publications, 1973, 第 10-11 章中描述。具有部分免疫化学相同性的多肽是指以部分相同的方式与抗血清反应的多肽, 如沉淀物的部分融合, 部分相同的沉淀物形态, 和/或使用特定免疫化学技术的部分相同电泳迁移率。

抗体也可以是单克隆抗体。单克隆抗体可依据 E. Harlow 和 D. Lane 编辑, 1988, 《抗体: 实验室手册》, Cold Spring Harbor Press, Cold Spring Harbor, New York 的方法制备和使用。

本发明多肽具有至少 20%、优选至少 40%、更优选至少 60%、甚至更优选至少 80%、甚至更优选至少 90%、并且最优选至少 100%的 SEQ ID NO: 2 成熟多肽的葡糖淀粉酶活性。

本发明多肽可以得自任何属的微生物。就发明目的而言, 文中与给定来源结合使用的术语“获自”是指由核苷酸序列编码的多肽是由此来源或者由插入了来自此来源的核苷酸序列的细胞产生的。在优选的实施方案中, 多肽被分泌到细胞外。

本发明的多肽可以是细菌多肽。例如, 所述多肽可以是革兰氏阳性细菌多肽如芽孢杆菌属 (*Bacillus*) 多肽, 如, 嗜碱芽孢杆菌 (*Bacillus alkalophilus*)、解淀粉芽孢杆菌 (*Bacillus amyloliquefaciens*)、短芽孢杆菌 (*Bacillus brevis*)、环状芽孢杆菌 (*Bacillus circulans*)、凝结芽孢杆菌 (*Bacillus coagulans*)、*Bacillus lautus*、迟缓芽孢杆菌 (*Bacillus lentus*)、地衣芽孢杆菌 (*Bacillus licheniformis*)、巨大芽孢杆菌 (*Bacillus*

megaterium)、嗜热脂肪芽孢杆菌 (*Bacillus stearothermophilus*)、枯草芽孢杆菌 (*Bacillus subtilis*) 或苏云金芽孢杆菌 (*Bacillus thuringiensis*) 多肽; 或链霉菌属 (*Streptomyces*) 多肽, 如浅青紫链霉菌 (*Streptomyces lividans*) 或鼠灰链霉菌 (*Streptomyces murinus*) 多肽; 或者格兰氏阴性细菌多肽, 如大肠杆菌 (*E. coli*) 或假单胞菌属 (*Pseudomonas sp*) 多肽。

本发明的多肽可以是真菌多肽, 并且更优选为酵母多肽, 如念珠菌属 (*Candida*)、克鲁维酵母属 (*Kluyveromyces*)、毕赤氏酵母属 (*Pichia*)、糖酵母属 (*Saccharomyces*)、裂殖酵母属 (*Schizosaccharomyces*) 或假丝酵母属 (*Yarrowia*) 多肽; 或者更优选丝状真菌多肽如枝顶孢霉属 (*Acremonium*)、曲霉属 (*Aspergillus*)、短梗霉属 (*Aureobasidium*)、隐球菌属 (*Cryptococcus*)、*Filibasidium*、镰刀菌属 (*Fusarium*)、腐质霉属 (*Humicola*)、稻梨孢属 (*Magnaporthe*)、毛霉属 (*Mucor*)、*Neocallimastix*、链孢霉属 (*Neurospora*)、拟青霉属 (*Paecilomyces*)、青霉属 (*Penicillium*)、*Piromyces*、裂褶菌属 (*Schizophyllum*)、篮状菌属 (*Talaromyces*)、嗜热子囊菌属 (*Thermoascus*)、梭孢壳属 (*Thielavia*)、弯颈霉属 (*Tolyptocladium*)、或木霉属 (*Trichoderma*) 多肽。

在优选的实施方案中, 多肽是卡尔斯伯酵母 (*Saccharomyces carlsbergensis*)、酿酒酵母 (*Saccharomyces cerevisiae*)、糖化酵母 (*Saccharomyces diastaticus*)、*Saccharomyces douglasii*、克鲁弗酵母 (*Saccharomyces kluyveri*)、诺地酵母 (*Saccharomyces norbensis*) 或卵形酵母 (*Saccharomyces oviformis*) 多肽。

在另一个优选的实施方案中, 多肽是棘孢曲霉 (*Aspergillus aculeatus*) 入泡盛曲霉 (*Aspergillus awamori*) 入臭曲霉 (*Aspergillus foetidus*) 入日本曲霉 (*Aspergillus japonicus*) 入构巢曲霉 (*Aspergillus nidulans*) 入黑曲霉 (*Aspergillus niger*) 入米曲霉 (*Aspergillus oryzae*) 入杆孢状镰刀菌 (*Fusarium bactridioides*) 入谷类镰刀菌 (*Fusarium cereals*) 入 *Fusarium crookwellense*、大刀镰刀菌 (*Fusarium culmorum*) 入禾谷镰刀菌 (*Fusarium graminearum*) 入禾赤镰刀菌 (*Fusarium graminum*) 入异孢镰刀菌 (*Fusarium heterosporum*) 入合欢木镰刀菌 (*Fusarium negundi*) 入尖镰刀菌 (*Fusarium oxysporum*) 入

网状镰刀菌 (*Fusarium reticulatum*)、玫瑰色镰刀菌 (*Fusarium roseum*)、接骨木镰刀菌 (*Fusarium sambucinum*)、肤色镰刀菌 (*Fusarium sarcochromum*)、拟分枝镰刀菌 (*Fusarium sporotrichioides*)、硫色镰刀菌 (*Fusarium sulphureum*)、*Fusarium torulosum*、*Fusarium trichothecioides*、*Fusarium venenatum*、孤独腐质霉 (*Humicola insolens*)、柔毛腐质霉 (*Humicola lanuginosa*)、米赫毛霉 (*Mucor miehei*)、嗜热毁丝霉 (*Myceliophthora thermophila*)、粗糙链孢霉 (*Neurospora crassa*)、产紫青霉 (*Penicillium purpurogenum*)、哈茨木霉 (*Trichoderma harzianum*)、康氏木霉 (*Trichoderma koningii*)、长柄木霉 (*Trichoderma longibrachiatum*)、里氏木霉 (*Trichoderma reesei*) 或绿色木霉 (*Trichoderma viride*) 多肽。

在另一个优选的实施方案中，多肽是 *Thielavia achromatica*、*Thielavia albomyces*、*Thielavia albopilosa*、*Thielavia appendiculata*、*Thielavia arenaria*、*Thielavia australiensis*、基生梭孢壳 (*Thielavia basicola*)、*Thielavia californica*、*Thielavia fimeti*、*Thielavia fragilis*、*Thielavia heterothallica*、*Thielavia hyrcaniae*、*Thielavia kirilenkoeae*、*Thielavia kiwaitensis*、*Thielavia leptoderma*、*Thielavia microspora*、*Thielavia minuta*、*Thielavia octospora*、*Thielavia ovispora*、*Thielavia aperuviana*、瘤孢梭孢壳 (*Thielavia sepedonium*)、毛梭孢壳 (*Thielavia setosa*)、*Thielavia spirotricha*、*Thielavia subthermophila*、*Thielavia tanzanica*、*Thielavia terrestris*、*Thielavia terricola*、*Thielavia tetraspora*、*Thielavia thermophila*、*Thielavia tortuosa*、*Thielavia variospora* 或 *Thielavia wareingii* 多肽。

在更为优选的实施方案中，多肽是罗氏阿太菌多肽，例如，具有 SEQ ID NO: 2 的氨基酸序列的多肽。

应当理解对于前述物种而言，本发明涵盖了有性和无性两种阶段，以及其他分类学的等同物如无性型 (anamorphs)，而不管他们以何种种名著称。本领域的技术人员将会容易的识别正确等同物的身份。例如，梭孢壳属等同物的分类由 Morgan-Jones, 1974, 《加拿大植物学杂志》 (*Canadian Journal of Botany*) 52: 429-431 和 Glenn 等人, 1996, 《真菌学》 (*Mycologia*) 88: 369-38 定义。例如，*Thielavia terrestris* 的无性形态

被称为亚拉巴马枝顶孢霉 (*Acremonium alabamense*)。

这些物种的菌株公众可自众多的培养物保藏中心得到, 例如, 美国典型培养物保藏中心(ATCC)、德意志微生物和细胞培养物保藏中心 (DSM), 真菌菌种保藏中心(CBS)和农业研究服务专利培养物保藏中心、北方研究中心 (NRRL)。

进一步的, 使用上面提及的探针可从其他来源包括从自然界(如, 土壤、堆肥、水等)分离的微生物鉴定并得到这些多肽。从自然环境中分离微生物的技术是本领域公知的。然后通过类似地筛选另一个微生物基因组或 cDNA 文库可得到所述核酸序列。一旦用探针检测到编码多肽的核酸序列后, 则通过应用本领域普通技术人员公知的技术可以分离或克隆该序列(如参见 Sambrook 等人, 1989, 同上)。

如文中定义的, “分离的”多肽是指基本上没有其它非葡糖淀粉酶多肽的多肽, 例如, 如 SDS-PAGE 所测定的那样, 至少大约 20%纯, 优选至少大约 40%纯, 更优选至少大约 60%纯, 甚至更优选大约 80%纯, 最优选大约 90%纯, 并且甚至最优选大约 95% 纯。

由本发明核酸序列编码的多肽还包括融合多肽或可切割的融合多肽, 其中在多肽或其片段的 N-端或 C-端融合了另一多肽。融合多肽是通过将编码另一多肽的核酸序列(或其一部分)与本发明的核酸序列(或其一部分)融合而生产的。生产融合多肽的技术是本领域公知的, 并且包括连接编码所述多肽的编码序列, 从而他们在同一读框内, 且融合多肽的表达处于相同的启动子和终止子的控制之下。

核酸序列

本发明还涉及编码本发明多肽的分离的多核苷酸或核苷酸序列。在优选的实施方案中, 核酸序列如 SEQ ID NO :1 所示。在另一个优选的实施方案中, 核酸序列是 SEQ ID NO :1 中的成熟多肽编码区。本发明也涵盖这样的核酸序列, 其编码的多肽具有 SEQ ID NO :2 或其成熟多肽的氨基酸序列, 并由于遗传密码的简并而与 SEQ ID NO: 1 不同。本发明还涉及 SEQ ID NO: 1 的亚序列, 其编码具有葡糖淀粉酶活性的 SEQ ID NO : 2 片

段。

SEQ ID NO: 1 的亚序列是包含在 SEQ ID NO :1 核酸序列中,但在 5' 和/或 3 端'缺失一个或多个核苷酸。优选的,亚序列含有至少 1650 个核苷酸,更优选至少 1680 个核苷酸,并且最优选至少 1737 个核苷酸。

本发明还涉及在 SEQ ID NO :1 的成熟多肽编码序列中具有至少一个突变的突变核酸序列,其中该突变核酸序列编码的多肽由 SEQ ID NO: 2 的 19 至 471 位氨基酸组成。

用来分离或克隆编码多肽的核酸序列的技术是本领域公知的,并包括从基因组 DNA 分离、从 cDNA 制备或者其组合。例如,通过使用众所周知的聚合酶链式反应 (PCR) 或表达文库抗体筛选以检测具有同样结构特征的克隆 DNA 片段,可实现本发明核酸序列从此类基因组 DNA 中的克隆。如参见, Innis 等人,1990,《PCR: 方法及应用指南》(PCR: A Guide to Methods and Application), Academic Press, New York。可使用其他核苷酸扩增方法如连接酶链式反应 (LCR)、连接激活转录 (LAT) 和以核酸序列为基础的扩增 (NASBA)。可从梭孢壳属、阿太菌属菌株或另一或相关生物体中克隆所述核酸序列,例如,所述核酸可以是核酸序列中多肽编码区的等位或物种变体。

文中使用的术语“分离的核酸序列”是指基本上不含其他核酸序列的核酸序列,例如,如由琼脂糖电泳所测定的那样,至少大约 20% 纯,优选至少大约 40% 纯,更优选至少大约 60% 纯,甚至更优选至少大约 80% 纯,并且最优选至少大约 90% 纯。例如,通过遗传工程中所用的标准克隆操作可以得到分离的核酸序列,以便将所述核酸序列从其原始区域重新置于可复制的不同位点。所述克隆操作可包括切除和分离包含编码所述多肽的核酸序列的期望核酸片段,将该片段插入到载体分子中,并将重组载体整合进能复制多拷贝或克隆的核苷酸序列的宿主细胞中。核酸序列可以是基因组、cDNA、RNA、半合成或合成来源的,或其任意组合。

本发明还涉及这样的核酸序列,其与 SEQ ID NO :1 中成熟多肽编码序列的同一性程度为至少约 75%、优选约 80%、优选约 85%、更优选约 90%、甚至更优选约 95%、并且最优选约 97% 同一性,并且其编码活性葡

糖淀粉酶多肽。为了本发明的目的，通过 Wilbur-Lipman 法(Wilbur 和 Lipman, 1983, 《美国国家科学院学报》(*Proceedings of the National Academy of Science USA*) 80: 726-730)确定两核酸序列之间的同源性程度，利用 LASERGENE™ MEGALIGN™ 软件 (DNASTAR, Inc., Madison, WI)和同一性表及如下的多重比对参数: 缺口罚分为 10 且缺口长度罚分为 10。成对比对参数是 Ktuple=3, 缺口罚分=3, 而窗口=20。

修饰编码本发明多肽的核酸序列以合成与所述多肽基本上相似的多肽可能是必要的。术语与所述多肽“基本上相似”是指非天然存在的多肽形式。这些多肽可能在某些改造的方面与自然来源分离的多肽不同，如，比活、热稳定性、最适 pH 值等不同的变体。变体序列可以根据 SEQ ID NO: 1 的多肽编码部分呈现的核酸序列如其亚序列来构建，和/或通过引入不引起核苷酸序列编码另一种多肽氨基酸序列、但对应于旨在生产该酶的宿主生物的密码子使用的核苷酸取代来构建，或通过引入可产生不同氨基酸序列的核苷酸取代来构建。有关核苷酸取代的一般描述参见如 Ford 等人, 1991, 《蛋白表达和纯化》(*Protein Expression and Purification*) 2: 95-107。

对本领域技术人员显而易见的是，可以在分子的功能关键区之外进行这样的取代仍产生活性多肽。可依据本领域公知的方法鉴定对本发明的分离核苷酸序列编码的多肽的活性所必须的，因此优选不进行取代的氨基酸残基，如定点诱变或丙氨酸扫描诱变(参见 Cunningham 和 Wells, 1989, 《科学》(*Science*) 244: 1081-1085)。在后一种技术里，在分子里每一个带正电残基处引入突变，并且测定所得突变分子的葡糖淀粉酶活性以确定分子活性的关键氨基酸残基。还可通过对诸如核磁共振分析、晶体学、光亲和标记这样的技术测定的三维结构进行分析来确定底物-酶相互作用位点(参见如 de Vos 等人, 1992, 《科学》(*Science*) 255: 306-312; Smith 等人, 1992, 《分子生物学杂志》(*Journal of Molecular Biology*) 224 : 899-904; Wlodaver 等人, 1992, 《FEBS 通讯》(*FEBS Letters*) 309: 59-64)。

本发明还涉及编码本发明多肽的分离核酸序列，在极低严谨条件，优选低严谨条件，更优选中严谨条件，更优选中-高严谨条件，甚至更优选高严谨条件，并且最优选极高严谨条件下与这样的核酸探针杂交，所述核

酸探针在相同条件下与 SEQ ID NO :1 核酸序列或它的互补链、或其等位变体及亚序列(Sambrook 等人, 1989,同上)杂交, 如文中所定义的那样。

制备突变核酸序列的方法

本发明还涉及制备突变核苷酸序列的方法, 包括在 SEQ ID NO :1 的成熟多肽编码序列或其亚序列中引入至少一个突变, 其中所述突变核酸序列编码由 SEQ ID NO: 2 的 19 - 471 位氨基酸组成的多肽或其具有葡糖淀粉酶活性的片段。

使用本领域公知的任何方法通过定点诱变可以实现在核苷酸序列中引入突变, 以将一个核苷酸调换为另一个核苷酸。特别有用的是利用超螺旋、具有目的插入物的双链 DNA 载体以及包括期望突变的两个合成引物的操作。所述寡核苷酸引物分别与载体的对应链一一互补, 通过 *Pfu* DNA 聚合酶在温度循环期间延伸。引物一经掺入, 就生成了含有粘性缺口的突变质粒。随着温度循环, 用特异于甲基化和半甲基化 DNA 的 *DpnI* 处理产物, 以消化亲代 DNA 模板并选择含有突变的合成 DNA。还可以使用本领域公知的其他操作。

核酸构建体

本发明还涉及包含有效连接于一个或多个控制序列的本发明核酸序列的核酸构建体, 所述控制序列在与控制序列相容的条件指导所述编码序列在适合的宿主细胞中表达。表达应理解为包括生产多肽所涉及的任何步骤, 包括但不限于: 转录、转录后修饰、翻译、翻译后修饰和分泌。

“核酸构建体”在文中定义为一核酸分子, 既可以是单链的也可以是双链的, 其分离自天然产生的基因, 或已被修饰以含有否则自然界将不会存在的方式组合和毗邻的核酸片段。当核苷酸建构体包含表达本发明编码序列所必需的所有控制序列时, 术语核酸构建体与术语表达盒同义。术语“编码序列”在文中定义为直接规定其蛋白质产物的氨基酸序列的核酸序列。基因组编码序列的边界一般由位于 mRNA 的 5'端紧邻开放阅读框上游的核糖体结合位点(原核生物)或 ATG 起始密码子(真核生物)以及

mRNA 的 3'端紧邻开放阅读框下游的转录终止子决定。编码序列包括但不限于 DNA、cDNA 和重组核酸序列。

可以通过多种方式操作编码本发明多肽的分离核酸序列以保证多肽的表达。取决于表达载体，在将核酸序列插入载体之前对其操作可能是优选的或必须的。利用重组 DNA 方法修饰核酸序列的技术是本领域熟知的。

术语“控制序列”在文中定义为包括表达本发明多肽所必需的或对其有利的所有组件。编码多肽的核酸序列的每一个控制序列可以是天然的或外源的。这样的控制序列包括，但不限于：前导序列、多聚腺苷酸序列、肽原序列、启动子、信号肽序列和转录终止子。控制序列至少包括启动子、转录和翻译终止信号。控制序列可具有以引入特异性限制位点的接头，便于控制序列与编码多肽的核酸序列的编码区连接。术语“有效连接”在文中定义为适当地将控制序列放置在相对于 DNA 序列中编码序列的某个位置上，从而控制序列可指导多肽表达的一种结构。

控制序列可以是合适的启动子序列，即被宿主细胞识别用于表达所述核酸序列的一种核酸序列。启动子序列含有介导多肽表达的转录控制序列。启动子可以在所选宿主细胞中有转录活性的任何核酸序列，包括突变、截短和杂合启动子，并且所述启动子可得自与所述宿主细胞同源或异源的编码胞外或胞内多肽的基因。

用于指导本发明核酸构建体在丝状真菌宿主细胞中转录的合适启动子的实例是得自米曲霉 TAKA 淀粉酶、米赫根毛霉 (*Rhizomucor miehei*) 天门冬氨酸蛋白酶、黑曲霉中性 α -淀粉酶、黑曲霉酸稳定的 α -淀粉酶、黑曲霉或者泡盛曲霉葡糖淀粉酶(*glaA*)、米赫根毛霉脂肪酶、米曲霉碱性蛋白酶、米曲霉磷酸丙糖异构酶、构巢曲霉乙酰胺酶和尖孢镰孢霉胰蛋白酶样蛋白酶(WO 96/00787)基因的启动子，还有 NA2-tpi 启动子 (来自黑曲霉中性 α -淀粉酶和米曲霉磷酸丙糖异构酶基因的杂合启动子)，及其突变、截短和杂合启动子。

控制序列还可以是合适的转录终止子序列，即可被宿主细胞识别以终止转录的序列。终止子序列有效连接于编码所述多肽的核酸序列的 3' 端。在所选宿主细胞中有功能的任何终止子都可用于本发明。

用于丝状真菌宿主细胞的优选终止子来自米曲霉 TAKA 淀粉酶、黑曲霉葡糖淀粉酶、构巢曲霉邻氨基苯甲酸合酶、黑曲霉 α -葡萄糖苷酶和尖孢镰孢霉胰蛋白酶样蛋白酶的基因。

所述控制子还可以是合适的前导序列,即 mRNA 上对宿主细胞翻译很重要的非翻译区。前导序列有效连接于编码所述多肽的核酸序列的 5'端。在所选宿主细胞中有功能的任何前导序列都可用于本发明。

丝状真菌宿主细胞的前导序列优选来自米曲霉 TAKA 淀粉酶和构巢曲霉磷酸丙糖异构酶的基因。

控制序列还可以是多聚腺苷酸序列,即有效连接于核酸序列 3'端的序列,并且当转录时,其被宿主细胞识别为向转录的 mRNA 中添加多聚腺苷酸残基的信号。在所选宿主细胞有功能的任何多聚腺苷酸序列都可以用于本发明。

丝状真菌宿主细胞的多聚腺苷酸优选来自米曲霉 TAKA 淀粉酶、黑曲霉葡糖淀粉酶、构巢曲霉邻氨基苯甲酸合酶、黑曲霉 α -葡萄糖苷酶和尖孢镰孢霉胰蛋白酶样蛋白酶的基因。

控制序列还可以是信号肽编码区,其编码与多肽氨基末端连接的氨基酸序列并指导所编码的多肽进入细胞的分泌路径。核酸序列 5'端的编码序列可以固有地包含一个信号肽编码区,在翻译阅读框内自然地与编码分泌多肽的编码区段连接。或者,5'的编码序列可以包含对于编码序列是外源的信号肽编码区。在编码序列并不天然地包含信号肽编码区时,可能需要外源信号肽编码区。或者,外源信号肽编码区可以简单地代替天然的信号肽编码区,以加强所述多肽的分泌。但是,指导表达的多肽进入所选宿主细胞的分泌途径的任何信号肽编码区都可用于本发明。

丝状真菌宿主细胞的有效信号肽编码区有来自米曲霉 TAKA 淀粉酶、黑曲霉中性淀粉酶、黑曲霉葡糖淀粉酶、米赫根毛霉天门冬氨酸蛋白酶、孤独腐质霉纤维素酶和柔毛腐质霉 (*Humicola lanuginosa*) 脂肪酶的信号肽编码区。

在优选的实施方案中,信号肽编码区是 SEQ ID NO: 1 的 1-54 位核苷酸,其编码 SEQ ID NO: 2 的 1-18 位氨基酸;优选信号肽包含的氨基

酸序列与 SEQ ID NO: 2 的 1-18 位 (包括二者) 所示的序列具有至少 95 % 同一性。

控制序列还可以是编码位于多肽氨基末端的氨基酸序列的肽原编码区。所得多肽称之为酶原 (proenzyme) 或多肽原 (或者, 在某些情况下为酶原 (zymogen))。多肽原一般无活性, 且经从多肽原上催化或自身催化切割肽原, 可转化为成熟有活性的多肽。肽原编码区可以来自枯草芽孢杆菌碱性蛋白酶(*aprE*)、枯草芽孢杆菌中性蛋白酶(*nprT*)、酿酒糖酵母 α -因子、米赫根毛霉天门冬氨酸蛋白酶和嗜热毁丝霉漆酶(WO 95/33836)的基因。

当信号肽和肽原区都存在于多肽的氨基末端时, 肽原区与多肽氨基末端紧邻, 而信号肽区与肽原区的氨基末端紧邻。

可能也需要加入相对于宿主细胞生长而调节多肽的表达的调控序列。调控系统的实例是应答化学或物理刺激 (包括调控化合物的存在) 而引起基因表达的开启或关闭的那些系统。丝状真菌中的调控系统有 TAKA α -淀粉酶启动子、黑曲霉葡糖淀粉酶启动子, 并且米曲霉葡糖淀粉酶启动子可作为调控序列。调控序列的其它实例是允许基因扩增的那些序列。在真核系统里, 这些包括在甲胺蝶呤存在下扩增的二氢叶酸还原酶基因和有重金属时扩增的金属硫蛋白基因。在这些情况下, 编码多肽的核苷酸序列可操作的与调控序列连接。

本发明还涉及用以改变编码本发明的多肽的内源基因的表达的核酸构建体。所述构建体可以包含为改变内源基因的表达所必需最少数量的组分。在一个实施方案中, 核酸构建体优选包含 (a) 靶向序列、(b) 调控序列、(c) 外显子和 (d) 剪接供体位点。所述核苷酸构建体一经引入细胞, 构建体就通过同源重组在内源基因位点处插入到细胞基因组中。靶向序列指导元件 (a)-(d) 整合进内源基因中, 从而元件 (b)-(d) 有效地与内源基因连接。在另一个实施方案中, 核酸构建体包括 (a) 靶向序列、(b) 调控序列、(c) 外显子、(d) 剪接供体位点、(e) 内含子和 (f) 剪接受体位点, 其中靶向序列指导元件 (a)-(f) 的整合, 从而元件 (b)-(f) 与内源基因有效连接。不过, 构建体可以含有附加组分如选择性标记。

在这两个实施方案中，引入这些组分导致产生一个改变了内源基因表达的新转录单元。本质上，新转录单元是靶向构建体所引入的序列和内源基因的融合产物。在一个实施方案中，改变了内源基因，该基因被活化。在此实施方案中，通过插入导致基因明显地比相应的亲本细胞更高水平表达的调控序列，利用同源重组取代、破坏或失活与亲本细胞内源基因正常结合的调控区。使用本领域熟知的方法，通过在构建体中纳入可扩增的选择性标记基因，可进一步扩增活化的基因（参见如，美国专利号 No. 5,641,670）。在另一个实施方案中，改变内源基因，该基因的表达降低。

靶向序列可以在内源基因之内、直接邻近于所述基因、在上游基因之内或在内源基因的上游以及较远的距离。可使用一个或多个靶向序列。例如，环形质粒或 DNA 片段优选采用单个靶向序列，而线性质粒或 DNA 片段优选采用两个靶向序列。

构建体的调控序列可包含一个或多个启动子、增强子、骨架粘结区或基质粘附位点、负调控元件、转录结合位点或这些序列的组合。

构建体还包含所述内源基因的一个或多个外显子。外显子定义为复制成 RNA 并出现在成熟 mRNA 分子中的 DNA 序列，从而外显子序列与内源基因的编码区读框一致。外显子可以任选地含有编码一个或多个氨基酸和/或部分编码一个氨基酸的 DNA。或者，外显子含有与 5'非编码区相对应的 DNA。当外源外显子或外显子编码一个或多个氨基酸和/或一个氨基酸的一部分时，这样设计核酸构建体，从而经转录和剪接，阅读框与外源基因的编码区一致，使得不改变来自第二外显子的 mRNA 部分的正确阅读框。

构建体的所述剪接供体位点指导一个外显子剪接到另一个外显子。典型地，第一外显子位于第二外显子的 5'，而与第一外显子 3'端重叠且侧接的剪接供体位点识别侧接在第二外显子 5'端的剪接受体位点。剪接受体位点，与剪接供体位点一样，是指导一个外显子剪接到另一个外显子的序列。与剪接供体位点联合作用，剪接机器利用剪接受体位点实现内含子的切除。

表达载体

本发明还涉及包含本发明的核酸序列、启动子、转录和翻译终止信号的丝状真菌细胞重组表达载体。可以将上述种种核酸和控制序列连接在一起产生可包括一个或多个便利的限制位点的重组表达载体，以允许编码所述多肽的核酸序列在此位点处的插入或取代。或者，通过将核酸序列或包含所述序列的核酸构建体插入适宜的表达载体可表达本发明的所述核酸序列。在创建表达载体的过程中，将编码序列置于载体中与适当的控制序列有效连接而用于表达。

重组表达载体可以是任何载体（如，质粒或病毒），其能方便地进行重组 DNA 操作并能使核酸序列表达。载体的选择将典型地取决于载体与引入该载体的宿主细胞之间的相容性。所述载体可为线性或闭合的环形质粒。

载体可以是自主复制的载体，即，以染色体外实体形式存在的载体，其复制独立于染色体复制，如，质粒、染色体外元件、小染色体或人工染色体。所述载体可以包括任何保证自我复制的器件。或者，所述载体可以是这样的一种载体，当引入宿主细胞时，其被整合进基因组并与整合了该载体的染色体一起复制。此外，可使用单个载体或质粒，或一起包含了待引入宿主细胞基因组的总 DNA 的两个或两个以上的载体或质粒，或者转座子。

本发明所述载体优选含有一个或以上选择性标记，其允许易于选择转化细胞。选择性标记是这样的基因，其产物提供生物杀灭剂或病毒抗性、重金属抗性、原养型向营养缺陷型的转变等等。在丝状真菌宿主细胞中使用的选择性标记包括，但不限于：*amdS*（乙酰氨酶）、*arg8*（鸟氨酸氨甲酰转移酶）、*bar*（磷丝菌素乙酰转移酶）、*hph*（潮霉素磷酸转移酶）、*niaD*（硝酸还原酶）、*pyrG*（乳清苷-5'-磷酸脱羧酶）、*sC*（硫酸腺苷酰转移酶）和 *trpC*（邻氨基苯甲酸合成酶）及其等同物。优选在曲霉属细胞中使用构巢曲霉或米曲霉的 *amdS* 和 *pyrG* 基因，以及吸水链霉菌（*Streptomyces hygroscopicus*）的 *bar* 基因。

本发明的载体优选含有使得载体整合进宿主细胞基因组或载体在细胞内独立于基因组自我复制的元件。

为整合进宿主细胞基因组，载体可依赖于编码多肽的核酸序列或载体

所其它任何元件，通过同源或非同源重组将所述载体整合进载体。或者，所述载体可含有指导通过同源重组整合进宿主细胞基因组的其它核酸序列。所述其它核苷酸序列能在染色体的精确位置将载体整合进宿主细胞的基因组。为提高在精确位置整合的可能性，整合元件应当优选包括足够量的核酸，如 100 到 10,000 碱基对、优选 400 到 10,000 碱基对、并且最优选 800 到 10,000 碱基对，它们与相应靶序列高度同源以提高同源重组的可能性。所述整合元件可以是与宿主细胞基因组的靶序列同源的任何序列。另一方面，可以通过非同源重组将载体整合进宿主细胞的基因组中。

为进行自主复制，所述载体还可包含能使载体在所研究的宿主细胞里自主复制的复制起点。Ama1-序列适用于丝状真菌宿主细胞。

可将本发明核酸序列的一个以上的拷贝插入宿主细胞以提高基因产物的产量。通过在宿主细胞基因组中整合至少另外一个拷贝的序列，或者在细胞含有可扩增拷贝的选择性标记基因的情况下，通过在核酸序列中纳入选择性标记基因，可以实现核酸序列拷贝数的增加，由此可以通过在合适的选择性试剂的存在下培养细胞来选择另外拷贝的核酸序列。

用于连接上述元件以构建本发明重组表达载体的方法是本领域技术人员熟知的(参见如, Sambrook 等人, 1989, 同上)。

宿主细胞

本发明还涉及重组宿主细胞，包含编码本发明葡糖淀粉酶多肽的核酸序列，其可有利地在多肽的重组生产中使用。将包含本发明的核酸序列的载体引入进宿主细胞，因此所述载体作为染色体构成部分或如前述作为自我复制的染色体外载体得以维持。术语“宿主细胞”包括任何因发生在复制期间的突变而与亲本细胞不一致的亲本细胞的子代。宿主细胞选择的在很大程度上取决于编码多肽的基因及其来源。

在优选的实施方案中，宿主细胞是真菌细胞。文中所用的“真菌”包括子囊菌门 (*Ascomycota*)、担子菌门 (*Basidiomycota*)、壶菌门 (*Chytridiomycota*) 和接合菌门 (*Zygomycota*) (如 Hawksworth 等人在《Ainswoffh 和 Bisby 氏真菌词典》(*Ainswoffh and Bisby's Dictionary of*

The Fungi) 中所定义的那样, 第 8 版, 1995, CAB 国际, 大学出版社, Cambridge, UK) 以及卵菌门(*Oomycota*) (如 Hawksworth 等人, 1995, 同上, 第 171 页所引述的那样) 以及所有的有丝分裂真菌真菌 (*mitosporic fungi*) (Hawksworth 等, 1995, 同上)。

在另一个更为优选的实施方案中, 所述真菌宿主细胞是丝状真菌细胞。

“丝状真菌”包括真菌亚门 (*Eumycota*) 和卵菌亚门 (*Oomycota*) (如 Hawksworth 等人所定义的那样, 1995, 同上) 的所有丝状形式。丝状真菌一般特征是由壳质素、纤维素、葡聚糖、壳聚糖、甘露聚糖以及其它复合多糖组成的菌丝体壁。通过菌丝延长进行营养生长, 且碳分解代谢是专性需氧型的。相反地, 酵母如酿酒酵母的是通过叶状单细胞萌芽进行营养生长, 并且碳分解代谢可能是发酵型的。

在甚至更为优选的实施方案中, 真丝宿主细胞是, 但不限于: 枝顶孢霉属 (*Acremonium*)、曲霉属 (*Aspergillus*)、镰刀菌属 (*Fusarium*)、腐质霉属 (*Humicola*)、毛霉属 (*Mucor*)、毁丝霉属 (*Myceliophthora*)、链孢霉属 (*Neurospora*)、青霉属 (*Penicillium*)、梭孢壳属 (*Thielavia*)、弯颈霉属 (*Tolyocladium*) 或木霉属 (*Trichoderma*) 细胞。

在最为优选的实施方案中, 丝状真菌宿主细胞是泡盛曲霉 (*Aspergillus awamori*)、臭曲霉 (*Aspergillus foetidus*)、日本曲霉 (*Aspergillus japonicus*)、构巢曲霉 (*Aspergillus nidulans*)、黑曲霉 (*Aspergillus niger*) 或米曲霉 (*Aspergillus oryzae*) 细胞。在另一最为优选的实施方案中, 丝状真菌宿主细胞是杆孢状镰刀菌 (*Fusarium bactridioides*)、谷类镰刀菌 (*Fusarium cerealis*)、*Fusarium crookwellense*、大刀镰刀菌 (*Fusarium culmorum*)、禾谷镰刀菌 (*Fusarium graminearum*)、禾赤镰刀菌 (*Fusarium graminum*)、异孢镰刀菌 (*Fusarium heterosporum*)、合欢木镰刀菌 (*Fusarium negundi*)、尖镰刀菌 (*Fusarium oxysporum*)、网状镰刀菌 (*Fusarium reticulatum*)、玫瑰色镰刀菌 (*Fusarium roseum*)、接骨木镰刀菌 (*Fusarium sambucinum*)、肤色镰刀菌 (*Fusarium sarcochroum*)、拟分枝镰刀菌 (*Fusarium sporotrichioides*)、硫色镰刀菌 (*Fusarium sulphureum*)、*Fusarium torulosum*、*Fusarium trichothecioides* 或 *Fusarium venenatum* 细

胞。在甚至最优选的实施方案中，丝状真菌亲本细胞是 *Fusarium venenatum* (Nirenberg 新种) 细胞。在另一最为优选的实施方案中，丝状真菌宿主细胞是孤独腐质霉 (*Humicola insolens*)、柔毛腐质霉 (*Humicola lanuginosa*)、米赫毛霉 (*Mucor miehei*)、嗜热毁丝霉 (*Myceliophthora thermophila*)、粗糙链孢霉 (*Neurospora crassa*)、产紫青霉 (*Penicillium purpurogenum*)、哈茨木霉 (*Trichoderma harzianum*)、康氏木霉 (*Trichoderma koningii*)、长柄木霉 (*Trichoderma longibrachiatum*)、里氏木霉 (*Trichoderma reesei*) 或绿色木霉 (*Trichoderma viride*) 细胞。

真菌细胞可以其本身已知的方式通过涉及原生质体形成、原生质体转化以及细胞壁再生的过程转化。转化曲霉菌属宿主细胞合适的方法描述于 EP 238 023 和 Yelton 等人, 1984, 《美国国家科学院学报》(*Proceedings of the National Academy of Science USA*) 81: 1470-1474 中。转化镰刀菌属合适的方法描述于 Malardier 等人, 1989, 《基因》(*Gene*) 78: 147-156 以及 WO 96/00787 中。

生产方法

本发明还涉及重组生产葡糖淀粉酶的方法，所述方法包括在丝状真菌宿主细胞中表达编码具有葡糖淀粉酶活性的多肽(E.C. 3.2.1.3)的多核苷酸的步骤，并且还包括(a)在利于生产多肽的条件下培养宿主细胞；和(b)回收所述多肽。

本发明还涉及生产本发明多肽的方法，包括 (a) 在利于生产多肽的条件下培养宿主细胞，其中所述宿主细胞包含在 SEQ ID NO: 1 成熟多肽编码区中具有至少一个突变的突变核酸序列，其中所述突变核酸序列编码由 SEQ ID NO: 2 的 19 至 471 位氨基酸组成的多肽，以及(b)回收所述多肽。

本发明还涉及生产本发明多肽的方法，包括(a) 在利于生产多肽的条件下培养同源重组细胞，其中整合了新的转录单元，所述转录单元包括调控序列、外显子和/或剪接供体位点，所述剪接供体位点与编码所述多肽的内源核酸序列的第二外显子有效连接；以及 (b)回收所述多肽。该方法基于基因激活技术的应用，例如，如美国专利号 No. 5,641,670 中所述。

在本发明的生产方法中，使用本领域公知的方法在适于生产所述多肽的营养基质中培养细胞。例如，可在合适的培养基中和允许多肽表达和/或分离的条件下，通过摇瓶培养和实验室或工业发酵罐中进行的小规模或大规模发酵（包括连续、批次、流加批次或固态发酵）培养细胞。在包括碳和氮源以及无机盐的合适营养基质中，使用本领域公知的方法进行所述培养。合适的培养基可由商业供应商获得或根据已公开的组成成分（如，美国典型培养物保藏中心培养目录）制备。如果多肽分泌到营养基质中，可以从培养基中直接回收所述多肽。如果多肽没有分泌，则可以从细胞裂解物中回收。

可以使用本领域公知的特异于所述多肽的方法检测多肽。这些检测方法可以包括特异性抗体的应用、酶产物的形成或酶底物的消失。例如，酶分析可以用于测定文中所述多肽的活性。

使用本领域公知的纯化和/或回收方法可以纯化或回收所得多肽。例如，通过常规方法包括，但不限于离心、过滤、萃取、喷雾干燥、蒸发或沉淀可从营养基质中回收所述多肽。

使用本领域公知的多种纯化可以纯化本发明所述多肽，包括，但不限于：色谱法（如，离子交换、亲和、疏水、聚焦层析和大小排阻）、电泳方法（如，制备等电点聚焦）、差示溶解度（如，硫酸铵沉淀）、SDS-PAGE 或萃取（如参见，《蛋白质纯化》（*Protein Purification*），J. -C. Janson 和 Lars Ryden 编辑，VCH Publishers, New York, 1989）。

用途

本发明也涉及使用具有葡糖淀粉酶活性的多肽的方法。所属多肽可用于淀粉转化方法，特别是用于制备葡萄糖和果糖糖浆（美国专利 US3,912,590）以及挤压淀粉的预处理（如米和麦面条的挤压）、生产低碳水化合物含量的啤酒（Manners, *The Brewers Digest*, 1974 年 12 月, 56）、来自生淀粉发酵的酒精（DE 3638529 C），以及根据本领域已确立的方法造酒。本发明的葡糖淀粉酶还可用于清洁剂，包括洗衣清洁剂、洗盘清洁剂和硬表面清洗组合物。本发明葡糖淀粉酶的其他用途是混合肥料和生物废

物处理、用于食品添加剂的植物提取物的纯化、化妆品和药品、烘焙工业（包括面包和蛋糕生产）、寡糖生产方法、燃料酒精和饮用酒精（便携式酒精）生产方法、饮料生产、生产有机化合物的发酵方法（如柠檬酸、抗坏血酸、赖氨酸、谷氨酸）、生产专用糖浆的方法以及宠物食品生产中的用途。

本发明所述葡糖淀粉酶变体还以固定形式使用。这适于并常用于生产专用糖浆如麦芽糖糖浆，并且还用于与果糖糖浆生产有关的寡糖残液流（raffinate stream）。

酒精生产

在一个实施方案中，发明的方法可以是包括如下步骤的酒精生产方法，其中在预糖化和/或发酵过程中加入植酸酶活性。应当理解可以在酵母细胞繁殖过程和/或之后的实际发酵过程中加入根据本发明的葡糖淀粉酶。同样地考虑饮料生产如啤酒或葡萄酒的生产。

从全谷生产酒精，特别是生产乙醇，可以分为4个主要步骤：

- 粉碎
- 液化
- 糖化
- 发酵

粉碎

在一个实施方案中，粉碎（全）谷是为了打开结构并利于进一步加工。依据本发明优选两种工艺：湿磨和干磨。乙醇生产优选干磨，其中磨出全仁并用于工艺的其它部分。还可使用湿磨，其允许很好地分离种子和粗磨粉（淀粉颗粒和蛋白质），而且除了少数例外，可应用于并行生产糖浆的场所。干磨和湿磨都是例如乙醇生产领域所熟知的。

液化

在本发明液化步骤的实施方案中，将磨成胶状的全谷原料分解（水解）成大多数DE高于4的麦芽糖糊精（糊精）。所述水解通过酸处理进行或通

过 α -淀粉酶酶促进行,具体地,是用如在下文将要进一步描述的芽孢杆菌 α -淀粉酶。酸水解在有限的基础上使用。本发明一个实施方案中的粉碎原料是(全)谷。不过也可以使用来自淀粉加工的侧流(side stream)。

在本发明的一个实施方案中,酶促液化作为三步热浆法进行。将浆加热到60-95°C之间,优选80-85°C,并添加酶以起始液化(稀释),优选至少添加 α -淀粉酶。接着将浆喷射加热到温度95-140°C之间、优选105-125°C以完成浆的糊化。接着将浆冷却到60-95°C,并添加更多的酶结束水解(二次液化)。液化工艺在pH 4.5-6.5进行,特别是pH 5-6之间。磨碎和液化的全谷称之为醪液。

液化步骤可以在得自微生物或植物的 α -淀粉酶的存在下进行。优选的 α -淀粉酶源自真菌或细菌。芽孢杆菌 α -淀粉酶(常称之为“高温淀粉酶样- α -淀粉酶”),其变体和杂合体是根据本发明特别考虑的。已知的高温淀粉酶样- α -淀粉酶包括源自地衣芽孢杆菌(*B. licheniformis*)(作为TermamylTM可由商业上获得)、解淀粉芽孢杆菌(*B. amyloliquefaciens*)菌株的 α -淀粉酶,以及嗜热脂肪芽孢杆菌(*B. stearothermophilus*)的 α -淀粉酶(BSG)。其他高温淀粉酶样- α -淀粉酶包括源自芽孢杆菌属菌株NCIB 12289、NCIB 12512、NCIB 12513或DSM 9375的 α -淀粉酶,它们全部都详细描述在WO 95/26397中,以及由Tsukamoto等人,《生物化学和生物物理研究通讯》(Biochemical and Biophysical Research Communications),151(1988),第25-31页描述的 α -淀粉酶。在本发明上下文中,高温淀粉酶样- α -淀粉酶是如WO99/19467第3页、第18行到第6页、第27行所定义的 α -淀粉酶。设想的变体和杂合体在WO 96/23874、WO 97/41213和WO 99/19467中描述。设想的 α -淀粉酶得自曲霉属菌株,包括米曲霉和黑曲霉 α -淀粉酶。商业 α -淀粉酶产品和含有 α -淀粉酶的产品包括TERMAMYLTM SC、FUNGAMYLTM、LIQUOZYMETM和SANTM SUPER。

真菌 α -淀粉酶可以按0.001-1.0 AFAU/g DS的量添加,优选0.002-0.5 AFAU/g DS,优选0.02-0.1 AFAU/g DS。

芽孢杆菌 α -淀粉酶可以按本领域技术人员众所周知的有效量添加。

糖化

为生产能够由酵母代谢的低分子糖 $DP_{1,3}$ ，液化的麦芽糖糊精必须进一步水解。水解典型地用葡糖淀粉酶酶促进行，或者，可使用 α -葡萄糖苷酶或酸性 α -淀粉酶。完整的糖化步骤可持续 72 小时，然而，一般仅进行典型的 40-90 分钟的预糖化，随后在发酵期间完成糖化(SSF)。糖化典型地在 30-65℃，典型地在 60℃ 左右并在 pH 4.5 时进行。

发酵

向醪液中添加典型地来自糖酵母属 (*Saccharomyces* spp.) 的酵母，并进行 24-96 小时的发酵，例如典型的 35-60 小时。温度介于 26-34℃ 之间，特别是约 32℃，且 pH 3-6，优选 pH 4-5 附近。

注意最广泛使用的工艺是同步糖化和发酵(SSF)工艺，其中没有保持糖化的阶段，这意味着发酵生物体如酵母，以及酶是一起加入的。当进行 SSF 时，一般恰好在发酵之前，在温度 50℃ 以上引入预糖化步骤。

蛋白酶添加提高 FAN (游离氨基氮)水平，并提高酵母的代谢速率，并进一步提供更高的发酵效率。合适的蛋白酶包括真菌和细菌蛋白酶。优选蛋白酶是酸性蛋白酶，即，在低于 pH7 的酸性条件下蛋白酶表现出水解蛋白质的能力。合适的酸性真菌蛋白酶包括得自曲霉属 (*Aspergillus*)、毛霉属 (*Mucor*)、根霉属 (*Rhizopus*)、念珠菌属 (*Candida*)、革盖菌属 (*Coriolus*)、内座壳属 (*Endothia*)、*Enthomophtra*、耙齿菌属 (*Irpex*)、青霉属 (*Penicillium*)、麦角菌属 (*Sclerotium*) 和球拟酵母属 (*Torulopsis*) 的真菌蛋白酶。更优选的是得自黑曲霉(如参见 Koaze 等人, (1964), 《日本农业生物化学》(Agr. Biol. Chem. Japan), 28, 216)、*Aspergillus saitoi* (如参见 Yoshida, (1954) 《日本农业化学学会杂志》(J. Agr. Chem. Soc. Japan), 28, 66)、泡盛曲霉 (*Aspergillus awamori*) (Hayashida 等人, (1977) 《农业生物化学》(Agric. Biol. Chem.), 42 (5), 927-933)、棘孢曲霉 (*Aspergillus aculeatus*) (WO 95/02044)或米曲霉的蛋白酶; 以及来自微小毛霉 (*Mucor pusillus*) 或米赫毛霉 (*Mucor miehei*) 的酸性蛋白酶。不是酸性蛋白酶的

细菌蛋白酶，包括商业上可获得的产品 Alcalase®和 Neutrase® (可由 Novozymes A/S 获得)。在一个实施方案中，可按 10^{-7} 到 10^{-5} 克活性蛋白酶蛋白/g DS 的量添加蛋白酶，特别是 10^{-7} 到 5×10^{-6} 克活性蛋白酶蛋白/g DS。

蒸馏

任选地在发酵之后，可通过蒸馏醪液提取例如乙醇。在终产物是根据本发明的工艺获得的乙醇的情况下，它可以用于，例如，燃料酒精、饮用酒精如便携式中性酒精或工业酒精。

淀粉转化

本发明提供了使用本发明的葡糖淀粉酶由淀粉生产葡萄糖等的方法。一般地，该方法包括这样的步骤：在 α -淀粉酶的存在下部分水解淀粉前体，随后在葡糖淀粉酶存在下通过切割 α - (1, 4) 和 α - (1, 6) 糖苷键由淀粉或相关寡糖及多聚糖分子的非还原端进一步水解释放 D-葡萄糖。

利用 α -淀粉酶对淀粉前体的部分水解，通过水解内在 α - (1, 4) - 键对淀粉分子提供了最初的分解。在商业应用上，使用 α -淀粉酶的初始水解是在大约 105°C 的温度下进行的。处理极高的淀粉浓度，通常为 30% 到 40% 的固体。初始水解通常在这种升高的温度下进行 5 分钟。部分水解的淀粉随后转移到第二个罐中，并在 $85 - 98^{\circ}\text{C}$ 的温度下温育大约 1-2 小时，以得到 10 到 15 的葡萄糖当量(D.E.)。

在葡糖淀粉酶存在下由淀粉或相关寡糖及多聚糖分子的非还原端进一步水解释放 D-葡萄糖是在独立的罐中、在 $30 - 62^{\circ}\text{C}$ 之间的降低的温度下进行的。底物液体的温度优选降至 55 和 60°C 之间。溶液 pH 值从大约 5.5 到 6.5 降至 3 和 5.5 之间的范围内。优选地，溶液 pH 值为 4-4.5。向溶液中添加葡糖淀粉酶并进行 24-72 小时的反应，优选 36-48 小时。

通过使用热稳定的葡糖淀粉酶的糖化工艺可以在比传统批次糖化工艺更高的温度下进行。依照本发明，糖化可以在高于 $60 - 80^{\circ}\text{C}$ 的温度范围内进行，优选 $63 - 75^{\circ}\text{C}$ 。这既适用于传统的批次工艺 (上述)，又适用于连

续的糖化工艺。

实际上，包括一个或多个膜分离步骤即过滤步骤的连续糖化工艺，为了能够维持膜上合理的高流量或最小化微生物污染，必须在 60℃ 以上的温度下进行。因此，热稳定的葡糖淀粉酶提供了在工业糖化工艺可接受的一定时期内以合适的价格和/或以较低的蛋白质用量进行大规模连续糖化工艺的可能性。依据本发明，糖化时间甚至还可以缩短。

来自葡糖淀粉酶、酸性淀粉酶以及支链淀粉酶的典型糖化试验的葡萄糖产率是 95.5-96.5%。其余的碳水化合物典型地包括 1% 麦芽糖、1.5-2% 异麦芽糖和 1-1.5% 高级寡糖。二糖之所以产生，这是因为葡糖淀粉酶在高浓度葡萄糖和高干物质水平下具有形成转化产物的趋势。

对液化后溶液之中存在的糖以及糖化期间所形成的糖比活增加的葡糖淀粉酶将是有利的，因为那是可应用减少的酶用量或缩短加工时间。一般而言，与短链糖相比，葡糖淀粉酶对由更长的糖组成的底物具有偏好性，因此其对麦芽七糖比对麦芽糖的比活高大约 6 倍。对短链糖如麦芽糖的比活增加（不降低对寡糖的活性）因此还允许使用降低的酶量和/或缩短加工时间。

进一步，高葡萄糖产率可由具有提高的 $\alpha-1, 4$ 水解活性（如果 $\alpha-1, 6$ 活性没有变化或甚至降低的话）的葡糖淀粉酶变体获得，由于使用的是减少量的酶蛋白，因此 $\alpha-1, 6$ 转化产物的形成下降（更少异麦芽糖）。

本发明的葡糖淀粉酶可与仅水解具有至少 4 个葡糖残基的分子中的 $\alpha-(1, 6)$ -葡糖键的酶联合用于本发明的工艺之中。优选地，本发明的葡糖淀粉酶可与支链淀粉酶或异淀粉酶联合使用。异淀粉酶和支链淀粉酶用于去枝的用途、酶的分子特性以及所述酶和葡糖淀粉酶的潜在用途在 G.M.A. van Beynum 等人，《淀粉转化技术》（*Starch Conversion Technology*），Marcel Dekker, New York, 1985, 101-142 中得到阐述。

信号肽

本发明还涉及这样的核酸构建体，其包含编码有效地与由 SEQ ID NO: 1（编码由 SEQ ID NO: 2 的 1 至 18 位氨基酸组成的信号肽）的 1 至 54 位

核苷酸组成的核酸序列连接蛋白的基因，其中该基因是所述核酸序列的外源基因。

本发明还涉及包含此类核酸构建体的重组表达载体核重组宿主细胞。

本发明还涉及生产蛋白质的方法，包括(a)在适于生产蛋白质的条件下培养这样的重组宿主细胞；以及(b)回收蛋白质。

核酸序列可以有效地与具有其它控制序列的外源基因连接。这样的其它控制序列在前文有描述。

蛋白质可以是宿主细胞天生的或是异源的。术语“蛋白质”在文中不是指特定长度的编码产物，因此，其包括肽、寡肽和蛋白质。术语“蛋白质”还包括经组合以形成编码产物的两个或多个多肽。蛋白质还包括杂合多肽，其包括得自至少两个不同蛋白质的部分或完全多肽序列的组合，其中一个或多个是宿主细胞天生的或是异源的。蛋白质还包括上述蛋白质和杂合蛋白质天然发生的等位和改造变异。

优选地，所述蛋白质是激素或其变体、酶、受体或其部分、抗体或其部分或者报告子。在更优选的实施方案中，所述蛋白质是氧化还原酶、转移酶、水解酶、裂解酶、异构酶或连接酶。在甚至更优选的实施方案中，蛋白质是氨基肽酶、淀粉酶、糖酶、羧肽酶、过氧化氢酶、纤维素酶、几丁质酶、角质酶(cutinase)、环糊精糖基转移酶、脱氧核糖核酸酶、酯酶、 α -半乳糖苷酶、 β -半乳糖苷酶、葡糖淀粉酶、 α -葡萄糖苷酶、 β -葡萄糖苷酶、转化酶、漆酶、脂肪酶、甘露糖苷酶、变构酶(mutanase)、氧化酶、果胶水解酶(pectinolytic enzyme)、过氧化物酶、植酸酶、多酚氧化酶、蛋白水解酶、核糖核酸酶、谷氨酰氨转移酶或木聚糖酶。

基因可以来自任何原核生物、真核生物或其他来源。

进一步通过如下实施例对本发明进行更详细的说明，其不应当解释为限制本发明的范围。

实施例

实施例 1

将罗氏阿太菌基因组 DNA 克隆进曲霉载体。

利用可由公共数据库(如, EMBL 登录号 D49448)获得的 AMG 序列设计的特异性 PCR 引物从罗氏阿太菌菌株的基因组 DNA 克隆 AMG 编码基因。所述基因组克隆含有 9 个 EMBL 序列中未出现或未标出的内含子; 编码罗氏阿太菌的 AMG 的完整序列显示在 SEQ ID NO: 1 中, 所编码的 AMG 氨基酸序列显示在 SEQ ID NO: 2 中。

正向克隆引物包含 *Bam*H1 位点; 即 DCrF1 (SEQ ID NO: 3):

5'acgtacggatccacaatgtttcgttcactctgg

反向克隆引物包含 *Sal*1 位点; 即 DCrR1 (SEQ ID NO: 4):

5'gtacgtgtcgacctagagaaacaagatagg

以罗氏阿太菌基因组 DNA 为模板, 用克隆引物 DCrF1 和 DCrR1 以及高保真 ReddyMiX™ (Extensor Hi-Fidelity PCR Master mix ReddyMix™, 目录号 AB0794, Abgene®, UK) 进行 PCR 扩增。PCR 产物用 1% 琼脂糖凝胶检测, 从凝胶中切除 1876 bp 处的 PCR 片段, 并用来自 Amersham Pharmacia (目录号 27-9602-01) 的 GFX™ PCR DNA 和凝胶带纯化试剂盒进行纯化。用两个限制性酶 *Sal*1 和 *Bam*H1 在 2 个独立的反应中消化 PCR 片段。在消化物中使用 Amersham MicroSpin™ 层析柱纯化所述片段。

下面所述曲霉表达载体 pDAu71 以 *Xho*1 和 *Bam*H1 同时切割, 并从 1% 琼脂糖电泳中分离。使用 Amersham GFX™ 试剂盒切除并纯化 7857 bp 处的带。

用 T₄ DNA 连接酶连接纯化的载体和 PCR 片段并转化到大肠杆菌 XL1 蓝细胞中。从所得的克隆中分离质粒 DNA 并且测序以验证正确的构建体。

表达质粒 pMT 2188 的构建如下进行。曲霉表达质粒 pCaHj527 (WO 0070064 中描述) 包括这样的表达盒, 所述表达盒是以与构巢曲霉丙糖磷酸盐异构酶非翻译的前导序列 (Pna2/tpi) 以及黑曲霉淀粉葡萄糖苷酶终止子 (Tamg) 融合的黑曲霉中性淀粉酶 II 启动子为基础的。还存在于质粒中的是

曲霉选择性标记 *amdS*，其来自构巢曲霉，能够使得曲霉以乙酰胺作为唯一氮源能生长，以及来自酿酒酵母 (*Saccharomyces cerevisiae*) 的 *URA3* 标记，其能使 *pyrF* 缺陷大肠杆菌菌株 DB6507 (ATCC 35673) 得以生长。以如下方式利用酿酒酵母 *URA 3* 基因作为选择性标记进行大肠杆菌 DB6507 的转化：通过 Mandel 和 Higa (Mandel, M. 和 A. Higa (1970) 《分子生物学杂志》(J. Mol. Biol.) 45, 154) 的方法使大肠杆菌 DB6507 成为感受态的。在固体 M9 培养基上选择转化株 (Sambrook 等人 (1989) 《分子克隆：实验室手册》，第 2 版，冷泉港实验室出版社)，补充有 1 g/l 酪蛋白氨基酸、500 μ g/l 维生素 B₁ 和 10 mg/l 卡那霉素。pCaHj527 上存在的 Pna2/*tpi* 启动子通过简单的 PCR 方法进行定点诱变；利用诱变引物将核苷酸 134-144 从 GTACTAAAACC 改变为 CCGTTAAATTT，并利用另一个诱变引物将核苷酸 423-436 从 ATGCAATTAAACT 改变为 CGGCAATTTAACGG。所得的质粒被称为 pMT 2188。

质粒 pJaL719 的构建如下进行。将来自 pMT2188 的 6352 bp *EcoRI*-*BamHI* 片段和来自 pJaL676 的 617 bp *EcoRI*-*BamHI* 片段连接在一起得到质粒 pJaL719。

质粒 pJaL721 构建如下。质粒 pJaL721 是 pJaL719 的衍生物，其中复制了来自 7 到 510 位的部分 NA2 启动子。通过 PCR 扩增来自 pJaL676 的 538 bp DNA 片段、纯化并用核酸内切酶 *EcoRI* 消化，得到 520 bp DNA 片段。所述 520 bp DNA 片段与来自 pJaL719 的 6355 bp *EcoRI* DNA 片段连接在一起得到质粒 pJaL721。

质粒 pCaHj607 构建如下。将来自 pUC19 的氨苄青霉素抗性基因以这种方式插入到 pMT2188 的 *URA3* 基因中，从而氨苄青霉素抗性基因与 *Not I* 位点侧接，并且使用 *Not I* 通过切除氨苄青霉素抗性基因可以重建 *URA3* 基因。使用 PCR 引物 Amp 5' 和 Amp 3' 从 pUC 19 扩增氨苄青霉素抗性基因。使用引物 *URA amp 5'* 和 *URA 5 3'* 非从 pMT2188 扩增 *URA3* 基因的 5' 部分。使用引物 *URA 3 5'* 非和 *URA amp 3'* 从 pMT2188 扩增 *URA3* 基因的 3' 部分：

amp 3' (SEQ ID NO: 5): caaagagacatgggcgccgcaggatcttcacctagatcc

Amp 5' (SEQ ID NO: 6): cgtaaccttcatecgggcccgcgatgtatccgctcatgagac

Ura 3'非 (SEQ ID NO: 7):gtctcatgagcggatacatgcggcccgcgatgaaggttacg

URA 5'3'非 (SEQ ID NO: 8):ggatctagggtgaagatcctgcggcccgcceatgtctctttg

URA amp 3' (SEQ ID NO: 9): taatcggttaagecgagttgc

URA amp 5' (SEQ ID NO: 10): ttacgaatgcacacgggtgt

形成的 PCR 片段再次使用扩增用引物 URA amp 5'和 URA amp 3'利用剪接通过重叠法进行融合。将形成的 PCR 片段作为 Stu I-EcoR I 片段克隆进 pMT2188 得到质粒 pCaHj607。

曲霉表达载体 pDAu71 的构建如下进行。用 KspI 和 SspI 限制性酶切割曲霉表达载体 pJaL721。消化产物跑 1 %琼脂糖凝胶，切下 6558bp 载体带，并使用 Jetsorb™ DNA 提取凝胶试剂盒从琼脂糖凝胶中(Genomed®, 德国)纯化。PCR 以 pJaL721 为模板使用引物 DAuP105 和 DAuP106A 进行，在 SspI 位点引入突变。

DAuP105 (SEQ ID NO: 11): 5'gctggtgattggctggct

DAuP106 (SEQ ID NO: 12): 5'atgttgaatagctcgccc

Expand™ 高保真 PCR 系统(Roche™ 德国)根据厂家的说明书使用。使用 QIAquick™ DNA 清除系统 (Quiagen®)从引物和核苷酸中纯化扩增片段。然后用 KspI 限制酶消化 PCR 产物并跑 2% 琼脂糖凝胶。使用 Jetsorb™ DNA 提取试剂盒从琼脂糖凝胶中纯化 317 bp 的带。然后将 317 bp 片段克隆进 KspI-SspI 切割的载体 pJaL721 中,导致 SspI 位点的诱变,这样识别位点不再存在。将这个中间载体命名为 pDAu58。

使用 BamHI 和 XhoI 限制性酶切割 pDAu58, 消化产物跑 1% 琼脂糖凝胶, 并使用 Jetsorb™ DNA 提取试剂盒纯化所述切割载体, 将填充 DNA 序列克隆进 pDAu58。将也用 BamHI 和 XhoI 切割的填充 DNA 片段克隆进 pDAu58A。将这个中间载体命名为 pDAu68。

然后将来自质粒 pCaHJ607、包括插入在 URA3 编码区中的氨苄青霉素抗性基因的抗生素选择盒克隆进 pDAu68,以替代 URA3 选择标记和启动子区 NA2 负 TATA 盒。使用 XcmI 和 EcoRI 限制性酶从 pCaHJ607 切下抗生素选择盒。在 1%琼脂糖凝胶上分离 XcmI-EcoRI 片段,并用 Jetsorb™ DNA 提取试剂盒进行纯化。以 XcmI 和 EcoRI 切割质粒 pDAu68,并用 Jetsorb™ DNA 提取试剂盒纯化 1%琼脂糖凝胶。所得到的中间质粒命名为 pDAu70。

然后将启动子区 NA2 负 TATA 盒克隆回 pDAu70 中。用 EcoRI 质粒切割 pJaL721,用 1%琼脂糖凝胶分离启动子区 NA2 负 TATA 盒,并用 Jetsorb™ DNA 提取试剂盒纯化。将启动子区克隆进用 EcoRI 切割、并用 Jetsorb™ DNA 提取试剂盒由 1%琼脂糖凝胶纯化的 pDAu68 中。所得到的质粒命名为 dpDAu71。

实施例 2

制备没有接头和淀粉结合结构域(SBD)的 G2 型罗氏阿太菌 AMG。

设计克隆引物,其在 AMG 编码基因中编码氨基酸 IALP 的序列之后引入终止密码子。

反向引物包含 *Sac*II 位点;即 240303P1 (SEQ ID NO: 13):

5'gggcccccgcgctagggagagcgatcgtggcactc

以罗氏阿太菌的 cDNA(即无内含子)为模板,用克隆引物 DCrF1 (SEQ ID NO: 3) 和反向引物 240303P1 (SEQ ID NO: 13),以及来自 Qiagene®的 ProofStart™ 聚合酶进行 PCR 扩增。用 1%琼脂糖凝胶检测 PCR 产物,从凝胶中切下 PCR 片段带,并使用来自 Amersham Pharmacia (目录号 27-9602-01)的 GFX™ PCR DNA 和凝胶带纯化试剂盒进行纯化。用两个限制性酶 *Sac*II 和 *Bam*HI 消化所述 PCR 片段。

同时使用 *Sac*II 和 *Bam*HI 切割曲霉表达载体 pEN12516 (其构建详述于共同未决的专利申请 PA 2003 00169 中,该专利于 2003 年 2 月 6 日递交于

丹麦专利局)，接着从1%琼脂糖凝胶分离。在7857 bp处的带被切下，并使用 Amersham GFX™ 试剂盒纯化。

已纯化的载体和 PCR 片段使用 T₄ DNA 连接酶连接，并转化进大肠杆菌 One Shot® TOP10 (Invitrogen™) 细胞。从所得的克隆中分离质粒 DNA，并测序以验证正确的构建。

实施例 3

黑曲霉的转化。

在 90 ml 蔗糖培养基 + 1 ml 1 M NaNO₃ + 10 ml YPD (Sherman 等人, (1981), 《酵母遗传学方法》(Methods in Yeast Genetics), 冷泉港实验室) 中接种黑曲霉菌株 HowB112 (描述于 WO 99/28448 A1 中) 的孢子，并振动温浴大约 24 小时。穿过神奇滤布 (miracloth) 过滤收集菌丝体，并以 200 ml 的 0.6 M MgSO₄ 冲洗。将该菌丝体悬浮于 15 ml 的 1.2 M MgSO₄, 10 mM NaH₂PO₄, pH 5.8 中。将悬浮液置于冰上冷，并加入 40 mg Novozym™ 234 (Novozymes™)。5 分钟后，加入 1 ml 的 12 mg/ml BSA (Sigma™ type H25) 轻摇 37℃ 温育持续 1.5-2 小时，直到在显微镜下检测到样品中大量可见的原生质体。

然后悬浮液经神奇滤布过滤器过滤，将滤液转移到灭菌管中并用 5 ml 的 0.6 M 山梨醇覆盖。1000 G 离心 15 分钟，从 MgSO₄ 缓冲液顶部收集原生质体。向原生质体悬浮液中加入 2 体积 SC (1.2 M 山梨醇, 10 mM CaCl₂)，并使该混合物在 1000G 离心 5 分钟。用 5 ml SC 重悬该原生质体沉淀，并重新沉淀。重复此操作。最后，用 2 ml SC 重悬该原生质体。

100 ml 的原生质体悬浮液与 5 mg 罗氏阿太菌克隆的 DNA 制备物、及 1mg 选择质粒 pToC90 (描述于 WO 9117243 A1 中) 混合。混合液室温放置约 25 分钟。加入 0.3 ml 60% PEG 4000 和 10 mM CaCl₂ 并小心混合(两次)。混合液室温放置约 25 分钟。最后将原生质体涂布在基本培养板上 (Cove, (1966), 《生物化学生物物理学报》(Biochem. Biophys. Acta) 113, 51-56), 其中含有 1.0 M 蔗糖, pH 7.0, 作为氮源的 10 mM 乙酰胺和抑制本底生长的 20 mM CsCl。37℃ 温育约 7 天后，挑取孢子，接种在 YPM 中

并涂板筛选单克隆。在 34°C 温育 3 天后测量 YPM 上清的 AMG 活性。这个过程重复两遍，并将在第三遍再分离后的单克隆孢子作为明确的转化株储存。

实施例 4

罗氏阿太菌葡糖淀粉酶的糖化试验

通过将用 LIQUOZYME XTM (Novozymes A/S)液化的玉米淀粉所制备的 DE 11 麦芽糖糊精溶解在 Milli-QTM 水中来制备糖化底物，并调节干物质含量 (DS) 至大约 30%。糖化试验在配备有磁力搅拌器的密封 50ml 蓝帽烧瓶中进行。在开始糖化时调整溶液的 pH 值，在糖化温度下校准 pH 电极。使用了如下的酶：

罗氏阿太菌 GA	16.5 AGU/ml
DEXTROZYME TM GA (基准)	357 AGU/g

标准反应条件

培养基浓度	30.0 % w/w (初始)
温度	60°C
pH (初始)	4.3
酶的剂量	0.2 AGU/g DS

以设定的间隔取样并在沸水中加热 15 分钟以灭活酶。冷却后，在用 HPLC 分析之前，将样品稀释到大约 5% DS 并过滤 (Sartorius MINISARTTM NML 0.2 微米)。作为总可溶性碳水化合物%值的葡萄糖水平在下表 1 中给出。

表 1

样品	24 小时	48 小时	72 小时
罗氏阿太菌 GA	90.74	95.91	96.29
DEXTROZYME TM GA	90.22	94.87	95.68

结果表明从罗氏阿太菌纯化的葡糖淀粉酶比含有葡糖淀粉酶和酸稳定性淀粉酶两种活性的商业酶产品 DEXTROZYME™ GA (Novozymes A/S; 产自黑曲霉)给出更高的葡萄糖产率。

实施例 5

以葡糖淀粉酶和支链淀粉酶进行糖化试验

如上所述制备糖化底物，并在配备有磁力搅拌器的密封 50ml 蓝盖烧瓶中进行糖化反应。在糖化开始时调节 pH 值，在糖化温度下校准 pH 电极。使用的酶如下：

罗氏阿太菌 GA	1.93 AGU/ml
DEXTROZYME™ GA (基准)	357 AGU/g
PROMOZYME™ D2 (支链淀粉酶)	1480 NPUN/g

标准反应条件

培养基浓度	30.0 % w/w (初始)
温度	60°C
pH (初始)	4.5
酶的剂量	见表 2

表 2

样品号	酶	Dex GA AGU/g DS	Promo D2 NPUN/g DS	罗氏阿太菌 GA AGU/g DS
1	Dextrozyme™ GA	0.20		
2	罗氏阿太菌+ Promozyme™ D2		0.3	0.15

以设定的间隔取样并在沸水中加热 15 分钟以灭活酶。冷却后，在用 HPLC 分析之前将样品稀释到约 5% DS 并过滤 (Sartorius MINISART™

NML 0.2 微米)。作为总可溶性碳水化合物%值的葡萄糖水平在下表 3 中给出:

表 3

样品号	24 小时	48 小时	72 小时
1	93.56	95.88	95.95
2	89.53	96.13	96.46

这些结果显示当来自罗氏阿太菌的纯化葡糖淀粉酶与支链淀粉酶联合时,得到比商业产品 DEXTROZYME™ GA 更高的葡萄糖产率,尽管罗氏阿太菌制备物的葡糖淀粉酶浓度(酶的剂量) 低于 DEXTROZYMETM™ GA。

实施例 6

以罗氏阿太菌葡糖淀粉酶和黑曲霉 α -淀粉酶进行糖化试验

如上所述制备糖化培养基,并在配备有磁力搅拌器的密封 50ml 蓝盖烧瓶中进行糖化反应。在糖化开始时调节 pH 值,在糖化温度下校准 pH 电极。使用的酶及活性如下:

罗氏阿太菌 GA	2.05 AGU/ml
DEXTROZYME™ GA (基准)	357 AGU/g
黑曲霉 α -淀粉酶	65.4 FAU (A)/g

标准反应条件

培养基浓度	30.0 % w/w (初始)
温度	60 °C
pH (初始)	4.3
酶的剂量	见下文表 4

表 4

样品号	酶	Dex GA AGU/g DS	罗氏阿太菌 GA AGU/g DS	黑曲霉 α -淀粉酶 FAU(A)/g DS
1	Dextrozyme™ GA	0.20		
2	罗氏阿太菌+ 黑曲霉 α -淀粉酶		0.175	0.042
3	罗氏阿太菌+ 黑曲霉 α -淀粉酶		0.20	0.048

以设定间隔取样并在沸水中加热 15 分钟以灭活酶。冷却后，在用 HPLC 分析之前将样品稀释到约 5% DS 并过滤 (Sartorius MINISART™ NML 0.2 微米)。作为总可溶性碳水化合物%值的葡萄糖水平在下表 5 中给出：

表 5

样品号	24 小时	48 小时	72 小时
1	93.87	96.05	96.10
2	88.92	96.74	96.83
3	92.02	96.79	96.60

这些结果显示当来自罗氏阿太菌的纯化葡糖淀粉酶与黑曲霉酸稳定的 α -淀粉酶以与见于在商业产品 DEXTROZYME™ GA 中的相同的葡糖淀粉酶/ α -淀粉酶的比例联合时，能够获得比 DEXTROZYME™ GA 显著更高的葡萄糖产率。

实施例 7

依据发明通过离子中等分配色谱法 (Ion Moderated Partition Chromatography)、根据由 Scobell 等人, 《谷物化学》(Cereal Chemistry) 54 (4), 1977 年 7-8 月, 第 905-917 页所建立和公开的方法测量葡萄糖(DX) %值。为进行这样的分析, 推荐在 85℃ 达到热稳态的 Aminex HPX 87C 柱。使用去离子水作为洗脱液, 可实现单糖如右旋糖 (葡萄糖)和果糖良好的分离。根据区域标准化程序测定单糖和寡糖 (麦芽糖、异麦芽糖、潘糖、和

高级糖)，而无需进行校准和干物质测定。通过折射法测量的不同成分的量表达为% w/w(g/100g 干物质)。

<110> 诺和酶股份有限公司(Novozymes A/S)
<120> 生产葡糖淀粉酶的方法及其用途
<130> 10362.000
<160> 13
<170> PatentIn version 3.2
<210> 1
<211> 2427
<212> DNA
<213> 罗氏阿太菌 (Athelia rolfsii)

<220>
<221> CDS
<222> (1)..(208)

<220>
<221> 内含子
<222> (209)..(283)

<220>
<221> CDS
<222> (284)..(354)

<220>
<221> 内含子
<222> (355)..(410)

<220>
<221> misc_feature
<222> (367)..(367)
<223> 内含子中的核苷酸-任何核苷酸

<220>
<221> misc_feature
<222> (392)..(392)
<223> 内含子中的核苷酸-任何核苷酸

<220>
<221> CDS
<222> (411)..(557)

<220>
<221> 内含子
<222> (558)..(616)

<220>
<221> CDS
<222> (617)..(770)

<220>
<221> 内含子
<222> (771)..(825)

<220>
<221> CDS
<222> (826)..(986)

<220>
<221> 内含子
<222> (987)..(1058)

<220>
<221> CDS
<222> (1059)..(1331)

<220>
<221> 内含子
<222> (1332)..(1409)

<220>
 <221> CDS
 <222> (1410)..(1713)

<220>
 <221> 内含子
 <222> (1714)..(1787)

<220>
 <221> CDS
 <222> (1788)..(1958)

<220>
 <221> 内含子
 <222> (1959)..(2020)

<220>
 <221> CDS
 <222> (2021)..(2116)

<220>
 <221> 内含子
 <222> (2117)..(2173)

<220>
 <221> CDS
 <222> (2174)..(2325)

<400> 1

atg ttt cgt tca ctc ctg gcc ttg gct gcg tgt gca gtc gcc tct gta	48
Met Phe Arg Ser Leu Leu Ala Leu Ala Ala Cys Ala Val Ala Ser Val	
1 5 10 15	
tct gca cag tct gcg tct gcg aca gca tat ctt acc aag gaa tct gca	96
Ser Ala Gln Ser Ala Ser Ala Thr Ala Tyr Leu Thr Lys Glu Ser Ala	
20 25 30	
gtt gcc aag aat ggc gta ctt tgc aac att ggt agc cag gga tgc atg	144
Val Ala Lys Asn Gly Val Leu Cys Asn Ile Gly Ser Gln Gly Cys Met	
35 40 45	
tct gag ggt gcc tat agc ggt att gtg atc gca tct ccc tct aaa act	192
Ser Glu Gly Ala Tyr Ser Gly Ile Val Ile Ala Ser Pro Ser Lys Thr	
50 55 60	
agc cct gac tat ctc t gtgagtatta ttgttaaagt agcctcactg atagtacatt	248
Ser Pro Asp Tyr Leu	
65	
ttctgagttc tgttacaacc ctggtattat aatag at acc tgg act cgc gac	300
Tyr Thr Trp Thr Arg Asp	
75	
tcg tcg ctc gtc ttc aag atg tta att gac caa tac aca aat ggc ctg	348
Ser Ser Leu Val Phe Lys Met Leu Ile Asp Gln Tyr Thr Asn Gly Leu	
80 85 90	
gat acg gtagtgcca tcngcgttcc ggctgcctc aaagatgnaa aattgatgtt	404
Asp Thr	
tcttag aca ctg cgc act ctc att gac gag ttt gtc tct gcg gaa gcc	452
Thr Leu Arg Thr Leu Ile Asp Glu Phe Val Ser Ala Glu Ala	
95 100 105	
acc att caa caa acc agt aac cca tct ggt acc gtc tct acc ggt ggt	500
Thr Ile Gln Gln Thr Ser Asn Pro Ser Gly Thr Val Ser Thr Gly Gly	
110 115 120	
ctc ggc gaa ccc aaa ttc aat atc gac gag acg gca ttt acg ggc gca	548
Leu Gly Glu Pro Lys Phe Asn Ile Asp Glu Thr Ala Phe Thr Gly Ala	
125 130 135	
tgg ggt cgt gtaagctacc aatacacaat caaaatcgac catctgtatt	597
Trp Gly Arg	
140	

tactatctat aatttctag ccc caa cgt gat ggt ccc gcc ctc cgt gca acc Pro Gln Arg Asp Gly Pro Ala Leu Arg Ala Thr 145 150	649
gca atc atg acc tat gcg acg tat ctg tac aac aat ggc aac act tcc Ala Ile Met Thr Tyr Ala Thr Tyr Leu Tyr Asn Asn Gly Asn Thr Ser 155 160 165	697
tac gtg acc aac acc ctt tgg cct atc atc aag ctc gac ctt gac tat Tyr Val Thr Asn Thr Leu Trp Pro Ile Ile Lys Leu Asp Leu Asp Tyr 170 175 180 185	745
gtc aac tcg gac tgg aac cag acc a gtaagcgaat ttctaggggg Val Asn Ser Asp Trp Asn Gln Thr 190	790
acttatctaa aacagcatat tcaaccagta aatag cg ttt gac ctc tgg gaa Thr Phe Asp Leu Trp Glu 195	842
gaa gtt gac tcg tct tct ttc ttt acg act gcc gtt cag cac cgt gct Glu Val Asp Ser Ser Ser Phe Phe Thr Thr Ala Val Gln His Arg Ala 200 205 210 215	890
ctt gtt cag ggc gca gcc ttt gct acc ctc atc ggc caa act tcg tct Leu Val Gln Gly Ala Ala Phe Ala Thr Leu Ile Gly Gln Thr Ser Ser 220 225 230	938
gct tcg act tac tcc gcc acg gcc cct agc att ctc tgc ttc ttg cag Ala Ser Thr Tyr Ser Ala Thr Ala Pro Ser Ile Leu Cys Phe Leu Gln 235 240 245	986
gtgagataaa aatctttcta tgtaattggt tttcccctc aaattgaaat tgacatattt	1046
gcgatccaat ag tct tac tgg aac acc aac gga tac tgg acg gcc aac act Ser Tyr Trp Asn Thr Asn Gly Tyr Trp Thr Ala Asn Thr 250 255 260	1097
ggt ggc gga cgt tcc gcc aag gac gcc aac acc ata ctc gct tct atc Gly Gly Gly Arg Ser Gly Lys Asp Ala Asn Thr Ile Leu Ala Ser Ile 265 270 275	1145
cac aca ttt gac gcc agc gcc ggc tgc tct gct gcc acg tct caa cca His Thr Phe Asp Ala Ser Ala Gly Cys Ser Ala Ala Thr Ser Gln Pro 280 285 290	1193
tgc tct gac gta gca ttg gcc aac ctg aag gta tac gtt gac tct ttc Cys Ser Asp Val Ala Leu Ala Asn Leu Lys Val Tyr Val Asp Ser Phe 295 300 305	1241
cgt agt att tat acg atc aac agc ggt att tcc tct acc tcg ggt gtt Arg Ser Ile Tyr Thr Ile Asn Ser Gly Ile Ser Ser Thr Ser Gly Val 310 315 320	1289
gct act ggt cgc tac ccc gaa gat tcg tat tac aat ggc aac Ala Thr Gly Arg Tyr Pro Glu Asp Ser Tyr Tyr Asn Gly Asn 325 330 335	1331
gtacgtattt atctaatttt tccaagacag tcaaagtitta tgttcatectg ccccccttta	1391
cctgtacatt caaatag ccc tgg tac ctc tgc aca ctc gcc gtc gcc gag Pro Trp Tyr Leu Cys Thr Leu Ala Val Ala Glu 340 345	1442
cag ctc tat gat gct ctc atc gta tgg aag gct gcc ggg gag ctc aac Gln Leu Tyr Asp Ala Leu Ile Val Trp Lys Ala Ala Gly Glu Leu Asn 350 355 360 365	1490
gtc acc tcc gtc tcg ctc gcg ttc ttc cag caa ttc gac tcg agc atc Val Thr Ser Val Ser Leu Ala Phe Phe Gln Gln Phe Asp Ser Ser Ile 370 375 380	1538
acc gcc ggc act tac gcc tcc tcg tcg agc gta tac act tcg ctc atc Thr Ala Gly Thr Tyr Ala Ser Ser Ser Ser Val Tyr Thr Ser Leu Ile 385 390 395	1586

tct gac atc cag gcg ttc gca gac gag ttt gtt gac att gtt gcc aag Ser Asp Ile Gln Ala Phe Ala Asp Glu Phe Val Asp Ile Val Ala Lys 400 405 410	1634
tac acg cct tcg tct ggc ttc ttg tct gag cag tat gat aag tcc acg Tyr Thr Pro Ser Ser Gly Phe Leu Ser Glu Gln Tyr Asp Lys Ser Thr 415 420 425	1682
ggt gct cag gat tcg gct gct aac ttg act t gtaagtcac tattgttca Gly Ala Gln Asp Ser Ala Ala Asn Leu Thr 430 435	1733
ttctattcct tttcaaaaa aaaagtgatg ctaatgattt ttggcggaaa ccag gg Trp	1789
tcc tat gct gct gct atc acc gct tac caa gcc cgc aat ggc ttc aca Ser Tyr Ala Ala Ala Ile Thr Ala Tyr Gln Ala Arg Asn Gly Phe Thr 445 450 455	1837
ggt gct tcg tgg ggt gct aag gga gtt tct acc tcc tgc tcg act ggt Gly Ala Ser Trp Gly Ala Lys Gly Val Ser Thr Ser Cys Ser Thr Gly 460 465 470	1885
get aca agc ccg ggt ggc tcc tcg ggt agt gtc gag gtc act ttc gac Ala Thr Ser Pro Gly Gly Ser Ser Gly Ser Val Glu Val Thr Phe Asp 475 480 485	1933
ggt tac gct acc aca gta tat ggc c gtaagcactt gactagcttc Val Tyr Ala Thr Thr Val Tyr Gly 490 495	1978
aaaccatact tcacatgct gataaacaaa aaaatgaaac ag ag aac atc tat Gln Asn Ile Tyr 500	2031
atc acc ggt gat gtg agt gag ctc ggc aac tgg aca ccc gcc aat ggt Ile Thr Gly Asp Val Ser Glu Leu Gly Asn Trp Thr Pro Ala Asn Gly 505 510 515	2079
ggt gca ctc tct tct gct aac tac ccc acc tgg agt g gtaagttgac Val Ala Leu Ser Ser Ala Asn Tyr Pro Thr Trp Ser 520 525	2126
ccttaccagt atcttgacag acattgatat tgacttccgc aatacag cc acg atc Ala Thr Ile 530	2181
gct ctc ccc gct gac acg aca atc cag tac aag tat gtc aac att gac Ala Leu Pro Ala Asp Thr Thr Ile Gln Tyr Lys Tyr Val Asn Ile Asp 535 540 545	2229
ggc agc acc gtc atc tgg gag gat gct atc agc aat cgc gag atc acg Gly Ser Thr Val Ile Trp Glu Asp Ala Ile Ser Asn Arg Glu Ile Thr 550 555 560	2277
acg ccc gcc agc ggc aca tac acc gaa aaa gac act tgg gat gaa tct Thr Pro Ala Ser Gly Thr Tyr Thr Glu Lys Asp Thr Trp Asp Glu Ser 565 570 575	2325
taaactgctg aacttgaac gcttgcaaaa gcgaatggtg tagaaaataa acgaagattt	2385
tgattgcitt gttttgtttc tcttctate ttgtttctct ag	2427
<210> 2	
<211> 579	
<212> PRT	
<213> 罗氏阿太菌 (Athelia rolfsii)	
<400> 2	
Met Phe Arg Ser Leu Leu Ala Leu Ala Ala Cys Ala Val Ala Ser Val 1 5 10 15	

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

Thr Ser Gly Val Ala Thr Gly Arg Tyr Pro Glu Asp Ser Tyr Tyr Asn
 325 330 335
 Gly Asn Pro Trp Tyr Leu Cys Thr Leu Ala Val Ala Glu Gln Leu Tyr
 340 345 350
 Asp Ala Leu Ile Val Trp Lys Ala Ala Gly Glu Leu Asn Val Thr Ser
 355 360 365
 Val Ser Leu Ala Phe Phe Gln Gln Phe Asp Ser Ser Ile Thr Ala Gly
 370 375 380
 Thr Tyr Ala Ser Ser Ser Ser Val Tyr Thr Ser Leu Ile Ser Asp Ile
 385 390 395 400
 Gln Ala Phe Ala Asp Glu Phe Val Asp Ile Val Ala Lys Tyr Thr Pro
 405 410 415
 Ser Ser Gly Phe Leu Ser Glu Gln Tyr Asp Lys Ser Thr Gly Ala Gln
 420 425 430
 Asp Ser Ala Ala Asn Leu Thr Trp Ser Tyr Ala Ala Ala Ile Thr Ala
 435 440 445
 Tyr Gln Ala Arg Asn Gly Phe Thr Gly Ala Ser Trp Gly Ala Lys Gly
 450 455 460
 Val Ser Thr Ser Cys Ser Thr Gly Ala Thr Ser Pro Gly Gly Ser Ser
 465 470 475 480
 Gly Ser Val Glu Val Thr Phe Asp Val Tyr Ala Thr Thr Val Tyr Gly
 485 490 495
 Gln Asn Ile Tyr Ile Thr Gly Asp Val Ser Glu Leu Gly Asn Trp Thr
 500 505 510
 Pro Ala Asn Gly Val Ala Leu Ser Ser Ala Asn Tyr Pro Thr Trp Ser
 515 520 525
 Ala Thr Ile Ala Leu Pro Ala Asp Thr Thr Ile Gln Tyr Lys Tyr Val
 530 535 540
 Asn Ile Asp Gly Ser Thr Val Ile Trp Glu Asp Ala Ile Ser Asn Arg
 545 550 555 560
 Glu Ile Thr Thr Pro Ala Ser Gly Thr Tyr Thr Glu Lys Asp Thr Trp
 565 570 575

Asp Glu Ser

<210> 3
 <211> 34
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物 DCrF1

<400> 3 acgtacggat ccacaatggt tcgttcactc ctgg	34
<210> 4 <211> 30 <212> DNA <213> 人工序列	
<220> <223> 引物 DCrR1	
<400> 4 gtacgtgtcg acctagagaa acaagatagg	30
<210> 5 <211> 40 <212> DNA <213> 人工序列	
<220> <223> 引物 amp 3'	
<400> 5 caaagagaca tgggcggccg caggatcttc acctagatcc	40
<210> 6 <211> 40 <212> DNA <213> 人工序列	
<220> <223> 引物 Amp 5'	
<400> 6 cgtaaccttc atcgcggccg catgtatcgg ctcatgagac	40
<210> 7 <211> 40 <212> DNA <213> 人工序列	
<220> <223> 引物 Ura 3 5' 非	
<400> 7 gtctcatgag cggatacatg cggccgcgat gaaggttacg	40
<210> 8 <211> 40 <212> DNA <213> 人工序列	
<220> <223> 引物 URA 5 3' 非	
<400> 8 ggatctaggt gaagatcctg cggccgccca tgtctctttg	40
<210> 9 <211> 19 <212> DNA <213> 人工序列	
<220> <223> 引物 URA amp 3'	
<400> 9 taatcggtaa gcgagttgc	19

<210> 10
 <211> 19
 <212> DNA
 <213> 人工序列

 <220>
 <223> 引物 URA amp 5'

 <400> 10
 ttacgaatgc acacgggtg 19

<210> 11
 <211> 18
 <212> DNA
 <213> 人工序列

 <220>
 <223> 引物 DAuP105

 <400> 11
 gctggtgatt ggctggct 18

<210> 12
 <211> 18
 <212> DNA
 <213> 人工序列

 <220>
 <223> 引物 DAuP106

 <400> 12
 atggtgaata gctcgccc 18

<210> 13
 <211> 37
 <212> DNA
 <213> 人工序列

 <220>
 <223> 引物 240303P1

 <400> 13
 gggccccgc ggctagggga gagcgatcgt ggcactc 37