



(12) 发明专利申请

(10) 申请公布号 CN 103667215 A

(43) 申请公布日 2014. 03. 26

(21) 申请号 201310362275. 5 *C12N 1/15* (2006. 01)
(22) 申请日 2005. 02. 04 *C12P 21/00* (2006. 01)
(30) 优先权数据 *C12N 15/10* (2006. 01)
60/542, 614 2004. 02. 06 US *C12N 15/82* (2006. 01)
(62) 分案原申请数据 *C12N 5/10* (2006. 01)
200580011917. 0 2005. 02. 04 *A01H 5/00* (2006. 01)
(83) 生物保藏信息 *C11D 3/386* (2006. 01)
NRRL B-30704 2004. 01. 30 *C12P 19/14* (2006. 01)
(71) 申请人 诺维信股份有限公司
地址 美国加利福尼亚州
(72) 发明人 威廉·多森 詹妮弗·格里尼尔
丁晗澍
(74) 专利代理机构 北京市柳沈律师事务所
11105
代理人 封新琴
(51) Int. Cl.
C12N 9/42 (2006. 01)
C12N 15/56 (2006. 01)
C12N 15/70 (2006. 01)
C12N 15/80 (2006. 01)
C12N 1/21 (2006. 01)

权利要求书4页 说明书51页
序列表4页 附图3页

(54) 发明名称

具有增强分解纤维素活性的多肽和编码所述多肽的多核苷酸

(57) 摘要

本发明涉及具有增强分解纤维素活性的多肽和编码所述多肽的多核苷酸, 具体的涉及具有增强分解纤维素活性的分离多肽和编码所述多肽的分离核酸。本发明还涉及包含所述核酸的核酸构建体、载体、和宿主细胞, 以及生产和使用所述多肽的方法。

1. 选自下列的具有增强分解纤维素活性的分离多肽：
 - (a) 具有如下氨基酸序列的多肽,该氨基酸序列与 SEQ ID NO:2 第 23 至 250 位氨基酸具有至少 70% 的同一性；
 - (b) 由如下多核苷酸编码的多肽,该多核苷酸在至少低严紧条件下与 (i)SEQ ID NO:1 第 67 至 796 位核苷酸, (ii)SEQ ID NO:1 第 67 至 796 位核苷酸所含 cDNA 序列,或 (iii), (i) 或 (ii) 的互补链发生杂交 ;和
 - (c) 在 SEQ ID NO:2 第 23 至 250 位氨基酸中包含一个或多个氨基酸的保守取代、缺失、和 / 或插入的变体。
2. 权利要求 1 的多肽,具有与 SEQ ID NO:2 第 23 至 250 位氨基酸具有至少 70% 同一性的氨基酸序列。
3. 权利要求 2 的多肽,具有与 SEQ ID NO:2 第 23 至 250 位氨基酸具有至少 75% 同一性的氨基酸序列。
4. 权利要求 3 的多肽,具有与 SEQ ID NO:2 第 23 至 250 位氨基酸具有至少 80% 同一性的氨基酸序列。
5. 权利要求 4 的多肽,具有与 SEQ ID NO:2 第 23 至 250 位氨基酸具有至少 90% 同一性的氨基酸序列。
6. 权利要求 5 的多肽,具有与 SEQ ID NO:2 第 23 至 250 位氨基酸具有至少 95% 同一性的氨基酸序列。
7. 权利要求 6 的多肽,具有与 SEQ ID NO:2 第 23 至 250 位氨基酸具有至少 97% 同一性的氨基酸序列。
8. 权利要求 1-7 中任一项的多肽,包含 SEQ ID NO:2 的氨基酸序列。
9. 权利要求 1-8 中的多肽,由 SEQ ID NO:2 或其具有增强纤维素分解活性的片段组成。
10. 权利要求 9 的多肽,由 SEQ ID NO:2 组成。
11. 权利要求 9 的多肽,由 SEQ ID NO:2 第 23 至 250 位氨基酸组成。
12. 权利要求 1 的多肽,由在至少中等严紧条件下与 (i)SEQ ID NO:1 第 67 至 796 位核苷酸, (ii)SEQ ID NO:1 第 67 至 796 位核苷酸所含 cDNA 序列,或 (iii), (i) 或 (ii) 的互补链发生杂交的多核苷酸编码。
13. 权利要求 1 的多肽,由在至少中等—高严紧条件下与 (i)SEQ ID NO:1 第 67 至 796 位核苷酸, (ii)SEQ ID NO:1 第 67 至 796 位核苷酸所含 cDNA 序列,或 (iii), (i) 或 (ii) 的互补链发生杂交的多核苷酸编码。
14. 权利要求 1 的多肽,由在至少高严紧条件下与 (i)SEQ ID NO:1 第 67 至 796 位核苷酸, (ii)SEQ ID NO:1 第 67 至 796 位核苷酸所含 cDNA 序列,或 (iii), (i) 或 (ii) 的互补链发生杂交的多核苷酸编码。
15. 权利要求 1 的多肽,其中所述多肽是在 SEQ ID NO:2 第 23 至 250 位氨基酸中包含一个或多个氨基酸的保守取代、缺失、和 / 或插入的变体。
16. 权利要求 1 的多肽,由质粒 pDZA2-7 所含多核苷酸编码,所述质粒包含在大肠杆菌 NRRL B-30704 中。
17. 包含编码权利要求 1-16 任一项的多肽的核苷酸序列的分离多核苷酸。
18. 权利要求 17 的分离多核苷酸,在 SEQ ID NO:1 的成熟多肽编码序列中具有至少一

个突变,其中突变的核苷酸序列编码由 SEQ ID NO:2 第 23 至 250 位氨基酸组成的多肽。

19. 包含权利要求 17 或 18 的多核苷酸的核酸构建体,所述多核苷酸与指导多肽在表达宿主中产生的一种或多种控制序列可操作连接。

20. 包含权利要求 18 的核酸构建体的重组表达载体。

21. 包含权利要求 18 的核酸构建体的重组宿主细胞。

22. 生产权利要求 1-16 任一项的多肽的方法,包括:(a) 在有利于多肽产生的条件下培养细胞,该细胞的野生型形式能够产生所述多肽;并(b) 回收所述多肽。

23. 生产权利要求 1-16 任一项的多肽的方法,包括:(a) 在有利于多肽产生的条件下培养包含如下核酸构建体的宿主细胞,该核酸构建体包含编码所述多肽的核苷酸序列;并(b) 回收所述多肽。

24. 生产亲本细胞突变体的方法,包括破坏或删除编码权利要求 1-16 任一项的多肽的核苷酸序列,导致突变体产生的多肽要比亲本细胞少。

25. 通过权利要求 24 的方法产生的突变细胞。

26. 权利要求 24 的突变细胞,还包含编码天然或异源蛋白的基因。

27. 生产蛋白质的方法,包括:(a) 在有利于蛋白质产生的条件下培养权利要求 26 的突变细胞;并(b) 回收所述蛋白质。

28. 通过如下获得的分离多核苷酸,即(a) 在中等严紧条件下使 DNA 群体与(i) SEQ ID NO:1 第 67 至 796 位核苷酸,(ii) SEQ ID NO:1 第 67 至 796 位核苷酸所含 cDNA 序列,或(iii), (i) 或(ii) 的互补链进行杂交;并(b) 分离发生杂交的多核苷酸,其编码具有增强分解纤维素活性的多肽。

29. 权利要求 28 的分离多核苷酸,其通过如下获得,即(a) 在中等—高严紧条件下使 DNA 群体与(i) SEQ ID NO:1 第 67 至 796 位核苷酸,(ii) SEQ ID NO:1 第 67 至 796 位核苷酸所含 cDNA 序列,或(iii), (i) 或(ii) 的互补链进行杂交;并(b) 分离发生杂交的多核苷酸,其编码具有增强分解纤维素活性的多肽。

30. 权利要求 29 的分离的多核苷酸,其通过如下获得,即(a) 在高严紧条件下使 DNA 群体与(i) SEQ ID NO:1 第 67 至 796 位核苷酸,(ii) SEQ ID NO:1 第 67 至 796 位核苷酸所含 cDNA 序列,或(iii), (i) 或(ii) 的互补链进行杂交;并(b) 分离发生杂交的多核苷酸,其编码具有增强分解纤维素活性的多肽。

31. 产生具有突变核苷酸序列的多核苷酸的方法,包括:(a) 在 SEQ ID NO:1 的成熟多肽编码序列中引入至少一个突变,其中突变的核苷酸序列编码由 SEQ ID NO:2 第 23 至 250 位氨基酸组成的多肽;并(b) 回收包含突变核苷酸序列的多核苷酸。

32. 通过权利要求 31 的方法产生的突变多核苷酸。

33. 生产多肽的方法,包括:(a) 在有利于多肽产生的条件下培养包含编码所述多肽的权利要求 32 的突变多核苷酸的细胞;并(b) 回收所述多肽。

34. 包含编码蛋白质的基因的核酸构建体,其中所述基因可操作连接由 SEQ ID NO:1 第 1 至 66 位核苷酸组成的编码信号肽的核苷酸序列,其中所述基因对于所述核苷酸序列而言是异源的。

35. 包含权利要求 34 的核酸构建体的重组表达载体。

36. 包含权利要求 34 的核酸构建体的重组宿主细胞。

37. 生产蛋白质的方法,包括:(a)在有利于蛋白质产生的条件下培养权利要求36的重组宿主细胞;并(b)回收所述蛋白质。

38. 生产权利要求1-16任一项的多肽的方法,包括:(a)在有利于多肽产生的条件下培养包含如下多核苷酸的转基因植物或植物细胞,该多核苷酸编码本发明具有增强分解纤维素活性的多肽;并(b)回收所述多肽。

39. 经编码权利要求1的多肽的多核苷酸转化的转基因植物、植物部分或植物细胞。

40. 含有权利要求1-16任一项的具有增强分解纤维素活性的多肽、纤维素分解活性、和表面活性剂的清洁剂组合物。

41. 降解或者转化纤维素材料的方法,包括:在存在有效量的权利要求1-16任一项的具有增强分解纤维素活性的多肽时,用有效量的分解纤维素的蛋白质处理纤维素材料,其中与不存在具有增强分解纤维素活性的多肽时相比,具有增强分解纤维素活性的多肽的存在增加纤维素材料的降解。

42. 权利要求41的方法,其中纤维素材料选自草本植物材料、农业残余物、林业残余物、城市固体废物、废纸、以及纸浆和造纸厂的残余物。

43. 权利要求41的方法,其中纤维素材料是玉米秸秆。

44. 权利要求41的方法,其中一种或多种纤维素分解酶选自纤维素酶、内切葡聚糖酶、纤维二糖水解酶、和 β -葡糖苷酶。

45. 权利要求41的方法,还包括用有效量的一种或多种选自半纤维素酶、酯酶、蛋白酶、漆酶、或过氧化物酶的酶处理纤维素材料。

46. 权利要求41的方法,其中所述方法是预处理方法。

47. 权利要求41的方法,其中所述方法是同步糖化与发酵方法(SSF)中的步骤。

48. 权利要求41的方法,其中所述方法是混合水解与发酵方法(HHF)中的步骤。

49. 权利要求41的方法,还包括回收降解的纤维素材料。

50. 权利要求49的方法,其中降解的纤维素材料是糖类。

51. 权利要求50的方法,其中所述糖类选自葡萄糖、木糖、甘露糖、半乳糖、和阿拉伯糖。

52. 权利要求41的方法,其中分解纤维素的蛋白质和/或具有增强分解纤维素活性的多肽是含有或者不含细胞的发酵液形式。

53. 生产有机物质的方法,包括:

(a)在存在有效量的权利要求8的具有增强分解纤维素活性的多肽时,用有效量的分解纤维素的蛋白质糖化纤维素材料,其中与不存在具有增强分解纤维素活性的多肽时相比,具有增强分解纤维素活性的多肽的存在增加纤维素材料的降解;

(b)用一种或多种发酵微生物对步骤(a)的已糖化纤维素材料进行发酵;并

(c)由发酵回收有机物质。

54. 权利要求53的方法,其中纤维素材料选自草本植物材料、农业残余物、林业残余物、城市固体废物、废纸、以及纸浆和造纸厂的残余物。

55. 权利要求53的方法,其中纤维素材料是玉米秸秆。

56. 权利要求53的方法,其中一种或多种纤维素分解酶选自纤维素酶、内切葡聚糖酶、纤维二糖水解酶、和 β -葡糖苷酶。

57. 权利要求 53 的方法,还包括用有效量的一种或多种选自半纤维素酶、酯酶、蛋白酶、漆酶、或过氧化物酶的酶处理纤维素材料。
58. 权利要求 57 的方法,其中酯酶是脂肪酶、磷脂酶、角质酶、或其混合物。
59. 权利要求 53 的方法,其中步骤 (a) 和 (b) 在同步糖化和发酵中同时进行。
60. 权利要求 53 的方法,其中有机物质是醇、有机酸、酮、氨基酸、或气体。
61. 权利要求 60 的方法,其中醇是阿拉伯糖醇、丁醇、乙醇、甘油、甲醇、1,3-丙二醇、山梨糖醇、或木糖醇。
62. 权利要求 60 的方法,其中有机酸是乙酸、醋酮酸、己二酸、抗坏血酸、柠檬酸、2,5-二酮-D-葡糖酸、甲酸、富马酸、葡糖二酸、葡糖酸、葡糖醛酸、戊二酸、3-羟基丙酸、衣康酸、乳酸、苹果酸、丙二酸、草酸、丙酸、琥珀酸、或木糖酸。
63. 权利要求 60 的方法,其中酮是丙酮。
64. 权利要求 60 的方法,其中氨基酸是天冬氨酸、谷氨酸、甘氨酸、赖氨酸、丝氨酸、或苏氨酸。
65. 权利要求 60 的方法,其中气体是甲烷、氢、二氧化碳、或一氧化碳。
66. 权利要求 53 的方法,其中分解纤维素的蛋白质和 / 或具有增强分解纤维素活性的多肽是含有或者不含细胞的发酵液形式。

具有增强分解纤维素活性的多肽和编码所述多肽的多核苷酸

[0001] 本申请是申请日为 2005 年 02 月 04 日,申请号为 200580011917.0(国际申请号为 PCT/US2005/003802),发明名称为“具有增强分解纤维素活性的多肽和编码所述多肽的多核苷酸”的发明专利申请的分案申请。

[0002] 关于受到联邦政府资助的研究与开发的发明权利的声明

[0003] 本发明是在受到政府资助的由能源部签订的主要合同 DE-AC36-98G010337 的 NREL 转包合同 NO. ZCO-30017-02 下做出的。政府拥有本发明的一些权利。

[0004] 发明背景

发明领域

[0005] 本发明涉及具有增强分解纤维素活性的分离多肽和编码所述多肽的分离多核苷酸。本发明还涉及包含所述多核苷酸的核酸构建体、载体和宿主细胞以及生产和使用所述多肽的方法。

[0006] 相关领域的描述

[0007] 纤维素是单糖葡萄糖通过 β -1,4-连接键共价键合的聚合物。许多微生物产生水解 β -连接葡聚糖的酶。这些酶包括内切葡聚糖酶、纤维二糖水解酶和 β -葡糖苷酶。内切葡聚糖酶在任意位置消化纤维素聚合物,使其易受纤维二糖水解酶的攻击。纤维二糖水解酶从纤维素聚合物的末端顺序释放纤维二糖分子。纤维二糖是葡萄糖的水溶性 β -1,4-连接二聚体。 β -葡糖苷酶将纤维二糖水解成葡萄糖。

[0008] 将纤维素原料转变成乙醇具有如下优点,即易于获得大量原料、避免燃烧或填埋物质、以及乙醇燃料的清洁性。现在认为木材、农业残余物、草本作物、和城市固体废物是生产乙醇的原料。这些物质主要由纤维素、半纤维素和木质素组成。一旦将纤维素转变成葡萄糖,就容易通过酵母将葡萄糖发酵成乙醇。

[0009] 在本领域中提高纤维素原料的转化将是有利的。

[0010] 本发明的一个目的是提供具有增强分解纤维素活性的分离多肽和编码所述多肽的分离核酸序列以提高纤维素原料的转化。

[0011] 发明概述

[0012] 本发明涉及选自下列的具有增强分解纤维素活性的分离多肽:

[0013] (a) 具有如下氨基酸序列的多肽,该氨基酸序列与 SEQ ID NO:2 第 23 至 250 位氨基酸具有至少 70% 的同一性;

[0014] (b) 由如下核酸序列编码的多肽,该核酸序列在中等严谨条件下与 (i) SEQ ID NO:1 第 67 至 796 位核苷酸, (ii) SEQ ID NO:1 第 67 至 796 位核苷酸所含 cDNA 序列,或 (iii), (i) 或 (ii) 的互补链发生杂交;和

[0015] (c) 在 SEQ ID NO:2 第 23 至 250 位氨基酸中含有一个或多个氨基酸的保守缺失、插入、和 / 或替代的变体。

[0016] 本发明还涉及选自下列的编码具有增强分解纤维素活性的多肽的分离多核苷

酸：

[0017] (a) 编码具有如下氨基酸序列的多肽的多核苷酸，该氨基酸序列与 SEQ ID NO:2 第 23 至 250 位氨基酸具有至少 70% 的同一性；

[0018] (b) 与 SEQ ID NO:1 第 67 至 796 位核苷酸具有至少 70% 同一性的多核苷酸；和

[0019] (c) 在至少中等严谨条件下与 (i) SEQ ID NO:1 第 67 至 796 位核苷酸，(ii) SEQ ID NO:1 第 67 至 796 位核苷酸所含 cDNA 序列，或 (iii)，(i) 或 (ii) 的互补链发生杂交的多核苷酸。

[0020] 本发明还涉及包含所述多核苷酸的核酸构建体、重组表达载体、和重组宿主细胞。

[0021] 本发明还涉及生产这些具有增强分解纤维素活性的多肽的方法，包括 (a) 在有助于产生所述多肽的条件下培养包含如下核酸构建体的重组宿主细胞，该核酸构建体包含编码所述多肽的多核苷酸；并 (b) 回收所述多肽。

[0022] 本发明还涉及包含编码蛋白质的基因的核酸构建体，其中所述基因可操作连接由 SEQ ID NO:1 第 1 至 66 位核苷酸组成的编码信号肽的核苷酸序列，其中所述基因对于所述核苷酸序列而言是外源的。

[0023] 本发明还涉及降解或者转化纤维素材料的方法，包括：在存在有效量的具有增强分解纤维素活性的多肽时，用有效量的分解纤维素的蛋白质处理纤维素材料，其中与不存在该具有增强分解纤维素活性的多肽时相比，该具有增强分解纤维素活性的多肽的存在增加纤维素材料的降解。

[0024] 本发明还涉及生产有机物质的方法，包括：

[0025] (a) 在存在有效量的具有增强分解纤维素活性的多肽时，用有效量的分解纤维素的蛋白质糖化纤维素材料，其中与不存在该具有增强分解纤维素活性的多肽时相比，该具有增强分解纤维素活性的多肽的存在增加纤维素材料的降解；

[0026] (b) 用一种或多种发酵微生物对步骤 (a) 的已糖化纤维素材料进行发酵；并

[0027] (c) 由发酵回收有机物质。

[0028] 附图简述

[0029] 图 1 显示了橙色嗜热子囊菌 (*Thermoascus aurantiacus*) 的增强分解纤维素活性的基因组 DNA 序列和推导的氨基酸序列 (分别为 SEQ ID NO:1 和 2)。预测的内含子以斜体表示。预测的信号肽 (SignalP) 标以下划线。编码序列为 799bp，包括终止密码子且间含一个 56bp 的内含子。预测的成熟多肽含有 228 个氨基酸。

[0030] 图 2 显示了 pAILo1 的限制性图谱。

[0031] 图 3 显示了 pBANE10 的限制性图谱。

[0032] 图 4 显示了 pAILo2 的限制性图谱。

[0033] 图 5 显示了 pDZA2 的限制性图谱。

[0034] 定义

[0035] 增强分解纤维素活性 (cellulolytic enhancing activity): 术语“增强分解纤维素活性”在本文中定义为这样的生物活性，其促进具有纤维素分解活性的蛋白质水解纤维素材料。对本发明来说，增强分解纤维素活性是通过在如下条件下测量分解纤维素的蛋白质水解纤维素材料产生的还原糖的增量而测定的：在存在和不存在 0.01–2.5mg 增强分解纤维素活性 /g PCS 中的纤维素时，2.5mg 分解纤维素的蛋白质 /g PCS 于 50°C 保温 5–7

天,与用没有增强分解纤维素活性的等量总蛋白质加载(5.01-7.5mg 分解纤维素的蛋白质/g PCS 中的纤维素)进行的对照水解进行比较。在一个优选的方面,使用从 Novozymes A/S (**Bagsværd**, 丹麦)获得的、由表达米曲霉(*Aspergillus oryzae*) β -葡糖苷酶的瑞氏木霉(*Trichoderma reesei*) 发酵得到的、下称 Tr/AoBG (W002/095014) 的纤维素酶制备物的纤维素酶蛋白质加载作为分解纤维素活性的来源。

[0036] 本发明的多肽具有由 SEQ ID NO:2 第 23 至 250 位氨基酸所示氨基酸序列组成的多肽的至少 20%、优选至少 40%、更优选至少 50%、更优选至少 60%、更优选至少 70%、更优选至少 80%、甚至更优选至少 90%、最优选至少 95%、且甚至最优选至少 100% 的增强分解纤维素活性。

[0037] 纤维素材料(cellulosic material):术语“纤维素材料”在本文中定义为任何含有纤维素的材料。通常在例如植物的茎、叶、壳、皮和穗轴或者树的叶、枝、和木质发现有纤维素。纤维素材料还可以是,但不限于,草本植物材料、农业残余物、林业残余物、城市固体废物、废纸、以及纸浆和造纸厂残余物。应当理解本文的纤维素可以是木质纤维素的形式,即在混合基质中含有木质素、纤维素、和半纤维素的植物细胞壁材料。

[0038] 在一个优选的方面,纤维素材料是玉米秸秆。在另一个优选的方面,纤维素材料是玉米纤维。在另一个优选的方面,纤维素材料是稻草。在另一个优选的方面,纤维素材料是纸和纸浆加工废料。在另一个优选的方面,纤维素材料是木本或草本植物。在另一个优选的方面,纤维素材料是甘蔗渣。

[0039] 可以使用本领域已知的常规方法来利用纤维素材料,进行或者可进行预处理。例如,物理预处理技术可包括各种类型的碾磨、照射、蒸/蒸汽喷发(steam explosion)、和湿热分解作用(hydrothermolysis);化学预处理技术可包括用稀酸、碱、有机溶剂、氨、二氧化硫、二氧化碳、和 pH 控制的湿热分解作用;及生物预处理技术可包括应用溶解木质素的微生物(参见例如 Hsu, T-A., 1996, Pretreatment of biomass, 在《Handbook on Bioethanol: Production and Utilization》中, Wyman, C. E. 编, Taylor&Francis, Washington, DC, 179-212; Ghosh, P. 和 Singh, A., 1993, Physicochemical and biological treatments for enzymatic/microbial conversion of lignocellulosic biomass, *Adv. Appl. Microbiol.* 39 :295-333; McMillan, J. D., 1994, Pretreating lignocellulosic biomass: a review, 在《Enzymatic Conversion of Biomass for Fuels Production》, Himmel, M. E.、Baker, J. O. 和 Overend, R. P. 编, ACS Symposium Series 566, American Chemical Society, Washington, DC, chapter 15; Gong, C. S.、Cao, N. J.、Du, J. 和 Tsao, G. T., 1999, Ethanol production from renewable resources, 在《Advances in Biochemical Engineering/Biotechnology》, Scheper, T. 编, Springer-Verlag Berlin Heidelberg, Germany, 65 :207-241; Olsson, L. 和 Hahn-Hagerdal, B., 1996, Fermentation of lignocellulosic hydrolysates for ethanol production, *Enz. Microb. Tech.* 18 :312-331; 及 Vallander, L. 和 Eriksson, K.-E. L., 1990, Production of ethanol from lignocellulosic materials: State of the art, *Adv. Biochem. Eng./Biotechnol.* 42 :63-95)。

[0040] 纤维素分解活性:术语“纤维素分解活性”在本文中定义为水解纤维素材料的生物学活性。对本发明来说,纤维素分解活性是通过在如下条件下测量分解纤维素的混合物水

解纤维素材料的增量而测定的 :1-10mg 分解纤维素的蛋白质 /g PCS 中的纤维素于 50℃ 保温 5-7 天,与未添加分解纤维素的蛋白质的对照水解进行比较。在一个优选的方面,使用从 Novozymes A/S (**Bagsværd**, 丹麦) 获得的、由表达米曲霉 β - 葡糖苷酶的瑞氏木霉发酵得到的、下称 Tr/AoBG (W002/095014) 的载有纤维素酶蛋白质的纤维素酶制备物作为分解纤维素活性的来源。

[0041] 经过预处理的玉米秸秆 :术语“经过预处理的玉米秸秆”或“PCS”在本文中定义为通过用加热和稀酸处理玉米秸秆得到的纤维素材料。对本发明来说,PCS 是通过实施例 9 描述的方法或其改变时间、温度和酸量的变体方法制备的。

[0042] 糖苷水解酶家族 61 (family61 glycoside hydrolase) :术语“糖苷水解酶 61 家族”或“GH61 家族”在本文中定义为根据 B. Henrissat, 1991, A classification of glycosyl hydrolases based on amino-acid sequence similarities, *Biochem. J.* 280 :309-316 及 Henrissat B. 和 Bairoch A., 1996, Updating the sequence-based classification of glycosyl hydrolases, *Biochem. J.* 316 :695-696 归入糖苷水解酶 61 家族的多肽。目前, Henrissat 将 GH61 家族列入未分类的,指示属于这个家族的多肽的某些性质,诸如机制、催化性亲核体 / 碱、催化性质子供体、和三维结构是未知的。

[0043] 分离的多肽 :术语“分离的多肽”在用于本文时指根据 SDS-PAGE 的测定,至少 20% 纯、优选至少 40% 纯、更优选至少 60% 纯、还更优选至少 80% 纯、最优选至少 90% 纯、且甚至最优选至少 95% 纯的多肽。

[0044] 基本上纯的多肽 (substantially pure polypeptide) :术语“基本上纯的多肽”在本文中指以重量计含有至多 10%、优选至多 8%、更优选至多 6%、更优选至多 5%、更优选至多 4%、更优选至多 3%、甚至更优选至多 2%、最优选至多 1%、且甚至最优选至多 0.5% 与该多肽天然相伴随的其它多肽物质的多肽制备物。因此,优选的是,基本上纯的多肽以重量计制备物中存在的全部多肽物质是至少 92% 纯的、优选至少 94% 纯的、更优选至少 95% 纯的、更优选至少 96% 纯的、更优选至少 96% 的纯的、更优选至少 97% 纯的、更优选至少 98% 纯的、甚至更优选至少 99% 纯的、最优选至少 99.5% 纯的、且甚至最优选 100% 纯的。

[0045] 本发明的多肽优选是基本上纯的形式。具体而言,优选的是多肽是“本质上 (essentially) 纯的形式”,即多肽制备物本质上不含与该多肽天然相伴随的其它多肽物质。例如,这可通过采用众所周知的重组方法或采用经典的纯化方法制备多肽来实现。

[0046] 在本文中,术语“基本上纯的多肽”与术语“分离的多肽”和“分离形式的多肽”含义相同。

[0047] 同一性 (identity) :本文用参数“同一性”来描述两种氨基酸序列之间或两种核苷酸序列之间的相关性。

[0048] 对本发明来说,两种氨基酸序列之间的同一性程度可使用 Paracel BioView Workbench 软件 (Paracel, Pasadena, CA) 及 blosum62 矩阵通过 blastp 算法 (Higgins, 1989, *CABIOS* 5 :151-153) 来测定。采用缺口罚分,存在 :10 和延伸 :10,作为配对比对 (pairwise alignment) 参数。

[0049] 对本发明来说,两种核苷酸序列之间的同一性程度使用 LASERGENE™ MEGALIGN™ 软件 (DNASTAR 公司, Madison, WI) 及同一性表和后面的多重比对参数通过 Wilbur-Lipman 法 (Wilbur 和 Lipman, 1983, *Proceedings of the National Academy of Science USA* 80 :

726-730)来测定;缺口罚分 10 和缺口长度罚分 10。配对比对参数:Ktuple=3,缺口罚分=3,和窗口(windows)=20。

[0050] 多肽片段:术语“多肽片段”在本文中定义为从 SEQ ID NO:2 第 23 至 250 位氨基酸或其同源序列的氨基和 / 或羧基末端删除了一个或多个氨基酸的多肽,其中所述片段具有增强分解纤维素活性。优选的是,SEQ ID NO:2 第 23 至 250 位氨基酸的片段含有至少 175 个氨基酸残基、更优选至少 190 个氨基酸残基、且最优选至少 205 个氨基酸残基。

[0051] 子序列:术语“子序列”在本文中定义为由 SEQ ID NO:1 第 67 至 796 位核苷酸的 5' 和 / 或 3' 末端删除了一个或多个核苷酸的核苷酸序列,或其同源序列,其中子序列编码具有增强分解纤维素活性的多肽片段。优选的是,SEQ ID NO:1 第 67 至 796 位核苷酸的子序列含有至少 525 个核苷酸、更优选至少 570 个核苷酸、且最优选至少 615 个核苷酸。

[0052] 等位变体:术语“等位变体”在本文中占据同一染色体基因座的基因的两种或多种可选形式中的任一种。等位变异由突变天然引起,并可导致种群内多态性。基因突变可以是沉默的(所编码多肽没有变化),或者可编码氨基酸序列改变的多肽。多肽的等位变体是由基因的等位变体编码的多肽。

[0053] 分离的多核苷酸:术语“分离的多核苷酸”在用于本文时指根据琼脂糖电泳的测定,至少 20% 纯的、优选至少 40% 纯的、更优选至少 60% 纯的、甚至更优选至少 80% 纯的、最优选至少 90% 纯的、且甚至最优选至少 95% 纯的多核苷酸。

[0054] 基本上纯的多核苷酸:术语“基本上纯的多核苷酸”在用于本文时指不含其它外来或不需要的核苷酸且以适于在基因工程蛋白质生产系统内使用的形式存在的多核苷酸制备物。因此,基本上纯的多核苷酸以重量计含有至多 10%、优选至多 8%、更优选至多 6%、更优选至多 5%、更优选至多 4%、更优选至多 3%、甚至更优选至多 2%、最优选至多 1%、且甚至最优选至多 0.5% 与该多核苷酸天然相伴随的其它多核苷酸物质。然而,基本上纯的多核苷酸可包括天然存在的 5' 和 3' 非翻译区,诸如启动子和终止子。优选的是,基本上纯的多核苷酸以重量计是至少 90% 纯的、优选至少 92% 纯的、更优选至少 94% 纯的、更优选至少 95% 纯的、更优选至少 96% 纯的、更优选至少 97% 纯的、甚至更优选至少 98% 纯的、最优选至少 99% 纯的、且甚至最优选至少 99.5% 纯的。优选的是,本发明的多核苷酸是基本上纯的形式。具体而言,优选的是,本文公开的多核苷酸是“本质上纯的形式”,即多核苷酸制备物本质上不含与该多核苷酸天然相伴随的其它多核苷酸物质。在本文中,术语“基本上纯的多核苷酸”与术语“分离的多核苷酸”和“分离形式的多核苷酸”含义相同。多核苷酸可以是基因组、cDNA、RNA、半合成、合成来源或其任意组合。

[0055] cDNA:术语“cDNA”在本文中定义为可通过逆转录由真核细胞获得的成熟的、经过剪接的 mRNA 分子而制备的 DNA 分子。cDNA 缺乏通常在相应基因组 DNA 中存在的内含子序列。最初的初级 RNA 转录本是 mRNA 的前体,它在成为成熟的、经过剪接的 mRNA 之前要经过一系列加工步骤。这些步骤包括通过称为剪接的过程除去内含子序列。因此,由 mRNA 衍生的 cDNA 缺乏任何内含子序列。

[0056] 核酸构建体:术语“核酸构建体”在用于本文时指单链或双链的核酸分子,它是由天然存在的基因分离的,或是经过修饰而以自然界中不存在的方式含有核酸区段。当核酸构建体含有表达本发明编码序列所需要的控制序列时,术语核酸构建体与术语“表达盒”含义相同。

[0057] 控制序列:术语“控制序列”在本文中定义为包括对表达编码本发明多肽的多核苷酸所必需的或有利的的所有成分。每种控制序列对编码多肽的核苷酸序列来说可以是天然的或外来的,或者每种控制序列相对彼此也可以是天然的或外来的。这些控制序列包括,但不限于,前导序列、聚腺苷酸化序列、前肽(propeptide)序列、启动子、信号肽序列和转录终止子。至少,控制序列包括启动子以及转录和翻译终止信号。为了引入特异的限制性位点以便于控制序列与编码多肽的核苷酸序列编码区连接,控制序列可以带有接头(linkers)。

[0058] 可操作连接:术语“可操作连接”在本文中指这样一种构造,其中控制序列置于相对于多核苷酸序列的编码序列的适当位置,使得控制序列指导多肽编码序列的表达。

[0059] 编码序列:在用于本文时,术语“编码序列”意指直接规定其蛋白质产物的氨基酸序列的核苷酸序列。编码序列的边界一般由开放读码框确定,它通常起始于 ATG 起始密码子或备选起始密码子诸如 GTG 和 TTG,并终止于终止密码子诸如 TAA、TAG 和 TGA。编码序列可以是 DNA、cDNA 或重组核苷酸序列。

[0060] 表达:术语“表达”包括多肽产生涉及的所有步骤,包括但不限于:转录、转录后修饰、翻译、翻译后修饰、和分泌。

[0061] 表达载体:术语“表达载体”在本文中定义为如下线状或环状 DNA 分子,它包含编码本发明多肽的多核苷酸,且其与有助于其表达的其它核苷酸可操作连接。

[0062] 宿主细胞:术语“宿主细胞”在用于本文时包括易于用包含本发明多核苷酸的核酸构建体或表达载体转化、转染、转导、等等的任何细胞类型。

[0063] 修饰:术语“修饰”在本文中意指对包含 SEQ ID NO:2 第 23 至 250 位氨基酸或其同源序列或由其组成的多肽的任何化学修饰以及对编码该多肽的 DNA 的基因操作。修饰可以是一个或多个氨基酸的替代、删除和/或插入,以及一个或多个氨基酸侧链的替换。

[0064] 人工变体:在用于本文时,术语“人工变体”意指具有增强分解纤维素活性的多肽,它是由表达经修饰的核苷酸序列 SEQ ID NO:1 或其同源序列或其成熟编码区的生物体产生的。经修饰的核苷酸序列可通过人工干预对 SEQ ID NO:1 中所公开的核苷酸序列或其同源序列或其成熟编码区进行修饰来获得。

[0065] 发明详述

[0066] 具有增强分解纤维素活性的多肽

[0067] 在第一个方面,本发明涉及具有如下氨基酸序列的具有增强分解纤维素活性的分离多肽,该氨基酸序列与 SEQ ID NO:2 第 23 至 250 位氨基酸(即成熟多肽)具有至少 70%、优选至少 75%、更优选至少 80%、更优选至少 85%、甚至更优选至少 90%、最优选至少 95%、且甚至最优选至少 97%、98% 或 99% 的同一性程度(下文的“同源多肽”)。在一个优选的方面,同源多肽的氨基酸序列与 SEQ ID NO:2 第 23 至 250 位氨基酸相差 10 个氨基酸、优选 5 个氨基酸、更优选 4 个氨基酸、甚至更优选 3 个氨基酸、最优选 2 个氨基酸、且甚至最优选 1 个氨基酸。

[0068] 本发明的多肽优选包含氨基酸序列 SEQ ID NO:2 或其等位变体;或其具有增强分解纤维素活性的片段。在一个优选的方面,多肽包含氨基酸序列 SEQ ID NO:2。在另一个优选的方面,多肽包含 SEQ ID NO:2 第 23 至 250 位氨基酸或其等位变体;或其具有增强分解纤维素活性的片段。在另一个优选的方面,多肽包含 SEQ ID NO:2 第 23 至 250 位氨基酸。在另一个优选的方面,多肽由氨基酸序列 SEQ ID NO:2 或其等位变体;或其具有增强分解纤

纤维素活性的片段组成。在另一个优选的方面,多肽由氨基酸序列 SEQ ID NO:2 组成。在另一个优选的方面,多肽由 SEQ ID NO:2 第 23 至 250 位氨基酸或其等位变体;或其具有增强分解纤维素活性的片段组成。在另一个优选的方面,多肽由 SEQ ID NO:2 第 23 至 250 位氨基酸组成。

[0069] 在第二个方面,本发明涉及由如下多核苷酸编码的具有增强分解纤维素活性的分离多肽,该多核苷酸在很低严紧条件、优选低严紧条件、更优选中等严紧条件、更优选中等-高严紧条件、甚至更优选高严紧度条件、且最优选很高严紧条件下与 (i) SEQ ID NO:1 第 67 至 796 位核苷酸;(ii) SEQ ID NO:1 第 67 至 796 位核苷酸所含 cDNA 序列;(iii), (i) 或 (ii) 的子序列,或 (iv), (i)、(ii) 或 (iii) 的互补链发生杂交(J. Sambrook、E. F. Fritsch 和 T. Maniatus, 1989, 《Molecular Cloning, A Laboratory Manual》, 第 2 版, Cold Spring Harbor, New York)。SEQ ID NO:1 的子序列含有至少 100 个连续核苷酸或优选至少 200 个连续的核苷酸。另外,子序列可编码具有增强分解纤维素活性的多肽片段。

[0070] 根据本领域众所周知的方法,可使用核苷酸序列 SEQ ID NO:1 或其子序列以及氨基酸序列 SEQ ID NO:2 或其片段来设计核酸探针,用于从不同属或种的菌株鉴定和克隆编码具有增强分解纤维素活性的多肽的 DNA。具体而言,遵循标准 Southern 印迹程序,可将这些探针用于与目的属或种的基因组或 cDNA 进行杂交,从而鉴定和分离其中的相应基因。这些探针可以比完整序列短得多,但是长度应当是至少 14 个、优选至少 25 个、更优选至少 35 个、且最优选至少 70 个核苷酸。然而,优选的是,核酸探针的长度是至少 100 个核苷酸。例如,核酸探针的长度可以是至少 200 个核苷酸、优选至少 300 个核苷酸、更优选至少 400 个核苷酸、或最优选至少 500 个核苷酸。可使用甚至更长的探针,例如长度是至少 550 核苷酸、至少优选至少 600 个核苷酸、更优选至少 650 个核苷酸、或最优选至少 700 个核苷酸的核酸探针。DNA 和 RNA 探针都可使用。通常将探针进行标记,用于检测相应的基因(例如用 ^{32}P 、 ^3H 、 ^{35}S 、生物素、或亲和素)。本发明涵盖这些探针。

[0071] 因此,可从这样的其它生物体制备的基因组 DNA 或 cDNA 文库中筛选与上述探针发生杂交且编码具有增强分解纤维素活性的多肽的 DNA。来自这样的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凝胶电泳或其它分离技术来分离。可将来自文库的 DNA 或分离的 DNA 转移并固定在硝酸纤维素或其它适宜载体材料上。为了鉴定与 SEQ ID NO:1 或其子序列同源的克隆或 DNA,将载体材料用于 Southern 印迹。

[0072] 对本发明而言,杂交是指在很低至很高的严紧条件下将核苷酸序列与经标记的核酸探针进行杂交,其中核酸探针对应于 SEQ ID NO:1 所示核苷酸序列、SEQ ID NO:1 所含 cDNA 序列、其互补链、或其子序列。在这些条件下与核酸探针发生杂交的分子可用 X 射线胶片检测。

[0073] 在一个优选的方面,核酸探针是编码多肽 SEQ ID NO:2 或其子序列的核酸序列。在另一个优选的方面,核酸探针是 SEQ ID NO:1。在另一个优选的方面,核酸探针是 SEQ ID NO:1 的成熟多肽编码区。在另一个优选的方面,核酸探针是包含在质粒 pDZA2-7 中的核酸序列,所述质粒包含在大肠杆菌 NRRL B-30704 中,其中所述核酸序列编码具有增强分解纤维素活性的多肽。在另一个优选的方面,核酸探针是包含在质粒 pDZA2-7 中的成熟多肽编码区,所述质粒包含在大肠杆菌 NRRL B-30704 中。

[0074] 对于长度为至少 100 个核苷酸的长探针来说,很低至很高的严谨条件定义为遵循

标准 Southern 印迹程序,于 42°C,在 5X SSPE、0.3%SDS、200 μg/ml 经剪切且经变性的鲑精 DNA、和用于很低和低严谨度的 25% 甲酰胺、用于中等和中等 - 高严谨度的 35% 甲酰胺、或用于高和很高严谨度的 50% 甲酰胺中进行最佳 12-24 小时的预杂交和杂交。

[0075] 对于长度为至少 100 个核苷酸的长探针来说,最后优选于至少 45°C(很低严谨度)、更优选于至少 50°C(低严谨度)、更优选于至少 55°C(中等严谨度)、更优选于至少 60°C(中等 - 高严谨度)、甚至更优选于至少 65°C(高严谨度)、且最优选于至少 70°C(很高严谨度)将载体材料用 2×SSC、0.2%SDS 清洗 3 次,每次 15 分钟。

[0076] 对于长度为大约 15 个核苷酸至大约 70 个核苷酸的短探针来说,严谨条件定义为遵循标准 Southern 印迹程序,于比根据 Bolton 和 McCarthy (1962, Proceedings of the National Academy of Sciences USA48:1390) 的计算方法计算得到的 T_m 值低大约 5°C 至大约 10°C 的温度,在 0.9M NaCl、0.09M Tris-HCl pH7.6、6mM EDTA、0.5%NP-40、1X Denhardt 氏溶液、1mM 焦磷酸钠、1mM 磷酸二氢钠、0.1mM ATP、和 0.2mg/ml 酵母 RNA 中进行最佳 12 至 24 小时的预杂交、杂交、和杂交后清洗。

[0077] 对于长度为大约 15 个核苷酸至大约 70 个核苷酸的短探针来说,于比 T_m 计算值低大约 5°C 至 10°C 的温度将载体材料用 6X SSC 加 0.1%SDS 清洗 1 次 15 分钟,然后用 6X SSC 清洗两次,每次 15 分钟。

[0078] 在第三个方面,本发明涉及包含 SEQ ID NO:2 或其同源序列或其成熟多肽的一个或多个氨基酸的保守取代、删除、和 / 或插入的人工变体。优选的是,氨基酸改变的性质较小,也就是说不会明显影响蛋白质折叠和 / 或活性的保守氨基酸替代;一般是 1 个至大约 30 个氨基酸的小段删除;小段氨基或羧基末端延伸,诸如氨基末端的甲硫氨酸残基;可达大约 20-25 个残基的小段接头肽;或通过改变净电荷或另一功能而易于纯化的小段延伸,诸如多组氨酸序列段、抗原性表位、或结合域。

[0079] 保守取代的例子在碱性氨基酸(精氨酸、赖氨酸和组氨酸)、酸性氨基酸(谷氨酸和天冬氨酸)、极性氨基酸(谷氨酰胺和天冬酰胺)、疏水性氨基酸(亮氨酸、异亮氨酸和缬氨酸)、芳香族氨基酸(苯丙氨酸、色氨酸和酪氨酸)、和小氨基酸(甘氨酸、丙氨酸、丝氨酸、苏氨酸和甲硫氨酸)各组内。通常不改变特异活性的氨基酸替代是本领域已知的,例如 H. Neurath 和 R. L. Hill 在《The Proteins》中所述(1979, Academic Press, New York)。最常发生的替代是 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。

[0080] 除 20 种标准氨基酸之外,也可用非标准氨基酸(诸如 4- 羟基脯氨酸、6-N- 甲基赖氨酸、2- 氨基异丁酸、异缬氨酸、和 α- 甲基丝氨酸)来替代野生型多肽的氨基酸残基。可用有限数目的非保守氨基酸、并非由遗传密码编码的氨基酸、和非天然氨基酸来替代氨基酸残基。“非天然氨基酸”在蛋白质合成后受到修饰,和 / 或在它们的侧链上具有与标准氨基酸不同的化学结构。可化学合成非天然氨基酸,优选的是可通过商业途径购买,包括哌啶酸、噻唑烷羧酸、脱氢脯氨酸、3- 和 4- 甲基脯氨酸、和 3, 3- 二甲基脯氨酸。

[0081] 或者,氨基酸改变的本质是改变多肽的物理化学特性。例如,氨基酸改变可提高多肽的热稳定性、改变底物特异性、改变最佳 pH 等等。

[0082] 可根据本领域已知程序来确定亲本多肽中的关键氨基酸,诸如定点诱变或丙氨酸

扫描诱变(alanine-scanning mutagenesis) (Cunningham 和 Wells, 1989, Science 244 : 1081-1085)。在后一种技术中,在分子中的每个残基处引入单一的丙氨酸突变,并对得到的突变分子测试生物学活性(即增强分解纤维素活性)以鉴定对分子活性至关重要的氨基酸残基。也可参见 Hilton 等人, 1996, J. Biol. Chem. 271 : 4699-4708。也可结合假定接触位点氨基酸的突变,通过诸如核磁共振、结晶学、电子衍射、或光亲和标记等技术的测定,对结构进行物理学分析,从而确定酶的活性位点或其他生物学相互作用。例如参见 de Vos 等人, 1992, Science 255 : 306-312 ; Smith 等人, 1992, J. Mol. Biol. 224 : 899-904 ; Wlodaver 等人, 1992, FEBS Lett. 309 : 59-64。也可从与本发明多肽相关的多肽的同一性分析来推断关键氨基酸的身份。

[0083] 可用已知的诱变、重组、和 / 或改组(shuffling) 方法来进行单一或多个氨基酸替代并检测,随后进行相关筛选程序,诸如 Reidhaar-Olson 和 Sauer, 1988, Science 241 : 53-57 ; Bowie 和 Sauer, 1989, Proc. Natl. Acad. Sci. USA 86 : 2152-2156 ; W095/17413 ; 或 W095/22625 所公开的。可使用的其它方法包括易错 PCR、噬菌体展示(例如 Lowman 等人, 1991, Biochem. 30 : 10832-10837 ; 美国专利号 5, 223, 409 ; W092/06204)、和定区诱变(region-directed mutagenesis) (Derbyshire 等人, 1986, Gene 46 : 145 ; Ner 等人, 1988, DNA 7 : 127)。

[0084] 可将诱变 / 改组方法与高通量、自动筛选方法相结合,用于检测由宿主细胞表达的克隆的、经诱变的多肽的活性(Ness 等人, 1999, Nature Biotechnology 17 : 893-896)。可使用本领域的标准方法从宿主细胞回收编码活性多肽的经诱变 DNA 分子并快速测序。这些方法可快速测定目的多肽中单个氨基酸残基的重要性,并可应用于结构未知的多肽。

[0085] SEQ ID NO:2 第 23 至 250 位氨基酸中氨基酸替代、缺失、和 / 或插入的总数为 10、优选 9、更优选 8、更优选 7、更优选至多 6、更优选 5、更优选 4、甚至更优选 3、最优选 2、且甚至最优选 1。

[0086] 具有增强分解纤维素活性的多肽的来源

[0087] 可从任何属的微生物中获得本发明的多肽。对本发明来说,术语“从... 获得”在本文与指定的来源联用时,应当意指由核苷酸序列编码的多肽是由该来源产生的,或是来自于该来源的核苷酸序列已经插入其中的菌株产生的。在一个优选的方面,从指定来源获得的多肽分泌到细胞外。

[0088] 本发明的多肽可以是细菌多肽。例如,多肽可以是革兰氏阳性细菌的多肽,诸如芽孢杆菌属(Bacillus)的多肽,例如嗜碱芽孢杆菌(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)、灿烂芽孢杆菌(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)的多肽;或链霉菌属(Streptomyces)的多肽,例如浅青紫链霉菌(Streptomyces lividans)或鼠灰链霉菌(Streptomyces murinus)的多肽;或革兰氏阴性细菌的多肽,例如大肠杆菌(E. coli)或假单胞菌(Pseudomonas sp.)的多肽。

[0089] 本发明的多肽还可以是真菌多肽,且更优选酵母多肽,诸如假丝酵母属

(*Candida*)、克鲁维氏酵母属(*Kluyveromyces*)、毕赤氏酵母属(*Pichia*)、糖酵母属(*Saccharomyces*)、裂殖糖酵母属(*Schizosaccharomyces*)、或 *Yarrowia* 属的多肽；或更优选丝状真菌的多肽，诸如 *Acremonium* 属、曲霉属(*Aspergillus*)、短柄霉属(*Aureobasidium*)、隐球酵母属(*Cryptocoecus*)、*Filibasidium* 属、镰孢属(*Fusarium*)、腐质霉属(*Humicola*)、*Magnaporthe* 属、毛霉属(*Mucor*)、毁丝霉属(*Myceliophthora*)、*Neocallimastix* 属、脉孢菌属(*Neurospora*)、拟青霉属(*Paecilomyces*)、青霉属(*Penicillium*)、*Piromyces* 属、裂褶菌属(*Schizophyllum*)、*Talaromyces* 属、嗜热子囊菌属(*Thermoascus*)、梭孢壳属(*Thielavia*)、*Tolypocladium* 属、或木霉属(*Trichoderma*) 的多肽。

[0090] 在一个优选的方面，多肽是卡尔斯伯糖酵母(*Saccharomyces carlsbergensis*)、酿酒酵母(*Saccharomyces cerevisiae*)、糖化糖酵母(*Saccharomyces diastaticus*)、道拉斯糖酵母(*Saccharomyces douglasii*)、克鲁弗糖酵母(*Saccharomyces kluyveri*)、诺地糖酵母(*Saccharomyces norbensis*)、或卵形糖酵母(*Saccharomyces oviformis*)的具有增强分解纤维素活性的多肽。

[0091] 在另一个优选的方面，多肽是棘孢曲霉(*Aspergillus aculeatus*)、泡盛曲霉(*Aspergillus awamori*)、烟曲霉(*Aspergillus fumigatus*)、臭曲霉(*Aspergillus foetidus*)、日本曲霉(*Aspergillus japonicus*)、构巢曲霉(*Aspergillus nidulans*)、黑曲霉(*Aspergillus niger*)、米曲霉(*Aspergillus oryzae*)、灰盖鬼伞(*Coprinus cinereus*)、棉色二孢(*Diplodia gossypina*)、杆孢状镰孢(*Fusarium bactridioides*)、谷孢镰孢(*Fusarium cereals*)、*Fusarium crookwellense*、黄色镰孢(*Fusarium culmorum*)、禾谷镰孢(*Fusarium graminearum*)、禾赤镰孢(*Fusarium graminum*)、异孢镰孢(*Fusarium heterosporium*)、合欢木镰孢(*Fusarium negundi*)、尖孢镰孢(*Fusarium oxysporum*)、网状镰孢(*Fusarium reticulatum*)、粉红镰孢(*Fusarium roseum*)、接骨木镰孢(*Fusarium sambucinum*)、肤色镰孢(*Fusarium sarcochromum*)、拟枝孢镰孢(*Fusarium sporotrichioides*)、硫色镰孢(*Fusarium sulphureum*)、近念珠状镰孢(*Fusarium torulosum*)、单端孢镰孢(*Fusarium trichothecioides*)、有毒镰孢(*Fusarium venenatum*)、*Humicola insolens*、*Humicola lanuginosa*、*Magnaporthe grisea*、米赫毛霉(*Mucor miehei*)、嗜热毁丝霉(*Myceliophthora thermophila*)、粗糙脉孢菌(*Neurospora crassa*)、产紫青霉(*Penicillium purpurogenum*)、*Phanerochaete chrysosporium*、淡黑假黑盘菌(*Pseudoplectania nigrella*)、橙色嗜热子囊菌(*Thermoascus aurantiacus*)、哈茨木霉(*Trichoderma harzianum*)、康宁木霉(*Trichoderma koningii*)、长枝木霉(*Trichoderma longibrachiatum*)、瑞氏木霉(*Trichoderma reesei*)、绿色木霉(*Trichoderma viride*)、或 *Trichophaea saccata* 的多肽。

[0092] 在一个更优选的方面，多肽是橙色嗜热子囊菌的多肽，例如具有氨基酸序列 SEQ ID NO:2 的多肽。

[0093] 应当理解，对于上述物种，不管已知的种名是什么，本发明既包括完全阶段(perfect state)和不完全阶段(imperfect state)，还包括其它分类学等价物，例如无性型。本领域技术人员能容易的识别适当等价物的身份。

[0094] 这些物种的菌株可容易的在许多培养物保藏机构为公众所得到，诸如美国典型培

养物保藏中心(American Type Culture Collection, ATCC)、德国微生物和培养物保藏中心(Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, DSM)、荷兰培养物保藏中心(Centraalbureau Voor Schimmelcultures, CBS)、和美国农业研究机构专利培养物保藏中心北部研究中心(Agricultural Research Service Patent Culture Collection Northern Regional Research Center, NRRL)。

[0095] 此外,利用上述探针可从其它来源,包括从自然界(例如土壤、堆肥、水等)分离的微生物中鉴定和获得这些多肽。从自然生境中分离微生物的技术在本领域技术是众所周知的。然后可通过类似的筛选这样一种微生物的基因组或 cDNA 文库来获得多核苷酸。一旦用探针检测出编码多肽的多核苷酸序列,就可通过本领域普通技术人员所熟知的技术(例如参见 Sambrook 等人,1989,同上)分离或克隆此多核苷酸。

[0096] 本发明的多肽还包括融合多肽或可切割的融合多肽,其中另一多肽融合在所述多肽或其片段的 N-末端或 C-末端。可通过将编码另一多肽的核苷酸序列(或其部分)与本发明的核苷酸序列(或其部分)融合来产生融合的多肽。用于产生融合多肽的技术在本领域是已知的,包括连接编码多肽的编码序列使得它们处于同一读码框中且融合多肽的表达受相同启动子和终止子的控制。

[0097] 多核苷酸

[0098] 本发明还涉及具有编码本发明多肽的核苷酸序列的分离多核苷酸。

[0099] 在一个优选的方面,核苷酸序列为 SEQ ID NO:1 所示。在另一个更优选的方面,核苷酸序列是质粒 pDZA2-7 中所含的序列,所述质粒包含在大肠杆菌 NRRL B-30704 中。在另一个优选的方面,核苷酸序列是 SEQ ID NO:1 的成熟多肽编码区。在另一个更优选的方面,核苷酸序列是质粒 pDZA2-7 中所含的成熟多肽编码区,所述质粒包含在大肠杆菌 NRRL B-30704 中。本发明还涵盖编码具有氨基酸序列 SEQ ID NO:2 的多肽或其成熟多肽的核苷酸序列,它与 SEQ ID NO:1 因遗传密码的简并性而不同。本发明还涉及编码具有增强分解纤维素活性的 SEQ ID NO:2 之片段的 SEQ ID NO:1 之子序列。

[0100] 本发明还涉及在 SEQ ID NO:1 的成熟多肽编码序列中包含至少一处突变的突变多核苷酸,其中突变核苷酸序列编码由 SEQ ID NO:2 第 23 至 250 位氨基酸组成的多肽。

[0101] 用于分离或克隆编码多肽的多核苷酸的技术在本领域是已知的,包括从基因组 DNA 分离,从 cDNA 制备,或者两者的组合。例如,通过利用众所周知的聚合酶链式反应(PCR)或用抗体筛选表达文库以检测出具有共同结构特征的克隆 DNA 片段,可以实现从这些基因组 DNA 克隆本发明的多核苷酸序列。例如参见 Innis 等人,1990,《PCR:A Guide to Methods and Application》,Academic Press, New York。还可以使用其它核酸扩增程序,诸如连接酶链式反应(LCR)、连接激活转录(LAT)、和基于核苷酸序列的扩增(NASBA)。可从梭孢壳属菌株或者另外的或相关的生物体克隆多核苷酸,因此它可以是例如核苷酸序列中多肽编码区的等位或种间变体。

[0102] 本发明还涉及具有如下核苷酸序列的编码活性多肽的多核苷酸,该核苷酸序列与 SEQ ID NO:1 的成熟多肽编码序列(即第 67 至 796 位核苷酸)具有至少 70%、优选至少 75%、更优选至少 80%、更优选至少 85%、甚至更优选至少 90%、最优选至少 95%、且甚至最优选至少 97%、98% 或 99% 的同一性程度。

[0103] 为了合成与本发明多肽基本上类似的多肽,可能必需修饰编码多肽的核苷酸序

列。术语与多肽“基本上类似”是指多肽的非天然存在形式。这些多肽可能因某些改造而区别于从其天然来源分离的多肽,例如在比活、热稳定性、最佳 pH 等方面不同的人工变体。可以在 SEQ ID NO:1 多肽编码区所呈现的核苷酸序列基础上构建变体序列,例如其子序列,和 / 或通过引入不会导致由核苷酸序列编码的多肽具有另一种氨基酸序列、而是使之符合意图用来生产酶的宿主生物体的密码子使用率的核苷酸替代,或通过引入可导致不同氨基酸序列的核苷酸替代。对于核苷酸替代的一般说明可参见例如 Ford 等人,1991, *Protein Expression and Purification* 2:95-107。

[0104] 对本领域技术人员而言,显而易见的是这些种替代可在分子功能的关键区域之外进行并且仍然可得到活性多肽。可根据本领域的已知程序来鉴别对于由本发明的分离多核苷酸所编码的多肽的活性来说至关重要的因此优选不进行替代的氨基酸残基,诸如定点诱变或丙氨酸扫描诱变(例如参见 Cunningham 和 Wells,1989, *Science* 244:1081-1085)。在后一种技术中,在分子中每一带正电荷残基处引入突变,然后对得到的突变分子测试增强分解纤维素活性以鉴定对分子活性至关重要的氨基酸残基。底物-酶相互作用位点也可通过三维结构分析来确定,其中三维结构可由诸如核磁共振分析、晶体学、或光亲和标记等技术来测定(例如参见 de Vos 等人,1992, *Science* 255:306-312; Smith 等人,1992, *Journal of Molecular Biology* 224:899-904; Wlodaver 等人,1992, *FEBS Letters* 309:59-64)。

[0105] 本发明还涉及编码本发明多肽的分离多核苷酸,它在很低严紧条件、优选低严紧条件、更优选中等严紧条件、更优选中等-高严紧条件、甚至更优选高严紧度条件、且最优选很高严紧条件下与 (i) SEQ ID NO:1 第 67 至 796 位核苷酸, (ii) SEQ ID NO:1 第 67 至 796 位核苷酸所含 cDNA 序列,或 (iii), (i) 或 (ii) 的互补链;或其等位变体和子序列发生杂交 (Sambrook 等人,1989, 同上),正如本文所定义的。

[0106] 本发明还涉及通过如下获得的编码具有增强分解纤维素活性的多肽的分离多核苷酸,即 (a) 在很低、低、中等、中等-高、高、或很高严紧条件下将 DNA 群体与 (i) SEQ ID NO:1 第 67 至 796 位核苷酸, (ii) SEQ ID NO:1 第 67 至 796 位核苷酸所含 cDNA 序列,或 (iii), (i) 或 (ii) 的互补链进行杂交;并 (b) 分离发生杂交的多核苷酸。

[0107] 在一个优选的方面,DNA 群体在低严紧条件下进行杂交。在一个更优选的方面,DNA 群体是在中等严紧条件下进行杂交。在一个甚至更优选的方面,DNA 群体在中等-高严紧条件下进行杂交。在一个最优选的方面,DNA 群体在高严紧条件下进行杂交。

[0108] 核酸构建体

[0109] 本发明还涉及包含与一种或多种控制序列可操作连接的本发明的分离多核苷酸的核酸构建体,所述控制序列在适当宿主细胞中在与控制序列相容的条件下指导编码序列的表达。

[0110] 可利用多种方法操作编码本发明多肽的分离多核苷酸以提供多肽的表达。根据表达载体的不同,可能希望或必需在将多核苷酸序列插入载体之前对其进行操作。利用重组 DNA 方法修饰多核苷酸序列的技术是为本领域众所周知的。

[0111] 控制序列可以是适当的启动子序列,即受到宿主细胞识别来表达编码本发明多肽的多核苷酸的核苷酸序列。启动子序列包含调节多肽表达的转录控制序列。启动子可以在所选宿主细胞中表现出转录活性的任何核苷酸序列,包括突变的、截短的、和杂合的启动子,而且可从编码与宿主细胞同源或异源的细胞外或细胞内多肽的基因中获得。

[0112] 用于指导本发明核酸构建体转录的适当启动子的实例,尤其是在细菌宿主细胞中,是从大肠杆菌 lac 操纵子、天蓝色链霉菌(*Streptomyces coelicolor*) 琼脂酶基因(dagA)、枯草芽孢杆菌果聚糖蔗糖酶基因(sacB)、地衣芽孢杆菌 α -淀粉酶基因(amyL)、嗜热脂肪芽孢杆菌生麦芽糖淀粉酶基因(amyM)、解淀粉芽孢杆菌 α -淀粉酶基因(amyQ)、地衣芽孢杆菌青霉素酶基因(penP)、枯草芽孢杆菌 xylA 和 xylB 基因、和原核生物的 β -内酰胺酶基因(Villa-Kamaroff 等人,1978, Proceedings of the National Academy of Sciences USA75 :3727-3731) 中获得的启动子,以及 tac 启动子(DeBoer 等人,1983, Proceedings of the National Academy of Sciences USA80 :21-25)。还有“Useful proteins from recombinant bacteria”, Scientific American,1980,242 :74-94 及 Sambrook 等人,1989,同上中描述的其它启动子。

[0113] 用于在丝状真菌宿主细胞中指导本发明核酸构建体转录的适当启动子的实例是从米曲霉 TAKA 淀粉酶、米赫根毛霉(*Rhizomucor miehei*) 天冬氨酸蛋白酶、黑曲霉中性 α -淀粉酶、黑曲霉酸稳定 α -淀粉酶、黑曲霉或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、有毒镰孢淀粉葡糖苷酶(W000/56900)、有毒镰孢 Daria (W000/56900)、有毒镰孢 Quinn (W000/56900)、尖孢镰孢胰蛋白酶样蛋白酶(W096/00787)、瑞氏木霉 β -葡糖苷酶、瑞氏木霉纤维二糖水解酶 I、瑞氏木霉纤维二糖水解酶 II、瑞氏木霉内切葡聚糖酶 I、瑞氏木霉内切葡聚糖酶 II、瑞氏木霉内切葡聚糖酶 III、瑞氏木霉内切葡聚糖酶 IV、瑞氏木霉内切葡聚糖酶 V、瑞氏木霉木聚糖酶 I、瑞氏木霉木聚糖酶 II、瑞氏木霉 β -木糖苷酶基因中获得的启动子,以及 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);及其突变的、截短的、和杂合的启动子。

[0114] 在酵母宿主中,有用的启动子是从酿酒酵母烯醇化酶(ENO-1)、酿酒酵母半乳糖激酶(GAL1)、酿酒酵母醇脱氢酶 / 甘油醛 -3- 磷酸脱氢酶(ADH1、ADH2/GAP)、酿酒酵母磷酸丙糖异构酶(TPI)、酿酒酵母金属硫蛋白(CUP1)、和酿酒酵母 3- 磷酸甘油酸激酶基因中所获得的。还有 Romanos 等人,1992, Yeast8 :423-488 中描述了用于酵母宿主细胞的其它有用启动子。

[0115] 控制序列也可以是适当的转录终止子序列,即由宿主细胞识别来终止转录的序列。终止子序列可操作连接于编码多肽的核苷酸序列的 3' 端。在所选择的宿主细胞中起作用的任何终止子都可用于本发明。

[0116] 对于丝状真菌宿主细胞优选的终止子是从米曲霉 TAKA 淀粉酶、黑曲霉葡糖淀粉酶、构巢曲霉邻氨基苯甲酸(anthranilate) 合酶、黑曲霉 α -葡糖苷酶、和尖孢镰孢胰蛋白酶样蛋白酶的基因中获得的。

[0117] 对于酵母宿主细胞优选的终止子是从酿酒酵母烯醇化酶、酿酒酵母细胞色素 C(CYC1)、和酿酒酵母甘油醛 -3- 磷酸脱氢酶基因中获得的。还有 Romanos 等人,1992,同上中描述的用于酵母宿主细胞的其它有用终止子。

[0118] 控制序列还可以是适当的前导序列,即对宿主细胞翻译很重要的 mRNA 非翻译区。前导序列可操作连接于编码多肽的核苷酸序列的 5' 端。在所选择的宿主细胞中起作用的任何前导序列都可用于本发明。

[0119] 对于丝状真菌宿主细胞优选的前导序列是从米曲霉 TAKA 淀粉酶和构巢曲霉磷酸

丙糖异构酶基因中获得的。

[0120] 对于酵母宿主细胞优选的前导序列是从酿酒酵母烯醇化酶(ENO-1)、酿酒酵母 3-磷酸甘油酸激酶、酿酒酵母 α -因子、和酿酒酵母乙醇脱氢酶 / 甘油醛-3-磷酸脱氢酶(ADH2/GAP) 基因中获得的。

[0121] 控制序列还可以是聚腺苷酸化序列,即可操作连接于核苷酸序列 3' 端的序列,当转录时可由宿主细胞识别为将聚腺苷残基添加到转录的 mRNA 上的信号。在所选择的宿主细胞中起作用的任何聚腺苷酸化序列都可用于本发明。

[0122] 对于丝状真菌宿主细胞优选的聚腺苷酸化序列是从米曲霉 TAKA 淀粉酶、黑曲霉葡糖淀粉酶、构巢曲霉邻氨基苯甲酸合酶、尖孢镰孢胰蛋白酶样蛋白酶、和黑曲霉 α -葡糖苷酶的基因获得的。

[0123] Guo 和 Sherman 等人,1995, *Molecular Cellular Biology* 15 :5983-5990 中描述了对酵母宿主细胞有用的聚腺苷酸化序列。

[0124] 控制序列还可以是信号肽编码区,它编码的氨基酸序列连接于多肽的氨基末端并指导编码的多肽进入细胞分泌途径。核苷酸序列的编码序列的 5' 端可固有的包含信号肽编码区,它在翻译读码框中与编码分泌多肽的编码区段天然连接。或者,编码序列的 5' 端可以包含对编码序列是外来的信号肽编码区。如果编码序列天然不含信号肽编码区,则可能需要外来的信号肽编码区。或者,外来的信号肽编码区可简单的替换天然信号肽编码区以便增强多肽的分泌。然而,指导表达的多肽进入所选择宿主细胞分泌途径的任何信号肽编码区均可用于本发明。

[0125] 用于细菌宿主细胞有效的信号肽编码区是从芽孢杆菌 NCIB11837 的生麦芽糖淀粉酶、嗜热脂肪芽孢杆菌的 α -淀粉酶、地衣芽孢杆菌的枯草杆菌蛋白酶、地衣芽孢杆菌的 β -内酰胺酶、嗜热脂肪芽孢杆菌的中性蛋白酶类(nprT、nprS、nprM)、和枯草芽孢杆菌的 prsA 基因获得的信号肽编码区。还有 Simonen 和 Palva,1993, *Microbiological Reviews* 57 :109-137 中描述的其它信号肽。

[0126] 对于丝状真菌宿主细胞有效的信号肽编码区是从米曲霉 TAKA 淀粉酶、黑曲霉中性淀粉酶、黑曲霉葡糖淀粉酶、米赫根毛霉天冬氨酸蛋白酶、*Humicola insolens* 纤维素酶、*Humicola insolens* 内切葡聚糖酶 V、和 *Humicola lanuginosa* 脂肪酶的基因获得的信号肽编码区。

[0127] 在一个优选的方面,信号肽编码区是编码 SEQ ID NO:2 第 1 至 22 位氨基酸的 SEQ ID NO:1 第 1 至 66 位核苷酸。

[0128] 对于酵母宿主细胞有用的信号肽是从酿酒酵母 α -因子和酿酒酵母转化酶的基因获得的。还有 Romanos 等人,1992,同上中描述了其它的有用的信号肽编码区。

[0129] 控制序列还可以是前肽(propeptide) 编码区,它编码位于多肽氨基末端的氨基酸序列。所得的多肽称为酶原(proenzyme) 或多肽原(propolypeptide) (或有时候称作酶原(zymogen))。多肽原一般是没有活性的,且可通过前肽的催化或自催化切割从多肽原转变成成熟有活性的多肽。前肽编码区可从枯草芽孢杆菌碱性蛋白酶(aprE)、枯草芽孢杆菌中性蛋白酶(nprT)、酿酒酵母 α -因子、米赫根毛霉天冬氨酸蛋白酶、和嗜热毁丝霉漆酶(W095/33836) 的基因获得。

[0130] 当信号肽和前肽区都存在于多肽的氨基末端时,前肽区位于靠近多肽氨基末端的

位置,而信号肽区位于靠近前肽区的氨基末端位置。

[0131] 还可能需要加上调节序列以使多肽的表达可以相对于宿主细胞的生长进行调节。调节系统的实例是能响应化学或物理刺激(包括调节化合物的存在)而引起基因表达的开启或关闭的那些。在原核系统中,调节系统包括 lac、tac、和 trp 操纵子系统。在酵母中,可使用 ADH2 系统或 GAL1 系统。在丝状真菌中,可使用 TAKA α -淀粉酶启动子、黑曲霉葡萄糖淀粉酶启动子、和米曲霉葡萄糖淀粉酶启动子作为调节序列。调节序列的其它实例是允许基因扩增的那些。在真核系统中,这些包括在氨甲喋呤存在时扩增的二氢叶酸还原酶基因,以及有重金属存在时扩增的金属硫蛋白基因。在这些情况中,编码多肽的核苷酸序列将与调节序列可操作连接。

[0132] 表达载体

[0133] 本发明还涉及包含本发明的多核苷酸、启动子、及转录和翻译终止信号的重组表达载体。可以将本文所述的各种核酸和控制序列连接起来以产生重组表达载体,此重组表达载体可包括一个或多个便利的限制性位点使得在这些位点可插入或替换编码多肽的核苷酸序列。或者,本发明的核苷酸序列可通过将核苷酸序列或含有序列的核酸构建体插入适当的表达载体进行表达。构建表达载体时,将编码序列置于载体中以使编码序列与适当的表达控制序列可操作连接。

[0134] 重组表达载体可以是可方便的接受重组 DNA 操作并能引起核苷酸序列表达的任何载体(例如质粒或病毒)。载体的选择一般取决于载体与引入此载体的宿主细胞之间的相容性。载体可以是线性的或闭合环状的质粒。

[0135] 载体可以是自主复制载体,即载体可以染色体外实体的形式存在,其复制独立于染色体的复制,例如质粒、染色体外元件、微型染色体、或人工染色体。载体可包含任何用于确保自我复制的手段。或者,载体可以是这样的一种载体,它在导入宿主细胞后,整合到基因组中并与整合的染色体一起复制。此外,可使用单一载体或质粒,或是一起含有待引入宿主细胞基因组的全部 DNA 的两种或多种载体或质粒,或者转座子。

[0136] 本发明载体优选包含一种或多种选择标记,所述选择标记允许易于选择转化的、转染的、转导的等细胞。选择标记是一种基因,其产物有助于产生抗微生物剂或病毒抗性、重金属抗性、营养缺陷型变成原养型等等。

[0137] 细菌选择标记的实例是来自枯草芽孢杆菌或地衣芽孢杆菌的 *dal* 基因,或者赋予抗生素抗性诸如氨苄青霉素、卡那霉素、氯霉素或四环素抗性的标记。适于酵母宿主细胞的标记有 ADE2、HIS3、LEU2、LYS2、MET3、TRP1、和 URA3。在丝状宿主细胞中使用的选择标记包括但不限于:amdS(乙酰胺酶)、argB(鸟氨酸氨甲酰基转移酶)、bar(磷丝菌素乙酰基转移酶)、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(sulfate adenylyltransferase))、和 trpC(邻氨基苯甲酸合酶)及其等价物。优选用于曲霉属细胞的是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。

[0138] 本发明的载体优选包含允许载体整合到宿主细胞基因组中或者允许载体在细胞中不依赖于基因组自主复制的元件。

[0139] 对整合到宿主细胞基因组来说,载体可依赖编码多肽的多核苷酸序列或任何其它载体元件通过同源或非同源重组整合到基因组中。或者,载体可包含附加的核苷酸序列用

于指导通过同源重组在染色体精确位置整合进入宿主细胞基因组中。为了提高在精确位置整合的可能性,整合元件应优选包含足够数目的与对应靶序列具有高度同一性的核酸,诸如 100-10,000 个碱基对、优选 400-10,000 个碱基对、且最优选 800-10,000 个碱基对,从而提高同源重组的几率。整合元件可以是与宿主细胞基因组中的靶序列同源的任何序列。此外,整合元件可以是非编码或编码核苷酸序列。另一方面,载体可通过非同源重组整合到宿主细胞的基因组中。

[0140] 对自主复制来说,载体还可包含能使载体在所讨论的宿主细胞中自主复制的复制起点。复制起点可以是在细胞中起作用的调节自主复制的任何质粒复制因子。术语“复制起点”或“质粒复制子”在本文中定义为能使质粒或载体在体内复制的核苷酸序列。

[0141] 细菌复制起点的实例是使得在大肠杆菌中复制的质粒 pBR322、pUC19、pACYC177 和 pACYC184 的复制起点,以及使得在芽孢杆菌中复制的 pUB110、pE194、pTA1060 和 pAM β 1 的复制起点。

[0142] 用于酵母宿主细胞的复制起点的实例是 2 μ m 复制起点、ARS1、ARS4、ARS1 和 CEN3 的组合、及 ARS4 和 CEN6 的组合。

[0143] 可用于丝状真菌细胞的复制起点的实例有 AMA1 和 ANS1(Gems 等人,1991,Gene98 : 61-67 ;Cullen 等人,1987,Nucleic Acids Research15 :9163-9175 ;W000/24883)。AMA1 基因的分离和包含该基因的质粒或载体的构建可根据 W000/24883 中所公开的方法来完成。

[0144] 可将本发明多核苷酸一个以上的拷贝插入宿主细胞中以提高基因产物的产量。提高多核苷酸拷贝数可通过将序列的至少一个附加拷贝整合进宿主细胞基因组中或通过可将扩增的选择标记基因加入多核苷酸来获得,在其中包含选择标记基因的扩增拷贝的细胞,以及因此多核苷酸的附加拷贝可在存在适当的选择剂时通过培养此细胞选择得出。

[0145] 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所熟知的(例如参见 Sambrook 等人,1989,同上)。

[0146] 宿主细胞

[0147] 本发明还涉及包含本发明多核苷酸的重组宿主细胞,它们可有利的用于多肽的重组生产。如前所述,可将包含本发明多核苷酸的载体引入宿主细胞以使载体保持为染色体的整合体或作为自主复制的染色体外载体。术语“宿主细胞”包括所有那些由于在复制期间出现的突变而不同于亲本细胞的亲本细胞的后代。宿主细胞的选择在很大程度上依赖于编码多肽的基因和它的来源。

[0148] 宿主细胞可以是单细胞微生物,例如原核生物,或非单细胞微生物,例如真核生物。

[0149] 有用的单细胞微生物是细菌细胞,诸如革兰氏阳性细菌,包括但不限于芽孢杆菌属细胞,例如嗜碱芽孢杆菌、解淀粉芽孢杆菌、短芽孢杆菌、环状芽孢杆菌、*Bacillus clausii*、凝结芽孢杆菌、灿烂芽孢杆菌、迟缓芽孢杆菌、地衣芽孢杆菌、巨大芽孢杆菌、嗜热脂肪芽孢杆菌、枯草芽孢杆菌、和苏云金芽孢杆菌;或链霉菌属细胞,例如浅青紫链霉菌和鼠灰链霉菌,或者革兰氏阴性细菌,诸如大肠杆菌和假单胞菌。在一个优选的方面,细菌宿主细胞是迟缓芽孢杆菌、地衣芽孢杆菌、嗜热脂肪芽孢杆菌、或枯草杆菌细胞。在另一个优选的方面,芽孢杆菌属细胞是嗜碱的芽孢杆菌。

[0150] 将载体导入细菌宿主细胞可通过例如原生质体转化(参见例如 Chang 和 Cohen,

1979, *Molecular General Genetics* 168:111-115)、利用感受态细胞(参见例如 Young 和 Spizizen, 1961, *Journal of Bacteriology* 81:823-829; 或 Dubnau 和 Davidoff-Abelson, 1971, *Journal of Molecular Biology* 56:209-221)、电穿孔(参见例如 Shigekawa 和 Dower, 1988, *Biotechniques* 6:742-751)、或接合(参见例如 Koehler 和 Thorne, 1987, *Journal of Bacteriology* 169:5771-5278)来实现。

[0151] 宿主细胞还可以是真核细胞, 诸如哺乳动物、昆虫、植物、或真菌细胞。

[0152] 在一个优选的方面, 宿主细胞是真菌细胞。“真菌”在用于本文时包括子囊菌门(Acomycota)、担子菌门(Basidiomycota)、壶菌门(Chytridiomycota)、和接合菌门(Zygomycota)(如 Hawksworth 等人, 在《Ainsworth and Bisby's Dictionary of The Fungi》, 第八版, 1995, CAB International, University Press, Cambridge, UK 中所定义的), 以及卵菌门(Oomycota)(如 Hawksworth 等人, 1995, 同上第 171 页中所引用的)和所有有丝分裂孢子真菌(Hawksworth 等人, 1995, 同上)。

[0153] 在一个更优选的方面, 真菌宿主细胞是酵母细胞。“酵母”在用于本文时包括产子囊酵母(内孢霉目 Endomycetales)、产担孢子酵母、和属于不完全真菌的酵母(芽生菌纲 Blastomycetes)。由于酵母的分类今后还会变化, 对于本发明来说, 酵母应该是如在《Biology and Activities of Yeast》(Skinner, F. A.、Passmore, S. M. 和 Davenport, R. R. 编, Soc. App. Bacteriol. Symposium Series No. 9, 1980)中所定义的。

[0154] 在一个更优选的方面, 酵母宿主细胞是假丝酵母属、汉逊氏酵母属(Hansenula)、克鲁维氏酵母属、毕赤氏酵母属、糖酵母属、裂殖糖酵母属、或 Yarrowia 属细胞。

[0155] 在一个最优选的方面, 酵母宿主细胞是卡尔斯伯糖酵母、酿酒酵母、糖化糖酵母、道拉斯糖酵母、克鲁弗糖酵母、诺第糖酵母、或卵形糖酵母细胞。在另一个最优选的方面, 酵母宿主细胞是乳克鲁维氏酵母(Kluyveromyces lactis)细胞。在另一个最优选的方面, 酵母宿主细胞是 Yarrowia lipolytica 细胞。

[0156] 在另一个更优选的方面, 真菌宿主细胞是丝状真菌细胞。“丝状真菌”包括真菌门(Eumycota)和卵菌门亚门的所有丝状体(如 Hawksworth 等人, 1995, 同上中所定义的)。丝状真菌通常是以由壳多糖、纤维素、葡聚糖、壳聚糖、甘露聚糖、及其他复合多糖组成的菌丝壁为特征。营养生长表现为菌丝伸长, 碳代谢是专性需氧的。相反, 酵母诸如酿酒酵母的营养生长表现为单细胞菌体的出芽而碳代谢可以是发酵性的。

[0157] 在一个甚至更优选的方面, 丝状真菌宿主细胞是 Acremonium 属、曲霉属、短梗霉属、烟管菌属(Bjerkandera)、Ceriporiopsis 属、鬼伞属(Coprinus)、革盖菌属(Coriolus)、隐球酵母属、Filibasidium 属、镰孢属、腐质霉属、Magnaporthe 属、毛霉属、毁丝霉属、Neocallimastix 属、脉孢霉菌、拟青霉属、青霉属、Phanerochaete 属、射脉菌属(Phlebia)、Piromyces 属、侧耳菌属(Pleurotus)、裂褶菌属、Talaromyces 属、嗜热子囊菌属、梭孢壳属、Tolyocladium 属、栓菌属(Trametes)、或木霉属的细胞。

[0158] 在一个最优选的方面, 丝状真菌宿主细胞是泡盛曲霉、烟曲霉、臭曲霉、日本曲霉、构巢曲霉、黑曲霉、或米曲霉细胞。在另一个最优选的方面, 丝状真菌宿主细胞是杆孢状镰孢、谷孢镰孢、Fusarium crookwellense、黄色镰孢、禾谷镰孢、禾赤镰孢、异孢镰孢、合欢木镰孢、尖孢镰孢、网状镰孢、粉红镰孢、接骨木镰孢、肤色镰孢、拟枝孢镰孢、硫色镰孢、近念珠状镰孢、单端孢镰孢、或有毒镰孢细胞。在另一个最优选的方面, 丝状真菌宿主细胞是黑

刺烟管菌(*Bjerkandera adusta*)、*Ceriporiopsis aneirina*、*Ceriporiopsis caregiea*、*Ceriporiopsis gilvescens*、*Ceriporiopsis pannocinta*、*Ceriporiopsis rivulosa*、*Ceriporiopsis subrufa*、*Ceriporiopsis subvermispora*、灰色鬼伞、毛革盖菌(*Coriolus hirsutus*)、*Humicola insolens*、*Humicola lanuginosa*、米赫毛霉、嗜热毁丝霉、粗糙脉孢菌、产紫青霉、*Phanerochaete chrysosporium*、辐射射脉菌(*Phlebia radiata*)、*Pleurotus eryngii*、*Thielavia terrestris*、*Trametes villosa*、*Trametes versicolor*、哈茨木霉、康宁木霉、长枝木霉、瑞氏木霉、或绿色木霉细胞。

[0159] 真菌细胞可以本身已知的方式通过涉及原生质体形成、原生质体转化、和细胞壁再生的过程进行转化。EP238023 和 Yelton 等人,1984, *Proceedings of the National Academy of Sciences USA*81:1470-1474 中描述了适合曲霉属和木霉属宿主细胞转化的过程。Malardier 等人,1989, *Gene*78:147-159 和 W096/00787 中描述了适合转化镰孢属物种的方法。可利用 Becker 和 Guarente 在 Abelson, J. N. 和 Simon, M. I. 编的《*Guide to Yeast Genetics and Molecular Biology*》, *Methods in Enzymology*, Vol. 194, pp182-187, Academic Press 公司, New York; Ito 等人,1983, *Journal of Bacteriology*153:163; 及 Hinnen 等人,1978, *Proceedings of the National Academy of Sciences USA*75:1920 中所描述的程序转化酵母。

[0160] 生产方法

[0161] 本发明还涉及生产本发明多肽的方法,包括:(a) 在有助于产生所述多肽的条件下培养细胞,所述细胞能够在其野生型产生多肽;并(b) 回收所述多肽。优选的是,所述细胞是嗜热子囊菌属细胞,更优选的是橙色嗜热子囊菌。

[0162] 本发明还涉及生产本发明多肽的方法,包括:(a) 在有助于产生所述多肽的条件下培养宿主细胞;并(b) 回收所述多肽。

[0163] 本发明还涉及生产本发明多肽的方法,包括:(a) 在有助于产生所述多肽的条件下培养包含如下突变核苷酸序列的宿主细胞,所述突变核苷酸序列在 SEQ ID NO:1 的成熟多肽编码区中具有至少一处突变,其中突变核苷酸序列编码由 SEQ ID NO:2 第 23 至 250 位氨基酸组成的多肽,并(b) 回收所述多肽。

[0164] 在本发明的生产方法中,利用本领域众所周知的方法在适于产生所述多肽的培养基中培养细胞。例如,细胞培养可在实验室或工业发酵罐中在适当的培养基和可使所述多肽表达和/或分离的条件下,通过摇瓶培养、小规模或大规模发酵(包括连续的、分批的、补料-分批、或固态发酵)来进行。培养可通过本领域已知的程序在含有碳和氮源以及无机盐的合适营养培养基中进行。适宜的培养基可从商业供应商获得,或者可根据已公布的配方来制备(例如参见美国典型培养物保藏中心的目录)。如果多肽分泌到培养基中,那么可直接从培养基中回收多肽。如果多肽不分泌,那么可从细胞裂解物中进行回收。

[0165] 可利用本文描述的方法来检测具有增强分解纤维素活性的多肽。

[0166] 所得多肽可通过本领域已知的方法回收。例如,多肽可通过常规流程从培养基中回收,包括但不限于:离心、过滤、萃取、喷射-干燥、蒸发、或沉淀。

[0167] 本发明的多肽可通过本领域已知的多种程序纯化,包括但不限于:层析(例如离子交换、亲和、疏水、层析聚焦、和大小排阻)、电泳(例如制备性等电聚焦)、差异溶解(例如硫酸铵沉淀)、SDS-PAGE、或萃取(例如参见《*Protein Purification*》, J.-C. Janson 和 Lars

Ryden 编, VCH Publishers, New York, 1989), 以获得基本上纯的多肽。

[0168] 植物

[0169] 本发明还涉及经编码本发明具有增强分解纤维素活性的多肽的核苷酸序列转化而以可回收量表达和产生多肽的转基因植物、植株部分、或植物细胞。所述多肽可从植物或植株部分回收。或者, 将含有重组多肽的植物或植株部分就此用于改善食物或饲料的质量, 例如提高营养价值、可口性和流变性、或破坏抗营养因子。

[0170] 转基因植物可以是双子叶的(双子叶植物)或单子叶的(单子叶植物)。单子叶植物的实例有草(grasses), 例如草地草(蓝草, 早熟禾属(Poa)); 草料草, 诸如羊茅属(Festuca)、黑麦草属(Lolium); 温带草, 例如剪股颖属(Agrostis); 和谷类, 例如小麦、燕麦、黑麦、大麦、水稻、高粱、和玉蜀黍(玉米)。

[0171] 双子叶植物的实例有烟草, 豆科作物诸如羽扇豆、马铃薯、甜菜、豌豆、菜豆、和大豆, 和十字花科植物(十字花科 Brassicaceae), 诸如花椰菜、油菜籽、和亲缘很近的模式生物体拟南芥(*Arabidopsis thaliana*)。

[0172] 植株部分的实例有茎、愈伤组织、叶、根、果实、种子、和块茎, 以及构成这些部分的单独组织, 例如表皮、叶肉、薄壁组织、脉管组织、分生组织。特殊的植物细胞小室, 诸如叶绿体、质外体、线粒体、液泡、过氧化物酶体、和细胞质也认为是植株部分。此外, 任何植物细胞, 无论什么组织来源, 也认为是植株部分。类似的, 植株部分, 诸如有利于本发明应用的特定的分离组织和细胞也认为是植株部分, 例如胚、胚乳、糊粉、和种皮。

[0173] 这些植物、植株部分、和植物细胞的后代也包括在本发明范围之内。

[0174] 表达本发明多肽的转基因植物或植物细胞的构建可根据本领域已知的方法进行。简而言之, 植物或植物细胞的构建可通过将一种或多种编码本发明多肽的表达构建体掺入植物宿主基因组中并将所得的经修饰植物或植物细胞繁殖成转基因植物或植物细胞来进行。

[0175] 表达构建体方便的是核酸构建体, 它含有编码本发明多肽的多核苷酸且该多核苷酸可操作连接在所选定的植物或植株部分中表达此核苷酸序列所需的适当调控序列。而且, 表达构建体可包含用于鉴定整合了此表达构建体的宿主细胞的选择标记, 和将此构建体导入所讨论的植物所需要的 DNA 序列(后者取决于所用导入 DNA 的方法)。

[0176] 例如, 根据希望在何时、在何处、和怎样表达多肽来选择调控序列, 诸如启动子和终止子序列和任意的信号或转运序列。例如, 编码本发明多肽的基因的表达可以是组成性的或可诱导的, 或者是发育、阶段、或组织特异的, 基因产物可以是靶向特定组织或植株部分诸如种子或叶。调控序列有例如 Tague 等人, 1988, *Plant Physiology* 86:506 所述。

[0177] 对于组成性表达, 可使用 35S-CaMV、玉蜀黍泛蛋白 1、和水稻肌动白 1 启动子(Franck 等人, 1980, *Cell* 21:285-294; Christensen 等人, 1992, *Plant Mol. Biol.* 18:675-689; Zhang 等人, 1991, *Plant Cell* 3:1155-1165)。器官特异性启动子可以是例如来自贮藏库组织(storage sink tissues) 诸如种子、马铃薯块茎、和果实(Edwards 和 Coruzzi, 1990, *Ann Rev Genet* 24:275-303), 或来自于代谢库(metabolic sink tissues) 组织诸如分生组织(Ito 等人, 1994, *Plant Mol Biol* 24:863-878), 种子特异性启动子诸如来自水稻的谷蛋白、醇溶谷蛋白、球蛋白、或清蛋白启动子(Wu 等人, 1998, *Plant and Cell Physiology* 39:885-889), 来自蚕豆(*Vicia faba*) 豆球蛋白 B4 和未知种子蛋白基因的蚕

豆启动子(Conrad 等人,1998, Journal of Plant Physiology152 :708-711),来自种子油体(seed oil body)蛋白的启动子(Chen 等人,1998, Plant and Cell Physiology39 :935-941),来自欧洲油菜(*Brassica napus*)的 贮藏蛋白 napA 启动子,或者本领域已知的任何其它种子特异启动子,例如 W091/14772 中所述。此外,启动子可以是叶特异性启动子,诸如来自水稻或番茄的 *rbcs* 启动子(Kyozuka 等人,1993, Plant Physiology102 :991-1000),小球藻病毒腺嘌呤甲基转移酶基因启动子(Mitra 和 Higgins,1994, Plant Molecular Biology26 :85-93),或来自水稻的 *aldP* 基因启动子(Kagaya 等人,1995, Molecular and General Genetics248 :668-674),或者伤口诱导性启动子诸如马铃薯 *pin2* 启动子(Xu 等人,1993, Plant Molecular Biology22 :573-588)。同样,启动子可通过非生物处理例如温度、干旱、或盐度变化来诱导,或通过外源施加能激活启动子的物质,例如乙醇、雌激素、植物激素诸如乙烯、脱落酸和赤霉素以及重金属来诱导。

[0178] 还可使用启动子增强子元件,从而在植物中获得本发明多肽的更高表达。例如,启动子增强子元件可以是位于启动子和编码本发明多肽的核苷酸序列之间的内含子。例如 Xu 等人,1993,同上中公开了利用水稻肌动蛋白 1 基因的的第一个内含子来增强表达。

[0179] 选择标记基因和表达构建体的任何其它部分可选自本领域可利用的那些范围。

[0180] 可根据本领域已知的常规技术将核酸构建体掺入植物基因组中,包括土壤杆菌介导的转化、病毒介导的转化、显微注射、粒子轰击、生物射弹转化(biolistic transformation)、和电穿孔(Gasser 等人,1990, Science244 :1293 ;Potrykus,1990, Bio/Technology8 :535 ;Shimamoto 等人,1989, Nature338 :274)。

[0181] 目前,根瘤农杆菌(*Agrobacterium tumefaciens*)介导的基因转移是用于产生转基因双子叶植物的首选方法(其综述可参见 Hooykas 和 Schilperoort,1992, Plant Molecular Biology19 :15-38),而且也可用于转化单子叶植物,尽管这些植物常使用其它转化方法。目前,选择用于产生转基因单子叶植物的方法是粒子轰击(以转化用 DNA 包被的金或钨微粒)胚的愈伤组织或发育中的胚(Christou,1992, Plant Journal2 :275-281 ; Shimamoto,1994, Current Opinion Biotechnology5 :158-162 ;Vasil 等人,1992, Bio/Technology10 :667-674)。如 Omirulleh 等人,1993,Plant Molecular Biology21 :415-428 中所描述的,单子叶植物转化的替代方法是基于原生质体转化来进行的。

[0182] 在转化后,根据本领域熟知的方法选出已掺入了表达构建体的转化体,并使其再生成为完整植株。通常转化程序设计成在再生期间或在后代中选择性清除选择基因,例如通过两种不同 T-DNA 构建体的共转化或通过特异重组酶进行的选择基因的位点特异切除来进行。

[0183] 本发明还涉及生产本发明多肽的方法,包括(a)在有助于产生所述多肽的条件下,培养含有如下多核苷酸的转基因植物或植物细胞,所述多核苷酸编码本发明具有增强分解纤维素活性的多肽;并(b)回收所述多肽。

[0184] 增强分解纤维素活性的消除或降低

[0185] 本发明还涉及生产亲本细胞突变体的方法,包括破坏或删除编码本发明多肽的多核苷酸序列或其部分,这导致突变体细胞产生的多肽比在相同条件下培养的亲本细胞少。

[0186] 可使用本技术领域众所周知的方法通过降低或消除编码本发明多肽的核苷酸序列的表达来构建突变体细胞,例如使用插入、破坏、替换、或删除。在一个优选的方面,将核

核苷酸序列失活。例如,要修饰或失活的核苷酸序列可以是编码区或对或其活性必需的部分,或编码区表达所需要的调控元件。这种调控或控制序列的实例可以是启动子序列或其功能部分,即足以影响核苷酸序列表达的部分。可修饰的其它控制序列包括但不限于前导序列、聚腺苷酸化序列、前肽序列、信号肽序列、转录终止子、和转录激活因子。

[0187] 核苷酸序列的修饰或失活可通过将亲本细胞诱变,并选择所述核苷酸序列的表达降低或消除的突变细胞来进行。所述诱变可以是特异的或随机的,例如通过使用适宜的物理或化学诱变剂、通过使用适宜的寡核苷酸、或通过对 DNA 序列 PCR 产生诱变来进行。此外,诱变可通过使用这些诱变剂的任意组合来进行。

[0188] 适于本发明目的的物理或化学的诱变剂的实例包括紫外线(UV)照射、羟胺、N-甲基-N'-硝基-N-亚硝基胍(MNNG)、O-甲基羟胺、亚硝酸、甲磺酸乙酯(EMS)、亚硫酸氢钠、甲酸、和核苷酸类似物。

[0189] 在使用这些试剂时,诱变一般是这样进行的:在适宜条件下,在存在所选的诱变处理试剂下保温用于诱变的亲本细胞,并筛选和/或选择展示基因表达降低或不表达的突变细胞来。

[0190] 核苷酸序列的修饰或失活可通过在基因中或在其转录或翻译所需要的调控元件中引入、替换、或消除一个或多个核苷酸来进行。例如,可插入或消除核苷酸从而形成终止密码子的引入、起始密码子的消除、或开放读码框的改变。这些修饰或灭活可根据本领域已知的方法通过定点诱变或 PCR 产生诱变来完成。虽然,原则上,修饰可在体内进行,即直接在表达要修饰的核苷酸序列的细胞上进行,优选的是如下面实例中所示在体外进行修饰。

[0191] 适宜的消除或降低细胞核苷酸序列表达的方法的实例是以基因替换、基因删除、或基因破坏技术为基础的。例如,在基因破坏方法中,将相当于内源核苷酸序列的核酸在体外进行诱变以产生缺陷的核酸序列,然后将其转化到亲本细胞中以产生缺陷基因。通过同源重组,缺陷的核酸序列替换内源核苷酸序列。所希望的是,缺陷的核苷酸序列还编码可用于选择核苷酸序列已经修饰或破坏了的转化体的标记。在一个特别优选的实施方案中,诸如用本文所述选择标记来破坏核苷酸序列。

[0192] 或者,核苷酸序列的修饰或失活可通过已建立的反义或 RNAi 技术使用与核苷酸序列互补的序列来进行。更具体的说,细胞核苷酸序列的表达可通过引入与基因核苷酸序列互补的序列来降低或消除,其中该引入的序列可在细胞中转录并能与细胞中所产生的 mRNA 发生杂交。在允许互补的反义核苷酸序列与 mRNA 发生杂交的条件下,所翻译蛋白质的量由此降低或消除。

[0193] 本发明此外涉及在编码所述多肽的多核苷酸序列或其控制序列中包含破坏或删除的亲本细胞的突变细胞,这导致突变细胞产生的所述多肽比亲本细胞少或不产生多肽。

[0194] 这样构建的多肽缺陷的突变细胞作为宿主细胞用于天然和/或异源蛋白质的表达特别有用。因此,本发明还涉及生产天然或异源蛋白质的方法,包括(a)在有助于产生蛋白质的条件下培养突变细胞;并(b)回收所述多肽。术语“异源蛋白质”在本文中定义为对宿主细胞来说不是天然的蛋白质,经过修饰改变了天然序列的天然蛋白质,或是经过重组 DNA 技术处理宿主细胞从而使其表达发生了量变的天然蛋白质。

[0195] 在另一个方面,本发明涉及通过发酵生成本发明多肽和目的蛋白质产物二者的细胞来生产本质上没有增强分解纤维素活性的蛋白质产物的方法,其中在发酵完成之前、当

中、或之后通过往发酵液中加入有效量的可抑制增强分解纤维素活性的试剂，并从发酵液中回收目的产物，以及任选进行回收产物的进一步纯化。

[0196] 在另一个方面，本发明涉及通过下列步骤来生产本质上没有增强分解纤维素的蛋白质产物的方法，在允许产物表达的条件下培养细胞，并使所得培养液进行 pH 和温度联合处理以便显著地降低增强分解纤维素活性，并从培养液中回收产物。或者，pH 和温度联合处理可在从培养液中回收的酶制备物上进行。任选地，pH 和温度联合处理增强可与分解纤维素抑制剂处理结合。

[0197] 根据本发明的这个方面，可能消除至少 60%、优选至少 75%、更优选至少 85%、仍更优选至少 95%、且最优选至少 99% 的增强分解纤维素活性。通过使用这种方法可实现完全消除增强分解纤维素活性。

[0198] pH 和温度联合处理优选在 pH 为 4-5 和温度为 80-90°C 进行足够的一段时间以达到预期的效果。

[0199] 可使用本领域已知的方法进行培养和目的产物的纯化。

[0200] 本发明用于生产本质上不含增强分解纤维素的产物的方法在真核多肽，特别是真菌蛋白质，诸如酶的生产中特别令人感兴趣。例如，酶可选自淀粉分解酶、脂肪分解酶、蛋白水解酶、纤维素分解酶、氧化还原酶、或植物细胞壁降解酶。这些酶的实例包括氨肽酶、淀粉酶、淀粉葡糖苷酶、糖酶、羧肽酶、过氧化氢酶、纤维二糖水解酶、纤维素酶、壳多糖酶、角质酶、环糊精糖基转移酶、脱氧核糖核酸酶、内切葡聚糖酶、酯酶、半乳糖苷酶、 β -半乳糖苷酶、葡糖淀粉酶、葡萄糖氧化酶、 α -或 β -葡糖苷酶、卤代过氧化物酶(haloperoxidase)、半纤维素酶、转化酶、异构酶、漆酶、连接酶、脂肪酶、裂解酶、甘露糖苷酶、氧化酶、果胶分解酶、过氧化物酶、植酸酶、酚氧化酶、多酚氧化酶、蛋白水解酶、核糖核酸酶、转移酶、转谷氨酰胺酶、或木聚糖酶。增强分解纤维素缺陷的细胞还可用于表达具有药学价值的异源蛋白质，诸如激素、生长因子、受体等等。

[0201] 应理解的是术语“真核多肽”不仅包括天然多肽，还包括经过氨基酸替代、删除或添加的修饰或是用以提高活性、热稳定性、pH 耐受性等等其它类似修饰的多肽，例如酶。

[0202] 在另一个方面，本发明涉及通过本发明方法生产的本质上没有增强分解纤维素活性的蛋白质产物。

[0203] 组合物

[0204] 本发明还涉及包含本发明多肽的组合物。优选的是，组合物富集了这类多肽。术语“富集”是指组合物的增强分解纤维素活性已经被提高，例如富集因子为至少 1.1。

[0205] 所述组合物可包含本发明的多肽作为主要组分，例如单一组分组合物。或者，组合物还可含有一种或多种酶活性，诸如氨肽酶、淀粉酶、糖酶、羧肽酶、过氧化氢酶、纤维二糖水解酶、纤维素酶、壳多糖酶、角质酶、环糊精糖基转移酶、脱氧核糖核酸酶、内切葡聚糖酶、酯酶、 α -半乳糖苷酶、 β -半乳糖苷酶、葡糖淀粉酶、 α -葡糖苷酶、 β -葡糖苷酶、卤代过氧化物酶、转化酶、漆酶、脂肪酶、甘露糖苷酶、氧化酶、果胶分解酶、转谷氨酰胺酶、过氧化物酶、植酸酶、多酚氧化酶、蛋白水解酶、核糖核酸酶、转谷氨酰胺酶、或木聚糖酶。通过属于下列属的微生物可产生附加的酶，例如曲霉属，优选棘孢曲霉、泡盛曲霉、烟曲霉、臭曲霉、日本曲霉、构巢曲霉、黑曲霉、或米曲霉；镰孢属，优选杆孢状镰孢、谷孢镰孢、Fusarium crookwellense、黄色镰孢、禾谷镰孢、禾赤镰孢、异孢镰孢、合欢木镰孢、尖孢镰孢、网状镰

孢、粉红镰孢、接骨木镰孢、肤色镰孢、硫色镰孢、近念珠状镰孢、单端孢镰孢、或有毒镰孢；腐质霉属，优选 *Humicola insolens* 或 *Humicola lanuginosa*；或木霉属，优选哈茨木霉、康宁木霉、长枝木霉、瑞氏木霉、或绿色木霉。

[0206] 可根据本领域已知的方法来制备多肽组合物，并且它可以是液体或干燥组合物。例如，多肽组合物可以是颗粒或微粒的形式。可以本领域已知的方法稳定包含在组合物中的多肽。

[0207] 下文给出了本发明多肽组合物的优选用途的实施例。本发明多肽组合物的剂量和使用组合物的其它条件可根据本领域已知的方法来确定。

[0208] 生物材料降解或转化为单糖、二糖、和多糖

[0209] 本发明还涉及降解或转化纤维素材料的方法，包括：在存在有效量的具有增强分解纤维素活性的多肽时，用有效量的分解纤维素的蛋白质处理纤维素材料，其中所述具有增强分解纤维素活性的多肽的存在与不存在所述具有增强分解纤维素活性的多肽时相比增加了纤维素材料的降解。

[0210] 本发明的多肽和宿主细胞可用于从生物质生产单糖、二糖、和多糖，作为化学或发酵原料，用于生产乙醇、塑料、或其它产品或中间体。具体而言，通过纤维素或半纤维素的部分或完全溶解，本发明的多肽和宿主细胞可用于提高加工残余物（干酒糟、酿酒的酒糟、甘蔗渣等等）的价值。如下所述，在分解纤维素的蛋白质促进纤维素材料加工成葡萄糖、木糖、甘露糖、半乳糖、和阿拉伯糖的过程中，也产生了它们的聚合物或由其衍生的产物。该多肽可以是含有或不含细胞的粗制发酵液的形式，或者是半纯化或者纯化的酶制备物的形式。该增强分解纤维素的蛋白质可以是单组分制备物，例如家族 61 蛋白质，多组分蛋白质制备物，例如许多家族 61 蛋白质，或者多组分和单组分蛋白质制备物的组合。增强分解纤维素的蛋白质可在酸性、中性、或碱性 pH 范围内提高分解纤维素的蛋白质的活性。或者，本发明的宿主细胞可在生物材料的发酵过程中用作该多肽的来源。宿主细胞还可以含有天然的或异源的编码分解纤维素的蛋白质以及对加工生物材料有用的其它酶的基因。

[0211] 生物材料可包括但不限于：木材资源、城市固体废物、废纸、农作物、和农作物残余物（参见例如 Wiseloge 等人，1995，在《Handbook on Bioethanol》中，Charles E. Wyman 编，pp. 105-118，Taylor&Francis，Washington D. C.；Wyman，1994，Bioresource Technology 50：3-16；Lynd，1990，Applied Biochemistry and Biotechnology 24/25：695-719；Mosier 等人，1999，Recent Progress in Bioconversion of Lignocellulosics，在《Advances in Biochemical Engineering/Biotechnology》中，T. Scheper 主编，Vol. 65，pp. 23-40，Springer-Verlag，New York）。

[0212] 生物质原生（primary）细胞壁中的主要多糖是纤维素，第二丰富的是半纤维素，而第三是果胶。细胞停止生长后产生的次生（secondary）细胞壁也含有多糖，而且它通过与半纤维素共价交联的聚合木质素得到强化。纤维素是脱水纤维二糖的均聚物，因此为线状的 β -(1-4)-D-葡聚糖，而半纤维素包含多种化合物，诸如具有多种取代基的复杂分支结构的木聚糖、木糖葡聚糖、阿糖木聚糖、和甘露聚糖。虽然纤维素通常是多形的，但在植物组织中发现的纤维素主要是以平行葡聚糖链的不溶性晶体基质的形式存在。半纤维素通常与纤维素以及其它半纤维素以氢键结合，这有助于稳定细胞壁基质。

[0213] 在本发明的方法中，分解纤维素的蛋白质可以是参与将纤维素材料加工成葡萄糖

或将半纤维素加工成木糖、甘露糖、半乳糖和阿拉伯糖、它们的聚合物、或由其衍生的产物的任何蛋白质,正如下文所述。分解纤维素的蛋白质可以是单组分制备物,例如纤维素酶,多组分制备物,例如下文定义的内切葡聚糖酶、纤维二糖水解酶、葡糖水解酶、 β -葡萄糖苷酶,或多组分和单组分蛋白质制备物的组合。分解纤维素的蛋白质可在酸性、中性或碱性 pH 范围内具有活性,即水解纤维素。

[0214] 分解纤维素的蛋白质可以是真菌或细菌来源,可以从已知能够产生纤维素分解酶的微生物中获得或分离和纯化,例如芽孢杆菌属、假单胞菌属、腐质霉属、鬼伞属、梭孢壳属、镰孢属、毁丝霉属、*Acremonium* 属、头孢属 (*Cephalosporium*)、柱顶孢属 (*Scytalidium*)、青霉属或曲霉属的物种(参见例如 EP458162),尤其是那些由选自如下物种的菌株产生的那些:*Humicola insolens* (重新分类为嗜热柱顶孢 (*Scytalidium thermophilum*),参见例如美国专利号 4,435,307)、灰盖鬼伞、尖孢镰孢、嗜热毁丝霉、*Meripilus giganteus*、土生梭孢壳 (*Thielavia terrestris*)、*Acremonium* sp.、*Acremonium persicinum*、*Acremonium acremonium*、*Acremonium brachyphenium*、*Acremonium dichromosporum*、*Acremonium obclavatum*、*Acremonium pinkertoniae*、*Acremonium roseogriseum*、*Acremonium incoloratum* 和 *Acremonium furatum*;优选来自如下物种:*Humicola insolens* DSM1800、尖孢镰孢 DSM2672、嗜热毁丝霉 CBS117.65、头孢 RYM-202、*Acremonium* sp. CBS478.94、*Acremonium* sp. CBS265.95、*Acremonium persicinum* CBS169.65、*Acremonium acremonium* AHU9519、头孢 CBS535.71、*Acremonium brachyphenium* CBS866.73、*Acremonium dichromosporum* CBS683.73、*Acremonium obclavatum* CBS311.74、*Acremonium pinkertoniae* CBS157.70、*Acremonium roseogriseum* CBS134.56、*Acremonium incoloratum* CBS146.62 和 *Acremonium furatum* CBS299.70H。分解纤维素的蛋白质也可由木霉属(特别是绿色木霉、瑞氏木霉和康宁木霉)、嗜碱性芽孢杆菌(参见例如美国专利号 3,844,890 和 EP458162)和链霉菌属(参见例如 EP458162)获得。还包括化学修饰或蛋白质工程突变体。

[0215] 尤其合适的分解纤维素的蛋白质是碱性或中性纤维素酶。这样的纤维素酶的实例有 EP495,257、EP531,372、W096/11262、W096/29397、W098/08940 中描述的纤维素酶。其它实例有诸如 W094/07998、EP531,315、美国专利号 4,435,307、美国专利号 5,457,046、美国专利号 5,648,263、美国专利号 5,686,593、美国专利号 5,691,178、美国专利号 5,763,254、美国专利号 5,776,757、W089/09259、W095/24471、W098/12307 和 PCT/DK98/00299 中描述的那些纤维素酶变体。

[0216] 本发明方法中使用的分解纤维素的蛋白质和增强分解纤维素的蛋白质可通过使用本领域已知的程序将上文所述微生物菌株在含有合适的碳源和氮源和无机盐的营养培养基中进行发酵来生产(参见例如 Bennett, J. W. 和 LaSure, L. 编,《More Gene Manipulations in Fungi》,Academic Press,CA,1991)。合适的培养基可从供应商处获得,或者可根据发表的组成(例如在美国典型培养物保藏中心的目录中)来制备。适于生长和生产分解纤维素的蛋白质的温度范围和其它条件是本领域已知的(参见例如 Bailey, J. E. 和 Ollis, D. F.,《Biochemical Engineering Fundamentals》,McGraw-Hill Book Company, NY, 1986)。

[0217] 发酵可以是培养细胞而导致分解纤维素的蛋白质或增强分解纤维素的蛋白质表

达或分离的任何方法。因此,可将发酵理解为包括摇瓶培养、在实验室或工业发酵罐中在合适的培养基中和在允许分解纤维素的蛋白质或增强分解纤维素的蛋白质表达或分离的条件下进行的小规模或大规模发酵(包括连续、分批、补料-分批或固态发酵)。

[0218] 通过上述方法产生的所得分解纤维素的蛋白质或增强分解纤维素的蛋白质可从发酵培养基中通过常规程序回收,包括但不限于离心、过滤、喷雾-干燥、蒸发或沉淀。然后可通过多种层析,例如离子交换层析、凝胶过滤层析、亲和层析等将回收的蛋白质进一步纯化。

[0219] 分解纤维素的蛋白质可水解或水解羧甲基纤维素(CMC),由此降低保温混合物的粘度。所得粘度降低可通过振动粘度计(vibration viscosimeter)(例如来自法国 Sofraser 的 MIVI3000)来测定。依照纤维素酶粘度单位(CEVU)测量的纤维素酶活性的测定通过测量样品降低羧甲基纤维素(CMC)溶液粘度的能力来定量样品中存在的催化活性的量。该测定法以适于分解纤维素的蛋白质和底物的温度和 pH 进行。对于 Celluzyme™ (Novozymes A/S, Bagsværd, 丹麦),测定法于 40°C 在 0.1M 磷酸盐 pH9.0 缓冲液中以 CMC 作为底物(33.3g/L 羧甲基纤维素 Hercules7LFD)和大约 3.3-4.2CEVU/ml 的酶浓度进行 30 分钟。相对于公告的酶标准物,诸如 Celluzyme™ Standard17-1194(来自 Novozymes A/S),计算 CEVU 活性。

[0220] 适用于本发明的分解纤维素的制备物的实例包括例如 CELLUCLAST™ (可从 Novozymes A/S 获得)和 NOVOZYM™188 (可从 Novozymes A/S 获得)。包含可使用的纤维素酶的其他商品化制备物包括 CELLUZYME™、CEREFLO™ 和 ULTRAFLO™ (Novozymes A/S)、LAMINEX™ 和 SPEZYME™CP (Genencor Int.)及 ROHAMENT™7069W (Röhm GmbH)。纤维素酶以固体重量的大约 0.001% 至大约 5.0%、更优选固体重量的大约 0.025% 至大约 4.0%、和最优选固体重量的大约 0.005% 至大约 2.0% 的有效量加入。

[0221] 如上所述,用于本发明方法的分解纤维素的蛋白质或增强分解纤维素的蛋白质可以是单组分制备物,即本质上不含其它分解纤维素组分的组分。该单组分可以是重组组分,即通过克隆编码该单组分的 DNA 序列随后用该 DNA 序列转化细胞并在宿主中表达而产生的(参见例如 W091/17243 和 W091/17244)。单组分分解纤维素的蛋白质的其它实例包括但不限于 JP-07203960-A 和 W0-9206209 中公开的那些。所述宿主优选是异源宿主(酶对于宿主来说是外源的),但在某些条件下所述宿主也可以是同源宿主(酶对于宿主来说是天然的)。单组分分解纤维素的蛋白质还可以通过从发酵培养基中纯化所述蛋白来制备。

[0222] 可用于实施本发明方法的单组分分解纤维素的蛋白质的实例包括但不限于内切葡聚糖酶、纤维二糖水解酶、葡糖水解酶和 β -葡萄糖苷酶。

[0223] 术语“内切葡聚糖酶”在本文中定义为内切-1,4-(1,3;1,4)- β -D-葡聚糖 4-葡聚糖水解酶(E.C.No.3.2.1.4),它催化纤维素、纤维素衍生物(诸如羧甲基纤维素和羟乙基纤维素)和地衣淀粉中的 1,4- β -D-糖苷键,混合的 β -1,3 葡聚糖诸如谷类 β -D-葡聚糖或木糖葡聚糖,和含有纤维素组分的其它植物原料中的 β -1,4 键的内水解(endohydrolysis)。对本发明来说,内切葡聚糖酶活性根据 Ghose,1987,Pure and Appl.Chem.59:257-268 的程序,使用羧甲基纤维素(CMC)水解来测定。

[0224] 外切-1,4- β -D-葡聚糖酶包括纤维二糖水解酶和葡糖水解酶。

[0225] 术语“纤维二糖水解酶”在本文中定义为 1,4- β -D-葡聚糖纤维二糖水解酶

(E. C. 3. 2. 1. 91), 它催化纤维素、纤维寡糖 (cellooligosaccharides)、或含有 β -1, 4- 连接葡萄糖的任何聚合物中的 1, 4- β -D- 糖苷键的水解, 从链的还原性或非还原性末端释放纤维二糖。对本发明来说, 纤维二糖水解酶活性根据 Lever 等人, 1972, *Anal. Biochem.* 47 : 273-279 和 van Tilbeurgh 等人, 1982, *FEBS Letters* 149 : 152-156 ; van Tilbeurgh 和 Claeysens, 1985, *FEBS Letters* 187 : 283-288 描述的程序来测定。在本发明中, 采用 Lever 等人的方法来评估玉米秸秆中的纤维素水解, 而将 van Tilbeurgh 等人的方法用于测定对荧光二糖衍生物的纤维二糖水解酶活性。

[0226] 术语“葡糖水解酶”在本文中定义为 1, 4- β -D- 葡聚糖葡糖水解酶 (E. C. 3. 2. 1. 74), 它催化 1, 4- β -D- 葡聚糖中 1, 4- 键 (O- 糖基键) 的水解, 以除去连续的葡萄糖单元。对本发明来说, 外切葡聚糖酶活性根据 Himmel 等人, 1986, *J. Biol. Chem.* 261 : 12948-12955 描述的方法来测定。

[0227] 术语“ β - 葡萄糖苷酶”在本文中定义为 β -D- 葡萄糖苷葡糖水解酶 (E. C. 3. 2. 1. 21), 它催化末端非还原性 β -D- 葡萄糖残基的水解并释放 β -D- 葡萄糖。对本发明来说, β - 葡萄糖苷酶活性根据 Venturi 等人, 2002, *J. Basic Microbiol.* 42 : 55-66 描述的基本程序来测定, 只是如本文所述采用不同的条件。一个单位的 β - 葡萄糖苷酶活性定义为于 50°C、pH5 在 100mM 柠檬酸钠、0.01% Tween-20 中从作为底物的 4mM 对硝基苯基- β -D- 吡喃型葡萄糖苷每分钟产生 1.0 μ mole 对硝基苯酚。

[0228] 用本发明的多肽结合分解纤维素的蛋白质来降解生物质底物的纤维素组分 (参见例如 Brigham 等人, 1995, 在《Handbook on Bioethanol》中, Charles E. Wyman 编, pp. 119-141, Taylor&Francis, Washington D. C. ; Lee, 1997, *Journal of Biotechnology* 56 : 1-24)。

[0229] 具有增强分解纤维素活性的多肽和分解纤维素的蛋白质的最佳用量取决于几个因素, 包括但不限于分解纤维素的蛋白质组分的混合物、纤维素底物、纤维素底物的浓度、纤维素底物的预处理、温度、时间、pH、和发酵生物 (例如同时进行糖化和发酵的酵母) 的引入。术语“分解纤维素的蛋白质” (cellulolytic proteins) 在本文中定义为在测试条件下证明能够水解或转化或降解纤维素的蛋白质或蛋白质混合物。其用量通常是通过普通的测定法诸如 BCA (双辛可宁酸 (bicinchoninic acid), P. K. Smith 等人, 1985, *Anal. Biochem.* 150 : 76) 来测量的, 而且优选的添加量与水解的生物质的量成比例。

[0230] 在一个优选的方面, 每 g 纤维素材料中具有增强分解纤维素活性的多肽的量是每 g 纤维素材料大约 0.01 至大约 2.0mg、优选大约 0.025 至大约 1.5mg、更优选大约 0.05 至大约 1.25mg、更优选大约 0.075 至大约 1.25mg、更优选大约 0.1 至大约 1.25mg、甚至更优选大约 0.15 至大约 1.25mg、且最优选大约 0.25 至大约 1.0mg。

[0231] 在另一个优选的方面, 每 g 纤维素材料中分解纤维素的蛋白质的量是每 g 纤维素材料大约 0.5 至大约 50mg、优选大约 0.5 至大约 40mg、更优选大约 0.5 至大约 25mg、更优选大约 0.75 至大约 20mg、更优选大约 0.75 至大约 15mg、甚至更优选大约 0.5 至大约 10mg、且最优选大约 2.5 至大约 10mg。

[0232] 在一个优选的方面, 每 g 分解纤维素的蛋白质中具有增强分解纤维素活性的多肽的量是每 g 分解纤维素的蛋白质大约 0.005 至大约 1.0g、优选大约 0.01 至大约 1.0g、更优选大约 0.15 至大约 0.75g、更优选大约 0.15 至大约 0.5g、更优选大约 0.1 至大约 0.5g、甚

至更优选大约 0.1 至大约 0.5g、且最优选大约 0.05 至大约 0.2g。

[0233] 本发明的方法可用于将纤维素材料加工成许多有用的有机产品、化学品和燃料。除了乙醇之外,能够由纤维素生产的一些物品和专用化学品包括木糖、丙酮、乙酸盐或酯 (acetate)、甘氨酸、赖氨酸、有机酸(例如乳酸)、1,3-丙二醇、丁二醇、甘油、乙二醇、糠醛、聚羟基链烷酸盐或酯 (polyhydroxyalkanoates)、顺,顺-粘康酸 (cis,cis-muconic acid)、和动物饲料(Lynd, L. R.、Wyman, C. E. 和 Gerngross, T. U.,1999, Biocommodity Engineering, Biotechnol. Prog. 15 :777-793 ;Philippidis, G. P.,1996, Cellulose bioconversion technology, 在《Handbook on Bioethanol:Production and Utilization》中, Wyman, C. E. 编, Taylor&Francis, Washington, DC,179-212 ;及 Ryu, D. D. Y. 和 Mandels, M.,1980, Cellulases: biosynthesis and applications, Enz. Microb. Technol. 2 :91-102)。可能的共生产效益不仅仅是由可发酵的碳水化合物合成多种有机产品。可将生物学加工之后剩余的富含木质素的残余物转化为源自木质素的化学品,或用于发电。

[0234] 根据本发明的方法用于加工纤维素材料的常规方法对于本领域技术人员而言是完全理解的。可使用经过配置以依照本发明运作的任何常规生物材料加工设备来实施本发明的方法。

[0235] 这样的设备可包括分批搅拌反应器 (batch-stirred reactor)、带超滤的连续流搅拌反应器 (continuous flow stirred reactor with ultrafiltration)、连续活塞流柱反应器 (continuous plug-flow column reactor)(Gusakov, A. V. 和 Sinit syn, A. P.,1985, Kinetics of the enzymatic hydrolysis of cellulose:1.A mathematical model for a batch reactor process, Enz. Microb. Technol. 7 :346-352)、磨碎反应器 (attrition reactor) (Ryu, S. K. 和 Lee, J. M.,1983, Bioconversion of waste cellulose by using an attrition bioreactor, Biotechnol. Bioeng. 25 :53-65) 或具有由电磁场诱导的强搅拌的反应器(Gusakov, A. V.、Sinit syn, A. P.、Davyd kin, I. Y.、Davyd kin, V. Y.、Protas, O. V.,1996, Enhancement of enzymatic cellulose hydrolysis using a novel type of bioreactor with intensive stirring induced by electromagnetic field, Appl. Biochem. Biotechnol. 56 :141-153)。

[0236] 所述常规方法包括但不限于糖化、发酵、单独水解和发酵(SHF)、同时糖化和发酵(SSF)、同时糖化和共发酵(SSCF)、杂合水解和发酵(hybrid hydrolysis and fermentation, HHF)、及直接微生物转化(DMC)。

[0237] SHF 使用单独的加工步骤,首先用酶将纤维素水解成葡萄糖,然后将葡萄糖发酵成乙醇。在 SSF 中,将纤维素的酶促水解和葡萄糖变成乙醇的发酵组合在一个步骤中(Philippidis, G. P.,1996, Cellulose bioconversion technology, 在《Handbook on Bioethanol:Production and Utilization》中, Wyman, C. E. 编, Taylor&Francis, Washington, DC,179-212)。SSCF 包括多种糖的共发酵(Sheehan, J. 和 Himmel, M.,1999, Enzymes, energy and the environment:Astrategic perspective on the U. S. Department of Energy's research and development activities for bioethanol, Biotechnol. Prog. 15 :817-827)。HHF 包括在同一反应器中但在不同温度进行的两个单独的步骤,即高温酶促糖化,后续在发酵菌株能耐受的较低温度进行的 SSF。DMC 将所有三种过程(纤维

素酶生成、纤维素水解和发酵)组合在一个步骤中(Lynd, L. R.、Weimer, P. J.、van Zyl, W. H. 和 Pretorius, I. S., 2002, *Microbial cellulose utilization: Fundamentals and biotechnology*, *Microbiol. Mol. Biol. Reviews* 66 :506-577)。

[0238] “发酵”或“发酵方法”指任何发酵方法或包含发酵步骤的任何方法。发酵方法包括但不限于用于生产发酵产物的发酵方法,所述发酵产物包括醇类(例如阿拉伯糖醇、丁醇、乙醇、甘油、甲醇、1,3-丙二醇、山梨糖醇和木糖醇);有机酸(例如乙酸、醋酮酸、己二酸、抗坏血酸、柠檬酸、2,5-二酮-D-葡萄糖酸、甲酸、富马酸、葡萄糖二酸、葡萄糖酸、葡萄糖醛酸、戊二酸、3-羟基丙酸、衣康酸、乳酸、苹果酸、丙二酸、草酸、丙酸、琥珀酸和木糖酸);酮类(例如丙酮);氨基酸(例如天冬氨酸、谷氨酸、甘氨酸、赖氨酸、丝氨酸和苏氨酸);气体(例如甲烷、氢(H₂)、二氧化碳(CO₂)和一氧化碳(CO))。发酵方法还包括用于食用醇工业(例如啤酒和葡萄酒)、乳品加工业(例如发酵的乳制品)、皮革工业和烟草工业的发酵方法。

[0239] 本发明还涉及生产有机物质的方法,包括:(a)在存在有效量的具有增强分解纤维素活性的多肽时,用有效量的分解纤维素的蛋白质将纤维素材料糖化,其中所述具有增强分解纤维素活性的多肽的存在与不存在所述具有增强分解纤维素活性的多肽相比增加了纤维素材料的降解;(b)用一种或多种发酵微生物将步骤(a)糖化的纤维素材料发酵;并(c)从发酵物中回收有机物质。具有增强分解纤维素活性的多肽可以是含有或没有细胞的粗制发酵液的形式,或是半纯化或纯化的酶制备物的形式。增强分解纤维素的蛋白质可以是单组分制备物,例如一种家族61蛋白质,多组分蛋白质制备物,例如多种家族61蛋白质,或者多组分和单组分蛋白质制备物的组合。

[0240] 所述有机物质可以是发酵得到的任何物质。在一个优选的方面,所述有机物质是醇。可理解的是术语“醇”涵盖含有一个或多个羟基部分的有机物。在一个更优选的方面,所述醇是阿拉伯糖醇。在另一个更优选的方面,所述醇是丁醇。在另一个更优选的方面,所述醇是乙醇。在另一个更优选的方面,所述醇是甘油。在另一个更优选的方面,所述醇是甲醇。在另一个更优选的方面,所述醇是1,3-丙二醇。在另一个更优选的方面,所述醇是山梨糖醇。在另一个更优选的方面,所述醇是木糖醇。参见例如 Gong, C. S.、Cao, N. J.、Du, J. 和 Tsao, G. T., 1999, *Ethanol production from renewable resources*, 在《*Advances in Biochemical Engineering/Biotechnology*》中, Scheper, T. 编, Springer-Verlag Berlin Heidelberg, Germany, 65 :207-241; Silveira, M. M., 和 Jonas, R., 2002, *The biotechnological production of sorbitol*, *Appl. Microbiol. Biotechnol.* 59 :400-408; Nigam, P., 和 Singh, D., 1995, *Processes for fermentative production of xylitol-a sugar substitute*, *Process Biochemistry* 30 (2) :117-124; Ezeji, T. C.、Qureshi, N. 和 Blaschek, H. P., 2003, *Production of acetone, butanol and ethanol by Clostridium beijerinckii BA101 and in situ recovery by gas stripping*, *World Journal of Microbiology and Biotechnology* 19 (6) :595-603。

[0241] 在另一个优选的方面,有机物质是有机酸。在另一个更优选的方面,所述有机酸是乙酸。在另一个更优选的方面,所述有机酸是丙酮酸。在另一个更优选的方面,所述有机酸是己二酸。在另一个更优选的方面,所述有机酸是抗坏血酸。在另一个更优选的方面,所述有机酸是柠檬酸。在另一个更优选的方面,所述有机酸是2,5-二酮-D-葡萄糖酸。在另一个更优选的方面,所述有机酸是甲酸。在另一个更优选的方面,所述有机酸是富马酸。在另

一个更优选的方面,所述有机酸是葡糖二酸。在另一个更优选的方面,所述有机酸是葡糖酸。在另一个更优选的方面,所述有机酸是葡糖醛酸。在另一个更优选的方面,所述有机酸是戊二酸。在另一个更优选的方面,所述有机酸是3-羟基丙酸。在另一个更优选的方面,所述有机酸是衣康酸。在另一个更优选的方面,所述有机酸是乳酸。在另一个更优选的方面,所述有机酸是苹果酸。在另一个更优选的方面,所述有机酸是丙二酸。在另一个更优选的方面,所述有机酸是草酸。在另一个更优选的方面,所述有机酸是丙酸。在另一个更优选的方面,所述有机酸是琥珀酸。在另一个更优选的方面,所述有机酸是木糖酸。参见例如 Chen, R., 和 Lee, Y. Y., 1997, Membrane-mediated extractive fermentation for lactic acid production from cellulosic biomass, *Appl. Biochem. Biotechnol.* 63-65 :435-448。

[0242] 在另一个更优选的方面,有机物质是酮。可以理解的是术语“酮”涵盖含有一个或多个酮部分的有机物。在另一个更优选的方面,所述酮是丙酮。参见例如 Qureshi 和 Blaschek, 2003, 同上。

[0243] 在另一个更优选的方面,有机物质是氨基酸。在另一个更优选的方面,所述有机酸是天冬氨酸。在另一个更优选的方面,所述氨基酸是谷氨酸。在另一个更优选的方面,所述氨基酸是甘氨酸。在另一个更优选的方面,所述氨基酸是赖氨酸。在另一个更优选的方面,所述氨基酸是丝氨酸。在另一个更优选的方面,所述氨基酸是苏氨酸。参见例如 Richard, A., 和 Margaritis, A., 2004, Empirical modeling of batch fermentation kinetics for poly(glutamic acid) production and other microbial biopolymers, *Biotechnology and Bioengineering* 87(4) :501-515。

[0244] 在另一个更优选的方面,有机物质是气体。在另一个更优选的方面,所述气体是甲烷。在另一个更优选的方面,所述气体是 H₂。在另一个更优选的方面,所述气体是 CO₂。在另一个更优选的方面,所述气体是 CO。参见例如 Kataoka, N., A. Miya, 和 K. Kiriya, 1997, Studies on hydrogen production by continuous culture system of hydrogen-producing anaerobic bacteria, *Water Science and Technology* 36(6-7) : 41-47; 及 Gunaseelan V.N., *Biomass and Bioenergy* Vol. 13(1-2), pp. 83-114, 1997, Anaerobic digestion of biomass for methane production: A review。

[0245] 由纤维素材料生产有机物质通常需要四个主要步骤。这四个步骤是预处理、酶促水解、发酵和回收。下面列举的是生产乙醇的方法,但应理解的是类似的方法可用于生产其它有机物质,例如上文所述物质。

[0246] 预处理。在预处理或预水解步骤中,将纤维素材料加热以破坏木质素和碳水化合物结构,使大部分半纤维素溶解,并使纤维素级分可被纤维素分解酶接近。直接用蒸气或者在浆液中进行加热,在所述浆液中也可将催化剂加入原料以加速反应。催化剂包括强酸,诸如硫酸和 SO₂, 或碱,诸如氢氧化钠。预处理阶段的目的在于促进酶和微生物的透过 (penetration)。也可将纤维素生物材料进行湿热蒸汽喷发 (hydrothermal steam explosion) 预处理 (参见美国专利申请号 20020164730)。

[0247] 糖化。在酶促水解 (也称为糖化) 步骤中,将本文所述酶加入预处理的原料以将纤维素级分转化为葡萄糖和 / 或其它糖。糖化通常在 pH、温度和混合条件受控的搅拌槽反应器 (stirred-tank reactor) 和发酵罐中进行。糖化步骤可持续长达 200 个小时。糖化可于大约 30°C 至大约 65°C、特别是大约 50°C 的温度,和大约 4 至大约 5, 尤其是大约 pH 4.5 的

pH 进行。为了生产能被酵母代谢的葡萄糖,通过在存在 β -葡萄糖苷酶时进行水解。

[0248] 发酵。在发酵步骤中,将作为预处理和酶促水解步骤的结果由纤维素材料释放的糖通过发酵生物诸如酵母发酵成乙醇。发酵也可与酶促水解在同样 pH、温度和混合条件受控的同一容器中同时进行。当糖化和发酵在同一容器中同时进行时,该过程通常称为同时糖化和发酵或 SSF。

[0249] 任何合适的纤维素底物或原材料可用于本发明的发酵过程。通常根据预期的发酵产品,即要由发酵得到的有机物质,和采用的方法来选择底物,正如本领域众所周知的。适用于本发明方法的底物的实例包括含有纤维素的材料,诸如木材或植物残余物或由加工的纤维素材料得到的低分子糖 DP1-3,它们可由发酵微生物代谢,且可通过直接添加到发酵培养基中来提供。

[0250] 术语“发酵培养基”将理解为指发酵微生物加入之前的培养基,诸如由糖化过程得到培养基,以及用于同时糖化和发酵过程(SSF)的培养基。

[0251] “发酵微生物”指适用于期望发酵过程的任何微生物。根据本发明的合适的发酵微生物能够发酵,即直接或间接将糖,诸如葡萄糖、木糖、阿拉伯糖、甘露糖、半乳糖或寡糖转化为所需发酵产物。发酵微生物的实例包括真菌生物体,诸如酵母。优选的酵母包括糖酵母菌种的菌株,特别是酿酒酵母。商品化的酵母包括例如 Red Star®/™/Lesaffre Ethanol Red (可从美国的 Red Star/Lesaffre 获得)、FALI (可从美国的 Burns Philp Food Inc. 的 Fleischmann's Yeast 部门获得)、SUPERSTART (可从 Alltech 获得)、GERT STRAND (可从瑞典的 Gert Strand AB 获得)和 FERMIOL (可从 DSM Specialties 获得)。

[0252] 在一个优选的方面,所述酵母是糖酵母菌种。在一个更优选的方面,酵母是酿酒酵母。在另一个更优选的方面,所述酵母是糖化糖酵母。在另一个更优选的方面,所述酵母是葡萄汁糖酵母 (*Saccharomyces uvarum*)。在另一个优选的方面,所述酵母是克鲁维氏酵母属。在另一个更优选的方面,所述酵母是马克斯克鲁维氏酵母 (*Kluyveromyces marxianus*)。在另一个更优选的方面,所述酵母是脆壁克鲁维氏酵母 (*Kluyveromyces fragilis*)。在另一个优选的方面,所述酵母为假丝酵母属。在另一个更优选的方面,所述酵母是 *Candida pseudotropicalis*。在另一个更优选的方面,所述酵母是 *Candida brassicae*。在另一个优选的方面,所述酵母是 *Clavispora* 属。在另一个更优选的方面,所述酵母是 *Clavispora lusitaniae*。在另一个更优选的方面,所述酵母是 *Clavispora opuntiae*。在另一个优选的方面,所述酵母是管囊酵母属 (*Pachysolen*)。在另一个更优选的方面,所述酵母是管囊酵母 (*Pachysolen tannophilus*)。在另一个优选的方面,所述酵母是酒香酵母属 (*Brettanomyces*)。在另一个更优选的方面,所述酵母是克劳森酒香酵母 (*Brettanomyces clausenii*) (Philippidis, G. P., 1996, Cellulose bioconversion technology, 在《Handbook on Bioethanol: Production and Utilization》中, Wyman, C. E. 编, Taylor&Francis, Washington, DC, 179-212)。

[0253] 能有效将葡萄糖发酵为乙醇的细菌包括例如运动发酵单孢菌 (*Zymomonas mobilis*) 和热纤维梭菌 (*Clostridium thermocellum*) (Philippidis, 1996, 同上)。

[0254] 本领域众所周知上述生物体还可用于生产其它有机物质,正如本文所述。

[0255] 在酿酒酵母中 (Chen, Z., Ho, N. W. Y., 1993, Cloning and improving the expression of *Pichia stipitis* xylose reductase gene in *Saccharomyces*

cerevisiae, Appl. Biochem. Biotechnol. 39-40 :135-147 ;Ho, N. W. Y. 、Chen, Z. 、Brainard, A. P. ,1998, Genetically engineered Saccharomyces yeast capable of effectively cofermenting glucose and xylose, Appl. Environ. Microbiol. 64 :1852-1859), 或在细菌诸如大肠杆菌(Beall, D. S. 、Ohta, K. 、Ingram, L. O. ,1991, Parametric studies of ethanol production from xylose and other sugars by recombinant Escherichia coli, Biotech. Bioeng. 38 :296-303)、产酸克雷伯氏菌(Klebsiella oxytoca) (Ingram, L. O. 、Gomes, P. F. 、Lai, X. 、Moniruzzaman, M. 、Wood, B. E. 、Yomano, L. P. 、York, S. W. ,1998, Metabolic engineering of bacteria for ethanol production, Biotechnol. Bioeng. 58 :204-214) 和运动发酵单孢菌(Zhang, M. 、Eddy, C. 、Deanda, K. 、Finkelstein, M. 和 Picataggio, S. ,1995, Metabolic engineering of a pentose metabolism pathway in ethanologenic Zymomonas mobilis, Science 267 :240-243 ;Deanda, K. 、Zhang, M. 、Eddy, C. 和 Picataggio, S. ,1996, Development of an arabinose-fermenting Zymomonas mobilis strain by metabolic pathway engineering, Appl. Environ. Microbiol. 62 :4465-4470) 中克隆异源基因导致能将己糖和戊糖转化为乙醇(共发酵)的生物体的构建。

[0256] 通常将酵母或另一种微生物加入降解的纤维素或水解产物,而发酵进行大约 24 至大约 96 小时,诸如大约 35 至大约 60 小时。温度通常在大约 26°C 至大约 40°C,特别是在大约 32°C,和在大约 pH3 至大约 pH6,特别是 pH4-5 左右。

[0257] 在一个优选的方面,将酵母或另一种微生物施用于降解的纤维素或水解物,而发酵进行大约 24 至大约 96 小时,诸如通常 35-60 小时。在一个优选的方面,温度通常在大约 26 至大约 40°C 之间,特别是在大约 32°C,而 pH 通常为大约 pH3 至大约 pH6,优选 pH4-5 左右。优选将酵母或另一种微生物以每 ml 发酵培养液(fermentation broth) 大约 10^5 至 10^{12} 、优选大约 10^7 至 10^{10} 、尤其是大约 5×10^7 活菌计数的量施用。在乙醇生产阶段期间,酵母细胞计数应优选在大约 10^7 至 10^{10} 的范围内,尤其是大约 2×10^8 左右。关于使用酵母发酵的进一步指导可见于例如《The Alcohol Textbook》(K. Jacques、T. P. Lyons 和 D. R. Kelsall 编, Nottingham University Press, United Kingdom, 1999), 将其引入作为参考。

[0258] 本领域中最广泛使用的方法是同时糖化和发酵(SSF)方法,其中没有糖化的保持阶段(holding stage),这意味着酵母和酶一起加入。

[0259] 对于乙醇生产而言,发酵后将醪液(mash)蒸馏以提取乙醇。根据本发明的方法得到的乙醇可用作例如燃料乙醇;饮料乙醇,即可饮用酒精(potable neutral spirits);或工业乙醇。

[0260] 发酵刺激物(fermentation stimulator)可与本文描述的任何酶促过程组合使用以进一步改进发酵方法,特别是发酵微生物的性能,诸如速率增加和乙醇产量。“发酵刺激物”指用于发酵微生物特别是酵母的生长的刺激物。优选的用于生长的发酵刺激物包括维生素和矿物质。维生素的实例包括多种维生素(multivitamins)、生物素、泛酸、烟酸、内消旋肌醇(meso-inositol)、硫胺素、吡哆醇、对氨基苯甲酸、叶酸、核黄素、及维生素 A、B、C、D 和 E。参见例如 Alfenore 等人, Improving ethanol production and viability of Saccharomyces cerevisiae by a vitamin feeding strategy during fed-batch process, Springer-Verlag, 2002, 将其引入本文作为参考。矿物质的实例包括能够提供营

养的矿物质和矿物质盐,包括 P、K、Mg、S、Ca、Fe、Zn、Mn 和 Cu。

[0261] **回收**。从发酵的纤维素材料分离乙醇,并通过蒸馏的常规方法纯化。可得到纯度为高达大约 96vol. % 乙醇的乙醇,它可用作例如燃料乙醇、饮料乙醇即可饮用酒精、或工业乙醇。

[0262] 对于其它有机物质来说,可使用本领域已知的任何方法,包括但不限于层析(例如离子交换、亲和、疏水、层析聚焦和大小排阻)、电泳(例如制备性等电聚焦)、溶解度差异(例如硫酸铵沉淀)、SDS-PAGE、蒸馏或萃取。

[0263] 在本发明的方法中,可用一种或多种附加的酶活性补充分解纤维素酶的蛋白质和增强分解纤维素的多肽以改进纤维素材料的降解。优选的附加酶是半纤维素酶、酯酶(例如脂肪酶、磷脂酶和 / 或角质酶)、蛋白酶、漆酶、过氧化物酶或其混合物。

[0264] 在本发明的方法中,可将附加的酶在发酵前或发酵过程中加入,包括发酵微生物繁殖的过程中或之后。

[0265] 本文所指的酶可源自或得自任何合适的来源,包括细菌、真菌、酵母或哺乳动物来源。术语“得到”在本文中指酶可分离自天然产生该酶作为固有酶(native enzyme)的生物体。术语“得到”在本文中还可指酶可在宿主生物体中重组产生,其中所述重组产生的酶对于宿主生物体来说是天然的或外源的或具有经过修饰的氨基酸序列,例如删除、插入和 / 或替代一个或多个氨基酸,即重组产生的酶,它是天然氨基酸序列的突变体和 / 或片段,或是通过本领域已知的核酸改组(shuffling)方法产生的酶。涵盖在天然酶的意义中的有天然变体,而涵盖在外源酶的意义中的有诸如通过定点诱变或改组而重组得到的变体。

[0266] 酶还可以是纯化的。术语“纯化的”在用于本文时涵盖不含来自该酶的来源生物体的其它组分的酶。术语“纯化的”还涵盖不含来自得到所述酶的天然生物体的组分的酶。酶可以是纯化的,只有少量的其它蛋白质存在。表述“其它蛋白质”具体涉及其它酶。术语“纯化的”在用于本文时还指去除其它组分,具体是其它蛋白质,且最具体的是存在于本发明的酶的来源细胞中的其它酶。酶可以是“基本上纯的”,即不含来自产生所述酶的生物体的其它组分,即例如用于重组产生酶的宿主生物体。在一个优选的方面,酶是至少 75%(w/w)、优选至少 80%、更优选至少 85%、更优选至少 90%、更优选至少 95%、更优选至少 96%、更优选至少 97%、甚至更优选至少 98%、或最优选至少 99% 纯的。在另一个优选的方面,酶是 100% 纯的。

[0267] 用于本发明的酶可以是适用于本文描述的方法的任何形式,诸如例如含有或不含细胞的粗制发酵液、干粉或颗粒、非粉化颗粒(non-dusting granulate)、液体、稳定液体(stabilized liquid)或受保护的酶的形式。颗粒可例如如美国专利号 4,106,991 和 4,661,452 中所公开的来生产,且可任选通过本领域已知的方法包衣。液体酶制备物可例如根据已确定的方法通过加入稳定剂,诸如糖、糖醇或其它多元醇,和 / 或乳酸或其它有机酸来稳定。受保护的酶可根据 EP238,216 中公开的方法来制备。

[0268] **半纤维素酶**

[0269] 可通过多种真菌和细菌进行半纤维素的酶促水解。与纤维素降解类似,半纤维素水解需要许多酶的协调作用。可将半纤维素酶分成三种基本类型:攻击多糖链内内部键的内作用酶(endo-acting enzyme),从多糖链的还原性或非还原性末端向前起作用的外作用酶(exo-acting enzyme),和附属酶,水解木质素糖苷键的乙酰酯酶(acetylcetase)

和酯酶,诸如香豆酸(coumaric acid)酯酶和阿魏酸酯酶(Wong, K. K. Y.、Tan, L. U. L. 和 Saddler, J. N., 1988, Multiplicity of β -1,4-xylanase in microorganisms: Functions and applications, Microbiol. Rev. 52 :305-317 ;Tenkanen, M. 和 Poutanen, K., 1992, Significance of esterases in the degradation of xylans, 在《Xylans and Xylanases》, Visser, J.、Beldman, G.、Kuster-van Someren, M. A. 和 Voragen, A. G. J. 编, Elsevier, New York, NY, 203-212 ;Coughlan, M. P. 和 Hazlewood, G. P., 1993, 《Hemicellulose and Hemicellulases》, Portland, London, UK ;Brigham, J. S.、Adney, W. S. 和 Himmel, M. E., 1996, Hemicellulases: Diversity and applications, 在《Handbook on Bioethanol: Production and Utilization》, Wyman, C. E. 编, Taylor&Francis, Washington, DC, 119-141)。

[0270] 半纤维素酶包括木聚糖酶、阿拉伯呋喃糖苷酶、乙酰木聚糖酯酶、葡糖醛酸糖苷酶、内切-半乳聚糖酶、甘露聚糖酶、内切或外切阿拉伯聚糖酶(arabinase)、外切-半乳聚糖酶及其混合物。内作用半纤维素酶和附属酶的实例包括内阿拉伯聚糖酶(endoarabinanase)、内切阿拉伯糖半乳聚糖酶、内切葡聚糖酶、内切甘露聚糖酶、内切木聚糖酶和 feraxan 内切木聚糖酶。外作用半纤维素酶和附属酶的实例包括 α -L-阿拉伯糖苷酶、 β -L-阿拉伯糖苷酶、 α -1,2-L-岩藻糖苷酶、 α -D-半乳糖苷酶、 β -D-半乳糖苷酶、 β -D-葡萄糖苷酶、 β -D-葡糖醛酸糖苷酶、 β -D-甘露糖苷酶、 β -D-木糖苷酶、外切葡萄糖苷酶、外切纤维二糖水解脱酶、外切甘露二糖水解脱酶、外切甘露聚糖酶、外切木聚糖酶、木聚糖 α -葡糖醛酸糖苷酶、和松柏苷 β -葡萄糖苷酶。酯酶的实例包括乙酰酯酶(乙酰半乳聚糖酯酶、乙酰甘露聚糖酯酶和乙酰木聚糖酯酶)和芳基酯酶(香豆酸酯酶和阿魏酸酯酶)。

[0271] 优选的是,半纤维素酶是外作用半纤维素酶,而且更优选在低于 pH7 的酸性条件下具有水解半纤维素能力的外作用半纤维素酶。适用于本发明的半纤维素酶的实例包括 VISCOZYME™ (可从 Novozymes A/S 获得)。将半纤维素酶以固体重量的大约 0.001% 至大约 5.0%、更优选固体重量的大约 0.025% 至大约 4.0%、且最优选固体重量的大约 0.005% 至大约 2.0% 的有效量加入。

[0272] 木聚糖酶(E. C. 3. 2. 1. 8)可由任何合适的来源获得,包括真菌和细菌生物体,诸如曲霉属、Disporotrichum、青霉属、脉孢菌属、镰孢属、木霉属、腐质霉属、嗜热霉属(Thermomyces)和芽孢杆菌属。含有木聚糖酶的优选商品化制备物包括 SHEARZYME®、BIOFEED WHEAT®、BIO-FEED Plus®

L、CELLUCLAST®、ULTRAFLO®、VISCOZYME®、PENTOPAN MONO® BG 和 PLUPZYME® HC (Novozymes A/S); 及 LAMINEX® 和 SPEZYME® CP (Genencor Int.)。

[0273] 酯酶

[0274] 可用于纤维素的生物转化的酯酶包括乙酰酯酶,诸如乙酰半乳聚糖酯酶、乙酰甘露聚糖酯酶和乙酰木聚糖酯酶,以及水解木质素糖苷键的酯酶,诸如香豆酸酯酶和阿魏酸酯酶。

[0275] 在用于本文时,“酯酶”也称为羧酸脂水解酶,指作用于酯键上的酶,包括根据酶命名法(《Enzyme Nomenclature》, 1992, Academic Press, San Diego, California, 及分别

在 Eur. J. Biochem. 223 :1-5, 1994 ;Eur. J. Biochem. 232 :1-6, 1995 ;Eur. J. Biochem. 237 :1-5, 1996 ;Eur. J. Biochem. 250 :1-6, 1997 ; 和 Eur. J. Biochem. 264 :610-650, 1999 中的补遗 1(1993)、补遗 2(1994)、补遗 3(1995)、补遗 4(1997) 和补遗 5) 归入 EC3. 1. 1 羧酸脂水解酶的酶。酯酶的非限制性实例包括芳基酯酶、三酰甘油脂肪酶、乙酰酯酶、乙酰胆碱酯酶、胆碱酯酶、托品酯酶 (tropin esterase)、果胶酯酶、甾醇酯酶、叶绿素酶、L-阿拉伯糖酸内酯酶、葡糖酸内酯酶、糖醛酸内酯酶、鞣酸酶、棕榈酸视黄酯酯酶、羟基丁酸酯-二聚物水解酶、酰基甘油脂肪酶、3-氧代己二酸烯醇内酯酶、1, 4-内酯酶、半乳糖脂肪酶、4-吡哆醇内酯酶、酰基肉毒碱水解酶、氨基酰基-tRNA 水解酶、D-阿拉伯糖酸内酯酶、6-磷酸葡糖酸内酯酶、磷脂酶 A1、6-乙酰葡糖脱乙酰基酶、脂蛋白脂肪酶、二氢香豆素脂肪酶、柠檬素-D-环-内酯酶、类固醇内酯酶、三乙酸内酯酶、放线菌素内酯酶、苔色酸缩酚酸水解酶、头孢菌素-C 脱乙酰基酶、氯原酸水解酶、 α -氨基酸酯酶、4-草酰乙酸甲酯酯酶 (4-methyloxaloacetate esterase)、羧甲烯丁烯羟酸内酯酶 (carboxymethylenebutenolidase)、脱氧柠檬酸 A-环-内酯酶、2-乙酰基-1-烷基甘油磷酸胆碱酯酶 (2-acetyl-1-alkylglycerophosphocholine)、镰孢氨酸-C 鸟氨酸酯酶、芥子碱酯酶、蜡酯水解酶、佛波醇二酯水解酶、磷脂酰肌醇脱乙酰基酶、唾液酸 O-乙酰酯酶、乙酰氧基丁基并噻吩脱乙酰基酶 (acetoxabutynylbithiophene deacetylase)、乙酰水杨酸脱乙酰基酶、乙酸甲基伞酮脱乙酰基酶 (methylumbelliferyl-acetate deacetylase)、2-吡喃酮-4, 6-二羧化物内酯酶 (2-pyrone-4, 6-dicarboxylate lactonase)、N-乙酰半乳糖氨基聚糖脱乙酰基酶 (N-acetylgalactosaminoglycan deacetylase)、保幼激素酯酶、双(2-乙基己基)邻苯二甲酸酯酯酶 (bis(2-ethylhexyl)phthalate esterase)、蛋白质-谷氨酸酯甲基酯酶 (protein-glutamate methylesterase)、11-顺式-棕榈酸视黄酯水解酶、所有-反式-棕榈酸视黄酯水解酶、L-鼠李糖-1, 4-内酯酶、5-(3, 4-双乙酰基丁-1-基)-2, 2'-并噻吩脱乙酰基酶 (5-(3, 4-diacetoxybut-1-ynyl)-2, 2'-bithiophene deacetylase)、脂肪-酰基-乙基-酯合酶 (fatty-acyl-ethyl-ester synthase)、木质-1, 4-内酯酶 (xylono-1, 4-lactonase)、N-乙酰基葡糖胺基磷脂酰肌醇脱乙酰基酶、西曲酸酯苯甲基酯酶 (cetraxate benzylesterase)、乙酰基烷基甘油乙酰基水解酶、和乙酰木聚糖酯酶。

[0276] 用于本发明的优选酯酶是脂肪分解酶, 诸如脂肪酶(归入 EC3. 1. 1. 3、EC3. 1. 1. 23 和 / 或 EC3. 1. 1. 26)和磷脂酶(归入 EC3. 1. 1. 4 和 / 或 EC3. 1. 1. 32, 包括归入 EC3. 1. 1. 5 的溶血磷脂酶)。其它优选的酯酶有角质酶(归入 EC3. 1. 1. 74)。

[0277] 可将酯酶以有效量加入从而得到预期的优势以改进发酵微生物的性能, 例如改变发酵微生物内部和 / 或外部, 或发酵微生物的细胞膜中的脂质组成 / 浓度, 以在发酵过程中产生溶质进入和 / 或移出发酵微生物的运动中的改进, 和 / 或以提供更多可代谢的能量来源(诸如例如通过将诸如来自谷物底物的油等组分转化为发酵微生物有用的组分, 例如不饱和的脂肪酸和甘油), 以增加乙醇产量。有效量的酯酶的实例是大约 0.01 至大约 400LU/g DS (干固体)。优选的是, 以大约 0.1 至大约 100LU/g DS、更优选大约 0.5 至大约 50LU/g DS、和甚至更优选大约 1 至大约 20LU/g DS 的量使用酯酶。使用本领域已知的标准程序, 下文可获得酯酶量的进一步优化。

[0278] 一个脂肪酶单位(LU)是以三丁精作为底物和阿拉伯树胶作为乳化剂, 于 30℃、

pH7.0 (磷酸盐缓冲液),每分钟释放 $1.0 \mu\text{mol}$ 可滴定脂肪酸的酶量。

[0279] 在一个优选的方面,酯酶是脂肪分解酶,更优选脂肪酶。在用于本文时,“脂肪分解酶”指脂肪酶和磷脂酶(包括溶血磷脂酶)。脂肪分解酶优选是微生物来源的,特别是细菌、真菌或酵母来源。所用脂肪分解酶可源自任何来源,包括例如犁头霉属 (*Absidia*) 的菌株,特别是 *Absidia blakesleena* 和伞枝犁头霉 (*Absidia corymbifera*),无色杆菌属 (*Achromobacter*) 的菌株,特别是解毒无色杆菌 (*Achromobacter iophagus*),气单胞菌属 (*Aeromonas*) 的菌株,链格孢属 (*Alternaria*) 的菌株,特别是 *Alternaria brassiciola*,曲霉属的菌株,特别是黑曲霉、米曲霉、烟曲霉和黄曲霉,无色杆菌属的菌株,特别是解毒无色杆菌,短梗霉属的菌株,特别是出芽短梗霉 (*Aureobasidium pullulans*),芽孢杆菌属的菌株,特别是短小芽孢杆菌 (*Bacillus pumilus*)、嗜热脂肪芽孢杆菌和枯草芽孢杆菌,白僵菌属 (*Beauveria*) 的菌株,索丝菌属 (*Brochothrix*) 的菌株,特别是热杀索丝菌 (*Brochothrix thermosohata*),假丝酵母属的菌株,特别是 *Candida cylindracea* (皱落假丝酵母 (*Candida rugosa*)), *Candida paralipolytica* 和 *Candida antarctica*,色杆菌属 (*Chromobacter*) 的菌株,特别是粘稠色杆菌 (*Chromobacter viscosum*),鬼伞属的菌株,特别是灰盖鬼伞,镰孢属的菌株,特别是禾谷镰孢、尖孢镰孢、腐皮镰孢 (*Fusarium solani*)、*Fusarium solani pisi*、黄色粉红镰孢 (*Fusarium roseum culmorum*) 和有毒镰孢, *Geotricum* 属的菌株,特别是 *Geotricum penicillatum*,汉逊氏酵母属的菌株,特别是异常汉逊酵母 (*Hansenula anomala*),腐质霉属的菌株,特别是 *Humicola brevispora*、*Humicola brevis var. thermoidea* 和 *Humicola insolens*, *Hyphozyma* 属的菌株,乳杆菌属 (*Lactobacillus*) 的菌株,特别是完全乳杆菌 (*Lactobacillus curvatus*),绿僵菌属 (*Metarhizium*) 的菌株,毛霉属的菌株,拟青霉属的菌株,青霉属的菌株,特别是圆弧青霉 (*Penicillium cyclopium*)、皮落青霉 (*Penicillium crustosum*) 和扩展青霉 (*Penicillium expansum*),假单胞菌属的菌株,特别是铜绿假单胞菌 (*Pseudomonas aeruginosa*)、产碱假单胞菌 (*Pseudomonas alcaligenes*)、葱头假单胞菌 (*Pseudomonas cepacia*) (同义词洋葱伯克霍尔德氏菌 (*Burkholderia cepacia*))、荧光假单胞菌 (*Pseudomonas fluorescens*)、莓实假单胞菌 (*Pseudomonas fragi*)、嗜麦芽假单胞菌 (*Pseudomonas maltophilia*)、门多萨假单胞菌 (*Pseudomonas mendocina*)、*Pseudomonas mephitica lipolytica*、产碱假单胞菌、植物假单胞菌 (*Pseudomonas plantari*)、类产碱假单胞菌 (*Pseudomonas pseudoalcaligenes*)、恶臭假单胞菌 (*Pseudomonas putida*)、施氏假单胞菌 (*Pseudomonas stutzeri*) 和 *Pseudomonas wisconsinensis*,丝核菌属 (*Rhizooctonia*) 的菌株,特别是立枯丝核菌 (*Rhizooctonia solani*)、根毛霉属的菌株,特别是米赫根毛霉,根霉属的菌株,特别是日本根霉 (*Rhizopus japonicus*)、小孢根霉 (*Rhizopus microsporus*) 和 *Rhizopus nodosus*,红冬孢酵母属 (*Rhodosporeidium*) 的菌株,特别是红冬孢酵母 (*Rhodosporeidium toruloides*),红酵母属 (*Rhodotorula*) 的菌株,特别是红酵母 (*Rhodotorula glutinis*),掷孢酵母属 (*Sporobolomyces*) 的菌株,特别是 *Sporobolomyces shibatanus*,嗜热霉属的菌株,特别是细毛嗜热霉 (*Thermomyces lanuginosus*) (以前称为细毛腐质霉 *Humicola lanuginosa*), *Thiarosporella* 的菌株,特别是 *Thiarosporella phaseolina*,木霉属的菌株,特别是哈茨木霉和瑞氏木霉,和 / 或轮枝孢属 (*Verticillium*) 的菌株。

[0280] 在一个优选的方面,脂肪分解酶源自曲霉属、无色杆菌属、芽孢杆菌属、假丝酵母

属、色杆菌属、镰孢属、腐质霉属、Hyphozyma 属、假单胞菌属、根毛霉属、根霉属或嗜热霉属的菌株。

[0281] 在更优选的方面,脂肪分解酶是脂肪酶。在本文中可应用脂肪酶在例如由谷物底物产生的发酵培养基(包括发酵酵母)中修饰甘油三酸酯油和脂肪的结构和组成的能力。脂肪酶催化不同类型的甘油三酸酯转化,诸如水解、酯化和酯交换(transesterification)。合适的脂肪酶包括酸性、中性和碱性脂肪酶,正如本领域众所周知的,虽然酸性脂肪酶(诸如例如可从 Amano 获得的脂肪酶 G AMANO50)与中性或碱性脂肪酶相比,看起来在脂肪酶的较低浓度更有效。优选用于本发明的脂肪酶包括 *Candida antarctica* 脂肪酶和 *Candida cylindracea* 脂肪酶。更优选的脂肪酶是纯化的脂肪酶,诸如 *Candida antarctica* 脂肪酶(脂肪酶 A)、*Candida antarctica* 脂肪酶(脂肪酶 B)、*Candida cylindracea* 脂肪酶和沙门柏干酪青霉(*Penicillium camembertii*)脂肪酶。

[0282] 脂肪酶可以是 EP258,068-A 中公开的一种,或者可以是脂肪酶变体,诸如 W000/60063 或 W000/32758 中公开的变体,引入本文作为参考。优选的商品化脂肪酶包括 LECITASE™、LIPOLASE™ 和 LIPEX™(可从 Novozymes A/S 获得)和 G AMANO™50(可从 Amano 获得)。

[0283] 脂肪酶优选以大约 1 至大约 400LU/g DS、优选大约 1 至大约 10LU/gDS、和更优选大约 1 至大约 5LU/g DS 的量加入。

[0284] 在本发明的另一个优选方面,所述酯酶是角质酶。角质酶是能够降解角质的酶。角质酶可源自任何来源。在一个优选的方面,所述角质酶源自曲霉属的菌株,特别是米曲霉,链格孢属的菌株,特别是 *Alternaria brassiciola*, 镰孢属的菌株,特别是腐皮镰孢、*Fusarium solani pisi*、黄色粉红镰孢或 *Fusarium roseum sambucium*, 长蠕孢属(*Helminthosporium*)的菌株,特别是麦根腐长蠕孢(*Helminthosporium sativum*), 腐质霉属的菌株,特别是 *Humicola insolens*, 假单胞菌属的菌株,特别是门多萨假单胞菌或恶臭假单胞菌,丝核菌属的菌株,特别是立枯丝核菌,链霉菌属的菌株,特别是疮痂病链霉菌(*Streptomyces scabies*), 或单格孢属(*Ulocladium*), 特别是 *Ulocladium consortiale*。在一个最优选的方面,角质酶源自 *Humicola insolens* 的菌株,特别是菌株 *Humicola insolens* DSM1800。 *Humicola insolens* 角质酶在 W096/13580 中有描述,将其引入本文作为参考。角质酶可以是变体,诸如 W000/34450 和 W001/92502 中公开的变体之一,将其引入本文作为参考。优选的角质酶变体包括 W001/92502 的实施例 2 中列出的变体,将其明确引入本文作为参考。角质酶的有效量为大约 0.01 至大约 400LU/g DS、优选大约 0.1 至大约 100LU/g DS、且更优选大约 1 至大约 50LU/g DS。使用本领域已知的标准程序,下文可获得角质酶量的进一步优化。

[0285] 在另一个优选的方面,酯酶是磷脂酶。在用于本文时,术语“磷脂酶”是具有针对磷脂的活性例如水解活性的酶。磷脂,诸如卵磷脂或磷脂酰胆碱,由外部(sn-1)和中间(sn-2)位置以两个脂肪酸酯化且第三个位置以磷酸酯化的甘油组成。可将磷酸酯化成氨基醇。可区分几种类型的磷脂酶活性,包括磷脂酶 A1 和 A2,它们分别在 sn-1 和 sn-2 位置水解一个脂肪酰基以形成溶血磷脂;和溶血磷脂酶(或磷脂酶 B),它水解溶血磷脂中残余的脂肪酰基。磷脂酶 C 和磷脂酶 D(磷酸二酯酶)分别释放二酰基甘油(diacyl glycerol)或磷脂酸(phosphatidic acid)。

[0286] 术语“磷脂酶”包括具有磷脂酶活性的酶,例如磷脂酶 A (A1 或 A2)、磷脂酶 B 活性、磷脂酶 C 活性、或磷脂酶 D 活性。术语“磷脂酶 A”在用于本文时意图涵盖具有磷脂酶 A1 和 / 或磷脂酶 A2 活性的酶。也可通过具有其它活性的酶来提供磷脂酶活性,诸如例如具有磷脂酶活性的脂肪酶。磷脂酶活性可例如来自具有磷脂酶副活性 (side activity) 的脂肪酶。在其它方面,通过本质上只具有磷脂酶活性的酶来提供磷脂酶活性,而且其中磷脂酶活性不是副活性。

[0287] 磷脂酶可以是任何来源,例如动物来源(例如哺乳动物,例如牛或猪胰)、或蛇毒或蜂毒。或者,磷脂酶可以是微生物来源,例如来自丝状真菌、酵母或细菌,诸如曲霉属,例如泡盛曲霉、臭曲霉、日本曲霉、黑曲霉或米曲霉,网柄菌属 (*Dictyostelium*), 例如盘基网柄菌 (*D. discoideum*); 镰孢属,例如黄色镰孢、禾谷镰孢、异孢镰孢、腐皮镰孢、尖孢镰孢或有毒镰孢;毛霉属,例如,爪哇毛霉 (*M. javanicus*)、大毛霉 (*M.ucedo*) 或细孢毛霉 (*M. subtilissimus*); 脉孢霉属,例如粗糙脉孢菌;根毛霉属,例如,微小根毛霉 (*R. pusillus*); 根霉属,例如少根根霉 (*R. arrhizus*)、日本根霉或匍枝根霉 (*R. stolonifer*); 核盘菌属 (*Sclerotinia*), 例如大豆核盘菌 (*S. libertiana*); 发癣菌属 (*Trichophyton*), 例如红色发癣菌 (*T. rubrum*); Whetzelinia 属,例如 *W. sclerotiorum*; 芽孢杆菌属,例如巨大芽孢杆菌或枯草芽孢杆菌;柠檬酸杆菌属 (*Citrobacter*), 例如弗氏柠檬酸杆菌 (*C. freundii*); 肠杆菌属 (*Enterobacter*), 例如产气肠杆菌 (*E. aerogenes*) 或阴沟肠杆菌 (*E. cloacae*); 爱德华氏菌属 (*Edwardsiella*), 迟钝爱德华氏菌 (*E. tarda*); 欧文氏菌属 (*Erwinia*), 例如草生欧文氏菌 (*E. herbicola*); 埃希氏菌属 (*Escherichia*), 例如大肠杆菌 (*E. coli*); 克雷伯氏菌属 (*Klebsiella*), 例如肺炎克雷伯氏菌 (*K. pneumoniae*); 变形菌属 (*Proteus*), 例如普通变形菌 (*P. vulgaris*); 普罗威登斯菌属 (*Providencia*), 例如斯氏普罗威登斯菌 (*P. stuartii*); 沙门氏菌属 (*Salmonella*), 例如鼠伤寒沙门氏菌 (*S. typhimurium*); 沙雷氏菌属 (*Serratia*), 例如液化沙雷氏菌 (*S. liquefasciens*)、粘质沙雷氏菌 (*S. marcescens*); 志贺氏菌属 (*Shigella*), 例如弗氏志贺氏菌 (*S. flexneri*); 链霉菌属,例如,紫红链霉菌 (*S. violeceoruber*); 或耶尔森氏菌属 (*Yersinia*), 例如小肠结肠炎耶尔森氏菌 (*Y. enterocolitica*)。

[0288] 优选的商品化磷脂酶包括 LECITASE™ 和 LECITASE™ULTRA(可从 Novozymes A/S 获得)。

[0289] 磷脂酶的有效量为大约 0.01 至大约 400LU/g DS、优选大约 0.1 至大约 100LU/g DS、且更优选大约 1 至大约 50LU/g DS。使用本领域已知的标准程序,下文可获得磷脂酶量的进一步优化。

[0290] 蛋白酶

[0291] 在本发明的另一个优选方面,将至少一种表面活性剂和至少一种碳水化合物生成酶与至少一种蛋白酶组合使用。可使用蛋白酶例如来消化蛋白质以产生游离的氨基酸 (FAN)。所述游离的氨基酸起到酵母营养物的作用,从而增强酵母的生长,并因此增强乙醇的生产。

[0292] 可通过在存在至少一种蛋白酶时繁殖发酵微生物来产生用于发酵过程的发酵微生物。虽然不限于任何一种操作理论,但相信当发酵微生物后续用于发酵过程时,在含有有效量的至少一种蛋白酶的情况下繁殖发酵微生物与不加蛋白酶在相同条件下繁殖的发酵

微生物相比缩短了发酵微生物的迟滞时间 (lag time)。相信蛋白酶在繁殖过程中的作用是直接或间接分别导致在发酵过程中对于发酵微生物有害或有益基因的抑制或表达,从而缩短了迟滞时间并得到更快的发酵循环。

[0293] 蛋白酶是本领域众所周知的,并指催化切割肽键的酶。合适的蛋白酶包括真菌和细菌蛋白酶。优选的蛋白酶是酸性蛋白酶,即特征在于在低于 pH7 的酸性条件下水解蛋白质的能力的蛋白酶。合适的酸性真菌蛋白酶包括源自曲霉属、毛霉属、根霉属、假丝酵母属、革盖菌属、内座壳属 (Endothia)、Entomophtra、耙菌属 (Irpex)、青霉属、小核菌属 (Sclerotium) 和球拟酵母属 (Torulopsis) 的真菌蛋白酶。尤其预期源自黑曲霉(参见例如 Koaze 等人,1964, Agr. Biol. Chem. Japan28 :216)、斋藤曲霉 (Aspergillus saitoi) (参见例如 Yoshida,1954, J. Agr. Chem. Soc. Japan28 :66)、泡盛曲霉(Hayashida 等人,1977, Agric. Biol. Chem. 42 :927-933)、棘孢曲霉(W095/02044) 或米曲霉的蛋白酶;和源自微小毛霉或米赫毛霉的酸性蛋白酶。

[0294] 不是酸性蛋白酶的细菌蛋白酶包括商品化产品 ALCALASE™ 和 NEUTRASE™ (可从 Novozymes A/S 获得)。其它蛋白酶包括来自 Genencor Int, Inc. (美国) 的 GC106 和来自 Novozymes A/S 的 NOVOZYM™50006。

[0295] 优选的是,蛋白酶是天冬氨酸蛋白酶,正如例如《Handbook of Proteolytic Enzymes》(A. J. Barrett, N. D. Rawlings 和 J. F. Woessner 编, Academic Press, San Diego, 1998, 第 270 章) 中所述。天冬氨酸蛋白酶的合适实例包括例如 Berka 等人,1990, Gene96 :313 ;Berka 等人,1993, Gene125 :195-198; 及 Gomi 等人,1993, Biosci. Biotech. Biochem. 57 :1095-1100 中公开的那些。

[0296] 过氧化物酶

[0297] 具有过氧化物酶活性的其它化合物可以是任何过氧化物酶 (EC1. 11. 1. 7), 或具有由其衍生的过氧化物酶活性、显示过氧化物酶活性的任何片段。

[0298] 优选的是,过氧化物酶由植物(例如辣根或大豆过氧化物酶) 或微生物诸如真菌或细菌产生。

[0299] 一些优选的真菌包括属于半知菌亚门 (subdivision Deuteromycotina) 丝孢纲 (class Hyphomycetes) 的菌株,例如镰孢属、腐质霉属、木霉属、漆斑菌属 (Myrothecium)、轮枝孢属 (Verticillium)、Arthromyces 属、卡尔黑霉属 (Caldariomyces)、单隔孢属、Embellisia 属、支孢属 (Cladosporium) 或 Dreschlera, 特别是尖孢镰孢 (DSM2672)、Humicola insolens、瑞氏木霉、疣孢漆斑菌 (Myrothecium verrucaria) (IF06113)、黄菱轮枝孢菌 (Verticillium albo-atrum)、大丽花轮枝孢菌 (Verticillium dahliae)、Arthromyces ramosus (FERMP-7754)、Caldariomyces fumago、纸单隔孢 (Ulocladium chartarum)、Embellisia alli 或 Dreschlera halodes。

[0300] 其它优选的真菌包括属于担子菌亚门 (subdivision Basidiomycotina) 担子菌纲 (class Basidiomycetes) 的菌株,例如鬼伞属、Phanerochaete 属、革盖菌属或栓菌属,特别是 Coprinus cinereus f. microsporus (IF08371)、长根鬼伞 (Coprinus macrorrhizus)、Phanerochaete chrysosporium (例如 NA-12) 或栓菌属(以前称为多孔菌属 (Polyporus)), 例如 T. versicolor (例如 PR428-A)。

[0301] 其它优选的真菌包括属于接合菌亚门 (subdivision Zygomycotina) Mycoraceae

纲的菌株,例如根霉属或毛霉属,特别是冻土毛霉 (*Mucor hiemalis*)。

[0302] 一些优选的细菌包括放线菌目 (order Actinomycetales) 的菌株,例如类球形链霉菌 (*Streptomyces spheroides*) (ATTC23965)、热紫链霉菌 (*Streptomyces thermoviolaceus*) (IF012382) 或轮枝链霉菌轮枝亚种 (*Streptoverticillum verticillium* ssp. *Verticillium*)。

[0303] 其它优选的细菌包括类球红细菌 (*Rhodobacter sphaeroides*)、*Rhodomonas palustri*、乳链球菌 (*Streptococcus lactis*)、*Pseudomonas purrocina* (ATCC15958)、荧光假单胞菌 (NRRL B-11) 和芽孢杆菌属菌株,例如短小芽孢杆菌 (ATCC12905) 和嗜热脂肪芽孢杆菌。

[0304] 其它优选的细菌包括属于粘球菌属 (*Myxococcus*) 的菌株,例如变绿粘球菌 (*M. virescens*)。

[0305] 过氧化物酶也可以是通过如下方法产生的一种,所述方法包括在培养基中在允许过氧化物酶表达的条件下培养用重组 DNA 载体转化的宿主细胞,所述重组 DNA 载体携带编码过氧化物酶的 DNA 序列以及用于表达编码过氧化物酶的 DNA 序列的 DNA 序列,并从培养物中回收过氧化物酶。

[0306] 在一个优选的方面,重组产生的过氧化物酶是源自鬼伞菌种的过氧化物酶,特别是根据 W092/16634 的长根鬼伞或灰盖鬼伞。

[0307] 在本发明中,具有过氧化物酶活性的化合物包含过氧化物酶和源自细胞色素、血红蛋白或过氧化物酶的过氧化物酶活性片段。

[0308] 一个过氧化物酶单位 (POXU) 是在如下条件下每分钟催化转化 $1 \mu\text{mole}$ 过氧化氢的酶量,所述条件为于 30°C , 在 0.1M 磷酸盐缓冲液 pH7.0、 0.88mM 过氧化氢、和 1.67mM 2,2'-连氨基-二(3-乙基苯并噻唑啉-6-磺酸盐) (2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonate)) (ABTS)。监测反应 60 秒 (混合后 15 秒),即 418nm 吸光度的改变,它应在 0.15 至 0.30 的范围内。为了计算活性,使用氧化 ABTS 的吸光系数 $36\text{mM}^{-1}\text{cm}^{-1}$ 和化学计量关系每氧化 $2 \mu\text{mole}$ ABTS 转化 $1 \mu\text{mole}$ H_2O_2 。

[0309] 漆酶

[0310] 在本发明中,漆酶和漆酶相关酶包括归入 EC1.10.3.2 的任何漆酶、归入 EC1.10.3.1 的任何儿茶酚氧化酶、归入 EC1.3.3.5 的任何胆红素氧化酶、或归入 EC1.14.18.1 的任何一元酚单加氧酶。

[0311] 上文所述酶可以是微生物的,即得自细菌或真菌(包括丝状真菌和酵母),或者它们可源自植物。

[0312] 来自真菌的合适实例包括得自如下菌株的漆酶:曲霉属,脉孢菌属,例如粗糙脉孢菌,柄孢壳属 (*Podospora*),葡萄孢属 (*Botrytis*),金钱菌属 (*Collybia*),层孔菌属 (*Fomes*),香菇属 (*Lentinus*),侧耳属,栓菌属,例如 *T. villosa* 和 *T. versicolor*,丝核菌属,例如立枯丝核菌,鬼伞属,例如灰盖鬼伞、毛头鬼伞 (*C. comatus*)、费赖斯鬼伞 (*C. friesii*) 和褶纹鬼伞 (*C. plicatilis*),小脆柄菇属 (*Psathyrella*),例如黄盖小脆柄菇 (*P. condelleana*),斑褶菇属 (*Panaeolus*),例如蝶形斑褶菇 (*P. papilionaceus*),毁丝霉属,例如嗜热毁丝霉,*Schytalidium* 属,例如 *S. thermophilum*,多孔菌属,例如 *P. pinsitus*,密孔菌属 (*Pycnoporus*),例如朱红密孔菌 (*P. cinnabarinus*),射脉菌属 (*Phlebia*),例如射

脉菌 (*P. radita*) (W092/01046), 或革盖菌属, 例如毛革盖菌 (JP2-238885)。

[0313] 来自细菌的合适实例包括得自芽孢杆菌菌株的漆酶。

[0314] 优选得自鬼伞属、毁丝霉属、多孔菌属、密孔菌属、柱顶孢属或丝核菌属的漆酶; 特别是得自灰盖鬼伞、嗜热毁丝霉、*Polyporus pinsitus*、朱红密孔菌、嗜热柱顶孢或立枯丝核菌的漆酶。

[0315] 商品化漆酶有 NS51001 (*Polyporus pinsitius* 漆酶, 可从 Novozymes A/S 获得) 和 NS51002 (嗜热毁丝霉漆酶, 可从 Novozymes A/S 获得)。

[0316] 漆酶或漆酶相关酶也可以是通过如下方法产生的一种, 所述方法包括在培养基中在允许漆酶表达的条件下培养用重组 DNA 载体转化的宿主细胞, 所述重组 DNA 载体携带编码漆酶的 DNA 序列以及用于表达编码漆酶的 DNA 序列的 DNA 序列, 并从培养物中回收漆酶。

[0317] 漆酶活性 (LACU) 由需氧条件下于 pH5.5 的丁香醛连氮 (*syringaldazin*) 氧化来测定。产生的紫颜色在 530nm 用光度计测定。分析条件为 19mM 丁香醛连氮、23mM 乙酸盐缓冲液 pH5.5, 30°C, 1 分钟反应时间。一个漆酶单位 (LACU) 是上述条件下每分钟催化 1.0 μ mole 丁香醛连氮转化的酶量。

[0318] 漆酶活性 (LAMU) 由需氧条件下于 pH7.5 的丁香醛连氮氧化来测定。产生的紫颜色在 530nm 用光度计测定。分析条件为 19mM 丁香醛连氮、23mM Tris/马来酸盐 pH7.5, 30°C, 1 分钟反应时间。一个漆酶单位 (LAMU) 是上述条件下每分钟催化 1.0 μ mole 丁香醛连氮转化的酶量。

[0319] 本发明的多肽可以联合上文所述酶和 / 或分解纤维素的蛋白质用于进一步降解生物材料底物的纤维素组分 (参见例如 Brigham 等人, 1995, 在《Handbook on Bioethanol》中, Charles E. Wyman 编, pp. 119-141, Taylor&Francis, Washington D. C.; Lee, 1997, Journal of Biotechnology 56:1-24)。

[0320] 洗涤剂组合物

[0321] 本发明具有增强分解纤维素活性的多肽可添加到洗涤剂组合物中并因此成为它的一种组分。

[0322] 例如, 可以将本发明的洗涤剂组合物配制成手洗或机洗洗衣洗涤剂组合物, 包括适用于预处理污染织物的洗衣添加剂组合物和漂洗时添加的织物柔软剂组合物, 或配制成用于普通家庭硬面清洗操作的洗涤剂组合物, 或配制成用于手洗或机洗洗碟操作的洗涤剂组合物。

[0323] 在一个具体的方面, 本发明提供了含有本发明具有增强分解纤维素活性的多肽的洗涤剂添加剂。洗涤剂添加剂和洗涤剂组合物中可包含一种或多种其它的酶, 诸如蛋白酶、脂肪酶、角质酶、淀粉酶、糖酶、纤维素酶、果胶酶、甘露聚糖酶、阿拉伯糖酶、半乳聚糖酶、木聚糖酶、氧化酶 (如漆酶)、和 / 或过氧化物酶。

[0324] 通常, 所选酶的性质应该与所选洗涤剂相容 (即最佳 pH、与其它酶和非酶组分的相容性等等), 并且所述酶应该以有效量存在。

[0325] 纤维素酶: 适宜的纤维素酶包括细菌或真菌起源的那些。包括经化学修饰的或蛋白质工程改造的突变体。适宜的纤维素酶包括来自下列属的纤维素酶: 芽孢杆菌属、假单胞菌属、腐质霉属、镰孢属、梭孢壳属、*Acremonium* 属, 例如在 US4, 435, 307、US5, 648, 263、US5, 691, 178、US5, 776, 757、和 W089/09259 中公开的由 *Humicola insolens*、嗜热毁丝霉、

和尖孢镰孢产生的真菌纤维素酶。

[0326] 尤其适宜的纤维素酶是有益于保护颜色的碱性或中性纤维素酶。这些纤维素酶的实例有 EP0495257、EP0531372、W096/11262、W096/29397、W098/08940 中描述的纤维素酶。其它实例有诸如在 W094/07998、EP0531315、US5,457,046、US5,686,593、US5,763,254、W095/24471、W098/12307、和 PCT/DK98/00299 中描述的纤维素酶变体。

[0327] 市场上可购买到的纤维素酶包括 Celluzyme™ 和 Carezyme™ (Novozymes A/S)、Clazinase™ 和 Puradax HA™ (Genencor International Inc.)、及 KAC-500(B)™ (Kao Corporation)。

[0328] **蛋白酶**:适宜的蛋白酶包括那些动物、植物或微生物起源的。优选微生物起源的。包括经化学修饰或蛋白质工程改造的突变体。蛋白酶可以是丝氨酸蛋白酶或金属蛋白酶,优选碱性微生物蛋白酶或胰蛋白酶样蛋白酶。碱性蛋白酶的实例是枯草杆菌蛋白酶,尤其是那些来源于芽孢杆菌属的那些,例如枯草杆菌蛋白酶 Novo、枯草杆菌蛋白酶 Carlsberg、枯草杆菌蛋白酶 309、枯草杆菌蛋白酶 147、和枯草杆菌蛋白酶 168 (描述在 W089/06279 中)。胰蛋白酶样蛋白酶的实例是胰蛋白酶(例如起源于猪或牛的)和镰孢属蛋白酶(描述于 W089/06270 和 W094/25583 中)。

[0329] 有用的蛋白酶的实例是 W092/19729、W098/20115、W098/20116、和 W098/34946 中所描述的变体,尤其是在一个或多个如下位置发生替代的变体:27、36、57、76、87、97、101、104、120、123、167、170、194、206、218、222、224、235、和 274。

[0330] 优选的市场上可买到的蛋白酶包括 Alcalase™、Savinase™、Primase™、Duralase™、Esperase™ 和 Kannase™ (Novozymes A/S)、Maxatase™、Maxacal™、Maxapem™、Properase™、Purafect™、Purafect OxP™、FN2™ 和 FN3™ (Genencor International Inc.)。

[0331] **脂肪酶**:适宜的脂肪酶包括细菌或真菌起源的那些。包括经化学修饰的或经蛋白质工程改造的变体。有用的脂肪酶的实例包括来自腐质霉属(同义词嗜热霉属 *Thermomyces*) 的脂肪酶,例如 EP258068 和 EP305216 中所描述的 *H. lanuginosa* (*T. lanuginosus*) 或 W096/13580 中所描述的 *H. insolens*;假单胞菌属脂肪酶,例如产碱假单胞菌(*P. alkaligenes*) 或类产碱假单胞菌(*P. pseudoalcaligenes*) (EP218272)、葱头假单胞菌(*P. cepacia*) (EP331376)、施氏假单胞菌(*P. stutzeri*) (GB1,372,034)、荧光假单胞菌(*P. fluorescens*)、假单胞菌种菌株 SD705 (W095/06720) 和 W096/27002)、*P. wisconsinensis* (W096/12012);芽孢杆菌属脂肪酶,例如枯草芽孢杆菌(Dartois 等人,1993, *Biochemica et Biophysica Acta*,1131,253-360)、嗜热脂肪芽孢杆菌(JP64/744992)、或短小芽孢杆菌(W091/16422)。

[0332] 其它实例有诸如在 W092/05249、W094/01541、EP407225、EP260105、W095/35381、W096/00292、W095/30744、W094/25578、W095/14783、W095/22615、W097/04079、和 W097/07202 中所描述的脂肪酶变体。

[0333] 优选的市场上可买到的脂肪酶包括 Lipolase™ 和 Lipolase Ultra™ (Novozymes A/S)。

[0334] **淀粉酶**:适宜的淀粉酶(α 和 / 或 β) 包括细菌和真菌起源的那些。包括经化学修饰的或蛋白质工程改造的突变体。淀粉酶包括例如芽孢杆菌属例如地衣芽孢杆菌特定菌株来源的 α -淀粉酶,详细说明描述在 GB1,296,839 中。

[0335] 有用的淀粉酶的实例是在 W094/02597、W094/18314、W096/23873、和 W097/43424 中所描述的变体,尤其是在一个或多个如下位置发生替代的变体:15、23、105、106、124、128、133、154、156、181、188、190、197、202、208、209、243、264、304、305、391、408、和 444。

[0336] 市场上可买到的淀粉酶有 Duramyl™、Termamyl™、Fungamyl™ 和 BAN™ (Novozymes A/S)、Rapidase™ 和 Purastar™ (Genencor International Inc.)。

[0337] 过氧化物酶/氧化酶:适宜的过氧化物酶/氧化酶包括源自植物、细菌、或真菌的那些。包括经化学修饰或蛋白质工程改造的突变体。有用的过氧化物酶的实例包括来自鬼伞属例如灰盖鬼伞的过氧化物酶及其变体,如在 W093/24618、W095/10602、和 W098/15257 中所描述的那些。

[0338] 市场上可买到的过氧化物酶包括 Guardzyme™ (Novozymes A/S)。

[0339] 可通过添加含一种或多种酶的独立的添加剂或通过添加包含所有这些酶的组合添加剂将酶组分包括在洗涤剂组合物中。本发明的洗涤剂添加剂,即独立的添加剂或组合添加剂,可配制成例如颗粒、液体、浆体等形式。优选的洗涤剂添加剂制剂为颗粒剂,特别是非粉化的颗粒剂;液体,特别是稳定的液体;或浆体。

[0340] 例如可如在美国专利 US4,106,991 和 4,661,452 中所公开的方法制备非粉化颗粒剂,并可任选使用本领域已知的方法进行包衣。蜡质包衣材料的实例有平均摩尔质量为 1000-20000 的聚(环氧乙烷)产物(聚乙二醇,PEG);具有 16-50 个环氧乙烷单元的乙氧基化壬基酚;具有 15-80 个环氧乙烷单元的乙氧基化脂肪醇,其中醇含 12-20 个碳原子;脂肪醇;脂肪酸;以及脂肪酸的单-、双-和三甘油酯。GB1483591 中给出了适于通过流化床技术应用的成膜包衣材料的实例。例如,液体酶制备物可根据既定的方法通过加入多元醇诸如丙二醇、糖或糖醇、乳酸或硼酸来稳定。受保护的酶可根据 EP238,216 中所公开的方法制备。

[0341] 本发明的洗涤剂组合物可采用任一种方便的形式,例如条、片、粉末、颗粒、膏、或液体。液体洗涤剂可以是含水的,一般含可高达 70% 的水和 0-30% 的有机溶剂,或者是不含水的。

[0342] 洗涤剂组合物可含有一种或多种表面活性剂,可以是非离子的,包括半极性和/或阴离子和/或阳离子和/或两性离子的。表面活性剂的含量一般为以重量计的 0.1%-60%。

[0343] 当洗涤剂包含阴离子表面活性剂时,它通常含有大约 1% 到大约 40% 的阴离子表面活性剂,诸如线状的烷基苯磺酸酯、 α -石蜡磺酸酯、烷基硫酸酯(脂肪醇硫酸酯)、醇乙氧基硫酸酯、二级烷基磺酸酯、 α -磺基脂肪酸甲酯、烷基-或烯基琥珀酸、或皂。

[0344] 当洗涤剂中包含非离子型表面活性剂时,它通常含有大约 0.2% 到大约 40% 的非离子型表面活性剂,诸如醇乙氧基化物、壬基苯酚乙氧基化物、烷基多苷、烷基二甲基胺氧化物、乙氧基脂肪酸单乙醇酰胺、脂肪酸单乙醇酰胺、多羟基烷基脂肪酸酰胺、或葡糖胺的 N-酰基 N-烷基衍生物(“葡糖酰胺”)。

[0345] 洗涤剂可包含 0-65% 的洗涤剂助洗剂或络合剂,诸如沸石、二磷酸盐、三磷酸盐、膦酸酯、碳酸盐、柠檬酸盐、氨三乙酸、乙二胺四乙酸、二亚乙基三胺五乙酸、烷基-或烯基琥珀酸、可溶性硅酸盐、或层状硅酸盐(例如购自 Hoechst 的 SKS-6)。

[0346] 洗涤剂可包含一种或多种聚合物。实例有羧甲基纤维素、聚(乙烯吡咯烷酮)、聚(乙二醇)、聚(乙烯醇)、聚(乙烯吡啶-N-氧化物)、聚(乙烯咪唑)、聚羧酸酯诸如聚丙烯

烯酸酯、马来酸 / 丙烯酸共聚物、和甲基丙烯酸月桂酯 / 丙烯酸共聚物。

[0347] 洗涤剂可包含漂白系统,该系统可含有 H₂O₂ 来源诸如过硼酸盐或过碳酸盐,可将其与形成过酸的漂白活化剂诸如四乙酰基乙二胺或壬酰氧基苯磺酸酯相组合。或者,漂白系统可包含过氧酸,例如酰胺、亚胺、或砒型。

[0348] 可使用常规的稳定剂稳定本发明洗涤剂组合物的酶,例如多元醇,诸如丙二醇或丙三醇;糖或糖醇;乳酸;硼酸,或硼酸衍生物,例如芳香族硼酸酯,或苯基硼酸衍生物诸如 4-甲酰苯基硼酸;且所述的组合物可如 W092/19709 和 W092/19708 中所描述的方法来配制。

[0349] 洗涤剂还可包含其它的常规洗涤剂成分,诸如包括粘土的织物调节剂、泡沫促进剂、抑泡剂、防蚀剂、污垢悬浮剂、抗污垢再沉积剂、染色剂、杀菌剂、光漂白剂、水溶助长剂、酶暗抑制剂、或芳香剂。

[0350] 在洗涤剂组合物中,任何酶都可以相当于每升洗涤液 0.01-100mg 酶蛋白质的量加入,优选每升洗涤液 0.05-5mg 酶蛋白质,特别是每升洗涤液 0.1-1mg 酶蛋白质。

[0351] 在清洁剂组合物中,本发明具有增强分解纤维素活性的多肽可以相当于每升洗涤液 0.001-100mg 蛋白质、优选 0.005-50mg 蛋白质、更优选 0.01-25mg 蛋白质、甚至更优选 0.05-10mg 蛋白质、最优选 0.05-5mg 蛋白质、且甚至最优选 0.01-1mg 蛋白质的量加入。

[0352] 还可将本发明具有增强分解纤维素活性的多肽添加到如 W097/07202 所公开的洗涤剂制剂中,这里收入所述文献作为参考。

[0353] 其他用途

[0354] 一般而言,可以通过补充本发明具有增强分解纤维素活性的多肽来增强对任何植物细胞壁材料的处理。

[0355] 信号肽

[0356] 本发明还涉及包含编码蛋白质的基因且该基因与如下核苷酸序列可操作连接的核酸构建体,所述核苷酸序列由 SEQ ID NO:1 第 1 至 66 位核苷酸组成并编码由 SEQ ID NO:2 第 1 至 22 位氨基酸组成的、容许所述蛋白质分泌到培养基中的信号肽,其中所述基因对于所述核酸序列而言是外源的。

[0357] 本发明还涉及含有这种核酸构建体的重组表达载体和重组宿主细胞。

[0358] 本发明还涉及生产蛋白质的方法,包括(a)在有助于产生所述蛋白质的条件下培养所述重组宿主细胞;并(b)回收所述蛋白质。

[0359] 对宿主细胞所述蛋白质可以是天然的或异源的。术语“蛋白质”在本文中不是指规定长度的编码产物,而是涵盖肽、寡肽、和蛋白质。术语“蛋白质”还涵盖两种或多种多肽组合形成编码产物。蛋白质还包括杂合多肽,它包括来自于至少两种不同蛋白质的部分或完整多肽序列的组合,其中所述一种或多种蛋白质对宿主细胞可以是异源的或天然的。蛋白质还包括上述蛋白质和杂合蛋白的天然存在的等位和改造(engineered)变异。

[0360] 优选的是,蛋白质是激素或其变体、酶、受体或其部分、抗体或其部分、或报道分子。在一个更优选的方面,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。在一个甚至更优选的方面,蛋白质是氨肽酶、淀粉酶、糖酶、羧肽酶、过氧化氢酶、纤维二糖水解酶、纤维素酶、壳多糖酶、角质酶、环糊精糖基转移酶、脱氧核糖核酸酶、内切葡聚糖酶、酯酶、 α -半乳糖苷酶、 β -半乳糖苷酶、葡糖淀粉酶、 α -葡糖苷酶、 β -葡糖苷酶、转化

酶、漆酶、脂肪酶、甘露糖苷酶、突变酶、氧化酶、果胶分解酶、过氧化物酶、植酸酶、多酚氧化酶、蛋白水解酶、核糖核酸酶、转谷氨酰胺酶、或木聚糖酶。

[0361] 所述基因可从任何原核的、真核的、或其它来源中获得。

[0362] 本发明可通过下列实施例进一步说明,这些实施例不应视为限制本发明的范围。

实施例

[0363] 材料

[0364] 用作缓冲液和底物的化学制品至少是试剂级的商品。

[0365] 菌株

[0366] 使用橙色嗜热子囊菌菌株 NN044936T002-5 作为具有增强分解纤维素活性的 GH61 家族多肽的来源。使用米曲霉 JaL250 菌株(W099/61651)来表达具有增强分解纤维素活性的橙色嗜热子囊菌多肽。

[0367] 培养基

[0368] 用于橙色嗜热子囊菌的常规涂板和培养的固体马铃薯右旋糖培养基的组成是每升 39 克马铃薯右旋糖琼脂(Difco BD Biosciences, Franklin Lakes, NJ)。用于橙色嗜热子囊菌培养物的培养的液体培养基的组成是每升 24 克马铃薯右旋糖肉汤。

[0369] Luria-Bertani (LB)培养基的组成是每升 10g 胰胨、5g 酵母提取物、和 10g NaCl。同样也制备用于涂布的 LB 培养基,只是每升添加 15g 琼脂。将 5'-溴-4-氯-3-吡啶-β-D-吡喃半乳糖苷(X-Gal;Gold Biotechnology, Louis, MO)以每毫升 40 μg 的储液浓度溶解在 N,N'-二甲基甲酰胺中,常规涂布在 LB 平板上用于差异菌落筛选。氨苄青霉素以每升 50-100mg 的终浓度添加到 LB 培养基中,用于选择性生长条件。

[0370] NNCYP 培养基的组成是每升 5.0g NH_4NO_3 、0.5g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 、0.3g CaCl_2 、2.5g 柠檬酸、5.0g 细菌蛋白胨、1.0g 酵母提取物、COVE 痕量金属、和足够的 K_2HPO_4 以达到最终 pH 大约 5。Cove 痕量金属溶液的组成是每升 0.04g $\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}$ 、0.4g $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ 、1.2g $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ 、0.7g $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ 、0.8g $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ 、和 10g $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ 。

[0371] 实施例 1:基因组 DNA 文库构建

[0372] 从橙色嗜热子囊菌生长的 PDA 平板上取一块琼脂,接种装有 200ml 含 3% 葡萄糖、pH5.0 的 NNCYP 培养基的 500 毫升带挡板烧瓶。将培养物于 45°C 振荡培养过夜,转速 170rpm。通过 Whatman#1 滤纸(Whatman Inc, Clifton NJ)过滤收集菌丝体并在液氮中冷冻。在已冷却的研钵中将冷冻的菌丝体研磨成粉末并分装到带螺旋帽的试管中。将粉末悬浮在总体积 40ml 的含 0.5% 十二烷基硫酸锂和 0.5mM EDTA 的 50mM CAPS-NaOH 缓冲液中。将悬浮液置于 60°C 并周期性地颠倒混合 2 小时。再加入等体积的经中和的苯酚:氯仿(1:1),并将试管在转台上于 37°C 混合 2 小时。在 Sorvall H1000B 转头(Kendro, Asheville, NC)中以 2500rpm 离心 10 分钟后,再提取出水相并如上所述进行离心。向第二次抽提的水相加入乙酸铵至 2.5M 并置于 -20°C 直到冻结。融化后,将提取物以 15,000×g 离心 20 分钟。弃去沉淀物,通过添加 0.7 倍体积的异丙醇使上清液中的核酸沉淀。以 15,000×g 离心后,将沉淀物用 70% 乙醇漂洗三次,风干,并溶于 1.0ml 10.1X TE 中。通过添加乙酸铵到 2.0M 和乙醇到 63% 使溶解的沉淀物再一次沉淀。将沉淀物用 70% 乙醇漂洗两次,干燥,并溶于 200 μl 10.1X TE 中。一旦溶解,向 DNA 加 KCl 到 20mM。

[0373] 利用 **TOPO**[®] Shotgun Subcloning Kit 即亚克隆试剂盒 (Invitrogen, Inc, Carlsbad, CA) 构建橙色嗜热子囊菌的基因组文库。根据制造商的推荐,在氩气下通过喷雾随机剪切基因组 DNA。通过 TAE 缓冲液中 1% 琼脂糖凝胶上的制备性凝胶电泳,挑选 2.5–5.0kb 大小的插入物。

[0374] 根据制造商的说明,将基因组 DNA 克隆到 **pCR**[®] 4Blunt-TOPO 中,并用于转化大肠杆菌 TOP10 细胞 (Invitrogen, Inc, Carlsbad, CA)。通过在上文所述补充有氨苄青霉素和 X-Gal 的 LB 琼脂上的初始涂布,估计产生了大约 64,000 个克隆。在测定重组克隆的滴度后立即扩增文库。将每个转化反应用于接种 200ml 含 100 μ g/ml 氨苄青霉素的 LB 培养基。然后于 37°C 过夜培养摇瓶培养物至饱和,再通过碱裂解法提取质粒 DNA 并利用 Qiagen Plasmid Maxi 试剂盒进行纯化。

[0375] 随后将扩增的文库通过再转化并将大肠杆菌 TOP10 细胞涂布在上文所述补充有氨苄青霉素和 X-Gal 的 LB 琼脂培养基上来产生用于克隆分离克隆的菌落。根据制造商的说明,利用 Q-Pix 机械臂进行菌落挑取。最初将总共 22,353 个单独的菌落挑取到 96 孔板中,然后在上文培养基部分概述的选择性条件下培养,并于 -80°C 以甘油原种的形式保存。

[0376] 实施例 2:滚环扩增和微阵列的印制

[0377] 开发了用于滚环扩增 (RCA) 和产物稀释的机械臂法,以用于 Biomek FX 机械臂 (Beckman Coulter, Brea, CA)。将 TempliPhi DNA 测序模板扩增试剂盒 (Amersham Biosciences Corp., Piscataway, NJ) 用于 RCA,并在机械臂法中遵照制造商的方案。

[0378] 对来自文库的 9,972 个克隆进行 RCA 反应。从文库的 9,600 个 RCA 产物中制备出双份 384 孔稀释平板 (Genemate, Kaysville, UT),并将这些克隆印在聚 L-赖氨酸包被的载玻片上从而得到微阵列,每个微阵列呈现大约 33Mb 克隆的 DNA。另外,将橙色嗜热子囊菌 *cbh1* (登记号 AX657575) 和 *xynA* (登记号 AJ132635) 基因点在上面作为阳性对照,并包括空的 **pCR**[®] 4Blunt-TOPO 载体和从未转化大肠杆菌细胞制备的 RCA 作为阴性对照。利用美国专利号 5,807,522 描述的方法进行阵列印制。

[0379] 实施例 3:发酵和 RNA 提取

[0380] 所有的发酵都是在 2 升 Applikon 发酵罐中进行的,分批补料 5.2% (w/v) 葡萄糖或 5.2% 纤维素碳源。基础培养基是 NNCYP,且纤维素培养物还含有 0.4% 葡萄糖以促进初始细胞生长。在发酵期间通过添加氢氧化铵或磷酸来控制 pH。发酵于 42°C pH5 进行大约 120 小时。在接种后第 1–5 天采取菌丝体样品。通过 Miracloth[™] (Calbiochem, San Diego, CA) 过滤迅速从培养基中分离菌丝体样品,然后在液氮中冷冻并保存于 -80°C。

[0381] 遵照制造商的方案,用 **FastRNA**[®] 试剂盒 (Q•BIOgene, Carlsbad, CA) 从菌丝体样品中提取 RNA。通过在 TBE 缓冲液在中在 1% 琼脂糖凝胶上于 55V 进行 1–1.5 小时的电泳,或通过使用 Agilent2100 生物分析仪 (Agilent Technologies, Santa Clara, CA) 进行的毛细管电泳分析,评估 RNA 的质量。通过紫外分光光度法和 / 或 Agilent2100 生物分析仪的分析进行定量测定。只有来自发酵第 2、3 和 4 天的样品一贯产生足够品质以用于微阵列实验的 RNA。

[0382] 实施例 4:荧光探针构建和微阵列杂交

[0383] 橙色嗜热子囊菌样品的低 RNA 产量必需利用线性放大方法 (aRNA 扩增) 以产生足

够数量以用于微阵列杂交的 RNA。遵循制造商的说明,用氨基烯丙基 MessageAmp™ aRNA 试剂盒(Ambion, Inc., Austin, TX)进行 RNA 的扩增和荧光标记。

[0384] 简单的说,通过以 T7Oligo (dT)引发的反转录,从 2 μg 总 RNA 产生 cDNA 第一链。随后用 T7 启动子引物合成 cDNA 第二条链。纯化双链 cDNA,并加到体外转录反应中以产生 aRNA 的多拷贝。氨基烯丙基-dUTP 在体外转录中掺入 aRNA,便于随后荧光花青染料的标记。纯化所得 aRNA,并根据制造商的说明,通过将 Cy3 和 Cy5 荧光团(CyeDye Post 标记反应染料, Amersham Pharmacia Biotech, Arlington Heights, IL)直接偶联到 aRNA 上的经氨基烯丙基修饰的 UTP 残基上来进行标记。

[0385] 将荧光探针合并,纯化,并在真空下干燥,然后将其重悬浮于 15.5 μl 水中并加入:3.6 μl 120X SSC、2.5 μl 1250mM HEPES(pH7.0)、1.8 μl poly-dA (500 μg/ml)、和 0.54 μl 110%SDS。杂交前,用 0.22 μm 滤器过滤溶液,加热至 100°C 2 分钟并冷却至室温。

[0386] 在盖玻片下,将探针(Cy3- 和 Cy5- 标记的 aRNA 各 5 μg)施加到微阵列上,并置于湿润的小室中。于 63°C 杂交过夜(15-16 小时)。扫描前,将阵列用含 0.03%SDS 的 1X SSC、0.2X SSC、和 0.05X SSC 连续顺序洗涤,并在桌面式离心机(Sorvall R7, RTH-250 转子; Asheville, NC)中以 500rpm 离心 2 分钟以除去多余的液体。

[0387] 用 Axon GenePix® 4000B 扫描仪(Axon Instruments, Inc., Union City, CA)对微阵列载玻片进行成像。然后根据制造商的说明用 GenePix® Pro5.0 软件进行影像和数据分析。用 GenePix® 软件测量微阵列斑点的荧光强度值,并在减去默认背景后计算每个斑点的 Cy5 与 Cy3 强度的比值。挑选重复阵列中 Cy5/Cy3 的强度比为 2.0 或更大的斑点用于 DNA 序列分析。

[0388] 微阵列分析后,从保存于 -80°C 的甘油细胞悬液中挑选并分离感兴趣的克隆。通过碱裂解法制备用于测序的 DNA。用位于 pCR® 4Blunt-TOPO 载体的多接头侧翼的 T7 和 M13 反向引物(MWG Biotech, Inc., High Point, NC)对每个克隆进行测序。用全新的序列设计引物,以延长已知序列,随后通过此引物步移策略得到全长克隆。首先在 DNA 水平上利用 blast(n) 算法(Altschul 等人,1990, Journal of Molecular Biology 215:403-410)寻找同一性。随后将 DNA 序列在六个可能的读码框中的假定翻译物与公众数据库(例如 SwissProt、Swall、Trembl、Genpept 和 Genesep)条目利用 fasty (Pearson 等人,1997, Genomics 46:24-36)和 tblast(x) (Altschul 等人,1990,同上)进行对比。

[0389] 实施例 5:pAllo2 表达载体的构建

[0390] 表达载体 pAllo1 是通过修饰 pBANE6 (美国专利 6,461,837)而构建的,它包含 NA2-tpi 启动子、黑曲霉淀粉葡萄糖苷酶终止子序列(AMG 终止子)、和构巢曲霉乙酰胺酶基因(amdS)。pBANE6 的修饰是如下进行的,即首先通过定点诱变从 amdS 选择标记除去位于 2051、2722、和 3397bp 位置的三个 NcoI 限制位点。所有改变设计为“沉默的”,使得 amdS 基因的实际蛋白质序列不改变。三个位点的清除是利用 GeneEditor 定点诱变试剂盒(Promega, Madison, WI)根据制造商的说明使用下面的引物(标有下划线的核苷酸代表改变的碱基)同时进行的:

[0391] AMDS3NcoMut (2050):

[0392] 5' -GTGCCCCATGATACGCCTCCGG-3' (SEQ ID NO:3)

[0393] AMDS2NcoMut (2721) :

[0394] 5' -GAGTCGTATTTCCAAGGCTCCTGACC-3' (SEQ ID NO:4)

[0395] AMDS1NcoMut (3396) :

[0396] 5' -GGAGGCCATGAAGTGGACCAACGG-3' (SEQ ID NO:5)

[0397] 然后将包含所有三种期望序列改变的质粒利用 QuickChange 诱变试剂盒 (Stratagene, La Jolla, CA) 进行定点诱变, 以消除位于第 1643 位的 AMG 终止子末端的 NcoI 限制位点。将下列引物 (标有下划线的核苷酸代表改变的碱基) 用于诱变 :

[0398] 用于诱变黑曲霉淀粉葡萄糖苷酶 (AMG) 终止子序列的上游引物 :

[0399] 5' -CACCGTGAAAGCCATGCTCTTTCCTTCGTGTAGAAGACCAGACAG-3' (SEQ ID NO:6)

[0400] 用于诱变黑曲霉淀粉葡萄糖苷酶 (AMG) 终止子序列的下游引物 :

[0401] 5' -CTGGTCTTCTACACGAAGGAAAGAGCATGGCTTTCACGGTGTCTG-3' (SEQ ID NO:7)

[0402] 修饰 pBAnE6 的最后一步是利用 QuickChange 诱变试剂盒和下列引物 (标有下划线的核苷酸代表改变的碱基) 在多接头始端插入一个新的 NcoI 限制位点, 从而产生 pAILo1 (图 2)。

[0403] 用于诱变黑曲霉淀粉酶启动子 (NA2-tpi) 的上游引物 :

[0404] 5' -CTATATACACAACTGGATTTACCATGGGCCCGCGCCGCAGATC-3' (SEQ ID NO:8)

[0405] 用于诱变黑曲霉淀粉酶启动子 (NA2-tpi) 的下游引物 :

[0406] 5' -GATCTGCGGCCCGGGCCCATGGTAAATCCAGTTGTGTATATAG-3' (SEQ ID NO:9)

[0407] 用构巢曲霉 pyrG 基因替换 pAILo1 的 amdS 基因。将质粒 pBAnE10 (图 3) 用作 pyrG 基因的来源。分析 pBAnE10 的序列表明 pyrG 标记包含在 NsiI 限制片段中, 并且不含 NcoI 或 PacI 限制位点。因为 amdS 的也为 NsiI 限制位点所侧翼包夹, 所以用于转换选择标记的策略是进行 NsiI 限制片段的简单替换。用限制酶 NsiI 消化来自 pAILo1 和 pBAnE10 的质粒 DNA, 并利用标准程序通过琼脂糖凝胶电泳纯化产物。将源自 pBAnE10 的含有 pyrG 基因的 NsiI 片段连接到 pAILo1 骨架中以替换包含 amdS 基因的原始 NsiI DNA 片段。通过限制消化分析重组克隆以测定它们是否含有正确的插入片段和方向是否正确。选择以反时针方向转录 pyrG 基因的克隆。新的质粒称为 pAILo2 (图 4)。

[0408] 实施例 6 : 米曲霉表达载体的构建

[0409] 设计如下的两条合成寡核苷酸引物, 从编号 3 基因组克隆 PCR 扩增编码假定 GH61A 家族的橙色嗜热子囊菌基因。无需限制消化和连接, 使用 InFusion 克隆试剂盒 (BD Biosciences, Palo Alto, CA) 将片段直接克隆到表达载体 pAILo2 中。

[0410] 正向引物 : 5' -CACAACTGGATTTACC **ATGTCCTTTTCCAAG** -3' (SEQ ID NO:10)

[0411] 反向引物 : 5' -AGTCACCTCTAGTTA **TTAACCAGTATACAG** -3' (SEQ ID NO:11)

[0412] 粗体字母代表编码序列。剩余序列与 pAILo2 的插入位点是同源的。

[0413] 将上述每种引物各 200pmol 用于终体积 50 μ l 的 PCR 反应, 其组成是代表含有 GH61A 编码序列的原始橙色嗜热子囊菌编号 3 基因组克隆的 DNA、10mM KCl、20mM Tris-HCl pH8.8、10mM (NH₄)₂SO₄、2mM MgSO₄、0.1% Triton X-100、200 μ l 每种脱氧核糖核苷酸 (dATP、dTTP、dGTP 和 dCTP)、和 2 个单位 Vent[®] DNA 聚合酶 (所有 PCR 相关的酶和试剂都来自 New England Biolabs, Inc., Beverly, MA)。扩增条件为 1 个循环的 95°C 1 分钟 ; 30 个循环每个

94°C 30 秒、50°C 30 秒、72°C 1 分钟；及最后一个循环 72°C 7 分钟。然后将加热块转到 4°C 浸泡 (soak) 循环。

[0414] 然后用 InFusion 克隆试剂盒根据制造商的说明将片段克隆到 pA1Lo2 表达载体中。用 NcoI 和 PacI 在制造商推荐的条件下消化载体。用 MinElute™Reaction Cleanup Kit (QIAGEN, Valencia, CA) 纯化片段。将基因片段和线性化载体在反应中连接起来,产生表达质粒 pDZA2(图 5),其中 GH61A 家族基因的转录处于 NA2-tpi 启动子的控制之下。质粒 pDZA2 缺失 pA1Lo2 的 152bp。连接反应含有大约 100ng 经 NcoI 和 PacI 消化的 pA1Lo2 及 100ng 纯化的橙色嗜热子囊菌 GH61A PCR 产物。连接条件按照 InFusion 克隆试剂盒制造商的方案。用连接产物转化大肠杆菌 XL1-Blue 亚克隆级感受态细胞(Stratagene, La Jolla, CA)。通过对由大肠杆菌纯化的质粒的 GH61A 编码序列进行 DNA 测序来确认构建体的身份。将一个含有重组质粒的克隆命名为大肠杆菌 pDZA2-7。

[0415] 实施例 7:编码具有增强分解纤维素活性的 GH61 家族多肽的橙色嗜热子囊菌基因组序列的鉴定

[0416] 使用 Applied Biosystems3700 型自动 DNA 测序仪利用 3.1 版 Big Dye 终止物化学法和 dGTP 化学法(Applied Biosystems, Inc., Foster City, CA) 和引物步移策略对橙色嗜热子囊菌 GH61A 基因组克隆 15 进行 DNA 测序。利用 Vector NTI8 套装软件包(Informax, Inc., Frederick, MD) 的 ContigExpress 组件装配和比较核苷酸序列。

[0417] 图 1 显示了具有增强分解纤维素活性的橙色嗜热子囊菌多肽的核苷酸序列(SEQ ID NO:1) 和推导的氨基酸序列(SEQ ID NO:2)。该编码序列编码 250 个氨基酸的蛋白质。该编码序列有 799bp, 包含终止密码子, 中间有 46bp 的单一内含子。编码区有 48.7% 的 G+C。利用 SignalP 程序(Nielsen 等人, 1997, Protein Engineering10:1-6), 预测了 22 个氨基酸残基的信号肽, 表明成熟多肽含有 228 个氨基酸。

[0418] 利用 blastp 算法(Higgins, 1989, 同上) 使用 Paracel BioView Workbench 软件(Paracel, Pasadena, CA) 及 blosum62 矩阵测定了氨基酸序列的比较性对比。成对比对参数采用的是缺口罚分, 存在(existence):11 和延伸(extension):1。该对比表明编码具有增强分解纤维素活性的 GH61 家族多肽的橙色嗜热子囊菌基因的推导氨基酸序列与来自构巢曲霉推测蛋白 AN1041. 2(登录号 EAA65609)、构巢曲霉推测蛋白 AN7891. 2(登录号 EAA59545)、和构巢曲霉推测蛋白 AN9524. 2(登录号 EAA66740) 的 61 家族蛋白质推导氨基酸序列分别享有 67%、63%、和 58% 的同源性。

[0419] 含有质粒 pDZA2-7 的大肠杆菌 XL1-Blue (Stratagene, La Jolla, CA) 保藏于农业研究机构专利培养物保藏中心, 北方区域研究中心(Agricultural Research Service Patent Culture Collection, Northern Regional Research Center, 1815 University Street, Peoria, Illinois, 61604), 保藏编号 NRRL B-30704, 保藏日期 2004 年 1 月 30 日。

[0420] 实施例 8:编码具有增强分解纤维素活性的 GH61A 家族多肽的橙色嗜热子囊菌基因在米曲霉 JaL250 中的表达

[0421] 根据 Christensen 等人, 1988, Bio/Technology6:1419-1422 的方法制备米曲霉 JaL250 原生质体。用大约 1.5 μg pDZA2 转化米曲霉 JAL250。以质粒 pA1Lo2 作为对照。

[0422] 用 pDZA2 转化米曲霉 JaL250 获得大约 11 个转化体。将 10 个转化体分离到单独的 PAD 板上。

[0423] 用 5ml 0.01% 吐温 20 洗涤所有转化体的铺满的 PDA 板, 分别接种 125ml 玻璃摇瓶中的 25ml MDU2BP 培养基, 并于 34°C 以 250rpm 进行培养。保温 6 天后, 使用 8-16% Tris- 甘氨酸 SDS-PAGE 凝胶 (Invitrogen, Inc., Carlsbad, CA) 根据制造商的说明分析取自每份培养物的 5 μ l 上清液。SDS-PAGE 分析显示, 10 个转化体中有 9 个转化体的 GH61A 带以略微大于 25kDa 的表观分子量进行迁移。

[0424] 实施例 9: 具有增强分解纤维素活性的橙色嗜热子囊菌 GH61A 的鉴定

[0425] 在美国能源部国家再生能源实验室 (the U. S. Department of Energy National Renewable Energy Laboratory, NREL) 用稀硫酸对玉米秸秆进行预处理。预处理条件如下: 于 190°C, 0.048g 硫酸 /g 干生物量, 25%w/w 干固体, 大约 1 分钟。根据 NREL, 在经过预处理的玉米秸秆 (PCS) 中不溶于水的固体含有 52% 纤维素、3.6% 半纤维素和 29.8% 木质素。纤维素和半纤维素是利用 NREL 标准分析程序 #002 通过两级硫酸水解及后续高效液相色谱对糖类的分析而测定的。木质素是利用 NREL 标准分析程序 #003 在用硫酸水解纤维素和半纤维素成分后通过重量测定而测定的。在酶促水解前, 用大量的 DDI 水洗涤 PCS; 水洗后的 PCS 的干重为 20.6%。

[0426] 如实施例 8 所述在米曲霉中表达橙色嗜热子囊菌 GH61A 多肽, 将培养液以 9500x g 离心, 然后利用装备有 PM10 膜 (Millipore, Billerica, MA) 的 Amicon 搅拌器浓缩上清液, 并利用 Econo-Pac10DG 柱 (BioRad Laboratories, Hercules, CA) 脱盐。

[0427] 用 1.1ml Immunoware 微量试管 (Pierce, Rockford, IL) 进行 PCS (每 ml 50mM pH5.0 乙酸钠缓冲液 10mg) 水解, 总反应体积 1.0ml。对橙色嗜热子囊菌 GH61A 多肽测试其增强纤维素酶制备物水解能力的能力, 所述纤维素酶制备物是从 Novozymes A/S (Bagsværd, 丹麦) 获得的、由表达米曲霉 β - 葡糖苷酶的瑞氏木霉 (W002/095014) 发酵得到的, 下称 Tr/AoBG。PCS 水解的进行是采用每克 PCS 2.5mg Tr/AoBG, 且每克 PCS 补充 0.2mg 橙色嗜热子囊菌 GH61A 多肽。于 50°C (TS Autoflow CO₂ 夹套保温箱) 进行 PCS 水解。进行两份同样的反应, 而且在水解期间采集小样。将每份水解产物的 20 μ l 小样与 180 μ l 0.11M NaOH (终止试剂) 混和, 使 PCS 水解反应终止。对每份样品进行适当的连续稀释, 利用对羟基苯甲酸酰肼 (PHBAH, Sigma Chemical Co., St. Louis, MO) 测定法来测定还原糖含量, 该测定法经修改适应 96 孔微量滴定板格式, 如下所述。简单的说, 将适当稀释样品的 90 μ l 小样置于 96 孔锥底微量滴定板中。向每个孔中加入 60 μ l 溶于 2%NaOH 的 1.5% (w/v) PHBAH 以起动反应。将平板于 95°C 加热 10 分钟, 不加盖。使平板冷却至室温 (RT), 并向每个孔中加入 50 μ l 蒸馏 H₂O。从每个孔中取出 100 μ l 小样转移到平底 96 孔板中, 并用 SpectraMax 微板读数仪 (Molecular Devices, Sunnyvale, CA) 测量 A_{410nm} 的吸光率。用葡萄糖标准品 (0.1-0.0125mg/ml, 用 0.4% 氢氧化钠稀释) 制备标准曲线, 将得到的 A_{410nm} 值转换成葡萄糖的相当量。用所得的相当量计算每个反应 PCS 纤维素转化的百分比。

[0428] 利用下面的方程式计算纤维素转化成还原糖的程度 (转化, %):

[0429] 转化 (%) = $RS_{(mg/ml)} * 100 * 162 / (纤维素_{(mg/ml)} * 180) =$

[0430] $= RS_{(mg/ml)} * 100 / (纤维素_{(mg/ml)} * 1.111)$

[0431] 在此方程式中, RS 是溶液中按葡萄糖相当量 (mg/ml) 计量的还原糖浓度, 而系数 1.111 反映了纤维素转化成葡萄糖时的重量增量。

[0432] 表 1 概括给了通过单独的 Tr/AoBG (2.5mg/g PCS) 或者补充橙色嗜热子囊菌 GH61A

多肽(每克 PCS0.2mg)的纤维素转化。

[0433] 表 1:由单独的 Tr/AoBG 或补充橙色嗜热子囊菌 GH61A 多肽的 Tr/AoBG 于 50℃ pH5.0 保温 115 小时的纤维素转化

[0434]

试验编号	名称	加载量 mg/g PCS	115 小时转化 %
1	Tr/AoBG	2.5	74.6
2	橙色嗜热子囊菌 GH61A	0.2	<0.7%
3	Tr/AoBG + 橙色嗜热子囊菌 GH61A	2.5+0.2	83.8
4	Tr/AoBG	3.5	82.4

[0435] 表 1 显示了橙色嗜热子囊菌 GH61A 多肽增强了 Tr/AoBG 对 PCS 的活性。在 115 小时后,橙色嗜热子囊菌 GH61A 多肽自身(每克 PCS0.2mg)仅产生低于 0.7% 的纤维素转化。补充 0.2mg 橙色热质囊菌 GH61A 多肽的 2.5mg Tr/AoBG 所产生的纤维素转化要高于 3.5mg Tr/AoBG,表明橙色嗜热子囊菌 GH61A 多肽增强了 Tr/AoBG 对 PCS 的活性,且 Tr/AoBG 和橙色嗜热子囊菌 GH61A 多肽之间有协同效应。

[0436] 生物材料的保藏

[0437] 下列生物材料已按照布达佩斯条约保藏于农业研究机构专利培养物保藏中心,北方区域研究中心(Agricultural Research Service Patent Culture Collection, Northern Regional Research Center, 1815 University Street, Peoria, Illinois, 61604),并给出下列登记号:

[0438] 保藏物 保藏号 保藏日期

[0439] 大肠杆菌菌株 pDZA2-7 NRRL B-30704 2004-01-30

(*Escherichia coli*)

[0440] 该菌株已经在如下条件下进行保藏,即根据 37 C.F.R. § 1.14 和 35 U.S.C § 122,确保在本专利申请的审查过程中由授权的专利与商标官员所指定的人员能够获得该培养物。所述保藏物是所述保藏菌株的基本上纯的培养物。在提交本申请副本或其后代的国家中,根据外国专利法的要求可以提供上述保藏物。但是,应该理解的是可以获得保藏物并不构成这样的许可,即实施本发明侵害了由政府授予的专利权。

[0441] 本文描述并要求的发明并不限于本文所公开的具体实施方案的范围,因为这些实施方案意图在于举例说明本发明的几个方面。任何等同的实施方案意图包括在本发明的范围内。当然,根据上文描述,除本文所显示和描述的之外,对本发明的各种修改对本领域技术人员而言应是显而易见的。这些修改也应包括在所附权利要求的范围内。一旦出现冲突,将以包括定义在内的本公开为准。

[0442] 本文引用了多份参考文献,完整引入其公开书作为参考。

[0443]

申请人或代理人档案号	国际申请号 PCT/US2005/003802
------------	-------------------------

[0444] 关于微生物保藏的说明

[0445] (细则 13 之二)

[0446]

A. 对说明书第 <u>73</u> 页, 第 <u>1-5</u> 行所述的微生物的说明。	
B. 保藏事项 其它保藏在补充页中 <input type="checkbox"/>	
农业研究机构保藏中心 保藏单位名称 Agricultural Research Service Patent Culture Collection (NRRL)	
保藏单位地址 (包括邮政编码和国名) 美国 伊利诺斯州 皮奥里亚市 大学街 1815 号 北方区域研究中心 邮政编码: 61604 Northern Regional Research Center, 1815 University Street, Peoria, IL 61604, US	
保藏日期 2004 年 1 月 30 日	保藏编号 B-30704
C. 补充说明 (必要时) 本栏内容有补充页 <input type="checkbox"/>	
在专利申请期间, 关于欧洲和澳大利亚专利申请的指定, 所保藏的微生物的样本只向由要求此样本的人所指定的独立鉴定人提供 (Rule 28(4) EPC/Regulation 3.25 of Australia Statutory Rule 1991 No.71)。	
D. 本说明是为下列指定国作的 (如果说明不是为所有指定国而作的)	
E. 补充说明 (必要时)	
下列说明将随后向国际局提供 (写出说明的类别, 例如: “保藏的编号”)	

由受理局填写
<input type="checkbox"/> 本页已经和国际申请一起收到
受权官员

由国际局填写
<input type="checkbox"/> 国际局收到本页日期:
受权官员

[0447] PCT/R0/134 表 (1992 年 7 月)

[0001]

序列表

- <110> 诺维信股份有限公司 (Novozymes Biotech, Inc.)
- <120> 具有增强分解纤维素活性的多肽和编码所述多肽的多核苷酸
- <130> 10602.204-WO
- <150> 60/542, 614
- <151> 2004-02-06
- <160> 11
- <170> PatentIn version 3.2
- <210> 1
- <211> 799
- <212> DNA
- <213> 橙色热子囊菌 (Thermoascus aurantiacus)

```

<400> 1
atgtcctitt ccaagataat fgctactgcc ggcgftcttg cctctgcttc tctagggct    60
ggccatggct tegtccagaa catcgtgatt gatggtaaaa agtatgtcat tgcaagacgc    120
acataagcgg caacagctga caatcgacag itatggcggg tatctagtga accagiatcc    180
atacatgtcc aatcctccag aggtcatcgc ctggctctact acggcaactg atcttggatt    240
tgtggacggt actggafacc aaaccccaga tatcactcgc cataggggcg ccaagecctgg    300
agccctgact gctccagtct ctccaggagg aactgttgag ctccaatgga ctccatggcc    360
tgaitctcac catggcccag itatcaacta ccttgcctcg tgcaatggig attgttccac    420
tgtggataag acccaattag aattcttcaa aattgcccag agcggctca tcaatgatga    480
caatctctct gggatctggg ctfcagacaa tctgatagca gccacaaca gctggactgt    540
caccattcca accacaattg caccfgaaa ctatgtctcg aggcattgaga ttattgctct    600
tactcagct cagaaccagg atggtgccca gaactatecc cagtgcatea atctgcaggt    660
cacfgaggt ggttctgata accctgctgg aactcttggg acggcactct accaegatac    720
cgaicctgga altcigatea acaiclatca gaaacttcc agctatatea tccctggtec    780
tcctctgtat actggttaa    799
    
```

- <210> 2
- <211> 250
- <212> PRT
- <213> 橙色热子囊菌 (Thermoascus aurantiacus)

```

<400> 2
Met Ser Phe Ser Lys Ile Ile Ala Thr Ala Gly Val Leu Ala Ser Ala
1           5           10           15
    
```

[0002]

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

<210> 3
 <211> 22
 <212> DNA
 <213> 黑曲霉(Aspergillus niger)

<400> 3
 gtgccccatg atacgcctcc gg

22

<210> 4
 <211> 26
 <212> DNA

[0003]

<213> 黑曲霉 (<i>Aspergillus niger</i>)	
<400> 4	
gagtcgtatt tccaaggctc cfgacc	26
<210> 5	
<211> 24	
<212> DNA	
<213> 黑曲霉 (<i>Aspergillus niger</i>)	
<400> 5	
ggaggccatg aagtggaacca acgg	24
<210> 6	
<211> 45	
<212> DNA	
<213> 黑曲霉 (<i>Aspergillus niger</i>)	
<400> 6	
caccctgaaa gccatgcctc ttccctcgtg tagaagacca gacag	45
<210> 7	
<211> 45	
<212> DNA	
<213> 黑曲霉 (<i>Aspergillus niger</i>)	
<400> 7	
ctggtcttct acacgaagga aagagcatgg ctctcaggt gctcg	45
<210> 8	
<211> 44	
<212> DNA	
<213> 黑曲霉 (<i>Aspergillus niger</i>)	
<400> 8	
ctatatacac aactggattt accatgggcc cgcggcgcga gatc	44
<210> 9	
<211> 44	
<212> DNA	
<213> 黑曲霉 (<i>Aspergillus niger</i>)	
<400> 9	
gatctgcggc cgcgggccca tggtaaatec agttigtat atag	44
<210> 10	
<211> 31	
<212> DNA	
<213> 橙色热子囊菌 (<i>Thermoascus aurantiacus</i>)	
<400> 10	
cacaactgga tttaaccatgt ccttttccaa g	31

[0004]

<210> 11
<211> 30
<212> DNA
<213> 橙色热子囊菌 (*Thermoascus aurantiacus*)

<400> 11
agtcacctct agttattaac cagtatacag 30

```

      M S F S K I I A T A G V L A S A S
1  ATGTCCTTTT CCAAGATAAT TGCTACTGCC GGCCTTCTTG CCTCTGCTTC
   L V A G H G F V Q N I V I D G K
51  TCTAGTGGCT GGCCATGGCT TCGTTCAGAA CATCGTGATT GATGGTAAAA
   K Y
101 AGTATGTTCAT TGCAAGACGC ACATAAGCCG CAACAGCTCA CAATCCACAC
   Y G G Y L V N Q Y P Y M S N P P
151 TTATGGCGGG TATCTAGTGA ACCAGTATCC ATACATGTCC AATCCTCCAG
   E V I A W S T T A T D L G F V D G
201 AGGTCATCGC CTGGTCTACT ACGGCAACTG ATCTTGGATT TGTGGACGGT
   T G Y Q T P D I I C H R G A K P G
251 ACTGGATACC AAACCCGAGA TATCATCTGC CATAGGGGCG CCAAGCCTGG
   A L T A P V S P G G T V E L Q W
301 AGCCCTGACT GCTCCAGTCT CTCCAGGAGG AACTGTTGAG CTTCAATGGA
   T P W P D S H H G P V I N Y L A P
351 CTCCATGGCC TGATTCTCAC CATGGCCCAG TTATCAACTA CCTTGCTCCG
   C N G D C S T V D K T Q L E F F K
401 TGCAATGGTG ATTGTTCCAC TGTGGATAAG ACCCAATTAG AATTCTTCAA
   I A E S G L I N D D N P P G I W
451 AATTGCCGAG AGCGGTCTCA TCAATGATGA CAATCCTCCT GGGATCTGGG
   A S D N L I A A N N S W T V T I P
501 CTTCAGACAA TCTGATAGCA GCCAACAAACA GCTGGACTGT CACCATTCCA
   T T I A P G N Y V L R H E I I A L
551 ACCACAATTG CACCTGGAAA CTATGTTCTG AGGCATGAGA TTATTGCTCT
   H S A Q N Q D G A Q N Y P Q C I
601 TCACTCAGCT CAGAACCAGG ATGGTGCCCA GAACTATCCC CAGTGCATCA
   N L Q V T G G G S D N P A G T L G
651 ATCTGCAGGT CACTGGAGGT GGTTCGTGATA ACCCTGCTGG AACTCTTGGA
   T A L Y H D T D P G I L I N I Y Q
701 ACGGCACTCT ACCACGATAC CGATCCTGGA ATTCTGATCA ACATCTATCA
   K L S S Y I I P G P P L Y T G *
751 GAAACTTTCC AGCTATATCA TCCCTGGTCC TCCTCTGTAT ACTGGTTAA

```

图 1

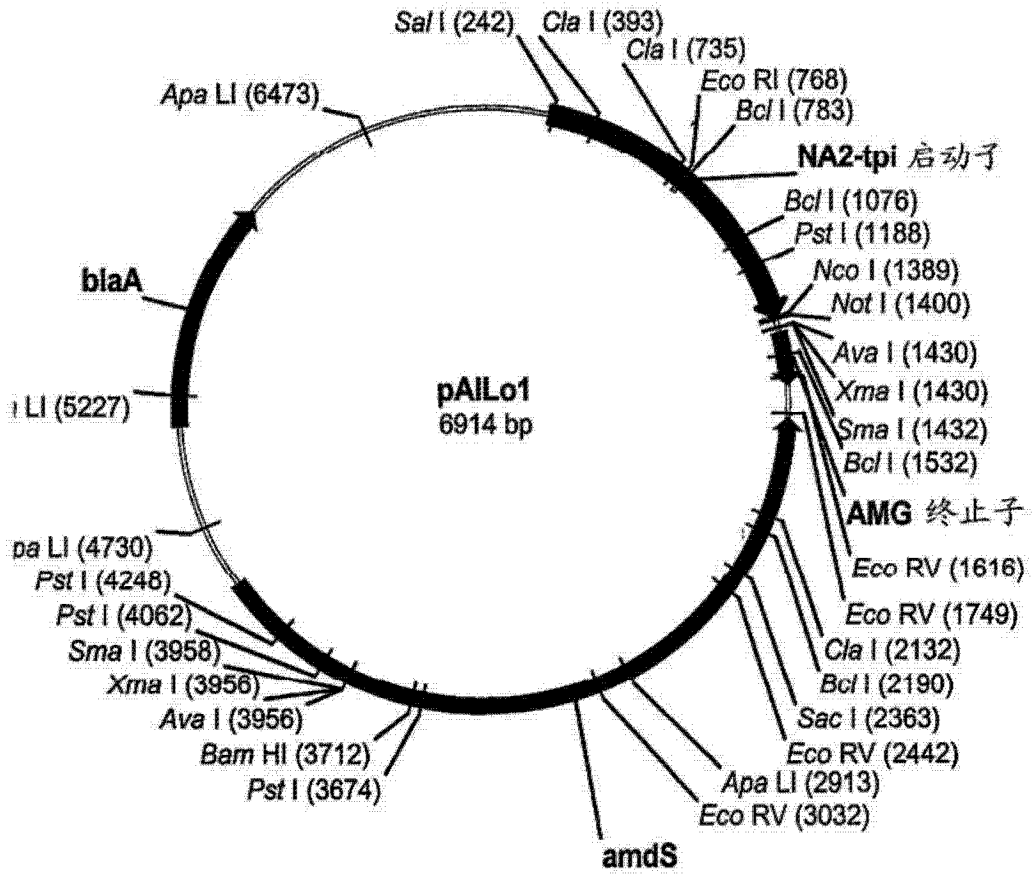


图 2

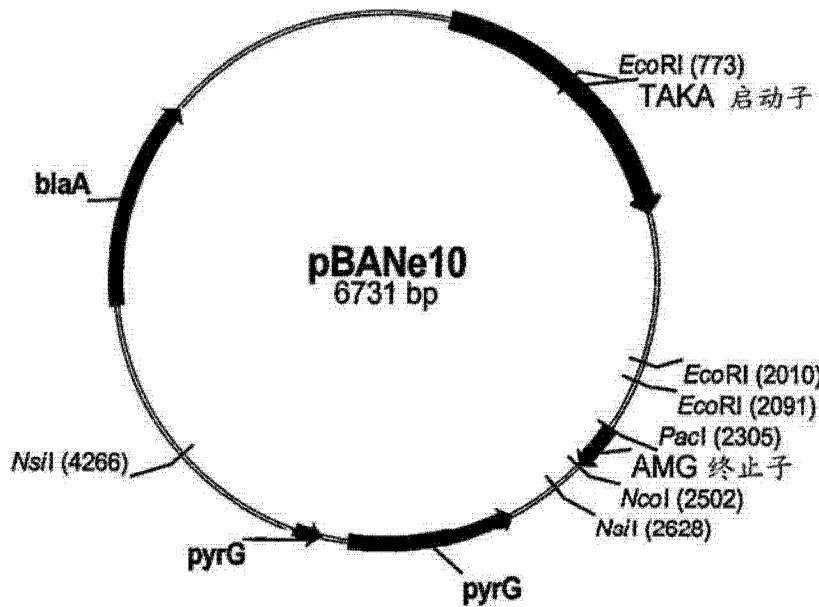


图 3

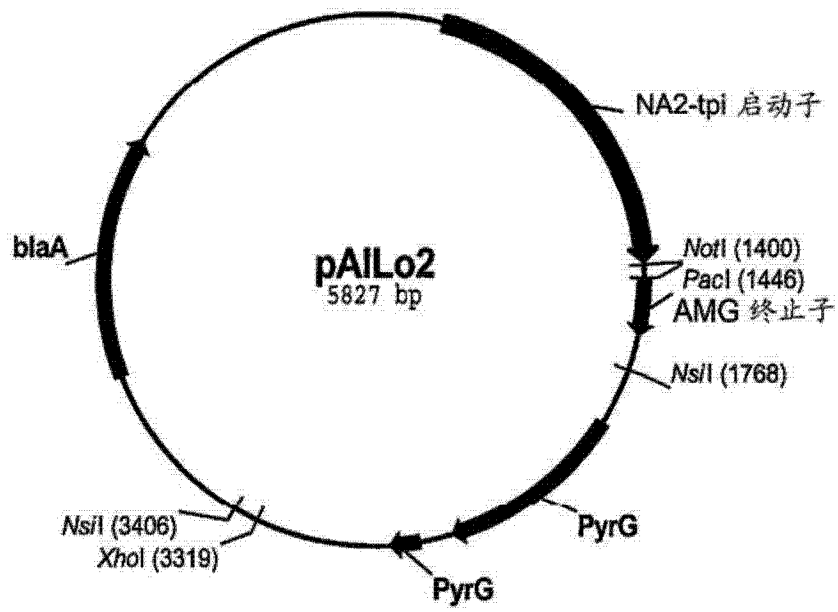


图 4

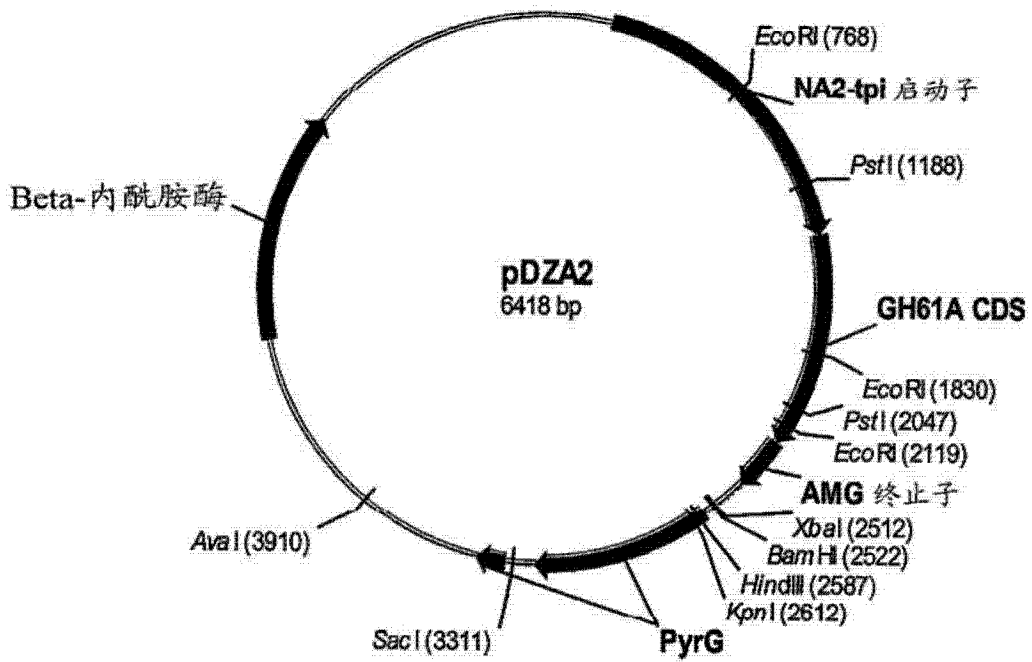


图 5