



(12) 发明专利申请

(10) 申请公布号 CN 104540955 A

(43) 申请公布日 2015. 04. 22

(21) 申请号 201380038353. 4

(51) Int. Cl.

(22) 申请日 2013. 07. 18

C12P 7/06(2006. 01)

(30) 优先权数据

C12N 9/14(2006. 01)

61/673, 556 2012. 07. 19 US

(85) PCT国际申请进入国家阶段日

2015. 01. 19

(86) PCT国际申请的申请数据

PCT/US2013/051084 2013. 07. 18

(87) PCT国际申请的公布数据

W02014/018368 EN 2014. 01. 30

(71) 申请人 诺维信公司

地址 丹麦鲍斯韦

(72) 发明人 Y·陈 J·刘 T·格林

M·V·史蒂文斯 H·许

(74) 专利代理机构 北京坤瑞律师事务所 11494

代理人 封新琴

权利要求书3页 说明书70页

序列表94页 附图5页

(54) 发明名称

用于增加纤维素材料的酶水解的方法

(57) 摘要

本发明涉及用于增加预处理的纤维素材料的水解的方法,这些方法包括在适于水解该预处理的纤维素材料的条件下,使该预处理的纤维素材料经受一种纤维素分解酶组合物;一种具有纤维素分解增强活性的多肽;一种过氧化物酶;以及一种非离子表面活性剂和/或一种阳离子表面活性剂。本发明还涉及用于生产发酵产品的工艺,这些工艺包括本发明的水解步骤并且涉及一种适合在本发明的方法中使用的组合物。

1. 一种用于水解预处理的纤维素材料的方法,该方法包括使该预处理的纤维素材料经受:

- (a) 一种纤维素分解酶组合物;
 - (b) 一种多肽,该多肽具有纤维素分解增强活性;
 - (c) 一种过氧化物酶;以及
 - (d) 一种非离子表面活性剂和/或一种阳离子表面活性剂,
- 在适于水解该预处理的木质纤维素材料的条件下。

2. 如权利要求1所述的方法,其中该纤维素分解酶组合物来源于拉克淖金孢子菌、特异腐质霉、嗜热毁丝霉、或里氏木霉。

3. 如权利要求1或2所述的方法,其中具有纤维素分解增强活性的该多肽是一种GH61多肽,例如一种来源于嗜热子囊菌属的多肽,例如金黄色嗜热子囊菌菌株,例如描述于在此的SEQ ID NO:14中的多肽;或一种来源于青霉属的菌株的多肽,例如埃默森青霉菌菌株,例如披露于在此的SEQ ID NO:72中的多肽。

4. 如权利要求1-3中任一项所述的方法,其中该纤维素分解酶组合物进一步包括一种或多种(若干种)选自下组的酶,该组由以下各项组成:半纤维素酶、酯酶、蛋白酶、以及漆酶。

5. 如权利要求1-4中任一项所述的方法,其中该过氧化物酶选自下组,该组包括过氧化物酶或过氧化物分解酶,包括但不限于以下各项:E.C. 1. 11. 1. 1NADH 过氧化物酶;E.C. 1. 11. 1. 2NADPH 过氧化物酶;E.C. 1. 11. 1. 3 脂肪酸过氧化物酶;E.C. 1. 11. 1. 5 细胞色素c 过氧化物酶;E.C. 1. 11. 1. 5;E.C. 1. 11. 1. 6 过氧化氢酶;E.C. 1. 11. 1. 7 过氧化物酶;E.C. 1. 11. 1. 8 碘化物过氧化物酶;E.C. 1. 11. 1. 9 谷胱甘肽过氧化物酶;E.C. 1. 11. 1. 10 氯化物过氧化物酶;E.C. 1. 11. 1. 11L- 抗坏血酸过氧化物酶;E.C. 1. 11. 1. 12 磷脂过氧化氢谷胱甘肽过氧化物酶;E.C. 1. 11. 1. 13 锰过氧化物酶;E.C. 1. 11. 1. 14 木质素过氧化物酶;E.C. 1. 11. 1. 15 过氧化物氧还蛋白;E.C. 1. 11. 1. 16 通用过氧化物酶;E.C. 1. 11. 1. B2 氯化物过氧化物酶;E.C. 1. 11. 1. B6 碘化物过氧化物酶(含钒);E.C. 1. 11. 1. B7 溴化物过氧化物酶;E.C. 1. 11. 1. B8 碘化物过氧化物酶。

6. 如权利要求1-5中任一项所述的方法,其中该过氧化物酶来源于一种微生物,例如真菌生物体,例如酵母或丝状真菌,或细菌;或植物。

7. 如权利要求1-6中任一项所述的方法,其中该过氧化物酶来源于鬼伞属的菌株,例如灰盖鬼伞菌株,例如示于在此的SEQ ID NO:71中的过氧化物酶,或一种与在此的SEQ ID NO:71具有至少60%、优选至少65%、更优选至少70%、更优选至少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%、并且甚至最优选至少96%、至少97%、至少98%、或至少99%序列一致性的过氧化物酶。

8. 如权利要求1-7中任一项所述的方法,其中该非离子表面活性剂是烷基或芳基:甘油醚、乙二醇醚、乙醇酰胺、磺酰胺、醇、酰胺、醇乙氧基化物、甘油酯、乙二醇酯、甘油酯和乙二醇酯的乙氧基化物、糖基烷基多糖苷、聚氧乙烯化脂肪酸、烷醇胺缩合物、烷醇酰胺、叔烷基乙二醇、聚氧乙烯化硫醇、羧酸酯、以及聚氧乙烯化聚氧丙烯乙二醇,例如EO/PO嵌段共聚物(EO是环氧乙烷,PO是环氧丙烷)、EO聚合物和共聚物、多胺、和聚乙烯吡咯烷酮。

9. 如权利要求1-8中任一项所述的方法,其中该非离子表面活性剂是一种直链伯或仲

或具有化学式 $\text{RO}(\text{CH}_2\text{CH}_2\text{O})_n\text{H}$ 的支链醇乙氧基化物, 其中 R 是烃链长度并且 n 是环氧乙烷的摩尔平均数, 例如其中 R 是范围从 C9 至 C16 的直链伯或支链仲烃链长度并且 n 范围从 6 至 13, 例如醇乙氧基化物, 其中 R 是直链 C9-C11 烃链长度, 并且 n 是 6。

10. 如权利要求 1-9 中任一项所述的方法, 其中该阳离子表面活性剂是一种伯、仲或叔胺, 例如二盐酸奥替尼啶; 烷基三甲基铵盐, 例如鲸蜡基三甲基溴化铵 (CTAB) (又名十六烷基三甲基溴化铵)、十六烷基三甲基氯化铵 (CTAC)、氯化十六烷基吡啶 (CPC)、苯扎氯铵 (BAC)、苜索氯铵 (BZT)、5-溴-5-硝基-1,3-二噁烷、二甲基双十八烷基氯化铵、双十八烷基二甲基溴化铵 (DODAB)。

11. 一种用于生产发酵产品的工艺, 该工艺包括

- (a) 水解如在权利要求 1-10 的任一项中定义的预处理的纤维素材料;
- (b) 用一种或多种 (若干种) 发酵微生物对该材料进行发酵, 以产生该发酵产品; 并且
- (c) 任选地从发酵中回收该发酵产品。

12. 如权利要求 11 所述的工艺, 其中该发酵微生物能够将己糖和 / 或戊糖发酵为一种所希望的发酵产品。

13. 如权利要求 11 或 12 所述的工艺, 其中该发酵产品是乙醇。

14. 一种组合物, 包括或其组成为:

- i) 一种多肽, 该多肽具有纤维素分解增强活性;
- ii) 一种过氧化物酶;
- iii) 一种非离子表面活性剂和 / 或一种阳离子表面活性剂。

15. 如权利要求 14 所述的组合物, 其中具有纤维素分解增强活性的该多肽与在此的 SEQ ID NO:14 或在此的 SEQ ID NO:72 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性。

16. 如权利要求 14 或 15 所述的组合物, 其中该过氧化物酶选自下组, 该组包括过氧化物酶或过氧化物分解酶, 包括但不限于以下各项: E. C. 1. 11. 1. 1NADH 过氧化物酶; E. C. 1. 11. 1. 2NADPH 过氧化物酶; E. C. 1. 11. 1. 3 脂肪酸过氧化物酶; E. C. 1. 11. 1. 5 细胞色素 c 过氧化物酶; E. C. 1. 11. 1. 5; E. C. 1. 11. 1. 6 过氧化氢酶; E. C. 1. 11. 1. 7 过氧化物酶; E. C. 1. 11. 1. 8 碘化物过氧化物酶; E. C. 1. 11. 1. 9 谷胱甘肽过氧化物酶; E. C. 1. 11. 1. 10 氯化物过氧化物酶; E. C. 1. 11. 1. 11L- 抗坏血酸过氧化物酶; E. C. 1. 11. 1. 12 磷脂过氧化氢谷胱甘肽过氧化物酶; E. C. 1. 11. 1. 13 锰过氧化物酶; E. C. 1. 11. 1. 14 木质素过氧化物酶; E. C. 1. 11. 1. 15 过氧化物氧还蛋白; E. C. 1. 11. 1. 16 通用过氧化物酶; E. C. 1. 11. 1. B2 氯化物过氧化物酶; E. C. 1. 11. 1. B6 碘化物过氧化物酶 (含钒); E. C. 1. 11. 1. B7 溴化物过氧化物酶; E. C. 1. 11. 1. B8 碘化物过氧化物酶。

17. 如权利要求 14-16 中任一项所述的组合物, 其中该过氧化物酶来源于鬼伞属的菌株, 例如灰盖鬼伞菌株, 例如示于在此的 SEQ ID NO:71 中的过氧化物酶, 其中该过氧化物酶是示于在此的 SEQ ID NO:71 中的过氧化物酶或一种与在此的 SEQ ID NO:71 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的过氧化物酶。

18. 如权利要求 14-17 中任一项所述的组合物, 其中该非离子表面活性剂是烷基或芳基: 甘油醚、乙二醇醚、乙醇酰胺、磺酰胺、醇、酰胺、醇乙氧基化物、甘油酯、乙二醇酯、甘油酯和乙二醇酯的乙氧基化物、糖基烷基多糖苷、聚氧乙烯化脂肪酸、烷醇胺缩合物、烷醇酰胺、叔炔属乙二醇、聚氧乙烯化硫醇、羧酸酯、以及聚氧乙烯化聚氧丙烯乙二醇, 例如 E0/P0 嵌段共聚物 (E0 是环氧乙烷, P0 是环氧丙烷)、E0 聚合物和共聚物、多胺、和聚乙烯吡咯烷酮。

19. 如权利要求 14-18 中任一项所述的组合物, 其中该非离子表面活性剂是一种直链伯或仲或具有化学式 $\text{RO}(\text{CH}_2\text{CH}_2\text{O})_n\text{H}$ 的支链醇乙氧基化物, 其中 R 是烃链长度并且 n 是环氧乙烷的摩尔平均数, 例如其中 R 是范围从 C9 至 C16 的直链伯或支链仲烃链长度并且 n 范围从 6 至 13, 例如醇乙氧基化物, 其中 R 是直链 C9-C11 烃链长度, 并且 n 是 6。

20. 如权利要求 14-19 中任一项所述的组合物, 其中该阳离子表面活性剂是一种伯、仲或叔胺, 例如二盐酸奥替尼啶; 烷基三甲基铵盐, 例如鲸蜡基三甲基溴化铵 (CTAB) (又名十六烷基三甲基溴化铵)、十六烷基三甲基氯化铵 (CTAC)、氯化十六烷基吡啶 (CPC)、苯扎氯铵 (BAC)、苄索氯铵 (BZT)、5- 溴 -5- 硝基 -1, 3- 二噁烷、二甲基双十八烷基氯化铵、双十八烷基二甲基溴化铵 (DODAB)。

用于增加纤维素材料的酶水解的方法

[0001] 对序列表的引用

[0002] 本申请含有一个序列表, 将其通过引用结合在此。

[0003] 发明背景

发明领域

[0004] 本发明涉及用酶组合物增加纤维素材料的水解的方法以及包括本发明的方法的工艺。本发明还涉及一种用于在本发明的方法或工艺中使用的共混物组合物。

[0005] 相关技术说明

[0006] 纤维素是单糖葡萄糖通过 β -1, 4- 键连接的一种聚合物。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、以及 β - 葡萄糖苷酶。内切葡聚糖酶在任意位置消化纤维素聚合物, 使其打开而被纤维二糖水解酶攻击。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一种水溶性 β -1, 4- 连接二聚体。 β - 葡萄糖苷酶将纤维二糖水解成葡萄糖。

[0007] WO 2005/067531 披露了一种用纤维素分解酶在至少一种选自下组的表面活性剂的存在下降解木质纤维素材料的方法, 该组由以下各项组成: 仲醇乙氧基化物、脂肪醇乙氧基化物、壬基酚乙氧基化物、十三烷基乙氧基化物以及聚氧乙烯醚。

[0008] WO 2010/080408 涉及通过在一种具有过氧化物酶活性的多肽的存在下用一种酶组合物处理纤维素材料而降解或转化所述纤维素材料的方法。

[0009] 本发明提供了使用纤维素分解酶来改进预处理的纤维素材料的水解的方法以及用于从水解产物中生产发酵产品的工艺。

[0010] 发明概述

[0011] 在此描述了用于降解 / 水解预处理的纤维素材料的方法, 这些方法包括使该预处理的纤维素材料经受:

[0012] - 一种纤维素分解酶组合物;

[0013] - 一种多肽, 该多肽具有纤维素分解增强活性;

[0014] - 一种过氧化物酶; 以及

[0015] - 一种非离子表面活性剂和 / 或一种阳离子表面活性剂,

[0016] 在适于水解该预处理的木质纤维素材料的条件下。

[0017] 可以使用本发明的方法将预处理的纤维素材料水解 / 糖化为可发酵糖。可以将可发酵糖转化为许多有用的所希望的物质, 例如燃料、可饮用乙醇和 / 或发酵产品 (例如, 酸、醇、酮、气体等)。

[0018] 该降解的 / 水解的预处理的纤维素材料可以是或可以含有以下糖, 这些糖可以用在用于生产糖浆 (例如, 高果糖玉米糖浆 (HFCS)) 和 / 或塑料 (例如, 聚乙烯、聚苯乙烯以及聚丙烯)、聚乳酸 (例如, 用于生产 PET) 的工艺中。

[0019] 本发明还涉及用于生产发酵产品的方法, 这些方法包括

[0020] (a) 根据本发明的方法水解一种预处理的纤维素材料;

[0021] (b) 用一种或多种(若干种)发酵微生物对该材料进行发酵,以产生该发酵产品;并且

[0022] (c) 任选地从发酵中回收该发酵产品。

[0023] 最后,本发明涉及包括或其组成为以下各项的组合物:

[0024] i) 具有纤维素分解增强活性的多肽;

[0025] ii) 一种过氧化物酶;

[0026] iii) 一种非离子表面活性剂和/或一种阳离子表面活性剂。

[0027] 在一个实施例中,该组合物还包括一种纤维素分解酶组合物。

[0028] 附图简要说明

[0029] 图1示出了CiP过氧化物酶与非离子表面活性剂之间的协同。

[0030] 图2示出了GH61a水平对表面活性剂与过氧化物酶协同的影响。

[0031] 图3示出了PeGH61a(埃默森青霉菌GH61多肽)与TaGH61a(金黄色嗜热子囊菌GH61多肽)的比较。

[0032] 图4示出了非离子表面活性剂与过氧化物酶之间的协同作用。

[0033] 图5示出了阳离子表面活性剂与过氧化物酶(HB:十六烷基三甲基溴化铵;BC:氯化十六烷基吡啶)之间的协同作用。

[0034] 图6示出了表面活性剂剂量对协同作用的影响。

[0035] 图7示出了不同纤维素分解酶组合物对协同作用的影响。

[0036] 图8示出了在不同木质纤维素材料上的CiP与表面活性剂之间的协同作用。

[0037] 图9示出了过氧化物酶(大豆过氧化物酶、王棕(royal palm)过氧化物酶、木质素过氧化物酶以及辣根过氧化物酶)与表面活性剂(李维本™(LEVAPON™))之间的协同作用。

[0038] 图10示出了过氧化物酶(大豆过氧化物酶、王棕过氧化物酶、木质素过氧化物酶以及辣根过氧化物酶)与表面活性剂(李维本™)之间的协同作用。

[0039] 定义

[0040] 过氧化物酶:在此定义了术语“过氧化物酶”,包括具有过氧化物酶活性的酶和过氧化物分解酶。

[0041] 过氧化物酶活性:在此将术语“过氧化物酶活性”定义为将过氧化物(例如,过氧化氢)转化为较少氧化的种类(例如,水)的酶活性。在此应该理解的是,具有过氧化物酶活性的多肽涵盖过氧化物分解酶(在下文定义)。

[0042] 过氧化物分解酶:在此将术语“过氧化物分解酶”定义为一种供体:催化还原底物($2e^- + R_2OOR'$) → 氧化底物 + $ROH + R'OH$ 反应的过氧化物氧化还原酶(E.C. 编号 1.11.1.x);例如催化苯酚 + H_2O_2 → 醌 + H_2O 反应的辣根过氧化物酶,和催化 $H_2O_2 + H_2O_2$ → $O_2 + 2H_2O$ 反应的过氧化氢酶。除过氧化氢之外,其他过氧化物也可以被这些酶分解。

[0043] 纤维素分解活性:在此将术语“纤维素分解活性”定义为水解一种纤维素材料的生物活性。用于测量纤维素分解活性的两种基本方法包括:(1) 测量总纤维素分解活性,以及(2) 测量单独的纤维素分解活性(内切葡聚糖酶、纤维二糖水解酶、以及 β -葡萄糖苷酶),如在张(Zhang)等人,2006,纤维素酶改进的展望:筛选和选择策略(Outlook for cellulase improvement: Screening and selection strategies),生物技术进展

(Biotechnology Advances)24:452-481 中所综述的。通常使用不溶性底物,包括沃特曼(Whatman) No 1 滤纸、微晶纤维素、细菌纤维素、藻类纤维素、棉花、预处理的木质纤维素等,测量总纤维素分解活性。最常用的总纤维素分解活性测定是使用沃特曼 No 1 滤纸作为底物的滤纸测定。该测定是由国际纯粹与应用化学联合会(IUPAC)(高斯(Ghose),1987,纤维素酶活性的测量(Measurement of cellulase activities),纯粹与应用化学(Pure Appl. Chem.)59:257-68)确立的。

[0044] 出于本发明的目的,通过测量在以下条件下由一种或多种纤维素分解酶进行的纤维素材料水解的增加来测定纤维素分解活性:1-20mg 的纤维素分解蛋白/g 于 PCS 中的纤维素,在 50℃-65℃ 下持续 3-7 天,与未添加纤维素分解蛋白的对照水解相比较。典型条件为:1ml 反应,洗涤或未洗涤的 PCS,5% 不溶性固形物,50mM 乙酸钠(pH 5),1mM MnSO_4 , 50℃-65℃,72 小时,通过 AMINEX® HPX-87H 柱(伯乐实验室有限公司(Bio-Rad Laboratories, Inc.),赫拉克勒斯,加州,美国)进行的糖分析。

[0045] 内切葡聚糖酶:在此将术语“内切葡聚糖酶”定义为一种内切-1,4-(1,3; 1,4)- β -D-葡聚糖 4-葡聚糖水解酶(E.C. 3.2.1.4),其催化纤维素、纤维素衍生物(如羧甲基纤维素和羟乙基纤维素)、地衣多糖中的 1,4- β -D-糖苷键和混合 β -1,3 葡聚糖如谷类 β -D-葡聚糖或木葡聚糖以及含有纤维素组分的其他植物材料中的 β -1,4 键的内切水解。可以基于底物粘度的减小或通过还原糖测定确定还原端增加来确定内切葡聚糖酶活性(张(Zhang)等人,2006,生物技术进展(Biotechnology Advances)24:452-481)。出于本发明的目的,根据高斯(Ghose),1987,纯粹与应用化学(Pure and Appl. Chem.)59:257-268 的程序,使用羧甲基纤维素(CMC)水解确定内切葡聚糖酶活性。

[0046] 纤维二糖水解酶:在此将术语“纤维二糖水解酶”定义为一种 1,4- β -D-葡聚糖纤维二糖水解酶(E.C. 3.2.1.91),其催化纤维素、纤维寡糖、或任何含 β -1,4-连接的葡萄糖的聚合物中的 1,4- β -D-糖苷键水解,从该链的还原端或非还原端释放纤维二糖(泰里(Teeri),1997,晶态纤维素降解:纤维二糖水解酶功能的新见解(Crystalline cellulose degradation:New insight into the function of cellobiohydrolases),生物技术趋势(Trends in Biotechnology)15:160-167;泰里(Teeri)等人,1998,里氏木霉纤维二糖水解酶:为何对晶态纤维素如此有效?(*Trichoderma reesei* cellobiohydrolases:why so efficient on crystalline cellulose?),生物化学学会学报(Biochem. Soc. Trans.)26:173-1780。出于本发明的目的,根据范·帝伯赫(van Tilbeurgh)等人,1982,欧洲生化学会联合会快报(FEBS Letters),149:152-156 以及范·帝伯赫和克莱森斯(Claeysens),1985,欧洲生化学会联合会快报,187:283-288 所述的程序,使用一种荧光二糖衍生物 4-甲基伞形酮基- β -D-乳糖确定纤维二糖水解酶活性。

[0047] β -葡糖苷酶:在此将术语“ β -葡糖苷酶”定义为一种 β -D-葡糖苷葡萄糖水解酶(E.C. 3.2.1.21),其催化末端非还原性 β -D-葡萄糖残基的水解,并释放 β -D-葡萄糖。出于本发明的目的,根据文丘里(Venturi)等人,2002,来自嗜热毛壳菌嗜粪变种的胞外 β -D-葡糖苷酶:产生、纯化以及一些生物化学特性(Extracellular beta-D-glucosidase from *Chaetomium thermophilum* var. *coprophilum*:production, purification and some biochemical properties),基础微生物学杂志(J. Basic Microbiol.)42:55-66 描述的基本程序确定 β -葡糖苷酶活性。将一个单位的 β -葡糖苷酶活性定义为在 40℃、pH 5 下,

在含有 0.01%TWEEN® 20 的 100mM 柠檬酸钠中从作为底物的 1mM 对硝基苯基-β-D-吡喃葡萄糖苷每分钟产生 1.0 微摩尔的对硝基苯酚。

[0048] 纤维素分解增强活性:在此将术语“纤维素分解增强活性”定义为由具有纤维素分解活性的多肽增强纤维素材料的水解的生物活性。出于本发明的目的,通过测量在以下条件下来自由纤维素分解蛋白水解纤维素材料的还原糖的增加或纤维二糖与葡萄糖总量的增加来测定纤维素分解增强活性:1-50mg 总蛋白 /g 于 PCS 中的纤维素,其中总蛋白由 50% -99.5% w/w 纤维素分解蛋白和 0.5% -50% w/w 具有纤维素分解增强活性的蛋白构成,在 50℃ -65℃ 持续 1-7 天,与不具有纤维素分解增强活性的相等总蛋白装载的对照水解(1-50mg 纤维素分解蛋白 /g 于 PCS 中的纤维素)相比较。在一个优选方面中,使用在总蛋白质重量的 3%的米曲霉 β-葡萄糖苷酶(根据 WO 02/095014 在米曲霉中重组产生)或总蛋白质重量的 3%的烟曲霉 β-葡萄糖苷酶(如 WO 02/095014 中所描述在米曲霉中重组产生)的纤维素酶蛋白负载量存在的情况下 CELLUCLAST® 1.5L(诺维信公司(Novozymes A/S),丹麦巴格斯瓦尔德(Bagsværd, Denmark))的混合物作为纤维素分解活性的来源。

[0049] 这些具有纤维素分解增强活性的多肽通过将达到相同的水解程度所需要的纤维素分解酶的量降低优选至少 1.01 倍、更优选至少 1.05 倍、更优选至少 1.10 倍、更优选至少 1.25 倍、更优选至少 1.5 倍、更优选至少 2 倍、更优选至少 3 倍、更优选至少 4 倍、更优选至少 5 倍、甚至更优选至少 10 倍、以及最优选至少 20 倍,来增强由具有纤维素分解活性的蛋白催化的纤维素材料的水解。

[0050] 家族 61 糖苷水解酶:在此将术语“家族 61 糖苷水解酶”或“GH61”或“家族 GH61”定义为根据亨利萨特(Henrissat),1991,基于氨基酸序列相似性的糖基水解酶的分类(A classification of glycosyl hydrolases based on amino-acid sequence similarities),生物化学杂志(Biochem. J.)280:309-316,和亨利萨特和贝洛赫(Bairoch),1996,修正糖基水解酶的基于序列的分类(Updating the sequence-based classification of glycosyl hydrolases),生物化学杂志 316:695-696,属于糖苷水解酶家族 61 的多肽。目前,亨利萨特将 GH61 家族列为未分类的,表明属于这一家族的多肽的特性(例如机制、催化亲核试剂/碱和催化质子供体)是未知的。

[0051] 木聚糖降解活性:在此将术语“木聚糖降解活性”或“木聚糖分解活性”定义为水解含有木聚糖的材料生物活性。用于测量木聚糖分解活性的两种基本方法包括:(1)测量总木聚糖分解活性,并且(2)测量单独的木聚糖分解活性(内切木聚糖酶、β-木糖苷酶、阿拉伯呋喃糖苷酶、α-葡萄糖醛酸糖苷酶、乙酰木聚糖酯酶、阿魏酸酯酶、以及 α-葡萄糖醛酸酯酶)。在若干出版物中对木聚糖分解酶的测定的最近进展进行了总结,包括:别雷(Biely)和普查德(Puchard),2006,木聚糖分解酶的测定的测定的最新进展(Recent progress in the assays of xylanolytic enzymes),食品与农业科学杂志(Journal of the Science of Food and Agriculture)86(11):1636-1647;斯潘妮科娃(Spanikova)和别雷,2006,葡萄糖醛酸酯酶-由裂褶菌生成的新型碳水化合物酯酶(Glucuronoyl esterase-Novel carbohydrate esterase produced by Schizophyllum commune),欧洲生化学会联合会快报(FEBS Letters)580(19):4597-4601;赫尔曼(Herrmann)等人,1997,里氏木霉的 β-D-木糖苷酶是一种多功能 β-D-木聚糖木糖水解酶(The beta-D-xylosidase of Trichoderma reesei is a multifunctional beta-D-xylan xylohydrolase),生物化

学杂志 (Biochemical Journal) 321:375-381。

[0052] 总木聚糖降解活性可以通过确定由不同类型的木聚糖 (包括燕麦 (oat spelt) 木聚糖、山毛榉木木聚糖、以及落叶松木木聚糖) 形成的还原糖, 或通过光度法确定从不同共价染色的木聚糖释放的染色的木聚糖片段来测量。最常见的总木聚糖分解活性测定基于由聚合 4-0- 甲基葡萄糖醛酸木聚糖产生的还原糖, 如描述于别雷 (Bailey), 别雷, 坡泰恩 (Poutanen), 1992, 用于木聚糖酶活性测定的多个实验室测试方法 (Interlaboratory testing of methods for assay of xylanase activity), 生物技术杂志 (Journal of Biotechnology) 23(3):257-270 中。

[0053] 出于本发明的目的, 通过测量在以下典型条件下由一种或多种木聚糖降解酶引起的桦木木聚糖 (西格玛化学有限公司 (Sigma Chemical Co., Inc.), 圣路易斯, 密苏里州, 美国) 水解的增加来确定木聚糖降解活性: 1ml 反应、5mg/ml 底物 (总固体)、5mg 木聚糖分解蛋白 /g 底物、50mM 乙酸钠 (pH 5)、50℃、24 小时, 使用如里弗 (Lever), 1972, 用于碳水化合物的比色测定的新反应 (A new reaction for colorimetric determination of carbohydrates), 分析生物化学 (Anal. Biochem) 47:273-279 所描述的对羟基苯甲酸酰肼 (PHBAH) 测定进行糖分析。

[0054] 木聚糖酶活性: 在此将术语“木聚糖酶活性”定义为 1, 4- β -D- 木聚糖 - 木糖水解酶 (E. C. 3. 2. 1. 8) 活性, 其催化木聚糖中的 1, 4- β -D- 木糖苷键的内切水解。出于本发明的目的, 使用桦木木聚糖作为底物确定木聚糖酶活性。将一个单位的木聚糖酶活性定义为在 50℃、pH 5 下的初始期水解中, 在含有 0.01% **TWEEN®** 20 的 50mM 乙酸钠中从作为底物的 2g 桦木木聚糖每升每分钟产生 1.0 μ 摩尔还原糖 (以如里弗 (Lever), 1972, 用于碳水化合物的比色测定的新反应 (A new reaction for colorimetric determination of carbohydrates), 分析生物化学 (Anal. Biochem) 47:273-279 所述的葡萄糖当量测量)。

[0055] β - 木糖苷酶活性: 在此将术语“ β - 木糖苷酶活性”定义为一种 β -D- 木糖苷木糖水解酶 (E. C. 3. 2. 1. 37), 其催化短 β (1 $\rightarrow$ 4)- 木寡糖的外水解, 以从非还原性末端去除连续的 D- 木糖残基。出于本发明的目的, 将一个单位的 β - 木糖苷酶活性定义为在 40℃、pH 5 下在含有 0.01% **TWEEN®** 20 的 100mM 柠檬酸钠中从作为底物的 1mM 对硝基苯基 - β -D- 木糖苷每分钟产生 1.0 微摩尔的对硝基苯酚。

[0056] 乙酰木聚糖酯酶活性: 在此将术语“乙酰木聚糖酯酶活性”定义为一种羧酸酯酶 (EC 3. 1. 1. 72) 活性, 其催化乙酰基自聚合木聚糖、乙酰化木糖、乙酰化葡萄糖、 α - 萘基乙酸酯、和对硝基苯基乙酸酯的水解。出于本发明的目的, 使用 0.5mM 对硝基苯基乙酸酯作为底物, 在含有 0.01% **TWEEN™** 20 的 50mM 乙酸钠 (pH 5.0) 中, 对乙酰木聚糖酯酶活性进行确定。将一个单位的乙酰木聚糖酯酶活性定义为在 pH 5、25℃ 下每分钟能够释放 1 微摩尔对硝基酚根阴离子的酶的量。

[0057] 阿魏酸酯酶 (Feruloyl esterase) 活性: 在此将术语“阿魏酸酯酶活性”定义为 4- 羟基 -3- 甲氧基肉桂酰基 - 糖水水解酶 (EC 3. 1. 1. 73) 活性, 其催化 4- 羟基 -3- 甲氧基肉桂酰基 (阿魏酰基) 基团从酯化的糖 (其在“天然”底物中通常为阿拉伯糖) 的水解, 以产生阿魏酸酯 (4- 羟基 -3- 甲氧基肉桂酸酯)。阿魏酸酯酶 (Feruloyl esterase) 也被称为阿魏酸酯酶 (ferulic acid esterase)、羟基肉桂酰基酯酶、FAE-III、肉桂酸酯水解酶、FAEA、cinnAE、FAE-I、或 FAE-II。出于本发明的目的, 在 50mM 乙酸钠 (pH 5.0) 中使用 0.5mM 阿

魏酸对硝基苯酯作为底物来确定阿魏酸酯酶活性。一个单位的阿魏酸酯酶活性等于,在 pH 5、25℃,每分钟能够释放 1 微摩尔的对硝基酚根阴离子的酶的量。

[0058] α -葡萄糖醛酸糖苷酶活性:在此将术语“ α -葡萄糖醛酸糖苷酶活性”定义为一种 α -D-葡萄糖苷酸葡萄糖醛酸水解酶 (EC 3.2.1.139) 活性,其催化 α -D-葡萄糖苷酸水解为 D-葡萄糖醛酸酯和醇。出于本发明的目的,根据德弗里斯 (de Vries),1998,细菌学杂志 (J. Bacteriol.) 180:243-249 来测定 α -葡萄糖醛酸糖苷酶活性。一个单位的 α -葡萄糖醛酸糖苷酶活性等于能够在 pH 5、40℃ 下每分钟释放 1 微摩尔的葡萄糖醛酸或 4-O-甲基葡萄糖醛酸的酶的量。

[0059] α -L-阿拉伯呋喃糖苷酶活性:在此将术语“ α -L-阿拉伯呋喃糖苷酶活性”定义为一种 α -L-阿拉伯呋喃糖苷阿拉伯呋喃水解酶 (EC 3.2.1.55) 活性,其催化 α -L-阿拉伯糖苷中的末端非还原性 α -L-阿拉伯呋喃糖苷残基的水解。该酶活性对 α -L-阿拉伯呋喃糖苷、含有 (1,3)-和 / 或 (1,5)-键的 α -L-阿拉伯聚糖、阿拉伯糖基木聚糖、以及阿拉伯半乳糖起作用。 α -L-阿拉伯呋喃糖苷酶还被称为阿拉伯糖苷酶、 α -阿拉伯糖苷酶、 α -L-阿拉伯糖苷酶、 α -阿拉伯呋喃糖苷酶、多糖 α -L-阿拉伯呋喃糖苷酶、 α -L-阿拉伯呋喃糖苷水解酶、L-阿拉伯糖苷酶、或 α -L-阿拉伯聚糖酶。出于本发明的目的,使用每 ml 的 100mM 乙酸钠 (pH 5) 中 5mg 的中等粘度小麦阿拉伯糖基木聚糖 (麦格酶国际爱尔兰股份有限公司 (Megazyme International Ireland, Ltd.), 爱尔兰威克洛郡布瑞公司 (Bray, Co. Wicklow, Ireland) 以总体积 200 μ l 在 40℃ 下持续 30 分钟,接着通过 AMINEX® HPX-87H 柱层析 (伯乐实验室有限公司 (Bio-Rad Laboratories, Inc.), 赫拉克勒斯,加州,美国) 进行阿拉伯糖分析来测定 α -L-阿拉伯呋喃糖苷酶活性。

[0060] 含有木聚糖的材料:在此将术语“含有木聚糖的材料”定义为包含含有 β -(1-4) 连接的木糖残基主链的植物细胞壁多糖的任何材料。陆生植物的木聚糖是具有 β -(1-4)-D-吡喃木糖主链的杂聚物,其通过短的碳水化合物链分支。它们包括 D-葡萄糖醛酸或其 4-O-甲基醚、L-阿拉伯糖、和 / 或不同的低聚糖,这些低聚糖由 D-木糖、L-阿拉伯糖、D-或 L-半乳糖、以及 D-葡萄糖构成。可以将木聚糖类型的多糖分成同源木聚糖 (homoxylan) 和异源木聚糖 (heteroxylan),包括葡萄糖醛酸木聚糖、(阿拉伯糖)葡萄糖醛酸木聚糖、(葡萄糖醛酸)阿拉伯糖基木聚糖、阿拉伯糖基木聚糖、以及复杂的异源木聚糖。参见,例如,埃伯林格罗瓦 (Ebringerova) 等人,2005,聚合物科学进展 (Adv. Polym. Sci.) 186:1-67。

[0061] 在本发明的方法中,可以使用含有木聚糖的任何材料。在一个优选方面,含有木聚糖的材料是木质纤维素。

[0062] 含有木聚糖的材料:在此将术语“含有木聚糖的材料”定义为包含含有 β -(1-4) 连接的木糖残基主链的植物细胞壁多糖的任何材料。陆生植物的木聚糖是具有 β -(1-4)-D-吡喃木糖主链的杂聚物,其通过短的碳水化合物链分支。它们包括 D-葡萄糖醛酸或其 4-O-甲基醚、L-阿拉伯糖、和 / 或不同的低聚糖,这些低聚糖由 D-木糖、L-阿拉伯糖、D-或 L-半乳糖、以及 D-葡萄糖构成。可以将木聚糖类型的多糖分成同源木聚糖 (homoxylan) 和异源木聚糖 (heteroxylan),包括葡萄糖醛酸木聚糖、(阿拉伯糖)葡萄糖醛酸木聚糖、(葡萄糖醛酸)阿拉伯糖基木聚糖、阿拉伯糖基木聚糖、以及复杂的异源木聚糖。参见,例如,埃伯林格罗瓦 (Ebringerova) 等人,2005,聚合物科学进展 (Adv. Polym.

Sci.) 186:1-67。

[0063] 在本发明的方法中,可以使用含有木聚糖的任何材料。在一个优选方面中,含有木聚糖的材料是木质纤维素。

[0064] 分离的多肽:如在此使用的术语“分离的多肽”是指从一个来源分离的多肽。在一个优选方面中,该多肽是至少 1% 纯的,优选至少 5% 纯的,更优选至少 10% 纯的,更优选至少 20% 纯的,更优选至少 40% 纯的,更优选至少 60% 纯的,甚至更优选至少 80% 纯的,并且最优选至少 90% 纯的,如通过 SDS-PAGE 所确定。

[0065] 基本上纯的多肽:术语“基本上纯的多肽”在此指一种包含按重量计至多 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%,并且甚至最优选至多 0.5% 的该多肽与其天然地或重组地相关的其他多肽材料的多肽制剂。因此,优选的是该基本上纯的多肽按存在于该制剂中的所有多肽材料的重量计是至少 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 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,并且甚至最优选 100% 纯的。这些多肽优选处于基本上纯的形式,即,该多肽制剂基本不含该多肽与其天然地或重组地相关的其他多肽材料。例如,这可以通过采用熟知的重组方法或采用经典的纯化方法制备多肽来完成。

[0066] 成熟多肽:在此将术语“成熟多肽”定义为在翻译和任何翻译后修饰(如 N 末端加工、C 末端截短、糖基化、磷酸化等)之后呈其最终形式的一种多肽。

[0067] 成熟多肽编码序列:在此将术语“成熟多肽编码序列”定义为编码成熟多肽的核苷酸序列。

[0068] 一致性:两个氨基酸序列之间或两个核苷酸序列之间的相关性由参数“一致性”描述。

[0069] 出于本发明的目的,使用尼德曼-翁施 (Needleman-Wunsch) 算法(尼德曼和翁施,1970,分子生物学杂志 (J. Mol. Biol.) 48:443-453) 来确定两个氨基酸序列之间的一致性程度,该算法如 EMBOSS 软件包 (EMBOSS:欧洲分子生物学开放软件套件 (The European Molecular Biology Open Software Suite),赖斯 (Rice) 等人,2000,遗传学趋势 (Trends in Genetics) 16:276-277) (优选 3.0.0 版或更新版本) 的尼德尔 (Needle) 程序所实施的。所使用的这些任选参数是空位开放罚分 10、空位延伸罚分 0.5,及 EBLOSUM62 (BLOSUM62 的 EMBOSS 版本) 取代矩阵。尼德尔标注的“最长的一致性”的输出(使用 -非简化选项获得)被用作百分比一致性,并且如下计算:

[0070] $(\text{一致的残基} \times 100) / (\text{比对长度} - \text{比对中的空位总数})$

[0071] 出于本发明的目的,使用尼德曼-翁施算法(尼德曼 (Needleman) 和翁施 (Wunsch),1970,见上文) 来确定两个脱氧核糖核苷酸序列之间的一致性的程度,如在 EMBOSS 软件包 (EMBOSS:欧洲分子生物学开放软件套件,赖斯 (Rice) 等人,2000,见上文) (优选 3.0.0 版或更新版本) 的尼德尔程序中所实施的。所使用的这些任选参数是空位开放罚分 10、空位延伸罚分 0.5,及 EDNAFULL (NCBI NUC4.4 的 EMBOSS 版本) 取代矩阵。标记为“最长一致性”的尼德尔的输出(使用 -nobrief 选项获得)被用作百分比一致性并且被计算如下:

[0072] $(\text{一致的脱氧核糖核苷酸} \times 100) / (\text{比对长度} - \text{比对中的空位总数})$

[0073] 同源序列：在此将术语“同源序列”定义为在用感兴趣的多肽的 tfasty 检索（皮尔森 (Pearson), W. R., 1999, 在生物信息学方法和协议 (Bioinformatics Methods and Protocols) 中, S. 麦森纳 (Misener) 和 S. A. 克拉韦茨 (Krawetz), 编辑, 185-219 页) 中具有小于 0.001 的 E 值（或期望分数）的预测蛋白。

[0074] 多肽片段：在此将术语“多肽片段”定义为使一个或多个（若干个）氨基酸从成熟多肽或其同源序列的氨基和 / 或羧基末端缺失的多肽, 其中该片段具有生物活性。

[0075] 子序列：在此将术语“子序列”定义为使一个或多个（若干个）核苷酸从成熟多肽编码序列或其同源序列的 5' 端和 / 或 3' 端缺失的核苷酸序列, 其中该子序列编码具有生物活性的一个多肽片段。

[0076] 等位基因变体：术语“等位基因变体”在此指占用同一染色体位点的一种基因的两个或更多个替代形式中的任一者。等位基因变异由突变天然产生, 并且可以导致群体内的多态性。基因突变可以是沉默的（在所编码的多肽中没有改变）或可编码具有改变的氨基酸序列的多肽。多肽的等位基因变体是由基因的等位基因变体编码的多肽。

[0077] 分离的多核苷酸：如在此使用的术语“分离的多核苷酸”是指从一个来源分离的多核苷酸。在一个优选方面中, 该多核苷酸是至少 1% 纯的, 优选至少 5% 纯的, 更优选至少 10% 纯的, 更优选至少 20% 纯的, 更优选至少 40% 纯的, 更优选至少 60% 纯的, 甚至更优选至少 80% 纯的, 并且最优选至少 90% 纯的, 如通过琼脂糖电泳所确定。

[0078] 基本上纯的多核苷酸：如在此使用的术语“基本上纯的多核苷酸”是指不含其他外部或不想要的核苷酸并且处在适用于基因工程化蛋白生产系统内部的形式多核苷酸制剂。因此, 基本上纯的多核苷酸包含按重量计至多 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%, 并且甚至最优选至多 0.5% 的该多肽与其天然地或重组地相关的其他多核苷酸材料。然而, 基本上纯的多核苷酸可以包括天然存在的 5' 和 3' 非翻译区, 如启动子和终止子。优选的是该基本上纯的多核苷酸按重量计是至少 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, 并且甚至最优选 99.5% 纯的。这些多核苷酸优选处于基本上纯的形式, 即, 该多核苷酸制剂基本不含该多核苷酸与其天然地或重组地相关的其他多核苷酸材料。这些多核苷酸可以是基因组、cDNA、RNA、半合成、合成来源的, 或其任意组合。

[0079] 编码序列：当在此使用时, 术语“编码序列”意指直接指明其蛋白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阅读框架决定, 该开放阅读框架通常以 ATG 起始密码子或替代性起始密码子（例如 GTG 和 TTG）开始, 并且以终止密码子（例如 TAA、TAG、和 TGA）结束。编码序列可以是 DNA、cDNA、合成的或重组的核苷酸序列。

[0080] cDNA：在此将术语“cDNA”定义为可以通过从成熟的剪接的 mRNA 分子逆转录制备的 DNA 分子, 该 mRNA 分子是从真核细胞中获得。cDNA 缺少可存在于相应基因组 DNA 中的内含子序列。早先的初始 RNA 转录本是 mRNA 的前体, 其在呈现为成熟的剪接的 mRNA 之前要经一系列的步骤进行加工。这些步骤包括通过称为剪接的过程除去内含子序列。因此, 衍生自 mRNA 的 cDNA 缺乏所有内含子序列。

[0081] 核酸构建体：如在此使用的术语“核酸构建体”是指从天然存在的基因中分离的、

或以自然界中不会另外出现的方式被修饰成包含核酸区段的、或合成的单链或双链的核酸分子。当核酸构建体含有表达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。

[0082] 控制序列：在此将术语“控制序列”定义为包括对于表达编码多肽的多核苷酸所必需的所有组分。各个控制序列对于编码所述多肽的核苷酸序列可以是天然的或外源的，或各个控制序列对于彼此可以是天然的或外源的。此类控制序列包括但不限于前导子、聚腺苷酸化序列、前肽序列、启动子、信号肽序列、以及转录终止子。至少，控制序列包括启动子，以及转录和翻译终止信号。出于引入有利于将这些控制序列与编码一种多肽的核苷酸序列的编码区连接的特异性限制酶切位点的目的，这些控制序列可以提供有多个接头。

[0083] 可操作地连接：术语“可操作地连接”在此指一种配置，其中一个控制序列相对于多核苷酸序列的编码序列放置在一个适当位置处，以使得控制序列指导多肽的编码序列的表达。

[0084] 表达：术语“表达”包括涉及多肽产生的任何步骤，包括但不限于，转录、转录后修饰、翻译、翻译后修饰、以及分泌。

[0085] 表达载体：在此将术语“表达载体”定义为包括编码多肽的多核苷酸并且可操作地与提供了其表达的另外的核苷酸相连接的线性或环状 DNA 分子。

[0086] 宿主细胞：术语“宿主细胞”，如在此使用，包括对于用包括本发明多核苷酸的核酸构建体或表达载体进行的转化、转染、转导等是易感的任何细胞类型。

[0087] 修饰：术语“修饰”在此意指一种多肽的任何化学修饰以及对编码该多肽的 DNA 所进行的遗传操作。该修饰可以是一个或多个（若干个）氨基酸的取代、缺失和 / 或插入，连同一个或多个（若干个）氨基酸侧链的替换。

[0088] 人工变体：当在此使用时，术语“人工变体”意指一种由表达编码多肽变体的修饰的多核苷酸序列的生物体产生的多肽。修饰的核苷酸序列经由多核苷酸序列的修饰通过人为干涉而获得。

[0089] 发明详细说明

[0090] 本发明涉及通过水解预处理的纤维素材料来将该预处理的纤维素材料降解 / 水解为糖的改进方法。本发明还涉及用于从该预处理的纤维素材料生产发酵产品的工艺。

[0091] 发明方法

[0092] 在第一方面中，本发明涉及用于降解 / 水解预处理的纤维素材料的方法，这些方法包括使该预处理的纤维素材料经受：

[0093] - 一种纤维素分解酶组合物；

[0094] - 一种多肽，该多肽具有纤维素分解增强活性；

[0095] - 一种过氧化物酶；以及

[0096] - 一种非离子表面活性剂和 / 或一种阳离子表面活性剂，

[0097] 在适于水解该预处理的木质纤维素材料的条件下。

[0098] 该组分可以通过添加至本发明的方法中而存在。根据本发明，在降解 / 水解过程中添加的这些组分可以作为一种组合物而添加，但是也可以作为两种或更多种单组分或多组分组合物而添加。例如，该纤维素分解酶组合物和该多肽可以作为一种组合物而添加，而该过氧化物酶和该一种或多种表面活性剂可以分开添加。在一个实施例中，该纤维素分

解酶组合物、具有纤维素分解增强活性的该多肽和该过氧化物酶可以作为一种组合物而添加,而该一种或多种表面活性剂可以分开添加。根据本发明,考虑了任何组合。还考虑了在降解/水解之前添加这些组分中的一种或多种。

[0099] 该降解的/水解的预处理的纤维素材料包括糖。可以将这些糖用在用于生产糖浆(例如,高果糖玉米糖浆(HFCS))和/或塑料(例如,聚乙烯、聚苯乙烯以及聚丙烯)、聚乳酸(例如,用于生产PET)的工艺中。这些糖还可以被发酵微生物发酵为发酵产品(如乙醇),该发酵微生物是如酵母,例如来自酵母属的菌株,如能够将C5糖(戊糖)和/或C6糖(己糖)转化为所希望的终产品(如乙醇)的酿酒酵母菌株。下面描述了考虑的产品(包括发酵产品)的非穷尽性列表。下面还描述了适合的发酵微生物的实例。

[0100] 根据本发明,该预处理的纤维素材料可以是农业废弃物、草本材料(包括能源作物)、城市固体废物、纸浆和造纸厂废弃物、废纸、或木材(包括林业废弃物)、或芦竹、甘蔗渣、竹子、玉米芯、玉米纤维、玉米秸秆、芒属、橙皮、稻秸、柳枝稷或麦秸。根据本发明,可以在水解和/或发酵之后回收该降解的预处理的纤维素材料,例如糖或转化为发酵产品的糖。

[0101] 这些糖可以是来自下组的一种糖,该组由以下各项组成:葡萄糖,木糖,甘露糖,半乳糖以及阿拉伯糖。当该终产品是一种发酵产品时,它可以是醇(例如尤其是乙醇)、有机酸、酮、氨基酸或气体。

[0102] 根据本发明,该预处理的纤维素材料可以按任何适合的方式预处理。该纤维素材料的预处理优选可以作为化学预处理、物理预处理、或化学预处理与物理预处理来进行。预处理方法和预处理条件在本领域是熟知的。

[0103] 在一个实施例中,用一种酸预处理该纤维素材料,例如稀酸预处理。在一个优选实施例中,通过在高温、高压下用一种酸(例如稀酸)预处理来进行该纤维素材料的预处理。

[0104] 在一个实施例中,使用乙酸或硫酸进行酸预处理。

[0105] 在一个实施例中,该纤维素材料的预处理是一种碱性预处理,例如铵预处理,例如温和的铵预处理。

[0106] 在另一个实施例中,该预处理是热化学预处理。

[0107] 在一个另外的实施例中,使用有机溶剂预处理来预处理该纤维素材料,例如Acetosolv和Acetocell工艺。

[0108] 在一个优选实施例中,该材料是稀酸预处理的玉米秸秆。在另一个实施例中,该预处理的材料是稀酸预处理的玉米芯。

[0109] 在本发明的背景下,降解预处理的纤维素材料与水解预处理的纤维素材料相同。

[0110] 水解法条件

[0111] 适合的方法条件是本领域普通技术人员熟知的或可以由本领域普通技术人员容易地确定。在一个实施例中,可以在10%-50%(w/w)TS(总固体)下,例如在15%-40%TS下,例如在15%-30%TS下,例如在大约20%TS下进行水解。水解可以进行12-240小时,例如24-192小时,例如48-144小时,例如大约96小时。水解过程中的温度可以在30℃-70℃之间,例如40℃-60℃,例如45℃-55℃,例如大约50℃。水解过程中的pH可以在4-7之间,例如pH 4.5-6,例如大约pH 5。

[0112] 在一个更具体实施例中,本发明涉及用于降解预处理的纤维素材料的方法,这些

方法包括使该预处理的纤维素材料经受：

[0113] - 一种纤维素分解酶组合物；

[0114] - 具有纤维素分解增强活性的多肽，优选是来源于金黄色嗜热子囊菌的如在此的 SEQ ID NO:14 所示的多肽，和 / 或来源于埃默森青霉菌的示于在此的 SEQ ID NO:72 中的多肽，或一种具有纤维素分解增强活性、与在此的 SEQ ID NO:14 或在此的 SEQ ID NO:72 具有至少 60%、至少 70%、至少 80%、至少 90%、至少 95%、至少 97%、至少 99% 序列一致性的多肽；

[0115] - 一种分类为 EC 1.11.1.7 过氧化物酶的过氧化物酶，优选是来源于灰盖鬼伞示于在此的 SEQ ID NO:71 中的过氧化物酶 (CiP)；或一种具有过氧化物酶活性、与在此的 SEQ ID NO:71 具有至少 60%、至少 70%、至少 80%、至少 90%、至少 95%、至少 97%、至少 99% 一致性的多肽；

[0116] - 一种非离子表面活性剂和 / 或一种阳离子表面活性剂；

[0117] 在适于水解该预处理的木质纤维素材料的条件下。

[0118] 纤维素分解酶组合物，酶和多肽

[0119] 在下面的“酶”部分披露了可以适合在本发明的用于降解预处理的纤维素材料的方法中或在本发明的用于生产发酵产品的工艺中使用的纤维素分解酶组合物、酶和多肽的非穷尽性披露。根据本发明，至少一种纤维素分解酶组合物；一种具有纤维素分解增强活性的多肽；一种过氧化物酶；以及一种非离子表面活性剂和 / 或一种阳离子表面活性剂存在或在水解之前和 / 或水解过程中被添加。

[0120] 纤维素分解酶组合物、酶、和具有纤维素分解增强活性的多肽、过氧化物酶及非离子和 / 或阳离子表面活性剂的最佳量取决于若干因素，包括但不限于纤维素分解酶、纤维素分解底物、纤维素分解底物的浓度、纤维素分解底物 / 材料的一个或多个预处理、温度、时间、pH，并且包括发酵微生物。

[0121] 根据本发明，任何纤维素分解酶组合物可以用于水解。在水解过程中装载的纤维素分解酶组合物或总酶和多肽的有效量可以在约 0.1 至约 25mg 之间，例如约 1-10mg，例如约 2 至约 8mg，例如大约 4mg 蛋白 / g 纤维素材料。

[0122] 在一个实施例中，具有针对纤维素材料的纤维素分解增强活性的多肽的量是约 0.01 至约 20mg，例如约 0.01 至约 10mg，例如约 0.01 至约 5mg，例如约 0.025 至约 1.5mg，例如约 0.05 至约 1.25mg，例如约 0.075 至约 1.25mg，例如约 0.1 至约 1.25mg，例如约 0.15 至约 1.25mg，以及例如约 0.25 至约 1.0mg/g 纤维素材料。

[0123] 在一个实施例中，针对纤维素材料的过氧化物酶的量是约 0.001 至约 20mg，例如约 0.01 至约 15mg，例如约 0.02 至约 10mg，例如约 0.05 至约 5mg/g 纤维素材料。

[0124] 该纤维素分解酶组合物可以包括一种或多种（若干种）选自下组的酶，该组由以下各项组成：内切葡聚糖酶、纤维二糖水解酶 (CBH) 以及 β -葡萄糖苷酶。该纤维素分解酶组合物还可以包括其他对产生纤维素分解酶的供体或宿主细胞而言天然或外源的酶和 / 或多肽。例如，该纤维素分解酶组合物可以由一种产生纤维素分解酶并且进一步产生一种或多种另外的重组酶的宿主细胞产生，例如像，一种对该宿主细胞而言外源的具有纤维素分解增强活性的 GH61 多肽以及其他对该宿主细胞而言外源的酶（如 β -葡萄糖苷酶）。

[0125] 在水解过程中使用的该纤维素分解酶组合物可以来源于或由以下产生：木霉属

的菌株,优选里氏木霉菌株;或腐质霉属的菌株,例如特异腐质霉菌株;或金孢子菌属的菌株,例如拉克淖金孢子菌(*Chrysosporium lucknowense*)菌株;或毁丝霉属的菌株,例如嗜热毁丝霉菌株。

[0126] 根据本发明,具有纤维素分解增强活性的多肽可以存在或在水解过程中添加。具有纤维素分解增强活性的该多肽可以与该纤维素分解酶组合物分开添加(例如,一种重组或单组分多肽),但是也可以作为所述组合物的一部分(例如,在产生纤维素分解酶的生产/宿主细胞中重组地产生)。具有纤维素分解增强活性的该多肽可以是一种GH61多肽。在一个优选实施例中,该GH61多肽可以来源于嗜热子囊菌属,如金黄色嗜热子囊菌菌株,如在例如WO 2005/074656中描述为SEQ ID NO:2或在此的SEQ ID NO:14的多肽;或来源于梭孢壳属的多肽,如土生梭孢壳菌株,如在例如WO 2005/074647中描述为SEQ ID NO:7(DNA)和SEQ ID NO:8(氨基酸)或在此的SEQ ID NO:8的多肽;或来源于曲霉属的菌株的多肽,如烟曲霉菌株,如在例如WO 2010/138754中描述为SEQ ID NO:1和SEQ ID NO:2的多肽;或来源于衍生自青霉属的菌株的多肽,如埃默森青霉菌菌株,如在例如WO 2011/041397中披露为SEQ ID NO:2或在此的SEQ ID NO:72的多肽。

[0127] 在一个实施例中,具有纤维素分解增强活性的该多肽与在此的SEQ ID NO:14具有至少60%、优选至少65%、更优选至少70%、更优选至少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%、并且甚至最优选至少96%、至少97%、至少98%、或至少99%序列一致性。

[0128] 在一个实施例中,具有纤维素分解增强活性的该多肽与在此的SEQ ID NO:72具有至少60%、优选至少65%、更优选至少70%、更优选至少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%、并且甚至最优选至少96%、至少97%、至少98%、或至少99%序列一致性。

[0129] 在本发明的一个实施例中, β -葡糖苷酶可以存在或在水解过程中添加。该 β -葡糖苷酶可以作为一种单独酶(例如,一种重组或单组分酶)或作为该纤维素分解酶组合物的一部分(例如,在产生纤维素分解酶的生产/宿主细胞中重组地产生)而添加至水解中。在一个实施例中,该 β -葡糖苷酶可以是一种来源于曲霉属的菌株的 β -葡糖苷酶,如米曲霉,如披露于例如WO 02/095014中的 β -葡糖苷酶,或披露于例如WO 2008/057637中具有 β -葡糖苷酶活性的融合蛋白,或烟曲霉,如在WO 2005/047499中披露为SEQ ID NO:2或在此的SEQ ID NO:78的 β -葡糖苷酶,或披露于例如WO 2012/044915中例如具有以下突变:F100D、S283G、N456E、F512Y(在此使用SEQ ID NO:78用于编号)的烟曲霉 β -葡糖苷酶变体;或棘孢曲霉菌株(例如,WO 2012/030845)或青霉属的菌株,如披露于例如WO 2007/019442中或在此的SEQ ID NO:62的巴西青霉菌株,或木霉属的菌株,如里氏木霉菌株。

[0130] 在一个实施例中,该 β -葡糖苷酶来自曲霉属的菌株,如烟曲霉菌株,如烟曲霉 β -葡糖苷酶(在此的SEQ ID NO:78),该 β -葡糖苷酶包括一个或多个选自下组的取代,该组由以下各项组成:L89M、G91L、F100D、I140V、I186V、S283G、N456E以及F512Y;例如其具有以下取代的变体:

[0131] -F100D+S283G+N456E+F512Y;

[0132] -L89M+G91L+I186V+I140V;

[0133] -I186V+L89M+G91L+I140V+F100D+S283G+N456E+F512Y(在此使用 SEQ ID NO:78 用于编号)。

[0134] 在一个实施例中,取代的数目在 1 与 10 之间,例如 1 与 8 之间,例如 1 与 6 之间,例如 1 与 4 之间,例如 1、2、3、4、5、6、7、8、9 或 10 个取代。

[0135] 在一个实施例中,该 β -葡萄糖苷酶是一种与在此的 SEQ ID NO:78 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的 β -葡萄糖苷酶。

[0136] 在一个实施例中,该 β -葡萄糖苷酶变体是一种与在此的 SEQ ID NO:78 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的 β -葡萄糖苷酶变体。

[0137] 在本发明的一个实施例中,木聚糖酶可以存在或在水解过程中添加。该木聚糖酶可以作为一种单独酶(例如,一种重组或单组分酶)或作为该纤维素分解酶组合物的一部分(例如,在产生纤维素分解酶的生产/宿主细胞中重组地产生)而添加至水解中。在一个优选实施例中,该木聚糖酶是一种 GH10 木聚糖酶。在一个实施例中,该木聚糖酶来源于曲霉属的菌株,例如来自烟曲霉的菌株,例如在 WO 2006/078256 中披露为 SEQ ID NO:6(Xyl III) 或在此的 SEQ ID NO:75 的木聚糖酶,或来自棘孢曲霉的菌株,例如在 WO 94/21785 中披露为例如 SEQ ID NO:5(Xyl II) 或在此的 SEQ ID NO:74 的木聚糖酶。

[0138] 在一个实施例中,该木聚糖酶是一种与在此的 SEQ ID NO:74 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的木聚糖酶。

[0139] 在一个实施例中,该木聚糖酶是一种与在此的 SEQ ID NO:75 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的木聚糖酶。

[0140] 在本发明的一个实施例中, β -木糖苷酶可以存在或在水解过程中添加。该 β -木糖苷酶可以作为一种单独酶(例如,一种重组或单组分酶)或作为该纤维素分解酶组合物的一部分(例如,在产生纤维素分解酶的生产/宿主细胞中重组地产生)而添加至水解中。在一个实施例中,该 β -木糖苷酶是一种来源于曲霉属的菌株的 β -木糖苷酶,例如烟曲霉菌株,例如披露于未决美国临时编号 61/526833 或 WO 2013/028928(实例 16 和 17)(通过引用而结合)中或在此的 SEQ ID NO:73 的 β -木糖苷酶,或来源于木霉属的菌株,例如里氏木霉菌株,例如 WO 2011/057140 中的 SEQ ID NO:58 的成熟多肽。

[0141] 在一个实施例中,该 β -木糖苷酶是一种与在此的 SEQ ID NO:73 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的 β -木糖苷酶。

[0142] 在本发明的一个实施例中,纤维二糖水解酶 I(CBH I) 可以存在或在水解过程中添

加。该纤维二糖水解酶 I (CBH I) 可以作为一种单独酶 (例如, 一种重组或单组分酶) 或作为该纤维素分解酶组合物的一部分 (例如, 在产生纤维素分解酶的生产 / 宿主细胞中重组地产生) 而添加至水解中。在一个实施例中, 纤维二糖水解酶 I (CBH I) 是一种来源于以下的纤维二糖水解酶 I, 曲霉属的菌株, 例如烟曲霉菌株, 例如披露于例如 WO 2011/057140 中的 SEQ ID NO:6 或在此的 SEQ ID NO:76 的 Cel7a CBH I, 或木霉属的菌株, 例如里氏木霉菌株。

[0143] 在一个实施例中, 该 CBH I 是一种与在此的 SEQ ID NO:76 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的 CBH I。

[0144] 在本发明的一个实施例中, 纤维二糖水解酶 II (CBH II) 可以存在或在水解过程中添加。该纤维二糖水解酶 II (CBH II) 可以作为一种单独酶 (例如, 一种重组或单组分酶) 或作为该纤维素分解酶组合物的一部分 (例如, 在产生纤维素分解酶的生产 / 宿主细胞中重组地产生) 而添加至水解中。在一个实施例中, 该纤维二糖水解酶 II (CBH II) 是一种来源于以下的纤维二糖水解酶 II, 曲霉属的菌株, 例如烟曲霉菌株, 例如在 WO 2011/057140 中示为 SEQ ID NO:18 或在此的 SEQ ID NO:77 的纤维二糖水解酶 II; 木霉属的菌株, 例如里氏木霉; 或梭孢壳属的菌株, 例如土生梭孢壳菌株, 例如来自土生梭孢壳的纤维二糖水解酶 II CEL6A。

[0145] 在一个实施例中, 该 CBH II 是一种与在此的 SEQ ID NO:77 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的 CBH II。

[0146] 在本发明的一个实施例中, 该纤维素分解酶组合物可以是一种里氏木霉纤维素分解酶组合物并且具有纤维素分解增强活性的该多肽是金黄色嗜热子囊菌 GH61A (例如, WO 2005/074656 中的 SEQ ID NO:2 或在此的 SEQ ID NO:14)。在一个实施例中, β -葡萄糖苷酶也存在或在水解过程中添加。该 β -葡萄糖苷酶可以优选是一种米曲霉 β -葡萄糖苷酶融合蛋白 (例如, 2008/057637 中的 SEQ ID NO:74 或 76 或在此的 SEQ ID NO:68 或 70)。在一个实施例中, 该 β -葡萄糖苷酶可以优选是一种棘孢曲霉 β -葡萄糖苷酶, 例如披露于在此的 SEQ ID NO:66 中的 β -葡萄糖苷酶。

[0147] 在一个实施例中, 该纤维素分解酶组合物是一种里氏木霉纤维素分解酶组合物并且具有纤维素分解增强活性的该多肽是在 WO 2011/041397 中披露为 SEQ ID NO:2 的埃默森青霉菌 GH61A 多肽 (在此的 SEQ ID NO:72)。在一个实施例中, β -葡萄糖苷酶也可以存在或在水解过程中添加。该 β -葡萄糖苷酶可以是一种烟曲霉 β -葡萄糖苷酶 (例如, WO 2005/047499 的 SEQ ID NO:2 或在此的 SEQ ID NO:78) 或其具有以下突变的变体: F100D、S283G、N456E、F512Y (使用 SEQ ID NO>78 用于编号) (参见 WO 2012/044915)。

[0148] 在一个优选实施例中, 根据本发明的用于降解预处理的纤维素材料的方法使用的该纤维素分解酶组合物可以是一种里氏木霉纤维素分解酶组合物并且其中存在或添加以下组分中的一种或多种:

[0149] (i) 一种烟曲霉纤维二糖水解酶 I;

[0150] (ii) 一种烟曲霉纤维二糖水解酶 II；

[0151] (iii) 一种烟曲霉 β -葡萄糖苷酶或其变体,例如具有以下取代中的一个或多个: F100D、S283G、N456E、F512Y(在此使用 SEQ ID NO:78 用于编号)(参见 WO 2012/044915); 以及

[0152] (iv) 一种具有纤维素分解增强活性的青霉属(埃默森(emersonii))GH61 多肽;或其同系物。

[0153] 在一个实施例中,存在或还添加一种木聚糖酶(例如,来源于烟曲霉并且在 WO 2006/078256 中披露为 SEQ ID NO:6(Xyl III)或在此的 SEQ ID NO:75,或来源于棘孢曲霉并且在 WO 94/21785 中披露为 SEQ ID NO:5(Xyl II)(在此的 SEQ ID NO:74)),和/或一种 β -木糖苷酶(例如,来源于烟曲霉并且披露于未决美国临时编号 61/526833 或 WO 2013/028928 或在此的 SEQ ID NO:73 中)。

[0154] 根据本发明的方法,该纤维素分解酶组合物可以与一种或多种(若干种)选自下组的酶一起添加或存在,该组由以下各项组成:半纤维素酶、酯酶、蛋白酶以及漆酶。

[0155] 根据本发明,添加或存在的该纤维素分解酶组合物可以进一步包括一种或多种(若干种)选自下组的酶,该组由以下各项组成:木聚糖酶、乙酰木聚糖酯酶、阿魏酸酯酶、阿拉伯呋喃糖苷酶、木糖苷酶、葡萄糖醛酸糖苷酶及其组合。

[0156] 过氧化物酶

[0157] 在本发明的降解预处理的纤维素材料的方法中,一种过氧化物酶与一种纤维素分解酶组合物;一种具有纤维素分解增强活性的多肽;以及一种非离子表面活性剂和/或一种阳离子表面活性剂一起存在或在水解过程中添加。

[0158] 根据本发明,术语“过氧化物酶”是一种过氧化物酶或过氧化物分解酶。该过氧化物酶可以选自下组,该组包括过氧化物酶或过氧化物分解酶,包括但不限于以下各项: E. C. 1. 11. 1. 1NADH 过氧化物酶;E. C. 1. 11. 1. 2NADPH 过氧化物酶;E. C. 1. 11. 1. 3 脂肪酸过氧化物酶;E. C. 1. 11. 1. 5 细胞色素 c 过氧化物酶;E. C. 1. 11. 1. 5;E. C. 1. 11. 1. 6 过氧化氢酶;E. C. 1. 11. 1. 7 过氧化物酶;E. C. 1. 11. 1. 8 碘化物过氧化物酶;E. C. 1. 11. 1. 9 谷胱甘肽过氧化物酶;E. C. 1. 11. 1. 10 氯化物过氧化物酶;E. C. 1. 11. 1. 11L- 抗坏血酸过氧化物酶;E. C. 1. 11. 1. 12 磷脂过氧化氢谷胱甘肽过氧化物酶;E. C. 1. 11. 1. 13 锰过氧化物酶;E. C. 1. 11. 1. 14 木质素过氧化物酶;E. C. 1. 11. 1. 15 过氧化物氧还蛋白;E. C. 1. 11. 1. 16 通用过氧化物酶;E. C. 1. 11. 1. B2 氯化物过氧化物酶;E. C. 1. 11. 1. B6 碘化物过氧化物酶(含钒);E. C. 1. 11. 1. B7 溴化物过氧化物酶;E. C. 1. 11. 1. B8 碘化物过氧化物酶。

[0159] 在一个优选实施例中,该过氧化物酶是一种 E. C. 1. 11. 1. 7 过氧化物酶。

[0160] 该过氧化物酶可以来源于任何微生物,例如一种真菌生物体,例如酵母或丝状真菌,或一种细菌;或一种植物。

[0161] 在一个优选实施例中,该过氧化物酶是一种来源于鬼伞属的菌株,例如灰盖鬼伞菌株的过氧化物酶(E. C. 1. 11. 1. 7),例如在此示为 SEQ ID NO:71 的过氧化物酶(CiP)。在一个实施例中,该过氧化物酶与在此的 SEQ ID NO:71 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性。

[0162] 表面活性剂

[0163] 在本发明的用于降解预处理的纤维素材料的方法中,一种非离子表面活性剂、一种阳离子表面活性剂或其混合物可以与一种纤维素分解酶组合物;一种具有纤维素分解增强活性的多肽;以及一种过氧化物酶一起存在或在水解过程中添加。

[0164] 非离子表面活性剂:

[0165] 非离子表面活性剂是本领域熟知的表面活性剂。根据本发明,可以使用任何非离子表面活性剂。该非离子表面活性剂可以是一种烷基或一种芳基。非离子表面活性剂的实例包括甘油醚、乙二醇醚、乙醇酰胺、磺酰胺 (sulfoanlyamide)、醇、酰胺、醇乙氧基化物、甘油酯、乙二醇酯、甘油酯和乙二醇酯的乙氧基化物、糖基烷基多糖苷、聚氧乙烯化脂肪酸、烷醇胺缩合物、烷醇酰胺、叔炔属乙二醇 (tertiary acetylenic glycol)、聚氧乙烯化硫醇、羧酸酯、以及聚氧乙烯化聚氧丙烯乙二醇 (例如 EO/PO 嵌段共聚物 (EO 是环氧乙烷, PO 是环氧丙烷))、EO 聚合物和共聚物、多胺、和聚乙烯吡咯烷酮。

[0166] 在一个实施例中,该非离子表面活性剂是一种直链伯或仲或具有化学式 $RO(CH_2CH_2O)_nH$ 的支链醇乙氧基化物,其中 R 是烃链长度并且 n 是环氧乙烷的摩尔平均数,例如其中 R 是范围从 C9 至 C16 的直链伯或支链仲烃链长度并且 n 范围从 6 至 13,例如醇乙氧基化物,其中 R 是直链 C9-C11 烃链长度,并且 n 是 6。

[0167] 在一个优选实施例中,该非离子表面活性剂是壬基酚乙氧基化物。在一个优选实施例中,该非离子表面活性剂是 $C_{14}H_{22}O(C_2H_4O)_n$ 。在一个优选实施例中,该非离子表面活性剂是 C_{13} -醇聚乙二醇醚 (10EO)。在一个优选实施例中,该非离子表面活性剂是 EO、PO 共聚物。在一个优选实施例中,该非离子表面活性剂是烷基聚乙二醇醚。在一个优选实施例中,该非离子表面活性剂是 $RO(EO)_5H$ 。在一个优选实施例中,该非离子表面活性剂是 $HOCH_2(EO)_nCH_2OH$ 。在一个优选实施例中,该非离子表面活性剂是 $HOCH_2(EO)_nCH_2OH$ 。

[0168] 阳离子表面活性剂:

[0169] 阳离子表面活性剂是本领域熟知的表面活性剂。根据本发明,可以使用任何阳离子表面活性剂。在一个实施例中,该阳离子表面活性剂是一种伯、仲或叔胺,例如二盐酸奥替尼啶;烷基三甲基铵盐,例如鲸蜡基三甲基溴化铵 (cetyl trimethylammonium bromide) (CTAB) (又名十六烷基三甲基溴化铵 (hexadecyl trimethyl ammonium bromide))、十六烷基三甲基氯化铵 (CTAC)、氯化十六烷基吡啶 (CPC)、苯扎氯铵 (BAC)、苄索氯铵 (BZT)、5-溴-5-硝基-1,3-二噁烷、二甲基双十八烷基氯化铵、双十八烷基二甲基溴化铵 (DODAB) 以及十六烷基三甲基溴化铵 (hexadecyltrimethylammonium bromide)。

[0170] 在一个优选实施例中,该阳离子表面活性剂是 $C_{21}H_{38}NCl$ 。在一个优选实施例中,该阳离子表面活性剂是 $CH_3(CH_2)_{15}N(CH_3)_3Br$

[0171] 发明工艺

[0172] 在此方面中,本发明涉及用于生产发酵产品的工艺,这些工艺包括

[0173] (a) 根据本发明的方法,水解预处理的纤维素材料;

[0174] (b) 用一种或多种(若干种)发酵微生物对该材料进行发酵,以产生该发酵产品;并且

[0175] (c) 任选地从发酵中回收该发酵产品。

[0176] 在水解步骤(还称为糖化)中,将该预处理的纤维素材料水解,以将纤维素以及可

替代地还有半纤维素分解成可发酵糖,如葡萄糖、纤维二糖、木糖、木酮糖、阿拉伯糖、甘露糖、半乳糖、和 / 或可溶性寡糖。在易于由本领域普通技术人员确定的条件下、在适合的含水环境中进行水解。在一个优选方面中,水解在适合于一种或多种酶的活性,即对于该一种或多种酶而言最佳的条件下进行。水解可以作为分批补料过程或连续过程进行,其中将预处理的纤维素材料(底物)逐渐补料至例如含酶的水解溶液中。可以在受控的 pH、温度以及混合条件下在搅拌釜反应器或发酵罐中进行水解。适合的处理时间、温度以及 pH 条件可以由本领域普通技术人员容易地确定。适合的水解条件的实例可以发现于以上“水解法条件”部分中。

[0177] 在一个实施例中,水解步骤(a)和水解步骤(b)顺序或同时进行。在一个实施例中,水解步骤(a)和水解步骤(b)作为分开的水解和发酵(SHF)进行。在一个实施例中,水解步骤(a)和水解步骤(b)作为同时糖化和发酵(SSF)进行。在一个实施例中,水解步骤(a)和水解步骤(b)作为同时糖化和共发酵(SSCF)进行。在一个实施例中,水解步骤(a)和水解步骤(b)作为混合水解和发酵(HHF)进行。在一个实施例中,水解步骤(a)和水解步骤(b)作为分开的水解和共发酵(SHCF)进行。在一个实施例中,水解步骤(a)和水解步骤(b)作为混合水解和共发酵(HHCF)进行。在一个实施例中,水解步骤(a)和水解步骤(b)作为直接微生物转化(DMC)(有时也称为整合生物加工(CBP))进行。SHF 使用分开的处理步骤以首先将纤维素材料酶促水解为可发酵糖,例如,葡萄糖、纤维二糖、纤维三糖、以及戊糖,并且然后将这些可发酵糖发酵成乙醇。在 SSF 中,纤维素材料的酶水解和糖发酵成例如乙醇被组合在一个步骤中(菲利皮迪斯 • G. P. (Philippidis, G. P.), 1996, 纤维素生物转化技术 (Cellulose bioconversion technology), 生物乙醇手册:生产和利用 (Handbook on Bioethanol: Production and Utilization), 怀曼 • C. E. (Wyman, C. E.) 编辑, 泰勒 - 弗朗西斯出版集团 (Taylor & Francis), 华盛顿特区 (Washington, DC), 179-212)。SSCF 涉及多种糖的共发酵(希恩 (Sheehan) 和希默尔 (Himmel), 1999, 酶、能量与环境:美国能源部研究和开发生物乙醇活性的战略性观点 (Enzymes, energy and the environment: A strategic perspective on the U.S. Department of Energy's research and development activities for bioethanol), 生物技术进展 (Biotechnol. Prog.) 15:817-827)。HHF 涉及一个分开的水解步骤,并且另外涉及一个同时糖化和水解步骤,这些步骤可以在同一反应器中进行。HHF 过程中的步骤可以在不同的温度下进行,即高温酶糖化,接着在发酵菌株能够耐受的更低温度下进行 SSF。DMC 将全部三个过程(酶生产、水解、和发酵)组合在一个或多个(若干个)步骤中,其中使用了相同的微生物来产生用于转化纤维素材料为可发酵糖并且用于转化可发酵糖为终产物的酶(林德 (Lynd) 等人, 2002, 微生物纤维素利用:基本原理和生物技术 (Microbial cellulose utilization: Fundamentals and biotechnology), 微生物学与分子生物学评论 (Microbiol. Mol. Biol. Reviews) 66:506-577)。在此应理解的是,本领域中已知的包括预处理、酶水解(糖化)、发酵、或其组合的任何方法,可以用于实施本发明的方法。

[0178] 使用的常规装置包括补料分批搅拌反应器、分批搅拌反应器、具有超过滤的连续流搅拌反应器、和 / 或连续活塞流动柱反应器(费尔南达德卡斯蒂约斯柯瑞芝 (Fernanda de Castilhos Corazza) 等人, 2003, 用于纤维二糖水解的补料分批反应器中的最佳控制 (Optimal control in fed-batch reactor for the cellobiose hydrolysis), 技术学报

(Acta Scientiarum. Technology) 25:33-38; 古萨科夫 (Gusakov) 和辛涅特西 (Sinityn), 1985, 纤维素酶水解动力学: 1. 用于分批反应器过程的数学模型 (Kinetics of the enzymatic hydrolysis of cellulose: 1. A mathematical model for a batch reactor process), 酶与微生物技术 (Enz. Microb. Technol.) 7:346-352), 研磨反应器 (隆 (Ryu) 和李 (Lee), 1983, 通过使用研磨生物反应器, 生物转化废物纤维素 (Bioconversion of waste cellulose by using an attrition bioreactor), 生物技术和生物工程 (Biotechnol. Bioeng) 25:53-65), 或具有电场诱导的有力搅拌的反应器 (古萨科夫 (Gusakov) 等人, 1996, 使用具有电场诱导的有力搅拌的新颖类型的生物反应器的纤维素酶水解的增强 (Enhancement of enzymatic cellulose hydrolysis using a novel type of bioreactor with intensive stirring induced by electromagnetic field), 应用生物化学与生物技术 (Appl. Biochem. Biotechnol.) 56:141-153)。另外的反应器类型包括: 用于水解和 / 或发酵的流化床反应器、升流层 (upflow blanket) 反应器、固定化反应器、以及挤出机型反应器。

[0179] 根据本发明, 可以使用一种微生物 (例如酵母或细菌) 进行发酵。在一个实施例中, 该发酵微生物能够将己糖和 / 或戊糖发酵成所希望的发酵产品。在一个优选实施例中, 该发酵微生物是酵母, 例如酵母属的菌株, 例如酿酒酵母菌株。适合的发酵微生物的实例可以发现于下面的“发酵微生物”部分中。

[0180] 在一个实施例中, 在约 26°C 至约 60°C 之间的温度, 例如约 32°C 或 50°C, 和约 pH 3 至约 pH 8, 例如 pH 4-5、6 或 7 下进行发酵。

[0181] 当该发酵微生物是酵母 (例如酵母属的菌株, 特别是酿酒酵母菌株) 时, 可以在从 20°C -40°C 的温度, 例如 26°C -34°C, 优选大约 32°C 下进行发酵, 尤其是当该所希望的发酵产品是乙醇时。在一个实施例中, 在 pH 3-7, 例如 pH 4-6 下进行发酵。在一个实施例中, 发酵进行约 12 至约 96 小时, 例如典型地是 24-60 小时。在一个优选实施例中, 该发酵产品是一种醇, 例如乙醇。

[0182] 纤维素材料

[0183] 在本发明的方法或工艺中使用的该纤维素材料可以是任何含纤维素材料。生物质的初生细胞壁中的主要多糖是纤维素, 第二丰富的是半纤维素, 而第三丰富的是果胶。细胞停止生长后产生的次生细胞壁也包含多糖, 并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, 因此是一种线性 β -(1-4)-D- 葡聚糖, 而半纤维素包括多种化合物, 如具有一系列取代基以复杂支链结构存在的木聚糖、木葡聚糖、阿拉伯糖基木聚糖、以及甘露聚糖。尽管纤维素一般为多态的, 但发现其在植物组织中主要以平行葡聚糖链的不溶性晶体基质存在。半纤维素通常氢键合至纤维素以及其他半纤维素, 这有助于稳定细胞壁基质。

[0184] 纤维素通常见于例如植物的茎、叶、壳、皮、以及穗轴, 或树的叶、枝、以及木材中。纤维素材料可以是, 但不限于: 草本材料、农业废弃物、林业废弃物、城市固体废物、废纸、以及纸浆和造纸厂废弃物 (参见, 例如, 维色洛戈尔 (Wiselogle) 等人, 1995, 于生物乙醇手册 (Handbook on Bioethanol) (查尔斯·E·怀曼 (Charles E. Wyman) 编辑), 第 105-118 页, 泰勒弗朗西斯出版集团 (Taylor&Francis), 华盛顿特区 (Washington D.C.); 怀曼 (Wyman), 1994, 生物资源技术 (Bioresource Technology) 50:3-16; 林德 (Lynd), 1990, 应

用生物化学与生物技术 (Applied Biochemistry and Biotechnology) 24/25:695-719; 莫思尔 (Mosier) 等人, 1999, 木质纤维素的生物转化的最近进展 (Recent Progress in Bioconversion of Lignocellulosics), 生物化学工程 / 生物技术的进展 (Advances in Biochemical Engineering/Biotechnology), T·谢伯 (T. Scheper) 主编, 第 65 卷, 第 23-40 页, 纽约斯普林格出版社 (Springer-Verlag, New York)。在此应该理解的是, 纤维素可以处于木素纤维素, 在混合基质中包含木质素、纤维素、和半纤维素的植物细胞壁材料的形式。在一个优选方面中, 该纤维素材料是木质纤维素。

[0185] 在一个方面中, 该纤维素材料是草本材料。在另一个方面中, 该纤维素材料是农业废弃物。在另一个方面中, 该纤维素材料是林业废弃物。在另一个方面中, 该纤维素材料是城市固体废物。在另一个方面中, 该纤维素材料是废纸。在另一个方面中, 该纤维素材料是纸浆和造纸厂废弃物。

[0186] 在另一个方面中, 该纤维素材料是玉米秸秆。在另一个方面中, 该纤维素材料是玉米纤维。在另一个方面中, 该纤维素材料是玉米芯。在另一个方面中, 该纤维素材料是橙皮。在另一个方面中, 该纤维素材料是稻秸。在另一个方面中, 该纤维素材料是麦秸。在另一个方面中, 该纤维素材料是柳枝稷。在另一个方面中, 该纤维素材料是芒属。在另一个方面中, 该纤维素材料是甘蔗渣。

[0187] 在另一个方面中, 该纤维素材料是微晶纤维素。在另一个方面中, 该纤维素材料是细菌纤维素。在另一个方面中, 该纤维素材料是海藻纤维素。在另一个方面中, 该纤维素材料是棉短绒。在另一个方面中, 该纤维素材料是无定形磷酸处理的纤维素。在另一个方面中, 该纤维素材料是滤纸。

[0188] 该纤维素材料可以按原样使用或可以使用本领域已知的常规方法进行预处理, 如在此所描述。在一个优选方面中, 该纤维素材料进行了预处理。

[0189] 在一个优选实施例中, 该预处理的纤维素材料是预处理的玉米秸秆或作为用热和稀硫酸处理的玉米秸秆的“PCS”。

[0190] 预处理

[0191] 在实践本发明的方法或工艺中, 制备该预处理的纤维素材料, 可以使用本领域已知的任何预处理工艺来破坏植物细胞壁的纤维素材料组分 (钱德拉 (Chandra) 等人, 2007, 底物预处理: 木质纤维素的有效酶水解的关键? (Substrate pretreatment: The key to effective enzymatic hydrolysis of lignocellulosics?), 生物化学工程 / 生物技术的进展 (Adv. Biochem. Engin. / Biotechnol.) 108:67-93; 盖博 (Galbe) 和小林 (Zacchi), 2007, 用于高效生物乙醇生产的木质纤维素材料的预处理 (Pretreatment of lignocellulosic materials for efficient bioethanol production), 生物化学工程 / 生物技术的进展 108:41-65; 亨德里克斯 (Hendriks) 和季曼 (Zeeman), 2009, 用以增强木质纤维素生物质的消化性的预处理 (Pretreatments to enhance the digestibility of lignocellulosic biomass), 生物资源技术 (Bioresource Technol.) 100:10-18; 莫热 (Mosier) 等人, 2005, 用于木质纤维素生物质的预处理的有前景技术的特征 (Features of promising technologies for pretreatment of lignocellulosic biomass), 生物资源技术 96:673-686; 塔希尔扎德 (Taherzadeh) 和卡里米 (Karimi), 2008, 预处理木质纤维素废弃物以改善乙醇和生物气体产生: 综述 (Pretreatment of lignocellulosic wastes

to improve ethanol and biogas production:A review), 分子科学国际杂志 (Int. J. of Mol. Sci.) 9:1621-1651; 杨 (Yang) 和怀曼, 2008, 预处理: 释放低成本纤维素乙醇的关键 (Pretreatment: the key to unlocking low-cost cellulosic ethanol), 生物燃料、生物产品与生物精炼 (Biofuels Bioproducts and Biorefining-Biofpr.) 2:26-40)。

[0192] 纤维素材料也可以在预处理之前使用本领域中已知的方法进行颗粒尺寸减缩、预浸泡、润湿、洗涤、或调理。

[0193] 常规预处理包括但不限于蒸汽预处理 (用或不用爆发)、稀酸预处理、热水预处理、碱预处理、石灰预处理、湿氧化、湿爆发、氨纤维爆发、有机溶剂预处理、和生物预处理。另外的预处理包括氨渗滤、超声、电穿孔、微波、超临界 CO₂、超临界 H₂O、臭氧、以及 γ 辐射预处理。

[0194] 可以在水解和 / 或发酵之前对纤维素材料进行预处理。优选在水解前进行预处理。可替代地, 可以同时用酶水解进行预处理, 以释放可发酵糖, 例如葡萄糖、木糖、和 / 或纤维二糖。在多数情况下, 预处理步骤自身导致将生物质转化为可发酵糖 (即使在没有酶的情况下)。

[0195] 蒸汽预处理:

[0196] 在蒸汽预处理中, 加热纤维素材料以破坏植物细胞壁组分, 包括木质素、半纤维素、以及纤维素, 以使酶可接触纤维素和其他级分, 例如, 半纤维素。使纤维素材料经过或通过反应容器, 其中注入蒸汽以将温度增加至所需的温度和压力, 并且在其中保持所希望的反应时间。蒸汽预处理优选在 140 °C -230 °C, 更优选 160 °C -200 °C, 并且最优选 170 °C -190 °C 下执行, 其中最佳温度范围取决于化学催化剂的任意添加。蒸汽预处理的停留时间优选是 1-15 分钟, 更优选 3-12 分钟, 并且最优选 4-10 分钟, 其中最佳停留时间取决于温度范围和化学催化剂的任意添加。蒸汽预处理允许相对较高的固体加载量, 这样使得纤维素材料在预处理过程中通常仅变得潮湿。蒸汽预处理经常与预处理后的材料的爆发放料组合, 这被称为蒸汽爆发, 即, 迅速急骤蒸发至大气压和材料湍流, 以通过破碎增加可及的表面积 (达夫 (Duff) 和默里 (Murray), 1996, 生物资源技术 855:1-33; 盖博 (Galbe) 和小林 (Zacchi), 2002, 应用微生物学与生物技术 (Appl. Microbiol. Biotechnol.) 59:618-628; 美国专利申请号 2002/0164730)。在蒸汽预处理过程中, 半纤维素乙酰基被裂解, 并且所得到的酸自催化半纤维素部分水解成单糖和寡糖。仅在有限的程度上去除木质素。

[0197] 经常在蒸汽预处理之前添加催化剂, 如 H₂SO₄ 或 SO₂ (典型地 0.3% 至 3% w/w), 这降低时间和温度、增加回收率、并且改善酶水解 (巴列斯特罗斯 (Ballesteros) 等人, 2006, 应用生物化学与生物技术 129-132:496-508; 瓦尔加 (Varga) 等人, 2004, 应用生物化学与生物技术 113-116:509-523; 萨斯那 (Sassner) 等人, 2006, 酶与微生物技术 (Enzyme Microb. Technol.) 39:756-762)。

[0198] 化学预处理:

[0199] 术语“化学处理”是指能促进纤维素、半纤维素、和 / 或木质素的分离和 / 或释放的任何化学预处理。适合的化学预处理方法的实例包括: (例如) 稀酸预处理、石灰预处理、湿氧化、氨纤维 / 冷冻爆发 (AFEX)、氨渗滤 (APR)、以及有机溶剂预处理。

[0200] 在稀酸预处理中, 将纤维素材料与稀酸 (典型地 H₂SO₄) 和水混合以形成浆料, 通过蒸汽加热至所希望的温度, 并且在停留时间后急骤蒸发至大气压。可以用许多反应器设

计来进行稀酸预处理,例如,活可采用多种反应器设计来进行稀酸预处理,例如活塞流反应器、逆流反应器或连续逆流收缩床反应器(达夫(Duff)和默里(Murray),1996,同上;谢尔(Schell)等人,2004,生物资源技术(Bioresource Technol.)91:179-188;李(Lee)等人,1999,生物化学工程/生物技术进展(Adv. Biochem. Engin./Biotechnol.)65:93-115)。

[0201] 还可以使用在碱性条件下的几种预处理方法。这些碱性预处理包括但不限于:石灰预处理、湿氧化、氨渗滤(APR)、以及氨纤维/冷冻爆发(AFEX)。

[0202] 用碳酸钙、氢氧化钠、或氨,在85℃-150℃的低温下进行石灰预处理,并且停留时间为从1小时到几天(怀曼(Wyman)等人,2005,生物资源技术(Bioresource Technol.)96:1959-1966;莫热(Mosier)等人,2005,生物资源技术96:673-686)。WO 2006/110891、WO 2006/110899、WO 2006/110900、以及WO 2006/110901披露了使用氨的预处理方法。

[0203] 湿氧化是一种热预处理,其典型地在添加氧化剂(如过氧化氧或过压氧)的情况下在180℃-200℃下持续5-15分钟进行(施密特(Schmidt)和汤姆森(Thomsen),1998,生物资源技术64:139-151;帕罗内(Palonen)等人,2004,应用生物化学与生物技术117:1-17;瓦尔加等人,2004,生物技术与生物工程(Biotechnol. Bioeng.)88:567-574;马丁(Martin)等人,2006,化学技术与生物技术杂志(J. Chem. Technol. Biotechnol.)81:1669-1677)。预处理优选以1%至40%干物质,更优选2%至30%干物质,并且最优选5%至20%干物质进行,并且经常通过添加碱如碳酸钠来增加初始pH。

[0204] 被称为湿爆发(湿氧化和蒸汽爆发的组合)的湿氧化预处理方法的修改方案能够处理高达30%的干物质。在湿爆发中,在某一停留时间后,在预处理期间引入氧化试剂。然后通过急骤蒸发至大气压结束预处理(WO2006/032282)。

[0205] 氨纤维爆发(AFEX)涉及在如90℃-100℃的中等温度和如17至20巴的高压下,用液体或气态氨处理纤维素材料5至10分钟,其中干物质含量可以高达60%(格拉帕里(Gollapalli)等人,2002,应用生物化学与生物技术(Appl. Biochem. Biotechnol.)98:23-35;俊达瓦特(Chundawat)等人,2007,生物技术与生物工程(Biotechnol. Bioeng.)96:219-231;阿里扎德(Alizadeh)等人,2005,应用生物化学与生物技术121:1133-1141;泰莫里(Teymouri)等人,2005,生物资源技术(Bioresource Technol.)96:2014-2018)。AFEX预处理导致纤维素的解聚和半纤维素的部分水解。木质素-碳水化合物复合物被裂解。

[0206] 有机溶剂预处理通过使用含水乙醇(40%-60%乙醇)在160℃-200℃下萃取30-60分钟而将纤维素材料脱木质素(潘(Pan)等人,2005,生物技术与生物工程(Biotechnol. Bioeng.)90:473-481;潘等人,2006,生物技术与生物工程94:851-861;库拉比(Kurabi)等人,2005,应用生物化学与生物技术(Appl. Biochem. Biotechnol.)121:219-230)。通常加入硫酸作为催化剂。在有机溶剂预处理中,大部分半纤维素被去除。

[0207] 适合的预处理方法的其他实例由谢尔(Schell)等人,2003,应用生物化学与生物技术(Appl. Biochem. and Biotechnol.)105-108:69-85,和马塞尔(Mosier)等人,2005,生物资源技术(Bioresource Technology)96:673-686,以及美国专利申请2002/0164730进

行描述。

[0208] 在一个方面中,化学预处理优选是以酸处理进行,并且更优选持续稀酸和 / 或弱酸处理。酸通常是硫酸,但也可以使用其他酸,如乙酸、柠檬酸、硝酸、磷酸、酒石酸、琥珀酸、氯化氢、或其混合物。弱酸处理优选在 1-5,更优选 1-4,并且最优选 1-3 的 pH 范围内进行。在一个方面中,该酸浓度在优选从 0.01 至 20wt. % 酸,更优选 0.05 至 10wt. % 酸,甚至更优选 0.1 至 5wt. % 酸,并且最优选 0.2 至 2.0wt. % 酸的范围中。使酸与纤维素材料相接触,并在优选 160℃ -220℃,并且更优选 165℃ -195℃ 范围内的温度下保持从几秒到几分钟,例如 1 秒至 60 分钟范围内的时间。

[0209] 在另一个方面中,预处理是作为氨纤维爆发步骤 (AFEX 预处理步骤) 进行。

[0210] 在另一个方面中,预处理在水性浆料中进行。在优选的方面中,在预处理过程中纤维素材料以优选 10-80wt. % 之间,更优选 20-70wt. % 之间,并且最优选 30-60wt. % 之间,例如 50wt. % 左右的量存在。预处理的纤维素材料可以不洗涤或使用本领域已知的任何方法洗涤,例如,用水洗涤。

[0211] 机械预处理:

[0212] 术语“机械预处理”是指各种类型的研磨或碾磨 (例如,干磨、湿磨、或振动球磨)。

[0213] 物理预处理:

[0214] 术语“物理预处理”是指促进纤维素、半纤维素、和 / 或木质素从纤维素材料中分离和 / 或释放的任何预处理。例如,物理预处理可包括辐射 (例如微波辐射)、汽蒸 / 蒸汽爆发、水热解及其组合。

[0215] 物理预处理可以包括高压和 / 或高温 (蒸汽爆发)。在一个方面中,高压意指在优选约 300 至约 600psi,更优选约 350 至约 550psi,并且最优选约 400 至约 500psi,例如 450psi 左右的范围中的压力。在另一个方面中,高温意指在约 100℃ 至约 300℃,优选约 140℃ 至约 235℃ 范围中的温度。在一个优选方面,在分批过程中,以蒸汽枪水解器系统 (例如从顺智公司 (Sunds Defibrator AB),瑞典 (Sweden) 可获得的 Sunds 水解器 (Sunds Hydrolyzer)) 来进行机械预处理,该系统使用如上所定义的高压和高温。

[0216] 组合的物理和化学预处理:

[0217] 纤维素材料可以物理地且化学地预处理。例如,该预处理步骤可包括稀酸或弱酸预处理以及高温和 / 或高压处理。这些物理预处理和化学预处理可以根据需要顺序地进行或同时进行。也可以包括机械预处理。

[0218] 因此,在一个优选方面中,使纤维素材料经受机械、化学或物理预处理、或其任何组合,以促进纤维素、半纤维素、和 / 或木质素的分离和 / 或释放。

[0219] 生物预处理:

[0220] 术语“生物预处理”是指促进纤维素、半纤维素、和 / 或木质素从纤维素材料中分离和 / 或释放的任何生物预处理。生物预处理技术可以涉及应用溶解木质素的微生物 (参见,例如,舒 • T. -A. (Hsu, T. -A.), 1996, 生物质的预处理 (Pretreatment of biomass), 生物乙醇手册:生产和利用,怀曼 • C. E. 编辑,泰勒 - 弗朗西斯出版集团,华盛顿特区,179-212;高希 (Ghosh) 和辛格 (Singh), 1993, 用于纤维素生物质的酶 / 微生物转化的物理化学与生物处理 (Physicochemical and biological treatments for enzymatic/microbial conversion of cellulosic biomass), 应用微生物学进展 (Adv.

Appl. Microbiol.) 39:295-333; 麦克米兰 • J. D. (McMillan, J. D.), 1994, 预处理木质纤维素生物质: 综述 (Pretreating lignocellulosic biomass: a review), 用于燃料生产的生物质的酶转化 (Enzymatic Conversion of Biomass for Fuels Production), 希默尔 • M. E.、贝克 • J. O.、以及奥弗伦 • R. P. (Overend, R. P.) 编辑, 美国化学学会讨论会系列 566 (ACS Symposium Series 566), 美国化学学会 (American Chemical Society), 华盛顿特区, 第 15 章; 贡 • C. S. (Gong, C. S.)、卡奥 • N. J. (Cao, N. J.)、杜 • J. (Du, J.)、以及曹 • G. T. (Tsao, G. T.), 1999, 由可再生资源生产乙醇 (Ethanol production from renewable resources), 生物化学工程 / 生物技术的进展, 舍佩尔 • T. 编辑, 施普林格出版社德国海德堡柏林 (Berlin Heidelberg, Germany), 65:207-241; 奥尔森 (Olsson) 和哈恩-哈格达尔 (Hahn-Hagerdal), 1996, 用于乙醇生产的木质纤维素水解物的发酵 (Fermentation of lignocellulosic hydrolysates for ethanol production), 酶与微生物技术 (Enz. Microb. Tech.) 18:312-331; 以及瓦蓝德 (Vallander) 和埃里克松 (Eriksson), 1990, 由木质纤维素材料生产乙醇: 技术现状 (Production of ethanol from lignocellulosic materials: State of the art), 生物化学工程 / 生物技术的进展 42:63-95。

[0221] 发酵

[0222] 根据本发明的方法和工艺, 可发酵糖从水解预处理的纤维素材料获得。通过能够将糖直接或间接发酵 / 转化为所希望的发酵产品的一种或多种 (若干种) 发酵微生物可以发酵所述糖。术语“发酵”是指包括发酵步骤的任何工艺。发酵条件取决于所希望的发酵产品和发酵微生物。发酵条件可以由本领域普通技术人员容易地确定。

[0223] 在发酵步骤中, 水解导致的从该预处理的纤维素材料释放的糖被发酵微生物 (例如酵母) 发酵成一种所希望的产品, 例如乙醇。水解 (糖化) 和发酵可以是分开的 (SHF) 或同时的 (SSF), 或如上所述的。

[0224] 在实践本发明中, 任何适合的水解的预处理的纤维素材料都可以用于发酵中。一般基于所希望的发酵产品来选择该材料。

[0225] 术语“发酵培养基”在此可理解为指在添加发酵微生物之前的培养基, 例如, 由水解产生的培养基, 以及在同时糖化和发酵 (SSF) 中使用的培养基。

[0226] 发酵微生物

[0227] 根据本发明的工艺, 使用一种或多种发酵微生物来根据本发明的方法发酵 / 转化通过水解预处理的纤维素材料而产生的糖。术语“发酵微生物”是指适用于本发明的工艺的任何微生物, 包括细菌和真菌生物体。发酵微生物可以是 C_6 或 C_5 发酵微生物、或其组合。 C_6 和 C_5 发酵微生物二者都是本领域熟知的。适合的发酵微生物能够将糖 (如葡萄糖、木糖、木酮糖、阿拉伯糖、麦芽糖、甘露糖、半乳糖、或寡糖) 直接或间接地发酵 (即, 转化) 成所希望的发酵产品。

[0228] 产生乙醇的细菌和真菌发酵微生物的实例由林 (Lin) 等人, 2006, 应用微生物学与生物技术 (Appl. Microbiol. Biotechnol.) 69:627-642 描述。

[0229] 能够发酵 C_6 糖的发酵微生物的实例包括细菌生物和真菌生物, 如酵母。优选的酵母包括酵母菌属菌株, 优选酿酒酵母。

[0230] 能够发酵 C_5 糖的发酵微生物的实例包括细菌生物和真菌生物, 如酵母。优选的 C_5 发酵酵母包括毕赤酵母属的菌株, 优选树干毕赤酵母, 例如树干毕赤酵母 CBS 5773;

假丝酵母属的菌株,优选博伊丁假丝酵母 (*Candida boidinii*)、芸薹假丝酵母 (*Candida brassicae*)、休哈塔假丝酵母 (*Candida sheatae*)、迪丹斯假丝酵母 (*Candida diddensii*)、假热带假丝酵母 (*Candida pseudotropicalis*)、或产朊假丝酵母 (*Candida utilis*)。

[0231] 其他发酵微生物包括发酵单胞菌属的菌株,例如运动发酵单胞菌;汉逊酵母属,例如异常汉森酵母;克鲁维酵母属,例如脆壁克鲁维酵母 (*K. fragilis*);裂殖酵母属,例如粟酒裂殖酵母 (*S. pombe*);以及大肠杆菌,尤其已被遗传修饰以提高乙醇产率的大肠杆菌菌株。

[0232] 在一个优选方面中,酵母是酵母菌属种。在一个更优选方面中,酵母是酿酒酵母。在另一个更优选方面中,酵母是糖化酵母 (*Saccharomyces distaticus*)。在另一个更优选方面中,酵母是葡萄汁酵母 (*Saccharomyces uvarum*)。在另一个优选方面中,酵母是克鲁维酵母属。在另一个更优选方面中,酵母是马克斯克鲁维酵母。在另一个更优选方面中,酵母是脆壁克鲁维酵母。在另一个优选方面中,酵母是假丝酵母属。在另一个更优选方面中,酵母是博伊丁假丝酵母。在另一个更优选方面中,酵母是芸薹假丝酵母。在另一个更优选方面中,酵母是迪丹斯假丝酵母。在另一个更优选方面中,酵母是假热带假丝酵母。在另一个更优选方面中,酵母是产朊假丝酵母。在另一个优选方面中,酵母是棒孢酵母属 (*Clavispora*)。在另一个更优选方面中,酵母是葡萄牙棒孢酵母 (*Clavispora lusitaniae*)。在另一个更优选方面中,酵母是仙人掌棒孢酵母 (*Clavispora opuntiae*)。在另一个优选方面中,酵母是管囊酵母属。在另一个更优选方面中,酵母是嗜鞣管囊酵母。在另一个优选方面中,酵母是毕赤酵母属。在另一个优选方面中,酵母是树干毕赤酵母。在另一个优选方面中,酵母是酒香酵母属 (*Bretanomyces*)。在另一个更优选方面中,酵母是克劳森酒香酵母 (菲利皮迪斯·G.P. (Philippidis, G.P.), 1996, 纤维素生物转化技术 (Cellulose bioconversion technology), 生物乙醇手册:生产和利用 (Handbook on Bioethanol: Production and Utilization), 怀曼·C.E (Wyman, C.E.) 编辑, 泰勒-弗朗西斯出版集团 (Taylor&Francis), 华盛顿特区, 179-212)。

[0233] 能够有效地将己糖和戊糖发酵成乙醇的细菌包括,例如,运动发酵单胞菌和热纤维梭菌 (*Clostridium thermocellum*) (菲利皮迪斯, 1996, 见上文)。

[0234] 在一个优选方面中,细菌是发酵单胞菌属。在一个更优选方面中,细菌是运动发酵单胞菌。在另一个优选方面中,细菌是梭菌属。在另一个更优选方面中,细菌是热纤维梭菌。

[0235] 适于乙醇生产的可商购的酵母包括,例如乙醇红酵母 (ETHANOL RED™ yeast) (可从富酶泰斯/乐斯富 (Fermentis/Lesaffre), USA 获得)、FALI™ (可从弗莱施曼酵母 (Fleischmann's Yeast), USA 获得)、SUPERSTART™ 和 THERMOSACC™ 新鲜酵母 (可从乙醇技术 (Ethanol Technology), WI, USA)、BIOFERM™ AFT 和 XR (可从 NABC- 北美生物产品公司 (North American Bioproducts Corporation), GA, USA 获得)、GERT STRAND™ (可从 Gert Strand AB, 瑞典获得)、以及 FERMIOL™ (可从帝斯曼食品配料部 (DSM Specialties) 获得)。

[0236] 在一个优选方面中,发酵微生物已经经过遗传修饰,以提供发酵戊糖的能力,如利用木糖的微生物、利用阿拉伯糖的微生物、以及共同利用木糖和阿拉伯糖的微生物。

[0237] 将异源基因克隆到多种发酵微生物中已经构建出能够将己糖和戊糖转化成乙醇 (共发酵) 的微生物 (陈 (Chen) 和霍 (Ho), 1993, 酿酒酵母中毕赤酵母木糖还原酶基因的克隆和改善表达 (Cloning and improving the expression of *Pichia stipitis*

xylose reductase gene in *Saccharomyces cerevisiae*), 应用生物化学与生物技术 (Appl. Biochem. Biotechnol.) 39-40:135-147; 霍等人, 1998, 能够有效共发酵葡萄糖和木糖的遗传工程化的酵母菌属酵母 (Genetically engineered *Saccharomyces* yeast capable of effectively cofermenting glucose and xylose), 应用与环境微生物学 (Appl. Environ. Microbiol.) 64:1852-1859; 科特 (Kotter) 和西拉塞 (Ciriacy), 1993, 酿酒酵母发酵木糖 (Xylose fermentation by *Saccharomyces cerevisiae*), 应用微生物学与生物技术 (Appl. Microbiol. Biotechnol.) 38:776-783; 维尔福森 (Walfridsson) 等人, 1995, 过表达编码戊糖磷酸途径酶转酮醇酶和转醛醇酶的 TKL1 和 TAL1 基因の木糖代谢酿酒酵母菌株 (Xylose-metabolizing *Saccharomyces cerevisiae* strains overexpressing the TKL1 and TAL1 genes encoding the pentose phosphate pathway enzymes transketolase and transaldolase), 应用与环境微生物学 61:4184-4190; 凯珀 (Kuyper) 等人, 2004, 用于有效厌氧木糖发酵的酿酒酵母的最小代谢工程化: 原理的证实 (Minimal metabolic engineering of *Saccharomyces cerevisiae* for efficient anaerobic xylose fermentation: a proof of principle), 欧洲微生物学会联合会酵母研究 (FEMS Yeast Research) 4:655-664; 比尔 (Beall) 等人, 1991, 通过重组大肠杆菌从木糖和其他糖产生乙醇的参数研究 (Parametric studies of ethanol production from xylose and other sugars by recombinant *Escherichia coli*), 生物技术与生物工程 (Biotech. Bioeng.) 38:296-303; 英格拉姆 (Ingram) 等人, 1998, 用于乙醇产生的细菌的代谢工程化 (Metabolic engineering of bacteria for ethanol production), 生物技术与生物工程 58:204-214; 张 (Zhang) 等人, 1995, 产生乙醇的运动发酵单胞菌中戊糖代谢途径的代谢工程化 (Metabolic engineering of a pentose metabolism pathway in ethanologenic *Zymomonas mobilis*), 科学 (Science) 267:240-243; 狄安妲 (Deanda) 等人, 1996, 通过代谢途径工程化开发发酵阿拉伯糖的运动发酵单胞菌菌株 (Development of an arabinose-fermenting *Zymomonas mobilis* strain by metabolic pathway engineering), 应用与环境微生物学 62:4465-4470; WO 03/062430, 木糖异构酶)。

[0238] 在一个实施例中, C₅发酵微生物是一种酿酒酵母修饰菌株, 包括如披露于 WO 03/062340、WO 2004/099381 或 WO 2006/009434 中的木糖异构酶基因。

[0239] 在一个优选方面中, 遗传修饰的发酵微生物是酿酒酵母。在另一个优选方面中, 遗传修饰的发酵微生物是运动发酵单胞菌。在另一个优选方面中, 遗传修饰的发酵微生物是大肠杆菌。在另一个优选方面中, 遗传修饰的发酵微生物是产酸克雷伯氏菌 (*Klebsiella oxytoca*)。在另一个优选方面中, 遗传修饰的发酵微生物是克鲁维酵母属。

[0240] 本领域中熟知的是, 以上所描述的微生物还可以用于产生其他物质, 如在此所描述。

[0241] 典型地向降解的预处理的纤维素材料或水解产物中添加该发酵微生物, 并且发酵进行约 12 至约 96 小时, 例如约 24 至约 60 小时。温度典型地在约 26°C 至约 60°C 之间, 具体地约 32°C 或 50°C, 并且在约 pH 3 至约 pH 8, 例如 pH 4 至 5、6、或 7 左右。

[0242] 在一个优选方面中, 对降解的预处理的纤维素材料施用酵母和 / 或另一种微生物, 并且如上所述, 发酵进行约 12 至约 96 小时, 如典型地 24 至 60 小时。在一个优选方面中, 温度优选在约 20°C 至约 60°C 之间, 更优选约 25°C 至约 50°C, 并且最优选约 32°C 至约 50°C,

具体地约 32℃或 50℃,并且 pH 通常是从约 pH 3 至约 pH 7,优选地在 pH 4-7 左右。然而,一些发酵微生物(例如细菌)具有更高的最适发酵温度。

[0243] 酵母或另一种微生物优选以每 ml 发酵液大约 10^5 至 10^{12} ,优选从大约 10^7 至 10^{10} ,特别是大约 2×10^8 个活细胞计数的量施用。关于使用酵母进行发酵的进一步指南可以见于例如“醇教材”(“The Alcohol Textbook”)(K·雅克(K. Jacques)、T.P. 里昂(T.P. Lyons)以及 D.R. 凯尔索尔(D.R. Kelsall)编辑,诺丁汉大学出版社(Nottingham University Press),联合王国(United Kingdom)1999),其通过引用结合在此。

[0244] 对于乙醇生产,在发酵之后,蒸馏发酵的浆料以萃取乙醇。根据本发明的工艺获得的乙醇可用作例如燃料乙醇、饮用乙醇,即可饮用的浓酒精,或工业乙醇。

[0245] 发酵刺激剂可以与在此所描述的任何工艺组合使用,以进一步改进发酵,并且特别是改进发酵微生物的性能,如,提高速度和产品产率。“发酵刺激剂”是指用于发酵微生物(特别是酵母)生长的刺激剂。用于生长的优选发酵刺激剂包括维生素和矿物质。维生素的实例包括多种维生素、生物素、泛酸、烟酸、内消旋肌醇、硫胺、吡哆醇、对氨基苯酸、叶酸、核黄素、以及维生素 A、B、C、D 和 E,例如参见阿尔弗雷多(Alfenore)等人,通过在进料分批方法过程中的一种维生素进料策略改进乙醇产生和酿酒酵母的存活力(Improving ethanol production and viability of *Saccharomyces cerevisia* by a vitamin feeding strategy during fed-batch process),施普林格(2002),其通过引用结合在此。矿物质的实例包括可以供应包含 P、K、Mg、S、Ca、Fe、Zn、Mn 和 Cu 营养素的矿物质和矿物盐。

[0246] 发酵产品

[0247] (所希望的)发酵产品可以是来源于本发明的工艺(包括发酵步骤)的任何物质。该发酵产品可以是(不限制)一种醇(例如,阿拉伯糖醇、丁醇、乙醇、甘油、甲醇、1,3-丙二醇、山梨糖醇、以及木糖醇);一种有机酸(例如,乙酸、醋酸、己二酸、抗坏血酸、柠檬酸、2,5-二酮-D-葡萄糖酸、甲酸、反丁烯二酸、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、草酰乙酸、丙酸、琥珀酸、以及木糖酸);一种酮(例如,丙酮);一种氨基酸(例如,天冬氨酸、谷氨酸、甘氨酸、赖氨酸、丝氨酸、以及苏氨酸);以及一种气体(例如,甲烷、氢气(H_2)、二氧化碳(CO_2)、以及一氧化碳(CO))。发酵产品还可以是作为高价值产品的蛋白。

[0248] 在一个优选实施例中,该发酵产品是一种醇。应理解的是,术语“醇”涵盖包含一个或多个羟基部分的物质。在一个更优选方面中,该醇是阿拉伯糖醇。在另一个更优选方面中,该醇是丁醇。在另一个更优选方面中,该醇是乙醇。在另一个实施例中,该醇是甘油。在另一个优选实施例中,该醇是甲醇。在另一个更优选方面中,该醇是 1,3-丙二醇。在另一个更优选方面中,该醇是山梨糖醇。在另一个更优选方面中,该醇是木糖醇。参见,例如,贡·C.S. (Gong, C. S.)、卡奥·N. J. (Cao, N. J.,)、杜·J. (Du, J.)、以及曹·G. T. (Tsao, G. T.), 1999,由可再生资源生产乙醇,生物化学工程/生物技术的进展,舍佩尔·T. 编辑,施普林格,德国海德堡柏林,65:207-241;西尔韦拉·M. (Silveira, M.) 和乔纳斯(Jonas),2002,山梨糖醇的生物技术生产(The biotechnological production of sorbitol),应用微生物学与生物技术 59:400-408;尼加姆(Nigam)和辛格(Singh),1995,用于木糖醇-一种糖替代品的发酵生产工艺(Processes for fermentative production of xylitol-a sugar substitute),加工生物化学(Process Biochemistry)30(2):117-124;伊泽吉(Ezeji)

等人,2003,通过拜季林斯基羧菌 BA101 生产丙酮、丁醇和乙醇并且通过气提原位回收 (Production of acetone, butanol and ethanol by *Clostridium beijerinckii* BA101 and in situ recovery by gas stripping), 微生物与生物技术世界杂志 (World Journal of Microbiology and Biotechnology) 19(6):595-603。

[0249] 在另一个优选实施例中,该发酵产品是一种有机酸。在另一个更优选实施例中,该有机酸是乙酸。在另一个更优选实施例中,该有机酸是醋酮酸。在另一个更优选实施例中,该有机酸是己二酸。在另一个更优选实施例中,该有机酸是抗坏血酸。在另一个更优选实施例中,该有机酸是柠檬酸。在另一个更优选实施例中,该有机酸是 2,5-二酮-D-葡萄糖酸。在另一个更优选实施例中,该有机酸是甲酸。在另一个更优选实施例中,该有机酸是富马酸。在另一个更优选实施例中,该有机酸是葡萄糖二酸。在另一个更优选实施例中,该有机酸是葡萄糖酸。在另一个更优选实施例中,该有机酸是葡萄糖醛酸。在另一个更优选实施例中,该有机酸是戊二酸。在另一个更优选实施例中,该有机酸是 3-羟基丙酸。在另一个更优选实施例中,该有机酸是衣康酸。在另一个更优选实施例中,该有机酸是乳酸。在另一个更优选实施例中,该有机酸是苹果酸。在另一个更优选实施例中,该有机酸是丙二酸。在另一个更优选实施例中,该有机酸是草酸。在另一个更优选实施例中,该有机酸是丙酸。在另一个更优选实施例中,该有机酸是琥珀酸。在另一个更优选实施例中,该有机酸是木糖酸。参见例如陈 (Chen) 和李 (Lee), 1997, 用于从纤维素生物质生产乳酸的膜介导的萃取发酵 (Membrane-mediated extractive fermentation for lactic acid production from cellulosic biomass) 应用生物化学与生物技术 (Appl. Biochem. Biotechnol.), 63-65:435-448。

[0250] 在另一个优选实施例中,该发酵产品是一种酮。应理解的是,术语“酮”涵盖包含一个或多个酮部分的物质。在另一个更优选方面,该酮是丙酮。参见,例如,库雷希和布拉舍克,2003,见上文。

[0251] 在另一个优选实施例中,该发酵产品是一种氨基酸。在另一个更优选实施例中,该有机酸是天冬氨酸。在另一个更优选实施例中,该氨基酸是谷氨酸。在另一个更优选实施例中,该氨基酸是甘氨酸。在另一个更优选实施例中,该氨基酸是赖氨酸。在另一个更优选实施例中,该氨基酸是丝氨酸。在另一个更优选实施例中,该氨基酸是苏氨酸。参见例如理查德 (Richard) 和马加里蒂 (Margaritis), 2004, 用于聚(谷氨酸)生产和其他微生物生物聚合物的分批发酵动力学的实验建模 (Empirical modeling of batch fermentation kinetics for poly(glutamic acid) production and other microbial biopolymers), 生物技术和生物工程 (Biotechnology and Bioengineering) 87(4):501-515。

[0252] 在另一个优选实施例中,该发酵产品是一种气体。在另一个更优选实施例中,该气体是甲烷。在另一个更优选实施例中,该气体是 H_2 。在另一个更优选实施例中,该气体是 CO_2 。在另一个更优选实施例中,该气体是 CO。参见例如片冈 (Kataoka) 等人, 1997, 关于通过产氢气厌氧细菌的连续培养系统的氢气生产的研究 (Studies on hydrogen production by continuous culture system of hydrogen-producing anaerobic bacteria), 水科学与技术 (Water Science and Technology) 36(6-7):41-47; 以及古那森兰 (Gunaseelan), 1997, 用于甲烷生产的厌氧消化:一篇综述 (Anaerobic digestion of biomass for methane production: A review), 生物质与生物能源 (Biomass and Bioenergy) 13(1-2):83-114。

[0253] 回收

[0254] 可以使用本领域已知的任何方法,任选地从发酵中回收发酵产品,这些方法包括但不限于:层析法、电泳程序、差别溶解度、蒸馏、或萃取。例如,一种醇(例如,乙醇)可以与根据本发明水解并发酵的纤维素材料分离并且任选地通过蒸馏的常规方法纯化。可以获得具有高达约 96vol. % 的纯度的乙醇,这可以用作例如燃料乙醇、饮用乙醇,即可饮用的浓酒精、或工业乙醇。

[0255] 酶

[0256] 具有纤维素分解增强活性的多肽

[0257] 在本发明的用于降解预处理的纤维素材料的方法中,一种具有纤维素分解增强活性的多肽与一种纤维素分解酶组合物;一种过氧化物酶;以及一种非离子表面活性剂和/或一种阳离子表面活性剂一起存在或在水解过程中添加。

[0258] 在一个实施例中,具有纤维素分解增强活性的该多肽包括以下基序:

[0259] [ILMV]-P-X(4,5)-G-X-Y-[ILMV]-X-R-X-[EQ]-X(4)-[HNQ] 和 [FW]-[TF]-K-[AIV],

[0260] 其中X是任何氨基酸,X(4,5)是处于4个或5个连续位置的任何氨基酸,并且X(4)是处于4个连续位置的任何氨基酸。

[0261] 包括以上提到的基序的多肽可以进一步包括:

[0262] H-X(1,2)-G-P-X(3)-[YW]-[AILMV],

[0263] [EQ]-X-Y-X(2)-C-X-[EHQN]-[FILV]-X-[ILV]、或

[0264] H-X(1,2)-G-P-X(3)-[YW]-[AILMV] 以及 [EQ]-X-Y-X(2)-C-X-[EHQN]-[FILV]-X-[ILV],

[0265] 其中X是任何氨基酸,X(1,2)是处于1个位置或2个连续位置的任何氨基酸,X(3)是处于3个连续位置的任何氨基酸,并且X(2)是处于2个连续位置的任何氨基酸。在以上基序中,采用可接受的IUPAC单一字母氨基酸缩写。

[0266] 在一个优选实施例中,具有纤维素分解增强活性的该多肽进一步包括H-X(1,2)-G-P-X(3)-[YW]-[AILMV]。在另一个优选方面中,具有纤维素分解增强活性的分离多肽进一步包括[EQ]-X-Y-X(2)-C-X-[EHQN]-[FILV]-X-[ILV]。在另一个优选实施例中,具有纤维素分解增强活性的该多肽进一步包括H-X(1,2)-G-P-X(3)-[YW]-[AILMV]和[EQ]-X-Y-X(2)-C-X-[EHQN]-[FILV]-X-[ILV]。

[0267] 在另一个实施例中,具有纤维素分解增强活性的该多肽包括以下基序:

[0268] [ILMV]-P-x(4,5)-G-x-Y-[ILMV]-x-R-x-[EQ]-x(3)-A-[HNQ],

[0269] 其中x是任何氨基酸,x(4,5)是处于4个或5个连续位置的任何氨基酸,并且x(3)是处于3个连续位置的任何氨基酸。在以上基序中,采用可接受的IUPAC单一字母氨基酸缩写。

[0270] 在一个实施例中,具有纤维素分解增强活性的该多肽包括一个氨基酸序列,该氨基酸序列与SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:14、或SEQ ID NO:16的成熟多肽具有优选至少60%、更优选至少65%、更优选至少70%、更优选至少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%、并且甚至最优选至少96%、至少97%、至少98%、或至少99%的一

致性程度。在一个优选方面中,该成熟多肽序列是 SEQ ID NO:2 的氨基酸 20 至 326、SEQ ID NO:4 的氨基酸 18 至 239、SEQ ID NO:6 的氨基酸 20 至 258、SEQ ID NO:8 的氨基酸 19 至 226、SEQ ID NO:10 的氨基酸 20 至 304、SEQ ID NO:12 的氨基酸 16 至 317、SEQ ID NO:14 的氨基酸 23 至 250、或 SEQ ID NO:16 的氨基酸 20 至 249。

[0271] 具有纤维素分解增强活性的多肽优选包括 SEQ ID NO:2 的氨基酸序列或其等位基因变体;或其具有纤维素分解增强活性的片段。在一个优选方面中,该多肽包括 SEQ ID NO:2 的氨基酸序列。在另一个优选方面中,该多肽包括 SEQ ID NO:2 的成熟多肽。在另一个优选方面中,该多肽包括 SEQ ID NO:2 的氨基酸 20 至 326 或其等位基因变体;或其具有纤维素分解增强活性的片段。在另一个优选方面中,该多肽包括 SEQ ID NO:2 的氨基酸 20 至 326。在另一个优选方面中,该多肽由 SEQ ID NO:2 的氨基酸序列或其等位基因变体;或其具有纤维素分解增强活性的片段组成。在另一个优选方面中,该多肽由 SEQ ID NO:2 的氨基酸序列组成。在另一个优选方面中,该多肽由 SEQ ID NO:2 的成熟多肽多肽组成。在另一个优选方面中,该多肽由 SEQ ID NO:2 的氨基酸 20 至 326 或其等位基因变体;或其具有纤维素分解增强活性的片段组成。在另一个优选方面中,该多肽由 SEQ ID NO:2 的氨基酸 20 至 326 组成。

[0272] 具有纤维素分解增强活性的多肽优选包括 SEQ ID NO:4 的氨基酸序列或其等位基因变体;或其具有纤维素分解增强活性的片段。在一个优选方面中,该多肽包括 SEQ ID NO:4 的氨基酸序列。在另一个优选方面中,该多肽包括 SEQ ID NO:4 的成熟多肽。在另一个优选方面中,该多肽包括 SEQ ID NO:4 的氨基酸 18 至 239 或其等位基因变体;或其具有纤维素分解增强活性的片段。在另一个优选方面中,该多肽包括 SEQ ID NO:4 的氨基酸 18 至 239。在另一个优选方面中,该多肽由 SEQ ID NO:4 的氨基酸序列或其等位基因变体;或其具有纤维素分解增强活性的片段组成。在另一个优选方面中,该多肽由 SEQ ID NO:4 的氨基酸序列组成。在另一个优选方面中,该多肽由 SEQ ID NO:4 的成熟多肽多肽组成。在另一个优选方面中,该多肽由 SEQ ID NO:4 的氨基酸 18 至 239 或其等位基因变体;或其具有纤维素分解增强活性的片段组成。在另一个优选方面中,该多肽由 SEQ ID NO:4 的氨基酸 18 至 239 组成。

[0273] 具有纤维素分解增强活性的多肽优选包括 SEQ ID NO:6 的氨基酸序列或其等位基因变体;或其具有纤维素分解增强活性的片段。在一个优选方面中,该多肽包括 SEQ ID NO:6 的氨基酸序列。在另一个优选方面中,该多肽包括 SEQ ID NO:6 的成熟多肽。在另一个优选方面中,该多肽包括 SEQ ID NO:6 的氨基酸 20 至 258 或其等位基因变体;或其具有纤维素分解增强活性的片段。在另一个优选方面中,该多肽包括 SEQ ID NO:6 的氨基酸 20 至 258。在另一个优选方面中,该多肽由 SEQ ID NO:6 的氨基酸序列或其等位基因变体;或其具有纤维素分解增强活性的片段组成。在另一个优选方面中,该多肽由 SEQ ID NO:6 的氨基酸序列组成。在另一个优选方面中,该多肽由 SEQ ID NO:6 的成熟多肽多肽组成。在另一个优选方面中,该多肽由 SEQ ID NO:6 的氨基酸 20 至 258 或其等位基因变体;或其具有纤维素分解增强活性的片段组成。在另一个优选方面中,该多肽由 SEQ ID NO:6 的氨基酸 20 至 258 组成。

[0274] 具有纤维素分解增强活性的多肽优选包括 SEQ ID NO:8 的氨基酸序列或其等位基因变体;或其具有纤维素分解增强活性的片段。在一个优选方面中,该多肽包括 SEQ ID

NO:8 的氨基酸序列。在另一个优选方面中,该多肽包括 SEQ ID NO:8 的成熟多肽。在另一个优选方面中,该多肽包括 SEQ ID NO:8 的氨基酸 19 至 226 或其等位基因变体;或其具有纤维素分解增强活性的片段。在另一个优选方面中,该多肽包括 SEQ ID NO:8 的氨基酸 19 至 226。在另一个优选方面中,该多肽由 SEQ ID NO:8 的氨基酸序列或其等位基因变体;或其具有纤维素分解增强活性的片段组成。在另一个优选方面中,该多肽由 SEQ ID NO:8 的氨基酸序列组成。在另一个优选方面中,该多肽由 SEQ ID NO:8 的成熟多肽多肽组成。在另一个优选方面中,该多肽由 SEQ ID NO:8 的氨基酸 19 至 226 或其等位基因变体;或其具有纤维素分解增强活性的片段组成。在另一个优选方面中,该多肽由 SEQ ID NO:8 的氨基酸 19 至 226 组成。

[0275] 具有纤维素分解增强活性的多肽优选包括 SEQ ID NO:10 的氨基酸序列或其等位基因变体;或其具有纤维素分解增强活性的片段。在一个优选方面中,该多肽包括 SEQ ID NO:10 的氨基酸序列。在另一个优选方面中,该多肽包括 SEQ ID NO:10 的成熟多肽。在另一个优选方面中,该多肽包括 SEQ ID NO:10 的氨基酸 20 至 304 或其等位基因变体;或其具有纤维素分解增强活性的片段。在另一个优选方面中,该多肽包括 SEQ ID NO:10 的氨基酸 20 至 304。在另一个优选方面中,该多肽由 SEQ ID NO:10 的氨基酸序列或其等位基因变体;或其具有纤维素分解增强活性的片段组成。在另一个优选方面中,该多肽由 SEQ ID NO:10 的氨基酸序列组成。在另一个优选方面中,该多肽由 SEQ ID NO:10 的成熟多肽多肽组成。在另一个优选方面中,该多肽由 SEQ ID NO:10 的氨基酸 20 至 304 或其等位基因变体;或其具有纤维素分解增强活性的片段组成。在另一个优选方面中,该多肽由 SEQ ID NO:10 的氨基酸 20 至 304 组成。

[0276] 具有纤维素分解增强活性的多肽优选包括 SEQ ID NO:12 的氨基酸序列或其等位基因变体;或其具有纤维素分解增强活性的片段。在一个优选方面中,该多肽包括 SEQ ID NO:12 的氨基酸序列。在另一个优选方面中,该多肽包括 SEQ ID NO:12 的成熟多肽。在另一个优选方面中,该多肽包括 SEQ ID NO:12 的氨基酸 16 至 317 或其等位基因变体;或其具有纤维素分解增强活性的片段。在另一个优选方面中,该多肽包括 SEQ ID NO:12 的氨基酸 16 至 317。在另一个优选方面中,该多肽由 SEQ ID NO:12 的氨基酸序列或其等位基因变体;或其具有纤维素分解增强活性的片段组成。在另一个优选方面中,该多肽由 SEQ ID NO:12 的氨基酸序列组成。在另一个优选方面中,该多肽由 SEQ ID NO:12 的成熟多肽多肽组成。在另一个优选方面中,该多肽由 SEQ ID NO:12 的氨基酸 16 至 317 或其等位基因变体;或其具有纤维素分解增强活性的片段组成。在另一个优选方面中,该多肽由 SEQ ID NO:12 的氨基酸 16 至 317 组成。

[0277] 具有纤维素分解增强活性的多肽优选包括 SEQ ID NO:14 的氨基酸序列或其等位基因变体;或其具有纤维素分解增强活性的片段。在一个优选方面中,该多肽包括 SEQ ID NO:14 的氨基酸序列。在另一个优选方面中,该多肽包括 SEQ ID NO:14 的成熟多肽。在另一个优选方面中,该多肽包括 SEQ ID NO:14 的氨基酸 23 至 250 或其等位基因变体;或其具有纤维素分解增强活性的片段。在另一个优选方面中,该多肽包括 SEQ ID NO:14 的氨基酸 23 至 250。在另一个优选方面中,该多肽由 SEQ ID NO:14 的氨基酸序列或其等位基因变体;或其具有纤维素分解增强活性的片段组成。在另一个优选方面中,该多肽由 SEQ ID NO:14 的氨基酸序列组成。在另一个优选方面中,该多肽由 SEQ ID NO:14 的成熟多肽多肽组成。

在另一个优选方面中,该多肽由 SEQ ID NO:14 的氨基酸 23 至 250 或其等位基因变体;或其具有纤维素分解增强活性的片段组成。在另一个优选方面中,该多肽由 SEQ ID NO:14 的氨基酸 23 至 250 组成。

[0278] 具有纤维素分解增强活性的多肽优选包括 SEQ ID NO:16 的氨基酸序列或其等位基因变体;或其具有纤维素分解增强活性的片段。在一个优选方面中,该多肽包括 SEQ ID NO:16 的氨基酸序列。在另一个优选方面中,该多肽包括 SEQ ID NO:16 的成熟多肽。在另一个优选方面中,该多肽包括 SEQ ID NO:16 的氨基酸 20 至 249 或其等位基因变体;或其具有纤维素分解增强活性的片段。在另一个优选方面中,该多肽包括 SEQ ID NO:16 的氨基酸 20 至 249。在另一个优选方面中,该多肽由 SEQ ID NO:16 的氨基酸序列或其等位基因变体;或其具有纤维素分解增强活性的片段组成。在另一个优选方面中,该多肽由 SEQ ID NO:16 的氨基酸序列组成。在另一个优选方面中,该多肽由 SEQ ID NO:16 的成熟多肽多肽组成。在另一个优选方面中,该多肽由 SEQ ID NO:16 的氨基酸 20 至 249 或其等位基因变体;或其具有纤维素分解增强活性的片段组成。在另一个优选方面中,该多肽由 SEQ ID NO:16 的氨基酸 20 至 249 组成。

[0279] 优选地,SEQ ID NO:2 的成熟多肽的一个片段含有至少 277 个氨基酸残基,更优选至少 287 个氨基酸残基,并且最优选至少 297 个氨基酸残基。优选地,SEQ ID NO:4 的成熟多肽的一个片段含有至少 185 个氨基酸残基,更优选至少 195 个氨基酸残基,并且最优选至少 205 个氨基酸残基。优选地,SEQ ID NO:6 的成熟多肽的一个片段含有至少 200 个氨基酸残基,更优选至少 212 个氨基酸残基,并且最优选至少 224 个氨基酸残基。优选地,SEQ ID NO:8 的成熟多肽的一个片段含有至少 175 个氨基酸残基,更优选至少 185 个氨基酸残基,并且最优选至少 195 个氨基酸残基。优选地,SEQ ID NO:10 的成熟多肽的一个片段含有至少 240 个氨基酸残基,更优选至少 255 个氨基酸残基,并且最优选至少 270 个氨基酸残基。优选地,SEQ ID NO:12 的成熟多肽的一个片段含有至少 255 个氨基酸残基,更优选至少 270 个氨基酸残基,并且最优选至少 285 个氨基酸残基。优选地,SEQ ID NO:14 的成熟多肽的一个片段含有至少 175 个氨基酸残基,更优选至少 190 个氨基酸残基,并且最优选至少 205 个氨基酸残基。优选地,SEQ ID NO:16 的成熟多肽的一个片段含有至少 200 个氨基酸残基,更优选至少 210 个氨基酸残基,并且最优选至少 220 个氨基酸残基。

[0280] 优选地,SEQ ID NO:1 的成熟多肽编码序列的一个子序列含有至少 831 个核苷酸,更优选至少 861 个核苷酸,并且最优选至少 891 个核苷酸。优选地,SEQ ID NO:3 的成熟多肽编码序列的一个子序列含有至少 555 个核苷酸,更优选至少 585 个核苷酸,并且最优选至少 615 个核苷酸。优选地,SEQ ID NO:5 的成熟多肽编码序列的一个子序列含有至少 600 个核苷酸,更优选至少 636 个核苷酸,并且最优选至少 672 个核苷酸。优选地,SEQ ID NO:7 的成熟多肽编码序列的一个子序列含有至少 525 个核苷酸,更优选至少 555 个核苷酸,并且最优选至少 585 个核苷酸。优选地,SEQ ID NO:9 的成熟多肽编码序列的一个子序列含有至少 720 个核苷酸,更优选至少 765 个核苷酸,并且最优选至少 810 个核苷酸。优选地,SEQ ID NO:11 的成熟多肽编码序列的一个子序列含有至少 765 个核苷酸,更优选至少 810 个核苷酸,并且最优选至少 855 个核苷酸。优选地,SEQ ID NO:13 的核苷酸 67 至 796 的成熟多肽编码序列的一个子序列含有至少 525 个核苷酸,更优选至少 570 个核苷酸,并且最优选至少 615 个核苷酸。优选地,SEQ ID NO:15 的成熟多肽编码序列的一个子序列含有至少 600

个核苷酸,更优选至少 630 个核苷酸,并且最优选至少 660 个核苷酸。

[0281] 在一个第四方面中,具有纤维素分解增强活性的该多肽由以下多核苷酸编码,该多核苷酸在至少非常低严谨度条件、优选至少低严谨度条件、更优选至少中严谨度条件、更优选至少中-高严谨度条件、甚至更优选至少高严谨度条件、以及最优选至少非常高严谨度条件下与以下各项杂交:(i)SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13、或 SEQ ID NO:15 的成熟多肽编码序列,(ii) 包含在 SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:5、或 SEQ ID NO:13 的成熟多肽编码序列中的 cDNA 序列、或包括 SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11、或 SEQ ID NO:15 的成熟多肽编码序列的基因组 DNA 序列,(iii) (i) 或 (ii) 的子序列,或 (iv) (i)、(ii)、或 (iii) 的全长互补链(Sambrook(萨姆布鲁克)等人,1989,见上文)。SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13、或 SEQ ID NO:15 的成熟多肽编码序列的一个子序列含有至少 100 个连续核苷酸或优选至少 200 个连续核苷酸。此外,该子序列可以编码具有纤维素分解增强活性的多肽片段。在一个优选方面中,该成熟多肽编码序列是 SEQ ID NO:1 的核苷酸 388 至 1332、SEQ ID NO:3 的核苷酸 98 至 821、SEQ ID NO:5 的核苷酸 126 至 978、SEQ ID NO:7 的核苷酸 55 至 678、SEQ ID NO:9 的核苷酸 58 至 912、SEQ ID NO:11 的核苷酸 46 至 951、SEQ ID NO:13 的核苷酸 67 至 796、或 SEQ ID NO:15 的核苷酸 77 至 766。

[0282] 可以根据本领域熟知的方法,使用 SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13 或 SEQ ID NO:15 的核苷酸序列或其子序列;以及 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:14 或 SEQ ID NO:16 的氨基酸序列或其片段来设计用于从不同属或种的菌株中鉴定并克隆编码具有纤维素分解增强活性的多肽的 DNA 的核酸探针。具体地说,这类探针可以用于按照标准 DNA 印迹程序与感兴趣的属或种的基因组或 cDNA 杂交,以便鉴定并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。可以使用甚至更长的探针,例如长度优选是至少 600 个核苷酸、更优选至少 700 个核苷酸、甚至更优选至少 800 个核苷酸、或最优选至少 900 个核苷酸的核酸探针。DNA 和 RNA 探针两者都可使用。典型地将探针进行标记(例如,用 ^{32}P 、 ^3H 、 ^{35}S 、生物素、或抗生物素蛋白),以检测相应的基因。本发明涵盖此类探针。

[0283] 因此,可以针对与以上所描述的探针杂交并且编码具有纤维素分解增强活性的多肽的 DNA 对由这类其他菌株制备的基因组 DNA 或 cDNA 文库进行筛选。来自这类其他菌株的基因组 DNA 或其他 DNA 可以通过琼脂糖或聚丙烯酰胺凝胶电泳,或其他分离技术来分离。来自文库的 DNA 或分离的 DNA 可转移到并固定在硝酸纤维素或其他适合的载体材料上。为识别与 SEQ ID NO:1 或其子序列同源的克隆或 DNA,在 Southern 印迹法中优选使用载体材料。

[0284] 出于本发明的目的,杂交指示核苷酸序列与一种标记的核酸探针在如上所述的非常低至非常高严谨度条件下杂交,该标记的核酸探针对应于 SEQ ID NO:1、SEQ ID NO:3、SEQ

ID NO:5、SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13 或 SEQ ID NO:15 的成熟多肽编码序列,包含在 SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:5 或 SEQ ID NO:13 的成熟多肽编码序列中的 cDNA 序列,或包括 SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11 或 SEQ ID NO:15 的成熟多肽编码序列中的基因组 DNA 序列,其全长互补链,或其子序列。

[0285] 在一个优选方面中,该核酸探针是 SEQ ID NO:1 的成熟多肽编码序列。在另一个优选方面中,该核酸探针是 SEQ ID NO:1 的核苷酸 388 至 1332。在另一个优选方面中,该核酸探针是编码 SEQ ID NO:2 的多肽的多核苷酸序列或其子序列。在另一个优选方面中,该核酸探针是 SEQ ID NO:1。在另一个优选方面中,该核酸探针是包含在质粒 pEJG120 中的多核苷酸序列,该质粒包含在大肠杆菌 NRRL B-30699 中,其中其多核苷酸序列编码具有纤维素分解增强活性的多肽。在另一个优选方面中,该核酸探针是包含在质粒 pEJG120 中的成熟多肽编码序列,该质粒包含在大肠杆菌 NRRL B-30699 中。

[0286] 在另一个优选方面中,该核酸探针是 SEQ ID NO:3 的成熟多肽编码序列。在另一个优选方面中,该核酸探针是 SEQ ID NO:3 的核苷酸 98 至 821。在另一个优选方面中,该核酸探针是编码 SEQ ID NO:4 的多肽的多核苷酸序列或其子序列。在另一个优选方面中,该核酸探针是 SEQ ID NO:3。在另一个优选方面中,该核酸探针是包含在质粒 pTter61C 中的多核苷酸序列,该质粒包含在大肠杆菌 NRRL B-30813 中,其中其多核苷酸序列编码具有纤维素分解增强活性的多肽。在另一个优选方面中,该核酸探针是包含在质粒 pTter61C 中的成熟多肽编码序列,该质粒包含在大肠杆菌 NRRL B-30813 中。

[0287] 在另一个优选方面中,该核酸探针是 SEQ ID NO:5 的成熟多肽编码序列。在另一个优选方面中,该核酸探针是 SEQ ID NO:5 的核苷酸 126 至 978。在另一个优选方面中,该核酸探针是编码 SEQ ID NO:6 的多肽的多核苷酸序列或其子序列。在另一个优选方面中,该核酸探针是 SEQ ID NO:5。在另一个优选方面中,该核酸探针是包含在质粒 pTter61D 中的多核苷酸序列,该质粒包含在大肠杆菌 NRRL B-30812 中,其中其多核苷酸序列编码具有纤维素分解增强活性的多肽。在另一个优选方面中,该核酸探针是包含在质粒 pTter61D 中的成熟多肽编码序列,该质粒包含在大肠杆菌 NRRL B-30812 中。

[0288] 在另一个优选方面中,该核酸探针是 SEQ ID NO:7 的成熟多肽编码序列。在另一个优选方面中,该核酸探针是 SEQ ID NO:7 的核苷酸 55 至 678。在另一个优选方面中,该核酸探针是编码 SEQ ID NO:8 的多肽的多核苷酸序列或其子序列。在另一个优选方面中,该核酸探针是 SEQ ID NO:7。在另一个优选方面中,该核酸探针是包含在质粒 pTter61E 中的多核苷酸序列,该质粒包含在大肠杆菌 NRRL B-30814 中,其中其多核苷酸序列编码具有纤维素分解增强活性的多肽。在另一个优选方面中,该核酸探针是包含在质粒 pTter61E 中的成熟多肽编码序列,该质粒包含在大肠杆菌 NRRL B-30814 中。

[0289] 在另一个优选方面中,该核酸探针是 SEQ ID NO:9 的成熟多肽编码序列。在另一个优选方面中,该核酸探针是 SEQ ID NO:9 的核苷酸 58 至 912。在另一个优选方面中,该核酸探针是编码 SEQ ID NO:10 的多肽的多核苷酸序列或其子序列。在另一个优选方面中,该核酸探针是 SEQ ID NO:9。在另一个优选方面中,该核酸探针是包含在质粒 pTter61G 中的多核苷酸序列,该质粒包含在大肠杆菌 NRRL B-30811 中,其中其多核苷酸序列编码具有纤维素分解增强活性的多肽。在另一个优选方面中,该核酸探针是包含在质粒 pTter61G 中的成熟多肽编码序列,该质粒包含在大肠杆菌 NRRL B-30811 中。

[0290] 在另一个优选方面中,该核酸探针是 SEQ ID NO:11 的成熟多肽编码序列。在另一个优选方面中,该核酸探针是 SEQ ID NO:11 的核苷酸 46 至 951。在另一个优选方面中,该核酸探针是编码 SEQ ID NO:12 的多肽的多核苷酸序列或其子序列。在另一个优选方面中,该核酸探针是 SEQ ID NO:11。在另一个优选方面中,该核酸探针是包含在质粒 pTter61F 中的多核苷酸序列,该质粒包含在大肠杆菌 NRRL B-50044 中,其中其多核苷酸序列编码具有纤维素分解增强活性的多肽。在另一个优选方面中,该核酸探针是包含在质粒 pTter61F 中的成熟多肽编码区,该质粒包含在大肠杆菌 NRRL B-50044 中。

[0291] 在另一个优选方面中,该核酸探针是 SEQ ID NO:13 的成熟多肽编码序列。在另一个优选方面中,该核酸探针是 SEQ ID NO:13 的核苷酸 67 至 796。在另一个优选方面中,该核酸探针是编码 SEQ ID NO:14 的多肽的多核苷酸序列或其子序列。在另一个优选方面中,该核酸探针是 SEQ ID NO:13。在另一个优选方面中,该核酸探针是包含在质粒 pDZA2-7 中的多核苷酸序列,该质粒包含在大肠杆菌 NRRL B-30704 中,其中其多核苷酸序列编码具有纤维素分解增强活性的多肽。在另一个优选方面中,该核酸探针是包含在质粒 pDZA2-7 中的成熟多肽编码序列,该质粒包含在大肠杆菌 NRRL B-30704 中。

[0292] 在另一个优选方面中,该核酸探针是 SEQ ID NO:15 的成熟多肽编码序列。在另一个优选方面中,该核酸探针是 SEQ ID NO:15 的核苷酸 77 至 766。在另一个优选方面中,该核酸探针是编码 SEQ ID NO:16 的多肽的多核苷酸序列或其子序列。在另一个优选方面中,该核酸探针是 SEQ ID NO:15。在另一个优选方面中,该核酸探针是包含在质粒 pTr333 中的多核苷酸序列,该质粒包含在大肠杆菌 NRRL B-30878 中,其中其多核苷酸序列编码具有纤维素分解增强活性的多肽。在另一个优选方面中,该核酸探针是包含在质粒 pTr333 中的成熟多肽编码序列,该质粒包含在大肠杆菌 NRRL B-30878 中。

[0293] 对于长度为至少 100 个核苷酸的长探针而言,非常低至非常高严谨度条件定义为最佳地按照标准 DNA 印迹程序,在 42℃ 下在 5X SSPE、0.3% SDS、200 μ 克 /ml 剪切并变性的鲑鱼精子 DNA 中,以及对于非常低和低严谨度在 25% 甲酰胺中,对于中和中-高严谨度在 35% 甲酰胺中,或对于高和非常高严谨度在 50% 甲酰胺中预杂交和杂交 12 至 24 小时。

[0294] 对于长度为至少 100 个核苷酸的长探针而言,载体材料最终使用 2XSSC、0.2% SDS,优选在 45℃ 下(非常低严谨度),更优选在 50℃ 下(低严谨度),更优选在 55℃ 下(中严谨度),更优选在 60℃ 下(中-高严谨度),甚至更优选在 65℃ 下(高严谨度),并且最优选在 70℃ 下(非常高严谨度)洗涤三次,每次 15 分钟。

[0295] 对于长度为约 15 个核苷酸至约 70 个核苷酸的短探针而言,严谨度条件被定义为根据标准 DNA 印迹程序,在比使用根据博尔顿 (Bolton) 和麦卡锡 (McCarthy) (1962, 美国国家科学院院刊 (Proceedings of the National Academy of Sciences USA) 48:1390) 的计算计算的 T_m 低约 5℃ 至约 10℃ 下,在 0.9M NaCl、0.09M Tris-HCl (pH 7.6)、6mM EDTA、0.5% NP-40、1X 登哈特氏溶液 (Denhardt's solution)、1mM 焦磷酸钠、1mM 磷酸二氢钠、0.1mMATP、以及每 ml 0.2mg 的酵母 RNA 中预杂交、杂交及杂交后洗涤持续最佳 12 至 24 小时。

[0296] 对于长度为约 15 个核苷酸至约 70 个核苷酸的短探针而言,将载体材料在比计算的 T_m 低 5℃ 至 10℃ 下,在 6X SSC 加 0.1% SDS 中洗涤一次,持续 15 分钟并且使用 6X SSC 洗涤两次,每次 15 分钟。

[0297] 在一个第五方面中,具有纤维素分解增强活性的该多肽由以下多核苷酸编码,该多核苷酸包括或其组成为以下核苷酸序列,该核苷酸序列与 SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13、或 SEQ ID NO:15 的成熟多肽编码序列具有优选至少 60%、更优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 的一致性程度。

[0298] 在一个优选方面中,该成熟多肽编码序列是 SEQ ID NO:1 的核苷酸 388 至 1332、SEQ ID NO:3 的核苷酸 98 至 821、SEQ ID NO:5 的核苷酸 126 至 978、SEQ ID NO:7 的核苷酸 55 至 678、SEQ ID NO:9 的核苷酸 58 至 912、SEQ ID NO:11 的核苷酸 46 至 951、SEQ ID NO:13 的核苷酸 67 至 796、或 SEQ ID NO:15 的核苷酸 77 至 766。

[0299] 在一个第六方面中,具有纤维素分解增强活性的该多肽是一种人工变体,该人工变体包括 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:12、或 SEQ ID NO:14 或 SEQ ID NO:16 的成熟多肽的一个或多个(或若干个)氨基酸的取代、缺失和/或插入;或其同源序列。用于制备这样一种人工变体的方法如上文所述。

[0300] SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:12 或 SEQ ID NO:14 或 SEQ ID NO:16 的成熟多肽的氨基酸取代、缺失和/或插入的总数是 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,并且甚至最优选 1。

[0301] 具有纤维素分解增强活性的多肽可以从任何属的微生物获得。在一个优选方面中,从指定来源获得的多肽分泌到细胞外。

[0302] 具有纤维素分解增强活性的多肽可以是一种细菌多肽。例如,该多肽可以是革兰氏阳性细菌多肽,如具有纤维素分解增强活性的芽孢杆菌属、链球菌属、链霉菌属、葡萄球菌属、肠球菌属、乳酸杆菌属、乳球菌属、梭菌属、土芽孢杆菌属、或海洋芽孢杆菌属多肽;或革兰氏阴性细菌多肽,如具有纤维素分解增强活性的大肠杆菌、假单胞菌属、沙门氏菌属、弯曲杆菌属、螺旋杆菌属、黄杆菌属、梭杆菌属、泥杆菌属、奈瑟氏菌属、或脲原体属多肽。

[0303] 在一个优选方面中,该多肽是具有纤维素分解增强活性的嗜碱芽孢杆菌、解淀粉芽孢杆菌、短芽孢杆菌、环状芽孢杆菌、克劳氏芽孢杆菌、凝结芽孢杆菌、坚硬芽孢杆菌、灿烂芽孢杆菌、迟缓芽孢杆菌、地衣芽孢杆菌、巨大芽孢杆菌、短小芽孢杆菌、嗜热脂肪芽孢杆菌、枯草芽孢杆菌、或苏云金芽孢杆菌多肽。

[0304] 在另一个优选方面中,该多肽是具有纤维素分解增强活性的似马链球菌、酿脓链球菌、乳房链球菌、或马链球菌兽疫亚种多肽。

[0305] 在另一个优选方面中,该多肽是具有纤维素分解增强活性的不产色链霉菌、阿维链霉菌、天蓝色链霉菌、灰色链霉菌、或变铅青链霉菌多肽。

[0306] 具有纤维素分解增强活性的多肽还可以是真菌多肽,并且更优选地是一种具有纤维素分解增强活性的酵母多肽,如假丝酵母属、克鲁维酵母属、毕赤酵母属、酵母菌属、裂殖酵母属、或耶氏酵母属多肽;或更优选地是一种具有纤维素分解增强活性的丝状真菌多肽,如枝顶孢霉属、伞菌属、链格孢属、曲霉属、短梗霉属、葡萄座腔菌属(Botryosphaeria)、拟蜡菌属、毛喙壳属、金孢子菌属、麦角菌属、旋孢腔菌属、鬼伞属、乳白蚁属、棒囊壳属、隐丛赤壳菌属、隐球菌属、色二孢属、黑耳属、线黑粉酵母属、镰刀菌属、赤霉属、全鞭毛虫属、

腐质霉属、耙齿菌属、蘑燕属、小腔球菌属、梨孢菌属、黑果菌属 (*Melanocarpus*)、多孔菌属、毛霉属、毁丝霉属、新美鞭菌属、链孢菌属、拟青霉属、青霉菌属、平革菌属、瘤胃壶菌属、*Poitrasia*、假黑盘菌属、假披发虫属 (*Pseudotrichonympha*)、根毛霉菌属、裂褶菌属、柱顶孢属、踝节菌属、嗜热子囊菌属、梭孢壳霉属、弯颈霉属、木霉属、长毛盘菌属、轮枝孢属、小包脚菇属、或炭角菌属多肽。

[0307] 在一个优选方面中,该多肽是具有纤维素分解增强活性的卡尔斯伯酵母、酿酒酵母、糖化酵母、格拉斯酵母、克鲁维酵母、诺地酵母、或卵形酵母多肽。

[0308] 在另一个优选方面中,该多肽是具有纤维素分解增强活性的解纤维枝顶孢霉、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、米曲霉、嗜角质金孢子菌、拉克淖金孢子菌、热带金孢子菌、粪状金孢子菌 (*Chrysosporium merdarium*)、狭边金孢子菌 (*Chrysosporium inops*)、租金孢子菌、节状金孢子菌 (*Chrysosporium queenslandicum*)、带纹金孢子菌 (*Chrysosporium zonatum*)、杆孢状镰孢、谷类镰孢、库威镰孢、大刀镰孢、禾谷镰孢、禾赤镰孢、异孢镰孢、合欢木镰孢、尖镰孢、多枝镰孢、粉红镰孢、接骨木镰孢、肤色镰孢、拟分枝孢镰孢、硫色镰孢、圆镰孢、拟丝孢镰孢、镶片镰孢、灰腐质霉、特异腐质霉、疏棉状腐质霉、白耙齿菌、米黑毛霉、嗜热毁丝霉、粗糙脉孢菌、绳状青霉菌、产紫青霉菌、黄孢平革菌、无色梭孢壳霉 (*Thielavia achromatica*)、*Thielavia albomyces*、白毛梭孢壳霉 (*Thielavia albopilosa*)、澳洲梭孢壳霉、*Thielavia fimeti*、小孢梭孢壳霉、卵孢梭孢壳霉、秘鲁梭孢壳霉 (*Thielavia peruviana*)、瘤孢梭孢壳霉、毛梭孢壳霉、耐热梭孢壳 (*Thielavia subthermophila*)、土生梭孢壳、哈茨木霉、康宁木霉、长枝木霉、里氏木霉、绿色木霉、或褐孢长毛盘菌 (*Trichophaea saccata*) 多肽。

[0309] 将理解的是,对于以上提到的物种而言,本发明涵盖完全状态和不完全状态 (perfect and imperfect states) 二者、以及其他分类学等效物,例如无性型,而不管它们已知的物种名称是什么。本领域的普通技术人员将容易地识别适当等效物的身份。

[0310] 这些物种的菌株可以容易地在许多培养物保藏中心为公众所获得,如美国典型培养物保藏中心 (ATCC)、德国微生物菌种保藏中心 (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, DSM)、荷兰菌种保藏中心 (Centraalbureau Voor Schimmelcultures, CBS)、以及美国农业研究菌种保藏中心北方地区研究中心 (NRRL)。

[0311] 此外,可以使用上述探针,从其他来源,包括从自然界 (例如,土壤、堆肥、水,等等) 分离的微生物鉴定并获得具有纤维素分解增强活性的该多肽。用于从自然生境分离微生物的技术是本领域熟知的。然后可通过类似地筛选这样一种微生物的基因组或 cDNA 文库来获得多核苷酸。一旦用这种或这些探针检测到编码一种多肽的多核苷酸,就可以通过使用本领域普通技术人员众所周知的的技术分离或克隆该多核苷酸 (参见,例如,萨姆布鲁克 (Sambrook) 等人,1989,见上文)。

[0312] 包括编码具有纤维素分解增强活性的多肽的核苷酸序列的多核苷酸可以被分离并利用来表达具有纤维素分解增强活性的该多肽,用于在本发明的方法中进行评估,如在此所描述。

[0313] 这些多核苷酸包括以下核苷酸序列,这些核苷酸序列与 SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13、或 SEQ ID

NO:15 的成熟多肽编码序列具有优选至少 60%、更优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 的一致性程度,这些多核苷酸编码具有纤维素分解增强活性的多肽。

[0314] 该多核苷酸还可以是一种编码具有纤维素分解增强活性的多肽的多核苷酸,该多核苷酸在至少非常低严谨度条件、优选至少低严谨度条件、更优选至少中严谨度条件、更优选至少中-高严谨度条件、甚至更优选至少高严谨度条件、以及最优选至少非常高严谨度条件下与以下各项杂交:(i)SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13、或 SEQ ID NO:15 的成熟多肽编码序列,(ii) 包含在 SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:5、或 SEQ ID NO:13 的成熟多肽编码序列中的 cDNA 序列、或包括 SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11、或 SEQ ID NO:15 的成熟多肽编码序列的基因组 DNA 序列,或 (iii) (i) 或 (ii) 的全长互补链;或其等位基因变体和子序列(Sambrook(萨姆布鲁克)等人,1989,见上文),如在此所描述。在一个优选方面中,该成熟多肽编码序列是 SEQ ID NO:1 的核苷酸 388 至 1332、SEQ ID NO:3 的核苷酸 98 至 821、SEQ ID NO:5 的核苷酸 126 至 978、SEQ ID NO:7 的核苷酸 55 至 678、SEQ ID NO:9 的核苷酸 58 至 912、SEQ ID NO:11 的核苷酸 46 至 951、SEQ ID NO:13 的核苷酸 67 至 796、或 SEQ ID NO:15 的核苷酸 77 至 766。

[0315] 如先前所述,用来分离或克隆编码多肽的多核苷酸的技术在本领域中是已知的,并且包括从基因组 DNA 分离、从 cDNA 制备或其组合。

[0316] 过氧化物酶

[0317] 一种过氧化物酶与一种纤维素分解酶组合物;一种具有纤维素分解增强活性的多肽;以及一种非离子表面活性剂和/或一种阳离子表面活性剂一起存在或在本发明的用于降解预处理的纤维素材料的方法过程中添加。

[0318] 在本发明的方法中,具有过氧化物酶活性的该多肽可以是任何具有过氧化物酶活性的多肽。该过氧化物酶可以作为一种酶活性存在于该酶组合物中和/或作为添加至该组合物中一种或多种(若干种)蛋白组分。在一个优选方面中,具有过氧化物酶活性的该多肽对于该纤维素分解酶组合物的一种或多种(若干种)组分而言是外源的。

[0319] 过氧化物酶的实例是过氧化物酶和过氧化物分解酶,包括但不限于优选各项:

[0320] E. C. 1. 11. 1. 1NADH 过氧化物酶;

[0321] E. C. 1. 11. 1. 2NADPH 过氧化物酶;

[0322] E. C. 1. 11. 1. 3 脂肪酸过氧化物酶;

[0323] E. C. 1. 11. 1. 5 细胞色素 c 过氧化物酶;

[0324] E. C. 1. 11. 1. 6 过氧化氢酶;

[0325] E. C. 1. 11. 1. 7 过氧化物酶;

[0326] E. C. 1. 11. 1. 8 碘化物过氧化物酶;

[0327] E. C. 1. 11. 1. 9 谷胱甘肽过氧化物酶;

[0328] E. C. 1. 11. 1. 10 氯化物过氧化物酶;

[0329] E. C. 1. 11. 1. 11L- 抗坏血酸过氧化物酶;

[0330] E. C. 1. 11. 1. 12 磷脂过氧化氢谷胱甘肽过氧化物酶;

[0331] E. C. 1. 11. 1. 13 锰过氧化物酶；

[0332] E. C. 1. 11. 1. 14 木质素过氧化物酶；

[0333] E. C. 1. 11. 1. 15 过氧化物氧还蛋白；

[0334] E. C. 1. 11. 1. 16 通用过氧化物酶；

[0335] E. C. 1. 11. 1. B2 氯化物过氧化物酶；

[0336] E. C. 1. 11. 1. B6 碘化物过氧化物酶；

[0337] E. C. 1. 11. 1. B7 溴化物过氧化物酶；

[0338] E. C. 1. 11. 1. B8 碘化物过氧化物酶；

[0339] EC 编号和名称可以发现于例如 www.brenda-enzymes.org。

[0340] 在一个方面中,该过氧化物酶是一种 NADH 过氧化物酶。在另一个方面中,该过氧化物酶是一种 NADPH 过氧化物酶。在另一个方面中,该过氧化物酶是一种脂肪酸过氧化物酶。在另一个方面中,该过氧化物酶是一种细胞色素 c 过氧化物酶。在另一个方面中,该过氧化物酶是一种过氧化氢酶。在另一个方面中,该过氧化物酶是一种过氧化物酶。在另一个方面中,该过氧化物酶是一种碘化物过氧化物酶。在另一个方面中,该过氧化物酶是一种谷胱甘肽过氧化物酶。在另一个方面中,该过氧化物酶是一种氯化物过氧化物酶。在另一个方面中,该过氧化物酶是一种 L- 抗坏血酸过氧化物酶。在另一个方面中,该过氧化物酶是一种磷脂过氧化氢谷胱甘肽过氧化物酶。在另一个方面中,该过氧化物酶是一种锰过氧化物酶。在另一个方面中,该过氧化物酶是一种木质素过氧化物酶。在另一个方面中,该过氧化物酶是一种过氧化物氧还蛋白。在另一个方面中,该过氧化物酶是一种通用过氧化物酶。在另一个方面中,该过氧化物酶是一种氯化物过氧化物酶。在另一个方面中,该过氧化物酶是一种碘化物过氧化物酶。在另一个方面中,该过氧化物酶是一种溴化物过氧化物酶。在另一个方面中,该过氧化物酶是一种碘化物过氧化物酶。

[0341] 在一个优选实施例中,该过氧化物酶是一种 E. C. 1. 11. 1. 7 过氧化物酶。

[0342] 过氧化物酶的实例包括但不限于:灰盖鬼伞过氧化物酶(巴恩斯戈德(Baunsgaard)等人,1993,长根鬼伞过氧化物酶的氨基酸序列和编码灰盖鬼伞过氧化物酶的 cDNA 序列,一种新真菌过氧化物酶家族(Amino acid sequence of *Coprinus macrorrhizus* peroxidase and cDNA sequence encoding *Coprinus cinereus* peroxidase. A new family of fungal peroxidases), 欧洲生物化学杂志(Eur. J. Biochem.)213(1):605-611(登录号 P28314)或在此的 SEQ ID NO:71);辣根过氧化物酶(藤山(Fujiyama)等人,1988,辣根过氧化物酶同工酶 C 基因的结构(Structure of the horseradish peroxidase isozyme C genes),欧洲生物化学杂志173(3):681-687(登录号 P15232));过氧化物氧还蛋白(辛格(Singh)和志知(Shichi),1998,牛眼中的一种新颖谷胱甘肽过氧化物酶,序列分析、mRNA 水平和翻译(A novel glutathione peroxidase in bovine eye. Sequence analysis, mRNA level, and translation), 生物化学杂志(J. Biol. Chem.)273(40):26171-26178(登录号 077834));乳过氧化物酶(Dull(达尔)等人,1990,编码牛和人乳过氧化物酶的 cDNA 的分子克隆(Molecular cloning of cDNAs encoding bovine and human lactoperoxidase), DNA 与细胞生物学(DNA Cell Biol.)9(7):499-509(登录号 P80025));嗜酸性粒细胞过氧化物酶(弗恩海姆(Fornhem)等人,1996,猪阳离子嗜酸性粒细胞颗粒蛋白的分离与表征(Isolation

and characterization of porcine cationic eosinophilgranule proteins), 过敏与免疫学国际档案 (Int. Arch. Allergy Immunol.) 110(2):132-142(登录号 P80550)); 通用过氧化物酶 (鲁伊斯-杜埃尼亚斯 (Ruiz-Duenas) 等人, 1999, 一种分离自木质素分解真菌杏鲍菇的新颖过氧化物酶的分子表征 (Molecular characterization of a novel peroxidase isolated from the ligninolytic fungus *Pleurotus eryngii*), 分子微生物学 (Mol. Microbiol.) 31(1):223-235(登录号 094753)); 芜菁过氧化物酶 (马萨 (Mazza) 和韦林德 (Welinder), 1980, 芜菁过氧化物酶 7 的共价结构, 溴化氰片段、完整结构及与辣根过氧化物酶 C 的比较 (Covalent structure of turnip peroxidase 7. Cyanogen bromide fragments, complete structure and comparison to horseradish peroxidase C), Eur. J. Biochem. 108(2):481-489(登录号 P00434)); 髓过氧化物酶 (森下 (Morishita) 等人, 1987, 人类髓过氧化物酶的染色体基因结构及由粒细胞集落刺激因子调控其表达 (Chromosomal gene structure of human myeloperoxidase and regulation of its expression by granulocyte colony-stimulating factor), 生物化学杂志 262(31):15208-15213(登录号 P05164)); 过氧蛋白 (peroxidasin) 和过氧蛋白同系物 (堀越 (Horikoshi) 等人, 1999, 从 p53 依赖性凋亡细胞中分离差异表达的 cDNA: 果蝇过氧蛋白基因的人类同系物的激活 (Isolation of differentially expressed cDNAs from p53-dependent apoptotic cells: activation of the human homologue of the *Drosophila* peroxidasin gene), 生物化学与生物物理研究通讯 (Biochem. Biophys. Res. Commun.) 261(3):864-869(登录号 Q92626)); 木质素过氧化物酶 (田 (Tien) 和屠 (Tu), 1987, 来自黄孢原毛平革菌的木质素酶的 cDNA 克隆与测序 (Cloning and sequencing of a cDNA for a ligninase from *Phanerochaete chrysosporium*), 自然 (Nature) 326(6112):520-523(登录号 P06181)); 锰过氧化物酶 (奥尔特 (Orth) 等人, 1994, 编码来自黄孢原毛平革菌的锰过氧化物酶的 cDNA 的表征: 木质素和锰过氧化物酶编码基因的基因组组织 (Characterization of a cDNA encoding a manganese peroxidase from *Phanerochaete chrysosporium*: genomic organization of lignin and manganese peroxidase-encoding genes) 基因 (Gene) 148(1):161-165(登录号 P78733)); α -双加氧酶, 双氧化酶, 过氧蛋白, 无脊椎动物细胞黏附蛋白, 短 peroxidockerin, 乳过氧化物酶, 髓过氧化物酶, 非哺乳动物脊椎动物过氧化物酶, 过氧化氢酶, 过氧化氢酶-脂肪氧合酶融合, 二血红素细胞色素 c 过氧化物酶, 甲胺利用蛋白, DyP 型过氧化物酶, 卤代过氧化物酶, 抗坏血酸过氧化物酶, 过氧化氢酶过氧化物酶, 混合抗坏血酸-细胞色素 c 过氧化物酶, 木质素过氧化物酶, 锰过氧化物酶, 通用过氧化物酶, 其他 II 类过氧化物酶, III 类过氧化物酶, 烷基羟基过氧化物酶 (alkylhydroperoxidase) D, 其他烷基羟基过氧化物酶, 无血红素、无金属卤代过氧化物酶, 无血红素钒卤代过氧化物酶, 锰过氧化氢酶, NADH 过氧化物酶, 谷胱甘肽过氧化物酶, 半胱氨酸过氧化物酶, 硫氧还蛋白依赖型巯基过氧化物酶, 以及 AhpE 样过氧化物氧还蛋白 (帕沙德 (Passard) 等人, 2007, 植物化学 (Phytochemistry) 68:1605-1611)。

[0343] 该过氧化物酶活性可以从任何属的微生物获得。在一个方面中, 获得自给定来源的多肽被分泌到细胞外。

[0344] 该过氧化物酶活性可以是一种细菌多肽。例如, 该多肽可以是革兰氏阳性细菌多

肽,如具有过氧化物酶活性的芽孢杆菌属、链球菌属、链霉菌属、葡萄球菌属、肠球菌属、乳酸杆菌属、乳球菌属、梭菌属、土芽孢杆菌属、或海洋芽孢杆菌属多肽;或革兰氏阴性细菌多肽,如具有过氧化物酶活性的大肠杆菌、假单胞菌属、沙门氏菌属、弯曲杆菌属、螺旋杆菌属、黄杆菌属、梭杆菌属、泥杆菌属、奈瑟氏菌属、或脲原体属多肽。

[0345] 在一个实施例中,该过氧化物酶来源于以下各项的菌株:嗜碱芽孢杆菌、解淀粉芽孢杆菌、短杆菌、环状芽孢杆菌、克劳氏芽孢杆菌、凝结芽孢杆菌、坚硬芽孢杆菌、灿烂芽孢杆菌、迟缓芽孢杆菌、地衣芽孢杆菌、巨大芽孢杆菌、短小芽孢杆菌、嗜热脂肪芽孢杆菌、枯草芽孢杆菌或苏云金芽孢杆菌。

[0346] 在另一个实施例中,该过氧化物酶来源于以下各项的菌株:似马链球菌、酿脓链球菌、乳房链球菌、或马链球菌兽疫亚种。

[0347] 在另一个方面中,该过氧化物酶来源于以下各项的菌株:不产色链霉菌、阿维链霉菌、天蓝色链霉菌、灰色链霉菌、或变铅青链霉菌。

[0348] 该过氧化物酶活性还可以是一种真菌多肽,并且更优选是一种酵母多肽,如来源于假丝酵母属、克鲁维酵母属、毕赤酵母属、酵母菌属、裂殖酵母属、或耶氏酵母属具有过氧化物酶活性的多肽;或更优选是一种丝状真菌多肽,如枝顶孢霉属、伞菌属、链格孢属、曲霉属、短梗霉属、葡萄座腔菌属、拟蜡菌属、毛喙壳属、金孢子菌属、麦角菌属、旋孢腔菌属、鬼伞属、乳白蚁属、棒囊壳属、隐丛赤壳菌属、隐球菌属、色二孢属、黑耳属、线黑粉酵母属、镰刀菌属、赤霉属、全鞭毛虫属、腐质霉属、耙齿菌属、蘑燕属、小腔球菌属、梨孢菌属、黑果菌属、多孔菌属、毛霉属、毁丝霉属、新美鞭菌属、链孢菌属、拟青霉属、青霉菌属、平革菌属、瘤胃壶菌属、Poitrasia、假黑盘菌属、假披发虫属、根毛霉菌属、裂褶菌属、柱顶孢属、踝节菌属、嗜热子囊菌属、梭孢壳霉属、弯颈霉属、木霉属、长毛盘菌属、轮枝孢属、小包脚菇属、或炭角菌属。

[0349] 在另一个方面中,该过氧化物酶来源于以下各项的菌株:卡尔斯伯酵母、酿酒酵母、糖化酵母、格拉斯酵母、克鲁维酵母、诺地酵母、或卵形酵母。

[0350] 在另一个方面中,该过氧化物酶来源于以下各项的菌株:解纤维枝顶孢霉、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、米曲霉、嗜角质金孢子菌、拉克淖金孢子菌、热带金孢子菌、粪状金孢子菌、狭边金孢子菌、毡金孢子菌、节状金孢子菌、带纹金孢子菌、杆孢状镰孢、禾谷镰孢、库威镰孢、黄色镰刀菌、禾谷镰刀菌、禾赤镰孢、异孢镰孢、合欢木镰孢、尖孢镰刀菌、多枝镰孢、粉红镰孢菌、接骨木镰孢、肤色镰孢、拟分枝孢镰刀菌、硫色镰孢、圆镰孢、拟丝孢镰刀菌、镶片镰孢菌、灰腐质酶、特异腐质霉、疏棉状腐质霉、白囊耙齿菌、米黑毛霉、嗜热毁丝霉、粗糙脉孢菌、绳状青霉菌、产紫青霉、黄孢原毛平革菌、无色梭孢壳、*Thielavia albomyces*、白毛梭孢壳、澳洲梭孢壳、*Thielavia fimeti*、小孢梭孢壳、卵孢梭孢壳、秘鲁梭孢壳、瘤孢梭孢壳、毛梭孢壳、耐热梭孢壳、土生梭孢壳、哈茨木霉、康宁木霉、长枝木霉、里氏木霉或绿色木霉。

[0351] 在另一个方面中,该过氧化物酶是辣根过氧化物酶。在另一个方面中,该过氧化物酶是灰盖鬼伞过氧化物酶,例如示于在此的 SEQ ID NO:71 中的灰盖鬼伞过氧化物酶。在一个实施例中,该过氧化物酶与在此的 SEQ ID NO:71 (即, CiP) 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99%

序列一致性。

[0352] 用来分离或克隆编码具有过氧化物酶活性的多肽的多核苷酸的技术在本领域中是已知的,并且包括从基因组 DNA 分离、从 cDNA 制备或其组合。可以例如通过使用熟知的聚合酶链反应 (PCR) 或表达文库的抗体筛选来检测具有共有结构特征的克隆 DNA 片段,实现从这样的基因组 DNA 克隆本发明的多核苷酸。参见例如,伊尼斯 (Innis) 等人,1990, PCR: 方法和应用指南 (PCR: A Guide to Methods and Application), 学术出版社 (Academic Press), 纽约。可以使用其他核酸扩增程序,例如连接酶链式反应 (LCR)、连接激活转录 (LAT) 和基于核苷酸序列的扩增 (NASBA)。

[0353] 纤维素分解酶组合物

[0354] 在本发明的方法或工艺中,该纤维素分解酶组合物可以包括在将预处理的纤维素材料加工为葡萄糖和 / 或纤维二糖,或将半纤维素加工成木糖、甘露糖、半乳糖和 / 或阿拉伯糖中涉及的任何蛋白。

[0355] 该纤维素分解酶组合物典型地包括具有纤维素分解活性的酶。在一个方面中,该纤维素分解酶组合物包括一种或多种(若干种)纤维素分解酶。在一个方面中,该纤维素分解酶组合物进一步包括一种或多种(若干种)木聚糖降解酶。在另一个方面中,该纤维素分解酶组合物包括一种或多种(若干种)纤维素分解酶和一种或多种(若干种)木聚糖降解酶。

[0356] 该一种或多种(若干种)纤维素分解酶优选是选自下组,该组由以下各项组成:内切葡聚糖酶、纤维二糖水解酶、以及 β -葡萄糖苷酶。该一种或多种(若干种)木聚糖降解酶优选是选自下组,该组由以下各项组成:木聚糖酶、乙酰木聚糖酯酶、阿魏酸酯酶、阿拉伯呋喃糖苷酶、木糖苷酶、以及葡萄糖醛酸糖苷酶。

[0357] 在另一个方面中,该纤维素分解酶组合物可以进一步或甚至进一步包括一种或多种(若干种)另外的酶活性以改善含纤维素材料的降解。优选的另外的酶是半纤维素酶(例如 α -D-葡萄糖醛酸糖苷酶、 α -L-阿拉伯呋喃糖苷酶、内切-甘露聚糖酶、 β -甘露糖苷酶、 α -半乳糖苷酶、内切- α -L-阿拉伯聚糖酶、 β -半乳糖苷酶)、碳水化合物-酯酶(例如乙酰基-木聚糖酯酶、乙酰基-甘露聚糖酯酶、阿魏酸酯酶、香豆酸酯酶、葡萄糖醛酸酯酶)、果胶酶、蛋白酶、木质素分解酶(例如漆酶、锰过氧化物酶、木质素过氧化物酶、 H_2O_2 生产酶、氧化还原酶)、扩张蛋白、膨胀素、或其混合物。在本发明的方法中,可以在发酵之前或过程中,例如在糖化过程中,或在发酵微生物的繁殖过程中或之后添加这一种或多种另外的酶。

[0358] 该纤维素分解酶组合物的一种或多种(若干种)组分可以是野生型蛋白、重组蛋白、或野生型蛋白与重组蛋白的组合。例如,一种或多种(若干种)组分可以是用作宿主细胞以重组表达该酶组合物的一种或多种(若干种)其他组分的细胞的原生蛋白质。该酶组合物的一种或多种(若干种)组分可以作为单组分产生,然后将这些单组分组合以形成该酶组合物。该纤维素分解酶组合物可以是多组分和单组分蛋白制剂的组合。

[0359] 用于本发明的方法或工艺中的酶可以处于适合在此描述的方法或工艺中使用的任何形式,例如像去除或未去除细胞的粗发酵液、具有或不具有细胞碎片的细胞裂解液、半纯化的或纯化的酶制剂、或作为酶来源的宿主细胞。该纤维素分解酶组合物可以是干粉或颗粒、非尘颗粒、液体、稳定化的液体、或稳定化的受保护的酶。可以根据已建立的方法例如通过添加如糖、糖醇或其它多元醇等稳定剂和 / 或乳酸,对液体酶制剂进行稳定化。

[0360] 具有纤维素分解酶活性或木聚糖降解活性的多肽可以是细菌多肽。例如,该多肽可以是革兰氏阳性细菌多肽,如具有纤维素分解酶活性或木聚糖降解活性的芽孢杆菌属、链球菌属、链霉菌属、葡萄球菌属、肠球菌属、乳酸杆菌属、乳球菌属、梭菌属、土芽孢杆菌属、或海洋芽孢杆菌属多肽;或革兰氏阴性细菌多肽,如具有纤维素分解酶活性或木聚糖降解活性的大肠杆菌、假单胞菌属、沙门氏菌属、弯曲杆菌属、螺旋杆菌属、黄杆菌属、梭杆菌属、泥杆菌属、奈瑟氏菌属、或脲原体属多肽。

[0361] 在一个优选方面中,该多肽是具有纤维素分解酶活性或木聚糖降解活性的嗜碱芽孢杆菌、解淀粉芽孢杆菌、短芽孢杆菌、环状芽孢杆菌、克劳氏芽孢杆菌、凝结芽孢杆菌、坚强芽孢杆菌、灿烂芽孢杆菌、迟缓芽孢杆菌、地衣芽孢杆菌、巨大芽孢杆菌、短小芽孢杆菌、嗜热脂肪芽孢杆菌、枯草芽孢杆菌、或苏云金芽孢杆菌多肽。

[0362] 在另一个优选方面中,该多肽是具有纤维素分解酶活性或木聚糖降解活性的似马链球菌、酿脓链球菌、乳房链球菌和马链球菌兽瘟亚种多肽。

[0363] 在另一个优选方面中,该多肽是具有纤维素分解酶活性或木聚糖降解活性的不产色链霉菌、阿维链霉菌、天蓝色链霉菌、灰色链霉菌、或变铅青链霉菌多肽。

[0364] 具有纤维素分解酶活性或木聚糖降解活性的该多肽还可以是真菌多肽,并且更优选地是一种具有纤维素分解酶活性或木聚糖降解活性的酵母多肽,如假丝酵母属、克鲁维酵母属、毕赤酵母属、酵母菌属、裂殖酵母属、或耶氏酵母属多肽;或更优选地是一种具有纤维素分解酶活性或木聚糖降解活性的丝状真菌多肽,如枝顶孢霉属、伞菌属、链格孢属、曲霉属、短梗霉属、葡萄座腔菌属、拟蜡菌属、毛喙壳属、金孢子菌属、麦角菌属、旋孢腔菌属、鬼伞属、乳白蚁属、棒囊壳属、隐丛赤壳菌属、隐球菌属、色二孢属、黑耳属、线黑粉酵母属、镰刀菌属、赤霉属、全鞭毛虫属、腐质霉属、耙齿菌属、磨燕属、小腔球菌属、梨孢菌属、黑果菌属、多孔菌属、毛霉属、毁丝霉属、新美鞭菌属、链孢菌属、拟青霉属、青霉菌属、平革菌属、瘤胃壶菌属、Poitrasia、假黑盘菌属、假披发虫属、根毛霉菌属、裂褶菌属、柱顶孢属、踝节菌属、嗜热子囊菌属、梭孢壳霉属、弯颈霉属、木霉属、长毛盘菌属、轮枝孢属、小包脚菇属、或炭角菌属多肽。

[0365] 在一个优选方面中,该多肽是具有纤维素分解酶活性或木聚糖降解活性的卡尔斯伯酵母、酿酒酵母、糖化酵母、格拉斯酵母、克鲁维酵母、诺地酵母、或卵形酵母多肽。

[0366] 在另一个优选方面中,该多肽是具有纤维素分解酶活性或木聚糖降解活性的解纤维枝顶孢霉、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、米曲霉、嗜角质金孢子菌、拉克淖金孢子菌、热带金孢子菌、粪状金孢子菌、狭边金孢子菌、租金孢子菌、节状金孢子菌、带纹金孢子菌、杆孢状镰孢、谷类镰孢、库威镰孢、大刀镰孢、禾谷镰孢、禾赤镰孢、异孢镰孢、合欢木镰孢、尖镰孢、多枝镰孢、粉红镰孢、接骨木镰孢、肤色镰孢、拟分枝孢镰孢、硫色镰孢、圆镰孢、拟丝孢镰孢、镶片镰孢、灰腐质霉、特异腐质霉、疏棉状腐质霉、白耙齿菌、米黑毛霉、嗜热毁丝霉、粗糙脉孢菌、绳状青霉菌、产紫青霉菌、黄孢平革菌、无色梭孢壳霉、*Thielavia albomyces*、白毛梭孢壳霉、澳洲梭孢壳霉、*Thielavia fimetii*、小孢梭孢壳霉、卵孢梭孢壳霉、秘鲁梭孢壳霉、瘤孢梭孢壳霉、毛梭孢壳霉、耐热梭孢壳、土生梭孢壳、哈茨木霉、康宁木霉、长枝木霉、里氏木霉、绿色木霉、或褐孢长毛盘菌多肽。

[0367] 还可以使用具有纤维素分解酶活性或木聚糖降解活性的多肽的化学修饰或蛋白质工程化的突变体。

[0368] 酶组合物的一种或多种(若干种)组分可以是重组组分,即,通过克隆编码该单一组分的 DNA 序列并且随后用该 DNA 序列转化细胞并且在宿主中表达产生(参见,例如,WO 91/17243 和 WO 91/17244)。优选宿主是异源宿主(酶对于宿主而言是异源的),但是在某些条件下,宿主还可以是同源宿主(酶对于宿主而言是原生的)。还可以通过纯化来自发酵液的这样一种蛋白来制备单组分纤维素分解蛋白。

[0369] 适合在本发明中使用的商业化纤维素分解酶组合物的实例包括例如 CELLIC™ Ctec(诺维信公司)、CELLIC™ Ctec2(诺维信公司)、CELLIC™ Ctec3(诺维信公司); CELLUCLAST™(诺维信公司)、NOVOZYM™188(诺维信公司)、CELLUZYME™(诺维信公司)、CEREFLO™(诺维信公司)、和 ULTRAFLO™(诺维信公司)、ACCELERASE™(杰能科公司(Genencor Int.))、LAMINEX™(杰能科公司)、SPEZYME™ CP(杰能科公司);ROHAMENT™7069W(罗姆公司(Röhm GmbH))、FIBREZYME® LDI(并矢国际公司(Dyadic International, Inc.))、FIBREZYME® LBR(并矢国际公司)、VISCOSTAR® 150L(并矢国际公司)或 AlternaFuel® CMAX3™(并矢国际公司)。以总固体的从约 0.001 至约 5.0wt.%,更优选总固体的从约 0.025 至约 4.0wt.%,并且最优选总固体的从约 0.005 至约 2.0wt.%的有效量添加这些纤维素分解酶组合物。以总固体的从约 0.001 至约 5.0wt.%,更优选总固体的从约 0.025 至约 4.0wt.%,并且最优选总固体的从约 0.005 至约 2.0wt.%的有效量添加这些纤维素分解酶组合物。

[0370] 内切葡聚糖酶

[0371] 用于本发明的方法或工艺中的纤维素分解酶组合物可以包括任何内切葡聚糖酶。可以在本发明的方法中使用的细菌内切葡聚糖酶的实例包括但不限于:解纤维热酸菌(*Acidothermus cellulolyticus*)内切葡聚糖酶(WO 91/05039;WO 93/15186;美国专利号 5,275,944;WO 96/02551;美国专利号 5,536,655;WO 00/70031;WO 2005/093050);褐色高温双歧菌(*Thermobifida fusca*)内切葡聚糖酶 III(WO 2005/093050);以及褐色高温双歧菌内切葡聚糖酶 V(WO 2005/093050)。

[0372] 可以用于本发明的方法中的真菌内切葡聚糖酶的实例包括但不限于:里氏木霉内切葡聚糖酶 I(潘特拉(Penttilä)等人,1986,基因(Gene)45:253-263;GENBANK™登录号 M15665);里氏木霉内切葡聚糖酶 II(萨洛黑莫(Saloheimo)等人,1988,基因 63:11-22;GENBANK™登录号 M19373);里氏木霉内切葡聚糖酶 III(冈田(Okada)等人,1988,应用与环境微生物学(Appl. Environ. Microbiol.)64:555-563;GENBANK™登录号 AB003694);棘孢曲霉内切葡聚糖酶(黄(Ooi)等人,1990,核酸研究(Nucleic Acids Research)18:5884);白曲霉内切葡聚糖酶(坂本(Sakamoto)等人,1995,当代遗传学(Current Genetics)27:435-439);胡萝卜软腐欧文氏菌内切葡聚糖酶(斯拉哈蒂(Saarilahti)等人,1990,基因 90:9-14);尖孢镰刀菌内切葡聚糖酶(GENBANK™登录号 L29381);灰腐质霉高温变种(*Humicola grisea* var. *thermoidea*)内切葡聚糖酶(GENBANK™登录号 AB003107);热白丝菌(*Melanocarpus albomyces*)内切葡聚糖酶(GENBANK™登录号 MAL515703);粗糙脉孢菌内切葡聚糖酶(GENBANK™登录号 XM_324477);特异腐质霉内切葡聚糖酶 V(SEQ ID NO:20);特异腐质霉内切葡聚糖酶 V 核(斯库莱茵(Schulein),1997,生物技术杂志(J. Biotechnology)57:71-81-213 氨基酸)(即,EG V 核);嗜热毁丝霉 CBS 117.65 内切葡聚糖酶(SEQ ID NO:22);担子菌纲 CBS 495.95 内切葡聚糖

酶 (SEQ ID NO:24) ;担子菌纲 CBS494.95 内切葡聚糖酶 (SEQ ID NO:26) ;土生梭孢霉 NRRL 8126CEL6B 内切葡聚糖酶 (SEQ ID NO:28) ;土生梭孢霉 NRRL 8126CEL6C 内切葡聚糖酶 (SEQ ID NO:30) ;土生梭孢霉 NRRL 8126CEL7C 内切葡聚糖酶 (SEQ ID NO:32) ;土生梭孢霉 NRRL 8126CEL7E 内切葡聚糖酶 (SEQ ID NO:34) ;土生梭孢霉 NRRL 8126CEL7F 内切葡聚糖酶 (SEQ ID NO:36) ;Cladorrhinum foecundissimum ATCC 62373CEL7A 内切葡聚糖酶 (SEQ ID NO:38) ;以及里氏木霉菌株号 VTT-D-80133 内切葡聚糖酶 (SEQ ID NO:40 ;GENBANK™ 登录号 M15665)。上述 SEQ ID NO:20、SEQ ID NO:22、SEQ ID NO:24、SEQ ID NO:26、SEQ ID NO:28、SEQ ID NO:30、SEQ ID NO:32、SEQ ID NO:34、SEQ ID NO:36、SEQ ID NO:38 以及 SEQ ID NO:40 的内切葡聚糖酶分别由 SEQ ID NO:19、SEQ ID NO:21、SEQ ID NO:23、SEQ ID NO:25、SEQ ID NO:27、SEQ ID NO:29、SEQ ID NO:31、SEQ ID NO:33、SEQ ID NO:35、SEQ ID NO:37 以及 SEQ ID NO:39 的成熟多肽编码序列编码。

[0373] 纤维二糖水解酶

[0374] 用于本发明的方法或工艺中的纤维素分解酶组合物可以包括任何纤维二糖水解酶。

[0375] 可用于本发明的方法中的纤维二糖水解酶的实例包括但不限于 :里氏木霉纤维二糖水解酶 I (SEQ ID NO:42) ;里氏木霉纤维二糖水解酶 II (SEQ ID NO:44) ;特异腐质霉纤维二糖水解酶 I (SEQ ID NO:46) ,嗜热毁丝霉纤维二糖水解酶 II (SEQ ID NO:48 和 SEQ ID NO:50) ,土生梭孢霉纤维二糖水解酶 II (CEL6A) (SEQ ID NO:52) ,嗜热毛壳菌纤维二糖水解酶 I (SEQ ID NO:54) ,以及嗜热毛壳菌纤维二糖水解酶 II (SEQ ID NO:56)。上述 SEQ ID NO:40、SEQ ID NO:42、SEQ ID NO:44、SEQ ID NO:46、SEQ ID NO:48、SEQ ID NO:50、SEQ ID NO:52 以及 SEQ ID NO:54 的纤维二糖水解酶分别由 SEQ ID NO:41、SEQ ID NO:43、SEQ ID NO:45、SEQ ID NO:47、SEQ ID NO:49、SEQ ID NO:51、SEQ ID NO:53 以及 SEQ ID NO:55 的成熟多肽编码序列编码。

[0376] β -葡糖苷酶

[0377] 用于本发明的方法或工艺中的纤维素分解酶组合物可以包括任何 β -葡糖苷酶。

[0378] 可用于本发明的方法中的 β -葡糖苷酶的实例包括但不限于 :米曲霉 β -葡糖苷酶 (SEQ ID NO:58) ;烟曲霉 β -葡糖苷酶 (SEQ ID NO:60) ;巴西青霉 IBT 20888 β -葡糖苷酶 (SEQ ID NO:62) ;黑曲霉 β -葡糖苷酶 (SEQ ID NO:64) ;以及棘孢曲霉 β -葡糖苷酶 (SEQ ID NO:66)。上述 SEQ ID NO:58、SEQ ID NO:60、SEQ ID NO:62、SEQ ID NO:64 以及 SEQ ID NO:66 的 β -葡糖苷酶分别由 SEQ ID NO:57、SEQ ID NO:59、SEQ ID NO:61、SEQ ID NO:63 以及 SEQ ID NO:65 的成熟多肽编码序列编码。

[0379] 具有 β -葡糖苷酶活性的米曲霉多肽可根据 WO 02/095014 获得。具有 β -葡糖苷酶活性的烟曲霉菌多肽可根据 WO 2005/047499 获得。具有 β -葡糖苷酶活性的巴西青霉多肽可根据 WO 2007/019442 或在此的 SEQ ID NO:62 获得。具有 β -葡糖苷酶活性的黑曲霉多肽可根据丹 (Dan) 等人,2000,生物化学杂志 (J. Biol. Chem.) 275:4973-4980 获得。具有 β -葡糖苷酶活性的棘孢曲霉多肽可根据川口 (Kawaguchi) 等人,1996,基因 (Gene) 173:287-288 获得。在一个实施例中,该 β -葡糖苷酶可以是一种棘孢曲霉 β -葡糖苷酶,例如披露于在此的 SEQ ID NO:66 中的 β -葡糖苷酶。

[0380] 在一个实施例中, β -葡糖苷酶融合蛋白是一种与在此的 SEQ ID NO:66 具有至

少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的 β - 葡糖苷酶融合蛋白。

[0381] 该 β - 葡糖苷酶可以是一种融合蛋白。在一个方面中, 该 β - 葡糖苷酶是在此的 SEQ ID NO:68 的米曲霉 β - 葡糖苷酶变体 BG 融合蛋白或在此的 SEQ ID NO:70 的米曲霉 β - 葡糖苷酶融合蛋白。在另一个方面中, 该米曲霉 β - 葡糖苷酶变体 BG 融合蛋白由在此的 SEQ ID NO:67 的多核苷酸编码或该米曲霉 β - 葡糖苷酶融合蛋白由在此的 SEQ ID NO:69 编码。

[0382] 在一个实施例中, β - 葡糖苷酶融合蛋白是一种与在此的 SEQ ID NO:68 或 70 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的 β - 葡糖苷酶融合蛋白。

[0383] 在另一个实施例中, 该 β - 葡糖苷酶可以是一种来源于烟曲霉的 β - 葡糖苷酶, 例如, 示于 WO 2005/047499 的 SEQ ID NO:5 或在此的 SEQ ID NO:78 中的 β - 葡糖苷酶或其变体, 例如, 具有以下取代: F100D、S283G、N456E、F512Y (使用 SEQ ID NO:78 用于编号)。

[0384] 在一个实施例中, 该 β - 葡糖苷酶来自曲霉属的菌株, 例如烟曲霉菌株, 例如烟曲霉 β - 葡糖苷酶 (在此的 SEQ ID NO:78), 该 β - 葡糖苷酶包括一个或多个选自下组的取代, 该组由以下各项组成: L89M、G91L、F100D、I140V、I186V、S283G、N456E 以及 F512Y; 例如具有以下取代的其变体:

[0385] -F100D+S283G+N456E+F512Y;

[0386] -L89M+G91L+I186V+I140V;

[0387] -I186V+L89M+G91L+I140V+F100D+S283G+N456E+F512Y (在此使用 SEQ ID NO:78 用于编号)。

[0388] 在一个实施例中, 取代的数目在 1 与 10 之间, 例如在 1 与 8 之间, 例如在 1 与 6 之间, 例如在 1 与 4 之间, 例如 1、2、3、4、5、6、7、8、9 或 10 个取代。

[0389] 在一个实施例中, 该 β - 葡糖苷酶是一种与在此的 SEQ ID NO:78 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的 β - 葡糖苷酶。

[0390] 在一个实施例中, 该 β - 葡糖苷酶变体是一种与在此的 SEQ ID NO:78 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的 β - 葡糖苷酶变体。

[0391] 其他内切葡聚糖酶、纤维二糖水解酶、以及 β - 葡糖苷酶披露于使用根据以下的分类的许多糖基水解酶家族中: 亨利萨特 (Henrissat), 1991, 基于氨基酸序列相似性的糖基水解酶的分类 (A classification of glycosyl hydrolases based on amino-acid sequence similarities), 生物化学杂志 (Biochem. J.) 280:309-316; 以及亨利萨特和贝洛赫 (Bairoch), 1996, 修正糖基水解酶的基于序列的分类 (Updating the sequence-based classification of glycosyl hydrolases), 生物化学杂志 316:695-696。

[0392] 可用在本发明中的其他纤维素分解酶描述于以下中:EP 495, 257、EP531, 315、EP 531, 372、WO 89/09259、WO 94/07998、WO 95/24471、WO 96/11262、WO 96/29397、WO 96/034108、WO 97/14804、WO 98/08940、WO 98/12307、WO 98/13465、WO 98/15619、WO 98/15633、WO 98/28411、WO 99/06574、WO 99/10481、WO 99/25846、WO 99/25847、WO 99/31255、WO 00/09707、WO 02/050245、WO 02/076792、WO 02/101078、WO 03/027306、WO 03/052054、WO 03/052055、WO 03/052056、WO 03/052057、WO 03/052118、WO 2004/016760、WO 2004/043980、WO 2004/048592、WO 2005/001065、WO 2005/028636、WO 2005/093050、WO 2005/093073、WO 2006/074005、WO 2006/117432、WO 2007/071818、WO 2007/071820、WO 2008/008070、WO 2008/008793、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、以及美国专利号 5, 776, 757。

[0393] 木聚糖酶

[0394] 用于本发明的方法或工艺中的纤维素分解酶组合物可以包括任何木聚糖酶。

[0395] 适用于本发明的商业化木聚糖降解酶制剂的实例包括例如 SHEARZYME™(诺维信公司)、CELLIC™ HTec(诺维信公司)、VISCOZYME®(诺维信公司)、ULTRAFLO®(诺维信公司)、PULPZYME® HC(诺维信公司)、MULTIFECT®木聚糖酶(杰能科(Genencor))、ECOPULP® TX-200A(AB 酶制剂公司(AB Enzymes))、HSP 6000 木聚糖酶(帝斯曼)、DEPOL™333P(生物催化剂有限公司(Biocatalysts Limit), 英国威尔士(Wales, UK))、DEPOL™740L(生物催化剂有限公司, 英国威尔士)以及 DEPOL™762P(生物催化剂有限公司, 英国威尔士)。

[0396] 可用于本发明的方法中的木聚糖酶的实例包括但不限于:棘孢曲霉木聚糖酶(GeneSeqP:AAR63790;WO 94/21785), 烟曲霉木聚糖酶(例如, 在 WO 2006/078256 中示为 SEQ ID NO:6 或在此的 SEQ ID NO:75 的 Xyl III), 以及土生梭孢霉 NRRL 8126 木聚糖酶(WO 2009/079210)。

[0397] β-木糖苷酶

[0398] 用于本发明的方法或工艺中的纤维素分解酶组合物可以包括任何 β-木糖苷酶。

[0399] 可用于本发明的方法的 β-木糖苷酶的实例包括但不限于:里氏木霉 β-木糖苷酶(UniProtKB/TrEMBL 登录号 Q92458)、埃默森踝节菌(SwissProt 登录号 Q8X212)、以及粗糙脉孢菌(SwissProt 登录号 Q7S0W4)。

[0400] 乙酰木聚糖酯酶

[0401] 用于本发明的方法或工艺中的纤维素分解酶组合物可以包括任何乙酰木聚糖酯酶。

[0402] 可用于本发明的方法的乙酰木聚糖酯酶的实例包括但不限于:红褐肉座菌(Hypocrea jecorina), 乙酰木聚糖酯酶(WO 2005/001036), 粗糙脉孢菌乙酰木聚糖酯酶(UniProt 登录号 q7s259), 土生梭孢霉 NRRL 8126 乙酰木聚糖酯酶(WO 2009/042846), 球毛壳菌乙酰木聚糖酯酶(Uniprot 登录号 Q2GWX4), 细毛壳菌(Chaetomium gracile)乙酰木聚糖酯酶(GeneSeqP 登录号 AAB82124), 颖枯壳针孢(Phaeosphaeria nodorum)乙酰木聚糖酯酶(Uniprot 登录号 Q0UHJ1), 以及特异腐质霉 DSM 1800 乙酰木聚糖酯酶(WO 2009/073709)。

[0403] 阿魏酸酯酶

[0404] 用于本发明的方法或工艺中的纤维素分解酶组合物可以包括任何阿魏酸酯酶。

[0405] 可用于本发明的方法的阿魏酸酯酶的实例包括但不限于：特异腐质霉 DSM 1800 阿魏酸酯酶 (WO 2009/076122)、粗糙脉孢菌阿魏酸酯酶 (UniProt 登录号 Q9HGR3)、以及费希新萨托菌 (*Neosartorya fischeri*) 阿魏酸酯酶 (UniProt 登录号 A1D9T4)。

[0406] 阿拉伯呋喃糖苷酶

[0407] 用于本发明的方法或工艺中的纤维素分解酶组合物可以包括任何阿拉伯呋喃糖苷酶。

[0408] 可用于本发明的方法中的阿拉伯呋喃糖苷酶的实例包括但不限于：特异腐质霉 DSM 1800 阿拉伯呋喃糖苷酶 (WO 2009/073383) 以及黑曲霉阿拉伯呋喃糖苷酶 (GeneSeqP 登录号 AAR94170)。

[0409] α -葡糖醛酸糖苷酶

[0410] 用于本发明的方法或工艺中的纤维素分解酶组合物可以包括任何 α -葡糖醛酸糖苷酶。

[0411] 可用于本发明的方法的 α -葡糖醛酸糖苷酶的实例包括但不限于：棒曲霉 α -葡糖醛酸糖苷酶 (UniProt 登录号 alccl2)、里氏木霉 α -葡糖醛酸糖苷酶 (UniProt 登录号 Q99024)、埃默森踝节菌 α -葡糖醛酸糖苷酶 (UniProt 登录号 Q8X211)、黑曲霉 α -葡糖醛酸糖苷酶 (UniProt 登录号 Q96WX9)、土曲霉 α -葡糖醛酸糖苷酶 (SwissProt 登录号 Q0CJP9)、以及烟曲霉 α -葡糖醛酸糖苷酶 (SwissProt 登录号 Q4WW45)。

[0412] 酶和多肽的生产

[0413] 有用于本发明的方法的酶和蛋白质可以通过在含有适合碳源和氮源以及无机盐的营养培养基上,使用本领域中已知的程序发酵上述微生物菌株来产生(参见,例如,班尼特·J.W. (Bennett, J.W.) 和拉热·L. (LaSure, L.) (编辑),真菌中的更多基因操纵 (More Gene Manipulations in Fungi), 学术出版社,加利福尼亚州,1991)。适合的培养基可从商业供应商获得或可以根据公开的组成(例如,在美国典型培养物保藏中心的目录中)制备。适合用于生长和酶产生的温度范围和其他条件在本领域中是已知的(参见,例如,贝利·J.E. (Bailey, J.E.) 和奥利斯·D.F. (Ollis, D.F.), 生物化学工程基础 (Biochemical Engineering Fundamentals), 麦格劳-希尔图书公司 (McGraw-Hill Book Company), 纽约, 1986)。

[0414] 本发明的组合物

[0415] 在一个最终方面中,本发明涉及一种组合物。该组合物是一种具有至少三种组分的共混物或混合物。根据本发明的方法或工艺,可以在进行水解之前和/或水解过程中添加该组合物。在其中本发明的组合物不包括一种如在此定义的纤维素分解酶组合物的实施例中,它可以与一种纤维素分解酶组合物一起添加至水解中。它典型地与该纤维素分解酶组合物同时和/或其后添加,但是也可以在水解之前添加。

[0416] 更确切而言,本发明的组合物包括或其组成为:

[0417] i) 一种多肽,该多肽具有纤维素分解增强活性;

[0418] ii) 一种过氧化物酶;

[0419] iii) 一种非离子表面活性剂和/或一种阳离子表面活性剂。

[0420] 具有纤维素分解增强活性的多肽可以是一种披露于以上“具有纤维素分解增强活性的多肽”部分中的多肽。

[0421] 该过氧化物酶可以是一种披露于以上“过氧化物酶”部分中的过氧化物酶。

[0422] 该非离子和阳离子表面活性剂可以是一种披露于以上“非离子表面活性剂”或“阳离子表面活性剂”部分中的非离子和阳离子表面活性剂。

[0423] 在一个实施例中,具有纤维素分解增强活性的多肽是一种 GH61 多肽。在一个实施例中,具有纤维素分解增强活性的该多肽是一种来源于嗜热子囊菌属,如金黄色嗜热子囊菌菌株的多肽,例如,在 WO 2005/074656 中描述为 SEQ ID NO:2 或在此的 SEQ ID NO:14 的多肽。

[0424] 在一个实施例中,具有纤维素分解增强活性的该多肽与在此的 SEQ ID NO:14 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性。

[0425] 在一个实施例中,具有纤维素分解增强活性的该多肽是一种来源于梭孢壳属,例如土生梭孢壳菌株的多肽,例如,在 WO 2005/074647 中描述为 SEQ ID NO:7 和 SEQ ID NO:8 的多肽。在一个实施例中,具有纤维素分解增强活性的该多肽是一种来源于曲霉属的菌株,例如烟曲霉菌株的多肽,例如,在 WO 2010/138754 中描述为 SEQ ID NO:1 和 SEQ ID NO:2 的多肽。在一个实施例中,具有纤维素分解增强活性的该多肽是一种来源于青霉属的菌株,例如埃默森青霉菌菌株的多肽,例如,在 WO 2011/041397 中披露为 SEQ ID NO:2 或在此的 SEQ ID NO:72 的多肽。

[0426] 在一个实施例中,该过氧化物酶选自下组,该组包括过氧化物酶或过氧化物分解酶,包括但不限于以下各项:E.C. 1.11.1.1 NADH 过氧化物酶;E.C. 1.11.1.2 NADPH 过氧化物酶;E.C. 1.11.1.3 脂肪酸过氧化物酶;E.C. 1.11.1.5 细胞色素 c 过氧化物酶;E.C. 1.11.1.5;E.C. 1.11.1.6 过氧化氢酶;E.C. 1.11.1.7 过氧化物酶;E.C. 1.11.1.8 碘化物过氧化物酶;E.C. 1.11.1.9 谷胱甘肽过氧化物酶;E.C. 1.11.1.10 氯化物过氧化物酶;E.C. 1.11.1.11 L-抗坏血酸过氧化物酶;E.C. 1.11.1.12 磷脂过氧化氢谷胱甘肽过氧化物酶;E.C. 1.11.1.13 锰过氧化物酶;E.C. 1.11.1.14 木质素过氧化物酶;E.C. 1.11.1.15 过氧化物氧还蛋白;E.C. 1.11.1.16 通用过氧化物酶;E.C. 1.11.1.B2 氯化物过氧化物酶;E.C. 1.11.1.B6 碘化物过氧化物酶(含钒);E.C. 1.11.1.B7 溴化物过氧化物酶;E.C. 1.11.1.B8 碘化物过氧化物酶。

[0427] 在一个优选实施例中,该过氧化物酶是一种 E.C. 1.11.1.7 过氧化物酶。

[0428] 在一个实施例中,该过氧化物酶来源于一种微生物,例如真菌生物体,例如酵母或丝状真菌,或细菌;或植物。

[0429] 在一个实施例中,该过氧化物酶来源于鬼伞属的菌株,例如灰盖鬼伞菌株,例如分类为 EC 1.11.1.7 的过氧化物酶,例如示于在此的 SEQ ID NO:71 中的过氧化物酶(即, CiP)。在一个实施例中,该过氧化物酶与在此的 SEQ ID NO:71 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性。

[0430] 在一个实施例中,该非离子表面活性剂是烷基或芳基。在一个实施例中,该非离子表面活性剂选自以下各项的组:甘油醚、乙二醇醚、乙醇酰胺、磺酰胺、醇、酰胺、醇乙氧基化物、甘油酯、乙二醇酯、甘油酯和乙二醇酯的乙氧基化物、糖基烷基多糖苷、聚氧乙烯化脂肪酸、烷醇胺缩合物、烷醇酰胺、叔炔属乙二醇、聚氧乙烯化硫醇、羧酸酯、以及聚氧乙烯化聚氧丙烯乙二醇(例如EO/PO嵌段共聚物(EO是环氧乙烷,PO是环氧丙烷))、EO聚合物和共聚物、多胺、和聚乙烯吡咯烷酮。

[0431] 在一个实施例中,该非离子表面活性剂是一种直链伯或仲或具有化学式 $\text{RO}(\text{CH}_2\text{CH}_2\text{O})_n\text{H}$ 的支链醇乙氧基化物,其中R是烃链长度并且n是环氧乙烷的摩尔平均数,例如其中R是范围从C9至C16的直链伯或支链仲烃链长度并且n范围从6至13,例如醇乙氧基化物,其中R是直链C9-C11烃链长度,并且n是6。

[0432] 在一个实施例中,该阳离子表面活性剂选自以下各项的组:伯、仲或叔胺,例如二盐酸奥替尼啶;烷基三甲基铵盐,例如鲸蜡基三甲基溴化铵(CTAB)(又名十六烷基三甲基溴化铵)、十六烷基三甲基氯化铵(CTAC)、氯化十六烷基吡啶(CPC)、苯扎氯铵(BAC)、苜蓿氯铵(BZT)、5-溴-5-硝基-1,3-二噁烷、二甲基双十八烷基氯化铵、双十八烷基二甲基溴化铵(DODAB)。

[0433] 在一个实施例中,本发明的组合物进一步包括一种纤维素分解酶组合物。

[0434] 在一个实施例中,本发明的组合物包括一种 β -葡糖苷酶。

[0435] 在一个实施例中,该纤维素分解酶组合物包括一种 β -葡糖苷酶,优选来源于以下各项的 β -葡糖苷酶:曲霉属的菌株,如米曲霉,如披露于WO 02/095014中的 β -葡糖苷酶或披露于WO 2008/057637中具有 β -葡糖苷酶活性的融合蛋白,或烟曲霉,如披露于WO 2005/047499,例如在此的SEQ ID NO:78中的 β -葡糖苷酶,或披露于WO 2012/044915中的烟曲霉 β -葡糖苷酶变体(参见以上变体);或青霉属的菌株,如披露于WO 2007/019442或在此的SEQ ID NO:62中的巴西青霉菌株,或木霉属的菌株,如里氏木霉菌株。

[0436] 在一个实施例中,该纤维素分解酶组合物来源于里氏木霉、特异腐质霉、或拉克淖金孢子菌、或嗜热毁丝霉。

[0437] 在一个更具体实施例中,本发明的组合物包括或其组成为:

[0438] i) 一种具有纤维素分解增强活性的多肽,优选是来源于金黄色嗜热子囊菌的如在此的SEQ ID NO:14所示的多肽,和/或来源于埃默森青霉菌的示于在此的SEQ ID NO:72中的多肽,或一种具有纤维素分解增强活性与在此的SEQ ID NO:14或在此的SEQ ID NO:72具有至少60%、至少70%、至少80%、至少90%、至少95%、至少97%、至少99%序列一致性的多肽:

[0439] ii) 一种分类为EC 1.11.1.7过氧化物酶的过氧化物酶,优选是来源于灰盖鬼伞示于在此的SEQ ID NO:71中的过氧化物酶;或一种具有过氧化物酶活性与在此的SEQ ID NO:71具有至少60%、至少70%、至少80%、至少90%、至少95%、至少97%、至少99%一致性的多肽:

[0440] iii) 一种非离子表面活性剂和/或一种阳离子表面活性剂。

[0441] 在一个实施例中,该组合物还包括一种纤维素分解酶组合物,尤其是在此定义的纤维素分解酶组合物。

[0442] 通过以下实例进一步描述本发明,这些实例不应当解释为限制本发明的范围。

[0443] 材料与方法

[0444] 半纤维素酶 3 (“HEMI 3”)是由里氏木霉菌株重组产生的半纤维素酶组合物并且含有一种来源于烟曲霉的热稳定木聚糖酶 GH10 和一种烟曲霉 β -木糖苷酶。

[0445] 纤维素分解酶组合物由里氏木霉菌株产生。

[0446] 纤维素分解酶组合物 2 (“Tr Cel 2”);具有 Af CBHI 和 Af CBHII 的里氏木霉

[0447] 购自西格玛公司 (SIGMA) (P2088-10KU) 的辣根过氧化物酶 (“HrP”) (254 单位 / mg 固体)

[0448] 用于 HrP 的西格玛单位定义

[0449] 一个邻苯三酚单位在 pH 6.0 下,在 20℃下,将在 20sec 内从邻苯三酚形成 1.0mg 红棕酚。

[0450] 购自西格玛公司 (42603-10MG-F) 的木质素过氧化物酶 (“LiP”) (0.1 单位 /mg 固体)

[0451] 用于 LiP 的西格玛单位定义

[0452] 一个单位对应于在 pH 3.0 和 30℃下,每分钟氧化 1 微摩尔 3,4-二甲氧基苯基醇的酶量

[0453] 大豆过氧化物酶 (“Soy P”)

[0454] 王棕过氧化物酶 (“RpP”) 示于 SEQ ID NO:79 中。

[0455] 表 1 酶概述

[0456]

酶	遗传来源	缩写
内切葡聚糖酶 V 核	特异腐质霉	EG V 核
CBH I	烟曲霉	AfCBH I
CBH II	烟曲霉	AfCBH II
GH61a	埃默森青霉菌	PeGH61a
GH61a	金黄色嗜热子囊菌	TaGH61a
β -葡糖苷酶	棘孢曲霉	AaBG
β -葡糖苷酶变体 (F100D、S283G、N456E、F512Y 取代)	烟曲霉	AfBG 4M
过氧化物酶 (EC 1.11.1.7)	灰盖鬼伞	CiP

[0457] 表 2 表面活性剂概述

[0458]

表面活性剂	结构	供应商	分类
意格倍 (Igepal) CO 730	壬基酚乙氧基化物	罗地亚公司 (Rhodia)	非离子
Triton X 100 (目录号 T9284)	$C_{14}H_{22}O(C_2H_4O)_n$	西格玛公司	非离子
诺威 (Novel) II TDA-10	C13-醇聚乙二醇醚 (10 EO)	萨索尔公司 (Sasol)	非离子
杰福克斯 (Jeffox) WL-5000	EO, PO 共聚物	亨斯迈公司 (Huntsman)	非离子
Levapon 李维本 150N	烷基聚乙二醇醚	拜耳公司 (Bayer)	非离子
鲁腾索 (Lutensol) TO5	$RO(EO)_5H$	BASF	非离子
PEG 8000 (目录号 89510)	$HOCH_2(EO)_nCH_2OH$	西格玛公司	非离子
PEG 35000 (目录号 94646)	$HOCH_2(EO)_nCH_2OH$	西格玛公司	非离子
氯化十六烷基吡啶 (目录号	$C_{21}H_{38}NCl$	西格玛公司	阳离子

[0459]

C0732)			
十六烷基三甲基溴化铵 (目录号 H9151)	$CH_3(CH_2)_{15}N(CH_3)_3Br$	西格玛公司	阳离子

[0460] 预处理的玉米秸秆的制备

[0461] 在美国能源部国家可再生能源实验室 (NREL), 使用稀硫酸预处理玉米秸秆。使用以下条件用于预处理: 5% 硫酸 (w/w, 在干玉米秸秆基础上), 在 180°C 下持续 4 分钟。通过遵循由 NREL 研发的标准分析程序确定该预处理的玉米秸秆 (PCS) 的组成和不溶性固体分数 (FIS) (斯勒伊特 (Sluiter) 等人, 2008a, 生物质中的总固体和液态处理样品中的总溶解固体的确定 (Determination of Total Solids in Biomass and Total Dissolved Solids in Liquid Process Samples) NREL/TP-510-42621, 国家可再生研究实验室 (National Renewable Research Laboratory), 戈尔登 (Golden), 哥伦比亚州 www.nrel.gov/biomass/pdfs/42621.pdf。

[0462] 斯勒伊特等人 2008b, 生物质中的结构性碳水化合物和木质素的确定 (Determination of structural carbohydrates and lignin in biomass), 实验室分析程序 (Laboratory Analytical Procedures), NREL/TP-510-42618, 国家可再生研究实验室 (National Renewable Research Laboratory), 戈尔登 (Golden), 哥伦比亚州 www.nrel.gov/biomass/pdfs/42618.pdf。

[0463] 斯勒伊特等人, 2008c, 生物质中的总固体和液态处理样品中的总溶解固体的确

定。实验室分析程序 (Laboratory Analytical Procedures), NREL/TP-510-42621, 国家可再生能源研究实验室 (National Renewable Research Laboratory), 戈尔登 (Golden), 哥伦比亚州 www.nrel.gov/biomass/pdfs/42621.pdf)。

[0464] PCS 中的水不溶性固体含有 62% 葡聚糖、2% 木聚糖以及 29.7% 酸不溶性木质素。发现 PCS 的 FIS 为 56%。

[0465] PCS 的酶水解

[0466] 在 50mL 耐洁 (Nalgene) 聚碳酸酯离心管 (赛默飞世尔科技公司 (Thermo Scientific), 匹兹堡, 宾夕法尼亚) 中进行批次酶水解。将 PCS 与 50mM 补充有酶、表面活性剂 (按需) 以及用于阻止微生物生长的 2.5mg/L lactol 的乙酸钠缓冲液 (pH 5.0) 混合。在此研究中使用的所有酶和表面活性剂都概述于表 1 和 2 中。除非另外说明, 否则最终的总固体浓度为 20% (w/w, 在干重基础上)。将反应混合物 (20g) 在 50°C 下于杂交培养箱 (Combi-D24, FINEPCR®, 钟杨 (Yang-Chung), 首尔, 韩国) 中搅拌 120 小时。在水解结束时, 将 600 μ L 的水解产物转移至 Costar Spin-X 离心过滤管 (科尔-帕默公司 (Cole-Parmer), 弗农希尔斯, 伊利诺伊州) 并且在离心 (14,000rpm, 20 分钟) 过程中通过 0.2 μ m 尼龙过滤器过滤。将上清液用 5 μ L 的 40% (w/v) 硫酸酸化, 以钝化剩余的酶活性并且通过高效液相层析 (HPLC) 分析糖浓度。

[0467] 糖的分析

[0468] 用一个 HPLC 系统 (1200 系列 LC 系统, 安捷伦科技公司 (Agilent Technologies Inc.), 帕洛阿尔托, 加利福尼亚州) 分析自 PCS 的水解释放的糖, 该系统配备有一个 Rezex ROA- 有机酸 H⁺ 柱 (8%) (7.8 $\times$ 300mm) (菲罗门公司 (Phenomenex Inc.), 托伦斯, 加利福尼亚州)、0.2 μ m 线内过滤器、一个自动取样器、一个梯度泵以及一个屈光指数检测器。使用的流动相是处于 0.9ml/min 的流速的 5mM 硫酸。将处于 0、10、30 和 50mg/L 的浓度的单体糖用作标准品。使用类似于由朱 (Zhu) 等人, 2011, 计算生物物质的高固体水解的糖产率 (Calculating sugar yields in high solids hydrolysis of biomass), 生物资源技术 (Bioresource Technology) 102(3):2897-2903 公开的方法, 基于酶水解上清液中的糖和新进料的生物质组成计算来自预处理和水解的总体葡聚糖 / 木聚糖转化。

[0469] 实例

[0470] 实例 1

[0471] 使用过氧化物酶和非离子表面活性剂增强自预处理的玉米秸秆 (PCS) 生产糖

[0472] 在 50°C、pH 5 下, 在 20% (w/w, 在干重基础上) 总固体装载下进行 PCS 的水解。使用三种酶混合物: 纤维素分解酶组合物, 一种含有 10% EG V 核、40% AfCBHI、30% AfCBHII、5% AfBG 4M、10% TaGH61a 和 5% 半纤维素酶的纤维素酶混合物, 以及一种含有 10% EG、40% AfCBHI、30% AfCBHII、5% AaBG、10% TaGH61a 和 5% 半纤维素酶的纤维素酶混合物。GH61、纤维素酶和半纤维素酶的总蛋白剂量是 4mg/g PCS 纤维素。样品还补充有 CiP (120 μ g/g PCS 纤维素)、李维本非离子表面活性剂 (2% w/w, 在干 PCS 基础上) 和处于类似于表 3 中概括的剂量的过氧化物酶和非离子表面活性剂的组合。在第 72 和 120 小时取出样品并且如通过 HPLC 所述进行分析。

[0473] 表 3 实验设计: 由 CiP 和非离子表面活性剂增强水解产率

[0474]

样品 ID	纤维素分解酶组合物%	EG V 核%	AfCB HI %	AfCB H II %	AaB G %	TaGH61 a %	AfBG 4M %	半纤维素 3 %	CiP %	Levapon 李维本 150N %
1	100									
2	100								3	
3	100									2
4	100								3	2
5		10	40	30		10	5	5		
6		10	40	30		10	5	5	3	
7		10	40	30		10	5	5		2
8		10	40	30		10	5	5	3	2
9		10	40	30	5	10		5		
10		10	40	30	5	10		5	3	
11		10	40	30	5	10		5		2
12		10	40	30	5	10		5	3	2

[0475] 如示于图 1 中的结果证明,添加 CiP (120 μ g/g PCS 纤维素) 和李维本非离子表面活性剂 (2% w/w, 在干 PCS 基础上) 两者在孵育 120 小时后, 分别将葡萄糖的水解产率增加 3-5g/L 和 4-7g/L。然而, 在过氧化物酶与非离子表面活性剂之间存在协同作用。当一起给予过氧化物酶和表面活性剂两者时, 总葡萄糖产率增加 14-18g/L, 这显著高于由单独的过氧化物酶或表面活性剂的助强作用 (boosting effect) 的组合。

[0476] 实例 2

[0477] 过氧化物酶与非离子表面活性剂之间的协同作用对 GH61 水平的依赖性

[0478] 在 50°C、pH 5 下, 在 20% (w/w, 在干重基础上) 总固体装载下进行 PCS 的水解。水解酶是处于不同比例的 EG V 核、AfCBHI、AfCBHII、AaBG 及半纤维素酶的组合。TaGH61a 的浓度如概述于表 4 中在 0-20% 之间变化。GH61、纤维素酶和半纤维素酶的总蛋白剂量是 4mg/g PCS 纤维素。样品还补充有 CiP (120 μ g/g PCS 纤维素)、李维本非离子表面活性剂 (2% w/w, 在干 PCS 基础上) 和处于类似剂量的过氧化物酶和非离子表面活性剂的组合 (表 4)。在第 72 和 120 小时取出样品并且如通过 HPLC 所述进行分析。

[0479] 表 4 实验设计: 由 CiP 和非离子表面活性剂增强水解产率

[0480]

样品 ID	EG V 核%	AfCBH I %	AfCBH II %	AaBG %	半纤维素 酶 3 %	TaGH61a %	CiP %	Levapon 李维本 150 %
1	11.11	44.44	33.33	5.56	5.56	0		
2	11.11	44.44	33.33	5.56	5.56	0	3	
3	11.11	44.44	33.33	5.56	5.56	0		2
4	11.11	44.44	33.33	5.56	5.56	0	3	2
5	10.56	42.22	31.67	5.28	5.28	5		
6	10.56	42.22	31.67	5.28	5.28	5	3	
7	10.56	42.22	31.67	5.28	5.28	5		2
8	10.56	42.22	31.67	5.28	5.28	5	3	2
9	10.00	40.00	30.00	5.00	5.00	10		
10	10.00	40.00	30.00	5.00	5.00	10	3	
11	10.00	40.00	30.00	5.00	5.00	10		2
12	10.00	40.00	30.00	5.00	5.00	10	3	2
13	8.89	35.56	26.67	4.44	4.44	20		
14	8.89	35.56	26.67	4.44	4.44	20	3	
15	8.89	35.56	26.67	4.44	4.44	20		2
16	8.89	35.56	26.67	4.44	4.44	20	3	2

[0481] 如示于图 2 中的结果证明,当 GH61a 水平的范围从 0-20%,存在协同作用。含有 5% GH61a 的酶显示出最大协同。当一起给予过氧化物酶和表面活性剂两者时,葡萄糖浓度增加 23g/L,这显著高于由单独的过氧化物酶(大约 7g/L)或表面活性剂(大约 6.7g/L)的助强作用的组合。

[0482] 实例 3

[0483] GH61 的来源对过氧化物酶与非离子表面活性剂之间的协同作用的影响

[0484] 在 50℃、pH 5 下,在 20% (w/w,在干重基础上)总固体装载下进行 PCS 的水解。水解酶是处于示于表 5 中的比例的 EG V 核、AfCBHI、AfCBHII、AaBG、半纤维素酶以及来自金黄色嗜热子囊菌或埃默森青霉菌的 GH61a 的组合。GH61、纤维素酶和半纤维素酶的总蛋白剂量是 3mg/g PCS 纤维素。样品还补充有 CiP(90 μg/g PCS 纤维素)、李维本非离子表面活性剂(2% w/w,在干 PCS 基础上)和处于类似剂量的过氧化物酶和非离子表面活性剂的组合(表 5)。在第 72 和 120 小时取出样品并且如通过 HPLC 所述进行分析。

[0485] 表 5 实验设计:来自不同遗传来源的 GH61a 的比较

[0486]

样品 ID	EG V 核%	AfCBH I %	AfCBH II %	AaBG %	半纤维素 酶 3 %	TaGH61 a %	PeGH6 1 %	CiP %	Levapon 李维 本 150 %
1	10	37.5	37.5	5	5	5			
2	10	37.5	37.5	5	5	5			2
3	10	37.5	37.5	5	5	5		3	
4	10	37.5	37.5	5	5	5		3	2
7	10	37.5	37.5	5	5		5		
8	10	37.5	37.5	5	5		5		2
9	10	37.5	37.5	5	5		5	3	
10	10	37.5	37.5	5	5		5	3	2

[0487] 图 3 示出了水解 120 小时后的结果。含有金黄色嗜热子囊菌或埃默森青霉菌 GH61a 的酶混合物都观察到了协同作用。

[0488] 实例 4

[0489] 表面活性剂的化学结构对协同作用的影响

[0490] 在 50℃、pH 5 下,在 20% (w/w,在干重基础上) 总固体装载下进行 PCS 的水解。对于非离子表面活性剂,使用 4mg/g 纤维素的纤维素分解酶。样品还补充有 CiP(120 μg/g PCS 纤维素)、非离子表面活性剂 (2% w/w,在干 PCS 基础上) (表 2) 和处于类似剂量的过氧化物酶和非离子表面活性剂的组合 (表 6)。对于阳离子表面活性剂,水解酶是 EG V 核、AfCBHI、AfCBHII、AaBG、TaGH61a 及半纤维素酶的组合 (表 7)。将总酶剂量维持在 3mg/g 纤维素。样品还补充有 CiP(90 μg/g PCS 纤维素)、阳离子表面活性剂 (2% w/w,在干 PCS 基础上) (表 2) 和处于类似剂量的过氧化物酶和阳离子表面活性剂的组合 (表 7)。在第 72 和 120 小时取出样品并且如通过 HPLC 所述进行分析。

[0491] 表 6 实验设计:非离子表面活性剂的结构对协同作用的影响

[0492]

样品 ID	纤维素 分解酶 组合物	意格倍 (Igepal CO_730)	瑾 璨 (Triton X100)	诺威龙 (Novel 11) 10 %	鲁腾索 (Lutensol)	杰福克斯 (Jeffox) WL-5000	PEG 8K %	PEG 35K %	Levapon 李维本 150N %	CiP %
-------	-------------------	---------------------------	-------------------------	---------------------------	-----------------------	-----------------------------	----------------	-----------------	--------------------------	-------

[0493]

	%		%		TO-5 %	%				
1	100									
2	100	2								
3	100	2								3
4	100		2							
5	100		2							3
6	100			2						
7	100			2						3
8	100				2					
9	100				2					3
10	100					2				
11	100					2				3
12	100						2			
13	100						2			3
14	100							2		
15	100							2		3
16	100								2	
17	100								2	3
18	100									3

[0494] 表 7 实验设计 :阳离子表面活性剂的结构对协同作用的影响

[0495]

样品 ID	EG V 核 %	AfCBH I %	AfCBHI I %	AaB G %	半纤维素酶 3 %	TaGH6 1a %	CiP%	CH ₃ (CH ₂) ₁₅ N(CH ₃) ₃ Br %	C ₂₁ H ₃₈ NCl %
1	10	37.5	37.5	5	5	5		1	
2	10	37.5	37.5	5	5	5		2	
3	10	37.5	37.5	5	5	5			1
4	10	37.5	37.5	5	5	5			2
5	10	37.5	37.5	5	5	5	3	1	
6	10	37.5	37.5	5	5	5	3	2	
7	10	37.5	37.5	5	5	5	3		1
8	10	37.5	37.5	5	5	5	3		2

[0496] 结果显示,在所有非离子表面活性剂与过氧化物酶之间都存在协同作用(图4)。在所测试的阳离子表面活性剂与过氧化物酶之间也观察到了类似结果(图5)。阳离子表面

活性剂的协同作用较不显著。

[0497] 实例 5

[0498] 表面活性剂的剂量对过氧化物酶与非离子表面活性剂之间的协同作用的影响

[0499] 在 50℃、pH 5 下,在 20% (w/w,在干重基础上) 总固体装载下进行 PCS 的水解。使用含有处于 4mg/g 纤维素的 EG V 核、AfCBHI、AfCBHII、AaBG、TaGH61a 及半纤维素酶的水解酶 (表 8)。样品还补充有 CiP (120 μg/g PCS 纤维素)、李维本非离子表面活性剂 (0-2% w/w,在干 PCS 基础上) 和处于类似剂量的过氧化物酶和非离子表面活性剂的组合 (表 8)。在第 72 和 120 小时取出样品并且如通过 HPLC 所述进行分析。

[0500] 表 8 实验设计:表面活性剂剂量对协同作用的影响

[0501]

样品 ID	EG V 核 %	AfCB HI %	AfCB HII %	AaBG %	半纤维素酶 3 %	TaGH61 a %	CiP %	Levapon 李维本 150N %
1	10	40	30	5	5	10		
2	10	40	30	5	5	10	3	
3	10	40	30	5	5	10		0.25
4	10	40	30	5	5	10	3	0.25
5	10	40	30	5	5	10		0.5
6	10	40	30	5	5	10	3	0.5
7	10	40	30	5	5	10		1
8	10	40	30	5	5	10	3	1
9	10	40	30	5	5	10		1.5
10	10	40	30	5	5	10	3	1.5
11	10	40	30	5	5	10		2.0
12	10	40	30	5	5	10	3	2.0

[0502] 结果 (图 6) 显示,所有测试的非离子表面活性剂剂量都存在协同作用。当非离子表面活性剂在 1% -2% 之间时,观察到最显著协同。

[0503] 实例 6

[0504] 纤维素酶组成对过氧化物酶与非离子表面活性剂之间的协同作用的影响

[0505] 在 50℃,pH 5 下,用含有 EG V 核、AfCBH、AaBG、TaGH61a 及半纤维素酶的纤维素酶单组分混合物进行水解。将总 CBH 剂量维持在 70% 的总纤维素酶和半纤维素酶剂量 (3mg/g 纤维素)。CBHI 与 CBHII 的比例从 0:70 至 70:0 变化 (表 9)。样品还补充有 CiP (90 μg/g PCS 纤维素)、李维本非离子表面活性剂 (2% w/w,在干 PCS 基础上) 和处于类似剂量的过氧化物酶和非离子表面活性剂的组合。在第 72 和 120 小时取出样品并且如通过 HPLC 所述进行分析。

[0506] 表 9 实验设计:纤维素酶组成对协同作用的影响

[0507]

样品 ID	EG 核 %	V AfCBHI %	AfCBHI I %	AaBG %	半纤维素酶 3 %	TaGH61a %	CiP %	Levapon 李维本 150N %
1	10	0	70	5	5	10		
2	10	0	70	5	5	10		2
3	10	0	70	5	5	10	3	
4	10	0	70	5	5	10	3	2
5	10	15	55	5	5	10		
6	10	15	55	5	5	10		2
7	10	15	55	5	5	10	3	
8	10	15	55	5	5	10	3	2
9	10	35	35	5	5	10		
10	10	35	35	5	5	10		2
11	10	35	35	5	5	10	3	
12	10	35	35	5	5	10	3	2
13	10	55	15	5	5	10		
14	10	55	15	5	5	1		2
15	10	55	15		5	10	3	
16	10	55	15	5	5	10	3	2
17	10	70	0	5	5	10		
18	10	70	0	5	5	10		2
19	10	70	0	5	5	10	3	
20	10	70	0	5	5	10	3	2

[0508] 图 7 示出了水解 120 小时后的结果。含有不同量的 CBHI 和 CBHII 的所有纤维素酶混合物都观察到了协同作用。

[0509] 实例 7

[0510] 50℃下,不同木质纤维素底物上的过氧化物酶与非离子表面活性剂之间的协同作用

[0511] 表 10 概述了预处理方法和在此研究中测试的木质纤维素底物的组成。在预处理与水解之间不进行底物的洗涤。在不同固体装载下,用 5mg/g 纤维素的纤维素分解酶进行不同底物的水解(表 11)。5mg/g 纤维素的酶是基于针对芦竹和混合木材的预处理底物中的纤维素,而在热水和稀酸处理的玉米秸秆中,它是基于原料玉米秸秆中的纤维素(38%)。样品还补充有 CiP(150 μg/gPCS 纤维素)、非离子表面活性剂(2% w/w,在干底物基础上)和处于类似剂量的过氧化物酶和非离子表面活性剂的组合(表 11)。在第 72 和 120 小时取出样品并且如通过 HPLC 所述进行分析。

[0512] 表 10 木质纤维素底物的组成

[0513]

样品 ID	预处理	不溶性固体 分数	纤维素	木聚糖	酸不溶性木质素
芦竹	两段式热水	72.3%	47.4%	14.9%	36.2%
玉米秸 秆	稀酸	68.0%	54.1%	5.2%	29.0%
玉米秸 秆	热水	73.0%	50.9%	12.7%	23.5%
混合木 材	稀酸	78.60%	51.40%	6.5%	34.1%

[0514] 表 11 实验设计：不同木质纤维素材料上的协同作用

[0515]

样品 ID	底物	水解中的总 固体装载%	纤维素分解 酶组合物%	CiP %	表面活性剂%
1	芦竹	15	100		
2	芦竹	15	100	3	
3	芦竹	15	100		2
4	芦竹	15	100	3	2
5	玉米秸秆（稀酸）	20	100		
6	玉米秸秆（稀酸）	20	100	3	
7	玉米秸秆（稀酸）	20	100		2
8	玉米秸秆（稀酸）	20	100	3	2
9	玉米秸秆（热水）	15	100		

[0516]

10	玉米秸秆（热水）	15	100	3	
11	玉米秸秆（热水）	15	100		2
12	玉米秸秆（热水）	15	100	3	2
13	混合木材	10	100		
14	混合木材	10	100	3	
15	混合木材	10	100		2
16	混合木材	10		3	2

[0517] 图 8 示出了水解 120 小时后的结果。所有木质纤维素材料都观察到了协同作用。

[0518] 实例 8

[0519] 不同过氧化物酶和非离子表面活性剂对 GH61 水平的影响

[0520] 在 50℃、pH 5 下，在 20%（w/w，在干重基础上）总固体装载下进行 PCS 的水解。

使用的酶是处于如概述于表 12 中的不同比例的里氏木霉纤维素酶 2 (Tr Cel 2)、棘孢曲霉 β -葡萄糖苷酶 (AaBG)、半纤维素酶 3 (Hemi 3)、金黄色嗜热子囊菌 GH61 多肽 (TaGH61a)、灰盖鬼伞过氧化物酶 (CiP)、大豆过氧化物酶 (Soy P)、王棕过氧化物酶 (RpP)、木质素过氧化物酶 (LiP) 及辣根过氧化物酶 (HrP) 的组合。GH61、纤维素酶和半纤维素酶的总蛋白剂量是 3mg/g PCS 纤维素, 李维本非离子表面活性剂 (1% w/w, 在于 PCS 基础上) 和过氧化物酶与非离子表面活性剂的组合处于类似剂量 (表 12)。在第 72 和 144 小时取出样品并且如通过 HPLC 所述进行分析。

[0521] 表 12 实验设计:

[0522]

编号	Tr cel 2 %	AaBG %	Hemi 3 %	TaGH61a %	CiP	Soy P	王棕 P	LiP	HrP	levapon 李维本 150 %
					基于纤维素酶的%			单位/g 纤维素酶	单位/g 纤维素酶	
1	85	5	5	5						
2	85	5	5	5						1
3	85	5	5	5	3					
4	85	5	5	5	3					1
5	85	5	5	5		0.5				

[0523]

6	85	5	5	5		1				
7	85	5	5	5		0.5				1
8	85	5	5	5		1				1
9	85	5	5	5			2			
10	85	5	5	5			4			
11	85	5	5	5			2			1
12	85	5	5	5			4			1
13	85	5	5	5				0.01		
14	85	5	5	5				0.04		
15	85	5	5	5				0.01		1
16	85	5	5	5				0.04		1
17	85	5	5	5					25	
18	85	5	5	5					100	
19	85	5	5	5					25	1
20	85	5	5	5					100	1
21	85	5	5	5						
22	85	5	5	5						1

[0524] 结果示于图 9 和 10 中。

[0525] 在此描述并且要求的本发明不应限于在此披露的具体方面的范围,因为这些方面旨在作为本发明若干方面的说明。预期任何等效方面都处于本发明的范围内。实际上,除在此所示和描述的那些之外,本发明的不同修改对于本领域普通技术人员而言从前述描述将变得清楚。这类修改也旨在落入所附权利要求书的范围内。在有冲突的情况下,以包括定义的本披露为准。

[0526] 在以下编号的段落中进一步描述本发明：

[0527] 1. 一种用于降解 / 水解预处理的纤维素材料的方法,该方法包括使该预处理的纤维素材料经受：

[0528] - 一种纤维素分解酶组合物；

[0529] - 一种多肽,该多肽具有纤维素分解增强活性；

[0530] - 一种过氧化物酶；以及

[0531] - 一种非离子表面活性剂和 / 或一种阳离子表面活性剂,

[0532] 在适于水解该预处理的木质纤维素材料的条件下。

[0533] 2. 如段落 1 所述的方法,其中在 10 % -50 % TS,例如 15 % -40 % TS,例如 15 % -30 % TS,例如大约 20 % TS 下进行该水解。

[0534] 3. 如段落 1 或 2 所述的方法,其中该水解进行 12-240 小时,例如 24-192 小时,例如 48-144,例如大约 96 小时。

[0535] 4. 如段落 1-3 中任一项所述的方法,其中水解过程中的温度在 30℃ -70℃ 之间,例

如 40℃ -60℃,例如 45℃ -55℃,例如大约 50℃。

[0536] 5. 如段落 1-4 中任一项所述的方法,其中水解过程中的 pH 在 4-7 之间,例如 4.5-6,例如大约 pH 5。

[0537] 6. 如段落 1-5 中任一项所述的方法,其中在水解过程中装载的该纤维素分解酶组合在约 0.1 至约 25mg 之间,例如约 1-10mg,例如约 2 至约 8mg,例如大约 4mg 蛋白 /g 纤维素材料。

[0538] 7. 如段落 1-6 中任一项所述的方法,其中该纤维素分解酶组合包括一种或多种(若干种)选自下组的酶,该组由以下各项组成:内切葡聚糖酶、纤维二糖水解酶(CBH)和 β -葡糖苷酶。

[0539] 8. 如段落 1-7 中任一项所述的方法,其中该纤维素分解酶组合来源于拉克淖金孢子菌、特异腐质霉、嗜热毁丝霉或里氏木霉。

[0540] 9. 如段落 1-8 中任一项所述的方法,其中具有纤维素分解增强活性的该多肽是一种 GH61 多肽,如来源于嗜热子囊菌属的多肽,如金黄色嗜热子囊菌菌株,如在 WO 2005/074656 中描述为 SEQ ID NO:2 或在此的 SEQ ID NO:14 的多肽;或来源于梭孢壳属的多肽,如土生梭孢壳菌株,如在 WO 2005/074647 中描述为 SEQ ID NO:7 和 SEQ ID NO:8 的多肽;或来源于曲霉属的菌株的多肽,如烟曲霉菌株,如在 WO 2010/138754 中描述为 SEQ ID NO:1 和 SEQ ID NO:2 的多肽;或来源于衍生自青霉属的菌株的多肽,如埃默森青霉菌菌株,如在 WO 2011/041397 中披露为 SEQ ID NO:2 或在此的 SEQ ID NO:72 的多肽。

[0541] 10. 如段落 1-9 中任一项所述的方法,其中具有纤维素分解增强活性的该多肽是一种与在此的 SEQ ID NO:14 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的多肽。

[0542] 11. 如段落 1-9 中任一项所述的方法,其中具有纤维素分解增强活性的该多肽是一种与在此的 SEQ ID NO:72 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的多肽。

[0543] 12. 如段落 1-11 中任一项所述的方法,其中该纤维素分解酶组合包括一种 β -葡糖苷酶,优选来源于曲霉属的菌株的 β -葡糖苷酶,如米曲霉,如披露于 WO 02/095014 中的 β -葡糖苷酶或在例如 WO 2008/057637 中披露为 SEQ ID NO:68 或在此的 SEQ ID NO:70 的具有 β -葡糖苷酶活性的融合蛋白;棘孢曲霉,如披露于在此的 SEQ ID NO:66 中的 β -葡糖苷酶,或烟曲霉,如披露于 WO 2005/047499,例如在此的 SEQ ID NO:78 中的 β -葡糖苷酶;或披露于 WO 2012/044915 中的烟曲霉 β -葡糖苷酶变体(例如,F100D、S283G、N456E、F512Y);或青霉属的菌株,如披露于 WO 2007/019442 或在此的 SEQ ID NO:62 中的巴西青霉菌株,或木霉属的菌株,如里氏木霉菌株。

[0544] 13. 如段落 12 所述的方法,其中该 β -葡糖苷酶变体来自曲霉属的菌株,例如烟曲霉菌株,例如烟曲霉 β -葡糖苷酶(在此的 SEQ ID NO:78),该 β -葡糖苷酶包括一个或多个选自下组的取代,该组由以下各项组成:L89M、G91L、F100D、I140V、I186V、S283G、N456E 以及 F512Y。

[0545] 14. 如段落 13 所述的方法,其中该 β -葡糖苷酶变体具有以下取代:

[0546] -F100D+S283G+N456E+F512Y;

[0547] -L89M+G91L+I186V+I140V;

[0548] -I186V+L89M+G91L+I140V+F100D+S283G+N456E+F512Y(在此使用 SEQ ID NO:78 用于编号)。

[0549] 15. 如段落 12-14 中任一项所述的方法,其中该 β -葡萄糖苷酶变体具有在 1 与 10 之间的取代数目,例如在 1 与 8 之间,例如在 1 与 6 之间,例如在 1 与 4 之间,例如 1、2、3、4、5、6、7、8、9 或 10 个取代。

[0550] 16. 如段落 12-15 中任一项所述的方法,其中该 β -葡萄糖苷酶是一种与在此的 SEQ ID NO:78 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的 β -葡萄糖苷酶。

[0551] 17. 如段落 12-16 中任一项所述的方法,其中该 β -葡萄糖苷酶变体是一种与在此的 SEQ ID NO:78 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的 β -葡萄糖苷酶变体。

[0552] 18. 如段落 1-17 中任一项所述的方法,其中该纤维素分解酶组合物包括一种木聚糖酶,优选一种 GH10 木聚糖酶,例如来源于曲霉属的菌株的木聚糖酶,例如来自烟曲霉的菌株,例如在 WO 2006/078256 中披露为 SEQ ID NO:6(Xyl III) 或在此的 SEQ ID NO:75 的木聚糖酶,或来自棘孢曲霉的菌株,例如在 WO 94/21785 中披露为 SEQ ID NO:5(Xyl II) 或在此的 SEQ ID NO:74 的木聚糖酶。

[0553] 19. 如段落 18 所述的方法,其中该木聚糖酶是一种与在此的 SEQ ID NO:74 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的木聚糖酶。

[0554] 20. 如段落 18 所述的方法,其中该木聚糖酶是一种与在此的 SEQ ID NO:75 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的木聚糖酶。

[0555] 21. 如段落 1-20 中任一项所述的方法,其中该纤维素分解酶组合物包括一种 β -木糖苷酶,例如一种来源于曲霉属的菌株的 β -木糖苷酶,例如烟曲霉菌株,例如披露于未决美国临时编号 61/526833 或 WO 2013/028928(实例 16 和 17) 中或在此的 SEQ ID NO:73 的 β -木糖苷酶,或来源于木霉属的菌株,例如里氏木霉菌株,例如 WO 2011/057140 中的 SEQ ID NO:58 的成熟多肽。

[0556] 22. 如段落 21 所述的方法,其中该 β -木糖苷酶是一种与在此的 SEQ ID NO:73 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的 β -木糖苷酶。

[0557] 23. 如段落 1-22 中任一项所述的方法,其中该纤维素分解酶组合物包括一种纤维二糖水解酶 I(CBH I),例如一种来源于以下的纤维二糖水解酶 I,曲霉属的菌株,例如烟曲

霉菌株,例如披露于 WO 2011/057140 中的 SEQ ID NO:6 或在此的 SEQ ID NO:76 的 Cel7a CBH I,或木霉属的菌株,例如里氏木霉菌株。

[0558] 24. 如段落 23 所述的方法,其中该 CBH I 是一种与在此的 SEQ ID NO:76 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的 CBH I。

[0559] 25. 如段落 1-24 中任一项所述的方法,其中该纤维素分解酶组合物包括一种纤维二糖水解酶 II (CBH II),例如一种来源于以下的纤维二糖水解酶 II,曲霉属的菌株,例如烟曲霉菌株,例如在 WO 2011/057140 中示为 SEQ ID NO:18 或在此的 SEQ ID NO:77 的纤维二糖水解酶 II;或木霉属的菌株,例如里氏木霉;或梭孢壳属的菌株,例如土生梭孢壳菌株,例如来自土生梭孢壳的纤维二糖水解酶 II CEL6A。

[0560] 26. 如段落 25 所述的方法,其中该 CBH II 是一种与在此的 SEQ ID NO:77 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的 CBH II。

[0561] 27. 如段落 1-26 中任一项所述的方法,其中该纤维素分解酶组合物是一种里氏木霉纤维素分解酶组合物并且具有纤维素分解增强活性的该多肽是金黄色嗜热子囊菌 GH61A (WO 2005/074656 中的 SEQ ID NO:2 或在此的 SEQ ID NO:14),例如一种与在此的 SEQ ID NO:14 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的多肽。

[0562] 28. 如段落 27 所述的方法,进一步其中存在或添加一种 β -葡糖苷酶,例如在 WO 2008/057637 中示为 SEQ ID NO:74 或 76 或在此的 SEQ ID NO:68 或 70 的米曲霉 β -葡糖苷酶融合蛋白。

[0563] 29. 如段落 28 所述的方法,其中该 β -葡糖苷酶是一种与在此的 SEQ ID NO:68 或 70 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的 β -葡糖苷酶。

[0564] 30. 如段落 1-29 中任一项所述的方法,其中该纤维素分解酶组合物是一种里氏木霉纤维素分解酶组合物并且具有纤维素分解增强活性的该多肽是在 WO 2011/041397 中披露为 SEQ ID NO:2 的埃默森青霉菌 GH61A 多肽,例如一种与在此的 SEQ ID NO:72 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的多肽。

[0565] 31. 如段落 30 所述的方法,进一步其中存在或添加一种 β -葡糖苷酶,例如烟曲霉 β -葡糖苷酶 (WO 2005/047499 的 SEQ ID NO:2 或在此的 SEQ ID NO:76) 或其具有以下取代的变体:F100D、S283G、N456E、F512Y (WO 2012/044915)。

[0566] 32. 如段落 1-31 中任一项所述的方法,其中该纤维素分解酶组合物是一种里氏木霉纤维素分解酶组合物并且其中存在或添加以下组分中的一种或多种:

[0567] (i) 一种烟曲霉纤维二糖水解酶 I；

[0568] (ii) 一种烟曲霉纤维二糖水解酶 II；

[0569] (iii) 一种烟曲霉 β -葡萄糖苷酶或其变体,例如具有以下取代中的一个或多个: F100D、S283G、N456E、F512Y(在此使用 SEQ ID NO:78 用于编号);以及

[0570] (iv) 一种具有纤维素分解增强活性的青霉属 GH61 多肽;或其同系物。

[0571] 33. 如段落 1-32 中任一项所述的方法,其中该纤维素分解酶组合物进一步包括一种或多种(若干种)选自下组的酶,该组由以下各项组成:半纤维素酶、酯酶、蛋白酶以及漆酶。

[0572] 34. 如段落 1-33 中任一项所述的方法,其中该纤维素分解酶组合物进一步包括一种或多种(若干种)选自下组的酶,该组由以下各项组成:木聚糖酶、乙酰木聚糖酯酶、阿魏酸酯酶、阿拉伯呋喃糖苷酶、木糖苷酶、葡萄糖醛酸糖苷酶及其组合。

[0573] 35. 如段落 1-34 中任一项所述的方法,其中该过氧化物酶选自下组,该组包括过氧化物酶或过氧化物分解酶,包括但不限于以下各项:E.C. 1. 11. 1. 1NADH 过氧化物酶;E.C. 1. 11. 1. 2NADPH 过氧化物酶;E.C. 1. 11. 1. 3 脂肪酸过氧化物酶;E.C. 1. 11. 1. 5 细胞色素 c 过氧化物酶;E.C. 1. 11. 1. 5;E.C. 1. 11. 1. 6 过氧化氢酶;E.C. 1. 11. 1. 7 过氧化物酶;E.C. 1. 11. 1. 8 碘化物过氧化物酶;E.C. 1. 11. 1. 9 谷胱甘肽过氧化物酶;E.C. 1. 11. 1. 10 氯化物过氧化物酶;E.C. 1. 11. 1. 11L- 抗坏血酸过氧化物酶;E.C. 1. 11. 1. 12 磷脂过氧化氢谷胱甘肽过氧化物酶;E.C. 1. 11. 1. 13 锰过氧化物酶;E.C. 1. 11. 1. 14 木质素过氧化物酶;E.C. 1. 11. 1. 15 过氧化物氧还蛋白;E.C. 1. 11. 1. 16 通用过氧化物酶;E.C. 1. 11. 1. B2 氯化物过氧化物酶;E.C. 1. 11. 1. B6 碘化物过氧化物酶(含钒);E.C. 1. 11. 1. B7 溴化物过氧化物酶;E.C. 1. 11. 1. B8 碘化物过氧化物酶。

[0574] 36. 如段落 1-35 中任一项所述的方法,其中该过氧化物酶来源于一种微生物,例如真菌生物体,例如酵母或丝状真菌,或细菌;或植物。

[0575] 37. 如段落 1-36 中任一项所述的方法,其中该过氧化物酶来源于鬼伞属的菌株,例如灰盖鬼伞菌株,例如示于在此的 SEQ ID NO:71 中的过氧化物酶(即,CiP),或一种与在此的 SEQ ID NO:71 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的过氧化物酶。

[0576] 38. 如段落 1-37 中任一项所述的方法,其中该非离子表面活性剂是烷基或芳基:甘油醚、乙二醇醚、乙醇酰胺、磺酰胺、醇、酰胺、醇乙氧基化物、甘油酯、乙二醇酯、甘油酯和乙二醇酯的乙氧基化物、糖基烷基多糖苷、聚氧乙烯化脂肪酸、烷醇胺缩合物、烷醇酰胺、叔炔属乙二醇、聚氧乙烯化硫醇、羧酸酯、以及聚氧乙烯化聚氧丙烯乙二醇(例如 E0/P0 嵌段共聚物(E0 是环氧乙烷,P0 是环氧丙烷))、E0 聚合物和共聚物、多胺、和聚乙烯吡咯烷酮。

[0577] 39. 如段落 1-38 中任一项所述的方法,其中该非离子表面活性剂是一种直链伯或仲或具有化学式 $\text{RO}(\text{CH}_2\text{CH}_2\text{O})_n\text{H}$ 的支链醇乙氧基化物,其中 R 是烃链长度并且 n 是环氧乙烷的摩尔平均数,例如其中 R 是范围从 C9 至 C16 的直链伯或支链仲烃链长度并且 n 范围从 6 至 13,例如醇乙氧基化物,其中 R 是直链 C9-C11 烃链长度,并且 n 是 6。

[0578] 40. 如段落 1-39 中任一项所述的方法,其中该阳离子表面活性剂是一种伯、仲或叔胺,例如二盐酸奥替尼啶;烷基三甲基铵盐,例如鲸蜡基三甲基溴化铵(CTAB)(又名十六

烷基三甲基溴化铵)、十六烷基三甲基氯化铵(CTAC)、氯化十六烷基吡啶(CPC)、苯扎氯铵(BAC)、苜蓿素氯铵(BZT)、5-溴-5-硝基-1,3-二噁烷、二甲基双十八烷基氯化铵、双十八烷基二甲基溴化铵(DODAB)。

[0579] 41. 如段落 1-40 中任一项所述的方法,其中该水解的预处理的纤维素材料是一种糖。

[0580] 42. 如段落 1-41 中任一项所述的方法,其中该预处理的纤维素材料是农业废弃物、草本材料(包括能源作物)、城市固体废物、纸浆和造纸厂废弃物、废纸、或木材(包括林业废弃物)、或芦竹、甘蔗渣、竹子、玉米芯、玉米纤维、玉米秸秆、芒属、橙皮、稻秸、柳枝稷或麦秸。

[0581] 43. 如段落 1-42 中任一项所述的方法,其中这些糖由一种发酵微生物发酵为一种发酵产品。

[0582] 44. 如段落 1-43 中任一项所述的方法,进一步包括回收该水解的纤维素材料,例如糖或发酵产品。

[0583] 45. 如段落 44 所述的方法,其中该糖选自下组,该组由以下各项组成:葡萄糖、木糖、甘露糖、半乳糖以及阿拉伯糖。

[0584] 46. 如段落 43-45 中任一项所述的方法,其中该发酵产品是醇(例如乙醇)、有机酸、酮、氨基酸或气体。

[0585] 47. 如段落 1-46 中任一项所述的方法,其中具有过氧化物酶活性的多肽的 K_m 处于优选 0.0001 至 50mM、更优选 0.001 至 10mM、甚至更优选 0.005 至 1mM、并且最优选 0.01 至 0.1mM 的范围内。

[0586] 48. 如段落 1-47 中任一项所述的方法,其中该预处理的纤维素材料通过化学预处理、物理预处理、或化学预处理和物理预处理进行预处理。

[0587] 49. 如段落 1-48 中任一项所述的方法,其中该预处理是碱性预处理,例如铵预处理,例如温和的铵预处理。

[0588] 50. 如段落 1-49 中任一项所述的方法,其中热化学预处理该纤维素材料。

[0589] 51. 如段落 1-50 中任一项所述的方法,其中预处理该纤维素材料包括用一种酸预处理,例如稀酸预处理。

[0590] 52. 如段落 1-51 中任一项所述的方法,其中通过在高温、高压下用一种酸(例如稀酸)预处理该纤维素材料来制备该预处理的纤维素材料。

[0591] 53. 如段落 51 或 52 所述的方法,其中使用乙酸进行酸预处理。

[0592] 54. 如段落 1-53 中任一项所述的方法,其中已经通过使用有机溶剂(organosolv)预处理来预处理纤维素材料制备了该预处理的纤维素材料,例如丙酮溶剂(Acetosolv)和丙酮池(Acetocell)工艺。

[0593] 55. 一种用于生产发酵产品的工艺,该工艺包括

[0594] (a) 水解/降解如在段落 1-54 任一项中定义的该预处理的纤维素材料;

[0595] (b) 用一种或多种(若干种)发酵微生物对该材料进行发酵,以产生该发酵产品;并且

[0596] (c) 任选地从发酵中回收该发酵产品。

[0597] 56. 如段落 55 所述的工艺,其中水解步骤(a)和水解步骤(b)顺序或同时;作为

分开的水解和发酵 (SHF) ;同时糖化和发酵 (SSF) ;同时糖化和共发酵 (SSCF) ;混合水解和发酵 (HHF) ;分开的水解和共发 (SHCF) ;混合水解和共发酵 (HHCF) ;或直接微生物转化 (DMC),有时也称为整合生物加工 (CBP) 而进行。

[0598] 57. 如段落 55 或 56 所述的工艺,其中使用一种酵母或细菌进行发酵。

[0599] 58. 如段落 55-57 中任一项所述的工艺,其中该发酵微生物能够将己糖和 / 或戊糖发酵为一种所希望的发酵产品。

[0600] 59. 如段落 58 所述的工艺,其中该发酵微生物是一种酵母,例如酵母属的菌株,例如酿酒酵母菌株。

[0601] 60. 如段落 55-59 中任一项所述的工艺,其中在约 26℃ 至约 60℃ 之间的温度,例如约 32℃ 或 50℃,和约 pH 3 至约 pH 8,例如 pH 4-5、6 或 7 下进行该发酵。

[0602] 61. 如段落 55-60 中任一项所述的工艺,其中当该发酵微生物是酵母,例如酵母属的菌株,特别是酿酒酵母菌株时,在从 20℃ -40℃ 的温度,例如 26℃ -34℃,优选大约 32℃ 下进行该发酵,尤其是当该发酵产品是乙醇时。

[0603] 62. 如段落 55-61 中任一项所述的工艺,其中在 pH 3-7,例如 pH 4-6 下进行该发酵。

[0604] 63. 如段落 55-62 中任一项所述的工艺,其中该发酵进行约 12 至约 96 小时,例如典型地是 24-60 小时。

[0605] 64. 如段落 55-63 中任一项所述的工艺,其中该发酵产品是乙醇。

[0606] 65. 一种组合物,包括或其组成为:

[0607] i) 一种多肽,该多肽具有纤维素分解增强活性;

[0608] ii) 一种过氧化物酶;

[0609] iii) 一种非离子表面活性剂和 / 或一种阳离子表面活性剂。

[0610] 66. 如段落 65 所述的组合物,其中具有纤维素分解增强活性的该多肽是一种 GH61 多肽,如来源于嗜热子囊菌属的多肽,如金黄色嗜热子囊菌菌株,如在 WO 2005/074656 中描述为 SEQ ID NO:2 或在此的 SEQ ID NO:14 的多肽;或来源于梭孢壳属的多肽,如土生梭孢壳菌株,如在 WO 2005/074647 中描述为 SEQ ID NO:8 或在此的 SEQ ID NO:8 的多肽;或来源于曲霉属的菌株的多肽,如烟曲霉菌株,如在 WO 2010/138754 中描述为 SEQ ID NO:1 和 SEQ ID NO:2 的多肽;或来源于衍生自青霉属的菌株的多肽,如埃默森青霉菌菌株,如在 WO 2011/041397 中披露为 SEQ ID NO:2 或在此的 SEQ ID NO:72 的多肽。

[0611] 67. 如段落 65 或 66 所述的组合物,其中具有纤维素分解增强活性的该多肽与在此的 SEQ ID NO:14 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性。

[0612] 68. 如段落 65-67 中任一项所述的组合物,其中具有纤维素分解增强活性的该多肽与在此的 SEQ ID NO:72 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性。

[0613] 69. 如段落 65-68 中任一项所述的组合物,其中该过氧化物酶选自下组,该组包括过氧化物酶或过氧化物分解酶,包括但不限于以下各项:E. C. 1. 11. 1. 1NADH 过氧化物酶;

E. C. 1. 11. 1. 2NADPH 过氧化物酶 ;E. C. 1. 11. 1. 3 脂肪酸过氧化物酶 ;E. C. 1. 11. 1. 5 细胞色素 c 过氧化物酶 ;E. C. 1. 11. 1. 5 ;E. C. 1. 11. 1. 6 过氧化氢酶 ;E. C. 1. 11. 1. 7 过氧化物酶 ;E. C. 1. 11. 1. 8 碘化物过氧化物酶 ;E. C. 1. 11. 1. 9 谷胱甘肽过氧化物酶 ;E. C. 1. 11. 1. 10 氯化物过氧化物酶 ;E. C. 1. 11. 1. 11L- 抗坏血酸过氧化物酶 ;E. C. 1. 11. 1. 12 磷脂过氧化氢谷胱甘肽过氧化物酶 ;E. C. 1. 11. 1. 13 锰过氧化物酶 ;E. C. 1. 11. 1. 14 木质素过氧化物酶 ;E. C. 1. 11. 1. 15 过氧化物氧还蛋白 ;E. C. 1. 11. 1. 16 通用过氧化物酶 ;E. C. 1. 11. 1. B2 氯化物过氧化物酶 ;E. C. 1. 11. 1. B6 碘化物过氧化物酶 (含钒) ;E. C. 1. 11. 1. B7 溴化物过氧化物酶 ;E. C. 1. 11. 1. B8 碘化物过氧化物酶。

[0614] 70. 如段落 65-69 中任一项所述的组合物, 其中该过氧化物酶来源于一种微生物, 例如真菌生物体, 例如酵母或丝状真菌, 或细菌 ; 或植物。

[0615] 71. 如段落 65-70 中任一项所述的组合物, 其中该过氧化物酶来源于鬼伞属的菌株, 例如灰盖鬼伞菌株。

[0616] 72. 如段落 65-71 中任一项所述的组合物, 其中该过氧化物酶是示于在此的 SEQ ID NO:71 中的过氧化物酶或一种与在此的 SEQ ID NO:71 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的过氧化物酶。

[0617] 73. 如段落 65-72 中任一项所述的组合物, 其中该非离子表面活性剂是烷基或芳基 : 甘油醚、乙二醇醚、乙醇酰胺、磺酰胺、醇、酰胺、醇乙氧基化物、甘油酯、乙二醇酯、甘油酯和乙二醇酯的乙氧基化物、糖基烷基多糖苷、聚氧乙烯化脂肪酸、烷醇胺缩合物、烷醇酰胺、叔炔属乙二醇、聚氧乙烯化硫醇、羧酸酯、以及聚氧乙烯化聚氧丙烯乙二醇 (例如 EO/PO 嵌段共聚物 (EO 是环氧乙烷, PO 是环氧丙烷))、EO 聚合物和共聚物、多胺、和聚乙烯吡咯烷酮。

[0618] 74. 如段落 65-73 中任一项所述的组合物, 其中该非离子表面活性剂是一种直链伯或仲或具有化学式 $\text{RO}(\text{CH}_2\text{CH}_2\text{O})_n\text{H}$ 的支链醇乙氧基化物, 其中 R 是烃链长度并且 n 是环氧乙烷的摩尔平均数, 例如其中 R 是范围从 C9 至 C16 的直链伯或支链仲烃链长度并且 n 范围从 6 至 13, 例如醇乙氧基化物, 其中 R 是直链 C9-C11 烃链长度, 并且 n 是 6。

[0619] 75. 如段落 65-74 中任一项所述的组合物, 其中该阳离子表面活性剂是一种伯、仲或叔胺, 例如二盐酸奥替尼啶 ; 烷基三甲基铵盐, 例如鲸蜡基三甲基溴化铵 (CTAB) (又名十六烷基三甲基溴化铵)、十六烷基三甲基氯化铵 (CTAC)、氯化十六烷基吡啶 (CPC)、苯扎氯铵 (BAC)、苄索氯铵 (BZT)、5- 溴 -5- 硝基 -1, 3- 二噁烷、二甲基双十八烷基氯化铵、双十八烷基二甲基溴化铵 (DODAB)。

[0620] 76. 如段落 65-75 中任一项所述的组合物, 进一步包括一种纤维素分解酶组合物。

[0621] 77. 如段落 76 所述的组合物, 包括一种 β - 葡糖苷酶。

[0622] 78. 如段落 65-77 中任一项所述的组合物, 其中该纤维素分解酶组合物包括一种 β - 葡糖苷酶, 优选一种来源于以下的 β - 葡糖苷酶, 曲霉属的菌株, 如米曲霉, 如披露于 WO 02/095014 中的 β - 葡糖苷酶或披露于 WO 2008/057637 或在此的 SEQ ID NO:68 或 70 中、具有 β - 葡糖苷酶活性的融合蛋白, 或烟曲霉, 如在 WO 2005/047499 中披露为 SEQ ID NO:2 或在此的 SEQ ID NO:78 的 β - 葡糖苷酶, 或披露于 WO 2012/044915 中的烟曲霉 β - 葡糖

苷酶变体(例如,F100D、S283G、N456E、F512Y);或青霉属的菌株,如披露于 WO 2007/019442 中、示于在此的 SEQ ID NO:62 中的巴西青霉菌株,或木霉属的菌株,如里氏木霉菌株。

[0623] 79. 如段落 78 所述的组合物,其中该 β -葡糖苷酶变体来自曲霉属的菌株,例如烟曲霉菌株,例如烟曲霉 β -葡糖苷酶(在此的 SEQ ID NO:78),该 β -葡糖苷酶包括一个或多个选自下组的取代,该组由以下各项组成:L89M、G91L、F100D、I140V、I186V、S283G、N456E 以及 F512Y。

[0624] 80. 如段落 77-79 中任一项所述的组合物,其中该 β -葡糖苷酶变体具有以下取代:

[0625] -F100D+S283G+N456E+F512Y;

[0626] -L89M+G91L+I186V+I140V;

[0627] -I186V+L89M+G91L+I140V+F100D+S283G+N456E+F512Y(在此使用 SEQ ID NO:78 用于编号)。

[0628] 81. 如段落 77-80 中任一项所述的组合物,其中该 β -葡糖苷酶变体具有在 1 与 10 之间的取代数目,例如在 1 与 8 之间,例如在 1 与 6 之间,例如在 1 与 4 之间,例如 1、2、3、4、5、6、7、8、9 或 10 个取代。

[0629] 82. 如段落 77-81 中任一项所述的组合物,其中该 β -葡糖苷酶是一种与在此的 SEQ ID NO:78 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的 β -葡糖苷酶。

[0630] 83. 如段落 77-82 中任一项所述的组合物,其中该 β -葡糖苷酶变体是一种与在此的 SEQ ID NO:78 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性的 β -葡糖苷酶变体。

[0631] 84. 如段落 77-83 中任一项所述的组合物,其中该纤维素分解酶组合物来源于里氏木霉、特异腐质霉、或拉克淖金孢子菌、或嗜热毁丝霉。

[0632] 85. 如段落 65-84 中任一项所述的组合物,包括:

[0633] i) 一种具有纤维素分解增强活性的多肽,该多肽与在此的 SEQ ID NO:14 或在此的 SEQ ID NO:72 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性;

[0634] ii) 一种过氧化物酶,该过氧化物酶与在此的 SEQ ID NO:71 具有至少 60%、优选至少 65%、更优选至少 70%、更优选至少 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%、并且甚至最优选至少 96%、至少 97%、至少 98%、或至少 99% 序列一致性;

[0635] iii) 一种非离子表面活性剂和/或一种阳离子表面活性剂。

[0636] 86. 如段落 85 所述的组合物,其中该非离子表面活性剂是烷基或芳基:甘油醚、乙二醇醚、乙醇酰胺、磺酰胺、醇、酰胺、醇乙氧基化物、甘油酯、乙二醇酯、甘油酯和乙二醇酯的乙氧基化物、糖基烷基多糖苷、聚氧乙烯化脂肪酸、烷醇胺缩合物、烷醇酰胺、叔炔属乙二醇、聚氧乙烯化硫醇、羧酸酯、以及聚氧乙烯化聚氧丙烯乙二醇(例如 EO/PO 嵌段共聚物(EO

是环氧乙烷, PO 是环氧丙烷))、EO 聚合物和共聚物、多胺、和聚乙烯吡咯烷酮。

[0637] 87. 如段落 85 或 86 所述的组合物, 其中该非离子表面活性剂是一种直链伯或仲或具有化学式 $\text{RO}(\text{CH}_2\text{CH}_2\text{O})_n\text{H}$ 的支链醇乙氧基化物, 其中 R 是烃链长度并且 n 是环氧乙烷的摩尔平均数, 例如其中 R 是范围从 C9 至 C16 的直链伯或支链仲烃链长度并且 n 范围从 6 至 13, 例如醇乙氧基化物, 其中 R 是直链 C9-C11 烃链长度, 并且 n 是 6。

[0638] 88. 如段落 85-87 中任一项所述的组合物, 其中该阳离子表面活性剂是一种伯、仲或叔胺, 例如二盐酸奥替尼啶; 烷基三甲基铵盐, 例如鲸蜡基三甲基溴化铵 (CTAB) (又名十六烷基三甲基溴化铵)、十六烷基三甲基氯化铵 (CTAC)、氯化十六烷基吡啶 (CPC)、苯扎氯铵 (BAC)、苄索氯铵 (BZT)、5- 溴 -5- 硝基 -1, 3- 二噁烷、二甲基双十八烷基氯化铵、双十八烷基二甲基溴化铵 (DODAB)。

[0639] 89. 如段落 85-88 中任一项所述的组合物, 其中该非离子表面活性剂选自以下各项的组: 壬基酚乙氧基化物; $\text{C}_{14}\text{H}_{22}\text{O}(\text{C}_2\text{H}_4\text{O})_n$; C_{13} -醇聚乙二醇醚 (10EO); EO, PO 共聚物; 烷基聚乙二醇醚; $\text{RO}(\text{EO})_5\text{H}$; $\text{HOCH}_2(\text{EO})_n\text{CH}_2\text{OH}$; 以及 $\text{HOCH}_2(\text{EO})_n\text{CH}_2\text{OH}$ 。

[0640] 90. 如段落 85-89 中任一项所述的组合物, 其中该阳离子表面活性剂选自以下各项的组: $\text{C}_{21}\text{H}_{38}\text{NCl}$ 和 $\text{CH}_3(\text{CH}_2)_{15}\text{N}(\text{CH}_3)_3\text{Br}$ 。

[0001]

序列表

<110> Chen, Ye Liu, Jiyin Green, Terry Stevens, Mark V Xu, Hui V		
<120> 用于增加纤维素材料的酶水解的方法		
<130> 12436-W0-PCT		
<160> 79		
<170> PatentIn 3.5 版		
<210> 1		
<211> 1846		
<212> DNA		
<213> 土生梭孢壳(Thielavia terrestris)		
<400> 1		
aattgaagga gggagtggcg gagtggccac caagtcaggc ggctgtcaac taaccaagga	60	
tggaacagt tcgctcgcc ttgccgagg gcagcggtcc ctgatggga egaaccatgg	120	
gaciggggc agcigtgta taaaagtta aatcgatgat ctctcagatg gcgctgtgg	180	
ggtgtttcgc gttttccat cctcgcaacc tggatatcca ctagtccagc gttcggcacc	240	
atgaagfcgt tcaccattgc cgccttggca gccctatggg eccaggaggc cgccgcccac	300	
gcgaccttc aggacctctg gattgatgga gtcgactacg gctcgcaatg tgtcgcctc	360	
ccggcgfcca actccccgt caccaatgtt gcgtcgacg atatccgatg caatgtcggc	420	
acctcgagge ccaccgtcaa gtgccggtc aaggccggct ccaeggtcac gatcgagatg	480	
caccagggtc gcagcctct ctgcgtaggc ccccagcta ctatatggca ctaacacgac	540	
ctccagcaac ctggcgaccg gtcttgcgc aacgaggeta tggcgggcga ccactacggc	600	
ccglaalgg tglacatgc caagtcgat gacgcggiga cagccgacgg tlcactgggc	660	
tggttcaagg tgttcagga cagctgggc aagaacctgt cgggttcgac gggcgacgac	720	
gactactggg gcaccaagga cctcaactcg tcttcggca agatgaacgt caagatcccc	780	
gaagacateg agccggcgca ctacctgtc cgcgcgagg ttatcgctgt gcacgtggcc	840	
gccagctcgg gcggcgcgca gttctacatg tctgtctacc agctgacctg gacgggtcc	900	
ggcagcgcca cccctcgac cgtgaatttc ccggcgccct actcggccag cgaccgggc	960	
atctgatca acatccagc gcccatgtcg acctacgtcg tcccggccc gacctgtac	1020	
gcggcgggcf cgaccaagtc ggcgtgcagc tctgtctcgg gctgcgaggc gacctgcacg	1080	
gttggttcgg gccccagcgc gacctgaag cagcccaact ccaccggac cgcgacctcc	1140	
gccctggcg gcggcggtc cggctgcacg gcggccaagt accagcagtg cggcggcaac	1200	
ggctacactg ggtgcaccac ctgcctgta agttccctcg tgatatgag cggaacaccg	1260	
tctggactgt ttgttaact cgcgtcgtag tccgggtcta cctgcagcgc cgtctcgccf	1320	
ccgtactact cgcagtccct ctaagccggg agcgttctgt cagcgggtg ctgtgaagga	1380	
gtccatgtc cccatgcgc catggccgga gtaccgggt gagegcccac ttctgtata	1440	
tagttgagtt ttcccaatca tgaatacata tgcactgca tggactgttg cgtcgtcagt	1500	
ctacatctt tgcctcactg aactgtgaga ccccaigca tccggacct tcgacggig	1560	

[0002]

ctcgctctac catctcgggt gatgggtctg ggcttgagag tcactggcac gtcctcggcg 1620
 gtaatgaaat gtggaggaaa gtgtgagctg tetgacgcac tcggcgctga tgagacgttg 1680
 agcgcggccc acactgggtg tctgtaagcc agcacacaaa agaatactcc aggatggccc 1740
 atagcggcaa atatacagta tcagggatgc aaaaagtga aaagtaaggg gctcaatcgg 1800
 ggategaacc cgagacctg cacatgaatt attcaagtc aggggt 1846

<210> 2
 <211> 326
 <212> PRT
 <213> 土生梭孢壳(*Thielavia terrestris*)

<400> 2

Met Lys Ser Phe Thr Ile Ala Ala Leu Ala Ala Leu Trp Ala Gln Glu
 1 5 10 15
 Ala Ala Ala His Ala Thr Phe Gln Asp Leu Trp Ile Asp Gly Val Asp
 20 25 30
 Tyr Gly Ser Gln Cys Val Arg Leu Pro Ala Ser Asn Ser Pro Val Thr
 35 40 45
 Asn Val Ala Ser Asp Asp Ile Arg Cys Asn Val Gly Thr Ser Arg Pro
 50 55 60
 Thr Val Lys Cys Pro Val Lys Ala Gly Ser Thr Val Thr Ile Glu Met
 65 70 75 80
 His Gln Gln Pro Gly Asp Arg Ser Cys Ala Asn Glu Ala Ile Gly Gly
 85 90 95
 Asp His Tyr Gly Pro Val Met Val Tyr Met Ser Lys Val Asp Asp Ala
 100 105 110
 Val Thr Ala Asp Gly Ser Ser Gly Trp Phe Lys Val Phe Gln Asp Ser
 115 120 125
 Trp Ala Lys Asn Pro Ser Gly Ser Thr Gly Asp Asp Asp Tyr Trp Gly
 130 135 140
 Thr Lys Asp Leu Asn Ser Cys Cys Gly Lys Met Asn Val Lys Ile Pro
 145 150 155 160
 Glu Asp Ile Glu Pro Gly Asp Tyr Leu Leu Arg Ala Glu Val Ile Ala
 165 170 175
 Leu His Val Ala Ala Ser Ser Gly Gly Ala Gln Phe Tyr Met Ser Cys
 180 185 190
 Tyr Gln Leu Thr Val Thr Gly Ser Gly Ser Ala Thr Pro Ser Thr Val
 195 200 205
 Asn Phe Pro Gly Ala Tyr Ser Ala Ser Asp Pro Gly Ile Leu Ile Asn
 210 215 220
 Ile His Ala Pro Met Ser Thr Tyr Val Val Pro Gly Pro Thr Val Tyr
 225 230 235 240
 Ala Gly Gly Ser Thr Lys Ser Ala Gly Ser Ser Cys Ser Gly Cys Glu
 245 250 255
 Ala Thr Cys Thr Val Gly Ser Gly Pro Ser Ala Thr Leu Thr Gln Pro
 260 265 270
 Thr Ser Thr Ala Thr Ala Thr Ser Ala Pro Gly Gly Gly Gly Ser Gly
 275 280 285

[0003]

Cys Thr Ala Ala Lys Tyr Gln Gln Cys Gly Gly Thr Gly Tyr Thr Gly
290 295 300

Cys Thr Thr Cys Ala Ser Gly Ser Thr Cys Ser Ala Val Ser Pro Pro
305 310 315 320

Tyr Tyr Ser Gln Cys Leu
325

<210> 3
<211> 880
<212> DNA
<213> 土生梭孢壳(*Thielavia terrestris*)

<400> 3
accccgat cactgccct aggaaccagc acacctcgtt ccaatcatgc ggttcgacgc 60
ectctccgcc ctctctcttg cgcgccttgt ggcctggccac ggcgccttga ccagctacat 120
catcgccggc aaaacctatc ccggtacga ggctctctcg ctgcctcga gcccgccgac 180
gatccagtac cagtggcccg actacaacc gacctgagc gtgaccgacc cgaagatgcg 240
ctgcaacggc ggcacctcgg cagagctcag cgcgccctc caggccggcg agaactgac 300
ggcgtcttgg aagcagtgga ccaccagca aggcgccgtc atggtcttga tgttcaagt 360
ccccggcgac ttctctctgt gccacggcga cggcaagggc tggttcaaga tcgaccagct 420
gggcctgttg ggcaacaacc tcaactcga caactggggc accgcgatcg tctacaagac 480
cctccagttg agcaaccga tcccaagaa cctcgcgcgc ggcaactacc tcatcgcga 540
cgagctcttc gccctgcacc aggccaacac gccgcagttc tacgccgagt gcgccagct 600
ggtctgtctc ggcagcggct ccgccctgcc cccgtccgac tacctctaca gcaccccg 660
ctacgcgcc cagaacgacc ccggcatcac cgtgagtggg ctctcgttcc ggcgcgagct 720
ctgtggaaf ctgtctgacg atggctagg ttgacatcta caacggcggg ctacctctc 780
acacccgcc cggcgcccc gctctgtctg gcttcgagtt ttagcgcat tgagtcgggg 840
gctacgaggg gaaggcatct gttecatga gcgtgggtac 880

<210> 4
<211> 239
<212> PRT
<213> 土生梭孢壳(*Thielavia terrestris*)

<400> 4
Met Arg Phe Asp Ala Leu Ser Ala Leu Ala Leu Ala Pro Leu Val Ala
1 5 10 15
Gly His Gly Ala Val Thr Ser Tyr Ile Ile Gly Gly Lys Thr Tyr Pro
20 25 30
Gly Tyr Glu Gly Phe Ser Pro Ala Ser Ser Pro Pro Thr Ile Gln Tyr
35 40 45
Gln Trp Pro Asp Tyr Asn Pro Thr Leu Ser Val Thr Asp Pro Lys Met
50 55 60
Arg Cys Asn Gly Gly Thr Ser Ala Glu Leu Ser Ala Pro Val Gln Ala
65 70 75 80
Gly Glu Asn Val Thr Ala Val Trp Lys Gln Trp Thr His Gln Gln Gly
85 90 95
Pro Val Met Val Trp Met Phe Lys Cys Pro Gly Asp Phe Ser Ser Ser

[0004]

100	105	110
His Gly Asp Gly Lys Gly Trp Phe Lys Ile Asp Gln Leu Gly Leu Trp 115	120	125
Gly Asn Asn Leu Asn Ser Asn Asn Trp Gly Thr Ala Ile Val Tyr Lys 130	135	140
Thr Leu Gln Trp Ser Asn Pro Ile Pro Lys Asn Leu Ala Pro Gly Asn 145	150	155
Tyr Leu Ile Arg His Glu Leu Leu Ala Leu His Gln Ala Asn Thr Pro 165	170	175
Gln Phe Tyr Ala Glu Cys Ala Gln Leu Val Val Ser Gly Ser Gly Ser 180	185	190
Ala Leu Pro Pro Ser Asp Tyr Leu Tyr Ser Ile Pro Val Tyr Ala Pro 195	200	205
Gln Asn Asp Pro Gly Ile Thr Val Asp Ile Tyr Asn Gly Gly Leu Thr 210	215	220
Ser Tyr Thr Pro Pro Gly Gly Pro Val Trp Ser Gly Phe Glu Phe 225	230	235
<210> 5 <211> 1000 <212> DNA <213> 土生梭孢壳(Thielavia terrestris)		
<400> 5 ctctctgttcc tgggcacacg cttgttgcct gaactattgg tagagtgggt ctattgctag 60 agttggccat gttttcaca tcagtcctcg gctcggctgc cctgcttgc agcggcgtg 120 cggcacacgg cggcgtgac agctacatca tcgccggcna gaattaccgg ggggtggtag 180 ctgattattg agggcgctt caaggttcat accggtgtgc atggctgaca accggtgtgc 240 agataccaag gctttttcc tgcgaactcg ccgaacgtca tccaatggca atggcatgac 300 tacaaccccg tcttgcgtg cagcgactcg aagcttcgct gcaacggcgg cagtcggcc 360 accctgaacg ccacggccgc accggcgac accatcacgg ccatctgggc gcagtggacg 420 cacagccagg gcccatactt gggtggatg tacaagtgc cgggtctctt cagctctctg 480 gacggctccg gcgctggctg gttcaagatc gacgaggcgg gcttcacgg cgacggcgtc 540 aaggctttcc tcgacaccga gaacccgtcc ggctgggaca tggccaagct cgtcggcggc 600 aacaagcagt ggagcagcaa ggtccccgag ggcctcgcgc ccggcaacta cctcgtccgc 660 cagcagttga tcgcctgca ccaggccaac aacccgcagt tctacccgga gtgcgcccag 720 gtctcatca ccggtccgg caccgcgcag ccgatgcct catacaagga ggctatcccc 780 ggctactgca accagaatga cccgaacatc aaggtgagat ccaggcgtaa tgcagtctac 840 tgcgtgaaag aaagtggctc aagctaaacc gcgtccagg tgcccatcaa cgaccactcc 900 atccctcaga cctacaagat tccggccct cccgtcttca agggcacccg cagcaagaag 960 gccegggaet tcaccgctg aagtgttga atcgatggag 1000		
<210> 6 <211> 258 <212> PRT <213> 土生梭孢壳(Thielavia terrestris)		

[0005]

<400> 6

Met Leu Leu Thr Ser Val Leu Gly Ser Ala Ala Leu Leu Ala Ser Gly
 1 5 10 15
 Ala Ala Ala His Gly Ala Val Thr Ser Tyr Ile Ile Ala Gly Lys Asn
 20 25 30
 Tyr Pro Gly Tyr Gln Gly Phe Ser Pro Ala Asn Ser Pro Asn Val Ile
 35 40 45
 Gln Trp Gln Trp His Asp Tyr Asn Pro Val Leu Ser Cys Ser Asp Ser
 50 55 60
 Lys Leu Arg Cys Asn Gly Gly Thr Ser Ala Thr Leu Asn Ala Thr Ala
 65 70 75 80
 Ala Pro Gly Asp Thr Ile Thr Ala Ile Trp Ala Gln Trp Thr His Ser
 85 90 95
 Gln Gly Pro Ile Leu Val Trp Met Tyr Lys Cys Pro Gly Ser Phe Ser
 100 105 110
 Ser Cys Asp Gly Ser Gly Ala Gly Trp Phe Lys Ile Asp Glu Ala Gly
 115 120 125
 Phe His Gly Asp Gly Val Lys Val Phe Leu Asp Thr Glu Asn Pro Ser
 130 135 140
 Gly Trp Asp Ile Ala Lys Leu Val Gly Gly Asn Lys Gln Trp Ser Ser
 145 150 155 160
 Lys Val Pro Glu Gly Leu Ala Pro Gly Asn Tyr Leu Val Arg His Glu
 165 170 175
 Leu Ile Ala Leu His Gln Ala Asn Asn Pro Gln Phe Tyr Pro Glu Cys
 180 185 190
 Ala Gln Val Val Ile Thr Gly Ser Gly Thr Ala Gln Pro Asp Ala Ser
 195 200 205
 Tyr Lys Ala Ala Ile Pro Gly Tyr Cys Asn Gln Asn Asp Pro Asn Ile
 210 215 220
 Lys Val Pro Ile Asn Asp His Ser Ile Pro Gln Thr Tyr Lys Ile Pro
 225 230 235 240
 Gly Pro Pro Val Phe Lys Gly Thr Ala Ser Lys Lys Ala Arg Asp Phe
 245 250 255
 Thr Ala

<210> 7

<211> 681

<212> DNA

<213> 土生梭孢壳(Thielavia terrestris)

<400> 7

atgctcgcaa acggtgccat cgtcttcctg gccgccgcc tcggcgctcag tggccaactac 60
 acctggccac gggttaacga cggcgccgac tggcaacagg tccgtaaggc ggacaactgg 120
 caggacaacg gctacgtcgg ggaatgtcac tcgccacaga tccgctgttt ccaggcgacc 180
 ccgtccccgg ccccatccgt cctcaacacc acggccgget cgaecgtgac ctactgggcc 240
 aaeccegcag tctaccaccc cgggectgtg cagttttaca tggcccgcgt gcccgatggc 300
 gaggacatca actcgtggaa cggcgacggc gccgtgiggt tcaaggtgta cgaggaccat 360
 cctacettig gcgctcaget cacatggccc agcacgggca agagctcgtt cgcggttccc 420

[0006]

```

atccccccgt gcatcaagtc cggctactac ctctccggg cggagcaaat cggcctgcac      480
gtcggcccaga gcgtaggcgg agcgcagttc tacatctcat gcgcccagct cagcgtcacc      540
ggcggcggca gcaccgagcc gcgaacaag gtggccctcc cggcgctta cagtgcgacg      600
gacccgggca ttctgatcaa catctactac cctgttccca cgtccctacca gaaccccggc      660
ccggccgctct tcagctgctg a                                             681

```

<210> 8
 <211> 226
 <212> PRT
 <213> 土生梭孢壳(*Thielavia terrestris*)

<400> 8

```

Met Leu Ala Asn Gly Ala Ile Val Phe Leu Ala Ala Ala Leu Gly Val
1              5              10              15

Ser Gly His Tyr Thr Trp Pro Arg Val Asn Asp Gly Ala Asp Trp Gln
20            25            30

Gln Val Arg Lys Ala Asp Asn Trp Gln Asp Asn Gly Tyr Val Gly Asp
35            40            45

Val Thr Ser Pro Gln Ile Arg Cys Phe Gln Ala Thr Pro Ser Pro Ala
50            55            60

Pro Ser Val Leu Asn Thr Thr Ala Gly Ser Thr Val Thr Tyr Trp Ala
65            70            75            80

Asn Pro Asp Val Tyr His Pro Gly Pro Val Gln Phe Tyr Met Ala Arg
85            90            95

Val Pro Asp Gly Glu Asp Ile Asn Ser Trp Asn Gly Asp Gly Ala Val
100           105           110

Trp Phe Lys Val Tyr Glu Asp His Pro Thr Phe Gly Ala Gln Leu Thr
115           120           125

Trp Pro Ser Thr Gly Lys Ser Ser Phe Ala Val Pro Ile Pro Pro Cys
130           135           140

Ile Lys Ser Gly Tyr Tyr Leu Leu Arg Ala Glu Gln Ile Gly Leu His
145           150           155           160

Val Ala Gln Ser Val Gly Gly Ala Gln Phe Tyr Ile Ser Cys Ala Gln
165           170           175

Leu Ser Val Thr Gly Gly Gly Ser Thr Glu Pro Pro Asn Lys Val Ala
180           185           190

Phe Pro Gly Ala Tyr Ser Ala Thr Asp Pro Gly Ile Leu Ile Asn Ile
195           200           205

Tyr Tyr Pro Val Pro Thr Ser Tyr Gln Asn Pro Gly Pro Ala Val Phe
210           215           220

Ser Cys
225

```

<210> 9
 <211> 960
 <212> DNA
 <213> 土生梭孢壳(*Thielavia terrestris*)

```

<400> 9
atgaagggaac ttttcagtgc cgcgccttc tcctggccg tcggccaggc ttcggcccat      60

```

[0007]

```

tacatcttcc agcaactctc catcaacggg aaccagtttc cgggtgtacca atatattcgc      120
aagaacacca attataacag tcccgttacc gatctcacgt cgcacgatct tcggtgcaat      180
gtcggcgccc aggggtctgg gacagacacc gtcacggtga aggccggcga ccagttcacc      240
ttcacccttg acaceccctgt ttaccaccag gggcccctct ccatctacat gtccaaggcc      300
ccgggcgcgg cgtcagacta cgaatggcagc ggcggctggt tcaagatcaa ggactggggc      360
ccgactttca acgccgacgg caccggccacc tgggacatgg ccggctcata cactacaaac      420
ateccgacct gcatcccca cggcgactat ctgctccga tccagtcgt ggccatccac      480
aacccttggc cggcgggcat cccgcagttc tacatctcct gcgccagat caccgtgacc      540
ggcggcgcca accgcaacc tggcccagc gccctcctcc ccggcgccct caaggacacc      600
gacccgggct acacggtgaa catctacaac aacttcaca actacaagg tccccggccg      660
gaggtcttca gctgcaacgg cggcgcteg aaccggccc ccgggtgag tagcagcagc      720
cccgcgacca cgacgttgt cagctcagc cgcaccacgt cctccacgt ctcggcctcg      780
acgccggcct cgaccggcg ctgcaccgt gccaaatgg gccagtggc cggcaacggg      840
tacaccggct gcaagactg cgcggccggg tccacctgca gcaagcagaa cgactactac      900
tcgcagtgt tgtaaggag gccgcaaagc atgaggtgtt tgaagaggag gagagggtc      960

```

<210> 10
 <211> 304
 <212> PRT
 <213> 土生梭孢壳(Thielavia terrestris)

<400> 10

```

Met Lys Gly Leu Phe Ser Ala Ala Ala Leu Ser Leu Ala Val Gly Gln
1           5           10          15
Ala Ser Ala His Tyr Ile Phe Gln Gln Leu Ser Ile Asn Gly Asn Gln
20          25          30
Phe Pro Val Tyr Gln Tyr Ile Arg Lys Asn Thr Asn Tyr Asn Ser Pro
35          40          45
Val Thr Asp Leu Thr Ser Asp Asp Leu Arg Cys Asn Val Gly Ala Gln
50          55          60
Gly Ala Gly Thr Asp Thr Val Thr Val Lys Ala Gly Asp Gln Phe Thr
65          70          75          80
Phe Thr Leu Asp Thr Pro Val Tyr His Gln Gly Pro Ile Ser Ile Tyr
85          90          95
Met Ser Lys Ala Pro Gly Ala Ala Ser Asp Tyr Asp Gly Ser Gly Gly
100         105         110
Trp Phe Lys Ile Lys Asp Trp Gly Pro Thr Phe Asn Ala Asp Gly Thr
115        120        125
Ala Thr Trp Asp Met Ala Gly Ser Tyr Thr Tyr Asn Ile Pro Thr Cys
130        135        140
Ile Pro Asp Gly Asp Tyr Leu Leu Arg Ile Gln Ser Leu Ala Ile His
145        150        155        160
Asn Pro Trp Pro Ala Gly Ile Pro Gln Phe Tyr Ile Ser Cys Ala Gln
165        170        175
Ile Thr Val Thr Gly Gly Gly Asn Gly Asn Pro Gly Pro Thr Ala Leu
180        185        190

```

[0008]

Ile Pro Gly Ala Phe Lys Asp Thr Asp Pro Gly Tyr Thr Val Asn Ile
 195 200 205
 Tyr Thr Asn Phe His Asn Tyr Thr Val Pro Gly Pro Glu Val Phe Ser
 210 215 220
 Cys Asn Gly Gly Gly Ser Asn Pro Pro Pro Pro Val Ser Ser Ser Thr
 225 230 235 240
 Pro Ala Thr Thr Thr Leu Val Thr Ser Thr Arg Thr Thr Ser Ser Thr
 245 250 255
 Ser Ser Ala Ser Thr Pro Ala Ser Thr Gly Gly Cys Thr Val Ala Lys
 260 265 270
 Trp Gly Gln Cys Gly Gly Asn Gly Tyr Thr Gly Cys Thr Thr Cys Ala
 275 280 285
 Ala Gly Ser Thr Cys Ser Lys Gln Asn Asp Tyr Tyr Ser Gln Cys Leu
 290 295 300

<210> 11
 <211> 954
 <212> DNA
 <213> 土生梭孢壳(*Thielavia terrestris*)

<400> 11
 atgaagggcc tcagcctcct cgcgcgtgag tcggcagcga ctgctcatac catcttcgtg 60
 cagctcaggt cagggggaac gacctatccg gtatcctacg gcatccggga ccttagctac 120
 gacggtccca tcaccgacgt cacctccgac tcaactggtt gcaatgggtc cccgaacccc 180
 acgacgccgt ccccgtaaat catcaacgtc accgcgcgca ccacggtcgc ggcgatctgg 240
 aggcacaccc tcacatccgg ccccgacgat gtcattggac ccagccacaa ggggccgacc 300
 ctggcctacc tcaagaaggt cgatgatgcc ttgaccgaca cgggtatcgg cggcgctgg 360
 ttcaagatcc aggaggccgg tiacgacaat ggcaattggg ctaccagcac ggigatcacc 420
 aacggtggct tcaatatat tgacatcccc gccctgcatc ccaacggcca gtatctgtc 480
 cgcgcgaga tgatgcgct ccacgcgcc agcagcagg ggggtgcca gctctacatg 540
 gagtgcgcgc agatcaacgt ggtgggcggc tccgcagcgc ccagcccga gacgtacagc 600
 atcccgggca tctaccaggc aaccgacccg ggcctgtgta tcaacateta ctccatgacg 660
 ccgtccagcc agtacaccat tccgggtccg cccctgttca cctgcagcgg cagcggcaac 720
 aacggcggcg gcagcaaccc gtccggcggg cagaaccaga cggcgaagcc caccagcagc 780
 acggcgggga cgaccacetc ctccgcgct cctaccagea gccagggggg cagcagcgg 840
 tgcaccgttc ccagtgga gacgtgcggt ggcatctcgt tcaccggctg caccacctgc 900
 gcggcggggt acacctgcaa gtatctgaac gactattact cgcaatgcca gtaa 954

<210> 12
 <211> 317
 <212> PRT
 <213> 土生梭孢壳(*Thielavia terrestris*)

<400> 12
 Met Lys Gly Leu Ser Leu Leu Ala Ala Ala Ser Ala Ala Thr Ala His
 1 5 10 15
 Thr Ile Phe Val Gln Leu Glu Ser Gly Gly Thr Thr Tyr Pro Val Ser

[0009]

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

<210> 13
 <211> 799
 <212> DNA
 <213> 金黄色嗜热子囊菌(Thermoascus aurantiacus)

<400> 13
 atgtcctttt ccaagataat tgetactgcc ggcgttcttg cctctgcttc tctagtggct 60
 ggccatggct tcgttcagaa catcgtgatt gatggtaaaa agtatgtcat tgcaagacgc 120
 acataagegg caacagctga caatcgacag ttatggcggg tatctagtga accagtatcc 180
 atacatgtcc aatccctccag aggicacgc ctggtctact acggcaactg atcttggatt 240
 tctggacggt actggatacc aaaccccgaga tateatctgc cataggggcg ccaagcctgg 300

[0010]

```

agccctgaact gctccagtct ctccaggagg aactgttgag ctccaatgga ctccatgccc 360
tgattctcac catggccag ttatcaacta ccttgctccg tgcaatgggt attgttcac 420
tgtggataag acccaattag aattcttcaa aattgccgag ageggctctca tcaatgatga 480
caatctctct gggatciggg ctccagacaa tctgatagca gccacaaca gctggactgt 540
caccattcca accacaattg cacttgaaa ctaigtcttg aggcattgaga ttattgctct 600
tcactcagct cagaaccagg atgggcecca gaactatccc cagtgcattc atctgcaggt 660
cactggaggt ggttctgata accctgctgg aactcttga acggcactct accacgatac 720
cgatcttgga attctgatca acatctatca gaaacttccc agctatatca tccctggtcc 780
tcctctgtat actggttaa 799

```

<210> 14
 <211> 250
 <212> PRT
 <213> 金黄色嗜热子囊菌(*Thermoascus aurantiacus*)

<220>
 <221> 信号
 <222> (1)..(22)

<220>
 <221> 成熟肽
 <222> (23)..(250)

<400> 14

```

Met Ser Phe Ser Lys Ile Ile Ala Thr Ala Gly Val Leu Ala Ser Ala
  -20              -15              -10
Ser Leu Val Ala Gly His Gly Phe Val Gln Asn Ile Val Ile Asp Gly
  -5              -1 1              5              10
Lys Lys Tyr Tyr Gly Gly Tyr Leu Val Asn Gln Tyr Pro Tyr Met Ser
              15              20              25
Asn Pro Pro Glu Val Ile Ala Trp Ser Thr Thr Ala Thr Asp Leu Gly
              30              35              40
Phe Val Asp Gly Thr Gly Tyr Gln Thr Pro Asp Ile Ile Cys His Arg
              45              50              55
Gly Ala Lys Pro Gly Ala Leu Thr Ala Pro Val Ser Pro Gly Gly Thr
              60              65              70
Val Glu Leu Gln Trp Thr Pro Trp Pro Asp Ser His His Gly Pro Val
              75              80              85              90
Ile Asn Tyr Leu Ala Pro Cys Asn Gly Asp Cys Ser Thr Val Asp Lys
              95              100              105
Thr Gln Leu Glu Phe Phe Lys Ile Ala Glu Ser Gly Leu Ile Asn Asp
              110              115              120
Asp Asn Pro Pro Gly Ile Trp Ala Ser Asp Asn Leu Ile Ala Ala Asn
              125              130              135
Asn Ser Trp Thr Val Thr Ile Pro Thr Thr Ile Ala Pro Gly Asn Tyr
              140              145              150
Val Leu Arg His Glu Ile Ile Ala Leu His Ser Ala Gln Asn Gln Asp
              155              160              165              170
Gly Ala Gln Asn Tyr Pro Gln Cys Ile Asn Leu Gln Val Thr Gly Gly
              175              180              185

```

[0011]

Gly Ser Asp Asn Pro Ala Gly Thr Leu Gly Thr Ala Leu Tyr His Asp
 190 195 200
 Thr Asp Pro Gly Ile Leu Ile Asn Ile Tyr Gln Lys Leu Ser Ser Tyr
 205 210 215
 Ile Ile Pro Gly Pro Pro Leu Tyr Thr Gly
 220 225

<210> 15
 <211> 1172
 <212> DNA
 <213> 里氏木霉(Trichoderma reesei)

<400> 15
 ggatctaagc cccatcgata tgaagtcctg cgccattctt gcagccettg gctgtcttgc 60
 cgggagcggt ctcggccatg gacaagtcca aaatttcacg atcaatggac aatacaatca 120
 gggtttcatt ctcgattact aciatcagaa gcagaatact ggtaacttcc ccaacgttgc 180
 tggctgttac gccgaggacc tagaactggg ctccatctcc cctgaccaat acaccacgcc 240
 cgacattgtc tgcacaaga acgcggcccc aggtgccatt tctgccactg cagcgccggg 300
 cagcaacatc gtcttccaat ggggccctgg cgtctggcct caccctacg gtcccatcgt 360
 taactacgtg gctgagtgea gccgatcgtg cagcaccgtg aacaagaaca acctgcgctg 420
 ggtaagatt caggagcgcc gcataacta taacaccena gtctgggcgc agcaggatct 480
 galcaaccag ggcaacaagt ggaactgtga gatcccgteg agctcaggc ccggaacta 540
 tgtcttcgc catgaacttc ttctgccc aigtgcctct agtgcaacg gcattgcagaa 600
 ctatcctcag tgcgtgaaca tcgcctcac aggtcgggc acgaaagcg tccctgcggg 660
 aactctgca actcagctct acaagccac tgacctggc atcttgtta acccttacac 720
 aacaatcac agctacacca tccctggccc agccctgtgg caaggetaga tccaggggta 780
 cggctgttgc gttcgtgaag tcggagctgt tgacaaggat atctgatgat gaacggagag 840
 gactgatggg cgtgactgag tgtatatatt ttgatgacc aaattgtata cgaaatccga 900
 acgcatgggt atcattgttt atccctgtag tatattgtct ccaggctgct aagagccac 960
 cgggtgtatt acggcaacaa agtcaggaat ttgggtggca atgaacgcag gtctccatga 1020
 atgtatatgt gaagaggcat cggctggcat gggcattacc agatatagge cctgtgaaac 1080
 atatatgtact tgaacgtgct actggaacgg atcataagca agtcataac atgtgaaaaa 1140
 aaactacatg taataaaaaa aaaaaaaa aa 1172

<210> 16
 <211> 249
 <212> PRT
 <213> 里氏木霉(Trichoderma reesei)

<400> 16
 Met Lys Ser Cys Ala Ile Leu Ala Ala Leu Gly Cys Leu Ala Gly Ser
 1 5 10 15
 Val Leu Gly His Gly Gln Val Gln Asn Phe Thr Ile Asn Gly Gln Tyr
 20 25 30
 Asn Gln Gly Phe Ile Leu Asp Tyr Tyr Tyr Gln Lys Gln Asn Thr Gly
 35 40 45

[0012]

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

<210> 17

<211> 2576

<212> DNA

<213> 特异腐质霉(*Humicola insolens*)

<400> 17

agctacagct tccctgggccc cgtctgaacc aacctctctgg gaccaagtgg tgagatctgg 60
 cggcacaacc atgaagtccc tcggccgtat tggggcgacc gcccttgcgg cgtegtgta 120
 tctcacatca ggcgcgcgc aagccactgg tgatgcgtac accgaatcgg aaacaggcat 180
 taagtccag acctggccc cggtaccga gttactttt gcccttgcgc tgccgcggga 240
 tgccctggag aaggatgcca ctgagtacat tggctctctc cgtgcacca ggcccgaccc 300
 atccgaaccc ggtaactgg gtcctctcca tggccaggtc gccagatga cgcagtcgt 360
 gcttctcgtg gccctggccc acgagaacca ggtctacacg tcgttccgt acgccaccgg 420
 ctacaccctc ccgggtctgt acaccggcaa cgttaagctg acccagctct ccgtcaacat 480
 caccgacacc agcttcgagc tcatctaccg ctgcgagaa tgcctctctg gggagcacga 540
 aggcagcacc ggaatctagt egacctcca ggctatctc gtcctcggtc gtgcttccgc 600
 ccgcgcgggc gtcgtggccc cgaattgccc ggacaaggcc accttgggt tccacgacaa 660
 tggttcctgt cagtggggtg ttggtctga gaatgcgtt tcggagcagt attctgagt 720
 ggcttcgtg ccgggtctga ctgttgagac caactgcgaa ggatccggcc ctggtgagcc 780
 gcagtcgtg cctgcccctg aggagactta tgactatatt gttgttggtg ctggcgccgg 840

[0013]

```

cgggtattcct gtcgccgaca agctgagcga ggccggccac aaggtttctg tcatcgagaa    900
gggtcccccgc tcgacgggce gctggcaggg taccatgaag cccgagtgge ttgaaggcac    960
tgacctcact cggttcgatg tgcccggcct ttgcaaccag atctgggttg actcggctgg    1020
cattgcctgc actgatactg atcagatggc tggtctcgtc ttggcggtg gcacggccgt    1080
taatgctggc ctgtgtgga agccattga cctcgactgg gatgagaact tccctgaggg    1140
ctggcactcg caggatctcg ccggcgagc cgagcgctg ttgagcgca tcccggcac    1200
ctggcaccgc tccatgatg gcaagctgta ccgtgacgaa ggctacaagg ttctctccag    1260
cggctgtggt gagtctggt ggaaggaggt tgtggccaac gaggttccca acgagaagaa    1320
ccgcactttc gccacacccc acttcatggt cgtggcgga gacgtaacg ggctctctg    1380
cacttacctg gtctctgccc atgcccgga gaacttctg ccttgacca aactgtgt    1440
tcgcgcgct gtcgcactg gtggcaaggt cacaggtgc gagctcagt gctgactga    1500
tgggcgctac agcggcattg ttaagctcaa tgaggcggt ggctcatct tctcgccgg    1560
tgcttttgt teggccaagc tctcttccg cagcgctatc ggccctgagg atcagctccg    1620
cgttgtgtgc tctcttaagg acggagagga cttcatcgac gagaaggact ggattaagct    1680
ccccgtcggc tacaacctga ttgaccacct taacaetgac ctcactctca ctcacccga    1740
tgtctcttc taegacttct atgaggcctg gaccaccccg atcgaggccg acaagcagct    1800
gtaccttgag cagcgctctg gcactcttgc ccagctgct cctaacattg gcccattgat    1860
gtgggagcag gtacacccct cggacggcat taccgcgcaa ttccagtgga cggctcgcgt    1920
cgaggcgagc agccgcttca ccaactcttc tcatgcatg actctcagc agtacctcg    1980
ccgtgtgtgc gtgtcgcgcg gtgcgccac caccaccag ggtctcgtca ccaccgtggc    2040
tgagcaccgc taactccaca acgcggcgga caaggagcc gtcattcagg gcatcaagaa    2100
ctcatttgag tctcttaacg tgattcccaa cactacttgg gtctgcgc ctcctgtag    2160
cactgtcgag gaatacgtcg attcgtcct cgtctccgc tcggctcgtc gctegaacca    2220
ctggatgggc acggccaagc tgggtaetga tgatgcgcg tacggcggt cttcggtcgt    2280
cgacctcgac accaaggtct acggcaccga taacctgttc gtgttgatg ctccatctt    2340
ccctggcatg tcgacggca acccgctcgc tatgatcgtg attgcctg agcaggtgc    2400
ggagcgcatt cttagctga ggaagtaaga aggggagaga ggatggagg atgacattga    2460
ggaaaatagg gttatgagtt gatgagttat gggcgaaatgt gtcagccagt gtacttgact    2520
tattacctga gttaaacaac acgacgtgct tgatgtgta aaaaaaaaaa aacttt    2576

```

<210> 18

<211> 785

<212> PRT

<213> 特异腐质霉(*Humicola insolens*)

<400> 18

```

Met Lys Phe Leu Gly Arg Ile Gly Ala Thr Ala Leu Ala Ala Ser Leu
1           5           10          15
Tyr Leu Thr Ser Gly Ala Ala Gln Ala Thr Gly Asp Ala Tyr Thr Asp
                20          25          30

```

[0014]

Ser Glu Thr Gly Ile Lys Phe Gln Thr Trp Ser Pro Asp Pro Gln Phe
 35 40 45
 Thr Phe Gly Leu Ala Leu Pro Pro Asp Ala Leu Glu Lys Asp Ala Thr
 50 55 60
 Glu Tyr Ile Gly Leu Leu Arg Cys Thr Arg Ala Asp Pro Ser Asp Pro
 65 70 75 80
 Gly Tyr Cys Gly Leu Ser His Gly Gln Val Gly Gln Met Thr Gln Ser
 85 90 95
 Leu Leu Leu Val Ala Trp Ala Tyr Glu Asn Gln Val Tyr Thr Ser Phe
 100 105 110
 Arg Tyr Ala Thr Gly Tyr Thr Leu Pro Gly Leu Tyr Thr Gly Asn Ala
 115 120 125
 Lys Leu Thr Gln Leu Ser Val Asn Ile Thr Asp Thr Ser Phe Glu Leu
 130 135 140
 Ile Tyr Arg Cys Glu Asn Cys Phe Ser Trp Glu His Glu Gly Ser Thr
 145 150 155 160
 Gly Ser Ser Ser Thr Ser Gln Gly Tyr Leu Val Leu Gly Arg Ala Ser
 165 170 175
 Ala Arg Arg Gly Val Val Gly Pro Thr Cys Pro Asp Thr Ala Thr Phe
 180 185 190
 Gly Phe His Asp Asn Gly Phe Gly Gln Trp Gly Val Gly Leu Glu Asn
 195 200 205
 Ala Val Ser Glu Gln Tyr Ser Glu Trp Ala Ser Leu Pro Gly Leu Thr
 210 215 220
 Val Glu Thr Thr Cys Glu Gly Ser Gly Pro Gly Glu Ala Gln Cys Val
 225 230 235 240
 Pro Ala Pro Glu Glu Thr Tyr Asp Tyr Ile Val Val Gly Ala Gly Ala
 245 250 255
 Gly Gly Ile Pro Val Ala Asp Lys Leu Ser Glu Ala Gly His Lys Val
 260 265 270
 Leu Leu Ile Glu Lys Gly Pro Pro Ser Thr Gly Arg Trp Gln Gly Thr
 275 280 285
 Met Lys Pro Glu Trp Leu Gln Gly Thr Asp Leu Thr Arg Phe Asp Val
 290 295 300
 Pro Gly Leu Cys Asn Gln Ile Trp Val Asp Ser Ala Gly Ile Ala Cys
 305 310 315 320
 Thr Asp Thr Asp Gln Met Ala Gly Cys Val Leu Gly Gly Gly Thr Ala
 325 330 335
 Val Asn Ala Gly Leu Trp Trp Lys Pro Ile Asp Leu Asp Trp Asp Glu
 340 345 350
 Asn Phe Pro Glu Gly Trp His Ser Gln Asp Leu Ala Ala Ala Thr Glu
 355 360 365
 Arg Val Phe Glu Arg Ile Pro Gly Thr Trp His Pro Ser Met Asp Gly
 370 375 380
 Lys Leu Tyr Arg Asp Glu Gly Tyr Lys Val Leu Ser Ser Gly Leu Ala
 385 390 395 400
 Glu Ser Gly Trp Lys Glu Val Val Ala Asn Glu Val Pro Asn Glu Lys
 405 410 415
 Asn Arg Thr Phe Ala His Thr His Phe Met Phe Ala Gly Gly Glu Arg

[0015]

	420	425	430
Asn Gly Pro Leu Ala Thr Tyr Leu Val Ser Ala Asp Ala Arg Glu Asn			
435	440	445	
Phe Ser Leu Trp Thr Asn Thr Ala Val Arg Arg Ala Val Arg Thr Gly			
450	455	460	
Gly Lys Val Thr Gly Val Glu Leu Glu Cys Leu Thr Asp Gly Gly Tyr			
465	470	475	480
Ser Gly Ile Val Lys Leu Asn Glu Gly Gly Gly Val Ile Phe Ser Ala			
485	490	495	
Gly Ala Phe Gly Ser Ala Lys Leu Leu Phe Arg Ser Gly Ile Gly Pro			
500	505	510	
Glu Asp Gln Leu Arg Val Val Ala Ser Ser Lys Asp Gly Glu Asp Phe			
515	520	525	
Ile Asp Glu Lys Asp Trp Ile Lys Leu Pro Val Gly Tyr Asn Leu Ile			
530	535	540	
Asp His Leu Asn Thr Asp Leu Ile Leu Thr His Pro Asp Val Val Phe			
545	550	555	560
Tyr Asp Phe Tyr Glu Ala Trp Thr Thr Pro Ile Glu Ala Asp Lys Gln			
565	570	575	
Leu Tyr Leu Glu Gln Arg Ser Gly Ile Leu Ala Gln Ala Ala Pro Asn			
580	585	590	
Ile Gly Pro Met Met Trp Glu Gln Val Thr Pro Ser Asp Gly Ile Thr			
595	600	605	
Arg Gln Phe Gln Trp Thr Ala Arg Val Glu Gly Asp Ser Arg Phe Thr			
610	615	620	
Asn Ser Ser His Ala Met Thr Leu Ser Gln Tyr Leu Gly Arg Gly Val			
625	630	635	640
Val Ser Arg Gly Arg Ala Thr Ile Thr Gln Gly Leu Val Thr Thr Val			
645	650	655	
Ala Glu His Pro Tyr Leu His Asn Ala Gly Asp Lys Glu Ala Val Ile			
660	665	670	
Gln Gly Ile Lys Asn Leu Ile Glu Ser Leu Asn Val Ile Pro Asn Ile			
675	680	685	
Thr Trp Val Leu Pro Pro Pro Gly Ser Thr Val Glu Glu Tyr Val Asp			
690	695	700	
Ser Leu Leu Val Ser Ala Ser Ala Arg Arg Ser Asn His Trp Met Gly			
705	710	715	720
Thr Ala Lys Leu Gly Thr Asp Asp Gly Arg Tyr Gly Gly Thr Ser Val			
725	730	735	
Val Asp Leu Asp Thr Lys Val Tyr Gly Thr Asp Asn Leu Phe Val Val			
740	745	750	
Asp Ala Ser Ile Phe Pro Gly Met Ser Thr Gly Asn Pro Ser Ala Met			
755	760	765	
Ile Val Ile Ala Ala Glu Gln Ala Ala Glu Arg Ile Leu Lys Leu Arg			
770	775	780	
Lys			
785			

<210> 19

[0016]

<211>	923	
<212>	DNA	
<213>	特异腐质霉(<i>Humicola insolens</i>)	
<400>	19	
atgcgttccct ccccccctct ccgcctccgc gttgtggccg cctgcccgt gttggccctt	60	
gccgctgatg gcaggteccac ccgtactgg gactgetgca agccttcgtg cggetgggccc	120	
aagaaggctc ccgtgaacca gccgtgtttt tctgcaacg ccaacttcca gcgtatcaag	180	
gacttcgaag ccaagtcagg ctgcgagccg ggccgtgtcg cctactcgtg cgccgaccag	240	
accccatggg ctgtgaacga cgacttcgg ctgggttttg ctgccacctc tattgccggc	300	
agcaatgagg cgggctgggt ctgcgcctgc tacgagctca cttcacatc cggtcctgtt	360	
gctggcaaga agatggctgt ccagtcacc agcactggcg gtgatcttgg cagcaaccac	420	
ttcgatctca acatcccgg cgccggcgtc ggcatcttcg acggatgcac tcccagttc	480	
ggcggtctgc ccggccagcg ctacggcggc atctcgtccc gcaacgagtg cgatcggttc	540	
cccgacgccc tcaagccgg ctgctactgg cgttcgact gggtcaagaa cgccgacaat	600	
ccgagcttca gcttcgta ggtccagtc ccagccgagc tcgtcgctcg caccggatgc	660	
cgccgcaacg acgacggcaa ctccctgcc gtccagatcc cctccagcag caccagctct	720	
ccggtcaacc agcctaccag caccagcacc acgtccacct ccaccacctc gagcccgcca	780	
gtccagccta cgactcccag cggctgcaat gctgagaggt gggtcagtg cggcggcaat	840	
ggctggagcg gctgcaccac ctgcgtcgtt gccagcactt gcacgaagat taatgactgg	900	
taccatcagt gccgttagaa ttc	923	
<210>	20	
<211>	305	
<212>	PRT	
<213>	特异腐质霉(<i>Humicola insolens</i>)	
<400>	20	
Met Arg Ser Ser Pro Leu Leu Arg Ser Ala Val Val Ala Ala Leu Pro		
1 5 10 15		
Val Leu Ala Leu Ala Ala Asp Gly Arg Ser Thr Arg Tyr Trp Asp Cys		
20 25 30		
Cys Lys Pro Ser Cys Gly Trp Ala Lys Lys Ala Pro Val Asn Gln Pro		
35 40 45		
Val Phe Ser Cys Asn Ala Asn Phe Gln Arg Ile Thr Asp Phe Asp Ala		
50 55 60		
Lys Ser Gly Cys Glu Pro Gly Gly Val Ala Tyr Ser Cys Ala Asp Gln		
65 70 75 80		
Thr Pro Trp Ala Val Asu Asp Asp Phe Ala Leu Gly Phe Ala Ala Thr		
85 90 95		
Ser Ile Ala Gly Ser Asn Glu Ala Gly Trp Cys Cys Ala Cys Tyr Glu		
100 105 110		
Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Lys Met Val Val Gln		
115 120 125		
Ser Thr Ser Thr Gly Gly Asp Leu Gly Ser Asn His Phe Asp Leu Asn		
130 135 140		
Ile Pro Gly Gly Gly Val Gly Ile Phe Asp Gly Cys Thr Pro Gln Phe		

[0017]

145	150	155	160
Gly Gly Leu Pro	Gly Gln Arg Tyr Gly Gly Ile Ser Ser Arg Asn Glu		
	165	170	175
Cys Asp Arg Phe Pro Asp Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe			
	180	185	190
Asp Trp Phe Lys Asn Ala Asp Asn Pro Ser Phe Ser Phe Arg Gln Val			
	195	200	205
Gln Cys Pro Ala Glu Leu Val Ala Arg Thr Gly Cys Arg Arg Asn Asp			
	210	215	220
Asp Gly Asn Phe Pro Ala Val Gln Ile Pro Ser Ser Ser Thr Ser Ser			
	225	230	235
Pro Val Asn Gln Pro Thr Ser Thr Ser Thr Thr Ser Thr Ser Thr Thr			
	245	250	255
Ser Ser Pro Pro Val Gln Pro Thr Thr Pro Ser Gly Cys Thr Ala Glu			
	260	265	270
Arg Trp Ala Gln Cys Gly Gly Asn Gly Trp Ser Gly Cys Thr Thr Cys			
	275	280	285
Val Ala Gly Ser Thr Cys Thr Lys Ile Asn Asp Trp Tyr His Gln Cys			
	290	295	300
Leu			
305			

<210> 21
 <211> 1188
 <212> DNA
 <213> 嗜热毁丝霉(Myceliophthora thermophila)

<400> 21
 cgacttgaaa cgcccacaaat gaagtcctcc atcctcgcca gcgtcttcgc caccggcgccc 60
 gtggctcaaa gtgtccgtg gcagcaatgt ggtggcctcg gatggcaagg atcgaccgac 120
 tgtgtgtcgg gctaccactg cgtctaccag aacgatttgt acagccagtg cgtgccctggc 180
 gcggcgctga caacgtgca gacatcgacc acgtccaggc ccaccgccac cagcacccgc 240
 cctccgttgt ccaccacctc gcttagcaag ggcaagctga agtggctcgg cagcaacgag 300
 tcgggcgcgc agttcgggga gggcaattac cccggcctct ggggcaagca ctctatcttc 360
 ccgtgcattt cggcgattca gacgtctatc aatgatggat acaacatctt ccggatcgac 420
 ttctcgatgg agcgtcttgt gcccaaccag ttgacgtcgt ccttcgacca gggttacctc 480
 cgcaacctga ccgaggttgt caacttcgtg acgaacgcgg gcaagtacgc cgtccctggac 540
 ccgcacaact acggccggtg ctacggcaac atcatcacgg acacgaacgc gtcccgacc 600
 ttctggacca acctggccaa gcagttcgcc tccaactcgc tcgtcatctt cgacaccaac 660
 aacgagtaaa caacgatgga ccagaccctg gtgtcaacc tcaaccaggc cgccatcgac 720
 ggcatccggg ccgccggcgc gacctcgag tacatcttcg tcgagggcaa cgcgtggagc 780
 ggggcctgga gctggaacac gaccaacacc aacatggcgg cctgacgga ccgcagaaac 840
 aagatcgtgt acgagatgca ccagtacctc gactcggaca gtcggggcac ccacgccgag 900
 tgcgtcagca gcaccatcgg cgtccagcgc gtctcgggag ccaccagtg gctccgcgcc 960
 aacgggaagc tcggcgtcct cggcgagttc gcggcgggcg ccaacgccgt ctgccagcag 1020

[0018]

gccgtcaccg gcctcctcga ccacctccag gacaacagcg acgtctggct gggtgcectc 1080
 tggtagggcgg ccggtccttg gtggggcgac tacatgtact cgttcgagcc tecttcgggc 1140
 accggctatg tcaactacaa ctcgatcttg aagaagtact tgccgtaa 1188

<210> 22
 <211> 389
 <212> PRT
 <213> 嗜热毁丝霉(*Myceliophthora thermophila*)

<400> 22

Met Lys Ser Ser Ile Leu Ala Ser Val Phe Ala Thr Gly Ala Val Ala
 1 5 10 15
 Gln Ser Gly Pro Trp Gln Gln Cys Gly Gly Ile Gly Trp Gln Gly Ser
 20 25 30
 Thr Asp Cys Val Ser Gly Tyr His Cys Val Tyr Gln Asn Asp Trp Tyr
 35 40 45
 Ser Gln Cys Val Pro Gly Ala Ala Ser Thr Thr Leu Gln Thr Ser Thr
 50 55 60
 Thr Ser Arg Pro Thr Ala Thr Ser Thr Ala Pro Pro Ser Ser Thr Thr
 65 70 75 80
 Ser Pro Ser Lys Gly Lys Leu Lys Trp Leu Gly Ser Asn Glu Ser Gly
 85 90 95
 Ala Glu Phe Gly Glu Gly Asn Tyr Pro Gly Leu Trp Gly Lys His Phe
 100 105 110
 Ile Phe Pro Ser Thr Ser Ala Ile Gln Thr Leu Ile Asn Asp Gly Tyr
 115 120 125
 Asn Ile Phe Arg Ile Asp Phe Ser Met Glu Arg Leu Val Pro Asn Gln
 130 135 140
 Leu Thr Ser Ser Phe Asp Gln Gly Tyr Leu Arg Asn Leu Thr Glu Val
 145 150 155 160
 Val Asn Phe Val Thr Asn Ala Gly Lys Tyr Ala Val Leu Asp Pro His
 165 170 175
 Asn Tyr Gly Arg Tyr Tyr Gly Asn Ile Ile Thr Asp Thr Asn Ala Phe
 180 185 190
 Arg Thr Phe Trp Thr Asn Leu Ala Lys Gln Phe Ala Ser Asn Ser Leu
 195 200 205
 Val Ile Phe Asp Thr Asn Asn Glu Tyr Asn Thr Met Asp Gln Thr Leu
 210 215 220
 Val Leu Asn Leu Asn Gln Ala Ala Ile Asp Gly Ile Arg Ala Ala Gly
 225 230 235 240
 Ala Thr Ser Gln Tyr Ile Phe Val Glu Gly Asn Ala Trp Ser Gly Ala
 245 250 255
 Trp Ser Trp Asn Thr Thr Asn Thr Asn Met Ala Ala Leu Thr Asp Pro
 260 265 270
 Gln Asn Lys Ile Val Tyr Glu Met His Gln Tyr Leu Asp Ser Asp Ser
 275 280 285
 Ser Gly Thr His Ala Glu Cys Val Ser Ser Thr Ile Gly Ala Gln Arg
 290 295 300
 Val Val Gly Ala Thr Gln Trp Leu Arg Ala Asn Gly Lys Leu Gly Val
 305 310 315 320

[0019]

Leu Gly Glu Phe Ala Gly Gly Ala Asn Ala Val Cys Gln Gln Ala Val
 325 330 335
 Thr Gly Leu Leu Asp His Leu Gln Asp Asn Ser Asp Val Trp Leu Gly
 340 345 350
 Ala Leu Trp Trp Ala Ala Gly Pro Trp Trp Gly Asp Tyr Met Tyr Ser
 355 360 365
 Phe Glu Pro Pro Ser Gly Thr Gly Tyr Val Asn Tyr Asn Ser Ile Leu
 370 375 380
 Lys Lys Tyr Leu Pro
 385

<210> 23
 <211> 1232
 <212> DNA
 <213> 担子菌纲(Basidiomycete) CBS 495.95

<400> 23
 ggatccacit agtaacggcc gccagtgtgc tggaaagcat gaagtctctc ttctctgtcac 60
 ttgtagcgac cgtcgcgtc agctcgccag tattctctgt cgcagtctgg gggcaatgeg 120
 gcggcattgg cttcagcggc agcacgtct gtgatgcagg cgcggctgt gtgaagctca 180
 acgactatta ctctcaatgc caaccggcg ctccactgc tacatccgcg gcgccaagta 240
 gcaacgcacc gtcggcact tcgacggcct cggccccctc ctccagcctt tgccttgcca 300
 gccgcacgcc gttccagttc ttccgtgtca acgaatccgg cgcggagttc ggcaacctga 360
 acatccccgg tgttctgggc accgaactaca cctggccgtc gccatccagc attgacttct 420
 tcatgggcaa gggaatgaat accttccgta ttccgttctt eatggagcgt ctgttcccc 480
 ctgccactgg catcacagga cctctcgacc agacgtacit ggcgggcctg cagacgattg 540
 tcaactacat caccggcaaa ggcggctttg ctctcattga ccgcacaaac ttatgatct 600
 acaatggcca gacgatctc agtaccagcg acttccagaa gttctggcag aacctcgag 660
 gagtgtttaa atcgaacagt cagctcatct tcgatgttat gaacgagcct cagcatatc 720
 ccgccagac cgtgttccaa ctgaaccaag ccgtgtcaa tggcatccgt gcgagcggtg 780
 cgacgtcgca gctcattctg gtcgagggca caagctggac tggagcctgg acctggacga 840
 cctctggcaa cagcgatgca ttccgtgcca ttaaggatcc caacaacaa gtcgcatcc 900
 agatgcata gttacctgat agcgatggct ctggcacttc gcagacctgc gigtctccca 960
 ccatcgggtc cgagcggttg caggctgcca ctcaatggtt gaagcagaac aacctcaagg 1020
 gcttcttggg cgagatcggc gccgcteta actccgcttg catcagcgt gtgcagggtg 1080
 cgttgtgttc gatgcagcaa tctgggtgtt ggctcgggc tctctgggtg gctgcgggac 1140
 cgttgtgggg cgactactac cagtcacatg agccgccctc tggccccggc gtgtccgca 1200
 tectccgca ggccctgtg cgttccggt aa 1232

<210> 24
 <211> 397
 <212> PRT
 <213> 担子菌纲(Basidiomycete) CBS 495.95
 <400> 24

[0020]

```

Met Lys Ser Leu Phe Leu Ser Leu Val Ala Thr Val Ala Leu Ser Ser
1           5           10           15

Pro Val Phe Ser Val Ala Val Trp Gly Gln Cys Gly Gly Ile Gly Phe
20           25           30

Ser Gly Ser Thr Val Cys Asp Ala Gly Ala Gly Cys Val Lys Leu Asn
35           40           45

Asp Tyr Tyr Ser Gln Cys Gln Pro Gly Ala Pro Thr Ala Thr Ser Ala
50           55           60

Ala Pro Ser Ser Asn Ala Pro Ser Gly Thr Ser Thr Ala Ser Ala Pro
65           70           75           80

Ser Ser Ser Leu Cys Ser Gly Ser Arg Thr Pro Phe Gln Phe Phe Gly
85           90           95

Val Asn Glu Ser Gly Ala Glu Phe Gly Asn Leu Asn Ile Pro Gly Val
100          105          110

Leu Gly Thr Asp Tyr Thr Trp Pro Ser Pro Ser Ser Ile Asp Phe Phe
115          120          125

Met Gly Lys Gly Met Asn Thr Phe Arg Ile Pro Phe Leu Met Glu Arg
130          135          140

Leu Val Pro Pro Ala Thr Gly Ile Thr Gly Pro Leu Asp Gln Thr Tyr
145          150          155          160

Leu Gly Gly Leu Gln Thr Ile Val Asn Tyr Ile Thr Gly Lys Gly Gly
165          170          175

Phe Ala Leu Ile Asp Pro His Asn Phe Met Ile Tyr Asn Gly Gln Thr
180          185          190

Ile Ser Ser Thr Ser Asp Phe Gln Lys Phe Trp Gln Asn Leu Ala Gly
195          200          205

Val Phe Lys Ser Asn Ser His Val Ile Phe Asp Val Met Asn Glu Pro
210          215          220

His Asp Ile Pro Ala Gln Thr Val Phe Gln Leu Asn Gln Ala Ala Val
225          230          235          240

Asn Gly Ile Arg Ala Ser Gly Ala Thr Ser Gln Leu Ile Leu Val Glu
245          250          255

Gly Thr Ser Trp Thr Gly Ala Trp Thr Trp Thr Thr Ser Gly Asn Ser
260          265          270

Asp Ala Phe Gly Ala Ile Lys Asp Pro Asn Asn Asn Val Ala Ile Gln
275          280          285

Met His Gln Tyr Leu Asp Ser Asp Gly Ser Gly Thr Ser Gln Thr Cys
290          295          300

Val Ser Pro Thr Ile Gly Ala Glu Arg Leu Gln Ala Ala Thr Gln Trp
305          310          315          320

Leu Lys Gln Asn Asn Leu Lys Gly Phe Leu Gly Glu Ile Gly Ala Gly
325          330          335

Ser Asn Ser Ala Cys Ile Ser Ala Val Gln Gly Ala Leu Cys Ser Met
340          345          350

Gln Gln Ser Gly Val Trp Leu Gly Ala Leu Trp Trp Ala Ala Gly Pro
355          360          365

Trp Trp Gly Asp Tyr Tyr Gln Ser Ile Glu Pro Pro Ser Gly Pro Ala
370          375          380

Val Ser Ala Ile Leu Pro Gln Ala Leu Leu Pro Phe Ala

```

[0021]

385	390	395
<210> 25 <211> 1303 <212> DNA <213> 担子菌纲(Basidiomycete) CBS 495.95 <400> 25 ggaaagegtc agtatggtga aatttgcgct tgtggcaact gtcggcgcaa tcttgagcgc 60 ttctggggcc aatgcggtt ctatctacca gcaatgtgga ggcattgat ggtctgggtc 120 cactgtttgc gagccggtc tcgcttgcgt tatcctcaat gcgtactact ttcagtgtt 180 gagcccgcc gcgggccaga caacgacggg ctccggcgca ccggcgtaaa catcaacctc 240 tcaactcaacg gtcactacgg ggagctcaca ctcaacaacc gggacgacgg cgacgaaaac 300 aactaccaet ccgtcgacca ccaagacctt acccgccatc tctgtgtctg gtccgctctg 360 ctctggctcc aggacgaagt tcaagttctt cgggtlgaat gaaagcggcg ccgaattcgg 420 gaacactget tggccagggc agctcgggaa agactataca tggccttcgc ctacgagcgt 480 ggactacttc atgggggctg gattcaatac attccgtatc acctctctga tggagcgtat 540 gagccctccg gctaccggac tcactggccc attcaaccag acgtacctgt cgggcctcac 600 caccattgtc gactacatca cgaacaaagg aggatacgtt cttattgacc cccacaactt 660 catgcgttac aacaacggca taatcagcag cacatcigac ttcgcgactt ggiggagcaa 720 tttggccact gtattcaaat ccacgaagaa cgcctcttc gacatccaga acgagccgta 780 cggaaatgat gcgagaccg tatacgaact gaatcaagct gccatcaatt cgatccgcgc 840 cgttggcgct acgtcacagt tgattctggt tgaaggaac tcatacactg gagcttggac 900 gtgggtctcg tccggaacg gagctgtttt cgcggcgtt acggatcctt acaacaacac 960 ggcaantgaa atgcaccaat acctcgacag cgacggttct gggacaaacg aagactgtgt 1020 ctctccacc attggtctgc aacgtctcca agctgccact gcgtgctgc aacaaacagg 1080 actcaaggga ttctcggag agacgggtgc tgggtcgaat tcccagtga tcgacgccgt 1140 gttcgatgaa ctttctata tgcaacagca aggcggctcc tggatcgggt cactctgggt 1200 ggttgcgggt ccttgggtgg gcacgtacat ttactcgatt gaacctccga gcggtgccgc 1260 tatcccagaa gtcttcttc aggtctctgc tccattctc tag 1303		
<210> 26 <211> 429 <212> PRT <213> 担子菌纲(Basidiomycete) CBS 495.95 <400> 26 Met Val Lys Phe Ala Leu Val Ala Thr Val Gly Ala Ile Leu Ser Ala 1 5 10 15 Ser Ala Ala Asn Ala Ala Ser Ile Tyr Gln Gln Cys Gly Gly Ile Gly 20 25 30 Trp Ser Gly Ser Thr Val Cys Asp Ala Gly Leu Ala Cys Val Ile Leu 35 40 45 Asn Ala Tyr Tyr Phe Gln Cys Leu Thr Pro Ala Ala Gly Gln Thr Thr 50 55 60		

[0022]

Thr Gly Ser Gly Ala Pro Ala Ser Thr Ser Thr Ser His Ser Thr Val
 65 70 75 80
 Thr Thr Gly Ser Ser His Ser Thr Thr Gly Thr Thr Ala Thr Lys Thr
 85 90 95
 Thr Thr Thr Pro Ser Thr Thr Thr Thr Leu Pro Ala Ile Ser Val Ser
 100 105 110
 Gly Arg Val Cys Ser Gly Ser Arg Thr Lys Phe Lys Phe Phe Gly Val
 115 120 125
 Asn Glu Ser Gly Ala Glu Phe Gly Asn Thr Ala Trp Pro Gly Gln Leu
 130 135 140
 Gly Lys Asp Tyr Thr Trp Pro Ser Pro Ser Ser Val Asp Tyr Phe Met
 145 150 155 160
 Gly Ala Gly Phe Asn Thr Phe Arg Ile Thr Phe Leu Met Glu Arg Met
 165 170 175
 Ser Pro Pro Ala Thr Gly Leu Thr Gly Pro Phe Asn Gln Thr Tyr Leu
 180 185 190
 Ser Gly Leu Thr Thr Ile Val Asp Tyr Ile Thr Asn Lys Gly Gly Tyr
 195 200 205
 Ala Leu Ile Asp Pro His Asn Phe Met Arg Tyr Asn Asn Gly Ile Ile
 210 215 220
 Ser Ser Thr Ser Asp Phe Ala Thr Trp Trp Ser Asn Leu Ala Thr Val
 225 230 235 240
 Phe Lys Ser Thr Lys Asn Ala Ile Phe Asp Ile Gln Asn Glu Pro Tyr
 245 250 255
 Gly Ile Asp Ala Gln Thr Val Tyr Glu Leu Asn Gln Ala Ala Ile Asn
 260 265 270
 Ser Ile Arg Ala Ala Gly Ala Thr Ser Gln Leu Ile Leu Val Glu Gly
 275 280 285
 Thr Ser Tyr Thr Gly Ala Trp Thr Trp Val Ser Ser Gly Asn Gly Ala
 290 295 300
 Ala Phe Ala Ala Val Thr Asp Pro Tyr Asn Asn Thr Ala Ile Glu Met
 305 310 315 320
 His Gln Tyr Leu Asp Ser Asp Gly Ser Gly Thr Asn Glu Asp Cys Val
 325 330 335
 Ser Ser Thr Ile Gly Ser Gln Arg Leu Gln Ala Ala Thr Ala Trp Leu
 340 345 350
 Gln Gln Thr Gly Leu Lys Gly Phe Leu Gly Glu Thr Gly Ala Gly Ser
 355 360 365
 Asn Ser Gln Cys Ile Asp Ala Val Phe Asp Glu Leu Cys Tyr Met Gln
 370 375 380
 Gln Gln Gly Gly Ser Trp Ile Gly Ala Leu Trp Trp Ala Ala Gly Pro
 385 390 395 400
 Trp Trp Gly Thr Tyr Ile Tyr Ser Ile Glu Pro Pro Ser Gly Ala Ala
 405 410 415
 Ile Pro Glu Val Leu Pro Gln Gly Leu Ala Pro Phe Leu
 420 425

<210> 27
 <211> 1580
 <212> DNA

[0023]

<213> 土生梭孢壳(*Thielavia terrestris*)

<400> 27

```

agccccccgt tcaggcacac ttggcatcag atcagcttag cagcgccctgc acagcatgaa      60
gctctcgtag tcggccgcgc tggcggcact caccgcgacg gcgtctgcgc cccctctgcc      120
cacgacgcgc caggcgcga ggcaggttc agccggctgc tegtctgcgg tcacgtctga      180
cgccagcacc aacgtttgga agaagtacac gctgcacccc aacagctact accgcaagga      240
ggttgaggcc gcggtggcgc agatctogga ccgggaccte gccgccaagg ccaagaaggt      300
ggccgacgtc ggcaccttcc ttgtgctga ctogacgag aacatcgga agctggagcc      360
ggcgatccag gacgtgccct gcgagaacat cctgggctgt gtcctctac acctgccggg      420
ccgcgactgc gcgccaagg cgtccaacgg cgagetcaag gtggcgaga tcgaccgcta      480
caagaccgag tacatcgaca gtgagtgtg cccccgggt tcgagagag cgtgggggaa      540
agggaagggt ttgactgact gacacggcgc actgcagaga tcgtgtgat cctcaaggca      600
caccccaaca cggcgcttgc gctggtcac gagccggact cgtgcccac cctggtgacc      660
aacagcaact tggacacgtg ctgagcgc gcgtcgggt accggaagg cgtggcttac      720
gcccctaaga acctcaacct gcccacgtg atcatgtacc tcgacgccg ccacggggc      780
tggtctggtt gggacgcaa cctgcagccc ggcgcgcagg agctagcaa ggcgtacaag      840
aaegccggtt cgcgaagca gctccgggc ttctgacaa acgtggcgg ctggaaetcc      900
tggtgagctt ttctcattt catttctct tcttctctc tcttctctc cactctgag      960
ccccccctcc cccaagcacc cactggcgtt ccggttgtt gactcggcct cctttccccc      1020
gggcaccagg gateaatgc ccgggaatt ctccaggcg tcgacgcca agtacaacaa      1080
gtgccagAAC gagaagatct acgtcagcgc ctccgctcc gcgtccagt cggccggcat      1140
gcccaccac gccatcgtc acacgggcc caacggcgtc accgcctgc gcaaggagt      1200
gggtgactgg tgcaacgtca acggtgcagg ttctgtgtt tcttttctc ctcttttgtt      1260
tgcacgtcgt ggtcctttt aagcagcgtt gtttggttg gggagatgga ctccggtga      1320
tgltctgtct cctctctagg ctctggcgtg cgcgcgacga gcaacacggg cctcagctg      1380
gccgacgcgt tcgtgtgggt caagcccgcc ggcgagtcgg acggcaccag cgacagctcg      1440
tcgcgcgctt acgacagctt ctgcggcaag gacgacgcct tcaagccctc gcccgaggcc      1500
ggcaccttga acgaggccta cttcgagatg ctgctcaaga acgctgtgc gtcgttctaa      1560
gacggtccag cactcatcgg      1580

```

<210> 28

<211> 396

<212> PRT

<213> 土生梭孢壳(*Thielavia terrestris*)

<400> 28

```

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

```

[0024]

Lys Lys Tyr Thr Leu His Pro Asn Ser Tyr Tyr Arg Lys Glu Val Glu
 50 55 60
 Ala Ala Val Ala Gln Ile Ser Asp Pro Asp Leu Ala Ala Lys Ala Lys
 65 70 75 80
 Lys Val Ala Asp Val Gly Thr Phe Leu Trp Leu Asp Ser Ile Glu Asn
 85 90 95
 Ile Gly Lys Leu Glu Pro Ala Ile Gln Asp Val Pro Cys Glu Asn Ile
 100 105 110
 Leu Gly Leu Val Ile Tyr Asp Leu Pro Gly Arg Asp Cys Ala Ala Lys
 115 120 125
 Ala Ser Asn Gly Glu Leu Lys Val Gly Glu Ile Asp Arg Tyr Lys Thr
 130 135 140
 Glu Tyr Ile Asp Lys Ile Val Ser Ile Leu Lys Ala His Pro Asn Thr
 145 150 155 160
 Ala Phe Ala Leu Val Ile Glu Pro Asp Ser Leu Pro Asn Leu Val Thr
 165 170 175
 Asn Ser Asn Leu Asp Thr Cys Ser Ser Ser Ala Ser Gly Tyr Arg Glu
 180 185 190
 Gly Val Ala Tyr Ala Leu Lys Asn Leu Asn Leu Pro Asn Val Ile Met
 195 200 205
 Tyr Leu Asp Ala Gly His Gly Gly Trp Leu Gly Trp Asp Ala Asn Leu
 210 215 220
 Gln Pro Gly Ala Gln Glu Leu Ala Lys Ala Tyr Lys Asn Ala Gly Ser
 225 230 235 240
 Pro Lys Gln Leu Arg Gly Phe Ser Thr Asn Val Ala Gly Trp Asn Ser
 245 250 255
 Trp Asp Gln Ser Pro Gly Glu Phe Ser Gln Ala Ser Asp Ala Lys Tyr
 260 265 270
 Asn Lys Cys Gln Asn Glu Lys Ile Tyr Val Ser Thr Phe Gly Ser Ala
 275 280 285
 Leu Gln Ser Ala Gly Met Pro Asn His Ala Ile Val Asp Thr Gly Arg
 290 295 300
 Asn Gly Val Thr Gly Leu Arg Lys Glu Trp Gly Asp Trp Cys Asn Val
 305 310 315 320
 Asn Gly Ala Gly Phe Gly Val Arg Pro Thr Ser Asn Thr Gly Leu Glu
 325 330 335
 Leu Ala Asp Ala Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly
 340 345 350
 Thr Ser Asp Ser Ser Ser Pro Arg Tyr Asp Ser Phe Cys Gly Lys Asp
 355 360 365
 Asp Ala Phe Lys Pro Ser Pro Glu Ala Gly Thr Trp Asn Glu Ala Tyr
 370 375 380
 Phe Glu Met Leu Leu Lys Asn Ala Val Pro Ser Phe
 385 390 395

<210> 29

<211> 1203

<212> DNA

<213> 土生梭孢壳(Thielavia terrestris)

[0025]

<400> 29
atgaagtacc tcaacctcct cgcagctctc ctgcgcgtcg ctctctctct cctcgtctga 60
cccagcatcg aggccagaca gtcgaaagtc aaccataca tcggcaagag cccgctcggt 120
attaggtcgt acgccccaaa gcttgaggag accgtcagga ccttccagca acgtggcgac 180
cagctcaacg ctgcgaggac aeggacggtg cagaaegtig cgaacttctc ctggatctcg 240
gataccaatg gtattggagc cattcagcct ctcatccaag atgtctctcg ccagcaggt 300
cgcactggac agaaggtcat cgtccaaatc gtctcttaca acctcccaga tcgcgactgc 360
tgttccaaag cctcgactgg agagttcacc gtaggaaacg acgtctctaa ccgatacaag 420
aactttgtca acaccatctc ccgcgagctc tcgaactgtg acgtgacaa gtccacttt 480
gcccctctcc tcgaaccga cgcacttgcc aacctcttca ccaacgcgaa tgcctccagg 540
tgccgaatcg ccgtctccgc ttacaaggag ggtatcgctt acacctctc cactttgtcc 600
aagcccaacg tcgacgtcta catcgacgcc gccaacggtg gctgctctgg ctggaacgac 660
aacctccgcc ccttcccgca actcttcaag gaagctcacc acctcgcgc ccgcataac 720
cccaacgcca aggtcccgcg cgtcccgctc aacgtctcca actacaacca gtaccgcgt 780
gaagtcgcgc agccttctac cgagtggag gacgcctggg acgagagcgc ctacgtcaac 840
gtctcacc ccgaacctca cgcgctcggc ttctcgcgc acttcatctg tgaccagga 900
cgcggtggca agggcggtat caggacggag tggggccagt ggtgcaact taggaacgt 960
gggttcggta tcaggectac tcggatcag ggcgtgctcc agaaccgaa tgtggatcg 1020
attgtgtggg ttaagccggg tggagagtcg gatggcacga gtgattgaa ctggaacagg 1080
tatgacctta cgtgcaggag tcgggtggcg catgttcccg ctccagaggc tggccagtgg 1140
ttcaacgagt atgttgttaa cctcgttttg aacgtcaacc cccctcttga gcctacctgg 1200
taa 1203

<210> 30
<211> 400
<212> PRT
<213> 土生梭孢壳(*Thielavia terrestris*)

<400> 30
Met Lys Tyr Leu Asn Leu Leu Ala Ala Leu Leu Ala Val Ala Pro Leu
1 5 10 15
Ser Leu Ala Ala Pro Ser Ile Glu Ala Arg Gln Ser Asn Val Asn Pro
20 25 30
Tyr Ile Gly Lys Ser Pro Leu Val Ile Arg Ser Tyr Ala Gln Lys Leu
35 40 45
Glu Glu Thr Val Arg Thr Phe Gln Gln Arg Gly Asp Gln Leu Asn Ala
50 55 60
Ala Arg Thr Arg Thr Val Gln Asn Val Ala Thr Phe Ala Trp Ile Ser
65 70 75 80
Asp Thr Asn Gly Ile Gly Ala Ile Arg Pro Leu Ile Gln Asp Ala Leu
85 90 95
Ala Gln Gln Ala Arg Thr Gly Gln Lys Val Ile Val Gln Ile Val Val
100 105 110
Tyr Asn Leu Pro Asp Arg Asp Cys Ser Ala Asn Ala Ser Thr Gly Glu

[0026]

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

<210> 31
 <211> 1501
 <212> DNA
 <213> 土生梭孢壳(Thielavia terrestris)

<400> 31	
gccgttgtea agatgggcca gaagacgtcg cacggattcg ccgccacggc fttggccgtt	60
ctcccctttg tgaaggctea gcagcccggc aacttcacgc cggaggtgca cccgcaactg	120
ccaacgtgga agtgcacgac cgcggcgggc tgcgttcagc aggacacttc ggtggtgctc	180
gaetggaact accgttgat ccacaatgcc gaeggeacgc cctcgtgcac gacgtccage	240
ggggtcgacc acacgtgtg tccagatgag gcgacctgcg cgaagaactg cttcgtgga	300
ggcgteaact acacgagcag cgggtgcacc acatccggca gticgtgcac gatgaggcag	360
tatttcaagg ggagcaacgg gcagaccaac agcgtttcgc ctcgtctcta cctgctcggc	420

[0027]

```

tcggatggaa actacgtaat gctcaagctg ctcggccagg agctgagctt cgatgtcgat      480
ctctccacgc tcccttgcgg cgagaacggc gcgctgtacc tgtccgagat ggacgcgacc      540
ggtggcagga accagtacaa caccggcggc gccaaactacg gctcgggcta ctgtgacgcc      600
cagtgctccc tgacagcgtg gatgaacggc acgctgaaca ccaacgggca gggctaactgc      660
tgcaacgaga tggacatcct cgaggccaac tcccgcgcca acgcgatgac acctcacccc      720
tgcgccaacg gcagctggga caagagcggg tgcggactca acccttaccg cgagggtac      780
aagagctact acggaccggg cctcacggtt gacacgtcga agcccttcac catcattacc      840
cgcttcaica ccgacgacgg cagcaccagc ggcacctca accagatcca gcggatctat      900
gtgcagaatg gcaagacggt cgcgtcggct gcgtccggag gcgacatcat cacggcatcc      960
ggctgaacct cggcccaggc gtccggcggg ctggccaaca tgggcggcgc gcttggacgg     1020
ggcatggtgc tgaccttcag catctggaac gacgtggggg gctacatgaa ctggtctgac     1080
agcggcaaca acggcccggt cagcagcacc gagggaacc cgtccaacat cctggccaac     1140
tacccgaca cccacgtggt ctctccaac atccgtggg gagacatcgg ctgcacggtc     1200
caggtctcgg gaggcggaac cggcggtcgc accaccacca cgtcgaccac cacgtgagg     1260
acctcgacea cgaccaccac caccgcccgc acggccaatg ccacgaactg gggacaatgc     1320
ggcggaatcg ggttacgta accgctcctt gcattctgtt gaggaagtta actaacgtgg     1380
cctacgcagt ggactggacc gaccgtctgc gaatgccgtt acgcatgcaa ggagctgaac     1440
ccctggtact accagtgctt ctaaagtatt gcagigaagc catactccgt gctcggcatg     1500
g                                                                           1501

```

<210> 32
 <211> 464
 <212> PRT
 <213> 土生梭孢壳(*Thielavia terrestris*)
 <400> 32

```

Met Gly Gln Lys Thr Leu His Gly Phe Ala Ala Thr Ala Leu Ala Val
1           5           10           15

Leu Pro Phe Val Lys Ala Gln Gln Pro Gly Asn Phe Thr Pro Glu Val
20          25          30

His Pro Gln Leu Pro Thr Trp Lys Cys Thr Thr Ala Gly Gly Cys Val
35          40          45

Gln Gln Asp Thr Ser Val Val Leu Asp Trp Asn Tyr Arg Trp Ile His
50          55          60

Asn Ala Asp Gly Thr Ala Ser Cys Thr Thr Ser Ser Gly Val Asp His
65          70          75          80

Thr Leu Cys Pro Asp Glu Ala Thr Cys Ala Lys Asn Cys Phe Val Glu
85          90          95

Gly Val Asn Tyr Thr Ser Ser Gly Val Thr Thr Ser Gly Ser Ser Leu
100         105         110

Thr Met Arg Gln Tyr Phe Lys Gly Ser Asn Gly Gln Thr Asn Ser Val
115         120         125

Ser Pro Arg Leu Tyr Leu Leu Gly Ser Asp Gly Asn Tyr Val Met Leu
130         135         140

```

[0028]

Lys Leu Leu Gly Gln Glu Leu Ser Phe Asp Val Asp Leu Ser Thr Leu
 145 150 155 160
 Pro Cys Gly Glu Asn Gly Ala Leu Tyr Leu Ser Glu Met Asp Ala Thr
 165 170 175
 Gly Gly Arg Asn Gln Tyr Asn Thr Gly Gly Ala Asn Tyr Gly Ser Gly
 180 185 190
 Tyr Cys Asp Ala Gln Cys Pro Val Gln Thr Trp Met Asn Gly Thr Leu
 195 200 205
 Asn Thr Asn Gly Gln Gly Tyr Cys Cys Asn Glu Met Asp Ile Leu Glu
 210 215 220
 Ala Asn Ser Arg Ala Asn Ala Met Thr Pro His Pro Cys Ala Asn Gly
 225 230 235 240
 Ser Cys Asp Lys Ser Gly Cys Gly Leu Asn Pro Tyr Ala Glu Gly Tyr
 245 250 255
 Lys Ser Tyr Tyr Gly Pro Gly Leu Thr Val Asp Thr Ser Lys Pro Phe
 260 265 270
 Thr Ile Ile Thr Arg Phe Ile Thr Asp Asp Gly Thr Thr Ser Gly Thr
 275 280 285
 Leu Asn Gln Ile Gln Arg Ile Tyr Val Gln Asn Gly Lys Thr Val Ala
 290 295 300
 Ser Ala Ala Ser Gly Gly Asp Ile Ile Thr Ala Ser Gly Cys Thr Ser
 305 310 315 320
 Ala Gln Ala Phe Gly Gly Leu Ala Asn Met Gly Ala Ala Leu Gly Arg
 325 330 335
 Gly Met Val Leu Thr Phe Ser Ile Trp Asn Asp Ala Gly Gly Tyr Met
 340 345 350
 Asn Trp Leu Asp Ser Gly Asn Asn Gly Pro Cys Ser Ser Thr Glu Gly
 355 360 365
 Asn Pro Ser Asn Ile Leu Ala Asn Tyr Pro Asp Thr His Val Val Phe
 370 375 380
 Ser Asn Ile Arg Trp Gly Asp Ile Gly Ser Thr Val Gln Val Ser Gly
 385 390 395 400
 Gly Gly Asn Gly Gly Ser Thr Thr Thr Thr Ser Thr Thr Thr Leu Arg
 405 410 415
 Thr Ser Thr Thr Thr Thr Thr Thr Ala Pro Thr Ala Thr Ala Thr His
 420 425 430
 Trp Gly Gln Cys Gly Gly Ile Gly Trp Thr Gly Pro Thr Val Cys Glu
 435 440 445
 Ser Pro Tyr Ala Cys Lys Glu Leu Asn Pro Trp Tyr Tyr Gln Cys Leu
 450 455 460

<210> 33

<211> 1368

<212> DNA

<213> 土生梭孢壳(Thielavia terrestris)

<400> 33

accgatacgc tcgaagatgg cgcccaagtc tacagttctg gccgcctggc tgctctcctc 60

gctggccgcg gccagcaga tcggcaaagc cgtgcccgag gtccacccca aactgacaac 120

gcagaagtgc actctccgcg ggggtgcaa gccgtccgc aectcggctg tgctcgactc 180

[0029]

```

gtccgcgcgc tcgctgcaca aggtcgggga ccccaacacc agctgcagcg tcggcggcga      240
cctgtgctcg gacgcgaagt cgtgcggcaa gaactgcgcg ctcgaggcg tcgactacgc      300
ggcccacggc gtggcgacca agggcgacgc cctcacgtg caccagtggc tcaagggggc      360
cgacggcacc tacaggaccg tctcgcgcg egtatacctc ctgggcgagg acgggaagaa      420
ctacaggac ttcaagctgc tcaacgcga gctcagcttc gacgtcgaeg tgteccagct      480
cgtctgcggc atgaacggcg cctgtactt ctcgagatg gagatggacg gcggcgcag      540
cccgtgaac ccggcggcg ccacgtacgg caccggctac tgcgacgcg agtgccecaa      600
gttggacttt atcaacggcg aggtattict tctctcttct gttttcttt tccatcgctt      660
tttctgaccg gaatccgcc tcttagctca acaccaacca cactacggg gcgtgctgca      720
acgagatgga catctgggag gccaacgcgc tggcgaggc gctcacgcg caccctgca      780
acgcgacgcg gggttacaag tgcgacacgg cggacgagt cgggcagccg gtggcggtg      840
gcgacgaatg ggggtgcfcg tacaaccctt ccaacttcgg ggtcaaggac tactacggc      900
gcaacctgac ggtagacacg aaccgcaagt tcacggtag gacgcagttc gtgacgtcca      960
acgggcgggc ggacggcgag ctgaccgaga tccggcggt gtacgtgcag gacggcgtag      1020
tgatccagaa ccacgcggt acggcggcg ggcgcacgta cgacagcatc acggacggt      1080
tctgcaacgc gacggccacc tggacgcag agcggggcgg gctcgcgcg atgggcgagg      1140
ccatcgccgc cggcatggg ctcacttcca gctgtgggt tgacaacggc ggcttcatga      1200
actggctega cagcgcaac gccggccct gcaacgccac cgaggcgac ccggccctga      1260
tcttcgaca gcaccggac gccagcgtca cttctecaa cctccgatg ggcgagateg      1320
gcagcacgta caagagcgag tgcagccact agagtagagc ttgtaatt      1368

```

<210> 34
 <211> 423
 <212> PRT
 <213> 土生梭孢壳(*Thielavia terrestris*)

<400> 34

```

Met Ala Pro Lys Ser Thr Val Leu Ala Ala Trp Leu Leu Ser Ser Leu
1           5           10          15
Ala Ala Ala Gln Gln Ile Gly Lys Ala Val Pro Glu Val His Pro Lys
20          25          30
Leu Thr Thr Gln Lys Cys Thr Leu Arg Gly Gly Cys Lys Pro Val Arg
35          40          45
Thr Ser Val Val Leu Asp Ser Ser Ala Arg Ser Leu His Lys Val Gly
50          55          60
Asp Pro Asn Thr Ser Cys Ser Val Gly Gly Asp Leu Cys Ser Asp Ala
65          70          75          80
Lys Ser Cys Gly Lys Asn Cys Ala Leu Glu Gly Val Asp Tyr Ala Ala
85          90          95
His Gly Val Ala Thr Lys Gly Asp Ala Leu Thr Leu His Gln Trp Leu
100         105         110
Lys Gly Ala Asp Gly Thr Tyr Arg Thr Val Ser Pro Arg Val Tyr Leu
115         120         125

```

[0030]

Leu Gly Glu Asp Gly Lys Asn Tyr Glu Asp Phe Lys Leu Leu Asn Ala
 130 135 140
 Glu Leu Ser Phe Asp Val Asp Val Ser Gln Leu Val Cys Gly Met Asn
 145 150 155 160
 Gly Ala Leu Tyr Phe Ser Glu Met Glu Met Asp Gly Gly Arg Ser Pro
 165 170 175
 Leu Asn Pro Ala Gly Ala Thr Tyr Gly Thr Gly Tyr Cys Asp Ala Gln
 180 185 190
 Cys Pro Lys Leu Asp Phe Ile Asn Gly Glu Leu Asn Thr Asn His Thr
 195 200 205
 Tyr Gly Ala Cys Cys Asn Glu Met Asp Ile Trp Glu Ala Asn Ala Leu
 210 215 220
 Ala Gln Ala Leu Thr Pro His Pro Cys Asn Ala Thr Arg Val Tyr Lys
 225 230 235 240
 Cys Asp Thr Ala Asp Glu Cys Gly Gln Pro Val Gly Val Cys Asp Glu
 245 250 255
 Trp Gly Cys Ser Tyr Asn Pro Ser Asn Phe Gly Val Lys Asp Tyr Tyr
 260 265 270
 Gly Arg Asn Leu Thr Val Asp Thr Asn Arg Lys Phe Thr Val Thr Thr
 275 280 285
 Gln Phe Val Thr Ser Asn Gly Arg Ala Asp Gly Glu Leu Thr Glu Ile
 290 295 300
 Arg Arg Leu Tyr Val Gln Asp Gly Val Val Ile Gln Asn His Ala Val
 305 310 315 320
 Thr Ala Gly Gly Ala Thr Tyr Asp Ser Ile Thr Asp Gly Phe Cys Asn
 325 330 335
 Ala Thr Ala Thr Trp Thr Gln Gln Arg Gly Gly Leu Ala Arg Met Gly
 340 345 350
 Glu Ala Ile Gly Arg Gly Met Val Leu Ile Phe Ser Leu Trp Val Asp
 355 360 365
 Asn Gly Gly Phe Met Asn Trp Leu Asp Ser Gly Asn Ala Gly Pro Cys
 370 375 380
 Asn Ala Thr Glu Gly Asp Pro Ala Leu Ile Leu Gln Gln His Pro Asp
 385 390 395 400
 Ala Ser Val Thr Phe Ser Asn Ile Arg Trp Gly Glu Ile Gly Ser Thr
 405 410 415
 Tyr Lys Ser Glu Cys Ser His
 420

<210> 35
 <211> 1000
 <212> DNA
 <213> 土生梭孢壳(Thielavia terrestris)

<400> 35
 atgacctac ggctccctgt catcagcctg ctggcctcgc tggcagcagg cgccgtcgtc 60
 gteccacggg cggagtttca cccccctctc ccgaattgga aatgcaegac ctccgggggc 120
 tgctgcagc agaacaccag cgtcgtcctg gaccgtgact cgaagtacgc cgcacacagc 180
 gccggetcgc ggacggaate ggattacgcg gcaatgggag tgtccacttc gggcaatgcc 240
 gtgacgtctg accactacgt caagaccaac ggcaccctcg tcccccttc gccgcgcac 300

[0031]

tacctcctgg gcgcggacgg caagtaactg ctatggacc tctcaacca ggagctgtcg 360
 gtggacgtcg acttctcgge gtgcccgtgc ggcgagaacg gggcctteta cctgtccgag 420
 atggcggcgg acggggcggg egacgggggg gcgggcgacg ggtactgega cgcgcagtgc 480
 cagggetact gctgcaacga gatggacatc ctcgaggeca actcgatggc gacggccatg 540
 acgcgcgacc cgtgcaaggg caacaactgc gaccgcagcg gctgcggcta caaccgtag 600
 gccagcgccc agcgcggctt ctacgggccc ggcaagacgg tcgacacgag caagcccttc 660
 accgtcgtca cgcagttcgc cgccagcggc ggcaagctga ccagatcac ccgcaagtac 720
 atccagaacg gccgggagat cggcggcggc ggcaceatct ccagctgcgg ctccgagtct 780
 tcgacggcgg gcttgaccgg catgggcgag gcgctggggc gcggaatggt gctggccatg 840
 agcatctgga acgacgcggc ccaggagatg gcatggctcg atcccgcaa caacggccct 900
 tgcgccagtg gccagggcag cccgtccgtc attcagtcgc agcatccga caccacgtc 960
 gtcttctcca acatcaggtg gggcgacatc ggggtctacca 1000

<210> 36
 <211> 336
 <212> PRT
 <213> 土生梭孢壳(Thielavia terrestris)

<400> 36

Met Thr Leu Arg Leu Pro Val Ile Ser Leu Leu Ala Ser Leu Ala Ala
 1 5 10 15
 Gly Ala Val Val Val Pro Arg Ala Glu Phe His Pro Pro Leu Pro Thr
 20 25 30
 Trp Lys Cys Thr Thr Ser Gly Gly Cys Val Gln Gln Asn Thr Ser Val
 35 40 45
 Val Leu Asp Arg Asp Ser Lys Tyr Ala Ala His Ser Ala Gly Ser Arg
 50 55 60
 Thr Glu Ser Asp Tyr Ala Ala Met Gly Val Ser Thr Ser Gly Asn Ala
 65 70 75 80
 Val Thr Leu Tyr His Tyr Val Lys Thr Asn Gly Thr Leu Val Pro Ala
 85 90 95
 Ser Pro Arg Ile Tyr Leu Leu Gly Ala Asp Gly Lys Tyr Val Leu Met
 100 105 110
 Asp Leu Leu Asn Gln Glu Leu Ser Val Asp Val Asp Phe Ser Ala Leu
 115 120 125
 Pro Cys Gly Glu Asn Gly Ala Phe Tyr Leu Ser Glu Met Ala Ala Asp
 130 135 140
 Gly Arg Gly Asp Ala Gly Ala Gly Asp Gly Tyr Cys Asp Ala Gln Cys
 145 150 155 160
 Gln Gly Tyr Cys Cys Asn Glu Met Asp Ile Leu Glu Ala Asn Ser Met
 165 170 175
 Ala Thr Ala Met Thr Pro His Pro Cys Lys Gly Asn Asn Cys Asp Arg
 180 185 190
 Ser Gly Cys Gly Tyr Asn Pro Tyr Ala Ser Gly Gln Arg Gly Phe Tyr
 195 200 205
 Gly Pro Gly Lys Thr Val Asp Thr Ser Lys Pro Phe Thr Val Val Thr

[0032]

210	215	220
Gln Phe Ala Ala Ser Gly Gly Lys Leu Thr Gln Ile Thr Arg Lys Tyr 225 230 235 240		
Ile Gln Asn Gly Arg Glu Ile Gly Gly Gly Gly Thr Ile Ser Ser Cys 245 250 255		
Gly Ser Glu Ser Ser Thr Gly Gly Leu Thr Gly Met Gly Glu Ala Leu 260 265 270		
Gly Arg Gly Met Val Leu Ala Met Ser Ile Trp Asn Asp Ala Ala Gln 275 280 285		
Glu Met Ala Trp Leu Asp Ala Gly Asn Asn Gly Pro Cys Ala Ser Gly 290 295 300		
Gln Gly Ser Pro Ser Val Ile Gln Ser Gln His Pro Asp Thr His Val 305 310 315 320		
Val Phe Ser Asn Ile Arg Trp Gly Asp Ile Gly Ser Thr Thr Lys Asn 325 330 335		
<210> 37		
<211> 1480		
<212> DNA		
<213> Cladorrhinum foecundissimum		
<400> 37		
gatccgaatt cctcctctcg ttcttttagtc acagaccaga catctgceca cgatgggtca	60	
caagttcgec ctccctaccc geetgcgcgc ctccctcgca tetgcccage agateggcac	120	
cgctgtcccc gagtctcacc ccaagcttcc caccaagcgc tgcaactctcg ccggtggetg	180	
ccagaccgtc gacacctcca tctctatcga cgccttccag cgtccctccc acaagatcgg	240	
cgaccttccc actccttggc tctgtggcgg ccctctctgc ccgacgcca agtctctgcg	300	
tgagaaatgc gcctctgagg gtgtcgacta tgcctctctgg ggcataaaga ccgagggcga	360	
cgcctaact ctcaaccagt gcatgcccga ccggcgcaac cctggccagt acaagacgac	420	
tactccccgt acttaccttg ttgttgagga cggcaagaac taaggaggatg tgaagctctt	480	
ggctaaggag atctcgtttg atccgatgt cagcaacctt cctgtcggca tgaacggtgc	540	
tttctaattg tetgagatgt tgatggatgg tggacgtggc gacctcaacc ctgtcgtgtc	600	
cgagtatggt accggttaact gtagtcgcca gtgcttcaag ttggatttca tcaacggcga	660	
ggccaacatc gacaaaaage acggcgcttg ctgcaacgaa atggacattt tgaatccaa	720	
ctcgcgcgcc aagaccttgc tccccaccc ctgcaacatc acgcaggctct acaagtgcga	780	
aggcgaagac gagtgcggcc agcccgctgg cgtgtgcgac aagtgggggt gcggttcaa	840	
cgagtacaaa tggggcgctg agtcttctca cggccggggc tgcagttcg ccctcgactc	900	
ctccaagaag ttacccgtca ccacgcagtt cctgaccgac aacggcaagg aggacggcgt	960	
cctgtctgag atccgccgat tgtggcacea ggatggcaag ctgatcaaga acaccgctat	1020	
ccaggttgag gagaactaca gcacggactc ggtgagcacc gagttctgcg agaagactgc	1080	
ttttttcacc atgcagcgcg gtgtctctca ggcatgggc gaggtctatc gtcgtggtat	1140	
ggtgtctggtt ttacgactct gggcggtatg ttccgggttt atgaactggt tggatgcgga	1200	
gggtaattgc cttgcagcg cgaetgaggg cgatccgaag gagattgtca agaataagcc	1260	
ggatgctagg gttacgttct caaacattag gatgtgtgag gttggtagca cgtatgctcc	1320	

[0033]

gggtgggaag tgcggtgta agagcagggt tgctaggggg cttaactgctt ctaaggggg 1380
 gtgtgaagag aggaggaggt gtgttgagg gtggagatg ataattgggc gagatgggt 1440
 agagcggggt ggttgatat gaatacgttg aattggatgt 1480

<210> 38
 <211> 440
 <212> PRT
 <213> Cladorrhinum foecundissimum

<400> 38

Met Val His Lys Phe Ala Leu Leu Thr Gly Leu Ala Ala Ser Leu Ala
 1 5 10 15
 Ser Ala Gln Gln Ile Gly Thr Val Val Pro Glu Ser His Pro Lys Leu
 20 25 30
 Pro Thr Lys Arg Cys Thr Leu Ala Gly Gly Cys Gln Thr Val Asp Thr
 35 40 45
 Ser Ile Val Ile Asp Ala Phe Gln Arg Pro Leu His Lys Ile Gly Asp
 50 55 60
 Pro Ser Thr Pro Cys Val Val Gly Gly Pro Leu Cys Pro Asp Ala Lys
 65 70 75 80
 Ser Cys Ala Glu Asn Cys Ala Leu Glu Gly Val Asp Tyr Ala Ser Trp
 85 90 95
 Gly Ile Lys Thr Glu Gly Asp Ala Leu Thr Leu Asn Gln Trp Met Pro
 100 105 110
 Asp Pro Ala Asn Pro Gly Gln Tyr Lys Thr Thr Thr Pro Arg Thr Tyr
 115 120 125
 Leu Val Ala Glu Asp Gly Lys Asn Tyr Glu Asp Val Lys Leu Leu Ala
 130 135 140
 Lys Glu Ile Ser Phe Asp Ala Asp Val Ser Asn Leu Pro Cys Gly Met
 145 150 155 160
 Asn Gly Ala Phe Tyr Leu Ser Glu Met Leu Met Asp Gly Gly Arg Gly
 165 170 175
 Asp Leu Asn Pro Ala Gly Ala Glu Tyr Gly Thr Gly Tyr Cys Asp Ala
 180 185 190
 Gln Cys Phe Lys Leu Asp Phe Ile Asn Gly Glu Ala Asn Ile Asp Gln
 195 200 205
 Lys His Gly Ala Cys Cys Asn Glu Met Asp Ile Phe Glu Ser Asn Ser
 210 215 220
 Arg Ala Lys Thr Phe Val Pro His Pro Cys Asn Ile Thr Gln Val Tyr
 225 230 235 240
 Lys Cys Glu Gly Glu Asp Glu Cys Gly Gln Pro Val Gly Val Cys Asp
 245 250 255
 Lys Trp Gly Cys Gly Phe Asn Glu Tyr Lys Trp Gly Val Glu Ser Phe
 260 265 270
 Tyr Gly Arg Gly Ser Gln Phe Ala Ile Asp Ser Ser Lys Lys Phe Thr
 275 280 285
 Val Thr Thr Gln Phe Leu Thr Asp Asn Gly Lys Glu Asp Gly Val Leu
 290 295 300
 Val Glu Ile Arg Arg Leu Trp His Gln Asp Gly Lys Leu Ile Lys Asn

[0034]

305	310	315	320
Thr Ala Ile Gln Val Glu Glu Asn Tyr Ser Thr Asp Ser Val Ser Thr			
	325	330	335
Glu Phe Cys Glu Lys Thr Ala Ser Phe Thr Met Gln Arg Gly Gly Leu			
	340	345	350
Lys Ala Met Gly Glu Ala Ile Gly Arg Gly Met Val Leu Val Phe Ser			
	355	360	365
Ile Trp Ala Asp Asp Ser Gly Phe Met Asn Trp Leu Asp Ala Glu Gly			
	370	375	380
Asn Gly Pro Cys Ser Ala Thr Glu Gly Asp Pro Lys Glu Ile Val Lys			
	385	390	395
Asn Lys Pro Asp Ala Arg Val Thr Phe Ser Asn Ile Arg Ile Gly Glu			
	405	410	415
Val Gly Ser Thr Tyr Ala Pro Gly Gly Lys Cys Gly Val Lys Ser Arg			
	420	425	430
Val Ala Arg Gly Leu Thr Ala Ser			
	435	440	
<210> 39			
<211> 1380			
<212> DNA			
<213> 里氏木霉(Trichoderma reesei)			
<400> 39			
atggcgccct cagttacaact gccgttgacc acggccatcc tggccattgc ccggtctgtc			60
gccgcccagc aaccgggtac cagcaccccc gaggtccatc ccaagttgac aacctacaag			120
tgtacaaagt cgggggggtg cgtggcccag gacacctcgg tggctcctga ctgggaactac			180
cgtctggatgc acgaacgcaa ctacaactcg tgcacctga acggcggcgt caacaccacg			240
ctctgccctg acgaggcgac ctgtggcaag aactgcttca tcgaggcgt cgactacgcc			300
gcctcggggc tcacgacctc gggcagcagc ctcccatga accagtacat gccacgacg			360
tctggcggt acagcagcgt ctctctcgg ctgtatctcc tggactctga cggtagtac			420
gtgatctga agctcaacgg ccaggagctg agcttcgac tcgacctc tgctctgccg			480
tgtggagaga acggctcgt ctacctgtct cagatggacg agaacgggg cgccaaccag			540
tataaacagg ccggtgccaa ctacgggagc ggctactcgc atgctcagtg ccccgctcag			600
acatggagga acggcaccct caactactag caccagggt tctgtgcaa cgagatggat			660
atcctggagg gcaactcgag ggcgaatgcc ttagccctc actcttgac ggccacggcc			720
tgcgactctg ccggttgagg cttaacccc tatggcagcg gctacaaaag ctactacggc			780
cccgagata ccgttgacac ctccaagacc ttaccatca tcaccagtt caacacggac			840
aacggctcgc cctcgggcaa ccttgtgagc atcaccgcga agtaccagca aaacggcgtc			900
gacatcccca gcgccagcc cggcggcgac accatctgt cctgcctgt cgcctcagcc			960
tacggcggtc tcgccaccat gggcaaggcc ctgagcagcg gcatgggtgt cgtgttcagc			1020
atttggaacg acaacagcca gtacatgaac tggctcgaca ggggcaacgc cggccctgtc			1080
agcagcaccg agggcaaccc atccaacatc ctggccaaca accccaacac gcacgtctgc			1140
ttctccaaca tcgctgggg agacattggg tctactacga actcgactgc gcccctgccc			1200

[0035]

ccgcctgcgt ccagcagcac gttttcgact acacggagga gctcgacgac ttcgagcagc 1260
 ccgagctgca cgcagaactca ctgggggcag tgcggtggca ttgggtacag cgggtgcaag 1320
 acgtgcacgt cgggcactac gtgccagtat agcaacgact actactcgca atgccttttag 1380

 <210> 40
 <211> 459
 <212> PRT
 <213> 里氏木霉(Trichoderma reesei)

 <400> 40
 Met Ala Pro Ser Val Thr Leu Pro Leu Thr Thr Ala Ile Leu Ala Ile
 1 5 10 15
 Ala Arg Leu Val Ala Ala Gln Gln Pro Gly Thr Ser Thr Pro Glu Val
 20 25 30
 His Pro Lys Leu Thr Thr Tyr Lys Cys Thr Lys Ser Gly Gly Cys Val
 35 40 45
 Ala Gln Asp Thr Ser Val Val Leu Asp Trp Asn Tyr Arg Trp Met His
 50 55 60
 Asp Ala Asn Tyr Asn Ser Cys Thr Val Asn Gly Gly Val Asn Thr Thr
 65 70 75 80
 Leu Cys Pro Asp Glu Ala Thr Cys Gly Lys Asn Cys Phe Ile Glu Gly
 85 90 95
 Val Asp Tyr Ala Ala Ser Gly Val Thr Thr Ser Gly Ser Ser Leu Thr
 100 105 110
 Met Asn Gln Tyr Met Pro Ser Ser Ser Gly Gly Tyr Ser Ser Val Ser
 115 120 125
 Pro Arg Leu Tyr Leu Leu Asp Ser Asp Gly Glu Tyr Val Met Leu Lys
 130 135 140
 Leu Asn Gly Gln Glu Leu Ser Phe Asp Val Asp Leu Ser Ala Leu Pro
 145 150 155 160
 Cys Gly Glu Asn Gly Ser Leu Tyr Leu Ser Gln Met Asp Glu Asn Gly
 165 170 175
 Gly Ala Asn Gln Tyr Asn Thr Ala Gly Ala Asn Tyr Gly Ser Gly Tyr
 180 185 190
 Cys Asp Ala Gln Cys Pro Val Gln Thr Trp Arg Asn Gly Thr Leu Asn
 195 200 205
 Thr Ser His Gln Gly Phe Cys Cys Asn Glu Met Asp Ile Leu Glu Gly
 210 215 220
 Asn Ser Arg Ala Asn Ala Leu Thr Pro His Ser Cys Thr Ala Thr Ala
 225 230 235 240
 Cys Asp Ser Ala Gly Cys Gly Phe Asn Pro Tyr Gly Ser Gly Tyr Lys
 245 250 255
 Ser Tyr Tyr Gly Pro Gly Asp Thr Val Asp Thr Ser Lys Thr Phe Thr
 260 265 270
 Ile Ile Thr Gln Phe Asn Thr Asp Asn Gly Ser Pro Ser Gly Asn Leu
 275 280 285
 Val Ser Ile Thr Arg Lys Tyr Gln Gln Asn Gly Val Asp Ile Pro Ser
 290 295 300
 Ala Gln Pro Gly Gly Asp Thr Ile Ser Ser Cys Pro Ser Ala Ser Ala
 305 310 315 320

[0036]

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

<210> 41
 <211> 1545
 <212> DNA
 <213> 里氏木霉(*Trichoderma reesei*)

<400> 41
 atgtatcgga agttggccgt catctcggcc ttcttggcca cagctcgtgc tcagtcggcc 60
 tgcactctcc aatcggagac tcaccgcct ctgacatggc agaaatgctc gtctgtgtggc 120
 acgtgcacac aacagacagg ctccgtggtc atcgacgcca actggcgctg gactcacgtc 180
 acgaacagca gcacgaactg ctacgatggc aacacttga gctcgaccct atgtcctgac 240
 aacgagacct gcgcgaagaa ctgctgtctg gacggtgccg cctacgcgtc cacttacgga 300
 gttaccagca gcggtaacag cctctccatt ggctttgtca cccagtctgc gcagaagaac 360
 gttagcgctc gcctttacct tatggcgagc gacaagaact accaggaatt caccctgctt 420
 ggcaacgagt tctctttcga tgttgatgtt tcgcagctgc cgtgcggcct gaacggagct 480
 ctctacttct gtctcatgga cgcggatggt ggcgtgagca agtatccac caacaccgtc 540
 ggcgccaaat acggcacggg gtactgtgac agccagtgtc ccccgatct gaagttcacc 600
 aatggccagg ccaacgttga ggcctgggag ccgtcatcca acaacgcgaa cactggcatt 660
 ggaggacacg gaagctgctg ctctgagatg gatctctggg aggcacaact catctccagc 720
 gctcttacc cccacccttg caagactgtc ggccaggaga tctgcagggt tgatgggtgc 780
 ggcggaactt actccgataa cagatatggc ggcacttgcg atcccgatgg ctgcgactgg 840
 aaccataacc gcctgggcaa caccagcttc tacggccctg gctcaagctt taccctcgat 900
 accaccaaga aattgaccgt tgtcacccag ttcgagaagt cgggtgccat caaccgatac 960
 tatgtccaga atggcgteac ttccagcag eccaacgccg agcttggtag ttactctggc 1020
 aacgagctca acgatgattia ctgcacagct gaggagcgag aattcggcgg atcctcttct 1080
 tcagacaagg gcggcctgac tcagtccaag aaggetacct ctggcggeat ggttctggtc 1140

[0037]

atgagtctgt gggatgatta ctacgccaac atgctgtggtc tggactccac ctaccgcaca 1200
 aacgagacct cctccacacc cgggtccgtg cggggaagct gctccaccag ctccggtgtc 1260
 cctgctcagg tcgaatctca gtctcccaac gccaaagtca ccttctccaa catcaagttc 1320
 ggacccattg gcagcaccgg caaccctagc ggcggcaacc ctcccggcgg aaaccgcctc 1380
 ggcaccacca ccaccgccg cccagccact accactggaa gtctctccgg acctaccag 1440
 tctcactacg gccagtgcgg cggtattggc tacagcgccc ccacggctg cggcagcggc 1500
 acaacttgcc aggtcctgaa cccttactac tctcagtgcg tgtaa 1545

<210> 42
 <211> 514
 <212> PRT
 <213> 里氏木霉(*Trichoderma reesei*)

<400> 42

Met Tyr Arg Lys Leu Ala Val Ile Ser Ala Phe Leu Ala Thr Ala Arg
 1 5 10 15
 Ala Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr
 20 25 30
 Trp Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser
 35 40 45
 Val Val Ile Asp Ala Asn Trp Arg Trp Thr His Ala Thr Asn Ser Ser
 50 55 60
 Thr Asn Cys Tyr Asp Gly Asn Thr Trp Ser Ser Thr Leu Cys Pro Asp
 65 70 75 80
 Asn Glu Thr Cys Ala Lys Asn Cys Cys Leu Asp Gly Ala Ala Tyr Ala
 85 90 95
 Ser Thr Tyr Gly Val Thr Thr Ser Gly Asn Ser Leu Ser Ile Gly Phe
 100 105 110
 Val Thr Gln Ser Ala Gln Lys Asn Val Gly Ala Arg Leu Tyr Leu Met
 115 120 125
 Ala Ser Asp Thr Thr Tyr Gln Glu Phe Thr Leu Leu Gly Asn Glu Phe
 130 135 140
 Ser Phe Asp Val Asp Val Ser Gln Leu Pro Cys Gly Leu Asn Gly Ala
 145 150 155 160
 Leu Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Val Ser Lys Tyr Pro
 165 170 175
 Thr Asn Thr Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp Ser Gln
 180 185 190
 Cys Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Val Glu Gly
 195 200 205
 Trp Glu Pro Ser Ser Asn Asn Ala Asn Thr Gly Ile Gly Gly His Gly
 210 215 220
 Ser Cys Cys Ser Glu Met Asp Ile Trp Glu Ala Asn Ser Ile Ser Glu
 225 230 235 240
 Ala Leu Thr Pro His Pro Cys Thr Thr Val Gly Gln Glu Ile Cys Glu
 245 250 255
 Gly Asp Gly Cys Gly Gly Thr Tyr Ser Asp Asn Arg Tyr Gly Gly Thr
 260 265 270

[0038]

Cys Asp Pro Asp Gly Cys Asp Trp Asn Pro Tyr Arg Leu Gly Asn Thr
 275 280 285
 Ser Phe Tyr Gly Pro Gly Ser Ser Phe Thr Leu Asp Thr Thr Lys Lys
 290 295 300
 Leu Thr Val Val Thr Gln Phe Glu Thr Ser Gly Ala Ile Asn Arg Tyr
 305 310 315 320
 Tyr Val Gln Asn Gly Val Thr Phe Gln Gln Pro Asn Ala Glu Leu Gly
 325 330 335
 Ser Tyr Ser Gly Asn Glu Leu Asn Asp Asp Tyr Cys Thr Ala Glu Glu
 340 345 350
 Ala Glu Phe Gly Gly Ser Ser Phe Ser Asp Lys Gly Gly Leu Thr Gln
 355 360 365
 Phe Lys Lys Ala Thr Ser Gly Gly Met Val Leu Val Met Ser Leu Trp
 370 375 380
 Asp Asp Tyr Tyr Ala Asn Met Leu Trp Leu Asp Ser Thr Tyr Pro Thr
 385 390 395 400
 Asn Glu Thr Ser Ser Thr Pro Gly Ala Val Arg Gly Ser Cys Ser Thr
 405 410 415
 Ser Ser Gly Val Pro Ala Gln Val Glu Ser Gln Ser Pro Asn Ala Lys
 420 425 430
 Val Thr Phe Ser Asn Ile Lys Phe Gly Pro Ile Gly Ser Thr Gly Asn
 435 440 445
 Pro Ser Gly Gly Asn Pro Pro Gly Gly Asn Pro Pro Gly Thr Thr Thr
 450 455 460
 Thr Arg Arg Pro Ala Thr Thr Thr Gly Ser Ser Pro Gly Pro Thr Gln
 465 470 475 480
 Ser His Tyr Gly Gln Cys Gly Gly Ile Gly Tyr Ser Gly Pro Thr Val
 485 490 495
 Cys Ala Ser Gly Thr Thr Cys Gln Val Leu Asn Pro Tyr Tyr Ser Gln
 500 505 510
 Cys Leu

<210> 43

<211> 1611

<212> DNA

<213> 里氏木霉(Trichoderma reesei)

<400> 43

atgattgtcg gcaftctcac cacgtgggt acgtggcca cactgcagc tagtgtcct 60
 ctagaggagc ggcaagcttg ctcaagcgtc tggttaattat gtgaaccctc tcaagagacc 120
 caaatactga gatatgtcaa ggggccaatg tgggtggccag aattggtcgg gtccgacttg 180
 ctgtgcttcc ggaagcacat gcgtctactc caacgactat tactcccagt gtcttccgg 240
 cgtcgaagc tcaagctgt ccacgcgcgc cgcgtcgacg acttctcgag tatccccac 300
 aacatccgg tcgagctccg cgacgectcc acctggttct actactacca gactacctcc 360
 agtcggatcg ggaaccgcta cgtattcagg caacctttt gttggggctc ctcttgggc 420
 caatgcatai tacgcctctg aagttagcag cctcgtatt cctagcttga ctggagccat 480
 ggccaetgct gcagcagctg tcgcaaaggt tccctctttt atgtggctgt aggtcctccc 540

[0039]

```

ggaaccaagg caatctgita ctgaaggctc atcattcact gcagagatac tcttgacaag      600
acccctctca tggagcaaac ctiggccgac atccgcaccg ccaacaagaa tggcggtaac      660
tatgccggac agtttggtgt gtatgacttg ccggatcgcg attgcgctgc ccttgccctg      720
aatggcgaat actctatigc cgaatggggc gtcgccaaat ataagaacta tatcgacacc      780
attcgtcaaa ttgtcgtgga atattccgat atccggaccc tcctgggttat tggatatagt      840
ttaaacacct gectcccccc ccccttccct tcctttcccg ccggcatctt gtcgttgtgc      900
taactatigt tccctcttcc agagccigac tctcttgcca acctgggtgac caacctcggt      960
actccaaagt gtgccaatgc tcagtcagcc taccttgagt gcatacaata cgccgtcaca     1020
cagctgaacc ttccaaatgt tgcgatgtat ttggacgctg gccatgcagg atggcctggc     1080
tggeccggca accaagaccc ggccgctcag ctatttgcaa atgtttacaa gaatgcacg     1140
tctccgagag ctcttcgagg attggcaacc aatgtcgcca actacaacgg gtggaacatt     1200
accagccccc catcgctacac gcaaggcaac gctgtctaca acgagaagct gtacatccac     1260
gctattggac gtctcttgc caatcacgga tggtecaacg ccttcttcat cactgatcaa     1320
ggtcgatcgg gaaagcagcc taccggacag caacagtggg gagactggig caatgtgatc     1380
ggcaccggat ttggtattcg cccatccgca aacactgggg actegttgct ggattcggtt     1440
gtctgggtca agccaggcgg cgagtgtgac ggcaccagcg acagcagtgc gccacgattt     1500
gactcccact gtgcgtccc agatgccttg caaccggcgc ctcaagctgg tgccttggtc     1560
caagcctact ttgtgcagct tctcacaac gcaaacccat cgttcctgta a              1611

```

<210> 44
 <211> 471
 <212> PRT
 <213> 里氏木霉(*Trichoderma reesei*)

<400> 44

```

Met Ile Val Gly Ile Leu Thr Thr Leu Ala Thr Leu Ala Thr Leu Ala
1           5           10           15

Ala Ser Val Pro Leu Glu Glu Arg Gln Ala Cys Ser Ser Val Trp Gly
                20           25           30

Gln Cys Gly Gly Gln Asn Trp Ser Gly Pro Thr Cys Cys Ala Ser Gly
          35           40           45

Ser Thr Cys Val Tyr Ser Asn Asp Tyr Tyr Ser Gln Cys Leu Pro Gly
          50           55           60

Ala Ala Ser Ser Ser Ser Ser Thr Arg Ala Ala Ser Thr Thr Ser Arg
65           70           75           80

Val Ser Pro Thr Thr Ser Arg Ser Ser Ser Ala Thr Pro Pro Pro Gly
          85           90           95

Ser Thr Thr Thr Arg Val Pro Pro Val Gly Ser Gly Thr Ala Thr Tyr
          100          105          110

Ser Gly Asn Pro Phe Val Gly Val Thr Pro Trp Ala Asn Ala Tyr Tyr
          115          120          125

Ala Ser Glu Val Ser Ser Leu Ala Ile Pro Ser Leu Thr Gly Ala Met
          130          135          140

Ala Thr Ala Ala Ala Ala Val Ala Lys Val Pro Ser Phe Met Trp Leu
145          150          155          160

```

[0040]

Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile
 165 170 175
 Arg Thr Ala Asn Lys Asn Gly Gly Asn Tyr Ala Gly Gln Phe Val Val
 180 185 190
 Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly Glu
 195 200 205
 Tyr Ser Ile Ala Asp Gly Gly Val Ala Lys Tyr Lys Asn Tyr Ile Asp
 210 215 220
 Thr Ile Arg Gln Ile Val Val Glu Tyr Ser Asp Ile Arg Thr Leu Leu
 225 230 235 240
 Val Ile Glu Pro Asp Ser Leu Ala Asn Leu Val Thr Asn Leu Gly Thr
 245 250 255
 Pro Lys Cys Ala Asn Ala Gln Ser Ala Tyr Leu Glu Cys Ile Asn Tyr
 260 265 270
 Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu Asp Ala
 275 280 285
 Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp Pro Ala Ala
 290 295 300
 Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro Arg Ala Leu
 305 310 315 320
 Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Gly Trp Asn Ile Thr
 325 330 335
 Ser Pro Pro Ser Tyr Thr Gln Gly Asn Ala Val Tyr Asn Glu Lys Leu
 340 345 350
 Tyr Ile His Ala Ile Gly Arg Leu Leu Ala Asn His Gly Trp Ser Asn
 355 360 365
 Ala Phe Phe Ile Thr Asp Gln Gly Arg Ser Gly Lys Gln Pro Thr Gly
 370 375 380
 Gln Gln Gln Trp Gly Asp Trp Cys Asn Val Ile Gly Thr Gly Phe Gly
 385 390 395 400
 Ile Arg Pro Ser Ala Asn Thr Gly Asp Ser Leu Leu Asp Ser Phe Val
 405 410 415
 Trp Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Ser Ser Ala
 420 425 430
 Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro Ala
 435 440 445
 Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu Thr
 450 455 460
 Asn Ala Asn Pro Ser Phe Leu
 465 470

<210> 45

<211> 2046

<212> DNA

<213> 特异腐质霉(*Humicola insolens*)

<400> 45

g c g t g a c c t t g c g c g c t t t g g g t g g c g g t g g c g a g t e g t g g a c g g t g c t t g c t g g t c g c 60

c g g c c t t c c c g g c g a t c c g c g t g a t g a g a g g g c c a c c a a t g g c g g g a t g a t g c t c c a t g g 120

g g a a c t t c c c a t g g a g a a g a g a g a a a c t t g c g g a g c c g t a t c t g g g a a a g a t g c t 180

[0041]

ccgtgtctcg tctatataac tctagttctc ccgagccctc aacaccacca gctctgatct	240
cacctcccc atcgacaatc acgcaaacac agcagttgtc gggccattcc ttcagacaca	300
tcagtcaccc tctttcaaaa tgcgtaccgc caagttcggc accctcgccg cctttgtggc	360
ctcgcccgcc gcccagcagg cgtgcagtct caccacccag aggcaccctt cctctctttg	420
gaacaagtgc accgccggcg gccagtgcga gaccgtccag gcttccatca ctctcgactc	480
caactggcgc tggactcacc aggtgtctgg etccaccaac tctacacgg genacaagtg	540
ggatacttag atctgcactg atgccaagtc gtgcgtctag aactgtctgc tcatgtgtgc	600
cgactacacc agcacctatg gcatcaccac caacggtgat tecttgagcc tcaagttcgt	660
caccaagggc cagcactega ccaacgtcgg ctccggtacc tacttgatgg acggcgagga	720
caagtatcag agtacgttct atcttcagcc ttctcgcgc ttgaatcctg gctaacgttt	780
acacttcaca gccctcgagc tcttcggcaa cgagttcacc ttcatgtctg atgtctccaa	840
catcggtctc ggtctcaacg gcgccttgta ctctgtctcc atggacgcgc atgggtgtct	900
cagccgctat cctggcaaca agcctgtgtc caagtacggt accggtact gcgatgctea	960
gtgccccgt gacatcaagt tcatcaacgg cagggccaac attgagggt ggaccggctc	1020
caccaacgac cccaacgcgc gcgcgggcgc ctatgttacc tctgtctctg agatggatat	1080
ctgggaagcc aacaacatgg ctactgcctt cactctcac ccttgccaca tcattggcca	1140
gagccgctgc gagggcgact cgigcgttgg cacttacagc aacgagcgt acgccggcgt	1200
ctgcgacccc gatggctctg acttcaactc gtaccgccag ggcaacaaga ccttctacgg	1260
caaggcatg accgtcgaca ccaccaagaa gatcactgtc gtcaccagc tcttcaagga	1320
tgccaacggc gatctcggcg agatcaacgc ctctacgtc caggatggca agatcatccc	1380
caactccgag tccaccatcc ccggcgtcga gggcaattcc atcaccagc actggtcga	1440
ccgccagaag gtgacctttg gcgacattga cgacttcaac cgcaaggcg gcatgaagca	1500
gatgggcaag gccctcgccg gccccaaggt cctggctatg tccatctggg atgaccacgc	1560
ctccaacatg ctctggctcg actcgacctt cctgtctgat gccgtggca agcccggcg	1620
cgagcgggt gcctcccca ccacctcggg tctccctgct gaggttgagg ccgaggcccc	1680
caacagcaac gtctgtttct ccaacatccg ctctggcccc atcggtctga ccgttctgtg	1740
tctccccggc ggggcaacg gcggcaacaa cggcggaac ccccccccc ccaccaccac	1800
cactctctcg gctccggcca ccaccaccac cgcagcgtt ggcccccaagg ctggccgctg	1860
gcagcagtc gccggcatcg gcttcaactg ccgaccagc tgcgaggagc cctacatttg	1920
caccaagctc aacgactggt actctcagtg cctgtaatt ctgagtcgt gactcgacga	1980
tcacggccgg tttttgatg aaaggaaaca aacgaccgcg ataaaaatgg agggtaatga	2040
gatgtc	2046

<210> 46
 <211> 525
 <212> PRT
 <213> 特异腐质霉(Humicola insolens)
 <400> 46

[0042]

```

Met Arg Thr Ala Lys Phe Ala Thr Leu Ala Ala Leu Val Ala Ser Ala
1           5           10           15
Ala Ala Gln Gln Ala Cys Ser Leu Thr Thr Glu Arg His Pro Ser Leu
20           25           30
Ser Trp Asn Lys Cys Thr Ala Gly Gly Gln Cys Gln Thr Val Gln Ala
35           40           45
Ser Ile Thr Leu Asp Ser Asn Trp Arg Trp Thr His Gln Val Ser Gly
50           55           60
Ser Thr Asn Cys Tyr Thr Gly Asn Lys Trp Asp Thr Ser Ile Cys Thr
65           70           75           80
Asp Ala Lys Ser Cys Ala Gln Asn Cys Cys Val Asp Gly Ala Asp Tyr
85           90           95
Thr Ser Thr Tyr Gly Ile Thr Thr Asn Gly Asp Ser Leu Ser Leu Lys
100          105          110
Phe Val Thr Lys Gly Gln His Ser Thr Asn Val Gly Ser Arg Thr Tyr
115          120          125
Leu Met Asp Gly Glu Asp Lys Tyr Gln Thr Phe Glu Leu Leu Gly Asn
130          135          140
Glu Phe Thr Phe Asp Val Asp Val Ser Asn Ile Gly Cys Gly Leu Asn
145          150          155          160
Gly Ala Leu Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Leu Ser Arg
165          170          175
Tyr Pro Gly Asn Lys Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp
180          185          190
Ala Gln Cys Pro Arg Asp Ile Lys Phe Ile Asn Gly Glu Ala Asn Ile
195          200          205
Glu Gly Trp Thr Gly Ser Thr Asn Asp Pro Asn Ala Gly Ala Gly Arg
210          215          220
Tyr Gly Thr Cys Cys Ser Glu Met Asp Ile Trp Glu Ala Asn Asn Met
225          230          235          240
Ala Thr Ala Phe Thr Pro His Pro Cys Thr Ile Ile Gly Gln Ser Arg
245          250          255
Cys Glu Gly Asp Ser Cys Gly Gly Thr Tyr Ser Asn Glu Arg Tyr Ala
260          265          270
Gly Val Cys Asp Pro Asp Gly Cys Asp Phe Asn Ser Tyr Arg Gln Gly
275          280          285
Asn Lys Thr Phe Tyr Gly Lys Gly Met Thr Val Asp Thr Thr Lys Lys
290          295          300
Ile Thr Val Val Thr Gln Phe Leu Lys Asp Ala Asn Gly Asp Leu Gly
305          310          315          320
Glu Ile Lys Arg Phe Tyr Val Gln Asp Gly Lys Ile Ile Pro Asn Ser
325          330          335
Glu Ser Thr Ile Pro Gly Val Glu Gly Asn Ser Ile Thr Gln Asp Trp
340          345          350
Cys Asp Arg Gln Lys Val Ala Phe Gly Asp Ile Asp Asp Phe Asn Arg
355          360          365
Lys Gly Gly Met Lys Gln Met Gly Lys Ala Leu Ala Gly Pro Met Val
370          375          380
Leu Val Met Ser Ile Trp Asp Asp His Ala Ser Asn Met Leu Trp Leu

```

[0043]

385	390	395	400
Asp Ser Thr Phe	Pro Val Asp Ala Ala Gly Lys	Pro Gly Ala Glu Arg	
	405	410	415
Gly Ala Cys Pro Thr Thr Ser Gly Val Pro Ala Glu Val Glu Ala Glu			
	420	425	430
Ala Pro Asn Ser Asn Val Val Phe Ser Asn Ile Arg Phe Gly Pro Ile			
	435	440	445
Gly Ser Thr Val Ala Gly Leu Pro Gly Ala Gly Asn Gly Gly Asn Asn			
	450	455	460
Gly Gly Asn Pro Pro Pro Pro Thr Thr Thr Thr Ser Ser Ala Pro Ala			
	465	470	475
Thr Thr Thr Thr Ala Ser Ala Gly Pro Lys Ala Gly Arg Trp Gln Gln			
	485	490	495
Cys Gly Gly Ile Gly Phe Thr Gly Pro Thr Gln Cys Glu Glu Pro Tyr			
	500	505	510
Ile Cys Thr Lys Leu Asn Asp Trp Tyr Ser Gln Cys Leu			
	515	520	525

<210> 47
 <211> 1812
 <212> DNA
 <213> 嗜热毁丝霉(Myceliophthora thermophila)

<400> 47
 atggccaaga agcttttcat caccgcgcgc cttgcggctg cagtgttggc ggcceccgctc 60
 attgaggagc gccagaactg cggcgcgtgtg tggtaagaaa gcccggtctg agtttcccat 120
 gacttttctc tccagtaatg gcataaggcc caccctctcg actgactgtg agaatcgatc 180
 aaatccagga ctcaatgcgg cggcaacggg tggcaggggc ccacatgctg cgcctcgggc 240
 tccacctgag ttgcgcagaa cagatgggtac tctcagtgcg tgcctcaacaa tcaggtagacg 300
 agttccaaca ctccgtcgtc gacttccacc tgcagcgcga gcagcagcac ctccagcagc 360
 agcaccagga gcggcagctc ctctctctcc accaccacgc cccctcccgt ctccagcccc 420
 gtgactagca ttcccgggcg tgcgaccacc acggcgagct actctggaac ccccttctcg 480
 ggcgctcggc tcttcgcca cgaactactac aggtccgagg tccacaatct cgccttctct 540
 agcatgaccg gfacctgtgc ggcgaaggct tccgcgctg cgaagtcctc tagcttccag 600
 tggctcgacc ggaacgtcac catcgacacc ctgatgggtc agactctgtc ccagatccgg 660
 gctgccaata atgcgggtgc caatctctcc tatgttgggt agttacatgg cggcgacttg 720
 ccttctcgtc cccacacctt ctgacggga tgggttacct gacctggagg caaaacaaaa 780
 ccagcccaac ttgtcgtcta cgaactcccc gaccgtgact gcgccgcgcg tgcgtccaac 840
 ggcgagtttt cgattgcaaa cggcggcgcc gccaaactaca ggagctacat cgacgetatc 900
 cgaagcaca tcattgagta ctccgacatc cggatcatcc tggttatcga gcccgactcg 960
 atggccaaca tggtagccaa catgaacgtg gccaaagtga gcaacgccgc gtcgacgtac 1020
 cagcagttga ccgtgtacgc gctcaagcag ctgaacctgc ccaacgtcgc catgtatctc 1080
 gacgccggcc acgcgggctg gctcggctgg cccgccaaca tccagccccg cggcgacctg 1140
 ttgcccggca tctacaatga cgcggccaag ccggtgccc tccggggcct ggccaactaac 1200

[0044]

```

gtcgccaact acaacgccg gagtatcgt tcggccccgt cglacacgt cctaaccct 1260
aactacgacg agaagcaeta catcgaggoc ttcageccgc tctgaacgc ggccggcttc 1320
cccgcacgct tcatgttgcg cactggccgc aacggcaaac aacctaccgg tatggttttt 1380
ttcttttttt ttctctgttc cctccccct tccccctcag ttggcgctca caaggctctt 1440
tagttttgct tcttctcgga ccaaccttcc cccaccccc aaacgcaccg cccacaaccg 1500
ttcgacteta tactcttggg aatgggcgcg gaaactgacc gttegacagg ccaacaacag 1560
tggggtgact ggtgcaatgt caagggcact ggetttggcg tgcgcccgc ggccaacacg 1620
ggccacgacc tggtcgatgc ctttgtctgg gtcaagcccg gcgctgagtc cgacggcaca 1680
agegacacca gcgcgcgcgc ctacgaactac cactgggcgc tctcgatgc cctgcagcct 1740
gtctcggagg ctggacagtg gtccagggc tacttcgagc agctgctcac caacgccaac 1800
ccgcccctct aa 1812

```

<210> 48
 <211> 482
 <212> PRT
 <213> 嗜热毁丝霉(*Myceliophthora thermophila*)

<400> 48

```

Met Ala Lys Lys Leu Phe Ile Thr Ala Ala Leu Ala Ala Val Leu
1      5      10     15
Ala Ala Pro Val Ile Glu Glu Arg Gln Asn Cys Gly Ala Val Trp Thr
20     25     30
Gln Cys Gly Gly Asn Gly Trp Gln Gly Pro Thr Cys Cys Ala Ser Gly
35     40     45
Ser Thr Cys Val Ala Gln Asn Glu Trp Tyr Ser Gln Cys Leu Pro Asn
50     55     60
Asn Gln Val Thr Ser Ser Asn Thr Pro Ser Ser Thr Ser Thr Ser Gln
65     70     75     80
Arg Ser Ser Ser Thr Ser Ser Ser Ser Thr Arg Ser Gly Ser Ser Ser
85     90     95
Ser Ser Thr Thr Thr Pro Pro Pro Val Ser Ser Pro Val Thr Ser Ile
100    105    110
Pro Gly Gly Ala Thr Thr Thr Ala Ser Tyr Ser Gly Asn Pro Phe Ser
115    120    125
Gly Val Arg Leu Phe Ala Asn Asp Tyr Tyr Arg Ser Glu Val His Asn
130    135    140
Leu Ala Ile Pro Ser Met Thr Gly Thr Leu Ala Ala Lys Ala Ser Ala
145    150    155    160
Val Ala Glu Val Pro Ser Phe Gln Trp Leu Asp Arg Asn Val Thr Ile
165    170    175
Asp Thr Leu Met Val Gln Thr Leu Ser Gln Ile Arg Ala Ala Asn Asn
180    185    190
Ala Gly Ala Asn Pro Pro Tyr Ala Ala Gln Leu Val Val Tyr Asp Leu
195    200    205
Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn Gly Glu Phe Ser Ile
210    215    220
Ala Asn Gly Gly Ala Ala Asn Tyr Arg Ser Tyr Ile Asp Ala Ile Arg

```

[0045]

225	230	235	240
Lys His Ile Ile Glu Tyr Ser Asp Ile Arg Ile Ile Leu Val Ile Glu			
	245	250	255
Pro Asp Ser Met Ala Asn Met Val Thr Asn Met Asn Val Ala Lys Cys			
	260	265	270
Ser Asn Ala Ala Ser Thr Tyr His Glu Leu Thr Val Tyr Ala Leu Lys			
	275	280	285
Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu Asp Ala Gly His Ala			
	290	295	300
Gly Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro Ala Ala Asp Leu Phe			
305	310	315	320
Ala Gly Ile Tyr Asn Asp Ala Gly Lys Pro Ala Ala Val Arg Gly Leu			
	325	330	335
Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser Ile Ala Ser Ala Pro			
	340	345	350
Ser Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu Lys His Tyr Ile Glu			
	355	360	365
Ala Phe Ser Pro Leu Leu Asn Ala Ala Gly Phe Pro Ala Arg Phe Ile			
	370	375	380
Val Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr Gly Gln Gln Gln Trp			
385	390	395	400
Gly Asp Trp Cys Asn Val Lys Gly Thr Gly Phe Gly Val Arg Pro Thr			
	405	410	415
Ala Asn Thr Gly His Asp Leu Val Asp Ala Phe Val Trp Val Lys Pro			
	420	425	430
Gly Gly Glu Ser Asp Gly Thr Ser Asp Thr Ser Ala Ala Arg Tyr Asp			
	435	440	445
Tyr His Cys Gly Leu Ser Asp Ala Leu Gln Pro Ala Pro Glu Ala Gly			
	450	455	460
Gln Trp Phe Gln Ala Tyr Phe Glu Gln Leu Leu Thr Asn Ala Asn Pro			
465	470	475	480

Pro Phe

<210> 49
 <211> 1725
 <212> DNA
 <213> 里氏木霉(Trichoderma reesei)

<400> 49	
gagggcagct cactgaaga gcttgtaag atcacctct gtgtattgca ccatgattgt	60
cggcattctc accacgctgg ctacgctgga cacactcgca gctagtgtgc ctctagagga	120
gcggcaagct tgcicaagcg tcggggcca atgtggtggc cagaatttgt egggtccgac	180
ttgctgtget tccggaagca catgegtcta ctccaagac tattactccc agtgtcttcc	240
cggcgctgca agctcaagct cgtccacgag cggcgctgag acgacttctc gagtatcccc	300
cacaacatcc eggtcagct ccgcgacgcc tccacctggt tetactacta ccagagtacc	360
tccagtcgga tcgggaaccg ctacgtattc aggcacccct tttgttgggg tcactccttg	420
ggccaatgca tattacgeet ctgaagttag cagctctgct attcctagct tgactggagc	480

[0046]

```

catggccact gctgcagcag ctgtcgcaaa ggttccctct tttatgtggc tagatactct 540
tgacaagacc cctctcatgg agcaaacctt gcccgacatc cgcaccgcca acaagaatgg 600
cggtaactat gccggacagt ttgtgggtga tgacttgccg gatcgcgatt gcgctgcctt 660
tgccctgaat ggcgaaatct ctattgccga tgggtggcgtc gccaaatata agaactatat 720
cgacaccatt cgtcaaatgg tctgtgaata ttccgataac cggaccctcc tggttattga 780
gcctgactct ctigccaacc tggtagacca cctcggtact ccaaagtgtg ccaatgctca 840
gtcagcctac cttagtgca tcaactacgc cgtcacacag ctgaaccttc caaatgttgc 900
gatgtatttg gacgttggcc atgcaggatg gcttggctgg ccggcaaacc aagaccgggc 960
cgctcagcta ttgtcaaatg ttacaagaa tgcctcgtct ccgagagctc ttccgggatt 1020
ggcaaccaat gtgcgaact acaacgggtg gaacattacc agcccccat cgtacacgca 1080
aggcaacgct gctacaacg agaagctgta catccacgct attggacctc ttcttgccaa 1140
tcacggctgg tccaacgctt tcttcatcac tgatcaaggt cgtcgggaa agcagcctac 1200
cggacagcaa cagtggggag actgtgtcaa tgtgatcggc accgatttg gtattcgccc 1260
atccgcaaac actggggact cgttgcgtga ttctttgtc tgggtcaagc caggcggcga 1320
gtgtgacgge accagcgaca gcagtgcgc acgatttgac tcccactgtg cgctccaga 1380
tgccctgcaa ccggcgctc aagctgggtc ttggttccaa gctactttg tgcagcttct 1440
cacaacgcaa aaccatcgt tctgtgaagg ctttcgtgac cgggttccaa acaatgaigt 1500
gcgatgtgtt ggttcccggt tggcggagtc ttgtctact ttggttgtct gtcgcaggtc 1560
ggtagaccgc aaatgagcaa ctgatggatt gttgccagcg atactataat tcacaiggt 1620
ggtctttgtc gatcagtacg tagtgagaga gagagaacat ctatccacaa tgtcaggtgt 1680
ctattagaca tactccgaga aaaaaaaaaa aaaaaaaaaa aaaaa 1725

```

<210> 50

<211> 471

<212> PRT

<213> 里氏木霉(Trichoderma reesei)

<400> 50

```

Met Ile Val Gly Ile Leu Thr Thr Leu Ala Thr Leu Ala Thr Leu Ala
1           5           10           15
Ala Ser Val Pro Leu Glu Glu Arg Gln Ala Cys Ser Ser Val Trp Gly
20           25           30
Gln Cys Gly Gly Gln Asn Trp Ser Gly Pro Thr Cys Cys Ala Ser Gly
35           40           45
Ser Thr Cys Val Tyr Ser Asn Asp Tyr Tyr Ser Gln Cys Leu Pro Gly
50           55           60
Ala Ala Ser Ser Ser Ser Ser Thr Arg Ala Ala Ser Thr Thr Ser Arg
65           70           75           80
Val Ser Pro Thr Thr Ser Arg Ser Ser Ser Ala Thr Pro Pro Pro Gly
85           90           95
Ser Thr Thr Thr Arg Val Pro Pro Val Gly Ser Gly Thr Ala Thr Tyr
100          105          110
Ser Gly Asn Pro Phe Val Gly Val Thr Pro Trp Ala Asn Ala Tyr Tyr
115          120          125

```

[0047]

Ala Ser Glu Val Ser Ser Leu Ala Ile Pro Ser Leu Thr Gly Ala Met
 130 135 140
 Ala Thr Ala Ala Ala Val Ala Lys Val Pro Ser Phe Met Trp Leu
 145 150 155 160
 Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile
 165 170 175
 Arg Thr Ala Asn Lys Asn Gly Gly Asn Tyr Ala Gly Gln Phe Val Val
 180 185 190
 Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly Glu
 195 200 205
 Tyr Ser Ile Ala Asp Gly Gly Val Ala Lys Tyr Lys Asn Tyr Ile Asp
 210 215 220
 Thr Ile Arg Gln Ile Val Val Glu Tyr Ser Asp Ile Arg Thr Leu Leu
 225 230 235 240
 Val Ile Glu Pro Asp Ser Leu Ala Asn Leu Val Thr Asn Leu Gly Thr
 245 250 255
 Pro Lys Cys Ala Asn Ala Gln Ser Ala Tyr Leu Glu Cys Ile Asn Tyr
 260 265 270
 Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu Asp Ala
 275 280 285
 Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp Pro Ala Ala
 290 295 300
 Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro Arg Ala Leu
 305 310 315 320
 Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Gly Trp Asn Ile Thr
 325 330 335
 Ser Pro Pro Ser Tyr Thr Gln Gly Asn Ala Val Tyr Asn Glu Lys Leu
 340 345 350
 Tyr Ile His Ala Ile Gly Pro Leu Leu Ala Asn His Gly Trp Ser Asn
 355 360 365
 Ala Phe Phe Ile Thr Asp Gln Gly Arg Ser Gly Lys Gln Pro Thr Gly
 370 375 380
 Gln Gln Gln Trp Gly Asp Trp Cys Asn Val Ile Gly Thr Gly Phe Gly
 385 390 395 400
 Ile Arg Pro Ser Ala Asn Thr Gly Asp Ser Leu Leu Asp Ser Phe Val
 405 410 415
 Trp Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Ser Ser Ala
 420 425 430
 Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro Ala
 435 440 445
 Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu Thr
 450 455 460
 Asn Ala Asn Pro Ser Phe Leu
 465 470

<210> 51
 <211> 1446
 <212> DNA
 <213> 土生梭孢壳(Thielavia terrestris)

[0048]

```

<400> 51
atggctcaga agctccttct cggcgccgcc ctggcggcca gcgccctcgc tgcctccgtc      60
gtcgaggagc gccagaactg cgtttccgtc tggagccaat gggcgggcat tggttggtcc      120
ggcgcgacct gctgcgttc ggccaatacc tgcgttgagc tgaaccgta ctactcgcag      180
tgcttgccca acagccaggt gactacctcg accagcaaga ccacctccac caccaccagg      240
agcagcaacca ccagccacag cagcgttccc accagcacga gcaccaccac caccagcagt      300
cccgtggta ctaaccggcc gactacctcc atcccggcg gtgcctcgtc aacggccagc      360
tggttcggca acccgttctc gggcgtgcag atgtgggcca acgactacta cgcctccgag      420
gtctcgtcgc tggcctccc cagcatgacg ggcgccatgg ccaccaaggc ggccgaggtg      480
gccaaagtgc ccagcttcca gtggctfagc cgcacgtca ccatcgacac gctgttcgcc      540
cacacgtgtg cgcagatccg cgcggccaac cagaaaggcg ccaaccgcc ctacgcgggc      600
atcttcgtgg tctacgacct tcggaccgc gactgcgcg ccgcccgctc caacggcgag      660
ttctccatcg cgaacaacgg ggcggccaac taagaagct acatcgacgc gatccggagc      720
ctcgtcatcc agtactcaga catccgcatc atcttcgtca tcgagccoga ctcgttcgcc      780
aacatggtga ccaacctgaa cgtggccaag tgcgccaag ccgagtcgac ctacaaggag      840
ttgaacgtct acgcgttcca gacgtgaac ctgcccacg tgcccatgta cctggacgcc      900
ggccaagccg gctggctcgg ctggcccgcc aacatccagc cggccgcca cctcttcgcc      960
gagatctaca cgagcgccgg caagccggcc gccgtgcgcg gcctcgccac caacgtggcc      1020
aactacaacg gctggagcct ggccacgcc cctctgtaac ccaggggcga ccccaactac      1080
gacgagagcc actacgtcca ggcctcgcgc ccgtgctca ccgccaacgg ctccccgcc      1140
cacctcatca ccgacaccgg ccgcaacggc aagcagccga ccgacaacg gcaatgggga      1200
gactggtgca acgttatcgg aactggcttc gccgtgcgcg cgacgacaaa caccggcctc      1260
gacatcgagg acgccttcgt ctgggtcaag ccgcgcgcg agtcgacgg cagagcaac      1320
acgacctctc cccgtacga ctaccactgc gccctgtcgg acgcgtgca gccgtctcgc      1380
gaggccggca cttggttcca ggcctacttc gacgactcc tgaccaacgc caaccgcccc      1440
ttttaa                                           1446

```

```

<210> 52
<211> 481
<212> PRT
<213> 土生梭孢壳(Thielavia terrestris)

```

```

<400> 52
Met Ala Gln Lys Leu Leu Leu Ala Ala Ala Leu Ala Ala Ser Ala Leu
1          5          10          15
Ala Ala Pro Val Val Glu Glu Arg Gln Asn Cys Gly Ser Val Trp Ser
20          25          30
Gln Cys Gly Gly Ile Gly Trp Ser Gly Ala Thr Cys Cys Ala Ser Gly
35          40          45
Asn Thr Cys Val Glu Leu Asn Pro Tyr Tyr Ser Gln Cys Leu Pro Asn
50          55          60
Ser Gln Val Thr Thr Ser Thr Ser Lys Thr Thr Ser Thr Thr Thr Arg
65          70          75          80

```

[0049]

Ser	Ser	Thr	Thr	Ser	His	Ser	Ser	Gly	Pro	Thr	Ser	Thr	Ser	Thr	Thr	
				85					90						95	
Thr	Thr	Ser	Ser	Pro	Val	Val	Thr	Thr	Pro	Pro	Ser	Thr	Ser	Ile	Pro	
				100				105						110		
Gly	Gly	Ala	Ser	Ser	Thr	Ala	Ser	Trp	Ser	Gly	Asn	Pro	Phe	Ser	Gly	
		115					120					125				
Val	Gln	Met	Trp	Ala	Asn	Asp	Tyr	Tyr	Ala	Ser	Glu	Val	Ser	Ser	Leu	
	130					135					140					
Ala	Ile	Pro	Ser	Met	Thr	Gly	Ala	Met	Ala	Thr	Lys	Ala	Ala	Glu	Val	
145					150					155					160	
Ala	Lys	Val	Pro	Ser	Phe	Gln	Trp	Leu	Asp	Arg	Asn	Val	Thr	Ile	Asp	
				165					170					175		
Thr	Leu	Phe	Ala	His	Thr	Leu	Ser	Gln	Ile	Arg	Ala	Ala	Asn	Gln	Lys	
			180					185						190		
Gly	Ala	Asn	Pro	Pro	Tyr	Ala	Gly	Ile	Phe	Val	Val	Tyr	Asp	Leu	Pro	
		195					200					205				
Asp	Arg	Asp	Cys	Ala	Ala	Ala	Ala	Ser	Asn	Gly	Glu	Phe	Ser	Ile	Ala	
	210					215					220					
Asn	Asn	Gly	Ala	Ala	Asn	Tyr	Lys	Thr	Tyr	Ile	Asp	Ala	Ile	Arg	Ser	
225					230					235					240	
Leu	Val	Ile	Gln	Tyr	Ser	Asp	Ile	Arg	Ile	Ile	Phe	Val	Ile	Glu	Pro	
			245						250					255		
Asp	Ser	Leu	Ala	Asn	Met	Val	Thr	Asn	Leu	Asn	Val	Ala	Lys	Cys	Ala	
		260						265					270			
Asn	Ala	Glu	Ser	Thr	Tyr	Lys	Glu	Leu	Thr	Val	Tyr	Ala	Leu	Gln	Gln	
		275					280					285				
Leu	Asn	Leu	Pro	Asn	Val	Ala	Met	Tyr	Leu	Asp	Ala	Gly	His	Ala	Gly	
	290					295					300					
Trp	Leu	Gly	Trp	Pro	Ala	Asn	Ile	Gln	Pro	Ala	Ala	Asn	Leu	Phe	Ala	
305					310					315					320	
Glu	Ile	Tyr	Thr	Ser	Ala	Gly	Lys	Pro	Ala	Ala	Val	Arg	Gly	Leu	Ala	
				325					330					335		
Thr	Asn	Val	Ala	Asn	Tyr	Asn	Gly	Trp	Ser	Leu	Ala	Thr	Pro	Pro	Ser	
		340						345					350			
Tyr	Thr	Gln	Gly	Asp	Pro	Asn	Tyr	Asp	Glu	Ser	His	Tyr	Val	Gln	Ala	
		355					360					365				
Leu	Ala	Pro	Leu	Leu	Thr	Ala	Asn	Gly	Phe	Pro	Ala	His	Phe	Ile	Thr	
	370					375					380					
Asp	Thr	Gly	Arg	Asn	Gly	Lys	Gln	Pro	Thr	Gly	Gln	Arg	Gln	Trp	Gly	
385					390					395					400	
Asp	Trp	Cys	Asn	Val	Ile	Gly	Thr	Gly	Phe	Gly	Val	Arg	Pro	Thr	Thr	
			405						410					415		
Asn	Thr	Gly	Leu	Asp	Ile	Glu	Asp	Ala	Phe	Val	Trp	Val	Lys	Pro	Gly	
		420						425					430			
Gly	Glu	Cys	Asp	Gly	Thr	Ser	Asn	Thr	Thr	Ser	Pro	Arg	Tyr	Asp	Tyr	
		435					440					445				
His	Cys	Gly	Leu	Ser	Asp	Ala	Leu	Gln	Pro	Ala	Pro	Glu	Ala	Gly	Thr	
	450					455						460				

[0050]

Trp Phe Gln Ala Tyr Phe Glu Gln Leu Leu Thr Asn Ala Asn Pro Pro	
465 470 475 480	
Phe	
<210> 53	
<211> 1593	
<212> DNA	
<213> 嗜热毛壳菌(<i>Chaetomium thermophilum</i>)	
<400> 53	
atgatgtaca agaagttcgc egetetecgc gccctcgtgg ctggcgccgc cggccagcag	60
gettgtcccc tcaccactga gaaccacccc agaetcaett ggaagcgctg cacctctggc	120
ggcaactgct cgaccgtgaa cggcgcgcgc accatcgatg ccaactggcg ctggactcac	180
actgtttccg gctcgaccaa ctgctacacc ggcaacgagt gggataccctc catctgctct	240
gatggcaaga gctgcgccea gaccigtgcg gtgcagcgcg ctgactactc ttcgacctat	300
ggatcaccaa ccagcggtga ctcctgaac ctcaagttcg tcaccaagca ccagcacggc	360
accaatgtcg gctctcgtgt ctacctgatg gagaacgaca ccaagtacca gatgttcgag	420
ctctctggca acgagttcac cttcgatgtc gatgtctcta acctgggctg cggctctaac	480
ggcgccctct acttcgtctc catggacget gatggtggtg tgagcaagta ctctggcaac	540
aaggctggcg ccaagtacgg taccggctac tgcgatgtc agtgcccgcg cgaccttaag	600
ttcatcaacg gcgaggccaa catlgagaac tggaccctt cgaccaatga tgccaacgcc	660
ggtttcggcc gtatggcag ctgctgctct gagatggata tctgggatgc caacaacatg	720
gtacttgctt tcactctctc cccttgcaac attatcgccc agagccgctg cgaggccaac	780
agctgcgggt gcacctacag ctctgagcgc tatgtgggtg ttgcatcc tgatggctgc	840
gacttcaacg ctaccgcca gggcgacaag acctctctac gcaaggscat gaccgtcgac	900
accaccaaga agatgacctg cgtcacccag ttccacaaga actcggttg cgctctcagc	960
gagatcaage gcttctactg tcaggacggc aagatcattg ccaacgccga gtccaagatc	1020
cccgcaaac ccggcaactc caccacccag gatgggtgcg atgcccagaa ggctgccttc	1080
ggtgacatcg atgacttcaa ccgcaagggc ggtatggctc agatgagcaa ggccctcgag	1140
ggccctatgg tctgggtcat gtcgtctgg gatgaccact acgccaacat gctctggctc	1200
gactcgacct accccattga caaggccgcg acccccgcg ccgagcgcg tgcttgcccg	1260
accacctcgg gtgtccctgc cgagattgag gccaggttc ccaacagcaa cgttatcttc	1320
ttcaaacatc gcttcggccc catcggtctg accgtccctg gcctcgacgg cagcaccccc	1380
agcaaccga ccgccacctg tctctctccc acttctacca ccaccagctg gagaagcagc	1440
actactcaga ttccacccc gactagccag cccggcggt gcaccacca gaagtggggc	1500
cagtgcgggt gtatcggtc caccggtgc actaacgctg ttgtgggac taactgcact	1560
gagctcaacc cctggtacag ccagtgcctg taa	1593
<210> 54	
<211> 530	
<212> PRT	
<213> 嗜热毛壳菌(<i>Chaetomium thermophilum</i>)	

[0051]

<400> 54

```

Met Met Tyr Lys Lys Phe Ala Ala Leu Ala Ala Leu Val Ala Gly Ala
1           5           10           15
Ala Ala Gln Gln Ala Cys Ser Leu Thr Thr Glu Thr His Pro Arg Leu
20           25           30
Thr Trp Lys Arg Cys Thr Ser Gly Gly Asn Cys Ser Thr Val Asn Gly
35           40           45
Ala Val Thr Ile Asp Ala Asn Trp Arg Trp Thr His Thr Val Ser Gly
50           55           60
Ser Thr Asn Cys Tyr Thr Gly Asn Glu Trp Asp Thr Ser Ile Cys Ser
65           70           75           80
Asp Gly Lys Ser Cys Ala Gln Thr Cys Cys Val Asp Gly Ala Asp Tyr
85           90           95
Ser Ser Thr Tyr Gly Ile Thr Thr Ser Gly Asp Ser Leu Asn Leu Lys
100          105          110
Phe Val Thr Lys His Gln His Gly Thr Asn Val Gly Ser Arg Val Tyr
115          120          125
Leu Met Glu Asn Asp Thr Lys Tyr Gln Met Phe Glu Leu Leu Gly Asn
130          135          140
Glu Phe Thr Phe Asp Val Asp Val Ser Asn Leu Gly Cys Gly Leu Asn
145          150          155          160
Gly Ala Leu Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Met Ser Lys
165          170          175
Tyr Ser Gly Asn Lys Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp
180          185          190
Ala Gln Cys Pro Arg Asp Leu Lys Phe Ile Asn Gly Glu Ala Asn Ile
195          200          205
Glu Asn Trp Thr Pro Ser Thr Asn Asp Ala Asn Ala Gly Phe Gly Arg
210          215          220
Tyr Gly Ser Cys Cys Ser Glu Met Asp Ile Trp Asp Ala Asn Asn Met
225          230          235          240
Ala Thr Ala Phe Thr Pro His Pro Cys Thr Ile Ile Gly Gln Ser Arg
245          250          255
Cys Glu Gly Asn Ser Cys Gly Gly Thr Tyr Ser Ser Glu Arg Tyr Ala
260          265          270
Gly Val Cys Asp Pro Asp Gly Cys Asp Phe Asn Ala Tyr Arg Gln Gly
275          280          285
Asp Lys Thr Phe Tyr Gly Lys Gly Met Thr Val Asp Thr Thr Lys Lys
290          295          300
Met Thr Val Val Thr Gln Phe His Lys Asn Ser Ala Gly Val Leu Ser
305          310          315          320
Glu Ile Lys Arg Phe Tyr Val Gln Asp Gly Lys Ile Ile Ala Asn Ala
325          330          335
Glu Ser Lys Ile Pro Gly Asn Pro Gly Asn Ser Ile Thr Gln Glu Trp
340          345          350
Cys Asp Ala Gln Lys Val Ala Phe Gly Asp Ile Asp Asp Phe Asn Arg
355          360          365
Lys Gly Gly Met Ala Gln Met Ser Lys Ala Leu Glu Gly Pro Met Val
370          375          380

```

[0052]

Leu Val Met Ser Val Trp Asp Asp His Tyr Ala Asn Met Leu Trp Leu
 385 390 395 400
 Asp Ser Thr Tyr Pro Ile Asp Lys Ala Gly Thr Pro Gly Ala Glu Arg
 405 410 415
 Gly Ala Cys Pro Thr Thr Ser Gly Val Pro Ala Glu Ile Glu Ala Gln
 420 425 430
 Val Pro Asn Ser Asn Val Ile Phe Ser Asn Ile Arg Phe Gly Pro Ile
 435 440 445
 Gly Ser Thr Val Pro Gly Leu Asp Gly Ser Thr Pro Ser Asn Pro Thr
 450 455 460
 Ala Thr Val Ala Pro Pro Thr Ser Thr Thr Thr Ser Val Arg Ser Ser
 465 470 475 480
 Thr Thr Gln Ile Ser Thr Pro Thr Ser Gln Pro Gly Gly Cys Thr Thr
 485 490 495
 Gln Lys Trp Gly Gln Cys Gly Gly Ile Gly Tyr Thr Gly Cys Thr Asn
 500 505 510
 Cys Val Ala Gly Thr Thr Cys Thr Glu Leu Asn Pro Trp Tyr Ser Gln
 515 520 525
 Cys Leu
 530

<210> 55
 <211> 1434
 <212> DNA
 <213> 嗜热毛壳菌(*Chaetomium thermophilum*)

<400> 55
 atggctaagc agctgctgct cactgccgct cttgcggcca cttegtggc tgccctctc 60
 cttgaggagc gccagagctg ctctccgctc tggggtaaat gcggtggcat caattacaac 120
 ggcccgacct gctgccagtc cggcagtgtt tgcacttacc tgaatgacfg gtacagccag 180
 tgcattcccg gtcaggctca gcccggcacg actagcacca cggctcggac caccagcacc 240
 agcaccacca gcacttcgtc ggtccgcccg accacctega ataccctgt gacgactgct 300
 cccccagca ccaccatccc ggccggcgcc tcgagcacgg ccagctacaa cggcaaccgc 360
 ttttcgggtg ttcaactttg ggccaacacc tactactcgt ccgaggtgca cactttggcc 420
 atccccagct tgtctctga gctggtgcc aaggccgcca aggtcgtga ggttcccagc 480
 ttccagtggc tcgaccgcaa tggactgtt gacactctct tctccggcac tcttgccgaa 540
 atccgcgcgc ccaaccagcg cggcgccaac ccgccttatg ccggcatttt cgtggtttat 600
 gacttaccag accgtgattg cgcgctgct gcttcgaacg gcgagtggc tatcgccaac 660
 aatggtgcca acaactacaa gcctacatc gaccggatcc gtgagctect tatccagtac 720
 tccgatatcc gcactattct ggtcattgaa cctgattccc tggccaacat ggtcaccaac 780
 atgaacgtcc agaagtgtc gaacgtgcc tccacttaca aggagcttac tgtctatgcc 840
 ctcaaacagc tcaatcttcc tcacgttgcc atgtacatgg atgctggcca cgctggctgg 900
 cttggttggc ccgccaacat ccagctgct gctgagctct ttgctcaaat ctaccgcgac 960
 gctggcagge ccgtgctgt ccgggtctt gcgaccaacg ttgccaacta caatgcttgg 1020
 tcgatcgcca gcctccgct ctacaactct cctaaccgga actacgacga gaagcactat 1080

[0053]

```

attgaggcct ttgtctctct tctccgaac cagggtttcg acgcaaagtt catcgtcgac      1140
accggccgta acggcaagca gccactggc cagcttgaat ggggtcacatg gtgcaatgic      1200
aagggaactg gcttcgggtgt gcgccctact gctaacactg ggcattgaact tgttgatget      1260
ttcgtgtggg tcaagcccg tgccgagtc gacggcaacca gtccggacac cagcgctget      1320
cgttatgact atcactgcgg cctttccgac gcactgacte cggcgccctga ggctggccaa      1380
tggttcagg cttatttga acagctgcte atcaatgcca accctccgct ctga      1434

```

<210> 56
 <211> 477
 <212> PRT
 <213> 嗜热毛壳菌(*Chaetomium thermophilum*)

<400> 56

```

Met Ala Lys Gln Leu Leu Leu Thr Ala Ala Leu Ala Ala Thr Ser Leu
1           5           10           15
Ala Ala Pro Leu Leu Glu Glu Arg Gln Ser Cys Ser Ser Val Trp Gly
20          25          30
Gln Cys Gly Gly Ile Asn Tyr Asn Gly Pro Thr Cys Cys Gln Ser Gly
35          40          45
Ser Val Cys Thr Tyr Leu Asn Asp Trp Tyr Ser Gln Cys Ile Pro Gly
50          55          60
Gln Ala Gln Pro Gly Thr Thr Ser Thr Thr Ala Arg Thr Thr Ser Thr
65          70          75          80
Ser Thr Thr Ser Thr Ser Ser Val Arg Pro Thr Thr Ser Asn Thr Pro
85          90          95
Val Thr Thr Ala Pro Pro Thr Thr Thr Ile Pro Gly Gly Ala Ser Ser
100         105         110
Thr Ala Ser Tyr Asn Gly Asn Pro Phe Ser Gly Val Gln Leu Trp Ala
115         120         125
Asn Thr Tyr Tyr Ser Ser Glu Val His Thr Leu Ala Ile Pro Ser Leu
130         135         140
Ser Pro Glu Leu Ala Ala Lys Ala Ala Lys Val Ala Glu Val Pro Ser
145         150         155         160
Phe Gln Trp Leu Asp Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly
165         170         175
Thr Leu Ala Glu Ile Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro Pro
180         185         190
Tyr Ala Gly Ile Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala
195         200         205
Ala Ala Ala Ser Asn Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn
210         215         220
Asn Tyr Lys Arg Tyr Ile Asp Arg Ile Arg Glu Leu Leu Ile Gln Tyr
225         230         235         240
Ser Asp Ile Arg Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn
245         250         255
Met Val Thr Asn Met Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Thr
260         265         270
Tyr Lys Glu Leu Thr Val Tyr Ala Leu Lys Gln Leu Asn Leu Pro His

```

[0054]

275	280	285
Val Ala Met Tyr Met Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro		
290	295	300
Ala Asn Ile Gln Pro Ala Ala Glu Leu Phe Ala Gln Ile Tyr Arg Asp		
305	310	315 320
Ala Gly Arg Pro Ala Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn		
	325	330 335
Tyr Asn Ala Trp Ser Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn		
	340 345	350
Pro Asn Tyr Asp Glu Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu		
	355 360	365
Arg Asn Gln Gly Phe Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn		
	370 375	380
Gly Lys Gln Pro Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val		
385	390	395 400
Lys Gly Thr Gly Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu		
	405 410	415
Leu Val Asp Ala Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly		
	420 425	430
Thr Ser Ala Asp Thr Ser Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu		
	435 440	445
Ser Asp Ala Leu Thr Pro Ala Pro Glu Ala Gly Gln Trp Phe Gln Ala		
	450 455	460
Tyr Phe Glu Gln Leu Leu Ile Asn Ala Asn Pro Pro Leu		
465	470	475

<210> 57
 <211> 2586
 <212> DNA
 <213> 米曲霉(Aspergillus oryzae)

<400> 57
 atgaagcttg gttggatcga ggtggccgca ttggcggtcg cctcagtagt cagtgccaaag 60
 gatgatctcg cgtaccccc tcefttctac cctccccat gggcagatgg tcagggtgaa 120
 tggcggaag tatacaaacg cgtgttagac atagtctccc agatgacgtt gacagagaaa 180
 gtcaacttaa cgactggaac aggatggcaa ctagagaggt gfgttggaca aactggcagt 240
 gtccccagac tcaacatccc cagcttgtgt ttgcaggata gtctcttgg tattcgtttc 300
 tcggactaca atcagctttt cctgcgggtt gtaaatgtcg ctgccacctg ggacaagacg 360
 ctgcctacc ttctgtgtca ggcaatgggt gaggagtcca gtgataaggg tattgacgtt 420
 cagctgggtc ctgctgtctg ccctctcggt gctatccgg atggcggtag aaactgggaa 480
 ggtttctcac cagatccage cctcaccggt gtactttttg cggagacgat taagggtatt 540
 caagatgctg gtgtcattgc gacagctaag cattatatca tgaacgaaca agagcatttc 600
 cgccaacaac ccgaggttgc gggttacgga ttcaacgtaa gcgacagttt gagttccaac 660
 gttgatgaca agactatgca tgaattgtac ctctggccct tcgcggatgc agtacgcgt 720
 ggagtcggtg ctgtcatgtg ctcttacaac caaatcaaca acagctacgg itgcgagaat 780
 agcgaaactc tgaacaaget ttgaaggcg gagcttggtt tccaaggctt cgtcatgagt 840

[0055]

gattggaccg ctcatcacag cggcgtaggg gctgctttag caggctctgga tatgtcgaig 900
 cccggtgatg ttacctttega tagtggtaag tctttctggg gtgcaaacit gacggctcgg 960
 gtcccttaacg gtacaatccc ccaatggcgt gttgatgaca tggctgtccg tatcatggcc 1020
 gcttattaca aggttggcgg cgacaccaaa tacacccctc ccaacttcag ctgctggacc 1080
 agggacgaat atggtttcgc gcataaccat gtttcggaag gtgcttacga gagggtaac 1140
 gaattcgtgg acgtgcaacg cgatcatgcc gacctaatcc gtgcgacggc cgcgcagagg 1200
 actgtttctg tgaagaacaa ggggtgcctg cccttgagcc gcaaggaaaa gctggctgcc 1260
 cttctgggag aggatgcggg ttccaactcg tggggcgcta acggctgtga tgaccgtggt 1320
 tgcgataacg gtaccttgc catggcctgg ggtagcggta ctgcgaattt cccatacctc 1380
 gtgacaccag agcagggcat tcagaacgaa gtcttcagg gccgiggtaa tgtcttcgcc 1440
 gtgaecgaca gttgggcgtc cgacaagatc gctggcgtg cccgccagcg cagcgtatct 1500
 ctgctgttcg tcaactccga ctcaggagaa ggctatctta gtgtggatgg aatgagggc 1560
 gatcgtaaca acatcacctt gtggaagaac ggcacaatg tggtaagac cgcagcgaat 1620
 aactglaaca acaccgttgt catcatccac tccgtcggac cagttttgat cgatgaatgg 1680
 tatgaccacc ccaatgtcac tggatattct tgggtcggtc tgcagggcca ggagtctggt 1740
 aactccattg ccgatgtgct gtacggtcgt gtcaacctg gcgccaagtc tcccttcaet 1800
 tggggcaaga cccgggagtc gtaiggttct cccttggtca aggatgcaa caatggcaac 1860
 ggagcgcgcc agtctgattt caccacgggt gttttcatcg attaccgcca ttctgataag 1920
 ttcaatgaga cccctatcta cgagtttggc tacgcttga gtacacccac ctctgagctc 1980
 tccgacctcc atgttcagcc cctgaacggc tccgataca ctcccaccag tggcaigact 2040
 gaagctgcaa agaactttgg tgaatttggc gatgcgtcgg agtacgtgta tccggagggg 2100
 ctggaaagga tccatgagtt tatctatccc tggatcaact ctaccgacct gaaggcatcg 2160
 tctgacgatt ctaactacgg ctgggaagac tccaagtata ttcccgaagg cggcacggat 2220
 gggctctgcc agccccgttt gcccgctagt ggtgggtgcc gaggaacccc cggctctgac 2280
 gaggatcttt tcccgctctc tgtgaaggtc aagaacacgg gcaatgtcgc cggatgatga 2340
 gttctctcagc tgtacgttcc cctagggcgc ccgaatgagc ccaaggtggt actgcgcaag 2400
 tttagcgtga ttcacttggc ccttcgcag gaggcctgt ggacaacgac ccttaccctg 2460
 cgtgaccttg caaactggga cgtttcggt caggactgga ccgtcactcc ttaccccaag 2520
 acgatctacg ttggaaactc ctacgggaaa ctgcgcctcc aggcctcgtt gcttaaggcc 2580
 cagtaa 2586

<210> 58

<211> 861

<212> PRT

<213> 米曲霉(*Aspergillus oryzae*)

<400> 58

Met Lys Leu Gly Trp Ile Glu Val Ala Ala Leu Ala Ala Ala Ser Val
1 5 10 15

Val Ser Ala Lys Asp Asp Leu Ala Tyr Ser Pro Pro Phe Tyr Pro Ser
20 25 30

[0056]

Pro Trp Ala Asp Gly Gln Gly Glu Trp Ala Glu Val Tyr Lys Arg Ala
 35 40 45
 Val Asp Ile Val Ser Gln Met Thr Leu Thr Glu Lys Val Asn Leu Thr
 50 55 60
 Thr Gly Thr Gly Trp Gln Leu Glu Arg Cys Val Gly Gln Thr Gly Ser
 65 70 75 80
 Val Pro Arg Leu Asn Ile Pro Ser Leu Cys Leu Gln Asp Ser Pro Leu
 85 90 95
 Gly Ile Arg Phe Ser Asp Tyr Asn Ser Ala Phe Pro Ala Gly Val Asn
 100 105 110
 Val Ala Ala Thr Trp Asp Lys Thr Leu Ala Tyr Leu Arg Gly Gln Ala
 115 120 125
 Met Gly Glu Glu Phe Ser Asp Lys Gly Ile Asp Val Gln Leu Gly Pro
 130 135 140
 Ala Ala Gly Pro Leu Gly Ala His Pro Asp Gly Gly Arg Asn Trp Glu
 145 150 155 160
 Gly Phe Ser Pro Asp Pro Ala Leu Thr Gly Val Leu Phe Ala Glu Thr
 165 170 175
 Ile Lys Gly Ile Gln Asp Ala Gly Val Ile Ala Thr Ala Lys His Tyr
 180 185 190
 Ile Met Asn Glu Gln Glu His Phe Arg Gln Gln Pro Glu Ala Ala Gly
 195 200 205
 Tyr Gly Phe Asn Val Ser Asp Ser Leu Ser Ser Asn Val Asp Asp Lys
 210 215 220
 Thr Met His Glu Leu Tyr Leu Trp Pro Phe Ala Asp Ala Val Arg Ala
 225 230 235 240
 Gly Val Gly Ala Val Met Cys Ser Tyr Asn Gln Ile Asn Asn Ser Tyr
 245 250 255
 Gly Cys Glu Asn Ser Glu Thr Leu Asn Lys Leu Leu Lys Ala Glu Leu
 260 265 270
 Gly Phe Gln Gly Phe Val Met Ser Asp Trp Thr Ala His His Ser Gly
 275 280 285
 Val Gly Ala Ala Leu Ala Gly Leu Asp Met Ser Met Pro Gly Asp Val
 290 295 300
 Thr Phe Asp Ser Gly Thr Ser Phe Trp Gly Ala Asn Leu Thr Val Gly
 305 310 315 320
 Val Leu Asn Gly Thr Ile Pro Gln Trp Arg Val Asp Asp Met Ala Val
 325 330 335
 Arg Ile Met Ala Ala Tyr Tyr Lys Val Gly Arg Asp Thr Lys Tyr Thr
 340 345 350
 Pro Pro Asn Phe Ser Ser Trp Thr Arg Asp Glu Tyr Gly Phe Ala His
 355 360 365
 Asn His Val Ser Glu Gly Ala Tyr Glu Arg Val Asn Glu Phe Val Asp
 370 375 380
 Val Gln Arg Asp His Ala Asp Leu Ile Arg Arg Ile Gly Ala Gln Ser
 385 390 395 400
 Thr Val Leu Leu Lys Asn Lys Gly Ala Leu Pro Leu Ser Arg Lys Glu
 405 410 415

[0057]

Lys Leu Val Ala Leu Leu Gly Glu Asp Ala Gly Ser Asn Ser Trp Gly
 420 425 430
 Ala Asn Gly Cys Asp Asp Arg Gly Cys Asp Asn Gly Thr Leu Ala Met
 435 440 445
 Ala Trp Gly Ser Gly Thr Ala Asn Phe Pro Tyr Leu Val Thr Pro Glu
 450 455 460
 Gln Ala Ile Glu Asn Glu Val Leu Gln Gly Arg Gly Asn Val Phe Ala
 465 470 475 480
 Val Thr Asp Ser Trp Ala Leu Asp Lys Ile Ala Ala Ala Ala Arg Gln
 485 490 495
 Ala Ser Val Ser Leu Val Phe Val Asn Ser Asp Ser Gly Glu Gly Tyr
 500 505 510
 Leu Ser Val Asp Gly Asn Glu Gly Asp Arg Asn Asn Ile Thr Leu Trp
 515 520 525
 Lys Asn Gly Asp Asn Val Val Lys Thr Ala Ala Asn Asn Cys Asn Asn
 530 535 540
 Thr Val Val Ile Ile His Ser Val Gly Pro Val Leu Ile Asp Glu Trp
 545 550 555 560
 Tyr Asp His Pro Asn Val Thr Gly Ile Leu Trp Ala Gly Leu Pro Gly
 565 570 575
 Gln Glu Ser Gly Asn Ser Ile Ala Asp Val Leu Tyr Gly Arg Val Asn
 580 585 590
 Pro Gly Ala Lys Ser Pro Phe Thr Trp Gly Lys Thr Arg Glu Ser Tyr
 595 600 605
 Gly Ser Pro Leu Val Lys Asp Ala Asn Asn Gly Asn Gly Ala Pro Gln
 610 615 620
 Ser Asp Phe Thr Gln Gly Val Phe Ile Asp Tyr Arg His Phe Asp Lys
 625 630 635 640
 Phe Asn Glu Thr Pro Ile Tyr Glu Phe Gly Tyr Gly Leu Ser Tyr Thr
 645 650 655
 Thr Phe Glu Leu Ser Asp Leu His Val Gln Pro Leu Asn Ala Ser Arg
 660 665 670
 Tyr Thr Pro Thr Ser Gly Met Thr Glu Ala Ala Lys Asn Phe Gly Glu
 675 680 685
 Ile Gly Asp Ala Ser Glu Tyr Val Tyr Pro Glu Gly Leu Glu Arg Ile
 690 695 700
 His Glu Phe Ile Tyr Pro Trp Ile Asn Ser Thr Asp Leu Lys Ala Ser
 705 710 715 720
 Ser Asp Asp Ser Asn Tyr Gly Trp Glu Asp Ser Lys Tyr Ile Pro Glu
 725 730 735
 Gly Ala Thr Asp Gly Ser Ala Gln Pro Arg Leu Pro Ala Ser Gly Gly
 740 745 750
 Ala Gly Gly Asn Pro Gly Leu Tyr Glu Asp Leu Phe Arg Val Ser Val
 755 760 765
 Lys Val Lys Asn Thr Gly Asn Val Ala Gly Asp Glu Val Pro Gln Leu
 770 775 780
 Tyr Val Ser Leu Gly Gly Pro Asn Glu Pro Lys Val Val Leu Arg Lys
 785 790 795 800
 Phe Glu Arg Ile His Leu Ala Pro Ser Gln Glu Ala Val Trp Thr Thr

[0058]

	805	810	815
Thr Leu Thr Arg Arg Asp Leu Ala Asn Trp Asp Val Ser Ala Gln Asp			
	820	825	830
Trp Thr Val Thr Pro Tyr Pro Lys Thr Ile Tyr Val Gly Asn Ser Ser			
	835	840	845
Arg Lys Leu Pro Leu Gln Ala Ser Leu Pro Lys Ala Gln			
	850	855	860
<210> 59			
<211> 3060			
<212> DNA			
<213> 烟曲霉(Aspergillus fumigatus)			
<400> 59			
atgagattcg gttggctcga ggtggcgcct ctgacggcgc cttctgtagc caatgccag			60
gtttgtgatg ctttcccgtc attgtttcgg atatatgtga caatagtcac ggaataatc			120
aggaattggc ttctcttcca ccattctacc cttcgccllg ggcctgaggc caggagagat			180
gggcagatgc ccacgcagc gcgctcgaga tctgtttctc gatgacactg gcggagaagg			240
ttaaccttac aacgggtact ggggtgggtg cgaacttttt gttgacagtg agctttcttc			300
actgaccatc tacacagaig ggaatggac cgaatgcgtc gtcaaacagg cagcgttccc			360
aggtaaagctt gcaattctgc aacaacgtgc aagtgtagtt gctaaaacgc ggtggtgcag			420
acttggtatc aactggggtc ttgtggcca ggattccctt ttgggtatcc gtttctgtga			480
gctataccgc cggagtcctt cagtccttgt attatgigt ctgatgtgc tctgtatagc			540
tgacctcaac tccgccttc ctgctggtae taatgtgcgc gcgacatggg acaagacact			600
cgcctacctt cgtggcaagg ccattgggtga ggaattcaac gacaaggcgc tggacatttt			660
gctggggcct gctgctggte ctctcggaac ataccggac ggcggcagaa tctggzaagg			720
cttctctcct gatccggttc tcactgggtt acttttcgcc gaaactatca aggtatcca			780
agacgcgggt gctgattgta ctgccaagca ttacattctg aatgaacagg agcatttccg			840
acaggttggc gaggcccagg gatatggtta caacatcacg gagacgatea gctccaacgt			900
ggatgacaag accatgcacg agttgtacct ttggtgagta gttgacactg caaatgagga			960
ccctgattga ttgactgac ctggaatgca ggccttttgc agatgcigtg cgcgtaaga			1020
tttccgtag acttgacctc gcgacgaaga aatcgctgac gaaccatcgt agctggcglt			1080
ggcgtgtgca tgtgttccca caatcaaatc aacaacagct acggttgtca aaacagtcaa			1140
actctcaaca agctctctca ggtgagctg ggcctccaag gcttcgtcat gagtgaactg			1200
agcgtcaccc acagcgggtg cgcgcgtgcc ctgctgggtt tggatafctc gatgcctgga			1260
gacatttctt tgcacgacgg actctctctc tggggcacga acctaacgtg cagtgttctt			1320
aacggcaccc ttccagcctg gcgtgtcgat gacatggctg ttctgtatcat gaccgcgtac			1380
tacaaggttg gtcgtgaccg tcttcgtatt cccctaaact tcagctcctg gaccgggat			1440
gagtaeggct gggagcattc tctgtctctc gagggagcct ggaccaaggt gaacgacttc			1500
gteaatgtgc agcgcagtea ctctcagatc atccgtgaga ttggtgccgc tagtacagtg			1560
ctcttgaaga acacgggtgc tcttcttttg accggcaagg aggttaaagt ggggtgtctc			1620
ggtaagacgc ctggttccaa cccgtggggt gctaacggct gccccgaccg cggtgtgat			1680

[0059]

```

aacggcactc ttgctatggc ctgggtagt ggtactgcca acttcctta ccttgtcacc 1740
cccgagcagg ctatccagcg agaggtaac agcaacggcg gcaatgtctt tgcgtgact 1800
gataacgggg ctctcagcca gatggcagat gtgcatctc aatccagggt agtgcgggct 1860
cttagaaaaa gaacgttctc tgaatgaagt tttttaacca ttgcgaacag cgtgtctttg 1920
gtgtttgtca acgcccactc tggagagggt ttcatcagtg tegacggcaa cgagggtgac 1980
cgcaaaaate tcaactgtg gaagaacggc gaggcgctca ttgacctgt tgtcagccac 2040
tgcaacaaca cgatttgtgt tattcacagt gtgggcccgc tcttgatcga ccggtggtat 2100
gataacccca acgtcactgc catcatctgg gccggcttgc ccggtcagga gagtggcaac 2160
tcccgtgtcg acgtgtctta tggccgcgtc aaccccagcg ccaagacccc gttcacctgg 2220
ggcaagactc gggagtctta cggggtctcc ttgtcaccg agcctaacaa tggcaatggt 2280
gtcctcccagg atgatttcaa cgaggcgctc ttcatgact accgtcacit tgacaagcgc 2340
aatgagaccc ccatttatga gtttgccat ggttgagct acaccacctt tggtaactct 2400
caccttcggg ttcagccct caatagtctg agtctggcat atgtcccgac tagcggagag 2460
accaagcctg cgccaacctt tggtagatc gtagtgccg ccgactacct gtatcccgag 2520
ggctcaaaaa gaattaccaa gtttatttac ccttggctca actcgaccga cctcgaggat 2580
tcttctgacg acccgaacta cggctgggag gactcggagt acattcccga aggcgctagg 2640
gatgggtctc ctcaacccct cctgaaggct ggccggcgtc ctggtggtaa ccctaccctt 2700
tatcaggatc ttgttaggtt gtcggccacc ataaccaaca ctggtaacgt cgcgggttat 2760
gaagtccttc aattggtgag tgaccgcctt gttccttggc ttgcaatttg gctaaactgc 2820
ttctagtatg ttctactggg cggaccgaac gacctcggg tegtctctcg caagttcgac 2880
cgaatcttcc tggtccttgg ggagcaaaag gtttgacca cgactcttaa ccgtcgtgat 2940
ctcgccaatt gggatgtgga ggtcaggac tgggtcatca caagtaccc caagaaagtg 3000
cacgtcgcca gctcctcgcg taagctgcct ctgagagcgc ctctgccccg tgtctactag 3060

```

<210> 60
 <211> 863
 <212> PRT
 <213> 烟曲霉(*Aspergillus fumigatus*)

<400> 60

```

Met Arg Phe Gly Trp Leu Glu Val Ala Ala Leu Thr Ala Ala Ser Val
1           5           10           15
Ala Asn Ala Gln Glu Leu Ala Phe Ser Pro Pro Phe Tyr Pro Ser Pro
20           25           30
Trp Ala Asp Gly Gln Gly Glu Trp Ala Asp Ala His Arg Arg Ala Val
35           40           45
Glu Ile Val Ser Gln Met Thr Leu Ala Glu Lys Val Asn Leu Thr Thr
50           55           60
Gly Thr Gly Trp Glu Met Asp Arg Cys Val Gly Gln Thr Gly Ser Val
65           70           75           80
Pro Arg Leu Gly Ile Asn Trp Gly Leu Cys Gly Gln Asp Ser Pro Leu
85           90           95

```

[0060]

Gly Ile Arg Phe Ser Asp Leu Asn Ser Ala Phe Pro Ala Gly Thr Asn
 100 105 110
 Val Ala Ala Thr Trp Asp Lys Thr Leu Ala Tyr Leu Arg Gly Lys Ala
 115 120 125
 Met Gly Glu Glu Phe Asn Asp Lys Gly Val Asp Ile Leu Leu Gly Pro
 130 135 140
 Ala Ala Gly Pro Leu Gly Lys Tyr Pro Asp Gly Gly Arg Ile Trp Glu
 145 150 155 160
 Gly Phe Ser Pro Asp Pro Val Leu Thr Gly Val Leu Phe Ala Glu Thr
 165 170 175
 Ile Lys Gly Ile Gln Asp Ala Gly Val Ile Ala Thr Ala Lys His Tyr
 180 185 190
 Ile Leu Asn Glu Gln Glu His Phe Arg Gln Val Gly Glu Ala Gln Gly
 195 200 205
 Tyr Gly Tyr Asn Ile Thr Glu Thr Ile Ser Ser Asn Val Asp Asp Lys
 210 215 220
 Thr Met His Glu Leu Tyr Leu Trp Pro Phe Ala Asp Ala Val Arg Ala
 225 230 235 240
 Gly Val Gly Ala Val Met Cys Ser Tyr Asn Gln Ile Asn Asn Ser Tyr
 245 250 255
 Gly Cys Gln Asn Ser Gln Thr Leu Asn Lys Leu Leu Lys Ala Glu Leu
 260 265 270
 Gly Phe Gln Gly Phe Val Met Ser Asp Trp Ser Ala His His Ser Gly
 275 280 285
 Val Gly Ala Ala Leu Ala Gly Leu Asp Met Ser Met Pro Gly Asp Ile
 290 295 300
 Ser Phe Asp Asp Gly Leu Ser Phe Trp Gly Thr Asn Leu Thr Val Ser
 305 310 315 320
 Val Leu Asn Gly Thr Val Pro Ala Trp Arg Val Asp Asp Met Ala Val
 325 330 335
 Arg Ile Met Thr Ala Tyr Tyr Lys Val Gly Arg Asp Arg Leu Arg Ile
 340 345 350
 Pro Pro Asn Phe Ser Ser Trp Thr Arg Asp Glu Tyr Gly Trp Glu His
 355 360 365
 Ser Ala Val Ser Glu Gly Ala Trp Thr Lys Val Asn Asp Phe Val Asn
 370 375 380
 Val Gln Arg Ser His Ser Gln Ile Ile Arg Glu Ile Gly Ala Ala Ser
 385 390 395 400
 Thr Val Leu Leu Lys Asn Thr Gly Ala Leu Pro Leu Thr Gly Lys Glu
 405 410 415
 Val Lys Val Gly Val Leu Gly Glu Asp Ala Gly Ser Asn Pro Trp Gly
 420 425 430
 Ala Asn Gly Cys Pro Asp Arg Gly Cys Asp Asn Gly Thr Leu Ala Met
 435 440 445
 Ala Trp Gly Ser Gly Thr Ala Asn Phe Pro Tyr Leu Val Thr Pro Glu
 450 455 460
 Gln Ala Ile Gln Arg Glu Val Ile Ser Asn Gly Gly Asn Val Phe Ala
 465 470 475 480
 Val Thr Asp Asn Gly Ala Leu Ser Gln Met Ala Asp Val Ala Ser Gln

[0061]

485	490	495
Ser Ser Val Ser Leu Val Phe Val Asn Ala Asp Ser Gly Glu Gly Phe 500 505 510		
Ile Ser Val Asp Gly Asn Glu Gly Asp Arg Lys Asn Leu Thr Leu Trp 515 520 525		
Lys Asn Gly Glu Ala Val Ile Asp Thr Val Val Ser His Cys Asn Asn 530 535 540		
Thr Ile Val Val Ile His Ser Val Gly Pro Val Leu Ile Asp Arg Trp 545 550 555 560		
Tyr Asp Asn Pro Asn Val Thr Ala Ile Ile Trp Ala Gly Leu Pro Gly 565 570 575		
Gln Glu Ser Gly Asn Ser Leu Val Asp Val Leu Tyr Gly Arg Val Asn 580 585 590		
Pro Ser Ala Lys Thr Pro Phe Thr Trp Gly Lys Thr Arg Glu Ser Tyr 595 600 605		
Gly Ala Pro Leu Leu Thr Glu Pro Asn Asn Gly Asn Gly Ala Pro Gln 610 615 620		
Asp Asp Phe Asn Glu Gly Val Phe Ile Asp Tyr Arg His Phe Asp Lys 625 630 635 640		
Arg Asn Glu Thr Pro Ile Tyr Glu Phe Gly His Gly Leu Ser Tyr Thr 645 650 655		
Thr Phe Gly Tyr Ser His Leu Arg Val Gln Ala Leu Asn Ser Ser Ser 660 665 670		
Ser Ala Tyr Val Pro Thr Ser Gly Glu Thr Lys Pro Ala Pro Thr Tyr 675 680 685		
Gly Glu Ile Gly Ser Ala Ala Asp Tyr Leu Tyr Pro Glu Gly Leu Lys 690 695 700		
Arg Ile Thr Lys Phe Ile Tyr Pro Trp Leu Asn Ser Thr Asp Leu Glu 705 710 715 720		
Asp Ser Ser Asp Asp Pro Asn Tyr Gly Trp Glu Asp Ser Glu Tyr Ile 725 730 735		
Pro Glu Gly Ala Arg Asp Gly Ser Pro Gln Pro Leu Leu Lys Ala Gly 740 745 750		
Gly Ala Pro Gly Gly Asn Pro Thr Leu Tyr Gln Asp Leu Val Arg Val 755 760 765		
Ser Ala Thr Ile Thr Asn Thr Gly Asn Val Ala Gly Tyr Glu Val Pro 770 775 780		
Gln Leu Tyr Val Ser Leu Gly Gly Pro Asn Glu Pro Arg Val Val Leu 785 790 795 800		
Arg Lys Phe Asp Arg Ile Phe Leu Ala Pro Gly Glu Gln Lys Val Trp 805 810 815		
Thr Thr Thr Leu Asn Arg Arg Asp Leu Ala Asn Trp Asp Val Glu Ala 820 825 830		
Gln Asp Trp Val Ile Thr Lys Tyr Pro Lys Lys Val His Val Gly Ser 835 840 845		
Ser Ser Arg Lys Leu Pro Leu Arg Ala Pro Leu Pro Arg Val Tyr 850 855 860		

<210> 61

[0062]

<211> 2800	
<212> DNA	
<213> 巴西青霉(Penicillium brasilianum)	
<400> 61	
tgaaaatgca gggttctaca atctttctgg ctttgcctc atgggegagc caggttgetg	60
ccattgcga gccatcacag aagcacgagg ttgttttlat ctgtctcatg gacgtgctt	120
gacttgacta attgttttac atacagcccg gatctctgca cgggcccaca gccatagaat	180
cgttctcaga accgtttctac ccgtgcctt ggatgaatcc tcaegccgag ggctgggagg	240
ccgcatacca gaaagctcaa gattttgtct cgcaactcac tatcttgag aaaataaacc	300
tgaccaccgg tgttgggtaa gtctctccga ctgtctcagg gtcacggtgc gacgagccac	360
tgactttttg aagctgggaa aatgggcctt gtgtaggaaa cactggatca attctctgtc	420
tcggattcaa aggattttgt acccaggatt caccacaggg tgttcggttc gcagattatt	480
ctctccgttt cacatctagc caaatggccg ccgcaacatt tgaccgtcca attctttatc	540
aacgaggcca agccatggca caggaacaca aggctaaggg tatcacaatt caattgggcc	600
ctgttgcggg ccctctcggt cgcctcccg agggcggccg caactgggaa ggattctccc	660
ctgatcctgt cttgactggt atagccatgg ctgagacaat taaggscatg caggatactg	720
gagtgtatgc ttgcctaaa cattatattg gaaacgagca ggagcacttc cgtcaagtgg	780
gtgaagctgc gggtcacgga tacactattt ccgatactat ttcactetaat attgacgacc	840
gtgctatgca tgagctatac ttgtggccat ttgtctatgc cgttccgctt ggtgtgggtt	900
ctttcatgtg ctcatactct cagatcaaca actctacagg atgccaaaac agtcagacc	960
tcacaagct cctcnaagc gaattgggtt tccaaggctt tgtcatgagc gattgggggtg	1020
cccatcactc tggagtgtca tcggcgctag ctggacttga tatgagcatg ccgggtgata	1080
ccgaatttga ttctggcttg agcttctggg gctctaacct caccatttga attctgaacg	1140
gcacggttcc cgaatggcgc ctggatgaca tgccgatgag aattatggct gcatacttca	1200
aagttggcct tactatttag gatcaaccag atgtcaactt caatgcctgg acccatgaca	1260
cctacggata taataacgtt tatagcaagg aagattacga gcaggccaac tggcatgtcg	1320
atgttcgcag cgaccacaat aagctcattc gcgagactgc cgcgaagggt acagtcttgc	1380
tgaagaacaa ctttcatgtt ctccctctga agcagcccag gttcgtggcc gtcgttggtc	1440
aggatgccgg gccaaacccc aagggcccta acggctgcgc agaccgagga tgcgaccaag	1500
gcactctcgc aatgggatgg ggcctcagggt ctaccgaatt cctttacctg gtcactcccg	1560
acactgttat tcagtcaaag gtctctgaat acgggggtcg atacgagatg atttttgata	1620
actatgacga caatgtatc ttgtcgcttg tctcacagcc tgatgcaacc tgtatcgttt	1680
ttgcaaatgc cgattccggt gaaggctaca tcaactgtga caacaactgg ggtgaecgca	1740
acaatctgac cctctggcaa aatgcccagc aagtgattag cactgtcagc tcgcgatgca	1800
acaacacaat cgttgttctc cactctgtcg gaccagtgtt gctaaatggt atatatgagc	1860
acccgaacat cacagctatt gtctggcgag ggtatgccag cgaagaatct ggcaatgctc	1920
tcgttgatat tctttggggc aatgttaacc ctgccggtcg cactccgttc acctgggcca	1980
aaagtcgaga ggactatggc actgatataa tgtacgagcc caacaacggc cagcgtgcgc	2040

[0063]

```

ctcagcagga tttcaccgag agcatctacc tcgactaccg ccatttcgac aaagetggta 2100
tcgagccaat ttacgagttt ggattcggcc tctcctatac caccttegaa tactctgacc 2160
tcctgtttgt gaagaagtat gtccaacat acagtcccac gaccggcacc ggtgctcaag 2220
caccttccat cggacagcca cctagccaga acctggatac ctacaagttc cctgctacat 2280
acaagfacat caaaaccctc atttatecct acctgaacag cactgtctcc ctcgcgctg 2340
cttccaagga tcccgaatac ggtcgtacag accttatecc accccacgeg cgtgatggct 2400
ccctcaacc tctcaacccc gtggagacc cagtggccag tggtggaac aacatgctct 2460
acgacgaact ttacgaggtc actgcacaga tcaaaaacac tggcgacgtg gccggegacg 2520
aagtcgtcca gctttacgta gatctcgggg gtgacaaccc gcctcgtcag ttgagaaact 2580
ttgacagggt ttatctgctg cccggtcaga gctcaacatt cggggtaca ttgacgcgcc 2640
gtgatttgag caactgggat atigaggcgc agaactggcg agttacggaa tcgctaaga 2700
gagtgtatgt tggacggteg agtcgggatt tgcgctgag ctcacaattg gactaatgat 2760
catgtctacc aatagatggt gaatgtctgg tgtggatatt 2800

```

<210> 62
 <211> 878
 <212> PRT
 <213> 巴西青霉(Penicillium brasilianum)

<400> 62

```

Met Gln Gly Ser Thr Ile Phe Leu Ala Phe Ala Ser Trp Ala Ser Gln
1           5           10          15
Val Ala Ala Ile Ala Gln Pro Ile Gln Lys His Glu Pro Gly Phe Leu
20          25          30
His Gly Pro Gln Ala Ile Glu Ser Phe Ser Glu Pro Phe Tyr Pro Ser
35          40          45
Pro Trp Met Asn Pro His Ala Glu Gly Trp Glu Ala Ala Tyr Gln Lys
50          55          60
Ala Gln Asp Phe Val Ser Gln Leu Thr Ile Leu Glu Lys Ile Asn Leu
65          70          75          80
Thr Thr Gly Val Gly Trp Glu Asn Gly Pro Cys Val Gly Asn Thr Gly
85          90          95
Ser Ile Pro Arg Leu Gly Phe Lys Gly Phe Cys Thr Gln Asp Ser Pro
100         105         110
Gln Gly Val Arg Phe Ala Asp Tyr Ser Ser Ala Phe Thr Ser Ser Gln
115         120         125
Met Ala Ala Ala Thr Phe Asp Arg Ser Ile Leu Tyr Gln Arg Gly Gln
130         135         140
Ala Met Ala Gln Glu His Lys Ala Lys Gly Ile Thr Ile Gln Leu Gly
145         150         155         160
Pro Val Ala Gly Pro Leu Gly Arg Ile Pro Glu Gly Gly Arg Asn Trp
165         170         175
Glu Gly Phe Ser Pro Asp Pro Val Leu Thr Gly Ile Ala Met Ala Glu
180         185         190
Thr Ile Lys Gly Met Gln Asp Thr Gly Val Ile Ala Cys Ala Lys His
195         200         205

```

[0064]

Tyr Ile Gly Asn Glu Gln Glu His Phe Arg Gln Val Gly Glu Ala Ala
 210 215 220
 Gly His Gly Tyr Thr Ile Ser Asp Thr Ile Ser Ser Asn Ile Asp Asp
 225 230 235 240
 Arg Ala Met His Glu Leu Tyr Leu Trp Pro Phe Ala Asp Ala Val Arg
 245 250 255
 Ala Gly Val Gly Ser Phe Met Cys Ser Tyr Ser Gln Ile Asn Asn Ser
 260 265 270
 Tyr Gly Cys Gln Asn Ser Gln Thr Leu Asn Lys Leu Leu Lys Ser Glu
 275 280 285
 Leu Gly Phe Gln Gly Phe Val Met Ser Asp Trp Gly Ala His His Ser
 290 295 300
 Gly Val Ser Ser Ala Leu Ala Gly Leu Asp Met Ser Met Pro Gly Asp
 305 310 315 320
 Thr Glu Phe Asp Ser Gly Leu Ser Phe Trp Gly Ser Asn Leu Thr Ile
 325 330 335
 Ala Ile Leu Asn Gly Thr Val Pro Glu Trp Arg Leu Asp Asp Met Ala
 340 345 350
 Met Arg Ile Met Ala Ala Tyr Phe Lys Val Gly Leu Thr Ile Glu Asp
 355 360 365
 Gln Pro Asp Val Asn Phe Asn Ala Trp Thr His Asp Thr Tyr Gly Tyr
 370 375 380
 Lys Tyr Ala Tyr Ser Lys Glu Asp Tyr Glu Gln Val Asn Trp His Val
 385 390 395 400
 Asp Val Arg Ser Asp His Asn Lys Leu Ile Arg Glu Thr Ala Ala Lys
 405 410 415
 Gly Thr Val Leu Leu Lys Asn Asn Phe His Ala Leu Pro Leu Lys Gln
 420 425 430
 Pro Arg Phe Val Ala Val Val Gly Gln Asp Ala Gly Pro Asn Pro Lys
 435 440 445
 Gly Pro Asn Gly Cys Ala Asp Arg Gly Cys Asp Gln Gly Thr Leu Ala
 450 455 460
 Met Gly Trp Gly Ser Gly Ser Thr Glu Phe Pro Tyr Leu Val Thr Pro
 465 470 475 480
 Asp Thr Ala Ile Gln Ser Lys Val Leu Glu Tyr Gly Gly Arg Tyr Glu
 485 490 495
 Ser Ile Phe Asp Asn Tyr Asp Asp Asn Ala Ile Leu Ser Leu Val Ser
 500 505 510
 Gln Pro Asp Ala Thr Cys Ile Val Phe Ala Asn Ala Asp Ser Gly Glu
 515 520 525
 Gly Tyr Ile Thr Val Asp Asn Asn Trp Gly Asp Arg Asn Asn Leu Thr
 530 535 540
 Leu Trp Gln Asn Ala Asp Gln Val Ile Ser Thr Val Ser Ser Arg Cys
 545 550 555 560
 Asn Asn Thr Ile Val Val Leu His Ser Val Gly Pro Val Leu Leu Asn
 565 570 575
 Gly Ile Tyr Glu His Pro Asn Ile Thr Ala Ile Val Trp Ala Gly Met
 580 585 590
 Pro Gly Glu Glu Ser Gly Asn Ala Leu Val Asp Ile Leu Trp Gly Asn

[0065]

595	600	605
Val Asn Pro Ala Gly Arg Thr Pro Phe Thr Trp Ala Lys Ser Arg Glu 610	615	620
Asp Tyr Gly Thr Asp Ile Met Tyr Glu Pro Asn Asn Gly Gln Arg Ala 625	630	635 640
Pro Gln Gln Asp Phe Thr Glu Ser Ile Tyr Leu Asp Tyr Arg His Phe 645	650	655
Asp Lys Ala Gly Ile Glu Pro Ile Tyr Glu Phe Gly Phe Gly Leu Ser 660	665	670
Tyr Thr Thr Phe Glu Tyr Ser Asp Leu Arg Val Val Lys Lys Tyr Val 675	680	685
Gln Pro Tyr Ser Pro Thr Thr Gly Thr Gly Ala Gln Ala Pro Ser Ile 690	695	700
Gly Gln Pro Pro Ser Gln Asn Leu Asp Thr Tyr Lys Phe Pro Ala Thr 705	710	715 720
Tyr Lys Tyr Ile Lys Thr Phe Ile Tyr Pro Tyr Leu Asn Ser Thr Val 725	730	735
Ser Leu Arg Ala Ala Ser Lys Asp Pro Glu Tyr Gly Arg Thr Asp Phe 740	745	750
Ile Pro Pro His Ala Arg Asp Gly Ser Pro Gln Pro Leu Asn Pro Ala 755	760	765
Gly Asp Pro Val Ala Ser Gly Gly Asn Asn Met Leu Tyr Asp Glu Leu 770	775	780
Tyr Glu Val Thr Ala Gln Ile Lys Asn Thr Gly Asp Val Ala Gly Asp 785	790	795 800
Glu Val Val Gln Leu Tyr Val Asp Leu Gly Gly Asp Asn Pro Pro Arg 805	810	815
Gln Leu Arg Asn Phe Asp Arg Phe Tyr Leu Leu Pro Gly Gln Ser Ser 820	825	830
Thr Phe Arg Ala Thr Leu Thr Arg Arg Asp Leu Ser Asn Trp Asp Ile 835	840	845
Glu Ala Gln Asn Trp Arg Val Thr Glu Ser Pro Lys Arg Val Tyr Val 850	855	860
Gly Arg Ser Ser Arg Asp Leu Pro Leu Ser Ser Gln Leu Glu 865	870	875

<210> 63

<211> 2583

<212> DNA

<213> 黑曲霉(Aspergillus niger)

<400> 63

atgaggttca ctttgatcga ggcggtggct ctgactgccg tctcgtggc cagcgtgat	60
gaattggcct actcccaccc gtattaccca tccccttggg ccaatggcca gggegactgg	120
gcgcaggcat accagcgcgc tgttgatatt gtctgcgaaa tgacattgga tgagaaggtc	180
aatctgacca caggaaactg atgggaattg gaactaigtg ttggtcagac tggegggtgtt	240
ccccgattgg gagttccggg aatgtgttta caggatagcc ctctgggcgt tcgcgactcc	300
gaetacaaet ctgctttccc tgceggcatg aacgtggetg caacctggga caagaatctg	360
gcataacctt cgcccaaggc tatgggtcag gaatttagtg acaagggtgc cgatatccaa	420

[0066]

ttgggtccag ctgccgccc tctcggtaga agtcccgcag gtggtcgtaa ctgggagggc	480
ttctccccag accctgccct aagtggigtg ctttttgcg agaccatcaa gggatccaa	540
gatgctgggtg tggttgcgac ggctaagcac tacattgctt acgagcaaga gcatttcgt	600
caggcgccctg aagcccaagg ttttggattt aatatttccg agagtggaa gtcgaacctc	660
gatgataaga ctatgcacga gctgtacctc tggeccctcg eggatgccat ccgtgcaggt	720
gctggcgctg tgatgtgtc ctacaaccag atcaacaaca gttatggctg ccagaacage	780
tacactctga acaagctgct caaggccgag ctgggtctcc agggctttgt catgagtgat	840
tgggtgctgc accatgctgg tgtgagtgtt gctttggcag gatggatat gtctatgcca	900
ggagacgtcg actacgacag tggtaacctc tactgggta caaacttgac cattagcgtg	960
ctcaacggaa cgggtcccca atggcgtgtt gatgacatgg ctgtccgcat catggccgcc	1020
tactacaagg tcggccgtga ccgtctgttg actcctccca acttcagctc atggaccaga	1080
gatgaatacg gctacaagta ctactacgtg tcggagggac cgtacgagaa ggtaaccag	1140
tacgtgaatg tgcaacgcaa ccacagcgaa ctgattcgcc gcattggagc ggacagcacg	1200
gtgtcctca agaaccgacgg cgtcttgcct ttgacttgta aggagcgctt ggtcgcgtt	1260
atcgagaag atgcgggctc caacccttat ggtgccaacg gctgcagtga ccgtggatgc	1320
gacaatggaa catfggcgat gggctgggga agtggtaactg ccaacttccc atacctggtg	1380
accccgagc aggcctctc aaacgaggtg ctttaagcaca agaattggtt attcaccgcc	1440
accgataact gggctatcga tcagattgag gcgtttgcta agaccgccag tgtctctctt	1500
gtctttgtca acgccgactc tggtaggggt tacatcaatg tggacggaaa cctgggtgac	1560
cgcaggaacc tgacctgtg gaggaacggc gataatgtga tcaaggtgc tgctagcaac	1620
tgcaacaaca caatcgtgt cattcactct gtcggaccag tcttggttaa cgagtggfac	1680
gacaacccca atgttaccgc tatcctctgg ggtggtttgc ccggtcagga gtctggcaac	1740
tctcttgcg acgtctctc tggecgtgc aaccccggtg ccaagtcgcc ctttacctg	1800
ggcaagactc gtgaggccta ccaagactac ttggtcaecg agcccaaca cggcaacgga	1860
gcccctcagg aagactttgt cgaggcgctc ttcattgact accgtggatt tgacaagcgc	1920
aacgagaacc cgaictacga gttcggctat ggtctgagct acaccacttt caactactcg	1980
aaccttgagg tgcaggtgct gagegcccc gctacgagc ctgttcggg tgagaccgag	2040
gcagcgccaa ccttcggaga ggttggaaat gcgtcggatt acctctacc cagcggattg	2100
cagagaaata ccaagttcat ctacccctgg ctcaacgta ccgactcga ggcactctcc	2160
gggatgcta gctacggca ggactcctcc gactatctc ccgagggagc caccgatggc	2220
tctgcgaac cgactctgcc tgccggiggc ggtcttggcg gcaaccctcg cctgtacgac	2280
gagctcatcc gcgtgtcagt gaccatcaag aacaccggca aggttctgtg tgatgaagtt	2340
ccccactgt atgtttccct tggcggtccc aatgagccca agatcgtgct gcgtcaattc	2400
gagcgcatca cgctgcagcc gtcggaggag acgaagtga gcacgactct gacgcgccgt	2460
gaccttgcaa actggaatgt tgagaagcag gactgggaga ttacgtcgt tcccaagatg	2520
gtgtttgtcg gaagctctc gcggaagctg ccgtccggg cgtctctgcc tactgttca	2580

[0067]

taa 2583

<210> 64
 <211> 860
 <212> PRT
 <213> 黑曲霉(*Aspergillus niger*)

<400> 64

Met Arg Phe Thr Leu Ile Glu Ala Val Ala Leu Thr Ala Val Ser Leu
 1 5 10 15

Ala Ser Ala Asp Glu Leu Ala Tyr Ser Pro Pro Tyr Tyr Pro Ser Pro
 20 25 30

Trp Ala Asn Gly Gln Gly Asp Trp Ala Gln Ala Tyr Gln Arg Ala Val
 35 40 45

Asp Ile Val Ser Gln Met Thr Leu Asp Glu Lys Val Asn Leu Thr Thr
 50 55 60

Gly Thr Gly Trp Glu Leu Glu Leu Cys Val Gly Gln Thr Gly Gly Val
 65 70 75 80

Pro Arg Leu Gly Val Pro Gly Met Cys Leu Gln Asp Ser Pro Leu Gly
 85 90 95

Val Arg Asp Ser Asp Tyr Asn Ser Ala Phe Pro Ala Gly Met Asn Val
 100 105 110

Ala Ala Thr Trp Asp Lys Asn Leu Ala Tyr Leu Arg Gly Lys Ala Met
 115 120 125

Gly Gln Glu Phe Ser Asp Lys Gly Ala Asp Ile Gln Leu Gly Pro Ala
 130 135 140

Ala Gly Pro Leu Gly Arg Ser Pro Asp Gly Gly Arg Asn Trp Glu Gly
 145 150 155 160

Phe Ser Pro Asp Pro Ala Leu Ser Gly Val Leu Phe Ala Glu Thr Ile
 165 170 175

Lys Gly Ile Gln Asp Ala Gly Val Val Ala Thr Ala Lys His Tyr Ile
 180 185 190

Ala Tyr Glu Gln Glu His Phe Arg Gln Ala Pro Glu Ala Gln Gly Phe
 195 200 205

Gly Phe Asn Ile Ser Glu Ser Gly Ser Ala Asn Leu Asp Asp Lys Thr
 210 215 220

Met His Glu Leu Tyr Leu Trp Pro Phe Ala Asp Ala Ile Arg Ala Gly
 225 230 235 240

Ala Gly Ala Val Met Cys Ser Tyr Asn Gln Ile Asn Asn Ser Tyr Gly
 245 250 255

Cys Gln Asn Ser Tyr Thr Leu Asn Lys Leu Leu Lys Ala Glu Leu Gly
 260 265 270

Phe Gln Gly Phe Val Met Ser Asp Trp Ala Ala His His Ala Gly Val
 275 280 285

Ser Gly Ala Leu Ala Gly Leu Asp Met Ser Met Pro Gly Asp Val Asp
 290 295 300

Tyr Asp Ser Gly Thr Ser Tyr Trp Gly Thr Asn Leu Thr Ile Ser Val
 305 310 315 320

Leu Asn Gly Thr Val Pro Gln Trp Arg Val Asp Asp Met Ala Val Arg
 325 330 335

[0068]

```

Ile Met Ala Ala Tyr Tyr Lys Val Gly Arg Asp Arg Leu Trp Thr Pro
340                               345                               350
Pro Asn Phe Ser Ser Trp Thr Arg Asp Glu Tyr Gly Tyr Lys Tyr Tyr
355                               360                               365
Tyr Val Ser Glu Gly Pro Tyr Glu Lys Val Asn Gln Tyr Val Asn Val
370                               375                               380
Gln Arg Asn His Ser Glu Leu Ile Arg Arg Ile Gly Ala Asp Ser Thr
385                               390                               395                               400
Val Leu Leu Lys Asn Asp Gly Ala Leu Pro Leu Thr Gly Lys Glu Arg
405                               410                               415
Leu Val Ala Leu Ile Gly Glu Asp Ala Gly Ser Asn Pro Tyr Gly Ala
420                               425                               430
Asn Gly Cys Ser Asp Arg Gly Cys Asp Asn Gly Thr Leu Ala Met Gly
435                               440                               445
Trp Gly Ser Gly Thr Ala Asn Phe Pro Tyr Leu Val Thr Pro Glu Gln
450                               455                               460
Ala Ile Ser Asn Glu Val Leu Lys His Lys Asn Gly Val Phe Thr Ala
465                               470                               475                               480
Thr Asp Asn Trp Ala Ile Asp Gln Ile Glu Ala Leu Ala Lys Thr Ala
485                               490                               495
Ser Val Ser Leu Val Phe Val Asn Ala Asp Ser Gly Glu Gly Tyr Ile
500                               505                               510
Asn Val Asp Gly Asn Leu Gly Asp Arg Arg Asn Leu Thr Leu Trp Arg
515                               520                               525
Asn Gly Asp Asn Val Ile Lys Ala Ala Ala Ser Asn Cys Asn Asn Thr
530                               535                               540
Ile Val Val Ile His Ser Val Gly Pro Val Leu Val Asn Glu Trp Tyr
545                               550                               555                               560
Asp Asn Pro Asn Val Thr Ala Ile Leu Trp Gly Gly Leu Pro Gly Gln
565                               570                               575
Glu Ser Gly Asn Ser Leu Ala Asp Val Leu Tyr Gly Arg Val Asn Pro
580                               585                               590
Gly Ala Lys Ser Pro Phe Thr Trp Gly Lys Thr Arg Glu Ala Tyr Gln
595                               600                               605
Asp Tyr Leu Val Thr Glu Pro Asn Asn Gly Asn Gly Ala Pro Gln Glu
610                               615                               620
Asp Phe Val Glu Gly Val Phe Ile Asp Tyr Arg Gly Phe Asp Lys Arg
625                               630                               635                               640
Asn Glu Thr Pro Ile Tyr Glu Phe Gly Tyr Gly Leu Ser Tyr Thr Thr
645                               650                               655
Phe Asn Tyr Ser Asn Leu Glu Val Gln Val Leu Ser Ala Pro Ala Tyr
660                               665                               670
Glu Pro Ala Ser Gly Glu Thr Glu Ala Ala Pro Thr Phe Gly Glu Val
675                               680                               685
Gly Asn Ala Ser Asp Tyr Leu Tyr Pro Ser Gly Leu Gln Arg Ile Thr
690                               695                               700
Lys Phe Ile Tyr Pro Trp Leu Asn Gly Thr Asp Leu Glu Ala Ser Ser
705                               710                               715                               720
Gly Asp Ala Ser Tyr Gly Gln Asp Ser Ser Asp Tyr Leu Pro Glu Gly

```

[0069]

725	730	735
Ala Thr Asp Gly Ser Ala Gln Pro Ile Leu Pro Ala Gly Gly Gly Pro 740	745	750
Gly Gly Asn Pro Arg Leu Tyr Asp Glu Leu Ile Arg Val Ser Val Thr 755	760	765
Ile Lys Asn Thr Gly Lys Val Ala Gly Asp Glu Val Pro Gln Leu Tyr 770	775	780
Val Ser Leu Gly Gly Pro Asn Glu Pro Lys Ile Val Leu Arg Gln Phe 785	790	800
Glu Arg Ile Thr Leu Gln Pro Ser Glu Glu Thr Lys Trp Ser Thr Thr 805	810	815
Leu Thr Arg Arg Asp Leu Ala Asn Trp Asn Val Glu Lys Gln Asp Trp 820	825	830
Glu Ile Thr Ser Tyr Pro Lys Met Val Phe Val Gly Ser Ser Ser Arg 835	840	845
Lys Leu Pro Leu Arg Ala Ser Leu Pro Thr Val His 850	855	860
<210> 65 <211> 2583 <212> DNA <213> 棘孢曲霉 (<i>Aspergillus aculeatus</i>)		
<400> 65		
atgaagctca gttgcttga ggcggctgcc ttgacggctg cttcagtcgt cagcgtgat	60	
gaactggcgt tctctcctcc ttctacccc tctccgiggc ccaatggcca gggagagtgg	120	
gcggaagcct accagcgtgc agtggccatt gtatccaga tgactctgga tgagaaggtc	180	
aacctgacca ccggaactgg atgggagctg gagaagtgcg tcggtcagac tgggtgtgtc	240	
ccaagaactg acatcggtgg catgtgtctt caggacagtc ccttggaat tcgtgatagt	300	
gactacaatt cggtttccc tctgtgtgtc aacgttctg cgacatggga caagaacctt	360	
gcttatctac gtggtcaggc tatgggtcaa gatttcagtg acaaggaaat tgatgttcaa	420	
ttgggacagg ccggcggtcc cctcggcagg agccctgatg gaggtcgcaa ctgggaaggt	480	
ttctctccag acccggtctt tacttggtgtg ctctttgcgg agacgattaa ggggtattcaa	540	
gacgtggtg tcgtggcgac agccaagcat tacattctca atgagcaaga gcatttccgc	600	
caggtcgcag aggttcgggg ctacggattc aatctctccg acacgatcag ctctaacttt	660	
gatgacaaga caattcatga aatgtacctc tggcccttcg cggatgccgt tcgcgcgggc	720	
gttggcgcca tcatgtgttc ctacaaccag atcaacaaca gctacggttg ccagaacagt	780	
tacactctga acaagcttct gaagccgag ctoggtcttc aggtctttgt gatgtctgac	840	
tgggtgtctc accacagtgg tgttggtctt gctttggcgg gcttgatat gtaaatgcct	900	
ggcgatalea ccttcgattc tgccactagt ttctggggta ccaacctgac cattgtctgt	960	
ctcaacggta ccgtcccgca gtggcggtt gacgacatgg ctgtccgtat catggtgtgc	1020	
tactacaagg ttggcccgca ccgctgtac cagccgccta acttcagctc ctggactcgc	1080	
gatgaatacg gcttcaagta ttctacccc caggaagggc cctatgagaa ggtcaatcac	1140	
tttgtcaatg tgcagcgcaa ccacagcgag gttattcgca agttgggagc agacagtaact	1200	

[0070]

gttctactga agaacaacaa tgccttgcgc ctgaccggaa aggagcgcaa agttgcgatic	1260
ctgggtgaag atgctggatc caactcgtac ggtgccaatg gctgctctga ccgtggctgt	1320
gacaacggta ctcttgctat ggcttggggg agcggcactg ccgaattccc atatctcgtg	1380
acccctgagc aggcatttca agccgagggt ctcaagcata agggcagcgt ctacgccatc	1440
acggacaact gggcgctgag ccagggtggag accctcgcta aacaagccag tgtctctctt	1500
gtatttgta actcggacgc gggagagggc tatatctecg tggacggaaa cgaggcgac	1560
cgcaacaace tcacctctg gaagaacggc gacaacctca tcaaggetgc tgcaacaac	1620
tgaacaaca ccactgtgt caicccatcc gttggaccig ttttgggtga cgagiggat	1680
gaccacccca acgttactgc catctcttgg ggggcttgc ctggccagga gtctggcaac	1740
tcttggctg acgtgcteta cggccgcgtc aaccggggcg ccaaatctcc attcaccitg	1800
ggcaagacga gggaggcgta cggggattac ctgtccgtg agctcaacaa cgcaacggga	1860
gtcccccaag atgatttctc ggaaggigt ttcattgact accgcggatt cgacaagcgc	1920
aatgagacc cgaatcaga gtccggacat ggtctgagct acaccacttt caactactct	1980
ggccttcaac tccaggttct caacgcctcc tccaacgctc aagtagccac tgagactggc	2040
gcgctccca ccttcggaca agtcggcaat gcctctgact acgtgtaccc tgagggtattg	2100
accagaatca gcaagttcat ctatccctgg cttaattcca cagacctgaa ggcctcatct	2160
ggcgaccctg actatggagt cgacacgcgc gageacgtgc ccgagggtgc tactgatggc	2220
tctccgcage ccgttctgcc tgcgggtggg ggtctctggg gtaacccgcg cctctacgat	2280
gagttagtcc gtgtttcggg gacagtcaag aacactggtc gtgttgcgg tgatctgtg	2340
ccfcaatigt atgtttccct tggtagacc ccatgagccca aggttgigt ggcgaattc	2400
gaccgcctca cctcaagcc ctccgaggag acggtgtgga cgactacct gaccgcgcg	2460
gatctgteta actgggaagt tgcggctcag gactgggtca tcacttctta ccgaagaag	2520
gtccatgttg gtagctcttc ggtcagctg ccccttcacg cggcgctccc gaaggtgcaa	2580
tga	2583

<210> 66
 <211> 860
 <212> PRT
 <213> 棘孢曲霉(Aspergillus aculeatus)

<400> 66

```

Met Lys Leu Ser Trp Leu Glu Ala Ala Ala Leu Thr Ala Ala Ser Val
1           5           10           15

Val Ser Ala Asp Glu Leu Ala Phe Ser Pro Pro Phe Tyr Pro Ser Pro
20           25           30

Trp Ala Asn Gly Gln Gly Glu Trp Ala Glu Ala Tyr Gln Arg Ala Val
35           40           45

Ala Ile Val Ser Gln Met Thr Leu Asp Glu Lys Val Asn Leu Thr Thr
50           55           60

Gly Thr Gly Trp Glu Leu Glu Lys Cys Val Gly Gln Thr Gly Gly Val
65           70           75           80

Pro Arg Leu Asn Ile Gly Gly Met Cys Leu Gln Asp Ser Pro Leu Gly
85           90           95

```

[0071]

```

Ile Arg Asp Ser Asp Tyr Asn Ser Ala Phe Pro Ala Gly Val Asn Val
100                               105                               110

Ala Ala Thr Trp Asp Lys Asn Leu Ala Tyr Leu Arg Gly Gln Ala Met
115                               120                               125

Gly Gln Glu Phe Ser Asp Lys Gly Ile Asp Val Gln Leu Gly Pro Ala
130                               135                               140

Ala Gly Pro Leu Gly Arg Ser Pro Asp Gly Gly Arg Asn Trp Glu Gly
145                               150                               155                               160

Phe Ser Pro Asp Pro Ala Leu Thr Gly Val Leu Phe Ala Glu Thr Ile
165                               170                               175

Lys Gly Ile Gln Asp Ala Gly Val Val Ala Thr Ala Lys His Tyr Ile
180                               185                               190

Leu Asn Glu Gln Glu His Phe Arg Gln Val Ala Glu Ala Ala Gly Tyr
195                               200                               205

Gly Phe Asn Ile Ser Asp Thr Ile Ser Ser Asn Val Asp Asp Lys Thr
210                               215                               220

Ile His Glu Met Tyr Leu Trp Pro Phe Ala Asp Ala Val Arg Ala Gly
225                               230                               235                               240

Val Gly Ala Ile Met Cys Ser Tyr Asn Gln Ile Asn Asn Ser Tyr Gly
245                               250                               255

Cys Gln Asn Ser Tyr Thr Leu Asn Lys Leu Leu Lys Ala Glu Leu Gly
260                               265                               270

Phe Gln Gly Phe Val Met Ser Asp Trp Gly Ala His His Ser Gly Val
275                               280                               285

Gly Ser Ala Leu Ala Gly Leu Asp Met Ser Met Pro Gly Asp Ile Thr
290                               295                               300

Phe Asp Ser Ala Thr Ser Phe Trp Gly Thr Asn Leu Thr Ile Ala Val
305                               310                               315                               320

Leu Asn Gly Thr Val Pro Gln Trp Arg Val Asp Asp Met Ala Val Arg
325                               330                               335

Ile Met Ala Ala Tyr Tyr Lys Val Gly Arg Asp Arg Leu Tyr Gln Pro
340                               345                               350

Pro Asn Phe Ser Ser Trp Thr Arg Asp Glu Tyr Gly Phe Lys Tyr Phe
355                               360                               365

Tyr Pro Gln Glu Gly Pro Tyr Glu Lys Val Asn His Phe Val Asn Val
370                               375                               380

Gln Arg Asn His Ser Glu Val Ile Arg Lys Leu Gly Ala Asp Ser Thr
385                               390                               395                               400

Val Leu Leu Lys Asn Asn Asn Ala Leu Pro Leu Thr Gly Lys Glu Arg
405                               410                               415

Lys Val Ala Ile Leu Gly Glu Asp Ala Gly Ser Asn Ser Tyr Gly Ala
420                               425                               430

Asn Gly Cys Ser Asp Arg Gly Cys Asp Asn Gly Thr Leu Ala Met Ala
435                               440                               445

Trp Gly Ser Gly Thr Ala Glu Phe Pro Tyr Leu Val Thr Pro Glu Gln
450                               455                               460

Ala Ile Gln Ala Glu Val Leu Lys His Lys Gly Ser Val Tyr Ala Ile
465                               470                               475                               480

```

[0072]

Thr Asp Asn Trp Ala Leu Ser Gln Val Glu Thr Leu Ala Lys Gln Ala
 485 490 495
 Ser Val Ser Leu Val Phe Val Asn Ser Asp Ala Gly Glu Gly Tyr Ile
 500 505 510
 Ser Val Asp Gly Asn Glu Gly Asp Arg Asn Asn Leu Thr Leu Trp Lys
 515 520 525
 Asn Gly Asp Asn Leu Ile Lys Ala Ala Ala Asn Asn Cys Asn Asn Thr
 530 535 540
 Ile Val Val Ile His Ser Val Gly Pro Val Leu Val Asp Glu Trp Tyr
 545 550 555 560
 Asp His Pro Asn Val Thr Ala Ile Leu Trp Ala Gly Leu Pro Gly Gln
 565 570 575
 Glu Ser Gly Asn Ser Leu Ala Asp Val Leu Tyr Gly Arg Val Asn Pro
 580 585 590
 Gly Ala Lys Ser Pro Phe Thr Trp Gly Lys Thr Arg Glu Ala Tyr Gly
 595 600 605
 Asp Tyr Leu Val Arg Glu Leu Asn Asn Gly Asn Gly Ala Pro Gln Asp
 610 615 620
 Asp Phe Ser Glu Gly Val Phe Ile Asp Tyr Arg Gly Phe Asp Lys Arg
 625 630 635 640
 Asn Glu Thr Pro Ile Tyr Glu Phe Gly His Gly Leu Ser Tyr Thr Thr
 645 650 655
 Phe Asn Tyr Ser Gly Leu His Ile Gln Val Leu Asn Ala Ser Ser Asn
 660 665 670
 Ala Gln Val Ala Thr Glu Thr Gly Ala Ala Pro Thr Phe Gly Gln Val
 675 680 685
 Gly Asn Ala Ser Asp Tyr Val Tyr Pro Glu Gly Leu Thr Arg Ile Ser
 690 695 700
 Lys Phe Ile Tyr Pro Trp Leu Asn Ser Thr Asp Leu Lys Ala Ser Ser
 705 710 715 720
 Gly Asp Pro Tyr Tyr Gly Val Asp Thr Ala Glu His Val Pro Glu Gly
 725 730 735
 Ala Thr Asp Gly Ser Pro Gln Pro Val Leu Pro Ala Gly Gly Gly Ser
 740 745 750
 Gly Gly Asn Pro Arg Leu Tyr Asp Glu Leu Ile Arg Val Ser Val Thr
 755 760 765
 Val Lys Asn Thr Gly Arg Val Ala Gly Asp Ala Val Pro Gln Leu Tyr
 770 775 780
 Val Ser Leu Gly Gly Pro Asn Glu Pro Lys Val Val Leu Arg Lys Phe
 785 790 795 800
 Asp Arg Leu Thr Leu Lys Pro Ser Glu Glu Thr Val Trp Thr Thr Thr
 805 810 815
 Leu Thr Arg Arg Asp Leu Ser Asn Trp Asp Val Ala Ala Gln Asp Trp
 820 825 830
 Val Ile Thr Ser Tyr Pro Lys Lys Val His Val Gly Ser Ser Ser Arg
 835 840 845
 Gln Leu Pro Leu His Ala Ala Leu Pro Lys Val Gln
 850 855 860

[0073]

<210>	67					
<211>	3294					
<212>	DNA					
<213>	米曲霉(<i>Aspergillus oryzae</i>)					
<400>	67					
atgcgttctt	ccccctctt	ccgctccgcc	gttgtggccg	ccttgcgggt	gttggccctt	60
gccgctgatg	gcaggctccac	ccgctactgg	gactgtctga	agccttcgtg	cggtcgggcc	120
aagaaggctc	ccgtgaacca	gcctgtcttt	tcttgcaacg	ccaaactcca	gcgtatcacg	180
gaattcgacg	ccaagtccgg	ctgcgagccg	ggcgggtgtc	cctactcgtg	cgccgaccag	240
accccatggg	ctgtgaacga	cgacttcgcg	ctcggttttg	ctgccacctc	tattgccggc	300
agcaatgagg	cgggctggtg	ctgcgectgc	tacgagctca	ccttcacatc	cggtctcgtt	360
gctggcaaga	agatggctct	ccagtcacac	agcacigggc	gtgatcttgg	cagcaaccac	420
ttcgatctca	acatcccggg	cggcggcgct	ggcatcttcg	acggatgcac	tccccagttc	480
ggttggtctg	ccggccagcg	ctacggcggc	atctcgtccc	gcaacgagtg	cgatcggttc	540
ccccagcccc	tcaagccggg	ctgctactgg	cgttcgact	ggttcaagaa	cgcgcgacaat	600
ccgagcttca	gcttccgtca	ggctccagtc	ccagccgagc	tcgtcgtctg	caccggatgc	660
cgccgcaacg	acgacggcaa	cttcctctgc	gtccagatcc	ccatgcgttc	ctccccctc	720
ctccgctccg	ccgttgtggc	cgcctctccg	gttgtggccc	ttgccaagga	tgatctcgcg	780
tactccccct	ctttctaccc	tcccccatgg	gcagatggtc	aggggtgaatg	ggcgggaagta	840
tacaaacgcg	ctgttagacat	agtttcccag	atgacgttga	cagagaaagt	caacttaacg	900
actggaacag	gatggcaact	agagagggtg	gttggacaaa	ctggcagtgt	tcccagactc	960
aacatcccca	gcttgtgttt	gcaggatagt	cctcttggtt	ttcgtttctc	ggactacaat	1020
tcagctttcc	ctcggggtgt	taatgtcgct	gccaccgggg	acaagacgct	cgcctacctt	1080
cgttggtcagg	caatgggtga	ggagttcagt	gataagggta	ttgacgttca	gctgggtcct	1140
gctgtgtggc	ctctcggtgc	tcatecggat	ggcggtagaa	actgggaagg	tttctcacca	1200
gatccagccc	tcaccggtgt	actttttgcg	gagacgatta	agggtattca	agatgtgtgt	1260
gtcattgcga	cagctaagca	ttatatcatg	aacgaacaag	agcattttcc	ccaacaaccc	1320
gaggetgcgg	gttacggatt	caacgtaagc	gacagtttga	gttccaacgt	tgatgacaag	1380
actatgcatg	aattgtacct	ctggcccttc	gcggatgcag	tacgcgctgg	agtcggtgct	1440
gtcatgtgt	cttacaacca	aateacaac	agctacgggt	gcgagaatag	cgaaactctg	1500
aacaagcttt	tgaaggcgga	gtttgggttc	caaggcttcg	tcatgagtga	ttggaccgct	1560
catcacagcg	gcgtaggcgc	tgccttagca	ggtctggata	tgtcgatgcc	cggtgatgtt	1620
accttcgata	gtggtacgtc	ttctcgggg	gcaaacitga	cggtcgggtg	ccttaacggt	1680
acaatccccc	aatggcggtg	tgatgacatg	gtgtccgta	tcatggccgc	ttattacaag	1740
gttggccgcg	acaccaata	cacctctccc	aacttcagct	cgtggaccag	ggacgaatat	1800
ggttttcgcg	ataacctagt	ttcggaaggt	gcttaegaga	gggtcaacga	attcgtggac	1860
gtgcaacgcg	atcatgccga	cctaaccgt	cgcacggcg	cgcagagcac	tgttctgctg	1920
aagaacaagg	gtgccttgcc	cttgagccgc	aaggaaaagc	tggctgcctt	tctgggagag	1980
gatgcgggtt	ccaactcgtg	ggcgctaac	ggcgtgatg	accgtgggtt	cgataacggt	2040

[0074]

```

acccttgcca tggcctgggg tagcgtact gcgaattfcc catacctcgt gacaccagag 2100
caggegaatc agaacgaagt tcttcagggc cgtggtaatg tcttcgccgt gaccgacagt 2160
tgggcgctcg acaagatcgc tgcggctgcc cgcagagcca gcgtatctct cgtgttcgtc 2220
aactccgact caggagaagg ctatcttagt gtggatggaa atgagggcga tcgtaacaac 2280
atcactctgt ggaagaacgg cgacaatgtg gtcaagaccg cagcgaataa ctgtaacaac 2340
accgttgcca tcatccactc cgtcggacca gttttgatcg atgaatggta tgaccacccc 2400
aatgtcactg gtattctctg ggctggctcg ccaggccagg agtctggtaa ctcattgccc 2460
gatgtgctgt acggtcgtgt caacctggc gccaaagtctc ctctcacttg gggcaagacc 2520
cgggagtcgt atggtttctc ctgggtcaag gatgccaaca atggcaacgg agcgcgccag 2580
tctgatttca ccagggtgt tttcatcgat taccgccatt tcgataagtt caatgagacc 2640
ctatctacg agtttgcta cgcttgagc tacaccacct tcgagctctc cgacctccat 2700
gttcagcccc tgaacgcgtc ccgatacact ccaccagtg gcatactga agctgcnaag 2760
aactttgggt aaattggcga tgcgtcggag tacgtgtatc cggaggggct ggaaggatc 2820
catgagtfta tctatccctg gatcaactct accgacctga aggcacgtc tgacgattct 2880
aactacggct gggaagactc caagtatatt ccgaaggcg ccacggatgg gtctgcccag 2940
ccccgttgc ccgctagtgg tgggtccgga ggaaacccc gtctgtacga ggatcttttc 3000
cgctctctg tgaaggtaaa gaacacgggc aatgtcgccg gtgatgaagt tcttcagctg 3060
tacgtttccc taggggccc gaatgagccc aagggtgtac tgcgcaagtt tgagcgtatt 3120
cacttgcccc ctctgcagga ggccgtgtgg acaacgcccc ttaccgctcg tgaccttgca 3180
aactgggacg ttctggctca ggactggacc gtcaactctt accecaagac gatctacgtt 3240
ggaaactcct caccgaaact gccgtccag gccctcgtgc ctaaggccca gtaa 3294

```

<210> 68
 <211> 1097
 <212> PRT
 <213> 米曲霉(*Aspergillus oryzae*)

<400> 68

```

Met Arg Ser Ser Pro Leu Leu Arg Ser Ala Val Val Ala Ala Leu Pro
1           5           10           15
Val Leu Ala Leu Ala Ala Asp Gly Arg Ser Thr Arg Tyr Trp Asp Cys
20          25          30
Cys Lys Pro Ser Cys Gly Trp Ala Lys Lys Ala Pro Val Asn Gln Pro
35          40          45
Val Phe Ser Cys Asn Ala Asn Phe Gln Arg Ile Thr Asp Phe Asp Ala
50          55          60
Lys Ser Gly Cys Glu Pro Gly Gly Val Ala Tyr Ser Cys Ala Asp Gln
65          70          75          80
Thr Pro Trp Ala Val Asn Asp Asp Phe Ala Leu Gly Phe Ala Ala Thr
85          90          95
Ser Ile Ala Gly Ser Asn Glu Ala Gly Trp Cys Cys Ala Cys Tyr Glu
100         105         110
Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Lys Met Val Val Gln

```

[0075]

115	120	125
Ser Thr Ser Thr Gly Gly Asp Leu Gly Ser Asn His Phe Asp Leu Asn 130 135 140		
Ile Pro Gly Gly Gly Val Gly Ile Phe Asp Gly Cys Thr Pro Gln Phe 145 150 155 160		
Gly Gly Leu Pro Gly Gln Arg Tyr Gly Gly Ile Ser Ser Arg Asn Glu 165 170 175		
Cys Asp Arg Phe Pro Asp Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe 180 185 190		
Asp Trp Phe Lys Asn Ala Asp Asn Pro Ser Phe Ser Phe Arg Gln Val 195 200 205		
Gln Cys Pro Ala Glu Leu Val Ala Arg Thr Gly Cys Arg Arg Asn Asp 210 215 220		
Asp Gly Asn Phe Pro Ala Val Gln Ile Pro Met Arg Ser Ser Pro Leu 225 230 235 240		
Leu Arg Ser Ala Val Val Ala Ala Leu Pro Val Leu Ala Leu Ala Lys 245 250 255		
Asp Asp Leu Ala Tyr Ser Pro Pro Phe Tyr Pro Ser Pro Trp Ala Asp 260 265 270		
Gly Gln Gly Glu Trp Ala Glu Val Tyr Lys Arg Ala Val Asp Ile Val 275 280 285		
Ser Gln Met Thr Leu Thr Glu Lys Val Asn Leu Thr Thr Gly Thr Gly 290 295 300		
Trp Gln Leu Glu Arg Cys Val Gly Gln Thr Gly Ser Val Pro Arg Leu 305 310 315 320		
Asn Ile Pro Ser Leu Cys Leu Gln Asp Ser Pro Leu Gly Ile Arg Phe 325 330 335		
Ser Asp Tyr Asn Ser Ala Phe Pro Ala Gly Val Asn Val Ala Ala Thr 340 345 350		
Trp Asp Lys Thr Leu Ala Tyr Leu Arg Gly Gln Ala Met Gly Glu Glu 355 360 365		
Phe Ser Asp Lys Gly Ile Asp Val Gln Leu Gly Pro Ala Ala Gly Pro 370 375 380		
Leu Gly Ala His Pro Asp Gly Gly Arg Asn Trp Glu Gly Phe Ser Pro 385 390 395 400		
Asp Pro Ala Leu Thr Gly Val Leu Phe Ala Glu Thr Ile Lys Gly Ile 405 410 415		
Gln Asp Ala Gly Val Ile Ala Thr Ala Lys His Tyr Ile Met Asn Glu 420 425 430		
Gln Glu His Phe Arg Gln Gln Pro Glu Ala Ala Gly Tyr Gly Phe Asn 435 440 445		
Val Ser Asp Ser Leu Ser Ser Asn Val Asp Asp Lys Thr Met His Glu 450 455 460		
Leu Tyr Leu Trp Pro Phe Ala Asp Ala Val Arg Ala Gly Val Gly Ala 465 470 475 480		
Val Met Cys Ser Tyr Asn Gln Ile Asn Asn Ser Tyr Gly Cys Glu Asn 485 490 495		
Ser Glu Thr Leu Asn Lys Leu Leu Lys Ala Glu Leu Gly Phe Gln Gly 500 505 510		

[0076]

Phe Val Met Ser Asp Trp Thr Ala His His Ser Gly Val Gly Ala Ala
 515 520 525
 Leu Ala Gly Leu Asp Met Ser Met Pro Gly Asp Val Thr Phe Asp Ser
 530 535 540
 Gly Thr Ser Phe Trp Gly Ala Asn Leu Thr Val Gly Val Leu Asn Gly
 545 550 555 560
 Thr Ile Pro Gln Trp Arg Val Asp Asp Met Ala Val Arg Ile Met Ala
 565 570 575
 Ala Tyr Tyr Lys Val Gly Arg Asp Thr Lys Tyr Thr Pro Pro Asn Phe
 580 585 590
 Ser Ser Trp Thr Arg Asp Glu Tyr Gly Phe Ala His Asn His Val Ser
 595 600 605
 Glu Gly Ala Tyr Glu Arg Val Asn Glu Phe Val Asp Val Gln Arg Asp
 610 615 620
 His Ala Asp Leu Ile Arg Arg Ile Gly Ala Gln Ser Thr Val Leu Leu
 625 630 635 640
 Lys Asn Lys Gly Ala Leu Pro Leu Ser Arg Lys Glu Lys Leu Val Ala
 645 650 655
 Leu Leu Gly Glu Asp Ala Gly Ser Asn Ser Trp Gly Ala Asn Gly Cys
 660 665 670
 Asp Asp Arg Gly Cys Asp Asn Gly Thr Leu Ala Met Ala Trp Gly Ser
 675 680 685
 Gly Thr Ala Asn Phe Pro Tyr Leu Val Thr Pro Glu Gln Ala Ile Gln
 690 695 700
 Asn Glu Val Leu Gln Gly Arg Gly Asn Val Phe Ala Val Thr Asp Ser
 705 710 715 720
 Trp Ala Leu Asp Lys Ile Ala Ala Ala Ala Arg Gln Ala Ser Val Ser
 725 730 735
 Leu Val Phe Val Asn Ser Asp Ser Gly Glu Gly Tyr Leu Ser Val Asp
 740 745 750
 Gly Asn Glu Gly Asp Arg Asn Asn Ile Thr Leu Trp Lys Asn Gly Asp
 755 760 765
 Asn Val Val Lys Thr Ala Ala Asn Asn Cys Asn Asn Thr Val Val Ile
 770 775 780
 Ile His Ser Val Gly Pro Val Leu Ile Asp Glu Trp Tyr Asp His Pro
 785 790 795 800
 Asn Val Thr Gly Ile Leu Trp Ala Gly Leu Pro Gly Gln Glu Ser Gly
 805 810 815
 Asn Ser Ile Ala Asp Val Leu Tyr Gly Arg Val Asn Pro Gly Ala Lys
 820 825 830
 Ser Pro Phe Thr Trp Gly Lys Thr Arg Glu Ser Tyr Gly Ser Pro Leu
 835 840 845
 Val Lys Asp Ala Asn Asn Gly Asn Gly Ala Pro Gln Ser Asp Phe Thr
 850 855 860
 Gln Gly Val Phe Ile Asp Tyr Arg His Phe Asp Lys Phe Asn Glu Thr
 865 870 875 880
 Pro Ile Tyr Glu Phe Gly Tyr Gly Leu Ser Tyr Thr Thr Phe Glu Leu
 885 890 895

[0077]

Ser Asp Leu His Val Gln Pro Leu Asn Ala Ser Arg Tyr Thr Pro Thr
 900 905 910
 Ser Gly Met Thr Glu Ala Ala Lys Asn Phe Gly Glu Ile Gly Asp Ala
 915 920 925
 Ser Glu Tyr Val Tyr Pro Glu Gly Leu Glu Arg Ile His Glu Phe Ile
 930 935 940
 Tyr Pro Trp Ile Asn Ser Thr Asp Leu Lys Ala Ser Ser Asp Asp Ser
 945 950 955 960
 Asn Tyr Gly Trp Glu Asp Ser Lys Tyr Ile Pro Glu Gly Ala Thr Asp
 965 970 975
 Gly Ser Ala Gln Pro Arg Leu Pro Ala Ser Gly Gly Ala Gly Gly Asn
 980 985 990
 Pro Gly Leu Tyr Glu Asp Leu Phe Arg Val Ser Val Lys Val Lys Asn
 995 1000 1005
 Thr Gly Asn Val Ala Gly Asp Glu Val Pro Gln Leu Tyr Val Ser
 1010 1015 1020
 Leu Gly Gly Pro Asn Glu Pro Lys Val Val Leu Arg Lys Phe Glu
 1025 1030 1035
 Arg Ile His Leu Ala Pro Ser Gln Glu Ala Val Trp Thr Thr Thr
 1040 1045 1050
 Leu Thr Arg Arg Asp Leu Ala Asn Trp Asp Val Ser Ala Gln Asp
 1055 1060 1065
 Trp Thr Val Thr Pro Tyr Pro Lys Thr Ile Tyr Val Gly Asn Ser
 1070 1075 1080
 Ser Arg Lys Leu Pro Leu Gln Ala Ser Leu Pro Lys Ala Gln
 1085 1090 1095

<210> 69

<211> 3294

<212> DNA

<213> 米曲霉(Aspergillus oryzae)

<400> 69

atgcgttccct ccccccctct ccgcctccgcc gttgtggecg cctgcccgt gttggccctt 60
 gccgctgatg gcaggtccac ccgcctactgg gactgetgca agccttcgtg cggetgggcc 120
 aagaaggctc ccgtgaacea gcctgtcttt tcttgeaacg ccaacttcca gcgtatcacg 180
 gacttcgacg ccaagtccgg ctgcgagccg ggcgggtgic cctactcgtg cgccgaccag 240
 accccatggg ctgtgaacga cgacttcgeg ctgggttttg ctgccacctc tattgcgggc 300
 agcaatgagg cgggctgggt ctgugccfge tacgagctca ctttcacatc cggfcccgtt 360
 gctggcaaga agatggctgt ccagtcacac agcactggcg gtgatcttgg cagcaaccac 420
 ttgatctca acatcccgg cgccggcgct ggcctctcg acggatgcac tccccagttc 480
 ggtggctctg cggccagcg ctacggcggc atctegtcgc gcaacgagtg cgateggttc 540
 cccgacgccc tcaagcccgg ctgcctactgg cgcttcgact ggttcaagaa cgccgacaat 600
 ccgagcttca gcttcgctca ggtccagtc ccagccgagc tegtgcgtcg caccggatgc 660
 cgccgeaacg acgacggcaa ctcccctgcc gtccagatcc ccatgcgttc ctccccctc 720
 ctccgctccg ccgttggtgg cgccctgccg gtgttggecc ttgccaagga tgatctcgcg 780
 tactccctc ctttctaccc ttcccctatg gcagatggtc aggggtgaat ggcggaagta 840

[0078]

tacaaacgcg ctgtagacat agtttccag atgacgttga cagagaaagt caacttaacg	900
actggaacag gatggcaact agagaggtgt gttggacaaa ctggcagtgt tcccagactc	960
aacatcccca gcttgtgttt gcaggatagt cctcttggta ttcttttctc ggactacaat	1020
tcagcttttc ctgcgggtgt taatgtcgtt gccacctggg acaagacgct cgcctacctt	1080
cgtggtcagg caatgggtga ggagttcagt gataagggtt ttgacgttca gctgggtcct	1140
gctgtcggcc ctctcgggtc tcattcggat ggcggtagaa actgggaaag ttcttcacca	1200
gatccagccc tcaccggtgt actttttgcg gagaegatta agggatttca agatgctggt	1260
gtcattgoga cagctaagca ttatatcatg aacgaacaag agcatttccg ccaacaaccc	1320
gaggetgcgg gttacggatt caacgtaagc gacagttaga gttccaacgt tgatgacaag	1380
actatgeatg aattgtacct ctggcccttc gcggatgcag tacgcgtcgg agtcggtgct	1440
gttatgtgct cttaacaacca aatcaacaac agctacggtt gcgagaatag cgaactctg	1500
aacaagcttt tgaaggcga gcttggtttc caaggtcttc tcattgagta ttggaccgct	1560
caacacagcg gcgtaggcgc tgccttagca ggtctggata tgcgatgcc cggtagtgtt	1620
accttcgata gtgttacgtc ttcttgggtt gcaacttga cggtcggtgt ccttaacggt	1680
acaatccccc aatggcgtgt tgatgacatg gctgtccgta tcattggccc ttattacaag	1740
gttggccgcg acaacaaata cacccttccc aacttcagct cgtggaccag ggacgaatat	1800
ggtttcgcgc ataaccatgt ttcggaaggt gtttacgaga gggtaacga attcgtggac	1860
gtgcaacgcg atcatgccga cctaatcctt cgcctcggcg cgcagagcac tgttctgctg	1920
aagaacaagg gtgccttgcc cttagaccgc aaggaaaagc tggtcgccct tcitggagag	1980
gatgcgggtt ccaactcgtg gggcgctaac ggctgtgatg accgtggttg cgataacggt	2040
acctttgcca tggcctgggg tagcggtaact gcaattttcc catacctcgt gacaccagag	2100
caggcgattc agaacgaagt tcttcagggc cgttgtaatg tcttcgccgt gaccgacagt	2160
tgggcgctcg acaagatcgc tgcggctgcc cgcaggcca gcgtatctct cgtgttcgtc	2220
aactccgact caggagaagg ctatcttagt gtggatggaa atgaggcgca tcgtaacaac	2280
atcactctgt ggaagaacgg cgacaatgtg gtcaagaccg cagcgaataa ctgtaacaac	2340
accgttgtca tcattcactc cgtcggacca gttttgatcg atgaatgta tgaccacccc	2400
aatgtcaetg gtattctctg ggtgtgtctg ccaggccagg agtctggtaa ctccattgcc	2460
gatgtgtgt acggtcgtgt caaccttggc gccaaagtct ctctcacttg gggaagacc	2520
cgggagtcgt atggttctcc cttggtcaag gatgccaaca atggcaacgg agcgcgccag	2580
tcgtatttca cccagggtgt ttctatcgat taccgccatt tcgataagtt caatgagacc	2640
cctatctacg agtttggcta cggcttgagc tacaccacct tcgagctctc cgacctccat	2700
gttcagcccc tgaacgcgtc ccgatacaact cccaccagtg gcatgactga agctgcaaag	2760
aactttgggt aaattggcga tgcgtcggag tacgtgtatc cggaggggct ggaaaggatc	2820
catgagttaa tctatccctg gatcaactct accgacctga aggcctcgtc tgacgattct	2880
aactacggct gggaagactc caagtatatt cccgaaggcg ccacggatgg gtctgccag	2940
ccccgtttgc ccgtatgtg tggtgccgga ggaaaccccg gtctgtacga ggatcttttc	3000

[0079]

cgcgtctctg tgaaggicaa gaacacgggc aatgtcgccg gtgatgaagt tectcagctg 3060
 tacgtttccc taggcggccc gaatgagccc aaggtagtac tgcgcaagtt tgagcgtatt 3120
 cacttggccc ctctgcagga gcccggtggg acaacgaccc ttaccgctcg tgaccttgca 3180
 aactgggaag ttctggctca ggactggacc gtcactcett accecaagac gatctacgtt 3240
 ggaaactcct caccgaaact gccgctccag gctcgtctgc ctaaggceca gtaa 3294

<210> 70
 <211> 1097
 <212> PRT
 <213> 米曲霉(*Aspergillus oryzae*)

<400> 70

Met Arg Ser Ser Pro Leu Leu Arg Ser Ala Val Val Ala Ala Leu Pro
 1 5 10 15
 Val Leu Ala Leu Ala Ala Asp Gly Arg Ser Thr Arg Tyr Trp Asp Cys
 20 25 30
 Cys Lys Pro Ser Cys Gly Trp Ala Lys Lys Ala Pro Val Asn Gln Pro
 35 40 45
 Val Phe Ser Cys Asn Ala Asn Phe Gln Arg Ile Thr Asp Phe Asp Ala
 50 55 60
 Lys Ser Gly Cys Glu Pro Gly Gly Val Ala Tyr Ser Cys Ala Asp Gln
 65 70 75 80
 Thr Pro Trp Ala Val Asn Asp Asp Phe Ala Leu Gly Phe Ala Ala Thr
 85 90 95
 Ser Ile Ala Gly Ser Asn Glu Ala Gly Trp Cys Cys Ala Cys Tyr Glu
 100 105 110
 Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Lys Met Val Val Gln
 115 120 125
 Ser Thr Ser Thr Gly Gly Asp Leu Gly Ser Asn His Phe Asp Leu Asn
 130 135 140
 Ile Pro Gly Gly Gly Val Gly Ile Phe Asp Gly Cys Thr Pro Gln Phe
 145 150 155 160
 Gly Gly Leu Pro Gly Gln Arg Tyr Gly Gly Ile Ser Ser Arg Asn Glu
 165 170 175
 Cys Asp Arg Phe Pro Asp Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe
 180 185 190
 Asp Trp Phe Lys Asn Ala Asp Asn Pro Ser Phe Ser Phe Arg Gln Val
 195 200 205
 Gln Cys Pro Ala Glu Leu Val Ala Arg Thr Gly Cys Arg Arg Asn Asp
 210 215 220
 Asp Gly Asn Phe Pro Ala Val Gln Ile Pro Met Arg Ser Ser Pro Leu
 225 230 235 240
 Leu Arg Ser Ala Val Val Ala Ala Leu Pro Val Leu Ala Leu Ala Lys
 245 250 255
 Asp Asp Leu Ala Tyr Ser Pro Pro Phe Tyr Pro Ser Pro Trp Ala Asp
 260 265 270
 Gly Gln Gly Glu Trp Ala Glu Val Tyr Lys Arg Ala Val Asp Ile Val
 275 280 285
 Ser Gln Met Thr Leu Thr Glu Lys Val Asn Leu Thr Thr Gly Thr Gly

[0080]

290	295	300
Trp Gln Leu Glu Arg Cys Val Gly Gln Thr Gly Ser Val Pro Arg Leu 305 310 315 320		
Asn Ile Pro Ser Leu Cys Leu Gln Asp Ser Pro Leu Gly Ile Arg Phe 325 330 335		
Ser Asp Tyr Asn Ser Ala Phe Pro Ala Gly Val Asn Val Ala Ala Thr 340 345 350		
Trp Asp Lys Thr Leu Ala Tyr Leu Arg Gly Gln Ala Met Gly Glu Glu 355 360 365		
Phe Ser Asp Lys Gly Ile Asp Val Gln Leu Gly Pro Ala Ala Gly Pro 370 375 380		
Leu Gly Ala His Pro Asp Gly Gly Arg Asn Trp Glu Ser Phe Ser Pro 385 390 395 400		
Asp Pro Ala Leu Thr Gly Val Leu Phe Ala Glu Thr Ile Lys Gly Ile 405 410 415		
Gln Asp Ala Gly Val Ile Ala Thr Ala Lys His Tyr Ile Met Asn Glu 420 425 430		
Gln Glu His Phe Arg Gln Gln Pro Glu Ala Ala Gly Tyr Gly Phe Asn 435 440 445		
Val Ser Asp Ser Leu Ser Ser Asn Val Asp Asp Lys Thr Met His Glu 450 455 460		
Leu Tyr Leu Trp Pro Phe Ala Asp Ala Val Arg Ala Gly Val Gly Ala 465 470 475 480		
Val Met Cys Ser Tyr Asn Gln Ile Asn Asn Ser Tyr Gly Cys Glu Asn 485 490 495		
Ser Glu Thr Leu Asn Lys Leu Leu Lys Ala Glu Leu Gly Phe Gln Gly 500 505 510		
Phe Val Met Ser Asp Trp Thr Ala Gln His Ser Gly Val Gly Ala Ala 515 520 525		
Leu Ala Gly Leu Asp Met Ser Met Pro Gly Asp Val Thr Phe Asp Ser 530 535 540		
Gly Thr Ser Phe Trp Gly Ala Asn Leu Thr Val Gly Val Leu Asn Gly 545 550 555 560		
Thr Ile Pro Gln Trp Arg Val Asp Asp Met Ala Val Arg Ile Met Ala 565 570 575		
Ala Tyr Tyr Lys Val Gly Arg Asp Thr Lys Tyr Thr Pro Pro Asn Phe 580 585 590		
Ser Ser Trp Thr Arg Asp Glu Tyr Gly Phe Ala His Asn His Val Ser 595 600 605		
Glu Gly Ala Tyr Glu Arg Val Asn Glu Phe Val Asp Val Gln Arg Asp 610 615 620		
His Ala Asp Leu Ile Arg Arg Ile Gly Ala Gln Ser Thr Val Leu Leu 625 630 635 640		
Lys Asn Lys Gly Ala Leu Pro Leu Ser Arg Lys Glu Lys Leu Val Ala 645 650 655		
Leu Leu Gly Glu Asp Ala Gly Ser Asn Ser Trp Gly Ala Asn Gly Cys 660 665 670		
Asp Asp Arg Gly Cys Asp Asn Gly Thr Leu Ala Met Ala Trp Gly Ser 675 680 685		

[0081]

Gly Thr Ala Asn Phe Pro Tyr Leu Val Thr Pro Glu Gln Ala Ile Gln
 690 695 700
 Asn Glu Val Leu Gln Gly Arg Gly Asn Val Phe Ala Val Thr Asp Ser
 705 710 715 720
 Trp Ala Leu Asp Lys Ile Ala Ala Ala Ala Arg Gln Ala Ser Val Ser
 725 730 735
 Leu Val Phe Val Asn Ser Asp Ser Gly Glu Gly Tyr Leu Ser Val Asp
 740 745 750
 Gly Asn Glu Gly Asp Arg Asn Asn Ile Thr Leu Trp Lys Asn Gly Asp
 755 760 765
 Asn Val Val Lys Thr Ala Ala Asn Asn Cys Asn Asn Thr Val Val Ile
 770 775 780
 Ile His Ser Val Gly Pro Val Leu Ile Asp Glu Trp Tyr Asp His Pro
 785 790 795 800
 Asn Val Thr Gly Ile Leu Trp Ala Gly Leu Pro Gly Gln Glu Ser Gly
 805 810 815
 Asn Ser Ile Ala Asp Val Leu Tyr Gly Arg Val Asn Pro Gly Ala Lys
 820 825 830
 Ser Pro Phe Thr Trp Gly Lys Thr Arg Glu Ser Tyr Gly Ser Pro Leu
 835 840 845
 Val Lys Asp Ala Asn Asn Gly Asn Gly Ala Pro Gln Ser Asp Phe Thr
 850 855 860
 Gln Gly Val Phe Ile Asp Tyr Arg His Phe Asp Lys Phe Asn Glu Thr
 865 870 875 880
 Pro Ile Tyr Glu Phe Gly Tyr Gly Leu Ser Tyr Thr Thr Phe Glu Leu
 885 890 895
 Ser Asp Leu His Val Gln Pro Leu Asn Ala Ser Arg Tyr Thr Pro Thr
 900 905 910
 Ser Gly Met Thr Glu Ala Ala Lys Asn Phe Gly Glu Ile Gly Asp Ala
 915 920 925
 Ser Glu Tyr Val Tyr Pro Glu Gly Leu Glu Arg Ile His Glu Phe Ile
 930 935 940
 Tyr Pro Trp Ile Asn Ser Thr Asp Leu Lys Ala Ser Ser Asp Asp Ser
 945 950 955 960
 Asn Tyr Gly Trp Glu Asp Ser Lys Tyr Ile Pro Glu Gly Ala Thr Asp
 965 970 975
 Gly Ser Ala Gln Pro Arg Leu Pro Ala Ser Gly Gly Ala Gly Gly Asn
 980 985 990
 Pro Gly Leu Tyr Glu Asp Leu Phe Arg Val Ser Val Lys Val Lys Asn
 995 1000 1005
 Thr Gly Asn Val Ala Gly Asp Glu Val Pro Gln Leu Tyr Val Ser
 1010 1015 1020
 Leu Gly Gly Pro Asn Glu Pro Lys Val Val Leu Arg Lys Phe Glu
 1025 1030 1035
 Arg Ile His Leu Ala Pro Ser Gln Glu Ala Val Trp Thr Thr Thr
 1040 1045 1050
 Leu Thr Arg Arg Asp Leu Ala Asn Trp Asp Val Ser Ala Gln Asp
 1055 1060 1065

[0082]

Trp Thr Val Thr Pro Tyr Pro Lys Thr Ile Tyr Val Gly Asn Ser
 1070 1075 1080
 Ser Arg Lys Leu Pro Leu Gln Ala Ser Leu Pro Lys Ala Gln
 1085 1090 1095
 <210> 71
 <211> 363
 <212> PRT
 <213> 灰盖鬼伞(Coprinus cinereus)
 <220>
 <221> 信号
 <222> (1)..(20)
 <220>
 <221> 成熟肽
 <222> (21)..(363)
 <400> 71
 Met Lys Leu Ser Leu Leu Ser Thr Phe Ala Ala Val Ile Ile Gly Ala
 -20 -15 -10 -5
 Leu Ala Leu Pro Gln Gly Pro Gly Gly Gly Gly Ser Val Thr Cys Pro
 -1 1 5 10
 Gly Gly Gln Ser Thr Ser Asn Ser Gln Cys Cys Val Trp Phe Asp Val
 15 20 25
 Leu Asp Asp Leu Gln Thr Asn Phe Tyr Gln Gly Ser Lys Cys Glu Ser
 30 35 40
 Pro Val Arg Lys Ile Leu Arg Ile Val Phe His Asp Ala Ile Gly Phe
 45 50 55 60
 Ser Pro Ala Leu Thr Ala Ala Gly Gln Phe Gly Gly Gly Gly Ala Asp
 65 70 75
 Gly Ser Ile Ile Ala His Ser Asn Ile Glu Leu Ala Phe Pro Ala Asn
 80 85 90
 Gly Gly Leu Thr Asp Thr Val Glu Ala Leu Arg Ala Val Gly Ile Asn
 95 100 105
 His Gly Val Ser Phe Gly Asp Leu Ile Gln Phe Ala Thr Ala Val Gly
 110 115 120
 Met Ser Asn Cys Pro Gly Ser Pro Arg Leu Glu Phe Leu Thr Gly Arg
 125 130 135 140
 Ser Asn Ser Ser Gln Pro Ser Pro Pro Ser Leu Ile Pro Gly Pro Gly
 145 150 155
 Asn Thr Val Thr Ala Ile Leu Asp Arg Met Gly Asp Ala Gly Phe Ser
 160 165 170
 Pro Asp Glu Val Val Asp Leu Leu Ala Ala His Ser Leu Ala Ser Gln
 175 180 185
 Glu Gly Leu Asn Ser Ala Ile Phe Arg Ser Pro Leu Asp Ser Thr Pro
 190 195 200
 Gln Val Phe Asp Thr Gln Phe Tyr Ile Glu Thr Leu Leu Lys Gly Thr
 205 210 215 220
 Thr Gln Pro Gly Pro Ser Leu Gly Phe Ala Glu Glu Leu Ser Pro Phe
 225 230 235
 Pro Gly Glu Phe Arg Met Arg Ser Asp Ala Leu Leu Ala Arg Asp Ser
 240 245 250

[0083]

Arg Thr Ala Cys Arg Trp Gln Ser Met Thr Ser Ser Asn Glu Val Met
 255 260 265
 Gly Gln Arg Tyr Arg Ala Ala Met Ala Lys Met Ser Val Leu Gly Phe
 270 275 280
 Asp Arg Asn Ala Leu Thr Asp Cys Ser Asp Val Ile Pro Ser Ala Val
 285 290 295 300
 Ser Asn Asn Ala Ala Pro Val Ile Pro Gly Gly Leu Thr Val Asp Asp
 305 310 315
 Ile Glu Val Ser Cys Pro Ser Glu Pro Phe Pro Glu Ile Ala Thr Ala
 320 325 330
 Ser Gly Pro Leu Pro Ser Leu Ala Pro Ala Pro
 335 340
 <210> 72
 <211> 253
 <212> PRT
 <213> 扩展青霉(Penicillium expansum)
 <400> 72
 Met Leu Ser Ser Thr Thr Arg Thr Leu Ala Phe Thr Gly Leu Ala Gly
 1 5 10 15
 Leu Leu Ser Ala Pro Leu Val Lys Ala His Gly Phe Val Gln Gly Ile
 20 25 30
 Val Ile Gly Asp Gln Phe Tyr Ser Gly Tyr Ile Val Asn Ser Phe Pro
 35 40 45
 Tyr Glu Ser Asn Pro Pro Pro Val Ile Gly Trp Ala Thr Thr Ala Thr
 50 55 60
 Asp Leu Gly Phe Val Asp Gly Thr Gly Tyr Gln Gly Pro Asp Ile Ile
 65 70 75 80
 Cys His Arg Asn Ala Thr Pro Ala Pro Leu Thr Ala Pro Val Ala Ala
 85 90 95
 Gly Gly Thr Val Glu Leu Gln Trp Thr Pro Trp Pro Asp Ser His His
 100 105 110
 Gly Pro Val Ile Thr Tyr Leu Ala Pro Cys Asn Gly Asn Cys Ser Thr
 115 120 125
 Val Asp Lys Thr Thr Leu Glu Phe Phe Lys Ile Asp Gln Gln Gly Leu
 130 135 140
 Ile Asp Asp Thr Ser Pro Pro Gly Thr Trp Ala Ser Asp Asn Leu Ile
 145 150 155 160
 Ala Asn Asn Asn Ser Trp Thr Val Thr Ile Pro Asn Ser Val Ala Pro
 165 170 175
 Gly Asn Tyr Val Leu Arg His Glu Ile Ile Ala Leu His Ser Ala Asn
 180 185 190
 Asn Lys Asp Gly Ala Gln Asn Tyr Pro Gln Cys Ile Asn Ile Glu Val
 195 200 205
 Thr Gly Gly Gly Ser Asp Ala Pro Glu Gly Thr Leu Gly Glu Asp Leu
 210 215 220
 Tyr His Asp Thr Asp Pro Gly Ile Leu Val Asp Ile Tyr Glu Pro Ile
 225 230 235 240
 Ala Thr Tyr Thr Ile Pro Gly Pro Pro Glu Pro Thr Phe
 245 250

[0084]

<210> 73
 <211> 792
 <212> PRT
 <213> 烟曲霉(*Aspergillus fumigatus*)
 <400> 73
 Met Ala Val Ala Lys Ser Ile Ala Ala Val Leu Val Ala Leu Leu Pro
 1 5 10 15
 Gly Ala Leu Ala Gln Ala Asn Thr Ser Tyr Val Asp Tyr Asn Val Glu
 20 25 30
 Ala Asn Pro Asp Leu Thr Pro Gln Ser Val Ala Thr Ile Asp Leu Ser
 35 40 45
 Phe Pro Asp Cys Glu Asn Gly Pro Leu Ser Lys Thr Leu Val Cys Asp
 50 55 60
 Thr Ser Ala Arg Pro His Asp Arg Ala Ala Ala Leu Val Ser Met Phe
 65 70 75 80
 Thr Phe Glu Glu Leu Val Asn Asn Thr Gly Asn Thr Ser Pro Gly Val
 85 90 95
 Pro Arg Leu Gly Leu Pro Pro Tyr Gln Val Trp Ser Glu Ala Leu His
 100 105 110
 Gly Leu Asp Arg Ala Asn Phe Thr Asn Glu Gly Glu Tyr Ser Trp Ala
 115 120 125
 Thr Ser Phe Pro Met Pro Ile Leu Thr Met Ser Ala Leu Asn Arg Thr
 130 135 140
 Leu Ile Asn Gln Ile Ala Thr Ile Ile Ala Thr Gln Gly Arg Ala Phe
 145 150 155 160
 Asn Asn Val Gly Arg Tyr Gly Leu Asp Val Tyr Ala Pro Asn Ile Asn
 165 170 175
 Ala Phe Arg Ser Ala Met Trp Gly Arg Gly Gln Glu Thr Pro Gly Glu
 180 185 190
 Asp Ala Tyr Cys Leu Ala Ser Ala Tyr Ala Tyr Glu Tyr Ile Thr Gly
 195 200 205
 Ile Gln Gly Gly Val Asp Pro Gln His Leu Lys Leu Val Ala Thr Ala
 210 215 220
 Lys His Tyr Ala Gly Tyr Asp Leu Glu Asn Trp Asp Gly His Ser Arg
 225 230 235 240
 Leu Gly Asn Asp Met Asn Ile Thr Gln Gln Glu Leu Ser Glu Tyr Tyr
 245 250 255
 Thr Pro Gln Phe Leu Val Ala Ala Arg Asp Ala Lys Val His Ser Val
 260 265 270
 Met Cys Ser Tyr Asn Ala Val Asn Gly Val Pro Ser Cys Ala Asn Ser
 275 280 285
 Phe Phe Leu Gln Thr Leu Leu Arg Asp Thr Phe Gly Phe Val Glu Asp
 290 295 300
 Gly Tyr Val Ser Ser Asp Cys Asp Ser Ala Tyr Asn Val Trp Asn Pro
 305 310 315 320
 His Glu Phe Ala Ala Asn Ile Thr Gly Ala Ala Ala Asp Ser Ile Arg
 325 330 335
 Ala Gly Thr Asp Ile Asp Cys Gly Thr Thr Tyr Gln Tyr Tyr Phe Gly

[0085]

340	345	350
Glu Ala Phe Asp Glu Gln Glu Val Thr Arg Ala Glu Ile Glu Arg Gly 355	360	365
Val Ile Arg Leu Tyr Ser Asn Leu Val Arg Leu Gly Tyr Phe Asp Gly 370	375	380
Asn Gly Ser Val Tyr Arg Asp Leu Thr Trp Asn Asp Val Val Thr Thr 385	390	395
Asp Ala Trp Asn Ile Ser Tyr Glu Ala Ala Val Glu Gly Ile Val Leu 405	410	415
Leu Lys Asn Asp Gly Thr Leu Pro Leu Ala Lys Ser Val Arg Ser Val 420	425	430
Ala Leu Ile Gly Pro Trp Met Asn Val Thr Thr Gln Leu Gln Gly Asn 435	440	445
Tyr Phe Gly Pro Ala Pro Tyr Leu Ile Ser Pro Leu Asn Ala Phe Gln 450	455	460
Asn Ser Asp Phe Asp Val Asn Tyr Ala Phe Gly Thr Asn Ile Ser Ser 465	470	475
His Ser Thr Asp Gly Phe Ser Glu Ala Leu Ser Ala Ala Lys Lys Ser 485	490	495
Asp Val Ile Ile Phe Ala Gly Gly Ile Asp Asn Thr Leu Glu Ala Glu 500	505	510
Ala Met Asp Arg Met Asn Ile Thr Trp Pro Gly Asn Gln Leu Gln Leu 515	520	525
Ile Asp Gln Leu Ser Gln Leu Gly Lys Pro Leu Ile Val Leu Gln Met 530	535	540
Gly Gly Gly Gln Val Asp Ser Ser Ser Leu Lys Ser Asn Lys Asn Val 545	550	555
Asn Ser Leu Ile Trp Gly Gly Tyr Pro Gly Gln Ser Gly Gly Gln Ala 565	570	575
Leu Leu Asp Ile Ile Thr Gly Lys Arg Ala Pro Ala Gly Arg Leu Val 580	585	590
Val Thr Gln Tyr Pro Ala Glu Tyr Ala Thr Gln Phe Pro Ala Thr Asp 595	600	605
Met Ser Leu Arg Pro His Gly Asn Asn Pro Gly Gln Thr Tyr Met Trp 610	615	620
Tyr Thr Gly Thr Pro Val Tyr Glu Phe Gly His Gly Leu Phe Tyr Thr 625	630	635
Thr Phe His Ala Ser Leu Pro Gly Thr Gly Lys Asp Lys Thr Ser Phe 645	650	655
Asn Ile Gln Asp Leu Leu Thr Gln Pro His Pro Gly Phe Ala Asn Val 660	665	670
Glu Gln Met Pro Leu Leu Asn Phe Thr Val Thr Ile Thr Asn Thr Gly 675	680	685
Lys Val Ala Ser Asp Tyr Thr Ala Met Leu Phe Ala Asn Thr Thr Ala 690	695	700
Gly Pro Ala Pro Tyr Pro Asn Lys Trp Leu Val Gly Phe Asp Arg Leu 705	710	715
Ala Ser Leu Glu Pro His Arg Ser Gln Thr Met Thr Ile Pro Val Thr 725	730	735

[0086]

Ile Asp Ser Val Ala Arg Thr Asp Glu Ala Gly Asn Arg Val Leu Tyr
740 745 750

Pro Gly Lys Tyr Glu Leu Ala Leu Asn Asn Glu Arg Ser Val Val Leu
755 760 765

Gln Phe Val Leu Thr Gly Arg Glu Ala Val Ile Phe Lys Trp Pro Val
770 775 780

Glu Gln Gln Gln Ile Ser Ser Ala
785 790

<210> 74
 <211> 406
 <212> PRT
 <213> 棘孢曲霉(*Aspergillus aculeatus*)

<400> 74

Met Val Gly Leu Leu Ser Ile Thr Ala Ala Leu Ala Ala Thr Val Leu
1 5 10 15

Pro Asn Ile Val Ser Ala Val Gly Leu Asp Gln Ala Ala Val Ala Lys
20 25 30

Gly Leu Gln Tyr Phe Gly Thr Ala Thr Asp Asn Pro Glu Leu Thr Asp
35 40 45

Ile Pro Tyr Val Thr Gln Leu Asn Asn Thr Ala Asp Phe Gly Gln Ile
50 55 60

Thr Pro Gly Asn Ser Met Lys Trp Asp Ala Thr Glu Pro Ser Gln Gly
65 70 75 80

Thr Phe Thr Phe Thr Lys Gly Asp Val Ile Ala Asp Leu Ala Glu Gly
85 90 95

Asn Gly Gln Tyr Leu Arg Cys His Thr Leu Val Trp Tyr Asn Gln Leu
100 105 110

Pro Ser Trp Val Thr Ser Gly Thr Trp Thr Asn Ala Thr Leu Thr Ala
115 120 125

Ala Leu Lys Asn His Ile Thr Asn Val Val Ser His Tyr Lys Gly Lys
130 135 140

Cys Leu His Trp Asp Val Val Asn Glu Ala Leu Asn Asp Asp Gly Thr
145 150 155 160

Tyr Arg Thr Asn Ile Phe Tyr Thr Thr Ile Gly Glu Ala Tyr Ile Pro
165 170 175

Ile Ala Phe Ala Ala Ala Ala Ala Ala Asp Pro Asp Ala Lys Leu Phe
180 185 190

Tyr Asn Asp Tyr Asn Leu Glu Tyr Gly Gly Ala Lys Ala Ala Ser Ala
195 200 205

Arg Ala Ile Val Gln Leu Val Lys Asn Ala Gly Ala Lys Ile Asp Gly
210 215 220

Val Gly Leu Gln Ala His Phe Ser Val Gly Thr Val Pro Ser Thr Ser
225 230 235 240

Ser Leu Val Ser Val Leu Gln Ser Phe Thr Ala Leu Gly Val Glu Val
245 250 255

Ala Tyr Thr Glu Ala Asp Val Arg Ile Leu Leu Pro Thr Thr Ala Thr
260 265 270

Thr Leu Ala Gln Gln Ser Ser Asp Phe Gln Ala Leu Val Gln Ser Cys

[0087]

275	280	285
Val Gln Thr Thr Gly Cys	Val Gly Phe Thr Ile Trp Asp Trp Thr Asp	
290	295	300
Lys Tyr Ser Trp Val Pro Ser Thr Phe Ser Gly Tyr Gly Ala Ala Leu		
305	310	315
Pro Trp Asp Glu Asn Leu Val Lys Lys Pro Ala Tyr Asn Gly Leu Leu		
325	330	335
Ala Gly Met Gly Val Thr Val Thr Thr Thr Thr Thr Thr Thr Ala		
340	345	350
Thr Ala Thr Gly Lys Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr		
355	360	365
Gly Thr Thr Ala Ala His Trp Gly Gln Cys Gly Gly Leu Asn Trp Ser		
370	375	380
Gly Pro Thr Ala Cys Ala Thr Gly Tyr Thr Cys Thr Tyr Val Asn Asp		
385	390	395
Tyr Tyr Ser Gln Cys Leu		
405		
<210> 75		
<211> 397		
<212> PRT		
<213> 烟曲霉(Aspergillus funigatus)		
<400> 75		
Met Val His Leu Ser Ser Leu Ala Ala Ala Leu Ala Ala Leu Pro Leu		
1	5	10
Val Tyr Gly Ala Gly Leu Asn Thr Ala Ala Lys Ala Lys Gly Leu Lys		
20	25	30
Tyr Phe Gly Ser Ala Thr Asp Asn Pro Glu Leu Thr Asp Ser Ala Tyr		
35	40	45
Val Ala Gln Leu Ser Asn Thr Asp Asp Phe Gly Gln Ile Thr Pro Gly		
50	55	60
Asn Ser Met Lys Trp Asp Ala Thr Glu Pro Ser Gln Asn Ser Phe Ser		
65	70	75
Phe Ala Asn Gly Asp Ala Val Val Asn Leu Ala Asn Lys Asn Gly Gln		
85	90	95
Leu Met Arg Cys His Thr Leu Val Trp His Ser Gln Leu Pro Asn Trp		
100	105	110
Val Ser Ser Gly Ser Trp Thr Asn Ala Thr Leu Leu Ala Ala Met Lys		
115	120	125
Asn His Ile Thr Asn Val Val Thr His Tyr Lys Gly Lys Cys Tyr Ala		
130	135	140
Trp Asp Val Val Asn Glu Ala Leu Asn Glu Asp Gly Thr Phe Arg Asn		
145	150	155
Ser Val Phe Tyr Gln Ile Ile Gly Pro Ala Tyr Ile Pro Ile Ala Phe		
165	170	175
Ala Thr Ala Ala Ala Ala Asp Pro Asp Val Lys Leu Tyr Tyr Asn Asp		
180	185	190
Tyr Asn Ile Glu Tyr Ser Gly Ala Lys Ala Thr Ala Ala Gln Asn Ile		
195	200	205

[0088]

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

<210> 76
 <211> 532
 <212> PRT
 <213> 烟曲霉(*Aspergillus fumigatus*)

<220>
 <221> 信号
 <222> (1)..(26)

<220>
 <221> 成熟肽
 <222> (27)..(532)

<400> 76

Met Leu Ala Ser Thr Phe Ser Tyr Arg Met Tyr Lys Thr Ala Leu Ile
 -25 -20 -15
 Leu Ala Ala Leu Leu Gly Ser Gly Gln Ala Gln Gln Val Gly Thr Ser
 -10 -5 -1 1 5
 Gln Ala Glu Val His Pro Ser Met Thr Trp Gln Ser Cys Thr Ala Gly
 10 15 20
 Gly Ser Cys Thr Thr Asn Asn Gly Lys Val Val Ile Asp Ala Asn Trp
 25 30 35
 Arg Trp Val His Lys Val Gly Asp Tyr Thr Asn Cys Tyr Thr Gly Asn
 40 45 50
 Thr Trp Asp Thr Thr Ile Cys Pro Asp Asp Ala Thr Cys Ala Ser Asn
 55 60 65 70
 Cys Ala Leu Glu Gly Ala Asn Tyr Glu Ser Thr Tyr Gly Val Thr Ala
 75 80 85

[0089]

```

Ser Gly Asn Ser Leu Arg Leu Asn Phe Val Thr Thr Ser Gln Gln Lys
    90                      95                      100

Asn Ile Gly Ser Arg Leu Tyr Met Met Lys Asp Asp Ser Thr Tyr Glu
    105                      110                      115

Met Phe Lys Leu Leu Asn Gln Glu Phe Thr Phe Asp Val Asp Val Ser
    120                      125                      130

Asn Leu Pro Cys Gly Leu Asn Gly Ala Leu Tyr Phe Val Ala Met Asp
    135                      140                      145                      150

Ala Asp Gly Gly Met Ser Lys Tyr Pro Thr Asn Lys Ala Gly Ala Lys
    155                      160                      165

Tyr Gly Thr Gly Tyr Cys Asp Ser Gln Cys Pro Arg Asp Leu Lys Phe
    170                      175                      180

Ile Asn Gly Gln Ala Asn Val Glu Gly Trp Gln Pro Ser Ser Asn Asp
    185                      190                      195

Ala Asn Ala Gly Thr Gly Asn His Gly Ser Cys Cys Ala Glu Met Asp
    200                      205                      210

Ile Trp Glu Ala Asn Ser Ile Ser Thr Ala Phe Thr Pro His Pro Cys
    215                      220                      225                      230

Asp Thr Pro Gly Gln Val Met Cys Thr Gly Asp Ala Cys Gly Gly Thr
    235                      240                      245

Tyr Ser Ser Asp Arg Tyr Gly Gly Thr Cys Asp Pro Asp Gly Cys Asp
    250                      255                      260

Phe Asn Ser Phe Arg Gln Gly Asn Lys Thr Phe Tyr Gly Pro Gly Met
    265                      270                      275

Thr Val Asp Thr Lys Ser Lys Phe Thr Val Val Thr Gln Phe Ile Thr
    280                      285                      290

Asp Asp Gly Thr Ser Ser Gly Thr Leu Lys Glu Ile Lys Arg Phe Tyr
    295                      300                      305                      310

Val Gln Asn Gly Lys Val Ile Pro Asn Ser Glu Ser Thr Trp Thr Gly
    315                      320                      325

Val Ser Gly Asn Ser Ile Thr Thr Glu Tyr Cys Thr Ala Gln Lys Ser
    330                      335                      340

Leu Phe Gln Asp Gln Asn Val Phe Glu Lys His Gly Gly Leu Glu Gly
    345                      350                      355

Met Gly Ala Ala Leu Ala Gln Gly Met Val Leu Val Met Ser Leu Trp
    360                      365                      370

Asp Asp His Ser Ala Asn Met Leu Trp Leu Asp Ser Asn Tyr Pro Thr
    375                      380                      385                      390

Thr Ala Ser Ser Thr Thr Pro Gly Val Ala Arg Gly Thr Cys Asp Ile
    395                      400                      405

Ser Ser Gly Val Pro Ala Asp Val Glu Ala Asn His Pro Asp Ala Tyr
    410                      415                      420

Val Val Tyr Ser Asn Ile Lys Val Gly Pro Ile Gly Ser Thr Phe Asn
    425                      430                      435

Ser Gly Gly Ser Asn Pro Gly Gly Gly Thr Thr Thr Thr Thr Thr
    440                      445                      450

Gln Pro Thr Thr Thr Thr Thr Thr Ala Gly Asn Pro Gly Gly Thr Gly
    455                      460                      465                      470

Val Ala Gln His Tyr Gly Gln Cys Gly Gly Ile Gly Trp Thr Gly Pro

```

[0090]

	475	480	485
Thr Thr Cys Ala Ser Pro Tyr Thr Cys Gln Lys Leu Asn Asp Tyr Tyr			
	490	495	500
Ser Gln Cys Leu			
	505		
<210>	77		
<211>	454		
<212>	PRT		
<213>	烟曲霉(<i>Aspergillus fumigatus</i>)		
<220>			
<221>	信号		
<222>	(1)..(19)		
<220>			
<221>	成熟肽		
<222>	(20)..(454)		
<400>	77		
Met Lys His Leu Ala Ser Ser Ile Ala Leu Thr Leu Leu Leu Pro Ala			
	-15	-10	-5
Val Gln Ala Gln Gln Thr Val Trp Gly Gln Cys Gly Gly Gln Gly Trp			
	-1 1	5	10
Ser Gly Pro Thr Ser Cys Val Ala Gly Ala Ala Cys Ser Thr Leu Asn			
	15	20	25
Pro Tyr Tyr Ala Gln Cys Ile Pro Gly Ala Thr Ala Thr Ser Thr Thr			
	30	35	40
Leu Thr Thr Thr Thr Ala Ala Thr Thr Thr Ser Gln Thr Thr Thr Lys			
	50	55	60
Pro Thr Thr Thr Gly Pro Thr Thr Ser Ala Pro Thr Val Thr Ala Ser			
	65	70	75
Gly Asn Pro Phe Ser Gly Tyr Gln Leu Tyr Ala Asn Pro Tyr Tyr Ser			
	80	85	90
Ser Glu Val His Thr Leu Ala Met Pro Ser Leu Pro Ser Ser Leu Gln			
	95	100	105
Pro Lys Ala Ser Ala Val Ala Glu Val Pro Ser Phe Val Trp Leu Asp			
	110	115	120
Val Ala Ala Lys Val Pro Thr Met Gly Thr Tyr Leu Ala Asp Ile Gln			
	130	135	140
Ala Lys Asn Lys Ala Gly Ala Asn Pro Pro Ile Ala Gly Ile Phe Val			
	145	150	155
Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly			
	160	165	170
Glu Tyr Ser Ile Ala Asn Asn Gly Val Ala Asn Tyr Lys Ala Tyr Ile			
	175	180	185
Asp Ala Ile Arg Ala Gln Leu Val Lys Tyr Ser Asp Val His Thr Ile			
	190	195	200
Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Leu Val Thr Asn Leu Asn			
	210	215	220
Val Ala Lys Cys Ala Asn Ala Gln Ser Ala Tyr Leu Glu Cys Val Asp			
	225	230	235
Tyr Ala Leu Lys Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu Asp			

[0091]

240	245	250
Ala Gly His Ala Gly Trp 255	Leu Gly Trp Pro Ala Asn 260	Leu Gly Pro Ala 265
Ala Thr Leu Phe Ala Lys Val Tyr Thr Asp Ala Gly Ser Pro Ala Ala 270	275	280 285
Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser Leu 290	295	300
Ser Thr Cys Pro Ser Tyr Thr Gln Gly Asp Pro Asn Cys Asp Glu Lys 305	310	315
Lys Tyr Ile Asn Ala Met Ala Pro Leu Leu Lys Glu Ala Gly Phe Asp 320	325	330
Ala His Phe Ile Met Asp Thr Ser Arg Asn Gly Val Gln Pro Thr Lys 335	340	345
Gln Asn Ala Trp Gly Asp Trp Cys Asn Val Ile Gly Thr Gly Phe Gly 350	355	360 365
Val Arg Pro Ser Thr Asn Thr Gly Asp Pro Leu Gln Asp Ala Phe Val 370	375	380
Trp Ile Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asn Ser Thr Ser 385	390	395
Pro Arg Tyr Asp Ala His Cys Gly Tyr Ser Asp Ala Leu Gln Pro Ala 400	405	410
Pro Glu Ala Gly Thr Trp Phe Gln Ala Tyr Phe Glu Gln Leu Leu Thr 415	420	425
Asn Ala Asn Pro Ser Phe 430	435	
<210> 78		
<211> 863		
<212> PRT		
<213> 烟曲霉(Aspergillus fumigatus)		
<400> 78		
Met Arg Phe Gly Trp Leu Glu Val Ala Ala Leu Thr Ala Ala Ser Val 1	5	10 15
Ala Asn Ala Gln Glu Leu Ala Phe Ser Pro Pro Phe Tyr Pro Ser Pro 20	25	30
Trp Ala Asp Gly Gln Gly Glu Trp Ala Asp Ala His Arg Arg Ala Val 35	40	45
Glu Ile Val Ser Gln Met Thr Leu Ala Glu Lys Val Asn Leu Thr Thr 50	55	60
Gly Thr Gly Trp Glu Met Asp Arg Cys Val Gly Gln Thr Gly Ser Val 65	70	75 80
Pro Arg Leu Gly Ile Asn Trp Gly Leu Cys Gly Gln Asp Ser Pro Leu 85	90	95
Gly Ile Arg Phe Ser Asp Leu Asn Ser Ala Phe Pro Ala Gly Thr Asn 100	105	110
Val Ala Ala Thr Trp Asp Lys Thr Leu Ala Tyr Leu Arg Gly Lys Ala 115	120	125
Met Gly Glu Glu Phe Asn Asp Lys Gly Val Asp Ile Leu Leu Gly Pro 130	135	140

[0092]

Ala Ala Gly Pro Leu Gly Lys Tyr Pro Asp Gly Gly Arg Ile Trp Glu
 145 150 155 160
 Gly Phe Ser Pro Asp Pro Val Leu Thr Gly Val Leu Phe Ala Glu Thr
 165 170 175
 Ile Lys Gly Ile Gln Asp Ala Gly Val Ile Ala Thr Ala Lys His Tyr
 180 185 190
 Ile Leu Asn Glu Gln Glu His Phe Arg Gln Val Gly Glu Ala Gln Gly
 195 200 205
 Tyr Gly Tyr Asn Ile Thr Glu Thr Ile Ser Ser Asn Val Asp Asp Lys
 210 215 220
 Thr Met His Glu Leu Tyr Leu Trp Pro Phe Ala Asp Ala Val Arg Ala
 225 230 235 240
 Gly Val Gly Ala Val Met Cys Ser Tyr Asn Gln Ile Asn Asn Ser Tyr
 245 250 255
 Gly Cys Gln Asn Ser Gln Thr Leu Asn Lys Leu Leu Lys Ala Glu Leu
 260 265 270
 Gly Phe Gln Gly Phe Val Met Ser Asp Trp Ser Ala His His Ser Gly
 275 280 285
 Val Gly Ala Ala Leu Ala Gly Leu Asp Met Ser Met Pro Gly Asp Ile
 290 295 300
 Ser Phe Asp Asp Gly Leu Ser Phe Trp Gly Thr Asn Leu Thr Val Ser
 305 310 315 320
 Val Leu Asn Gly Thr Val Pro Ala Trp Arg Val Asp Asp Met Ala Val
 325 330 335
 Arg Ile Met Thr Ala Tyr Tyr Lys Val Gly Arg Asp Arg Leu Arg Ile
 340 345 350
 Pro Pro Asn Phe Ser Ser Trp Thr Arg Asp Glu Tyr Gly Trp Glu His
 355 360 365
 Ser Ala Val Ser Glu Gly Ala Trp Thr Lys Val Asn Asp Phe Val Asn
 370 375 380
 Val Gln Arg Ser His Ser Gln Ile Ile Arg Glu Ile Gly Ala Ala Ser
 385 390 395 400
 Thr Val Leu Leu Lys Asn Thr Gly Ala Leu Pro Leu Thr Gly Lys Gln
 405 410 415
 Val Lys Val Gly Val Leu Gly Glu Asp Ala Gly Ser Asn Pro Trp Gly
 420 425 430
 Ala Asn Gly Cys Pro Asp Arg Gly Cys Asp Asn Gly Thr Leu Ala Met
 435 440 445
 Ala Trp Gly Ser Gly Thr Ala Asn Phe Pro Tyr Leu Val Thr Pro Glu
 450 455 460
 Gln Ala Ile Gln Arg Glu Val Ile Ser Asn Gly Gly Asn Val Phe Ala
 465 470 475 480
 Val Thr Asp Asn Gly Ala Leu Ser Gln Met Ala Asp Val Ala Ser Gln
 485 490 495
 Ser Ser Val Ser Leu Val Phe Val Asn Ala Asp Ser Gly Glu Gly Phe
 500 505 510
 Ile Ser Val Asp Gly Asn Glu Gly Asp Arg Lys Asn Leu Thr Leu Trp
 515 520 525
 Lys Asn Gly Glu Ala Val Ile Asp Thr Val Val Ser His Cys Asn Asn

[0093]

530	535	540
Thr Ile Val Val Ile His Ser Val Gly Pro Val Leu Ile Asp Arg Trp 545	550	555 560
Tyr Asp Asn Pro Asn Val Thr Ala Ile Ile Trp Ala Gly Leu Pro Gly 565	570	575
Gln Glu Ser Gly Asn Ser Leu Val Asp Val Leu Tyr Gly Arg Val Asn 580	585	590
Pro Ser Ala Lys Thr Pro Phe Thr Trp Gly Lys Thr Arg Glu Ser Tyr 595	600	605
Gly Ala Pro Leu Leu Thr Glu Pro Asn Asn Gly Asn Gly Ala Pro Gln 610	615	620
Asp Asp Phe Asn Glu Gly Val Phe Ile Asp Tyr Arg His Phe Asp Lys 625	630	635 640
Arg Asn Glu Thr Pro Ile Tyr Glu Phe Gly His Gly Leu Ser Tyr Thr 645	650	655
Thr Phe Gly Tyr Ser His Leu Arg Val Gln Ala Leu Asn Ser Ser Ser 660	665	670
Ser Ala Tyr Val Pro Thr Ser Gly Glu Thr Lys Pro Ala Pro Thr Tyr 675	680	685
Gly Glu Ile Gly Ser Ala Ala Asp Tyr Leu Tyr Pro Glu Gly Leu Lys 690	695	700
Arg Ile Thr Lys Phe Ile Tyr Pro Trp Leu Asn Ser Thr Asp Leu Glu 705	710	715 720
Asp Ser Ser Asp Asp Pro Asn Tyr Gly Trp Glu Asp Ser Glu Tyr Ile 725	730	735
Pro Glu Gly Ala Arg Asp Gly Ser Pro Gln Pro Leu Leu Lys Ala Gly 740	745	750
Gly Ala Pro Gly Gly Asn Pro Thr Leu Tyr Gln Asp Leu Val Arg Val 755	760	765
Ser Ala Thr Ile Thr Asn Thr Gly Asn Val Ala Gly Tyr Glu Val Pro 770	775	780
Gln Leu Tyr Val Ser Leu Gly Gly Pro Asn Glu Pro Arg Val Val Leu 785	790	795 800
Arg Lys Phe Asp Arg Ile Phe Leu Ala Pro Gly Glu Gln Lys Val Trp 805	810	815
Thr Thr Thr Leu Asn Arg Arg Asp Leu Ala Asn Trp Asp Val Glu Ala 820	825	830
Gln Asp Trp Val Ile Thr Lys Tyr Pro Lys Lys Val His Val Gly Ser 835	840	845
Ser Ser Arg Lys Leu Pro Leu Arg Ala Pro Leu Pro Arg Val Tyr 850	855	860
<210> 79		
<211> 304		
<212> PRT		
<213> 王棕(Royal Palm)		
<400> 79		
Asp Leu Gln Ile Gly Phe Tyr Asn Thr Ser Cys Pro Thr Ala Glu Ser 1	5	10 15

[0094]

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

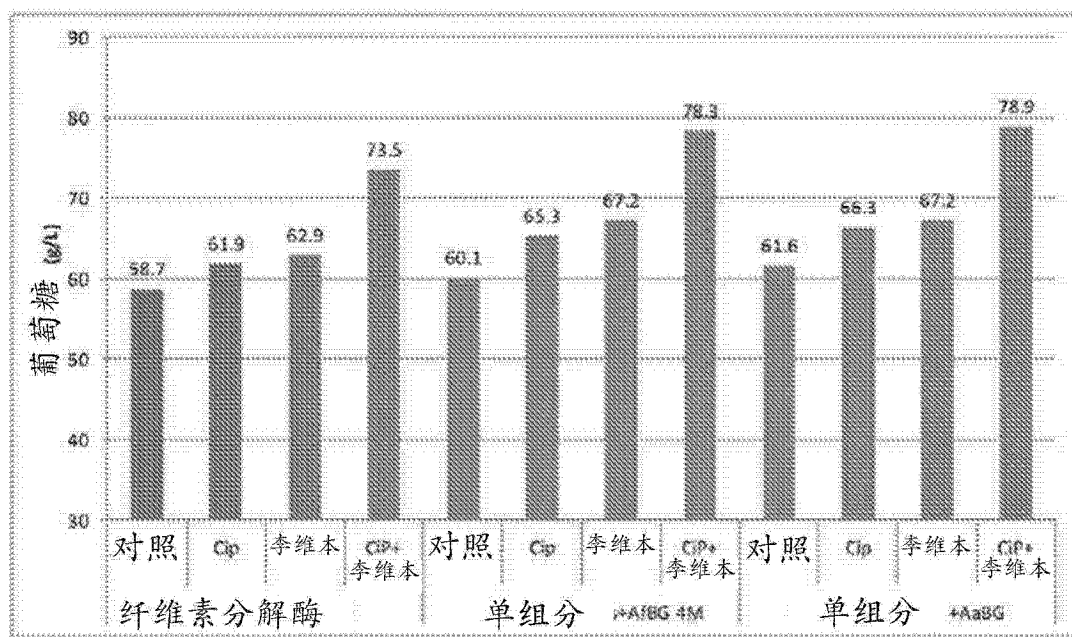


图 1

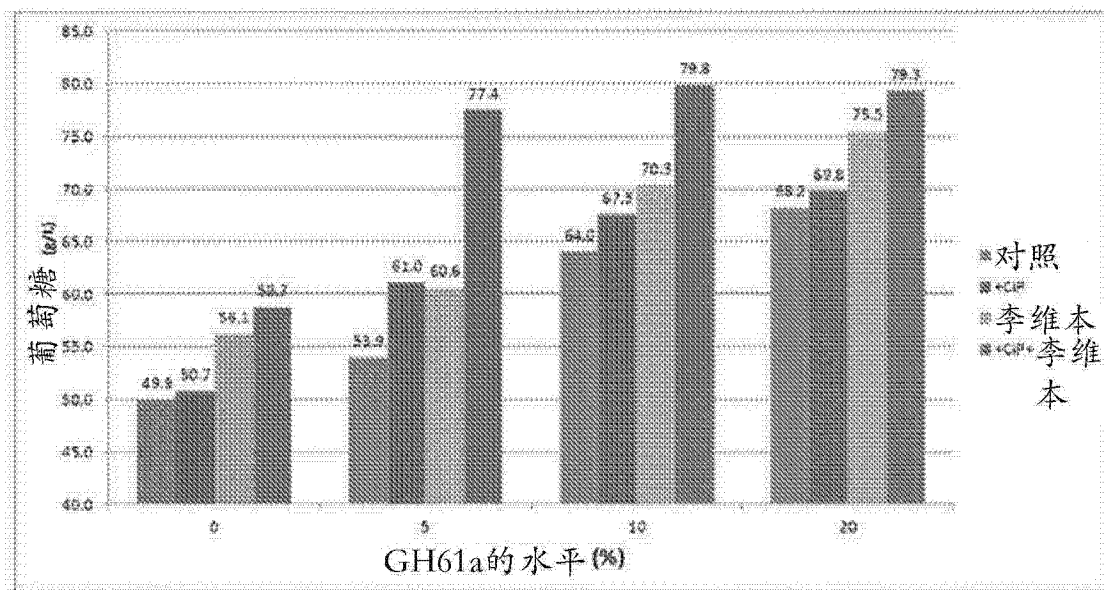


图 2

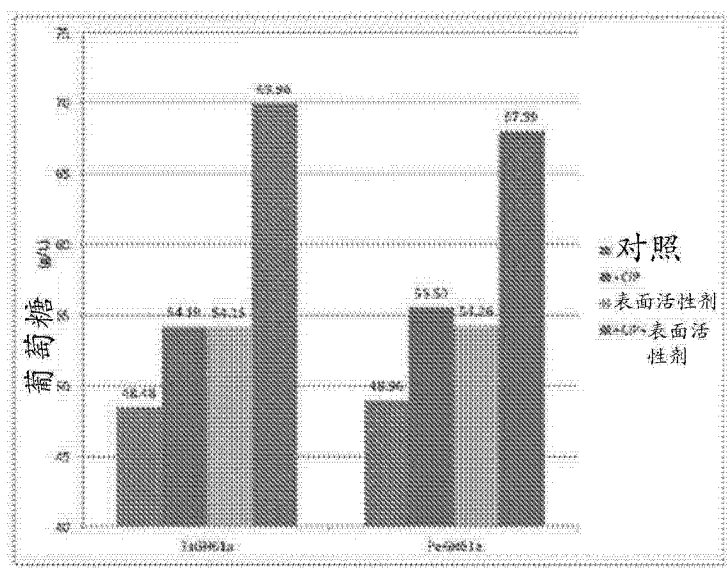


图 3

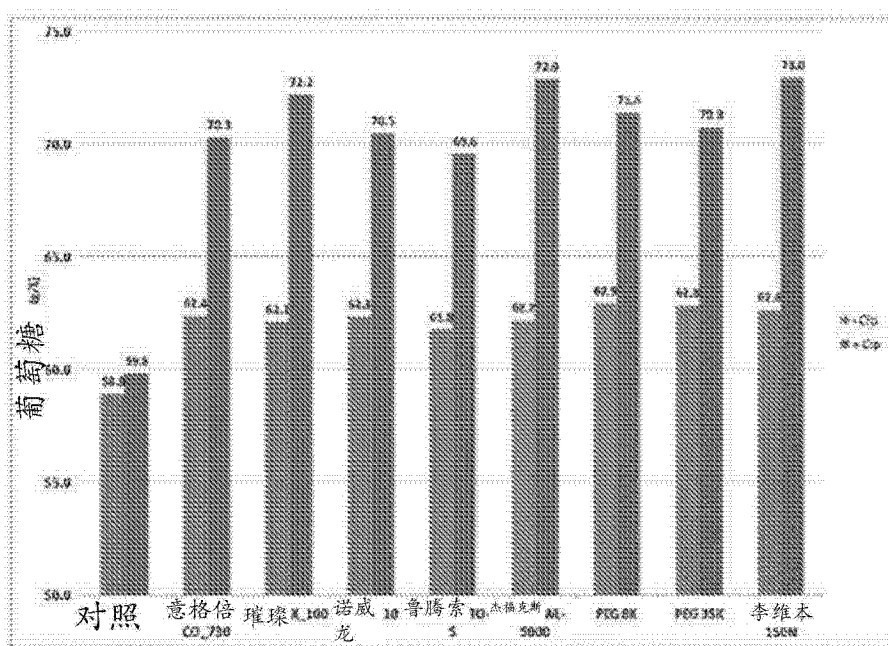


图 4

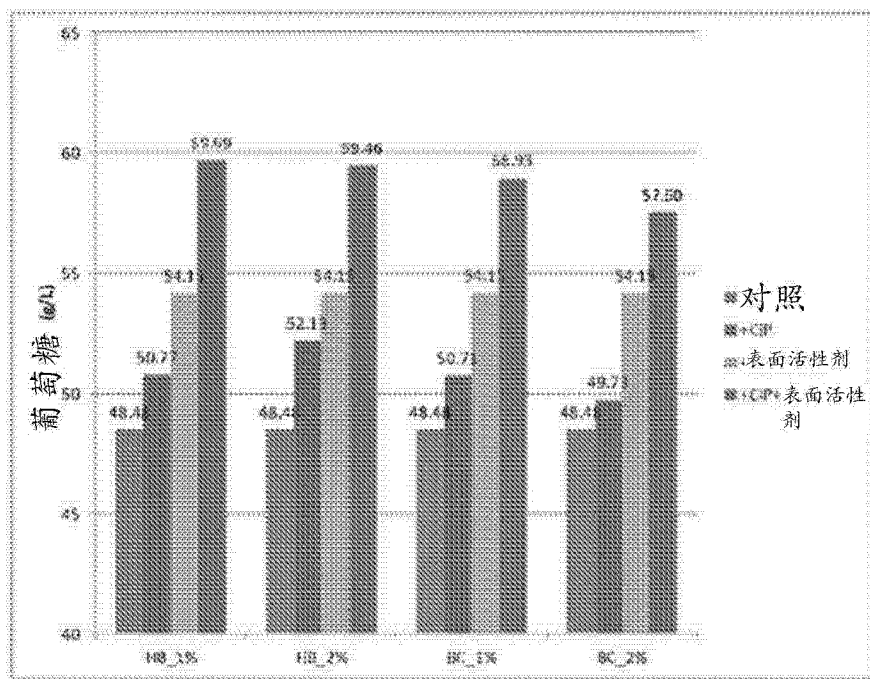


图 5

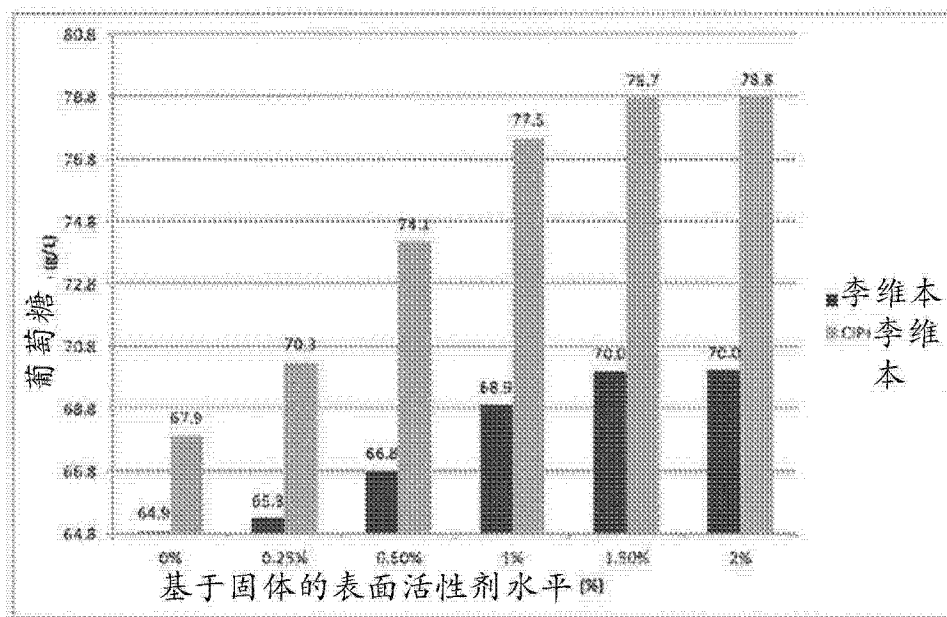


图 6

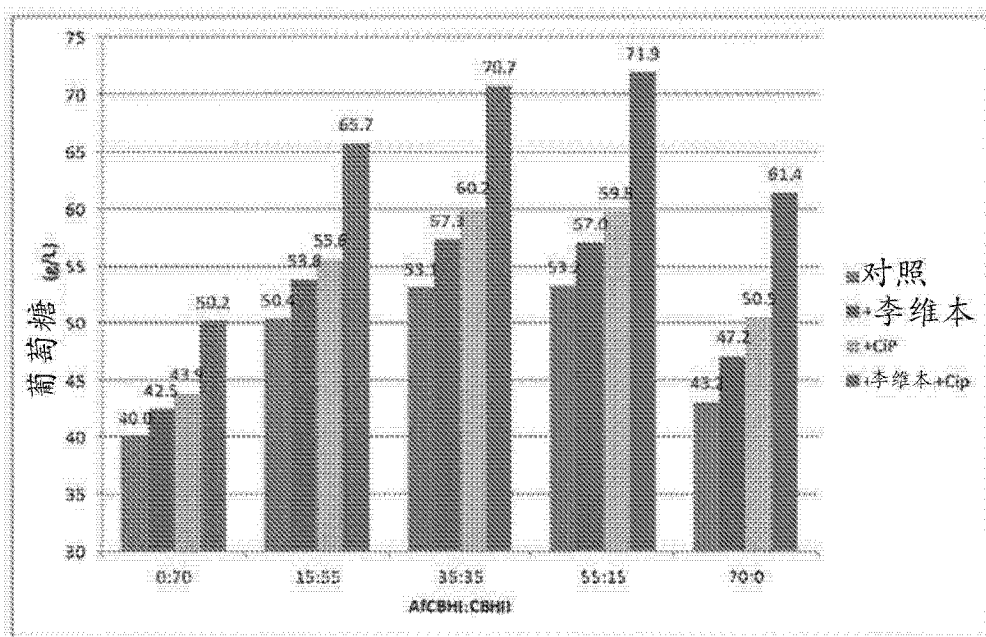


图 7

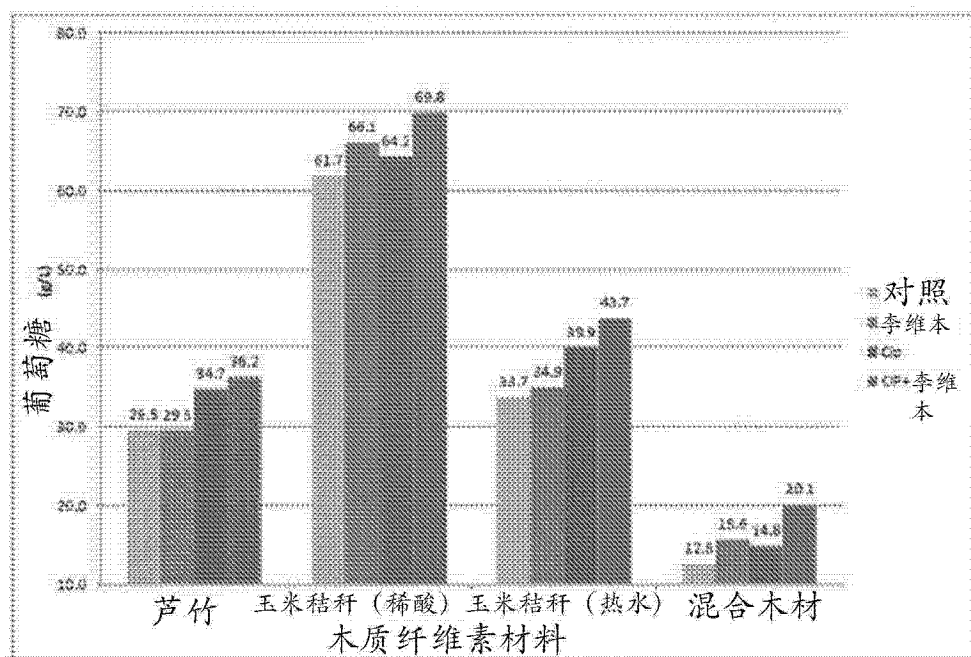


图 8

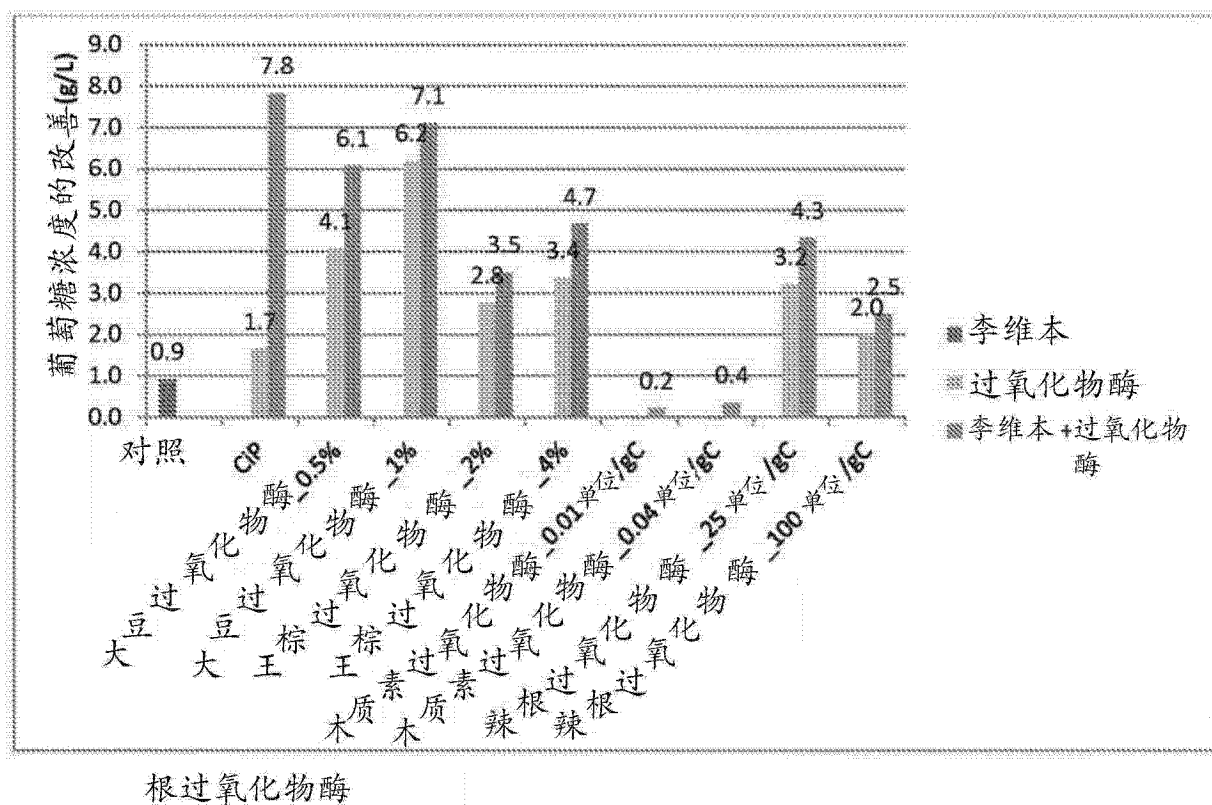


图 9

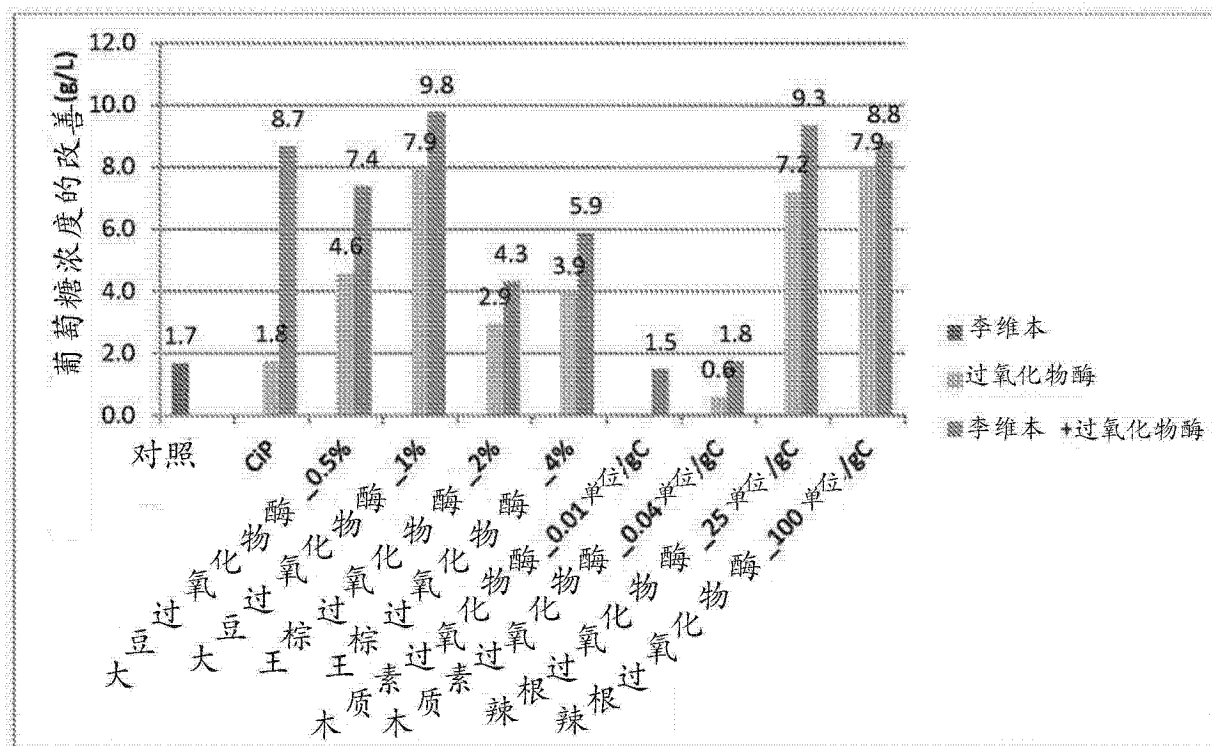


图 10