



(12) 发明专利申请

(10) 申请公布号 CN 103097529 A

(43) 申请公布日 2013. 05. 08

(21) 申请号 201180043537. 0

(51) Int. Cl.

(22) 申请日 2011. 09. 08

C12N 9/42 (2006. 01)

(30) 优先权数据

C12P 19/14 (2006. 01)

61/381, 455 2010. 09. 10 US

(85) PCT申请进入国家阶段日

2013. 03. 11

(86) PCT申请的申请数据

PCT/US2011/050846 2011. 09. 08

(87) PCT申请的公布数据

W02012/033926 EN 2012. 03. 15

(71) 申请人 先正达参股股份有限公司

地址 瑞士巴塞尔

(72) 发明人 J·史蒂芬斯 P·奥勒 Y·巴隆

(74) 专利代理机构 中国国际贸易促进委员会专

利商标事务所 11038

代理人 傅宇昌

权利要求书3页 说明书31页

序列表24页

(54) 发明名称

在纤维素生物质的预处理期间有活性的木聚糖酶

(57) 摘要

提供了多种用一种具有木聚糖酶活性的木聚糖酶来处理木质纤维素材料的组合物以及方法。该酶在增加的 pH 与温度下是稳定的并且有活性的。本发明因此提供用于水解木质纤维素材料的方法尤其是水解纤维素与半纤维素的方法,它们是非木本植物与木本植物的细胞壁的主要组分。这些用于水解纤维素与半纤维素的方法可以用于任何植物、木材或木制品、木材废料、纸浆、纸产品或纸废料或副产品。

1. 一种加工木质纤维素材料的方法,该方法包括使木质纤维素材料与具有木聚糖酶活性的至少一种酶接触的步骤,用来水解该木质纤维素材料,其中该酶具有选自下组一个氨基酸序列,该组由 SEQ ID 号:6、8 或 10 或其一个活性变体或片段组成,并且其中该接触是在约 pH6 至约 pH12 的 pH 条件下。

2. 如权利要求 1 所述的方法,其中这些条件包括约 95° C 的一个反应温度。

3. 如权利要求 1 或 2 所述的方法,其中该木聚糖酶以从约 0.01 单位至约 1000 单位存在。

4. 如权利要求 1 或 2 所述的方法,其中该木聚糖酶以约 500 单位存在。

5. 如权利要求 1 或 2 所述的方法,其中该 pH 随时间降低至约 pH6 的 pH。

6. 如权利要求 1 或 2 所述的方法,其中该接触持续约 12 小时至约 24 小时。

7. 如权利要求 1 或 2 所述的方法,其中该接触持续约 6 小时。

8. 如权利要求 1 或 2 所述的方法,其中该接触是在选自下组的至少一种额外的酶存在下进行的,该组由以下各项组成:一种纤维素酶、半纤维素酶、木质酶、果胶酶和蛋白酶。

9. 如权利要求 8 所述的方法,其中该纤维素酶选自下组,该组由以下各项组成:一种甘露聚糖内切-1,4- β -甘露糖苷酶、1,3- β -D-葡聚糖葡聚糖水解酶、1,3- β -葡聚糖葡聚糖水解酶、1,3-1,4- β -D-葡聚糖葡聚糖水解酶和 1,6- β -D-葡聚糖葡聚糖水解酶。

10. 如权利要求 8 所述的方法,其中该半纤维素酶选自下组,该组由以下各项组成:一种 α -L-阿拉伯呋喃糖苷酶、 α -葡萄糖醛酸酶、乙酰基甘露聚糖酯酶、乙酰基木聚糖酯酶、 α -半乳糖苷酶、 β -葡萄糖苷酶、 β -1,4-木糖苷酶、内切-半乳聚糖酶、内切- β -1,4-甘露聚糖酶、内切- α -1,5-阿拉伯聚糖酶、外切- β -1,4-甘露糖苷酶、外切- β -1,4-木糖苷酶、阿魏酸酯酶(feruloyl esterase)、阿魏酸酯酶(ferulic acid esterase)、对位-熏草酸酯酶、葡萄糖醛酸木聚糖木聚糖水解酶和木葡聚糖内糖基转移酶。

11. 如权利要求 8 所述的方法,其中该木质酶选自下组,该组由以下各项组成:一种二芳基丙烷过氧化酶、葡萄糖氧化酶、乙二醛氧化酶、木素过氧化酶、锰过氧化酶、甲醇氧化酶、甲醇氧化还原酶、苯酚氧化酶、苯酚过氧化酶和藜芦基醇氧化酶。

12. 如权利要求 8 所述的方法,其中该果胶酶是选自下组,该组由以下各项组成:一种果胶降解酶、pectozyme 酶和多聚半乳糖醛酸酶。

13. 如权利要求 8 所述的方法,其中该蛋白酶选自下组,该组由以下各项组成:一种萝藦蛋白酶、菠萝酶、caricain 酶、木瓜凝乳蛋白酶、胶原酶、甘氨酸肽链内切酶、胃蛋白酶、链霉菌蛋白酶、枯草杆菌蛋白酶和嗜热菌蛋白酶。

14. 如权利要求 1、2、或 8 所述的方法,其中该方法包括用至少一种选自下组的酶进行加工,该组由以下各项组成:一种淀粉酶、过氧化氢酶、角质酶、葡聚糖酶、葡糖淀粉酶、葡萄糖异构酶、脂肪酶、植酸酶、支链淀粉酶和木糖异构酶。

15. 如权利要求 1、2、8、或 14 所述的方法,其中该方法进一步包括用至少一种选自下组的试剂进行加工,该组由以下各项组成:一种氯、洗涤剂、次氯酸盐、过氧化氢、草酸、过酸、pH-调节试剂、磷酸三钠、亚氯酸钠、硝酸钠、表面活性剂、脲和水。

16. 如权利要求 1、2、8、14、或 15 所述的方法,其中该方法进一步包括用一种产乙醇微生物进行加工,该产乙醇微生物例如是一种细菌或酵母。

17. 如权利要求 1 直至 16 中任一项所述的方法,其中在步骤 (a) 之前,通过生物预处理

理、化学预处理或物理预处理对该木质纤维素材料进行预处理。

18. 如权利要求 17 所述的方法,其中该生物预处理包括使用木素增溶微生物。

19. 如权利要求 18 所述的方法,其中该化学预处理选自下组,该组由使用以下处理组成:碱处理、氨处理、稀酸处理、有机溶剂处理、臭氧处理、二氧化硫处理、二氧化碳处理和 pH 控制的水热处理。

20. 如权利要求 18 所述的方法,其中该物理预处理选自下组,该组由使用以下项组成:水热处理、辐射、研磨、气烘/蒸汽喷发和声处理(超声处理)。

21. 一种加工木质纤维素材料的方法,该方法包括使木质纤维素材料与具有木聚糖酶活性的至少一种酶接触的步骤,用来水解该木质纤维素材料,其中在 10 或更大的 pH 下,该酶保持至少 60% 的最佳木聚糖酶活性。

22. 一种包括具有木聚糖酶活性的酶的组合物,其中该酶具有选自下组的氨基酸序列,该组由 SEQ ID 号:6、8 或 10 或其一个活性变体或片段组成,并且其中该组合物具有至少约 pH10 的一个 pH。

23. 如权利要求 22 所述的组合物,其中所述木聚糖酶具有选自下组的氨基酸序列,该组由以下各项组成:

a) 与 SEQ ID 号:6、8、或 10 中列出的氨基酸序列具有至少约 90% 序列一致性的一个氨基酸序列;

b) 与 SEQ ID 号:6、8、或 10 中列出的氨基酸序列具有至少约 95% 序列一致性的一个氨基酸序列;以及

c) 与 SEQ ID 号:6、8、或 10 中列出的氨基酸序列具有至少约 99% 序列一致性的一个氨基酸序列。

24. 如权利要求 22 所述的组合物,其中该组合物具有约 pH12 的 pH。

25. 如权利要求 24 所述的组合物,其中该酶在高达约 95° C 的温度下是稳定的。

26. 如权利要求 22 或 23 所述的组合物,进一步包括选自下组的至少一种酶,该组由以下各项组成:一种纤维素酶、半纤维素酶、木质酶、果胶酶和蛋白酶。

27. 如权利要求 26 所述的组合物,其中该纤维素酶选自下组,该组由以下各项组成:一种甘露聚糖内切-1,4- β -甘露糖苷酶、1,3- β -D-葡聚糖葡聚糖水解酶、1,3- β -葡聚糖葡聚糖水解酶、1,3-1,4- β -D-葡聚糖葡聚糖水解酶和 1,6- β -D-葡聚糖葡聚糖水解酶。

28. 如权利要求 26 所述的组合物,其中该半纤维素酶选自下组,该组由以下各项组成:一种 α -L-阿拉伯呋喃糖苷酶、 α -葡萄糖醛酸酶、乙酰基甘露聚糖酯酶、乙酰基木聚糖酯酶、 α -半乳糖苷酶、 β -葡糖苷酶、 β -1,4-木糖苷酶、内切-半乳聚糖酶、内切- β -1,4-甘露聚糖酶、内切- α -1,5-阿拉伯聚糖酶、外切- β -1,4-甘露糖苷酶、外切- β -1,4-木糖苷酶、阿魏酸酯酶(feruloyl esterase)、阿魏酸酯酶(ferulic acid esterase)、对位-熏草酸酯酶、葡糖醛酸木聚糖木聚糖水解酶和木葡聚糖内糖基转移酶。

29. 如权利要求 26 所述的组合物,其中该木质酶选自下组,该组由以下各项组成:一种二芳基丙烷过氧化酶、葡萄糖氧化酶、乙二醛氧化酶、木素过氧化酶、锰过氧化酶、甲醇氧化酶、甲醇氧化还原酶、苯酚氧化酶、苯酚过氧化酶和藜芦基醇氧化酶。

30. 如权利要求 26 所述的组合物,其中该果胶酶是选自下组,该组由以下各项组成:一种果胶降解酶、pectozyme 酶和多聚半乳糖醛酸酶。

31. 如权利要求 26 所述的组合物,其中该蛋白酶选自下组,该组由以下各项组成:一种萝藦蛋白酶、菠萝酶、caricain 酶、木瓜凝乳蛋白酶、胶原酶、甘氨酸肽链内切酶、胃蛋白酶、链霉蛋白酶、枯草杆菌蛋白酶和嗜热菌蛋白酶。

32. 如权利要求 22、23、或 26 所述的组合物,进一步包括至少一种选自下组的酶,该组由以下各项组成:一种淀粉酶、过氧化氢酶、角质酶、葡聚糖酶、葡糖淀粉酶、葡萄糖异构酶、脂肪酶、植酸酶、支链淀粉酶和木糖异构酶。

33. 如权利要求 22、23、26、或 32 所述的组合物,进一步包括至少一种选自下组的试剂,该组由以下各项组成:一种氯、洗涤剂、次氯酸盐、过氧化氢、草酸、过酸、pH- 调节试剂、磷酸三钠、亚氯酸钠、硝酸钠、表面活性剂、脲和水。

34. 如权利要求 22、23、26、32、或 33 所述的组合物,进一步包括一种产乙醇微生物,例如一种细菌或酵母。

35. 用一种核酸构建体稳定转化的一种宿主细胞,该构建体包括可操作地连接到一个核酸序列的一个启动子,该核酸序列编码一个多肽,该多肽具有选自下组的一个氨基酸序列,该组由 SEQ ID 号:6、8 或 10 或其一个变体或片段组成,其中该细胞是一种细菌细胞、真菌细胞、昆虫细胞、哺乳动物细胞或植物细胞,并且其中该启动子关于该核苷酸序列是异源的。

36. 如权利要求 35 所述的宿主细胞,其中该木聚糖酶具有选自下组的一个氨基酸序列,该组由以下各项组成:SEQ ID 号:6、8 或 10,或其一个活性变体或片段。

37. 如权利要求 36 所述的宿主细胞。其中该核苷酸序列选自下组,该组由以下各项组成:SEQ ID 号:5、7 和 9。

在纤维素生物质的预处理期间有活性的木聚糖酶

[0001] 相关申请的交叉引用

[0002] 本申请要求于 2010 年 9 月 10 日提交的美国临时专利申请序列号 61/381,455 的权益,由此通过引用将其以其全文结合在此。

[0003] 电子提交的序列列表参考

[0004] 序列列表的官方副本是 2010 年 7 月 1 日生成的经由 EFS-Web 以 ASCII 格式的序列列表以名为“72840SequenceListing.txt”的文件进行电子提交的,该副本的大小为 76kb 并且与本说明书同时提交,该副本是于此的一部分并且如果提出则以其全文通过引用结合在此。

技术领域

[0005] 本发明总体上涉及使用具有木聚糖酶活性的酶类来加工木质纤维素材料的方法。

[0006] 发明背景

[0007] 半纤维素是一类植物衍生的杂多糖,它们与纤维素以及木素有关联。最普通的半纤维素是木聚糖、葡糖醛酸木聚糖、阿拉伯木聚糖、葡甘露聚糖以及木葡聚糖。

[0008] 木聚糖酶催化了木聚糖(一种木糖杂聚物)的 1,4- β 骨架的内部水解。木聚糖酶是由以下生物体自然产生的,例如藻类、细菌、真菌、腹足类与原生动物,它们可以利用木糖作为一种碳源用于细胞新陈代谢。参见,Prade (布雷德) (1995)《生物工业遗传工程综述》(Biotech. Genet. Eng. Rev.) 13:100-131。

[0009] 木聚糖酶在饲料、食品 / 饮料和工业产业方面的商业应用(例如,生物质应用)多样。在饲料产业中,木聚糖酶在单胃与反刍动物饲料中用来增加很差地可降解的饲料(例如大麦、青贮饲料以及小麦)的消化率与营养价值。在食品 / 饮料产业中,木聚糖酶在水果与蔬菜加工中用来加工制造蜜、泥以及汁;在酿造和酿酒中用来水解谷物中的粘液物质用于发酵;并且在烘焙中用来改进生面团的品质(例如,弹性以及强度)。在工业产业中,木聚糖酶在造纸中用来减少在木纸浆的漂白过程中的氯消耗与毒物排放;在纺织品加工中用来减少或替代化学沤麻;并且在生物除污 / 生物转化中用来处理 / 再循环废弃物并且加工生物燃料与精细化工产品。

[0010] 在这些应用的许多中,木聚糖酶是与多种其他的木质纤维素酶结合使用,例如纤维素酶、半纤维素酶、木质酶、果胶酶以及蛋白酶。木质纤维素酶因此在饲料、食品 / 饮料与工业行业具有重要的潜在应用。需要用于有效加工木质纤维素材料的新组合物和方法。

[0011] 发明概述

[0012] 提供了多种用于加工木质纤维素材料的组合物以及方法。这些组合物包括至少一种酶,其中该酶具有木聚糖酶活性并且其中该酶在高温与高 pH 下具有活性。木聚糖酶包括耐热的并且嗜碱的木聚糖酶,特别是家族 10 木聚糖酶。由于这些酶的嗜碱性质,它们可以制备为被进一步合并到在高 pH 下(典型地, pH8.5 或更高)进行的方法中的组合物或配制品。这些酶在高温下保持活性。这些组合物还可以包括其他的用于加工木质纤维素的酶类,例如纤维素酶、半纤维素酶、木质酶、果胶酶以及蛋白酶。

[0013] 本发明的这些木聚糖酶可以用来转化感兴趣的生物体。这类生物体包括细菌、真菌、以及植物。因此,披露了已经被改性为表达本发明的这些木聚糖酶中的至少一种的植物、植物部分以及种子。认识到改性为表达木聚糖酶的植物可以展示可以通过从家族 10 类酶选择木聚糖酶而改善的负农学表型。

[0014] 这些方法涉及木聚糖酶或在此描述的组合物在不同饲料、食品 / 饮料和技术应用中的木质纤维素材料中水解木聚糖的用途。这些木聚糖酶对于将木质纤维素材料进行生物转化为更简单多糖或甚至单糖而言是有用的。这些单糖或多糖可以用于例如,饲料或食品 / 饮料的生产,例如基于谷类的动物饲料、麦芽汁或啤酒、奶或奶产品、水果或蔬菜产品;精细化工产品的生产,例如丁醇、乙醇、甲醇和 / 或丙醇;或燃料的生产,包括生物燃料,例如生物乙醇、生物醚、生物柴油以及合成气。木聚糖酶在木质纤维素材料的生物转化中是有用的,因为它是热稳定性的并且可以在高 pH 下使用。

[0015] 以下实施例涵盖在本发明中。

[0016] 1. 包括至少一种木聚糖酶的一种组合物,其中该组合物具有至少约 pH8.5 的 pH,并且其中所述木聚糖酶是一种耐热的并且嗜碱的家族 10 木聚糖酶。

[0017] 2. 如实施例 1 所述的组合物,其中所述木聚糖酶具有选自下组的一个氨基酸序列,该组由以下各项组成:

[0018] a) SEQ ID 号 :6、8 或 10 中列出的一个氨基酸序列;

[0019] b) 与 SEQ ID 号 :6、8 或 10 中列出的氨基酸序列具有至少约 90% 序列一致性的一个氨基酸序列;

[0020] c) 与 SEQ ID 号 :6、8 或 10 中列出的氨基酸序列具有至少约 95% 序列一致性的一个氨基酸序列;

[0021] d) 与 SEQ ID 号 :6、8 或 10 中列出的氨基酸序列具有至少约 99% 序列一致性的一个氨基酸序列;以及,

[0022] e) SEQ ID 号 :6、8 或 10 的一个片段,其中所述片段保持木聚糖酶活性。

[0023] 3. 如实施例 1 所述的组合物,其中该组合物具有约 pH12 的 pH。

[0024] 4. 如实施例 3 所述的组合物,其中该酶在高达约 95°C 的温度下是稳定的。

[0025] 5. 如实施例 1 或 2 所述的组合物,进一步包括选自下组的至少一种酶,该组由以下各项组成:一种纤维素酶、半纤维素酶、木质酶、果胶酶和蛋白酶。

[0026] 6. 如实施例 5 所述的组合物,其中该纤维素酶选自下组,该组由以下各项组成:一种甘露聚糖内切-1,4- β -甘露糖苷酶、1,3- β -D-葡聚糖葡聚糖水解酶、1,3- β -葡聚糖葡聚糖水解酶、1,3-1,4- β -D-葡聚糖葡聚糖水解酶和 1,6- β -D-葡聚糖葡聚糖水解酶。

[0027] 7. 如实施例 5 所述的组合物,其中该半纤维素酶选自下组,该组由以下各项组成:一种 α -L-阿拉伯呋喃糖苷酶、 α -葡萄糖醛酸酶、乙酰基甘露聚糖酯酶、乙酰基木聚糖酯酶、 α -半乳糖苷酶、 β -葡糖苷酶、 β -1,4-木糖苷酶、内切-半乳聚糖酶、内切- β -1,4-甘露聚糖酶、内切- α -1,5-阿拉伯聚糖酶、外切- β -1,4-甘露糖苷酶、外切- β -1,4-木糖苷酶、阿魏酸酯酶(feruloyl esterase)、阿魏酸酯酶(ferulic acid esterase)、对位-熏草酸酯酶、葡糖醛酸木聚糖木聚糖水解酶和木葡聚糖内糖基转移酶。

[0028] 8. 如实施例 5 所述的组合物,其中该木质酶选自下组,该组由以下各项组成:一种二芳基丙烷过氧化酶、葡萄糖氧化酶、乙二醛氧化酶、木素过氧化酶、锰过氧化酶、甲醇氧化

酶、甲醇氧化还原酶、苯酚氧化酶、苯酚过氧化酶和藜芦基醇氧化酶。

[0029] 9. 如实施例 5 所述的组合物,其中该果胶酶是选自下组,该组由以下各项组成:一种果胶降解酶、pectozyme 酶和多聚半乳糖醛酸酶。

[0030] 10. 如实施例 5 所述的组合物,其中该蛋白酶选自下组,该组由以下各项组成:一种萝藦蛋白酶、菠萝酶、caricain 酶、木瓜凝乳蛋白酶、胶原酶、甘氨酸肽链内切酶、胃蛋白酶、链霉蛋白酶、枯草杆菌蛋白酶和嗜热菌蛋白酶。

[0031] 11. 如实施例 1、2、或 5 所述的组合物,进一步包括至少一种选自下组的酶,该组由以下各项组成:一种淀粉酶、过氧化氢酶、角质酶、葡聚糖酶、葡糖淀粉酶、葡萄糖异构酶、脂肪酶、植酸酶、支链淀粉酶和木糖异构酶。

[0032] 12. 如实施例 1、2、5、或 11 所述的组合物,进一步包括至少一种选自下组的试剂,该组由以下各项组成:一种氯、洗涤剂、次氯酸盐、过氧化氢、草酸、过酸、pH- 调节试剂、磷酸三钠、亚氯酸钠、硝酸钠、表面活性剂、脲和水。

[0033] 13. 如实施例 1、2、5、11、或 12 所述的组合物,进一步包括一种产乙醇微生物,例如一种细菌或酵母。

[0034] 14. 用核酸构建体转化的一种宿主细胞,包括任选地连接到编码木聚糖酶的核苷酸序列的一个启动子,其中该细胞是一种细菌细胞、真菌细胞、昆虫细胞、哺乳动物细胞或植物细胞,其中该启动子关于该核苷酸序列是异源的,并且其中该木聚糖酶是一种耐热的并且嗜碱的家族 10 木聚糖酶。

[0035] 15. 如实施例 14 所述的宿主细胞,其中该木聚糖酶具有选自下组的一个氨基酸序列,该组由以下各项组成:SEQ ID 号:6、8 或 10,或其一个活性变体或片段。

[0036] 16. 如实施例 14 所述的宿主细胞。其中该核苷酸序列选自下组,该组由以下各项组成:SEQ ID 号:5、7 和 9。

[0037] 17. 一种加工木质纤维素材料的方法,所述方法包括在其中 pH 为至少 pH8.5 的条件下,使木质纤维素材料与至少一种木聚糖酶接触,并且其中所述木聚糖酶是一种耐热的并且嗜碱的家族 10 木聚糖酶。

[0038] 18. 如实施例 17 所述的方法,其中所述木聚糖酶具有选自下组的一个氨基酸序列,该组由以下各项组成:SEQ ID 号:6、8 或 10 或其一个活性变体或片段。

[0039] 19. 如实施例 17 所述的方法,其中这些条件包括约 95°C 的反应温度。

[0040] 20. 如实施例 17 或 18 所述的方法,其中该木聚糖酶以从约 0.01 单位至约 1000 单位存在。

[0041] 21. 如实施例 17 或 18 所述的方法,其中该木聚糖酶以约 500 单位存在。

[0042] 22. 如实施例 17 或 18 所述的方法,其中该 pH 随时间降低至约 pH6 的 pH。

[0043] 23. 如实施例 17 或 18 所述的方法,其中该接触持续约 12 小时至约 24 小时。

[0044] 24. 如实施例 17 或 18 所述的方法,其中该接触持续约 6 小时。

[0045] 25. 如实施例 17 或 18 所述的方法,其中该接触是在选自下组的至少一种额外的酶存在下进行的,该组由以下各项组成:一种纤维素酶、半纤维素酶、木质酶、果胶酶和蛋白酶。

[0046] 26. 如实施例 25 所述的方法,其中该纤维素酶选自下组,该组由以下各项组成:甘露聚糖内切-1,4- β -甘露糖苷酶、1,3- β -D-葡聚糖葡聚糖水解酶、1,3- β -葡聚糖葡聚糖

水解酶、1, 3-1, 4- β -D-葡聚糖葡聚糖水解酶和 1, 6- β -D-葡聚糖葡聚糖水解酶。

[0047] 27. 如实施例 25 所述的方法, 其中该半纤维素酶选自下组, 该组由以下各项组成: 一种 α -L-阿拉伯呋喃糖苷酶、 α -葡萄糖醛酸酶、乙酰基甘露聚糖酯酶、乙酰基木聚糖酯酶、 α -半乳糖苷酶、 β -葡糖苷酶、 β -1, 4-木糖苷酶、内切-半乳聚糖酶、内切- β -1, 4-甘露聚糖酶、内切- α -1, 5-阿拉伯聚糖酶、外切- β -1, 4-甘露糖苷酶、外切- β -1, 4-木糖苷酶、阿魏酸酯酶(feruloyl esterase)、阿魏酸酯酶(ferulic acid esterase)、对位-熏草酸酯酶、葡糖醛酸木聚糖木聚糖水解酶和木葡聚糖内糖基转移酶。

[0048] 28. 如实施例 25 所述的方法, 其中该木质酶选自下组, 该组由以下各项组成: 一种二芳基丙烷过氧化酶、葡萄糖氧化酶、乙二醛氧化酶、木素过氧化酶、锰过氧化酶、甲醇氧化酶、甲醇氧化还原酶、苯酚氧化酶、苯酚过氧化酶和藜芦基醇氧化酶。

[0049] 29. 如实施例 25 所述的方法, 其中该果胶酶是选自下组, 该组由以下各项组成: 果胶降解酶、pectozyme 酶和多聚半乳糖醛酸酶。

[0050] 30. 如实施例 25 所述的方法, 其中该蛋白酶选自下组, 该组由以下各项组成: 一种萝藦蛋白酶、菠萝酶、caricain 酶、木瓜凝乳蛋白酶、胶原酶、甘氨酸肽链内切酶、胃蛋白酶、链霉菌蛋白酶、枯草杆菌蛋白酶和嗜热菌蛋白酶。

[0051] 31. 如实施例 17、18、或 25 所述的方法, 其中该方法包括用至少一种选自下组的酶进行加工, 该组由以下各项组成: 一种淀粉酶、过氧化氢酶、角质酶、葡聚糖酶、葡糖淀粉酶、葡萄糖异构酶、脂肪酶、植酸酶、支链淀粉酶和木糖异构酶。

[0052] 32. 如实施例 17、18、25、或 31 所述的组合物, 其中该方法进一步包括用至少一种选自下组的试剂进行加工, 该组由以下各项组成: 一种氯、洗涤剂、次氯酸盐、过氧化氢、草酸、过酸、pH-调节试剂、磷酸三钠、亚氯酸钠、硝酸钠、表面活性剂、脲和水。

[0053] 33. 如实施例 17、18、25、31、或 32 所述的方法, 其中该方法进一步包括用一种产乙醇微生物进行加工, 该产乙醇微生物例如是一种细菌或酵母。

[0054] 34. 如实施例 17 至 33 中任一项所述的方法, 其中在步骤 (a) 之前, 通过生物预处理、化学预处理或物理预处理对该木质纤维素材料进行预处理。

[0055] 35. 如实施例 34 所述的方法, 其中该生物预处理包括使用木素增溶微生物。

[0056] 36. 如实施例 35 所述的方法, 其中该化学预处理选自下组, 该组由使用以下处理组成: 碱处理、氨处理、稀酸处理、有机溶剂处理、臭氧处理、二氧化硫处理、二氧化碳处理和 pH 控制的水热处理。

[0057] 37. 如实施例 35 所述的方法, 其中该物理预处理选自下组, 该组由使用以下项组成: 水热处理、辐射、研磨、气烘 / 蒸汽喷发和声处理(超声处理)。

[0058] 从以下描述, 本发明的这些和其他优点、特征、以及目的将变得更好理解。示例性实施例的描述并不旨在将本发明限制在所披露的二聚体形式, 而是相反, 意图是打算覆盖落在如以上实施例和以下权利要求书所限定的本发明的精神和范围内的所有替代方案、等价物和修改。因此应参考用于解释本发明的范围的以上实施例和以下权利要求书。

[0059] 序列表概述

[0060] SEQ ID 号 :1 :衍生自热纤维梭菌的多核苷酸序列, 基因 XynZ 全酶

[0061] SEQ ID 号 :2 :由 SEQ ID 号 :1 编码的多肽

[0062] SEQ ID 号 :3 :衍生自热纤维梭菌的多核苷酸序列, 基因 XynZ 阿魏酸酯酶

(ferrulic acid esterase)

[0063] SEQ ID 号 :4 :由 SEQ ID 号 :3 编码的多肽

[0064] SEQ ID 号 :5 :衍生自海栖热孢菌(*Thermotoga maritime*)的多核苷酸序列,基因 XynA

[0065] SEQ ID 号 :6 :由 SEQ ID 号 :5 编码的多肽

[0066] SEQ ID 号 :7 :衍生自海栖热孢菌的多核苷酸序列,基因 XynB

[0067] SEQ ID 号 :8 :由 SEQ ID 号 :7 编码的多肽

[0068] SEQ ID 号 :9 :衍生自芽胞杆菌属的多核苷酸序列,基因 XynA

[0069] SEQ ID 号 :10 :由 SEQ ID 号 :9 编码的多肽

[0070] 发明详细说明

[0071] 概述

[0072] 在此描述的工作首先示出分离自微生物(例如栖热孢菌)的木聚糖酶可以在木质纤维素材料的化学预处理过程中具有活性。即,在预处理过程中,木聚糖酶水解了木聚糖。这些酶在高 pH 与高温下保持了显著的木聚糖酶活性。木聚糖是半纤维素的一种主要组分,并且与纤维素和木素一起共同称为木质纤维素,构成非木本与木本植物的组要成分。

[0073] 如在此使用,“木质纤维素”或“木质纤维素材料”是指包括多糖聚合物(例如,纤维素、半纤维素以及木素)的非木本和木本植物材料,例如作物、植物和树木生物质。木质纤维素材料还可以包括作物、植物或树木废料产品,包括农业废弃物例如玉米秸秆以及甘蔗渣、专门的能源作物废弃物、木材废料例如锯木厂以及造纸厂废弃物、以及城市废纸。

[0074] 在此描述的这些组合物与方法因此可以用来将木质纤维素材料加工为许多有用的有机化学品、燃料以及产品。例如,一些可以从木质纤维素材料生产的商品与特种化学品包括但不限于,丙酮、乙酸盐、丁二醇、顺式-己二烯二酸、乙醇、乙二醇、糠醛、丙三醇、甘氨酸、赖氨酸、有机酸(例如,乳酸)、1,3-丙二醇、聚羟基烷酸酯、以及木糖。同样地,动物饲料以及不同的食品/饮料可以从木质纤维素材料来生产。总体参见, Lynd (林德) 等人 (1999) 《生物技术进展》(Biotechnol. Prog.) 15:777-793; Philippidis (菲力皮迪斯), 《生物技术手册:生产与利用》(Handbook on Bioethanol: Production and Utilization) 中的 179-212 页“纤维素生物转化技术”(“Cellulose bioconversion technology”), 编辑 Wyman (怀曼) (泰勒与弗朗西斯出版社(Taylor&Francis) 1996); 以及 Ryu (柳) & Mandels (曼德尔斯) (1980) 《酶微生物技术》(Enz. Microb. Technol.) 2:91-102。潜在的共同生产效益超出了在木质纤维素材料中从发酵性糖合成多种有机产品,因为加工之后剩下的富含木素的残余物可以转化为木素衍生的化学品或可以用于电力生产。

[0075] 组合物

[0076] 酶组合物

[0077] 本发明的组合物包括至少一种木聚糖酶或其活性变体或片段,其中这些组合物是在高 pH 下有活性。“高 pH”指一个至少约 pH8.5 的 pH,包括约 pH9.0、约 pH9.5、约 pH10.0、约 pH10.5、约 pH11.0、约 pH11.5、约 pH12.0、约 pH12.5 或更高。其他组分可以包括在包括其他酶类的这些组合物中。

[0078] 本发明的木聚糖酶是耐热的并且嗜碱的家族 10 木聚糖酶。“耐热的”是指该酶在 60°C 30 分钟以后,保持其活性的至少 70%。“嗜碱的”是指该酶在 8.5 的 pH30 分钟以后,保

持其活性的至少 80%。

[0079] 木聚糖酶是水解酶的一个大家族。如在此使用,“木聚糖酶”(xylanase)或“木聚糖酶类”(xylanases)是指一种能够至少将木聚糖水解为木二糖与木三糖的酶。由于木聚糖的异质性与复杂性,木聚糖酶在其折叠模式(即一级序列变体)、水解活性(即,产量、速率以及产物)、作用机制、底物特异性以及物理化学特征方面是相当多样的。参见,《植物生物化学与分子生物学方法》(Methods in Plant Biochemistry and Molecular Biology)(Dashek (达谢克)编辑,CRC 出版社(CRC Press)1997);以及 Collins (柯林斯)等人(2005)《FEMS 微生物学综述》(FEMS Microbiol. Rev.) 29:3-23。木聚糖酶的正式名称是例如内切-1,4- β -木聚糖酶;然而,它在本领域还称为内切木聚糖酶、1,4- β -D-木聚糖-木聚糖水解酶、内切-1,4- β -D-木聚糖酶、 β -1,4-木聚糖酶和 β -木聚糖酶。

[0080] 对木聚糖酶与其他糖苷酶通过催化结构域的一级结构对比而进行分类并且分组进入相关序列的家族。参见,Henrissat (亨利萨特)等人(1989)《基因》(Gene)81:83-95;以及 Henrissat (亨利萨特)等人(2001)《酶学方法》(Methods Enzymol.) 330:183-201。在这一分类系统内,木聚糖酶典型地被限制于家族 10 和 11。由于木聚糖酶在饲料、食物/饮料以及工业产业方面内的商业需求,一些极端嗜性木聚糖酶已经得以分离,特别是来自嗜酸、嗜碱以及嗜热细菌的木聚糖酶。

[0081] 在此感兴趣的是高温(即,热稳定的)以及高 pH (即,嗜碱的)耐受性的分离自微生物(例如,细菌)的展示出木聚糖酶活性的木聚糖酶。这类木聚糖酶是家族 10 木聚糖酶。在一个实施例中,木聚糖酶可以来自一种栖热孢菌《环境微生物学》(Environ. Microbiol.) 61:4403-4408。SEQ ID 号:5 和 7 分别列出了该酶的核苷酸和氨基酸序列。

[0082] 这类木聚糖酶可以自其分离的其他细菌的实例包括但并不局限于芽孢杆菌(例如坚硬芽孢杆菌(GenBank® 登录号 AAQ83581)、耐盐嗜碱芽孢杆菌 S7 (GenBank® 登录号 AAV98623)、耐盐嗜碱芽孢杆菌 C-125 (GenBank® 登录号 BAA00055)、芽孢杆菌菌株 NG-27 (GenBank® 登录号 AAB70918)、以及嗜热脂肪芽孢杆菌 T-6 (GenBank® 登录号 ABI49951));和栖热孢菌(例如海栖热孢菌(GenBank® 登录号 AAD35164))。还参见 Chang (常)等人,(2004)《生物化学与生物物理研究通讯》(Biochem. Biophys. Res. Commun.) 319:1017-1025;Gat (加特)等人,(1994)《应用与环境微生物学》(Appl. Environ. Microbiol.), 60:1889-1896;Gupta (格普塔)等人,(2000)《应用与环境微生物学》(Appl. Environ. Microbiol.), 66:2631-2635;Hamamoto (滨本)等人,(1987)《农业与生物化学》(Agr. Biol. Chem.), 51:953-955;Horikoshi (堀越)(1999)《微生物学与分子生物学综述》(Microbiol. Mol. Biol. Rev.), 63:735-750;Jiang (江)等人,(2006)《应用微生物学与生物技术》(Appl. Microbiol. Biotechnol.), 70:65-71;以及 Mamo (马莫)等人,(2006)《分子生物技术》(Mol. Biotechnol.), 33:149-159。

[0083] 这些酶在高温下显示出最佳木聚糖酶活性并且是热稳定的。在 85°C,这些酶在 pH6.8 与 pH9.3 时的半衰期可以分别是约 24 小时与 6.5 小时。这些酶还可以在一个广泛的 pH 范围显示出实质性的木聚糖酶活性并且是嗜碱的。

[0084] 这些木聚糖酶是耐热的并且热活性的。如在此使用,“耐热的”描述了具有大于 60°C 的熔化温度的酶。如在此使用,“热活性的”描述了在大于 60°C 具有最佳活性的酶。耐热的木聚糖酶在约 60°C 保温 30 分钟后保持至少约 70% 活性,在约 70°C 保温 30 分钟后保持

至少约 65% 活性,在约 80℃ 保温 30 分钟后保持至少约 60% 活性。这些酶在高达约 85℃、86℃、87℃、88℃、89℃、90℃、91℃、92℃、93℃、94℃、95℃、96℃、97℃、98℃、99℃、100℃、101℃、102℃、103℃、104℃、105℃、106℃、107℃、108℃、109℃、110℃、或更高的温度下保持木聚糖酶活性。

[0085] 如在此使用,“嗜碱的”是指这些酶在 pH8.5 保持至少约 70% 活性,在 pH9.0 保持至少 65% 活性,在 pH10.0 保持至少 60% 活性。该木聚糖酶在 pH12.0 或更高时保持活性。

[0086] 如在此使用,“约”是指一个值处于统计学上有意义的范围内,例如一种说明的浓度、时间范围、分子量、体积、温度或 pH。这样的范围可以在给定值或给定范围的一种数量级之内,典型地在 20% 之内,更典型地在 10% 之内,并且甚至更典型地在 5% 之内。“约”所涵盖的可允许的变化将取决于研究的具体系统,并且可以易为本领域一个技术人员所理解。

[0087] 虽然一种全长酶可以用于这些组合中,但是其活性变体或片段也可以使用。如在此所使用,“变体”是指具有与参比酶序列(例如 SEQ ID 号:6、8 或 10)实质性相似的氨基酸序列的一种酶。变体还包括具有与编码该参比酶的核酸(例如 SEQ ID 号:5、7 或 9)实质性相似的核苷酸序列的一个核酸分子。对于分子(例如酶),变体可以包括在参比酶的氨基酸序列中的一个或多个位点添加或删除一个或多个氨基酸,和 / 或在参比酶的氨基酸序列中的一个或多个位点取代一个或多个氨基酸残基。如在此所述,变体可以是活性的,并且因此继续拥有参比酶的所希望的活性,例如木聚糖酶活性。该变体可以产生于例如遗传多态性或产生于人的操作。如通过序列比对程序以及在此的其他地方描述的参数所确定的,该参比酶的活性变体与该参比酶的氨基酸序列可以具有至少约 40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、或更高的序列一致性。活性变体可以不同于该参比酶序列,不同处可少至 1-15 个氨基酸残基、少至 1-10、少至 6-10、少至 5、少至 4、3、2、或甚至 1 个氨基酸残基。

[0088] 可以用不同方式从参比酶改变成变体,这些方式包括核苷酸或氨基酸删除、插入、取代和平截。用于这类操作的方法在本领域是熟知的。

[0089] 例如,该参比酶的氨基酸序列变体可以通过将编码该酶的核苷酸序列进行突变来制备。用于诱变与核苷酸序列改变的方法在本领域是熟知的。参见,例如, Kunkel (昆克) (1985) 《美国科学院院刊》(Proc. Natl. Acad. Sci. USA) 82:488-492; Kunkel (昆克) 等人(1987) 《酶学方法》(Methods in Enzymol.) 154:367-382; 以及《分子生物学技术》(Techniques in Molecular Biology) (Walker (沃克) Gastra (嘉仕堡) 编辑,麦克米兰出版公司(MacMillan Publishing Co. 1983) 以及其中引用的参考文献;连同美国专利号 4,873,192。不影响参考酶的生物活性的适当的氨基酸取代的指导可以发现于 Dayhoff (戴霍夫) 等人(1978) 《蛋白序列与结构图谱》(Atlas of Protein Sequence and Structure) (国家生物医药研究基金会,华盛顿特区) (National Biomedical Research Foundation, Washington, D. C.) 的模型中。

[0090] 并未预期酶的删除、插入和取代产生其特征中的根本性变化。然而,当在这样做以前很难预测取代、删除或插入的确切效果时,本领域的一个技术人员将理解可以通过木聚糖酶活性测定来评估效果。

[0091] 木聚糖酶活性测定在本领域是熟知的。参见,例如, Bailey (贝利) 等人(1992), 《生物技术杂志》(J. Biotechnol.) 23:257-270; Ge (葛) 等人(2007), 《分析生物化学》

(Anal. Biochem.) 362:63-68 ;Gibbs (吉布斯) 等人(1995), 如以上 ;Jeffries (杰夫瑞斯) 等人(1998), 《应用生物化学与生物技术》(Appl. Biochem. Biotechnol.) 70-72:257-265 ; Kenealy (基尼利)&Jeffries (杰夫瑞斯) (2003), 《生物技术通讯》(Biotechnol. Lett.) 25:1619-1623 ;Miller(米勒)(1959), 《分析化学》(Anal. Chem.)31:426-428 ;Taguchi(田口) 等人(1996), 《生物科学生物技术生物化学》(Biosci. Biotechnol. Biochem.)60:983-985 ; Teather (特泽)&Wood (伍德) (1982), 《应用与环境微生物学》(Appl. Environ. Microbiol.) 43:777-780 ;以及 Wang (王)&Broda (布罗达) (1992), 《应用与环境微生物学》(Appl. Environ. Microbiol.) 58:3433-3436。同样地, 用于测定木聚糖酶活性的试剂盒是可商购的, 例如, 购自 Invitrogen 公司(卡尔斯巴德(Carlsbad), 加利福尼亚州) 以及 Megazyme Int'l Ireland Ltd 公司(威克洛(Wicklow), 爱尔兰)。

[0092] 如以上所指出, 这些组合物可以使用该木聚糖酶的活性片段。如在此使用的, “片段”是指该参考酶的一部分, 该部分保留了木聚糖酶活性。片段还指为编码该参比酶的核酸分子的一部分。该酶的活性片段可以例如通过将表达该酶的经编码的片段的酶编码核酸分子的一部分进行分离(例如, 利用体外重组体表达), 并评估该片段的活性来制备。编码此类片段的核酸分子可以是至少约 150、200、250、300、350、400、450、500、550、600、650、700、800、900、1,000、1,100、1,200、1,300 或 1,400 个连续核苷酸, 或高达存在于在此披露的全长酶编码核酸分子中的核苷酸数目。按照这样, 多肽片段可以是至少约 50、60、70、80、90、100、125、150、175、200、225 或 250 个连续氨基酸残基, 或高达存在于该全长酶中的氨基酸残基总数目。

[0093] 确定任何两种序列之间的序列一致性百分比可以使用一种数学算法来完成。此类数学算法的非限制性实例是 Myers (迈尔斯)与 Miller (米勒) (1988)CABIOS4:11-17 的算法 ;Smith (史密斯) 等人(1981) 《应用数学进展》(Adv. Appl. Math.) 2:482-489 的局部比对算法 ;Needleman (尼德曼)与 Wunsch (翁施) (1970) 《分子生物学杂志》(J. Mol. Biol.) 48:443-453 的全局比对算法 ;Pearson (皮尔逊)&Lipman (利普曼) (1988) 《美国科学院院刊》85:2444-2448 的局部搜索比对法(search-for-local alignment method);Karlin (卡尔林)与 Altschul (阿尔丘尔) (1990) 《美国科学院院刊》87:2264-2268 的算法, 如在 Karlin (卡尔林)与 Altschul (阿尔丘尔) (1993) 《美国科学院院刊》90:5873-5877 中的修改。

[0094] 这些数学算法的电脑执行能用以比较序列以确定序列一致性。此类执行包括但不限于 :PC/Gene 程序中的 CLUSTAL (Intelligenetics 公司, 山景城, 加利福尼亚州 (Mountain View, CA));ALIGN 程序(版本 2.0 ;Corpet (科佩特) 等人(1988) 《核酸研究》(Nucleic Acids Res.) 16:10881-10890 ;Higgins (希金斯) 等人(1988) 《基因》(Gene) 73:237-244 ;Higgins (希金斯) 等人(1989) CABIOS5:151-153 ;Huang (黄) 等人(1992) CABIOS8:155-165 ;以及 Pearson (皮尔逊) 等人(1994) 《分子生物学方法》(Meth. Mol. Biol.) 24:307-331) 以及 GCG 威斯康辛遗传软件包(GCG Wisconsin Genetics Software Package) 第 10 版中的 GAP、BESTFIT、BLAST、FASTA 与 TFASTA (Accelrys Inc. 公司 ;圣迭哥, 加利福尼亚州 (San Diego, CA))。利用这些程序的比对可以使用缺省参数来进行。

[0095] 如在此使用, 在上下文中两种核苷酸或氨基酸序列的“序列一致性”或“一致性”是指, 当对一个特定区域的最大相符性进行比对时, 这两种序列中的残基是相同的。在对于蛋

白例如酶使用序列一致性的百分比时,公认地,不一致的残基位置通常因保守氨基酸取代而不同,其中氨基酸残基被取代为其他具有相似化学特性(例如,电荷或疏水性)的氨基酸残基并且因此并不改变该酶的功能特性。当序列在保守取代方面不同时,该序列一致性百分比可以向上调整以对该取代的保守性质进行校正。因此类保守取代而不同的序列被认为是具有“序列相似性”或“相似性”。用于进行这种调整的手段是本领域一个技术人员熟知的。保守取代的分数是例如在 PC/GENE 程序(PC/GENE Program)中实施的进行计算。

[0096] 如在此使用,“序列一致性百分比”是指一种通过在指定区域比较两种最佳比对的序列而确定的值,其中在该指定区域中的核苷酸或氨基酸序列部分与该参比序列(它不包括添加或删除)相比,可以包括用于这两种序列最佳比对的添加或删除(即,缺口)。该百分比是通过以下计算的:确定位置的数目(在这些位置,该一致的核苷酸或氨基酸残基在两种序列中都存在,以产生匹配位置的数目)、用比较窗口中的位置总数目除以该匹配位置数目、并且将该结果乘以 100 以产生序列一致性百分比。

[0097] 如此,在该组合物中,木聚糖酶可以作为一种部分或完全纯化的全长酶、或作为其活性变体或片段来提供,或甚至可以作为一种产酶微生物、植物、植物部分或植物宿主细胞来提供。在任何情况下都不要要求完全纯化。该酶、变体或片段因此可以是至少约 5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99% 纯,或更纯。

[0098] 纯化多肽与蛋白(例如酶类)的方法在本领域是熟知的。参见,例如, Cesar (恺撒) & Mrša (1996),《酶与微生物技术》(Enzyme Microb. Tech.) 19:289-296 ;Chivero (切沃罗) 等人(2001),《食品化学》(Food Chem.) 72:179-185 ;de Albuquerque Lucena-Neto 等人(2004),《英国微生物杂志》(Br. J. Microb.) 35:86-90 ;Ehle (艾尔) & Horn (霍恩)(1990),《生物分离》(Bioseparation) 1:97-110 ;Gupta (格普塔) 等人(2002),《生物技术通讯》(Biotechnol. Lett.) 24:2005-2009 ;Hengen (亨根)(1995),《生物化学科学趋势》(Trends Biochem Sci.) 20:285-286 ;Kanda (神田) 等人(1985),《生物化学杂志》(J. Biochem.) 98:1545-1554 ;Khasin (哈辛) 等人(1993),《应用与环境微生物学》(Appl. Environ. Microb.) 59:1725-1730 ;Kudo (工藤) 等人(1985),《普通微生物学杂志》(J. Gen. Microbiol.) 131:2825-2830 ;Regnier (雷古纳)(1983),《科学》(Science) 222:245-252 ;Roy (罗伊) & Uddin (乌丁)(2004),《巴基斯坦生物科学杂志》(Pak. J. Biol. Sci.) 7:372-379 ;Royer (罗耶) & Nakas (纳卡斯)(1991),《欧洲生物化学杂志》(Eu. J. Biochem.) 202:521-529 ;Sá-Pereira 等人(2003),《分子生物技术》(Mol. Biotechnol.) 24:257-281 ;Shaw (萧),《分子生物学方法》第 32 卷中的通过反相 HPLC (高效液相色谱法)的肽纯化(“Peptide purification by reverse-phase HPLC”257-287 In: Methods in Molecular Biology, Vol. 32) (Walker (沃克) 编辑, Humana 出版社, 1994) (Walker ed., Humana Press 1994) ;Simpson (辛普森) 等人(1991),《生物化学杂志》(Biochem. J.) 277:419-417 ;《化学技术百科全书》第 52 版(Encyclopedia of Chemical Technology, 52nd ed.) (Kirk (科克) -Othmer (奥斯默) 编辑, 威利交互科学, 2007) (Kirk-Othmer ed., Wiley-Interscience 2007) ;蛋白纯化与分析基本方法:实验室手册(Basic Methods in Protein Purification and Analysis: A Laboratory Manual) (Simpson (辛普森) 等人编辑,冷泉港实验室出版社(Cold Spring Harbor Laboratory Press), 2008) ;《酶纯化

与相关技术》(Enzyme Purification and Related Techniques)。《酶学方法》第 22 卷 (Methods in Enzymology, Vol. 22)(Jakoby (雅各布)编辑, 学术出版公司(Academic Press Inc.)1971)《酶学方法:亲和技术-酶纯化:B 部分:第 34 卷:亲和技术 B 部分》(Methods in Enzymology:Affinity Techniques - Enzyme Purification:Part B:Vol. 34:Affinity Techniques Part B)(Kaplan (卡普兰)等人编辑, 爱思唯尔出版公司(Elsevier)1974); 以及美国专利申请号 2009/0137022 与 2009/0239262; 以及美国专利申请号 4, 347, 322; 4, 634, 673; 4, 725, 544 与 5, 437, 992。适合酶类的纯化技术的实例包括但不限于, 沉淀, 例如硫酸铵沉淀; 基于分子大小的分离, 例如凝胶过滤; 基于电荷的分离, 例如离子交换层析; 基于与其他生物分子特定相互作用的分离, 例如生物亲和层析或氨基酸序列的抗体识别; 基于其他原理的分离, 例如疏水相互作用层析或羟磷灰石层析。

[0099] 无论这些组合物包括一种全长木聚糖酶, 或其一个活性变体或片段, 它们可以包括每克要处理的木质纤维素材料约 0.01 单位至约 1000 单位、约 0.1 单位至约 500 单位、或约 1 单位至约 50 单位的酶活性。可替代地, 这些组合物可以包括每克要处理的木质纤维素材料约 1 单位至约 10 单位、约 10 单位至约 100 单位、约 100 单位至约 200 单位、约 200 单位至约 300 单位、约 300 单位至约 400 单位、约 400 单位至约 500 单位、约 500 单位至约 600 单位、约 600 单位至约 700 单位、约 700 单位至约 800 单位、约 800 单位至约 900 单位、约 900 大内至约 1000 单位、或更多的酶活性。一单位可以定义为在一个测定温度下每分钟释放 $1 \mu\text{mol}$ 木糖所要求的酶量。如果使用一个用于确定木聚糖酶活性的还原糖测定, 一单位可以定义为在一个测定温度下每分钟(minut) 释放 1 微摩尔木糖等效物所要求的酶量。

[0100] 这些组合物可以仅包括该木聚糖酶、其变体或片段, 或可以是一种混合物(即, “鸡尾酒”(“cocktail”)), 该混合物具有该木聚糖酶、其变体或片段以及至少一种或多种其他酶, 这些其他酶包括其他的在加工木质纤维素材料中有用的木质纤维素酶类, 例如其他的纤维素酶、半纤维素酶、木质酶、果胶酶以及蛋白酶。如在此使用的, “木质纤维素酶”是指一种酶, 该酶加工木质纤维素材料, 例如纤维素、半纤维素与木素, 以及其他多糖和蛋白组分, 它们可以是木质纤维素的一部分或与木质纤维素有关。

[0101] 纤维素酶的实例包括但不限于, 甘露聚糖内切-1, 4- β -甘露糖苷酶、1, 3- β -D-葡聚糖葡聚糖水解酶、1, 3- β -葡聚糖葡聚糖水解酶、1, 3-1, 4- β -D-葡聚糖葡聚糖水解酶和 1, 6- β -D-葡聚糖葡聚糖水解酶。

[0102] 半纤维素酶的实例包括但不限于, α -L-阿拉伯呋喃糖苷酶、 α -葡萄糖醛酸酶、乙酰基甘露聚糖酯酶、乙酰基木聚糖酯酶、 α -半乳糖苷酶、 β -葡糖苷酶、 β -1, 4-木糖苷酶、内切-半乳聚糖酶、内切- β -1, 4-甘露聚糖酶、内切- α -1, 5-阿拉伯聚糖酶、外切- β -1, 4-甘露糖苷酶、外切- β -1, 4-木糖苷酶、阿魏酸酯酶(feruloyl esterase)、阿魏酸酯酶(ferulic acid esterase)、对位-熏草酸酯酶、葡糖醛酸木聚糖木聚糖水解酶和木葡聚糖内糖基转移酶。

[0103] 木质酶的实例包括但不限于, 二芳基丙烷过氧化酶、葡萄糖氧化酶、乙二醛氧化酶、木素过氧化物酶(LiP)、锰过氧化酶、甲醇氧化酶、甲醇氧化还原酶、苯酚氧化酶(虫漆酶)、苯酚过氧化酶和藜芦基醇氧化酶。

[0104] 果胶酶的实例包括但不限于, 果胶降解酶、pectozyme 酶和多聚半乳糖醛酸酶。

[0105] 蛋白酶的实例包括但不限于, 萝藦蛋白酶、菠萝酶、caricain 酶、木瓜凝乳蛋白

酶、胶原酶、甘氨酸肽链内切酶、胃蛋白酶、链霉菌蛋白酶、枯草杆菌蛋白酶和嗜热菌蛋白酶。

[0106] 这些组合物还可以包括其他酶类,例如淀粉酶、过氧化氢酶、角质酶、葡聚糖酶、葡糖淀粉酶、葡萄糖异构酶、脂肪酶、植酸酶、支链淀粉酶和木糖异构酶。

[0107] 如以上,这些酶中的任一种都可以作为部分或完全纯化的全长酶、或其活性变体或片段来提供,或者可以作为一种产酶微生物来提供。此外,可以按有效水解它们的底物的一个量来提供这些酶中的任何一种,这些量例如从约 0.001wt.% 至约 50wt.%、从约 0.01wt.% 至约 50wt.%、从约 0.1wt.% 至约 50wt.%、从约 1wt.% 至约 50wt.%、从约 10wt.% 至约 50wt.%、从约 20wt.% 至约 50wt.%、从约 30wt.% 至约 50wt.%、从约 40wt.% 至约 50wt.%、或更多。

[0108] 这些组合物还可以包括在加工木质纤维素材料中典型地使用的试剂,例如一种氯、洗涤剂、次氯酸盐、过氧化氢、草酸、过酸、pH- 调节试剂、磷酸三钠、亚氯酸钠、硝酸钠、表面活性剂、脲和水。

[0109] 洗涤剂的实例包括但不局限于,阴离子、阳离子或中性洗涤剂,例如诺纳(Nonidet)(N)P-40、十二烷基硫酸钠(SDS)、月桂基硫酸钠(SLS)、磺基甜菜碱、正-辛基葡糖苷、脱氧胆酸盐、**Triton®X-100**(道化学公司(Dow Chemical Co.));米德兰,密歇根州(Midland, MI))以及**Tween® 20**(ICI 美国有限公司(ICI Americas, Inc.));布里奇沃特,新泽西州(Bridgewater, NJ))。

[0110] 过酸的实例包括但不局限于,间氯过氧苯甲酸、过苯甲酸、高氯酸、邻羧基过苯甲酸、过马来酸、过乙酸、过甲酸、过氧丙酸(perpropionic acid)和对硝基过苯甲酸。

[0111] 表面活性剂的实例包括但并不局限于,一种仲醇乙氧基化物、脂肪醇乙氧基化物、壬基酚乙氧基化物、脂肪醇的磷酸酯、聚氧乙烯醚、聚乙二醇、聚氧乙烯化的烷基酚以及硬脂酸以及十三烷基乙氧基化物。

[0112] 这些试剂中的任一种都可以做为部分或完全纯化的来提供。此外,可以按从约 0.001wt.% 至约 50wt.%、从约 0.01wt.% 至约 50wt.%、从约 0.1wt.% 至约 50wt.%、从约 1wt.% 至约 50wt.%、从约 10wt.% 至约 50wt.%、从约 20wt.% 至约 50wt.%、从约 30wt.% 至约 50wt.%、从约 40wt.% 至约 50wt.%、或更多的量提供这些试剂中的任一种。

[0113] 这些组合物还可以包括微生物,尤其是产乙醇的和 / 或木素增溶的微生物,例如细菌或酵母。总体参见,Burchhardt (布哈特)&Ingram (英格拉姆)(1992),《应用与环境微生物学》58:1128-1133 ;Dien (迪恩)等人(1998),《酶与微生物技术》23:366-371 ;Keating (基廷)等人(2004),《酶与微生物技术》35:242-253 ;Lawford (劳福德)&Rousseau (卢梭)(1997),《应用生物化学与生物技术》63-65:221-241 ;《生物技术手册:生产与利用》(Wyman (怀曼)编辑,CRC 出版社 1996);连同美国专利申请号 2009/0246841 与 2009/0286293 ;以及美国专利申请号 6,333,181。此类微生物可以表达在加工木质纤维素材料中辅助的酶类,例如醇脱氢酶、丙酮酸脱羧酶、转醛醇酶、转酮醇酶丙酮酸脱羧酶(transketolasepyruvate decarboxylase)、木糖还原酶、木糖醇脱氢酶或木糖异构酶木酮糖激酶。此类微生物的实例包括但并不局限于,以下属中的成员,例如假丝酵母属、欧文菌属、埃希氏菌属、克雷白氏菌属、毕赤酵母属、酵母菌属、链霉菌属和发酵单胞菌属。参见,例如,Dien (迪恩)(1998),如以上 ;Ingram (英格拉姆)&Conway (康威)((1988),《应用与环境微生物学》54:397-404 ;Jarboe (扎博)等人(2007),《生物化学工程 / 生物技术进展》(Adv. Biochem. Engin. /

Biotechnol.) 108:237-261 ;Keating (基廷) 等人(2004),《工业微生物生物技术杂志》(J. Indust. Microbiol. Biotech.)31:235-244 ;Keating (基廷)等人(2006),《生物技术与生物工程学》(Biotechnol. Bioeng.)93:1196-1206 ;Pasti (帕斯提)等人(1990),《应用与环境微生物学》56:2213-2218 ;以及 Zhang (张)等人(1995),《科学》(Science)267:240-243。

[0114] 具有该酶、其变体或片段,连同以上描述的任何其他酶类和 / 或试剂的组合物可以制备为液体、膏剂、固体或凝胶。

[0115] 转基因生物

[0116] 本发明的组合物包括用编码至少一种木聚糖酶、或其具有木聚糖酶活性的变体或片段的核酸构建体转化的植物、植物部分和植物宿主细胞。如在此使用,“植物部分”是指非木本和木本植物细胞、植物原生质体、植物可以由之再生的植物细胞组织培养物、植物愈伤组织、植物簇(plant clumps)、以及在植物或以下植物的部分中完整的植物细胞,这些植物的部分例如是胚、花粉、胚珠、种子、叶、花、枝、果实、核、穗、穗轴、外壳、茎、根、根尖、花药、块茎、根茎、以及类似物。转化的植物旨在用于生产纤维素乙醇。本发明的转基因植物将用作在木质纤维素转化过程中木聚糖酶活性的来源。可以结合含有其他酶的其他转基因植物(例如 AMY797E,一种表达热稳定 α -淀粉酶的转基因植物)使用本发明的转基因植物。

[0117] 植物的实例包括但并不局限于,大麦、豆类(beans in general)、芸薹属、三叶苜蓿、可可、咖啡、棉花、亚麻、玉米、粟、花生、油菜 / 卡罗拉(canola)、稻、黑麦、红花、高粱、大豆、甘蔗、甜菜、向日葵、甘薯、茶以及小麦 ;蔬菜类例如西兰花、甘蓝小包菜、甘蓝、胡萝卜、木薯、花椰菜、葫芦科植物、滨豆、莴苣、豌豆、胡椒、马铃薯、萝卜以及番茄 ;草类例如紫花苜蓿、百慕达草、象草、盖氏虎尾草、高羊茅草、高小麦草、芒属以及柳枝稷 ;果树类例如苹果、杏、鳄梨、香蕉、柑橘、椰子、梨、桃、菠萝以及核桃 ;以及花卉例如康乃馨、兰花和玫瑰。

[0118] 为了防止与植物中的木聚糖酶表达相关的负表型,认识到将需要小心地控制植物中木聚糖酶的表达、活化、或定位。如在此进一步讨论,控制负表型的一种方法是在植物中隔绝表达的木聚糖酶。特别是,信号序列或靶向序列可以用于将该酶靶向至细胞器,例如液泡、内质网、叶绿体、以及类似物。以这种方式,将一种信号序列或靶向序列可操作地连接至该木聚糖酶上。因此,用于在植物中表达该酶的一种 DNA 构建体或表达盒将包括一种可操作地连接至该木聚糖酶的编码序列的信号序列。此类靶向序列会将该木聚糖酶靶向至液泡、内质网、叶绿体、以及类似物。

[0119] 此类序列的实例包括但不局限于,针对酰基载体蛋白的转运肽、二磷酸核酮糖羧化酶(RUBISCO)的小亚基、植物 EPSP 合酶、以及类似物。参见例如 Archer (阿彻) 等人(1990)《生物能与生物膜杂志》(J. Bioenerg. Biomemb.) 22:789-810), Clark (克拉克) 等人(1989)《生物化学杂志》(J. Biol. Chem.) 264:17544-17550 ;Daniell (丹尼尔) (1999)《自然生物技术》(Nat. Biotech.)17:855-856 ;de Castro Silva Filho 等人(1996)《植物分子生物学》(Plant Mol. Biol.)30:769-780 ;Della (黛拉)-Cioppa (塞奥帕)等人(1987)《植物生理学》(Plant Physiol.) 84:965-968 ;Lamppa (兰帕) 等人(1988)《生物化学杂志》263:14996-14999 ;Lawrence (劳伦斯)等人(1997)《生物化学杂志》272:20357-20363 ;Romer (罗默) 等人(1993)《生物化学与生物物理研究通讯》196:1414-1421 ;Schmidt (施密特) 等人(1993)《生物化学杂志》268:27447-27457 ;Schnell (施耐尔) 等人(1991)《生物化学杂志》266:3335-3342);Shah et al. (1986)《科学》233:478-481 ;Von Heijne (冯

海耶)等人(1991)《植物分子生物学报》(Plant Mol. Biol. Rep.) 9:104-126;以及 Zhao (赵)等人(1995)《生物化学杂志》270:6081-6087;连同美国专利号 6, 338, 168。

[0120] 内质网信号肽包括描述于 Raikhel (莱克尔)&Chrispeels (克里斯皮尔斯)《植物生物化学与分子生物学》中的“蛋白分选与囊泡运输”(“Protein sorting and vesicle traffic”In:Biochemistry and Molecular Biology of Plants) (Buchanan (布坎南)等人编辑,美国植物生理学家学会(American Society of Plant Physiologists)2000)中的那些。参见,例如, Denecke 等人(1992),《EMBO 杂志》(EMBO J.) 11:2345-2355;Denecke (丹耐克)等人(1993),《实验植物学杂志》(J. Exp. Bot.) 44:213-221;Gomord (谷莫德)等人(1996),《植物生理学与生物化学》(Plant Physiol. Biochem.) 34:165-181;Lehmann (莱曼)等人(2001),《植物生理学》127:436-449;Munro (芒罗)&Pelham (佩勒姆)(1986),《细胞》(cell) 46:291-300;Munro (芒罗)&Pelham (佩勒姆)(1987),《细胞》48:899-907;Vitale (维塔尔)等人(1993),《实验植物学杂志》44:1417-1444;以及 Wandelt (万德尔特)等人(1992),《植物杂志》(Plant J.) 2:181-192。

[0121] 在一些情况下,令人希望的是将该木聚糖酶定位于液泡内。为了木聚糖酶的液泡靶向表达,将植物用载体进行转化,这些载体包括一种液泡靶向序列,例如来自烟草甲壳酶基因的靶向序列。液泡分选信号序列包括大麦多胺氧化酶 2 (BPAO2) 信号序列(Cervelli (瑟维里)等人(2004)《植物杂志》(The Plant Journal) 40:410-418)。对于其他的液泡分选信号,参见例如美国专利申请号 12/359, 421 以及 Raikhel (莱克尔)&Chrispeels (克里斯皮尔斯)(2000),如以上。

[0122] 质体靶向序列在本领域是熟知的并且可以用于本发明的惯例中。参见,例如, Clark (克拉克)等人(1989),《生物化学杂志》264:17544-17550;Della (黛拉)-Cioppa (塞奥帕)等人(1987),如以上;Romer (罗默)等人(1993),《生物化学与生物物理学研究通讯》196:1414-1421;Shah (沙阿)等人(1986),《科学》233:478-481;以及 Von Heijne (冯海耶)等人(1991),《植物分子生物学报》9:104-126。其他的靶向叶绿体序列包括核酮糖-1, 5-二磷酸羧化酶(Rubisco)的叶绿体小亚基(de Castro Silva Filho 等人(1996),《植物分子生物学》30:769-780;以及 Schnell (施耐尔)等人(1991),《生物化学杂志》266:3335-3342);5-(烯醇丙酮)莽草酸-3-磷酸合酶(EPSPS) (Archer (阿彻)等人(1990),如以上);色氨酸合酶(Zhao (赵)等人(1995),如以上);质体蓝素(Lawrence (劳伦斯)等人(1997),如以上);分支酸合酶(Schmidt (施密特)等人(1993),如以上);以及集光叶绿素 a/b 结合蛋白(LHBP) (Lamppa (兰帕)等人(1988),如以上)。本领域一个技术人员还可以设想产生在其中该叶绿体已经得以转化的转基因植物,以过表达该木聚糖酶。参见,例如,Daniell (丹尼尔)(1999),《自然生物技术》17:855-856;连同美国专利号 6, 338, 168。

[0123] 用于转化叶绿体的方法在本领域是熟知的。参见,例如 Svab (斯瓦博)等人(1990),《美国科学院院刊》87:8526-8530;Svab (斯瓦博)&Maliga (马利加)(1993),《美国科学院院刊》90:913-917;Svab (斯瓦博)&Maliga (马利加)(1993),《EMBO 杂志》12:601-606。同样地,可以针对在该叶绿体中的表达将有待靶向至该叶绿体的核苷酸序列进行最优化,以考虑植物细胞核与这种细胞器之间在密码子使用上的不同。以这种方式,可以使用叶绿体偏好密码子合成感兴趣的多核苷酸。参见,例如,美国专利号 5, 380, 831。

[0124] 为了该木聚糖酶定位于该质外体内,可以使用一种质外体靶向序列,例如一种玉

米 γ -玉米素 N 端信号序列(美国专利号 7, 102, 057)。

[0125] 额外的有关靶向多肽的指导可以发现于,例如 Bruce (布鲁斯) ((2001), *Biochim Biophys Acta*1541:2-21 ;Emanuelsson (埃曼纽尔森) 等人(2000),《分子生物学杂志》300:1005-1016 ;Emanuelsson(埃曼纽尔森)&von Heijne(冯海耶)(2001),*Biochim Biophys Acta*1541:114-119 ;Hadlington (哈丁顿)&Denecke (丹耐克) (2000),《植物生物学现行观点》(*Curr. Opin. Plant Biol.*)3:461-468 ;Nicchitta (尼基塔) (2002),《植物生物学现行观点》14:412-416 ;以及 Silva-Filho (2003),《植物生物学现行观点》6:589-595。

[0126] 简言之,在植物部分(例如植物宿主细胞)中表达一种核苷酸序列的方法包括,用例如一种核酸构建体转化植物宿主细胞,该构建体是例如一种表达盒,该表达盒具有针对一种木聚糖酶、其变体或片段的可操作地连接至一种或多种调节序列上的编码序列。用于一些实施例的木聚糖酶编码序列包括 SEQ ID 号 :5、7 或 9 中列出的核苷酸序列,或其片段或变体,或 SEQ ID 号 :6、8 或 10 中列出的多肽序列或其片段或变体。转化的植物宿主细胞是在该核苷酸序列得以在该植物宿主细胞中表达的条件下生长的。在一些情况下,这些方法包括一种额外的步骤,该步骤是将该经转化的植物宿主细胞再生成为一种形态学上正常的可育植物。

[0127] 如在此使用,“表达盒”是指一种核酸分子,该核酸分子具有至少一种可操作地连接至感兴趣(例如木聚糖酶)的编码序列上的控制序列。以这种方式,可操作地与针对该木聚糖酶的核苷酸序列相互作用的植物启动子在表达盒中提供,用于在一种植物、植物部分或植物宿主细胞中表达。

[0128] 表达该木聚糖酶的一种植物、植物部分或植物宿主细胞因此可以通过将针对一种木聚糖酶的编码序列引入该植物、植物部分或植物宿主细胞来获得。优选地,该编码序列可以存在于一个表达盒中或另一核酸构建体(例如载体)中。

[0129] 感兴趣的编码序列在表达盒中的表达可以是在一个组成型启动子或一个诱导型启动子的控制下,只有当该宿主细胞暴露于一些具体外部刺激时,该启动子才引发转录。可替代地,该表达盒可以处在一种组织特异性启动子的控制下,该启动子在具体的组织、器官、或者甚至在具体的发育阶段起作用。

[0130] 为了本发明的目的,对于宿主细胞或对于彼此,调节区(即启动子、转录调节区、以及转录终止区)可以是天然的 / 同功的。可替代地,对于宿主细胞或对于彼此,调节区可以是异源的。如在此使用,提及一种序列的“异源的”意为一种序列,该序列源自一种外来物种,或者如果源自相同物种,则是通过有意的人工介入对其天然形式在构成和 / 或基因座方面进行了实质性的修改。例如,可操作地连接至一种异源多核苷酸上的启动子是来自一种物种,该物种与该多核苷酸从其衍生而来的物种不同,或者,如果来自相同 / 类似物种,则一者或两者在其原始形式和 / 或基因座方面进行了修改,或者该启动子对于该可操作地连接的多核苷酸而言不是天然的启动子。

[0131] 有待使用的启动子的选择取决于几个因素,包括但不限于 :细胞 - 或组织 - 特异性表达、所希望的表达水平、效率、可诱导性以及可选择性。例如,当希望在特定组织或器官中表达时,可以使用组织特异性启动子。相反,当希望响应刺激表达时,可以使用可诱导性启动子。当希望遍布植物的细胞中连续表达时,可以使用组成型启动子。对于本领域一个技术人员而言,通过相对于一种核苷酸序列来适当地选择并且定位启动子以及其他的调节

区而调整该核苷酸序列的表达是一种惯例。

[0132] 因此, 在一些情况下, 可以使用组成型启动子。组成型启动子的实例包括但不限于, 水稻肌动蛋白 1 启动子(Wang (王) 等人(1992), 《分子和细胞生物学》(Mol. Cell. Biol.)12:3399-3406; 以及美国专利号 5,641,876)、CaMV35S 启动子(Ode11 (奥德尔) 等人(1985), 《自然》(Nature)313:810-812)、CaMV19S 启动子(Lawton (劳顿) 等人(1987), 《植物分子生物学》9:315-324)、nos 启动子(Ebert (埃伯特) 等人(1987), 《美国科学院院刊》84:5745-5749)、Adh 启动子(Walker(沃克) 等人(1987), 《美国科学院院刊》84:6624-6629)、蔗糖合酶启动子(Yang (杨)&Russell (拉塞尔) (1990), 《美国科学院院刊》87:4144-4148)、以及泛素启动子。

[0133] 此外, 在植物中已经报道了几个组织特异性调节基因和 / 或启动子。一些报道的组织特异性基因包括对种子贮藏蛋白(例如 β -伴大豆球蛋白、十字花科蛋白、油菜籽蛋白和菜豆素), 玉米素或油体蛋白(例如 oleosin), 或脂肪酸生物合成中涉及的蛋白(包括酰基载体蛋白、硬脂酰 -ACP 去饱和酶和脂肪酸去饱和酶(fad2-1)), 以及在胚发育中表达的其他基因(例如 Bce4, 参见例如 Kridl (克里德) 等人(1991) 《种子科学研究》(Seed Sci. Res.)1:209-219; 连同欧洲专利号 255378)。组织特异性启动子的额外实例包括但不限于, 凝集素启动子(Lindstrom (林斯特龙) 等人(1990), Der. Genet. 11:160-167; 以及 Vodkin (沃德金) (1983), 《临床生物学研究进展》(Prog. Clin. Biol. Res.)138:87-98)、玉米醇脱氢酶 1 启动子(Dennis (丹尼斯) 等人(1984), 《核酸研究》12:3983-4000)、玉米集光复合体启动子(Bansal (班赛尔) 等人(1992), 《美国科学院院刊》89:3654-3658)、玉米热休克蛋白启动子(O' De11 (欧德尔) 等人(1985), 《EMBO 杂志》5:451-458; 以及 Rochester (罗彻斯特) 等人(1986), 《EMBO 杂志》5:451-458)、豌豆小亚基 RuBP 羧化酶启动子(Cashmore (卡什莫尔), 《植物遗传工程》29-39 中的编码核酮糖 -1,5-二磷酸羧化酶小亚基的核基因("Nuclear genes encoding the small subunit of ribulose-1,5-bisphosphate carboxylase"29-39In:Genetic Engineering of Plants)(Hollaender(霍兰德)编辑, 普力纳姆出版社(Plenum Press)1983); 以及 Poulsen(波尔森) 等人(1986), 《分子和普通遗传学》(Mol. Gen. Genet.)205:193-200)、Ti 质粒甘露氨酸合酶启动子(Langridge (兰格里奇) 等人(1989), 《美国科学院院刊》86:3219-3223)、Ti 质粒胭脂氨酸合酶启动子(Langridge (兰格里奇) 等人(1989), 如以上)、矮牵牛查尔酮异构酶启动子(van Tunen (范杜能) 等人(1988), 《EMBO 杂志》7:1257-1263)、大豆甘氨酸富含蛋白 1 启动子(Keller (凯勒) 等人(1989), 《基因与发育》(Genes Dev.)3:1639-1646)、截短的 CaMV35S 启动子(O' De11 (欧德尔) 等人(1985), 《自然》313:810-812)、马铃薯块茎储藏蛋白启动子(Wenzler (温茨勒) 等人(1989), 《植物分子生物学》13:347-354)、根细胞启动子(Yamamoto (山本) 等人(1990), 《核酸研究》18:7449)、玉米玉米素启动子(Kriz (柯立兹) 等人(1987), 《分子和普通遗传学》207:90-98; Langridge (兰格里奇) 等人(1983), 《细胞》34:1015-1022; Reina (雷纳) 等人(1990), 《核酸研究》18:6425; Reina (雷纳) 等人(1990), 《核酸研究》18:7449; 以及 Wandelt (万德尔特) 等人(1989), 《核酸研究》17:2354)、球蛋白 -1 启动子(Belanger (贝朗格) 等人(1991), 《遗传学》(Genetics)129:863-872)、 α -微管素 cab 启动子(Sullivan (苏利文) 等人(1989) 《分子和普通遗传学》215:431-440)、PEPCase 启动子(Hudspeth (哈得斯佩斯) &Gru1a (格鲁拉) (1989), 《植物分子生物学》(Plant Mol. Biol.)12:579-589), R 基因复合

体相关启动子(Chandler (钱德勒)等人(1989)《植物细胞》(Plant Cell)1:1175-1183),以及查耳酮合酶启动子(Franken (弗兰肯)等人(1991)《EMBO 杂志》10:2605-2612)。对于种子特异性表达特别有用的是豌豆球蛋白启动子(Czako (查克)等人(1992),《分子和普通遗传学》235:33-40;以及美国专利号 5,625,136)。用于在成熟的叶子中表达的其他有用的启动子是在衰老的开始时打开的那些启动子,例如来自拟南芥的 SAG 启动子(Gan (甘)等人,(1995)《科学》(Science) 270:1986-1988)。

[0134] 在一些情况下,可以使用可诱导型启动子。可诱导型启动子的实例包括但不局限于,四环素阻抑物系统启动子、Lac 抑制物系统启动子、可诱导的铜系统启动子、可诱导的水杨酸盐系统启动子(例如 PR1a 系统)、可诱导的糖皮质激素系统启动子(Aoyama (青山) T. 等人(1997),《植物杂志》11:605-612)和可诱导的蜕皮激素系统启动子。其他可诱导型的启动子包括可诱导的 ABA 启动子和可诱导的膨胀启动子、生长素结合蛋白基因的启动子(Schwob (施沃布)等人(1993),《植物杂志》4:423-432)、UDP 葡萄糖类黄酮糖基转移酶基因启动子(Ralston (罗尔斯顿)等人(1988),《遗传学》(Genetics) 119:185-197)、MPI 蛋白酶抑制剂启动子(Cordero (科尔德罗)等人(1994),《植物杂志》6:141-150)、以及甘油醛-3-磷酸脱氢酶基因启动子(Kohler (科勒)等人(1995),《植物分子生物学》29:1293-1298;Martinez (马丁内斯)等人(1989),《分子生物学杂志》208:551-565;以及 Quigley (奎格利)等人(1989),《分子进化学杂志》(J. Mol. Evol.) 29:412-421)。还包括苯磺酰胺可诱导的系统(美国专利号 5,364,780)以及醇可诱导的系统(国际专利申请公开号 W097/06269 和 W097/06268)以及谷胱甘肽 S-转移酶启动子。同样地,可以使用描述于以下文献之中的任何可诱导型启动子:Gatz (盖兹)(1996),《生物技术现行观点》(Current Opinion Biotechnol.) 7:168-172 以及 Gatz (盖兹)(1997),《植物生理学与植物分子生物学年度综述》(Annu. Rev. Plant Physiol. Plant Mol. Biol.) 48:89-108。

[0135] 除了以上所描述的启动子,该表达盒还可以包括其他调节序列。如在此使用的,“调节序列”意为位于一种编码序列的上游(5' 非编码序列)、内部、或下游(3' 非编码序列)的核苷酸序列,并且这些核苷酸序列影响了相关编码序列的转录、RNA 加工或稳定性、或翻译。调节顺序包括但不局限于,增强子、内含子、翻译前导序列以及多腺苷酸化信号序列。

[0136] 该表达盒还可以包括在植物中起作用的一种转录和/或翻译终止区(即,终止区)。对于在表达盒中使用而言,多种转录终止子是可利用的并且负责在超出该转基因时的转录终止以及正确的 mRNA 聚腺苷酸化。该终止区可以是与该转录引发区天然在一起的、可以是与感兴趣的可操作地连接的核苷酸序列天然在一起的、可以是与该宿主植物天然在一起的、或者可以是源自另一种来源(即,对于该启动子、该感兴趣的核苷酸序列、该宿主植物、或它们的任何组合是外来的或异源的)。适当的转录终止子包括但不局限于,CAMV35S 终止子、tml 终止子、胭脂碱合酶终止子以及豌豆 rbcS E9 终止子。这些终止子可以在单子叶植物和双子叶植物中使用。此外,可以使用一种编码序列的天然转录终止子。

[0137] 如以上讨论,为了防止与植物中木聚糖酶相关的负表型,可以将信号序列可操作地连接到该酶,用来将该酶导入细胞区室。以这种方式,该表达盒将包括一种针对该木聚糖酶的编码序列,该序列可操作地连接至该信号序列的核酸序列上。该信号序列可以是可操作地连接在该酶的 N- 或 C- 末端上。

[0138] 不考虑一种或多种调节序列的类型,可以将它们可操作地连接至该木聚糖酶的编

码序列上。如在此使用的，“可操作地连接”是指，对一种核酸构建体（例如一种表达盒）的元件进行配置以执行它们的常规功能。因此，可操作地连接至一种编码序列上的控制序列（即，启动子）能够影响该编码序列的表达。这些控制序列不需要与该编码序列邻接，只要它们起功能指导其表达。因此，例如，介入未翻译的，已转录的，序列可以在一种启动子与一种编码序列之间存在，并且该启动子序列仍可以被认为“可操作地连接”至该编码序列。

[0139] 该表达盒还可以包括一种用于可选择性标记的核苷酸序列，它可以用于选择转化的植物宿主细胞。如在此使用，“可选择性标记”是指赋予表达该标记的宿主细胞独特表型的一个分子，并且因此允许这样转化的宿主细胞与不具有该标记的细胞区别开。这样的核苷酸序列可以编码一种可选择的亦或可筛选的标记，这取决于该标记是否给予一种可以通过化学方法而被选择的性状，例如通过使用一种选择性试剂（例如，一种抗生素、除草剂、或类似物），或者取决于该标记是否仅仅是一种可以通过观察或测试而鉴别的性状，例如通过筛选（例如，R 基因性状）。当然，适合的可选择性标记的许多实例在本领域中是已知的并且可以用于在此描述的这些表达盒中。

[0140] 可选择性标记的实例包括但不限于，一个 neo 或 nptII 基因，它赋予对卡那霉素、G418、以及类似物的抗性（Potrykus（珀特里库斯）等人（1985），《分子和普通遗传学》199:183-188）；一个 bar 基因，它赋予对草丁膦的抗性；一个改变的 5-烯醇丙酮莽草酸-3-磷酸合酶（EPSP）基因，它赋予对草甘膦的抗性（Hinchee（亨基）等人（1988），《生物技术》（Biotech.）6:915-922）；一种胍水解酶基因（例如来自臭鼻克雷白氏杆菌的 bxn），它赋予对溴草膦的抗性（Stalker（斯托克）等人（1988），《科学》242:419-423）；一种改变的乙酰乳酸合酶（ALS）基因，它赋予对咪唑啉酮、磺酰脲或其他 ALS-抑制化学品的抗性（欧洲专利申请号 154204）；一种甲氨蝶呤-抗性的二氢叶酸还原酶（DHFR）基因（Thillet（赛利特）等人（1988），《生物化学杂志》263:12500-12508）；一种茅草枯脱卤素酶基因，它赋予对茅草枯的抗性；一种甘露糖-6-磷酸异构酶（也称为磷酸甘露糖异构酶（PMI））基因，它赋予代谢甘露糖的能力（美国专利号 5,767,378 与 5,994,629）；一种改变的邻氨基苯甲酸盐合酶基因，它赋予对 5-甲基色氨酸的抗性；或一种 hph 基因，它赋予对潮霉素的抗性。本领域一个普通技术人员能够选择一种适合于表达盒的可选择性标记。

[0141] 额外的可选择性标记包括但不限于，编码一种酶的一个 β -葡萄糖醛酸酶或 uidA 基因（GUS），它的不同显色底物是已知的；编码一种产物的一个 R-基因座基因，它调节植物组织中的花色苷色素（红色）生产（Dellaporta（德拉普塔）等人，“通过用 Ac 转座子标记的玉米 R-nj 的等位基因分子克隆（Molecular cloning of the maize R-nj allele by transposon-tagging with Ac）”263-282 在《染色体结构和功能：新概念的影响（Chromosome Structure and Function: Impact of New Concepts）》中，第 18 届斯塔德勒遗传学研讨会（18th Stadler Genetics Symposium）（Gustafson（古斯塔夫森）& Appels（阿佩尔斯）编辑，Plenum Press 出版社 1988））；编码一种酶的一个 β -内酰胺酶基因，它的不同显色底物是已知的（例如 PADAC，一种显色头孢菌素）（Sutcliffe（苏利夫）（1978）《美国科学院院刊》75:3737-3741）；编码一种儿茶酚加双氧酶的一个 xylE 基因（Zukowsky（祖科斯基）等人（1983）《美国科学院院刊》80:1101-1105）；编码一种酶（该酶能够将酪氨酸氧化为 DOPA 和多巴醌）的一个酪氨酸酶基因，多巴醌转而缩合以形成黑色素（Katz（卡茨）等人（1983）《普通微生物学杂志》129:2703-2714）；编码一种酶的一个 β -半乳糖苷酶基因，

对于它存在显色底物；允许生物发光检测的一个荧光素酶(lux)基因(Ow (欧)等人(1986)《科学》234:856-859);可以用于钙敏感生物发光检测的一个水母素基因(Prasher (普鲁切)等人(1985)《生物化学与生物物理研究通讯》126:1259-1268);或一种绿色荧光蛋白基因(Niedz (涅茨)等人(1995)《植物细胞报告》(Plant Cell Reports)14:403-406)。

[0142] 该表达盒还可以包括针对其他木质纤维素酶的编码序列。此类序列可以与任何的核苷酸序列组合叠加,以创造出具有所希望表型的植物、植物部分或植物宿主细胞。叠加的组合可以通过任何方法来产生,这些方法包括但不局限于,通过任何常规的方法学或通过遗传转化的杂交育种植物。如果是通过遗传转化这些植物来进行叠加的,所感兴趣的核苷酸序列可以在任何时间并且以任何次序进行组合。例如,包括一种或多种所希望的性状的转基因植物可以用作通过后续转化而引入另外的性状的靶标。这些核苷酸序列可以在一种共转化方案中与由表达盒的任何组合提供的针对该木聚糖酶的编码序列同时引入。例如,如果将引入两段核苷酸序列,则它们可以包含在分开的盒(反式)中或包含在相同的盒(顺式)中。这些核苷酸序列的表达可以通过相同的启动子或通过不同的启动子来驱动。进一步认识到的是,核苷酸序列可以在一个所希望的基因组位置处使用位点特异性重组系统进行叠加。参见,例如,国际专利申请公开号 W099/25821 ;W099/25854 ;W099/25840 ;W099/25855 以及 W099/25853。

[0143] 该表达盒还可以包括有关农学性状的一种或多种多肽的编码序列,这些主要地是对种子公司、种植者、或谷粒处理机有益的,例如细菌性病原体抗性、真菌抗性、除草剂抗性、昆虫抗性、线虫抗性以及病毒抗性。参见,例如,美国专利号 5,304,730 ;5,495,071 ;5,569,823 ;6,329,504 以及 6,337,431。该性状还可以是提高植物活力或产量(包括允许植物在不同的温度、土壤条件以及日光和沉淀水平下生长的性状)的性状、或是允许对展现感兴趣性状的植物鉴定的性状(例如,可选择性标记基因、种皮颜色等)。不同的感兴趣的性状,连同用于引这些性状进入植物的方法描述于例如,美国专利号 4,761,373 ;4,769,061 ;4,810,648 ;4,940,835 ;4,975,374 ;5,013,659 ;5,162,602 ;5,276,268 ;5,304,730 ;5,495,071 ;5,554,798 ;5,561,236 ;5,569,823 ;5,767,366 ;5,879,903,5,928,937 ;6,084,155 ;6,329,504 与 6,337,431 ;以及美国专利申请公开号 2001/0016956。还参见,万维网(World Wide Web)上的 lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/。

[0144] 众多核苷酸序列已知于增强来自转录单位之内的基因表达,并且这些序列可以与针对该木聚糖酶的编码序列结合使用以增加其在转基因植物中的表达。例如,已经显示玉米 Adh1 基因的内含子与内含子 1 增强了基因表达。参见,例如,Callis (卡利斯)等人(1987),《基因与发育》(Genes Develop.)1:1183-1200。

[0145] 同样地,还已知一些从病毒得到的未翻译的前导序列增强了基因表达。确切地,已经显示来自烟草花叶病毒(TMV,“W-序列”)、玉米枯黄斑点病毒(MCMV)以及苜蓿花叶病毒(AMV)的前导序列有效增强了表达(Gallie(加利)等人(1987)《核酸研究》15:8693-8711;以及 Skuzeski (斯库泽斯基)等人(1990)《植物分子生物学》15:65-79)。本领域已知的其他前导序列类包括但不局限于:小 RNA 病毒病毒前导序列(leader),例如脑心肌炎(EMCV)5' 非编码区前导序列(Elroy (埃尔罗伊)-Stein (斯坦因)等人(1989),《美国科学院院刊》86:6126-6130;马铃薯 Y 病毒属前导序列,例如烟草蚀纹病毒(TEV)前导序列(Allison (艾利森)等人(1986),《病毒学》(Virology)154:9-20);玉米矮花叶病毒(MDMV)前导序

列;(Allison (艾利森)等人(1986),如以上);人类免疫球蛋白重链结合蛋白(BiP)前导序列,(Macejak (麦切扎克)&Samow (萨摩)(1991),《自然》353:90-94);来自苜蓿花叶病毒(AMV)的外壳蛋白 mRNA 的非翻译前导序列(AMV RNA4;Jobling (乔布灵)&Gehrke (格尔克)(1987),《自然》325:622-625);烟草花叶病毒(TMV)前导序列,(Gallie (加利)等人(1989),《RNA 分子生物学》(Molecular Biology of RNA)237-256);以及 MCMV 前导序列(Lommel (洛梅尔)等人(1991),《病毒学》81:382-385)。参见 Della-Cioppa 等人(1987) Plant Physiol. 《植物生理学》84:965-968。

[0146] 一旦针对该木聚糖酶的编码序列被克隆进入该表达盒中,它可以被转化进入一种植物、植物部分或植物宿主细胞中。如在此使用,“植物”是指任何植物,具体地指种子植物,并且“植物宿主细胞”是该植物的一个结构单元和生理单元,该单元包括一个细胞壁并且还可以指一个原生质体。植物宿主细胞可以处于一种分离的单细胞形式,或者可以是一种培养细胞,或者可以是作为较高级的组织单位(例如,一种植物组织或一种植物器官)的一部分。如在此使用,“转化”指将一个核酸分子或片段转到一个宿主细胞的基因组中,从而导致基因上稳定遗传。包含该转化的核酸分子或片段的宿主细胞称为“转基因”细胞,并且包含转基因细胞的生物称为“转基因生物”。

[0147] 转化植物、植物部分、植物宿主细胞的方法的实例包括但不限于,农杆菌介导的转化(De Blaere (德布莱热)等人(1987)《酶学方法》(Meth. Enzymol.)153:277-293),以及粒子轰击介导的转化(Klein (克莱因)等人(1987)《自然》(Nature)327:70-73;连同美国专利号 4,945,050)。

[0148] 因此能以本领域熟知的任何数目的方式将表达盒引入植物、植物部分或植物宿主细胞。如在此使用,在上下文中,核酸构建体(例如表达盒)的“引入”是指按它获得进入植物细胞内部的这样一种方式,将如在此披露的编码木聚糖酶的核酸序列呈递给植物、植物部分或植物宿主细胞。在有待引入一种以上的核苷酸序列的情况下,它们可以作为一种单一核酸构建体的部分、或者作为分开的核酸构建体而进行组装,并且可以位于相同的或不同的核酸构建体上。

[0149] 因此,可以在单一的转化事件中、在分开的转化事件中,或例如作为育种方案的一部分在植物中,将这些核苷酸序列引入感兴趣的宿主细胞中。本发明的这些方法并不取决于用于将一种或多种核苷酸引入植物中的具体方法,仅仅取决于它们获得进入该植物至少一个细胞内部的途径。在本领域中已知的用于将核苷酸序列引入植物中的方法包括但不限于瞬时转化方法、稳定转化方法、以及病毒介导的方法。

[0150] 如在此使用,“瞬时转化”或“瞬时转化的”是指将核酸构建体(例如表达盒)引入植物中并且并不整合到植物的基因组中。

[0151] 如在此使用,“稳定转化”或“稳定转化的”是指将核酸构建体(例如表达盒)引入植物,整合到植物的基因组中,并且能够通过它的子代,更具体地,通过多个连续代的子代而遗传。

[0152] 除了表达盒,众多植物转化载体在对于本领域一个技术人员而言是熟知的,并且在此描述的这些核苷酸序列可以与任何这样的载体结合使用。载体的选择将取决于优选的转化技术以及用于转化的目标物种。对于某些目标物种,如在本文中其他地方讨论的,不同的抗生素或除草剂选择标记可以是优选的。

[0153] 例如,已经使用 Ti 质粒载体以及直接 DNA 摄入、脂质体、电穿孔、微注射、以及微弹来递送外源 DNA。此外,可以使用来自农杆菌属的细菌转化植物细胞。以下是用于转化单子叶植物(单子叶)和双子叶植物(双子叶)的代表性方法以及代表性质体转化方法的说明。

[0154] 对于使用根癌农杆菌的转化而言,很多载体是可获得的。这些载体典型地携带至少一个 T-DNA 边界序列并且包括例如 pBIN19 (Bevan (贝文) (1984),《核酸研究》12:8711-8721)的载体。美国专利申请公开号 2006/0260011 描述了用于构建在农杆菌介导的转化中有用的载体的方法。

[0155] 不使用根癌农杆菌的转化避开了在选择转化载体中对于 T-DNA 序列的要求,并且因此还可以使用缺少这些序列的载体。不依赖农杆菌的转化技术包括经由微注射、粒子轰击、以及原生质体摄入(例如 PEG 和电穿孔)的转化。如以上所指出,载体的选择很大程度上取决于对于被转化的物种的优先选择。

[0156] 转化单子叶植物的方法在本领域是熟知的,并且使用 PEG- 或电穿孔介导的方法、连同粒子轰击直接将基因转移进入愈伤组织。可以用单一 DNA 种类或多种 DNA 种类(即,共转化)进行转化,并且这两种方法都适合在此使用。共转化可能具有避免完全型载体构建以及生成对于感兴趣的基因与选择标记具有非伴性基因座的转基因植物的优点,使得能够在随后的世代中去除该选择标记(如果这被认为是所希望的话)。

[0157] 欧洲专利申请号 EP0292435 与 0392225 连同国际专利申请公开号 W093/07278 描述了用于从良种近交系玉米制备愈伤组织和原生质体、使用 PEG 或电穿孔转化原生质体、以及从转化的原生质体再生玉米植物的方法。参见,Fromm (弗洛姆)等人(1990),《生物/技术》(Bio/Technology)8:833-844;以及 Gordon (戈登)-Kamm (卡姆)等人(1990),《植物细胞》2:603-618,它们描述了用于使用粒子轰击转化源于 A188 的玉米系的方法。此外,国际专利申请公开号 W093/07278 和 Koziel (科泽尔)等人(1993)《生物技术》(Biotechnology) 11:194-200 描述了用于通过粒子轰击转化良种近交系玉米的方法。这些方法使用从授粉之后 14-15 天的玉米穗切除的 1.5mm-2.5mm 长的未成熟玉米胚以及一台用于轰击的 PDS-1000He Biolistics 装置。在每一情况下,可以使用本领域熟知的方法将这些转化的细胞再生为完整的植物。

[0158] 用于双子叶植物的转化方法在本领域也是熟知的,并且包括农杆菌介导的方法以及不要求农杆菌的方法。非农杆菌介导的方法涉及直接通过原生质体或细胞的外源基因材料的摄入。这种方法可以通过 PEG 或电穿孔介导的摄入、粒子轰击介导的递送、或微注射来实现。这些方法的实例描述于例如,由 Klein (克莱因)等人(1987),《自然》327:70-73; Paszkowski (帕克沃斯基)等人(1984),《EMBO 杂志》3:2717-2722; Potrykus (珀特里库斯)等人(1985),《分子和普通遗传学》199:169-177;以及 Reich (里奇)等人(1986),《生物技术》(Biotechnology) 4:1001-1004。在每一情况下,可以使用本领域熟知的方法将这些转化的细胞再生为完整的植物。

[0159] 对于转化双子叶植物,农杆菌介导的转化是一种优选的方法,因为它的高转化效率以及它对于许多不同物种的广泛实用性。农杆菌介导的转化典型地涉及将携带感兴趣的外源 DNA 的二元载体转移至合适的农杆菌菌株,这可以依赖于由宿主农杆菌菌株携带的、或者在共同存在的 Ti 质体上的或染色体的 vir 基因的互补(Uknes(乌肯内斯)等人(1993),《植物细胞》5:159-169)。将该重组二元载体转移至农杆菌可以使用携带该重组二元载体的

大肠杆菌、一种辅助大肠杆菌的菌株(该辅助菌株携带一种质粒,该质粒能够将该重组二元载体移动到目标农杆菌菌株中)通过一种三亲本交配程序实现。可替代地,可以通过 DNA 转化将该重组二元载体转移到农杆菌中(Höfgen & Willmitzer (威廉米泽)(1988),《核酸研究》16:9877)。

[0160] 通过重组体农杆菌转化目标植物物种通常涉及该农杆菌与来自该植物的外植体的共培养,并且遵循本领域熟知的方法。将转化的组织在存在于二元质粒 T-DNA 边界之间的携带有抗生素标记或除草剂抗性标记的可选择性培养基上再生。

[0161] 另一种用于转化植物细胞的方法涉及在植物组织和细胞上推进惰性的或生物学活性的粒子。参见,例如,美国专利号 4,945,050 ;5,036,006 以及 5,100,792。总体上,这种方法涉及在有效穿透该细胞的外表面并提供掺入其内部中的条件下在细胞上推进惰性的或生物学活性的粒子。当利用惰性粒子时,可以通过用含有所希望的编码序列的载体包被这些粒子以将该载体引入细胞中。可替代地,该目标细胞可以被该载体围绕使得该载体通过该粒子的激发而被带入该细胞中。也可以将生物学活性的粒子(例如,干酵母细胞、干细菌或噬菌体,各自包含被试图引入的 DNA)注入到植物细胞组织中。

[0162] 经由用感兴趣的编码序列的转化获得的植物可以是各种各样的植物物种中的任何一种,包括单子叶植物和双子叶植物;然而,本发明的方法中所使用的植物优先选自以上列出的农学上重要的目标作物列表。感兴趣的核苷酸序列的表达组合其他的对于产量和品质重要的特征可以通过育种结合进入植物系中。育种的途径与方法在本领域是已知的。参见,例如,Welsh (威尔士),《植物遗传与育种基础》(Fundamentals of Plant Genetics and Breeding) (约翰威利父子公司(John Wiley & Sons) 1981);《作物育种》(Crop Breeding) (Wood (伍德) 编辑,美国农学学会(American Society of Agronomy) 1983);Mayo (梅奥),《植物育种理论第二版》(The Theory of Plant Breeding, 2nd ed.) (克拉伦登出版社(Clarendon Press) 1987);Singh (辛格),《针对疾病与虫害抗性的育种》(Breeding for Resistance to Diseases and Insect Pests) (斯普林格出版社(Springer-Verlag) 1986);以及 Wricke (里克) & Weber (韦伯),《数量遗传学与选着性植物育种》(Quantitative Genetics and Selection Plant Breeding) (Walter de Gruyter and Co. 1986)。

[0163] 通过本领域熟知的方法可以从转基因细胞再生整个植物。参见,例如,Fromm (弗洛姆) 等人(1990),如以上;以及 Gordon (戈登) -Kamm (卡姆) 等人(1990),如以上。

[0164] 同样地,工程化进入以上描述的转基因种子与植物中的遗传特性可以通过有性生殖或营养生长被传递,并且因此可以在子代植物中被维持和繁殖。总体而言,维持和繁殖利用了已知的农业方法,这些方法被开发为适合特定的目的(如收获、播种或耕作)。

[0165] 这些木聚糖酶还可以通过育种或通过普通的基因工程技术结合进入或维持在植物系中。参见,同前。相关的方法包括但不限于:非整倍体技术、回交育种、双单倍体近交、杂交、近交、种间杂交、多系育种、品种共混(variety blend)、等。杂交技术还包括植物的绝育以通过生物化学、化学、遗传(包括转基因)或机械方法产生雄性或雌性不育植物。

[0166] 本发明的组合物还包括用编码木聚糖酶、其具有木聚糖酶活性的变体或片段的核酸构建体转化的非植物宿主细胞和生物。如在此使用,“宿主细胞”是指一种在其中该表达盒(包括一种包含该表达盒的载体)可以被繁殖与表达的细胞。宿主细胞还包括受试宿主细胞或其衍生物的任何子代。应理解,所有子代可以不与母细胞完全相同,因为在复制过程中

可能会发生突变;然而,此种子代是包括在“宿主细胞”内。用于在此使用的宿主细胞的实例包括但不局限于,原核细胞,例如细菌细胞,以及真核细胞,例如真菌(例如,酵母细胞),昆虫(例如,果蝇细胞)以及哺乳动物细胞。

[0167] 简而言之,在非植物宿主细胞或非植物生物中表达核苷酸序列的方法包括,用例如一种核酸构建体转化该宿主细胞或生物,该核酸构建体具有一种针对该木聚糖酶、或其变体或片段的可操作地连接至一种或多种调节序列上的编码序列。针对木聚糖酶的编码序列可以是 SEQ ID 号 :5、7 或 9 中列出的核苷酸序列,或其片段或变体,或编码 SEQ ID 号 :6、8 或 10 或其片段或变体的多肽序列。转化的宿主细胞在该核苷酸序列得以表达的条件下生长。

[0168] 用于构建表达盒的方法在以上进行了描述。一旦构建,可以将该表达盒结合进入一种载体以辅助该表达盒引入该宿主细胞中。

[0169] 将核酸构建体(例如表达盒与载体)引入宿主细胞中的方法在本领域是熟知的并且取决于所选择的宿主细胞而变化。参见,例如, Davis (戴维斯) 等人,《分子生物学基本方法》(Basic Methods in Molecular Biology) 爱思唯尔出版公司 1986 (Elsevier Press 1986)。用于将核酸构建体(例如表达盒或载体)引入宿主细胞中的方法的实例包括但不局限于,农杆菌介导的转化、磷酸钙介导的转化、环糊精介导的转化、电穿孔、脂质体介导的转化、纳米颗粒介导的转化、粒子轰击、聚合物介导的转化、以及病毒介导的转导。参见,例如, Bacchetti (巴切蒂)&Graham (格拉汉姆) (1977),《美国科学院院刊》74:1590-1594; Bertram (伯特伦) (2006),《当前药学生物技术》(Current Pharm. Biotechnol.) 7:277-228; Fischer (费舍尔) 等人(2002),《生物偶联物化学》(Bioconjugate Chem.) 13:1124-1133; Graham (格拉汉姆)&van der Eb (范德伊布) (1973),《病毒学》52:456-467; Menuel (曼纽尔) 等人(2008),《生物偶联物化学》19:2357-2362; Soltani (索尔塔尼) 等人,《农杆菌属:从生物学到生物技术》649-675 中的“非植物生物的农杆菌介导的转化”(“Agrobacterium-mediated transformation of non-plant organisms”649-675 In: Agrobacterium: From Biology to Biotechnology) (Tzfira (茨菲拉)&Citovsky (契托夫斯基) 编辑,斯普林格出版社(Springer) 2008); 以及 Tsukakoshi (冢越) 等人(1984),《应用物理学 B- 光物理学与激光化学》(Appl. Physics B-Photophys. Laser Chem.) 35:135-140。

[0170] 如以上,宿主细胞可以被瞬时转化或稳定转化。

[0171] 非植物宿主细胞的实例包括但不局限于,以上讨论的产乙醇生物,连同真菌,例如酵母,尤其是短梗霉属、隐球酵母属、克鲁维酵母属、赤酵母属、酵母菌属以及掷孢酵母属。

[0172] 一旦该木聚糖酶、其活性变体或片段通过该宿主细胞或生物得以表达,它可以从该宿主细胞或生物中纯化以用于在此描述的这些组合物中。

[0173] 该木聚糖酶可以被分离并且配制为用于木质纤维素生物加工的组合物。以此方式,可以在一个高 pH,典型地在 pH8.5 或更高,包括 pH10、pH11、pH12 和更高,优选 pH12 配制这些组合物。其他组分可以包括在该组合物内,这些组分包括但不局限于,其他的生物加工酶类或化学药品连同稳定剂。

[0174] 方法

[0175] 本发明的方法包括将木质纤维素材料与一种如在此描述的酶组合物接触,以水解其中的纤维素、半纤维素以及木素聚合物为用于后续在不同饲料、食品 / 饮料与工业产业

中使用的单体。特别是,这些方法在高温、高 pH 应用中找到了用途。

[0176] 简而言之,将木质纤维素材料生物转化为有用的、更高价值的产物通常要求多步骤加工,这些步骤包括:(1) 预处理(例如生物的、化学的或机械的)木质纤维素材料,以及(2) 通过使它们与木聚糖酶、其变体或片段接触,来水解纤维素、半纤维素和 / 或木素聚合物以产生更简单的、仍易于可代谢的分子(例如己糖或戊糖)。可任选地进行额外的步骤,例如可能在饲料、食品 / 饮料与工业行业中要求的步骤。这类额外的步骤可以包括:(3) 这些分子的生物利用以支持微生物生长或以生产化学产物,以及(4) 这些化学产物的分离与纯化。

[0177] 木质纤维素可以在例如非木本植物的茎、叶、皮、壳以及穗轴,或者木本植物(例如树木)的叶、枝与木中找到。木质纤维素植物材料还可以从木质纤维素废弃物,例如植物残体与废纸来获得。植物残体的实例包括但不限于,茎、叶、皮、壳、穗轴等等,以及木、木屑、木浆与锯屑。废纸的实例包括但不限于,丢弃的复印纸、计算机打印纸、笔记本纸、记事本纸、打字机纸等等,以及丢弃的报纸、杂志、纸板以及纸质包装材料。木质纤维素植物材料还可以来自农业废弃物、森林枯落物(*forestry residue*)、草质材料、城市固体废物、以及浆料与纸厂残余物。

[0178] 这些方法可以包括预处理木质纤维素材料的一个步骤。尽管木质纤维素材料可以“按原样”使用,但是还可以用本领域熟知的方法预处理,这些方法包括生物预处理、化学预处理或物理预处理。生物预处理的实例包括但不限于,使用木素增溶微生物(例如梭菌属或链霉菌属)。化学预处理的实例包括但不限于碱处理、氨处理、稀酸处理、有机溶剂处理、臭氧处理、二氧化硫处理、二氧化碳处理和 pH 控制的水热处理。物理预处理的实例包括但不限于不同类型的水热处理、辐射、研磨、气烘 / 蒸汽喷发和声处理(超声处理)。大体参见, Hsu (许),《生物乙醇手册:生产与利用》179-212 中的生物质预处理(“*Pretreatment of biomass*”179-212 In: *Handbook on Bioethanol: Production and Utilization*) (Wyman (怀曼) 编辑,泰勒弗朗西斯出版社(Taylor&Francis) 1996); Ghosh (高希) & Singh (辛格) (1993),《应用微生物学进展》(*Adv. Appl. Microbiol.*) 39:295-333; McMillan (麦克米伦),《用于燃料生产的酶转化》第 15 章中的木质纤维素生物质预处理:综述(“*Pretreating lignocellulosic biomass: a review*” Chapter 15 In: *Enzymatic Conversion of Biomass for Fuels Production*) (Himmel (西摩尔) 等人编辑,美国化学学会(American Chemical Society) 1994); Gong (龚) 等人(1999),《生物化学工程 / 生物技术进展》(*Adv. Biochem. Eng. / Biotechnol.*) 65:207-241; Olsson (奥尔森) & Hahn (哈恩) -Hagerdal (哈格代尔) (1996),《酶微生物学技术》(*Enz. Microb. Tech.*) 18:312-331; 以及 Vallander (瓦伦德尔) & Eriksson (埃里克森) (1990),《生物化学工程 / 生物技术进展》42:63-95; 以及美国专利号 5, 733, 741 与 6, 333, 181。

[0179] 这些方法包括使具有纤维素、半纤维素和 / 或木素的木质纤维素材料与木聚糖酶组合物接触,来将该材料水解成更简单的、更易于可利用的分子。如在此使用,“接触”是指使酶、其变体或片段与木质纤维素材料在一起,由此促进酶、其变体或片段水解木质纤维素材料或对该材料起作用所必需的分子间相互作用。取决于该酶、和 / 或其变体或片段的来源,接触可以发生在体外(*in vitro*)、活体外(*ex vivo*)或活体内(*in vivo*)。

[0180] 总体上,合适的酶剂量是约 0.01 单位至 1000 单位每克的干木质纤维素材料,并且

更优选 0.1 单位至 500 单位每克。该酶组合物的活性可以如以下来确定：向 0.5ml 木聚糖溶液(1%；西格玛，圣路易斯，密苏里州(Sigma, St. Louis, MO)；在 50mM 磷酸盐缓冲液中制备，pH7)添加 0.5ml 适当稀释的处于相同缓冲液中的酶。将该溶液在 70℃ 孵育约 10 分钟。然后将该反应通过添加 1ml DNS 试剂(3, 5-二硝基水杨酸盐 10g/L；Na/K 酒石酸盐 300g/L；NaOH16g/L)来终止，并且可以通过煮沸该样品 5 分钟显出颜色。参见，例如，Ghose (高斯)(1987)，如以上；以及 Miller (米勒)(1959)，如以上。然后可以在 540nm 的波长处测量吸收性。一酶单位在测定条件下释放一 μmol 作为每分钟木糖计算的还原糖。该活性可以从在测定条件下释放 4 μmol 的还原糖的酶稀释液来计算。

[0181] 如果该酶组合物的活性是不足的，则可以将它在使用前浓缩。浓缩多肽与蛋白例如酶类的方法在本领域是熟知的。参见，例如，Ahmed (艾哈迈德)，《蛋白提取、纯化与鉴定的原理与方法》(Principles and Reactions of Protein Extraction, Purification, and Characterization) (CRC 出版社 2004)；Dennison (丹尼森)，《蛋白分离指南》(A Guide to Protein Isolation) 第二版，(克鲁乌学士出版社 2003 (Kluwer Academic Publishers 2003))；Degerli (德葛利) & Akpınar (阿克帕纳尔) (2001)，《分析生物化学》297:192-194；以及《蛋白方法》(Protein Methods) 第二版，(Bollag (伯莱格) 等人编辑，威利出版社(Wiley Publishers)1996)。例如，简单超滤法可以用于浓缩该酶的上清液 5-25 倍。

[0182] 接触时间可以并且将取决于以下因素而变化，例如所希望的结果、所使用的木质纤维素材料的类型、所使用的木质纤维素材料的量、所使用的酶的量、该酶组合物的特异活性、温度、pH、等等。然而，该步骤可以进行一段时间，这段时间足以减少木质纤维素材料的纤维素、半纤维素和 / 或木素含量。如此，木质纤维素材料和酶、其变体或片段可以接触持续以下时间：约 5 分钟、约 10 分钟、约 15 分钟、约 30 分钟、约 45 分钟、约 1 小时、约 2 小时、约 3 小时、约 4 小时、约 5 小时、约 6 小时、约 7 小时、约 8 小时、约 9 小时、约 10 小时、约 12 小时、约 14 小时、约 16 小时、约 18 小时、约 20 小时、约 22 小时、约 24 小时、约 26 小时、约 28 小时、约 30 小时、约 32 小时、约 34 小时、约 36 小时、约 38 小时、约 40 小时、约 42 小时、约 44 小时、约 46 小时、约 48 小时、或更长。

[0183] 如以上所指出，优选该接触是在以下温度以及 pH 下进行，该温度与 pH 在体外增强了该酶的活性。温度的范围可以从约 10℃ 至约 100℃，其中约 65℃ 至 95℃ 是优选的。对该步骤而言优选的 pH 可以从约 pH6 至约 pH12。在此描述的酶组合物的的一个特征是，它可以在宽广的 pH 范围上具有活性，连同在 pH6-pH12 具有高活性，这是对于饲料、食品 / 饮料以及工业产业(例如，在制浆厂所用的存储塔中)中的许多加工所要求的 pH。因为该酶在宽广的 pH 范围上具有活性，该 pH 在这些方法中可以变化。如以上所指出，该酶在约 pH6 下具有木聚糖酶活性，并且在甚至高至约 pH12 时保留活性。因此，该 pH 可以在约 pH6 开始，其中存在木聚糖酶活性，并且然后稳定地变化至更碱性的 pH，例如 pH12，其中活性将继续。可替代地，pH 可以在约 pH12，并且然后稳定地变化至一个更酸性的 pH，例如 pH6。

[0184] 这些酶组合物或改性的生物可以用于不同方法中。这些木聚糖酶具有以下优点，它们可以在高温以及高 pH 下使用。在一个实施例中，这些方法可以用在饲料产业中以制造动物饲料。这些方法提供用于在动物消费之前或之中水解饲料中的木聚糖，这些方法包括以下步骤：(a) 获得包括木质纤维素材料(例如纤维素、半纤维素以及木素)的饲料材料；

(b) 将足够量的具有一种木聚糖酶(例如, SEQ ID 号 :6、8 或 10) 或其一种具有木聚糖酶活性的变体或片段的酶组合物与该饲料材料接触并且持续足够时段, 以引起木聚糖的水解, 由此在动物消费之前或之中水解饲料中的这些木聚糖。这些方法还可以包括 (c) 将该饲料材料给予该动物, 其中在消费之后, 在该动物的消化道中该木聚糖酶活性引起该饲料中的木聚糖的水解。该饲料可以是例如一种谷物、玉米、籽粒、以及类似物。

[0185] 在另一个实施例中, 这些方法可以用在食品 / 饮料产业中以加工果实或蔬菜材料。这些方法提供了在人类消费之前水解食品或饮料中的木聚糖, 这些方法包括以下步骤: (a) 提供包含木质纤维素材料(例如纤维素、半纤维素以及木素)的水果或蔬菜材料, 例如水果或蔬菜材料; 并且 (b) 将一种足够量的具有一种木聚糖酶(例如, SEQ ID 号 :6、8 或 10) 或其一种具有木聚糖酶活性的变体或片段的酶组合物与该水果或蔬菜材料接触并且持续足够时段, 以引起木聚糖的水解, 由此加工该水果或蔬菜材料。该食品可以是例如一种果实或蔬菜汁、啤酒、葡萄酒、糖浆、酱或提取物以及类似物。

[0186] 在另一个实施例中, 这些方法可以用于工业产业中以处理浆料。这些方法提供了水解非木本或木本植物材料中的木聚糖, 这些方法包括以下步骤: (a) 提供包括纤维素、半纤维素以及木素的非木本或木本植物材料, 例如作物、植物或树木生物质; 并且 (b) 将一种足够量的具有一种木聚糖酶(例如, SEQ ID 号 :6、8 或 10) 或其一种具有木聚糖酶活性的变体或片段的酶组合物与该非木本或木本植物材料接触并且持续足够时段, 以引起木聚糖的水解, 由此加工该非木本或木本植物材料。这些方法还可以在步骤 (b) 之前包括, 当漂白浆状材料时, 将化学浆料暴露于一种二氧化氯漂白阶段以产生一种部分漂白的浆料; (b) 将该部分漂白的浆料与包括该酶、其具有木聚糖酶活性的变体或片段的该酶组合物相接触; 并且 (c) 将该经处理的浆料暴露于一种第二二氧化氯漂白中, 而没有碱性提取阶段。

[0187] 在另一个实施例中, 这些方法可以用在工业产业中以生产乙醇。这些方法提供了水解非木本或木本植物材料中的木聚糖, 这些方法包括以下步骤: (a) 提供包括纤维素、半纤维素以及木素的非木本或木本植物材料, 例如作物、植物或树木生物质; (b) 将一种足够量的具有一种木聚糖酶(例如, SEQ ID 号 :6、8 或 10) 或其一种具有木聚糖酶活性的变体或片段的酶组合物与该非木本或木本植物材料接触并且持续足够时段, 以引起木聚糖的水解, 由此加工该非木本或木本植物材料用于随后的乙醇生产。

[0188] 在另一个实施例中, 这些方法可以用以制造生物燃料。这些方法提供了水解非木本或木本植物材料中的木聚糖, 这些方法包括以下步骤: (a) 提供包括纤维素、半纤维素以及木素的非木本或木本植物材料, 例如作物、植物或树木生物质; (b) 将一种足够量的具有一种木聚糖酶(例如, SEQ ID 号 :6、8 或 10) 或其一种具有木聚糖酶活性的变体或片段的酶组合物与该非木本或木本植物材料接触并且持续足够时段, 以引起木聚糖的水解, 由此加工该非木本或木本植物材料用于随后的生物燃料生产。

[0189] 实验

[0190] 以下实例是出于说明, 而非限制的目的而提供的。

[0191] 实例 1 : 具有热稳定性的并且嗜碱的木聚糖酶活性的木聚糖酶。

[0192] 该实例示出木聚糖酶可以具有热稳定性的并且嗜碱的木聚糖酶活性。

[0193] 可以使用大肠杆菌作为宿主生产木聚糖酶多肽。使用放置在一个诱导型 T7lac 启动子下游的框内核苷酸序列的限制性酶位点, 将编码木聚糖酶的选择的核苷酸序列克隆到

pET24a 大肠杆菌表达载体中。这类核苷酸序列包括嗜热脂肪芽孢杆菌木聚糖酶(BsXynA ; SEQ ID 号 :9)和海栖热孢菌木聚糖酶(TmXynB ;SEQ ID 号 :7)。通过使用对于大肠杆菌优选的密码子回译该酶的多肽序列生成编码特定酶的核苷酸序列。将表达载体引入一种大肠杆菌表达株 BL21Star (DE3 ;Invitrogen 公司)中。

[0194] 在含有 50ug/mL 卡那霉素的标准 LB 琼脂上对含有修饰的 pET24a 表达载体的重组大肠杆菌分离株进行选择。在 37°C 下在摇动下在 20mL 含有 50ug/mL 卡那霉素的 LB 培养基中培养分离株 8 小时至过夜。

[0195] 将 20ml 大肠杆菌培养物转移到含有 25ug/ml 卡那霉素的 1L 自诱导培养基中 (9.57g 胰蛋白胨、4.8g 酵母提取物、2ml1M MgSO₄、1ml1000X 痕量金属、20ml50X5052、以及 20ml50X M) (1000X 痕量金属 :36ml 无菌水、50ml 在 0.12M HCl 中的 0.1M FeCl₃、2ml1M CaCl₂、1ml1M MnCl₂·4H₂O、1ml1M ZnSO₄·7H₂O、1ml0.2M CoCl₂·6H₂O、2ml0.1M CuCl₂·2H₂O、1ml0.2M NiCl₂·6H₂O、2ml0.1M Na₂MoO₄·2H₂O、以及 2ml0.1M H₃BO₃) (50X5052 :25g 甘油、73mlH₂O、2.5g 葡萄糖、以及 10g α - 乳糖一水合物) (50X M:80ml H₂O、17.75gNa₂HPO₄、17.0g KH₂PO₄、13.4g NH₄Cl、和 3.55g Na₂SO₄) ,并且在 28°C 摇荡生长过夜。

[0196] 通过在 10,000X g 下离心 15 分钟从该自诱导培养基收获这些大肠杆菌细胞,并且将这些收集的细胞在 -80°C 冷冻。通过将细胞聚合体重新悬浮于一种含有 DNase (1U/ml 缓冲液)的缓冲液 (50mM Tris、pH7.5、50mM NaCl、0.1mM EDTA)中以及使用弗氏压碎器来溶化细胞。通过在 4°C 12,000X g 离心 20 分钟来除去不可溶的碎片。

[0197] 对于 TmXynB (SEQ ID 号 :7 和 8),将含有总可溶蛋白以及重组酶的上清液转移至新的 40ml 离心管并且在 75°C 孵育 30 分钟。这些木聚糖酶仍然可溶并且通过 12,000X g 离心 30 分钟与沉淀的蛋白分开,这些收集的蛋白可用于另外的表征。

[0198] 对于 BsXynA (SEQ ID 号 :9 和 10),在弗氏压碎器溶化后,将上清液倾倒至一个阴离子亲和柱上,使用 ÄKTA™快速蛋白液相色谱法 (FPLC ;通用电器生命科学公司 (GE Healthcare Life Sciences),乌普萨拉(Uppsala),瑞典)并且使用 50mM Tris,pH8.0,具有从 0 至 2M 的 NaCl 梯度,洗脱为多个部分。

[0199] 使用 SDS-PAGE 与光密度测定法确定所有木聚糖酶纯度。

[0200] 通过结合如以上所述的木聚糖酶制剂与桦树底物基生物物质,来评估氨水预处理过程中的木聚糖酶活性。该预处理混合物(cocktail)包括 0.75% 桦树木聚糖以及不同浓度的氨水。每一处理的 pH 在 pH11.3 和 pH12.4 之间。

[0201] 通过一种剂量响应曲线来确定所使用的木聚糖酶制剂的量,在其中将不同的酶量添加至一种利用 0.75% 桦树木聚糖以及 50mM Tris (pH8.0) 来进行的测定中,并且孵育 1 小时。木聚糖酶的最佳量被确定为是一个不使该桦树木聚糖底物饱和的最大量。对于 BsXynA (SEQ ID 号 :10),这些量是 0.6 μ g 每 mg 的木聚糖,并且对于 TmXynB (SEQ ID 号 :8) 是 0.25 μ g 每 mg 的木聚糖。

[0202] 该预处理鸡尾酒包含 0.75% 桦树木聚糖、木聚糖酶制剂以及氨水。通过如在以下标题“木聚糖酶测定”标题下所述的 DNS 显色测定,在对 BsXynA (SEQ ID 号 :10)和 TmXynB (SEQ ID 号 :8) 测量木聚糖酶活性以前,在不同温度下孵育预处理混合物 1 小时。

[0203] 木聚糖酶测定 :将细胞在 8000x g 下离心 5 分钟。弃去该上清液,并且将细胞在 0.1M 甘氨酸缓冲液 (pH12,含有 7%NH₄OH)中重新悬浮至 OD_{600nm}=200。将 100 μ l 的重新悬浮

的细胞在 60°C 孵育 1 小时。1 小时孵育之后,将在该甘氨酸缓冲液 (pH12) 中制备的 200 μ l 的 5%w/v 燕麦木聚糖 (西格玛) 溶液作为底物添加,并且允许该反应在 60°C 继续 1 小时。1 小时的孵育之后,针对还原糖用 DNS 比色测定对酶活性进行检测,如描述于 Miller (米勒) (1959),如以上。

[0204] DNS 测定被用于测量由木聚糖酶产生的木糖多糖的还原端的产生;因此,该方法测量了木聚糖酶活性。如以上所指出,该测定基于还原端与 DNS 的反应;可以比色法使用光谱测定法测量产生的产物。参见例如 Miller (米勒) (1959),如以上。经由标准曲线将比色值翻译为还原糖浓度,其中在 A540 的吸光度被确定用于 0-10mM 木糖的范围。

[0205] 还在将酶制剂暴露于氨水中的预处理以确定预处理条件不可逆地改变酶到什么程度以后,评估木聚糖酶活性。为了确定大肠杆菌生产的木聚糖酶的稳定性,将上述木聚糖酶制剂在 60°C 下在不同浓度的氨水中孵育,随后中和并且随后进行木聚糖酶活性测定。作为额外对照,分析还可以用来自绿色木霉的 M1 木聚糖酶 (Megazyme Int'l Ireland Ltd. 公司) 来进行。如在上述那些内容中一样,在这些测定中使用相同量的木聚糖酶制剂;按 1 μ g 每 mg 的木聚糖使用 M1。

[0206] 通过将木聚糖酶在桦树木聚糖底物的存在下、在 60°C、在不同浓度的氨水中孵育进行预处理。将对照样品即刻 (AAPT0 小时) 中和,同时将测试样品在该氨水中孵育 1 小时。将这些样品通过与 DowexTMWeak Acid Resin (道化学公司;米德兰,密歇根州) 加上另外的桦树木聚糖孵育 0 小时 (ACT0 小时对照) 或 1 小时 (ACT1 小时测试样品) 来中和。通过如在以上描述的 DNS 测定来确定木聚糖酶活性。

[0207] 表 1 和表 2 示出了在氨水预处理过程中,每一酶的相对木聚糖酶活性。这些值被计算为在 60°C 和 0.01% 氨水条件下观察到的酶活性 (此酶活性被设定为 100%) 的百分比。

[0208] 表 1:在氨水预处理期间的 BsXynA (SEQ ID 号:10) 木聚糖酶活性。

[0209]

氨水的百分比	50°C	60°C	80°C
0.01	53.5	100.0	34.9
0.1	34.3	77.6	5.6
1	31.5	50.8	0.0
4	20.6	15.7	0.0
8	8.0	0.4	0.0

[0210] 表 2:在氨水预处理期间的 TmXynB (SEQ ID 号:8) 木聚糖酶活性。

[0211]

氨水的百分比	50°C	60°C	80°C
0.01	29.7	100.0	193.2
0.1	16.5	74.1	141.8
1	19.6	42.2	4.5
4	9.4	27.0	3.6
8	3.5	4.7	3.5

[0212] 因此这些木聚糖酶在氨水预处理过程中是活性的。在于氨水中孵育过程中,这些木聚糖酶能够水解该桦树木聚糖底物,从而表明了这些酶可以在纤维素生物质的预处理过程中使用,以转化该生物质为可发酵糖。在对增加浓度的氨水的耐受性和对更高温度的耐受性中,分析的不同酶示出一些差异。

[0213] 当选择一种适当的木聚糖酶用于纤维素生物质的预处理时,在提高氨水浓度的同时,这些酶对温度的灵敏度是重要考虑的。

[0214] 表 3 至表 5 示出在生物质底物存在下,在氨基预处理条件下孵育后,不同木聚糖保持木聚糖酶活性。

[0215] 表 3 :用氨水预处理(AAPT)、随后是中和以及一小时活性孵育(AI)的 M1 木聚糖酶活性(mM 木糖当量)。

氨水的百分比	AAPT: 0	AAPT: 0	AAPT: 1	AAPT: 1
	hr, AI: 0 hr	hr, AI: 1 hr	hr, AI: 0 hr	hr, AI: 1 hr
[0216] 0.01	37.8	100.0	18.2	19.3
1	72.7	129.7	10.8	9.6
4	56.1	107.1	7.3	6.2
8	51.7	97.1	5.4	5.3

[0217] 表示为百分比,其中 100%= 在 0.01%AA 的活性, AAPT :0hr, 作用 :1hr.

[0218] 表 4 :用氨水预处理(AAPT)、随后是中和以及一小时活性孵育(AI)的 BsXynA (SEQ ID 号 :10) 木聚糖酶活性(mM 木糖当量)。

[0219]

氨水的百分比	AAPT: 0	AAPT: 0	AAPT: 1	AAPT: 1
	hr, AI: 0 hr	hr, AI: 1 hr	hr, AI: 0 hr	hr, AI: 1 hr
0.01	3.6	100.0	56.7	187.1
1	3.8	82.6	20.1	65.1
4	3.1	70.4	3.6	3.7
8	2.4	75.0	2.4	4.7

[0220] 表示为百分比,其中 100%= 在 0.01%AA 的活性, AAPT :0hr, 作用 :1hr

[0221] 表 5 :用氨水预处理(AAPT)、随后是中和以及一小时活性孵育(AI)的 TmXynB (SEQ ID 号 :8) 木聚糖酶活性(mM 木糖当量)。

[0222]

氨水的百分比	AAPT: 0 hr, AI: 0 hr	AAPT: 0 hr, AI: 1 hr	AAPT: 1 hr, AI: 0 hr	AAPT: 1 hr, AI: 1 hr
0.01	4.0	100.0	31.2	136.5
1	3.4	91.6	18.3	107.7
4	2.5	84.1	8.3	68.8
8	2.8	92.3	2.7	12.3

[0223] 表示为百分比,其中 100%= 在 0.01%AA 的活性, AAPT :0hr, 作用 :1hr

[0224] 因此这些木聚糖酶可以在基于氨的预处理过程中使用并且在生物质的纤维素转变为可发酵糖加工过程中的生物质预处理之后继续起作用。对暴露于氨基预处理, M1 酶(对照)示出很少至没有耐受性,同时对暴露于氨基预处理, BsXynA (SEQ ID 号 :10) 和 TmXynB (SEQ ID 号 :8) 这二者都示出不同水平的耐受性。当考虑开发一种包括了使用能够分解生物质的酶的纤维素生物质转化过程时,这些酶可以是一种选用酶。

[0225] 实例 2 :木聚糖酶 10 蛋白和对植物叶组织的作用

[0226] 用表达来自木聚糖酶家族 11 类酶的木聚糖酶的转基因植物的以上经验导致其中表达家族 11 木聚糖酶的负农学表型的关联。表达家族 11 木聚糖酶的转基因植物(其中酶在植物的种子中表达)经常展示负种子表型,例如枯萎的种子。此外,具有半纯化木聚糖酶家族 11 酶的渗透的植物叶导致坏死叶损伤。本领域存在对木聚糖酶的需求,这些木聚糖酶可以表达在转基因植物中,而没有关联的负表型,例如坏死损伤或枯萎的种子。以下实例描述了与植物叶中的家族 10 内切 1, 4 β 木聚糖酶存在关联的植物表型。

[0227] 通过用半纯化的酶制剂渗透植物叶(烟草和玉米),测试了木聚糖酶家族 10 酶制剂引起植物坏死表型的能力。这些植物观察到坏死损伤的发展。已经指出,当该酶被渗透到叶中时,测试的家族 10 木聚糖酶中没有一个是与植物叶中的坏死损伤关联。

[0228] 将编码家族 10 内切 1, 4 β 木聚糖酶多肽的四个多核苷酸序列克隆到表达载体 pET24(Novagen 公司)中。这四个核苷酸序列是 TmXynA(基因库 ID 号 :AAD35155 ;SEQ ID 号 :5), TmXynB (基因库 ID 号 :AAD35164 ;SEQ ID 号 :7), CtXynZ 全酶(基因库 ID 号 :ABN53181 ;SEQ ID 号 :1), 以及含有阿魏酸酯酶(ferrulic acid esterase)活性的 CtXynZ 的一个片段(SEQ ID 号 :3)。在每一个重组 pET24 表达载体中,在 His 标签(它是 pET24 表达载体的一个组分)之前引入一个终止密码子。该终止密码子确保重组 pET24 表达载体产生的多肽不包含一个 His 标签。

[0229] 使用标准分子生物学技术,将重组 pET24 表达载体(包含编码木聚糖酶 10 蛋白的多核苷酸序列)转化到大肠杆菌宿主 BL21[DE3] 中。为了从重组大肠杆菌宿主纯化蛋白,在 50mL 的自诱导培养基(9.57g 胰蛋白胨、4.8g 酵母提取物、2mL1M MgSO₄、1mL1000X 痕量金属、20mL50X5052、20mL50X M) (1000X 痕量金属 :36mL 无菌水、50mL0.1M FeCl₃ 在 0.12M HCl 中、2mL1M CaCl₂、1mL1M MnCl₂4H₂O、1mL1M ZnSO₄7H₂O、1mL0.2M CoCl₂6H₂O、2mL0.1M

CuCl₂·2H₂O)、1mL 0.2M NiCl₂·6H₂O、2mL 0.1M Na₂MoO₄·2H₂O、2mL 0.1M H₃BO₃) (50X5052 :25g 甘油、73mL H₂O、2.5g 葡萄糖、10g α-乳糖一水合物) (50X M :80mL H₂O、17.75g Na₂HPO₄、17.0g KH₂PO₄、13.4g NH₄Cl、3.55g Na₂SO₄) 中生长,该培养基中还含有 25ug/mL 卡那霉素作为用于重组 pET24 质粒的选择剂。通过离心从这 50mL 的培养物中回收细菌细胞,并且然后重新悬浮在无蛋白酶抑制剂的缓冲液(50mM Tris pH7.5、50mM NaCl、0.1mM EDTA)中。通过使该溶液穿过弗氏压碎器来溶化重悬浮的细胞,并且通过离心回收细胞碎片(弃去团块,保留上清液)。通过在 70℃ 孵育 30 分钟进一步热处理上清液,并且通过离心除去生成的变性蛋白(弃去团块,保留上清液)。进一步浓缩上清液并且通过穿过一个交联葡萄糖 G-25 中型基质柱(GE Healthcare 公司,catalog#17-0851-01)脱盐,并且洗脱结合的蛋白到缓冲液 10mM 磷酸钾缓冲液 pH5.8 中。

[0230] 使用牛 γ 球蛋白(Pierce 公司)以产生标准曲线的 Bradford 测定,来量化上述制剂中的蛋白。通过进行其中考马斯亮蓝染色蛋白的 SDS-PAGE,随后是使用具有 Quantity One 软件包,版本 4.6.3 (Quantity One1-D 分析软件,Catalog#:170-9600,版本 4.6.5, Build094) 的 BioRad 图像仪的光密度测定法,来确定上述木聚糖酶制剂的纯度 CtXynZ 蛋白(基因库 ID 号 :ABN53181 ;SEQ ID 号 :2)为 41.2% 纯, TmXynA (基因库 ID 号 :AAD35155 ;SEQ ID 号 :6)为 78% 纯并且 TmXynB (基因库 ID 号 :AAD35164 ;SEQ ID 号 :8)为 96% 纯。此外,通过基本如以下所述的蜡质活性测定证明木聚糖酶活性 ;然而用于该测定的标准为 0、20mM、40mM、60mM、80mM 和 100mM 木糖,并且该测定允许进行 4 小时。基于蜡质活性测定(除了仅有负对照的 CtXynA FAE 组分(SEQ ID 号 :4)),所有酶制剂包含木聚糖酶活性。

[0231] 将木聚糖酶制剂渗透到来自烟草或玉米的叶中,以确定木聚糖酶 10 酶是否会引引起植物组织中的坏死表型。将来自含有一个来自 TEV 的突变的 P1/HC-Pro 基因(该基因抑制转录后基因沉默(Mallory (马洛里)等人,《自然生物技术》20:622 (2002)))的转基因 TEV-B 烟草植物(在烟草栽培品种 Xanthi 中制造)的烟草叶用于瞬时表达选择的酶。使用一个 5mL 注射器通过将该注射器的尖端(无针头)压到叶子的背轴面表面上,在玉米(约 3 周龄)和 TEV-B 烟草植物(约 4 周龄)上进行单独的叶的渗透。用两种不同剂量的木聚糖酶制剂(150ug/mL 亦或 25ug/mL)渗透植物 ;在 1mg/ml 或 0.1mg/ml 渗透小麦阿拉伯木聚糖对照,同时在 10mM 或 1mM 渗透消化的小麦阿拉伯木聚糖。将已渗透的植物保持在 22℃ -25℃ 下其中光周期为 16 小时光照和 8 小时黑暗。浸润后 7 天后收获植物组织用于随后的分析。

[0232] 分析用木聚糖酶 10 渗透的来自烟草植物和玉米植物这二者的叶的视觉症状,例如通过叶组织的枯萎和褐变所证明的坏死损伤。木聚糖酶 10 蛋白渗透的组织中没有发展坏死或损伤 ;制剂包括 CtXynA 全酶(SEQ ID 号 :2), CtXynA FAE 组分(SEQ ID 号 :4), TmXynA (SEQ ID 号 :6), 以及 TmXynB (SEQ ID 号 :8)。作为阴性对照,底物小麦阿拉伯木聚糖被渗透到烟草和玉米叶中,基本上如以上对渗透木聚糖 10 蛋白的描述。还将消化的小麦阿拉伯木聚糖进行渗透以确定木聚糖酶分解产物是否诱导植物表型。小麦阿拉伯木聚糖并不诱导植物表型 ;然而,消化的小麦阿拉伯木聚糖制剂在烟草和玉米叶这二者中都不诱导坏死损伤。作为纯制剂,有可能被用于消化小麦阿拉伯木聚糖的 M1 木聚糖酶并未从分解产物适当地滤掉,并且负责引起渗透烟草和玉米叶中的坏死。当渗透到烟草或玉米叶中时, M1 木聚糖酶已经诱导了损伤。

[0233] 此外,使用蜡质活性测定对来自用木聚糖酶 10 酶制剂渗透的植物叶的提取物进

行测定,以确定渗透的木聚糖酶制剂是否保持了木聚糖酶活性。所有渗透到植物叶中的木聚糖酶制剂包含木聚糖酶活性,除了只有阴性对照 CtXynA FAE 组分(SEQ ID 号 :4),它并不具有木聚糖酶活性。

[0234] 蜡质活性测定 :

[0235] 与 50uL 的小麦阿拉伯木聚糖(如 Megazyme 所述制备)一起孵育 10uL 的酶制剂。通过结合 50uL 的阿拉伯木聚糖与 10uL 的木糖标准产生标准。不搅拌,将这些制剂在 40°C 孵育 3 小时。通过添加 50uL 的 DNS 量化这些反应(溶解 5.0g 的 3,5-二硝基水杨酸和 150g 酒石酸钠钾四水化合物于 900mL 的 0.4M 氢氧化钠中。转移至 1L 体积的烧瓶中并且用 0.4M 氢氧化钠调整体积至 1L)并且在 95°C 孵育 10 分钟。确定生成的制剂的在 540nm 的吸光度。木糖标准被用于建立标准曲线,并且基于标准曲线计算来自木聚糖酶制剂的吸光度读数。

[0236] 本说明书中提到的所有公开文献以及专利申请都指示了本发明涉及的本领域内的那些技术人员的水平。所有公开物和专利申请均通过引用结合在此,其程度如同每个单独的公开物或专利申请被确切地并单独地指明通过引用而被结合。

[0237] 尽管已经为了清楚理解的目的通过解释和实例详细地描述了以上发明,显而易见的是在所附权利要求的范围内可以实施某些改变和变更。

[0001]

序列表

<110> 先正达公司
 <120> 在纤维素生物质的预处理期间有活性的木聚糖酶
 <130> 不可用
 <160> 12
 <170> 3.5版本专利
 <210> 1
 <211> 2427
 <212> DNA
 <213> 热纤维梭菌
 <220>
 <221> CDS
 <222> (1)..(2427)
 <223> 木聚糖酶 Z, 来自热纤维梭菌 - 全酶
 <400> 1
 atg agc ctg ccg aca atg cct ccg agc ggt tat gat cag gtt cgt aat 48
 Met Ser Leu Pro Thr Met Pro Pro Ser Gly Tyr Asp Gln Val Arg Asn
 1 5 10 15
 ggt gtt ccg cgt ggt cag gtt gtt aat att age tat ttt agc acc gcc 96
 Gly Val Pro Arg Gly Gln Val Val Asn Ile Ser Tyr Phe Ser Thr Ala
 20 25 30
 acc aat agc acc cgt ccg gca cgt gtt tat ctg cct ccg ggt tat agc 144
 Thr Asn Ser Thr Arg Pro Ala Arg Val Tyr Leu Pro Pro Gly Tyr Ser
 35 40 45
 aaa gat aaa aaa tat agc gtg ctg tat ctg ctg cat ggt att ggt ggt 192
 Lys Asp Lys Lys Tyr Ser Val Leu Tyr Leu Leu His Gly Ile Gly Gly
 50 55 60
 agc gaa aat gat tgg ttt gaa ggt ggt ggt cgt gca aat gtt att gcc 240
 Ser Glu Asn Asp Trp Phe Glu Gly Gly Gly Arg Ala Asn Val Ile Ala
 65 70 75 80
 gat aat ctg att gcc gaa ggc aaa atc aaa ccg ctg att att gtt acc 288
 Asp Asn Leu Ile Ala Glu Gly Lys Ile Lys Pro Leu Ile Ile Val Thr
 85 90 95
 ccg aat acc aat gca gcc ggt ccg ggt att gca gat ggc tat gaa aat 336
 Pro Asn Thr Asn Ala Ala Gly Pro Gly Ile Ala Asp Gly Tyr Glu Asn
 100 105 110
 ttt acc aaa gat ctg ctg aat agc ctg att ccg tat att gaa agc aat 384
 Phe Thr Lys Asp Leu Leu Asn Ser Leu Ile Pro Tyr Ile Glu Ser Asn
 115 120 125
 tat agc gtg tat acc gat cgc gaa cat cgt gca att gcc ggt ctg agc 432
 Tyr Ser Val Tyr Thr Asp Arg Glu His Arg Ala Ile Ala Gly Leu Ser
 130 135 140
 atg ggt ggt ggt cag agc ttt aat att ggc ctg acc aat ctg gat aaa 480
 Met Gly Gly Gly Gln Ser Phe Asn Ile Gly Leu Thr Asn Leu Asp Lys
 145 150 155 160
 ttt gcc tat atc ggt ccg att agc gca gca ccg aat acc tat ccg aat 528
 Phe Ala Tyr Ile Gly Pro Ile Ser Ala Ala Pro Asn Thr Tyr Pro Asn
 165 170 175
 gaa cgt ctg ttt ccg gat ggt ggt aaa gca gca cgt gaa aaa ctg aaa 576
 Glu Arg Leu Phe Pro Asp Gly Gly Lys Ala Ala Arg Glu Lys Leu Lys
 180 185 190
 ctg ctg ttt att gca tgt ggc acc aat gat agc ctg att ggt ttt ggt 624
 Leu Leu Phe Ile Ala Cys Gly Thr Asn Asp Ser Leu Ile Gly Phe Gly
 195 200 205

[0002]

cag cgt gtg cac gaa tat tgc gtg gcc aat aat att aat cat gtg tat Gln Arg Val His Glu Tyr Cys Val Ala Asn Asn Ile Asn His Val Tyr 210 215 220	672
tgg ctg att cag ggt ggt ggc cat gat ttt aat gtt tgg aaa ccg ggt Trp Leu Ile Gln Gly Gly Gly His Asp Phe Asn Val Trp Lys Pro Gly 225 230 235 240	720
ctg tgg aat ttt ctg cag atg gca gaf gaa gcc ggt ctg acc cgt gat Leu Trp Asn Phe Leu Gln Met Ala Asp Glu Ala Gly Leu Thr Arg Asp 245 250 255	768
ggc aat aca ccg gtt ccg acc ccg tca ccg aaa ccg gca aat acc cgt Gly Asn Thr Pro Val Pro Thr Pro Ser Pro Lys Pro Ala Asn Thr Arg 260 265 270	816
att gaa gcc gaa gat tat gat ggc att aat agc agc agc att gaa att Ile Glu Ala Glu Asp Tyr Asp Gly Ile Asn Ser Ser Ser Ile Glu Ile 275 280 285	864
att ggt gtt ccg cct gaa ggt ggt cgt ggt att ggt tat att acc agc Ile Gly Val Pro Pro Glu Gly Gly Arg Gly Ile Gly Tyr Ile Thr Ser 290 295 300	912
ggc gat tat ctg gtg tat aaa agc att gat ttt ggc aat ggt gcg acc Gly Asp Tyr Leu Val Tyr Lys Ser Ile Asp Phe Ile Asn Gly Ala Thr 305 310 315 320	960
agc ttt aaa gca aaa gtg gcc aat gca aat acc agc aat att gaa ctg Ser Phe Lys Ala Lys Val Ala Asn Ala Asn Thr Ser Asn Ile Glu Leu 325 330 335	1008
cgt ctg aat ggt ccg aat ggc acc ctg att ggc acc ctg agc gtt aaa Arg Leu Asn Gly Pro Asn Gly Thr Leu Ile Gly Thr Leu Ser Val Lys 340 345 350	1056
agc acc ggt gat tgg aat acc tat gaa gaa cag acc tgt agc att agc Ser Thr Gly Asp Trp Asn Thr Tyr Glu Glu Gln Thr Cys Ser Ile Ser 355 360 365	1104
aaa gtg acc ggc att aat gat ctg tat ctg gtg ttt aaa ggt ccg gtg Lys Val Thr Gly Ile Asn Asp Leu Tyr Leu Val Phe Lys Gly Pro Val 370 375 380	1152
aat att gat tgg ttt acc ttt ggt gtt gaa agc agc agc acc ggt ctg Asn Ile Asp Trp Phe Thr Phe Gly Val Glu Ser Ser Ser Thr Gly Leu 385 390 395 400	1200
ggc gat ctg aat ggt gat ggc aat att aat agc agt gat ctg cag gca Gly Asp Leu Asn Gly Asp Gly Asn Ile Asn Ser Ser Asp Leu Gln Ala 405 410 415	1248
ctg aaa cgt cat ctg ctg ggt att tct ccg ctg acc ggt gaa gca ctg Leu Lys Arg His Leu Leu Gly Ile Ser Pro Leu Thr Gly Glu Ala Leu 420 425 430	1296
ctg cgt gca gat gtt aat cgt agc ggt aaa gtt gat agc acc gat tat Leu Arg Ala Asp Val Asn Arg Ser Gly Lys Val Asp Ser Thr Asp Tyr 435 440 445	1344
tct gtg ctg aaa cgc tat att ctg cgc att att acc gaa ttt ccg ggt Ser Val Leu Lys Arg Tyr Ile Leu Arg Ile Ile Thr Glu Phe Pro Gly 450 455 460	1392
cag ggt gat gtt cag acc ccg aat ccg agc gtt acc ccg aca cag aca Gln Gly Asp Val Gln Thr Pro Asn Pro Ser Val Thr Pro Thr Gln Thr 465 470 475 480	1440
ccg att ccg acc att agc ggt aat gca ctg cgt gat tat gcc gaa gca Pro Ile Pro Thr Ile Ser Gly Asn Ala Leu Arg Asp Tyr Ala Glu Ala 485 490 495	1488
cgt ggt att aaa att ggc acc tgt gtg aat tat ccg ttt tat aat aat Arg Gly Ile Lys Ile Gly Thr Cys Val Asn Tyr Pro Phe Tyr Asn Asn 500 505 510	1536

[0003]

agc gat ccg acc tat aat agc att ctg cag cgc gaa ttt agc atg gtt Ser Asp Pro Thr Tyr Asn Ser Ile Leu Gln Arg Glu Phe Ser Met Val 515 520 525	1584
gtg tgc gaa aat gaa atg aaa ttt gat gea ctg caa ccg cgt cag aat Val Cys Glu Asn Glu Met Lys Phe Asp Ala Leu Gln Pro Arg Gln Asn 530 535 540	1632
gtt ttt gat ttt agc aaa ggc gat cag ctg ctg gca ttt gca gaa cgt Val Phe Asp Phe Ser Lys Gly Asp Gln Leu Leu Ala Phe Ala Glu Arg 545 550 555 560	1680
aat ggt atg cag atg cgt ggt cat acc ctg att tgg cat aat cag aat Asn Gly Met Gln Met Arg Gly His Thr Leu Ile Trp His Asn Gln Asn 565 570 575	1728
ccg agc tgg ctg acc aat ggt aat tgg aat cgt gat agc ctg ctg gca Pro Ser Trp Leu Thr Asn Gly Asn Trp Asn Arg Asp Ser Leu Leu Ala 580 585 590	1776
gtg atg aaa aat cat att acc acc gtg atg acc cat tat aaa ggc aaa Val Met Lys Asn His Ile Thr Thr Val Met Thr His Tyr Lys Gly Lys 595 600 605	1824
att gtg gaa tgg gat gtt gcc aat gaa tgt atg gat gat agc ggt aat Ile Val Glu Trp Asp Val Ala Asn Glu Cys Met Asp Asp Ser Gly Asn 610 615 620	1872
ggt ctg cgt agc agc att tgg cgt aat gtt att ggc cag gat tat ctg Gly Leu Arg Ser Ser Ile Trp Arg Asn Val Ile Gly Gln Asp Tyr Leu 625 630 635 640	1920
gat tat gcc ttt cgt tat gca cgt gaa gca gat ccg gat gca ctg ctg Asp Tyr Ala Phe Arg Tyr Ala Arg Glu Ala Asp Pro Asp Ala Leu Leu 645 650 655	1968
ttt tat aat gat tat aat att gaa gat ctg ggt ccg aaa agc aat gcc Phe Tyr Asn Asp Tyr Asn Ile Glu Asp Leu Gly Pro Lys Ser Asn Ala 660 665 670	2016
gtg ttt aat atg att aaa agc atg aaa gaa cgt ggt gtt ccg att gat Val Phe Asn Met Ile Lys Ser Met Lys Glu Arg Gly Val Pro Ile Asp 675 680 685	2064
ggt gtt ggt ttt cag tgc cat ttt att aat ggc atg tct ccg gaa tat Gly Val Gly Phe Gln Cys His Phe Ile Asn Gly Met Ser Pro Glu Tyr 690 695 700	2112
ctg gca agc att gat cag aat atc aaa cgc tat gcc gaa att ggt gtg Leu Ala Ser Ile Asp Gln Asn Ile Lys Arg Tyr Ala Glu Ile Gly Val 705 710 715 720	2160
att gtg agc ttt acc gaa att gat att cgt att ccg cag agc gaa aat Ile Val Ser Phe Thr Glu Ile Asp Ile Arg Ile Pro Gln Ser Glu Asn 725 730 735	2208
ccg gca acc gca ttt cag gtt cag gcc aat aat tat aaa gaa ctg atg Pro Ala Thr Ala Phe Gln Val Gln Ala Asn Asn Tyr Lys Glu Leu Met 740 745 750	2256
aaa att tgt ctg gcc aat ccg aat tgt aat acc ttt gtg atg tgg ggc Lys Ile Cys Leu Ala Asn Pro Asn Cys Asn Thr Phe Val Met Trp Gly 755 760 765	2304
ttt acc gat aaa tat acc tgg att ccg ggt aca ttt ccg ggt tat ggt Phe Thr Asp Lys Tyr Thr Trp Ile Pro Gly Thr Phe Pro Gly Tyr Gly 770 775 780	2352
aat ccg ctg att tat gat agc aat tat aat ccg aaa ccg get tat aat Asn Pro Leu Ile Tyr Asp Ser Asn Tyr Asn Pro Lys Pro Ala Tyr Asn 785 790 795 800	2400
gcc att aaa gaa gcc ctg atg ggc tat Ala Ile Lys Glu Ala Leu Met Gly Tyr 805	2427

[0004]

<210> 2
 <211> 809
 <212> PRT
 <213> 热纤维梭菌
 <400> 2
 Met Ser Leu Pro Thr Met Pro Pro Ser Gly Tyr Asp Gln Val Arg Asn
 1 5 10 15
 Gly Val Pro Arg Gly Gln Val Val Asn Ile Ser Tyr Phe Ser Thr Ala
 20 25 30
 Thr Asn Ser Thr Arg Pro Ala Arg Val Tyr Leu Pro Pro Gly Tyr Ser
 35 40 45
 Lys Asp Lys Lys Tyr Ser Val Leu Tyr Leu Leu His Gly Ile Gly Gly
 50 55 60
 Ser Glu Asn Asp Trp Phe Glu Gly Gly Gly Arg Ala Asn Val Ile Ala
 65 70 75 80
 Asp Asn Leu Ile Ala Glu Gly Lys Ile Lys Pro Leu Ile Ile Val Thr
 85 90 95
 Pro Asn Thr Asn Ala Ala Gly Pro Gly Ile Ala Asp Gly Tyr Glu Asn
 100 105 110
 Phe Thr Lys Asp Leu Leu Asn Ser Leu Ile Pro Tyr Ile Glu Ser Asn
 115 120 125
 Tyr Ser Val Tyr Thr Asp Arg Glu His Arg Ala Ile Ala Gly Leu Ser
 130 135 140
 Met Gly Gly Gly Gln Ser Phe Asn Ile Gly Leu Thr Asn Leu Asp Lys
 145 150 155 160
 Phe Ala Tyr Ile Gly Pro Ile Ser Ala Ala Pro Asn Thr Tyr Pro Asn
 165 170 175
 Glu Arg Leu Phe Pro Asp Gly Gly Lys Ala Ala Arg Glu Lys Leu Lys
 180 185 190
 Leu Leu Phe Ile Ala Cys Gly Thr Asn Asp Ser Leu Ile Gly Phe Gly
 195 200 205
 Gln Arg Val His Glu Tyr Cys Val Ala Asn Asn Ile Asn His Val Tyr
 210 215 220
 Trp Leu Ile Gln Gly Gly Gly His Asp Phe Asn Val Trp Lys Pro Gly
 225 230 235 240
 Leu Trp Asn Phe Leu Gln Met Ala Asp Glu Ala Gly Leu Thr Arg Asp
 245 250 255
 Gly Asn Thr Pro Val Pro Thr Pro Ser Pro Lys Pro Ala Asn Thr Arg
 260 265 270

[0005]

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

[0006]

Pro Ser Trp Leu Thr Asn Gly Asn Trp Asn Arg Asp Ser Leu Leu Ala
 580 585 590

Val Met Lys Asn His Ile Thr Thr Val Met Thr His Tyr Lys Gly Lys
 595 600 605

Ile Val Glu Trp Asp Val Ala Asn Glu Cys Met Asp Asp Ser Gly Asn
 610 615 620

Gly Leu Arg Ser Ser Ile Trp Arg Asn Val Ile Gly Gln Asp Tyr Leu
 625 630 635 640

Asp Tyr Ala Phe Arg Tyr Ala Arg Glu Ala Asp Pro Asp Ala Leu Leu
 645 650 655

Phe Tyr Asn Asp Tyr Asn Ile Glu Asp Leu Gly Pro Lys Ser Asn Ala
 660 665 670

Val Phe Asn Met Ile Lys Ser Met Lys Glu Arg Gly Val Pro Ile Asp
 675 680 685

Gly Val Gly Phe Gln Cys His Phe Ile Asn Gly Met Ser Pro Glu Tyr
 690 695 700

Leu Ala Ser Ile Asp Gln Asn Ile Lys Arg Tyr Ala Glu Ile Gly Val
 705 710 715 720

Ile Val Ser Phe Thr Glu Ile Asp Ile Arg Ile Pro Gln Ser Glu Asn
 725 730 735

Pro Ala Thr Ala Phe Gln Val Gln Ala Asn Asn Tyr Lys Glu Leu Met
 740 745 750

Lys Ile Cys Leu Ala Asn Pro Asn Cys Asn Thr Phe Val Met Trp Gly
 755 760 765

Phe Thr Asp Lys Tyr Thr Trp Ile Pro Gly Thr Phe Pro Gly Tyr Gly
 770 775 780

Asn Pro Leu Ile Tyr Asp Ser Asn Tyr Asn Pro Lys Pro Ala Tyr Asn
 785 790 795 800

Ala Ile Lys Glu Ala Leu Met Gly Tyr
 805

<210> 3
 <211> 1185
 <212> DNA
 <213> 热纤维梭菌

<220>
 <221> CDS
 <222> (1)..(1185)
 <223> CxYnz的阿魏酸酯酶 (ferrulic acid esterase)

<400> 3
 atg gcc gcc tcc etc ccg acc atg ccg ccg tcc ggc tac gac cag gtg
 Met Ala Ala Ser Leu Pro Thr Met Pro Pro Ser Gly Tyr Asp Gln Val
 1 5 10 15

48

[0007]

序 列 表

egc aac ggc gtg ccg cgc ggc cag gtg gtg aac atc tcc tac ttc tcc Arg Asn Gly Val Pro Arg Gly Gln Val Val Asn Ile Ser Tyr Phe Ser 20 25 30	96
acc gcc acc aac tcc acc cgc ccg gcc cgc gtg tac ctc ccg ccg ggc Thr Ala Thr Asn Ser Thr Arg Pro Ala Arg Val Tyr Leu Pro Pro Gly 35 40 45	144
tac tcc aag gac aag aag tac tcc gtg ctc tac ctc ctc cac ggc atc Tyr Ser Lys Asp Lys Lys Tyr Ser Val Leu Tyr Leu Leu His Gly Ile 50 55 60	192
ggc ggc tcc gag aac gac tgg ttc gag gcc ggc ggc cgc gcc aac gtg Gly Gly Ser Glu Asn Asp Trp Phe Glu Gly Gly Gly Arg Ala Asn Val 65 70 75 80	240
atc gcc gac aac ctc atc gcc gag gcc aag atc aag ccg ctc atc atc Ile Ala Asp Asn Ile Ile Ala Glu Gly Lys Ile Lys Pro Leu Ile Ile 85 90 95	288
gtg acc ccg aac acc aac gcc gcc gcc ccg gcc atc gcc gac gcc tac Val Thr Pro Asn Thr Asn Ala Ala Gly Pro Gly Ile Ala Asp Gly Tyr 100 105 110	336
gag aac ttc acc aag gac ctc ctc aac tcc ctc atc ccg tac atc gag Glu Asn Phe Thr Lys Asp Leu Leu Asn Ser Leu Ile Pro Tyr Ile Glu 115 120 125	384
tcc aac tac tcc gtg tac acc gac cgc gag cac cgc gcc atc gcc gcc Ser Asn Tyr Ser Val Tyr Thr Asp Arg Glu His Arg Ala Ile Ala Gly 130 135 140	432
ctc tct atg ggc ggc ggc cag tcc ttc aac atc ggc ctc acc aac ctc Leu Ser Met Gly Gly Gly Gln Ser Phe Asn Ile Gly Leu Thr Asn Leu 145 150 155 160	480
gac aag ttc gcc tac atc gcc ccg atc tcc gcc gcc ccg aac acc tac Asp Lys Phe Ala Tyr Ile Gly Pro Ile Ser Ala Ala Pro Asn Thr Tyr 165 170 175	528
ccg aac gag cgc ctc ttc ccg gac gcc gcc aag gcc gcc cgc gag aag Pro Asn Glu Arg Leu Phe Pro Asp Gly Gly Lys Ala Ala Arg Glu Lys 180 185 190	576
ctc aag ctc ctc ttc atc gcc tgc gcc acc aac gac tcc ctc atc gcc Leu Lys Leu Leu Phe Ile Ala Cys Gly Thr Asn Asp Ser Leu Ile Gly 195 200 205	624
ttc gcc cag cgc gtg cac gag tac tgc gtg gcc aac aac atc aac cac Phe Gly Gln Arg Val His Glu Tyr Cys Val Ala Asn Asn Ile Asn His 210 215 220	672
gtg tac tgg ctc atc cag gcc gcc gcc cae gac ttc aac gtg tgg aag Val Tyr Trp Leu Ile Gln Gly Gly Gly His Asp Phe Asn Val Trp Lys 225 230 235 240	720
ccg gcc ctc tgg aac ttc ctc cag atg gcc gac gag gcc gcc ctc acc Pro Gly Leu Trp Asn Phe Leu Gln Met Ala Asp Glu Ala Gly Leu Thr 245 250 255	768
egc gac gcc aac acc ccg gtg ccg acc ccg tcc ccg aag ccg gcc aac Arg Asp Gly Asn Thr Pro Val Pro Thr Pro Ser Pro Lys Pro Ala Asn 260 265 270	816
acc cgc atc gag gcc gag gac tac gac gcc atc aac tcc tcc tcc atc Thr Arg Ile Glu Ala Glu Asp Tyr Asp Gly Ile Asn Ser Ser Ser Ile 275 280 285	864
gag atc atc gcc gtg ccg ccg gag gcc gcc cgc gcc atc gcc tac atc Glu Ile Ile Gly Val Pro Pro Glu Gly Gly Arg Gly Ile Gly Tyr Ile 290 295 300	912
acc tcc gcc gac tac ctc gtg tac aag tcc atc gac ttc gcc aac gcc Thr Ser Gly Asp Tyr Leu Val Tyr Lys Ser Ile Asp Phe Gly Asn Gly 305 310 315 320	960

[0008]

```

gcc acc tcc ttc aag gcc aag gtg gcc aac gcc aac acc tcc aac atc      1008
Ala Thr Ser Phe Lys Ala Lys Val Ala Asn Ala Asn Thr Ser Asn Ile
                325                      330                      335

gag ctt cgc ctc aac ggc ceg aac ggc acc ctc atc ggc acc ctc tcc      1056
Glu Leu Arg Leu Asn Gly Pro Asn Gly Thr Leu Ile Gly Thr Leu Ser
                340                      345                      350

gtg aag tcc acc ggc gac tgg aac acc tac gag gag cag acc tgc tcc      1104
Val Lys Ser Thr Gly Asp Trp Asn Thr Tyr Glu Glu Gln Thr Cys Ser
                355                      360                      365

atc tcc aag gtg acc ggc atc aac gac ctc tac ctc gtg ttc aag ggc      1152
Ile Ser Lys Val Thr Gly Ile Asn Asp Leu Tyr Leu Val Phe Lys Gly
                370                      375                      380

ccg gtg aac atc gac tgg ttc acc ttc ggc gtg                          1185
Pro Val Asn Ile Asp Trp Phe Thr Phe Gly Val
385                      390                      395

<210> 4
<211> 395
<212> PRT
<213> 热纤维梭菌

<400> 4

Met Ala Ala Ser Leu Pro Thr Met Pro Pro Ser Gly Tyr Asp Gln Val
1 5 10 15

Arg Asn Gly Val Pro Arg Gly Gln Val Val Asn Ile Ser Tyr Phe Ser
20 25 30

Thr Ala Thr Asn Ser Thr Arg Pro Ala Arg Val Tyr Leu Pro Pro Gly
35 40 45

Tyr Ser Lys Asp Lys Lys Tyr Ser Val Leu Tyr Leu Leu His Gly Ile
50 55 60

Gly Gly Ser Glu Asn Asp Trp Phe Glu Gly Gly Gly Arg Ala Asn Val
65 70 75 80

Ile Ala Asp Asn Leu Ile Ala Glu Gly Lys Ile Lys Pro Leu Ile Ile
85 90 95

Val Thr Pro Asn Thr Asn Ala Ala Gly Pro Gly Ile Ala Asp Gly Tyr
100 105 110

Glu Asn Phe Thr Lys Asp Leu Leu Asn Ser Leu Ile Pro Tyr Ile Glu
115 120 125

Ser Asn Tyr Ser Val Tyr Thr Asp Arg Glu His Arg Ala Ile Ala Gly
130 135 140

Leu Ser Met Gly Gly Gly Gln Ser Phe Asn Ile Gly Leu Thr Asn Leu
145 150 155 160

Asp Lys Phe Ala Tyr Ile Gly Pro Ile Ser Ala Ala Pro Asn Thr Tyr
165 170 175

Pro Asn Glu Arg Leu Phe Pro Asp Gly Gly Lys Ala Ala Arg Glu Lys
180 185 190
    
```

[0009]

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

Phe Gly Gln Arg Val His Glu Tyr Cys Val Ala Asn Asn Ile Asn His
 210 215 220

Val Tyr Trp Leu Ile Gln Gly Gly Gly His Asp Phe Asn Val Trp Lys
 225 230 235 240

Pro Gly Leu Trp Asn Phe Leu Gln Met Ala Asp Glu Ala Gly Leu Thr
 245 250 255

Arg Asp Gly Asn Thr Pro Val Pro Thr Pro Ser Pro Lys Pro Ala Asn
 260 265 270

Thr Arg Ile Glu Ala Glu Asp Tyr Asp Gly Ile Asn Ser Ser Ser Ile
 275 280 285

Glu Ile Ile Gly Val Pro Pro Glu Gly Gly Arg Gly Ile Gly Tyr Ile
 290 295 300

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

Ala Thr Ser Phe Lys Ala Lys Val Ala Asn Ala Asn Thr Ser Asn Ile
 325 330 335

Glu Leu Arg Leu Asn Gly Pro Asn Gly Thr Leu Ile Gly Thr Leu Ser
 340 345 350

Val Lys Ser Thr Gly Asp Trp Asn Thr Tyr Glu Glu Gln Thr Cys Ser
 355 360 365

Ile Ser Lys Val Thr Gly Ile Asn Asp Leu Tyr Leu Val Phe Lys Gly
 370 375 380

Pro Val Asn Ile Asp Trp Phe Thr Phe Gly Val
 385 390 395

<210> 5
 <211> 3090
 <212> DNA
 <213> 海栖热孢菌

<220>
 <221> CDS
 <222> (1)..(3090)
 <223> XynA, 来自海栖热孢菌

<400> 5
 atg gca agc ggt gtt ctg agc ttt ggt aaa gaa gca agc agc aaa ggc 48
 Met Ala Ser Gly Val Leu Ser Phe Gly Lys Glu Ala Ser Ser Lys Gly
 1 5 10 15

gat agc agc ctg gaa ace gtt ctg gca ctg agc ttt gaa ggc acc acc 96
 Asp Ser Ser Leu Glu Thr Val Leu Ala Leu Ser Phe Glu Gly Thr Thr
 20 25 30

gaa ggt gtt gtt cgg ttt ggt aaa gaf gtt gtt ctg acc gca agc cag 144
 Glu Gly Val Val Pro Phe Gly Lys Asp Val Val Leu Thr Ala Ser Gln
 35 40 45

[0010]

gat gtt gca gca gat ggt gaa tat agc ctg aaa gtg gaa aat cgt acc Asp Val Ala Ala Asp Gly Glu Tyr Ser Leu Lys Val Glu Asn Arg Thr 50 55 60	192
agc ccg tgg gat ggt gtt gaa att gat ctg acc ggc aaa gtt aaa agc Ser Pro Trp Asp Gly Val Glu Ile Asp Leu Thr Gly Lys Val Lys Ser 65 70 75 80	240
ggt gca gat tat ctg ctg tcc ttt cag gtt tat cag agc agt gat gca Gly Ala Asp Tyr Leu Leu Ser Phe Gln Val Tyr Gln Ser Ser Asp Ala 85 90 95	288
ccg cag ctg ttt aat gtt gtt gca cgc acc gaa gat gaa aaa ggt gaa Pro Gln Leu Phe Asn Val Val Ala Arg Thr Glu Asp Glu Lys Gly Glu 100 105 110	336
cgc tat gat gtg atc ctg gat aaa gtt gtt gtt agc gac cac tgg aaa Arg Tyr Asp Val Ile Leu Asp Lys Val Val Val Ser Asp His Trp Lys 115 120 125	384
gaa att ctg gtt ccg ttt agc ccg acc ttt gaa ggt aca ccg gca aaa Glu Ile Leu Val Pro Phe Ser Pro Thr Phe Glu Gly Thr Pro Ala Lys 130 135 140	432
tac agc ctg att atc gtg gca agc aaa aac acc aac ttt aat ttt tat Tyr Ser Leu Ile Ile Val Ala Ser Lys Asn Thr Asn Phe Asn Phe Tyr 145 150 155 160	480
ctg gat aaa gtc cag gtt ctg gca ccg aaa gaa tct ggt ccg aaa gtg Leu Asp Lys Val Gln Val Leu Ala Pro Lys Glu Ser Gly Pro Lys Val 165 170 175	528
att tat gaa acc agc ttt gaa aat ggt gtt ggt gat tgg cag cct cgt Ile Tyr Glu Thr Ser Phe Glu Asn Gly Val Gly Asp Trp Gln Pro Arg 180 185 190	576
ggt gat gtt aat att gaa gcc agc agc gaa gtt gca cat agc ggt aaa Gly Asp Val Asn Ile Glu Ala Ser Ser Glu Val Ala His Ser Gly Lys 195 200 205	624
agc agc ctg ttt att agc aat cgt cag aaa ggt tgg cag ggt gca cag Ser Ser Leu Phe Ile Ser Asn Arg Gln Lys Gly Trp Gln Gly Ala Gln 210 215 220	672
atc aat ctg aaa ggc att ctg aaa acc ggt aaa acc tat gca ttt gag Ile Asn Leu Lys Gly Ile Leu Lys Thr Gly Lys Thr Tyr Ala Phe Glu 225 230 235 240	720
gca tgg gtt tat cag aat agc ggt cag gat cag acc att att atg acc Ala Trp Val Tyr Gln Asn Ser Gly Gln Asp Gln Thr Ile Ile Met Thr 245 250 255	768
atg cag cgc aaa tat agc tct gat gca agc acc cag tat gaa tgg att Met Gln Arg Lys Tyr Ser Ser Asp Ala Ser Thr Gln Tyr Glu Trp Ile 260 265 270	816
aaa agc gca acc gtt ccg agc ggt cag tgg gtt cag ctg tct ggc acc Lys Ser Ala Thr Val Pro Ser Gly Gln Trp Val Gln Leu Ser Gly Thr 275 280 285	864
tat acc att ccg gca ggc gtt acc gtt gaa gat ctg acc ctg tat ttc Tyr Thr Ile Pro Ala Gly Val Thr Val Glu Asp Leu Thr Leu Tyr Phe 290 295 300	912
gaa agc cag aat ccg acc ctg gaa ttt tat gtg gac gac gtg aaa att Glu Ser Gln Asn Pro Thr Leu Glu Phe Tyr Val Asp Asp Val Lys Ile 305 310 315 320	960
gtt gat acc acc tct gcc gaa atc aaa att gag atg gaa ccg gaa aaa Val Asp Thr Thr Ser Ala Glu Ile Lys Ile Glu Met Glu Pro Glu Lys 325 330 335	1008
gaa att ccg gca ctg aaa gaa gtg ctg aaa gat tac ttt aaa gtt ggt Glu Ile Pro Ala Leu Lys Glu Val Leu Lys Asp Tyr Phe Lys Val Gly 340 345 350	1056

[0011]

gtt gca ctg ccg agc aaa gtt ttt ctg aac ccg aaa gat att gag ctg Val Ala Leu Pro Ser Lys Val Phe Leu Asn Pro Lys Asp Ile Glu Leu 355 360 365	1104
att acg aaa cac ttt aac agc atc acc gca gaa aat gaa atg aaa ccg Ile Thr Lys His Phe Asn Ser Ile Thr Ala Glu Asn Glu Met Lys Pro 370 375 380	1152
gaa agc ctg ctg gca ggc att gaa aaf ggc aaa ctg aaa ttc cgt ttc Glu Ser Leu Leu Ala Gly Ile Glu Asn Gly Lys Leu Lys Phe Arg Phe 385 390 395 400	1200
gaa acc gca gat aaa tat atc cag ttt gtg gaa gaa aat ggc atg gtt Glu Thr Ala Asp Lys Tyr Ile Gln Phe Val Glu Glu Asn Gly Met Val 405 410 415	1248
att cgt ggt cat acc ctg gtt tgg cat aat cag aca ccg gat tgg ttt Ile Arg Gly His Thr Leu Val Trp His Asn Gln Thr Pro Asp Trp Phe 420 425 430	1296
ttc aaa gat gag aac ggc aat ctg ctg tct aaa gaa gca atg acc gaa Phe Lys Asp Glu Asn Gly Asn Leu Leu Ser Lys Glu Ala Met Thr Glu 435 440 445	1344
cgc ctg aaa gaa tat atc cat acc gtg gtg ggc cat ttt aaa ggt aaa Arg Leu Lys Glu Tyr Ile His Thr Val Val Gly His Phe Lys Gly Lys 450 455 460	1392
gtg tat gcc tgg gat gtt gtt aat gaa gcc gtt gat ccg aat cag ccg Val Tyr Ala Trp Asp Val Val Asn Glu Ala Val Asp Pro Asn Gln Pro 465 470 475 480	1440
gat ggt ctg cgt cgt agc acc tgg tat cag att atg ggt ccg gat tat Asp Gly Leu Arg Arg Ser Thr Trp Tyr Gln Ile Met Gly Pro Asp Tyr 485 490 495	1488
att gaa ctg gcc ttt aaa ttt gca cgt gaa gca gat ccg gat gcc aaa Ile Glu Leu Ala Phe Lys Phe Ala Arg Glu Ala Asp Pro Asp Ala Lys 500 505 510	1536
ctg ttt tac aac gat tat aac acc ttt gaa cct ccg aaa ccg gat att Leu Phe Tyr Asn Asp Tyr Asn Thr Phe Glu Pro Arg Lys Arg Asp Ile 515 520 525	1584
atc tat aat ctg gtg aaa gat ctg aaa gaa aaa ggc ctg att gat ggt Ile Tyr Asn Leu Val Lys Asp Leu Lys Glu Lys Gly Leu Ile Asp Gly 530 535 540	1632
att ggt atg cag tgt cat att agc ctg gcc acc gat att aaa caa atc Ile Gly Met Gln Cys His Ile Ser Leu Ala Thr Asp Ile Lys Gln Ile 545 550 555 560	1680
gaa gaa gcc att aaa aaa ttc agc acc att ccg ggt att gaa atc cat Glu Glu Ala Ile Lys Lys Phe Ser Thr Ile Pro Gly Ile Glu Ile His 565 570 575	1728
atc acc gaa ctg gat atg agc gtt tat cgt gat agc agc agc aat tat Ile Thr Glu Leu Asp Met Ser Val Tyr Arg Asp Ser Ser Ser Asn Tyr 580 585 590	1776
ccg gaa gca ccg cgt acc gca ctg att gaa cag gca cac aaa atg atg Pro Glu Ala Pro Arg Thr Ala Leu Ile Glu Gln Ala His Lys Met Met 595 600 605	1824
cag ctg ttt gaa att ttt aag aag tat agc aac gtc att acg aac gtt Gln Leu Phe Glu Ile Phe Lys Lys Tyr Ser Asn Val Ile Thr Asn Val 610 615 620	1872
acc ttt tgg ggt ctg aaa gat gac tat agc tgg cgt gca acc cgt cgt Thr Phe Trp Gly Leu Lys Asp Asp Tyr Ser Trp Arg Ala Thr Arg Arg 625 630 635 640	1920
aat gat tgg cct ctg ate ttc gat aaa gat cat cag gca aaa ctg gca Asn Asp Trp Pro Leu Ile Phe Asp Lys Asp His Gln Ala Lys Leu Ala 645 650 655	1968

[0012]

tat tgg gca att gtt gca ccg gaa gtt ctg cct ccg ctg ccg aaa gaa Tyr Trp Ala Ile Val Ala Pro Glu Val Leu Pro Pro Leu Pro Lys Glu 660 665 670	2016
tct cgt att agc gaa ggt gaa gca gtt gtt gtt ggc atg atg gat gat Ser Arg Ile Ser Glu Gly Glu Ala Val Val Val Gly Met Met Asp Asp 675 680 685	2064
agc tac ctg atg tct aaa ccg att gaa atc ctg gat gaa gag ggt aat Ser Tyr Leu Met Ser Lys Pro Ile Glu Ile Leu Asp Glu Glu Gly Asn 690 695 700	2112
ggt aaa gca acc att cgt gcc gtt tgg aaa gat age acc atc tat atc Val Lys Ala Thr Ile Arg Ala Val Trp Lys Asp Ser Thr Ile Tyr Ile 705 710 715 720	2160
tat ggt gag gtg cag gac aaa acc aaa aaa ccg gca gaa gat ggc gtt Tyr Gly Glu Val Gln Asp Lys Thr Lys Lys Pro Ala Glu Asp Gly Val 725 730 735	2208
gcc att ttt att aac ccg aac aat gaa cgt acc ccg tat ctg cag ccg Ala Ile Phe Ile Asn Pro Asn Asn Glu Arg Thr Pro Tyr Leu Gln Pro 740 745 750	2256
gat gat acc tat gca gtt ctg tgg acc aat tgg aaa acc gaa gtg aat Asp Asp Thr Tyr Ala Val Leu Trp Thr Asn Trp Lys Thr Glu Val Asn 755 760 765	2304
cgt gaa gat gtt cag gtg aaa aaa ttt gtg ggt ccg ggt ttt cgt cgt Arg Glu Asp Val Gln Val Lys Lys Phe Val Gly Pro Gly Phe Arg Arg 770 775 780	2352
tat agc ttc gaa atg agc att aca att ccg ggt gtg gag ttc aaa aaa Tyr Ser Phe Glu Met Ser Ile Thr Ile Pro Gly Val Glu Phe Lys Lys 785 790 795 800	2400
gat age tac atc ggt ttt gat gca gcc gtt atc gat gat ggt aaa tgg Asp Ser Tyr Ile Gly Phe Asp Ala Ala Val Ile Asp Asp Gly Lys Trp 805 810 815	2448
tat age tgg tcc gat acc acc aat age cag aaa acc aac acc atg aat Tyr Ser Trp Ser Asp Thr Thr Asn Ser Gln Lys Thr Asn Thr Met Asn 820 825 830	2496
tat ggc acc ctg aaa ctg gaa ggt att atg gtt gca acc gcc aaa tat Tyr Gly Thr Leu Lys Leu Glu Gly Ile Met Val Ala Thr Ala Lys Tyr 835 840 845	2544
ggt aca ccg gtg att gat ggc gaa att gat gaa att tgg aat acc acc Gly Thr Pro Val Ile Asp Gly Glu Ile Asp Glu Ile Trp Asn Thr Thr 850 855 860	2592
gaa gag att gaa acc aaa gca gtt gca atg ggt age ctg gat aaa aat Glu Glu Ile Glu Thr Lys Ala Val Ala Met Gly Ser Leu Asp Lys Asn 865 870 875 880	2640
gcc acc gca aaa gtt cgt gtt ctg tgg gat gag aac tat ctg tat gtt Ala Thr Ala Lys Val Arg Val Leu Trp Asp Glu Asn Tyr Leu Tyr Val 885 890 895	2688
ctg gcc att gtt aaa gat ccg gtg ctg aac aaa gat aat agc aat ccg Leu Ala Ile Val Lys Asp Pro Val Leu Asn Lys Asp Asn Ser Asn Pro 900 905 910	2736
tgg gaa cag gat age gtg gaa atc ttt atc gat gaa aat aat cat aaa Trp Glu Gln Asp Ser Val Glu Ile Phe Ile Asp Glu Asn Asn His Lys 915 920 925	2784
acc ggc tat tat gaa gat gat gat gcc cag ttt cgc gtg aat tat atg Thr Gly Tyr Tyr Glu Asp Asp Ala Gln Phe Arg Val Asn Tyr Met 930 935 940	2832
aac gaa cag acc ttt ggc acc ggt ggt tet ccg gca cgt ttt aaa acc Asn Glu Gln Thr Phe Gly Thr Gly Gly Ser Pro Ala Arg Phe Lys Thr 945 950 955 960	2880

[0013]

```

gca gtg aaa ctg att gaa ggt ggc tat att gtt gaa gca gcc att aaa      2928
Ala Val Lys Leu Ile Glu Gly Gly Tyr Ile Val Glu Ala Ala Ile Lys
          965                               970                               975

tgg aaa acc att aaa ccg acc ccg aat acc gtt att ggc ttt aac atc      2976
Trp Lys Thr Ile Lys Pro Thr Pro Asn Thr Val Ile Gly Phe Asn Ile
          980                               985                               990

cag gtg aat gat gcc aat gaa aaa ggt cag cgt gtg ggt att att agc      3024
Gln Val Asn Asp Ala Asn Glu Lys Gly Gln Arg Val Gly Ile Ile Ser
          995                               1000                               1005

tgg tct gat ccg aca aat aat tct tgg cgt gac ccg agc aaa ttt      3069
Trp Ser Asp Pro Thr Asn Asn Ser Trp Arg Asp Pro Ser Lys Phe
          1010                               1015                               1020

ggt aat ctg cgc ctg att aaa      3090
Gly Asn Leu Arg Leu Ile Lys
          1025                               1030

<210> 6
<211> 1030
<212> PRT
<213> 海栖热孢菌

<400> 6

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

Asp Ser Ser Leu Glu Thr Val Leu Ala Leu Ser Phe Glu Gly Thr Thr
20          25          30

Glu Gly Val Val Pro Phe Gly Lys Asp Val Val Leu Thr Ala Ser Gln
35          40          45

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

Ser Pro Trp Asp Gly Val Glu Ile Asp Leu Thr Gly Lys Val Lys Ser
65          70          75          80

Gly Ala Asp Tyr Leu Leu Ser Phe Gln Val Tyr Gln Ser Ser Asp Ala
85          90          95

Pro Gln Leu Phe Asn Val Val Ala Arg Thr Glu Asp Glu Lys Gly Glu
100         105         110

Arg Tyr Asp Val Ile Leu Asp Lys Val Val Val Ser Asp His Trp Lys
115        120        125

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

Tyr Ser Leu Ile Ile Val Ala Ser Lys Asn Thr Asn Phe Asn Phe Tyr
145        150        155        160

Leu Asp Lys Val Gln Val Leu Ala Pro Lys Glu Ser Gly Pro Lys Val
165        170        175

Ile Tyr Glu Thr Ser Phe Glu Asn Gly Val Gly Asp Trp Gln Pro Arg
180        185        190
    
```

[0014]

Gly Asp Val Asn Ile Glu Ala Ser Ser Glu Val Ala His Ser Gly Lys
 195 200 205
 Ser Ser Leu Phe Ile Ser Asn Arg Gln Lys Gly Trp Gln Gly Ala Gln
 210 215 220
 Ile Asn Leu Lys Gly Ile Leu Lys Thr Gly Lys Thr Tyr Ala Phe Glu
 225 230 235 240
 Ala Trp Val Tyr Gln Asn Ser Gly Gln Asp Gln Thr Ile Ile Met Thr
 245 250 255
 Met Gln Arg Lys Tyr Ser Ser Asp Ala Ser Thr Gln Tyr Glu Trp Ile
 260 265 270
 Lys Ser Ala Thr Val Pro Ser Gly Gln Trp Val Gln Leu Ser Gly Thr
 275 280 285
 Tyr Thr Ile Pro Ala Gly Val Thr Val Glu Asp Leu Thr Leu Tyr Phe
 290 295 300
 Glu Ser Gln Asn Pro Thr Leu Glu Phe Tyr Val Asp Asp Val Lys Ile
 305 310 315 320
 Val Asp Thr Thr Ser Ala Glu Ile Lys Ile Glu Met Glu Pro Glu Lys
 325 330 335
 Glu Ile Pro Ala Leu Lys Glu Val Leu Lys Asp Tyr Phe Lys Val Gly
 340 345 350
 Val Ala Leu Pro Ser Lys Val Phe Leu Asn Pro Lys Asp Ile Glu Leu
 355 360 365
 Ile Thr Lys His Phe Asn Ser Ile Thr Ala Glu Asn Glu Met Lys Pro
 370 375 380
 Glu Ser Leu Leu Ala Gly Ile Glu Asn Gly Lys Leu Lys Phe Arg Phe
 385 390 395 400
 Glu Thr Ala Asp Lys Tyr Ile Gln Phe Val Glu Glu Asn Gly Met Val
 405 410 415
 Ile Arg Gly His Thr Leu Val Trp His Asn Gln Thr Pro Asp Trp Phe
 420 425 430
 Phe Lys Asp Glu Asn Gly Asn Leu Leu Ser Lys Glu Ala Met Thr Glu
 435 440 445
 Arg Leu Lys Glu Tyr Ile His Thr Val Val Gly His Phe Lys Gly Lys
 450 455 460
 Val Tyr Ala Trp Asp Val Val Asn Glu Ala Val Asp Pro Asn Gln Pro
 465 470 475 480
 Asp Gly Leu Arg Arg Ser Thr Trp Tyr Gln Ile Met Gly Pro Asp Tyr
 485 490 495

[0015]

Ile Glu Leu Ala Phe Lys Phe Ala Arg Glu Ala Asp Pro Asp Ala Lys
 500 505 510

Leu Phe Tyr Asn Asp Tyr Asn Thr Phe Glu Pro Arg Lys Arg Asp Ile
 515 520 525

Ile Tyr Asn Leu Val Lys Asp Leu Lys Glu Lys Gly Leu Ile Asp Gly
 530 535 540

Ile Gly Met Gln Cys His Ile Ser Leu Ala Thr Asp Ile Lys Gln Ile
 545 550 555 560

Glu Glu Ala Ile Lys Lys Phe Ser Thr Ile Pro Gly Ile Glu Ile His
 565 570 575

Ile Thr Glu Leu Asp Met Ser Val Tyr Arg Asp Ser Ser Ser Asn Tyr
 580 585 590

Pro Glu Ala Pro Arg Thr Ala Leu Ile Glu Gln Ala His Lys Met Met
 595 600 605

Gln Leu Phe Glu Ile Phe Lys Lys Tyr Ser Asn Val Ile Thr Asn Val
 610 615 620

Thr Phe Trp Gly Leu Lys Asp Asp Tyr Ser Trp Arg Ala Thr Arg Arg
 625 630 635 640

Asn Asp Trp Pro Leu Ile Phe Asp Lys Asp His Gln Ala Lys Leu Ala
 645 650 655

Tyr Trp Ala Ile Val Ala Pro Glu Val Leu Pro Pro Leu Pro Lys Glu
 660 665 670

Ser Arg Ile Ser Glu Gly Glu Ala Val Val Val Gly Met Met Asp Asp
 675 680 685

Ser Tyr Leu Met Ser Lys Pro Ile Glu Ile Leu Asp Glu Glu Gly Asn
 690 695 700

Val Lys Ala Thr Ile Arg Ala Val Trp Lys Asp Ser Thr Ile Tyr Ile
 705 710 715 720

Tyr Gly Glu Val Gln Asp Lys Thr Lys Lys Pro Ala Glu Asp Gly Val
 725 730 735

Ala Ile Phe Ile Asn Pro Asn Asn Glu Arg Thr Pro Tyr Leu Gln Pro
 740 745 750

Asp Asp Thr Tyr Ala Val Leu Trp Thr Asn Trp Lys Thr Glu Val Asn
 755 760 765

Arg Glu Asp Val Gln Val Lys Lys Phe Val Gly Pro Gly Phe Arg Arg
 770 775 780

Tyr Ser Phe Glu Met Ser Ile Thr Ile Pro Gly Val Glu Phe Lys Lys
 785 790 795 800

[0016]

Asp Ser Tyr Ile Gly Phe Asp Ala Ala Val Ile Asp Asp Gly Lys Trp
805 810 815

Tyr Ser Trp Ser Asp Thr Thr Asn Ser Gln Lys Thr Asn Thr Met Asn
820 825 830

Tyr Gly Thr Leu Lys Leu Glu Gly Ile Met Val Ala Thr Ala Lys Tyr
835 840 845

Gly Thr Pro Val Ile Asp Gly Glu Ile Asp Glu Ile Trp Asn Thr Thr
850 855 860

Glu Glu Ile Glu Thr Lys Ala Val Ala Met Gly Ser Leu Asp Lys Asn
865 870 875 880

Ala Thr Ala Lys Val Arg Val Leu Trp Asp Glu Asn Tyr Leu Tyr Val
885 890 895

Leu Ala Ile Val Lys Asp Pro Val Leu Asn Lys Asp Asn Ser Asn Pro
900 905 910

Trp Glu Gln Asp Ser Val Glu Ile Phe Ile Asp Glu Asn Asn His Lys
915 920 925

Thr Gly Tyr Tyr Glu Asp Asp Asp Ala Gln Phe Arg Val Asn Tyr Met
930 935 940

Asn Glu Gln Thr Phe Gly Thr Gly Gly Ser Pro Ala Arg Phe Lys Thr
945 950 955 960

Ala Val Lys Leu Ile Glu Gly Gly Tyr Ile Val Glu Ala Ala Ile Lys
965 970 975

Trp Lys Thr Ile Lys Pro Thr Pro Asn Thr Val Ile Gly Phe Asn Ile
980 985 990

Gln Val Asn Asp Ala Asn Glu Lys Gly Gln Arg Val Gly Ile Ile Ser
995 1000 1005

Trp Ser Asp Pro Thr Asn Asn Ser Trp Arg Asp Pro Ser Lys Phe
1010 1015 1020

Gly Asn Leu Arg Leu Ile Lys
1025 1030

<210> 7
<211> 987
<212> DNA
<213> 海栖热孢菌

<220>
<221> CDS
<222> (1)..(987)
<223> XynB, 来自海栖热孢菌

<400> 7
atg cag aat gta tct ctg aga gaa ctc gca gaa aag ctg aac atc tat 48
Met Gln Asn Val Ser Leu Arg Glu Leu Ala Glu Lys Leu Asn Ile Tyr
1 5 10 15

[0017]

att ggt ttt gcc gca atc aac aac ttt tgg tct ctt tcc gac gca gaa Ile Gly Phe Ala Ala Ile Asn Asn Phe Trp Ser Leu Ser Asp Ala Glu 20 25 30	96
aag tac atg gaa gtt gca aga aga gaa ttc aac atc ctg acc cct gag Lys Tyr Met Glu Val Ala Arg Arg Glu Phe Asn Ile Leu Thr Pro Glu 35 40 45	144
aac cag atg aag tgg gat acg att cag cca gaa aga gac aga tac aat Asn Gln Met Lys Trp Asp Thr Ile His Pro Glu Arg Asp Arg Tyr Asn 50 55 60	192
ttc act ccc gct gaa aaa cac gtt gag ttt gca gaa gaa aac gac atg Phe Thr Pro Ala Glu Lys His Val Glu Phe Ala Glu Glu Asn Asp Met 65 70 75 80	240
atc gtg cat gga cac act ctt gtc tgg cac aac cag ett cct gga tgg Ile Val His Gly His Thr Leu Val Trp His Asn Gln Leu Pro Gly Trp 85 90 95	288
atc act ggt aga gaa tgg aca aag gaa gaa ctt ttg aac gtt ctt gaa Ile Thr Gly Arg Glu Trp Thr Lys Glu Glu Leu Leu Asn Val Leu Glu 100 105 110	336
gac cac ata aaa acg gtg gtg tct cat ttc aaa ggt aga gtg aag atc Asp His Ile Lys Thr Val Val Ser His Phe Lys Gly Arg Val Lys Ile 115 120 125	384
tgg gat gtg gtg aac gaa gcg gtg age gat tct gga acc tac agg gaa Trp Asp Val Val Asn Glu Ala Val Ser Asp Ser Gly Thr Tyr Arg Glu 130 135 140	432
agc gtg tgg tac aag acg atc ggt cct gaa tac att gaa aaa gcg ttc Ser Val Trp Tyr Lys Thr Ile Gly Pro Glu Tyr Ile Glu Lys Ala Phe 145 150 155 160	480
aga tgg gca aaa gaa gcc gat cca gat gcg att ctc atc tac aac gac Arg Trp Ala Lys Glu Ala Asp Pro Asp Ala Ile Leu Ile Tyr Asn Asp 165 170 175	528
tac agc ata gaa gaa atc aac gca aaa tgg aac ttc gtc tac aac atg Tyr Ser Ile Glu Glu Ile Asn Ala Lys Ser Asn Phe Val Tyr Asn Met 180 185 190	576
ata aaa gag ctg aaa gaa aag gga gta cct gtt gat gga ata gga ttt Ile Lys Glu Leu Lys Glu Lys Gly Val Pro Val Asp Gly Ile Gly Phe 195 200 205	624
cag atg cac ata gac tac aga ggg ctc aat tat gac agt ttc aga agg Gln Met His Ile Asp Tyr Arg Gly Leu Asn Tyr Asp Ser Phe Arg Arg 210 215 220	672
aat ttg gag aga ttt gcg aaa ctc ggt ctt caa ata tac atc aca gag Asn Leu Glu Arg Phe Ala Lys Leu Gly Leu Gln Ile Tyr Ile Thr Glu 225 230 235 240	720
atg gat gtg aga att cct ctc agt ggt tgg gag gag tat tat ttg aaa Met Asp Val Arg Ile Pro Leu Ser Gly Ser Glu Glu Tyr Tyr Leu Lys 245 250 255	768
aaa cag gct gaa gtt tgt gcg aag atc ttc gat ata tgc ttg gac aac Lys Gln Ala Glu Val Cys Ala Lys Ile Phe Asp Ile Cys Leu Asp Asn 260 265 270	816
cct gca gtt aaa gcg atc cag ttt tgg gga ttc aca gac aaa tac tcc Pro Ala Val Lys Ala Ile Gln Phe Trp Gly Phe Thr Asp Lys Tyr Ser 275 280 285	864
tgg gtt ccc gcg ttt ttc aaa ggg tac ggg aaa gcg ttg ctc ttc gat Trp Val Pro Gly Phe Phe Lys Gly Tyr Gly Lys Ala Leu Leu Phe Asp 290 295 300	912
gag aat tac aac ccc aag cct tgt tat tac gcg ata aaa gag gig ctg Glu Asn Tyr Asn Pro Lys Pro Cys Tyr Tyr Ala Ile Lys Glu Val Leu 305 310 315 320	960

[0018]

gag aaa aag ata gaa gaa aga aaa tga 987
 Glu Lys Lys Ile Glu Glu Arg Lys
 325

<210> 8
 <211> 328
 <212> PRT
 <213> 海栖热孢菌

<400> 8

Met Gln Asn Val Ser Leu Arg Glu Leu Ala Glu Lys Leu Asn Ile Tyr
 1 5 10 15

Ile Gly Phe Ala Ala Ile Asn Asn Phe Trp Ser Leu Ser Asp Ala Glu
 20 25 30

Lys Tyr Met Glu Val Ala Arg Arg Glu Phe Asn Ile Leu Thr Pro Glu
 35 40 45

Asn Gln Met Lys Trp Asp Thr Ile His Pro Glu Arg Asp Arg Tyr Asn
 50 55 60

Phe Thr Pro Ala Glu Lys His Val Glu Phe Ala Glu Glu Asn Asp Met
 65 70 75 80

Ile Val His Gly His Thr Leu Val Trp His Asn Gln Leu Pro Gly Trp
 85 90 95

Ile Thr Gly Arg Glu Trp Thr Lys Glu Glu Leu Leu Asn Val Leu Glu
 100 105 110

Asp His Ile Lys Thr Val Val Ser His Phe Lys Gly Arg Val Lys Ile
 115 120 125

Trp Asp Val Val Asn Glu Ala Val Ser Asp Ser Gly Thr Tyr Arg Glu
 130 135 140

Ser Val Trp Tyr Lys Thr Ile Gly Pro Glu Tyr Ile Glu Lys Ala Phe
 145 150 155 160

Arg Trp Ala Lys Glu Ala Asp Pro Asp Ala Ile Leu Ile Tyr Asn Asp
 165 170 175

Tyr Ser Ile Glu Glu Ile Asn Ala Lys Ser Asn Phe Val Tyr Asn Met
 180 185 190

Ile Lys Glu Leu Lys Glu Lys Gly Val Pro Val Asp Gly Ile Gly Phe
 195 200 205

Gln Met His Ile Asp Tyr Arg Gly Leu Asn Tyr Asp Ser Phe Arg Arg
 210 215 220

Asn Leu Glu Arg Phe Ala Lys Leu Gly Leu Gln Ile Tyr Ile Thr Glu
 225 230 235 240

Met Asp Val Arg Ile Pro Leu Ser Gly Ser Glu Glu Tyr Tyr Leu Lys
 245 250 255

[0019]

Lys Gln Ala Glu Val Cys Ala Lys Ile Phe Asp Ile Cys Leu Asp Asn
 260 265 270

Pro Ala Val Lys Ala Ile Gln Phe Trp Gly Phe Thr Asp Lys Tyr Ser
 275 280 285

Trp Val Pro Gly Phe Phe Lys Gly Tyr Gly Lys Ala Leu Leu Phe Asp
 290 295 300

Glu Asn Tyr Asn Pro Lys Pro Cys Tyr Tyr Ala Ile Lys Glu Val Leu
 305 310 315 320

Glu Lys Lys Ile Glu Glu Arg Lys
 325

<210> 9
 <211> 1143
 <212> DNA
 <213> 嗜热脂肪芽孢杆菌

<220>
 <221> CDS
 <222> (1)..(1143)
 <223> XynA, 来自芽孢杆菌

<400> 9
 atg aaa aat gcc gat agc tat gcc aaa aaa ccg cat att agc gca ctg 48
 Met Lys Asn Ala Asp Ser Tyr Ala Lys Lys Pro His Ile Ser Ala Leu
 1 5 10 15

aat gca ccg cag ctg gat cag cgt tat aaa aat gaa ttt acc att ggt 96
 Asn Ala Pro Gln Leu Asp Gln Arg Tyr Lys Asn Glu Phe Thr Ile Gly
 20 25 30

gca gca gtt gaa ccg tat cag ctg cag aat gaa aaa gat gtg cag atg 144
 Ala Ala Val Glu Pro Tyr Gln Leu Gln Asn Glu Lys Asp Val Gln Met
 35 40 45

ctg aaa cgc cac ttt aat agc att gtg gcc gaa aat gtg atg aaa ccg 192
 Leu Lys Arg His Phe Asn Ser Ile Val Ala Glu Asn Val Met Lys Pro
 50 55 60

att agc att cag ccg gaa gaa ggc aaa ttt aat ttt gaa cag gcc gat 240
 Ile Ser Ile Gln Pro Glu Glu Gly Lys Phe Asn Phe Glu Gln Ala Asp
 65 70 75 80

cgc att gtt aaa ttt gcc aaa gcc aat ggt atg gat att cgc ttt cat 288
 Arg Ile Val Lys Phe Ala Lys Ala Asn Gly Met Asp Ile Arg Phe His
 85 90 95

acc ctg gtt tgg cat agc cag gtt ccg cag tgg ttt ttt ctg gat aaa 336
 Thr Leu Val Trp His Ser Gln Val Pro Gln Trp Phe Phe Leu Asp Lys
 100 105 110

gaa ggc aaa ccg atg gtt aat gaa acc gat ccg gtt aaa cgc gaa cag 384
 Glu Gly Lys Pro Met Val Asn Glu Thr Asp Pro Val Lys Arg Glu Gln
 115 120 125

aat aaa cag ctg ctg ctg aaa cgt ctg gaa acc cat att aaa acc att 432
 Asn Lys Gln Leu Leu Leu Lys Arg Leu Glu Thr His Ile Lys Thr Ile
 130 135 140

gtg gaa cgc tat aaa gat gac att aaa tat tgg gat gtg gtg aat gaa 480
 Val Glu Arg Tyr Lys Asp Asp Ile Lys Tyr Trp Asp Val Val Asn Glu
 145 150 155 160

gtt gtt ggt gat gat ggt aaa ctg cgt aat agc ccg tgg tat cag att 528
 Val Val Gly Asp Asp Gly Lys Leu Arg Asn Ser Pro Trp Tyr Gln Ile
 165 170 175

[0020]

gca ggc att gat tat att aaa gtg gcc ttt cag gca gca cgt aaa tat Ala Gly Ile Asp Tyr Ile Lys Val Ala Phe Gln Ala Ala Arg Lys Tyr 180 185 190	576
ggt ggc gat aat att aaa ctg tat atg aat gat tat aat acc gaa gtg Gly Gly Asp Asn Ile Lys Leu Tyr Met Asn Asp Tyr Asn Thr Glu Val 195 200 205	624
gaa ccg aaa cgt acc gca ctg tat aat etg gtg aaa cag ctg aaa gaa Glu Pro Lys Arg Thr Ala Leu Tyr Asn Leu Val Lys Gln Leu Lys Glu 210 215 220	672
gaa ggc gtt ccg att gac ggt att ggt cat cag agc cat att cag att Glu Gly Val Pro Ile Asp Gly Ile Gly His Gln Ser His Ile Gln Ile 225 230 235 240	720
ggt tgg ccg agc gaa gca gaa att gaa aaa acc att aat atg ttt gca Gly Trp Pro Ser Glu Ala Glu Ile Glu Lys Thr Ile Asn Met Phe Ala 245 250 255	768
gca ctg ggt ctg gat aat cag att acc gaa ctg gat gtt agc atg tat Ala Leu Gly Leu Asp Asn Gln Ile Thr Glu Leu Asp Val Ser Met Tyr 260 265 270	816
ggt tgg cct ccg cgt gca tat ccg acc tat gat gca att ccg aaa cag Gly Trp Pro Pro Arg Ala Tyr Pro Thr Tyr Asp Ala Ile Pro Lys Gln 275 280 285	864
aaa ttt ctg gat cag gca gcc cgt tat gat cgt ctg ttt aaa ctg tat Lys Phe Leu Asp Gln Ala Ala Arg Tyr Asp Arg Leu Phe Lys Leu Tyr 290 295 300	912
gaa aaa ctg agc gat aaa att agc aat gtg acc ttt tgg ggt att gca Glu Lys Leu Ser Asp Lys Ile Ser Asn Val Thr Phe Trp Gly Ile Ala 305 310 315 320	960
gat aat cat acc tgg ctg gat agc cgt gca gat gtt tat tat gat gcc Asp Asn His Thr Trp Leu Asp Ser Arg Ala Asp Val Tyr Tyr Asp Ala 325 330 335	1008
aat ggc aat gtt gtt gtt gat ccg aat gca ccg tat gcc aaa gtg gaa Asn Gly Asn Val Val Val Asp Pro Asn Ala Pro Tyr Ala Lys Val Glu 340 345 350	1056
aaa ggc aaa ggt aaa gat gca ccg ttt gtt ttt ggt ccg gat tat aaa Lys Gly Lys Gly Lys Asp Ala Pro Phe Val Phe Gly Pro Asp Tyr Lys 355 360 365	1104
gtg aaa ccg gca tat tgg gcc att att gat cat aaa tga Val Lys Pro Ala Tyr Trp Ala Ile Ile Asp His Lys 370 375 380	1143
<210> 10	
<211> 380	
<212> PRT	
<213> 嗜热脂肪芽孢杆菌	
<400> 10	
Met Lys Asn Ala Asp Ser Tyr Ala Lys Lys Pro His Ile Ser Ala Leu 1 5 10 15	
Asn Ala Pro Gln Leu Asp Gln Arg Tyr Lys Asn Glu Phe Thr Ile Gly 20 25 30	
Ala Ala Val Glu Pro Tyr Gln Leu Gln Asn Glu Lys Asp Val Gln Met 35 40 45	
Leu Lys Arg His Phe Asn Ser Ile Val Ala Glu Asn Val Met Lys Pro 50 55 60	

[0021]

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

[0022]

Val	Lys	Pro	Ala	Tyr	Trp	Ala	Ile	Ile	Asp	His	Lys					
370						375					380					
<210> 11 <211> 999 <212> DNA <213> 嗜热网球菌																
<220> <221> CDS <222> (1)..(999) <223> XynA, 来自嗜热网球菌																
<400> 11																
atg	ttt	ctg	aaa	gaa	gaa	gcc	aaa	ggt	atg	gaa	att	ccg	agc	ctg	aaa	48
Met	Phe	Leu	Lys	Glu	Glu	Ala	Lys	Gly	Met	Glu	Ile	Pro	Ser	Leu	Lys	
1			5					10					15			
gaa	gtg	tat	aaa	gat	lac	ttt	acc	atc	ggt	gca	gca	gtt	agc	cat	ctg	96
Glu	Val	Tyr	Lys	Asp	Tyr	Phe	Thr	Ile	Gly	Ala	Ala	Val	Ser	His	Leu	
			20					25					30			
aat	att	tat	cat	tat	gag	aac	ctg	ctg	aaa	aaa	cat	ttt	aat	agc	ctg	144
Asn	Ile	Tyr	His	Tyr	Glu	Asn	Leu	Lys	Lys	His	Phe	Asn	Ser	Leu		
			35				40					45				
aca	ccg	gaa	aat	cag	atg	aaa	tgg	gaa	gtg	att	cat	ccg	aaa	ccg	tat	192
Thr	Pro	Glu	Asn	Gln	Met	Lys	Trp	Glu	Val	Ile	His	Pro	Lys	Pro	Tyr	
		50				55					60					
gtg	tat	gat	ttt	ggt	ccg	gca	gat	gaa	att	gtt	gat	ttt	gcc	atg	aaa	240
Val	Tyr	Asp	Phe	Gly	Pro	Ala	Asp	Glu	Ile	Val	Asp	Phe	Ala	Met	Lys	
65					70					75				80		
aat	ggc	atg	aaa	gtt	cgt	ggt	cat	acc	ctg	gtt	tgg	cat	aat	cag	aca	288
Asn	Gly	Met	Lys	Val	Arg	Gly	His	Thr	Leu	Val	Trp	His	Asn	Gln	Thr	
				85				90						95		
ccg	ggt	tgg	gtt	tat	gca	ggc	acc	aaa	gat	gaa	att	ctg	gca	cgc	ctg	336
Pro	Gly	Trp	Val	Tyr	Ala	Gly	Thr	Lys	Asp	Glu	Ile	Leu	Ala	Arg	Leu	
			100					105					110			
aaa	gaa	cat	att	aaa	gaa	gtg	gtg	ggc	cat	tat	aaa	ggt	aaa	gtg	tat	384
Lys	Glu	His	Ile	Lys	Glu	Val	Val	Gly	His	Tyr	Lys	Gly	Lys	Val	Tyr	
			115					120					125			
gcc	tgg	gat	gtt	ggt	aat	gaa	gcc	ctg	agc	gat	aat	ccg	aat	gaa	ttt	432
Ala	Trp	Asp	Val	Val	Asn	Glu	Ala	Leu	Ser	Asp	Asn	Pro	Asn	Glu	Phe	
		130				135						140				
ctg	cgt	cgt	gca	ccg	tgg	tat	gat	att	tgt	ggc	gaa	gaa	gtg	att	gaa	480
Leu	Arg	Arg	Ala	Pro	Trp	Tyr	Asp	Ile	Cys	Gly	Glu	Glu	Val	Ile	Glu	
145					150					155					160	
aaa	gcc	ttt	att	tgg	gca	cat	gaa	ggt	gat	ccg	gat	gcc	aaa	ctg	ttt	528
Lys	Ala	Phe	Ile	Trp	Ala	His	Glu	Val	Asp	Pro	Asp	Ala	Lys	Leu	Phe	
				165						170				175		
tat	aat	gat	tat	aat	ctg	gaa	gat	ccg	att	aaa	cgc	gaa	aaa	gcc	tat	576
Tyr	Asn	Asp	Tyr	Asn	Leu	Glu	Asp	Pro	Ile	Lys	Arg	Glu	Lys	Ala	Tyr	
			180					185						190		
aaa	ctg	gtg	aaa	aaa	ctg	aaa	gat	aaa	ggc	ggt	ccg	att	cat	ggt	att	624
Lys	Leu	Val	Lys	Lys	Leu	Lys	Asp	Lys	Gly	Val	Pro	Ile	His	Gly	Ile	
			195				200					205				
ggt	att	cag	ggt	cat	tgg	acc	ctg	gca	tgg	ccg	acc	ccg	aaa	atg	ctg	672
Gly	Ile	Gln	Gly	His	Trp	Thr	Leu	Ala	Trp	Pro	Thr	Pro	Lys	Met	Leu	
			210				215					220				
gaa	gat	agc	att	aaa	cgt	ttt	gcc	gaa	ctg	ggt	ggt	gaa	ggt	cag	gtg	720
Glu	Asp	Ser	Ile	Lys	Arg	Phe	Ala	Glu	Leu	Gly	Val	Glu	Val	Gln	Val	
225					230						235				240	

[0023]

```

acc gag ttc gac atc agc atc tac tat gat cgc aat gaa aac aat aac      768
Thr Glu Phe Asp Ile Ser Ile Tyr Tyr Asp Arg Asn Glu Asn Asn Asn
                245                                250                                255

ttt aaa gtg cct ccg gaa gat cgt ctg gaa cgt cag gca cag ctg tat      816
Phe Lys Val Pro Pro Glu Asp Arg Leu Glu Arg Gln Ala Gln Leu Tyr
                260                                265                                270

aaa gaa gcc ttt gaa att ctg cgc aaa tac aaa ggc atc gtg acc ggc      864
Lys Glu Ala Phe Glu Ile Leu Arg Lys Tyr Lys Gly Ile Val Thr Gly
                275                                280                                285

gtg acc ttt tgg ggt gtt gca gat gat tat acc tgg ctg tat ttt tgg      912
Val Thr Phe Trp Gly Val Ala Asp Asp Tyr Thr Trp Leu Tyr Phe Trp
                290                                295                                300

ccg gtt cgt ggt cgt gaa gat tat ccg ctg ctg ttt gat aaa aat cat      960
Pro Val Arg Gly Arg Glu Asp Tyr Pro Leu Leu Phe Asp Lys Asn His
305                                310                                315                                320

aat ccg aaa aaa gcc ttt tgg gaa att gtg aaa ttt taa                999
Asn Pro Lys Lys Ala Phe Trp Glu Ile Val Lys Phe
                325                                330

<210> 12
<211> 332
<212> PRT
<213> 嗜热网球菌

<400> 12
Met Phe Leu Lys Glu Glu Ala Lys Gly Met Glu Ile Pro Ser Leu Lys
1                                5                                10                                15

Glu Val Tyr Lys Asp Tyr Phe Thr Ile Gly Ala Ala Val Ser His Leu
20                                25                                30

Asn Ile Tyr His Tyr Glu Asn Leu Leu Lys Lys His Phe Asn Ser Leu
35                                40                                45

Thr Pro Glu Asn Gln Met Lys Trp Glu Val Ile His Pro Lys Pro Tyr
50                                55                                60

Val Tyr Asp Phe Gly Pro Ala Asp Glu Ile Val Asp Phe Ala Met Lys
65                                70                                75                                80

Asn Gly Met Lys Val Arg Gly His Thr Leu Val Trp His Asn Gln Thr
85                                90                                95

Pro Gly Trp Val Tyr Ala Gly Thr Lys Asp Glu Ile Leu Ala Arg Leu
100                               105                               110

Lys Glu His Ile Lys Glu Val Val Gly His Tyr Lys Gly Lys Val Tyr
115                               120

Ala Trp Asp Val Val Asn Glu Ala Leu Ser Asp Asn Pro Asn Glu Phe
130                               135                               140

Leu Arg Arg Ala Pro Trp Tyr Asp Ile Cys Gly Glu Glu Val Ile Glu
145                               150                               155                               160

Lys Ala Phe Ile Trp Ala His Glu Val Asp Pro Asp Ala Lys Leu Phe
165                               170                               175

```

[0024]

Tyr Asn Asp Tyr Asn Leu Glu Asp Pro Ile Lys Arg Glu Lys Ala Tyr
 180 185 190

Lys Leu Val Lys Lys Leu Lys Asp Lys Gly Val Pro Ile His Gly Ile
 195 200 205

Gly Ile Gln Gly His Trp Thr Leu Ala Trp Pro Thr Pro Lys Met Leu
 210 215 220

Glu Asp Ser Ile Lys Arg Phe Ala Glu Leu Gly Val Glu Val Gln Val
 225 230 235 240

Thr Glu Phe Asp Ile Ser Ile Tyr Tyr Asp Arg Asn Glu Asn Asn Asn
 245 250 255

Phe Lys Val Pro Pro Glu Asp Arg Leu Glu Arg Gln Ala Gln Leu Tyr
 260 265 270

Lys Glu Ala Phe Glu Ile Leu Arg Lys Tyr Lys Gly Ile Val Thr Gly
 275 280 285

Val Thr Phe Trp Gly Val Ala Asp Asp Tyr Thr Trp Leu Tyr Phe Trp
 290 295 300

Pro Val Arg Gly Arg Glu Asp Tyr Pro Leu Leu Phe Asp Lys Asn His
 305 310 315 320

Asn Pro Lys Lys Ala Phe Trp Glu Ile Val Lys Phe
 325 330