

[19]中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12N 15/55

C12N 9/14 C02F 1/50

C12N 15/80 C12N 1/15

[12] 发明专利申请公开说明书

[21] 申请号 99814190.9

[43] 公开日 2002 年 6 月 19 日

[11] 公开号 CN 1354795A

[22] 申请日 1999.11.5 [21] 申请号 99814190.9

[30] 优先权

[32] 1998.11.10 [33] US [31] 09/189,497

[86] 国际申请 PCT/US99/26227 1999.11.5

[87] 国际公布 WO00/28043 英 2000.5.18

[85] 进入国家阶段日期 2001.6.7

[71] 申请人 诺沃奇梅兹生物技术有限公司

地址 美国加利福尼亚

[72] 发明人 R·M·博卡

M·W·雷

[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

代理人 黄革生

权利要求书 4 页 说明书 50 页 附图页数 5 页

[54] 发明名称 具有内酯水解酶活性的多肽和编码该多肽的核酸

[57] 摘要

本发明涉及具有内酯水解酶活性的分离的多肽和编码该多肽的分离的核酸序列。本发明还涉及含有该核酸序列的核酸结构、载体和宿主细胞以及制备和使用该多肽的方法。本发明进一步涉及防止微生物生物膜形成的方法。

ISSN 1008-4274

权 利 要 求 书

1. 具有内酯水解酶活性的分离的多肽，其选自：
 - (a) 具有与 SEQ ID NO.2 第 18-400 位氨基酸有至少 95%一致性的氨基酸序列的多肽；
 - (b) 与 SEQ ID NO.1 第 90-1238 位核苷酸有至少 95%同源性的核酸序列所编码的多肽；
 - (c) 具有 SEQ ID NO.2 氨基酸序列的多肽的变体，其含有一或多个氨基酸替换、缺失和/或插入；
 - (d) (a)或(b)的等位变体；和
 - (e) (a)、(b)、或(d)的具有内酯水解酶活性的片段。
2. 权利要求 1 的多肽，其具有与 SEQ ID NO.2 第 18-400 位氨基酸有至少 95%一致性的氨基酸序列。
3. 权利要求 2 的多肽，其具有与 SEQ ID NO.2 第 18-400 位氨基酸有至少 97%一致性的氨基酸序列。
4. 权利要求 1-3 之任意一项的多肽，其含有 SEQ ID NO.2 的氨基酸序列。
5. 权利要求 1-4 之任意一项的多肽，其由 SEQ ID NO.2 的氨基酸序列或其片段组成。
6. 权利要求 5 的多肽，其由 SEQ ID NO.2 的氨基酸序列组成。
7. 权利要求 6 的多肽，其由 SEQ ID NO.2 的第 18-400 位氨基酸组成。
8. 权利要求 1 的多肽，其由与 SEQ ID NO.1 的第 90-1238 位核苷酸

有至少 95%同源性的核酸序列编码。

9. 权利要求 8 的多肽，其由与 SEQ ID NO. 1 的第 90-1238 位核苷酸有至少 97%同源性的核酸序列编码。

10. 权利要求 1 的多肽，其中所述多肽是具有 SEQ ID NO. 2 氨基酸序列的多肽的含有一或多个氨基酸替换、缺失和/或插入的变体。

11. 权利要求 1 的多肽，其由大肠杆菌 NRRL B-30074 中的质粒 pFA0576 所包含的核酸序列编码。

12. 权利要求 1-11 之任意一项的多肽，其具有 SEQ ID NO. 2 的内酯水解酶活性的至少 20%。

13. 具有与权利要求 1-12 之任意一项的多肽相同的内酯水解酶活性的多肽。

14. 含有编码权利要求 1-13 之任意一项的多肽的核酸序列的分离核酸序列。

15. 含有在 SEQ ID NO. 1 的成熟多肽编码序列中具有至少一个突变的核酸序列的分离核酸序列，其中该突变核酸序列编码由 SEQ ID NO. 2 第 18-400 位氨基酸组成的多肽。

16. 含有与在适合表达宿主中指导多肽产生的一个或多个控制序列可操作地连接的权利要求 14 的核酸序列的核酸结构。

17. 含有权利要求 16 的核酸结构的重组表达载体。

18. 含有权利要求 16 的核酸结构的重组宿主细胞。

19. 制备突变核酸序列的方法，其包括 (a) 向 SEQ ID NO. 1 的成熟多肽编码序列中引入至少一个突变，其中该突变核酸序列编码由 SEQ ID NO. 2 的第 18-400 位氨基酸组成的多肽；和 (b) 回收该突变核酸序列。

20. 通过权利要求 19 的方法制备的突变核酸序列。

21. 制备多肽的方法，其包括 (a) 培养含有编码该多肽的权利要求 20 的突变核酸序列的菌株以产生含有该多肽的上清液；和 (b) 回收该多肽。

22. 制备权利要求 1-13 之任意一项的多肽的方法，其包括 (a) 培养菌株以产生含有该多肽的上清液；和 (b) 回收该多肽。

23. 制备权利要求 1-13 之任意一项的多肽的方法，其包括 (a) 在适合该多肽产生的条件下培养包含含有该多肽编码核酸序列的核酸结构的宿主细胞；和 (b) 回收该多肽。

24. 制备多肽的方法，包括 (a) 在有利于该多肽产生的条件下培养宿主细胞，其中该宿主细胞含有在 SEQ ID NO. 1 的成熟多肽编码序列中具有至少一个突变的突变核酸序列，其中该突变核酸序列编码由 SEQ ID NO. 2 的第 18-400 位氨基酸组成的多肽，和 (b) 回收该多肽。

25. 制备权利要求 1-13 之任意一项的多肽的方法，其包括 (a) 在有利于该多肽产生的条件下，培养其中已导入了一个新转录单元的同源重组细胞，其中该新转录单元包含与编码该多肽的内源核酸序列的第二外显子可操作地连接的调节序列、外显子和/或剪接供体位点；和 (b) 回收该多肽。

26. 制备细胞突变体的方法，其包括破坏或缺失编码权利要求 1-13 之任意一项的多肽的核酸序列或其控制序列，由此导致该多肽的产生比所述细胞少的突变体。

27. 通过权利要求 26 的方法制备的突变体。

28. 权利要求 27 的突变体，其进一步含有编码异源蛋白质的核酸序列。

29. 制备异源多肽的方法，其包括 (a) 在有利于该多肽产生的条件下培养权利要求 28 的突变体；和 (b) 回收该多肽。

30. 含有可操作地与编码由 SEQ ID NO. 1 第 38-89 位核苷酸组成的信号肽的核酸序列相连的蛋白编码基因的核酸结构，其中所述基因对于所述核酸序列而言是外源的。

31. 含有权利要求 30 的核酸结构的重组表达载体。

32. 含有权利要求 30 的核酸结构的重组宿主细胞。

33. 制备蛋白质的方法，包括 (a) 在适合该蛋白产生的条件下培养权利要求 32 的重组宿主细胞；和 (b) 回收该蛋白。

34. 用于防止一种或多种微生物在液相-固相界面上形成生物膜的方法，其包括向液相-固相界面施用有效量的含有一种或多种具有内酯水解酶活性的多肽和载体的组合物，以降解由该一种或多种微生物产生的一种或多种内酯，其中生物膜的形成涉及该一种或多种内酯。

说明书

具有内酯水解酶活性的多肽和编码该多肽的核酸

发明背景

发明领域

本发明涉及具有内酯水解酶 (lactonohydrolase) 活性的分离的多肽和编码这种多肽的分离的核酸序列。本发明还涉及包含该核酸序列的核酸结构、载体和宿主细胞, 以及制备和使用这种多肽的方法。

相关领域的描述

内酯水解酶可逆地催化内酯化合物水解成羧酸, 即该酶介导羧基羧酸的内酯和酸形式之间的相互转化。

Shimizu 等 (1992, 欧洲生物化学杂志 (European Journal of Biochemistry) 209: 383-390) 公开了获自尖镰孢 (Fusarium oxysporum) 的内酯水解酶。该酶制品立体特异性地水解醛糖酸内酯如 D-半乳糖酸- γ -内酯和 D-葡萄糖酸- δ -内酯。此外, 尖镰孢的内酯水解酶催化 D-泛酰内酯的不对称水解, D-泛酰内酯可用作合成 D-泛酸盐的手性结构单元 (Shimazu 和 Kataoka, 1996, 纽约科学院年报 (Annals of the New York Academy of Sciences) 799: 650-658; Kataoka 等, 1995, 应用微生物生物技术 (Appl. Microbiol. Biotechnol.) 44: 333-338; Kataoka 等, 1996, Enzyme Microb. Technol. 19: 307-310)。而且, 内酯水解酶不可逆地水解许多芳香族内酯, 如二氢香豆素和尿黑酸内酯。

尖镰孢的内酯水解酶基因的克隆和表达已经公开 (WO 97/10341)。

本发明的目的是提供改善的具有内酯水解酶活性的多肽和编码该多肽的核酸。

本发明概要

本发明涉及具有内酯水解酶活性的分离的多肽, 所述多肽选自:

(a) 具有与 SEQ ID NO. 2 第 18 至 400 位氨基酸有至少 95%一致性的氨基酸序列的多肽;

(b) 与 SEQ ID NO. 1 的第 90 至 1238 位核苷酸有至少 95%同源性的核酸序列所编码的多肽;

(c) 具有 SEQ ID NO. 2 之氨基酸序列的多肽的包含一个或多个氨基酸替代、缺失和/或插入的变体;

(d) (a) 或 (b) 的等位变体 (allelic variant); 和

(e) (a)、(b) 或 (d) 的具有内酯水解酶活性的片段。

本发明还涉及编码所述多肽的分离的核酸序列, 并涉及包含该核酸序列的核酸结构、载体和宿主细胞, 以及制备和使用所述多肽的方法。

本发明进一步涉及防止微生物生物膜形成的方法。

附图简述

图 1A 和 1B 显示 *Fusarium venenatum* 内酯水解酶的 cDNA 序列和推导的氨基酸序列 (分别是 SEQ ID NO. 1 和 2)。

图 2 显示 pDM181 的限制性图谱。

图 3 显示 pSheB1 的限制性图谱。

图 4 显示 pTriggs1 的限制性图谱。

本发明详述

具有内酯酶活性的多肽

术语“内酯水解酶活性”本文定义为催化醛糖酸内酯 (aldonate lactone) 和芳香族内酯水解成相应羧酸的水解酶活性。为了本发明的目的, 内酯水解酶活性按照 Fishbein 和 Bessman, 1996, 生物化学杂志 (Journal of Biological Chemistry) 241: 4835-4841 中描述的程序来测定, 其中测量 D-半乳糖酸- γ -内酯的水解。

在第一个实施方案中, 本发明涉及所具有的氨基酸序列与 SEQ ID NO. 2 的第 18 至 400 位氨基酸 (即成熟多肽) 具有至少约 95%, 优选至少约 97%

一致性程度，并具有内酯水解酶活性的分离的多肽（以下称为“同源多肽”）。在一个优选实施方案中，该同源多肽的氨基酸序列与 SEQ ID NO. 2 的第 18 至 400 位氨基酸有 5 个氨基酸不同，优选有 4 个氨基酸不同，更优选有 3 个氨基酸不同，甚至更优选有 2 个氨基酸不同，最优选有 1 个氨基酸不同。为了本发明的目的，两个氨基酸序列之间的一致性程度采用 Clustal 法 (Higgins, 1989, CABIOS 5: 151-153) 和 LASERGENE™ MEGALIGN™ 软件 (DNASTAR 公司, Madison, WI) 并使用同一性表和下列多序列对比 (multiple alignment) 参数来确定：空位罚分为 10，空位长度罚分 (gap penalty) 为 10。两两对比 (Pairwise alignment) 参数是：Ktuple=1，空位罚分=3，窗口 (windows) =5，对角线 (diagonals) =5。

优选地，本发明多肽包含 SEQ ID NO. 2 的氨基酸序列或其等位变体；或它们的具有内酯水解酶活性的片段。在一个更优选的实施方案中，本发明多肽包含 SEQ ID NO. 2 的氨基酸序列。在另一个优选实施方案中，本发明多肽包含 SEQ ID NO. 2 的第 18 至 400 位氨基酸，或其等位变体；或它们的具有内酯水解酶活性的片段。在另一个优选实施方案中，本发明多肽包含 SEQ ID NO. 2 的第 18 至 400 位氨基酸。在另一个优选实施方案中，本发明的多肽由 SEQ ID NO. 2 的氨基酸序列或其等位变体组成；或由它们的具有内酯水解酶活性的片段组成。在另一个优选实施方案中，本发明的多肽由 SEQ ID NO. 2 的氨基酸序列组成。在另一个优选实施方案中，所述多肽由 SEQ ID NO. 2 的第 18 至 400 位氨基酸或其等位变体组成；或由它们的具有内酯水解酶活性的片段组成。在另一个优选实施方案中，所述多肽由 SEQ ID NO. 2 的第 18 至 400 位氨基酸组成。

SEQ ID NO. 2 的片段是从该氨基酸序列的氨基和/或羧基末端缺失了一个或多个氨基酸的多肽。优选地，片断含有至少 310 个氨基酸残基，更优选至少 340 个氨基酸残基，最优选至少 370 个氨基酸残基。

等位变体是指一个基因占据同一个染色体座位的两个或多个可替代形式中的任意一个。等位变异通过突变自然发生，并可导致种群内的多态现象。基因突变可以是沉默的（在编码的多肽中没有变化），也可编码

具有改变的氨基酸序列的多肽。多肽的等位变体是基因的等位变体编码的多肽。

在第二个实施方案中，本发明涉及在极低严谨条件，优选低严谨条件，更优选中等严谨条件，更优选中高严谨条件，甚至更优选高严谨条件，最优选极高严谨条件下和在相同条件下与下列序列杂交的核酸探针杂交的核酸序列所编码的具有内酯水解酶活性的分离多肽：(i) SEQ ID NO. 1 的第 90-1238 位核苷酸，(ii) 含有 SEQ ID NO. 1 第 90-1238 位核苷酸的基因组序列，(iii) (i) 或 (ii) 的子序列，或 (iv) (i)、(ii) 或 (iii) 的互补链 (J. Sambrook, E. F. Fritsch, 和 T. Maniatus, 1989, 分子克隆实验手册 (Molecular Cloning, A Laboratory Manual), 第二版, Cold Spring Harbor, 纽约)。SEQ ID NO. 1 的子序列可以是至少 100 个核苷酸，优选至少 200 个核苷酸。而且，该子序列可编码具有内酯水解酶活性的多肽片段。本发明多肽也可以是该多肽的具有内酯水解酶活性的等位变体或片段。

根据本领域熟知的方法，可使用 SEQ ID NO. 1 的核酸序列或其子序列，以及 SEQ ID NO. 2 的氨基酸序列或其片段，设计核酸探针以从不同属或种的株系鉴定和克隆编码具有内酯水解酶活性的多肽的 DNA。具体地，该探针可用于与感兴趣的属或种的基因组或 cDNA 杂交，随后进行标准 Southern 印迹程序，以鉴定和分离其中的相应基因。该探针可比完整序列短得多，但应长至少 15 个，优选至少 25 个，更优选至少 35 个核苷酸。也可使用更长的探针。既可使用 DNA 探针，也可使用 RNA 探针。典型地，标记该探针以检测相应基因（例如用 ^{32}P 、 ^3H 、 ^{35}S 、生物素或抗生物素蛋白进行标记）。该探针包括在本发明的范围之内。

因此，可筛选从这些其它生物制备的基因组 DNA 或 cDNA 文库，以获得与上述探针杂交并编码具有内酯水解酶活性的多肽的 DNA。这些其它生物的基因组或其它 DNA 可通过琼脂糖或聚丙烯酰胺凝胶电泳或其它分离技术分离。来自文库的 DNA 或分离的 DNA 可被转移并固定至硝酸纤维素或其它适当的载体材料上。为了鉴定与 SEQ ID NO. 1 或其子序列同源的克隆或 DNA，将该载体材料用于 Southern 印迹。对于本发明的目的，杂

交是指核酸序列与相应于 SEQ ID NO. 1 所示核酸序列、其互补序列或其子序列的标记核酸探针在极低到极高的严谨条件下杂交。在这些条件下与该核酸探针杂交的分子用 X 光胶片检测。

在一个优选实施方案中，核酸探针是编码 SEQ ID NO. 2 之多肽的核酸序列或其子序列。在另一个优选实施方案中，核酸探针是 SEQ ID NO. 1。在另一个优选实施方案中，核酸探针是 SEQ ID NO. 1 的成熟多肽编码区。在另一个优选实施方案中，核酸探针是大肠杆菌 (*E. coli*) NRRL B-30074 中的质粒 pFA0576 所包含的该核酸序列，其中该核酸序列编码具有酸性磷酸酶活性的多肽。在另一个优选实施方案中，核酸探针是大肠杆菌 NRRL B-30074 中的质粒 pFA0576 所包含的该成熟多肽编码区。

对于至少 100 核苷酸长的长探针，极低到极高严谨条件定义为在 $5 \times$ SSPE、0.3% SDS、 $200 \mu\text{g/ml}$ 剪切的变性鲑精 DNA 和或者 25% 甲酰胺（对于极低和低严谨条件）或者 35% 甲酰胺（对于中等和中高严谨条件）或者 50% 甲酰胺（对于高和极高严谨条件）中于 42°C 进行预杂交和杂交，然后进行标准 Southern 印迹程序。

对于至少 100 个核苷酸长的长探针，载体材料最后用 $2 \times$ SSC、0.2% SDS，优选至少在 45°C （极低严谨条件），更优选至少在 50°C （低严谨条件），更优选至少在 55°C （中等严谨条件），更优选至少在 60°C （中高严谨条件），甚至更优选至少在 65°C （高严谨条件），最优选至少在 70°C （极高严谨条件）洗涤三次，每次 15 分钟。

对于长约 15 个核苷酸到约 70 个核苷酸的短探针，严谨条件定义为在使用 Bolton 和 McCarthy (1962, 美国国家科学院院刊 (Proceedings of the National Academy of Sciences USA) 48:1390) 的计算方法计算的 T_m 值以下 5°C – 10°C 于 0.9 M NaCl、0.09 M Tris-HCl pH 7.6、6 mM EDTA、0.5% NP-40、 $1 \times$ Denhardt's 溶液、1 mM 焦磷酸钠、1 mM 磷酸二氢钠 (sodium monobasic phosphate)、0.1 mM ATP 和 0.2 mg 酵母 RNA/ml 中预杂交、杂交和杂交后洗涤，然后进行标准 Southern 印迹程序。

对于长约 15 个核苷酸到约 70 个核苷酸的短探针，载体材料在计算的 T_m 以下 5°C – 10°C 用 $6 \times$ SCC 加 0.1% SDS 洗涤一次，时间 15 分钟，用 $6 \times$

SSC 洗涤两次，每次 15 分钟。

在第三个实施方案中，本发明涉及具有 SEQ ID NO. 2 之氨基酸序列的多肽的含有一或多个氨基酸替代、缺失和/或插入的变体。

该变体多肽的氨基酸序列与 SEQ ID NO. 2 或其成熟多肽的氨基酸序列的差别可以是一个或多个氨基酸残基的插入或缺失和/或一个或多个氨基酸残基被不同氨基酸残基替代。优选地，氨基酸的改变在性质上是较小的，即不显著影响该蛋白质的折叠和/或活性的保守氨基酸替换；小的缺失，典型的是 1 到约 30 个氨基酸的缺失；小的氨基或羧基末端突出，如氨基末端的甲硫氨酸残基；不超过约 20-25 个残基的小接头肽；或通过改变净电荷或其它功能便于纯化的小突出，如聚组氨酸片段、抗原表位或结合域。

保守替代的例子是碱性氨基酸（精氨酸、赖氨酸和组氨酸），酸性氨基酸（谷氨酸和天冬氨酸），极性氨基酸（谷氨酰胺和天冬酰胺），疏水性氨基酸（亮氨酸、异亮氨酸和缬氨酸），芳香族氨基酸（苯丙氨酸、色氨酸和酪氨酸）和小氨基酸（甘氨酸、丙氨酸、丝氨酸、苏氨酸和甲硫氨酸）。通常不改变比活性的氨基酸替代是本领域已知的，并由例如 H. Neurath 和 R. L. Hill, 1979 在蛋白质 (The Proteins), Academic Press, 纽约中作了描述。最常发生的交换是 Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu 和 Asp/Gly, 以及它们的相反替代。

在第四个实施方案中，本发明涉及与具有 SEQ ID NO. 2 之氨基酸序列的多肽或其成熟多肽具有免疫化学同一性或部分免疫化学同一性的分离的多肽。免疫化学同一性通过用熟知的乌赫特朗尼氏双向免疫扩散 (Ouchterlony double immunodiffusion) 程序进行免疫交叉反应同一性试验来确定。具体地，根据 Harboe 和 Ingild 在 N. H. Axelsen, J. Krøll 和 B. Weeks 编著，定量免疫电泳手册 (A Manual of Quantitative Immunoelectrophoresis), Blackwell Scientific Publications, 1973, 第 23 章，或 Johnstone 和 Thorpe, 实用免疫化学 (Immunochemistry in

Practice), Blackwell Scientific Publications, 1982 (更具体的见第 27-31 页) 中描述的程序, 通过免疫兔 (或其它啮齿类动物) 制备含有与具有 SEQ ID NO. 2 之氨基酸序列的多肽或其成熟多肽的表位发生免疫反应或结合的多克隆抗体的抗血清。具有免疫化学同一性的多肽是使用特定免疫化学技术以相同方式如沉淀物的完全融合、相同的沉淀物形态和/或相同的电泳迁移率, 与该抗血清反应的多肽。Axelsen, Bock, 和 KrØll 在 N.H. Axelsen, J. KrØll 和 B. Weeks 编著, 定量免疫电泳手册, Blackwell Scientific Publications, 1973, 第 10 章描述了免疫化学同一性的其它例子。具有部分免疫化学同一性的多肽是使用特定免疫化学技术以部分相同方式如沉淀物的部分融合、部分相同的沉淀物形态和/或部分相同的电泳迁移率, 与该抗血清反应的多肽。Bock 和 Axelsen 在 N.H. Axelsen, J. KrØll 和 B. Weeks 编著, 定量免疫电泳手册, Blackwell Scientific Publications, 1973, 第 11 章描述了部分免疫化学同一性的其它例子。

所述抗体也可以是单克隆抗体。单克隆抗体可根据例如 E. Harlow 和 D. Lane 编著, 抗体实验手册 (Antibodies, A Laboratory Manual), 冷泉港出版社 (Cold Spring Harbor Press), 冷泉港, 纽约的方法制备和使用。

本发明的多肽具有至少 20%, 优选至少 40%, 优选至少 60%, 甚至更优选至少 80%, 甚至更优选至少 90%, 最优选至少 100% 的 SEQ ID NO. 2 之成熟多肽的内酯水解酶活性。

本发明多肽可获自任何属的微生物。为了本发明的目的, 本文中给定来源相连使用的术语“获自”是指本发明核酸序列编码的多肽由该来源或已插入了该来源核酸序列的细胞产生。在一个优选实施方案中, 该多肽被分泌于细胞外。

本发明的多肽可以是细菌多肽。例如, 该多肽可是革兰氏阳性细菌多肽, 如芽孢杆菌属 (Bacillus) 多肽, 如 *Bacillus alkalophilus*, 解淀粉芽孢杆菌 (*Bacillus amyloliquefaciens*)、短芽孢杆菌 (*Bacillus brevis*)、环状芽孢杆菌 (*Bacillus circulans*)、凝结芽孢杆菌 (*Bacillus*

coagulans)、灿烂芽孢杆菌(*Bacillus lautus*)、迟缓芽孢杆菌(*Bacillus lentus*)、地衣芽孢杆菌(*Bacillus licheniformis*)、巨大芽孢杆菌(*Bacillus megaterium*)、嗜热脂肪芽孢杆菌(*Bacillus stearothermophilus*)、枯草芽孢杆菌(*Bacillus subtilis*)或苏云金芽孢杆菌(*Bacillus thuringiensis*)多肽;或链霉菌属(*Streptomyces*)多肽,如浅青紫链霉菌(*Streptomyces lividans*)或鼠灰链霉菌(*Streptomyces murinus*)多肽;或革兰氏阴性细菌多肽,如大肠杆菌或假单胞菌属(*Pseudomonas*)种的多肽。

本发明的多肽可以是真菌多肽,更优选酵母多肽,如假丝酵母属(*Candida*)、克鲁维酵母属(*Kluyveromyces*)、毕赤酵母属(*Pichia*)、酵母属(*Saccharomyces*)、裂殖酵母属(*Schizosaccharomyces*)或 *Yarrowia* 多肽;更优选丝状真菌多肽如 *Acremonium*、曲霉属(*Aspergillus*)、短梗霉属(*Aureobasidium*)、隐球酵母属(*Cryptococcus*)、*Filibasidium*、镰孢霉属(*Fusarium*)、腐质霉属(*Humicola*)、*Magnaporthe*、毛霉属(*Mucor*)、毁丝霉属(*Myceliophthora*)、*Neocallimastix*、脉孢菌属(*Neurospora*)、拟青霉属(*Paecilomyces*)、青霉属(*Penicillium*)、*Piromyces*、裂褶菌属(*Schizophyllum*)、*Talaromyces*、嗜热子囊菌属(*Thermoascus*)、梭孢壳属(*Thielavia*)、*Tolypocladium* 或木霉属(*Trichoderma*)多肽。

在一个优选实施方案中,所述多肽是卡尔酵母(*Saccharomyces carlsbergensis*)、酿酒酵母(*Saccharomyces cerevisiae*)、糖化酵母(*Saccharomyces diastaticus*)、*Saccharomyces douglasii*、克鲁弗酵母(*Saccharomyces kluyveri*)、诺地酵母(*Saccharomyces norbensis*)或 *Saccharomyces oviformis* 多肽。

在另一个优选实施方案中,该多肽是棘孢曲霉(*Aspergillus aculeatus*)、泡盛曲霉(*Aspergillus awamori*)、臭曲霉(*Aspergillus foetidus*)、日本曲霉(*Aspergillus japonicus*)、构巢曲霉(*Aspergillus nidulans*)、黑曲霉(*Aspergillus niger*)、米曲霉(*Aspergillus oryzae*)、*Humicola insolens*、*Humicola lanuginosa*、米黑毛霉(*Mucor miehei*)、

Myceliophthora thermophila、粗糙脉孢菌 (*Neurospora crassa*)、产紫青霉 (*Penicillium purpurogenum*)、*Trichoderma harzianum*、康宁木霉 (*Trichoderma kőningii*)、*Trichoderma longibrachiatum*、*Trichoderma reesei* 或绿色木霉 (*Trichoderma viride*) 多肽。

在另一个优选实施方案中，该多肽是杆孢状镰孢 (*Fusarium bactridioides*)、*Fusarium cerealis*、*Fusarium crookwellense*、大刀镰孢 (*Fusarium culmorum*)、禾本科镰孢 (*Fusarium graminearum*)、禾赤镰孢 (*Fusarium graminum*)、异孢镰孢 (*Fusarium heterosporum*)、合欢木镰孢 (*Fusarium negundi*)、尖镰孢、多枝镰孢 (*Fusarium reticulatum*)、粉红镰孢 (*Fusarium roseum*)、接骨木镰孢 (*Fusarium sambucinum*)、肤色镰孢 (*Fusarium sarcochroum*)、拟分枝孢镰孢 (*Fusarium sporotrichioides*)、*Fusarium sulphureum*、*Fusarium torulosum*、*Fusarium trichothecioides* 或 *Fusarium venenatum* 多肽。

在一个更优选的实施方案中，*Fusarium venenatum* 细胞是 *Fusarium venenatum* A3/5，其最初作为禾本科镰孢 ATCC 20334 保藏，最近由 Yoder 与 Christianson (1998, 真菌遗传学和生物学 (Fungal Genetics and Biology) 23:62-80) 和 O'Donnell 等 (1998, 真菌遗传学和生物学 23:57-67) 重新划分为 *Fusarium venenatum*；以及 *Fusarium venenatum* 的分类学等同物，无论其目前被称为何种种名。在另一个优选实施方案中，正如 WO 97/26330 中所公开的，*Fusarium venenatum* 细胞是 *Fusarium venenatum* A3/5 或 *Fusarium venenatum* ATCC 20334 的形态突变体。

应该理解，对于前面提及的种，本发明既包括完全和不完全阶段，又包括其它分类学等同物，例如无性型，而不论它们被称为何种种名。本领域技术人员将容易识别出适合的等同物。例如，D. L. Hawksworth, P. M. Kirk, B. C. Sutton 和 D. N. Pegler (编者) 1995 年在《Ainsworth & Bisby 的真菌字典》(Ainsworth & Bisby's Dictionary of the Fungi) 第八版 (CAB international, University Press, Cambridge, 英国) 第 173-174 页中定义了镰孢霉属的分类学等同物。

公众可容易地从许多培养物保藏中心获得这些种的菌株，这些培养物

保藏中心例如美国典型培养物保藏中心 (ATCC)、德意志微生物保藏中心 (DSM)、真菌菌种保藏中心 (CBS) 和农业研究机构保藏中心 (NRRL)。

而且, 采用上述探针可从其它来源鉴定并获得这种多肽, 这些来源包括从自然环境 (如土壤、堆肥、水等) 分离的微生物。从自然生活环境中分离微生物的技术是本领域熟知的。然后可通过对另一种微生物的基因组或 cDNA 文库的相似筛选获得核酸序列。一旦用探针检测到编码多肽的核酸序列, 即可利用本领域普通技术人员已知的技术 (见例如 Sambrook 等, 1989, 同上) 分离或克隆该序列。

正如本文所定义的, “分离的” 多肽是基本上不含有其它非内酯水解酶多肽的多肽, 例如其具有通过 SDS-PAGE 确定的至少约 20% 的纯度, 优选至少约 40% 的纯度, 更优选约 60% 的纯度, 甚至更优选约 80% 的纯度, 最优选约 90% 的纯度, 甚至最优选约 95% 的纯度。

本发明核酸序列编码的多肽还包括融合多肽或可切割的融合多肽, 其中另一个多肽融合在该多肽或其片段的 N 端或 C 端。融合多肽通过将编码另一个多肽的核酸序列 (或其部分) 与本发明核酸序列 (或其部分) 融合来产生。产生融合多肽的技术是本领域已知的, 包括连接编码这些多肽的编码序列以使它们符合阅读框, 并且融合多肽的表达处于相同的一个或多个启动子和终止子的控制之下。

核酸序列

本发明还涉及编码本发明多肽的分离的核酸序列。在一个优选的实施方案中, 该核酸序列显示在 SEQ ID NO.1 中。在另一个更优选的实施方案中, 该核酸序列是大肠杆菌 NRRL B-30074 中的质粒 pFA0576 所包含的该序列。在另一个优选实施方案中, 该核酸序列是 SEQ ID NO.1 的成熟多肽编码区。在另一个更优选的实施方案中, 该核酸序列是大肠杆菌 NRRL B-30074 中的质粒 pFA0576 所包含的该成熟多肽编码区。本发明还包括编码具有 SEQ ID NO.2 氨基酸序列的多肽或其成熟多肽的核酸序列, 由于遗传密码的简并性该核酸序列不同于 SEQ ID NO.1。本发明还涉及编码具有内酯水解酶活性的 SEQ ID NO.2 片段的 SEQ ID NO.1 子序列。

SEQ ID NO. 1 的子序列是指 SEQ ID NO. 1 所包含的核酸序列，只是已从 5' 和/或 3' 端缺失了一个或多个核苷酸。优选地，子序列含有至少 930 个核苷酸，更优选至少 1020 个核苷酸，最优选至少 1110 个核苷酸。

本发明还涉及在 SEQ ID NO. 1 的成熟多肽编码序列中含有至少一个突变的突变核酸序列，其中该突变核酸序列编码由 SEQ ID NO. 2 的第 18-400 位氨基酸组成的多肽。

用于分离或克隆编码多肽的核酸序列的技术是本领域已知的，包括从基因组 DNA 分离，从 cDNA 制备，或它们的组合。从该基因组 DNA 克隆本发明核酸序列可通过例如采用熟知的聚合酶链式反应 (PCR) 或抗体筛选表达文库以检测具有共同结构特征的克隆 DNA 片段来实现。见例如，Innis 等，1990，PCR: 方法和应用指南 (PCR: A Guide to Methods and Application)，Academic Press，纽约。可使用其它核酸扩增程序如连接酶链式反应 (LCR)、连接的激活转录 (ligated activated transcription, LAT) 和基于核酸序列的扩增 (nucleic acid sequence-based amplification, NASBA)。该核酸序列可从镰孢霉属株系或另一个或相关生物克隆，因此，可以是该核酸序列多肽编码区的等位变体或种的变体。

本文所用术语“分离的核酸序列”是指基本没有其它核酸序列的，如用琼脂糖电泳确定的至少约 20% 纯，优选至少约 40% 纯，更优选至少约 60% 纯，甚至更优选至少约 80% 纯，最优选至少约 90% 纯的核酸序列。例如，分离的核酸序列可通过遗传工程中使用的标准克隆程序将该核酸序列从其天然位置重新安置到将被复制的不同位置来获得。克隆程序可涉及包含所述多肽编码核酸序列的期望核酸片段的切割和分离、此片段插入载体分子和重组载体引入将在其中复制该核酸序列的多个拷贝或克隆的宿主细胞中。该核酸序列可以是基因组、cDNA、RNA、半合成、合成或其任意的组合来源的。

本发明还涉及与 SEQ ID NO. 1 的成熟多肽编码序列 (即第 90 到 1238 位核苷酸) 具有至少约 95%，优选约 97% 同源性程度并编码活性多肽的核酸序列。为了本发明的目的，两个核酸序列之间的同源性程度通过

Wilbur-Lipman 方法 (Wilbur 和 Lipman, 1983, 美国国家科学院院刊 (Proceedings of the National Academy of Science USA) 80: 726-730) 采用 LASERGENE™ MEGALIGN™ 软件 (DNASTAR, Inc., Madison, WI) 并使用同一性表和下列多序列对比参数进行确定: 空位罚分 10, 空位长度罚分 10. 两两对比参数为 Ktuple=3, 空位罚分=3, 窗口=20.

编码本发明多肽的核酸序列的修饰对于合成与该多肽基本类似的多肽而言是必需的。术语与该多肽“基本类似”是指该多肽的非天然存在形式。这些多肽可以以某种被改造的方式与从其天然来源分离的多肽不同, 如比活性、热稳定性、最适 pH 等不同的变体。该变体序列可基于作为 SEQ ID NO. 1 的多肽编码部分如其子序列的核酸序列, 和/或通过导入不产生该核酸序列编码多肽的另一种氨基酸序列、但符合用于生产所述酶的宿主生物的密码子使用的核苷酸替代, 或者通过导入可能产生不同氨基酸序列的核苷酸替代来构建。对于核苷酸替代的一般描述, 见例如 Ford 等 1991, 蛋白表达和纯化 (Protein Expression and Purification) 2: 95-107.

本领域技术人员将明了, 这些替代可在对分子功能关键的区域之外产生, 并仍导致活性多肽。对于本发明分离核酸序列所编码多肽的活性必需的、并因此优选不进行替代的氨基酸残基可根据本领域已知的方法例如定点诱变或丙氨酸扫描诱变 (alanine-scanning mutagenesis) (见例如 Cunningham 和 Wells, 1989, 科学 (Science) 244:1081-1085) 来鉴定。在后面的这个技术中, 在该分子的每个带正电荷残基处引入突变, 并检测所获突变分子的内酯水解酶活性以鉴定对该分子活性关键的氨基酸残基。也可通过分析用例如核磁共振分析、晶体学或光亲和标记等技术所测定的三维结构来确定底物-酶相互作用的位点 (见例如 de Vos 等, 1992, 科学 255:306-312; Smith 等, 1992, 分子生物学杂志 (Journal of Molecular Biology) 224:899-904; Wlodaver 等, 1992, FEBS Letters 309:59-64)。

本发明还涉及编码本发明多肽的分离核酸序列, 这些序列在极低严谨条件下, 优选低严谨条件下, 更优选中等严谨条件下, 更优选中高严谨

条件下, 甚至更优选高严谨条件下, 最优选极高严谨条件下与在相同条件下和 SEQ ID NO.1 的核酸序列或其互补链杂交的核酸探针杂交; 或本文所定义的这些分离核酸序列的等位变体或子序列 (Sambrook 等, 1989, 见上)。

本发明还涉及通过如下步骤制备的分离的核酸序列: (a) 在极低、低、中等、中高、高或极高的严谨条件下与 (i) SEQ ID NO.1 的第 90-1238 位核苷酸, (ii) 含有 SEQ ID NO.1 的第 90-1238 位核苷酸的基因组序列, (iii) (i) 或 (ii) 的子序列, 或 (iv) (i)、(ii) 或 (iii) 的互补链杂交 DNA; 和 (b) 分离该核酸序列。所述子序列优选是至少 100 个核苷酸的序列, 例如编码具有内酯水解酶活性的多肽片段的序列。

制备突变核酸序列的方法

本发明还涉及制备突变核酸序列的方法, 该方法包括向 SEQ ID NO.1 的成熟多肽编码序列或其子序列中引入至少一个突变, 其中所述突变核酸序列编码由 SEQ ID NO.2 的第 18-400 位氨基酸组成的多肽或其具有内酯水解酶活性的片段。

可采用本发明领域已知的任何方法通过定点诱变实现向所述核酸序列引入突变, 以将一个核苷酸更换成另一个核苷酸。利用带有目的插入片段的超螺旋双链 DNA 载体和含有期望突变的两个合成引物的方法是尤其有用的。每个均与载体相反链互补的这两个寡核苷酸引物在温度循环过程中依靠 Pfu DNA 聚合酶进行延伸。通过引物的引入, 产生含有交错切口的突变质粒。在温度循环之后, 用对甲基化和半甲基化 DNA 特异的 DpnI 处理产物, 以消化亲本 DNA 模板和筛选含有突变的合成 DNA。还可使用其它本领域已知的方法。

核酸结构

本发明还涉及含有与一个或多个控制序列可操作地连接的本发明核酸序列的核酸结构, 其中所述控制序列可在适合的宿主细胞中在与该控制序列相容的条件下指导编码序列的表达。表达将理解为包括多肽所涉及

产生的任何步骤，包括但不限于转录、转录后修饰、翻译、翻译后修饰和分泌。

“核酸结构”在本文中定义为单链或双链核酸分子，该核酸分子可分离自天然存在的基因，或已被修饰而含有以天然不存在的方式联合和并置在一起的核酸片段。当核酸结构含有表达本发明编码序列所需的所有控制序列时，术语“核酸结构”与术语“表达盒”同义。术语“编码序列”在本文中定义为直接规定其蛋白产物的氨基酸序列的核酸序列。编码序列的界限一般由恰好位于 mRNA 5'末端开放阅读框上游的核糖体结合位点（原核生物）或 ATG 起始密码子（真核生物）以及恰好位于 mRNA 3'末端开放阅读框下游的转录终止子来确定。编码序列可包括但不限于 DNA、cDNA 和重组核酸序列。

可通过各种方式操作编码本发明多肽的分离核酸序列以保证多肽的表达。取决于表达载体，在插入载体之前对该核酸序列进行操作是期望的或是必需的。利用重组 DNA 方法修饰核酸序列的技术是本领域熟知的。

术语“控制序列”在本文中定义为包括所有对本发明多肽表达必需或有利的成分。每个控制序列对于编码该多肽的核酸序列而言可以是自身的或是外源的。这些控制序列包括但不限于前导序列、聚腺苷酸化序列、前肽序列、启动子、信号肽序列和转录终止子。最少地，控制序列包括启动子以及转录和翻译终止信号。控制序列可与接头一起提供以引入特异的限制性位点，利于控制序列与编码多肽的核酸序列编码区连接。术语“可操作地连接”在本文中定义为一种构型，其中控制序列被适当地置于相对该 DNA 序列的编码序列的一个位置上以致该控制序列指导多肽的表达。

控制序列可以是适合的启动子序列，即宿主细胞识别以表达所述核酸序列的一种核酸序列。启动子序列含有介导多肽表达的转录控制序列。启动子可以是任何在所选宿主细胞中表现出转录活性的核酸序列，包括突变的、截断的和杂合的启动子，并且可从编码与宿主细胞同源或异源的细胞外或细胞内多肽的基因获得。

特别在细菌宿主细胞中指导本发明核酸结构转录的适合启动子的例子

是获自大肠杆菌 lac 操纵子、天蓝色链霉菌 (*Streptomyces coelicolor*) 琼脂糖酶基因 (dagA)、枯草芽孢杆菌果聚糖蔗糖酶 (levansucrase) (sacB)、地衣芽孢杆菌 α -淀粉酶基因 (amyL)、嗜热脂肪芽孢杆菌产麦芽糖淀粉酶 (maltogenic amylase) 基因 (amyM)、解淀粉芽孢杆菌 α -淀粉酶基因 (amyQ)、地衣芽孢杆菌青霉素酶基因 (penP)、枯草芽孢杆菌 xylA 和 xylB 基因和原核生物的 β 内酰胺酶基因的启动子 (Villa-Kamaroff 等, 1978, 美国国家科学院院刊 75:3727-3731), 以及 tac 启动子 (DeBoer 等, 1983, 美国国家科学院院刊 80:21-25)。更多的启动子描述于《科学美国人》(Scientific American) 中的“从重组细菌获得的有用蛋白 (Useful proteins from recombinant bacteria)”, 1980, 242:74-94; 和 Sambrook 等, 1989, 见上。

在丝状真菌宿主细胞中指导本发明核酸结构转录的适合启动子的例子是从米曲霉 TAKA 淀粉酶、*Rhizomucor miehei* 天冬氨酸蛋白酶、黑曲霉中性 α -淀粉酶、黑曲霉酸稳定的 α -淀粉酶、黑曲霉或泡盛曲霉葡萄糖淀粉酶 (glaA)、*Rhizomucor miehei* 脂肪酶、米曲霉碱性蛋白酶、米曲霉丙糖磷酸异构酶、构巢曲霉乙酰胺酶和尖镰孢胰蛋白酶样蛋白酶 (WO 96/00787) 的基因获得的启动子, 以及 NA2-tpi 启动子 (黑曲霉中性 α -淀粉酶和米曲霉丙糖磷酸异构酶的基因启动子的杂合物), 和它们的突变的、截短的和杂合的启动子。

在酵母宿主中, 有用的启动子可从酿酒酵母烯醇化酶 (ENO-1)、酿酒酵母半乳糖激酶 (GAL1)、酿酒酵母醇脱氢酶/甘油醛-3-磷酸脱氢酶 (ADH2/GAP) 和酿酒酵母 3-磷酸甘油酸激酶的基因获得。用于酵母宿主细胞的其它有用启动子描述于 Romanos 等, 1992, 酵母 (Yeast) 8: 423-488。

控制序列也可以是适当的转录终止子序列, 即被宿主细胞识别以终止转录的序列。终止子序列可操作地与编码所述多肽的核酸序列的 3' 末端相连接。任何在所选宿主细胞中起作用的终止子都可用于本发明。

用于丝状真菌宿主细胞的优选终止子可获自米曲霉 TAKA 淀粉酶、黑曲霉葡萄糖淀粉酶、构巢曲霉邻氨基苯甲酸合成酶、黑曲霉 α -葡萄糖苷酶和

尖镰孢胰蛋白酶样蛋白酶的基因。

用于酵母宿主细胞的优选终止子序列从酿酒酵母烯醇化酶、酿酒酵母细胞色素 C (CYC1) 和酿酒酵母甘油醛-3-磷酸脱氢酶的基因获得。用于酵母宿主细胞的其它有用的终止子序列描述于 Romanos 等, 1992, 同上。

控制序列也可以是适当的前导序列, 一种对宿主细胞翻译重要的 mRNA 非翻译区。前导序列可操作地与编码所述多肽的核酸序列的 5' 端连接。任何在所选宿主细胞中有功能的前导序列都可用于本发明。

用于丝状真菌宿主细胞的优选前导序列可从米曲霉 TAKA 淀粉酶和构巢曲霉丙糖磷酸异构酶的基因获得。

适用于酵母宿主细胞的前导序列可从酿酒酵母烯醇化酶 (ENO-1)、酿酒酵母 3-磷酸甘油酸激酶、酿酒酵母 α -因子和酿酒酵母醇脱氢酶/甘油醛-3-磷酸脱氢酶 (ADH2/GAP) 的基因获得。

控制序列也可以是聚腺苷酸化序列, 一种与所述核酸序列 3' 末端可操作地连接的、当转录时被宿主细胞作为向转录的 mRNA 添加聚腺苷酸残基 (polyadenosine residues) 的信号而识别的序列。任何在所选宿主细胞中有作用的聚腺苷酸化序列均可用于本发明。

用于丝状真菌宿主细胞的优选聚腺苷酸化序列可从米曲霉 TAKA 淀粉酶、黑曲霉葡糖淀粉酶、构巢曲霉邻氨基苯甲酸合成酶、尖镰孢胰蛋白酶样蛋白酶和黑曲霉 α -葡糖苷酶。

用于酵母宿主细胞的有用聚腺苷酸化序列描述于 Guo 和 Sherman, 1995, 分子细胞生物学 (Molecular Cellular Biology) 15: 5983-5990。

控制序列还可以是编码与多肽氨基末端连接的氨基酸序列、并指导编码的多肽进入细胞分泌途径的信号肽编码区。所述核酸序列编码区的 5' 端本身可含有在翻译阅读框中天然与编码分泌多肽的编码区片段连接的信号肽编码区。或者, 编码序列的 5' 端可含有对该编码序列而言外源的信号肽编码区。在编码序列并不天然含有信号肽编码区的时候可能需要外源信号肽编码区。或者, 为了增强多肽的分泌, 外源信号肽编码区可简单替代天然信号肽编码区。然而, 任何指导表达多肽进入所选宿主细胞分泌途径的信号肽编码区均可用于本发明。

用于细菌宿主细胞的有效信号肽编码区是从芽孢杆菌属 NCIB 11837 产麦芽糖淀粉酶、嗜热脂肪芽孢杆菌 α -淀粉酶、地衣芽孢杆菌枯草杆菌蛋白酶、地衣芽孢杆菌 β -内酰胺酶、嗜热脂肪芽孢杆菌中性蛋白酶(nprT, nprS, nprM) 和枯草芽孢杆菌 prsA 的基因获得。更多的信号肽描述于 Simonen 和 Palva, 1993, 微生物学评论 (Microbiological Reviews) 57: 109-137。

用于丝状真菌宿主细胞的有效信号肽编码区是从米曲霉 TAKA 淀粉酶、黑曲霉中性淀粉酶、黑曲霉葡萄糖淀粉酶、Rhizomucor miehei 天冬氨酸蛋白酶、Humicola insolens 纤维素酶和 Humicola lanuginosa 脂肪酶的基因获得的信号肽编码区。

在一个优选实施方案中, 信号肽编码区是 SEQ ID NO. 1 中编码 SEQ ID NO. 2 第 1-17 位氨基酸的第 38-89 位核苷酸。

用于酵母宿主细胞的有用信号肽可从酿酒酵母 α -因子和酿酒酵母转化酶(invertase)的基因获得。其它有用的信号肽编码区描述于 Romanos 等, 1992, 同上。

控制序列还可以是编码位于多肽氨基末端的氨基酸序列的前肽编码区。所获得的多肽被称为酶原(proenzyme)和多肽原(有时也叫酶原(zymogen))。多肽原通常是无活性的, 并能通过前肽的催化切割或自身催化切割从多肽原转化成成熟的活性多肽。前肽编码区可从枯草芽孢杆菌碱性蛋白酶(aprE)、枯草芽孢杆菌中性蛋白酶(nprT)、Saccharomyces cerevisiae α -因子、Rhizomucor miehei 天冬氨酸蛋白酶和 Myceliophthora thermophila 漆酶(WO 95/33836)的基因获得。

当多肽的氨基末端同时存在信号肽和前肽区时, 前肽区紧接在多肽氨基末端之后, 而信号肽区紧接在前肽区的氨基末端之后。

也可期望加入允许相对于宿主细胞生长调节所述多肽表达的调节序列。调节系统的例子是那些使基因表达响应化学或物理刺激包括调节化合物的存在而打开或关闭的调节系统。原核系统中的调节系统包括 lac、tac 和 trp 操纵子系统。在酵母中, 可使用 ADH2 系统或 GAL1 系统。在丝状真菌中, 可使用 TAKA α -淀粉酶启动子、黑曲霉葡萄糖淀粉酶启动子和

米曲霉葡萄糖淀粉酶启动子作为调节序列。调节序列的其它例子是那些允许基因扩增的调节序列。在真核系统中，这些包括在氨甲蝶呤时存在扩增的二氢叶酸还原酶基因和在重金属存在时扩增的金属硫蛋白基因。在这些情况下，编码所述多肽的核酸序列将与调节序列可操作地连接。

本发明还涉及用于改变编码本发明多肽的内源基因的表达的核酸结构。这些结构可包含改变该内源基因表达所必需的最少数目的成分。在一个实施方案中，这些核酸结构优选包含 (a) 导向序列，(b) 调节序列，(c) 外显子，和 (d) 剪接供体位点。该核酸结构导入细胞后，该结构即通过同源重组在内源基因部位插入至细胞基因组中。导向序列指导元件 (a) - (d) 整合至内源基因中，使元件 (b) - (d) 与内源基因可操作地连接。在另一个实施方案中，该核酸结构包含 (a) 靶向序列，(b) 调节序列，(c) 外显子，(d) 剪接供体位点，(e) 内含子，和 (f) 剪接受体位点，其中，靶向序列指导元件 (a) - (f) 的整合，使元件 (b) - (f) 与内源基因可操作地连接。然而，该结构还可包含其它成分如选择标记。

在这两个实施方案中，这些成分的导入均导致产生改变了内源基因表达的新的转录单元。大体上，此新转录单元是通过导向结构引入的序列和内源基因的融合产物。在一个改变了内源基因的实施方案中，该基因被激活。在该实施方案中，通过插入使基因的水平表达比在相应亲本细胞中更高的调节序列，采用同源重组替代、破坏或失效正常与亲代细胞的内源基因有关的调节区。采用本领域熟知的方法（见例如，美国专利 5,641,670），激活的基因可通过在该结构中包含可扩增的选择标记基因来进一步扩增。在另一个改变了内源基因的实施方案中，该基因的表达降低。

导向序列可位于所述内源基因内，紧临该基因，在上游基因内，或在该内源基因上游并相隔一段距离。可使用一个或多个导向序列。例如，环状质粒或 DNA 片段优选使用单个导向序列，而线性质粒或 DNA 片段优选使用两个导向序列。

所述结构的调节序列可由一个或多个启动子、增强子、支架结构附着

所述结构的剪接供体位点指导一个外显子向另一个外显子的剪接。典型地，第一外显子位于第二外显子的 5'，而且重叠和位于第一外显子 3' 侧翼的剪接供体位点识别位于第二外显子 5' 侧翼的剪接受体位点。象剪接供体位点一样，剪接受体位点是指导一个外显子向另一个外显子剪接的序列。联合剪接供体位点一起作用，剪接装置使用剪接受体位点实现内含子的去除。

本发明还涉及包含本发明核酸序列、启动子和转录与翻译终止信号的重组表达载体。以上描述的各种核酸和控制序列可联合在一起，产生可包括一个或多个方便的限制性位点以允许编码所述多肽的核酸序列在这些位点插入或替代的重组表达载体。或者，本发明的核酸序列可通过将该核酸序列或含有该序列的核酸结构插入至适当表达载体中来表达。在建立表达载体时，编码序列位于该载体中使得该编码序列与适当表达控制序列可操作地连接。

该载体可以是自主复制型载体，即作为染色体外实体而存在、其复制

不依赖于染色体复制的载体，如质粒、染色体外因子、微型染色体或人工染色体。该载体可包含任何保证自我复制的序列。或者，该载体可以是当导入宿主细胞时整合至基因组中并和它所整合的染色体一起复制的载体。而且，可使用单个载体或质粒，或一起含有待导入宿主细胞基因组的全部 DNA 的两个或多个载体或质粒，或转座子。

本发明的载体优选含有一个或多个允许方便选择被转化细胞的选择标记。选择标记是其产物提供杀生物剂或病毒抗性、重金属抗性、相对营养缺陷型的原养型等的基因。细菌选择标记的例子是枯草芽孢杆菌或地衣芽孢杆菌的 *dal* 基因，或赋予抗生素抗性如氨苄青霉素、卡那霉素、氯霉素或四环素抗性的标记。用于酵母宿主细胞的适当标记是 *ADE2*、*HIS3*、*LEU2*、*LYS2*、*MET3*、*TRP1* 和 *URA3*。用于丝状真菌宿主细胞的选择标记包括但不限于 *amdS*（乙酰胺酶）、*argB*（鸟氨酸氨甲酰基转移酶）、*bar*（磷丝菌素乙酰基转移酶（phosphinothricin acetyltransferase））、*hygB*（潮霉素磷酸转移酶）、*niaD*（硝酸还原酶）、*pyrG*（乳清酸核苷-5'-磷酸脱羧酶）、*sC*（硫酸腺苷酰转移酶（sulfate adenylyltransferase））、*trpC*（邻氨基苯甲酸合成酶），以及它们的等同物。优选用于曲霉属细胞中的是构巢曲霉或米曲霉的 *amdS* 和 *pyrG* 基因以及吸水链霉菌（*Streptomyces hygroscopicus*）的 *bar* 基因。

本发明的载体优选包含允许该载体稳定整合至宿主细胞基因组中或允许该载体在该细胞中不依赖于基因组而自主复制的一个或多个元件。

为了整合至宿主细胞基因组中，该载体可以依赖于编码所述多肽的核酸序列或任何其它用于通过同源或非同源重组使该载体稳定整合至基因组中的载体元件。或者，该载体可包含通过同源重组指导整合至宿主细胞基因组中的额外核酸序列。这些额外核酸序列使载体在一条或多条染色体的一个或多个精确位置整合至宿主细胞基因组中。为了增加在精确位置整合的可能性，整合元件应优选含有足够数目的与相应靶序列同源以增强同源重组可能性的核酸，如 100 - 10,000 碱基对，优选 400 - 10,000 碱基对，最优选 800 - 10,000 碱基对。整合元件可以是任何与宿主细胞基因组中的靶序列同源的序列。而且，该整合元件可以是非编码或编码

核酸序列。另一方面，该载体可以通过非同源重组整合至宿主细胞的基因组中。

为了自主复制，该载体可进一步包含使载体能在所述宿主细胞中自主复制的复制起点。细菌复制起点的例子是允许在大肠杆菌复制的 pBR322、pUC19、pACYC177 和 pACYC184 质粒和允许在芽孢杆菌属中复制的 pUB110、pE194、pTA1060 和 pAM β 1 质粒的复制起点。用于酵母宿主细胞中的复制起点的例子是 2 μ 复制起点、ARS1、ARS4、ARS1 和 CEN3 的组合以及 ARS4 和 CEN6 的组合。复制起点可以是含有使其功能在宿主细胞中温度敏感的突变的复制起点（见例如，Ehrlich, 1978, 美国国家科学院院刊 75: 1433）。

可向宿主细胞插入一个拷贝以上的本发明核酸序列，以增加基因产物的产生。本发明核酸序列拷贝数的增加可通过如下方法获得，所述方法是将该序列的至少一个额外拷贝整合至宿主细胞基因组中，或包括带有本发明核酸序列的可扩增选择标记基因，其中可通过在适当选择剂存在时培养细胞来选择含有扩增拷贝的选择标记基因、从而含有额外拷贝的本发明核酸序列的细胞。

用于连接上述元件以构建本发明的重组表达载体的程序是本领域技术人员熟知的（见例如，Sambrook 等，1989，同上）。

宿主细胞

本发明还涉及包含本发明核酸序列的重组宿主细胞，其可有利地用于所述多肽的重组产生。将含有本发明核酸序列的载体导入宿主细胞，使得该载体作为前面描述的染色体整合体或自我复制的染色体外载体而存在。术语“宿主细胞”包括任何由于在复制过程中发生的突变而与亲代细胞不同的亲代细胞后代。宿主细胞的选择在很大程度上取决于编码所述多肽的基因及其来源。

宿主细胞可以是单细胞微生物，如原核生物，或非单细胞生物，如真核生物。

有用的单细胞细胞是细菌细胞，如革兰氏阳性细菌，包括但不限于芽

孢杆菌属细胞，如 *Bacillus alkalophilus*、解淀粉芽孢杆菌、环状芽孢杆菌、*Bacillus clausii*、凝结芽孢杆菌、灿烂芽孢杆菌、迟缓芽孢杆菌、地衣芽孢杆菌、巨大芽孢杆菌、嗜热脂肪芽孢杆菌、枯草芽孢杆菌和苏云金芽孢杆菌；或链霉菌属细胞，如浅青紫链霉菌或鼠灰链霉菌，或革兰氏阴性细菌，如大肠杆菌和假单胞杆菌属的种。在一个优选实施方案中，细菌宿主细胞是迟缓芽孢杆菌、地衣芽孢杆菌、嗜热脂肪芽孢杆菌或枯草芽孢杆菌细胞。在另一个优选实施方案中，芽孢杆菌属细胞是嗜碱性的芽孢杆菌属细胞。

例如，载体导入至细菌宿主细胞可通过原生质体转化（见例如 Chang 和 Cohen, 1979, 分子普通遗传学 (Molecular General Genetics) 168: 111-115）、使用感受态细胞（见例如 Young 和 Spizizin, 1961, 细菌学杂志 (Journal of Bacteriology) 81: 823 - 829, 或 Dubnau 和 Davidoff-Abelson, 1971, 分子生物学杂志 (Journal of Molecular Biology) 56: 209-221）、电穿孔（见例如 Shigekawa 和 Dower, 1988, 生物技术 (Biotechniques) 6: 742 - 751）或接合（见例如 Koehler 和 Thorne, 1987, 细菌学杂志 169: 5771 - 5278）来实施。

宿主细胞可以是真核生物，如哺乳动物、昆虫、植物或真菌细胞。

在一个优选的实施方案中，宿主细胞是真菌细胞。这里所用“真菌”包括 phyla Ascomycota, Basidiomycota, Chytridiomycota 和 Zygomycota（如 Hawksworth 等在 Ainsworth and Bisby's Dictionary of The Fungi, 第 8 版, 1995, CAB International, University Press, Cambridge, UK 中所定义的）和以及 Oomycota（见 Hawksworth 等, 1995, 同上, 等 171 页），所有有丝分裂孢子真菌（Hawksworth 等, 1995, 见上）。

在一个更优选的实施方案中，真菌宿主细胞是酵母细胞。这里所用“酵母”包括产子囊酵母（内孢霉目 (Endomycetales)）、产担孢子（basidiosporogenous）酵母和属于半知菌类 (Fungi Imperfecti)（芽孢纲 (Blastomycetes)）的酵母。因为酵母的分类将来可能改变，为了本发明的目的，酵母应按《酵母的生物学和活动》(Biology and Activities of Yeast)（Skinner, F.A., Passmore, S.M., 和 Davenport, R.R. 编

著, Soc. App. Bacteriol. Symposium Series No. 9, 1980) 中的描述来定义。

在一个甚至更优选的实施方案中, 酵母宿主细胞是假丝酵母属、汉逊酵母属 (*Hansenula*)、克鲁维酵母属、毕赤酵母属、酵母属、裂殖酵母属或 *Yarrowia* 细胞。

在一个最优选的实施方案中, 酵母宿主细胞是卡尔酵母、酿酒酵母、糖化酵母、*Saccharomyces douglasii*、克鲁弗酵母、诺地酵母或 *Saccharomyces oviformis* 细胞。在另一个最优选的实施方案中, 酵母宿主细胞是乳酸克鲁维酵母 (*Kluyveromyces lactis*) 细胞。在另一个最优选的实施方案中, 酵母宿主细胞是 *Yarrowia lipolytica* 细胞。

在另一个更优选的实施方案中, 真菌宿主细胞是丝状真菌细胞。“丝状真菌”包括所有丝状形式的真菌门 (*Eumycota*) 和 *Oomycota* (正如 Hawksworth 等, 1995, 见上, 所定义的)。丝状真菌的特征是菌丝体壁由几丁质、纤维素、葡聚糖、脱乙酰壳多糖、甘露聚糖和其它复杂多糖构成。营养生长是通过菌丝延长来进行的, 碳代谢为专性需氧的。相反, 酵母如酿酒酵母的营养生长是通过单细胞菌体的出芽来进行的, 而且碳代谢可以是发酵的。

在一个甚至更优选的实施方案中, 丝状真菌宿主细胞是但不限于 *Acremonium*、曲霉属、镰孢霉属、腐质霉属、毛霉属、毁丝霉属、脉孢菌属、青霉属、梭孢壳属、*Tolypocladium* 或木霉属的种的细胞。

在一个最优选的实施方案中, 丝状真菌宿主细胞是泡盛曲霉、臭曲霉、日本曲霉、构巢曲霉、黑曲霉或米曲霉细胞。在另一个最优选的实施方案中, 丝状真菌宿主细胞是杆孢状镰孢、*Fusarium cerealis*、*Fusarium crookwellense*、大刀镰孢、禾本科镰孢、禾赤镰孢、异孢镰孢、合欢木镰孢、尖镰孢、多枝镰孢、粉红镰孢、接骨木镰孢、肤色镰孢、拟分枝孢镰孢、*Fusarium sulphureum*、*Fusarium torulosum*、*Fusarium trichothecioides* 和 *Fusarium venenatum* 细胞。在一个甚至最优选的实施方案中, 丝状真菌亲代细胞是 *Fusarium venenatum* (Nirenberg sp. nov.) 细胞。在另一个最优选的实施方案中, 丝状真菌宿主细胞是 *Humicola*

节序列、外显子和/或剪接供体位点；和（b）回收该多肽。该方法基于基因激活技术的使用，例如美国专利 5,641,670 所描述的技术。

在本发明的制备方法中，采用本领域已知的方法在适于产生所述多肽的营养培养基中培养细胞。例如，可通过在适当培养基中、在允许所述多肽表达和/或分离的条件下进行摇瓶培养、实验室或工业发酵罐中的小规模或大规模发酵（包括连续、批式、补料分批或固态发酵），来培养细胞。该培养在含有碳和氮源以及无机盐的适合营养培养基中采用本领域已知的程序进行。适合的培养基可从商业厂家获得，也可根据公开的组成（如美国典型培养物保藏中心目录中的）制备。如果该多肽分泌至营养培养基中，则该多肽可直接从培养基中回收。如果该多肽是不分泌的，则它可从细胞裂解液中回收。

所述多肽可采用该多肽特异的本领域已知方法来检测。这些检测方法可包括使用特异抗体、形成酶产物、或酶底物消失。例如，可按本文所述使用酶方法确定该多肽的活性。

所获得的多肽可通过本领域已知的方法回收。例如，可通过传统方法，包括但不限于离心、过滤、提取、喷雾干燥、蒸发或沉淀，从营养培养基中回收所述多肽。

本发明的多肽可通过本领域已知的多种方法来纯化，这些方法包括但不限于层析（如离子交换、亲和、疏水、层析聚焦和大小排阻层析）、电泳方法（如制备型等电聚焦）、溶解度差异（如硫酸铵沉淀）、SDS-PAGE 或提取（见例如，蛋白质纯化（Protein Purification），J.-C. Janson 和 Lars Ryden 编著，VCH Publishers，纽约，1989）。

植物

本发明还涉及已被编码具有内酯水解酶活性的本发明多肽的核酸序列转化以致可以以能回收的量表达和产生该多肽的转基因植物、植物部分或植物细胞。该多肽可从该植物或植物部分回收。或者，含有重组多肽的植物或植物部分可用于改善食物或饲料的品质，如改善营养价值、可口性和流变性质，或破坏抗营养因子。

该转基因植物可以是双子叶的（双子叶植物）或单子叶的（单子叶植物）。单子叶植物的例子是草，如草地草（早熟禾属，*Poa*），饲料草如羊茅属（*festuca*），黑麦草，温带牧草（temperate grass），如剪股颖属（*Agrostis*），和谷类植物，如小麦、燕麦、黑麦、大麦、稻、高粱和玉米。

双子叶植物的例子是烟草，豆科植物，如羽扇豆、马铃薯、甜菜、豌豆、菜豆和大豆，和十字花科植物（十字花科（*Brassicaceae*）），如花椰菜、油籽油菜，和亲缘关系密切的模式生物拟南芥（*Arabidopsis thaliana*）。

植物部分的例子是茎、愈伤组织、叶、根、果实、种子和块茎。具体的植物组织如叶绿体、质外体、线粒体、液泡、过氧化物酶体和细胞质也认为是植物部分。而且，任何植物细胞，无论何种组织来源，均被认为是植物部分。

这些植物、植物部分和植物细胞的后代也包括在本发明的范围之内。

表达本发明多肽的转基因植物或植物细胞可根据本领域已知方法来构建。简而言之，该植物或植物细胞通过将一个或多个编码本发明多肽的表达结构引入植物宿主基因组，并将产生的修饰植物或植物细胞繁殖成转基因植物或植物细胞来构建。

简便地讲，该表达结构是包含编码本发明多肽的核酸序列的核酸结构，其中该核酸序列与在所选植物或植物部分中表达该核酸序列所必需的适当调节序列可操作地连接。而且，该表达结构可包含对鉴定已整合了该表达结构的宿主细胞有用的选择标记，和为该结构导入至所述植物所必需的 DNA 序列（后者取决于所用的 DNA 导入方法）。

调节序列如启动子和终止子序列以及可任选的信号或转移序列的选择基于例如该多肽期望表达的时间、部位以及如何表达。例如，编码本发明多肽的基因的表达可以是组成型的或诱导型的，或者可以是发育、阶段或组织特异的，而且该基因产物可被导向特定组织或植物部分如种子或叶。调节序列描述于例如 Tague 等，1988，植物生理学（*Plant Physiology*）86: 506。

对于组成型表达, 可使用 35S-CaMV 启动子 (Franck 等, 1980, 细胞 (Cell) 21: 285-294)。器官特异性启动子可以是例如贮藏沉积组织 (storage sink tissue) 如种子、马铃薯块茎和果实 (Edwards & Coruzzi, 1990, Ann. Rev. Genet. 24: 275-303) 或代谢沉积组织如分生组织 (Ito 等, 1994, Plant Mol. Biol. 24:863-878) 的启动子, 种子特异性启动子如稻的麦谷蛋白、谷醇溶蛋白 (prolamin)、球蛋白和清蛋白启动子 (Wu 等, 1998, 植物和细胞生理学 (Plant and cell Physiology) 39: 885-889)、来自蚕豆 (*Vicia faba*) 豆球蛋白 B4 和未知种子蛋白基因的蚕豆启动子 (Conrad 等, 1998, 植物生理学杂志 152: 708-711)、种子油体蛋白的启动子 (Chen 等, 1998, 植物和细胞生理学 39: 935-941)、*Brassica napus* 的贮藏蛋白 napA 启动子、或任何其它本领域已知的种子特异性启动子, 如 WO 91/14772 中所描述的。而且, 该启动子可以是叶子特异性启动子如稻或番茄的 *rbcS* 启动子 (Kyoizuka 等, 1993, 植物生理学 102: 991-1000), 小球藻病毒腺嘌呤碱基转移酶 (adenine methyltransferase) 基因启动子 (Mittra 和 Higgins, 1994, 植物分子生物学 (Plant Molecular Biology) 26: 85-93), 或稻的 *aldP* 基因启动子 (Kagaya 等, 1995, 分子和普通遗传学 (Molecular and General Genetics) 248: 668-674), 或创伤诱导的启动子如马铃薯 *pin2* 启动子 (Xu 等, 1993, 植物分子生物学 22: 573-588)。

也可使用启动子增强子元件实现所述酶在植物中的更高表达。例如, 启动子增强子元件可以是位于该启动子和编码本发明多肽的核酸序列之间的内含子。例如, Xu 等, 1993, 见上, 公开了使用稻肌动蛋白 1 基因的第一个内含子以增强表达。

选择标记基因和所述表达结构的任何其它部分均可从本领域现有的选取。

根据本领域已知的传统技术, 将所述核酸结构导入植物基因组中, 这些技术包括农杆菌 (*Agrobacterium*) 介导的转化、病毒介导的转化、显微注射、颗粒轰击、生物轰击转化和电穿孔 (Gasser 等, 1990, 科学 244: 1293; Potrykus, 1990, Bio/Technology 8: 535; Shimamoto 等, 1989,

自然 (Nature) 338: 274)。

目前, 根癌农杆菌 (*Agrobacterium tumefaciens*) 介导的基因转移是选择用于产生转基因双子叶植物的方法 (综述见 Hooykas 和 Schilperoort, 1992, 植物分子生物学 19: 15-38)。然而, 它也可用于转化单子叶植物, 尽管对这些植物通常优选其它转化方法。目前, 选择用于产生转基因单子叶植物的方法是对胚胎愈伤组织或发育中的胚胎进行颗粒轰击 (包被了转化 DNA 的微型金或钨颗粒) (Christou, 1992, 植物杂志 (Plant Journal) 2: 275-281; Shimamoto, 1994, Current opinion Biotechnology 5: 158-162; Vasil 等, 1992, Bio/Technology 10: 668-674)。转化单子叶植物的另一个方法基于 Omirulleh 等, 1993, 植物分子生物学 21: 415-428 所描述的原生质体转化。

转化之后, 根据本领域熟知的方法, 选择其中已导入了所述表达结构的转化体并将其再生成全株。

本发明还涉及制备本发明多肽的方法, 其包括 (a) 在有益于产生该多肽的条件下, 培养含有编码具有内酯水解酶活性的本发明多肽的核酸序列的转基因植物或植物细胞; 和 (b) 回收该多肽。

内酯水解酶活性的去除或降低

本发明还涉及制备亲本细胞的突变细胞的方法, 包括打断或缺失编码所述多肽的核酸序列或其控制序列, 导致当在相同条件下培养时产生的多肽比亲本细胞少的突变细胞。

具有降低内酯水解酶活性的菌株的构建可方便地通过修饰或失活具有内酯水解酶活性的多肽在该细胞中表达必需的核酸序列来实现。待修饰或失活的核酸序列可以是例如编码该多肽或其表现内酯水解酶活性所必需的部分的核酸序列, 或该核酸序列可具有从该核酸序列的编码序列表达该多肽所必需的调节功能。这种调节或控制序列的例子可以是启动子序列或其功能部分, 即足于影响多肽表达的部分。其它可能进行修饰的控制序列在上面已描述过。

所述核酸序列的修饰或失活可通过诱变细胞并选择或筛选内酯水解酶

优选地，根据本发明方法待修饰的细胞是微生物来源的，例如适合产生对于细胞而言同源或异源的期望蛋白产物的真菌株。

本发明还涉及亲本细胞的突变细胞，该突变细胞包含编码所述多肽的核酸序列或其控制序列的破坏或缺失，这导致产生的该多肽比亲本细胞少的突变细胞。

由此产生的多肽缺陷性突变细胞作为表达同源和/或异源多肽的宿主细胞特别有用。因此，本发明还涉及制备同源或异源多肽的方法，该方法包括 (a) 在有益于所述多肽产生的条件下培养突变细胞；并 (b) 回收所述多肽。术语“异源多肽”在本文中定义为对于宿主细胞非天然的多肽，已经过修饰改变了天然序列的天然多肽，或其表达由于用重组 DNA 技术对宿主细胞的操作而在数量上改变的天然蛋白。

在进一步的方面，本发明涉及通过将产生本发明多肽和目的多肽产物的细胞发酵制备基本无内酯水解酶活性的蛋白质产物的方法，该方法包括在发酵完成之前、过程中或之后向发酵培养基中加入有效量的能抑制内酯水解酶活性的试剂，从发酵培养基回收目的产物，并选择性地将回收的产物进行进一步纯化。

在进一步的方面，本发明涉及制备基本上无内酯水解酶活性的蛋白产物的方法，该方法是在允许所述产物表达的条件下培养细胞，将所获培养物接受组合 pH 和温度处理以便相当大地降低内酯水解酶活性，并从所述培养物回收产物。或者，可以对从培养物回收的酶制品进行该组合的 pH 和温度处理。该组合 pH 和温度处理可选择性地联合内酯水解酶抑制剂处理一起使用。

根据本发明的该方面，可以去除至少 60%、优选至少 75%、更优选至少 85%、仍更优选至少 95%、最优选至少 99% 的内酯水解酶活性。可通过使用该方法实现内酯水解酶活性的完全去除。

组合 pH 和温度处理优选在 6.5-7.5 的 pH 范围内以及 40-70℃ 的温度范围内进行足够长的时间，以获得期望的结果，典型地，30-60 分钟足够。

用于培养和纯化目的产物的方法可通过本领域已知的方法进行。

用于产生基本上无内酯水解酶的产物的方法在真核多肽，特别是真菌

蛋白如酶的生产中尤其是有意义的。所述酶可选自例如淀粉分解酶、脂分解酶、蛋白水解酶、细胞裂解酶、氧化还原酶或植物细胞壁降解酶。这些酶的例子包括氨肽酶、淀粉酶、淀粉葡萄糖苷酶、糖酶、羧肽酶、过氧化氢酶、纤维素酶、几丁质酶、角质酶、环糊精糖基转移酶、脱氧核糖核酸酶、酯酶、半乳糖苷酶、 β 半乳糖苷酶、葡糖淀粉酶、葡糖氧化酶、葡糖苷酶、haloperoxidase、半纤维素酶、转化酶、异构酶、漆酶、连接酶、脂肪酶、裂解酶、甘露糖苷酶、氧化酶、果胶分解酶、过氧化物酶、肌醇六磷酸酶、酚氧化酶、多酚氧化酶、蛋白水解酶、核糖核酸酶、转移酶、谷氨酰胺转移酶或木聚糖酶。这些内酯水解酶缺陷性细胞还可用于表达药学上有意义的异源蛋白质例如激素、生长因子、受体等。

应该理解，术语“真核多肽”不仅包括天然多肽，还包括那些经氨基酸替换、缺失或添加，或其它增强活性、热稳定性、pH 耐受性等的修饰方式修饰的多肽，如酶。

在进一步的方面，本发明涉及通过本发明方法制备的基本无内酯水解酶活性的蛋白产物。

生物膜

本发明多肽可用于防止微生物生物膜（又称粘液）的产生。

生物膜是在水环境中的固体表面产生并维持的生物学薄膜，其来源于微生物细胞在固体表面的吸附（Palmer 和 White, 1997, 微生物学进展 (Trends in Microbiology) 5:435-440; Costerton 等, 1987, 微生物学年评 (Annual Reviews of Microbiology) 41:435-464; Mueller, 1994, TAPPI Proceedings, 1994 生物科学专题讨论会 195-201)。这种吸附可以为微生物提供一种竞争优势，因为藉此它们可进行繁殖、容易到达更宽的营养和氧环境、不被冲走、并且对抗微生物剂更不敏感。生物膜的形成也可通过外聚物 (exo-polymeric materials) (多糖、多糖醛酸、藻酸盐、糖蛋白和蛋白质) 的产生来实现，这些外聚物与细胞一起形成被充水间隔分离的厚分化结构层 (McEldowney 和 Fletcher, 1986, 普通微生物学杂志 (Journal of General Microbiology) 132:513-523;

Sutherland, 原核细胞的表面碳水化合物 (Surface Carbohydrates of the Prokaryotic Cell), Academic Press, 纽约, 1977, 第 27-96 页)。驻留微生物可以是单个种类的微生物细胞或微生物细胞混合群落, 可包括需氧和厌氧细菌、藻类、原生动物和真菌。

因此, 生物膜是一种埋植在由驻留微生物分泌的一或多种基质聚合物构成的有机结构中的活微生物复杂系统。

生物膜可发展成几毫米或厘米厚的宏观结构并可覆盖大面积的表面。这些形成物可限制或完全阻断管道系统中的液流、降低热交换器中的热传导, 或在城市供水、食品加工、医药装置 (如导管、矫形外科装置、植入物) 中引起致病问题, 并常常通过植入微生物介导的腐蚀作用减少材料寿命。在工业水加工系统、纸浆和纸生产工艺、冷却水系统、油回收用的喷射井、冷却塔、多孔介质 (沙和土)、海洋环境、和空调系统以及任何密式水再循环系统中, 这种生物学污垢是一个严重的经济问题。

传统上, 生物膜的清除或防止需要使用分散剂、表面活性剂、去污剂、酶制剂、抗微生物剂、杀生物剂、煮沸操作和/或腐蚀性化学制品如碱。采用这些措施的方法是本领域所熟知的。例如, 在纸浆和纸工业中传统上造纸机器中生物膜积累物的清除需要一个沉积物控制程序, 包括正确的设备管理和保养工作以避免表面有溅出的原料、新鲜水和添加物的抗微生物处理、使用杀生物剂以减少机器上微生物的生长、以及定期的煮沸以清除形成的沉积物。

铜绿假单胞菌的生物膜形成涉及至少两个参与细胞-细胞间通讯的细胞外信号的产生 (WO 98/58075)。这两个细胞-细胞间信号系统是 lasR-lasI 和 rhlR-rhlI (又称 vsmR-vsmI) 系统 (Davies 等, 1998, 科学 280:295-298)。lasI 基因指导可扩散细胞外信号 N-(3-氧代十二烷酰)-L-高丝氨酸内酯的合成。lasR 产物是一个转录调节物, 该调节物需要足够水平的 N-(3-氧代十二烷酰)-L-高丝氨酸内酯以激活大量毒性基因 (virulence gene) 包括 lasI 和 rhlR-rhlI 系统。rhlI 基因指导细胞外信号 N-丁酰 buytryl)-L-高丝氨酸内酯的合成, 这是 rhlR 基因产物激活毒性基因和表达稳定期因子 RpoS 所必需的。这种类型的基因调节已被

称作定额感知和应答 (quorum sensing and response)。Davis 等已阐明, lasR-lasI 系统参与细胞膜形成的分化。WO 98/58075 提供了一种通过 lasR-lasI 系统操纵细菌中的细胞-细胞间通讯以控制生物膜构造和结构完整性的方法。

本发明还涉及用于防止液相-固相界面上一种或多种微生物引起的生物膜形成的方法, 包括向液相-固相界面施用有效量的包含具有内酯水解酶活性的一种或多种多肽和载体的组合物, 以降解由这一种或多种微生物产生的一种或多种内酯, 其中生物膜的形成涉及此一种或多种内酯。

所述内酯可以是生物膜形成所涉及的任何内酯。在一个优选实施方案中, 所述内酯是高丝氨酸内酯。在一个更优选的实施方案中, 所述内酯是 N-(3-氧代十二烷酰)-L-高丝氨酸内酯。在另一个更优选的实施方案中, 所述内酯是 N-丁酰-L-高丝氨酸内酯。

所述液相-固相界面可以是任何易于形成生物膜的系统, 尤其是上面所列出的系统。

生物膜可由两种或多种微生物的综合群落或由一种微生物产生。所述微生物可以是生物膜形成所涉及的任何微生物, 包括需氧或厌氧细菌 (革兰氏阳性和革兰氏阴性)、真菌 (酵母或丝状真菌)、藻类和原生动物。在一个优选的实施方案中, 所述细菌是需氧细菌。在另一个优选的实施方案中, 所述细菌是厌氧细菌。在一个更优选的实施方案中, 所述需氧细菌是假单胞菌属。在另一个更优选的实施方案中, 所述需氧细菌是黄杆菌属 (Flavobacterium)。在另一个更优选的实施方案中, 所述厌氧细菌是脱硫弧菌属 (Desulfovibrio)。在一个最优选的实施方案中, 所述需氧细菌是铜绿假单胞菌 (Pseudomonas aeruginosa)。在另一个最优选的实施方案中, 所述厌氧细菌是脱硫脱硫弧菌 (Desulfovibrio desulfuricans)。

含有一种或多种具有内酯水解酶活性的多肽的组合物可以是液体、喷雾或粉末制剂。为了降解、清除或防止生物膜的形成该组合物可增加一种或多种试剂。这些试剂包括但不限于分散剂、表面活性剂、去污剂、酶制剂、抗微生物剂和杀生物剂。

在一个优选的实施方案中，所述试剂是表面活性剂。在一个更优选的实施方案中，所述表面活性剂是十二烷基硫酸钠、季铵化合物、烷基碘化吡啶鎓、Tween 80、Tween 85、Triton X-100、Brij 56、生物学表面活性剂、鼠李糖脂、枯草菌表面活性剂、粘液菌素（viscosin）或磺酸盐。

在一个优选的实施方案中，所述试剂是一种或多种酶。在一个更优选的实施方案中，此一种或多种酶选自蛋白酶、藻酸盐裂解酶和糖酶。

本发明还涉及用于防止生物膜形成的组合物。而且，该组合物可以是消毒剂组合物。对于来自包括但不限于假单胞菌科（*Pseudomonadaceae*）、固氮菌科（*Azotobacteraceae*）、根瘤菌科（*Rhizobiaceae*）、甲基球菌科（*Methylococcaceae*）、盐杆菌科（*Halobacteriaceae*）、军团菌科（*Legionellaceae*）、奈瑟氏球菌科（*Neisseriaceae*）的革兰氏阴性细菌而言，该消毒剂组合物作为消毒剂可能是有用的。

其它用途

本发明还涉及使用具有内酯水解酶活性的多肽的其它方法。

本发明多肽可用于将醛糖酸 γ -或 δ -内酯（例如 D-古洛糖酸- γ -内酯、L-甘露糖酸- γ -内酯、D-半乳糖酸- γ -内酯、或 D-葡萄糖酸- δ -内酯）或芳香族 γ -或 δ -内酯（例如二氢香豆素和尿黑酸内酯）水解成相应的酸。

根据 WO 97/10341 中描述的方法，本发明多肽还可用于通过 D-选择性不对称水解进行 D,L-泛酰内酯（pantolactone）的光学拆分以产生 D-泛酸盐。然而，对于酶学光学拆分，采用本领域熟知的方法，该酶可以是细胞结合的、固定在载体上、或以游离形式使用。

通过水解醛糖酸或芳香族 γ -或 δ -内酯例如柠檬苦素，本发明多肽还可用于去除柠檬汁的苦味。

本发明多肽还可用于污染了棒曲霉素的苹果汁的去毒。

信号肽

本发明还涉及含有可操作地与由 SEQ ID NO.1 第 38-89 位核苷酸组

成的核酸序列连接的蛋白质编码基因的核酸结构，其中所述 SEQ ID NO. 1 第 38-89 位核苷酸编码由 SEQ ID NO. 2 的第 1-17 位氨基酸组成的信号肽，而且所述基因对于所述核酸序列而言是外源的。

本发明还涉及含有这种核酸结构的重组表达载体和重组宿主细胞。

本发明还涉及制备蛋白质的方法，其包括 (a) 在适合蛋白产生的条件下培养这样的重组宿主细胞；和 (b) 回收该蛋白质。

所述核酸序列可以可操作地与带有其它控制序列的外源基因相连，这种其它控制序列描述于上文。

对于宿主细胞而言，该蛋白质可是自身的或异源的。术语“蛋白质”在此并非旨在指特定长度的编码产物，因此包括肽、寡肽和蛋白质。术语“蛋白质”还包括联合形成编码产物的两个或多个多肽。该蛋白质也包括含有获自至少两种不同蛋白（其中对于宿主细胞而言一或多种可以是异源的或自身的）的部分或全部多肽序列之组合的杂合多肽。蛋白质还包括上述蛋白或杂合蛋白的天然存在的等位和工程性变体。

优选地，所述蛋白质是激素、激素变体、酶、受体或其片段、抗体或其片段、或报道分子。在一个更优选的实施方案中，所述蛋白质是氧化还原酶、转移酶、水解酶、裂解酶、异构酶或连接酶。在一个甚至更优选的实施方案中，所述蛋白质是氨肽酶、淀粉酶、糖酶、羧肽酶、过氧化氢酶、纤维素酶、几丁质酶、角质酶、环糊精糖基转移酶、脱氧核糖核酸酶、酯酶、 α -半乳糖苷酶、 β -半乳糖苷酶、葡糖淀粉酶、 α -葡萄糖苷酶、 β -葡萄糖苷酶、转化酶、漆酶、脂肪酶、甘露糖苷酶、mutanase、氧化酶、果胶分解酶、过氧化物酶、肌醇六磷酸酶、多酚氧化酶、蛋白水解酶、核糖核酸酶、谷氨酰胺转移酶或木聚糖酶。

所述基因可从任何原核、真核或其它来源获得。

本发明通过如下实施例作进一步描述，这些实施例不应理解为是对本发明范围的限制。

实施例

用作缓冲液和培养基的化学制品是至少试剂级别的商品产品。

培养基和溶液

COVE 痕量金属溶液每升由 0.04g $\text{NaB}_4\text{O}_7 \cdot 10\text{H}_2\text{O}$ 、0.4g $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ 、1.2g $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ 、0.7g $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ 、0.8g $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ 和 10g $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ 组成。

50× COVE 盐溶液每升由 26g KCl 、26g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 、76g KH_2PO_4 和 50ml COVE 痕量金属溶液组成。

COVE 培养基每升由 342.3 g 蔗糖、20ml 50× COVE 盐溶液、10ml 1M 乙酰胺、10ml 1.5M CsCl_2 和 25g 纯净琼脂组成。

50× Vogels 培养基每升由 150g 柠檬酸钠、250g KH_2PO_4 、10g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 、10g $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ 、2.5ml 生物素贮存液和 5.0ml AMG 痕量金属溶液。

COVE 顶层琼脂每升由 20ml 50× COVE 盐、0.8M 蔗糖、1.5M 氯化铯、1.0M 乙酰胺、20 g 低熔点琼脂糖组成，调节 pH 至 6.0。

RA 孢子形成培养基每升由 50 g 琥珀酸、12.1 g NaNO_3 、1 g 蔗糖、20ml 50× Vogels 和 0.5ml 10mg/ml NaMoSO_4 贮存液组成，调节 pH 至 6.0。

YEPG 培养基每升由 10g 酵母提取物、20g 蛋白胨和 20g 葡萄糖组成。

STC 由 0.8 M 山梨糖醇、25mM Tris pH 8、25mM CaCl_2 组成。

SPTC 由 40% PEG 4000、0.8 M 山梨糖醇、25 mM Tris pH 8、25 mM CaCl_2 组成。

M400Da 培养基每升由 50 g 麦芽糖糊精、2 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 、2g KH_2PO_4 、4g 柠檬酸、8g 酵母提取物、2g 尿素和 1ml COVE 痕量金属溶液组成。

实施例 1: 发酵和菌丝体组织

采用补料分批发酵方案，使用作为碳源的 NUTRIOSE™ (Roquette Freres, S. A., Beinheim, 法国) 和酵母提取物，将镰孢霉属株系 ATCC 20334 的形态突变体 *Fusarium venenatum* CC1-3 (Wiebe 等, 1991, Mycol. Research 95: 1284-1288) 生长在实验室规模的两升发酵罐中。pH 维持

在 6-6.5, 温度保持在 30℃, 含有活性溶解氧 (positive dissolved oxygen)。

在接种后第 2、4、6 和 8 天收获菌丝体样品, 并迅速冷冻于液氮中。将这些样品保存在 -80℃, 直到将其破碎用于 RNA 提取。

实施例 2: cDNA 文库构建

根据 Timberlake 和 Barnard 的方法 (1981, 细胞 (Cell) 26: 29-37) 从实施例 1 中描述的菌丝体样品提取总细胞 RNA, 并在从 1% 甲醛-琼脂糖凝胶印迹之后通过 Northern 杂交分析该 RNA 样品 (Davis 等, 1986, 分子生物学基本方法 (Basic Methods in Molecular Biology), Elsevier Science Publishing Co., Inc., 纽约)。使用 mRNA 分离试剂盒 (mRNA Separator Kit™) (Clontech Laboratories, Inc., Palo Alto, CA) 根据厂商说明从总 RNA 分离聚腺苷酸化 mRNA 部分。使用大约 5μg poly(A)⁺mRNA 根据 Gubler 和 Hoffman 的方法 (1983, 基因 (Gene) 25:263-269) 合成双链 cDNA, 不同的是使用 NotI-(dT)18 引物 (Pharmacia Biotech, Inc., Piscataway, NJ) 起始第一链合成。使用绿豆核酸酶 (Boehringer Mannheim Corporation, Indianapolis, IN) 处理该 cDNA, 并用 T4 DNA 聚合酶 (New England Biolabs, Beverly, MA) 使末端平端化。

用 NotI 消化 cDNA, 通过琼脂糖凝胶电泳按大小选择片段 (大约 0.7-4.5 kb), 并与用 NotI 和 EcoRV 切割并经小牛小肠碱性磷酸酶 (Boehringer Mannheim Corporation, Indianapolis, IN) 去酸化的 pZER0-2.1 (Invitrogen Corporation, Carlsbad, CA) 连接。使用该连接混合物转化大肠杆菌 TOP10 感受态细胞 (Invitrogen Corporation, Carlsbad, CA)。在含有 50μg/ml 终浓度卡那霉素的 2YT 琼脂平板上筛选转化体 (Miller, 1992, 细菌遗传学简短教程。关于大肠杆菌和相关细菌的实验室指南和手册 (A Short Course in Bacterial Genetics. A laboratory Manual and Handbook for Escherichia coli and Related Bacteria), Cold Spring Harbor Press, Cold Spring Harbor, 纽约)。

使用质粒克隆载体 pZErO-2.1 构建两个独立的定向 (directional) cDNA 文库。文库 A 使用第 4 天收获的菌丝体的 mRNA 制备, 文库 B 使用从第 6 天时间点获得的 mRNA 构建。为了检验细胞中基因表达谱的代表性“瞬态图 (snapshot)”, 两个 cDNA 文库均没有扩增。相反地, 将文库铺板, 标定滴度, 并通过 DNA 测序分析每个文库获得的独立克隆。

文库 A (第 4 天的细胞) 由大约 7.5×10^4 个独立克隆组成, 文库 B (第 6 天的细胞) 由大约 1.2×10^5 个克隆组成。每个文库中从 40 个克隆分离小量制备的 DNA (Miniprep DNA), 并检查 cDNA 插入片段的存在和大小。在该分析中, 来自文库 A 的 40 个克隆中有 39 个 (97.5%) 含有大小在 600bp 至 2200bp 之间 (平均=1050bp) 的插入片段。相似地, 来自文库 B 的 40 个克隆中 39 个 (97.5%) 含有大小在 800bp 至 3600bp 之间 (平均=1380bp) 的插入片段。

实施例 3: 模板的制备和核苷酸序列测定

从实施例 2 中描述的每个 cDNA 文库中, 直接从转化平板将 1192 个转化体克隆挑取至含有包含 $50 \mu\text{g}/\text{ml}$ 卡那霉素的 $200 \mu\text{l}$ 2YT 培养液 (Miller, 1992, 见上) 的 96 孔微量滴定板中。将微量滴定板在 37°C 不摇动孵育过夜。孵育之后, 向每个孔加入 $100 \mu\text{l}$ 无菌的 50% 甘油。将转化体复制到第二深盘型 (deep dish) 96 孔微量培养板 (Advanced Genetic Technologies Corporation, Gaithersburg, MD) 中, 该微量培养板的每孔中含有 1ml 补充了 $50 \mu\text{g}/\text{ml}$ 卡那霉素的 Magnificent Broth™ (MacConnell Research, San Diego, CA)。原始的微量滴定板冷冻储存在 -80°C 。第二深盘型培养板于 37°C 旋转摇床上剧烈振荡 (300rpm) 孵育过夜。为防止溅出和交叉污染, 以及允许充分的通气, 用聚丙烯衬垫 (Advanced Genetic Technologies Corporation, Gaithersburg, MD) 和塑料微量滴定盘盖覆盖每个第二培养板。

使用 Utterback 等所修改的 Advanced Genetic Technologies 公司 (Gaithersburg, MD) 96 孔 Miniprep 试剂盒程序 (1995, 基因组科学技术 (Genome Sci. Technol.) 1:1-8), 从每个孔分离 DNA。使用 Perkin-Elmer

Applied Biosystems 377 XL 型自动 DNA 测序仪 (Perkin-Elmer Applied Biosystems 公司, Foster City, CA), 采用染料终止物化学法 (dye-terminator chemistry) (Giesecks 等, 1992, 病毒学方法杂志 (Journal of Virology Methods) 38:47-60) 以及反向 lac 测序引物进行单次 (single-pass) DNA 测序。

实施例 4: DNA 序列数据分析

仔细检查核苷酸序列数据的质量, 给出错误间隔或超过 2% 模糊水平的样品被丢弃或进行重新测序。借助于 FACTURA™ 软件 (Perkin Elmer Applied Biosystems, Inc., Foster City, CA) 手工修剪载体序列。此外, 当模糊碱基信号数目增加, 在每个样品的末端截断序列。使用 AutoAssembler™ 软件 (Perkin Elmer Applied Biosystems, Inc., Foster City, CA) 将所有序列互相比对以确定多重性。最后, 按三种框架翻译所有序列, 并使用采用修改的 Smith-Waterman 算法的 GeneAssist™ 软件 (Perkin Elmer Applied Biosystems, Inc., Foster City, CA) 对非冗余数据库 (NRDB) 进行搜索, 其中所述算法使用阈值为 70 的 BLOSUM 62 矩阵。NRDB 是由 Genpept、Swiss-Prot 和 PIR 数据库组合而来。

实施例 5: 内酯水解酶 cDNA 克隆的鉴定

如实施例 4 所述, 通过使用 Applied Biosystems 377 XL 型自动 DNA 测序仪根据厂商说明进行随机 cDNA 克隆的部分测序, 并将推导的氨基酸序列与尖镰孢内酯水解酶 (Swissprot 保藏号 W21857) 的氨基酸序列比较, 鉴定推测的内酯水解酶克隆。在以此方式发现的几个克隆中, 基于其与尖镰孢内酯水解酶氨基酸序列的对比, 以及使用 signal-P 计算机程序 (Nielsen 等, 1997, 蛋白质工程 (Protein Engineering) 10:1-6) 检测的可能信号肽的存在, 推测一个是全长的。选择含有 pFA0576、定名为大肠杆菌 FA0576 的克隆进行核苷酸序列分析和表达研究。

实施例 6: 大肠杆菌 FA0576 来源的 *Fusarium venenatum* 内酯水解酶 cDNA

的核苷酸序列测定和特性分析

使用染料终止物化学法用 Applied Biosystems 377 XL 型自动 DNA 测序仪进行 DNA 测序。使用转座子插入策略(Primer Island Transposition 试剂盒, Perkin-Elmer/Applied Biosystems, Inc., Foster City, CA) 产生连续序列。对大肠杆菌 FA0567 来源的内酯水解酶克隆进行测序, 达到平均冗余 5.9。

该内酯水解酶克隆编码一个 1200bp 的开放阅读框, 该阅读框编码一个 400 个氨基酸的多肽。其核苷酸序列 (SEQ ID NO. 1) 和推导的氨基酸序列 (SEQ ID NO. 2) 显示在图 1 中。使用 SignalP 程序 (Nielson 等, 1997, 蛋白质工程 10:1-6), 预测了一个 17 个残基的信号肽。因此, 成熟的内酯水解酶由 383 个氨基酸组成。

使用 LASERGENE™ MEGALIGN™ 软件 (DNASTAR, Inc., Madison, WI) 和 Clustal 方法 (Higgins, 1989, CABIOS 5:151-153), 采用同一性表和如下多序列对比参数进行内酯水解酶序列的比较性对比: 空位罚分为 10, 空位长度罚分为 10。两两对比参数是 Ktuple=1, 空位罚分=3, 窗口=5, 对角线=5。

比较性对比显示, *Fusarium venenatum* 的内酯水解酶与尖镰孢的内酯水解酶 (WO 97/10341) 具有 87% 的一致性。在 *Fusarium venenatum* 的内酯水解酶中有 5 个潜在的 N-连接的糖基化位点 (Asn-X-Ser/Thr), 然而这些位点中的一个含有一个内在的 Pro 残基, 因此不可能被利用。所有剩余的糖基化位点在尖镰孢的内酯水解酶中保守, 已知在第 28、106、179 和 277 位 Asn 残基上携带 3 种类型的 N-连接的高甘露糖碳水化合物 (WO 97/10341)。

实施例 7: pDM181 的构建

采用拼接重叠延伸技术构建质粒 pDM181 以将 1.2kb 尖镰孢胰蛋白酶启动子 (SP387) 与 1.1kb 尖镰孢胰蛋白酶终止子 (SP387) 融合在一起。作为重叠 PCR 策略的一部分, 将含有 *Swa*I、*Kpn*I 和 *Pac*I 限制性位点的多位点接头插入该启动子和终止子之间。在启动子的 5' 末端添加一个 *Xho*I

位点, 并保留原来的 EcoRI 位点。在终止子的 3' 末端, 通过 PCR 反应引入 EcoRI、HindIII 和 NsiI 位点。

使用如下引物从质粒 pJRoy20 (Royer 等, 1995, 生物技术 13:1479-1483) 产生含有尖镰孢胰蛋白酶启动子第 -1208 至 -1 位和 25 个碱基对多位点接头的 PCR 片段:

引物 1 (有义): 5'-GAGCTCGAGGAATTCTTACAAACCTTCAAC-3' (SEQ ID NO. 3)

XhoI EcoRI

引物 2 (反义): 5'-TTAATTAAGGTACCTGAATTTAAATGGTGAAGAGATAGATATCCAAG-3' (SEQ ID NO. 4)

PacI KpnI SwaI

100 μ l PCR 反应物含有 1 $\times$ Pwo 缓冲液 (Boehringer Mannheim, Indianapolis, IN), dATP、dCTP、dGTP 和 dTTP 各 200 μ M, 10ng pJRoy20 和 5 单位 Pwo DNA 聚合酶 (Boehringer Mannheim, Indianapolis, IN)。所用的 PCR 条件是 95 $^{\circ}$ C 3 分钟, 之后 25 个循环, 每个循环包括 95 $^{\circ}$ C 30 秒、50 $^{\circ}$ C 1 分钟和 72 $^{\circ}$ C 1 分钟。最后的延伸循环是 72 $^{\circ}$ C 5 分钟。

使用相同的 PCR 条件和如下引物, 从质粒 pJRoy20 产生含有尖镰孢胰蛋白酶启动子的第 -5 到 -1 位碱基对、25 个碱基对多位点接头和尖镰孢胰蛋白酶基因 3' 非翻译区的 1060 个碱基对 (终止子区) 的第二个 PCR 片段:

引物 3 (有义):

5'-TCACCATTAAATTCAGGTACCTTAATTAATTCCTTGTGGAAGCGTCGA-3' (SEQ ID NO. 5)

SwaI KpnI PacI

引物 4 (反义):

5'-TGGTATGCATAAGCTTGAATTCAGGTAAACAAGATATAATTT-3' (SEQ ID NO. 6)

NsiI HindIII EcoRI

使用 0.2 μ l 第一次 PCR (启动子) 反应产物和 3 μ l 第二次 (终止子) 反应产物作为模板, 以及引物 1 和 4, 获得含有尖镰孢胰蛋白酶启动子第 -1208 到 -1 位、25 个碱基对多位点接头和尖镰孢胰蛋白酶终止子的 1060

碱基对的最终 2.3kb 重叠 PCR 片段。所用 PCR 条件是 95℃ 3 分钟，之后 30 个循环，每个循环包括 95℃ 30 秒、62℃ 1 分钟和 72℃ 3 分钟。最后的延伸循环是 72℃ 5 分钟。该反应也使用 Pwo DNA 聚合酶。

将所获的含有胰蛋白酶启动子、多位点接头和胰蛋白酶终止子的 2.3kb 片段用 EcoRI 消化，并连入 EcoRI 消化的含有 bar 基因 (WO 97/26330) 的载体 pMT1612 中产生 pDM181 (图 2)。

实施例 8: 质粒 pSheB1 的构建

通过修饰 pDM181 产生 *Fusarium venenatum* 表达载体 pSheB1 (图 3)。修饰包括 (a) 去除 pDM181 序列中的两个 NcoI 位点，和 (b) 重新恢复尖镰孢胰蛋白酶启动子的天然翻译起始 (在 ATG 起始密码子处重新构建一个 NcoI 位点)。

使用 QuikChange™ 定点诱变试剂盒 (Stratagene Cloning Systems, La Jolla, CA)，根据厂商说明使用如下诱变引物对实现 pDM181 序列中两个 NcoI 位点的去除：

5'-dCAGTGAATTGGCCTCGATGGCCGCGGCCGCGAATT-3' (SEQ ID NO. 7) 和

5'-dAATTCGCGGCCGCGGCCATCGAGGCCAATTCAGT-3' (SEQ ID NO. 8)

5'-dCACGAAGGAAAGACGATGGCTTTCACGGTGTCTG-3' (SEQ ID NO. 9) 和

5'-dCAGACACCGTGAAAGCCATCGTCTTTCCTTCGTG-3' (SEQ ID NO. 10)

使用 QuikChange™ 定点诱变试剂盒以及如下诱变引物对也实现了尖镰孢胰蛋白酶启动子天然翻译起始的重新恢复：

5'-dCTATCTCTTCACCATGGTACCTTAATTAAATACCTTGTGGAAGCG-3' (SEQ ID NO. 11) 和

5'-dCGCTTCCAACAAGGTATTTAATTAAAGGTACCATGGTGAAGAGATAG-3' (SEQ ID NO. 12)

所有定点改变均通过合适载体区的 DNA 序列分析确证。

实施例 9: 表达载体 pTriggs1 的构建

使用如下程序构建内酯水解酶表达载体 pTriggs1 (图 4)。使用如下

引物对从克隆 FA0567 扩增内酯水解酶编码区:

5'-dCGGCATGCCTTCCACCATTTGCTG-3' (正向) (SEQ ID NO. 13)

5'-dTAAATTAAGTAGTCATATAACTTGGGTCC-3' (反向) (SEQ ID NO. 14).

正向引物在起始密码子处引入 SphI 限制性位点, 反向引物在终止密码子之后引入 PacI 位点。

扩增反应 (50μl) 含有如下成分: 0.8mg 克隆 FA0567 cDNA、40pmol 正向引物、40pmol 反向引物、dATP、dCTP、dGTP 和 dTTP 各 200μM、1 × Pwo DNA 聚合酶缓冲液和 2.5 单位 Pwo DNA 聚合酶。反应物在 Perkin-Elmer 480 型热循环仪中孵育 30 个循环, 每个循环包括 95℃ 3 分钟、58℃ 2 分钟和 72℃ 2 分钟。反应产物在 1.5% 琼脂糖凝胶 (Eastman Kodak, Rochester, NY) 上分离, 从胶上切下 1.2kb 产物条带, 并使用 Qiaex II (Qiagen, Chatsworth, CA) 根据厂商说明进行纯化。

扩增的内酯水解酶片段用 SphI 进行消化, 然后在 dNTP 存在时用 DNA 聚合酶 I (Klenow 片段; Boehringer Mannheim, Indianapolis, IN) 处理。该酶的 3→5' 外切酶活性去除了 SphI 粘性末端的 4 个核苷酸, 产生平端 DNA 片段。然后用 PacI 消化该 Klenow 处理片段, 并采用标准方法通过凝胶电泳纯化 (见 Sambrook 等, 1989, 见上)。

将该纯化 DNA 片段连在之前已按上述方法用 NcoI 切割、DNA 聚合酶 I (Klenow 片段) 处理、然后用 PacI 消化的载体 pSheB1 上。用 Klenow 片段处理 NcoI 消化的载体获得 NcoI 粘性末端的“补平”, 由此使其平端化并与 Klenow 处理的内酯水解酶 DNA 片段 SphI 位点相容。所获的表达质粒命名为 pTriggs1 (图 4)。

实施例 10: 内酯水解酶 cDNA 在 *Fusarium venenatum* 中的表达

通过用从添加了 2.5% 葡萄糖和 2.5mM 硝酸钠的 1 × Vogels 培养基平板 (2.5% 纯净琼脂) 获得的 10 个孢塞 (plug) 接种一个含有 500ml RA 孢子形成培养基的三角瓶中, 并在 28℃、150rpm 孵育 2-3 天, 产生 *Fusarium venenatum* CC1-3 (MLY-3) 的孢子。通过 Miracloth (Calbiochem, San Diego, CA) 收获孢子, 并在 Sorvall RC-5B 离心机 (E. I. DuPont De Nemours and

Co., Wilmington, DE) 内以 7000rpm 离心 20 分钟。用无菌蒸馏水洗涤沉淀孢子两次, 重悬在少量水中, 然后采用血细胞计数器计数。

通过将 4×10^7 个 *Fusarium venenatum* CC1-3 孢子接种在 100ml YEPG 培养基中, 于 24℃、150rpm 孵育 16 小时, 制备原生质体。培养物在 Sorvall RT 6000D (E. I. DuPont De Nemours and Co., Wilmington, DE) 中以 3500rpm 离心 7 分钟。用 30ml 1M MgSO_4 洗涤沉淀两次, 并重悬在 15ml 含有 5mg/ml NOVOZYME 234™ (batch PPM 4356, Novo Nordisk A/S, Bagsvaerd, 丹麦) 的 1M MgSO_4 中。培养物于 24℃、150rpm 孵育, 直到原生质体形成。向原生质体消化物中加入 35ml 体积的 2M 山梨糖醇, 该混合物以 2500rpm 离心 10 分钟。重悬沉淀, 用 STC 洗涤两次, 并以 2000rpm 离心 10 分钟沉淀原生质体。用血细胞计数器计数原生质体, 并将其重悬在 8:2:0.1 的 STC:SPTC:DMSO 溶液中至终浓度 1.25×10^7 个原生质体/ml。在 Nalgene Cryo 1℃ 冷冻箱 (VWR Scientific, Inc., San Francisco, CA) 中控制速率地进行冷冻后, 将该原生质体保存在 -80℃。

在冰上融解冷冻的 *Fusarium venenatum* CC1-3 原生质体。在 50ml 无菌聚丙烯管中加入 5 μg 实施例 9 中描述的 pTriggs1 和 5 μl 肝素 (每 ml STC 含 5mg)。加入 100 μl 原生质体, 轻柔混合, 并在冰上孵育 30 分钟。加入 1ml SPTC 并在室温孵育 20 分钟。加入 25ml 40℃ COVE 顶层琼脂糖后, 将混合物铺在空的 150mm 直径平板上, 并于室温孵育过夜。然后在平板之上倒含有 10mg BASTA™/ml 的另外 25ml 40℃ COVE 顶层琼脂糖, 并在室温孵育长达 14 天。除草剂 BASTA™ 中的活性成分是膦丝菌素。BASTA™ 从 AgrEvo (Hoechst Schering, Rodovre, 丹麦) 获得, 并在使用前用酚:氯仿:异戊醇 (25:24:1) 抽提两次, 用氯仿:异戊醇 (24:1) 抽提一次。

从筛选平板 (COVE 底层, COVE-BASTA™ 覆盖层) 直接挑取 3 个转化株至含有 25ml 添加了 1mM CaCl_2 和 100 $\mu\text{g}/\text{ml}$ 氯苄青霉素 (防止细菌污染) 的 M400Da 培养基的 125ml 摇瓶中, 并于 28℃、200rpm 在平台摇床上孵育 6 天。还包括未转化受体株作为阴性对照。

在第 6 天从摇瓶取样。通过离心除去细胞, 然后将每个 10 μl 上清

样品与等体积 SDS-PAGE 样品缓冲液 (Novex Experimental Technology, San Diego, CA) 加热至 95℃ 5 分钟。在 10-20% 剃度凝胶 (Novex Experimental Technology, San Diego, CA) 上分离变性的上清液蛋白, 并用考马斯兰染色。

SDS-PAGE 分析显示, 产生内酯水解酶的转化体分泌一种显著的表现分子量为大约 55kDa 的多肽。该种类的表现分子量 (大约 55,000) 相当大地高于成熟内酯水解酶的预测分子量 (大约 41,600), 提示该酶可能被极大程度地糖基化。观察到的分子量与 Shimizu 和 Kataoka (1996, 纽约科学院年报 (Annals of the New York Academy of Sciences) 799:650-658) 报道的尖镰孢内酯水解酶的亚基分子量 (60,000) 非常符合。根据报道, 活性内酯水解酶是二聚体 (Shimizu 和 Kataoka, 1996, 见上)。

实施例 11: 重组制备的内酯水解酶的纯化

按实施例 9 描述获得的摇瓶培养物通过 Miracloth 过滤, 然后冻存。融化的样品通过 0.45 μm 注射滤器过滤, 并稀释和调节 pH 至 2.4M 和 pH 7.5。将样品装载在 Q-Sepharose Big Beads 柱 (90ml) 上, 该柱已用含有 1mM CaCl_2 的 20mM Tris-HCl pH7.5 进行过预平衡。洗涤柱子至达到基线吸光度。用大于 6.67 柱床体积的在含有 1mM CaCl_2 的 20mM Tris-HCl pH7.5 中的 0-0.3M NaCl 线性梯度过柱。采用如下分析方法分析各部分。

在 140 μl 50mM 磷酸钾 pH7.0 以及溶于 50mM 磷酸钾 pH7.0 的 40 μl 100mM D-半乳糖醛酸 γ 内酯中加入 20 μl 体积的酶样品。反应物孵育 34 分钟。将 50 μl 该溶液样品和 50 μl 1:1 的 2N 羟胺-盐酸:3.5N NaOH 置于另一个孔中。然后加入 100 μl 1:1 的乙醇:溶于 4N HCl 的 10% w/v 三氯化铁。然后测量 520nm 处的吸光度。520nm 处的吸光度降低指示内酯底物浓度降低。

通过 SDS PAGE 确定稀释酶是均一的。

实施例 12: 泛解酸的内酯化

通过在 2ml 5N NaOH 中室温水解泛酰内酯 (130mg) 15 分钟制备泛解酸。通过加入约 1.5ml 6M HCl 将溶液 pH 调节至 5, 并加入水至终体积 10ml 以获得 100mM 终浓度泛解酸。

底物与培养物于 30℃ 孵育 30 分钟, 并通过下文的标准方法分析内酯形成。与 0.25M 乙酸钠 pH5 一起孵育获得比 20mM 磷酸钠 pH6.4 多四倍的内酯化。

生物材料的保藏

如下生物材料已根据布达佩斯条约保藏在农业研究机构保藏中心, 1815 University Street, Peoria, Illinois, 61640, 并给予了如下保藏号:

保藏物	保藏号	保藏日期
-----	-----	------

大肠杆菌 TOP10 (pFA0576)	NRRL B-30074	1998 年 10 月 27 日
----------------------	--------------	------------------

该菌株在如下条件下保藏, 即确保在本专利申请未决期间由专利和商标局局长依照 37 C.F.R. §1.14 和 35 U.S.C. §122 授权的人可获得该培养物。该保藏物代表该保藏菌株基本上纯的培养物。如果提出本主题申请的等同申请或其派生申请的国家该外国专利法要求, 可以获得该保藏物。但是, 应当理解, 可获得保藏物并不构成以损失政府法令所授予专利权利而实施本主题发明的许可。

本文描述和要求保护的本发明的范围并不被本文公开的具体实施方案限制, 因为这些实施方案旨在用作本发明几个方面的举例说明。任何等同的实施方案都包括在本发明的范围内。实际上, 从前面的描述来看, 本文给出和描述的之外的本发明各种修改将为本领域技术人员所明了。这些修改也包括在所附权利要求的范畴内。在冲突的情况下, 包括定义在内的本公开将优先。

本文引用了各种参考文献, 它们的公开以参考文献的方式完整地并入本文。

序列表

<110> Randy M. Berka
Michael W. Rey

<120> 具有内酯水解酶活性的多肽和编码该多肽的核酸

<130> 5721.204-WO

<140> 待定

<141> 1999-11-05

<150> 09/189,497

<151> 1998-11-10

<160> 14

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 1334

<212> DNA

<213> 镰孢霉属

<400> 1

ctccacacc	actcagtttc	acttctacct	cattcgccat	gccttccacc	attgctgccc	60
ttgttgctgg	gatttggtggc	gttgccgctg	ccaaacttcc	ctccacggct	caggctcattg	120
atcagaagtc	tttcaatggt	ttgaaggatg	tcccaccacc	ttccgtggct	aatgacacac	180
tggtctttac	atggcccggg	gtgacagagg	aatctctcgt	cgagaaaccc	ttccatgttt	240
atgatgatga	attcctcgat	gtcatcgga	aggatccctc	tctgacactt	gttgctacgt	300
cagaaagcga	ccccatcttc	cacgaggctg	tagtctggta	cccacctaca	gacgaggttt	360
ttttcgtgca	aaatgcgggt	gctcctgctg	caggcactgg	cctgaacaag	tcttccatca	420
tccagaagat	ttctctcaaa	gatgcagagg	ctttgcgcaa	gggaacccta	ggcaaggatg	480
aagtgaaggt	gacagtcggt	gacacagcta	accctcaagt	cattaacccc	aatggtggca	540
tttactacaa	gggcgaaatc	atctttgctg	gtgaaggcca	aggtgacgaa	gttccctcgg	600
ccctttaccg	catgaacccc	ttgcctccat	acaacacaag	caccctcctc	aacaactact	660
ttggccgcca	gttcaactcc	ttgaacgacg	ttggcatcaa	ccccaggaat	gggtgacttgt	720
acttcaccga	cactctctac	ggctatctcc	aagacttccg	tcctgtccct	gggtctgcgaa	780
accaagtgtg	ccgatacaac	ttcgacactg	gtgctgtaac	tggtgtcgct	gatgacttta	840
ctcttcccaa	cggtattggg	tttgtccctg	atggaaagcg	tgtctatgtc	accgacactg	900
gcctcgctct	tggcttctac	ggccgtaacc	tttctcccc	cgcctctgtt	tactctttcg	960
acgtgaacaa	ggatggtacc	cttgagaacc	gcaagacttt	tgcttacgta	gcttctttca	1020
tcccagacgg	tggtcatacc	gattccaagg	gtcgtgtcta	tgctgggtgt	ggtgacgggtg	1080
tccatgtctg	gaacccctct	ggcaagctca	ttggcaagat	ctataccggg	atcactgctg	1140
ccaacttcca	atgtgctgga	aaaggaagat	tgattatcac	tggtcagact	aagctgttct	1200
acgttaccct	ggctgcttca	ggacccaagt	tatatgacta	gaactccct	gtggcagtat	1260
agaaacagat	atgttaccgt	attgatagaa	gataatatta	attattaatc	gataaaaaaa	1320
aaaaaaaaaa	aaaa					1334

<210> 2

<211> 400

<212> PRT

<213> 镰孢霉属

<400> 2

Met	Pro	Ser	Thr	Ile	Ala	Ala	Leu	Val	Val	Gly	Ile	Cys	Gly	Val	Ala
1				5					10					15	
Ala	Ala	Lys	Leu	Pro	Ser	Thr	Ala	Gln	Val	Ile	Asp	Gln	Lys	Ser	Phe
			20					25					30		
Asn	Val	Leu	Lys	Asp	Val	Pro	Pro	Pro	Ser	Val	Ala	Asn	Asp	Thr	Leu
		35					40					45			
Val	Phe	Thr	Trp	Pro	Gly	Val	Thr	Glu	Glu	Ser	Leu	Val	Glu	Lys	Pro
	50				55						60				
Phe	His	Val	Tyr	Asp	Asp	Glu	Phe	Leu	Asp	Val	Ile	Gly	Lys	Asp	Pro
65				70					75					80	
Ser	Leu	Thr	Leu	Val	Ala	Thr	Ser	Glu	Ser	Asp	Pro	Ile	Phe	His	Glu
				85				90						95	
Ala	Val	Val	Trp	Tyr	Pro	Pro	Thr	Asp	Glu	Val	Phe	Phe	Val	Gln	Asn
			100					105					110		
Ala	Gly	Ala	Pro	Ala	Ala	Gly	Thr	Gly	Leu	Asn	Lys	Ser	Ser	Ile	Ile
		115					120					125			
Gln	Lys	Ile	Ser	Leu	Lys	Asp	Ala	Glu	Ala	Leu	Arg	Lys	Gly	Thr	Leu
	130					135					140				
Gly	Lys	Asp	Glu	Val	Lys	Val	Thr	Val	Val	Asp	Thr	Ala	Asn	Pro	Gln
145					150					155				160	
Val	Ile	Asn	Pro	Asn	Gly	Gly	Ile	Tyr	Tyr	Lys	Gly	Glu	Ile	Ile	Phe
				165				170						175	
Ala	Gly	Glu	Gly	Gln	Gly	Asp	Glu	Val	Pro	Ser	Ala	Leu	Tyr	Arg	Met
			180				185						190		
Asn	Pro	Leu	Pro	Pro	Tyr	Asn	Thr	Ser	Thr	Leu	Leu	Asn	Asn	Tyr	Phe
		195					200					205			
Gly	Arg	Gln	Phe	Asn	Ser	Leu	Asn	Asp	Val	Gly	Ile	Asn	Pro	Arg	Asn
	210					215					220				
Gly	Asp	Leu	Tyr	Phe	Thr	Asp	Thr	Leu	Tyr	Gly	Tyr	Leu	Gln	Asp	Phe
225					230					235				240	
Arg	Pro	Val	Pro	Gly	Leu	Arg	Asn	Gln	Val	Tyr	Arg	Tyr	Asn	Phe	Asp
				245				250						255	
Thr	Gly	Ala	Val	Thr	Val	Val	Ala	Asp	Asp	Phe	Thr	Leu	Pro	Asn	Gly
			260					265					270		
Ile	Gly	Phe	Ala	Pro	Asp	Gly	Lys	Arg	Val	Tyr	Val	Thr	Asp	Thr	Gly
	275					280						285			
Ile	Ala	Leu	Gly	Phe	Tyr	Gly	Arg	Asn	Leu	Ser	Ser	Pro	Ala	Ser	Val
	290					295					300				
Tyr	Ser	Phe	Asp	Val	Asn	Lys	Asp	Gly	Thr	Leu	Glu	Asn	Arg	Lys	Thr
305					310					315				320	
Phe	Ala	Tyr	Val	Ala	Ser	Phe	Ile	Pro	Asp	Gly	Val	His	Thr	Asp	Ser
				325				330						335	
Lys	Gly	Arg	Val	Tyr	Ala	Gly	Cys	Gly	Asp	Gly	Val	His	Val	Trp	Asn
			340					345				350			
Pro	Ser	Gly	Lys	Leu	Ile	Gly	Lys	Ile	Tyr	Thr	Gly	Ile	Thr	Ala	Ala
		355					360					365			
Asn	Phe	Gln	Phe	Ala	Gly	Lys	Gly	Arg	Leu	Ile	Ile	Thr	Gly	Gln	Thr
	370					375					380				
Lys	Leu	Phe	Tyr	Val	Thr	Leu	Ala	Ala	Ser	Gly	Pro	Lys	Leu	Tyr	Asp
385					390					395					400

<210> 3
 <211> 30
 <212> DNA

<213> 镰孢霉属
 <400> 3
 gagctcgagg aattccttaca aaccttcaac 30
 <210> 4
 <211> 47
 <212> DNA
 <213> 镰孢霉属
 <400> 4
 ttaattaagg tacctgaatt taaatgggtga agagatagat atccaag 47
 <210> 5
 <211> 51
 <212> DNA
 <213> 镰孢霉属
 <400> 5
 tcaccattta aattcaggtga ccttaattaa attccttggt ggaagcgctcg a 51
 <210> 6
 <211> 42
 <212> DNA
 <213> 镰孢霉属
 <400> 6
 tggatgcat aagcttgaat tcaggtaaac aagatataat tt 42
 <210> 7
 <211> 35
 <212> DNA
 <213> 镰孢霉属
 <400> 7
 cagtgaattg gcctcgatgg ccgcgggccgc gaatt 35
 <210> 8
 <211> 35
 <212> DNA
 <213> 镰孢霉属
 <400> 8
 aattcgcggc cgcgccatc gaggccaatt cactg 35
 <210> 9
 <211> 34
 <212> DNA
 <213> 镰孢霉属
 <400> 9
 cacgaaggaa agacgatggc tttcacggtg tctg 34
 <210> 10
 <211> 34
 <212> DNA

<213> 镰孢霉属

<400> 10

cagacaccgt gaaagccatc gtctttcctt cgtg

34

<210> 11

<211> 46

<212> DNA

<213> 镰孢霉属

<400> 11

ctatctcttc accatgggtac cttaattaaa taccttggtg gaagcg

46

<210> 12

<211> 46

<212> DNA

<213> 镰孢霉属

<400> 12

cgtttccaac aaggtattta attaaggtac catgggtgaag agatag

46

<210> 13

<211> 23

<212> DNA

<213> 镰孢霉属

<400> 13

cggcatgcct tccaccattg ctg

23

<210> 14

<211> 29

<212> DNA

<213> 镰孢霉属

<400> 14

ttaattaact agtcatataa cttgggtcc

29

说明书附图

CTCCACACCACTCAGTTTCACTTCTACCTCATTCGCCCATCGCCATTGCCACCATTTGTTGTCGGGATTGTGGC 80
M P S T I A A L V V G I C G

GTTGCGCTGCCAAACTTCCCTCCACGGCTCAGGTCAATGATCAGAAAGTCTTTCAATGTTTGAAGGATGTCACCAACC 160
V A A A K L P S T A Q V I D Q K S F N V L K D V P P P

TTCCGTGGCTAATGACACACTGGTCTTTACATGGCCCGAGTGACAGAGGAATCTCTCGTCGAGAAACCCCTTCCATGTTT 240
S V A N D T L V F T W P G V T E E S L V E K P F H V

ATGATGATGAATTCTCTGATGTCTATCGGAAAGGATCCCTCTCTGACACTTGTGTCTACGTACAGAAAGCGACCCCATCTTC 320
Y D D E F L D V I G K D P S L T L V A T S E S D P I F

CACGAGGCTGTAGTCTGGTACCCACCTACAGACGAGGTTTCTTCGTGCAAAATGCGGGTGTCTCCTGCGGCAGGCACCTGG 400
H E A V V W Y P P T D E V F F V Q N A G A P A A G T G

CCTGAACAAGTCTTCCATCATCCAGAAAGATTCTCTCAAAAGATGCAGAGGCTTTCGCGCAAGGAAACCTAGGCAAGGATG 480
L N K S S I I Q K I S L K D A E A L R K G T L G K D

AAGTGAAGGTGACAGTCTGTTGACACAGCTAACCTCAAGTCATTAAACCCCAATGGTGGCATTACTACAAGGGCGAAATC 560
E V K V T V V D T A N P Q V I N P N G G I Y Y K G E I

ATCTTTGCTGTGAAGGCCAAGGTGACGAAGTTCCCTCGGCCCTTTACCGCATGAACCCCTTGCTCCATACACACAAG 640
I F A G E G Q G D E V P S A L Y R M N P L P P Y N T S

CACCCCTCTCAACAACACTACTTGGCCGCCAGTTTCAACTCTTGAACGACGTGGCATCAACCCAGGAATGGTGACTTGT 720
T L L N N Y F G R Q F N S L N D V G I N P R N G D L

ACTTACCGACACTCTACGGCTATCTCCAAGACTTCCGTCTGTCCTGTCGCGAAACCAAGTGTACCGATACAAC 800
Y F T D T L Y G Y L Q D F R P V P G L R N Q V Y R Y N

TTGACACTGGTGTAACTGTGTGCTGATGACTTTACTCTTCCCAACGGTATTGTTTGTCTCTCTGATGGAAGCG 880
F D T G A V T V V A D D F T L P N G I G F A P D G K R

图 1A

TGCTATGTCA	CGGACACTG	GCATCGCTCT	TGGCTTCTAC	GGCCGTAACCT	TTCTCCCCCGCCT	CTGTTTACTCT	TTTCG	960
V Y V T D T	G I A L G F	Y G R N L S	S P A S V Y	S F				
ACGTGAACA	AGGATGGT	ACCTTGAGAA	CCGCAAGACT	TTTGCCCTAC	GTAGCTTCTTT	CATCCAGACG	GTGTTCA	TACC 1040
D V N K D G	T L E N R K	T F A Y V A	S F I P D G	V H T				
GATTCGAAG	GGTGTCTAT	GCTGGTTGTG	TGACGGTGT	CCATGTC	TGGAAACCCCT	CTGGCAAGCT	CATTGGCA	AGAT 1120
D S K G R V	Y A G C G D	G V H V W N	P S G K L I	G K I				
CTATACCGG	GATCACTG	CGCAACTTCC	AATTTGCTG	GAAAGGAAG	ATTGATTATCA	CTGTCAGACT	AAAGCTGT	CT 1200
Y T G I T A	A N F Q F A	G K G R L I	I T G Q T K	L F				
ACGTTACCC	TGGCTTCAG	GACCCAAAGT	TATATGACT	AGAACTCC	CTGTGGCAGT	ATAGAAA	CAGATATTT	TACCGT 1280
Y V T L A A	S G P K L Y	D						
ATTGATAGA	AGATAATAT	TAAATTATTA	ATCGATAAAA	AAAAAAAAAAAA	AAAAA	1334		

- 2 -

图 1B

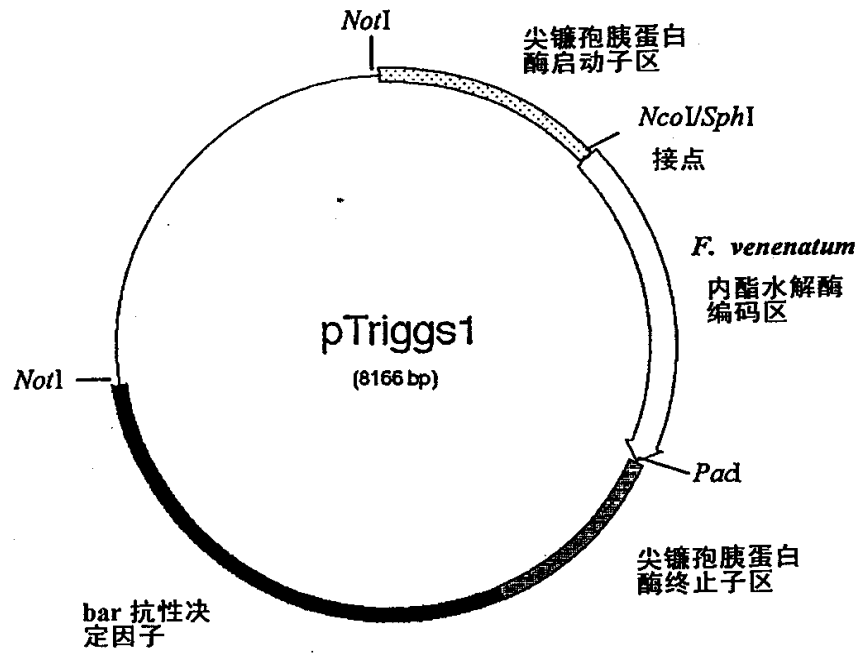


图 2

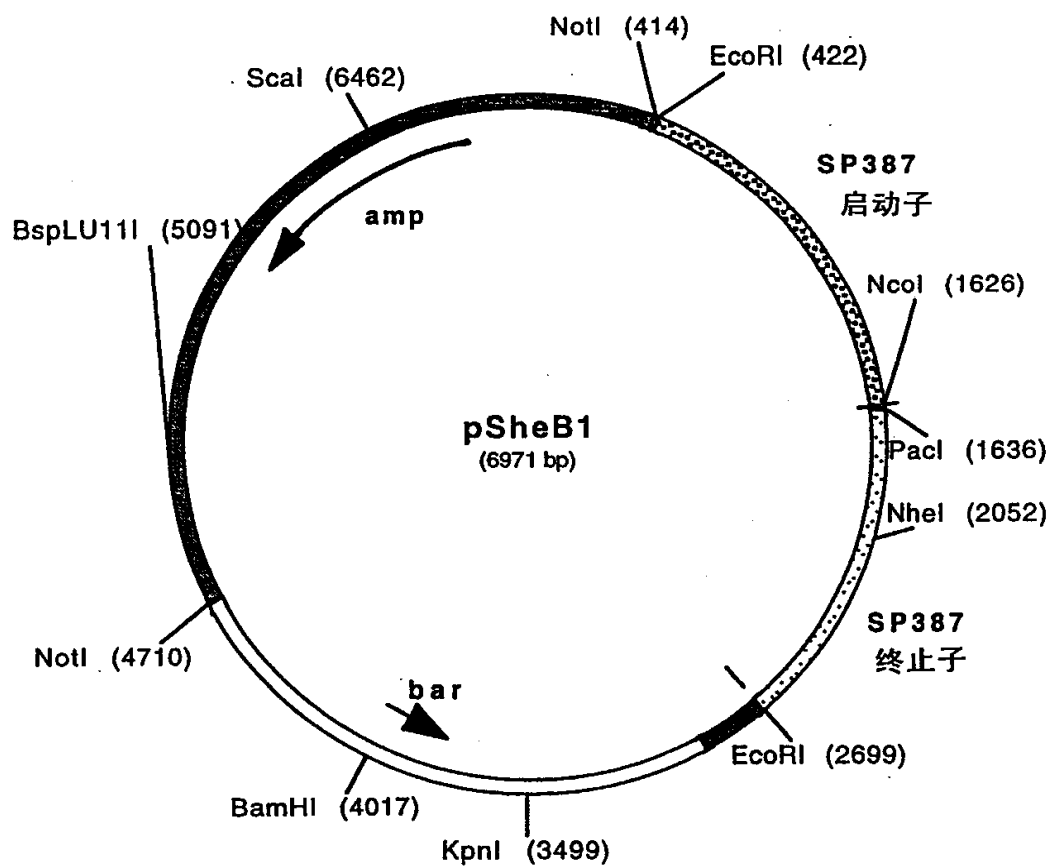


图 3

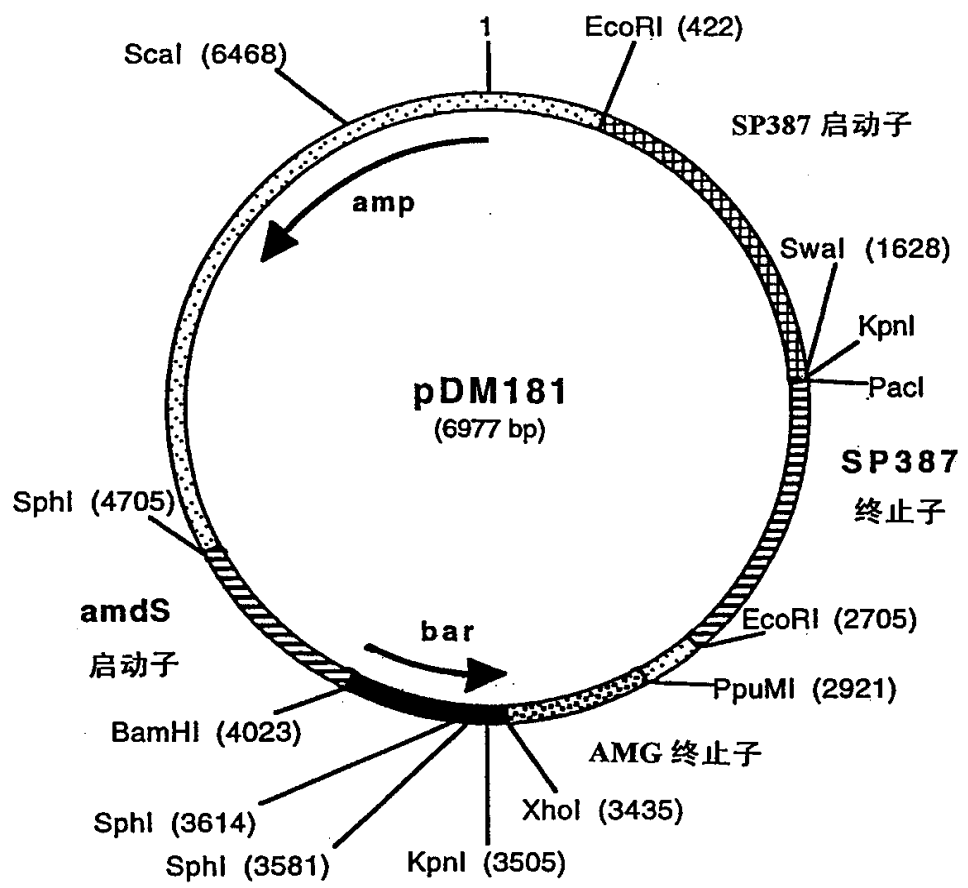


图 4