

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.  
C12N 9/48 (2006.01)



# [12] 发明专利说明书

专利号 ZL 98812237.5

[45] 授权公告日 2006年3月29日

[11] 授权公告号 CN 1247777C

[22] 申请日 1998.11.13 [21] 申请号 98812237.5

[30] 优先权

[32] 1997.12.16 [33] DK [31] 1465/97

[32] 1997.12.16 [33] US [31] 60/069,719

[32] 1998.5.15 [33] DK [31] PA199800670

[86] 国际申请 PCT/DK1998/000495 1998.11.13

[87] 国际公布 WO1999/031226 英 1999.6.24

[85] 进入国家阶段日期 2000.6.15

[71] 专利权人 诺沃奇梅兹有限公司

地址 丹麦鲍斯韦

共同专利权人 诺沃奇梅兹生物技术有限公司

日本烟草产业株式会社

[72] 发明人 A·布林克夫斯基 T·S·伯温

A·V·克劳兹 A·斯洛马

K·布朗 M·汤 藤井幹夫

丸本千草 L·V·考弗德

审查员 辜学英

[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 48 页 附图 13 页

[54] 发明名称

具有氨肽酶活性的多肽及其编码核酸

[57] 摘要

本发明涉及具有氨肽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构建体、载体和宿主细胞以及生产和应用这些多肽的方法。

1. 一种具有氨肽酶活性的分离的多肽，其选自下组：
  - (a) 有与 SEQ ID NO: 2 的第 33 至 673 位氨基酸至少 70% 相同的氨基酸序列的多肽；
  - (b) 由这样一个核酸序列编码的多肽，该核酸序列在中度严谨条件下与 SEQ ID NO: 1 的第 670 - 2592 位核苷酸所示核酸序列或者其互补链可杂交，其中中度严谨条件定义如下：在  $5 \times$  SSPE、0.3% SDS、200  $\mu$ g/ml 经剪切和变性的鲑鱼精 DNA 和 35% 甲酰胺中于 42 $^{\circ}$ C 预杂交和杂交，在  $2 \times$  SSC 和 0.2% SDS 中于 55 $^{\circ}$ C 洗涤；和
  - (c) (a) 或 (b) 的酶促活性片段。
2. 权利要求 1 的多肽，其具有与 SEQ ID NO: 2 的第 33 至 673 位氨基酸至少 70% 相同的氨基酸序列。
3. 权利要求 2 的多肽，其具有与 SEQ ID NO: 2 的第 33 至 673 位氨基酸至少 75% 相同的氨基酸序列。
4. 权利要求 3 的多肽，其具有与 SEQ ID NO: 2 的第 33 至 673 位氨基酸至少 80% 相同的氨基酸序列。
5. 权利要求 4 的多肽，其具有与 SEQ ID NO: 2 的第 33 至 673 位氨基酸至少 85% 相同的氨基酸序列。
6. 权利要求 5 的多肽，其具有与 SEQ ID NO: 2 的第 33 至 673 位氨基酸至少 90% 相同的氨基酸序列。
7. 权利要求 6 的多肽，其具有与 SEQ ID NO: 2 的第 33 至 673 位氨基酸至少 95% 相同的氨基酸序列。
8. 权利要求 1 的多肽，其中包含 SEQ ID NO: 2 的氨基酸序列。
9. 权利要求 1 的多肽，其由 SEQ ID NO: 2 的氨基酸序列或其酶促活性片段组成，所述片段由 SEQ ID NO: 2 的第 33 至 673 位氨基酸组成。
10. 权利要求 9 的多肽，其由 SEQ ID NO: 2 中第 33 - 673 位氨基酸所示序列组成。
11. 权利要求 1 的多肽，其由在中度严谨条件下可与 SEQ ID NO: 1

中第 670 - 2592 位核苷酸所示序列或其互补链相杂交的核酸序列所编码, 其中中度严谨条件定义如下: 在  $5 \times$  SSPE、0.3%SDS、200 $\mu$ g/ml 经剪切和变性的鲑鱼精 DNA 和 35% 甲酰胺中于 42 $^{\circ}$ C 预杂交和杂交, 在  $2 \times$  SSC 和 0.2% SDS 中于 55 $^{\circ}$ C 洗涤。

12. 权利要求 1 的多肽, 其由在高度严谨条件下可与 SEQ ID NO: 1 中第 670 - 2592 位核苷酸所示序列或其互补链相杂交的核酸序列所编码, 其中高度严谨条件定义如下: 在  $5 \times$  SSPE、0.3%SDS、200 $\mu$ g/ml 经剪切和变性的鲑鱼精 DNA 和 50% 甲酰胺中于 42 $^{\circ}$ C 预杂交和杂交, 在  $2 \times$  SSC 和 0.2% SDS 中于 65 $^{\circ}$ C 洗涤。

13. 权利要求 1 的多肽, 其 (a) 在 37 $^{\circ}$ C 测得, 在 pH 5.0-8.5 的范围内具有氨肽酶活性, (b) 等电点在 7.4-8.5 的范围内; (c) pH7.5 时用 Gly-pNA 的 Tris-HCl 缓冲系统测得, 在 20 - 55 $^{\circ}$ C 的温度范围内具有氨肽酶活性; (d) 可水解 Ala-pNA, Gly-pNA, Leu-pNA, Glu-pNA, Asp-pNA, Lys-pNA, Ile-pNA 和 Val-pNA; (e) 不能水解 Phe-pNA 和 Pro-pNA; (f) 不受苯甲磺酰氟抑制, 和 (g) 由 SDS-PAGE 测得分子量为 67.5 kDa。

14. 权利要求 1 - 13 中任一项的多肽, 其来自鞘氨醇单胞菌属 (*Sphingomonas*) 菌株。

15. 权利要求 14 的多肽, 其来自荚膜鞘氨醇单胞菌 (*Sphingomonas capsulata*) 菌株。

16. 权利要求 1 的多肽, 其由包含在大肠杆菌 NRRL B-30032 内的质粒 pMRT004.1-14 中所含有的核酸序列编码。

17. 一种分离的核酸, 其中包含编码权利要求 1-16 中任一项所述多肽的核酸序列。

18. 一种分离核酸, 其中包含在 SEQ ID NO: 1 成熟多肽编码核酸序列中有至少一个突变的核酸序列, 所述突变核酸序列编码由 SEQ ID NO: 2 的第 33 至 673 位氨基酸所组成的多肽。

19. 一种分离的核酸, 其中包含编码具有氨肽酶活性的多肽的核酸序列, 所述核酸序列通过如下获得:

(a) 在低度严谨条件下, 将 DNA 与 SEQ ID NO: 1 所示序列或其互补

链杂交；和(b)分离所述核酸序列。

20. 权利要求 19 的分离的核酸，其中步骤(a)中的杂交是在中度严谨条件下进行的。

21. 权利要求 19 的分离的核酸，其中步骤(a)中的杂交是在高度严谨条件下进行的。

22. 一种核酸构建体，其中包含权利要求 17-21 中任一项的核酸，以及与之可操作连接、可指导多肽在适当表达宿主中生产的一或多个控制序列。

23. 一种重组表达载体，其中包含权利要求 22 的核酸构建体。

24. 一种重组宿主细胞，该重组宿主细胞中包含权利要求 22 的核酸构建体。

25. 产生突变核酸的方法，包括(a)向 SEQ ID NO: 1 的成熟多肽编码序列中导入至少一个突变，其中该突变核酸编码由 SEQ ID NO: 2 中第 33 至 673 位氨基酸组成的多肽；和(b)回收该突变核酸。

26. 根据权利要求 25 的方法产生的突变核酸。

27. 一种生产具有氨肽酶活性的多肽的方法，包括(a)在适合于产生所述多肽的条件下培养包含权利要求 26 的编码多肽的突变核酸的宿主细胞；和(b)回收这种多肽。

28. 一种生产权利要求 1-16 中任一项的多肽的方法，包括(a)在适合于产生所述多肽的条件下培养野生型状态下能生成所述多肽的菌株；和(b)回收该多肽。

29. 由权利要求 28 的方法生产的多肽。

30. 生产权利要求 1-16 中任一项的多肽的方法，包括(a)在适于产生该多肽的条件下，培养含有核酸构建体的宿主细胞，该核酸构建体包含编码该多肽的核酸序列；和(b)回收该多肽。

31. 由权利要求 30 的方法生产的多肽。

32. 生产多肽的方法，包括(a)在有利于产生该多肽的条件下培养宿主细胞，其中该宿主细胞包含 SEQ ID NO: 1 的突变核酸序列，该突变核酸序列在 SEQ ID NO: 1 的成熟多肽编码序列中带有至少一个突变，并编

码由 SEQ ID NO: 2 中第 33 至 673 位氨基酸组成的多肽; 和 (b) 回收该多肽。

33. 由权利要求 32 的方法生产的多肽。

34. 生产亲代细胞的突变体的方法, 包括对编码权利要求 1-16 中任一项的多肽的核酸序列或其控制序列进行灭活, 从而得到生产该多肽的水平比其亲代细胞少的突变细胞。

35. 由权利要求 34 的方法生产的突变细胞。

36. 权利要求 33 的突变细胞, 其中进一步包括编码异源蛋白质的核酸序列。

37. 生产异源多肽的方法, 包括 (a) 在有利于产生该多肽的条件下培养权利要求 36 的突变细胞; 和 (b) 回收该多肽。

## 具有氨肽酶活性的多肽及其编码核酸

### 发明背景

#### 发明领域

本发明涉及具有氨肽酶活性的分离的多肽和编码这些多肽的分离的核酸序列。本发明还涉及包含所述核酸序列的核酸构建体、载体和宿主细胞以及所述多肽的制备和应用方法。

#### 相关技术的描述

许多食品，例如汤、调味汁和佐料都含有通过水解蛋白类材料而获得的提味剂。传统上，这种水解过程是采用浓盐酸来进行，随后用氢氧化钠进行中和。但是，这种化学水解方法导致水解过程中得到的氨基酸严重降解，并且该化学反应过程中会形成有害的副产品。对化学法水解得到的提味剂的日益担忧引发了酶法水解工艺的开发。

蛋白类原料的酶法水解过程目的在于达到较高的水解度（DH），这一点通常是通过使用非特异性作用的蛋白水解酶复合体（即非特异性作用的内肽酶和外肽酶）实现的。例如，W094/25580 描述了一种使用得自米曲霉（*Aspergillus oryzae*）的非特异性作用的酶制剂来水解蛋白质的方法。目前还没有特异性作用的蛋白水解酶用于这一目的，因为这样的酶只能产生不完全水解。

具有氨肽酶活性的多肽能催化从肽、多肽和蛋白质的 N-端切下 1 或更多个氨基酸残基。这类多肽被分类在国际生物化学和分子生物学会制定的酶学分类编号 E. C. 3. 4. 11. - 中。

W096/28542 公开了一种分子量为 35kDa 的氨肽酶。JP-7-5034631 (Noda) 公开了一种得自 yellow koji 霉菌（包括米曲霉）的亮氨酸氨肽酶。JP-7-4021798 (Zaidan Hojin Noda Sangyo) 公开了通过添加亮氨酸氨肽酶 II 来制备日本豆酱的方法，其中亮氨酸氨肽酶 II 是通过培养多个菌株，包括米曲霉 460 (ATCC 20386) 和 IAM 2616 制成的。已知

米曲霉 460 能产生多种亮氨酸氨肽酶，其中的三种通过凝胶过滤测得分子量为 26.5、56 和 61kDa（分别见 Nakada 等，1972，农业和生物化学 (Agricultural and Biological Chemistry) 37: 757-765; Nakada 等，1972，农业和生物化学 37: 767-774; 和 Nakada 等，1972，农业和生物化学 37: 775-782)。柑橘青霉 (*Penicillium citrium*) 能产生一种胞内亮氨酸氨肽酶，通过 SDS-PAGE 测得分子量为 65kDa (Kwon 等，1996，工业微生物学杂志 17:30-35)。

W097/04108 (Roehm) 公开了编码酱油曲霉 (*Aspergillus sojiae*) 亮氨酸氨肽酶的 DNA。Chang 和 Smith(1989，生物学化学杂志 264:6979-6983) 公开了得自酿酒酵母 (*Saccharomyces cerevisiae*) 的液泡氨肽酶编码基因的分子克隆和测序。Chang 等(1992，生物学化学杂志 267:8007-8011) 公开了得自酿酒酵母的甲硫氨酸氨肽酶编码基因的分子克隆和测序。

通常需要使用有肽酶活性的混合体系来制备具有良好感官品质和较高水解度的蛋白质水解产物。人们很希望能提供一种单一组分的肽酶，可利用其活性，将它单独或与其他酶结合使用以改善食品中所用蛋白质水解产物的感官品质和水解度。

本发明的一个目的是提供具有氨肽酶活性的改善的多肽，和这些多肽的编码核酸。

### 发明概述

本发明涉及具有氨肽酶活性的分离的多肽，它们选自：

(a) 有与 SEQ ID NO: 2 的第 33 至 673 位氨基酸至少 50% 相同的氨基酸序列的多肽；

(b) 由这样一个核酸序列编码的多肽，该核酸序列在低度严谨条件下与 (i) SEQ ID NO: 1 的核酸序列，(ii) 其互补链，或其至少含 100 个核苷酸的亚序列可杂交；

(c) (a) 或 (b) 的等位变体；

(d) (a)、(b) 或 (c) 的片段，该片段具有氨肽酶活性；和

(e) 具有如下特性的多肽：(i) 在 37℃ 测得，在 pH 5.0-8.5 的范围内具有氨肽酶活性，(ii) 等电点在 7.4-8.5 的范围内；(iii) pH 7.5 时用

Gly-pNA 的 Tris-HCl 缓冲系统测得, 在 20 - 55°C 的温度范围内具有氨肽酶活性; (iv) 可水解 Ala-pNA, Gly-pNA, Leu-pNA, Glu-pNA, Asp-pNA, Lys-pNA, Ile-pNA 和 Val-pNA; (v) 不能水解 Phe-pNA 和 Pro-pNA; (vi) 不受苯甲磺酰氟(phenylmethanesulfonyl fluoride)抑制, 受 EDTA、二异丙基氟磷酸(di-isopropyl fluoro phosphate)、对氯汞苯甲酸(p-chloromercuribenzoic acid)和碘醋酸(iodoacetic acid)的轻度抑制, 受邻菲咯啉(o-phenanthroline)的彻底抑制, 和/或(vii)得自鞘氨醇单胞菌属的菌株, 分子量为 67 至 68 kDa。

本发明还涉及编码所述多肽的分离的核酸序列, 和包含该核酸序列的核酸构建体、载体和宿主细胞以及制备和应用所述多肽的方法。

#### 附图简述

图 1 显示在室温下用含 100 mg/ml Ala-pNA 的 DMSO 溶液测得的荚膜鞘氨醇单胞菌的氨肽酶活性随 pH 的变化。

图 2 显示在 50 mM 磷酸钠缓冲液 pH7.5 中, 荚膜鞘氨醇单胞菌的氨肽酶活性随温度的变化。

图 3 显示在 50 mM 磷酸钠缓冲液 pH7.5 中, 荚膜鞘氨醇单胞菌的氨肽酶活性随温度的变化。

图 4 显示分别在低剂量和高剂量 Flavourzyme™ 背景下, 加入荚膜鞘氨醇单胞菌的酶粗提物和荚膜鞘氨醇单胞菌氨肽酶对 DH 的影响。

图 5 显示分别在低剂量和高剂量 Flavourzyme™ 背景下, 加入荚膜鞘氨醇单胞菌的酶粗提物和荚膜鞘氨醇单胞菌的氨肽酶对 DH 的影响。

图 6 显示 Flavourzyme™、Alcalase® 和荚膜鞘氨醇单胞菌氨肽酶的不同组合所得的 DH。

图 7 显示 Flavourzyme™、Alcalase® 和荚膜鞘氨醇单胞菌氨肽酶的不同组合所得的 DH。

图 8 显示 pSJ1678 的限制性图谱。

图 9 显示荚膜鞘氨醇单胞菌 IF0 12533 的氨肽酶的核酸序列及推定的氨基酸序列(分别为 SEQ ID NO: 1 和 2)。

图 10 显示 pMRT014-1 的限制性图谱。



## 发明详述

### 具有氨肽酶活性的多肽

术语“氨肽酶活性”在文中定义为能催化从肽、寡肽或蛋白质的 N-端切下氨基酸的肽酶活性。一般来说，氨肽酶活性能从肽、多肽或蛋白质的 N-端切除氨基酸 X，其中 X 可以代表选自：丙氨酸、精氨酸、天冬酰胺、天冬氨酸、半胱氨酸、谷氨酰胺、谷氨酸、甘氨酸、组氨酸、异亮氨酸、亮氨酸、赖氨酸、甲硫氨酸、苯丙氨酸、脯氨酸、丝氨酸、苏氨酸、色氨酸、酪氨酸和缬氨酸的任何氨基酸残基，但至少是丙氨酸、甘氨酸、亮氨酸、谷氨酸、天冬氨酸、赖氨酸、异亮氨酸和/或缬氨酸。应当理解，本发明所述具有氨肽酶活性的分离的多肽对要裂解的肽、多肽或蛋白质的氨基酸序列可以是非特异性的。为了本发明的目的，氨肽酶活性通过测量 405 nm 处对丙氨酰对硝基苯胺的起始水解速率而确定。

在第一个实施方案中，本发明涉及具有氨肽酶活性的分离的多肽，其氨基酸序列与 SEQ ID NO: 2 中第 33 至 673 位氨基酸（即成熟多肽）的相同程度至少是大约 50%、优选至少是大约 55%、优选至少是大约 60%，优选至少是大约 65%，优选至少是大约 70%，更优选至少是大约 80%，更优选至少是大约 90%，最优选至少是大约 95%，最最优选的是至少大约 97%（此后称为“同源多肽”）。在一个优选实施方案中，该同源多肽具有的氨基酸序列与 SEQ ID NO: 2 氨基酸序列有 5 个氨基酸不同，优选 4 个氨基酸不同，更优选 3 个不同，还要优选的是 2 个不同，最优选 1 个氨基酸不同。为了本发明的目的，两个或更多氨基酸序列之间的相同程度通过 BLAST 2.0 蛋白质数据库查询程序（Altschul 等，1997，核酸研究 25: 3389 - 3402）并采用以下参数确定：blastall -p blastp -a 4 -e 10 -E 0 -v 500 -b 250 -I [查询文档] -d prot\_all，其中 -p 指程序名称，-a 指将要用到的服务器数，-e 指期望值，-E 指延伸缺口的代价，-v 指单线描述（one-line description）数，-b 指将要显示的比对数，-I 指查询文档，-d 指用于查询的数据库。

优选地，本发明的多肽包含 SEQ ID NO: 2 的氨基酸序列或其等位变体；或其有氨肽酶活性的片段。在一个更优选的实施方案中，本发明所

述多肽包含 SEQ ID NO: 2 的氨基酸序列。在另一个优选实施方案中, 本发明的多肽包含 SEQ ID NO: 2 中第 33 至 673 位氨基酸 (即 SEQ ID NO: 2 的成熟多肽), 或其等位变体; 或其有氨肽酶活性的片段。在另一个优选实施方案中, 本发明的多肽包含 SEQ ID NO: 2 中第 33 至 673 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; 或其具有氨肽酶活性的片段组成。在另一个优选的实施方案中, 本发明的多肽由 SEQ ID NO: 2 的氨基酸序列组成。在另一个优选实施方案中, 多肽由 SEQ ID NO: 2 中第 33 至 673 位氨基酸或其等位变体; 或其有氨肽酶活性的片段组成。在另一个优选实施方案中, 多肽由 SEQ ID NO: 2 中第 38 至 654 位氨基酸组成。

SEQ ID NO: 2 的片段是在该氨基酸序列的氨基末端和/或羧基末端缺失了一或多个氨基酸的多肽。优选地, 片段含有至少 524 个氨基酸残基, 更优选至少 574 个氨基酸残基, 最优选有至少 624 个氨基酸残基。

等位变体表示占据相同染色体位点的一个基因的两种或多种可替换形式中任意一种。等位基因变异通过突变自然产生, 并可能导致种群内的多态现象。基因突变可以是沉默的 (编码的多肽没有变化) 或者可以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所编码的多肽。

同源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不同之处可能在于插入或缺失 1 或多个氨基酸残基和/或有 1 或多个氨基酸残基被不同氨基酸残基取代。优选地, 氨基酸改变是性质改变较小的变化, 即是不会显著影响蛋白质的折叠和/或活性的保守性氨基酸取代; 小片段缺失, 通常是 1 到大约 30 个氨基酸的缺失; 小的氨基或羧基末端延伸, 如氨基端添加的甲硫氨酸残基; 有多达大约 20-25 个残基的小连接肽; 或可通过改变净电荷或者其它功能而有助于纯化的小延伸如多聚组氨酸片段、抗原表位或结合区。

保守性取代的例子是在碱性氨基酸 (精氨酸、赖氨酸和组氨酸)、酸性氨基酸 (谷氨酸和天冬氨酸)、极性氨基酸 (谷氨酰胺和天冬酰胺)、疏水氨基酸 (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)内进行的取代。通常不会改变特异活性的氨基酸取代是本领域已知的,并且由例如 H. Neurath 和 R. L. Hill, 1979, 在《蛋白质》一书, Academic Press, New York 中描述过。最常见的替换是 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 以及反向进行的替换。

在第二个实施方案中,本发明涉及具有氨肽酶活性的分离的多肽,其中编码它们的核酸序列在低度严谨条件下,更优选在中度严谨条件下,更优选在高度严谨条件下,最优选在极高度严谨条件下与核酸探针(所述探针在相同的条件下,能与 SEQ ID NO: 1 的核酸序列或其互补链;或其编码氨肽酶活性多肽片段的亚序列杂交)能够杂交(J. Sambrook, E. F. Fritsch, 和 T. Maniatus, 1989, 分子克隆: 实验指南, 第 2 版, Cold Spring Harbor, New York); 或这些多肽的等位变体和有氨肽酶活性的片段。

可以根据本领域熟知的方法,利用 SEQ ID NO: 1 的核酸序列或其亚序列、以及 SEQ ID NO: 2 的氨基酸序列或其片段来设计核酸探针,以便从不同属或种的株系中鉴定和克隆到编码有氨肽酶活性的多肽的 DNA。具体来说,这些探针可用来与所研究的属或种的基因组或 cDNA 杂交,随后进行标准 Southern 印迹操作,以便鉴定和分离其中的相应基因。这些探针可以比完整序列短得多,但应至少有 15 个、优选至少 25 个,更优选至少 35 个核苷酸长。也可以使用更长的探针。DNA 探针和 RNA 探针都可使用。通常对探针进行标记(例如,用  $^{32}\text{P}$ 、 $^3\text{H}$ 、 $^{35}\text{S}$ 、生物素或抗生物素蛋白)以检测相应基因。这些探针均包含在本发明中。

因此,可以从得自这样的其他生物的基因组 DNA 或 cDNA 文库筛选能与上述探针杂交、并编码有氨肽酶活性的多肽的 DNA。可以通过琼脂糖或聚丙烯酰胺凝胶电泳,或者其他分离技术来分离来自其他生物的基因组 DNA 或其他 DNA。可以将来自文库的 DNA 或分离的 DNA 转移并固定到硝酸纤维素膜或其他合适的载体材料上。为了鉴定出与 SEQ ID NO: 1 同源的

克隆或 DNA，在 Southern 印迹中使用了载体材料。为了本发明的目的，杂交指的是该核酸序列可与 SEQ ID NO: 1 所示核酸序列、其互补链、或其亚序列相应的核酸探针在从低至极高度严谨条件下杂交。该核酸探针在这些条件下所杂交的分子均用 X 射线胶片检测。

在一个实施方案中，核酸探针是编码 SEQ ID NO: 2 的多肽或其亚序列的核酸序列。在另一个实施方案中，核酸探针为 SEQ ID NO: 1。在另一个实施方案中，核酸探针是 SEQ ID NO: 1 的第 670 至 2592 位核苷酸，它编码具有氨肽酶活性的成熟多肽。在另一个实施方案中，核酸探针是包含在质粒 pMRT004.1-14 中的核酸序列，该质粒包含在大肠杆菌 NRRL B-30032 中，其中所述核酸序列编码 SEQ ID NO: 2 的多肽。在另一个实施方案中，核酸探针是大肠杆菌 NRRL B-30032 中所含质粒 pMRT004.1-14 中所具有的编码 SEQ ID NO: 2 的成熟多肽（即第 33 至 674 位氨基酸）的核酸序列。

对至少 100 个核苷酸的长探针而言，低度至极高度条件规定为：在 42℃、5X SSPE、0.3% SDS、200 μg/ml 剪切和变性的鲑精 DNA，分别为 25%、35% 或 50% 甲酰胺（分别对应低度、中度、高度和极高度严谨度）中按标准的 Southern 印迹程序进行预杂交和杂交。

对至少 100 个核苷酸的长探针而言，载体材料最后洗 3 次，每次用 2X SSC、0.2% SDS，优选至少在 45℃（极低度严谨）、更优选至少 50℃（低度严谨）、更优选至少 55℃（中度严谨）、更优选至少 60℃（中高度严谨）、更优选 65℃（高度严谨）、最优选 70℃（极高度严谨）洗 15 分钟。

对约 15 至约 70 个核苷酸长的短探针而言，严谨条件规定为：在比按 Bolton & McCarthy 的计算方法（1962，美国国家科学院学报 48:1390）计算的  $T_m$  低 5-10℃ 的温度下，在 0.9 M NaCl, 0.09 M Tris-HCl pH 7.6, 6 mM EDTA, 0.5% NP-40, 1X Denhardt's 溶液, 1 mM 焦磷酸钠, 1 mM 磷酸二氢钠, 0.1 mM ATP, 和 0.2 mg/ml 酵母 RNA 中按标准 Southern 印迹程序进行预杂交、杂交和杂交后洗涤。

对约 15 至 70 个核苷酸长的短探针而言，载体材料在比计算的  $T_m$  低 5-10℃ 的温度下用添加了 0.1% SDS 的 6X SSC 洗一次（15 分钟），并用

6X SSC 洗两次, 每次 15 分钟。

在第三个实施方案中, 本发明涉及具有以下物理化学特性的分离的多肽: (a) 在 37°C 测得, 在 pH 5.0-8.5 的范围内具有氨肽酶活性, (b) 等电点 (pI) 在 7.4-8.5 的范围内; (c) 在 pH7.5 时用 Gly-pNA 的 Tris-HCl 缓冲系统测得, 在 20-55°C 的温度范围内具有氨肽酶活性; (d) 可水解 Ala-pNA, Gly-pNA, Leu-pNA, Glu-pNA, Asp-pNA, Lys-pNA, Ile-pNA 和 Val-pNA; (e) 不能水解 Phe-pNA 和 Pro-pNA; (f) 不受苯甲磺酰氟抑制, EDTA、二异丙基氟磷酸、对氯汞苯甲酸和碘醋酸可轻度抑制, 邻菲洛啉可彻底抑制, 和/或 (g) 得自鞘氨醇单胞菌属的菌株, 分子量为 67 ±5 kDa。

在一个优选实施方案中, 多肽于 37°C 测得在 pH5.0-8.5 的范围内具有氨肽酶活性, 更优选在 pH6.5-8.0 有较高氨肽酶活性, 最优选在 pH7.0-7.5 的范围内具有最高氨肽酶活性。

氨肽酶的等电点不相同, 因为新鲜制备的酶的等电点用活性染色法估计为 8.4, 但在纯化和/或保存过程中等电点在 7.4-8.5 的范围内变化。等电点不稳定可能是由于氨肽酶自其氨基端自我消化所致。

在另一个实施方案中, 多肽分子量为 67 ±5 kDa、优选 67 ±2 kDa, 来自鞘氨醇单胞菌属、更优选荚膜鞘氨醇单胞菌、更优选荚膜鞘氨醇单胞菌 IFO 12533 株。分子量是用 10-15% 的梯度胶在 Pharmacia (Uppsala, Sweden) 的 FAST 系统上经 SDS-聚丙烯酰胺凝胶电泳 (SDS-PAGE) 测定的。酶的分子量根据以下分子量标记的回归线估计: 磷酸化酶 b (94 kDa)、白蛋白 (67 kDa)、卵白蛋白 (43 kDa)、碳酸酐酶 (30 kDa)、胰酶抑制物 (20.1 kDa) 和 α-乳白蛋白 (14.4 kDa)。

在第四个实施方案中, 本发明涉及与具有 SEQ ID NO: 2 氨基酸序列的多肽在免疫化学特性上相同或部分相同的分离的多肽。免疫化学特性是采用熟知的 Ouchterlony 双向免疫扩散操作通过免疫交叉反应相同性检测实验确定的。具体说, 是依照 Harboe 和 Ingild (在 N. H. Axelsen, J. Kroll, 和 B. Weeks 编的“定量免疫电泳手册”, Blackwell Scientific Publications, 1973, 23 章) 或 Johnstone 和 Thorpe (实用免疫化学, Blackwell Scientific Publications, 1982 (具体在 27-31

页)) 描述的步骤, 通过免疫兔(或其他啮齿动物) 制备含有多克隆抗体的抗血清, 该抗体与具有 SEQ ID NO: 2 氨基酸序列之多肽的表位具有免疫反应性或可与之结合。有相同免疫化学特性的多肽是在使用特定免疫化学技术时, 与该抗血清的反应方式相同的多肽(如沉淀物的完全融合、相同的沉淀形态和/或相同的电泳迁移率)。Axelsen、Bock 和 Kroll 在 N. H. Axelsen, J. Kroll, 和 B. Weeks 编“定量免疫电泳手册”(Blackwell Scientific Publications, 1973) 10 章中对免疫化学相同性作了更详细解释。免疫化学特性部分相同的多肽是采用特定的免疫化学技术时与抗血清反应的方式部分相同(如沉淀物的部分融合、部分相同的沉淀形态和/或部分相同的电泳迁移率)的多肽。Axelsen、Bock 在 N. H. Axelsen, J. Kroll, 和 B. Weeks 所编“定量免疫电泳手册”(Blackwell Scientific Publications, 1973) 11 章中对免疫化学特性部分相同作了更详细解释。

抗体也可以是单克隆抗体。单克隆抗体可按如 E. Harlow & D. Lane 编, 1988, 抗体: 实验室指南, Cold Spring Harbor 出版社, Cold Spring Harbor, New York 的方法制备并应用。

本发明的多肽具有 SEQ ID NO: 2 之氨肽酶活性的至少 20%, 优选至少 40%, 更优选至少 60%, 更优选至少 80%, 更优选至少 90%, 最优选至少 100%。

可以从任何属的微生物中获得本发明的多肽。例如, 多肽可以是革兰氏阳性细菌多肽如芽孢杆菌多肽, 例如嗜碱芽孢杆菌、解淀粉芽孢杆菌、短芽孢杆菌、环状芽孢杆菌、凝结芽孢杆菌、灿烂芽孢杆菌、缓慢芽孢杆菌、地衣芽孢杆菌、巨大芽孢杆菌、嗜热脂肪芽孢杆菌、枯草芽孢杆菌或苏云金芽孢杆菌的多肽; 或链霉菌多肽如浅青紫链霉菌或鼠灰链霉菌多肽; 或者革兰氏阴性细菌多肽, 例如大肠杆菌或假单胞菌多肽。

在另一个优选实施方案中, 多肽是鞘氨醇单胞菌多肽, 如粘连鞘氨醇单胞菌、荚膜鞘氨醇单胞菌、类少动鞘氨醇单胞菌、少动鞘氨醇单胞菌或矢野口鞘氨醇单胞菌的多肽。

在一个更优选实施方案中, 荚膜鞘氨醇单胞菌多肽是荚膜鞘氨醇单胞

菌 IFO 12533 的多肽，如具有 SEQ ID NO: 2 的氨基酸序列的多肽。

应理解，对上述物种而言，本发明包括其它分类学等价物，而不管其已知物种名称如何。本领域内技术人员能轻易认识到相应等价名称的一致性。例如，荚膜鞘氨醇单胞菌也称为荚膜黄杆菌，少动鞘氨醇单胞菌也称为贪食黄杆菌和少动假单胞菌，鞘氨醇单胞菌之种也称为蓝黑色杆菌。参见 Yabuuchi 等，1990，微生物学与免疫学 34: 99 - 119 有关鞘氨醇单胞菌分类学的讨论。

公众可从很多培养物保藏中心，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) 及农业研究机构保藏中心 (NRRL) 的北部地区研究中心轻易获得这些物种的菌株。

另外，用前面提到的探针可以从其他来源，包括从自然界（如土壤、沉积物、水体等）分离到的微生物中鉴别并获得这些多肽。用于从其自然生存环境分离微生物的技术是本领域已知的。类似地，通过筛选另一种微生物的基因组或 cDNA 文库可以得到所述核酸序列。一旦用探针检测到了编码所述多肽的核酸序列，就可以采用本领域技术人员已知的技术（参见，例如 Sambrook 等，1989，出处同前）对序列进行分离或克隆。

如文中定义，“分离的”多肽是基本上不含其他非氨肽酶多肽的多肽，例如通过 SDS-PAGE 确定出至少大约 20% 纯，优选至少大约 40% 纯，更优选大约 60% 纯，还要优选的是大约 80% 纯，最优选大约 90% 纯，最最优选大约 95% 纯的多肽。

本发明之核酸序列编码的多肽还包括在多肽或其片段的 N-末端或 C-末端融合了另一个多肽的融合多肽或可裂解的融合多肽。将编码另一个多肽的核酸序列（或其部分）与本发明的核酸序列（或其部分）融合就可产生融合多肽。产生融合多肽的技术为领域内已知，包括连接编码多肽的编码序列，从而使它们在同一读框中，并且融合多肽的表达受控于相同的启动子和终止子。

## 核酸序列

本发明还涉及编码本发明所述多肽的分离的核酸序列。在一个优选实

实施方案中，该核酸序列示于 SEQ ID NO: 1。在另一个更优选的实施方案中，所述核酸序列是质粒 pMRT004.1-14（该质粒包含在大肠杆菌 NRRL B-30032）中所含序列。在另一个更优选实施方案中，所述核酸序列是编码 SEQ ID NO: 2 之成熟多肽（即第 33 至 673 位氨基酸）的序列，该序列包含在大肠杆菌 NRRL B-30032 内所含质粒 pMRT004.1-14 中。在另一个优选实施方案中，所述核酸序列是 SEQ ID NO: 1 中编码成熟多肽的第 670 至 2592 位核苷酸。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，它与 SEQ ID NO: 1 之间因遗传密码的简并性而有所不同。本发明还涉及 SEQ ID NO: 1 的亚序列，所述亚序列编码 SEQ ID NO: 2 的具有氨肽酶活性的片段。

SEQ ID NO: 1 的亚序列是 SEQ ID NO: 1 涵盖的核酸序列，但从 5' 末端和/或 3' 末端缺失掉了 1 或多个核苷酸。优选地，亚序列含有至少 1572 个核苷酸，更优选至少 1722 个核苷酸，最优选至少 1872 个核苷酸。

本发明还涉及在 SEQ ID NO: 1 的成熟多肽编码区至少有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 33 至 673 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 中分离、由 cDNA 制备或将它们组合运用。为了从这类基因组 DNA 克隆本发明所述核酸序列，可以使用已知的聚合酶链式反应（PCR）或用抗体筛选表达文库来检测有共同结构特征的克隆 DNA 片段。参见例如，Innis 等，1990，PCR: 方法和应用指南，Academic Press, New York。还可以使用其他核酸扩增操作，如连接酶链反应（LCR）、连接活化转录（LAT）和基于核酸序列的扩增方法（NASBA）。可以从鞘氨醇单胞菌属的菌株或者另一种微生物或相关微生物中克隆得到所述核酸序列，因此，该核酸序列例如可以是所述核酸序列的多肽编码区之等位变体或种变体。

文中使用的术语“分离的核酸序列”是指基本上不含其他核酸序列的核酸序列，例如，经琼脂糖电泳确定出至少大约 20% 纯，优选至少大约 40% 纯，更优选至少大约 60% 纯，还要优选的是至少大约 80% 纯，最优选



至少大约 90% 纯。例如，可以通过标准克隆步骤获得分离的核酸序列，在基因工程中常用这些步骤来将核酸序列从其天然位置转移到不同的位点并在此复制。克隆步骤可以包括切割和分离所需的包含多肽编码核酸序列的核酸片段，将该片段插入载体分子以及将重组载体导入宿主细胞并在其中复制出多拷贝或多克隆的该核酸序列。核酸序列可以是基因组 DNA、cDNA、RNA、半合成的、合成的或者是它们的任何组合方式。

本发明还涉及编码活性多肽、与 SEQ ID NO: 1 的成熟多肽编码序列（即第 670 至 2592 位核苷酸）有一定同源性的核酸序列，同源程度至少约为 50%、优选约 55%、优选约 60%、优选约 65%、优选约 70%、优选约 80%、更优选约 90%，还要优选的是约 95%，最优选的是大约 97% 同源。就本发明目的而言，两个核酸序列间的同源程度是通过 Wilbur-Lipman 方法（Wilbur & Lipman, 1983, 美国国家科学院学报 80: 726 - 730）用 LASERGENE™ MEGALIGN™ 软件（DNASTAR, Inc., Madison, WI）以及相同性表和以下多重对比参数：缺口罚分和缺口长度罚分均为 10 确定的。成对对比参数为  $K_{\text{tuple}} = 3$ ，缺口罚分 = 3，窗口 (windows) = 20。

合成与本发明所述多肽基本相似的多肽时，可能有必要对编码本发明多肽的核酸序列进行修饰。术语与所述多肽“基本相似”是指非天然形式的多肽。这些多肽可能与从天然来源分离到的多肽在某些加工方式上不同。例如，可能希望用例如定点诱变来合成多肽的变体，这些变体有不同的比活性、热稳定性或最适 pH 等。可以在 SEQ ID NO: 1 之多肽编码部分的核酸序列，例如其亚序列基础上构建出类似序列；并/或通过引入核苷酸取代而构建，所述取代不会使产生的氨基酸序列与原核酸序列编码的多肽不同，但它们符合酶制备所用宿主生物对密码子的使用习惯；或通过引入会产生不同氨基酸序列的核苷酸取代而构建。关于核苷酸取代的综述，参见例如 Ford 等，1991，蛋白质表达和纯化 2: 95-107。

对本领域技术人员显而易见的是，可以在分子功能关键区以外作这些取代，而仍可以得到活性多肽。那些为本发明所述分离核酸序列所编码多肽的活性所必需、因此最好不进行取代的氨基酸残基可以依照本领域的已知操作，例如定点诱变或丙氨酸扫描诱变（参见例如，Cunningham

和 Wells, 1989, 科学 244: 1081-1085) 进行鉴定。后一种技术中, 在分子中每个带正电荷的残基上导入突变, 然后检测所得突变分子的氨肽酶活性从而鉴定出对分子活性比较关键的氨基酸残基。通过对用如核磁共振、晶体学或光亲和标记测定三维结构进行分析, 可以确定底物-酶相互作用的位点(参见例如, de Vos 等, 1992, 科学 255: 306-312; Smith 等, 1992, 分子生物学杂志 224: 899-904; Wlodaver 等, 1992, FEBS 快报 (FEBS Letters) 309: 59-64)。

本发明还涉及编码本发明所述多肽的分离的核酸序列, 这些序列在低度严谨条件, 更优选在中度严谨条件, 更优选在高度严谨条件, 最优选在极高度严谨条件下能与这样的核酸探针杂交, 该探针在相同的条件下能与 SEQ ID NO: 1 的核酸序列或其互补链, 或其等位变体和亚序列杂交 (Sambrook 等, 1989, 出处同前), 如本文所述。

本发明还涉及经以下方法产生的分离的核酸序列: (a) 将 DNA 与 SEQ ID NO: 1 或其互补链, 或其编码氨肽酶活性多肽片段的亚序列在低度、中度、高度或极高度严谨条件下杂交; 和 (b) 分离该核酸序列。

#### 产生突变核酸序列的方法

本发明还涉及产生突变核酸序列的方法, 包括向 SEQ ID NO: 1 的成熟多肽编码序列或其亚序列中导入至少一个突变, 其中该突变核酸序列编码由 SEQ ID NO: 2 的第 33 至 673 位氨基酸组成的多肽或其有氨肽酶活性的片段。

向核酸序列中导入突变使一个核苷酸变成另一个核苷酸的过程, 可通过应用本领域内已知的任何方法经定点诱变完成。应用带有所需插入片段的超螺旋双链 DNA 载体及两个含有所需突变的合成引物的技术十分有用。分别互补于载体相反链的核苷酸引物在温度循环期间经 *Pfu* DNA 聚合酶延伸。一旦插入引物便产生带有交错切口的突变质粒。温度循环后, 产物用特异于甲基化和半甲基化 DNA 的 *DpnI* 处理, 以消化亲代 DNA 模板并选择出带突变的合成 DNA。也可应用本领域内已知的其它技术。

#### 核酸构建体

本发明还涉及包含本发明所述核酸序列及与之可操作连接的 1 或多个

调控序列的核酸构建体，所述调控序列在其相容条件下能指导编码序列在合适的宿主细胞中进行表达。表达应理解为包括多肽生产中所涉及的任何步骤，包括，但不限于转录、转录后修饰、翻译、翻译后修饰和分泌。

“核酸构建体”在文中定义为单链或双链核酸分子，它们分离自天然基因，或者经修饰而含有以非天然方式组合和并列的核酸片段。当核酸构建体包含表达本发明所述编码序列必需的所有调控序列时，术语核酸构建体与表达盒同义。术语“编码序列”在文中定义为核酸序列中直接确定其蛋白产物的氨基酸序列的部分。编码序列的边界通常是由紧邻 mRNA 5'端开放读码框上游的核糖体结合位点(对于原核细胞)和紧邻 mRNA 3'端开放读码框下游的转录终止序列确定。编码序列可以包括，但不限于 DNA、cDNA 和重组核酸序列。

可以以多种方式操作编码本发明所述多肽的分离的核酸序列，使其表达所述多肽。可能期望或必须在插入载体之前对核酸序列进行加工，这取决于表达载体。应用重组 DNA 方法修饰核酸序列的技术为本领域所熟知。

本文中术语“控制序列”定义为包括表达本发明多肽所必需或有利的所有组分。每个调控序列对于编码多肽的核酸序列可以是天然含有的或外来的。这些调控序列包括，但不限于，前导序列、多聚腺苷酸化序列、前肽序列、启动子、信号序列和转录终止子。最低限度，调控序列要包括启动子以及转录和翻译的终止信号。为了导入特定的限制位点以便将调控序列与编码多肽的核酸序列的编码区进行连接，可以提供带接头的调控序列。术语“可操作连接”在文中定义为这样一种构象，其中调控序列位于相对 DNA 序列之编码序列的适当位置，以使调控序列指导多肽的表达。

调控序列可以是合适的启动子序列，即可被表达核酸序列的宿主细胞识别的核酸序列。启动子序列含有介导多肽表达的转录调控序列。启动子可以是在所选宿主细胞中有转录活性的任何核酸序列，包括突变的、截短的和杂合的启动子，可以得自编码与宿主细胞同源或异源的胞外或

胞内多肽的基因。

指导转录本发明之核酸构建体的合适启动子的例子(尤其在细菌宿主细胞中)是得自以下来源的启动子:大肠杆菌 lac 操纵子、天蓝色链霉菌琼脂糖酶基因(dagA)、枯草芽孢杆菌果聚糖酶基因(sacB)、地衣芽孢杆菌 $\alpha$ -淀粉酶基因(amyL)、嗜热脂肪芽孢杆菌麦芽糖源淀粉酶基因(amyM)、解淀粉芽孢杆菌 $\alpha$ -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,原核细胞的 $\beta$ -内酰胺酶基因(Villa-Kamaroff 等,1978,美国国家科学院学报 75: 3727- 3731),以及 tac 启动子(DeBoer 等,1983,美国国家科学院学报,80:21-25)。

“来自重组细菌的有用蛋白质”(科学美国人,1980,242: 74-94)一文以及 Sambrook 等(1989,出处同前)描述过其他启动子。

调控序列还可以是合适的转录终止序列,即能被宿主细胞识别从而终止转录的一段序列。终止序列可操作连接在编码多肽的核酸序列的 3' 末端。在所选宿主细胞中可发挥功能的任何终止子都可以用于本发明。

调控序列还可以是合适的前导序列,即对宿主细胞的翻译十分重要的 mRNA 非翻译区。前导序列可操作连接于编码多肽的核酸序列的 5' 末端。在所选宿主细胞中可发挥功能的任何前导序列均可用于本发明。

调控序列还可以是信号肽编码区,该区编码一段连在多肽氨基端的氨基酸序列,能引导编码多肽进入细胞分泌途径。核酸序列编码区的 5' 端可能天然含有翻译读框一致地与分泌多肽的编码区片段自然连接的信号肽编码区。或者,编码区的 5' 端可含有对编码序列是外来的信号肽编码区。当编码序列在正常情况下不含有信号肽编码区时,可能需要添加外来信号肽编码区。或者,可以用外来的信号肽编码区简单地替换天然的信号肽编码区以增强多肽分泌。但是,任何能引导表达后的多肽进入所用宿主细胞的分泌途径的信号肽编码区都可以用于本发明。

有效于细菌宿主细胞的信号肽编码区可以是得自芽孢杆菌 NCIB 11837 的麦芽糖原淀粉酶基因、嗜热脂肪芽孢杆菌 $\alpha$ -淀粉酶基因、地衣芽孢杆菌枯草杆菌蛋白酶基因、地衣芽孢杆菌 $\beta$ -内酰胺酶基因、嗜热脂肪芽孢杆菌中性蛋白酶基因(nprT, nprS, nprM),或枯草芽孢杆菌 prsA

基因的信号肽编码区。Simonen & Palva, 1993, 微生物学综述 57: 109 - 137 描述了其它信号肽。

调控序列还可以是肽原编码区, 该区编码位于多肽氨基末端的一段氨基酸序列。所得多肽被称为酶原或多肽原。多肽原通常没有活性, 可以通过催化或自我催化而从多肽原切割肽原而转化为成熟的活性多肽。肽原编码区可以得自枯草芽孢杆菌碱性蛋白酶基因 (aprE)、或枯草芽孢杆菌中性蛋白酶基因 (nprT)。

在多肽的氨基末端既有信号肽又有肽原区时, 肽原区紧邻多肽的氨基末端, 而信号肽区则紧邻肽原区的氨基末端。

添加能根据宿主细胞的生长情况来调节多肽表达的调控序列可能也是需要的。调控系统的例子是那些能对化学或物理刺激物 (包括在有调控化合物的情况下) 作出反应, 从而开放或关闭基因表达的系统。原核细胞体系中的调控系统包括 lac、tac 和 trp 操纵子系统。调控序列的其他例子是那些能使基因扩增的调控序列。在这些例子中, 应将编码多肽的核酸序列与调控序列可操作连接在一起。

### 表达载体

本发明还涉及包含本发明核酸序列、启动子和转录及翻译终止信号的重组表达载体。可以将上述各种核酸和调控序列连接在一起来制备重组表达载体, 该载体可以包括 1 或多个方便的限制位点, 以便在这些位点插入或取代编码多肽的核酸序列。或者, 可以通过将核酸序列或包含该序列的核酸构建体插入适当表达载体而表达本发明所述核酸序列。制备表达载体时, 可使编码序列位于载体中以便与适当的表达调控序列可操作连接。

重组表达载体可以是任何便于进行重组 DNA 操作并表达核酸序列的载体 (例如质粒或病毒)。载体的选择通常取决于载体与它将要导入的宿主细胞的相容性。载体可以是线性或闭环质粒。

载体可以是自主复制型载体 (即存在于染色体外的完整结构, 可独立于染色体进行复制), 例如质粒、染色体外元件、微小染色体或人工染色体。载体可包含保证自我复制的任何机制。或者, 载体是一个当导入宿

主细胞时，将整合到基因组中并与所整合到的染色体一起复制的载体。此外，可应用单个载体或质粒，或总体包含将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。

优选本发明所述载体含有 1 或多个便于选择转化细胞的选择标记。选择标记是这样一个基因，其产物赋予对杀生物剂或病毒的抗性、对重金属的抗性，或赋予营养缺陷体原养型等。细菌选择标记的例子如枯草芽孢杆菌或地衣芽孢杆菌的 *dal* 基因，或者抗生素如氨基青霉素、卡那霉素、氯霉素或四环素的抗性标记。

优选本发明所述载体包含能使载体稳定整合到宿主细胞基因组中，或保证载体在细胞中独立于细胞基因组而进行自主复制的元件。

就整合到宿主细胞基因组的情况而言，载体可以依赖于编码多肽的核酸序列或载体的其他元件，以便载体通过同源或非同源重组稳定地整合到基因组中。或者，载体可含有用于引导通过同源重组整合到宿主细胞基因组中的附加核酸序列。附加的核酸序列能使载体整合到宿主细胞基因组中染色体上的精确位点。为了增加在精确位点整合的可能性，优选整合元件应含有足够数量的核酸，如 100 到 1500 个碱基对，优选 400 到 1500 个碱基对，最优选 800 到 1500 个碱基对，而且与相应靶序列高度同源，以增加同源重组的可能性。整合元件可以是与宿主细胞基因组中靶序列同源的任何序列。另外，整合元件可以是非编码或编码核酸序列。另一方面，载体可以通过非同源重组整合到宿主细胞的基因组中。

就进行自主复制的情况而言，载体还可以包含复制起点，使载体能在目标宿主细胞中自主地复制。细菌复制起点的例子如可在大肠杆菌中复制的质粒 pBR322、pUC19、pACYC177 和 pACYC184 的复制起点，以及可在芽孢杆菌中复制的 pUB110、pE194、pTA1060、和 pAMB1 的复制起点。复制起点可以带有使其在宿主细胞中成为温度敏感型的突变（参见例如，Ehrlich, 1978, 美国国家科学院学报 75: 1433）。

可以向宿主细胞插入 1 个以上拷贝的本发明核酸序列以提高该基因产物的产量。该核酸序列的拷贝数增加可以通过将该序列的至少 1 个附加拷贝插入宿主细胞基因组中，或者与该核酸序列一起插入一个可扩增的

选择标记，通过在有合适选择试剂存在下培养细胞，挑选出含有扩增拷贝的选择标记基因，从而含有附加拷贝核酸序列的细胞。

用于连接上述各元件来构建本发明所述重组表达载体的操作是本领域技术人员熟知的（参见例如，Sambrook等，1989，出处同前）。

#### 宿主细胞

本发明还涉及包含便于用来重组生产多肽的本发明所述核酸序列的重组宿主细胞。可将包含本发明之核酸序列的载体导入宿主细胞，从而使该载体以上述染色体整合体或自我复制的染色体外载体形式得以维持。术语“宿主细胞”涵盖任何由于复制期间发生的突变而与亲本细胞不同的后代。宿主细胞的选择很大程度上取决于多肽编码基因及其来源。

宿主细胞可以是单细胞微生物（例如原核细胞）。可利用的单细胞是细菌细胞，如革兰氏阳性细菌，包括，但不限于芽孢杆菌属的细胞，例如：嗜碱芽孢杆菌、解淀粉芽孢杆菌、短芽孢杆菌、环状芽孢杆菌、*Bacillus clausii*、凝结芽孢杆菌、灿烂芽孢杆菌、缓慢芽孢杆菌、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、苏云金芽孢杆菌；或者链霉菌属的细胞，例如浅青紫链霉菌或鼠灰链霉菌；或者是革兰氏阴性细菌，例如大肠杆菌或假单胞菌。在一个优选实施方案中，细菌宿主细胞是缓慢芽孢杆菌、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，芽孢杆菌细胞是嗜碱芽孢杆菌的细胞。

可以通过例如原生质体转化（参见，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（参见，例如，Young和Spizizin, 1961, 细菌学杂志 81:823-829, 或Dubnau和Davidoff-Abelson, 1971, 分子生物学杂志 56:209-221）、电穿孔（参见，例如，Shigekawa和Dower, 1988, 生物技术 6:742-751）或接合（参见，例如，Koehler和Thorne, 1987, 细菌学杂志 169: 5771-5278）将载体导入细菌宿主细胞。

#### 制备方法

本发明还涉及制备本发明多肽的方法，该方法包括（a）培养野生型

形式能产生所述多肽的菌株，从而得到包含该多肽的上清液；和（b）回收多肽。优选地，该菌株是鞘氨醇单胞菌，更优选为荚膜鞘氨醇单胞菌。

本发明还涉及制备本发明多肽的方法，该方法包括（a）在有利于所述多肽产生的条件下培养宿主细胞；和（b）回收该多肽。

本发明还涉及制备本发明多肽的方法，包括（a）在有利于该多肽产生的条件下培养宿主细胞，其中宿主细胞中含有在 SEQ ID NO: 1 的核酸序列中有至少有 1 个突变的突变核酸序列，该突变核酸序列编码具有 SEQ ID NO: 2 的氨基酸序列的多肽；和（b）回收该多肽。

在本发明所述制备方法中，用本领域已知方法在适合多肽产生的营养培养基中培养细胞。例如，可以在合适的培养基中，在允许多肽表达和/或分离的条件下，通过摇瓶培养、实验室或工业发酵罐中小规模或大规模发酵（包括连续、分批、分批加料或固态发酵）来培养细胞。在包含碳和氮源以及无机盐的合适的培养基中，采用本领域已知的步骤进行培养。合适的培养基可由供应商提供或者可以参照公开的组成（例如，美国典型培养物保藏中心的目录中所述）来制备。如果多肽被分泌到培养基中，则可以直接从培养基中回收多肽。如果多肽不分泌，可以从细胞裂解物中回收。

可以用本领域已知针对所述多肽的特异方法检测多肽。这些检测方法包括特异性抗体的使用、酶产物的形成或酶底物的消失。例如，可以用酶法确定多肽的活性。用于确定氨肽酶活性的步骤是本领域已知的。如前述，氨肽酶活性可通过在 405nm 测量丙氨酰对硝基苯胺的起始水解率而确定。

可以用本领域已知方法回收所产生的多肽。例如，可以通过常规操作（包括，但不限于离心、过滤、抽提、喷雾干燥、蒸发或沉淀）从培养基中回收多肽。

可以通过各种本领域已知的操作来纯化本发明所述多肽，这些操作包括，但不限于层析（例如，离子交换层析、亲和层析、疏水作用层析、层析聚焦、和大小排阻层析）、电泳（例如，制备性等电点聚焦）、差示溶解度（例如硫酸铵沉淀）、SDS-PAGE 或抽提（参见例如，蛋白质纯化，



J. C. Janson 和 Lars Ryden 编, VCH Publishers, New York, 1989)。

## 植物

本发明还涉及转基因植物、植物部分或植物细胞, 它们转化有编码本发明之氨肽酶活性多肽的核酸序列, 因而可表达并产生足以回收的多肽。多肽可从植物或植物部分中回收。或者, 包含重组体多肽的植物或植物部分可被用于提高食物或饲料的质量, 如提高营养价值、改进口味和流变特性, 或破坏不利于营养的因素。

转基因植物可为双子叶植物或单子叶植物(缩写为 dicot, monocot)。单子叶植物实例有禾本科植物如草地早熟禾, 饲用牧草如羊茅、黑麦草, 寒地型牧草如翦股颖, 以及谷类如小麦、燕麦、黑麦、大麦、稻、高粱和玉米。双子叶植物实例有烟草, 豆类如羽扇豆、马铃薯、甜菜、豌豆、菜豆和大豆, 十字花科植物(芸苔科)如花椰菜、油籽油菜以及密切相关的模型植物拟南芥。

植物部分的实例有茎、愈伤组织、叶、根、果实、种子和块茎。在本文中, 叶绿体、质外体、线粒体、液泡、过氧化物酶体和细胞质等特定植物组织均被认为是植物部分。此外, 任何植物细胞, 不管其组织来源是什么, 均被认为是植物部分。

本发明还包括这些植物、植物部分和植物细胞的后代。

表达本发明多肽的转基因植物或植物细胞可按领域内已知方法构建。简而言之, 植物或植物细胞的构建可以通过将 1 或多个编码本发明多肽的表达构建体导入植物宿主基因组并使修饰后植物或植物细胞增殖成转基因植物或植物细胞而完成。

为方便起见, 表达构建体可以是其中含编码本发明多肽的核酸序列可操作连接了在所选植物或植物部分中表达该核酸序列必需的适当调控序列的核酸构建体。而且, 表达构建体可能包含可用于对表达构建体已整合其中的宿主细胞进行鉴别的选择标记以及将构建体导入目标植物所必需的 DNA 序列(后者取决于所用 DNA 导入方法)。

调控序列(如启动子和终止子序列及任选信号序列或转位序列)的选择取决于如希望何时、何地和以何种方式表达该多肽。例如, 编码本

发明多肽的核酸序列的表达可以是组成型或诱导型，或者发育阶段或组织特异型，表达产物可定位于特定组织或植物部分如种子或叶片。调控序列描述于如 Tague 等，植物生理学 86, 506, 1988 中。

组成型表达中可用 35S - CaMV 启动子 (Franck 等, 1980, 细胞 21: 285 - 294)。器官特异性启动子可以是例如贮藏 sink 组织如种子、马铃薯块茎和果实的启动子 (Edwards & Coruzzi, 1990. 遗传学年鉴 24: 275 - 303)、或代谢 sink 组织如分生组织的启动子 (Ito 等, 1994, 植物分子生物学 24: 863 - 878)、种子特异性启动子如稻的谷蛋白、谷醇溶蛋白、球蛋白或白蛋白启动子 (Wu 等, 植物和细胞生理学, 39 卷, 第 8 期, 第 885 - 889 页 (1998))、豆球蛋白 B4 和蚕豆的未知种子蛋白基因的蚕豆启动子 (Conrad U 等, 植物生理学杂志, 卷 152, 第 6 期, 第 708 - 711 页 (1998) 述)、种子油体蛋白的启动子 (Chen 等, 植物和细胞生理学, 卷 39, 第 9 期, 第 935 - 941 页 (1998))、蔓菁的贮藏蛋白 napA 启动子或领域内已知的其它种子特异性启动子, 如 WO 91/14772 所述。此外, 启动子可以是叶特异性启动子如大米或番茄的 rbcS 启动子 (Kyoizuka 等, 植物生理学, 卷 102, 第 3 期, 第 991 - 1000 页 (1993))、绿藻病毒的腺嘌呤甲基转移酶基因启动子 (Mittra, A. & Higgins, DW, 植物分子生物学, 卷 26, 第 1 期, 第 85 - 93 页 (1994))、或大米的 aldP 基因启动子 (Kagaya 等, 分子和普通遗传学, 卷 248, 第 6 期, 第 668 - 674 页 (1995))、或创伤诱导型启动子如马铃薯 pin2 启动子 (Xu 等, 植物分子生物学, 卷 22, 第 4 期, 第 573 - 588 页 (1993))。

可应用启动子增强子元件以获得植物体内多肽的高水平表达。例如, 启动子增强子元件可以是位于该启动子与编码多肽的核酸序列之间的一个内含子。例如, Xu 等 (出处同前) 公开了大米肌动蛋白 1 基因的第一内含子在增强表达方面的应用。

选择标记基因及表达构建体的任何其它部分均可选自本领域内可供使用的那些。

按本领域内已知的常规方法将核酸构建体导入植物基因组, 包括农杆菌介导的转化、病毒介导的转化、显微注射、粒子轰击、biolistic

转化以及电穿孔 (Gasser 等, 科学, 244, 1293; Potrykus, 生物/技术, 8, 535, 1990; Shimamoto 等, 自然, 338, 274, 1989)。

近来, 根癌农杆菌介导的基因转移已成为制作双子叶转基因植物的首选方法 (综述见 Hooykas & Schilperoort, 1992, 植物分子生物学, 19: 15 - 38), 而它也可用于转化单子叶植物, 只不过单子叶植物常优选其它转化方法。目前, 单子叶植物转基因操作的首选方法是对胚性愈伤组织或正在发育的胚进行粒子轰击 (包裹了转化 DNA 的金或钨微粒) (Christou, 1992, 植物杂志 2: 275 - 281; Shimamoto, 1994, 生物技术现状 (Curr. Opin. Biotechnol.) 5: 158 - 162; Vasil 等, 1992, 生物/技术 10: 667 - 674)。另一种转化单子叶植物的方法是基于原生质体的转化, 如 Omirulleh S 等, 植物分子生物学, 卷 21, 第 3 期, 第 415 - 428 页 (1993) 所述。

转化后, 按领域内已知方法选择导入了表达构建体的转化子并再生成完整植株。

本发明还涉及产生本发明多肽的方法, 包括 (a) 在有助于多肽生产的条件下培养含有编码本发明之氨肽酶活性多肽的核酸序列的转基因植物或植物细胞; 和 (b) 回收多肽。

#### 去除或降低氨肽酶活性

本发明还涉及制备亲本细胞的突变细胞的方法, 包括破坏或删除编码该多肽的核酸序列或者它的调控序列, 从而使突变细胞在相同培养条件下时多肽产量比亲本细胞少。

可以对有氨肽酶活性的多肽在细胞中表达所必需的核酸序列进行修饰或使它失活来方便地构建出氨肽酶活性降低的菌株。要进行修饰或失活的核酸序列可以是, 例如编码该多肽或表现氨肽酶活性所必需的其多肽部分的核酸序列, 或者核酸序列中可以具有其编码序列表达为多肽时所需要的调控功能。这种调控或调节序列的例子可以是启动子序列或它的功能性部分 (即足够影响多肽表达的那部分)。其他可能被修饰的调控序列如上述。

对核酸序列的修饰或失活可通过对细胞进行诱变处理并挑选出其氨

肽酶产生能力降低的细胞来完成。特异性或随机诱变可以通过例如使用合适的物理或化学诱变剂、使用合适的核苷酸、或使 DNA 序列经 PCR 诱变而完成。另外，可以将这些诱变剂任意组合起来进行诱变。

适用于此目的的物理或化学诱变剂的例子包括紫外线(UV)照射、羟胺、N-甲基-N'-硝基-N-亚硝基胍(MNNG)、O-甲基羟胺、亚硝酸、甲磺酸乙酯(EMS)、亚硫酸氢钠、甲酸和核苷酸类似物。

使用这类试剂时，诱变通常是通过在适当条件下有所选诱变剂时温育待诱变细胞，并挑选出氨肽酶活性降低或产量减少的细胞而完成。

可以通过在编码多肽的核酸序列或者其转录或翻译所需调控元件中导入、取代或去除 1 或多个核苷酸来完成本发明所述多肽生产的修饰或失活。例如，可以插入或删除核苷酸来导入终止子、去掉起始密码子或改变开放读码框。可以依照本领域已知的方法通过定点突变或 PCR 诱变来达到这种修饰或失活的目的。尽管一般来说，修饰可以在体内进行，即直接对待修饰核酸序列的表达细胞进行，但优选体外修饰，如下述实例。

使所选宿主细胞中的表达失活或降低的方便方法的例子是基于基因取代或基因中断技术。例如，在基因中断法中，将相应于目标内源基因或基因片段的核酸序列经体外诱变产生缺陷的核酸序列，然后将该序列转化到宿主细胞中产生缺陷型基因。经过同源重组，缺陷型核酸序列将取代内源基因或基因片段。最好缺陷基因或基因片段还编码一种标记，以使用该标记挑选出其中多肽编码基因已被修饰或破坏了的转化子。

或者，可以用互补于多肽编码序列的核苷酸序列经成熟的反义技术使编码本发明多肽的核酸序列修饰或失活。更具体地说，可以导入与编码多肽的核酸序列互补的核苷酸序列，该序列在细胞中可以进行转录并能与细胞中产生的多肽 mRNA 杂交，从而降低或消除细胞的多肽产量。在能使互补的反义核苷酸序列与多肽 mRNA 杂交的条件下，翻译出的多肽产量就会因此降低或消除。

优选地，欲按照本发明所述方法进行修饰的细胞是微生物来源的，例如是适于产生同源或异源目标蛋白质产物的真菌菌株。

本发明还涉及亲本细胞的突变细胞，其中所含编码多肽的核酸序列或其调控序列被破坏或缺失，从而导致突变细胞中的多肽产量少于亲本细胞。

这样得到的多肽缺陷型突变细胞特别适合作为表达同源和/或异源多肽的宿主细胞。因此，本发明还涉及同源或异源多肽的制备方法，包括（a）在有利于多肽产生的条件下培养突变细胞；和（b）回收多肽。术语“异源多肽”此处定义为宿主细胞的非天然多肽、经修饰而改变了天然序列的天然蛋白质，或者因对宿主细胞进行了重组 DNA 技术操作而使其表达量有改变的天然蛋白质。

另一方面，本发明涉及通过对既能产生本发明所述多肽又能产生目标蛋白质产物的细胞进行发酵来生产基本没有氨肽酶活性的蛋白质产物的方法。该方法包括在发酵期间或发酵完成后，向发酵液中添加足够量的能抑制氨肽酶活性的试剂，从发酵液中回收目标产物，并任选地将回收的产物进一步纯化。该方法详述于以下实施例。

本发明另一方面涉及生产基本上没有氨肽酶活性的蛋白质产物的方法，其中目标蛋白质产物由细胞中编码本发明所述多肽的 DNA 序列编码。该方法包括在允许产物表达的条件下培养细胞，将得到的培养液进行 pH 和温度的组合处理以便实质性降低氨肽酶活性，并从培养液中回收产物。或者，可以针对从培养液中回收的酶制剂进行 pH 和温度的组合处理。pH 和温度的组合处理可任选与氨肽酶抑制剂处理联合进行。

根据本发明的这一方面，有可能除去至少 60%、优选至少 75%、更优选至少 85%、还要优选至少 95%、最优选至少 99% 的氨肽酶活性。预计通过使用这些方法可以完全消除氨肽酶活性。

优选在 pH6.5-7.5 和 40-60℃ 的范围内进行足够长时间的 pH 和温度组合处理以达到预期的效果，通常 30-60 分钟即可。

用于培养和纯化目标产物的方法可以采用本领域已知的方法。

本发明所述用于生产基本没有氨肽酶活性的产物的方法对于生产真核细胞多肽（尤其是真菌蛋白质，比如酶）特别有意义。所述的酶可以选自，例如，淀粉分解酶、脂类分解酶、蛋白水解酶、纤维素水解酶、

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(cyclodextrin glycosyltransferase)、脱氧核糖核酸酶、酯酶、半乳糖苷酶、 $\beta$ -半乳糖苷酶、葡糖淀粉酶、葡萄糖氧化酶、葡糖苷酶、卤素过氧化物酶(haloperoxidase)、半纤维素酶、转化酶、异构酶、漆酶、连接酶、脂酶、裂合酶、甘露糖苷酶、氧化酶、果胶水解酶、过氧化物酶、植酸酶、酚氧化酶、多酚氧化酶、蛋白水解酶、核糖核酸酶、转移酶、转谷氨酰胺酶或木聚糖酶。氨肽酶缺陷型细胞也可用于表达有药用价值的异源蛋白质,如激素、生长因子、受体等。

术语“真核细胞多肽”应理解为不仅包括天然多肽,也包括通过氨基酸取代、缺失或添加或者其他此类修饰使它们的活性、热稳定性、pH耐受能力等增强的多肽(例如酶)。

本发明另一个方面涉及通过本发明方法制备的基本上没有氨肽酶活性的蛋白质产物。

#### 应用

本发明所述多肽可用于蛋白水解产物生产中以提高水解程度并改善口味。

本发明所述多肽还可用于使酶失活。

另外,本发明所述多肽可以用于多种需要特异性水解肽序列的目的中。例如,某些蛋白质或肽是以非活性的前体形式(在成熟蛋白质的N-端包含许多附加氨基酸残基)合成的。本发明所述多肽可以提供必要的翻译后加工从而活化这类前体蛋白质。

#### 信号肽

本发明还涉及这样的核酸构建体,其中含有与编码SEQ ID NO: 2中第1至32位氨基酸所组成信号肽的SEQ ID NO: 1中第574至669位核苷酸所组成之核酸序列可操作连接的蛋白质编码基因,该基因相对于该核酸序列是外源的。

本发明还涉及包含这类核酸构建体的重组表达载体和重组宿主细胞。

本发明还涉及产生蛋白质的方法，包括(a)在适于目的蛋白质产生的条件下培养含有核酸构建体的重组宿主细胞，其中所述核酸构建体包含与由 SEQ ID NO: 1 中第 574 至 669 位核苷酸组成的信号肽编码区可操作连接的蛋白质编码基因，该基因相对于该核酸序列为外源性；和(b)回收蛋白质。

核酸序列可以可操作连接带有其它调控序列的外源基因。这类其它调控序列如上述。

该蛋白质可以是任何蛋白质。此外，该蛋白质可以是宿主细胞的天然或异源蛋白。术语“蛋白质”在此并非指特定长度的编码产物，因此包含肽、寡肽和蛋白质。术语“蛋白质”还包含组合成编码产物的两个或多个多肽。蛋白质还包括杂合多肽，其中有至少来自两种不同蛋白质（其中一或多种为宿主细胞的异源或天然蛋白质）的部分或完整多肽序列的组合。蛋白质还包括上述蛋白质和杂合蛋白质的天然等位变体和基因工程化变异体。

优选该蛋白质为激素、激素变体、酶、受体或其部分、抗体或其部分、或报道分子。在一个更优选实施方案中，该蛋白质是氧化还原酶、转移酶、水解酶、裂合酶、异构酶或连接酶。在一个更优选实施方案中，该蛋白质是氨肽酶、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、 $\alpha$ -半乳糖苷酶、 $\beta$ -半乳糖苷酶、葡糖淀粉酶、 $\alpha$ -葡糖苷酶、 $\beta$ -葡糖苷酶、转化酶、漆酶、脂酶、甘露糖苷酶、齿斑葡聚糖酶(mutanase)、氧化酶、果胶水解酶、过氧化物酶、植酸酶、多酚氧化酶、蛋白水解酶、核糖核酸酶、转谷氨酰胺酶或木聚糖酶。

基因可来自任何原核细胞、真核细胞或其它来源。为本发明之目的，术语“来自”在本文中与所述特定来源联用时，指的是蛋白质由该来源产生或由插入了该来源的基因的细胞产生。

通过以下实施例进一步描述本发明，但不应将其理解为对本发明范围的限制。

## 实施例

### 实施例 1: 荚膜鞘氨醇单胞菌的粗酶粉的制备

在内含 3 升培养液的 5 升发酵罐中 30℃ 通气并搅动培养荚膜鞘氨醇单胞菌 IF0 12533 9 小时, 该培养液含有 0.5% 蔗糖、1% 明胶、0.5% 酵母浸膏、1% 玉米浸液、0.3% NaCl、0.2%  $K_2HPO_4$  和 0.1%  $MgSO_4 \cdot 7H_2O$ 。自培养液中收集细胞, 在 10mM Tris-HCl pH8.0 缓冲液中洗两次, 得到湿重为 43g 的细胞团。将细胞团悬浮于 10mM Tris-HCl pH8.0 缓冲液 200ml 中, 加入溶菌酶和 Triton X-100 至终浓度分别为 1 mg/ml 和 0.1% 并将溶液于 37℃ 温育 1 小时以进行裂解。对溶液进行超声破碎, 再在 15,000 x g 离心 10 分钟。回收上清 (200ml), 加入硫酸鱼精蛋白至终浓度为 0.1% 以沉淀核酸。经相同方式的离心弃去沉淀, 所得溶液作为酶的粗提物使用。

测定了酶粗提物对各种合成肽的活性, 如表 1 所示。结果证实酶粗提物具有二肽酰肽酶 IV、亮氨酸氨肽酶、甘氨酸氨肽酶和脯氨酰寡肽酶活性。

表 1. 酶粗提物溶液中各种肽酶的活性

肽酶	底物	活性 (%)
脯氨酰寡肽酶	Z-Ala-Ala-Pro-pNA	46.0
脯氨酰寡肽酶	Z-Gly-Pro-NA	68.7
二肽酰肽酶 IV	Gly-Pro-pNA	21.9
甘氨酸氨肽酶	Gly-pNA	100
亮氨酸氨肽酶	Leu-pNA	67.8

酶粗提物加入等体积的冷丙酮沉淀。将酶沉淀物溶于少量 10mM 磷酸盐 pH6.0 缓冲液中, 不溶物离心除去。酶经冻干变成粗酶粉。

### 实施例 2: 从粗酶粉中纯化荚膜鞘氨醇单胞菌氨肽酶 I

将实施例 1 所述粗酶粉溶于 20mM 磷酸盐 pH7.0 缓冲液中, 并稀释至样品的导电率与上样缓冲液 20mM 磷酸盐 pH7.0 的相等。将样品上样于已用 20mM 磷酸 pH7.0 缓冲液预平衡的 Pharmacia Q-Sepharose 或 Mono Q 层析柱。按下述过程用丙氨酰-对硝基苯胺 (Ala-pNA) 分析这些层析流



出物中氨肽酶的活性。

将 Ala-pNA 的 100 mg/ml 二甲亚砜贮存液用 50mM 磷酸钠 pH7.5 缓冲液稀释至终浓度为 2 mg/ml。将 10 - 50  $\mu$ l 酶液加入微滴板小孔中的 200  $\mu$ l 底物溶液中启动氨肽酶与该酰基-对硝基苯胺的反应。该酰基-对硝基苯胺水解的起始速率分析在 THERMOmax Microplate Reader (Molecular Devices Corp., Sunnyvale, CA, USA) 中、室温下于 405 nm 处进行监控。

流出物中有氨肽酶活性。用配备了 DIAFLO®PM 10 超滤膜 (Amicon, Inc., USA) 的 Amicon Spiral 超滤系统 (Amicon, New Bedford, MA, USA) 浓缩流出物, 再将其 pH 用 70 mM 醋酸钠 pH4.0 缓冲液调整至 6.0。

将浓缩的流出物上样于已用 50 mM MES pH6.0 缓冲液预平衡的 7 x 50 mm 的 Pharmacia Mono S 5/5 预装柱 (Pharmacia Biotech AB, Uppsala, Sweden)。如上述分析级分的氨肽酶活性。结合的氨肽酶活性用所含 NaCl 经 0 而递增至 0.2 M 的梯度的 50 mM MES pH6.0 缓冲液洗脱。有显著活性的级分再次用 PM 10 超滤膜浓缩, 然后用含 0.5 M 硫酸铵的 50 mM 磷酸盐缓冲液 pH7.0 平衡。

然后将浓缩的样品上样于已用含 0.5 M 硫酸铵的 50 mM 磷酸盐 pH7.0 缓冲液预平衡的 Phenyl Superose 层析柱。酶用含 0.5M 硫酸铵的 50mM 磷酸盐缓冲液减至无硫酸铵的 50 mM 磷酸盐缓冲液梯度洗脱。将活性级分汇总。

对纯化氨肽酶的 SDS - PAGE 分析显示一条分子量为 67 kDa 的带。纯化的氨肽酶被命名为荚膜鞘氨醇单胞菌氨肽酶 I。

实施例 3: 从粗制酶粉末中纯化荚膜鞘氨醇单胞菌氨肽酶 I

将实施例 1 所述酶粗粉溶于蒸馏水并上样于已用 15 mM 磷酸钠缓冲液 pH 6.0 预平衡的 CM Sepharose CL-6B 层析柱 (Pharmacia Biotech, Uppsala, Sweden) 中。氨肽酶 I 活性用 15 - 60mM 线性梯度的磷酸钠缓冲液 pH 6.0 洗脱。按 7.4 ml 的级分收集并如实施例 2 所述分析氨肽酶活性。

汇总活性级分并对 15 mM 磷酸缓冲液 pH6.0 透析。随后将透析产物上样于已用 15 mM 磷酸钠缓冲液 pH6.0 平衡的羟基磷灰石层析柱。氨肽

酶 I 活性用 15 - 300mM 线性梯度的磷酸钠缓冲液 pH 6.0 洗脱。收集级分并分析氨肽酶活性。汇总活性级分，对蒸馏水透析，冻干。

#### 实施例 4: 纯化荚膜鞘氨醇单胞菌氨肽酶 I

在起始 pH 为 7.45 的 1.5 升培养液中、在 31°C 以 250 rpm 培养荚膜鞘氨醇单胞菌 IF0 12533 共 15 小时，每升培养液含有 10g 细菌用蛋白胨、5g 酵母浸膏、3g NaCl、2g K<sub>2</sub>HPO<sub>4</sub>、0.1g MgSO<sub>4</sub> · 7H<sub>2</sub>O 和 5g 葡萄糖（分别高压灭菌）。从培养上清中纯化荚膜鞘氨醇单胞菌氨肽酶 I。

氨肽酶 I 活性用丙氨酰 - 对硝基苯胺 (Ala-pNA) 如实施例 2 所述测定。

经离心、Whatman 玻璃微纤维滤膜 (Whatman, Maidstone, England) 过滤、配备 0.22 mm 滤膜的 Nalgene Filterware 过滤、并用配备 PM 10 超滤膜的 Amicon Spiral 超滤系统浓缩而制备的培养液上清 (约 1 升)，在 10 mM 磷酸钠 pH 6.0 缓冲液中平衡至电导率和 pH 与上样缓冲液 50 mM MES pH 6.0 相同。将滤过液上样于已用 50 mM MES pH 6.0 预平衡的约含 180 ml SP- Sepharose 的 Fast Flow (Pharmacia Biotech AB, Uppsala, Sweden) 24 × 390 mm 层析柱。用所含 NaCl 经 0 而递增至 0.2 M 的 50 mM MES pH 6.0 缓冲液 240 ml 梯度洗脱具有氨肽酶 I 活性的蛋白质。汇总具有氨肽酶 I 活性的级分，经 PM 10 膜脱盐，并用 20 mM 磷酸钠 pH 7.0 缓冲液平衡。

将汇总液上样于已用 20 mM 磷酸钠 pH 7.0 缓冲液预平衡的 Pharmacia MonoQ Beads 层析柱中。具有氨肽酶 I 活性的蛋白质不与柱结合，在流过液中收集。如上述经 PM 10 膜系统浓缩流过液，将 pH 用 70 mM 醋酸钠 pH 4.0 缓冲液调至 6.0。

将浓缩的流过液上样至已用 50 mM MES pH 6.0 缓冲液预平衡的 Pharmacia Mono S 层析柱上。氨肽酶用所含 NaCl 经 0 而浓度递增至 0.2 M 的 50 mM MES pH 6.0 缓冲液 60 ml 梯度洗脱。有显著 Ala-pNA 活性的级分经汇总、如上述经 PM 10 膜系统浓缩、并用含 0.5 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 的 50 mM 磷酸 pH 7.0 缓冲液平衡。

最后，将浓缩样品上样至已用含 0.5 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 的 50 mM 磷酸 pH 7.0

缓冲液预平衡的已装好 Pharmacia Phenyl Superose 5/5 的 7 x 50 mm 柱 (Pharmacia Biotech AB, Uppsala, Sweden)。有氨肽酶 I 活性的蛋白质用所含  $(\text{NH}_4)_2\text{SO}_4$  从 0.5 M 而浓度递减至 0 的 50 mM 磷酸 pH 7.0 缓冲液 30 ml 梯度洗脱。含氨肽酶 I 活性的级分经 SDS - PAGE 分析后汇总。

对纯化荚膜鞘氨醇单胞菌氨肽酶 I 的 SDS - PAGE 分析显示了一条分子量为 67 kDa 的带。

#### 实施例 5: 荚膜鞘氨醇单胞菌氨肽酶 I (67 kDa) 的氨基酸测序

实施例 2 所述荚膜鞘氨醇单胞菌氨肽酶 I 的纯化样品在 8 - 16 % Tris-甘氨酸 SDS - 聚丙烯酰胺凝胶上电泳, 然后用 10 mM CAPS (3 - [环己基氨] - 1 - 丙烷磺酸) pH 11 的 10 % 甲醇液印迹转移 2 小时至 PVDF 膜上。PVDF 膜用 0.1 % 考马斯亮兰 R - 250 的 40 % 甲醇 / 1 % 醋酸液染色 20 秒, 在 50 % 乙醇中脱色, 观察蛋白质带。切下 67 kDa 主带, 在 Applied Biosystems Model 476A Protein Sequencer 上用印迹板 (blot cartridge) 和液相 TFA 分送法按说明书进行氨基端测序。结果发现此蛋白质的 N - 末端被封闭, 不能用 Edman 化学测序法测序。

实施例 2 所述氨肽酶 I 的纯化样品 1.0 ml 在 Savant Speed Vac AS160 上干燥, 再用 70 % 甲酸 (水溶液) 300  $\mu$ l 重新溶解。加入少量溴化氰结晶, 室温暗室温育过夜。样品在 Speed Vac 中重新干燥, 溶于 Tricine 样品缓冲液 (Novex, San Diego, CA, USA)。溴化氰裂解片段在 10 - 20 % Tricine SDS - 聚丙烯酰胺凝胶上分离为 42、30、17、15、10、6 和 4 kDa 的几条带, 印迹转移至 PVDF 膜上。切下 6、10、15、17、30 和 42 kDa 的带, 进行氨基端测序。得到 15、10 和 6 kDa 带的 N - 端序列如下:

15 kDa: AVPIAHGVPDAQDVPYYPG (SEQ ID NO:2)

10 kDa: AVNGDAYDADKLGKAITNAKTGNPGAGRPI (SEQ ID NO:2)

6 kDa: GSIGLEHHRSSSENQQEPKSLTDWAA (SEQ ID NO:2)

纯化的氨肽酶还用内切蛋白酶 Glu-C 进行部分消化, 如下: 7.5  $\mu$ l 含 2.5 % SDS 的 0.125 M Tris-HCl pH 6.7 加入 60  $\mu$ l 浓缩氨肽酶的 0.125 M Tris-HCl pH 6.7 溶液中。将样品混合, 煮沸 2 分钟, 然后在 23 $^{\circ}$ C 冷却 15 分钟。然后加入内切蛋白酶 Glu-C 测序级 (Boehringer Mannheim,

Indianapolis, IN) 的浓度为  $400 \mu\text{g/ml}$  的  $0.125 \text{ M}$  Tris-HCl pH 6.7 溶液  $10 \mu\text{l}$ 。样品在  $37^\circ\text{C}$  温育 2 小时。温育后, 加入  $2\text{X}$  Tricine SDS 样品缓冲液 (Novex, San Diego, CA, USA)  $45 \mu\text{l}$ , 将样品煮沸 5 分钟。

肽片段在  $10-20\%$  Novex Tris-tricine 胶上在还原条件下分离。切下在 40、30、25、22、20、17、10、6、5 和 4 kDa 处观察到的带。对 40、30 和 10 kDa 带的测序显示它们为混合物。17 和 22 kDa 带也是混合物, 但其序列主要是

FKDEPNPYDKARMADAKVLSLFLNSLGVTLDKDGKV (SEQ ID NO:2)

#### 实施例 6: 荚膜鞘氨醇单胞菌氨肽酶 I (67 kDa) 的定性

$100 \text{ mg/ml}$  各种酰基对硝基苯胺的二甲亚砷贮存液用  $50 \text{ mM}$  磷酸钠 pH 7.5 缓冲液稀释至浓度为  $2 \text{ mg/ml}$ 。如果底物不能完全溶解, 则使用它们的悬浮液 (以下表 2 中以\*表示)。加入含荚膜鞘氨醇单胞菌氨肽酶 I 的  $50 \text{ mM}$  磷酸钠 pH 7.0 溶液  $10 \mu\text{l}$  至 96 孔板上的  $190 \mu\text{l}$  底物溶液中启动酶与每种酰基对硝基苯胺的反应。在  $405 \text{ nm}$  及  $25^\circ\text{C}$  下在 THERMOmax Microplate Reader (Molecular Devices Corp., Sunnyvale, CA, USA) 中分析酰基对硝基苯胺的起始水解速率。结果 (表 2) 显示, 氨肽酶 I 优选水解 Ala-pNA, 也水解 Gly-pNA。但在二肽底物中, 氨肽酶 I 水解 Gly-Phe-pNA 的速率快过水解 Ala-Ala-pNA 的速率。

表 2 荚膜鞘氨醇单胞菌氨肽酶 I 的底物特异性

氨基酸酰基对硝基苯胺	相对活性 (%)
Ala	100
Met	24
Gly*	20
Leu	18.5
Glu*	6
Asp	4.5
Lys	6
Ile	0.5
Val	0.5
Phe*	0
Pro	0

67 kDa 氨肽酶 I 的最佳 pH 用以下方案确定: 向 4.25 g 醋酸钠和 3.78 g Tris (游离碱) 中加蒸馏水至终体积为 250 ml, 用 36.5 - 38% HCl 调整 pH 至 8.5, 8.0, 7.5, 7.0, 6.5, 5.80, 5.5 和 5.0, 得到各种不同 pH 的缓冲液。从每种 pH 液中取出 10 ml。将 100 mg/ml Ala-pNA 的 DMSO 溶液 20  $\mu$ l 加至 980  $\mu$ l 各种缓冲液中。加入底物后测得缓冲液的 pH 为 8.5, 8.0, 7.57, 7.30, 7.09, 6.68, 6.14 和 5.25。另配制含 2 mg/ml 底物的各 pH 溶液。将用 10 mM Tris-HCl pH 7.5 稀释 5 倍的酶溶液 10  $\mu$ l 加至不同 pH 的底物液 200  $\mu$ l 中, 室温下启动反应, 监测 405 nm 处的变化。

示于表 3 的结果证实, 氨肽酶 I 的最佳 pH 在 5.3 - 8.5 的范围内, 优选 7.0 - 7.5 (参见图 1)。在 pH 5.25 - 8.5 未测得 Ala-pNA 的自我水解。pH 5.5 时, 氨肽酶 I 催化的 Ala-pNA 水解不遵循 Michaelis-Menten 动力学 (此时表观  $K_m$  和  $V_{max}$  因酶被底物活化而为负值)。pH 7.5 时,  $K_m$  为 2.5 mM。

表 3 荚膜鞘氨醇单胞菌氨肽酶 I 的 pH 曲线

pH	相对活性 (%)
5.25	24.9
6.14	56.4
6.68	78.4
7.09	99.4
7.30	100
7.57	92.8
8.0	61
8.49	24.2

最佳温度用 Ala-pNA 作底物在 50 mM 磷酸钠 pH 7.5 缓冲液中从 30  $^{\circ}$ C 至 55  $^{\circ}$ C 进行测定。结果显示, pH 7.5 时氨肽酶 I 的最佳温度在 35  $^{\circ}$ C - 46  $^{\circ}$ C 的范围内, 更优选在 40  $^{\circ}$ C - 45  $^{\circ}$ C 范围内 (参见图 2)。

氨肽酶 I 的温度稳定性测定如下: 在 35  $^{\circ}$ C - 55  $^{\circ}$ C 的温度范围内将酶在 50 mM 磷酸钠 pH 7.5 缓冲液中温育 20 分钟, 然后以冰冷却, 如上述

用 Ala-pNA 作底物在 pH 7.5 测量残留活性。结果显示(参见图 3), pH 7.5 时氨肽酶 I 直至 40℃ 仍几乎 100% 稳定, 45℃ 时酶活性还残留近 60%, 而 50℃ 时酶彻底失活。

实施例 7: 荚膜鞘氨醇单胞菌氨肽酶 I 的粗制品和纯制品之间的 DH 增加效应比较

荚膜鞘氨醇单胞菌氨肽酶 I (67 kDa) 的 DH 增加效应用大豆蛋白质水解作用评估, 并与实施例 1 所述酶粗粉的性能进行比较。

大豆蛋白的水解程度如下测定: 如 Adler-Nissen (1986, 食物蛋白质的酶水解, Elsevier Applied Science Publishers) 所定义的 DH 用上清液与 OPA (邻苯二醛, Sigma Chemical Co., St. Louis, MO) 的反应测定, 该反应基本如 Church 等, 1983, 奶制品学杂志 66: 1219 - 1227 所述进行, 并参照 Adler-Nissen, 1979, 农业与食品化学 27: 1256 - 1262 确定的校正因子。

将氨肽酶 I 以递增剂量加至 FLAVOURZYME 1000L™ 的低本底溶液 (1.5% 底物蛋白) 或高本底溶液 (6% 底物蛋白) 和 1.5% 底物蛋白的 ALCALASE 2.4L® 溶液中。实验如下进行:

每 10 ml 反应物 50℃ 水解 18 小时。起始 pH 为 7, 水解过程中不调整 pH。酶于 85℃ 水浴中处理 3 分钟而失活。底物为大豆粉的 2% 蛋白液。底物于 85℃ 水浴处理 3 分钟。与上述 FLAVOURZYME 1000L™ 和 ALCALASE 2.4L® 联合使用的酶为荚膜鞘氨醇单胞菌氨肽酶 I 粗粉 (参见实施例 1) 和实施例 4 所述荚膜鞘氨醇单胞菌氨肽酶 I 纯化制品。

所用氨肽酶 I 的剂量根据 pH 7.5 时水解 Ala-pNA 测得的丙氨酰 - 氨肽酶单位 (AAU) 而定。规定氨肽酶 I 的一个 AAU 单位是 pH 7.5、22℃、离子强度为 0.05 时 9.1 mM Ala-pNA 溶液中每分钟水解 1.0 微摩尔 Ala-pNA 的酶用量。

氨肽酶 I 纯制品的比活性为约 12 AAU/mg 酶。因此, 例如, 每 200 mg 大豆蛋白用量为 0.24 AAU 相当于每 kg 大豆蛋白用 0.1 g 酶。

低本底 FLAVOURZYME™ 时 DH 作为 AAU 剂量的函数示于图 4, 高本底 FLAVOURZYME™ 时 DH 作为 AAU 剂量的函数示于图 5。

图 4 和 5 显示, 高剂量 AAU 时氨基酸酶 I 活性的水解产物接近饱和。酶粗制品加入低剂量 FLAVOURZYME™ 中能使 DH 从 44% 增至 72%。氨基酸酶 I 纯制品可使 DH 增至 61-62%。因此, 氨基酸酶 I 粗制品的 DH 增加效应中酶纯制品促成的占  $(62-44)/(72-44) \times 100\% = 64\%$ 。酶粗制品加入高剂量 FLAVOURZYME™ 中能使 DH 从 61% 增至 77%。氨基酸酶 I 纯制品可使 DH 增至 71%。因此, 酶粗制品的增加效应中氨基酸酶 I 纯制品促成的占  $(71-61)/(77-61) \times 100\% = 63\%$ 。

如果不加氨基酸酶 I, 可选择加更多 FLAVOURZYME™。在基本剂量 1.5% 中另加 0.5% FLAVOURZYME™ 可使 DH 从 44% 增至 48%。加入氨基酸酶 I 的效应经计算为酶粗制品效应的  $(62-48)/(72-48) \times 100\% = 58\%$ 。

相似的, 在基本剂量 6% 中另加 1% FLAVOURZYME™ 可使 DH 从 61% 增至 63%。此时氨基酸酶 I 的效应经计算为酶粗制品效应的  $(71-63)/(77-63) \times 100\% = 57\%$ 。

所得最高 DH 为 77%。DH 值基于的是总蛋白而非可溶性蛋白。蛋白质溶解度约为 85% 左右。因此, 可溶性蛋白的 DH 约为 91% 左右, 非常接近 100%。

因加入 AAU 而造成的各游离氨基酸的相对增加率 (100%) 示于表 4。Hyp、甲硫氨酰磺酸 (methioinin sulphonic acid)、Trp 因结果波动很大、不确定而未包括在表内。

表 4 因加入氨基酸酶所致的 FAA 相对增加 %

FAA	1.5% FLAVOURZYME™		6.0% FLAVOURZYME™		1.5% FLAVOURZYME™ + 67 kDa AP 相比于 6.0% FLAVOURZYME™
	+ 粗制品	+ 67 kDa AP	+ 粗制品	+ 67 kDa AP	
Asp	250.3	164.5	74.8	24.8	24.2
Glu	195.6	145.9	70.7	31.3	28.7
Asn	80.5	32.3	36.9	7.6	-15.2
Ser	149.5	69.3	78.7	41.7	12.5
Gln	198.6	138.9	51.6	28.1	9.9
Gly	353.6	305.7	117.4	92.9	90.6
His	64.0	30.3	18.8	7.5	-6.8
Arg	59.6	37.4	18.1	-31.4	-2.9
Thr	84.3	34.7	21.2	2.4	-18.3
Ala	200.2	130.3	71.2	36.5	27.3

Pro	73.2	32.4	-9.0	-11.5	-39.7
Tyr	75.4	47.6	13.8	8.5	-6.9
Val	114.1	44.0	32.5	6.3	-17.7
Met	47.9	-20.0	42.5	21.1	-46.7
Cys	5.0	12.6	17.5	9.5	6.3
Ile	133.6	48.7	34.0	7.2	-21.0
Leu	67.4	35.2	16.4	5.2	-9.1
Phe	62.6	41.8	13.3	7.2	-4.5
Lys	105.5	62.4	28.2	13.8	-1.4
total	110.1	64.5	36.9	10.3	0.0
DH	59.6	34.8	24.4	14.2	-2.2

结果显示, 当加入荚膜鞘氨醇单胞菌氨肽酶 I 的粗制品或纯制品时, Gly 显示是增加最多的氨基酸。其它显示有增加的氨基酸是 Ala, Glu, Gln, Asp 和 Ser, 但用酶粗制品时这些氨基酸的释放量表现得比用氨肽酶 I 纯制品时多, 可能是由于酶粗制品中有其它氨肽酶的存在。

应用高剂量 FLAVOURZYME™ 可产生高 DH 水解产物。应用添加了 67 kDa 氨肽酶 I 或酶粗制品的低剂量 FLAVOURZYME™ 可获得相同高水平 DH 和 FAA 释放水平。根据表 4, 应用添加了 67 kDa 氨肽酶 I 的低剂量 FLAVOURZYME™ 所得水解产物与只应用高剂量 FLAVOURZYME™ 所得的水解产物相比, 其 Gly 的水平特别高, Ala, Glu 和 Asp 的水平也高一些, 而 Met, Pro 和 Ile 的水平则更低。

#### 实施例 8: 明胶水解中荚膜鞘氨醇单胞菌氨肽酶 I 的 DH 增加效应

荚膜鞘氨醇氨肽酶 I 的 DH 增加效应在明胶中进行检验。

如实施例 7 将氨肽酶 I 以不断增加的剂量加至低或高本底的 FLAVOURZYME™ 和 ALCALASE® 中。

水解在 Eppendorf 管中的 200  $\mu$ l 反应物中进行。底物明胶 (Merck) 在 85°C 蒸馏水中经 5 分钟溶解。冷却至 50°C 后, 将 pH 调至 6.5。加入酶后将明胶终浓度调至 2%。每份反应物中底物经计算为 4 mg。

与 FLAVOURZYME™ 和 ALCALASE® 联合使用的酶为荚膜鞘氨醇氨肽酶 I 粗粉和实施例 7 中的荚膜鞘氨醇氨肽酶 I 纯制品。酶粗制品和纯制品的甘氨酸氨肽酶单位 (GAPU) 在 37°C 下、50 mM Tris-HCl pH 7.5 缓冲液



中用 Gly-pNA 作底物而测定。氨基酶 I 在 37°C 下在 50 mM Tris-HCl (pH 7.5) 中分析。将 Gly-pNA (Bachem Feinchemikalien AG, Bubendorf, Switzerland) 溶于 40% 二噁烷中, 浓度为 2.5 mg/ml, 加入 1.5 倍体积的 50 mM Tris-HCl (pH 7.5) 使底物浓度为 1 mg/ml。酶样品用相同缓冲液 500  $\mu$ l 稀释, 加入底物溶液 100  $\mu$ l, 37°C 温育 5 分钟。加入 1 M 醋酸钠 pH 4.0 缓冲液 300  $\mu$ l 终止反应。在空白对照反应中, 将终止液加入酶溶液中, 37°C 温育 5 分钟, 然后加入底物溶液。对样品和空白反应物的 410 nm 吸光度进行测量。用 pNA 在 410 nm 处的分子消光系数 9480  $M^{-1} cm^{-1}$  计算酶活性。一个 GAPU 单位定义为在 37°C、上述条件下每分钟水解 1 微摩尔 Gly-pNA 的酶用量。所用氨基酶 I 粗制品和纯制品的比活性分别为 0.725 GAPU/mg 和 6.45 GAPU/mg。按表 5 的方案, 酶用量为 4 mg 蛋白质。

表 5

管号	FLAVOURZYME™ ( $\mu$ g)	ALCALASE® ( $\mu$ g)	氨基酶 I ( $\mu$ g)
1	80 (2%)	40 (1%)	
2	"	"	
3	"	"	
4	"	"	
5	"	"	8 (0.2%) $\approx$ 0.05 U
6	"	"	31 (0.77%) $\approx$ 0.2 U
7	"	"	78 (2%) $\approx$ 0.5 U
8	"	"	155 (4%) $\approx$ 1.0 U
9	280 (7%)	40 (1%)	
10	"	"	
11	"	"	
12	"	"	
13	"	"	8 (0.2%) $\approx$ 0.05 U
14	"	"	31 (0.77%) $\approx$ 0.2 U
15	"	"	78 (2%) $\approx$ 0.5 U
16	"	"	155 (4%) $\approx$ 1.0 U

表中, 酶量 (以重量表示) 后的圆括号内为酶/底物比。

水解反应在 50°C 进行 18 小时。酶在 85°C 处理 5 分钟使失活, 随后

离心。DH 用 OPA 法依据水解产物的总蛋白含量计算。

示于图 6 和 7 的氨肽酶 I 纯制品的 DH 增加效应证实，将酶的粗制品和纯制品（0.2U/4 mg 明胶）加至低 FLAVOURZYME™ 本底中可使 DH 从 38% 分别增至 67% 和 61%。在该剂量下 DH 增加效应几乎达到饱和。

当 FLAVOURZYME™ 的用量增加至 7% 时，对照 DH 增至 43%。但用氨肽酶 I 纯制品得到的 DH 为 61%，这几乎与低 FLAVOURZYME™ 本底时的 DH 相同。

这些结果说明，荚膜鞘氨醇氨肽酶 I 对蛋白质的水解可达到很高的 DH 水平。而且，应用氨肽酶 I 后，可减少其它蛋白水解酶的用量而获得相同水平的 DH。

实施例 9：用于荚膜鞘氨醇单胞菌 B46 氨肽酶 I 基因鉴定的探针的制备

根据以下所示荚膜鞘氨醇单胞菌 B46 氨肽酶 I 的内部氨基酸序列，合成简并引物以便通过聚合酶链式反应（PCR）扩增探针，用于鉴定荚膜鞘氨醇单胞菌 B46 氨肽酶 I 基因。

内部肽 AB0713: AVNGDAYDADKLKGAITNAKTGNPGAGRPI (SEQ ID NO:2)

内部肽 AB0781: FKDEPNPYDKARMADAKVLSLFLNSLGVTLDDKDGKV (SEQ ID NO:2)

以下所示、命名为 550-39-1、550-23-4 和 550-39-2 的引物将用于下述扩增反应中。

550-23-4: 5'-gcrtctangertcnc-3' (SEQ ID NO:3)

550-39-1: 5'-aargaygarccnaaycc-3' (SEQ ID NO:4)

550-39-2: 5'-acytтыtctctcttctc-3' (SEQ ID NO:5)

制备体积为 50 μl 的扩增反应物，其中有 50 pmol 引物 550-39-1 和 550-23-4 或者 550-39-1 和 550-39-2，1 μg 荚膜鞘氨醇单胞菌 B46 染色体 DNA 为模板，1X PCR 缓冲液（Perkin-Elmer, Foster City, CA），dATP、dCTP、dGTP、dTTP 各 200 μM，以及 AmpliTaq Gold（Perkin-Elmer, Foster City, CA）0.5 U。荚膜鞘氨醇单胞菌 B46 染色体 DNA 用《Qiagen 基因组手册》（Qiagen, Inc., Chatsworth, CA）中所述细菌分离方案进行分离。

反应物在 Stratagene Robocycler 40 (Stratagene, La Jolla, CA) 中温育, 程序设定为 95°C 10 分钟的 1 个循环, 95°C 1 分钟、44°C 1 分钟、72°C 1 分钟 35 循环, 72°C 7 分钟的 1 个循环。

用引物 550-39-1 和 550-39-2 进行的扩增产生一段 100 bp 的产物, 命名为 100 bp, 用引物 550-39-1 和 550-23-4 进行的扩增产生一段 191bp 的产物, 命名为 191bp。采用 TOPO/TA 克隆试剂盒 (Invitrogen, Inc., La Jolla, CA) 按说明书将两种 PCR 产物分别克隆至载体 pCR2.1/TOPO 中。Applied Biosystems Model 377 Sequencer (Applied Biosystems, Foster City, CA) 测序结果显示, 100 bp 产物的序列包含在 191 bp 产物内, 这表明氨基酸序列 AB0713 位于氨基酸序列 AB0781 上游约 30 个氨基酸处。

下述命名为 550-71-2 和 550-79-1 的一套非简并性引物随后被用于 PCR 扩增 DIG 标记的 191 bp 之探针, 扩增用 Genius System PCR DIG 探针合成试剂盒 (Boehringer-Mannheim Corporation, Indianapolis, IN) 按说明书在 Stratagene Robocycler 40 (Stratagene, La Jolla, CA) 上进行, 程序设定为 95°C 2 分钟的 1 个循环, 95°C 1 分钟、56°C 1 分钟、72°C 1 分钟的 25 循环, 72°C 7 分钟的 1 个循环。

550-71-2: 5'-cttttcgtccttgccagc-3' (SEQ ID NO:6)

550-79-1: 5'-gcgtcatatgcgtctcc-3' (SEQ ID NO:7)

#### 实施例 10: 染色体文库的筛选

用实施例 9 所述探针 191 bp 筛选荚膜鞘氨醇单胞菌 B46 的染色体文库。文库的构建通过将荚膜鞘氨醇单胞菌 B46 染色体 DNA 的 *Sau* 3A 部分消化片段 (5-7kb) 连接至载体 pSJ1678 的 *Bam* HI 位点而完成。用染色体文库转化大肠杆菌 XL1 Blue MR (Stratagene, La Jolla, CA), 用 Genius DIG 非放射性核酸标记及检测系统 (Boehringer Mannheim Biochemicals, Indianapolis, IN) 按说明书制备 DIG 标记的 191 bp 探针, 用此探针筛选菌落浸提物 (lift)。筛选了近 5000 个菌落后, 找到 5 个可与探针杂交的菌落, 所得质粒命名为 pMRT004.1-7、pMRT004.1-14、pMRT004.1-15、pMRT004.1-16 和 pMRT004.1-17。用 QIAGEN Miniprep 质粒试剂盒 (Qiagen, Chatsworth, CA) 制备所有阳性克隆的质粒 DNA。用限制性内切酶 *Pst* I

对质粒 DNA 进行的限制性分析在 1% 琼脂糖凝胶上显示, 质粒 pMRT004.1-7 包含一段近 6 kb 的插入片段, 质粒 pMRT004.1-14 包含一段近 5.6 kb 的插入片段, 质粒 pMRT004.1-15 包含一段近 8.5 kb 的插入片段, 质粒 pMRT004.1-16 包含一段近 8.6 kb 的插入片段, 质粒 pMRT004.1-17 包含一段近 7.2 kb 的插入片段。此凝胶进一步用 DIG 标记的 191 bp 探针采用 Genius DIG 非放射性核酸标记及检测系统按说明书进行 Southern 杂交分析。结果说明, DIG 标记的 191 bp 探针与 pMRT004.1-7、pMRT004.1-14、pMRT004.1-16、pMRT004.1-17 中的约 3500 bp 片段能杂交, 且只与 pMRT004.1-15 中的一段约 1600 bp 片段杂交。

pMRT004.1-7 和 pMRT004.1-14 的质粒 DNA 用 QIAGEN Maxiprep 质粒试剂盒 (Qiagen, Chatsworth, CA) 制备。随后对 pMRT004.1-7 的 DNA 进行测序。DNA 测序在 Applied Biosystems Model 377 XL 自动 DNA 测序仪上用染料-终止物化学法进行。用转座子插入方案 (Primer Island Transposition Kit, Perkin-Elmer/Applied Biosystems, Inc., Foster City, CA) 按说明书产生相邻序列。结果说明, 该克隆丢失了蛋白质 N-端的氨肽酶 I 基因的近三分之一。根据 *Ava* I、*Sph* I 和 *Eco* RI 的限制性分析, pMRT004.1-14 克隆似乎包含完整的氨肽酶 I 基因, 因此通过对该克隆测序而获得氨肽酶 I 基因的其余序列。

#### 实施例 11: 荚膜鞘氨醇单胞菌氨肽酶 I 基因经 PCR 进行分离

采用 PCR 从实施例 10 所述的 pMRT004.1-14 克隆以 *Nru* I/*Bgl* II 片段形式分离荚膜鞘氨醇单胞菌氨肽酶 I 基因。扩增用引物 591-35-1 和 591-36-1 进行如下。扩增反应物制备成体积 50  $\mu$ l, 其中含有 50 pmol 引物 591-35-1 和 591-36-1, 10 ng pMRT004.1-14 质粒 DNA 为模板, 1X PCR 缓冲液 (Perkin-Elmer, Foster City, CA), dATP、dCTP、dGTP、dTTP 各 200  $\mu$ M, 1  $\mu$ l DMSO 以及 AmpliTaq Gold (Perkin-Elmer, Foster City, CA) 0.5 U。反应物在 Stratagene Robocycler 40 (Stratagene, La Jolla, CA) 中温育, 程序设定为 95 $^{\circ}$ C 10 分钟的 1 个循环, 95 $^{\circ}$ C 1.5 分钟、55 $^{\circ}$ C 1.5 分钟、72 $^{\circ}$ C 3 分钟的 30 循环, 72 $^{\circ}$ C 7 分钟的 1 个循环。

591-35-1: 5'-cgaatcggccagatctccatcg-3' (SEQ ID NO:8)

591-36-1: 5'-gatcctggcgcagctcgcgaaggcgcagg-3' (SEQ ID NO:9)

扩增产生一段2100 bp的PCR产物。如已述,采用Invitrogen TOPO/TA克隆试剂盒按说明书将该2100 bp产物克隆至载体pCR2.1/TOPO中并测序。这一新克隆记为pMRT010-4。比较pMRT010-4的推测氨基酸序列与pMRT004.1-14的氨肽酶I的已知氨基酸序列,表明它们相同。但在DNA水平上,pMRT010-4在第170位氨基酸处包含一个沉默点突变,将编码甘氨酸的ggt密码子变为仍然编码甘氨酸的ggc密码子。

氨肽酶I克隆有一段2019 bp的开放读框,它编码一段含673个氨基酸的多肽。其核苷酸序列(SEQ ID NO: 1)和推测的氨基酸序列(SEQ ID NO: 2)示于图9。用SignalP程序(Nielsen等,1997,蛋白质工程10: 1-6)结合分别为0.42, 0.34, 0.95和0.55的最大C、最大Y、最大S和平均S的域值,测知有一段32个残基的信号肽,从而指示分泌的氨肽酶I分子量近70,600道尔顿。因此,成熟氨肽酶I包含641个氨基酸。

氨肽酶序列的对应比较用BLAST 2.0蛋白质数据库查询程序(Altschul等,1997,核酸研究25: 3389-3402)结合以下参数进行:  
blastall -p blastp -a 4 -e 10 -E 0 -v 500 -b 250 -I[查询卷号] -d prot\_all,其中-p指程序名,-a指所用服务器数,-e指期望值,-E指扩展一个缺口的代价,-v指单线描述数,-b指将显示的对比较数,-I指查询文档,-d指用于查询的数据库。

BLAST查询揭示,荚膜鞘氨醇单胞菌氨肽酶I与集胞蓝细菌的推测67 kDa蛋白质(录入号Q55449)有23%的相同区。

实施例12: 将荚膜鞘氨醇单胞菌氨肽酶I基因克隆至芽孢杆菌表达载体pPL2419CAsub2-P<sub>ter</sub>/498

在枯草芽孢杆菌中表达并分泌氨肽酶I的策略是,将编码氨肽酶成熟区的基因片段克隆在芽孢杆菌整合载体中强启动子和革兰氏阳性信号序列之后,处于同一个读框中。选择质粒pCAsub2-P<sub>ter</sub>/498进行表达和分泌,这是因为它编码强启动子P<sub>ter</sub>以及芽孢杆菌蛋白酶PD498的革兰氏阳性信号序列(美国专利第5,621,089号)。该质粒已按如下方法构建:用

*SphI* 消化质粒 p498-5 (美国专利第 5,621,089), 将包含  $P_{ter}/498$  表达盒的 2500 bp 片段克隆至 pIC20R (Marsh 等, 1984, 基因 32: 481-485) 的 *SphI* 位点, 产生 pIC20R- $P_{ter}/498$ 。然后用 *HindIII* 和 *BamHI* 对 pIC20R- $P_{ter}/498$  进行消化, 将包含  $P_{ter}/498$  表达盒的 2500 bp 片段克隆至 pSJ2882 (WO 98/22598) 的 *HindIII/BamHI* 位点, 产生 pSJ2882- $P_{ter}/498$ 。用 *HindIII* 消化 pSJ2882- $P_{ter}/498$ , 经 Klenow 片段处理产生平端, 再用 *BamHI* 消化。然后将包含  $P_{ter}/498$  表达盒的 2500 bp 片段克隆至整合型载体 pCAsub2 (WO 98/22598) 的 *NotI* 位点 (用 Klenow 片段填平) / *BamHI* 位点, 产生 pCAsub2- $P_{ter}/498$ 。

为构建包含氨肽酶 I 基因的最终整合载体, 用 *MscI* 和 *BglII* 消化 pCAsub2- $P_{ter}/498$  的质粒 DNA, 用 *NruI* 和 *BglII* 消化 pMRT010-4 的质粒 DNA。pCAsub2- $P_{ter}/498$  的 8000 bp *MscI/BglII* 片段和 pMRT010-4 的 2100 bp *NruI/BglII* 片段在 1% 琼脂糖凝胶上分离, 用 QIAGEN 琼脂糖 DNA 凝胶提取试剂盒 II (Qiagen, Inc., Chatsworth, CA) 按说明书提取。用快速 DNA 连接试剂盒 (Boehringer-Mannheim Corp., Indianapolis, IN) 连接这两个片段, 产生整合载体 pMRT014-1 (图 10), 按 Petit 等, 1990, 出处同上的方法用它直接转化枯草芽孢杆菌 PL1801 *spoIIE* 细胞。筛选转化子的氯霉素抗性。枯草芽孢杆菌 PL1801 *spoIIE* 是枯草芽孢杆菌 168 (芽孢杆菌保藏中心, Columbus, OH) 缺失了基因 *apr* 和 *npr*, 并在 *spoIIE* 基因中插入了 Tn917 的变株。将转化子铺板于含氯霉素 5  $\mu$ g/ml 的 Tryptone blood agar base (TBAB) 平板上, 37 $^{\circ}$ C 温育过夜。命名为枯草芽孢杆菌 1801IIIe:: pMRT014-1 的转化子的基因组 DNA 用《Qiagen 基因组手册》一书 (Qiagen, Inc., Chatsworth, CA) 中所述细菌分离方案进行分离。将枯草芽孢杆菌 1801IIIe:: pMRT014-1 的基因组 DNA 如上述转化至枯草芽孢杆菌 164 $\Delta$ 5 (WO 98/22598) 中。一个命名为枯草芽孢杆菌 A164 $\Delta$ 5:: pMRT014-1 的转化子在含氯霉素 30  $\mu$ g/ml 的 TBAB 琼脂平板上扩增培养 3 次。

实施例 13: 荚膜鞘氨醇单胞菌 B46 氨肽酶 I 基因在枯草芽孢杆菌中的表达

在含 50ml PS-1 培养基的双份振荡烧瓶中、37℃、250rpm 振荡培养枯草芽孢杆菌 A164Δ5::pMRT014-1 共 4 天，该培养基的成分是：10% 蔗糖、4% 豆粉、0.42% 无水磷酸氢二钠、0.01% pluronic acid、0.5% 碳酸钙。此外，用包含整合载体的枯草芽孢杆菌 A164Δ5::pCAsub2 作为阴性对照。

在 24、48 和 72 小时的时间点从每个烧瓶中取 1 ml 样品。铺板于 Luria-Bertani (LB) 平板，再以小片贴养于含氯霉素 30 μg/ml 的 TBAB 上，以证实氨肽酶 I 整合的活力和稳定性。在每例中，贴养于 TBAB 平板的菌落的过夜生长显示，整合为 100% 稳定。

从每个烧瓶中取 12 μl 上清，用 Novex 8-16% Tris-甘氨酸预制胶(上样孔为 1.0 mm ×12 孔)和 Novex DryEase Mini 凝胶干燥系统(Novel Experimental Technology, San Diego, CA) 按说明书进行 SDS-PAGE 分析，确定表达水平。被检菌株的表达水平还可以 L-Ala-pNA 为底物通过氨肽酶平板试验进行分光光度法测定。SDS-PAGE 凝胶数据和分光光度法测定的数据均证实，枯草芽孢杆菌 A164Δ5::pMRT014-1 表达了氨肽酶 I，而且在 48 小时的表达水平最高。

### 生物材料的保藏

以下生物材料已遵循 Budapest 条约的规定，保藏在 Agricultural Research Service Patent Culture Collection, Northern Regional Research Center, 1815 University Street, Peoria, Illinois, 61604, 保藏号如下：

保藏物	保藏号	保藏日期
大肠杆菌 XL1 Blue MR	pMRT004.1-14 NRRL B-30032	1998 年 6 月 10 日

本文所描述和要求保护的发明并不限定在所公开的具体实施方案的范围内，因为这些实施方案意在举例说明本发明的几个方面。任何等价的实施方案均包括在本发明的范围内。事实上，除了本文所示和所述，对本发明的修改方案在参照前面的描述后对本领域内技术人员而言是显

---

而易见的。这些修改也包含在所附权利要求书的范围内。有冲突时，以本文公开内容，包括定义为准。

本文引用了很多文献，它们均全文引入作为参考。



## 序列表

<110> Blinkovsky, Alexander

Byun, Tony S.

Klotz, Alan V1

Sloma, Alan

Brown, Kimberly

Tang, Maria

Fujii, Mikio

Marumoto, Chigusa

<120> 具有氨肽酶活性的多肽及其编码核酸

<130> 5379.200-US

<150> 60/069,719

<151> 1997-12-16

<150> 1465/97

<151> 1997-12-16

<150> PA 1998 00670

<151> 1998-05-15

<160> 9

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 3000

<212> DNA

<213> 鞘氨醇单胞菌

<400> 1

atcggccatg	accagcagtt	ccacgtcggc	ctgceggttcg	gccgcgccga	tcacttcggc	60
atcgatctgg	tgaaactggc	gatagcggcc	cttttgcggg	cgctcatagc	ggaacagggc	120
cccgtgcgtc	gcgatcctga	gcggggcgtg	ctgctgccag	ccattggtga	gataggcgcg	180
ggcgagggcg	gcggtgaatt	cgggcccag	cgtcagcgat	tcgccgccgc	gatcctcgaa	240
cgatacatt	tccttcgata	ccacgtcggg	ggtttcgccc	agcgagcgcg	agaacaccgt	300
ggtcttttcg	aagaccggca	ttccaccgc	gcgaaagcga	tagagcttgc	gcacgcgctc	360
gaacgtttcc	acgacatggc	caaaggcctc	ggcctcaagc	ccgaaaatgt	cctgggtgcc	420
acgaatagcc	ttgggtgtcg	gaatggtttg	cttgctcatg	gcgcgcgcg	atagcggctt	480
tcgcgcggtg	ggggaagcat	ccgtggcgat	ccgcgcgta	ttgtgcccat	ctggccggtg	540
cagatgagac	cgcttgggcg	catgaaaggc	gcatgvcgca	aaacccccca	aggcattggc	600
ctgctttccg	ccctttccac	gtccaccttg	gcacttgcca	ccctgatcct	ggcgcagccc	660
ggcctggcgc	aggtgcagcc	ggcgagcaac	agccgcccg	tggcagtgcc	gatcgctcat	720
gggggtgccg	atgcgcagga	cgtgccctat	cccggcacga	tcggggtgca	gatcgatgcc	780
accgatctgg	ccaccggggc	gttcggggtg	gtggaaccgc	tgccgggtgg	ggccgatgcc	840
aaggaactga	tcctgcaact	gccggcctgg	ctgcggggtg	agcatggcaa	tcgcccggcc	900

gtggccgagc	tggccggcat	cacgtttgaa	gccaaagggcc	agaagctggc	ctggaccgpc	960
gacccgggtg	aagtgaacgc	gttccacatc	cccctgcccg	ccggcaccag	cgaagtgggtg	1020
gcccgtttca	tccacacctc	gccgctgccc	gacagcgaag	gccgcatcac	cgttacgcgc	1080
gaaatgctca	acgtgcagtg	ggagaagatg	agcctctatc	ccgcccgtca	ctatgtgcgg	1140
cagatcaag	tgcgtcccac	cgtcagcttc	ccgcagggct	ggaccgtggt	caccgcgctg	1200
gatggcaaga	cgcagagcgg	cgcgggcaat	accgtgacct	gggcccgaac	cgactatgaa	1260
accctggctg	attcgccgat	ctttgcccgg	ctctatgccc	cgcgccatga	tctgggcccac	1320
aacgtctatt	tccgatctggt	ggccgacaag	cccagctgct	tggcgatcaa	gccggaaaac	1380
ctggcccgcct	atcgcaacct	ggccgacgaa	gccgtggggc	cattcgggcg	gcgccatttc	1440
gatcactacg	atttcctgct	cgcgctgacc	gatcgcatgg	gcagcatcgg	cctggaaacac	1500
caccgtttcca	gcgaaaacca	gcaggaacct	aagagcctga	ccgactgggc	cgccatgac	1560
tgggaccgca	acgtgatcgc	ccacgaattc	agccacagct	gggatggcaa	gtatcgccgc	1620
tccggccaagc	tgtggacgcc	cgactatcgc	cagccgatgc	aggacaacct	gctgtgggtc	1680
tatgaagggc	agacgcagtt	ctggggcctg	gtcctggccg	cagctcggg	cggtgcagagc	1740
aaggacgtgg	tcttgggcag	cctcgccaac	tatgcggca	cgttcaccca	gaccgcggg	1800
cgcgactggc	gctcgggtga	agacacgacg	atggatccca	tcttcgccc	ccgcaagccc	1860
aagccctatt	cctcgtttac	ccgtaacgag	gactattaca	ccgaaggcgc	gctggtgtgg	1920
ctggaagcgg	accagatcat	ccgcgatggc	accggcggca	agaagggcct	ggatgatttc	1980
gccaaaggcgt	tctttggcgt	gcgcgacggc	gattggggcg	tgctgacct	tgaattcgat	2040
gacgtggtca	agaccctcaa	cgcgctctat	ccctatgact	gggcccagtt	cctcaagacc	2100
cgctcgcaga	cgcccggcca	gccggtgccc	ctcggcggga	tccagcgcgg	cggtcacaag	2160
ctggaattca	aggacgagcc	caaccctat	gacaaggcgc	gcatggccga	tgccaaggtg	2220
ctcagcctgt	tcaactcgct	ggcgctgacg	ctggacaagg	acggcaaat	caccgcctcg	2280
cgctgggatg	gcccggcgtt	caaggcgggg	ctggtttcgg	gcatgcaggt	gatggccctg	2340
aacggcgacg	cctatgacgc	ggacaagctc	aagggcgcga	tcaccaatgc	caagaccggc	2400
aaccccggcg	ccggccgccc	gatcgaactg	ctggtcaagc	gtgacgatcg	ctttgtcacg	2460
ctgccgatca	cctatgcca	tggcctgccc	tggcctggc	tggtgcgcac	ggcgccggg	2520
acggcaccga	ccgggctgga	caagctgctg	gccccgcacg	ccagcaagct	gcccgtgggc	2580
aaggctgcca	agtgatgtca	ggggccgggc	caagcctgca	tctgggccc	ctggccccac	2640
cccgttcat	cctgttctctg	gtgctgctga	tgggcccgcac	cggggctggtg	tgggtgtggc	2700
atccccctac	ccgcaacagt	ggcgaactgg	ccgattcgct	ggccatgggc	tttgattttg	2760
cgggcctggc	attcctgttg	tccgtggtgc	cactgctgcg	ctgtgcccga	gccgatacga	2820
tgcggcaaa	cgcggtggac	aacgatgcc	accgcgtgct	ggtgctgtgc	atcaccaccg	2880
ttctgaccg	agtggctgatg	gcacgatcgc	ccggcgagct	gccccgcgcg	gcacatggcg	2940
attcgctggc	aaagctccgg	ctgatcggga	cgctgggtct	gacctggctt	tttgccaaca	3000

<210> 2

<211> 673

<212> PRT

<213> 鞘氨醇单胞菌

<400> 2

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

Glu Lys Met Ser Leu Tyr Pro Ala Gly His Tyr Val Arg Gln Ile Lys  
 180 185 190  
 Val Arg Pro Thr Val Ser Phe Pro Gln Gly Trp Thr Val Phe Thr Ala  
 195 200 205  
 Leu Asp Gly Lys Thr Gln Ser Gly Ala Gly Asn Thr Val Thr Trp Ala  
 210 215 220  
 Glu Thr Asp Tyr Glu Thr Leu Val Asp Ser Pro Ile Phe Ala Gly Leu  
 225 230 235 240  
 Tyr Ala Ala Arg His Asp Leu Gly His Asn Val Tyr Phe Asp Leu Val  
 245 250 255  
 Ala Asp Lys Pro Glu Leu Leu Ala Ile Lys Pro Glu Asn Leu Ala Ala  
 260 265 270  
 Tyr Arg Asn Leu Ala Asp Glu Ala Val Gly Ala Phe Gly Ala Arg His  
 275 280 285  
 Phe Asp His Tyr Asp Phe Leu Leu Ala Leu Thr Asp Arg Met Gly Ser  
 290 295 300  
 Ile Gly Leu Glu His His Arg Ser Ser Glu Asn Gln Gln Glu Pro Lys  
 305 310 315 320  
 Ser Leu Thr Asp Trp Ala Ala Tyr Asp Trp Asp Arg Asn Val Ile Ala  
 325 330 335  
 His Glu Phe Ser His Ser Trp Asp Gly Lys Tyr Arg Arg Ser Ala Lys  
 340 345 350  
 Leu Trp Thr Pro Asp Tyr Arg Gln Pro Met Gln Asp Asn Leu Leu Trp  
 355 360 365  
 Val Tyr Glu Gly Gln Thr Gln Phe Trp Gly Leu Val Leu Ala Ala Arg  
 370 375 380  
 Ser Gly Val Gln Ser Lys Asp Val Val Leu Gly Ser Leu Ala Asn Tyr  
 385 390 395 400  
 Ala Gly Thr Phe Thr Gln Thr Ala Gly Arg Asp Trp Arg Ser Val Glu  
 405 410 415  
 Asp Thr Thr Met Asp Pro Ile Phe Ala Ala Arg Lys Pro Lys Pro Tyr  
 420 425 430  
 Ser Ser Leu Thr Arg Asn Glu Asp Tyr Tyr Thr Glu Gly Ala Leu Val  
 435 440 445  
 Trp Leu Glu Ala Asp Gln Ile Ile Arg Asp Gly Thr Gly Gly Lys Lys  
 450 455 460  
 Gly Leu Asp Asp Phe Ala Lys Ala Phe Phe Gly Val Arg Asp Gly Asp  
 465 470 475 480  
 Trp Gly Val Leu Thr Tyr Glu Phe Asp Asp Val Val Lys Thr Leu Asn  
 485 490 495  
 Gly Val Tyr Pro Tyr Asp Trp Ala Thr Phe Leu Lys Thr Arg Leu Gln  
 500 505 510  
 Thr Pro Gly Gln Pro Val Pro Leu Gly Gly Ile Glu Arg Gly Tyr  
 515 520 525  
 Lys Leu Glu Phe Lys Asp Glu Pro Asn Pro Tyr Asp Lys Ala Arg Met  
 530 535 540  
 Ala Asp Ala Lys Val Leu Ser Leu Phe Asn Ser Leu Gly Val Thr Leu  
 545 550 555 560  
 Asp Lys Asp Gly Lys Val Thr Ala Ser Arg Trp Asp Gly Pro Ala Phe  
 565 570 575  
 Lys Ala Gly Leu Val Ser Gly Met Gln Val Met Ala Val Asn Gly Asp  
 580 585 590  
 Ala Tyr Asp Ala Asp Lys Leu Lys Gly Ala Ile Thr Asn Ala Lys Thr  
 595 600 605  
 Gly Asn Pro Gly Ala Gly Arg Pro Ile Glu Leu Leu Val Lys Arg Asp  
 610 615 620  
 Asp Arg Phe Val Thr Leu Pro Ile Thr Tyr Ala Asp Gly Leu Arg Trp  
 625 630 635 640  
 Pro Trp Leu Val Arg Thr Ala Pro Gly Thr Ala Pro Thr Gly Leu Asp  
 645 650 655  
 Lys Leu Leu Ala Pro His Ala Ser Lys Leu Pro Val Gly Lys Ala Ala  
 660 665 670  
 Lys

<210> 3

<211> 17

<212> DNA

<213> 鞘氨醇单胞菌

<400> 3		
gertcrtang crtcncc		17
<210> 4		
<211> 17		
<212> DNA		
<213> 鞘氨醇单胞菌		
<400> 4		
aargaygarc cnaaycc		17
<210> 5		
<211> 17		
<212> DNA		
<213> 鞘氨醇单胞菌		
<400> 5		
acyttyttert cyttrtc		17
<210> 6		
<211> 19		
<212> DNA		
<213> 鞘氨醇单胞菌		
<400> 6		
cttttcgtcc ttgtccagc		19
<210> 7		
<211> 17		
<212> DNA		
<213> 鞘氨醇单胞菌		
<400> 7		
gcgtcatatg cgtctcc		17
<210> 8		
<211> 22		

<212> DNA

<213> 鞘氨醇单胞菌

<400> 8

cgaatcggcc agatctccat cg

22

<210> 9

<211> 29

<212> DNA

<213> 鞘氨醇单胞菌

<400> 9

gatcctggcg cagctcgca aggcgcagg

29

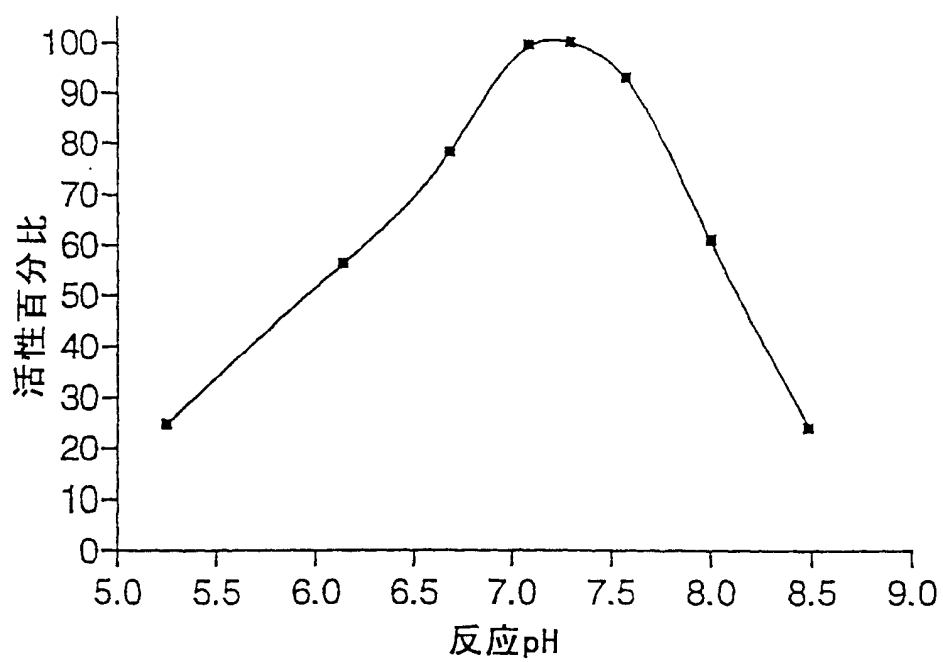


图 1

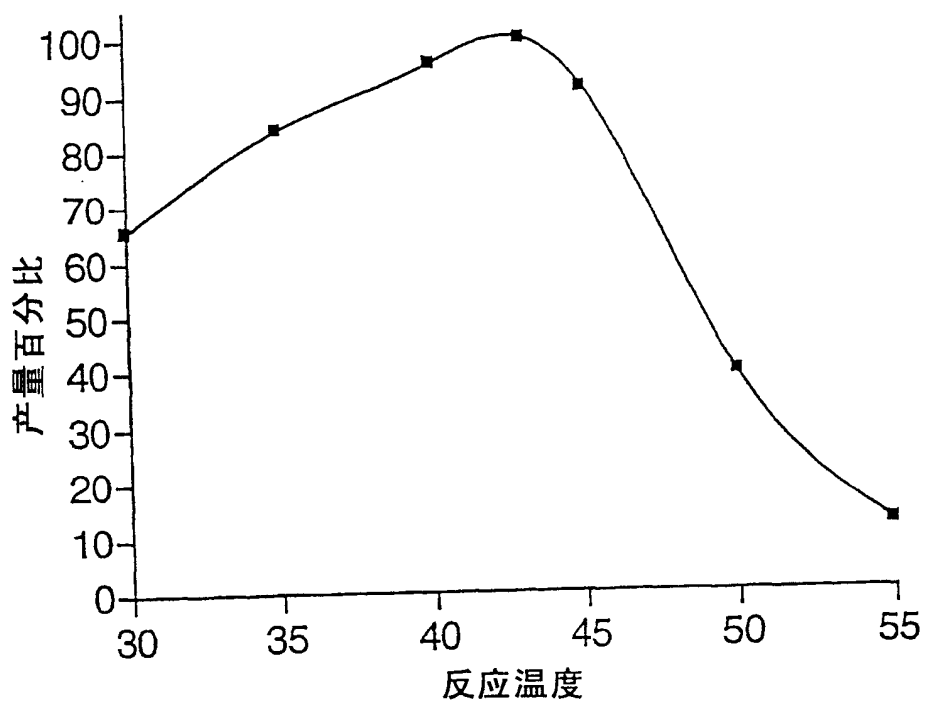


图 2

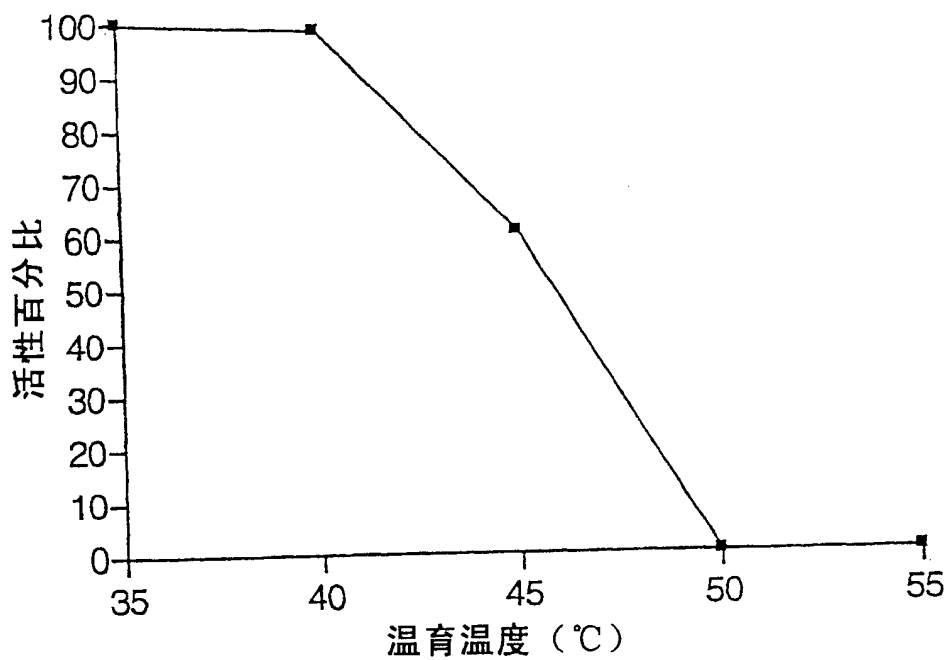


图 3



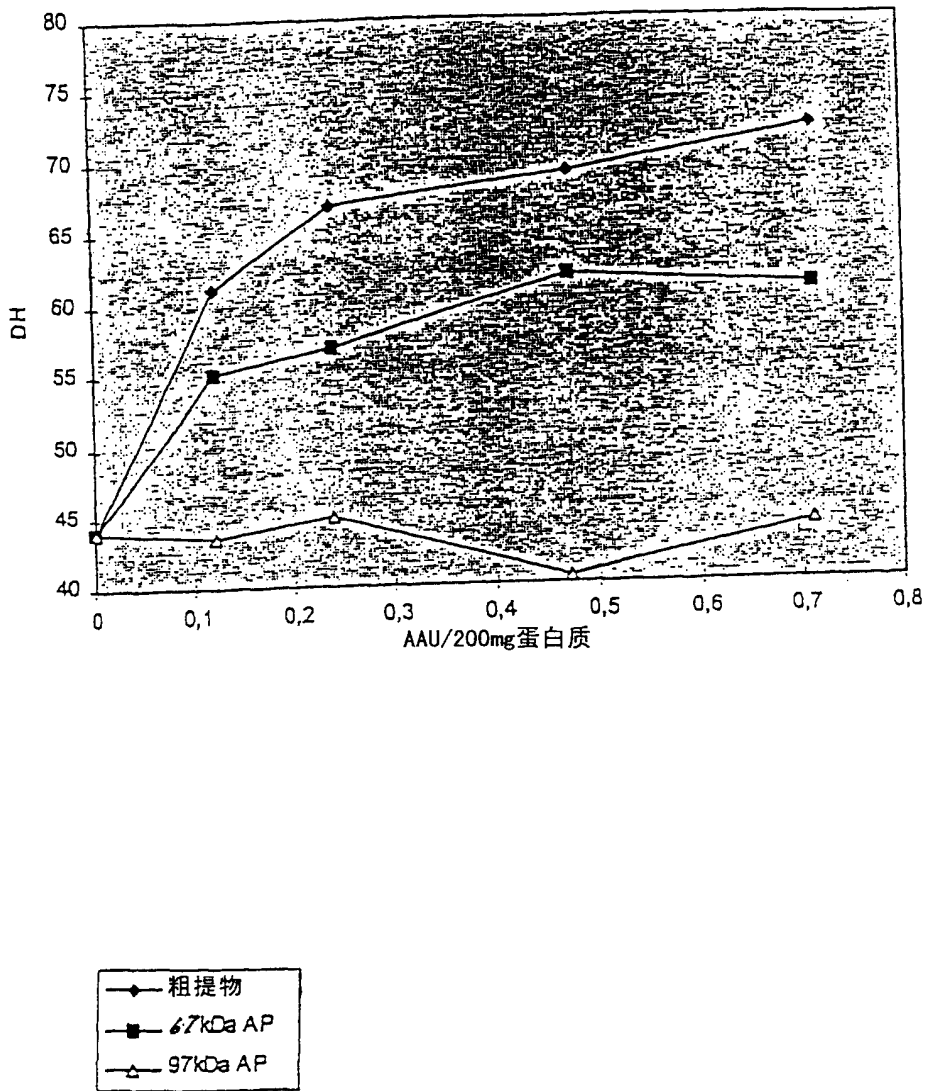


图 4

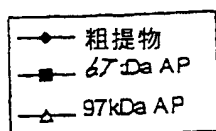
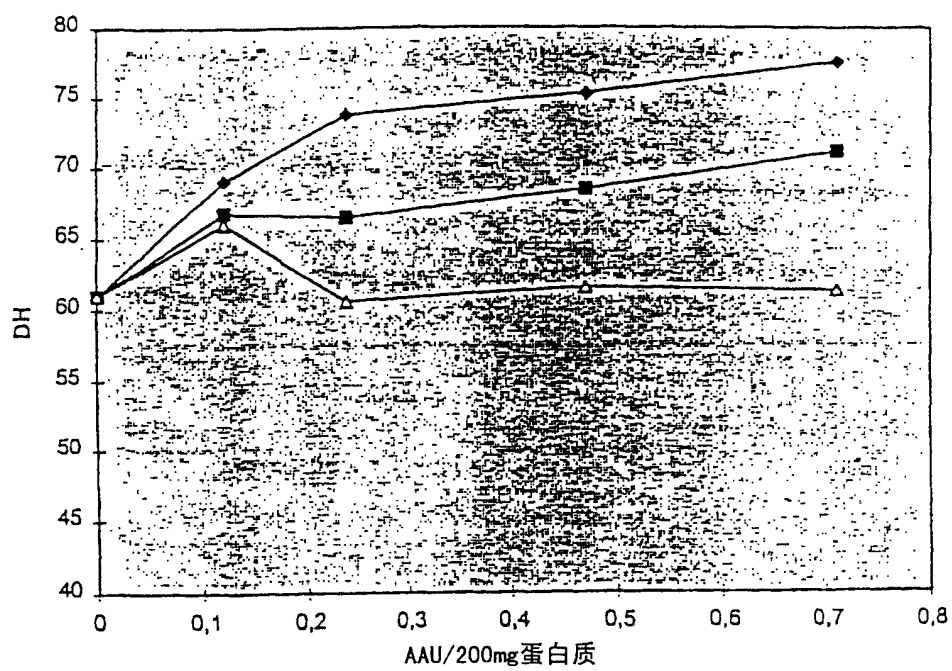


图 5

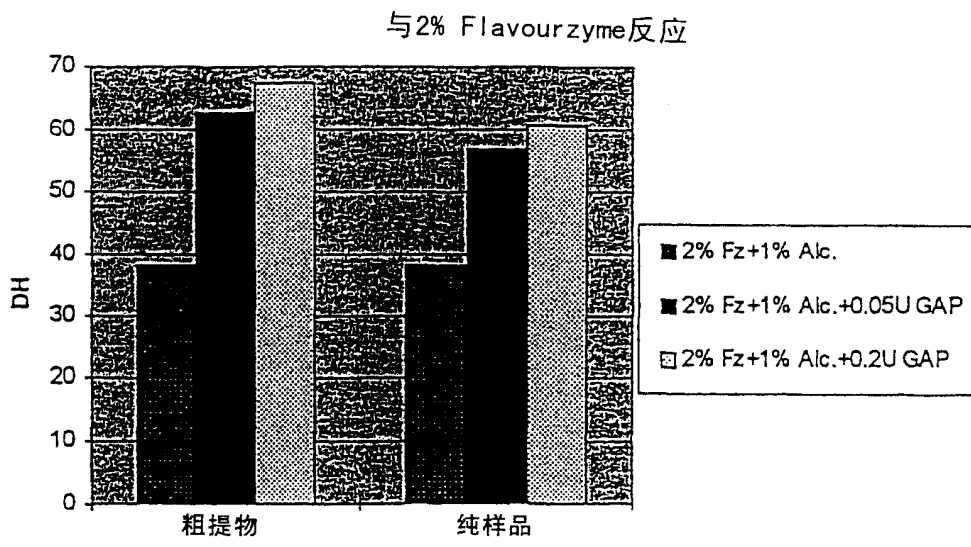


图 6

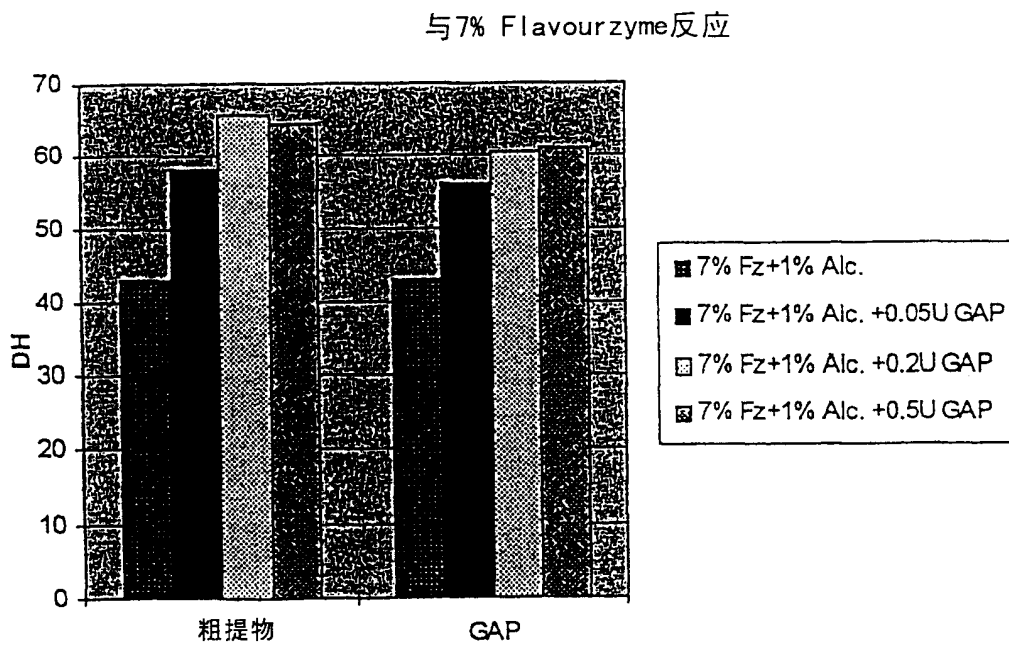


图7

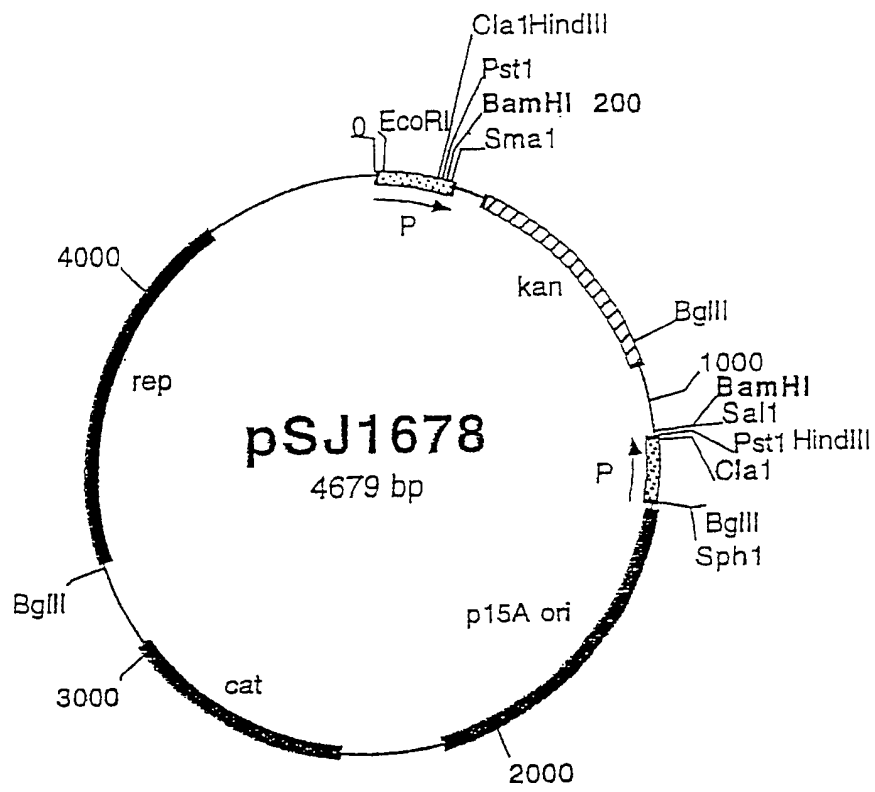


图 8

ATCGGCCATGACCAGCAGTTCCACGTCGGCCCTGCGGTTTCGGCCGGCCGATCATTTCGGCATCGATCTGG 70  
 TGAAAACGTGGCGGATAGGGCCCTTTTTCGGGGCGCTCATAGCGGAACAGGGCCCCCGTCCGCGATCTTGA 140  
 GCGGGCGTGTCTGCCAGCCATTGGTGAGATAGCGGGCGGAGGCGGGCGGTGAATTTCGGGCCCGCAG 210  
 CGTCAGCGATTTCGGCCCGCGGATCCTCGAACGAATACATTTCCATCCGATACCACGTCGGTGGTTTCGCC 280  
 AGCGAGCGGAGAACACCCGTGGTCTTTTCGAAGACCCGGCATTTCCACCCCGCGAAAGCGGATAGAGCTTGC 350  
 GCACGGCTCGAACGTTTCCACGACATGGCCCAAGGCCCTCGGCCCTCAAGCCCCGAAAATGTCTCCCTGGGTGCC 420  
 ACGAATAGCCTTGGGTGTTCGGAATGGTTTGTCTGCTCATGGCGCGCGGATAGCGGCTTTTCGGCCGGGTG 490  
 GGGGAAGCATCCGTGGCGATCCCGCCCGGTAATGTGCCCATCTGGCCGTTGCAGATGAGACCCTTGGCGG 560  
 CATGAAAGGGCCCATGCGCAAAAACCCCAAGGCATTGGCCCTTCCGCCCTTTCCACGTCACACCTTG 630  
 M R K T P Q G I G L S A L S T S T L  
 GCAC TTGCCACCCCTGATCCTGGCGCAGCCCGCTGGCGCAGGTGCAGCCGGCGGAGCAACAGCCGCCCGA 700  
 A L A T L I L A Q P A L A Q V Q P A S N S R P  
 TGGCAGTCCCGATCGCTCATGGGTGCCCGATGCCGACGACGTCGCCCTATCCCGGCACGATCGGGCTGCA 770  
 M A V P I A H G V P D A Q D V P Y P G T I G L Q  
 GATCGATGCCACCGATCTGGCCACCGGGCGTTCCGGGTGGTGGAAAACCGTGCCTGGCGGGCGGATGCC 840  
 I D A T D L A T G A F R V V E T V P V A D A  
 AAGGAACTGATCCTGCAACTGCCGGCCTGGCTGCCGGGTGAGCATGGCAATCGCGGCCCTGGCCGAGC 910  
 K E L I L Q L P A W L P G E H G N R G P V A E  
 TGGCCGCGCATCACGTTTGAAGCCCAAGGGCCAGAAAGCTGGCCTGGACCCCGGACCCCGGTGGAAGTGAACGC 980  
 L A G I T F E A K G Q K L A W T R D P V E V N A

图 9A

GTTCCACATCCCCCTGCCCGCCGGCACCCAGCGAAGTGGTGGCCCGCTTCATCCACACACCTCGCCCGCTGCCG 1050  
 FHILPLPAGTSEVVARFIRFTSPLR  
 GACAGCGAAGCGCATCACCGTTACGGCGGAAATGCTCAACGTGCAGTGGAGAAAGATGAGCCCTCTATC 1120  
 DSEGRITVTREMLNVQWELMSLY  
 CCGCCGTCACTATGTGGCCAGATCAAGTGGTCCACCGTCCAGCTTCCGCGAGGGCTGGACCGTGT 1190  
 PAGHYVRQIKVRRPTVSTFPQGWTVF  
 CACCGCGTGGATGGCAAGACGCAGAGCGCGGGCAATACCGTGACCTGGGCCGAAACCGACTATGAA 1260  
 TALDGGKTQSGAGNTVTWAETDYE  
 ACCCTGGTCCGATCTTTGCCGGCTATGCCGGCGGATGATCTGGGCCACAAACGTCATTT 1330  
 TLVDSPIFAGLYAARHDLGHNVY  
 TCGATCTGGTGGCCGACAAGCCCGAGTGTGGCGATCAAGCCGGAAACCTGGCCGCTATCGCAACCT 1400  
 FDLVADKPELLAIAIKPENLAAYRNL  
 GGCCGACGAAGCCGTGGCGCATTCGGCGCGCCATTTGATCACTACGATTCCTGCTCGCGCTGACC 1470  
 ADAVGAFAFRHFDHYDFLLALT  
 GATCGCATGGCAGCATCGGCCTGGAACACCCCGTCCAGCGAAACCAGCAGGAACCCAAAGACCTGA 1540  
 DRMGSLIGLEHHRSSSENQOEPSL  
 CCGACTGGCCCTATGACTGGACCGCAACGTGATCGCCACGAAATTCAGCCACAGCTGGGATGGCAA 1610  
 TDWAAAYDWDNRNVIAHEFSEHSWDGK  
 GTATCGCCGCTCGGCCAAGCTGTGGACGCCCGACTATCGCCAGCCGATGCAGGACAACCTGCTGTGGGTC 1680  
 YRRSAKLLWTPDYRQPMQDNLWV  
 TATGAAGGCGAGCAGTTCGGGCTGGTCTCGCCGCGCAGCTCGGGCGTGCAGAGCAAGGACGTGG 1750  
 YEGQTQFWGLVLAARSGVQSKDV  
 TCTTGGGCGAGCCTCGCCAACTATGCCGGCACGTTACCCAGACCCCGGGCGGACTGGCGCTCGGTGGA 1820  
 VLGSLAN YAGTF TQTAGRDWRSE  
 AGACACGAGGATGCCATCTTCGCCCGCCGCAAGCCCAAGCCCTATTCCTCGCTTACCCTAACCGAG 1890  
 DTMDFIAARAKPKPYSSLTRNE  
 GACTATTACCCGAAAGCGCGTGGTGTGGTGGAAAGCGGACCCAGATCATCCGCGATGGCACCCGGCGCA 1960  
 DYYTEGALLVWLEADQIIIRDGTGG

图 9B





ATCACCAACCGTTCTGACCGCAGTGGTGTGGCATCGATCGCCGGCGAGGCTGCCCCGGCGGCCACATGGCG 2940  
ATTCGCTGGCAAAGCTCCGGCTGATCGGGACGCTGGTGTGCTGACCTGGCTTTTGGCCACA 3000

图 9D

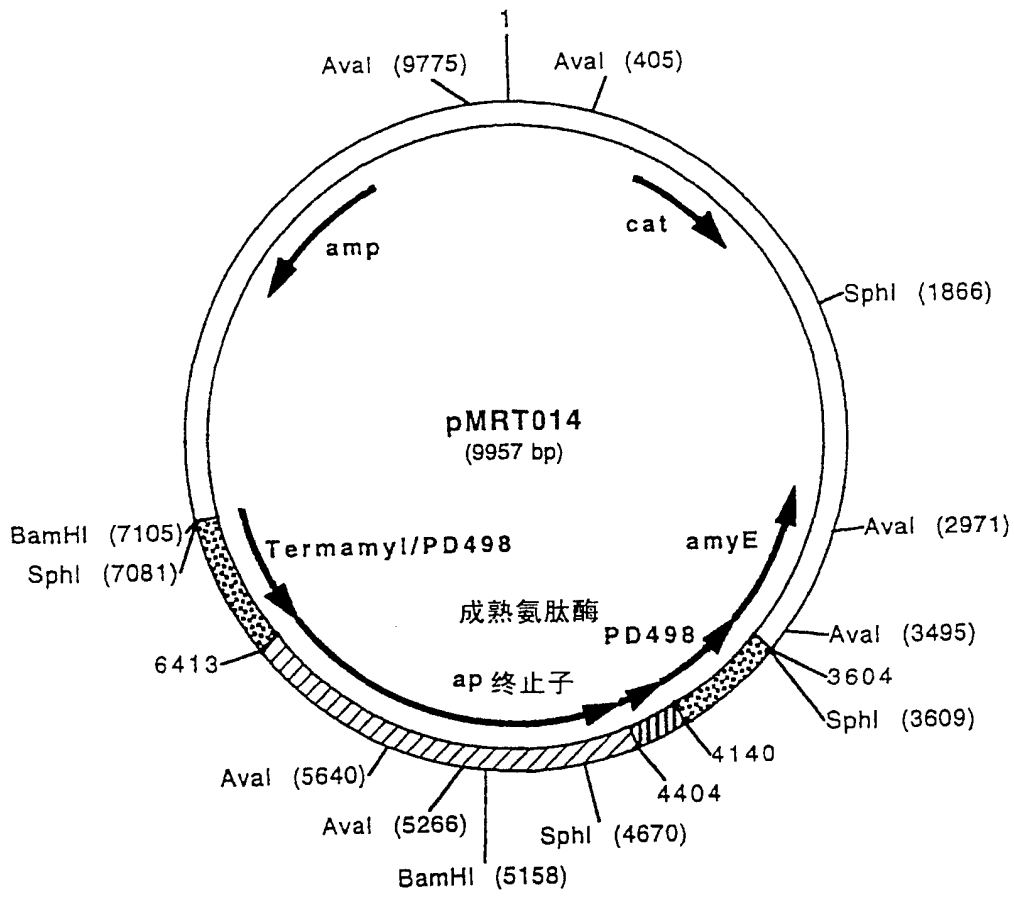


图 10