

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12N 15/57



[12] 发明专利说明书

C12N 9/48 C12N 1/15

C12N 1/19 C12P 21/06

A23J 3/30 A21D 2/26

[21] ZL 专利号 98805174.5

[45] 授权公告日 2004 年 8 月 25 日

[11] 授权公告号 CN 1163607C

[22] 申请日 1998.5.15 [21] 申请号 98805174.5

[30] 优先权

[32] 1997.5.16 [33] US [31] 08/857,886

[32] 1997.10.20 [33] US [31] 60/062,893

[86] 国际申请 PCT/US1998/009940 1998.5.15

[87] 国际公布 WO1998/051804 英 1998.11.19

[85] 进入国家阶段日期 1999.11.16

[71] 专利权人 诺沃奇梅兹生物技术有限公司

地址 美国加利福尼亚

共同专利权人 诺沃奇梅兹有限公司

[72] 发明人 A·布林克沃斯凯 K·布朗

E·格莱特里 T·博恩

L·V·科弗德

审查员 张秀丽

[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 3 页 说明书 69 页 附图 6 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

本发明涉及得自米曲霉、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制备和使用所述多肽的方法。

ISSN 1008-4274

1. 一种具有氨肽酶活性的分离多肽，其包含有 SEQ ID NO: 2 中第 16 - 496 位氨基酸所示序列，或者在 SEQ ID NO: 2 第 16 - 496 位氨基酸所示序列中有一或多个保守性取代、插入和/或缺失的序列。

2. 权利要求 1 的多肽，它含有 SEQ ID NO: 2 中氨基酸 16 到 496 的氨基酸序列。

3. 权利要求 1 的多肽，它由 SEQ ID NO: 2 中氨基酸 16 到 496 的氨基酸序列所组成。

4. 权利要求 1 的多肽，其包含在 SEQ ID NO: 2 第 16 - 496 位氨基酸所示序列中有一或多个保守性取代、插入和/或缺失的序列。

5. 权利要求 1 的多肽，其得自曲霉属 (*Aspergillus*) 菌株。

6. 权利要求 5 的多肽，其得自一株米曲霉 (*Aspergillus oryzae*)。

7. 权利要求 1 的多肽，它是由包含在大肠杆菌 NRRL B-21677 中的质粒 pEJG18 中的核酸序列编码的。

8. 一种具有氨肽酶活性的多肽，它具有以下物理化学特性：(a) 在有 N-丙氨酰对硝基苯胺条件下，于环境温度确定的最佳 pH 范围为 pH7.27 到 pH10.95；(ii) 没有底物情况下于 60℃ 保温 20 分钟后，测得 pH7.5 时相对于起始活性，热稳定性为至少 90%；以及 (iii) 对 Xaa-对硝基苯胺有活性，其中 Xaa 选自由 Leu、Glu、Gly、Ala 和 Pro 组成之组；并且所述多肽包含有 SEQ ID NO: 2 中第 16 - 496 位氨基酸所示序列，或者在 SEQ ID NO: 2 第 16 - 496 位氨基酸所示序列中有一或多个保守性取代、插入和/或缺失的序列。

9. 权利要求 8 的多肽，其得自曲霉属菌株。

10. 权利要求 9 的多肽，其得自一株米曲霉。

11. 制备权利要求 1 所述多肽的方法，包括 (a) 培养包含编码权利要求 1 所述多肽的核酸序列的菌株，以生产多肽；和 (b) 回收该多肽。

12. 一种包含编码权利要求 1 所述多肽的核酸序列的分离核酸序列。

13. 一种核酸构建体，它包含权利要求 12 所述核酸序列，该序列可操

纵性连接了指导在合适的表达宿主中产生多肽的调控序列。

14. 一种重组表达载体，它包含权利要求 13 所述核酸构建体、启动子和转录及翻译终止信号。

15. 一种包含权利要求 13 所述核酸构建体的重组宿主细胞。

16. 多肽的生产方法，包括 (a) 在适合该多肽产生的条件下培养权利要求 15 所述宿主细胞；和 (b) 回收该多肽。

17. 突变细胞的制备方法，包括破坏或缺失编码权利要求 1 所述多肽的核酸序列，从而得到与原细胞相比所述多肽的产量有所减少的突变细胞。

18. 由权利要求 17 的方法得到的突变细胞。

19. 制备异源多肽的方法，包括 (a) 在适合该多肽产生的条件下培养权利要求 18 所述突变细胞；和 (b) 回收该多肽。

20. 由蛋白类底物制备水解产物的方法，包括使底物受到权利要求 1 所述多肽和内肽酶的作用。

21. 权利要求 20 的方法，其中的水解产物富含亮氨酸、甘氨酸、谷氨酸、丝氨酸、天冬氨酸、天冬酰胺、脯氨酸、半胱氨酸、丙氨酸和谷氨酰胺。

22. 权利要求 20 的方法，其中的水解产物富含甘氨酸。

23. 由权利要求 20 所述方法制备的蛋白水解产物。

24. 权利要求 23 的蛋白水解产物，其中的水解产物富含亮氨酸、甘氨酸、谷氨酸、丝氨酸、天冬氨酸、天冬酰胺、脯氨酸、半胱氨酸、丙氨酸和谷氨酰胺。

25. 权利要求 23 的蛋白水解产物，其中的水解产物富含甘氨酸。

26. 包含权利要求 23 所述蛋白水解产物的食品。

27. 由蛋白类底物获得富含游离谷氨酸和/或肽键合谷氨酸残基的蛋白质水解产物的方法，包括使底物进行脱酰胺作用和受到权利要求 1 所述多肽的作用。

28. 权利要求 27 的方法，还包括使底物受到一或多种非特异作用的内肽酶和/或外肽酶酶类的作用。

-
29. 由权利要求 28 的方法得到的蛋白质水解产物。
 30. 包含权利要求 29 所述蛋白质水解产物的食品。
 31. 包含权利要求 1 所述多肽和合适载体的改善风味组合物。
 32. 包含权利要求 1 所述多肽和烘烤成分、用于生面团的预混合物。

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方法。本发明进一步涉及获得可作为提味剂的蛋白质水解产物的方法。

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原料而获得的提味剂。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切和外切肽酶)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EC3.4.11.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

氨酸酶。JP-7-4021798 (Zaidan Hojin Noda Sangyo) 公开了通过添加亮氨酸氨酸酶 II 来制备酱油的方法, 其中亮氨酸氨酸酶 II 是通过培养多个菌株, 包括米曲霉 460 (ATCC20386) 和 IAM2616 制成的。已知米曲霉 460 能产生多种亮氨酸氨酸酶, 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酸酶, 通过 SDS-PAGE 测得分子量为 65kDa (Kwon 等, 1996, 工业微生物学杂志 (Journal of Industrial Microbiology) 17:30-35)。

WO97/04108 (Roehm) 公开了编码酱油曲霉 (*A. sojae*) 亮氨酸氨酸酶的 DNA。Chang 和 Smith (1989, 生物学化学杂志 (Journal of Biological Chemistry) 264:6979-6983) 公开了得自酿酒酵母 (*Saccharomyces cerevisiae*) 的液泡氨酸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酸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酸酶活性的改善的多肽, 和获得有良好感官品质和高水解度的蛋白质水解产物的方法。

发明概述

本发明涉及具有氨酸酶活性的分离的多肽, 它们选自:

(a) 有与 SEQ ID NO: 2 氨基酸序列至少 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, 该核酸序列在中度严谨条

件下与 (i) SEQ ID NO: 1 的核酸序列, (ii) 其互补链, 或 (iii) 其亚序列可杂交;

(c) (a) 或 (b) 的等位基因变体; 和

(d) (a)、(b) 或 (c) 的一个片段, 该片段具有氨肽酶活性; 和

(e) 有氨肽酶活性和如下物理化学特性的多肽: (i) 在有 N-丙氨酰对硝基苯胺的情况下, 于室温确定出最佳 pH 范围在大约 pH 7.27 到 pH 10.95; (ii) 在没有底物情况下, 60℃ 保温 20 分钟后于 pH 7.5 确定出对温度的稳定性为初始活性的 90% 或以上; 和 (iii) 对 Xaa-pNA (Xaa- β -nitroanilide) 有活性, 其中 Xaa 选自亮氨酸、谷氨酸、甘氨酸、丙氨酸和脯氨酸。

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所述多肽的方法。

本发明还涉及由蛋白类底物获得水解产物的方法, 该方法包括将蛋白类材料与单独的、或结合内肽酶的有氨肽酶活性的多肽进行接触, 还涉及由该方法获得的水解产物。

本发明还涉及由蛋白类底物获得富含游离谷氨酸和/或肽键合谷氨酸残基的水解产物的方法, 该方法包括将底物进行脱酰胺反应, 和接受具有氨肽酶活性的多肽的作用。

本发明进一步涉及包含有氨肽酶活性的多肽的提味组合物。该组合物还可以包含另外的酶活性。

最后一个方面, 本发明的方法可以用于食品相关的应用 (如烘烤) 中, 以改善风味。可选择地, 通过添加由本发明所述方法获得的水解产物可以改善食品的风味。

附图简述

图 1 显示米曲霉 ATCC20386 氨肽酶的核酸序列和推导出的氨基酸序列 (分别是 SEQ ID NO: 1 和 2)。

图 2 显示 pMWR3 的限制图谱。

图 3 显示 pEJG19 的限制图谱。

图 4 显示 pDM181 的限制图谱。

图 5 显示 pEJG28 的限制图谱。

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，但至少是亮氨酸、谷氨酸、甘氨酸、丙氨酸和/或脯氨酸。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。

在第一个实施方案中，本发明涉及分离的多肽，该多肽具有氨肽酶活性，其氨基酸序列与 SEQ ID NO: 2 氨基酸序列的相同程度至少是大约 50%、优选至少是大约 60%、优选至少是大约 70%，更优选至少是大约 80%，还要优选的是至少大约 90%，最优选至少是大约 95%，最最优选的是至少大约 97%（以下称为“同源多肽”）。在一个优选实施方案中，同源多肽具有的氨基酸序列与 SEQ ID NO: 2 氨基酸序列有 5 个不同的氨基酸，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。用于本发明的目的，通过 Clustal 方法（Higgins, 1989, CABIOS5: 151-153）用相同性表、缺口罚分和缺口长度罚分均为 10 来确定两个氨基酸序列之间的相同程度。

在又一个实施方案中，本发明涉及一种具有氨肽酶活性的分离多肽，其包含有 SEQ ID NO: 2 中第 16-496 位氨基酸所示序列，或者在 SEQ ID NO: 2 第 16-496 位氨基酸所示序列中有一或多个保守性取代、插入和/或缺失的序列。

优选地，本发明所述多肽包含 SEQ ID NO: 2 的氨基酸序列或等位基因变体；和它们的一个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，本发明所述多肽具有 SEQ ID NO: 2 的氨基酸序列或它的一个有氨肽酶活性的片段。SEQ ID NO: 2 的片段是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。在一个最优选的实施方案中，多肽具有 SEQ ID NO: 2 的氨基酸序列。

优选地，多肽片段含有至少 330 个氨基酸残基，更优选至少 380 个氨基酸残基，最优选有至少 430 个氨基酸残基。

等位基因变体表示一个基因占据相同染色体位点的两个或多个可替换形式中任意一种。通过突变自然产生等位基因变体，并可能导致种群内的表现型多态性。基因突变可以是沉默的（编码的多肽没有变化）或者可以编码氨基酸序列改变的多肽。术语“等位基因变体”也用来表示等位基因变体所编码的蛋白质。

同源多肽的氨基酸序列与 SEQ ID NO: 2 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/或用不同的氨基酸残基取代 1 或多个氨基酸残基。优选地，氨基酸改变是小的变化，即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；或者这样一个小延伸，它通过改变净电荷有助于纯化或者有其它功能（如多聚组氨酸尾巴、抗原表位或结合结构域）。

保守性取代的例子是在碱性氨基酸（如精氨酸、赖氨酸和组氨酸）、酸性氨基酸（如谷氨酸和天冬氨酸）、极性氨基酸（如谷氨酰胺和天冬酰胺）、疏水氨基酸（如亮氨酸、异亮氨酸和缬氨酸）、芳香族氨基酸（如苯丙氨酸、色氨酸和酪氨酸）和小氨基酸（如丙氨酸、甘氨酸、丝氨酸、苏氨酸和甲硫氨酸）内进行的取代。通常不会改变特异性作用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) 等（《蛋白质》，Academic Press, New York）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，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，最优选在高度严谨条件下，与寡核苷酸探针（这些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杂交）能够杂交（J. Sambrook, E. F. Fritsch, 和 T. Maniatus, 分子克隆：实验指南，第 2 版，Cold Spring Harbor, New York）；或所述多肽的等位基因变体和具有氨肽酶活性的片段。

杂交是指按照标准 Southern 印迹操作进行实验，核酸序列与对应于 SEQ ID NO: 1 所示核酸序列中多肽编码部分的寡核苷酸探针在低度到高度严谨条件下（即，在 $5 \times$ SSPE、0.3% SDS、 $200 \mu\text{g/ml}$ 剪切变性的鲑精 DNA、甲酰胺（对低度、中度和高度严谨条件分别为 25、35 或 50%）中于 42°C 预杂交和杂交）能发生杂交。

可以根据本领域熟知的方法，用 SEQ ID NO: 2 的氨基酸序列或其部分序列来设计寡核苷酸引物，或者可以用编码本发明所述多肽的核酸序列（如 SEQ ID NO: 1 的核酸序列，或其亚序列）从不同属或种的株系中鉴定和克隆到编码具有氨肽酶活性的多肽的 DNA。具体来说，可以按照标准 Southern 印迹操作，用这些探针与所研究的属或种的基因组或 cDNA 杂交，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 40 个核苷酸长。也可以使用更长的探针。DNA 和 RNA 探针都可使用。通常为了检测相应基因要将探针进行标记（例如，用 ^{32}P 、 ^3H 、 ^{35}S 、生物素或抗生物素蛋白）。

因此，可以从得自其他生物的基因组、cDNA 或组合化学文库筛选能与上述探针杂交、并编码具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物质上。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，并

且最终用 $2 \times \text{SSC}$ 、0.2%SDS 将该载体材料洗 3 次，每次 30 分钟，洗涤温度优选至少 50°C ，更优选至少 55°C ，更优选至少 60°C ，更优选至少 65°C ，还要优选的是至少 70°C ，最优选的是至少 75°C 。用 X-光胶片检测出在这些条件下与寡核苷酸探针杂交的分子。

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：(i) 在有 N-丙氨酰对硝基苯胺的情况下，于室温确定的最佳 pH 在大约 pH7.27 到大约 pH10.95；(ii) 没有底物的情况下， 60°C 保温 20 分钟后于 pH7.5 确定出对温度的稳定性为起始活性的 90%或以上；和 (iii) 对 N-Xaa 酰基-对硝基苯胺有活性，其中的 Xaa 选自亮氨酸、谷氨酸、甘氨酸、丙氨酸和脯氨酸。本发明所述多肽还有水解其他底物的能力。

在一个优选实施方案中，在有 N-丙氨酰-脯氨酰 对硝基苯胺的情况下，于室温温育 5 分钟后，确定出最佳 pH 在大约 pH7.27 到大约 pH10.95，更优选在大约 pH8.03 到 pH10.95，最优选在大约 pH8.62 到 pH 10.51。

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) 制备抗血清，该抗血清所含抗体与具有 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这样的多肽，在使用特定免疫化学技术时，它们与抗血清反应的方式相同 (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 (N. H. Axelsen, J. Kroll, 和 B. Weeks 编“定量免疫电泳手册”，Blackwell Scientific Publications,

1973, 10章)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(N. H. Axelsen, J. Kroll, 和 B. Weeks 编“定量免疫电泳手册”, Blackwell Scientific Publications, 1973, 11章)对免疫化学特性部分相同作了更详细解释。

可以从任何属的微生物中获得能和这样的寡核苷酸探针杂交的核酸序列所编码的多肽(该寡核苷酸探针能与 SEQ ID NO: 1 核酸序列、其互补链或者等位基因变体和亚序列杂交); 该多肽的等位基因变体和片段; 或者同源多肽和免疫化学特性相同或部分相同的多肽。

在一个优选实施方案中, 这些多肽可以来源于细菌。例如, 可以从革兰氏阳性细菌, 如芽孢杆菌或链霉菌中得到这些多肽。芽孢杆菌的例子如嗜碱芽孢杆菌、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、苏云金芽孢杆菌; 链霉菌的例子如浅青紫链霉菌或鼠灰链霉菌; 或者从革兰氏阴性细菌, 例如大肠杆菌或假单胞菌中得到。

所述多肽可以来源于真菌, 更优选得自酵母菌, 如假丝酵母属、克鲁维酵母属、毕赤酵母属、酵母属、许旺酵母属 (*Schizosaccharomyces*) 或 *Yarrowia*; 或者来源于丝状真菌, 如支顶孢属 (*Acremonium*)、曲霉属、短梗霉属 (*Aureobasidium*)、隐球酵母属、*Filibasidium*、镰孢属、腐质霉属、*Magnaporthe*、毛霉属、毁丝霉属、*Neocallimastix*、链孢霉属、拟青霉属 (*Paecilomyces*)、青霉属、*Piromyces*、裂褶菌属 (*Schizophyllum*)、踝节菌属 (*Talaromyces*)、热子囊菌属 (*Thermoascus*)、草根霉属 (*Thielavia*)、*Tolyocladium* 或木霉属的菌株。

在一个优选实施方案中, 从卡尔酵母、酿酒酵母、糖化酵母、*S. douglasii*、克鲁弗酵母、诺地酵母、卵形糖酵母 (*S. oviformis*) 菌株中获得所述多肽。

在另一个优选实施方案中，所述多肽来源于杆孢状镰孢、*Fusarium cerealis*、*Fusarium crookwellense*、大刀镰孢、禾谷镰孢、禾赤镰孢、异孢镰孢、合欢镰孢(*Fusarium negundi*)、尖镰孢、多枝镰孢、粉红镰孢、接骨木镰孢、肤色镰孢、拟分枝孢镰孢、硫色镰孢(*Fusarium sulphureum*)、*Fusarium torulosum*、*Fusarium trichothecioides*、*Fusarium venenatum*、*Humicola insolens*、*Humicola lanuginosa*、米赫毛霉(*Mucor miehei*)、嗜热毁丝霉(*Myceliophthora thermophila*)、粗糙脉孢菌、产紫青霉、*Trichoderma harzianum*、康宁木霉、*Trichoderma longibrachiatum*、*Trichoderma reesei*或绿色木霉菌株。

优选本发明所述多肽得自曲霉属的某些菌种，这些种包括但不限于，棘孢曲霉、泡盛曲霉、臭曲霉、日本曲霉、构巢曲霉、黑曲霉、米曲霉。

在一个更优选的实施方案中，本发明多肽，例如具有 SEQ ID NO: 2 氨基酸序列的多肽，得自米曲霉菌株、最优选得自米曲霉 ATCC20386 或它的突变株。

应理解，就上述菌株而言，本发明包括其良好和不良状态，以及其他分类学上的等同体(例如变形体)，而无论其已知种名为何。本领域技术人员容易认识到适当的等同体是相同的。本发明所述多肽还可以得自与“曲霉”同义的微生物，如 Raper, K. D. 和 Fennel, D. I. 定义的(1965,《曲霉属》，The Wilkins Company, Baltimore)。Aspergilli 是有丝分裂孢子(mitosporic)真菌，其特征是 aspergillum 包含没发现其有性形态的分生孢子柄，它的末端形成一个小泡，后者又有 1 或 2 层同时形成的分化细胞，分别称为担孢子梗或瓶梗，无性过程形成的孢子称为分生孢子。曲霉属的已知有性型包括散囊菌属、*Neosartorya* 和赤孢壳属(*Emericella*)。公众可很容易从许多培养物保藏中心(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)和农业研究机构保藏中心(NRRL))获得曲霉属菌株及其有性型。

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定和获得这些多肽。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）对序列进行分离和克隆。

就本发明的目的而言，文中与特定来源连在一起使用的术语“得自”意味着所述多肽由该来源产生或由插入了来自该来源的基因的细胞产生。

如文中定义，“分离的”多肽是基本上没有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。

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。在一个优选实施方案中，编码多肽的核酸序列得自曲霉，例如米曲霉，在一个更优选的实施方案中，该核酸序列，例如 SEQ ID NO: 1 的核酸序列得自米曲霉 ATCC20386。在另一个更优选的实施方案中，所述核酸序列是质粒 pEJG18（该质粒包含在大肠杆菌 NRRL B-21677）所含序列。本发明还包括编码具有 SEQ ID NO: 2 氨基酸序列的多肽的核酸序列，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，这些亚序列编码 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 990 个核苷酸，更优选至少 1140 个核苷酸，最优选至少 1290 个核苷酸。

所述核酸序列可以得自在分类学上等同于曲霉的微生物

(Raper, K. D. 和 Fennel, D. I. 定义过的, 出处同前), 不论这些微生物名称如何。

用于分离和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)。可以从曲霉属的一个或另一个菌株或者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基因或种变体。

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, 该序列在此被复制。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 核酸序列有一定同源性的核酸序列, 同源程度至少约为 50%、优选约 60%、优选约 70%、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 Clustal 方法 (Higgins, 1989, 出处同前) 用相同性表, 缺口罚分和缺口长度罚分均为 10 确定的。

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设计方式上不同。例如，可能希望用例如定点诱变来合成多肽的变体，这些变体有不同的特异性作用、热稳定性或最佳 pH 等。可以在 SEQ ID NO: 1 中多肽编码部分所示核酸序列的基础上构建出类似序列，例如其亚序列；并/或通过引入这样的核苷酸取代而构建，这些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，但它们符合制备酶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，蛋白质表达和纯化 (Protein Expression and Purification) 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。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 (Science) 244: 1081-1085)，鉴定出那些为本发明所述分离的核酸序列编码的多肽活性所必需、因此最好不进行取代的氨基酸残基。后一种技术中，向分子中每个带正电荷的残基导入突变，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。通过分析例如用核磁共振、晶体学或光亲和标记确定的立体结构，也可以确定底物-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 (Journal of Molecular Biology) 224: 899-904; Wlodaver 等, 1992, FEBS 快报 (FEBS Letters) 309: 59-64)。

本发明所述多肽还包括融合多肽或可裂解的融合多肽，这些融合多肽是在所述多肽或其片段的 N-或 C-末端融合上了另一个多肽。融合多肽是通过将编码另一个多肽的核酸序列 (或者其一部分) 与本发明的核酸序列 (或者其一部分) 融合在一起制备的。用于制备融合多肽的技术是本领域已知的，包括读框一致地连接多肽编码序列，并且使融合多肽的表达受到相同的启动子和终止子的调控。

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，最优选在高度严谨条件下能与这样的寡核苷酸探针杂交，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，或其等位基因变体和亚序列杂交 (Sambrook 等, 1989, 出处同前)。

核酸构建体

本发明还涉及包含本发明所述核酸序列的核酸构建体，在这些核酸构建体中，核酸序列可操纵地连接了 1 或多个调控序列，该调控序列在与其相容的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，或者已被修饰而含有以非天然方式组合和排列的核酸片段的基因。当核酸构建体包含所有表达本发明所述编码序列必需的调控序列时，术语核酸构建体与表达盒同义。术语“编码序列”在文中定义为能转录为 mRNA，并翻译成本发明所述多肽的序列。编码序列的范围通常是由位于 mRNA 5'端开放读码框上游的核糖体结合位点（对于原核细胞）或 ATG 起始密码子（对于真核细胞）和位于 mRNA 3'端开放读码框下游的转录终止序列确定的。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。在将编码多肽的核酸序列插入载体中之前对其进行加工可能比较理想，或者是必需的，这取决于具体的表达载体。利用克隆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”定义为包括本发明多肽的表达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终止信号。为了导入特定的限制位点以便于将调控序列与编码多肽的核酸序列的编码区连接起来，可以提供带有接头的调控序列。术语“可操纵地连接”在文中定义为这样一种结构，其中调控序列相对于DNA序列的编码序列适当地位于能指导产生多肽的位置上。

调控序列可以是合适的启动子序列、用来表达核酸序列的宿主细胞能识别的核酸序列。启动子序列含有介导多肽表达的转录调控序列。启动子可以是任何在所选用的宿主细胞中有转录活性的核酸序列，包括突变的、截短的和杂合的启动子，可以得自编码对于宿主细胞是内源或外源、胞外或胞内多肽的基因。

指导转录本发明所述核酸构建体（特别是在细菌宿主细胞中）的合适启动子的例子，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 α -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 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，原核细胞的 β -内酰胺酶基因（Villa-Kamaroff 等，1978，美国科学院学报（Proceedings of the National Academy of Sciences USA）75: 3727-3731），以及 tac 启动子（DeBoer 等，1983，美国科学院学报，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。

用于在丝状真菌宿主细胞中指导转录本发明所述核酸构建体的合适启动子的例子得自编码以下酶的基因：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 α -淀粉酶、黑曲霉或泡盛曲霉葡糖淀粉酶（glaA）、*Rhizomucor miehei* 脂酶、米曲霉碱性蛋白酶、米曲霉磷酸丙糖异构酶、构巢曲霉乙酰胺酶、尖镰孢胰蛋白酶样蛋白酶（美国专利 4288627）和它们的突变、截短和杂合启动子。特别优选用于丝状真菌宿主细胞的启动子

是 TAKA 淀粉酶、NA2-tpi(来源于编码黑曲霉中性 α -淀粉酶的基因和米曲霉磷酸丙糖异构酶的基因的杂合启动子), 以及 glaA 启动子。

可用于酵母宿主的启动子得自酿酒酵母烯醇化酶(ENO-1)基因、半乳糖激酶基因(GAL1)、乙醇脱氢酶/甘油醛-3-磷酸脱氢酶基因(ADH2/GAP)和3-磷酸甘油酸激酶基因。Romanos等(1992, 酵母(Yeast) 8: 423-488)描述过其他可用于酵母宿主细胞的启动子。可用于哺乳动物宿主细胞的启动子包括病毒启动子, 如来自猴病毒40(SV40)、劳氏肉瘤病毒(RSV)、腺病毒、牛乳头瘤病毒(BPV)和人巨细胞病毒(CMV)的病毒启动子。

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。终止序列可操纵地连接在编码多肽的核酸序列的3'末端。任何适用于所用宿主细胞的终止子都可以用于本发明。

用于丝状真菌宿主细胞的优选终止子得自编码以下酶的基因: 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、黑曲霉 α -葡糖苷酶和尖镰孢胰蛋白酶样蛋白酶。

用于酵母宿主细胞的优选终止子得自编码以下酶的基因: 酿酒酵母烯醇化酶、酿酒酵母细胞色素C(CYC1)、或酿酒酵母甘油醛-3-磷酸脱氢酶。Romanos等(1992, 出处同前)描述过其他可用于酵母宿主细胞的终止子。用于哺乳动物宿主细胞的终止序列为本领域所熟知。

调控序列还可以是合适的前导序列(mRNA的非翻译区, 对宿主细胞翻译非常重要)。前导序列可操纵地连接在编码多肽的核酸序列的5'末端。任何适用于所用宿主细胞的前导序列都可以用于本发明。

用于丝状真菌宿主细胞的优选前导序列得自编码以下酶的基因: 米曲霉 TAKA 淀粉酶和构巢曲霉磷酸丙糖异构酶。

适用于酵母宿主细胞的前导序列得自以下基因: 酿酒酵母烯醇化酶(ENO-1)基因、酿酒酵母3-磷酸甘油酸激酶基因、酿酒酵母 α -因子和酿酒酵母乙醇脱氢酶/甘油醛-3-磷酸脱氢酶基因(ADH2/GAP)。

调控序列还可以是多聚腺苷酸化序列, 该序列可操纵地连接在核酸序列的3'末端, 当翻译时, 它能被宿主细胞识别为向转录后的 mRNA

添加多聚腺苷酸残基的信号。任何适用于所用宿主细胞的多聚腺苷酸化序列都可以用于本发明。

用于丝状真菌宿主细胞的优选多聚腺苷酸化序列得自编码以下酶的基因：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和黑曲霉 α -葡糖苷酶。

Guo 和 Sherman (1995, 分子细胞生物学 15: 5983-5990) 描述过可用于酵母宿主细胞的多聚腺苷酸化序列。用于哺乳动物宿主细胞的多聚腺苷酸化序列为本领域所熟知。

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、能引导编码多肽进入细胞分泌途径的氨基酸序列。核酸序列编码区的 5'端可能天然含有信号肽编码区，该编码区与分泌多肽的编码区片段翻译读框一致地连接在一起。可选择地，核酸序列编码区的 5'端可含有对编码序列是外来的信号肽编码区。在正常情况下不含有信号肽编码区的编码序列，可能需要添加外来信号肽编码区。可选择地，为了增强多肽分泌，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。信号肽编码区可以得自曲霉属内某些种的葡糖淀粉酶或淀粉酶基因、根霉属内某些种的脂酶或蛋白酶基因、酿酒酵母 α -因子的基因、芽孢杆菌属内某些种的淀粉酶或蛋白酶基因，或者小牛前凝乳酶原基因。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。

用于细菌宿主细胞的有效信号肽编码区是得自芽孢杆菌 NCIB11837 的麦芽糖源淀粉酶基因、嗜热脂肪芽孢杆菌的 α -淀粉酶基因、地衣芽孢杆菌的枯草蛋白酶基因、地衣芽孢杆菌的 β -内酰胺酶基因、嗜热脂肪芽孢杆菌的中性蛋白酶基因 (nprT、nprS、nprM) 或枯草芽孢杆菌 PrsA 基因的信号肽编码区。Simonen 和 Palva (1993, 微生物学综述 (Microbiological Reviews) 57: 109-137) 描述过其他的信号肽。

用于丝状真菌宿主细胞的有效信号肽编码区是得自米曲霉 TAKA 淀粉酶基因、黑曲霉中性淀粉酶基因、*Rhizomucor miehei* 天冬氨酸蛋

白酶基因、*Humicola lanuginosa* 纤维素酶或 *Humicola lanuginosa* 脂酶基因的信号肽编码区。

可用于酵母宿主细胞的信号肽得自酿酒酵母 α -因子和酿酒酵母转化酶的编码基因。Romanos 等(1992, 出处同前)描述过其他有用的信号肽编码区。

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。这样产生的多肽已知是酶原或多肽原(或者在某些情况中是酶原)形式。多肽原通常没有活性, 可以由多肽原催化性地或自身催化性地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*)、枯草芽孢杆菌中性蛋白酶基因(*nprT*)、酿酒酵母 α -因子基因、*Rhizomucor miehei* 天冬氨酸蛋白酶基因或嗜热毁丝霉漆酶基因(W095/33836)。

在多肽的氨基末端有信号肽和肽原区的情况中, 肽原区与多肽的氨基末端相邻, 而信号肽区与肽原区的氨基末端相邻。

本发明所述核酸构建体还可以包含 1 或多个核酸序列, 这些序列编码 1 或多个有助于指导多肽表达的因子, 例如转录激活因子(例如, 反式作用因子)、伴侣蛋白和加工蛋白酶。任何可用于所选宿主细胞的因子都可以用于本发明。编码 1 或多个这类因子的核酸不必与编码多肽的核酸序列串联在一起。

转录激活因子是一种能激活编码多肽的核酸序列进行转录的蛋白质(Kudla 等, 1990, *EMBO Journal* 9:1355-1364; Jarai 和 Buxton, 1994, *现代遗传学(Current Genetics)* 26:2238-2244; Verdier, 1990, *酵母* 6:271-291)。编码转录激活因子的核酸序列可以得自编码以下蛋白质的基因: 嗜热脂肪芽孢杆菌 *NprA*(*nprA*)、酿酒酵母血红蛋白激活蛋白 1(*hap1*)、酿酒酵母半乳糖代谢蛋白 4(*gal4*)、构巢曲霉氨调节蛋白(*areA*)和米曲霉 α -淀粉酶激活因子(*amyR*)。另外的例子可参见 Verdier, 1990, 出处同前和 MacKenzie 等, 1993, *普通微生物学杂志(Journal of General Microbiology)* 139: 2295-2307。

伴侣蛋白是能协助其他多肽进行正确折叠的蛋白质(Hartl 等,

1994, TIBS19: 20-25; Bergeron 等, 1994, TIBS19: 124-128; Demolder 等, 1994, 生物技术杂志 (Journal of Biotechnology) 32: 179-189; Craig, 1993, 科学 260: 1902-1903; Gething 和 Sambrook, 1992, 自然 (Nature) 355: 33-45; Puig 和 Gilbert, 1994, 生物学化学杂志 269: 7764-7771; Wang 和 Tsou, 1993, FASEB 杂志 7: 1515-11157; Robinson 等, 1994, 生物/技术 (Bio/Technology) 1: 381-384; Jacobs 等, 1993, 分子微生物学 (Molecular Microbiology) 8: 957-966)。编码伴侣蛋白的核酸序列可以得自编码以下蛋白质的基因: 枯草芽孢杆菌 GroE 蛋白、枯草芽孢杆菌 PrsA、米曲霉蛋白质二硫键异构酶、酿酒酵母钙联接蛋白、酿酒酵母 BiP/GRP78 和酿酒酵母 Hsp70。另外的例子可参见 Gething 和 Sambrook, 1992, 出处同前和 Hartl 等, 1994, 出处同前。

加工蛋白酶是能切割肽原从而产生成熟的有生物化学活性的多肽的蛋白酶 (Enderlin 和 Ogrydziak, 1994, 酵母 10: 67-79; Fuller 等, 1989, 美国科学院学报 86: 1434-1438; Julius 等, 1984, 细胞 (Cell) 37: 1075-1089; Julius 等, 1983, 细胞 32: 839-852; 美国专利 5702934)。编码加工蛋白酶的核酸序列可以得自编码下述蛋白质的基因: 酿酒酵母二肽基氨肽酶、酿酒酵母 Kex2、*Yarrowia lipolytica* 双碱性加工内蛋白酶 (xpr6) 和尖镰孢金属蛋白酶 (p45 基因)。

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的表达的系统。原核细胞体系中的调控系统包括 lac、tac 和 trp 操纵子系统。在酵母中, 可以使用 ADH2 系统或 GAL1 系统。在丝状真菌中, 可以用 TAKA α -淀粉酶启动子、黑曲霉葡糖淀粉酶启动子和米曲霉葡糖淀粉酶启动子作为调控序列。调控序列的其他例子是那些能使基因发生扩增的调控序列。在真核细胞体系中, 这类调控序列包括二氢叶酸还原酶基因 (该基因在有氨甲蝶呤的情况下发生扩增) 和金属硫蛋白基因 (重金属可使其扩增)。在这些情况中, 应将编码多肽的核酸序列与调控序

列可操纵地连接在一起。

本发明还涉及使编码本发明所述多肽的内源基因的表达情况发生改变的核酸构建体。构建体可以含有使内源基因表达改变所必需的最少量成分。在一个实施方案中，核酸构建体优选含有(a)靶向序列、(b)调控序列、(c)外显子和(d)剪接供体位点。一旦将核酸构建体导入了细胞，该构建体将通过同源重组插入到细胞基因组中的内源基因位点。靶向序列引导元件(a)-(d)整合到内源基因中，从而使元件(b)-(d)与内源基因可操纵地连接在一起。在另一个实施方案中，核酸构建体包含(a)靶向序列、(b)调控序列、(c)外显子和(d)剪接供体位点、(e)内含子和(f)剪接受体位点，其中的靶向序列引导元件(a)-(f)整合到内源基因中，从而使元件(b)-(f)与内源基因可操纵地连接在一起。但是，构建体可以含有其他成分，如选择标记。

在这两个实施方案中，这些成分的导入将导致产生一个新的转录单位，在这个转录单位中内源基因的表达情况会有所改变。本质上来说，这个新转录单位是靶向构建体导入的序列与内源基因的融合产物。在一个内源基因发生改变的实施方案中，该基因是被激活。在这个实施方案中，利用同源重组插入一个调控序列，代替、破坏或失活正常情况下与亲本细胞的内源基因相关的调控区，从而导致基因表达水平高于相应的亲本细胞。可以进一步利用本领域已知方法(参见，例如美国专利 5641670)，使构建体含有扩增性选择标记基因，从而使被激活的基因发生扩增。在内源基因发生改变的另一个实施方案中，该基因的表达减少。

靶向序列可以位于内源基因之内、紧邻该基因、在上游基因内部或与内源基因相距较远。可以使用1或多个靶向序列。例如，环状质粒或DNA片段优选使用单个靶向序列，而线性质粒或DNA片段优选使用两个靶向序列。

构建体中的调控序列可以包含1或多个启动子、增强子、或斯加弗德结合区或基质结合位点、负调控元件、转录结合位点或这些序列的组合形式。

构建体还可含有内源基因的 1 或多个外显子。外显子定义为能被拷贝成 RNA 并存在于成熟 mRNA 分子中、使外显子序列合乎内源基因的编码区读框的 DNA 序列。外显子可以任选地含有编码 1 或多个氨基酸和/或部分编码氨基酸的 DNA。可选择地，外显子含有对应于 5'非编码区的 DNA。当这个（些）外源外显子编码 1 或多个氨基酸和/或氨基酸的一部分时，应这样设计核酸构建体，以便经转录和剪接后，其读码框与内源基因的编码区一致，这样来源于第二个外显子的 mRNA 部分中合适的读码框部分不被改变。

构建体的剪接供体位点指导一个外显子与另一个外显子进行剪接。通常，第一个外显子位于第二外显子的 5'方向，在第一个外显子的 3'方向与之部分重叠和侧接的剪接供体位点识别侧接在第二外显子 5'方向的剪接受体位点。剪接受体位点与剪接供体位点一样，都是指导一个外显子与另一个外显子进行剪接的序列。剪接结构用剪接受体位点（与剪接供体位点结合发挥作用）来去除内含子。

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，以便可以在这些位点插入或取代编码多肽的核酸序列。可选择地，通过将核酸序列或包含该序列的核酸构建体插入用于表达的合适载体中，可以表达本发明所述核酸序列。为了进行表达并可能分泌，在制备表达载体时，将编码序列放在载体中的一定位置以便编码序列与适当的调控序列可操纵地连接在一起。

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）。选择载体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（即以染色体外个体存在的载体，其复制独立于染色体的复制），例如质粒、染色体外元件、微小染色体或人工染色体。载体

可以含有任何保证自我复制的手段。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。载体系统可以是单个载体或质粒，或者是两个或多个载体或质粒，它们一同含有要导入到宿主细胞基因组中的全部DNA，或者是转座子。

优选本发明所述载体含有1或多个使得容易挑选出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、相对营养缺陷型的原养型等。细菌选择标记的例子是来自枯草芽孢杆菌或地衣芽孢杆菌的dal基因，或者能赋予抗生素抗性（如氨基青霉素、卡那霉素、氯霉素或四环素抗性）的标记。适用于哺乳动物细胞的标记是二氢叶酸还原酶(dfhr)、潮霉素磷酸转移酶(hygB)、氨基糖苷磷酸转移酶II和腐草霉素抗性基因。适用于酵母宿主细胞的标记是ADE2、HIS3、LEU2、LYS2、MET3、TRP1和URA3。用于丝状真菌宿主细胞的选择标记可选自(包括,但不限于): amdS(乙酰胺酶)、argB(鸟氨酸氨甲酰转移酶)、bar(磷丝菌素乙酰转移酶)、hygB(潮霉素磷酸转移酶)、niaD(硝酸盐还原酶)、pyrG(乳清酸核苷-5'-磷酸脱羧酶)、sC(硫酸腺苷酰转移酶)、trpC(邻氨基苯甲酸合成酶)以及来自其他种的等同物质。优选用于曲霉细胞的是来自构巢曲霉或米曲霉的amdS和pyrG基因以及吸水链霉菌(*S. hygroscopicus*)的bar基因。另外,可以通过共转化进行选择,例如W091/17243中描述过的,其中选择标记位于独立的载体上。

优选本发明所述载体包含这样的元件,它(们)允许载体稳定地整合到宿主细胞基因组中,或者保证载体在细胞中独立于细胞的基因组进行自主复制。

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。可选择地,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，它们与相应的目标序列高度同源，从而增加了同源重组的可能性。整合元件可以是与宿主细胞基因组中的目标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是以下质粒的复制起点：允许在大肠杆菌中复制的 pBR322、pUC19、pACYC177 和 pACYC184 的复制起点，以及允许在芽孢杆菌中复制的 pUB110、pE194、pTA1060、和 pAM β 1 的复制起点。用于酵母宿主细胞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, 美国科学院学报 75: 1433）。

可以向宿主细胞插入 1 个以上拷贝的编码本发明所述多肽的核酸序列来扩大表达该核酸序列。可以通过将至少 1 个附加拷贝的序列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，来使核酸序列稳定扩增，后一情况中，可以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。

用于连接前面描述的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

本发明还涉及包含本发明所述核酸序列的重组宿主细胞，这些细胞适合用来重组地制备多肽。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。

将包含本发明所述核酸序列的载体导入宿主细胞，使载体保持染色体整合的或染色体外自我复制的形态。通常认为整合有优势，因为这

样核酸序列更可能稳定地保持在细胞内。载体可以通过如上描述的同源或非同源重组整合到宿主的染色体中。

选择宿主细胞很大程度上取决于编码多肽的基因及其来源。宿主细胞可以是单细胞微生物（例如原核细胞）或非单细胞微生物（例如真核细胞）。有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。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 (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（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。

宿主细胞可以是真核细胞如哺乳动物细胞、昆虫细胞、植物细胞或真菌细胞。有用的哺乳动物细胞包括中国仓鼠卵巢细胞 (CHO)、Hela 细胞、幼仓鼠肾 (BHK) 细胞、COS 细胞或任何其他可以从例如美国典型培养物保藏中心得到的永生化细胞系。

在一个优选实施方案中，宿主细胞是一种真菌细胞。文中所用的“真菌”包括子囊菌门 (Ascomycota)、担子菌门 (Basidiomycota)、壶菌门 (Chytridiomycota) 和接合菌门 (Zygomycota) (Hawksworth 等, 在 Ainsworth and Bisby's Dictionary of The Fungi, 第 8 版, CAB International, University Press, Cambridge, UK 中定义的)

以及卵菌门 (Oomycota) (Hawksworth 等, 1995, 出处同前, 171 页) 和所有有丝分裂孢子真菌 (Hawksworth 等, 1995, 出处同前)。子囊菌门的代表性组包括例如脉孢菌属、真青霉属 (=青霉属)、赤孢壳属 (*Emericella*) (=曲霉属)、散囊菌属 (=曲霉属) 和以下所列的真酵母。担子菌门的例子包括例如蘑菇、锈菌和黑粉菌。壶菌门的代表性组包括例如异水霉属 (*Allomyces*)、小芽枝霉属 (*Blastocladiella*)、雕蚀菌属 (*Coelomomyces*) 和水生真菌。卵菌门的代表性组包括例如水霉科 (Saprolegniomycetous) 水生真菌 (水霉), 如绵霉属 (*Achlya*)。有丝分裂孢子真菌的例子有曲霉属、青霉属、假丝酵母属和链格孢属。接合菌门的代表性组包括例如根霉属和毛霉属。

在一个更优选的实施方案中, 真菌宿主细胞是一种酵母细胞。文中使用的“酵母”包括子囊孢子酵母 (内孢霉目)、担子孢子酵母和属于半知菌类 (芽孢纲) 的酵母。子囊孢子酵母分为蚀精霉科 (*Spermophthoraceae*) 和酵母科 (*Saccharomycetaceae*) 两个科。后者包括四个亚科: 裂殖酵母亚科 (*Schizosaccharomycoideae*) (例如裂殖酵母属)、拿逊酵母亚科 (*Nadsonioideae*)、油脂酵母亚科 (*Lipomycoideae*) 和酵母亚科 (*Saccharomycoideae*) (例如毕赤酵母属、克鲁维酵母属和酵母属)。担子孢子酵母包括白冬孢酵母属 (*Leucosporidium*)、红冬孢酵母属 (*Rhodospodium*)、锁掷酵母属 (*Sporidiobolus*)、*Filobasidium* 和 *Filobasidiella*。属于半知菌类的酵母分为两个科: 掷孢酵母科 (*Sporobolomycetaceae*) (例如 *Sporobolomyces* 属和布勒弹孢酵母属 (*Bullera*)) 和隐球酵母科 (*Cryptococcaceae*) (例如假丝酵母属)。因为酵母的分类在将来可能会改变, 就本发明的目的而言, 应按照《酵母的生物学和活性》(Skinner, F. A., Passmore, S. M., 和 Davenport, R. P. 编, Soc. App. Bacteriol, 论文集 No. 9, 1980) 中的描述来定义酵母。酵母的生物学和遗传学操作是本领域所熟知的 (参见例如, 《酵母的生物化学和遗传学》, Bacil, M., Horecker, B. J. 和 Stopani, A. O. M. 编, 第二版,

1987; 《酵母》, Rose, A. H. 和 Harrison, J. S. 编, 第二版, 1987; 以及《酵母属酵母的分子生物学》, Strathern 等编, 1981)。

在一个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、克鲁维酵母属、酵母属、裂殖酵母属、毕赤酵母属或 *Yarrowia* 属内的种的细胞。

在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、诺地酵母或卵形酵母 (*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, 真菌宿主细胞是一种丝状真菌细胞。“丝状真菌”包括真菌门 (*Eumycota*) 和卵菌门亚支的所有丝状形态 (如 Hawksworth 等定义的, 1995, 出处同前)。丝状真菌的特点在于其菌丝壁由壳多糖、纤维素、葡聚糖、脱乙酰壳多糖、甘露聚糖和其他复杂多糖组成。它们通过菌丝的延伸进行营养生长, 碳代谢是专性好氧的。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进行的, 碳代谢可以是发酵。在一个更优选的实施方案中, 丝状真菌宿主细胞是以下种但不限于这些种的细胞: 枝顶孢属 (*Acremonium*)、曲霉属、镰孢属、腐质霉属、毁丝霉属、毛霉属、脉孢菌属、青霉属、梭孢壳属、*Tolyposcladium* 属及木霉属。

在一个甚至更优选的实施方案中, 该丝状真菌宿主细胞是曲霉属细胞。在另一个更优选的实施方案中, 丝状真菌宿主细胞是枝顶孢属细胞。在另一个更优选的实施方案中, 丝状真菌宿主细胞是镰孢属细胞。在另一个更优选的实施方案中, 丝状真菌宿主细胞是腐质霉属细胞。在另一个更优选的实施方案中, 丝状真菌宿主细胞是毁丝霉属细胞。在另一个更优选的实施方案中, 丝状真菌宿主细胞是毛霉属细胞。在另一个更优选的实施方案中, 丝状真菌宿主细胞是脉孢菌属细胞。在另一个更优选的实施方案中, 丝状真菌宿主细胞是青霉属细胞。在另

一个更优选的实施方案中，丝状真菌宿主细胞是梭孢壳属细胞。在另一个更优选的实施方案中，丝状真菌宿主细胞是 *Tolyposcladium* 属细胞。在另一个更优选的实施方案中，丝状真菌宿主细胞是木霉属细胞。

在一个最优选的实施方案中，该丝状真菌宿主细胞是泡盛曲霉细胞、臭曲霉细胞、日本曲霉细胞、黑曲霉细胞或米曲霉细胞。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、禾谷镰孢、禾赤镰孢、异孢镰孢、合欢镰孢、尖镰孢、多枝镰孢、粉红镰孢、接骨木镰孢、肤色镰孢、拟分枝孢镰孢、硫色镰孢、*Fusarium torulosum*、*Fusarium trichothecioides* 或 *Fusarium venenatum*。在一个甚至更最优选的实施方案中，丝状真菌宿主细胞是 *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 insolens* 或 *Humicola lanuginosa* 细胞。在另一个最优选的实施方案中，丝状真菌宿主细胞是米赫毛霉细胞。在另一个最优选的实施方案中，丝状真菌宿主细胞是嗜热毁丝霉细胞。在另一个最优选的实施方案中，丝状真菌宿主细胞是粗糙链孢菌细胞。在另一个最优选的实施方案中，丝状真菌宿主细胞是产紫青霉细胞。在另一个最优选的实施方案中，丝状真菌宿主细胞是 *Thielavia terrestris* 细胞。在另一个最优选的实施方案中，木霉细胞是 *Trichoderma harzianum*、康宁木霉、*Trichoderma longibrachiatum*、*Trichoderma reesei* 或绿色木霉细胞。

可以通过已知方式，利用包括原生质体形成、原生质体转化和细胞壁再生的过程来转化真菌细胞。在 EP238023 中和 Yelton 等 (1984, 美国科学院学报 81: 1470-1474) 中描述了转化曲霉宿主细胞的合适步骤。Malardier 等 (1989, 基因 78: 147-156) 或 W096/00787 描述了转化镰孢的合适方法。可以使用 Becker 和 Guarente (在 Abelson, J.N. 和 Simon, M. I. 编, 酵母遗传和分子生物学指南, 《酶学方法》, 194 卷, 182-187 页, Academic Press, Inc., New York)、Ito 等 (1983, 细菌学杂志 153:163) 和 Hinnen 等 (1978, 美国科学院学报 75:1920)

描述的步骤转化酵母。对于哺乳动物细胞可以采用 Graham 和 Van der Eb 的磷酸钙沉淀法 (1978, 病毒学 52: 546) 通过直接摄取进行转化。

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) 培养能产生所述多肽的野生型菌株, 从而得到包含该多肽的上清液; 和 (b) 纯化多肽。优选地, 该菌株是曲霉属菌株。

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) 纯化该多肽。

本发明进一步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含有一个新转录单位插入其中的同源重组细胞, 该转录单位包括调控序列、一个外显子和/或剪接供体位点与编码多肽的内源核酸序列的第二外显子可操纵连接; 和 (b) 纯化该多肽。这些方法基于使用例如美国专利 5641670 中描述的基因激活技术。基因激活技术基于激活一个正常情况下在细胞中不表达的基因, 或者增加细胞内低水平表达的基因的表达。基因激活技术包括那些将一个含有调控序列、外显子和/或剪接供体位点的外源 DNA 构建体插入细胞基因组 DNA 的方法, 而插入方式使得产生一个新的转录单位, 其中的调控序列、外显子和/或剪接供体位点与内源基因可操纵连接并能激活它进行表达。

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基质中培养细胞。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、在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 (参见, 例如细菌和酵母的参考书; Bennett, J. W. 和 LaSure, L. 编, 《真菌中的基因操作》, Academic Press, CA, 1991)。合适的培养基有供应商提供或者可以参照公开的成分 (例如, 美国典型培养物保藏中心的目录) 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提纯多肽。如果多肽不分泌，可以从细胞裂解物中纯化。

可以用本领域已知的所述多肽的特异方法检测多肽。这些检测方法包括使用特异抗体、形成酶产物或酶底物的消失。例如，可以用酶测定法确定多肽的活性。用于确定氨肽酶活性的步骤是本领域已知的，包括，例如在 405nm 测量 N-苄酰基对硝基苯胺的起始水解率。

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、萃取、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、疏水层析、层析聚焦、和大小排阻层析）、电泳操作（例如，制备性等电点聚焦）、分级溶解（例如硫酸铵沉淀）、SDS-PAGE 或萃取（参见例如，蛋白质纯化，J.C.Janson 和 Lars Ryden 编，VCH Publishers, New York, 1989）。

去除或减少氨肽酶活性

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，从而得到多肽产量比亲本减少的突变细胞。

可以将有氨肽酶活性的多肽在细胞中进行表达所必需的核酸序列修饰或使它失活来方便地构建出氨肽酶活性减少的菌株。要进行修饰或失活的核酸序列可以是，例如编码所述多肽或多肽具有氨肽酶活性所必需的部分的核酸序列，或者是由核酸序列的编码序列表达为多肽时所需要的有调控功能的核酸序列。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包括，但不限于前导序列、多聚腺苷酸化序列、前肽序列、信号序列和终止子。

可以对细胞进行诱变处理并挑选出其氨肽酶产生能力减少或消失

的细胞，从而使核酸序列被修饰或失活。可以通过例如使用合适的物理或化学诱变剂、使用合适的寡核苷酸、或将 DNA 序列做 PCR 进行（特异性或随机）诱变。另外，可以将这些诱变剂组合起来进行诱变。

适用于本发明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，通常于合适的条件下，在有突变剂存在的情况下培养细胞使其突变，并挑选出氨肽酶活性呈现减少或不表达的细胞。

可以通过在编码多肽的核酸序列或者其转录或翻译所必需的调控元件中导入、取代或删除 1 或多个核苷酸来完成本发明所述多肽的修饰或失活。例如，可以插入或删除核苷酸来导入一个终止子、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，即直接在待修饰核酸序列的表达细胞上进行，但优选象下面的例子一样在体外进行修饰。

使所用宿主细胞表达失活或减少的方便方法的例子是基于基因取代或基因破坏技术。例如，在基因破坏方法中，在体外将对应于目标内源基因或基因片段的核酸序列进行诱变，从而产生缺陷的核酸序列，然后将该序列转化到宿主细胞中，以产生缺陷型基因。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。可能希望缺陷基因或基因片段还编码标记物，以使用该标记物挑选出编码多肽的基因已经被修饰或破坏的转化子。

可选择地，利用成熟的反义技术，用与多肽编码序列互补的核苷酸序列将编码本发明所述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，从而减少或消除细胞的多肽产量。在能使得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减少或消除。

优选地，按照本发明所述方法进行修饰的细胞是微生物来源的，例

如是适合用来产生（对于细胞是同源或异源的）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。在本文中，术语“异源多肽”此处定义为宿主细胞的非天然多肽、其中有修饰而使天然序列改变的天然蛋白质，或者因采用重组 DNA 技术操作宿主细胞而使其表达量有改变的天然蛋白质。

另一方面，本发明涉及通过发酵能产生本发明所述多肽和目标蛋白质产物的细胞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，以及任选将回收的产物进一步纯化。

本发明另一可选择的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，其中的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较大地减少氨肽酶活性，以及从培养液中回收产物。可选择地，针对从培养液中回收的酶制剂进行 pH 和温度的组合处理。任选地，可以将 pH 和温度的组合处理与用氨肽酶抑制剂进行的处理结合使用。

优选在 pH9-11、40-70℃ 的范围内，使 pH 和温度的组合处理进行充分的时间以达到预期的效果，通常 30-60 分钟即可。

根据本发明的这一方面，可以将氨肽酶活性消除至少 60%、优选至少 75%、更优选至少 85%、还要优选的是至少 95%，最优选至少 99%。预计通过使用这些方法可以完全消除氨肽酶活性。

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、淀粉葡萄糖苷酶、糖酶、羧肽酶、过氧化氢酶、纤维素酶、壳多糖酶、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、 β -半乳糖苷酶、葡萄糖淀粉酶、葡萄糖氧化酶、葡萄糖苷酶、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、裂解酶、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。可以用氨肽酶缺陷的细胞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修饰使它们的活性、热稳定性、pH耐受能力等增强的其它一些多肽（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。

蛋白质水解产物的制备方法

可以将本发明所述多肽用来生产蛋白质水解产物以便提高水解程度和改善风味。

本发明还涉及将本发明所述多肽与内肽酶结合使用来由富含蛋白质的原材料生产高度水解产物的方法。该方法包括用所述多肽和内肽酶处理蛋白类底物。可以用诸酶同时或连续处理底物。

将本发明所述多肽加入到蛋白类底物中，加入量是蛋白质水解过程中常规采用的有效量，优选每 100g 蛋白质在大约 0.1 到 100000 氨肽酶单位，更优选每 100g 蛋白质在大约 1 到 10000 氨肽酶单位。本文将一个氨肽酶单位（APU）定义为在特定条件下，每分钟从 N-亮氨酸对

硝基苯胺 (Leu-p-nitroanilide, Leu-pNA) (Sigma Chemical Co., St. Louis MO) 中释放 1 微摩尔 N-某酰基对硝基苯胺 (p-nitroanilide) 所需的酶量。可选择地, 氨肽酶的用量优选在大约 0.5 到 500LAPU/g 蛋白质, 更优选在大约 5 到 50LAPU/g 蛋白质。LAPU 定义为按照 AF298/1-GB (可以向 Novo Nordisk A/S, Denmark 索取) 中所述确定的亮氨酸氨肽酶活性。

内肽酶可以得自芽孢杆菌属的菌株(优选地衣芽孢杆菌或枯草芽孢杆菌), 葡萄球菌属的菌株(优选金黄色葡萄球菌), 链霉菌属的菌株[优选热普通链霉菌 (*S. thermovularis*) 或灰色链霉菌], 放线菌菌株, 曲霉属菌株(优选棘孢曲霉、泡盛曲霉、臭曲霉、构巢曲霉、黑曲霉或米曲霉), 或者镰孢属菌株(优选 *Fusarium venenatum*)。

将内肽酶以在蛋白质水解过程中常规采用的有效量加入蛋白类底物中, 加入量优选在大约 0.05 到 15AU/100g 蛋白质, 更优选在大约 0.1 到 8AU/100g 蛋白质。一个 AU (Anson Unit) 定义为在标准条件下(即, 25°C, pH7.5 和反应 10 分钟)降解血红蛋白的起始速率使得每分钟释放出 TCA 可溶性产物的量与酚试剂产生如 1 毫当量酪氨酸相同颜色的酶量。分析方法 AF4/5 可以向 Novo Nordisk A/S, Denmark 索取, 该文此处引作参考文献。

所述酶法处理, 即底物与酶制剂的共同保温, 可以在不会使酶制剂失活的任何方便温度进行, 优选在大约 20°C 到 70°C。根据成熟的作法, 可以通过将保温体系的温度升高到酶的失活温度(例如升高到大约 70°C 以上), 或类似地, 通过将保温体系的 pH 降低到酶失活的 pH 值(例如, 低于约 pH4.0), 来使酶制剂适当地失活。

另外, 本发明所述方法使蛋白类底物水解度增加。文中使用的水解度 (DH) 是蛋白质中被蛋白水解酶水解的氨基键占全部氨基键的百分比。在一个优选实施方案中, 蛋白质水解产物中的亮氨酸、甘氨酸、谷氨酸、丝氨酸、天冬氨酸、天冬酰胺、脯氨酸、半胱氨酸、丙氨酸和/或谷氨酰胺含量增加, 例如至少增加了 1.1 倍。在一个更优选的实施方案中, 蛋白质水解产物的亮氨酸含量增加。在另一个更优选的实

施方案中，蛋白质水解产物的 Gly 含量增加。在另一个更优选的实施方案中，蛋白质水解产物的 Glu 含量增加。在另一个更优选的实施方案中，蛋白质水解产物的 Ser 含量增加。在另一个更优选的实施方案中，蛋白质水解产物的 Asp 含量增加。在另一个更优选的实施方案中，蛋白质水解产物的 Asn 含量增加。在另一个更优选的实施方案中，蛋白质水解产物的 Pro 含量增加。在另一个更优选的实施方案中，蛋白质水解产物的 Cys 含量增加。在另一个更优选的实施方案中，蛋白质水解产物的 Ala 含量增加。在另一个更优选的实施方案中，蛋白质水解产物的 Gln 含量增加。

本发明还涉及用于获得富含游离谷氨酸和/或肽键合的谷氨酸残基的蛋白质水解产物的方法，该方法包括：

- (a) 使底物进行脱酰胺过程；和
- (b) 用具有氨肽酶活性的多肽作用于底物。

这两个步骤可以同时进行，或者在第一步骤后进行第二步骤。

本发明所述方法能产生有极佳风味的蛋白质水解产物，因为无论是游离还是肽键合形式的谷氨酸 (Glu) 都对蛋白质水解产物的风味和口味有重要影响。这些方法还产生功能改善的蛋白质水解产物，具体说是溶解性提高，乳化性能提高、水解程度增加和发泡性能改善。

通过释放氨将酰胺类化合物 (谷氨酰胺或天冬酰胺) 转化为带电荷的酸 (谷氨酸或天冬氨酸) 称为脱酰胺反应。脱酰胺反应可以以非酶法或酶法脱酰胺过程进行。

在一个优选实施方案中，以酶法脱酰胺过程进行脱酰胺反应，例如使底物接受转谷氨酰胺酶和/或肽谷氨酰胺酶的作用。

转谷氨酰胺酶可以是任何方便的来源，包括哺乳动物 (参见例如 JP1050382 和 JP5023182)，包括活化的因子 XIII (参见例如 W093/15234)；来源于鱼 (参见例如 EP555649)；以及得自微生物的转谷氨酰胺酶 (参见例如 EP379606, W096/06931 和 W096/22366)。在一个优选的实施方案中，转谷氨酰胺酶得自卵菌纲细胞，包括疫霉属 (*Phytophthora*) 菌株，优选恶疫霉，或腐霉属 (*Pythium*) 菌株，

优选畸雌腐霉、腐霉菌、中间型腐霉、终极腐霉、周雄腐霉（或缠器腐霉）。在另一个实施方案中，所述转谷氨酰胺酶是细菌来源的，得自芽孢杆菌属菌株（优选枯草芽孢杆菌），链轮丝菌属（*Streptoverticillium*），优选茂原链轮丝菌（*Streptoverticillium mobarraensis*），灰肉色链轮丝菌（*Streptoverticillium griseocarneum*），肉桂色链轮丝菌（*Streptoverticillium cinnamoneum*），和链霉菌属菌株，优选利迪链霉菌。

所述肽谷氨酰胺酶可以是肽谷氨酰胺酶 I（肽基谷氨酰胺酶；EC3.5.1.43）或肽谷氨酰胺酶 II（蛋白-谷氨酰胺谷氨酰胺酶；EC3.5.1.44）或者它们的任何混合物。该肽谷氨酰胺酶可以得自曲霉属的菌株（优选日本曲霉）、芽孢杆菌属的菌株（优选环状芽孢杆菌）、隐球酵母菌株（优选浅白隐球酵母）或德巴利酵母属菌株，（优选克洛德巴利酵母（*Debaryomyces kloecheri*））。

将脱酰胺过程中常规采用的有效量的转谷氨酰胺酶加入蛋白类底物中，加入量优选在大约 0.01 到 5%（w/w），更优选在大约 0.1 到 1%（w/w）（酶制剂相对底物的用量）。

将脱酰胺过程中常规采用的有效量的肽谷氨酰胺酶加入蛋白类底物中，加入量优选在大约 0.01 到 100,000 PG 酶单位/100g 底物，更优选在大约 0.1 到 10,000 PG 酶单位/100g 底物。

可以根据 Cedrangoro 等的步骤（1965，酶学（Enzymologia）29:143）来确定肽谷氨酰胺酶活性。按照该步骤，将 0.5ml 酶样品用 1N NaOH 调节到 pH6.5，倒入一个小容器中。然后向容器中加入 1ml 硼酸盐缓冲液（pH10.8）。用 5N 硫酸吸收释放出的氨，并用 Nessler's 试剂使混合体系呈色，于 420nm 测量。一个 PG 酶单位是能在这些条件下每分钟产生 1 微摩尔氨的酶量。

可选择地，根据美国专利 3857967 或下文实施例 20 中描述的步骤来确定肽谷氨酰胺酶活性。

在本发明所述方法的（b）步，用本发明所述多肽作用于底物。按照蛋白质水解过程中常规采用的有效量将本发明所述多肽加入蛋白类

底物中，优选用量在大约 0.001 到 0.5AU/100g 底物，更优选在大约 0.01 到 0.1 AU/100g 底物。

在另一个实施方案中，本发明所述用于制备富含游离谷氨酸和/或肽键合谷氨酸残基的水解产物的方法还包括：

(c) 使底物接受 1 或多种非特异性内肽酶和/或外肽酶的作用。

该步骤可以与步骤 (a) 和 (b) 同时进行，或者在 (a) 和 (b) 后进行。

在一个优选实施方案中，所述非特异性内肽酶和/或外肽酶得自曲霉属的菌株（优选黑曲霉、米曲霉或酱油曲霉）、或芽孢杆菌属的菌株（优选解淀粉芽孢杆菌、缓慢芽孢杆菌、地衣芽孢杆菌或枯草芽孢杆菌）。

按照蛋白质水解过程中常规采用的有效量将非特异性作用的内肽酶和/或外肽酶加入底物中，优选用量在大约 0.05 到 15CPU/100g 底物，更优选在大约 0.1 到 5CPU/100g 底物。一个 CPU（酪蛋白蛋白酶单位）定义为在标准条件下（即，25℃，pH9.5 保温 30 分钟）每分钟从酪蛋白中释放出 1 微摩尔伯氨基团（通过与标准丝氨酸比较确定）的酶量。分析方法 AF228/1（该文此处引作参考文献）可以向 Novo Nordisk A/S, Bagsvaerd, Denmark 索取。

可以在酶制剂不会失活的任何温度下进行各种酶法处理，优选在大约 20℃ 到 70℃ 进行。可以通过将温度升高到例如大约 70℃ 以上或者将 pH 降低到例如 pH4.0 以下来使酶制剂失活。

用于本发明所述方法的蛋白类底物可以由完整蛋白质、预先水解的蛋白质（即肽）或者它们的混合物组成。所述蛋白类底物可以来源于植物或动物。优选地，蛋白类底物来源于植物，例如大豆蛋白质、谷物蛋白质（例如小麦谷蛋白、玉米谷蛋白、大麦、黑麦、燕麦、大米）玉米蛋白、羽扇豆、棉籽蛋白、油菜籽蛋白、花生、紫花苜蓿蛋白、豌豆蛋白、蚕豆蛋白、芝麻种子蛋白或向日葵。来源于动物的蛋白类底物可以是乳清蛋白质、酪蛋白、肉蛋白、鱼蛋白、血红细胞、蛋清、明胶或乳白蛋白。

本发明还涉及由这些方法制备的蛋白质水解产物。

其他用途

本发明还涉及用本发明所述多肽使酶失活的方法。

另外，本发明所述多肽可以用于多种需要特异性裂解肽序列的目的。例如，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）合成某些蛋白质或肽。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。

组合物

本发明更进一步的方面涉及包含本发明所述多肽的多肽组合物。优选地，这些组合物富含本发明所述多肽。在本文中，术语“富含”是指多肽组合物的氨肽酶活性提高了（富集系数为1.1）。

这种多肽组合物可以包含本发明所述多肽，并且是以该多肽为主要的酶组分，例如单组分多肽组合物。可选择地，组合物可包含多种酶活性，如氨肽酶、淀粉酶、糖酶、羧肽酶、过氧化氢酶、纤维素酶、壳多糖酶、角质酶、环糊精糖基转移酶、脱氧核糖核酸酶、酯酶、 α -半乳糖苷酶、 β -半乳糖苷酶、葡糖淀粉酶、 α -葡糖苷酶、 β -葡糖苷酶、卤素过氧化物酶、转化酶、漆酶、脂酶、甘露糖苷酶、氧化酶、果胶水解酶、肽谷氨酰胺酶、过氧化物酶、植酸酶、多酚氧化酶、蛋白水解酶、核糖核酸酶、转谷氨酰胺酶或木聚糖酶。其他的酶可以通过以下属的微生物来制备：曲霉属（优选棘孢曲霉、泡盛曲霉、黑曲霉或米曲霉）或木霉属、腐质霉属（优选 *Humicola insolens*）或镰孢属（优选杆孢状镰孢、*Fusarium cerealis*、*Fusarium crookwellense*、大刀镰孢、禾谷镰孢、禾赤镰孢、异孢镰孢、合欢镰孢、尖镰孢、多枝镰孢、粉红镰孢、接骨木镰孢、肤色镰孢、拟分枝孢镰孢、硫色镰孢、*Fusarium torulosum*、*Fusarium trichothecioides* 或 *Fusarium venenatum*）。

在一个优选实施方案中，本发明涉及一种改善风味的组合物，它包

含具有氨肽酶活性的多肽和合适的载体。可以使用任何本领域已知的任何合适载体。在另一个优选实施方案中，这种改善风味的组合物还包含内肽酶。在另一个优选实施方案中，这种改善风味的组合物还包含1或多种非特异性作用的内肽酶和/或外肽酶。在另一个优选实施方案中，这种改善风味的组合物还包含1或多种特异性作用的内肽酶和/或外肽酶。

在一个优选实施方案中，所述特异性作用的蛋白水解酶是一种内肽酶，比如谷氨酰内肽酶(EC3.4.21.19)；赖氨酰内肽酶(EC3.4.21.50)；亮氨酰内肽酶(EC3.4.21.57)；甘氨酰内肽酶(EC3.4.22.25)；脯氨酰内肽酶(EC3.4.21.26)；胰蛋白酶(EC3.4.21.4)或胰蛋白酶样(赖氨酸/精氨酸特异性的)内肽酶；或肽酰-Asp 金属内肽酶(EC3.4.24.33)。

优选谷氨酰内肽酶(EC3.4.21.19)得自芽孢杆菌属菌株(尤其是地衣芽孢杆菌和枯草芽孢杆菌)、葡萄球菌属菌株(尤其是金黄色葡萄球菌)、链霉菌属菌株(尤其是热普通链霉菌和变灰链霉菌)或放线菌属菌株。

优选赖氨酰内肽酶(EC3.4.21.50)得自无色杆菌属菌株(尤其是水解无色杆菌)、溶杆菌属菌株(尤其是产酶溶杆菌)或者假单胞菌属菌株(尤其是铜绿假单胞菌)。

亮氨酰内肽酶(EC3.4.21.57)可以来源于植物。

优选甘氨酰内肽酶(EC3.4.22.25)得自木瓜植物(番木瓜)；

优选脯氨酰内肽酶(EC3.4.21.26)得自黄杆菌属的菌株，或者来源于植物。

优选胰蛋白酶样内肽酶得自镰孢属的菌株(尤其是例如W089/06270或W094/25583中描述的尖镰孢)。

优选肽酰-Asp 金属内肽酶(EC3.4.24.33)得自假单胞菌(尤其是莓实假单胞菌)。

在另一个优选实施方案中，所述特异作用的蛋白水解酶是可以从肽的任一端开始作用的内肽酶。

在一个优选实施方案中，所述特异作用的蛋白水解酶是一种氨肽酶，比如亮氨酸氨肽酶（EC3.4.11.1）或三肽氨肽酶（EC3.4.11.4）。

在另一个优选实施方案中，所述特异作用的蛋白水解酶是一种羧肽酶，比如脯氨酸羧肽酶（EC3.4.16.2）、羧肽酶 A（EC3.4.17.1）、羧肽酶 B（EC3.4.17.2）、羧肽酶 C（EC3.4.16.5）、羧肽酶 D（EC3.4.16.6）、赖氨酸（精氨酸）羧肽酶（EC3.4.17.3）、甘氨酸羧肽酶（EC3.4.17.4）、丙氨酸羧肽酶（EC3.4.17.6）、谷氨酸羧肽酶（EC3.4.17.11）、肽酰二肽酶 A（EC3.4.15.1）或肽酰二肽酶（EC3.4.15.5）。

可以依照本领域已知的方法来制备液态或干燥形式的所述多肽组合物。通过本领域已知方法可使多肽稳定化。

本发明还涉及食品，例如包含由本领域方法得到的蛋白水解产物的烤制食品。这种食品的感官特性提高，比如风味、口味、口感、香味和外皮色泽都有改善。

在本文中，术语“烤制食品”包括任何从生面团生产的松软或酥脆的食物。可以通过本发明方便地生产出的烤制食品的例子（无论是白色、浅色或黑色类的）是面包（尤其是白色、全麦或燕麦面包，通常做成块状或面包卷的形式）、法国长方形面包、比萨饼面包、玉米面豆卷、蛋糕、薄烤饼、饼干、脆面包等。

这些烤制食品通常是用含有面粉和水的生面团生产的，并且通常要经过发酵。可以用多种方式发酵生面团，比如通过加入碳酸氢钠等，或者通过加入酵面（发酵过的生面团），但是优选通过加入合适的酵母培养物，比如酿酒酵母（面包酵母）来发酵。可以选用任何可以购买到的酿酒酵母菌株。

另外，用于生产烤制食品的生面团可以是新鲜或冷冻的。K. Kulp 和 K. Lorenz 在《冷冻和冷藏的生面和奶油面糊》中描述了冷冻生面团的生产方法。在生面团中通常包含 0.01-5%，更优选 0.1-3%的本发明所述改善风味的组合物。

在本发明所述方法中，可以单独或同时将本发明所述多肽、内肽

酶、转谷氨酰胺酶、肽谷氨酰胺酶、1或多种特异和/或非特异作用的内肽酶和/或外肽酶、和/或1或多种以上具体指出的酶类加入制备生面团的混合物中或加入用来制备生面团的任何成分（例如面粉）中。

本发明还涉及一种制备生面团或由生面团生产烤制食品的预混合物，例如，采取面粉组合物的形式，其中所述的预混合物包含本发明所述多肽或改善风味组合物以及载体或烤制成分，并且任选含有1或多种上述具体指出的其他酶类。

在另一个实施方案中，这种预混合物包含由本发明所述方法获得的水解产物。

可以通过将相关酶类与合适的载体（比如面粉、淀粉、糖或盐）混合来制备预混合物。这种预混合物可以含有其他面团改善和/或面包改善添加剂。

本文中，术语“预混合物”是烘烤物（通常包括面粉）的混合物，它是预先制备好的，可以在指定条件下储存，并可以给生面团的制备过程带来便利。这种预混合物可以方便地用于工业和商业中烘烤面包的工厂和作坊，以及零售面包店。

本发明还涉及用本发明所述方法制备的水解产物作为食品（比如烤制食品）添加剂来提高感官品质，比如风味、口味和香味。

可以将通过本发明所述方法得到的富含游离谷氨酸和/或肽键合谷氨酸残基的水解产物用于多种工业应用，尤其是需要含有功能性蛋白质的应用。

例如，本发明还涉及食品，这些食品包含由本发明所述方法得到的富含游离谷氨酸和/或肽键合谷氨酸残基的水解产物，还涉及包含由本发明所述方法得到的富含游离谷氨酸和/或肽键合谷氨酸残基的水解产物的动物饲料添加剂。

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: 纯化 FLAVOURZYME™ 氨肽酶 II

从 FLAVOURZYME™ 培养液 (Novo Nordisk A/S, Bagsvaerd, Denmark) 中纯化氨肽酶。FLAVOURZYME™ 培养液是通过在含有碳和氮源以及微量金属的培养基中培养米曲霉菌株 1568 (ATCC20386) 得到的。首先, 用 180ml 20mM 磷酸钠 (pH7.0) 缓冲液稀释培养液 (20ml, 含有 720mg 蛋白质) 并用配备 0.45 μ m 滤膜 (Nalgene, Rochester, NY) 的 Nalgene Filterware 装置将其过滤。将滤过液上样到含有 31ml Q-Sepharose, Big Beads (Pharmacia Biotech AB, Uppsala, Sweden) 的 24 \times 130mm 柱子中, 这个柱子已用 20mM 磷酸钠 (pH7.0) 缓冲液预平衡过。利用从 7.0 (20mM 磷酸钠缓冲液) 到 5.0 (20mM 醋酸钠缓冲液)、从 5.0 到 3.5 (20mM 醋酸钠缓冲液) 然后从 3.5 到 3.0 (20mM 醋酸钠缓冲液) 的 pH 梯度洗脱蛋白质。收集 pH 3.5 到 3.0 之间洗脱的组分, 将其汇合, 并用 PM10 滤膜 (Amicon, New Bedford, MA) 经超滤浓缩到 20ml。

浓缩好的溶液用 100ml 20mM 磷酸钠 (pH7.0) 缓冲液稀释, 然后上样到含有 Pharmacia MonoQ Beads (Pharmacia Biotech AB, Uppsala, Sweden) 的 20 \times 100mm 柱子中, 这个柱子已用 20mM 磷酸钠 (pH7.0) 缓冲液预平衡过。用 20mM 磷酸钠 (pH7.0) 缓冲液中的 0 到 0.4M NaCl 梯度洗脱蛋白质。收集 0.330 到 0.343M NaCl 之间的组分, 将其汇合, 并用对 20mM 醋酸钠缓冲液 (pH4.0) 的超滤进行浓缩。

经 SDS-PAGE 分析发现纯化过的制品含有三个主要条带。样品由分子量为大约 65、50 和 33kDa 的几种组分构成。

实施例 2: 氨肽酶 II 的氨基酸测序

将 1 份实施例 1 中描述的纯氨肽酶 II 制剂电泳, 随后用在 10% 甲醇中的 10mM CAPS (3-(环己基氨基)-1-丙磺酸) pH11 将其转印到 PVDF 膜 (Novex, San Diego, CA) 上 2 小时。用含 0.1% 考马斯蓝 R-250 的 40% 甲醇/1% 醋酸将 PVDF 膜染色 20 秒, 然后在 50% 乙醇中脱色来观察蛋白质条带。切下 65、50 和 33kDa 的三个组分, 并依照制造商的说明用印迹箱和液相 TFA 送递法在 Applied Biosystems 476A 型蛋白质测

序仪 (Applied Biosystems, Inc., Foster City, CA) 上做氨基端测序。三个组分都得到相同的氨基端序列：
RALVSPDEFPEDIQLEDLLEGSQQLEDFAY (SEQ ID NO:2).

将 300 μ l 蛋白质样品在 Savant Speed Vac AS160 (Savant Instruments, Farmingdale, NY) 中干燥，然后用 300 μ l 70% 蚁酸 (水性) 重新溶解。加入几粒溴化氰晶体，并置暗处于室温保温过夜。将样品再次在 Speed Vac 中干燥，并在 Tricine 样品缓冲液 (Novex, San Diego, CA) 中重新溶解。用 10-20% Tricine SDS-PAGE 将溴化氰裂解的片段分离成 6、10、15、22、27、40 和 50kDa 的条带，并转印到 PVDF 膜上。切下 6、10、15 和 22 kDa 的条带，进行氨基端测序。

15 和 22 kDa 条带的氨基端序列与上述氨基端序列相同，而 6 和 10 kDa 条带确定出含有序列：

TYSPSVEVTADVAVVKNLGTSEADYPDVEGKVAL (SEQ ID NO:2).

实施例 3: 分离米曲霉菌株 1568 的 RNA

在装有每升含 7.5g 马铃薯淀粉、10g 大豆粉、2g KH_2PO_4 、5g $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$ 和 0.1g $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ 的培养基的发酵罐中培养米曲霉菌株 1568。于 30 $^\circ\text{C}$ 生长 5 天后取出 2 升样品，收集菌丝，在液 N_2 中冷冻，保存在 -80 $^\circ\text{C}$ 。用硫氰酸胍进行抽提，随后经 5.7M 氯化铯超速离心，从冷冻的粉碎菌丝中制备总 RNA (Chirgwin 等, 1979, 生物化学 18: 5294-5299)。依照 Aviv 和 Leder (1972, 美国科学院学报 69: 1408-1412) 描述的寡(dT)-纤维素亲和层析分离 Poly(A) + RNA。

实施例 4: 构建 cDNA 文库

采用 Gubler 和 Hoffman (1983, 基因 25:263-269) 以及 Sambrook 等 (1989, 分子克隆: 实验指南, Cold Spring Harbor Press, Cold Spring Harbor, New York) 描述的步骤，由 5 μ g 实施例 3 中得到的米曲霉菌株 1568 Poly(A) + RNA 来合成双链 cDNA，但在第一链反应中使用了寡聚(dT)-NotI 锚定引物，而不是寡聚(dT) 12-18 引物。合成结束后，

用绿豆核酸酶 (Life Technologies, Gaithersburg, MD) 处理 cDNA, 用 T4 DNA 聚合酶 (Boehringer Mannheim, Indianapolis, IN) 形成平末端, 并使其与非回文结构的 BstXI 衔接子 (Invitrogen, San Diego, CA) 连接 (用大约 50 倍摩尔过量的衔接子)。用 NotI 消化连接好的 cDNA, 经琼脂糖凝胶电泳根据体积筛分出 1.2-3.0kb cDNA 片段, 并连接到用 BstXI/ NotI 裂解过的 pYES2.0 (Invitrogen, San Diego, CA) 中。依照制造商的说明将连接体系转化到电感受态的大肠杆菌 DH10B 细胞 (Life Technologies, Gaithersburg, MD) 中。将由 1×10^6 个独立克隆组成的文库于 20% 甘油中以单个池 (25000-30000 个集落形成单位/池) 的形式保存在 -80°C , 双链 cDNA 和连接体系保存在 -20°C 。

实施例 5: 提取基因组 DNA

米曲霉 1568 在 25ml 0.5% 酵母提取物-2% 葡萄糖 (YEG) 培养基中于 37°C 、250rpm 生长 24 小时。经 Miracloth (Calbiochem, La Jolla, CA) 过滤收集菌丝, 并用 25ml 10mM Tris-1mM EDTA (TE) 缓冲液洗一次。吸掉菌丝制剂上的多余缓冲液, 随后将其冷冻在液氮中。将冷冻的菌丝制品在电动咖啡磨中磨成细粉末, 将粉末加入含有 20ml TE 缓冲液和 5ml 20% w/v 十二烷基磺酸钠 (SDS) 的一次性塑料离心管中, 将混合液轻轻颠倒几次以确保混合均匀, 用等体积酚: 氯仿: 异戊醇 (25: 24: 1 v/v/v) 抽提两次。向抽提过的样品中加入醋酸钠 (3M 溶液) 至终浓度为 0.3M, 随后用 2.5 体积的冰乙醇沉淀 DNA。将离心管于 $15000 \times g$ 离心 30 分钟来沉淀 DNA。将 DNA 沉淀风干 30 分钟后重新悬浮于 0.5ml TE 缓冲液中。将不含 DNase 的核酸酶 A 加入重悬的 DNA 沉淀中至浓度为 $100\mu\text{g}/\text{ml}$, 然后将混合液于 37°C 保温 30 分钟。加入蛋白酶 K ($200\mu\text{g}/\text{ml}$), 离心管于 37°C 再保温 1 小时。最后, 用酚: 氯仿: 异戊醇抽提两次, 用乙醇沉淀 DNA。沉淀的 DNA 用 70% 乙醇洗涤, 真空下干燥, 重新悬浮在 TE 缓冲液中, 保存在 4°C 。

实施例 6: PCR 扩增米曲霉 1568 氨肽酶 II

根据实施例 2 中描述的米曲霉 1568 氨肽酶 II 部分肽的氨基酸序列, 依照制造商的说明用 Applied Biosystems 394 型 DNA/RNA 合成仪合成以下简并寡核苷酸引物, 以便从米曲霉 1568 基因组 DNA 经 PCR 来扩增氨肽酶 II 基因。

正向引物: 5'-CCIGAYGARTTYCCIGARGA-3' (SEQ ID NO:3)

反向引物: 5'-RTTYTTIACIACIGCIACRTCIGCIGTIACYTCIAC-3' (SEQ ID NO:4)

(R=A 或 G, Y=C 或 T, N=G 或 A 或 C 或 T, H=A 或 C 或 T, I=次黄苷)

将按照实施例 5 所述制备的大约 1 μ g 米曲霉 1568 基因组 DNA 作为模板准备扩增反应体系 (50 μ l)。每份反应体系含有如下成分: 1 μ g 基因组 DNA、40pmol 正向引物、40 pmol 反向引物、dATP、dCTP、dGTP 和 dTTP 各 200 μ M、1 \times Taq 聚合酶缓冲液 (Perkin-Elmer Corp., Branchburg, NJ.) 和 2.5 单位 Taq 聚合酶 (Perkin-Elmer Corp., Branchburg, NJ.)。将反应体系在 Perkin-Elmer 480 型热循环仪中保温, 如下进行反应: 第 1 个循环 94 $^{\circ}$ C 5 分钟, 50 $^{\circ}$ C 2 分钟, 72 $^{\circ}$ C 2 分钟; 第 2 到 26 个循环 94 $^{\circ}$ C 2 分钟, 50 $^{\circ}$ C 1 分钟, 72 $^{\circ}$ C 2 分钟。在 1.5% 琼脂糖凝胶 (Eastman Kodak, Rochester, NY) 上分离出反应产物, 从胶上切下 309bp 产物条带, 并依照制造商的说明用 Qiaex II (Qiagen, Chatsworth, CA) 将其纯化。随后将纯化好的 PCR 产物克隆到 pCRII 载体 (Invitrogen, San Diego, CA) 中, 用 lac 正向和反向引物 (New England BioLabs, Beverly, MA) 确定 DNA 序列。

用上述氨肽酶 II 特异性 PCR 引物由米曲霉 1568 扩增含有 103 个密码子的氨肽酶 II 基因片段 (309bp)。DNA 序列分析表明扩增得到的基因片段编码相应米曲霉 1568 氨肽酶 II 基因的一部分。用氨肽酶 II 基因片段探测实施例 5 中描述的米曲霉 1568 cDNA 文库。

实施例 7: 鉴定米曲霉 1568 氨肽酶 II 克隆

将米曲霉 1568 cDNA 文库铺在加有 50 μ g/ml 羧苄青霉素的 Luria 琼脂平板上。将大约 10000 个菌落做菌落转移 (Maniatis 等, 1982, 分

子克隆: 实验指南, Cold Spring Harbor Process, Cold Spring Harbor, New York), 并用 UV Stratalinker (Stratagene, La Jolla, CA) 将 DNA 交联到膜 (HybondN+, Amersham, Arlington Heights, IL) 上。将膜于 45°C, 在含有 5×SSPE、0.3%SDS、50%甲酰胺和 10μg/ml 剪切变性鲑精 DNA 的杂交溶液中浸泡 3 小时。用随机引发 DNA 标记试剂盒 (Boehringer Mannheim, Mannheim, Germany) 将实施例 6 中描述的从米曲霉 1568 分离到的氨肽酶 II 基因片段放射标记, 通过加入 NaOH 至终浓度 0.1M 使其变性, 然后加入到杂交溶液中 (放射活性大约 1×10^6 cpm/ml 杂交液)。杂交体系于 45°C 水浴震荡过夜。温育后, 于 55°C 将膜在有 0.2% SDS 的 2×SSC 中洗 3 次。然后将膜在印迹滤纸上干燥 15 分钟, 包在 SaranWrap™ 中, 于 -70°C 用增强屏 (Kodak, Rochester, NY) 对 X-光胶片曝光 48 小时。

有 11 个菌落与探针产生强杂交信号。将这 11 个菌落接种到 5ml 加有 50μg/ml 羧苄青霉素的 LB 培养基中, 并于 37°C 生长过夜。用 Wizard 373 DNA 纯化试剂盒 (Promega, Madison, WI) 由每个菌落制备小量 DNA。通过 DNA 测序证明克隆 9 和 10 含有氨肽酶 II 编码序列。其中克隆 9 (pEJG18) 是全长序列。将质粒 pEJG18 亚克隆到大肠杆菌 DH5α 细胞中得到大肠杆菌 DH5α EJG18。

实施例 8: 米曲霉氨肽酶 II 基因的 DNA 序列分析

利用 Applied Biosystems 373A 型自动 DNA 测序仪 (Applied Biosystems, Inc., Foster City, CA), 采用引物步行技术通过染料终止化学法 (Giesecke 等, 1992, 病毒学方法杂志 (Journal of Virology Methods) 38:47-60) 在双链上测定实施例 7 所述大肠杆菌 DH5α EJG18 中的 pEJG18 所含氨肽酶 II 基因的 DNA 序列。将寡核苷酸引物设计成氨肽酶 II 基因的互补序列, 并依照制造商的说明在 Applied Biosystems 394 型 DNA/RNA 合成仪上合成。

图 1 显示了米曲霉 1568 氨肽酶 II 编码基因的核苷酸序列和由其推测出的氨基酸序列 (分别是 SEQ ID NO: 1 和 2)。对克隆插入片段进

行的序列分析揭示出一个有 1488 个核苷酸(不包括终止密码)的大开放读码框,它编码一个 496 个氨基酸序列的蛋白质(SEQ ID NO: 2)。该开放读码框的 G+C 含量是 58%。根据 van Heijne(van Heijne, 1984, 分子生物学杂志 173: 243-251)的规则,起首的 15 个氨基酸可能包含引导新合成多肽进入内质网的分泌信号肽(图 1 中加双下划线部分)。

实施例 2 中描述的来源于纯化氨肽酶 II 的部分肽,其氨基酸序列在图 1 中加了下划线,它们与米曲霉 1568 氨肽酶 II cDNA 的推测氨基酸序列(SEQ ID NO: 2)中发现的一致。

用 Clustal 比对程序(Higgins, 1989, 出处同前)比较米曲霉 1568 氨肽酶 II 的推测氨基酸序列和酿酒酵母氨肽酶 II Y 的氨基酸序列(SEQ ID NO: 5),观察到相同百分比为 33.7%。

实施例 9: 构建用于曲霉宿主的米曲霉 1568 氨肽酶 II 表达载体

设计了如下两个合成寡核苷酸引物来由质粒 pEJG18(大肠杆菌 DH5 α -EJG18)经 PCR 扩增米曲霉 1568 氨肽酶 II 基因编码序列,以使用于在曲霉宿主中进行亚克隆和表达。

正向引物: 5'-ATGATGAGGTCGCTTTTGTGGGC-3'(SEQ ID NO:6)

反向引物: 5'-GGGATGCATCTATGCCTCGACTT-3'(SEQ ID NO:7)

粗体字代表编码序列。

为了方便将基因片段亚克隆到定名为 pMWR3(图 2)的表达载体中,在氨肽酶 II 基因的 3'端导入 NsiI 限制酶位点。给 5'端加上 ATG 使其成平末端以便插入 SwaI 位点。载体 pMWR3 含有作为调控序列的 TAKA 启动子和终止子。由于质粒不含用于真菌转化的选择标记,用含有可作为选择标记的 amdS 的 pToC90(W091/17243)与其共转化。

在 PCR 反应体系(50 μ l)中使用了各 50 pmol 的上述引物,该反应体系含有 70ng pEJG18(pYES2 中的米曲霉 1568 cDNA 克隆); 1 \times Pwo 缓冲液(Boehringer Mannheim, Indianapolis, IN); 8 μ l 10mM 的 dATP、dTTP、dGTP、dCTP 混合液; 2.5 单位 PwoII(Boehringer Mannheim, Indianapolis, IN)。扩增是在 94 $^{\circ}$ C 2 分钟、55 $^{\circ}$ C 30 秒、72 $^{\circ}$ C 1

分钟的情况下做1个循环；94℃ 15秒、55℃ 30秒、72℃ 1分钟的情况下做9个循环；94℃ 15秒、55℃ 30秒、72℃ 1分钟，每个循环中延长20秒，这种情况下做15个循环；最后一个循环94℃ 15秒、55℃ 30秒、72℃ 7分钟。然后将加热块转入4℃浸泡循环。经凝胶电泳和QiaexII纯化扩增好的1500bp DNA片段。用NsiI消化氨肽酶克隆（使用制造商确定的条件）。将片段进行酚-氯仿抽提和乙醇沉淀。将切下的片段克隆到预先用SwaI和NsiI切割好的pMWR3中，得到表达质粒pEJG19（图3），在该载体中氨肽酶II基因的转录受到TAKA启动子的调控。将质粒pEJG19转化到大肠杆菌DH5α（Life Technologies, Gaithersburg, MD）中。分离到一株含有pEJG19质粒的大肠杆菌转化子，根据Sambrook等描述的步骤（1989，出处同前）制备质粒DNA。

实施例10：在米曲霉中表达米曲霉1568氨肽酶II基因

用以下原生质体转化方法将质粒pEJG19导入碱性蛋白酶缺陷的米曲霉宿主JaL142-6中。进行转化的原生质体浓度为约 2×10^7 个原生质体/ml。将100μl原生质体和约5μg pEJG19、5μg pTOC90置于冰上；加入250μl 60% PEG4000、10mM Tris-HCl（pH7.5）、10mM CaCl₂，然后将原生质体于37℃保温30分钟。加入3ml STC（1.2M 山梨醇、10mM Tris-HCl（pH7.5）和10mM CaCl₂）。将溶液轻轻混匀，倒在COVE转化平板（每升：0.52g KCl、0.52g MgSO₄-7H₂O、1.52g KH₂PO₄、1ml下述微量金属、342.3g蔗糖、25g Noble琼脂、10ml 1M 乙酰胺、10ml 3M CsCl）上。每升该微量金属溶液（1000×）含有22g ZnSO₄-7H₂O、11g H₃BO₃、5g MnCl₂-4H₂O、5g FeSO₄-7H₂O、1.6g CoCl₂-5H₂O、1.6g (NH₄)₆Mo₇O₂₄和50g Na₄EDTA。将平板于37℃保温7天。将转化子转移到含相同培养基的平板上，于37℃保温2天。根据转化子在以乙酰胺为唯一氮源的COVE培养基上的生长能力总共得到140个转化子。

转化子在每孔含有1ml用75% MY50盐稀释的25% MY50培养基之24孔平板中，于34℃、200rpm生长4天。每升MY50中含有50g麦芽

糊精、2.0g $\text{MgSO}_4\cdot 7\text{H}_2\text{O}$ 、10g KH_2PO_4 、2g 柠檬酸、10g 酵母提取物、2.0g 尿素、2g K_2SO_4 和0.5ml 微量元素溶液(pH调至6.0)。每升该微量元素溶液含有14.3g $\text{ZnSO}_4\cdot 7\text{H}_2\text{O}$ 、2.5g $\text{CuSO}_4\cdot 5\text{H}_2\text{O}$ 、0.5g $\text{NiCl}_2\cdot 6\text{H}_2\text{O}$ 、13.8g $\text{FeSO}_4\cdot 7\text{H}_2\text{O}$ 、8.5g $\text{MnSO}_4\cdot \text{H}_2\text{O}$ 、3g 柠檬酸。每升MY50盐含有2.0g $\text{MgSO}_4\cdot 7\text{H}_2\text{O}$ 、10g KH_2PO_4 、2g 柠檬酸和2g K_2SO_4 (pH6.0)。

用Leu-pNA(盐酸盐)作为底物分析140个孔的氨肽酶II活性。在96孔的微量平板中,将4 μl 上清液加入100 μl 溶在50mM磷酸钠缓冲液(pH7.5)中的1mg/ml Leu-pNA中。监测405nm的吸光度。

氨肽酶II活性最高的4个转化子(20、88、90和137)在含有25ml MY50培养基的125ml 摇瓶中于34 $^{\circ}\text{C}$ 生长4天。

在第2、3和4天通过将100 μl 10倍稀释的上清液与100 μl 溶在50mM磷酸钠缓冲液(pH7.5)中的2mg/ml Leu-pNA混合来分析样品的氨肽酶II活性。转化子20、90和137产量最高。为了纯化氨肽酶II,将转化子20和/或90在如上的摇瓶或含有合适碳、氮源的发酵培养基中培养。

实施例 11: 纯化曲霉产生的重组米曲霉 1568 氨肽酶 II

将由实施例 10 所述摇瓶培养液得到的上清液合并(100ml 中约有100mg 蛋白质),并稀释到3.7mS,调pH至7.0。然后将稀释过的样品上样到用20mM磷酸钠缓冲液(pH7.0)预平衡过的Q-Sepharose, Big Beads中。用于20mM磷酸钠缓冲液(pH7.0)中的0-0.4M NaCl 梯度洗脱氨肽酶II,随后用0.4M NaCl 洗涤。通过将100 μl 各组分与100 μl 每ml 含2mg Leu-pNA的50mM磷酸钠缓冲液(pH7.5)混合来检测氨肽酶II活性。检测结果表明氨肽酶II在梯度最后和0.4M NaCl 洗涤过程中被洗脱。SDS-PAGE 分析显示这种酶是均一的。

实施例 12: 构建用于镰孢的米曲霉 1568 氨肽酶 II 表达载体

设计了如下两个合成寡核苷酸引物来由质粒 pEJG18(大肠杆菌 DH5 α -EJG18)经 PCR 扩增米曲霉 1568 氨肽酶 II 基因编码序列,以便

用于在镰孢宿主中进行亚克隆和表达。

正向引物：5'-ATTTAAATcaccATGAGGTCGCTTTTGTGGGC-3' (SEQ ID NO:8)

反向引物：5'-GGGTTAATTA ACTATGCCTCGACTTGAGAATG-3' (SEQ ID NO:9)

粗体字代表编码序列。小写字母代表用于增强表达的 Kozak 共有序列。(Kozak, 1981, 核酸研究 (Nucleic Acid Research) 12: 857-872)。

为了方便将基因片段亚克隆到定名为 pDM181 (图 4) 的表达载体中, 在氨肽酶 II 基因的 5' 和 3' 端分别导入 *Swa*I 和 *Pac*I 限制酶位点。载体 pDM181 含有作为调控序列的尖镰孢胰蛋白酶样蛋白酶 (SP387) 启动子和终止子 (W096/00787)。质粒还含有 *bar* 基因作为真菌转化的选择标记 (de Block 等, 1987, EMBO 杂志 6: 2513-2518)。

在 PCR 反应体系 (50 μ l) 中使用了各 50 pmol 的上述引物, 该反应体系含有 70ng pEJG18; 1 \times Pwo 缓冲液; 5 μ l 10mM 的 dATP、dTTP、dGTP、dCTP 混合液; 2.5 单位 PwoI。扩增是 94 $^{\circ}$ C 2 分钟、55 $^{\circ}$ C 30 秒、72 $^{\circ}$ C 1 分钟情况下做 1 个循环; 94 $^{\circ}$ C 15 秒、55 $^{\circ}$ C 30 秒、72 $^{\circ}$ C 1 分钟情况下做 9 个循环; 94 $^{\circ}$ C 15 秒、55 $^{\circ}$ C 30 秒、72 $^{\circ}$ C 1 分钟, 每个循环中延长 20 秒, 这种情况下做 15 个循环; 最后一个循环 94 $^{\circ}$ C 15 秒、55 $^{\circ}$ C 30 秒、72 $^{\circ}$ C 7 分钟。然后将加热转入 4 $^{\circ}$ C 浸泡循环。经凝胶电泳和 QiaexII 纯化扩增好的 1500bp DNA 片段, 然后亚克隆至 pCRII TOPO TA 克隆载体 (Stratagene, San Diego, CA) 中。用 *Swa*I 和 *Pac*I 消化 pCRII 氨肽酶克隆 (使用制造商规定的条件)。通过凝胶电泳和 QiaexII 纯化片段。将切下的片段克隆到预先用 *Swa*I 和 *Pac*I 切割过的 pDM181 中, 得到表达质粒 pEJG28 (图 5), 在该载体中氨肽酶 II 基因的转录受到尖镰孢胰蛋白酶样蛋白酶启动子的调控。将质粒 pEJG28 转化到大肠杆菌 ABLE K 细胞 (Stratagene, San Diego, CA) 中。分离到一株含有 pEJG28 质粒的大肠杆菌转化子, 根据 Sambrook 等描述的步骤 (1989, 出处同前) 制备质粒 DNA。

实施例 13: 镰孢 CC1-3 的转化和转化子分析

镰孢 CC1-3 是镰孢菌株 A3/5 (ATCC20334) 之高度分枝的形态突变株 (Wiebe 等, 1992, 真菌学研究 (Mycological Research) 96:55-562; Wiebe 等, 1991, 真菌学研究 95:1284-1288; Wiebe 等, 1991, 真菌学研究 96:555-562)。将其在含有 Vogel's 盐 (Vogel, 1964, Am. Nature 98:435-446)、25mM NaNO₃ 和 1.5% 葡萄糖的液体培养基中, 于 28℃、150rpm 生长 4 天。经 4 层干酪皮, 最后通过一层 Miracloth 过滤纯化分生孢子。经离心使分生孢子悬液浓缩。50ml YPG 培养基 (含有 1% 酵母提取物、2% 细菌蛋白胨和 2% 葡萄糖) 接种大约 10⁸ 个分生孢子, 于 24℃、150rpm 培养 14 小时。用无菌 0.4mm 滤器截取得到的菌丝, 相继用无菌蒸馏水和 1.0 M MgSO₄ 洗涤。将菌丝重悬于 10ml NOVOZYM234™ 溶液 (在 1.0M MgSO₄ 中 2-10mg/ml), 于 34℃、80rpm 搅动消化 15-30 分钟。NOVOZYM234™ 得自 Novo Nordisk A/S, Bagsvaerd, Denmark。相继经 4 层干酪皮和一层 Miracloth 过滤从得到的原生质体悬液中除去未消化的菌丝物质。将 20ml 1M 山梨醇与原生质体溶液合并。混匀后, 经离心沉淀原生质体并相继经过在 20ml 1M 山梨醇、20ml STC (0.8M 山梨醇、0.05M Tris (pH8.0)、0.05M CaCl₂) 中重悬和离心来洗涤原生质体。洗过的原生质体以 5 × 10⁷ 个/ml 的浓度重悬于 4 份 STC 和 1 份 SPTC (0.8M 山梨醇、40%PEG4000、0.05M Tris (pH8.0)、0.05M CaCl₂) 中。

将 100ml 原生质体悬液加入含有 10μg pEJG28 的聚丙烯管 (17 × 100mm) 中, 混匀并在冰上放置 30 分钟。将 1ml SPTC 轻轻地混入原生质体悬液中, 继续于室温保温 20 分钟。将 12.5ml 含有 1 × Vogel's 盐、25mM NaNO₃、0.8M 蔗糖和 1% 低熔点琼脂糖 (Sigma Chemical Company, St. Louis, MO) 的融化好的溶液 (冷却到 40℃) 与原生质体混匀, 然后铺到 100mm 的空培养皿上。继续于室温保温 10 到 14 天。在室温保温 24 小时后, 取 12.5ml 加有 10mg BASTA™ (Hoechst Schering, Rodovre, Denmark) 的相同培养基覆盖在培养皿上。用酚: 氯仿: 异戊

醇(25:24:1)将 BASTA™ 抽提两次,在使用前再用氯仿:异戊醇(24:1)抽提一次。

两周后,可以看到两个明显的转化子(定名为#1和#2)。将每个转化子边缘的菌丝片段转移到含有 Vogel's /BASTA™ 培养基之 24 孔板的孔中。该培养基每升含有 25g 蔗糖、25g Nobel 琼脂、20ml 50× Vogel's 盐(Vogel, 1964, 出处同前)、25mM NaNO₃ 和 10g BASTA™。将平板密封在一个塑料袋中以保持潮湿,于室温培养大约一周。

实施例 14: 在镰孢中表达米曲霉 1568 氨肽酶 II 基因

将从实施例 13 所述 2 个镰孢 CC1-3 转化子得到的菌丝片段接种在 20ml M400Da 培养基中,并于 30℃、150rpm 保温 7 天。每升 M400Da 培养基含有 50g 麦芽糊精、2.0g MgSO₄-7H₂O、2.0g KH₂PO₄、4.0g 柠檬酸、8.0g 酵母提取物、2.0g 尿素和 0.5ml 微量金属溶液。用 5N NaOH 将培养基 pH 调至 6.0。每升该微量金属溶液含有 14.3g ZnSO₄-7H₂O、2.5g CuSO₄-5H₂O、0.5g NiCl₂-6H₂O、13.8g FeSO₄-7H₂O、8.5g MnSO₄-H₂O 和 3.0g 柠檬酸。未转化的宿主作为对照。在第 7 天收集 1ml 培养物上清液并进行保存和分析。通过将 10μl 上清液与 200μl 每 ml 含有 2mg N-亮氨酸对硝基苯胺的 50mM 磷酸钠缓冲液(pH7.5)之底物储液混匀,并在 4 分钟内监测 405nm 吸光度的变化。两个转化子都显示出对 Leu-pNA 的活性要高于未转化的对照。

将实施例 13 所述原代镰孢转化子#2 在 125ml 摇瓶(含有 25ml M400Da 培养基)中,于 30℃ 培养 5 天。用双层 Miracloth 将全部培养液过滤。回收滤过液,于-20℃ 冷冻。

实施例 15: 纯化镰孢产生的重组米曲霉 1568 氨肽酶 II

将 20ml 实施例 14 所述镰孢的 5 天培养液经 0.45μm 滤器过滤。然后将样品在 20mM 磷酸钠缓冲液(pH7.0)中稀释 8 倍。样品的导电率和 pH 分别为 3.2mS 和 7.10。将样品上样到用 20mM 磷酸钠缓冲液(pH7.0)预平衡过的含有 60ml Q-Sepharose, Big Beads 的 XK-26 (Pharmacia

Biotech AB, Uppsala, Sweden) 柱子中。将柱子洗至达到基线, 然后用于 20mM 磷酸钠缓冲液 (pH7.0) 中的 0-0.5M NaCl 线性梯度经 8.3 柱体积洗脱氨肽酶 II, 流速为 5ml/min。用 Leu-pNA 作为底物, 通过将 10 μ l 各组分与 90 μ l 50mM 磷酸钠缓冲液 (pH7.5) 和 100 μ l 于每 ml 50mM 磷酸钠缓冲液 (pH7.5) 中含 2mg Leu-pNA 的底物储液混合, 并监测 4 分钟内 405nm 处吸收值的变化, 从而检测组分。将所有对 Leu-pNA 有活性的组分汇合, 稀释并利用带有 PM10 膜的 Amicon 超滤装置将其浓缩。

将浓缩过的样品上样到用 20mM 磷酸钠缓冲液 (pH7.0) 预平衡过的 MonoQ16/10 柱子 (Pharmacia Biotech AB, Uppsala, Sweden) 中。用 0.15M NaCl 洗柱子。然后在 10 个柱体积内进行 0.15-0.5M NaCl 梯度洗脱, 流速为 2ml/min。用含有 1.7M (NH₄)₂SO₄ 的 50mM 磷酸钠缓冲液 (pH7.0) 平衡活性组分。

将样品上样到用含有 1.7M (NH₄)₂SO₄ 的 50mM 磷酸钠缓冲液 (pH7.0) 预平衡过的 Phenyl Superose5/5 柱子 (Pharmacia Biotech AB, Uppsala, Sweden) 中。用含有 1.7M (NH₄)₂SO₄ 的 50mM 磷酸钠缓冲液 (pH7.0) 洗柱子至达到基线。然后在 30 个柱体积内进行 1.7-0M (NH₄)₂SO₄ 梯度洗脱酶, 流速为 0.5ml/min。流出液与洗脱下的组分一样对 Leu-pNA 有活性。根据 SDS-PAGE 分析, 酶似乎有一系列不同的糖苷化形式。根据制造商的建议用内 糖苷酶 F/N 糖苷酶 F (BoehringerMannheim, Indianapolis, IN) 处理各种形式的酶时, 所有分析样品都显示出分子量 ~ 58kDa 的单一一条带。将不同糖苷化形式的酶汇合, 用 50mM 磷酸钠缓冲液 (pH7.5) 脱盐并进行生化鉴定。

实施例 16: 重组米曲霉 1568 氨肽酶 II 的鉴定

以下鉴定中使用实施例 11 中描述的纯化氨肽酶 II。

确定氨肽酶 II 对几种 N-某酰对硝基苯胺 (pNA) (包括 Leu-pNA、Gly-pNA、Ala-pNA 和 Pro-pNA (Sigma Chemical Co. St. Louis, MO 或 Bachem, Torrance, CA)) 的动力学参数。用 50mM 磷酸钾缓冲液

(pH7.0)将每 ml 二甲基亚砷含有 100mg 各种 pNA 的储液稀释至 0.0064 到 9.56mM。应注意，底物的溶解性不是总能足以使得浓度与 K_m 相称，这可能会造成比正常预期更高的误差。将 100 μ l 溶在 50mM 磷酸钾缓冲液 (pH7.0)中的酶溶液加入于微量板上孔中的 100 μ l 底物溶液时，氨肽酶 II 即开始与 pNA 反应，用 THERMOmax Microplate Reader (Molecular Devices Corp., Sunnyvale, CA)于 25 $^{\circ}$ C、405nm 监测。pNA 水解的起始速率实验得到表 1 所示结果：

表 1 氨肽酶在 pH7.0、25 $^{\circ}$ C 的动力学参数

底物	K_m (mM)	V_{max} (相对单位)
Leu-pNA	7	6400
Gly-pNA	0.3	1670
Ala-pNA	11	1200
Pro-pNA	2	120

这些结果表明氨肽酶 II 具有底物特异性，其中对丙氨酸的特异性比对甘氨酸的差得多。

用在 50mM Tris 缓冲液 (pH7.5)中的 Leu-pNA 做底物，在 405nm 监测水解反应来评价 1,10-二氮杂菲对氨肽酶 II 的抑制。制备在甲醇中的 200mM 1,10-二氮杂菲溶液。通过将 100 μ l 每 ml 含有 2mg Leu-pNA 的 50mM 磷酸钠缓冲液 (pH7.5)、10 μ l 1,10-二氮杂菲溶液和 100 μ l 在 50mM 磷酸钠缓冲液 (pH7.5)中稀释 5 倍的氨肽酶 II 混合来进行抑制反应。对照中用 10 μ l 20mM Tris 缓冲液 (pH7.6)代替 1,10-二氮杂菲溶液。

结果表明，1,10-二氮杂菲溶液能抑制氨肽酶 II，这暗示氨肽酶 II 是一种金属蛋白酶。在有 1,10-二氮杂菲溶液的情况下，Leu-pNA 的水解速率从 285mOD/分钟降低到 21mOD/分钟。

以下鉴定中使用实施例 15 中描述的纯化氨肽酶 II。

用 Ala-pNA 作为底物，在含有 0.125M 柠檬酸、0.125M 磷酸二氢钠和 0.125M 硼酸的普通缓冲液 (用 10M NaOH 将 pH 调至 4.5-11, pH 增

量幅度为 0.5) 中确定氨肽酶的最佳 pH。将 100mg Ala-pNA 溶解在 1ml DMSO 中, 并取 20 μ l Ala-pNA/ DMSO 加入 980 μ l 各种 pH 值的普通缓冲液中来制备 Ala-pNA 底物。于环境温度, 将 15 μ l 在 50mM 磷酸钠缓冲液 (pH7.5) 中的氨肽酶 II 加入 200 μ l 各种 pH 值的 2mg/ml Ala-pNA 中起始实验。监测 405nm 的吸光度 5 分钟内的变化。将 15 μ l 50mM 磷酸钠缓冲液 (pH7.5) 加入各种 pH 值的 200 μ l 2 mg/ml Ala-pNA 中确定底物的自水解, 以便作为对照。

下表 2 所示结果表明氨肽酶 II 在 pH4.91 到 10.91 的测量范围内对 Ala-pNA 底物有活性, 最佳活性在 pH ~ 9.5-10。在 pH11 及更小时没有观察到底物的自水解。

表 2

pH	Avg. 活性	相对活性
4.42	0 mOD/min	0
4.91	2	0.006
5.41	7.8	0.024
5.89	13.9	0.043
6.40	16.37	0.051
6.90	23.48	0.0727
7.27	41.24	0.128
7.59	69.15	0.214
8.03	145.6	0.45
8.62	245.99	0.761
9.25	306.97	0.95
9.68	323.15	1.0
10.51	270.8	0.838
10.95	197.86	0.612

用以下方案确定氨肽酶 II 对温度的稳定性: 将 1.7ml Eppendorf 管中的 480 μ l 50mM 磷酸钠缓冲液 (pH7.5) 于 37 $^{\circ}$ C、45 $^{\circ}$ C、55 $^{\circ}$ C、60

℃、65℃、70℃和75℃预温育30分钟。然后加入20μl纯化过的氨肽酶II，样品再保温20分钟。然后将样品置于冰上。所有温度的温育过程一结束，即用Leu-pNA作为底物检测样品的活性。

通过将100μl各个温度的温育体系与100μl含有2mg/ml Leu-pNA的50mM磷酸钠缓冲液(pH7.5)在环境温度中混合来进行检测。监测5分钟405nm吸光度。

表3所示结果表明于60℃、pH7.5温育20分钟后，氨肽酶II保持90%的活性。

表3

温度(℃)	相对37℃的活性百分比
37	100
45	101
55	99
60	90
65	73.7
70	64.6
75	46

用以下方案确定氨肽酶II对其各个底物的动力学参数。使用纯化的氨肽酶II($A_{280}=0.581$)。将每个底物溶解在DMSO中至浓度为100mg/ml，然后在50mM磷酸钠缓冲液(pH7.5)中稀释50倍至浓度为2mg/ml。底物包括Leu-pNA、Glu-pNA(Bachem, Torrance, CA)和Ala-pNA。在96孔微量板中，将10μl纯化的氨肽酶II与以下所示各种底物一起保温，但Glu-pNA用50μl纯化的氨肽酶II，监测4分钟405nm的吸光度：

1. 200μl 2mg/ml 底物+0μl 50mM磷酸钠缓冲液(pH7.5)
2. 100μl 2mg/ml 底物+100μl 50mM磷酸钠缓冲液(pH7.5)
3. 50μl 2mg/ml 底物+150μl 50mM磷酸钠缓冲液(pH7.5)
4. 25μl 2mg/ml 底物+175μl 50mM磷酸钠缓冲液(pH7.5)

做 Lineweaver-Burke 曲线确定每种底物的 K_m 和 K_{cat} , 不同糖苷化形式都使用平均分子量 97kDa.

对 Leu-pNA, 确定出 K_m 和 K_{cat} 分别为 5.78mM 和 230.9min^{-1} .

对 Glu-pNA, 确定出 K_m 和 K_{cat} 分别为 1.17mM 和 8.217min^{-1} .

对 Ala-pNA, 确定出 K_m 和 K_{cat} 分别为 1.49mM 和 34.638min^{-1} .

实施例 17: 用米曲霉 1568 氨肽酶 II 制备蛋白质水解产物

用大豆、小麦谷蛋白和酪蛋白作为底物, 根据以下步骤检测实施例 11 所述纯化氨肽酶 II 的水解能力:

水解度 (DH) 检测是用大豆粉片、小麦谷蛋白粉片和酪蛋白酸钠 (浓度为 2%、pH 调至 7, 如果需要, 在水解期间不调 pH) 做的 10ml 小规模水解, 于 50°C 进行 18 小时。在 85°C 水浴 3 分钟使水解产物失活。使用的酶是 FLAVOURZYME™ 和氨肽酶 II。酶用量情况如下: 对于大豆, 加入 2 LAPU 和 5 LAPU 氨肽酶 II (重组), 与 3 LAPU FLAVOURZYME™ 比较。对谷蛋白, 加入 2 LAPU 和 5 LAPU 氨肽酶 II (重组), 与 3 LAPU FLAVOURZYME™ 比较。对酪蛋白, 加入 1 LAPU 和 2 LAPU 氨肽酶 II (重组), 与 3 LAPU FLAVOURZYME™ 比较。一个 LAPU (亮氨酸氨肽酶单位) 是在如下条件每分钟分解 1 微摩尔 L-Leu-pNA 所需的酶量: 在 0.1M Tris 缓冲液 (pH8.0) 中的 26mM L-亮氨酸-pNA 于 40°C 反应 10 分钟。水解过程中, 释放出 pNA 使溶液变黄, 监测 405nm。

水解度 (DH) 如 Adler-Nissen (1986, 食物蛋白质的酶法水解, Elsevier Applied Science Publishers) 所定义和描述, 是根据以下步骤, 通过上清液与 OPA (邻苯二醛, Sigma Chemical Co., St. Louis, MO) 的反应确定的。将水解产物用蒸馏水稀释 100 倍。然后取 $120\mu\text{l}$ 转移到 $900\mu\text{l}$ OPA 试剂中。制备 OPA 试剂的过程, 是将 160 mg OPA 溶解在 4ml 乙醇中并转移到 200ml 体积的烧瓶中, 烧瓶含有 7.62g 四硼酸二钠十水合物、200mg 十二烷基磺酸钠和 176mg 二硫苏糖醇, 用水补至 200ml。然后将溶液摇匀, 恰好 2 分钟后, 测量 340nm 吸光度, 并在减去空白值 (水与 OPA 试剂反应) 后与 0.95mM L-丝氨酸 (蒸馏

水)溶液的吸光度进行比较。为了确定 DH 的实际数值,用 Adler-Nissen 建议的用于三硝基苯磺酸方法 (Adler-Nissen, 1979, 农业和食品化学 (Agricultural and Food Chemistry) 17:1256) 的指数修正在水解产物中测量的丝氨酸当量,在所述方法中产生与文中描述的 OPA 方法相同的效果。根据水解混合体系中的蛋白质总量而不是溶解蛋白质来计算 DH。

在一个微量培养板孔中将 25 μ l 适当稀释的上清液与 200 μ l OPA 试剂混匀,并于 25 $^{\circ}$ C 恰好反应 2 分钟。在微量培养板读数器中测量 340nm 的吸光度,减去空白值(水与 OPA 试剂反应)后与 95mM L-丝氨酸标准溶液的吸光度进行比较。为了确定 DH 的实际数值,用 Adler-Nissen 建议的用于三硝基苯磺酸方法 (Adler-Nissen, 1979, 农业和食品化学 17:1256) 的指数修正在上清中测量的丝氨酸当量,在所述方法中产生与文中描述的 OPA 方法相同的反应。根据水解混合体系中的蛋白质总量而不是溶解蛋白质来计算 DH。

对于大豆,向 3 LAPU FLAVOURZYMETM 加入 2LAPU 和 5 LAPU 氨肽酶 II 分别使绝对 DH 相对于只用 3 LAPU FLAVOURZYMETM 的样品升高至少 8%和 10%。

对于谷蛋白,向 3 LAPU FLAVOURZYMETM 加入 2LAPU 和 5 LAPU 氨肽酶 II 分别使绝对 DH 升高 6%和 9%。

对于白明胶,向 3 LAPU FLAVOURZYMETM 加入 2LAPU 和 5 LAPU 氨肽酶 II 分别使绝对 DH 升高 4.9%和 5.3%。

对于酪蛋白,向 3 LAPU FLAVOURZYMETM 加入 1LAPU 和 2 LAPU 氨肽酶 II 分别使绝对 DH 相对于只用 3 LAPU FLAVOURZYMETM 的样品升高 7%和 9%。

实施例 18: 用米曲霉氨肽酶 II 水解大豆蛋白质

以 10ml 规模(微量水解)水解大豆蛋白质,起始 pH 为 7.0,蛋白质浓度为 2%。水解时间和温度分别为 18 小时和 50 $^{\circ}$ C。将酶于 85 $^{\circ}$ C 5 分钟失活,离心水解产物。用 OPA 方法分析上清液的 DH。通过上清液

与 OPA (邻苯二醛, Sigma Chemical Co., St. Louis, MO) 的反应来确定 Adler-Nissen (1986, 食物蛋白质的酶法水解, Elsevier Applied Science Publishers) 所定义和描述的水解度 (DH)。制备 OPA 试剂的过程是将 160 mg OPA 溶解在 4ml 乙醇中并转移到 200ml 体积的烧瓶中, 烧瓶含有 7.62g 四硼酸二钠十水合物、200mg 十二烷基磺酸钠和 176mg 二硫苏糖醇, 用水补至 200ml。用 PicoTag HPLC 方法 (Waters Associates, Milford, MA) 依照制造商的说明分析所选样品的游离氨基酸含量。

每个含有 200mg 大豆蛋白质的水解摇瓶中酶用量如下表 4 所示。氨肽酶 II 是按照实施例 10 所述在米曲霉中重组产生并纯化的。该氨肽酶 II 溶液 A_{280} 为 8.1, 由氨基酸测定的结果估计蛋白质含量为 5mg/ml。

DH 分析的结果列在表 4 中。由总蛋白质浓度 2% 而非溶解蛋白质含量来计算 DH。

表 4 水解产物的 DH 结果

	FLAVOURZYME™ 1000L %	氨肽酶 II g 酶蛋白/kg 大豆蛋白	DH %
1	1.5	0	45.1
2	1.5	0.03	50.9
3	1.5	0.06	51.0
4	1.5	0.12	51.3
5	1.5	0.25	55.7
6	1.5	0.50	58.0
7	2.0	0	51.9
8	6.0	0	62.8
9	6.0	0.03	62.9
10	6.0	0.06	62.9
11	6.0	0.12	63.6
12	6.0	0.25	68.5
13	6.0	0.50	67.8
14	7.0	0	63.2

*这项计算使用的氨肽酶 II 浓度为 5mg/ml

表 5 显示添加最大量的氨肽酶 II (0.5g/kg 大豆蛋白) 时, 与基量 FLAVOURZYME™ 相比每种氨基酸的相对增量%

表 5 添加氨肽酶 II 导致的游离氨基酸的相对增量%

氨基酸	1.5%FLAVOURZYME™ +氨肽酶 II	6%FLAVOURZYME™ +氨肽酶 II
天冬氨酸	123.4	26.0
谷氨酸	54.2	24.5
天冬酰胺	115.1	-7.0
丝氨酸	123.8	0.0
谷氨酰胺	91.7	19.8
甘氨酸	145.7	22.4
组氨酸	31.9	2.8
精氨酸	24.9	6.1
苏氨酸	40.4	8.8
丙氨酸	77.4	18.5
脯氨酸	87.5	59.2
酪氨酸	51.3	10.5
缬氨酸	41.4	12.7
甲硫氨酸	36.8	9.7
半胱氨酸	74.4	23.2
异亮氨酸	24.6	13.0
亮氨酸	22.4	7.5
苯丙氨酸	22.8	8.0
赖氨酸	49.0	10.3
总量	49.1	10.6
DH	28.7	8.0

结果表明向低剂量 FLAVOURZYME™ (1.5%) 中加入氨肽酶 II 时, 甘氨酸显出最高的相对增量, 然后是丝氨酸、天冬氨酸、天冬酰胺、脯氨酸、半胱氨酸和丙氨酸。当向高剂量 FLAVOURZYME™ (6%) 加入氨肽酶 II 时, 脯氨酸显出最高的相对增量, 然后是天冬氨酸、谷氨酸、半胱氨酸、甘氨酸和谷氨酰胺。

实施例 19: 脱酰胺作用使蛋白质溶解性及谷氨酸释放增加

小麦谷蛋白得自 Cargill(JOB5141), 脱酰胺小麦谷蛋白(DWG)得自 StaPro Consultancy B.V., Lemdijk32, 9422TH Smilde, NL。将 11g 谷蛋白和 89g 水混合在一起得到 8% 的蛋白质悬液。用 NaOH 将 pH 调到 6.5。将谷氨酸/天冬氨酸特异的蛋白酶(SP446)(可根据 W091/13554 的描述得到), 或者赖氨酸/精氨酸特异的蛋白酶(SP387)(可根据 W089/06270 的描述得到)加入悬液。蛋白质的用量对 SP446 是 0.01AU/g 蛋白质, 对 SP387 为 0.006AU/g 蛋白质。以 20LAPU/g 蛋白质的剂量将 FLAVOURZYME™ (一种可得自 Novo Nordisk A/S, Bagsvaerd, Denmark 的非特异性蛋白酶制剂, 含有内肽酶和外肽酶活性, 是通过米曲霉发酵产生的) 加入某些水解产物中。

水解于 50℃ 进行 18 小时, 不需进一步调 pH。于 85℃ 加热 15 分钟使酶失活。将 pH 调到 5, 并将水解产物离心。确定上清液中的蛋白质和游离谷氨酸含量。

经 Kjeldahl 分析确定蛋白质含量, 所用 Kjeldahl 因子为 6.25。

用谷氨酸定量试剂盒(Boehringer Mannheim, Indianapolis, IN), 依照制造商的说明确定游离谷氨酸的含量。将方法略做改进以适用于微滴培养板。

就小麦谷蛋白(WG)和脱酰胺小麦谷蛋白(DWG)进行比较时, 表 6 所示结果表明脱酰胺化使谷蛋白对特异蛋白酶的易感性增加, 因此有更多的蛋白质溶解。将 FLAVOURZYME™ 与一种特异蛋白酶同时加入时, 由于脱酰胺化而使谷氨酸的释放量加倍。

表 6

水解产物	蛋白质溶解度%		谷氨酸含量 mg/ml	
	WG	DWG	WG	DWG
SP446	18	54	0	0
SP387	35	44	0	0
SP446+FLAVOURZYME™	34	87	1000	2000

实施例 20: 酶法脱酰胺以及谷氨酸释放

使环状芽孢杆菌 ATCC21590 于 30℃、270rpm 混合在摇瓶 (400ml) 中生长 20 小时来生产肽谷氨酰胺酶 II, 摇瓶中含有 200ml 由 1%蛋白胨、0.5%乳糖、0.025%MgSO₄-7H₂O、0.005%FeSO₄-7H₂O、0.025%KH₂PO₄ 和 17% Na₂HPO₄-12H₂O (将 pH 调至 7.2) 组成的培养基。在 4000rpm 离心收获 1 升摇瓶中的细胞。然后将细胞冷冻。

得自环状芽孢杆菌的肽谷氨酰胺 II 的纯化过程在室温下进行。将冷冻的环状芽孢杆菌细胞融化, 悬浮在裂解缓冲液 (50mM Tris/HCl; 25%(w/v)蔗糖; 1mM EDTA, pH8.0) 中直到获得均匀的悬液, 每升裂解缓冲液有 100g 湿细胞。将溶菌酶 (10mg/ml) 和 DNase I (Sigma DN-25, 10mg/ml) 溶解在裂解缓冲液中。然后在每升细胞悬液中加入 100ml 溶菌酶溶液、10ml 1.0M MgCl₂ 和 1ml DNase I 溶液。将酶放置反应 1 小时。

将悬液经蔡氏深度滤板过滤, 滤过液在 Sephadex G25 柱子 (Pharmacia Biotech AB, Uppsala, Sweden) 上转移到 10mM KH₂PO₄/NaOH (pH8.0, 缓冲液 A) 中。将酶溶液上样到在缓冲液 A 中平衡过的 SOURCE Q 柱子 (Pharmacia Biotech AB, Uppsala, Sweden) 上, 并利用溶于缓冲液 A 的线性 NaCl 梯度 (0-500mM) 洗脱。按以下描述分析从柱子上流下的组分的肽谷氨酰胺酶 II 活性, 并将有活性的组分汇合。汇合组分在 280nm 的吸光度为 1.78, 据此估计蛋白质的纯度为 1.8mg/ml。

经 SDS-PAGE 凝胶判断肽谷氨酰胺酶 II 汇合液中的蛋白质纯度接近 25%。因此, 制品中含有大约 0.5mg/ml 纯肽谷氨酰胺酶 II。

用一个 Boehringer-Mannheim 定氮试剂盒 (目录号 1112732)。通

过测量 N-叔丁氧羰基-Gln-Pro (N-t-BOC-Gln-Pro; Sigma No. B-4403) 中的 γ -羧基酰胺在水解过程中形成的氨来确定肽谷氨酰胺酶的活性。在这个试剂盒中, 是通过确定谷氨酸脱氢酶消耗 NADH 的量来测量氨含量的, 还要使用不加 N-t-BOC-Gln-Pro 的空白以便消除其他消耗 NADH 的酶类的影响。

将总量 200mg 的小麦谷蛋白加入 9ml 沸水中, 冷却后, 将 pH 调至 7.0。然后加入上述肽谷氨酰胺酶 II 制剂 (PEP)。以 0.04AU/g 蛋白质的用量加入实施例 19 中描述的谷氨酸/天冬氨酸特异性蛋白酶, 以 20LAPU/g 蛋白质的用量加入实施例 19 中描述的 FLAVOURZYME™。

水解反应于 50℃ 进行 18 小时 (不调 pH)。同时做不加肽谷氨酰胺酶的对照。离心水解产物, 按照实施例 19 的描述测量谷氨酸含量。按照实施例 18 的描述确定 DH。

下表 7 所示结果表明, 用肽谷氨酰胺酶制剂进行的水解使得 DH 和谷氨酸的释放量都增加。

表 7

水解反应	DH%	谷氨酸 mg/l
无 PEP	40	131
加 PEP	43	171

生物材料的保藏

按照布达佩斯条约将以下生物材料保藏在农业研究机构保藏中心, 北方地区研究中心, 1815 University Street, Peoria, Illinois, 61604, 所给编号如下:

保藏物	编号	保藏日
大肠杆菌 DH5 α pEJG18	NRRL B-21677	1997 年 4 月 4 日

序列表

(1) 一般资料:

(i) 申请人:

(A) 姓名: Novo Nordisk Biotech, Inc.

(B) 街道: 1445 Drew Avenue

(C) 城市: Davis, Californis

(E) 国家: 美国

(F) 邮政编码 (ZIP): 95616-4880

(G) 电话: (530) 757-8100

(H) 传真: (530) 758-0317

(ii) 发明题目: 具有氨肽酶活性的多肽及其编码核酸

(iii) 序列数: 9

(iv) 通讯地址:

(A) 收件人: Novo Nordisk of North America, Inc.

(B) 街道: 405 Lexington Avenue

(C) 城市: New York

(D) 国家: 美国

(E) ZIP: 10174

(v) 计算机可读形式:

(A) 介质类型: 软盘

(B) 计算机: IBM PC 可兼容机

(C) 操作系统: PC-DOS/MS-DOS

(D) 软件: PatentIn Release #1.0, Version #1.30

(vi) 目前的申请资料:

(A) 申请号: 待定

(B) 申请日: 1998年5月13日

(C) 分类号:

(vii) 在先申请资料:

(A) 申请号:

(B) 申请日:

(viii) 代理人/代理资料:

(A) 姓名: Gregg, Valeta A

(B) 登记号: 35,127

(C) 资料/文档号: 5253,204-W0

(ix) 通讯资料:

(A) 电话: 212-867-0123

(B) 传真: 212-878-9655

(2) SEQ ID NO: 1 的资料:

(i) 序列特征:

(A) 长度: 1491 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(ix) 特征:

(xi) 序列描述: SEQ ID NO: 1:

ATGAGGTCGC	TTTTGTGGGC	TCGTTGCTT	TCGGCGTGT	TGGCTGGGAG	GGCGCTTGT	60
TCGCCGGATG	AGTTCCCCGA	GGATATTCAG	TTGGAAGATC	TGCTGGAAGG	ATCCCAACAG	120
CTTGAGGACT	TCGCCTATGC	CTACCCCGAG	CGCAATCGCG	TCTTTGGTGG	TAAAGCCCAC	180
GACGACACGG	TTAACTATCT	CTACGAGGAG	CTGAAGAAGA	CTGGCTACTA	TGATGTCTAC	240
AAGCAGCCTC	AGGTGCACCT	GTGGAGCAAT	GCCGACCAGA	CGCTCAAGGT	GGCGGATGAG	300
GAAATCGAGG	CGAAGACCAT	GACCTACAGT	CCCAGCGTGC	AGGTCACCGC	CGATGTAGCC	360

GTCGTCAAGA	ACCTGGGATG	CAGCGAGGCG	GATTACCCAT	CCGATGTCGA	GGGCAAGGTC	420
GCCCTGATCA	AGCGTGGAGA	ATGCCCGTTC	GGCGACAAGT	CGGTTCTCGC	TGCCAAAGCC	480
AAGGCCGCGG	CTTCGATTGT	CTATAACAAT	GTGGCCGGAT	CCATGGCGGG	CACCCTTGGC	540
GCGGCGCAGA	GTGATAAGGG	ACCGTATTCG	GCCATTGTCTG	GTATCAGCTT	GGAGGATGGC	600
CAGAAGCTGA	TCAAGCTTGC	TGAGGCTGGA	TCGGTATCTG	TGGATCTGTG	GGTGGATAGT	660
AAGCAGGAGA	ACCGTACGAC	GTATAACGTT	GTCGCGCAGA	CGAAGGGCGG	CGATCCGAAC	720
AACGTCGTCG	CGCTGGGTGG	CCACACGGAC	TCAGTCGAGG	CGGGCCCTGG	TATCAACGAC	780
GATGGCTCGG	GCATTATTAG	CAACTTGGTC	ATTGCCAAAG	CGCTCACGCA	GTACTCCGTC	840
AAGAATGCCG	TGCGCTTCCT	CTTCTGGACA	GCAGAGGAGT	TCGGTCTGCT	GGGCAGCAAC	900
TACTACGCTC	CCCATCTGAA	TGCCACCGAG	CTGAACAAGA	TCCGACTGTA	CCTGAACCTC	960
GACATGATCG	CCTCACCTAA	CTACGCCCTC	ATGATCTATG	ACGGTGATGG	ATCGGCGTTC	1020
AACCAGAGCG	GACCGGCCCG	TTCCGCCCAG	ATCGAGAAAC	TGTTCGAGGA	CTACTACGAC	1080
TCCATCGACC	TGCCTCATAT	CCCCACCCAG	TTTGACGGAC	GTTCCGACTA	CGAGGCCTTT	1140
ATCCTGAACG	GCATTCCGTC	CGGTGGACTC	TTCACGGGCG	CCGAGGGCAT	CATGTCCGAA	1200
GAGAACGCAA	GCCGCTGGGG	AGGTCAAGCC	GGCGTGGCCT	ACGACGCCAA	CTACCACGCC	1260
GCGGGAGACA	ACATGACCAA	CCTCAACCAT	GAAGCCTTCC	TGATCAACTC	CAAAGCCACC	1320
GCCTTCGCCG	TCGCCACCTA	CGCCAACGAC	CTCTCCTCGA	TCCCCAAACG	GAATACCACA	1380
TCCTCCTTGC	ACCGACGAGC	CCGCACCATG	CGACCATTTC	GCAAGAGAGC	TCCGAAGACA	1440
CACGCTCAGC	TATCAGGATC	CGGATGCTGG	CATTCTCAAG	TCGAGGCATA	G	1491

(2) SEQ ID NO:2 的资料:

(i) 序列特征:

(A) 长度: 496 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(xi) 序列描述: SEQ ID NO:2:

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

Gly	Thr	Leu	Gly	Ala	Ala	Gln	Ser	Asp	Lys	Gly	Pro	Tyr	Ser	Ala	Ile
			180					185					190		
Val	Gly	Ile	Ser	Leu	Glu	Asp	Gly	Gln	Lys	Leu	Ile	Lys	Leu	Ala	Glu
		195					200					205			
Ala	Gly	Ser	Val	Ser	Val	Asp	Leu	Trp	Val	Asp	Ser	Lys	Gln	Glu	Asn
		210				215					220				
Arg	Thr	Thr	Tyr	Asn	Val	Val	Ala	Gln	Thr	Lys	Gly	Gly	Asp	Pro	Asn
225				230						235				240	
Asn	Val	Val	Ala	Leu	Gly	Gly	His	Thr	Asp	Ser	Val	Glu	Ala	Gly	Pro
			245						250					255	
Gly	Ile	Asn	Asp	Asp	Gly	Ser	Gly	Ile	Ile	Ser	Asn	Leu	Val	Ile	Ala
		260						265				270			
Lys	Ala	Leu	Thr	Gln	Tyr	Ser	Val	Lys	Asn	Ala	Val	Arg	Phe	Leu	Phe
		275					280					285			
Trp	Thr	Ala	Glu	Glu	Phe	Gly	Leu	Leu	Gly	Ser	Asn	Tyr	Tyr	Val	Ser
		290				295					300				
His	Leu	Asn	Ala	Thr	Glu	Leu	Asn	Lys	Ile	Arg	Leu	Tyr	Leu	Asn	Phe
305					310					315					320
Asp	Met	Ile	Ala	Ser	Pro	Asn	Tyr	Ala	Leu	Met	Ile	Tyr	Asp	Gly	Asp
				325					330					335	
Gly	Ser	Ala	Phe	Asn	Gln	Ser	Gly	Pro	Ala	Gly	Ser	Ala	Gln	Ile	Glu
			340					345					350		
Lys	Leu	Phe	Glu	Asp	Tyr	Tyr	Asp	Ser	Ile	Asp	Leu	Pro	His	Ile	Pro
		355					360					365			
Thr	Gln	Phe	Asp	Gly	Arg	Ser	Asp	Tyr	Glu	Ala	Phe	Ile	Leu	Asn	Gly
		370				375					380				
Ile	Pro	Ser	Gly	Gly	Leu	Phe	Thr	Gly	Ala	Glu	Gly	Ile	Met	Ser	Glu
385					390					395					400
Glu	Asn	Ala	Ser	Arg	Trp	Gly	Gly	Gln	Ala	Gly	Val	Ala	Tyr	Asp	Ala
				405					410					415	
Asn	Tyr	His	Ala	Ala	Gly	Asp	Asn	Met	Thr	Asn	Leu	Asn	His	Glu	Ala
			420					425					430		
Phe	Leu	Ile	Asn	Ser	Lys	Ala	Thr	Ala	Phe	Ala	Val	Ala	Thr	Tyr	Ala
		435					440					445			
Asn	Asp	Leu	Ser	Ser	Ile	Pro	Lys	Arg	Asn	Thr	Thr	Ser	Ser	Leu	His
		450				455					460				
Arg	Arg	Ala	Arg	Thr	Met	Arg	Pro	Phe	Gly	Lys	Arg	Ala	Pro	Lys	Thr
465					470					475					480
His	Ala	His	Val	Ser	Gly	Ser	Gly	Cys	Trp	His	Ser	Gln	Val	Glu	Ala
				485					490					495	

(2) SEQ ID NO:3 的资料:

(i) 序列特征:

- (A) 长度: 20 个氨基酸
- (B) 类型: 氨基酸
- (C) 链型: 单链
- (D) 拓扑结构: 线性

(ii) 分子类型: 无

(xi) 序列描述: SEQ ID NO:3:

```

Cys Cys Ile Gly Ala Tyr Gly Ala Arg Thr Thr Tyr Cys Cys Ile Gly
 1                               5                               10                               15
Ala Arg Gly Ala
                20

```

(2) SEQ ID NO:4 的资料:

(i) 序列特征:

(A) 长度: 36 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(xi) 序列描述: SEQ ID NO:4:

```

Arg Thr Thr Tyr Thr Thr Ile Ala Cys Ile Ala Cys Ile Gly Cys Ile
 1                               5                               10                               15
Ala Cys Arg Thr Cys Ile Gly Cys Ile Gly Thr Ile Ala Cys Tyr Thr
                20                               25                               30
Cys Ile Ala Cys
                35

```

(2) SEQ ID NO:5 的资料:

(i) 序列特征:

(A) 长度: 537 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(xi) 序列描述: SEQ ID NO:5:

```

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

```

```

145          150          155          160
Ala Asn Thr Thr Ala Phe Ala Leu Ser Pro Pro Val Asp Gly Phe Val
          165          170          175
Gly Lys Leu Val Glu Ile Pro Asn Leu Gly Cys Glu Glu Lys Asp Tyr
          180          185          190
Ala Ser Val Val Pro Pro Arg His Asn Glu Lys Gln Ile Ala Leu Ile
          195          200          205
Glu Arg Gly Lys Cys Pro Phe Gly Asp Lys Ser Asn Leu Ala Gly Lys
          210          215          220
Phe Gly Phe Thr Ala Val Val Ile Tyr Asp Asn Glu Pro Lys Ser Lys
225          230          235          240
Glu Gly Leu His Gly Thr Leu Gly Glu Pro Thr Lys His Thr Val Ala
          245          250          255
Thr Val Gly Val Pro Tyr Lys Val Gly Lys Lys Leu Ile Ala Asn Ile
          260          265          270
Ala Leu Asn Ile Asp Tyr Ser Leu Tyr Phe Ala Met Asp Ser Tyr Val
          275          280          285
Glu Phe Ile Lys Thr Gln Asn Ile Ile Ala Asp Thr Lys His Gly Asp
          290          295          300
Pro Asp Asn Ile Val Ala Leu Gly Ala His Ser Asp Ser Val Glu Glu
305          310          315          320
Gly Pro Gly Ile Asn Asp Asp Gly Ser Gly Thr Ile Ser Leu Ile Asn
          325          330          335
Val Ala Lys Gln Leu Thr His Phe Lys Ile Asn Asn Lys Val Arg Phe
          340          345          350
Ala Trp Trp Ala Ala Glu Glu Glu Gly Leu Leu Gly Ser Asn Phe Tyr
          355          360          365
Ala Tyr Asn Leu Thr Lys Glu Glu Asn Ser Lys Ile Arg Val Phe Met
          370          375          380
Asp Tyr Asp Met Met Ala Ser Pro Asn Tyr Glu Tyr Glu Ile Tyr Asp
385          390          395          400
Ala Asn Asn Lys Glu Asn Pro Lys Gly Ser Glu Glu Leu Lys Asn Leu
          405          410          415
Tyr Val Asp Tyr Tyr Lys Ala His His Leu Asn Tyr Thr Leu Val Pro
          420          425          430
Phe Asp Gly Arg Ser Asp Tyr Val Gly Phe Ile Asn Asn Gly Ile Pro
          435          440          445
Ala Gly Gly Ile Ala Thr Gly Ala Glu Lys Asn Asn Val Asn Asn Gly
          450          455          460
Lys Val Leu Asp Arg Cys Tyr His Gln Leu Cys Asp Asp Val Ser Asn
465          470          475          480
Leu Ser Trp Asp Ala Phe Ile Thr Asn Thr Lys Leu Ile Ala His Ser
          485          490          495
Val Ala Thr Tyr Ala Asp Ser Phe Glu Gly Phe Pro Lys Arg Glu Thr
          500          505          510
Gln Lys His Lys Glu Val Asp Ile Leu Asn Ala Gln Gln Pro Gln Phe
          515          520          525
Lys Tyr Arg Ala Asp Phe Leu Ile Ile
          530          535

```

(2) SEQ ID NO:6 的资料:

(i) 序列特征:

(A) 长度: 23 个碱基对

- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线性

(xi) 序列描述: SEQ ID NO:6:

ATGATGAGGT CGCTTTTGTG GGC

23

(2) SEQ ID NO:7 的资料:

(i) 序列特征:

- (A) 长度: 23 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线性

(xi) 序列描述: SEQ ID NO:7:

GGGATGCATC TATGCCTCGA CTT

23

(2) SEQ ID NO:8 的资料:

(i) 序列特征:

- (A) 长度: 32 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线性

(xi) 序列描述: SEQ ID NO:8:

ATTTAAATCA CCATGAGGTC GCTTTTGTGG GC

32

(2) SEQ ID NO:9 的资料:

(i) 序列特征:

- (A) 长度: 32 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO:9:

GGGTTAATTA ACTATGCCTC GACTTGAGAA TG

32

ATGAGGTGGCTTTCGGGGCTTGGTTGCTTTCGGGGCGTGTGGCTGGAGGGCGGCTTGTTCGCCCGGATGAGTTCCCGGAGGATAATTCAG 90
M R S L L W A S L L S G V L A G R A L V S P D E F P E D I Q
 TTGGMAGATCTGCTGGMAGGATCCCAACAGCTTGAGGACTTCGCCATATGCCATACCCGAGGCGMTCGGGCTTTTGGTGGTAAAGCCAC 180
L E D L L E G S Q Q L E D F A Y A Y P E R N R V F G G K A H
 GACGACAGGTTAAGTATCTACGAGGAGCTGMMAGACTGGCTACTATATGATGCTACAMGACGCCTCAGGTGCAGGTTGTGGAGCMT 270
D D T V N Y L Y E E L K K T G Y Y D V Y K Q P Q V H L W S H
 GCGGACAGAGCCCAAGGTGGCGGATGAGGAAATCGAGGGMAGACCAIAGCCATACAGTCCAGGTCGAGGTCACCGCGGATGATGCC 360
A D D T L K V G D E E I E A K T M T Y S P S V E V T A D V A
 GTGGTCAAGAACTTGGGATGACAGGAGCGGGATACCCATCCGATCGGATGTCGAGGGGCMGCTCGCCCIGATCAGGGTGGAGMTCGCCGCTC 450
V V K N L G C S E A D Y P S D V E G K V A L I K R G F C P F
 GGGCACAGTCCGGTCTCGCTGCCMAGCCMAGCCCGGGCTTCGATTGCTATMCMATGTGGCCGGATCCATGGCGGGCACCCCTGGC 540
G D K S V I A A K A A A S I V Y N N V A G S M A G T L G
 GCGGGCCAGAGTGGATMAGGACCGTATTCGGCCATTTGTCGGTATCAGCTTGCAGGATGGCCAGMAGCTGATCAGCTTGCATGAGGCTGGA 630
A A Q S D K G P Y S A I V G I S L E D G Q K I I K L A I A G
 TCGGATCTGTCGGAICGTGGTGGATAGTAGCAGGAGMCCGTAACGAGGTATMCGTTGTCGGCCAGACGMAAGGCGCGGATCCGMAC 720
S V S V D L H V D S K Q E N R T T Y N V V A Q T K G G D P M
 AACGTGCTCCGGCTCGGTCGCCACACCGGACTCAGTCGAGGCGGGCCCTGGTATCAGCAGGATGGCTCGGGCATTAATAGCMACTTGGTC 810
H V V A L G G H T D S V E A G P G I N D D G S G I I S N L V

图 1A

ATTGCCAAGCGCTCACGCCAGTACTCCGTCAMGAMTGGCGTGGCTTCCCTCTTCTGGACAGCAGAGGAGITCGGTCGTGGTGGCAGCAAC 900
 I A K A L T Q Y S V K N A V R F L F W T A E E F G L L G S N
 TACTAGGTCICCCATCTGMA TGCCACCGAGCTGACAGATCCGACTGTACCTGMACTTCGACATGATCGCCTCACCTAACTACGCCCTC 990
 Y Y V S H L N A T E L N K I R L Y L N F D M I A S P N Y A I
 ATGATCTATGACGGTGATGGATCGGGGTTCAACAGAGGGACCGGCCGTTCCGCCAGATCGAGAACGTGTCGAGGACTACTACGAC 1080
 M J Y D G D G S A F N Q S G P A G S A Q I E K L F E D Y Y D
 TCCATCGACCTGCCTCATAICCCACCCAGTTTGACGGACGTTCCGACTACGAGGCTTATCCTGMA CGGCATTCCGTCGGGTGGACIC 1170
 S I D L P H I P T Q F D G R S D Y E A F I L N G I P S G G L
 TTCACGGGGCGGGGCATCATGTCGMAAGAGAACGCCCGCTGGGAGGTCAMGCCGGGTGGCCTACGACGCCAACACTACACGGCC 1260
 F T G A E G I M S E E N A S R W G G Q A G V A Y D A N Y H A
 GCGGAGACACATGACCACTCACCATGAMGCCTTCTGATCMACTCCAMGCCACC GCCCTCCCGTGGCCACCCTACGCCAACGAC 1350
 A G D N M T N L N H E A F L I N S K A T A F A V A T Y A N D
 CTCCTCTGATCCCCAMCGGATACACATCCTCTTGCACCGAGCCCGCACCATGGACCATTCGGCAAGAGAGCTCCGGAGACA 1440
 L S S I P K R N I T S S L H R R A R I M R P F G K R A P K I
 CAGGCTCACGTATCAGGATCCGGATGCTGGCATTCMAGTCGAGGCATAG 1491
 H A H V S G S G C W H S Q V E A

图 1B

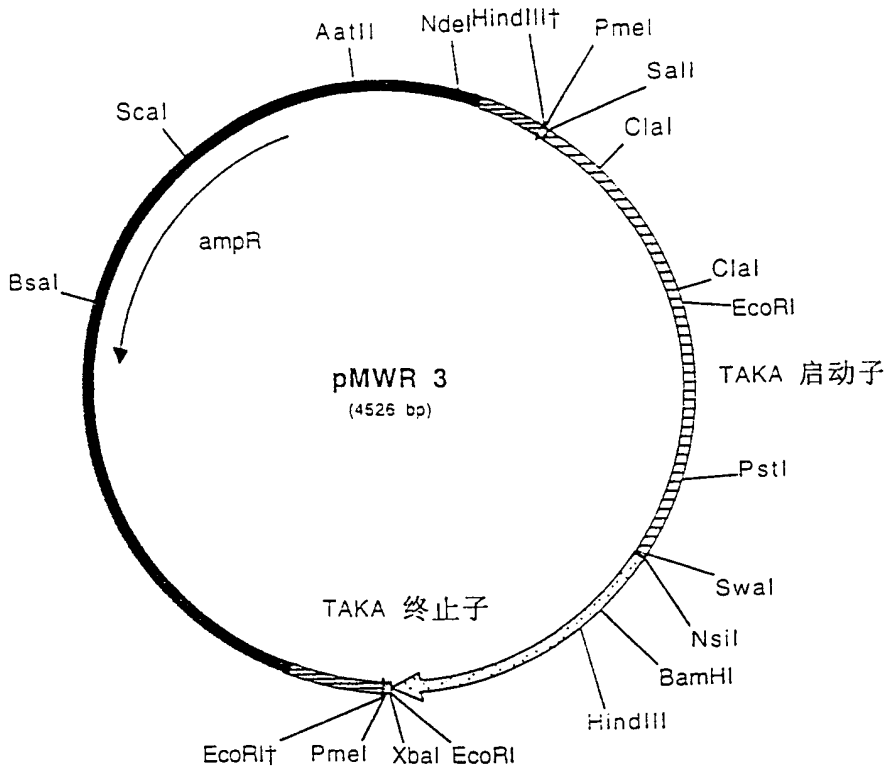


图 2

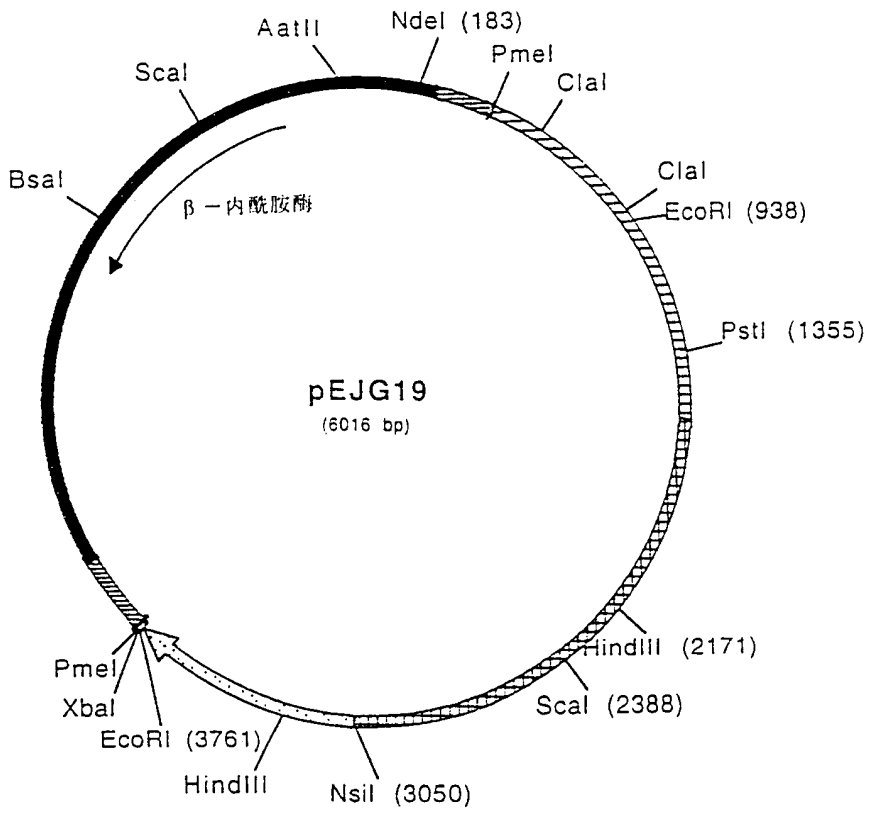


图 3

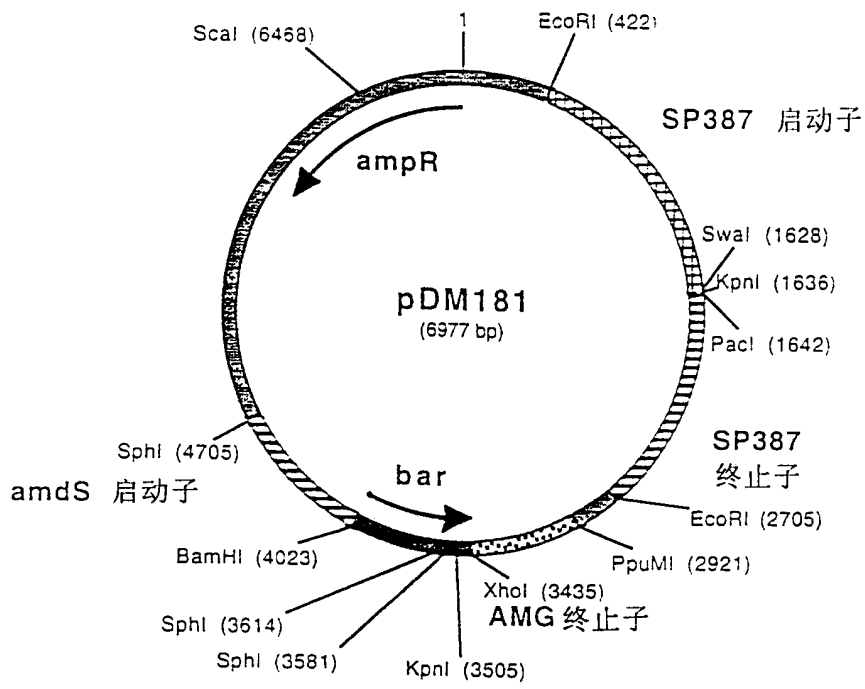


图 4

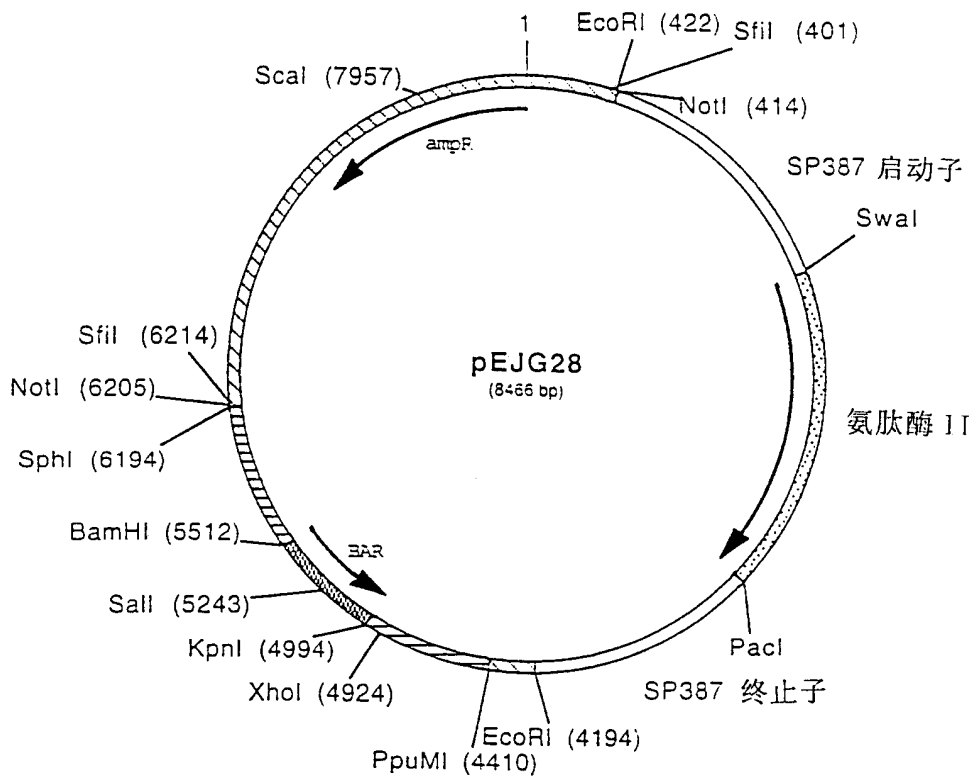


图 5