

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公开说明书

[21] 申请号 200480017079.3

[51] Int. Cl.

*C12N 9/58 (2006.01)*

*C12N 9/52 (2006.01)*

*C12N 15/57 (2006.01)*

*A23J 3/34 (2006.01)*

[43] 公开日 2006年7月26日

[11] 公开号 CN 1809634A

[22] 申请日 2004.6.21

[21] 申请号 200480017079.3

[30] 优先权

[32] 2003.6.19 [33] DK [31] PA200300913

[32] 2003.10.10 [33] DK [31] PA200301492

[32] 2004.3.1 [33] DK [31] PA200400332

[86] 国际申请 PCT/DK2004/000433 2004.6.21

[87] 国际公布 WO2004/111221 英 2004.12.23

[85] 进入国家阶段日期 2005.12.19

[71] 申请人 诺维信公司

地址 丹麦鲍斯韦

[72] 发明人 索伦·F·拉森 卡斯滕·肖霍姆

彼得·R·奥斯特加德

[74] 专利代理机构 北京市柳沈律师事务所

代理人 张文辉 巫肖南

权利要求书 2 页 说明书 64 页 序列表 8 页

[54] 发明名称

蛋白酶

[57] 摘要

与来自拟诺卡氏菌的蛋白酶同源的高特异活性蛋白酶，及其在野生型，和包括转基因植物和非人转基因动物的重组宿主细胞中的生产。蛋白酶在动物饲料和洗涤剂中是有效的。公开了与蛋白酶的高特异活性有关的特有结构特征，该蛋白酶属于肽酶家族 S2A 或 S1E。

1. 具有蛋白酶活性, 且在pH 7.5及25°C时对血红蛋白具有至少39 AU/g 特异活性的分离的多肽, 其中多肽选自下组:
  - 5 (a) 具有与SEQ ID NO: 2的1至188位氨基酸, 和/或SEQ ID NO: 6的1-192位氨基酸有至少65%同一性的氨基酸序列的多肽;
    - (b) 由在低严谨条件下与SEQ ID NO: 1的502-1065位核苷酸, 和/或SEQ ID NO: 5的568-1143位核苷酸杂交的核酸序列所编码的多肽;
    - (c) 由与 SEQ ID NO: 1 的 502-1065 位核苷酸, 和/或 SEQ ID NO: 5 的
  - 10 568-1143 位核苷酸具有至少 74%同一性的核酸序列所编码的多肽。
    2. 包含编码权利要求 1 的多肽的核酸序列的分离的核酸序列。
    3. 分离的核酸序列, 其包含编码具有蛋白酶活性, 且在 pH 7.5 及 25°C 时对血红蛋白具有至少 39 AU/g 特异活性的多肽的核酸序列的, 其中该核酸序列
  - 15 (a) 在低严谨条件下与 SEQ ID NO: 1 的 502-1065 位核苷酸, 和/或 SEQ ID NO: 5 的 568-1143 位核苷酸中任何一个杂交;
    - (b) 与 SEQ ID NO: 1 的 502-1065 位核苷酸, 和/或 SEQ ID NO: 5 的 568-1143 位核苷酸中任何一个具有至少 74%的同一性; 和/或
    - (c) 编码与 SEQ ID NO: 2 的 1 至 188 位氨基酸, 和/或 SEQ ID NO: 6
  - 20 的 1-192 位氨基酸有至少 65%同一性的多肽。
    4. 包含可操作地连接至一个或多个控制序列的权利要求 2-3 中任一个的核酸序列的核酸构建体, 控制序列引导多肽在适当表达宿主中的生产。
    5. 包含权利要求 4 的核酸构建体的重组表达载体。
    6. 包含权利要求 4 的核酸构建体或权利要求 5 的载体的重组宿主细胞。
    - 25 7. 生产权利要求 1 的多肽的方法, 该方法包含: (a)培养权利要求 6 的重组宿主细胞以生产包含所述多肽的上清; 和(b)回收该多肽。
    8. 转基因植物, 或植物部分, 其能够表达权利要求 1 所述的多肽。
    9. 转基因非人动物, 或者产品, 或其部分, 其可以表达权利要求 1 的多肽。
    - 30 10. 生产权利要求 1 的多肽的方法, 该方法包含
      - (a) 培养下述株系中任何一个:

(i) *Nocardiosis dassonvillei* 亚种 *dassonvillei* DSM 43235,

(ii) *Nocardiosis alba* DSM 15647; 和

(b) 回收所述多肽。

5 11. 权利要求1中定义的至少一种多肽(i)在动物饲料中; (ii)在制备用于动物饲料的组合物中; (iii)用于改善动物饲料的营养价值; (iv)用于增加动物食物中的可消化和/或可溶性蛋白; (v)用于提高动物食物中蛋白质的水解程度; 和/或(vi)用于蛋白质处理的用途。

12. 动物饲料添加剂, 其包含权利要求1定义的至少一种多肽; 和

(a) 至少一种脂溶性维生素, 和/或

10 (b) 至少一种水溶性维生素, 和/或

(c) 至少一种微量矿质。

13. 粗蛋白含量为 50 至 800 g/kg 的动物饲料组合物, 其包含至少一种权利要求 1 定义的多肽, 或者至少一种权利要求 12 的饲料添加剂。

15 14. 包含权利要求 1 定义的至少一种多肽的组合物, 其还同时包含至少一种选自淀粉酶、肌醇六磷酸酶、木聚糖酶、半乳聚糖酶、 $\alpha$ -半乳糖苷酶、蛋白酶、磷脂酶、和/或 $\beta$ -葡聚糖酶的其他酶。

15. 权利要求 1 所定义的至少一种多肽在洗涤剂中的用途。

16. *Nocardiosis alba* DSM 15647。

## 蛋白酶

### 5 技术领域

本发明涉及具有蛋白酶活性的与拟诺卡氏菌属(*Nocardiopsis*)蛋白酶同源的分多肽,还涉及编码该多肽的分离的核酸序列。此外,本发明涉及核酸构建体,载体,以及包含这些核酸序列的包括转基因植物和非人动物的宿主细胞,还涉及生产及使用此蛋白酶的方法,尤其在动物饲料中。

10 本发明的蛋白酶具有高度的特异活性。公开了与蛋白酶的高特异活性有关的特有结构特征,该蛋白酶属于肽酶家族 S2A 或 S1E。

### 发明背景

WO88/03947 中公开了源于拟诺卡氏菌(*Nocardiopsis sp.*)NRRL 18262 和  
15 达松维尔拟诺卡氏菌(*Nocardiopsis dassonvillei*) NRRL 18133 的蛋白酶。DK 申请 no. 1996 00013 中公开了源于拟诺卡氏菌 NRRL 18262 的蛋白酶的 DNA 和氨基酸序列。WO 01/58276 公开了与源于拟诺卡氏菌 NRRL 18262 的蛋白酶有关的酸稳定性蛋白酶,以及源于 *Nocardiopsis alba* DSM 14010 的蛋白酶在动物饲料中的应用。然而,这些蛋白酶具有低的特异活性。

20 JP 2-255081-A 公开了源于拟诺卡氏菌株系 OPC-210 (FERM P-10508) 的蛋白酶,但没有序列信息。该株系(strain)已不能获得,因为其保藏已撤消。

DD20043218 公开了源于 *Nocardiopsis dassonvillei* 株系 ZIMET 43647 的蛋白水解制剂,然而没有序列信息。该株系看起来已无法获得。

公布于本发明首次申请日之后的 JP 2003284571-A 公开了源于拟诺卡氏  
25 菌 TOA-1 (FERM P-18676)的蛋白酶的氨基酸序列和相应的 DNA 序列。该序列已登录到 GENESEQP, 条目号为 no. ADF43564。

最为同源的已有非拟诺卡氏菌蛋白酶是 SapII\_ *Streptomyces*\_ sp\_ sptrembl\_q55353, 其成熟部分与 SEQ ID NO: 2 和 6 的成熟部分分别具有  
61.5%和 63.5%的氨基酸同一性。相应的 DNA 与 SEQ ID NO: 1 和 5 分别具  
30 有 70.3%和 72.7%的同一性。相关的链霉菌(*Streptomyces*)蛋白酶的成熟部分

即 SapII\_ Streptomyces\_sp\_sptrembl\_q55352 与 SEQ ID NO: 1 的成熟部分具有稍高的同一性百分率, 即 70.8%。

本发明的目的是提供具有高特异活性的与拟诺卡氏菌蛋白酶同源的蛋白酶, 尤其是具有用于动物饲料和/或洗涤剂潜力的蛋白酶。

5

### 发明概述

已分离并鉴定了具有高特异活性的蛋白酶, 即源于 *Nocardiosis alba* DSM 15647 (见 SEQ ID NO: 1 和 2) 的蛋白酶, 和源于 *Nocardiosis dassonvillei* subsp. *dassonvillei* DSM 43235 (见 SEQ ID NO: 5 和 6) 的蛋白酶。

10 第一个方面, 本发明涉及具有蛋白酶活性的分离的多肽, 且在 pH 7.5 和 25°C 下具有至少 39 AU/g 的针对血红蛋白的特异活性, 其中多肽选自下组: (a) 具有与 SEQ ID NO: 2 的 1 至 188 位氨基酸, 和/或 SEQ ID NO: 6 的 1-192 位氨基酸具有至少 65% 同一性的氨基酸序列的多肽; (b) 由在低严谨条件下与 SEQ ID NO: 1 的 502-1065 位核苷, 和/或 SEQ ID NO: 5 的 568-1143  
15 位核苷酸杂交的核酸序列所编码的多肽; (c) 由与 SEQ ID NO: 1 的 502-1065 位核苷, 和/或 SEQ ID NO: 5 的 568-1143 位核苷具有至少 74% 同一性的核酸序列所编码的多肽。本发明还涉及编码这样的蛋白酶的分离的核酸序列; 核酸构建体, 载体, 和包含此核酸序列的宿主细胞; 还涉及生产此蛋白酶的方法, 以及在尤其是动物饲料方面使用此蛋白酶。

20 第二个方面, 本发明涉及:

A. 肽酶家族 S2A 和/或肽酶家族 S1E 具有蛋白酶活性的分离的多肽, 且在指定位置具有包含至少一个下述氨基酸的氨基酸序列: 25S, 38T, 42P, 44S, 49Q, 54R, 62S, 89S, 91S, 92S, 95A, 99Q, 100I, 114V, 120T, 125Q, 129Q, 131L, 135N, 147F, 151S, 165S, 166F, 171Y, 176N, 179L, 180S, 184L, 和/或 185T; 优选  
25 25S, 38T, 42P, 44S, 54R, 62S, 125Q, 131 L, 165S, 171 Y, 176N, 179L, 180S, 184L, 和/或 185T; 更优选同时具有至少一个 24A, 51V, 53E, 86A, 87T, 96I, 和/或 186L; 和/或同时具有 (H35 + D61 + S143); 其中每个位置对应于 SEQ ID NO: 2 的位置。

B. 包含在指定位置具有至少一个下述氨基酸的 A 所述的多肽: 38T, 92S, 120T, 125Q, 131L, 135N, 147F, 151S, 165S, 和/或 171Y。  
30

C. 包含在指定位置具有至少一个下述氨基酸的 A 所述的多肽: 25S, 42P,

44S, 49Q, 54R, 62S, 89S, 91S, 95A, 99Q, 100I, 114V, 129Q, 166F, 176N, 179L, 180S, 184L, 和/或 185T。

**D.** A、B、或 C 中任何一个多肽，其具有用 DSC 在 10mM 磷酸钠、50 mM 氯化钠、pH 7.0 测定时至少 78°C 的 T<sub>m</sub>；在 pH 9 及 80°C 时至少 0.40 的相对活性；和/或在 pH 7.5 及 25°C 时对血红蛋白至少 39 AU/g 的特异活性。

**E.** A、B、C 或 D 中任何一个多肽，其与 SEQ ID NO: 2 的-167 至 188 位氨基酸，优选 1 至 188 位氨基酸，和/或 SEQ ID NO: 6 的-160 至 192 位氨基酸，优选 1-192 位氨基酸具有至少 65% 的同一性百分数，更优选与 SEQ ID NO: 2 的 1-188 位氨基酸，或-167(167)至 188 位氨基酸具有至少 50% 的同一性百分数。

**F.** A、B、C、D 或 E 中任何一个多肽，其是 i) 细菌蛋白酶；ii) 放线菌 (*Actinobacteria*) 门的蛋白酶；iii) 放线菌纲的 (the class *Actinobacteria*)；iv) 放线菌 (*Actinomycetales*) 目的；v) 拟诺卡氏菌科 (*Nocardiopsaceae*) 的；vi) 拟诺卡氏菌属 (*Nocardiopsis*) 的；和/或源于 vii) 拟诺卡氏菌种如 *Nocardiopsis alba*，南极拟诺卡氏菌 (*Nocardiopsis antarctica*)，*Nocardiopsis prasina*，*Nocardiopsis composta*，*Nocardiopsis exhalans*，*Nocardiopsis halophila*，*Nocardiopsis halotolerans*、*Nocardiopsis kunsanensis*、*Nocardiopsis listeri*、*Nocardiopsis lucentensis*、*Nocardiopsis metallicus*、*Nocardiopsis synnemataformans*、*Nocardiopsis trehalosi*，热带拟诺卡氏菌 (*Nocardiopsis tropica*)、*Nocardiopsis umidischolae*、新疆拟诺卡氏菌 (*Nocardiopsis xinjiangensis*)、或 *Nocardiopsis dassonvillei*；和任选地也源于 *Nocardiopsis alkaliphila*，例如源于 *Nocardiopsis alba* 的蛋白酶，例如，*Nocardiopsis alba* DSM 15647、或源于 *Nocardiopsis dassonvillei* 的蛋白酶、例如 *Nocardiopsis dassonvillei* subsp. *dassonvillei* DSM 43235，例如具有 SEQ ID NO: 2 的-167 至 188，优选 1-188 位氨基酸序列的多肽，和/或具有 SEQ ID NO: 6 的-160 至 192，优选 1-192 位氨基酸序列的多肽。

**G.** 包含编码 A、B、C、D、E 或 F 中任何一个多肽的核酸序列的分离的核酸序列。

**H.** 包含与一个或多个控制序列可操作相连的 G 的核酸序列的核酸构建体，控制序列引导多肽在适当的表达宿主中生产。

**I.** 包含 H 的核酸构建体的重组表达载体。

- J. 包含 H 的核酸构建体或 I 的载体的重组宿主细胞。
- K. 生产 A、B、C、D、E 或 F 中任一多肽的方法，该方法包括：(a) 培养 J 的重组宿主细胞以生产包含该多肽的上清；和(b)回收该多肽。
- L. 能够表达 A、B、C、D、E 或 F 中任一多肽的转基因植物、或植物  
5 部分。
- M. 能够表达 A、B、C、D、E 或 F 中任一多肽的转基因非人动物、或产物、或其部分(element)。
- N. A、B、C、D、E 或 F 任一个所定义的至少一种多肽 (i)在动物饲料中；(ii)在制备用于动物饲料的组合物中；(iii)用于提高动物饲料的营养价值；  
10 (iv)用于增加动物食物中的可消化和/或可溶性蛋白；(v)用于提高动物食物中蛋白质的水解程度；和/或(vi)用于处理蛋白质的用途。
- O. 包含 A、B、C、D、E 或 F 中任一个所定义的至少一种多肽；和(a)至少一种脂溶性维生素，和/或(b)至少一种水溶性维生素，和/或(c)至少一种微量矿质元素(mineral)的动物饲料添加剂。
- 15 P. 粗蛋白含量为 50 至 800 g/kg 的动物饲料组合物，其包含 A、B、C、D、E 或 F 任一个所定义的至少一种多肽，或包含至少一种 O 的饲料添加剂。
- Q. 包含 A、B、C、D、E 或 F 中任一个所定义的至少一种多肽的组合物，其还同时包含至少一种选自 $\alpha$ -淀粉酶(EC 3.2. 1.1)、肌醇六磷酸酶(EC 3.1.  
20 3.8 or 3.1. 3.26)；木聚糖酶(EC 3.2. 1.8)；半乳聚糖酶(galactanase) (EC 3.2. 1.89)； $\alpha$ -半乳糖苷酶 (EC 3.2. 1.22)；蛋白酶(EC 3.4.-.-)，磷脂酶 A1 (EC 3.1. 1.32)；磷脂酶 A2 (EC 3.1. 1.4)；溶血磷脂酶(EC 3.1. 1.5)；磷脂酶 C (3.1. 4.3)；磷脂酶 D (EC 3.1. 4.4)；和/或 $\beta$ -葡聚糖酶(EC 3.2. 1.4 或 EC 3.2. 1.6)中的至少一种其他酶。
- 25 R. A、B、C、D、E 或 F 中任一个所定义的至少一种多肽在洗涤剂中的用途。

第三个方面，本发明涉及：

- a. 具有蛋白酶活性的分离的多肽，选自下组：(a)具有与 SEQ ID NO: 2 的 1  
30 至 188 位氨基酸有至少 86%同一性，和/或与 SEQ ID NO: 6 的 1-192 位氨基酸有至少 72%同一性的氨基酸序列的多肽；(b) 由在中等 - 高严谨条件下与 (i)SEQ ID NO: 1 的 502-1065 位核苷，和/或 SEQ ID NO: 5 的 568-1143 位核

- 5 昔中任意一个, (ii) 至少 100 个核苷的序列(i)的子序列; 和/或(iii) (i)-(ii)任意一个的互补链杂交的核酸序列所编码的多肽; (c)具有 SEQ ID NO: 2 的 1 至 188 位氨基酸, 或 SEQ ID NO: 6 的 1-192 位氨基酸的多肽的变异体, 包含一个或多个氨基酸的取代, 删除, 延伸, 和/或插入; (d) (a)、(b)、或(c)的等位变异体; 和(e) 具有蛋白酶活性的(a)、(b)、(c)、或(d)的片段;
- 10 b. 分离的核酸序列, 其为包含(a)编码 *a* 的多肽的核酸序列; (b)编码具有蛋白酶活性的多肽, 且其在中等 - 高严谨条件下与(i)SEQ ID NO: 1 的 502-1065 位核苷, 和/或 SEQ ID NO: 5 的 568-1143 位核苷中任意一个, (ii) 序列(i)的至少 100 个核苷; 和/或(iii) (i)-(ii)任意一个的互补链杂交的核酸序列; 和/或(c)编码具有蛋白酶活性的多肽, 且其与(i) SEQ ID NO: 1 的 502-1065 位核苷具有至少 86% 的同一性, 和/或与 SEQ ID NO: 5 的 568-1143 位核苷具有至少 82% 的同一性的核酸序列的分离的核酸序列。
- 15 c. 分离的核酸序列, 其为通过(a)在中等 - 高严谨条件下将 DNA 与(i) SEQ ID NO: 1 的 502-1065 位核苷, 和/或 SEQ ID NO: 5 的 568-1143 位核苷中的任意一个; (ii)序列(i)的至少 100 个核苷; 和/或(iii) (i)-(ii)中任一个的互补链杂交; 和(b)分离核酸序列, 来生产分离的核酸序列;
- d. 包含与一个或多个控制序列可操作相连的 *b* 或 *c* 的核酸序列的核酸构建体, 控制序列引导多肽在适当的表达宿主中生产;
- e. 包含 *d* 的核酸构建体的重组表达载体;
- 20 f. 包含 *d* 的核酸构建体或 *e* 的载体的重组宿主细胞;
- g. 生产 *a* 的多肽的方法, 该方法包括: (a)培养 *f* 的重组宿主细胞以生产包含该多肽的上清(supernatant); 和(b)回收该多肽;
- h. 能够表达 *a* 的多肽的转基因植物、或植物部分;
- i. 能够表达 *a* 的多肽的转基因非人动物、或产物、或其部分;
- 25 j. 生产 *a* 的多肽的方法, 该方法包含(a)培养下述株系中的任一个: (i) *Nocardiosis dassonvillei* subsp. *dassonvillei* DSM 43235, 或 *Nocardiosis alba* DSM 15647; 和(b)回收多肽;
- k. *a* 中定义的至少一种多肽(i)在动物饲料中; (ii)在制备用于动物饲料的组合物中; (iii)用于提高动物饲料的营养价值; (iv)用于增加动物食物中的可消化和/或可溶性蛋白; (v)用于提高动物食物中蛋白质的水解程度; 和/或(vi)用于处理蛋白质的用途;
- 30

- l. 包含 *a* 中定义的至少一种多肽; 和(a)至少一种脂溶维生素, 和/或(b)至少一种水溶维生素, 和/或(c)至少一种微量矿质元素的动物饲料添加剂;
- m. 粗蛋白含量为 50 至 800 g/kg 的动物饲料组合物, 其包含 *a* 中定义的至少一种多肽, 或包含至少一种 *l* 的饲料添加剂;
- 5 n. 包含 *a* 中定义的至少一种多肽的组合物, 其还同时包含至少一种选自 $\alpha$ -淀粉酶(EC 3.2. 1.1)、肌醇六磷酸酶(EC 3.1. 3.8 or 3.1. 3.26); 木聚糖酶(EC 3.2. 1.8); 半乳聚糖酶(EC 3.2. 1.89);  $\alpha$ -半乳糖苷酶 (EC 3.2. 1.22); 蛋白酶(EC 3.4.--), 磷脂酶 A1 (EC 3.1. 1.32); 磷脂酶 A2 (EC 3.1. 1.4); 溶血磷脂酶(EC 3.1. 1.5); 磷脂酶 C (3.1. 4.3); 磷脂酶 D (EC 3.1. 4.4); 和/或 $\beta$ -葡聚糖酶(EC 3.2. 1.4 或 EC 3.2. 1.6)中的至少一种其他酶; 以及
- 10 o. *a* 中定义的至少一种多肽在洗涤剂中的用途。

- 第四个方面, 本发明涉及: 具有蛋白酶活性的分离的多肽, 选自下组:
- (a)具有与 SEQ ID NO:2 的 1 至 188 位氨基酸有至少 84%同一性的氨基酸序列的多肽;(b) 具有与 SEQ ID NO:2 的-167 至 188 位氨基酸有至少 78%同一性的氨基酸序列的多肽;(c)由在中等-高严谨条件下与(i) 可使用引物 SEQ ID NO.3 和 4 由 *Nocardiosis alba* DSM 15647 基因组 DNA 获得的编码蛋白酶的 DNA; (ii) SEQ ID NO: 1 的 502-1065 位核苷; (iii) SEQ ID NO: 1 的 1-1065 位核苷; (iv)至少 100 个核苷的(i)或(ii)或(iii)的子序列; 和/或(v) (i)、(ii)、(iii)或(iv)的互补链杂交的核酸序列所编码的多肽; (d)具有 SEQ ID NO: 20 2 的 1 至 188, 或-167 至 188 位氨基酸序列的多肽的变体, 其包含一个或多个氨基酸的取代、删除、延伸、和/或插入; (e) (a)、(b)或(c)的等位变体; 和(f)具有蛋白酶活性的(a)、(b)、(c)、(d)或(e)的片段。

- 具有蛋白酶活性的分离的多肽, 其在用差示扫描量热法(Differential Scanning Calorimetry, DSC)在 10mM 磷酸钠、50 mM 氯化钠缓冲液、pH 7.0、扫描速率常数 1.5°C/min 的条件下测定时, 具有至少 78°C 的解链温度( $T_m$ ), 其中多肽选自下组: (a) 具有与 SEQ ID NO: 2 的 1 至 188 位氨基酸有至少 50%同一性的氨基酸序列的多肽; (b) 具有与 SEQ ID NO: 2 的-167 至 188 位氨基酸有所述同一性的氨基酸序列的多肽; (c)由在低严谨条件下与(i) 可使用引物 SEQ ID NO.3 和 4 由 *Nocardiosis alba* DSM 15647 基因组 DNA 获得的编码蛋白酶的 DNA; (ii) SEQ ID NO: 1 的 502-1065 位核苷; (iii) SEQ ID NO: 1 的 1-1065 位核苷; (iv)至少 100 个核苷的(i)或(ii)或(iii)的子序列; 和/
- 25
- 30

或(v) (i)、(ii)、(iii)或(iv)的互补链杂交的核酸序列所编码的多肽; (d)具有 SEQ ID NO: 2 的 1 至 188, 或-167 至 188 位氨基酸序列的多肽的变异体, 其包含一个或多个氨基酸的取代、删除、延伸、和/或插入; (e) (a)、(b)或(c)的等位变异体; 和(f)具有蛋白酶活性的(a)、(b)、(c)、(d)或(e)的片段。

- 5 包含(a)编码以上定义的多肽的核酸序列; (b)编码具有蛋白酶活性的多肽, 且在中等-高严谨条件下与(i)可使用引物 SEQ ID NO.3 和 4 由 *Nocardiosis alba* DSM 15647 基因组 DNA 获得的编码蛋白酶的 DNA; (ii) SEQ ID NO: 1 的 502-1065 位或 1-1065 位核苷; (iii)至少 100 个核苷的(i)或(ii)的子序列; 和/或(iv) (i)、(ii)或(iii)的互补链杂交的核酸序列; (c)编码具有蛋白酶活性的多肽, 且与 SEQ ID NO: 1 的 502-1065 位核苷有至少 86%同一性的核酸序列; (d)编码具有蛋白酶活性的多肽, 且与 SEQ ID NO: 1 的 1-1065 位核苷有至少 82%同一性的核酸序列的分离的核酸序列。

- 包含编码具有蛋白酶活性的多肽的核酸序列的分离的核酸序列, 其在用差示扫描量热法(DSC)在 10mM 磷酸钠、50 mM 氯化钠缓冲液、pH 7.0、扫描速率常数 1.5°C/min 的条件下测定时, 具有至少 78°C 的解链温度( $T_m$ ), 其中所述核酸序列(a)编码如上定义的  $T_m$  至少为 78°C 的多肽; (b)在低严谨条件下与(i)可使用引物 SEQ ID NO.3 和 4 由 *Nocardiosis alba* DSM 15647 基因组 DNA 获得的编码蛋白酶的 DNA; (ii) SEQ ID NO: 1 的 502-1065 位或 1-1065 位核苷; (iii)至少 100 个核苷的(i)或(ii)的子序列; 和/或(iv) (i)、(ii)或(iii)的互补链杂交; (c)与 SEQ ID NO: 1 的 502-1065 位核苷具有至少 50% 的同一性; 和/或(d) 与 SEQ ID NO: 1 的 1-1065 位核苷具有至少 50%的同一性。

- 分离的核酸序列, 其通过(a)在中等-高严谨条件下将 DNA 与(i)可使用引物 SEQ ID NO.3 和 4 由 *Nocardiosis alba* DSM 15647 基因组 DNA 获得的编码蛋白酶的 DNA; (ii) SEQ ID NO: 1 的 502-1065 位核苷; (iii)至少 100 个核苷的(i)或(ii)的子序列; 或(iv) (i)、(ii)或(iii)的互补链杂交; 和(b)分离核酸序列而生产。

- 核酸构建体, 其包含紧接本段的上面三段中任何一段所定义的一条核酸序列之一, 其中所述核酸序列与一个或多个引导多肽在适当的表达宿主中生产的控制序列可操作相连。

包含所述核酸构建体的重组表达载体。

包含所述核酸构建体或载体的重组宿主细胞。

生产上述定义的多肽的方法，该方法包含：(a)培养权利要求8的重组宿主细胞以生产包含所述多肽的上清；和(b)回收所述多肽。

能够表达上述定义的多肽的转基因植物，或植物部分。

5 能够表达如上定义的多肽的转基因非人动物，或产物，或其部分。

上述定义的至少一种多肽(i)在动物饲料中；(ii)在制备用于动物饲料的组合物中；(iii)用于改善动物饲料的营养价值；(iv)用于增加动物食物(diet)中的可消化和/或可溶性蛋白；(v)用于提高动物食物中蛋白质的水解程度；和/或(vi)用于处理蔬菜蛋白质的用途。

10 包含如上定义的至少一种多肽；和(a)至少一种脂溶维生素，和/或(b)至少一种水溶维生素，和/或(c)至少一种微量矿质元素的动物饲料添加剂。

粗蛋白含量为50至800 g/kg的动物饲料组合物，其包含如上定义的至少一种多肽，或包含如上定义的至少一种饲料添加剂。

包含如上定义的至少一种多肽的组合物，其同时还包含至少一种选自  
15  $\alpha$ -淀粉酶(EC 3.2. 1.1)、肌醇六磷酸酶(EC 3.1. 3.8 or 3.1. 3.26)；木聚糖酶(EC 3.2. 1.8)；半乳聚糖酶(EC 3.2. 1.89)； $\alpha$ -半乳糖苷酶 (EC 3.2. 1.22)；蛋白酶(EC 3.4.-.-)，磷脂酶 A1 (EC 3.1. 1.32)；磷脂酶 A2 (EC 3.1. 1.4)；溶血磷脂酶(EC 3.1. 1.5)；磷脂酶 C (3.1. 4.3)；磷脂酶 D (EC 3.1. 4.4)；和/或 $\beta$ -葡聚糖酶(EC 3.2. 1.4 或 EC 3.2. 1.6)中的至少一种其他酶。

20 如上定义的至少一种多肽在洗涤剂中的用途。

上述第二、第三、和第四个方面的实施方案相互独立，也是本发明第一个方面的优选亚方面，同时也互为优选的亚方面。

### 发明详述

25 具有蛋白酶(protease)活性的多肽，或蛋白酶，有时候也称肽酶、蛋白酶(proteinases)、肽水解酶、或蛋白水解酶。蛋白酶可以从肽的任何一端开始水解肽的外切型，或者是在多肽链内部起作用的内切型(肽内切酶)。肽内切酶在N和C端阻断的肽底物上显示活性，肽底物与所述蛋白酶的特异性相关。

30 此处术语“蛋白酶”定义为水解肽键的酶。其包括属于EC 3.4的酶组酶(包括其13个亚类中的任何一类)中的任何酶。EC编号参考NC-IUBMB，

Academic Press, San Diego, California 的酶命名法(Enzyme Nomenclature)1992, 包括分别发表于 Eur. J. Biochem. 1994,223, 1-5; Eur. J. Biochem. 1995,232, 1-6; Eur. J. Biochem. 1996,237, 1-5; Eur. J. Biochem. 1997,250, 1-6; 和 Eur. J. Biochem. 1999,264, 610-650 的增刊 1-5。该命名法  
5 定期补充更新; 参见如 World Wide Web (WWW)网址 <http://www.chem.qmw.ac.uk/iubmb/enzyme/index.html>。

以其催化机制为基础将蛋白酶分为下述的组: 丝氨酸蛋白酶(S)、半胱氨酸蛋白酶(C)、天冬氨酸蛋白酶(A)、金属蛋白酶(M)和未知的或未分类的蛋白酶(U), 参见 Handbook of Proteolytic Enzymes, A. J.Barrett, N. D. Rawlings,  
10 J. F. Woessner (eds), Academic Press (1998), 尤其是一般介绍部分。

在特定实施方案中, 本发明的蛋白酶和根据本发明的用途选自下组:

- (a) 属于 EC 3.4.-.- 酶组的蛋白酶;
- (b) 属于上述手册中 S 组的丝氨酸蛋白酶;
- (c) 肽酶家族 S2A 的丝氨酸蛋白酶; 和/或
- 15 (d) Biochem. J. 290: 205-218 (1993)和 MEROPS 蛋白酶数据库, 版本 6.20, March 24,2003, ([www.merops.ac.uk](http://www.merops.ac.uk))中描述的肽酶家族 S1E 的丝氨酸蛋白酶。Rawlings, N. D. , O'Brien, E. A. & Barrett, A. J. (2002) MEROPS: the protease database. Nucleic Acids Res. 30,343-346 描述了该数据库。

为了测定给定的蛋白酶是否丝氨酸蛋白酶, 及 S2A 家族蛋白酶, 将上述的手册及其指定的原则作为参考。这样的测定可用于所有类型的蛋白酶, 无论其是天然的还是野生型的蛋白酶; 或者是基因工程化或合成的蛋白酶。

可以采用任何的测试方法来测量蛋白酶活性, 所述测试方法中使用底物, 该底物包括与所讨论的蛋白酶的特异性有关的肽键。pH-测试和温度-测试同样适合于所讨论的蛋白酶。pH 值-测试的例子为  
25 pH2,3,4,5,6,7,8,9,10,11,或 12。温度-测试的例子为 30,35,37,40,45,50, 55,60,65,70, 80,90,或 95°C。

非限制性的蛋白酶底物的例子为酪蛋白, 例如 Azurine-交联酪蛋白 (AZCL-Casein), 和血红蛋白。用于测定本发明蛋白酶的特异活性目的的底物是血红蛋白, 实施方案 3 公开了适合的测试方法。实施方案 2 描述了两个其他的蛋白酶, 其中任何一个都可以用于测定大致的蛋白酶活性。所谓的  
30 pNA 测试是用于除特异活性测定之外的目的的优选测试方法。

pH 7.5, 25°C 下, 本发明的蛋白酶显示了针对血红蛋白的至少 39AU/g 的特异活性。可按实施方案 3 所述测定此特异活性。本发明的蛋白酶可以显示出至少 40,41, 42, 43,44, 45,46, 47,48, 49,50, 51,52, 53,或至少 54AU/g 的特异活性。

5 特异性测定, 包括纯化蛋白酶的蛋白质含量及蛋白酶活性的测定是大家熟知的。

在特定实施方案中, 通过氨基酸分析, 例如通过蛋白酶的酸水解, 之后分离并量化释放的氨基酸来测定蛋白质含量, 优选在 Biochrom 20 Plus 氨基酸分析仪上测定。

10 以下是蛋白酶活性测定中特定的、任选的技术特征: (i)将血红蛋白底物变性; (ii)以 0.65%w/w 的量使用血红蛋白底物; (ii)测试缓冲液是  $\text{KH}_2\text{PO}_4/\text{NaOH}$  缓冲液, pH 7.50; (ii)蛋白酶的反应时间为 10 分钟; (iii)酶反应后, 用三氯乙酸(TCA)沉淀未消化的血红蛋白并将其去除, 优选通过过滤去除; (iv)测定滤液中的 TCA-可溶性血红蛋白降解产物, 优选用 Folin &  
15 Ciocalteu's 苯酚试剂测定; (v) 参考 ALCALASE™ 酶标准测量并确定活性单位(AU); (vi)用 EB-SM-0349 测试, 优选 EB-SM-0349 02/01 测量活性单位(AU); 和/或(vii)所用的测试为实施方案 3 所公开的“蛋白酶活性测试(AU/ml)”。ALCALASE™ 标准和 EM-SM-0349 测试可从 Novozymes A/S, Krogshoejvej 36, DK-2880 Bagsvaerd, Denmark (分别提供" your ref. Patent  
20 10476"和"EB- SM-0349", 或"EB-SM-0349 02/01,")获得。

与 SEQ ID NO: 2 和 6 的蛋白酶相关的高特异活性蛋白酶的筛选可按如下方法进行: 第一步, 用 SEQ ID NO: 3,4, 7, 或 8 的引物, 或者优选用 SEQ ID NO:1 和 5 的成熟肽编码部分筛选 DNA 文库; 在合适的株系例如芽孢杆菌(*Bacillus*)或大肠杆菌(*E. coli.*)中表达杂交克隆。下一步, 纯化所表达的与  
25 SEQ ID NO: 2 和/或 6 相关的蛋白酶, 优选在微-纯化方法中纯化(参见, 如 WO 03/037914), 在下面的步骤, 用酶的强抑制剂, 以熟知的活性定点滴定方法(active site titration ,AST)测定每个候选蛋白酶的活性蛋白酶的量。这样做是考虑到在接下来的最后步骤中能够比较同样摩尔量的每个蛋白酶, 所述最后步骤是使用任何适当的测试方法测定目前已知量的蛋白酶的蛋白酶  
30 活性, 例如实施方案 2 所用的 pNA 测试。该程序的大部分可以是自动化的,

如果需要的话可以在机器人的协助下进行。按照此处实验部分的描述，通过例如纯化蛋白酶及确定特异活性，完成高度特异活性的鉴定。

本发明蛋白酶的来源和/或其根据本发明的用途没有限制。如此，术语蛋白酶不仅包括由任何种类的微生物获得的天然的或野生型的蛋白酶，还包括其显示蛋白酶活性的任何突变体、变异体、片段等等，还包括合成的蛋白酶，例如改组蛋白酶(shuffled proteases)，共有蛋白酶(consensus proteases)。这样的基因工程蛋白酶可以按照本领域熟知的方法制备，例如通过定点突变，通过PCR(在PCR反应中将包含所需突变的PCR片段用作其中一个引物)，或者通过随机突变。例如EP 897985中描述了共有蛋白酶的制备。例如WO95/22625和WO 96/00343中简要描述了基因改组(gene shuffling)。可以通过如Ness, J. E.等在Nature Biotechnology, Vol. 20 (12), pp. 1251-1255, 2002中所述的合成改组，由亲本的特定序列独立地完成蛋白酶基因的重组。按照参考文献所述，设计在其DNA序列中简并的合成寡核苷酸并组装基因，在其DNA序列中简并的合成寡核苷酸提供了在该组亲本蛋白酶中发现所有氨基酸的可能性。可以对全长序列进行改组，也可以只对该序列的一部分进行改组，然后再与基因的剩余部分组合而得到全长序列。具有包含SEQ ID NO: 2和6任一个的成熟部分的氨基酸序列的蛋白酶是这种可用于上述改组的亲本蛋白酶的特定例子，如果需要，还可同时加上源于如拟诺卡氏菌NRRL 18262的蛋白酶以提供本发明额外的蛋白酶。此处与给定来源连用的术语“由……获得的”(obtained from)意思应为，核酸序列所编码的多肽由该来源生产，或者由其中存在该来源的核酸序列的细胞生产。在一个优选实施方案中，该多肽是胞外分泌的。

在特定实施方案中，此蛋白酶为低致敏性变异体，其被设计用于在接触动物包括人时激发降低的免疫反应。术语免疫反应应理解为由接触该蛋白酶的动物的免疫系统引起的任何反应。免疫反应的一种类型为导致被接触(exposed)动物的IgE水平升高的过敏反应。可以使用本领域已知的技术制备低致敏性变异体。例如，可以将蛋白酶与涉及免疫反应的蛋白酶的多聚体部分(moiety)屏蔽部分或者抗原决定簇结合。与多聚体的结合可能涉及如WO 96/17929, WO 98/30682, WO 98/35026,和/或WO 99/00489中所描述的多聚体与蛋白酶的体外化学连接。结合可以额外地或可选地设计体内多聚体与蛋白酶结合。这样的结合可以通过对编码该蛋白酶的核苷序列实施基因

工程，插入编码蛋白酶中额外的糖基化位点的共有序列，以及在能够使蛋白酶糖基化的宿主中表达该蛋白酶，参见，例如 WO 00/26354。另外一种提供低致敏性变异体的方法是对编码该蛋白酶的核苷序列实施基因工程，以使该蛋白酶自我寡聚化，使蛋白酶单体能够屏蔽其他蛋白酶单体的抗原决定簇，并因此降低寡聚体的抗原性。例如 WO 96/16177 中描述了这样的产物及其制备。可以通过多种方法例如 WO 00/26230 和 WO 01/83559 中描述的噬菌体展示方法，或者 EP 561907 中描述的随机方法来鉴定免疫反应中涉及的抗原决定簇。一旦鉴定了抗原决定簇(epitope)，可以通过已知的基因操作技术如定点突变(参见，如 WO 00/26230, WO 00/26354 和/或 WO 00/22103)来改变其氨基酸序列，以生产免疫特性被改变的蛋白酶，或者可以使多聚体与抗原决定簇足够近地结合以使多聚体屏蔽该抗原决定簇。

根据本发明任一方面的多肽，可以包含与 SEQ ID NO: 2 或 6 中任一个的成熟肽部分，例如 SEQ ID NO: 2 的 1 至 188 位氨基酸，和/或 SEQ ID NO: 6 的 1 至 192 位氨基酸(成熟肽部分)具有例如约 65%同一性的氨基酸序列，且其具有蛋白酶活性(以下称“同源多肽”)。在特定实施方案中，与 SEQ ID NO: 2 或 6 中任一个的成熟肽部分的同一性至少为约 66%，67%，68%，69%，70%，71%，72%，73%，74%，75%，76%，77%，78%，79%，80%，81%，82%，83%，84%，85%，86%，87%，88%，89%，90%，91%，92%，93%，94%，95%，96%，97%，98%，或者 99%。在可选实施方案中，同一性至少为约 50%，51%，52%，53%，54%，55%，56%，57%，58%，59%，60%，61%，62%，63%，或者 64%。

本发明也涉及包含与 SEQ ID NO: 2 的-167 至 188 位氨基酸，和/或与 SEQ ID NO: 6 的-160 至 192 位氨基酸具有例如至少约 78%同一性的氨基酸序列的分离的多肽，且其具有蛋白酶活性。在特定实施方案中，与 SEQ ID NO: 2 的-167 至 188 位氨基酸具有至少约 80%，85%，86%，87%，88%，89%，90%，91%，92%，93%，94%，95%，96%，97%，98%，或者 99%的同一性。在可选实施方案中，同一性至少为约 50%，51%，52%，53%，54%，55%，56%，57%，58%，59%，60%，61%，62%，63%，64%，65%，66%，67%，68%，69%，70%，71%，72%，73%，74%，75%，76%，77%，78%，79%，80%，81%，82%，83%，或者 84%。

在特定实施方案中，本发明的多肽 i)具有上述提到的任何同一性的氨基酸序列；或者 ii)由具有上述提到的任何同一性的氨基酸序列组成。

为达到本发明的目的，使用 Needleman-Wunsch(例如整体比对)“比对”程序测定两个氨基酸序列之间以及两个核酸序列之间的同一性。该程序用于多肽的比对，也同时用于核苷序列比对。用默认的计分阵列 BLOSUM50 进行多肽比对，默认的一致性阵列用于核苷比对。对多肽来说，空位(gap)的  
5 的第一个残基的罚分(penalty)为-12，对核苷来说为-16。空位的其他残基的罚分，多肽为-2，核苷为-4。

“比对”是版本号为 v20u6 的 FASTA 软件包(参见 W. R. Pearson and D. J. Lipman (1988), "Improved Tools for Biological Sequence Analysis", PNAS 85: 2444-2448, 和 W. R. Pearson (1990) "Rapid and Sensitive Sequence  
10 Comparison with FASTP and FASTA, "Methods in Enzymology 183 : 63-98)的一部分。FASTA 蛋白质比对使用 Smith- Waterman 运算法则，对空位大小没有限制(参见"Smith-Waterman algorithm", T. F. Smith and M. S. Waterman (1981) J. Mol. Biol. 147: 195-197)。

在特定实施方案中，将两个氨基酸序列的成熟肽部分、或者预期或所  
15 希望的成熟肽部分用于比对。可选地，选择正在测定其与 SEQ ID NO: 2 的成熟肽部分的同一性的序列，依据按照下述方法所作的多序列比对结果，其与 SEQ ID NO: 2 的成熟肽部分，如通过多序列比对所鉴定的相应氨基酸残基最为相似。

在此上下文中，为氨基酸残基编号(或者指定位置号，参考本发明的第  
20 二个方面)的基础是 SEQ ID NO: 2，起始于 A1 结束于 T188。可选地，该基础是源于拟诺卡氏菌 NRRL 18262 的蛋白酶的 1-188 位氨基酸(WO 01/58276 中公开的 SEQ ID NO: 1, 优选 WO 01/58276 中公开的 A87 被 T87 取代的 SEQ ID NO: 1)。

蛋白酶可包含相对于成熟肽部分延伸，即在其 N 端和/或 C 端的延伸。  
25 如果要给任一延伸的氨基酸编号的话，将依据本领域的惯例对其编号，例如，对于 C 端延伸：189,190,191 等等，对于 N 端延伸-1,-2,-3 等。

对于与参照序列，例如 SEQ ID NO: 2 比对的每一个蛋白酶中的每一个氨基酸残基，如上面解释的那样(用于测定同一性目的)，直接无疑义地指定其在参照序列如 SEQ ID NO: 2 中对应的氨基酸残基是可能的。参照例如  
30 SEQ ID NO: 2，将相应的残基指定为相同的位置或编号。

对于另一蛋白酶中的每一氨基酸残基,可以找出参照序列如 SEQ ID NO: 2 的相应位置,如下:

5 将其他蛋白酶的氨基酸序列指定为 SEQ X。按如下方法找出相应于 SEQ ID NO: 2 的位置 N 的位置:如上所述,将 SEQ X 与 SEQ ID NO: 2 比对。使用下述的原则,可由比对结果直接无疑义地得出 SEQ X 中相应于 SEQ ID NO: 2 的位置 N 的位置。

10 SEQ X 可以是所讨论的蛋白酶的成熟部分,或者也可以包括信号肽部分,或者是具有蛋白酶活性的成熟蛋白酶的片段,例如与 SEQ ID NO: 2 长度相同的片段,和/或可以是与此处描述的 SEQ ID NO: 2 比对时从 A1 延伸至 T188 的片段。

下面插入了三个比对即表 I、II 和 III。通过如上描述的将另一蛋白酶(分别为 SEQ X1, SEQ X2 和 SEQ X3)与 SEQ ID NO: 2 比对得到这些比对结果。其显示了每个蛋白酶的大约 50 个氨基酸残基。

15 首先看表 I 的比对,很明显,SEQ ID NO: 2 的 P42 对应于 SEQ X1 的 Q42,因为在此比对中这些残基上下对齐。均为其指定编号 42,即 SEQ ID NO: 2 中相应残基的编号。由此比对,同样很明显,例如 SEQ X1 不包含 25S, 38T, 42P, 44S,或 49Q 的任何一个。

**Table I**

	ADIIGGLAYT	MGGRCSVGFA	ATNASGQPGF	VTAGHCCTVG	TPVSIGNGQG	SEQ ID NO: 2
20	ADIIGGLAYT	MGGRCSVGFA	ATNAAGQPGF	VTAGHCGRVG	TQVTIGNGRG	SEQ X1
	10	20	30	40	50	

表 II 和 III 是在两个序列任何一个中产生空位的比对的例子。

25 在表 II 的比对中,SEQ X2 中产生了空位。为了当前的目的,将 SEQ X2 中突出显示的氨基酸残基 P 指定为 P28,尽管这样的位置在 SEQ X 中为 P25。

**Table II**

ADIIGGLAYT	MGGRCSVGFA	ATNASGQPGF	VTAGHCGTVG	TPVSI <del>GN</del> GQG	SEQ ID NO: 2
ADIIGGLAYT	MGGRCSVGFA	ATNA--- <u>P</u> GF	VTAGHCGRVG	TQVTIGNGRG	SEQ X2
10	20	30	40	50	

- 5 表 III 的比对中，在 SEQ X3 中产生了空位。当在 SEQ ID NO: 2 的位置编号为 nn 和(nn+1)的氨基酸之间产生空位时，将空位的每个位置指定为前面位置号例如 nn 的小写字母或下标：a, b, c 等。相应地，空位的每一个位置编号为 nna, nnb 等。为当前的目的，将 SEQ X3 中突出显示的氨基酸残基 R 指定为 R33a，尽管在 SEQ X3 中这样的位置为 R34。

10

**Table III**

ADIIGGLAYT	MGGRCSVGFA	ATNASGQPGF	VTA--GHCCT	VGTPVSI <del>GN</del> GQG	SEQ ID NO: 2
ADIIGGLAYT	MGGRCSVGFA	ATNAAGQPGF	VT <u>A</u> RS <del>G</del> HCGR	VGTTQVTIGNGRG	SEQ X3
10	20	30	40	50	

- 15 在本发明任何方面的另外的特定实施方案中，通过差示扫描量热法 (DSC)测定，本发明的多肽具有至少 75°C, 76°C, 77°C, 78°C, 79°C, 80°C, 81°C, 82°C, 83°C, 84°C, 85°C, 86°C, 87°C, 88°C, 89°C, 90°C, 91°C, 92°C, 93°C, 94°C, 或者 95°C 的解链温度 T<sub>m</sub>。DSC 测定在 10 mM 磷酸钠、50 mM 氯化钠缓冲液、pH 7.0 的条件下进行。扫描速率为常数，例如 1.5°C/min。扫描间隔可  
20 以从 20 至 100°C。

T<sub>m</sub> 没有上限，但是，目前预期 T<sub>m</sub> 可能在 150°C, 145°C, 140°C, 135°C, 130°C, 125°C, 120°C, 115°C, 110°C, 105°C, 或者 100°C 以下。

在可选实施方案中，选择另一缓冲液用于扫描，例如 pH 5.0, 5.5, 6.0, 或 pH 6.5。

- 25 在另外的可选实施方案中，可以使用更高或更低的扫描速率，例如低扫描速率 1.4°C/min, 1.3°C/min, 1.2°C/min, 1.1°C/min, 1.0°C/min, 或 0.9°C/min。

有关扫描程序的更详细资料可参见作为参考的实施方案 2。

在特定实施方案中，与源于拟诺卡氏菌 NRRL 18262 的蛋白酶相比，本发明的蛋白酶显示了改变的温度活性谱。例如，pH 9 及 80°C 条件下，本发明的蛋白酶可以显示出至少 0.40，优选至少 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90，或者 0.95 的相对活性，术语“相对”指相对于测定的所述蛋白酶的最大活性。对于源于拟诺卡氏菌 NRRL 18262 的蛋白酶，70°C 时的活性设定为 1.000(100%)，参见实施方案 2。如另一实施方案，pH 9 及 90°C 条件下，本发明的蛋白酶显示了至少 0.10，优选至少 0.15, 0.20, 0.25, 0.30，或者 0.35 的相对活性。在特定实施方案中，用实施方案 2 的 Protazyme AK 测试测量蛋白酶活性。

10 本发明还涉及本发明多肽的动物饲料用途。

也可以用具有同一性表的 LASERGENE™ MEGALIGN™ 软件 (DNASTAR, Inc., Madison, WI)，通过 Clustal 方法(Higgins, 1989, CABIOS 5: 151-153)测定两个氨基酸序列之间的同一性，多序列比对参数如下：空位罚分为 10、空位长度罚分为 10。成对比对参数为 Ktuple=1、空位罚分=3、窗口=5、对角线(diagnosis)=5。可以用与上述相同的运算法则和软件包测定两个核苷序列之间的同一性，设定如下：空位罚分为 10，空位长度罚分为 10。成对比对参数为 Ktuple=3、空位罚分=3 及窗口=20。

在特定实施方案中，同源肽具有与(a) SEQ ID NO: 2 或 6 中任一个的成熟肽部分，或者(b)其前体形式(不包括信号肽部分，包括成熟肽部分)不同的氨基酸序列，所述不同具有(i)不超过 50,49, 48,47, 46,45, 44,43, 42,41, 40,39, 38,37, 36,35, 34, 33,32, 31,30, 29,28, 27,26, 25,24, 23,22, 21,或者 20 个氨基酸；(ii)不超过 20,19,18,17,16,15,14,13,12,或者 11 个氨基酸；(iii)不超过 10,9,8,7,6,5,4,3,2,或 1 个氨基酸；(iv)10,9,8,7,6,5 个氨基酸；或者(v)4,3,2,或者 1 个氨基酸的差异。

25 在特定实施方案中，本发明的多肽包含 SEQ ID NO: 2 或 6 中任一个的成熟肽部分的氨基酸序列，或其等位变异体；或其具有蛋白酶活性的片段。

在另一优选实施方案中，本发明的多肽由 SEQ ID NO: 2 或 6 中任一个的成熟肽部分，或其等位变异体；或其具有蛋白酶活性的片段组成。

30 片段是从这些氨基酸序列的氨基和/或羧基端删除一个或多个氨基酸的多肽。在一个实施方案中，片段包含至少 75 个氨基酸残基，或至少 100 个氨基酸残基，或至少 125 个氨基酸残基，或至少 150 个氨基酸残基，或至

少 160 个氨基酸残基，或至少 165 个氨基酸残基，或至少 170 个氨基酸残基，或至少 175 个氨基酸残基。

等位变异体指占据相同染色体位点的基因的两种或多种可选形式中的任何一个。等位变异通过突变自然发生，可能在群体中导致多态。基因突  
5 变可以是沉默的(在所编码的多肽中没有变化)或者可以编码具有改变的氨基酸序列的多肽。多肽的等位变异体是由基因的等位变异体编码的多肽。

本发明也涉及具有蛋白酶活性的分离的多肽，其由在极低的、或低的、  
或中等的、或者中等-高的(medium-high)、或高的、或极高的严谨条件下与  
核酸探针杂交的核酸序列编码，该探针在相同条件下与(a)SEQID NO: 1 的  
10 1-1065 位，优选 502-1065 位核苷，和/或 SEQID NO: 5 的 1-1143，优选  
568-1143 位核苷；(b)a 的子序列，或者(c) (a)或(b)的互补链(J. Sambrook, E. F. Fritsch, and T. Maniatis, 1989, Molecular Cloning, A Laboratory Manual, 2nd  
edition, Cold Spring Harbor, New York)杂交。在一个特定的实施方案中，核  
酸探针选自上述(a), (b), 或(c)的核酸序列。

15 上述(a)下面提到的核苷子序列可以为至少 100 个核苷，或者在另一实  
施方案中为至少 200 个核苷。另外，该子序列(subsequence)可以编码具有蛋  
白酶活性的多肽片段。

上述(a)的核酸序列，或者其子序列，以及 SEQ ID NO: 2 或 6 的氨基酸  
序列的相应部分，或其片段，可以根据本领域熟知的方法，用来设计核酸  
20 探针以鉴定并克隆来自不同种属株系的编码具有蛋白酶活性的多肽的  
DNA。特别地，可以依照标准的 Southern 印迹法用这样的探针与感兴趣的  
种属的基因组 DNA 或 cDNA 杂交，以鉴定并分离其中的相应基因。这样的  
探针可以比全序列短得多，但至少应当为 15 个，优选至少 25 个，更优选  
至少 35 个核苷长度。也可以使用更长的探针。DNA 和 RNA 探针均可使用。  
25 典型地，探针被标记用于探测相应的基因(例如，用  $^{32}\text{P}$ ,  $^3\text{H}$ ,  $^{35}\text{S}$ , 生物素或抗  
生物素蛋白)。本发明也包含这样的探针。

如此，可以筛选从这样的其他生物制备的基因组 DNA 或 cDNA 文库以  
得到与上述探针杂交的 DNA，其编码具有蛋白酶活性的多肽。可以用琼脂  
糖或聚丙烯酰胺凝胶电泳或其他分离技术分离来自这样的其他生物的基  
30 因组 DNA 或者其他的 DNA。可以将来自所述文库的 DNA 或者分离的 DNA  
转移并固定在硝化纤维或其他适当的载体材料上。为了鉴定与 SEQ ID NO:

1 或 5 中任一个, 或其子序列同源的克隆或 DNA, 在 Southern 印迹中使用所述载体材料。为了本发明的目的, 杂交指在极低的至极高的严谨条件下, 将核酸序列与相应于 SEQ ID NO: 1 或 5 中任一个的核酸序列, 其互补链, 或其子序列的标记的核酸探针杂交。用 X-射线胶片探测在这些条件下核酸

5 探针所杂交的分子。

对于至少 100 个核苷长度的长探针, 极低的至极高的严谨条件定义为 42°C 下, 在 5X SSPE, 0.3% SDS, 200 $\mu$ g/ml 经剪切的变性鲑精 DNA, 以及对于极低的和低的严格度为 25% 的甲酰胺, 对于中等的和中等高的严格度为 35% 的甲酰胺, 或者对于极高的严格度为 50% 的甲酰胺中, 按照标准的

10 Southern 印迹法进行预杂交和杂交。

对于至少 100 个核苷长度的长探针, 最后用 0.2  $\times$  SSC, 0.2% SDS, 20% 甲酰胺冲洗载体材料三次, 每次 15 分钟, 优选在至少 45°C 下冲洗(极低严格度), 更优选在至少 50°C 下冲洗(低严格度), 更优选在至少 55°C 下冲洗(中等严格度), 更优选在至少 60°C 下冲洗(中等 - 高严格度), 再优选在至少 65°C

15 下冲洗(高严格度), 最优选在至少 70°C 下冲洗(极高严格度)。

对于约 15 个至约 70 个核苷长度的短探针, 严谨条件定义为, 在低于依照 Bolton 和 McCarthy 的计算方法(1962, Proceedings of the National Academy of Sciences USA48 : 1390)所计算的 T<sub>m</sub> 值 5°C 至 10°C 条件下, 在 0.9 M NaCl, 0.09 M Tris-HCl pH 7.6, 6 mM EDTA, 0.5% NP-40, 1X Denhardt's

20 溶液, 1 mM 焦磷酸钠, 1 mM 单碱磷酸钠, 0.1 mM ATP, 以及每毫升 0.2 mg 酵母 RNA 中, 按照标准的 Southern 印迹法进行预杂交, 杂交和杂交后冲洗。

对于约 15 个至约 70 个核苷长度的短探针, 低于所计算的 T<sub>m</sub> 值 5°C 至 10°C 条件下, 在 6X SSC 加 0.1% SDS 中冲洗载体材料一次, 15 分钟, 用 6X SSC 冲洗两次, 每次 15 分钟。

25 本发明还涉及具有 SEQ ID NO: 2 和 6 中任一个的成熟部分氨基酸序列的多肽的变体, 其包含一个或多个氨基酸的取代、删除、和/或插入。

变异多肽的氨基酸序列可以由于一个或多个氨基酸残基的插入或删除和/或用不同的氨基酸残基取代一个或多个氨基酸残基, 而与 SEQ ID NO: 2 和 6 中任一个的成熟部分的氨基酸序列不同。在特定实施方案中, 氨基酸

30 变化是具有较小的属性变化, 也就是没有明显影响到蛋白质的折叠和/或活性的保守性氨基酸取代; 是小的删除, 典型地为一个至约 30 个氨基酸; 是

小的氨基或羧基端延伸,例如氨基端的蛋氨酸,直至约 20-25 个残基的小肽;或者是小的延伸,此小的延伸通过改变净电荷或其他属性以方便纯化,如多组氨酸片段,抗原表位或者结合域。

保守性取代的例子发生在碱性氨基酸(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),和小氨酸(甘氨酸、丙氨酸、丝氨酸、苏氨酸和甲硫氨酸)组内。相应地,例如,本发明涉及具有或包含 SEQ ID NO: 2 或 6 中任一一个所列出的序列,优选其成熟部分的多肽,或其活性片段,其中保守性氨基酸取代包含在碱性氨基酸(精氨酸、赖氨酸和组氨酸)中、酸性氨基酸(谷氨酸和天冬氨酸)中、极性氨基酸(谷氨酰胺和天冬酰胺)中、疏水氨基酸(亮氨酸、异亮氨酸和缬氨酸)中、芳香族氨基酸(苯丙氨酸、色氨酸和酪氨酸)中、和小氨酸(甘氨酸、丙氨酸、丝氨酸、苏氨酸和甲硫氨酸)中的置换,一个取代另一个,或其任意组合。

通常不改变特异活性的氨基酸取代在本领域是已知的,例如 H. Neurath and R. L. Hill, 1979,在 *The Proteins*, Academic Press, New York 中所描述的。最常发生的改变为 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 以及相反的改变。

在特定实施方案中,本发明的以及用于本发明所述用途的多肽是酸稳定的。为此目的,术语酸稳定的指在 pH 2.0、pH 2.5 或 pH 3.0 及 37°C 下保温 2 小时后,与 pH 9.0 及 5°C 下保温 2 小时后的相应样品相比,其残余活性至少为 50%。在特定实施方案中,残余活性至少为 60%, 70%, 80% 或 90%。测定酸稳定性的合适的测试方法为实施方案 2 的 pH 稳定性测试方法。

在特定实施方案中,本发明的以及用于本发明所述用途的多肽在 pH 7.0 条件下,具有至少 0.10, 0.15, 0.20, 0.25, 0.30 或 0.35 的相对活性。实施方案 2 的 pH 谱测试方法用于本测定。

在另外的特定实施方案中,本发明的以及用于本发明所述用途的多肽具有 i) 在 60°C 及 pH 9 的条件下,至少 0.05, 0.10, 0.15 或 0.20 的相对活性;和/或 ii) 70°C 时至少 0.40, 0.50 或 0.56 的相对活性。实施方案 2 的温度谱测试用于这些测定。

在另外的特定实施方案中，本发明的以及用于本发明所述用途的多肽在用 DSC 测定时，具有至少 78°C、或者至少 79、80、81、82 或 83°C 的 T<sub>m</sub>。在 pH 7.0 条件下按实施方案 2 所述测定 T<sub>m</sub>。

5 本发明的以及用于本发明所述用途的多肽可以是细菌或真菌多肽。真菌多肽可以源于丝状真菌或酵母。

在特定实施方案中，本发明的多肽是 i) 细菌蛋白酶；ii) 放线菌门的蛋白酶；iii) 放线菌纲的；iv) 放线菌目的；v) 拟诺卡氏菌科的；vi) 拟诺卡氏菌属的；和/或源于 vii) 拟诺卡氏菌如 *Nocardioopsis dassonvillei*、*Nocardioopsis alkaliphila*、南极拟诺卡氏菌、*Nocardioopsis prasina*、*Nocardioopsis composta*、  
10 *Nocardioopsis exhalans*、*Nocardioopsis halophila*、*Nocardioopsis halotolerans*、*Nocardioopsis kunsanensis*、*Nocardioopsis listeri*、*Nocardioopsis lucentensis*、*Nocardioopsis metallicus*、*Nocardioopsis synnemataformans*、*Nocardioopsis trehalosi*，热带拟诺卡氏菌、*Nocardioopsis umidischolae*、新疆拟诺卡氏菌、或 *Nocardioopsis alba* 的蛋白酶，例如源于 *Nocardioopsis alba* DSM 15647 的蛋  
15 白酶，例如具有 SEQ ID NO: 2 的 1 至 188 或 -167 至 188 位氨基酸序列的多肽，或者源于 *Nocardioopsis dassonvillei* 亚种 *dassonvillei* DSM 4235 的多肽，如具有 SEQ ID NO: 6 的 1-192 或 -160 至 192 位氨基酸序列的多肽。在特定实施方案中，蛋白酶源于 *Nocardioopsis alba*。

上述分类法依据 G. M. Garrity & J. G. Holt 的 *Bergey's Manual of Systematic Bacteriology*, 2001, second edition, volume 1, David R. Boone, Richard W. Castenholz 中 The road map to the Manual 一章。

对于前述物种，应当理解本发明同时包含理想和不理想状况的、以及其他的分类等同物，例如，变形，而不考虑其已知的种名。本领域技术人员可以很容易识别近似等同物的同一性。

25 公众可以很容易在几个培养物保藏中心得到这些物种的株系，例如美国典型培养物保藏中心(American Type Culture Collection ,ATCC)、德国微生物与细胞培养物保藏中心(Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, DSM)、霉菌培养物中心局(Centraalbureau Voor Schimmelcultures, CBS)和 Agricultural Research Service Patent Culture  
30 Collection, Northern Regional Research Center, NRRL)。例如公众能够由 DSMZ(德国微生物与细胞培养物保藏中心, Braunschweig, 德国)获得

*Nocardiopsis dassonvillei* 亚种 *dassonvillei* DSM 43235。该株系也保藏在其他的保藏机构如：ATCC 23219、IMRU 1250、NCTC 10489。

此外，也可以用上述的探针从其他来源包括从自然界(例如，土壤、堆肥(composts)、水等)分离的微生物鉴定并获得这样的多肽。从自然的生活环境分离微生物的技术是本领域熟知的。然后可以类似地通过筛选另一微生物的基因组或 cDNA 文库来得到核酸序列。一旦用探针探测到编码多肽的核酸序列，就可以利用本领域普通技术人员已知的技术(参见，例如，Sambrook 等，1989，如上)分离或克隆该序列。

此处定义的“分离的”或“纯化的”多肽是基本上独立于其他多肽的多肽，例如通过 SDS-PAGE 测定为至少约 20%纯化的，优选至少约 40%纯化的，更优选约 60%纯化的，再优选约 80%纯化的，最优选约 90%纯化的，甚至最优选约 95%纯化的多肽。

本发明的核酸序列编码的多肽也包括融合的多肽或可裂解的融合多肽，其中将另一多肽融合于所述多肽或其片段的 N 端或 C 端。通过将编码另一多肽的核酸序列(或其部分)与本发明的核酸序列(或其部分)融合来生产融合多肽。生产融合多肽的技术是本领域已知的，例如，PCR，或者将编码多肽的编码序列置于框中，并且将融合多肽的表达置于相同启动子和终止子的控制之下。

在其他的特定实施方案中，本发明不包括来自(i)WO 88/03947 中公开的 *Nocardiopsis dassonvillei* NRRL 18133; (ii) JP 2- 255081-A 中公开的拟诺卡氏菌株系 OPC-210 (FERM P-10508); (iii) DD 20043218 中公开的 *Nocardiopsis dassonvillei* 的株系 ZIMET 43647; (iv) JP 2003284571 中公开的拟诺卡氏菌 TOA-1(FERM-P-18676); 和/或(v)相应 DNA 的一种或多种蛋白酶。

## 25 核酸序列

本发明还涉及编码本发明多肽的分离的核酸序列。本发明特定的核酸序列为 SEQ ID NO: 1 的 502-1065 或 1-1065 位核苷，以及 SEQ ID NO: 5 的 1-1143，优选 568-1143 核苷，其中 SEQ ID NO: 1 的 502-1065 位核苷相应于成熟肽编码区。本发明还包含编码具有 SEQ ID NO: 2 的-167 至 188，优选 1-188 位，或者 SEQ ID NO: 6 的-160 至 192，优选 1-192 位氨基酸序列的核酸序列，其由于遗传密码的简并性而与 SEQ ID NO: 1 的相应部分不同。本

发明还涉及 SEQ ID NO: 1 或 5 的子序列,其分别编码具有蛋白酶活性的 SEQ ID NO: 2 和 6 的片段。

SEQ ID NO: 1 或 5 的子序列包含于 SEQ ID NO: 1 或 5 中任一个,只是 5'和/或 3'端的一个或多个核苷被删除。优选地,子序列包含至少 225 个核苷,更优选至少 300 个核苷,更加优选至少 375、450、500、531、600、700、800、900 或 1000 个核苷。

本发明还涉及与(i) SEQ ID NO: 1 的 502-1065 位核苷,和/或 SEQ ID NO: 5 的 568-1143 位核苷,或者与(ii) SEQ ID NO: 1 的 1-1065 位核苷,和/或 SEQ ID NO: 5 的 88-1143 位核苷具有至少 74%同一性的核苷序列。在特定实施方案中,与(i)或(ii)任一个的核苷的同一性为至少 74%,75%,76%,77%,78%,79%,80%,81%,82%,83%,84%,85%,86%,87%,88%,89%,90%,91%,92%,93%,94%,95%,96%,97%,98%,或 99%。在其他的特定实施方案中,与(i)或(ii)任一个的核苷的同一性为至少 60%,61%,62%,63%,64%,65%,66%,67%,68%,69%,70%,71%,72%,或 73%。

本发明还涉及在 SEQ ID NO: 1 或 5 中,优选在其成熟肽编码部分中包含至少一个突变的突变核酸序列,其中突变核酸序列编码的多肽(i)由 SEQ ID NO: 2 的-167 至 188,优选 1-188 位,或者 SEQ ID NO: 6 的-160 至 192,优选 1-192 位氨基酸组成;或者(ii)是(i)中任一序列的变异体,其中变异体包含一个或多个氨基酸的取代、删除、和/或插入;或者(iii)是(i)中任一序列的等位变异体,或者(iv)是(i)中任一序列的片段。

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,其包括从基因组 DNA 分离,从 cDNA 制备,或其组合。可以使用例如已知的多聚酶链式反应法(PCR)或表达文库抗体筛选法,从这样的基因组 DNA 中克隆本发明的核酸序列,以探测具有共有结构特征的克隆的 DNA 片段。参见,例如 Innis 等,1990,PCR: A Guide to Methods and Application, Academic Press, New York。可以使用其他的核酸扩增方法例如连接酶链式反应法(ligase chain reaction, LCR)、ligated activated transcription (LAT) 和核酸基础序列扩增法(nucleic acid sequence-based amplification, NASBA)。可以从拟诺卡氏菌株系或另外的或相关的生物克隆该核酸序列,这可能是,例如,核酸序列的多肽编码区的等位或物种变异体。

此处术语“分离的核酸序列”指基本上独立于其他核酸序列的核酸序列，例如通过琼脂糖凝胶电泳测定为至少约 20%纯化的，优选至少约 40%纯化的，更优选约 60%纯化的，再优选约 80%纯化的，最优选约 90%纯化的核酸序列。例如，可以使用基因工程中标准的克隆方法获得，所述基因工程

5 用于将核酸序列从其天然位置重新定位于其将要被复制的不同的位置。所述克隆方法可能涉及将包含编码多肽的核酸序列的所需核酸片段切除并分离，将此片段插入载体分子，将重组载体整合至宿主细胞中，核酸序列的多拷贝或克隆将在该宿主细胞中复制。该核酸序列可以是基因组，cDNA，RNA，半合成的，合成来源的，或其任意组合。

10 本发明编码多肽的核酸序列的修饰对于与所述多肽基本类似的多肽的合成可能是必要的。术语与所述多肽“基本类似的”是指非天然发生形式的多肽。这些多肽在某些工程方面可能与从其天然来源分离的多肽不同，例如，在特异活性、热稳定性、最佳 pH、致敏性或类似方面有所差异的变异体。所述变异序列可以以作为 SEQ ID NO: 1 和/或 5 的多肽编码部分形式存在

15 的核酸序列，例如其子序列为基础来构建，和/或通过引入不产生具有其他氨基酸序列的多肽，但与宿主生物倾向于生产蛋白酶的密码利用相符的核苷取代来构建，或者通过引入可能产生不同氨基酸序列的核苷取代来构建。如需要有关核苷取代的一般说明，可参见，例如，Ford 等，1991, Protein Expression and Purification 2: 95-107。例如，可以如上所述制备低致敏性多

20 肽。

这样的取代可在对分子功能至关重要的区域之外进行，这样仍然产生具有活性的多肽，这对本领域熟练技术人员来说是很显然的。可以依照本领域已知的方法，例如定点突变或丙氨酸扫描突变(参见，例如 Cunningham and Wells, 1989, Science 244: 1081-1085)来鉴定涉及多肽活性的基本氨基酸

25 残基，并因此优选不取代涉及多肽活性的基本氨基酸残基，所述多肽由本发明的分离的核酸序列所编码。在后来的技术中，将突变引入分子中的每一个带正电的残基，测试所得突变分子的蛋白酶活性，以鉴定对于分子活性至关重要的氨基酸残基。也可以通过使用核磁共振分析、结晶学或光亲和标记方法(参见，例如，de Vos 等，1992, Science 255: 306-312; Smith 等，1992,

30 Journal of Molecular Biology 224: 899-904; Wlodaver 等，1992, FEBS Letters 309: 59-64)测定三维结构来测定底物-蛋白酶反应的位置。

本发明还涉及编码本发明多肽的分离的核酸序列，其在极低严谨条件下，优选在低严谨条件下，更优选在中等严谨条件下，更优选在中等-高严谨条件下，再优选在高严谨条件下，最优选在极高严谨条件下与核酸探针杂交，所述核酸探针在相同条件下与 SEQ ID NO: 1 和/或 5 核酸序列，或其互补链杂交；或此处定义的其等位变异体和子序列(Sambrook 等, 1989, 如上)。

本发明还涉及通过(a)在极低、低、中等、中等高、高、或极高严谨条件下将 DNA 与(i) SEQ ID NO: 1 的 1-1065, 优选 502-1065 位核苷, 或者 SEQ ID NO: 5 的 1-1143, 或 88-1143, 优选 568-1143 位核苷, (ii) (i)的子序列, 或者(iii) (i)或(ii)的互补链杂交; 和(b)分离核酸序列, 而生产的分离核酸序列。所述子序列优选为具有至少 100 个核苷的序列, 以使这样的序列所编码的多肽片段具有蛋白酶活性。

#### 生产突变核酸序列方法

本发明进一步涉及生产突变核酸序列的方法, 包括将至少一个突变引入 SEQ ID NO: 1 和/或 5 的成熟多肽编码序列, 或其子序列, 其中突变的核酸序列编码由 SEQ ID NO: 2 的-167 至 188, 优选 1-188 位氨基酸, 或者 SEQ ID NO: 6 的-160 至 192, 优选 1-192 位氨基酸; 或其具有蛋白酶活性的片段所组成的多肽。

可以使用本领域已知的任何方法, 通过定点突变向核酸序列引入突变, 用一个核苷交换另一个核苷。利用插入有感兴趣序列的超螺旋、双链 DNA 载体和包含所需突变的两个合成引物的方法是尤其有用的。与载体的相对链分别互补寡核苷酸引物, 在温度循环期间通过 Pfu DNA 多聚酶方式进行延伸。引物的组合使用产生了包含交叉缺口的质粒。温度循环之后, 用 DpnI 处理产物, DpnI 特异于甲基化和半甲基化的 DNA, 用于消化亲本 DNA 模板并选择包含突变的合成 DNA。也可使用其他本领域已知的方法。

#### 核酸构建体

本发明还涉及包含与一个或多个控制序列可操作相连的本发明的核酸序列的核酸构建体, 所述控制序列在与控制序列相容的条件下引导编码序

列在适当的宿主细胞中表达。表达应被理解为包括涉及多肽生产的任何步骤，其包括但不限于转录、转录后修饰、翻译、翻译后修饰和分泌。

5 此处“核酸构建体”定义为单链或双链的核酸分子，其分离于天然产生的基因，或者是经过修饰而包含了以自然界不存在的方式组合和邻接(juxtaposed)的核酸片段。当核酸构建体包含了所有本发明的编码序列表达所需的控制序列时，术语核酸构建体与术语表达盒是同义的。此处术语“编码序列”定义为直接特异于其蛋白质产物的氨基酸序列的核酸序列。通过位于mRNA 5'端开放读码框正上游的核糖体结合位置(原核生物)或ATG起始密码子(真核生物)，以及位于mRNA 3'端开放读码框正下游的转录终止子序  
10 列来确定编码序列的界限。编码序列可以包括但不限于DNA、cDNA、以及重组核酸序列。

可以以多种方法操作编码本发明多肽的分离的核酸序列，以供多肽表达之用。依据表达载体的情况，在核酸序列插入载体之前对其进行操作是需要或必要的。利用重组DNA方法修饰核酸序列的技术是本领域熟知的。

15 此处定义的术语“控制序列”包括表达本发明多肽所必须或有利的所有组件。每个控制序列对于编码多肽的核酸序列来说可以是天然的或者是外源的。这样的控制序列包括但不限于前导序列(leader)、聚腺苷酸化(polyadenylation)序列、前导肽(propeptide)序列、启动子、信号肽序列、和转录终止子。控制序列最少包括启动子、以及转录和翻译终止信号。可以  
20 为控制序列提供接头(linker)，目的是为了引入特定限制性位点以便于控制序列与编码多肽的核酸序列的编码区连接。此处术语“可操作相连”定义为一种结构，在这种结构中，将控制序列适当地置于与DNA序列的编码序列相关的位置，以使控制序列引导多肽的表达。

控制序列可以是合适的启动子序列，由宿主细胞所识别而使核酸序列  
25 表达的核酸序列。启动子序列包含介导多肽表达的转录控制序列。启动子可以是在所选择的宿主细胞中显示转录活性的任何核酸序列，包括突变的、截断的和杂交的启动子，其可由编码胞外或胞内多肽的，与宿主细胞同源或异源的基因获得。

引导本发明的核酸构建体在尤其是细菌宿主细胞中转录的合适的启动  
30 子的例子为：由大肠杆菌的lac操纵元、天蓝色链霉菌(*Streptomyces coelicolor*)的琼胶酶基因(dagA)、枯草芽孢杆菌(*Bacillus subtilis*)的果聚糖合酶

(levansucrase)基因(sacB), 地衣芽孢杆菌(*Bacillus licheniformis*)的 $\alpha$ -淀粉酶基因(amyL)、嗜热脂肪芽胞杆菌(*Bacillus stearothermophilus*)的麦芽糖淀粉酶基因(amyM)、解淀粉芽孢杆菌(*Bacillus amyloliquefaciens*)的 $\alpha$ -淀粉酶基因(amyQ)、地衣芽孢杆菌的青霉素酶基因(penP)、枯草芽孢杆菌的xylA和xylB基因以及原核生物的 $\beta$ -内酰胺酶基因(Villa-Kamaroff等, 1978, Proceedings of the National Academy of Sciences USA 75: 3727-3731)所获得的启动子, 还有tac启动子(DeBoer等, 1983, Proceedings of the National Academy of Sciences USA 80: 21-25)。Scientific American, 1980, 242: 74-94, "Useful proteins from recombinant bacteria"中和Sambrook等, 1989, 如上描述了更多的启动子。

10 引导本发明的核酸构建体在丝状真菌宿主细胞中转录的合适的启动子的例子为: 由米曲霉(*Aspergillus oryzae*) TAKA淀粉酶、米赫根毛霉(*Rhizomucor miehei*)天冬氨酸蛋白酶、黑曲霉(*Aspergillus niger*)中性 $\alpha$ -淀粉酶、黑曲霉酸稳定性 $\alpha$ -淀粉酶、黑曲霉或泡盛曲霉(*Aspergillus awamori*)葡萄糖淀粉酶(glaA)、米赫根毛霉脂肪酶、米曲霉碱性蛋白酶、米曲霉丙糖磷酸异构酶(triose phosphate isomerase)、构巢曲霉(*Aspergillus nidulans*)乙酰胺酶、和尖孢镰刀霉(*Fusarium oxysporum*)胰蛋白酶样蛋白酶基因获得的启动子(WO 96/00787)、还有NA2-tpi启动子(黑曲霉中性 $\alpha$ -淀粉酶与米曲霉丙糖磷酸异构酶基因的启动子的杂合体)、及其突变、截短以及杂交的启动子。

15 在酵母宿主中, 有用的启动子由啤酒酵母(*Saccharomyces cerevisiae*)烯醇酶(ENO-1)、啤酒酵母半乳糖激酶(GAL1)、啤酒酵母醇脱氢酶/甘油醛-3-磷酸脱氢酶(ADH2/GAP)以及啤酒酵母3-磷酸甘油酸激酶的基因获得。其他有用的酵母宿主细胞启动子描述于Romanos等, 1992, Yeast 8: 423-488。

20 控制序列也可以是适当的转录终止子序列, 由宿主细胞所识别而终止转录。终止子序列与编码多肽的核酸序列3'端可操作相连。任何在所选择的宿主细胞中发挥功能的终止子均可用于本发明。

用于丝状真菌宿主细胞的优选终止子由米曲霉TAKA淀粉酶、黑曲霉葡萄糖淀粉酶、构巢曲霉邻氨基苯甲酸合成酶、黑曲霉 $\alpha$ -葡萄糖苷酶、尖孢镰刀菌胰蛋白酶(trypsin)样蛋白酶的基因获得。

25 用于酵母宿主细胞的优选终止子由啤酒酵母烯醇酶、啤酒酵母细胞色素C(CYC1)和啤酒酵母甘油醛-3-磷酸脱氢酶的基因获得。Romanos等, 1992, 如上描述了其他用于酵母宿主细胞的有用终止子。

用于细菌宿主细胞例如芽孢杆菌的优选终止子为来自地衣芽孢杆菌的 $\alpha$ -淀粉酶基因(amyL)、嗜热脂肪芽胞杆菌的麦芽糖淀粉酶基因(amyM)、或者解淀粉芽孢杆菌的 $\alpha$ -淀粉酶基因(amyQ)的终止子。

5 控制序列也可以是前导序列，其是mRNA的非翻译区，但对宿主细胞的翻译是重要的。前导序列与编码多肽的核酸序列5'端可操作相连。任何在所选择的宿主细胞中发挥功能的前导序列均可用于本发明。

用于丝状真菌宿主细胞的优选前导序列由米曲霉TAKA淀粉酶、构巢曲霉丙糖磷酸异构酶基因获得。

10 用于酵母宿主细胞的合适的前导序列由啤酒酵母烯醇酶(ENO-1)、啤酒酵母3-磷酸甘油酸激酶、啤酒酵母 $\alpha$ -因子、和啤酒酵母醇脱氢酶/甘油醛-3-磷酸脱氢酶(ADH2/GAP)的基因获得。

15 控制序列也可以是聚腺苷酸化序列，其与核酸序列的3'端可操作相连，当转录时，其作为添加多聚腺苷残基的信号由宿主细胞所识别，以转录mRNA。任何在所选择的宿主细胞中发挥功能的聚腺苷酸化序列均可用于本发明。

用于丝状真菌宿主细胞的优选聚腺苷酸化序列由米曲霉TAKA淀粉酶、黑曲霉葡萄糖淀粉酶、构巢曲霉邻氨基苯甲酸合成酶、和尖孢镰刀霉胰蛋白酶样蛋白酶、黑曲霉 $\alpha$ -葡萄糖苷酶的基因获得。

20 Guo和Sherman, 1995, Molecular Cellular Biology 15: 5983-5990描述了对酵母宿主细胞有用的聚腺苷酸化序列。

25 控制序列也可以是信号肽编码区，其编码连接于多肽氨基端的氨基酸序列，引导所编码的多肽进入细胞分泌途径。核酸序列的编码序列5'端可以固有地包含信号肽编码区，该信号肽编码区在翻译阅读框中天然地与编码分泌多肽的编码区片段相连。可选地，编码区的5'端可以包含对编码序列来说为外源的信号肽编码区。编码序列本身不包括信号肽编码区时，就需要外源的信号肽编码区。可选地，外源信号肽编码区可以简单地替换本身的信号肽编码区以增强多肽的分泌。然而，引导表达的多肽进入宿主细胞分泌途径的任何信号肽编码区均可用于本发明。

30 用于细菌宿主细胞的有效信号肽编码区由芽孢杆菌NCIB 11837麦芽糖淀粉酶、嗜热脂肪芽胞杆菌 $\alpha$ -淀粉酶、地衣芽孢杆菌的枯草杆菌蛋白酶、地衣芽孢杆菌的 $\alpha$ -淀粉酶、嗜热脂肪芽胞杆菌的中性蛋白酶(nprT, nprS,

nprM)、和枯草芽孢杆菌的 *prsA* 的基因获得。Simonen 和 Palva, 1993, *Microbiological Reviews* 57: 109-137 中描述了更多的信号肽。

5 用于丝状真菌宿主细胞的有效信号肽编码区为由米曲霉 TAKA 淀粉酶、黑曲霉中性淀粉酶、黑曲霉葡萄糖淀粉酶、米赫根毛霉的天冬氨酸蛋白酶、*Humicola insolens* 的纤维素酶和 *Humicola lanuginosa* 的脂肪酶基因获得的信号肽编码区。

对酵母宿主细胞有用的信号肽由啤酒酵母  $\alpha$ -因子和啤酒酵母转化酶基因获得。Romanos 等, 1992, 如上描述了其他有用信号肽编码区。

10 控制序列也可以是编码位于多肽氨基端的氨基酸序列的前导肽编码区。已知所得的多肽称为前体酶(proenzyme)或前体多肽(propolypeptide)(有些时候也叫酶原(zymogen))。前体多肽通常是没有活性的, 可以由前体多肽通过前导肽催化或自体催化裂解转变为成熟的活性多肽。可以从枯草芽孢杆菌碱性蛋白酶(*aprE*)、枯草芽孢杆菌中性蛋白酶(*aprT*)、啤酒酵母  $\alpha$ -因子、米赫根毛霉天冬氨酸蛋白酶、嗜热黄霉(*Myceliophthora thermophila*)的漆酶  
15 (WO 95/33836) 基因获得前导肽编码区。

在优选实施方案中, 前导肽编码区为 SEQ ID NO: 1 的 1-501 位核苷, 或者 SEQ ID NO: 6 的 30-189 位核苷。

当多肽的氨基端同时存在信号肽和前导肽编码区时, 前导肽区位于多肽的氨基端, 信号肽区位于前导肽区的氨基端。

20 也可能需要添加允许调控宿主细胞生长相关多肽表达的调控序列。调控系统的例子为, 因应化学或物理刺激, 包括调控化合物的存在, 而使基因表达开放或关闭的调控系统。原核系统中的调控系统包括 *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, TAKA  $\alpha$ -淀粉酶启动子、黑曲霉葡萄糖淀粉酶启动子和米曲霉葡萄糖淀粉  
25 酶启动子可以作为调控序列。调控序列的其他例子为允许基因扩增的调控序列。在真核系统中, 其包括存在氨甲蝶呤时扩增的二氢叶酸还原酶基因, 以及存在重金属时扩增的金属硫蛋白基因。在这些情况下, 编码多肽的核酸序列将与调控序列可操作相连。

### 30 表达载体

本发明还涉及包含本发明的核酸序列、启动子、以及转录及翻译终止

5 信号的重组表达载体。上述的各种核酸和控制序列可连接在一起而得到重组表达载体，其可以包括一个或多个方便的限制性位点，以允许在这些位点插入或取代编码多肽的核酸序列。可选地，可以通过将核酸序列或包括该序列的核酸构建体插入用于表达的适当的表达载体来表达本发明的核酸序列。在构建表达载体时，将编码序列置于载体中，以将编码序列与适当的表达控制序列可操作相连。

10 重组表达载体可以是能够方便地用于重组DNA方法，并可引发核酸序列表达的任何载体(例如，质粒或病毒)。典型地，载体的选择依赖于载体与其将要被导入的宿主细胞的兼容性。该载体可以是线性的质粒或封闭的环状质粒。

15 该载体可以自主复制载体，例如，以染色体外实体存在的载体，其复制不依赖于染色体的复制，例如，质粒、染色体外成分、微小染色体、或者人工染色体。载体可以包含任何确保自我复制的方式。可选地，载体可以是这样的载体，当其被导入宿主细胞中时，其整合到基因组中，并与其所整合的染色体一起复制。另外，也可以使用同时包含将要导入宿主细胞基因组的全部DNA的单一载体或质粒，或者两个或多个载体或质粒，或者可以使用转座子。

20 本发明的载体优选包含一个或多个可选择的标记，其允许很容易地选择转化的细胞。可选择标记为基因，其提供生物杀虫或病毒抗性、抗重金属、对营养缺陷型的原养型及类似的产物。细菌的可选择标记的例子为来自枯草芽孢杆菌或地衣芽孢杆菌的*dal*基因。用于酵母宿主细胞的合适标记为ADE2, HIS3, LEU2, LYS2, MET3, TRP1, 和URA3。用于丝状真菌宿主细胞的可选择标记包括但不限于amdS(乙酰胺酶)、argB(鸟氨酸氨基甲酰转移酶)、bar(草丁膦乙酰转移酶)、hygB(潮霉素磷酸转移酶)、niaD(硝酸还原酶)、25 pyrG(乳清酸核苷-5'-磷酸脱羧酶)、sC(硫酸腺苷酰基转移酶)、trpC(邻氨基苯甲酸合成酶)，及其等同物。优选用于曲霉菌细胞的为构巢曲霉或者米曲霉的amdS和pyrG基因以及吸水链霉菌(*Streptomyces hygroscopicus*)的bar基因。

30 本发明的载体优选包含允许载体稳定地整合于宿主细胞基因组或者允许载体在细胞中不依赖于基因组而自主复制的成分。

为了整合于宿主细胞基因组，载体可能依赖于编码多肽的核酸序列，

或者其他任何用于通过同源和不同源重组将载体稳定地整合于基因组的载体成分。可选地，载体可以包含通过同源重组引导向宿主细胞基因组整合的额外的核酸序列。额外的核酸序列使载体能够整合于宿主细胞基因组中精确的染色体位置。为了增加在精确位点整合的可能性，整合元件应优选

5 包含足够数量的核苷酸，例如100至1,500个碱基对，优选400至1,500个碱基对，最优选800至1,500个碱基对，这些碱基对与相应的目标序列高度同源，以增加同源重组的概率。整合成分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，整合成分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至宿主细胞基因组。

10 为了自主复制，载体可以进一步包含使载体能够在所述的宿主细胞中自主复制的复制起源。细菌的复制起源的例子为允许在大肠杆菌中复制的质粒pBR322, pUC19, pACYC177,和pACYC184,和允许在芽孢杆菌中复制的质粒pUB110, pE194, pTA1060, 和pAMB1的复制起源。用于酵母宿主细胞的复制起源的例子为2 micron复制起源, ARS1, ARS4, ARS1与CEN3的组合,

15 和ARS4与CEN6的组合。复制起源可以是具有突变的复制起源, 该突变使其在宿主细胞中发挥温度敏感性功能(参见, 例如, Ehrlich, 1978, Proceedings of the National Academy of Sciences USA 75: 1433)。

可以将超过一个拷贝编码本发明的核酸序列插入宿主细胞以增强基因产物的生产。可以通过将至少一个额外的序列拷贝整合于宿主细胞基因组

20 中以增加核酸序列的拷贝数, 或者通过在包含扩增的可选择标记基因拷贝的细胞中, 使核酸序列包含可扩增的可选择标记基因, 并因此可以在存在适当的可选择试剂的情况下, 通过培养细胞来选择额外拷贝的核酸序列, 以用于增加核酸序列的拷贝数。

用于连接上述组分以构建本发明的重组表达载体的方法是本领域技术

25 人员所熟知的(参见, 例如, Sambrook等, 1989, 如上)。

也可以将蛋白酶与至少一种其他的感兴趣的酶共表达以用于动物饲料, 例如淀粉酶, 如 $\alpha$ -淀粉酶(EC 3.2.1.1), 肌醇六磷酸酶(EC 3.1.3.8或3.1.3.26); 木聚糖酶(EC 3.2.1.8); 半乳聚糖酶 (EC 3.2.1.89);  $\alpha$ -半乳糖苷酶 (EC 3.2.1.22); 蛋白酶(EC 3.4.-.-), 磷脂酶A1 (EC 3.1.1.32); 磷脂酶 A2 (EC

30 3.1.1.4); 溶血磷脂酶(EC 3.1.1.5); 磷脂酶C (3.1.4.3); 磷脂酶 D (EC 3.1.4.4); 和/或 $\beta$ -葡聚糖酶(EC 3.2.1.4或EC 3.2.1.6)。

可以从不同的载体，从一个载体，或者两种技术的组合来共表达酶。当使用不同的载体时，载体可以具有不同的可选择标记，以及不同的复制起源。当只使用一个载体时，可以从一个或多个启动子表达基因。如果在一个启动子(二或多顺反子)的调控下克隆，所克隆的基因的顺序可能影响蛋白质的表达水平。蛋白酶也可以作为融合蛋白表达，例如，在读码框中，编码蛋白酶的基因被融合至编码另一蛋白的基因。该蛋白可能是另一种酶或另一种酶的功能区。

### 宿主细胞

10 本发明还涉及包含本发明核酸序列的重组宿主细胞，其可方便地用于多肽的重组生产。将包含本发明核酸序列的载体导入宿主细胞，以使载体作为染色体成分或作为前述的自我复制的染色体外载体存在。术语“宿主细胞”包含与复制过程中发生突变的亲本细胞不同的亲本细胞的任何后裔。宿主细胞的选择在很大程度上将依赖于编码多肽的基因及其来源。

15 宿主细胞可以是单细胞微生物，例如，原核生物，或者是非单核微生物，例如，真核生物。

有用的单细胞为细菌细胞如革兰氏阳性细菌包括但不限于芽孢杆菌属细胞，或链霉菌属细胞，或乳酸菌细胞；或者为革兰氏阴性细菌例如大肠杆菌和假单胞菌(*Pseudomonas*)类的乳酸细菌，包括但不限于乳球菌属(*Lactococcus*)、乳杆菌属(*Lactobacillus*)、明串珠菌属(*Leuconostoc*)、链球菌属(*Streptococcus*)、片球菌属(*Pediococcus*)、肠球菌属(*Enterococcus*)的物种。

20 可以通过例如原生质体转化(参见，例如，Chang和Cohen, 1979, *Molecular General Genetics* 168: 111-115)、使用感受态细胞(参见，例如，Young和Spizizin, 1961, *Journal of Bacteriology* 81: 823-829，或者Dubnau和Davidoff-Abelson, 1971, *Journal of Molecular Biology* 56: 209-221)、电穿孔法(参见，例如，Shigekawa和Dower, 1988, *Biotechniques* 6: 742-751)、接合法(参见，例如，Koehler和Thorne, 1987, *Journal of Bacteriology* 169: 5771-5278)将载体导入细菌宿主细胞。

30 宿主细胞可以是真核细胞，例如非人动物细胞，昆虫细胞、植物细胞或者真菌细胞。

在特定实施方案中，宿主细胞是真菌细胞。此处所用“真菌(fungi)”包括

子囊菌门(Ascomycota)、担子菌门(Basidiomycota)、壶菌门(Chytridiomycota)、和接合菌门(Zygomycota) (如Hawksworth等, 在Ainsworth and Bisby's Dictionary of The Fungi, 8th edition, 1995, CAB International, University Press, Cambridge, UK中所定义的)、以及Oomycota(如Hawksworth等, 1995, 如上, 5 page 171所引用的)和所有的有丝分裂孢子真菌(Hawksworth等, 1995, 如上)。

在另一特定实施方案中, 真菌宿主细胞是酵母细胞。此处所用的“酵母”包括产子囊酵母(ascosporogenous) (内孢霉目(Endomycetales)), 产担子孢子囊酵母(basidiosporogenous), 和属于半知菌(Fungi Imperfecti) (芽孢纲(Blastomycetes))的酵母。由于酵母的分类将来可能会有所变化, 为了本发明的目的, 应该按照Biology and Activities of Yeast(Skinner, F.A., Passmore, 10 S.M., and Davenport, R.R., eds, Soc. App. Bacteriol. Symposium Series No. 9, 1980)中所述定义酵母。

酵母宿主细胞可以是假丝酵母(*Candida*)、汉逊酵母(*Hansenula*)、克鲁维酵母(*Kluyveromyces*)、毕赤酵母(*Pichia*)、酵母(*Saccharomyces*)、裂殖酵母 15 (*Schizosaccharomyces*)、或者亚罗威阿酵母(*Yarrowia*)细胞。

真菌宿主细胞可以是丝状真菌细胞。“丝状真菌”包括真菌门(Eumycota)和卵菌门(Oomycota) (如Hawksworth等, 1995, 以上所定义的)所有丝状形式的亚类。丝状真菌以由壳多糖、纤维素、葡聚糖、脱乙酰壳多糖、甘露聚糖、以及其他复合多糖组成的菌丝体壁为特征。营养生长依靠菌丝的延长, 20 碳分解是需氧的。相反, 酵母例如啤酒酵母的营养生长通过单细胞叶状体的萌发进行, 碳分解可以是发酵。

丝状真菌宿主细胞的例子是但不限于支顶孢霉(*Acremonium*), 曲霉, 镰刀霉(*Fusarium*), 腐质霉(*Humicola*), 毛霉(*Mucor*), 黄霉(*Myceliophthora*), 脉孢菌 25 (*Neurospora*), 青霉(*Penicillium*), 梭孢壳菌(*Thielavia*), *Tolypocladium*, 或木霉(*Trichoderma*)上述属物种的细胞。

可以通过涉及原本已知的原生质体形成、原生质体转化、以及细胞壁再生等方式的方法转化真菌细胞。EP 238 023和Yelton等, 1984, Proceedings of the National Academy of Sciences USA 81: 1470-1474中描述了转化曲霉宿主细胞的适当的方法。Malardier等, 1989, Gene 78: 147-156和WO 96/00787 30 中描述了转化镰刀霉物种的适当方法。可以用Becker and Guarente, In Abelson, J.N. and Simon, M.I., editors, Guide to Yeast Genetics and Molecular

Biology, Methods in Enzymology, Volume 194, pp 182-187, Academic Press, Inc., New York; Ito等, 1983, Journal of Bacteriology 153: 163; 和Hinnen等, 1978, Proceedings of the National Academy of Sciences USA 75: 1920中描述的方法转化酵母。

5

### 生产方法

本发明还涉及生产本发明的多肽的方法, 包含(a)培养菌株, 该菌株的野生型能够生产所述多肽; 和(b)回收多肽。优选该菌株属于拟诺卡氏菌属, 更优选为 *Nocardiosis prasina*, 南极拟诺卡氏菌, *Nocardiosis dassonvillei* 或  
10 *Nocardiosis alba*, 最优选为 *Nocardiosis alba* DSM 15647, 或者 *Nocardiosis dassonvillei* 亚种 *dassonvillei* DSM 43235。

本发明还涉及生产本发明的多肽的方法, 包含(a)在有利于多肽生产的条件下培养宿主细胞; 和(b)回收多肽。

本发明还涉及生产本发明的多肽的方法, 包含(a) 在有利于多肽生产的  
15 条件下培养宿主细胞, 其中宿主细胞包含突变的核酸序列, 该突变的核酸序列在SEQ ID NO: 1的502-1065或1-1065位核苷, 或者在SEQ ID NO: 5的1-1143、88-1143, 优选568-1143位核苷包含至少一个突变, 其中的突变核酸序列编码多肽, 该多肽(i)由SEQ ID NO: 2的1至188, 或者-167至188, 或者  
20 SEQ ID NO: 6的-160至192, 优选1-192氨基酸组成, 或者(ii)是(i)中任一序列的变异体, 其中该变异体包含一个或多个氨基酸的取代、删除、和/或插入, 或者(iii)是(i)中任一序列的等位变异体, 或者(iv)是(i)中任一序列的片段。

在本发明的生产方法中, 细胞培养于适合用本领域已知方法生产多肽的营养培养基中。例如, 可以在实验室或者工业发酵罐中, 在适当的培养基中以及允许多肽表达和/或分离的条件下, 通过摇瓶培养、小规模或大规模  
25 发酵(包括连续的、分批的、补料-分批的、或者固态发酵)来培养细胞。在包含碳源和氮源以及无机盐的适当营养培养基中以本领域已知的方法进行培养。合适的培养基可以从供应商处获得, 或者可以用公布的组合物(例如, 美国典型培养物保藏中心目录中的)制备。如果多肽分泌到营养培养基中, 则可以直接从培养基回收多肽。如果多肽没有分泌, 则可以通过细胞  
30 裂解来回收。

可以用本领域已知的专门用于所述多肽的方法探测多肽。这些探测方

法可能包括使用特异性抗体、形成产物、或者底物的消失。例如，可以用蛋白酶测试来测定此处所述多肽的活性。

5 可以用本领域已知的方法回收所得多肽。例如，可以通过传统的方法包括但不限于离心、过滤、提取、喷雾干燥、蒸发、或沉淀从营养培养基中回收多肽。

可以用本领域已知的方法包括但不限于层析(例如，离子交换层析、亲和层析、疏水层析、色谱聚焦层析、大小排阻层析)、电泳(例如，制备性等电聚焦电泳)、差别溶解度(differential solubility) (例如，硫酸铵沉淀)、SDS-PAGE、或者提取(参见，例如，Protein Purification, J.-C. Janson and Lars  
10 Ryden, editors, VCH Publishers, New York, 1989)来纯化本发明的多肽。

### 植物

本发明还涉及用编码本发明具有蛋白酶活性的多肽的核酸序列转化的转基因植物、植物部分或者植物细胞，其用于表达生产可回收量的多肽。

15 可以从植物或植物部分回收所述多肽。可选地，包含重组多肽的植物或植物部分可用于如改善食物或饲料的质量，例如，提高营养价值、改善口味和改善流变性状，或者用于破坏抗营养因子。

在特定实施方案中，所述多肽定位于种子的胚乳存储液泡中。可以通过用适当的信号肽以前体形式合成多肽来达到这种目的，参见Horvath等，  
20 PNAS, Feb. 15, 2000, vol. 97, no. 4, p. 1914-1919。

转基因植物可以是双子叶的(双子叶植物)或者单子叶的(单子叶植物)或其工程变异体。单子叶植物的例子为草，例如早熟禾(meadow grass, blue grass, *Poa*)，饲草如羊茅(*Festuca*)、黑麦草(*Lolium*)，temperate grass如剪股颖属(*Agrostis*)，和谷类例如小麦、燕麦、黑麦、大麦、水稻、高粱、黑小麦(triticale)(小麦(*Triticum*)和黑麦(*Secale*)的稳定化杂种)、和玉米(corn)。双  
25 子叶植物的例子是烟草，豆科植物例如羽扇豆(legume)、马铃薯、糖用甜菜、豌豆(pea)、菜豆(bean)和大豆(soybean)，和十字花科植物(family Brassicaceae)例如向日葵(*Helianthus*)、棉花 (*Gossypium*)、花椰菜、油菜籽、和非常相关的模式生物拟南芥(*Arabidopsis thaliana*)。例如US专利no. 5,689,054和US专  
30 利no. 6,111,168中描述的肌醇六磷酸低含量植物为工程植物的例子。

植物部分的例子是茎、愈伤组织、叶、根、果实、种子和块茎，还有

包含这些部分的单独的组织，例如表皮、叶肉、薄壁组织、维管组织、分生组织。特定的细胞区室例如叶绿体、质外体、线粒体、液泡、过氧化物酶体、和细胞质被认为是植物部分。另外，任何植物细胞，无论什么组织来源，都被认为是植物部分。同样，植物部分例如为方便本发明利用而分离的特定组织和细胞也被认为是植物部分，例如，胚、胚乳、糊粉(aleurone)、种皮。

这些植物、植物部分和植物细胞的后代也包括在本发明的范围之内。

可以按照本领域已知的方法构建表达本发明的多肽的转基因植物或植物细胞。简单地说，将一个或多个编码本发明多肽的表达构建体整合至植物宿主基因组，并将所得的经改造的植物或植物细胞繁殖为转基因植物或植物细胞，以此来构建植物或植物细胞。

方便地，所述表达构建体为核酸构建体，其包含与适当的调控序列可操作相连的编码本发明多肽的核酸序列，所述调控序列是在所选植物或植物细胞中表达核酸序列所需要的。另外，表达构建体可以包含可选择标记，该标记可用于鉴定整合了表达构建体的宿主细胞，其序列对于引导所述构建体进入所述植物是必须的(后者依赖于所使用的DNA导入方法)。

调控序列，例如启动子和终止子序列和可选的信号或转运序列的选择以例如需要何时、何地及如何表达所述多肽为基础来确定。例如，编码本发明多肽的基因的表达可以是组成性的或可诱导的，或者可以是发育、场所或组织特异的，基因产物可以定向至特定的组织或植物部分例如种子或叶子。例如，Tague等, 1988, *Plant Physiology* 86: 506描述了调控序列。

对于组成性表达，可以使用以下的启动子：35S-CaMV启动子(Franck等, 1980, *Cell* 21: 285-294)，玉米ubiquitin 1 (Christensen AH, Sharrock RA and Quail 1992. *Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation*)，或者水稻actin 1启动子(*Plant Mo. Biol.* 18, 675-689.; Zhang W, McElroy D. and Wu R 1991, *Analysis of rice Act1 5' region activity in transgenic rice plants. Plant Cell* 3, 1155-1165)。器官特异性启动子可以是，例如，来自贮藏组织如种子、马铃薯块茎、和果实的启动子(Edwards & Coruzzi, 1990, *Ann. Rev. Genet.* 24: 275-303)，或者来自代谢组织如分生组织的启动子(Ito等, 1994, *Plant Mol. Biol.* 24: 863-878)，可以是种子特异的启

5 动子如来自水稻的glutelin, prolamin, globulin, 和albumin的启动子(Wu等, 1998, Plant and Cell Physiology 39: 885-889), 来自legumin B4和蚕豆的未知种子蛋白基因的蚕豆启动子(Conrad等, 1998, Journal of Plant Physiology 152: 708-711), 来自种子油体蛋白的启动子(Chen等, 1998, Plant and Cell  
10 Physiology 39: 935-941), 来自甘蓝型油菜(*Brassica napus*)贮藏蛋白napA启动子, 或者本领域已知的任何其他种子特异性启动子, 例如WO 91/14772所描述的。另外, 启动子可以是叶特异性启动子, 例如来自水稻或番茄的rbcS启动子(Kyozuka等, 1993, Plant Physiology 102: 991-1000), 小球藻病毒腺嘌呤甲基转移酶基因启动子(Mitra and Higgins, 1994, Plant Molecular Biology  
15 26: 85-93), 或者水稻的aldP基因启动子(Kagaya等, 1995, Molecular and General Genetics 248: 668-674), 或者创伤诱导启动子例如马铃薯pin2启动子(Xu等, 1993, Plant Molecular Biology 22: 573-588)。同样, 启动子可以是能够由非生物处理如温度、干旱、盐度改变所诱导的, 或者是能够由激活启动子的外来物质如乙醇、雌激素、植物激素像乙烯、脱落酸、赤霉素、和/或  
20 重金属所诱导的。

也可以使用启动子增强子组件在植物中获得更高的蛋白酶表达。例如, 启动子增强子组件可以是位于启动子和编码本发明多肽的核苷序列之间的内含子。例如, Xu等, 1993, 以上公开了水稻actin 1基因的第一个内含子用于增强表达的用途。

20 更进一步, 可以优化密码子利用率以提升所述的植物物种的表达(参见, 以上Horvath等所述)。

可以选择本领域可获得的可选择标记基因和表达构建体的任何其他部分。

25 根据本领域已知的传统技术将核酸构建体整合至植物基因组, 包括农杆菌(*Agrobacterium*)介导的转化、病毒介导的转化、微注射、粒子轰击、基因枪法转化、和电穿孔法(Gasser等, 1990, Science 244: 1293; Potrykus, 1990, Bio/Technology 8: 535; Shimamoto等, 1989, Nature 338: 274)。

30 目前, 土壤农杆菌(*Agrobacterium tumefaciens*)介导的基因转移法是用于生产转基因双子叶植物的选择(for a review, see Hooykas and Schilperoort, 1992, Plant Molecular Biology 19: 15-38), 其也可以用于转化单子叶植物, 尽管其他的转化方法对于这些植物来说一般是优选的。目前, 选择用于生产

转基因单子叶植物的方法是对胚胎愈伤组织或发育中的胚胎(Christou, 1992, *Plant Journal* 2: 275-281; Shimamoto, 1994, *Current Opinion Biotechnology* 5: 158-162; Vasil等, 1992, *Bio/Technology* 10: 667-674)进行粒子轰击(转化DNA包被的显微金或钨粒子), 加农杆菌方法。用于单子叶植物  
5 转化的可选方法基于Omirulleh等, 1993, *Plant Molecular Biology* 21: 415-428 所述的原生质体转化。

转化之后, 根据本领域熟知的方法选择其中整合了表达构建体的转化体并使其再生为完整的植物。

本发明还涉及生产本发明的多肽的方法, 包含(a)在有利于生产所述多  
10 肽的条件下, 培养包含编码本发明多肽的核酸序列的转基因植物或植物细胞, 所述多肽具有蛋白酶活性; 和(b)回收所述多肽。

### 动物

本发明还涉及转基因非人动物及产品, 或其部分, 其例子为例如奶和  
15 血液、器官、肉、以及动物细胞。在例如哺乳动物细胞中表达蛋白质的技术是本领域已知的, 参见, 例如the handbook *Protein Expression: A Practical Approach*, Higgins and Hames (eds), Oxford University Press (1999), 以及涉及  
基因转录、RNA处理、翻译后处理的该系列的三本其他手册。一般来讲,  
为制备转基因动物, 用编码本发明具有蛋白酶活性的多肽的核酸序列转化  
20 所选动物的选定细胞来表达生产该多肽。可以从动物例如雌性动物的奶中  
回收多肽, 或者多肽的表达可能对动物自身是有利的, 例如帮助动物消化。  
在下面标题为动物饲料的部分提到了动物的例子。

为了生产转基因动物以从动物的奶中回收蛋白酶, 可以将编码蛋白酶的  
基因插入所述动物的受精卵, 例如通过使用含有适当的奶蛋白启动子和  
25 编码蛋白酶的基因的转基因表达载体来实现。将转基因表达载体微注射到  
受精卵, 优选永久整合至染色体。一旦卵开始生长分裂, 就将潜在的胚胎  
植入代孕母亲, 并鉴定携带有转化基因的动物。可以通过传统的饲养方法  
繁殖所得的动物。可以从动物的奶中纯化多肽, 参见, 例如Meade, H.M. 等  
(1999): *Expression of recombinant proteins in the milk of transgenic animals*,  
30 *Gene expression systems: Using nature for the art of expression*. J. M. Fernandez  
and J. P. Hoeffler (eds.), Academic Press.

在可选方案中，为了生产在其体细胞和/或生殖细胞基因组中携带包括异源转基因构建体的核酸序列的转基因非人动物，所述转基因构建体包括编码蛋白酶的转基因，可以如WO 2000064247所公开的，将该转基因与用于蛋白酶唾腺特异性表达的第一个调控序列可操作相连。

5

### 组合物

在另一方面，本发明涉及包含本发明多肽的组合物。

可以依照本领域已知的方法制备多肽组合物，该多肽组合物可以是液体形式的或干燥形式的组合物。例如，多肽组合物可以是颗粒或微小颗粒形式。可以依据本领域已知的方法把要包括在该组合物中的多肽稳定化。

10

下面给出了本发明的多肽或多肽组合物的优选用途的例子。

### 动物饲料

本发明还涉及在动物饲料中使用本发明多肽的方法，也涉及包含本发明多肽的饲料组合物和饲料添加剂。

15

术语动物包括所有的动物，包括人。动物的例子为非反刍动物和反刍动物。反刍动物包括例如绵羊、山羊、马、和牛例如肉牛、母牛、小牛犊。在特定实施方案中，动物是非反刍动物。非反刍动物包括单胃动物例如猪(pigs或swine)(包括但不限于猪仔、生长猪、和大母猪); 家禽例如火鸡、鸭、鸡(包括但不限于肉鸡、蛋鸡); young calves; 和鱼(包括但不限于鲑鱼、鳟鱼、罗非鱼、鲶鱼、鲤鱼); 和甲壳类(包括但不限于虾(shrimp)、对虾)。

20

术语饲料或者饲料组合物指任何适合或者希望动物摄入的化合物、制剂、混合物或者组合物。

在本发明的用途中，可以在食物之前、之后饲喂动物，或者与食物同时饲喂动物。后者是优选的。

25

在特定的实施方案中，蛋白酶添加到饲料中或者包括于饲料添加剂中的形式是明确的。“明确的”(well-defined)是指通过大小排阻层析测定，蛋白酶制剂的纯度至少为50%。在其他的特定实施方案中，以此方法测定的蛋白酶制剂纯度至少为60, 70, 80, 85, 88, 90, 92, 94,或95%。

30

明确的蛋白酶制剂是有利的。例如，向饲料中添加正确剂量的蛋白酶而基本不受其他蛋白酶的干扰或污染是非常容易地。术语添加正确剂量尤

其是指获得一致或恒定结果的目的，和基于所需效果而优化剂量的能力。

然而，用于动物饲料用途时，蛋白酶没有必要是那样的纯度；例如它可以包括其他的酶，这种情况下其可以称为蛋白酶制剂。

蛋白酶制剂可以(a)直接添加到饲料中(直接用于蛋白质的处理过程)，或者(b)可以用于一种或多种中间组合物如饲料添加剂或者预混料的生产，随后添加到饲料中(或者用于处理过程中)。上述的纯度指最初的蛋白酶制剂的纯度，无论其是用于上述的(a)或是(b)中。

特别是具有这种纯度数量级的蛋白酶制剂是可以使用重组生产的方法获得的，而使用传统的发酵方法生产时，是不能够如此容易获得的，并且有非常高的批量(batch-to-batch)变异。

当然，这样的蛋白酶制剂可以与其他的酶混合。

在特定实施方案中，用于本发明的用途的蛋白酶能够溶解蛋白质。实施方案5中公开了测定溶解蛋白质的合适的测试方法。

蛋白质可以是动物蛋白质，例如肉和骨粉，和/或鱼粉；或者可以是植物蛋白质。此处所用术语植物蛋白质指包括得自或源于植物的至少一种蛋白质的任何化合物、组合物、制剂或混合物，包括改造的蛋白质和蛋白质衍生物。在特定实施方案中，植物蛋白质的蛋白质含量至少为10, 20, 30, 40, 50,或60% (w/w)。

植物蛋白质可以得自植物蛋白质来源，例如豆类和谷类，例如豆科(*Fabaceae*, *Leguminosae*)、十字花科(*Cruciferaceae*)、藜科(*Chenopodiaceae*)和*Poaceae*植物家族的原料，例如大豆粉、羽扇豆(lupin)粉和油菜籽粉。

在特定实施方案中，植物蛋白质来源是一种或多种豆科植物例如大豆、羽扇豆、豌豆、或豆(bean)的原料。

在另外的特定实施方案中，植物蛋白质来源是一种或多种藜科植物例如甜菜、糖用甜菜、菠菜或奎奴亚藜(quinoa)的原料。

其他植物蛋白质来源的例子是油菜籽、向日葵籽、棉花籽和甘蓝(cabbage)。

大豆是优选的植物蛋白质来源。

其他的植物蛋白质来源的例子是谷类例如大麦、小麦、黑麦、燕麦、玉米(corn)、水稻、黑小麦和高粱。

根据本发明用至少一种本发明的蛋白酶对蛋白质的处理使蛋白质的溶

解增强。

下面是在单胃动物体外模型中，能够用本发明的蛋白酶得到的溶解蛋白质%的例子：相对于空白为至少101%、或者至少102%，103%，104%，105%，106%，107%。用实施方案5的单胃动物体外模型测定溶解蛋白质的百分数。

5 术语蛋白质的溶解意思基本上指使蛋白质成为溶液。这样的溶解可以归因于蛋白质从常见复合天然组合物如饲料的其他成分中的释放，这种释放由蛋白酶介导。

在另外的特定实施方案中，用于根据本发明的用途中的蛋白酶能够增加可消化蛋白质的量。下面是在单胃动物体外模型中，能够用本发明的蛋白酶得到的已消化的或可消化的蛋白质%的例子：相对于空白来说，至少为101%，或者至少为102%。用实施方案5的体外模型测定已消化的或可消化的蛋白质的百分数。

15 在其他的特定实施方案中，用于本发明的用途中的蛋白酶能够提高蛋白质的水解程度(DH)。在特定实施方案中，水解程度相对于空白至少为101%，102%，103%，104%，或者105%。水解程度由实施方案5的体外模型测定。

20 在本发明(预)处理方法的特定实施方案中，所述的蛋白酶影响(或作用于、或发挥其溶解作用)蛋白质或蛋白质来源。为达到此目的，典型地，将蛋白质或蛋白质来源悬浮于溶剂，例如水溶剂如水，并且根据所述酶的特征调整pH和温度。例如处理可以在当前蛋白酶能够发挥其至少40%，50%，60%，70%，80%或90%活性的pH值下进行。同样，例如处理可以在当前蛋白酶能够发挥其至少40%，50%，60%，70%，80%或90%活性的温度下进行。上述的活性百分数是相对于最大活性而言的。使酶反应持续至得到所需结果为止，此后可以通过将酶灭活，如通过热处理步骤来终止酶反应，也可不终止。

25 在本发明处理方法的特定实施方案中，蛋白酶作用是缓释的，意思是，例如，蛋白酶加进了蛋白质或蛋白质来源，但其溶解作用可以说直至以后需要时才打开，此时建立了适当的溶解条件，或者灭活了任何的酶抑制剂，或者取消了已用于延迟酶作用的任何其他方式。

30 在一个实施方案中，所述处理是对动物饲料的或用于动物饲料的蛋白质的预处理。

术语改善动物饲料的营养价值意思是改善蛋白质的可用性和/或可消化性，由此使来自食物成分的蛋白质提取物增加，得到更高的蛋白质产量，蛋白质降解增强和/或蛋白质利用率提高。因此提高了饲料的营养价值，动物的性状如生长率和/或增重量(weight gain)和/或饲料转化率(例如，摄取的  
5 饲料与增重量的对比)得到提高。

在特定实施方案中，饲料转化率提高至少1%，2%，3%，4%，5%，6%，7%，8%，9% 或10%。在另外的特定实施方案中，增重量提高至少2%，3%，4%，5%，6%，7%，8%，9%，10% 或11%。这些数字相对于没有添加蛋白酶的对照实验。

10 可以如EEC (1986) :Directive de la Commission du 9 avril 1986 fixant la méthode de calcul de la valeur énergétique des aliments composés destinés à la volaille. Journal Officiel des Communautés Européennes, L130, 53 – 54中所述计算饲料转化率(FCR)和增重量。

可以将蛋白酶以任何形式加入饲料，可以是相对纯的蛋白酶，或者是  
15 与其他要添加到动物饲料的成分一起的混合物，例如，以饲料添加剂的形式，例如所谓的动物饲料预混料。

另一方面，本发明涉及用于动物饲料的组合物，例如动物饲料，和动物饲料添加剂，例如预混料。

除了本发明的蛋白酶，本发明的动物饲料添加剂还包含至少一种脂溶性维生素，和/或至少一种水溶性维生素，和/或至少一种微量矿质元素。饲料添加剂也可以包含至少一种大量矿质元素。  
20

另外，可选地，饲料添加剂成分为色素制剂，例如，类胡萝卜素如β-胡萝卜素、虾青素(astaxanthin)、叶黄素(lutein)；芳香化合物；稳定剂；抗菌肽；多聚不饱和脂肪酸；活性氧生成物质(species)；和/或选自淀粉酶，例如α-淀粉酶(EC 3.2.1.1)，肌醇六磷酸酶 (EC 3.1.3.8或3.1.3.26)；木聚糖酶  
25 (EC 3.2.1.8)；半乳聚糖酶(EC 3.2.1.89)；α-半乳糖苷酶(EC 3.2.1.22)；蛋白酶(EC 3.4.-.-)，磷脂酶A1 (EC 3.1.1.32)；磷脂酶 A2 (EC 3.1.1.4)；溶血磷脂酶(EC 3.1.1.5)；磷脂酶 C (3.1.4.3)；磷脂酶 D (EC 3.1.4.4)；β-葡聚糖酶 (EC 3.2.1.4或EC 3.2.1.6)的至少一种其他酶。

30 在特定实施方案中，这些其他酶是明确的(如上述针对蛋白酶制剂所定义的)。

抗菌肽的例子(AMP's)为CAP18、Leucocin A、Tritrpticin、Protegrin-1、死亡肽(Thanstin)、防卫肽(Defensin)、乳铁蛋白(Lactoferrin)、乳铁蛋白肽(Lactoferricin)、和Ovispirin例如Novispirin (Robert Lehrer, 2000)、Plectasins、和Statins, 包括WO 03/044049和WO 03/048148中公开的化合物和多肽, 以及保留抗菌活性的上述肽的变异体或片段。

抗真菌多肽(AFP's)的例子为WO 94/01459和WO 02/090384中所公开的*Aspergillus giganteus*和黑曲霉肽, 以及其保留抗真菌活性的变异体和片段。

多聚不饱和脂肪酸的例子为C18, C20和C22多聚不饱和脂肪酸, 例如花生四烯酸、二十二碳六烯酸(docosohexaenoic acid)、二十碳五烯酸和 $\gamma$ -亚麻酸。

活性氧生成物质的例子是化学药品例如过硼酸盐(perborate)、过硫酸盐(persulphate)、或者过碳酸盐(percarbonate); 和酶例如氧化酶、加氧酶或者合酶(synthase)。

常见脂溶性和水溶性维生素, 以及微量矿质元素构成欲往饲料中添加的所谓预混料的一部分, 而大量矿质元素通常单独添加到饲料中。富含本发明的蛋白酶的预混料是本发明的动物饲料添加剂的例子。

在特定的实施方案中, 要添加到动物食物或饲料中的本发明的动物饲料添加剂在0.01至10.0%水平, 更特别地在0.05至5.0%; 或0.2至1.0%(%指g添加剂每100 g饲料)。在预混料中尤其如此。

下面是这些成分实例的非排除性清单:

脂溶性维生素的例子为维生素A、维生素D3、维生素E、和维生素K, 例如维生素K3。

水溶性维生素的例子为维生素B12、生物素和胆碱、维生素B1、维生素B2、维生素B6、烟酸、叶酸和泛酸盐(panthothenate)例如Ca-D-泛酸。

微量矿质元素的例子为锰、锌、铁、铜、碘、硒和钴。

大量矿质元素的例子为钙、磷和钠。

WO 01/58275的表A列出了对这些成分的营养需求(以家禽和仔猪/猪为例)。营养需求意思是应当在食物中提供指定浓度的这些成分。

可选地, 本发明的动物饲料添加剂包含WO 01/58275的表A中列出的单一成分中的至少一种。至少一种指两者之一、一种或多种、一种、或两种、或三种、或四种等等, 直至所有13种, 或者直至所有15种单一成分。更特

别地，这种至少一种单一成分在本发明添加剂中的量提供表A第四栏、或第五栏、或第六栏所示范围之内的饲料中浓度。

本发明还涉及动物饲料组合物。动物饲料组合物或食物具有相关高的蛋白质含量。家禽和猪食物可以WO 01/58275的表B第2-3栏所示为特征。鱼食物可以表B第4栏所示为特征。另外，这样的鱼食物通常具有200-310 g/kg的粗脂肪含量。WO 01/58275相应于US 09/77334，此处一并作为参考。

根据本发明的动物饲料组合物含有50-800 g/kg的粗蛋白，另外还包括至少一种此处所要求的蛋白酶。

此外，或者可选地(对于上述指定的粗蛋白含量)，本发明的动物饲料组合物含有10-30 MJ/kg的可代谢能量；和/或0.1-200 g/kg的钙；和/或0.1-200 g/kg的可用磷；和/或0.1-100 g/kg甲硫氨酸；和/或0.1-150 g/kg的甲硫氨酸加半胱氨酸；和/或0.5-50 g/kg的赖氨酸。

在特定实施方案中，可代谢能量、粗蛋白、钙、磷、甲硫氨酸、甲硫氨酸加半胱氨酸、和/或赖氨酸的含量在WO 01/58275 (R. 2-5)的表B中2, 3, 4或5中任意一个范围内。

以氮(N)乘以因数6.25来计算粗蛋白，例如，粗蛋白(g/kg)= N (g/kg) x 6.25。以Kjeldahl法(A.O.A.C., 1984, Official Methods of Analysis 14th ed., Association of Official Analytical Chemists, Washington DC)测定氮含量。

可以以NRC publication Nutrient requirements in swine, ninth revised edition 1988, subcommittee on swine nutrition, committee on animal nutrition, board of agriculture, national research council. National Academy Press, Washington, D.C., pp. 2-6, and the European Table of Energy Values for Poultry Feed-stuffs, Spelderholt centre for poultry research and extension, 7361 DA Beekbergen, The Netherlands. Grafisch bedrijf Ponsen & looijen bv, Wageningen. ISBN 90-71463-12-5为基础计算可代谢能量。

以饲料表例如 Veevoedertabel 1997, gegevens over chemische samenstelling, verteerbaarheid en voederwaarde van voedermiddelen, Central Veevoederbureau, Runderweg 6, 8219 pk Lelystad. ISBN 90-72839-13-7为基础计算完全动物食物中钙、可用磷和氨基酸的含量。

在特定实施方案中，本发明的动物饲料组合物至少包含一种上述定义的蛋白质或蛋白质来源。也可以包含动物蛋白，例如肉和骨粉，和/或鱼粉，

典型地，用量为0-25%。

在另外的特定实施方案中，本发明的动物饲料组合物包含0-80%的玉米；和/或0-80%的高粱；和/或0-70%的小麦；和/或0-70%的大麦；和/或0-30%的燕麦；和/或0-40%的大豆粉；和/或0-25%的鱼粉；和/或0-25%的肉和骨粉；  
5 和/或0-20%的乳清。

动物食物可以制成例如糊状饲料(非颗粒状)或粒状(pelleted)饲料。典型地，混合磨碎的饲料-原料，按照对所述物种的说明加入足够量的基本维生素和矿质元素。酶可以以固体或液体酶制剂的方式加入。例如，固体酶制剂典型地在混合步骤之前或混合步骤中加入；液体酶制剂典型地以制粒步骤之后加入。也可以将酶加至饲料添加剂或者预混料中。  
10

食物中的最终酶浓度在0.01-200 mg酶蛋白每kg食物的范围内，例如在0.5-25 mg酶蛋白每kg动物食物的范围内。当然，要以有效量使用蛋白酶，例如，足够增强饲料的溶解性和/或提高其营养价值的用量。目前考虑以一个或多个下述的用量(剂量范围)加入酶：0.01-200；0.01-100；0.5-100；1-50；  
15 5-100；10-100；0.05-50；或者0.10-10 - 所有这些范围均为mg蛋白酶蛋白每kg饲料(ppm)。

为了测定mg酶蛋白每kg饲料，从饲料组合物中纯化蛋白酶，用相关性测试(参见下面的蛋白酶活性，底物，和测试)测定纯化蛋白酶的特异活性。这样的饲料组合物的蛋白酶活性也用同样的测试方法测定，以这两个测定  
20 为基础，来计算以mg酶蛋白每kg饲料计的剂量。

使用同样的方法来测定饲料添加剂中的mg酶蛋白。当然，如果可以获得用于制备饲料添加剂或者饲料的蛋白酶，就由该样品测定特异活性(无需从饲料组合物或添加剂中纯化蛋白酶)。

下面的实施方案进一步描述本发明，不应将其解释为对本发明的范围  
25 的限定。

### 洗涤剂组合物

本发明的蛋白酶可以添加到洗涤剂组合物中，这样就成为洗涤剂组合物的成分。

30 本发明的洗涤剂组合物可以制成例如手洗或机洗洗涤剂组合物，其包括适合于玷污织物预处理的洗涤添加剂组合物和漂洗加入的织物软化剂组

合物，或者可以制成用于常规家用物品硬表面清洗操作的洗涤剂组合物，或者可以制成用于手或机器洗碗盘操作的洗涤剂组合物。

在特定方面，本发明提供了包含本发明的蛋白酶的洗涤剂添加剂。洗涤剂添加剂以及洗涤剂组合物可以包含一种或多种其他的酶，例如另一种  
5 蛋白酶，如源于芽孢杆菌的碱性蛋白酶、脂肪酶、角质酶、淀粉酶、糖酶、纤维素酶、果胶酶、甘露聚糖酶、阿拉伯聚糖酶、半乳聚糖酶、木聚糖酶、氧化酶，例如漆酶、和/或过氧化物酶。

一般来讲，所选择的酶的特性应与选择的洗涤剂相容，(例如，最适pH，与其他酶学和非酶学组分的相容性等)，并且酶应以有效量存在。

10 合适的脂肪酶包括细菌或真菌来源的脂肪酶。包括化学修饰的或蛋白质工程化变异体。有用的脂肪酶的例子包括来自腐质霉(同义词 *Thermomyces*)的脂肪酶，例如EP 258068和EP 305216所述的来自胎毛腐质霉 (*H. lanuginosa*) (*T. lanuginosus*)或者WO 96/13580所述的来自 *H. insolens*的脂肪酶，假单胞菌脂肪酶，例如来自产碱假单胞菌(*P. alcaligenes*)或者类产碱  
15 假单胞菌(*P. pseudoalcaligenes*) (EP 218272)、*P. cepacia* (EP 331376)、施氏假单胞菌 (*Pseudomonas stutzeri*) (GB 1,372,034)、荧光假单胞菌 (*P. fluorescens*)、假单胞菌株系SD 705 (WO 95/06720 和WO 96/27002)、*P. wisconsinensis* (WO 96/12012)的脂肪酶，芽孢杆菌脂肪酶，例如来自枯草芽孢杆菌(Dartois等，(1993), *Biochemica et Biophysica Acta*, 1131, 253-360)、嗜  
20 热脂肪芽孢杆菌(JP 64/744992)或者短小芽孢杆菌(*B. pumilus*) (WO 91/16422)的脂肪酶。其他的例子为例如WO 92/05249, WO 94/01541, EP 407225, EP 260105, WO 95/35381, WO 96/00292, WO 95/30744, WO 94/25578, WO 95/14783, WO 95/22615, WO 97/04079 和WO 97/07202中描述的脂肪酶变异体。商业上可获得的优选脂肪酶为Lipolase<sup>TM</sup> 和Lipolase  
25 Ultra<sup>TM</sup> (Novozymes A/S)。

合适的淀粉酶( $\alpha$ -和/或 $\beta$ -)包括细菌或真菌来源的淀粉酶。包括化学修饰的或者蛋白质工程化变异体。淀粉酶包括，例如，由芽孢杆菌例如地衣芽孢杆菌的特定株系获得的 $\alpha$ -淀粉酶，GB 1,296,839中有更为详细的描述。有用的淀粉酶的例子为WO 94/02597, WO 94/18314, WO 95/26397, WO  
30 96/23873, WO 97/43424, WO 00/60060, 和WO 01/66712中描述的变异体，尤其是在下述位置具有一个或多个取代的变异体：15, 23, 105, 106, 124, 128,

133, 154, 156, 181, 188, 190, 197, 202, 208, 209, 243, 264, 304, 305, 391, 408, 和 444。商业上可获得的淀粉酶为Natalase<sup>TM</sup>, Supramyl<sup>TM</sup>, Stainzyme<sup>TM</sup>, Duramyl<sup>TM</sup>, Termamyl<sup>TM</sup>, Fungamyl<sup>TM</sup>和BAN<sup>TM</sup> (Novozymes A/S), Rapidase<sup>TM</sup> 和Purastar<sup>TM</sup> (来自Genencor International Inc.)。

- 5 合适的纤维素酶包括细菌或真菌来源的纤维素酶。包括化学修饰的或者蛋白质工程化变异体。合适的纤维素酶包括来自芽孢杆菌属、假单胞菌属、腐质霉属、镰刀霉属、梭孢壳属(*Thielavia*)、顶孢霉属的纤维素酶, 例如US 4,435,307, US 5,648,263, US 5,691,178, US 5,776,757和WO 89/09259中公开的从*Humicola insolens*、嗜热黄霉(*Myceliophthora thermophila*)和尖孢镰刀霉(*Fusarium oxysporum*)生产的真菌纤维素酶。尤其适合的纤维素酶是具有颜色养护(care)作用的碱性或中性纤维素酶。EP 0 495257, EP 531372, WO 96/11262, WO 96/29397, WO 98/08940中公开了这样的纤维素酶的例子。其他的例子为纤维素酶变异体, 例如WO 94/07998, EP 0 531 315, US 5,457,046, US 5,686,593, US 5,763,254, WO 95/24471, WO 98/12307和WO 99/01544中
- 10 所描述的。商业上可获得的纤维素酶包括Celluzyme<sup>TM</sup>, 和Carezyme<sup>TM</sup> (Novozymes A/S), Clazinase<sup>TM</sup>, 和Puradax HA<sup>TM</sup> (Genencor International Inc.), 以及KAC-500(B)<sup>TM</sup> (Kao Corporation)。

合适的过氧化物酶/氧化酶包括植物、细菌或真菌来源的过氧化物酶/氧化酶。包括化学修饰的或者蛋白质工程化变异体。有用的过氧化物酶的例子包括WO 93/24618, WO 95/10602, 和WO 98/15257中所描述的来自鬼伞(*Coprinus*), 例如灰盖鬼伞(*C. cinereus*)的过氧化物酶及其变异体。商业上可获得的过氧化物酶包括Guardzyme<sup>TM</sup> (Novozymes)。

20

可以通过添加包含一种或多种酶的单独的添加剂, 或者通过添加包含所有这些酶的组合添加剂将洗涤剂酶加入洗涤剂组合物中。本发明的洗涤剂添加剂, 例如单独的或组合的添加剂, 可以制成例如颗粒、液体、浆等。优选的洗涤剂添加剂制剂(formulations)为颗粒尤其是无粉尘颗粒、液体尤其是稳定化液体、或者浆。

25

可以如US 4,106,991和4,661,452中所述生产无粉尘颗粒, 可选地, 用本领域已知的方法包被。蜡样的包被材料的例子为具有1000至20000平均摩尔重量的聚氧化乙烯产品(聚乙二醇, PEG); 具有16至50个环氧乙烷单位的乙氧基化壬基苯酚; 乙氧基化脂肪醇, 其中醇包含12至20个碳原子, 且其具

30

有15至80个环氧乙烷单位；脂肪醇；脂肪酸；和脂肪酸的单甘油酯、甘油二酯以及甘油三酸酯。GB 1483591给出了适合于流化床技术使用的膜形成包被材料。可以根据已建立的方法通过例如加入多羟基化合物如丙二醇、糖或糖醇、乳酸或者硼酸将液态酶制品稳定化。可以根据EP 238216中公开的方法制备被保护的酶。

本发明的洗涤剂组合物可以为任何方便的形式，例如，条、药片、粉末、颗粒、糊或液体。液体洗涤剂可以是水化的，典型地包含直至70%的水以及0-30%的有机溶剂，或者是非水化的。

洗涤剂组合物包含一种或多种的表面活性剂，其可以是不带离子的包括半极性的和/或阴离子的和/或阳离子的和/或两性离子的。表面活性剂典型地以0.1%至60%的重量水平存在。

当其中包含阴离子表面活性剂时，通常洗涤剂将包含约1%至约40%的阴离子表面活性剂，例如烷基苯磺酸盐、 $\alpha$ -烯烴磺酸盐(olefinsulfonate)、烷基磺酸盐(脂肪酸磺酸盐)、醇乙氧基硫酸酯、二级烷基磺酸(secondary alkanesulfonate)、 $\alpha$ -磺基脂肪酸甲酯、烷基或链烯基琥珀酸或者肥皂。

当其中包含不带离子的表面活性剂时，通常洗涤剂将包含约0.2%至约40%的不带离子的表面活性剂，例如醇乙氧基酯类化合物、壬基苯酚乙氧基酯类化合物(nonylphenol ethoxylate)、烷基糖苷(alkylpolyglycoside)、烷基二甲基氨基氧化物(alkyldimethylamineoxide)、乙氧基脂肪酸单乙醇胺、脂肪酸单乙醇胺、多羟基烷基脂肪酸胺、或者葡糖胺的N-酰基 N-烷基衍生物("glucamides")。

洗涤剂可以包含0-65 %的洗涤剂助洗剂或者复合试剂如沸石、二磷酸盐、三磷酸盐、磷酸盐、碳酸盐、柠檬酸盐、次氨基三乙酸、乙二胺四乙酸、二乙烯三胺五乙酸、烷基或链烯基琥珀酸、可溶性硅酸盐或层状硅酸盐(例如，Hoechst的SKS-6)。

洗涤剂可以包含一种或多种多聚体。例子为羧甲基纤维素、多聚乙烯吡咯烷酮、聚乙二醇、聚乙烯醇、聚乙烯基吡啶N氧化物(poly(vinylpyridine-N-oxide))、聚乙烯基咪唑、聚羧酸酯如聚丙烯酸酯、马来酸/丙烯酸共聚物和月桂酸甲基丙烯酸酯/丙烯酸共聚物。

洗涤剂可以包含漂白系统，漂白系统可以包含 $H_2O_2$ 源例如过硼酸盐或者过碳酸盐，其可以与过酸形成漂白激活剂例如四乙酰乙二胺或壬酰氧苯

磺酸盐(nonanoyloxybenzenesulfonate)组合。可选地,漂白系统可以包含例如氨、亚胺、硼类型的过氧酸。

5 可以用传统的稳定化试剂来使本发明的洗涤剂组合物的酶稳定化,稳定化试剂为例如,多羟基化合物如丙二醇或丙三醇、糖或糖醇、乳酸、硼酸,或者为硼酸衍生物,例如芳族硼酸酯、或苯基硼酸衍生物如4-甲酰基苯基硼酸,可以按照例如WO 92/19709和 WO 92/19708所述方法制备组合物。

洗涤剂也可以包含其他的传统洗涤剂成分,例如织物护理剂,包括粘土(clays)、增泡剂、抑泡剂、防腐剂、污垢悬浮剂、防污垢再沉积剂、染料、杀菌剂、荧光增白剂、助水溶剂(hydrotropes)、防失泽剂、或香料。

10 目前,预期可以以相当于0.01-100 mg酶蛋白每升洗涤水,优选0.05-5 mg酶蛋白每升洗涤水,尤其是0.1-1 mg酶蛋白每升洗涤水的量向洗涤剂组合物中加入任何的酶,尤其是本发明的酶。

另外,本发明的酶可以加至WO 97/07202所公开的洗涤剂配方中。

### 15 生物材料的保藏

下述由2001年在丹麦收集的土壤样品中分离的生物材料,已经按照布达佩斯条约的条款保藏于德国微生物与细胞培养物保藏中心, Mascheroder Weg 1b, D-38124 Braunschweig, 并给出下面的保藏号:

	保藏物	保藏号	保藏日期
20	<i>Nocardiopsis alba</i>	DSM 15647	May 30, 2003

25 该株系的保藏保证了在本专利申请未决期间,由专利与商标专员依据37 C.F.R. §1.14和35 U.S.C. §122所授权的人能够得到该株系。该保藏物为所保藏株系的基本上纯化的培养物。此保藏物可以在提交了这些申请的副本或其后续文本的国家依据该外国专利法的要求获得。然而,应当理解,保藏物的获得并不构成对实施本发明的许可,实施本发明将侵犯由政府行为所授予的专利权。

30 由于本发明的实施方案是要作为本发明几个方面的示例,因而本发明此处的描述和权利要求并不是要以此处公开的特定实施方案来限定其范围。任何等同的实施方案将被认为在本发明的范围之内。事实上,对本发明除此处所列出和描述的之外,通过前面的描述,很多的修改对于本领域技术人员来讲都将是很明显的。这种修改也将落在所附的权利要求的范围

之内。对于有冲突的情况，将由本发明的说明书包括定义来解释。

此处引用的多个参考文献，其公开的内容整体一并作为参考。

### 实施例

5

**实施例1: 源于白色拟诺卡菌DSM 15647的蛋白酶的克隆与表达**

#### 试剂与培养基

- LB 琼脂 Ausubel, F. M. 等(eds.) "Current protocols in Molecular Biology". John Wiley and Sons, 1995中所述
- 10 LB-PG琼脂 补充了0.5%葡萄糖和0.05 M 磷酸钾, pH 7.0的LB琼脂
- PS-1 10% 蔗糖, 4% 大豆粉, 1%  $\text{Na}_3\text{PO}_4 \cdot 12\text{H}_2\text{O}$ , 0.5%  $\text{CaCO}_3$ , 和0.01% pluronic acid
- TE 10 mM Tris-HCl, pH 7.4  
1 mM EDTA, pH 8.0
- 15 TEL TE-缓冲液中50 mg/ml的Lysozym
- 硫氰酸盐 5M 硫氰酸胍(guanidium thiocyanate)  
100 mM EDTA  
0.6 % w/v N-月桂醇肌氨酸, 钠盐
- 20 60 g 硫氰酸盐, 20 ml 0.5 M EDTA, pH 8.0, 20 ml  $\text{H}_2\text{O}$ 在65°C溶解。冷却至室温(RT), 加入0.6 g N-月桂醇肌氨酸。加水至100 ml, 用0.2  $\mu$ 消毒的过滤器过滤。
- $\text{NH}_4$  Ac 7.5 M  $\text{CH}_3\text{COONH}_4$
- TER TE-buffer缓冲液中1  $\mu\text{g}/\text{ml}$ 的Rnase A
- CIA 氯仿/异戊醇 24:1

25

#### 试验程序

SEQ ID NO: 1是编码来自Nocardiopsis alba DSM 15647的蛋白酶前体形式的DNA序列。502-1065位核苷对应于成熟肽编码部分。

- 30 SEQ ID NO: 2是由SEQ ID NO: 1推导出来的氨基酸序列。-167至-1位氨基酸是前肽, 氨基酸1至188是成熟肽。

SEQ ID NO: 1克隆

野生型在收获前于下述培养基中30°C生长3天:

- 5      Trypticase            20 g  
       酵母提取物        5 g  
       氯化铁            6 mg  
       硫酸镁            15 mg  
       蒸馏水加至        1000 ml  
       通过加入碳酸钠将pH调至9
- 10     按照以下方法分离Nocardiosis alba DSM 15647的基因组DNA:  
       收获1.5 ml培养物, 重悬于100 µl TEL。37°C保温30 min。  
       加入500 µl硫氰酸盐缓冲液于室温放置10 min。  
       加入250 µl NH<sub>4</sub>Ac置于冰上10 min。  
       加入500 µl CIA, 混合。
- 15     转至微型离心管离心10 min。  
       将上清转至新的Eppendorf管, 加入0.54体积的冰冻异丙醇。彻底混合。  
       用70 % EtOH旋转冲洗DNA颗粒。  
       于100 µl TER中重悬基因组DNA。  
       基因组DNA用作PCR扩增的模板, 以下述的SEQ ID NO. 3和4为引物。
- 20     在0.7%琼脂糖凝胶上分离PCR片段。

引物:

- 1421: 5'-GTT CAT CGA TCG CAT CGG CTG CGA CCG GCC CCC TCC  
 CCC AGT C-3' (SEQ ID NO: 3)
- 25     1604: 5'- GCG GAT CCT ATC AGG TGC GCA GGG TCA GAC C-3'  
       (SEQ ID NO: 4)

- 30     将消化与纯化的PCR片段连接于Cla I和BamH I消化的质粒  
       pDG268NeoMCS-PrmyQ/PreryIII/cryIIIAstab/Sav (US Patent No. 5,955,310)  
       上。

将连接所得的混合物转化至*E. coli* TOP10F' (Invitrogen BV, The

Netherlands), 选择若干克隆用于微量制备(miniprep)(QIAprep spin, QIAGEN GmbH, Germany)。转化前, 检查纯化的质粒, 用破坏的apr、npr和pel基因(Diderichsen等, (1990), J. Bacteriol., 172, 4315-4321)将其插入源于枯草芽孢杆菌DN 1885的枯草芽孢杆菌株系。基因的破坏基本上按照"*Bacillus subtilis* and other Gram-Positive Bacteria," American Society for Microbiology, p.618, eds. A.L. Sonenshein, J.A. Hoch and Richard Losick (1993)所述方法进行。将转化的细胞在添加了6 µg/ml 氯霉素的1% 脱脂乳LB-PG 琼脂板上划平板。划平板的细胞于37°C保温过夜, 通过周围的空白区(clearing)鉴定包含蛋白酶的克隆。选择蛋白酶阳性克隆, 通过DNA序列分析从表达构建体中确定所表达的酶的编码序列。

### 发酵

将上述转化的枯草芽孢杆菌宿主细胞置于500 ml带隔板的Erlenmeyer瓶中, Erlenmeyer瓶中含有添加了6 µg/ml 氯霉素的100 ml PS-1培养基, 将其放在旋转摇床(250 r.p.m.)上, 37°C发酵16小时, 再于26°C额外发酵4天。

### **实施例2: 来自Nocardiopsis alba DSM 15647的蛋白酶的纯化与鉴定**

#### 蛋白酶测试

##### 1) pNA测试:

20 pNA 底物: Suc-AAPF-pNA (Bachem L-1400).

温度: 室温(25°C)

测试缓冲液: 100mM 琥珀酸, 100mM HEPES, 100mM CHES, 100mM CABS, 1mM CaCl<sub>2</sub>, 150mM KCl, 0.01% Triton X-100 用HCl或NaOH将pH-值调整至2.0, 2.5, 3.0, 3.5, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0, 11.0, 和12.0。

25 将20µl蛋白酶(稀释于0.01% Triton X-100中)与100µl测试缓冲液混合。加入100µl pNA底物(50mg溶解于1.0ml DMSO, 进一步用0.01% Triton X-100稀释45x)以开始测试。监测OD<sub>405</sub>的增强, 作为蛋白酶活性的测量值。

##### 2) Protazyme AK测试:

30 底物: Protazyme AK 药片 (交联并染色的酪蛋白;来自Megazyme)

温度: 受控的温度(测试温度)

测试缓冲液: 100mM琥珀酸, 100mM HEPES, 100mM CHES, 100mM CABS, 1mM CaCl<sub>2</sub>, 150mM KCl, 0.01% Triton X-100用HCl或NaOH将pH-值调整至2.0, 2.5, 3.0, 3.5, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0和11.0。

轻微搅动, 使Protazyme AK 药片悬浮于2.0ml 0.01% Triton X-100中。

- 5 在Eppendorf管中将500μl的此悬浮液与500μl测试缓冲液混合, 置于冰上。加入20μl蛋白酶样品(稀释于0.01% Triton X-100中)。将Eppendorf管移至Eppendorf热混合器以开始测试, 其中Eppendorf热混合器设定在测试温度。在Eppendorf热混合器的最高摇动速率(1400 rpm)将Eppendorf管保温15分钟。将Eppendorf管转至冰浴, 以终止保温。然后Eppendorf管在冰冻离心机中离心几分钟, 将200μl的上层转至微量滴定盘(microtiter plate)。OD<sub>650</sub>处读数作为蛋白酶活性测量值。测试中包括空白缓冲液(buffer blind)(代替酶)。

- 离心(20000 x g, 20 min)实施例1中的蛋白酶发酵物, 从沉淀物上小心移走上清。用Seitz EKS盘过滤合并的上清, 以除去残留的拟诺卡氏菌细胞。将EKS滤液转至G25 sephadex柱上的50mM H<sub>3</sub>BO<sub>3</sub>, 5mM 琥珀酸, 1mM CaCl<sub>2</sub>, pH 为7的溶液, 得到混浊的溶液。再次用Seitz EKS盘过滤以除去混浊物。将澄清的滤液置于相同缓冲液平衡过的杆菌肽(bacitracin)硅胶柱。用平衡缓冲液大致冲洗柱子后, 用100mM H<sub>3</sub>BO<sub>3</sub>, 10mM琥珀酸, 2mM CaCl<sub>2</sub>, 1M NaCl, 25% 异丙醇, pH 7的溶液逐步洗脱蛋白酶。将杆菌肽洗脱液转至G25 sephadex柱上的50mM H<sub>3</sub>BO<sub>3</sub>, 5mM琥珀酸, 1mM CaCl<sub>2</sub>, pH 为7的溶液, 在配备有5000Da临界点膜的Amicon浓缩室中超滤浓缩至最小体积。将浓缩的酶置于在100mM H<sub>3</sub>BO<sub>3</sub>, 10mM 琥珀酸, 2mM CaCl<sub>2</sub>, 200mM NaCl, pH为7的溶液中平衡过的Superdex 75大小排阻层析柱上, 用相同的缓冲液洗脱。用柱子上的组分分析蛋白酶活性(37°C, pH 9条件下用Protazyme AK测试分析), 用SDS-PAGE进一步分析活性组分。将在考马斯染色的SDS-PAGE胶上仅看到一个条带的组分取下来, 作为纯化的制品, 用于进一步的鉴定。

#### pH-活性, pH-稳定性, 以及温度-活性

- 用pNA测试获得pH-活性谱以及pH-稳定性谱。对pH-稳定性谱, 将蛋白酶在测试缓冲液中稀释10 x, 于37°C保温2小时。保温后, 在测试残余活性之前, 将蛋白酶样品稀释于pH 9的测试缓冲液中, 使之转变为相同的pH-pH 9。

Protazyme AK测试用于获得pH 9时的温度-活性谱。结果列于下面的表1-3。

表 1: pH-活性谱

pH	源于 <i>Nocardiosis alba</i> DSM 15647的蛋白酶	源于 <i>Nocardiosis sp.</i> NRRL 18262的蛋白酶
2	0.00	-
3	0.00	0.00
4	0.01	0.02
5	0.06	0.07
6	0.18	0.21
7	0.37	0.44
8	0.69	0.67
9	0.99	0.88
10	1.00	1.00
11	0.95	0.93

5

表 2: pH-稳定性谱

pH	源于 <i>Nocardiosis alba</i> DSM 15647 的蛋白酶	源于 <i>Nocardiosis</i> <i>sp. NRRL 18262</i> 的蛋白 酶
2.0	1.04	0.78
2.5	1.05	1.00
3.0	1.00	1.03
3.5	1.00	0.98
4.0	0.93	0.99
5.0	1.00	1.02
6.0	1.00	1.00
7.0	1.00	1.01
8.0	1.03	0.98
9.0	1.02	0.99
10.0	0.96	0.99
11.0	0.95	0.86
12.0	0.88	-
9.0, 5 °C 2 小时后	1.00	1.00

表 3: 温度活性谱

温度 (°C)	源于 <i>Nocardiosis alba</i> DSM 15647的蛋白酶	源于 <i>Nocardiosis sp.</i> NRRL 18262的蛋白酶
15	0.02	0.02
25	0.05	0.02
37	0.10	0.07
50	0.27	0.20
60	0.56	0.51
70	1.00	1.00
80	0.49	0.39
90	-	-

发现蛋白酶被苯基甲基磺酰氟所抑制。其相对分子量经SDS-PAGE测定为 $M_r = 19\text{kDa}$ 。

5

#### 差示扫描量热法(DSC)

DSC用于测定pH 7.0时源于*Nocardiosis alba* 和*Nocardiosis sp.* NRRL 18262的蛋白酶的温度稳定性。纯化的蛋白酶置于10 mM 磷酸钠, 50 mM 氯化钠, pH 7.0的溶液中4°C透析过夜, 在VP-DSC 装置(Micro Cal)上以1.5°C/min的恒定扫描速率从20至100°C测定。用MicroCal Origin软件处理数据。

10

所得的变性或解链温度 $T_m$ 为: 对于本发明源于*Nocardiosis alba* 的蛋白酶为78.3°C; 对于源于*Nocardiosis sp.* NRRL 18262的蛋白酶为76.5°C。

#### 15 **实施例3: 源于*Nocardiosis alba* 的蛋白酶的特异活性**

用实施例2所述的纯化蛋白酶制品测定特异活性。SDS-PAGE(按照WO 01/58275中实施例2A所述方法测定)分析显示制品的纯度在95%以上。将蛋白酶样品一分为二。一部分通过氨基酸分析来分析其蛋白质含量, 另一部分用于分析蛋白酶活性。

20

#### 氨基酸分析(AAA)/(mg/ml)

用酸水解蛋白酶样品的肽键，然后在Biochrom 20 Plus氨基酸分析仪上根据生产商的说明分离并量化释放的氨基酸，该分析仪在商业上可从Bie & Berntsen A/S, Sandbaekvej 5-7, DK-2610 Roedovre, Denmark获得。将蛋白质样品于真空离心机中干燥，溶解于18.5% (体积/体积) HCl + 0.1% (体积/体积) 苯酚中，110°C保温16小时，用于酸水解。保温之后，样品于真空离心机中再次干燥，溶解于上样缓冲液(0.2 M 柠檬酸钠, pH 2.2)，上样到Biochrom 20 Plus氨基酸分析仪。

为了进行量化，将水解的样品上样到钠形式的阳离子交换树脂UltroPac no. 8上，该柱子可通过商业途径从Bie & Berntsen A/S获得，商品目录号为no. 80-2104-15。根据生产商的上述说明，使改变pH(pH 1至pH 8)和离子强度的缓冲液通过柱子，分离不同的氨基酸。同样根据生产商的说明，精确地控制柱子温度(从 53°C到92°C再回到53 °C)，以确保所需的分离。将柱子洗脱液与茚三酮试剂(Bie & Berntsen, 商品目录号no. 80-2038-07)混合，使所得混合物通过氨基酸分析仪的高温反应盘管。在反应盘管中，茚三酮与氨基酸反应形成有色化合物，有色化合物的量直接与所存在的氨基酸的量成比例。

#### 蛋白酶活性测试(AU/ml)

变性的血红蛋白(在6.7mM  $\text{KH}_2\text{PO}_4/\text{NaOH}$  缓冲液中占0.65% (w/w), pH 7.50)于25°C用蛋白酶降解10分钟，用三氯乙酸(TCA)沉淀未消化的血红蛋白，过滤除去。用Folin & Ciocalteu's 苯酚试剂测定滤液中的TCA可溶性血红蛋白降解产物，有几种氨基酸会显示蓝色。参考ALCALASE™ 标准测量并确定活性单位(AU)，ALCALASE™ 标准可请求Novozymes A/S, Krogshoejvej 36, DK-2880 Bagsvaerd, Denmark(测试号为 no. EB-SM-0349.02/01)提供。

按下述计算特异活性: 特异活性(AU/g) = (活性 (AU/ml) / AAA (mg/ml))  
x 1000 (mg/g)

源于Nocardiosis alba DSM 15647的蛋白酶的特异活性为53.5 AU/g，与之对照，源于Nocardiosis sp. NRRL 18262的蛋白酶的特异活性为38.3 AU/g。

**实施例4: 蛋白酶 L1a**

如实施例1所述, 使*Nocardiosis dassonvillei* subsp. *dassonvillei* DSM 43235生长于野生型胰蛋白酶培养基, 分离基因组DNA。

用下述引物1424和1485在基因组DNA上扩增成熟蛋白酶L1a前体的编  
5 码序列(SEQ ID NO: 1的88-1143位核苷):

引物1485 (SEQ ID NO: 7):

5'-gcttttagttcatcgatcgatcgctgacgaccgtaccggccgagccag-3'

引物1424 (SEQ ID NO: 8):

5'-ggagcggattgaacatgcatgactaaccggtcaccagggacagcc-3'

10 用PCR手段将L1 a多核苷酸在阅读框中与编码Sav信号肽(SEQ ID NO: 9)的异源DNA片段融合。

通过在枯草芽孢杆菌MB1053宿主细胞基因组(WO03/95658)上的同源重组合并基因(包括信号肽编码部分), 构建了定名为Sav-L1a的枯草芽孢杆菌菌株系。在由地衣芽孢杆菌 $\alpha$ -淀粉酶基因(*amyL*)启动子、解淀粉芽孢杆菌  
15  $\alpha$ -淀粉酶基因(*amyQ*)启动子和苏云金芽孢杆菌(*Bacillus thuringiensis*) *cryIII*A启动子包括稳定化序列组成的三倍启动子系统(如WO 99/43835所述)的控制下表达基因。氯霉素乙酰转移酶用作标记(如Diderichsen,B.; Poulsen,G.B.; Joergensen,S.T.; A useful cloning vector for *Bacillus subtilis*. Plasmid 30:312 (1993)中所述)。

20 按实施例1所述检查了氯霉素抗性转化体的蛋白酶活性, 并选择转化体用于序列确定, 然后同样按照实施例1的方法发酵, 只是于26°C下发酵6天。

离心培养物肉汤(20000 x g, 20 min), 小心地将上清从沉淀上转移走。将合并的上清通过Seitz EKS盘过滤, 以除去芽孢杆菌宿主细胞的剩余物质。将EKS滤液转至G25 sephadex柱上的50mM H<sub>3</sub>BO<sub>3</sub>, 5mM 琥珀酸, 1mM  
25 CaCl<sub>2</sub>, pH为7的溶液。从G25 sephadex柱向酶溶液加入固体硫酸铵, 酶溶液中(NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>的终浓度为1.6M。加入(NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>期间用磁力搅拌器轻轻混合酶溶液, 添加完成后继续搅动30分钟, 使系统平衡。然后将酶溶液上样至在100mM H<sub>3</sub>BO<sub>3</sub>, 10mM 琥珀酸, 2mM CaCl<sub>2</sub>, 1.6M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, pH为7的溶液中平衡过的Butyl Toyopearl柱。用平衡缓冲液彻底冲洗柱子后, 在相同的缓冲  
30 液中用线性(NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 梯度(1.6至0 M)洗脱蛋白酶。取出含有蛋白酶的组分, 转至G25 sephadex 柱上的20mM HEPES, pH为8的溶液, 上样至相同缓冲液

平衡过的Q sepharose FF柱。用相同缓冲液彻底冲洗柱子后，在相同缓冲液中用线性NaCl梯度(0至0.5M)洗脱蛋白酶。用从柱子得到的组分分析蛋白酶活性(用Suc-AAPF-pNA 测试，pH 9)，进一步用SDS-PAGE分析活性组分。将仅有一个条带的组分(用考马斯染色的SDS-PAGE胶判断)取下来，作为纯化的制品用于进一步的鉴定。

L1a蛋白酶是 $\alpha$ -裂解蛋白酶样的酶(S1E肽酶家族，旧符号为S2A)，发现其被苯基甲基磺酰氟(PMSF)和链霉菌枯草杆菌蛋白酶抑制剂(SSI)所抑制。

通过SDS-PAGE测定的其相对分子量为 $M_r = 22\text{kDa}$ 。

按照实施例3所述方法测定的L1a蛋白酶的特异活性为49.8 AU/g。

10

#### 实施例5: 单胃动物体外测试结果

在体外模型中测试了实施例2所述的纯化蛋白酶刺激单胃动物消化的性能。特别地，测定了蛋白酶提高玉米/-SBM (玉米/-大豆粉)蛋白质溶解和消化的能力。体外系统由10个瓶子组成，先将瓶中的玉米/-SBM底物与HCl/胃蛋白酶一起保温-刺激胃内消化-然后再与胰酶一起保温-刺激肠内消化。肠内阶段起始时在5个瓶子中加入蛋白酶，而其余的五个瓶子作为空白。肠内保温阶段结束时，取出体外消化物样品用于分析溶解和消化的蛋白质。

15

大致的体外消化程序

添加的成分	pH	温度	时间进程	经刺激的消化阶段
10 g 玉米/-SBM 底物(6:4), 41 ml HCl (0.105M)	3.0	40°C	t=0 分钟	混合
5 ml HCl (0.105M) / 胃蛋白酶 (3000 U/g 底物), 1 ml 蛋白酶 (提供100 mg 蛋白酶蛋白每kg底 物)	3.0	40°C	t=30 分钟	胃内消化
16 ml H <sub>2</sub> O	3.0	40°C	t=1.0 小时	胃内消化
7 ml NaOH (0.39M)	6.8	40°C	t=1.5 小时	肠内消化
5 ml NaHCO <sub>3</sub> (1M) / 胰酶 (8 mg/g 食物)	6.8	40°C	t=2.0 小时	肠内消化
终止保温	7.0	40°C	t=6.0 小时	

条件

- 5 底物: 4 g SBM, 6 g 玉米 (预混的)  
pH: 3.0 胃步骤 / 6.8-7.0 肠步骤  
HCl: 0.105 M, 1.5 小时(例如30 min HCl-底物预混合)  
胃蛋白酶: 3000 U/g 食物1小时  
胰酶: 8 mg/g食物4小时
- 10 温度: 40°C  
重复数: n

溶液

- 0.39 M NaOH  
15 0.105 M HCl

0.105 M HCl, 每5 ml含有6000 U胃蛋白酶

1 M NaHCO<sub>3</sub>, 每ml含有16 mg胰酶

125 mM NaAc-缓冲液, pH 6.0

- 5 以A<sub>280</sub> 的值和氨基酸序列(氨基酸组合物)为基础, 用S.C.Gill & P.H. von Hippel, Analytical Biochemistry 182, 319-326, (1989)中概述的原理计算蛋白酶蛋白(EP)的量。

实验程序依据上述的概要。在1、2.5、和5.5小时时测量pH。6小时后终止保温, 取出30 ml的样品先放置在冰上, 然后离心(10000 x g, 10 min, 4°C)。

- 10 移出上清, -20°C保存。

通过胶过滤分析所有样品中溶解的和消化的蛋白质的含量, 并用OPA方法分析其水解程度(DH)。

#### 用OPA-方法测定DH

- 15 用基于比色方法的半自动微量滴定盘(Nielsen,P.M.; Petersen,D.; Dambmann,C. Improved method for determining food protein degree of hydrolysis. J.Food Sci. 2001, 66, 642-646)测定不同样品中蛋白质的水解程度。如下制备OPA试剂: 将7.620 g十水四硼酸二钠和200 mg十二烷基硫酸钠(SDS)溶于150 ml去离子水中。试剂完全溶解后再继续。将160 mg 97%的邻苯二醛(OPA)溶解于4 ml乙醇。将OPA溶液定量转移至上述溶液, 用去离子水冲洗。在溶液中加入176 mg 99%的二巯苏糖醇(DTT), 用去离子水加至
- 20 200 ml。将50 mg丝氨酸(Merck, Germany)溶解于500 ml去离子水中制备丝氨酸标准(0.9516 meqv/l)。

- 25 通过将每个样品稀释至吸光值(280 nm)约0.5来制备样品溶液。通常, 用自动Tecan稀释台(Männedorf, Switzerland)稀释(100 ×)上清。所有其他的分光光度计读数均在340 nm进行, 以去离子水为对照。将25 μl样品、标准和空白分配到微量滴定盘。将微量滴定盘插入iEMS MF读数器(Labsystems, Finland), 自动分配200μl OPA试剂。测定吸光值前摇动微量滴定盘(2 min; 700 rpm)。最后, 计算DH。所有样品的测定均进行五次。

30

#### 估算溶解和消化的蛋白质

通过胶过滤HPLC定量粗蛋白(CP),以估算体外消化样品的上清中溶解蛋白的含量。解冻上清,用0.45 μm聚碳酸酯过滤器过滤,用H<sub>2</sub>O稀释(1:50, v/v)。稀释的样品用Superdex Peptide PE (7.5 x 300 mm)胶过滤柱(Global)进行HPLC层析。用于等度洗脱的洗脱液为含有150 mM NaCl的50 mM磷酸钠缓冲液(pH 7.0)。每次洗脱的洗脱液总体积为26 ml,流速为0.4 ml/min。在214 nm处记录洗脱曲线,测定洗脱曲线下的总面积。为了由总面积估算蛋白质含量,用体外消化的参照玉米/-SBM样品的系列稀释制作校准曲线( $R^2=0.9993$ ),其中玉米/-SBM样品的总蛋白含量已知。此参照样品中蛋白质的测定用Kjeldahl法(%氮的测定; A.O.A.C. (1984) Official Methods of Analysis 14th ed., Washington DC)进行。

累加对应于具有1500道尔顿或以下分子量的肽和氨基酸的层析图面积(Savoie,L.; Gauthier,S.F. Dialysis Cell For The In-vitro Measurement Of Protein Digestibility. J. Food Sci. 1986, 51, 494-498; Babinszky,L.; Van,D.M.J.M.; Boer,H.; Den,H.L.A. An In-vitro Method for Prediction of The Digestible Crude Protein Content in Pig Feeds. J. Sci. Food Agr. 1990, 50, 173-178; Boisen,S.; Eggum,B.O. Critical Evaluation of In-vitro Methods for Estimating Digestibility in Simple-Stomach Animals. Nutrition Research Reviews 1991, 4, 141-162)来估算消化的蛋白质的含量。为测定1500道尔顿分界线,以细胞色素 C (Boehringer, Germany)、aprotinin、gastrin I、和substance P (Sigma Aldrich, USA)作为分子量标准校准胶过滤柱。

以下表4和5列出的结果显示,相对于空白而言,蛋白酶显著地提高了可消化蛋白质的水平,以及水解程度。

表 4: 水解程度(DH)

酶 (剂量, mg EP/kg饲料)	n	相对于空白	
		%DH	%CV
空白	5	100.0	a 2.05
实施例2的蛋白酶	5	104.8	b 1.57

同一栏中的不同字母代表显著差异(1-way ANOVA, Tukey-Kramer test,  $P < 0.05$ )。SD = 标准误。%CV = 变异系数 = (SD/平均值) x 100%

5

表 5: 溶解和消化的粗蛋白

酶	n	相对于空白			
		%可消化CP	CV%	%可溶性CP	CV%
空白	5	100.0	a 2.3	100.0	a 2.4
实施例2的蛋白酶	5	104.7	b 0.7	101.8	a 0.5

同一栏中的不同字母代表显著差异(1-way ANOVA, Tukey-Kramer test,  $P < 0.05$ )。SD = 标准误。%CV = 变异系数 = (SD/平均值) x 100%

#### 10 实施例6: 动物饲料与动物饲料添加剂

按实施例2所述制备含蛋白酶的动物饲料添加剂, 维生素与矿质元素预混料形式的动物饲料添加剂由下述表6所示组成。维生素和类胡萝卜素商业上可以从DSM Nutritional Products获得。所有用量均以g/kg计。

表 6: 预混料组合物

维生素 A	ROVIMIX A 500	4.00
维生素 D3	ROVIMIX D3 500	1.00
维生素 E	ROVIMIX E 50 Ads	8.00
维生素 B2	ROVIMIX B2 80-SD	1.0
	CAROPHYLL Yellow	10.0
	氯化胆碱(Choline chloride)50%, min.	300.0
矿质元素	Mn氧化物	60.0
	Zn氧化物	12.0
	一水硫酸铁	20.0
	Cu氧化物	2.0
	Co硫酸盐	0.2
酶	实施例2的蛋白酶	10.0
	小麦粗粉	571.8

产蛋鸡的食物中包括表6的预混料和下面表7中显示的组合物。每种成分的量以均为% (w/w)。食物中L2a蛋白酶的含量为100 mg蛋白酶蛋白每kg食物。

5

表 7: 产蛋鸡食物

玉米	55.00
小麦	10.00
燕麦	7.50
大豆	20.00
石灰	7.50
表6的预混料	1.00

## 原始文件 (用于提交) (该页不是也不作为国际申请的一部分)

0-1	PCT/RO/134 表 (SAFE) 对保藏的微生物或其它生物物质的说明 (PCT 细则 13 之二)	
0-1-1	备用 (prepared using)	PCT-SAFE [EASY mode] Version 3.50 (Build 0002.162)
0-2	国际申请号	
0-3	申请人或代理人档案号	10476.204-WO

1	以下说明是针对在说明书被提及的保藏的微生物或其它生物物质:	
1-1	页数	48
1-2	行数	20
1-3	保藏鉴定	
1-3-1	保藏机构	DSMZ DSMZ-Deutsche Sammlung von Mikroorganismen and Zellkulturen GmbH
1-3-2	保藏机构的地址	Mascheroder Weg 1b, D-38124 Braunschweig, Germany
1-3-3	保藏日期	30 May 2003 (30.05.2003)
1-3-4	保藏编号	DSMZ 15647
1-4	补充说明	Nocardiosis alba
1-5	本说明所针对的指定国	All designations

5

## 由受理局填写

0-4	本表格已经和国际申请一起收到 (是或否)	是
0-4-1	受理官员	

10

## 由国际局填写

0-5	国际局收到本表格日期	2004/04/13
0-5-1	受理官员	

<110> 诺维信公司 (Novozymes A/S)  
 <120> 蛋白酶  
 <130> 10476.204-WO  
 <160> 9  
 <170> PatentIn version 3.2  
 <210> 1  
 <211> 1068  
 <212> DNA  
 <213> Nocardiosis alba DSM 15647  
  
 <220>  
 <221> CDS  
 <222> (1)..(1065)  
  
 <220>  
 <221> mat\_peptide  
 <222> (502)..(1065)  
  
 <400> 1  
 gcg acc ggc ccc ctc ccc cag tcc ccc acc ccg gat gaa gcc gag 45  
 Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Asp Glu Ala Glu  
 -165 -160 -155  
  
 gcc acc acc atg gtc gag gcc ctc cag cgc gac ctc ggc ctg tcc 90  
 Ala Thr Thr Met Val Glu Ala Leu Gln Arg Asp Leu Gly Leu Ser  
 -150 -145 -140  
  
 ccc tct cag gcc gac gag ctc ctc gag gcg cag gcc gag tcc ttc 135  
 Pro Ser Gln Ala Asp Glu Leu Leu Glu Ala Gln Ala Glu Ser Phe  
 -135 -130 -125  
  
 gag atc gac gag gcc gcc acc gcg gcc gca gcc gac tcc tac ggc 180  
 Glu Ile Asp Glu Ala Ala Thr Ala Ala Ala Ala Asp Ser Tyr Gly  
 -120 -115 -110  
  
 ggc tcc atc ttc gac acc gac agc ctc acc ctg acc gtc ctg gtc acc 228  
 Gly Ser Ile Phe Asp Thr Asp Ser Leu Thr Leu Thr Val Leu Val Thr  
 -105 -100 -95  
  
 gac gcc tcc gcc gtc gag gcg gtc gag gcc gcc ggc gcc gag gcc aag 276  
 Asp Ala Ser Ala Val Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys  
 -90 -85 -80  
  
 gtg gtc tcg cac ggc atg gag ggc ctg gag gag atc gtc gcc gac ctg 324  
 Val Val Ser His Gly Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu  
 -75 -70 -65 -60  
  
 aac gcg gcc gac gct cag ccc ggc gtc gtg ggc tgg tac ccc gac atc 372  
 Asn Ala Ala Asp Ala Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile  
 -55 -50 -45  
  
 cac tcc gac acg gtc gtc ctc gag gtc ctc gag ggc tcc ggt gcc gac 420  
 His Ser Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp  
 -40 -35 -30  
  
 gtg gac tcc ctg ctc gcc gac gcc ggt gtg gac acc gcc gac gtc aag 468  
 Val Asp Ser Leu Leu Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys  
 -25 -20 -15  
  
 gtg gag agc acc acc gag cag ccc gag ctg tac gcc gac atc atc ggc 516  
 Val Glu Ser Thr Thr Glu Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly  
 -10 -5 -1 1 5

ggt ctc gcc tac acc atg ggt ggg cgc tgc tgc gtc ggc ttc gcg gcc 564  
 Gly Leu Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala  
 10 15 20

acc aac gcc tcc ggc cag ccc ggg ttc gtc acc gcc ggc cac tgc ggc 612  
 Thr Asn Ala Ser Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly  
 25 30 35

acc gtc ggc acc ccg gtc agc atc ggc aac ggc cag ggc gtc ttc gag 660  
 Thr Val Gly Thr Pro Val Ser Ile Gly Asn Gly Gln Gly Val Phe Glu  
 40 45 50

cgt tcc gtc ttc ccc ggc aac gac tcc gcc ttc gtc cgc ggc acc tcg 708  
 Arg Ser Val Phe Pro Gly Asn Asp Ser Ala Phe Val Arg Gly Thr Ser  
 55 60 65

aac ttc acc ctg acc aac ctg gtc agc cgc tac aac acc ggt ggt tac 756  
 Asn Phe Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr  
 70 75 80 85

gcg acc gtc tcc ggc tcc tcg cag gcg gcg atc ggc tcg cag atc tgc 804  
 Ala Thr Val Ser Gly Ser Gln Ala Ala Ile Gly Ser Gln Ile Cys  
 90 95 100

cgt tcc ggc tcc acc acc ggc tgg cac tgc ggc acc gtc cag gcc cgc 852  
 Arg Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Val Gln Ala Arg  
 105 110 115

ggc cag acg gtg agc tac ccc cag ggc acc gtg cag aac ctg acc cgc 900  
 Gly Gln Thr Val Ser Tyr Pro Gln Gly Thr Val Gln Asn Leu Thr Arg  
 120 125 130

acc aac gtc tgc gcc gag ccc ggt gac tcc ggc ggc tcc ttc atc tcc 948  
 Thr Asn Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ser  
 135 140 145

ggc agc cag gcc cag ggc gtc acc tcc ggt ggc tcc ggc aac tgc tcc 996  
 Gly Ser Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser  
 150 155 160 165

ttc ggt ggc acc acc tac tac cag gag gtc aac ccg atg ctg agc agc 1044  
 Phe Gly Gly Thr Thr Tyr Tyr Gln Glu Val Asn Pro Met Leu Ser Ser  
 170 175 180

tgg ggt ctg acc ctg cgc acc tga 1068  
 Trp Gly Leu Thr Leu Arg Thr  
 185

<210> 2  
 <211> 355  
 <212> PRT  
 <213> Nocardiosis alba DSM 15647

<400> 2

Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Asp Glu Ala Glu  
 -165 -160 -155

Ala Thr Thr Met Val Glu Ala Leu Gln Arg Asp Leu Gly Leu Ser  
 -150 -145 -140

Pro Ser Gln Ala Asp Glu Leu Leu Glu Ala Gln Ala Glu Ser Phe  
 -135 -130 -125

Glu Ile Asp Glu Ala Ala Thr Ala Ala Ala Ala Asp Ser Tyr Gly  
 -120 -115 -110

Gly Ser Ile Phe Asp Thr Asp Ser Leu Thr Leu Thr Val Leu Val Thr  
 -105 -100 -95

Asp Ala Ser Ala Val Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys  
 -90 -85 -80

Val Val Ser His Gly Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu  
 -75 -70 -65 -60

Asn Ala Ala Asp Ala Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile  
 -55 -50 -45

His Ser Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp  
 -40 -35 -30

Val Asp Ser Leu Leu Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys  
 -25 -20 -15

Val Glu Ser Thr Thr Glu Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly  
 -10 -5 -1 1 5

Gly Leu Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala  
 10 15 20

Thr Asn Ala Ser Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly  
 25 30 35

Thr Val Gly Thr Pro Val Ser Ile Gly Asn Gly Gln Gly Val Phe Glu  
 40 45 50

Arg Ser Val Phe Pro Gly Asn Asp Ser Ala Phe Val Arg Gly Thr Ser  
 55 60 65

Asn Phe Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr  
 70 75 80 85

Ala Thr Val Ser Gly Ser Ser Gln Ala Ala Ile Gly Ser Gln Ile Cys  
 90 95 100

Arg Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Val Gln Ala Arg  
 105 110 115

Gly Gln Thr Val Ser Tyr Pro Gln Gly Thr Val Gln Asn Leu Thr Arg  
 120 125 130

Thr Asn Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ser  
 135 140 145

Gly Ser Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser

150	155	160	165	
Phe Gly Gly Thr	Thr Tyr Tyr Gln Glu Val Asn Pro Met Leu Ser Ser			
	170	175	180	
Trp Gly Leu Thr	Leu Arg Thr			
	185			
<210>	3			
<211>	43			
<212>	DNA			
<213>	合成的			
<220>				
<221>	misc-feature			
<223>	引物			
<400>	3			
gttcatcgat cgcacggct gcgaccggcc cctccccca gtc				43
<210>	4			
<211>	31			
<212>	DNA			
<213>	合成的			
<220>				
<221>	misc-feature			
<223>	引物			
<400>	4			
gcggatccta tcaggtgcgc agggtcagac c				31
<210>	5			
<211>	1146			
<212>	DNA			
<213>	Nocardiosis dassonvillei 亚种 dassonvillei DSM 43235			
<220>				
<221>	CDS			
<222>	(1)..(1143)			
<220>				
<221>	sig-peptide			
<222>	(1)..(87)			
<220>				
<221>	mat-peptide			
<222>	(568)..(1143)			
<400>	5			
atg cga ccc tcc ccc gct atc tcc gct atc ggc acc ggc gca ctc				45
Met Arg Pro Ser Pro Ala Ile Ser Ala Ile Gly Thr Gly Ala Leu				
	-185	-180	-175	
gcg ttc ggt ctg gcg ttc tcc gtg acg ccg ggc gcc agt gcg gcg				90
Ala Phe Gly Leu Ala Phe Ser Val Thr Pro Gly Ala Ser Ala Ala				
	-170	-165	-160	
acc gta ccg gcc gag cca gcg agc gag gcc cag acg atg atg gaa				135
Thr Val Pro Ala Glu Pro Ala Ser Glu Ala Gln Thr Met Met Glu				
	-155	-150	-145	

gcg ctg cag aga gac ctc ggc ctc acc ccg ctc ggg gcc gag gag Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro Leu Gly Ala Glu Glu -140 -135 -130	180
ctg ctc tcg gcg cag gaa gag gcg atc gag acc gac gcc gag gcc Leu Leu Ser Ala Gln Glu Glu Ala Ile Glu Thr Asp Ala Glu Ala -125 -120 -115	225
acc gag gcc gcg gga gcg tcc tac ggc ggc tcc ctg ttc gac acc Thr Glu Ala Ala Gly Ala Ser Tyr Gly Gly Ser Leu Phe Asp Thr -110 -105 -100	270
gag acc ctc cag ctc acc gtg ctg gtg acc gac gcc tcg gcc gtc gag Glu Thr Leu Gln Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val Glu -95 -90 -85	318
gcg gtg gag gcc acc ggc gcc gag gcc acc gtg gtc tca cac ggc gca Ala Val Glu Ala Thr Gly Ala Glu Ala Thr Val Val Ser His Gly Ala -80 -75 -70	366
gag ggc ctg gcc gag gtg gtc gac gcg ctc gac gag acc ggc ggc cgg Glu Gly Leu Ala Glu Val Val Asp Ala Leu Asp Glu Thr Gly Gly Arg -65 -60 -55	414
gaa ggg gtc gtc ggc tgg tac ccg gac gtg gag agc gac acc gtc gtg Glu Gly Val Val Gly Trp Tyr Pro Asp Val Glu Ser Asp Thr Val Val -50 -45 -40	462
gtc cag gtc gcc gag ggc gcc agc gcc gac ggc ctc atc gag gcc gcg Val Gln Val Ala Glu Gly Ala Ser Ala Asp Gly Leu Ile Glu Ala Ala -35 -30 -25 -20	510
ggc gtg gac ccc tcc gcc gtc cgg gtg gag gag acc agt gag act ccg Gly Val Asp Pro Ser Ala Val Arg Val Glu Glu Thr Ser Glu Thr Pro -15 -10 -5	558
cgc ctg tac gcc gac atc gtc ggc ggc gag gcg tac tac atg ggc ggc Arg Leu Tyr Ala Asp Ile Val Gly Gly Glu Ala Tyr Tyr Met Gly Gly -1 1 5 10	606
gga cgc tgc tcg gtc ggg ttc gcc gtg acc gac ggc tcc ggc gcg ggc Gly Arg Cys Ser Val Gly Phe Ala Val Thr Asp Gly Ser Gly Ala Gly 15 20 25	654
ggc ttc gtg acg gcg ggc cac tgc ggc acc gtc ggc acc ggc gcc gag Gly Phe Val Thr Ala Gly His Cys Gly Thr Val Gly Thr Gly Ala Glu 30 35 40 45	702
agc tcc gac ggc agc ggc tcc gga acc ttc cag gag tcc gtc ttc ccg Ser Ser Asp Gly Ser Gly Ser Gly Thr Phe Gln Glu Ser Val Phe Pro 50 55 60	750
ggc agc gac ggc gcc ttc gtc gcg gcc acc tcc aac tgg aac gtg acc Gly Ser Asp Gly Ala Phe Val Ala Ala Thr Ser Asn Trp Asn Val Thr 65 70 75	798
aac ctg gtc agc cgg tac gac tcc ggc agc ccc cag gcg gtg tcg ggt Asn Leu Val Ser Arg Tyr Asp Ser Gly Ser Pro Gln Ala Val Ser Gly 80 85 90	846
tcc agc cag gcc ccg gag ggc tcg gcg gtg tgc cgc tcc ggc tcc acc Ser Ser Gln Ala Pro Glu Gly Ser Ala Val Cys Arg Ser Gly Ser Thr 95 100 105	894
acc ggc tgg cac tgc ggg acc atc gag gcc cgc ggc cag acg gtg aac Thr Gly Trp His Cys Gly Thr Ile Glu Ala Arg Gly Gln Thr Val Asn 110 115 120 125	942

tac ccg cag ggc acg gtc cag gac ctg acc cgg acg gac gtg tgc gcc 990  
 Tyr Pro Gln Gly Thr Val Gln Asp Leu Thr Arg Thr Asp Val Cys Ala  
 130 135 140

gag ccc ggt gac tcc ggc ggc tcg ttc atc gcc ggt tcg cag gcc cag 1038  
 Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ala Gly Ser Gln Ala Gln  
 145 150 155

ggc gtc acc tcc ggc ggc tcg ggc aac tgc acc tcc ggc ggc acg acc 1086  
 Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Thr Ser Gly Gly Thr Thr  
 160 165 170

tac tac cag gag gtc act ccc ctg ctg agc agc tgg ggg ctg tcc ctg 1134  
 Tyr Tyr Gln Glu Val Thr Pro Leu Leu Ser Ser Trp Gly Leu Ser Leu  
 175 180 185

gtg acc ggt tag 1146  
 Val Thr Gly  
 190

<210> 6  
 <211> 381  
 <212> PRT  
 <213> *Nocardiosis dassonvillei* 亚种 *dassonvillei* DSM 43235

<400> 6

Met Arg Pro Ser Pro Ala Ile Ser Ala Ile Gly Thr Gly Ala Leu  
 -185 -180 -175

Ala Phe Gly Leu Ala Phe Ser Val Thr Pro Gly Ala Ser Ala Ala  
 -170 -165 -160

Thr Val Pro Ala Glu Pro Ala Ser Glu Ala Gln Thr Met Met Glu  
 -155 -150 -145

Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro Leu Gly Ala Glu Glu  
 -140 -135 -130

Leu Leu Ser Ala Gln Glu Glu Ala Ile Glu Thr Asp Ala Glu Ala  
 -125 -120 -115

Thr Glu Ala Ala Gly Ala Ser Tyr Gly Gly Ser Leu Phe Asp Thr  
 -110 -105 -100

Glu Thr Leu Gln Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val Glu  
 -95 -90 -85

Ala Val Glu Ala Thr Gly Ala Glu Ala Thr Val Val Ser His Gly Ala  
 -80 -75 -70

Glu Gly Leu Ala Glu Val Val Asp Ala Leu Asp Glu Thr Gly Gly Arg  
 -65 -60 -55

Glu Gly Val Val Gly Trp Tyr Pro Asp Val Glu Ser Asp Thr Val Val  
 -50 -45 -40

Val Gln Val Ala Glu Gly Ala Ser Ala Asp Gly Leu Ile Glu Ala Ala



---

<400> 7  
gcttttagtt catcgatcgc atcggctgcg accgtaccgg ccgagccag 49

<210> 8  
<211> 46  
<212> DNA  
<213> 合成的

<220>  
<221> misc.feature  
<223> 引物

<400> 8  
ggagcggatt gaacatgcga ttactaaccg gtcaccaggg acagcc 46

<210> 9  
<211> 81  
<212> DNA  
<213> Bacillus clausii

<220>  
<221> sig\_peptide  
<222> (1)..(81)

<400> 9  
atgaagaaac cgttggggaa aattgtcgca agcaccgcac tactcatttc tgttgctttt 60  
agttcatcga tcgcatcggc t 81