

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12N 15/57

C12P 21/02

C12N 1/19



[12] 发明专利申请公开说明书

[21] 申请号 03825284.8

[43] 公开日 2005 年 11 月 23 日

[11] 公开号 CN 1701119A

[22] 申请日 2003.9.19 [21] 申请号 03825284.8

[30] 优先权

[32] 2002.9.20 [33] DK [31] PA200201391

[86] 国际申请 PCT/DK2003/000609 2003.9.19

[87] 国际公布 WO2004/027067 英 2004.4.1

[85] 进入国家阶段日期 2005.5.19

[71] 申请人 CYMIP 有限公司

地址 丹麦哥本哈根

[72] 发明人 L·荣松 J·F·雷费尔德

A·H·约翰森

[74] 专利代理机构 北京市中咨律师事务所

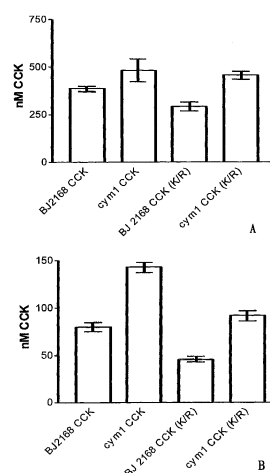
代理人 黄革生 林柏楠

权利要求书 2 页 说明书 52 页 序列表 68 页
附图 17 页

[54] 发明名称 增加宿主细胞重组多肽产量的方法

[57] 摘要

本发明提供了包含编码重组多肽的核酸序列的宿主细胞，其中天然存在的包含序列 HXXEH 的金属蛋白酶的产生通过遗传操作得到降低或抑制。本发明还涉及通过破坏金属蛋白酶的合成或活性而提高细胞产生多肽的方法。本发明尤其涉及提高自宿主细胞例如但不限于酵母和细菌细胞分泌重组多肽的方法。金属蛋白酶为蛋白酶 pitrilysin 亚家族的成员，其特征在于包含序列 HXXEH，其中 X 为任一氨基酸。



1. 用于在宿主细胞中产生目的蛋白质的方法，其中所述宿主细胞已经遗传修饰，从而与相应的未经修饰细胞相比，在相同条件下培养时表达显著降低水平的包含 HXXEH 基元 (SEQ ID NO 1) 的金属蛋白酶，所述方法包括：

- a) 向宿主细胞中导入编码目的蛋白质的核酸序列；
- b) 在适当的生长培养基中培养步骤 (a) 的宿主细胞，以产生目的蛋白质；以及
- c) 分离目的蛋白质。

2. 根据权利要求 1 的方法，其中金属蛋白酶还包含位于 HXXEH 基元第二个组氨酸残基 C-端 70 和 80 个氨基酸之间的谷氨酸残基。

3. 根据前述权利要求中任意一项所述的方法，其中金属蛋白酶还包含位于 HXXEH 基元第一个组氨酸残基 N-端 3 个氨基酸的甘氨酸残基。

4. 根据前述权利要求中任意一项所述的方法，其中金属蛋白酶还包含位于 HXXEH 基元第二个组氨酸残基 C-端 5 个氨基酸的甘氨酸残基。

5. 根据前述权利要求中任意一项所述的方法，其中金属蛋白酶还包含位于 HXXEH 基元第二个组氨酸残基 C-端 8 个氨基酸的赖氨酸残基。

6. 根据前述权利要求中任意一项所述的方法，其中金属蛋白酶还包含位于 HXXEH 基元第二个组氨酸残基 C-端 9 个氨基酸的酪氨酸残基。

7. 根据前述权利要求中任意一项所述的方法，其中金属蛋白酶还包含位于 HXXEH 基元第二个组氨酸残基 C-端 10 个氨基酸的脯氨酸残基。

8. 根据前述权利要求中任意一项所述的方法，其中金属蛋白酶还包含共有序列 SEQ ID NO 2。

9. 根据前述权利要求中任意一项所述的方法，其中金属蛋白酶还包含共有序列 SEQ ID NO 3。

10. 根据前述权利要求中任意一项所述的方法，其中金属蛋白酶还在 HXXEH 基元第二个组氨酸残基 C-端 20 和 30 个氨基酸之间包含

NAXTXXXXT 基元。

11. 根据前述权利要求中任意一项所述的方法，其中金属蛋白酶选自：

- i) 由 SEQ ID NO 4 到 15 组成的组中的任何一个序列；以及
- ii) 与 SEQ ID NO 4 到 15 中任一序列至少 80 % 相同的序列。

12. 根据前述权利要求中任意一项所述的方法，其中金属蛋白酶与 SEQ ID NO:4 至少 80 % 相同。

13. 根据前述权利要求中任意一项所述的方法，其中目的蛋白质的总量与相应的未经修饰细胞相比在相同条件下培养时至少增加 5 %。

14. 根据前述权利要求中任意一项所述的方法，其中目的蛋白质的总量与相应的未经修饰细胞相比在相同条件下培养时至少增加 50 %。

15. 根据前述权利要求中任意一项所述的方法，其中宿主细胞为原核细胞。

16. 根据权利要求 1-14 中任意一项所述的方法，其中宿主细胞为真核细胞。

17. 根据权利要求 16 的方法，其中宿主细胞为非丝状真菌细胞。

18. 根据权利要求 16 的方法，其中宿主细胞为丝状真菌细胞。

19. 根据权利要求 17 的方法，其中宿主细胞为酵母属菌株。

20. 根据权利要求 19 的方法，其中宿主细胞为酿酒酵母。

21. 用于表达目的蛋白质的宿主细胞，其中所述细胞已经遗传修饰，从而与相应的未经修饰细胞相比，当在相同条件下培养时表达显著降低水平的包含 HXXEH 基元 (SEQ ID NO 1) 的金属蛋白酶。

22. 根据权利要求 21 的宿主细胞，其中金属蛋白酶还包含共有序列 SEQ ID NO 3。

增加宿主细胞重组多肽产量的方法

发明领域

本发明涉及通过破坏 ME 集团 (M16 家族) 的金属蛋白酶的合成或活性来提高细胞中多肽产量的方法。本发明尤其涉及提高自宿主细胞的重组多肽分泌的方法, 提高例如但不限于酵母和细菌细胞中重组多肽分泌的方法。

发明背景

缩胆囊素 (CCK) 为肠和脑组织中表达的脊椎动物神经内分泌肽激素。生物活性 CCK 肽的成熟依赖于翻译后的酪氨酸硫酸盐化作用、内肽酶切割、外切蛋白酶修剪和羧基端酰胺化。N-端内切蛋白酶加工过程已经鉴定出 CCK-83、-58、-39、-33、-22、-8 和 -5。大多数 CCK 肽在单一的精氨酸残基处切割后得到合成, 然而 CCK-22 需要单一的赖氨酸残基之后的加工。

许多重组多肽已经在酵母中以融合蛋白的形式, 表达为酿酒酵母 α -因子前原肽 (prepropeptide) 以通过分泌途径指导分泌。特征最清楚的酵母蛋白酶为丝氨酸内切蛋白酶 Kex2p (Fuller 等, 1989), 其参与 α -交配信息素和杀伤毒素的成熟过程 (Julius 等, 1984)。另一个酵母蛋白酶为属于糖基磷脂酰肌醇 (GPI)-锚定的天冬氨酰蛋白酶 yapsin 家族的 Yps1p, 其在 *kex2* 突变体中过量表达能拯救交配缺陷 (Egel-Mitani 等, 1990)。外源蛋白质的表达表明 Yps1p 和 Yps2p 含有内切蛋白酶活性。

宿主细胞用于重组多肽的表达大大简化了大量商业价值多肽的生产, 否则其只能通过从天然来源纯化而获得。目前可获得的表达系统有多种选择, 可从中选择以生产任一特定的多肽, 包括真细菌和真核宿主。选择适

当的表达系统中一个重要因素为宿主细胞产生充足多肽产量的能力。然而，常常遇到的问题是特定宿主细胞或培养基中产生的高水平蛋白水解酶。因此，需要提高宿主细胞中重组多肽产量的其他方法。

金属蛋白酶为四种主要蛋白酶类型中最为多样的一种，目前已经鉴定出 30 多个家族。在这些酶中，二价阳离子通常为锌活化水分子。锌金属蛋白酶可基于锌结合位点而分为例如 Zincins 和 Inverzincins (Hooper, N.M. 1994)。金属离子通过通常数量为 3 个的氨基酸配基固定。已知的金属配基为组氨酸、谷氨酸、天冬氨酸或赖氨酸和至少一个其他残基，其起到催化所需的亲电作用。已知的金属蛋白酶中大约一半含有 HEXXH 基元，晶体研究表明其形成部分金属结合位点。

研究表明活性依赖于二价阳离子的许多蛋白酶属于一个家族，即肽酶 M16 家族。这包括胰岛素酶、线粒体加工蛋白酶、pitrilysin、nardilysin 和许多细菌蛋白质。这些蛋白质共有不多序列相似性的区域，但其中最值得注意的为 N-端部分。这一区域包括保守的组氨酸，其后为两个残基，然后为一个谷氨酸和另一个组氨酸。已经表明 pitrilysin 中的此 HXXEH 基元参与了酶活性 (Becker 等, 1992)；两个组氨酸结合锌，谷氨酸为催化活性必需的。X 可为任一氨基酸。此家族中的无活性成员丢失这些活性位点残基中的一到三个残基。

以前在例如 US 5,861,280(WO 98/12300)中建议提供这样的产生蛋白质的宿主细胞和方法，其在丝状真菌宿主细胞中通过表达水平显著降低的遗传修饰来表达水平显著降低的含有 HEXXH 基元的金属蛋白酶。

其他人还提供了蛋白酶缺乏的丝状真菌，其特征不在于丝状真菌含有造成该丝状真菌金属蛋白酶活性降低的 DNA 位点选择性破坏，并提供了从丝状真菌中获得的编码具有金属蛋白酶活性蛋白质的经分离 DNA 序列 (WO 97/46689)。另外，所述金属蛋白酶含有 HEXXH 基元。

然而，这样的金属蛋白酶还尚未描述，其可通过遗传修饰降低以在非丝状真菌宿主细胞和含有非 HEXXH 基元的其他细胞中表达水平显著降低的所述金属蛋白酶。

发明概述

研究重组宿主细胞中多种蛋白酶在加工 proCCK 过程中所起的作用的同时,本发明者惊奇地注意到 *CYM1* 的缺失/破坏能提高重组多肽的产量和分泌。另一方面,本发明者发现 *Cym1p* 属于金属蛋白酶家族,其活性下调可提高宿主细胞生产的重组多肽水平。

因此,本发明首先在一方面提供了包含编码重组多肽的核酸序列的宿主细胞,其中包含 SEQ ID NO:1 所述序列的天然存在金属蛋白酶的产量通过遗传操作被减少或抑制。

宿主细胞可以为任何天然状态下拥有金属蛋白酶的细胞。因此宿主细胞可为真核或原核细胞。优选的真核细胞例子包括但不限于哺乳动物细胞、植物细胞和真菌细胞。在优选的实施方案中,宿主细胞为酵母细胞。更优选地,酵母细胞选自但不限于酵母属 (*Saccharomyces sp.*) 的某些种例如酿酒酵母 (*Saccharomyces cerevisiae*)、奇异酵母 (*Saccharomyces paradoxus*)、*Saccharomyces mikatae*、贝酵母 (*Saccharomyces bayanus*)、*Saccharomyces castellii* 和克鲁弗酵母 (*Saccharomyces kluyveri*)、裂殖酵母属的某些种 (*Schizosaccharomyces sp.*) 例如粟酒裂殖酵母 (*Schizosaccharomyces pombe*)、乳酸克鲁维酵母 (*Kluyveromyces lactis*)、多形汉逊酵母 (*Hansenula polymorpha*)、巴斯德毕赤酵母 (*Pichia pastoris*)、甲醇毕赤酵母 (*Pichia methanolica*)、克鲁弗毕赤酵母 (*Pichia kluyveri*)、*Yarrowia lipolytica*、产朊假丝酵母 (*Candida utilis*)、可可假丝酵母 (*Candida cacaoi*) 和发酵地霉 (*Geotrichum fermentans*)。

金属蛋白酶属于水解酶,其中对肽键的亲核攻击由水分子介导。这是天冬氨酸蛋白酶共有的特征,但在金属蛋白酶中通常为锌有时为钴或锰的二价阳离子活化水分子。金属离子通常通过数量为 3 个的氨基酸配基固定,金属蛋白酶中已知的金属配基为组氨酸、谷氨酸、天冬氨酸或赖氨酸残基。

金属蛋白酶可依照催化所需的金属离子分成两个主要的组,在文献中金属蛋白酶分成至少 8 个不同的集团: MA、MB、MC、MD、ME、MF、

MG 和 MH。由此表明这组蛋白酶的复杂多样性。这种分类基于配基结合的不同共有序列，并且这样每一个家族具有不同的生物底物和/或功能。

根据本发明下调的金属蛋白酶为蛋白酶 pitrilysin 亚家族 (ME) 的一员，其特征在于包含 HXXEH (SEQ ID NO:1) 序列，其中 X 为任一氨基酸。目前有 180 多个成员为 Swissprot 的 ME 集团。这样最优选的实施方案中金属蛋白酶包含 SEQ ID NO:1 提供的序列。另一个实施方案中，金属蛋白酶包含 SEQ ID NO:2 提供的序列。甚至更优选的，金属蛋白酶包含 SEQ ID NO:3 提供的序列。此外，金属蛋白酶优选地包含 SEQ ID NO:1 和第二个组氨酸残基 C-端 70-80 位氨基酸之间的谷氨酸残基。此外，金属蛋白酶优选地包含选自 SEQ ID NO:1、SEQ ID NO:2 和 SEQ ID NO:3 的序列以及选自以下的序列：

- i) 由 SEQ ID NO: 4 到 15 组成的组中的任何一个序列和
- ii) 与 SEQ ID NO: 4 到 15 任何一个至少 80 % 相同的序列。

金属蛋白酶更优选地包含 SEQ ID NO:1 或 SEQ ID NO:2 或 SEQ ID NO:3 以及选自下组中的序列：

- i) SEQ ID NO: 4 或 5 的任何一个序列和
- ii) 与 SEQ ID NO: 4 或 5 任何一个至少 80 % 相同的序列。

金属蛋白酶优选地包含与 SEQ ID NO: 4 或 15 中的任何一个序列至少 85 % 相同的序列，例如至少 90 % 相同，例如至少 95 % 相同，例如至少 99 % 相同。

在尤其优选的实施方案中，金属蛋白酶包含 SEQ ID NO:4 提供的序列，或者与其至少 80 % 相同的序列，例如至少 90 %，例如至少 95 %，例如至少 99 % 相同的序列。

宿主细胞可通过任何本领域已知的方法进行遗传操作，只要与未进行遗传操作的亲本宿主细胞相比金属蛋白酶得到降低或抑制。此类遗传操作宿主细胞的方法包括但不限于：基因敲除、基因破坏、随机或定点诱变、引入显性失活的金属蛋白酶、使用 dsRNA 的 RNA 干扰 (RNAi)、催化核酸（例如核酶和 DNA 酶）和反义核酸。遗传操作优选地直接作用于编

码金属蛋白酶的基因、基因转录的 mRNA 或产生改变金属蛋白酶活性的蛋白质，例如与金属蛋白酶竞争结合于底物但例如不具有催化活性的显性失活突变体。然而，可对宿主细胞进行遗传操作，从而使其间接影响金属蛋白酶的产量或活性。例如，遗传操作可靶向于参与转录金属蛋白酶基因编码的 mRNA 的转录因子，这样至少降低所操作的宿主细胞产生的金属蛋白酶水平。

此外，宿主细胞可进一步遗传操作以使其缺少至少一个其他天然存在的宿主细胞金属蛋白酶，或者至少一个其他天然存在的宿主细胞金属蛋白酶具有降低的活性。蛋白酶可为其酶活性的抑制能增加宿主细胞重组多肽产量的任何酶。蛋白酶可为内肽酶、氨肽酶或羧肽酶。优选的蛋白酶包括丝氨酸蛋白酶、天冬氨酸蛋白酶、半胱氨酸蛋白酶和其他金属蛋白酶。

在一个实施方案中，宿主细胞为酵母细胞，其他天然存在的蛋白酶为任何以下蛋白酶基因编码的至少一个蛋白酶，所述蛋白酶基因选自 KEX2、YPS1 (以前已知的 YAP3)、YPS2 (以前已知的 MKC7)、YPS3、YPS6、YPS7、BAR1、STE13、KEX1、PRC1、PEP4 (也称为 PRA1)、APE1、APE2、APE3 和 CPS1。宿主细胞优选地为酵母细胞并且 KEX2 的产生得到破坏。除了在此特别描述的金属蛋白酶之外，其他非酵母宿主细胞中相似的天然存在蛋白酶也可破坏和/或遗传操作以获得进一步的产量提高。

重组多肽可为宿主细胞中能产生的任一目的多肽。重组多肽可包含天然存在的序列或人为干涉产生的序列（例如天然存在蛋白质的突变体或截短，或至少两个不同多肽的融合）。重组多肽一般具有商业价值，例如在疾病治疗中具有商业价值。

重组多肽可为任意大小。重组多肽一般地大小范围为约 30 个氨基酸到约 4,500 个氨基酸。在一个实施方案中，重组多肽长度在约 30 个到约 200 个氨基酸之间。

至少在一些产生重组多肽的宿主细胞表达系统中，需要指导重组多肽从宿主细胞中分泌。这样在优选的实施方案中，核酸包含编码指导重组多肽从宿主细胞分泌的信号序列。信号序列优选地为 N-端疏水信号序列。此

类 N-端疏水信号序列为本领域已知, 并包括例如但不限于来源于真菌淀粉葡萄糖苷酶 (AG) 基因的前导序列, 例如来源于曲霉菌某些种 (*Aspergillus sp.*) 的 galA (包括 18 和 24 氨基酸形式); 酵母属某些种 (*Saccharomyces sp.*) 和克鲁维酵母某些种 (*Kluyveromyces sp.*) 酵母的 α -因子基因; 裂殖酵母某些种 (*Schizosaccharomyces sp.*) 的 P-因子和芽孢杆菌属某些种 (*Bacillus sp.*) α -淀粉酶基因。在一个实施方案中, 重组多肽以 N-端疏水信号序列和编码与信号序列不同来源的重组多肽的第二多肽序列的融合形式表达。

编码重组多肽的核酸可使用本领域任何已知的技术提供给宿主细胞。在一个实施方案中, 核酸使用例如基于同源重组的技术插入宿主细胞的基因组。在另一个实施方案中, 核酸在保留于染色体外的表达构建体中转染或转化入宿主细胞。例如, 表达构建体可以为质粒或病毒。此外, 载体优选地包含选择性标记, 其可用于有选择地繁殖包含载体的宿主细胞。此类选择性标记和其用途也为本领域已知。

第二方面, 本发明提供了产生重组多肽的方法, 方法包括在适当的条件下培养根据第二方面的宿主细胞以产生重组多肽, 以及收获重组多肽。

因为金属蛋白酶的蛋白水解作用降低或抑制, 本发明第二方面的方法提高了宿主细胞产生的重组多肽的稳定性。

在优选的实施方案中, 重组多肽从宿主细胞中分泌。此外, 分泌的蛋白质优选地在包含宿主细胞的培养物的指数生长阶段收获。

所收获的重组多肽的量优选地高于使用亲本宿主细胞的情况。更优选地, 所收获的重组多肽的量比使用亲本宿主细胞的情况高 50%。

第三方面, 本发明提供了在碱性残基处切割多肽的方法, 方法包括在二价阳离子存在的条件下使多肽与包含选自下组序列的金属蛋白酶接触:

- i) SEQ ID NO:4 到 15 中的任何一个序列;
- ii) 与 SEQ ID NO:4 到 15 任何一个至少 80% 相同的序列。

在尤其优选的实施方案中, 金属蛋白酶包含如 SEQ ID NO:4 提供的序列, 或与其至少 80% 相同的序列, 例如至少 90%, 例如至少 95% 和例如

99%相同。

金属蛋白酶优选地在选自 Lys、Arg、ArgArg、LysLys、ArgLys 和 LysArg 的氨基酸或氨基酸序列的 C-端侧切割多肽，因此，多肽优选地包含 Lys、Arg、ArgArg、LysLys、ArgLys 或 LysArg。其他序列要求也为切割所必需，然而这些可通过常规实验容易地确定。

二价阳离子优选地选自 Zn^{2+} 、 Co^{2+} 和 Mn^{2+} 。

第三方面的方法可在产生金属蛋白酶的重组宿主细胞体内进行，或在适当的反应条件下体外进行。鉴于本公布，熟练的技术人员可以容易地进行第三方面的方法。在此提供了用特定的金属蛋白酶切割多肽的体外系统的实例。在这种情况下，将多肽与在 0.1M NaH_2PO_4 (pH 4.5) 并存在 1 mM Mn^{2+} 和 1 mM 苯丁抑制素的酵母细胞粗提取物中提供的金属蛋白酶接触。在另一个实例中，金属蛋白酶可以带有适当“标记”的融合蛋白形式重组产生，例如能简化融合蛋白纯化的 His-标记。此类“标记”优选地在金属蛋白酶暴露于底物多肽之前去掉（例如通过酶促切割）。

第四方面，本发明提供了鉴定抑制包含 SEQ ID NO:1 所述序列的金属蛋白酶活性的试剂的方法，方法包括下列步骤：

- a) 在二价阳离子和适当底物存在的条件下将金属蛋白酶与试剂孵育；
- b) 确定金属蛋白酶对底物的活性；
- c) 将步骤 b) 中获得的活性与未与所述试剂孵育的对照样品的活性比较；
- d) 选择抑制金属蛋白酶活性的试剂。

底物可为被金属蛋白酶和检测到的切割事件切割的任一多肽。在此公开的一个实施例使用 CCK 作为底物，其中切割事件通过 CCK-22 的产生检测。可以容易地开发相似的鉴定法用于其他底物。

第四方面优选的实施方案中，金属蛋白酶包含如 SEQ ID NO:4 提供的序列，或包含与其至少 80% 相同的序列，例如至少 90%，例如至少 95% 和例如 99% 相同的序列。

第五方面，本发明提供了产生重组多肽的方法，方法包括在适当的条件下培养包含编码重组多肽的核酸序列的宿主细胞以产生重组多肽。以及收获重组多肽，其中所述的培养包括金属蛋白酶抑制剂的存在，所述金属蛋白酶包含 SEQ ID NO:1 提供的序列。

抑制剂优选地根据第四方面的方法鉴定。

本发明一个方面优选的特征和特性显然可以应用于本发明其他方面。

整篇说明书中，单词“包含”或例如其复数或动名词的变化形式应理解为指包括成分、整数或步骤、或成分、整数或步骤的组，但不排除任一其他成分、整数或步骤、或成分、整数或步骤的组。

本发明以下列非限制性的图和实施例形式描述。

附图简述

图 1. 缩胆囊素表达构建体。显示了 PreproMf α 1p-proCCK 融合蛋白质在融合位点附近和原始切割位点的氨基酸序列。下面显示了分泌的 CCK 主要形式的 N-端和 C-端氨基酸残基。

图 2. CCK-22 在细胞中和培养基中的成熟过程作为细胞生长的函数。表达 preproMf α 1p-proCCK 融合蛋白的 BJ2168。使用 Ab 89009 通过 RIA 测量 CCK-22 免疫反应性，并在胰蛋白酶切割后用 Ab 89009 测量总 CCK 含量。空心圆代表分泌的 CCK-22 级分而 CCK-22 胞内级分以实心三角形表示。细胞生长通过 OD₆₀₀（空心方框）测量。数据代表两个独立实验的平均值。

图 3. BJ2168 中分泌的正常 CCK 和 K $\rightarrow$ A 突变 CCK 的层析分析。pRS426 preproMf α 1p-proCCK 和 pRS426 preproMf α 1p-proCCK (K $\rightarrow$ A) 转化的酵母培养基用于 G-50 凝胶层析并用对甘氨酸延伸 CCK 特异性的 Ab7270 (A 和 C) 和 CCK-22 N-端特异性的 Ab89009 (B 和 D) 测量 CCK 免疫反应性。

图 4. 包括抑制剂和激活剂的体外蛋白酶鉴定法。CCK-22 级分计算为将使用 Ab 89009 的免疫反应性除以胰蛋白酶处理后用 Ab 89009 测量的成

热 CCK-22 和 N-端延伸 CCK-22 的总量。A, 不同抑制剂的影响。B, 蛋白酶活性用 1 mM EDTA 抑制的提取物中通过加入 1.2 mM 二价金属离子再活化蛋白酶。数据代表三个独立实验的平均值 $\pm$ SD。

图 5. 使用 Zn^{2+} 和 Mn^{2+} 的蛋白酶再活化。用 LJY123 细胞提取物进行体外蛋白酶测定, 其中活性用 1 mM EDTA (实心方框) 抑制并通过加入 1.2 mM Mn^{2+} (空心圆) 或 1.2 mM Zn^{2+} (实心圆) 再活化。活性测定为孵育 30、60 和 120 分钟后的成熟 CCK-22 级分。CCK-22 级分计算为将使用 Ab 89009 的免疫反应性除以胰蛋白酶处理后用 Ab 89009 测量的成熟 CCK-22 和 N-端延伸 CCK-22 的总量。数据代表平均值 $\pm$ SD (n=3)。

图 6. yapsin 家族成员引起的胞外 CCK-22 成熟过程。完整细胞加工胞外 CCK 的能力如“实验方法”所述分析 BY4705 和同基因 *yps1*、*yps1 yps3* 和 *yps1 yps2 yps3* 菌株。CCK-22 级分计算为将使用 Ab 89009 的免疫反应性除以胰蛋白酶处理后用 Ab 89009 测量的成熟 CCK-22 和 N-端延伸 CCK-22 的总量。数据代表平均值 $\pm$ SD (n=4)。统计使用实验方法所述的非成对 t 检验进行。(***) = $P < 0.001$, (**) = $P < 0.01$ 和 (*) = $P < 0.05$ 。

图 7. Cym1p 过量表达之后的蛋白水解增加。用 pRS425 空质粒 (A) 和含有 *CYM1* 的 pRS425 (B) 转化的 BJ2168 的细胞提取物进行体外蛋白酶测定。使用 Ab 89009 (实心方框和圆) 随时间测量 CCK-22 免疫反应性并在胰蛋白酶处理之后用 Ab 89009 测量成熟的 CCK-22 和 N-端延伸的 CCK-22 总量 (空心方框和圆)。数据代表三个独立实验的平均值 $\pm$ SD。

图 8. *KEX2* 和 *CYM1* 缺失对 proCCK 分泌和 CCK-22 成熟过程的影响。在指数期收获 proCCK 表达构建体转化的酵母细胞并收集培养基。胰蛋白酶处理之后用 Ab 89009 测量总 CCK 的胞内量 (A) 和胞外量 (B)。胞内 CCK-22 级分 (C) 和分泌的 CCK-22 级分 (D) 计算为将胰蛋白酶切割前使用 Ab 89009 的免疫反应性除以在 (A) 和 (B) 测量的 CCK 总量。*kex2*、*cym1* 和 *kex2 cym1* 菌株是 BJ2168 的同基因。数据以平均值 $\pm$ SD (n=4) 给出。统计使用非成对 t 检验进行 (***) = $P < 0.001$, (**) = $P < 0.01$ 和 (*) = $P < 0.05$ 。括号中的星号为 *kex2* 和 *kex2 cym1* 菌株间的比较。

图 9. CCK 的胞内降解依赖于 Cym1 对 CCK-22 的切割。在 BJ2168 和与 BJ2168 同基因的 *CYM1* 受破坏菌株中表达野生型 CCK、preproMf α 1p-proCCK 和 CCK 突变体 preproMf α 1p-proCCK (K $\rightarrow$ A)。指数生长过程中沉淀细胞并在胰蛋白酶和羧肽酶 B 处理后用对 Gly-延伸 CCK 的特异性抗体 Ab 7270 测量 CCK 总量(带阴影的框)。成熟 Gly-延伸 CCK 的量(取决于移位到分泌途径、Kex2p 和羧肽酶活性)测定为胰蛋白酶切割和羧肽酶 B 处理之前使用 Ab 7270 的免疫反应性。数据以平均值 $\pm$ SD (n=3)给出。

图 10. 天冬酰胺蛋白酶参与了 CCK-22 的成熟过程。在 BY4705 和同基因 yapsin 缺失菌 *YPS1*、*YPS2* 和 *YPS3* 株中表达野生型 proCCK。在指数生长过程中测量合成的 CCK-22 的胞内级分(A)和胞外级分(B)。成熟 CCK-22 级分计算为将胰蛋白酶切割前使用 Ab 89009 的免疫反应性除以在 CCK-22 的总量。数据以平均值 $\pm$ SD (n=3)表示。统计如实验方法所述使用非成对 t 检验进行。(**= P<0.001, **= P<0.01 和*= P<0.05)。

图 11. Cym1p 对 Lys 和 Arg 残基 C 端的加工。CYM1 缺失将分泌的 CCK 数量提高到野生型 CCK 和 Lys⁶¹ $\rightarrow$ Arg⁶¹ 突变体的两倍多。在 BJ2168 和 BJ2168 同基因的 *CYM1* 破坏体中表达野生型 CCK、preproMf α 1p-proCCK 和 CCK 突变体 preproMf α 1p-proCCK (Lys⁶¹ $\rightarrow$ Arg⁶¹)。指数期收获酵母细胞并收集培养基。胰蛋白酶处理后用 Ab 89009 测量总 CCK 的胞内数量(A)和胞外数量(B)。数据以平均值 $\pm$ SD (n=3)给出。

图 12. 通过质谱鉴定分泌的 proCCK 片段。CCK-数字指 C-端酰胺化 CCK。除了*表示平均值外分子量以单一同位素值给出。菌株 A, 液泡蛋白酶缺陷菌株(BJ2168); 菌株 B, *KEX1 KEX2* 破坏的同基因菌株(LJY22)。

图 13. 产生 C-端延伸 CCK (A) 和 GLP2(B)的模型。这些融合肽的表达应在 *sec61* 突变体中进行或者应去掉 α -交配因子的前序列(pre-sequence)以避免移位到内质网。显示了融合位点附近的氨基酸序列。下划线为 Gly-延伸 CCK-22 和 GLP1 的 N-端和 C-端氨基酸。

图 14. A. preproMf α 1p-proBNP 表达构建体。

B. preproMf α 1p-KREAEA-BNP-32 表达构建体。

C. preproMf α 1p-KR-BNP-32 表达构建体。

图 15. A. preproMf α 1p- proBNP 表达构建体转化 BJ2168、LJY430(*cym1* 突变体)、LJY431 (*yps1* 突变体)和 LJY432 (*cym1yps1* 双突变体)。对到达稳定期的细胞使用 proBNP N-端特异的抗体 Ab 98192 分析培养基。*cym1*、*yps1* 和 *cym1yps1* 菌株为 BJ2168 同基因。数据以平均值 $\pm$ SD (n=3)给出。统计如实验方法所述使用非成对 t 检验进行。 (**= P<0.001, ** = P<0.01 和 * = P<0.05)。括号中的星号为野生型菌株 BJ2168 和 *cym1*、BJ2168 和 *yps1*、*yps1* 和 *yps1cym1* 之间的比较。

(ns, = 不显著)。

B. 分析 *cym1* 突变体中分泌的 proCCK 片段。含 10 皮摩尔 proBNP 的培养基用于 Akta 纯化系统的 Superdex 200 柱。使用 proBNP N-端特异的抗体 98192 测量所收集级分中的 proBNP 含量。

序列表说明:

SEQ ID NO:1 –pitrilysin 蛋白酶的共有序列。

SEQ ID NO:2 – 至少为一些 pitrilysin 蛋白酶的共有序列。

SEQ ID NO:3 -至少为一些 pitrilysin 蛋白酶的共有序列。

SEQ ID NO:4 – 酿酒酵母 Cym1p (Swissprot 登录号 P32898)。

SEQ ID NO:5 -粟酒裂殖酵母 C119.7 (Swissprot 登录号 O42908)。

SEQ ID NO:6 -产气荚膜梭菌 (*Clostridium perfringens*) HypA 蛋白质(Swissprot 登录号 Q46205)。

SEQ ID NO:7 -布氏疏螺旋体 (*Borrelia burgdorferi*) 蛋白质 BB0228 (Swissprot 登录号 O51246)。

SEQ ID NO:8 -秀丽隐杆线虫 (*Caenorhabditis elegans*) C05D11.1 蛋白质(Swissprot 登录号 P48053)。

SEQ ID NO:9 – 大肠杆菌 (*E. coli*) 蛋白酶III(Swissprot 登录号

P05458)。

SEQ ID NO:10 – 大鼠 NRD 转换酶(Swissprot 登录号 P47245)。

SEQ ID NO:11 –人胰岛素崩解酶 (Swissprot 登录号 P14735)。

SEQ ID NO:12 – 拟南芥 (*Arabidopsis thaliana*) CPE (Genbank 登录号 T03302)。

SEQ ID NO:13 –人金属蛋白酶 I 部分序列 (GenBank 登录号 AAH01150) ，人金属蛋白酶 I 全长序列(Swissprot 登录号 O95204)。

SEQ ID NO:14 – 枯草芽孢杆菌 (*Bacillus subtilis*) 锌蛋白酶 ymxG (GenBank 登录号 Q04805)。

SEQ ID NO:15 –结核分枝杆菌(*Mycobacterium tuberculosis*)锌蛋白酶 Rv2782c (GenBank 登录号 O33324)。

SEQ ID NO 16 至 42 – 寡核苷酸。

SEQ ID NO 43 至 52 – 图 12 提供的序列。

SEQ ID NO 53 至 55 –图 1 提供的序列。

SEQ ID NO 56 至 65 – 寡核苷酸。

SEQ ID NO:66 – 至少为一些 pitrilysin 蛋白酶的共有序列。

SEQ ID NO:67 -至少为一些 pitrilysin 蛋白酶的共有序列。

SEQ ID NO:68 -至少为一些 pitrilysin 蛋白酶的共有序列。

发明详述

本发明提供了对表达多肽有用的宿主细胞，所述细胞经遗传操作以至少产生与亲本细胞相比降低水平的特定金属蛋白酶。这样宿主细胞能由于金属蛋白酶的蛋白水解作用得到降低或抑制从而以较高量表达目的蛋白质，这也提高了目的蛋白质的稳定性。

通过本发明的方法，金属蛋白酶的蛋白水解作用得到降低或抑制，因此改善了所获产物的稳定性。

这样本发明的一个实施方案涉及对表达目的蛋白质有用的宿主细胞，其中所述的细胞经遗传修饰以在相同条件下表达与相应未修饰细胞相比显

著降低水平的包含 HXXEH 基元 (SEQ ID NO:1) 的金属蛋白酶。

需要根据本发明下调的金属蛋白酶不共有许多序列相似性区域,最值得注意的在 N-端部分。这个区域包括保守的组氨酸,其后两个残基为一个谷氨酸和另一个组氨酸。pitrilysin 中表明此 HXXEH 基元参与酶促活性。两个组氨酸结合锌,谷氨酸为催化活性所必需。此家族的非活性成员丢失这些活性位点残基中的 1 到 3 个残基。

根据本发明下调的金属蛋白酶家族目前分类为 ME 集团 M16 家族的成员。此家族通常分为 4 个亚家族:

包含 pitrilysin 的 M16A

包含线粒体加工肽酶- β 亚基 (酿酒酵母) 的 M16B

包含 eupitrilysin (人 (*Homo sapiens*)) 的 M16C

包含痘苗病毒型金属内肽酶 (痘苗病毒) 的 M44

这些蛋白质的序列比对表明存在几个序列相似。这些序列相似高度保守并可用于区分此家族成员与非成员。

在这些序列相似中,几个独立的氨基酸高度保守并易于在距离 HXXEH 基元的特定位点处识别。

这样本发明的一个实施方案涉及这样的宿主细胞,其中金属蛋白酶包含位于 HXXEH 基元第二个组氨酸残基 C-端 70 至 80 个氨基酸之间的谷氨酸残基。

本发明另一个实施方案涉及这样的宿主细胞,其中金属蛋白酶包含位于 HXXEH 基元第一个组氨酸残基 N-端 3 个氨基酸的甘氨酸残基。

本发明另一个实施方案涉及这样的宿主细胞,其中金属蛋白酶包含位于 HXXEH 基元第二个组氨酸残基 C-端 5 个氨基酸的甘氨酸残基。

本发明另一个实施方案涉及这样的宿主细胞,其中金属蛋白酶包含位于 HXXEH 基元第二个组氨酸残基 C-端 8 个氨基酸的赖氨酸残基。

本发明的一个实施方案也涉及这样的宿主细胞,其中金属蛋白酶包含位于 HXXEH 基元第二个组氨酸残基 C-端 9 个氨基酸的酪氨酸残基。

此外,本发明涉及这样的宿主细胞,其中金属蛋白酶包含位于 HXXEH

基元第二个组氨酸残基 C-端 10 个氨基酸的脯氨酸残基。

序列相似中有几个氨基酸区域也高度保守并易于识别。这样在目前优选的实施方案中，本发明涉及这样的宿主细胞，其中金属蛋白酶包含 SEQ ID NO 2 的共有序列。

在另一个目前优选的实施方案中，本发明涉及这样的宿主细胞，其中金属蛋白酶包含 SEQ ID NO 3 的共有序列。

在目前最优选的实施方案中，本发明涉及这样的宿主细胞，其中金属蛋白酶包含位于 HXXEH 基元第二个组氨酸残基 C-端 20 和 30 个氨基酸之间的 NAXTXXXXT 基元。

在目前另一个实施方案中，本发明涉及这样的宿主细胞，其中金属蛋白酶包含 SEQ ID NO 66-68 的共有序列。本发明的一个实施方案涉及对表达目的蛋白质有用的宿主细胞，其中所述的细胞经遗传操作以表达与亲本细胞相比降低水平的金属蛋白酶，其与 SEQ ID NO: 4-15 的任何一个至少 80 % 相同。

本篇中，术语“目的蛋白质”涉及大量天然存在的极复杂底物中的任何一个，例如但不限于由通过肽键连接的氨基酸残基组成的蛋白质、酶和/或抗体。本发明优选实施方案的目的就是提供此类为宿主细胞自身产物的天然蛋白质和/或异源蛋白质、融和蛋白、重组蛋白质、真核蛋白质、原核蛋白质、溶酶体蛋白质、液泡蛋白质、前体蛋白、酶原蛋白质、前原蛋白和分泌蛋白质。

所要求权利的方法的优选实施方案由于目的蛋白质的较高产量很有利，这样，与在相同条件下培养的相应未修饰细胞中产量相比，如在此描述的经修饰宿主细胞中生产的任何目的蛋白质的量增加都在本发明的范围之内。

熟练技术人员可以评价在此所述修饰的宿主细胞中目的蛋白质提高的产量并与相应未修饰细胞的产量比较的鉴定法，通过在相同条件下培养两种不同宿主细胞并通过使用目的蛋白质特异抗体的放射免疫测定法测量产生的目的蛋白质的量。本说明书的实施例中更详细地描述了一个这样的鉴

定法。

本发明的一个实施方案涉及这样的宿主细胞，其中在相同条件下培养时，目的蛋白质总量与相应未修饰的细胞相比增加至少 5%，例如在相同条件下培养时与相应未修饰的细胞相比增加至少 10%，例如在相同条件下培养时与相应未修饰的细胞相比增加至少 20%，例如在相同条件下培养时与相应未修饰的细胞相比增加至少 50%，例如在相同条件下培养时与相应未修饰的细胞相比增加至少 100%，例如在相同条件下培养时与相应未修饰的细胞相比增加至少 200%，甚至在相同条件下培养时与相应未修饰的细胞相比增加至少 1000%。

本文中，术语“宿主细胞”涉及能产生目的蛋白质的任何细胞。这样在优选的实施方案中，宿主为原核细胞。在另一个优选的实施方案中，宿主细胞为真核细胞，例如但不限于丝状真菌细胞和非丝状真菌细胞。非局限性的例子为酵母菌株，尤其是酿酒酵母菌株。

在此描述的涉及本发明方法的所有特征也适用于涉及宿主细胞的实施方案，并且反之亦然。

本申请中描述的方法涉及宿主细胞中目的蛋白质的产量，其中所述的宿主细胞经遗传修饰，当在相同条件下培养时，这包括：

- a) 将编码目的蛋白质的核酸序列导入宿主细胞；
- b) 在适当的生长培养基中培养步骤 (a) 的宿主细胞以产生目的蛋白质；
- c) 分离目的蛋白质，

与亲本细胞相比，其表达显著降低水平的与 SEQ ID NO:4 至少 80% 相同的金属蛋白酶。

本发明的一个实施方案涉及在宿主细胞中产生目的蛋白质的方法，其中的宿主细胞通过选自基因敲除、基因破坏、随机或定点诱变、引入显性失活的金属蛋白酶、使用 dsRNA 的 RNA 干涉、催化核酸（如核酶和 DNA 酶）、反义核酸或其组合的方法遗传修饰。

在目前最优选的实施方案中，宿主细胞基本上不含任何金属蛋白酶活

性。

本发明一个优选的实施方案涉及在宿主细胞中产生目的蛋白质的方法，其中目的蛋白质为真核蛋白质，其选自胰岛素、生长激素、胰高血糖素、促生长素抑制素、干扰素、促肾上腺皮质激素、血管紧张素原、心房肽、强啡肽、内啡肽、甘丙肽、胃分泌素、胃分泌素释放肽、神经肽Y片段、胰抑制素、胰多肽、促胰液素、血管活性肠肽、生长激素释放因子、促黑激素、神经降压素、肾上腺肽、甲状旁腺激素及相关肽、生长激素抑制素及相关肽、脑钠肽、降钙素、促肾上腺皮质激素释放因子(CRF)、受可卡因和安非他明调节的转录产物(CART)、胸腺素、硬骨鱼紧张肽、胰高血糖素及胰高血糖素样肽(GLP-1和GLP-2)、生长激素抑制素、干扰素、血管内皮细胞生长因子(VEGF)、血小板衍生生长因子(PDGF)、因子VII、因子VIII、因子V、因子IX、白细胞介素、尿激酶、促红细胞生成素(EPO)、凝乳酶、组织纤维蛋白溶酶原激活剂、正辅因子2谷氨酰胺/富含Q的相关蛋白(Positive cofactor 2 glutamine/Q-rich-associated protein, PCAP)、肽酪氨酸酪氨酸(PYY)、血液生长素、增食因子、 β -新内啡肽-强啡肽前体、CCK或血清白蛋白。

本发明另一个实施方案涉及在宿主细胞中生产目的蛋白质的方法，其中目的蛋白质为选自淀粉分解酶、 α -淀粉酶、 β -淀粉酶、糖淀粉酶、 α -半乳糖苷酶、溶细胞酶、脂肪分解酶、木聚糖水解酶、蛋白水解酶、氧化还原酶、过氧化物酶、漆酶、果胶酶、或角质酶的真菌来源蛋白质。

本发明另一个优选实施方案涉及在宿主细胞中产生目的蛋白质的方法，其中目的蛋白质为选自淀粉分解酶、 α -淀粉酶、 β -淀粉酶、糖淀粉酶、 β -半乳糖苷酶、溶细胞酶、脂肪分解酶、木聚糖水解酶、蛋白水解酶、氧化还原酶、过氧化物酶、漆酶、果胶酶、或角质酶的细菌蛋白质。

本发明特别的实施方案涉及在宿主细胞中产生目的蛋白质的方法，其中目的蛋白质为前体，即酶原、杂合蛋白、以前序列或原前序列(prepro sequence)获得的或未成熟形式蛋白质。

普通分子生物学

除非特别说明, 本发明使用的重组 DNA 技术为本领域技术人员所熟知的常规方法。此类技术在整篇的下列来源中描述和解释并在此引用作为参考, 例如 J. Perbal, *A Practical Guide to Molecular Cloning*, John Wiley and Sons (1984), J. Sambrook 等, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbour Laboratory Press (1989), T.A.Brown(编辑), *Essential Molecular Biology: A Practical Approach*, 卷 1 和 2, IRL Press (1991), D.M. Glover 和 B.D. Hames(编辑), *DNA Cloning: A Practical Approach*, 卷 1-4, IRL Press (1995 和 1996), F.M. Ausubel 等 (编辑), *Current Protocols in Molecular Biology*, Greene Pub. Associates 和 Wiley-Interscience (1988, 包括至今所有更新), *Methods in Enzymology*. Vol 194. *Guide to Yeast Genetics and Molecular Biology*. (1991) Ed Gunthrie and Fink Academic Press, *Methods in Microbiology* Vol. 26. *Yeast Gene Analysis*. (1998) Brown 和 Tuite 编辑. Academic Press, Miller, J.H.(1992) *A Short Course in Bacterial Genetics* (Manual, L.编辑), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, Johnston, J.R.(1994) *Molecular Genetics of Yeast* (A Practical Approach) Oxford University Press, Oxford, 以及 *Molecular Genetics of Yeast: A Practical Approach*, Ed. J.R. Johnston 编辑, IRL Press (1994)。

多肽同源性百分比通过缺口创建罚分=5 和缺口延伸罚分=0.3 的 GAP(Needleman 和 Wunsch, 1970)分析 (GCG 程序) 确定。所查询序列长度至少为 15 个氨基酸, GAP 分析比对两序列至少 15 个氨基酸的区域。更优选地所查询序列长度至少为 50 个氨基酸, GAP 分析比对两序列至少 50 个氨基酸的区域。甚至更优选地, 所查询序列长度至少为 100 个氨基酸, GAP 分析比对两序列至少 100 个氨基酸的区域。更优选地, 所查询序列长度至少为 250 个氨基酸, GAP 分析比对两序列至少 250 个氨基酸的区域。甚至更优选地, 所查询序列长度至少为 500 个氨基酸, GAP 分析比对两序列至少 500 个氨基酸的区域。

金属蛋白酶 pitrilysin 亚家族

金属蛋白酶 pitrilysin 亚家族的特征在于存在 HXXEH 基元 (SEQ ID NO:1)。Rawlings 和 Barrett (1995)提供了这个亚家族的综述。此家族成员包括但不限于酿酒酵母 Cym1p (SEQ ID NO:4) (Swissprot 登录号 P32898)、粟酒裂殖酵母 C119.7 (SEQ ID NO:5) (Swissprot 登录号 O42908)、产气荚膜梭菌 (*Clostridium perfringens*) HypA 蛋白质 (SEQ ID NO:6) (Swissprot 登录号 Q46205)、布氏疏螺旋体蛋白质 BB0228 (SEQ ID NO:7) (Swissprot 登录号 O51246)、秀丽隐杆线虫 C05D11.1 蛋白质 (SEQ ID NO:8) (Swissprot 登录号 P48053)、大肠杆菌蛋白酶 III (也称为 pitrilysin) (SEQ ID NO:9) (Swissprot 登录号 P05458)、大鼠 NRD 转化酶 (SEQ ID NO:10) (Swissprot 登录号 P47245)、人胰岛溶素 (SEQ ID NO:11) (Swissprot 登录号 P14735)、拟南芥 CPE (SEQ ID NO:12) (Genbank 登录号 T03302)、人金属蛋白酶 I (部分) (SEQ ID NO:13) (GenBank 登录号 AAH01150) (全长序列: Swissprot 登录号 O95204)、枯草芽孢杆菌 (*Bacillus subtilis*) 锌蛋白酶 ymxG (SEQ ID NO:14) (GenBank 登录号 Q04805) 和结核分歧杆菌 (*Mycobacterium tuberculosis*) 锌蛋白酶 Rv2782c (SEQ ID NO:15) (GenBank 登录号 O33324)。对大肠杆菌金属蛋白酶 III (SEQ ID NO:9), 表明 SEQ ID NO:1 中的组氨酸残基和谷氨酸-169 参与二价阳离子的结合同时与组氨酸残基相邻的谷氨酸残基为催化残基。

编码 pitrilysin 金属蛋白酶的基因可通过杂交编码全部或部分金属蛋白酶的核酸序列的筛选方法很容易地鉴定, 例如使用合成寡核苷酸探针, 其可基于 cDNA 序列, 如根据常规技术制备的编码 SEQ ID NO: 4 到 15 的金属蛋白酶中的任何一个的核苷酸序列。

遗传操作

本发明经遗传操作以产生降低水平的特定金属蛋白酶的宿主细胞可使用本领域技术人员已知的常规重组 DNA 技术来修饰。引起金属蛋白酶产生的基因序列可以失活或全部消除。

在特定的实施方案中, 本发明的宿主细胞在金属蛋白酶基因的编码或调节区域进行遗传操作。已知并有用的技术包括但不限于基因敲除、基因

破坏、随机或定点诱变、引入显性失活金属蛋白酶、使用 dsRNA 的 RNA 干涉、催化核酸（例如核酶和 DNA 酶）和反义核酸或其组合。

诱变使用适当的物理或化学诱变剂进行。适于此目的的物理或化学诱变剂的例子包括紫外线（UV）照射、羟胺、N-甲基-N'-硝基-N-亚硝基胍(MNNG)、O-甲基羟胺、硝酸、甲基磺酸乙酯(EMS)、亚硫酸氢钠、蚁酸和核苷类似物。当使用此类诱变剂时，一般通过在适于诱变发生的条件下在所选诱变剂中孵育待诱变细胞，并选择显著降低金属蛋白酶产量的突变细胞进行诱变。

遗传操作也可通过引入、替代或去除金属蛋白酶编码序列或其转录或翻译所需的调节元件中一个或多个核苷酸来完成。例如可插入或删除核苷酸以造成终止密码子的引入，起始密码子的去除或开放阅读框的变化。结构序列或调节元件的修饰或失活可根据本领域已知的方法通过定点诱变或 PCR 产生诱变而完成。

失活或降低宿主细胞金属蛋白酶产量的便利方法基于基因间断原则。此方法包括使用与待破坏内源基因或基因片段对应的 DNA 序列。DNA 序列为在体外突变成缺陷基因并转化宿主细胞。缺陷基因通过同源重组替代内源基因或基因片段。可能需要编码用于重组体选择的标记的缺陷基因或基因片段，其中编码金属蛋白酶的基因已经得到修饰或破坏。

在此使用的术语“反义”指与特定核酸序列互补的核苷酸序列。反义分子可通过任何方法生产，其包括通过将目的基因与允许互补链合成的病毒启动子反向连接来合成。一旦导入宿主细胞，其经转录链与细胞产生的编码金属蛋白酶的天然序列结合形成双链体。这些双链体阻止进一步的转录或者进一步的翻译。突变表型可以这种方式产生。

术语“催化核酸”指特异识别特定底物并催化此底物化学修饰的 DNA 分子、含 DNA 的分子（也是本领域已知的“脱氧核酶”）、RNA 或者含 RNA 的分子（也是本领域已知的“核酶”）。催化核酸中的核酸碱基可以为碱基 A、C、G、T 和 U，也可为其衍生物。这些碱基的衍生物为本领域所熟知。

催化核酸一般地含有特异识别目的核酸的反义序列和切割性酶促活性的核酸，在此也称为“催化结构域”。本发明中尤其有用的核酶类型为锤头状核酶(Haseloff 和 Gerlach, 1988)和发卡状核酶(Shippy 等,1999)。

对本发明方法有用的核酶和编码核酶的 DNA 可使用本领域熟知的方法化学合成。核酶也可以从可操作地与 RNA 聚合酶启动子连接的 DNA 分子(一旦转录产生 RNA 分子)制备，其中所述启动子例如 T7RNA 聚合酶或 SP6RNA 聚合酶的启动子。当载体还含有可操作地与 DNA 分子连接的 RNA 聚合酶启动子时，可通过与 RNA 聚合酶和核苷酸孵育体外产生核酶。在独立的实施方案中，DNA 可插入表达盒或转录盒。合成之后，RNA 分子可通过与具有稳定核酶并使其抵抗 RNA 酶能力的 DNA 分子连接来修饰。作为一种选择，核酶可修饰成用于脂质体递送系统中的硫代磷酸(phosphothio)类似物。这种修饰也使核酶具有抵抗核酸内切酶的活性。

dsRNA (RNAi)对特异抑制特定蛋白质产生尤其有用。此技术依赖于含有与目的基因的 mRNA 基本上一致序列的 dsRNA 分子的存在，在这种情况下所述 mRNA 编码金属蛋白酶。dsRNA 可方便地在重组载体或宿主细胞中的一个开放阅读框中产生，其中有义与反义序列与无关序列相邻，其能使有义与反义序列杂交以形成 dsRNA 分子，而无关序列形成环状结构。用于遗传操作的适当 dsRNA 分子的设计和产生为本领域技术人员能力范围之内所熟知的，尤其见 Waterhouse 等(1998)，Elbashir 等 (2001)，WO 99/32619，WO 99/53050，WO 99/49029 和 WO 01/34815。

本发明的宿主细胞由于遗传操作表达显著降低水平的金属蛋白酶。在优选的实施方案中宿主细胞表达的金属蛋白酶水平降低大于约 25%，例如大于约 30%，例如大于约 35%，例如大于约 40%，例如大于约 45%，例如大于约 50%，例如大于约 55%，例如大于约 60%，例如大于约 65%，例如大于约 70%，例如大于约 75%，例如大于约 80%，例如大于约 85%，例如大于约 90%，例如大于约 95%，例如大于约 98%和例如大于约 99%。

在目前最优选的实施方案中，宿主细胞表达的产物基本不具有任何所述的金属蛋白酶活性。

在本文中，术语“基本不具有”指这样的宿主，其中所述宿主细胞表达的金属蛋白酶降低到所述金属蛋白酶的功能对目的蛋白质产生没有生物学上使所述蛋白质显著降低的影响。

目的蛋白质

术语“多肽”、“蛋白质”和“肽”在此可替换地使用并且在本文中涉及大量天然存在的极复杂物质中的任何一个，例如但不限于由肽键连接的氨基酸残基组成的酶或抗体，其含有碳元素、氢元素、氮元素、氧元素，通常含有硫元素并且偶尔含有其他元素例如但不限于磷或铁；为所有活细胞的必需成分；在植物中自原材料天然合成但在动物中以单个氨基酸同化；虽然对许多所述物质已得到结晶但天然状态下其为酸性、碱性并通常为胶体，且可被酸、碱、蛋白水解酶和腐败细菌水解成多肽、简单肽并最终水解成 α -氨基酸。

正如在此定义的，“重组多肽”为宿主细胞的非天然蛋白质，或天然蛋白质中已进行修饰以改变天然序列，或对天然蛋白质表达需要的天然调节序列例如启动子、核酶结合位点等等的操作或者其他通过重组 DNA 技术对宿主细胞的操作，造成表达得到量上的改变的天然蛋白质。

由于所述金属蛋白酶活性的消失或降低，宿主细胞表达的至少一部分多肽可能为前体蛋白质，例如酶原、杂合蛋白质、以前序列或原前序列或未成熟形式获得的蛋白质。

在更特定的实施方案中，重组多肽为真核来源如胰岛素、促肾上腺皮质激素、血管紧张素原、心房钠尿肽、强啡肽、内啡肽、甘丙肽、胃分泌素、胃分泌素释放肽、神经肽 Y 片段、胰抑制素、胰多肽、促胰液素、血管活性肠肽、生长激素释放因子、促黑激素、神经降压素、肾上腺肽、甲状旁腺激素及相关肽、生长激素抑制素及相关肽、脑钠肽、降钙素、促肾上腺皮质激素释放因子(CRF)、受可卡因和安非他明调节的转录产物(CART)、胸腺素、硬骨鱼紧张肽、胰高血糖素及胰高血糖素样肽 (GLP-1 和 GLP-2)、生长激素抑制素、干扰素、血管内皮细胞生长因子(VEGF)、血小板衍生生长因子 (PDGF)、因子VII、因子VIII、因子V、因子IX、白细

胞介素、尿激酶、促红细胞生成素 (EPO)、凝乳酶、组织纤维蛋白溶酶原激活剂、CCK 或血清白蛋白。

特别要提到“胰高血糖素及胰高血糖素样肽”，在此使用的此术语指人或其他动物来源的和重组或半合成来源的多肽，其包括胰高血糖素家族全部成员，例如 GRPP(胰高血糖素样肽相关多肽)、胰高血糖素、GLP-1(胰高血糖素样肽 1)和 GLP-2(胰高血糖素样肽 2)，其包括截短和/或 N-端延伸的形式例如 GLP-1(7-36)并包括类似物 GLP-1(7-35)R36A GLP-2 F22Y、GLP-2 A19T+34Y、GLP2 A2G 和 GLP-2 A19T，以及其他具有 1 到 3 个氨基酸改变、添加和/或缺失的类似物。

宿主细胞及由其表达的重组多肽

本发明使用的宿主细胞可为原核或真核宿主细胞。本发明使用的真核宿主细胞可为例如真菌、哺乳动物、植物或昆虫细胞。宿主细胞优选地为酵母细胞。

为产生目的多肽，本发明的宿主细胞包含编码重组多肽的核酸序列和指导目的产物表达的调节序列例如包含如转录、翻译和终止必需的核苷酸序列的区域。本发明宿主细胞的遗传设计可由本领域技术人员使用转化或转染宿主细胞的常规重组 DNA 技术完成。

宿主细胞优选地通过本领域已知的导入例如质粒或病毒载体中的、包含编码重组多肽的核酸的适当表达盒方法来修饰。

表达盒可通过多种技术引入宿主细胞，包括但不限于作为自主复制质粒或整合入染色体。表达盒可含有调节序列，例如转录控制序列、翻译控制序列、复制起点和与重组细胞相容以及控制编码重组多肽的核酸表达的其他调节序列。本发明的重组核酸分子尤其包括转录控制序列。转录控制序列为控制转录起始、延伸和终止的序列。尤其重要的转录控制序列为控制转录起始的序列，例如启动子、增强子、操纵子和阻遏子序列。适当的转录控制序列包括至少可在本发明的一种重组细胞中发挥功能的任何转录控制序列。多种这样的转录控制序列为本领域技术人员已知。优选的转录控制序列包括在细菌、酵母、节肢动物和哺乳动物细胞中发挥功能的序列，

例如但不限于 tac、lac、trp、trc、oxy-pro、omp/lpp、rrnB、 λ 噬菌体、T7 噬菌体、T7lac、T3 噬菌体、SP6 噬菌体、SP01 噬菌体、金属硫蛋白、 α -交配因子、毕赤酵母醇氧化酶、甲病毒亚基因组启动子（例如辛德毕斯病毒亚基因组启动子）、抗生素抗性基因、杆状病毒、玉米夜蛾（*Heliothis Zea*）昆虫病毒、痘苗病毒、疱疹病毒、浣熊痘病毒、其他痘病毒、腺病毒、巨细胞病毒（如中早期启动子）、猴病毒 40、逆转录病毒、肌动蛋白、逆转录病毒长末端重复、劳斯肉瘤病毒、热休克、磷酸盐和硝酸盐转录控制序列和能在原核或真核细胞中控制基因表达的其他序列。本发明的转录控制序列最优选地为天然存在的酵母相关控制序列。酿酒酵母适当的启动子包括 MF α 1 启动子、半乳糖诱导型启动子例如 GAL1、GAL7 和 GAL10 启动子、包括 TPI1 和 PGK1 启动子、TRP1 启动子、CYC1 启动子、CUP1 启动子、PHOS 启动子、ADH1 启动子和 HSP 启动子的糖酵解酶启动子。毕赤酵母属中适当的启动子为 AOX1（甲醇利用）启动子。

本发明的重组多肽也可含有（a）分泌信号，使表达的多肽从产生所述多肽的细胞分泌和/或（b）引起融合蛋白表达的融合序列。适当的信号片段的例子包括能指导融合蛋白分泌的任何信号片段。优选的信号片段包括但不限于组织纤溶酶原激活剂（t-PA）、干扰素、白细胞介素、生长激素、组织相容性和病毒包膜糖蛋白信号片段、来源于真菌淀粉葡萄糖苷酶（AG）基因例如 galA 的前导序列（例如来源于曲霉属某些种（*Aspergillus sp.*）的 18 个和 24 个氨基酸形式）、来源于酵母如芽殖酵母某些种和克鲁维酵母某些种的 α -因子基因、裂殖酵母某些种的 P-因子和芽孢杆菌某些种的 α -淀粉酶基因以及天然信号序列。

克隆载体也可包含选择性标记，例如其产物弥补宿主细胞缺陷的基因或给予抗生素抗性的基因例如氨苄青霉素、卡那霉素、氯霉素或四环素抗性的基因。

分别用于连接本发明的 DNA 构建体、启动子、终止子和其他元件以及将其插入含有复制必需信息的适当克隆载体的方法为本领域技术人员熟知。

可使用重组 DNA 技术通过如操作宿主细胞中核酸分子的拷贝数量、操作转录核酸分子的效率、操作得到的转录物的翻译效率和操作翻译后修饰的效率提高经转化核酸分子的表达。对本发明方法有用的增加核酸分子表达的重组技术包括但不限于可操纵地将核酸分子与高拷贝数质粒连接、将核酸分子整合入一条或多条宿主细胞染色体、在质粒中加入载体稳定序列、替代或修饰转录控制信号（如启动子、操纵子和增强子）、替代或修饰翻译控制信号（如核糖体结合位点、SD 序列）、修饰核酸分子使其与宿主细胞密码子使用对应和去掉使转录物不稳定的序列。本发明表达的重组多肽的活性可通过片段化、修饰或衍生化编码此类蛋白质的多聚核苷酸分子来提高。

产生重组多肽的方法

已经用编码重组多肽的核酸转染或转化的宿主细胞在本领域技术人员查阅在此公开的资料已知的适当条件下培养充足的时间以允许多肽的产生，并且优选地允许多肽的分泌，然后收获目的产物。

此外，由于金属蛋白酶活性降低，宿主细胞表达的重组多肽可以前体蛋白质形式获得，即酶原、杂合蛋白质、以前序列或原前序列或未成熟形式获得的蛋白质

用于培养的肉汤或培养基为适于提到的宿主细胞生长的任何常规培养基，并可根据现有技术的原理合成。培养基优选地含有碳源和氮源以及其他无机盐。适当的培养基如基本培养基或复合培养基可获自供应商或可根据出版的方法制备。

谈及酵母宿主细胞，在二倍体酵母培养中产生异源多肽经常是有利的，因为可能的遗传缺陷可在单倍体酵母培养如生产规模的连续发酵中为表型形式表达，还因为其产量更高。Molecular Genetics of Yeast: A Practical Approach, Ed. J.R. Johnston, IRL Press (1994)中描述了在酵母宿主细胞中产生重组多肽，其在此引用作为参考。

培养之后，蛋白质通过从培养肉汤中分离和纯化蛋白质的常规方法收获。众所周知的纯化方法包括通过离心或过滤从培养基中分离细胞、通过

盐如硫酸铵方法沉淀蛋白质成分和层析方法如离子交换层析、凝胶过滤层析、亲和层析等。

本发明通过显示 CCK 和 proBNP 总量与未修饰宿主细胞相比分别增加约 60 % 和 100 % 作为例证。所述实例由此表明本发明的宿主细胞能增加多种蛋白质的产量

此外，实施例公开其他蛋白酶的额外破坏提高目的蛋白质的产量，例如 *KEX2* 和 *CYMI* 的破坏造成 CCK 生产中产量上的加性效应（几乎 100 %）。

此外，熟练的技术人员应理解一些所述基元中的特殊氨基酸，尤其是组氨酸、谷氨酸、天冬氨酸或赖氨酸通过作为金属配基起作用而为金属蛋白酶功能所必需。

此外，应认识到本发明的金属蛋白酶在本文中作为 pitrilysin 家族、胰岛素降解酶家族、胰岛素酶家族、 inverzincin 家族和 ME 集团 M16 亚家族成员广泛地做过解释。

因此，应理解与任何这些注释的不同家族相关的上述任何特征和/或方面同样可应用于文中所述的金属蛋白酶，其均包括 HXXEH 基元。

然而，请注意单独的 pitrilysin（其后不接家族）是指大肠杆菌金属蛋白酶 ME 集团的特定成员。

实施例

材料和方法

酵母菌株和生长条件

使用的酵母菌株列于表 I 中。菌株构建使用两步基因破坏技术 (Rothstein, 1991) 或 (Brachmann 等, 1998) 基于 PCR 的方法进行。培养基购自 Difco，氨基酸和其他添加成分购自 Sigma-Aldrich。酵母细胞 30°C 下在 YPD(1% 酵母提取物, 2% 蛋白胨和 2% 葡萄糖) 或在基于酵母氮碱的具有硫酸铵、丁二酸、NaOH 和适当氨基酸的合成完全培养基 (SC) 中生长。使用改良的所述醋酸锂方法 (Gietz 等, 1995) 用线性 DNA 或质粒进行

转化。由于 CCK-22 生物合成的连续性，对异源表达的 CCK 的分析是在生长至对数期的酵母中进行的，其与来自稳定期酵母的结果相反（图 2）。从细胞提取物和 5 A₆₀₀ 单位细胞/毫升合成的完全培养基中分析 proCCK 加工。细胞生长通过 600nm 吸光度跟踪。

表 I. 本研究使用的酿酒酵母菌株。假定的金属蛋白酶无效突变体通过基因型的 ORF 命名，（*）代表线粒体蛋白酶。

菌株	基因型	来源
BY4705	<i>MATα ade2Δ::hisG his3Δ200 leu2Δ0 lys2Δ0 ura3Δ0</i>	(Brachmann 等, 1998)
LJY13	<i>MATα ade2Δ::hisG his3Δ200 leu2Δ0 lys2Δ0 ura3Δ0 yps1::TRP1</i>	本研究
LJY14	<i>MATα ade2Δ::hisG his3Δ200 leu2Δ0 lys2Δ0 ura3Δ0 yps1::TRP1 yps3::LEU2</i>	本研究
LJY15	<i>MATα ade2Δ::hisG his3Δ200 leu2Δ0 lys2Δ0 ura3Δ0 yps1::TRP1 yps3::LEU2 yps2::URA3</i>	本研究
BJ2168	<i>MATa prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2</i>	(Jones, 1991)
LJY21	<i>MATa prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 kex1::LEU2</i>	本研究
LJY22	<i>MATa prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 kex1::LEU2 kex2::TRP1</i>	本研究
LJY23	<i>MATa prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 kex2::TRP1</i>	本研究
LJY122	<i>MATα ape1::KANMX ape2::LYS2 his3Δ0 leu2Δ0 lys2Δ0 ura3Δ0</i>	本研究
LJY123	<i>MATα ape1::KANMX ape2::LYS2 ape3::LEU2 his3Δ0 leu2Δ0 lys2Δ0 ura3Δ0</i>	本研究
LJY201	<i>MATa prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52 gal2 axl1::LEU2</i>	本研究
LJY202	<i>MATa prc1-407 prb1-1122 pep4-3 leu2 trp1</i>	本研究

	<i>ura3-52 gal2 ste24::LEU2</i>				
LJY203	<i>MATa prc1-407 prb1-1122 pep4-3 leu2 trp1</i>	本研究			
	<i>ura3-52 gal2 prd1:LEU2</i>				
LJY204	<i>MATa prc1-407 prb1-1122 pep4-3 leu2 trp1</i>	本研究			
	<i>ura3-52 gal2 yil108w::LEU2</i>				
Y15298	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i>	Euroscarf			
	<i>ste23::KANMX</i>				
Y11874	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i>	Euroscarf			
	<i>aap1::KANMX</i>				
Y10148	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i>	Euroscarf			
(*)	<i>afg3::KANMX</i>				
Y14953	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i>	Euroscarf			
	<i>apel::KANMX</i>				
Y16224	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i>	Euroscarf			
(*)	<i>rca1::KANMX</i>				
Y14984	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i>	Euroscarf			
(*)	<i>mip1::KANMX</i>				
Y17144	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i>	Euroscarf			
	<i>yme1::KANMX</i>				
Y13211	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i>	Euroscarf			
	<i>ybr074w::KANMX</i>				
Y13801	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i>	Euroscarf			
	<i>ydl104c::KANMX</i>				
Y11941	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i>	Euroscarf			
	<i>yhr113w::KANMX</i>				
Y11960	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i>	Euroscarf			
	<i>yhr132c::KANMX</i>				
Y12296	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i>	Euroscarf			
	<i>yil137c::KANMX</i>				
Y15370	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i>	Euroscarf			
	<i>ynl045w::KANMX</i>				
10864B	<i>MATα ura3-Δ851 leu2-Δ1 his3Δ200 lys2Δ202</i>	Euroscarf			

	<i>ykr035c-ykr038c::URA3</i>	
Y11749	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i> <i>yol057w::KANMX</i>	Euroscarf
Y16248	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i> <i>yol098c::KANMX</i>	Euroscarf
10231B	<i>MATα his3Δ200 leu2Δ1 trp1Δ63 ura3-52</i> <i>yol154w(4,744)::KANMX</i>	Euroscarf
Y14266	<i>MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i> <i>ydr430c::KANMX</i>	Euroscarf
LJY430	<i>MATa prc1-407 prb1-1122 pep4-3 leu2 trp1</i> <i>ura3-52 gal2 ydr430c::LEU2</i>	本研究
LJY432	<i>MATa prc1-407 prb1-1122 pep4-3 leu2 trp1</i> <i>ura3-52 gal2 kex2::TRP1 ydr430c::LEU2</i>	本研究

DNA 提取和扩增

酵母基因组 DNA 按所述(Philippsen 等, 1991)分离。聚合酶链式反应 (PCR)使用 *Pwo* 聚合酶或基于 *Taq*、*Pwo* 和 *Pfu* 聚合酶(扩展长距离 PCR 试剂盒, XL-PCR)的混合酶, 均为 Roche 生产。所有 PCR 产品通过琼脂糖凝胶电泳观察, PCR 产物或者用凝胶提取试剂盒 (Qiagen) 从凝胶中纯化或者通过 PCR 纯化旋转柱 (GENOMED) 从反应混合物中纯化。基于 PCR 的一步基因破坏使用 50ng pRS400 系列质粒(Brachmann 等, 1998)作为模板进行。使用具有 20 个质粒核苷酸和 50 个与目的基因相邻的核苷酸的寡核苷酸进行标记的扩增(表 2)。所有其他 DNA 操作根据标准方法进行 (Sambrook 等, 1989)。

质粒构建

proCCK 的表达在 pRS426 [2μ *URA3*] (Brachmann 等, 1998) 中使用磷酸甘油酸激酶启动子(*PGK1p*)和终止子(*PGK1t*)进行。*PGK1* 启动子使用 100ng 酵母基因组 DNA 作为模板用 *PGK1p* 5'*HindIII* 和 *PGK1p* 3'*MCS* 扩增(表 2)并随后克隆到 pGEM-11 (Promega)的 *HindIII* 和 *SacI* 限制性酶位点中。终止子用 *PGK1t* 5'*BglII* 和 *PGK1t* 3'*SacI* 扩增(表 2)并连接到在

SacI 和 *EcoRI* 限制性酶位点处含有启动子的质粒。然后此构建体 pGEM-11 *PGK1pMCSPGK1t* 含有 *PGK1*-启动子、带有限制性酶位点 *EcoRI*、*BamHI*、*XbaI* 和 *BglII* 的多克隆位点 (MCS) 然后为 *PGK1* 终止子。preproMf α 1p-proCCK 融合物(Rourke 等, 1997)(图 1)亚克隆到 pGEM-11 *PGK1pMCSPGK1t* 的 *EcoRI* 和 *XbaI* 位点处, 最后全长基因克隆到 pRS423 和 pRS426 中以分别完成酵母 CCK 表达构建体 pRS423 preproMf α 1p-proCCK 和 pRS426 preproMf α 1p-proCCK。多拷贝质粒上 *CYMI* 的表达通过扩增 *CYMI* 的开放阅读框(ORF)以及额外的 5'端 926bp 和 3'端 703bp 构建。扩增使用 100ng 酵母基因组 DNA 和寡核苷酸 *CYMI* 5'*ApaI* 和 *CYMI* 3'*XhoI* (表 2) 通过 XL-PCR 进行。PCR 产物在旋转柱上纯化并随后克隆入 pRS425 的 *ApaI* 和 *XhoI* 限制性酶位点中。

表 2. 所用的寡核苷酸

寡聚物	寡核苷酸序列 (5'-3')	用途
<i>PGK1p5'HindIII</i>	AATAGAAGCTTGTCTGACTGAT CTATCCAAAAGTCTG (SEQ ID NO: 16)	表达构建体
<i>PGK1p3'MCS</i>	AAAAGAGCTCGGCCAGATCT TCTAGAGGATCCAA GAATTCTGTTTTATATTTGTT GTAAAAAGTAG (SEQ ID NO: 17)	表达构建体
<i>PGK1t5'BglII</i>	TTTTGAATTCCAAGATCTCCC ATGTCTCTACTGGTGG (SEQ ID NO: 18)	表达构建体
<i>PGK1t3'SacI</i>	CCCCGAGCTCGTCGACCCTT CTCGAAAGCTTTAACGAACG C (SEQ ID NO: 19)	表达构建体
5'MF α 1- <i>EcoRI</i>	TTTTGAATTCAAAGAATGAGA TTTCCTTCAATTTTACTGCA G (SEQ ID NO: 20)	preproMf α 1p-pro CCK
CCK3'- <i>XbaI</i>	TTTTTCTAGACTAGGAGGGGT	preproMf α 1p-pro

	ACTCATACTCCTCGGC (SEQ ID NO: 21)	CCK
CCK-22 K→A (S)	CGAATGTCCATCGTT <u>GCGAA</u> CCTGCAGAACCTG (SEQ ID NO: 22)	Lys ⁶¹ →Ala ⁶¹ 突变
CCK-22 K→A (AS)	CAGGTTCTGCAGGTT <u>CCTAAC</u> GATGGACATTCG (SEQ ID NO: 23)	Lys ⁶¹ →Ala ⁶¹ 突变
CCK-22 K→R (S)	CGAATGTCCATCGTT <u>AGGAA</u> CCTGCAGAACCTG (SEQ ID NO: 24)	Lys ⁶¹ →Arg ⁶¹ 突变
CCK-22 K→R (AS)	CAGGTTCTGCAGGTT <u>CCTAAC</u> GATGGACATTCG (SEQ ID NO: 25)	Lys ⁶¹ →Arg ⁶¹ 突变
CCK-22 seq	TCGCAGAGAACGGATGGC (SEQ ID NO: 26)	测序
CYM15' <i>Apa</i> I	TTTTGGGCCCTTCATGGTGAT ACGGTATCTCTTGGC (SEQ ID NO: 27)	克隆 <i>CYM1</i>
CYM13' <i>Xho</i> I	TTTTCTCGAGAAGGTGGAAC ATACTGCCCTGGGATGG (SEQ ID NO: 28)	克隆 <i>CYM1</i>
<i>KEX25'</i>	TTTTGAGCTCGTTTAGGAAAC GTCCTTGGCGGAGATGC (SEQ ID NO: 29)	克隆 <i>KEX2</i>
<i>KEX23'</i>	TTTTTCTAGACACTGCGAATC CATGGTATAAACCAAAACC (SEQ ID NO: 30)	<i>KEX2</i> 克隆
<i>KEX2DC5'</i>	GTCGTTGTTTCATGGACATACC TCC (SEQ ID NO: 31)	克隆 Δ <i>kex2</i>
<i>KEX2DC3'</i>	TACAAATGTTCTTCTGCCATT TCTGG (SEQ ID NO: 32)	控制 Δ <i>kex2</i>
<i>TRP15' Nde</i> I	GGTTCATATGCGCCGGAGCT	克隆 <i>TRP1</i>

	CCTCGACAGCAG (SEQ ID NO: 33)	
<i>TRP13' AvrII</i>	GGTTCCTAGGATCCGCAAGT TTGATTCCATTGCGGTG (SEQ ID NO: 34)	克隆 <i>TRP1</i>
<i>KEX15'GD400</i>	TTAAAGAGTACCTTGGCTATA GAATACCGTAGAGATAAAGA CCTGAATAGAGATTGTACTGA GAGTGCAC (SEQ ID NO: 35)	删除 <i>KEX1</i>
<i>KEX13'GD400</i>	AGGTATTATAACTATTTTTCT GTATTTTTTATATATTTTTATT TGCCAAGCTGTGCGGTATTTC ACACCG (SEQ ID NO: 36)	删除 <i>KEX1</i>
<i>KEX15'DC400</i>	CTTTGGTTAAAGAGTACCTTG GC (SEQ ID NO: 37)	控制 $\Delta kex1$
<i>KEX13'DC400</i>	TACTACGAAAAGCGTGTGCG AGG (SEQ ID NO: 38)	控制 $\Delta kex1$
<i>CYM15'GD400</i>	TAGAAGGCTACTCAAAAGAA TAAAGTTACTATAAAATATAC TGCGGTATATAGATTGTACTG AGAGTGCAC (SEQ ID NO: 39)	删除 <i>CYM1</i>
<i>CYM13'GD400</i>	GATCGGCAAGAACTTTGAA GCAGTATATTTACAGGATTAA ATTATATATCTGTGCGGTATT TCACACCG (SEQ ID NO: 40)	删除 <i>CYM1</i>
<i>CYM15'DC400</i>	CGGAGGGGCTCTATGATAAA GG (SEQ ID NO: 41)	控制 $\Delta cym1$
<i>CYM13'DC400</i>	GAGTAACTAGGGCTTCTCTTC CC (SEQ ID NO: 42)	控制 $\Delta cym1$

认为对蛋白水解 proCCK 释放 CCK-22 很关键的 Lys⁶¹ 残基通过定点

诱变用丙氨酸替换(Horton等,1993)。替换通过使用 *Pwo* 聚合酶(Boehringer Mannheim)的 PCR 进行,其中两个产物用寡核苷酸组 *PGK1p5' HindIII / CCK-22 K→A* (反义) 和 *CCK-22 K→A* (有义) / *PGK1t3' SacI* (表2)和每个反应 50ng pRS426 preproMfα1p-proCCK 作为模板扩增。两产物进行琼脂糖凝胶电泳,切出每一产物的约 1mm² 在第三次 PCR 反应中直接作为模板使用。在这次反应中,编码融合蛋白的全长 cDNA 使用 *PGK1p5' HindIII* 和 *PGK1t3' SacI* 扩增(表2)。PCR 产物亚克隆到 pCR-Blunt II (Invitrogen) 中并用 CCK 特异引物 CCK-22seq. 测序。最后, *PGK1p* preproMfα1p-proCCK (K→A) *PGK1t* 产物克隆到 pRS426 的 *HindIII* 和 *SacI* 位点中构建表达质粒 proCCK (K→A)。用精氨酸替代赖氨酸如上所述通过用 CCK-22 K→R (反义) 和 CCK-22 K→R (有义)替换 CCK 特异引物构建 proCCK (K→R)载体进行。

菌株构建

部分 *KEX2* 破坏的构建通过使用 100ng 酵母基因组 DNA 作为模板和寡核苷酸 *KEX2 5'SacI* 和 *KEX2 3'XbaI*(表2)的 XL-PCR 扩增全长 *KEX2* 基因 ORF 每一位点的 1000bp 在 BJ2168 中进行。PCR 产物在旋转柱上纯化并克隆到 pCR-Blunt II 中(Invitrogen)。*TRP1* 的扩增通过使用 *TRP1 5'NdeI* 和 *TRP1 3'AvrII* (表2) 在 ORF5'端 925bp 引入 *NdeI* 位点和终止子 3'端 212bp 引入 *AvrII* 位点进行 XL-PCR。纯化 PCR 产物并亚克隆到去除 *KEX2* 2018bp 和启动子 170bp 的 *KEX2* 中的 *NdeI* 和 *AvrII* 位点处。*Kex2::TRP1* 构建体使用限制性酶 *NotI* 和 *SpeI* 从 pCR-Blunt II 切下并随后转化 BJ2168。在 SC-Trp 平板上选择转化体,然后通过使用覆盖全部标记加上 *KEX2* 每一位点的额外 1200bp 的 *kex2 DC5'*和 *kex2 DC3'* (表2)寡核苷酸进行菌落 PCR 筛选,以检测正确的整合。*kex2 kex1* 菌株的构建通过使用 *LEU2* 标记的两步基因破坏技术(Rothstein, 1991)进行。*LEU2* 的扩增通过使用 50ng pRS405 作为模板以及 *kex15'GD400* 和 *kex13'GD400* (表2)的 XL-PCR 进行。PCR 产物用 PCR 纯化旋转柱纯化 (GENPMED) 并随后转化 LJY23。在 SC-Leu 平板上选择转化体,正确

的整合通过使用 *kex15'*DC 以及 *kex13'*DC (表 2)的基于 PCR 的菌落筛选来检测。

BJ2168 背景下的 $\Delta cym1::LEU2$ (LJY430)菌株通过上述用于 $\Delta kex1$ 菌株的一步基因破坏技术构建, 其使用寡核苷酸 *CYM15'*GD400 和 *CYM13'*GD400(表 2)用于基因破坏, 使用 *cym15'*DC 和 *cym13'*DC (表 2)用于破坏控制。通过此方法建立的所有无效突变体用针对与 5'和 3'UTR 相邻的 50 碱基并带有与含有多种标记的 pRS400 系列载体特异的 3'端设计的寡核苷酸制备(Brachmann 等, 1998)。在适当的琼脂板上选择转化体然后通过使用覆盖全部标记加上每一位点的额外 200bp 的寡核苷酸进行菌落 PCR 筛选, 以检测正确的整合。表 2 只显示了不位于上述位置的寡核苷酸。

使用 PCR 破坏技术(Brachmann 等, 1998)和以 pRS405 [*LEU2*]作为模板在 BJ2168 中进行 *STE24*、*AXL1*、*PRD1* 和 *YIL108w* 的基因缺失。

含有 *APE1*、-2 和-3 基因缺失的 LJY123 用 PCR 破坏技术从 Y14953 衍生(Brachmann 等, 1998)。*APE2* 最初用 *LYS2* (pRS317 [*cen*; *LYS2*])代替, 其中 PCR 产物在转化之前从琼脂糖凝胶中纯化并且 *APE3* 用 *LEU2* 标记(pRS405 [*LEU2*])代替。

yps1 yps2 yps3 三联突变体 (LJY15) 在 BY4705 中使用 PCR 破坏技术构建(Brachmann 等, 1998)。*YPS1* 的 ORF 最初通过插入 *TRP1* 基因座 (pRS404) 来删除, 以产生 LJY13。然后此菌株用作宿主, 通过插入 *LEU2* 标记(pRS405)删除 *YPS3*, 最后通过插入 pRS406[*URA3*]扩增的 *URA3* 标记删除 *YPS2*, 从而构建 LJY15 (表 I)。

CCK 和 *CYM1* 的表达

人 proCCK 以酵母 α -交配因子前原前导序列和 proCCK 之间的融合蛋白形式表达(preproMf α 1p-proCCK)。融合构建体在多拷贝质粒上表达, 具有自磷酸甘油酸激酶启动子进行的组成型基因转录。“ProCCK 表达”指使用 pRS426 preproMf α 1p-proCCK 的表达, 其用于除了 BY4705 和从 pRS423 preproMf α 1p-proCCK 表达 proCCK 的同基因 yapsins 缺失菌株之外的所有酵母菌株中。*CYM1* 表达通过自身启动子驱动。所用的质粒构

建体和寡核苷酸列于表 2。

酶促处理

使用 1mg/ml 胰蛋白酶(Worthington Biochemical Corporation)在 50mM 磷酸钠缓冲液 (pH7.5) 中室温 30 分钟进行胰蛋白酶处理并通过浸入沸水 10 分钟终止处理。在 0.1mM 磷酸钠缓冲液(pH 7.5)中室温 30 分钟进行 4 μ g/ml 终浓度的羧肽酶 B(Boehringer Mannheim)处理。反应通过浸入沸水 10 分钟终止。

凝胶层析

生长到晚对数期的酵母转化体于 15000g 离心收集细胞, 4°C 下将 500 μ l 培养基直接加样至 Sephadex G-50 特级(Pharmacia)柱(1 $\times$ 100 cm)。样品用 VBA 缓冲液 (20 mM 巴比妥缓冲液, 0.11% 牛血清白蛋白和 0.6 mM 硫柳汞) 以 3.5ml/小时的流速洗脱并每 17 分钟收集级分。通过包括 125 I-白蛋白(V_0)和 22 NaCl (V_t)进行校准。 V_e 处洗脱的峰的洗脱常数 K_d 按 $K_d = (V_e - V_0)/(V_t - V_0)$ 计算。

放射免疫测定法

两种不同的抗血清用于确定加工的缩胆囊素量。Ab89009 (Paloheimo 等, 1994)对 CCK-22 的 N-端特异, Ab7270 (Hilsted 等, 1986)对甘氨酸-延伸的 CCK 特异。加工成 CCK-22 的 CCK 级分通过将 Ab89009 测量的免疫反应性除以样品用胰蛋白酶处理后用相同抗体测量的数量来计算而测量 N-端延伸的 CCK-22 总量。

酵母提取物和蛋白酶测定法

生长到对数期的酵母细胞 10 个 A_{600} 单位通过 3000g 离心 5 分钟沉淀, 用 25mlH₂O 洗一次并转入 2 ml Eppendorf 管中。加入等量酸洗过的玻璃珠 (Sigma-Aldrich) 然后加入包括多种抑制因子(150 μ M 苯丁酸抑制素, 30 μ M E-64, 10 μ M 亮抑酶肽, 1 μ M 胃酶抑制剂 A, 1 mM 苯甲基磺酰氟(PMSF)、1 mM EDTA 和 1 mM 1,10-邻二氮杂菲或每 2.5ml 0.1M NaH₂PO₄ (Boehringer Mannheim)含或不含 EDTA 的 1 片完全抑制剂的 0.1M NaH₂PO₄ (pH 4.5)200 μ l。细胞通过 3 $\times$ 20 秒涡旋破碎然后提取物通过

15000g 离心 10 分钟澄清。所有步骤在 4°C 进行。蛋白酶分析使用 20pmol 合成酰胺化 CCK-33(Peninsula Laboratorie Europe, Merseyside, England) 或 Ac-CCK-33-甘氨酸(Cambridge Research Biochemicals, Stockton, England)作为底物、20μl 酵母提取物、多种抑制因子和激活剂以总体积 30μl 进行。混合物在 30°C 孵育 1 小时, 通过加入 500 μl VBA 缓冲液终止反应然后立即浸入沸水浴 10 分钟。

使用金属蛋白酶缺陷菌株进行蛋白酶分析

分析如上所述进行, 但加入 1 mM 苯丁酸抑制素和 1 mM Mn^{2+} 以减少 N-端降解。

使用完整酵母细胞进行蛋白酶分析

沉淀对数生长细胞的 5 个 A_{600} 单位, 在 5 ml H_2O 中洗一次并在 SC 培养基(pH 6.0)中洗一次, 然后细胞重悬于 25μl SC 培养基中。蛋白酶分析通过加入 20 pmol 合成 Ac-CCK-33-甘氨酸作为底物 30°C 轻摇孵育混合物 1 小时进行。通过加入 500 μl VBA 终止反应, 通过离心去除细胞然后上清浸入沸水 10 分钟。

通过 MALDI-TOF 分析分泌的 CCK

CCK 转化的酵母细胞的 50 个 A_{600} 单位加入 25ml 新鲜培养基中, 然后孵育 3 小时。15000g 离心 10 分钟去除细胞, 500 μl 培养基通过使用 ZipTip C_{18} 柱 (Millipore)反向浓缩和除盐。肽用 10μl 50% 乙腈洗脱。纯化的多肽在使用时间滞后调焦(延时提取)的反射模式操作的基质辅助激光吸附/离子化飞行时间质谱仪(Biflex, Bruker-Franzen, Bremen, Germany)中分析。0.5μl 样品与 0.5μl 基质溶液(溶于乙腈/甲醇的 α -氰-4-羟基肉桂酸, Hewlett Packard)混合用于分析。然后将 0.5 μl 混合物加入探测器并在导入质谱仪之前允许其干燥。

统计学分析

使用非成对 students t-检验分析表达 proCCK 的野生型酵母和野生型菌株同基因突变体的 proCCK 表达的变化或成熟 CCK-22 级分的变化是否是统计学上显著的而进行统计学计算。所计算的 yapsin 突变体间 CCK-22

加工的 P-值为将 BY4705 与每一突变体进行的比较, 而括号分别代表 BY4705 *yps1* 和 BY4705 *yps1 yps3*, 以及 BY4705 *yps1 yps3* 和 BY4705 *yps1 yps2 yps3* 之间的比较。

在酿酒酵母中表达 proBNP - *yps1* 突变体的构建 *preproBNP* 的克隆

使用 Quickprep Micro mRNA 纯化试剂盒(Amersham Pharmacia Biotech)从 500 mg 人心脏活组织中分离信使 RNA。第一链 cDNA 从 2 μ g mRNA 在含有 2.5 μ l 10 $\times$ 第一链缓冲液 (Promega)、2.5 μ l 100 mM DTT、2.5 μ l 10 mM dNTP、2.5 μ l 焦磷酸钠、10 pmol Oligo(dT)₁₈、10 个单位反转录酶 AMV (Promega), 并补充 H₂O 到 25 μ l 的反应中制备。加热信使 RNA 和 Oligo(dT)₁₈ 到 70°C 5 分钟, cDNA 合成之前将其冰上冷却 5 分钟。42°C 60 分钟进行第一链 cDNA 合成。

使用 Pwo 聚合酶(Roche), 1 μ l 第一链 cDNA, 5 μ l 包括 MgCl₂ 的 10 $\times$ Pwo 缓冲液(Roche), 5 μ l 2.5 mM dNTP, 每一引物 30 pmol (BNP5'*EcoRI* 和 BNP3'*XbaI*)扩增编码 *preproBNP* 的 cDNA。编码 *preproBNP* 的 PCR 产物克隆到 pBluescript II (Stratagene)中。所有后续 PCR 反应如上所述进行。

编码 α -交配因子前原序列和 *proBNP* 的 cDNA 的融合使用突出端延伸 PCR 进行, 其中建立两个 PCR 反应。一个使用 50 ng pRS426 *preproMf α 1p-proCCK* 作为模板, MF α 15'*EcoRI* 和 MF1BNP (AS) 作为引物; 另一个使用 50 ng 克隆到 pBluescript 中的 *preproBNP* 并且引物为 MF1BNP (S)和 BNP3'*XbaI*。第三个 PCR 反应中两初始 PCR 中每一 PCR 产物约 50 ng 使用凝胶提取试剂盒(Qiagen)从琼脂糖凝胶中纯化并用作两个寡核苷酸 MF α 15'*EcoRI* 和 BNP3'*XbaI* 的模板。 *preproMf α 1p-proBNP* 编码构建体亚克隆到 pCR-Blunt II (Invitrogen)中, 并在亚克隆到 pGEM-11 *PGK1pMCSPGK1t* 的 *EcoRI* 和 *XbaI* 位点之前用载体特异的寡核苷酸测序。最后将完整基因克隆到 pRS426 中完成酵母 *proBNP* 表达构建体 pRS426 *preproMf α 1p-proBNP*。

另外, 构建两个其他构建体, 其中已去除了 proBNP 片段(1-76)。这些构建体与 preproMf α 1p-proBNP 的构建相似但仅合成 BNP-32。第一个构建体中, Kex2p 切割位点和 preproMf α 1p 的间隔肽不变(KREAEA)(图 14B), 而另一个构建体中去除间隔肽(图 14C)。野生型酵母和同基因 *CYMI* 破坏体的 BNP-32 表达分析使用 Electra-Box Diagnostica ApS 的 Shionoria-BNP 系统通过 RIA 分析。此鉴定法为 BNP-32 特异的。

proCCK, proBNP 和 Cym1p 的表达

人 proCCK 以酵母 α -交配因子前原前导序列和 proCCK 的融合蛋白形式表达(preproMf α 1p-proCCK)。融合构建体在多拷贝质粒中表达, 自磷酸甘油酸激酶启动子起始进行组成型基因转录。“ProCCK 表达”指使用 pRS426 preproMf α 1p-proCCK 的表达, 其用于除 BY4705 和 proCCK 从 pRS423 preproMf α 1p-proCCK 表达的同基因 yapsins 缺失菌株外的所有酵母菌株中。人 proBNP 也以酵母 α -交配因子前原前导序列和 proBNP (preproMf α 1p-proBNP) 的融合蛋白形式表达(图 14A)。CYMI 在自身启动子驱动下表达。质粒构建体和所用的寡核苷酸列于表 2。

BNP 放射免疫测定

98192 抗体对 proBNP N-端特异(Götze 等, 2002)。

层析

FPLC 层析在 Äkta 净化装置(Amersham Pharmacia Biotech)的 Superdex 200 柱上进行。50 mM 磷酸钠缓冲液中包括 100mM NaCl 和 6M 胍。

菌株构建

BJ2168 背景中的 $\Delta yps1::TRP1$ (LJY440)和 $\Delta cym1::LEU2 \Delta yps1::TRP1$ (LJY431) 菌株如上所述使用用于基因破坏的寡核苷酸 *YPS15'*GD400 和 *YPS13'*GD400(表 2)通过一步基因破坏技术构建。PCR 产物转化入 BJ2168 和 LJY430。通过使用 *yps15'*DC 和 *yps13'*DC(表 2)的 PCR 分析确认破坏盒的正确整合。

鉴定巴斯德毕赤酵母或甲醇毕赤酵母中编码 Cym1 直向同源的基因

从巴斯德毕赤酵母和甲醇毕赤酵母中鉴定编码酿酒酵母 Cym1p 直向同源物的未知基因使用下面提到的相同系列的简并引物以类似方式进行。巴斯德毕赤酵母和甲醇毕赤酵母在下文中以毕赤酵母形式提及。

通过比对克鲁弗酵母和粟酒裂殖酵母的直向同源 Cym1 蛋白质和酿酒酵母的 Cym1p 鉴定出许多相同的氨基酸序列。从这些序列可合成与所有三个物种 *CYM1* 的互补 DNA 链结合的简并寡聚核苷酸(表 3), 并且也和毕赤酵母的 *CYM* 基因结合。基因组序列的扩增最初通过使用高质量基因组 DNA 作为模板, *Pichia-CYM1-Ia* 和 *Pichia-CYM1-Ib* 和 *Pwo* 聚合酶(Roche)进行。扩增的预期为 525bp 大小的序列克隆到 pBlunt 或类似载体中并用载体特异引物测序。如果最初扩增未出现带, 使用第一次 PCR 反应 1 μ l 作为模板用两个嵌套引物 *Pichia-CYM1-IIa* 和 *Pichia-CYM1-IIb* 进行第二轮 PCR 循环。预期产物约为 270bp 然后克隆到 pBlunt 中并用 M13 正向和 M13 反向引物测序。从获得的序列可合成序列特异性引物、两个嵌套正义特异性引物和两个嵌套反义特异性引物。使用正义引物之一可能使用高质量基因组 DNA 作为模板获得带有 *Pichia-CYM1-IIIb* 的 PCR 产物。克隆并测序此 ~ 2100bp 产物。如果未产生 ~ 2100bp 的带, 就必需如 Clontech Marathon cDNA Amplification Kit 操作手册(BD (Becton, Dickinson and Company))中所述从毕赤酵母中分离 mRNA, 产生双链 cDNA 并在末端连上接头。使用两个序列特异性正义引物可能获得 mRNA 3'端约 2600bp 序列并且使用序列特异性反义引物扩增包括编码假定活性部位 HXXEH 基元序列的 5'端。通过从 5'和 3'非翻译区合成序列特异性寡核苷酸, 可克隆编码毕赤酵母 Cym1 直向同源物的全长 cDNA。

表 3

氨基酸序列	K Y P V R D P
寡聚物 <i>Pichia-CYM1-Ia</i>	5' AARTAYCCXGTXMGXGAYCC 3'
氨基酸序列	H P S N A K
寡聚物 <i>Pichia-CYM1-Ib</i>	3' GTRGGXWSXTTRCGXTTY 5'
氨基酸序列	D P F F K M
寡聚物 <i>Pichia-CYM1-IIa</i>	5' GAYCCXTTYTTTYAARATG 3'
氨基酸序列	G V V Y N E M
寡聚物 <i>Pichia-CYM1-IIb</i>	3' CCXCAXCAXATRTTRCTYTAC 5'
氨基酸序列	E K G G A Y G
寡聚物 <i>Pichia-CYM1-IIIb</i>	3' CTYTTYCCXCCXCGXATRCC 5'

X-次黄嘌呤核苷，简并寡核苷酸按照国际生物化学协会设计 (<http://www.chem.qmul.ac.uk/iubmb/misc/naseq.html>)。

巴斯德毕赤酵母或甲醇毕赤酵母中 *Cym1* 直向同源物的基因破坏

将编码毕赤酵母 *Cym1* 直向同源物的序列克隆到像 pBluescript-Ⅱ的载体的 *Hind*Ⅲ和 *Sac* I 中或类似的载体中，条件是其不存在于 *ORF* 中。将 *ORF* 插入 *Hind*Ⅲ和 *Sac* I 位点中去掉了载体多克隆位点中的大多数，这使得易于发现只存在于 *ORF* 中的限制性酶位点。克隆 *ORF* 中的 *KanMX* 盒，优选地这样毕赤酵母 *CYM1* 的 1000 bp 存在于 *KanMX* 盒的一侧而产生毕赤酵母 *CYM1* 破坏盒。然后此构建体使用毕赤酵母 *CYM1* 基因 5'端和 3'端特异性引物通过 PCR 扩增。将 PCR 产物转化毕赤酵母菌株然后在含 100 μg/ml 遗传霉素(G-418)的 YPD 平板上选择转化体。使用与 *KanMX* 盒 5'端和 3'端结合的毕赤酵母序列特异性 *CYM1* 引物通过菌落 PCR 确认在毕赤酵母基因组中正确的整合。从 PCR 产物的大小能分辨整合事件是否

正确。

在巴斯德毕赤酵母中表达处源蛋白质和肽

为表达肽可使用 Invitrogen 的表达构建体 pPIC α (诱导型表达)或者 pGAPZ α (组成型表达)。这两个载体都使用酿酒酵母 α -交配因子的前原序列通过分泌途径指导融合肽。 α -交配因子的前原序列 在高尔基体中去掉然后目的肽释放于培养基中。如果表达蛋白质, 当异源表达的蛋白质为胞质定位并可从完整细胞中分离时, 可使用没有 α -交配因子前原序列的两个上面所提及载体(pPICZ 和 pGAPZ)。

在甲醇毕赤酵母中表达外源蛋白质和肽

在甲醇毕赤酵母中表达蛋白质和肽以如巴斯德毕赤酵母中类似的方式进行, 其中质粒可使用用于蛋白质胞内表达和分泌到培养基的质粒。胞内表达的蛋白质可克隆到 pMET(Invitrogen)中而分泌蛋白质克隆到 pMET α (Invitrogen)中。pMET α 含有如用于在巴斯德毕赤酵母表达的酿酒酵母 α -交配因子的前原序列。此系统中通过甲醇诱导表达。

结果

生长条件对 CCK-22 加工的影响

从表达 proCCK 的 BJ2168 中测量 CCK-22 的胞内和胞外级分。无论细胞处于指数生长期还是到达稳定期, 胞内级分都保持不变(图 2)。但分泌的 CCK-22 的相对量在细胞到达稳定期时明显改变。在指数生长期 CCK-22 级分为 23%, 但在稳定期(270 分钟后)此级分增加到 37% (图 2)。所以在此所述的实验中仅使用指数生长的细胞。

赖氨酸残基在 CCK-22 释放中的意义

为评价赖氨酸残基在蛋白水解中的作用, 将用两个表达构建体 proCCK 和 proCCK (K $\rightarrow$ A)转化的 BJ2168 转化体生长到晚指数期并收集培养基。对每一菌株的培养基进行凝胶层析并用 Ab7270 测量收集的级分中 Gly-延伸 CCK 含量。野生型培养基中的 CCK 在 $K_d=0.8$ 和 1.1 的两个主峰中洗脱下来(图 3A), 这与以前建立的 CCK-22-Gly 和 CCK-8-Gly (Cantor 等, 1987; Rourke 等, 1997)的洗脱位置一致, 而 proCCK (K $\rightarrow$ A)

产生 CCK-8-Gly 和在 $K_d = 0.6$ 洗脱下来的较大形式 (图 3C)。野生型构建体中在 $K_d = 0.7$ 和 0.8 洗脱下来的两个峰 (图 3B) 分别与 C 端延伸的 CCK-22 和 CCK-22-Gly 对应 (Rourke 等, 1997), 而在这些位置处未观察到 proCCK (K→A) 构建体的 CCK-22 免疫反应性 (图 3D)。然而在 $K_d = 0.55$ 观察到免疫反应性小峰, 其可能由于 Ab 89009 与较大未加工片段的轻微交叉反应性 (Paloheimo 等, 1994)。为研究精氨酸取代赖氨酸的影响, 用 proCCK (K→R) 转化 BJ2168。在胰蛋白酶切割之前和之后分析转化体的培养基。加工成 CCK-22 的 proCCK 级分与在野生型 CCK 中观察到的类似。

通过质谱分析分泌的 CCK 肽

收集自 proCCK 转化的 BJ2168 培养基通过质谱分析 (图 12)。获得的片段与加工的 CCK-39、CCK-22 和 CCK-8 相对应。Lys⁶¹ 的 N-端鉴定出了两个肽 (Tyr⁴⁵-Val⁶⁰, 1805.0 Da 和 Tyr⁴⁵-Lys⁶¹, 1932.2 Da)。表明前者可能为后者的羧肽酶降解产物。为说明这个问题并尝试鉴定 C-端延伸的 CCKs, 本发明者产生了这样的破坏菌株, 其中编码负责加工成 CCK-8 的丝氨酸蛋白酶的 *KEX2* 和 *KEX1* 编码的羧肽酶都发生了突变。将 proCCK 转化入此 *kex2 kex1* 菌株 (LJY22) 并随后分析分泌的肽之后, 本发明者仅发现了与 Tyr⁴⁵-Lys⁶¹ 对应的峰。使用 *KEX1* 和 *KEX2* 单基因破坏表达 proCCK 过程中看到了同样的模式, 即仅有与 Tyr⁴⁵-Lys⁶¹ 对应的峰 (数据未显示)。这样 Tyr⁴⁵-Val⁶⁰ 肯定为与来源于 Lys⁶¹ 后切割的 CCK-22 一致的降解产物。其他的片段通过在与 CCK-61 (未在哺乳动物中鉴定出)、CCK-58、C-端延伸的 CCK-39 和 C-端延伸的 CCK-22 (图 12) 加工相应的 *kex2 kex1* 菌株中表达发现, 而与我们前面工作一致的是未鉴定出与 CCK-8 相应的肽, 这表明 Kex2p 参与了此加工过程 (Rourke 等, 1997)。

Kex2p 参与了 CCK-22 的生物合成

以前对 *kex2* 菌株分泌的 CCK 肽的分析和质谱获得的结果表明在 Lys⁶¹ 处切割释放 CCK-22 可在没有 Kex2p 的参与下发生。然而 *kex2* 菌株显示出 CCK-22 浓度降低。ProCCK 在液泡蛋白酶缺陷菌株和同基因 *kex2* 菌株 (BJ2168 和 LJY23) 中表达并且测量指数生长细胞中加工的胞内和胞外

CCK-22 级分。BJ2168 中约 28% 的胞内 CCK 内容物在 Lys⁶¹ 之后得到加工，而 *kex2* 菌株中仅 6% 得到加工。对分泌的 CCK 肽分析表明从表达 proCCK 的 BJ2168 中收集的培养基含有约 20% 的 CCK-22，而 *kex2* 突变体中，数量减少到 5%。这些结果表明 Kex2p 参与了生成 CCK-22 的加工过程。然而还有其他蛋白酶能进行 Lys⁶¹ 切割。

Lys⁶¹ 切割的体外鉴定法

为研究除 Kex2p 外能在 proCCK 单一 Lys⁶¹ 之后进行内肽酶切割的其它蛋白酶本质，使用酿酒酵母粗制品且以合成 CCK-33 作为底物建立了体外鉴定法。

使用液泡蛋白酶缺陷菌株 BJ2168 的提取物，存在普遍的 N-端降解并且可测量的 CCK 回收少于无酵母提取物对照的 10%。因为鉴定法依赖于抗体结合的 CCK-22 完整 N 端，本发明者建立了其中一些已知酿酒酵母氨肽酶缺失的菌株。Y14953 菌株 (*ape1*) 用作亲代菌株，其也缺失了 *APE2* 和 *APE3* 基因。使用此 LJY123 菌株制备细胞提取物，与用 BJ2168 观察到的回收相比有 2-3 倍增加的免疫反应性。

加工成 CCK-22 依赖于金属离子

蛋白酶切割合成的人 CCK-33 产生 CCK-22 的本质通过包括多种不同抑制因子用 LJY123 提取物来分析。结果表明仅金属螯合剂的加入抑制 CCK-33 蛋白水解成 CCK-22 (图 4A)。

活性最初被 1mM EDTA 的加入抑制之后，体外检测蛋白酶的金属依赖性。通过加入 0.2mM 过量的不同二价阳离子检测活性重建导致的 CCK-22 成熟。与已知金属蛋白酶只能由 Zn²⁺、Co²⁺ 和 Mn²⁺ 活化的性质一致，Zn²⁺、Co²⁺ 和 Mn²⁺ 的加入能重建蛋白酶活性而 Ca²⁺、Cu²⁺ 或 Mg²⁺ 的加入没有影响 (图 4B)。使用渐增的 Zn²⁺ 浓度再活化表明为一个双相模式，其中 Zn²⁺ 在 5mM 浓度以上执行抑制作用 (数据未显示)。

1 mM EDTA 最初抑制之后使用 LJY123 细胞提取物分析 Zn²⁺ 和 Mn²⁺ 再活化的金属蛋白酶的 CCK-切割时间进程。通过加入 1.2 mM Zn²⁺ 或 Mn²⁺ 然后孵育 30、60 和 120 分钟进行再活化。在此鉴定法和下面体外蛋

白酶鉴定法中发明者使用 N-端乙酰化的 CCK-33-Gly(Ac-CCK-33-Gly)作为底物,其引起非常缓慢的非特异性降解。使用 Ab89009 测量胰蛋白酶切割前后的 CCK-22 免疫反应性表明 30 和 60 分钟时 Zn^{2+} 和 Mn^{2+} 没有活化能力的区别,然而 120 分钟后测量到使用 Mn^{2+} 作为活化剂与 Zn^{2+} 相比 CCK-22 免疫反应性增加了 10% 以上(图 5)。免疫反应性的增加可能由于在此加入 Mn^{2+} 之后抑制了降解,也可能由于表 4 中使用的酵母细胞提取物。

表 4. 酿酒酵母金属蛋白酶。搜索在 Swiss-Prot 序列检索系统(SRS) <http://www.expasy.ch/> 中进行。蛋白酶分析使用金属蛋白酶缺陷菌株提取物的两个独立分析(A 和 B)中进行。用 Ab89009 测量 CCK-22 量并在胰蛋白酶切割之后用 Ab89009 测量 CCK 总量。假定的金属蛋白酶用*标记。

名称	Swiss-Prot acc #	ORF	CCK-22 [nM]		总 CCK [nM]		级分 CCK-22	
			A ₁	B ₁	A ₂	B ₂	A ₁ /A ₂	B ₁ /B ₂
AAP1	P37898	YHR047	3.2	3.2	36	32	0.09	0.10
AFG3	P39925	YER017	2.8	2.4	23	24	0.12	0.10
APE1	P14904	YKL103	4.4	3.7	34	28	0.13	0.13
APE2	P32454	YKL157						
APE3	P37302	YBR286						
DPP3	Q08225	YOL057	4.3	4.6	31	35	0.14	0.13
LTA4	Q10740	YNL045	3.9	4.0	32	29	0.12	0.13
MIP1	P35999	YKL134	3.4	3.3	34	33	0.10	0.10
PRD1	P25375	YCL057	2.4	2.8	28	26	0.09	0.11
QRI7	P43122	YDL104	2.8	2.9	23	24	0.12	0.12
RCA1	P40341	YMR089	4.2	3.5	31	27	0.14	0.13
STE2	Q06010	YLR389	2.6	2.3	25	20	0.10	0.12
STE2	P47154	YJR117	3.4	2.8	19	17	0.18	0.16
YBS4	P38244	YBR074	2.6	2.7	27	29	0.10	0.09
YHR3	P38821	YHR113	2.5	2.4	26	26	0.10	0.09
YHT2	P38836	YHR132	3.8	3.3	34	32	0.11	0.10
YIK8	P40483	YIL108	5.7	5.2	43	39	0.13	0.13
YIN7	P40462	YIL137c	3.5	2.8	23	20	0.15	0.14
YK18	P36132	YKR038	2.9	2.1	23	20	0.13	0.11
YME	P32795	YPR024	2.3	2.7	25	25	0.09	0.11

MAS2	P11914	YHR024	ND, 致死基因					
MAS	P10507	YLR163c						
AXL1	P40851	YPR122	3.5	3.4	31	30	0.11	0.11
CYM	P32898	YDR430	0.6	0.4	42	39	0.01	0.01
YOJ8	Q12496	YOL098	3.6	3.8	35	34	0.10	0.12

胞外 yapsin 活性

为研究蛋白酶活性为分泌的还是与质膜胞外联系，在培养基中用完整酵母细胞分析了蛋白酶活性。使用指数生长的 LJY123 细胞 30°C 孵育 1 小时后未发生 CCK-33 的降解，这与以前的观察一致 (Rourke 等, 1997)。在与完整酵母细胞孵育过程中，可测量暴露 CCK-22 N-端的切割 (图 6)，然而这种蛋白酶活性不能被所研究的抑制剂消除 (数据未显示)。通过使用包含 *YPS1*、*YPS2* 和 *YPS3* 基因破坏 (表 I) 的完整细胞，加工的 CCK-22 级分通过缺失三个天冬氨酰蛋白酶的任一个从而与野生型细胞相比减少 (图 6)。这些数据表明三个蛋白酶都具有胞外蛋白酶活性，其能在 proCCK 中 Lys⁶¹ 处切割。初步结果表明 *YPS7* 基因破坏与 *YPS1* 缺失相比，前者降低胞外 Lys⁶¹ 加工的量 (未发表结果)。

在金属蛋白酶缺陷菌株中表达 proCCK

基于前面所述酿酒酵母中金属蛋白酶的内肽酶活性 (Adames 等, 1995; Schmidt 等, 2000), *AXL1* (LJY201) 和 *STE24* (LJY202) 基因缺失菌株最初在 BJ2168 中制备。ProCCK 在这些菌株中的表达表明与野生型相比 Lys⁶¹ 之后的蛋白水解未改变并决定检测余下的金属蛋白酶缺陷菌株 LJY123、LJY203、LJY204 和通过 Euroscarf 获得的金属蛋白酶缺陷菌株 (表 I) 分泌 CCK-22 的能力 (不包括线粒体肽酶)。产生 CCK 的金属蛋白酶缺陷菌株中 CCK-22 免疫反应性未显著改变 (数据未显示)，并且通过此方法未鉴定出将异源表达的 proCCK 加工成 CCK-22 的蛋白酶。

CYM1 编码能释放 CCK-22 游离 N-端的蛋白酶

制备每一可育金属蛋白酶缺陷菌株的细胞提取物并在体外蛋白酶鉴定法中检测以研究是否有可测量的蛋白水解的减少。在此鉴定法中，加入

Ac-CCK-33-Gly 之前使用 1 mM Mn^{2+} 和 1 mM 苯丁抑制素, 因为发现加入这些氨肽酶抑制剂时回收为 80-90 % 而不加入这些氨肽酶抑制剂时回收为 30 % (数据未显示)。CYM1 的缺失几乎消除了蛋白酶活性, 而其他金属蛋白酶缺陷菌株均未表明 CCK-22 生物合成的显著变化 (表 4)。

多拷贝质粒上表达 CYM1 在体外增加成熟 CCK-22 级分

为确定在体外合成的 CCK-22 的数量是否与 Cym1p 的数量相关, 将 Cym1p 在多拷贝质粒上表达并随着时间分析合成的 CCK-22 级分。pRS425 CYM1 转化的 BJ2168 细胞提取物和 pRS425 空载体转化的对照用于含 1 mM Mn^{2+} 的体外蛋白酶鉴定法中, 其中反应在 15、30、45 和 60 分钟后终止。CCK-22 免疫反应性用 Ab89009 测量, 余下的 CCK-33 在胰蛋白酶切割之后用相同的抗体测量 (图 7)。多拷贝质粒上 CYM1 的表达提高 CCK-22 产生速度数倍。然而, 发明者还观察到 CCK-33 和 CCK-22 降解的增加 (图 7B)。当在 pH 6.0 和 pH 7.5 下进行相同实验时, 降解有非常明显的增加并且 30 分钟孵育之后在 pH 6.0 中检测不到 CCK 免疫反应性 (数据未显示)。这些结果表明 CCK-22 体外成熟过程中的赖氨酸特异性切割依赖于 Cym1p 的量。

cym1 突变体菌株中 proCCK 的表达提高 CCK 的分泌

为说明 CYM1 在 CCK-22 体内合成中的作用, 在液泡蛋白酶缺陷菌株 BJ2168 和同基因 *kex2* 菌株中制备 CYM1 基因缺失。CYM1 的缺失造成细胞中 proCCK 总量增加约 40 % (图 8A) 同时 CCK-22 不依赖于 KEX2 破坏有类似的减少 (图 8C)。同样 *cym1* 菌株中总 CCK 分泌量增加 60 % 以上 (图 8B), 但与胞内 CCK-22 级分减少不同, CCK-22 胞外级分与液泡蛋白酶缺陷菌株和同基因 *kex2* 菌株相比有所增加 (图 8D)。

与 *cym1* 菌株中野生型 CCK 累积相比, CCK K→A 突变体的表达引起胞内 CCK 的累积

CYM1 的基因破坏造成 CCK 胞内浓度增加 (图 8A) 的观察提出这样一个问题, Cym1p 的蛋白水解活性是否在移位到内质网之前引起 CCK-22 的降解。因此本发明者在表达 CCK 的菌株中检测了胞内 CCK 含量, 其中

通过使用 $\text{Lys}^{61} \rightarrow \text{Ala}^{61}$ 突变体消除了 CCK-22 的成熟。胰蛋白酶和羧肽酶 B 处理之后使用 Ab7270 分析此 CCK 突变体在液泡蛋白酶缺陷菌株 BJ2168 和同基因 *cym1* 菌株中的转化体并且与野生型 CCK 表达相比此构建体的胞内 CCK 免疫反应性有所增加（图 9）。分别在 BJ2168 和 *CYM1* 破坏菌株中表达时，突变体 CCK(K $\rightarrow$ A) 和野生型 CCK 转化体造成胞内 proCCK 浓度的增加。胞内 proCCK 的增加并非附加表明 Cym1p 的蛋白水解活性在移位到内质网之前引起 CCK-22 的降解。

在天冬氨酰蛋白酶缺陷菌株中表达 proCCK

在无效突变体 *YPS1*、-2 和 -3 中分析 proCCK 的 Lys^{61} -特异性切割，其中测量指数生长的野生型酵母细胞 BY4705 和用 proCCK 转化 LJY13、-14 和 -15 的同基因天冬氨酰蛋白酶缺陷菌株中 CCK-22 的胞内和胞外数量。BY4705 的胞内和胞外 CCK 免疫反应性比液泡蛋白酶缺陷菌株 BJ2168 相比低 10 倍多（数据未显示）。CCK-22 胞内级分从野生型细胞中约 28% 显著减少到 *yps1* 菌株中的 17%，而通过 *YPS2* 和 *YPS3* 基因破坏未测量到额外的减少（图 10A）。然而 CCK-22 胞外级分确实表明 Yps1p、Yps2p 和 Yps3p 都参与了 CCK-22 的生物合成并且三联突变体降低 CCK-22 到野生型酵母的 2/3（图 10B）。

CYM1 破坏引起分泌的野生型 CCK 和 CCK K $\rightarrow$ R 突变体总量增加两倍

为说明 Cym1p 是否切割单一精氨酸残基的 C-端，在液泡蛋白酶缺陷菌株 BJ2168 和同基因 *cym1* 菌株中表达 CCK(K $\rightarrow$ R) 突变体。比较胞内和胞外 CCK 浓度与这些菌株中表达的野生型 CCK 浓度。突变 CCK(K $\rightarrow$ R) 总量与野生型 CCK 相比在胞内和胞外都有增加（图 11）。在 *cym1* 菌株中表达时，野生型 CCK 和 $\text{Lys}^{61} \rightarrow \text{Arg}^{61}$ 突变体都表明胞外 CCK 可测量数量增加两倍（图 11B）。

在肽合成中利用 Cym1p 活性

本发明的另一方面为使用 Cym1p 活性产生数量增加的胞质 Cym1p 活性，其中 Cym1p 从其自身启动子表达或从强组成型启动子例如 *PGK1*、

ADH1 或 *TPI1* 或诱导型 *GAL1* 启动子表达。如前面提到的, 当在 2 μ 质粒上 *CYM1* 从其自身启动子转录时 CCK-22 的合成显著增加 (图 7)。转录可从含有启动子、*CYM1* 和终止子的质粒进行或者通过异源重组取代 *CYM1* 的内源启动子在基因组中引入目的启动子进行。

活性可用于细胞内以产生不需经分泌途径的翻译后修饰的肽, 例如二硫键形成、N-和 O-糖基化或外切蛋白酶活性。

野生型细胞中 CCK-22 的生物合成与 *CYM1* 破坏同基因菌株的比较表明了 Cym1p 胞质活性在胞内肽合成中的作用, 其表明 CCK-22 数量的显著增加 (图 8C)。目的肽的合成以避免移位到内质网 (ER) 的方式进行。这可通过从蛋白质中去掉导致翻译后进入内质网的疏水氨基端信号序列进行或在温度敏感分泌突变体如 *sec61* 中的表达来进行, 当温度提高到 37°C 时其能消除分泌肽移位到内质网中。

然后目的前肽或原前肽 (prepropeptide) 可为胞质定位并为 Cym1p 的潜在底物。肽从其前体的释放通过引入 proCCK 中观察到的切割位点通过 Cym1p 活性进行, 其造成内肽酶解加工 C-端 Lys⁶¹ (Ser-Ile-Val-Lys⁶¹ ↓) 后 Gly-延伸 CCK-22 的释放 (图 13A)。如果目的肽为 GLP1, 合成可以与 Cym1p 切割位点融合的形式进行, 其为部分 proCCK (图 13B)。然后目的肽将在胞质中累积并可从裂解之后的沉淀细胞中纯化。

ProBNP 在 *cym1* 突变体菌株中的表达提高 proBNP 的分泌

为说明 *CYM1* 在 proBNP 体外生物合成中的作用, 在液泡缺陷菌株 BJ2168 和三个蛋白酶缺陷同基因菌株 $\Delta cym1::LEU2$ (LJY430)、 $\Delta yps1::TRP1$ (LJY440) 和 $\Delta cym1::LEU2 \Delta yps1::TRP1$ (LJY431) 中表达 proBNP。CYM1 的缺失造成分泌的 proBNP 总量增加约 100%, 而 proBNP 的分泌不依赖于编码天冬氨酰蛋白酶 Yps1p 的基因破坏并因此与野生型菌株相似 (图 15A)。预期 Cym1 和 Yps1 的破坏都与 *cym1* 突变体中的分泌量相似 (图 15A)。

proBNP 从 *cym1* 突变体中分泌的分析

为分析酿酒酵母中表达的 proBNP, *cym1* 突变体的培养基用于 FPLC

层析并使用抗体 98192 通过 RIA 分析。从 34-39 级分洗脱的峰与完整 proBNP 对应, 而 53-62 级分中洗脱的峰很可能是 proBNP 1-76 片段的 proBNP 加工形式(图 15B)。1-76 片段和 BNP-32 从 proBNP 的释放是由于单一精氨酸残基之后的切割并可能由于 Kex2 或 Yps1 的活性。

讨论

分泌的多肽随生长条件而改变, 当培养到达稳定期时 CCK-22 级分增加, 而胞内加工级分在胁迫条件下保持不变。胞外切割成 CCK-22 随着细胞进入稳定期而增加表明能将 proCCK 加工成 CCK-22 的胞外内切蛋白酶为分泌的或在细胞膜上表达。已知天冬氨酰蛋白酶 Yps1p 和 Yps2p 表现细胞表面活性(Komano 等, 1998)。此外, 以前表明异源肽在 *yps1* 菌株中的表达通过抑制单碱性残基 C-端的蛋白水解提高对像白蛋白、胰高血糖素、GLP1、GLP2 和 CART 的蛋白质和肽的回收(Egel-Mitani 等, 2000; Kerry-Williams 等, 1998)。这样最近的研究((Egel-Mitani 等, 2000; Kerry-Williams 等, 1998)和本发明者的研究)表明了指数生长期收集分泌肽以避免额外胞外加工的重要性。

液泡蛋白酶缺陷菌株中表达的 ProCCK 在 proCCK 的 Lys⁶¹ 处表现 30 % 的胞内加工。而胞外 Lys⁶¹-加工级分减少到完整酵母细胞中观察到级分的 2/3, 其显示了分泌前 CCK-22 的胞内降解。有限营养来源下胞外蛋白水解的增加可能由于如观察到的稳定期 *YPS1* 转录上调的有限营养来源下胞外蛋白酶转录的激活或上调(Gasch 等, 2000)。Yapsins 中 Yps1p、Yps2p 和 Yps3p 具有部分细胞表面活性, 但是甚至在三联突变体中仍保留了一些胞外活性仍。

本研究中发明者显示了 *KEX2* 的缺失引起胞内和胞外 Lys⁶¹ 切割都降低 5 倍。表达 proCCK 的 *kex2* 菌株不仅改变了 proCCK 中的 Lys⁶¹ 切割还改变了 CCK 的胞内存留时间, 因为与野生型酵母相比 CCK 肽的胞内浓度降低了 60 % 以上而胞外 CCK 浓度增加了几乎 60 %。此外 *kex1 kex2* 双突变体和 *kex2* 突变体中分泌的 CCK 肽分析表明 Tyr⁴⁵-Val⁶⁰ 降解产物消失。这样表明通过高尔基体外侧网络提高的分泌速度在 *kex2* 菌株中消除了通

过 Kex1p 去掉 Lys⁶¹。这些结果和 proCCK 快速分泌的观察暗示了 Kex2p 造成的胞内存留引起了野生型酵母中通过 yas1p 和可能某种程度上通过 Kex2p 的 CCK-22 合成的增加。

在体外蛋白酶鉴定法中使用酿酒酵母粗提取物研究了引起 CCK-22 胞内成熟的蛋白酶类型,以分析合成的人 CCK-33 在不同抑制剂存在条件下加工成 CCK-22 的过程。通过在蛋白酶活性的提取过程中不包括去垢剂避免了 Kex2p 和 GPI-锚定的 yapsins 活性(Azaryan 等, 1993; Fuller 等, 1989; Komano 等, 1999)。在受检测的抑制剂中,只有 EDTA 和 1,10 邻二氮杂菲抑制蛋白水解并且活性可通过加入二价阳离子 Zn^{2+} 、 Co^{2+} 和 Mn^{2+} 恢复。这表明金属蛋白酶参与了 CCK-22 的成熟过程。

候选金属蛋白酶都不含有指导蛋白质到内质网的明显信号肽。因此本发明者研究了线粒体蛋白酶之外的每一种金属蛋白酶缺陷菌株。proCCK 在每一菌株中的表达造成与野生型酵母中看到的相似不变的 CCK-22 成熟过程。然而,通过使用体外蛋白酶鉴定法,本发明者鉴定出 Cym1p 为进行 CCK-33 赖氨酸后切割的内肽酶。Cym1p 可切割 proCCK 中的 Lys⁶¹ 通过过量表达研究验证,其表明酶活性增加数倍。

cym1 菌株中 CCK-22 胞内合成减少同时总 proCCK 浓度增加。相反地,胞外 CCK-22 级分与总的 CCK 平行增加的野生型酵母相比增加。这些发现与像大多数胰岛素降解酶一样的 Cym1p 活性的胞质定位一致(Bai 等, 1996)并表明其在移位到内质网之前作用于 preproMf α 1p-proCCK 构建体。这样 proCCK 的移位前降解通过 *CYM1* 破坏减少并且总产量增加。

在 *cym1* Δ 突变体中 proBNP 以与 preproMf α 1p 序列的融合肽形式表达显示出胞外 proBNP 含量与野生型菌株相比增加了两倍。层析分析分泌的 proBNP 揭示了存在两种主要形式。一种为完整 proBNP,而另一种为 proBNP 1-76 片段,这样也合成了生物活性的 BNP-32,但是在本鉴定法中未检测到。proBNP 1-76 片段的释放很可能依赖于 Kex2p 活性,然而这在本鉴定法中未检测,因为 proBNP 的释放依赖于 Kex2p 和 Kex1p 两者。

所有以上讨论的出版物在此全部引用。

本领域技术人员理解如详细实施方案中显示的, 对本发明所做的大量改变和/或变化并不偏离广泛描述的本发明的核心或范围。因此本实施方案在所有方面应理解为的说明性的而并非限制性的。

参考文献

- Adames, N., Blundell, K., Ashby, M.N.和 Boone, C. (1995). *Science* 270:464-67.
- Azaryan, A.V., Wong, M., Friedman, T.C., Cawley, N.X., stivariz, F.E., Chen, H.C.和 Loh, Y.P. (1993). *J. Biol. Chem.* 68:11968-75.
- Bai, J.P., Hong, H.J., Rothenberger, D.A., Wong, W.D.和 Buls, J.G. (1996). *J. Pharm. Pharmacol.* 48:1180-84.
- Becker, A.B 和 Roth, R.A. (1992) *PNAS* 89: 3835-3839.
- Brachmann, C.B., Davies, A., Cost, G.J., Caputo, E., Li, J., Hieter, P.和 Boeke, J.D. (1998). *Yeast* 14:115-132.
- Cantor, P. 和 Rehfeld, J.F. (1987). *Clin. Chim. Acta* 168:153-58.
- Egel-Mitani, M., Andersen, A.S., Diers, I., Hach, M., Thim, L., Hastrup, S.和 Vad, K. (2000). *Enzyme Microb. Technol.* 26:671-77.
- Egel-Mitani, M., Flygenring, H.P.和 Hansen, M.T. (1990). *Yeast* 6:127-137.
- Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K.和 Tuschl, T. (2001). *Nature* 411, 494-98.
- Fuller, R.S., Brake, A.和 Thorner, J. (1989). *Proc. Natl. Acad. Sci. USA.* 86:1434-38.
- Gasch, A.P., Spellman, P.T., Kao, C.M., Carmel-Harel, O., Eisen, M.B., Storz, G., Botstein, D.和 Brown, P.O. (2000). *Mol. Biol. Cell* 11:4241-57.
- Gietz, R.D., Schiestl, R.H., Willems, A.R.和 Woods, R.A. (1995). *Yeast* 11:355-60.

Götze JP, Kastrup J, Pedersen F 和 Rehfeld JF. (2002) 使用不依赖于前体加工的分析法定量 pro-B-型尿钠肽及其在人血浆中的产物, Clin Chem. Jul;48(7):1035-42).

Haseloff, J. 和 Gerlach, W.L. (1988). Nature, 334: 585-91.

Hilsted, L.和 Rehfeld, J.F. (1986). Anal. Biochem. 152:119-26.

Hooper N.M. (1994) FEBS Letters 354: 1-6.

Horton, R.M., Ho, S.N., Pullen, J.K., Hunt, H.D., Cai, Z.和 Pease, L.R. (1993). Methods Enzymol. 217:270-79.

Jones, E.W. (1991). Methods Enzymol. 194:428-53.

Julius, D., Schekman, R.和 Thorner, J. (1984). Cell 36:309-18.

Kerry-Williams, S.M., Gilbert, S.C., Evans, L.R.和 Balance, D.J. (1998). Yeast 14:161-69.

Komano, H., Rockwell, N., Wang, G.T., Krafft, G.A.和 Fuller, R.S. (1999). J. Biol. Chem. 274:24431-37.

Komano, H., Seeger, M., Gandy, S., Wang, G.T., Krafft, G.A. 和 Fuller, R.S. (1998). J. Biol. Chem. 273:31648-51.

Needleman, S.B.和 Wunsch, C.D. (1970). J. Mol. Biol. 48: 443-53.

Paloheimo, L.I.和 Rehfeld, J.F. (1994). Clin. Chim. Acta 229:49-65.

Philippsen, P., Stotz, A.和 Scherf, C. (1991). Methods Enzymol. 194:169-82.

Rawlings, N.D.和 Barrett, A.J. (1995). Meth. Enzymol. 248:183-228.

Rothstein, R. (1991). Methods Enzymol. 194:281-301.

Rourke, I.J., Johnsen, A.H., Din, N., Petersen, J.G.L.和 Rehfeld, J.F. (1997). J. Biol. Chem. 272:9720-27.

Sambrook, J., Fritsch, E.F.和 Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 第二版, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Schmidt, W.K., Tam, A.和 Michaelis, S. (2000). J. Biol. Chem.

275:6227-33.

Shippy, R., Lockner, R., Farnsworth, M.和 Hampel, A. (1999). Mol. Biotech. 12: 117-29.

Waterhouse, P.M., Graham, M.W.和 Wang, M.-B. (1998). Proc Natl Acad Sci U S A, 95: 13959-64.

- <110> CYM1P 有限公司
- 5 <120> 增加宿主细胞重组多肽产量的方法
- <130> P32077PC01
- 10 <150> DK/PA 2002 01391
<151> 2002-09-20
- <160> 68
- 15 <170> 用于 Windows 版本 4.0 的 FastSEQ
- <210> 1
<211> 5
- 20 <212> PRT
<213> 人工序列
- <220>
<223> Pitrilysin 共有序列
- 25 <220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> X = 任一氨基酸
- 30 <220>
<221> MISC_FEATURE
<222> (3)..(3)
<223> X = 任一氨基酸
- 35 <400> 1
- His Xaa Xaa Glu His
40 1 5
- <210> 2
<211> 46
- 45 <212> PRT
<213> 人工序列
- <220>
<223> Pitrilysin 共有序列
- 50 <220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> X = 任一氨基酸

<220>
<221> MISC_FEATURE
5 <222> (3)..(3)
<223> X = 任一氨基酸
<220>
<221> MISC_FEATURE
<222> (5)..(5)
10 <223> X = 任一氨基酸

<220>
<221> MISC_FEATURE
15 <222> (6)..(6)
<223> X = 任一氨基酸

<220>
20 <221> MISC_FEATURE
<222> (9)..(9)
<223> X = 任一氨基酸

25 <220>
<221> MISC_FEATURE
<222> (10)..(10)
<223> X = 任一氨基酸

30
<220>
<221> MISC_FEATURE
<222> (11)..(11)
<223> X = 任一氨基酸
35

<220>
<221> MISC_FEATURE
<222> (12)..(12)
40 <223> X = 任一氨基酸

<220>
<221> MISC_FEATURE
45 <222> (14)..(14)
<223> X = 任一氨基酸

<220>
50 <221> MISC_FEATURE
<222> (15)..(15)
<223> X = 任一氨基酸

55 <220>
<221> MISC_FEATURE
<222> (16)..(16)
<223> X = 任一氨基酸

```
<220>
<221> MISC_FEATURE
5 <222> (17)..(17)
<223> X = 任一氨基酸
<220>
<221> MISC_FEATURE
<222> (18)..(18)
10 <223> X = 任一氨基酸

<220>
<221> MISC_FEATURE
15 <222> (19)..(19)
<223> X = 任一氨基酸

<220>
20 <221> MISC_FEATURE
<222> (20)..(20)
<223> X = 任一氨基酸

<220>
25 <221> MISC_FEATURE
<222> (21)..(21)
<223> X = 任一氨基酸

30
<220>
<221> MISC_FEATURE
<222> (22)..(22)
<223> X = 任一氨基酸
35

<220>
<221> MISC_FEATURE
<222> (23)..(23)
40 <223> X = 任一氨基酸

<220>
<221> MISC_FEATURE
45 <222> (24)..(24)
<223> X = 任一氨基酸

<220>
50 <221> MISC_FEATURE
<222> (25)..(25)
<223> X = 任一氨基酸

55 <220>
<221> MISC_FEATURE
<222> (26)..(26)
<223> X = 任一氨基酸
```

```
<220>
<221> MISC_FEATURE
5 <222> (27)..(27)
<223> X = 任一氨基酸
<220>
<221> MISC_FEATURE
<222> (28)..(28)
10 <223> X = 任一氨基酸

<220>
<221> MISC_FEATURE
15 <222> (29)..(29)
<223> X = 任一氨基酸

<220>
20 <221> MISC_FEATURE
<222> (30)..(30)
<223> X = 任一氨基酸

25 <220>
<221> MISC_FEATURE
<222> (31)..(31)
<223> X = 任一氨基酸

30
<220>
<221> MISC_FEATURE
<222> (32)..(32)
<223> X = 任一氨基酸
35

<220>
<221> MISC_FEATURE
<222> (33)..(33)
40 <223> X = 任一氨基酸

<220>
<221> MISC_FEATURE
45 <222> (34)..(34)
<223> X = 任一氨基酸

<220>
50 <221> MISC_FEATURE
<222> (35)..(35)
<223> X = 任一氨基酸

55 <220>
<221> MISC_FEATURE
<222> (36)..(36)
<223> X = 任一氨基酸或不存在
```

```

    <220>
    <221> MISC_FEATURE
5  <222> (37)..(37)
    <223> X = 任一氨基酸或不存在
    <220>
    <221> MISC_FEATURE
    <222> (40)..(40)
10 <223> X = 任一氨基酸

    <220>
    <221> MISC_FEATURE
15 <222> (42)..(42)
    <223> X = 任一氨基酸

    <220>
    <221> MISC_FEATURE
20 <222> (43)..(43)
    <223> X = 任一氨基酸

    <220>
    <221> MISC_FEATURE
25 <222> (44)..(44)
    <223> X = 任一氨基酸

    <220>
    <221> MISC_FEATURE
30 <222> (45)..(45)
    <223> X = 任一氨基酸

35
    <400> 2

Gly Xaa Xaa His Xaa Xaa Glu His Xaa Xaa Xaa Xaa Gly Xaa Xaa Xaa
40 1          5          10          15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
    20          25          30
45

Xaa Xaa Xaa Xaa Xaa Asn Ala Xaa Thr Xaa Xaa Xaa Xaa Thr
    35          40          45
50
    <210> 3
    <211> 46
    <212> PRT
    <213> 人工序列
55
    <220>
    <223> Pitrilysin 共有序列

```

<220>
<221> MISC_FEATURE
<222> (2)..(2)
5 <223> X = 任一氨基酸

<220>
<221> MISC_FEATURE
<222> (3)..(3)
10 <223> X = 任一氨基酸

<220>
<221> MISC_FEATURE
15 <222> (5)..(5)
<223> X = 任一氨基酸

<220>
20 <221> MISC_FEATURE
<222> (6)..(6)
<223> X = 任一氨基酸

25 <220>
<221> MISC_FEATURE
<222> (9)..(9)
<223> X = 任一氨基酸

30
<220>
<221> MISC_FEATURE
<222> (10)..(10)
<223> X = 任一氨基酸

35
<220>
<221> MISC_FEATURE
<222> (11)..(11)
40 <223> X = 任一氨基酸

<220>
<221> MISC_FEATURE
45 <222> (12)..(12)
<223> X = 任一氨基酸

<220>
50 <221> MISC_FEATURE
<222> (14)..(14)
<223> X = S or T

55 <220>
<221> MISC_FEATURE
<222> (15)..(15)
<223> X = 任一氨基酸

<220>
<221> MISC_FEATURE
5 <222> (19)..(19)
<223> X = 任一氨基酸
<220>
<221> MISC_FEATURE
<222> (20)..(20)
10 <223> X = 任一氨基酸

<220>
<221> MISC_FEATURE
15 <222> (21)..(21)
<223> X = 任一氨基酸

<220>
20 <221> MISC_FEATURE
<222> (22)..(22)
<223> X = 任一氨基酸

25 <220>
<221> MISC_FEATURE
<222> (23)..(23)
<223> X = 任一氨基酸

30
<220>
<221> MISC_FEATURE
<222> (24)..(24)
<223> X = 任一氨基酸

35
<220>
<221> MISC_FEATURE
<222> (25)..(25)
40 <223> X = 任一氨基酸

<220>
45 <221> MISC_FEATURE
<222> (26)..(26)
<223> X = 任一氨基酸

<220>
50 <221> MISC_FEATURE
<222> (27)..(27)
<223> X = 任一氨基酸或不存在

55 <220>
<221> MISC_FEATURE
<222> (28)..(28)
<223> X = 任一氨基酸或不存在

<220>
<221> MISC_FEATURE
5 <222> (30)..(30)
<223> X = 任一氨基酸
<220>
<221> MISC_FEATURE
<222> (31)..(31)
10 <223> X = 任一氨基酸

<220>
<221> MISC_FEATURE
15 <222> (32)..(32)
<223> X = 任一氨基酸

<220>
<221> MISC_FEATURE
<222> (33)..(33)
<223> X = 任一氨基酸

25 <220>
<221> MISC_FEATURE
<222> (34)..(34)
<223> X = 任一氨基酸

30
<220>
<221> MISC_FEATURE
<222> (35)..(35)
<223> X = 任一氨基酸

35
<220>
<221> MISC_FEATURE
<222> (36)..(36)
40 <223> X = 任一氨基酸

<220>
<221> MISC_FEATURE
45 <222> (37)..(37)
<223> X = 任一氨基酸或不存在

<220>
50 <221> MISC_FEATURE
<222> (40)..(40)
<223> X = 任一氨基酸

55 <220>
<221> MISC_FEATURE
<222> (42)..(42)
<223> X = 任一氨基酸

```

    <220>
    <221> MISC_FEATURE
5  <222> (43)..(43)
    <223> X = 任一氨基酸
    <220>
    <221> MISC_FEATURE
    <222> (44)..(44)
10 <223> X = D or E

    <220>
    <221> MISC_FEATURE
15 <222> (45)..(45)
    <223> X = 任一氨基酸

    <400> 3
20
Gly Xaa Xaa His Xaa Xaa Glu His Xaa Xaa Xaa Xaa Gly Xaa Xaa Lys
1      5      10      15

25 Tyr Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa
      20      25      30

    Xaa Xaa Xaa Xaa Xaa Asn Ala Xaa Thr Xaa Xaa Xaa Xaa Thr
30      35      40      45

    <210> 4
    <211> 989
35 <212> PRT
    <213> 酿酒酵母 (Saccharomyces cerevisiae)

    <400> 4

40 Met Leu Arg Phe Gln Arg Phe Ala Ser Ser Tyr Ala Gln Ala Gln Ala
1      5      10      15

    Val Arg Lys Tyr Pro Val Gly Gly Ile Phe His Gly Tyr Glu Val Arg
45      20      25      30

    Arg Ile Leu Pro Val Pro Glu Leu Arg Leu Thr Ala Val Asp Leu Val
      35      40      45
50

    His Ser Gln Thr Gly Ala Glu His Leu His Ile Asp Arg Asp Asp Lys
      50      55      60

55
Asn Asn Val Phe Ser Ile Ala Phe Lys Thr Asn Pro Pro Asp Ser Thr
65      70      75      80

```

Gly Val Pro His Ile Leu Glu His Thr Thr Leu Cys Gly Ser Val Lys
 85 90 95

5 Tyr Pro Val Arg Asp Pro Phe Phe Lys Met Leu Asn Lys Ser Leu Ala
 100 105 110

10 Asn Phe Met Asn Ala Met Thr Gly Pro Asp Tyr Thr Phe Phe Pro Phe
 115 120 125

15 Ser Thr Thr Asn Pro Gln Asp Phe Ala Asn Leu Arg Gly Val Tyr Leu
 130 135 140

Asp Ser Thr Leu Asn Pro Leu Leu Lys Gln Glu Asp Phe Asp Gln Glu
 145 150 155 160

20 Gly Trp Arg Leu Glu His Lys Asn Ile Thr Asp Pro Glu Ser Asn Ile
 165 170 175

25 Val Phe Lys Gly Val Val Tyr Asn Glu Met Lys Gly Gln Ile Ser Asn
 180 185 190

30 Ala Asn Tyr Tyr Phe Trp Ser Lys Phe Gln Gln Ser Ile Tyr Pro Ser
 195 200 205

35 Leu Asn Asn Ser Gly Gly Asp Pro Met Lys Ile Thr Asp Leu Arg Tyr
 210 215 220

Gly Asp Leu Leu Asp Phe His His Lys Asn Tyr His Pro Ser Asn Ala
 225 230 235 240

40 Lys Thr Phe Thr Tyr Gly Asn Leu Pro Leu Val Asp Thr Leu Lys Gln
 245 250 255

45 Leu Asn Glu Gln Phe Ser Gly Tyr Gly Lys Arg Ala Arg Lys Asp Lys
 260 265 270

50 Leu Leu Met Pro Ile Asp Leu Lys Lys Asp Ile Asp Val Lys Leu Leu
 275 280 285

55 Gly Gln Ile Asp Thr Met Leu Pro Pro Glu Lys Gln Thr Lys Ala Ser
 290 295 300

Met Thr Trp Ile Cys Gly Ala Pro Gln Asp Thr Tyr Asp Thr Phe Leu
 305 310 315 320

Leu Lys Val Leu Gly Asn Leu Leu Met Asp Gly His Ser Ser Val Met
 325 330 335
 5 Tyr Gln Lys Leu Ile Glu Ser Gly Ile Gly Leu Glu Phe Ser Val Asn
 340 345 350

Ser Gly Val Glu Pro Thr Thr Ala Val Asn Leu Leu Thr Val Gly Ile
 355 360 365
 10

Gln Gly Val Ser Asp Ile Glu Ile Phe Lys Asp Thr Val Asn Asn Ile
 370 375 380

15 Phe Gln Asn Leu Leu Glu Thr Glu His Pro Phe Asp Arg Lys Arg Ile
 385 390 395 400

20 Asp Ala Ile Ile Glu Gln Leu Glu Leu Ser Lys Lys Asp Gln Lys Ala
 405 410 415

25 Asp Phe Gly Leu Gln Leu Leu Tyr Ser Ile Leu Pro Gly Trp Thr Asn
 420 425 430

Lys Ile Asp Pro Phe Glu Ser Leu Leu Phe Glu Asp Val Leu Gln Arg
 435 440 445
 30

Phe Arg Gly Asp Leu Glu Thr Lys Gly Asp Thr Leu Phe Gln Asp Leu
 450 455 460

35 Ile Arg Lys Tyr Ile Val His Lys Pro Cys Phe Thr Phe Ser Ile Gln
 465 470 475 480

40 Gly Ser Glu Glu Phe Ser Lys Ser Leu Asp Asp Glu Glu Gln Thr Arg
 485 490 495

45 Leu Arg Glu Lys Ile Thr Ala Leu Asp Glu Gln Asp Lys Lys Asn Ile
 500 505 510

Phe Lys Arg Gly Ile Leu Leu Gln Glu Lys Gln Asn Glu Lys Glu Asp
 515 520 525
 50

Leu Ser Cys Leu Pro Thr Leu Gln Ile Lys Asp Ile Pro Arg Ala Gly
 530 535 540

55 Asp Lys Tyr Ser Ile Glu Gln Lys Asn Asn Thr Met Ser Arg Ile Thr
 545 550 555 560

Asp Thr Asn Gly Ile Thr Tyr Val Arg Gly Lys Arg Leu Leu Asn Asp
 565 570 575
 Ile Ile Pro Phe Glu Leu Phe Pro Tyr Leu Pro Leu Phe Ala Glu Ser
 580 585 590
 5

Leu Thr Asn Leu Gly Thr Thr Thr Glu Ser Phe Ser Glu Ile Glu Asp
 595 600 605
 10

Gln Ile Lys Leu His Thr Gly Gly Ile Ser Thr His Val Glu Val Thr
 610 615 620
 15

Ser Asp Pro Asn Thr Thr Glu Pro Arg Leu Ile Phe Gly Phe Asp Gly
 625 630 635 640
 20

Trp Ser Leu Asn Ser Lys Thr Asp His Ile Phe Glu Phe Trp Ser Lys
 645 650 655
 25

Ile Leu Leu Glu Thr Asp Phe His Lys Asn Ser Asp Lys Leu Lys Val
 660 665 670
 30

Leu Ile Arg Leu Leu Ala Ser Ser Asn Thr Ser Ser Val Ala Asp Ala
 675 680 685
 35

Gly His Ala Phe Ala Arg Gly Tyr Ser Ala Ala His Tyr Arg Ser Ser
 690 695 700
 40

Gly Ala Ile Asn Glu Thr Leu Asn Gly Ile Glu Gln Leu Gln Phe Ile
 705 710 715 720
 45

Asn Arg Leu His Ser Leu Leu Asp Asn Glu Glu Thr Phe Gln Arg Glu
 725 730 735
 50

Val Val Asp Lys Leu Thr Glu Leu Gln Lys Tyr Ile Val Asp Thr Asn
 740 745 750
 55

Asn Met Asn Phe Phe Ile Thr Ser Asp Ser Asp Val Gln Ala Lys Thr
 755 760 765
 60

Val Glu Ser Gln Ile Ser Lys Phe Met Glu Arg Leu Pro His Gly Ser
 770 775 780
 65

Cys Leu Pro Asn Gly Pro Lys Thr Ser Asp Tyr Pro Leu Ile Gly Ser
 785 790 795 800
 70

Lys Cys Lys His Thr Leu Ile Lys Phe Pro Phe Gln Val His Tyr Thr

805 810 815
 Ser Gln Ala Leu Leu Gly Val Pro Tyr Thr His Lys Asp Gly Ser Ala
 820 825 830

5
 Leu Gln Val Met Ser Asn Met Leu Thr Phe Lys His Leu His Arg Glu
 835 840 845

10 Val Arg Glu Lys Gly Gly Ala Tyr Gly Gly Gly Ala Ser Tyr Ser Ala
 850 855 860

15 Leu Ala Gly Ile Phe Ser Phe Tyr Ser Tyr Arg Asp Pro Gln Pro Leu
 865 870 875 880

20 Lys Ser Leu Glu Thr Phe Lys Asn Ser Gly Arg Tyr Ile Leu Asn Asp
 885 890 895

Ala Lys Trp Gly Val Thr Asp Leu Asp Glu Ala Lys Leu Thr Ile Phe
 900 905 910

25 Gln Gln Val Asp Ala Pro Lys Ser Pro Lys Gly Glu Gly Val Thr Tyr
 915 920 925

30 Phe Met Ser Gly Val Thr Asp Asp Met Lys Gln Ala Arg Arg Glu Gln
 930 935 940

35 Leu Leu Asp Val Ser Leu Leu Asp Val His Arg Val Ala Glu Lys Tyr
 945 950 955 960

40 Leu Leu Asn Lys Glu Gly Val Ser Thr Val Ile Gly Pro Gly Ile Glu
 965 970 975

Gly Lys Thr Val Ser Pro Asn Trp Glu Val Lys Glu Leu
 980 985

45
 <210> 5
 <211> 882
 <212> PRT
 <213> 栗酒裂殖酵母 (Schizosaccharomyces pombe)

50
 <400> 5

Met Asn Tyr Ala Lys Leu Ser Ile Ala Phe Ser Lys Lys Thr Ile Lys
 1 5 10 15

55 Thr His Asn Cys Arg Leu Phe Gln Arg Trp Leu His Val Gly Asp Lys
 20 25 30

Val His Asp Phe Arg Val Val Asp Thr Lys Lys Val Pro Glu Leu Gln
 35 40 45

5 Leu Asn Tyr Thr Arg Leu Lys His Glu Pro Thr Asn Ala Asp Met Ile
 50 55 60

10 His Leu Asp Arg Glu Asp Pro Asn Ser Val Phe Ser Ile Gly Phe Gln
 65 70 75 80

15 Thr Pro Ala Glu Asn Asp Glu Gly Ile Pro His Ile Leu Glu His Thr
 85 90 95

20 Thr Leu Cys Gly Ser Asn Lys Tyr Pro Val Arg Asp Pro Phe Phe Lys
 100 105 110

25 Met Leu Asn Arg Ser Leu Ala Thr Phe Met Asn Ala Phe Thr Ala Ser
 115 120 125

30 Asp Phe Thr Phe Tyr Pro Phe Ala Thr Val Asn Thr Thr Asp Tyr Lys
 130 135 140

35 Asn Leu Arg Asp Val Tyr Leu Asp Ala Thr Leu Phe Pro Lys Leu Arg
 145 150 155 160

40 Lys Leu Asp Phe Leu Gln Glu Gly Trp Arg Phe Glu His Ala Asp Val
 165 170 175

45 Asn Asp Lys Lys Ser Pro Ile Ile Phe Asn Gly Val Val Tyr Asn Glu
 180 185 190

50 Met Lys Gly Gln Val Ser Asp Ser Ser Tyr Ile Phe Tyr Met Leu Phe
 195 200 205

55 Gln Gln His Leu Phe Gln Gly Thr Ala Tyr Gly Phe Asn Ser Gly Gly
 210 215 220

60 Asp Pro Leu Ala Ile Pro Asp Leu Lys Tyr Glu Glu Leu Val Lys Phe
 225 230 235 240

65 His Arg Ser His Tyr His Pro Ser Asn Ala Lys Ile Leu Ser Tyr Gly
 245 250 255

70 Ser Phe Pro Leu Glu Asp Asn Leu Ser Ala Leu Ser Glu Thr Phe Arg
 260 265 270

75 Pro Phe Ser Lys Arg Glu Leu Asn Leu Pro Asn Thr Phe Leu Lys Glu

	275	280	285
5	Phe Asp Gln Glu Lys Arg Val Val Glu Tyr Gly Pro Leu Asp Pro Val 290 295 300		
10	Met Ala Pro Gly Arg Gln Val Lys Thr Ser Ile Ser Phe Leu Ala Asn 305 310 315 320		
15	Asp Thr Ser Asn Val Tyr Glu Thr Phe Ala Leu Lys Val Leu Ser Lys 325 330 335		
20	Leu Cys Phe Asp Gly Phe Ser Ser Pro Phe Tyr Lys Ala Leu Ile Glu 340 345 350		
25	Ser Gly Leu Gly Thr Asp Phe Ala Pro Asn Ser Gly Tyr Asp Ser Thr 355 360 365		
30	Thr Lys Arg Gly Ile Phe Ser Val Gly Leu Glu Gly Ala Ser Glu Glu 370 375 380		
35	Ser Leu Ala Lys Ile Glu Asn Leu Val Tyr Ser Ile Phe Asn Asp Leu 385 390 395 400		
40	Ala Leu Lys Gly Phe Glu Asn Glu Lys Leu Glu Ala Ile Leu His Gln 405 410 415		
45	Met Glu Ile Ser Leu Lys His Lys Ser Ala His Phe Gly Ile Gly Leu 420 425 430		
50	Ala Gln Ser Leu Pro Phe Asn Trp Phe Asn Gly Ala Asp Pro Ala Asp 435 440 445		
55	Trp Leu Ser Phe Asn Lys Gln Ile Glu Trp Leu Lys Gln Lys Asn Ser 450 455 460		
60	Asp Gly Lys Leu Phe Gln Lys Leu Ile Lys Lys Tyr Ile Leu Glu Asn 465 470 475 480		
65	Lys Ser Arg Phe Val Phe Thr Met Leu Pro Ser Ser Thr Phe Pro Gln 485 490 495		
70	Arg Leu Gln Glu Ala Glu Ala Lys Lys Leu Gln Glu Arg Thr Ser Lys 500 505 510		
75	Leu Thr Asp Glu Asp Ile Ala Glu Ile Glu Lys Thr Ser Val Lys Leu 515 520 525		

71

Thr Gln His Glu Val Val Glu Lys Ala Leu Gln Lys Phe Met Lys Ser
 770 775 780

5

Arg Gly Val Asn Gln Gln Thr Gln Thr Lys Ser Thr Ser Lys Glu Arg
 785 790 795 800

10

Asn Gly Ile Asn Ser Ile Lys Thr Tyr His Glu Leu Pro Phe Gln Thr
 805 810 815

15

Tyr Phe Ala Ala Lys Ser Cys Leu Gly Val Pro Tyr Thr His Pro Asp
 820 825 830

20

Gly Ala Pro Leu Gln Ile Leu Ser Ser Leu Leu Thr His Lys Tyr Leu
 835 840 845

His Gly Glu Ile Arg Glu Lys Gly Gly Ala Tyr Gly Ala Gly Leu Ser
 850 855 860

25

Tyr Ser Gly Ile Asp Gly Val Leu Ser Phe Phe Thr Tyr Arg Asp Ser
 865 870 875 880

30 Asp Pro

<210> 6
 35 <211> 973
 <212> PRT
 <213> 产气荚膜梭菌 (Clostridium perfringens)
 <400> 6

40

Met Asn Phe Lys Glu Asn Asn Ile Tyr Ser Gly Phe Lys Leu Leu Asn
 1 5 10 15

45

Ile Glu Asn Leu Asn Glu Ile Gly Gly Val Gly Leu Arg Phe Glu His
 20 25 30

50

Glu Lys Thr Lys Ala Lys Leu Ile Lys Ile Leu Ser Glu Asp Asp Asn
 35 40 45

Lys Cys Phe Ala Ile Gly Phe Arg Thr Pro Pro Glu Asn Ser Thr Gly
 50 55 60

55

Val Pro His Ile Leu Glu His Ser Val Leu Cys Gly Ser Arg Lys Phe
 65 70 75 80

	Asn Thr Lys Glu Pro Phe Val Glu Leu Leu Lys Gly Ser Leu Asn Thr	85	90	95
5	Phe Leu Asn Ala Met Thr Tyr Pro Asp Lys Thr Ile Tyr Pro Val Ala	100	105	110
10	Ser Arg Asn Glu Lys Asp Phe Met Asn Leu Met Asp Val Tyr Leu Asp	115	120	125
15	Ala Val Leu Tyr Pro Asn Ile Tyr Lys His Lys Glu Ile Phe Met Gln	130	135	140
20	Glu Gly Trp His Tyr Tyr Ile Glu Asn Lys Glu Asp Glu Leu Lys Tyr	145	150	155
	Asn Gly Val Val Tyr Asn Glu Met Lys Gly Ala Tyr Ser Ser Pro Asp	165	170	175
25	Ser Ile Leu Tyr Arg Lys Ile Pro Gln Thr Ile Tyr Pro Asp Thr Cys	180	185	190
30	Tyr Ala Leu Ser Ser Gly Gly Asp Pro Asp Glu Ile Pro Asn Leu Thr	195	200	205
35	Tyr Glu Glu Phe Val Glu Phe His Lys Lys Tyr Tyr His Pro Ser Asn	210	215	220
40	Ser Tyr Ile Phe Leu Tyr Gly Asn Gly Asp Thr Glu Lys Glu Leu Glu	225	230	235
	Phe Ile Asn Glu Glu Tyr Leu Lys Asn Phe Glu Tyr Lys Glu Ile Asp	245	250	255
45	Ser Glu Ile Lys Glu Gln Lys Ser Phe Glu Ser Met Lys Glu Glu Ser	260	265	270
50	Phe Thr Tyr Gly Ile Ala Glu Ser Glu Asp Leu Asn His Lys Ser Tyr	275	280	285
55	Tyr Ser Leu Asn Phe Val Ile Gly Asp Ala Thr Asp Gly Glu Lys Gly	290	295	300
	Leu Ala Phe Asp Val Leu Ala Tyr Leu Leu Thr Arg Ser Thr Ala Ala	305	310	315
	Pro Leu Lys Lys Ala Leu Ile Asp Ala Gly Ile Gly Lys Ala Val Ser			

	325	330	335
5	Gly Asp Phe Asp Asn Ser Thr Lys Gln Ser Ala Phe Thr Val Leu Val 340 345 350		
10	Lys Asn Ala Glu Leu Asn Lys Glu Glu Glu Phe Lys Lys Val Val Met 355 360 365		
15	Asp Thr Leu Lys Asp Leu Val Glu Asn Gly Ile Asp Lys Glu Leu Ile 370 375 380		
20	Gly Ser Tyr Pro Asn Gly Leu Ile Tyr Tyr Leu Lys Val Met Asp Ser 405 410 415		
25	Trp Leu Tyr Asp Gly Asp Pro Tyr Val His Leu Glu Tyr Glu Lys Asn 420 425 430		
30	Leu Glu Lys Ile Lys Ser Ala Leu Thr Ser Asn Tyr Phe Glu Asp Leu 435 440 445		
35	Ile Glu Arg Tyr Met Ile Asn Asn Thr His Ser Ser Leu Val Ser Leu 450 455 460		
40	His Pro Glu Lys Gly Ile Asn Glu Lys Lys Ser Ala Glu Leu Lys Lys 465 470 475 480		
45	Lys Leu Glu Glu Ile Lys Asn Ser Phe Asp Glu Lys Thr Leu Asn Glu 485 490 495		
50	Ile Ile Asp Asn Cys Lys Lys Leu Lys Glu Arg Gln Ser Thr Pro Asp 500 505 510		
55	Lys Lys Glu Asp Leu Glu Ser Ile Pro Met Leu Ser Leu Glu Asp Ile 515 520 525		
	Asp Lys Glu Ala Thr Lys Ile Pro Thr Glu Glu Lys Glu Ile Asp Gly 530 535 540		
	Ile Thr Thr Leu His His Asp Phe His Thr Asn Lys Ile Asp Tyr Val 545 550 555 560		
	Asn Phe Phe Phe Asn Thr Asn Ser Val Pro Glu Asp Leu Ile Pro Tyr 565 570 575		

Val Gly Leu Leu Cys Asp Ile Leu Gly Lys Cys Gly Thr Glu Asn Tyr
 580 585 590
 5

Asp Tyr Ser Lys Leu Ser Asn Ala Ile Asn Ile Ser Thr Gly Gly Ile
 595 600 605
 10

Ser Phe Gly Ala Ile Thr Phe Ala Asn Leu Lys Lys Asn Asn Glu Phe
 610 615 620
 15

Arg Pro Tyr Leu Glu Ile Ser Tyr Lys Ala Leu Ser Ser Lys Thr Asn
 625 630 635 640
 20

Lys Ala Ile Glu Leu Val Asp Glu Ile Val Asn His Thr Asp Leu Asp
 645 650 655
 25

Asp Met Asp Arg Ile Met Gln Ile Ile Arg Glu Lys Arg Ala Arg Leu
 660 665 670
 30

Glu Gly Ala Ile Phe Asp Ser Gly His Arg Ile Ala Met Lys Lys Val
 675 680 685
 35

Leu Ser Tyr Ser Thr Asn Arg Gly Ala Tyr Asp Glu Lys Ile Ser Gly
 690 695 700
 40

Leu Asp Tyr Tyr Asp Phe Leu Val Asn Ile Glu Lys Glu Asp Lys Lys
 705 710 715 720
 45

Ser Thr Ile Ser Asp Ser Leu Lys Lys Val Arg Asp Leu Ile Phe Asn
 725 730 735
 50

Lys Gly Asn Met Leu Ile Ser Tyr Ser Gly Lys Glu Glu Glu Tyr Glu
 740 745 750
 55

Asn Phe Lys Glu Lys Val Lys Tyr Leu Ile Ser Lys Thr Asn Asn Asn
 755 760 765
 60

Asp Phe Glu Lys Glu Glu Tyr Asn Phe Glu Leu Gly Lys Lys Asn Glu
 770 775 780
 65

Gly Leu Leu Thr Gln Gly Asn Val Gln Tyr Val Ala Lys Gly Gly Asn
 785 790 795 800
 70

Tyr Lys Thr His Gly Tyr Lys Tyr Ser Gly Ala Leu Ser Leu Leu Glu
 805 810 815
 75

Ser Ile Leu Gly Phe Asp Tyr Leu Trp Asn Ala Val Arg Val Lys Gly
 820 825 830

5

Gly Ala Tyr Gly Val Phe Ser Asn Phe Arg Arg Asp Gly Gly Ala Tyr
 835 840 845

10

Ile Val Ser Tyr Arg Asp Pro Asn Ile Lys Ser Thr Leu Glu Ala Tyr
 850 855 860

15

Asp Asn Ile Pro Lys Tyr Leu Asn Asp Phe Glu Ala Asp Glu Arg Glu
 865 870 875 880

20

Met Thr Lys Tyr Ile Ile Gly Thr Ile Arg Lys Tyr Asp Gln Pro Ile
 885 890 895

Ser Asn Gly Ile Lys Gly Asp Ile Ala Val Ser Tyr Tyr Leu Ser Asn
 900 905 910

25

Phe Thr Tyr Glu Asp Leu Gln Lys Glu Arg Glu Glu Ile Ile Asn Ala
 915 920 925

30

Asp Val Glu Lys Ile Lys Ser Phe Ala Pro Met Ile Lys Asp Leu Met
 930 935 940

35

Lys Glu Asp Tyr Ile Cys Val Leu Gly Asn Glu Glu Lys Ile Lys Glu
 945 950 955 960

Asn Lys Asp Leu Phe Asn Asn Ile Lys Ser Val Ile Lys
 965 970

40

<210> 7
 <211> 971
 <212> PRT
 45 <213> 布氏疏螺旋体 (*Borrelia burgdorferi*)
 <400> 7

50

Met Lys Lys Lys Lys Ile Phe Lys Leu Ile Ser Lys Thr Tyr Leu Glu
 1 5 10 15

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

55

Val Phe His Leu Lys Ser Asp Ser Phe Lys Glu Asn Ala Phe Cys Ile
 35 40 45

	Ala	Phe	Lys	Thr	Ile	Pro	Ser	Asn	Asn	Thr	Gly	Val	Ala	His	Val	Leu
	50						55					60				
5	Glu	His	Thr	Ile	Phe	Cys	Gly	Ser	Ser	Lys	Tyr	Lys	Ile	Lys	Asp	Pro
	65					70					75					80
10	Phe	Leu	Tyr	Leu	Leu	Lys	Gly	Ser	Leu	Asn	Thr	Phe	Leu	Asn	Ala	Met
				85						90					95	
15	Thr	Phe	Pro	Asp	Lys	Thr	Ile	Tyr	Pro	Ala	Ala	Ser	Thr	Ile	Glu	Lys
				100					105					110		
20	Asp	Tyr	Phe	Asn	Leu	Phe	Asn	Ile	Tyr	Ala	Asp	Ser	Ile	Phe	Asn	Pro
			115					120					125			
25	Leu	Leu	Lys	Lys	Glu	Ser	Phe	Met	Gln	Glu	Gly	Tyr	Asn	Ile	Asn	Pro
	130						135					140				
30	Lys	Asp	Phe	Lys	Val	Ser	Gly	Ile	Val	Phe	Asn	Glu	Met	Lys	Gly	Ser
	145					150					155					160
35	Tyr	Ser	Asn	Lys	Asn	Ser	Leu	Ile	Asn	Glu	Ile	Val	Ser	Ser	Ser	Leu
				165						170					175	
40	Phe	Glu	Glu	Gly	Ala	Tyr	Lys	Tyr	Asp	Ser	Gly	Gly	Ile	Pro	Thr	Asn
				180					185					190		
45	Ile	Ile	Asp	Leu	Thr	Tyr	Glu	Ser	Phe	Leu	Asp	Phe	Tyr	Lys	Lys	Tyr
			195					200					205			
50	Tyr	Thr	Leu	Glu	Asn	Cys	Lys	Ile	Phe	Leu	Cys	Gly	Asn	Thr	Gln	Thr
	210						215					220				
55	Glu	Lys	Asn	Leu	Asn	Phe	Ile	Glu	Lys	Tyr	Ile	Ile	Arg	Pro	Tyr	Lys
	225					230					235					240
60	Lys	Glu	Lys	Ser	Asn	Val	Asn	Ile	Asn	Ile	Glu	Asn	Val	Lys	Arg	Trp
					245				250						255	
65	Glu	Lys	Gly	Lys	Lys	Leu	Thr	Tyr	Lys	Ile	Pro	Lys	Glu	Asn	Asp	Asn
				260					265				270			
70	Ser	Leu	Gly	Val	Tyr	Thr	Ile	Asn	Trp	Leu	Cys	Thr	Glu	Ile	Asn	Asn
		275						280					285			
75	Ile	Glu	Asp	Ser	Ile	Gly	Leu	Glu	Ile	Leu	Ser	Glu	Ile	Leu	Leu	Asp

	290	295	300	
5	Asp Ser Cys Ser Phe Thr Ile Asn Ile Leu Lys Ser Gly Ile Gly Glu 305 310 315 320			
10	Asp Ile Ala His Ile Ser Gly Ile Asn Thr Asp Leu Lys Glu Ser Ile 325 330 335			
15	Phe Ser Phe Gly Leu Gln Asn Val Val Glu Asn Lys Glu Lys Glu Phe 340 345 350			
20	Lys Asn Leu Val Phe Ser Glu Leu Lys Asn Leu Val Lys Asn Lys Ile 355 360 365			
25	Pro Lys Glu Leu Ile Lys Gly Ile Leu Phe Gly Tyr Glu Phe Ala Leu 370 375 380			
30	Lys Glu Glu Lys Gly Gln Asn Phe Pro Ile Ala Leu Met Ile Lys Ser 385 390 395 400			
35	Phe Lys Gly Trp Leu Asn Gly Leu His Pro Ile Lys Thr Leu Gln Thr 405 410 415			
40	Ser Tyr Tyr Ile Asn Glu Ile Thr Asn Lys Leu Glu Lys Gly Ile Tyr 420 425 430			
45	Tyr Phe Glu Asn Leu Ile Glu Lys Tyr Leu Ile Phe Asn Asn His Tyr 435 440 445			
50	Thr Leu Ile Ser Phe Ile Pro Ser His Asp Thr Glu Lys Glu Met Glu 450 455 460			
55	Glu Glu Ile Glu Lys Lys Leu Met Ala Arg Glu Ile Glu Ile Lys Gln 465 470 475 480			
60	Asn Pro Glu Glu Phe Leu Gln Phe Lys Lys Asp Tyr Asn Gln Phe Lys 485 490 495			
65	Lys Tyr Gln Asn Lys Lys Asp Ser Lys Ala Asp Ile Ala Lys Leu Pro 500 505 510			
70	Leu Leu Lys Ile Glu Asp Leu Pro Lys Gln Ile Glu Lys Ser Leu Asp 515 520 525			
75	Leu Asn Glu Ile Lys Glu Leu Asn Leu His Ser Phe Lys Phe Lys Ser 530 535 540			

```

      Asn Asn Ile Phe Asn Val Asn Leu Phe Phe Lys Leu Asp Phe Leu Glu
545                               550                     555                   560
5
      Lys Glu Asp Tyr Ile Tyr Leu Ser Leu Phe Lys Arg Ala Leu Gln Asp
                               565                     570                   575

10
      Leu Ser Thr Lys Asn Tyr Ser Tyr Ile Asn Ile Asn Asn Lys Ile Gln
                               580                     585                   590

15
      Asn Thr Leu Gly Gln Ile Asn Ile Ser Glu Ser Tyr Asp Glu Asp Ile
                               595                     600                   605

20
      Asp Gly Asn Ile Leu Asn Ser Phe Asn Ile Ser Phe Lys Ser Phe Asn
610                               615                     620

      Asn Lys Val Lys Glu Ser Phe Glu Leu Ile Lys Glu Ile Leu Ile Asn
625                               630                     635                   640
25
      Ile Asn Phe His Asp Tyr Glu Arg Leu Lys Glu Ile Thr Leu Ser Leu
                               645                     650                   655

30
      Lys Asn Asp Phe Lys Ser Leu Leu Ile Pro Lys Gly His Leu Leu Ala
                               660                     665                   670

35
      Met Leu Arg Ser Lys Ser Lys Leu Lys Leu Asn Glu Tyr Leu Lys Glu
                               675                     680                   685

40
      Leu Gln Asn Gly Ile Thr Gly Arg Glu Phe Trp Gln Lys Ala Lys Thr
690                               695                     700

      Asp Thr Glu Ser Leu Lys Glu Ile Ala Asn Lys Leu Asp Asn Leu Lys
705                               710                     715                   720
45
      Asn Lys Ile Ile Leu Lys Asn Asn Leu Ser Ala Leu Ile Met Gly Asn
                               725                     730                   735
      Thr Asp Asp Ile Leu Lys Asn Leu Glu Asn Glu Phe Phe Asn Leu Lys
50                               740                     745                   750

      Glu Ser Leu Glu Glu Ser Asn His Tyr Asn Gly Leu Leu Asn Leu Asp
755                               760                     765

55
      Ala Asn Ser Lys Ala Leu Arg Glu Ile Ile Ile Ile Gln Ser Lys Val
770                               775                     780

```

Ala Phe Asn Ala Ile Cys Phe Pro Ser Tyr Lys Ile Asn Asp Glu Asn
 785 790 795 800

5 Tyr Pro Lys Ala Asn Phe Leu Glu His Val Leu Arg Ser Gly Ile Phe
 805 810 815

10 Trp Glu Lys Ile Arg Val Met Gly Gly Ala Tyr Gly Ala Ser Ala Ser
 820 825 830

15 Ile Ala Asn Gly Ile Phe Ser Phe Ala Ser Tyr Arg Asp Pro Asn Phe
 835 840 845

20 Thr Lys Thr Tyr Gln Ala Phe Glu Lys Ser Leu Glu Glu Leu Ala Asn
 850 855 860

Asn Lys Met Thr Asp Asp Glu Ile Tyr Thr Tyr Leu Ile Gly Leu Ile
 865 870 875 880

25 Gly Thr Asn Ile Tyr Val Lys Thr Lys Ala Thr Glu Ala Leu Gln Ser
 885 890 895

30 Tyr Arg Arg Lys Met Leu Asn Ile Ser Asp Ser Leu Arg Gln Asp Ile
 900 905 910

35 Arg Asn Ala Tyr Phe Thr Ile Thr Pro Gln Asp Ile Lys Glu Ile Ser
 915 920 925

40 Thr Lys Ile Leu Thr Gln Ile Arg Gln His Asn Ser Ile Ala Ser Leu
 930 935 940

Val Asn Asn Gln Ile Tyr Glu Glu Glu Lys Asn Asn Leu Glu Lys Leu
 945 950 955 960

45 Ile Gly Lys Glu Tyr Ser Leu Lys Lys Ile Tyr
 965 970

<210> 8
 <211> 995
 50 <212> PRT
 <213> 秀丽隐杆线虫 (Caenorhabditis elegans)

<400> 8

55 Met Ser Ala Ser Lys Leu Trp Ser Cys Thr Glu Thr Val Leu Asn Gly
 1 5 10 15

Gly Ile Lys Leu Phe Leu Tyr Ser Ser Lys Asn Thr Lys Leu Arg Val

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

	260	265	270
5	Arg Gly Ala Val Glu Val Ala Trp Phe Ala His Ser Pro Ser Asp Leu 275 280 285		
10	Glu Thr His Ser Ser Leu His Val Leu Phe Asp Tyr Leu Ser Asn Thr 290 295 300		
15	Ser Val Ala Pro Leu Gln Lys Asp Phe Ile Leu Leu Glu Asp Pro Leu 305 310 315 320		
	Ala Ser Ser Val Ser Phe His Ile Ala Glu Gly Val Arg Cys Asp Leu 325 330 335		
20	Arg Leu Asn Phe Ala Gly Val Pro Val Glu Lys Leu Asp Glu Cys Ala 340 345 350		
25	Pro Lys Phe Phe Asp Lys Thr Val Arg Glu His Leu Glu Glu Ala Asn 355 360 365		
30	Phe Asp Met Glu Arg Met Gly Tyr Leu Ile Asp Gln Thr Ile Leu Asn 370 375 380		
	Glu Leu Val Lys Leu Glu Thr Asn Ala Pro Lys Asp Ile Met Ser His 385 390 395 400		
35	Ile Ile Gly His Gln Leu Phe Asp Asn Glu Asp Glu Glu Leu Phe Lys 405 410 415		
40	Lys Arg Thr Asn Glu Ile Asp Phe Leu Lys Lys Leu Lys Ser Glu Pro 420 425 430		
45	Ala Ser Tyr Trp Val Gln Leu Val Asn Lys Tyr Phe Thr Ala Pro Ser 435 440 445		
	Ala Thr Val Ile Gly Val Pro Asn Glu Glu Leu Val Asp Lys Ile Ala 450 455 460		
50	Glu Glu Glu Glu Lys Arg Ile Ala Ala Gln Cys Glu Lys Leu Gly Lys 465 470 475 480		
55	Lys Gly Leu Glu Glu Ala Gly Lys Ser Leu Glu Ala Ala Ile Leu Glu 485 490 495		
	Asn Thr Ala Asn His Pro Ser Ala Glu Leu Leu Asp Gln Leu Ile Val 500 505 510		

Leu Ser Ser Asp Leu Trp Thr Trp Val Gln Ala Asp Pro Arg Phe Gly
 755 760 765

5

Pro Gly His Gln Phe Ser Ala Glu Ala Gly Glu Asn Val Ser Leu Glu
 770 775 780

10

Leu Gly Lys Glu Leu Leu Ile Gly Val Gly Gly Ser Glu Ser Ser Phe
 785 790 795 800

15

Ile Tyr Gln Thr Ser Phe Leu Asp Ala Asn Trp Asn Ser Glu Glu Leu
 805 810 815

20

Ile Pro Ala Met Ile Phe Gly Gln Tyr Leu Ser Gln Cys Glu Gly Pro
 820 825 830

Leu Trp Arg Ala Ile Arg Gly Asp Gly Leu Ala Tyr Gly Ala Asn Val
 835 840 845

25

Phe Val Lys Pro Asp Arg Lys Gln Ile Thr Leu Ser Leu Tyr Arg Cys
 850 855 860

30

Ala Gln Pro Ala Val Ala Tyr Glu Arg Thr Arg Asp Ile Ile Arg Lys
 865 870 875 880

35

Ile Val Glu Ser Gly Glu Ile Ser Lys Ala Glu Phe Glu Gly Ala Lys
 885 890 895

40

Arg Ser Thr Val Phe Glu Met Met Lys Arg Glu Gly Thr Val Ser Gly
 900 905 910

45

Ala Ala Lys Ile Ser Ile Leu Asn Asn Phe Arg Gln Thr Pro His Pro
 915 920 925
 Phe Asn Ile Asp Leu Cys Arg Arg Ile Trp Asn Leu Thr Ser Glu Glu
 930 935 940

50

Met Val Lys Ile Gly Gly Pro Pro Leu Ala Arg Leu Phe Asp Glu Lys
 945 950 955 960

Cys Phe Val Arg Ser Ile Ala Val His Pro Ser Lys Leu Asn Glu Met
 965 970 975

55

Lys Lys Ala Phe Pro Gly Ser Ser Lys Ile Lys Ile Ser Asp Leu Gln
 980 985 990

Phe Ala Cys
995

5 <210> 9
 <211> 962
 <212> PRT
 <213> 大肠杆菌 (Escherichia coli)

10 <400> 9

Met Pro Arg Ser Thr Trp Phe Lys Ala Leu Leu Leu Leu Val Ala Leu
 1 5 10 15

15
 Trp Ala Pro Leu Ser Gln Ala Glu Thr Gly Trp Gln Pro Ile Gln Glu
 20 25 30

20 Thr Ile Arg Lys Ser Asp Lys Asp Asn Arg Gln Tyr Gln Ala Ile Arg
 35 40 45

25 Leu Asp Asn Gly Met Val Val Leu Leu Val Ser Asp Pro Gln Ala Val
 50 55 60

30 Lys Ser Leu Ser Ala Leu Val Val Pro Val Gly Ser Leu Glu Asp Pro
 65 70 75 80

Glu Ala Tyr Gln Gly Leu Ala His Tyr Leu Glu His Met Ser Leu Met
 85 90 95

35
 Gly Ser Lys Lys Tyr Pro Gln Ala Asp Ser Leu Ala Glu Tyr Leu Lys
 100 105 110

40 Met His Gly Gly Ser His Asn Ala Ser Thr Ala Pro Tyr Arg Thr Ala
 115 120 125

45 Phe Tyr Leu Glu Val Glu Asn Asp Ala Leu Pro Gly Ala Val Asp Arg
 130 135 140

Leu Ala Asp Ala Ile Ala Glu Pro Leu Leu Asp Lys Lys Tyr Ala Glu
 145 150 155 160

50
 Arg Glu Arg Asn Ala Val Asn Ala Glu Leu Thr Met Ala Arg Thr Arg
 165 170 175

55 Asp Gly Met Arg Met Ala Gln Val Ser Ala Glu Thr Ile Asn Pro Ala
 180 185 190

His Pro Gly Ser Lys Phe Ser Gly Gly Asn Leu Glu Thr Leu Ser Asp

	195	200	205
5	Lys Pro Gly Asn Pro Val 210	Gln Gln Ala Leu Lys 215	Asp Phe His Glu Lys 220
10	Tyr Tyr Ser Ala Asn Leu Met Lys Ala Val 225	Ile Tyr Ser Asn Lys Pro 235	
15	Leu Pro Glu Leu Ala Lys Met Ala Ala Asp Thr Phe Gly Arg Val Pro 245		255
20	Asn Lys Glu Ser Lys Lys Pro Glu Ile Thr Val Pro Val Val Thr Asp 260		270
25	Ala Gln Lys Gly Ile Ile Ile His Tyr Val Pro Ala Leu Pro Arg Lys 275		285
30	Val Leu Arg Val Glu Phe Arg Ile Asp Asn Asn Ser Ala Lys Phe Arg 290		300
35	Ser Lys Thr Asp Glu Leu Ile Thr Tyr Leu Ile Gly Asn Arg Ser Pro 305		320
40	Gly Thr Leu Ser Asp Trp Leu Gln Lys Gln Gly Leu Val Glu Gly Ile 325		335
45	Ser Ala Asn Ser Asp Pro Ile Val Asn Gly Asn Ser Gly Val Leu Ala 340		350
50	Ile Ser Ala Ser Leu Thr Asp Lys Gly Leu Ala Asn Arg Asp Gln Val 355		365
55	Val Ala Ala Ile Phe Ser Tyr Leu Asn Leu Leu Arg Glu Lys Gly Ile 370		380
60	Asp Lys Gln Tyr Phe Asp Glu Leu Ala Asn Val Leu Asp Ile Asp Phe 385		400
65	Arg Tyr Pro Ser Ile Thr Arg Asp Met Asp Tyr Val Glu Trp Leu Ala 405		415
70	Asp Thr Met Ile Arg Val Pro Val Glu His Thr Leu Asp Ala Val Asn 420		430
75	Ile Ala Asp Arg Tyr Asp Ala Lys Ala Val Lys Glu Arg Leu Ala Met 435		445

Met Thr Pro Gln Asn Ala Arg Ile Trp Tyr Ile Ser Pro Lys Glu Pro
 450 455 460
 5

His Asn Lys Thr Ala Tyr Phe Val Asp Ala Pro Tyr Gln Val Asp Lys
 465 470 475 480

10
 Ile Ser Ala Gln Thr Phe Ala Asp Trp Gln Lys Lys Ala Ala Asp Ile
 485 490 495

15 Ala Leu Ser Leu Pro Glu Leu Asn Pro Tyr Ile Pro Asp Asp Phe Ser
 500 505 510

20 Leu Ile Lys Ser Glu Lys Lys Tyr Asp His Pro Glu Leu Ile Val Asp
 515 520 525

Glu Ser Asn Leu Arg Val Val Tyr Ala Pro Ser Arg Tyr Phe Ala Ser
 530 535 540
 25

Glu Pro Lys Ala Asp Val Ser Leu Ile Leu Arg Asn Pro Lys Ala Met
 545 550 555 560

30
 Asp Ser Ala Arg Asn Gln Val Met Phe Ala Leu Asn Asp Tyr Leu Ala
 565 570 575

35 Gly Leu Ala Leu Asp Gln Leu Ser Asn Gln Ala Ser Val Gly Gly Ile
 580 585 590

Ser Phe Ser Thr Asn Ala Asn Asn Gly Leu Met Val Asn Ala Asn Gly
 595 600 605
 40 Tyr Thr Gln Arg Leu Pro Gln Leu Phe Gln Ala Leu Leu Glu Gly Tyr
 610 615 620

45 Phe Ser Tyr Thr Ala Thr Glu Asp Gln Leu Glu Gln Ala Lys Ser Trp
 625 630 635 640

Tyr Asn Gln Met Met Asp Ser Ala Glu Lys Gly Lys Ala Phe Glu Gln
 645 650 655
 50

Ala Ile Met Pro Ala Gln Met Leu Ser Gln Val Pro Tyr Phe Ser Arg
 660 665 670
 55

Asp Glu Arg Arg Lys Ile Leu Pro Ser Ile Thr Leu Lys Glu Val Leu
 675 680 685

Ala Tyr Arg Asp Ala Leu Lys Ser Gly Ala Arg Pro Glu Phe Met Val
 690 695 700

5

Ile Gly Asn Met Thr Glu Ala Gln Ala Thr Thr Leu Ala Arg Asp Val
 705 710 715 720

10

Gln Lys Gln Leu Gly Ala Asp Gly Ser Glu Trp Cys Arg Asn Lys Asp
 725 730 735

15

Val Val Val Asp Lys Lys Gln Ser Val Ile Phe Glu Lys Ala Gly Asn
 740 745 750

20

Ser Thr Asp Ser Ala Leu Ala Ala Val Phe Val Pro Thr Gly Tyr Asp
 755 760 765

25

Glu Tyr Thr Ser Ser Ala Tyr Ser Ser Leu Leu Gly Gln Ile Val Gln
 770 775 780

30

Pro Trp Phe Tyr Asn Gln Leu Arg Thr Glu Glu Gln Leu Gly Tyr Ala
 785 790 795 800

35

Val Phe Ala Phe Pro Met Ser Val Gly Arg Gln Trp Gly Met Gly Phe
 805 810 815

40

Leu Leu Gln Ser Asn Asp Lys Gln Pro Ser Phe Leu Trp Glu Arg Tyr
 820 825 830

45

Lys Ala Phe Phe Pro Thr Ala Glu Ala Lys Leu Arg Ala Met Lys Pro
 835 840 845

50

Asp Glu Phe Ala Gln Ile Gln Gln Ala Val Ile Thr Gln Met Leu Gln
 850 855 860

55

Ala Pro Gln Thr Leu Gly Glu Glu Ala Ser Lys Leu Ser Lys Asp Phe
 865 870 875 880

60

Asp Arg Gly Asn Met Arg Phe Asp Ser Arg Asp Lys Ile Val Ala Gln
 885 890 895

65

Ile Lys Leu Leu Thr Pro Gln Lys Leu Ala Asp Phe Phe His Gln Ala
 900 905 910

70

Val Val Glu Pro Gln Gly Met Ala Ile Leu Ser Gln Ile Ser Gly Ser
 915 920 925

Gln Asn Gly Lys Ala Glu Tyr Val His Pro Glu Gly Trp Lys Val Trp
 930 935 940

5 Glu Asn Val Ser Ala Leu Gln Gln Thr Met Pro Leu Met Ser Glu Lys
 945 950 955 960

Asn Glu
 10

<210> 10
 <211> 1161
 15 <212> PRT
 <213> 人 (Homo sapiens)
 <400> 10

20 Met Leu Arg Arg Val Ala Val Ala Ala Val Phe Ala Thr Gly Arg Lys
 1 5 10 15

25 Leu Arg Cys Glu Ala Gly Arg Asp Val Thr Ala Val Gly Arg Ile Glu
 20 25 30

Ala Arg Gly Leu Cys Glu Glu Ser Ala Lys Pro Phe Pro Thr Leu Thr
 35 40 45
 30

Met Pro Gly Arg Asn Lys Ala Lys Ser Thr Cys Ser Cys Pro Asp Leu
 50 55 60

35 Gln Pro Asn Gly Gln Asp Leu Gly Glu Ser Gly Arg Val Ala Arg Leu
 65 70 75 80
 Gly Ala Asp Glu Ser Glu Glu Glu Gly Arg Ser Leu Ser Asn Val Gly
 85 90 95

40 Asp Pro Glu Ile Ile Lys Ser Pro Ser Asp Pro Lys Gln Tyr Arg Tyr
 100 105 110

45 Ile Lys Leu Gln Asn Gly Leu Gln Ala Leu Leu Ile Ser Asp Leu Ser
 115 120 125

50 Asn Val Glu Gly Lys Thr Gly Asn Ala Thr Asp Glu Glu Glu Glu Glu
 130 135 140

Glu Glu Glu Glu Glu Glu Gly Glu Glu Glu Glu Glu Glu Glu Asp
 55 145 150 155 160

Asp Asp Asp Asp Asp Asp Glu Asp Ser Gly Ala Glu Ile Gln Asp Asp
 165 170 175

Asp Glu Glu Gly Phe Asp Asp Glu Glu Glu Phe Asp Asp Asp Glu His
 180 185 190
 5

Asp Asp Asp Asp Leu Asp Asn Glu Glu Asn Glu Leu Glu Glu Leu Glu
 195 200 205
 10

Glu Arg Val Glu Ala Arg Lys Lys Thr Thr Glu Lys Gln Ser Ala Ala
 210 215 220
 15

Ala Leu Cys Val Gly Val Gly Ser Phe Ala Asp Pro Asp Asp Leu Pro
 225 230 235 240
 20

Gly Leu Ala His Phe Leu Glu His Met Val Phe Met Gly Ser Leu Lys
 245 250 255
 25

Tyr Pro Asp Glu Asn Gly Phe Asp Ala Phe Leu Lys Lys His Gly Gly
 260 265 270
 30

Ser Asp Asn Ala Ser Thr Asp Cys Glu Arg Thr Val Phe Gln Phe Asp
 275 280 285
 35

Val Gln Arg Lys Tyr Phe Lys Glu Ala Leu Asp Arg Trp Ala Gln Phe
 290 295 300
 40

Phe Ile His Pro Leu Met Ile Arg Asp Ala Ile Asp Arg Glu Val Glu
 305 310 315 320
 Ala Val Asp Ser Glu Tyr Gln Leu Ala Arg Pro Ser Asp Ala Asn Arg
 325 330 335
 45

Lys Glu Met Leu Phe Gly Ser Leu Ala Arg Pro Gly His Pro Met Gly
 340 345 350
 50

Lys Phe Phe Trp Gly Asn Ala Glu Thr Leu Lys His Glu Pro Lys Lys
 355 360 365
 55

Asn Asn Ile Asp Thr His Ala Arg Leu Arg Glu Phe Trp Met Arg Tyr
 370 375 380
 Tyr Ser Ala His Tyr Met Thr Leu Val Val Gln Ser Lys Glu Thr Leu
 385 390 395 400
 Asp Thr Leu Glu Lys Trp Val Thr Glu Ile Phe Ser Gln Ile Pro Asn
 405 410 415

Asn Gly Leu Pro Lys Pro Asn Phe Ser His Leu Thr Asp Pro Phe Asp
 420 425 430

5

Thr Pro Ala Phe Asn Lys Leu Tyr Arg Val Val Pro Ile Arg Lys Ile
 435 440 445

10

His Ala Leu Thr Ile Thr Trp Ala Leu Pro Pro Gln Gln Gln His Tyr
 450 455 460

15

Arg Val Lys Pro Leu His Tyr Ile Ser Trp Leu Val Gly His Glu Gly
 465 470 475 480

20

Lys Gly Ser Ile Leu Ser Tyr Leu Arg Lys Lys Cys Trp Ala Leu Ala
 485 490 495

Leu Phe Gly Gly Asn Gly Glu Thr Gly Phe Glu Gln Asn Ser Thr Tyr
 500 505 510

25

Ser Val Phe Ser Ile Ser Ile Thr Leu Thr Asp Glu Gly Tyr Glu His
 515 520 525

30

Phe Tyr Glu Val Ala His Thr Val Phe Gln Tyr Leu Lys Met Leu Gln
 530 535 540

35

Lys Leu Gly Pro Glu Lys Arg Val Phe Glu Glu Ile Gln Lys Ile Glu
 545 550 555 560
 Asp Asn Glu Phe His Tyr Gln Glu Gln Thr Asp Pro Val Glu Tyr Val
 565 570 575

40

Glu Asn Met Cys Glu Asn Met Gln Leu Tyr Pro Arg Gln Asp Phe Leu
 580 585 590

45

Thr Gly Asp Gln Leu Leu Phe Glu Tyr Lys Pro Glu Val Ile Ala Glu
 595 600 605

50

Ala Leu Asn Gln Leu Val Pro Gln Lys Ala Asn Leu Val Leu Leu Ser
 610 615 620

Gly Ala Asn Glu Gly Arg Cys Asp Leu Lys Glu Lys Trp Phe Gly Thr
 625 630 635 640

55

Gln Tyr Ser Ile Glu Asp Ile Glu Asn Ser Trp Thr Glu Leu Trp Lys
 645 650 655

	Ser	Asn	Phe	Asp	Leu	Asn	Ser	Asp	Leu	His	Leu	Pro	Ala	Glu	Asn	Lys	
				660					665					670			
5	Tyr	Ile	Ala	Thr	Asp	Phe	Thr	Leu	Lys	Ala	Phe	Asp	Cys	Pro	Glu	Thr	
			675					680					685				
10	Glu	Tyr	Pro	Ala	Lys	Ile	Val	Asn	Thr	Pro	Gln	Gly	Cys	Leu	Trp	Tyr	
		690					695					700					
15	Lys	Lys	Asp	Asn	Lys	Phe	Lys	Ile	Pro	Lys	Ala	Tyr	Ile	Arg	Phe	His	
	705					710					715					720	
	Leu	Ile	Ser	Pro	Leu	Ile	Gln	Lys	Ser	Ala	Ala	Asn	Val	Val	Leu	Phe	
					725					730					735		
20	Asp	Ile	Phe	Val	Asn	Ile	Leu	Thr	His	Asn	Leu	Ala	Glu	Pro	Ala	Tyr	
				740					745					750			
25	Glu	Ala	Asp	Val	Ala	Gln	Leu	Glu	Tyr	Lys	Leu	Val	Ala	Gly	Glu	His	
			755					760						765			
30	Gly	Leu	Ile	Ile	Arg	Val	Lys	Gly	Phe	Asn	His	Lys	Leu	Pro	Leu	Leu	
	770						775					780					
	Phe	Gln	Leu	Ile	Ile	Asp	Tyr	Leu	Thr	Glu	Phe	Ser	Ser	Thr	Pro	Ala	
	785					790					795					800	
35	Val	Phe	Thr	Met	Ile	Thr	Glu	Gln	Leu	Lys	Lys	Thr	Tyr	Phe	Asn	Ile	
					805					810					815		
40	Leu	Ile	Lys	Pro	Glu	Thr	Leu	Ala	Lys	Asp	Val	Arg	Leu	Leu	Ile	Leu	
				820					825					830			
45	Glu	Tyr	Ser	Arg	Trp	Ser	Met	Ile	Asp	Lys	Tyr	Arg	Ala	Leu	Met	Asp	
			835					840					845				
	Gly	Leu	Ser	Leu	Glu	Ser	Leu	Leu	Asn	Phe	Val	Lys	Asp	Phe	Lys	Ser	
		850					855					860					
50	Gln	Leu	Phe	Val	Glu	Gly	Leu	Val	Gln	Gly	Asn	Val	Thr	Ser	Thr	Glu	
	865					870					875					880	
55	Ser	Met	Asp	Phe	Leu	Arg	Tyr	Val	Val	Asp	Lys	Leu	Asn	Phe	Val	Pro	
					885					890					895		
	Leu	Glu	Arg	Glu	Met	Pro	Val	Gln	Phe	Gln	Val	Val	Glu	Leu	Pro	Ser	

	900	905	910
5	Gly His His Leu Cys Lys Val Arg Ala Leu Asn Lys Gly Asp Ala Asn 915 920 925		
10	Ser Glu Val Thr Val Tyr Tyr Gln Ser Gly Thr Arg Ser Leu Arg Glu 930 935 940		
15	Tyr Thr Leu Met Glu Leu Leu Val Met His Met Glu Glu Pro Cys Phe 945 950 955 960		
20	Asp Phe Leu Arg Thr Lys Gln Thr Leu Gly Tyr His Val Tyr Pro Thr 965 970 975		
25	Cys Arg Asn Thr Ser Gly Ile Leu Gly Phe Ser Val Thr Val Gly Thr 980 985 990		
30	Gln Ala Thr Lys Tyr Asn Ser Glu Thr Val Asp Lys Lys Ile Glu Glu 995 1000 1005		
35	Phe Leu Ser Ser Phe Glu Glu Lys Ile Glu Asn Leu Thr Glu Asp 1010 1015 1020		
40	Ala Phe Asn Thr Gln Val Thr Ala Leu Ile Lys Leu Lys Glu Cys 1025 1030 1035		
45	Glu Asp Thr His Leu Gly Glu Glu Val Asp Arg Asn Trp Asn Glu 1040 1045 1050		
50	Val Val Thr Gln Gln Tyr Leu Phe Asp Arg Leu Ala His Glu Ile 1055 1060 1065		
55	Glu Ala Leu Lys Ser Phe Ser Lys Ser Asp Leu Val Ser Trp Phe 1070 1075 1080		
60	Lys Ala His Arg Gly Pro Gly Ser Lys Met Leu Ser Val His Val 1085 1090 1095		
65	Val Gly Tyr Gly Lys Tyr Glu Leu Glu Glu Asp Gly Ala Pro Val 1100 1105 1110		
70	Cys Glu Asp Pro Asn Ser Arg Glu Gly Met Gln Leu Ile Tyr Leu 1115 1120 1125		
75	Pro Pro Ser Pro Leu Leu Ala Glu Ser Thr Thr Pro Ile Thr Asp 1130 1135 1140		

Ile Arg Ala Phe Thr Ala Thr Leu Ser Leu Phe Pro Tyr His Lys
 1145 1150 1155
 5
 Ile Val Lys
 1160
 10
 <210> 11
 <211> 1019
 <212> PRT
 <213> 人
 15
 <400> 11
 Met Arg Tyr Arg Leu Ala Trp Leu Leu His Pro Ala Leu Pro Ser Thr
 1 5 10 15
 20
 Phe Arg Ser Val Leu Gly Ala Arg Leu Pro Pro Pro Glu Arg Leu Cys
 20 25 30
 25
 Gly Phe Gln Lys Lys Thr Tyr Ser Lys Met Asn Asn Pro Ala Ile Lys
 35 40 45
 30 Arg Ile Gly Asn His Ile Thr Lys Ser Pro Glu Asp Lys Arg Glu Tyr
 50 55 60
 Arg Gly Leu Glu Leu Ala Asn Gly Ile Lys Val Leu Leu Met Ser Asp
 65 70 75 80
 35
 Pro Thr Thr Asp Lys Ser Ser Ala Ala Leu Asp Val His Ile Gly Ser
 85 90 95
 40 Leu Ser Asp Pro Pro Asn Ile Ala Gly Leu Ser His Phe Cys Glu His
 100 105 110
 45 Met Leu Phe Leu Gly Thr Lys Lys Tyr Pro Lys Glu Asn Glu Tyr Ser
 115 120 125
 Gln Phe Leu Ser Glu His Ala Gly Ser Ser Asn Ala Phe Thr Ser Gly
 130 135 140
 50
 Glu His Thr Asn Tyr Tyr Phe Asp Val Ser His Glu His Leu Glu Gly
 145 150 155 160
 55
 Ala Leu Asp Arg Phe Ala Gln Phe Phe Leu Cys Pro Leu Phe Asp Glu
 165 170 175

Ser Cys Lys Asp Arg Glu Val Asn Ala Val Asp Ser Glu His Glu Lys
 180 185 190

5 Asn Val Met Asn Asp Ala Trp Arg Leu Phe Gln Leu Glu Lys Ala Thr
 195 200 205

10 Gly Asn Pro Lys His Pro Phe Ser Lys Phe Gly Thr Gly Asn Lys Tyr
 210 215 220

Thr Leu Glu Thr Arg Pro Asn Gln Glu Gly Ile Asp Val Arg Gln Glu
 225 230 235 240

15 Leu Leu Lys Phe His Ser Ala Tyr Tyr Ser Ser Asn Leu Met Ala Val
 245 250 255

20 Cys Val Leu Gly Arg Glu Ser Leu Asp Asp Leu Thr Asn Leu Val Val
 260 265 270

25 Lys Leu Phe Ser Glu Val Glu Asn Lys Asn Val Pro Leu Pro Glu Phe
 275 280 285

30 Pro Glu His Pro Phe Gln Glu Glu His Leu Lys Gln Leu Tyr Lys Ile
 290 295 300
 Val Pro Ile Lys Asp Ile Arg Asn Leu Tyr Val Thr Phe Pro Ile Pro
 305 310 315 320

35 Asp Leu Gln Lys Tyr Tyr Lys Ser Asn Pro Gly His Tyr Leu Gly His
 325 330 335

40 Leu Ile Gly His Glu Gly Pro Gly Ser Leu Leu Ser Glu Leu Lys Ser
 340 345 350

Lys Gly Trp Val Asn Thr Leu Val Gly Gly Gln Lys Glu Gly Ala Arg
 355 360 365

45 Gly Phe Met Phe Phe Ile Ile Asn Val Asp Leu Thr Glu Glu Gly Leu
 370 375 380

50 Leu His Val Glu Asp Ile Ile Leu His Met Phe Gln Tyr Ile Gln Lys
 385 390 395 400

55 Leu Arg Ala Glu Gly Pro Gln Glu Trp Val Phe Gln Glu Cys Lys Asp
 405 410 415

Leu Asn Ala Val Ala Phe Arg Phe Lys Asp Lys Glu Arg Pro Arg Gly

	420	425	430
5	Tyr Thr Ser Lys Ile Ala Gly Ile Leu His Tyr Tyr Pro Leu Glu Glu 435 440 445		
10	Val Leu Thr Ala Glu Tyr Leu Leu Glu Glu Phe Arg Pro Asp Leu Ile 450 455 460		
15	Glu Met Val Leu Asp Lys Leu Arg Pro Glu Asn Val Arg Val Ala Ile 465 470 475 480		
	Val Ser Lys Ser Phe Glu Gly Lys Thr Asp Arg Thr Glu Glu Trp Tyr 485 490 495		
20	Gly Thr Gln Tyr Lys Gln Glu Ala Ile Pro Asp Glu Val Ile Lys Lys 500 505 510		
25	Trp Gln Asn Ala Asp Leu Asn Gly Lys Phe Lys Leu Pro Thr Lys Asn 515 520 525		
30	Glu Phe Ile Pro Thr Asn Phe Glu Ile Leu Pro Leu Glu Lys Glu Ala 530 535 540 Thr Pro Tyr Pro Ala Leu Ile Lys Asp Thr Val Met Ser Lys Leu Trp 545 550 555 560		
35	Phe Lys Gln Asp Asp Lys Lys Lys Lys Pro Lys Ala Cys Leu Asn Phe 565 570 575		
40	Glu Phe Phe Ser Pro Phe Ala Tyr Val Asp Pro Leu His Cys Asn Met 580 585 590		
	Ala Tyr Leu Tyr Leu Glu Leu Leu Lys Asp Ser Leu Asn Glu Tyr Ala 595 600 605		
45	Tyr Ala Ala Glu Leu Ala Gly Leu Ser Tyr Asp Leu Gln Asn Thr Ile 610 615 620		
50	Tyr Gly Met Tyr Leu Ser Val Lys Gly Tyr Asn Asp Lys Gln Pro Ile 625 630 635 640		
55	Leu Leu Lys Lys Ile Ile Glu Lys Met Ala Thr Phe Glu Ile Asp Glu 645 650 655		
	Lys Arg Phe Glu Ile Ile Lys Glu Ala Tyr Met Arg Ser Leu Asn Asn 660 665 670		

Phe Arg Ala Glu Gln Pro His Gln His Ala Met Tyr Tyr Leu Arg Leu
 675 680 685
 5

Leu Met Thr Glu Val Ala Trp Thr Lys Asp Glu Leu Lys Glu Ala Leu
 690 695 700
 10

Asp Asp Val Thr Leu Pro Arg Leu Lys Ala Phe Ile Pro Gln Leu Leu
 705 710 715 720
 15

Ser Arg Leu His Ile Glu Ala Leu Leu His Gly Asn Ile Thr Lys Gln
 725 730 735
 20

Ala Ala Leu Gly Ile Met Gln Met Val Glu Asp Thr Leu Ile Glu His
 740 745 750
 25

Ala His Thr Lys Pro Leu Leu Pro Ser Gln Leu Val Arg Tyr Arg Glu
 755 760 765
 30

Val Gln Leu Pro Asp Arg Gly Trp Phe Val Tyr Gln Gln Arg Asn Glu
 770 775 780
 Val His Asn Asn Cys Gly Ile Glu Ile Tyr Tyr Gln Thr Asp Met Gln
 785 790 795 800
 35

Ser Thr Ser Glu Asn Met Phe Leu Glu Leu Phe Cys Gln Ile Ile Ser
 805 810 815
 40

Glu Pro Cys Phe Asn Thr Leu Arg Thr Lys Glu Gln Leu Gly Tyr Ile
 820 825 830
 45

Val Phe Ser Gly Pro Arg Arg Ala Asn Gly Ile Gln Ser Leu Arg Phe
 835 840 845
 50

Ile Ile Gln Ser Glu Lys Pro Pro His Tyr Leu Glu Ser Arg Val Glu
 850 855 860
 55

Ala Phe Leu Ile Thr Met Glu Lys Ser Ile Glu Asp Met Thr Glu Glu
 865 870 875 880
 Ala Phe Gln Lys His Ile Gln Ala Leu Ala Ile Arg Arg Leu Asp Lys
 885 890 895
 Pro Lys Lys Leu Ser Ala Glu Cys Ala Lys Tyr Trp Gly Glu Ile Ile
 900 905 910

Ser Gln Gln Tyr Asn Phe Asp Arg Asp Asn Thr Glu Val Ala Tyr Leu
 915 920 925

5

Lys Thr Leu Thr Lys Glu Asp Ile Ile Lys Phe Tyr Lys Glu Met Leu
 930 935 940

10

Ala Val Asp Ala Pro Arg Arg His Lys Val Ser Val His Val Leu Ala
 945 950 955 960

15

Arg Glu Met Asp Ser Cys Pro Val Val Gly Glu Phe Pro Cys Gln Asn
 965 970 975

20

Asp Ile Asn Leu Ser Gln Ala Pro Ala Leu Pro Gln Pro Glu Val Ile
 980 985 990

Gln Asn Met Thr Glu Phe Lys Arg Gly Leu Pro Leu Phe Pro Leu Val
 995 1000 1005

25

Lys Pro His Ile Asn Phe Met Ala Ala Lys Leu
 1010 1015

30

<210> 12
 <211> 1265
 <212> PRT
 <213> 拟南芥 (*Arabidopsis thaliana*)

35

Met Ala Ser Ser Ser Ser Ser Ile Phe Thr Gly Val Lys Phe Ser Pro
 1 5 10 15

40

Ile Leu Ala Pro Phe Asn Ser Gly Asp Ser Arg Arg Ser Arg Tyr Leu
 20 25 30

45

Lys Asp Ser Arg Asn Lys Val Arg Phe Asn Pro Ser Ser Pro Arg Leu
 35 40 45

50

Thr Pro His Arg Val Arg Val Glu Ala Pro Ser Leu Ile Pro Tyr Asn
 50 55 60

Gly Leu Trp Ala Ala Gln Pro Asn Ser His Lys Gly Arg Leu Lys Arg
 65 70 75 80

55

Asn Ile Val Ser Gly Lys Glu Ala Thr Gly Ile Ser Leu Ser Gln Gly
 85 90 95

	Arg	Asn	Phe	Cys	Leu	Thr	Cys	Lys	Arg	Asn	Gln	Ala	Gly	Ile	Arg	Arg	
				100					105					110			
5	Ala	Leu	Pro	Ser	Ala	Phe	Val	Asp	Arg	Thr	Ala	Phe	Ser	Leu	Ser	Arg	
			115					120					125				
10	Ser	Ser	Leu	Thr	Ser	Ser	Leu	Arg	Lys	His	Ser	Gln	Ile	Val	Asn	Ala	
		130					135					140					
15	Thr	Leu	Gly	Pro	Asp	Glu	Pro	His	Ala	Ala	Gly	Thr	Ala	Trp	Pro	Asp	
	145					150					155					160	
	Gly	Ile	Val	Ala	Glu	Arg	Gln	Asp	Leu	Asp	Leu	Leu	Pro	Pro	Glu	Ile	
					165					170					175		
20	Asp	Ser	Ala	Glu	Leu	Glu	Ala	Phe	Leu	Gly	Cys	Glu	Leu	Pro	Ser	His	
				180					185					190			
25	Pro	Lys	Leu	His	Arg	Gly	Gln	Leu	Lys	Asn	Gly	Leu	Arg	Tyr	Leu	Ile	
			195					200					205				
30	Leu	Pro	Asn	Lys	Val	Pro	Pro	Asn	Arg	Phe	Glu	Ala	His	Met	Glu	Val	
		210						215				220					
	His	Val	Gly	Ser	Ile	Asp	Glu	Glu	Glu	Asp	Glu	Gln	Gly	Ile	Ala	His	
	225					230					235					240	
35	Met	Ile	Glu	His	Val	Ala	Phe	Leu	Gly	Ser	Lys	Lys	Arg	Glu	Lys	Leu	
					245					250					255		
40	Leu	Gly	Thr	Gly	Ala	Arg	Ser	Asn	Ala	Tyr	Thr	Asp	Phe	His	His	Thr	
				260					265					270			
45	Val	Phe	His	Ile	His	Ser	Pro	Thr	His	Thr	Lys	Asp	Ser	Glu	Asp	Asp	
			275					280					285				
50	Leu	Phe	Pro	Ser	Val	Leu	Asp	Ala	Leu	Asn	Glu	Ile	Ala	Phe	His	Pro	
		290					295					300					
	Lys	Phe	Leu	Ser	Ser	Arg	Val	Glu	Lys	Glu	Arg	Arg	Ala	Ile	Leu	Ser	
	305					310					315					320	
55	Glu	Leu	Gln	Met	Met	Asn	Thr	Ile	Glu	Tyr	Arg	Val	Asp	Cys	Gln	Leu	
				325						330					335		

Leu Gln His Leu His Ser Glu Asn Lys Leu Gly Arg Arg Phe Pro Ile
 340 345 350

5 Gly Leu Glu Glu Gln Ile Lys Lys Trp Asp Val Asp Lys Ile Arg Lys
 355 360 365

10 Phe His Glu Arg Trp Tyr Phe Pro Ala Asn Ala Thr Leu Tyr Ile Val
 370 375 380

15 Gly Asp Ile Asp Asn Ile Pro Arg Ile Val His Asn Ile Glu Ala Val
 385 390 395 400

Phe Gly Lys Asn Gly Leu Asp Asn Glu Ser Thr Pro Ser Ser Pro Ser
 405 410 415

20 Pro Gly Ala Phe Gly Ala Met Ala Asn Phe Leu Val Pro Lys Leu Pro
 420 425 430

25 Ala Gly Leu Gly Gly Thr Phe Ser Asn Glu Lys Thr Asn Thr Ala Asp
 435 440 445
 Gln Ser Lys Met Ile Lys Arg Glu Arg His Ala Ile Arg Pro Pro Val
 450 455 460

30 Glu His Asn Trp Ser Leu Pro Gly Thr Ser Val Asp Leu Lys Pro Pro
 465 470 475 480

35 Gln Ile Phe Lys His Glu Leu Leu Gln Asn Phe Ala Ile Asn Met Phe
 485 490 495

40 Cys Lys Ile Pro Val Ser Lys Val Gln Thr Phe Gly Asp Leu Arg Asn
 500 505 510

Val Leu Met Lys Arg Ile Phe Leu Ser Ala Leu His Phe Arg Ile Asn
 515 520 525

45 Thr Arg Tyr Lys Ser Ser Asn Pro Pro Phe Thr Ser Val Glu Leu Asp
 530 535 540

50 His Ser Asp Ser Gly Arg Glu Gly Cys Thr Val Thr Thr Leu Thr Val
 545 550 555 560

55 Thr Ala Glu Pro Gln Asn Trp Gln Asn Ala Val Lys Val Ala Val Gln
 565 570 575

Glu Val Arg Arg Leu Lys Glu Phe Gly Val Thr Arg Gly Glu Leu Thr

	580	585	590
5	Arg Tyr Met Asp Ala Leu Leu Lys Asp Ser Glu His Leu Ala Ala Met 595 600 605		
10	Ile Asp Asn Val Ser Ser Val Asp Asn Leu Asp Phe Ile Met Glu Ser 610 615 620		
15	Asp Ala Leu Ser His Thr Val Met Asp Gln Thr Gln Gly His Glu Thr 625 630 635 640		
20	Leu Val Ala Val Ala Gly Thr Val Thr Leu Glu Glu Val Asn Thr Val 645 650 655		
25	Gly Ala Lys Val Leu Glu Phe Ile Ser Asp Phe Gly Arg Pro Thr Ala 660 665 670		
30	Leu Leu Pro Ala Ala Ile Val Ala Cys Val Pro Thr Lys Val His Val 675 680 685		
35	Asp Gly Val Gly Glu Ser Asp Phe Asn Ile Ser Pro Asp Glu Ile Ile 690 695 700		
40	Glu Ser Val Lys Ser Gly Leu Leu Ala Pro Ile Glu Ala Glu Pro Glu 705 710 715 720		
45	Leu Glu Val Pro Lys Glu Leu Ile Ser Gln Ser Gln Leu Lys Glu Leu 725 730 735		
50	Thr Leu Gln Arg Asn Pro Cys Phe Val Pro Ile Pro Gly Ser Gly Leu 740 745 750		
55	Thr Lys Leu His Asp Lys Glu Thr Gly Ile Thr Gln Leu Arg Leu Ser 755 760 765		
60	Asn Gly Ile Ala Val Asn Tyr Lys Lys Ser Thr Thr Glu Ser Arg Ala 770 775 780		
65	Gly Val Met Arg Leu Ile Val Gly Gly Gly Arg Ala Ala Glu Thr Ser 785 790 795 800		
70	Asp Ser Lys Gly Ala Val Val Val Gly Val Arg Thr Leu Ser Glu Gly 805 810 815		
75	Gly Arg Val Gly Asp Phe Ser Arg Glu Gln Val Glu Leu Phe Cys Val 820 825 830		

Asn His Leu Ile Asn Cys Ser Leu Glu Ser Thr Glu Glu Phe Ile Ala
 835 840 845
 5

Met Glu Phe Arg Phe Thr Leu Arg Asp Asn Gly Met Gln Ala Ala Phe
 850 855 860
 10

Gln Leu Leu His Met Val Leu Glu Arg Ser Val Trp Leu Glu Asp Ala
 865 870 875 880
 15

Phe Asp Arg Ala Arg Gln Leu Tyr Leu Ser Tyr Phe Arg Ser Ile Pro
 885 890 895
 20

Lys Ser Leu Glu Arg Ala Thr Ala His Lys Leu Met Ile Ala Met Leu
 900 905 910
 25

Asn Gly Asp Glu Arg Phe Val Glu Pro Thr Pro Lys Ser Leu Gln Ser
 915 920 925
 Leu Asn Leu Glu Ser Val Lys Asp Ala Val Met Ser His Phe Val Gly
 930 935 940
 30

Asp Asn Met Glu Val Ser Ile Val Gly Asp Phe Ser Glu Glu Glu Ile
 945 950 955 960
 35

Glu Arg Cys Ile Leu Asp Tyr Leu Gly Thr Val Lys Ala Ser His Asp
 965 970 975
 Ser Ala Lys Pro Pro Gly Ser Glu Pro Ile Leu Phe Arg Gln Pro Thr
 980 985 990
 40

Ala Gly Leu Gln Phe Gln Gln Val Phe Leu Lys Asp Thr Asp Glu Arg
 995 1000 1005
 45

Ala Cys Ala Tyr Ile Ala Gly Pro Ala Pro Asn Arg Trp Gly Phe
 1010 1015 1020
 50

Thr Val Asp Gly Asp Asp Leu Phe Gln Ser Val Ser Lys Leu Pro
 1025 1030 1035
 55

Val Ala His Asp Gly Leu Leu Lys Ser Glu Glu Gln Leu Leu Glu
 1040 1045 1050
 Gly Gly Asp Arg Glu Leu Gln Lys Lys Leu Arg Ala His Pro Leu
 1055 1060 1065

[illegible]

```

<400> 13

  Met Arg Pro Asp Asp Lys Tyr His Glu Lys Gln Ala Gln Val Glu Ala
5  1              5              10              15

  Thr Lys Leu Lys Gln Lys Val Glu Ala Leu Ser Pro Gly Asp Arg Gln
10              20              25              30

  Gln Ile Tyr Glu Lys Gly Leu Glu Leu Arg Ser Gln Gln Ser Lys Pro
              35              40              45

15
  Gln Asp Ala Ser Cys Leu Pro Ala Leu Lys Val Ser Asp Ile Glu Pro
              50              55              60

20 Thr Ile Pro Val Thr Glu Leu Asp Val Val Leu Thr Ala Gly Asp Ile
   65              70              75              80

  Pro Val Gln Tyr Cys Ala Gln Pro Thr Asn Gly Met Val Tyr Phe Arg
              85              90              95

25
  Ala Phe Ser Ser Leu Asn Thr Leu Pro Glu Glu Leu Arg Pro Tyr Val
              100              105              110

30
  Pro Leu Phe Cys Ser Val Leu Thr Lys Leu Gly Cys Gly Leu Leu Asp
              115              120              125

35 Tyr Arg Glu Gln Ala Gln Gln Ile Glu Leu Lys Thr Gly Gly Met Ser
   130              135              140

  Ala Ser Pro His Val Leu Pro Asp Asp Ser His Met Asp Thr Tyr Glu
40  145              150              155              160

  Gln Gly Val Leu Phe Ser Ser Leu Cys Leu Asp Arg Asn Leu Pro Asp
              165              170              175

45
  Met Met Gln Leu Trp Ser Glu Ile Phe Asn Asn Pro Cys Phe Glu Glu
              180              185              190

50
  Glu Glu His Phe Lys Val Leu Val Lys Met Thr Ala Gln Glu Leu Ala
              195              200              205

55 Asn Gly Ile Pro Asp Ser Gly His Leu Tyr Ala Ser Ile Arg Ala Gly
   210              215              220

  Arg Thr Leu Thr Pro Ala Gly Asp Leu Gln Glu Thr Phe Ser Gly Met

```

	225		230		235		240
5	Asp Gln Val Arg Leu Met Lys Arg Ile Ala Glu Met Thr Asp Ile Lys		245		250		255
10	Pro Ile Leu Arg Lys Leu Pro Arg Ile Lys Lys His Leu Leu Asn Gly		260		265		270
15	Asp Asn Met Arg Cys Ser Val Asn Ala Thr Pro Gln Gln Met Pro Gln		275		280		285
	Thr Glu Lys Ala Val Glu Asp Phe Leu Arg Ser Ile Gly Arg Ser Lys		290		295		300
20	Lys Glu Arg Arg Pro Val Arg Pro His Thr Val Glu Lys Pro Val Pro		305		310		315
	Ser Ser Ser Gly Gly Asp Ala His Val Pro His Gly Ser Gln Val Ile		325		330		335
25	Arg Lys Leu Val Met Glu Pro Thr Phe Lys Pro Trp Gln Met Lys Thr		340		345		350
30	His Phe Leu Met Pro Phe Pro Val Asn Tyr Val Gly Glu Cys Ile Arg		355		360		365
35	Thr Val Pro Tyr Thr Asp Pro Asp His Ala Ser Leu Lys Ile Leu Ala		370		375		380
40	Arg Leu Met Thr Ala Lys Phe Leu His Thr Glu Ile Arg Glu Lys Gly		385		390		395
	Gly Ala Tyr Gly Gly Gly Ala Lys Leu Ser His Asn Gly Ile Phe Thr		405		410		415
45	Leu Tyr Ser Tyr Arg Asp Pro Asn Thr Ile Glu Thr Leu Gln Ser Phe		420		425		430
50	Gly Lys Ala Val Asp Trp Ala Lys Ser Gly Lys Phe Thr Gln Gln Asp		435		440		445
55	Ile Asp Glu Ala Lys Leu Ser Val Phe Ser Thr Val Asp Ala Pro Val		450		455		460
	Ala Pro Ser Asp Lys Gly Met Asp His Phe Leu Tyr Gly Leu Ser Asp		465		470		475
							480

Glu Met Lys Gln Ala His Arg Glu Gln Leu Phe Ala Val Ser His Asp
 485 490 495
 5

Lys Leu Leu Ala Val Ser Asp Arg Tyr Leu Gly Thr Gly Lys Ser Thr
 500 505 510
 10

His Gly Leu Ala Ile Leu Gly Pro Glu Asn Pro Lys Ile Ala Lys Asp
 515 520 525
 15 Pro Ser Trp Ile Ile Arg
 530

20 <210> 14
 <211> 409
 <212> PRT
 <213> 枯草芽孢杆菌 (Bacillus subtilis)
 25 <400> 14

Met Ile Lys Arg Tyr Thr Cys Pro Asn Gly Val Arg Ile Val Leu Glu
 1 5 10 15
 30

Asn Asn Pro Thr Val Arg Ser Val Ala Ile Gly Val Trp Ile Gly Thr
 20 25 30
 35 Gly Ser Arg His Glu Thr Pro Glu Ile Asn Gly Ile Ser His Phe Leu
 35 40 45

Glu His Met Phe Phe Lys Gly Thr Ser Thr Lys Ser Ala Arg Glu Ile
 40 50 55 60
 45

Ala Glu Ser Phe Asp Arg Ile Gly Gly Gln Val Asn Ala Phe Thr Ser
 65 70 75 80
 45

Lys Glu Tyr Thr Cys Tyr Tyr Ala Lys Val Leu Asp Glu His Ala Asn
 85 90 95
 50

Tyr Ala Leu Asp Val Leu Ala Asp Met Phe Phe His Ser Thr Phe Asp
 100 105 110
 55 Glu Asn Glu Leu Lys Lys Glu Lys Asn Val Val Tyr Glu Glu Ile Lys
 115 120 125

Met Tyr Glu Asp Ala Pro Asp Asp Ile Val His Asp Leu Leu Ser Lys

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

	370	375	380
5	Asn Gly Leu Ala Arg Gln Leu Phe Thr Glu Asp Tyr Ala Leu Ala Leu 385 390 395 400		
10	Ile Ser Pro Ser Gly Asn Met Pro Ser 405		
	<210> 15 <211> 438 <212> PRT		
15	<213> 结核分枝杆菌 (Mycobacterium tuberculosis)		
	<400> 15 Met Pro Arg Arg Ser Pro Ala Asp Pro Ala Ala Ala Leu Ala Pro Arg 1 5 10 15		
20	Arg Thr Thr Leu Pro Gly Gly Leu Arg Val Val Thr Glu Phe Leu Pro 20 25 30		
25	Ala Val His Ser Ala Ser Val Gly Val Trp Val Gly Val Gly Ser Arg 35 40 45		
30	Asp Glu Gly Ala Thr Val Ala Gly Ala Ala His Phe Leu Glu His Leu 50 55 60		
35	Leu Phe Lys Ser Thr Pro Thr Arg Ser Ala Val Asp Ile Ala Gln Ala 65 70 75 80		
40	Met Asp Ala Val Gly Gly Glu Leu Asn Ala Phe Thr Ala Lys Glu His 85 90 95		
45	Thr Cys Tyr Tyr Ala His Val Leu Gly Ser Asp Leu Pro Leu Ala Val 100 105 110		
	Asp Leu Val Ala Asp Val Val Leu Asn Gly Arg Cys Ala Ala Asp Asp 115 120 125		
50	Val Glu Val Glu Arg Asp Val Val Leu Glu Glu Ile Ala Met Arg Asp 130 135 140		
55	Asp Asp Pro Glu Asp Ala Leu Ala Asp Met Phe Leu Ala Ala Leu Phe 145 150 155 160		
	Gly Asp His Pro Val Gly Arg Pro Val Ile Gly Ser Ala Gln Ser Val 165 170 175		

5 Ser Val Met Thr Arg Ala Gln Leu Gln Ser Phe His Leu Arg Arg Tyr
 180 185 190

10 Thr Pro Glu Arg Met Val Val Ala Ala Ala Gly Asn Val Asp His Asp
 195 200 205

15 Gly Leu Val Ala Leu Val Arg Glu His Phe Gly Ser Arg Leu Val Arg
 210 215 220

20 Gly Arg Arg Pro Val Ala Pro Arg Lys Gly Thr Gly Arg Val Asn Gly
 225 230 235 240
 Ser Pro Arg Leu Thr Leu Val Ser Arg Asp Ala Glu Gln Thr His Val
 245 255

25 Ser Leu Gly Ile Arg Thr Pro Gly Arg Gly Trp Glu His Arg Trp Ala
 260 265 270

30 Leu Ser Val Leu His Thr Ala Leu Gly Gly Gly Leu Ser Ser Arg Leu
 275 280 285

35 Phe Gln Glu Val Arg Glu Thr Arg Gly Leu Ala Tyr Ser Val Tyr Ser
 290 295 300

40 Ala Leu Asp Leu Phe Ala Asp Ser Gly Ala Leu Ser Val Tyr Ala Ala
 305 310 315 320

45 Cys Leu Pro Glu Arg Phe Ala Asp Val Met Arg Val Thr Ala Asp Val
 325 330 335

50 Leu Glu Ser Val Ala Arg Asp Gly Ile Thr Glu Ala Glu Cys Gly Ile
 340 345 350

55 Ala Lys Gly Ser Leu Arg Gly Gly Leu Val Leu Gly Leu Glu Asp Ser
 355 360 365

60 Ser Ser Arg Met Ser Arg Leu Gly Arg Ser Glu Leu Asn Tyr Gly Lys
 370 375 380

65 His Arg Ser Ile Glu His Thr Leu Arg Gln Ile Glu Gln Val Thr Val
 385 390 395 400

70 Glu Glu Val Asn Ala Val Ala Arg His Leu Leu Ser Arg Arg Tyr Gly
 405 410 415

	Ala Ala Val Leu Gly Pro His Gly Ser Lys Arg Ser Leu Pro Gln Gln	
	420 425 430	
5	Leu Arg Ala Met Val Gly	
	435	
10	<210> 16	
	<211> 34	
	<212> DNA	
	<213> 人工序列	
15	<220>	
	<223> 寡核苷酸	
	<400> 16	
20	aatagaagct tgtcgactga tctatccaaa actg	34
	<210> 17	
	<211> 66	
25	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 寡核苷酸	
30	<400> 17	
	aaaagagctc ggccagatct tctagaggat ccaagaattc tgttttatat ttgttgtaaa	60
	aagtag	66
35	<210> 18	
	<211> 37	
	<212> DNA	
40	<213> 人工序列	
	<220>	
	<223> 寡核苷酸	
45	<400> 18	
	ttttgaattc caagatctcc catgtctcta ctggtgg	37
50	<210> 19	
	<211> 41	
	<212> DNA	
	<213> 人工序列	
55	<220>	
	<223> 寡核苷酸	

	<400> 19	
	ccccgagctc gtcgaccctt ctcgaaagct ttaacgaacg c	41
5	<210> 20	
	<211> 43	
	<212> DNA	
	<213> 人工序列	
10	<220>	
	<223> 寡核苷酸	
	<400> 20	
	ttttgaattc aaagaatgag atttccttca atttttactg cag	43
15	<210> 21	
	<211> 37	
	<212> DNA	
	<213> 人工序列	
20	<220>	
	<223> 寡核苷酸	
	<400> 21	
25	tttttctaga ctaggagggg tactcatact cctcggc	37
	<210> 22	
	<211> 33	
	<212> DNA	
30	<213> 人工序列	
	<220>	
	<223> 寡核苷酸	
35	<400> 22	
	cgaatgtcca tcgttgcgaa cctgcagaac ctg	33
	<210> 23	
	<211> 33	
40	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 寡核苷酸	
45	<400> 23	
	caggttctgc aggttcctaa cgatggacat tcg	33
	<210> 24	
50	<211> 33	
	<212> DNA	
	<213> 人工序列	
	<220>	
55	<223> 寡核苷酸	
	<400> 24	
	cgaatgtcca tcgttaggaa cctgcagaac ctg	33

	<210>	25	
	<211>	33	
5	<212>	DNA	
	<213>	人工序列	
	<220>		
10	<223>	寡核苷酸	
	<400>	25	
		caggttctgc aggttcctaa cgatggacat tcg	33
	<210>	26	
15	<211>	18	
	<212>	DNA	
	<213>	人工序列	
	<220>		
20	<223>	寡核苷酸	
	<400>	26	
		tcgcagagaa cggatggc	18
25			
	<210>	27	
	<211>	36	
	<212>	DNA	
	<213>	人工序列	
30			
	<220>		
	<223>	寡核苷酸	
	<400>	27	
35		ttttgggcc ttcatggtga tacggtatct cttggc	36
	<210>	28	
	<211>	37	
40	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	寡核苷酸	
45			
	<400>	28	
		ttttctcgag aaggtggaac atactgccct gggatgg	37
50	<210>	29	
	<211>	38	
	<212>	DNA	
	<213>	人工序列	
55	<220>		
	<223>	寡核苷酸	
	<400>	29	

	ttttgagctc gtttaggaaa cgtccttggc ggagatgc	38
5	<210> 30 <211> 40 <212> DNA <213> 人工序列 <220> <223> 寡核苷酸	
10	<400> 30 tttttctaga cactgcgaat ccatggtata aacccaaaacc	40
15	<210> 31 <211> 24 <212> DNA <213> 人工序列 <220> <223> 寡核苷酸	
20	<400> 31 gtcgttggttc atggacatac ctcc	24
25	<210> 32 <211> 26 <212> DNA <213> 人工序列 <220> <223> 寡核苷酸	
30	<400> 32 tacaaatggt cttctgccat ttctgg	26
40	<210> 33 <211> 32 <212> DNA <213> 人工序列 <220> <223> 寡核苷酸	
45	<400> 33 ggttcatatg cgccggagct cctcgacagc ag	32
50	<210> 34 <211> 37 <212> DNA <213> 人工序列 <220> <223> 寡核苷酸	
55	<400> 34	

	ggttcctagg atccgcaagt ttgattccat tgcggtg	37
	<210> 35	
5	<211> 70	
	<212> DNA	
	<213> 人工序列	
	<220>	
10	<223> 寡核苷酸	
	<400> 35	
	ttaaagagta ccttggctat agaataccgt agagataaag acctgaatag agattgtact	60
	gagagtgcac	70
15		
	<210> 36	
	<211> 70	
	<212> DNA	
20	<213> 人工序列	
	<220>	
	<223> 寡核苷酸	
25	<400> 36	
	aggtattata actatatttttc tgtatatttttt atatatttttt atttgccaaag ctgtgcggta	60
	tttcacaccg	70
30		
	<210> 37	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
35	<220>	
	<223> 寡核苷酸	
	<400> 37	
40	ctttggttaa agagtacctt ggc	23
	<210> 38	
	<211> 23	
45	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 寡核苷酸	
50	<400> 38	
	tactacgaaa agcgtgtgcg agg	23
55	<210> 39	
	<211> 71	
	<212> DNA	
	<213> 人工序列	

```

    <220>
    <223> 寡核苷酸

5  <400>  39
   tagaaggcta ctcaaaagaa taaagttact ataaaatata ctgcggtata tagattgtac   60
   tgagagtgca c                                                                71
10

    <210>  40
    <211>  70
15 <212>  DNA
    <213>  人工序列

    <220>
    <223> 寡核苷酸
20
    <400>  40
   gatcggcaag aaactttgaa gcagtatatt tacaggatta aattatatat ctgtgcggtta   60
   tttcacaccg                                                                70
25

    <210>  41
    <211>  22
    <212>  DNA
30 <213>  人工序列

    <220>
    <223> 寡核苷酸

35 <400>  41
   cggaggggct ctatgataaa gg                                                                22

    <210>  42
40 <211>  23
    <212>  DNA
    <213>  人工序列

    <220>
45 <223> 寡核苷酸

    <400>  42
   gagtaactag ggcttctctt ccc                                                                23
50

    <210>  43
    <211>  85
    <212>  PRT
    <213>  人
55
    <400>  43

   Ser Gly Leu Gln Arg Ala Glu Glu Ala Pro Arg Arg Gln Leu Arg Val

```

```

      1              5              10              15

5   Ser Gln Arg Thr Asp Gly Glu Ser Arg Ala His Leu Gly Ala Leu Leu
      20              25              30

10  Ala Arg Tyr Ile Gln Gln Ala Arg Lys Ala Pro Ser Gly Arg Met Ser
      35              40              45

15  Ile Val Lys Asn Leu Gln Asn Leu Asp Pro Ser His Arg Ile Ser Asp
      50              55              60

15  Arg Asp Tyr Met Gly Trp Met Asp Phe Gly Arg Arg Ser Ala Glu Glu
      65              70              75              80

20  Tyr Glu Tyr Pro Ser
      85

25  <210> 44
    <211> 22
    <212> PRT
    <213> 人

30  <400> 44

    Gln Leu Arg Val Ser Gln Arg Thr Asp Gly Glu Ser Arg Ala His Leu
    1              5              10              15

35  Gly Ala Leu Leu Ala Arg
      20

40  <210> 45
    <211> 19
    <212> PRT
    <213> 人

45  <400> 45

    Val Ser Gln Arg Thr Asp Gly Glu Ser Arg Ala His Leu Gly Ala Leu
    1              5              10              15

50  Leu Ala Arg

55  <210> 46
    <211> 51
    <212> PRT
    <213> 人

```

```

<400> 46

5 Tyr Ile Gln Gln Ala Arg Lys Ala Pro Ser Gly Arg Met Ser Ile Val
  1             5             10             15

Lys Asn Leu Gln Asn Leu Asp Pro Ser His Arg Ile Ser Asp Arg Asp
10      20      25      30
  Tyr Met Gly Trp Met Asp Phe Gly Arg Arg Ser Ala Glu Glu Tyr Glu
    35      40      45

15 Tyr Pro Ser
   50

<210> 47
<211> 17
20 <212> PRT
   <213> 人

<400> 47

25 Tyr Ile Gln Gln Ala Arg Lys Ala Pro Ser Gly Arg Met Ser Ile Val
   1             5             10             15

30 Lys

<210> 48
<211> 16
35 <212> PRT
   <213> 人

<400> 48

40 Tyr Ile Gln Gln Ala Arg Lys Ala Pro Ser Gly Arg Met Ser Ile Val
   1             5             10             15

<210> 49
<211> 13
45 <212> PRT
   <213> 人

<400> 49

50 Asn Leu Gln Asn Leu Asp Pro Ser His Arg Ile Ser Asp
   1             5             10

55 <210> 50
   <211> 23
   <212> PRT
   <213> 人

```

<400> 50
 Asn Leu Gln Asn Leu Asp Pro Ser His Arg Ile Ser Asp Arg Asp Tyr
 5 1 5 10 15
 Met Gly Trp Met Asp Phe Gly
 20
 10 <210> 51
 <211> 34
 <212> PRT
 <213> 人
 15 <400> 51
 Asn Leu Gln Asn Leu Asp Pro Ser His Arg Ile Ser Asp Arg Asp Tyr
 1 5 10 15
 20 Met Gly Trp Met Asp Phe Gly Arg Arg Ser Ala Glu Glu Tyr Glu Tyr
 20 25 30
 25 Pro Ser
 <210> 52
 30 <211> 20
 <212> PRT
 <213> 人
 <400> 52
 35 Asp Tyr Met Gly Trp Met Asp Phe Gly Arg Arg Ser Ala Glu Glu Tyr
 1 5 10 15
 Glu Tyr Pro Ser
 20
 40 <210> 53
 <211> 12
 <212> PRT
 45 <213> 人工序列
 <220>
 <223> Portion of fusion protein
 50 <400> 53
 Lys Arg Glu Ala Glu Ala Ser Gly Leu Gln Arg Ala
 1 5 10
 55 <210> 54
 <211> 9
 <212> PRT
 <213> 人

- <400> 54
- 5 Arg Met Ser Ile Val Lys Asn Leu Gln
1 5
- <210> 55
10 <211> 13
<212> PRT
<213> 人
- <400> 55
15 Asp Arg Asp Tyr Met Gly Trp Met Asp Phe Gly Arg Arg
1 5 10
- 20 <210> 56
<211> 48
<212> DNA
<213> 人工序列
- 25 <220>
<223> 寡核苷酸 MFa1BNP (S)
- <400> 56
30 ggataaaaga gaggctgaag ctcacccgct gggcagcccc ggttcagc 48
- <210> 57
<211> 48
<212> DNA
35 <213> 人工序列
- <220>
<223> MF1aBNP (AS)
- 40 <400> 57
gctgaaccgg ggctgcccag cgggtgagct tcagcctctc ttttatcc 48
- <210> 58
<211> 34
45 <212> DNA
<213> 人工序列
- <220>
<223> 寡核苷酸 BNP5'EcoRI
- 50
- <400> 58
ttttgaattc atggatcccc agacagcacc ttcc 34
- 55 <210> 59
<211> 33
<212> DNA
<213> 人工序列

	<220>		
	<223> 寡核苷酸 BNP3'Xba1		
5	<400> 59		
	tttttctaga ttaatgccgc ctcagcactt tgc		33
	<210> 60		
10	<211> 48		
	<212> DNA		
	<213> 人工序列		
	<220>		
15	<223> 寡核苷酸 MF1BNP (S)		
	<400> 60		
	ggataaaaga gaggctgaag ctcacccgct gggcagcccc gggtcagc		48
20	<210> 61		
	<211> 48		
	<212> DNA		
	<213> 人工序列		
25	<220>		
	<223> 寡核苷酸 MF1BNP (AS)		
	<400> 61		
30	gctgaaccgg ggctgcccag cgggtgagct tcagcctctc ttttatcc		48
	<210> 62		
	<211> 69		
35	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 寡核苷酸 YPS15'GD400		
40	<400> 62		
	aaaaagataa ggtgaacacc aagcatatag tataatatta cctaccacat gattgtactg 60		
	agagtgcac		69
45	<210> 63		
	<211> 70		
	<212> DNA		
	<213> 人工序列		
50	<220>		
	<223> 寡核苷酸 YPS13'GD400		
	<400> 63		
	aactccaact ggcttgaga tgtgaatgtc taaactttgt gcaacgggtt ctgtgcggta 60		
55	tttcacaccg		70
	<210> 64		
	<211> 20		

```

    <212> DNA
    <213> 人工序列

    <220>
5  <223> 寡核苷酸 YPS15'DC

    <400> 64
    tcgtttcact gatgtgtccg
                                                                20

10 <210> 65
    <211> 21
    <212> DNA
    <213> 人工序列

15 <220>
    <223> 寡核苷酸 YPS15'DC

    <400> 65
    gattataggc catatcccag g
                                                                21
20

    <210> 66
    <211> 13
    <212> PRT
    <213> 人工序列

25

    <220>
    <223> Pitrilysin 共有序列

30 <221> 变体
    <222> (1)...(13)
    <223> Xaa = 任一氨基酸

    <221> 变体
35 <222> (1)...(13)
    <223> Xaa = 任一氨基酸

    <221> 变体
    <222> (1)...(13)
40 <223> Xaa = 任一氨基酸

    <221> 未确定
    <222> (0)...(0)

45 <400> 66
    Gly Xaa Xaa His Xaa Xaa Glu His Xaa Xaa Xaa Xaa Gly
      1              5              10

50 <210> 67
    <211> 44
    <212> PRT
    <213> 人工序列

55 <220>
    <223> Pitrilysin 共有序列

```

```

    <221> 变体
    <222> (1)...(44)
    <223> Xaa = 任一氨基酸
5
    <221> 变体
    <222> (1)...(44)
    <223> Xaa = 任一氨基酸

10 <221> 变体
    <222> (1)...(44)
    <223> Xaa = 任一氨基酸

    <221> 变体
15 <222> (34)...(34)
    <223> Xaa = 任一氨基酸或不存在

    <221> 变体
    <222> (35)...(35)
20 <223> Xaa = 任一氨基酸或不存在

    <400> 67
    Gly Xaa Xaa His Xaa Xaa Glu His Xaa Xaa Xaa Xaa Gly Xaa Xaa Xaa
      1              5              10              15
25 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
      20              25              30
    Xaa Xaa Xaa Asn Ala Xaa Thr Xaa Xaa Xaa Xaa Thr
      35              40

30
    <210> 68
    <211> 44
    <212> PRT
    <213> 人工序列
35
    <220>
    <223> Pitrilysin 共有序列

40 <221> 变体
    <222> (1)...(44)
    <223> Xaa = 任一氨基酸

    <221> 变体
45 <222> (1)...(44)
    <223> Xaa = 任一氨基酸

    <221> 变体
    <222> (1)...(44)
50 <223> Xaa = 任一氨基酸

    <221> 变体
    <222> (34)...(34)
    <223> Xaa = 任一氨基酸或不存在
55
    <221> 变体

```

<222> (35)...(35)

<223> Xaa = 任一氨基酸或不存在

<400> 68

```

5 Gly Xaa Xaa His Xaa Xaa Glu His Xaa Xaa Xaa Xaa Gly Xaa Xaa Lys
  1          5          10          15
  Tyr Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa Xaa Xaa
    20          25          30
10 Xaa Xaa Xaa Asn Ala Xaa Thr Xaa Xaa Xaa Xaa Thr
    35          40

```

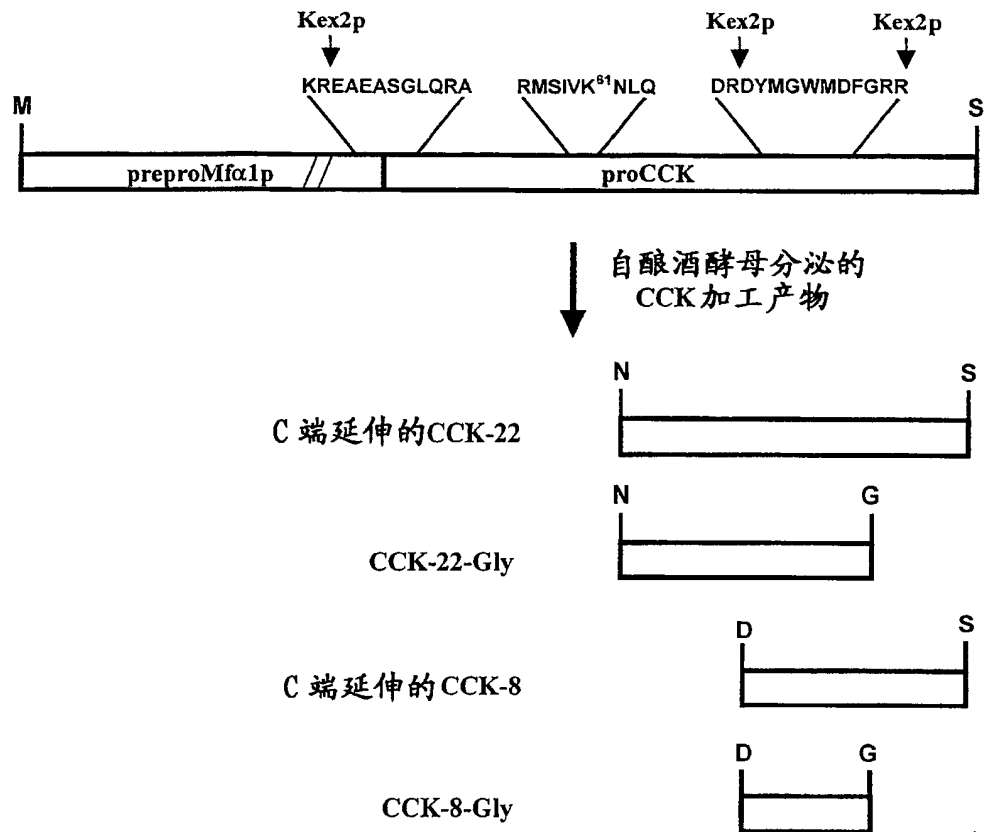


图 1

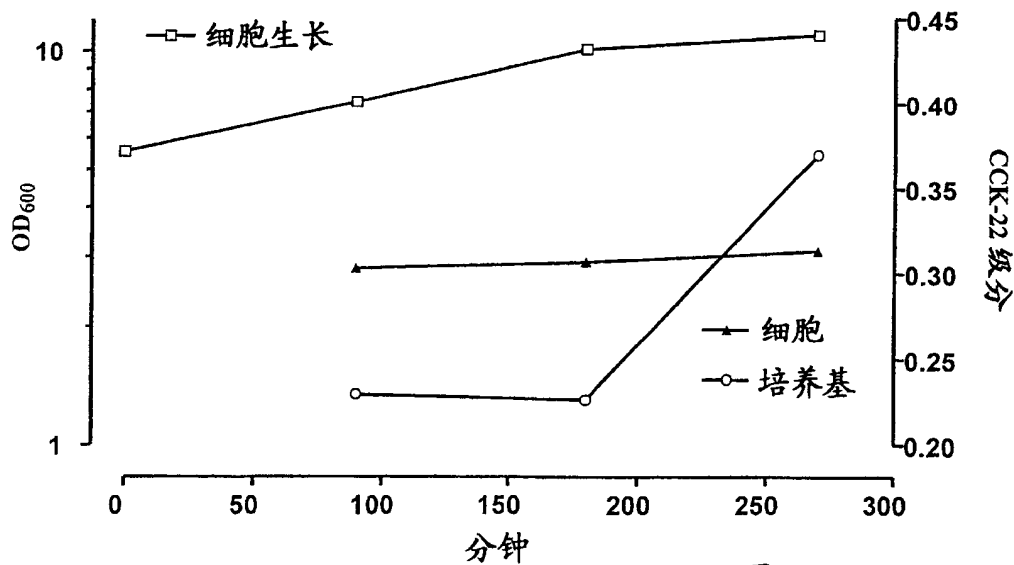


图 2

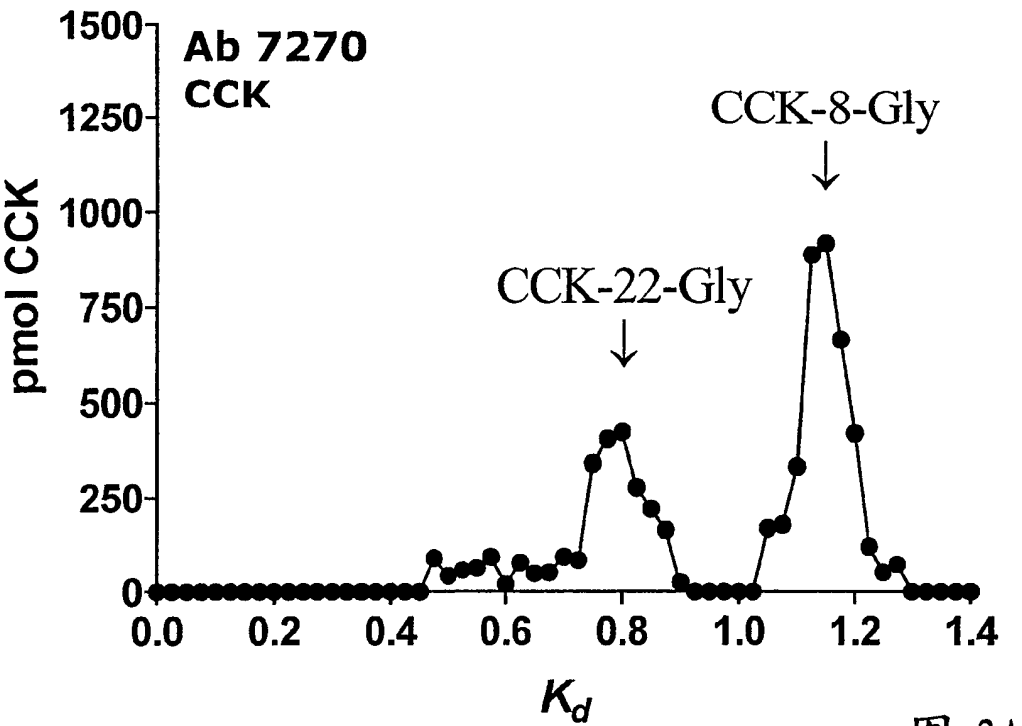


图 3A

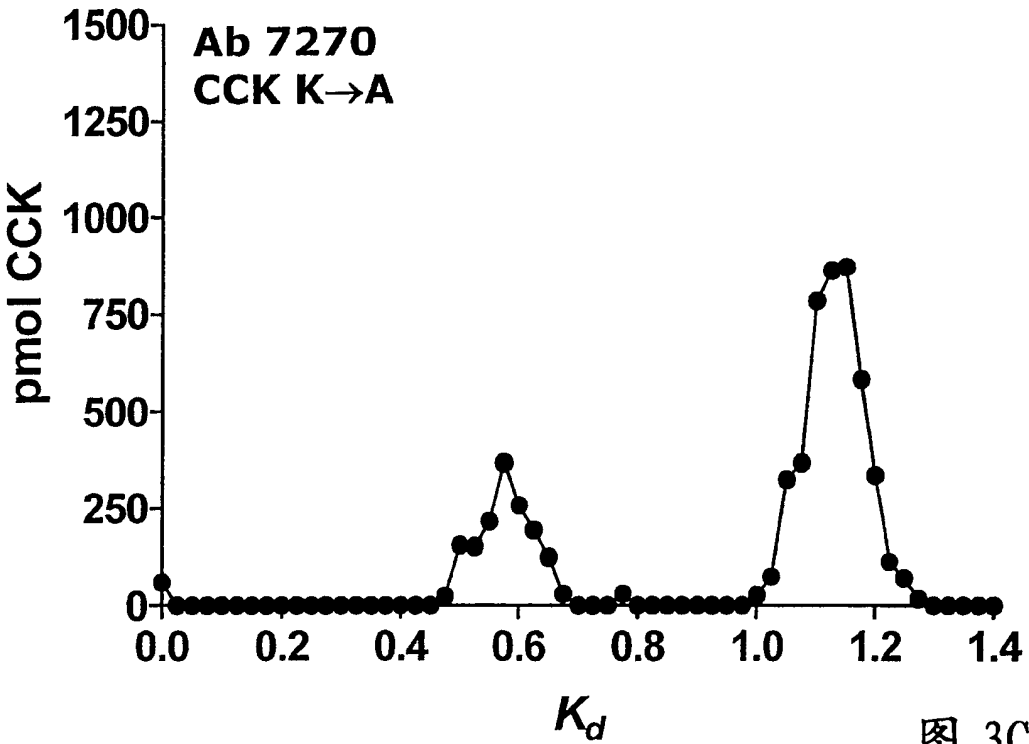


图 3C

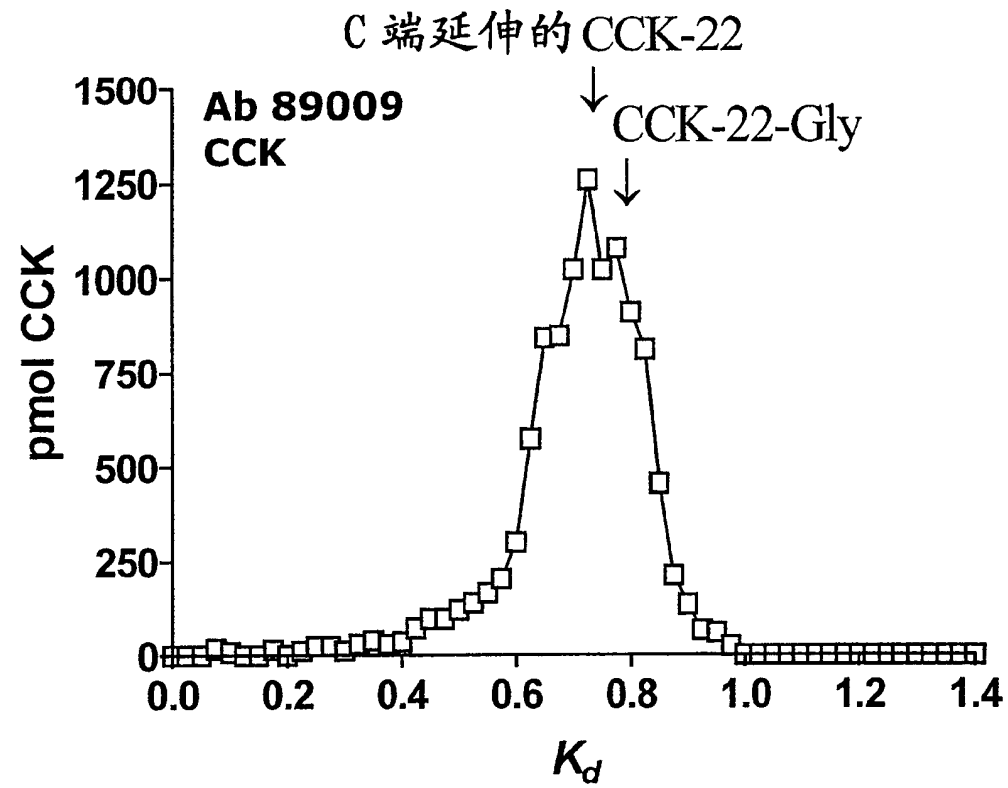


图 3B

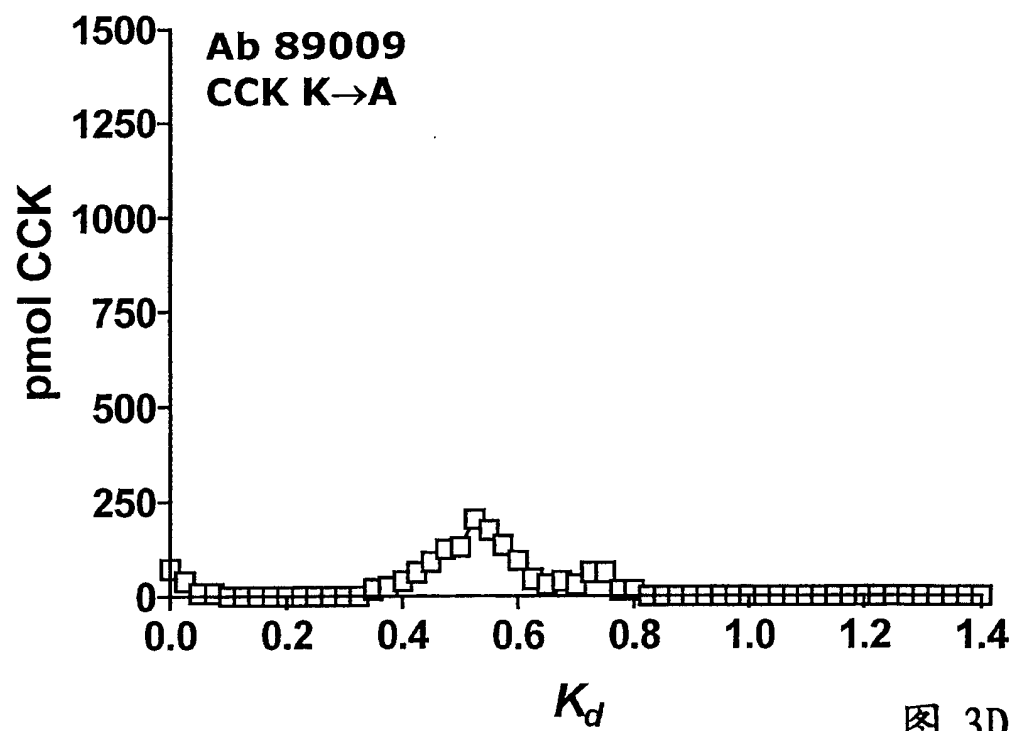


图 3D

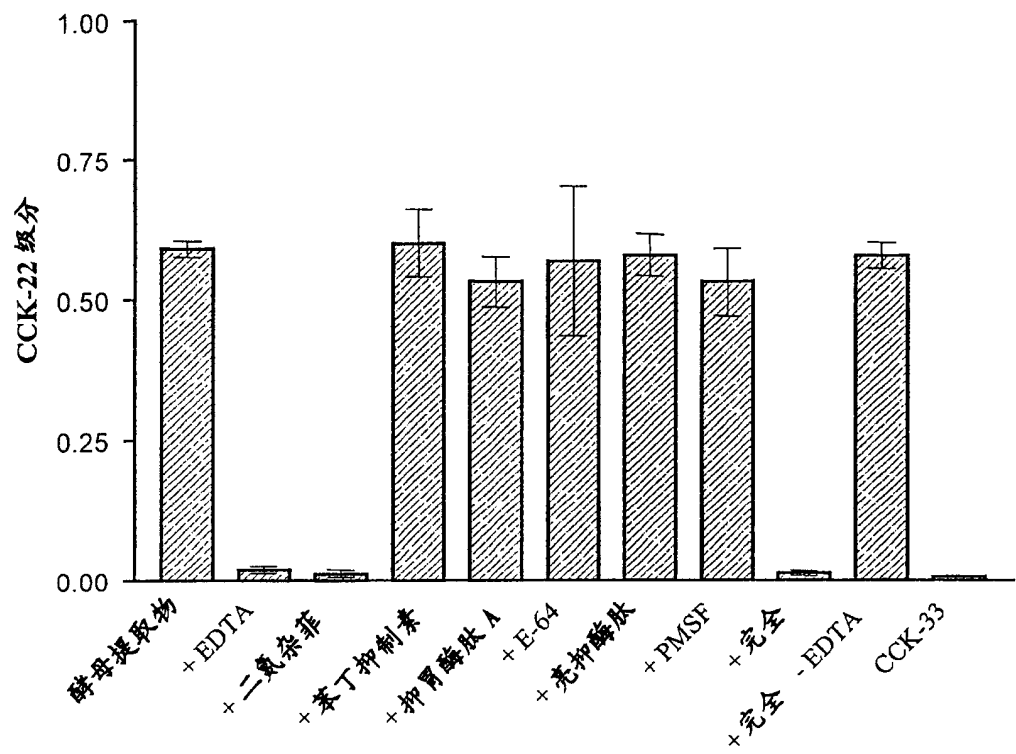


图 4A

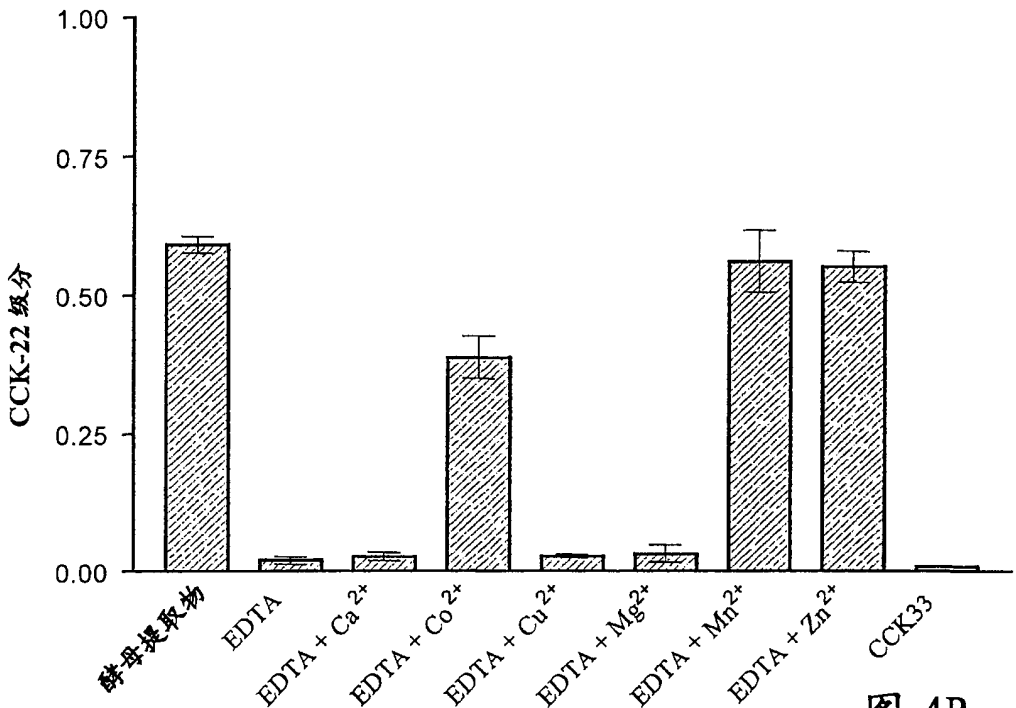


图 4B

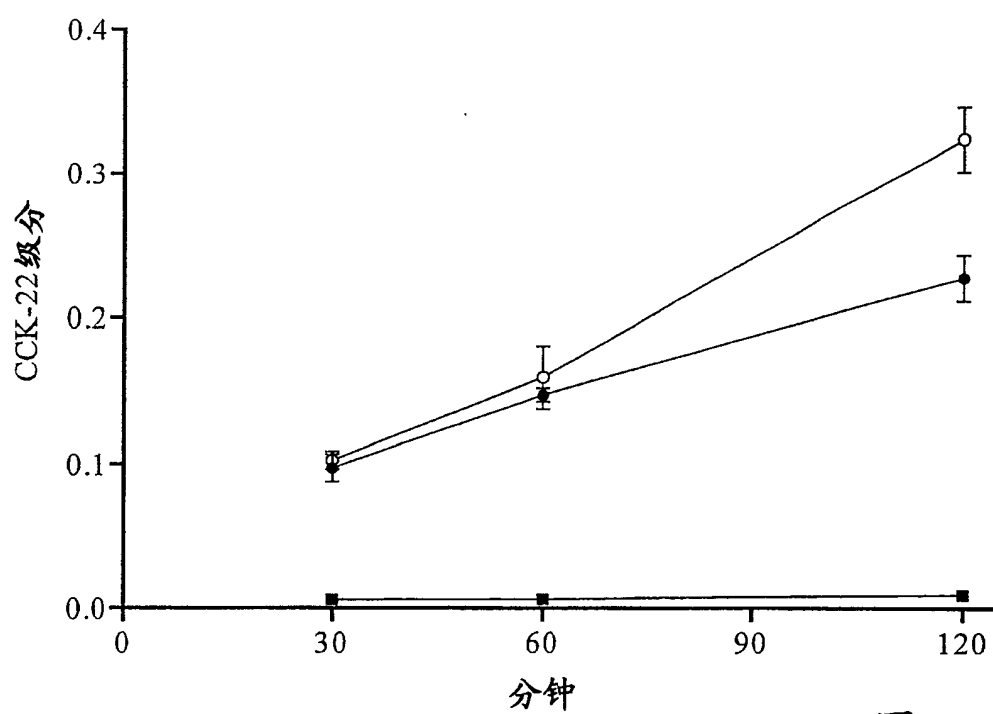


图 5

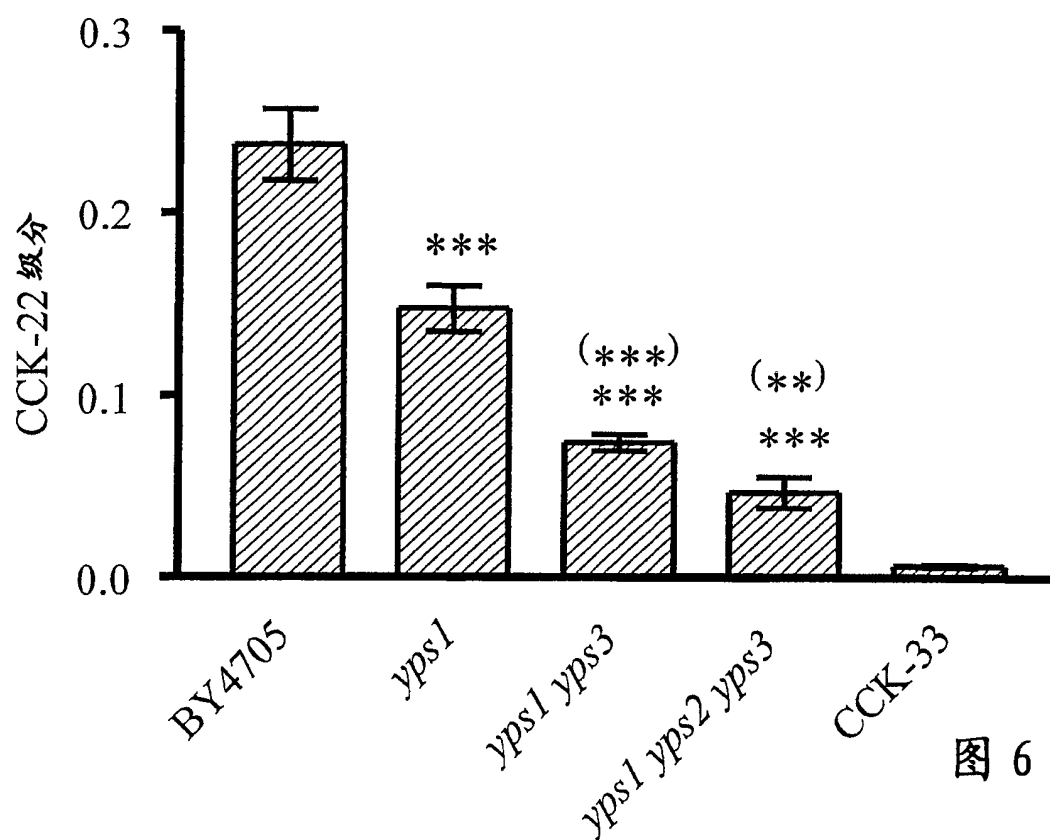


图 6

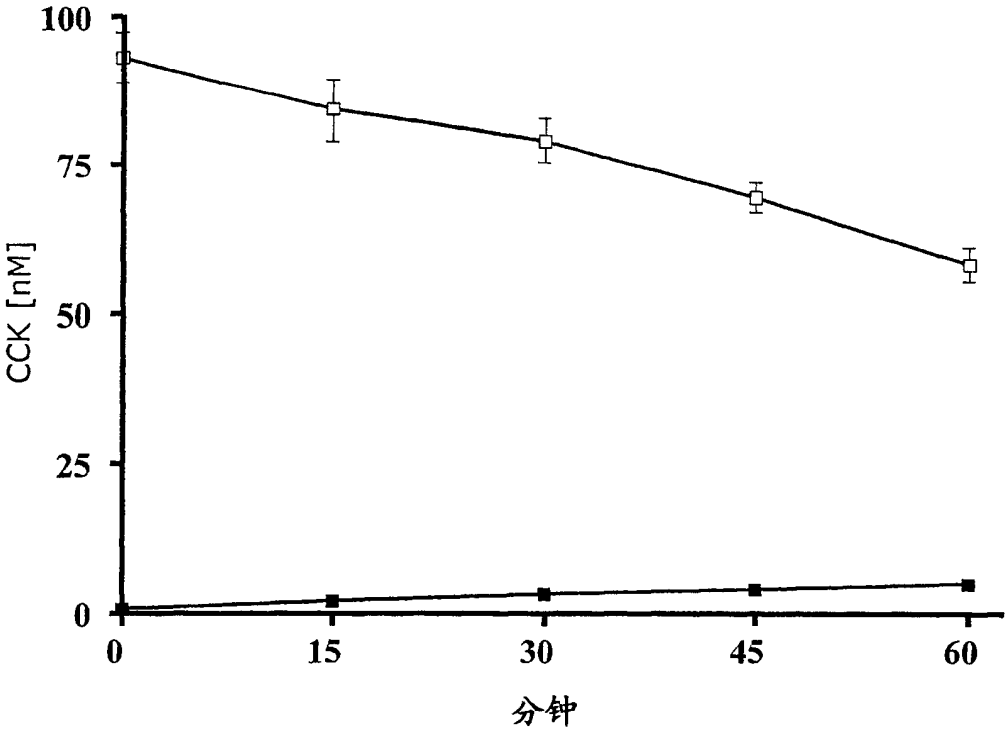


图 7A

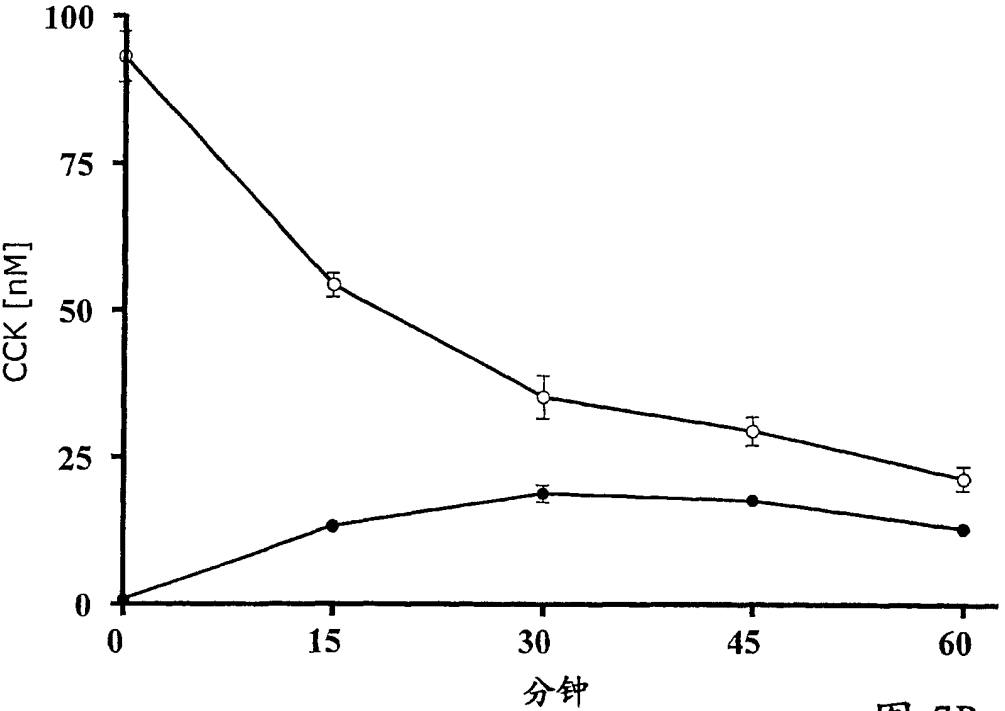


图 7B

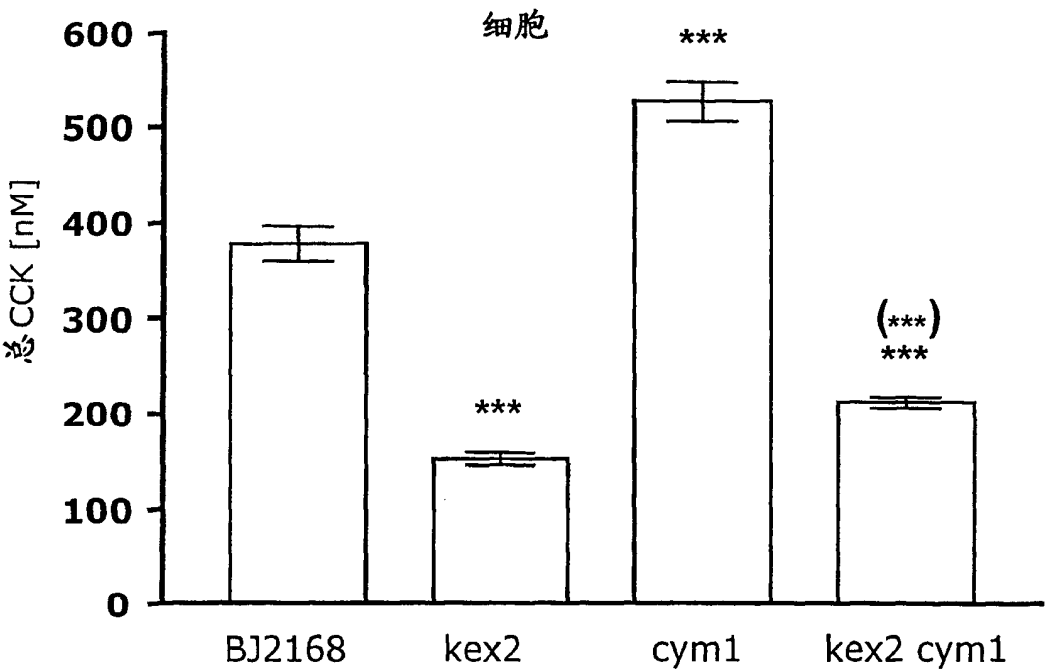


图 8A

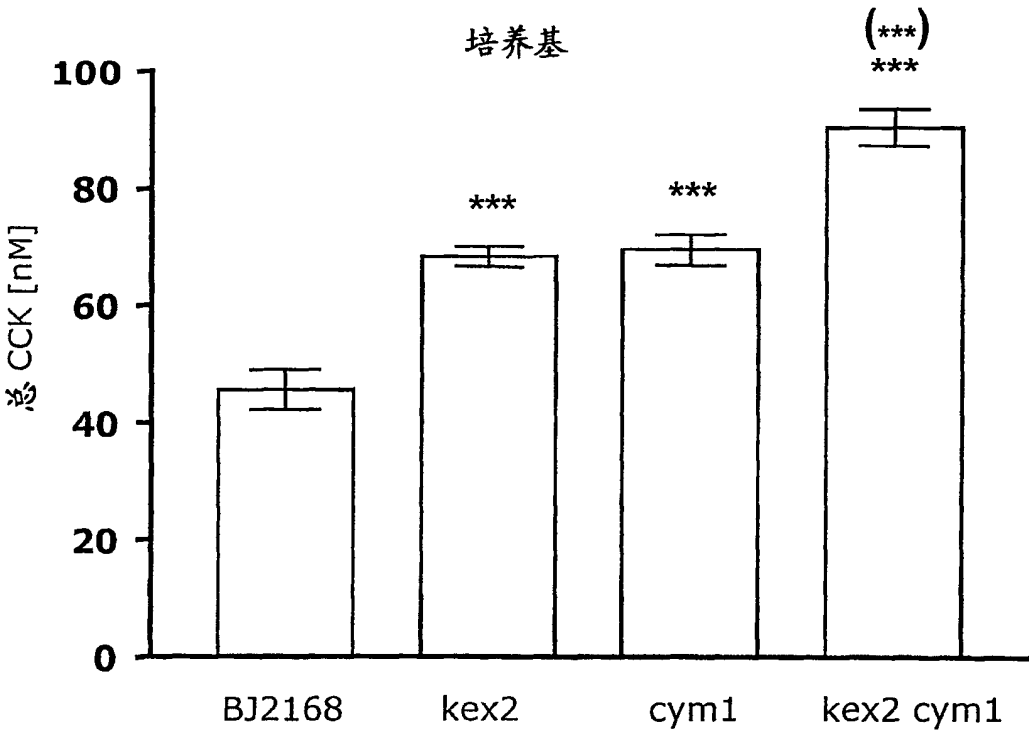


图 8B

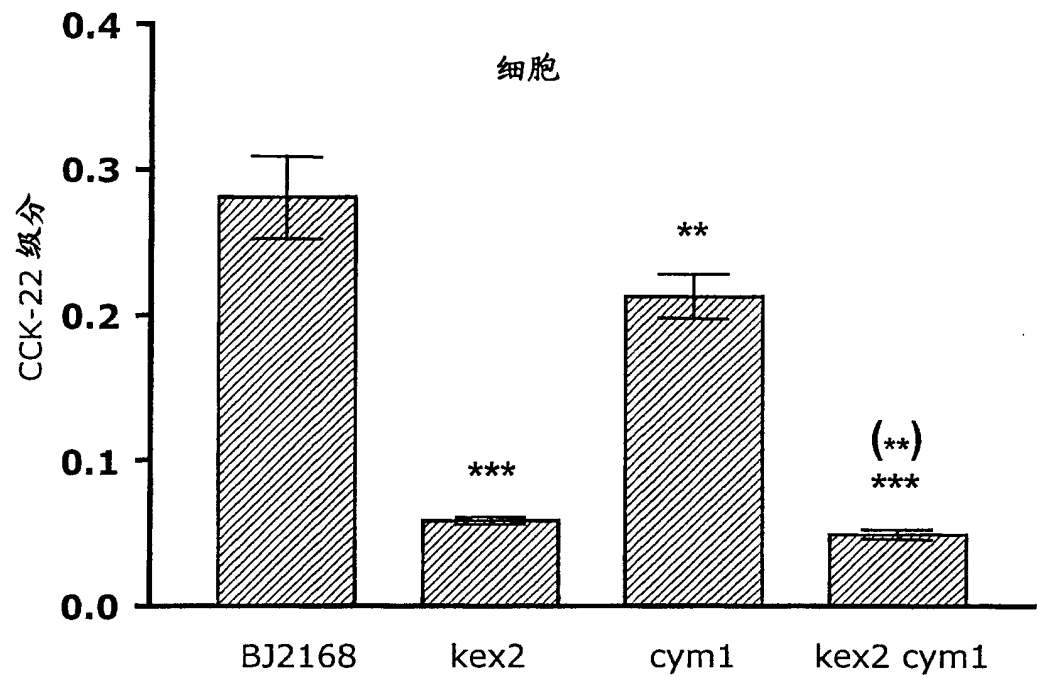


图 8C

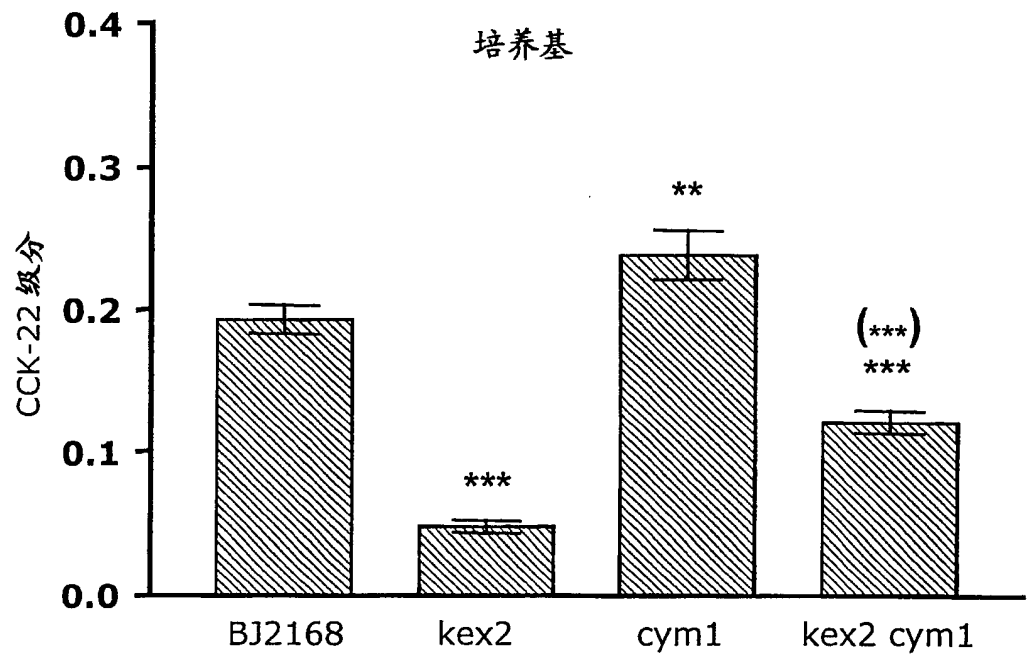


图 8D

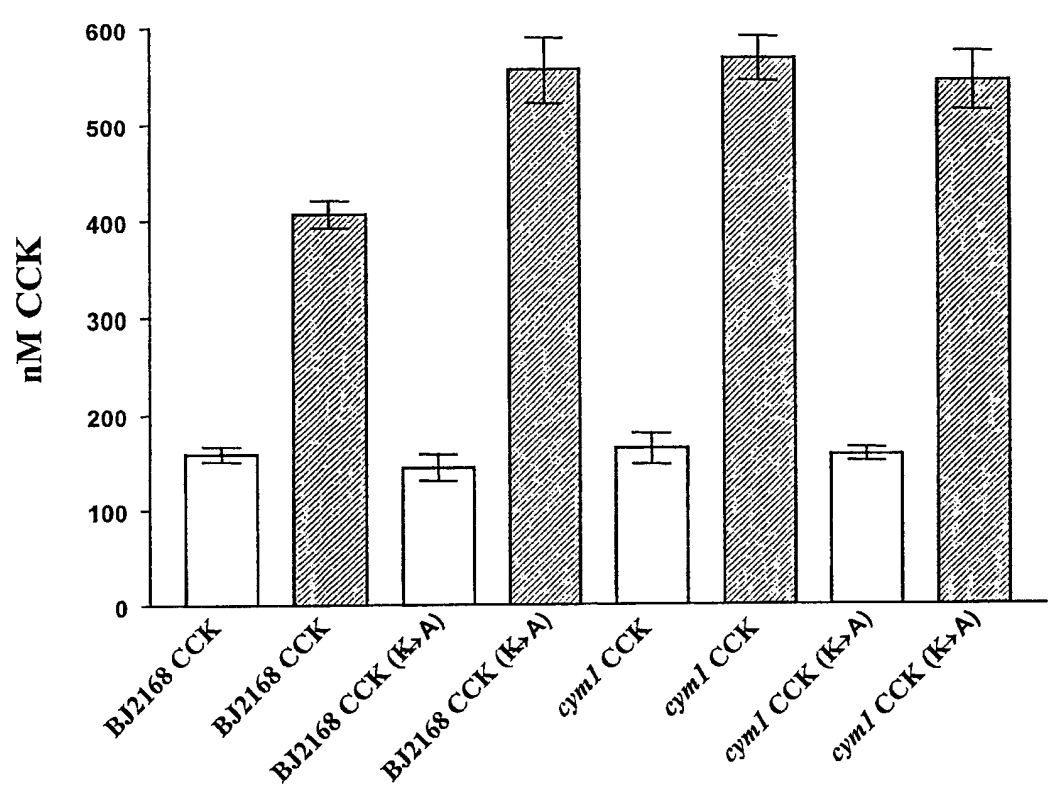
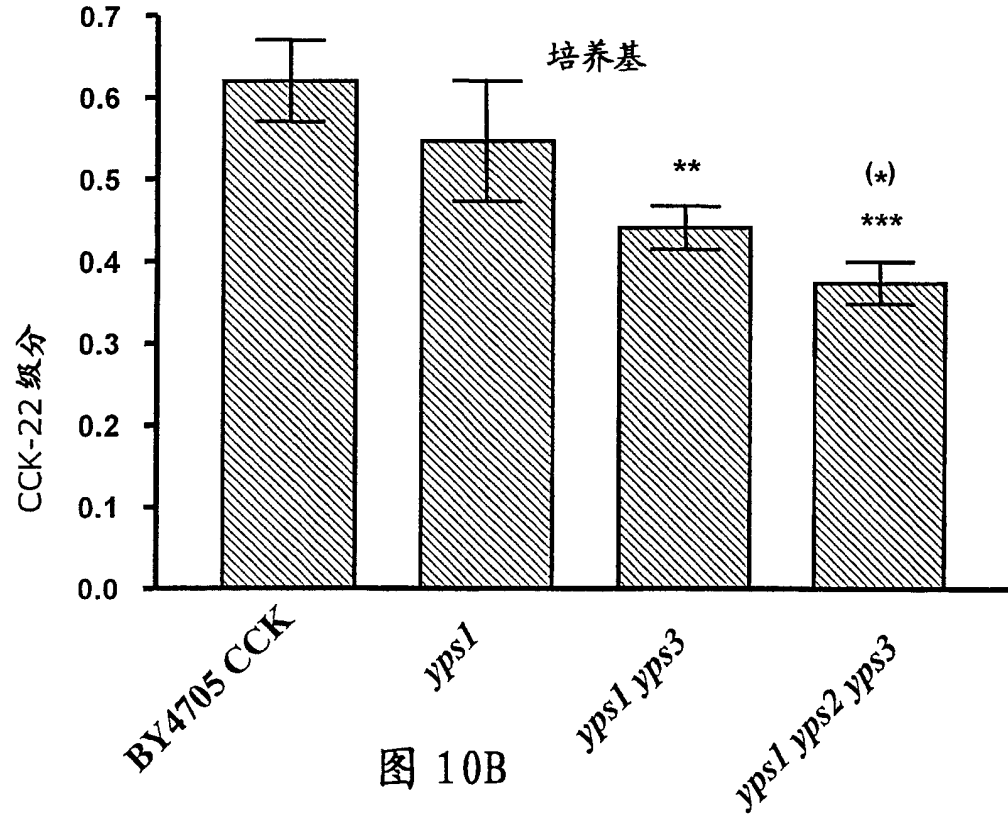
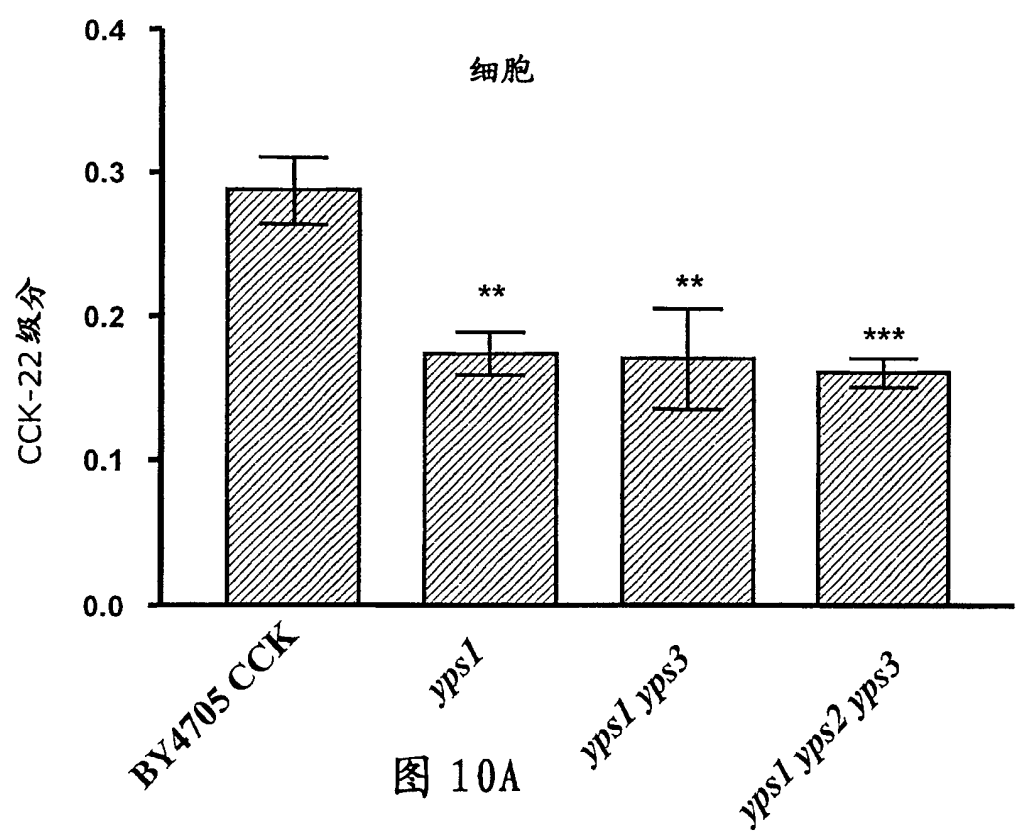
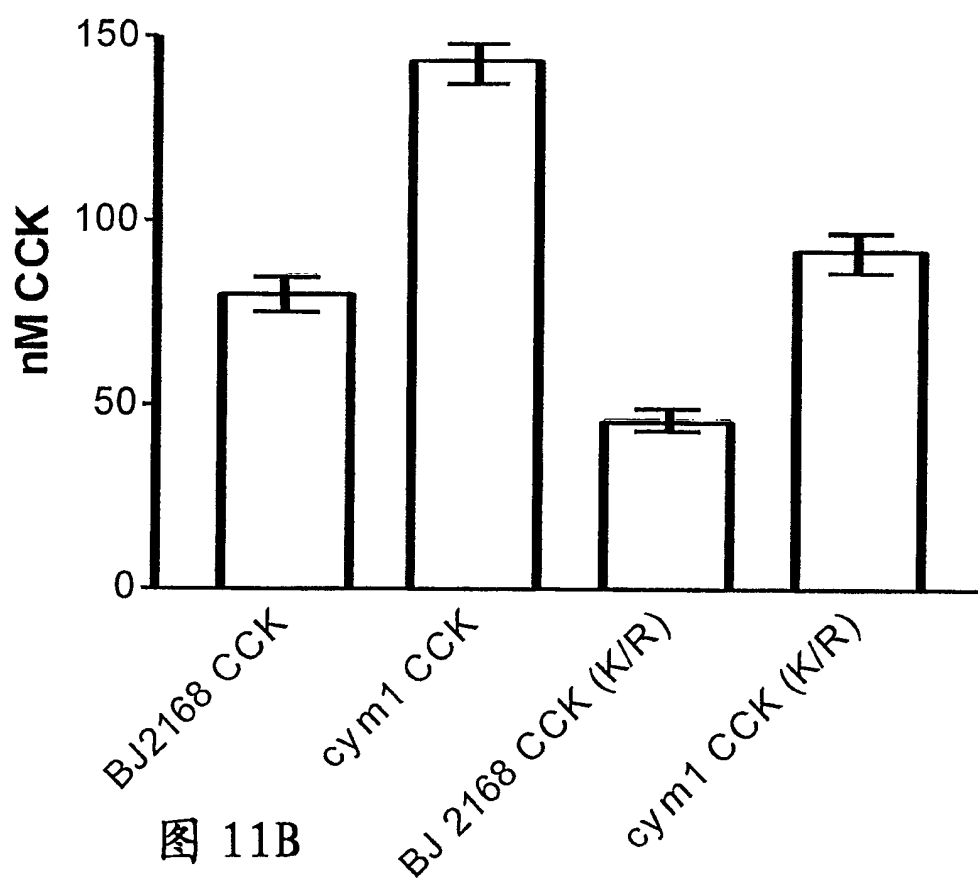
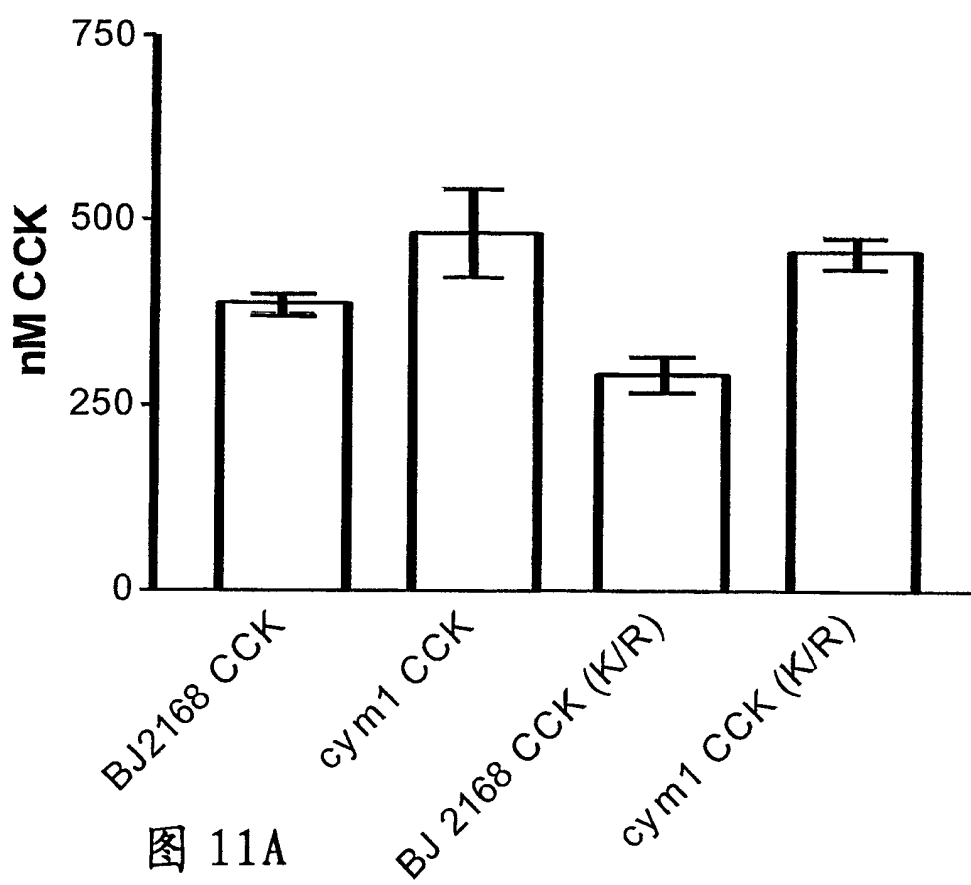


图 9





ProCCK	分子量	菌株	SEQ ID No.
CCK-61 CCK-58			
CCK-39		CCK-22 CCK-8	
Prepromfulp // .SGLQRAEEAPRRQLRVSQRTDGESRAHLGALLARYIQQARKAPSGRMSIVKNLQNLDPSHRISDRDYMGMWDFGRRSAEEYEYPS			43
QLRVSQRTDGESRAHLGALLAR	2433.3 2433.5	B	44
VSQRTDGESRAHLGALLAR	2036.1 2036.1	B	45
YIQQARKAPSGRMSIVKNLQNLDPSHRISDRDYMGMWDFGRRSAEEYEYPS	6051.9* 6051.6*	B	46
YIQQARKAPSGRMSIVK	1932.1 1932.2	A & B	47
YIQQARKAPSGRMSIV	1805.2 1805.0	A	48
NLQNLDPSHRISD	1509.0 1508.7	A	49
NLQNLDPSHRISDRDYMGMWDFG	2766.2 2766.1	A	50
NLQNLDPSHRISDRDYMGMWDFGRRSAEEYEYPS	4133.8 4133.9	A & B	51
DYMGWMDFGRRSAEEYEYPS	2488.0 2488.1	A	52

图 12

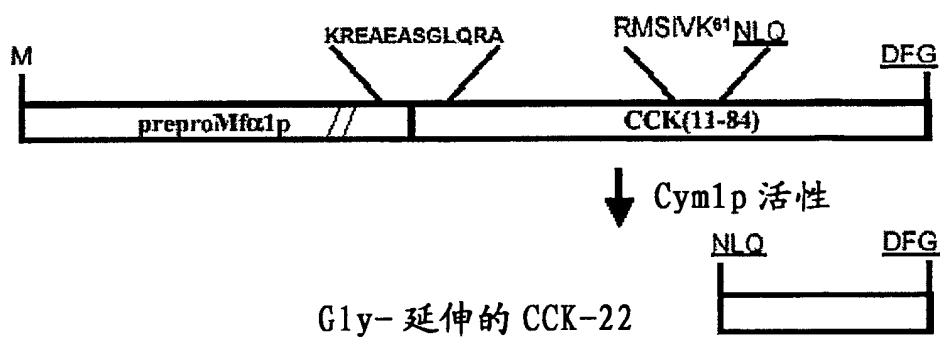


图 13A

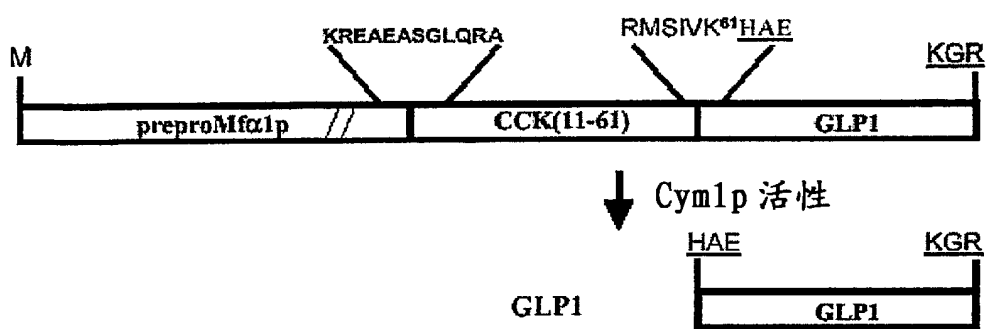


图 13B

```

ATGAGATTTCTTCAATTTTACTGCAGTTTTATTTCGCAGCATCCTCCGCATTAGCTGCT
1  -----+-----+-----+-----+-----+-----+ 60
TACTCTAAAGGAAGTTAAAAATGACGTCAAATAAGCGTCGTAGGAGGCGTAATCGACGA
M R F P S I F T A V L F A A S S A L A A -
CCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCGGCTGAAGCTGTCATCGGT
61  -----+-----+-----+-----+-----+-----+ 120
GGTCAGTTGTGATGTTGTCTTCTACTTTGCCGTGTTAAGGCCGACTTCGACAGTAGCCA
P V N T T T E D E T A Q I P A E A V I G -
TACTTAGATTTAGAAGGGGATTTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAAT
121 -----+-----+-----+-----+-----+-----+ 180
ATGAATCTAAATCTTCCCTAAAGCTACAACGACAAAACGGTAAAAGGTTGTCGTGTTTA
Y L D L E G D F D V A V L P F S N S T N -
AACGGGTTATTGTTTATAAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTA
181 -----+-----+-----+-----+-----+-----+ 240
TTGCCCAATAACAAATATTTATGATGATAACGGTCGTAACGACGATTTCTTCTTCCCAT
N G L L F I N T T I A S I A A K E E G V -
TCTTTGGATAAAAGAGAGGCTGAAGCTCACCCGCTGGGCAGCCCCGGTTCAGCCTCGGAC
241 -----+-----+-----+-----+-----+-----+ 300
AGAAACCTATTTTCTCTCCGACTTCGAGTGGGCGACCCGTCGGGGCCAAGTCGGAGCCTG
S L D K R E A E A H P L G S P G S A S D -
TTGGAACGTCCGGGTTACAGGAGCAGCGCAACCATTTGCAGGGCAAACGTGTCGGAGCTG
301 -----+-----+-----+-----+-----+-----+ 360
AACCTTTGCAGGCCCAATGTCTCGTCGCGTTGGTAAACGTCCCGTTTGACAGCCTCGAC
L E T S G L Q E Q R N H L Q G K L S E L -
CAGGTGGAGCAGACATCCCTGGAGCCCCTCCAGGAGAGCCCCCGTCCACAGGTGTCTGG
361 -----+-----+-----+-----+-----+-----+ 420
GTCCACCTCGTCTGTAGGGACCTCGGGGAGGTCTCTCGGGGGCAGGGTGTCCACAGACC
Q V E Q T S L E P L Q E S P R P T G V W -
AAGTCCCGGGAGGTAGCCACCGAGGGCATCCGTGGGCACCGCAAATGGTCCTCTACACC
421 -----+-----+-----+-----+-----+-----+ 480
TTCAGGGCCCTCCATCGGTGGCTCCCGTAGGCACCCGTGGCGTTTTACCAGGAGATGTGG
K S R E V A T E G I R G H R K M V L Y T -
CTGCGGGCACCACGAAGCCCCAAGATGGTGCAAGGGTCTGGCTGCTTTGGGAGGAAGATG
481 -----+-----+-----+-----+-----+-----+ 540
GACGCCCGTGGTGTCTCGGGGTTCTACCAGTTCCAGACCGACGAAACCCTCCTTCTAC
L R A P R S P K M V Q G S G C F G R K M -
GACCGGATCAGCTCCTCCAGTGGCCTGGGCTGCAAAGTGCTGAGGCGGCATTAA
541 -----+-----+-----+-----+-----+-----+ 594
CTGGCCTAGTCGAGGAGGTACCGGACCCGACGTTTACGACTCCGCCGTAATT
D R I S S S S G L G C K V L R R H * -

```

图 14A

```

ATGAGATTTCTTCAATTTTACTGCAGTTTATTTCGCAGCATCTCCGCATTAGCTGCT
1  -----+-----+-----+-----+-----+-----+ 60
TACTCTAAAGGAAGTTAAAAATGACGTCAAAATAAGCGTCGTAGGAGGCGTAATCGACGA

M R F P S I F T A V L F A A S S A L A A -

CCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCGGCTGAAGCTGTCATCGGT
61  -----+-----+-----+-----+-----+-----+ 120
GGTCAGTTGTGATGTTGTCTTCTACTTTGCCGTGTTTAAGGCCGACTTCGACAGTAGCCA

P V N T T T E D E T A Q I P A E A V I G -

TACTTAGATTTAGAAAGGGGATTTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAAT
121 -----+-----+-----+-----+-----+-----+ 180
ATGAATCTAAATCTTCCCCTAAAGCTACAACGACAAAACGGTAAAAGGTTGTCGTGTTTA

Y L D L E G D F D V A V L P F S N S T N -

AACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTA
181 -----+-----+-----+-----+-----+-----+ 240
TTGCCCAATAACAAATATTTATGATGATAACGGTCGTAACGACGATTCTTCTTCCCAT

N G L L F I N T T I A S I A A K E E G V -

TCTTTGGATAAAAGAGAGGCTGAAGCTAGCCCCAAGATGGTGCAAGGGTCTGGCTGCTTT
241 -----+-----+-----+-----+-----+-----+ 300
AGAAACCTATTTTCTCTCCGACTTCGATCGGGGTTCTACCACGTTCCAGACCGACGAAA

S L D K R E A E A S P K M V Q G S G C F -

GGGAGGAAGATGGACCGGATCAGCTCCTCCAGTGGCCTGGGCTGCAAAGTGCTGAGGCGG
301 -----+-----+-----+-----+-----+-----+ 360
CCCTCCTTCTACCTGGCCTAGTCGAGGAGGTCACCGGACCCGACGTTTCACGACTCCGCC

G R K M D R I S S S S G L G C K V L R R -

CATTA
361 ----- 366
GTAATT

H * -

```

图 14B

```

      ATGAGATTTCTTCAATTTTTACTGCAGTTTTATTTCGCAGCATCCTCCGCATTAGCTGCT
1  -----+-----+-----+-----+-----+-----+ 60
      TACTCTAAAGGAAGTTAAAAATGACGTCAAAATAAGCGTCGTAGGAGGCGTAATCGACGA

      M R F P S I F T A V L F A A S S A L A A -

      CCAGTCAACACTACAACAGAAGATGAAACGGCACAATTCGGCTGAAGCTGTCATCGGT
61 -----+-----+-----+-----+-----+-----+ 120
      GGTCA GTTGTGATGTTGTCTTCTACTTTGCCGTGTTTAAGGCCGACTTCGACAGTAGCCA

      P V N T T T E D E T A Q I P A E A V I G -

      TACTTAGATTTAGAAGGGGATTTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAAT
121 -----+-----+-----+-----+-----+-----+ 180
      ATGAATCTAAATCTTCCCCTAAAGCTACAACGACAAAACGGTAAAAGGTTGTCGTGTTTA

      Y L D L E G D F D V A V L P F S N S T N -

      AACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTA
181 -----+-----+-----+-----+-----+-----+ 240
      TTGCCCAATAACAAATATTTATGATGATAACGGTCGTAACGACGATTTCTTCTTCCCAT

      N G L L F I N T T I A S I A A K E E G V -

      TCTTTGGATAAAAGAAGCCCCAAGATGGTGCAAGGGTCTGGCTGCTTTGGGAGGAAGATG
241 -----+-----+-----+-----+-----+-----+ 300
      AGAAACCTATTTCTTCGGGGTTCTACCACGTTCCCAGACCGACGAAACCCTCCTTCTAC

      S L D K R S P K M V Q G S G C F G R K M -

      GACCGGATCAGCTCCTCCAGTGGCCTGGGCTGCAAAGTGCTGAGGCGGCATTAA
301 -----+-----+-----+-----+-----+-----+ 354
      CTGGCCTAGTCGAGGAGGTCACCGGACCCGACGTTTCACGACTCCGCCGTAATT

      D R I S S S S G L G C K V L R R H * -

```

图 14C

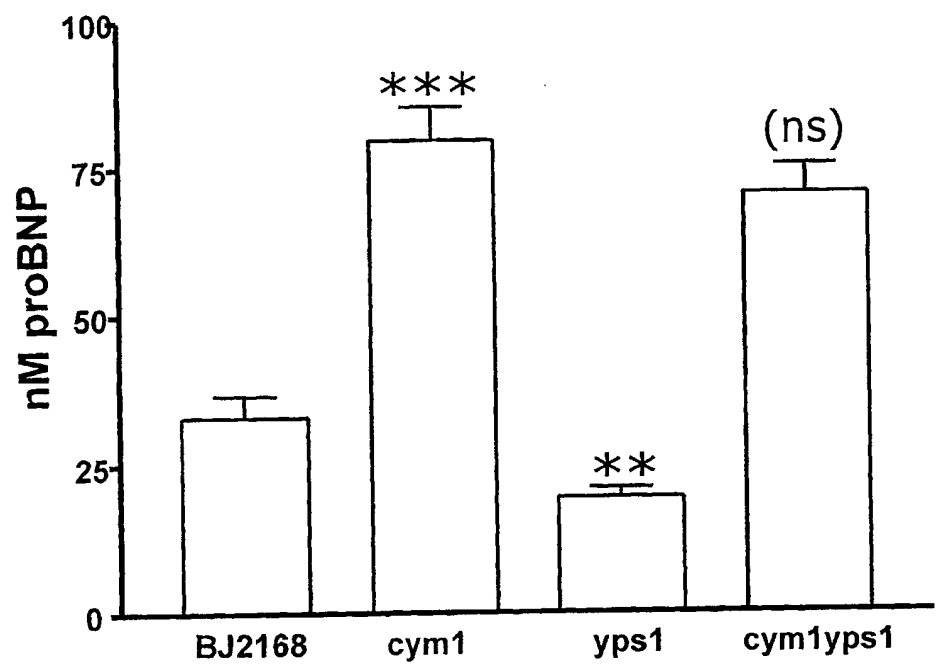


图 15A

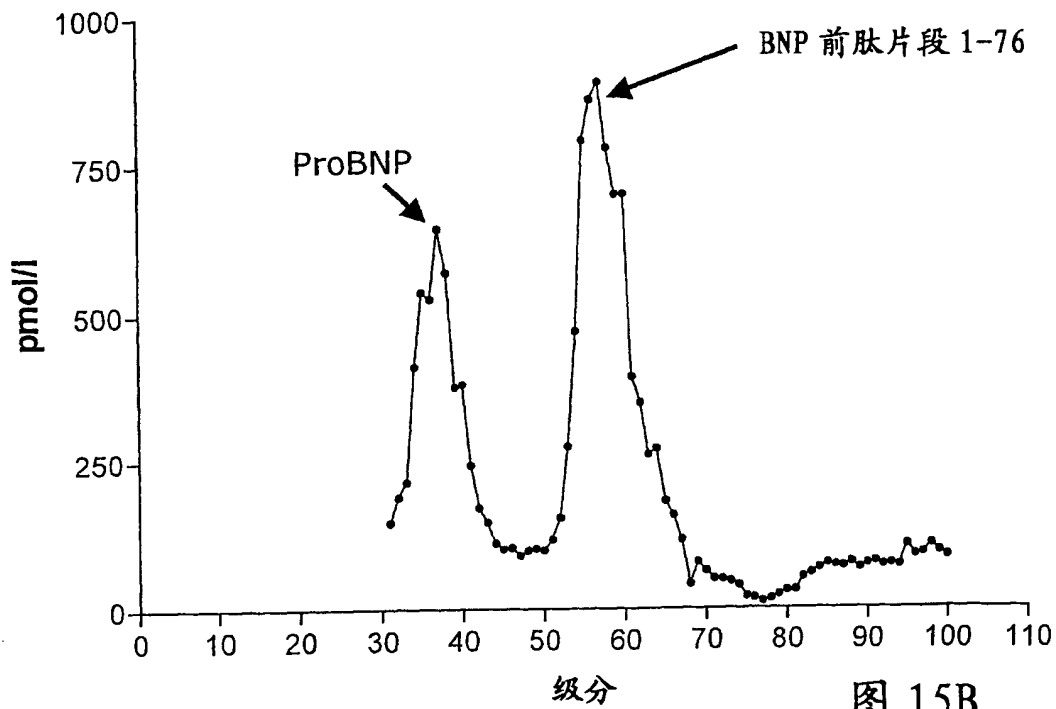


图 15B