

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

C12N 15/81 (2006.01)

C12N 15/62 (2006.01)



## [12] 发明专利申请公开说明书

[21] 申请号 200480026522.3

[43] 公开日 2006年10月25日

[11] 公开号 CN 1852983A

[22] 申请日 2004.9.15

[21] 申请号 200480026522.3

[30] 优先权

[32] 2003.9.16 [33] DE [31] 10342794.5

[86] 国际申请 PCT/EP2004/010346 2004.9.15

[87] 国际公布 WO2005/033316 德 2005.4.14

[85] 进入国家阶段日期 2006.3.15

[71] 申请人 巴斯福股份公司

地址 德国路德维希港

[72] 发明人 K·奥斯特曼 G·勒德尔

[74] 专利代理机构 北京市中咨律师事务所

代理人 黄革生 隗永良

权利要求书 3 页 说明书 35 页 序列表 36 页  
附图 9 页

[54] 发明名称

蛋白质从酵母的分泌

[57] 摘要

本发明涉及包含核酸序列的表达构建体，所述核酸序列编码可通过酵母细胞加工的穿梭肽构建体；涉及包含此类构建体的对应表达载体；涉及在其帮助下重组制备靶蛋白的方法；涉及用其转化的宿主；涉及穿梭肽及编码所述穿梭肽的核酸序列；涉及编码与外源蛋白质融合的此类穿梭肽的核酸序列；涉及用此种穿梭肽制备的疏水蛋白及涉及疏水蛋白用于涂布物体例如皮革的用途。

1. 表达构建体，其包含编码穿梭肽构建体的核酸序列，所述穿梭肽构建体是酵母细胞可加工的，具有式

(Sig-SP)，

并且所述表达构建体在 5'-3' 方向包含核酸序列，其编码

a) 信号肽(Sig)，其可加工地连接

b) 所述酵母细胞分泌的至少一种穿梭肽(SP)。

2. 权利要求 1 中要求保护的表达构建体，其中穿梭肽构建体(Sig-SP)来自通过裂殖酵母属，特别是粟酒裂殖酵母加工的多肽。

3. 前面权利要求任意一项要求保护的表达构建体，其中穿梭肽构建体(Sig-SP)来自酵母的信息素前蛋白质，所述信息素(Pher)来自该前蛋白质并且可通过 N-及 C-末端加工分泌。

4. 权利要求 3 中要求保护的表达构建体，其中信号多肽(Sig)为信息素前蛋白质的可蛋白酶解除去的天然信号多肽。

5. 权利要求 4 中要求保护的表达构建体，其中经 C 末端加工的信息素(Pher)包含 C 末端蛋白酶切割位点。

6. 前面权利要求中任意一项要求保护的表达构建体，其还包含编码与穿梭肽构建体(Sig-SP)的 C 末端可加工地连接的同源或异源靶蛋白(Targ)的核酸序列。

7. 前面权利要求任意一项中要求保护的表达构建体，其包含编码融合蛋白的核酸序列，所述融合蛋白可通过酵母细胞加工且具有式

$$\text{Sig-L1}_n\text{-Pher-L2}_m\text{-Targ}$$

其中

Sig、Pher 及 Targ 如以上定义，

L1 及 L2 为可加工的接头并且

n 及 m 相互独立地为 0 或 1。

8. 前面权利要求中任意一项要求保护的表达构建体，其中编码穿梭肽

构建体(Sig-SP)的核酸序列包含根据 SEQ ID NO: 32 的信号多肽编码序列或其功能等同物, 其有效连接编码成熟信息素蛋白(P 因子)的根据 SEQ ID NO: 5 的核酸序列或其功能等同物。

9. 前面权利要求中任意一项要求保护的表达构建体, 其中编码穿梭肽构建体的核酸序列包含根据 SEQ ID NO:1 的序列, 其在适当时在 3'端由编码靶蛋白(Targ)的序列延伸。

10. 前面权利要求中任意一项要求保护的表达构建体, 其中靶蛋白为疏水蛋白, 特别是I类疏水蛋白。

11. 如权利要求 10 中要求保护的表达构建体, 其中疏水蛋白选自 SEQ ID NO: 14(DewA)、SEQ ID NO: 19(RdIA)、SEQ ID NO: 20(RdIB)、SEQ ID NO: 21(HYP1)及 SEQ ID NO: 22(HYP4)或通过根据 SEQ ID NO: 13 的核酸序列编码。

12. 表达载体, 其包含前面权利要求中任意一项要求保护的表达构建体, 该表达构建体与至少一种调节核酸序列有效连接。

13. 重组微生物, 其包含适当时稳定整合到宿主基因组的至少一个如权利要求 12 中要求保护的表达载体或如权利要求 1 至 11 中任意一项要求保护的表达构建体。

14. 如权利要求 13 中要求保护的微生物, 其选自酵母。

15. 如权利要求 14 中要求保护的微生物, 其选自裂殖酵母属, 特别是粟酒裂殖酵母的酵母。

16. 通过酵母细胞可加工且来自酵母的信息素前蛋白质的穿梭肽构建体(Sig-SP), 其中信息素来自所述前蛋白质且可通过 N-及 C-末端加工分泌。

17. 如权利要求 16 中要求保护的穿梭肽构建体, 其包含与 C 末端可加工的信息素多肽的 N 末端可加工地连接的信号多肽。

18. 如权利要求 17 中要求保护的穿梭肽构建体, 其中信号多肽为信息素前蛋白质的可蛋白酶解除去的天然信号多肽。

19. 如权利要求 17 中要求保护的穿梭肽构建体, 其中 C 末端加工的信息素多肽包含 C 末端蛋白酶切割位点。

20. 如在权利要求 16 至 19 中任意一项要求保护的穿梭肽构建体,其包含如 SEQ ID NO: 2 中定义的氨基酸序列或其功能等同物。

21. 重组制备靶蛋白的方法,其包括培养如权利要求 13 至 15 中任意一项的微生物、表达编码所述靶蛋白的核酸序列及分离分泌到培养基中的靶蛋白。

22. 如权利要求 21 中要求保护的方法,其中靶蛋白为如权利要求 10 或 11 中定义的疏水蛋白。

23. 核酸,其编码如权利要求 16 至 20 中任意一项要求保护的穿梭肽构建体。

24. 核酸,其如权利要求 1 至 11 中任意一项定义。

25. 疏水蛋白,其可以通过如权利要求 22 中要求保护的方法获得。

26. 如权利要求 25 中要求保护的疏水蛋白用于表面处理的用途。

27. 如权利要求 26 中要求保护的用途,其中处理选自玻璃、纤维、织物、皮革、上漆物品、膜及正面的物体的表面。

28. 如权利要求 10 或 11 中定义的疏水蛋白的用途,用于纤维、织物及皮革的表面处理。

## 蛋白质从酵母的分泌

本发明涉及包含核酸序列的表达构建体，所述核酸序列编码可通过酵母细胞加工的穿梭肽构建体；涉及包含此类构建体的对应表达载体；涉及在其帮助下重组制备靶蛋白的方法；涉及用其转化的宿主；涉及穿梭肽及编码所述穿梭肽的核酸序列；涉及编码与外源蛋白质融合的此类穿梭肽的核酸序列；涉及用此种穿梭肽制备的疏水蛋白及涉及疏水蛋白用于涂布物体例如皮革的用途。

### 现有技术

#### a) 酵母中的表达

酵母广泛用作用于异源蛋白质表达的宿主。原因是酵母表达系统具有几个优点，因为与细菌及其它真核细胞相比，酵母可以较高密度生长且能够进行蛋白质糖基化及翻译后修饰。此外，由于酵母对细胞裂解具有高度抗性且生长培养基一般包含低量外源蛋白质，通过酵母产生及分泌的产物可以简单方式纯化。此外，酵母在便宜营养培养基中以高密度比其它真核细胞更快地生长。

在现有技术中可发现用于在酵母中表达及分泌异源蛋白质的多种不同方法。因此，例如，US 5,642,487描述了用于在酵母中重组产生蛋白质的方法，所述方法包括用编码结构元件的表达盒转化酵母，所述结构元件的表达盒编码动物肽神经激素的前导序列、产生 $\alpha$ -螺旋结构的衔接头序列、加工信号及结构基因。

自现有技术进一步知道用通过酵母产生的信息素 $\alpha$ -因子基因的调节元件控制酵母中异源蛋白质的表达。因此，例如， $\alpha$ -因子信号前导肽序列用于表达异源蛋白质(参照例如，US 5,010,182)。

此外，出版的美国专利申请 US 2003/0077831 公开了用于在酵母中表达异源蛋白质的表达载体，所述表达载体包含，侧翼为合适转录起始及翻译起始及终止序列的杂合前体多肽的编码序列，所述杂合前体多肽的元件包含酵母分泌的蛋白质的信号肽及前导肽和异源蛋白质，其侧翼为所述异源蛋白质 N 末端及 C 末端前肽序列的异源蛋白质。

#### b) 疏水蛋白

疏水蛋白为小的富含半胱氨酸的蛋白质，其包含大约 100 个氨基酸残基且具有令人感兴趣的技术性质。它们能使疏水性表面亲水。它们使亲水性表面成为疏水的。

然而，有许多关于疏水蛋白及其应用的专利：因此，例如，WO-A-96/41882 描述了来自可食用真菌的疏水蛋白(参照 SEQ ID NO: 21 及 22)。WO-A-00/58342 涉及通过相提取对包含疏水蛋白的融合蛋白的纯化。WO-A-01/57066 描述了由于亚硫酸盐处理，疏水蛋白的稳定化、溶解及其相关经改进的应用。WO-A-01/57076 描述了通过吸附到特氟隆(Teflon)珠且通过去污剂如吐温(Tween)于低温洗脱纯化疏水蛋白。WO-A-01/57528 描述了通过应用吐温及达 85 摄氏度的温度将疏水蛋白固定到表面。

WO-A-01/74864 描述了丝状细菌，特别是链霉菌种(*Streptomyces* sp.) 的非典型疏水蛋白(仅仅一个二硫桥)，称为 RdIA 及 RdIB(参照 SEQ ID NO: 19 及 20)。疏水蛋白用于多种物体如窗户、隐形眼镜、车辆的表面处理。进一步建议在重组宿主中产生于此描述的蛋白质，所述宿主释放所述蛋白质至培养基中。去除宿主后，预计包含疏水蛋白的培养基适合用于表面涂布。没有提供实际表达及分泌的实验证据。

#### 发明概述

本发明的一个目的是提供方法，所述方法可能使酵母中特别是粟酒裂殖酵母(*Schizosaccharomyces pombe*)中表达的同源或特别是异源蛋白质自酵母细胞分泌到周围培养基中。特别地，提供了能自宿主细胞分泌重组产生的疏水蛋白的方法。

我们发现通过提供包含编码穿梭肽构建体的核酸序列的表达构建体能够实现该目标，所述穿梭肽构建体可通过酵母细胞加工，具有式

(Sig-SP)

并且所述表达构建体在 5'-3'方向包含核酸序列，其编码

- a) 信号肽(Sig)，其可加工地连接
- b) 所述酵母细胞分泌的至少一种穿梭肽(SP)。

通过构巢曲霉(*Aspergillus nidulans*)疏水蛋白 DewA(具有根据 SEQ ID NO:13 编码序列的根据 SEQ IN NO: 14 的成熟蛋白质；具有根据 SEQ IN NO:15 编码核酸序列的信号序列 SEQ ID NO: 16 的前蛋白质)作为异源靶蛋白(Targ)的实例以模型方式示例实现以上目的。该蛋白质为I类疏水蛋白，即能自组装分泌的真菌外被蛋白的代表。

特别地，编码靶蛋白(DewA)的 DNA 序列(SEQ ID NO: 13)与编码粟酒裂殖酵母肽信息素(P 因子，成熟 P 因子的根据 SEQ ID NO: 6 的氨基酸序列)的 DNA 序列(成熟 P 因子的 SEQ ID NO: 5)的 3'端融合。产生的融合蛋白包含用于分泌所述信息素所需的所有信号序列及与其融合的靶蛋白，特别是可去除的信号肽(SEQ ID NO:4)。分泌包含融合蛋白的蛋白酶解加工。结果，信息素(P 因子)(SEQ ID NO:6)及靶蛋白(疏水蛋白，SEQ ID NO:14)独自分泌到培养基中。

本发明的发现是令人惊奇的，因为 P 因子前蛋白质(成熟信息素的 N 末端)的实际调节元件控制显然不足以控制酵母细胞对靶蛋白的分泌。只有应用其中额外的共分泌蛋白质组分(成熟信息素)可加工地位于待分泌的靶蛋白上游的构建体才能使所述靶蛋白以所希望的方式分泌到培养基中。

## 发明详述

### a) 一般信息

在说明书及附图中蛋白质序列一般以“单字母密码”标明。

根据本发明，通过宿主细胞，特别是通过酵母细胞内表达且通过内源性细胞机制通过细胞膜从细胞分泌，优选分泌至周围培养基中的蛋白质为

“可分泌的”。

根据本发明，如果蛋白质前体可通过在宿主细胞内和/或细胞外酶解加工转化为成熟形式，那么蛋白质前体(即原来表达形式的蛋白质，例如在成熟的经加工蛋白质中不再存在 N 和/或 C 末端肽序列的前蛋白质)为“可加工的”。

如果待加工蛋白质中的个别蛋白质区通过肽键连接，该肽键可以通过宿主细胞蛋白酶切割，则存在“可加工的键”。

“加工”可在成熟经加工蛋白质(靶蛋白)序列的 N 末端发生，且适当时也在 C 末端发生。

虽然“同源”靶蛋白原来在根据本发明应用的宿主中表达且因此为宿主内源性蛋白质，由于用本发明的表达构建体转化所述宿主，它通过宿主细胞分泌。

“异源”靶蛋白原来在根据本发明应用的宿主中不表达且因此不为内源性宿主蛋白质，但由于用本发明的表达构建体转化所述宿主，它通过宿主细胞分泌。

“穿梭肽”为在根据本发明应用的宿主细胞中可加工的“穿梭肽构建体”的部分。它与其在 C 和/或 N 末端，优选 N 末端连接的一个或多个可加工的调节肽片段，如信号序列、前导序列一起形成所述穿梭肽构建体。与信号肽相反，例如，穿梭肽为通过宿主细胞分泌的多肽。优选在细胞内进行调节元件的加工。甚至优选穿梭肽在 C 末端与靶蛋白可加工地融合时，穿梭肽也保持可分泌的。此 C 末端加工，即在分泌过程中蛋白酶解除去靶蛋白优选例如在通过宿主细胞包膜期间，或在细胞外间隙，例如在周围培养基中，通过内源性细胞蛋白酶进行。

根据本发明的“表达构建体”或“表达盒”包含在特定宿主系统如，特别是酵母细胞中控制表达所需的用于转录的起始及终止信号及适当时翻译信号，其有效连接如以上定义的可加工的穿梭肽构建体的编码核酸序列。表达构建体特别包含转录因子的结合位点。编码序列的 5'上游包含在宿主细胞中可操作的组成型或诱导型、自身或异源的、天然或合成的启动子。

此外表达构建体包含许多限制酶切位点，例如用于将所述构建体插入到表达载体中的限制酶切位点。此外，表达构建体可包含选择标记基因。

“表达载体”描述了通过将本发明的表达盒导入到复制子，例如导入到质粒、粘粒或病毒中可获得的构建体。此种载体能自主复制或能整合到宿主基因组中且包含的所需的控制序列，所述所需的控制序列用于控制编码如以上定义的可加工穿梭肽构建体的根据本发明的核酸序列的转录和(适当时)翻译。

#### b) 优选实施方案

本发明首先涉及表达构建体，所述表达构建体包含编码可通过酵母细胞加工，具有式

(Sig-SP)

的穿梭肽构建体的核酸序列

并且在 5'-3' 方向包含核酸序列，其编码

a) 信号肽(Sig)，其可加工地连接

b) 通过所述酵母细胞分泌的至少一种穿梭肽(SP)；

和适当时，促进加工和/或分泌并位于编码信号肽序列的 5'- 或 3' - 端的一个或多个核酸序列。

编码 SP 及 Sig 的序列在相同读框内且，此外，在翻译期间在 Sig 的 C 末端及 SP 的 N 末端之间形成可加工序列。所述可加工序列的实例为人工导入的、可蛋白酶解切割的天然或合成的衔接头序列，然而，所述衔接头序列优选为 Sig 的 C 末端或 SP 的 N 末端的部分。可对衔接头序列进行加工使经切割序列完全或部分存在于 Sig 的 C 末端或 SP 的 N 末端。只要对 SP 的分泌能力没有实质的不利作用且特别不会防止 SP 分泌的能力，那么经切割序列完全或部分存在于 SP 的 N 末端是可能的。

本发明特别涉及编码表达构建体，所述表达构建体编码来自更广义上通过酵母加工的多肽的可加工穿梭肽。所述酵母为特别选自子囊菌(ascmycetes)的酵母。优选选自原子囊菌(Archiascomycetes)纲裂殖酵母(Schizosaccaromycetales)目的酵母，且酵母特别优选选自裂殖酵母

(*Schizosaccharomyces*)属, 如粟酒裂殖酵母。虽然有数据表明负型细胞也分泌 P 因子, 但是优选与交配因子(信息素)匹配的菌株(即正型细胞用于正因子(P 因子)且负型细胞用于负因子(M 因子)。

可加工的穿梭肽构建体特别来自酵母的信息素前蛋白质, 所述信息素从所述前蛋白质通过 N 及 C 末端加工产生。信息素优选具有通过加工可去除的 N 末端多肽, 所述多肽特别包含加工和/或分泌前蛋白质所需的元件, 如信号肽和适当时, 前导肽, 以及所需的蛋白酶切割位点。

真菌信息素是已知的且例如对担子菌(basidiomycetes)如玉蜀黍黑粉菌(*Ustilago maydis*)(Urban, M., Kahmann, R. 和 Bolker, M. (1996) *The biallelic a mating type locus of Ustilago maydis: remnants of an additional pheromone gene indicate evolution from a multiallelic ancestor (Mol Gen Genet 250(4): 414-420)*)或 *Coprinopsis cinerea*(Halsall, J. R., Milner, M. J. 和 Casselton, L. A. (2000) *Three subfamilies of pheromone and receptor genes generate multiple B mating specificities in the mushroom Coprineus cinereus (Genetics 154(3):1115-1123)*)及对于子囊菌如粟酒裂殖酵母 (Imai, Y 和 Yamamoto, M. (1995) *The fission yeast mating pheromone P-factor: its molecular structure, gene structure, and ability to induce gene expression and G1 arrest in the mating partner (Genes Dev 8(3):328-338)*, Davey, J. (1992) *Mating pheromones of the fission yeast Schizosaccharomyces pombe: purification and structural characterization of M-factor and isolation and analysis of two genes encoding the pheromone (EMBO J 11(3): 951-960)*)、酿酒酵母 (*Saccharomyces cerevisiae*)(Michaelis, S.和 Herskowitz, I. (1988) *The a-factor pheromone of Saccharomyces cerevisiae is essential for mating (MOL cell Biol 8(3):1309-1318)*)、Kurjan, J. 和 Herskowitz, I. (1982) *Structure of a yeast pheromone gene (MF-alpha): a putative alpha-factor precursor contains four tandem copies of mature alpha-factor (Cell 30(3):933-943)*)、*Kluyveromyces delphensis* (Wong, S., Fares, M.A., Zimmermann, W.,

Butler, G.和 Wolfe, K. H. (2003) Evidence from comparative genomics for a complete sexual cycle in the 'asexual' pathogenic yeast *Candida glabrata* (Genome Biol 4 (2) R 10)) 及克鲁维酵母 (*Saccharomyces kluyveri*)(Egel-Mitani, M.和 Hansen, M.T. (1987) Nucleotide sequence of the gene encoding the *Saccharomyces kluyveri* alpha mating pheromone (Nucleic Acids Res 15 (15) 6303))进行描述。

根据本发明合适的信息素为相对小肽(例如, 5 至 40 个或 8 至 30 个氨基酸)。它们的一级序列一般不显示显著同源性。它们形成为前蛋白质, 其经蛋白酶解加工并释放到培养基中。

特别合适的信息素或对应前蛋白质的实例为来自粟酒裂殖酵母的“P 因子”及“M 因子”及它们的前蛋白质(参见 Imai, Y 和 Yamamoto, M. (1995) The fission yeast mating pheromone P-factor: its molecular structure, gene structure, and ability to induce gene expression and G1 arrest in the mating partner (Genes Dev 8(3):328-338), Davey, J. (1992) Mating pheromones of the fission yeast *Schizosaccharomyces pombe*: purification and structural characterization of M-factor and isolation and analysis of two genes encoding the pheromone (EMBO J 11(3): 951-960), Kjaerulff, S., Davey, J.和 Nielsen, O. (1994) Analysis of the structural genes encoding M-factor in the fission yeast *Schizosaccharomyces pombe*: identification of a third gene, *mfm3* (Mol Cell Biol 14(6) 3895-3905))。

例如, P 因子前蛋白质具有根据 SEQ ID NO: 9 的 DNA 序列及根据 SEQ ID NO:10 的蛋白质序列。前蛋白质包含与四个连续信息素肽序列桥接的 N 末端信号肽序列, 它们通过加工可以分离(见图 3)。

在本发明优选的构建体中, 设计可加工的穿梭肽构建体以包含与 C 末端可加工的信息素多肽(Pher)的 N 末端可加工地连接的信号多肽(Sig)。

信号多肽特别包含信息素前蛋白质的可蛋白酶解除去的天然信号肽(例如通过 SEQ ID NO:3 编码的 SEQ ID NO:4)或与其相同。

进一步优选包含 C 末端蛋白酶切割位点的 C 末端加工的信息素多肽。

表达构建体进一步优选包含编码与穿梭肽构建体(Sig-SP)的C末端可加工连接的同源或异源靶蛋白(Targ)的核酸序列。

本发明优选涉及上述类型的表达构建体,所述表达构建体包含编码融合蛋白的核酸序列,所述融合蛋白可通过酵母细胞加工且具有式



其中

Sig、Pher 及 Targ 如以上定义,

L1 及 L2 为可加工接头或衔接头序列且

n 及 m 相互独立地为 0 或 1。然而,优选 n 为 1 且 m 为 0。

L1 及 L2 可为天然或合成接头。它们包含至少一个可蛋白酶解加工的肽序列。适当时,具有促进例如加工、分泌、转录和/或翻译功能的其它效应子可与 L1 和/或 L2 相关。

特别优选表达构建体,其中编码可加工的穿梭肽构建体的核酸序列包含与编码成熟 P 因子信息素(Pher)的根据 SEQ ID NO: 5 核酸序列或其功能等同物有效连接的根据 SEQ ID NO: 3 的信号多肽(Sig)编码序列或其功能等同物。

优选地,不存在 L2 接头。相比,优选提供 L1 接头,其包含编码根据 SEQ ID NO:10 中的氨基酸残基 21 至 30 的多肽的序列。这里 L1 桥接信号肽与前激素的第一个信息素构件(SEQ ID NO:10 中的 31 至 57 位)。L1 的 C 末端对应于蛋白酶解加工所需的蛋白酶识别序列。

在特别优选的实施方案中,编码可加工的穿梭肽构建体的核酸序列包含根据 SEQ ID NO:1 的序列。

在粟酒裂殖酵母中存在的第二种信息素 M 因子的帮助下,原则上也可应用相同方法且用于表达任意同源及异源靶蛋白。有三种基因组基因(mfm1<sup>+</sup>, SEQ ID NO: 42; mfm2<sup>+</sup>, SEQ ID NO: 45; mfm3<sup>+</sup> SEQ ID NO: 48), 它们每种情况下编码 M 因子——具有负交配类型的细胞的信息素。首先从每种基因产生在分泌过程中加工的前蛋白质(SEQ ID NO: 43、46 及 49)。最后,分别通过 SEQ ID NO: 44、47 及 50 编码的 M 因子(YTPKVPYMC;

SEQ ID NO: 51)作为成熟信息素释放到培养基中(参照图 9)。

因此根据本发明合适的其他穿梭肽构建体可来自编码与 M 因子信息素功能性连接的 M 因子信号肽的根据 SEQ ID NO: 42、45 或 48 的编码序列。对应编码穿梭肽序列的非限制性实例包括, 例如, 根据 SEQ ID NO: 42 的 1 至 117 位核苷酸残基; 根据 SEQ ID NO: 45 的 1 至 123 位核苷酸残基; 或根据 SEQ ID NO: 48 的 1 至 114 位核苷酸残基; 或从它们得到的功能等同构建体, 其控制 M 因子信息素的分泌及加工和与所述信息素 C 末端且可蛋白酶解除去地连接的同源或异源靶蛋白的分泌及加工。功能等同物可包含未改变或经修饰(例如通过缺失单个或多个核酸残基)形式的位于成熟 M 因子(SEQ ID NO: 44、47 或 50)编码序列的 5'上游的序列片段, 且因此编码穿梭肽, 其具有改变的氨基酸序列并例如将成熟 M 因子肽序列与例如 C 末端截短的信号序列片段功能性连接。

只要根据本发明表达的靶蛋白(Targ)可通过宿主细胞以根据本发明的方式作为具有穿梭肽(SP)的融合蛋白部分进行表达、分泌及加工, 所述靶蛋白就可来自任意原核或真核生物, 特别是人、动物或酵母。经分泌及加工的产物可具有治疗用途或可具有其它有益的可应用性质。可在这里谈及的有治疗用途的蛋白质的实例为免疫球蛋白、肽激素、生长因子、淋巴因子、蛋白酶抑制剂等等。可在这里谈及的具有令人感兴趣应用的其它性质的靶蛋白的实例尤其是疏水蛋白。

在本发明特别优选的实施方案中, 靶蛋白为疏水蛋白, 特别是I类疏水蛋白。

典型疏水蛋白为具有 8 个半胱氨酸的保守基序(X<sub>2-38</sub>-C-X<sub>5-9</sub>-C-C-C<sub>11-39</sub>-C-X<sub>8-23</sub>-C-X<sub>5-9</sub>-C-C-X<sub>6-18</sub>-C-X<sub>2-13</sub>)的相对小(100 ± 25 个氨基酸)的中等疏水蛋白质。疏水蛋白可在亲水-疏水界面装配以产生蛋白质膜。此种I类疏水蛋白的团聚体在 SDS 中是不可溶的, 而II类疏水蛋白团聚体在 SDS 中是可溶的(Wessels, J. G. H. (1997) *Hydrophobins: Proteins that change the nature of the fungal surface. Adv Microb Physiol* 38:1-45)。

根据本发明可应用的疏水蛋白特别来自真菌，例如来自子囊菌如曲霉属，特别是构巢曲霉的子囊菌。

有用的疏水蛋白从上述现有技术中也是已知的且不限于来自真菌的疏水蛋白。

有用的疏水蛋白的非限制性实例选自 SEQ ID NO: 14(DewA)、SEQ ID NO: 19(RdIA)、SEQ ID NO: 20(RdIB)、SEQ ID NO: 21(HYP1)、SEQ ID NO: 22(HYP4)及 SEQ ID NO: 56(RodA)。

p52750 (DewA)

MRFIVSLLAF TAAATATALP ASAAKNAKLA TSAAFKQAE GTTCNVGSIA  
CCNSPAETNN DSLLSGLLGA GLLNGLSGNT GSACAKASLI DQLGLLALVD  
HTEEGPVCKN IVACCPEGTT  
NCVAVDNAGA GTKAE

q9l190 (RdIA)

MLKKAMVAAA AAASVIGMSA AAPQALAIG DDNGPAVANG NGAESAFGNS  
ATKGDMSPLSLVEGTLNKP CLGVEDVNVA VINLVPIQDI NVLADDLNQQ  
CADNSTQAKR DGALSHVLED LSVLSANGEG R  
q934f8 (RdIB)

MIKKVVAYAA IAASVMGASA AAPQAMAIG DDSGPVSANG NGASQYFGNS  
MTTGNMSPQM ALIQGSFNKP CIAVSDIPVS VIGLVPIQDL NVLGDDMNQQ  
CAENSTQAKR DGALAHLLLED  
VSILSSNGEG GKG

HYP1\_AGABI (P49072)

MISRVLVAAL VALPALVTAT PAPGKPKASS QCDVGEIHCC DTQQTPDHTS  
AAASGLLGVP INLGAFLGFD CTPISVLGVG GNNCAAQPVC CTGNQFTALI  
NALDCSPVNV NL

HYP4\_AGABI (O43122)

MVSTFITVAK TLLVALLFVN INIVVGTATT GKHCSSTGPIE CCKQVMDSKS  
PQATELLTKN GLGLGVLAVG KGLVGANCSPTAIGIGSGS QCSGQTVCCQ  
NNNFNGVVAI CTPINANV

RodA

LPPAHDSQFA GNGVGNKGNV NVKFPVPEV TVKQASDKCG DQAQLSCCNK  
ATYAGDTTTV DEGLLSGALS GLIGAGSGAE GLGLFDQCSK LDVAVLIGIQ  
DLVNQKCKQN IACCQNSPSS ADGNLIGVGL PCVALGSIL

**RodA 蛋白质与 DewA 蛋白质一起为构巢曲霉的外层孢子囊的部分。**

本发明还涉及表达载体，所述表达载体包含与至少一个调节核酸序列有效连接的如以上定义的表达构建体。

本发明还涉及重组微生物，其包含如以上定义的至少一个表达载体或表达构建体，其适当时稳定整合到宿主基因组。

根据本发明的“重组微生物”包含至少一个本发明的表达载体或本发明的表达构建体且在最广义上来自酵母。所述酵母特别来自子囊菌。优选的酵母选自原子囊菌纲裂殖酵母目，且特别优选选自裂殖酵母属，如粟酒裂殖酵母。

本发明还涉及可通过酵母细胞加工且来自酵母信息素前蛋白质的穿梭肽构建体，信息素可从所述前蛋白质通过 N 及 C 末端加工得到并且是可分泌的。

优选包含与 C 末端加工的信息素多肽 N 末端可加工地连接的信号多肽的穿梭肽构建体。

所述信号多肽优选为信息素前蛋白质的可蛋白酶解除去的天然信号多肽，且 C 末端加工的信息素多肽包含 C 末端蛋白酶切割位点。

优选的穿梭肽构建体来自酵母的信息素前蛋白质，特别是粟酒裂殖酵母 P 因子及 M 因子的前蛋白质。特别优选的穿梭肽包含根据 SEQ ID NO: 2 的氨基酸序列或其功能等同物。

本发明还涉及重组制备靶蛋白的方法，所述方法包括培养如以上定义的重组微生物，表达编码所述靶蛋白的核酸序列及分离分泌到培养基中的靶蛋白，例如如以上定义的疏水蛋白。

本发明还涉及编码如以上定义的穿梭肽构建体的核酸；还涉及编码如以上定义的融合蛋白的核酸，所述融合蛋白可通过酵母细胞加工且包含靶蛋白。

本发明还涉及通过本发明方法可获得的疏水蛋白。

最后，本发明涉及此类疏水蛋白用于表面处理的用途，其包含处理具体选自玻璃、纤维、织品、皮革、上漆物品，例如机动车辆、膜、正面(façade)物体的表面。

本发明还涉及疏水蛋白的用途，其用于处理纤维、织品及皮革表面。

### c) 发明的其它实施方案

#### c1) 多肽/蛋白质

本发明也包含特别公开或应用的多肽/蛋白质的“功能等同物”。这用于立即产生的融合蛋白及其组分，即靶蛋白(Targ)、穿梭肽(SP)，如信息素(Pher)，也用于信号肽(Sig)及接头。用于多肽/蛋白质的一般术语下文将只为“多肽”。

在本发明范围内，特别公开多肽的“功能等同物”或类似物为与其不同但是具有目的生物学活性的多肽。类似穿梭肽将还适合用于控制靶蛋白的分泌及加工。对应地，穿梭肽组分的功能等同物，如信号多肽、信息素、接头也欲还具有融合蛋白有效分泌及加工与靶蛋白释放所需的性质。

本发明多肽的“功能等同物”，如靶蛋白、穿梭肽在C和/或N末端可包含特别是自蛋白酶解切割产生的天然接头或衔接头序列的剩余物。

根据本发明，“功能等同物”特别表示突变蛋白质，所述突变蛋白质在上述特定序列的至少一个序列位置具有不同于特别谈及的氨基酸的至少一个氨基酸，但仍然具有一种上述生物活性。因此“功能等同物”包含通过一个或多个氨基酸添加、替代(参照下表中的实例)、缺失和/或倒位可获得的突变蛋白质，只要它们产生具有本发明性质谱的突变蛋白质，在任意

序列位置发生所述修饰是可能的。

适合用于氨基酸替代的残基的实例为：

最初残基	替代实例
Ala	Ser
Arg	Lys
Asn	Gln; His
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Pro
His	Asn; Gln
Ile	Leu; Val
Leu	Ile; Val
Lys	Arg; Gln; Glu
Met	Leu; Ile
Phe	Met; Leu; Tyr
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp; Phe
Val	Ile; Leu

当突变体与未改变多肽具有质量上匹配的活性图式时也特别是功能等

同的。这表示,例如,经修饰穿梭肽在相同宿主中以较高或较低效率表达或分泌相同靶蛋白;或经修饰靶蛋白具有升高或降低的药理学作用或经修饰的可应用的性质。

“功能等同物”在以上意义中也包含所描述的多肽前体和所述多肽的功能衍生物及盐。表达“盐”表示本发明蛋白质分子的羧基的盐及氨基的酸加成盐。羧基盐可以本身已知的方法制备且包含无机盐例如,钠、钙、铵、铁及锌盐且也包括与有机碱例如,胺,如三乙醇胺、精氨酸、赖氨酸、哌啶等等形成的盐。本发明同样涉及酸加成盐,例如与无机酸如盐酸或硫酸形成的盐及与有机酸如乙酸及草酸形成的盐。

本发明多肽的“功能衍生物”可同样在已知技术的帮助下由功能性氨基酸侧基或它们的N或C末端制备。此类衍生物包含,例如,羧酸基团的脂族酯、与铵或伯胺或仲胺反应可获得的羧酸基团的酰胺;通过与酰基反应制备的游离氨基的N-酰基衍生物;或通过与酰基反应制备的游离羧基的O-酰基衍生物。

当然,“功能等同物”也包含从与特别谈及的生物不同的生物可获得的多肽和天然存在变体。例如,可能通过同源序列区的序列比较区建立并根据本发明的特定指导确定等同酶。

“功能等同物”也包含片段,优选例如具有目的生物学功能的本发明多肽的个别结构域或序列基序。

“功能等同物”也为融合蛋白,所述融合蛋白具有一个上述多肽序列或从其衍生的功能等同物和功能性N或C末端连接(即具有融合蛋白部分的相互可忽略的功能损害)与所述多肽序列功能不同的至少一个其他异源序列。此类异源序列的非限制性实例为例如信号肽、酶、免疫球蛋白、表面抗原、受体或受体配体。

本发明也包含的“功能等同物”为特别谈及的多肽的同系物。根据Pearson和Lipman, Proc. Natl. Acad. Sci. (USA) 85 (8), 1988, 2444-2448的算法计算,所述同系物与任意特别公开的序列具有至少60%,优选至少75%,特别是至少85%,例如90%、95%或99%同源性。本发明同源多肽

的同源性百分数特别表示基于文中特别描述的任意氨基酸序列总长度的氨基酸残基的同一性百分数。

对于可能的蛋白质糖基化,本发明的等同物包含去糖基化或糖基化形式的多肽及可通过改变糖基化模式获得的经修饰形式。

本发明蛋白质或多肽的同系物可通过诱变,例如通过蛋白质点突变或截短以本身已知的方式产生。

### c2)核酸序列

本发明的核酸序列,特别是编码任意以上多肽和它们的功能等同物的核酸序列包含单链及双链 DNA 及 RNA 序列,例如也包含 cDNA 及 mRNA。

文中谈及的所有核酸序列为天然来源或可以本身已知的方式通过核苷酸构件的化学合成,例如通过个体重叠的、互补核酸构件的片段缩合制备。

可以本身已知的方式,例如根据亚磷酰胺方法(Voet, Boet, 2<sup>nd</sup> edition, Wiley Press New York, 896-897 页)进行寡核苷酸的化学合成。利用 DNA 聚合酶 Klenow 片段及连接反应合成寡核苷酸的装配及缺口的填补及一般克隆方法在 Sambrook 等人(1989), *Molecular Cloning: A laboratory manual*, Cold Spring Harbor Laboratory Press 中描述。

本发明还涉及编码任意以上多肽及它们的功能等同物的核酸序列,所述核酸序列可以例如通过用人工核苷酸类似物获得。

本发明涉及编码本发明多肽或其生物学活性部分的分离的核酸分子且涉及适合用作例如用于鉴定或扩增本发明编码核酸的杂交探针或引物的核酸片段。

本发明的核酸分子还可包含编码基因区 3'和/或 5'端的非翻译序列。

“分离的”核酸分子自所述核酸的天然来源中存在的其它核酸分子除去且另外,通过重组技术制备时,基本上不含其它细胞材料或培养基或化学合成时不含化学前体或其它化学品。

可通过分子生物学标准技术及根据本发明提供的序列信息分离本发明的核酸分子。例如,可通过应用任意特别公开的全序列或其片段作为杂交探针以及标准杂交技术(例如,在 Sambrook, J., Fritsch, E. F.和 Maniatis, T.

**Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> edition, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989** 描述的)从合适 cDNA 文库分离 cDNA。此外,可能用基于所述序列产生的寡核苷酸引物,通过聚合酶链反应分离包含任意公开序列或其片段的核酸分子。可将以该方式扩增的核酸克隆到合适载体并通过 DNA 序列分析表征。

本发明进一步包含与特别描述的核苷酸序列或其片段互补的核酸分子。

所述核苷酸序列能产生可用于在其它细胞类型及生物中鉴定和/或克隆同源序列的探针及引物。此类探针或引物一般包含在严格条件下与本发明核酸序列的有义链或对应反义链的至少约 12 个,优选至少约 25 个,例如约 40、50 或 75 个保守核苷酸杂交的核苷酸序列区。

本发明的其他核酸序列来自特别公开序列且通过单个或多个核苷酸的添加、替代、插入或缺失而与其不同,但仍编码具有目的性质谱的多肽。

与特别谈及序列以及其天然存在变体例如剪接变体或等位变体比较,本发明还包括包含“沉默”突变或根据特定来源或宿主生物的密码子使用进行修饰的核酸序列。本发明还涉及可通过保守核苷酸替代(用相同电荷、大小、极性和/或溶解性的氨基酸置换所述氨基酸)可获得的序列。

本发明还涉及来自由于序列多态性从特别公开的核酸得到的分子。这些遗传多态性可由于自然变异存在于群体中个体之间。这些自然变异一般引起基因核苷酸序列中 1 至 5% 的变化。

本发明还包含与上述编码序列杂交或与其互补的核酸序列。筛选基因组或 cDNA 文库时可发现这些多核苷酸,且适当时,用合适引物通过 PCR 扩增且例如随后用合适探针分离这些多核苷酸。

能与多核苷酸“杂交”的性质表示在严格条件下,多核苷酸或寡核苷酸结合近乎互补序列的能力,而在这些条件下非互补配偶体之间不发生非特异结合。为此目的,序列应 70-100%,优选 90-100% 互补。能互相特异结合的互补序列的性质用于例如 RNA 印迹或 DNA 印迹技术或用于 PCR 或 RT-PCR 中的引物结合。一般长度为 30 或 30 个以上碱基对的寡核苷酸用

于此目的。例如在 RNA 印迹技术中, 严格条件表示于 50-70°C, 优选 60-65°C 用洗涤溶液, 例如含有 0.1% SDS 的 0.1% SSC 缓冲液(20X SSC: 3M NaCl, 0.3M 柠檬酸钠, pH7.0)洗脱非特异杂交的 cDNA 探针或寡核苷酸。如上述, 这表示只有高度互补的核酸保持互相结合。设定严格条件是本领域技术人员已知的, 且例如, 在 Ausubel 等人, *Current Protocols in Molecular Biology*, John Wiley & Sons, N. Y. (1989), 6.3.1-6.3.6 中描述。

### c3) 表达构建体及载体

本发明还涉及表达构建体, 其包含在调节核酸序列的遗传控制下的编码根据本发明的待表达多肽的核酸序列; 且涉及包含至少一个所述表达构建体的载体。

此种本发明的构建体优选包含特定编码序列 5' 上游启动子及 3' 下游终止子序列和适当时, 其他常见调节元件, 其在每种情况下有效连接编码序列。

“有效连接”表示启动子、编码序列、终止子及适当时, 其它调节元件的顺序排列, 从而每个调节元件可以根据需要完成在表达编码序列中的功能。

有效连接序列的实例为靶定序列以及增强子、多腺苷酸化信号等等。其它调节元件包含选择标记、扩增信号、复制起点等等。合适调节序列例如在 Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990) 中描述。

基因构建体可包含一个或多个拷贝的编码核酸序列。

可应用的启动子的实例为酵母启动子 ADC1、MF $\alpha$ 、AC、P-60、CYC1、GAPDH、nmt1、nmt41 及 nmt81。

可谈及的用于粟酒裂殖酵母的合适启动子的实例为 nmt1、nmt41、nmt81、adh1、fbp1、SV40 或 CaMV。进一步信息参见 (<http://pingu.salk.edu/~forsburg/vectors.htm#exp>)。启动子在它们的转录速率方面不同。选择取决于所希望的表达水平。这因此用于其它酵母。

合适的酵母启动子, 例如在此处特别作为参考的已公布的美国专利中

请 2003/0077831 中描述。

也可以应用诱导型启动子，例如，光诱导型启动子，尤其是温度诱导型启动子。

谈及的调节序列意在使得可能定向表达所述核酸序列及蛋白质。取决于宿主生物，这表示例如只在诱导后基因表达或过表达或它立即表达或过表达。

在这方面，调节序列或调节因子可以主动方式优选影响表达且因此增加或降低表达。因此，调节元件可通过应用强转录信号如启动子和/或“增强子”在转录水平有益地增强。然而，除此之外，也可能例如通过提高 mRNA 稳定性增强翻译。

通过合适启动子与合适编码核苷酸序列及与终止子信号或聚腺苷酸化信号融合制备表达盒。为此目的，应用一般重组及克隆技术，例如在 T. Maniatis, E. F. Fritsch 和 J. Sambrook, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982) 及在 T. J. Silhavy, M. L. Berman 和 L. W. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) 及在 Ausubel, F. M. 等人, *Current Protocols in Molecular Biology*, Greene Publishing Assoc. and Wiley Interscience(1987)中描述的技术。

通过将重组核酸构建体或基因构建体有益地插入到使得所述构建体在宿主中最佳表达可能的宿主特异载体，在合适宿主生物中所述表达重组核酸构建体或基因构建体。载体是本领域技术人员已知的且例如在“克隆载体”(Pouwels P. H. 等人, eds., Elsevier, Amsterdam-New York-Oxford, 1985)中找到。除质粒外，载体还表示本领域技术人员已知的其它载体，例如，噬菌体、病毒，如 SV40、CMV、杆状病毒及腺病毒、转座子、IS 元件、噬菌粒、粘粒及线性或环状 DNA。这些载体可在宿主生物中自主复制或随染色体复制。

可在这里特别谈及的根据本发明合适的表达载体的实例为适于酵母粟酒裂殖酵母的构建体（参见，例如

(<http://pingu.salk.edu/~forsburg/vectors.html#exp>).

其它实例为:

pART1(McLeod, M., Stein, M.和 Beach, D. (1987) The product of the *mei3+* gene expressed under control of the mating type locus, induces meiosis and sporulation in fission yeast. *EMBO J.* 6:729-736,

pCHY21 ( Hoffman, C. S.和 Winston, F. (1991). Glucose repression of transcription of the *Schizosaccharomyces pombe* *fbp1* gene occurs by a cAMP signaling pathway. *Genes Dev.* 5: 561-571 ),

REP1、REP3、REP4 ( Maundrell, K. (1990). *nmt1* of fission yeast: a highly transcribed gene completely repressed by thiamine. *J. Biol. Chem.* 265: 10857-10864 ),

REP41、REP42、PEP81、PEP82(Basi, G., Schmid, E 和 Maundrell, K. (1993). TATA box mutations in the *Schizosaccharomyces pombe* *nmt1* promoter affect transcription efficiency but not the transcription start point or thiamine repressibility. *Gene* 123: 131-136).

用于在酿酒酵母中表达的酵母表达载体为如 pYEpSec1 (Baldari 等人, (1987)*Embo J.* 6: 229-234)、pMFa ( kurjan 和 Herskowitz (1982) *Cell* 30: 933-943) 、 pJRY88(Schultz 等人 (1987)*Gene* 54: 113-123) 及 pYES2(Invitrogen Corporation, San Diego, CA)。合适用于在其它真菌如丝状真菌中的载体及构建载体的方法包括在 van den Hondel, C.A.M.J.J.&Punt, P. J. (1991) “Gene transfer systems and vector development for filamentous fungi, in: *Applied Molecular Genetics of Fungi*, J.F. Peberdy 等人, eds., pp. 1-28, Cambridge University Press: Cambridge 中详细描述载体及方法。

其它合适的表达系统在 Sambrook, J., Fritsch, E. F.和 Maniatis, T., *Molecular cloning: A Laboratory Manual*, 2<sup>nd</sup> edition, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989 的第 16 及 17 章中描述。

#### c4)重组微生物

在本发明载体的帮助下，可能制备例如用至少一个本发明载体转化且可用于产生本发明多肽的重组微生物。

有利地，将以上描述的本发明重组构建体导入到合适宿主系统中并表达。优选地，这包括用本领域技术人员已知的熟悉的克隆及转染方法，例如，共沉淀、原生质体融合、电穿孔、反转录病毒转染等等，以便使谈及的核酸在特定表达系统中表达。合适的系统例如在 *Current Protocols in Molecular Biology*, F. Ausubel 等人，编者，Wiley Interscience, New York 1997，或 Sambrook 等人，*Molecular cloning: A Laboratory Manual*, 2<sup>nd</sup> edition, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989 中描述。

合适宿主生物原则上为能使本发明的核酸、它们的等位变体、它们的功能等同物或衍生物表达的任意生物。优选的宿主生物为酵母。

用于将外源 DNA 导入到酵母细胞的方法是现有技术中已知的。所述导入可例如通过根据 Hinnen 等人(1978, *Proc. Natl. Acad. Sci., USA*, 75: 1919-1935)的原生质球转化进行。例如用于粟酒裂殖酵母的化学转化方法可参见 Alfa 等人(Alfa, C., Fantes, P., Hyams, J., McLeod, M.和 Warbrick, E. (1993) *Experiments with fission yeast*. Cold Spring Harbor Laboratory Press, New York)或用于酿酒酵母的化学转化方法可参见 Kaiser 等人(Kaiser, C., Michaelis, S.和 Mitchell, A. (1994) *Methods in Yeast Genetics*. Cold Spring Harbor Laboratory Press, New York)。

在酵母中常常应用营养缺陷型标记物选择转化体。在这种情况下，待转化的菌株缺乏产生特定代谢产物所需的蛋白质。通过应用的载体将对应的活性蛋白质导入到细胞中。常常应用的标记物为例如尿嘧啶、亮氨酸、组氨酸或色氨酸生物合成的基因。

可通过载体中或在表达盒中也包括的标记基因选择成功转化的生物。此类标记基因的实例为抗生素抗性基因及酶的基因，所述酶的基因催化产色反应，其导致所转化的细胞染色。然后可通过自动细胞分选选择染色的

转化细胞。可通过包含对应抗生素的培养基或营养床(nutrient bed)筛选用载体成功转化的且携带适当抗生素抗性基因(例如 G418 或潮霉素)的微生物。可通过亲和层析将暴露于细胞表面的标记物蛋白质用于筛选。

宿主生物与用于所述生物的适当载体如质粒、病毒或噬菌体, 例如包含 RNA 聚合酶/启动子系统的质粒、噬菌体 8 或其它温和噬菌体或转座子和/或其它有益调节序列的组合组成表达系统。

#### c5) 靶蛋白的重组制备

本发明还涉及用于重组制备如以上定义的靶蛋白的方法。

根据已知方法可培养及发酵重组微生物。例如对于粟酒裂殖酵母, 合适的培养条件在 Alfa 等人(Alfa, C., Fantes, P., Hyams, J., McLeod, M.和 Warbrick, E. (1993) Experiments with fission yeast. Cold Spring Harbor Laboratory Press, New York)及 Gutz 等人(Gutz, H., Heslot, H., Leupold, U.和 Loprieno, U. (1974) Schizosaccharomyces pombe. In: Handbook of Genetics 1, pp 395-446, Plenum Press New York)中详细描述, 或对于酿酒酵母的合适培养条件在 Kaiser 等人(Kaiser, C., Michaelis, S.和 Mitchell, A. (1994) Methods in Yeast Genetics. Cold Spring Harbor Laboratory Press, New York) 中详细描述。

如果靶蛋白分泌到培养物上清液中, 根据已知蛋白质分离方法, 自上清液中除去细胞并自上清液中获得靶蛋白。

用已知的层析方法如分子筛层析(凝胶过滤)、离子交换层析如 Q-Sepharose 层析及疏水层析及用其它常规方法如超滤、结晶、盐析、透析及非变性凝胶电泳可对靶蛋白进行纯化。合适的方法例如在 Cooper, F. G., Biochemische Arbeitsmethoden [最初标题: The Tools of Biochemistry], Verlag Walter de Gruyter, Berlin, New York 或在 Scopes, R., Protein Purification, Springer Verlag, New York, Heidelberg, Berlin 中描述。

通过用简化纯化的编码经改变的多肽或融合蛋白的载体系统也有助于重组蛋白质的分离。此类合适修饰的实例为作为锚的“标记物”, 例如已知作为六组氨酸锚或可通过抗体作为抗原识别的表位的修饰(例如, 在

Harlow, E.和 Lane, D. 1988, *Antibodies: A Laboratory Manual*. Cold Spring Harbor (N.Y.) Press 中描述)。这些锚可用于将蛋白质附着到固相支持体, 例如, 可装入层析柱或可在微量滴定板或另一支持体应用上的聚合物基质。

这些锚也可同时用于识别蛋白质。此外, 通过用常规标记物如荧光染料、与底物反应后形成可检测反应产物的酶标记物或放射性标记, 单独或联合用于衍生蛋白质的所述锚可识别所述蛋白质。

#### c6)用疏水蛋白表面处理

用疏水蛋白处理表面以改变例如疏水或亲水表面性质原则上是已知的, 且现在通过本发明极大简化, 本发明的疏水蛋白重组制备提供充足的起始材料。

考虑到现有技术的教导(例如 WO-A-01/57066, 其描述由于亚硫酸盐处理对疏水蛋白的稳定化、溶解及其相关改进应用; 或 WO-A-01/57076, 其描述通过吸附到特氟隆珠且于低温通过去污剂如吐温洗脱纯化疏水蛋白; 或 WO-A-01/57528, 其描述通过应用吐温及达 85 摄氏度的温度将疏水蛋白固定到表面), 非常多的固体材料如玻璃、纤维、织物、皮革、上漆物品、膜、正面可用疏水蛋白涂布。

以下非限制性实例参考附图更详细描述了本发明, 所述附图中:

图 1 描述了根据本发明制备的用于自粟酒裂殖酵母分泌疏水蛋白的多种构建体。

图 2A)描述了 DewA 基因(SEQ ID NO: 39)的基因组序列; 两个内含子的序列用下划线标出; B)描述了氨基酸序列且圆括号中为构巢曲霉 DewA 蛋白质对应的 DNA 序列; 信号序列用粗体印刷且信号序列后的部分序列对应于成熟 DewA 序列; C)描述了氨基酸序列且圆括号中为 HA-Tag 的对应 DNA 序列。

图 3A)描述了氨基酸序列且圆括号中为 P 因子前蛋白质的对应 DNA 序列; 信号序列用粗体印刷; 信号序列后用下划线标出的部分序列对应于四个成熟信息素肽的序列; 邻近信号肽的信息素称作 P 因子; B)描述了氨基

酸序列且在圆括号中为可除去信号肽的对应 DNA 序列, 且其下游为 P 因子前蛋白质的 6 个氨基酸(用下划线标出); C)描述了氨基酸序列且在圆括号中为本发明“P 穿梭”的对应 DNA 序列; 信号序列用粗体印刷, 信号序列后用下划线标出的部分序列对应于成熟 P 因子的序列。

图 4 描述了本发明的融合蛋白, 其包含“P 穿梭”序列(粗体标出信号序列; 下划线标出成熟 P 蛋白质)、成熟 DewA(双下划线标出)及 C 末端融合的 HA 标记(SEQ ID NO: 18; 通过 SEQ IN NO: 17 编码)。

图 5 描述了粟酒裂殖酵母中疏水蛋白的免疫学检测。对于免疫学检测, 将构巢曲霉疏水蛋白基因 DewA 及 RodA 与 HA 标记融合, 克隆到 pJR1-3XL 表达载体并转化到粟酒裂殖酵母中。“膜级分”及“胞质蛋白质”通过 SDS-PAGE 分级分离。用 HA 抗体进行 Western 分析检测。大小标准(KDa)在左边标明。A: 应用包含无插入物载体(pJR1-3XL, 阴性对照)、HA 标记的对照蛋白质(阳性对照)及包含具有内含子的 HA 标记的 DewA 基因的载体(DewA-HA(+ 内含子)的培养物的样品。B: 每种情况下应用包含载体(pJR1-3XL, 阴性对照), 包含无内含子的 HA 标记 DewA 基因的载体(DewA-HA(- 内含子)或包含有内含子的 HA 标记的 RodA 基因(RodA-HA(+ 内含子)的载体的培养样品。

图 6 描述了在粟酒裂殖酵母中疏水蛋白表达的免疫学检测。PDewAHA 蛋白质在粟酒裂殖酵母中表达。收获细胞, 将培养物上清液分为等分试样且一份用 TCA 沉淀。SDA-PAGE 及在 HA 抗体帮助下的蛋白质印迹之后对蛋白质进行检测。通过\*标明的带对应于前体蛋白质(大约 18 kD, 上面的带)及成熟形式(大约 17 kD, 下面的带)。

图 7 描述了在粟酒裂殖酵母中疏水蛋白表达的检测。粟酒裂殖酵母细胞用通过强启动子(pJR1-3XL)或弱启动子(pJR1-81XL)表达 P + 6DewA 的质粒转化。细胞染色体上携带含有 c-myc 标记的 prp1 基因形式, c-myc 标记作为对照以排除裂解细胞对培养物上清液的污染。收获细胞(沉淀物), 培养物上清液用 TCA 沉淀(上清液)。SDA-PAGE 及在针对 HA(A)或针对 c-myc(B)抗体帮助下进行蛋白质印迹之后对蛋白质进行检测。

图 8 描述了在“P 穿梭”方法帮助下分泌的检测。粟酒裂殖酵母细胞用通过弱启动子(pJR1-81XL)表达 PfakDewA 的质粒转化。收获细胞(沉淀物)且培养物上清液用 TCA 沉淀(SN)。SDA-PAGE 及在针对 HA 抗体帮助下进行蛋白质印迹之后对蛋白质进行检测。

图 9 描述了每种情况下编码粟酒裂殖酵母 M 因子(对于成熟因子为 SEQ ID NO: 51)的三种基因: A)描述了 mfm1<sup>+</sup>基因的序列; B)描述了 mfm2<sup>+</sup>基因的序列; C)描述了 mfm3<sup>+</sup>基因的序列。

图 10 描述了 RodA 基因。RodA 基因的基因组序列(SEQ ID NO: 52)包含在对应编码 ORF(SEQ ID NO: 53)中不存在的两个内含子(下划线标出)。前蛋白质(SEQ ID NO: 54)包含在成熟蛋白质(SEQ ID NO: 56; 通过 SEQ IN NO: 55 编码)中不存在的可除去的信号序列(用粗体印刷)。

## 实验部分

### 一般实验详情

#### a) 一般克隆方法

除非另外注明,用于本发明目的进行的克隆步骤,例如,限制性切割、琼脂糖凝胶电泳、DNA 片段的纯化、核酸转移到硝酸纤维素膜及尼龙膜、DNA 片段的连接、大肠杆菌(*E. coli*)细胞的转化、细菌培养、噬菌体增殖及重组 DNA 的序列分析如通过在 Sambrook 等人(1989)上述引文中描述的进行。

根据生产商信息,通过 NucleoSpin 提取试剂盒(Extract Kit)(Machery-Nagel, Düren, 德国)自反应混合物或在凝胶电泳后纯化 DNA 且在 NucleoSpin 质粒快速纯化试剂盒(Plasmid Quick Pure Kit)(Machery-Nagel, Düren, 德国)帮助下自大肠杆菌分离质粒 DNA。

根据生产商信息应用限制酶(Invitrogen)。根据生产商信息在 T4 连接酶(Promega, Mannheim, 德国)帮助下进行 DNA 连接。

根据生产商信息,用 Gene Pulser II 仪器(BIO-RAD, Munich, 德国)及 2-mm 电穿孔小杯(Biozym Diagnostik, Hess. Oldendorf, 德国)通过电穿

孔进行大肠杆菌转化。在含有氨苄青霉素(150mg/l)的 LB 培养基(Lennox, 1955, *Virology* 1:190)上选择转化体。

b) 聚合酶链反应

根据生产商信息, 在 Combizyme DNA 聚合酶(Invitex, Berlin, 德国)帮助下进行 PCR 扩增。每种情况下每 100  $\mu$ l 反应体积用 1 pmol 的适当引物。

c) 培养

如在 Alfa 等人(Alfa, C., Fantes, P., Hyams, J., McLeod, M. 和 Warbrick, E. (1993) *Experiments with fission yeast*. Cold Spring Harbor Laboratory Press, New York)及 Gutz 等人(Gutz, H., Heslot, H., Leupold, U. 和 Loprieno, U. (1974) *Schizosaccharomyces pombe*. In: *Handbook of Genetics* 1, pp 395-446, Plenum Press, New York)中的描述培养粟酒裂殖酵母。

如在 Alfa 等人(Alfa, C., Fantes, P., Hyams, J., McLeod, M. 和 Warbrick, E. (1993) *Experiments with fission yeast*. Cold Spring Harbor Laboratory Press, New York)及 Gutz 等人(Gutz, H., Heslot, H., Leupold, U. 和 Loprieno, U. (1974) *Schizosaccharomyces pombe*. In: *Handbook of Genetics* 1, pp 395-446, Plenum Press, New York)中的描述培养重组菌株。

d) 细胞破碎

对于表达的快速控制, 通过离心(5 分钟, 3, 500  $\times$  g)除去细胞且细胞沉淀物直接用 Laemmli 缓冲液(Laemmli, U. K. (1970) *Cleavage of structural proteins during the assembly of the head of bacteriophage T4* (*Nature* 227: 680-685))吸收用于凝胶电泳。

对于细胞破碎, 通过于 3, 500  $\times$  g 离心 5 分钟收获的细胞。在 1ml 1  $\times$  PBS 中重新悬浮细胞沉淀物并加入 1 个体积的玻璃珠。将混合物涡旋 5 分钟, 并除去玻璃珠上的上清液。

e) 应用的生物

对于大肠杆菌工作应用菌株 DH5 $\alpha$ (Invitrogen) 、

**XL10-Gold(Stratagene)或 BL21(BioLabs)。**

粟酒裂殖酵母菌株来自德国 Institute for Genetics of Technische Universität Dresden 的 G. Rödel 博士教授小组的裂殖酵母收藏中心。

**实施例 1: DewA 及 DewAHA 表达构建体的制备及克隆到载体 pJR1-3XL 中**

**a) DewA 基因的基因组 DNA 序列的分离及内含子的除去**

如在 Kaiser 等人(Kaiser, C., Michaelis, S.和 Mitchell, A. (1994) *Methods in Yeast Genetics*. Cold Spring Harbor Laboratory Press, New York)中分离的构巢曲霉染色体 DNA 由 A. Brakhage 博士教授(Hanover, 德国)友情提供。

用染色体 DNA 作为模板和引物 SpDewBamrev 及 ScDewBamfor 对大约 550 bp 基因组 DNA 片段进行 PCR 扩增。

**ScDewBamfor:**

**5'-TAA TAA GGA TCC ATG CGC TTC ATC GTC TCT CTC C-3'**  
(SEQ ID NO: 41)

**SpDewBamrev:**

**5'-TAA TAA GGA TCC TTA CTC AGC CTT GGT ACC GGC-3'**  
(SEQ ID NO: 28)

反应混合物通过凝胶电泳分级分离且如以上描述洗脱对应的 DNA 带。侧翼为通过引物导入的 BamHI切割位点的片段根据生产商信息用限制性内切酶 BamHI(Invitrogen)切割并自反应混合物纯化(参见上文)。

载体 pUC18(Yanisch-Pron, C., Vieira, J.和 Messing, J. (1985) *Improved M13 phage cloning vectors and host strains: Nucleotide sequences of M13mp18 and pUC19 vectors*. *Gene* 33: 103)同样用 BamHI切割, 通过凝胶电泳分级分离且随后自凝胶洗脱(参见上文)。

连接载体及片段(参见上文)且将连接混合物转化到大肠杆菌中。在质粒制备及随后的限制性消化后对重组质粒进行鉴定。克隆后, 所克隆 PCR

产物的正确 DNA 序列通过测序证实，且对于下文制备的所有构建体也通过测序证实。根据 Sanger 等人(Sanger. F., Nicklen, S.和 Coulson, A. R. (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74: 5463-5467.)进行测序反应。在“Thermo-Sequenase fluorescent labeled primer cycle sequencing kit with 7-deaza-dGTP” (Amersham Pharmacia Biotech, Freiburg, 德国)及 5'IRD800 标记引物(MWG Biotech AG, Ebersberg, 德国)帮助下进行测序反应。对产物分级分离并用自动测序系统 LI-COR 4000/4200(MWG Biotech AG, Ebersberg, 德国)对序列进行分析。

含有内含子、克隆到 pUC18 载体的 BamHI 切割位点中的基因组 DewA 基因的构建体称作 pDewAgen。

通过“重叠延伸 PCR”(OEP)(参考 Pogulis, R.j., Vallejo, A. N.和 Pease, L. R. In vitro recombination and mutagenesis by overlap extension PCR. Methods Mol Biol. 1996; 57: 167-76)除去基因组 DewA DNA 中存在的两个内含子(参见基因组 DewA 序列 SEQ ID NO:39)。

用 pDewAgen 构建体的 DNA 作为模板，在引物对 ScDewBamfor/DewI1rev、DewI1for/DewI2rev 及 DewI2for/SpDewBamrev 的帮助下，对 DewA 的可读框(ORF)的亚片段进行第一次 PCR 扩增。

**ScDewBamfor:**

5' - TAA TAA GGA TCC ATG CGC TTC ATC GTC TCT CTC C - 3' (SEQ ID NO:41)

**DewI1rev:**

5' - GT GTG GTC GAC GAG AGC GAG CAG ACC CAG CTG - 3' (SEQ ID NO:24)

**DewI1for:**

5' - CAG CTG GGT CTG CTC GCT CTC GTC GAC CAC AC - 3' (SEQ ID NO:23)

**DewI2rev:**

5' - GTC GAC GGC AAC ACA GTT GGT GGT TCC CTC - 3' (SEQ ID NO:26)

**DewI2for:**

5' - GAG GGA ACC ACC AAC TGT GTT GCC GTC GAC - 3' (SEQ ID NO:25)

**SpDewBamrev:**

5' - TAA TAA GGA TCC TTA CTC AGC CTT GGT ACC GGC - 3' (SEQ ID NO:28)

这些亚片段通过凝胶电泳分级分离并纯化。在最后 PCR 中，用所述亚片段作为模板及远侧引物 ScDewBamfor 及 ScDewBamrev 扩增无内含子的 ORF。大约 410 bp PCR 产物通过凝胶电泳分级分离、纯化且用限制性内切酶 BamHI 切割。通过远侧引物导入适当的切割位点。对切割片段进行纯化并克隆到同样用 BamHI 切割的 pUC18 载体。将载体及片段连接(参见上文)并将连接混合物转化到大肠杆菌中。质粒小量制备后，鉴定重组质粒且通过测序证实所克隆的 ORF 的正确序列。以这种方式获得的构建体 (pDewAORF) 作为模板用于对应表达质粒的构建。

**b) DewHA(+内含子)及 DewHA(-内含子)表达构建体的制备及 C 末端血凝素标记的导入**

由于不能获得针对 DewA 的特异抗体，通过 OEP 将 DewA 与 HA 表位融合以检测异源表达。首先，在初次 PCR 中，利用引物对 SpDewXhfor/DewAHArev 及 DewAHAfor/DewAHANcorev。

**SpDewXhofor:**

5' - TAA TAA CTC GAG ATG CGC TTC ATC GTC TCT CTC C - 3' (SEQ ID NO:27)

**DewAHArev:**

5' - TCC ACG CGG AAC CAG CTC AGC CTT GGT ACC - 3' (SEQ ID NO:30)

**DewAHAfor:**

5' - GGT ACC AAG GCT GAG CTG GTT CCG CGT GGA - 3' (SEQ ID NO:29)

**DewAHANcorev:**

5' - ATT ATT CCA TGG CTA TTA GCG GCC GCA CTG AGC AGC - 3' (SEQ ID NO:31)

pDewAgen 构建体的 DNA 作为模板用于制备 DewHA(+内含子), 且 pDewAORF 构建体的 DNA 作为模板用于制备 DewHA(-内含子)。携带 HA 标记 DNA 序列的载体 yEP351HA(Kettner, K., Friederichs, S., Schlapp, T. 和 Rödel G(2001)Expression of a VEGP-like protein from Parapoxvirus ovis in the yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. *Protein Expr Purif Aug*; 22(3): 479-83) 用作模板与 DewAHAfor/DewAHANcorev 用于 PCR。通过引物对 SpDewShofor/DewAHANcorev 各自亚片段在最终 PCR 中将编码 HA 表位的 DNA 与特定 DewA-DNA 融合。以这种方式扩增的片段 5'侧翼为 XhoI 限制性切割位点且 3'侧翼为 NcoI限制性切割位点, 用远侧引物导入所述 XhoI及 NcoI两个限制性切割位点。片段通过凝胶电泳分级分离、纯化且用限制性内切酶 XhoI及 NcoI切割并自反应混合物中纯化。

载体 pJR1-3XL(Moreno, M. B., Duran, A.和 Ribas,J. C. A family of multifunctional thiamine-repressible expression vectors for fission yeast. *Yeast*. 2000 Jun 30:16(9):861-72)同样用限制酶 XhoI及 NcoI切割, 通过凝胶电泳分级分离并纯化。将载体与片段连接(参见上文)且将连接混合物通过电穿孔转化到大肠杆菌中。质粒少量制备后, 鉴定重组质粒且所克隆 ORF 的正确序列通过测序证实。在以此种方式获得的 DewA-HA(+内含子)及 DewA-HA(-内含子)表达质粒中, 粟酒裂殖酵母中融合构建体的表达处于强 nmt1 启动子控制下。

c) DewA-HA(+内含子)及 DewA-HA(-内含子)的表达

如通过 Schiestl 及 Gietz(Schiestl, R. H.和 Gietz, R. D. (1989) High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr Genet* 16: 339-346)描述的, 将根据 a)及 b)获得的 DewA-HA(+内含子)及 DewA-HA(-内含子)载体转化到粟酒裂殖酵母宿主菌株 KO103( $h^{-s}ade6$ -M210leu1-32his7-366)中。通过 leu1-32 突变引起的粟酒裂殖酵母菌株的亮氨酸营养缺陷型通过存在于表达载体上的酿酒酵母 LEU2 基因互补。因此可能在没有亮氨酸的基本培养基上筛选转化体。通过蛋白质印迹分析研究在对应酵母转化体中融合蛋白的表达。

抗 HA 抗体(编号 1583816, 抗 HA(12CA5)小鼠单克隆抗体)及抗 c-myc 抗体(编号 1667149, 抗 c-myc 抗体)购自 Roche Diagnostics(Mannheim, 德国)。

培养后, 收获酵母转化体, 用玻璃珠破碎, 且除去  $3,500 \times g$  离心上清液。每种情况下应用  $50 \mu g$   $20,000 \times g$  沉淀物(“膜级分”)及  $50 \mu g$  上清液(“胞质蛋白质”)并在 SDS-PAGE 中分级分离。用 HA 抗体进行蛋白质印迹分析中的检测。图 5 描述了结果。大小标准(kDa)在左边标出。在图 5A 中, 应用含有无插入物的载体(pJR1-3XL, 阴性对照)、HA 标记的对照蛋白质(阳性对照)或包含有内含子的 HA 标记的 DewA 基因(DewA-HA(+内含子))的载体的培养物样品。在图 5B 中, 每种情况下应用含有载体(pJR1-3XL, 阴性对照)或包含无内含子的 HA 标记的 DewA 基因(DewA-HA(-内含子))或有内含子的 HA 标记的 RodA 基因(Rod-AHA(+内含子))的培养物样品。

类似地根据实施例 1a)及 1b)的信息制备 RodAHA(+内含子)。RodA 是来自构巢曲霉的另一疏水蛋白。

**实施例 2: 制备用于分泌包含 PDewAHA 构建体的表达的 DewA 载体的表达载体**

a) 包含 P 因子信号肽编码序列的 PDewAHA 构建体的制备

为了优化蛋白质从粟酒裂殖酵母细胞的分泌, 将在裂殖酵母中无效的构巢曲霉蛋白质的真正分泌信号首先用粟酒裂殖酵母 P 因子的可除去信

号肽替换。P 因子作为肽信息素通过细胞分泌到培养基中。它在细胞中作为前体蛋白质(前蛋白质)合成,所述前体蛋白质由可除去的 N 末端信号序列及四个拷贝的 P 因子组成,每种情况下通过短间隔序列分隔,且在分泌过程中,经历成熟,包括信号序列的除去及四个 P 因子肽的蛋白酶解释放。

P 因子信号序列首先通过 PCR 且以粟酒裂殖酵母基因组 DNA 作为模板,用引物对 SigPXhovor/PDewArev 扩增,并纯化对应的 PCR 产物。

**SigPXhovor:**

5' - TAA TTT CTC GAG ATG AAG ATC ACC GCT GTC ATT GCC CTT TTA TTC TCA  
C - 3' (SEQ ID NO:34)

**PDewArev:**

5' - GGC AGA GGC CGG GAG TGG AAT AGG TGA GGC - 3' (SEQ ID NO:33)

以引物对 PDewfor/DewAHANcorev 及 DewAHA(-内含子)构建体 DNA 作为模板得到的 PCR 产物同样通过凝胶电泳分级分离并纯化。

**PDewfor:**

5' - GCC TCA CCT ATT CCA CTC CCG GCC TCT GCC - 3' (SEQ ID NO:32)

**DewAHANcorev:**

5' - ATT ATT CCA TGG CTA TTA GCG GCC GCA CTG AGC AGC - 3' (SEQ ID  
NO:31)

在使用远侧引物 SigPXhovor/DewAHANcorev 的最终 PCR 中,这两种初次 PCR 产物用作模板。以这种方式扩增的 PDewAHA 片段 5' 侧翼为 XhoI 限制性切割位点且 3' 端为 NcoI 限制性切割位点,所述两个酶切位点均借助以上引物导入。在通过这个片段(PDewAHA)编码的融合蛋白中,构巢曲霉 DewA 的可除去的信号序列通过 P 因子前体蛋白质的可除去的信号肽替换。

PDewAHA 片段用限制性内切酶 XhoI 及 NcoI 切割,通过凝胶电泳分级分离并连接到用相同限制性内切酶切割的 pJR1-3XL 载体中(参见上文)。将载体及片段连接(参见上文)且将连接混合物通过电穿孔转化到大肠杆菌中。质粒小量制备后,鉴定重组质粒且通过测序证实所克隆的 ORF 的正确序列。获得的构建体称作 PDewAHA。

**b) 表达**

与实施例 1c) 类似进行实验。

**PDewAHA** 蛋白质在粟酒裂殖酵母中表达。收获细胞，将培养物上清液等分且将一份用 TCA 沉淀。TCA 沉淀物用 Laemmli 缓冲液吸收 (Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4 (Nature 227: 680-685))。研究细胞沉淀物、上清液及 TCA 沉淀的上清液。SDA-PAGE 及用 HA 抗体进行的蛋白质印迹后检测蛋白质。图 6 描述了结果。通过\*标明的带对应于前体蛋白质(大约 18kD, 上面的带)及成熟形式(大约 17kD, 下面的带)。

分析显示没有观察到有效分泌。可除去 P 因子信号肽的 N 末端融合不足够用于分泌融合的疏水蛋白。

**实施例 3: 制备用于分泌包含 P+6DewAHA 构建体的所表达的 DewA 载体的表达载体**

a) 包含 P 因子信号肽编码序列, C 末端延伸 6 个氨基酸的 P+6DewAHA 构建体的制备

为保证对分泌重要的信号肽的真正序列环境, 通过 OEP, 在初次 PCR 反应中使用引物对 **SigPXhofor/P+6DewArev** 及 **P+6DewAfor/DewAHANcorev** 与作为模板的 **PDewAHA** 构建体 DNA, 通过 6 个 C 末端连接的氨基酸(P+6DewA)延伸融合蛋白中所述信号肽的序列 **SigPXhofor:**

5' - TAA TTT CTC GAG ATG AAG ATC ACC GCT GTC ATT GCC CTT TTA TTC TCA  
C - 3' (SEQ ID NO:34)

**P+6DewArev:**

5' - CAC ACC AGG ATC GGC AAC TGG AAT AGG TGA GGC - 3' (SEQ ID NO:36)

**P+6DewAfor:**

5' - GTT GCC GAT CCT GGT GTG CTC CCG GCC TCT GCC - 3' (SEQ ID NO:35)

**DewAHANcorev:**

5' - ATT ATT CCA TGG CTA TTA GCG GCC GCA CTG AGC AGC - 3' (SEQ ID NO:31)

且在最终 PCR 反应中使用引物对 **SigPXhofor/DewAHANcorev**。

**P+6DewA** 片段用限制性内切酶 **XhoI**及 **NcoI**切割, 通过凝胶电泳分级

分离并连接到用相同限制性内切酶切割的 pJR1-3XL 载体中(参见上文), 且将连接混合物通过电穿孔转化到大肠杆菌中。质粒小量制备后, 鉴定重组质粒且通过测序证实克隆的 ORF 的正确序列。将 P+6DewA 片段也克隆到 pJR-81XL 载体中。这里, 融合基因的转录处于弱 nmt81 启动子的控制下。此构建体欲用于检测 pJR1-3XL 构建体中非常高的转录对分泌的负面影响。

对应的构建体称作 P+6DewA/pJR1-3XL 及 P+6DewA/pJR1-81XL。

与实施例 2a 类似进行实验。

与实施例 2a 类似进行 pJR1-3XL 中扩增序列的克隆。

#### b) 表达

与实施例 2b)类似进行表达。

用通过强启动子(pJR1-3XL)及弱启动子(pJR1-81XL)表达 P+6DewA 的两种质粒转化粟酒裂殖酵母细胞。细胞染色体上携带具有 c-myc 标记的 prp1 基因形式, 所述标记作为对照以排除裂解细胞对培养物上清液的污染。收获(沉淀)细胞且将培养物上清液用 TCA 沉淀(上清液)。沉淀物用 Laemmli 缓冲液吸收并同样分析。SDA-PAGE 及用针对 HA 的抗体(图 7A)及针对 c-myc 的抗体(Roche Diagnosites)(图 7B)进行蛋白质印迹后对蛋白质进行检测。

如图 7 示出的, 此构建体也不合适用于有效分泌。

**实施例 4: 制备用于分泌包含 PfakDewAHA 构建体的表达的 DewA 载体的表达载体**

a) 制备包含编码包括 P 因子信号肽的成熟的第一种 P 因子序列的 PfakDewAHA 构建体

DewAHA 通过 OEP 与成熟 P 因子序列的羧基末端融合。用引物对 SigPXhofor/PfakDewArev 及粟酒裂殖酵母基因组 DNA 作为模板

**SigPXhovor:**

5' - TAA TTT CTC GAG ATG AAG ATC ACC GCT GTC ATT GCC CTT TTA TTC TCA  
C - 3' (SEQ ID NO:34)

**PfakDewArev:**

5' - GGC AGA GGC CGG GAG GCG CTT TTT CAA GTT GGG TC - 3' (SEQ ID  
NO:38)

和引物对 **PfakDewAfor/DewAHANcorev** 及 **P+6DewA/pJR-81XL** 构建体的 DNA 作为模板

**PfakDewAfor:**

5' - AAC TTG AAA AAG CGC CTC CCG GCC TCT GCC - 3' (SEQ ID NO:37)

**DewAHANcorev:**

5' - ATT ATT CCA TGG CTA TTA GCG GCC GCA CTG AGC AGC - 3' (SEQ ID  
NO:31)

的初次 PCR 反应中获得的 PCR 片段通过凝胶电泳分离，纯化且用作模板用于借助引物对 **SigPXhovor/DewAHANcorev** 的最终 PCR。以这种方式获得的 **PfakDewA** 片段用限制性内切酶 **XhoI** 及 **NcoI** 切割，通过凝胶电泳分级分离并连接到用相同限制性内切酶切割的 **pJR1-81XL** 载体中(参见上文)。将连接混合物通过电穿孔转化到大肠杆菌中。质粒小量制备后，鉴定重组质粒且通过测序证实所克隆的 ORF 的正确序列。此构建体称作 **PfakDewA/pJR-81XL**。在此构建体中，编码包括第一个氨基端信息素的 P 因子前蛋白质并编码疏水蛋白的融合序列处于 **nmt81** 启动子的控制下。

与实施例 2a 类似进行实验，但应用 **pJR1-81XL** 表达载体。

与实施例 2a 类似将扩增序列克隆到 **pJR1-81XL** 中。为此，将用限制性内切酶 **XhoI** 及 **NcoI** 切割的扩增 DNA 克隆到 **pJR1-81XL** 表达载体的 **XhoI** 及 **NcoI** 切割位点中。

## b) 表达

与实施例 3a) 类似进行表达。

沉淀细胞且将培养物上清液用 TCA 沉淀。沉淀物用 Laemmli 缓冲液吸收并同样地分析。SDA-PAGE 及使用针对 HA 的抗体进行蛋白质印迹后检测蛋白质。图 8 描述了结果，其表明借助此构建体获得的疏水蛋白有效分泌到培养基中。因此，对应的融合蛋白包含在它们的真正背景中分泌所

需的所有 P 因子前蛋白质序列。P 因子自身可蛋白酶解除去。

#### 实施例 5: 所表达的疏水蛋白吸附到特氟隆的显微镜检查

利用荧光标记的 HA 抗体(Molecular Probes, 目录号 A-21287), 显微镜检查所表达的疏水蛋白向特氟隆的吸附。

培养根据实施例 1 至 4 中任意一个制备的经转化宿主细胞。分别收获细胞及适当时收获上清液。用不含疏水蛋白基因的对应载体转化和培养的细胞和对应的培养物上清液作为参考样品。

将细胞机械破碎(振动研磨机)。将特氟隆板在细胞破碎溶液中或在上清液中于室温孵育 18 小时, 且用水冲洗( $3 \times 10$  分钟)。然后将经处理的特氟隆在 PBS 中与荧光标记的抗体孵育, 接着再用 PBS 冲洗( $3 \times 15$  分钟)且在氮气喷射器中干燥。最后, 在荧光显微镜下进行评估。

对参考样品没有观察到荧光(结果没有显示), 但在细胞匀浆物或培养物上清液(此时疏水蛋白通过表达细胞分泌)中孵育后斑点状荧光是明显的。

<110> 巴斯福股份公司  
 <120> 蛋白质从酵母的分泌  
 <130> M/44234  
 <160> 56  
 <170> PatentIn 版本 3.1  
 <210> 1  
 <211> 171  
 <212> DNA  
 <213> 粟酒裂殖酵母 (Schizosaccharomyces pombe)  
 <220>  
 <221> CDS  
 <222> (1).. (171)  
 <223>

<400> 1

atg aag atc acc gct gtc att gcc ctt tta ttc tca ctt gct gct gcc 48  
 Met Lys Ile Thr Ala Val Ile Ala Leu Leu Phe Ser Leu Ala Ala Ala  
 1 5 10 15

tca cct att cca gtt gcc gat cct ggt gtg gtt tca gtt agc aag tca 96  
 Ser Pro Ile Pro Val Ala Asp Pro Gly Val Val Ser Val Ser Lys Ser  
 20 25 30

tat gct gat ttc ctt cgt gtt tac caa agt tgg aac act ttt gct aat 144  
 Tyr Ala Asp Phe Leu Arg Val Tyr Gln Ser Trp Asn Thr Phe Ala Asn  
 35 40 45

cct gat aga ccc aac ttg aaa aag cgc 171  
 Pro Asp Arg Pro Asn Leu Lys Lys Arg  
 50 55

<210> 2

<211> 57

<212> PRT

<213> 粟酒裂殖酵母

<400> 2

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

Ser Pro Ile Pro Val Ala Asp Pro Gly Val Val Ser Val Ser Lys Ser  
 20 25 30

Tyr Ala Asp Phe Leu Arg Val Tyr Gln Ser Trp Asn Thr Phe Ala Asn  
 35 40 45

Pro Asp Arg Pro Asn Leu Lys Lys Arg  
 50 55

<210> 3  
 <211> 60  
 <212> DNA  
 <213> 粟酒裂殖酵母  
 <220>  
 <221> CDS  
 <222> (1)..(60)  
 <223>

<220>  
 <221> sig\_peptide  
 <222> (1)..(60)  
 <223>

<400> 3  
 atg aag atc acc gct gtc att gcc ctt tta ttc tca ctt gct gct gcc 48  
 Met Lys Ile Thr Ala Val Ile Ala Leu Leu Phe Ser Leu Ala Ala Ala  
 1 5 10 15

tca cct att cca 60  
 Ser Pro Ile Pro  
 20

<210> 4  
 <211> 20

<212> PRT

<213> 粟酒裂殖酵母

<400> 4

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

Ser Pro Ile Pro  
                   20

<210> 5

<211> 81

<212> DNA

<213> 粟酒裂殖酵母

<220>

<221> CDS

<222> (1)..(81)

<223>

<400> 5

aag tca tat gct gat ttc ctt cgt gtt tac caa agt tgg aac act ttt            48  
 Lys Ser Tyr Ala Asp Phe Leu Arg Val Tyr Gln Ser Trp Asn Thr Phe  
 1                    5                    10                    15

gct aat cct gat aga ccc aac ttg aaa aag cgc                            81  
 Ala Asn Pro Asp Arg Pro Asn Leu Lys Lys Arg  
                   20                    25

<210> 6

<211> 27

<212> PRT

<213> 粟酒裂殖酵母

<400> 6

Lys Ser Tyr Ala Asp Phe Leu Arg Val Tyr Gln Ser Trp Asn Thr Phe  
 1                    5                    10                    15

Ala Asn Pro Asp Arg Pro Asn Leu Lys Lys Arg  
 20 25

<210> 7

<211> 78

<212> DNA

<213> 粟酒裂殖酵母

<220>

<221> CDS

<222> (1)..(78)

<223>

<220>

<221> sig\_peptide

<222> (1)..(60)

<223>

<400> 7

atg aag atc acc gct gtc att gcc ctt tta ttc tca ctt gct gct gcc 48  
 Met Lys Ile Thr Ala Val Ile Ala Leu Leu Phe Ser Leu Ala Ala Ala  
 1 5 10 15

tca cct att cca gtt gcc gat cct ggt gtg 78  
 Ser Pro Ile Pro Val Ala Asp Pro Gly Val  
 20 25

<210> 8

<211> 26

<212> PRT

<213> 粟酒裂殖酵母

<400> 8

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

Ser Pro Ile Pro Val Ala Asp Pro Gly Val  
 20 25

<210> 9

<211> 606

<212> DNA

<213> 粟酒裂殖酵母

<220>

<221> CDS

<222> (1)..(606)

<223>

<400> 9

atg aag atc acc gct gtc att gcc ctt tta ttc tca ctt gct gct gcc 48

Met Lys Ile Thr Ala Val Ile Ala Leu Leu Phe Ser Leu Ala Ala Ala

1 5 10 15

tca cct att cca gtt gcc gat cct ggt gtg gtt tca gtt agc aag tca 96

Ser Pro Ile Pro Val Ala Asp Pro Gly Val Val Ser Val Ser Lys Ser

20 25 30

tat gct gat ttc ctt cgt gtt tac caa agt tgg aac act ttt gct aat 144

Tyr Ala Asp Phe Leu Arg Val Tyr Gln Ser Trp Asn Thr Phe Ala Asn

35 40 45

cct gat aga ccc aac ttg aaa aag cgc gaa ttc gaa gct gct ccc gca 192

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

50 55 60

aaa act tat gct gat ttc ctt cgt gct tat caa agt tgg aac act ttt 240

Lys Thr Tyr Ala Asp Phe Leu Arg Ala Tyr Gln Ser Trp Asn Thr Phe

65 70 75 80

gtt aat cct gac aga ccc aat ttg aaa aag cgt gag ttt gaa gct gcc 288

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

85 90 95

cca gag aag agt tat gct gat ttc ctt cgt gct tac cat agt tgg aac 336

Pro Glu Lys Ser Tyr Ala Asp Phe Leu Arg Ala Tyr His Ser Trp Asn

100 105 110

act ttt gtt aat cct gac aga ccc aac ttg aaa aag cgc gaa ttc gaa 384

Thr Phe Val Asn Pro Asp Arg Pro Asn Leu Lys Lys Arg Glu Phe Glu

115	120	125	
gct gct ccc gca aaa act tat gct gat ttc ctt cgt gct tac caa agt			432
Ala Ala Pro Ala Lys Thr Tyr Ala Asp Phe Leu Arg Ala Tyr Gln Ser			
130	135	140	
tgg aac act ttt gtt aat cct gac aga ccc aac ttg aaa aag cgc act			480
Trp Asn Thr Phe Val Asn Pro Asp Arg Pro Asn Leu Lys Lys Arg Thr			
145	150	155	160
gaa gaa gat gaa gag aat gag gaa gag gat gaa gaa tac tat cgc ttt			528
Glu Glu Asp Glu Glu Asn Glu Glu Glu Asp Glu Glu Tyr Tyr Arg Phe			
165	170	175	
ctt cag ttt tat atc atg act gtc cca gag aat tcc act att aca gat			576
Leu Gln Phe Tyr Ile Met Thr Val Pro Glu Asn Ser Thr Ile Thr Asp			
180	185	190	
gtc aat att act gcc aaa ttt gag agc taa			606
Val Asn Ile Thr Ala Lys Phe Glu Ser			
195	200		

<210> 10  
 <211> 201  
 <212> PRT  
 <213> 粟酒裂殖酵母  
 <400> 10

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

Ser Pro Ile Pro Val Ala Asp Pro Gly Val Val Ser Val Ser Lys Ser  
                   20                    25                    30

Tyr Ala Asp Phe Leu Arg Val Tyr Gln Ser Trp Asn Thr Phe Ala Asn  
                   35                    40                    45

Pro Asp Arg Pro Asn Leu Lys Lys Arg Glu Phe Glu Ala Ala Pro Ala  
                   50                    55                    60



<223> 待完成

<220>

<221> CDS

<222> (1)..(156)

<223>

<400> 11

ctg gtt ccg cgt gga tcc atc gaa ggt cgt ggc ggc cgc atc ttt tac 48  
 Leu Val Pro Arg Gly Ser Ile Glu Gly Arg Gly Gly Arg Ile Phe Tyr  
 1 5 10 15

cca tac gat gtt cct gac tat gcg ggc tat ccc tat gac gtc ccg gac 96  
 Pro Tyr Asp Val Pro Asp Tyr Ala Gly Tyr Pro Tyr Asp Val Pro Asp  
 20 25 30

tat gca gga tcc tat cca tat gac gtt cca gat tac gct gct cag tgc 144  
 Tyr Ala Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ala Gln Cys  
 35 40 45

ggc cgc taa tag 156  
 Gly Arg  
 50

<210> 12

<211> 50

<212> PRT

<213> 未知

<220>

<223> 待完成

<400> 12

Leu Val Pro Arg Gly Ser Ile Glu Gly Arg Gly Gly Arg Ile Phe Tyr  
 1 5 10 15

Pro Tyr Asp Val Pro Asp Tyr Ala Gly Tyr Pro Tyr Asp Val Pro Asp  
 20 25 30

Tyr Ala Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ala Gln Cys  
 35 40 45

Gly Arg  
 50

<210> 13  
 <211> 354  
 <212> DNA  
 <213> 构巢曲霉 (Aspergillus nidulans)  
 <220>  
 <221> CDS  
 <222> (1)..(354)  
 <223>

<400> 13  
 ctc ccg gcc tct gcc gca aag aac gcg aag ctg gcc acc tcg gcg gcc 48  
 Leu Pro Ala Ser Ala Ala Lys Asn Ala Lys Leu Ala Thr Ser Ala Ala  
 1 5 10 15  
 ttc gcc aag cag gct gaa ggc acc acc tgc aat gtc ggc tcg atc gct 96  
 Phe Ala Lys Gln Ala Glu Gly Thr Thr Cys Asn Val Gly Ser Ile Ala  
 20 25 30  
 tgc tgc aac tcc ccc gct gag acc aac aac gac agt ctg ttg agc ggt 144  
 Cys Cys Asn Ser Pro Ala Glu Thr Asn Asn Asp Ser Leu Leu Ser Gly  
 35 40 45  
 ctg ctc ggt gct ggc ctt ctc aac ggg ctc tcg ggc aac act ggc agc 192  
 Leu Leu Gly Ala Gly Leu Leu Asn Gly Leu Ser Gly Asn Thr Gly Ser  
 50 55 60  
 gcc tgc gcc aag gcg agc ttg att gac cag ctg ggt ctg ctc gct ctc 240  
 Ala Cys Ala Lys Ala Ser Leu Ile Asp Gln Leu Gly Leu Leu Ala Leu  
 65 70 75 80  
 gtc gac cac act gag gaa ggc ccc gtc tgc aag aac atc gtc gct tgc 288  
 Val Asp His Thr Glu Glu Gly Pro Val Cys Lys Asn Ile Val Ala Cys

---

	85	90	95	
tgc cct gag gga acc acc aac tgt gtt gcc gtc gac aac gct ggc gcc				336
Cys Pro Glu Gly Thr Thr Asn Cys Val Ala Val Asp Asn Ala Gly Ala				
	100	105	110	

ggt acc aag gct gag taa				354
Gly Thr Lys Ala Glu				
	115			

<210> 14  
 <211> 117  
 <212> PRT  
 <213> 构巢曲霉  
 <400> 14

Leu Pro Ala Ser Ala Ala Lys Asn Ala Lys Leu Ala Thr Ser Ala Ala			
1	5	10	15

Phe Ala Lys Gln Ala Glu Gly Thr Thr Cys Asn Val Gly Ser Ile Ala			
	20	25	30

Cys Cys Asn Ser Pro Ala Glu Thr Asn Asn Asp Ser Leu Leu Ser Gly			
	35	40	45

Leu Leu Gly Ala Gly Leu Leu Asn Gly Leu Ser Gly Asn Thr Gly Ser			
	50	55	60

Ala Cys Ala Lys Ala Ser Leu Ile Asp Gln Leu Gly Leu Leu Ala Leu			
65	70	75	80

Val Asp His Thr Glu Glu Gly Pro Val Cys Lys Asn Ile Val Ala Cys			
	85	90	95

Cys Pro Glu Gly Thr Thr Asn Cys Val Ala Val Asp Asn Ala Gly Ala			
	100	105	110

Gly Thr Lys Ala Glu  
115

<210> 15  
<211> 408  
<212> DNA  
<213> 构巢曲霉  
<220>  
<221> CDS  
<222> (1)..(408)  
<223>

<400> 15  
atg cgc ttc atc gtc tct ctc ctc gcc ttc act gcc gcg gcc acc gca 48  
Met Arg Phe Ile Val Ser Leu Leu Ala Phe Thr Ala Ala Ala Thr Ala  
1 5 10 15  
  
acc gcc ctc ccg gcc tct gcc gca aag aac gcg aag ctg gcc acc tcg 96  
Thr Ala Leu Pro Ala Ser Ala Ala Lys Asn Ala Lys Leu Ala Thr Ser  
20 25 30  
  
gcg gcc ttc gcc aag cag gct gaa ggc acc acc tgc aat gtc ggc tcg 144  
Ala Ala Phe Ala Lys Gln Ala Glu Gly Thr Thr Cys Asn Val Gly Ser  
35 40 45  
  
atc gct tgc tgc aac tcc ccc gct gag acc aac aac gac agt ctg ttg 192  
Ile Ala Cys Cys Asn Ser Pro Ala Glu Thr Asn Asn Asp Ser Leu Leu  
50 55 60  
  
agc ggt ctg ctc ggt gct ggc ctt ctc aac ggg ctc tcg ggc aac act 240  
Ser Gly Leu Leu Gly Ala Gly Leu Leu Asn Gly Leu Ser Gly Asn Thr  
65 70 75 80  
  
ggc agc gcc tgc gcc aag gcg agc ttg att gac cag ctg ggt ctg ctc 288  
Gly Ser Ala Cys Ala Lys Ala Ser Leu Ile Asp Gln Leu Gly Leu Leu  
85 90 95  
  
gct ctc gtc gac cac act gag gaa ggc ccc gtc tgc aag aac atc gtc 336  
Ala Leu Val Asp His Thr Glu Glu Gly Pro Val Cys Lys Asn Ile Val

---

	100	105	110	
	gct tgc tgc cct gag gga acc acc aac tgt gtt gcc gtc gac aac gct			384
	Ala Cys Cys Pro Glu Gly Thr Thr Asn Cys Val Ala Val Asp Asn Ala			
	115	120	125	
	ggc gcc ggt acc aag gct gag taa			408
	Gly Ala Gly Thr Lys Ala Glu			
	130	135		
<210>	16			
<211>	135			
<212>	PRT			
<213>	构巢曲霉			
<400>	16			
Met Arg Phe Ile Val Ser Leu Leu Ala Phe Thr Ala Ala Ala Thr Ala				
1	5	10	15	
Thr Ala Leu Pro Ala Ser Ala Ala Lys Asn Ala Lys Leu Ala Thr Ser				
	20	25	30	
Ala Ala Phe Ala Lys Gln Ala Glu Gly Thr Thr Cys Asn Val Gly Ser				
	35	40	45	
Ile Ala Cys Cys Asn Ser Pro Ala Glu Thr Asn Asn Asp Ser Leu Leu				
	50	55	60	
Ser Gly Leu Leu Gly Ala Gly Leu Leu Asn Gly Leu Ser Gly Asn Thr				
65	70	75	80	
Gly Ser Ala Cys Ala Lys Ala Ser Leu Ile Asp Gln Leu Gly Leu Leu				
	85	90	95	
Ala Leu Val Asp His Thr Glu Glu Gly Pro Val Cys Lys Asn Ile Val				
	100	105	110	

Ala Cys Cys Pro Glu Gly Thr Thr Asn Cys Val Ala Val Asp Asn Ala  
 115 120 125

Gly Ala Gly Thr Lys Ala Glu  
 130 135

<210> 17  
 <211> 678  
 <212> DNA  
 <213> 人工序列  
 <220>  
 <223> 融合蛋白

<220>  
 <221> CDS  
 <222> (1)..(678)  
 <223>

<400> 17  
 atg aag atc acc gct gtc att gcc ctt tta ttc tca ctt gct gct gcc 48  
 Met Lys Ile Thr Ala Val Ile Ala Leu Leu Phe Ser Leu Ala Ala Ala  
 1 5 10 15  
 tca cct att cca gtt gcc gat cct ggt gtg gtt tca gtt agc aag tca 96  
 Ser Pro Ile Pro Val Ala Asp Pro Gly Val Val Ser Val Ser Lys Ser  
 20 25 30  
 tat gct gat ttc ctt cgt gtt tac caa agt tgg aac act ttt gct aat 144  
 Tyr Ala Asp Phe Leu Arg Val Tyr Gln Ser Trp Asn Thr Phe Ala Asn  
 35 40 45  
 cct gat aga ccc aac ttg aaa aag cgc ctc ccg gcc tct gcc gca aag 192  
 Pro Asp Arg Pro Asn Leu Lys Lys Arg Leu Pro Ala Ser Ala Ala Lys  
 50 55 60  
 aac gcg aag ctg gcc acc tcg gcg gcc ttc gcc aag cag gct gaa ggc 240  
 Asn Ala Lys Leu Ala Thr Ser Ala Ala Phe Ala Lys Gln Ala Glu Gly

65	70	75	80	
acc acc tgc aat gtc ggc tgc atc gct tgc tgc aac tcc ccc gct gag				288
Thr Thr Cys Asn Val Gly Ser Ile Ala Cys Cys Asn Ser Pro Ala Glu				
	85	90	95	
acc aac aac gac agt ctg ttg agc ggt ctg ctc ggt gct ggc ctt ctc				336
Thr Asn Asn Asp Ser Leu Leu Ser Gly Leu Leu Gly Ala Gly Leu Leu				
	100	105	110	
aac ggg ctc tgc ggc aac act ggc agc gcc tgc gcc aag gcg agc ttg				384
Asn Gly Leu Ser Gly Asn Thr Gly Ser Ala Cys Ala Lys Ala Ser Leu				
	115	120	125	
att gac cag ctg ggt ctg ctc gct ctc gtc gac cac act gag gaa ggc				432
Ile Asp Gln Leu Gly Leu Leu Ala Leu Val Asp His Thr Glu Glu Gly				
	130	135	140	
ccc gtc tgc aag aac atc gtc gct tgc tgc cct gag gga acc acc aac				480
Pro Val Cys Lys Asn Ile Val Ala Cys Cys Pro Glu Gly Thr Thr Asn				
	145	150	155	160
tgt gtt gcc gtc gac aac gct ggc gcc ggt acc aag gct gag ctg gtt				528
Cys Val Ala Val Asp Asn Ala Gly Ala Gly Thr Lys Ala Glu Leu Val				
	165	170	175	
ccg cgt gga tcc atc gaa ggt cgt ggc ggc cgc atc ttt tac cca tac				576
Pro Arg Gly Ser Ile Glu Gly Arg Gly Gly Arg Ile Phe Tyr Pro Tyr				
	180	185	190	
gat gtt cct gac tat gcg ggc tat ccc tat gac gtc ccg gac tat gca				624
Asp Val Pro Asp Tyr Ala Gly Tyr Pro Tyr Asp Val Pro Asp Tyr Ala				
	195	200	205	
gga tcc tat cca tat gac gtt cca gat tac gct gct cag tgc ggc cgc				672
Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ala Gln Cys Gly Arg				
	210	215	220	
taa tag				678

&lt;210&gt; 18

&lt;211&gt; 224

<212> PRT  
 <213> 人工序列  
 <220>  
 <223> 融合蛋白

<400> 18

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

Ser Pro Ile Pro Val Ala Asp Pro Gly Val Val Ser Val Ser Lys Ser  
 20 25 30

Tyr Ala Asp Phe Leu Arg Val Tyr Gln Ser Trp Asn Thr Phe Ala Asn  
 35 40 45

Pro Asp Arg Pro Asn Leu Lys Lys Arg Leu Pro Ala Ser Ala Ala Lys  
 50 55 60

Asn Ala Lys Leu Ala Thr Ser Ala Ala Phe Ala Lys Gln Ala Glu Gly  
 65 70 75 80

Thr Thr Cys Asn Val Gly Ser Ile Ala Cys Cys Asn Ser Pro Ala Glu  
 85 90 95

Thr Asn Asn Asp Ser Leu Leu Ser Gly Leu Leu Gly Ala Gly Leu Leu  
 100 105 110

Asn Gly Leu Ser Gly Asn Thr Gly Ser Ala Cys Ala Lys Ala Ser Leu  
 115 120 125

Ile Asp Gln Leu Gly Leu Leu Ala Leu Val Asp His Thr Glu Glu Gly  
 130 135 140





Val Ile Gly Leu Val Pro Ile Gln Asp Leu Asn Val Leu Gly Asp Asp  
 85 90 95

Met Asn Gln Gln Cys Ala Glu Asn Ser Thr Gln Ala Lys Arg Asp Gly  
 100 105 110

Ala Leu Ala His Leu Leu Glu Asp Val Ser Ile Leu Ser Ser Asn Gly  
 115 120 125

Glu Gly Gly Lys Gly  
 130

<210> 21

<211> 112

<212> PRT

<213> Agaricus bisporus

<400> 21

Met Ile Ser Arg Val Leu Val Ala Ala Leu Val Ala Leu Pro Ala Leu  
 1 5 10 15

Val Thr Ala Thr Pro Ala Pro Gly Lys Pro Lys Ala Ser Ser Gln Cys  
 20 25 30

Asp Val Gly Glu Ile His Cys Cys Asp Thr Gln Gln Thr Pro Asp His  
 35 40 45

Thr Ser Ala Ala Ala Ser Gly Leu Leu Gly Val Pro Ile Asn Leu Gly  
 50 55 60

Ala Phe Leu Gly Phe Asp Cys Thr Pro Ile Ser Val Leu Gly Val Gly  
 65 70 75 80



Thr Pro Ile Asn Ala Asn Val  
115

<210> 23  
<211> 32  
<212> DNA  
<213> 人工序列  
<220>  
<223> PCR 引物

<400> 23  
cagctggggtc tgctcgctct cgtcgaccac ac 32

<210> 24  
<211> 32  
<212> DNA  
<213> 人工序列  
<220>  
<223> PCR 引物

<400> 24  
gtgtgggtcga cgagagcgag cagacccagc tg 32

<210> 25  
<211> 30  
<212> DNA  
<213> 人工序列  
<220>  
<223> PCR 引物

<400> 25  
gaggaacca ccaactgtgt tgccgtcgac 30

---

<210> 26  
<211> 30  
<212> DNA  
<213> 人工序列  
<220>  
<223> PCR 引物

<400> 26  
gtcgacggca acacagttgg tggttccctc 30

<210> 27  
<211> 34  
<212> DNA  
<213> 人工序列  
<220>  
<223> PCR 引物

<400> 27  
taataactcg agatgcgctt catcgtctct ctcc 34

<210> 28  
<211> 33  
<212> DNA  
<213> 人工序列  
<220>  
<223> PCR 引物

<400> 28  
taataaggat ccttactcag ccttggtagc ggc 33

<210> 29  
<211> 30  
<212> DNA  
<213> 人工序列  
<220>

---

<223> PCR 引物

<400> 29

ggtaccaagg ctgagctggt tccgcgtgga

30

<210> 30

<211> 30

<212> DNA

<213> 人工序列

<220>

<223> PCR 引物

<400> 30

tccacgcgga accagctcag ccttgggtacc

30

<210> 31

<211> 36

<212> DNA

<213> 人工序列

<220>

<223> PCR 引物

<400> 31

attattccat ggctattage ggccgcactg agcagc

36

<210> 32

<211> 30

<212> DNA

<213> 人工序列

<220>

<223> PCR 引物

<400> 32

gcctcaccta ttccactccc ggcctctgcc

30

<210> 33  
<211> 30  
<212> DNA  
<213> 人工序列  
<220>  
<223> PCR 引物

<400> 33  
ggcagaggcc gggagtggaa taggtgaggc 30

<210> 34  
<211> 49  
<212> DNA  
<213> 人工序列  
<220>  
<223> PCR 引物

<400> 34  
taatttctcg agatgaagat caccgctgtc attgcccttt tattctcac 49

<210> 35  
<211> 33  
<212> DNA  
<213> 人工序列  
<220>  
<223> PCR 引物

<400> 35  
gtgccgatc ctggtgtgct cccggcctct gcc 33

<210> 36  
<211> 33

<212> DNA  
 <213> 人工序列  
 <220>  
 <223> PCR 引物

<400> 36  
 cacaccagga tcggcaactg gaataggtga ggc 33

<210> 37  
 <211> 30  
 <212> DNA  
 <213> 人工序列  
 <220>  
 <223> PCR 引物

<400> 37  
 aactgaaaa agcgcctccc ggcctctgcc 30

<210> 38  
 <211> 35  
 <212> DNA  
 <213> 人工序列  
 <220>  
 <223> PCR 引物

<400> 38  
 ggcagaggcc gggaggcgct ttttcaagtt gggtc 35

<210> 39  
 <211> 552  
 <212> DNA  
 <213> 构巢曲霉  
 <220>  
 <221> CDS  
 <222> (1)..(288)

<223>

<220>

<221> CDS

<222> (508)..(549)

<223>

<220>

<221> 内含子

<222> (456)..(507)

<223>

<220>

<221> CDS

<222> (381)..(455)

<223>

<220>

<221> 内含子

<222> (289)..(380)

<223>

<400> 39

atg cgc ttc atc gtc tct ctc ctc gcc ttc act gcc gcg gcc acc gca	48
Met Arg Phe Ile Val Ser Leu Leu Ala Phe Thr Ala Ala Ala Thr Ala	
1                    5                    10                    15	

acc gcc ctc ccg gcc tct gcc gca aag aac gcg aag ctg gcc acc tcg	96
Thr Ala Leu Pro Ala Ser Ala Ala Lys Asn Ala Lys Leu Ala Thr Ser	
20                    25                    30	

gcg gcc ttc gcc aag cag gct gaa ggc acc acc tgc aat gtc ggc tcg	144
Ala Ala Phe Ala Lys Gln Ala Glu Gly Thr Thr Cys Asn Val Gly Ser	
35                    40                    45	

atc gct tgc tgc aac tcc ccc gct gag acc aac aac gac agt ctg ttg	192
Ile Ala Cys Cys Asn Ser Pro Ala Glu Thr Asn Asn Asp Ser Leu Leu	
50                    55                    60	

agc ggt ctg ctc ggt gct ggc ctt ctc aac ggg ctc tcg ggc aac act 240  
 Ser Gly Leu Leu Gly Ala Gly Leu Leu Asn Gly Leu Ser Gly Asn Thr  
 65 70 75 80

ggc agc gcc tgc gcc aag gcg agc ttg att gac cag ctg ggt ctg ctc 288  
 Gly Ser Ala Cys Ala Lys Ala Ser Leu Ile Asp Gln Leu Gly Leu Leu  
 85 90 95

ggtacgtgat cccactcag tcgctcccg agaggctgag ggaagacgag cgacggtcta 348

gaaatggtgt gctaatagat gcatgtgtgc ag ctc tcg tcg acc aca ctg agg 401  
 Leu Ser Ser Thr Thr Leu Arg  
 100

aag gcc ccg tct gca aga aca tcg tcg ctt gct gcc ctg agg gaa cca 449  
 Lys Ala Pro Ser Ala Arg Thr Ser Ser Leu Ala Ala Leu Arg Glu Pro  
 105 110 115

cca acg tacgtctttc agatctgcta caagtgagc gatcaaaact aacatattcc ag 507  
 Pro Thr  
 120

tgt gtt gcc gtc gac aac gct ggc gcc ggt acc aag get gag taa 552  
 Cys Val Ala Val Asp Asn Ala Gly Ala Gly Thr Lys Ala Glu  
 125 130 135

<210> 40  
 <211> 135  
 <212> PRT  
 <213> 构巢曲霉  
 <400> 40

Met Arg Phe Ile Val Ser Leu Leu Ala Phe Thr Ala Ala Ala Thr Ala  
 1 5 10 15

Thr Ala Leu Pro Ala Ser Ala Ala Lys Asn Ala Lys Leu Ala Thr Ser  
 20 25 30

Ala Ala Phe Ala Lys Gln Ala Glu Gly Thr Thr Cys Asn Val Gly Ser  
 35 40 45

Ile Ala Cys Cys Asn Ser Pro Ala Glu Thr Asn Asn Asp Ser Leu Leu  
 50 55 60

Ser Gly Leu Leu Gly Ala Gly Leu Leu Asn Gly Leu Ser Gly Asn Thr  
 65 70 75 80

Gly Ser Ala Cys Ala Lys Ala Ser Leu Ile Asp Gln Leu Gly Leu Leu  
 85 90 95

Leu Ser Ser Thr Thr Leu Arg Lys Ala Pro Ser Ala Arg Thr Ser Ser  
 100 105 110

Leu Ala Ala Leu Arg Glu Pro Pro Thr Cys Val Ala Val Asp Asn Ala  
 115 120 125

Gly Ala Gly Thr Lys Ala Glu  
 130 135

<210> 41  
 <211> 34  
 <212> DNA  
 <213> 人工序列  
 <220>  
 <223> PCR引物

<400> 41  
 taataaggat ccatgcgctt catcgtctct ctcc

34

<210> 42  
 <211> 129  
 <212> DNA  
 <213> 粟酒裂殖酵母

<220>

<221> CDS

<222> (1)..(126)

<223>

<400> 42

atg gac tca atg gct aac tcc gtt tct tcc tcc tct gtc gtc aac gct 48  
Met Asp Ser Met Ala Asn Ser Val Ser Ser Ser Ser Val Val Asn Ala  
1 5 10 15

ggc aac aag cct gct gaa act ctt aac aag acc gtt aag aat tat acc 96  
Gly Asn Lys Pro Ala Glu Thr Leu Asn Lys Thr Val Lys Asn Tyr Thr  
20 25 30

ccc aag gtt cct tac atg tgt gtc att gca taa 129  
Pro Lys Val Pro Tyr Met Cys Val Ile Ala  
35 40

<210> 43

<211> 42

<212> PRT

<213> 粟酒裂殖酵母

<400> 43

Met Asp Ser Met Ala Asn Ser Val Ser Ser Ser Ser Val Val Asn Ala  
1 5 10 15

Gly Asn Lys Pro Ala Glu Thr Leu Asn Lys Thr Val Lys Asn Tyr Thr  
20 25 30

Pro Lys Val Pro Tyr Met Cys Val Ile Ala  
35 40

<210> 44

<211> 27

<212> DNA

<213> 粟酒裂殖酵母

<400> 44  
tataaccccca aggttcctta catgtgt 27

<210> 45  
<211> 135  
<212> DNA  
<213> 粟酒裂殖酵母  
<220>  
<221> CDS  
<222> (1)..(132)  
<223>

<400> 45  
atg gac tcc att gca act aac act cat tct tca tcc att gtc aat gcc 48  
Met Asp Ser Ile Ala Thr Asn Thr His Ser Ser Ser Ile Val Asn Ala  
1 5 10 15

tac aac aac aat cct acc gat gtt gta aaa act caa aac att aaa aat 96  
Tyr Asn Asn Asn Pro Thr Asp Val Val Lys Thr Gln Asn Ile Lys Asn  
20 25 30

tat act cca aag gtt cct tat atg tgt gta att gct taa 135  
Tyr Thr Pro Lys Val Pro Tyr Met Cys Val Ile Ala  
35 40

<210> 46  
<211> 44  
<212> PRT  
<213> 粟酒裂殖酵母  
<400> 46

Met Asp Ser Ile Ala Thr Asn Thr His Ser Ser Ser Ile Val Asn Ala  
1 5 10 15

Tyr Asn Asn Asn Pro Thr Asp Val Val Lys Thr Gln Asn Ile Lys Asn  
20 25 30

Tyr Thr Pro Lys Val Pro Tyr Met Cys Val Ile Ala  
 35 40

<210> 47

<211> 27

<212> DNA

<213> 粟酒裂殖酵母

<400> 47

tatactccaa aggttcctta tatgtgt

27

<210> 48

<211> 126

<212> DNA

<213> 粟酒裂殖酵母

<220>

<221> CDS

<222> (1)..(123)

<223>

<400> 48

atg gac tca atg gct aac act gtt tct tcc tcc gtc gtt aac act ggc 48

Met Asp Ser Met Ala Asn Thr Val Ser Ser Ser Val Val Asn Thr Gly

1 5 10 15

aac aag cct tct gaa act ctt aac aag act gtt aag aat tat acc ccc 96

Asn Lys Pro Ser Glu Thr Leu Asn Lys Thr Val Lys Asn Tyr Thr Pro

20 25 30

aag gtt cct tac atg tgt gtc att gca taa 126

Lys Val Pro Tyr Met Cys Val Ile Ala

35 40

<210> 49

<211> 41

<212> PRT

<213> 粟酒裂殖酵母

<400> 49

Met Asp Ser Met Ala Asn Thr Val Ser Ser Ser Val Val Asn Thr Gly  
 1                    5                    10                    15

Asn Lys Pro Ser Glu Thr Leu Asn Lys Thr Val Lys Asn Tyr Thr Pro  
                   20                    25                    30

Lys Val Pro Tyr Met Cys Val Ile Ala  
                   35                    40

<210> 50

<211> 27

<212> DNA

<213> 粟酒裂殖酵母

<400> 50

tataccccca aggttcctta catgtgt

27

<210> 51

<211> 9

<212> PRT

<213> 粟酒裂殖酵母

<400> 51

Tyr Thr Pro Lys Val Pro Tyr Met Cys

1                    5

<210> 52

<211> 586

<212> DNA

<213> 构巢曲霉

<220>

<221> 内含子

<222> (471)..(530)

<223>

<220>

<221> 内含子

<222> (338)..(389)

<223>

<400> 52

```

atgaagttct ccattgctgc cgctgtcggt gctttcgccg cctccgctgc ggccctccct    60
cctgcccattg attcccagtt cgctggcaat ggtgttgcca acaagggcaa cagcaacgct    120
aagttccctg tccccgaaaa cgtgaccgtc aagcaggcct ccgacaagtg cggtgaccag    180
gcccagctct cttgctgcaa caaggccacg tacgccggtg acaccacaac cgttgatgag    240
ggtcttctgt ctggtgccct cagcggcctc atcgcgccg ggtctggtgc cgaaggtctt    300
ggtctcttcg atcagtgtc caagcttgat gttgctggtc agttcttcga aaatcacttt    360
cgtgatgccc caatgctaac aattaccagt cctcattggc atccaagatc ttgtcaacca    420
gaagtgaag caaacattg cctgctgcca gaactcccc tccagcggg tatgttcctt    480
tgttttacag cttattcact taaaccgatt aatctaaca cgctcacagg atggcaacct    540
tattggtgtc ggtctccctt gcgttgccct tggtccatc ctctaa                    586

```

<210> 53

<211> 474

<212> DNA

<213> 构巢曲霉

<220>

<221> CDS

<222> (1)..(471)

<223>

<400> 53

```

atg aag ttc tcc att gct gcc gct gtc gtt gct ttc gcc gcc tcc gtc    48
Met Lys Phe Ser Ile Ala Ala Ala Val Val Ala Phe Ala Ala Ser Val
1           5           10           15

```

gcg gcc ctc cct cct gcc cat gat tcc cag ttc gct ggc aat ggt gtt Ala Ala Leu Pro Pro Ala His Asp Ser Gln Phe Ala Gly Asn Gly Val 20 25 30	96
ggc aac aag ggc aac agc aac gtc aag ttc cct gtc ccc gaa aac gtg Gly Asn Lys Gly Asn Ser Asn Val Lys Phe Pro Val Pro Glu Asn Val 35 40 45	144
acc gtc aag cag gcc tcc gac aag tgc ggt gac cag gcc cag ctc tct Thr Val Lys Gln Ala Ser Asp Lys Cys Gly Asp Gln Ala Gln Leu Ser 50 55 60	192
tgc tgc aac aag gcc acg tac gcc ggt gac acc aca acc gtt gat gag Cys Cys Asn Lys Ala Thr Tyr Ala Gly Asp Thr Thr Thr Val Asp Glu 65 70 75 80	240
ggt ctt ctg tct ggt gcc ctc agc ggc ctc atc ggc gcc ggg tct ggt Gly Leu Leu Ser Gly Ala Leu Ser Gly Leu Ile Gly Ala Gly Ser Gly 85 90 95	288
gcc gaa ggt ctt ggt ctc ttc gat cag tgc tcc aag ctt gat gtt gct Ala Glu Gly Leu Gly Leu Phe Asp Gln Cys Ser Lys Leu Asp Val Ala 100 105 110	336
gtc ctc att ggc atc caa gat ctt gtc aac cag aag tgc aag caa aac Val Leu Ile Gly Ile Gln Asp Leu Val Asn Gln Lys Cys Lys Gln Asn 115 120 125	384
att gcc tgc tgc cag aac tcc ccc tcc agc gcg gat ggc aac ctt att Ile Ala Cys Cys Gln Asn Ser Pro Ser Ser Ala Asp Gly Asn Leu Ile 130 135 140	432
ggt gtc ggt ctc cct tgc gtt gcc ctt ggc tcc atc ctc taa Gly Val Gly Leu Pro Cys Val Ala Leu Gly Ser Ile Leu 145 150 155	474

<210> 54  
 <211> 157  
 <212> PRT  
 <213> 构巢曲霉  
 <400> 54

Met Lys Phe Ser Ile Ala Ala Ala Val Val Ala Phe Ala Ala Ser Val  
1 5 10 15

Ala Ala Leu Pro Pro Ala His Asp Ser Gln Phe Ala Gly Asn Gly Val  
20 25 30

Gly Asn Lys Gly Asn Ser Asn Val Lys Phe Pro Val Pro Glu Asn Val  
35 40 45

Thr Val Lys Gln Ala Ser Asp Lys Cys Gly Asp Gln Ala Gln Leu Ser  
50 55 60

Cys Cys Asn Lys Ala Thr Tyr Ala Gly Asp Thr Thr Thr Val Asp Glu  
65 70 75 80

Gly Leu Leu Ser Gly Ala Leu Ser Gly Leu Ile Gly Ala Gly Ser Gly  
85 90 95

Ala Glu Gly Leu Gly Leu Phe Asp Gln Cys Ser Lys Leu Asp Val Ala  
100 105 110

Val Leu Ile Gly Ile Gln Asp Leu Val Asn Gln Lys Cys Lys Gln Asn  
115 120 125

Ile Ala Cys Cys Gln Asn Ser Pro Ser Ser Ala Asp Gly Asn Leu Ile  
130 135 140

Gly Val Gly Leu Pro Cys Val Ala Leu Gly Ser Ile Leu  
145 150 155

<210> 55  
<211> 420  
<212> DNA

<213> 构巢曲霉

<220>

<221> CDS

<222> (1)..(417)

<223>

<400> 55

ctc cct cct gcc cat gat tcc cag ttc gct ggc aat ggt gtt ggc aac 48  
 Leu Pro Pro Ala His Asp Ser Gln Phe Ala Gly Asn Gly Val Gly Asn  
 1 5 10 15

aag ggc aac agc aac gtc aag ttc cct gtc ccc gaa aac gtg acc gtc 96  
 Lys Gly Asn Ser Asn Val Lys Phe Pro Val Pro Glu Asn Val Thr Val  
 20 25 30

aag cag gcc tcc gac aag tgc ggt gac cag gcc cag ctc tct tgc tgc 144  
 Lys Gln Ala Ser Asp Lys Cys Gly Asp Gln Ala Gln Leu Ser Cys Cys  
 35 40 45

aac aag gcc acg tac gcc ggt gac acc aca acc gtt gat gag ggt ctt 192  
 Asn Lys Ala Thr Tyr Ala Gly Asp Thr Thr Thr Val Asp Glu Gly Leu  
 50 55 60

ctg tct ggt gcc ctc agc ggc ctc atc ggc gcc ggg tct ggt gcc gaa 240  
 Leu Ser Gly Ala Leu Ser Gly Leu Ile Gly Ala Gly Ser Gly Ala Glu  
 65 70 75 80

ggt ctt ggt ctc ttc gat cag tgc tcc aag ctt gat gtt gct gtc ctc 288  
 Gly Leu Gly Leu Phe Asp Gln Cys Ser Lys Leu Asp Val Ala Val Leu  
 85 90 95

att ggc atc caa gat ctt gtc aac cag aag tgc aag caa aac att gcc 336  
 Ile Gly Ile Gln Asp Leu Val Asn Gln Lys Cys Lys Gln Asn Ile Ala  
 100 105 110

tgc tgc cag aac tcc ccc tcc agc gcg gat ggc aac ctt att ggt gtc 384  
 Cys Cys Gln Asn Ser Pro Ser Ser Ala Asp Gly Asn Leu Ile Gly Val  
 115 120 125

ggt ctc cct tgc gtt gcc ctt ggc tcc atc ctc taa 420  
 Gly Leu Pro Cys Val Ala Leu Gly Ser Ile Leu  
 130 135

<210> 56  
 <211> 139  
 <212> PRT  
 <213> 构巢曲霉  
 <400> 56

Leu Pro Pro Ala His Asp Ser Gln Phe Ala Gly Asn Gly Val Gly Asn  
 1                   5                   10                   15

Lys Gly Asn Ser Asn Val Lys Phe Pro Val Pro Glu Asn Val Thr Val  
           20                   25                   30

Lys Gln Ala Ser Asp Lys Cys Gly Asp Gln Ala Gln Leu Ser Cys Cys  
           35                   40                   45

Asn Lys Ala Thr Tyr Ala Gly Asp Thr Thr Thr Val Asp Glu Gly Leu  
       50                   55                   60

Leu Ser Gly Ala Leu Ser Gly Leu Ile Gly Ala Gly Ser Gly Ala Glu  
 65                   70                   75                   80

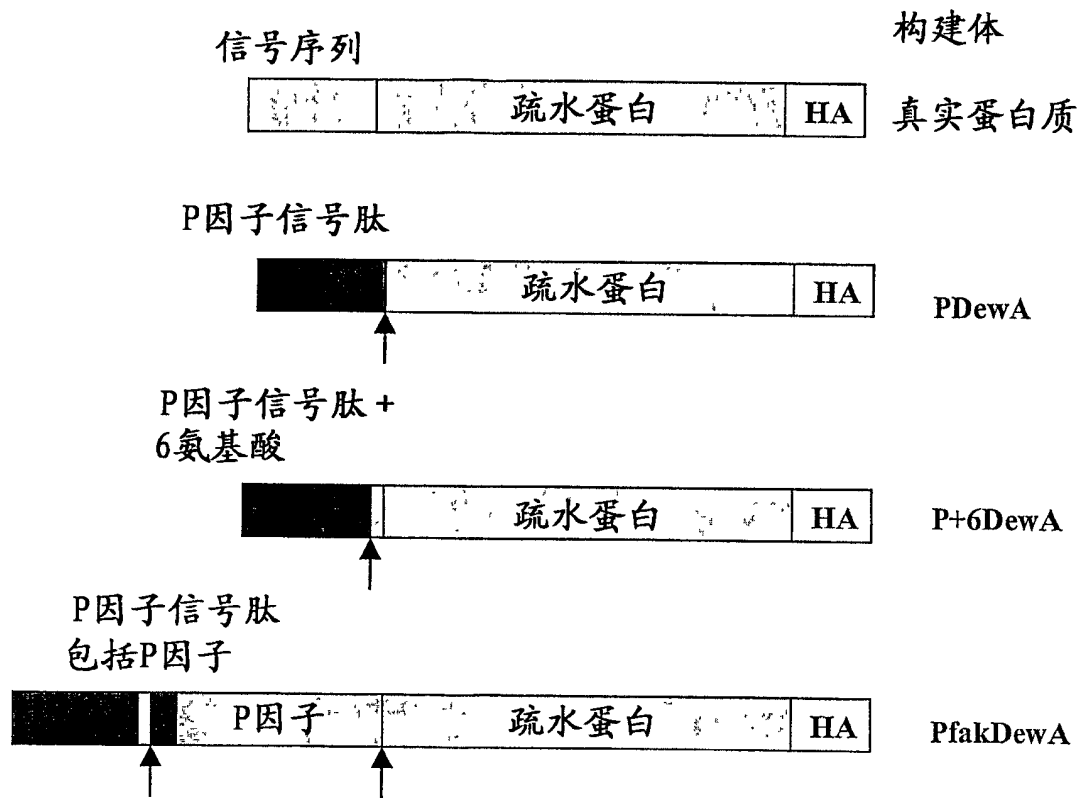
Gly Leu Gly Leu Phe Asp Gln Cys Ser Lys Leu Asp Val Ala Val Leu  
           85                   90                   95

Ile Gly Ile Gln Asp Leu Val Asn Gln Lys Cys Lys Gln Asn Ile Ala  
           100                   105                   110

Cys Cys Gln Asn Ser Pro Ser Ser Ala Asp Gly Asn Leu Ile Gly Val  
           115                   120                   125

Gly Leu Pro Cys Val Ala Leu Gly Ser Ile Leu  
       130                   135

用于粟酒裂殖酵母疏水蛋白分泌的构建体



↑ = 推测的蛋白酶切割位点

图1

## A) DewA基因的基因组序列:

(两个内含子的序列用下划线标出)

ATGCGCTTCA TCGTCTCTCT CCTCGCCTTC ACTGCCGCGG CCACCGCAAC CGCCCTCCCG  
 GCCTCTGCCG CAAAGAACGC GAAGCTGGCC ACCTCGGCGG CCTTCGCCAA GCAGGCTGAA  
 GGCACCACCT GCAATGTCGG CTCGATCGCT TGCTGCAACT CCCCCTGTA GACCAACAAC  
 GACAGTCTGT TGAGCGGTCT GCTCGGTGCT GGCTTCTCA ACGGGCTCTC GGGCAACT  
 GGCAGCGCCT GCGCCAAGGC GAGCTTGATT GACCAGCTGG GTCTGCTCGG TACGTGATCC  
CCACTCAGTC GCTCCCGGAG AGGCTGAGGG AAGACGAGCG ACGGTCTAGA AATGGTGTGC  
TAATAGATGC ATGTGTGCAG CTCTCGTCGA CCACACTGAG GAAGGCCCG TCTGCAAGAA  
 CATCGTCGCT TGCTGCCCTG AGGGAACCAC CAACGTACGT CTTTCAGATC TGCTACAAGT  
 GAGGCGATCA AAATAACAT ATTCCAGTGT GTTGCCGTCG ACAACGCTGG CGCCGGTACC  
 AAGGCTGAGT AA

## B) 构巢曲霉DewA蛋白质的序列:

MRFIVSLLAF TAAATATALP ASAAKNAKLA TSAAFKQAE GTTCNVGSIA CCNSPAETNN  
 DILLSGLLGA GLLNLSGNT GSACAKASLI DQLGLLALVD HTEEGPVCKN IVACCPEGTT  
 NCVAVDNAGA GTKAE

(ATGCGCTTCA TCGTCTCTCT CCTCGCCTTC ACTGCCGCGG CCACCGCAAC CGCCCTCCCG  
 GCCTCTGCCG CAAAGAACGC GAAGCTGGCC ACCTCGGCGG CCTTCGCCAA GCAGGCTGAA  
 GGCACCACCT GCAATGTCGG CTCGATCGCT TGCTGCAACT CCCCCTGTA GACCAACAAC  
 GACAGTCTGT TGAGCGGTCT GCTCGGTGCT GGCTTCTCA ACGGGCTCTC GGGCAACT  
 GGCAGCGCCT GCGCCAAGGC GAGCTTGATT GACCAGCTGG GTCTGCTCGC TCTCGTCGAC  
 CACTGAGG AAGGCCCGT CTGCAAGAAC ATCGTCGCTT GCTGCCCTGA GGAACCACC  
 AACTGTGTTG CCGTCGACAA CGCTGGCGCC GGTAACAAGG CTGAGTAA)

## C) HA-标记序列:

LVPRGSIEGR GGRIFYPYDV PDYAGYPYDV PDYAGSYPYD VPDYAAQCGR  
 (CTGGT TCCGCGTGGA TCCATCGAAG GTCGTGGCGG CCGCATCTTT TACCCATACG  
 ATGTTCTGTA CTATGCGGGC TATCCCTATG ACGTCCCGGA CTATGCAGGA TCCTATCCAT  
 ATGACGTTCC AGATTACGCT GCTCAGTGCG GCCGCTAATA G)

图2

## A) P因子前蛋白质的序列:

MKITAVIALL FSLAAASPIP VADPGVVSVS KSYADFLRVY QSWNTFANPD RPNLKKREFE  
AAPAKTYADF LRAYQSWNTF VNPDRPNLKK REFEEAPEKS YADFLRAYHS WNTFVNPDRP  
NLKKREFEAA PAKTYADFLR AYQSWNTFVN PDRPNLKKRT EEDEENEEED EEYYRFLQFY  
 IMTVPENSTI TDVNITAKFE S  
 (ATGAAGATCA CCGCTGTCAT TGCCCTTTTA TTCTCACTTG CTGCTGCCTC ACCTATTCCA  
 GTTGCCGATC CTGGTGTGGT TTCAGTTAGC AAGTCATATG CTGATTTCCCT TCGTGTTTAC  
 CAAAGTTGGA ACACTTTTGC TAATCCTGAT AGACCCAACCT TGAAAAAGCG CGAATTGCGAA  
 GCTGCTCCCG CAAAAACTTA TGCTGATTTTCT CTCGTGCTT ATCAAAGTTG GAACACTTTT  
 GTTAATCCTG ACAGACCCAA TTTGAAAAAG CGTGAGTTTG AAGCTGCCCC AGAGAAGAGT  
 TATGCTGATT TCCTTCGTGC TTACCATAGT TGGAACACTT TTGTTAATCC TGACAGACCC  
 AACTTGAAAA AGCGCGAATT CGAAGCTGCT CCCGCAAAAA CTTATGCTGA TTTCTTCGT  
 GCTTACCAA GTTGGAACAC TTTTGTTAAT CCTGACAGAC CCAACTTGAA AAAGCGCACT  
 GAAGAAGATG AAGAGAATGA GGAAGAGGAT GAAGAATACT ATCGCTTTCT TCAGTTTTAT  
 ATCATGACTG TCCCAGAGAA TTCCAATATT ACAGATGTCA ATATTACTGC CAAATTTGAG  
 AGCTAA)

## B) 可除去的信号肽及其下游前蛋白质6个氨基酸的序列:

MKITAVIALL FSLAAASPIP VADPGV  
 (ATGAAGATCA CCGCTGTCAT TGCCCTTTTA TTCTCACTTG CTGCTGCCTC ACCTATTCCA  
 GTTGCCGATC CTGGTGTGG)

## C) 用于“P-穿梭”的序列:

MKITAVIALL FSLAAASPIP VADPGVVSVS KSYADFLRVY QSWNTFANPD RPNLKKR  
 (ATGAAGATCA CCGCTGTCAT TGCCCTTTTA TTCTCACTTG CTGCTGCCTC ACCTATTCCA  
 GTTGCCGATC CTGGTGTGGT TTCAGTTAGC AAGTCATATG CTGATTTCCCT TCGTGTTTAC  
 CAAAGTTGGA ACACTTTTGC TAATCCTGAT AGACCCAACCT TGAAAAAGCG C)

图 3

包含“P-穿棱”序列、成熟DewA和C末端融合的HA-标记的融合蛋白:

MKITAVIALL FSLAAASPIP VADPGVSVS KSYADFLRVY QSWNTFANPD RPNLKKRLPA  
SAAKNAKLAT SAAFAKQAEQ TTCNVGSIAC CNSPAETNND SLISGLLGAG LLNGLSGNTG  
SACAKASLID QLGLLALVDH TEEGVPCKNI VACCPEGTTN CVAVDNAGAG TKAEIVPRGS  
 IEGRGGRIIFY PYDVPDYAGY PYDVPDYAGS YPYDVPDYAA QCGR  
 (ATGAAGATCA CCGCTGTCAT TGCCCTTTTA TTCTCACTTG CTGCTGCCTC ACCTATTCCA  
 GTTGCCGATC CTGGTGTGGT TTCAGTTAGC AAGTCATATG CTGATTCCT TCGTGTTTAC  
 CAAAGTTGGA ACACTTTTGC TAATCCTGAT AGACCCAAC TGA AAAAGCG CCTCCCGGCC  
 TCTGCCGCA AGAACGCGAA GCTGGCCACC TCGGCGGCCT TCGCCAAGCA GGCTGAAGGC  
 ACCACCTGCA ATGTCGGCTC GATCGCTTGC TGCAACTCCC CCGCTGAGAC CAACAACGAC  
 AGTCTGTTGA GCGGTCTGCT CGGTGCTGGC CTTCTCAACG GGCTCTCGGG CAACACTGGC  
 AGCGCCTGCG CCAAGGCGAG CTTGATTGAC CAGCTGGGTC TGCTCGCTCT CGTCGACCAC  
 ACTGAGGAAG GCCCGTCTG CAAGAACATC GTCGCTTGCT GCCCTGAGGG AACCACCAAC  
 TGTGTTGCCG TCGACAACGC TGGCGCCGGT ACCAAGGCTG AGCTGGTTCC GCGTGGATCC  
 ATCGAAGGTC GTGGCGGCCG CATCTTTTAC CCATACGATG TTCCTGACTA TCGGGGCTAT  
 CCCTATGACG TCCCGGACTA TGCAGGATCC TATCCATATG ACGTTCCAGA TTACGCTGCT  
 CAGTGC GGCC GCTAATAG)

图 4

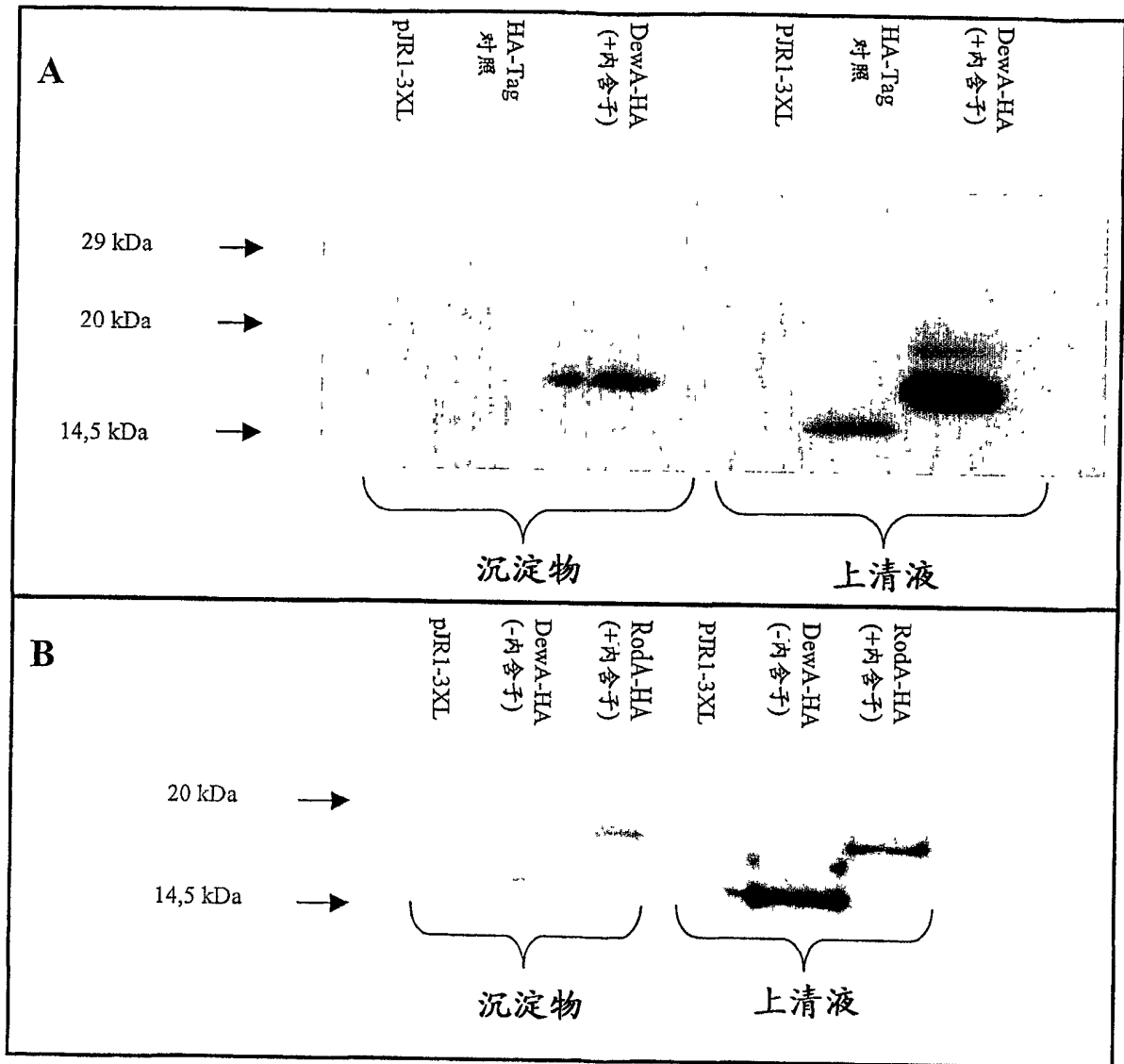


图 5

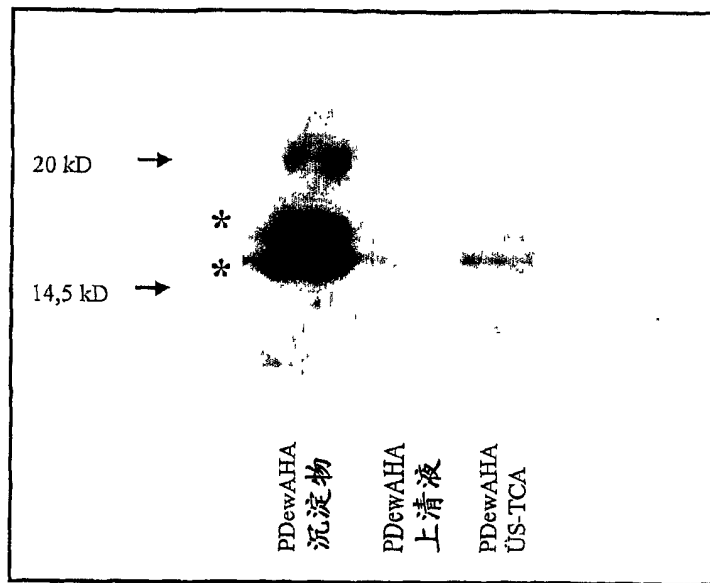


图6

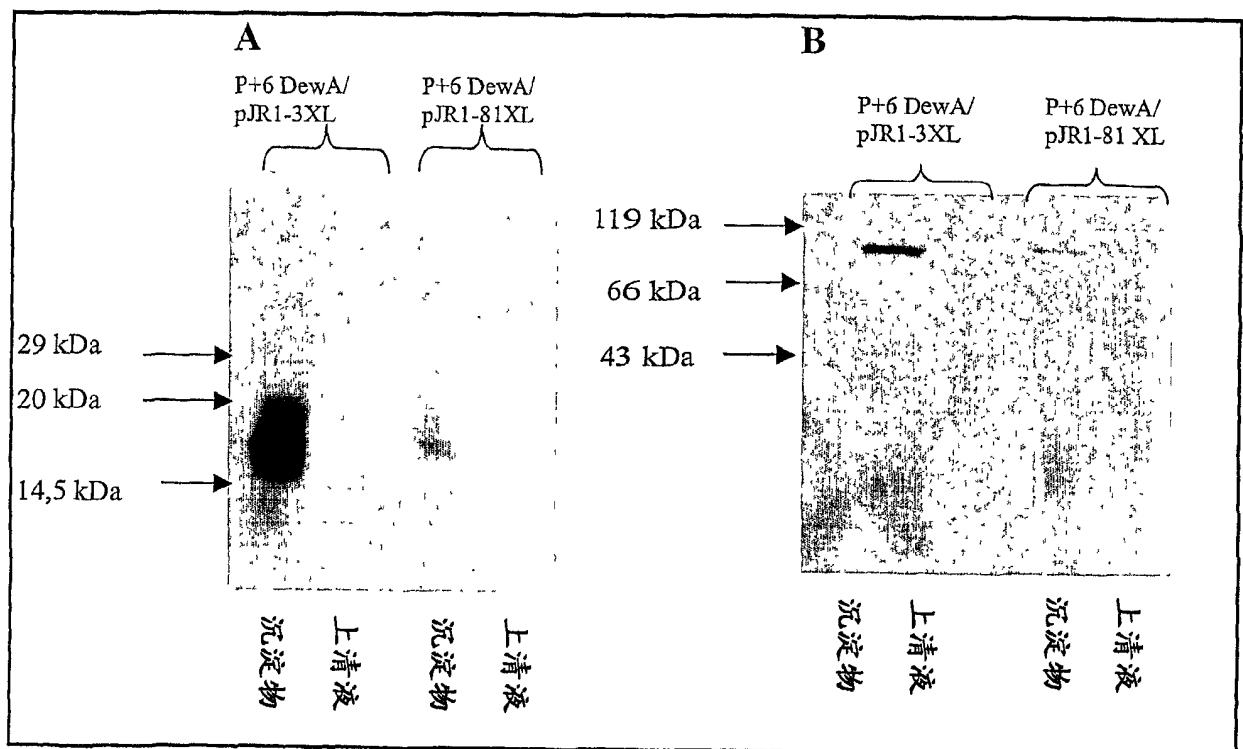


图7

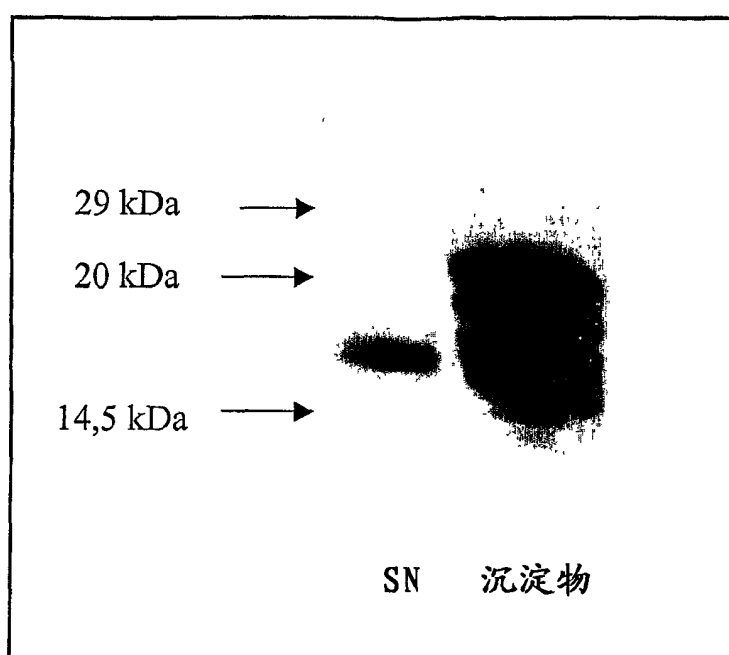


图8

**A) *mfm1*<sup>+</sup> 基因*****mfm1* 前蛋白质的序列**

MDSMANSVSSSSVFNAGNKPAETLNKTVKNYTPKVPYMCVIA

***mfm1*<sup>+</sup> 基因的序列**

atggactcaa tggctaactc cgtttcttcc tcctctgtcg tcaacgctgg caacaagcct  
gctgaaactc ttaacaagac cgtaagaat tataccccca aggttcctta catgtgtgtc attgcataa

***mfm1* 成熟M-信息素**

YTPKVPYMC

**成熟*mfm1* M信息素的DNA序列**

tataccccca aggttcctta catgtgt

**B) *mfm2*<sup>-</sup> 基因*****mfm2* 前蛋白质的序列**

MDSIATNTHSSSIVNAYNNNPTDVVKTQNIKNYTPKVPYMCVIA

***mfm2*<sup>+</sup> 基因的序列**

atggactcca ttgcaactaa cactcattct tcattccattg tcaatgccta caacaacaat  
cctaccgatg ttgtaaaaac tcaaaacatt aaaattata ctccaaaggt tccttatatg  
tgtgtaattg cttaa

***mfm2* 成熟M-信息素**

YTPKVPYMC

**成熟*mfm2* M信息素的DNA序列**

tata ctccaaaggt tccttatatg tgt

**C) *mfm3*<sup>+</sup> 基因*****mfm3* 前蛋白质的序列**

MDSMANIVSSSVVNTGNKPSETLNKTVKNYTPKVPYMCVIA

***mfm3*<sup>+</sup> 基因的序列**

atggactcaa tggctaacac tgtttcttcc tcctctgtta acaactggcaa caagccttct  
gaaactctta acaagactgt taagaattat accccaaggt ttccttacat gtgtgtcatt gcataa

***mfm3* 成熟M-信息素**

YTPKVPYMC

**成熟*mfm3* M信息素的DNA序列**

tat accccaaggt ttccttacat gtgt

图9

## RodA基因的基因组序列

ATGAAGTTCT	CCATFGCTGC	CGCTGTCGTT	GCTTTCGC CG	CCTCCGTCGC	GGCCCTCCCT	CCTGCCCATG
ATTCCCAGTT	CGCTGGCAAT	GGTGTGGCA	ACAAGGGCAA	CAGCAACGTC	AAGTTCCTG	TCCCCGAAAA
CGTGACCGTC	AAGCAGGCCT	CCGACAAGTG	CGGTGACCAG	GCCCAGCTCT	CTTGCTGCAA	CAAGGCCACG
TACGCCGGTG	ACACCACAAC	CGTTGATGAG	GGTCTTCTGT	CTGGTGCCCT	CAGCGGCCTC	ATCGGCGCCG
GGTCTGGTGC	CGAAGGTCTT	GGTCTCTTCG	ATCAGTGC TC	CAAGCTTGAT	GTTGCTGGTC	<u>AGTTCTTCGA</u>
<u>AAATCACTTT</u>	<u>CGTGATGCC</u>	<u>CAATGCTAAC</u>	<u>AATTACCA GT</u>	<u>CCTCATTTGGC</u>	<u>ATCCAAGATC</u>	<u>TTGTCAACCA</u>
GAAGTGCAAG	CAAAACATTG	CCTGCTGCCA	GAACTCCCC	TCCAGCGCGG	<u>TATGTTCCCT</u>	<u>TGTTTTACAG</u>
<u>CTTATTCACT</u>	<u>TAAACCGATT</u>	<u>AATCTAACAA</u>	<u>CGCTCACAGG</u>	ATGGCAACCT	TATTGGTGTC	GGTCTCCCTT
GCGTTGCCCT	TGGCTCCATC	CTCTAA				

## RodA基因的可读框 (ORF) 的DNA序列

ATGAAGTTCT	CCATFGCTGC	CGCTGTCGTT	GCTTTCGC CG	CCTCCGTCGC	GGCCCTCCCT	CCTGCCCATG
ATTCCCAGTT	CGCTGGCAAT	GGTGTGGCA	ACAAGGGCAA	CAGCAACGTC	AAGTTCCTG	TCCCCGAAAA
CGTGACCGTC	AAGCAGGCCT	CCGACAAGTG	CGGTGACCAG	GCCCAGCTCT	CTTGCTGCAA	CAAGGCCACG
TACGCCGGTG	ACACCACAAC	CGTTGATGAG	GGTCTTCTGT	CTGGTGCCCT	CAGCGGCCTC	ATCGGCGCCG
GGTCTGGTGC	CGAAGGTCTT	GGTCTCTTCG	ATCAGTGC TC	CAAGCTTGAT	GTTGCTGTCC	TCATTGGCAT
CCAAGATCTT	GTCAACCAGA	AGTGCAAGCA	AAACATTGCC	TGCTGCCAGA	ACTCCCCCTC	CAGCGCGGAT
GGCAACCTTA	TTGGTGTCCG	TCTCCCTTGC	GTTGCCCTTG	GCTCCATCCT	CTAA	

## RodA蛋白质的序列

MKFSIAAAV	AFAASVAALP	PAHDSQFAGN	GVTNKGNSNV	KFPVPENVTV	KQASDKCGDQ	AQLSCCNKAT
YAGDTTVDDE	GLLSGALSGL	IGAGSGAEGE	GLFDQCSKLD	VAVLIGIQDL	VNQKCKQNI	CCQNSPSSAD
GNLIGVGLPC	VALGSIL					

图10