

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
C12N 9/00 (2006.01)



[12] 发明专利说明书

专利号 ZL 01810182.8

[45] 授权公告日 2009年11月11日

[11] 授权公告号 CN 100558885C

[22] 申请日 2001.3.30 [21] 申请号 01810182.8

[30] 优先权

[32] 2000.3.31 [33] US [31] 60/193,707

[86] 国际申请 PCT/US2001/010481 2001.3.30

[87] 国际公布 WO2001/075077 英 2001.10.11

[85] 进入国家阶段日期 2002.11.26

[73] 专利权人 纳慕尔杜邦公司

地址 美国特拉华州威尔明顿

[72] 发明人 R·D·法伦 M·S·佩尼

S·乔汗 R·迪科西莫

约翰 E·加瓦甘

[56] 参考文献

US5858736A 1999.1.12

PURIFICATION AND CHARACTERIZATION
OF NITRILASERESPONSIBLE FOR THE ENAN-
TIOSELECTIVEHYD. AGRICULTURAL AND
BIOLOGICAL CHEMISTRY, Vol. 55 No. 6. 1991

审查员 李美宣

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 孟凡宏

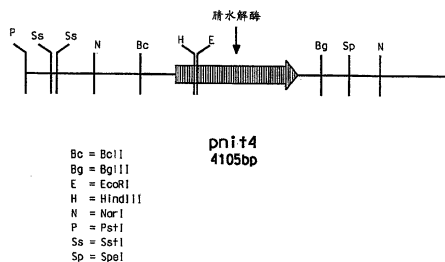
权利要求书5页 说明书61页 附图1页

[54] 发明名称

分离和表达敏捷食酸菌 72W 的腈水解酶基因

[57] 摘要

提供表达腈水解酶并可用作水解含腈底物的生物催化剂的重组微生物菌株。重组细胞用分离自敏捷食酸菌 72W 的编码热稳定腈水解酶的外源基因转化,所述热稳定腈水解酶在温和反应条件下将含腈底物催化水解为羧酸。提供所述腈水解酶基因的核苷酸序列以及所述腈水解酶基因编码的推定氨基酸序列。



1. 一种编码脲水解酶的分离核酸片段，该核酸片段选自：
 - (a) 编码具有脲水解酶活性的多肽的分离核酸片段，其中所述多肽的氨基酸序列示于 SEQ ID NO: 5；和
 - (b) 与(a)完全互补的分离核酸片段。
2. 一种编码脲水解酶的分离核酸片段，所述分离核酸片段选自：
 - (a) SEQ ID NO: 4 的分离核酸片段，其编码具有脲水解酶活性的多肽；和
 - (b) 与(a)完全互补的分离核酸片段。
3. 权利要求 1 或 2 的分离核酸片段，其中所述片段分离自食酸菌(*Acidovorax*)菌株。
 4. 一种多肽，其由权利要求 1 或 2 的核酸片段编码。
 5. 按照权利要求 4 的多肽，其中所述多肽的氨基酸序列示于 SEQ ID NO: 5。
 6. 权利要求 4 的多肽，其特征还在于对选自脂肪族脲和芳香族脲的含脲底物具有脲水解酶活性。
 7. 一种嵌合基因，它包含与合适的调节序列有效连接的权利要求 1、2 或 3 的分离核酸片段。
 8. 保藏号为 ATCC PTA-1175 的大肠杆菌 SW91 包含的质粒 pSW91、保藏号为 ATCC PTA-1176 的大肠杆菌 DH5 α :pnit4 包含的质粒 pnit4、或者保藏号为 ATCC PTA-1177 的大肠杆菌 SS1001 包含的质粒 pnitex2。
 9. 一种表达盒，其包含权利要求 7 的嵌合基因。
 10. 权利要求 9 的表达盒，其选自权利要求 8 所述的质粒。
 11. 一种转化微生物，其包含权利要求 7 的嵌合基因。
 12. 一种转化微生物，其包含权利要求 8 的质粒。

13. 一种转化微生物，其包含权利要求 9 的表达盒。
14. 权利要求 13 的转化微生物，其中所述表达盒为染色体整合型。
15. 权利要求 14 的转化微生物，它还含有合适的调节序列。
16. 权利要求 15 的转化微生物，其中所述合适调节序列包括：
 - a) 至少一个选自以下的启动子：大肠杆菌色氨酸操纵子启动子 P_{trp}、大肠杆菌乳糖操纵子启动子 P_{lac}、大肠杆菌 P_{tac} 启动子、λ噬菌体右启动子 P_R、λ噬菌体左启动子 P_L、T7 启动子、巴斯德毕赤酵母(*Pichia pastoris*) AOX1 基因启动子和巴斯德毕赤酵母 GAP 基因启动子；和
 - b) 至少一个得自λ噬菌体 CII 基因的核糖体结合位点或选自以下微生物基因的核糖体结合位点：丛毛单胞菌属、棒杆菌属、短杆菌属、红球菌属、固氮菌属、柠檬酸杆菌属、肠杆菌属、梭菌属、克鲁伯氏菌属、沙门氏菌属、乳杆菌属、曲霉属、酵母属、接合酵母属、毕赤酵母属、克鲁维酵母属、假丝酵母属、汉逊酵母属、*Dunaliella*、德巴利酵母属、白霉属、球拟酵母属、甲基营养菌属、芽孢杆菌属、埃希氏杆菌属、假单胞菌属、根瘤菌属和链霉菌属。
17. 权利要求 16 的转化微生物，其中所述宿主微生物为毕赤酵母属或埃希氏杆菌属。
18. 一种转化微生物，该微生物选自：
 - (a) 保藏号为 ATCC PTA-1175 的大肠杆菌 SW91；
 - (b) 保藏号为 ATCC PTA-1176 的大肠杆菌 DH5α:pnit4；以及
 - (c) 保藏号为 ATCC PTA-1177 的大肠杆菌 SS1001。
19. 一种将分子式为 NC-R-CN 的二腈酶促转变为羧酸的方法，其中 R 是具 1-10 个碳原子的亚烷基，该方法包括：
 - (a) 在合适条件下使表达权利要求 4 多肽的转化异源宿主接触所述二腈；和

- (b) 收集或不收集步骤(a)产生的羧酸。
20. 权利要求 19 的方法, 其中所述二腈为 2-甲基戊二腈。
21. 一种将分子式为 NC-R-CN 的二腈酶促转变为羧酸的方法, 其中 R 是具 1-10 个碳原子的亚烷基, 该方法包括:
- (a) 在合适条件下使包含权利要求 7 嵌合基因的转化异源宿主接触所述二腈, 从而制备得到羧酸; 和
- (b) 收集步骤(a)产生的羧酸。
22. 权利要求 21 的方法, 其中所述二腈为 2-甲基戊二腈。
23. 权利要求 21 的方法, 其中所述嵌合基因的合适调节序列包括诱导型启动子。
24. 权利要求 23 的方法, 所述步骤(a)的合适条件还包括存在诱导型启动子的诱导物。
25. 一种将 2-甲基戊二腈转变为相应羧酸的方法, 该方法包括:
- (a) 在合适条件下使命名为 ATCC PTA-1175 的大肠杆菌 SW91 接触 2-甲基戊二腈; 和
- (b) 收集步骤(a)产生的羧酸。
26. 一种由脂肪族 α, ω -二腈制备五元环内酰胺或六元环内酰胺的改进方法, 该方法包括:
- (a) 使脂肪族 α, ω -二腈在水性反应混合物中与酶催化剂接触, 从而将脂肪族 α, ω -二腈转变为 ω -氰羧酸铵盐, 所述脂肪族 α, ω -二腈的分子式为 $\text{NCCX}_a(\text{R})(\text{CH}_2)_n\text{CN}$, 其中 a 为 0 或 1, 当 a 为 1 时 X 为氢, 而 R 为 H、烷基或取代烷基、或链烯基或取代链烯基、或亚烷基或取代亚烷基, 而 n 为 1 或 2;
- (b) 使步骤(a)获得的水性产物混合物与氢和氢化催化剂接触, 从而将 ω -氰羧酸铵盐直接转变为相应的内酰胺, 而不用分离出中间体 ω -氰羧酸、 ω -氰羧酸铵盐、 ω -氨基羧酸或 ω -氨基羧酸铵盐; 和

(c) 由步骤(b)获得的水性产物混合物回收所述内酰胺,

所述改进包括在步骤(a)中使脂肪族 α, ω -二腈在水性反应混合物中与选自以下的酶催化剂接触:

- (1) 保藏号为 ATCC PTA-1175 的大肠杆菌 SW91;
- (2) 保藏号为 ATCC PTA-1176 的大肠杆菌 DH5 α :pnit4; 以及
- (3) 保藏号为 ATCC PTA-1177 的大肠杆菌 SS1001.

27. 权利要求 26 的方法, 其中所述 α, ω -二腈为 2-甲基戊二腈.

28. 权利要求 26 的方法, 其中所述脂肪族 α, ω -氰在 α -碳原子上不对称取代, 而所述酶催化剂的特征在于由于区域选择性水解 ω -氰基而产生 ω -氰羧酸铵盐的脂肪族腈水解酶活性, 由此在步骤(b)中仅产生两种可能的内酰胺产物中的一种内酰胺产物。

29. 权利要求 26 的方法, 该方法还包括在步骤(b)之前向包含 ω -氰羧酸铵盐的水性产物混合物中加入氢氧化铵、氨气或甲胺。

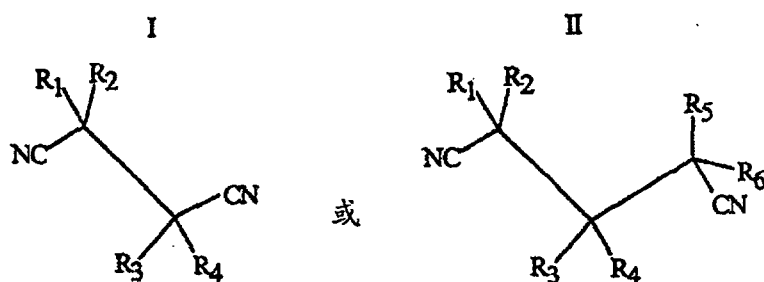
30. 权利要求 29 的方法, 其中所述加入到水性产物混合物中的氢氧化铵、氨气或甲胺的量相对于存在的 ω -氰羧酸铵盐量为 0-4 摩尔当量。

31. 权利要求 30 的方法, 其中在步骤(b)之前向含 ω -氰羧酸铵盐的水性产物混合物中加入甲胺, 而步骤(b)的产物为 N-甲基内酰胺。

32. 权利要求 26 的方法, 其中步骤(b)的水性产物混合物的温度为 45-200°C。

33. 权利要求 26 的由脂肪族 α, ω -二腈制备五元环内酰胺或六元环内酰胺的方法, 其中:

所述脂肪族 α, ω -二腈为以下任何一种结构式的二腈:



其中 R_1 和 R_2 都为 H, 而 R_3 、 R_4 、 R_5 和 R_6 各自独立选自 H、烷基或取代烷基、或链烯基或取代链烯基, 或者 R_3 和 R_4 一起为亚烷基或取代亚烷基, 或者 R_5 和 R_6 一起独立为亚烷基或取代亚烷基。

34. 权利要求 26 或 33 的方法, 其中所述酶催化剂为固定在不溶性支持体之中或之上的完整微生物细胞形式。

分离和表达敏捷食酸菌 72W 的腈水解酶基因

发明领域

本发明涉及分子生物学领域以及重组微生物表达目的基因和基因产物用途。更具体地说，本发明使用的基因是一种编码腈水解酶的新核酸片段，腈水解酶催化各种各样的含腈物质水解，产生相应的羧酸。本发明还提供表达腈水解酶活性并可用作水解含腈底物的生物催化剂的重组菌株。另外，本发明涉及帮助分离所述腈水解酶基因的特定核酸。

发明背景

腈容易通过各种化学方法转变为相应的羧酸，但这些方法通常需要强酸或强碱反应条件和高反应温度，并经常产生不需要的副产物和/或大量为不需要的无机盐废物。其中酶催化水解将含腈底物转变为相应羧酸的方法通常优于化学方法，因为这些方法 1)经常于室温进行；2)不需要使用强酸或强碱反应条件；和 3)不产生大量的不需要副产物。酶催化水解各种脂肪族或芳香族二腈特别优于化学水解之处在于高度区域选择性，即两个腈基团中仅有一个被水解为对应的羧酸铵盐。

可通过一步反应或两步反应(表 1)将腈底物酶催化水解为对应的羧酸。

表 1

	底物*	产物	酶
两步反应			
反应 1	$\text{RCN} + \text{H}_2\text{O}$	RC(O)NH_2	腈水合酶
反应 2	$\text{RC(O)NH}_2 + \text{H}_2\text{O}$	$\text{RC(O)OH} + \text{NH}_3$	酰胺酶
一步反应			
反应 1	$\text{RCN} + 2\text{H}_2\text{O}$	$\text{RC(O)OH} + \text{NH}_3$	腈水解酶

*R 表示特别针对选定酶的有机取代基变异谱。

各种细菌属总的来说具有不同的腈水合酶、酰胺酶或腈水解酶活性谱，所述细菌包括红球菌属(*Rhodococcus*)、假单胞菌属(*Pseudomonas*)、产碱菌属(*Alcaligenes*)、节杆菌属(*Arthrobacter*)、芽孢杆菌属(*Bacillus*)、杆菌属(*Bacteridium*)、短杆菌属(*Brevibacterium*)、棒杆菌属(*Corynebacterium*)、农杆菌属(*Agrobacterium*)、微球菌属(*Micrococcus*)和丛毛单胞菌属(*Comamonas*)。这些微生物的水性悬浮液以及分离的酶都已经用于将腈转变为羧酸。最近，Cowan 等(*Extremophiles* (1998) 2: 207-216)综述了这些酶的生物技术应用。

腈水解酶直接将腈转变为对应羧酸的水溶液，而没有形成酰胺中间体。多年以前就已经知道使用腈水解酶将芳香族腈水解为相应的羧酸铵盐，但仅是最近才报道使用腈水解酶转变脂肪族腈。Kobayashi 等(*Tetrahedron* (1990) 46: 5587-5590; *J. Bacteriology* (1990) 172: 4807-4815)描述由紫红红球菌(*Rhodococcus rhodochrous*) K22 分离脂肪族腈水解酶，该酶将脂肪族腈催化水解为相应的羧酸铵盐，还水解几种脂肪族 α, ω -二腈。已经由睾丸酮丛毛单胞菌(*Comamonas testosteroni*)分离出腈水解酶，该酶可以将一系列脂肪族 α, ω -二腈转变为相应的 ω -氰羧酸铵盐或二羧酸二铵盐(CA 2,103,616; 和 Lévy-Schil 等, *Gene* (1995) 161: 15-20)。

在由脂肪族 α, ω -二腈制备五元环或六元环内酰胺的方法中利用了未固定化敏捷食酸菌(*Acidovorax facilis*) 72W 细胞的腈水解酶活性(US 5,858,736)。在该方法中，首先使用具脂肪族腈水解酶(EC 3.5.5.7)活性的催化剂将脂肪族 α, ω -二腈转变为 ω -氰羧酸铵盐水溶液。然后通过在水溶液中氢化将 ω -氰羧酸铵盐直接转变为相应的内酰胺，而没有分离出中间体 ω -氰羧酸或 ω -氨基-羧酸。当脂肪族 α, ω -二腈还在 α -碳原子上不对称取代时，腈水解酶由于水解区域选择性大于 98% 的 ω -腈基团而产生 ω -氰羧酸铵盐，由此在随后的氢化过程中仅产生

两种可能的内酰胺产物中的一种。例如,用未固定化敏捷食酸菌 72W 细胞水解 2-甲基戊二腈(MGN),以 100%转化率、区域选择性超过 98% 产生 4-氰戊酸(4-CPA)铵盐。

已在外源系统(尤其是大肠杆菌)中克隆和表达了腈水解酶基因。Petre 等(美国专利第 5,635,391 号)公开了用大肠杆菌和恶臭假单胞菌(*Pseudomonas putida*)表达羧基酮丛毛单胞菌腈水解酶。还公开了一种通过共表达 GroE 伴侣蛋白改进大肠杆菌中可溶性腈水解酶蛋白水平的方法,该方法获得较高的腈水解酶比活性。使用大肠杆菌启动子 P_{lac} 和 P_{trp} 驱动腈水解酶编码序列表达。 P_{lac} 还在大肠杆菌中成功表达了粪产碱菌(*Alcaligenes faecalis*) JM3 (Kobayashi 等, *Proc. Natl. Acad. Sci.* (1993) 90: 247 和 JP#4-30663)、紫红红球菌 J1 (Kobayashi 等, *J. Biol. Chem.* (1992) 267: 20746)和紫红红球菌 K22 (Kobayashi 等, *Biochem.* (1992) 31, 9000)的腈水解酶编码序列。Stalker (美国专利第 4,810,648 号)公开了水解卤芳腈的腈水解酶基因可在大肠杆菌中在其天然启动子控制下表达。在美国专利第 5,602,014 号中, Mizumura 和 Yu 公开了一种用于在红串红球菌(*Rhodococcus erythropolis*)中表达腈水解酶基因的特化调节系统。

据报道,腈水解酶高度不稳定,并且不能大量获得(Kobayashi 等, *Tetrahedron* (1990) 46: 5587-5590; *J. Bacteriology* (1990) 172: 4807-4815)。与腈水合酶相反,腈水解酶特征在于低比活性和低反应速率(Nagasawa 等, *Appl Microbiol. Biotechnol.* (1993) 40: 189-195)。据报道,中温生物腈水解酶固有的热不稳定性限制了其工业应用(Cramp 等, *Microbiol.* (1997) 143: 2313-2320)。

目前问题是:在于温和反应条件(包括室温和没有过度的酸或碱条件)下获得高产产物且不产生较大量的不需要废物的应用(例如区域选择性水解脂肪族二腈为氰羧酸)中,缺少适用作含腈底物催化剂的工业上实用的热稳定性高产率腈水解酶。

发明概述

本发明提供食酸菌 72W 脲水解酶基因特有的分离 DNA 序列。本发明包括与所附序列表的完整序列互补的分离核酸片段以及基本类似的核酸序列。本发明包括编码 SEQ ID NO: 5 或 SEQ ID NO: 14 所列出的全部或大部分脲水解酶氨基酸序列的任何核酸片段。本发明还包括在以下杂交条件下与编码 SEQ ID NO: 5 或 SEQ ID NO: 14 的全部或大部分氨基酸序列的核酸片段杂交的分离核酸分子, 其中所述杂交条件为 $6 \times \text{SSC}$ (1 M NaCl)、40-45%甲酰胺、1% SDS、 37°C , 并在 $55-60^\circ\text{C}$ 用 $0.5 \times$ 至 $1 \times \text{SSC}$ 中清洗。本发明还要求保护与所列出的 SEQ ID NO: 1-16 互补的核酸片段(SEQ ID NO: 5 和 14 除外)。还提供由本发明核酸片段编码的多肽(例如 SEQ ID NO: 5 和 SEQ ID NO: 14)。本发明包括由以上描述的序列及其 cDNA 衍生物转录的 RNA 分子或反义 RNA 分子。

本发明还提供用转化微生物由这些脲水解酶编码序列生产活性酶蛋白的方法。由于与敏捷食酸菌 72W 相比转化体催化剂比活增加, 所以由本发明获得的改进是反应速率增加和反应物产率较高。用本文描述的遗传物质转化的完整细胞生产的活性脲水解酶可用于对含脲底物进行有效水解。含具嵌合基因的表达盒的转化微生物、具有含编码脲水解酶的分离核酸片段的质粒(每个单元都如本文所述)的转化微生物也是本发明的一部分。一个实施方案使用以下任一种特定酶催化剂与 α, ω - ω -羧酸接触: 大肠杆菌 SW91 (ATCC PTA-1175)、大肠杆菌 DH5 α :pnit4 (ATCC PTA-1176)、大肠杆菌 SS1001 (ATCC PTA-1177)、含质粒 pnitex2 的大肠杆菌 SS1002 或含质粒 pnitex2 的大肠杆菌 SS1011。另一个实施方案使用特定转化微生物以改进方法将 α, ω -二脲转变为 ω -羧酸, ω -羧酸是制备五元环或六元环内酰胺的中间体(见美国专利第 5,858,736 号)。

提供一种使用天然微生物(尤其是敏捷食酸菌 72W)基因获得突变微生物基因的方法, 所述天然微生物基因编码蛋白的特征在于对

含脘底物的脘水解酶活性，而突变微生物基因编码蛋白的特征在于对含脘底物的特异性脘水解酶活性增加和/或脘水解酶稳定性增加，其中一种或两种特征相对于天然微生物基因增加，生产所述突变微生物基因的方法包括以下步骤：

- (i) 使核苷酸序列混合物与限制性内切核酸酶接触，产生限制性片段混合物，所述核苷酸序列混合物包含：
 - a) 天然微生物基因；
 - b) 与(i)(a)的天然微生物基因核苷酸序列杂交的第一种核苷酸片段；
 - c) 不与(i)(a)的天然微生物基因核苷酸序列杂交的第二种核苷酸片段，
- (ii) 使步骤(i)的限制性片段混合物变性；
- (iii) 将步骤(ii)的限制性片段变性混合物与聚合酶温育；
- (iv) 重复步骤(ii)和(iii)足够的次数，以产生突变微生物基因，该基因编码蛋白的特征在于对含脘底物的特异性脘水解酶活性增加和/或脘水解酶稳定性增加，其中一种或两种特征相对于天然微生物基因增强。本发明包括用该方法产生的突变微生物基因。

附图、生物保藏和序列表简述

由合起来构成本申请的附图、生物保藏、所附序列描述、说明书详述以及权利要求书可更充分地理解本发明。

图 1 显示 pnit4 的 4.1 kb *Pst*I 片段的限制图谱。*Bcl*II-*Bgl*III 片段包含脘水解酶基因和侧翼染色体区。

本申请已经根据国际承认用于专利程序的微生物保存的布达佩斯条约的条款进行了以下的生物保藏：

保藏物分类命名	国际保藏号	保藏日
大肠杆菌 SW91	ATCC PTA-1175	2000 年 1 月 11 日
大肠杆菌 DH5 α : pnit4	ATCC PTA-1176	2000 年 1 月 11 日
大肠杆菌 SS1001	ATCC PTA-1177	2000 年 1 月 11 日

本文使用的“ATCC”是指位于 ATCC, 10801 University Blvd., Manassas, VA 20110-2209, USA 的国际保藏机构美国典型培养物保藏中心。“国际保藏号”是 ATCC 保藏的培养物的编号。

所列出的保藏物将在所指明的国际保藏机构保存至少 30 年，并可在其公开专利获准时让公众得到。保藏物的可利用性不等同于许可在损害政府机构授予的专利权的情况下实施本主题发明。

申请人提供的 32 个序列符合 37 C.F.R. 1.821-1.825 (“对含核苷酸序列和/或氨基酸序列公开内容的专利申请的要求-序列法规”)，并与世界知识产权组织(WIPO)标准 ST.25 (1998)以及 EPO 和 PCT 的序列要求(法规 5.2 和 49.5 (a-bis)以及管理章程第 208 章和附件 C)相一致。用于核苷酸和氨基酸序列数据的符号和格式遵从 37 C.F.R. § 1.822 给出的规则。

SEQ ID NO: 1 是正向引物(1F)的核苷酸序列，该序列得自位于细菌脲水解酶序列(可由 GenBank 数据获得)中的保守区，并用作简并 PCR 引物，以发现新脲水解酶基因。

SEQ ID NO: 2 是反向引物(7R)的核苷酸序列，该序列得自位于细菌脲水解酶序列(可由 GenBank 数据获得)中的保守区，并用作简并 PCR 引物，以发现新脲水解酶基因。

SEQ ID NO: 3 是 pJJ28-5 克隆敏捷食酸菌 72W 基因组部分的核苷酸序列，没有简并引物区。

SEQ ID NO: 4 是在 pnit4 的 4.1 kb 插入片段中鉴定出的敏捷食酸菌 72W 脲水解酶编码序列的核苷酸序列。

SEQ ID NO: 5 是由敏捷食酸菌 72W 脲水解酶编码序列编码的核苷酸序列(SEQ ID NO: 4, SEQ ID NO: 15)推定的氨基酸序列。

SEQ ID NO: 6 是用于扩增敏捷食酸菌 72W 基因组 DNA 的脲水解酶编码序列的正向引物的核苷酸序列。

SEQ ID NO: 7 是用于扩增敏捷食酸菌 72W 基因组 DNA 的脲水解酶编码序列的反向引物的核苷酸序列。

SEQ ID NO: 8 是用于扩增敏捷食酸菌 72W 基因组 DNA 的腓水解酶编码序列的正向引物的核苷酸序列。

SEQ ID NO: 9 是用于扩增敏捷食酸菌 72W 基因组 DNA 的腓水解酶编码序列的反向引物的核苷酸序列。

SEQ ID NO: 10 是用于扩增敏捷食酸菌 72W 基因组 DNA 的腓水解酶编码序列的正向引物的核苷酸序列。

SEQ ID NO: 11 是用于扩增敏捷食酸菌 72W 基因组 DNA 的腓水解酶编码序列的正向引物的核苷酸序列, 该序列还用于加入 *XhoI* 限制位点。

SEQ ID NO: 12 是用于扩增敏捷食酸菌 72W 基因组 DNA 的腓水解酶编码序列的反向引物的核苷酸序列, 该序列用于加入 *XhoI* 限制位点。

SEQ ID NO: 13 是用于在大肠杆菌中过表达的质粒 pnitex2 中的腓水解酶编码序列的核苷酸序列。

SEQ ID NO: 14 是由用于在大肠杆菌中过表达腓水解酶的质粒 pnitex2 中的腓水解酶编码序列编码的推定氨基酸序列。

SEQ ID NO: 15 是含德氏食酸菌(*A. delafieldii*)腓水解酶基因和侧翼染色体区的质粒 pnit4 中的 1776 bp *BclI-BglII* 基因组片段的核苷酸序列。

SEQ ID NO: 16 是合成形式的腓水解酶基因, 密码子选择对于在毕赤酵母(*Pichia*)中表达最佳。

SEQ ID NO: 17-32 是用于构建 SEQ ID NO: 16 的合成腓水解酶基因的寡核苷酸。

发明详述

申请人用涉及使用腓水解酶基因序列的发明解决了所述问题, 其中所述腓水解酶基因序列的编码蛋白用作将含腓底物水解为羧酸的生物催化剂。腓水解产物羧酸还可以为具有氨副产物或其它缓冲组分的羧酸盐形式。腓水解酶基因分离自敏捷食酸菌 72W。具体来

说, 表达活性敏捷食酸菌 72W 腈水解酶蛋白的重组菌株宿主细胞(例如大肠杆菌)对催化各种各样的含腈底物水解(包括高度区域选择性水解二腈)非常有用。典型的二腈分子式为 NC-R-CN, 其中 R 为具有约 1 至约 10 个碳原子的亚烷基基团。

该生物催化方法在工业上引人注目, 因为与先前已知的方法相比, 该方法产生的不需要副产物的量要小得多, 并且在温和得多的条件下操作。本发明的产物在化学、农业和医药工业中用作高价值的聚合物、溶剂和化学品的前体。

申请人对技术领域所存在问题的解决在于下文和实施例更详细论述的下列成果:

I. 提供用于鉴定细菌中是否存在未知腈水解酶基因的简并性 PCR 引物序列(SEQ ID NO: 1 和 2);

II. 用于筛选存在的任何腈水解酶基因的分离核酸序列;

III. 作图、鉴定和克隆了敏捷食酸菌 72W (ATCC 55746)的区域选择性热稳定腈水解酶的完整编码序列 (SEQ ID NO: 4 和 15);

IV. 构建重组 DNA 质粒 pSW91、pnit4 和 pnitex2、嵌合基因以及含如 III 所述位于敏捷食酸菌 72W (ATCC 55746)的 4.1 kb 染色体 DNA 片段中的编码序列的表达盒;

V. 构建表达活性敏捷食酸菌 72W 腈水解酶蛋白的重组大肠杆菌菌株;

VI. 证实使用本文描述的表达敏捷食酸菌 72W 腈水解酶蛋白的重组大肠杆菌完整细胞将 2-甲基戊二腈(MGN)区域选择性水解为 4-氰戊酸(4-CPA); 并指导使用天然微生物基因获得突变微生物基因, 所述天然微生物基因编码蛋白的特征在于对含腈底物(优选为 2-甲基戊二腈)的腈水解酶活性, 而突变微生物基因编码蛋白的特征在于腈水解酶活性改变和/或稳定性增加, 其中一种或两种特征相对于天然腈水解酶蛋白增加。

敏捷食酸菌 72W 腈水解酶被证明是由脂肪族或芳香族腈生产羧

酸的出乎意料强的催化剂。所有已知的脲水解酶(包括敏捷食酸菌 72W 脲水解酶)都在酶活性部位具有亲核半胱氨酸(Cowan 等, (1998), 同上), 并且全部对硫醇试剂(1.0 mM 浓度的氯化铜、硝酸银、醋酸汞或氯化铁, 每种都使 72W 脲水解酶活性大幅降低)灭活敏感。半胱氨酸残基还能够不可逆地氧化为亚磺酸, 导致酶活性丧失。尽管脲水解酶对各种灭活机制敏感, 但是用作催化剂的固定化敏捷食酸菌 72W 细胞产生的产物(4-CPA)仍高达 3.9×10^7 摩尔 4-CPA/摩尔脲水解酶(总转化量, TTN)。

在外源宿主细胞中表达活性敏捷食酸菌 72W 脲水解酶相比于制备和使用敏捷食酸菌 72W 细胞作为脲水解酶催化剂具有若干种其它优势。敏捷食酸菌 72W 的脲水解酶蛋白表达水平大约为总可溶性蛋白的 3.4% (实施例 14)并且为组成型表达; 充分筛选脲水解酶生产(脂肪族或芳香族脲或酰胺、内酰胺、尿素等)的潜在诱导物发现, 诱导物对敏捷食酸菌 72W 发酵过程中产生的脲水解酶水平不起作用。

相反, 大肠杆菌转化体表达的具酶活性的敏捷食酸菌 72W 脲水解酶水平高达总可溶性蛋白的 12%; 大肠杆菌转化体产生的脲水解酶总量(有活性或无活性)高达总可溶性蛋白的 58%。这种大肠杆菌转化细胞的脲水解酶表达水平升高使比活(脲水解酶活性单位/g 干细胞重量)相比于敏捷食酸菌 72W 细胞增加高达 2.4 倍之多, 而这又显著增加了催化剂产率(g 生产产物/g 干细胞重量/小时)(实施例 7 和 9)。而且, 敏捷食酸菌 72W 细胞在发酵培养时需要甘油, 甘油是一种相对昂贵的碳源, 而使用便宜的葡萄糖还没有成功培养过。相比之下, 大肠杆菌转化体可用葡萄糖在大约一半时间培养至和敏捷食酸菌 72W 细胞相同的细胞密度, 显著降低了生物催化剂的生产成本。

敏捷食酸菌 72W 细胞还含有脲水合酶以及酰胺酶, 他们都不具有区域选择性, 并且在将 α, ω -二脲转变为相应的 ω -氰羧酸铵盐时可产生不需要的副产物。敏捷食酸菌 72W 细胞需要热处理以灭活脲水合酶以及酰胺酶(美国专利第 5,814,508 号), 存在损失脲水解酶活性

的风险，并增加了生产成本。有人制备了没有脲水合酶的敏捷食酸菌突变菌株 72-PF-15 和 72-PF-17 (美国专利第 5,858,736 号)，但这些菌株仅具有大约一半的亲代敏捷食酸菌 72W 菌株脲水解酶比活。

相比之下，用表达脲水解酶的遗传物质转化的大肠杆菌不需要热处理步骤来灭活不需要的脲水合酶以及酰胺酶活性，削减了该工艺催化剂生产步骤的成本，并避免了脲水解酶活性的无意损失。申请人已按实施例 6 和 7 所述获得具有高脲水解酶比活性以及随之而来的优势的大肠杆菌转化体。

敏捷食酸菌 72W 的完整细胞脲水解酶活性在高达约 55℃ 时也非常稳定；在 0.10 M 磷酸缓冲液(pH 7.0)中的细胞悬浮液在 50℃ 时的脲水解酶半寿期为 22.7 小时(Gavagan 等, *Appl. Microbiol. Biotechnol.* (1999) 52: 654-659)。纯化酶也具有极佳的温度稳定性，10 mg/ml 纯化脲水解酶的 0.10 M 磷酸缓冲溶液在 45℃、24 小时后没有观察到活性损失。当以 50 mM 磷酸钾缓冲溶液(pH 7.0)于 5℃ 储存时，46 天后没有观察到纯化脲水解酶活性损失。尽管敏捷食酸菌 72W (一种中温细菌)最佳培养温度为 32℃，但 72W 脲水解酶自身具有热稳定性，优于嗜热细菌如苍白芽孢杆菌(*Bacillus pallidus*)菌株 DAC521 的脲水解酶(Cramp 等, *Microbiol.* (1997) 13: 2313-2320)。

当使用未固定化细胞作为 MGN 水解为 4-CPA 的催化剂时，需要离心或超滤回收未固定化细胞以便再使用。在产物浓度高(高达 29% (重量)4-CPA 铵盐)时，每次再使用未固定化细胞活性明显下降，并观察到细胞裂解。相反，固定化细胞简化了催化剂回收和再使用，改进了细胞对裂解的抵抗力，并且与使用未固定化细胞相比在连续分批反应重复利用时增加了固定化细胞酶活性的稳定性。

用角叉胶在相同条件下固定敏捷食酸菌 72W 完整细胞或大肠杆菌转化体 SS1001 产生的胶体在生产高浓度 4-CPA 铵盐(高达约 200 g/L)的连续分批反应中重复使用时是稳定的。首先将加热的细胞水性悬浮液(5%干细胞重量)和角叉胶(3%重量)于 50℃ 分散于加热的豆油

中, 通过将油温降低至角叉胶成胶温度以下使获得的液滴成胶(Audet等, *Process Biochem.* (1989) 24: 217), 由此制备胶珠。由豆油中分离平均直径为 0.5 mm-3 mm 的细胞/角叉胶珠, 然后用水性缓冲液清洗, 并用戊二醛和聚乙烯亚胺交联。

在用于生产 1.25 M 4-CPA 铵盐的相同条件下, 以平行反应比较在角叉胶珠中固定化的敏捷食酸菌 72W 细胞和大肠杆菌转化体 SS1001 细胞。在使用固定化大肠杆菌转化体 SS1001 时反应速率比使用固定化敏捷食酸菌 72W 细胞高 1.7 倍(分别为 310 mM 4-CPA 铵盐/小时和 184 mM 4-CPA 铵盐/小时; 实施例 9)。大肠杆菌转化体 SS1001 获得的反应速率增加的原因是固定化转化细胞催化剂的比活(脘水解酶活性单位/g 珠)较高, 显示了使用脘水解酶比活高于亲代敏捷食酸菌 72W 菌株的转化细胞的一个优势。当使用按照公开方法(Bucke, *Methods Enzymol.* (1987) 135: 175-189)用海藻酸钙珠固定的大肠杆菌转化体 SW91 细胞时, 反应速率比使用角叉胶固定的敏捷食酸菌 72W 细胞高 3.0 倍(分别为 739 mM 4-CPA 铵盐/小时和 184 mM 4-CPA 铵盐/小时; 实施例 9 和 11)。通过使用固定化转化催化剂使催化产率(g 产物/g 催化剂/小时)增加显著降低了生产成本。

本发明提供新脘水解酶编码核酸序列。该序列包含一个位于分离自敏捷食酸菌 72W 基因组 DNA 的 4.1 kb *Pst*I 片段上的可读框(ORF)。新确定的 ORF 编码一种可证明将脘转变为相应的羧酸铵盐的酶。鉴定所述 ORF 既基于表达活性脘水解酶, 也基于使用本领域众所周知的算法与公共数据库比较核酸序列以及推断的氨基酸序列。由所述 ORF 编码的蛋白显示与其它已知脘水解酶的一级氨基酸序列同一性高达 71%。

因此, 本发明优选多肽是那些与本文报道的氨基酸序列同一性至少为 80%的活性蛋白。更优选的氨基酸片段与本文序列至少 90%相同。最优选的氨基酸片段与本文报道的氨基酸片段的同一性至少为 90%。同样, 对应于所述 ORF 的优选核酸序列是那些编码活性蛋

白并与本文报道的核酸序列至少 80%相同的序列。更优选的核酸片段与本文序列至少 90%相同。最优选的核酸片段与本文报道的核酸片段至少 95%相同。

本发明的核酸片段可用于分离相同或其它细菌物种的编码同源酶的 cDNA 和基因。使用依赖于序列的方法分离同源基因是本领域众所周知的。依赖于序列的方法的实例包括但不限于核酸杂交方法以及以核酸扩增技术(例如 PCR、连接酶链式反应)的各种应用为代表的 DNA 和 RNA 扩增方法。

通常,对于推定多肽或核酸序列与已知蛋白或基因同源而言,必须 10 个或 10 个以上连续氨基酸的序列或者 30 个或 30 个以上核苷酸的序列。而且,就核苷酸序列来说,含 20-30 个连续核苷酸的基因特异性寡核苷酸探针可用于基因鉴定(例如 DNA 杂交)和分离(例如细菌菌落或噬菌斑的原位杂交)的序列依赖性方法。另外,12-15 个碱基的短寡核苷酸可用作 PCR 扩增引物,以便获得含所述引物的特定核酸片段。

例如,可通过使用本领域技术人员周知的方法,使用全部或部分本发明核酸片段作为 DNA 杂交探针筛选任何目的细菌文库,直接分离编码与本发明的酶相似的酶的基因,该基因或者为 cDNA 或者为基因组 DNA。可利用本领域已知的方法(Sambrook, J., Fritsch, E.F. 和 Maniatis, T., *Molecular Cloning: A Laboratory Manual* (1989), Cold Spring Harbor Laboratory Press, Cold Spring Harbor (本文统称为“Maniatis”))设计和合成基于本发明核酸序列的特异性寡核苷酸探针。而且,可利用技术人员已知的方法如随机引物 DNA 标记、缺口翻译或末端标记技术使用完整序列直接合成 DNA 探针,或者可利用可获得的体外转录系统使用完整序列直接合成 RNA 探针。另外,可设计特异性引物并用于扩增部分或全长本发明序列。获得的扩增产物可在扩增反应期间直接标记,或者在扩增反应后标记,并可用作在适当严格条件下分离全长 cDNA 或基因组片段的探针。

本发明 ORF 的两个短节段可用于 PCR 方法，以扩增编码 DNA 或 RNA 同源基因的较长核酸片段。另一方面，第二个引物序列可基于衍生自克隆载体的序列。例如，技术人员可按照 RACE 方法，通过使用 PCR 扩增位于转录物单个点和 3' 或 5' 末端之间的区域的拷贝，生产 cDNA。可由本发明序列设计 3'-5' 方向的引物。使用市售的 3' RACE 或 5' RACE 系统(BRL)，可分离特异性 3' 或 5' cDNA 片段(Ohara 等, *PNAS USA* (1989) 86: 5673; Loh 等, *Science* (1989) 243: 217)。

通常，在 PCR 型扩增技术中，引物具有不同的序列并且互相之间不互补。应当根据需要的测试条件设计所述引物序列，以便有效正确复制的靶核酸。PCR 引物设计方法是本领域众所周知的普通技术。(Thein 和 Wallace, “在遗传疾病诊断中使用寡核苷酸作为特异性杂交探针”，(1986) 第 33-50 页，载于: *Human Genetic Diseases: A Practical Approach*, K.E. Davis (主编), IRS Press, Herndon, Virginia); Rychlik, W., PCR 方案: 当前方法和应用, (1993) 15: 31-39, 载于: *Methods in Molecular Biology*, B.A. White, (主编), Humana Press, Inc., Totowa, NJ)。

另一方面，本发明序列可用作鉴定同源物的杂交试剂。核酸杂交测试的基本成分包括探针、怀疑含有目的基因或基因片段的样品以及特异性杂交方法。本发明的探针通常为单链核酸序列，与要检测的核酸序列互补。探针可与要检测的核酸序列(cDNA、基因组 DNA 或 RNA) “杂交”。探针长度可由 5 个碱基至数万碱基不等，长度取决于要进行的具体测试。探针分子仅有一部分需要与要检测的核酸序列互补。另外，探针和靶序列之间不需要绝对互补。在不完全互补的分子之间确实发生杂交，结果杂交区域的某一部分碱基不与正确互补的碱基配对。

杂交方法是经过充分定义的(Maniatis, 特别是第 11 章和表 11.1)。通常，探针和样品必须在允许核酸杂交的条件下混合。这包括在合适温度和离子强度压条件下存在无机盐或有机盐时使探针和

样品接触。探针和样品核酸必须接触足够长时间，使得探针和样品核酸之间任何可能的杂交都可发生。混合物中探针和靶的浓度将决定发生杂交所需要的时间。探针和靶的浓度越大，所需要的杂交温育时间就越短。可选地可添加离液剂。离液剂通过抑制核酸酶活性稳定核酸。而且，离液剂允许短寡核苷酸探针在室温下发生敏感且严格的杂交(Van Ness 等, *Nucl. Acids Res.* (1991) 19: 5143-5151)。合适的离液剂包括氯化胍、硫氰酸胍、硫氰酸钠、四氯乙酸锂、高氯酸钠、四氯乙酸铷、碘化钾和三氟乙酸铯等。通常，离液剂的终浓度约为 3M。如果需要，可将甲酰胺加入到杂交混合物中，通常为 30-50% (体积/体积)。

可以使用各种杂交溶液。通常，这些溶液包含约 20%-60% (优选 30%) 体积的极性有机溶剂。常用的杂交溶液使用约 30-50% (体积/体积) 的甲酰胺、约 0.15-1 M 的氯化钠、约 0.05-0.1 M 缓冲液(如柠檬酸钠、Tris-HCl、PIPES 或 HEPES (pH 范围约 6-9))、约 0.05-0.2% 变性剂如 SDS、或 0.5-20 mM EDTA、FICOLL (Pharmacia Inc.)(约 300-500 kd)、聚丙烯吡咯烷酮(约 250-500 kd)和血清白蛋白。典型的杂交溶液还包括约 0.1-5 mg/mL 未标记载体核酸、片段化核 DNA 例如牛胸腺 DNA 或鲑鱼精 DNA、或酵母 RNA 以及可选的约 0.5-2% (重量/体积) 甘氨酸。还可以包括其它添加剂，例如体积排阻剂，包括各种极性水溶性介质或可膨胀剂，例如聚乙二醇、阴离子聚合物如聚丙烯酸酯或聚甲基丙烯酸酯以及阴离子糖聚合物如硫酸葡聚糖。

核酸杂交适用于各种测定形式。最适宜的一种是夹心测定形式。夹心测定特别适合于非变性条件下的杂交。夹心型测定的主要组分是固相支持物。固相支持物吸附或共价偶联未标记并与一部分序列互补的固定化核酸探针。

本发明核苷酸序列和推定氨基酸序列的可用性促进了对 cDNA 表达文库的免疫筛选。可以合成代表部分本发明氨基酸序列的合成肽。这些肽可用于免疫动物以产生特异性针对含所述氨基酸序列的

肽或蛋白的多克隆或单克隆抗体。然后这些抗体可用于筛选 cDNA 表达文库，以分离想要的全长 cDNA 克隆(Lerner, *Adv. Immunol.* 36: 1 (1984); 和 Maniatis)。

异源宿主细胞:

可用异源宿主细胞(优选微生物细胞)生产活性敏捷食酸菌 72W 脘水解酶蛋白。对本发明特别有用的是容易适应大规模发酵方法的细胞。这样的微生物是工业生物工程领域众所周知的，其实例可见“Recombinant Microbes for Industrial and Agricultural Applications”，Murooka 等主编, Marcel Dekker, Inc., New York, NY (1994)，包括发酵细菌以及酵母和丝状真菌。宿主细胞可包括但不限于丛毛单胞菌属(*Comamonas sp.*)、棒杆菌属(*Corynebacterium sp.*)、短杆菌属(*Brevibacterium sp.*)、红球菌属(*Rhodococcus sp.*)、固氮菌属(*Azotobacter sp.*)、柠檬酸杆菌属(*Citrobacter sp.*)、肠杆菌属(*Enterobacter sp.*)、梭菌属(*Clostridium sp.*)、克鲁伯氏菌属(*Klebsiella sp.*)、沙门氏菌属(*Salmonella sp.*)、乳杆菌属(*Lactobacillus sp.*)、曲霉属(*Aspergillus sp.*)、酵母属(*Saccharomyces sp.*)、接合酵母属(*Zygosaccharomyces sp.*)、毕赤酵母属(*Pichia sp.*)、克鲁维酵母属(*Kluyveromyces sp.*)、假丝酵母属(*Candida sp.*)、汉逊酵母属(*Hansenula sp.*)、*Dunaliella*、德巴利酵母属(*Debaryomyces sp.*)、白霉属(*Mucor sp.*)、球拟酵母属(*Torulopsis sp.*)、甲基营养菌属(*Methylobacteria sp.*)、芽孢杆菌属(*Bacillus sp.*)、埃希氏杆菌属(*Escherichia sp.*)、假单胞菌属(*Pseudomonas sp.*)、根瘤菌属(*Rhizobium sp.*)和链霉菌属(*Streptomyces sp.*)的细胞。特别优选大肠杆菌。可表达脘水解酶基因的合适大肠杆菌宿主细胞的实例包括但不限于本文列举的宿主细胞和 MG1655 (ATCC 47076)、W3110 (ATCC 27325)、MC4100 (ATCC 35695)、W1485 (ATCC 12435)及其衍生物。

微生物表达系统和表达载体:

含控制外源蛋白高水平表达的调节序列的微生物表达系统和表达载体是本领域技术人员众所周知的。它们可用于构建用于生产本发明 4.1 kb 片段基因产物的嵌合基因。然后可通过转化将这些嵌合基因引入合适的微生物中, 以提供高水平表达的脲水解酶。本发明的核苷酸可用于生产其活性水平相对于天然基因序列增强或改变的基因产物。

另外, 嵌合基因有效改变宿主细胞的特性。例如, 将在合适启动子控制下编码本发明 ORF 的至少一个拷贝的嵌合基因引入到宿主细胞中, 获得的宿主细胞能够将 2-甲基戊二脲转变为 4-氰戊酸。本发明的嵌合基因包含适用于驱动本发明脲水解酶序列基因表达的调节序列。调节序列包括但不限于启动子、翻译前导序列以及核糖体结合位点。这些序列最好得自宿主生物; 但是, 技术人员会认识到还可使用异源调节序列。

在一个实施方案中, 所述调节序列包括启动子。启动子可以为组成型或诱导型。诱导型启动子一般对特异性刺激物(例如诱导 *lac* 启动子的 IPTG)起反应。诱导型启动子可对各种刺激物起反应, 包括化学物质、生活周期、温度改变、pH 改变以及重量摩尔渗透压浓度, 在此仅举几例。在本发明的一个优选实施方案中, 发现含脲水解酶编码区(例如 SEQ ID NO: 4、SEQ ID NO: 13 或 SEQ ID NO: 15)的嵌合基因在调节区含诱导型启动子时在没有诱导物的情况下有效表达或更有效表达。

通过将嵌合基因克隆到合适的表达载体中将其引入到合适的宿主中。用于转化合适宿主细胞的载体或表达盒本领域众所周知。通常, 所述载体或表达盒包含控制相关基因转录和翻译的序列、选择标记以及允许自主复制或染色体整合的序列。合适的载体包括具有控制转录起始的编码序列 5'区和控制转录终止的 DNA 片段 3'区。尽管所述控制区不需要得自选定为生产宿主的特定物种的天然基因,

但最优选两个控制区都得自宿主细胞同源基因。

用于驱动所述 ORF 在目的宿主细胞中表达的起始控制区或启动子为数众多, 本领域技术人员对此很熟悉, 包括但不限于 CYC1、HIS3、GAL1、GAL10、ADH1、PGK、PHO5、GAPDH、ADC1、TRP1、URA3、LEU2、ENO、TPI (用于在酵母表达); AOX1 (用于毕赤酵母表达) 以及 lac、trp、 IP_L 、 IP_R 、T7、tac、 P_{BAD} 和 trc (用于大肠杆菌表达)。实例包括至少一个选自以下的启动子: 大肠杆菌的色氨酸操纵子启动子 P_{trp}、大肠杆菌的乳糖操纵子启动子 Plac、大肠杆菌的 Ptac 启动子、 λ 噬菌体右启动子 PR、 λ 噬菌体左启动子 PL、T7 启动子、巴斯德毕赤酵母(*Pichia pastoris*) GAP 基因启动子, 或者为至少一个选自以下微生物的强启动子: 丛毛单胞菌属、棒杆菌属、短杆菌属、红球菌属、固氮菌属、柠檬酸杆菌属、肠杆菌属、梭菌属、克鲁伯氏菌属、沙门氏菌属、乳杆菌属、曲霉属、酵母属、毕赤酵母属、接合酵母属、克鲁维酵母属、假丝酵母属、汉逊酵母属、*Dunaliella*、德巴利酵母属、白霉属、球拟酵母属、甲基营养菌属、芽孢杆菌属、埃希氏杆菌属、假单胞菌属、根瘤菌属和链霉菌属。

终止控制区还可以得自优选宿主的各种天然基因。作为选择的是, 终止位点为非必须的, 但是最优选包括终止位点。

另外, 插入的遗传物质可包括核糖体结合位点。核糖体结合位点可得自 λ 噬菌体 CII 基因或选自以下微生物基因的核糖体结合位点: 丛毛单胞菌属、棒杆菌属、短杆菌属、红球菌属、固氮菌属、柠檬酸杆菌属、肠杆菌属、梭菌属、克鲁伯氏菌属、沙门氏菌属、乳杆菌属、曲霉属、酵母属、接合酵母属、毕赤酵母菌属、克鲁维酵母菌属、假丝酵母属、汉逊酵母属、*Dunaliella*、德巴利酵母属、白霉属、球拟酵母属、甲基营养菌属、杆菌属、埃希氏杆菌属、假单胞菌属、根瘤菌属和链霉菌属。

可选地, 所述基因产物可优选为转化宿主的分泌产物。分泌目的蛋白至培养基中简化了纯化步骤并降低了成本。分泌信号序列经

常用于促使表达蛋白穿过细胞膜主动转运。可通过将编码分泌信号的 DNA 序列掺入到宿主中产生能够分泌的转化宿主。选择合适信号序列的方法是本领域众所周知的(参阅例如 EP 546049、WO 9324631)。分泌信号 DNA 可位于表达控制 DNA 和所述编码序列或编码序列片段之间,并与后者一起位于读框中。

可以分批、补料分批或连续方式进行的发酵是本领域众所周知的普通技术(Thomas D. Brock, 载于: *Biotechnology: A Textbook of Industrial Microbiology*, 第二版 (1989) Sinauer Associates, Inc., Sunderland 等, *Appl. Biochem. Biotechnol.* (1992) 36: 227)。

本发明的核苷酸片段可用于生产活性水平增强或改变的基因产物。具体来说,所述核苷酸片段可用于生产突变微生物基因,该基因编码蛋白的特征在于对含腈底物的腈水解酶活性增加或酶稳定性增加,其中一种或两种特征相对于天然微生物基因的腈水解酶活性增加。已知多种突变天然基因序列以生产活性相对于天然基因序列改变或增加的基因产物的方法。这些方法包括但不限于定向进化、随机诱变、结构域交换(使用锌指结构域或限制酶)、合理设计、易错 PCR (Melnikov 等, *Nucleic Acids Research* (1999) 27 (4): 1056-1062)、定点诱变(Coomb 等, 载于: *Proteins* (1998) 259-311, R.H. Angeletti (主编), Academic Press, San Diego, CA)以及“基因改组”(美国专利第 5,605,793、5,811,238、5,830,721 和 5,837,458 号,这些专利通过引用结合到本文中)。

缩写和术语的定义:

在本说明书中,许多术语和缩写以以下定义的方式使用。

本说明书中的缩写等于检测单位、技术、特性或化合物,具体如下:“sec”是指秒;“min”是指分钟;“h”是指小时;“d”是指天;“ μ L”是指微升;“mL”是指毫升;“L”是指升;“mM”是指毫摩尔;“M”是指摩尔;“mmol”是指毫摩尔;“Ampr”是

指氨苄青霉素抗性；“Amps”是指氨苄青霉素敏感；“kb”是指千碱基；“kd”是指千道尔顿；“nm”是指纳米；而“wt”是指重量。“ORF”是指可读框；“PCR”是指聚合酶链反应；“SSC”是指柠檬酸钠盐水缓冲液；“HPLC”是指高效液相层析；“ca”是指大约；“dcw”是指干细胞重量；“O.D.”是指指定波长的光密度；“IU”是指国际单位；“MGN”是指 2-甲基戊二腈；“4-CPA”是指 4-氰戊酸，而“IPTG”是指异丙基 β -D-硫代吡喃半乳糖苷。

“酶催化剂”是指特征为对含腈底物具有特异性腈水解酶活性的催化剂。

“氢化催化剂”是指加速氢化但自身不被消耗或产生化学变化的物质。适用于本发明的氢化催化剂包括但不限于各种铂金属，例如铱、钼、铑、钌、铂和钯，还包括各种其它过渡金属，例如钴、铜、镍和锌。所述催化剂可以为非支持体型(例如为阮内镍或氧化铂)，或者可以为支持体型(例如为披钯碳、披铂氧化铝或披镍硅藻土)。

术语“宿主细胞”和“宿主生物”是指能够接受外源或异源基因、基因片段或 DNA 片段的细胞。

术语“中温菌”是指生活在接近温血动物温度范围的细菌，通常表现出的最佳生长温度为 25-40℃。

术语“重组生物”、“转化宿主”、“转化体”、“转基因生物”和“转化微生物宿主”是指已经用异源或外源 DNA 转化的宿主生物。本发明的重组生物表达编码活性腈水解酶的外源编码序列或基因。“转化”是指将 DNA 片段转移至宿主生物基因组获得遗传上稳定的遗传。“转化盒”是指包含一组遗传元件的特定 DNA 片段，这些遗传元件的排列便于插入到宿主细胞中，通常是作为质粒的一部分。“表达盒”是指包含一组遗传元件的特定 DNA 片段，这些遗传元件的排列便于插入到宿主细胞中，通常是作为质粒的一部分，还可增强在宿主中的基因表达。

术语“质粒”和“载体”是指其携带基因往往不是宿主细胞中

心代谢的一部分的染色体外元件, 该元件通常为环状双链 DNA 分子形式。这样的元件可以为自主复制序列、基因组整合序列、噬菌体或核苷酸序列, 它们为得自任何来源的线性或环状的单链或双链 DNA 或 RNA, 其中许多核苷酸序列连接或重组为独特的结构。

术语“核酸”是指活体细胞存在的高分子量复合化合物, 其基本单位是与磷酸桥连接在一起的核苷酸。核酸可再分为两种类型: 核糖核酸(RNA)和脱氧核糖核酸(DNA)。

字母“A”、“G”、“T”和“C”当表示核酸时分别是指嘌呤碱基(腺嘌呤($C_5H_5N_5$)和鸟嘌呤($C_5H_5N_5O$))和嘧啶碱基(胸腺嘧啶($C_5H_6N_2O_2$)和胞嘧啶($C_4H_5N_3O$))。

术语“互补”用于描述互相杂交的核苷酸碱基之间的关系。例如, 对 DNA 来说, 腺嘌呤与胸腺嘧啶互补, 而胞嘧啶与鸟嘌呤互补。cDNA 是与 RNA 互补的单链 DNA, cDNA 通过反转录酶由 RNA 体外合成。反义 RNA 是与另一个 RNA 互补的 RNA 分子。

术语“编码序列”或“编码区”是指编码特定氨基酸序列的 DNA 序列。术语“ORF”和“可读框”以及“编码序列”和“编码区”可交互使用, 用来指翻译为蛋白的 DNA 序列部分。通常用在 DNA 序列翻译为蛋白质序列时称为起始(起始密码子)的三个碱基对和称为终止(终止密码子)的三个碱基对划定 ORF 在序列中的界限。

术语“核酸片段”或“核苷酸片段”是指可编码基因和/或编码序列之前(5', 上游)或之后(3', 下游)的调节序列的 DNA 片段。“片段”为特定区段完整核酸序列的一部分。一个片段可以构成一个完整基因。

术语“限制性核酸内切酶”和“限制酶”是指催化双链 DNA 中特定核苷酸序列水解断裂的酶。

术语“寡核苷酸”是指引物、探针、要检测的寡聚物片段、标记的复制封闭探针以及寡聚物对照, 一般来说是指多脱氧核糖核苷酸(包括 2-脱氧-D-核糖)、多核糖核苷酸(包括 D-核糖)以及为嘌呤或

嘧啶碱基(核苷酸)或修饰的嘌呤或嘧啶碱基的 N-糖苷的任何多核苷酸。“寡核苷酸”定义还包括核酸类似物(例如肽核酸)以及结构上修饰(例如硫代磷酸酯键)的核酸(另见 Thuong 等, *Biochimie* (1985) 7-8 月 67 (7-8): 673-684)。没有特意区别“核酸”、“多核苷酸”或“寡核苷酸”的长度。

术语“引物”是指寡核苷酸(合成或天然), 当将所述寡核苷酸放置于通过聚合酶催化互补链合成的条件下时, 所述寡核苷酸起沿着互补链合成或复制核酸的起点作用。

术语“探针”是指寡核苷酸(合成或天然), 其与“片段”显著互补, 并通过与所述片段的至少一条链杂交形成双链结构。

“合适的调节序列”是指影响关联编码序列转录、RNA 加工、RNA 稳定性或翻译的核苷酸序列, 其位于编码序列的上游(5'非编码序列)、之中或下游(3'非编码序列)。调节序列可包括启动子、翻译前导序列、内含子以及多腺苷酸识别序列。

“启动子”是指能够控制编码序列或功能 RNA 表达的 DNA 序列。一般来说, 编码序列位于启动子序列的 3'。启动子可整体得自天然基因, 或者可由得自天然存在的不同启动子的不同元件组成, 甚至可包含合成 DNA 节段。本领域技术人员可以理解, 不同启动子可控制基因在不同组织或细胞类型中表达, 或在不同的发育阶段表达, 或在不同环境条件作用下表达。使基因在大部分时间或在大多数条件下在大多数细胞类型中表达的启动子通常称为“组成型启动子”。使基因仅在存在特定化合物或环境条件时表达的启动子通常称为“诱导型启动子”。因为在大多数情况下调节序列的确切边界没有完全确定, 所以不同长度的 DNA 片段可能具有相同的启动子活性。

术语“有效连接”是指单个核酸片段上的核酸序列连接, 使得一个序列的功能受另一个影响。例如, 如果启动子能够影响编码序列表达(即编码序列处于启动子转录控制之下), 那么启动子就与编码

序列有效连接。编码序列可以有义或反义方向有效连接至调节序列。

“3'非编码序列”是指位于编码序列下游的 DNA 序列，包括多腺苷酸识别序列和编码能够影响 mRNA 加工或基因表达的调节信号的其它序列。多腺苷酸信号的特征通常为影响多腺苷酸序列加入到 mRNA 前体的 3'末端。

“基因”是指表达特定蛋白的核酸片段，包括编码序列之前(5'非编码序列)和之后(3'非编码序列)的调节序列。“天然基因”是指其自身调节序列按其天然可见排列的基因。“嵌合基因”是指任何非天然基因的基因，包括未发现天然在一起的调节序列和编码序列。因此，嵌合基因可以包括得自不同来源的调节序列和编码序列，或者得自相同来源但以不同于天然可见方式排列的调节序列和编码序列。“内源基因”是指在天然生物基因组中在其天然位置的天然基因。“外源”基因是指在宿主生物中通常不存在但其通过基因转移引入到宿主生物中的基因。外源基因可包含插入到非天然生物中的天然基因或嵌和基因。“转基因”是已经通过转化方法引入到基因组中的基因。

“合成基因”可由使用本领域技术人员已知的方法化学合成的寡核苷酸构件组装而成。连接这些构件并退火，以形成基因节段，然后酶促组装以构建完整基因。“化学合成”在论及 DNA 序列时是指体外装配核苷酸组分。可使用已充分确立的方法完成 DNA 的人工化学合成，或者可使用众多市售机器中的一种进行自动化学合成。

技术人员完全了解使用给定氨基酸特有的核苷酸密码子时特定宿主细胞表现出的“密码子偏倚”。因此，当在宿主细胞中合成用于改进表达的基因时，理想的是所设计基因能使其密码子使用反映出宿主细胞的优选密码子偏倚。对得自可获得序列信息的宿主细胞的基因进行测定可确定其密码子偏倚。

术语“表达”是指由基因产物(通常为蛋白)序列编码基因转录和翻译为基因产物。

术语“蛋白”、“多肽”和“肽”可交互使用，是指表达的基因产物。“成熟”蛋白是指翻译加工后的多肽(即在一级翻译产物中存在的任何前肽或原肽已经去除)。“前体”蛋白是指 mRNA 的一级翻译产物(即仍存在前肽或原肽)。前肽或原肽可以是但不限于胞内定位信号。

术语“同一性百分率”是指根据序列对比测定的两种或多种多肽序列或两种或多种多核苷酸序列之间的关系。“同一性”还表示看情况根据序列排列之间的匹配测定的多肽或多核苷酸序列之间序列相关度。

“相似性”是一种描述性术语，仅表示两个序列根据某一标准互相之间相似，不具有其起源或祖先的含义。“同源性”特别是指由于共同祖先的遗传而具有的相似性。根据一组序列之间的相似性关系有可能推断同源性，但在明确的实验室模型系统之外，共同祖先的遗传仍是假说(States 等, *Similarity and Homogeneity* (1992) 89-92 页, 载于: *J. Sequence Analysis Primer*, Gribskov, M.和 Devereux, J., (主编), Freeman and Co.)。

“同一性”和“相似性”可容易地通过已知方法计算，所述方法包括但不限于描述于以下文献的方法：Computational Molecular Biology (Lesk, A.M., 主编) Oxford University Press, New York (1988); Biocomputing: Informatics and Genome Projects (Smith, D.W., 主编) Academic Press, New York (1993); Computer Analysis of Sequence Data, Part I (Griffin, A.M.和 Griffin, H.G. 主编) Humana Press, Totowa, New Jersey (1994); Sequence Analysis in Molecular Biology (von Heinje, G. 主编) Academic Press (1987); 以及 Sequence Analysis Primer (Gribskov, M.和 Devereux, J. 主编) Stockton Press, New York (1991)。测定同一性的优选方法用来获得测试序列的最佳匹配。

测定两个序列之间同一性和相似性的优选计算机程序方法包括但不限于 GCG 程序包中的 Wisconsin Package Version 9.0 和 10.0

Genetics Computer Group (GCG) Gap 程序、BLASTP、BLASTN 和 FASTA (Pearson 等, *Proc. Natl. Acad. Sci. USA.* (1988) 85: 2444-2448)。公众可由 NCBI 和其它来源获得 BLAST X 程序(BLAST 手册, Altschul 等, *Natl. Cent. Biotechnol. Inf.*, Natl. Library Med. (NCBI NLM) NIH, Bethesda, MD 20894; Altschul 等, *J. Mol. Biol.* (1990) 215: 403-410)。

术语“大致相似”是指其中一个或多个核苷酸碱基的改变导致一个或多个氨基酸的取代但不影响 DNA 序列编码蛋白功能特性的核酸片段。“大致相似”还指其中一个或多个核苷酸碱基的改变不影响核酸片段通过反义或共抑制技术介导基因表达改变的能力的核酸片段。“大致相似”还指不显著影响所获得转录物的功能特性的修饰的本发明核酸片段, 例如一个或多个核苷酸碱基的取代、缺失或插入。

“密码子简并”是指允许核苷酸序列改变但不影响编码多肽的氨基酸序列的遗传密码分散性。例如, 本领域众所周知三联体密码子 CTT、CTC、CTA 和 CTG 全都编码亮氨酸(Atlas, R., *Principles of Microbiology*)。本领域还广为知道, 在已知位点产生化学等同氨基酸但不影响编码蛋白功能特性的基因改变是常有的事。因此, 一种疏水性氨基酸丙氨酸的密码子可由编码另一个疏水性较差的残基(例如甘氨酸)或疏水性更强的残基(例如缬氨酸、亮氨酸或异亮氨酸)的密码子取代, 而不影响编码蛋白的功能特性。同样, 还可以预期用一个负电荷残基取代另一个(例如天冬氨酸取代谷氨酸)或用一个正电荷残基取代另一个(例如赖氨酸取代精氨酸)可产生功能上等同的产物。还可以预期使蛋白分子 N 端和 C 端部分改变的核苷酸变化不会改变蛋白活性。所提出的每一种修饰都完全属于本领域常规技术, 测定编码产物的生物活性是否保留同样如此。

而且, 技术人员会认识到, 本发明包含的大致相似的核苷酸序列还可由其在严格条件下与本文例举序列杂交的能力确定。

通常, 严格条件是指盐浓度低于约 1.5 M Na⁺ (通常约 0.01-1.0 M

Na⁺或其它盐), pH 7.0-8.3, 对于短探针(例如 10-50 个核苷酸)温度至少约 30℃, 对于长探针(例如 50 个以上核苷酸)温度至少约 60℃。严格条件还可以通过加入去稳定剂如甲酰胺获得。典型的低严格条件包括用 6 × SSC (1 M NaCl)、30-35%甲酰胺、1% SDS (十二烷基硫酸钠)的缓冲溶液于 37℃ 杂交, 并用 1 × 至 2 × SSC (20 × SSC=3.0 M NaCl/0.3 M 柠檬酸三钠)于 50-55℃ 清洗。典型的中等严格条件包括用 6 × SSC (1 M NaCl)、40-45%甲酰胺、1% SDS 于 37℃ 杂交, 并用 0.5 × 至 1 × SSC 于 55-60℃ 清洗。典型的高严格条件包括用 6 × SSC (1 M NaCl)、50%甲酰胺、1% SDS 于 37℃ 杂交, 并用 0.1 × SSC 于 60-65℃ 清洗。

“特异性”通常随杂交后清洗变化, 关键因素是离子强度和最终清洗溶液的温度。可计算探针-靶杂交物的解链温度 T_m , 以提供确定正确严格条件的起点。 T_m (在确定的离子强度和 pH 下)是在该温度下 50%互补靶序列与完全匹配探针杂交的温度。对于 DNA-DNA 杂交物, 可如下由 Meinkoth 和 Wahl 方程(*Anal. Biochem.* (1984) 138: 267-284)估算:

$T_m = 81.5^\circ\text{C} + 16.6 (\log M) + 0.41 (\% G + C) - 0.61 (\% \text{甲酰胺}) - 500/L$; 其中 M 是一价阳离子的摩尔浓度, % G + C 是 DNA 中鸟嘌呤和胞嘧啶核苷酸的百分率, %甲酰胺是杂交溶液中甲酰胺的百分比, 而 L 是杂交物的碱基对长度。

每发生 1%的错配, T_m 就降低约 1℃; 因此, 可调节 T_m 、杂交和/或清洗条件, 以与希望同一性的序列杂交。例如, 如果要寻找同一性大于 90%的序列, 那么可将 T_m 降低 10℃。一般来说, 选择在确定的离子强度和 pH 时比特定序列及其互补物的热解链点 T_m 低约 5℃的严格条件。但是, 可使用比热解链点 T_m 低 1、2、3 或 4℃的高严格条件杂交和/或清洗; 可使用比热解链点 T_m 低 6、7、8、9 或 10℃的中等严格条件杂交和/或清洗; 以及可使用比热解链点 T_m 低 11、12、13、14、15 或 20℃的低严格条件杂交和/或清洗。普通技术

人员应理解,使用方程、杂交和清洗组合物以及需要的 T_m 从本质上描述了杂交和/或清洗溶液严格性的变化。如果需要的错配程度导致 T_m 小于 45°C(水溶液)或 32°C(甲酰胺溶液),则优选增加 SSC 浓度,以便可使用较高的温度。

杂交需要两种含互补序列的核酸,尽管与杂交严格性有关,但碱基错配仍是可能的。核酸杂交的合适严格性取决于核酸长度以及互补度,其可变性本领域众所周知。两种核苷酸序列的相似性和同一性的程度越大,具有这些序列的核酸杂交物的 T_m 值就越高。核酸杂交的相对稳定性(对应于较高的 T_m)按以下顺序降低: RNA:RNA、DNA:RNA、DNA:DNA。对于长度超过 100 个核苷酸的杂交物,已经导出了计算 T_m 的方程(参见 Sambrook 等,同上,9.50-9.51)。对于与较短核酸(即寡核苷酸)杂交,错配的位置变得更重要,而寡核苷酸长度决定了其特异性(参见 Sambrook 等,同上,11.7-11.8)。在一个实施方案中,可杂交核酸的长度至少约 10 个核苷酸。优选可杂交核酸的最小长度至少约 15 个核苷酸;更优选至少约 20 个核苷酸;最优选长度为至少 30 个核苷酸。而且,技术人员会认识到,可根据需要按照例如探针长度和靶 DNA 的 G+C 组成这样的因素调节温度和清洗溶液盐浓度。

关于核酸杂交的广泛指导可见于 Tijssen, *Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes*, 第 I 篇,第 2 章“杂交原则以及核酸探针测定严格性的综述”(1993) Elsevier, New York, 以及载于: *Current Protocols in MOLECULAR Biology*, (1995)第 2 章, Ausubel 等(主编), Greene Publishing and Wiley-Interscience, New York.

一般方法

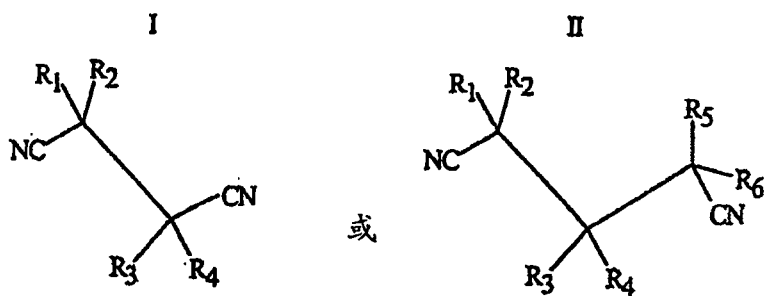
所有的 HPLC 方法都按照 Gavagan 等, *J. Org. Chem.* (1998) 63: 4792-4801 所述进行。

为组装和分析计算的核酸和蛋白序列(包括测定同一性和相似性), 申请人使用 GCG 程序包中的 Wisconsin Package Version 9.0 和 10.0 Genetics Computer Group (GCG) Gap 程序(使用 Needleman 和 Wunsch 算法, 采用标准默认值: 空隙起点罚分=50, 空隙延伸罚分=3 (Devereux 等, *Nucleic Acids Res.* (1984) 12: 387-395)), Sequencher Version and Vector NTI Deluxe v 4.0.3 软件以及数据库软件包。除非另有说明, 否则全部使用默认参数。

适用于维持和生长细菌培养物的材料和方法是本领域众所周知的。适用于以下实施例的技术可见于 Manual of Methods for General Bacteriology (1994) (Phillipp Gerhardt 等(主编), American Society for Microbiology, Washington, DC) 提出的方法或 Biotechnology: A Textbook of Industrial Microbiology (1989) 第二版, (Thomas D. Brock, Sinauer Associates, Inc., Sunderland, MA) 提出的方法。

适用于本发明的含腈底物是分子式为 NC-R-CN 的二腈, 其中 R 是具约 1 至约 10 个碳原子的亚烷基(alkylene)。更优选的含腈底物是分子式为 NCCXa(R)(CH₂)_nCN 的脂肪族 α , ω -二腈, 其中 a 为 0 或 1, 当 a 为 1 时 X 为氢, 而 R 为 H、烷基或取代烷基、或链烯基或取代链烯基、或次烷基或取代次烷基, n 为 1 或 2。用于本发明的最优选含腈底物为 2-甲基戊二腈。

在一个用含本文要求保护的生物单位的生物催化剂生产五元环或六元环内酰胺的改进方法中, 优选的含腈底物为以下任一式的脂肪族 α , ω -二腈:



其中 R_1 和 R_2 都为 H, 而 R_3 、 R_4 、 R_5 和 R_6 每个都独立选自 H、烷基或取代烷基、或链烯基或取代链烯基, 或者 R_3 和 R_4 合起来为次烷基或取代次烷基, 或者 R_5 和 R_6 合起来独立为次烷基或取代次烷基。更优选的含腓底物为在 α 碳原子上不对称取代的脂肪族 α , ω -二腓(参见美国专利第 5,858,736 号, 该专利通过引用结合到本文中)。

PCR 扩增、通过内切核酸酶和外切核酸酶进行 DNA 修饰以产生用于克隆 DNA 的需要末端、连接和细菌转化所需要的方法是本领域众所周知的。本文使用的标准重组 DNA 和分子克隆技术是本领域众所周知的, 描述于 Maniatis 和 T.J. Silhavy 等(载于: *Experiments with Gene Fusions*, (1984) Cold Spring Harbor Laboratory Press, Cold Spring, NY)以及 Ausubel 等(载于: *Current Protocols in Molecular Biology* (1994-1998) John Wiley & Sons, Inc., New York)。

A. 敏捷食酸菌 72W 腓水解酶的分离和部分氨基酸测序:

由敏捷食酸菌 72W (ATCC 55746)提取物分离本发明的腓水解酶并精制至大于 90%纯度。已知细菌腓水解酶一般含有一个亚单位(Novo 等, *FEBS Letters* (1995) 367: 275-279)。纯化腓水解酶的方法是本领域已知的(Bhalla 等, *Appl. Microbiol. Biotechnol.* (1992) 37: 184-190, Goldlust 等, *Biotechnol. Appl. Biochem.* (1989) 11: 581-601, Yamamoto 等, *Agric. Biol. Chem.* (1991) 55: 1459-1466)。通过穿过 Q-Sepharose 离子交换介质, 接着在 Hiload 16/60 Superdex 200 柱上凝胶过滤, 纯化所述腓水解酶。酶的纯化和分离方法是本领域已知的。(参见例如 Rudolph 等, *Chromatogr. Sci.* (1990) 51 (HPLC Biol. Macromol.): 333-50)。根据 5 mM 苺腓的 100 mM、pH 7.2 磷酸缓冲溶液于 245 nm 吸收增加的指示监测苺腓转变为苯甲酸的转化率, 由此监测纯化过程中的腓水解酶活性。

测定纯化的敏捷食酸菌 72W 腓水解酶蛋白的 N 末端氨基酸序列和肽消化产物序列。在胰蛋白酶消化后使用本领域已知的方法对凝

胶过滤纯化的脲水解酶进行测序, 所述方法包括通过还原进行半胱氨酸修饰, 接着用 4-乙烯嘧啶烷基化(参见例如 Matsudaira, *Methods Enzymol.* (1990) 182 (Guide Protein Purif.): 602-13 或 Allen, 蛋白和肽测序, 载于: *Laboratory Techniques in Biochemistry and Molecular Biology* (Burdon, R.H. 和 van Knippenberg, P.H., 主编), Elsevier, Amsterdam, New York, Oxford (1989))。

使用 Wisconsin Software Package 9.1 (Genetics Computer Group, Madison, WI) 在 GCG 上进行数据库检索。将测定的氨基酸序列与 SWISS-PROT 数据库包含的蛋白质序列信息对比。由纯化的敏捷食酸菌 72W 脲水解酶蛋白消化物分离的 10 个寡肽的氨基酸序列与其它已知脲水解酶蛋白的相似性超过 60%。这表明, 纯化蛋白含有脲水解酶的可能性很高。

B. 分离含脲水解酶基因的 DNA 片段并在大肠杆菌中表达:

为鉴定可能与已知脲水解酶基因同源的 DNA 序列, 设计和合成用作 PCR 引物的简并寡核苷酸引物。这些 PCR 引物的设计是基于在由 GenBank 数据库获得的细菌脲水解酶序列(GenBank 登录号 D12583、J03196、D13419、L32589、D67026)中发现的保守编码区。通过加入两轮苯酚-氯仿提取的改良型标准方法(Maniatis)分离敏捷食酸菌 72W (ATCC 55746)基因组 DNA。如此分离的 DNA 用作具有众多简并引物组合的 PCR 靶。将获得的扩增产物克隆入 pGem-T 载体, 并利用本领域常用的方法测序。由 10 个不同的 PCR 引物对实验可知, 仅有一个产物产生的 385 bp DNA 片段(SEQ ID NO: 3)与紫红红球菌 K22 的脲水解酶 DNA 序列(GenBank 登录号 D12583)的同一性为 74.1%, 该产物在质粒 pJJ28-5 中发现, 由引物 1F (SEQ ID NO: 1)和 7R (SEQ ID NO: 2)产生。另外, 由此 385 bp DNA 片段推定的蛋白产物含有许多与分离自纯化的 72W 脲水解酶蛋白消化物的序列匹配的肽序列。这是所鉴定的 385 bp 片段含有一部分目的 72 W 脲水解酶

基因的另一个证据。

由于发现此部分基因序列与其它已知脲水解酶基因具有高度同源性，所以申请人发现一段用作探针的 DNA，用于在任何 DNA 文库中鉴定包含 72W 脲水解酶基因的重组克隆。另外，成功使用 SEQ ID NO: 1 和 SEQ ID NO: 2 作为简并 PCR 引物寻找新脲水解酶基因表明，申请人发现的 DNA 序列具有用于鉴定和分离细菌菌株未知脲水解酶的通用用途。

为了将完整脲水解酶基因作图为单一限制片段，对敏捷食酸菌 72W 基因组 DNA 进行 DNA 杂交。使用 Qiagen genomic tip-100/G DNA 分离试剂盒(Qiagen Inc., USA)分离敏捷食酸菌 72W 细胞的高分子量基因组 DNA。用各种限制酶消化 1 μ g 基因组 DNA，在 1%琼脂糖凝胶上分离，然后转移至尼龙膜。用 385 bp 的部分脲水解酶片段(SEQ ID NO: 3)通过 DNA 杂交探测固定在尼龙膜上的限制性片段。当用以下任何一种限制酶消化基因组 DNA 时：*Bam*I、*Bcl*I、*Bgl*II、*Cla*I、*Eag*I、*Kpn*I、*Nar*I、*Nhe*I、*Not*I、*Nsi*I、*Pst*I、*Sal*I、*Spe*I、*Sst*I、*Xba*I、*Xho*I，将脲水解酶基因作图为单一限制片段。这些观察结果强有力地证实：在敏捷食酸菌 72W (ATCC 55746)中存在单一脲水解酶基因。其中含有 *Eco*RI 位点的已知部分基因序列(SEQ ID NO: 3)的 *Eco*RI 消化物产生两个条带。*Pst*I 消化获得与所述探针反应的单个 4.1 kb 条带，提示脲水解酶同源物存在于染色体的 4.1 kb *Pst*I 片段上。该信息对于选择限制酶(如果是 *Pst*I 的话)以产生基因组片段的丰富文库来分离完整脲水解酶基因很关键。

为构建文库，用 *Pst*I 消化敏捷食酸菌 72W 基因组 DNA，并通过凝胶电泳分离。由凝胶回收 2.5-7 kb 大小的片段，并连接入载体 pBluescript II SK(+)(Stratagene, La Jolla, CA, USA)。将连接的 DNA 电转化至大肠杆菌 DH10B electromax 细胞中(Life Technologies, Rockville, MD, USA)中。产生的文库代表 4×10^6 独立的敏捷食酸菌 72W *Pst*I 片段重组克隆。为筛选该文库，由从板上刮下的细胞制备

质粒。

将以上质粒文库转化入大肠杆菌 DH5 α 。使用含 385 bp pJJ28-5 脲水解酶基因片段的标记探针片段通过菌落杂交筛选所述文库。将显示与脲水解酶探针阳性杂交的三个菌落在存在氨苄青霉素的液体培养基中培养。这些菌落的质粒(pnit4、pnit5 和 pnit6)当用 *Pst*I、*Eco*RI 和 *Hind*III 消化时显示出相同的限制片段图谱，提示这些质粒包含相同的插入片段。*Pst*I 消化产生了 4.1 kb 片段，证实了 DNA 杂交的结果。在以上所有质粒中的插入片段都分别含有一个内部 *Eco*RI 位点和一个 *Hind*III 位点，同已鉴定的部分基因序列(SEQ ID NO: 3)的预测一致。

质粒 pnit4、pnit5 和 pnit6 中的插入片段具有相同的核苷酸序列(SEQ ID NO: 4)，该序列包含一个编码 369 个氨基酸的序列(SEQ ID NO: 5)的 ORF，在长度上与其它脲水解酶类似。一级氨基酸序列与紫红红球菌 K22 脲水解酶(GenBank 登录号 D12583)的同一性为 71%。ORF 起始密码子为 GTG，其产生以缬氨酸开始而不是以更常用的甲硫氨酸开始的蛋白。使用 Scanprosite (www.Espay.ch/tools/scnpsit1.html) 和 Profilescan (www.isrec.isb-sib.ch/software/PFSCAN-form.html) 在蛋白质家族和结构域的 PROSITE 数据库中扫描由以上 ORF (SEQ ID NO: 4) 编码的氨基酸序列(SEQ ID NO: 5)，发现在所有已知的脲水解酶中都包含有特征图谱。根据此信息，我们提出 164 位(SEQ ID NO: 5 和 14)的半胱氨酸可能是敏捷食酸菌 72W 脲水解酶的活性半胱氨酸位点。与已知脲水解酶的序列相似性以及存在脲水解酶特征这二者一起证实：脲水解酶基因位于 pnit4、pnit5 和 pnit6 的 4.1 kb *Pst*I 片段上的编码 369 个氨基酸序列(SEQ ID NO: 5)的 ORF(SEQ ID NO: 4)上。

本发明因此提供包含便于操作形式的完整敏捷食酸菌 72W 脲水解酶基因以及侧翼染色体 DNA 的质粒克隆 DH5 α :pnit4。

将在 pnit4 的 4.1 kb *Pst*I 片段中鉴定出的敏捷食酸菌 72W 脲水

解酶 ORF 克隆入许多基于用于过表达的 T7 启动子-T7 RNA 聚合酶系统的 pet 表达载体中(Studier 等, *Meth. Enzymol.* (1990) 185: 60-89)。更具体地说, 将所述 ORF 克隆入 pET-3c 和 pET-21a (二者均得自 Novagen, Madison, WI)中。pET-3c 腓水解酶构建物(pnitex2 (实施例 6)) 和 pET-21a 构建物(pSW91 (实施例 7))都用 ATG 取代了天然腓水解酶 ORF 的起始密码子(GTG)。另一个 pET-21a 构建物(pSW90 (实施例 7)) 产生一种在蛋白 N 端包含 11 个氨基酸的 T7 标记融合物的修饰形式。将质粒 pnitex2 转化入两种不同的宿主菌株 BL21 (DE3)和 BL21-SI 中, 分别产生菌株 SS1001 和 SS1002 (实施例 6)。这些菌株可以在有或没有 IPTG 诱导(菌株 SS1001, 实施例 6)或者 NaCl 诱导(菌株 SS1002, 实施例 6)的情况下表达腓水解酶。将质粒 pSW90 和 pSW91 转化入 BL21 (DE3), 分别产生菌株 SW90 和 SW91 (实施例 7), 它们都在有或没有 IPTG 诱导时表达酶活性腓水解酶。当按照生产商提供的方法诱导包含这些构建物的以上大肠杆菌菌株时, SS1001、SW90 和 SW91 生产预期分子量约 40 kd 的特定蛋白。另外, 在测试腓水解酶活性(由催 MGN 转变为 4-CPA 的能力表示)时, 所有的表达系统(SS1001、SS1002、SW90 和 SW91)都催化该反应, 表明基因工程过的大肠杆菌菌株能够生产活性敏捷食酸菌 72W 腓水解酶(实施例 6 和 7)。

C. 敏捷食酸菌 72W (ATCC 55746)和大肠杆菌表达菌株的生产 培养敏捷食酸菌 72W (ATCC 55746)

将一个冷冻种子批小瓶解冻, 将 1 mL 内含物接种在以下列出的 500 mL 无菌接种培养基上。于 30°C 在 2 L 摇瓶中以 250 rpm 振荡培养接种物 24-30 小时。

接种培养基

组分:	最终浓度:
磷酸二氢钾	1.5 g/L
磷酸氢二钾	3.4 g/L
硫酸铵	1.5 g/L
二水柠檬酸三钠	1 g/L
七水硫酸镁	0.4 g/L
痕量金属溶液(见下文)	1 mL/L
Amberex 695 (通用食品)	1 g/L
甘油(单独灭菌)	8 g/L

痕量金属溶液

组分:	母液浓度:
盐酸	10 mL/L
二水氯化钙	11.4 g/L
一水硫酸锰	1.23 g/L
五水硫酸铜	0.63 g/L
六水氯化钴	0.16 g/L
硼酸	0.91 g/L
七水硫酸锌	1.77 g/L
二水钼酸钠	0.05 g/L
二水硫酸氧钒	0.08 g/L
六水硝酸镍	0.04 g/L
亚硒酸钠	0.04 g/L
七水硫酸亚铁	6.0 g/L

无菌操作将摇瓶中的接种物转移至预先灭菌的含以下列出的发酵罐培养基的 Braun Biostat C 发酵罐中。

发酵罐培养基

组分:	终浓度:
磷酸二氢钾	0.39 mL/L
磷酸氢二钾	0.39 g/L
Difco 酵母提取物	5.0 g/L

培养在以下条件下进行: 32℃、pH 6.8-7.0、25%饱和溶解氧。在接种时, 发酵罐含有 8.5 L 发酵罐培养基和 218 g 营养供给溶液, 获得约 7 g/L 甘油的初始浓度。营养供给溶液包括以下分别灭菌并在冷却后组合的组分: 19.6 g 磷酸二氢钾的 0.25 L 去离子水溶液、3.3 g 七水硫酸镁和 4 mL 硫酸的 0.15 L 去离子水溶液、67 mL 痕量金属溶液以及 400 g 甘油的 0.80 L 去离子水溶液。在接种后第 18 小时, 开始流加营养供给溶液。开始, 营养供给溶液以 0.4 g 供给溶液/分钟(0.15 g 甘油/分钟)的速率加入。检测 550 nm 的培养物光密度(O.D. 550)为约 8-9。在第 26 小时, O.D. 550 为 16-18, 将流加速率增加至 0.9 g 供给溶液/分钟(0.3 g 甘油/分钟)。最后在第 34 小时将流加速率增加至 1.8 g 供给溶液/分钟(0.6 g 甘油/分钟)。该速率持续至发酵结束(约 42 小时)。最终的 O.D. 550 约为 65-75, 相当于 25-30 g dcw/L 的细胞密度。

离心回收细胞, 冷冻储存直至使用。为用作生物催化剂, 将在 0.35 M 磷酸缓冲溶液(pH 7.3)中的细胞加热至 50℃达 1 小时, 然后用于催化 MGN 为 4-CPA 的转变。

培养完整细胞活性的大肠杆菌细胞并固定化

将 25 mL 由单菌落培养的大肠杆菌菌株 SS1001 过夜培养物接种到 250 mL 新鲜 LB 培养基中。在有或没有 1 mM IPTG 诱导 1 小时的情况下将培养物培养至对数中期, 离心收获细胞。将菌株 SS1002 在 LBON (无 NaCl 的 LB)培养基中培养, 按照生产商对 BL21-SI 过表达

菌株的说明(Life Technologies, Rockville, MD, USA)于对数中期用 0.2 M NaCl 诱导 1 小时。将细菌细胞沉淀储存在冰上过夜, 然后检测腓水解酶活性。为固定化, 在 10 L 批发酵中用 LB 培养基将大肠杆菌菌株 SS1001 培养至 600 nm O.D.为 8.6。离心收获细胞并储存在湿冰上, 以便保存完整细胞活性和固定化。

对比如上所述的天然菌株产生的生物催化剂与实施例 6、7 和 15 的基因工程生物催化剂的重量-比活, 显然表达菌株 SS1001、SS1011 和 SW91 生产的生物催化剂的比活(实施例 6 的表 2、实施例 7 的表 3 和 4 以及实施例 15 的表 5)明显超过天然菌株(表 2、3、4 和 5)。

实施例 1

纯化腓水解酶蛋白

除非另有说明, 否则本方法中的所有步骤都在 5°C、pH 7.5 进行。

用 20 mM Tris pH 7.5、0.1 mM 苯甲基磺酰氟(PMSF)和 2.0 mM 二硫苏糖醇制备敏捷食酸菌 72W (ATCC 55746)湿细胞浆的 25% (重量)悬浮液。

按照本领域已知的方法通过弗氏细胞压碎器(American Instrucment Co., Silver Springs, MD, USA)制备所述悬浮液的提取物。以 27,500 g 离心 30 分钟去除细胞碎片后, 制备提取物的 20-55%硫酸铵流分, 然后在加入固体硫酸铵至 65%饱和度后通过过夜沉淀浓缩。使用最小体积的 20 mM Tris pH7.5 (缓冲液 A)溶解浓缩的蛋白沉淀, 并经含 Sephadex G-25 树脂(Pharmacia)的 PD10 柱脱盐。

脱盐后, 使用含 50 mL Q-Sepharose fast flow (Pharmacia)的柱通过阴离子交换层析分离浓缩的蛋白提取物。在浓缩的蛋白提取物上柱之后, 用 3 个柱体积的缓冲液 A 以 2 mL/分钟的流速流洗该柱, 以去除未吸附的蛋白。使用以缓冲液 A 制备的 0-0.5 M NaCl 梯度由该柱洗脱吸附的蛋白。于 280 nm 监测该柱的蛋白洗脱。在整个纯化过程中使用检测苜蓿腓水解产生苯甲酸的测定监测腓水解酶活性。以 0.4

M NaCl 洗脱活性脬水解酶。通过在还原条件(5% β -巯基乙醇)下于 10-15% SDS 聚丙烯酰胺凝胶上进行凝胶电泳(SDS-PAGE)分离 0.4 M NaCl 蛋白流分中的蛋白组分。0.4 M NaCl 蛋白流分的 50%以上由具有 39.7 kd 分子量亚单位的蛋白组成。这在脬水解酶的预期分子量范围之内(Cowan 等, *Extremophiles* (1998) 2: 207-216)。使用本领域已知的方法, 使用 Hiload 16/60 Superdex 200 柱(Pharmacia)以 20 mM 磷酸缓冲液 pH 7 凝胶过滤层析后测定脬水解酶的天然分子量为 570 kd, 其中所述柱已用凝胶过滤分子量标准(Pharmacia #17-0442-01)校准。在凝胶过滤后, 脬水解酶蛋白的纯度大于 90%。使用 2-甲基戊二脬作为底物于 25°C 测定纯化酶的比活为 35 IU/mg 蛋白。

实施例 2

分离显示与已知脬水解酶高度同源的 DNA 片段

通过加入两轮苯酚-氯仿提取的改良型标准方法(Maniatis)分离敏捷食酸菌 72W (ATCC 55746)基因组 DNA。如此分离的 DNA 用作使用引物 1F (SEQ ID NO: 1)和 7R (SEQ ID NO: 2)的 PCR 扩增靶。将获得的扩增产物克隆入 pGem-T 载体(Promega, Madison, WI), 并利用本领域常用的方法测序。克隆 pJJ28-5 产生的 385 bp DNA 片段(SEQ ID NO: 3)与紫红红球菌 K22 脬水解酶 DNA 序列(GenBank 登录号 D12583)的同一性为 74.1%。

实施例 3

定位包含与 385 bp 部分基因片段高度同源的序列的染色体片段

用生产商提供的在缓冲液中的限制酶 *Pst*I (Life Technologies (Gibco-BRL), Gaithersburg, MD, USA)消化 1-3 μ g 敏捷食酸菌 72W 的高分子量基因组 DNA 样品。在琼脂糖凝胶上分离消化的样品, 并在尼龙膜上进行 DNA 印迹。在插入片段侧翼的限制位点切割 pJJ28-5 制备含 385 bp 脬水解酶基因片段(SEQ ID NO: 3)的探针。于 60°C 进

行杂交, 然后进行高严格清洗(0.1 × SSC、0.1% SDS、60°C、15 分钟)。使用 ECL 随机引物标记和检测系统 II 型(Amersham International plc, Buckinghamshire, England)进行探针标记、杂交和点印迹检测以及随后的 DNA 印迹实验。

在与所述探针杂交时 *Pst*I 消化物产生单一的 4.1 kb 条带, 表明腓水解酶基因存在于敏捷食酸菌 72W 基因组的 4.1 kb *Pst*I 片段上。

实施例 4

构建敏捷食酸菌 72W (ATCC 55746)基因组文库

用 *Pst*I 消化 5 μg 敏捷食酸菌 72W 的高分子量基因组 DNA, 并通过制备型 1%琼脂糖凝胶电泳分离。由凝胶回收 2.5-7 kb 大小的片段, 并连接入 *Pst*I 消化和去磷酸化的 pBluescript II SK(+)载体(Stratagene, La Jolla, CA, USA)。将连接的 DNA 电转化至大肠杆菌 DH10B electromax 细胞(Life Technologies (Gibco-BRL), Gaithersburg, MD, USA)中, 将转化细胞接种在 LB+氨苄青霉素(100 μg/mL)+IPTG+X-Gal 平板上。获得的具有 95%重组质粒的文库代表 4×10^6 独立的重组克隆。为筛选和储存该文库, 由从所述板上刮下的细胞制备质粒。

将以上质粒文库转化入大肠杆菌 DH5α (Life Technologies, Rockville, MD, USA)细胞中, 并将转化体接种在 LB+氨苄青霉素(100 μg/mL)平板上。于 37°C 过夜温育后, 将完好分离的菌落转移至尼龙膜并原位裂解, 以将 DNA 固定在膜上。用实施例 3 描述的标记腓水解酶基因探针探测所述膜。将显示与腓水解酶探针阳性杂交的三个菌落在含有氨苄青霉素的液体培养基中培养。在分别用限制酶 *Pst*I、*Eco*RI 和 *Hind*III 消化时质粒 pnit4、pnit5 和 pnit6 显示出相同的图谱。这些结果证实 pnit4、pnit5 和 pnit6 中存在相同的插入片段。正如先前的 DNA 杂交所预测的一样, 限制性消化产生 4.1 kb *Pst*I 插入片段。

实施例 5

敏捷食酸菌 72W (ATCC 55746) 脲水解酶 ORF 序列

使用标准 Sanger 双脱氧链终止法对质粒 pnit4 的 4.1 kb 插入片段测序。使用 GCG MAP 和 Vector NTI 软件进行序列分析揭示, 存在一个 1110 bp 的可读框(SEQ ID NO: 4、SEQ ID NO: 15, 碱基 332-1441)。该 ORF 存在于 pnit4 插入片段的 *BclI*-*BglII* 片段(SEQ ID NO: 15 和图 1)中。检索 GenBank 序列数据库显示, 该可读框(SEQ ID NO: 4)与紫红红球菌 K22 脲水解酶 ORF(GenBank 登录号 D12583)的同一性为 68%。由所述可读框(SEQ ID NO: 4)推定的 369 个氨基酸的序列(SEQ ID NO: 5)与紫红红球菌 K22 脲水解酶蛋白的同一性为 71%。使用 Profilescan 和 Scanprosite 工具在蛋白质家族和结构域的 PROSITE 数据库中扫描由敏捷食酸菌 72W ORF (SEQ ID NO: 4)编码的氨基酸序列(SEQ ID NO: 5), 结果显示存在所有已知脲水解酶的保守特征图谱。

实施例 6

用大肠杆菌表达敏捷食酸菌 72W (ATCC 55746) 脲水解酶

使用寡核苷酸 SEQ ID NO: 6 和 SEQ ID NO: 7 以 PCR 反应扩增敏捷食酸菌 72W (ATCC 55746)基因组 DNA 的脲水解酶 ORF。用 *NdeI* 和 *BamHI* 消化 PCR 产物, 并克隆入 *NdeI*-*BamHI* 线性化的 pET-3c 中。所获得的表达质粒(pnitex2)中的脲水解酶 ORF 的起始密码子为 ATG (SEQ ID NO: 13)而不是天然的 GTG 密码子(SEQ ID NO: 4)。因此, pnitex2 表达的脲水解酶的氨基酸序列在位置 1 (SEQ ID NO: 14)为 M (甲硫氨酸), 而不是天然敏捷食酸菌 72W 脲水解酶(SEQ ID NO: 5)的 V (缬氨酸)。将质粒 pnitex2 和 pET-3c (对照)转化入大肠杆菌菌株 BL21(DE3) (Novagen, Madison, WI, USA)中, 分别产生 SS1001 和 BL21(DE3):pET-3c。

将 pnitex2 和 pET-3c 转化入 BL21-SI (Life Technologies, Rockville,

MD)所获得的菌株分别命名为 SS1002 和 BL21-SI:pET-3c。质粒 pnitex2 包含处于 T7 启动子控制之下的脲水解酶编码序列。通过对宿主菌株 BL21(DE3)加入 IPTG 和对宿主菌株 BL21-SI 加入 NaCl 诱导染色体定位的 T7 RNA 聚合酶转录, T7 RNA 聚合酶调节 T7 启动子控制的基因转录(在此情况下为嵌合脲水解酶基因)。在摇瓶或 10 L 发酵罐中培养以上的大肠杆菌转化体 SS1001 (以及对照菌株 BL21(DE3):pET-3c)和 SS1002 (以及对照菌株 BL21-SI: pET-3c), 并在有或没有诱导的情况下检测嵌合脲水解酶基因表达, 其中诱导是按照生产商提供的方法通过加入 IPTG 或 NaCl 完成。制备菌株 SS1001、SS1002 和 BL21(DE3):pET-3c 的粗提物, 并以每个泳道 5-20 μg 蛋白在 SDS-PAGE 凝胶上电泳。对 SDS-PAGE 凝胶进行分子量和激光光密度计分析, 测定出菌株 SS1001 和 SS1002 表达的特异性蛋白条带约为 40 kd 大小, 与预期的敏捷食酸菌 72W 脲水解酶蛋白大小相当。该条带占 SS1001 中总可溶性蛋白的 50%以上。BL21(DE3):pET-3c 粗提物中没有相应的条带。酶测定 SS1001 提取物显示, 总可溶性蛋白的 12%为酶活性脲水解酶。

还分别测试了大肠杆菌转化体 SS1001 和 SS1002 以及对照菌株 BL21(DE3):pET-3c 和 BL21-SI:pET-3c 完整细胞的脲水解酶活性。将以上的转化体培养至对数末期, 将有或没有诱导时完整细胞的脲水解酶活性水平与仅具有对照 pET-3c 载体的菌株活性进行对比(表 2)。通过 MGN 的 4-CPA 生产率检测脲水解酶活性。用 0.10 M 磷酸钾缓冲液 pH 7.0 制备 50 mg (干细胞重量)/mL 细胞悬浮液。于 25 $^{\circ}\text{C}$ 向装配有磁力搅拌棒的 20 mL 玻璃闪烁小瓶中加入 3.0 mL 0.40 M MGN 水溶液。于 25 $^{\circ}\text{C}$ 边搅拌边加入 1.0 mL 所述细胞悬浮液。在加入细胞悬浮液后的第 5、10 和 15 分钟, 由反应混合物中取出 180 μL 的等份溶液, 与 5 μL 6.0 N HCl 和 20 μL 0.75 M N-甲基丙酰胺的水溶液 (HPLC 外标)混合, 离心, 通过 HPLC 分析上清液的 4-CPA 铵盐生产率。

1 单位脲水解酶活性(IU)等于每分钟生产 1 μmol 4-CPA 铵盐。活性水平以单位/g 干细胞重量表示, 并与敏捷食酸菌 72W 的活性对比。

表 2: 大肠杆菌转化体的脲水解酶活性

转化体催化剂	诱导	脲水解酶活性 (IU/g 干细胞重量)
大肠杆菌 SS1001 (BL21(DE3):pnitex2)	无	360
大肠杆菌 SS1001 (BL21(DE3):pnitex2)	IPTG	466
大肠杆菌对照 (BL21(DE3):pET-3c)	IPTG	0
大肠杆菌 SS1002 (BL21-SI:pnitex2)	NaCl	288
大肠杆菌对照 (BL21-SI:pET-3c)	无	0
敏捷食酸菌 72W (ATCC 55746)	不适用	271

实施例 7

用大肠杆菌表达敏捷食酸菌 72W 脲水解酶基因

使用对应于测定的编码序列 5'和 3'末端 DNA 序列的引物通过 PCR 分离敏捷食酸菌 72W 的 72W 脲水解酶编码序列。使用标识为 SEQ ID NO: 8 和 SEQ ID NO: 9 的引物扩增编码序列, 并亚克隆入 pGEM-T (Promega, Madison, WI)。鉴定为 SEQ ID NO: 8 的引物将天然 GTG 起始密码子更改为 ATG。然后用 *Bam*HI 和 *Sac*I 取出 pGEM-T 的脲水解酶基因片段, 将其亚克隆入大肠杆菌表达载体 pET-21a (Novagen, Madison, WI) 的 *Bam*HI 和 *Sac*I 之间产生质粒 pSW90, 使得在脲水解酶编码序列的 N 端编码 11 个氨基酸的 T7 标记。还使用鉴定为 SEQ ID NO: 10 和 SEQ ID NO: 9 的引物扩增所述编码序列, 并亚克隆入 pGEM-T。鉴定为 SEQ ID NO: 10 的引物将天然 GTG 起始密码子更改为 ATG。然后用 *Ned*I 由 pGEM-T 中取出所述基因片段, 将其亚克隆入 pET-21a 的 *Ned*I 位点产生质粒 pSW91, 使得对天然脲水解酶编

码序列没有修饰(除了所述起始密码子之外)。将质粒 pSW90 和 pSW91 转化入大肠杆菌菌株 BL21(DE3)(Novagen, Madison, WI), 分别产生菌株 SW90 和 SW91。在对 SW90 和 SW91 进行标准培养和诱导(Novagen 建议)之后, 制备细胞提取物并通过 SDS-PAGE 检测, 观察到对应于 72W 腈水解酶蛋白的预期大小约 40 kDa 的主蛋白带。在 SW91 的情况下, PAGE 凝胶上的腈水解酶蛋白占总可溶性蛋白的 50% 以上。对照没有观察到相应的主带。

按实施例 6 所述通过检测 MGN 的 4-CPA 生产率, 测试表达敏捷食酸菌 72W (ATCC 55746)腈水解酶编码序列的大肠杆菌转化体 SW91 的腈水解酶活性。

在按生产商(Novagen, Madison, WI)建议培养和 IPTG 诱导之后, 检测 SW91 的腈水解酶活性水平, 并与具有无腈水解酶基因表达载体的对照转化体活性以及通常观察到的敏捷食酸菌 72W 活性进行对比。结果示于表 3。1 单位腈水解酶活性(IU)等于每分钟生产 1 μmol 4-CPA 铵盐, 并以单位/g 干细胞重量表示。

表 3: 标准培养和诱导之后的 SW91 腈水解酶活性

菌株	腈水解酶活性 (IU/g 干细胞重量)
大肠杆菌 SW91 (BL21(DE3):pSW91)	551
大肠杆菌对照(BL21(DE3):pET-21a)	0
敏捷食酸菌 72W (ATCC 55746)	271

还在无 IPTG 诱导的情况下用摇瓶在 LB 培养基(Maniatis)中于 37 $^{\circ}\text{C}$ 培养大肠杆菌转化体 SW91, 直至饱和(12-16 小时), 此后通过离心收获细胞, 并测定腈水解酶活性。结果示于表 4, 并与通常观察到的敏捷食酸菌 72W 活性进行对比。

表 4: 无诱导培养至饱和后的 SW91 脲水解酶活性

菌株	脲水解酶活性 (IU/g 干细胞重量)
大肠杆菌 SW91 (BL21(DE3):pSW91)	662
敏捷食酸菌 72W (ATCC 55746)	271

实施例 8

用角叉胶固定敏捷食酸菌 72W 或大肠杆菌转化体 SS1001 细胞

在快速搅拌下, 向 250 mL 培养瓶(装配有磁力搅拌棒并含有 50 °C 的 64.12 g 蒸馏去离子水)中缓慢加入 3.38 g FMC BioPolymer ISAGEL® RG300 角叉胶。在快速搅拌下将混合物加热至 75-80 °C, 直至角叉胶完全溶解, 用恒温水浴将获得的溶液冷却至 55-56 °C(成胶温度约 52 °C)。将敏捷食酸菌 72W 细胞或大肠杆菌转化体 SS1001 (12.5%干细胞重量)在 0.35 M 磷酸氢二钠缓冲液(45 mL 总体积, pH 7.3) 中的悬浮液加热至 50 °C 达 60 分钟(敏捷食酸菌 72W)或 12 分钟(大肠杆菌转化体 SS1001), 然后在搅拌下加入 55-56 °C 的角叉胶溶液。立即将细胞/角叉胶混合物缓慢加入到 50 °C 的 450 mL 豆油中, 在加入的同时使用搅拌器(overhead stirrer)搅拌。通过控制搅拌速率在油中产生需要大小的细胞/角叉胶液滴, 然后将油温降低至 35 °C 以使液滴成胶, 倾去获得的珠中的油, 所述珠用 0.10 M 碳酸氢钾缓冲液(pH 7.3)清洗。将 20 g 珠重悬浮在 48.8 mL 的 0.10 M 碳酸氢钾缓冲液(pH 7.3) 中, 加入 0.25 g 25%(重量)戊二醛水溶液, 于 25 °C 混合所述珠 1.0 小时。然后向混合物中加入 1.0 g 12.5%(重量)聚乙烯亚胺(BASF Lupasol® PR971L, 平均分子量约为 750,000)的水溶液, 再于 25 °C 混合所述珠 1 小时。然后用 50 mL 0.30 M 碳酸氢铵(pH 7.3)于 25 °C 清洗交联珠, 于 5 °C 储存在相同的缓冲液中。

实施例 9

对比角叉胶固定的敏捷食酸菌 72W 和大肠杆菌转化体 SS1001 细胞 作为生产 4-氰戊酸铵盐的催化剂的能力

在典型反应中，将 16.5 g 固定化细胞/角叉胶珠放置于 125 mL 夹套反应容器中，该容器用循环温浴控温在 30℃。向反应容器中加入 69.25 mL 水和 14.25 mL (13.54 g, 1.25 M) 2-甲基戊二腈，于 30℃ 搅拌混合物。将样品(0.100 mL)与 0.400 mL 水混合，然后将 0.360 mL 稀释样品与 0.040 mL 0.75 M N-甲基丙酰胺水溶液以及 0.020 mL 6.0 N HCl 混合。过滤(0.22 μm)所获得的混合物，通过 HPLC 分析滤过液中的 2-甲基戊二腈、4-氰戊酸和 2-甲基戊二酸水平。使用角叉胶固定的敏捷食酸菌 72W 和大肠杆菌转化体 SS1001 细胞时的 4-氰戊酸产率分别为 184 mM/小时和 310 mM/小时。在 2-甲基戊二腈完全转变时，每种催化剂都分别以 98.7%和 1.3%得率生产 4-氰戊酸铵盐和 2-甲基戊二酸二铵盐。

实施例 10

用海藻酸钙固定大肠杆菌转化体 SW91 细胞

在快速搅拌下，向 100 mL 培养瓶(装配有磁力搅拌棒并含有 50℃的 22.9 g 蒸馏去离子水)中缓慢加入 1.10 g FMC BioPolymer Protanal® LF 10/60 海藻盐。在快速搅拌下将混合物加热至 75-80℃，直至海藻酸盐完全溶解，用水浴将获得的溶液冷却至 25℃。将大肠杆菌转化体 SW91 (50%湿细胞重量，11.5%干细胞重量)在 0.15 M 乙酸钠缓冲液(16 mL 总体积，pH 7.0)中的悬浮液在搅拌下加入到 25℃ 的海藻酸盐溶液中。在搅拌下用注射器将细胞/海藻酸盐混合物滴加到 25℃的 213 mL 0.20 M 乙酸钙缓冲液(pH 7.0)中。搅拌 2 小时后，倾去所获得的珠中的缓冲液，于 25℃将所述珠重悬浮在 84 mL 的 0.20 M 乙酸钙缓冲液(pH 7.0)中。边搅拌边加入 0.88 g 25%(重量)戊二醛水溶液，于 25℃混合所述珠 1.0 小时。然后向混合物中加入 3.5 g

12.5%(重量)聚乙烯亚胺(BASF Lupasol® PR971L, 平均分子量约为750,000)水溶液, 再于25℃混合所述珠1小时。然后用84 mL 0.20 M 乙酸钙缓冲液(pH 7.0)于25℃清洗交联珠两次, 储存于5℃相同的缓冲液中。

实施例 11

海藻酸钙固定的大肠杆菌转化体 SW91 细胞作为生产 4-氰戊酸铵盐的催化剂

将16.5 g如实施例10所述制备的大肠杆菌 SW91/海藻酸盐珠置于125 mL 夹套反应容器中(用循环温浴控温在30℃)。向反应容器中加入68.25 mL 水、1.0 mL 0.20 M 乙酸钙缓冲液(pH 7.0)和14.25 mL (13.54 g, 1.25 M) 2-甲基戊二腈, 于30℃搅拌混合物。将样品(0.100 mL)与0.400 mL 水混合, 然后将0.360 mL 稀释样品与0.040 mL 0.75 M N-甲基丙酰胺水溶液以及0.020 mL 6.0 N HCl 混合。过滤(0.22 μm)所获得的混合物, 通过 HPLC 分析滤过液中的 2-甲基戊二腈、4-氰戊酸和 2-甲基戊二酸水平。4-氰戊酸产率为739 mM/小时, 反应在2.5小时以内结束。在2-甲基戊二腈完全转变时, 4-氰戊酸铵盐和 2-甲基戊二酸二铵盐的产率分别为98.5%和1.5%。

在反应结束时, 倾出催化剂珠上的产物混合物, 催化剂珠如上所述在另5个连续批反应中重复使用。在第6次反应中的4-氰戊酸产率为721 mM/小时, 反应在2.5小时以内结束。在2-甲基戊二腈完全转变时, 4-氰戊酸铵盐和 2-甲基戊二酸二铵盐的产率分别为98.0%和2.0%。

实施例 12

固定化敏捷食酸菌 72W 细胞腈水解酶的总转化数(TTN)

将16.5 g 固定化敏捷食酸菌细胞催化剂(如实施例8所述制备)置于125 mL 带夹套的反应容器中(控温在25℃)。向反应容器中加入72.1

mL 水和 11.4 mL (10.83 g, 1.0 M) MGN, 于 25°C 搅拌混合物。将样品 (0.100 mL) 与 0.400 mL 水混合, 然后将 0.360 mL 稀释样品与 0.040 mL 0.75 M N-甲基丙酰胺水溶液以及 0.020 mL 6.0 N HCl 混合。过滤(0.22 μm) 所获得的混合物, 通过 HPLC 分析滤过液。在 MGN 完全转变后, 将产物混合物和催化剂倾出至 250 mL 配衡烧杯中, 倾出催化剂中的水性产物混合物。剩余的催化剂和产物混合物残留物称重, 记录重量, 然后将水加入到催化剂中至最终总重(水和催化剂)为 88.6 g。将催化剂悬浮液转移回反应容器, 加入 11.4 ml MGN, 重复进行反应。

用催化剂再循环总共进行了 67 次连续批反应。在 67 次重复使用的反应中催化剂重量没有可检测的损失, 生产出 1001 g 4-CPA/g dcw 敏捷食酸菌 72W 细胞。初反应速率为 142 mm 4-CPA/小时, 相当于 237 IU 脲水解酶活性或 6.76 mg (1.7×10^{-7} 摩尔) 72W 脲水酶。67 次连续批反应产生的 4-CPA 量为 6.6 mol, 因此总转化数(TTN=mol 产物/mol 酶)为 $6.6/1.7 \times 10^{-7}$ 或 3.9×10^7 TTN。

实施例 13

用巴斯德毕赤酵母表达敏捷食酸菌 72W 脲水解酶编码序列

使用 16 个 90 bp 的寡聚物(SEQ ID NO: 17-32)和 PCR 介导的重叠延伸技术构建合成脲水解酶基因(SEQ ID NO: 16)(其编码蛋白序列与敏捷食酸菌 72W 中发现的蛋白序列(SEQ ID NO: 14)相同)。合成基因优化了巴斯德毕赤酵母所具有的密码子使用, 其与敏捷食酸菌 72W 脲水解酶基因序列具有的密码子使用明显不同。通过核苷酸测序证实以后, 将合成脲水解酶基因亚克隆入 pGAPZA 和 pPICZA (Invitrogen, San Diego, CA) 中的 *EcoRI* 位点。pGAPZA 使用组成型巴斯德毕赤酵母 GAP (甘油醛-3-磷酸脱氢酶) 启动子驱动外源基因表达; pPICZA 使用甲醇诱导型巴斯德毕赤酵母 AOX1 (乙醇氧化酶 I) 启动子驱动外源基因表达。基本如所述(Cregg 等, *Mol. Cell Biol.* (1985) 5: 3376-3385) 通过原生质球转化用质粒 pGAP::nit 和 pAOX::nit 将巴

斯德毕赤酵母 GS115 (Invitrogen)转化为 zeocin 抗性。合适培养和诱导(仅 pAOX::nit 需要诱导)后,通过对基本如所述(Sreekrishna 等, *Biochem* (1989) 28: 4117-4125)制备的细胞提取物进行 SDS-PAGE 蛋白分析以及通过如前所述的酶活性测定(实施例 6 和 7)鉴定生产腈水解酶的转化体。

实施例 14

测定敏捷食酸菌 72W 腈水解酶占总可溶性蛋白的百分比

如前所述(实施例 1)通过将 25%(重量)细胞悬浮液穿过弗氏细胞压碎器制备敏捷食酸菌 72W 粗提物。通过在还原条件(5% β -巯基乙醇)下于 10-15% SDS 聚丙烯酰胺凝胶上进行凝胶电泳(SDS-PAGE)分离等份粗提物可溶性蛋白流分中的蛋白组分。在电泳后,于 25°C 用由 0.2%考马斯亮蓝染料、30%甲醇和 10%乙酸组成的溶液处理凝胶 15 分钟。在用由 30%甲醇和 10%乙酸组成的溶液脱色之后目测凝胶上的蛋白条带。然后使用 LKB Ultrosan XL 增强型激光光密度计对凝胶中的蛋白条带积分。具有分子量约 40 kd 的亚单位的腈水解酶条带占凝胶上检测到的总可溶性蛋白的 3.4%。

实施例 15

SS1011 [MG1655(DE3):pnitex2]构建物及其腈水解酶活性

将 λ DE3 前噬菌体位点特异性整合入大肠杆菌菌株 MG1655 (ATCC 47076)染色体,获得菌株 MG1655(DE3)。为此按照生产商说明使用 λ DE3 溶原化试剂盒(目录号 69734-3, Novagen, Inc., Madison, WI)。用实施例 6 描述的质粒 pnitex2 转化 MG1655(DE3),获得 SS1011。在 LB 培养基中培养菌株 SS1011 和 SS1001 16-17 小时。使用基于 μ l 平板的分光光度测定检测完整细胞腈水解酶活性,结果示于表 5,其中所述分光光度测定于 35°C 进行,根据 245 nm 吸收增加指示检测苜蓿腈(9.5 mM, 0.1 M 磷酸缓冲液, pH 7.0)水解产生苯甲酸。

因为将较高测定温度(35℃对 25℃)和酶对苜蓿的相对较高的底物特异性组合在一起, 所以使用该测定检测的脲水解酶活性单位(IU)通常比使用先前所述的甲基戊二脲测定检测的活性单位高 4-5 倍。

表 5: 通过苜蓿水解检测测定的大肠杆菌转化体 SS1001 和 SS1011
脲水解酶活性

转化体催化剂	脲水解酶活性 (IU/g 干细胞重量)
大肠杆菌 SS1011 (MG1655(DE3):pnitex2)	3652
大肠杆菌 SS1001 (BL21(DE3):pnitex2)	4500
敏捷食酸菌 72W (ATCC 55746)	1491

序列表

<110> E. I. du Pont de Nemours & Company

<120> 分离和表达敏捷食酸菌 72W 的腓水解酶基因

<130> BC-1032 PCT

<140>

<141>

<150> 60/193,707

<151> 2000年3月31日

<160> 32

<170> Microsoft Office 97

<210> 1

<211> 17

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: 正向引物(1F)

<220>

<223> K= G 或 T, M= A 或 C, S= G 或 C, Y= C 或 T

<400> 1

tkkmtkccsg gctaycc

17

<210> 2

<211> 17

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: 反向引物(7R)

<220>

<223> S= G 或 C, H= A 或 C 或 T, M= A 或 C, R= A 或 G,
Y= C 或 T

<400> 2
ggccasshtg mrayrtg 17

<210> 3
<211> 385
<212> DNA
<213> 敏捷食酸菌(Acidovorax facilis)

<400> 3
ctattgggcg tggctcggcg acgtgaagta cagcctaagc ttactttcac gctatcacga 60
gaattcgttg gagctaggtg acgaccgtat gcgtcgcctc cagctggcgg cgcgcgcgcaa 120
caaaatcgca ctcgtcatgg gctattcggga gcggggaagcc ggatcgcgct atctgagcca 180
ggtgttcacg gacgagcgtg gogagatogt tgccaatcgg cgcaagctga agccccacaca 240
cgttgagcgt acgatctacg gogaaggcaa cggaaaccgat ttctcaccgc acgacttcgc 300
gttcggacgc gtcgggtggat tgaactgctg ggaacatttc caaccgctca gcaagttcat 360
gatgtacagc ctcggtgagc aggtc 385

<210> 4
<211> 1110
<212> DNA
<213> 敏捷食酸菌(Acidovorax facilis)

<400> 4
gtggtttcgt ataacagcaa gttcctcgcg gcaaccgttc aggcagagcc ggtatggctc 60
gacgcagacg caacgatcga caagtcgacg ggcacatcag aagaagotgc caaaagggc 120
gogagtctga tcgctttccc ggaagtattc attccgggct acccctattg ggctggctc 180
ggogacgtga agtacagcct aagctttact tcacgctatc acgagaatto gttggagcta 240
ggtgacgacc gtatgcgtcg cctccagctg gcgcgcgcgc gcaacaaaat cgcactcgtc 300
atgggctatt cggagcggga agccggatcg cgtatctga gccaggtgtt catcgacgag 360
cgtggcgaga tcggttccaa toggogcaag ctgaagccca cacacgttga gcgtacgac 420
tacggcgaag gcaacggaac cgatttcctc acgcacgact tcgcgttcgg acgctcggg 480
ggattgaaact gctgggaaca ttccaaccg ctcagcaagt tcatgatgta cagcctcgg 540
gagcaggtcc acgttgcatc gtggccggcg atgtcccctc ttcagccgga tgttttccaa 600
ctgagcatcg aagccaacgc gaoggtcacc cgcctcgtac caatcgaagg ccaaaccttt 660
gtgctttgct cgacgcaggt gatcggacct agcgcgatcg aaacgttctg cctcaacgac 720
gaacagcggc cactgttgcc gcaaggatgt ggctggggcg gcatttacgg cccgatgga 780
agcagccttg cgaagcctct ggoggaagat gctgagggga tctgttacgc agagatcga 840
ctggagcaga ttctgctggc gaaggctgga gccgatccgg tcgggcaacta ttgcggcct 900
gacgtgctgt cggctcagtt cgaccgcgcg aatcatacgc cagttcatcg catcggcatt 960
gacggtcgtc tggatgtgaa taccgcagc cgcgtggaga attccgact gcgacaagcg 1020
gctgagcagg agcgtcagcg atccaagcgg ctcggaacga aactctttga acaatccctt 1080
ctggctgaag aacoggtccc agcaaagtag 1110

<210> 5
<211> 369
<212> PRT

<213> 敏捷食酸菌(Acidovorax facilis)

<400> 5

Val Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
 1 5 10 15

Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
 20 25 30

Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
 35 40 45

Val Phe Ile Pro Gly Tyr Pro Tyr Trp Ala Trp Leu Gly Asp Val Lys
 50 55 60

Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
 65 70 75 80

Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
 85 90 95

Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
 100 105 110

Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
 115 120 125

Arg Lys Leu Lys Pro Thr His Val Glu Arg Thr Ile Tyr Gly Glu Gly
 130 135 140

Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
 145 150 155 160

Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met
 165 170 175

Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
 180 185 190

Pro Leu Gln Pro Asp Val Phe Gln Leu Ser Ile Glu Ala Asn Ala Thr
 195 200 205

Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
 210 215 220

Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
 225 230 235 240

Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr

	245	250	255
Gly Pro Asp	Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu		
	260	265	270
Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys			
	275	280	285
Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser			
	290	295	300
Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile			
	305	310	315
Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg			
	325	330	335
Leu Arg Gln Ala Ala Glu Gln Glu Arg Gln Ala Ser Lys Arg Leu Gly			
	340	345	350
Thr Lys Leu Phe Glu Gln Ser Leu Leu Ala Glu Glu Pro Val Pro Ala			
	355	360	365

Lys

<210> 6

<211> 27

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述:引物

<400> 6

gacgcatatg gtttcgtata acagcaa

27

<210> 7

<211> 28

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述:引物

<400> 7

ogacggatcc ttatggctac tttgctgg

28

<210> 8	
<211> 31	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列描述:引物	
<400> 8	
cggatccatg gtttcgtata acagcaagtt c	31
<210> 9	
<211> 23	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列描述:引物	
<400> 9	
ttatggctac tttgctggga ccg	23
<210> 10	
<211> 29	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列描述:引物	
<400> 10	
tacatatggt ttcgtataac agcaagttc	29
<210> 11	
<211> 26	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列描述:引物	
<400> 11	
catctcgaga tggtttcgta taacag	26

<210> 12
 <211> 23
 <212> DNA
 <213> 人工序列

<220>
 <223> 人工序列描述: 引物

<400> 12
 cactcgagct actttgctgg gac 23

<210> 13
 <211> 1110
 <212> DNA
 <213> 敏捷食酸菌(*Acidovorax facilis*)

<400> 13
 atggtttcgt ataacagcaa gttcctcgcg gcaaccgttc aggcagagcc ggtatggctc 60
 gacgcagacg caacgatcga caagtcgatc ggcatcatcg aagaagctgc ccaaaagggc 120
 gcgagtctga tcgctttccc ggaagtattc attccgggct acccctattg ggcgtggctc 180
 ggcgacgtga agtacagcct aagctttact tcacgctatc acgagaattc gttggagcta 240
 ggtgacgacc gtatgcgctg cctccagctg gccgcgcgcc gcaacaaaat cgcactcgtc 300
 atgggctatt cggagcggga agccggatcg cgctatctga gccagggtgtt catcgacgag 360
 cgtggcgaga tcgttgccaa tcggcgcaag ctgaagccca cacacgttga gcgtacgac 420
 tacggcgaag gcaacggaac cgatttctc acgcacgact tcgcgttcgg acgcgtcgg 480
 ggattgaact gctgggaaca ttccaaccg ctcagcaagt tcatgatgta cagcctcgg 540
 gagcaggctc acgttgcac gtggccggcg atgtccctc ttcagccgga tgttttccaa 600
 ctgagcatcg aagccaacgc gacggtcacc cgctcgtacg caatogaagg ccaaaccctt 660
 gtgctttgct cgacgcagg gatcggaacct agcgcgatcg aaacgttctg cctcaacgac 720
 gaacagcgcg cactgttgcc gcaaggatgt ggctgggcgc gcatttacgg cccgatgga 780
 agcgagcttg cgaagcctct ggcggaagat gctgagggga tctgtacgc agagatcgat 840
 ctggagcaga ttctgctgge gaagctgga gccgatccgg tcgggcacta ttgcggcct 900
 gacgtgctgt cggctcagtt cgaccgcgc aatcatacgc cagttcatcg catcggcatt 960
 gacggtcgct tggatgtgaa tacccgcagt cgcgtggaga attccgact gcgacaagcg 1020
 gctgagcagg agcgtcagge atccaagcgg ctcggaacga aactctttga acaatccct 1080
 ctggetgaag aaccggtctc agcaaagtag 1110

<210> 14
 <211> 369
 <212> PRT
 <213> 敏捷食酸菌(*Acidovorax facilis*)

<400> 14
 Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
 1 5 10 15

Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
 20 25 30

Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
 35 40 45

Val Phe Ile Pro Gly Tyr Pro Tyr Trp Ala Trp Leu Gly Asp Val Lys
 50 55 60

Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
 65 70 75 80

Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
 85 90 95

Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
 100 105 110

Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
 115 120 125

Arg Lys Leu Lys Pro Thr His Val Glu Arg Thr Ile Tyr Gly Glu Gly
 130 135 140

Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
 145 150 155 160

Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met
 165 170 175

Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
 180 185 190

Pro Leu Gln Pro Asp Val Phe Gln Leu Ser Ile Glu Ala Asn Ala Thr
 195 200 205

Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
 210 215 220

Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
 225 230 235 240

Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
 245 250 255

Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
 260 265 270

Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys

275 280 285
Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
290 295 300
Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
305 310 315 320
Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
325 330 335
Leu Arg Gln Ala Ala Glu Gln Glu Arg Gln Ala Ser Lys Arg Leu Gly
340 345 350
Thr Lys Leu Phe Glu Gln Ser Leu Leu Ala Glu Glu Pro Val Ser Ala
355 360 365

Lys

- <210> 15
- <211> 1776
- <212> DNA
- <213> 德氏食酸菌(Acidovorax delafieldii)

<400> 15
tgatcactcc tgaccacott gctgaaaaat tcagcagcgt agccgtcaac gtggetcaat 60
tttcagtggtg atccaccccc aaaatggccc agttttcggt gtgattcaac agcgtogtgt 120
ctatgaogtc tactccatac ttocgaaga aaaagggggg gaaatttttc attcccacat 180
tattagggag atoggtctaa tagtaaaggc caaacctga ttttttatta ggctagatgg 240
totaataatt aaatcagctc ggogaatgag tagcgtcgg gcaaccaga caaggcaatt 300
ctgacagtga caccocctctt aggagacgac cgtggtttcc tataacagca agttcctcgc 360
ggcaaccgtt caggcagagc oggtatggct cagcgcagac gcaacgatcg acaagtogat 420
cgcatcatc gaagaagctg cccaaaaggc cgcgagtctg atcgtttcc cggaagtatt 480
cattccgggc taccocattt gggcgtggct cggcgcagtg aagtacagcc taagctttac 540
ttcacgtat cacgagaatt cgttgagct aggtgacgac cgtatgcgtc gcctccagct 600
ggcgcgogc cgcacaaaaa tcgactcgt catgggctat tcggagoggg aagcggatc 660
gcgctatctg agccaggtgt tcotogaaga gctggcagag atcgttgcca atcggcgcaa 720
gctgaagccc acacaogttg agcgtacgat ctacggcgaa ggcaacggaa cogatttctt 780
cacgcacgac ttgcgttcg gacgcgtcgg tggattgaa tgctgggaa atttccaacc 840
gctcagcaag ttcatgatgt acagcctcgg tgagcaggtc cacgttgcac cgtggcgggc 900
gatgtcccct ctocagccgg atgtttcca actgagcacc gaagcacaac cgacggtcac 960
ccgtcgttac gcaatogaag gccaaacctt tgtgctttgc tcgacgcagg tgatcggacc 1020
tagcgcgac gaaacgttct gcctcaacga cgaacagcgc gcaactgttc cgcaaggatg 1080
tggtgaggog cgcatttacg gcccgatgg aagcagcctt gcgaagcctc tggcggaaaga 1140
tgctgagggg atottgtacg cagagatcga totggagcag attctgotgg cgaaggctgg 1200
agccgatccg gtcgggcaact attcgggccc tgacgtgctg tcggtccagt tcgaccgog 1260
caatcatacg ccagttcatc gcatcggcat tgacggctgc ttgatgtga ataccgcag 1320

```

tgcggtggag aatttcogac tgcgacaagc ggotgagcag gagcgtcagg catccaagcg 1380
gctcggaaacg aaactctttg aacaatccct tctggctgaa gaaccggtcc cagcaaagta 1440
gccataagtt gagagtogcg agatagtatc ggggaaagcc atctctggtc ttccccttta 1500
ttctccaagc ogacatcacc gctgaaagcg ggtttctttg ctaccccgag ttctgatccc 1560
goatcgccgt cgcgtgagat ttgcgtcaga goggacatte aagttgtgtg gcaaggctcg 1620
ccagactgtc cacggaaaat tcccagttct cactcggttc aaggctcagtc gtttgcgtcg 1680
ggcgtgttc ctgtggcgcg ctgaogaatg cgtcctcag gccacaacgt cgagcggctg 1740
ccaagtcato gttgtgcgcc gccaccatgc agatct 1776

```

<210> 16

<211> 1110

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: 合成型腈水解酶

<400> 16

```

atggtttctt acaactccaa gttcttggtc gotactgttc aagctgagcc agtttggttg 60
gacgcagacg ctactattga caagtctatc ggtatcatcg aagaagctgc caaaagggt 120
gcctctttga tcgctttccc agaagtttct attccaggtt acccatactg ggccctggtg 180
ggtgacgtta agtactcttt gtcctttact tccagatate acgagaactc tttggagttg 240
ggtgacgaca gaatgcgtag actgcaattg gotgcccgtg gaaacaaaat tgctttggtc 300
atgggttatt ccgagagaga agctggatct cgttacttgt cccaagtctt catogacgag 360
agagggtgaga ttgttgcaaa tcgtogtaag ttgaagccaa ctcaogttga gogtaccate 420
tacggagaag gtaacggaac cgatttcttg actcaccgact tcgccttcgg aagagttggt 480
ggattgaact gttgggaaca tttccaacct ctgtctaagt tcatgatgta ctcttggtg 540
gagcaagtcg acgttgcttc ttggccagct atgtcccctc ttcagccaga tgttttccaa 600
ttgtccatcg aagccaaocg caccgtcacc agatcctaog ccatcgaagg tcaaaactttt 660
gtcctttgct ctaccaggtt cattggacct tctgctatcg aaacctctg tctgaacgac 720
gaacagagag ctttgttgcc acaaggatgt ggttgggcaa gaatttacgg tccagatgga 780
tctgagcttg ccaagccttt ggotgaagat gctgagggtg ttttgtaocg tgagatcgat 840
ttggagcaaa ttctgctggc caaggctgga gccgatccag tcggtcacta ctccagacct 900
gacgtcttgt ccgtccagtt cgaccctaga aaccacactc cagttcacag aattggtatt 960
gacggtagat tggatgttaa caccagatcc agagtogaga acttcagact gagacaagct 1020
gctgagcagg agagacagcg ttctaagaga cttggaacta aacttttoga acaatctctt 1080
ttggetgaag aacctgtccc agccaagtaa 1110

```

<210> 17

<211> 84

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述: 合成寡核苷酸

<400> 17

catgaattca tggtttctta caactccaag ttcttggtg ctactgttca agctgagcca 60
gtttggttg acgcagacgc tact 84

<210> 18
<211> 90
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述: 合成寡核苷酸

<400> 18
tttgatcgtc ttcccagaag ttttcattcc aggttaccca tactgggcct gttgggtga 60
cgtaagtac tctttgtcct ttacttcag 90

<210> 19
<211> 90
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述: 合成寡核苷酸

<400> 19
aattggtcgc ccgtagaaac aaaattgctt tggatcggg ttattccgag agagaagctg 60
gatctcgtta cttgtcccaa gtcttcacg 90

<210> 20
<211> 90
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述: 合成寡核苷酸

<400> 20
gttgagcgtc ccatctacgg agaaggtaac ggaaccgatt tcttgactca cgacttcgcc 60
ttcggaagag ttggtggtgatt gaactggttg 90

<210> 21
<211> 90
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：合成寡核苷酸

<400> 21

```
agtccacggtt gcttcttggc cagctatgtc ccctcttcag ccagatggtt tccaattgtc 60
catcgaagcc aagccacog tcaccagatc 90
```

<210> 22

<211> 90

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：合成寡核苷酸

<400> 22

```
gaccttctgc tatcgaacc ttctgtctga acgacgaaca gagagcttg ttgccacaag 60
gatgtggttg ggcaagaatt tacggtccag 90
```

<210> 23

<211> 90

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：合成寡核苷酸

<400> 23

```
tacgctgaga tcgatttggg gcaaattctg ctggccaagg ctggagccga tccagtoggt 60
cactactcca gacctgacgt cttgtccgtc 90
```

<210> 24

<211> 90

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：合成寡核苷酸

<400> 24

```
tagattggat gttaacacca gatccagagt cgagaacttc agactgagac aagctgctga 60
gcaggagaga caggcttcta agagacttgg 90
```

<210> 25

<211> 84
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述: 合成寡核苷酸

<400> 25
catgaattct tacttggctg ggacaggttc ttcagccaaa agagattgtt cgaaaagttt 60
agttccaagt ctcttagaag cctg 84

<210> 26
<211> 90
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述: 合成寡核苷酸

<400> 26
tgggtttaac atccaatcta ccgtaatac caattctgtg aactggagtg tggtttctag 60
ggtogaactg gacggacaag acgtcaggtc 90

<210> 27
<211> 90
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述: 合成寡核苷酸

<400> 27
tccaaatcga tctcagcgtc caaaataccc tcagcatctt cagccaaagg cttggcaagc 60
tcagatccat ctggaccgta aattcttgcc 90

<210> 28
<211> 90
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述: 合成寡核苷酸

<400> 28
ggtttcgata gcagaaggtc caatgacctg ggtagagcaa aggacaaaag tttgacctc 60
gatggcgtag gatctggtga oggtggcgtt 90

<210> 29
<211> 90
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述: 合成寡核苷酸

<400> 29
gccagaagc aacgtggact tgctaccca aggagtacat catgaactta gacagaggtt 60
ggaaatgttc ccaacagttc aatccaccaa 90

<210> 30
<211> 90
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述: 合成寡核苷酸

<400> 30
ccgtagatgg tacgctcaac gtgagttgce ttcaacttac gacgatttgc aacaatctca 60
cctctctogt cgatgaagac ttgggacaag 90

<210> 31
<211> 90
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述: 合成寡核苷酸

<400> 31
gtttctacgg gcagccaatt gcagtctacg cattctgtcg tcaccaact ccaaagagtt 60
ctcgtgatat ctggaagtaa aggacaaaga 90

<210> 32
<211> 90
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述: 合成寡核苷酸

<220>

<400> 32

cttctgggaa agcgatcaaa gaggcacct tttgggcagc ttcttcgatg ataccgatag 60
acttgatcaat agtagcgtct ggcaccaacc 90

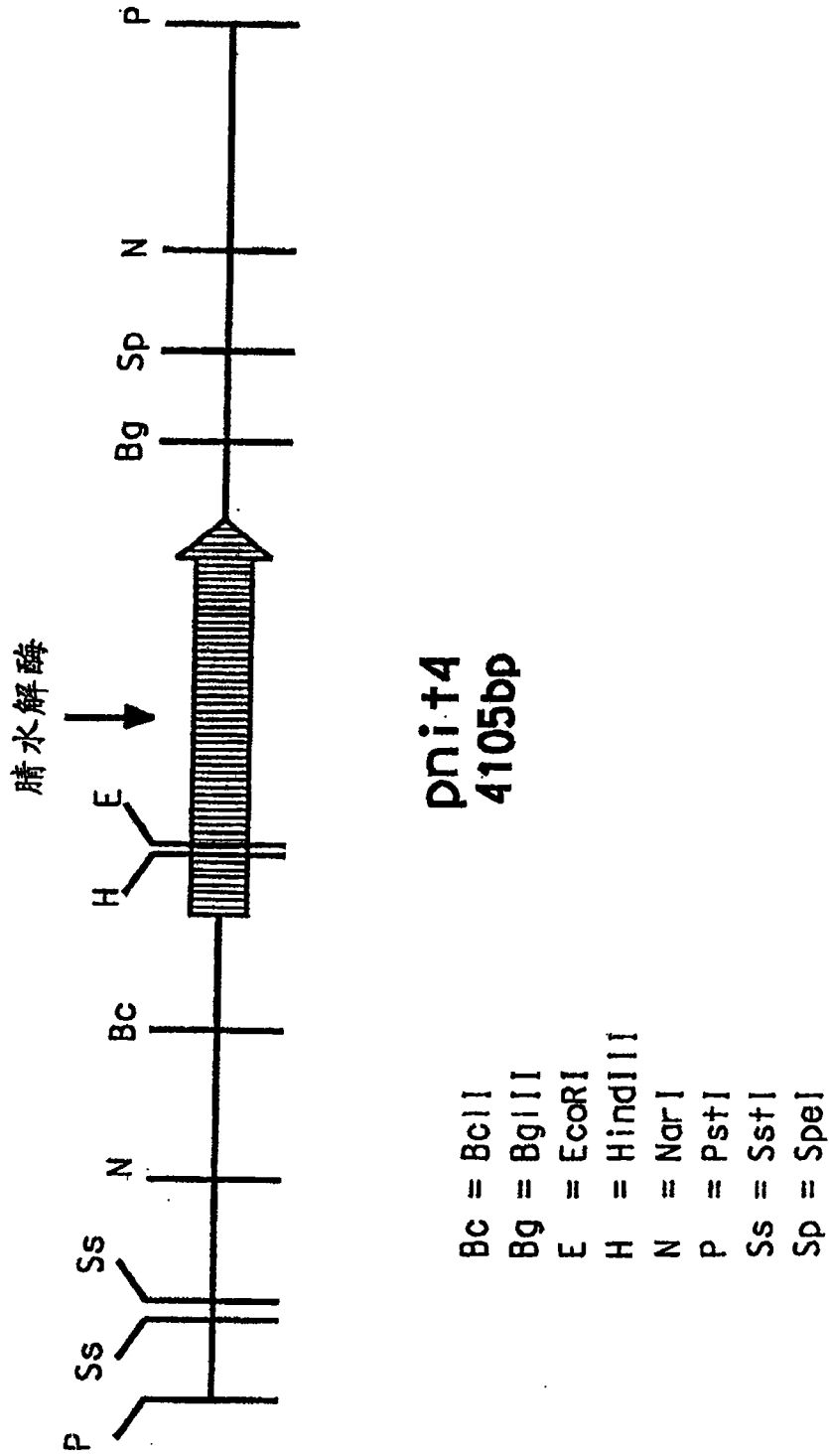


图 1