

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷



[12] 发明专利申请公开说明书

[21] 申请号 02811380.2

C12N 9/02

C12N 9/12

C12N 9/88

C12N 9/90

C12N 15/53

C12N 15/54

C12N 15/60

[43] 公开日 2005 年 6 月 22 日

[11] 公开号 CN 1630718A

[22] 申请日 2002.6.5 [21] 申请号 02811380.2

[30] 优先权

[32] 2001.6.6 [33] US [31] 60/296,299

[86] 国际申请 PCT/EP2002/006171 2002.6.5

[87] 国际公布 WO2002/099095 英 2002.12.12

[85] 进入国家阶段日期 2003.12.5

[71] 申请人 DSM IP 资产公司

地址 荷兰海尔伦

[72] 发明人 艾伦·贝里 沃纳·布雷策尔

马库斯·许布林

鲁阿尔·洛佩兹-尤利巴里

安妮·F·迈耶

亚历克西·耶利西夫

[74] 专利代理机构 北京市柳沈律师事务所

代理人 巫肖南 封新琴

C12N 15/61 C12N 1/20 C12N 1/21

C12P 7/02 C12P 17/06 C12N 9/10

C12P 23/00

权利要求书 5 页 说明书 99 页 序列表 142 页

[54] 发明名称 改进的类异戊二烯生产

[57] 摘要

提供了分离的多核苷酸，其编码具有甲羟戊酸途径各酶活性的多肽，所述酶例如，甲羟戊二酸单酰-CoA 还原酶，异戊烯二磷酸异构酶，甲羟戊二酸单酰-CoA 合成酶，甲羟戊酸激酶，磷酸甲羟戊酸激酶，或二磷酸甲羟戊酸脱羧酶，该多核苷酸可用于重组生产类异戊二烯化合物，如类胡萝卜素，例如八氢番茄红素，番茄红素， β -胡萝卜素，玉米黄质，角黄素，虾青素，福寿草黄素(adonixanthin)，隐黄素，海胆酮和福寿草红素(adonirubin)。还提供了表达载体，培养的细胞，以及生产类异戊二烯化合物的方法。

I S S N 1 0 0 8 - 4 2 7 4

1. 分离的多肽，其包括选自下述的氨基酸序列：
- (a) SEQ ID NO: 43 的 1-340 残基所示的氨基酸序列；
- 5 (b) SEQ ID NO: 45 的 1-349 残基所示的氨基酸序列；
- (c) SEQ ID NO: 47 的 1-388 残基所示的氨基酸序列；
- (d) SEQ ID NO: 49 的 1-378 残基所示的氨基酸序列；
- (e) SEQ ID NO: 51 的 1-305 残基所示的氨基酸序列；
- (f) SEQ ID NO: 53 的 1-332 残基所示的氨基酸序列；
- 10 (g) 选自 SEQ ID NO: 43, 45, 47, 49, 51 或 53 所示氨基酸序列的片段，其中所述片段至少具有 30 个连续的氨基酸残基；
- (h) 选自 SEQ ID NO: 43, 45, 47, 49, 51 或 53 所示多肽的片段的氨基酸序列，该片段具有羟甲戊二酰-CoA 还原酶(HMG-CoA 还原酶)，异戊烯二磷酸异构酶，羟甲戊二酰-CoA 合成酶(HMG-CoA 合成酶)，甲羟戊酸激酶，磷酸甲羟戊酸激酶，或二磷酸甲羟戊酸脱羧酶的活性；
- 15 (i) 严谨条件下能够与杂交探针杂交的多核苷酸所编码的多肽的氨基酸序列，所述杂交探针含有 SEQ ID NO: 42 或 SEQ ID NO: 42 的互补序列的至少 30 个连续核苷酸，其中所述多肽具有 HMG-CoA 还原酶，异戊烯二磷酸异构酶，HMG-CoA 合成酶，甲羟戊酸激酶，磷酸甲羟戊酸激酶，或二磷酸甲羟戊酸脱羧酶的活性；和
- 20 (j) SEQ ID NO: 43, 45, 47, 49, 51 或 53 的保守修饰的变体。
2. 分离的多肽，其包括选自下述的氨基酸序列：
- (a) SEQ ID NO: 159 的 1-287 残基所示氨基酸序列；
- (b) SEQ ID NO: 159 的至少 30 个连续的氨基酸残基；
- 25 (c) SEQ ID NO: 159 的片段的氨基酸序列，所述片段具有法呢基二磷酸合成酶(FPP 合成酶)的活性；
- (d) 在严谨条件下能够与杂交探针杂交的多核苷酸所编码的多肽的氨基酸序列，所述杂交探针含有跨 SEQ ID NO: 157 的 295-1158 位核苷酸或其互补序列的至少 30 个连续核苷酸，其中所述多肽具有 FPP 合成酶的活性；
- 30 和
- (e) SEQ ID NO: 159 的保守修饰的变体。

3. 分离的多肽, 其含有选自下述的氨基酸序列:
- (a) SEQ ID NO: 160 的残基 1-142 所示的氨基酸序列;
 - (b) SEQ ID NO: 160 的至少 30 个连续的氨基酸残基;
 - (c) SEQ ID NO: 160 的片段的氨基酸序列, 所述片段具有 1-脱氧木酮糖
- 5 -5-磷酸合成酶(DXPS)的活性;
- (d) 在严谨条件下能够与杂交探针杂交的多核苷酸所编码的多肽的氨基酸序列, 所述杂交探针含有 SEQ ID NO: 157 的 1185-1610 位核苷酸或其互补序列的至少 30 个连续核苷酸, 其中所述多肽具有 DXPS 的活性;
 - (e) SEQ ID NO: 160 的保守修饰的变体。
- 10 4. 分离的多肽, 其含有选自下述的氨基酸序列:
- (a) SEQ ID NO: 178 的 1-390 残基所示的氨基酸序列;
 - (b) SEQ ID NO: 178 的至少 30 个连续氨基酸残基;
 - (c) SEQ ID NO: 178 的多肽片段的氨基酸序列, 所述片段具有乙酰-CoA
- 15 乙酰基转移酶的活性;
- (d) 在严谨条件下能够与杂交探针杂交的多核苷酸所编码的多肽的氨基酸序列, 所述杂交探针含有 SEQ ID NO: 177 的 1-1170 位核苷酸或其互补序列的至少 30 个连续核苷酸, 其中所述多肽具有乙酰-CoA 乙酰基转移酶的活性; 和
 - (e) SEQ ID NO: 178 保守修饰的变体。
- 20 5. 分离的多肽, 其含有选自下述的氨基酸序列:
- (a) SEQ ID NO: 179 的 1-240 残基所示的氨基酸序列;
 - (b) SEQ ID NO: 179 的至少 30 个连续的氨基酸残基;
 - (c) SEQ ID NO: 179 的多肽片段的氨基酸序列, 所述片段具有乙酰乙酰
- 25 -CoA 还原酶的活性;
- (d) 在严谨条件下能够与杂交探针杂交的多核苷酸所编码的多肽氨基酸序列, 所述杂交探针含有 SEQ ID NO: 177 或其互补序列的 1258-1980 位核苷酸的至少 30 个连续核苷酸, 其中所述多肽具有乙酰乙酰-CoA 还原酶的活性; 和
 - (e) SEQ ID NO: 179 保守修饰的变体。
- 30 6. 分离的多核苷酸序列, 其含有选自下述的核苷酸序列: SEQ ID NO: 42; SEQ ID NO: 42 的变体, 该变体含有一个或多个根据副球菌属的菌株

R1534 密码子使用表的取代; SEQ ID NO: 42 的片段, 该片段编码具有选自下述活性的多肽: 羟甲戊二酰-CoA 还原酶(HMG-CoA 还原酶), 异戊烯二磷酸异构酶, 羟甲戊二酰-CoA 合成酶(HMG-CoA 合成酶), 甲羟戊酸激酶, 磷酸甲羟戊酸激酶, 或二磷酸甲羟戊酸脱羧酶; 或在严谨条件下能够与杂交
5 探针杂交的多核苷酸序列, 该探针的核苷酸序列由 SEQ ID NO: 42 或 SEQ ID NO: 42 的互补序列的至少 30 个连续核苷酸构成, 所述的多核苷酸编码具有选自下述活性的多肽: HMG-CoA 还原酶, 异戊烯二磷酸异构酶, HMG-CoA 合成酶, 甲羟戊酸激酶, 磷酸甲羟戊酸激酶或二磷酸甲羟戊酸脱羧酶。

7. 分离的多核苷酸序列, 其含有选自下述的多核苷酸序列: SEQ ID NO:
10 157 的核苷酸序列; SEQ ID NO: 157 变体, 该变体含有一个或多个根据副球菌属的菌株 R1534 密码子使用表的取代; SEQ ID NO: 157 的片段, 该片段编码具有下述活性的多肽: 法呢基二磷酸合成酶(FPP 合成酶)活性, 1-脱氧-D-木酮糖-5-磷酸合成酶活性或具有 XseB 活性的多肽; 或在严谨条件下能够与杂交探针杂交的多核苷酸序列, 该探针的核苷酸序列由 SEQ ID NO:
15 157 或 SEQ ID NO: 157 的互补序列的至少 30 个连续核苷酸构成, 所述的多核苷酸编码具有下述活性的多肽: FPP 合成酶活性, 1-脱氧-D-木酮糖-5-磷酸合成酶活性或 XseB 活性。

8. 分离的多核苷酸序列, 其含有选自下述的多核苷酸序列: SEQ ID NO:
177 的核苷酸序列; SEQ ID NO: 177 的变体; 该变体含有一个或多个根据副
20 球菌属的菌株 R1534 密码子使用表的取代; SEQ ID NO: 177 的片段, 该片段编码具有选自乙酰-CoA 乙酰基转移酶活性或乙酰乙酰-CoA 还原酶活性的多肽; 或在严谨条件下能够与杂交探针杂交的多核苷酸序列, 该探针的核苷酸序列由 SEQ ID NO: 177 或 SEQ ID NO: 177 的互补序列的至少 30 个连续核苷酸构成, 所述的多核苷酸编码具有选自乙酰-CoA 乙酰基转移酶活
25 性或乙酰乙酰-CoA 还原酶活性的多肽。

9. 分离的多核苷酸序列, 其包括选自下述的核苷酸序列: SEQ ID NO: 42, SEQ ID NO: 157, SEQ ID NO: 177 或其组合。

10. 含有权利要求 6, 7, 8, 或 9 的多核苷酸序列的表达载体。

11. 选自下述的表达载体: pBBR-K-mev-op16-1, pBBR-K-mev-op16-2,
30 pDS-mvaA, pDS-idi, pDS-hcs, pDS-mvk, pDS-pmk, pDS-mvd, pDS-His-mvaA, pDS-His-idi, pDS-His-hcs, pDS-His-mvk, pDS-His-pmk, pDS-His-mvd,

- pBBR-K-Zea4, pBBR-K-Zea4-up, pBBR-K-Zea4-down, pBBR-K-PcrtE-crtE-3, pBBR-tK-PcrtE-mvaA, pBBR-tK-PcrtE-idi, pBBR-tK-PcrtE-hcs, pBBR-tK-PcrtE-mvk, pBBR-tK-PcrtE-pmk, pBBR-tK-PcrtE-mvd, pBBR-K-PcrtE-mvaA-crtE-3, pDS-His-phaA, pBBR-K-PcrtE-crtW, 5 pBBR-K-PcrtE-crtWZ, pBBR-K-PcrtE-crtZW 或其组合。

12. 培养的细胞或该细胞的后代, 其中所述细胞含有权利要求 6, 7, 8 或 9 的多核苷酸序列或权利要求 10 或 11 的表达载体, 其中所述细胞表达由所述多核苷酸序列所编码的多肽。

13. 生产类胡萝卜素的方法, 包括在允许表达所述多核苷酸序列编码的多肽的条件下, 培养权利要求 12 的细胞, 并从细胞或者该细胞培养基中分离类胡萝卜素。

14. 制备产类胡萝卜素细胞的方法, 包括:

- (a) 将编码甲羟戊酸途径中的酶的多核苷酸序列导入细胞, 所述酶在该细胞中能够表达; 和
15 (b) 选择出含有步骤(a)多核苷酸序列的细胞, 并且该细胞所生产类胡萝卜素的水平为该细胞导入多核苷酸序列之前的约 1.1-1000 倍。

15. 一种改造细菌以使其生产类异戊二烯化合物的方法, 包括:

- (a) 在允许表达类异戊二烯化合物的条件下, 在培养基中培养亲代细菌, 并从该培养基中筛选出突变细菌, 所述突变细菌生产的类异戊二烯化合物比亲代细菌高约 1.1-1000 倍;
20

(b) 将含有 SEQ ID NO: 42 多核苷酸序列的表达载体导入突变细菌, 所述多核苷酸序列与表达调控序列可操作连接; 和

(c) 筛选出含有表达载体并且生产的类异戊二烯化合物比步骤(a)中的突变体高出至少约 1.1 倍的细菌。

- 25 16. 副球菌属的微生物, 所述微生物具有下述特征:

(i) 使用 GeneCompar v. 2.0 软件, 缺口罚分为 0%的条件下, 利用同源性计算获得的相似性的矩阵, 与 SEQ ID NO: 12 的序列相似性>97%; 在 81.5 °C, 使用 DNA:DNA 杂交, 与菌株 R-1512, R1534, R114 或 R-1506 的同源性 >70%;

- 30 其基因组 DNA 的 G+C 含量与 R114, R-1512, R1534 和 R-1506 的基因组 DNA 的 G+C 含量相比, 变化小于 1%; 和

使用实施例 2 的 AFLP 方法, 平均 DNA 指纹与菌株 R-1512, R1534, R114 和 R-1506 具有约 58%的相似性, 附带条件为该微生物不是副球菌属菌种(MBIC3966);

- 5 (ii) 含有至少约 75%的细胞膜总脂肪酸的 18:1w7c; 不具有使用福寿草醇, i-赤藓糖醇, 龙胆二糖, β -甲基葡萄糖苷, D-山梨糖醇, 木糖醇和奎尼酸作为碳源用于生长的能力; 和

能够使用 L-天冬酰胺和 L-天冬氨酸作为碳源用于生长, 附带条件为该微生物不是副球菌属菌种(MBIC3966); 或

- (iii) 能够在 40℃ 生长;
- 10 能够在含有 8% NaCl 的培养基中生长;
- 能够在 pH 9.1 的培养基中生长; 和
- 形成黄色-橙色色素的菌落, 附带条件为该微生物不是副球菌属菌种(MBIC3966)。

改进的类异戊二烯生产

5 本发明涉及新的用于类异戊二烯生物合成途径的多核苷酸和多肽序列。更具体地，本发明涉及重组制备的细胞，该细胞显示出改进的玉米黄质产量。还提供了制备和使用这种细胞系的方法。

类胡萝卜素是一种重要的 C-40 类异戊二烯化合物商品，可作为人类的营养性支持物、药物和食用色素，还可作为动物饲料的色素。目前重要的
10 工业类胡萝卜素的生产主要是通过化学合成(β -胡萝卜素，角黄素和虾青素)或者从天然材料中提取(叶黄素提取自万寿菊，辣椒红提取自红辣椒(paprika))。但是，使用微生物生产类胡萝卜素已经取得了几例成功。例如，使用真菌三孢布拉霉(*blakeslea trispora*)(US 5,328,845)进行发酵或使用耐盐藻类 *Dunaliella salina* 进行池塘养殖[Borowitzka, J. Biotechnol. 70:313-321
15 (1999)]都可生产 β -胡萝卜素。还有三孢布拉霉(*B.trispora*)生产番茄红素的报道(WO 00/77234)。使用酵母(*Phaffia rhodozyma*, (最近更名为 *Xanthophyllomyces dendrorous*))(US 6,015,684)发酵或使用藻类 *Haematococcus pluvialis* 在光生物反应器内或开放的池塘内可以生产虾青素[Lorenz 和 Cysewski, Trends Biotechnol. 18:160-167 (1999); Olaizola, J. Appl. Phycol. 12:
20 499-506 (2000)]。但是这种微生物生产系统生产的类胡萝卜素数量过低，工业规模生产则难以产生经济效益。

在 1960 年代中期，Hoffmann-La Roche 的科学家们分离了几种海洋细菌，其能够生产黄色的类胡萝卜素玉米黄质，该物质可用于家禽色素形成以及预防人类中与年龄相关的黄斑变性。一种细菌，其显示出极有潜力的
25 玉米黄质产率，被命名为 R-1512，已保藏在美国典型培养物保藏中心(American Type Culture Collection)(ATCC, Manassas, VA,USA)，保藏号为 ATCC 21588 (US 3,891,504)。使用当时公认的分类标准(按照 Eidgenossische Technische Hochschule (Zurich)和 National Collection of Industrial Bacteria, Torry Research Station (Aberdeen Scotland)进行的分类)，这个能够生产玉米
30 黄质的有机体被归类为黄杆菌属(*Flavobacterium*)，但是没有命名种名。

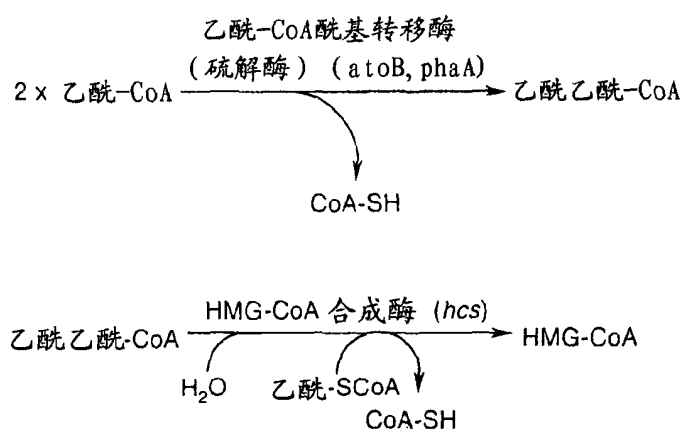
然后进行了大量的诱变和筛选过程，以分离得到具有更高玉米黄质产

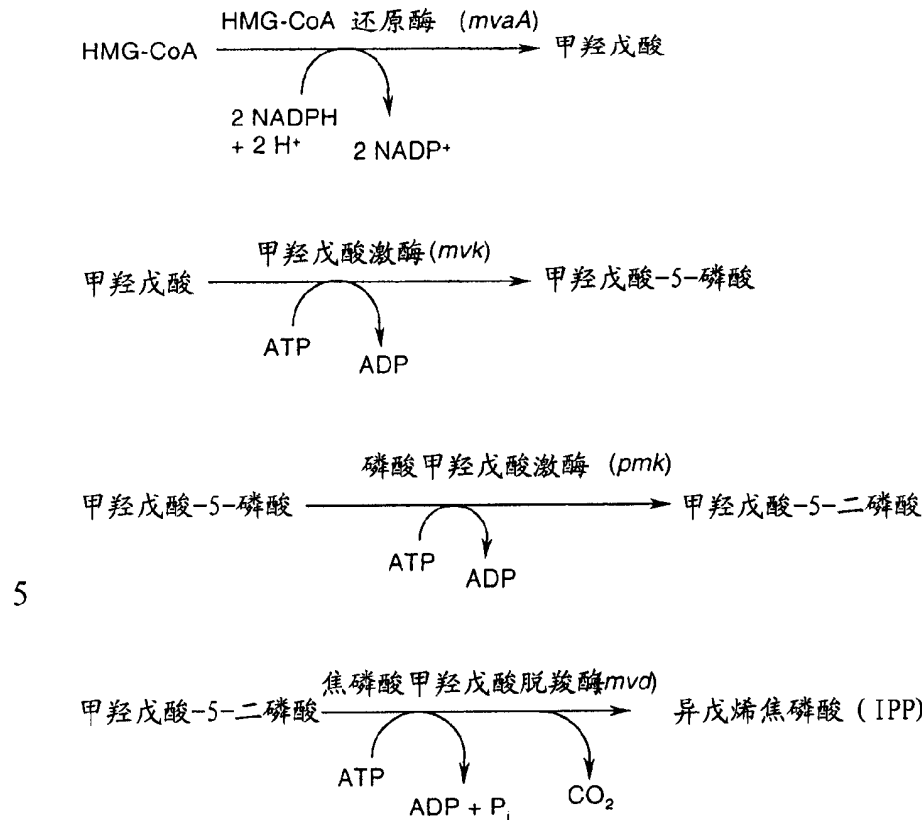
率的 R-1512 突变体。在本文描述的研究中,有两种这样的突变体比较显著。按照玉米黄质产率顺序,列出这两个突变体: R1534 和 R114。在过去几年中还用过各种其它的突变体,用于类胡萝卜素的生物合成的生化研究 [Goodwin, Biochem. Soc. Symp. 35:233-244 (1972); McDermott 等, Biochem.J. 134:1115-1117(1973); Britton 等, Arch.Microbiol. 113:33-37(1977); Mohanty 等, Helvetica Chimica Acta 83:2036-2053 (2000)]。

通过常规利用 R-1512 菌株衍生的突变体发展一种商业生命力的发酵过程来生产玉米黄质的早期尝试没有成功。但是,随着分子生物学的进展,发展一种更高玉米黄质产率菌株的可能性随之出现。为此,采取的第一个步骤是克隆并测序 R-1534 菌株的类胡萝卜素基因簇(US 6,087,152,将该文献全文并入作为参考)。US 6,087,152 公开了类胡萝卜素基因在大肠杆菌和枯草芽孢杆菌(*Bacillus Subtilis*)中功能性表达,使得这些宿主产生玉米黄质。US 6,087,152 还公开了通过修饰类胡萝卜素基因簇或加入自产虾青素细菌的基因,就有可能产生玉米黄质之外的各种类胡萝卜素(EP 872,554)。而且,EP 872,554 公开了通过向多拷贝质粒导入克隆的类胡萝卜素基因簇,可以提高 R1534 菌株的类胡萝卜素产量。

尽管类异戊二烯化合物存在大量的结构多样性,但是都是由一个共同的 C-5 前体,异戊烯焦磷酸(IPP)生物合成而来的。直到 1990 年代早期,才基本上接受了 IPP 是通过甲羟戊酸途径在所有有机体内合成,即使有些实验结果与这个生物发生模式不一致 [Eisenreich 等, Chemistry and Biology 5:R211-R233 (1998)]。

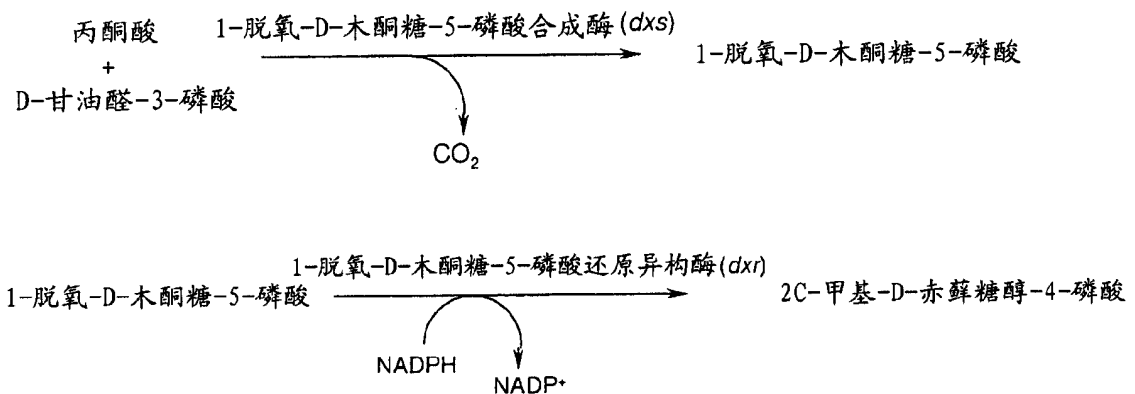
甲羟戊酸途径:

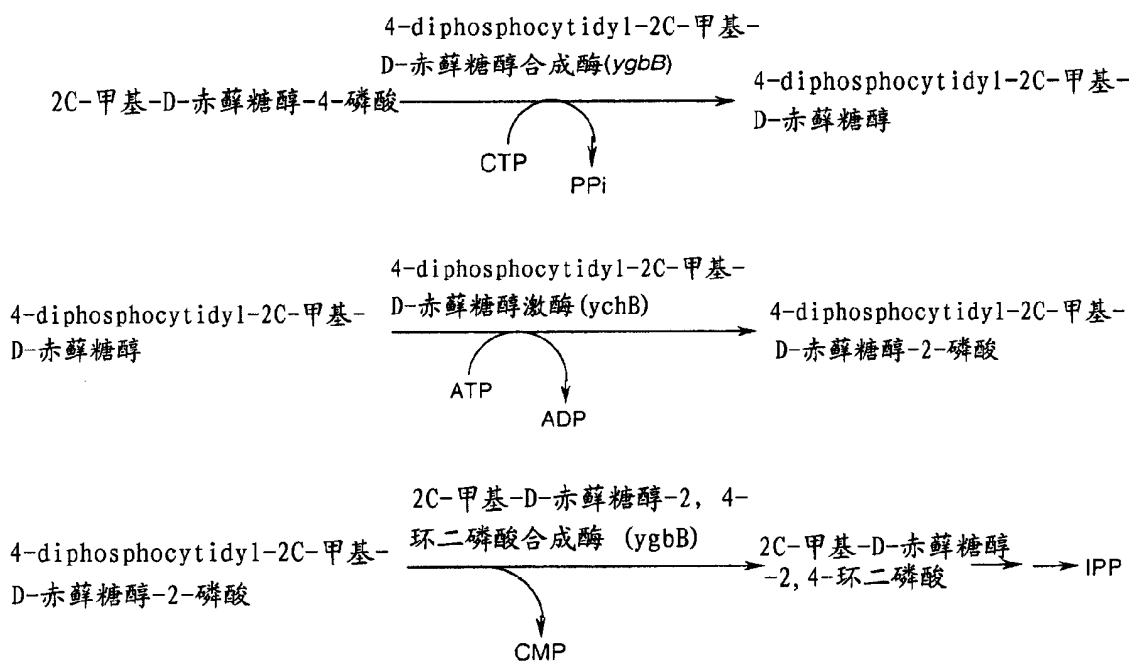




自从发现了 IPP 生物合成的替代性途径，脱氧木糖醇(DXP)途径之后，
 10 这种不一致就找到了解释(注释：在科学文献中，IPP 生物合成的替代性途径已被命名了多个名字(DXP 途径，DOXP 途径，MEP 途径，GAP/丙酮酸途径以及非甲羟戊酸途径。本文中为简化，仅使用 DXP 途径)。DXP 途径的前 5 个反应已被鉴定[Herz 等，Proc.Nat.Acad.Sci 97:2486-2490(2000)]，但是其后生成 IPP 的步骤尚不清楚。

15 DXP 途径：





5

McDermott 等(同上)和 Britton 等[J.Chem.Soc.Chem.Comm. p.27(1979)]揭示: 源自原 Roche 分离群的产玉米黄质突变株, 其粗提物中的玉米黄质掺入了标记的甲羟戊酸。没有理由再质疑这个通过甲羟戊酸途径进行 IPP 生物合成的证据, 这项研究是在发现 DXP 途径之前进行的, 已有报道称一些细菌(链霉菌属的种类)同时具备 IPP 合成的这两种途径, 并且这些途径的表达受时间(temporally)调控[Seto 等, Tetrahedron Lett. 37:7979-7982 (1996); Dairi 等, Mol.Gen.Genet. 262: 957-964 (2000)]。另外, 目前只有少数的真细菌显示具有 IPP 合成的甲羟戊酸途径。这些细菌中编码甲羟戊酸途径各种酶的基因已被克隆并测序[Wilding 等, J.Bacteriol. 182:4319-4327 (2000); Takagi 等, J.Bacteriol 182:4153-4157(2000)]。

已有几个范例显示: 代谢工程的应用已成功地改变或改进了微生物生产类胡萝卜素的产量[Lagarde 等, Appl.Env.Microbiol. 66:64-72 (2000); Wang 等, Biotechnol.Bioeng. 62:235-241(1999); Wang 等, Biotechnol. Prog. 16:922-926 (2000)(作为本文的参考); Sandmann 等, Trends Biotechnol. 17:233-237(2000); Misawa and Shimada, J.Biotechnol. 59:169-181(1998); Matthews and Wurtzel, Appl. Microbiol. Biotechnol. 53: 396-400(2000); Albrecht et al., Nature Biotechnol. 18:843-846 (2000); Schmidt-Dannert et al. Nature Biotechnol. 18:750-753 (2000)]。例如, 大肠杆菌, 不具备胡萝卜素形

成作用(non-carotenogenic)的细菌,可通过导入克隆的类胡萝卜素(crt)基因进行改造使细菌生产类胡萝卜素,所述细菌: *Agrobacterium aurantiacum*, 草生欧文氏菌(*Erwinia herbicola*)或菠萝欧文氏菌(*Erwinia uredovora*)(Misawa 和 Shimada, 同上)。Harker 和 Bramley[FEBS Lett. 448:115-119 (1999)]和 Matthews 和 Wurtzel(同上)公开:通过过量表达编码 1-脱氧-D-木糖醇-5-磷酸合成酶(DXPS)的基因,可提高这种改造过的大肠杆菌菌株中的类胡萝卜素产量,该酶是 DXP 途径中的第一个酶(大肠杆菌只有类异戊二烯生物合成的 DXP 途径,并不使用甲羟戊酸途径[Lange et al., Proc.Nat.Acad.Sci. 97:13172-13177 (2000)]。Harker和Bramley(同上)还公开了在过量产生 DXPS 的细胞中,类异戊二烯化合物泛醌-8 的升高。这些结果均支持如下的假说:由于 DXPS 体内活性不足,导致 IPP 有效性受限,这就限制了类胡萝卜素以及其它类异戊二烯化合物在改造菌株中的产量。使用类似的大肠杆菌体系, Kim 和 Keasling[Biotechnol. Bioeng. 72:408-415 (2001)]公开了组合的过量表达编码 DXPS 和 DXP 途径的第二个酶——DXP 还原异构酶(1-脱氧-D-木酮糖-5-磷酸还原异构酶)(1-deoxy-D-xylulose-5-phosphate reductoisomerase)的基因,所得到的类胡萝卜素产量高于仅仅过量表达编码 DXPS 的基因的产量。

所有这些研究都是对大肠杆菌进行改造使其生产类胡萝卜素。相应地,这些研究的一个不足之处在于:这些重组大肠杆菌菌株生产的类胡萝卜素产量即使跟工业化生产类胡萝卜素的非重组的微生物相比,也非常低。另外,通过遗传工程改造 IPP 生物合成途径在细菌中改进类胡萝卜素生产只应用 DXP 途径进行 IPP 合成的生物体中才能看到。对于经甲羟戊酸途径生成 IPP 的细菌则没有类似的研究报道。

已经报道了在酵母中通过甲羟戊酸的代谢工程以改进类异戊二烯化合物的生产。例如, WO 00/01649 公开了当 3-羟基-3-甲基戊二酸单酰辅酶 A 还原酶(HMG-CoA 还原酶)被过量表达时,类异戊二烯化合物在酿酒酵母(*Saccharomyces cerevisiae*)中的产量有所提高。但是,未显示该方案能够改进细菌中类胡萝卜素的产量,特别是它没有显示出通过扩增甲羟戊酸途径各基因的表达可以改进细菌中类胡萝卜素的产量。已经发现,一些真核生物 [Campos 等, Biochem.J. 353:59-67(2001)] 和链霉菌属细菌菌株 CL190(Takagi 等, 同上)的甲羟戊酸途径基因可以在大肠杆菌中表达,未报

道在这些菌株中类异戊二烯的产量有所提高。

除了形成 IPP[通过 DXP 或甲羟戊酸途径]以及将法呢基焦磷酸转换为各种其它类异戊二烯(如类胡萝卜素, 苯醌)的反应之外, 还已知两种反应涉及类异戊二烯的生物合成。IPP 异构酶使 IPP 及其异构体——二甲基烯丙基焦磷酸(DMAPP)两者之间相互转变。存在两种形式的 IPP 异构酶, 众所周知真核生物和一些细菌中的 1 型酶, 而新近鉴定的 2 型酶是 FMN-和 NADP(H)依赖性的[Kaneda 等, Proc.Nat.Acad.Sci. 98:932-937 (2001)]。

几篇报道公开了在改造的产类胡萝卜素大肠杆菌中, 扩增天然或异源的 1 型 IPP 异构酶(idi)基因可以促进类胡萝卜素的产量[Kajiwara 等, Biochem.J. 324:421-426 (1997);Verdoes 和 van Ooyen, Acta Bot.Gallica 146: 43-53 (1999);Wang 等, 同上]。在一篇报道(Wang 等, 同上)中进一步公开: 当 idi 和 crtE(GGPP 合成酶/geranyl-geranyl diphosphate synthase)基因组合过量表达时, 过量表达编码 FPP 合成酶(法呢基二磷酸合成酶(farnesyl diphosphate synthase))的 ispA 基因, 可提高改造的产类胡萝卜素大肠杆菌菌株的类胡萝卜素产量。但是, 对于 IPP 生物合成途径而言, 没有显示出过量表达编码 IPP 异构酶或 FPP 合成酶的基因可以改进天然产类胡萝卜素微生物的类胡萝卜素产量。同样, 上述的大肠杆菌菌株所产生的类胡萝卜素水平很低, 而且, 在原本类胡萝卜素产量较高的工业微生物中, 这种策略还未显示出作用。

总之, 没有任何现有证据证明: 提高编码甲羟戊酸途径的各酶基因的表达可以改进天然产类胡萝卜素(carotenogenic)细菌、或天然不生产类胡萝卜素而被改造使其生产类胡萝卜素的细菌的类胡萝卜素产量。

本发明的一个实施方案是分离的多肽, 其含有选自下列的氨基酸序列: (a)SEQ ID NO: 43 中 1-340 残基所示的氨基酸序列; (b)SEQ ID NO: 45 中 1-349 残基所示的氨基酸序列; (c) SEQ ID NO: 47 中 1-388 残基所示的氨基酸序列; (d)SEQ ID NO: 49 中 1-378 残基所示的氨基酸序列; (e)SEQ ID NO: 51 中 1-305 残基所示的氨基酸序列; (f)SEQ ID NO: 53 中 1-332 残基所示的氨基酸序列; (g)选自 SEQ ID NO: 43,45,47,49,51 和 53 所示氨基酸序列的片段, 其中所述片段至少具有 30 个连续的氨基酸残基; (h)选自 SEQ ID NO: 43,45,47,49,51 和 53 所示多肽片段的氨基酸序列, 所述片段具有 HMG-CoA 还原酶, 异戊烯二磷酸异构酶, 羟基甲基戊二酸单酰-CoA 合成酶(HMG-CoA

合成酶), 甲羟戊酸激酶, 磷酸甲羟戊酸激酶, 或二磷酸甲羟戊酸脱羧酶的活性; (i) 严谨条件下与杂交探针杂交的多核苷酸所编码的多肽氨基酸序列, 所述杂交探针含有 SEQ ID NO: 42 或其互补序列的至少 30 个连续核苷酸, 其中所述多肽具有 HMG-CoA 还原酶, 异戊烯二磷酸异构酶, HMG-CoA 合成酶, 异戊烯二磷酸异构酶, 甲羟戊酸激酶, 磷酸甲羟戊酸激酶, 或二磷酸甲羟戊酸脱羧酶的活性; 以及(j)SEQ ID NO: 43,45,47,49,51 或 53 经保守修饰而形成的变体。

如上所述, 本发明包括 SEQ ID NO: 43,45,47,49, 51 和 53, 其分别相应于下述甲羟戊酸途径各酶的多肽序列: 羟甲基戊二酸单酰 CoA(HMG-CoA)(hydroxymethyl glutaryl CoA)还原酶, 异戊烯二磷酸(IPP)异构酶, HMG-CoA 合成酶, 甲羟戊酸激酶, 磷酸甲羟戊酸激酶和二磷酸甲羟戊酸脱羧酶。本发明还包括每个鉴定序列的至少 30 个连续氨基酸, 或足够数目的连续氨基酸, 以定义生物活性分子。

本发明还包括选自 SEQ ID NO: 43,45,47,49,51 和 53 多肽的片段。该片段的长度至少应约为 30 个氨基酸长, 且必须具有鉴定多肽的活性, 例如对于 SEQ ID NO: 43 而言, 其落入本发明范围之内的片段具有 HMG-CoA 还原酶的活性。在本文中, 各片段活性的测定在实施例 1 中描述。实施例 1 所述的试验中具有高于背景活性的片段被认为具有生物学活性, 属于本发明的范围。

本发明还包括多核苷酸所编码多肽的氨基酸序列, 所述多核苷酸在严谨条件(如上面所定义)下能够与杂交探针杂交, 所述探针含有 SEQ ID NO: 42(即甲羟戊酸操纵子)或其互补序列的至少 30 个连续核苷酸。该多核苷酸必须编码甲羟戊酸途径的至少一种酶。本发明中, “杂交探针” 指含有 SEQ ID NO: 42 的大约 10-9066 核苷酸的多核苷酸序列。

在该实施方案中, 分离的多肽可以具有 SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51 或 SEQ ID NO: 53 的氨基酸序列。或者, 该分离的多肽可含有选自各个氨基酸序列区域的、大约 30 个连续的氨基酸, 与来自不同种类, 且具有相同功能的酶相比, 所述氨基酸序列具有最小的同源性。因此, 本发明多肽可含有: 例如 SEQ ID NO: 43 的 68-97, SEQ ID NO: 45 的 1-30, SEQ ID NO: 47 的 269-298, SEQ ID NO: 49 的 109-138, SEQ ID NO: 51 的 198-227 或 SEQ ID NO: 53 的 81-110 氨基

酸。

本发明的另一实施方案是分离的多肽，其具有选自下述的氨基酸序列：
(a)SEQ ID NO: 159 的 1-287 残基所示的氨基酸序列；(b)SEQ ID NO: 159 的
至少 30 个连续的氨基酸残基；(c)SEQ ID NO: 159 的片段的氨基酸序列，所
述片段具有法呢基-二磷酸合成酶(FPP 合成酶)的活性；(d)在严谨条件下与
杂交探针杂交的多核苷酸所编码的多肽氨基酸序列，所述杂交探针含有
ispA 基因(即 SEQ ID NO: 157 的 295-1158 核苷酸)或其互补序列的至少 30
个连续核苷酸，其中所述多肽具有 FPP 合成酶的活性；以及(e)SEQ ID NO:
159 经保守修饰而形成的变体。

10 这样，该实施方案中，氨基酸可由编码 FPP 合成酶的整个开放阅读框
架编码，即 SEQ ID NO: 159 的 1-287 残基，其至少 30 个连续的残基，或
SEQ ID NO: 159 的片段，其必须具有 FPP 合成酶活性，如实施例 1 中所描
述的试验所测定。另外，本发明的该实施方案还包括在严谨条件(如上定义)
下，能够与杂交探针杂交的多核苷酸所编码的氨基酸序列，所述探针含有
15 ispA 基因(即 SEQ ID NO: 157 的 295-1158 核苷酸)或其互补序列的至少 30
个连续核苷酸，其中所述多肽如上定义具有 FPP 合成酶的活性。

在一个优选实施方案中，该多肽具有 SEQ ID NO: 159 的氨基酸序列。

本发明的另一实施方案是分离的多肽，其具有选自下述的氨基酸序列：
(a)SEQ ID NO: 160 的 1-142 残基所示的氨基酸序列；(b)SEQ ID NO: 160 的
20 至少 30 个连续的氨基酸残基；(c)SEQ ID NO: 160 的片段的氨基酸序列，所
述片段具有 1-脱氧木糖醇-5-磷酸合成酶(DXPS)的活性；(d)在严谨条件下能
够与杂交探针杂交的多核苷酸所编码的多肽氨基酸序列，所述杂交探针含
有 SEQ ID NO: 157 或其互补序列 1185-1610 位的至少 30 个连续核苷酸，其
中所述多肽具有 DXPS 合成酶的活性；以及(e)SEQ ID NO: 160 经保守修饰
25 而形成的变体。

这样，该实施方案中，氨基酸可由编码 DXPS 合成酶的整个开放阅读
框架编码，即 SEQ ID NO: 160 的 1-142 残基，其至少 30 个连续的残基，或
SEQ ID NO: 160 的片段，其如实施例 1 中所描述的试验所测定具有 DXPS
合成酶活性。另外，本发明的该实施方案还包括在严谨条件(如上定义)下，
30 能够与杂交探针杂交的多核苷酸所编码的氨基酸序列，所述探针含有 DXPS
基因(即 SEQ ID NO: 157 的 1185-1610 核苷酸)或其互补序列的至少 30 个连

续核苷酸，其中所述多肽如上定义具有 DXPS 合成酶的活性。

在一个优选实施方案中，该多肽具有 SEQ ID NO: 160 的氨基酸序列。

本发明的另一实施方案是分离的多肽，其具有选自下述的氨基酸序列：

- (a)SEQ ID NO: 178 的 1-390 残基所示的氨基酸序列；(b)SEQ ID NO: 178 的
5 至少 30 个连续的氨基酸残基；(c)SEQ ID NO: 178 的片段的氨基酸序列，所述片段具有乙酰-CoA 乙酰基转移酶的活性；(d)在严谨条件下能够与杂交探针杂交的多核苷酸所编码的多肽氨基酸序列，所述杂交探针含有 phaA 基因(即 SEQ ID NO: 177 的 1-1179 核苷酸)或其互补序列的至少 30 个连续核苷酸，其中所述多肽具有乙酰-CoA 乙酰基转移酶的活性；以及(e)SEQ ID NO:
10 178 经保守修饰而形成的变体。

- 这样，该实施方案中，氨基酸可由编码乙酰-CoA 乙酰基转移酶整个开放阅读框架编码，即 SEQ ID NO: 178 的 1-143 残基，其至少 30 个连续的残基，或 SEQ ID NO: 178 的片段，其必须具有乙酰-CoA 乙酰基转移酶活性，如实施例 1 中所描述的试验所测定。另外，本发明的该实施方案还包括在
15 严谨条件(如上定义)下，能够与杂交探针杂交的多核苷酸所编码的氨基酸序列，所述探针含有 phaA 基因(即 SEQ ID NO: 177 的 1-1170 核苷酸)或其互补序列的至少 30 个连续核苷酸，其中所述多肽如上定义具有乙酰-CoA 乙酰基转移酶的活性。

在一个优选实施方案中，该多肽具有 SEQ ID NO: 178 的氨基酸序列。

- 20 本发明的另一实施方案是分离的多肽，其具有选自下述的氨基酸序列：
(a)SEQ ID NO: 179 的 1-240 残基所示的氨基酸序列；(b)SEQ ID NO: 179 的至少 30 个连续的氨基酸残基；(c)SEQ ID NO: 179 的片段的氨基酸序列，所述片段具有乙酰乙酰-CoA 还原酶的活性；(d)在严谨条件下能够与杂交探针杂交的多核苷酸所编码的多肽氨基酸序列，所述杂交探针含有 phaB 基因(即
25 SEQ ID NO: 177 的 1258-1980 核苷酸)或其互补序列的至少 30 个连续核苷酸，其中所述多肽具有乙酰乙酰-CoA 还原酶的活性；以及(e)SEQ ID NO: 179 经保守修饰而形成的变体。

- 这样，该实施方案中，氨基酸可由编码乙酰乙酰-CoA 还原酶整个开放阅读框架编码，即 SEQ ID NO: 179 的 1-240 残基，其至少 30 个连续的残基，或 SEQ ID NO: 179 的片段，该片段具有乙酰乙酰-CoA 还原酶活性，如
30 实施例 1 中所描述的试验所测定。另外，本发明的该实施例还包括在严谨

条件(如上定义)下, 能够与杂交探针杂交的多核苷酸所编码的氨基酸序列, 所述探针含有 phaB 基因(即 SEQ ID NO: 177 的 1258-1980 核苷酸)或其互补序列的至少 30 个连续核苷酸, 其中所述多肽如上定义具有乙酰乙酰-CoA 还原酶的活性。

5 在一个优选实施方案中, 该多肽具有 SEQ ID NO: 179 的氨基酸序列。

本文所用的术语“多肽”, “多肽序列”, “氨基酸”和“氨基酸序列”可互换使用, 指寡肽, 肽, 多肽, 或蛋白序列, 或它们的片段, 以及天然存在或合成的分子。在本文中, “片段”, “免疫原性片段”, 或“抗原性片段”指本文所定义的任一多肽的片段, 该片段长度至少约为 30 个氨基酸, 10 并具有所述多肽的某些生物活性或免疫活性。当用“氨基酸序列”描述天然蛋白分子的氨基酸序列时, “氨基酸序列”及其类似的术语并不是将氨基酸序列限定为与所述的蛋白分子相关的完全天然的氨基酸序列。

当用于描述多肽时, 术语“分离的”指蛋白或多肽已经从天然状态下与其混合(accompany)的各个组分中分离出来。当样品的至少约 60-75%呈现 15 为单一多肽序列时, 即分离得到了单体蛋白。分离的蛋白通常含有约 60-90%W/W 的蛋白样品, 更通常约为 95%, 优先高于约 99%纯度。蛋白纯度或均一性可利用本领域公知的许多方式表示, 如蛋白样品的聚丙烯酰胺凝胶电泳, 再通过染色凝胶观测单一多肽条带。为满足某些目的, 可使用 HPLC 或其它本领域公知的方法提供更高的纯化分辨率。

20 本文中术语“生物活性”指具有天然存在分子的结构, 调节性或生化功能的蛋白。同样, “免疫活性”指天然, 重组或合成多肽或其任一寡肽的性能, 该性能可在适当的动物或细胞内诱导特定免疫应答, 并结合特定抗体。

本发明的另一实施方案是分离的多核苷酸序列, 其具有甲羟戊酸操纵子 25 (SEQ ID NO: 42)的核苷酸序列, SEQ ID NO: 42 各变体的核苷酸序列, 所述变体含有一个或多个符合副球菌种(*Paracoccus* sp.)菌株 1534 密码子利用表(参见表 14)的取代, 或 SEQ ID NO: 42 的片段。SEQ ID NO: 42 的变体和片段必须编码具有选自下述活性的多肽: HMG-CoA 还原酶, 异戊烯二磷酸异构酶活性, 羟基甲基戊二酸单酰-CoA 合成酶(HMG-CoA 合成酶), 甲羟戊 30 酸激酶, 磷酸甲羟戊酸激酶和二磷酸甲羟戊酸脱羧酶。该实施方案还包括在如上定义的严谨条件下能够与杂交探针杂交的多核苷酸序列, 所述探针

的核苷酸序列由 SEQ ID NO: 42 的大约 10 到约 9066 个核苷酸构成, 优选由 SEQ ID NO: 42 或其互补序列的 30 个连续核苷酸构成, 该多核苷酸编码具有选自下述活性的多肽: HMG-CoA 还原酶, 异戊烯二磷酸异构酶, HMG-CoA 合成酶, 甲羟戊酸激酶, 磷酸甲羟戊酸激酶和二磷酸甲羟戊酸脱羧酶。

该实施方案还包括分离的多核苷酸序列, 其含有 SEQ ID NO: 42 的所述残基: 2622-3644, 3641-4690, 4687-5853, 5834-6970, 6970-7887, 7880-8878。这些序列的片段也包括在本发明的范围之内, 只要是这些片段能够分别编码具有 HMG-CoA 还原酶活性, 异戊烯二磷酸异构酶活性, HMG-CoA 合成酶活性, 甲羟戊酸激酶活性, 磷酸甲羟戊酸激酶活性和二磷酸甲羟戊酸脱羧酶活性的多肽。

该实施方案还包括在如上定义的严谨条件下, 能够与杂交探针杂交的多核苷酸序列, 所述探针选自下述由 SEQ ID NO: 42: 2622-3644, 3641-4690, 4687-5853, 5834-6970, 6970-7887, 7880-8878 或其互补序列的至少 30 个连续核苷酸残基所构成的核苷酸序列, 其中多核苷酸分别编码选自具有 HMG-CoA 还原酶活性, 异戊烯二磷酸异构酶活性, HMG-CoA 合成酶活性, 甲羟戊酸激酶活性, 磷酸甲羟戊酸激酶活性或二磷酸甲羟戊酸脱羧酶活性的多肽。

优选地, 该分离的多核苷酸由 SEQ ID NO: 42 的 2622-3644, 3641-4690, 4687-5853, 5834-6970, 6970-7887, 7880-8878 核苷酸构成。

本发明的另一实施方案是分离的多核苷酸序列, 其具有 SEQ ID NO: 157 的核苷酸序列, SEQ ID NO: 157 各变体的核苷酸序列, 所述变体含有一个或多个符合副球菌种菌株 1534 密码子利用表(参见表 14)的取代, 或 SEQ ID NO: 157 的片段, SEQ ID NO: 157 的变体和片段均编码具有下述活性的多肽: FPP 合成酶活性, 1-脱氧-D-木酮糖-5-磷酸合成酶活性或 XseB 活性。该实施方案还包括在如上定义的严谨条件下能够与杂交探针杂交的多核苷酸序列, 所述探针的核苷酸序列由 SEQ ID NO: 157 或 SEQ ID NO: 157 的互补序列的至少 30 个连续核苷酸构成, 其中该多核苷酸编码具有下述活性的多肽: FPP 合成酶活性, 1-脱氧-D-木酮糖-5-磷酸合成酶活性或 XseB 活性。

优选地, 该分离的多核苷酸由 SEQ ID NO: 157 的 59-292, 295-1158 或 1185-1610 位核苷酸构成。

还提供了具有选自下述核苷酸序列的分离多核苷酸序列：SEQ ID NO: 157 的 59-292 位核苷酸，SEQ ID NO: 157 各变体的核苷酸序列，所述变体含有一个或多个符合副球菌种菌株 R1534 密码子利用表(参见表 14)的取代，SEQ ID NO: 157 的 59-292 位核苷酸序列的片段，该片段编码具有 XseB 功
5 能的多肽，以及在严谨条件下能够与杂交探针杂交的多核苷酸序列，所述探针的核苷酸序列由 SEQ ID NO: 157 或其互补序列的 59-292 位的至少 30 个连续核苷酸构成，其中该多核苷酸编码具有 XseB 功能的多肽。

优选地，该分离的多核苷酸由 SEQ ID NO: 157 的 59-292 位核苷酸构成。

还提供了具有选自下述核苷酸序列的分离多核苷酸序列：SEQ ID NO: 157 的 295-1158 位核苷酸，SEQ ID NO: 157 的 295-1158 位核苷酸序列的各
10 变体，所述变体含有一个或多个符合副球菌种菌株 R1534 密码子利用表(参见表 14)的取代，SEQ ID NO: 157 的 295-1158 位核苷酸序列的片段，该片段编码 FPP 合成酶活性，以及在严谨条件下能够与杂交探针杂交的多核苷酸序列，所述探针的核苷酸序列由 SEQ ID NO: 157 或其互补序列的
15 295-1158 位的至少 30 个连续核苷酸构成，其中该多核苷酸编码具有 FPP 合成酶活性的多肽。

优选地，该分离的多核苷酸由 SEQ ID NO: 157 的 295-1158 位核苷酸构成。

本发明的另一实施方案是具有 SEQ ID NO: 157 的 1185-1610 位核苷酸
20 序列的分离多核苷酸序列，SEQ ID NO: 157 的 1185-1610 位核苷酸序列的各变体，所述变体含有一个或多个符合副球菌种菌株 1534 密码子利用表(参见表 14)的取代，或 SEQ ID NO: 157 的 1185-1610 位核苷酸序列的片段，该片段编码具有 1-脱氧木酮糖-5-磷酸合成酶活性的多肽。该实施方案还包括在
25 严谨条件下能够与杂交探针杂交的多核苷酸序列，所述探针的核苷酸序列由 SEQ ID NO: 157 或其互补序列的 1185-1610 位的至少 30 个连续核苷酸构成，其中该多核苷酸编码具有 1-脱氧木酮糖-5-磷酸合成酶活性的多肽。

优选地，该分离的多核苷酸由 SEQ ID NO: 157 的 1185-1610 位核苷酸构成。

本发明的另一实施方案是具有 SEQ ID NO: 177 的核苷酸序列的分离多
30 核苷酸序列，SEQ ID NO: 177 核苷酸序列的各变体，所述变体含有一个或多个符合副球菌种菌株 1534 密码子利用表(参见表 14)的取代，或 SEQ ID

NO: 177 核苷酸序列的片段, 该片段编码具有选自下述活性的多肽: 乙酰-CoA 乙酰基转移酶和乙酰乙酰-CoA 还原酶。该实施方案还包括在如上定义的严谨条件下能够与杂交探针杂交的多核苷酸序列, 所述探针的核苷酸序列由 SEQ ID NO: 177 或其互补序列的至少 30 个连续核苷酸构成, 其中该多核苷酸编码具有选自乙酰-CoA 乙酰基转移酶或乙酰乙酰-CoA 还原酶活性的多肽。

该实施方案中, 分离的多核苷酸序列还可包括 SEQ ID NO: 177 的 1-1170 位核苷酸序列, SEQ ID NO: 177 核苷酸序列的各变体, 所述变体含有一个或多个符合副球菌种菌株 1534 密码子利用表(参见表 14)的取代, 或 SEQ ID NO: 177 核苷酸序列的片段, 该片段编码具有乙酰-CoA 乙酰基转移酶活性的多肽。该实施方案还包括在严谨条件下能够与杂交探针杂交的多核苷酸序列, 所述探针的核苷酸序列由 SEQ ID NO: 177 或其互补序列的 1-1170 位的至少 30 个连续核苷酸构成, 其中该多核苷酸编码具有乙酰-CoA 乙酰基转移酶活性的多肽。

优选地, 该分离的多核苷酸序列由 SEQ ID NO: 177 的 1-1170 位核苷酸构成。

该实施方案中, 分离的多核苷酸序列还可以选择地是 SEQ ID NO: 177 的 1258-1980 位核苷酸序列, SEQ ID NO: 177 核苷酸序列的各变体, 所述变体含有一个或多个符合副球菌种菌株 1534 密码子利用表(参见表 14)的取代, 或 SEQ ID NO: 177 核苷酸序列的片段, 该片段编码具有乙酰乙酰-CoA 还原酶活性的多肽。该实施方案还包括在严谨条件下能够与杂交探针杂交的多核苷酸序列, 所述探针的核苷酸序列由 SEQ ID NO: 177 或其互补序列的 1258-1980 位的至少 30 个连续核苷酸构成, 其中该多核苷酸编码具有乙酰乙酰-CoA 还原酶活性的多肽。

优选地, 该分离的多核苷酸序列由 SEQ ID NO: 177 的 1258-1980 位核苷酸构成。

在另一优选实施方案中, 分离的多核苷酸序列具有选自 SEQ ID NO: 42, SEQ ID NO: 157, SEQ ID NO: 177 或其组合的核苷酸序列。本文的词组“或其组合”, 当用于指核苷酸序列时, 是指可对所提及序列的各种组合进行结合以形成该分离的多核苷酸序列。而且, 在本发明中, 可使用相同序列的多个拷贝, 即多连体(concatamers)。同样, 如同以下将要详细论述的, 含有

相同多核苷酸序列的质粒多拷贝可被转移到适当的宿主细胞中。

在本文中，“分离的”多核苷酸(如 RNA, DNA 或混合的多聚物)是实质上与其它细胞组分分离开的多核苷酸，而这些组分如核糖体，聚合酶，许多其它基因组序列和蛋白，在天然条件下是与天然序列或多肽相互混合的。

- 5 该术语包括从天然存在的环境中被分离出来的多核苷酸，还包括重组或克隆的 DNA 分离物，以及化学合成的类似物或由异源系统生物合成的类似物。

术语“核酸序列”指从 5'到 3'末端方向阅读的、单链或双链的脱氧核糖核苷酸或核糖核苷酸碱基多聚物。其包括染色体 DNA，自我复制的质粒，感染性 DNA 或 RNA 多聚体，以及具有一级结构功能的 DNA 或 RNA。

- 10 “表达调控序列”指核苷酸调控序列的排列，其指导着可操作性连接的核酸的转录。这种表达调控序列的实例为“启动子”。启动子包括在转录起始位点附近的必要的核酸序列。启动子还可任选地含有末端(distal)增强子或阻遏元件，增强子或阻遏元件可定位在转录起始位点的几千个碱基对之外。“组成型”启动子是指在多数环境条件和发育条件下都能够激活的启动子。而“诱导型”启动子是指在环境因素或发育因素的调节下被激活的启动子。
- 15 术语“可操纵性连接/可操作相连”指在核酸表达调控序列(如启动子或转录因子结合位点的排列)和第二个核酸序列之间的功能性连接，其中表达调控序列指导着相应的第二个序列核酸的转录。

- 如果一段多核苷酸序列来自异源物种，或者虽然来自相同物种，但已被修饰成与原来不同的形式，那么这段多核苷酸则与某种有机体或第二个多核苷酸序列是“异源的”。例如，启动子与异源编码序列可操纵连接是指编码序列的来源物种不同于启动子的来源物种，或者，来自相同物种，但是编码序列已被修饰成与天然存在的等位变体不同的形式。
- 20

- 当同时发生转基因的表达和外源基因的抑制(如通过反义作用，或有意
- 25 义抑制(sense suppression))时，本领域技术人员将会意识到插入的多核苷酸序列不必要完全相同于其所来源于的基因序列，但是必须是“基本上相同的”。

- 当插入的多核苷酸序列被转录以及翻译以生成功能性多肽时，本领域技术人员会认识到由于密码子简并性，大量的多核苷酸序列都可以编码相同的多肽。这些变体都具体包括在本发明的范围之内。另外，本发明还具
- 30 体包括彼此基本上相同(如下所述进行测定)的序列，并且其编码的多肽是野

生型多肽的突变体,或仍维持多肽的功能(如,该多肽氨基酸的保守取代而形成的)。另外,变体可以编码如下所述的显性阴性突变(negative mutant)。

按如下所述,在两个核酸序列或两个多肽序列是按照最大一致性排列时,如果两序列中核苷酸序列或氨基酸残基是分别相同的,可以说这两个核酸序列或这两个多肽是“相同的”。术语“相同”或“同一性”百分比,当用于描述两个或多个核酸或多肽序列时,指当在比较窗口下对比并排列其最大一致性时(按照下述序列对比算法之一进行测定,或通过人工排列和目测进行测定),两个或多个序列或亚序列是相同的,或者具有特定百分比的相同氨基酸残基或核苷酸。当序列同一性的百分比用于蛋白或多肽时,就认为不相同的残基位置通常因保守氨基酸取代而不同,其中氨基酸残基是被其它类似化学属性(如电荷或疏水性)的氨基酸取代,从而不改变该分子的功能。当序列以保守取代的形式而各不相同,序列同一性百分比可被上调,以校正取代的保守性性质。进行这种调整的方法是本领域技术人员所公知的。通常包括以部分而不是全部错配的方式对保守取代打分(scoring),从而提高序列同一性的百分比。这样,例如,当评定相同的氨基酸为1分时,非保守取代的得分为0,而保守取代的得分在0到1之间。保守取代的得分可根据例如 Meyers & Miller, Computer Applic. Biol. Sci. 4:11-17 (1988)的算法进行计算,如同在 PC/GENE(Intelligenetics, Mountain View, Calif., USA)程序中所执行的。

短语“实质上相同(substantially identical)”,当用于形容两个核酸或多肽时,指当在比较窗口下对比并排列最大一致性时(按照下述序列对比算法之一进行测定,或通过人工排列和目测进行测定),序列间或亚序列间具有至少60%、优选80%,最优选90-95%的核苷酸或氨基酸残基同一性。该定义还指一种序列,该序列的互补序列能够与检测序列杂交。

进行序列对比时,通常一个序列作为对照序列,将检测序列与其进行对比。当使用序列对比算法时,检测 and 对照序列都被输入计算机,如果必要的话,还可指定各个并列(coordinate)的亚序列,并制定序列算法程序的各个参数。可使用默认的程序参数,或指定可选参数。然后,序列对比算法可在程序参数基础上,计算检测序列相对于对照序列的序列同一性百分比。

本文所用的“对比窗口”,包括选自以下任何数目的连续位置片段参照体系:20-600,通常约50-约200,更通常约100-约150,其中序列可以与

相同数目的连续位置的对照序列在这两个序列被最优化排列之后，进行对比。排列序列用于对比的方法是本领域公知的。用于对比的序列优化排列可通过：如 Smith 和 Waterman, *Adv. Appl. Math.* 2: 482(1981)的局部同源性算法进行，如 Needleman 和 Wunsch, *J. Nat'l, Acad. Sci.USA* 85: 2444(1988)的同源性排列算法，这些算法(Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, Wis.中的 GAP, BESTFIT, FASTA, 和 TFASTA)的计算机化运行，或人工排列和目测(visualinspection)。

一种有用的算法实例是 PILEUP。PILEUP 从一组相关的序列中，使用渐进性成对排列的方法，产生出多个序列排列，以显示其相关性和序列同一性百分比。它还绘出树形图或树状图(dendogram)，以标明用于形成排列的聚类(clustering)关系。PILEUP 使用简化的 Feng 和 Doolittle, *J.Mol.Evol.* 35:351-360 (1987)渐进性排列方法。所用的该方法类似于 Higgins 和 Sharp, *CABIOS* 5:151-153 (1989)所描述的方法。该程序可以序列对比(align)达 300 个序列，每个序列的最大长度为 5000 个核苷酸或氨基酸。多元排列方法是以两个最近似序列的配对排列开始，形成了两个排列序列的组(cluster)。然后该组与下一个最相关的序列或排列序列的组相排列。两组序列通过两个单个序列的成对排列的简单延伸而排列。一系列渐进性配对排列之后就获得了最后的排列。通过指定用于序列比较区域的具体序列和其氨基酸或核苷酸序列匹配物(coordinate)，并指定程序的参数，就可运行该程序。例如，对照序列可与其它检测序列相比较，以测定序列同一性关系的百分比，使用以下参数：默认的 gap weight(3.00)，默认的 gap length weight(0.10)和 weighted end gaps。

另一种适用于测定序列同一性百分比以及序列相似性的算法的实例为 BLAST 算法[Altschul 等, *J.Mol.Biol.* 215:403-410 (1990)]。进行 BLAST 分析的软件可通过 National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>)获得。该算法涉及：首先通过在查询序列中鉴定出 W 长度的短序列(words)，当该序列与数据库序列中相同长度的序列对比时，能够匹配或符合某个以位置进行衡量(positive-valued)的阈值 T，以此鉴定出高分的序列对(HSP)。T 指临近序列(neighborhood word)的分数阈值(Altschul 等，同上)。这些起始性临近序列点作为种子(seeds)用于起始检索，以找到将之包含其中的更长的 HSP。序列点(hit)在每个序列的两个方向上均

延伸,以使累积的排列分数能够增加。当下述情况下,停止序列点在两个方向上的延伸:累积的排列分数从其获得的最大值处下降了 X 数量;由于一个或多个负分数残基加入排列后逐渐积累,使得累积的分数变为 0 或负数,或者任一序列已到达末尾。该 BLAST 算法参数 W, T 和 X 决定着排列的灵敏度和速度。使用 BLAST 程序的默认参数:序列长度(word length)(W)为 11, BLOSUM62 分数基质(参见 Henikoff 和 Henikoff, Proc.Natl.Acad.Sci USA 89: 10915 (1989))排列(B)为 50, 预期值(E)为 10, M=5, N=-4, 以及两条链的对比。

BLAST 算法还可对两条序列之间的相似性进行统计学分析[例如,参见 Karlin 和 Altschul, Proc.Nat'l.Acad.Sci. USA 90:5873-5787 (1993)]。BLAST 算法提供的相似性测定方法是概率的最小和(P(N)),其表明两个核苷酸或氨基酸序列之间偶然发生匹配的机率。例如,如果检测核酸与对照核酸之间的对比中,最小总和概率(smallest sum probability)小于约 0.2,更有选小于约 0.01,最优选小于约 0.001,则认为该核酸与对照核酸相似。

“保守修饰的变体”可用于描述氨基酸或核酸序列。当用于特定核酸序列时,保守修饰的变体指编码相同或基本相同氨基酸序列的核酸,或不编码氨基酸序列的核酸,基本相同的序列。由于遗传密码的简并性,许多功能相同的核酸密码子可编码任何指定的蛋白。例如,密码子 GCA, GCC, GCG 和 GCU 都编码丙氨酸。这样,在每个丙氨酸被某个密码子具体指定的各位置,该密码子都可被改变为任何相应所述的密码子,而不会改变被编码的多肽。这样的核酸变体是“沉默变体”(silent variations),其作为一种保守修饰的变体。本文中编码多肽的每个核酸序列还描述该核酸的各个可能沉默变体。本领域技术人员公知:核酸中的每个密码子(除了 AUG,其通常为蛋氨酸的唯一密码子)均可被修饰,以形成功能相同的分子。同样,编码多肽的核酸的每个沉默突变均已隐含在每个被描述的序列中。

当描述氨基酸序列时,本领域技术人员公知:核酸中单个的取代,缺失或插入,或肽,多肽或蛋白序列的取代,其改变了编码序列中的单个氨基酸或小百分率氨基酸(即,小于 20%,如 15%, 10%, 5%, 4%, 3%, 2% 或 1%),当这种改变形成了类似化学属性氨基酸的氨基酸取代时,则认为该核酸序列为“保守取代变体”。提供功能类似氨基酸的保守取代表是本领域公知的。

下述六组均含有彼此为保守取代的氨基酸:

丙氨酸(A), 丝氨酸(S), 苏氨酸(T);

天冬氨酸(D), 谷氨酸(E);

天冬酰胺(N), 谷氨酰胺(Q);

5 精氨酸(R), 赖氨酸(K);

异亮氨酸(I), 亮氨酸(L), 蛋氨酸(M), 缬氨酸(V); 和

苯丙氨酸(F), 酪氨酸(Y), 色氨酸(W)。(参见, 如 Creighton, *Proteins* (1984))。

指示两个核酸序列或多肽基本上相同的标识为: 第一个核酸编码的多肽, 与抗第二个核酸所编码多肽的抗体之间具有免疫交叉反应。这样, 例如, 当一个多肽与第二个多肽之间的区别仅仅在于保守取代时, 这两个多肽通常实质上(substantially)相同。指示两个核酸序列实质上相同的标识为: 这两个分子或其互补序列彼此之间能够在如下所述的严谨条件下杂交。

15 短语“与……特异性杂交”指当核酸序列存在于复杂混合物(如总细胞或 DNA 或 RNA 文库)中时, 在严谨杂交条件下, 某分子仅仅与该特定的核酸序列结合, 配对(duplexing)或杂交。

短语“严谨杂交条件”指在该条件下, 探针能够与其靶序列而非其它序列杂交, 所述靶序列通常存在于核酸序列的复杂混合物中。严谨杂交条件是序列依赖性的, 并在不同的环境条件下有所不同。长序列在较高的温度下才会特异杂交。核酸杂交的详细指南可参见 Tijssen, *Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Probes, “Overviews of Principles of Hybridization and the Strategy of Nucleic Acid Assays”* (1993)。通常, 高度严谨条件选择为: 在指定的离子强度和 pH 下, 比特定序列的热融点(T_m)低约 5-10°C。低严谨条件通常选择为 T_m 下约 15-30°C。 T_m 是指这样的温度: 在该温度下(在指定离子强度, pH, 和核酸浓度下), 有 50%互补于靶序列的探针以平衡状态与该靶序列杂交(当靶序列过量存在时, 在 T_m 下, 50%的探针被结合, 该状态维持平衡)。严谨条件指这样的条件: 在 pH7.0-8.3 下, 盐浓度低于约 1.0M 钠离子, 通常约 0.01-1.0M 钠离子浓度(或其它盐), 对于短探针(如 10-50 个核苷酸)而言, 温度至少约为 30°C, 而长探针(如大于 50 个核苷酸)至少约为 60°C。严谨条件还可加入去稳定剂如甲酰胺而获得。为获得选择性或特定杂交, 阳性信

号应至少是杂交背景的 2 倍, 优选 10 倍。

如果它们所编码的肽实质上是相同的, 在严谨条件下彼此并不杂交的核酸实质上仍然是相同的。例如当使用该遗传密码的最大密码子简并性形成核酸拷贝时, 这种情况会发生。在这种情形下, 核酸通常在中等程度的
5 严谨条件下杂交。

在本发明中, 含有本发明核酸的基因组 DNA 或 cDNA 可使用本发明公开的核酸序列, 在严谨条件下, 用标准 Southern 印迹鉴定。为此目的, 适于这种杂交的严谨条件包括在 37℃, 在 40% 甲酰胺, 1M NaCl, 1% 十二烷基硫酸钠(SDS)的缓冲液中杂交, 并在至少约 50℃, 通常为约 55℃-约 60℃,
10 用 0.2X SSC 至少进行 20 分钟的洗涤一次, 或等价条件进行。阳性杂交至少是背景的 2 倍。本领域普通技术人员公知其它可选的杂交条件和洗涤条件也可用于提供类似的严格度条件。

进一步指示两多核苷酸实质上相同的标识为: 如果用一对寡核苷酸引物扩增对照序列, 该对照序列在严谨杂交条件下用作探针, 用于从 cDNA
15 或基因组文库中分离检测序列, 或鉴定检测序列, 如在 northern 或 Southern 印迹中。

本发明还包括如上定义的表达载体。该表达载体包括一个或多个上述每种多核苷酸序列的拷贝。本发明的表达载体可含有本文所定义的任一多核苷酸序列, 如 SEQ ID NO: 42, 或 SEQ ID NO: 42 的下述残基: 2622-3644,
20 3641-4690, 4687-5853, 5834-6970, 6970-7887, 7880-8878, 以及 SEQ ID NO: 157 的残基 59-292, 295-1158 或 1185-1610, SEQ ID NO: 177 的残基: 1-1170 或 1258-1980。该表达载体可含有本文鉴定的多核苷酸序列的各组合, 如, SEQ ID NO: 42, SEQ ID NO: 157 和 SEQ ID NO: 177。

该表达载体中的多核苷酸还可如上定义, 任选地与表达调控序列可操作
25 作相连并如实施例中进行扩增。

本发明还包括以下表达载体: pBBR-K-mev-op16-1, pBBR-K-mev-op16-2, pDS-mvaA, pDS-idi, pDS-hcs, pDS-mvk, pDS-pmk, pDS-mvd, pDS-His-mvaA, pDS-His-idi, pDS-His-hcs, pDS-His-mvk, pDS-His-pmk, pDS-His-mvd, pBBR-K-Zea4, pBBR-K-Zea4-up,
30 pBBR-K-Zea4-down, pBBR-K-PcrtE-crtE-3, pBBR-tK-PcrtE-mvaA, pBBR-tK-PcrtE-idi, pBBR-tK-PcrtE-hcs, pBBR-tK-PcrtE-mvk,

pBBR-tK-PcrtE-pmk, pBBR-tK-PcrtE-mvd, pBBR-K-PcrtE-mvaA-crtE-3,
pDS-His-phaA, pBBR-K-PcrtE-crtW, pBBR-K-PcrtE-crtWZ,
pBBR-K-PcrtE-crtZW 及其组合。这些表达载体在以下的实施例中定义更为
详细。而且, 本发明还包括含有本文所定义任一序列的表达载体, 该表达
5 载体被用于在适当的宿主细胞内表达类异戊二烯化合物, 如类胡萝卜素,
优选玉米黄质。本文中, 短语“表达载体”是携带编码本文所述多核苷酸
序列的 DNA 序列, 并能够介导该序列表达的可复制载体。

在本文中, 术语“可复制”指在该载体已经导入的、指定类型的宿主
细胞中, 该载体能够复制。目的多核苷酸序列的直接上游可以提供一段编
10 码信号肽的序列, 该肽的存在可以确保携带有载体的宿主细胞表达该编码
多肽的分泌。该信号序列可以是与被选择多核苷酸序列天然相关连, 或来
自另外的来源。

该载体可以是任何可方便地进行重组 DNA 操作的载体, 载体的选择通
常依赖于该载体将要导入的宿主细胞。这样, 该载体可以是自主复制载体,
15 即, 作为染色体外实体存在的载体, 其复制独立于染色体的复制; 这种载
体的例子可包括质粒, 噬菌体, 粘粒或极微染色体。或者, 该载体可以是
这样的: 当其被导入宿主细胞时, 其可以整合到宿主染色体组中, 并随着
该被整合的染色体一起复制。适当的载体如实施例所示。本发明的表达载
体可以携带以下定义的本发明任一 DNA 序列, 以及用于表达本发明以下定
20 义的任何多肽。

本发明还包括含有本文所述的一个或多个多核苷酸序列和/或一个或多
个表达载体的培养细胞。本文中, “培养细胞”包括任何能够在指定条件下
生长并表达一个或多个本发明多核苷酸编码多肽的细胞。优选地, 该培养
细胞为酵母, 真菌, 细菌或藻类。更优选地, 该培养细胞为副球菌, 黄杆菌
25 属(Flavobacterium), 土壤杆菌属(Agrobacterium), 产碱杆菌属(Alcaligenes),
欧文氏杆菌属(Erwinia), 大肠杆菌(E.coli)或枯草杆菌(B.subtilis)。甚至更优
选地, 该细胞为副球菌, 如 R-1506, R-1512, R1534 或 R114。本发明还包括
本文所鉴定的、能够表达本文所公开多肽的任何细胞的后代。在本发明中,
如果使用实施例 2 所述条件, 某个细胞的 AFLP DNA 指纹与假定亲代细胞
30 的无区别, 则认为该细胞即为其后代。

这样, 本发明的培养细胞可包括, 例如 SEQ ID NO: 42 或 SEQ ID NO: 42

的下述残基：2622-3644，3641-4690，4687-5853，5834-6970，6970-7887，7880-8878，以及 SEQ ID NO: 157 的残基 59-292，295-1158 或 1185-1610，SEQ ID NO: 177 的残基：1-1170 或 1258-1980。这些序列可以单独地，或作为其它表达载体的一部分，被转移到细胞中。这些序列还可任选地与表达
5 调控序列可操作相连。该培养细胞还可含有本文所鉴定多核苷酸序列的各个组合，如 SEQ ID NO: 42, SEQ ID NO: 157 和 SEQ ID NO: 177。

本发明的培养细胞还可进一步含有编码一个或多个类胡萝卜素生物合成途径酶的多核苷酸。例如，本发明的培养细胞可含有一个或多个 SEQ ID NO: 180, 182 和 184 的拷贝，这些拷贝可以单独或与本文所定义的任一多核
10 苷酸序列一起包含在本发明培养细胞中。这样，本发明公开的多核苷酸序列可以单独地，或与其它能够提高类异戊二烯化合物靶物质(如象玉米黄质或虾青素的类胡萝卜素)产量的另一多核苷酸序列一起，被转移到培养细胞中。在这方面，本发明包括任何编码如类胡萝卜素生物合成相关多肽的多核苷酸的应用，所述多肽如 GGPP 合成酶， β -胡萝卜素- β 4-加氧酶(酮酶)，
15 和/或 β -胡萝卜素羟化酶。另外，还可将编码类胡萝卜素生物合成相关多肽的多核苷酸组合与一个或多个本文鉴定的多核苷酸在相同或不同的表达载体中联合使用。这种构建体可被转移到本发明的培养细胞中以形成能够表达目的类异戊二烯物质的细胞。

例如，本发明的培养细胞可含有一个或多个下述表达载体：
20 pBBR-K-mev-op16-1, pBBR-K-mev-op16-2, pDS-mvaA, pDS-idi, pDS-hcs, pDS-mvk, pDS-pmk, pDS-mvd, pDS-His-mvaA, pDS-His-idi, pDS-His-hcs, pDS-His-mvk, pDS-His-pmk, pDS-His-mvd, pBBR-K-Zea4, pBBR-K-Zea4-up, pBBR-K-Zea4-down, pBBR-K-PcrtE-crtE-3, pBBR-tK-PcrtE-mvaA, pBBR-tK-PcrtE-idi, pBBR-tK-PcrtE-hcs, pBBR-tK-PcrtE-mvk,
25 pBBR-tK-PcrtE-pmk, pBBR-tK-PcrtE-mvd, pBBR-K-PcrtE-mvaA-crtE-3, pDS-His-phaA, pBBR-K-PcrtE-crtW, pBBR-K-PcrtE-crtWZ, pBBR-K-PcrtE-crtZW 及其组合。

本发明的另一实施方案为生产类胡萝卜素的方法。该方法中，在能够表达如上定义多核苷酸序列所编码的多肽的条件下，培养如上定义的培养
30 细胞。能够表达多肽的培养条件如下述实施例所述，但如需要可进行修改，以适应特定目的需求。然后从细胞或者如果分泌，从细胞培养基中分离类

胡萝卜素。

本发明中，“类胡萝卜素”包括下述化合物：八氢番茄红素，番茄红素， β -胡萝卜素，玉米黄质，角黄素，虾青素，福寿草黄素(adonixanthin)，隐黄素，海胆酮和福寿草红素(adonirubin)及其组合。优选的类胡萝卜素为玉米黄质。

本发明的另一实施方案为制备产类胡萝卜素细胞的方法。该方法包括(a)将编码甲羟戊酸途径酶的多核苷酸序列导入细胞，所述酶在该细胞中能够表达；以及(b)选择出含有步骤(a)多核苷酸序列的细胞，并且该细胞所产生类胡萝卜素的水平为该细胞导入多核苷酸序列之前的 1.1-1000 倍。

本文中，短语“甲羟戊酸途径酶”指 IPP 生物合成的甲羟戊酸途径中相关的各种酶，并由 atoB 或 phaA, hcs, mvaA, mvk, pmk 和 mvd 基因编码。就本发明的目的而言，如果使用实施例 1 所示的任一活性试验能够检测到某种酶，则该酶“在细胞内表达”。检测类胡萝卜素产量的各方法是本领域公知的。实施例 1, 11 和 12 分别提供了鉴定玉米黄质，番茄红素和虾青素的存在常规试验方法。类似地，其它类胡萝卜素检测试验可用于检测细胞或培养基中如八氢番茄红素，角黄素，福寿草黄素，隐黄素，海胆酮和福寿草红素的存在。

这样，该方法可用于下述示范性类胡萝卜素的测定：八氢番茄红素，番茄红素， β -胡萝卜素，玉米黄质，角黄素，虾青素，福寿草黄素(adonixanthin)，隐黄素，海胆酮和福寿草红素(adonirubin)及其组合。该方法中，优选的类胡萝卜素为玉米黄质。

该方法包括制备一种细胞，该细胞生成类胡萝卜素的水平约为该细胞导入多核苷酸序列之前时的 1.1-1000 倍，优选约 1.5-500 倍，例如约 100 倍，或至少 10 倍。

该方法中，细胞生成类胡萝卜素的量约为 1mg/L-10g/L。优选细胞的类胡萝卜素产量为 100mg/L-约 9g/L，例如，约 500mg/L-8g/L，或约 1g/L-5g/L。

该方法中，细胞可选自酵母，真菌，细菌或藻类。优选地，该细胞为副球菌(Paracoccus)，黄杆菌属(Flavobacterium)，土壤杆菌属(Agrobacterium)，产碱杆菌属(Alcaligenes)，欧文氏杆菌属(Erwinia)，大肠杆菌(E.coli)或枯草杆菌(B.subtilis)。更优选地，该细胞为副球菌。

该方法中，细胞可以是突变细胞。本文中“突变细胞”是指任何含有

非天然多核苷酸序列、或含有已被改变(如通过重排或缺失或取代 1-100 个, 优选 20-50 个, 更优选低于 10 个核苷酸)为不同于天然形式的多核苷酸序列的细胞。这种非天然序列可通过随机诱变, 化学诱变, UV 辐射等等获得。优选地, 该突变导致一种或多种甲羟戊酸途径的基因表达量提高, 从而提高类胡萝卜素如玉米黄质的产量。制备, 筛选和鉴定这种突变细胞的方法是本领域公知的, 并如以下的实施例所示。这种突变体的例子为 R114 或 R1534。优选地, 该突变细胞为 R114。

该方法中, 多核苷酸序列为 SEQ ID NO: 42 或 SEQ ID NO: 42 的下述残基: 2622-3644, 3641-4690, 4687-5853, 5834-6970, 6970-7887, 7880-8878, 以及 SEQ ID NO: 157 的残基 59-292, 295-1158 或 1185-1610, SEQ ID NO: 177 的残基: 1-1170 或 1258-1980。这些序列可以单独地, 或作为其它表达载体的一部分使用。这些序列还可任选地与表达调控序列可操作相连。在该方法中, 还可以使用本文所鉴定多核苷酸序列的各个组合, 如 SEQ ID NO: 42, SEQ ID NO: 157 和 SEQ ID NO: 177。

可被选择用于该方法的表达载体包括: pBBR-K-mev-op16-1, pBBR-K-mev-op16-2, pDS-mvaA, pDS-idi, pDS-hcs, pDS-mvk, pDS-pmk, pDS-mvd, pDS-His-mvaA, pDS-His-idi, pDS-His-hcs, pDS-His-mvk, pDS-His-pmk, pDS-His-mvd, pBBR-K-Zea4, pBBR-K-Zea4-up, pBBR-K-Zea4-down, pBBR-K-PcrtE-crtE-3, pBBR-tK-PcrtE-mvaA, pBBR-tK-PcrtE-idi, pBBR-tK-PcrtE-hcs, pBBR-tK-PcrtE-mvk, pBBR-tK-PcrtE-pmk, pBBR-tK-PcrtE-mvd, pBBR-K-PcrtE-mvaA-crtE-3, pDS-His-phaA, pBBR-K-PcrtE-crtW, pBBR-K-PcrtE-crtWZ, pBBR-K-PcrtE-crtZW 及其组合。

该方法中, 可使用任何常规方法将多核苷酸序列导入细胞。将多核苷酸序列导入细胞的适当方法包括转化, 转导, 转染, 脂转染(lipofection), 电穿孔[参见如 Shigekawa 和 Dower, Biotechniques 6:742-751(1988)], 接合[参见如 Koehler 和 Thorne, Journal of Bacteriology 169:5771-5278 (1987)]和基因枪法(biolistics)。

接合在转移多核苷酸序列中的应用, 例如以表达载体的形式将多核苷酸转移到受体细菌中的应用通常是有效的, 也是熟知的方法(如 US5,985,623)。根据细菌的菌株, 更通常使用感受态细胞与纯化 DNA 进行

转化。

公知的电穿孔技术(体外和体内)是通过向定位于处理区域附近的电极施加短暂高压脉冲而进行的。在电极间所形成的电场会在细胞膜上产生瞬时形成的小孔,通过该小孔,外源物质就可进入细胞。已知的电穿孔应用中,该电场含有在约 100 μ s 的时期中,1000V/cm 数量级的单一方形波脉冲。这样的脉冲可通过例如 BTX Division of Genetronics, Inc 制造的 Elector Square Porator T820 的已知应用来产生。

基因枪法是运用微弹轰击技术,将多核苷酸运送到靶细胞的系统。通过加速将多核苷酸运送到靶细胞的方法的说明性实施方案为 Biolistics Particle Delivery System,其可用于推进 DNA 或细胞包覆的颗粒通过一个屏幕,如不锈钢或 Nytex 屏幕,到达覆盖有培养细胞的过滤器表面。该屏幕会使颗粒散开,这样颗粒不会过于集中地到达培养细胞。据认为居于离子轰击仪和待轰击细胞之间的屏幕会减少粒子凝集的大小,可以通过减少由于过大微弹攻击受体细胞而造成的损失,从而提高转化频率。

为进行轰击,悬浮液中的细胞优选被浓缩到过滤器或固体培养基上。或者,可将其它靶细胞处理到固体培养基上。待轰击的细胞被固定到微弹阻停板下面适当的距离处。如果需要,还可在加速装置和待轰击细胞之间固定一个或多个屏幕。利用这些已知的技术,可以获得高于 1000 或更多数目的瞬时表达标记基因的细胞焦点。焦点中,轰击后 48 小时表达外源基因产物的细胞数目通常为 1-10,平均为 1-3。

在轰击转化中,可将轰击前培养条件和轰击参数最优化,以产生最大数目的稳定转化体。该项技术中轰击的物理和生物学参数都很重要。物理因素是那些与处置多核苷酸/-微弹沉淀相关的因素,或者是影响大或小轰击弹飞行和速度的因素。生物学因素包括在轰击之前和之后立即操作细胞的相关所有步骤,帮助减轻与轰击相关创伤的靶细胞渗透压的调整,还包括转化 DNA 的性质,如线性 DNA 或完整的超螺旋质粒。

相应地,可以考虑小范围地调整各种轰击参数以达到条件的最优化。技术人员可能特别期望调整物理参数,如间隙距(gap distance),飞行距离,组织距离(tissue distance)和氨压。技术人员还可以通过修饰影响受体细胞生理状态的条件从而影响转化和整合的效率,将损伤简化因子(trajuma reduction factors)(TRFs)最小化。例如,可以调整渗透状态,组织水合作用和

受体细胞的传代培养阶段或细胞周期来优选转化。参照本文，本领域技术人员将公知其它常规的调整操作。

微粒介导的转化方法是本领域技术人员公知的。例如，US 5,015,580(将其引入本文作为参考)描述了使用这种技术进行的大豆转化。

- 5 本发明的另一实施方案是对细菌进行改造使其生产类异戊二烯化合物的方法。这样的细菌是按如下制备：(a)在能够表达类异戊二烯的条件下，在培养基中培养亲代细菌，并从该培养基中筛选出突变细菌，该细菌生产的类异戊二烯水平为亲代细菌的约 1.1-1000 倍；(b)将含有 SEQ ID NO: 42 多核苷酸序列的表达载体导入突变细菌，该多核苷酸序列与表达调控序列
10 可操作相连；以及(c)筛选出含有表达载体并且生产类异戊二烯的水平比步骤(a)中的突变体高出至少约 1.1 倍的细菌。

在该实施方案中，类异戊二烯化合物指以异戊烯二磷酸(IPP)单元的下述通式为结构基础的化合物。



这种化合物包括半萜，单萜，倍半萜，双萜，三萜(如植物甾醇，植物雌激素(phytoestrogens)，植物昆虫脱皮激素，雌激素，植物雌激素(phytoestrogens))，四萜(类胡萝卜素)和多萜。优选地，类异戊二烯为类胡萝卜素，例如上述鉴定的类胡萝卜素之一，特别是玉米黄质。

- 20 细菌可以是使用本发明公开方法能够生成类异戊二烯化合物的任何细菌。优选地，该细菌为副球菌(Paracoccus)，黄杆菌属(Flavobacterium)，土壤杆菌属，产碱杆菌属(Alcaligenes)，欧文氏菌属(Erwinia)，大肠杆菌或枯草杆菌。甚至更优选地，该细胞为副球菌。优选地，亲代细菌为 R-1506 或 R1512，突变细菌为 R1534 或 R114，优选为 R114。

- 25 在优化生产类异戊二烯的条件下，在培养基中培养细菌。培养基的选择和培养条件是本领域公知的。实施例 1，11 和 12 中所示分析试验提供测定培养基中某种类胡萝卜素的存在的示范性试验方法。通过优化培养条件以及测定靶类异戊二烯的产量，可以实现符合本文所述特定产量参数的突变株的培养和筛选。以这种方法，可以筛选出类异戊二烯生产水平比亲代
30 细菌高约 1.1-1000 倍的突变细菌。优选地，该突变细菌生产的类异戊二烯比亲代细菌高约 1.5-500 倍，例如，比亲代细菌高至少约 100 倍或至少约 10

倍。然后培养该细菌，用于后续的步骤。

筛选出能够以期望的水平生产类异戊二烯的突变细菌后，使用前述的或实施例中所描述的任一方法将表达载体导入该细菌。本文所定义的任一表达载体均可导入该突变细胞。优选地，该表达载体含有 SEQ ID NO: 42。一旦将表达载体导入突变细菌后，筛选出类异戊二烯产量比未转化突变体高出至少约 1.1 倍，例如约 5-20 倍的稳定转化体。然后将该选出的转化体在适于类异戊二烯生成的条件下培养，接着从细胞或培养基中分离出类异戊二烯。

该方法的进一步步骤为向突变细菌中引入突变，该突变能够提高细菌的类异戊二烯化合物的产量。该突变可选自下述的至少一种：失活聚羟基烷羧酸酯(polyhydroxyalkanoate)(PHA)途径，提高乙酰-CoA 乙酰基转移酶的表达，提高 FPP 合成酶的表达，提高类胡萝卜素生物合成途径中酶的表达，提高转化异戊烯二磷酸(IPP)为二甲基丙烯二磷酸(dimethylallyl diphosphate)(DMAPP)的酶的表达。

失活聚羟基烷羧酸酯(polyhydroxyalkanoate)(PHA)途径可通过选择出不表达 phaB(SEQ ID NO: 177 的 1258-1980 位核苷酸)编码多肽的突变细菌而完成，或者使用 SEQ ID NO: 177 或其片段进行同源重组，通过破坏野生型 phaB 基因的表达而完成。

该方法中，提高乙酰-CoA 乙酰基转移酶的表达可通过下述完成：将含有 SEQ ID NO: 175 或 SEQ ID NO: 177 的 1-1170 位核苷酸所示多核苷酸序列的载体导入突变细菌，所述多核苷酸序列与表达调控序列可操作相连。该方法中，提高 FPP 合成酶的表达可通过下述完成：将含有 SEQ ID NO: 157 的 295-1158 位核苷酸所示多核苷酸序列的载体导入突变细菌，所述多核苷酸序列与表达调控序列可操作相连。该方法中，提高类胡萝卜素基因的表达可通过下述完成：将载体导入突变细菌，所述载体含有编码类胡萝卜素生物合成途径中一种或多种酶的多核苷酸序列，如选自下述的多核苷酸序列：SEQ ID NO: 180, 182 和 184，所述多核苷酸序列与表达调控序列可操作相连。

该方法中，优选类异戊二烯化合物为异戊烯二磷酸(IPP)。还优选该类异戊二烯化合物为类胡萝卜素，例如八氢番茄红素，番茄红素， β -胡萝卜素，玉米黄质，角黄素，虾青素，福寿草黄素(adonixanthin)，隐黄素，海胆酮和

福寿草红素(adonirubin)及其组合。

本发明的另一实施方案为副球菌属(*Paracoccus*)的微生物, 该微生物具有下述特点: (a)使用 GeneCompar v2.0 软件, 缺口罚分(gap penalty)为 0%条件下, 利用同源性计算获得的相似度基质, 与 SEQ ID NO: 12 的序列相似性
5 大于 97%; (b)在 81.5°C, 使用 DNA:DNA 杂交, 与 R-1512, R1534, R114 或 R-1506 的相似性大于 70%; (c)基因组 DNA 的 G+C 含量与 R114, R-1512, R1534 和 R-1506 的染色体组 DNA 的 G+C 含量相比, 区别小于 1%; 以及(d)使用实施例 2 的 AFLP 方法, 平均 DNA 指纹与菌株 R-1512, R1534, R114 和 R-1506 具有 cluster 约 58%的相似性, 附带条件为该微生物不属于副球菌种(MBIC3966)。
10

测定上述每一特征的方法在实施例 2 中充分描述, 可以预期当使用这些方法时, 很容易地检测出符合上述标准的微生物。优选本发明的微生物具有上述的每项特征(即 a-d)。但是, a-d 任一特征的组合也属于本发明的范围, 只要该组合能够提供足够的信息可以确凿地从分类学角度描述与 R114,
15 R-1512, R1534 和 R-1506 属于相同种的微生物, 除副球菌种之外(MBIC3966)。

本发明的另一实施例是副球菌属的微生物, 该微生物具有下述特征: (a)含有至少约 75%的细胞膜总脂肪酸的 18:1w7c; (b)不能使用福寿草醇(adonitol), i-赤藓糖醇, 龙胆二糖, β -甲基葡萄糖苷, D-山梨糖醇, 木糖醇和奎尼酸作为碳源用于生长; 和(c)使用 L-天冬酰胺和 L-天冬氨酸作为碳源用于生长, 附带条件为该微生物不属于副球菌种(MBIC3966)。
20

实施例 2 也详细描述了测定每项这些特征的方法, 可以预期当使用这些方法时, 可以容易地检测出符合上述标准的微生物。优选本发明的微生物具有上述的每项特征(即 a-c)。但是, a-c 任一特征的组合也属于本发明的
25 范围, 只要该组合能够提供足够的信息可以确凿地从分类学角度描述与 R114, R-1512, R1534 和 R-1506 属于相同种的微生物, 除副球菌种之外(MBIC3966)。

本发明的另一实施方案是副球菌属的微生物, 该微生物具有下述特征: (a)能够在 40°C 生长; (b)能够在含有 8%NaCl 的培养基中生长; (c)能够在 pH9.1 的培养基中生长; 以及(d)黄色-橙色集落色素形成, 附带条件为该微生物不属于副球菌种(MBIC3966)。
30

实施例 2 也详细描述了测定上述每项特征的方法, 可以预期当使用这些方法时, 可以容易地检测出符合上述标准的微生物。优选本发明的微生物具有上述的每项特征(即 a-d)。但是, a-d 任一特征的组合也属于本发明的范围, 只要该组合能够提供足够的信息可以确凿地从分类学角度描述与

5 R114, R-1512, R1534 和 R-1506 属于相同种的微生物, 除副球菌种之外(MBIC3966)。

本发明的微生物还可以使用上述 11 项特征的组合进行定义, 只要该组合能够提供足够的信息可以确凿地从分类学角度描述与 R114, R-1512, R1534 和 R-1506 属于相同种的微生物, 除副球菌种之外(MBIC3966)。

10 综上所述, 本发明提供

(1) 含有选自下述氨基酸序列的分离多肽:

(a) SEQ ID NO: 43 的 1-340 残基所示的氨基酸序列, 特别是相应于 SEQ ID NO: 43 的 68-97 位的氨基酸序列;

15 (b) SEQ ID NO: 45 的 1-349 残基所示的氨基酸序列, 特别是相应于 SEQ ID NO: 45 的 1-30 位的氨基酸序列;

(c) SEQ ID NO: 47 的 1-388 残基所示的氨基酸序列, 特别是相应于 SEQ ID NO: 47 的 269-298 位的氨基酸序列;

(d) SEQ ID NO: 49 的 1-378 残基所示的氨基酸序列, 特别是相应于 SEQ ID NO: 49 的 109-138 位的氨基酸序列;

20 (e) SEQ ID NO: 51 的 1-305 残基所示的氨基酸序列, 特别是相应于 SEQ ID NO: 51 的 198-227 位的氨基酸序列;

(f) SEQ ID NO: 53 的 1-332 残基所示的氨基酸序列, 特别是相应于 SEQ ID NO: 53 的 81-110 位的氨基酸序列;

25 (g) 选自 SEQ ID NO: 43, 45, 47, 49, 51 或 53 的氨基酸序列的片段, 其中所述片段具有至少 30 个连续的氨基酸残基;

(h) 选自 SEQ ID NO: 43, 45, 47, 49, 51 或 53 的多肽片段的氨基酸序列, 该片段具有羟基甲基戊二酸单酰-CoA 还原酶(HMG-CoA 还原酶), 异戊烯二磷酸异构酶, 羟基甲基戊二酸单酰-CoA 合成酶(HMG-CoA 合成酶), 甲羟戊酸激酶, 磷酸甲羟戊酸激酶, 或二磷酸甲羟戊酸脱羧酶的活性;

30 (i) 在严谨条件下能够与杂交探针杂交的多核苷酸所编码的多肽氨基酸序列, 所述杂交探针含有至少 SEQ ID NO: 42 或 SEQ ID NO: 42 互补序列

的 30 个连续多核苷酸，其中所述多肽具有 HMG-CoA 还原酶，异戊烯二磷酸异构酶，HMG-CoA 合成酶，甲羟戊酸激酶，磷酸甲羟戊酸激酶，或二磷酸甲羟戊酸脱羧酶的活性；以及

(j) SEQ ID NO: 43, 45, 47, 49, 51 或 53 经保守修饰而形成的各变体。

5 (2) 含有选自下述氨基酸序列的分离多肽：

(a) SEQ ID NO: 159 的 1-287 残基所示氨基酸序列；

(b) SEQ ID NO: 159 的至少 30 个连续氨基酸残基；

(c) SEQ ID NO: 159 的片段的氨基酸序列，该片段具有法呢基二磷酸合成酶(FPP 合成酶)的活性；

10 (d) 在严谨条件下能够与杂交探针杂交的多核苷酸所编码的多肽氨基酸序列，所述杂交探针含有至少 SEQ ID NO: 157 或其互补序列的 295-1158 位的 30 个连续多核苷酸，其中所述多肽具有 FPP 合成酶的活性；以及

(e) SEQ ID NO: 159 经保守修饰而形成的变体。

(3) 含有选自下述氨基酸序列的分离多肽：

15 (a) SEQ ID NO: 160 的 1-142 残基所示氨基酸序列；

(b) SEQ ID NO: 160 的至少 30 个连续氨基酸残基；

(c) SEQ ID NO: 160 的片段的氨基酸序列，该片段具有 1-脱氧木酮糖-5-磷酸合成酶(DXPS)的活性；

20 (d) 在严谨条件下能够与杂交探针杂交的多核苷酸所编码的多肽氨基酸序列，所述杂交探针含有至少 SEQ ID NO: 157 或其互补序列的 1185-1610 位的 30 个连续多核苷酸，其中所述多肽具有 DXPS 的活性；以及

(e) SEQ ID NO: 160 经保守修饰而形成的各变体。

(4) 含有选自下述氨基酸序列的分离多肽：

(a) SEQ ID NO: 178 的 1-390 位残基所示氨基酸序列；

25 (b) SEQ ID NO: 178 的至少 30 个连续氨基酸残基；

(c) SEQ ID NO: 178 的多肽片段的氨基酸序列，该片段具有乙酰-CoA 乙酰基转移酶的活性；

30 (d)在严谨条件下能够与杂交探针杂交的多核苷酸所编码的多肽氨基酸序列，所述杂交探针含有至少 SEQ ID NO: 177 或其互补序列的 1-1170 位的 30 个连续多核苷酸，其中所述多肽具有乙酰-CoA 乙酰基转移酶的活性；以及

(e) SEQ ID NO: 178 经保守修饰而形成的变体。

(5) 含有选自下述氨基酸序列的分离多肽:

(a) SEQ ID NO: 179 的 1-240 位残基所示氨基酸序列;

(b) SEQ ID NO: 179 的至少 30 个连续氨基酸残基;

5 (c) SEQ ID NO: 179 的多肽片段的氨基酸序列, 该片段具有乙酰乙酰-CoA 还原酶的活性;

(d) 在严谨条件下能够与杂交探针杂交的多核苷酸所编码的多肽氨基酸序列, 所述杂交探针含有至少 SEQ ID NO: 177 或其互补序列的 1258-1980 位的 30 个连续多核苷酸, 其中所述多肽具有乙酰乙酰-CoA 还原酶的活性;

10 以及

(e) SEQ ID NO: 179 经保守修饰而形成的各变体。

(6) 含有选自下述核苷酸序列的分离多核苷酸序列: SEQ ID NO: 42, SEQ ID NO: 42 变体的核苷酸序列, 所述变体含有一个或多个符合副球菌种菌株 R1534 密码子利用表(参见表 14)的取代, SEQ ID NO: 42 的片段, 该片段编码具有选自下述活性的多肽: 羟基甲基戊二酸单酰-CoA 还原酶 (HMG-CoA 还原酶), 异戊烯二磷酸异构酶, 羟基甲基戊二酸单酰-CoA 合成酶 (HMG-CoA 合成酶), 甲羟戊酸激酶, 磷酸甲羟戊酸激酶和二磷酸甲羟戊酸脱羧酶, 在严谨条件下能够与杂交探针杂交的多核苷酸序列, 所述探针由 SEQ ID NO: 42 或其互补序列的至少 30 个连续核苷酸构成, 该多核苷酸
15 编码具有选自下述活性的多肽: HMG-CoA 还原酶, 异戊烯二磷酸异构酶, HMG-CoA 合成酶, 甲羟戊酸激酶, 磷酸甲羟戊酸激酶和二磷酸甲羟戊酸脱羧酶; 特别是

(a) 分离的多核苷酸序列, 其含有的多核苷酸序列选自: SEQ ID NO: 42 的 2622-3644 位核苷酸, SEQ ID NO: 42 的片段, 其编码具有 HMG-CoA 还原酶活性的多肽, 以及在严谨条件下能够与杂交探针杂交的多核苷酸序列, 所述探针由 SEQ ID NO: 42 或其互补序列 2622-3644 位残基的至少 30 个连续核苷酸构成, 所述多核苷酸编码具有 HMG-CoA 还原酶活性的多肽, 尤其是 SEQ ID NO: 42 的 2622-3644 位核苷酸构成的序列;

(b) 分离的多核苷酸序列, 其含有的多核苷酸序列选自: SEQ ID NO: 42 的 3641-4690 位核苷酸, 其各种变体, 所述变体含有一个或多个符合副球菌种菌株 R1534 密码子利用表(参见表 14)的取代, SEQ ID NO: 42 的片段, 该
30

片段编码具有异戊烯二磷酸异构酶活性的多肽, 以及在严谨条件下能够与杂交探针杂交的多核苷酸序列, 所述探针由 SEQ ID NO: 42 或其互补序列 3641-4690 位残基的至少 30 个连续核苷酸构成, 所述多核苷酸编码具有异戊烯二磷酸异构酶活性的多肽, 尤其是 SEQ ID NO: 42 的 3641-4690 位核苷酸构成的多核苷酸序列;

(c) 分离的多核苷酸序列, 其含有的多核苷酸序列选自: SEQ ID NO: 42 的 4687-5853 位核苷酸, 其各种变体, 所述变体含有一个或多个符合副球菌种菌株 R1534 密码子利用表(参见表 14)的取代, SEQ ID NO: 42 的片段, 该片段编码具有 HMG-CoA 合成酶活性的多肽, 以及在严谨条件下能够与杂交探针杂交的多核苷酸序列, 所述探针由 SEQ ID NO: 42 或其互补序列 4687-5853 位残基的至少 30 个连续核苷酸构成, 所述多核苷酸编码具有 HMG-CoA 合成酶活性的多肽, 尤其是 SEQ ID NO: 42 的 3641-4690 位核苷酸构成的序列;

(d) 分离的多核苷酸序列, 其含有的多核苷酸序列选自: SEQ ID NO: 42 的 5834-6970 位核苷酸, 其各种变体, 所述变体含有一个或多个符合副球菌种菌株 R1534 密码子利用表(参见表 14)的取代, SEQ ID NO: 42 的片段, 该片段编码具有甲羟戊酸激酶活性的多肽, 以及在严谨条件下能够与杂交探针杂交的多核苷酸序列, 所述探针由 SEQ ID NO: 42 或其互补序列 5834-6970 位残基的至少 30 个连续核苷酸构成, 所述多核苷酸编码具有甲羟戊酸激酶活性的多肽, 尤其是 SEQ ID NO: 42 的 3641-4690 位核苷酸构成的多核苷酸序列;

(e) 分离的多核苷酸序列, 其含有的多核苷酸序列选自: SEQ ID NO: 42 的 6970-7887 位核苷酸, 其各种变体, 所述变体含有一个或多个符合副球菌种菌株 R1534 密码子利用表(参见表 14)的取代, SEQ ID NO: 42 的片段, 该片段编码具有磷酸甲羟戊酸激酶活性的多肽, 以及在严谨条件下能够与杂交探针杂交的多核苷酸序列, 所述探针由 SEQ ID NO: 42 或其互补序列 6970-7887 位残基的至少 30 个连续核苷酸构成, 所述多核苷酸编码具有磷酸甲羟戊酸激酶活性的多肽, 尤其是 SEQ ID NO: 42 的 3641-4690 位核苷酸构成的多核苷酸序列;

(f) 分离的多核苷酸序列, 其含有的多核苷酸序列选自: SEQ ID NO: 42 的 7880-8878 位核苷酸, 其各种变体, 所述变体含有一个或多个符合副球菌

种菌株 R1534 密码子利用表(参见表 14)的取代, SEQ ID NO: 42 的片段, 该片段编码具有二磷酸甲羟戊酸脱羧酶活性的多肽, 以及在严谨条件下能够与杂交探针杂交的多核苷酸序列, 所述探针由 SEQ ID NO: 42 或其互补序列 7880-8878 位残基的至少 30 个连续核苷酸构成, 所述多核苷酸编码具有二磷酸甲羟戊酸脱羧酶活性的多肽, 尤其是 SEQ ID NO: 42 的 3641-4690 位核苷酸构成的多核苷酸序列;

(7) 含有选自下述多核苷酸序列的分离多核苷酸序列: SEQ ID NO: 157 的核苷酸序列, SEQ ID NO: 157 各变体, 所述变体含有一个或多个符合副球菌种菌株 R1534 密码子利用表(参见表 14)的取代, SEQ ID NO: 157 的片段, 该片段编码具有选自下述活性的多肽: 法呢基焦磷酸(FPP)合成酶活性, 1-脱氧-D-木酮糖-5-磷酸合成酶活性, 或编码具有 XseB 活性的多肽, 以及在严谨条件下能够与杂交探针杂交的多核苷酸序列, 所述探针由 SEQ ID NO: 157 或其互补序列的至少 30 个连续核苷酸构成, 该多核苷酸编码具有选自下述活性的多肽: FPP 合成酶活性, 1-脱氧-D-木酮糖-5-磷酸合成酶活性, 以及 XseB 活性, 特别是

(a) 分离的多核苷酸序列, 其含有的多核苷酸序列选自: SEQ ID NO: 157 的 59-292 位核苷酸序列, SEQ ID NO: 157 各变体, 所述变体含有一个或多个符合副球菌种菌株 R1534 密码子利用表(参见表 14)的取代, SEQ ID NO: 157 的 59-292 位核苷酸序列的片段, 该片段编码具有 XseB 功能的多肽, 以及在严谨条件下能够与杂交探针杂交的多核苷酸序列, 所述探针由 SEQ ID NO: 157 或其互补序列 59-292 位残基的至少 30 个连续核苷酸构成, 所述多核苷酸编码具有 XseB 功能的多肽, 尤其是 SEQ ID NO: 157 的 59-292 位核苷酸构成的序列;

(b) 分离的多核苷酸序列, 其含有的多核苷酸序列选自: SEQ ID NO: 157 的 295-1158 位核苷酸序列, SEQ ID NO: 157 的 295-1158 位核苷酸序列的变体, 所述变体含有一个或多个符合副球菌种菌株 R1534 密码子利用表(参见表 14)的取代, SEQ ID NO: 157 的 295-1158 位核苷酸序列的片段, 该片段编码具有 FPP 合成酶活性的多肽, 以及在严谨条件下能够与杂交探针杂交的多核苷酸序列, 所述探针由 SEQ ID NO: 157 或其互补序列 295-1158 位残基的至少 30 个连续核苷酸构成, 所述多核苷酸编码具有 FPP 合成酶活性多肽, 尤其是 SEQ ID NO: 157 的 295-1158 位核苷酸构成的多核苷酸序列;

(c) 分离的多核苷酸序列, 其含有的多核苷酸序列选自: SEQ ID NO: 157 的 1185-1610 位核苷酸序列, SEQ ID NO: 157 的 1185-1610 位核苷酸序列的各变体, 所述变体含有一个或多个符合副球菌种菌株 R1534 密码子利用表(参见表 14)的取代, SEQ ID NO: 157 的 1185-1610 位核苷酸序列的片段, 该片段编码具有 1-脱氧木酮糖-5-磷酸合成酶活性的多肽, 以及在严谨条件下能够与杂交探针杂交的多核苷酸序列, 所述探针由 SEQ ID NO: 157 或其互补序列 1185-1610 位残基的至少 30 个连续核苷酸构成, 所述多核苷酸编码具有 1-脱氧木酮糖-5-磷酸合成酶活性多肽, 尤其是 SEQ ID NO: 157 的 1185-1610 位核苷酸构成的序列;

- 10 (8) 含有选自下述多核苷酸序列的分离多核苷酸序列: SEQ ID NO: 177 的核苷酸序列, SEQ ID NO: 177 各变体, 所述变体含有一个或多个符合副球菌种菌株 R1534 密码子利用表(参见表 14)的取代, SEQ ID NO: 177 的片段, 该片段编码具有选自下述活性的多肽: 乙酰-CoA 乙酰基转移酶活性和乙酰乙酰-CoA 还原酶活性, 以及在严谨条件下能够与杂交探针杂交的多核苷酸序列, 所述探针由 SEQ ID NO: 177 或其互补序列的至少 30 个连续核苷酸构成, 该多核苷酸编码具有选自下述活性的多肽: 乙酰-CoA 乙酰基转移酶和乙酰乙酰-CoA 还原酶, 特别是

- 20 (a) 分离的多核苷酸序列, 其含有的多核苷酸序列选自: SEQ ID NO: 177 的 1-1170 位核苷酸序列, SEQ ID NO: 177 各变体, 所述变体含有一个或多个符合副球菌种菌株 R1534 密码子利用表(参见表 14)的取代, SEQ ID NO: 177 的片段, 该片段编码具有乙酰-CoA 乙酰基转移酶活性的多肽, 以及在严谨条件下能够与杂交探针杂交的多核苷酸序列, 所述探针由 SEQ ID NO: 177 或其互补序列 1-1170 位残基的至少 30 个连续核苷酸构成, 所述多核苷酸编码具有乙酰-CoA 乙酰基转移酶活性的多肽, 尤其是 SEQ ID NO: 25 177 的 1-1170 位核苷酸构成的序列;

- (b) 分离的多核苷酸序列, 其含有的多核苷酸序列选自: SEQ ID NO: 177 的 1258-1980 位核苷酸序列, SEQ ID NO: 177 各变体, 所述变体含有一个或多个符合副球菌种菌株 R1534 密码子利用表(参见表 14)的取代, SEQ ID NO: 177 的片段, 该片段编码具有乙酰乙酰-CoA 还原酶活性的多肽, 以及在严谨条件下能够与杂交探针杂交的多核苷酸序列, 所述探针由 SEQ ID NO: 177 或其互补序列 1258-1980 位残基的至少 30 个连续核苷酸构成, 所

述多核苷酸编码具有乙酰乙酰-CoA 还原酶活性的多肽, 尤其是 SEQ ID NO: 177 的 1258-1980 位核苷酸构成的序列;

(9) 含有选自下述核苷酸序列的分离多核苷酸序列: SEQ ID NO: 42, SEQ ID NO: 157, SEQ ID NO: 177 及其组合;

- 5 (10) 含有(6)(a)-(6)(f), (7)(a)-(7)(c), (8)(a), (8)(b)或(9)任一多核苷酸序列的表达载体, 特别是这样的表达载体: 其中多核苷酸序列与表达调控序列可操作相连, 例如进一步含有编码类胡萝卜素生物合成途径中酶的多核苷酸序列的表达载体, 更特别是其中所含多核苷酸序列选自下述的表达载体: SEQ ID NO: 180, SEQ ID NO: 182, SEQ ID NO: 184 及其组合, 该多核苷酸序列与表达调控序列可操作相连;

- (11) 选自下述的表达载体: pBBR-K-mev-op16-1, pBBR-K-mev-op16-2, pDS-mvaA, pDS-idi, pDS-hcs, pDS-mvk, pDS-pmk, pDS-mvd, pDS-His-mvaA, pDS-His-idi, pDS-His-hcs, pDS-His-mvk, pDS-His-pmk, pDS-His-mvd, pBBR-K-Zea4, pBBR-K-Zea4-up, pBBR-K-Zea4-down, pBBR-K-PcrtE-crtE-3, 15 pBBR-tK-PcrtE-mvaA, pBBR-tK-PcrtE-idi, pBBR-tK-PcrtE-hcs, pBBR-tK-PcrtE-mvk, pBBR-tK-PcrtE-pmk, pBBR-tK-PcrtE-mvd, pBBR-K-PcrtE-mvaA-crtE-3, pDS-His-phaA, pBBR-K-PcrtE-crtW, pBBR-K-PcrtE-crtWZ, pBBR-K-PcrtE-crtZW 及其组合, 特别是

- (a) 选自 pBBR-K-mev-op16-1 或 pBBR-K-mev-op16-2 的表达载体, 20 (b) 选自 pBBR-K-Zea4, pBBR-K-Zea4-up 或 pBBR-K-Zea4-down 的表达载体,

- (c) 选自 pBBR-K-PcrtE-crtE-3, pBBR-tK-PcrtE-mvaA, pBBR-tK-PcrtE-idi, pBBR-tK-PcrtE-hcs, pBBR-tK-PcrtE-mvk, pBBR-tK-PcrtE-pmk, pBBR-tK-PcrtE-mvd 或其组合的表达载体,

- 25 (d) 表达载体 pBBR-K-PcrtE-mvaA-crtE-3,

- (e) 表达载体 pDS-His-phaA, 或

- (f) 选自 pBBR-K-PcrtE-crtW, pBBR-K-PcrtE-crtWZ 或 pBBR-K-PcrtE-crtZW 的表达载体;

- (12) 含有(6)(a)-(f), (7)(a)-(c), (8)(a), (8)(b)或(9)任一多核苷酸序列的 30 培养细胞, 或含有(10)或(11)表达载体的培养细胞, 或该细胞的后代, 其中该细胞表达由所述多核苷酸编码的多肽, 特别是进一步具有选自下述特征

的细胞:

(a) 进一步含有编码类胡萝卜素生物合成途径中酶的多核苷酸序列, 特别是这样的培养细胞: 该细胞中编码类胡萝卜素生物合成途径酶的多核苷酸序列是选自: SEQ ID NO: 180, SEQ ID NO: 182, SEQ ID NO: 184, 或者是
5 该细胞的后代, 其中该细胞表达由所述多核苷酸序列编码的多肽, 以及

(b) 选自酵母, 真菌, 细菌和藻类, 特别是选自下述的细菌: 副球菌, 黄杆菌属(Flavobacterium), 土壤杆菌属(Agrobacterium), 产碱杆菌属(Alcaligenes), 欧文氏杆菌属(Erwinia), 大肠杆菌(E.coli)或枯草杆菌(B.subtilis), 更具体是副球菌, 尤其是选自 R-1506, R-1512, R1534 或 R114
10 的副球菌。

(13) 生产类胡萝卜素的方法, 包括: 在能够表达多核苷酸序列编码多肽的条件下培养(12)细胞, 并从细胞或该细胞的培养基中分离出类胡萝卜素;

(14) 制备产类胡萝卜素细胞的方法, 包括:

15 (a) 向细胞中导入编码甲羟戊酸途径中酶的多核苷酸序列, 所述酶在该细胞中表达; 以及

(b) 选择出含有步骤(a)多核苷酸序列的细胞, 并且该细胞所产生类胡萝卜素的水平为该细胞导入多核苷酸序列之前的 1.1-1000 倍, 特别是以下述特点为特征的方法:

20 (i) 筛选步骤包括选择出含有步骤(a)多核苷酸序列的细胞, 且该细胞所产生类胡萝卜素的水平约为该细胞导入多核苷酸序列之前的 1.5-500 倍, 特别是约 100 倍, 或至少约 10 倍;

(ii) 该细胞的类胡萝卜素产量约为 1mg/L-约 10g/L。

(ii) 该细胞选自酵母, 真菌, 细菌和藻类, 特别是选自下述的细菌: 副
25 球菌, 黄杆菌属(Flavobacterium), 土壤杆菌属(Agrobacterium), 产碱杆菌属(Alcaligenes), 欧文氏杆菌属(Erwinia), 大肠杆菌(E.coli)或枯草杆菌(B.subtilis), 更具体是副球菌;

(iii) 步骤(a)的细胞为突变细胞, 特别是选自 R114 和 R1534 的细胞, 特别是这样的突变细胞: 其类胡萝卜素化合物产量约为其非突变亲代的
30 1.1-1000 倍, 特别是约 1.5-500 倍, 更特别是至少约高 100 倍, 或至少高 10 倍;

(v) 选自(6)(a)-(f), (7)(a)-(c), (8)(a), (8)(b)或(9)的多核苷酸序列, 特别是与表达调控序列可操作相连的多核苷酸序列;

(vi) 表达载体(10)或(11)的多核苷酸序列;

(vii) 导入步骤选自: 转化, 转导, 转染, 脂转染, 电穿孔, 接合或基因枪法。

(viii) 类胡萝卜素选自: 八氢番茄红素, 番茄红素, β -胡萝卜素, 玉米黄质, 角黄素, 虾青素, 福寿草黄素(adonixanthin), 隐黄素, 海胆酮和福寿草红素(adonirubin)及其组合; 特别是类胡萝卜素为玉米黄质。

(15) 改造细菌使其生成类异戊二烯化合物的方法, 包括:

10 (a) 在能够表达类异戊二烯化合物的条件下, 在培养基中培养亲代细菌, 并从该培养基中筛选出突变细菌, 所述突变细菌类异戊二烯化合物产量比亲代细菌高约 1.1-1000 倍;

(b) 向该突变细菌中导入含有 SEQ ID NO: 42 所示多核苷酸序列的表达载体, 所述多核苷酸序列与表达调控序列可操作相连; 以及

15 (c) 筛选出含有该表达载体的细菌, 并且所述细菌的类异戊二烯化合物产量比步骤(a)的突变株至少高 1.1 倍, 尤其是

(i) 进一步向突变细菌导入突变的方法, 更特别是所述突变产生下述至少一种作用的方法: 使聚羟基链羧基酯(PHA)途径失活, 提高乙酰-CoA 乙酰基转移酶的表达, 提高法呢基焦磷酸(FPP)合成酶的表达, 提高类胡萝卜素途径酶的表达, 提高转化异戊烯二磷酸(IPP)为二甲基烯丙基焦磷酸(DMAPP)的酶的表达,

25 最尤其是使聚羟基链羧基酯(PHA)途径失活的方法, 包括筛选出突变细菌, 该细菌不表达 phaB(SEQ ID NO: 177 的 1258-1980 位核苷酸)编码的多肽, 或者该方法包括通过使用 SEQ ID NO: 177 或其片段进行同源重组, 从而破坏野生型 phaB 基因的表达, 或者

一种提高乙酰-CoA 乙酰基转移酶的表达的方法, 包括向突变细菌中导入含有 SEQ ID NO: 175 所示核苷酸序列或 SEQ ID NO: 177 1-1170 位核苷酸的载体, 所述核苷酸序列与表达调控序列可操作相连, 或

30 一种提高 FPP 合成酶的表达的方法, 包括向突变细菌中导入含有 SEQ ID NO: 157 的 295-1158 位核苷酸所示多核苷酸序列的载体, 所述核苷酸序列与表达调控序列可操作相连, 或

一种提高类胡萝卜素途径酶的表达的方法, 包括向突变细菌中导入载体, 所述载体含有选自下述的多核苷酸序列: SEQ ID NO: 180, 182 和 184, 且该多核苷酸与表达调控序列可操作相连,

(a) 一种方法, 其中类异戊二烯为异戊烯二磷酸(IPP)。

- 5 (b) 一种方法, 其中类异戊二烯为类胡萝卜素, 特别是类胡萝卜素选自下述的方法: 八氢番茄红素, 番茄红素, β -胡萝卜素, 玉米黄质, 角黄素, 虾青素, 福寿草黄素(adonixanthin), 隐黄素, 海胆酮和福寿草红素(adonirubin) 及其组合;

- (c) 一种方法, 其中亲代细菌为副球菌, 特别是 R1512 或 R-1506, 或
10 R1534 或 R114, 特别是突变株为 R114;

(16) 副球菌属微生物, 其具有下述特点:

- (i) 使用利用 GeneCompar v2.0 软件, 缺口罚分(gap penalty)为 0%条件下, 利用同源性计算获得的相似度基质(similarity matrix), 与 SEQ ID NO: 12 的序列相似性大于 97%; 在 81.5°C, 使用 DNA:DNA 杂交, 与 R-1512, R1534,
15 R114 或 R-1506 的同源大于 70%; 基因组 DNA 的 G+C 含量与 R114, R-1512, R1534 和 R-1506 的基因组 DNA 的相比, 区别小于 1%; 以及 使用实施例 2 的 AFLP 方法, 平均 DNA 指纹与菌株 R-1512, R1534, R114 和 R-1506 具有约 58%的相似性, 附带条件为该微生物不属于副球菌种(MBIC3966);

- (ii) 含有至少约 75%的细胞膜总脂肪酸的 18:1w7c; 能够抑制使用福草
20 糖醇(adonitol), i-赤藓糖醇, 龙胆二糖, β -甲基葡萄糖苷, D-山梨糖醇, 木糖醇和奎尼酸作为碳源用于生长; 和 能够使用 L-天冬酰胺和 L-天冬氨酸作为碳源用于生长, 附带条件为该微生物不属于副球菌种(MBIC3966), 或

- (iii) 能够在 40°C 生长; 能够在含有 8%NaCl 的培养基中生长; 能够在 pH9.1 的培养基中生长; 以及黄色-橙色集落色素形成, 附带条件为该微生物
25 物不属于副球菌种(MBIC3966)。

下述实施例用于进一步说明本发明各个方面。这些实施例仅为说明目的, 而非限制本发明的范围。

实施例 1: 分析方法和生化方法

- 30 (a) 类胡萝卜素的分析

样品制备 首先制备二甲基亚砷(DMSO)和四氢呋喃的 1: 1 溶剂混合物。加

入丁化羟基甲苯(butylated hydroxytoluene)(BHT, 0.5g/l 溶剂混合物)稳定溶剂混合物。向装有 0.4ml 细菌培养物的一次性 15ml 聚丙烯离心管中加入 4 毫升稳定了的 DMSO/THF 混合物(最终稀释因子为 1/11)。将这些离心管盖严, 每个离心管均用 Vortex 混合仪混合 10 秒钟。然后将该样品放入

5 Brinkmann Vibramix 振荡器内振荡 20 分钟。室温下, 4000rpm 离心 4 分钟, 将各份澄清的黄色/橙色上清液分别转移到棕色玻璃瓶中, 用于高效液相色谱(HPLC)。

HPLC 发展一种反向 HPLC 方法, 用于同时测定虾青素, 玉米黄质, 角黄素, β -胡萝卜素和番茄红素。该方法能够将玉米黄质的各主要顺式异构体分离。

10 使用带有恒温自动取样器和二极管阵列检测仪的 Agilent 1100 HPLC 系统进行色谱。该方法的各参数如下:

柱: YMC 类胡萝卜素 C30 柱, 微粒大小为 5 微米
250* 4.6mm I.D. 钢
YMC, Part No. CT99S052546WT)

15 Guard 柱: Pelliguard LC-18 cartridge, 20mm
(SUPELCO, Part No. 59654)

移动相: 甲醇(MeOH)/甲基特丁基醚(TBME)梯度

	%MeOH	%TBME
起始	80	20
10min	65	35
20min	10	90

进行时间: 28 分钟; 典型柱压力: 初始 90 巴; 流速: 1.0ml/min;

20 检测: 450nm 紫外光; 加入体积: 10 μ l; 柱温: 15°C

试剂 甲醇和 TBME 为 HPLC 级, 分别来自 EM Science 和 J.T.Baker。DMSO(Omnisolve)购自 EM Science。THF(HPLC 溶剂)来自 Burdick 和 Jackson。

计算 使用外标准(由 Hoffmann-La Roche, Basel, Switzerland 提供), 用二级校准(two level calibration)进行定量计算。计算基于峰面积。

25

选择性 该方法的选择性依据相关类胡萝卜素对照化合物标准溶液的加入而得到检验。靶化合物(全反式类胡萝卜素(all-trans-carotenoids))被彻底分

离，并未显示干涉。一些次要的顺式异构体可以共洗脱，尽管这些潜在的干涉异构体非常少量，可以在常规分析中忽略不计。组合物的滞留时间如表 1 所列。

表 1 类胡萝卜素的 HPLC 滞留时间

类胡萝卜素	滞留时间(min)	类胡萝卜素	滞留时间(min)
虾青素	6.99	角黄素	9.95
福寿草黄素	7.50	隐黄素	13.45
15-顺-玉米黄质	7.80	β-胡萝卜素	17.40
13-顺-玉米黄质	8.23	番茄红素	21.75
全-反-玉米黄质	9.11		

5

直线性 将 25 毫克全反式玉米黄质溶于 50ml DMSO/THF 混合物(玉米黄质终浓度为 500µg/ml)。制备稀释系列(玉米黄质终浓度为 250，100，50，10，5，1 和 0.1µg/ml)，并用上述的 HPLC 方法进行分析。发现线性范围在 0.1µg/ml-250µg/ml。相关系数为 0.9998。

10 检测的极限值 该方法对玉米黄质进行检测的检测下限被测定为 60µg/l。更高的注射体积以及优化整合参数可以将检测极限减低到约 5µg/l。

重现性 全顺式玉米黄质的滞留时间非常稳定(相对标准偏差(RSD)，0.2%)。根据对相同培养样品进行的 10 次重复性分析，峰面积重现性测定为：全反式玉米黄质 0.17%RSD，隐黄素 1.0%。

15 (b) 制备粗提物以及酶试验方法

制备粗提物 用 1ml 抽提缓冲液(视被试验的酶而定——组分与下述的各个酶学试验方法而具体说明)重悬浮洗涤的细胞沉淀，制备副球菌和大肠杆菌粗提物。将细胞悬浮液置于 2ml 塑料瓶中，并使用 Mini Bead Beater 8(Biospec Products, Bartlesville OK USA)玻璃珠振荡，破碎沉淀。然后在 4℃，使用培养基振荡器进行破裂。在 4℃下，21000x g 离心该破裂的沉淀制备溶液 20 分钟，使细胞碎片沉淀，上清液直接用于酶试验。

蛋白检测 使用 Bio-Rad Protein Assay Reagent(Bio-Rad, Hercules, CA, USA)，通过 Bradford[Anal. Biochem. 72: 248-2254(1976)]的方法测定粗提物中的蛋白浓度。使用牛血清白蛋白作为对照蛋白，用于构建标准曲线。

25 乙酰-CoA 乙酰基转移酶试验 粗提物在 pH8.0 的 150mM EPPS(N-[2-羟乙

基]piperizine-N'-[3-propanesulfonic acid])缓冲液中制备。根据 Slater 等人[J.Bacteriol. 180:1979-1987(1998)]的描述,在硫解作用方向上进行试验。该试验在 304nm 下,利用分光光度计测量乙酰乙酰-CoA 的消失。反应混合物含有 150mM EPPS 缓冲液(pH8.0), 50mM MgCl₂, 100μM CoA, 40μM 乙酰乙酰-CoA 和粗提物。反应在 30℃进行,通过加入粗提物起始反应。使用 SpectraMAX Plus 板式读数器(Molecular Devices Corp., Sunnyvale, CA, USA)和石英微量滴定板(也可使用任何标准分光光度计),在 304nm 下监测乙酰乙酰-CoA 的消失。使用乙酰乙酰-CoA(1 单位活性=1μmol 被消耗的乙酰乙酰-CoA/分钟)制备的标准曲线,计算活性(表达为 U/mg 蛋白)。乙酰-CoA 乙酰基转移酶活性的检测下限为 0.006U/mg。

HMG-CoA 合成酶试验 根据 Honda 等人[Hepatology 27: 154-159(1998)]的方法试验 HMG-CoA 合成酶。该试验中,从乙酰-CoA 和乙酰乙酰-CoA 形成的 HMG-CoA 通过由 HPLC 分离反应产物和底物直接进行测定。粗提物在 50mM Tris-HCl 缓冲液中(pH8.0)制备。反应混合物(0.1ml)含有 50mM Tris-HCl 缓冲液(pH8.0), 0.1mM EDTA, 20mM MgCl₂, 0.1mM 乙酰乙酰-CoA, 0.8mM 乙酰-CoA 和粗提物。加入粗提物之前,在 30℃预温育反应混合物 2 分钟。在 30℃反应 5 分钟后,加入 0.2ml 200mM 四丁基磷酸铵(TBAP, 溶解在甲醇-水(3: 2, 终 pH 为 5.5)中,并含有 0.2mM 丙酰-CoA 作为内部回收标准)终止反应。然后在 4℃, 21000x g 离心混合物 3 分钟,并随即保存在冰上,直到通过反相离子对 HPLC 进行分析。使用 Nova-Pak C18 柱(3.9x150mm, Waters Corporation, Milford, MA, USA),从乙酰-CoA 和乙酰乙酰-CoA 中分离出 HMG-CoA 和丙酰-CoA。注入体积为 20μl,移动相为溶于甲醇-水(1: 1, 终 pH 为 5.5)的 50mM TBAP,流速为 1.0ml/min。通过在 254nm 下的吸收值来检测 HMG-CoA 和丙酰-CoA。该反应中形成的 HMG-CoA 与使用可信 HMG-CoA 制备的标准曲线进行对比。活性定义为 U/mg 蛋白。1 单位活性=生成 1nmol HMG-CoA/分钟。HMG-CoA 合成酶的检测下限约为 1U/mg。

HMG-CoA 还原酶试验 粗提物在含有 50mM KCl, 1mM EDTA 和蛋白酶抑制剂混合物(protease inhibitor cocktail)(Sigma Chemical Co., St.Louis, MO, USA, catalog#P-2714)的 25mM 磷酸钾缓冲液(pH7.2)中制备。根据 Takahashi 等人[J.Bacteriol. 181:1256-1263(1999)]的方法进行试验。该方法在

- 340nm 下, 利用分光光度计测量 HMG-CoA 依赖性 NADPH 的氧化。反应混合物含有 25mM 磷酸钾缓冲液(pH7.2), 50mM KCl, 1mM EDTA, 5mM 二硫苏糖醇, 0.3mM NADPH, 0.3mM R,S-HMG-CoA 和粗提物。反应在 30℃ 下进行, 加入 HMG-CoA 起始反应。使用 SpectraMAX Plus 板式读数器
- 5 (Molecular Devices Corp., Sunnyvale, CA, USA)和石英微量滴定板(也可使用任何标准分光光度计), 在 340nm 下监测 HMG-CoA 依赖性 NADPH 氧化。使用 NADPH 制备的标准曲线, 计算酶活(表达为 U/mg 蛋白)(1 单位活性=氧化 NADPH 1μmol/分钟)。HMG-CoA 还原酶活性的检测下限为 0.03U/mg。
- 甲羟戊酸激酶, 磷酸甲羟戊酸激酶和甲羟戊酸二磷酸脱羧酶试验 制备底
- 10 物, 用于甲羟戊酸激酶、磷酸甲羟戊酸激酶和甲羟戊酸二磷酸脱羧酶的试验方法已在 Popják[Methods Enzymol. 15:393-425(1969)]中详细描述。所有的试验中, 1 单位酶活均定义为生成的 1μmol 产物/分钟。除了分光光度计和放射性色谱试验, 还可使用其他的方法, 如使用 HPLC 来分离反应底物和产物。甲羟戊酸激酶、磷酸甲羟戊酸激酶和甲羟戊酸二磷酸脱羧酶的检测
- 15 下限通常约为 0.001U/mg 蛋白。
- IPP 异构酶试验 粗提物被制备到 50mM Tris-HCl 缓冲液(pH7.5)中。使用 Spurgeon 等人[Arch.Biochem.Biophys. 230:445-545(1984)]的方法进行试验。该反析是基于 IPP 和 DMAPP 之间酸不稳定性的差异。反应混合物(终体积 0.1ml)含有 50mM Tris-HCl 缓冲液(pH7.5), 2mM 二硫苏糖醇, 5mM MgCl₂,
- 20 20μM [1-¹⁴C]-IPP 和粗提物。在 30℃ 下反应 15 分钟, 加入 0.3ml 浓缩的 HCl: 甲醇(4: 1)混合物终止反应, 并额外在 37℃ 下孵育 20min。加入己烷(0.9ml), 混合试管(使用涡旋混合器, 10 秒, 4 次)。离心(21000xg, 5 分钟)后, 将 0.6ml 的己烷层转移到闪烁瓶内, 加入闪烁液, 计数放射性。活性表达为 U/mg 蛋白。1 单位活性=1pmol 被掺入到酸不稳定性产物的[1-¹⁴C]-IPP/分钟)。IPP
- 25 异构酶活性的检测下限为 1U/mg。
- FPP 合成酶 粗提物被制备到 50mM Tris-HCl 缓冲液(pH8.0)中。FPP 合成酶试验方法类似于上述的 IPP 异构酶试验, 是基于 IPP 和 FPP 之间酸不稳定性的差异[Spurgeon 等, 同上]。反应混合物(终体积 0.1ml)含有 50mM Tris-HCl 缓冲液(pH 8.0), 2mM 二硫苏糖醇, 5mM MgCl₂, 20μM [1-¹⁴C]-IPP, 25μM
- 30 GPP(geranyl pyrophosphate)和粗提物。在 30℃ 下反应 15 分钟, 加入 0.3ml 浓缩的 HCl: 甲醇(4: 1)混合物终止反应, 并额外在 37℃ 下孵育 20min。加入

己烷(0.9ml), 混合试管(使用涡旋混合器, 10 秒, 4 次)。离心(21000xg, 5 分钟)后, 将 0.6ml 的己烷层转移到闪烁管内, 加入闪烁液, 计数放射性。

酶活性单位和检测下限, 与上述 IPP 异构酶的定义相同。当较高的 IPP 异构酶活性干扰了 FPP 合成酶活性的测定时, 可将粗提物在 5mM 碘乙酰胺的

5 存在下预温育 5 分钟, 以抑制 IPP 异构酶的活性。

GGPP 合成酶试验 粗提物被制备到含有 2mM 二硫苏糖醇的 50mM Tris-HCl(pH8.0)中。根据 Kuzuguchi 等人[J.Biol.Chem. 274:5888-5894 (1999)]的方法对 GGPP 合成酶进行分析。该试验是基于与上述 FPP 合成酶相同的原理。反应混合物(终体积 0.1ml)含有 50mM Tris-HCl 缓冲液(pH 8.0), 2mM 二硫苏糖醇, 5mM MgCl₂, 20μM [1-¹⁴C]-IPP, 25μM FPP 和粗提物。所有的反应条件和随后用于闪烁计数的样品处理均与上述 FPP 合成酶的相同。还可使用上面所述的碘乙酰胺处理抽提物, 以抑制 IPP 异构酶的活性。酶活单位和检测下限与上述 IPP 异构酶的定义相同。

乙酰乙酰-CoA 还原酶试验 粗提物被制备到含有 50mM KCl 和 5mM 二硫苏糖醇的 50mM Tris-HCl(pH8.0) 中。根据 Chohan 和 Copeland [Appl.Environ.Microbiol. 64: 2859-2863(1998)]的方法进行乙酰乙酰-CoA 还原酶试验。该试验在 340nm 下, 利用分光光度计测定乙酰乙酰-CoA 依赖性的 NADPH 氧化。反应混合物(1ml)含有 50mM Tris-HCl 缓冲液(pH 8.5), 15mM MgCl₂, 250μM NADPH, 和 100μm 乙酰乙酰-CoA。反应在 30℃, 在石英比色杯中进行, 加入乙酰乙酰-CoA 起始反应。使用 NADPH 制备的标准曲线, 计算酶活(表达为 U/mg 蛋白)(1 单位活性=1μmol 被氧化的 NADPH/分钟)。乙酰乙酰-CoA 还原酶活性的检测下限约为 0.01U/mg。

实施例 2: 从分类学角度将黄杆菌属菌种重新分类为副球菌

该实施例说明了将早期命名为黄杆菌属的菌株 R1512(ATCC 21588)从分类学角度重新分类为副球菌种菌株 R-1512(ATCC 21588)。全面的基因组和生化/生理分析由 Belgian Coordinated Collections of Microorganisms/Laboratorium voor Microbiologie, Universiteit Gent (BCCMTM/LMG)进行, 使用目前公知的本领域常规技术作为细菌分类科学标准。除副球菌种菌株 R-1512 之外, 其他几种属于副球菌属的细菌也包括在该研究中(归纳在表 2 中)。

表 2.用于分类学研究的细菌

细菌	菌株命名	参考来源
副球菌种	R-1512(ATCC 21588)	ATCC(环境分离)US 3,891,504
副球菌种	R1534	R-1512 的突变株; US 6,087,152
副球菌种	R114	R-1512 的突变株; 本文
副球菌种	R-1506	环境分离; 本文
副球菌种	MBIC3024	H. Kasai, Kamaishi Institute, Japan
副球菌种	MBIC3966	H. Kasai, Kamaishi Institute, Japan
副球菌种	MBIC4017	H. Kasai, Kamaishi Institute, Japan
副球菌种	MBIC4020	H. Kasai, Kamaishi Institute, Japan
<i>P.marcusii</i>	DSM 11574 ^T	Harker 等; 同上
<i>P. carotinifaciens</i>	E-396 ^T	Tsubokura 等; 同上
<i>P. solventivorans</i>	DSM 6637 ^T	Siller et. al., Int. J. Syst. Bacteriol. 46:1125-1130(1996)

菌株 R1534 和 R114 是通过经典诱变技术由菌株 R-1512 突变而形成，利用改进的玉米黄质产量而将其筛选出来。初级筛选是选出形成色彩强度最大的集落。次级筛选是根据实施例 1 所述 HPLC 方法在液体培养基中进行。菌株 R-1506 是从提供菌株 R-1512 的相同起始筛选中从环境微生物中独立分离而获得。菌株 MBIC3024, MBIC3966, MBIC4017, MBIC4020 根据其 16S rDNA 基因的核苷酸序列(DNA 序列存储在公共 EMBL 数据库中，参见表 5)，均被鉴定为副球菌属的成员。最近，*paracoccus marcusii* DSM 11574^T 和 *Paracoccus carotinifaciens* E-396^T 被认定为属于产类胡萝卜素细菌类型 [Harker 等，Int.J.Syst.Bacteriol. 48: 543-548(1998); Tsubokura 等，Int.J.Syst.Bacteriol. 49:277-282(1999)]。Paracoccus solventivorans DSM 6637^T 作为“对照”菌株，该菌株属于副球菌属，但是与所用的其它细菌亲缘关系较远。

初步的试验产生了下述结果。本文所述的各种方法能够清楚地鉴定有机体相似性的分类学关系或相关程度。Van Damme 等人 Microbiological Reviews 60:407-438(1996)和 Jassen 等人 Microbiology 142: 1881-1893(1996)中详细地描述并比较了这些方法及其在描述细菌门类上的应用。

- (1) 对菌株 R1534 和 R114 细胞膜的脂肪酸分析显示：这两个菌株高度相似,说明了这两个菌株与 *Paracoccus denitrificans* 和 *Rhodobacter capsulatus* 的分类学相关度。
- (2) 细胞蛋白的单向凝胶电泳显示在 R1534 和 R114 之间存在极高的相似度(即，种内水平相关度)，但是图谱均未表明这些菌株属于 *R.capsulatus* 或 *P.denitrificans*。
- (3) 菌株 R1534 与 *R.capsulatus* LMG2962^T 和 *P.denitrificans* LMG4218^T 之间的 DNA:DNA 杂交证实了菌株 R1534 既不是 *R.capsulatus* 也不是 *P.denitrificans* 。
- (4) 菌株 R1534 和 R114 的 16S rDNA 基因测序显示这些有机体属于副球菌属，但是它们代表了新的种。最高程度的序列相似性是副球菌种菌株 MBIC3966, MBIC4020 和 MBIC3024 的 16S rDNA 基因。
- (5) 使用扩增片段长度多态性(AFLPTM)进行的菌株 R1534 和 R114 DNA 指纹图谱显示这两个菌株基因组 DNA 之间综合相似性极高，表明种内相关度(即 AFLPTM可以在相同种内的两个成员间有所差异)。

下述部分是本发明对副球菌种菌株 R-1512(及其突变株衍生的 R1534 和 R114)所进行的大量分类学研究的结果和结论。

- 16S rDNA 测序和系统发育研究 表 2 所述的细菌培养在 LMG 培养基 185((TSA) BBL 11758 中，如果需要的话，可在培养基中补充 1.5%的 Difco Bacto 琼脂)。根据 Niemann 等人[J.Appl.Microbiol. 82:477-484(1997)]的方法，制备基因组 DNA。利用聚合酶链式反应(PCR)，使用表 3 所示引物，从菌株 R-1512, R1534 和 R-1506 的基因组 DNA 中扩增编码 16S rDNA 的基因。

表 3. PCR 扩增副球菌种菌株 R-1512, R1534 和 R-1506 中编码 16S rDNA 的 DNA 所用引物

引物名 ^a	序列 (5' → 3')	SEQ ID NO	位置 ^b
16F27	AGA GTT TGA TCC TGG CTC AG	SEQ ID NO:1	8-27
16F38	CTG GCT CAG GAC/T GAA CGC TG	SEQ ID NO:2	19-38
16R1522	AAG GAG GTG ATC CAG CCG CA	SEQ ID NO:3	1541-1522

- ^aF，正向引物；R，反向引物。正向引物 16F27(又称：PA)用于 R1534 和 R-1506，而正向引物 16F38(又称：ARI C/T)用于菌株 R-1512 和 R114。反向引物 16R1522(又称：PH)用于所有菌株。

^b 杂交位置指大肠杆菌 16S rDNA 基因序列编号。

使用 Qiaquick PCR Purification Kit(Qiagen GmbH, Hilden, Germany)纯化 PCR 扩增的 DNA。使用 Applied Biosystems, Inc.337 测序仪和操作手册 (Perkin-Elmer, Applied Biosystems Division, Foster City, CA, USA)所述方法，

5 使用 “ABI PRISMTM Big DyeTM Terminator Cycle Sequencing Ready Reaction Kit(用 AmpliTaq® DNA 聚合酶 Fs)”，进行完整的测序。DNA 测序所用引物如表 4 所示。

表 4.用于测序副球菌种菌株 R-1512, R1534, R114 和 R-1506 中编码 16S rDNA 基因的 PCR 扩增片段的引物

引物名 ^a	序列 (5' →3')	SEQ ID NO	位置 ^b
又名			
16F358/*γ	CTC CTA CGG GAG GCA GCA GT	SEQ ID NO:4	339-358
16F536/*PD	CAG CAG CCG CGG TAA TAC	SEQ ID NO:5	519-536
16F926/*O	AAC TCA AAG GAA TTG ACG G	SEQ ID NO:6	908-926
16F1112/*3	AGT CCC GCA ACG AGC GCA AC	SEQ ID NO:7	1093-1112
16F1241/*R	GCT ACA CAC GTG CTA CAA TG	SEQ ID NO:8	1222-1241
16R339/γ	ACT GCT GCC TCC CGT AGG AG	SEQ ID NO:9	358-339
16R519/PD	GTA TTA CCG CGG CTG CTG	SEQ ID NO:10	536-519
16R1093/3	GTT GCG CTC GTT GCG GGA CT	SEQ ID NO:11	1112-1093

10 ^aF，正向引物；R，反向引物。

^b 杂交位置指大肠杆菌 16S rDNA 基因序列编号。

使用 5 个正向引物和 3 个反向引物制备部分重叠的序列，确保得到高度可信的组合序列数据。将共有序列(来自 R-1512, R1534, R114 和 R-1506)包含在小核糖体亚基序列的排列之后(所述小核糖体亚基序列由国际核苷酸

15 序列文库 EMBL 收集)，使用 AutoAssembler 程序(Perkin-Elmer, Applied Biosystems Division, Foster City, CA, USA)进行序列组合。使用 GeneComparTM 软件包(v2.0, Applied Maths B.V.B.A., Kortrijk, Belgium)进行系统发育分析。使用 100%的 open gap penalty 和 0%的 unit gap penalty，成对地计算该排列。使用 0%的 gap penalty，去除未知的碱基后，通过同源性

20 计算形成相似性基质(similarity matrix)。使用邻接法(neighbor-joining method)构建结果树。

副球菌菌株 R-1512 的 16S rDNA 基因核苷酸序列被定义为 SEQ ID NO: 12。R-1512 及其亲缘关系最近的菌株之间的距离基质(distance matrix), 作为 16S rDNA 序列相似性的百分数, 如表 5 所示。菌株 R-1512 及其突变株衍生物 R1534 和 R114 的序列相同。R1506 的序列与后面的菌株的序列仅一个核苷酸不同。这说明菌株 R-1512 和 R-1506 在系统发育学上高度相关, 似乎属于相同的种(由 DNA:DNA 杂交证实, 见下文)。R-1512 和 R-1506 序列与那些可得的 EMBL 文库中序列所进行的对比, 将 R-1512 和 R-1506 定位在副球菌属种。但是, 观察到的、能够确凿地从所有目前的分类学角度说明副球菌种的序列相似性是小于 97%, 这个数值通常作为物种水平可能相关的极限值 [Stackebrandt 和 Goebel, Int.J.Syst.Bacteriol. 44: 846-849(1994)]。这说明 R-1512 菌株(及其突变衍生物)和 R-1506 属于一个或两个新的副球菌种。观察到四个未命名的副球菌菌株和菌株 R-1512, R1534, R114 和 R-1506 之间存在大于 97%的序列相似性(在物种水平上可能的相关性显著), 说明一个或多个未命名(MBIC)的菌株可能与菌株 R-1512 和 R-1506 在种水平上相关。根据簇分析(描述副球菌种菌株 R-1512, R1534, R114, R-1506, MBIC3966 以及副球菌属其它成员之间系统发育相关程度的系统发育树), 菌株 R-1512, R1534, R114 和 R-1506 和四个未命名的副球菌菌株(MBIC3024, MBIC3966, MBIC4017 和 MBIC4020)被选择用于分析物种相关度的 DNA:DNA 杂交试验以分析物种的相关性。

表 5.距离基质(distance matrix), 作为副球菌种菌株 R-1512 及其亲缘关系最近的菌株之间 16S rDNA 序列相似性的百分数

菌株 ^a	EMBL 登录号	%相似性
R-1512	-	100
R1534	-	100
R114	-	100
R-1506	-	99.9
副球菌种 MBIC3966	AB018688	100
副球菌种 MBIC3024	AB008115	98.2

5	副球菌种 MBIC4020	AB025191	98.1
	副球菌种 MBIC4036	AB025192	97.0
	副球菌种 MBIC4017	AB025188	96.9
	副球菌种 MBIC4019	AB025190	96.8
	副球菌种 MBIC4018	AB025189	96.4
10	<i>Paracoccus marcusii</i> DSM 11574 ^T	Y12703	96.2
	<i>Paracoccus carotinifaciens</i> E-396 ^T	AB006899	96.1
	<i>Paracoccus solventivorans</i> DSM 6637 ^T	Y07705	95.4
	<i>Paracoccus thiocyanaticus</i> THIO11 ^T	D32242	95.3
	<i>Paracoccus aminophilus</i> JCM 7686 ^T	D32239	95.1
15	<i>Paracoccus alcaliphilus</i> JCM 7364 ^T	D32238	95.0
	<i>Paracoccus pantotrophicus</i> ATCC 35512 ^T	Y16933	95.0
	<i>Paracoccus denitrificans</i> ATCC 17741 ^T	Y16927	94.8
	<i>Paracoccus versutus</i> IAM 12814 ^T	D32243	94.7
	<i>Paracoccus kocurii</i> JCM 7684 ^T	D32241	94.6
20	<i>Paracoccus aminovorans</i> JCM 7685 ^T	D32240	94.4
	<i>Paracoccus alkenifer</i> A901/1 ^T	Y13827	94.3
	<i>Rhodobacter capsulatus</i> ATCC 11166 ^T	D16428	92.9

^a 菌株类型后接 a^T

25 DNA:DNA 杂交及 G+C 含量测定 用 LMG 培养基 185 培养表 5 所述细菌。根据 Wilson 的方法[In Ausabel 等, (eds), Current Protocols in Molecular Biology, Greene Publishing and Wiley Interscience, New York, 2.4.1-2.4.5(1987)], 制备基因组 DNA。根据 Mesbach 等人[Int.J.Syst.Bacteriol. 39:159-167(1989)], 由 Logan 等人[Int.J.Syst.Bacteriol. 50: 1741-1753(2000)]进行的修改, 利用 HPLC 进行 DNA 的 G+C 含量测定。报道的数值是对相同 DNA 样品进行测量得到的平均值。使用 De Ley 等人描述

30 [Eur.J.Biochme.12:133-142(1970)]的起始复性率方法[initial renaturation rate

method], 进行 DNA:DNA 杂交。杂交温度为 81.5℃。Vauterin 等人 [Int.J.Syst.Bacteriol. 45:472-489(1995)]曾报道了该方法+/-5.8%的平均偏差。该细菌 DNA 的 G+C 含量以及 DNA 杂交试验结果归纳在表 6 中。

5 表 6.副球菌种各菌株的 DNA G+C 含量(mol%), 以及各菌株间 DNA 同源性百分比

10

菌株:	%G+C	%DNA 同源性							
R-1512	67.6	100							
R1534	67.7	96	100						
R114	67.5	100	97	100					
R-1506	67.5	94	90	88	100				
MBIC3024	65.4	31	nd ^a	nd	31	100			
MBIC3966	66.9	93	nd	nd	88	32	100		
MBIC4017	67.2	32	nd	nd	31	24	24	100	
MBIC4020	68.4	27	nd	nd	25	25	23	34	100

15

^a 未测定

菌株 R-1512, R1534, R114, R-1506 和 MBIC3966 之间 DNA 同源性大于 70%(通常接受用于物种描述的极限值[Wayne 等, Int.J.Syst.Bacteriol. 37:463-464(1987)]), 因而属于副球菌属中相同的种。这 5 个菌株的 G+C 含量在 66.9%到 67.7%之间, 差别小于 1%, 表征着精确定义的物种。另一方面, 菌株 MBIC3024, MBIC4017 及 MBIC4020 和 R-1512, R1534, R114, R-1506 及 MBIC3966 之间较低的 DNA 同源性说明, MBIC3024, MBIC4017 及 MBIC4020 属于副球菌属内不同的基因种。

使用 AFLPTM进行的 DNA 指纹 AFLPTM是基于 PCR 技术,通过选择性扩增和选择性限制性片段显色 [Vos 等, Nucleic Acids Research 23:4407-4414(1995);Janssen 等, 同上], 对整个染色体 DNA 进行指纹印迹。该试验中, 比较副球菌种菌株 R-1512, R1534, R114, R-1506, MBIC3966 和 Paracoccus marcusii DSM 11574^T, 评价其种内相关度。在 LMG 培养基 185 中培养这些细菌。根据 Wilson 的方法(同上)制备这些细菌的中各个基因组

DNA。用两种限制性酶——4-碱基酶切(cutter)酶和 6-碱基酶切酶，消化纯化的 DNA。通过这种方法，可以获得带有两种不同末端的、大小适于进行 PCR 的有限数目片段。含有一个相容性末端的接头(15-20bp 的小双链 DNA 分子)与适当的限制性片段“粘性”末端相连。两个接头都是限制性半位点特异性，具有不同的序列。这些接头作为 PCR 引物的结合位点。本文中，所用的限制性酶是 *ApaI*(6 切的酶，识别序列 GGGCC/C)和 *TaqI*(4 切的酶，识别序列 T/GCA)。与用限制性酶酶切后形成的粘性末端相连的接头序列如表 7 所示(SEQ ID NO: 13-22)。使用 PCR 进行限制性片段的选择性扩增。该 PCR 引物特异性地与限制性片段的接头末端退火。由于引物的 3'末端含有所谓的“选择性碱基”，这些碱基能够在限制性位点之外延伸到片段中，只有与限制性位点临近的、具备适当互补序列的限制性片段才能被扩增。所用的 6 个 PCR 引物序列组合也归纳在表 7 中。

表 7.用于 AFLP™分析的接头和 PCR 引物

	序列	SEQ ID NO
相应于限制性酶 <i>ApaI</i> 的接头		
接头 93A03	5'-TCGTAGACTGCGTACAGGCC-3'	SEQ ID NO:13
接头 93A04	3'-CATCTGACGCATGT-5'	SEQ ID NO:14
相应于限制性酶 <i>TaqI</i> 的接头		
接头 94A01	5'-GACGATGAGTCCTGAC-3'	SEQ ID NO:15
接头 94A02	3'-TACTCAGGACTGGC-5'	SEQ ID NO:16
	序列	SEQ ID NO
引物组合 1 (PCA)		
A01	5'GACTGCGTACAGGCCCA <u>3</u> '	SEQ ID NO:17
T01	5' CGATGAGTCCTGACCGAA <u>3</u> '	SEQ ID NO:18
引物组合 2 (PCB)		
A01	5' GACTGCGTACAGGCCCA <u>3</u> '	SEQ ID NO:17
T02	5' CGATGAGTCCTGACCGAC <u>3</u> '	SEQ ID NO:19
引物组合 3 (PCD)		
A02	5' GACTGCGTACAGGCCCC <u>3</u> '	SEQ ID NO:20
T01	5' CGATGAGTCCTGACCGAA <u>3</u> '	SEQ ID NO:18

引物组合 4 (PCI)		
A03	5' GACTGCGTACAGGCCCG 3'	SEQ ID NO:21
T03	5' CGATGAGTCCTGACCGAG 3'	SEQ ID NO:22
引物组合 5 (PCG)		
A03	5' GACTGCGTACAGGCCCG 3'	SEQ ID NO:21
T01	5' CGATGAGTCCTGACCGAA 3'	SEQ ID NO:18
引物组合 6 (PCH)		
A03	5' GACTGCGTACAGGCCCG 3'	SEQ ID NO:21
T02	5' CGATGAGTCCTGACCGAC 3'	SEQ ID NO:19

扩增后,使用 DNA 测序仪(ABI377),在高分辨率的聚丙烯酰胺凝胶上,根据各 PCR 产物的长度将其分离。由于相应引物 5'末端标记了 ³²P,含有限制性半位点(该位点由 6-bp cutter 形成)特异性接头的片段可通过放射自显影而进行目测。扫描电泳图案,用 GelComparTM 4.2 软件(Applied Maths, B.V.B.A., Kortrijk, Belgium)进行数字分析,使用 Person 曲线匹配系数(curve matching coefficient)和未加权的对群连接(unweighted pair group linking)分组[分组方法参见 Sneath 和 Sokal, In: Numerical Taxonomy. Freeman & Son, San Francisco (1973)]。

在所有的 6 个引物组合中(PC A-H, 表 7)副球菌属菌株 R-1512, 1534 和 R114 的 DNA 指纹即使并非相同,也是高度近似。观察到微小区别的情况下,未估算重现性。曾预期这三种菌株的相似性或相同性会很高,因为 R1534 和 R114 来源于 R-1512。运用所有的引物组合,菌株 R-1512, 1534 和 R114 能够与菌株 R-1506 和 MBIC3966 清晰地区分开,后两种菌株均属于新的副球菌种。但是,指纹图谱并没有清楚地表明菌株 R-1512, R1534 和 R114 与 R-1506 或 MBIC3966 的亲缘关系更为密切。在所用的条件下,新物种聚簇(cluster)的五个菌株以大约 58%(该值是六个 AFLPTM 试验(6 个引物组合)中新物种分支点 6 个数值的平均数)相似性的平均水平进行分组,该组可以清楚地与 *Paracoccus marcusii* DSM 11574^T 的图谱区分开,该典型菌株是系统发育学相关的产类胡萝卜素的副球菌种。*Paracoccus marcusii* DSM 11574^T 6 个分支点与新物种的平均相似性值约为 11%。

脂肪酸分析 比较副球菌种菌株 R-1512, R1534, R114, R-1506,

MBIC3966 与典型菌株 *P. marcusii* DSM 11574^T, *P. carotinifaciens* E-396^T 和 *P. solventivorans* DSM 6637^T 的细胞膜脂肪酸组成。在 28℃, 用 LMG 培养基 185 培养细菌。使用商业系统 MIDI(Microbial Identification System, Inc., DE, USA), 通过气相色谱测定脂肪酸组成。根据 MIDI 系统的建议进行脂肪酸的提取和分析。表 8 总结了所有受测菌株的结果。计算五个新副球菌种 (R-1512, R1534, R114, R-1506, MBIC3966) 菌株的平均图谱。所有 8 个有机体显示出可比较的细胞膜脂肪酸组成, 18: 1w7c 是主要化合物。在新的副球菌种和三种典型菌株之间仅观察到微小的区别。

应用碳源进行生长 为测试碳源的有氧应用, 使用含有 95 种底物的 BIOLOG-SF-N Microplate 微量滴定板(Biolog Inc., Hayward, CA, USA), 只有 E6 孔的底物是 D,L-乳酸甲酯, 而不是通常所用的 D,L-乳酸钠盐。在 28℃ 下, 表 9 所鉴定的每种菌株细胞均在 LMG 培养基 12(Marine Agar, Difco 0979)中培养 24 小时。用无菌水制备密度等于 0.5McFarland 单位的细胞悬浮液。从该悬浮液中转移 18 滴液体到 21ml 的 AUX 培养基(API 20NE, bioMérieux, France)中, 并轻轻混合。将 0.1 毫升悬浮液转移到 BIOLOG MicroPlate 的各孔中, 30℃ 孵育该板。在 48 小时和 6 天后目测检查各孔的生长情况。同样, 在第 6 天, 使用 BIOLOG 板式读数器, 通过阅读微量滴定板证实目测的分数。

BIOLOG 分析的结果如表 9。与没有底物的对照相比, 不断增加的混浊度证实了生长(阳性反应)。在良好生长(+), 微弱生长(±)和无生长(-)之间进行区别。括号内的结果是 6 天之后获得的, 如果该值与 48 小时后获得的值不相符。问号表明第 6 天的结果不明确。在 95 种受测的碳源中, 包括新的副球菌种(R-1512, R1534, R114, R-1506, MBIC3966)在内的所有五种菌株可以利用其中的 12 种用于生长, 47 种不能被利用。这 5 个菌株对剩下的 36 种底物表现出不同的生长反应。通过对 7 种碳源(福寿糖醇, i-赤藓糖醇, 龙胆二糖, β-甲基葡萄糖苷, D-山梨糖醇, 木糖醇和奎尼酸)的利用能力, 新的副球菌种可以与两种其它的产类胡萝卜素细菌(*P. marcusii* DSM 11574^T 和 *P. carotinifaciens* E-396^T)相区分。新的副球菌种 5 种菌株都可利用的两种碳源(L-天冬酰胺和 L-天冬氨酸)不能被 *P. marcusii* DSM 11574^T 用于生长。

表 8. 副球菌种菌株 R-1512, R1534, R114, R-1506, MBIC3966 与其它副球菌种的三种典型菌株, 即 *P. marcusii* DSM 11574^T, *P. carotinifaciens* E-396

^T和 *P.solventivorans* DSM 6637^T 的细胞膜脂肪酸组成

5		平均%	%		
	名称	R-1512, R1534, R114, R-1506 和 MBIC3966	DSM 11574 ^T	E-396 ^T	DSM 6637 ^T
	10:0 3OH	4.9 ± 1.1	6.2	3.4	3.6
	未命名 11.799	3.6 ± 0.5	4.9	2.8	3.0
	未命名 15.275	1.5 ± 0.3	2.9	1.1	ND ^a
10	16:0	0.3 ± 0.2	ND	0.3	0.7
	17:1 w8c	ND	ND	0.6	0.8
	17:0	0.1 ± 0.1	ND	0.3	1.3
	18:1 w7c	80.5 ± 1.8	80.3	84.0	79.0
		平均%	%		
15	名称	R-1512, R1534, R114, R-1506 和 MBIC3966	DSM 11574 ^T	E-396 ^T	DSM 6637 ^T
	18:0	3.6 ± 0.4	2.6	5.2	6.6
	18:0 3OH	0.6 ± 0.4	ND	ND	ND
	19:0	ND	ND	ND	0.7
	20:1 w7c	0.8 ± 0.2	ND	0.2	2.0
20	总计特点 2	2.7 ± 0.4	3.0	2.1	2.6
	总计特点 3	0.7 ± 0.5	ND	0.2	ND
	合计	99.3	99.9	100.2	100.3

^aND, 未检测。

生化试验 使用 API 20NE 条带(strip)(bioMérieux, France)测试选出的生化特点。在 28℃下, 表 10 所鉴定的每种细菌菌株细胞培养于 LMG 培养基 12 中 24 小时。根据制造商的说明书制备细胞悬浮液及接种条带。条带(strip)在 28℃下接种, 在 24 和 48 小时后测定结果。结果总结在表 10 中。受测试的 9 个特点中, 只有一个(尿素酶活性)特点在新副球菌中的 5 个菌株中有所区别。这 9 个特点在新副球菌种和 *P. marcusii* DSM 11574^T, *P.carotinifaciens* E-396^T之间没有差异。

表 9.副球菌种菌株利用碳源用于生长的情况

	R1512	R1534	R114	R1506	MBIC 3966	DSM 11574 ^T	E-396 ^T	DSM 6637 ^T
α-环糊精	-	-	-	-	-	-	-	-
糊精	-	-	-	-	-	-	-	-(±)
糖原	-	-	-	-	-	-	-	-
Tween 40	-	-	-	-	-	-	-(?)	-
Tween 80	-	-	-	-	-	-	-	-
GalNAc	-	-	-	-	-	-	-	-
	R1512	R1534	R114	R1506	MBIC 3966	DSM 11574 ^T	E-396 ^T	DSM 6637 ^T
GulNAc	-	-	-	-	-	-	-	-(?)
福寿草醇	-	-	-	-	-	+	+	-
L-阿拉伯糖	-	-	-	-	-	+	-	+
D-阿拉伯糖醇	+	+	+	+	±(+)	+	+	-
纤维二糖	±(+)	±(+)	-(?)	-(+)	-(±)	+	+	-(+)
i-赤藓糖醇	-	-	-	-	-	+	+	-
D-果糖	+	+	+	+	-	+	+	+
L-岩藻糖	-	-	-	-	-	-	+	-
D-半乳糖	+	+	+	±(+)	±(+)	+	+	-(±)
龙胆二糖	-	-	-	-	-	+	+	-(±)
α-D-葡萄糖	+	+	±(+)	-(+)	+	+	±(+)	+
m-肌醇	+	+	+	-(+)	-(+)	+	-(±)	-
α-乳糖	+	±(+)	-(+)	-(+)	-(+)	+	+	±(+)
Lactulose	-(±)	-(+)	-(+)	-(±)	-	+	+	-(+)
麦芽糖	+	+	-(+)	-(+)	-(±)	+	+	-(+)
D-甘露醇	+	+	+	+	-(+)	+	+	-(±)

D-甘露糖	+	+	+	+	-(±)	+	+	-(+)
D-蜜二糖	+	+	+	-(+)	-(+)	+	+	-(?)
β-Methylgluc	-	-	-	-	-	+	+	+
D-阿洛酮糖	-(+)	±(+)	±(+)	-	-(+)	-	±	-
D-棉子糖	-	-	-	-	-	-(+)	+	-
L-鼠李糖	-	-	-	-	-	-	-	-(?)
D-山梨醇	-	-	-	-	-	+	+	-
蔗糖	+	+	±(+)	-(+)	-	+	+	+
D-海藻糖	+	+	-(+)	-(+)	-(+)	+	+	+
	R1512	R1534	R114	R1506	MBIC 3966	DSM 11574 ^T	E-396 ^T	DSM 6637 ^T
松二糖	-(+)	-(+)	-	-	-	+	+	+
木糖醇	-	-	-	-	-	+	+	-
丙酮酸甲酯	±	-	±	-(?)	±	-	+	-(±)
MMSucc	±(+)	±	-(±)	-(+)	-(±)	-(+)	+	-
乙酸	-	-	±	-	-	-	-	+
顺-乌头酸	-	±	±	-	-	±	-	-
柠檬酸	-	±	±	-	-	±	-	-
甲酸	-	-	-	-	-	-	-	-
GalAlactone	-(±)	-(±)	-(±)	-	-	-	-(±)	-(?)
GalacturonicA	-	-	-	-	-	-(+)	-(±)	-
D-葡糖酸	+	+	+	-(±)	-(±)	+	+	+
Glucosaminic A	-	-	-	-	-	-	-	-
GlucuronicA	±	+	+	-(±)	-	±(+)	-	-
AHBA	-(±)	-	-(±)	-	-(+)	-	-	-
BHBA	+	+	+	-(±)	±	-(+)	+	+

GHBA	-	-	-	-	-	-	-	-
PHPAA	-	-	-	-	-	-	-	-(+)
衣康酸	-	-	-	-	-	-	-	-
AKBA	-	-	-	-	-	-	-	-(±)
AKGA	-	-	-	-(±)	-(?)	-(±)	-(+)	-(±)
AKVA	-	-	-	-	-	-	-	-
LAME	-	-	-	-	-	-	-	-
丙二酸	-	-	-	-	-	-	-	-
丙酸	-	±	±	-	-	±	+	+
	R1512	R1534	R114	R1506	MBIC 3966	DSM 11574 ^T	E-396 ^T	DSM 6637 ^T
奎尼酸	-	-	-	-	-	+	+	-
SaccA	-(+)	±	-	-(±)	-	-	-	-
癸二酸	-(+)	-(+)	-(+)	-(+)	-(±)	-	-(+)	-
琥珀酸	-	-	-	-	-	-(+)	±	-(?)
BromosuccA	-	-	-	-	-	±	-	-
琥珀酰胺酸	-	-	-	-	-	-(+)	-(+)	-
葡萄糖酰胺	-	-	-	-	-(±)	-	-	-
Alaninamide	-	-	-(+)	-	-	-	-	-
D-丙胺酸	-	-	-(+)	-	-	-	-	-
L-丙胺酸	+	+	+	+	-	-(+)	+	+
L-丙氨酰-甘氨酸	-(+)	-	-(+)	-	-	-	-(+)	-(?)
L-天冬酰胺	+	+	±(+)	+	±(+)	-	+	-
L-天冬氨酸	+	+	±(+)	-(+)	-(+)	-	+	-
L-谷氨酸	+	+	+	+	±(+)	-(+)	+	-(+)
GAA	-	-	-	-(±)	-	-	-	-

GGA	-(?)	-	-	-(?)	-	-	-(±)	-
L-组氨酸	-	-	-	-	-	-(?)	-	+
HydPro	-	-	-	-	-	-	-	+
L-亮氨酸	-(±)	-(+)	-(+)	-(+)	-	-(+)	-(?)	-(+)
L-鸟氨酸	-	-(+)	±(+)	-(±)	-	-	-(+)	-
L-苯丙氨酸	-	-	-	-	-	-	-	-
L-脯氨酸	+	+	+	+	-	-(+)	+	+
PyroGluA	+	+	+	+	±(+)	-(+)	+	-
D-丝氨酸	-	-	-	-	-	-	-	-
L-丝氨酸	±	±(+)	-(+)	- (±)	- (+)	- (+)	- (+)	+
	R1512	R1534	R114	R1506	MBIC 3966	DSM 11574 ^T	E-396 ^T	DSM 6637 ^T
L-苏氨酸	-	-	-	-	-	-(+)	-	-
D,L-肉毒碱	-	-	-	-	-	-	-	-
GABA	-	-	-	-	-(+)	-	-(+)	-(+)
尿刊酸	-	-	-	-	-	-	-	-(+)
次黄苷	-	-(±)	-	-	-	-(±)	-(+)	-(+)
尿苷	-	-	-	-	-	-(±)	-(+)	-
胸苷	-	-	-	-	-	-(±)	-(±)	-
PEA	-	-	-	-	-	-	-	-
Putresceine	-	-	-	-	-	-	-	-
2-氨基乙醇	-	-	-	-	-	-	-	-
2,3-丁二醇	-	-	-	-	-	-	-	-
甘油	+	+	-(+)	+	-	+	+	-
GlycP	-	-	-	-	-(±)	-	-	-
Gluc-1-P	-	-	-	-	-	-	-(±)	-

Gluc-6-P	-	-	-	-	-	-	-	-
----------	---	---	---	---	---	---	---	---

- GalNAc: N-乙酰-D-半乳糖胺; GluNAc: N-乙酰基-D-葡萄糖胺;
 β -methylguc: β -甲基葡萄糖苷; MMSucc: 单-丁二酸甲酯; GalAlactone: D-半乳糖酸内脂; GalacturonicA: D-半乳糖醛酸; GlucosaminicA: D-葡萄糖甲酸 (Glucosaminic acid); GlucuroinicA: D-葡萄糖醛酸; AHBA: α -羟基丁酸; BHBA:
 5 α -羟基丁酸; GHBA: γ -羟基丁酸; PHPAA: p-羟基苯乙酸; AKBA: α -丁酮酸; AKGA: α -酮戊二酸; AKVA: α -酮基戊酸; LAME: D,L-乳酸甲酯; SaccA: D-糖酸(saccharic acid); BromosuccA: 溴代丁二酸; GAA: 甘氨酸-L-天冬氨酸; GGA: 甘氨酸-L-谷氨酸; HydPro: 羟基-L-脯氨酸; PyroGluA: L-焦谷氨酸; GABA: γ -氨基丁酸; PEA: 苯乙胺; GlycP: D,L- α -磷酸甘油; Gluc-1-P: 葡萄糖-1-磷酸; Gluc-6-P: 葡萄糖-6-磷酸
 10

表 10.副球菌种菌株的生化特征: 12=R1512; 34=R1534; 14=R114; 06=R1506; 66=MBIC3966; 74=DSM 11574^T; 96=E-396^T; 37=DSM 6637^T

	12	34	14	06	66	74	96	37
还原硝酸盐为亚硝酸盐	-	-	-	-	-	-	-	+
还原硝酸盐为氮	-	-	-	-	-	-	-	+
来自色氨酸的吲哚	-	-	-	-	-	-	-	-
葡萄糖发酵	-	-	-	-	-	-	-	-
精氨酸水解酶	-	-	-	-	-	-	-	-
脲酶	S/+5	-	-	S/+5	+	-	-	-
七叶灵水解 ^a	弱	S/+5	S/+5	+	+	+	+	+
明胶水解 ^b	-	-	-	-	-	-	-	-
β -半乳糖苷酶	+	+	+	+	+	+	+	-

^a: β -葡萄糖苷酶; ^b: 蛋白酶; S/+5: Slow+5 天

- 生理试验 对新副球菌种的 5 个菌株以及 *paracoccus marcusii* DSM
 15 11574^T, *paracoccus carotinifaciens* E-396^T 和 *paracoccus solventivorans* DSM 6637^T 进行几项理化和形态学试验。每项试验所用方法如下:

生长的温度范围 细胞在 28℃, 在 LMG 培养基 12 上培养 24 小时。用无菌水制备密度在 1-2McFarland 单位的细胞悬浮液。从该悬浮液中转移 3

滴液体到 LMG 培养基 12 的琼脂表面。一滴用划线法稀释，其余两滴不破坏。将该板在 10℃，25℃，30℃，33℃，37℃和 40℃下，孵育在有氧条件下，并在 24、48 小时和 5 天后检查生长情况。通过与 30℃(即对照)下的生长情况相比较，可以目测其生长(各滴汇合，稀释的接种液在划的线上形成集落)。按下述打分(与对照板相比)：生长优秀(++)，生长较好(相当于对照)(+)，生长较差(±)，生长极差(±)，无生长(-)。括号内结果是在划线处观察到的，其与保持未破坏液滴处的汇合生长情况不相同。

耐盐性 细胞在 28℃，在 LMG 培养基 12 上培养 24 小时。用无菌水制备密度在 1-2McFarland 单位的细胞悬浮液。从该悬浮液中转移 3 滴液体到 LMG 培养基 12 的琼脂表面，所述培养基补充有 NaCl，其终浓度为 3%，6%和 8%。一滴用划线法稀释，其余两滴未破坏。将该板在 28℃下，温育在有氧条件下，并在 24、48 小时和 5 天后检查生长情况。通过与未加 NaCl(对照)的生长情况相比较，可以目测其生长(各滴汇合，稀释的接种液在划的线上形成集落)。按下述打分(与对照板相比)：生长优秀(++)，生长较好(相当于对照)(+)，生长较差(±)，生长极差(±)，无生长(-)。括号内结果是在划线处观察到的，其与保持未破坏液滴处的汇合生长情况不相同。

生长的 pH 范围 细胞在 28℃，在 LMG 培养基 12 上培养 24 小时。用无菌水制备密度在 1-2McFarland 单位的细胞悬浮液。从该悬浮液中转移 3 滴液体到含 10 ml LMG 培养基 12 液体培养基管中，所述培养基的 pH 值被调整，使灭菌后最终 pH 值为 pH6.1，pH6.3，pH7.0，pH7.7，pH8.1 和 pH9.1。在 28℃下有氧(振荡)培养该液体培养基。在 24、48 小时、第 3 天和第 6 天检查生长情况。通过与 pH7.0(对照)的生长情况相比较，按增加的混浊度测定其生长情况(使用 BIOLOG 浊度计，按透射%进行测定)。按下述打分(与对照板相比)：生长优秀(++)，生长较好(相当于对照)(+)，生长较差(±)，生长极差(±)，无生长(-)。

淀粉水解作用 细胞在 28℃，在 LMG 培养基 12 平板上培养 24 小时。从平板上取一环细胞转移到补充有 0.2%可溶性淀粉的 LMG 培养基 12 的琼脂表面，并划线。然后在 28℃下有氧培养该液体培养基。当这些菌株达到生长良好状态时(48 小时后)，用鲁戈尔碘液(lugol solution)(0.5%I₂ 和 1%KI 在蒸馏水中)覆盖该板。(淀粉未水解的区域琼脂的颜色为蓝色，与该区域对比)可以根据生长旁边的无色区域，测定水解程度，

反硝化作用 细胞在 28℃, 在 LMG 培养基 12 平板上培养 24 小时。从平板上取一环细胞插入含有半固体(0.1%琼脂)LMG 培养基 12 的试管中, 所述培养基补充有 1%KNO₃。在 28℃下孵育该板 5 天。根据穿刺处旁边形成的气体测定脱氮作用(硝酸盐产生的 N₂)。

- 5 **未加入电子接受体情况下, 无氧条件下的生长** 细胞在 28℃, 在 LMG 培养基 12 平板上培养 24 小时。从平板上取一环菌转移到 LMG 培养基 12 的琼脂表面, 并划线。在 30℃、无氧条件(约 10%CO₂+约 90%N₂)下孵育该板。24 小时后和 5 天后检查生长情况。通过与有氧条件(对照)相比较, 测定其生长情况。按下述打分(与对照板相比): 生长优秀(++), 生长较好(相当于对照)(+), 生长较差(±), 生长极差(±), 无生长(-)。

- 10 **加入葡萄糖(发酵)情况下, 无氧条件下的生长** 细胞在 28℃, 在 LMG 培养基 12 平板上培养 24 小时。从平板上取一环菌插入含有 Hugh 和 Leifson[J.Bacteriol. 66:24-26(1953)]基础琼脂培养基的试管中。向培养基上部加入石蜡油, 并在 30℃孵育该板。48 小时后和 5 天后检查生长情况。目测
15 其生长情况。按下述打分: 生长好(+), 生长差(±), 无生长(-)。

- 以 KNO₃ 作为电子接受体, 无氧条件下的生长情况** 细胞在 28℃, 在 LMG 培养基 12 平板上培养 24 小时。从平板上取一环菌转移到补充有 0.1%KNO₃ 的 LMG 培养基 12 的琼脂表面, 并划线。在 30℃、无氧条件(约 10%CO₂+约 90%N₂)下孵育该板并在 3 天后检查生长情况。通过与有氧条件
20 (对照)相比较, 目测其生长情况。按下述打分(与对照板相比): 生长优秀(++), 生长较好(相当于对照)(+), 生长较差(±), 生长极差(±), 无生长(-)。

- 过氧化氢酶和氧化酶反应** 细胞在 28℃, 在 LMG 培养基 12 平板上培养 24 小时。过氧化氢酶活性的阳性结果表现为: 将菌落悬浮在一滴 10%H₂O₂ 中, 会有气泡产生。氧化酶活性的阳性结果表现为: 将菌落涂布在浸透了
25 1%四甲基对苯二胺(tetramethylparaphenylene)的滤纸上, 会形成紫红色。

菌落色素形成 细胞在 28℃, 在 LMG 培养基 12 上培养 5 天。肉眼观察菌落颜色。

- 细胞形态学和游动性** 细胞在 28℃, 在 LMG 培养基 12 上培养 24 小时。用无菌水制备细胞悬浮液。使用配备有相衬光学(phase contrast optics)的
30 Olympus 光学显微镜(放大倍数 1000x), 观察细胞形态学和游动性。

 表 11 总结了形态学和游动性试验的结果。新副球菌种的 5 个菌株对进

行的所有理化或形态学试验中，反应基本相同。所有五个新副球菌种菌株反应相同、且能够使这些菌株与 *Paracoccus marcusii* DSM 11574^T 和/或 *Paracoccus carotinifaciens* E-396^T 相区分的试验是：在 40°C 生长，在 8% 的 NaCl 中生长，pH9.1 下生长和菌落色素沉着。

5 菌株 R-1512, R1534, R114 和 R-1506 的玉米黄质产量 在 ME 培养基中培养 R-1512, R1534, R114 和 R-1506 菌株，所述培养基含有(每升蒸馏水)：5g 葡萄糖，10g 酵母提取物，10g 胰蛋白胨，30g NaCl 和 5g MgSO₄·7H₂O。高压灭菌前用 5N NaOH 将培养基的 pH 调整为 7.2。所有的培养(25-ml 体积在 250ml 带挡板、带塑料盖的 Erlenmeyer 烧瓶中)均在 28°C 下，200rpm 振荡下进行。种子培养物是从冷冻的甘油化保藏物接种，并过夜培养。将等份种子培养物转移到试验烧瓶中，使其起始光学密度为 660nm(OD₆₆₀)下 0.16。然后在 28°C 下，200rpm 振荡培养。在整个培养过程检测生长情况，并在第 6, 10(或 R114 菌株时为 15)和 24 小时，取出等份培养物，用实施例 1 所述方法分析类胡萝卜素。

15 这些条件下，菌株 R-1512, R1534 和 R-1506 的倍增时间分别为 0.85 小时，1.15 小时和 1.05 小时。菌株 R114 再现性地显示出二相性的生长曲线；在初相菌株 R114 的倍增时间为 1.4 小时，而第二相中的倍增时间为 3.2 小时。

20 表 12 显示副球菌种菌株在 ME 培养基中的玉米黄质产量和特异性形成(将玉米黄质产量校正到 OD₆₆₀)。该数据是 4 种独立试验的平均值，其中每个试验的每个菌株均在两个烧瓶中进行试验。清楚显示出：经典衍生的(classically-derived)突变株 R1534 和 R114 与亲代菌株 R-1512 进行比较的玉米黄质产量。菌株 R-1506 的玉米黄质产量与菌株 R-1512 大致相同。没有在任何培养物中检测到其他类胡萝卜素。

25 表 11.副球菌种菌株的生理特性：12=R1512；34=R1534；14=R114；06=R1506；66=MBIC3966；74=DSM 11574^T；96=E-396^T；37=DSM 6637^T

30

时间[天]	12	34	14	06	66	74	96	37
	在 10°C 生长							
1	-	-	-	-	-	-	-	-
5	±(±)	±(±)	±(-)	±(±)	±(-)	±(±)	±(±)	±(±)

5		在 25℃ 生长							
	1	+	+	+	+	+(±)	+(±)	+(±)	+(-)
	5	+	+	+	+	+	+	+	+
		在 30℃ 生长							
	1	+	+	+	+	+	+	+	+
10	5	+	+	+	+	+	+	+	+
		在 33℃ 生长							
	1	+	+	+	+	+	+	+	+
	5	+	+	+	+	+	+	+	+
		在 37℃ 生长							
15	1	+	+	+(±)	+	+	±(-)	±(-)	+
	5	+	+	+	+	+	±(-)	±(±)	+
		在 40℃ 生长							
	1	+	+(±)	+(-)	+(±)	±(-)	-	-	+(*)
	5	+	+(±)	+(-)	+	+(-)	-	-	+(*)
20		3% NaCl 下生长							
	1	+	+	+	+	+	+	+	±
	5	+	+	+	+	+	+	+	+
		6% NaCl 下生长							
	1	+(±)	±(±)	±(±)	+	±(±)	±(-)	±(-)	-
25	5	+	+	+	+	+(*)	+(±)	+(±)	-
	时间[天]	12	34	14	06	66	74	96	37
		8% NaCl 下生长							
	1	+(±)	±(±)	±(-)	+(±)	±(±)	-	-	-
	5	+	+	+	+	+(*)	±(-)	±(-)	-
30		pH 6.1 下生长							
	1	+	+	+	+	-	-	-	-
	6	+	+	+	+	+	+	+	+
		pH 6.3 下生长							

5

10

15

20

25

30

1	+	+	+	+	+	±	+	±
6	+	+	+	+	+	+	+	+
pH 7.0 下生长								
1	+	+	+	+	+	+	+	+
6	+	+	+	+	+	+	+	+
pH 7.7 下生长								
1	+	+	+	+	+	±	±	±
6	+	+	+	+	+	+	+	+
pH 8.1 下生长								
1	+	+	+	+	+	-	-	±
6	+	+	+	+	+	+	+	+
pH 9.1 下生长								
1	±	+	-	-	+	-	-	-
6	+	+	+	+	+	-	+	+
淀粉水解								
	-	-	-	-	-	-	-	-
反硝化作用								
	-	-	-	-	-	-	-	+
不加入电子受体, 厌氧生长								
	-	-	-	-	-	-	-	-
	12	34	14	06	66	74	96	37
加入葡萄糖 (发酵), 厌氧生长								
	-	-	-	-	-	-	-	±
加入 KNO ₃ , 厌氧生长								
	-	-	-	-	-	-	-	-
过氧化酶反应								
	+	+	+	+	+	+	+	+
氧化酶反应								
	+	+	+	+	+	+	+	+
革兰氏染色								
	-	-	-	-	-	-	-	-

和 Bacher, In: Setlow(ed), Genetic Engineering, Principles and Methods, Kluwer Academic/Plenum Publishers, New York 22: 121-153 (2000)]. 这个预言性数据分析方法能够根据对单一下游天然产物的分析, 准确地评估葡萄糖的分解代谢。本文中, 该方法涉及在含有未标记葡萄糖和特定 ^{13}C -标记葡萄糖的各种二元混合物的培养基中培养细菌, 然后纯化生成的玉米黄质并通过 NMR 光谱学分析标记模式。下面详细介绍了所用方法以及试验结果。

培养用于 ^{13}C 标记试验的副球菌种菌株 R114 从 Fluka(Milwaukee, WI, USA)购买未标记的 D-葡萄糖一水化物。从 Isotec(Miamisburg, OH, USA)购买 $[\text{U}-^{13}\text{C}_6]$ -D-葡萄糖, 从 Cambridge Isotope Laboratories (Andover, MA, USA) 10 购买 $[1-^{13}\text{C}_1]$ -D-葡萄糖, $[2-^{13}\text{C}_1]$ -D-葡萄糖和 $[6-^{13}\text{C}_1]$ -D-葡萄糖。从 EM Science(Gibbstown, NJ, USA)购买酵母提取物和蛋白胨。所有的其它盐和溶剂为分析纯, 购自正规化学试剂供应商。

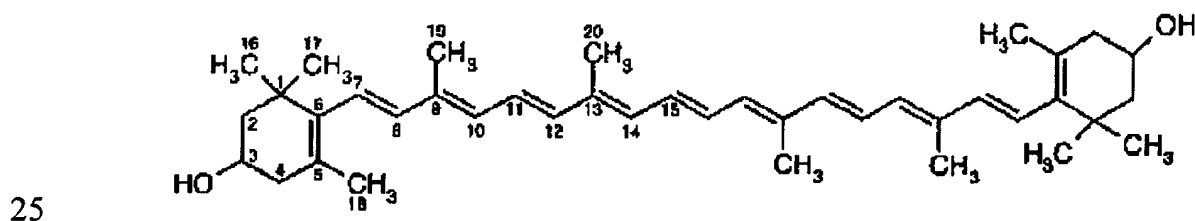
所有的培养均来自冻冷的细胞悬浮液(细胞密度为 12 OD_{660} 单位, 25% 甘油, 保藏在 -70°C)。使用 1ml 解冻的细胞悬浮液, 在含有 100ml 362F/2 培养基中, 进行接种前培养(500ml 带挡板的烧瓶), 所述培养基含有以下组分: 15 30g/l D-葡萄糖, 10g/l 酵母提取物, 10g/l 蛋白胨, 5g/l NaCl, 2.5g/l $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.75g/l $(\text{NH}_4)_2\text{HPO}_4$, 0.625g/l K_2HPO_4 , 187.5mg/l $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 0.2g/l $(\text{NH}_4)_2\text{Fe}(\text{SO}_4)_2 \cdot 6\text{H}_2\text{O}$, 15mg/l $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 12.5mg/l $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$, 5mg/l $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 0.5mg/l $\text{NiSO}_4 \cdot 6\text{H}_2\text{O}$, 15mg/l Na-EDTA 和 9.375 $\mu\text{l/l}$ HCl(37%储藏液)。该培养基的起始 pH 为 7.2。 20

在 28°C , 200rpm 振荡下进行 24 小时前培养, 前培养之后, OD_{660} 约为 22 吸收单位。在含有 362F/2 培养基的 Bioflo 3000 生物反应器(New Brunswick Scientific, Edison, NJ, USA)中进行主要培养, 所述培养基含有以下组分: 25 30g/l 总 D-葡萄糖(见下面 ^{13}C -标记: 未标记葡萄糖), 20g/l 酵母提取物, 10g/l 蛋白胨, 10g/l NaCl, 5g/l $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 1.5g/l $(\text{NH}_4)_2\text{HPO}_4$, 1.25g/l K_2HPO_4 , 0.4g/l $(\text{NH}_4)_2\text{Fe}(\text{SO}_4)_2 \cdot 6\text{H}_2\text{O}$, 375mg/l $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 30mg/l $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 25mg/l $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$, 10mg/l $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 1mg/l $\text{NiSO}_4 \cdot 6\text{H}_2\text{O}$, 30mg/l Na-EDTA 和 18.75 $\mu\text{l/l}$ HCl(37%储藏液)。四种试验中所用每种 ^{13}C -标记葡萄糖数量(表达为培养基中总共 30g/l 葡萄糖的百分比)为: 条件 1, 4% 30 $[\text{U}-^{13}\text{C}_6]$ -D-葡萄糖; 条件 2, 50% $[1-^{13}\text{C}_1]$ -D-葡萄糖; 条件 3, 25% $[2-^{13}\text{C}_1]$ -D-葡萄糖+1% $[\text{U}-^{13}\text{C}_6]$ -D-葡萄糖; 条件 4, 25% $[6-^{13}\text{C}_1]$ -D-葡萄糖+1%

[U- $^{13}\text{C}_6$]-D-葡萄糖。同样包括未标记葡萄糖的对照。对于条件 1 和 2(以及未标记的对照), 培养体积为 2l, 而条件 2 和 4 为 1l。用前培养物接种生物反应器(20ml/l 初始体积), 培养 22-24 小时, 该时间点培养基内已没有葡萄糖剩余。培养条件为: 28℃, pH7.2(对照为 25% H_3PO_4 和 28% NH_4OH), 溶解氧控制在(在格状物内搅拌(in a cascade with agitation))最小 40%, 搅拌速度和吹气速度分别为 300rpm(最小)和 1vvm。

纯化玉米黄质 培养末期, 将培养物冷却到 15℃。向每升培养物中加入 500ml 无水乙醇(absolute ethanol), 并以 100rpm 搅拌 20 分钟。将处理过的培养物在 5000x g 离心 20 分钟, 弃上清。然后用 5 倍体积的 THF 搅拌提取湿沉淀 20 分钟。提取的混合物经离心后, 保留上清, 用 1 倍体积的 THF 在相同条件下再提取所获沉淀一次, 离心。混合(提取物)上清, 并通过旋转蒸发, 浓缩到 50ml。向浓缩的 THF 溶液中加入 5ml 己烷。混合后形成了乳状液, 可通过离心将其分离。收集液体相, 用等体积的饱和 NaCl 溶液稀释, 用二氯甲烷再次提取。收集二氯甲烷相, 并与 THF/己烷相混合。然后通过旋转蒸发器再次浓缩该有机提取物的混合物, 以去除二氯甲烷。然后将该溶液应用于硅胶柱, 并用 n-己烷和乙醚混合物(1: 1)洗脱。去除最先洗脱的一小条浅黄色带。主要的玉米黄质产物是在柱子中移动缓慢的一条宽带中被洗脱出来的。产物大约需要 2 升溶剂彻底地洗脱主要带。将洗脱液收集在圆底烧瓶中, 并在 40℃ 旋转蒸发去除溶剂。将残余物在 40℃ 溶解于少量二氯甲烷中, 然后使溶液慢慢冷却。向混合物中一滴一滴地加入己烷, 直到出现混浊。在 4℃ 下, 48 小时内彻底完成结晶。将结晶收集在滤纸上, 用冷甲醇洗涤, 真空干燥。

NMR 研究 NMR 光谱学分析玉米黄质。为进行参考, 以下式说明玉米黄质的化学结构:



用 Bruker DRX 500 分光计, ^1H -NMR 和 ^{13}C -NMR 光谱分别记录在

500.13MHz 和 125.6MHz。一元试验和二元 NADEQUATE 试验的操作(acquisition)及处理参数均根据标准 Bruker 软件(XWINNMR)。溶剂氘化(deuterate)为氯仿。化学位移参照溶剂信号。

在相同条件下,记录同位素标记的玉米黄质样品的 ^{13}C -NMR 光谱,以及天然 ^{13}C 丰度(abundance)的玉米黄质样品的 ^{13}C -NMR 光谱。测定每个 ^{13}C -NMR 信号的积分(integrals),标记化合物中各个碳原子的信号积分与天然丰度材料的相参照,这样提供了标记分子物种中每个位点的相对 ^{13}C 丰度。然后在 1.71ppm, H-18 的 ^1H NMR 信号中,从 ^{13}C 偶联随体(coupling satellite),将该相对丰度转化为绝对丰度。在多元标记的玉米黄质样品的 ^{13}C NMR 光谱中,每个随体被分别积分。然后将每个随体对的积分与指定碳原子的总信号积分相参照。玉米黄质含有总共八个类异戊二烯部分(2DMAPP 单元和 6 IPP 单元);仅仅观察到 20 个化学位移简并性的 ^{13}C NMR 信号。

用[U- $^{13}\text{C}_6$]葡萄糖和未标记葡萄糖混合物(1: 7.5; w/w)的试验中,玉米黄质的所有碳原子都被标记,并显示出 $^{13}\text{C}^{13}\text{C}$ 偶联的随体(表 13)。4 个碳原子的信号具有高强度的 $^{13}\text{C}^{13}\text{C}$ 偶联随体(在指定原子的总 NMR 信号强度中为 $61.2\pm0.6\%$,表 13)。甲基原子 C-17/C-17'的信号只显示出微弱的 ^{13}C 偶联随体,其相对强度为 6%。中心的信号代表来自未标记葡萄糖的材料。这些信号未显示出远距离偶联的证据。很容易地从 $^{13}\text{C}^{13}\text{C}$ 偶联常数(表 13)和二元 INADEQUATE 试验发现碳的连通性(connectivity)。

三个碳原子获得了[6- $^{13}\text{C}_1$]葡萄糖的标记。其他两个碳原子被[2- $^{13}\text{C}_1$]葡萄糖标记。没有显著数量的标记来自[1- $^{13}\text{C}_1$]葡萄糖。

与未标记玉米黄质的光谱相比较,以及通过评价 ^1H NMR 光谱(表 13)中 $^1\text{H}^{13}\text{C}$ 偶联随体,来测定所有非等时碳原子的 ^{13}C 丰度。通过积分偶联随体,测定用[U- $^{13}\text{C}_6$]葡萄糖进行的试验中,连接转移的碳原子对所占部分。可根据重建 IPP 前体所发现的标准偏差所示,精确地重建 IPP 构建块的标记模式。DMAPP 和 IPP 的重构建标记模式在试验极限内相同。

表 13.提供了 ^{13}C 标记葡萄糖的副球菌种菌株 R114 所形成 ^{13}C 标记玉米黄质的 NMR 结果

5

10

15

20

位置	$\delta - ^{13}\text{C}$ ppm	J_{CC}, Hz	^{13}C 标记的葡萄糖前体				
			$[1-^{13}\text{C}]-$	$[2-^{13}\text{C}]-$	$[6-^{13}\text{C}]-$	$[\text{U}-^{13}\text{C}_6]$ 葡萄糖	
	134.08	44.2 (18, 18')	% ^{13}C	% ^{13}C	% ^{13}C	% ^{13}C	% $^{13}\text{C}^{13}\text{C}$
1, 1'	37.13	36.0 (16, 16')	1.10	10.71	2.22	3.47	61.2
2, 2'	48.46	35.8 (3, 3')	1.20	2.58	10.27	3.65	61.1
3, 3'	65.10	35.8 (2, 2')	1.12	12.47	2.38	3.64	60.4
4, 4'	42.57	37.1	1.27	2.59	10.63	3.89	8.4
5, 5'	126.17	44.2 (18, 18')	1.14	12.45	3.19	3.68	61.1
6, 6'	137.77	56.4 (7, 7')	1.30	2.15	9.98	3.60	60.4
7, 7'	125.59	56.2 (6, 6')	1.12	10.11	2.82	4.09	61.4
8, 8'	138.50	71.6, 55.7	1.28	2.24	9.95	3.92	4.3, 5.0
9, 9'	135.69	43.1 (19, 19')	1.12	9.53	2.95	3.84	61.7
10, 10'	131.31	59.7 (11, 11')	1.21	3.18	9.61	3.80	61.1
11, 11'	124.93	59.7 (10, 10')	1.10	8.79	2.70	4.02	61.0
12, 12'	137.57	70.5	1.20	2.01	8.80	3.59	5.1
13, 13'	136.48	43.1 (20, 20')	1.12	9.86	3.59	3.93	61.4
14, 14'	132.60	60.4 (15, 15')	1.21	2.83	10.51	3.77	59.5
15, 15'	130.08	60.4 (14, 14')	1.12	9.18	3.33	4.02	61.2
16, 16'	30.26	36.3 (1, 1')	1.27	3.19	12.31	3.91	62.0
17, 17'	28.73	34.9 (1, 1')	1.30	3.43	12.31	3.88	6.0
18, 18'	21.62	44.2 (5, 5')	1.27	3.01	11.66	3.70	62.0
19, 19'	12.82	43.1 (9, 9')	1.29	3.12	11.64	3.86	62.3
20, 20'	12.75	42.9 (13, 13')	1.33	3.21	11.99	3.75	62.1

25 以上测定的试验标记模式可以与各种预测相比，所述预测不仅考虑到用于类异戊二烯生物合成的甲羟戊酸途径与 DXP 途径，还考虑到不同的葡萄糖代谢途径。真细菌(Eubacteria)会典型地首先通过糖酵解代谢或 Entner-Doudoroff 途径利用葡萄糖。糖酵解会从葡萄糖产生两种丙糖磷酸分子。葡萄糖的 C-1 和 C-7 都转向了糖酵解过程中所形成的丙糖磷酸的 3-位。

30 另一方面，在 Entner-Doudoroff 途径中，葡萄糖转化为甘油醛 3-磷酸和丙酮酸的混合物。葡萄糖的 C-1 无一例外地均转向丙酮酸的 C-1，而葡萄糖的

C-6 则全部转向甘油醛 3-磷酸的 C-3。

糖酵解和 Entner-Doudoroff 途径的中间体和产物作为两种类异戊二烯生物合成途径的起始材料。关于甲羟戊酸途径, 丙酮酸和丙糖磷酸可转化为乙酰-CoA 的前体。通过糖酵解途径的葡萄糖代谢可将葡萄糖 C-1 和 C-6 5 的标记转到乙酰-CoA 的甲基。通过 Entner-Doudoroff 途径的葡萄糖代谢可以在转化丙酮酸为乙酰-CoA 的过程中, 将葡萄糖的 C-1 损耗。

试验中观察到的副球菌种菌株 R114 所生成玉米黄质的富集及其 $^{13}\text{C}^{13}\text{C}$ 偶联模式与 Entner-Doudoroff 途径和甲羟戊酸途径相结合时, 玉米黄质生物合成所需要的标记模式完全一致。如果两个糖酵解和 Entner-Doudoroff 途径 10 在所用试验条件下可同时进行, 至少一些来自 $[1-^{13}\text{C}_1]$ 葡萄糖的标记应该会进入玉米黄质。另外, 甲羟戊酸途径可以最多将两个碳原子单元分配给类萜, 而在 DXP 途径中, 通过丙糖磷酸前体可将三个碳原子单元运送到类异戊二烯。尽管这样的三碳单元被 DXP 途径相关的重排所分离, 三个标记碳原子的单元仍可通过远距离偶联识别。在培养的植物细胞(*Cantharantus roseus*)中, 相应的 $^{13}\text{C}-^{13}\text{C}$ 远距离偶联已在从 $[2,3,4,5-^{13}\text{C}_4]$ 1-脱氧-D-木酮糖到 15 类胡萝卜素叶黄素的生物合成中观察到[Arigoni 等, *Proc.Nat.Acad.Sci.* 94: 10600-10605(1997)]。副球菌种菌株 R114 形成的玉米黄质在本试验中未观察到这种长距离偶联。

应该注意, 这里的结果证实了类异戊二烯通过甲羟戊酸途径在副球菌 20 种菌株 R114 中形成, 该结果还说明: 在所采用的生长条件下, 几乎或完全没有糖酵解的葡萄糖代谢, 并不排除在 Entner-Doudoroff 途径之外, 通过戊糖磷酸途径进行一些葡萄糖代谢的可能性。可通过分析标记丙酮酸来源的氨基酸模式来定量测定以后两种途径进行的葡萄糖代谢(如对 *Paracoccus denitrificans* 进行的过程 [Dunstan 等, *Biomedical and Environ. Mass Spectrometry* 19: 369-381(1990)])。 25

实施例 4: 副球菌种菌株 R114 中编码 IPP 异构酶和甲羟戊酸途径中各酶的基因的克隆及测序

培养条件 在 28°C , F-培养基(10g/l 胰蛋白胨, 10g/l 酵母提取物, 30g/l NaCl, 10g/l D-葡萄糖, 5g/l $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, pH7.0)或在以上实施例 3 所述的前培养培养基中, 培养副球菌种菌株 R114。200rpm 在旋转摇床上培养液体 30 培养物。

分离基因组 DNA 将 600ml 副球菌种菌株 R114 的培养物在 4℃, 10000x g 离心 10 分钟, 用 200ml 溶解缓冲液(0.1M NaCl, 50mM EDTA, 10mM Tris-HCl, pH7.5)洗涤沉淀一次, 再用 100ml 溶解缓冲液洗涤一次。将最终获得的沉淀重悬浮于 20ml 含有 50mg 溶菌酶和 1mg RNase A(不含 DNase)的溶解缓冲液中。在 37℃ 孵育 15 分钟后, 加入 1.5ml 20% N-月桂酰-肌氨酸钠和 2.25mg 蛋白激酶 K。在 50℃ 孵育 30-60 分钟后, 用 1 倍体积的缓冲液-饱和苯酚, pH7.5-7.8(Life Technologies, Rockville, MD, USA), 轻轻地但彻底地混合, 提取溶解液。乳状液 30000x g 离心 20 分钟, 用苯酚提取液体相。如前述分离该相并用 1 倍体积的苯酚: 氯仿(1: 1)提取两次。该步骤中, 在 10 翻转旋转器内 3200x g 离心 20 分钟, 就可获得良好的相分离。用 1 倍体积的氯仿进行最后一次提取后, 加入 0.1 倍体积 3M 乙酸钠(pH5.2), 并用 2 倍体积冰冷的乙醇覆盖该溶液。用玻璃棒卷绕沉淀的 DNA, 浸入 70%的乙醇 5 分钟, 氯仿洗涤, 然后干燥 5-10 分钟。将 DNA 重悬在 5ml TE(10mM Tris-HCl, pH7.5, 1mM EDTA)中过夜。由于痕量的玉米黄质, 该溶液为黄色, 15 重复进行如上的有机提取和卷绕, 直到获得无色的制剂。

λ-DNA 的分离: 根据制造商的说明书, 使用 Qiagen® Lambda Kit(Qiagen, Hilden, Germany)。

聚合酶链式反应(PCR): 从 Life Technologies(Rockville, MD, USA)购买寡核苷酸。根据制造商的说明书, 使用 GC-rich PCR system(Roche Molecular Biochemicals, Mannheim, Germany), 在 GeneAmp®PCR system 9700(PE Applied Biosystems, Foster City, CA, USA)上进行 PCR。通常, 所用 MgCl₂ 浓度为 1.5mM, 加入溶解溶液的终浓度为 1M。20

DNA 标记和检测: 分别使用 PCR DIG Probe Synthesis Kit 和 DIG Luminescent Detection Kit(均购自 Roche Molecular Biochemicals, Mannheim, Germany)进行 DNA 标记和检测。25

DNA 测序: 根据制造商的说明书, 使用 BigDye®DNA 测序试剂盒(PE Applied Biosystems, Foster City, CA, USA)进行测序反应。测序反应在 DyeEx™ 旋转柱(Qiagen, Hilden, Germany)上纯化, 用 ABI Prism™ 310 Genetic Analyzer(PE Applied Biosystems, Foster City, CA, USA)进行片段分离和检测。30

λ-文库: 用 Sau3AI 部分酶切的副球菌种菌株 R114 DNA 在 λ FIX®II 中

的定制文库购自 Stratagene(La Jolla, CA, USA)。

副球菌种菌株 R114 甲羟戊酸途径基因簇的克隆, 测序和表征 甲羟戊酸途径各酶之一, 甲羟戊酸二磷酸脱羧酶, 含有高度保守的几个氨基酸长度的区域。从所有可获得的真细菌甲羟戊酸二磷酸脱羧酶的排列中选择三个这样的区域, 使用在副球菌种菌株 R1534 类胡萝卜素基因簇中发现的偏爱密码子(表 14)设计寡核苷酸。

为两个同源区域所设计的寡核苷酸如表 15 所示。为减低简并性程度, 每个多肽均设计了几组寡核苷酸。例如, 各个寡核苷酸 mvd-103a-d 仅在 3' 末端的第三个核苷酸有所区别, 每各寡核苷酸均具有一种可能的甘氨酸密码子(尽管 GGA 很罕见, 但还是将其容纳, 因为其与 3' 末端紧密相连)。通过设计寡核苷酸为两个残基来说明可替换的氨基酸, 如, 在多肽 1 的第二个位置(表 15), 寡核苷酸 mvd-101a 和 mvd-101b 分别是亮氨酸或异亮氨酸特异的。使用副球菌种菌株 R114 DNA 作为模板, 用寡核苷酸 mvd-101 和 mvd-104 或 mvd-106 进行 PCR, 得到预期大小的产物。在载体 pCR[®]2.1-TOPO(Invitrogen, Carlsbad, CA, USA)中克隆该 PCR 产物并测序。使用该克隆片段作为副球菌种菌株 R114 DNA Southern 分析的探针, 发现其与一段约 950bp 的 BamHI-SalI 片段杂交。用 BamHI 和 SalI 酶切副球菌种菌株 R114 DNA, 琼脂糖凝胶电泳分离其片段。分离到约 950bp 的片段, 并克隆到载体 pUC19 中。然后使用 mvd-PCR 片段作为探针, 筛选该部分文库, 测序插入的阳性克隆。并行地, 使用 mvd-PCR 片段作为探针, 筛选由副球菌种菌株 R114 DNA 制备的 λ 文库。从两个阳性 λ 克隆分离 DNA, 并用 BamHI 和 SalI 或 EcoRI 和 SalI 酶切。分离出大量的限制性片段, 将其克隆到载体 pUC19 中。几个片段含有与编码甲羟戊酸途径各蛋白基因同源的序列。使用来自克隆的限制性片段各序列的引物, 使用 λ 克隆的 DNA 作为模板, 通过 PCR, 制备出连接这些独立序列的克隆。由所有片段(SEQ ID NO: 42, 44, 46, 48, 50 和 52)和被编码蛋白各序列组装而成的序列如序列表(SEQ ID NO: 43, 45, 47, 49, 51 和 53)所示。由于 PatentIn Program 的限制, 重叠基因的操纵子不能作为单独的序列显示。这样, 在甲羟戊酸操纵子中的各基因, 每个基因操纵子的整个核苷酸序列 均重复。相应地, SEQ ID NO: 42, 44, 46, 48, 50 和 52 是相同的。为实现本发明, 我们使用 SEQ ID NO: 42 指代甲羟戊酸操纵子的核苷酸序列。

与已知的其它细菌甲羟戊酸基因簇相比,副球菌种菌株 R114 的甲羟戊酸途径基因排列是独一无二的。除了副球菌种菌株 R114,只有 *Borrelia burgdorferi* 和 *Streptomyces* 种菌株 CL190(Takagi 等, 同上)在单独的操纵子中具有所有的甲羟戊酸基因(Wilding 等, 同上)。在 *Streptococcus pyogenes* 中,所有的甲羟戊酸基因均集中在单一位点上,但是归属于两个操纵子。所有的其它物种都具有两种激酶的两个位点,甲羟戊酸二磷酸脱羧酶归属于一个操纵子,HMG-CoA 合成酶和 HMG-CoA 还原酶在第二个位点上,二者或构成一个操纵子(在 *Streptococcus pyogenes* 中),或为分离的转录单位。除了 *Streptococcus* 的成员外,所有的物种都具有额外的基因,与甲羟戊酸簇相连,该基因最近被鉴定为 IPP 异构酶(在 *Streptomyces* 种菌株 CL190 中 idi 基因)(Kaneda 等, 同上)。两个 *Enterococcus* 种和 *Staphylococcus haemolyticus* 具有与 HMG-CoA 还原酶相连的乙酰-CoA 乙酰基转移酶。在 *Enterococcus* 物种中,后两个基因是融合在一起的。

表 14: 在副球菌种菌株 R1534 中类胡萝卜素(crt)基因的密码子使用率

氨基酸	密码子	所用数目	%被利用
A - Ala	GCT	3	1.4
	GCC	96	46.2
	GCA	15	7.2
	GCG	94	45.2
C - Cys	TGT	0	0.0
	TGC	15	100.0
D - Asp	GAT	46	38.0
	GAC	75	62.0
E - Glu	GAA	17	25.4
	GAG	50	74.6
F - Phe	TTT	3	5.6
	TTC	51	94.4
G - Gly	GGT	16	10.8
	GGC	87	58.8
	GGA	5	3.4
	GGG	40	27.0
H - His	CAT	30	56.6
	CAC	23	43.4

5	I - Ile	ATT	5	6.4	
		ATC	72	92.3	
		ATA	1	1.3	
	K - Lys	AAA	4	14.3	
		AAG	24	85.7	
	L - Leu	TTA	0	0.0	
		TTG	5	2.9	
		CTT	15	8.7	
		CTC	11	6.4	
		CTA	1	0.6	
CTG		140	81.4		
10	M - Met	ATG	49	100.0	
	N - Asn	AAT	4	20.0	
AAC		16	80.0		
15	氨基酸	密码子	所用数目	%被利用	
	P - Pro	CCT	2	2.3	
		CCC	41	47.7	
		CCA	3	3.5	
		CCG	40	46.5	
	Q - Gln	CAA	6	11.3	
		CAG	47	88.7	
	20	R - Arg	CGT	11	7.3
			CGC	103	68.2
			CGA	2	1.3
CGG			26	17.2	
AGA			2	1.3	
AGG			7	4.6	
25	S - Ser	TCT	1	1.1	
		TCC	17	19.5	
		TCA	0	0.0	
		TCG	39	44.8	
		AGT	2	2.3	
		AGC	28	32.2	
30	T - Thr	ACT	2	2.7	
		ACC	36	48.9	
		ACA	4	5.3	
		ACG	33	44.0	

5

V - Val	GTT	6	5.7
	GTC	61	57.5
	GTA	1	0.9
	GTG	38	35.8
W - Trp	TGG	27	100.0
Y - Tyr	TAT	28	62.2
	TAC	17	37.8

表 15: 从两个保守的细菌 Mvd 多肽设计的寡核苷酸

10

15

20

25

30

			SEQ a
10	肽 1	AlaLeuIleLysTyrTrpGlyLys Ile ²	23
	核苷酸序列 ¹	CCSCTGATCAARTAYTGGGGBAARATC	24
15	寡核苷酸 mvd-101a (5' - 3')	GCSCTGATCAARTAYTGGGG	25
	寡核苷酸 mvd-101b (5' - 3')	GCSATCATCAARTAYTGGGG	26
	寡核苷酸 mvd-103a (5' - 3')	ATCAARTAYTGGGGTAA	27
	寡核苷酸 mvd-103b (5' - 3')	ATCAARTAYTGGGGCAA	28
	寡核苷酸 mvd-103c (5' - 3')	ATCAARTAYTGGGGGAA	29
	寡核苷酸 mvd-103d (5' - 3')	ATCAARTAYTGGGGAAA	30
20	肽 2	ThrMetAspAlaGlyProAsnVal Gln ²	31
	核苷酸序列 ¹ (5'-3')	ACSATGGAYGCSGGBCCSAAYGTS CAR	32
	互补序列 (3'-5') ¹ ,	TGSTACCTRCGSCCVGGSTTRCAS GTY	33
	寡核苷酸 mvd-104a (3' - 5')	TGGTACCTACGSCCVGG	34
25	寡核苷酸 mvd-104b (3' - 5')	TGGTACCTGCGSCCVGG	35
	寡核苷酸 mvd-104c (3' - 5')	TGCTACCTACGSCCVGG	36
	寡核苷酸 mvd-104d (3' - 5')	TGCTACCTGCGSCCVGG	37
	寡核苷酸 mvd-106a (3' - 5')	TACCTACGSCCVGGSTTRCA	38
30	寡核苷酸 mvd-106b (3' - 5')	TACCTGCGSCCVGGSTTRCA	39
	寡核苷酸 mvd-106c (3' - 5')	TACCTACGSCCVGGSGTYCA	40
	寡核苷酸 mvd-106d (3' - 5')	TACCTGCGSCCVGGSGTYCA	41

^a: SEQ ID NO:

¹: 使用副球菌种菌株 R1534 偏爱密码子, 见表 1

²: 在一些酶中存在的可替换氨基酸 S=C 或 G; R=A 或 G; Y=C 或 T; B=C 或 G 或 T; V=A 或 C 或 G

- 5 通过基因产物与综合数据库中蛋白的同源性来鉴定副球菌种菌株 R114 的甲羟戊酸操纵子基因。用链霉菌种菌株 CL190(SEQ ID NO: 54), *S.griseolosporeus*(SEQ ID NO: 55)和链霉菌种菌株 KO-3899(SEQ ID NO: 56)的细菌 I 类 HMG-CoA 还原酶进行副球菌种菌株 R114 HMG-CoA 还原酶的氨基酸排列(SEQ ID NO: 43)。EMBL/GenBank/DDBJ 数据库中, 链霉菌种菌株 CL190 的登录号为 q9z9n4, *S.griseolosporeus* 的为 q9znh1, 链霉菌种菌株 KO-3899 的为 q9znh0。存在两类 HMG-CoA 还原酶[Bochar 等, Mol.Genet.Metab. 66: 122-127 (1999); Boucher 等, Mol.Microbiol. 37:703-716(2000)]。真细菌的 HMG-CoA 还原酶通常属于第 II 类, 而在真核生物和古细菌(archaea)中发现了 I 类酶。链霉菌和副球菌的 HMG-CoA 还原酶以及来自霍乱弧菌(*Vibrio cholerae*)的酶是目前所知仅有的真细菌 I 类 HMG-CoA 还原酶。
- 10
- 15

- 用 EMBL 数据库中找到的高度同源物, 即草生欧文氏菌(*Erwinia herbicola*)(Q01335)(SEQ ID NO: 57), 布氏疏螺旋体(*Borrelia burgdorferi*)(O51627)(SEQ ID NO: 58), 集胞蓝细菌种(*Synechocystis*)PCC 6803 (P74287)(SEQ ID NO:59), 链霉菌种 CL190(Q9KWG2)(SEQ ID NO:60), *Streptomyces griseolosporeus*(Q9KWF6)(SEQ ID NO: 61), 硫磺矿硫化叶菌(*Sulfolobus solfataricus*) (P95997)(SEQ ID NO: 62), 普氏立克次氏体(*Rickettsia prowazekii*) (Q9ZD90)(SEQ ID NO: 63), 耐放射异常球菌(*Deinococcus radiodurans*) (Q9RVE2)(SEQ ID NO: 64), *Aeropyrum pernix* (Q9YB30)(SEQ ID NO: 65), 盐杆菌属菌种(*Halobacterium*)NRC-1(O54623)(SEQ ID NO: 66), 闪烁古生球菌(*Archaeoglobus fulgidus*) (O27997)(SEQ ID NO: 67), *Pyrococcus abyssi*(Q9UZS9)(SEQ ID NO: 68), *Pyrococcus horikoshii*(O58893)(SEQ ID NO: 69), 嗜热碱甲烷杆菌(*Methanobacterium thermoautotrophicum*) (O26154)(SEQ ID NO: 70), 詹氏甲烷球菌(*Methanococcus jannaschii*)(Q58272)(SEQ ID NO: 71), 嗜酸热原体(*Thermoplasma acidophilum*)(CAC11250)(SEQ ID NO: 72),
- 20
- 25
- 30

和 *Leishmania major* (Q9NDJ5)(SEQ ID NO: 73), 进行副球菌种菌株 R114 异戊烯二磷酸异构酶(IPP 异构酶)(idi)的氨基酸排列(SEQ ID NO: 45)。EMBL/GenBank/DDBJ 数据库会根据括号内的有机体名称, 指定登录号。最先的 9 个序列来自真细菌, 接下来的 8 个是来自古细菌(archaea)。有趣的是, 5 一个真细菌物种, 原生动物寄生虫 *Leishmania major*(SEQ ID NO: 73)也具有一个高度同源的蛋白。这出乎意料, 因为其它的真核生物具有不同的 idi, 命名为 1 型(Kaneda 等, 同上)。来自枯草杆菌的保守假定蛋白 YpgA, 也具有较高的同源性, 但比 2 型 idi 要小很多。

用 EMBL 数据库中找到的高度同源物, 即肺炎链球菌(*Streptococcus pneumoniae*) (AAG02453)(SEQ ID NO: 74), *Streptococcus pyogenes* (AAG02448)(SEQ ID NO: 75), 粪肠球菌(*Enterococcus faecalis*) (AAG02438) (SEQ ID NO: 76), 屎肠球菌(*Enterococcus faecium*)(AAG02443)(SEQ ID NO: 77), 溶血葡萄球菌(*Staphylococcus haemolyticus*) (AAG02427)(SEQ ID NO: 78), 表皮葡萄球菌(*Staphylococcus epidermis*) (AAG02433) (SEQ ID NO: 79), 15 金黄色葡萄球菌(*Staphylococcus aureus*) (AAG02422)(SEQ ID NO: 80), 肉葡萄球菌(*Staphylococcus carnosus*) (Q9ZB67)(SEQ ID NO: 81), 链霉菌种 CL190(Q9KWG1)(SEQ ID NO: 82), *Streptomyces griseolosporeus* (Q9KWF5)(SEQ ID NO: 83), 和布氏疏螺旋体(051626)(SEQ ID NO: 84), 进行副球菌种菌株 R114 细菌 HMG-CoA 合成酶的氨基酸排列(SEQ ID NO: 20 47)。EMBL/GenBank/DDBJ 数据库会根据括号内有机体的名称, 指定登录号。

使用来自其它细菌的直向同源(orthologous)蛋白, 即, 肺炎链球菌 (AAG02456)(SEQ ID NO: 85), *Streptococcus pyogenes* (AAG02451)(SEQ ID NO: 86), 粪肠球菌 (AAG02441)(SEQ ID NO: 87), 屎肠球菌 25 (AAG02446)(SEQ ID NO: 88), 溶血葡萄球菌(AAG02431)(SEQ ID NO: 89), 表皮葡萄球菌 (AAG02436)(SEQ ID NO: 90), 金黄色葡萄球菌 (AAG02425) (SEQ ID NO: 91), 链霉菌种 CL190(Q9KWG4)(SEQ ID NO: 92), *Streptomyces griseolosporeus*(Q9KWF8)(SEQ ID NO: 93), 和布氏疏螺旋体 (051629)(SEQ ID NO: 94), 进行副球菌种菌株 R114 细菌甲羟戊酸二磷酸脱羧酶的氨基酸 30 排列(SEQ ID NO: 53)。EMBL/GenBank/DDBJ 数据库会根据括号内有机体的名称, 指定登录号。

黄色粘球菌(*Myxococcus xanthus*)的两种蛋白, Tac 和 Taf(数据库登录号分别为 q9xb06 和 q9xb03)以及枯草杆菌蛋白 PksG(一种假定的多聚乙酰生物合成蛋白(数据库登录号为 p40830)), 与副球菌种菌株 R114 HMG-CoA 合成酶具有较高的同源性。副球菌种菌株 R114 HMG-CoA 合成酶与黄色粘球菌 Tac 和 Taf 蛋白的同源性高于副球菌种菌株 R114 HMG-CoA 合成酶与真核生物的同源性。细菌 HMG-CoA 合成酶和细菌甲羟戊酸二磷酸脱羧酶与其真核生物的直向同源物具有基本的特异性。古细菌(archaeal)HMG-CoA 合成酶组成了一个亲缘较远的酶组(Wilding 等, 同上), 而在古细菌内未找到甲羟戊酸二磷酸脱羧酶的直向同源物[Smit and Mushegian, *Genome Res.* 10: 1468-1484 (2000)]。

使用来自其它细菌的直向同源(orthologous)蛋白, 即, 肺炎链球菌(AAG02455)(SEQ ID NO: 95), *Streptococcus pyrogenes* (AAG02450)(SEQ ID NO: 96), 粪肠球菌 (AAG02440)(SEQ ID NO: 97), 屎肠球菌(AAG02445)(SEQ ID NO: 98), 溶血葡萄球菌 (AAG02430)(SEQ ID NO: 99), 表皮葡萄球菌 (AAG02435)(SEQ ID NO: 100), 金黄色葡萄球菌(AAG02424)(SEQ ID NO: 101), 链霉菌种 CL190(Q9KWG5)(SEQ ID NO: 102), *Streptomyces griseolosporeus*(Q9KWF9)(SEQ ID NO: 103), 和布氏疏螺旋体 (051631)(SEQ ID NO: 104)(Mvk), 和, 抗肺炎链球菌 (AAG02457)(SEQ ID NO: 105), *Streptococcus pyrogenes* (AAG02452)(SEQ ID NO: 106), 粪肠球菌(AAG02442)(SEQ ID NO: 107), 屎肠球菌(AAG02447)(SEQ ID NO: 108), 溶血葡萄球菌(AAG02432)(SEQ ID NO: 109), 表皮葡萄球菌(AAG02437)(SEQ ID NO: 110), 金黄色葡萄球菌 (AAG02426)(SEQ ID NO: 111), 链霉菌种 CL190(Q9KWG3)(SEQ ID NO: 112), *Streptomyces griseolosporeus*(Q9KWF7)(SEQ ID NO: 113), 和布氏疏螺旋体 (051630)(SEQ ID NO: 114) (pmk), 进行副球菌种菌株 R114 的甲羟戊酸激酶(Mvk)(SEQ ID NO: 49) 和 磷酸甲羟戊酸激酶 (Pmk)(SEQ ID NO: 51) 排列。EMBL/GenBank/DDBJ 数据库会根据括号内有机体的名称, 指定登录号。

细菌激酶之间的同源性要远远低于甲羟戊酸途径其它酶细菌直向同源物之间的同源性。副球菌种菌株 R114 的甲羟戊酸激酶(SEQ ID NO: 49)在氨基末端区域内具有一个 37 个氨基酸长度的插入, 这在其它甲羟戊酸激酶中是不存在的。细菌 Mvk 以及一些古细菌酶, 如来自闪烁古生球菌

(*Archaeoglobus fulgidus*), 嗜热碱甲烷杆菌和 *Pyrococcus abyssi* 的酶, 是与副球菌种菌株 R114 的 Mvk 同源性最高的酶。细菌磷酸甲羟戊酸激酶之间的同源性甚至低于细菌甲羟戊酸激酶之间的同源性。与副球菌种菌株 R114 Pmk(SEQ ID NO: 51)同源性最高的蛋白是来自古细菌的 Mvk, 如 *Aeropyrum* 5 *pernix*, *Pyrococcus horikoshii*, *M. thermoautotrophicum*, *P. abyssi* 和 *A.fulgidus*。由于在古细菌中未发现 Pmk(Smit and Mushegian, 同上), 这说明在两种磷酸化过程中可能是同一种激酶在起作用。

实施例 5: 副球菌种菌株 R114 甲羟戊酸途径各基因和 idi 基因在大肠杆菌的过表达

10 甲羟戊酸操纵子在大肠杆菌中的克隆和表达 一个λ克隆, 命名为克隆 16, 来自副球菌种菌株 R114 的λ文库(见实施例 4), 将其作为 PCR 扩增甲羟戊酸操纵子的模板。PCR 所用引物为 Mevop-2020 和 Mevop-9027(表 16)。

表 16.扩增副球菌种菌株 R114 甲羟戊酸操纵子所用引物

15

引物	序列 (5'-3')
Mevop-2020	GGGCAAGCTTGTCCACGGCACGACCAAGCA (SEQ ID NO:115)
Mevop-9027	CGTAATCCGCGGCCGCGTTTCCAGCGCGTC (SEQ ID NO:116)

20 将所获的 PCR 产物克隆到 TOPO-XL(Invitrogen, Carlsbad, CA, USA)中, 得到质粒 TOPO-XL-mev-op16。插入的部分带有甲羟戊酸操纵子, 用 HindIII 和 SacI 酶切该插入部分, 并克隆到 HindIII-SacI 处理过的载体 pBBR1MCS2[Kovach 等, Gene 166: 175-176(1995)] 中, 得到了质粒 pBBR-K-mev-op16。使用质粒 pBBR-K-mev-op16 转化电穿孔-感受态大肠杆

25 菌菌株 TG1[Stratagene, La Jolla, CA; Sambrook 等, In: Nolan, C.(ed), Molecular Cloning: A Laboratory Manual (Sencond Edition), p.A.12(1989)]。将两个代表性的阳性转化株(*E.coli* TG1/pBBR-K-mev-op16-1 和 *E.coli* TG1/pBBR-K-mev-op16-2)培养在含有 50mg/l 卡那霉素的 Luria Broth(LB, GibcoBRL, Life Technologies)中, 并使用实施例 1 所述方法, 检测 HMG-CoA

30 还原酶活性(由副球菌种菌株 R114 的 *mvaA* 基因编码)。大肠杆菌并不具有编码 HMG-CoA 还原酶的基因, 因此缺乏可检测的活性。 *E.coli*

TG1/pBBR-K-mev-op16 两个代表性转化株的粗提物具有可简易测定的 HMG-CoA 还原酶活性，说明克隆的 mvaA 基因得到了异源表达。

表 17.携带了克隆的副球菌种菌株 R114 甲羟戊酸基因簇的大肠杆菌 TG1 细胞粗提物中 HMG-CoA 还原酶活性

5

菌株	HMG-CoA 还原酶活性 (U/mg)
<i>E. coli</i> TG1	未测 ^a
<i>E. coli</i> TG1/ pBBR-K-mev-op16-1	0.25
<i>E. coli</i> TG1/ pBBR-K-mev-op16-2	0.78

10

^a: 小于 0.03U/mg

副球菌种菌株 R114 idi 基因和甲羟戊酸途径单个基因在大肠杆菌中的克隆与表达 使用表 18 所示引物, PCR 扩增副球菌种菌株 R114 甲羟戊酸操纵子各基因的编码区域。引物的设计使得 ATG 起始密码子构成了 NdeI 位点的后半(second)部分(切割识别位点 CATATG), 而在终止密码子后随即导入 BamHI 位点(GGATCC)。所有的 PCR 产物均克隆到 pCR[®]2.1-TOPO 载体中。所获载体的名字列在表 19 中。除了甲羟戊酸激酶基因, 所有的基因都含有 BamHI, NdeI 或 EcoRI 的限制性位点, 这些位点可以被消除, 以简化后面的克隆步骤。可使用 QuikChange[™] 定点突变试剂盒(Stratagene, La Jolla, CA, USA)和表 20 所示寡核苷酸, 通过导入沉默突变消除这些位点。用 BamHI 和 NdeI 从 TOPO 质粒中切除突变的编码区域, 并与 BamHI-NdeI 酶切的表达载体 pDS-His 和 pDS 相连。这些表达载体均来自 pDSNdeHis, 该载体在 EP 821,063 的实施例 2 中有所描述。通过缺失一段 857 个 bp 的 NheI 和 XbaI 片段(该片段带有沉默的氯霉素乙酰基转移酶基因), 从 pDSNdeHis 构建质粒 pDS-His。质粒 pDS 是通过用退火的引物 S/D-1 (5'AATTAAAGGAGGGTTTCATATGAA1 iCG)(SEQ ID NO: 117)和 S/D-2 (5'GATCCGAATTCATATGAAACCCTCCTTT)(SEQ ID NO: 118)替换一小段 EcoRI-BamHI 片段, 从 pDS-His 构建而来。

表 18.克隆甲羟戊酸操纵子基因的寡核苷酸

30

基因	正向引物		反向引物	
	名称	序列 (5'-3')	名称	序列 (5'-3')
5	<i>mvaA</i>	MvaA-Nde AAGGCCTCATATGATTTCC CCATACCCCGGT (SEQ ID NO:119)	MvaA-Bam	CGGGATCCTCATCGCTCCAT CTCCATGT (SEQ ID NO:120)
	<i>idi</i>	Idi-Nde AAGGCCTCATATGACCGA CAGCAAGGATCA (SEQ ID NO:121)	Idi-Bam	CGGGATCCTCATTGACGGAT AAGCGAGG (SEQ ID NO:122)
10	<i>hsc</i>	Hcs-Nde AAGGCCTCATATGAAAGT GCCTAAGATGA (SEQ ID NO:123)	Hcs-Bam	CGGGATCCTCAGGCCTGCCG GTCGACAT (SEQ ID NO:124)
	<i>mvk</i>	Mvk-Nde ¹ AAGGCCTCATATGAGCAC CGGCAGGCCTGAAGCA (SEQ ID NO:125)	Mvk-Bam ²	CGGGATCCTCATCCCTGCCC CGGCAGCGGTT (SEQ ID NO:126)
	<i>pmk</i>	Pmk-Nde AAGGCCTCATATGGATCA GGTCATCCGCGC (SEQ ID NO:127)	Pmk-Bam	CGGGATCCTCAGTCATCGAA AACAAAGTC (SEQ ID NO:128)
15	<i>mvd</i>	Mvd-Nde AAGGCCTCATATGACTGA TGCCGTCCGCGA (SEQ ID NO:129)	Mvd-Bam	CGGGATCCTCAACGCCCTC GAACGCGC (SEQ ID NO:130)

¹ 第二个密码子 TCA 改为 AGC(沉默突变-两个密码子均编码丝氨酸)。

² 最后密码子 GGC 改为 GGA(沉默突变-两个密码子均编码甘氨酸)。

表 19.表达质粒的名称和构建中间体

基因	在 pCR2.1-TOPO 中的 PCR 片段	第 1 个突变之后	第二个突变之后	pDS 中基因	pDS-His 中基因
25	<i>mvaA</i>	TOPO- <i>mvaA</i> -BB	TOPO- <i>mvaA</i> -B	pDS- <i>mvaA</i>	pDS-His- <i>mvaA</i>
	<i>idi</i>	TOPO-ORFX-B	TOPO- <i>idi</i>	pDS- <i>idi</i>	pDS-His- <i>idi</i>
	<i>hsc</i>	TOPO- <i>hcs</i> -EN	TOPO- <i>hcs</i> -N	pDS- <i>hcs</i>	pDS-His- <i>hcs</i>
	<i>mvk</i>	TOPO- <i>mvk</i>	n/a	pDS- <i>mvk</i>	pDS-His- <i>mvk</i>
	<i>pmk</i>	TOPO- <i>pmk</i> -B	TOPO- <i>pmk</i>	Nd	pDS-His- <i>pmk</i>
30	<i>mvd</i>	TOPO- <i>mvd</i> -B	TOPO- <i>mvd</i>	pDS- <i>mvd</i>	pDS-His- <i>mvd</i>

n/a: 不适用; nd: 未作

表 20: 定点突变所用寡核苷酸

	基因/位点	正向引物		反向引物	
		名称	序列 (5'-3')	名称	序列 (5'-3')
5	<i>mvaA</i> / <i>Bam</i> HI-1	Mva-B1up	CCGGCATTTCGGGCGGC ATCCAGGTCTCGCTG (SEQ ID NO:131)	Mva-B1down	CAGCGAGACCTGGATG CCGCCGAATGCCGG (SEQ ID NO:132)
	<i>mvaA</i> / <i>Bam</i> HI-2	Mva-B2up	CGTGCAGGGCTGGATT CTGTCCGAATACCCG (SEQ ID NO:133)	Mva-B2down	CGGGTATTCCGACAGA ATCCAGCCCTGCACG (SEQ ID NO:134)
10	<i>idi</i> / <i>Bam</i> HI	Idi-Bup2	GGGCTGCGCGCCGGCA TCCGGCATTTCGACG (SEQ ID NO:135)	Idi-Bdown2	CGTCGAAATGCCGGAT GCCGGCGCGCAGCCC (SEQ ID NO:136)
	<i>hcs</i> / <i>Eco</i> RI	Hcs-Eup	GGGTGCGACGGCGAG TTCTTCGATGCGCGG (SEQ ID NO:137)	Hcs-Edown	CCGCGCATCGAAGAAC TCGCCCCGTCGACCCC (SEQ ID NO:138)
15	<i>hcs</i> / <i>Nde</i> I	Hcs-Nup-c	CACGCCCCTCACATAC GACGAATACGTTGCC (SEQ ID NO:139)	Hcs-Ndown- c	GGCAACGTATTCGTCG TATGTGACGGCGGTG (SEQ ID NO:140)
	<i>pmk</i> / <i>Bam</i> HI	Pmk-Bup	GAGGCTCGGGCTTGGC TCCTCGGCGGCGGTG (SEQ ID NO:141)	Pmk-Bdown	CACCGCCGCCGAGGAG CCAAGCCCGAGCCTC (SEQ ID NO:142)
	<i>mvd</i> / <i>Bam</i> HI	Mvd-Bup	CGGCACGCTGCTGGAC CCGGGCGACGCCTTC (SEQ ID NO:143)	Mvd-Bdown	GAAGGCGTCGCCGGG TCCAGCAGCGTGCCG (SEQ ID NO:144)

20 用连接混合物转化携带 lacI(lac 抑制子)的大肠杆菌菌株 M15 [Villarejo 和 Zabin, J. Bacteriol. 120: 466-474(1974)]——所述大肠杆菌 M15 菌株含有质粒 pREP4(EMBL/GenBank 登录号 A25856)——并在含有 100mg/L 氨卞青霉素和 25mg/L 卡那霉素的 LB-Agar 平板上培养, 选择出重组细胞。再通过
25 PCR 验证含有正确甲羟戊酸操纵子基因插入的阳性克隆。

为表达插入的基因, 在含有 25mg/L 卡那霉素和 100mg/L 氨卞青霉素的 LB 培养基上, 37℃过夜培养各大肠杆菌菌株。第二天, 用 0.5ml 过夜培养物接种 25ml 新鲜培养基, 新的培养物在 37℃培养。当培养物的 OD₆₀₀ 达到 0.4 时, 加入异丙基-β-D-硫代半乳糖吡喃糖苷(IPTG), 使其终浓度为 1mM,
30 诱导克隆的基因表达, 继续(振荡)培养培养物 4 小时, 然后离心收集细胞。

粗提物制备, HMG-CoA 还原酶试验和 IPP 异构酶试验均按照实施例 1

进行。表 21 和 22 分别显示了重组大肠杆菌菌株体内 HMG-CoA 还原酶和 IPP 异构酶的活性。IPTG 诱导后，菌株 M15/pDS-mvaA 和 M15/pDS-idi 分别含有高水平的 HMG-CoA 还原酶或 IPP 异构酶活性。这说明大肠杆菌能够过量表达副球菌种菌株 R114 甲羟戊酸途径各基因(并以活性形式过量生产其同源基因的产物)。

表 21.过量表达来自副球菌种菌株 R114 的克隆 mvaA 基因的大肠杆菌菌株体内 HMG-CoA 还原酶活性的诱导

10

菌株/质粒	IPTG 诱导	HMG-CoA 还原酶活性 (U/mg)
M15/pDS-mvaA	-	8.34
M15/pDS-mvaA	+	90.0
M15/pDS-His-mvaA	-	1.74
M15/pDS-His-mvaA	+	2.95
M15/pDS-mvd ^a	-	0.05

15

^aM15/pDS-mvd 作为阴性对照

表 22.过量表达来自副球菌种菌株 R114 的克隆 idi 基因的大肠杆菌菌株体内 IPP 异构酶活性的诱导

接 92 页表 22

20

菌株/质粒	IPTG 诱导	IPP 异构酶活性(U/mg)
M15/pDS-idi	-	未测 ^b
M15/pDS-idi	+	22.0
M15/pDS-His-idi	-	未测
M15/pDS-His-idi	+	未测
M15/pDS-mvd ^a	-	未测

25

^aM15/pDS-mvd 作为阴性对照

^b<1U/mg

通过十二烷基硫酸钠聚丙烯酰胺凝胶电泳(SDS-PAGE)，分析用于酶试验的粗提物。对于菌株 E.coli M15/pDS-mvaA 和 E.coli M15/pDS-His-mvaA，预期分子量(36.3kD)蛋白的高度表达与否是与粗提物中所测的 HMG-CoA 还

30

原酶活性(表 21)相关联的。可这样解释 His-标记蛋白的缺乏: mRNA 或蛋白的不稳定性降低了转录或翻译水平的表达。E.coli M15/pDS-idi 和 E.coli M15/pDS-His-idi 的粗提物分别显示出预期分子量 37.3kD 和 39.0kD 蛋白的高度表达。但是, 只有 E.coli M15/pDS-idi 的 IPP 异构酶活性提高(表 22), 说明该酶的组氨酸标记形式在这些条件下不起作用。

通过对过量表达其它四种副球菌种菌株 R114 甲羟戊酸操纵子基因(hcs, pmk, mvk 和 mvd, 参照表 19)的大肠杆菌菌株粗提物的 SDS-PAGE 分析, IPTG 诱导后, 并未检测到该酶天然形式的高表达, 尽管有些表达不能被排除。另一方面, 均观察到所有四种 His-标记蛋白的高表达。

实施例 6: 通过过量表达 crtE 基因, 改进副球菌种菌株 R114 的玉米黄质产量

构建 pBBR-K-Zea4, pBBR-K-Zea4-up 和 pBBR-K-Zea4-down 以及这些质粒对副球菌种菌株 R114 玉米黄质产量的影响 从质粒 pZea-4 质粒中切出一段 8.3kb BamHI-EcoRI 片段, 该片段为副球菌种菌株 R1534 的类胡萝卜素 (crt)基因组簇。将含有 crt 基因的该片段连接到 BamHI 和 EcoRI 酶切过的载体 pBBR1MCS-2(GenBank 登录号#U23751)中, 得到 pBBR-K-Zea4。通过连接, 将质粒 pBBR-K-Zea4 导入副球菌种菌株 R114 中, 以检测其改进的玉米黄质产量。在摇瓶培养基(使用 362F/2 培养基, 见实施例 11)中, 测定对照菌株 R114 和 R114/pBBR-K-Zea4 的两个独立分离菌株玉米黄质产量。表 23 中的数据显示携带质粒 pBBR-K-Zea4 的两个重组菌株玉米黄质产量显著高于 R114, 且生产的比速率(specific rate)也更高(mg 玉米黄质/OD₆₆₀)。这说明在 pBBR-K-Zea4 的克隆插入片段中有一个或多个基因编码着限制副球菌种菌株 R114 玉米黄质产量的酶。

表 23.菌株 R114 和 R114/pBBR-K-Zea4 的玉米黄质产量

菌株	24 小时		48 小时		72 小时	
	ZXN ^a (mg/l)	Spec. Form. ^b	ZXN (mg/l)	Spec. Form.	ZXN (mg/l)	Spec. Form.
R114	54.5	2.2	81.7	4.1	78.1	4.5
R114/pBBR-K-Zea4 (克隆 4)	41.0	3.0	100.7	5.2	97.6	6.2
R114/pBBR-K-Zea4 (克隆 5)	41.1	3.1	110.5	5.7	102.1	6.5

^a 玉米黄质

^b 比形成(specific formation)(mg ZXN/1/OD₆₆₀)

为定位阳性作用，制备两种衍生质粒，使其含有 pBBR-K-Zea4 中存在的克隆插入片段的亚克隆区域。该 pBBR-K-Zea4 插入含有 ORF 5 和基因

5 atoB 和 crtE(Pasamontes 等，同上)，使该插入的“上游”区域被限制性酶 XbaI 和 AvrII 唯一位点侧接。用这两种酶酶切消化 pBBR-K-Zea4 并缺失“上游”区域，构建质粒 pBBR-K-Zea4-down。类似地，使用限制性酶 EcoRI 和 StuI，在 pBBR-K-Zea4 的克隆插入片段内缺失“下游”区域，构建质粒 pBBR-K-Zea4-up。通过接合，将这两个新的质粒转移到副球菌种菌株 R114

10 中。比较(摇瓶培养，与前述相同的条件)菌株 R114(宿主对照(host control))，R114/pBBR-K(空白载体对照)，R114/pBBR-K-Zea4-down 和 R114/pBBR-K-Zea4-up 的玉米黄质产量(表 24)。该数据清楚地表明：对玉米黄质产量的阳性作用是存在克隆片段多个拷贝的结果，所述片段含有 ORF 5,atoB 和 crtE，即是质粒 pBBR-K-Zea4-up 中存在的插入片段。从

15 pBBR-K-Zea4-up 构建一系列缺失质粒。将这些质粒逐一导入菌株 R114 并测定玉米黄质产量，结果表明：过量表达 crtE 基因引发了菌株 R114/pBBR-K-Zea4 和 pBBR-K-Zea4-up 玉米黄质产量的提高。该结果与副球菌种菌株 R114 体内限制玉米黄质产量的 GGPP 合成酶(由 crtE 编码)活性相一致。使用实施例 1 所述方法，发现 R114/pBBR-K-Zea4-up 粗提物的 GGPP

20 合成酶活性比 R114 高 2.6 倍。为直接证实改结果，按下述两部分，构建能够且仅仅能够过量表达 crtE 基因的新质粒。

表 24.带有 pBBR-K-Zea4 质粒缺失衍生物的菌株中玉米黄质产量

菌株	24 小时		48 小时		72 小时	
	ZXN ^a (mg/l)	Spec. Form. ^b	ZXN (mg/l)	Spec. Form.	ZXN (mg/l)	Spec. Form.
R114	35.0	1.2	75.7	4.1	73.9	4.4
R114/pBBR-K	32.0	1.5	59.3	3.1	63.3	3.9
R114/pBBR-K-Zea4-up	51.5	2.2	98.8	5.5	85.5	5.7
R114/pBBR-K-Zea4-down	41.6	1.8	63.4	3.3	66.4	3.9

^a 玉米黄质

^b 比形成(specific formation)(mg ZXN/1/OD₆₆₀)

- 构建表达载体 pBBR-K-PcrtE 和 pBBR-tK-PcrtE 用 BstXI 和 Bsu36I 酶切载体 pBBR1MCS-2, 将较大的片段与退火的寡核苷酸 MCS-2 上游 (5'TCAGAATTCGGTACCATATGAAGCTTGGATCCGGGG3')(SEQ ID NO: 145)和 MCS-2 下游(5'GGATCCAAGCTTCATATGGTACCGAATTC3')(SEQ ID NO: 146)相连, 得到载体 pBBR-K-Nde。副球菌种菌株 R114 中类胡萝卜素基因簇的 crtE 基因的 270bp 区上游片段含有假定的 crtE 启动子(PcrtE), 该启动子含有核糖体结合位点和 crtE 起始密码子(Pasamontes 等, 同上), 通过 PCR, 使用引物 crtE-up(5'GGAATTCGCTGCTGAACGCGATGGCG3')(SEQ ID NO: 147)和 crtE-down(5'GGGGTACCATATGTGCCTTCGTTGCGTCAGTC)(SEQ ID NO: 148), 从副球菌种菌株 R114 中扩增该上游片段。用 EcoRI 和 NdeI 酶切该 PCR 产物, 并插入 EcoRI-NdeI 酶切过的 pBBR-K-Nde 骨架, 形成质粒 pBBR-K-PcrtE。在引物 crtE-down 中包含一个 NdeI 位点, 该位点含有 crtE 的起始密码子 ATG。因此, 任何被导入的、含有存在于 NdeI 位点中的起始密码子的编码区域, 都可以使用 crtE 的核糖体结合位点被表达。用 BamHI 酶切 pBBR-K-PcrtE, 并插入退火的寡核苷酸 pha-t-up (5'GATCCGGCGTGTGCGCAATTAAATTGCGCACACGCCCCCTGCGTTTAAAC3')(SEQ ID NO: 149) 和 pha-t-down (5'GATCGTTTAAACGCAGGGGGCGTGTGCGCAATTAAATTGCGCACACGCCG3')(SEQ ID NO: 150)。测序验证插入的片段, 将下述的质粒命名为 pBBR-tK-PcrtE: 寡核苷酸插入的方向使得重组成的 BamHI 位点靠近 PcrtE 启动子。该插入序列带有在副球菌种菌株 R114 phaA 和 phaB 基因之间发现的(见实施例 10)、假定的转录终止子, 因此该序列应该能够适当地终止从 PcrtE 启动子开始的转录。
- 25 构建质粒 pBBR-K-PcrtE-crtE-3 为构建多元拷贝质粒用于提高 crtE 基因在副球菌种菌株 R114 宿主中的表达, 使用引物 crtE-Nde (5'AAGGCCTCATATGACGCCCAAGCAGCAATT3'J)(SEQ ID NO: 151)和 crtE-Bam(5'CGGGATCCTAGGCGCTGCGGCGGATG3')(SEQ ID NO: 152), 从质粒 p59-2(Pasamontes 等, 同上)中扩增 crtE 基因。将扩增的片段克隆到载体 pCR®2.1-TOPO 中, 得到质粒 TOPO-crtE。将 TOPO-crtE 的 NdeI-BamHI 片段亚克隆到 NdeI-BamHI 酶切过的质粒 pBBR-K-PcrtE 中, 得到

pBBR-K-PcrtE-crtE。最后，用 pBBR-K-Zea4-up 的较小 BglII 片段替换 pBBR-K-PcrtE-crtE 的较小 BglII 片段，构建 pBBR-K-PcrtE-crtE-3。通过电穿孔，将质粒 pBBR-K-PcrtE-crtE-3 转移到副球菌种菌株 R114 中。使用实施例 1 所述方法，测定 R114/pBBR-K-PcrtE-crtE-3 粗提物中 GGPP 合成酶活性比菌株 R114 高 2.9 倍。活性提高程度类似于在 R114/pBBR-K-Zea4-up 观察到的结果。表 25 显示菌株 R114/pBBR-K-PcrtE-crtE-3 玉米黄质测量与 R114/pBBR-K-Zea4-up 基本相同。

表 25.比较菌株 R114/pBBR-K-PcrtE-crtE-3 和 R114/pBBR-K-Zea4-up 的玉米黄质产量

菌株	24 小时		48 小时		72 小时	
	ZXN ^a (mg/l)	Spec. Form. ^b	ZXN (mg/l)	Spec. Form.	ZXN (mg/l)	Spec. Form.
R114	49.0	1.6	83.9	3.3	97.8	4.3
R114/pBBR-K	42.6	1.8	73.7	3.8	88.8	4.9
R114/pBBR-K-PcrtE-crtE-3	64.6	2.9	127.0	5.8	165.6	8.5
R114/pBBR-K-Zea4-up	64.7	2.9	118.0	5.9	158.0	10.1

^a 玉米黄质
^b 比形成(specific formation)(mg ZXN/1/OD₆₆₀)

实施例 7: 副球菌种菌株 R114 甲羟戊酸操纵子各独立基因在天然宿主副球菌种菌株 R114 中的表达

副球菌种菌株 R114 甲羟戊酸操纵子各独立基因在副球菌种菌株 R114 宿主中的表达 用 BamHI 和 NdeI 酶切 TOPO-质粒中突变的甲羟戊酸操纵子基因编码区域(见实施例 5)，并与 BamHI-NdeI 酶切过的载体 pBBR-tK-PcrtE-crtE(见实施例 6)相连。通过电穿孔，将获得的质粒 pBBR-tK-PcrtE-mvaA，pBBR-tK-PcrtE-idi，pBBR-tK-PcrtE-hcs, pBBR-tK-PcrtE-mvk, pBBR-tK-PcrtE-pmk 和 pBBR-tK-PcrtE-mvd 导入副球菌种菌株 R114 中。在含有 50mg/l 卡那霉素的琼脂培养基上选择转化株，并 PCR 验证。

为说明由质粒提供的甲羟戊酸途径各基因可在天然宿主副球菌种菌株 R114 内表达，比较菌株 R114/pBBR-K(对照)和 R114/pBBK-tK-PcrtE-mvaA 粗提物的 HMG-CoA 还原酶活性(所用方法如实施例 1 所述)。菌株

R114/pBBR-K 和 R114/pBBK-tK-PcrtE-mvaA 的 HMG-CoA 比活性分别为 2.37U/mg 和 6.0U/mg。这样在多元质粒上存在(以及从 PcrtE 启动子表达)mvaA 基因, 会使 HMG-CoA 还原酶活性由带有空载体 pBBR-K 的 R114 基础(即, 染色体编码的)活性提高 2.5 倍。

- 5 实施例 8: 构建“Mini-操纵子”用于同时过量表达甲羟戊酸途径的克隆基因和副球菌种菌株 R114 的 crtE 基因

质粒构建 如实施例 6 所示, 将质粒 pBBR-K-PcrtE-crtE-3 导入副球菌种菌株 R114, 引起玉米黄质产量提高, 说明 GGPP 合成酶活性是菌株 R114 中玉米黄质生物合成的限速因素。实施例 7 进一步显示了编码甲羟戊酸途径各酶的基因可在天然宿主副球菌种菌株 R114 中过量表达, 导致各编码酶活性提高。但是, 携带着质粒(所述质粒含有甲羟戊酸操纵子各个独立基因)的副球菌种菌株 R114 重组菌株均未显示出提高的玉米黄质产量。可能是玉米黄质途径中下游的“瓶颈”(GGPP 合成酶)屏蔽了甲羟戊酸途径操纵子各基因在副球菌种菌株 R114 中过量表达所带来的益处。制备出能够同时过量表达甲羟戊酸途径各个基因(或这些基因的组合)和 crtE 的质粒可以释放整个玉米黄质生物合成途径中的所有限速步骤, 从而提高玉米黄质产量。下面的部分描述了“Mini-操纵子”的构建, 用于使 crtE 和编码甲羟戊酸途径 5 种酶的各个基因共过量表达。

使用 BamHI 和 NdeI 酶, 从各个 TOPO-质粒(实施例 5 和 6 中所述)中切出 crtE, mvaA, idi 和 mvk 基因, 并与 BamHI-NdeI 酶切过的载体 pOCV-1(实施例 12 所述)连接。crtE 基因编码区域的最后核苷酸并非腺嘌呤, 而且, 具有 TAG 而不是 TGA 终止密码子, 终止密码子和 BamHI 位点之间的距离不合适。因此, crtE 的末端并不符合操纵子构建载体的要求(参见实施例 12), 而且 crtE 必须是用 pOCV-1-crtE 构建的任何操纵子中的最后一个基因。为满足将腺嘌呤作为第二个密码子的第一个核苷酸、且作为最后密码子的最后一个核苷酸的条件, 必须在甲羟戊酸操纵子的三个基因中引入突变。pmk 的第二个密码子, GAT, 编码 Asp, 被改变为编码 Asn 的 AAT。mvd 的最后密码子以 T 结尾, 而 pmk 和 hcs 以 C 结尾。将这些核苷酸突变为 A, 引起沉默突变, pmk 除外(pm k 的最后氨基酸由 Asp 变为 Glu)。设计寡核苷酸, 通过 PCR 导入必要的改变。这些 PCR 反应所用的寡核苷酸序列和模板列在表 26 中。将所有的 PCR 产物均克隆到 pCR[®]2.1-TOPO 载体中, 形成质粒

TOPO-mvd^{OCV}, TOPO-pmk^{OCV} 和 TOPO-hcs^{OCV}。用 NdeI 和 BamHI 酶切插入的部分, 并与 NdeI-BamHI 尘粒过的 pOCV-2(见实施例 2)相连。组装各个“Mini- 操纵子”的最后克隆步骤与前述类似, 可由构建 pBBR-K-PcrtE-mvaA-crtE-3 的代表性方案说明。

5 表 26.构建质粒 TOPO-mvd^{OCV}, TOPO-pmk^{OCV} 和 TOPO-hcs^{OCV} 的 PCR 所用寡核苷酸和模板

10

基因	正向引物		反向引物		模板
	名称	序列 (5'-3')	名称	序列 (5'-3')	
<i>Hcs</i>	Hcs-Nde	AAGGCCTCATATGAAA GTGCCTAAGATGA (SEQ ID NO:123)	Hcs-mut3	CCGGATCCTCATGCC TGCCGGTCGACATAG (SEQ ID NO:153)	pBBR-tK-PcrtE- <i>hcs</i>
<i>Pmk</i>	Pmk-mut5	GAAGGCACATATGAAT CAGGTCATCCGCGC (SEQ ID NO:154)	Pmk-mut3	GCCGGATCCTCATTC ATCGAAAACAAGTCC (SEQ ID NO:155)	pBBR-tK-PcrtE- <i>pmk</i>
<i>Mvd</i>	Mvd-Nde	AAGGCCTCATATGACT GATGCCGTCCGCGA (SEQ ID NO:129)	Mvd-mut3	ACGCCGGATCCTCAT CGCCCCTCGAACGGC (SEQ ID NO:156)	pBBR-tK-PcrtE- <i>mvd</i>

15

实施例 9: 副球菌种菌株 R114 中编码 FPP 合成酶的 *ispA* 基因的克隆和测序

由于 FPP 合成酶存在在副球菌种菌株 R114 玉米黄质生物合成中的中间途径中, 通过提高 *ispA* 基因的量从而提高该酶活性的做法对于改进玉米黄质的产量十分有效。因此, 按下述克隆并测序副球菌种菌株 R114 的 *ispA* 基因。从公共数据库中获得 6 种细菌 FPP 合成酶的氨基酸序列。这些序列具有几个高度保守的区域。表 27 列出了这些区域中的两个以及 PCR 所用寡核苷酸。使用副球菌种菌株 R114 DNA 作为模板, 与寡核苷酸 GTT-1 和 GTT-2 进行 PCR, 形成了预期大小的产物。将该 PCR 产物克隆到载体 pCR[®]2.1-TOPO 中并测序。使用克隆的片段作为副球菌种菌株 R114 DNA Southern 分析中的探针, 发现其与一段大约 1.9kb 的 BamHI-NcoI 片段杂交。用 BamHI 和 NcoI 酶切副球菌种菌株 R114 DNA, 琼脂糖凝胶电泳分离各片段。在 15.-2.1kb 之间的区域被分离, 并克隆到克隆载体的 BamHI 和 NcoI 位点。然后使用 *ispA*-PCR 片段作为探针, 筛选该部分文库, 分离得到两个阳性克隆。测序结果证实了两个克隆质粒均含有 *ispA* 基因。*ispA* 基因(SEQ

20

25

30

ID NO: 159)的上游是核酸外切酶 VII 的小亚基, XseB(SEQ ID NO: 158), 下游是编码 1-脱氧木酮糖-5-磷酸合成酶的 dxs 基因(SEQ ID NO: 160)。这与大肠杆菌中发现的基因排列相同。NcoI-BamHI 片段序列如 SEQ ID NO: 157 所示, XseB, IspA 和 Dxs 的氨基酸序列分别如 SEQ ID NO: 158, SEQ ID NO: 159 和 SEQ ID NO: 160 所示。ispA 的起始密码子可以是 GTG 或 ATC, 其可在天然 IspA 的氨基末端, 分别形成两个或一个甲硫氨酸残基。

使用实施例 5-7 所述的常规克隆策略, 可以构建一个新的质粒, pBBR-tK-PcrtE-2, 其能够在天然宿主副球菌种菌株 R114 中过量表达 ispA 基因。利用电穿孔, 将该质粒导入菌株 R114, 并 PCR 验证转化株。在 362F/2 培养基(实施例 11)上培养三种具有代表性的转化株和对照菌株 (R114/pBBR-K), 制备粗提物, 并根据实施例 1 所述方法, 试验测定 ispA 基因产物 FPP 合成酶的活性。菌株 R114/pBBR-K 中基础(染色体编码的)FPP 合成酶比活性为 62.6U/mg。三种转化株中 FPP 合成酶的比活性为 108.3U/mg(增长 73%), 98.5U/mg(增长 57%)和 83.8U/mg(增长 34%), 说明: 在副球菌种菌株 R114 中, 以活性形式过量表达 ispA 基因并且过量生成其产物 FPP 合成酶。

表 27.用于两种保守的细菌 IspA 基因多肽的寡核苷酸

20	肽 1	
	<i>Bradyrhizobium japonicum</i>	Val His Asp Asp Leu Pro (SEQ ID NO:161)
	<i>Rhizobium</i> sp. strain NGR234	Val His Asp Asp Leu Pro (SEQ ID NO:162)
	<i>Bacillus stearothermophilus</i>	Ile His Asp Asp Leu Pro (SEQ ID NO:163)
	<i>Bacillus subtilis</i>	Ile His Asp Asp Leu Pro (SEQ ID NO:164)
	<i>Escherichia coli</i>	Ile His Asp Asp Leu Pro (SEQ ID NO:165)
	<i>Haemophilus influenzae</i>	Ile His Asp Asp Leu Pro (SEQ ID NO:166)
25	Oligonucleotide GTT-1 (5'-3')	tc cay gay gay ctg cc (SEQ ID NO:167)
	肽 2	
	<i>Bradyrhizobium japonicum</i>	Asp Asp Ile Leu Asp (SEQ ID NO:168)
	<i>Rhizobium</i> sp. strain NGR234	Asp Asp Ile Leu Asp (SEQ ID NO:169)
30	<i>Bacillus stearothermophilus</i>	Asp Asp Ile Leu Asp (SEQ ID NO:170)
	枯草杆菌	Asp Asp Ile Leu Asp (SEQ ID NO:171)
	大肠杆菌	Asp Asp Ile Leu Asp (SEQ ID NO:172)
	<i>Haemophilus influenzae</i>	Asp Asp Ile Leu Asp (SEQ ID NO:173)
	寡核苷酸 GTT-2 (5'-3') 的反向互补序列	gay gay atc ctg gay (SEQ ID NO:174)

Y=C 或 T

实施例 10: 副球菌种菌株 R114 中编码乙酰-CoA 乙酰基转移酶各基因的克隆和测序

IPP 生物合成中第一个进行的步骤是利用 HMG-CoA 合成酶, 将乙酰
5 -CoA 和乙酰乙酰-CoA 缩合为羟基甲基戊二酸单酰-CoA(HMG-CoA)。底物
乙酰乙酰-CoA 是通过乙酰-CoA 乙酰基转移酶(又称为乙酰乙酰-CoA 硫解酶
或 β -酮硫解酶), 缩合两分子乙酰-CoA 而形成。由于该反应通过甲羟戊酸途
径, 将中间代谢(乙酰-CoA 处)和类异戊二烯的生物合成相连接, 因此, 通
10 过基因扩增提高乙酰-CoA 乙酰基转移酶的活性从而提高体内的碳流向类胡
萝卜素和其它类异戊二烯的方法十分有效。在副球菌种菌株 R114 中, 至少
有两种基因——atoB 和 phaA——编码乙酰-CoA 乙酰基转移酶。副球菌种
菌株 R1534(US 6,087,152)和 R114(本文)中, atoB 基因的末尾位于 crtE 起始
部位上游的 165 个核苷酸处。副球菌种菌株 R1534 中, 被编码的乙酰-CoA
乙酰基转移酶的 atoB 基因核苷酸序列和相应的氨基酸序列分别如 SEQ ID
15 NO: 175 和 SEQ ID NO: 176 所示。

使用实施例 5 所述的常规克隆策略, atoB 基因被克隆到质粒 pDS 和
pDS-His 中。将新的质粒——pDS-atoB 和 pDS-His-atoB——导入大肠杆菌
M15。在具有 IPTG 诱导(如实施例 5 所述)和不具有 IPTG 诱导的条件下, 培
养获得的菌株 M15/pDS-atoB 和 M15/pDS-His-atoB, 制备粗提物用于乙酰
20 -CoA 乙酰基转移酶试验(方法如实施例 1 所述)和 SDS-PAGE 分析。在
M15/pDS-atoB 和 M15/pDS-His-atoB(带诱导)的粗提物中, 乙酰-CoA 乙酰基
转移酶的比活性分别为 0.2U/mg 和 13.52U/mg。在不含质粒的大肠杆菌中测
得的基础活性为 0.006U/mg。IPTG 诱导后, atoB 基因产物——乙酰-CoA 乙
酰基转移酶在大肠杆菌 M15 中过量生成。天然(M15/pDS-atoB)和 His-标记
25 的(M15/pDS-His-atoB)形式中均过量生成。M15/pDS-His-atoB 过量生成的程
度要高得多, 该结果与两个(诱导的)菌株粗提物中所测到的乙酰-CoA 乙酰
基转移酶活性一致。

乙酰乙酰-CoA 还是聚羟基烷羧酸酯(PHA)生物合成所进行的第一个步
骤中的底物。在许多细菌中, 与 PHA 生物合成相关的基因都被分入各个操
30 纵子的组中[Madison 和 Huisman, Microbol. Mol. Biol. Rev., 63:21-53(1999)]。
在 *Paracoccus denitrificans* 中, phaA 和 phaB 基因, 分别编码乙酰-CoA 乙酰

基转移酶和乙酰乙酰-CoA 还原酶, 这两个基因成簇于同一个操纵子中 [Yabutani 等, FEMS Microbiol. Lett. 133: 85-90(1995)], 而编码该途径最后一个酶——聚(3-烷羧酸酯) (3-hydroxyalkanoate)合成酶——的 phaC 则不属于该操纵子 [Ueda 等, J.Bacteriol. 178: 774-779(1995)]。使用根据 *P.denitrificans* phaA 和 phaC 基因序列合成的引物, 获得含有副球菌种菌株 R1534 phaA 部分和副球菌种菌株 R114 phaC 的 PCR 片段。然后将该 PCR 片段作为探针用于筛选副球菌种菌株 R114 的 λ 文库(见实施例 4)。分离出了几个与 phaA 和 phaC 探针杂交的 λ 克隆, 通过序列分析证实 phaA 和 phaC 基因是否存在于插入片段中。通过亚克隆和测序, 进一步分析 phaA λ -克隆, 而在 phaA 基因的下流发现了 phaB。因此, 如同 *P.denitrificans* 的情形, phaA 和 phaB 成簇在一起而 phaC 基因位于基因组中其它的位置。副球菌种菌株 R114 phaAB 簇的核苷酸序列和推定的乙酰-CoA 乙酰基转移酶(PhaA)的氨基酸序列分别如 SEQ ID NO: 177, SEQ ID NO: 178 和 179 所示。PHA 生物合成操纵子中相关基因成簇说明: 当细胞生成聚(3-烷羧酸酯)时, 至少 phaA 和 phaB 是同时被表达的。另一方面, 在副球菌种菌株 R114 中, 在 phaA 和 phaB 基因之间发现了一个假定的转录停止信号, 而这在 *P.denitrificans* phaAB 操纵子中是不存在的(Yabutani 等, 同上)。这样, 这两个基因的表达在副球菌种菌株 R114 中可能不是相随的。

使用实施例 5 所述的常规克隆策略, phaA 基因被克隆到质粒 pDS-His 中。将新的质粒——pDS-His-phaA——导入大肠杆菌 M15。在具有 IPTG 诱导(如实施例 5 所述)和不具有 IPTG 诱导的条件下, 培养获得的菌株 M15/pDS-His-phaA, 并制备粗提物用于 SDS-PAGE 分析。在 IPTG 诱导后, 在大肠杆菌 M15 宿主中过量形成了克隆的 His 标记的副球菌种菌株 R114 PhaA。

以上已经提及了扩增 atoB 或 phaA 基因(编码乙酰-CoA 乙酰基转移酶)在玉米黄质产量上的潜在优势。另外, 降低或消除乙酰乙酰-CoA 还原酶(phaB 基因产物)的活性以避免一些在体内形成的乙酰乙酰-CoA 转向 PHA 途径, 也可以利于提高玉米黄质产量。可通过基因取代技术(特别是用无活性形式 phaB 基因取代野生型基因组中的 phaB 基因)或经典诱变和筛选技术, 制备得到缺乏 phaB 活性的一些副球菌种菌株 R114 突变株。

实施例 11: 使用副球菌种菌株 R114 衍生的突变株, 工业生产番茄红

素的范例

番茄红素是一种红色的类胡萝卜素，是由菌株 R-1512 及其突变株 R1534 和 R114 所代表的新副球菌种中，玉米黄质生物合成中的中间体。由于番茄红素本身是一种重要的商品，所以，十分有必要测试新副球菌种通过工业发酵生产番茄红素的潜能。为获得玉米黄质生物合成被阻断的、能够积累番茄红素的突变株，用紫外线(UV)对副球菌种菌株 R114 进行诱变，然后筛选红色菌落。紫外诱变按下述进行：在 ME 培养基(见实施例 2)中过夜培养菌株 R114。将该过夜培养物再次培养在新鲜的 ME 培养基中(起始 OD₆₁₀=0.1)，并在 28℃温育 3 小时。离心该培养物，用 20mM 磷酸钾溶液(pH7.2)洗涤沉淀。再次离心后，重悬浮沉淀，使终 OD₆₁₀ 为 0.1。将 4ml 细胞悬浮液置于灭菌的 100ml 玻璃烧杯中。以 1450μW/cm² 的通量对这一薄层细胞悬浮液进行预先测定最佳时间长度的 UV 照射。在照射中，通过烧杯内的曲别针和磁力搅拌器，混合该细胞悬浮液。将诱变后的细胞悬浮液(和未突变的对照)培养在 362/F2 琼脂培养基上(表 28)。在诱变前和诱变后，在悬浮液上进行三次存活平板计数。然后将该板在 28℃孵育 4-5 天，为菌落评分。通过再-划线(re-streaking)鉴定及纯化几个红色菌落(假定的产番茄红素菌落)。一个突变株，命名为 UV7-1，用于进一步评价番茄红素产量。

表 29 显示了对照菌株 R114 及其突变株 UV7-1 的玉米黄质产量和番茄红素产量。菌株 R114 只生成玉米黄质。突变株 UV7-1 生产大量的番茄红素，只生成少量玉米黄质，说明在 UV7-1 中(假设在 crtE 基因中)的突变阻断并不完全。这些结果说明可以从副球菌种菌株 R114 衍生出产番茄红素的菌株。

表 28.培养基 362F/2 的配方和制备

组份	数量
葡萄糖一水合物	33g
酵母提取物	10g
胰蛋白胨	10g
NaCl	5g
MgSO ₄ · 7H ₂ O	2.5g
琼脂(用于固体培养基)	20g
蒸馏水	To 932 ml
调 pH 为 7.4	

过滤(液体培养基)或高压(固体培养基)灭菌	
加入 2.5ml 各微量元素溶液、NKP 溶液和 CaFe 溶液	
微量元素溶液	每升蒸馏水内数量
$(\text{NH}_4)_2\text{Fe}(\text{SO}_4)_2 \cdot 6\text{H}_2\text{O}$	80g
$\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$	6g
$\text{MnSO}_4 \cdot \text{H}_2\text{O}$	2g
$\text{NiSO}_4 \cdot 6\text{H}_2\text{O}$	0.2g
EDTA	6g
过滤灭菌	
NKP 溶液	每升蒸馏水内数量
K_2HPO_4	250g
$(\text{NH}_4)_2\text{HPO}_4$	300g
过滤灭菌	
CaFe 溶液	每升蒸馏水内数量
$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$	75g
$\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$	5g
浓缩 HCl	3.75ml
过滤灭菌	

表 29.副球菌种菌株 R114 及其突变株 UV7-1 的玉米黄质产量和番茄红素产量

	玉米黄质 (mg/l)	番茄红素(mg/l)
24 小时		
5 R114	36.65	0
UV7-1	3.85	20.85
48 小时		
R114	72.95	0
UV7-1	5.75	85.95
72 小时		
10 R114	83.9	0
UV7-1	5.85	124.55

实施例 12: 使用副球菌种菌株 R114 衍生的突变株, 工业生产虾青素的范例

虾青素是商业中一种重要的类胡萝卜素, 主要用于水产养殖。EP 0872,552 说明: 通过导入含有副球菌种菌株 R1534 和 *Paracoccus carotinifaciens* E-396^T 克隆类胡萝卜素(crt)基因组合的质粒, 可在大肠杆菌中形成虾青素。同时, 克隆的 crt 基因(crtEBIYZ)和 crtW(β -胡萝卜素 β -4 加氧酶)编码了从 FPP 通过玉米黄质生成虾青素的整个生物合成途径。*P.carotinifaciens* E-396 crtW, 副球菌种 R1534 crtZ 和副球菌种 R1534 crtE 基因及编码多肽如(SEQ ID NO: 180 和 181(crtW); 182 和 184(crtZ); 及 184 和 185(crtW))所示。但是, 在副球菌种菌株 R114 宿主科(family)内未表现出可获得虾青素产量。为说明利用源自菌株 R114 的重组菌株进行虾青素合成, 按下述将克隆的 crtW 基因(SEQ ID NO: 180)导入菌株 R114 中。

表 30. 实施例 12 中用于本文所述的 PCR 引物

15

引物名称	序列
CrtW-Nde	5' AAGGCCTCATATGAGCGCACATGCCCTGCC 3' (SEQ ID NO:186)
CrtW-Bam	5' CGGGATCCTCATGCGGTGTCCCCCTTGG 3' (SEQ ID NO:187)
CrtZ-Nde	5' AAGGCCTCATATGAGCACTTGGGCCGCAAT 3' (SEQ ID NO:188)
CrtZ-Bam	5' AGGATCCTCATGTATTGCGATCCGCCCTT 3' (SEQ ID NO:189)

20

使用引物 crtW-Nde 和 crtW-Bam(表 30), 通过 PCR, 从 *Paracoccus carotinifaciens* 菌株 E-396^T(Tsubokura 等, 同上, EP 872,554)的克隆 crt 簇扩增 crtW 基因。引物的设计使得 ATG 起始密码子构成了 NdeI 位点的后半部分(second half) (酶切识别位点 CATATG), 而在终止密码子后随即导入 BamHI 位点(GGATCC)。PCR 产物被克隆到 pCR[®]2.1-TOPO 载体中, 得到质粒 TOPO-crtW。用 NdeI 和 BamHI 酶切 crtW 基因, 并亚克隆到 NdeI-BamHI 酶切过的载体 pBBR-K-PcrtE(实施例 6 所述)中, 得到 pBBR-K-PcrtE-crtW。

使用标准接合方法{大肠杆菌菌株 S17[Priefer 等, J.Bacteriol. 163: 324-330 (1985)]为供体有机体}, 将质粒 pBBR-K-PcrtE-crtW 转移到副球菌种菌株 R114 中。在含有 50mg/l 卡那霉素的 362F/2 琼脂培养基上选择转接合子, 并在相同培养基上通过划线将其纯化。PCR 验证质粒

pBBR-K-PcrtE-crtW 在菌株中的存在。除了使用 362F/2 培养基替代 ME 培养基之外，其余如实施例 1 和 2 所述，测定摇瓶培养物中菌株 R114，R114/pBBR-K(空载体对照)和 R114/pBBR-K-PcrtE-crtW 的类胡萝卜素产量。结果如表 31 所示。对照菌株 R114 和 R114/pBBR-K 仅生成玉米黄质。在菌株 R114/pBBR-K-PcrtE-crtW 中，玉米黄质完全被质粒编码的β-胡萝卜素β-4 加氧酶消耗。尽管生成了虾青素，但是，其它两种酮类胡萝卜素(ketocarotenoid)，福寿草黄素(adonixanthin)和角黄素积累的水平很高。这说明体内β-胡萝卜素羟化酶(菌株 R114 中由染色体 crtZ 基因编码)和克隆的β-胡萝卜素β-4 加氧酶(CrtW)不平衡。

为证实该假设，制备两种新的质粒，其在 Mini 操纵子中同时含有 crtZ 和 crtW 基因。使两种构建体中基因的顺序不相同(即 crtZ-crtW 和 crtW-crtZ)，以试图形成不同的 crtZ 和 crtW 基因表达比率。新载体的构建需要按下述组装特定的克隆载体组。设计一系列操纵子构建载体(根据载体 pCR®2.1-TOPO)，以促进将基因(在此 crtZ 和 crtW)组装到操纵子中。目的基因必须具有 ATG 起始密码子和 TGA 终止密码子，起始密码子应位于 NdeI 位点中(CATATG)，终止密码子应直接位于 BamHI 位点之后。

表 31.含有表达单独 crtW 基因、以及 crtW 和 crtZ 基因组合的质粒的副球菌种菌株 R114 中虾青素产量

菌株	24 小时					
	ZXN	ADN	CXN	AXN	Total	Sp. Form. ^a
R114	46.5	0	0	0	46.5	2.1
R114/pBBR-K	38.8	0	0	0	41.4	2.2
R114/pBBR-K-PcrtE -crtW	0	13.0	21.8	2.3	37.5	2.1
R114/pBBR-K-PcrtE -crtWZ	0	14.9	29.5	1.3	45.6	2.1
R114/pBBR-K-PcrtE -crtZW	0	18.0	20.4	7.3	45.65	2.1
菌株	48 小时					
	ZXN	ADN	CXN	AXN	Total	Sp. Form. ^a
R114	72.6	0	0	0	74.4	2.8
R114/pBBR-K	70.1	0	0	0	70.1	3.1
R114/pBBR-K-PcrtE -crtW	0	26.7	22.0	26.9	75.5	3.9
R114/pBBR-K-PcrtE -crtWZ	0	30.9	27.2	34.8	92.9	4.0
R114/pBBR-K-PcrtE -crtZW	0	15.7	11.2	58.3	85.1	4.1

5

	72 小时					
R114	82.5	0	0	0	82.5	5.3
R114/pBBR-K	82.9	0	0	0	82.9	5.1
R114/pBBR-K-P <i>crtE</i> - <i>crtW</i>	0	19.7	17.0	46.8	83.5	5.2
R114/pBBR-K-P <i>crtE</i> - <i>crtWZ</i>	0	28.7	26.4	43.8	98.8	6.1
R114/pBBR-K-P <i>crtE</i> - <i>crtZW</i>	0	18.3	14.4	66.3	98.9	5.9

^aZXN, 玉米黄质; AND 福寿草黄素(adonixanthin); CXN 角黄素; AXN 虾青素;

^b比形成, 表达为 mg/l 总类胡萝卜素/OD₆₆₀

10

另外, 起始密码子后的第一个核苷酸和终止密码子之前的最后一个核苷酸必须是腺嘌呤, 该基因必须缺乏至少一个下述酶的位点: BsgI, BseMII, BseRI 和 GsuI. 构建四个操纵子构建载体, 其区别在于其多接头序列(SEQ ID NO: 190-197)的排列。前两个酶的分裂位点在 NdeI 位点中。后两个酶的分裂位点在 BamHI 位点之前。pOCV-1 和 pOCV-4 中的 BseRI 位点并不唯一, 且不能用于构建操纵子。

15

首先, 将待组装入操纵子的各基因单独地插入到适当操纵子构建载体中的 NdeI 和 BamHI 位点之间。然后, 将带有想象操纵子上游基因的所获质粒用多接头末尾两个酶之一和一种在载体骨架中具有独特位点的酶进行酶切。用多接头末尾两个酶之一和用于第一个质粒(含有期望的上游基因)相同的酶(在载体骨架中具有独特位点), 酶切含有想象操纵子下游基因的质粒。分离出带有各基因的片段, 并连接, 获得在 NdeI 和 BamHI 位点之间带有这两个基因的 pOCV 质粒。可以类似的方式加入更多的基因。组装的各基因相互重叠, 使得上游基因 TGA 终止密码子的前两个核苷酸 TG 与下游基因起始密码子 ATG 的后两个核苷酸重合。在副球菌种菌株 R1534 类胡萝卜素(crt)操纵子(crtZYIB)的所有基因之间均发现了这种重叠(Pasamontes 等, 同上)。

20

25

30

pOCV 骨架来自 pCR[®]2.1-TOPO。通过定点突变, 将位点 CTCAG 改变为 CACAG, 消除了复制必须区域, ColE1 起始位点上游中的 BseMII 位点。通过去除一段 0.8kb 的 DdeI-Asp700 片段, 消除剩余的三个 BseMII 位点和一个 GsuI 位点。填平 DdeI 的凹缺末端后, 该载体变为平头末端连接。通过使寡核苷酸与适当的 5'突出端一起退火, 在 BamHI 和 XbaI 位点之间插入多

接头。

使用操纵子构建载体 pOCV-2, 按下述构建质粒 pBBR-K-PcrtE-crtZW。使用引物 crtZ-Nde 和 crtZ-Bam(表 30), 通过 PCR, 从副球菌种菌株 R114 扩增 crtZ 基因。引物的设计使得 ATG 起始密码子构成了 NdeI 位点的后半部分(酶切识别位点 CATATG), 而在终止密码子后随即导入 BamHI 位点 (GGATCC)。PCR 产物被克隆到 pCR[®]2.1-TOPO 载体中, 得到质粒 TOPO-crtZ。为将两个基因组装到一个 Mini 操纵子中, 用 NdeI 和 BamHI 酶切质粒 TOPO-crtZ 和 TOP-crtW 的 crtZ 和 crtW, 并亚克隆到 NdeI-BamHI 酶切过的载体 pOCV-2 中, 得到质粒 pOCV-2-crtZ 和 pOCV-2-crtW。用 BseRI 和 PstI(在卡那霉素抗性基因中存在一个独特的 PstI 位点)酶切质粒 pOCV-2-crtZ, 将 2.4kb 片段(含有 crtZ)与来自 pOCV-2-crtW 的含 crtW 的 1876bp BseRI-PstI 片段相连。所获的质粒 pOCV-2-crtZW 用 NdeI 和 BamHI 酶切, 并与 pBBR-K-PcrtE 的 NdeI-BamHI 骨架相连, 得到 pBBR-K-PcrtE-crtZW。以类似的方式构建质粒 pBBR-K-PcrtE-crtWZ。

表 31 中的数据表明: 与菌株 pBBR-K-PcrtE-crtW 相比, 菌株 R114/pBBR-K-PcrtE-crtWZ 中的福寿草黄素(adonixanthin), 角黄素和虾青素比率并没有显著地改变。但是, 在 pBBR-K-PcrtE-crtZW 中, 酮类胡萝卜素(ketocarotenoid)的产量转变为以虾青素为主。该结果说明表达水平依赖于基因在 Mini 操纵子中的位置, 并暗示着增强体内 β -胡萝卜素羟化酶的活性水平可以提高该酶和 β -胡萝卜素 β -4 加氧酶活性之间的平衡, 而 β -胡萝卜素 β -4 加氧酶对于将玉米黄质完全转化为虾青素更为有利。

该实施例所述的结果还显示: 通过适当的遗传操作, 不仅可以在副球菌种菌株 R114 或其亲缘菌株内产生虾青素, 还可以生成其它目的酮类胡萝卜素商品。例如, 在菌株 R114 的 crtZ 突变株(缺乏 β -胡萝卜素羟化酶活性)中表达编码 β -胡萝卜素 β -4 加氧酶的基因, 可以专一性地生产酮类胡萝卜素, 如, 海胆酮或角黄素, 而不会同时生成羟基化的类胡萝卜素。综上, 该实施例和实施例 11 的结果说明: 副球菌种菌株 R114 及其亲缘菌株在工业化生产重要类胡萝卜素中具有光明的应用前景。

实施例 13: 在培养过量表达甲羟戊酸途径各基因的副球菌种菌株 R114 过程中甲羟戊酸的积累

在副球菌种菌株 R114 中过量表达甲羟戊酸途径的各基因, 可以增强碳

流动通过甲羟戊酸途径。实施例 5 描述了质粒 pBBR-K-mev-op16-2 的构建。按下述构建质粒 pBBR-K-mev-op-up-4。对含有副球菌种菌株 R114 甲羟戊酸操纵子的 λ 克隆进行部分酶切后(参见实施例 4)，在 3.1kb SmaI-SalI 片段上获得了含有 mvaA 基因绝大部分以及整个 idi 和 hcs 基因的 DNA 片段。将该

5 片段亚克隆到 pUC19 中，得到质粒 pUC19mev-op-up'。为促进亚克隆，将含有甲羟戊酸各基因的 pUC19mev-op-up'的 KpnI-HindIII 片段再克隆到载体 pBluescriptKS⁺ 中，得到质粒 pBluKSp-mev-op-up'。然后将来自 pUC19mev-op-up'的一段 1.7kb SalI 片段克隆到质粒 2ES2-1 的 SalI 位点，该质粒为 pUC19 衍生的质粒，含有副球菌种菌株 R114 的克隆 SalI-EcoRI 片段 M(参见实施例 4)。这样形成了质粒 pUC19mev-op-up-2。然后将

10 pUC19mev-op-up-2 中带有甲羟戊酸操纵子起始部分的 BbsI-BsaI 片段与 pBluKSp-mev-op-up'中含有 idi 和 hcs 的 BbsI-BsaI 片段相结合，得到质粒 pUC19mev-op-up-3。另外，通过插入一段含有 MluI 限制性位点的退火引物，将单一的 MluI 位点导入载体 pBBR-1MCS-2(见实施例 5)的 NsiI 和 KpnI 位

15 点。用 MluI 和 KpnI 酶切新获得的克隆载体 pBBR-K-Mlu，并将 pUCmev-op-up-3 的 MluI-KpnI 片段(该片段含有甲羟戊酸操纵子的前三个基因)插入，获得质粒 pBBR-K-mev-op-up-3。然后，插入质粒 16SB3 的 SmaI 片段，构建质粒 pBBR-K-mev-op-up-4，所述 16SB3 含有 mvk 基因的绝大部分和 pmk 的 5'末端(质粒 16SB3 是衍生自 pUC19 的质粒，含有副球菌种菌株 R114 SalI-BamHI 片段 A；参见实施例 4)。pBBR-K-mev-op-up-4 的插入

20 部分含有推定的甲羟戊酸操纵子启动子区域，甲羟戊酸操纵子的前四个基因和 pmk 的 5'末端。

通过电穿孔，将质粒 pBBR-K-mev-op16-2 和 pBBR-K-mev-op-up-4 分别导入副球菌种菌株 R114。对比新菌株与对照菌株 R114 的玉米黄质和甲羟

25 戊酸产量。在带挡板摇瓶中，用液体培养基 362F/2(见实施例 11)培养各菌株 72 小时。对于菌株 R114/pBBR-K-mev-op16-2 和菌株 R114/pBBR-K-mev-op-up-4，还向培养基中加入卡那霉素(50mg/l)。培养温度为 28℃，摇瓶速度为 200rpm。通过实施例 1 所述方法测定玉米黄质，而培养物上清中甲羟戊酸的测定按下述进行：将 0.6ml 培养物样品 13000x g 离

30 心 4 分钟。将 400 微升上清加入到 400 微升甲醇中，并涡旋搅拌 1min。再次将混合物以 13000x g 离心 4 分钟。然后，使用 Lindemann 的方法

[J.Pharm.Biomed. Anal. 9:311-316 (1991)], 将所获上清直接用气相色谱法(GC)进行分析, 但其中几处微小变动如下所述: GC 是 Hewlett-Packard 6890+仪器(Hewlett-Packard, Avondale, PA, USA), 其配有柱上冷却(cool-on-column)注射器和火焰离子化检测器。将 1 微升如上制备的样品注入熔融石英毛细管柱(15m 长 x 0.32mm ID)中, 所述柱子已用 0.52 微米交联修饰聚乙二醇(HP-FFAP, Agilent Technologies, USA)薄膜覆盖。进口压力 0.6 巴, 以氦气作为载体气体。可程序化注射器的温度以 30℃/分钟的速率从 82℃向 250℃上升。柱子温度曲线为 80℃ 0.5 分钟, 然后以 15℃/分钟的线性温度梯度升到 250℃, 最后在 250℃维持 5 分钟。检测温度维持在 320℃。

在第一个试验中, 测定菌株 R114 和 R114/pBBR-K-mev-op16-2 的玉米黄质和甲羟戊酸产量(表 32)。两个菌株的玉米黄质产量类似, 但是菌株 R114/pBBR-K-mev-op16-2 的甲羟戊酸产量水平要高出 4 倍。这些结果显示: 在副球菌种菌株 R114 中过量表达甲羟戊酸途径各基因可以提高碳流动通过甲羟戊酸途径。预先估计到了甲羟戊酸会积累, 因为菌株 R114/pBBR-K-mev-op16-2 不具有过量表达的 crtE 基因, 而已知 crtE 基因的产物(GGPP 合成酶)是副球菌种菌株 R114 中玉米黄质生成的限速步骤(见实施例 6 和 8)。可以预期: 具有有限数量 GGPP 合成酶的细胞在过量生成甲羟戊酸途径各酶之后, 会积累 FPP, 而众所周知 FPP 是甲羟戊酸即酶的有效抑制剂[Dorsey 和 Porter, J. Biol. Chem. 243:4667-4670(1968); Gray 和 Kekwick, BBA 279: 290-296(1972); Hinson 等, J.Lipids Res. 38:2216-2223(1997)]。因此, 由于过量表达甲羟戊酸途径各基因而引起的 FPP 的积累, 将会导致对甲羟戊酸激酶的抑制作用, 而这反过来可以表明培养过程中甲羟戊酸的积累。

表 32. 菌株 R114 和 R114/pBBR-K-mev-op16-2 的玉米黄质和甲羟戊酸产量

菌株/质粒	甲羟戊酸(mg/l)	玉米黄质(mg/l)
R114	50.5	70.0
R114/pBBR-K-mev-op16-2	208.2	65.2

在第二个试验中, 测定菌株 R114 和 R114/pBBR-K-mev-op-up-4 的两个

独立分离株的玉米黄质和甲羟戊酸产量(表 33)。这些结果再次表明：过量表达甲羟戊酸途径各基因可以提高碳流动通过甲羟戊酸途径。

表 33.菌株 R114 和 R114/pBBR-K-mev-op-up-4 的玉米黄质和甲羟戊酸产量

菌株/质粒	甲羟戊酸(mg/l)	玉米黄质(mg/l)
R114	45	67.5
R114/pBBR-K-mev-op-up-4(分离 1)	133.2	53.7
R114/pBBR-K-mev-op-up-4(分离 2)	163.7	47.6

- 5 根据布达佩斯条约，下述生物材料保藏在美国典型培养物保藏中心 (American Type Culture Collection (ATCC)), 0801 University Blvd., Manassas VA 20110-2201, USA, 登录号如下：

菌株	登录号	保藏日
副球菌属菌种 R114	PTA-3335	April 24,2001
副球菌属菌种 R1534	PTA-3336	April 24,2001
副球菌属菌种 R-1506	PTA-3431	June 5,2001

以上引用的所有专利，专利申请和出版物均将其全部引入本文作为参考，如同将每篇在本文中全文引用一样。

- 10 本文所公开的发明，很明显可以通过各种方式对其进行修改。这种改变仍然属于本发明的精神和范围之内，这样的修改仍落入下面的权利要求范围中。

	<110>	罗奇维生素股份公司 (Roche Vitamins AG)	
	<120>	改进的类异戊二烯生产	
	<130>	20918	
5	<160>	197	
	<170>	PatentIn version 3.1	
	<210>	1	
	<211>	20	
10	<212>	DNA	
	<213>		
	<220>		
	<221>	misc_feature	
	<222>	(1)..(20)	
15	<223>	引物	
	<400>	1	
		agagtttgat cctggctcag	20
	<210>	2	
20	<211>	20	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<221>	misc_feature	
25	<222>	(12)..(12)	
	<223>	n 代表 C 或 T	
	<220>		
	<221>	misc_feature	
	<222>	(1)..(20)	
30	<223>	引物	
	<400>	2	
		ctggctcagg angaacgctg	20
	<210>	3	
35	<211>	20	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<221>	misc_feature	
40	<222>	(1)..(20)	
	<223>	引物	
	<400>	3	
		aaggaggtga tccagccgca	20
	<210>	4	
45	<211>	20	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<221>	misc_feature	
50	<222>	(1)..(20)	
	<223>	引物	
	<400>	4	
		ctcctacggg aggcagcagt	20
55	<210>	5	
	<211>	18	
	<212>	DNA	
	<213>	人工的	
	<220>		
60	<221>	misc_feature	

	<222> (1)..(18)	
	<223> 引物	
	<400> 5	
5	cagcagccgc ggtaatac	18
	<210> 6	
	<211> 19	
	<212> DNA	
	<213> 人工的	
10	<220>	
	<221> misc_feature	
	<222> (1)..(19)	
	<223> 引物	
	<400> 6	
15	aactcaaagg aattgacgg	19
	<210> 7	
	<211> 20	
	<212> DNA	
20	<213> 人工的	
	<220>	
	<221> misc_feature	
	<222> (1)..(20)	
	<223> 引物	
25	<400> 7	
	agtcccgcga cgagcgcaac	20
	<210> 8	
	<211> 20	
	<212> DNA	
30	<213> 人工的	
	<220>	
	<221> misc_feature	
	<222> (1)..(20)	
	<223> 引物	
35	<400> 8	
	gctacacacg tgctacaatg	20
	<210> 9	
	<211> 20	
40	<212> DNA	
	<213> 人工的	
	<220>	
	<221> misc_feature	
	<222> (1)..(20)	
45	<223> 引物	
	<400> 9	
	actgctgcct cccgtaggag	20
	<210> 10	
50	<211> 18	
	<212> DNA	
	<213> 人工的	
	<220>	
	<221> misc_feature	
55	<222> (1)..(18)	
	<223> 引物	
	<400> 10	
	gtattaccgc ggctgctg	18

	<210>	11		
	<211>	20		
	<212>	DNA		
	<213>	人工的		
5	<220>			
	<221>	misc_feature		
	<222>	(1)..(20)		
	<223>	引物		
	<400>	11		
10	gttgcgctcg ttgcgggact			20
	<210>	12		
	<211>	1404		
	<212>	DNA		
	<213>	副球菌属菌种(Paracoccus sp.)		
15	<400>	12		
	gcggcaggct taacacatgc aagtcgagcg aggtcttcgg acctagcggc ggacgggtga			60
	gtaacgcgtg ggaacgtgcc ctttgctacg gaatagtccc gggaaactgg gtttaatacc			120
	gtatgtgccc tacgggggaa agatttatcg gcaaaggatc ggcccgcgtt ggattaggta			180
	gttggtgggg taatggccta ccaagccgac gatccatagc tggtttgaga ggatgatcag			240
20	ccacactggg actgagacac ggcccagact cctacgggag gcagcagtgg ggaatcttag			300
	acaatggggg caaccctgat ctagccatgc cgcgtgagtg atgaaggccc tagggttgta			360
	aagctctttc agctgggaag ataatgacgg taccagcaga agaagccccg gctaactccg			420
	tgccagcagc cgcggttaata cggagggggc tagcgttggt cggaattact gggcgtaaa			480
	cgcacgtagg cggactggaa agttgggggt gaaatcccgg ggctcaacct cggaactgcc			540
25	tccaaaacta tcagtctgga gttcgagaga ggtgagtgga ataccgagt tagaggtgaa			600
	attcgtagat attcgggtgga acaccagtgg cgaaggcggc tctactggctc gatactgacg			660
	ctgaggtgcg aaagcgtggg gagcaaacag gattagatac cctggtagtc cagccgtaa			720
	acgatgaatg ccagtcgtcg ggttgcatgc aattcggtag cacacctaac ggattaagca			780
	ttccgccttg ggagtacggg cgcaagatta aaactcaaag gaattgacgg gggcccgac			840
30	aagcggtgga gcatgtggtt taattcgaag caacgcgcag aaccttacca acccttgaca			900
	tccctggaca tcccgagaga tcgggctttc acttcggtag ccaggagaca ggtgctgcat			960
	ggctgtcgtc agctcgtgtc gtgagatggt cggttaagtc cggcaacgag cgcaaccac			1020
	gtccctagtt gccagcattc agttgggcac tctatggaaa ctgccgatga taagtcggag			1080
	gaaggtgtgg atgacgtcaa gtcctcatgg cccttacggg ttgggctaca cacgtgctac			1140
35	aatggtggtg acagtgggtt aatcccaaaa agccatctca gttcggattg tcctctgcaa			1200
	ctcgagggca tgaagttgga atcgctagta atcgcggaac agcatgccgc ggtgaatacg			1260
	ttcccgggcc ttgtacacac cgcccgtcac accatgggag ttggttctac ccgacgacgc			1320
	tgcgctaacc cttcggggag gcaggcggcc acggtaggat cagcgactgg ggtgaagtgc			1380
	taacaaggta gccgtagggg aacc			1404
40				
	<210>	13		
	<211>	20		
	<212>	DNA		
	<213>	人工的		
45	<220>			
	<221>	misc_feature		
	<222>	(1)..(20)		
	<223>	衔接头		
	<400>	13		
50	tcgtagactg cgtacaggcc			20
	<210>	14		
55	<211>	14		
	<212>	DNA		
	<213>	人工的		
	<220>			
	<221>	misc_feature		

	<222> (1)..(14)	
	<223> 衔接头	
	<400> 14	
	catctgacgc atgt	14
5		
	<210> 15	
	<211> 16	
	<212> DNA	
	<213> 人工的	
10	<220>	
	<221> misc_feature	
	<222> (1)..(16)	
	<223> 衔接头	
	<400> 15	
15	gacgatgagt cctgac	16
	<210> 16	
	<211> 14	
	<212> DNA	
20	<213> 人工的	
	<220>	
	<221> misc_feature	
	<222> (1)..(14)	
	<223> 衔接头	
25	<400> 16	
	tactcaggac tggc	14
	<210> 17	
	<211> 17	
30	<212> DNA	
	<213> 人工的	
	<220>	
	<221> misc_feature	
	<222> (1)..(17)	
35	<223> 引物	
	<400> 17	
	gactgcgtac aggccca	17
	<210> 18	
	<211> 18	
40	<212> DNA	
	<213> 人工的	
	<220>	
	<221> misc_feature	
	<222> (1)..(18)	
45	<223> 引物	
	<400> 18	
	cgatgagtcc tgaccgaa	18
	<210> 19	
50	<211> 18	
	<212> DNA	
	<213> 人工的	
	<220>	
	<221> misc_feature	
55	<222> (1)..(18)	
	<223> 引物	
	<400> 19	
	cgatgagtcc tgaccgac	18

	<210>	20	
	<211>	17	
	<212>	DNA	
	<213>	人工的	
5	<220>		
	<221>	misc_feature	
	<222>	(1)..(17)	
	<223>	引物	
	<400>	20	
10		gactgcgtac aggcccc	17
	<210>	21	
	<211>	17	
	<212>	DNA	
15	<213>	人工的	
	<220>		
	<221>	misc_feature	
	<222>	(1)..(17)	
	<223>	引物	
20	<400>	21	
		gactgcgtac aggcccg	17
	<210>	22	
	<211>	18	
	<212>	DNA	
25	<213>	人工的	
	<220>		
	<221>	misc_feature	
	<222>	(1)..(18)	
	<223>	引物	
30	<400>	22	
		cgatgagtcg tgaccgag	18
	<210>	23	
	<211>	8	
35	<212>	PRT	
	<213>	人工的	
	<220>		
	<221>	MISC_FEATURE	
	<222>	(2)..(2)	
40	<223>	Xaa 代表 Leu 或 Ile	
	<220>		
	<221>	MISC_FEATURE	
	<222>	(1)..(8)	
	<223>	肽 1	
45	<400>	23	
		Ala Xaa Ile Lys Tyr Lys Gly Lys	
		1 5	
	<210>	24	
50	<211>	27	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<221>	misc_feature	
55	<222>	(3)..(3)	
	<223>	n 代表 C 或 G	
	<220>		
	<221>	misc_feature	
	<222>	(12)..(12)	
60	<223>	n 代表 A 或 G	

	<220>		
	<221>	misc_feature	
	<222>	(15)..(15)	
	<223>	n 代表 C 或 T	
5	<220>		
	<221>	misc_feature	
	<222>	(21)..(21)	
	<223>	n 代表 C 或 G	
	<220>		
10	<221>	misc_feature	
	<222>	(24)..(24)	
	<223>	n 代表 A 或 G	
	<220>		
	<221>	misc_feature	
15	<222>	(1)..(27)	
	<223>	寡核苷酸	
	<400>	24	
		ccnctgatca antantgggg naanatc	27
20	<210>	25	
	<211>	20	
	<212>	DNA	
	<213>	人工的	
	<220>		
25	<221>	misc_feature	
	<222>	(3)..(3)	
	<223>	n 代表 C 或 G	
	<220>		
	<221>	misc_feature	
30	<222>	(12)..(12)	
	<223>	n 代表 A 或 G	
	<220>		
	<221>	misc_feature	
	<222>	(15)..(15)	
35	<223>	n 代表 C 或 T	
	<220>		
	<221>	misc_feature	
	<222>	(1)..(20)	
	<223>	寡核苷酸	
40	<400>	25	
		gcnctgatca antantgggg	20
	<210>	26	
	<211>	20	
45	<212>	DNA	
	<213>	人工的	
	<220>		
	<221>	misc_feature	
	<222>	(3)..(3)	
50	<223>	n 代表 C 或 G	
	<220>		
	<221>	misc_feature	
	<222>	(12)..(12)	
	<223>	n 代表 A 或 G	
55	<220>		
	<221>	misc_feature	
	<222>	(15)..(15)	
	<223>	n 代表 C 或 T	
	<220>		
60	<221>	misc_feature	

	<222> (1)..(20)	
	<223> 寡核苷酸	
	<400> 26	
	gcnatcatca antantgggg	20
5	<210> 27	
	<211> 17	
	<212> DNA	
	<213> 人工的	
10	<220>	
	<221> misc_feature	
	<222> (6)..(6)	
	<223> n 代表 A 或 G	
	<220>	
15	<221> misc_feature	
	<222> (9)..(9)	
	<223> n 代表 C 或 T	
	<220>	
	<221> misc_feature	
20	<222> (1)..(17)	
	<223> 寡核苷酸	
	<400> 27	
	atcaantant ggggtaa	17
25	<210> 28	
	<211> 17	
	<212> DNA	
	<213> 人工的	
	<220>	
30	<221> misc_feature	
	<222> (6)..(6)	
	<223> n 代表 A 或 G	
	<220>	
	<221> misc_feature	
35	<222> (9)..(9)	
	<223> n 代表 C 或 T	
	<220>	
	<221> misc_feature	
	<222> (1)..(17)	
40	<223> 寡核苷酸	
	<400> 28	
	atcaantant ggggcaa	17
	<210> 29	
45	<211> 17	
	<212> DNA	
	<213> 人工的	
	<220>	
	<221> misc_feature	
50	<222> (6)..(6)	
	<223> n 代表 A 或 G	
	<220>	
	<221> misc_feature	
	<222> (9)..(9)	
55	<223> n 代表 C 或 T	
	<220>	
	<221> misc_feature	
	<222> (1)..(17)	
	<223> 寡核苷酸	
60	<400> 29	

	atcaantant gggggaa	17
	<210> 30	
	<211> 17	
5	<212> DNA	
	<213> 人工的	
	<220>	
	<221> misc_feature	
	<222> (6)..(6)	
10	<223> n 代表 A 或 G	
	<220>	
	<221> misc_feature	
	<222> (9)..(9)	
	<223> n 代表 C 或 T	
15	<220>	
	<221> misc_feature	
	<222> (1)..(17)	
	<223> 寡核苷酸	
	<400> 30	
20	atcaantant ggggaaa	17
	<210> 31	
	<211> 8	
	<212> PRT	
25	<213> 人工的	
	<220>	
	<221> MISC_FEATURE	
	<222> (7)..(7)	
	<223> Xaa 代表 Asn 或 Gln	
30	<220>	
	<221> MISC_FEATURE	
	<222> (1)..(8)	
	<223> 肽 2	
	<400> 31	
35	Thr Met Asp Ala Gly Pro Xaa Val	
	1 5	
	<210> 32	
	<211> 24	
40	<212> DNA	
	<213> 人工的	
	<220>	
	<221> misc_feature	
	<222> (3)..(3)	
45	<223> n 代表 C 或 G	
	<220>	
	<221> misc_feature	
	<222> (9)..(9)	
	<223> n 代表 C 或 T	
50	<220>	
	<221> misc_feature	
	<222> (12)..(12)	
	<223> n 代表 C 或 G	
	<220>	
55	<221> misc_feature	
	<222> (15)..(15)	
	<223> n 代表 C 或 G 或 T	
	<220>	
	<221> misc_feature	

	<222>	(18)..(18)	
	<223>	n 代表 C 或 G	
	<220>		
	<221>	misc_feature	
5	<222>	(21)..(21)	
	<223>	n 代表 C 或 T	
	<220>		
	<221>	misc_feature	
	<222>	(24)..(24)	
10	<223>	n 代表 C 或 G	
	<220>		
	<221>	misc_feature	
	<222>	(1)..(24)	
	<223>	寡核苷酸	
15	<400>	32	
		acnatggang cnggncnaa ngtn	24
	<210>	33	
	<211>	24	
20	<212>	DNA	
	<213>	人工的	
	<220>		
	<221>	misc_feature	
	<222>	(3)..(3)	
25	<223>	n 代表 C 或 G	
	<220>		
	<221>	misc_feature	
	<222>	(9)..(9)	
	<223>	n 代表 A 或 G	
30	<220>		
	<221>	misc_feature	
	<222>	(12)..(12)	
	<223>	n 代表 A 或 G	
	<220>		
35	<221>	misc_feature	
	<222>	(12)..(12)	
	<223>	n 代表 C 或 G	
	<220>		
	<221>	misc_feature	
40	<222>	(15)..(15)	
	<223>	n 代表 A 或 C 或 G	
	<220>		
	<221>	misc_feature	
	<222>	(18)..(18)	
45	<223>	n 代表 C 或 G	
	<220>		
	<221>	misc_feature	
	<222>	(21)..(21)	
	<223>	n 代表 A 或 G	
50	<220>		
	<221>	misc_feature	
	<222>	(24)..(24)	
	<223>	n 代表 C 或 G	
	<220>		
55	<221>	misc_feature	
	<222>	(1)..(24)	
	<223>	寡核苷酸	
	<400>	33	
		tgntacctnc gncnggntt ncan	24
60			

	<210>	34	
	<211>	17	
	<212>	DNA	
	<213>	人工的	
5	<220>		
	<221>	misc_feature	
	<222>	(12)..(12)	
	<223>	n 代表 C 或 G	
	<220>		
10	<221>	misc_feature	
	<222>	(15)..(15)	
	<223>	n 代表 A 或 C 或 G	
	<220>		
	<221>	misc_feature	
15	<222>	(1)..(17)	
	<223>	寡核苷酸	
	<400>	34	
		tggtacctac gnccngg	17
20			
	<210>	35	
	<211>	17	
25	<212>	DNA	
	<213>	人工的	
	<220>		
	<221>	misc_feature	
	<222>	(12)..(12)	
30	<223>	n 代表 C 或 G	
	<220>		
	<221>	misc_feature	
	<222>	(15)..(15)	
	<223>	n 代表 A 或 C 或 G	
35	<220>		
	<221>	misc_feature	
	<222>	(1)..(17)	
	<223>	寡核苷酸	
	<400>	35	
40		tggtacctgc gnccngg	17
	<210>	36	
	<211>	17	
	<212>	DNA	
45	<213>	人工的	
	<220>		
	<221>	misc_feature	
	<222>	(12)..(12)	
	<223>	n 代表 C 或 G	
50	<220>		
	<221>	misc_feature	
	<222>	(1)..(17)	
	<223>	寡核苷酸	
	<220>		
55	<221>	misc_feature	
	<222>	(15)..(15)	
	<223>	n 代表 A 或 C 或 G	
	<400>	36	
		tgctacctac gnccngg	17
60			

	<210>	37	
	<211>	17	
	<212>	DNA	
	<213>	人工的	
5	<220>		
	<221>	misc_feature	
	<222>	(12)..(12)	
	<223>	n 代表 C 或 G	
	<220>		
10	<221>	misc_feature	
	<222>	(15)..(15)	
	<223>	n 代表 A 或 C 或 G	
	<220>		
15	<221>	misc_feature	
	<222>	(1)..(17)	
	<223>	寡核苷酸	
	<400>	37	
		tgctacctgc gncnngg	17
20	<210>	38	
	<211>	20	
	<212>	DNA	
	<213>	人工的	
	<220>		
25	<221>	misc_feature	
	<222>	(9)..(9)	
	<223>	n 代表 C 或 G	
	<220>		
	<221>	misc_feature	
30	<222>	(12)..(12)	
	<223>	n 代表 A 或 C 或 G	
	<220>		
	<221>	misc_feature	
	<222>	(15)..(15)	
35	<223>	n 代表 C 或 G	
	<220>		
	<221>	misc_feature	
	<222>	(18)..(18)	
	<223>	n 代表 A 或 G	
40	<220>		
	<221>	misc_feature	
	<222>	(1)..(20)	
	<223>	寡核苷酸	
	<400>	38	
45		tacctacgnc cnggnttnca	20
50	<210>	39	
	<211>	20	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<221>	misc_feature	
55	<222>	(9)..(9)	
	<223>	n 代表 C 或 G	
	<220>		
	<221>	misc_feature	
	<222>	(12)..(12)	
60	<223>	n 代表 A 或 C 或 G	

	<220>		
	<221>	misc_feature	
	<222>	(15)..(15)	
	<223>	n 代表 C 或 G	
5	<220>		
	<221>	misc_feature	
	<222>	(18)..(18)	
	<223>	n 代表 A 或 G	
	<220>		
10	<221>	misc_feature	
	<222>	(1)..(20)	
	<223>	寡核苷酸	
	<400>	39	
		tacctgcgnc cnggnttnca	20
15	<210>	40	
	<211>	20	
	<212>	DNA	
	<213>	人工的	
20	<220>		
	<221>	misc_feature	
	<222>	(9)..(9)	
	<223>	n 代表 C 或 G	
	<220>		
25	<221>	misc_feature	
	<222>	(12)..(12)	
	<223>	n 代表 A 或 C 或 G	
	<220>		
	<221>	misc_feature	
30	<222>	(15)..(15)	
	<223>	n 代表 C 或 G	
	<220>		
	<221>	misc_feature	
	<222>	(18)..(18)	
35	<223>	n 代表 C 或 T	
	<220>		
	<221>	misc_feature	
	<222>	(1)..(20)	
	<223>	寡核苷酸	
40	<400>	40	
		tacctacgnc cnggngtnca	20
	<210>	41	
	<211>	20	
45	<212>	DNA	
	<213>	人工的	
	<220>		
	<221>	misc_feature	
	<222>	(9)..(9)	
50	<223>	n 代表 C 或 G	
	<220>		
	<221>	misc_feature	
	<222>	(12)..(12)	
	<223>	n 代表 A 或 C 或 G	
55	<220>		
	<221>	misc_feature	
	<222>	(15)..(15)	
	<223>	n 代表 C 或 G	
	<220>		
60	<221>	misc_feature	

	gttgacataa cttgcacgcg acgtctcgat tctgcccgcg aagaatgcga tgcattccaga	2580
	tgatgcagaa cgaagaagcg gaagcgcccg tgaaagacca g atg att tcc cat acc	2636
	Met Ile Ser His Thr	
	1 5	
5	ccg gtg ccc acg caa tgg gtc ggc ccg atc ctg ttc cgc ggc ccc gtc	2684
	Pro Val Pro Thr Gln Trp Val Gly Pro Ile Leu Phe Arg Gly Pro Val	
	10 15 20	
10	gtc gag ggc ccg atc agc gcg ccg ctg gcc acc tac gag acg ccg ctc	2732
	Val Glu Gly Pro Ile Ser Ala Pro Leu Ala Thr Tyr Glu Thr Pro Leu	
	25 30 35	
15	tgg ccc tcg acc gcg cgg ggg gca ggg gtt tcc cgg cat tcg ggc ggg	2780
	Trp Pro Ser Thr Ala Arg Gly Ala Gly Val Ser Arg His Ser Gly Gly	
	40 45 50	
20	atc cag gtc tcg ctg gtc gac gaa cgc atg agc cgc tcg atc gcg ctg	2828
	Ile Gln Val Ser Leu Val Asp Glu Arg Met Ser Arg Ser Ile Ala Leu	
	55 60 65	
25	cgg gcg cat gac ggg gcg gcg gcg acc gcc gcc tgg cag tcg atc aag	2876
	Arg Ala His Asp Gly Ala Ala Ala Thr Ala Ala Trp Gln Ser Ile Lys	
	70 75 80 85	
30	gcc cgc cag gaa gag gtc gcg gcc gtg gtc gcc acc acc agc cgc ttc	2924
	Ala Arg Gln Glu Glu Val Ala Ala Val Val Ala Thr Thr Ser Arg Phe	
	90 95 100	
35	gcc cgc ctt gtc gag ctg aat cgc cag atc gtg ggc aac ctg ctt tac	2972
	Ala Arg Leu Val Glu Leu Asn Arg Gln Ile Val Gly Asn Leu Leu Tyr	
	105 110 115	
40	atc cgc atc gaa tgc gtg acg ggc gac gcc tcg ggt cac aac atg gtc	3020
	Ile Arg Ile Glu Cys Val Thr Gly Asp Ala Ser Gly His Asn Met Val	
	120 125 130	
45	acc aag gcc gcc gag gcc gtg cag ggc tgg atc ctg tcg gaa tac ccg	3068
	Thr Lys Ala Ala Glu Ala Val Gln Gly Trp Ile Leu Ser Glu Tyr Pro	
	135 140 145	
50	atg ctg gcc tat tcc acg atc tcg ggg aac ctg tgc acc gac aag aag	3116
	Met Leu Ala Tyr Ser Thr Ile Ser Gly Asn Leu Cys Thr Asp Lys Lys	
	150 155 160 165	
55	gcg tcg gcg gtc aac ggc atc ctg ggc cgc ggc aaa tac gcc gtc gcc	3164
	Ala Ser Ala Val Asn Gly Ile Leu Gly Arg Gly Lys Tyr Ala Val Ala	
	170 175 180	
60	gag gtc gag atc ccg cgc aag atc ctg acc cgc gtg ctg cgc acc agc	3212
	Glu Val Glu Ile Pro Arg Lys Ile Leu Thr Arg Val Leu Arg Thr Ser	
	185 190 195	
65	gcc gag aag atg gtc cgc ctg aac tac gag aag aac tat gtc ggg ggt	3260
	Ala Glu Lys Met Val Arg Leu Asn Tyr Glu Lys Asn Tyr Val Gly Gly	
	200 205 210	
70	acg ctg gcg ggg tcg ctg cgc agt gcg aac gcg cat ttc gcc aac atg	3308
	Thr Leu Ala Gly Ser Leu Arg Ser Ala Asn Ala His Phe Ala Asn Met	
	215 220 225	

	ctg ctg ggc ttc tac ctg gcg acg ggg cag gac gcg gcc aac atc atc	3356
	Leu Leu Gly Phe Tyr Leu Ala Thr Gly Gln Asp Ala Ala Asn Ile Ile	
	230 235 240 245	
5	gag gcc agc cag ggc ttc gtc cat tgc gag gcc cgc ggc gag gat ctg	3404
	Glu Ala Ser Gln Gly Phe Val His Cys Glu Ala Arg Gly Glu Asp Leu	
	250 255 260	
10	tat ttc tcg tgc acg ctg ccc aac ctc atc atg ggc tcg gtc ggt gcc	3452
	Tyr Phe Ser Cys Thr Leu Pro Asn Leu Ile Met Gly Ser Val Gly Ala	
	265 270 275	
15	ggc aag ggc atc ccc tcg atc gag gag aac ctg tcg cgg atg ggc tgc	3500
	Gly Lys Gly Ile Pro Ser Ile Glu Glu Asn Leu Ser Arg Met Gly Cys	
	280 285 290	
20	cgc cag ccg ggc gaa ccc ggc gac aac gcg cgc cgt ctt gcg gcg atc	3548
	Arg Gln Pro Gly Glu Pro Gly Asp Asn Ala Arg Arg Leu Ala Ala Ile	
	295 300 305	
25	tgc gcg ggc gtc gtg ctg tgt ggt gaa ttg tcg ctg ctt gcg gcc cag	3596
	Cys Ala Gly Val Val Leu Cys Gly Glu Leu Ser Leu Leu Ala Ala Gln	
	310 315 320 325	
30	acc aac ccc gga gag ttg gtc cgc acc cac atg gag atg gag cga tga	3644
	Thr Asn Pro Gly Glu Leu Val Arg Thr His Met Glu Met Glu Arg	
	330 335 340	
35	ccgacagcaa ggatcaccat gtgcgcggggc gcaagotgga ccatctgcgt gcattggacg	3704
	acgatgcgga tatcgaccgg ggcgacagcg gcttcgaccg catcgcgctg acccatcgcg	3764
	ccctgcccga ggtggatttc gacgccatcg acacggcgac cagcttcctg ggccgtgaac	3824
	tgctcctccc gctgctgatac tcgtccatga ccggcgggcac cggcgaggag atcgagcgca	3884
	tcaaccgcaa cctggccgct ggtgccgagg agggccgcgt cgccatggcg gtgggctcgc	3944
	agcgcgtgat gttcaccgac ccctcggcgc gggccagctt cgacctgcgc gcccatgcgc	4004
	ccaccgtgcc gctgctggcc aatatcggcg cgggtgcagct gaacatgggg ctggggctga	4064
40	aggaatgcct ggccgcgatac gaggtgctgc aggcggacgg cctgtatctg cacctgaacc	4124
	ccctgcaaga ggccgtccag ccogaggggg atcgcgactt tgccgatctg ggcagcaaga	4184
	tcgcggccat cgcccgcgac gttcccgtgc ccgtcctgct gaaggagggtg ggctgcggcc	4244
	tgtcggcggc cgatatcgcc atcgggctgc gcgcggggat ccggcatttc gacgtggccg	4304
	gtcgcggcgg cacatccttg agccggatcg agtatcgccg ccgccagcgg gccgatgacg	4364
45	acctgggcct ggtcttccag gactggggcg tcgacaccgt ggacgcctg cgcgaggcgc	4424
	ggcccgcgct tgcggcccat gatggaacca gcgtgctgat cgccagcggc ggcatccgca	4484
	acggtgtcga catggcgaaa tgcgtcatcc tgggggcccga catgtgcggg gtcgccgcgc	4544
	ccctgctgaa agcggcccaa aactcgcgcg aggcggttgt atccgccatc cggaaactgc	4604
	atctggagtt ccggacagcc atgttcctcc tgggttgccg cacgcttgcc gacctgaagg	4664
50	acaattcctc gcttatccgt caatgaaagt gcctaagatg accgtgacag gaatcgaagc	4724
	gatcagcttc tacacccccc agaactacgt gggactggat atccttgccg cgcatacagg	4784
	gatcgacccc gagaagtctt cgaaggggat cgggcaggag aaaatcgcac tgcccggcca	4844
	tgacgaggat atcgtgacca tggccgcgca ggccgcgctg ccgatcatcg aacgcgcggg	4904
	cacgcagggc atcgacacgg ttctgttcgc caccgagagc gggatcgacc agtcgaaggc	4964
55	cgccgccatc tatctgcgcc gcctgctgga cctgtcgccc aactgccgtt gcgtcgagct	5024
	gaagccgagc tgctattccg cgacggcgcc gctgcagatg gcctgcgcgc atgtcgcccg	5084
	caagcccgac cgcaagggtgc tgggtgatgc gtccgatgtc gcgcgctatg accgcgaaag	5144
	ctcgggcgag gcgacgcagg gtgcggggcg cgtcgccatc cttgtcagcg ccgatcccaa	5204
	ggtggccgag atcggcaccg tctcggggct gttcaccgag gatatcatgg atttctggcg	5264
60	gocgaaccac cgccgcacgc ccctgttcga cggcaaggca tcgacgctgc gctatctgaa	5324
	cgcgctggtc gaggcgtgga acgactatcg cgcaaatggc ggccacgagt tcgccgattt	5384
	cgcgcatttc tgetatcacg tgccgtttct gcggatgggc gagaaggcga acagccacct	5444
	ggccaaggcg aacaagacgc cgggtggacat ggggcagggt cagacgggcc tgatctacaa	5504
	ccggcaggtc gggaactgct ataccgggtc gatctacctg gcattcgctt cgctgctgga	5564
	gaacgctcag gaggacctga ccggcgcgct ggtcgggtctg ttcagctatg gctcgggtgc	5624

	gacgggcgaa	ttcttcgatg	cgcggatcgc	gcccggttac	cgcgaccacc	tggttcgcgga	5684
	acgccatcgc	gaattgctgc	aggatcgcac	gcccgtcaca	tatgacgaat	acgttgccct	5744
	gtgggacgag	atcgacctga	cgcagggcgc	gcccgcacaag	gcgcgcggtc	gtttcaggct	5804
	ggcaggtatc	gaggacgaga	agcgcatac	tgctcgaccg	caggcctgaa	gcaggcgccc	5864
5	atgccccggg	caagctgatc	ctgtccgggg	aacattccgt	gctctatggt	gcgcccgcgc	5924
	ttgccatggc	catcgcccgc	tataccgagg	tgtggttcac	gccgcttggc	attggcgagg	5984
	ggatacgcac	gacattcgcc	aatctctcgg	gcggggcgac	ctattcgctg	aagctgctgt	6044
	cgggggtcaa	gtcgcggctg	gaccgcgggt	tcgagcagtt	cctgaacggc	gacctaaagg	6104
	tgcaacaagg	cctgacccat	cccgcagatc	tgccggtcta	tcgcctggcg	tcgcttctgc	6164
10	acgacaagcc	gccggggacc	gccgcgatgc	cgggcatcgg	cgcgatgcac	cacctgccgc	6224
	gaccgggtga	gctgggcagc	cggacggagc	tgcccatcgg	cgcgggcatg	gggtcgtctg	6284
	cggccatcgt	cgcggccacc	acggtcctgt	tcgagacgct	gctggaccgg	cccaagacgc	6344
	ccgaacagcg	cttcgaccgc	gtccgcttct	gcgagcgggt	gaagcacggc	aaggccggtc	6404
	ccatcgacgc	ggccagcgct	gtgcgcggcg	ggcttgctcg	cgtgggcggg	aacgggcccg	6464
15	gttcgatcag	cagcttcgat	ttgcccagg	atcacgacct	tgtcgcggga	cgcggctggt	6524
	actgggtact	gcacgggcgc	cccgtcagcg	ggaccggcga	atgcgtcagc	gcggtcgcgg	6584
	cggcgctagg	tcgcgatgcg	gcgctgtggg	acgccttcgc	agtctgcacc	cgcgcttggg	6644
	aggccgcgct	gctgtctggg	ggcagccccg	acgcgcccat	caccgagAAC	cagcgctgc	6704
	tggAACgcac	cggcgctcgt	ccggcagcga	cgcaggccct	cgtggcccag	atcgaggagg	6764
20	cgggtggcgc	ggccaagatc	tgccggcgag	gttcctgtcg	gggcgatcac	ggcggggcgg	6824
	tcctcgtgcg	gattgacgac	gcgcaggcga	tggtctcggg	catggcgcg	catcccgacc	6884
	tcgactgggc	gcccctgcgc	atgtcgcgca	cgggggcggc	acccggcccc	gcgcgcgctg	6944
	cgcAACcgct	gccggggcag	ggctgatgga	tcaggtcac	cgcgccagcg	cgcgggttc	7004
	ggcatgatc	acgggcgaac	atgccgtggg	ctatggacac	cgcgccatcg	tcgccgggat	7064
25	cgcgcagcgc	gcccattgta	cgatcgtccc	gcgtgccgac	cgcattgttc	gcacacctc	7124
	gcagatcggg	gcgcgcgagc	aggggtcgct	ggacgatctg	cctgcggggc	ggacctatcg	7184
	cttcgtgctg	gcccgcacgc	cgcgacacgc	gcccgaacct	ccttgccggg	tcgacatgga	7244
	catcacctcg	gggatcgatc	cgcggtcg	gcttggtacc	tcggcgcgcg	tgacggctgc	7304
	ctgcctcggc	gcgctgtcgc	ggctggcggg	gcgggggacc	gaggggctgc	atgacgacgc	7364
30	gctgcgcac	gtccgcgcca	tcaggggcag	gggcagcggg	gccgatctgg	cggccagcct	7424
	gcattggcgc	ttcgtcgctt	atcgcgcgcc	cgatggcggt	gccgcgcaga	tcgaggcgct	7484
	tcgggtgccc	cgggggcccgt	tcggcctgcg	ctatgcgggc	tacaagacc	cgcagccga	7544
	gggtgctgcg	cttggtggccg	atcggatggc	gggcaacgag	gccgctttcg	acgcgctcta	7604
	ctcccggtatg	ggcgcaagcg	cagatgcgcg	gatccgcgcg	gcgcaagggc	tggaactggg	7664
35	tgcatccac	gacgcgctga	acgaatacca	gcgcctgatg	gagcagctgg	gcgtgtccga	7724
	cgcacacgctg	gacgcgatca	tcgcgcaggc	gcgcgacgcg	ggcgcgcgag	tcgccaagat	7784
	ctccggctcg	gggctggggg	attgcgtgct	ggcagctggc	gaccagccca	agggtttcgt	7844
	gcccgcacag	attgccgaga	agggacttgt	tttcgatgac	tgatgccgtc	cgcgacatga	7904
	tcgcccgtgc	catggcgggc	gcgaccgaca	tcgcgacgag	cgcggcttat	gcgcccagca	7964
40	acatcgcgct	gtcgaaatac	tggggcaagc	gcgacgcgcg	gcggaacctt	ccgctgaaca	8024
	gctccgtctc	gatctcgttg	gcgaactggg	gctctcaca	gcgggtcgag	gggtccggca	8084
	cgggccacga	cgcgggtgat	cacaacggca	cgcgtgctga	tcggggcgac	gccttcgcgc	8144
	gccgcgcgct	ggcatctcgt	gacctgttcc	ggggggggag	gcacctgccg	ctgcggatca	8204
	cgcgcgcagaa	ctcgatcccc	acggcgggcg	ggcttgccct	gtcggccctc	gggttcgcgg	8264
45	cgcgtgacccg	tgccgtggcg	ggggcggttc	ggctggatct	ggacgacacg	gatctgagcc	8324
	gcattcgccc	gatcggcag	ggcagcgccg	acgcctcgat	ctggcacgcg	ttcgtccgct	8384
	ggaaccgggg	cgcggccgag	gatgggcag	ccagccacgg	cgtcccgcgt	gacctgcgct	8444
	ggcccggcct	ccgcacgcg	atcgtggccg	tggaacaagg	gcccgaacct	ttcagttcgc	8504
	gcgcagggcat	gaaccacacg	gtcgagacca	gcccgtgtgt	cccgccttgg	cctgcgcagg	8564
50	cggaaagcga	ttgccgcgct	atcgaggatg	cgcgcgcgcg	ccgcgacatg	gccgccttgg	8624
	gtccgcgggt	cgcggcggaac	gccccttgca	tgacgcgcac	gatgatggcc	gcgcgcccgc	8684
	cgcctcgtcta	cctgacgggc	ggcagctggc	aggtgctgga	acgcctgtgg	caggcccgcg	8744
	cgcgcgggct	tgccggcctt	gcgcagatgg	atgcggcccc	gaacgtcaag	ctgatcttcg	8804
	aggaaagcag	cgcgcgcgac	gtgctgtacc	tggtccccga	cgcgcgcctg	atcgcgcctg	8864
55	tcgagggggcg	ttgaacgcgt	aagacgacca	ctgggttaagg	ttctgcgcgc	cgtggtctcg	8924
	actgcctgca	aagaggtgct	tgagttgctg	cgtgactgcg	gcggccgact	tcgtgggact	8984
	tgcccgcacc	gctgacgcgc	tggaacgcgc	cccgcggatt	acgaccgcgt	cattgcccctg	9044
	aaccaatttc	ccgtcgggtcg	ac				9066

```

<210> 43
<211> 340
<212> PRT
<213> 副球菌属菌种(Paracoccus sp.)
5 <400> 43
Met Ile Ser His Thr Pro Val Pro Thr Gln Trp Val Gly Pro Ile Leu
1 5 10 15
Phe Arg Gly Pro Val Val Glu Gly Pro Ile Ser Ala Pro Leu Ala Thr
10 20 25 30
Tyr Glu Thr Pro Leu Trp Pro Ser Thr Ala Arg Gly Ala Gly Val Ser
35 40 45
15 Arg His Ser Gly Gly Ile Gln Val Ser Leu Val Asp Glu Arg Met Ser
50 55 60
Arg Ser Ile Ala Leu Arg Ala His Asp Gly Ala Ala Ala Thr Ala Ala
65 70 75 80
20 Trp Gln Ser Ile Lys Ala Arg Gln Glu Glu Val Ala Ala Val Val Ala
85 90 95
Thr Thr Ser Arg Phe Ala Arg Leu Val Glu Leu Asn Arg Gln Ile Val
25 100 105 110
Gly Asn Leu Leu Tyr Ile Arg Ile Glu Cys Val Thr Gly Asp Ala Ser
115 120 125
30 Gly His Asn Met Val Thr Lys Ala Ala Glu Ala Val Gln Gly Trp Ile
130 135 140
Leu Ser Glu Tyr Pro Met Leu Ala Tyr Ser Thr Ile Ser Gly Asn Leu
35 145 150 155 160
Cys Thr Asp Lys Lys Ala Ser Ala Val Asn Gly Ile Leu Gly Arg Gly
165 170 175
40 Lys Tyr Ala Val Ala Glu Val Glu Ile Pro Arg Lys Ile Leu Thr Arg
180 185 190
Val Leu Arg Thr Ser Ala Glu Lys Met Val Arg Leu Asn Tyr Glu Lys
195 200 205
45 Asn Tyr Val Gly Gly Thr Leu Ala Gly Ser Leu Arg Ser Ala Asn Ala
210 215 220
His Phe Ala Asn Met Leu Leu Gly Phe Tyr Leu Ala Thr Gly Gln Asp
50 225 230 235 240
Ala Ala Asn Ile Ile Glu Ala Ser Gln Gly Phe Val His Cys Glu Ala
245 250 255
55 Arg Gly Glu Asp Leu Tyr Phe Ser Cys Thr Leu Pro Asn Leu Ile Met
260 265 270
Gly Ser Val Gly Ala Gly Lys Gly Ile Pro Ser Ile Glu Glu Asn Leu
275 280 285
60

```

	Ser	Arg	Met	Gly	Cys	Arg	Gln	Pro	Gly	Glu	Pro	Gly	Asp	Asn	Ala	Arg
	290						295					300				
5	Arg	Leu	Ala	Ala	Ile	Cys	Ala	Gly	Val	Val	Leu	Cys	Gly	Glu	Leu	Ser
	305					310					315					320
	Leu	Leu	Ala	Ala	Gln	Thr	Asn	Pro	Gly	Glu	Leu	Val	Arg	Thr	His	Met
					325					330					335	
10	Glu	Met	Glu	Arg												
				340												
	<210>	44														
	<211>	9066														
15	<212>	DNA														
	<213>	副球菌属菌种(Paracoccus sp.)														
	<220>															
	<221>	CDS														
	<222>	(3641)..(4690)														
20	<223>															
	<400>	44														
	ggatccggca	gctcgacacg	ccgcagaacc	tgtacgaacg	tcccgccagc	cgcttcgctcg										60
	cggaattcgt	cgggcgcggg	acggtggtgc	ccgtgcaggc	ccatgacggc	gcgggcccgcg										120
	ccgcatacct	gggggcccag	gtggcggtga	acgccgcccc	gcaatcgcg	tttgtcgatc										180
25	acgtctgcct	gcgccccgag	aaccttgcca	tctccgagac	gggcgacctg	cgcgccaagg										240
	tcgcgcgcgt	cacctatctt	ggcgggaaat	acctgctgga	aaccgtgctg	gattgcggca										300
	cccggctggt	gaccgagacc	cgcgcccgc	tcgatacggg	cgcgagctt	ggcctgacca										360
	tcaacgcccc	ctgggccttt	gccgaggatt	gaatggacag	cgtgaagatc	ctttcgggca										420
	tgggcgtgaa	gggcccctgc	tgcatacagg	tggatgtcgg	cgggatgcgc	ctgatccctc										480
30	attgcggggac	cgggcccggac	gagggcgcg	agttcgaccc	cgcctggctg	gcggacgcgg										540
	atgcggtgct	gatcacccat	gaccacgtgg	accatatcgg	cggcgcgcgt	cacgcggtcg										600
	cggcggggct	gccgatccat	gcgacgcggc	agacggcggg	gttgctgccc	gcgggggcgg										660
	atctgcgcct	gctgcccga	cgcggtgtca	cgcggtatcg	cggggtcgat	ctgacgaccg										720
	gtcgcaacgg	gcatgcgcg	ggcggcgctc	ggatgcattt	cgacatgggc	gaggggctgt										780
35	tctattccgg	cgactgggtc	gaggaatccg	actggttcgc	cttcgatccg	ccccgcctg										840
	cggggacggc	gattctcgac	tgtctctatg	gcggtttcga	cgtggcgcaa	tcggattgca										900
	tcgcggacct	ggacgacctg	ctcgaggtgc	tgcgggggca	ggtactgctg	ccggtgccgc										960
	catccggccg	cgcgcccgag	ctggccctgc	ggctgatccg	ccgccacgga	ccgggcagcg										1020
	tgatggtcga	cgacgcctgc	ctgcgggcca	tcgcgcaact	gcccagaggc	cgcgactgg										1080
40	cctacgccac	cgaggcacgc	tttcttgtct	gcgacacgcc	gaacgccgaa	agccggcgcg										1140
	gcatggcggc	atctgcaagc	atggcgcgat	gcgggcaggc	tggggcgggg	cgcgcatgtc										1200
	gtcttcaccg	ggcacatgaa	cgtccatgcg	cgcgattctt	gcgaccgccc	cggcgggcat										1260
	ttccgcgcgt	ggaacgtgca	tccgcgcgtg	cgcgaccagc	gacggatgct	ggaacggctg										1320
	gccgcgcggc	gctttgcccc	ggccttctgc	cccgaacccg	agatctatct	ggcgctggac										1380
45	atgggcgcgc	aggtcttcat	gcaccaggag	gtgacgccat	gatccccgcc	cgcagcttct										1440
	gcctgatccg	ccacggcgaa	acgaccgcca	atgcaggggc	gatcatcgcg	ggcgcaaccg										1500
	atgtgcccct	gacgccaagg	ggccgcgatc	aggcccgcgc	cctggcaggg	cgcgatggc										1560
	catcgggcat	cgcgctgttc	gccagcccga	tgtcgcgtgc	ccgcgatacc	gcgctgctgg										1620
	cctttccggg	gcgcgaccac	cagcccgaac	ccgatctgcg	cgaacgcgac	tggggcatct										1680
50	tcgaggggacg	ccccgtcgcc	gatctgcccc	cgcgcgaaat	cacgccgcag	gggggcgagg										1740
	gctgggacga	cgtgatggcc	cgcgtggacc	gcgcgatccg	gcggatctgc	gcgacctcgg										1800
	gcgatgcgct	gccggtgctg	gtctgccatt	cgggcgtgat	ccgtgccgcg	cgcgtgctgt										1860
	ggaccaccgg	gatcggggc	gacgtccgcg	ccaacgccac	gccgatcctg	ttcagcccgg										1920
	acggcgaccg	attaaaggaa	ggaacgatat	gaccgccacc	acccctgcg	tcgtcttcga										1980
55	acgtggacgg	cacgcttgcc	gaattcgacg	ccgaccgcct	gggccatctt	gtccacggca										2040
	cgaccaagca	ctgggaagcc	ttccaccacg	cgatggccga	cgccccgccc	atccccgagg										2100
	tcgcccgcct	gatgcgcaag	ctgaaggagg	ggggcgagac	ggtcgtcatc	tgctcggggc										2160
	ggccccgcgg	ctggcaggat	cagacgatcg	catggctgcg	caagcacgac	ctgcccttcg										2220
	acgggatcta	tctgcgcccc	gaggatcagg	acggcgccag	cgaccccag	gtcaagcgcc										2280
60	gcgccttagc	cgagatgcgc	gccgacgggc	tggcgccctg	gctggtcgtg	gacgaccggc										2340

	gggtccgtcgt	ggatgcctgg	cgggcccagg	ggctgggtctg	cctgcaatgc	gcgcgcggggg	2400
	acttctaggg	ccgcgcgcacg	ggggcgcgga	caggctgggc	gggaaaccgc	cccgccacca	2460
	tgctcctgcac	gcgtcgaacc	gcccgtccga	cgccggtttc	cgcacggaaa	cgcgcgggcaa	2520
	gttgacataa	cttgacgcg	acgtctcgat	tctgcccgcg	aagaatgcga	tgcattccaga	2580
5	tgatgcagaa	cgaagaagcg	gaagcgccc	tgaaagacca	gatgatttcc	cataccccgg	2640
	tgccccacgca	atgggtcggc	ccgatcctgt	tccgcggccc	cgtcgtcgag	ggccccgatca	2700
	gcgcgcgcgt	ggccacctac	gagacgccgc	tctggccctc	gaccgcgcgg	ggggcagggg	2760
	tttcccggca	ttcgggcggg	atccaggtct	cgctgggtcga	cgaacgcgatg	agccgctcga	2820
	tcgcgctgcg	ggcgcgatgac	ggggcggcgg	cgaccgccgc	ctggcagtcg	atcaaggccc	2880
10	gccaggaaga	ggtcgcggcc	gtggtcgcca	ccaccagccg	cttcgcccgc	cttgctcgagc	2940
	tgaatcgcca	gatcgtgggc	aacctgcttt	acatccgcac	cgaatgcgtg	acgggcgacg	3000
	cctcgggtca	caacatggtc	accaaggccg	ccgaggccgt	gcagggctgg	atcctgtcgg	3060
	aatacccgat	gctggcctat	tccacgatct	cggggaacct	gtgcaccgac	aagaaggcgt	3120
	cggcgggtcaa	cggcatcctg	ggccgcggca	aatacgccgt	cgccgaggtc	gagatcccgc	3180
15	gcaagatcct	gaccgcggtg	ctgcgcacca	gcgcccagaaa	gatgggtccgc	ctgaactacg	3240
	agaagaacta	tgctgggggt	acgtctggcg	ggctcgtcgcg	cagtgcgaac	gcgcattttcg	3300
	ccaacatgct	gctgggcttc	tacctggcga	cggggcagga	cgcgccaac	atcatcgagg	3360
	ccagccaggg	cttcgtccat	tgcgaggccc	gcggcgagga	tctgtatttc	tcgtgcacgc	3420
	tgcccaacct	catcatgggc	tcggtcgtg	ccggcaaggg	catcccctcg	atcgaggaga	3480
20	acctgtcgcg	gatgggctgc	cgccagccgg	gcgaaccggg	cgacaacgcg	cgccgtcttg	3540
	cggcgatctg	cgcgggcgtc	gtgctgtgtg	gtgaattgtc	gctgcttgcg	gcccagacca	3600
	accccgagaga	gttggtccgc	accacatgg	agatggagcg	atg acc gac agc aag		3655
					Met Thr Asp Ser Lys		
					1 5		
25							
	gat cac cat gtc gcg ggg cgc aag ctg gac cat ctg cgt gca ttg gac						3703
	Asp His His Val Ala Gly Arg Lys Leu Asp His Leu Arg Ala Leu Asp						
		10		15		20	
30							
	gac gat gcg gat atc gac cgg ggc gac agc ggc ttc gac cgc atc gcg						3751
	Asp Asp Ala Asp Ile Asp Arg Gly Asp Ser Gly Phe Asp Arg Ile Ala						
		25		30		35	
	ctg acc cat cgc gcc ctg ccc gag gtg gat ttc gac gcc atc gac acg						3799
35	Leu Thr His Arg Ala Leu Pro Glu Val Asp Phe Asp Ala Ile Asp Thr						
		40		45		50	
	gcg acc agc ttc ctg ggc cgt gaa ctg tcc ttc ccg ctg ctg atc tcg						3847
40	Ala Thr Ser Phe Leu Gly Arg Glu Leu Ser Phe Pro Leu Leu Ile Ser						
		55		60		65	
	tcc atg acc ggc ggc acc ggc gag gag atc gag cgc atc aac cgc aac						3895
45	Ser Met Thr Gly Gly Thr Gly Glu Glu Ile Glu Arg Ile Asn Arg Asn						
		70		75		80	
	ctg gcc gct ggt gcc gag gag gcc cgc gtc gcc atg gcg gtg ggc tcg						3943
	Leu Ala Ala Gly Ala Glu Glu Ala Arg Val Ala Met Ala Val Gly Ser						
		90		95		100	
50							
	cag cgc gtg atg ttc acc gac ccc tcg gcg cgg gcc agc ttc gac ctg						3991
	Gln Arg Val Met Phe Thr Asp Pro Ser Ala Arg Ala Ser Phe Asp Leu						
		105		110		115	
55							
	cgc gcc cat gcg ccc acc gtg ccg ctg ctg gcc aat atc ggc gcg gtg						4039
	Arg Ala His Ala Pro Thr Val Pro Leu Leu Ala Asn Ile Gly Ala Val						
		120		125		130	
60							
	cag ctg aac atg ggg ctg ggg ctg aag gaa tgc ctg gcc gcg atc gag						4087
	Gln Leu Asn Met Gly Leu Gly Leu Lys Glu Cys Leu Ala Ala Ile Glu						

	135	140	145	
5	gtg ctg cag gcg gac ggc ctg tat ctg cac ctg aac ccc ctg caa gag Val Leu Gln Ala Asp Gly Leu Tyr Leu His Leu Asn Pro Leu Gln Glu 150 155 160 165			4135
10	gcc gtc cag ccc gag ggg gat cgc gac ttt gcc gat ctg ggc agc aag Ala Val Gln Pro Glu Gly Asp Arg Asp Phe Ala Asp Leu Gly Ser Lys 170 175 180			4183
	atc gcg gcc atc gcc cgc gac gtt ccc gtg ccc gtc ctg ctg aag gag Ile Ala Ala Ile Ala Arg Asp Val Pro Val Pro Val Leu Leu Lys Glu 185 190 195			4231
15	gtg ggc tgc ggc ctg tcg gcg gcc gat atc gcc atc ggg ctg cgc gcc Val Gly Cys Gly Leu Ser Ala Ala Asp Ile Ala Ile Gly Leu Arg Ala 200 205 210			4279
20	ggg atc cgg cat ttc gac gtg gcc ggt cgc gcc gcc aca tcc tgg agc Gly Ile Arg His Phe Asp Val Ala Gly Arg Gly Gly Thr Ser Trp Ser 215 220 225			4327
25	cgg atc gag tat cgc cgc cgc cag cgg gcc gat gac gac ctg ggc ctg Arg Ile Glu Tyr Arg Arg Arg Gln Arg Ala Asp Asp Asp Leu Gly Leu 230 235 240 245			4375
	gtc ttc cag gac tgg ggc ctg cag acc gtg gac gcc ctg cgc gag gcg Val Phe Gln Asp Trp Gly Leu Gln Thr Val Asp Ala Leu Arg Glu Ala 250 255 260			4423
30	cgg ccc gcg ctt gcg gcc cat gat gga acc agc gtg ctg atc gcc agc Arg Pro Ala Leu Ala Ala His Asp Gly Thr Ser Val Leu Ile Ala Ser 265 270 275			4471
35	ggc gcc atc cgc aac ggt gtc gac atg gcg aaa tgc gtc atc ctg ggg Gly Gly Ile Arg Asn Gly Val Asp Met Ala Lys Cys Val Ile Leu Gly 280 285 290			4519
40	gcc gac atg tgc ggg gtc gcc gcg ccc ctg ctg aaa gcg gcc caa aac Ala Asp Met Cys Gly Val Ala Ala Pro Leu Leu Lys Ala Ala Gln Asn 295 300 305			4567
45	tcg cgc gag gcg gtt gta tcc gcc atc cgg aaa ctg cat ctg gag ttc Ser Arg Glu Ala Val Val Ser Ala Ile Arg Lys Leu His Leu Glu Phe 310 315 320 325			4615
	cgg aca gcc atg ttc ctc ctg ggt tgc ggc acg ctt gcc gac ctg aag Arg Thr Ala Met Phe Leu Leu Gly Cys Gly Thr Leu Ala Asp Leu Lys 330 335 340			4663
50	gac aat tcc tcg ctt atc cgt caa tga aagtgcctaa gatgaccgtg Asp Asn Ser Ser Leu Ile Arg Gln 345			4710
55	acaggaatcg aagcgatcag cttctacacc ccccagaact acgtgggact ggatatcctt gccgcgcac acgggatcga ccccgagaag ttctcgaagg ggatcgggca ggagaaaatc gcaactgccg gccatgacga ggatatcgtg accatggccg ccgaggccgc gctgccgatc atcgaacgcg cgggcacgca gggcatcgac acggttctgt tcgccaccga gagcgggatc gaccagtoga aggcgcgcgc catctatctg cgccgcctgc tggacctgtc gcccaactgc cgttgcgtcg agctgaagca ggccgtgctat tccgcgacgg cggcgctgca gatggcctgc gcgcgatgtcg cccgcaagcc cgaccgcaag gtgctggtga tcgcgtccga tgtcgcgcgc			4770 4830 4890 4950 5010 5070 5130

	tatgaccg	aaagctcggg	cgaggcgacg	caggggtgcgg	gcgcgctgc	catccttgtc	5190
	agcgccgac	ccaaggtggc	cgagatcggc	accgtctcgg	ggctgttcac	cgaggatata	5250
	atggatttct	ggcgccgaa	ccaccgccc	acgcccctgt	tcgacggcaa	ggcatcgacg	5310
	ctgcgctatc	tgaacgcgct	ggtcgaggcg	tggaaacgact	atcgcgcaa	tggcgccac	5370
5	gagttcgccg	atttcgcgca	tttctgctat	cacgtgccgt	tctcgcgat	ggcgagaag	5430
	gcgaacagcc	acctggccaa	ggcgaacaag	acgcccgtgg	acatggggca	ggtgcagacg	5490
	ggcctgatct	acaaccggca	ggtcgggaac	tgctataccg	ggtcgatcta	cctggcattc	5550
	gcctcgctgc	tggagaacgc	tcaggaggac	ctgaccggcg	cgctggtcgg	tctgttcagc	5610
	tatggctcgg	gtgcgacggg	cgaattcttc	gatgcgcgga	tcgcgcccgg	ttaccgcgac	5670
10	cacctgttcg	cggaacgcca	tcgcgaattg	ctgcaggatc	gcacgcccgt	cacatatgac	5730
	gaatacgttg	ccctgtggga	cgagatcgac	ctgacgcagg	gcgcgcccga	caaggcgcg	5790
	ggtcgtttca	ggctggcagg	tatcgaggac	gagaagcgca	tctatgtcga	ccggcaggcc	5850
	tgaagcaggc	gccccatgcc	cgggcaagct	gacccgtgtc	ggggaacatt	ccgtgctcta	5910
	tgggtgcgcc	gcgcttgcca	tggccatcgc	ccgtataacc	gaggtgtggt	tcacgcgct	5970
15	tggcattggc	gaggggatac	gcacgacatt	cgccaatctc	tcgggcccgg	cgacctattc	6030
	gctgaagctg	ctgtcgggg	tcaagtcgcg	gctggaccgc	cggttcgagc	agttcctgaa	6090
	cggcgaccta	aaggtgcaca	aggtcctgac	ccatcccgcg	gatctggcgg	tctatgcgct	6150
	ggcgtcgctt	ctgcacgaca	agccgcgggg	gaccgcggcg	atgccgggca	tcggcgcgat	6210
	gcaccacctg	ccgcgaccgg	gtgagctggg	cagccggacg	gagctgccc	tcggcgcggg	6270
20	catggggctg	tctgcggcca	tcgtcgcggc	caccacggtc	ctgttcgaga	cgctgctgga	6330
	ccggcccaag	acgcccgaac	agcgcttcga	ccgcgtccgc	ttctgcgagc	ggttgaagca	6390
	cggcagggcc	ggtcccacgc	acgcggccag	cgctgcgcgc	ggcgggcttg	tcgcgctggg	6450
	cgggaacggg	ccgggttcga	tcagcagctt	cgatttgccc	gaggatcacg	accttgctgc	6510
	gggacgcggc	tggtaactgg	tactgcacgg	gcgcccgcgc	agcgggaccg	gcgaatgcgt	6570
25	cagcgcgctc	gcggcgcgcg	atggtcgcga	tgcgcgctg	tgggacgcct	tcgcagctctg	6630
	caccgcgcgc	ttggaggccg	cgctgctgtc	tgggggcagc	cccgcgcgcg	ccatcaccga	6690
	gaaccagcgc	ctgctggaac	gcatcggcgt	cgctgcggca	gcgacgcagg	ccctcgtggc	6750
	ccagatcgag	gaggcggggtg	gcgcggccaa	gatctgcggc	gcaggttcgc	tcggggcgca	6810
	tcacggcggg	gcggctcctg	tcgggattga	cgacgcgcag	gcgatggctt	cggatcatggc	6870
30	gcgccatccc	gacctcgact	gggcgcccct	gcgcattgtc	cgacgggggg	cggcaccgcg	6930
	ccccgcgcgc	cgctgcgcaac	cgctgcgcgg	gcagggtcga	tggatcaggt	catccgcgc	6990
	agcgcgccgg	gttcgggtcat	gatcacgggc	gaacatgcgc	tggtctatgg	acaccgcgcc	7050
	atcgctgcgc	ggatcgagca	gcgcgcccac	gtgacgatcg	tcgcgcgtgc	cgaccgcgat	7110
	tttcgcatca	cctcgagat	cggggcgccg	cagcaggggt	cgctggacga	tctgcctgcg	7170
35	ggcgggacct	atcgcttcgt	gctggccgcc	atcgcgcgac	acgcgcggga	cctgccttgc	7230
	gggttcgaca	tggacatcac	ctcggggatc	gatccgaggc	tcgggcttgg	atcctcggcg	7290
	gcggtgacgg	tcgcctgcct	cggcgcgctg	tcgcggctgg	cggggcgggg	gaccgagggg	7350
	ctgcatgacg	acgcgctgcg	catcgctccg	gccatccagg	gcaggggcag	cggggccgat	7410
	ctggcgggca	gcctgcatgg	cggcttcgtc	gcctatcgcg	cgcccgatgg	cgggtgccgcg	7470
40	cagatcgag	cgcttcgggt	gcccgcgggg	cggttcggcc	tgcgctatgc	gggtacaag	7530
	accccgacct	ccgaggtgct	gcgccttggt	gcgatcgga	tggcgggcaa	cgaggccgct	7590
	ttcgacgcgc	tctactcccg	gatggggcga	agcgagatg	ccgcgatccg	cgcgcgcaa	7650
	gggctggact	gggtgcatt	ccacgacgcg	ctgaacgaat	accagcgcc	gatggagcag	7710
	ctgggcgtgt	ccgacgacac	gctggacgcg	atcatccgcg	aggcgcgcg	cgcgggcgcc	7770
45	gcagtcgcca	agatctccgg	ctcggggctg	ggggattgctg	tgctggcact	ggcgaccag	7830
	cccaaggggt	tcgtgcccgc	aagcattgcc	gagaagggac	ttgttttcga	tgactgatgc	7890
	cgtccgcgac	atgatcgccc	gtgccatggc	gggcgcgacc	gacatccgag	cagccgaggc	7950
	ttatgcgccc	agcaacatcg	cgctgtcgaa	atactggggc	aagcgcgacg	ccgcgcggaa	8010
	ccttcgcgtg	aacagctccg	tctcgatctc	gttggcgaa	tggggctctc	atacgcggt	8070
50	cgaggggtcc	ggcacggggc	acgacgaggt	gcatacaca	ggcacgctgc	tggatccggg	8130
	cgacgccttc	gcgcgcgcgc	cgttgccatt	cgctgacctg	ttccgggggg	ggaggcacct	8190
	gccgctgcgg	atcacgacgc	agaactcgat	cccgcggcg	gcggggcttg	cctcgtcggc	8250
	ctcggggttc	gcggcgctga	cccgtgcgct	ggcgggggcg	ttcgggcttg	atctggacga	8310
	cacggatctg	agccgcacgc	cccggatcgg	cagtggcagc	gccgcccgc	cgatctggca	8370
55	cggcttcgtc	cgctggaacc	ggggcgaggc	cgaggatggg	catgacagcc	acggcgtccc	8430
	gctggacctg	cgtggcccgc	gcttcgcgat	cgcgatcgtg	gccgtggaca	aggggcccga	8490
	gcctttcagt	tcgcgcgacg	gcatgaacca	cacggtcgag	accagcccgc	tggtcccgc	8550
	ctggcctgcg	caggcggaag	cggattgccg	cgtcatcgag	gatgcgatcg	ccgcccgcga	8610
	catggccgcc	ctgggtccgc	gggtcgaggc	gaacgccctt	gcgatgcacg	ccacgatgat	8670
60	ggcgcgcgc	ccgcgcgtct	gctacctgac	ggcgggcagc	tggcaggtgc	tggaaacgcct	8730

	gtggcaggcc cgcgcggacg ggcttgcggc ctttgcgacg atggatgccg gcccgaacgt	8790
	caagctgac ttcgaggaaa gcagcgccgc cgacgtgctg tacctgttcc ccgacgccag	8850
	cctgatcgcg ccgttcgagg ggcgttgaac gcgtaagacg accactgggt aaggttctgc	8910
	cgcgcggtggt ctcgactgcc tgcaaagagg tgcttgagtt gctgcgtgac tgcggcggcc	8970
5	gacttcgtgg gacttgcccg ccacgctgac gcgctggaaa cgcgcccgcg gattacgacc	9030
	gcgtcattgc cctgaaccaa tttcccgtcg gtcgac	9066
	<210> 45	
	<211> 349	
10	<212> PRT	
	<213> 副球菌属菌种 (Paracoccus sp.)	
	<400> 45	
	Met Thr Asp Ser Lys Asp His His Val Ala Gly Arg Lys Leu Asp His	
	1 5 10 15	
15	Leu Arg Ala Leu Asp Asp Asp Ala Asp Ile Asp Arg Gly Asp Ser Gly	
	20 25 30	
	Phe Asp Arg Ile Ala Leu Thr His Arg Ala Leu Pro Glu Val Asp Phe	
20	35 40 45	
	Asp Ala Ile Asp Thr Ala Thr Ser Phe Leu Gly Arg Glu Leu Ser Phe	
	50 55 60	
25	Pro Leu Leu Ile Ser Ser Met Thr Gly Gly Thr Gly Glu Glu Ile Glu	
	65 70 75 80	
	Arg Ile Asn Arg Asn Leu Ala Ala Gly Ala Glu Glu Ala Arg Val Ala	
	85 90 95	
30	Met Ala Val Gly Ser Gln Arg Val Met Phe Thr Asp Pro Ser Ala Arg	
	100 105 110	
35	Ala Ser Phe Asp Leu Arg Ala His Ala Pro Thr Val Pro Leu Leu Ala	
	115 120 125	
	Asn Ile Gly Ala Val Gln Leu Asn Met Gly Leu Gly Leu Lys Glu Cys	
	130 135 140	
40	Leu Ala Ala Ile Glu Val Leu Gln Ala Asp Gly Leu Tyr Leu His Leu	
	145 150 155 160	
	Asn Pro Leu Gln Glu Ala Val Gln Pro Glu Gly Asp Arg Asp Phe Ala	
45	165 170 175	
	Asp Leu Gly Ser Lys Ile Ala Ala Ile Ala Arg Asp Val Pro Val Pro	
	180 185 190	
50	Val Leu Leu Lys Glu Val Gly Cys Gly Leu Ser Ala Ala Asp Ile Ala	
	195 200 205	
	Ile Gly Leu Arg Ala Gly Ile Arg His Phe Asp Val Ala Gly Arg Gly	
	210 215 220	
55	Gly Thr Ser Trp Ser Arg Ile Glu Tyr Arg Arg Arg Gln Arg Ala Asp	
	225 230 235 240	
60	Asp Asp Leu Gly Leu Val Phe Gln Asp Trp Gly Leu Gln Thr Val Asp	
	245 250 255	

	Ala	Leu	Arg	Glu	Ala	Arg	Pro	Ala	Leu	Ala	Ala	His	Asp	Gly	Thr	Ser	
				260					265					270			
5	Val	Leu	Ile	Ala	Ser	Gly	Gly	Ile	Arg	Asn	Gly	Val	Asp	Met	Ala	Lys	
			275					280					285				
	Cys	Val	Ile	Leu	Gly	Ala	Asp	Met	Cys	Gly	Val	Ala	Ala	Pro	Leu	Leu	
		290					295					300					
10	Lys	Ala	Ala	Gln	Asn	Ser	Arg	Glu	Ala	Val	Val	Ser	Ala	Ile	Arg	Lys	
	305				310						315					320	
	Leu	His	Leu	Glu	Phe	Arg	Thr	Ala	Met	Phe	Leu	Leu	Gly	Cys	Gly	Thr	
15					325					330					335		
	Leu	Ala	Asp	Leu	Lys	Asp	Asn	Ser	Ser	Leu	Ile	Arg	Gln				
				340					345								
20	<210>	46															
	<211>	9066															
	<212>	DNA															
	<213>	副球菌属菌种 (Paracoccus sp.)															
25	<220>																
	<221>	CDS															
	<222>	(4687) .. (5853)															
	<223>																
	<400>	46															
30	ggatccggca	gctcgacacg	ccgcagaacc	tgtacgaacg	tcccggccagc	cgcttcgctcg		60									
	cggaattcgt	cgggcgcggg	acgggtggtgc	ccgtgcagggc	ccatgacggc	gcggggccgcg		120									
	cccgcatacct	gggggcccag	gtggcggtga	acgcgcggcc	gcaatcgcg	tttgtcgatc		180									
	acgtctgcct	gcgccccgag	aaccttgcca	tctccgagac	ggcgacactg	cgcgccaagg		240									
	tcgcgcgcgt	cacctatctt	ggcgggaaat	acctgctgga	aaccgtgctg	gattgcgcca		300									
35	ccgggctggt	gaccgagacc	cgcgcccgtc	tcgatacggg	cgcgacgctt	ggcctgacca		360									
	tcaacgcccc	ctgggccttt	gccgaggatt	gaatggacag	cgtgaagatc	ctttcgggca		420									
	tgggcgtgaa	gggcccgtgc	tgcatacggc	tggatgtcgg	cgggatgcgc	ctgatccctg		480									
	attgcgggac	cgggcccggc	gagggcgcg	agttcgaccc	cgccctggctg	gcggacgcgg		540									
	atgcggtgct	gatcacccat	gaccacgtgg	accatatcgg	cggcgcgcgt	cacgcggctg		600									
40	cgggcggggt	gccgatccat	gcgacgcggc	agacggcggg	gttgctgccc	gcggggcgcg		660									
	atctgcgcct	gctgcccga	cgcggtgtca	cgcgatcgc	cggggtcgat	ctgacgaccc		720									
	gtcgcaacgg	gcatagcgcg	ggcggcgtct	ggatgcattt	cgacatgggc	gaggggctgt		780									
	tctattccgg	cgactggtcc	gaggaatccg	actggttcgc	cttcgatccg	ccccgcctg		840									
	cggggacggc	gattctcgac	tgtctctatg	gcggtttcga	cgtggcgcaa	tcggattgca		900									
45	tcgcggacct	ggacgacctg	ctcgagggtg	tgcgggggca	ggtaactgctg	ccggtgccgc		960									
	catccggccg	cgcgcccgag	ctggccctgc	ggctgatccg	ccgccacgga	ccgggcagcg		1020									
	tgatggtcga	cgacgcctgc	ctgccggcca	tcgcgcaact	gcccagaggc	cgcggaactg		1080									
	cctacgccac	cgaggcacgc	tttcttgtct	gcgacacgcc	gaacgccgaa	agccggcgcg		1140									
	gcatggcggc	atctgcaagc	atggcgcgat	gcgggcaggg	tggggcgggg	cgcgcatgtc		1200									
50	gtcttcaccg	ggcacatgaa	cgtccatgcg	cgcgattctt	gcgaccgccc	cggcgggcat		1260									
	ttccgcccgt	ggaacgtgca	tccgccgctg	cgcgaccagc	gacggatgct	ggaacggctg		1320									
	gccgcgcggc	gctttgcccc	ggcctttctg	cccgaacccg	agatctatct	ggcgtgggac		1380									
	atgggcgcgc	aggtcttcat	gcaccaggag	gtgacgccat	gatccccgcc	cgagctttct		1440									
	gcctgatccg	ccacggcgaa	acgaccgcca	atgcaggggc	gatcatcgcg	ggcgcaaccg		1500									
55	atgtgcccct	gacgccaagg	ggccgcgcat	aggcccgcgc	cctggcaggg	cgcgaaatgg		1560									
	catcgggcat	cgcgctgttc	gccagcccga	tgtcgcgtgc	ccgcgatacc	gcgctgctgg		1620									
	cctttccggg	gcgcgaccac	cagcccgaac	ccgatctgcg	cgaacgcgac	tggggcatct		1680									
	tcgagggacg	ccccgtcgcc	gatctgcccc	cgcgcgaaat	cacgcgcgag	gggggcgagg		1740									
	gctgggacga	cgtgatggcc	cgcgtggacc	gcgcgatccg	gcggatctgc	gcgacctcgg		1800									
60	gcgatgcgct	gccggtgctg	gtctgccatt	cgggcgtgat	ccgtgccgcg	cgcgctgctgt		1860									

	ggaccaccgg	cgatgcgggc	gatcgtcgc	ccaacgccac	gccgatcctg	ttcagcccgg	1920
	acggcgaccg	attaaaggaa	ggaacgatat	gaccgccacc	accccctgcg	tcgtcttcga	1980
	acgtggacgg	cacgcttgcc	gaattcgacg	ccgaccgcct	gggccatctt	gtccacggca	2040
	cgaccaagca	ctgggacgcc	ttccaccacg	cgatggccga	cgccccgccc	atccccgagg	2100
5	tcgcccgcct	gatgcgcaag	ctgaaggagg	ggggcgagac	ggtcgtcatc	tgctcggggc	2160
	ggccccgcgg	ctggcaggat	cagacgatcg	catggctgcg	caagcacgac	ctgcccttcg	2220
	acgggatcta	tctgcgcccc	gaggatcagg	acggcgccag	cgaccccag	gtcaagcgcc	2280
	gcgcccctagc	cgagatgcgc	gccgacgggc	tggcgcccctg	gctggtcgtg	gacgaccggc	2340
	ggtccgtcgt	ggatgcctgg	cgggccgagg	ggctggctctg	cctgcaatgc	gcgcgggggg	2400
10	acttctaggg	ccgcgcgacg	ggggcgcgga	caggctgggc	gggaaaccgc	cccgccacca	2460
	tgtcctgcac	gogtcgaacc	gcccgtccga	cgccgggttc	cgcacgaaa	cgcgcgcaa	2520
	gttgacataa	cttgacgcg	acgtctcgat	tctgcccgcg	aagaatgcga	tgcatccaga	2580
	tgatgcagaa	cgaagaagcg	gaagcgcccc	tgaagacca	gatgatttcc	cataccccgg	2640
	tgccacgca	atgggtcggc	ccgatcctgt	tccgcggccc	cgctcgtcag	ggcccgatca	2700
15	gcgcgcgcgt	ggccacctac	gagacgccgc	tctggccctc	gaccgcgcgg	ggggcagggg	2760
	tttcccggca	ttcggggcgg	atccaggtct	cgctggctga	cgaacgcatg	agccgctcga	2820
	tcgcgctgcg	ggcgcatgac	ggggcgggcg	cgaccgccc	ctggcagtcg	atcaaggccc	2880
	gccaggaaga	ggtcgcggcc	gtggtcgcca	ccaccagccg	cttcgcccgc	cttgctcgagc	2940
	tgaatcgcca	gatcgtgggc	aacctgcttt	acatccgcat	cgaatgcgtg	acgggcgacg	3000
20	cctcgggtca	caacatggtc	accaaggccg	ccgaggccgt	gcagggctgg	atcctgtcgg	3060
	aatacccgat	gctggcctat	tccacgatct	cggggaacct	gtgcaccgac	aagaaggcgt	3120
	cggcggtcaa	cgccatcctg	ggccgcggca	aataaccctg	cgccgaggtc	gagatcccgc	3180
	gcaagatcct	gaccgcgctg	ctgcgcacca	gcgccgagaa	gatggtccgc	ctgaactacg	3240
	agaagaacta	tgtcgggggt	acgctggcgg	ggctcgtcgc	cagtgcgaac	gcgcatttcg	3300
25	ccaacatgct	gctgggcttc	tacctggcga	cggggcagga	cgcgcccaac	atcatcgagg	3360
	ccagccaggg	cttcgtccat	tgcgaggccc	gcggcgagga	tctgtatttc	tcgtgcacgc	3420
	tgcccaacct	catcatgggc	tcggtcgggtg	ccggcaaggg	catcccctcg	atcgaggaga	3480
	acctgtcgcg	gatgggctgc	cgccagcccg	gcgaaccg	cgacaacgcg	cgccgtcttg	3540
	cggcgatctg	cgcgggcgtc	gtgctgtgtg	gtgaattgtc	gctgcttgcg	gccagacca	3600
30	accccggaga	gttgggtccg	accacatgg	agatggagcg	atgaccgaca	gcaaggatca	3660
	ccatgtcgcg	gggcgcaagc	tggaccatct	gcgtgcattg	gacgacgatg	cggatatcga	3720
	ccggggcgac	agcggcttcg	accgcatcgc	gctgacccat	cgcgccctgc	ccgaggtgga	3780
	tttcgacgcc	atcgacacgg	cgaccagctt	cctgggccgt	gaactgtcct	ttccgctgct	3840
	gatctcgtcc	atgaccggcg	gcaccggcga	ggagatcgag	cgcatcaacc	gcaacctggc	3900
35	cgctgggtgcc	gaggaggccc	gcgtcgccat	ggcgggtggc	tcgcagcgcg	tgatgttcac	3960
	cgacccctcg	gcgcgggcca	gcttcgacct	gcgcgcccat	gcgcccaccg	tgccgctgct	4020
	ggccaatatc	ggcgcggtgc	agctgaacat	ggggctgggg	ctgaagggaat	gcctggccgc	4080
	gatcgagggtg	ctgcaggcgg	acggcctgta	tctgcacctg	aaccccctgc	aagaggccgt	4140
	ccagcccagag	ggggatcgcg	actttgcca	tctgggcagc	aagatcgcg	ccatcgccc	4200
40	cgacgttccc	gtgcccgctc	tgctgaagga	ggtgggctgc	ggcctgtcgg	cggccgatat	4260
	cgccatcggg	ctgcgcgcgg	ggatccggca	tttcgacgtg	gcccgtcgcg	gcggcacatc	4320
	ctggagccgg	atcgagtatc	gccgcggcca	gcgggcccga	gacgacctgg	gcctgggtctt	4380
	ccaggactgg	ggcctgcaga	ccgtggacgc	cctgcgcgag	gcgcggcccc	cgcttgccgc	4440
	ccatgatgga	accagcgtgc	tgatcgccag	cggcggcatc	cgcaacggtg	tcgacatggc	4500
45	gaaatgcgtc	atcctggggg	ccgacatgtg	cggggtcgcc	gcgcccctgc	tgaagcggc	4560
	ccaaaactcg	cgcgaggcgg	ttgtatccgc	catccggaaa	ctgcatctgg	agttccggac	4620
	agccatgttc	ctcctgggtt	gcggcacgct	tgccgacctg	aaggacaatt	cctcgcttat	4680
	ccgtca atg	aaa gtg cct	aag atg acc	gtg aca gga	atc gaa gcg	atc	4728
	Met	Lys Val Pro	Lys Met Thr	Val Thr	Gly Ile Glu	Ala Ile	
50	1		5		10		
	agc ttc tac	acc ccc cag	aac tac gtg	gga ctg gat	atc ctt gcc	gcg	4776
	Ser Phe Tyr	Thr Pro Gln	Asn Tyr Val	Gly Leu Asp	Ile Leu Ala	Ala	
	15		20		25		30
55	cat cac ggg	atc gac ccc	gag aag ttc	tcg aag ggg	atc ggg cag	gag	4824
	His His Gly	Ile Asp Pro	Glu Lys Phe	Ser Lys Gly	Ile Gly Gln	Glu	
		35		40		45	
60	aaa atc gca	ctg ccc ggc	cat gac gag	gat atc gtg	acc atg gcc	gcc	4872

	Lys	Ile	Ala	Leu	Pro	Gly	His	Asp	Glu	Asp	Ile	Val	Thr	Met	Ala	Ala	
				50					55					60			
5	gag	gcc	gcg	ctg	ccg	atc	atc	gaa	cgc	gcg	ggc	acg	cag	ggc	atc	gac	4920
	Glu	Ala	Ala	Leu	Pro	Ile	Ile	Glu	Arg	Ala	Gly	Thr	Gln	Gly	Ile	Asp	
			65					70					75				
10	acg	gtt	ctg	ttc	gcc	acc	gag	agc	ggg	atc	gac	cag	tcg	aag	gcc	gcc	4968
	Thr	Val	Leu	Phe	Ala	Thr	Glu	Ser	Gly	Ile	Asp	Gln	Ser	Lys	Ala	Ala	
		80						85					90				
15	gcc	atc	tat	ctg	cgc	cgc	ctg	ctg	gac	ctg	tcg	ccc	aac	tgc	cgt	tgc	5016
	Ala	Ile	Tyr	Leu	Arg	Arg	Leu	Leu	Asp	Leu	Ser	Pro	Asn	Cys	Arg	Cys	
	95					100					105					110	
	gtc	gag	ctg	aag	cag	gcc	tgc	tat	tcc	gcg	acg	gcg	gcg	ctg	cag	atg	5064
	Val	Glu	Leu	Lys	Gln	Ala	Cys	Tyr	Ser	Ala	Thr	Ala	Ala	Leu	Gln	Met	
					115					120					125		
20	gcc	tgc	gcg	cat	gtc	gcc	cgc	aag	ccc	gac	cgc	aag	gtg	ctg	gtg	atc	5112
	Ala	Cys	Ala	His	Val	Ala	Arg	Lys	Pro	Asp	Arg	Lys	Val	Leu	Val	Ile	
				130					135					140			
25	gcg	tcc	gat	gtc	gcg	cgc	tat	gac	cgc	gaa	agc	tcg	ggc	gag	gcg	acg	5160
	Ala	Ser	Asp	Val	Ala	Arg	Tyr	Asp	Arg	Glu	Ser	Ser	Gly	Glu	Ala	Thr	
			145					150					155				
	cag	ggt	gcg	ggc	gcc	gtc	gcc	atc	ctt	gtc	agc	gcc	gat	ccc	aag	gtg	5208
	Gln	Gly	Ala	Gly	Ala	Val	Ala	Ile	Leu	Val	Ser	Ala	Asp	Pro	Lys	Val	
		160					165					170					
30	gcc	gag	atc	ggc	acc	gtc	tcg	ggg	ctg	ttc	acc	gag	gat	atc	atg	gat	5256
	Ala	Glu	Ile	Gly	Thr	Val	Ser	Gly	Leu	Phe	Thr	Glu	Asp	Ile	Met	Asp	
	175					180					185					190	
35	ttc	tgg	cgg	ccg	aac	cac	cgc	cgc	acg	ccc	ctg	ttc	gac	ggc	aag	gca	5304
	Phe	Trp	Arg	Pro	Asn	His	Arg	Arg	Thr	Pro	Leu	Phe	Asp	Gly	Lys	Ala	
					195					200					205		
40	tcg	acg	ctg	cgc	tat	ctg	aac	gcg	ctg	gtc	gag	gcg	tgg	aac	gac	tat	5352
	Ser	Thr	Leu	Arg	Tyr	Leu	Asn	Ala	Leu	Val	Glu	Ala	Trp	Asn	Asp	Tyr	
				210					215					220			
45	cgc	gcg	aat	ggc	ggc	cac	gag	ttc	gcc	gat	ttc	gcg	cat	ttc	tgc	tat	5400
	Arg	Ala	Asn	Gly	Gly	His	Glu	Phe	Ala	Asp	Phe	Ala	His	Phe	Cys	Tyr	
			225					230					235				
50	cac	gtg	ccg	ttc	tcg	cgg	atg	ggc	gag	aag	gcg	aac	agc	cac	ctg	gcc	5448
	His	Val	Pro	Phe	Ser	Arg	Met	Gly	Glu	Lys	Ala	Asn	Ser	His	Leu	Ala	
		240					245					250					
	aag	gcg	aac	aag	acg	ccg	gtg	gac	atg	ggg	cag	gtg	cag	acg	ggc	ctg	5496
	Lys	Ala	Asn	Lys	Thr	Pro	Val	Asp	Met	Gly	Gln	Val	Gln	Thr	Gly	Leu	
	255					260					265					270	
55	atc	tac	aac	cgg	cag	gtc	ggg	aac	tgc	tat	acc	ggg	tcg	atc	tac	ctg	5544
	Ile	Tyr	Asn	Arg	Gln	Val	Gly	Asn	Cys	Tyr	Thr	Gly	Ser	Ile	Tyr	Leu	
					275				280						285		
60	gca	ttc	gcc	tcg	ctg	ctg	gag	aac	gct	cag	gag	gac	ctg	acc	ggc	gcg	5592
	Ala	Phe	Ala	Ser	Leu	Leu	Glu	Asn	Ala	Gln	Glu	Asp	Leu	Thr	Gly	Ala	

	290	295	300	
5	ctg gtc ggt ctg ttc agc tat ggc tcg ggt gcg acg ggc gaa ttc ttc Leu Val Gly Leu Phe Ser Tyr Gly Ser Gly Ala Thr Gly Glu Phe Phe 305 310 315	5640		
10	gat gcg cgg atc gcg ccc ggt tac cgc gac cac ctg ttc gcg gaa cgc Asp Ala Arg Ile Ala Pro Gly Tyr Arg Asp His Leu Phe Ala Glu Arg 320 325 330	5688		
15	cat cgc gaa ttg ctg cag gat cgc acg ccc gtc aca tat gac gaa tac His Arg Glu Leu Leu Gln Asp Arg Thr Pro Val Thr Tyr Asp Glu Tyr 335 340 345 350	5736		
20	gtt gcc ctg tgg gac gag atc gac ctg acg cag ggc gcg ccc gac aag Val Ala Leu Trp Asp Glu Ile Asp Leu Thr Gln Gly Ala Pro Asp Lys 355 360 365	5784		
25	gcg cgc ggt cgt ttc agg ctg gca ggt atc gag gac gag aag cgc atc Ala Arg Gly Arg Phe Arg Leu Ala Gly Ile Glu Asp Glu Lys Arg Ile 370 375 380	5832		
30	tat gtc gac cgg cag gcc tga agcaggcgcc catgccccgg gcaagctgat Tyr Val Asp Arg Gln Ala 385	5883		
35	cctgtccggg gaacattccg tgctctatgg tgcgcccgcg cttgccatgg ccatcgccccg ctataccgag gtgtggttca cgccgcttgg cattggcgag gggatacgca cgacattcgc caatctctcg ggcggggcga cctattcgct gaagctgctg tcgggggttca agtcgcgggt ggaccgcccg ttcgagcagt tctgaacgg cgacctaaag gtgcacaagg tcttgaccca tcccagacgat ctggcggtct atgcgctggc gtcgcttctg cacgacaagc cgccggggac cgccgcgatg ccgggcatcg gcgcgatgca ccacctgccg cgaccgggtg agctgggcag ccggacggag ctgccatcg gcgcgggcat ggggtcgtct gcggccatcg tcgcggccac cacggtcctg ttcgagacgc tgctggaccg gcccaagacg cccgaacagc gcttcgaccg cgtccgcttc tgcgagcggg tgaagcacgg caaggccggt cccatcgacg cggccagcgt cgtgcgcggc gggcttgtcc gcgtgggcgg gaacggggccg gggttcgatca gcagcttcga tttgcggcag gatcacgacc ttgtcgcggg acgcggctgg tactgggtac tgcacgggcg ccccgtcagc gggaccggcg aatgcgtcag cgcggtcgcg gcggcgcatg gtcgcgatgc ggcgtgtgg gacgccttcg cagtctgcac ccgcgcgttg gaggccgcgc tgctgtctgg gggcagcccc gacgcgcgca tcaccgagaa ccagcgccgt ctggaacgca tcggcgctgt gcccgcagcg acgcaggccc tcgtggccca gatcgaggag gcgggtggcg cggccaagat ctgcggcgca ggttcctgtc gggcgatca cggcggggcg gtcctcgtgc ggattgacga cgcgcaggcg atggcttcgg tcatggcgcg ccatcccagc ctgcactggg gcgccctgcg catgtcgcgc acgggggcgg caccgcgcgt cgcgcaaccg tgcgggggca gggctgatgg atcaggtcat ccgcgccagc gcgcgggtt cgggtcatgat cacgggcgaa catgccgtgg tctatggaca ccgcgccatc gtgcgcggga tcgagcagcg cgcccatgtg acgatcgtcc cgcgtgccga ccgcattgtt cgcattacat cgcagatcgg ggccgcgag caggggtcgc tggacgatct gcctgcgggc gggacctatc gcttcgtgct ggccgccatc gcgcgacacg cgccggacct gccttgccgg ttcgacatgg acatcacctc ggggatcgat ccgaggctcg ggcttgatc ctgcgcggcg gtgacggtcg cctgcctcgg cgcgtgtcg cggttgccgg ggcgggggac cgaggggctg catgacgacg cgctgcgcat cgtccgcgcc atccagggca ggggcagcgg gccgatctg gcggccagcc tgcattggcg ctcggtgcg tatcgcgcgc ccgatggcgg tgccgcgcag atcgaggcgc ttccgggtgcc gccggggcgc ttcggcctgc gctatgcggg ctacaagacc ccgacagccg aggtgctgcg ccttgtggcc gatcggtatg cgggcaacga ggccgcttct gacgcgctct actcccggat gggcgcaagc gcagatgccg cgatccgcgc ggcgcaaggc ctggactggg ctgcattcca cgacgcgctg aacgaatacc agcgcctgat ggagcagctg ggctgtccg acgacacgct ggacgcgac atccgcgagg cgcgcgacgc gggcgccgca gtcgccaaag tctccggctc ggggctgggg gattgcgtgc tggcactggg cgaccagccc aagggttctg tgcccgaag cattgccgag aagggacttg ttttcgatga ctgatgccg ccgcgacatg atcgcccgtg ccatggcggg cgcgaccgac atccgagcag ccgaggctta tgcgcccagc aacatcgcgc tgtcgaaata	5943 6003 6063 6123 6183 6243 6303 6363 6423 6483 6543 6603 6663 6723 6783 6843 6903 6963 7023 7083 7143 7203 7263 7323 7383 7443 7503 7563 7623 7683 7743 7803 7863 7923 7983		

	ctggggcaag	cgcgacgccg	cgcggaacct	tccgctgaac	agctccgtct	cgatctcggt	8043
	ggcgaactgg	ggctctcata	cgcgggtcga	gggggtccggc	acggggccacg	acgaggtgca	8103
	tcacaacggc	acgtgtctgg	atccggggcga	cgccttcgcg	cgcgcgcgct	tggcattcgc	8163
	tgacctgttc	cgggggggga	ggcacctgcc	gctgcggatc	acgacgcaga	actcgatccc	8223
5	gacggcggcg	gggcttgcct	cgtcggcctc	ggggttcgcg	gcgctgaccc	gtgcgctggc	8283
	ggggcgcttc	gggctggatc	tggacgacac	ggatctgagc	cgcacgcgcc	ggatcggcag	8343
	tggcagcgcc	gcccgcctga	tctggcacgg	cttcgtccgc	tggaaaccggg	gcgaggccga	8403
	ggatgggcat	gacagccacg	gcgtcccgtc	ggacctgcgc	tggcccggct	tccgcacgc	8463
	gatcgtggcc	gtggacaagg	ggcccaagcc	tttcagttcg	cgcgacggca	tgaaccacac	8523
10	ggtcgagacc	agcccgtgtg	tcccgccttg	gcctgcgcag	gcggaagcgg	attgccgcgt	8583
	catcgaggat	gcgatcgccg	cccgcgacat	ggccgccttg	ggtccgcggg	tcgaggcgaa	8643
	cgcccttgcg	atgcacgcca	cgatgatggc	cgcgcgcccg	ccgctctgct	acctgacggg	8703
	cggcagctgg	caggtgctgg	aacgcctgtg	gcaggcccgc	gcggacgggc	ttgcggcctt	8763
	tgcgacgatg	gatgccggcc	cgaacgtcaa	gctgatcttc	gaggaaagca	gcgccgccga	8823
15	cgtgctgtac	ctgttccccg	acgccagcct	gatcgcgccg	ttcgaggggc	gttgaacgcg	8883
	taagacgacc	actgggtaag	gttctgccgc	gcgtggcttc	gactgcctgc	aaagaggtgc	8943
	ttgagttgct	gcgtgactgc	ggcggccgac	ttcgtgggac	ttgcccgcca	cgctgacgcg	9003
	ctggaacgcg	gcccgcggat	tacgaccgcg	tcattgccct	gaaccaattt	cccgtcggtc	9063
	gac						9066
20	<210> 47						
	<211> 388						
	<212> PRT						
	<213> 副球菌属菌种 (Paracoccus sp.)						
25	<400> 47						
	Met Lys Val Pro Lys Met Thr Val Thr Gly Ile Glu Ala Ile Ser Phe						
	1	5		10		15	
	Tyr Thr Pro Gln Asn Tyr Val Gly Leu Asp Ile Leu Ala Ala His His						
30		20		25		30	
	Gly Ile Asp Pro Glu Lys Phe Ser Lys Gly Ile Gly Gln Glu Lys Ile						
		35		40		45	
35	Ala Leu Pro Gly His Asp Glu Asp Ile Val Thr Met Ala Ala Glu Ala						
		50		55		60	
	Ala Leu Pro Ile Ile Glu Arg Ala Gly Thr Gln Gly Ile Asp Thr Val						
40		65		70		75	
	Leu Phe Ala Thr Glu Ser Gly Ile Asp Gln Ser Lys Ala Ala Ala Ile						
		85		90		95	
45	Tyr Leu Arg Arg Leu Leu Asp Leu Ser Pro Asn Cys Arg Cys Val Glu						
		100		105		110	
	Leu Lys Gln Ala Cys Tyr Ser Ala Thr Ala Ala Leu Gln Met Ala Cys						
		115		120		125	
50	Ala His Val Ala Arg Lys Pro Asp Arg Lys Val Leu Val Ile Ala Ser						
		130		135		140	
	Asp Val Ala Arg Tyr Asp Arg Glu Ser Ser Gly Glu Ala Thr Gln Gly						
55		145		150		155	
	Ala Gly Ala Val Ala Ile Leu Val Ser Ala Asp Pro Lys Val Ala Glu						
		165		170		175	
60	Ile Gly Thr Val Ser Gly Leu Phe Thr Glu Asp Ile Met Asp Phe Trp						

	180	185	190	
	Arg Pro Asn His Arg Arg Thr Pro Leu Phe Asp Gly Lys Ala Ser Thr			
	195	200	205	
5	Leu Arg Tyr Leu Asn Ala Leu Val Glu Ala Trp Asn Asp Tyr Arg Ala			
	210	215	220	
10	Asn Gly Gly His Glu Phe Ala Asp Phe Ala His Phe Cys Tyr His Val			
	225	230	235	240
	Pro Phe Ser Arg Met Gly Glu Lys Ala Asn Ser His Leu Ala Lys Ala			
	245	250	255	
15	Asn Lys Thr Pro Val Asp Met Gly Gln Val Gln Thr Gly Leu Ile Tyr			
	260	265	270	
20	Asn Arg Gln Val Gly Asn Cys Tyr Thr Gly Ser Ile Tyr Leu Ala Phe			
	275	280	285	
	Ala Ser Leu Leu Glu Asn Ala Gln Glu Asp Leu Thr Gly Ala Leu Val			
	290	295	300	
25	Gly Leu Phe Ser Tyr Gly Ser Gly Ala Thr Gly Glu Phe Phe Asp Ala			
	305	310	315	320
	Arg Ile Ala Pro Gly Tyr Arg Asp His Leu Phe Ala Glu Arg His Arg			
	325	330	335	
30	Glu Leu Leu Gln Asp Arg Thr Pro Val Thr Tyr Asp Glu Tyr Val Ala			
	340	345	350	
	Leu Trp Asp Glu Ile Asp Leu Thr Gln Gly Ala Pro Asp Lys Ala Arg			
35	355	360	365	
	Gly Arg Phe Arg Leu Ala Gly Ile Glu Asp Glu Lys Arg Ile Tyr Val			
	370	375	380	
40	Asp Arg Gln Ala			
	385			
	<210> 48			
	<211> 9066			
45	<212> DNA			
	<213> 副球菌属菌种(Paracoccus sp.)			
	<220>			
	<221> CDS			
	<222> (5834)..(6970)			
50	<223>			
	<400> 48			
	ggatccggca gctcgacacg ccgcagaacc tgtacgaacg tcccgccagc cgcttcgctcg			60
	cggaattcgt cgggcgcggg acggtggtgc ccgtgcaggc ccatgacggc gcgggcccgcg			120
	ccgcatacct gggggccgag gtggcgggtga acgcccggc gcaatcgcg tttgtcgatc			180
55	acgtctgcct gcgccccgag aaccttgcca tctccgagac gggcgacctg cgcgccaagg			240
	tcgcgcgcgt cacctatctt ggcgggaaat acctgctgga aaccgtgctg gattgcggca			300
	ccgggctggt gaccgagacc cgcgccgct tcgatacggg cgcgcagctt ggcctgacca			360
	tcaacgcccc ctgggccttt gccgaggatt gaatggacag cgtgaagatc ctttcgggca			420
	tgggcgtgaa gggccctgcc tgcacagggc tggatgtcgg cgggatgcgc ctgatcctcg			480
60	attgcgggac cggcccggac gagggcgcgg agttcgaccc cgcctggctg gcggacgcgg			540

	atgcgggtgct	gataccccat	gaccacgtgg	accatatcgg	cggcgcgctg	cacgcggtcg	600
	cggcgggggct	gccgatccat	gcgacgcggc	agacggcggg	gttgctgccc	gcggggggcgg	660
	atctgcgcct	gctgcccga	cgcggtgtca	cgcggatcgc	cggggtcgat	ctgacgaccg	720
	gtcgcaacgg	gcatgccg	ggcggcgtct	ggatgcattt	cgacatgggc	gaggggctgt	780
5	tctattccgg	cgactgggtc	gaggaatccg	actggttcgc	cttcgatccg	ccccgcctg	840
	cggggacggc	gattctcgac	tgtctctatg	gcggtttcga	cgtggcgcaa	tcggattgca	900
	tcgoggacct	ggacgacctg	ctcgaggtgc	tgccggggca	ggtactgctg	ccggtgccgc	960
	catccggccg	cgcggccgag	ctggccctgc	ggctgatccg	ccgccacgga	ccgggcagcg	1020
	tgatggtcga	cgacgcctgc	ctgcgggcca	tcgcgcaact	gcccgaaggc	cgcggactgg	1080
10	cctacgccac	cgaggcacgc	tttcttgtct	gcgacacgcc	gaacgccgaa	agccggcgcg	1140
	gcatggcggc	atctgcaagc	atggcgcgat	gcgggcaggc	tggggcggga	cgcgcagtgc	1200
	gtcttcaccg	ggcacatgaa	cgtccatg	cgcgcattct	gcgaccgccc	cggcgggcat	1260
	ttccgcgcgt	ggaacgtgca	tccgcgcgtg	cgcgaccagc	gacggatgct	ggaacggctg	1320
	gcgcgcggc	gctttgcccc	ggccttctgc	cccgaccccg	agatctatct	ggcgctggac	1380
15	atgggcgcgc	aggtcttcat	gcaccaggag	gtgacgccat	gatccccgcc	cgcagcttct	1440
	gcctgatccg	ccacggcgaa	acgaccgcca	atgcaggggc	gatcatcgcg	ggcgcaaccg	1500
	atgtgcccct	gacgccaaag	ggccgcgatc	aggcccgcg	cctggcaggg	cgcgaatggc	1560
	catcgggcat	cgcgctgttc	gccagcccga	tgtcgcgtgc	ccgcgatacc	gcgctgctgg	1620
	cctttccggg	gcgcgaccac	cagcccgaac	ccgatctg	cgaacgcgac	tggggcatct	1680
20	tcgagggacg	ccccgtcgcc	gatctgcccc	cgcgcgaaat	cacgcgcgag	gggggcgagg	1740
	gctgggacga	cgtgatggcc	cgcgtggacc	gcgcgatccg	gcggatctgc	gcgacctcgg	1800
	gcgatgcgct	gccgggtgctg	gtctgccatt	cgggcgtgat	ccgtgcccgc	cgcgtgctgt	1860
	ggaccaccg	cgatgcgggc	gatcgtccgc	ccaacgccac	gccgatcctg	ttcagcccgc	1920
	acggcgaccg	attaaaggaa	ggaacgatat	gaccgccacc	acccctgcg	tcgtcttcga	1980
25	acgtggacgg	cacgcttgcc	gaattcgacg	ccgaccgcct	gggccatctt	gtccacggca	2040
	cgaccaagca	ctgggacgcc	ttccaccacg	cgatggccga	cgcgccgccc	atccccgagg	2100
	tcgcccgcct	gatgcgcaag	ctgaaggagg	ggggcgagac	ggtcgtcatc	tgtcggggc	2160
	ggccccgcgg	ctggcaggat	cagacgatcg	catggctg	caagcacgac	ctgcccctcg	2220
	acgggatcta	tctgcgcccc	gaggatcagg	acggcgccag	cgacccccgag	gtcaagcgcc	2280
30	gcgccctagc	cgagatgcgc	gccgacgggc	tggcgccctg	gctggctcgtg	gacgacgggc	2340
	ggtccgtcgt	ggatgcctgg	cgggcccagg	ggctggctcgt	cctgcaatgc	gcgcggggg	2400
	acttctagg	ccgcgcgacg	ggggcgcgga	caggctgggc	gggaaaccgc	cccgccacca	2460
	tgtcctgcac	gcgtcgcaacc	gcccgtccga	cgcgggtttc	cgcacggaaa	cgcgcggcaa	2520
	gttgacataa	cttgacgcg	acgtctcgat	tctgcccgcg	aagaatgoga	tgcattccaga	2580
35	tgatgcagaa	cgaagaagcg	gaagcgcccc	tgaagagacca	gatgatttcc	cataccccgg	2640
	tgcccacgca	atgggtcggc	ccgatcctgt	tccgcggccc	cgtcgtcgag	ggcccgatca	2700
	gcgcgcgcgt	ggccacctac	gagacgcgc	tctggccctc	gacgcgcggg	ggggcagggg	2760
	tttcccggca	ttcgggcggg	atccaggctc	cgtcggctga	cgaacgcgatg	agccgctcga	2820
	tcgcgctg	ggcgcatgac	ggggcgggcg	cgaccgcgc	ctggcagtcg	atcaaggccc	2880
40	gccaggaaga	ggtcgcggcc	gtggtcgcca	ccaccagccg	cttcgcccgc	cttgtcgagc	2940
	tgaatcgcca	gatcgtgggc	aacctgcttt	acatccgc	cgaatgcgtg	acgggcgacg	3000
	cctcgggtca	caacatggtc	accaaggccg	ccgaggccgt	gcagggtcgtg	atcctgtcgg	3060
	aatacccgat	gctggcctat	tccacgatct	cggggaaact	gtgcaccgac	aagaaggcgt	3120
	cggcgggtcaa	cggcatcctg	ggccgcggca	aatacgccgt	cgcgcagggtc	gagatccgc	3180
45	gcaagatcct	gacccgcgtg	ctgcgcacca	gcgcgcgaga	gatgggtcgc	ctgaactacg	3240
	agaagaacta	tgtcgggggt	acgtcggcgg	ggtcgtcgtg	cagtgcgaac	gcgcatttctg	3300
	ccaacatgct	gctgggcttc	tacctggcga	cggggcagga	cgcggccaac	atcatcgagg	3360
	ccagccaggg	cttcgtccat	tgcgaggccc	gcggcgagga	tctgtatttc	tcgtgcacgc	3420
	tgcccacact	catcatgggc	tcggtcgggtg	ccggcaaggg	catccccctg	atcgaggaga	3480
50	acctgtcgcg	gatgggctgc	cgcacgcggg	gcgaaccggg	cgacaacgcg	cgcgctcttg	3540
	cggcgatctg	cgcgggcgtc	gtgctgtgtg	gtgaattgtc	gctgcttgcg	gccagacca	3600
	accccggaga	gttggtccgc	acccacatgg	agatggagcg	atgaccgaca	gcaaggatca	3660
	ccatgtcgcg	gggcgcaagc	tggaccatct	gcgtgcattg	gacgacgatg	cggatatcga	3720
	ccggggcgac	agcggcttcg	accgcacgc	gctgacccat	cgcgcctcgc	ccgagggtga	3780
55	tttcgacgcc	atcgacacgg	cgaccagctt	cctgggcccgt	gaactgtcct	ttccgctgct	3840
	gatctcgtcc	atgaccggcg	gcaccggcga	ggagatcgag	cgcacacacc	gcaacctggc	3900
	cgtcgtgccc	gaggaggccc	gcgtcgccat	ggcgggtggg	tcgcagcgcg	tgatgttcac	3960
	cgacccctcg	gcgcggggcca	gcttcgaact	gcgcgcccat	gcgccaccg	tgccgctgct	4020
	ggccaatatc	ggcgcgggtg	agctgaacat	ggggctgggg	ctgaagggaat	gcctggccgc	4080
60	gatcgagggtg	ctgcaggcgg	acggcctgta	tctgcacctg	aacccccctgc	aagaggccgt	4140

	ccagcccgag	ggggatcgcg	actttgccga	tctgggcagc	aagatcgcg	ccatcgcccg	4200
	cgacgttccc	gtgcccgtcc	tgctgaagga	ggtgggctgc	ggcctgtcgg	cggccgatat	4260
	cgccatcggg	ctgcgcgccg	ggatccggca	tttcgacgtg	gccggtcgcg	gcggcacatc	4320
	ctggagccgg	atcgagtatc	gccgcgcgca	gcggggccgat	gacgacctgg	gcctgggtctt	4380
5	ccaggactgg	ggcctgcaga	ccgtggacgc	cctgcgcgag	gcgcggcccg	cgcttgccggc	4440
	ccatgatgga	accagcgtgc	tgatcgccag	cggcggcatc	cgcaacggtg	tcgacatggc	4500
	gaaatgcgtc	atcctggggg	ccgacatgtg	cggggtcgcc	gcgcccctgc	tgaaagcgggc	4560
	ccaaaactcg	cgcgaggcg	ttgtatccgc	catccggaaa	ctgcatctgg	agttccggac	4620
	agccatgttc	ctcctgggtt	gcggcacgct	tgccgacctg	aaggacaatt	cctcgcttat	4680
10	ccgtcaatga	aagtgcctaa	gatgaccgtg	acaggaatcg	aagcgatcag	cttctacacc	4740
	ccccagaact	acgtgggact	ggatatectt	gccgcgcac	acgggatcga	ccccgagaag	4800
	ttctcgaagg	ggatcgggca	ggagaaaaat	gcactgccc	gcatgacga	ggatatcggtg	4860
	accatggccg	ccgagggcgc	gctgcccgtc	atcgaacgcg	cgggcacgca	gggcatcgac	4920
	acggttctgt	tcgccaccga	gagcgggac	gaccagtcga	aggccgccc	catctatctg	4980
15	cgccgcctgc	tggacctgtc	gcccactgc	cgttgctgc	agctgaagca	ggcctgctat	5040
	tccgcgacgg	cggcgctgca	gatggcctgc	gcgcatgtcg	cccgcaagcc	cgaccgcaag	5100
	gtgctggtga	tcgcgtccga	tgctgcgcgc	tatgaccgcg	aaagctcggg	cgagggcgacg	5160
	caggggtgcg	gcgcgcgtgc	catccttgct	agcgcgcgac	ccaaggtggc	cgagatcggc	5220
	accgtctcgg	ggctgttcac	cgaggatata	atggatttct	ggcggccgaa	ccaccgccc	5280
20	acgcccctgt	tcgacggcaa	ggcatcgacg	ctgcgctatc	tgaacgcgct	ggtcgaggcg	5340
	tggaacgact	atcgcgcgaa	tgccggccac	gagttcgccg	atttcgcgca	tttctgctat	5400
	cacgtgcccgt	tctcgccgat	gggcgagaag	gcgaacagcc	acctggccaa	ggcgacaag	5460
	acgcccgttg	acatggggca	ggtgcagacg	ggcctgatct	acaaccggca	ggtcgggaac	5520
	tgctataccg	ggtcgatcta	cctggcattc	gcctcgctgc	tggagaacgc	tcaggaggac	5580
25	ctgaccggcg	cgctggtcgg	tctgttcagc	tatggctcgg	gtgcgacggg	cgaattcttc	5640
	gatgcgcgga	tcgcgcccgg	ttaccgcgac	cacctgttcg	cggaacgcca	tcgcgaattg	5700
	ctgcaggatc	gcacgcccgt	cacatatgac	gaatacgttg	ccctgtggga	cgagatcgac	5760
	ctgacgcagg	gcgcgcccga	caaggcgcgc	ggtcgtttca	ggctggcagg	tatcgaggac	5820
30	gagaagcgca	tct atg tcg acc ggc agg cct gaa gca ggc gcc cat gcc					5869
		Met Ser Thr Gly Arg Pro Glu Ala Gly Ala His Ala					
		1 5 10					
	ccg ggc aag ctg atc ctg tcc ggg gaa cat tcc gtg ctc tat ggt gcg						5917
35	Pro Gly Lys Leu Ile Leu Ser Gly Glu His Ser Val Leu Tyr Gly Ala						
		15 20 25					
	ccc gcg ctt gcc atg gcc atc gcc cgc tat acc gag gtg tgg ttc acg						5965
40	Pro Ala Leu Ala Met Ala Ile Ala Arg Tyr Thr Glu Val Trp Phe Thr						
		30 35 40					
	ccg ctt ggc att ggc gag ggg ata cgc acg aca ttc gcc aat ctc tcg						6013
45	Pro Leu Gly Ile Gly Glu Gly Ile Arg Thr Thr Phe Ala Asn Leu Ser						
		45 50 55 60					
	ggc ggg gcg acc tat tcg ctg aag ctg ctg tcg ggg ttc aag tcg cgg						6061
50	Gly Gly Ala Thr Tyr Ser Leu Lys Leu Leu Ser Gly Phe Lys Ser Arg						
		65 70 75					
	ctg gac cgc cgg ttc gag cag ttc ctg aac ggc gac cta aag gtg cac						6109
55	Leu Asp Arg Arg Phe Glu Gln Phe Leu Asn Gly Asp Leu Lys Val His						
		80 85 90					
	aag gtc ctg acc cat ccc gac gat ctg gcg gtc tat gcg ctg gcg tcg						6157
60	Lys Val Leu Thr His Pro Asp Asp Leu Ala Val Tyr Ala Leu Ala Ser						
		95 100 105					
	ctt ctg cac gac aag ccg ccg ggg acc gcc gcg atg ccg ggc atc ggc						6205
	Leu Leu His Asp Lys Pro Pro Gly Thr Ala Ala Met Pro Gly Ile Gly						
		110 115 120					

	gcg atg cac cac ctg ccg cga ccg ggt gag ctg ggc agc cgg acg gag	6253
	Ala Met His His Leu Pro Arg Pro Gly Glu Leu Gly Ser Arg Thr Glu	
	125 130 135 140	
5	ctg ccc atc ggc gcg ggc atg ggg tcg tct gcg gcc atc gtc gcg gcc	6301
	Leu Pro Ile Gly Ala Gly Met Gly Ser Ser Ala Ala Ile Val Ala Ala	
	145 150 155	
10	acc acg gtc ctg ttc gag acg ctg ctg gac cgg ccc aag acg ccc gaa	6349
	Thr Thr Val Leu Phe Glu Thr Leu Asp Arg Pro Lys Thr Pro Glu	
	160 165 170	
15	cag cgc ttc gac cgc gtc cgc ttc tgc gag cgg ttg aag cac ggc aag	6397
	Gln Arg Phe Asp Arg Val Arg Phe Cys Glu Arg Leu Lys His Gly Lys	
	175 180 185	
20	gcc ggt ccc atc gac gcg gcc agc gtc gtg cgc gcc ggg ctt gtc cgc	6445
	Ala Gly Pro Ile Asp Ala Ala Ser Val Val Arg Gly Gly Leu Val Arg	
	190 195 200	
25	gtg ggc ggg aac ggg ccg ggt tcg atc agc agc ttc gat ttg ccc gag	6493
	Val Gly Gly Asn Gly Pro Gly Ser Ile Ser Ser Phe Asp Leu Pro Glu	
	205 210 215 220	
30	gat cac gac ctt gtc gcg gga cgc ggc tgg tac tgg gta ctg cac ggg	6541
	Asp His Asp Leu Val Ala Gly Arg Gly Trp Tyr Trp Val Leu His Gly	
	225 230 235	
35	cgc ccc gtc agc ggg acc ggc gaa tgc gtc agc gcg gtc gcg gcg gcg	6589
	Arg Pro Val Ser Gly Thr Gly Glu Cys Val Ser Ala Val Ala Ala Ala	
	240 245 250	
40	cat ggt cgc gat gcg gcg ctg tgg gac gcc ttc gca gtc tgc acc cgc	6637
	His Gly Arg Asp Ala Ala Leu Trp Asp Ala Phe Ala Val Cys Thr Arg	
	255 260 265	
45	gcg ttg gag gcc gcg ctg ctg tct ggg ggc agc ccc gac gcc gcc atc	6685
	Ala Leu Glu Ala Ala Leu Leu Ser Gly Gly Ser Pro Asp Ala Ala Ile	
	270 275 280	
50	acc gag aac cag cgc ctg ctg gaa cgc atc ggc gtc gtg ccg gca gcg	6733
	Thr Glu Asn Gln Arg Leu Leu Glu Arg Ile Gly Val Val Pro Ala Ala	
	285 290 295 300	
55	acg cag gcc ctc gtg gcc cag atc gag gag gcg ggt ggc gcg gcc aag	6781
	Thr Gln Ala Leu Val Ala Gln Ile Glu Glu Ala Gly Gly Ala Ala Lys	
	305 310 315	
60	atc tgc ggc gca ggt tcc gtg cgg ggc gat cac ggc ggg gcg gtc ctc	6829
	Ile Cys Gly Ala Gly Ser Val Arg Gly Asp His Gly Gly Ala Val Leu	
	320 325 330	
65	gtg cgg att gac gac gcg cag gcg atg gct tcg gtc atg gcg cgc cat	6877
	Val Arg Ile Asp Asp Ala Gln Ala Met Ala Ser Val Met Ala Arg His	
	335 340 345	
70	ccc gac ctc gac tgg gcg ccc ctg cgc atg tcg cgc acg ggg gcg gca	6925
	Pro Asp Leu Asp Trp Ala Pro Leu Arg Met Ser Arg Thr Gly Ala Ala	
	350 355 360	

	ccc ggc ccc gcg ccg cgt gcg caa ccg ctg ccg ggg cag ggc tga	6970
	Pro Gly Pro Ala Pro Arg Ala Gln Pro Leu Pro Gly Gln Gly	
	365 370 375	
5	tggatcaggt catccgcgcc agcgcgccgg gttcgggtcat gatcacgggc gaacatgccg	7030
	tgggtctatgg acaccgcgcc atcgtcgccg ggatcgagca gcgcgccc atgtacgatcg	7090
	tcccgcgtgc cgaccgcatg tttcgcatca cctcgagat cggggcgccg cagcaggggt	7150
	cgctggacga tctgcctgcg ggccggacat atcgcttcgt gctggccgcc atcgcgcgac	7210
10	acgcgcggga cctgccttgc gggttcgaca tggacatcac ctccgggatc gatccgaggc	7270
	tcgggcttgg atcctcgccg gcggtgacgg tcgcctgcct cggcgcgctg tcgcggctgg	7330
	cggggcgggg gaccgagggg ctgcatgacg acgcgtgcg catcgtcgcg gccatccagg	7390
	gcaggggacg cggggcgcat ctggcgccca gcctgcatgg cggcttcgtc gcctatcgcg	7450
	cgcccgatgg cgggtgcgcg cagatcgagg cgcttcgggt gccgcggggg ccgttcggcc	7510
15	tgcgctatgc gggctacaag accccgacag ccgaggtgct gcgccttg tgccgatcga	7570
	tggcgggcaa cgaggccgt ttcgacgcgc tctactccc gatggcgca agcgagatg	7630
	ccgcgatccg cgcggcgcaa gggctggact gggctgcatt ccacgacgcg ctgaacgaat	7690
	accagcgcct gatggagcag ctgggcgtgt ccgacgacac gctggacgcg atcatccgcg	7750
	aggcgcgca cgccggcgcc gcagtcgcca agatctccgg ctccgggctg ggggattgag	7810
20	tgctggcact gggcgaccag cccaaggggt tcgtgcccgc aagcattgcc gagaagggac	7870
	ttgttttcga tgactgatgc cgtccgcgac atgatcgccc gtgccatggc gggcgcgacc	7930
	gacatccgag cagccgaggc ttatgcgcc agcaacatcg cgtgtcgaa atactggggc	7990
	aagcgcgacg ccgcgcggaa ccttcgcgtg aacagctccg tctcgatctc gttggcgaa	8050
	tgggctctc atacgcgggt cgaggggtcc ggcacggggc acgacgaggt gcatcacaac	8110
25	ggcacgctgc tggatccggg cgacgccttc gcgcgccgcg cgttggcatt cgctgacctg	8170
	ttccgggggg ggaggcacct gccgctgcg atcacgacgc agaactcgat ccgacggcg	8230
	gcggggcttg cctcgtcgcc ctccgggttc gcggcgctga cccgtgcgct ggcggggggc	8290
	ttcgggctgg atctggacga cacggatctg agccgcatcg cccggatcgg cagtggcagc	8350
	gccgcccgtc cgatctggca cggcttcgtc cgtggaaacc gggcgagggc cgaggatggg	8410
30	catgacagcc acggcgtccc gctggacctg cgctggcccg gcttcgcgat cgcgatcgtg	8470
	gccgtggaca aggggcccga gcccttcagt tcgcgcgacg gcatgaacca caggtcgag	8530
	accagcccgc tgttcccgcg ctggcctgcg caggcggaag cggattgccg cgtcatcgag	8590
	gatgcgatcg ccgcccgcga catggccgcg ctgggtccgc gggtcgaggg gaacgccctt	8650
	gcgatgcacg ccacgatgat ggccgcgcgc ccgccgctct gctacctgac gggcggcagc	8710
35	tggcaggtgc tggaaacgct gtggcaggcc gcgcgggacg ggcttgcggc ctttgcgacg	8770
	atggatgccg gcccgaacgt caagctgatc ttcgaggaaa gcagcgccgc cgacgtgctg	8830
	tacctgttcc ccgacgccag cctgatcgcg ccgttcgagg ggcgttgaaac gcgtaagacg	8890
	accactgggt aaggttctgc cgcgcgtggt ctcgactgcc tgcaaagagg tgcttgagtt	8950
	gctgcgtgac tgcggcgcc gacttcgtgg gacttgcggc ccacgctgac gcgctggaaa	9010
40	cgcgcccgcg gattacgacc gcgtcattgc cctgaaccaa tttcccgtcg gtcgac	9066
	<210> 49	
	<211> 378	
	<212> PRT	
45	<213> 副球菌属菌种 (Paracoccus sp.)	
	<400> 49	
	Met Ser Thr Gly Arg Pro Glu Ala Gly Ala His Ala Pro Gly Lys Leu	
	1 5 10 15	
50	Ile Leu Ser Gly Glu His Ser Val Leu Tyr Gly Ala Pro Ala Leu Ala	
	20 25 30	
	Met Ala Ile Ala Arg Tyr Thr Glu Val Trp Phe Thr Pro Leu Gly Ile	
	35 40 45	
55	Gly Glu Gly Ile Arg Thr Thr Phe Ala Asn Leu Ser Gly Gly Ala Thr	
	50 55 60	
60	Tyr Ser Leu Lys Leu Leu Ser Gly Phe Lys Ser Arg Leu Asp Arg Arg	
	65 70 75 80	

	Phe	Glu	Gln	Phe	Leu	Asn	Gly	Asp	Leu	Lys	Val	His	Lys	Val	Leu	Thr
					85					90					95	
5	His	Pro	Asp	Asp	Leu	Ala	Val	Tyr	Ala	Leu	Ala	Ser	Leu	Leu	His	Asp
				100					105					110		
	Lys	Pro	Pro	Gly	Thr	Ala	Ala	Met	Pro	Gly	Ile	Gly	Ala	Met	His	His
			115					120					125			
10	Leu	Pro	Arg	Pro	Gly	Glu	Leu	Gly	Ser	Arg	Thr	Glu	Leu	Pro	Ile	Gly
		130					135					140				
	Ala	Gly	Met	Gly	Ser	Ser	Ala	Ala	Ile	Val	Ala	Ala	Thr	Thr	Val	Leu
15	145					150					155					160
	Phe	Glu	Thr	Leu	Leu	Asp	Arg	Pro	Lys	Thr	Pro	Glu	Gln	Arg	Phe	Asp
					165					170					175	
20	Arg	Val	Arg	Phe	Cys	Glu	Arg	Leu	Lys	His	Gly	Lys	Ala	Gly	Pro	Ile
				180					185					190		
	Asp	Ala	Ala	Ser	Val	Val	Arg	Gly	Gly	Leu	Val	Arg	Val	Gly	Gly	Asn
			195					200					205			
25	Gly	Pro	Gly	Ser	Ile	Ser	Ser	Phe	Asp	Leu	Pro	Glu	Asp	His	Asp	Leu
		210					215					220				
	Val	Ala	Gly	Arg	Gly	Trp	Tyr	Trp	Val	Leu	His	Gly	Arg	Pro	Val	Ser
30	225					230					235					240
	Gly	Thr	Gly	Glu	Cys	Val	Ser	Ala	Val	Ala	Ala	Ala	His	Gly	Arg	Asp
					245					250					255	
35	Ala	Ala	Leu	Trp	Asp	Ala	Phe	Ala	Val	Cys	Thr	Arg	Ala	Leu	Glu	Ala
				260					265					270		
	Ala	Leu	Leu	Ser	Gly	Gly	Ser	Pro	Asp	Ala	Ala	Ile	Thr	Glu	Asn	Gln
			275					280					285			
40	Arg	Leu	Leu	Glu	Arg	Ile	Gly	Val	Val	Pro	Ala	Ala	Thr	Gln	Ala	Leu
		290					295					300				
	Val	Ala	Gln	Ile	Glu	Glu	Ala	Gly	Gly	Ala	Ala	Lys	Ile	Cys	Gly	Ala
45	305					310					315					320
	Gly	Ser	Val	Arg	Gly	Asp	His	Gly	Gly	Ala	Val	Leu	Val	Arg	Ile	Asp
					325					330					335	
50	Asp	Ala	Gln	Ala	Met	Ala	Ser	Val	Met	Ala	Arg	His	Pro	Asp	Leu	Asp
				340					345					350		
	Trp	Ala	Pro	Leu	Arg	Met	Ser	Arg	Thr	Gly	Ala	Ala	Pro	Gly	Pro	Ala
			355					360					365			
55	Pro	Arg	Ala	Gln	Pro	Leu	Pro	Gly	Gln	Gly						
			370				375									
	<210>		50													
60	<211>		9066													

	gcaagatcct	gaccocgctg	ctgcgcacca	gcgcccagaa	gatggctcgc	ctgaactacg	3240
	agaagaacta	tgtcgggggt	acgctggcgg	ggctcgtcgc	cagtgcgaac	gcgcatttcg	3300
	ccaacatgct	gctgggcttc	tacctggcga	cggggcagga	cgcgccaac	atcatcgagg	3360
	ccagccagg	cttctgcat	tgcgaggccc	gcggcgagga	tctgtatttc	tcgtgcacgc	3420
5	tgccccacct	catcatgggc	tcggctcggg	ccggcaagg	catccccctg	atcgaggaga	3480
	acctgtcgcg	gatgggctgc	cgccagccgg	gcgaaccgg	cgacaacgcg	cgccgtcttg	3540
	cggcgatctg	cgcgggcgtc	gtgctgtgtg	gtgaattgtc	gctgcttgcg	gcccagacca	3600
	accccgagga	gttggctcgc	acccacatgg	agatggagcg	atgaccgaca	gcaaggatca	3660
	ccatgtcgcg	ggggcgcaag	tggaccatct	gcgtgcattg	gacgacgatg	cggatatcga	3720
10	ccggggcgac	agcggcttcg	accgcatcgc	gctgacccat	cgcgccctgc	ccgaggtgga	3780
	tttcgacgcc	atcgacacgg	cgaccagctt	cctgggcccgt	gaactgtcct	tcccgtctgt	3840
	gatctcgtcc	atgaccggcg	gcaccggcga	ggagatcgag	cgcatcaacc	gcaacctggc	3900
	cgctggtgcc	gaggaggccc	gcgtcgccat	ggcgggtggc	tcgcagcgcg	tgatgttcac	3960
	cgacccctcg	gcgcgggcca	gcttcgacct	gcgcgcccat	gcgcccaccg	tgcgctgct	4020
15	ggccaatatc	ggcgcggtgc	agctgaacat	ggggctgggg	ctgaaggaat	gcctggccgc	4080
	gatcgagggtg	ctgcaggcgg	acggcctgta	tctgcacctg	aacccccctg	aagaggccgt	4140
	ccagcccagag	ggggatcgcg	actttgccga	tctgggcagc	aagatcgcg	ccatcgcccg	4200
	cgacgttccc	gtgcccgtcc	tgtgaagga	ggtgggctgc	ggcctgtcgg	cgcccgatat	4260
	cgccatcggg	ctgcgcgcgc	ggatccggca	tttcgacgtg	gcgggtcgcg	gcggcacatc	4320
20	ctggagccgg	atcgagtatc	gccgcgcgca	gcgggcccgt	gacgacctgg	gcctggctct	4380
	ccaggactgg	ggcctgcaga	ccgtggacgc	cctgcgcgag	gcgcggcccg	cgcttgccgc	4440
	ccatgatgga	accagcgtgc	tgatcgccag	cggcgccatc	cgcaacggtg	tcgacatggc	4500
	gaaatgcgtc	atcctggggg	ccgacatgtg	cggggtcgcc	gcgcccctgc	tgaagcggc	4560
	ccaaaactcg	cgcgaggcgg	ttgtatccgc	catccggaaa	ctgcatctgg	agttccggac	4620
25	agccatgttc	ctcctggggt	gcggcacgct	tgccgacctg	aaggacaatt	cctcgcttat	4680
	ccgtcaatga	aagtgcctaa	gatgaccgtg	acaggaatcg	aagcgatcag	cttctacacc	4740
	ccccagaact	acgtgggact	ggatatcctt	gccgcgcatc	acgggatcga	ccccgagaag	4800
	ttctcgaagg	ggatcgggca	ggagaaaatc	gcaactgccc	gccaatgacga	ggatatcggt	4860
	accatggccg	ccgaggccgc	gctgccgatc	atcgaaacgc	cgggcacgca	gggcatcgac	4920
30	acggttctgt	tcgccaccga	gagcgggatc	gaccagtcga	aggccgccc	catctatctg	4980
	cgccgcctgc	tggacctgtc	gcccactgac	cggttgcgtc	agctgaagca	ggcctgctat	5040
	tcgcgcagcg	cggcgctgca	gatggcctgc	gcgcatgtcg	cccgaagcc	cgacgcgaag	5100
	gtgctggtga	tcgcgtccga	tgtcgcgcgc	tatgaccgcg	aaagctcggg	cgaggcgacg	5160
	cagggtgcgg	gcgcgcgtgc	catccttgtc	agcgcgcgat	ccaagggtgg	cgagatcggc	5220
35	accgtctcgg	ggctgttcac	cgaggatatc	atggatttct	ggcgcccgaa	ccaccgcgc	5280
	acgcccctgt	tcgacggcaa	ggcatcgacg	ctgcgctatc	tgaacgcgct	ggctcgaggcg	5340
	tggaaacgact	atcgcgcgaa	tggcgccac	gagttcgccg	atttcgcgca	tttctgctat	5400
	cacgtgccgt	tctcgcggat	ggcgagagaag	gcgaacagcc	acctggccaa	ggcgaacaag	5460
	acgcccgttg	acatggggca	ggtgcagacg	ggcctgatct	acaaccggca	ggctcggaac	5520
40	tgctataccg	ggtcgatcta	cctggcattc	gcctcgtcgc	tggagaacgc	tcaggaggac	5580
	ctgaccggcg	cgctggctcg	tctgttcagc	tatggctcgg	gtgcgacggg	cgaattcttc	5640
	gatgcgcgga	tcgcgcccg	ttaccgcgac	caactgttgc	cggaacgcca	tcgcgaattg	5700
	ctgcaggatg	gcacgcccgt	cacatatgac	gaatacgttg	ccctgtggga	cgagatcgac	5760
	ctgacgcagg	gcgcgcccga	caaggcgcg	ggctcgtttc	ggctggcagg	tatcgaggac	5820
45	gagaagcgca	tctatgtcga	ccggcaggcc	tgaagcaggc	gcccattgcc	cgggcaagct	5880
	gatcctgtcc	ggggaacatt	ccgtgctcta	tggctgcgcc	gcgcttgcca	tggccatcgc	5940
	ccgctatacc	gaggtgtggt	tcacgcgcgt	tggcattggc	gaggggatac	gcacgacatt	6000
	cgccaatctc	tcgggcgggg	cgacctatcc	gctgaagctg	ctgtcggggg	tcaagtgcgc	6060
	gctggaccgc	cggttcgagc	agttcctgaa	cggcgacctc	aaggtgcaca	aggtcctgac	6120
50	ccatcccagc	gatctggcgg	tctatgcgct	ggcgtcgctt	ctgcacgaca	agccgcggg	6180
	gaccgcgcgc	atgcggggca	tcggcgcgat	gcaccacctg	ccgcgaccgg	gtgagctggg	6240
	cagccggacg	gagctgcccc	tcggcgcggg	catggggctg	tctgcggcca	tcgtcgcggc	6300
	caccacggtc	ctgttcgaga	cgctgctgga	ccggcccaag	acgcccgaac	agcgtctcga	6360
	ccgcgtccgc	ttctgcgagc	gggtgaagca	cggaaggccc	ggctccatcg	acgcggccag	6420
55	cgctcgtcgc	ggcgggcttg	tcgcgctggg	cggaacggg	ccgggttcga	tcagcagctt	6480
	cgatttgccc	gaggatcacg	accttgctgc	gggacgcggc	tggctactgg	tactgcacgg	6540
	gcgccccgtc	agcgggaccg	gcgaatgcgt	cagcgcgggc	gcggcgggcg	atggctcgca	6600
	tgcggcgctg	tgggacgcct	tcgcagctcg	caccgcgcgc	ttggaggccg	cgctgctgtc	6660
	tgggggcagc	cccgacgcgc	ccatcaaccga	gaaccagcgc	ctgctggaac	gcacggcggt	6720
60	cgtgcgggca	gcgacgcagg	ccctcgtggc	ccagatcgag	gaggcggggtg	gcgcggccaa	6780

	225	230	235		
5	aac gaa tac cag cgc ctg atg gag cag ctg ggc gtg tcc gac gac acg Asn Glu Tyr Gln Arg Leu Met Glu Gln Leu Gly Val Ser Asp Asp Thr 240 245 250	7731			
10	ctg gac gcg atc atc cgc gag gcg cgc gac gcg ggc gcc gca gtc gcc Leu Asp Ala Ile Ile Arg Glu Ala Arg Asp Ala Gly Ala Ala Val Ala 255 260 265 270	7779			
15	aag atc tcc ggc tcg ggg ctg ggg gat tgc gtg ctg gca ctg ggc gac Lys Ile Ser Gly Ser Gly Leu Gly Asp Cys Val Leu Ala Leu Gly Asp 275 280 285	7827			
20	cag ccc aag ggt ttc gtg ccc gca agc att gcc gag aag gga ctt gtt Gln Pro Lys Gly Phe Val Pro Ala Ser Ile Ala Glu Lys Gly Leu Val 290 295 300	7875			
25	ttc gat gac tga tgccgtccgc gacatgatcg cccgtgccat ggcggggcgcg Phe Asp Asp 305	7927			
30	accgacatcc gagcagccga ggcttatgcyg cccagcaaca tcgcgctgtc gaaatactgg ggcaagcgcg acgcccgcgc gaaccttcgc ctgaacagct ccgtctcgat ctcggttgccg aactggggct ctcatacgcg ggtcgagggg tccggcacgg gccacgacga ggtgcatcac aacggcacgc tgctggatcc gggcgacgcc ttgcgcgcgc gcgcggttggc attcgctgac ctgttccggg gggggaggca cctgccgctg cggatcacga cgcagaactc gatcccgacg gcggcggggc ttgcctcgtc ggctcgggg ttgcgggcgc tgaccogtgc gctggcgggg gcgttcgggc tggatctgga cgacacggat ctgagccgca tcgcccggat cggcagtggc agcgccgccc gctcgatctg gcacggcttc gtccgctgga accggggcga ggccgaggat gggcatgaca gccacggcgt cccgctggac ctgcgctggc ccggcttcgc catcgcgatc gtggccgtgg acaaggggcc caagccttcc agttcgcgcg acggcatgaa ccacacggtc gagaccagcc cgctgttccc gccctggcct gcgcaggcgg aagcggattg ccgcgtcatc gaggatgcga tcgcccgcgc cgacatggcc gccctgggtc cgcgggtcga ggccaacgcc cttgcgatgc acgccacgat gatggccgcg cgcgcgcgc tctgctacct gacgggcggc agctggcagg tgctggaacg cctgtggcag gcccgcgcg acgggcttgc ggcctttgcyg acgatggatg ccggcccga cgtcaagctg atcttcgagg aaagcagcgc cgccgacgtg ctgtacctgt tccccgacgc cagcctgac gcgcgcttcg aggggcgttg aacgcgtaag acgaccactg ggtaagggtc tgccgcgcgt ggtctcgact gcctgcaaag aggtgcttga gttgctgcgt gactgcggcg gccgacttcg tgggaattgc ccgccacgct gacgcgctgg aaacgcgccc gcggattacg accgcgtcat tgccctgaac caatttccc tcggtcgac	7987 8047 8107 8167 8227 8287 8347 8407 8467 8527 8587 8647 8707 8767 8827 8887 8947 9007 9066			
45	<210> 51 <211> 305 <212> PRT <213> 副球菌属菌种(Paracoccus sp.) <400> 51 Met Asp Gln Val Ile Arg Ala Ser Ala Pro Gly Ser Val Met Ile Thr 1 5 10 15 Gly Glu His Ala Val Val Tyr Gly His Arg Ala Ile Val Ala Gly Ile 20 25 30 Glu Gln Arg Ala His Val Thr Ile Val Pro Arg Ala Asp Arg Met Phe 35 40 45 Arg Ile Thr Ser Gln Ile Gly Ala Pro Gln Gln Gly Ser Leu Asp Asp 50 55 60 Leu Pro Ala Gly Gly Thr Tyr Arg Phe Val Leu Ala Ala Ile Ala Arg				

	65				70					75					80	
	His	Ala	Pro	Asp	Leu	Pro	Cys	Gly	Phe	Asp	Met	Asp	Ile	Thr	Ser	Gly
					85					90					95	
5	Ile	Asp	Pro	Arg	Leu	Gly	Leu	Gly	Ser	Ser	Ala	Ala	Val	Thr	Val	Ala
				100					105					110		
	Cys	Leu	Gly	Ala	Leu	Ser	Arg	Leu	Ala	Gly	Arg	Gly	Thr	Glu	Gly	Leu
10			115					120					125			
	His	Asp	Asp	Ala	Leu	Arg	Ile	Val	Arg	Ala	Ile	Gln	Gly	Arg	Gly	Ser
		130					135					140				
15	Gly	Ala	Asp	Leu	Ala	Ala	Ser	Leu	His	Gly	Gly	Phe	Val	Ala	Tyr	Arg
	145					150					155					160
	Ala	Pro	Asp	Gly	Gly	Ala	Ala	Gln	Ile	Glu	Ala	Leu	Pro	Val	Pro	Pro
					165					170					175	
20	Gly	Pro	Phe	Gly	Leu	Arg	Tyr	Ala	Gly	Tyr	Lys	Thr	Pro	Thr	Ala	Glu
				180					185					190		
	Val	Leu	Arg	Leu	Val	Ala	Asp	Arg	Met	Ala	Gly	Asn	Glu	Ala	Ala	Phe
25			195					200					205			
	Asp	Ala	Leu	Tyr	Ser	Arg	Met	Gly	Ala	Ser	Ala	Asp	Ala	Ala	Ile	Arg
		210					215					220				
30	Ala	Ala	Gln	Gly	Leu	Asp	Trp	Ala	Ala	Phe	His	Asp	Ala	Leu	Asn	Glu
	225					230					235					240
	Tyr	Gln	Arg	Leu	Met	Glu	Gln	Leu	Gly	Val	Ser	Asp	Asp	Thr	Leu	Asp
					245					250					255	
35	Ala	Ile	Ile	Arg	Glu	Ala	Arg	Asp	Ala	Gly	Ala	Ala	Val	Ala	Lys	Ile
				260					265					270		
	Ser	Gly	Ser	Gly	Leu	Gly	Asp	Cys	Val	Leu	Ala	Leu	Gly	Asp	Gln	Pro
40			275					280					285			
	Lys	Gly	Phe	Val	Pro	Ala	Ser	Ile	Ala	Glu	Lys	Gly	Leu	Val	Phe	Asp
		290					295					300				
45	Asp															
	305															
	<210>	52														
	<211>	9066														
50	<212>	DNA														
	<213>	副球菌属菌种 (Paracoccus sp.)														
	<220>															
	<221>	CDS														
	<222>	(7880)..(8878)														
55	<223>															
	<400>	52														
	ggatccggca	gctcgacacg	ccgcagaacc	tgtacgaacg	tcccgccagc	cgcttcgctcg										60
	cggaattcgt	cgggcgcggg	acggtgggtgc	ccgtgcaggc	ccatgacggc	gcggggccgcg										120
	ccgcacatcct	gggggccgag	gtggcgggtga	acgcgcgcgc	gcaatcgcgc	tttgtcgatc										180
60	acgtctgcct	gcgccccgag	aaccttgcca	tctccgagac	gggcgacctg	cgcgcccaagg										240

	tgcgcgcgct	cacctatctt	ggcgggaaat	acctgctgga	aaccgtgctg	gattgcggca	300
	cccggctggg	gaccgagacc	cgcgcccgct	tcgatacggg	cgcgagctt	ggcctgacca	360
	tcaacgcccc	ctgggccttt	gccgaggatt	gaatggacag	cgtgaagatc	ctttcgggca	420
	tgggcgtgaa	gggcccgtgc	tgcatacagg	tggatgtcgg	cgggatgcgc	ctgatcctcg	480
5	attgcggggac	cgggcccgac	gagggcgcg	agttcgaccc	cgccctggctg	gcggacgcgg	540
	atgcgggtgct	gatcacccat	gaccacgtgg	accataatcg	cggcgcgcgt	cacgcggtcg	600
	cggcgggggct	gccgatccat	gcgacgcggc	agacggcggg	gttgctgccc	gcgggggctg	660
	atctgcgcct	gctgcccga	cgcggtgtca	cgcgatcgcc	cggggtcgat	ctgacgaccg	720
	gtcgcaacgg	gcatgccgcg	ggcggcgtct	ggatgcattt	cgacatgggc	gaggggctgt	780
10	tctattccgg	cgactggtcc	gaggaatccg	actggttcgc	cttcgatccg	ccccgcctg	840
	cggggacggc	gattctcgac	tgctcctatg	gcggtttcga	cgtggcgcaa	tcggattgca	900
	tcgcggacct	ggacgacctg	ctcgaggtgc	tgcgggggca	ggtagctgctg	ccggtgccgc	960
	catccggccg	cgcgcccgag	ctggccctgc	ggctgatccg	ccgccacgga	ccgggcagcg	1020
	tgatggtcga	cgacgcctgc	ctgccggcca	tcgcgcaact	gcccagggcg	cgcgactggg	1080
15	cctacgccac	cgaggcacgc	tttcttgtct	gcgacacgcc	gaacgccgaa	agccggcgcg	1140
	gcatggcggc	atctgcaagc	atggcgcgat	gcgggcaggc	tggggcgggg	cgcgcatgtc	1200
	gtcttcaccg	ggcacatgaa	cgtccatgog	cgcgatttct	gcgaccgccc	cggcgggcat	1260
	ttccgcgcgt	ggaacgtgca	tccgcgcgtg	cgcgaccagc	gacggatgct	ggaacggctg	1320
	gccgcgcggc	gctttgcccc	ggcctttctg	ccgaccccg	agatctatct	ggcgtggac	1380
20	atgggcgcgc	aggtcttcat	gcaccaggag	gtgacgccat	gatccccgcc	cgcagcttct	1440
	gcctgatccg	ccacggcgaa	acgaccgcca	atgcaggggc	gatcatcgcg	ggcgcaaccg	1500
	atgtgccccct	gacgccaaag	ggccgcgcgc	agggccgcgc	cctggcaggg	cgcaatggc	1560
	catcgggcat	cgcgctgttc	gccagcccga	tgtcgcgtgc	ccgcgatacc	gcgctgctgg	1620
	cctttccggg	gcgcgaccac	cagcccgaac	ccgatctgcg	cgaacgcgac	tggggcatct	1680
25	tcgagggaac	ccccgtcgcc	gatctgcccc	cgcgcgaaat	cacgccgcag	gggggcgagg	1740
	gctgggacga	cgtgatggcc	cgcgtaggac	gcgcgatccg	gcggatctgc	gcgacctcgg	1800
	gcgatgcgct	gcgggtgctg	gtctgccatt	cgggcggtgat	ccgtgccgcg	cgcgctgctgt	1860
	ggaccaccgg	cgatgcgggc	gatcgtccgc	ccaacgccac	gccgatccctg	ttcagcccgg	1920
	acggcgaccg	attaaaggaa	ggaacgatat	gaccgccacc	acccccctgcg	tcgtcttcga	1980
30	acgtggacgg	cacgcttgcc	gaattcgacg	ccgaccgcct	gggccatctt	gtccacggca	2040
	cgaccaagca	ctgggacgcc	ttccaccacg	cgatggccga	cgccccgcc	atccccgagg	2100
	tcgcccgcct	gatgcgcaag	ctgaaggagg	ggggcgagac	ggtcgtcatc	tgctcggggc	2160
	ggccccgcgg	ctggcaggat	cagacgatcg	catggtcgcg	caagcacgac	ctgcccttcg	2220
	acgggatcta	tctgcgcccc	gaggatcagg	acggcgccag	cgaccccgag	gtcaagcgcc	2280
35	gcgccttagc	cgagatgcgc	gccgacgggc	tggcgccctg	gctggctcgtg	gacgaccggc	2340
	ggtcgcgtgt	ggatgcctgg	cgggccgagg	ggctggctctg	cctgcaatgc	gcgccggggg	2400
	acttctaggg	ccgcgcgacg	ggggcgcgga	caggctgggc	gggaaaccgc	cccgccacca	2460
	tgtcctgcac	gcgtogaacc	gcccgtccga	cgccggtttc	cgcacggaaa	cgcgcggcaa	2520
	gttgacataa	cttgcaacgcg	acgtctcgat	tctgcccgcg	aagaatgcga	tgcatccaga	2580
40	tgatgcagaa	cgaagaagcg	gaagcgcccg	tgaaagacca	gatgatttcc	catacccccg	2640
	tgcccacgca	atgggtcggc	ccgatcctgt	tccgcggccc	cgtcgtcgag	ggcccgatca	2700
	gcgcgcgcgt	ggccacctac	gagacgcgcg	tctggccctc	gaccgcgcgcg	ggggcagggg	2760
	tttcccggca	ttcgggcggg	atccaggtct	cgctgggtcga	cgaacgcgatg	agccgctcga	2820
	tcgcgctgcg	ggcgcatgac	ggggcgggcg	cgaccgcgcg	ctggcagtcg	atcaaggccc	2880
45	gccaggaaga	ggtcgcggcc	gtggtcgcca	ccaccagccg	cttcgcccgc	cttgtcgagc	2940
	tgaatcgcca	gatcgtgggc	aacctgcttt	acatccgcgt	cgaatgcgtg	acgggcgacg	3000
	cctcggggtca	caacatggtc	accaaggccg	ccgaggccgt	gcagggtctg	atcctgtcgg	3060
	aatacccgat	gctggcctat	tccacgatct	cggggaacct	gtgcaccgac	aagaaggcgt	3120
	cggcgggtcaa	cggcatecctg	ggccgcggca	aatacgccgt	cgccgaggtc	gagatccccg	3180
50	gcaagatcct	gaccgcgcgtg	ctgcgcacca	gcgcgagaa	gatggtccgc	ctgaactacg	3240
	agaagaacta	tgtcgggggg	acgctggcgg	ggtcgtcgcg	cagtgcgaac	gcgcatttcg	3300
	ccaacatgct	gctgggcttc	tacctggcga	cggggcagga	cgcgcccaac	atcatcgagg	3360
	ccagccaggg	cttcgtccat	tgcgaggccc	gcggcgagga	tctgtatttc	tcgtgcacgc	3420
	tgcccacact	catcatgggc	tcggtcggtg	ccggcaaggg	catccccctg	atcgaggaga	3480
55	acctgtcgcg	gatgggctgc	cgccagccgg	gcgaaccgcg	cgacaacgcg	cgccgtcttg	3540
	cggcgatctg	cgcgggcgctc	gtgctgtgtg	gtgaattgtc	gctgcttgcg	gcccagacca	3600
	accccgagga	gttggtccgc	accacatgg	agatggagcg	atgaccgaca	gcaaggatca	3660
	ccatgtcgcg	gggcgcaagc	tggaccatct	gcgtgcattg	gacgacgatg	cggatatcga	3720
	ccggggcgac	agcggcttcg	accgcacgcg	gctgacccat	cgcgccctgc	ccgaggtgga	3780
60	tttcgacgcc	atcgacacgg	cgaccagctt	cctgggcccgt	gaactgtcct	tcccgcgtgct	3840

	gatctcgtcc	atgaccggog	gcaccggcga	ggagatcgag	cgcacacacc	gcaacctggc	3900
	cgctggtgcc	gaggaggccc	gcgtcgccat	ggcgggtggc	tcgcagcgcg	tgatgttcac	3960
	cgacccctcg	gcgcgggcca	gcttcgacct	gcgcgcccat	gcgcccaccg	tgccgctgct	4020
	ggccaatata	ggcgcggtgc	agctgaacat	ggggctgggg	ctgaaggaat	gcctggccgc	4080
5	gatcgaggtg	ctgcaggcgg	acggcctgta	tctgcacctg	aacccccctg	aagaggccgt	4140
	ccagcccgag	ggggatcgcg	actttgccga	tctgggcagc	aagatcgcg	ccatcgccc	4200
	cgacgttccc	gtgcccgctc	tgtgaagga	ggtgggctgc	ggcctgtcgg	cgcccgatat	4260
	cgccatcggg	ctgcgcgcgg	ggatccggca	tttcgacgtg	gccggtcgcg	gcggcacatc	4320
	ctggagccgg	atcgagtata	gccgccgcca	gcgggcccga	gacgacctgg	gcctggtcct	4380
10	ccaggactgg	ggcctgcaga	ccgtggacgc	cctgcgcgag	gcgcggccc	cgcttgcggc	4440
	ccatgatgga	accagcgtgc	tgatcgccag	cgcgccatc	cgcaacggtg	tcgacatggc	4500
	gaaatgcgtc	atcctggggg	ccgacatgtg	cggggtcgcc	gcgcccctgc	tgaaagcggc	4560
	ccaaaactcg	cgcgaggcgg	ttgtatccgc	catccggaaa	ctgcatctgg	agttccggac	4620
	agccatgttc	ctcctggggt	gcggcacgct	tgccgacctg	aaggacaatt	cctcgcttat	4680
15	ccgtcaatga	aagtgcctaa	gatgaccgtg	acaggaatcg	aagcgatcag	cttctacacc	4740
	ccccagaact	acgtgggact	ggatatcctt	gccgcgcac	acgggatcga	ccccgagaag	4800
	ttctcgaagg	ggatcgggca	ggagaaaatc	gcactgccc	gccatgacga	ggatatcggt	4860
	accatggccg	ccgaggccgc	gctgccgatc	atcgaaacgc	cgggcacgca	gggcatcgac	4920
	acggttctgt	tcgccaccga	gagcgggatc	gaccagtcca	aggccgccc	catctatctg	4980
20	cgcgcgctgc	tggacctgtc	gcccactgc	cgctgcgtcg	agctgaagca	ggcctgctat	5040
	tcgcgcagcg	cggcgctgca	gatggcctgc	gttgatgtcg	cccgaagcc	cgaccgaatg	5100
	gtgctggtga	tcgcgtccga	tgtcgcgcgc	tatgaccgcg	aaagctcggg	cgaggcgacg	5160
	cagggtgcgg	gcgcgcgtgc	catccttgtc	agcgcgcgac	ccaaggtggc	cgagatcggc	5220
	accgtctcgg	ggctgttcac	cgaggatata	atggatttct	ggcgcccgaa	ccaccgccc	5280
25	acgcccctgt	tcgacggcaa	ggcatcgacg	ctgcgctatc	tgaacgcgct	ggtcgaggcg	5340
	tggaacgact	atcgcgcgaa	tggcgccac	gagttcgccg	atttcgcgca	tttctgctat	5400
	cacgtgccgt	tctcgcggat	ggcgagaga	gcgaacagcc	acctggccaa	ggcgaacaag	5460
	acgcgggtgg	acatggggca	ggtgcagacg	ggcctgatct	acaaccggca	ggtcgggaa	5520
	tgctataacc	ggtcgatcta	cctggcattc	gcctcgctgc	tggagaacgc	tcaggaggac	5580
30	ctgaccggcg	cgctggtcgg	tctgttcagc	tatggtctcg	gtgcgacggg	cgaattcttc	5640
	gatgcgcgga	tcgcgcccg	ttaccgcgac	cacctgttcg	cggaacgcca	tcgcgaattg	5700
	ctgcaggatc	gcacgcccgt	cacatatgac	gaatacgttg	ccctgtggga	cgagatcgac	5760
	ctgacgcagg	gcgcgcccga	caaggcgcg	ggtcgtttca	ggctggcagg	tatcgaggac	5820
	gagaagcgca	tctatgtcga	ccggcaggcc	tgaagcaggc	gcccattgcc	cgggcaagct	5880
35	gatcctgtcc	ggggaacatt	ccgtgctcta	tgggtgcgcc	gcgcttgcca	tgcccatcgc	5940
	ccgctataacc	gaggtgtggt	tcacgcgcgt	tggcattggc	gaggggatac	gcacgacatt	6000
	cgccaatctc	tcgggcgggg	cgacctattc	gctgaagctg	ctgtcggggg	tcaagtcgcg	6060
	gctggaccgc	cggttcgagc	agttcctgaa	cggcgacctc	aaggtgcaca	aggtcctgac	6120
	ccatcccgcg	gatctggcgg	tctatgcgct	ggcgtcgctt	ctgcacgaca	agccgcccgg	6180
40	gaccgcccgc	atgccgggca	tcggcgcgat	gcaccacctg	ccgcgaccgg	gtgagctggg	6240
	cagccggacg	gagctgccca	tcggcgcggg	catggggtcg	tctgcggcca	tcgtcgcggc	6300
	caccacggtc	ctgttcgaga	cgctgctgga	ccggcccaag	acgcccgaac	agcgtttcga	6360
	ccgcgtccgc	ttctgcgagc	ggttgaagca	cggcaaggcc	ggtcccattc	acgcggccag	6420
	cgctcgtcgc	ggcgggcttg	tcgcgctggg	cgggaacggg	ccgggttcga	tcagcagctt	6480
45	cgatttgccc	gaggatcacg	accttgtcgc	gggacgcggc	tgggtactgg	tactgcacgg	6540
	gcgccccgct	agcgggaccg	gcgaatgcgt	cagcgcggtc	gcggcgggcg	atggctcgca	6600
	tgccgcgctg	tgggacgcct	tcgcagctcg	cacccgcgcg	ttggaggccg	cgctgctgtc	6660
	tgggggcagc	cccagcgccg	ccatcaccga	gaaccagcgc	ctgctggaac	gcacggtcgt	6720
	cgtgccggca	gcgacgcagg	ccctcgtggc	ccagatcgag	gaggcggggt	gcgcggccaa	6780
50	gatctgcggc	gcaggttcgg	tcggggcgga	tcacggcggg	gcggtcctcg	tcggtattga	6840
	cgacgcgcag	gcgatggctt	cggtcatggc	gcgccatccc	gacctcgact	gggcgcccct	6900
	gcgcattgtc	gcacggggg	cggcaccggg	ccccgcggcg	cgtgcgcaac	cgctgcgggg	6960
	gcagggttga	tggatcaggt	catccgcgcc	agcgcgcggg	gttcggtcat	gatcacgggc	7020
	gaacatgccg	tggctctatg	acaccgcgcc	atcgtcgccg	ggatcgagca	gcgcgcccac	7080
55	gtgacgatcg	tcccgcgtgc	cgaccgcatg	tttcgcatca	cctcgagat	cggggcggcg	7140
	cagcaggggt	cgctggacga	tctgcctgcg	ggcgggacct	atcgcttcgt	gctggccggc	7200
	atcgcgcgac	acgcgcggga	cctgccttgc	gggttcgaca	tggacatcac	ctcggggatc	7260
	gatccgaggc	tcgggcttgg	atcctcgcg	gcggtgacgg	tcgcctgcct	cggcgcgctg	7320
	tcgcggctgg	cggggcgggg	gaccgagggg	ctgcatgacg	acgcgctgcg	catcgctcgc	7380
60	gccatccagg	gcaggggacg	cggggccgat	ctggcgcgca	gcctgcatgg	cggcttcgctc	7440

5	gcctatcgcg	cgcccgatgg	cggtgccgcg	cagatcgagg	cgcttccggt	gccgccgggg	7500
	ccgttcggcc	tgcgctatgc	gggctacaag	accccgacag	ccgaggtgct	gcgccttggtg	7560
	gccgatcgga	tggcgggcaa	cgaggccgct	ttcgacgcgc	tctactcccg	gatgggcgca	7620
	agcgacagatg	ccgcgatccg	cgcgggcgcaa	gggctggact	gggctgcatt	ccacgacgcg	7680
	ctgaacgaat	accagcgcct	gatggagcag	ctgggcgtgt	ccgacgacac	gctggacgcg	7740
	atcatccgcg	aggcgcgcg	cgcgggcgcc	gcagtcgcca	agatctccgg	ctcggggctg	7800
	ggggattgcg	tgctggcact	gggcgaccag	cccaagggtt	tcgtgcccgc	aagcattgcc	7860
	gagaaggggac	ttgttttcg	atg act gat	gcc gtc cgc	gac atg atc	gcc cgt	7912
10			Met Thr Asp	Ala Val Arg	Asp Met Ile	Ala Arg	
		1		5		10	
15	gcc atg gcg ggc gcg acc gac atc cga gca gcc gag gct tat gcg ccc						7960
	Ala Met Ala Gly Ala Thr Asp Ile Arg Ala Ala Glu Ala Tyr Ala Pro	15		20		25	
20	agc aac atc gcg ctg tcg aaa tac tgg ggc aag cgc gac gcc gcg cgg						8008
	Ser Asn Ile Ala Leu Ser Lys Tyr Trp Gly Lys Arg Asp Ala Ala Arg	30		35		40	
25	aac ctt ccg ctg aac agc tcc gtc tcg atc tcg ttg gcg aac tgg ggc						8056
	Asn Leu Pro Leu Asn Ser Ser Val Ser Ile Ser Leu Ala Asn Trp Gly	45		50		55	
30	tct cat acg cgg gtc gag ggg tcc ggc acg ggc cac gac gag gtg cat						8104
	Ser His Thr Arg Val Glu Gly Ser Gly Thr Gly His Asp Glu Val His	60		65		70	75
35	cac aac ggc acg ctg ctg gat ccg ggc gac gcc ttc gcg cgc cgc gcg						8152
	His Asn Gly Thr Leu Leu Asp Pro Gly Asp Ala Phe Ala Arg Arg Ala	80		85		90	
40	ttg gca ttc gct gac ctg ttc cgg ggg ggg agg cac ctg ccg ctg cgg						8200
	Leu Ala Phe Ala Asp Leu Phe Arg Gly Gly Arg His Leu Pro Leu Arg	95		100		105	
45	atc acg acg cag aac tcg atc ccg acg gcg gcg ggg ctt gcc tcg tcg						8248
	Ile Thr Thr Gln Asn Ser Ile Pro Thr Ala Ala Gly Leu Ala Ser Ser	110		115		120	
50	gcc tcg ggg ttc gcg gcg ctg acc cgt gcg ctg gcg ggg gcg ttc ggg						8296
	Ala Ser Gly Phe Ala Ala Leu Thr Arg Ala Leu Ala Gly Ala Phe Gly	125		130		135	
55	ctg gat ctg gac gac acg gat ctg agc cgc atc gcc cgg atc ggc agt						8344
	Leu Asp Leu Asp Asp Thr Asp Leu Ser Arg Ile Ala Arg Ile Gly Ser	140		145		150	155
60	ggc agc gcc gcc cgc tcg atc tgg cac ggc ttc gtc cgc tgg aac cgg						8392
	Gly Ser Ala Ala Arg Ser Ile Trp His Gly Phe Val Arg Trp Asn Arg	160		165		170	
65	ggc gag gcc gag gat ggg cat gac agc cac ggc gtc ccg ctg gac ctg						8440
	Gly Glu Ala Glu Asp Gly His Asp Ser His Gly Val Pro Leu Asp Leu	175		180		185	
70	cgc tgg ccc ggc ttc cgc atc gcg atc gtg gcc gtg gac aag ggg ccc						8488
	Arg Trp Pro Gly Phe Arg Ile Ala Ile Val Ala Val Asp Lys Gly Pro	190		195		200	
75	aag cct ttc agt tcg cgc gac ggc atg aac cac acg gtc gag acc agc						8536

	Lys	Pro	Phe	Ser	Ser	Arg	Asp	Gly	Met	Asn	His	Thr	Val	Glu	Thr	Ser	
	205						210					215					
5	ccg	ctg	ttc	ccg	ccc	tgg	cct	gcg	cag	gcg	gaa	gcg	gat	tgc	cgc	gtc	8584
	Pro	Leu	Phe	Pro	Pro	Trp	Pro	Ala	Gln	Ala	Glu	Ala	Asp	Cys	Arg	Val	
	220					225					230					235	
10	atc	gag	gat	gcg	atc	gcc	gcc	cgc	gac	atg	gcc	gcc	ctg	ggt	ccg	cgg	8632
	Ile	Glu	Asp	Ala	Ile	Ala	Ala	Arg	Asp	Met	Ala	Ala	Leu	Gly	Pro	Arg	
					240					245					250		
15	gtc	gag	gcg	aac	gcc	ctt	gcg	atg	cac	gcc	acg	atg	atg	gcc	gcg	cgc	8680
	Val	Glu	Ala	Asn	Ala	Leu	Ala	Met	His	Ala	Thr	Met	Met	Ala	Ala	Arg	
				255					260					265			
20	ccg	ccg	ctc	tgc	tac	ctg	acg	ggc	ggc	agc	tgg	cag	gtg	ctg	gaa	cgc	8728
	Pro	Pro	Leu	Cys	Tyr	Leu	Thr	Gly	Gly	Ser	Trp	Gln	Val	Leu	Glu	Arg	
			270					275					280				
25	ctg	tgg	cag	gcc	cgc	gcg	gac	ggg	ctt	gcg	gcc	ttt	gcg	acg	atg	gat	8776
	Leu	Trp	Gln	Ala	Arg	Ala	Asp	Gly	Leu	Ala	Ala	Phe	Ala	Thr	Met	Asp	
		285				290						295					
30	gcc	ggc	ccg	aac	gtc	aag	ctg	atc	ttc	gag	gaa	agc	agc	gcc	gcc	gac	8824
	Ala	Gly	Pro	Asn	Val	Lys	Leu	Ile	Phe	Glu	Glu	Ser	Ser	Ala	Ala	Asp	
	300					305					310					315	
35	gtg	ctg	tac	ctg	ttc	ccc	gac	gcc	agc	ctg	atc	gcg	ccg	ttc	gag	ggg	8872
	Val	Leu	Tyr	Leu	Phe	Pro	Asp	Ala	Ser	Leu	Ile	Ala	Pro	Phe	Glu	Gly	
					320					325					330		
40	cgt	tga	acgcgtaaga	cgaccactgg	gtaaggttct	gccgcgcgtg	gtctcgactg										8928
	Arg																
45	cctgcaaaga	ggtgcttgag	ttgctgcgtg	actgcggcgg	ccgacttcgt	gggacttgcc											8988
	cgccacgctg	acgcgctgga	aacgcgcccg	cggattacga	ccgcgtcatt	gccctgaacc											9048
	aatttcccgt	cggtcgac															9066
	<210>	53															
	<211>	332															
	<212>	PRT															
	<213>	副球菌属菌种 (Paracoccus sp.)															
	<400>	53															
50	Met	Thr	Asp	Ala	Val	Arg	Asp	Met	Ile	Ala	Arg	Ala	Met	Ala	Gly	Ala	
	1				5					10					15		
55	Thr	Asp	Ile	Arg	Ala	Ala	Glu	Ala	Tyr	Ala	Pro	Ser	Asn	Ile	Ala	Leu	
				20					25					30			
60	Ser	Lys	Tyr	Trp	Gly	Lys	Arg	Asp	Ala	Ala	Arg	Asn	Leu	Pro	Leu	Asn	
			35				40						45				
65	Ser	Ser	Val	Ser	Ile	Ser	Leu	Ala	Asn	Trp	Gly	Ser	His	Thr	Arg	Val	
		50					55				60						
70	Glu	Gly	Ser	Gly	Thr	Gly	His	Asp	Glu	Val	His	His	Asn	Gly	Thr	Leu	
	65					70					75					80	
75	Leu	Asp	Pro	Gly	Asp	Ala	Phe	Ala	Arg	Arg	Ala	Leu	Ala	Phe	Ala	Asp	
					85					90					95		

Leu Phe Arg Gly Gly Arg His Leu Pro Leu Arg Ile Thr Thr Gln Asn
 100 105 110
 5 Ser Ile Pro Thr Ala Ala Gly Leu Ala Ser Ser Ala Ser Gly Phe Ala
 115 120 125
 Ala Leu Thr Arg Ala Leu Ala Gly Ala Phe Gly Leu Asp Leu Asp Asp
 130 135 140
 10 Thr Asp Leu Ser Arg Ile Ala Arg Ile Gly Ser Gly Ser Ala Ala Arg
 145 150 155 160
 Ser Ile Trp His Gly Phe Val Arg Trp Asn Arg Gly Glu Ala Glu Asp
 15 165 170 175
 Gly His Asp Ser His Gly Val Pro Leu Asp Leu Arg Trp Pro Gly Phe
 180 185 190
 20 Arg Ile Ala Ile Val Ala Val Asp Lys Gly Pro Lys Pro Phe Ser Ser
 195 200 205
 Arg Asp Gly Met Asn His Thr Val Glu Thr Ser Pro Leu Phe Pro Pro
 210 215 220
 25 Trp Pro Ala Gln Ala Glu Ala Asp Cys Arg Val Ile Glu Asp Ala Ile
 225 230 235 240
 Ala Ala Arg Asp Met Ala Ala Leu Gly Pro Arg Val Glu Ala Asn Ala
 245 250 255
 30 Leu Ala Met His Ala Thr Met Met Ala Ala Arg Pro Pro Leu Cys Tyr
 260 265 270
 Leu Thr Gly Gly Ser Trp Gln Val Leu Glu Arg Leu Trp Gln Ala Arg
 35 275 280 285
 Ala Asp Gly Leu Ala Ala Phe Ala Thr Met Asp Ala Gly Pro Asn Val
 290 295 300
 40 Lys Leu Ile Phe Glu Glu Ser Ser Ala Ala Asp Val Leu Tyr Leu Phe
 305 310 315 320
 Pro Asp Ala Ser Leu Ile Ala Pro Phe Glu Gly Arg
 325 330
 45 <210> 54
 <211> 353
 <212> PRT
 <213> 链霉菌属菌种 (Streptomyces sp.)
 50 <400> 54
 Met Thr Glu Thr His Ala Ile Ala Gly Val Pro Met Arg Trp Val Gly
 1 5 10 15
 Pro Leu Arg Ile Ser Gly Asn Val Ala Glu Thr Glu Thr Gln Val Pro
 55 20 25 30
 Leu Ala Thr Tyr Glu Ser Pro Leu Trp Pro Ser Val Gly Arg Gly Ala
 35 40 45
 60 Lys Val Ser Arg Leu Thr Glu Lys Gly Ile Val Ala Thr Leu Val Asp

	50					55						60					
	Glu	Arg	Met	Thr	Arg	Ser	Val	Ile	Val	Glu	Ala	Thr	Asp	Ala	Gln	Thr	
	65					70					75				80		
5	Ala	Tyr	Met	Ala	Ala	Gln	Thr	Ile	His	Ala	Arg	Ile	Asp	Glu	Leu	Arg	
					85					90					95		
	Glu	Val	Val	Arg	Gly	Cys	Ser	Arg	Phe	Ala	Gln	Leu	Ile	Asn	Ile	Lys	
				100					105					110			
10	His	Glu	Ile	Asn	Ala	Asn	Leu	Leu	Phe	Ile	Arg	Phe	Glu	Phe	Thr	Thr	
			115					120					125				
	Gly	Asp	Ala	Ser	Gly	His	Asn	Met	Ala	Thr	Leu	Ala	Ser	Asp	Val	Leu	
15		130					135					140					
	Leu	Gly	His	Leu	Leu	Glu	Thr	Ile	Pro	Gly	Ile	Ser	Tyr	Gly	Ser	Ile	
	145					150					155					160	
20	Ser	Gly	Asn	Tyr	Cys	Thr	Asp	Lys	Lys	Ala	Thr	Ala	Ile	Asn	Gly	Ile	
					165					170					175		
	Leu	Gly	Arg	Gly	Lys	Asn	Val	Ile	Thr	Glu	Leu	Leu	Val	Pro	Arg	Asp	
				180					185					190			
25	Val	Val	Glu	Asn	Asn	Leu	His	Thr	Thr	Ala	Ala	Lys	Ile	Val	Glu	Leu	
			195					200					205				
	Asn	Ile	Arg	Lys	Asn	Leu	Leu	Gly	Thr	Leu	Leu	Ala	Gly	Gly	Ile	Arg	
30		210					215					220					
	Ser	Ala	Asn	Ala	His	Phe	Ala	Asn	Met	Leu	Leu	Gly	Phe	Tyr	Leu	Ala	
	225					230					235					240	
35	Thr	Gly	Gln	Asp	Ala	Ala	Asn	Ile	Val	Glu	Gly	Ser	Gln	Gly	Val	Val	
					245					250					255		
	Met	Ala	Glu	Asp	Arg	Asp	Gly	Asp	Leu	Tyr	Phe	Ala	Cys	Thr	Leu	Pro	
				260					265					270			
40	Asn	Leu	Ile	Val	Gly	Thr	Val	Gly	Asn	Gly	Lys	Gly	Leu	Gly	Phe	Val	
			275					280					285				
	Glu	Thr	Asn	Leu	Ala	Arg	Leu	Gly	Cys	Arg	Ala	Asp	Arg	Glu	Pro	Gly	
45		290					295					300					
	Glu	Asn	Ala	Arg	Arg	Leu	Ala	Val	Ile	Ala	Ala	Ala	Thr	Val	Leu	Cys	
	305					310					315					320	
50	Gly	Glu	Leu	Ser	Leu	Leu	Ala	Ala	Gln	Thr	Asn	Pro	Gly	Glu	Leu	Met	
					325					330					335		
	Arg	Ala	His	Val	Gln	Leu	Glu	Arg	Asp	Asn	Lys	Thr	Ala	Lys	Val	Gly	
55				340					345					350			
	Ala																

```

<210> 55
<211> 353
<212> PRT
<213> 灰孢链霉菌(Streptomyces griseosporeus)
5 <400> 55
Met Thr Glu Ala His Ala Thr Ala Gly Val Pro Met Arg Trp Val Gly
1 5 10 15

Pro Val Arg Ile Ser Gly Asn Val Ala Thr Ile Glu Thr Gln Val Pro
10 20 25 30

Leu Ala Thr Tyr Glu Ser Pro Leu Trp Pro Ser Val Gly Arg Gly Ala
35 40 45

15 Lys Val Ser Arg Leu Thr Glu Lys Gly Ile Val Ala Thr Leu Val Asp
50 55 60

Glu Arg Met Thr Arg Ser Val Leu Val Glu Ala Thr Asp Ala Leu Thr
20 65 70 75 80

Ala Leu Ser Ala Ala Arg Thr Ile Glu Ala Arg Ile Asp Glu Leu Arg
85 90 95

25 Glu Leu Val Arg Gly Cys Ser Arg Phe Ala Gln Leu Ile Gly Ile Arg
100 105 110

His Glu Ile Thr Gly Asn Leu Leu Phe Val Arg Phe Glu Phe Ser Thr
115 120 125

30 Gly Asp Ala Ser Gly His Asn Met Ala Thr Leu Ala Ser Asp Val Leu
130 135 140
Leu Gln His Leu Leu Glu Thr Val Pro Gly Ile Ser Tyr Gly Ser Ile
145 150 155 160

35 Ser Gly Asn Tyr Cys Thr Asp Lys Lys Ala Thr Ala Ile Asn Gly Ile
165 170 175

Leu Gly Arg Gly Lys Asn Val Val Thr Glu Leu Leu Val Pro Arg Asp
40 180 185 190

Val Val Ala Asp Val Leu Asn Thr Thr Ala Ala Lys Ile Ala Glu Leu
195 200 205

45 Asn Leu Arg Lys Asn Leu Leu Gly Thr Leu Leu Ala Gly Gly Ile Arg
210 215 220

Ser Ala Asn Ala His Tyr Ala Asn Met Leu Leu Ala Phe Tyr Leu Ala
225 230 235 240

50 Thr Gly Gln Asp Ala Ala Asn Ile Val Glu Gly Ser Gln Gly Val Val
245 250 255

Thr Ala Glu Asp Arg Asp Gly Asp Leu Tyr Leu Ala Cys Thr Leu Pro
55 260 265 270

Asn Leu Ile Val Gly Thr Val Gly Asn Gly Lys Gly Leu Gly Phe Val
275 280 285

60 Glu Thr Asn Leu Asn Arg Leu Gly Cys Arg Ala Asp Arg Glu Pro Gly
290 295 300

```

Glu Asn Ala Arg Arg Leu Ala Val Ile Ala Ala Ala Thr Val Leu Cys
 305 310 315 320
 5 Gly Glu Leu Ser Leu Leu Ala Ala Gln Thr Asn Pro Gly Glu Leu Met
 325 330 335
 Arg Ala His Val Gln Leu Glu Arg Gly His Thr Thr Ala Lys Ala Gly
 340 345 350
 10 Val
 <210> 56
 15 <211> 353
 <212> PRT
 <213> 链霉菌属菌种 (Streptomyces sp.)
 <400> 56
 Met Thr Asp Thr His Ala Ile Ala Met Val Pro Met Lys Trp Val Gly
 20 1 5 10 15
 Pro Leu Arg Ile Ser Gly Asn Val Ala Thr Thr Glu Thr His Val Pro
 20 25 30
 25 Leu Ala Thr Tyr Glu Thr Pro Leu Trp Pro Ser Val Gly Arg Gly Ala
 35 40 45
 Lys Val Ser Met Leu Ser Glu Arg Gly Ile Ala Ala Thr Leu Val Asp
 50 55 60
 30 Glu Arg Met Thr Arg Ser Val Leu Val Glu Ala Thr Asp Ala Gln Thr
 65 70 75 80
 Ala Tyr Thr Ala Ala Arg Ala Ile Glu Ala Arg Ile Glu Glu Leu Arg
 35 85 90 95
 Ala Val Val Arg Thr Cys Ser Arg Phe Ala Glu Leu Leu Gln Val Arg
 100 105 110
 40 His Glu Ile Ala Gly Asn Leu Leu Phe Val Arg Phe Glu Phe Ser Thr
 115 120 125
 Arg Arg Pro Ser Gly His Asn Met Ala Thr Leu Ala Ser Asp Ala Leu
 130 135 140
 45 Leu Ala His Leu Leu Gln Thr Ile Pro Gly Ile Ser Tyr Gly Ser Ile
 145 150 155 160
 Ser Gly Asn Tyr Cys Thr Asp Lys Lys Ala Thr Ala Ile Asn Gly Ile
 50 165 170 175
 Leu Gly Arg Gly Lys Asn Val Val Thr Glu Leu Val Val Pro Arg Glu
 180 185 190
 55 Val Val Glu Arg Val Leu His Thr Thr Ala Ala Lys Ile Val Glu Leu
 195 200 205
 Asn Ile Arg Lys Asn Leu Leu Gly Thr Leu Leu Ala Gly Gly Ile Arg
 210 215 220
 60 Ser Ala Asn Ala His Tyr Ala Asn Met Leu Leu Gly Phe Tyr Leu Ala

	225		230		235		240
	Thr Gly Gln Asp	Ala Ala Asn Ile Val	Glu Gly Ser Gln Gly Val Thr				
		245	250			255	
5	Leu Ala Glu Asp	Arg Asp Gly Asp	Leu Tyr Phe Ser Cys Asn Leu Pro				
		260	265			270	
	Asn Leu Ile Val	Gly Thr Val Gly	Asn Gly Lys Gly Leu Glu Phe Val				
10		275	280			285	
	Glu Thr Asn Leu	Asn Arg Leu Gly Cys Arg	Glu Asp Arg Ala Pro Gly				
		290	295			300	
15	Glu Asn Ala Arg	Arg Leu Ala Val Ile	Ala Ala Ala Thr Val Leu Cys				
	305	310	315			320	
	Gly Glu Leu Ser	Leu Leu Ala Ala Gln	Thr Asn Pro Gly Glu Leu Met				
20		325	330			335	
	Arg Ala His Val	Glu Leu Glu Arg	Asp Asn Thr Thr Ala Glu Val Gly				
		340	345			350	
	Val						
25							
	<210>	57					
	<211>	347					
	<212>	PRT					
30	<213>	草生欧文氏菌(Erwinia herbicola)					
	<400>	57					
	Met Lys Asp Glu	Arg Leu Val Gln Arg	Lys Asn Asp His Leu Asp Ile				
	1	5	10			15	
35	Val Leu Asp Pro	Arg Arg Ala Val Thr	Gln Ala Ser Ala Gly Phe Glu				
		20	25			30	
	Arg Trp Arg Phe	Thr His Cys Ala Leu Pro	Glu Leu Asn Phe Ser Asp				
		35	40			45	
40	Ile Thr Leu Glu	Thr Thr Phe Leu Asn Arg	Gln Leu Gln Ala Pro Leu				
		50	55			60	
	Leu Ile Ser Ser	Met Thr Gly Gly Val	Glu Arg Ser Arg His Ile Asn				
45		65	70			75	80
	Arg His Leu Ala	Glu Ala Ala Gln Val	Leu Lys Ile Ala Met Gly Val				
		85	90			95	
	Gly Ser Gln Arg	Val Ala Ile Glu Ser	Asp Ala Gly Leu Gly Leu Asp				
50		100	105			110	
	Lys Thr Leu Arg	Gln Leu Ala Pro Asp	Val Pro Leu Leu Ala Asn Leu				
		115	120			125	
55	Gly Ala Ala Gln	Leu Thr Gly Arg Lys Gly	Ile Asp Tyr Ala Arg Arg				
		130	135			140	
	Ala Val Glu Met	Ile Glu Ala Asp Ala Leu	Ile Val His Leu Asn Pro				
60		145	150			155	160

	Leu	Gln	Glu	Ala	Leu	Gln	Pro	Gly	Gly	Asp	Arg	Asp	Trp	Arg	Gly	Arg	
					165					170					175		
5	Leu	Ala	Ala	Ile	Glu	Thr	Leu	Val	Arg	Glu	Leu	Pro	Val	Pro	Leu	Val	
				180					185					190			
	Val	Lys	Glu	Val	Gly	Ala	Gly	Ile	Ser	Arg	Thr	Val	Ala	Gly	Gln	Leu	
			195					200					205				
10	Ile	Asp	Ala	Gly	Val	Thr	Val	Ile	Asp	Val	Ala	Gly	Ala	Gly	Gly	Thr	
		210					215					220					
	Ser	Trp	Ala	Ala	Val	Glu	Gly	Glu	Arg	Ala	Ala	Thr	Glu	Gln	Gln	Arg	
	225					230					235					240	
15	Ser	Val	Ala	Asn	Val	Phe	Ala	Asp	Trp	Gly	Ile	Pro	Thr	Ala	Glu	Ala	
				245					250						255		
20	Leu	Val	Asp	Ile	Ala	Glu	Ala	Trp	Pro	Gln	Met	Pro	Leu	Ile	Ala	Ser	
				260					265					270			
	Gly	Gly	Ile	Lys	Asn	Gly	Val	Asp	Ala	Ala	Lys	Ala	Leu	Arg	Leu	Gly	
			275					280					285				
25	Ala	Cys	Met	Val	Gly	Gln	Ala	Ala	Ala	Val	Leu	Gly	Ser	Ala	Gly	Val	
		290					295					300					
30	Ser	Thr	Glu	Lys	Val	Ile	Asp	His	Phe	Asn	Val	Ile	Ile	Glu	Gln	Leu	
	305					310					315					320	
	Arg	Val	Ala	Cys	Phe	Cys	Thr	Gly	Ser	Arg	Ser	Leu	Ser	Asp	Leu	Lys	
				325						330					335		
35	Gln	Ala	Asp	Ile	Arg	Tyr	Val	Arg	Asp	Thr	Pro						
				340					345								
	<210>	58															
	<211>	360															
40	<212>	PRT															
	<213>	布氏疏螺旋体 (Borrelia burgdorferi)															
	<400>	58															
	Met	Met	Asp	Thr	Glu	Phe	Met	Gly	Ile	Glu	Pro	Asn	Ile	Leu	Glu	Asn	
	1				5					10					15		
45	Lys	Lys	Arg	His	Ile	Glu	Ile	Cys	Leu	Asn	Lys	Asn	Asp	Val	Lys	Gly	
				20					25					30			
	Gly	Cys	Asn	Phe	Leu	Lys	Phe	Ile	Lys	Leu	Lys	His	Asn	Ala	Leu	Ser	
50			35					40					45				
	Asp	Phe	Asn	Phe	Ser	Glu	Ile	Asn	Ile	Lys	Glu	Glu	Ile	Phe	Gly	Tyr	
		50					55					60					
55	Asn	Ile	Ser	Met	Pro	Val	Phe	Ile	Ser	Ser	Met	Thr	Gly	Gly	Ser	Lys	
	65					70					75					80	
	Glu	Gly	Asn	Asp	Phe	Asn	Lys	Ser	Leu	Val	Arg	Ile	Ala	Asn	Tyr	Leu	
					85					90					95		
60																	

```

      Lys Ile Pro Ile Gly Leu Gly Ser Phe Lys Leu Leu Phe Lys Tyr Pro
      100                               105                               110

5    Glu Tyr Ile Arg Asp Phe Thr Leu Lys Arg Tyr Ala His Asn Ile Pro
      115                               120                               125
      Leu Phe Ala Asn Val Gly Ala Val Gln Ile Val Glu Phe Gly Ile Ser
      130                               135                               140

10   Lys Ile Ala Glu Met Ile Lys Arg Leu Glu Val Asp Ala Ile Ile Val
      145                               150                               155                               160

      His Leu Asn Ala Gly Gln Glu Leu Met Lys Val Asp Gly Asp Arg Asn
      165                               170                               175

15   Phe Lys Gly Ile Arg Glu Ser Ile Ala Lys Leu Ser Asp Phe Leu Ser
      180                               185                               190

      Val Pro Leu Ile Val Lys Glu Thr Gly Phe Gly Ile Ser Pro Lys Asp
      195                               200                               205

20   Val Lys Glu Leu Phe Ser Leu Gly Ala Ser Tyr Val Asp Leu Ala Gly
      210                               215                               220

      Ser Gly Gly Thr Asn Trp Ile Leu Val Glu Gly Met Lys Ser Asn Asn
      225                               230                               235                               240

25   Leu Asn Ile Ala Ser Cys Phe Ser Asp Trp Gly Ile Pro Ser Val Phe
      245                               250                               255

      Thr Leu Leu Ser Ile Asp Asp Ser Leu Lys Ala Asn Ile Phe Ala Ser
      260                               265                               270

30   Gly Gly Tyr Glu Thr Gly Met Asp Ile Ala Lys Gly Ile Ala Leu Gly
      275                               280                               285

35   Ala Arg Leu Ile Gly Val Ala Ala Val Val Leu Arg Ala Phe Tyr Asp
      290                               295                               300

      Ser Gly Glu Asp Ala Val Phe Gly Leu Phe Ser Asp Tyr Glu His Ile
      305                               310                               315                               320

40   Leu Lys Met Ser Met Phe Leu Ser Gly Ser Lys Ser Leu Leu Glu Phe
      325                               330                               335

      Arg Asn Asn Lys Tyr Phe Leu Ser Ser Tyr Leu Leu Asp Glu Leu Gly
      340                               345                               350

50   Val Phe Lys Gln Phe Tyr Gly Thr
      355                               360

<210> 59
<211> 349
<212> PRT
55 <213> 集胞蓝细菌属菌种(Synechocystis sp.)
    <400> 59
    Met Asp Ser Thr Pro His Arg Lys Ser Asp His Ile Arg Ile Val Leu
    1                               5                               10                               15

60   Glu Glu Asp Val Val Gly Lys Gly Ile Ser Thr Gly Phe Glu Arg Leu

```

	20							25							30							
	Met	Leu	Glu	His	Cys	Ala	Leu	Pro	Ala	Val	Asp	Leu	Asp	Ala	Val	Asp						
	35							40							45							
5	Leu	Gly	Leu	Thr	Leu	Trp	Gly	Lys	Ser	Leu	Thr	Tyr	Pro	Trp	Leu	Ile						
	50							55							60							
10	Ser	Ser	Met	Thr	Gly	Gly	Thr	Pro	Glu	Ala	Lys	Gln	Ile	Asn	Leu	Phe						
	65							70							75							80
	Leu	Ala	Glu	Val	Ala	Gln	Ala	Leu	Gly	Ile	Ala	Met	Gly	Leu	Gly	Ser						
	85							90							95							
15	Gln	Arg	Ala	Ala	Ile	Glu	Asn	Pro	Asp	Leu	Ala	Phe	Thr	Tyr	Gln	Val						
	100							105							110							
	Arg	Ser	Val	Ala	Pro	Asp	Ile	Leu	Leu	Phe	Ala	Asn	Leu	Gly	Leu	Val						
20	115							120							125							
	Gln	Leu	Asn	Tyr	Gly	Tyr	Gly	Leu	Glu	Gln	Ala	Gln	Arg	Ala	Val	Asp						
	130							135							140							
25	Met	Ile	Glu	Ala	Asp	Ala	Leu	Ile	Leu	His	Leu	Asn	Pro	Leu	Gln	Glu						
	145							150							155							160
	Ala	Val	Gln	Pro	Asp	Gly	Asp	Arg	Leu	Trp	Ser	Gly	Leu	Trp	Ser	Lys						
	165							170							175							
30	Leu	Glu	Ala	Leu	Val	Glu	Ala	Leu	Glu	Val	Pro	Val	Ile	Val	Lys	Glu						
	180							185							190							
	Val	Gly	Asn	Gly	Ile	Ser	Gly	Pro	Val	Ala	Lys	Arg	Leu	Gln	Glu	Cys						
	195							200							205							
35	Gly	Val	Gly	Ala	Ile	Asp	Val	Ala	Gly	Ala	Gly	Gly	Thr	Ser	Trp	Ser						
	210							215							220							
	Glu	Val	Glu	Ala	His	Arg	Gln	Thr	Asp	Arg	Gln	Ala	Lys	Glu	Val	Ala						
40	225							230							235							240
	His	Asn	Phe	Ala	Asp	Trp	Gly	Leu	Pro	Thr	Ala	Trp	Ser	Leu	Gln	Gln						
	245							250							255							
45	Val	Val	Gln	Asn	Thr	Glu	Gln	Ile	Leu	Val	Phe	Ala	Ser	Gly	Gly	Ile						
	260							265							270							
	Arg	Ser	Gly	Ile	Asp	Gly	Ala	Lys	Ala	Ile	Ala	Leu	Gly	Ala	Thr	Leu						
	275							280							285							
50	Val	Gly	Ser	Ala	Ala	Pro	Val	Leu	Ala	Glu	Ala	Lys	Ile	Asn	Ala	Gln						
	290							295							300							
	Arg	Val	Tyr	Asp	His	Tyr	Gln	Ala	Arg	Leu	Arg	Glu	Leu	Gln	Ile	Ala						
55	305							310							315							320
	Ala	Phe	Cys	Cys	Asp	Ala	Ala	Asn	Leu	Thr	Gln	Leu	Ala	Gln	Val	Pro						
	325							330							335							
60	Leu	Trp	Asp	Arg	Gln	Ser	Gly	Gln	Arg	Leu	Thr	Lys	Pro									
	340							345														

```

<210> 60
<211> 361
<212> PRT
5 <213> 链霉菌属菌种(Streptomyces sp.)
<400> 60
Met Thr Ser Ala Gln Arg Lys Asp Asp His Val Arg Leu Ala Ile Glu
1 5 10 15
10 Gln His Asn Ala His Ser Gly Arg Asn Gln Asp Asp Val Ser Phe Val
20 25 30
His His Ala Leu Ala Gly Ile Asp Arg Pro Asp Val Ser Leu Ala Thr
35 40 45
15 Ser Phe Ala Gly Ile Ser Trp Gln Val Pro Ile Tyr Ile Asn Ala Met
50 55 60
Thr Gly Gly Ser Glu Lys Thr Gly Leu Ile Asn Arg Asp Leu Ala Thr
65 70 75 80
20 Ala Ala Arg Glu Thr Gly Val Pro Ile Ala Ser Gly Ser Met Asn Ala
85 90 95
Tyr Ile Lys Asp Pro Cys Ala Asp Thr Phe Arg Val Leu Arg Asp Glu
100 105 110
25 Asn Pro Asn Gly Phe Val Ile Ala Asn Ile Asn Ala Thr Thr Thr Val
115 120 125
30 Asp Asn Ala Gln Arg Ala Ile Asp Leu Ile Glu Ala Asn Ala Leu Gln
130 135 140
Ile His Ile Asn Thr Ala Gln Glu Thr Pro Met Pro Glu Gly Asp Arg
145 150 155
35 Ser Phe Ala Ser Trp Val Pro Gln Ile Glu Lys Ile Ala Ala Ala Val
165 170 175
Asp Ile Pro Val Ile Val Lys Glu Val Gly Asn Gly Leu Ser Arg Gln
180 185 190
40 Thr Ile Leu Leu Leu Ala Asp Leu Gly Val Gln Ala Ala Asp Val Ser
195 200 205
45 Gly Arg Gly Gly Thr Asp Phe Ala Arg Ile Glu Asn Gly Arg Arg Glu
210 215 220
Leu Gly Asp Tyr Ala Phe Leu His Gly Trp Gly Gln Ser Thr Ala Ala
225 230 235 240
50 Cys Leu Leu Asp Ala Gln Asp Ile Ser Leu Pro Val Leu Ala Ser Gly
245 250 255
Gly Val Arg His Pro Leu Asp Val Val Arg Ala Leu Ala Leu Gly Ala
260 265 270
55 Arg Ala Val Gly Ser Ser Ala Gly Phe Leu Arg Thr Leu Met Asp Asp
275 280 285
60

```

Gly Val Asp Ala Leu Ile Thr Lys Leu Thr Thr Trp Leu Asp Gln Leu
 290 295 300

5 Ala Ala Leu Gln Thr Met Leu Gly Ala Arg Thr Pro Ala Asp Leu Thr
 305 310 315 320

Arg Cys Asp Val Leu Leu His Gly Glu Leu Arg Asp Phe Cys Ala Asp
 325 330 335

10 Arg Gly Ile Asp Thr Arg Arg Leu Ala Gln Arg Ser Ser Ser Ile Glu
 340 345 350

Ala Leu Gln Thr Thr Gly Ser Thr Arg
 355 360

15 <210> 61
 <211> 364
 <212> PRT
 <213> 灰孢链霉菌 (Streptomyces griseosporus)
 20 <400> 61

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

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

Val His His Ala Leu Ala Gly Ile Asp Arg Pro Asp Val Arg Leu Ala
 35 40 45

30 Thr Thr Phe Ala Gly Ile Thr Trp Arg Leu Pro Leu Tyr Ile Asn Ala
 50 55 60

Met Thr Gly Gly Ser Ala Lys Thr Gly Ala Ile Asn Arg Asp Leu Ala
 65 70 75 80

35 Val Ala Ala Arg Glu Thr Gly Ala Ala Ile Ala Ser Gly Ser Met His
 85 90 95

40 Ala Phe Phe Arg Asp Pro Ser Cys Ala Asp Thr Phe Arg Val Leu Arg
 100 105 110

Thr Glu Asn Pro Asp Gly Phe Val Met Ala Asn Val Asn Ala Thr Ala
 115 120 125

45 Ser Val Asp Asn Ala Arg Arg Ala Val Asp Leu Ile Glu Ala Asn Ala
 130 135 140

Leu Gln Ile His Leu Asn Thr Ala Gln Glu Thr Pro Met Pro Glu Gly
 145 150 155 160

50 Asp Arg Ser Phe Gly Ser Trp Pro Ala Gln Ile Ala Lys Ile Thr Ala
 165 170 175

Ala Val Asp Val Pro Val Ile Val Lys Glu Val Gly Asn Gly Leu Ser
 180 185 190

55 Arg Gln Thr Leu Leu Ala Leu Pro Asp Leu Gly Val Arg Val Ala Asp
 195 200 205

60 Val Ser Gly Arg Gly Gly Thr Asp Phe Ala Arg Ile Glu Asn Ser Arg
 210 215 220

```

      Arg Pro Leu Gly Asp Tyr Ala Phe Leu His Gly Trp Gly Gln Ser Thr
      225                               230           235                     240

5   Pro Ala Cys Leu Leu Asp Ala Gln Asp Val Gly Phe Pro Leu Leu Ala
      245                               250

      Ser Gly Gly Ile Arg Asn Pro Leu Asp Val Ala Arg Ala Leu Ala Leu
      260                               265                     270

10  Gly Ala Gly Ala Val Gly Ser Ser Gly Val Phe Leu Arg Thr Leu Ile
      275                               280                     285

      Asp Gly Gly Val Ser Ala Leu Val Ala Gln Ile Ser Thr Trp Leu Asp
15  290                               295           300

      Gln Leu Ala Ala Leu Gln Thr Met Leu Gly Ala Arg Thr Pro Ala Asp
      305                               310           315                     320

20  Leu Thr Arg Cys Asp Val Leu Ile His Gly Pro Leu Arg Ser Phe Cys
      325                               330           335

      Thr Asp Arg Gly Ile Asp Ile Gly Arg Phe Ala Arg Arg Ser Ser Ser
25  340                               345                     350

      Ala Asp Ile Arg Ser Glu Met Thr Gly Ser Thr Arg
      355                               360

30  <210> 62
      <211> 368
      <212> PRT
      <213> 硫磺矿硫化叶菌(Sulfolobus solfataricus)
      <400> 62

35  Met Pro Asp Ile Val Asn Arg Lys Val Glu His Val Glu Ile Ala Ala
      1                               5           10                     15

      Phe Glu Asn Val Asp Gly Leu Ser Ser Ser Thr Phe Leu Asn Asp Val
      20                               25                     30

40  Ile Leu Val His Gln Gly Phe Pro Gly Ile Ser Phe Ser Glu Ile Asn
      35                               40                     45

      Thr Lys Thr Lys Phe Phe Arg Lys Glu Ile Ser Ala Pro Ile Met Val
45  50                               55           60

      Thr Gly Met Thr Gly Gly Arg Asn Glu Leu Gly Arg Ile Asn Arg Ile
      65                               70           75                     80

50  Ile Ala Glu Val Ala Glu Lys Phe Gly Ile Pro Met Gly Val Gly Ser
      85                               90                     95

      Gln Arg Val Ala Ile Glu Lys Ala Glu Ala Arg Glu Ser Phe Thr Ile
      100                              105                     110

55  Val Arg Lys Val Ala Pro Thr Ile Pro Ile Ile Ala Asn Leu Gly Met
      115                              120                     125

      Pro Gln Leu Val Lys Gly Tyr Gly Leu Lys Glu Phe Gln Asp Ala Ile
60  130                              135           140

```

	Gln	Met	Ile	Glu	Ala	Asp	Ala	Ile	Ala	Val	His	Leu	Asn	Pro	Ala	Gln
	145					150					155					160
5	Glu	Val	Phe	Gln	Pro	Glu	Gly	Glu	Pro	Glu	Tyr	Gln	Ile	Tyr	Ala	Leu
					165					170					175	
	Glu	Arg	Leu	Arg	Asp	Ile	Ser	Lys	Glu	Leu	Ser	Val	Pro	Ile	Ile	Val
				180					185					190		
10	Lys	Glu	Ser	Gly	Asn	Gly	Ile	Ser	Met	Glu	Thr	Ala	Lys	Leu	Leu	Tyr
			195					200					205			
	Ser	Tyr	Gly	Ile	Lys	Asn	Phe	Asp	Thr	Ser	Gly	Gln	Gly	Gly	Thr	Asn
		210					215					220				
15	Trp	Ile	Ala	Ile	Glu	Met	Ile	Arg	Asp	Ile	Arg	Arg	Gly	Asn	Trp	Lys
	225					230					235					240
	Ala	Glu	Ser	Ala	Lys	Asn	Phe	Leu	Asp	Trp	Gly	Val	Pro	Thr	Ala	Ala
20					245					250					255	
	Ser	Ile	Ile	Glu	Val	Arg	Tyr	Ser	Ile	Pro	Asp	Ala	Phe	Leu	Val	Gly
				260					265					270		
25	Ser	Gly	Gly	Ile	Arg	Ser	Gly	Leu	Asp	Ala	Ala	Lys	Ala	Ile	Ala	Leu
			275					280					285			
	Gly	Ala	Asp	Ile	Ala	Gly	Met	Ala	Leu	Pro	Val	Leu	Lys	Ser	Ala	Ile
		290					295					300				
30	Glu	Gly	Lys	Glu	Ser	Leu	Glu	Gln	Phe	Phe	Arg	Lys	Ile	Ile	Phe	Glu
	305					310					315					320
	Leu	Lys	Ala	Thr	Met	Met	Leu	Thr	Gly	Ser	Lys	Asn	Val	Glu	Ala	Leu
35					325					330					335	
	Lys	Arg	Ser	Ser	Ile	Val	Ile	Leu	Gly	Lys	Leu	Lys	Glu	Trp	Ala	Glu
				340					345					350		
40	Tyr	Arg	Gly	Ile	Asn	Leu	Ser	Ile	Tyr	Glu	Lys	Val	Arg	Lys	Arg	Glu
			355					360					365			
	<210>	63														
	<211>	342														
45	<212>	PRT														
	<213>	普氏立克次氏体 (Rickettsia prowazekii)														
	<400>	63														
	Met	Pro	Lys	Glu	Gln	Asn	Leu	Asp	Ile	Glu	Arg	Lys	Gln	Glu	His	Ile
	1				5					10					15	
50	Glu	Ile	Asn	Leu	Lys	Gln	Asn	Val	Asn	Ser	Thr	Leu	Lys	Ser	Gly	Leu
				20					25					30		
	Glu	Ser	Ile	Lys	Phe	Ile	His	Asn	Ala	Leu	Pro	Glu	Ile	Asn	Tyr	Asp
			35					40					45			
55	Ser	Ile	Asp	Thr	Thr	Thr	Thr	Phe	Leu	Gly	Lys	Asp	Met	Lys	Ala	Pro
		50						55				60				
	Ile	Leu	Ile	Ser	Ser	Met	Thr	Gly	Gly	Thr	Ala	Arg	Ala	Arg	Asp	Ile
60	65					70					75					80

	Asn	Tyr	Arg	Leu	Ala	Gln	Ala	Ala	Gln	Lys	Ser	Gly	Ile	Ala	Met	Gly	
				85						90					95		
5	Leu	Gly	Ser	Met	Arg	Ile	Leu	Leu	Thr	Lys	Pro	Asp	Thr	Ile	Lys	Thr	
				100					105					110			
	Phe	Thr	Val	Arg	His	Val	Ala	Pro	Asp	Ile	Pro	Leu	Leu	Ala	Asn	Ile	
			115					120					125				
10	Gly	Ala	Val	Gln	Leu	Asn	Tyr	Gly	Val	Thr	Pro	Lys	Glu	Cys	Gln	Tyr	
		130					135					140					
	Leu	Ile	Asp	Thr	Ile	Lys	Ala	Asp	Ala	Leu	Ile	Leu	His	Leu	Asn	Val	
15	145					150					155					160	
	Leu	His	Glu	Leu	Thr	Gln	Pro	Glu	Gly	Asn	Lys	Asn	Trp	Glu	Asn	Leu	
					165					170					175		
20	Leu	Pro	Lys	Ile	Lys	Glu	Val	Ile	Asn	Tyr	Leu	Ser	Val	Pro	Val	Ile	
				180					185					190			
	Val	Lys	Glu	Val	Gly	Tyr	Gly	Leu	Ser	Lys	Gln	Val	Ala	Lys	Lys	Leu	
			195					200					205				
25	Ile	Lys	Ala	Gly	Val	Lys	Val	Leu	Asp	Ile	Ala	Gly	Ser	Gly	Gly	Thr	
		210					215					220					
	Ser	Trp	Ser	Gln	Val	Glu	Ala	Tyr	Arg	Ala	Lys	Asn	Ser	Met	Gln	Asn	
30	225					230					235					240	
	Arg	Ile	Ala	Ser	Ser	Phe	Ile	Asn	Trp	Gly	Ile	Thr	Thr	Leu	Asp	Ser	
					245					250					255		
35	Leu	Lys	Met	Leu	Gln	Glu	Ile	Ser	Lys	Asp	Ile	Thr	Ile	Ile	Ala	Ser	
				260					265					270			
	Gly	Gly	Leu	Gln	Ser	Gly	Ile	Asp	Gly	Ala	Lys	Ala	Ile	Arg	Met	Gly	
40			275					280					285				
	Ala	Asn	Ile	Phe	Gly	Leu	Ala	Gly	Lys	Leu	Leu	Lys	Ala	Ala	Asp	Ile	
		290					295					300					
45	Ala	Glu	Ser	Leu	Val	Leu	Glu	Glu	Ile	Gln	Val	Ile	Ile	Glu	Gln	Leu	
	305					310					315					320	
	Lys	Ile	Thr	Met	Leu	Cys	Thr	Gly	Ser	Cys	Thr	Leu	Lys	Asp	Leu	Ala	
					325					330					335		
50	Lys	Ala	Glu	Ile	Met	Trp											
				340													
	<210>	64															
55	<211>	286															
	<212>	PRT															
	<213>	抗放射异常球菌 (Deinococcus radiodurans)															
	<400>	64															
60	Met	Arg	Leu	Asp	Thr	Val	Phe	Leu	Gly	Arg	Arg	Leu	Lys	Ala	Pro	Val	
	1				5					10					15		

Leu Ile Gly Ala Met Thr Gly Gly Ala Glu Lys Ala Gly Val Ile Asn
 20 25 30
 5 Arg Asn Leu Ala Thr Ala Ala Arg Asn Leu Gly Leu Gly Met Met Leu
 35 40 45
 Gly Ser Gln Arg Val Met Leu Glu His Pro Asp Ala Trp Glu Ser Phe
 50 55 60
 10 Asn Val Arg Glu Val Ala Pro Glu Ile Leu Leu Ile Gly Asn Leu Gly
 65 70 75 80
 Ala Ala Gln Phe Met Leu Gly Tyr Gly Ala Glu Gln Ala Arg Arg Ala
 85 90 95
 15 Val Asp Glu Val Met Ala Asp Ala Leu Ala Ile His Leu Asn Pro Leu
 100 105 110
 Gln Glu Ala Leu Gln Arg Gly Gly Asp Thr Arg Trp Gln Gly Val Thr
 115 120 125
 20 Tyr Arg Leu Lys Gln Val Ala Arg Glu Leu Asp Phe Pro Val Ile Ile
 130 135 140
 25 Lys Glu Val Gly His Gly Leu Asp Ala Ala Thr Leu Arg Ala Leu Ala
 145 150 155 160
 Asp Gly Pro Phe Ala Ala Tyr Asp Val Ala Gly Ala Gly Gly Thr Ser
 165 170 175
 30 Trp Ala Arg Val Glu Gln Leu Val Ala His Gly Gln Val His Ser Pro
 180 185 190
 Asp Leu Cys Glu Leu Gly Val Pro Thr Ala Gln Ala Leu Arg Gln Ala
 195 200 205
 35 Arg Lys Thr Leu Pro Gly Ala Gln Leu Ile Ala Ser Gly Gly Ile Arg
 210 215 220
 40 Ser Gly Leu Asp Ala Ala Arg Ala Leu Ser Leu Gly Ala Glu Val Val
 225 230 235 240
 Ala Val Ala Arg Pro Leu Leu Glu Pro Ala Leu Asp Ser Ser Glu Ala
 245 250 255
 45 Ala Glu Ala Trp Leu Arg Asn Phe Ile Gln Glu Leu Arg Val Ala Leu
 260 265 270
 Phe Val Gly Gly Tyr Arg Asp Val Arg Glu Val Arg Gly Gly
 275 280 285
 50 <210> 65
 <211> 361
 <212> PRT
 55 <213> Aeropyrum pernix
 <400> 65
 Met Ile Val Ser Ser Lys Val Glu Ser Arg Glu Ser Thr Leu Leu Glu
 1 5 10 15
 60 Tyr Val Arg Ile Val His Asn Pro Thr Pro Glu Val Asn Leu Gly Asp
 20 25 30

	Val	Ser	Leu	Glu	Ile	Asp	Phe	Cys	Gly	Gly	Arg	Leu	Arg	Ala	Pro	Leu	
			35					40					45				
5	Val	Ile	Thr	Gly	Met	Thr	Gly	Gly	His	Pro	Asp	Val	Glu	Trp	Ile	Asn	
		50					55					60					
	Arg	Glu	Leu	Ala	Ser	Val	Ala	Glu	Glu	Leu	Gly	Ile	Ala	Ile	Gly	Val	
	65					70					75					80	
10	Gly	Ser	Gln	Arg	Ala	Ala	Ile	Glu	Asp	Pro	Ser	Leu	Ala	Arg	Thr	Phe	
					85					90					95		
	Arg	Ala	Ala	Arg	Glu	Ala	Ala	Pro	Asn	Ala	Phe	Leu	Ile	Ala	Asn	Leu	
15				100					105					110			
	Gly	Ala	Pro	Gln	Leu	Ser	Leu	Gly	Tyr	Ser	Val	Arg	Glu	Val	Arg	Met	
			115					120					125				
20	Ala	Val	Glu	Met	Ile	Asp	Ala	Asp	Ala	Ile	Ala	Ile	His	Leu	Asn	Pro	
		130					135					140					
	Gly	Gln	Glu	Ala	Tyr	Gln	Pro	Glu	Gly	Asp	Pro	Phe	Tyr	Arg	Gly	Val	
	145					150					155					160	
25	Val	Gly	Lys	Ile	Ala	Glu	Ala	Ala	Glu	Ala	Ala	Gly	Val	Pro	Val	Ile	
					165					170					175		
	Val	Lys	Glu	Thr	Gly	Asn	Gly	Leu	Ser	Arg	Glu	Ala	Val	Ala	Gln	Leu	
30				180					185					190			
	Arg	Ala	Leu	Gly	Val	Arg	Cys	Phe	Asp	Val	Ala	Gly	Leu	Gly	Gly	Thr	
			195					200					205				
35	Asn	Trp	Ile	Lys	Ile	Glu	Val	Leu	Arg	Gly	Arg	Lys	Ala	Gly	Ser	Pro	
		210					215					220					
	Leu	Glu	Ala	Gly	Pro	Leu	Gln	Asp	Phe	Trp	Gly	Asn	Pro	Thr	Ala	Ala	
	225					230					235					240	
40	Ala	Leu	Met	Glu	Ala	Arg	Thr	Ala	Ala	Pro	Asp	Ala	Tyr	Ile	Ile	Ala	
					245					250					255		
45	Ser	Gly	Gly	Val	Arg	Asn	Gly	Leu	Asp	Ala	Ala	Arg	Ala	Ile	Ala	Leu	
				260					265					270			
	Gly	Ala	Asp	Ala	Ala	Gly	Val	Ala	Leu	Pro	Ala	Ile	Arg	Ser	Leu	Leu	
			275					280					285				
50	Ser	Gly	Gly	Arg	Gln	Ala	Thr	Leu	Lys	Leu	Leu	Lys	Ala	Ile	Glu	Tyr	
		290					295					300					
	Gln	Leu	Lys	Thr	Ala	Val	Tyr	Met	Val	Gly	Glu	Thr	Arg	Val	Arg	Gly	
55		305				310					315					320	
	Leu	Trp	Arg	Ala	Pro	Ile	Val	Val	Trp	Gly	Arg	Leu	Ala	Glu	Glu	Ala	
					325					330					335		
60	Glu	Ala	Arg	Gly	Ile	Asp	Pro	Arg	Trp	Tyr	Thr	Asn	Thr	Leu	Arg	Leu	

	340	345	350
	Glu Ala Leu Val Tyr Lys Asp Val Lys		
	355	360	
5	<210> 66 <211> 379 <212> PRT <213> 盐杆菌属菌种 (Halobacterium sp.)		
10	Met Gly Glu Ser Arg Tyr Asn Ser Ile Val Phe Pro Ser Leu Val Gln		
	1	5	10 15
15	Thr Arg Leu Met Thr Ala Gln Asp Ser Thr Gln Thr Glu Asp Arg Lys		
	20	25	30
	Asp Asp His Leu Gln Ile Val Gln Glu Arg Asp Val Glu Thr Thr Gly		
	35	40	45
20	Thr Gly Phe Asp Asp Val His Leu Val His Asn Ala Leu Pro Glu Leu		
	50	55	60
	Asp Tyr Asp Ala Ile Asp Pro Ser Ile Asp Phe Leu Gly His Asp Leu		
	65	70	75 80
25	Ser Ala Pro Ile Phe Ile Glu Ser Met Thr Gly Gly His His Asn Thr		
	85	90	95
	Thr Glu Ile Asn Arg Ala Leu Ala Arg Ala Ala Ser Glu Thr Gly Ile		
	100	105	110
30	Ala Met Gly Leu Gly Ser Gln Arg Ala Gly Leu Glu Leu Asp Asp Glu		
	115	120	125
	Arg Val Leu Glu Ser Tyr Thr Val Val Arg Asp Ala Ala Pro Asp Ala		
35	130	135	140
	Phe Ile Tyr Gly Asn Leu Gly Ala Ala Gln Leu Arg Glu Tyr Asp Ile		
	145	150	155 160
40	Glu Met Val Glu Gln Ala Val Glu Met Ile Asp Ala Asp Ala Leu Ala		
	165	170	175
	Val His Leu Asn Phe Leu Gln Glu Ala Thr Gln Pro Glu Gly Asp Val		
	180	185	190
45	Asp Gly Arg Asn Cys Val Ala Ala Ile Glu Arg Val Ser Glu Ala Leu		
	195	200	205
	Ser Val Pro Ile Ile Val Lys Glu Thr Gly Asn Gly Ile Ser Gly Glu		
50	210	215	220
	Thr Ala Arg Glu Leu Thr Ala Ala Gly Val Asp Ala Leu Asp Val Ala		
	225	230	235 240
55	Gly Lys Gly Gly Thr Thr Trp Ser Gly Ile Glu Ala Tyr Arg Ala Ala		
	245	250	255
	Ala Ala Asn Ala Pro Arg Gln Lys Gln Ile Gly Thr Leu Phe Arg Glu		
	260	265	270
60			

	Trp	Gly	Ile	Pro	Thr	Ala	Ala	Ser	Thr	Ile	Glu	Cys	Val	Ala	Glu	His	
			275					280					285				
5	Asp	Cys	Val	Ile	Ala	Ser	Gly	Gly	Val	Arg	Thr	Gly	Leu	Asp	Val	Ala	
		290					295					300					
10	Lys	Ala	Ile	Ala	Leu	Gly	Ala	Arg	Ala	Gly	Gly	Leu	Ala	Lys	Pro	Phe	
	305					310					315					320	
	Leu	Lys	Pro	Ala	Thr	Asp	Gly	Pro	Asp	Ala	Val	Ile	Glu	Arg	Val	Gly	
					325					330					335		
15	Asp	Leu	Ile	Ala	Glu	Leu	Arg	Thr	Ala	Met	Phe	Val	Thr	Gly	Ser	Gly	
				340					345					350			
	Ser	Ile	Asp	Glu	Leu	Gln	Gln	Val	Glu	Tyr	Val	Leu	His	Gly	Lys	Thr	
			355					360					365				
20	Arg	Glu	Tyr	Val	Glu	Gln	Arg	Thr	Ser	Ser	Glu						
		370					375										
	<210>	67															
	<211>	317															
25	<212>	PRT															
	<213>	闪烁古生球菌 (Archaeoglobus fulgidus)															
	<400>	67															
	Met	Met	Leu	Ile	His	Lys	Ala	Leu	Pro	Glu	Val	Asp	Tyr	Trp	Lys	Ile	
	1				5					10					15		
30	Asp	Thr	Glu	Ile	Glu	Phe	Phe	Gly	Lys	Lys	Leu	Ser	Phe	Pro	Leu	Leu	
				20					25					30			
35	Ile	Ala	Ser	Met	Thr	Gly	Gly	His	Pro	Glu	Thr	Lys	Glu	Ile	Asn	Ala	
			35					40					45				
	Arg	Leu	Gly	Glu	Ala	Val	Glu	Glu	Ala	Gly	Ile	Gly	Met	Gly	Val	Gly	
		50					55					60					
40	Ser	Gln	Arg	Ala	Ala	Ile	Glu	Asp	Glu	Ser	Leu	Ala	Asp	Ser	Phe	Thr	
	65					70					75					80	
	Val	Val	Arg	Glu	Lys	Ala	Pro	Asn	Ala	Phe	Val	Tyr	Ala	Asn	Ile	Gly	
					85					90					95		
45	Met	Pro	Gln	Val	Ile	Glu	Arg	Gly	Val	Glu	Ile	Val	Asp	Arg	Ala	Val	
				100					105					110			
	Glu	Met	Ile	Asp	Ala	Asp	Ala	Val	Ala	Ile	His	Leu	Asn	Tyr	Leu	Gln	
			115					120					125				
50	Glu	Ala	Ile	Gln	Pro	Glu	Gly	Asp	Leu	Asn	Ala	Glu	Lys	Gly	Leu	Glu	
		130					135					140					
	Val	Leu	Glu	Glu	Val	Cys	Arg	Ser	Val	Lys	Val	Pro	Val	Ile	Ala	Lys	
55	145					150					155					160	
	Glu	Thr	Gly	Ala	Gly	Ile	Ser	Arg	Glu	Val	Ala	Val	Met	Leu	Lys	Arg	
					165					170					175		
60	Ala	Gly	Val	Ser	Ala	Ile	Asp	Val	Gly	Gly	Lys	Gly	Gly	Thr	Thr	Phe	

	180							185							190						
	Ser	Gly	Val	Glu	Val	Tyr	Arg	Val	Asn	Asp	Glu	Val	Ser	Lys	Ser	Val					
			195					200					205								
5	Gly	Ile	Asp	Phe	Trp	Asp	Trp	Gly	Leu	Pro	Thr	Ala	Phe	Ser	Ile	Val					
		210					215					220									
10	Asp	Cys	Arg	Gly	Ile	Leu	Pro	Val	Ile	Ala	Thr	Gly	Gly	Leu	Arg	Ser					
	225					230					235					240					
	Gly	Leu	Asp	Val	Ala	Lys	Ser	Ile	Ala	Ile	Gly	Ala	Glu	Leu	Gly	Ser					
				245						250					255						
15	Ala	Ala	Leu	Pro	Phe	Leu	Arg	Ala	Ala	Val	Glu	Ser	Ala	Glu	Lys	Val					
				260					265					270							
	Arg	Glu	Glu	Ile	Glu	Tyr	Phe	Arg	Arg	Gly	Leu	Lys	Thr	Ala	Met	Phe					
			275					280					285								
20	Leu	Thr	Gly	Cys	Lys	Asn	Val	Glu	Glu	Leu	Lys	Gly	Leu	Lys	Val	Phe					
		290				295						300									
25	Val	Ser	Gly	Arg	Leu	Lys	Glu	Trp	Ile	Asp	Phe	Arg	Gly								
	305					310					315										
	<210>	68																			
	<211>	370																			
	<212>	PRT																			
30	<213>	Pyrococcus abyssi																			
	<400>	68																			
	Met	Glu	Glu	Gln	Thr	Ile	Leu	Arg	Lys	Phe	Glu	His	Ile	Lys	His	Cys					
	1				5					10					15						
35	Leu	Thr	Lys	Asn	Val	Glu	Ala	His	Val	Thr	Asn	Gly	Phe	Glu	Asp	Val					
			20						25					30							
	His	Leu	Ile	His	Lys	Ser	Leu	Pro	Glu	Ile	Asp	Lys	Asp	Glu	Ile	Asp					
			35					40					45								
40	Leu	Ser	Val	Lys	Phe	Leu	Gly	Arg	Lys	Phe	Asp	Tyr	Pro	Ile	Met	Ile					
		50					55					60									
	Thr	Gly	Met	Thr	Gly	Gly	Thr	Arg	Lys	Gly	Glu	Ile	Ala	Trp	Arg	Ile					
	65				70						75				80						
45	Asn	Arg	Thr	Leu	Ala	Gln	Ala	Ala	Gln	Glu	Leu	Asn	Ile	Pro	Leu	Gly					
					85					90					95						
50	Leu	Gly	Ser	Gln	Arg	Ala	Met	Ile	Glu	Lys	Pro	Glu	Thr	Trp	Glu	Ser					
				100					105					110							
	Tyr	Tyr	Val	Arg	Asp	Val	Ala	Pro	Asp	Val	Phe	Leu	Val	Gly	Asn	Leu					
			115					120					125								
55	Gly	Ala	Pro	Gln	Phe	Gly	Arg	Asn	Ala	Lys	Lys	Arg	Tyr	Ser	Val	Asp					
		130					135					140									
	Glu	Val	Leu	Tyr	Ala	Ile	Glu	Lys	Ile	Glu	Ala	Asp	Ala	Ile	Ala	Ile					
	145					150					155					160					
60																					

	His	Met	Asn	Pro	Leu	Gln	Glu	Ser	Ile	Gln	Pro	Glu	Gly	Asp	Thr	Thr
					165					170					175	
5	Phe	Ser	Gly	Val	Leu	Glu	Ala	Leu	Ala	Glu	Ile	Thr	Ser	Thr	Ile	Asp
				180					185					190		
	Tyr	Pro	Val	Ile	Ala	Lys	Glu	Thr	Gly	Ala	Gly	Val	Ser	Lys	Glu	Val
			195					200					205			
10	Ala	Val	Glu	Leu	Glu	Ala	Val	Gly	Val	Asp	Ala	Ile	Asp	Ile	Ser	Gly
		210					215					220				
	Leu	Gly	Gly	Thr	Ser	Trp	Ser	Ala	Val	Glu	Tyr	Tyr	Arg	Thr	Lys	Asp
15	225					230					235					240
	Gly	Glu	Lys	Arg	Asn	Leu	Ala	Leu	Lys	Phe	Trp	Asp	Trp	Gly	Ile	Lys
					245					250					255	
20	Thr	Ala	Ile	Ser	Leu	Ala	Glu	Val	Arg	Trp	Ala	Thr	Asn	Leu	Pro	Ile
				260					265					270		
	Ile	Ala	Ser	Gly	Gly	Met	Arg	Asp	Gly	Ile	Thr	Met	Ala	Lys	Ala	Leu
25			275					280					285			
	Ala	Met	Gly	Ala	Ser	Met	Val	Gly	Ile	Ala	Leu	Pro	Val	Leu	Arg	Pro
		290					295					300				
30	Ala	Ala	Lys	Gly	Asp	Val	Glu	Gly	Val	Ile	Arg	Ile	Ile	Lys	Gly	Tyr
	305					310					315					320
	Ala	Glu	Glu	Ile	Arg	Asn	Val	Met	Phe	Leu	Val	Gly	Ala	Arg	Asn	Ile
					325					330					335	
35	Lys	Glu	Leu	Arg	Lys	Val	Pro	Leu	Val	Ile	Thr	Gly	Phe	Val	Arg	Glu
				340					345					350		
	Trp	Leu	Leu	Gln	Arg	Ile	Asp	Leu	Asn	Ser	Tyr	Leu	Arg	Ala	Arg	Phe
40			355					360					365			
	Lys	Met														
	<210>	69														
45	<211>	371														
	<212>	PRT														
	<213>	Pyrococcus horikoshii														
	<400>	69														
50	Met	Lys	Glu	Glu	Leu	Thr	Ile	Leu	Arg	Lys	Phe	Glu	His	Ile	Glu	His
	1				5					10					15	
	Cys	Leu	Lys	Arg	Asn	Val	Glu	Ala	His	Val	Ser	Asn	Gly	Phe	Glu	Asp
				20					25					30		
55	Val	Tyr	Phe	Val	His	Lys	Ser	Leu	Pro	Glu	Ile	Asp	Lys	Asp	Glu	Ile
			35					40					45			
	Asp	Leu	Thr	Val	Glu	Phe	Leu	Gly	Arg	Lys	Phe	Asp	Tyr	Pro	Ile	Met
		50					55					60				
60	Ile	Thr	Gly	Met	Thr	Gly	Gly	Thr	Arg	Arg	Glu	Glu	Ile	Ala	Gly	Lys

	65				70					75					80
	Ile	Asn	Arg	Thr	Leu	Ala	Met	Ala	Ala	Glu	Glu	Leu	Asn	Ile	Pro Phe
					85					90					95
5	Gly	Val	Gly	Ser	Gln	Arg	Ala	Met	Ile	Glu	Lys	Pro	Glu	Thr	Trp Glu
				100					105					110	
	Ser	Tyr	Tyr	Val	Arg	Asp	Val	Ala	Pro	Asp	Ile	Phe	Leu	Ile	Gly Asn
10			115					120					125		
	Leu	Gly	Ala	Pro	Gln	Phe	Gly	Lys	Asn	Ala	Lys	Lys	Arg	Tyr	Ser Val
		130					135					140			
15	Lys	Glu	Val	Leu	Tyr	Ala	Ile	Glu	Lys	Ile	Glu	Ala	Asp	Ala	Ile Ala
	145					150					155				160
	Ile	His	Met	Asn	Pro	Leu	Gln	Glu	Ser	Val	Gln	Pro	Glu	Gly	Asp Thr
				165						170					175
20	Thr	Tyr	Ala	Gly	Val	Leu	Glu	Ala	Leu	Ala	Glu	Ile	Lys	Ser	Ser Ile
				180					185					190	
	Asn	Tyr	Pro	Val	Ile	Ala	Lys	Glu	Thr	Gly	Ala	Gly	Val	Ser	Lys Glu
25			195					200					205		
	Val	Ala	Ile	Glu	Leu	Glu	Ser	Val	Gly	Ile	Asp	Ala	Ile	Asp	Ile Ser
		210					215					220			
30	Gly	Leu	Gly	Gly	Thr	Ser	Trp	Ser	Ala	Val	Glu	Tyr	Tyr	Arg	Ala Lys
	225					230					235				240
	Asp	Ser	Glu	Lys	Arg	Lys	Ile	Ala	Leu	Lys	Phe	Trp	Asp	Trp	Gly Ile
				245						250					255
35	Lys	Thr	Ala	Ile	Ser	Leu	Ala	Glu	Val	Arg	Trp	Ala	Thr	Asn	Leu Pro
				260					265					270	
	Ile	Ile	Ala	Ser	Gly	Gly	Met	Arg	Asp	Gly	Val	Met	Met	Ala	Lys Ala
40			275					280					285		
	Leu	Ala	Met	Gly	Ala	Ser	Leu	Val	Gly	Ile	Ala	Leu	Pro	Val	Leu Arg
		290					295					300			
45	Pro	Ala	Ala	Arg	Gly	Asp	Val	Glu	Gly	Val	Val	Arg	Ile	Ile	Arg Gly
	305					310					315				320
	Tyr	Ala	Glu	Glu	Ile	Lys	Asn	Val	Met	Phe	Leu	Val	Gly	Ala	Arg Asn
50					325					330					335
	Ile	Arg	Glu	Leu	Arg	Arg	Val	Pro	Leu	Val	Ile	Thr	Gly	Phe	Val Arg
				340					345					350	
55	Glu	Trp	Leu	Leu	Gln	Arg	Ile	Asp	Leu	Asn	Ser	Tyr	Leu	Arg	Ser Arg
			355					360					365		
	Phe	Lys	His												
60		370													

```

<210> 70
<211> 349
<212> PRT
<213> 嗜热碱甲烷杆菌 (Methanobacterium thermoautotrophicum)
5 <400> 70
Met Ile Ser Asp Arg Lys Leu Glu His Leu Ile Leu Cys Ala Ser Cys
1 5 10 15
10 Asp Val Glu Tyr Arg Lys Lys Thr Gly Phe Glu Asp Ile Glu Ile Val
20 25 30
His Arg Ala Ile Pro Glu Ile Asn Lys Glu Lys Ile Asp Ile Ser Leu
35 40 45
15 Asp Phe Leu Gly Arg Glu Leu Ser Ser Pro Val Met Ile Ser Ala Ile
50 55 60
Thr Gly Gly His Pro Ala Ser Met Lys Ile Asn Arg Glu Leu Ala Arg
65 70 75 80
20 Ala Ala Glu Lys Leu Gly Ile Ala Leu Gly Leu Gly Ser Gln Arg Ala
85 90 95
Gly Val Glu His Pro Glu Leu Glu Gly Thr Tyr Thr Ile Ala Arg Glu
100 105 110
25 Glu Ala Pro Ser Ala Met Leu Ile Gly Asn Ile Gly Ser Ser His Ile
115 120 125
Glu Tyr Ala Glu Arg Ala Val Glu Met Ile Asp Ala Asp Ala Leu Ala
130 135 140
30 Val His Leu Asn Pro Leu Gln Glu Ser Ile Gln Pro Gly Gly Asp Val
145 150 155 160
35 Asp Ser Ser Gly Ala Leu Glu Ser Ile Ser Ala Ile Val Glu Ser Val
165 170 175
Asp Val Pro Val Met Val Lys Glu Thr Gly Ala Gly Ile Cys Ser Glu
180 185 190
40 Asp Ala Ile Glu Leu Glu Ser Cys Gly Val Ser Ala Ile Asp Val Ala
195 200 205
Gly Ala Gly Gly Thr Ser Trp Ala Ala Val Glu Thr Tyr Arg Ala Asp
210 215 220
45 Asp Arg Tyr Leu Gly Glu Leu Phe Trp Asp Trp Gly Ile Pro Thr Ala
225 230 235 240
50 Ala Ser Thr Val Glu Val Val Glu Ser Val Ser Ile Pro Val Ile Ala
245 250 255
Ser Gly Gly Ile Arg Ser Gly Ile Asp Ala Ala Lys Ala Ile Ser Leu
260 265 270
55 Gly Ala Glu Met Val Gly Ile Ala Leu Pro Val Leu Glu Ala Ala Gly
275 280 285
His Gly Tyr Arg Glu Val Ile Lys Val Ile Glu Gly Phe Asn Glu Ala
290 295 300
60

```

```

    Leu Arg Thr Ala Met Tyr Leu Ala Gly Ala Glu Thr Leu Asp Asp Leu
    305                310                315                320

5   Lys Lys Ser Pro Val Ile Ile Thr Gly His Thr Gly Glu Trp Leu Asn
    325                330                335

10  Gln Arg Gly Phe Glu Thr Lys Lys Tyr Ala Arg Arg Ser
    340                345

    <210> 71
    <211> 359
    <212> PRT
15  <213> 詹氏甲烷球菌 (Methanococcus jannaschii)
    <400> 71
    Met Val Asn Asn Arg Asn Glu Ile Glu Val Arg Lys Leu Glu His Ile
    1          5          10          15

20  Phe Leu Cys Ser Tyr Cys Asn Val Glu Tyr Glu Lys Thr Thr Leu Leu
    20          25          30

    Glu Asp Ile Glu Leu Ile His Lys Gly Thr Cys Gly Ile Asn Phe Asn
    35          40          45

25  Asp Ile Glu Thr Glu Ile Glu Leu Phe Gly Lys Lys Leu Ser Ala Pro
    50          55          60

    Ile Ile Val Ser Gly Met Thr Gly Gly His Ser Lys Ala Lys Glu Ile
30  65          70          75          80

    Asn Lys Asn Ile Ala Lys Ala Val Glu Glu Leu Gly Leu Gly Met Gly
    85          90          95

35  Val Gly Ser Gln Arg Ala Ala Ile Val Asn Asp Glu Leu Ile Asp Thr
    100         105         110

    Tyr Ser Ile Val Arg Asp Tyr Thr Asn Asn Leu Val Ile Gly Asn Leu
    115         120         125

40  Gly Ala Val Asn Phe Ile Val Asp Asp Trp Asp Glu Glu Ile Ile Asp
    130         135         140

    Lys Ala Ile Glu Met Ile Asp Ala Asp Ala Ile Ala Ile His Phe Asn
45  145         150         155         160

    Pro Leu Gln Glu Ile Ile Gln Pro Glu Gly Asp Leu Asn Phe Lys Asn
    165         170         175

50  Leu Tyr Lys Leu Lys Glu Ile Ile Ser Asn Tyr Lys Lys Ser Tyr Lys
    180         185         190
    Asn Ile Pro Phe Ile Ala Lys Gln Val Gly Glu Gly Phe Ser Lys Glu
    195         200         205

55  Asp Ala Leu Ile Leu Lys Asp Ile Gly Phe Asp Ala Ile Asp Val Gln
    210         215         220

    Gly Ser Gly Gly Thr Ser Trp Ala Lys Val Glu Ile Tyr Arg Val Lys
    225         230         235         240

60

```

	Glu	Glu	Glu	Ile	Lys	Arg	Leu	Ala	Glu	Lys	Phe	Ala	Asn	Trp	Gly	Ile	
					245					250					255		
5	Pro	Thr	Ala	Ala	Ser	Ile	Phe	Glu	Val	Lys	Ser	Val	Tyr	Asp	Gly	Ile	
				260					265					270			
	Val	Ile	Gly	Ser	Gly	Gly	Ile	Arg	Gly	Gly	Leu	Asp	Ile	Ala	Lys	Cys	
			275					280					285				
10	Ile	Ala	Ile	Gly	Cys	Asp	Cys	Cys	Ser	Val	Ala	Leu	Pro	Ile	Leu	Lys	
		290					295					300					
	Ala	Ser	Leu	Lys	Gly	Trp	Glu	Glu	Val	Val	Lys	Val	Leu	Glu	Ser	Tyr	
	305					310					315					320	
15	Ile	Lys	Glu	Leu	Lys	Ile	Ala	Met	Phe	Leu	Val	Gly	Ala	Glu	Asn	Ile	
					325					330					335		
	Glu	Glu	Leu	Lys	Lys	Thr	Ser	Tyr	Ile	Val	Lys	Gly	Thr	Leu	Lys	Glu	
20				340					345					350			
	Trp	Ile	Ser	Gln	Arg	Leu	Lys										
				355													
25	<210>	72															
	<211>	348															
	<212>	PRT															
	<213>	嗜酸热原体	(Thermoplasma acidophilum)														
	<400>	72															
30	Met	Ile	Gly	Lys	Arg	Lys	Glu	Glu	His	Ile	Arg	Ile	Ala	Glu	Asn	Glu	
	1				5					10					15		
	Asp	Val	Ser	Ser	Phe	His	Asn	Phe	Trp	Asp	Asp	Ile	Ser	Leu	Met	His	
				20					25					30			
35	Glu	Ala	Asp	Pro	Glu	Val	Asn	Tyr	Asp	Glu	Ile	Asp	Thr	Ser	Val	Asp	
			35					40					45				
	Phe	Leu	Gly	Lys	Lys	Leu	Lys	Phe	Pro	Met	Ile	Ile	Ser	Ser	Met	Thr	
	50						55					60					
40	Gly	Gly	Ala	Glu	Ile	Ala	Lys	Asn	Ile	Asn	Arg	Asn	Leu	Ala	Val	Ala	
	65					70				75					80		
	Ala	Glu	Arg	Phe	Gly	Ile	Gly	Met	Gly	Val	Gly	Ser	Met	Arg	Ala	Ala	
45				85						90					95		
	Ile	Val	Asp	Arg	Ser	Ile	Glu	Asp	Thr	Tyr	Ser	Val	Ile	Asn	Glu	Ser	
				100					105					110			
50	His	Val	Pro	Leu	Lys	Ile	Ala	Asn	Ile	Gly	Ala	Pro	Gln	Leu	Val	Arg	
			115					120					125				
	Gln	Asp	Lys	Asp	Ala	Val	Ser	Asn	Arg	Asp	Ile	Ala	Tyr	Ile	Tyr	Asp	
	130						135					140					
55	Leu	Ile	Lys	Ala	Asp	Phe	Leu	Ala	Val	His	Phe	Asn	Phe	Leu	Gln	Glu	
	145					150					155				160		
	Met	Val	Gln	Pro	Glu	Gly	Asp	Arg	Asn	Ser	Lys	Gly	Val	Ile	Asp	Arg	
60					165					170					175		

Ile Lys Asp Leu Ser Gly Ser Phe Asn Ile Ile Ala Lys Glu Thr Gly
 180 185 190
 5 Ser Gly Phe Ser Arg Arg Thr Ala Glu Arg Leu Ile Asp Ala Gly Val
 195 200 205
 Lys Ala Ile Glu Val Ser Gly Val Ser Gly Thr Thr Phe Ala Ala Val
 210 215 220
 10 Glu Tyr Tyr Arg Ala Arg Lys Glu Asn Asn Leu Glu Lys Met Arg Ile
 225 230 235 240
 Gly Glu Thr Phe Trp Asn Trp Gly Ile Pro Ser Pro Ala Ser Val Tyr
 15 245 250 255
 Tyr Cys Ser Asp Leu Ala Pro Val Ile Gly Ser Gly Gly Leu Arg Asn
 260 265 270
 20 Gly Leu Asp Leu Ala Lys Ala Ile Ala Met Gly Ala Thr Ala Gly Gly
 275 280 285
 Phe Ala Arg Ser Leu Leu Lys Asp Ala Asp Thr Asp Pro Glu Met Leu
 25 290 295 300
 Met Lys Asn Ile Glu Leu Ile Gln Arg Glu Phe Arg Val Ala Leu Phe
 305 310 315 320
 30 Leu Thr Gly Asn Lys Asn Val Tyr Glu Leu Lys Phe Thr Lys Lys Val
 325 330 335
 Ile Val Asp Pro Leu Arg Ser Trp Leu Glu Ala Lys
 340 345
 35
 <210> 73
 <211> 357
 <212> PRT
 <213> Leishmania major
 40 <400> 73
 Met Ser Ser Arg Asp Cys Thr Val Asp Arg Glu Ala Ala Val Gln Lys
 1 5 10 15
 Arg Lys Lys Asp His Ile Asp Ile Cys Leu His Gln Asp Val Glu Pro
 45 20 25 30
 His Lys Arg Arg Thr Ser Ile Trp Asn Lys Tyr Thr Leu Pro Tyr Lys
 35 40 45
 50 Ala Leu Pro Glu Val Asp Leu Gln Lys Ile Asp Thr Ser Cys Glu Phe
 50 55 60
 Met Gly Lys Arg Ile Ser Phe Pro Phe Phe Ile Ser Ser Met Thr Gly
 65 70 75 80
 55 Gly Glu Ala His Gly Arg Val Ile Asn Glu Asn Leu Ala Lys Ala Cys
 85 90 95
 Glu Ala Glu Lys Ile Pro Phe Gly Leu Gly Ser Met Arg Ile Ile Asn
 60 100 105 110

	Arg	Tyr	Ala	Ser	Ala	Val	His	Thr	Phe	Asn	Val	Lys	Glu	Phe	Cys	Pro	
			115					120					125				
5	Ser	Val	Pro	Met	Leu	Ala	Asn	Ile	Gly	Leu	Val	Gln	Leu	Asn	Tyr	Gly	
		130					135					140					
	Phe	Gly	Pro	Lys	Glu	Val	Asn	Asn	Leu	Val	Asn	Ser	Val	Arg	Ala	Asp	
	145					150					155					160	
10	Gly	Leu	Cys	Ile	His	Leu	Asn	His	Thr	Gln	Glu	Val	Cys	Gln	Pro	Glu	
				165						170					175		
	Gly	Asp	Thr	Asn	Phe	Glu	Gly	Leu	Ile	Glu	Lys	Leu	Arg	Gln	Leu	Leu	
				180					185					190			
15	Pro	His	Ile	Lys	Val	Pro	Val	Leu	Val	Lys	Gly	Val	Gly	His	Gly	Ile	
			195					200					205				
	Asp	Tyr	Glu	Ser	Met	Val	Ala	Ile	Lys	Ala	Ser	Gly	Val	Lys	Tyr	Val	
20		210					215					220					
	Asp	Val	Ser	Gly	Cys	Gly	Gly	Thr	Ser	Trp	Ala	Trp	Ile	Glu	Gly	Arg	
	225					230					235					240	
25	Arg	Gln	Pro	Tyr	Lys	Ala	Glu	Glu	Glu	Asn	Ile	Gly	Tyr	Leu	Leu	Arg	
					245					250					255		
	Asp	Ile	Gly	Val	Pro	Thr	Asp	Val	Cys	Leu	Arg	Glu	Ser	Ala	Pro	Leu	
				260					265					270			
30	Thr	Val	Asn	Gly	Asp	Leu	His	Leu	Ile	Ala	Gly	Gly	Gly	Ile	Arg	Asn	
			275					280					285				
	Gly	Met	Asp	Val	Ala	Lys	Ala	Leu	Met	Met	Gly	Ala	Glu	Tyr	Ala	Thr	
35		290					295					300					
	Ala	Ala	Met	Pro	Phe	Leu	Ala	Ala	Ala	Leu	Glu	Ser	Ser	Glu	Ala	Val	
	305					310					315					320	
40	Arg	Ala	Val	Ile	Gln	Arg	Met	Arg	Gln	Glu	Leu	Arg	Val	Ser	Met	Phe	
					325					330					335		
	Thr	Cys	Gly	Ala	Arg	Asn	Ile	Glu	Glu	Leu	Arg	Arg	Met	Lys	Val	Ile	
45				340					345					350			
	Glu	Leu	Gly	His	Leu												
				355													
50	<210>	74															
	<211>	398															
	<212>	PRT															
	<213>	肺炎链球菌 (Streptococcus pneumoniae)															
	<400>	74															
55	Met	Asn	Asp	Lys	Thr	Glu	Val	Asn	Met	Thr	Ile	Gly	Ile	Asp	Lys	Ile	
	1				5					10				15			
	Gly	Phe	Ala	Thr	Ser	Gln	Tyr	Val	Leu	Lys	Leu	Gln	Asp	Leu	Ala	Glu	
				20					25					30			
60																	

	Ala	Arg	Gly	Ile	Asp	Pro	Glu	Lys	Leu	Ser	Lys	Gly	Leu	Leu	Leu	Lys
			35					40					45			
5	Glu	Leu	Ser	Ile	Ala	Pro	Leu	Thr	Glu	Asp	Ile	Val	Thr	Leu	Ala	Ala
		50					55					60				
	Ser	Ala	Ser	Asp	Ser	Ile	Leu	Thr	Glu	Gln	Glu	Arg	Gln	Glu	Val	Asp
	65					70				75						80
10	Met	Val	Ile	Val	Ala	Thr	Glu	Ser	Gly	Ile	Asp	Gln	Ser	Lys	Ala	Ala
					85					90					95	
	Ala	Val	Phe	Val	His	Gly	Leu	Leu	Gly	Ile	Gln	Pro	Phe	Ala	Arg	Ser
				100					105					110		
15	Phe	Glu	Ile	Lys	Glu	Ala	Cys	Tyr	Gly	Ala	Thr	Ala	Ala	Leu	His	Tyr
			115					120					125			
	Ala	Lys	Leu	His	Val	Glu	Asn	Ser	Pro	Glu	Ser	Lys	Val	Leu	Val	Ile
20		130					135					140				
	Ala	Ser	Asp	Ile	Ala	Lys	Tyr	Gly	Ile	Glu	Thr	Pro	Gly	Glu	Pro	Thr
	145					150					155					160
25	Gln	Gly	Ala	Gly	Ser	Val	Ala	Met	Leu	Ile	Thr	Gln	Asn	Pro	Arg	Met
					165					170					175	
	Met	Ala	Phe	Asn	Asn	Asp	Asn	Val	Ala	Gln	Thr	Arg	Asp	Ile	Met	Asp
				180					185					190		
30	Phe	Trp	Arg	Pro	Asn	Tyr	Ser	Thr	Thr	Pro	Tyr	Val	Asn	Gly	Val	Tyr
			195					200					205			
	Ser	Thr	Gln	Gln	Tyr	Leu	Asp	Ser	Leu	Lys	Thr	Thr	Trp	Leu	Glu	Tyr
		210					215					220				
35	Gln	Lys	Arg	Tyr	Gln	Leu	Thr	Leu	Asp	Asp	Phe	Ala	Ala	Val	Cys	Phe
	225					230					235					240
	His	Leu	Pro	Tyr	Pro	Lys	Leu	Ala	Leu	Lys	Gly	Leu	Lys	Lys	Ile	Met
40					245					250					255	
	Asp	Lys	Asn	Leu	Pro	Gln	Glu	Lys	Lys	Asp	Leu	Leu	Gln	Lys	His	Phe
				260					265					270		
45	Asp	Gln	Ser	Ile	Leu	Tyr	Ser	Gln	Lys	Val	Gly	Asn	Ile	Tyr	Thr	Gly
			275					280					285			
	Ser	Leu	Phe	Leu	Gly	Leu	Leu	Ser	Leu	Leu	Glu	Asn	Thr	Asp	Ser	Leu
		290					295					300				
50	Lys	Ala	Gly	Asp	Lys	Ile	Ala	Leu	Tyr	Ser	Tyr	Gly	Ser	Gly	Ala	Val
	305					310					315					320
	Ala	Glu	Phe	Phe	Ser	Gly	Glu	Leu	Val	Glu	Gly	Tyr	Glu	Ala	Tyr	Leu
55					325					330					335	
	Asp	Lys	Asp	Arg	Leu	Asn	Lys	Leu	Asn	Gln	Arg	Thr	Ala	Leu	Ser	Val
				340					345					350		
60	Ala	Asp	Tyr	Glu	Lys	Val	Phe	Phe	Glu	Glu	Val	Asn	Leu	Asp	Glu	Thr

[illegible]

	Arg	Glu	Lys	Leu	Ile	Glu	Ala	Phe	Gln	Ala	Ser	Ile	Thr	Tyr	Ser	Lys
				260					265					270		
5	Gln	Ile	Gly	Asn	Ile	Tyr	Thr	Gly	Ser	Leu	Tyr	Leu	Gly	Leu	Leu	Ser
			275					280					285			
	Leu	Leu	Glu	Asn	Ser	Lys	Val	Leu	Gln	Ser	Gly	Asp	Lys	Ile	Gly	Phe
			290				295					300				
10	Phe	Ser	Tyr	Gly	Ser	Gly	Ala	Val	Ser	Glu	Phe	Tyr	Ser	Gly	Gln	Leu
	305					310					315					320
	Val	Ala	Gly	Tyr	Asp	Lys	Met	Leu	Met	Thr	Asn	Arg	Gln	Ala	Leu	Leu
					325					330					335	
15	Asp	Gln	Arg	Thr	Arg	Leu	Ser	Val	Ser	Lys	Tyr	Glu	Asp	Leu	Phe	Tyr
				340					345					350		
	Glu	Gln	Val	Gln	Leu	Asp	Asp	Asn	Gly	Asn	Ala	Asn	Phe	Asp	Ile	Tyr
20			355					360					365			
	Leu	Thr	Gly	Lys	Phe	Ala	Leu	Thr	Ala	Ile	Lys	Glu	His	Gln	Arg	Ile
		370					375					380				
25	Tyr	His	Thr	Asn	Asp	Lys	Asn									
	385					390										
	<210>	76														
	<211>	384														
30	<212>	PRT														
	<213>	粪肠球菌 (Enterococcus faecalis)														
	<400>	76														
	Glu	Met	Thr	Ile	Gly	Ile	Asp	Lys	Ile	Ser	Phe	Phe	Val	Pro	Pro	Tyr
	1			5					10					15		
35	Tyr	Ile	Asp	Met	Thr	Ala	Leu	Ala	Glu	Ala	Arg	Asn	Val	Asp	Pro	Gly
				20					25					30		
	Lys	Phe	His	Ile	Gly	Ile	Gly	Gln	Asp	Gln	Met	Ala	Val	Asn	Pro	Ile
			35					40					45			
40	Ser	Gln	Asp	Ile	Val	Thr	Phe	Ala	Ala	Asn	Ala	Ala	Glu	Ala	Ile	Leu
		50					55					60				
	Thr	Lys	Glu	Asp	Lys	Glu	Ala	Ile	Asp	Met	Val	Ile	Val	Gly	Thr	Glu
45	65					70				75					80	
	Ser	Ser	Ile	Asp	Glu	Ser	Lys	Ala	Ala	Ala	Val	Val	Leu	His	Arg	Leu
				85						90				95		
50	Met	Gly	Ile	Gln	Pro	Phe	Ala	Arg	Ser	Phe	Glu	Ile	Lys	Glu	Ala	Cys
				100					105					110		
	Tyr	Gly	Ala	Thr	Ala	Gly	Leu	Gln	Leu	Ala	Lys	Asn	His	Val	Ala	Leu
			115					120					125			
55	His	Pro	Asp	Lys	Lys	Val	Leu	Val	Val	Ala	Ala	Asp	Ile	Ala	Lys	Tyr
		130					135					140				
	Gly	Leu	Asn	Ser	Gly	Gly	Glu	Pro	Thr	Gln	Gly	Ala	Gly	Ala	Val	Ala
60	145					150					155					160

	Met	Leu	Val	Ala	Ser	Glu	Pro	Arg	Ile	Leu	Ala	Leu	Lys	Glu	Asp	Asn	
					165					170					175		
5	Val	Met	Leu	Thr	Gln	Asp	Ile	Tyr	Asp	Phe	Trp	Arg	Pro	Thr	Gly	His	
				180					185					190			
	Pro	Tyr	Pro	Met	Val	Asp	Gly	Pro	Leu	Ser	Asn	Glu	Thr	Tyr	Ile	Gln	
			195					200					205				
10	Ser	Phe	Ala	Gln	Val	Trp	Asp	Glu	His	Lys	Lys	Arg	Thr	Gly	Leu	Asp	
		210					215					220					
	Phe	Ala	Asp	Tyr	Asp	Ala	Leu	Ala	Phe	His	Ile	Pro	Tyr	Thr	Lys	Met	
15	225					230					235					240	
	Gly	Lys	Lys	Ala	Leu	Leu	Ala	Lys	Ile	Ser	Asp	Gln	Thr	Glu	Ala	Glu	
					245					250					255		
20	Gln	Glu	Arg	Ile	Leu	Ala	Arg	Tyr	Glu	Glu	Ser	Ile	Ile	Tyr	Ser	Arg	
				260					265					270			
	Arg	Val	Gly	Asn	Leu	Tyr	Thr	Gly	Ser	Leu	Tyr	Leu	Gly	Leu	Ile	Ser	
25			275					280					285				
	Leu	Leu	Glu	Asn	Ala	Thr	Thr	Leu	Thr	Ala	Gly	Asn	Gln	Ile	Gly	Leu	
		290					295					300					
30	Phe	Ser	Tyr	Gly	Ser	Gly	Ala	Val	Ala	Glu	Phe	Phe	Thr	Gly	Glu	Leu	
	305					310					315					320	
	Val	Ala	Gly	Tyr	Gln	Asn	His	Leu	Gln	Lys	Glu	Thr	His	Leu	Ala	Leu	
					325					330					335		
35	Leu	Asp	Asn	Arg	Thr	Glu	Leu	Ser	Ile	Ala	Glu	Tyr	Glu	Ala	Met	Phe	
				340					345					350			
	Ala	Glu	Thr	Leu	Asp	Thr	Asp	Ile	Asp	Gln	Thr	Leu	Glu	Asp	Glu	Leu	
40			355					360					365				
	Lys	Tyr	Ser	Ile	Ser	Ala	Ile	Asn	Asn	Thr	Val	Arg	Ser	Tyr	Arg	Asn	
		370					375					380					
45	<210>	77															
	<211>	384															
	<212>	PRT															
	<213>	屎肠球菌 (Enterococcus faecium)															
	<400>	77															
50	Met	Lys	Ile	Gly	Ile	Asp	Arg	Leu	Ser	Phe	Phe	Ile	Pro	Asn	Leu	Tyr	
	1				5					10					15		
	Leu	Asp	Met	Thr	Glu	Leu	Ala	Glu	Ser	Arg	Gly	Asp	Asp	Pro	Ala	Lys	
				20					25					30			
55	Tyr	His	Ile	Gly	Ile	Gly	Gln	Asp	Gln	Met	Ala	Val	Asn	Arg	Ala	Asn	
			35					40					45				
	Glu	Asp	Ile	Ile	Thr	Leu	Gly	Ala	Asn	Ala	Ala	Ser	Lys	Ile	Val	Thr	
60		50					55					60					

	Glu	Lys	Asp	Arg	Glu	Leu	Ile	Asp	Met	Val	Ile	Val	Gly	Thr	Glu	Ser
	65					70					75					80
5	Gly	Ile	Asp	His	Ser	Lys	Ala	Ser	Ala	Val	Ile	Ile	His	His	Leu	Leu
					85					90					95	
	Lys	Ile	Gln	Ser	Phe	Ala	Arg	Ser	Phe	Glu	Val	Lys	Glu	Ala	Cys	Tyr
				100					105					110		
10	Gly	Gly	Thr	Ala	Ala	Leu	His	Met	Ala	Lys	Glu	Tyr	Val	Lys	Asn	His
			115					120					125			
	Pro	Glu	Arg	Lys	Val	Leu	Val	Ile	Ala	Ser	Asp	Ile	Ala	Arg	Tyr	Gly
		130					135					140				
15	Leu	Ala	Ser	Gly	Gly	Glu	Val	Thr	Gln	Gly	Val	Gly	Ala	Val	Ala	Met
	145					150					155					160
	Met	Ile	Thr	Gln	Asn	Pro	Arg	Ile	Leu	Ser	Ile	Glu	Asp	Asp	Ser	Val
20					165					170					175	
	Phe	Leu	Thr	Glu	Asp	Ile	Tyr	Asp	Phe	Trp	Arg	Pro	Asp	Tyr	Ser	Glu
				180					185					190		
25	Phe	Pro	Val	Val	Asp	Gly	Pro	Leu	Ser	Asn	Ser	Thr	Tyr	Ile	Glu	Ser
			195					200					205			
	Phe	Gln	Lys	Val	Trp	Asn	Arg	His	Lys	Glu	Leu	Ser	Gly	Arg	Gly	Leu
		210				215						220				
30	Glu	Asp	Tyr	Gln	Ala	Ile	Ala	Phe	His	Ile	Pro	Tyr	Thr	Lys	Met	Gly
	225					230					235					240
	Lys	Lys	Ala	Leu	Gln	Ser	Val	Leu	Asp	Gln	Thr	Asp	Glu	Asp	Asn	Gln
35					245					250					255	
	Glu	Arg	Leu	Met	Ala	Arg	Tyr	Glu	Glu	Ser	Ile	Arg	Tyr	Ser	Arg	Arg
				260					265					270		
40	Ile	Gly	Asn	Leu	Tyr	Thr	Gly	Ser	Leu	Tyr	Leu	Gly	Leu	Thr	Ser	Leu
			275					280					285			
	Leu	Glu	Asn	Ser	Lys	Ser	Leu	Gln	Pro	Gly	Asp	Arg	Ile	Gly	Leu	Phe
45		290					295					300				
	Ser	Tyr	Gly	Ser	Gly	Ala	Val	Ser	Glu	Phe	Phe	Thr	Gly	Tyr	Leu	Glu
	305					310					315					320
50	Glu	Asn	Tyr	Gln	Glu	Tyr	Leu	Phe	Ala	Gln	Ser	His	Gln	Glu	Met	Leu
					325					330					335	
	Asp	Ser	Arg	Thr	Arg	Ile	Thr	Val	Asp	Glu	Tyr	Glu	Thr	Ile	Phe	Ser
				340					345					350		
55	Glu	Thr	Leu	Pro	Glu	His	Gly	Glu	Cys	Ala	Glu	Tyr	Thr	Ser	Asp	Val
			355					360					365			
	Pro	Phe	Ser	Ile	Thr	Lys	Ile	Glu	Asn	Asp	Ile	Arg	Tyr	Tyr	Lys	Ile
60		370					375					380				

```

<210> 78
<211> 388
<212> PRT
5 <213> 溶血葡萄球菌(Staphylococcus haemolyticus)
<400> 78
Met Ser Ile Gly Ile Asp Lys Ile Asn Phe Tyr Val Pro Lys Tyr Tyr
1 5 10 15
10 Val Asp Met Ala Lys Leu Ala Glu Ala Arg Gln Val Asp Pro Asn Lys
20 25 30
Phe Leu Ile Gly Ile Gly Gln Thr Gln Met Ala Val Ser Pro Val Ser
35 40 45
15 Gln Asp Ile Val Ser Met Gly Ala Asn Ala Ala Lys Asp Ile Ile Thr
50 55 60
20 Asp Asp Asp Lys Lys His Ile Gly Met Val Ile Val Ala Thr Glu Ser
65 70 75 80
Ala Ile Asp Asn Ala Lys Ala Ala Ala Val Gln Ile His Asn Leu Leu
85 90 95
25 Gly Val Gln Pro Phe Ala Arg Cys Phe Glu Met Lys Glu Ala Cys Tyr
100 105 110
Ala Ala Thr Pro Ala Ile Gln Leu Ala Lys Asp Tyr Ile Glu Lys Arg
115 120 125
30 Pro Asn Glu Lys Val Leu Val Ile Ala Ser Asp Thr Ala Arg Tyr Gly
130 135 140
Ile Gln Ser Gly Gly Glu Pro Thr Gln Gly Ala Gly Ala Val Ala Met
145 150 155 160
35 Leu Ile Ser Asn Asn Pro Ser Ile Leu Glu Leu Asn Asp Asp Ala Val
165 170 175
Ala Tyr Thr Glu Asp Val Tyr Asp Phe Trp Arg Pro Thr Gly His Lys
180 185 190
40 Tyr Pro Leu Val Ala Gly Ala Leu Ser Lys Asp Ala Tyr Ile Lys Ser
195 200 205
45 Phe Gln Glu Ser Trp Asn Glu Tyr Ala Arg Arg Glu Asp Lys Thr Leu
210 215 220
Ser Asp Phe Glu Ser Leu Cys Phe His Val Pro Phe Thr Lys Met Gly
225 230 235 240
50 Lys Lys Ala Leu Asp Ser Ile Ile Asn Asp Ala Asp Glu Thr Thr Gln
245 250 255
Glu Arg Leu Thr Ser Gly Tyr Glu Asp Ala Val Tyr Tyr Asn Arg Tyr
55 260 265 270
Val Gly Asn Ile Tyr Thr Gly Ser Leu Tyr Leu Ser Leu Ile Ser Leu
275 280 285
60 Leu Glu Asn Arg Ser Leu Lys Gly Gly Gln Thr Ile Gly Leu Phe Ser

```

	290	295	300
	Tyr Gly Ser Gly Ser Val Gly Glu Phe Phe Ser Ala Thr Leu Val Glu		
	305	310	315 320
5	Gly Tyr Glu Lys Gln Leu Asp Ile Glu Gly His Lys Ala Leu Leu Asn		
		325 330	335
10	Glu Arg Gln Glu Val Ser Val Glu Asp Tyr Glu Ser Phe Phe Lys Arg		
		340 345	350
	Phe Asp Asp Leu Glu Phe Asp His Ala Thr Glu Gln Thr Asp Asp Asp		
		355 360	365
15	Lys Ser Ile Tyr Tyr Leu Glu Asn Ile Gln Asp Asp Ile Arg Gln Tyr		
		370 375	380
20	His Ile Pro Lys		
	385		
	<210> 79		
	<211> 388		
	<212> PRT		
25	<213> 表皮葡萄球菌 (Staphylococcus epidermidis)		
	<400> 79		
	Met Asn Ile Gly Ile Asp Lys Ile Ser Phe Tyr Val Pro Lys Tyr Tyr		
	1	5 10	15
30	Val Asp Met Ala Lys Leu Ala Glu Ala Arg Gln Val Asp Pro Asn Lys		
		20 25	30
	Phe Leu Ile Gly Ile Gly Gln Thr Glu Met Thr Val Ser Pro Val Asn		
		35 40	45
35	Gln Asp Ile Val Ser Met Gly Ala Asn Ala Ala Lys Asp Ile Ile Thr		
		50 55	60
40	Glu Glu Asp Lys Lys Asn Ile Gly Met Val Ile Val Ala Thr Glu Ser		
		65 70	75 80
	Ala Ile Asp Asn Ala Lys Ala Ala Ala Val Gln Ile His His Leu Leu		
		85 90	95
45	Gly Ile Gln Pro Phe Ala Arg Cys Phe Glu Met Lys Glu Ala Cys Tyr		
		100 105	110
	Ala Ala Thr Pro Ala Ile Gln Leu Ala Lys Asp Tyr Leu Ala Gln Arg		
		115 120	125
50	Pro Asn Glu Lys Val Leu Val Ile Ala Ser Asp Thr Ala Arg Tyr Gly		
		130 135	140
	Ile His Ser Gly Gly Glu Pro Thr Gln Gly Ala Gly Ala Val Ala Met		
		145 150	155 160
55	Met Ile Ser His Asp Pro Ser Ile Leu Lys Leu Asn Asp Asp Ala Val		
		165 170	175
60	Ala Tyr Thr Glu Asp Val Tyr Asp Phe Trp Arg Pro Thr Gly His Gln		
		180 185	190

Tyr Pro Leu Val Ala Gly Ala Leu Ser Lys Asp Ala Tyr Ile Lys Ser
 195 200 205
 5 Phe Gln Glu Ser Trp Asn Glu Tyr Ala Arg Arg His Asn Lys Thr Leu
 210 215 220
 Ala Asp Phe Ala Ser Leu Cys Phe His Val Pro Phe Thr Lys Met Gly
 225 230 235 240
 10 Gln Lys Ala Leu Asp Ser Ile Ile Asn His Ala Asp Glu Thr Thr Gln
 245 250 255
 Asp Arg Leu Asn Ser Ser Tyr Gln Asp Ala Val Asp Tyr Asn Arg Tyr
 15 260 265 270
 Val Gly Asn Ile Tyr Thr Gly Ser Leu Tyr Leu Ser Leu Ile Ser Leu
 275 280 285
 20 Leu Glu Thr Arg Asp Leu Lys Gly Gly Gln Thr Ile Gly Leu Phe Ser
 290 295 300
 Tyr Gly Ser Gly Ser Val Gly Glu Phe Phe Ser Gly Thr Leu Val Asp
 305 310 315 320
 25 Gly Phe Lys Glu Gln Leu Asp Val Glu Arg His Lys Ser Leu Leu Asn
 325 330 335
 Asn Arg Ile Glu Val Ser Val Asp Glu Tyr Glu His Phe Phe Lys Arg
 30 340 345 350
 Phe Asp Gln Leu Glu Leu Asn His Glu Leu Glu Lys Ser Asn Ala Asp
 355 360 365
 35 Arg Asp Ile Phe Tyr Leu Lys Ser Ile Asp Asn Asn Ile Arg Glu Tyr
 370 375 380
 His Ile Ala Glu
 40 385
 <210> 80
 <211> 388
 <212> PRT
 45 <213> 金黄色葡萄球菌 (Staphylococcus aureus)
 <400> 80
 Met Thr Ile Gly Ile Asp Lys Ile Asn Phe Tyr Val Pro Lys Tyr Tyr
 1 5 10 15
 50 Val Asp Met Ala Lys Leu Ala Glu Ala Arg Gln Val Asp Pro Asn Lys
 20 25 30
 Phe Leu Ile Gly Ile Gly Gln Thr Glu Met Ala Val Ser Pro Val Asn
 35 40 45
 55 Gln Asp Ile Val Ser Met Gly Ala Asn Ala Ala Lys Asp Ile Ile Thr
 50 55 60
 Asp Glu Asp Lys Lys Lys Ile Gly Met Val Ile Val Ala Thr Glu Ser
 60 65 70 75 80

Ala Val Asp Ala Ala Lys Ala Ala Ala Val Gln Ile His Asn Leu Leu
 85 90 95
 5 Gly Ile Gln Pro Phe Ala Arg Cys Phe Glu Met Lys Glu Ala Cys Tyr
 100 105 110
 Ala Ala Thr Pro Ala Ile Gln Leu Ala Lys Asp Tyr Leu Ala Thr Arg
 115 120 125
 10 Pro Asn Glu Lys Val Leu Val Ile Ala Thr Asp Thr Ala Arg Tyr Gly
 130 135 140
 Leu Asn Ser Gly Gly Glu Pro Thr Gln Gly Ala Gly Ala Val Ala Met
 145 150 155 160
 Val Ile Ala His Asn Pro Ser Ile Leu Ala Leu Asn Glu Asp Ala Val
 165 170 175
 20 Ala Tyr Thr Glu Asp Val Tyr Asp Phe Trp Arg Pro Thr Gly His Lys
 180 185 190
 Tyr Pro Leu Val Asp Gly Ala Leu Ser Lys Asp Ala Tyr Ile Arg Ser
 195 200 205
 25 Phe Gln Gln Ser Trp Asn Glu Tyr Ala Lys Arg Gln Gly Lys Ser Leu
 210 215 220
 Ala Asp Phe Ala Ser Leu Cys Phe His Val Pro Phe Thr Lys Met Gly
 225 230 235 240
 30 Lys Lys Ala Leu Glu Ser Ile Ile Asp Asn Ala Asp Glu Thr Thr Gln
 245 250 255
 Glu Arg Leu Arg Ser Gly Tyr Glu Asp Ala Val Asp Tyr Asn Arg Tyr
 260 265 270
 35 Val Gly Asn Ile Tyr Thr Gly Ser Leu Tyr Leu Ser Leu Ile Ser Leu
 275 280 285
 40 Leu Glu Asn Arg Asp Leu Gln Ala Gly Glu Thr Ile Gly Leu Phe Ser
 290 295 300
 Tyr Gly Ser Gly Ser Val Val Glu Phe Tyr Ser Ala Thr Leu Val Val
 305 310 315 320
 45 Gly Tyr Lys Asp His Leu Asp Gln Ala Ala His Lys Ala Leu Leu Asn
 325 330 335
 Asn Arg Thr Glu Val Ser Val Asp Ala Tyr Glu Thr Phe Phe Lys Arg
 340 345 350
 50 Phe Asp Asp Val Glu Phe Asp Glu Glu Gln Asp Ala Val His Glu Asp
 355 360 365
 55 Arg His Ile Phe Tyr Leu Ser Asn Ile Glu Asn Asn Val Arg Glu Tyr
 370 375 380
 His Arg Pro Glu
 385
 60

```

<210> 81
<211> 389
<212> PRT
<213> 肉葡萄球菌 (Staphylococcus carnosus)
5 <400> 81
Met Thr Ile Gly Ile Asp Gln Leu Asn Phe Tyr Ile Pro Asn Phe Tyr
1 5 10 15

Val Asp Met Ala Glu Leu Ala Glu Ala Arg Gly Val Asp Pro Asn Lys
10 20 25 30

Phe Leu Ile Gly Ile Gly Gln Ser Gln Met Ala Val Ser Pro Val Ser
35 40 45

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

Glu Gln Asp Lys Lys Asp Ile Thr Met Val Ile Val Ala Thr Glu Ser
65 70 75 80

20 Ala Ile Asp Ser Ala Lys Ala Ser Ala Val Gln Ile His His Leu Leu
85 90 95

Gly Ile Gln Pro Phe Ala Arg Cys Phe Glu Met Lys Glu Ala Cys Tyr
100 105 110

25 Ala Ala Thr Pro Ala Ile Gln Leu Ala Lys Asp Tyr Leu Val Pro Arg
115 120 125

30 Pro Lys Glu Lys Val Leu Val Ile Ala Ser Asp Thr Ala Arg Tyr Gly
130 135 140

Leu Asn Ser Gly Gly Glu Pro Thr Gln Gly Ala Gly Ala Val Ala Met
145 150 155 160

35 Val Ile Ser His Asn Pro Ser Ile Leu Glu Leu His Asp Asp Ser Val
165 170 175

Ala Tyr Thr Glu Asp Val Tyr Asp Phe Trp Arg Pro Ser Gly Glu Ile
180 185 190

Tyr Pro Leu Val Ala Gly Lys Leu Ser Lys Asp Ala Tyr Ile Lys Ser
195 200 205

45 Phe Gln Glu Ser Trp Asn Glu Tyr Ala Lys Arg His His Lys Ser Leu
210 215 220

Ser Asp Phe Ala Ala Leu Cys Phe His Val Pro Phe Thr Lys Met Gly
225 230 235 240

50 Gln Lys Ala Leu Asp Ser Ile Leu Thr Asp Ser Ala Ser Glu Asp Thr
245 250 255

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

55 Tyr Val Gly Asn Val Tyr Thr Gly Ser Leu Tyr Leu Ser Leu Ile Ser
275 280 285

Leu Leu Glu Asn His Lys Leu Asn Gly Gly Asp Asn Ile Gly Leu Phe
290 295 300
60

```

Ser Tyr Gly Ser Gly Ser Val Gly Glu Phe Phe Ser Ala Thr Leu Val
 305 310 315 320

5 Asp Asn Tyr Gln Asp His Leu Asp Val Lys Ala His Lys Ala Met Leu
 325 330 335

Asp Asn Arg Lys Ala Leu Ser Val Glu Glu Tyr Glu Lys Phe Phe Asn
 340 345 350

10 Arg Phe Asp Asn Leu Glu Phe Asp Thr Glu Thr Glu Leu Glu Val Glu
 355 360 365

15 Pro Lys Gly Asn Phe Tyr Leu Lys Glu Ile Ser Asp Asn Ile Arg Tyr
 370 375 380

Tyr Asp Thr Val Lys
 385

20 <210> 82
 <211> 389
 <212> PRT
 <213> 链霉菌属菌种(*Streptomyces* sp.)
 <400> 82

25 Met Ser Ile Ser Ile Gly Ile His Asp Leu Ser Phe Ala Thr Thr Glu
 1 5 10 15
 Phe Val Leu Pro His Thr Ala Leu Ala Glu Tyr Asn Gly Thr Glu Ile
 20 25 30

30 Gly Lys Tyr His Val Gly Ile Gly Gln Gln Ser Met Ser Val Pro Ala
 35 40 45

Ala Asp Glu Asp Ile Val Thr Met Ala Ala Thr Ala Ala Arg Pro Ile
 50 55 60

35 Ile Glu Arg Asn Gly Lys Ser Arg Ile Arg Thr Val Val Phe Ala Thr
 65 70 75 80

Glu Ser Ser Ile Asp Gln Ala Lys Ala Gly Gly Val Tyr Val His Ser
 85 90 95

40 Leu Leu Gly Leu Glu Ser Ala Cys Arg Val Val Glu Leu Lys Gln Ala
 100 105 110

45 Cys Tyr Gly Ala Thr Ala Ala Leu Gln Phe Ala Ile Gly Leu Val Arg
 115 120 125

Arg Asp Pro Ala Gln Gln Val Leu Val Ile Ala Ser Asp Val Ser Lys
 130 135 140

50 Tyr Glu Leu Asp Ser Pro Gly Glu Ala Thr Gln Gly Ala Ala Ala Val
 145 150 155 160

Ala Met Leu Val Gly Ala Asp Pro Ala Leu Leu Arg Ile Glu Glu Pro
 165 170 175

55 Ser Gly Leu Phe Thr Ala Asp Val Met Asp Phe Trp Arg Pro Asn Tyr
 180 185 190

60 Leu Thr Thr Ala Leu Val Asp Gly Gln Glu Ser Ile Asn Ala Tyr Leu

	195	200	205
	Gln Ala Val Glu Gly Ala Trp Lys Asp Tyr Ala Glu Gln Asp Gly Arg		
	210	215	220
5	Ser Leu Glu Glu Phe Ala Ala Phe Val Tyr His Gln Pro Phe Thr Lys		
	225	230	235 240
10	Met Ala Tyr Lys Ala His Arg His Leu Leu Asn Phe Asn Gly Tyr Asp		
		245	250 255
	Thr Asp Lys Asp Ala Ile Glu Gly Ala Leu Gly Gln Thr Thr Ala Tyr		
		260	265 270
15	Asn Asn Val Ile Gly Asn Ser Tyr Thr Ala Ser Val Tyr Leu Gly Leu		
		275	280 285
20	Ala Ala Leu Leu Asp Gln Ala Asp Asp Leu Thr Gly Arg Ser Ile Gly		
		290	295 300
	Phe Leu Ser Tyr Gly Ser Gly Ser Val Ala Glu Phe Phe Ser Gly Thr		
		305	310 315 320
25	Val Val Ala Gly Tyr Arg Glu Arg Leu Arg Thr Glu Ala Asn Gln Glu		
		325	330 335
	Ala Ile Ala Arg Arg Lys Ser Val Asp Tyr Ala Thr Tyr Arg Glu Leu		
		340	345 350
30	His Glu Tyr Thr Leu Pro Ser Asp Gly Gly Asp His Ala Thr Pro Val		
		355	360 365
35	Gln Thr Thr Gly Pro Phe Arg Leu Ala Gly Ile Asn Asp His Lys Arg		
		370	375 380
	Ile Tyr Glu Ala Arg		
	385		
40	<210> 83		
	<211> 389		
	<212> PRT		
	<213> 灰抱链霉菌 (Streptomyces griseosporus)		
	<400> 83		
45	Met Pro Leu Ala Ile Gly Ile His Asp Leu Ser Phe Ala Thr Gly Glu		
	1	5	10 15
	Phe Val Leu Pro His Thr Ala Leu Ala Ala His Asn Gly Thr Glu Ile		
		20	25 30
50	Gly Lys Tyr His Ala Gly Ile Gly Gln Glu Ser Met Ser Val Pro Ala		
		35	40 45
	Ala Asp Glu Asp Ile Val Thr Leu Ala Ala Thr Ala Ala Ala Pro Ile		
		50	55 60
55	Val Ala Arg His Gly Ser Asp Arg Ile Arg Thr Val Val Leu Ala Thr		
		65	70 75 80
60	Glu Ser Ser Ile Asp Gln Ala Lys Ser Ala Gly Val Tyr Val His Ser		
		85	90 95

```

      Leu Leu Gly Leu Pro Ser Ala Thr Arg Val Val Glu Leu Lys Gln Ala
      100                                105                                110
5    Cys Tyr Gly Ala Thr Ala Gly Leu Gln Phe Ala Ile Gly Leu Val Gln
      115                                120                                125

      Arg Asp Pro Ala Gln Gln Val Leu Val Ile Ala Ser Asp Val Ser Lys
      130                                135                                140
10   Tyr Asp Leu Asp Ser Pro Gly Glu Ala Thr Gln Gly Ala Ala Ala Val
      145                                150                                155                                160

      Ala Met Leu Val Gly Ala Asp Pro Gly Leu Val Arg Ile Glu Asp Pro
      165                                170                                175
15   Ser Gly Leu Phe Thr Val Asp Val Met Asp Phe Trp Arg Pro Asn Tyr
      180                                185                                190

      Arg Thr Thr Ala Leu Val Asp Gly Gln Glu Ser Ile Gly Ala Tyr Leu
      195                                200                                205
20   Gln Ala Val Glu Gly Ala Trp Lys Asp Tyr Ser Glu Arg Gly Gly His
      210                                215                                220
25   Ser Leu Glu Gln Phe Ala Ala Phe Cys Tyr His Gln Pro Phe Thr Lys
      225                                230                                235                                240

      Met Ala His Lys Ala His Arg His Leu Leu Asn Tyr Cys Ser His Asp
      245                                250                                255
30   Ile His His Asp Asp Val Thr Arg Ala Val Gly Arg Thr Thr Ala Tyr
      260                                265                                270

      Asn Arg Val Ile Gly Asn Ser Tyr Thr Ala Ser Val Tyr Leu Gly Leu
      275                                280                                285
35   Ala Ala Leu Leu Asp Gln Ala Asp Asp Leu Thr Gly Glu Arg Ile Gly
      290                                295                                300
40   Phe Leu Ser Tyr Gly Ser Gly Ser Val Ala Glu Phe Phe Gly Gly Ile
      305                                310                                315                                320

      Val Val Ala Gly Tyr Arg Asp Arg Leu Arg Thr Ala Ala Asn Ile Glu
      325                                330                                335
45   Ala Val Ser Arg Arg Arg Pro Ile Asp Tyr Ala Gly Tyr Arg Glu Leu
      340                                345                                350
50   His Glu Trp Ala Phe Pro Ala Arg Arg Gly Ala His Ser Thr Pro Gln
      355                                360                                365

      Gln Thr Thr Gly Pro Phe Arg Leu Ser Gly Ile Ser Gly His Lys Arg
      370                                375                                380
55   Leu Tyr Arg Ala Cys
      385

```

```

<210> 84
<211> 407
<212> PRT
<213> 布氏疏螺旋体 (Borrelia burgdorferi)
5 <400> 84
Met Arg Ile Gly Ile Ser Asp Ile Arg Ile Phe Leu Pro Leu Asn Tyr
1 5 10 15
Leu Asp Phe Ser Val Leu Leu Glu Asn Pro Leu Tyr Phe Ser Asn Glu
10 20 25 30
Val Phe Phe Lys Lys Ile Asn Arg Ala Ile Asp Ala Thr Leu Gln Lys
35 40 45
15 Gly Phe Arg Phe Thr Ser Pro Asn Glu Asp Ser Val Thr Met Ala Ser
50 55 60
Ser Ala Val Lys Leu Ile Phe Asp Asn Asn Asn Leu Asp Leu Ser Lys
65 70 75 80
20 Ile Arg Ile Leu Leu Gly Gly Thr Glu Thr Gly Val Asp His Ser Lys
85 90 95
Ala Ile Ser Ser Tyr Val Phe Gly Ala Leu Lys Gln Ser Gly Ile Cys
100 105 110
25 Leu Gly Asn Asn Phe Leu Thr Phe Gln Val Gln His Ala Cys Ala Gly
115 120 125
Ala Ala Met Ser Leu His Thr Val Ala Ser Val Leu Ser His Ser Asn
130 135 140
Asn Ser Glu Tyr Gly Ile Val Phe Ser Ser Asp Ile Ala His Tyr Ser
145 150 155 160
35 Asn Leu Thr Thr Ala Glu Ile Thr Gln Gly Ala Gly Ala Thr Ala Ile
165 170 175
Leu Ile Glu Lys Asn Pro Lys Ile Leu Ser Ile Asn Leu Ser Glu Phe
180 185 190
40 Gly Val Tyr Thr Asp Asp Val Asp Asp Phe Phe Arg Pro Phe Gly Ser
195 200 205
Val Glu Ala Lys Val Arg Gly Gln Tyr Ser Val Glu Cys Tyr Asn Asn
210 215 220
45 Ala Asn Glu Asn Ala Leu Arg Asp Phe Ala Phe Lys Lys Gln Leu Ser
225 230 235 240
50 Met Lys Asp Leu Phe Ser Asn Tyr Arg Phe Val Leu His Val Pro Phe
245 250 255
Ala Lys Met Pro Ile Asp Ser Met His Tyr Ile Leu Lys Lys Tyr Tyr
260 265 270
55 Ser Asp Asp Glu Ser Val Arg Asn Ala Tyr Leu Glu Ser Ile Asp Phe
275 280 285
Tyr Asp Gly Val Glu Ala Ala Met Glu Val Gly Asn Leu Tyr Thr Gly
60 290 295 300

```

```

    Ser Ile Phe Leu Ser Leu Ala Phe Tyr Leu Lys Arg Val Phe Ser Lys
    305                      310                      315                      320
5   Lys Asp Ile Thr Gly Glu Lys Ile Leu Phe Cys Ser Tyr Gly Ser Gly
    325                      330                      335

    Asn Ile Met Ile Ile Tyr Glu Leu Thr Ile Glu Lys Ser Ala Phe Asp
    10                      340                      345                      350

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

    Ala Asn Phe Glu Glu Tyr Lys Asp Phe Phe Gln Asn Lys Ile Ile Pro
    15                      370                      375                      380

    Gly Glu Ser Arg Gly Phe Tyr Leu Lys Glu Leu Arg Asn Asp Gly Tyr
    20                      385                      390                      395                      400

    Arg Val Tyr Gly Tyr Arg Ala
    405

    <210> 85
    25 <211> 317
    <212> PRT
    <213> 肺炎链球菌 (Streptococcus pneumoniae)
    <400> 85
    30 Met Asp Arg Glu Pro Val Thr Val Arg Ser Tyr Ala Asn Ile Ala Ile
    1                      5                      10                      15

    Ile Lys Tyr Trp Gly Lys Lys Lys Glu Lys Glu Met Val Pro Ala Thr
    20                      25                      30

    35 Ser Ser Ile Ser Leu Thr Leu Glu Asn Met Tyr Thr Glu Thr Thr Leu
    35                      40                      45

    Ser Pro Leu Pro Ala Asn Val Thr Ala Asp Glu Phe Tyr Ile Asn Gly
    40                      50                      55                      60

    Gln Leu Gln Asn Glu Val Glu His Ala Lys Met Ser Lys Ile Ile Asp
    65                      70                      75                      80

    Arg Tyr Arg Pro Ala Gly Glu Gly Phe Val Arg Ile Asp Thr Gln Asn
    45                      85                      90                      95
    Asn Met Pro Thr Ala Ala Gly Leu Ser Ser Ser Ser Ser Gly Leu Ser
    100                      105                      110

    Ala Leu Val Lys Ala Cys Asn Ala Tyr Phe Lys Leu Gly Leu Asp Arg
    50                      115                      120                      125

    Ser Gln Leu Ala Gln Glu Ala Lys Phe Ala Ser Gly Ser Ser Ser Arg
    130                      135                      140

    55 Ser Phe Tyr Gly Pro Leu Gly Ala Trp Asp Lys Asp Ser Gly Glu Ile
    145                      150                      155                      160

    Tyr Pro Val Glu Thr Asp Leu Lys Leu Ala Met Ile Met Leu Val Leu
    165                      170                      175
60

```

	Glu Asp Lys Lys Lys Pro Ile Ser Ser Arg Asp Gly Met Lys Leu Cys	
	180	185 190
5	Val Glu Thr Ser Thr Thr Phe Asp Asp Trp Val Arg Gln Ser Glu Lys	
	195	200 205
	Asp Tyr Gln Asp Met Leu Ile Tyr Leu Lys Glu Asn Asp Phe Ala Lys	
	210	215 220
10	Ile Gly Glu Leu Thr Glu Lys Asn Ala Leu Ala Met His Ala Thr Thr	
	225	230 235 240
	Lys Thr Ala Ser Pro Ala Phe Ser Tyr Leu Thr Asp Ala Ser Tyr Glu	
	245	250 255
15	Ala Met Ala Phe Val Arg Gln Leu Arg Glu Lys Gly Glu Ala Cys Tyr	
	260	265 270
20	Phe Thr Met Asp Ala Gly Pro Asn Val Lys Val Phe Cys Gln Glu Lys	
	275	280 285
	Asp Leu Glu His Leu Ser Glu Ile Phe Gly Gln Arg Tyr Arg Leu Ile	
	290	295 300
25	Val Ser Lys Thr Lys Asp Leu Ser Gln Asp Asp Cys Cys	
	305	310 315
	<210> 86	
	<211> 314	
30	<212> PRT	
	<213> 酿脓链球菌 (Streptococcus pyogenes)	
	<400> 86	
	Met Asp Pro Asn Val Ile Thr Val Thr Ser Tyr Ala Asn Ile Ala Ile	
	1	5 10 15
35	Ile Lys Tyr Trp Gly Lys Glu Asn Gln Ala Lys Met Ile Pro Ser Thr	
	20	25 30
40	Ser Ser Ile Ser Leu Thr Leu Glu Asn Met Phe Thr Thr Thr Ser Val	
	35	40 45
	Ser Phe Leu Pro Asp Thr Ala Thr Ser Asp Gln Phe Tyr Ile Asn Gly	
	50	55 60
45	Val Leu Gln Asn Asp Glu Glu His Thr Lys Ile Ser Thr Ile Ile Asp	
	65	70 75 80
	Gln Phe Arg Gln Pro Gly Gln Ala Phe Val Lys Met Glu Thr Gln Asn	
	85	90 95
50	Asn Met Pro Thr Ala Ala Gly Leu Ser Ser Ser Ser Ser Gly Leu Ser	
	100	105 110
55	Ala Leu Val Lys Ala Cys Asp Gln Leu Phe Asp Thr Gln Leu Asp Gln	
	115	120 125
	Lys Ala Leu Ala Gln Lys Ala Lys Phe Ala Ser Gly Ser Ser Ser Arg	
	130	135 140
60	Ser Phe Phe Gly Pro Val Ala Ala Trp Asp Lys Asp Ser Gly Ala Ile	

	145				150					155					160	
	Tyr	Lys	Val	Glu	Thr	Asp	Leu	Lys	Met	Ala	Met	Ile	Met	Leu	Val	Leu
					165					170					175	
5	Asn	Ala	Ala	Lys	Lys	Pro	Ile	Ser	Ser	Arg	Glu	Gly	Met	Lys	Leu	Cys
				180					185					190		
	Arg	Asp	Thr	Ser	Thr	Thr	Phe	Asp	Glu	Trp	Val	Glu	Gln	Ser	Ala	Ile
			195					200					205			
10	Asp	Tyr	Gln	His	Met	Leu	Thr	Tyr	Leu	Lys	Thr	Asn	Asn	Phe	Glu	Lys
		210					215					220				
	Val	Gly	Gln	Leu	Thr	Glu	Ala	Asn	Ala	Leu	Ala	Met	His	Ala	Thr	Thr
15		225				230					235					240
	Lys	Thr	Ala	Asn	Pro	Pro	Phe	Ser	Tyr	Leu	Thr	Lys	Glu	Ser	Tyr	Gln
				245						250					255	
20	Ala	Met	Glu	Ala	Val	Lys	Glu	Leu	Arg	Gln	Glu	Gly	Phe	Ala	Cys	Tyr
				260					265					270		
	Phe	Thr	Met	Asp	Ala	Gly	Pro	Asn	Val	Lys	Val	Leu	Cys	Leu	Glu	Lys
			275					280					285			
25	Asp	Leu	Ala	Gln	Leu	Ala	Glu	Arg	Leu	Gly	Lys	Asn	Tyr	Arg	Ile	Ile
		290					295					300				
	Val	Ser	Lys	Thr	Lys	Asp	Leu	Pro	Asp	Val						
30		305				310										
	<210>	87														
	<211>	331														
	<212>	PRT														
35	<213>	粪肠球菌 (Enterococcus faecalis)														
	<400>	87														
	Met	Leu	Ser	Gly	Lys	Ala	Arg	Ala	His	Thr	Asn	Ile	Ala	Leu	Ile	Lys
	1				5					10					15	
40	Tyr	Trp	Gly	Lys	Ala	Asn	Glu	Glu	Tyr	Ile	Leu	Pro	Met	Asn	Ser	Ser
				20					25					30		
	Leu	Ser	Leu	Thr	Leu	Asp	Ala	Phe	Tyr	Thr	Glu	Thr	Thr	Val	Thr	Phe
			35					40					45			
45	Asp	Ala	His	Tyr	Ser	Glu	Asp	Val	Phe	Ile	Leu	Asn	Gly	Ile	Leu	Gln
		50					55					60				
	Asn	Glu	Lys	Gln	Thr	Lys	Lys	Val	Lys	Glu	Phe	Leu	Asn	Leu	Val	Arg
50		65				70					75					80
	Gln	Gln	Ala	Asp	Cys	Thr	Trp	Phe	Ala	Lys	Val	Glu	Ser	Gln	Asn	Phe
				85						90					95	
	Val	Pro	Thr	Ala	Ala	Gly	Leu	Ala	Ser	Ser	Ala	Ser	Gly	Leu	Ala	Ala
55				100					105					110		
	Leu	Ala	Gly	Ala	Cys	Asn	Val	Ala	Leu	Gly	Leu	Asn	Leu	Ser	Ala	Lys
			115					120					125			
60	Asp	Leu	Ser	Arg	Leu	Ala	Arg	Arg	Gly	Ser	Gly	Ser	Ala	Cys	Arg	Ser

	130		135		140														
	Ile	Phe	Gly	Gly	Phe	Ala	Gln	Trp	Asn	Lys	Gly	His	Ser	Asp	Glu	Thr			
	145					150					155				160				
5	Ser	Phe	Ala	Glu	Asn	Ile	Pro	Ala	Asn	Asn	Trp	Glu	Asn	Glu	Leu	Ala			
					165					170					175				
	Met	Leu	Phe	Ile	Leu	Ile	Asn	Asp	Gly	Glu	Lys	Asp	Val	Ser	Ser	Arg			
10				180					185					190					
	Asp	Gly	Met	Lys	Arg	Thr	Val	Glu	Thr	Ser	Ser	Phe	Tyr	Gln	Gly	Trp			
			195					200					205						
15	Leu	Asp	Asn	Val	Glu	Lys	Asp	Leu	Ser	Gln	Val	His	Glu	Ala	Ile	Lys			
	210						215					220							
	Thr	Lys	Asp	Phe	Pro	Arg	Leu	Gly	Glu	Ile	Ile	Glu	Ala	Asn	Gly	Leu			
	225					230					235					240			
20	Arg	Met	His	Gly	Thr	Thr	Leu	Gly	Ala	Val	Pro	Pro	Phe	Thr	Tyr	Trp			
					245					250					255				
	Ser	Pro	Gly	Ser	Leu	Gln	Ala	Met	Ala	Leu	Val	Arg	Gln	Ala	Arg	Ala			
25				260					265					270					
	Lys	Gly	Ile	Pro	Cys	Tyr	Phe	Thr	Met	Asp	Ala	Gly	Pro	Asn	Val	Lys			
			275					280					285						
30	Val	Leu	Val	Glu	Lys	Lys	Asn	Leu	Glu	Ala	Leu	Lys	Thr	Phe	Leu	Ser			
	290						295					300							
	Glu	His	Phe	Ser	Lys	Glu	Gln	Leu	Val	Pro	Ala	Phe	Ala	Gly	Pro	Gly			
	305					310					315					320			
35																			
	Ile	Glu	Leu	Phe	Glu	Thr	Lys	Gly	Met	Asp	Lys								
					325					330									
40	<210>	88																	
	<211>	325																	
	<212>	PRT																	
	<213>	屎肠球菌 (Enterococcus faecium)																	
	<400>	88																	
45	Met	Phe	Lys	Gly	Lys	Ala	Arg	Ala	Tyr	Thr	Asn	Ile	Ala	Leu	Ile	Lys			
	1				5					10				15					
	Tyr	Trp	Gly	Lys	Lys	Asn	Glu	Glu	Leu	Ile	Leu	Pro	Met	Asn	Asn	Ser			
				20					25					30					
50	Leu	Ser	Leu	Thr	Leu	Asp	Ala	Phe	Tyr	Thr	Glu	Thr	Glu	Val	Ile	Phe			
			35					40					45						
	Ser	Asp	Ser	Tyr	Met	Val	Asp	Glu	Phe	Tyr	Leu	Asp	Gly	Thr	Leu	Gln			
55		50					55					60							
	Asp	Glu	Lys	Ala	Thr	Lys	Lys	Val	Ser	Gln	Phe	Leu	Asp	Leu	Phe	Arg			
	65					70					75					80			
60	Lys	Glu	Ala	Gly	Leu	Ser	Leu	Lys	Ala	Ser	Val	Ile	Ser	Gln	Asn	Phe			

	85							90							95						
	Val	Pro	Thr	Ala	Ala	Gly	Leu	Ala	Ser	Ser	Ala	Ser	Gly	Leu	Ala	Ala					
				100					105					110							
5	Leu	Ala	Gly	Ala	Cys	Asn	Thr	Ala	Leu	Lys	Leu	Gly	Leu	Asp	Asp	Leu					
			115					120					125								
	Ser	Leu	Ser	Arg	Phe	Ala	Arg	Arg	Gly	Ser	Gly	Ser	Ala	Cys	Arg	Ser					
10		130					135					140									
	Ile	Phe	Gly	Gly	Phe	Val	Glu	Trp	Glu	Lys	Gly	His	Asp	Asp	Leu	Ser					
	145					150					155				160						
15	Ser	Tyr	Ala	Lys	Pro	Val	Pro	Ser	Asp	Ser	Phe	Glu	Asp	Asp	Leu	Ala					
					165				170						175						
	Met	Val	Phe	Val	Leu	Ile	Asn	Asp	Gln	Lys	Lys	Glu	Val	Ser	Ser	Arg					
				180					185					190							
20	Asn	Gly	Met	Arg	Arg	Thr	Val	Glu	Thr	Ser	Asn	Phe	Tyr	Gln	Gly	Trp					
			195					200					205								
	Leu	Asp	Ser	Val	Glu	Gly	Asp	Leu	Tyr	Gln	Leu	Lys	Gln	Ala	Ile	Lys					
	210						215					220									
25	Thr	Lys	Asp	Phe	Gln	Leu	Leu	Gly	Glu	Thr	Met	Glu	Arg	Asn	Gly	Leu					
	225					230					235					240					
	Lys	Met	His	Gly	Thr	Thr	Leu	Ala	Ala	Gln	Pro	Pro	Phe	Thr	Tyr	Trp					
30					245					250					255						
	Ser	Pro	Asn	Ser	Leu	Lys	Ala	Met	Asp	Ala	Val	Arg	Gln	Leu	Arg	Lys					
				260					265					270							
35	Gln	Gly	Ile	Pro	Cys	Tyr	Phe	Thr	Met	Asp	Ala	Gly	Pro	Asn	Val	Lys					
			275					280					285								
	Val	Leu	Val	Glu	Asn	Ser	His	Leu	Ser	Glu	Val	Gln	Glu	Thr	Phe	Thr					
	290					295						300									
40	Lys	Leu	Phe	Ser	Lys	Glu	Gln	Val	Ile	Thr	Ala	His	Ala	Gly	Pro	Gly					
	305					310					315					320					
	Ile	Ala	Ile	Ile	Glu																
45					325																
	<210> 89																				
	<211> 327																				
	<212> PRT																				
50	<213> 溶血葡萄球菌 (Staphylococcus haemolyticus)																				
	<400> 89																				
	Met	Lys	Lys	Ser	Gly	Lys	Ala	Arg	Ala	His	Thr	Asn	Ile	Ala	Leu	Ile					
	1				5					10					15						
55	Lys	Tyr	Trp	Gly	Lys	Ala	Asp	Glu	Ala	Leu	Ile	Ile	Pro	Met	Asn	Asn					
				20				25					30								
	Ser	Leu	Ser	Val	Thr	Leu	Asp	Arg	Phe	Tyr	Thr	Glu	Thr	Arg	Val	Thr					
			35					40					45								
60																					

	Phe	Asp	Glu	Thr	Leu	Thr	Glu	Asp	Gln	Leu	Ile	Leu	Asn	Gly	Glu	Ala
	50						55					60				
	Val	Asn	Ala	Lys	Glu	Ser	Ala	Lys	Ile	Gln	Arg	Tyr	Met	Glu	Met	Ile
	65					70					75					80
5	Arg	Lys	Glu	Ala	Gly	Ile	Ser	His	Glu	Ala	Leu	Ile	Glu	Ser	Glu	Asn
					85					90					95	
	Phe	Val	Pro	Thr	Ala	Ala	Gly	Leu	Ala	Ser	Ser	Ala	Ser	Ala	Tyr	Ala
10				100					105					110		
	Ala	Leu	Ala	Gly	Ala	Cys	Asn	Glu	Ala	Leu	Gln	Leu	Gly	Leu	Ser	Asp
				115				120					125			
15	Lys	Asp	Leu	Ser	Arg	Leu	Ala	Arg	Arg	Gly	Ser	Gly	Ser	Ala	Ser	Arg
	130						135					140				
	Ser	Ile	Tyr	Gly	Gly	Phe	Ala	Glu	Trp	Glu	Lys	Gly	Asn	Asp	Asp	Glu
	145					150					155					160
20	Thr	Ser	Phe	Ala	His	Arg	Val	Glu	Ala	Asp	Gly	Trp	Glu	Asn	Glu	Leu
					165					170					175	
	Ala	Met	Val	Phe	Val	Val	Ile	Asn	Asn	Lys	Ser	Lys	Lys	Val	Ser	Ser
25				180					185					190		
	Arg	Ser	Gly	Met	Ser	Leu	Thr	Arg	Asp	Thr	Ser	Arg	Phe	Tyr	Gln	Tyr
			195					200					205			
30	Trp	Leu	Asp	Asn	Val	Glu	Pro	Asp	Leu	Lys	Glu	Thr	Lys	Glu	Ala	Ile
	210						215					220				
	Ala	Gln	Lys	Asp	Phe	Lys	Arg	Met	Gly	Glu	Val	Ile	Glu	Ala	Asn	Gly
	225					230					235					240
35	Leu	Arg	Met	His	Ala	Thr	Asn	Leu	Gly	Ala	Gln	Pro	Pro	Phe	Thr	Tyr
					245					250					255	
	Leu	Val	Pro	Glu	Ser	Tyr	Asp	Ala	Met	Arg	Ile	Val	His	Glu	Cys	Arg
40				260					265					270		
	Glu	Ala	Gly	Leu	Pro	Cys	Tyr	Phe	Thr	Met	Asp	Ala	Gly	Pro	Asn	Val
			275					280					285			
45	Lys	Val	Leu	Ile	Glu	Lys	Lys	Asn	Gln	Gln	Ala	Ile	Val	Asp	Lys	Phe
	290						295					300				
	Leu	Gln	Glu	Phe	Asp	Gln	Ser	Gln	Ile	Ile	Thr	Ser	Asp	Ile	Thr	Gln
50	305					310					315					320
	Ser	Gly	Val	Glu	Ile	Ile	Lys									
					325											
55	<210>	90														
	<211>	327														
	<212>	PRT														
	<213>	表皮葡萄球菌 (Staphylococcus epidermidis)														
	<400>	90														
60	Met	Val	Lys	Ser	Gly	Lys	Ala	Arg	Ala	His	Thr	Asn	Ile	Ala	Leu	Ile

1	5	10	15
Lys Tyr Trp Gly 20	Lys Ala Asp Glu 25	Thr Tyr Ile Ile Pro Met Asn Asn 30	
5 Ser Leu Ser 35	Val Thr Leu Asp 40	Phe Tyr Thr Glu 45	Thr Lys Val Thr
10 Phe Asp Pro Asp 50	Phe Thr Glu Asp 55	Cys Leu Ile Leu 60	Asn Gly Asn Glu
Val Asn Ala Lys 65	Glu Lys Glu Lys 70	Ile Gln Asn Tyr 75	Met Asn Ile Val 80
15 Arg Asp Leu Ala 85	Gly Asn Arg Leu 90	His Ala Arg Ile 95	Glu Ser Glu Asn
Tyr Val Pro Thr 100	Ala Ala Gly Leu 105	Ala Ser Ser Ala 110	Ser Ala Tyr Ala
20 Ala Leu Ala 115	Ala Ala Cys Asn 120	Glu Ala Leu Ser 125	Asn Leu Ser Asp
25 Thr Asp Leu Ser 130	Arg Leu Ala Arg 135	Arg Gly Ser Gly 140	Ser Ala Ser Arg
Ser Ile Phe Gly Gly 145	Phe Ala Glu Trp 150	Glu Lys Gly His 155	Asp Asp Leu 160
30 Thr Ser Tyr Ala 165	His Gly Ile Asn 170	Ser Asn Gly Trp 175	Glu Lys Asp Leu 175
Ser Met Ile Phe 180	Val Val Ile Asn 185	Asn Gln Ser Lys 190	Lys Val Ser Ser 190
35 Arg Ser Gly 195	Met Ser Leu Thr 200	Arg Asp Thr Ser 205	Phe Tyr Gln Tyr 205
40 Trp Leu Asp His 210	Val Asp Glu Asp 215	Leu Asn Glu Ala 220	Lys Glu Ala Val 220
Lys Asn Gln Asp 225	Phe Gln Arg Leu 230	Gly Glu Val Ile 235	Glu Ala Asn Gly 240
45 Leu Arg Met His 245	Ala Thr Asn Leu 250	Gly Ala Gln Pro 255	Pro Phe Thr Tyr 255
Leu Val Gln Glu 260	Ser Tyr Asp Ala 265	Met Ala Ile Val 270	Glu Gln Cys Arg 270
50 Lys Ala Asn 275	Leu Pro Cys Tyr 280	Phe Thr Met Asp 285	Ala Gly Pro Asn Val 285
Lys Val Leu Val 290	Glu Lys Lys Asn 295	Lys Gln Ala Val 300	Met Glu Gln Phe 300
55 Leu Lys Val Phe 305	Asp Glu Ser Lys 310	Ile Ile Ala Ser 315	Asp Ile Ile Ser 320
60 Ser Gly Val Glu 325	Ile Ile Lys		

```

<210> 91
<211> 327
<212> PRT
5 <213> 金黄色葡萄球菌 (Staphylococcus aureus)
<400> 91
Met Ile Lys Ser Gly Lys Ala Arg Ala His Thr Asn Ile Ala Leu Ile
1 5 10 15
10 Lys Tyr Trp Gly Lys Lys Asp Glu Ala Leu Ile Ile Pro Met Asn Asn
20 25 30
Ser Ile Ser Val Thr Leu Glu Lys Phe Tyr Thr Glu Thr Lys Val Thr
35 40 45
15 Phe Asn Asp Gln Leu Thr Gln Asp Gln Phe Trp Leu Asn Gly Glu Lys
50 55 60
Val Ser Gly Lys Glu Leu Glu Lys Ile Ser Lys Tyr Met Asp Ile Val
65 70 75 80
20 Arg Asn Arg Ala Gly Ile Asp Trp Tyr Ala Glu Ile Glu Ser Asp Asn
85 90 95
Phe Val Pro Thr Ala Ala Gly Leu Ala Ser Ser Ala Ser Ala Tyr Ala
100 105 110
25 Ala Leu Ala Ala Ala Cys Asn Gln Ala Leu Asp Leu Gln Leu Ser Asp
115 120 125
30 Lys Asp Leu Ser Arg Leu Ala Arg Ile Gly Ser Gly Ser Ala Ser Arg
130 135 140
Ser Ile Tyr Gly Gly Phe Ala Glu Trp Glu Lys Gly Tyr Asn Asp Glu
145 150 155 160
35 Thr Ser Tyr Ala Val Pro Leu Glu Ser Asn His Phe Glu Asp Asp Leu
165 170 175
Ala Met Ile Phe Val Val Ile Asn Gln His Ser Lys Lys Val Pro Ser
180 185 190
40 Arg Tyr Gly Met Ser Leu Thr Arg Asn Thr Ser Arg Phe Tyr Gln Tyr
195 200 205
45 Trp Leu Asp His Ile Asp Glu Asp Leu Ala Glu Ala Lys Ala Ala Ile
210 215 220
Gln Asp Lys Asp Phe Lys Arg Leu Gly Glu Val Ile Glu Glu Asn Gly
225 230 235 240
50 Leu Arg Met His Ala Thr Asn Leu Gly Ser Thr Pro Pro Phe Thr Tyr
245 250 255
55 Leu Val Gln Glu Ser Tyr Asp Val Met Ala Leu Val His Glu Cys Arg
260 265 270
Glu Ala Gly Tyr Pro Cys Tyr Phe Thr Met Asp Ala Gly Pro Asn Val
275 280 285
60

```

Lys Ile Leu Val Glu Lys Lys Asn Lys Gln Gln Ile Ile Asp Lys Leu
 290 295 300
 5 Leu Thr Gln Phe Asp Asn Asn Gln Ile Ile Asp Ser Asp Ile Ile Ala
 305 310 315 320
 Thr Gly Ile Glu Ile Ile Glu
 325
 10 <210> 92
 <211> 350
 <212> PRT
 <213> 链霉菌属菌种 (Streptomyces sp.)
 <400> 92
 15 Met Arg Ser Glu His Pro Thr Thr Thr Val Leu Gln Ser Arg Glu Gln
 1 5 10 15
 Gly Ser Ala Ala Gly Ala Thr Ala Val Ala His Pro Asn Ile Ala Leu
 20 20 25 30
 Ile Lys Tyr Trp Gly Lys Arg Asp Glu Arg Leu Ile Leu Pro Cys Thr
 35 40 45
 25 Thr Ser Leu Ser Met Thr Leu Asp Val Phe Pro Thr Thr Thr Glu Val
 50 55 60
 Arg Leu Asp Pro Ala Ala Glu His Asp Thr Ala Ala Leu Asn Gly Glu
 65 70 75 80
 30 Val Ala Thr Gly Glu Thr Leu Arg Arg Ile Ser Ala Phe Leu Ser Leu
 85 90 95
 Val Arg Glu Val Ala Gly Ser Asp Gln Arg Ala Val Val Asp Thr Arg
 100 105 110
 35 Asn Thr Val Pro Thr Gly Ala Gly Leu Ala Ser Ser Ala Ser Gly Phe
 115 120 125
 Ala Ala Leu Ala Val Ala Ala Ala Ala Tyr Gly Leu Glu Leu Asp
 130 135 140
 40 Asp Arg Gly Leu Ser Arg Leu Ala Arg Arg Gly Ser Gly Ser Ala Ser
 145 150 155 160
 45 Arg Ser Ile Phe Gly Gly Phe Ala Val Trp His Ala Gly Pro Asp Gly
 165 170 175
 Thr Ala Thr Glu Ala Asp Leu Gly Ser Tyr Ala Glu Pro Val Pro Ala
 180 185 190
 50 Ala Asp Leu Asp Pro Ala Leu Val Ile Ala Val Val Asn Ala Gly Pro
 195 200 205
 Lys Pro Val Ser Ser Arg Glu Ala Met Arg Arg Thr Val Asp Thr Ser
 210 215 220
 55 Pro Leu Tyr Arg Pro Trp Ala Asp Ser Ser Lys Asp Asp Leu Asp Glu
 225 230 235 240
 60 Met Arg Ser Ala Leu Leu Arg Gly Asp Leu Glu Ala Val Gly Glu Ile
 245 250 255

Ala Glu Arg Asn Ala Leu Gly Met His Ala Thr Met Leu Ala Ala Arg
 260 265 270
 5 Pro Ala Val Arg Tyr Leu Ser Pro Ala Thr Val Thr Val Leu Asp Ser
 275 280 285
 Val Leu Gln Leu Arg Lys Asp Gly Val Leu Ala Tyr Ala Thr Met Asp
 290 295 300
 10 Ala Gly Pro Asn Val Lys Val Leu Cys Arg Arg Ala Asp Ala Glu Arg
 305 310 315 320
 Val Ala Asp Val Val Arg Ala Ala Ala Ser Gly Gly Gln Val Leu Val
 15 325 330 335
 Ala Gly Pro Gly Asp Gly Ala Arg Leu Leu Ser Glu Gly Ala
 340 345 350
 20 <210> 93
 <211> 331
 <212> PRT
 <213> 灰抱链霉菌 (Streptomyces griseosporeus)
 <400> 93
 25 Ala Thr Ala Val Ala Gln Pro Asn Ile Ala Leu Ile Lys Tyr Trp Gly
 1 5 10 15
 Lys Lys Asp Glu His Leu Val Leu Pro Arg Thr Asp Ser Leu Ser Met
 20 25 30
 30 Thr Leu Asp Ile Phe Pro Thr Thr Thr Arg Val Gln Leu Ala Pro Gly
 35 40 45
 Ala Gly Gln Asp Thr Val Ala Phe Asn Gly Glu Pro Ala Thr Gly Glu
 35 50 55 60
 Ala Glu Arg Arg Ile Thr Ala Phe Leu Arg Leu Val Arg Glu Arg Ser
 65 70 75 80
 40 Gly Arg Thr Glu Arg Ala Arg Val Glu Thr Glu Asn Thr Val Pro Thr
 85 90 95
 Gly Ala Gly Leu Ala Ser Ser Ala Ser Gly Phe Ala Ala Leu Ala Val
 100 105 110
 45 Ala Ala Ala Ala Ala Tyr Gly Leu Gly Leu Asp Ala Arg Gly Leu Ser
 115 120 125
 Arg Leu Ala Arg Arg Gly Ser Gly Ser Ala Ser Arg Ser Ile Phe Asp
 50 130 135 140
 Gly Phe Ala Val Trp His Ala Gly His Ala Gly Gly Thr Pro Glu Glu
 145 150 155 160
 55 Ala Asp Leu Gly Ser Tyr Ala Glu Pro Val Pro Ala Val Asp Leu Glu
 165 170 175
 Pro Ala Leu Val Val Ala Val Val Ser Ala Ala Pro Lys Ala Val Ser
 180 185 190
 60

Ser Arg Glu Ala Met Arg Arg Thr Val Asp Thr Ser Pro Leu Tyr Glu
 195 200 205
 Pro Trp Ala Val Ser Ser Arg Ala Asp Leu Ala Asp Ile Gly Ala Ala
 210 215 220
 5 Leu Ala Arg Gly Asn Leu Pro Ala Val Gly Glu Ile Ala Glu Arg Asn
 225 230 235 240
 10 Ala Leu Gly Met His Ala Thr Met Leu Ala Ala Arg Pro Ala Val Arg
 245 250 255
 Tyr Leu Ser Pro Ala Ser Leu Ala Val Leu Asp Gly Val Leu Gln Leu
 260 265 270
 15 Arg Arg Asp Gly Val Pro Ala Tyr Ala Thr Met Asp Ala Gly Pro Asn
 275 280 285
 Val Lys Val Leu Cys Pro Arg Ser Asp Ala Glu Arg Val Ala Glu Ala
 290 295 300
 20 Leu Arg Ala Ala Ala Pro Val Gly Ala Val His Ile Ala Gly Pro Gly
 305 310 315 320
 Arg Gly Ala Arg Leu Val Ala Glu Glu Cys Arg
 325 330
 25 <210> 94
 <211> 312
 <212> PRT
 30 <213> 布氏疏螺旋体(Borrelia burgdorferi)
 <400> 94
 Met Lys Ile Lys Cys Lys Val His Ala Ser Leu Ala Leu Ile Lys Tyr
 1 5 10 15
 35 Trp Gly Lys Lys Asp Val Phe Leu Asn Ile Pro Ala Thr Ser Ser Leu
 20 25 30
 Ala Val Ser Val Asp Lys Phe Tyr Ser Ile Ser Glu Leu Glu Leu Ser
 35 40 45
 40 Asn Arg Asp Glu Ile Ile Leu Asn Ser Lys Pro Val Ile Leu Lys Asn
 50 55 60
 Arg Glu Lys Val Phe Phe Asp Tyr Ala Arg Lys Ile Leu Asn Glu Pro
 65 70 75 80
 45 Asn Val Arg Phe Lys Ile Lys Ser Lys Asn Asn Phe Pro Thr Ala Ala
 85 90 95
 Gly Leu Ala Ser Ser Ser Ser Gly Phe Ala Ser Ile Ala Ala Cys Ile
 100 105 110
 50 Leu Lys Tyr Phe Asn Lys Tyr Ser Cys Asn Ser Ala Ser Asn Leu Ala
 115 120 125
 55 Arg Val Gly Ser Ala Ser Ala Ala Arg Ala Ile Tyr Gly Gly Phe Thr
 130 135 140
 Ile Leu Lys Glu Gly Ser Lys Glu Ser Phe Gln Leu Arg Asp Gln Ser
 145 150 155 160
 60

Tyr Phe Asn Asp Leu Arg Ile Ile Phe Ala Ile Ile Asp Ser Asn Glu
 165 170 175
 5 Lys Glu Leu Ser Ser Arg Ala Ala Met Asn Ile Cys Lys Arg His Lys
 180 185 190
 Phe Tyr Tyr Asp Ala Trp Ile Ala Ser Ser Lys Lys Ile Phe Lys Asp
 195 200 205
 10 Ala Leu Tyr Phe Phe Leu Lys Lys Asp Phe Ile His Phe Gly Ala Thr
 210 215 220
 Ile Val Lys Ser Tyr Gln Asn Met Phe Ala Leu Met Phe Ala Ser Ser
 225 230 235 240
 15 Ile Phe Tyr Phe Lys Asn Ser Thr Ile Asp Leu Ile Arg Tyr Ala Ala
 245 250 255
 Asp Leu Arg Asn Glu Gly Ile Phe Val Phe Glu Thr Met Asp Ala Gly
 260 265 270
 20 Pro Gln Val Lys Phe Leu Cys Leu Glu Glu Asn Leu Asn Thr Ile Leu
 275 280 285
 25 Lys Gly Leu Lys Gln Asn Phe Thr Gly Ile Asp Phe Ile Val Ser Lys
 290 295 300
 Val Gly Cys Asp Leu Glu Trp Ile
 30 305 310
 <210> 95
 <211> 292
 <212> PRT
 35 <213> 肺炎链球菌 (Streptococcus pneumoniae)
 <400> 95
 Met Thr Lys Lys Val Gly Val Gly Gln Ala His Ser Lys Ile Ile Leu
 1 5 10 15
 40 Ile Gly Glu His Ala Val Val Tyr Gly Tyr Pro Ala Ile Ser Leu Pro
 20 25 30
 Leu Leu Glu Val Glu Val Thr Cys Lys Val Val Ser Ala Glu Ser Pro
 35 40 45
 45 Trp Arg Leu Tyr Glu Glu Asp Thr Leu Ser Met Ala Val Tyr Ala Ser
 50 55 60
 Leu Glu Tyr Leu Asp Ile Thr Glu Ala Cys Val Arg Cys Glu Ile Asp
 65 70 75 80
 Ser Ala Ile Pro Glu Lys Arg Gly Met Gly Ser Ser Ala Ala Ile Ser
 85 90 95
 55 Ile Ala Ala Ile Arg Ala Val Phe Asp Tyr Tyr Gln Ala Asp Leu Pro
 100 105 110
 His Asp Val Leu Glu Ile Leu Val Asn Arg Ala Glu Met Ile Ala His
 115 120 125
 60

	Met	Asn	Pro	Ser	Gly	Leu	Asp	Ala	Lys	Thr	Cys	Leu	Ser	Asp	Gln	Pro
	130						135					140				
5	Ile	Arg	Phe	Ile	Lys	Asn	Val	Gly	Phe	Thr	Glu	Leu	Glu	Met	Asp	Leu
	145					150					155					160
	Ser	Ala	Tyr	Leu	Val	Ile	Ala	Asp	Thr	Gly	Val	Tyr	Gly	His	Thr	Arg
				165						170					175	
10	Glu	Ala	Ile	Gln	Val	Val	Gln	Asn	Lys	Gly	Lys	Asp	Ala	Leu	Pro	Phe
				180					185					190		
	Leu	His	Ala	Leu	Gly	Glu	Leu	Thr	Gln	Gln	Ala	Glu	Val	Ala	Ile	Ser
		195						200					205			
15	Gln	Lys	Tyr	Ala	Glu	Gly	Leu	Gly	Leu	Ile	Phe	Ser	Gln	Ala	His	Leu
	210						215					220				
	His	Leu	Lys	Glu	Ile	Gly	Val	Ser	Ser	Pro	Glu	Ala	Asp	Phe	Leu	Val
	225					230					235					240
20	Glu	Thr	Ala	Leu	Ser	Tyr	Gly	Ala	Leu	Gly	Ala	Lys	Met	Ser	Gly	Gly
				245						250					255	
	Gly	Leu	Gly	Gly	Cys	Ile	Ile	Ala	Leu	Val	Thr	Asn	Leu	Thr	His	Ala
25				260					265					270		
	Gln	Glu	Leu	Ala	Glu	Arg	Leu	Glu	Glu	Lys	Gly	Ala	Val	Gln	Thr	Trp
		275						280					285			
30	Ile	Glu	Ser	Leu												
	290															
	<210>	96														
	<211>	292														
35	<212>	PRT														
	<213>	酿脓链球菌 (Streptococcus pyogenes)														
	<400>	96														
	Met	Asn	Glu	Asn	Ile	Gly	Tyr	Gly	Lys	Ala	His	Ser	Lys	Ile	Ile	Leu
	1			5					10					15		
40	Ile	Gly	Glu	His	Ala	Val	Val	Tyr	Gly	Tyr	Pro	Ala	Ile	Ala	Leu	Pro
			20						25					30		
	Leu	Thr	Asp	Ile	Glu	Val	Val	Cys	His	Ile	Phe	Pro	Ala	Asp	Lys	Pro
45			35					40					45			
	Leu	Val	Phe	Asp	Phe	Tyr	Asp	Thr	Leu	Ser	Thr	Ala	Ile	Tyr	Ala	Ala
	50						55					60				
50	Leu	Asp	Tyr	Leu	Gln	Arg	Leu	Gln	Glu	Pro	Ile	Ala	Tyr	Glu	Ile	Val
	65				70					75						80
	Ser	Gln	Val	Pro	Gln	Lys	Arg	Gly	Met	Gly	Ser	Ser	Ala	Ala	Val	Ser
				85					90						95	
55	Ile	Ala	Ala	Ile	Arg	Ala	Val	Phe	Ser	Tyr	Cys	Gln	Glu	Pro	Leu	Ser
			100						105					110		
	Asp	Asp	Leu	Leu	Glu	Ile	Leu	Val	Asn	Lys	Ala	Glu	Ile	Ile	Ala	His
			115					120					125			
60																

	Thr	Asn	Pro	Ser	Gly	Leu	Asp	Ala	Lys	Thr	Cys	Leu	Ser	Asp	His	Ala	
	130						135					140					
5	Ile	Lys	Phe	Ile	Arg	Asn	Ile	Gly	Phe	Glu	Thr	Ile	Glu	Ile	Ala	Leu	
	145					150					155					160	
	Asn	Gly	Tyr	Leu	Ile	Ile	Ala	Asp	Thr	Gly	Ile	His	Gly	His	Thr	Arg	
				165						170					175		
10	Glu	Ala	Val	Asn	Lys	Val	Ala	Gln	Phe	Glu	Glu	Thr	Asn	Leu	Pro	Tyr	
				180					185					190			
	Leu	Ala	Lys	Leu	Gly	Ala	Leu	Thr	Gln	Ala	Leu	Glu	Arg	Ala	Ile	Asn	
			195					200					205				
15	Gln	Lys	Asn	Lys	Val	Ala	Ile	Gly	Gln	Leu	Met	Thr	Gln	Ala	His	Ser	
	210						215					220					
	Ala	Leu	Lys	Ala	Ile	Gly	Val	Ser	Ile	Ser	Lys	Ala	Asp	Gln	Leu	Val	
20	225					230					235					240	
	Glu	Ala	Ala	Leu	Arg	Ala	Gly	Ala	Leu	Gly	Ala	Lys	Met	Thr	Gly	Gly	
				245						250					255		
25	Gly	Leu	Gly	Gly	Cys	Met	Ile	Ala	Leu	Ala	Asp	Thr	Lys	Asp	Met	Ala	
				260					265					270			
	Glu	Lys	Ile	Ser	His	Lys	Leu	Lys	Glu	Glu	Gly	Ala	Val	Asn	Thr	Trp	
		275						280					285				
30	Ile	Gln	Met	Leu													
	290																
	<210>	97															
35	<211>	314															
	<212>	PRT															
	<213>	粪肠球菌 (Enterococcus faecalis)															
	<400>	97															
40	Met	Asn	Ile	Lys	Lys	Gln	Gly	Leu	Gly	Gln	Ala	Thr	Gly	Lys	Ile	Ile	
	1				5					10					15		
	Leu	Met	Gly	Glu	His	Ala	Val	Val	Tyr	Gly	Glu	Pro	Ala	Ile	Ala	Phe	
				20					25					30			
45	Pro	Phe	Gln	Ala	Thr	Glu	Ile	Thr	Ala	Val	Phe	Thr	Pro	Ala	Lys	Thr	
			35					40					45				
	Met	Gln	Ile	Asp	Cys	Ala	Tyr	Phe	Thr	Gly	Leu	Leu	Glu	Asp	Val	Pro	
	50						55					60					
50	Gln	Glu	Leu	Ala	Asn	Ile	Lys	Glu	Val	Val	Gln	Gln	Thr	Leu	His	Phe	
	65					70					75					80	
	Leu	Lys	Glu	Asp	Thr	Phe	Lys	Gly	Thr	Leu	Thr	Leu	Thr	Ser	Thr	Ile	
55					85					90					95		
	Pro	Ala	Glu	Arg	Gly	Met	Gly	Ser	Ser	Ala	Ala	Thr	Ala	Val	Ala	Ile	
				100					105					110			
60	Val	Arg	Ser	Leu	Phe	Asp	Tyr	Phe	Asp	Tyr	Ala	Tyr	Thr	Tyr	Gln	Glu	

	115	120	125
	Leu Phe Glu Leu Val Ser	Leu Ser Glu Lys Ile	Ala His Gly Asn Pro
	130	135	140
5	Ser Gly Ile Asp Ala Ala	Ala Thr Ser Gly Ala	Asp Pro Leu Phe Phe
	145	150	155 160
	Thr Arg Gly Phe Pro	Pro Thr His Phe	Ser Met Asn Leu Ser Asn Ala
10		165	170 175
	Tyr Leu Val Val Ala Asp	Thr Gly Ile Lys Gly	Gln Thr Arg Glu Ala
	180	185	190
15	Val Lys Asp Ile Ala Gln	Leu Ala Gln Asn Asn	Pro Thr Ala Ile Ala
	195	200	205
	Glu Thr Met Lys Gln Leu	Gly Ser Phe Thr Lys	Glu Ala Lys Gln Ala
	210	215	220
20	Ile Leu Gln Asp Asp Lys	Gln Lys Leu Gly Gln	Leu Met Thr Leu Ala
	225	230	235 240
	Gln Glu Gln Leu Gln Gln	Leu Thr Val Ser Asn	Asp Met Leu Asp Arg
	245	250	255
25	Leu Val Ala Leu Ser Leu	Glu His Gly Ala Leu	Gly Ala Lys Leu Thr
	260	265	270
	Gly Gly Gly Arg Gly Gly	Cys Met Ile Ala Leu	Thr Asp Asn Lys Lys
30		275 280	285
	Thr Ala Gln Thr Ile Ala	Gln Thr Leu Glu Glu	Asn Gly Ala Val Ala
	290	295	300
35	Thr Trp Ile Gln Ser Leu	Glu Val Lys Lys	
	305	310	
	<210> 98		
	<211> 314		
40	<212> PRT		
	<213> 屎肠球菌 (Enterococcus faecium)		
	<400> 98		
	Met Ala Asn Tyr Gly Gln	Gly Glu Ser Ser Gly	Lys Ile Ile Leu Met
	1	5	10 15
45	Gly Glu His Ala Val Val	Tyr Gly Glu Pro Ala	Ile Ala Phe Pro Phe
	20	25	30
	Tyr Ala Thr Lys Val Thr	Ala Phe Leu Glu Glu	Leu Asp Ala Met Asp
50		35 40	45
	Asp Gln Leu Val Ser Ser	Tyr Tyr Ser Gly Asn	Leu Ala Glu Ala Pro
	50	55	60
55	His Ala Leu Lys Asn Ile	Lys Lys Leu Phe Ile	His Leu Lys Lys Gln
	65	70	75 80
	His Asp Ile Gln Lys Asn	Leu Gln Leu Thr Ile	Glu Ser Thr Ile Pro
		85 90	95
60			

Ala Glu Arg Gly Met Gly Ser Ser Ala Ala Val Ala Thr Ala Val Thr
 100 105 110
 Arg Ala Phe Tyr Asp Tyr Leu Ala Phe Pro Leu Ser Arg Glu Ile Leu
 115 120 125
 5 Leu Glu Asn Val Gln Leu Ser Glu Lys Ile Ala His Gly Asn Pro Ser
 130 135 140
 Gly Ile Asp Ala Ala Ala Thr Ser Ser Leu Gln Pro Ile Tyr Phe Thr
 10 145 150 155 160
 Lys Gly His Pro Phe Asp Tyr Phe Ser Leu Asn Ile Asp Ala Phe Leu
 165 170 175
 15 Ile Val Ala Asp Thr Gly Ile Lys Gly Gln Thr Arg Glu Ala Val Lys
 180 185 190
 Asp Val Ala His Leu Phe Glu Thr Gln Pro His Glu Thr Gly Gln Met
 195 200 205
 20 Ile Gln Lys Leu Gly Tyr Leu Thr Lys Gln Ala Lys Gln Ala Ile Ile
 210 215 220
 Glu Asn Ser Pro Glu Thr Leu Ala Gln Thr Met Asp Glu Ser Gln Ser
 25 225 230 235 240
 Leu Leu Glu Lys Leu Thr Ile Ser Asn Asp Phe Leu Asn Leu Leu Ile
 245 250 255
 30 Gln Thr Ala Lys Asp Thr Gly Ala Leu Gly Ala Lys Leu Thr Gly Gly
 260 265 270
 Gly Arg Gly Gly Cys Met Ile Ala Leu Ala Gln Thr Lys Thr Lys Ala
 275 280 285
 35 Gln Glu Ile Ser Gln Ala Leu Glu Asp Ala Gly Ala Ala Glu Thr Trp
 290 295 300
 Ile Gln Gly Leu Gly Val His Thr Tyr Val
 40 305 310
 <210> 99
 <211> 307
 <212> PRT
 45 <213> 溶血葡萄球菌 (Staphylococcus haemolyticus)
 <400> 99
 Met Val Gln Arg Gly Tyr Gly Glu Ser Asn Gly Lys Ile Ile Leu Ile
 1 5 10 15
 50 Gly Glu His Ala Val Thr Phe Gly Glu Pro Ala Ile Ala Ile Pro Phe
 20 25 30
 Thr Ser Gly Lys Val Lys Val Leu Ile Glu Ser Leu Glu Lys Gly Asn
 35 40 45
 55 Tyr Ser Ala Ile Gln Ser Asp Val Tyr Asp Gly Pro Leu Tyr Asp Ala
 50 55 60
 Pro Glu His Leu Lys Ser Leu Ile Gly His Phe Val Glu Asn Lys Lys
 60 65 70 75 80

```

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

5  Arg Gly Leu Gly Ser Ser Ala Ala Val Ala Val Ala Phe Ile Arg Ala
      100                      105                      110

Ser Tyr Asp Tyr Leu Gly Leu Pro Leu Thr Asp Lys Glu Leu Leu Glu
      115                      120                      125
10
Asn Ala Asp Trp Ala Glu Arg Ile Ala His Gly Lys Pro Ser Gly Ile
      130                      135                      140

Asp Thr Lys Thr Ile Val Thr Asn Gln Pro Val Trp Tyr Gln Lys Gly
15  145                      150                      155                      160

Glu Val Glu Ile Leu Lys Thr Leu Asp Leu Asp Gly Tyr Met Val Val
      165                      170                      175

20  Ile Asp Thr Gly Val Lys Gly Ser Thr Lys Gln Ala Val Glu Asp Val
      180                      185                      190

His Gln Leu Cys Asp Asn Asp Lys Asn Tyr Met Gln Val Val Lys His
      195                      200                      205
25
Ile Gly Ser Leu Val Tyr Ser Ala Ser Glu Ala Ile Glu His His Ser
      210                      215                      220
Phe Asp Gln Leu Ala Thr Ile Phe Asn Gln Cys Gln Asp Asp Leu Arg
225                      230                      235                      240
30
Thr Leu Thr Val Ser His Asp Lys Ile Glu Met Phe Leu Arg Leu Gly
      245                      250                      255

Glu Glu Asn Gly Ser Val Ala Gly Lys Leu Thr Gly Gly Gly Arg Gly
35  260                      265                      270

Gly Ser Met Leu Ile Leu Ala Lys Glu Leu Gln Thr Ala Lys Asn Ile
      275                      280                      285

40  Val Ala Ala Val Glu Lys Ala Gly Ala Gln His Thr Trp Ile Glu Lys
      290                      295                      300

Leu Gly Gly
305
45
<210> 100
<211> 306
<212> PRT
<213> 表皮葡萄球菌(Staphylococcus epidermidis)
50 <400> 100
Met Thr Arg Gln Gly Tyr Gly Glu Ser Thr Gly Lys Ile Ile Leu Met
1      5      10      15

Gly Glu His Ala Val Thr Phe Gly Gln Pro Ala Ile Ala Ile Pro Phe
55      20      25      30

Asn Ala Gly Lys Ile Lys Val Leu Ile Glu Ser Leu Asp Glu Gly Asn
      35      40      45

60  Tyr Ser Ser Ile Thr Ser Asp Val Tyr Asp Gly Met Leu Tyr Asp Ala

```

	50	55	60	
	Pro Glu His Leu Lys Ser Ile Ile Asn Arg Phe Val Glu Lys Ser Gly			
	65	70	75	80
5	Val Lys Glu Pro Leu Ser Val Lys Ile Gln Thr Asn Leu Pro Pro Ser			
		85	90	95
	Arg Gly Leu Gly Ser Ser Ala Ala Val Ala Val Ala Phe Val Arg Ala			
10		100	105	110
	Ser Tyr Asp Phe Met Asp Gln Pro Leu Asp Asp Lys Thr Leu Ile Lys			
		115	120	125
	Glu Ala Asn Trp Ala Glu Gln Ile Ala His Gly Lys Pro Ser Gly Ile			
15		130	135	140
	Asp Thr Gln Thr Ile Val Ser Asn Lys Pro Val Trp Phe Lys Gln Gly			
	145	150	155	160
20	Gln Ala Glu Thr Leu Lys Ser Leu Lys Leu Asn Gly Tyr Met Val Val			
		165	170	175
	Ile Asp Thr Gly Val Lys Gly Ser Thr Lys Gln Ala Val Glu Asp Val			
		180	185	190
25	His Val Leu Cys Glu Ser Asp Glu Tyr Met Lys Tyr Ile Glu His Ile			
		195	200	205
	Gly Thr Leu Val His Ser Ala Ser Glu Ser Ile Glu Gln His Asp Phe			
30		210	215	220
	His His Leu Ala Asp Ile Phe Asn Ala Cys Gln Glu Asp Leu Arg His			
	225	230	235	240
35	Leu Thr Val Ser His Asp Lys Ile Glu Lys Leu Leu Gln Ile Gly Lys			
		245	250	255
	Glu His Gly Ala Ile Ala Gly Lys Leu Thr Gly Gly Gly Arg Gly Gly			
		260	265	270
40	Ser Met Leu Leu Leu Ala Glu Asn Leu Lys Thr Ala Lys Thr Ile Val			
		275	280	285
	Ala Ala Val Glu Lys Ala Gly Ala Ala His Thr Trp Ile Glu His Leu			
45		290	295	300
	Gly Gly			
	305			
50	<210> 101			
	<211> 306			
	<212> PRT			
	<213> 金黄色葡萄球菌 (Staphylococcus aureus)			
	<400> 101			
55	Met Thr Arg Lys Gly Tyr Gly Glu Ser Thr Gly Lys Ile Ile Leu Ile			
	1	5	10	15
	Gly Glu His Ala Val Thr Phe Gly Glu Pro Ala Ile Ala Val Pro Phe			
		20	25	30
60				

	Asn	Ala	Gly	Lys	Ile	Lys	Val	Leu	Ile	Glu	Ala	Leu	Glu	Ser	Gly	Asn	
			35					40					45				
5	Tyr	Ser	Ser	Ile	Lys	Ser	Asp	Val	Tyr	Asp	Gly	Met	Leu	Tyr	Asp	Ala	
		50					55					60					
	Pro	Asp	His	Leu	Lys	Ser	Leu	Val	Asn	Arg	Phe	Val	Glu	Leu	Asn	Asn	
	65					70					75					80	
10	Ile	Thr	Glu	Pro	Leu	Ala	Val	Thr	Ile	Gln	Thr	Asn	Leu	Pro	Pro	Ser	
					85					90					95		
	Arg	Gly	Leu	Gly	Ser	Ser	Ala	Ala	Val	Ala	Val	Ala	Phe	Val	Arg	Ala	
				100					105					110			
15	Ser	Tyr	Asp	Phe	Leu	Gly	Lys	Ser	Leu	Thr	Lys	Glu	Glu	Leu	Ile	Glu	
			115				120						125				
	Lys	Ala	Asn	Trp	Ala	Glu	Gln	Ile	Ala	His	Gly	Lys	Pro	Ser	Gly	Ile	
20		130					135					140					
	Asp	Thr	Gln	Thr	Ile	Val	Ser	Gly	Lys	Pro	Val	Trp	Phe	Gln	Lys	Gly	
	145					150					155					160	
25	Gln	Ala	Glu	Thr	Leu	Lys	Thr	Leu	Ser	Leu	Asp	Gly	Tyr	Met	Val	Val	
					165					170					175		
	Ile	Asp	Thr	Gly	Val	Lys	Gly	Ser	Thr	Arg	Gln	Ala	Val	Glu	Asp	Val	
				180					185					190			
30	His	Lys	Leu	Cys	Glu	Asp	Pro	Gln	Tyr	Met	Ser	His	Val	Lys	His	Ile	
			195					200					205				
	Gly	Lys	Leu	Val	Leu	Arg	Ala	Ser	Asp	Val	Ile	Glu	His	His	Asn	Phe	
35		210					215					220					
	Glu	Ala	Leu	Ala	Asp	Ile	Phe	Asn	Glu	Cys	His	Ala	Asp	Leu	Lys	Ala	
	225					230					235					240	
	Leu	Thr	Val	Ser	His	Asp	Lys	Ile	Glu	Gln	Leu	Met	Lys	Ile	Gly	Lys	
40					245					250					255		
	Glu	Asn	Gly	Ala	Ile	Ala	Gly	Lys	Leu	Thr	Gly	Ala	Gly	Arg	Gly	Gly	
				260					265					270			
45	Ser	Met	Leu	Leu	Leu	Ala	Lys	Asp	Leu	Pro	Thr	Ala	Lys	Asn	Ile	Val	
			275					280					285				
	Lys	Ala	Val	Glu	Lys	Ala	Gly	Ala	Ala	His	Thr	Trp	Ile	Glu	Asn	Leu	
		290					295					300					
50	Gly	Gly															
	305																
	<210>	102															
55	<211>	345															
	<212>	PRT															
	<213>	链霉菌属菌种(Streptomyces sp.)															
	<400>	102															
60	Met	Gln	Lys	Arg	Gln	Arg	Glu	Leu	Ser	Ala	Leu	Thr	Leu	Pro	Thr	Ser	
	1				5					10					15		

	Ala	Glu	Gly	Val	Ser	Glu	Ser	His	Arg	Ala	Arg	Ser	Val	Gly	Ile	Gly	
				20					25					30			
5	Arg	Ala	His	Ala	Lys	Ala	Ile	Leu	Leu	Gly	Glu	His	Ala	Val	Val	Tyr	
			35					40					45				
	Gly	Ala	Pro	Ala	Leu	Ala	Leu	Pro	Ile	Pro	Gln	Leu	Thr	Val	Thr	Ala	
		50					55					60					
10	Ser	Val	Gly	Trp	Ser	Ser	Glu	Ala	Ser	Asp	Ser	Ala	Gly	Gly	Leu	Ser	
	65					70				75					80		
	Tyr	Thr	Met	Thr	Gly	Thr	Pro	Ser	Arg	Ala	Leu	Val	Thr	Gln	Ala	Ser	
15					85					90					95		
	Asp	Gly	Leu	His	Arg	Leu	Thr	Ala	Glu	Phe	Met	Ala	Arg	Met	Gly	Val	
				100					105					110			
	Thr	Asn	Ala	Pro	His	Leu	Asp	Val	Ile	Leu	Asp	Gly	Ala	Ile	Pro	His	
20			115					120					125				
	Gly	Arg	Gly	Leu	Gly	Ser	Ser	Ala	Ala	Gly	Ser	Arg	Ala	Ile	Ala	Leu	
		130					135					140					
25	Ala	Leu	Ala	Asp	Leu	Phe	Gly	His	Glu	Leu	Ala	Glu	His	Thr	Ala	Tyr	
	145					150					155					160	
	Glu	Leu	Val	Gln	Thr	Ala	Glu	Asn	Met	Ala	His	Gly	Arg	Ala	Ser	Gly	
				165					170						175		
30	Val	Asp	Ala	Met	Thr	Val	Gly	Ala	Ser	Arg	Pro	Leu	Leu	Phe	Gln	Gln	
				180					185					190			
	Gly	Arg	Thr	Glu	Arg	Leu	Ala	Ile	Gly	Cys	Asp	Ser	Leu	Phe	Ile	Val	
35			195					200					205				
	Ala	Asp	Ser	Gly	Val	Pro	Gly	Ser	Thr	Lys	Glu	Ala	Val	Glu	Met	Leu	
		210					215					220					
40	Arg	Glu	Gly	Phe	Thr	Arg	Ser	Ala	Gly	Thr	Gln	Glu	Arg	Phe	Val	Gly	
	225					230					235					240	
	Arg	Ala	Thr	Glu	Leu	Thr	Glu	Ala	Ala	Arg	Gln	Ala	Leu	Ala	Asp	Gly	
				245						250					255		
45	Arg	Pro	Glu	Glu	Leu	Gly	Ser	Gln	Leu	Thr	Tyr	Tyr	His	Glu	Leu	Leu	
			260					265						270			
	His	Glu	Ala	Arg	Leu	Ser	Thr	Asp	Gly	Ile	Asp	Ala	Leu	Val	Glu	Ala	
50			275					280					285				
	Ala	Leu	Lys	Ala	Gly	Ser	Leu	Gly	Ala	Lys	Ile	Thr	Gly	Gly	Gly	Leu	
		290					295					300					
55	Gly	Gly	Cys	Met	Ile	Ala	Gln	Ala	Arg	Pro	Glu	Gln	Ala	Arg	Glu	Val	
	305					310					315					320	
	Thr	Arg	Gln	Leu	His	Glu	Ala	Gly	Ala	Val	Gln	Thr	Trp	Val	Val	Pro	
				325						330					335		
60																	

Leu Lys Gly Leu Asp Asn His Ala Gln
 340 345

5 <210> 103
 <211> 334
 <212> PRT
 <213> 灰链霉菌 (Streptomyces griseosporus)
 <400> 103

10 Met Thr Leu Pro Thr Ser Val Glu Glu Gly Ser Lys Ala His Arg Ala
 1 5 10 15
 Arg Ala Val Gly Thr Gly Arg Ala His Ala Lys Ala Ile Leu Leu Gly
 20 25 30
 15 Glu His Ala Val Val Tyr Gly Thr Pro Ala Leu Ala Met Pro Ile Pro
 35 40 45
 20 Gln Leu Ala Val Thr Ala Ser Ala Gly Trp Ser Gly Arg Ser Ala Glu
 50 55 60
 Ser Arg Gly Gly Pro Thr Phe Thr Met Thr Gly Ser Ala Ser Arg Ala
 65 70 75 80
 25 Val Thr Ala Gln Ala Leu Asp Gly Leu Arg Arg Leu Thr Ala Ser Val
 85 90 95
 Lys Ala His Thr Gly Val Thr Asp Gly Gln His Leu Asp Val Ser Leu
 100 105 110
 30 Asp Gly Ala Ile Pro Pro Gly Arg Gly Leu Gly Ser Ser Ala Ala Asn
 115 120 125
 Ala Arg Ala Ile Ile Leu Ala Leu Ala Asp Leu Phe Gly Arg Glu Leu
 130 135 140
 35 Thr Glu Gly Glu Val Phe Asp Leu Val Gln Glu Ala Glu Asn Leu Thr
 145 150 155 160
 40 His Gly Arg Ala Ser Gly Val Asp Ala Val Thr Val Gly Ala Thr Ala
 165 170 175
 Pro Leu Leu Phe Arg Ala Gly Thr Ala Gln Ala Leu Asp Ile Gly Cys
 180 185 190
 45 Asp Ala Leu Phe Val Val Ala Asp Ser Gly Thr Ala Gly Ser Thr Lys
 195 200 205
 Glu Ala Ile Glu Leu Leu Arg Ala Gly Phe Arg Ala Gly Ala Gly Lys
 210 215 220
 50 Glu Glu Arg Phe Met His Arg Ala Ala His Leu Val Asp Asp Ala Arg
 225 230 235 240
 Ala Ser Leu Ala Glu Gly Glu Pro Glu Ala Phe Gly Ser Cys Leu Thr
 245 250 255
 55 Glu Tyr His Gly Leu Leu Arg Gly Ala Gly Leu Ser Thr Asp Arg Ile
 260 265 270
 60 Asp Ala Leu Val Asp Ala Ala Leu Gln Ala Asp Ser Leu Gly Ala Lys

208

Gly Ile Lys Leu Gly Ala Leu Ser Gly Lys Leu Ser Gly Ala Gly Lys
 245 250 255
 5 Gly Gly Ala Phe Ile Phe Leu Phe Glu Ser Leu Ile Lys Ala Asn Ile
 260 265 270
 Val Gln Lys Glu Leu Asn Asn Met Leu Asp Ser Lys Ile Asp Leu Leu
 275 280 285
 10 Leu Lys Leu Lys Val Ile Glu Thr
 290 295
 <210> 105
 15 <211> 336
 <212> PRT
 <213> 肺炎链球菌 (Streptococcus pneumoniae)
 <400> 105
 20 Met Ile Ala Val Lys Thr Cys Gly Lys Leu Tyr Trp Ala Gly Glu Tyr
 1 5 10 15
 Ala Ile Leu Glu Pro Gly Gln Leu Ala Leu Ile Lys Asp Ile Pro Ile
 20 25 30
 25 Tyr Met Arg Ala Glu Ile Ala Phe Ser Asp Ser Tyr Arg Ile Tyr Ser
 35 40 45
 Asp Met Phe Asp Phe Ala Val Asp Leu Arg Pro Asn Pro Asp Tyr Ser
 50 55 60
 30 Leu Ile Gln Glu Thr Ile Ala Leu Met Gly Asp Phe Leu Ala Val Arg
 65 70 75 80
 Gly Gln Asn Leu Arg Pro Phe Ser Leu Lys Ile Cys Gly Lys Met Glu
 85 90 95
 35 Arg Glu Gly Lys Lys Phe Gly Leu Gly Ser Ser Gly Ser Val Val Val
 100 105 110
 40 Leu Val Val Lys Ala Leu Leu Ala Leu Tyr Asn Leu Ser Val Asp Gln
 115 120 125
 Asn Leu Leu Phe Lys Leu Thr Ser Ala Val Leu Leu Lys Arg Gly Asp
 130 135 140
 45 Asn Gly Ser Met Gly Asp Leu Ala Cys Ile Val Ala Glu Asp Leu Val
 145 150 155 160
 Leu Tyr Gln Ser Phe Asp Arg Gln Lys Ala Ala Ala Trp Leu Glu Glu
 165 170 175
 50 Glu Asn Leu Ala Thr Val Leu Glu Arg Asp Trp Gly Phe Phe Ile Ser
 180 185 190
 Gln Val Lys Pro Thr Leu Glu Cys Asp Phe Leu Val Gly Trp Thr Lys
 195 200 205
 55 Glu Val Ala Val Ser Ser His Met Val Gln Gln Ile Lys Gln Asn Ile
 210 215 220
 60 Asn Gln Asn Phe Leu Ser Ser Ser Lys Glu Thr Val Val Ser Leu Val

	225		230		235		240
	Glu Ala Leu Glu Gln Gly Lys Ala Glu Lys Val Ile Glu Gln Val Glu						
		245			250		255
5	Val Ala Ser Lys Leu Leu Glu Gly Leu Ser Thr Asp Ile Tyr Thr Pro						
		260			265		270
10	Leu Leu Arg Gln Leu Lys Glu Ala Ser Gln Asp Leu Gln Ala Val Ala						
		275			280		285
	Lys Ser Ser Gly Ala Gly Gly Gly Asp Cys Gly Ile Ala Leu Ser Phe						
		290			295		300
15	Asp Ala Gln Ser Ser Arg Asn Thr Leu Lys Asn Arg Trp Ala Asp Leu						
		305			310		315
							320
	Gly Ile Glu Leu Leu Tyr Gln Glu Arg Ile Gly His Asp Asp Lys Ser						
			325			330	335
20	<210> 106						
	<211> 335						
	<212> PRT						
	<213> 酿脓链球菌 (Streptococcus pyogenes)						
25	<400> 106						
	Met Ser Asn Tyr Cys Val Gln Thr Gly Gly Lys Leu Tyr Leu Thr Gly						
	1		5			10	15
30	Glu Tyr Ala Ile Leu Ile Pro Gly Gln Lys Ala Leu Ile His Phe Ile						
		20			25		30
	Pro Leu Met Met Thr Ala Glu Ile Ser Pro Ala Ala His Ile Gln Leu						
		35			40		45
35	Ala Ser Asp Met Phe Ser His Lys Ala Gly Met Thr Pro Asp Ala Ser						
		50			55		60
	Tyr Ala Leu Ile Gln Ala Thr Val Lys Thr Phe Ala Asp Tyr Leu Gly						
		65			70		75
40	Gln Ser Ile Asp Gln Leu Glu Pro Phe Ser Leu Ile Ile Thr Gly Lys						
		85			90		95
	Met Glu Arg Asp Gly Lys Lys Phe Gly Ile Gly Ser Ser Gly Ser Val						
		100			105		110
45	Thr Leu Leu Thr Leu Lys Ala Leu Ser Ala Tyr Tyr Gln Ile Thr Leu						
		115			120		125
50	Thr Pro Glu Leu Leu Phe Lys Leu Ala Ala Tyr Thr Leu Leu Lys Gln						
		130			135		140
	Gly Asp Asn Gly Ser Met Gly Asp Ile Ala Cys Ile Ala Tyr Gln Thr						
		145			150		155
55	Leu Val Ala Tyr Thr Ser Phe Asp Arg Glu Gln Val Ser Asn Trp Leu						
		165			170		175
	Gln Thr Met Pro Leu Lys Lys Leu Leu Val Lys Asp Trp Gly Tyr His						
		180			185		190
60							

```

      Ile Gln Val Ile Gln Pro Ala Leu Pro Cys Asp Phe Leu Val Gly Trp
      195                                200                                205

5    Thr Lys Ile Pro Ala Ile Ser Arg Gln Met Ile Gln Gln Val Thr Ala
      210                                215                                220

      Ser Ile Thr Pro Ala Phe Leu Arg Thr Ser Tyr Gln Leu Thr Gln Ser
      225                                230                                235                                240

10   Ala Met Val Ala Leu Gln Glu Gly His Lys Glu Glu Leu Lys Lys Ser
      245                                250                                255

      Leu Ala Gly Ala Ser His Leu Leu Lys Glu Leu His Pro Ala Ile Tyr
      260                                265                                270

15   His Pro Lys Leu Val Thr Leu Val Ala Ala Cys Gln Lys Gln Asp Ala
      275                                280                                285

20   Val Ala Lys Ser Ser Gly Ala Gly Gly Gly Asp Cys Gly Ile Ala Leu
      290                                295                                300

      Ala Phe Asn Gln Asp Ala Arg Asp Thr Leu Ile Ser Lys Trp Gln Glu
      305                                310                                315                                320

25   Ala Asp Ile Ala Leu Leu Tyr Gln Glu Arg Trp Gly Glu Asn Asp
      325                                330                                335

<210> 107
<211> 368
<212> PRT
<213> 粪肠球菌(Enterococcus faecalis)
<400> 107
Met Ile Glu Val Thr Thr Pro Gly Lys Leu Phe Ile Ala Gly Glu Tyr
35  1                                5                                10                                15

      Ala Val Val Glu Pro Gly His Pro Ala Ile Ile Val Ala Val Asp Gln
      20                                25                                30

40   Phe Val Thr Val Thr Val Glu Glu Thr Thr Asp Glu Gly Ser Ile Gln
      35                                40                                45

      Ser Ala Gln Tyr Ser Ser Leu Pro Ile Arg Trp Thr Arg Arg Asn Gly
      50                                55                                60

45   Glu Leu Val Leu Asp Ile Arg Glu Asn Pro Phe His Tyr Val Leu Ala
      65                                70                                75                                80

      Ala Ile His Leu Thr Glu Lys Tyr Ala Gln Glu Gln Asn Lys Glu Leu
      85                                90                                95

50   Ser Phe Tyr His Leu Lys Val Thr Ser Glu Leu Asp Ser Ser Asn Gly
      100                               105                               110

55   Arg Lys Tyr Gly Leu Gly Ser Ser Gly Ala Val Thr Val Gly Thr Val
      115                               120                               125

      Lys Ala Leu Asn Ile Phe Tyr Asp Leu Gly Leu Glu Asn Glu Glu Ile
      130                               135                               140

60

```

	Phe	Lys	Leu	Ser	Ala	Leu	Ala	His	Leu	Ala	Val	Gln	Gly	Asn	Gly	Ser
	145					150					155					160
	Cys	Gly	Asp	Ile	Ala	Ala	Ser	Cys	Tyr	Gly	Gly	Trp	Ile	Ala	Phe	Ser
					165					170					175	
5	Thr	Phe	Asp	His	Asp	Trp	Val	Asn	Gln	Lys	Val	Thr	Thr	Glu	Thr	Leu
				180					185					190		
10	Thr	Asp	Leu	Leu	Ala	Met	Asp	Trp	Pro	Glu	Leu	Met	Ile	Phe	Pro	Leu
			195					200					205			
	Lys	Val	Pro	Lys	Gln	Leu	Arg	Leu	Leu	Ile	Gly	Trp	Thr	Gly	Ser	Pro
		210					215					220				
15	Ala	Ser	Thr	Ser	Asp	Leu	Val	Asp	Arg	Val	His	Gln	Ser	Lys	Glu	Glu
	225					230					235					240
	Lys	Gln	Ala	Ala	Tyr	Glu	Gln	Phe	Leu	Met	Lys	Ser	Arg	Leu	Cys	Val
					245					250					255	
20	Glu	Thr	Met	Ile	Asn	Gly	Phe	Asn	Thr	Gly	Lys	Ile	Ser	Val	Ile	Gln
				260					265					270		
	Lys	Gln	Ile	Thr	Lys	Asn	Arg	Gln	Leu	Leu	Ala	Glu	Leu	Ser	Ser	Leu
25			275					280					285			
	Thr	Gly	Val	Val	Ile	Glu	Thr	Glu	Ala	Leu	Lys	Asn	Leu	Cys	Asp	Leu
		290					295					300				
30	Ala	Glu	Ser	Tyr	Thr	Gly	Ala	Ala	Lys	Ser	Ser	Gly	Ala	Gly	Gly	Gly
	305					310					315					320
	Asp	Cys	Gly	Ile	Val	Ile	Phe	Arg	Gln	Lys	Ser	Gly	Ile	Leu	Pro	Leu
					325					330					335	
35	Met	Thr	Ala	Trp	Glu	Lys	Asp	Gly	Ile	Thr	Pro	Leu	Pro	Leu	His	Val
				340					345					350		
40	Tyr	Thr	Tyr	Gly	Gln	Lys	Glu	Cys	Lys	Glu	Lys	His	Glu	Ser	Lys	Arg
			355					360					365			
	<210>		108													
	<211>		361													
	<212>		PRT													
45	<213>		屎肠球菌(Enterococcus faecium)													
	<400>		108													
	Met	Ile	Glu	Val	Ser	Ala	Pro	Gly	Lys	Leu	Tyr	Ile	Ala	Gly	Glu	Tyr
	1				5					10					15	
50	Ala	Val	Val	Glu	Thr	Gly	His	Pro	Ala	Val	Ile	Ala	Ala	Val	Asp	Gln
				20					25					30		
	Phe	Val	Thr	Val	Thr	Val	Glu	Ser	Ala	Arg	Lys	Val	Gly	Ser	Ile	Gln
			35					40					45			
55	Ser	Ala	Gln	Tyr	Ser	Gly	Met	Pro	Val	Arg	Trp	Thr	Arg	Arg	Asn	Gly
		50					55					60				
60	Glu	Leu	Val	Leu	Asp	Ile	Arg	Glu	Asn	Pro	Phe	His	Tyr	Ile	Leu	Ala
	65					70					75					80

Ala Ile Arg Leu Thr Glu Lys Tyr Ala Gln Glu Lys Asn Ile Leu Leu
 85 90 95

5 Ser Phe Tyr Asp Leu Lys Val Thr Ser Glu Leu Asp Ser Ser Asn Gly
 100 105 110

Arg Lys Tyr Gly Leu Gly Ser Ser Gly Ala Val Thr Val Ala Thr Val
 115 120 125

10 Lys Ala Leu Asn Val Phe Tyr Ala Leu Asn Leu Ser Gln Leu Glu Ile
 130 135 140

15 Phe Lys Ile Ala Ala Leu Ala Asn Leu Ala Val Gln Asp Asn Gly Ser
 145 150 155 160

Cys Gly Asp Ile Ala Ala Ser Cys Tyr Gly Gly Trp Ile Ala Phe Ser
 165 170 175

20 Thr Phe Asp His Pro Trp Leu Gln Glu Gln Glu Thr Gln His Ser Ile
 180 185 190

Ser Glu Leu Leu Ala Leu Asp Trp Pro Gly Leu Ser Ile Glu Pro Leu
 195 200 205

25 Ile Ala Pro Glu Asp Leu Arg Leu Leu Ile Gly Trp Thr Gly Ser Pro
 210 215 220

Ala Ser Thr Ser Asp Leu Val Asp Gln Val His Arg Ser Arg Glu Asp
 225 230 235 240

30 Lys Met Val Ala Tyr Gln Leu Phe Leu Lys Asn Ser Thr Glu Cys Val
 245 250 255

35 Asn Glu Met Ile Lys Gly Phe Lys Glu Asn Asn Val Thr Leu Ile Gln
 260 265 270

Gln Met Ile Arg Lys Asn Arg Gln Leu Leu His Asp Leu Ser Ala Ile
 275 280 285

40 Thr Gly Val Val Ile Glu Thr Pro Ala Leu Asn Lys Leu Cys Asn Leu
 290 295 300

Ala Glu Gln Tyr Glu Gly Ala Ala Lys Ser Ser Gly Ala Gly Gly Gly
 305 310 315 320

45 Asp Cys Gly Ile Val Ile Val Asp Gln Lys Ser Gly Ile Leu Pro Leu
 325 330 335

50 Met Ser Ala Trp Glu Lys Ala Glu Ile Thr Pro Leu Pro Leu His Val
 340 345 350

Tyr Ser Asp Gln Arg Lys Glu Asn Arg
 355 360

55 <210> 109
 <211> 358
 <212> PRT
 <213> 溶血葡萄球菌 (Staphylococcus haemolyticus)
 <400> 109

60 Met Ile Gln Val Lys Ala Pro Gly Lys Leu Tyr Val Ala Gly Glu Tyr

1	5	10	15
Ala Val Thr Glu Pro Gly Tyr Lys Ser Val Leu Ile Ala Val Asp Arg	20	25	30
5 Phe Val Thr Ala Ser Ile Glu Ala Ser Asn Ala Val Thr Ser Thr Ile	35	40	45
His Ser Lys Thr Leu His Tyr Glu Pro Val Thr Phe Asn Arg Asn Glu	50	55	60
10 Asp Lys Ile Asp Ile Ser Asp Ala Asn Ala Ala Ser Gln Leu Lys Tyr	65	70	75
Val Val Thr Ala Ile Glu Val Phe Glu Gln Tyr Ala Arg Ser Cys Asn	85	90	95
Val Lys Leu Lys His Phe His Leu Glu Ile Asp Ser Asn Leu Asp Asp	100	105	110
20 Ala Ser Gly Asn Lys Tyr Gly Leu Gly Ser Ser Ala Ala Val Leu Val	115	120	125
Ser Val Val Lys Ala Leu Asn Glu Phe Tyr Asp Met Gln Leu Ser Asn	130	135	140
25 Leu Tyr Ile Tyr Lys Leu Ala Val Ile Ser Asn Met Arg Leu Gln Ser	145	150	155
Leu Ser Ser Cys Gly Asp Ile Ala Val Ser Val Tyr Ser Gly Trp Leu	165	170	175
30 Ala Tyr Ser Thr Phe Asp His Asp Trp Val Lys Gln Gln Met Glu Glu	180	185	190
35 Thr Ser Val Asn Glu Val Leu Glu Lys Asn Trp Pro Gly Leu His Ile	195	200	205
Glu Pro Leu Gln Ala Pro Glu Asn Met Glu Val Leu Ile Gly Trp Thr	210	215	220
40 Gly Ser Pro Ala Ser Ser Pro His Leu Val Ser Glu Val Lys Arg Leu	225	230	235
Lys Ser Asp Pro Ser Phe Tyr Gly Arg Phe Leu Asp Gln Ser His Thr	245	250	255
Cys Val Glu Asn Leu Ile Tyr Ala Phe Lys Thr Asp Asn Ile Lys Gly	260	265	270
50 Val Gln Lys Met Ile Arg Gln Asn Arg Met Ile Ile Gln Gln Met Asp	275	280	285
Asn Glu Ala Thr Val Asp Ile Glu Thr Glu Asn Leu Lys Met Leu Cys	290	295	300
Asp Ile Gly Glu Arg Tyr Gly Ala Ala Ala Lys Thr Ser Gly Ala Gly	305	310	315
60 Gly Gly Asp Cys Gly Ile Ala Ile Ile Asp Asn Arg Ile Asp Lys Asn			

					325					330					335		
	Arg	Ile	Tyr	Asn	Glu	Trp	Ala	Ser	His	Gly	Ile	Lys	Pro	Leu	Lys	Phe	
				340					345					350			
5	Lys	Ile	Tyr	His	Gly	Gln											
			355														
	<210>	110															
10	<211>	358															
	<212>	PRT															
	<213>	表皮葡萄球菌 (Staphylococcus epidermidis)															
	<400>	110															
15	Met	Ile	Gln	Val	Lys	Ala	Pro	Gly	Lys	Leu	Tyr	Ile	Ala	Gly	Glu	Tyr	
	1				5					10					15		
	Ala	Val	Thr	Glu	Pro	Gly	Tyr	Lys	Ser	Ile	Leu	Ile	Ala	Val	Asn	Arg	
				20					25					30			
20	Phe	Val	Thr	Ala	Thr	Ile	Glu	Ala	Ser	Asn	Lys	Val	Glu	Gly	Ser	Ile	
			35					40					45				
	His	Ser	Lys	Thr	Leu	His	Tyr	Glu	Pro	Val	Lys	Phe	Asp	Arg	Asn	Glu	
		50					55					60					
25	Asp	Arg	Ile	Glu	Ile	Ser	Asp	Val	Gln	Ala	Ala	Lys	Gln	Leu	Lys	Tyr	
	65					70					75					80	
	Val	Val	Thr	Ala	Ile	Glu	Val	Phe	Glu	Gln	Tyr	Val	Arg	Ser	Cys	Asn	
30					85					90					95		
	Met	Asn	Leu	Lys	His	Phe	His	Leu	Thr	Ile	Asp	Ser	Asn	Leu	Ala	Asp	
				100					105					110			
35	Asn	Ser	Gly	Gln	Lys	Tyr	Gly	Leu	Gly	Ser	Ser	Ala	Ala	Val	Leu	Val	
			115					120					125				
	Ser	Val	Val	Lys	Ala	Leu	Asn	Glu	Phe	Tyr	Gly	Leu	Glu	Leu	Ser	Asn	
			130				135					140					
40	Leu	Tyr	Ile	Tyr	Lys	Leu	Ala	Val	Ile	Ala	Asn	Met	Lys	Leu	Gln	Ser	
	145					150					155					160	
	Leu	Ser	Ser	Cys	Gly	Asp	Ile	Ala	Val	Ser	Val	Tyr	Ser	Gly	Trp	Leu	
					165					170					175		
45	Ala	Tyr	Ser	Thr	Phe	Asp	His	Asp	Trp	Val	Lys	Gln	Gln	Met	Glu	Glu	
				180					185					190			
	Thr	Ser	Val	Asn	Asp	Val	Leu	Glu	Lys	Asn	Trp	Pro	Gly	Leu	His	Ile	
50			195					200					205				
	Glu	Pro	Leu	Gln	Ala	Pro	Glu	Asn	Met	Glu	Val	Leu	Ile	Gly	Trp	Thr	
		210					215					220					
55	Gly	Ser	Pro	Ala	Ser	Ser	Pro	His	Leu	Val	Ser	Glu	Val	Lys	Arg	Leu	
	225					230					235					240	
	Lys	Ser	Asp	Pro	Ser	Phe	Tyr	Gly	Asp	Phe	Leu	Asp	Gln	Ser	His	Ala	
					245					250					255		
60																	

	Cys	Val	Glu	Ser	Leu	Ile	Gln	Ala	Phe	Lys	Thr	Asn	Asn	Ile	Lys	Gly	
				260					265					270			
5	Val	Gln	Lys	Met	Ile	Arg	Ile	Asn	Arg	Arg	Ile	Ile	Gln	Ser	Met	Asp	
			275					280					285				
	Asn	Glu	Ala	Ser	Val	Glu	Ile	Glu	Thr	Asp	Lys	Leu	Lys	Lys	Leu	Cys	
		290					295					300					
10	Asp	Val	Gly	Glu	Lys	His	Gly	Gly	Ala	Ser	Lys	Thr	Ser	Gly	Ala	Gly	
	305					310					315					320	
	Gly	Gly	Asp	Cys	Gly	Ile	Thr	Ile	Ile	Asn	Lys	Val	Ile	Asp	Lys	Asn	
					325					330					335		
15	Ile	Ile	Tyr	Asn	Glu	Trp	Gln	Met	Asn	Asp	Ile	Lys	Pro	Leu	Lys	Phe	
				340					345					350			
20	Lys	Ile	Tyr	His	Gly	Gln											
			355														
	<210>	111															
	<211>	358															
25	<212>	PRT															
	<213>	金黄色葡萄球菌 (Staphylococcus aureus)															
	<400>	111															
	Met	Ile	Gln	Val	Lys	Ala	Pro	Gly	Lys	Leu	Tyr	Ile	Ala	Gly	Glu	Tyr	
	1				5					10					15		
30	Ala	Val	Thr	Glu	Pro	Gly	Tyr	Lys	Ser	Val	Leu	Ile	Ala	Leu	Asp	Arg	
				20					25					30			
	Phe	Val	Thr	Ala	Thr	Ile	Glu	Glu	Ala	Thr	Gln	Tyr	Lys	Gly	Thr	Ile	
35			35					40					45				
	His	Ser	Lys	Ala	Leu	His	His	Asn	Pro	Val	Thr	Phe	Ser	Arg	Asp	Glu	
		50					55					60					
40	Asp	Ser	Ile	Val	Ile	Ser	Asp	Pro	His	Ala	Ala	Lys	Gln	Leu	Asn	Tyr	
	65					70					75					80	
	Val	Val	Thr	Ala	Ile	Glu	Ile	Phe	Glu	Gln	Tyr	Ala	Lys	Ser	Cys	Asp	
					85					90					95		
45	Ile	Ala	Met	Lys	His	Phe	His	Leu	Thr	Ile	Asp	Ser	Asn	Leu	Asp	Asp	
				100					105					110			
	Ser	Asn	Gly	His	Lys	Tyr	Gly	Leu	Gly	Ser	Ser	Ala	Ala	Val	Leu	Val	
50			115					120					125				
	Ser	Val	Ile	Lys	Val	Leu	Asn	Glu	Phe	Tyr	Asp	Met	Lys	Leu	Ser	Asn	
		130					135					140					
55	Leu	Tyr	Ile	Tyr	Lys	Leu	Ala	Val	Ile	Ala	Asn	Met	Lys	Leu	Gln	Ser	
	145					150					155					160	
	Leu	Ser	Ser	Cys	Gly	Asp	Ile	Ala	Val	Ser	Val	Tyr	Ser	Gly	Trp	Leu	
					165					170					175		
60																	

```

      Ala Tyr Ser Thr Phe Asp His Glu Trp Val Lys His Gln Ile Glu Asp
              180              185              190
      Thr Thr Val Glu Glu Val Leu Ile Lys Asn Trp Pro Gly Leu His Ile
              195              200              205
5      Glu Pro Leu Gln Ala Pro Glu Asn Met Glu Val Leu Ile Gly Trp Thr
              210              215              220

      Gly Ser Pro Ala Ser Ser Pro His Phe Val Ser Glu Val Lys Arg Leu
10      225              230              235              240

      Lys Ser Asp Pro Ser Phe Tyr Gly Asp Phe Leu Glu Asp Ser His Arg
              245              250              255

15      Cys Val Glu Lys Leu Ile His Ala Phe Lys Thr Asn Asn Ile Lys Gly
              260              265              270

      Val Gln Lys Met Val Arg Gln Asn Arg Thr Ile Ile Gln Arg Met Asp
              275              280              285
20      Lys Glu Ala Thr Val Asp Ile Glu Thr Glu Lys Leu Lys Tyr Leu Cys
              290              295              300

      Asp Ile Ala Glu Lys Tyr His Gly Ala Ser Lys Thr Ser Gly Ala Gly
25      305              310              315              320

      Gly Gly Asp Cys Gly Ile Thr Ile Ile Asn Lys Asp Val Asp Lys Glu
              325              330              335

30      Lys Ile Tyr Asp Glu Trp Thr Lys His Gly Ile Lys Pro Leu Lys Phe
              340              345              350

      Asn Ile Tyr His Gly Gln
              355
35      <210> 112
          <211> 374
          <212> PRT
          <213> 链霉菌属菌种 (Streptomyces sp.)
40      <400> 112
      Met Thr Thr Gly Gln Arg Thr Ile Val Arg His Ala Pro Gly Lys Leu
      1              5              10              15

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

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

      Gln Ala Val Gly Trp Arg Trp His Asp Gly Arg Leu Val Val Arg Asp
      65              70              75              80

55      Pro Asp Asp Gly Gln Gln Ala Arg Ser Ala Leu Ala His Val Val Ser
              85              90              95

      Ala Ile Glu Thr Val Gly Arg Leu Leu Gly Glu Arg Gly Gln Lys Val
              100              105              110
60

```

```

    Pro Ala Leu Thr Leu Ser Val Ser Ser Arg Leu His Glu Asp Gly Arg
      115      120      125
5    Lys Phe Gly Leu Gly Ser Ser Gly Ala Val Thr Val Ala Thr Val Ala
      130      135      140
    Ala Val Ala Ala Phe Cys Gly Leu Glu Leu Ser Thr Asp Glu Arg Phe
      145      150      155      160
10   Arg Leu Ala Met Leu Ala Thr Ala Glu Leu Asp Pro Lys Gly Ser Gly
      165      170      175
    Gly Asp Leu Ala Ala Ser Thr Trp Gly Gly Trp Ile Ala Tyr Gln Ala
      180      185
15   Pro Asp Arg Ala Phe Val Leu Asp Leu Ala Arg Arg Val Gly Val Asp
      195      200      205
    Arg Thr Leu Lys Ala Pro Trp Pro Gly His Ser Val Arg Arg Leu Pro
      210      215      220
    Ala Pro Lys Gly Leu Thr Leu Glu Val Gly Trp Thr Gly Glu Pro Ala
      225      230      235      240
25   Ser Thr Ala Ser Leu Val Ser Asp Leu His Arg Arg Thr Trp Arg Gly
      245      250      255
    Ser Ala Ser His Gln Arg Phe Val Glu Thr Thr Thr Asp Cys Val Arg
      260      265      270
    Ser Ala Val Thr Ala Leu Glu Ser Gly Asp Asp Thr Ser Leu Leu His
      275      280      285
35   Glu Ile Arg Arg Ala Arg Gln Glu Leu Ala Arg Leu Asp Asp Glu Val
      290      295      300
    Gly Leu Gly Ile Phe Thr Pro Lys Leu Thr Ala Leu Cys Asp Ala Ala
      305      310      315      320
40   Glu Ala Val Gly Gly Ala Ala Lys Pro Ser Gly Ala Gly Gly Gly Asp
      325      330      335
    Cys Gly Ile Ala Leu Leu Asp Ala Glu Ala Ser Arg Asp Ile Thr His
      340      345      350
45   Val Arg Gln Arg Trp Glu Thr Ala Gly Val Leu Pro Leu Pro Leu Thr
      355      360      365
50   Pro Ala Leu Glu Gly Ile
      370
    <210> 113
    <211> 360
55   <212> PRT
    <213> 灰孢链霉菌 (Streptomyces griseosporus)
    <400> 113
    Met Thr Gly Pro Arg Ala Val Thr Arg Arg Ala Pro Gly Lys Leu Phe
      1      5      10      15
60

```

	Val	Ala	Gly	Glu	Tyr	Ala	Val	Val	Glu	Pro	Gly	Asn	Arg	Ala	Ile	Leu
			20						25					30		
5	Val	Ala	Val	Asp	Arg	Tyr	Val	Thr	Val	Thr	Val	Ser	Asp	Gly	Ala	Ala
			35					40					45			
	Pro	Gly	Val	Val	Val	Ser	Ser	Asp	Ile	Gly	Ala	Gly	Pro	Val	His	His
		50					55					60				
10	Pro	Trp	Gln	Asp	Gly	Arg	Leu	Thr	Gly	Gly	Thr	Gly	Thr	Pro	His	Val
	65					70					75					80
	Val	Ala	Ala	Val	Glu	Thr	Val	Ala	Arg	Leu	Leu	Ala	Glu	Arg	Gly	Arg
					85					90					95	
15	Ser	Val	Pro	Pro	Leu	Gly	Trp	Ser	Ile	Ser	Ser	Thr	Leu	His	Glu	Asp
				100					105					110		
	Gly	Arg	Lys	Phe	Gly	Leu	Gly	Ser	Ser	Gly	Ala	Val	Thr	Val	Ala	Thr
20			115					120					125			
	Val	Ser	Ala	Val	Ala	Ala	His	Cys	Gly	Leu	Glu	Leu	Thr	Ala	Glu	Glu
		130					135					140				
25	Arg	Phe	Arg	Thr	Ala	Leu	Ile	Ala	Ser	Ala	Arg	Ile	Asp	Pro	Arg	Gly
	145					150					155					160
	Ser	Gly	Gly	Asp	Ile	Ala	Thr	Ser	Thr	Trp	Gly	Gly	Trp	Ile	Ala	Tyr
					165					170					175	
30	Arg	Ala	Pro	Asp	Arg	Asp	Ala	Val	Leu	Asp	Leu	Thr	Arg	Arg	Gln	Gly
				180					185					190		
	Val	Asp	Glu	Ala	Leu	Arg	Ala	Pro	Trp	Pro	Gly	Phe	Ser	Val	Arg	Leu
35			195					200					205			
	Ser	Pro	Pro	Arg	Asn	Leu	Cys	Leu	Glu	Val	Gly	Trp	Thr	Gly	Asn	Pro
		210					215					220				
40	Val	Ser	Thr	Thr	Ser	Leu	Leu	Thr	Asp	Leu	His	Arg	Arg	Thr	Trp	Arg
	225					230					235					240
	Gly	Ser	Pro	Ala	Tyr	Arg	Arg	Tyr	Val	Gly	Ala	Thr	Gly	Glu	Leu	Val
					245					250					255	
45	Asp	Ala	Ala	Val	Ile	Ala	Leu	Glu	Asp	Gly	Asp	Thr	Glu	Gly	Leu	Leu
				260					265					270		
	Arg	Gln	Val	Arg	Arg	Ala	Arg	His	Glu	Met	Val	Arg	Leu	Asp	Asp	Glu
50			275					280					285			
	Val	Gly	Leu	Gly	Ile	Phe	Thr	Pro	Glu	Leu	Thr	Ala	Leu	Cys	Ala	Ile
		290					295					300				
55	Ala	Glu	Arg	Ala	Gly	Ala	Ala	Lys	Pro	Ser	Gly	Ala	Gly	Gly	Gly	Asp
	305					310					315					320
	Cys	Gly	Ile	Ala	Leu	Leu	Asp	Ala	Glu	Ala	Arg	Tyr	Asp	Arg	Ser	Pro
60					325					330					335	

```

Leu His Arg Gln Trp Ala Ala Ala Gly Val Leu Pro Leu Leu Val Ser
340 345 350

5 Pro Ala Thr Glu Gly Val Glu Glu
355 360

<210> 114
<211> 317
<212> PRT
10 <213> 布氏疏螺旋体 (Borrelia burgdorferi)
<400> 114
Met Asp Leu Ile Ser Phe Ser Val Pro Gly Asn Leu Leu Leu Met Gly
1 5 10 15

15 Glu Tyr Thr Ile Leu Glu Glu Lys Gly Leu Gly Leu Ala Ile Ala Ile
20 25 30

Asn Lys Arg Ala Phe Phe Ser Phe Lys Lys Ser Asp Ser Trp Arg Phe
35 40 45

20 Phe Ser Lys Lys Lys Lys Ile Asp Asp Phe Ser Leu Ile Glu Asn Arg
50 55 60

Ser Asp Phe Val Phe Lys Met Phe Ala Tyr Leu Ser Gln Asn Cys Phe
25 65 70 75 80

Phe Asn Leu Glu Asn Phe Ala Tyr Asp Val Tyr Ile Asp Thr Ser Asn
85 90 95

30 Phe Phe Phe Asn Asp Gly Thr Lys Lys Gly Phe Gly Ser Ser Ala Val
100 105 110

Val Ala Ile Gly Ile Val Cys Gly Leu Phe Leu Ile His Asn Ala Thr
115 120 125

35 Asn Val Val Glu Lys Gly Glu Ile Phe Lys Tyr Cys Leu Glu Ala Tyr
130 135 140
Arg Tyr Ser Gln Gly Gly Ile Gly Ser Gly Tyr Asp Ile Ala Thr Ser
145 150 155 160

40 Ile Phe Gly Gly Val Ile Glu Phe Glu Gly Gly Phe Asn Pro Lys Cys
165 170 175

Arg Gln Leu Gly Ala Val Glu Phe Asn Asp Phe Tyr Leu Met Gln Gly
45 180 185 190

Leu Gln Ala Ile Lys Thr Thr Thr Ser Ile Cys Glu Tyr Asn Lys His
195 200 205

50 Arg Asn Ser Ile Leu Asp Phe Ile Leu Lys Cys Asn Leu Glu Met Lys
210 215 220

Lys Leu Val Leu Asn Ala Ser Asn Ser Lys Ser Ala Leu Ile Ser Ser
225 230 235 240

55 Leu Arg Arg Ala Lys Glu Leu Gly Leu Ala Ile Gly Glu Ala Ile Gly
245 250 255

Val Ser Ala Ala Leu Pro Ser Ser Phe Asp His Leu Leu Gly Gln Cys
60 260 265 270

```



```

      Asp Leu Ile Lys Ala Leu Gly Ala Gly Asn Glu Thr Phe Leu Val Tyr
          275                                280                    285

5   Arg Pro Asn Ile Glu Ala Phe Asn Leu Ser Lys Ile Ile Ser Ile Val
      290                                295                    300

      Leu Glu Asn Glu Gly Ile Lys Phe Glu Ser Asp Lys Cys
      305                                310                    315

10  <210>  115
    <211>  30
    <212>  DNA
    <213>  人工的
15  <220>
    <221>  misc_feature
    <222>  (1)..(30)
    <223>  引物
    <400>  115
20  gggcaagctt gtccacggca cgaccaagca                                30
    <210>  116
    <211>  30
    <212>  DNA
    <213>  人工的
25  <220>
    <221>  misc_feature
    <222>  (1)..(30)
    <223>  引物
    <400>  116
30  cgtaatccgc ggccgcgttt ccagcgcgtc                                30
    <210>  117
    <211>  28
    <212>  DNA
    <213>  人工的
35  <220>
    <221>  misc_feature
    <222>  (1)..(28)
    <223>  引物
    <400>  117
40  aattaaagga gggtttcata tgaattcg                                28
    <210>  118
    <211>  28
    <212>  DNA
    <213>  人工的
45  <220>
    <221>  misc_feature
    <222>  (1)..(28)
    <223>  引物
50  <400>  118
    gatccgaatt catatgaaac cctccttt                                28
    <210>  119
55  <211>  30
    <212>  DNA
    <213>  人工的
    <220>
    <221>  misc_feature

```

	<222> (1)..(30)	
	<223> 引物	
	<400> 119	30
5	aaggcctcat atgatttccc ataccccggt	
	<210> 120	
	<211> 28	
	<212> DNA	
	<213> 人工的	
	<220>	
10	<221> misc_feature	
	<222> (1)..(28)	
	<223> 引物	
	<400> 120	28
	cgggatcctc atcgctccat ctccatgt	
15	<210> 121	
	<211> 30	
	<212> DNA	
	<213> 人工的	
20	<220>	
	<221> misc_feature	
	<222> (1)..(30)	
	<223> 引物	
	<400> 121	30
25	aaggcctcat atgaccgaca gcaaggatca	
	<210> 122	
	<211> 28	
	<212> DNA	
30	<213> 人工的	
	<220>	
	<221> misc_feature	
	<222> (1)..(28)	
	<223> 引物	
35	<400> 122	28
	cgggatcctc attgacggat aagcgagg	
	<210> 123	
	<211> 29	
40	<212> DNA	
	<213> 人工的	
	<220>	
	<221> misc_feature	
	<222> (1)..(29)	
45	<223> 引物	
	<400> 123	29
	aaggcctcat atgaaagtgc ctaagatga	
	<210> 124	
	<211> 28	
50	<212> DNA	
	<213> 人工的	
	<220>	
	<221> misc_feature	
	<222> (1)..(28)	
55	<223> 引物	
	<400> 124	28
	cgggatcctc aggcctgccg gtcgacat	
	<210> 125	
60	<211> 34	

	<212>	DNA	
	<213>	人工的	
	<220>		
	<221>	misc_feature	
5	<222>	(1)..(34)	
	<223>	引物	
	<400>	125	
		aaggcctcat atgagcaccg gcaggcctga agca	34
10	<210>	126	
	<211>	31	
	<212>	DNA	
	<213>	人工的	
	<220>		
15	<221>	misc_feature	
	<222>	(1)..(31)	
	<223>	引物	
	<400>	126	
		cgggatcctc atccctgccc cggcagcggt t	31
20	<210>	127	
	<211>	30	
	<212>	DNA	
	<213>	人工的	
25	<220>		
	<221>	misc_feature	
	<222>	(1)..(30)	
	<223>	引物	
	<400>	127	
30		aaggcctcat atggatcagg tcatccgcgc	30
	<210>	128	
	<211>	28	
	<212>	DNA	
	<213>	人工的	
35	<220>		
	<221>	misc_feature	
	<222>	(1)..(28)	
	<223>	引物	
	<400>	128	
40		cgggatcctc agtcacgaa aacaagtc	28
	<210>	129	
	<211>	30	
	<212>	DNA	
45	<213>	人工的	
	<220>		
	<221>	misc_feature	
	<222>	(1)..(30)	
	<223>	引物	
50	<400>	129	
		aaggcctcat atgactgatg ccgtccgcga	30
	<210>	130	
	<211>	28	
55	<212>	DNA	
	<213>	人工的	
	<220>		
	<221>	misc_feature	
	<222>	(1)..(28)	
60	<223>	引物	

	<400> 130		
	cgggatcctc aacgcccctc gaacggcg		28
	<210> 131		
5	<211> 31		
	<212> DNA		
	<213> 人工的		
	<220>		
	<221> misc_feature		
10	<222> (1)..(31)		
	<223> 引物		
	<400> 131		
	ccggcattcg ggcggcatcc aggtctcgct g		31
	<210> 132		
15	<211> 31		
	<212> DNA		
	<213> 人工的		
	<220>		
	<221> misc_feature		
20	<222> (1)..(31)		
	<223> 引物		
	<400> 132		
	cagcgagacc tggatgccgc ccgaatgccg g		31
25	<210> 133		
	<211> 31		
	<212> DNA		
	<213> 人工的		
	<220>		
30	<221> misc_feature		
	<222> (1)..(31)		
	<223> 引物		
	<400> 133		
	cgtgcagggc tggattctgt cggaataccc g		31
35	<210> 134		
	<211> 31		
	<212> DNA		
	<213> 人工的		
40	<220>		
	<221> misc_feature		
	<222> (1)..(31)		
	<223> 引物		
	<400> 134		
45	cggtattcc gacagaatcc agccctgcac g		31
	<210> 135		
	<211> 31		
	<212> DNA		
50	<213> 人工的		
	<220>		
	<221> misc_feature		
	<222> (1)..(31)		
	<223> 引物		
55	<400> 135		
	gggtgcgcg ccggcatccg gcatttcgac g		31
	<210> 136		
	<211> 31		
	<212> DNA		
60	<213> 人工的		

	<220>		
	<221>	misc_feature	
	<222>	(1)..(31)	
	<223>	引物	
5	<400>	136	31
		cgtcgaaatg ccggatgccg gcgcgcagcc c	
	<210>	137	
	<211>	31	
10	<212>	DNA	
	<213>	人工的	
	<220>		
	<221>	misc_feature	
	<222>	(1)..(31)	
15	<223>	引物	
	<400>	137	31
		gggtgcgacg ggcgagttct tcgatgcgcg g	
	<210>	138	
20	<211>	31	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<221>	misc_feature	
25	<222>	(1)..(31)	
	<223>	引物	
	<400>	138	31
		ccgcgcacg aagaactcgc ccgtcgcacc c	
30	<210>	139	
	<211>	31	
	<212>	DNA	
	<213>	人工的	
	<220>		
35	<221>	misc_feature	
	<222>	(1)..(31)	
	<223>	引物	
	<400>	139	31
		cacgcccgtc acatacgacg aatacgttgc c	
40	<210>	140	
	<211>	31	
	<212>	DNA	
	<213>	人工的	
	<220>		
45	<221>	misc_feature	
	<222>	(1)..(31)	
	<223>	引物	
	<400>	140	31
		ggcaacgtat tcgtcgtatg tgacgggcggt g	
50	<210>	141	
	<211>	31	
	<212>	DNA	
	<213>	人工的	
55	<220>		
	<221>	misc_feature	
	<222>	(1)..(31)	
	<223>	引物	
	<400>	141	31
60		gaggctcggg cttggctcct cggcggcggt g	

	<210>	142		
	<211>	31		
	<212>	DNA		
5	<213>	人工的		
	<220>			
	<221>	misc_feature		
	<222>	(1)..(31)		
	<223>	引物		
10	<400>	142		
		caccgccgcc gaggagccaa gcccgagcct c		31
	<210>	143		
	<211>	31		
15	<212>	DNA		
	<213>	人工的		
	<220>			
	<221>	misc_feature		
	<222>	(1)..(31)		
20	<223>	引物		
	<400>	143		
		cggcacgctg ctggaccgg ggcgcgcctt c		31
	<210>	144		
	<211>	31		
25	<212>	DNA		
	<213>	人工的		
	<220>			
	<221>	misc_feature		
	<222>	(1)..(31)		
30	<223>	引物		
	<400>	144		
		gaaggcgtcg cccgggtcca gcagcgtgcc g		31
	<210>	145		
35	<211>	36		
	<212>	DNA		
	<213>	人工的		
	<220>			
	<221>	misc_feature		
40	<222>	(1)..(36)		
	<223>	寡核苷酸		
	<400>	145		
		tcagaattcg gtaccatatg aagcttggat ccgggg		36
45	<210>	146		
	<211>	29		
	<212>	DNA		
	<213>	人工的		
	<220>			
50	<221>	misc_feature		
	<222>	(1)..(29)		
	<223>	寡核苷酸		
	<400>	146		
		ggatccaagc ttcatatggt accgaattc		29
55	<210>	147		
	<211>	26		
	<212>	DNA		
	<213>	人工的		
60	<220>			

```

    <221> misc_feature
    <222> (1)..(26)
    <223> 引物
    <400> 147
5  ggaattcgct gctgaacgcg atggcg                                26
    <210> 148
    <211> 32
    <212> DNA
    <213> 人工的
10 <220>
    <221> misc_feature
    <222> (1)..(32)
    <223> 引物
    <400> 148
15 ggggtaccat atgtgccttc gttgcgtcag tc                                32
    <210> 149
    <211> 50
    <212> DNA
    <213> 人工的
    <220>
    <221> misc_feature
    <222> (1)..(50)
    <223> 寡核苷酸
25 <400> 149
    gatccggcgt gtgcgcaatt taattgcgca cacgccccct gcgtttaaac        50
    <210> 150
    <211> 50
    <212> DNA
    <213> 人工的
    <220>
    <221> misc_feature
    <222> (1)..(50)
35 <223> 寡核苷酸
    <400> 150
    gatcgtttaa acgcaggggg cgtgtgcgca attaaattgc gcacacgccg        50
    <210> 151
    <211> 30
    <212> DNA
    <213> 人工的
    <220>
    <221> misc_feature
45 <222> (1)..(30)
    <223> 引物
    <400> 151
    aaggcctcat atgacgccca agcagcaatt                                30
    <210> 152
    <211> 26
    <212> DNA
    <213> 人工的
    <220>
    <221> misc_feature
55 <222> (1)..(26)
    <223> 引物
    <400> 152
    cgggatccta ggcgctgcgg cggatg                                26

```

	<210>	153		
	<211>	30		
	<212>	DNA		
	<213>	人工的		
5	<220>			
	<221>	misc_feature		
	<222>	(1)..(30)		
	<223>	引物		
	<400>	153		
10		cgggatactc atgcctgccg gtcgacatag		30
	<210>	154		
	<211>	30		
	<212>	DNA		
15	<213>	人工的		
	<220>			
	<221>	misc_feature		
	<222>	(1)..(30)		
	<223>	引物		
20	<400>	154		
		gaaggcacat atgaatcagg tcatccgcgc		30
	<210>	155		
	<211>	30		
25	<212>	DNA		
	<213>	人工的		
	<220>			
	<221>	misc_feature		
	<222>	(1)..(30)		
30	<223>	引物		
	<400>	155		
		gccggatcct cattcatcga aaacaagtcc		30
	<210>	156		
	<211>	30		
35	<212>	DNA		
	<213>	人工的		
	<220>			
	<221>	misc_feature		
	<222>	(1)..(30)		
40	<223>	引物		
	<400>	156		
		acgccggatc ctcacgcgcc ctcgaacggc		30
	<210>	157		
45	<211>	1612		
	<212>	DNA		
	<213>	副球菌属菌种(Paracoccus sp.)		
	<220>			
	<221>	基因		
50	<222>	(1)..(1612)		
	<223>			
	<400>	157		
		ccatggcatc cgggtcggat gccgtctatg ttggcccgaa caggcagcag gaggcccat		60
		gagcgatatc cagaccctct cgttcgagga agccatgcgc gagctggagg cgaccgtcgg		120
55		caagctggaa accggcgagg cgacgctcga ggactccatc gcgctctatg aacgcggggc		180
		ggcgtctgcgc gccatttgcg aaaccgcct gcgcgaggcc gaggagcggg tcgagaagat		240
		caccctggcc gcgaacgggc agccgtccgg aaccgagccc gccgagggcc tgtgatgcag		300
		gcccgcctgg ccgagatccg gcccttggtc gagggccgagc tgaacgcgc catcgacgcg		360
		ctgccgcgg gcgatctgtc ggatgcgatg cgctatgccg tgcagggcgg caagcggctg		420
60		cgcgcgttcc tggatgatga gtcggcgcgc ctgcacgggc tggacgacga cgcacgcgtg		480

	ccggtcgccg	ccgcggtcga	ggcgctgcac	gcctacagct	tgggtccatga	cgacctgccc	540
	gcgatggatg	acgacgacct	gcggcgcggt	cagcccaccg	tccacgtcaa	atggaccgag	600
	gcgaccgcga	tccttgcggg	cgatgcgctg	cagacgctgg	ccttccagct	gctggccgat	660
	ccgcgcgtgg	gcgacgatgc	ggcgcggtatg	cggctggtcg	gttcgctggc	gcaggcatcg	720
5	ggggctgcgg	gcatggtctg	gggccaggcg	ctggacatcg	cggccgagac	ctcgggctg	780
	ccgctggatc	tggacgcgat	catccgcctg	caggggtggca	agaccggcgc	gctgatccgc	840
	tttgccgcga	ccgccgggcc	gctgatggcg	ggggcggaac	ctgccgcgct	ggacgattat	900
	gcgcaggccg	tcgggctggc	cttccagatc	gcggacgaca	tcctggacgt	cgagggctgc	960
	gaggccgcga	ccggcaagcg	cgtcggcaag	gatgcggatg	ccaacaaggc	gaccttcgtc	1020
10	tcgctgctgg	gcctcgaggg	ggcgcggtcc	gaggcgcgtc	gcctggccga	tgcggggcag	1080
	gacgcgctgg	cgggttacgg	cgatgctgcg	gggaaccttc	gggacctggc	gcgcttcgtg	1140
	atcgaacgcg	acagctgac	gccgccttcc	cgccaagggg	caagatgatg	accgacggac	1200
	ccgcaacccc	gatcctggac	cgcgtccagc	agccatccga	cctggcatcg	ctggacgatg	1260
	cgcagctgcg	cctgctggcg	gacgagctgc	gggccgagac	catcgacatc	gtcagccgca	1320
15	cgggcggtca	cctggggcgcg	gggctggggc	tggctgaact	gacggtcgcc	ctgcacgccg	1380
	tctttcgggc	gcccgcgcgac	aagatcgctc	gggacgtggg	gcatcaatgc	tatccccaca	1440
	agatcctgac	gggcaggcgg	gaccggatgc	gcacgctgcg	catgggcggc	gggctgtcgg	1500
	ggttcaccaa	gcggcaggaa	agcgcgttcg	atccgttcgg	tgcggggcac	agctcgacct	1560
	cgatctcggc	ggcgctgggc	ttcgcgatgg	cgcgtgaact	tggcggggat	cc	1612
20							
	<210>	158					
	<211>	78					
	<212>	PRT					
	<213>	副球菌属菌种(Paracoccus sp.)					
25	<400>	158					
	Met Ser Asp Ile Gln Thr Leu Ser Phe Glu Glu Ala Met Arg Glu Leu						
	1	5	10	15			
	Glu Ala Thr Val Gly Lys Leu Glu Thr Gly Glu Ala Thr Leu Glu Asp						
30		20	25	30			
	Ser Ile Ala Leu Tyr Glu Arg Gly Ala Ala Leu Arg Ala His Cys Glu						
	35	40	45				
35	Thr Arg Leu Arg Glu Ala Glu Glu Arg Val Glu Lys Ile Thr Leu Ala						
	50	55	60				
	Ala Asn Gly Gln Pro Ser Gly Thr Glu Pro Ala Glu Gly Leu						
	65	70	75				
40							
	<210>	159					
	<211>	287					
	<212>	PRT					
	<213>	副球菌属菌种(Paracoccus sp.)					
45	<400>	159					
	Met Gln Ala Arg Leu Ala Glu Ile Arg Pro Leu Val Glu Ala Glu Leu						
	1	5	10	15			
	Asn Ala Ala Ile Asp Ala Leu Pro Ala Gly Asp Leu Ser Asp Ala Met						
50		20	25	30			
	Arg Tyr Ala Val Gln Gly Gly Lys Arg Leu Arg Ala Phe Leu Val Met						
	35	40	45				
55	Glu Ser Ala Arg Leu His Gly Leu Asp Asp Asp Ala Ser Leu Pro Val						
	50	55	60				
	Ala Ala Ala Val Glu Ala Leu His Ala Tyr Ser Leu Val His Asp Asp						
60	65	70	75	80			

	Leu	Pro	Ala	Met	Asp	Asp	Asp	Asp	Leu	Arg	Arg	Gly	Gln	Pro	Thr	Val	
					85					90					95		
5	His	Val	Lys	Trp	Thr	Glu	Ala	Thr	Ala	Ile	Leu	Ala	Gly	Asp	Ala	Leu	
				100					105					110			
	Gln	Thr	Leu	Ala	Phe	Gln	Leu	Leu	Ala	Asp	Pro	Arg	Val	Gly	Asp	Asp	
			115					120					125				
10	Ala	Ala	Arg	Met	Arg	Leu	Val	Gly	Ser	Leu	Ala	Gln	Ala	Ser	Gly	Ala	
		130					135					140					
	Ala	Gly	Met	Val	Trp	Gly	Gln	Ala	Leu	Asp	Ile	Ala	Ala	Glu	Thr	Ser	
15	145					150					155					160	
	Gly	Val	Pro	Leu	Asp	Leu	Asp	Ala	Ile	Ile	Arg	Leu	Gln	Gly	Gly	Lys	
					165					170					175		
20	Thr	Gly	Ala	Leu	Ile	Arg	Phe	Ala	Ala	Thr	Ala	Gly	Pro	Leu	Met	Ala	
				180					185					190			
	Gly	Ala	Asp	Pro	Ala	Ala	Leu	Asp	Asp	Tyr	Ala	Gln	Ala	Val	Gly	Leu	
			195					200					205				
25	Ala	Phe	Gln	Ile	Ala	Asp	Asp	Ile	Leu	Asp	Val	Glu	Gly	Cys	Glu	Ala	
		210					215					220					
	Ala	Thr	Gly	Lys	Arg	Val	Gly	Lys	Asp	Ala	Asp	Ala	Asn	Lys	Ala	Thr	
30	225					230					235					240	
	Phe	Val	Ser	Leu	Leu	Gly	Leu	Glu	Gly	Ala	Arg	Ser	Glu	Ala	Arg	Arg	
					245					250					255		
35	Leu	Ala	Asp	Ala	Gly	Gln	Asp	Ala	Leu	Ala	Gly	Tyr	Gly	Asp	Ala	Ala	
				260					265					270			
	Gly	Asn	Leu	Arg	Asp	Leu	Ala	Arg	Phe	Val	Ile	Glu	Arg	Asp	Ser		
			275					280					285				
40																	
	<210>	160															
	<211>	142															
	<212>	PRT															
	<213>	副球菌属菌种 (Paracoccus sp.)															
45	<400>	160															
	Met	Met	Thr	Asp	Gly	Pro	Ala	Thr	Pro	Ile	Leu	Asp	Arg	Val	Gln	Gln	
	1				5					10					15		
	Pro	Ser	Asp	Leu	Ala	Ser	Leu	Asp	Asp	Ala	Gln	Leu	Arg	Leu	Leu	Ala	
50				20					25					30			
	Asp	Glu	Leu	Arg	Ala	Glu	Thr	Ile	Asp	Ile	Val	Ser	Arg	Thr	Gly	Gly	
			35					40					45				
55	His	Leu	Gly	Ala	Gly	Leu	Gly	Val	Val	Glu	Leu	Thr	Val	Ala	Leu	His	
		50					55					60					
	Ala	Val	Phe	Arg	Ala	Pro	Arg	Asp	Lys	Ile	Val	Trp	Asp	Val	Gly	His	
	65					70					75					80	
60																	

Gln Cys Tyr Pro His Lys Ile Leu Thr Gly Arg Arg Asp Arg Met Arg
 85 90 95

5 Thr Leu Arg Met Gly Gly Gly Leu Ser Gly Phe Thr Lys Arg Gln Glu
 100 105 110

Ser Ala Phe Asp Pro Phe Gly Ala Gly His Ser Ser Thr Ser Ile Ser
 115 120 125

10 Ala Ala Leu Gly Phe Ala Met Ala Arg Glu Leu Gly Gly Asp
 130 135 140

<210> 161
 <211> 6
 15 <212> PRT
 <213> 大豆慢生根瘤菌 (*Bradyrhizobium japonicum*)
 <400> 161
 Val His Asp Asp Leu Pro
 1 5

20 <210> 162
 <211> 6
 <212> PRT
 <213> 根瘤菌属菌种 (*Rhizobium* sp.)
 25 <400> 162
 Val His Asp Asp Leu Pro
 1 5

<210> 163
 30 <211> 6
 <212> PRT
 <213> 嗜热脂肪芽孢杆菌 (*Bacillus stearothermophilus*)
 <400> 163
 Ile His Asp Asp Leu Pro
 35 1 5

<210> 164
 <211> 6
 <212> PRT
 40 <213> 枯草芽孢杆菌 (*Bacillus subtilis*)
 <400> 164
 Ile His Asp Asp Leu Pro
 1 5

45 <210> 165
 <211> 6
 <212> PRT
 <213> 大肠杆菌 (*Escherichia coli*)
 <400> 165
 50 Ile His Asp Asp Leu Pro
 1 5

<210> 166
 <211> 6
 55 <212> PRT
 <213> 流感嗜血菌 (*Haemophilus influenzae*)
 <400> 166
 Ile His Asp Asp Leu Pro
 1 5

60

	<210>	167	
	<211>	16	
	<212>	DNA	
	<213>	人工的	
5	<220>		
	<221>	misc_feature	
	<222>	(5)..(5)	
	<223>	n 代表 C 或 T	
	<220>		
10	<221>	misc_feature	
	<222>	(8)..(8)	
	<223>	n 代表 C 或 T	
	<220>		
	<221>	misc_feature	
15	<222>	(11)..(11)	
	<223>	n 代表 C 或 T	
	<220>		
	<221>	misc_feature	
	<222>	(1)..(16)	
20	<223>	寡核苷酸	
	<400>	167	
	tccanganga nctgcc		16
	<210>	168	
25	<211>	5	
	<212>	PRT	
	<213>	大豆慢生根瘤菌(<i>Bradyrhizobium japonicum</i>)	
	<400>	168	
	Asp Asp Ile Leu Asp		
30	1	5	
	<210>	169	
	<211>	5	
	<212>	PRT	
35	<213>	根瘤菌属菌种(<i>Rhizobium</i> sp.)	
	<400>	169	
	Asp Asp Ile Leu Asp		
	1	5	
40	<210>	170	
	<211>	5	
	<212>	PRT	
	<213>	嗜热脂肪芽孢杆菌(<i>Bacillus stearothermophilus</i>)	
	<400>	170	
45	Asp Asp Ile Leu Asp		
	1	5	
50	<210>	171	
	<211>	5	
	<212>	PRT	
	<213>	枯草芽孢杆菌(<i>Bacillus subtilis</i>)	
	<400>	171	
55	Asp Asp Ile Leu Asp		
	1	5	
	<210>	172	
	<211>	5	
60	<212>	PRT	

	<213> 大肠杆菌 (Escherichia coli)	
	<400> 172	
	Asp Asp Ile Leu Asp	
5	1 5	
	<210> 173	
	<211> 5	
	<212> PRT	
10	<213> 流感嗜血菌 Haemophilus influenzae)	
	<400> 173	
	Asp Asp Ile Leu Asp	
	1 5	
	<210> 174	
15	<211> 15	
	<212> DNA	
	<213> 人工的	
	<220>	
	<221> misc_feature	
20	<222> (3)..(3)	
	<223> n 代表 C 或 T	
	<220>	
	<221> misc_feature	
	<222> (6)..(6)	
25	<223> n 代表 C 或 T	
	<220>	
	<221> misc_feature	
	<222> (15)..(15)	
	<223> n 代表 C 或 T	
30	<220>	
	<221> misc_feature	
	<222> (1)..(15)	
	<223> 寡核苷酸	
	<400> 174	
35	ganganatcc tggan	15
	<210> 175	
	<211> 1176	
	<212> DNA	
40	<213> 副球菌属菌种 (Paracoccus sp.)	
	<220>	
	<221> CDS	
	<222> (1)..(1173)	
	<223>	
45	<400> 175	
	atg gac ccc atc gtc atc acc ggc gcg atg cgc acc ccg atg ggg gca	48
	Met Asp Pro Ile Val Ile Thr Gly Ala Met Arg Thr Pro Met Gly Ala	
	1 5 10 15	
50	ttc cag ggc gat ctt gcc gcg atg gat gcc ccg acc ctt ggc gcg gcc	96
	Phe Gln Gly Asp Leu Ala Ala Met Asp Ala Pro Thr Leu Gly Ala Ala	
	20 25 30	
	gcg atc cgc gcc gcg ctg aac ggc ctg tcg ccc gac atg gtg gac gag	144
55	Ala Ile Arg Ala Ala Leu Asn Gly Leu Ser Pro Asp Met Val Asp Glu	
	35 40 45	
	gtg ctg atg ggc tgc gtc ctg ccc gcg ggc cag ggt cag gca ccg gca	192
60	Val Leu Met Gly Cys Val Leu Pro Ala Gly Gln Gly Gln Ala Pro Ala	
	50 55 60	

5	cgt cag gcg gcg ctt gac gcc gga ctg ccg ctg tcg gcg ggc gcg acc	240
	Arg Gln Ala Ala Leu Asp Ala Gly Leu Pro Leu Ser Ala Gly Ala Thr	
	65 70 75 80	
10	acc atc aac aag atg tgc gga tcg ggc atg aag gcc gcg atg ctg ggc	288
	Thr Ile Asn Lys Met Cys Gly Ser Gly Met Lys Ala Ala Met Leu Gly	
	85 90 95	
15	cat gac ctg atc gcc gcg gga tcg gcg ggc atc gtc gtc gcc ggc ggg	336
	His Asp Leu Ile Ala Ala Gly Ser Ala Gly Ile Val Val Ala Gly Gly	
	100 105 110	
20	atg gag agc atg tcg aac gcc ccc tac ctg ctg ccc aag gcg cgg tcg	384
	Met Glu Ser Met Ser Asn Ala Pro Tyr Leu Leu Pro Lys Ala Arg Ser	
	115 120 125	
25	ggg atg cgc atg ggc cat gac cgt gtg ctg gat cac atg ttc ctc gac	432
	Gly Met Arg Met Gly His Asp Arg Val Leu Asp His Met Phe Leu Asp	
	130 135 140	
30	ggg ttg gag gac gcc tat gac aag ggc cgc ctg atg ggc acc ttc gcc	480
	Gly Leu Glu Asp Ala Tyr Asp Lys Gly Arg Leu Met Gly Thr Phe Ala	
	145 150 155 160	
35	gag gat tgc gcc ggc gat cac ggt ttc acc cgc gag gcg cag gac gac	528
	Glu Asp Cys Ala Gly Asp His Gly Phe Thr Arg Glu Ala Gln Asp Asp	
	165 170 175	
40	tat gcg ctg acc agc ctg gcc cgc gcg cag gac gcc atc gcc agc ggt	576
	Tyr Ala Leu Thr Ser Leu Ala Arg Ala Gln Asp Ala Ile Ala Ser Gly	
	180 185 190	
45	gcc ttc gcc gcc gag atc gcg ccc gtg acc gtc acg gca cgc aag gtg	624
	Ala Phe Ala Ala Glu Ile Ala Pro Val Thr Val Thr Ala Arg Lys Val	
	195 200 205	
50	cag acc acc gtc gat acc gac gag atg ccc ggc aag gcc cgc ccc gag	672
	Gln Thr Thr Val Asp Thr Asp Glu Met Pro Gly Lys Ala Arg Pro Glu	
	210 215 220	
55	aag atc ccc cat ctg aag ccc gcc ttc cgt gac ggt ggc acg gtc acg	720
	Lys Ile Pro His Leu Lys Pro Ala Phe Arg Asp Gly Gly Thr Val Thr	
	225 230 235 240	
60	gcg gcg aac agc tcg tcg atc tcg gac ggg gcg gcg gcg ctg gtg atg	768
	Ala Ala Asn Ser Ser Ser Ile Ser Asp Gly Ala Ala Ala Leu Val Met	
	245 250 255	
65	atg cgc cag tcg cag gcc gag aag ctg ggc ctg acg ccg atc gcg cgg	816
	Met Arg Gln Ser Gln Ala Glu Lys Leu Gly Leu Thr Pro Ile Ala Arg	
	260 265 270	
70	atc atc ggt cat gcg acc cat gcc gac cgt ccc ggc ctg ttc ccg acg	864
	Ile Ile Gly His Ala Thr His Ala Asp Arg Pro Gly Leu Phe Pro Thr	
	275 280 285	
75	gcc ccc atc ggc gcg atg cgc aag ctg ctg gac cgc acg gac acc cgc	912
	Ala Pro Ile Gly Ala Met Arg Lys Leu Leu Asp Arg Thr Asp Thr Arg	
	290 295 300	

	ctt ggc gat tac gac ctg ttc gag gtg aac gag gca ttc gcc gtc gtc	960
	Leu Gly Asp Tyr Asp Leu Phe Glu Val Asn Glu Ala Phe Ala Val Val	
	305 310 315 320	
5	gcc atg atc gcg atg aag gag ctt ggc ctg cca cac gat gcc acg aac	1008
	Ala Met Ile Ala Met Lys Glu Leu Gly Leu Pro His Asp Ala Thr Asn	
	325 330 335	
10	atc aac ggc ggg gcc tgc gcg ctt ggg cat ccc atc ggc gcg tcg ggg	1056
	Ile Asn Gly Gly Ala Cys Ala Leu Gly His Pro Ile Gly Ala Ser Gly	
	340 345 350	
15	gcg cgg atc atg gtc acg ctg ctg aac gcg atg gcg gcg cgg ggc gcg	1104
	Ala Arg Ile Met Val Thr Leu Leu Asn Ala Met Ala Ala Arg Gly Ala	
	355 360 365	
20	acg cgc ggg gcc gca tcc gtc tgc atc ggc ggg ggc gag gcg acg gcc	1152
	Thr Arg Gly Ala Ala Ser Val Cys Ile Gly Gly Gly Glu Ala Thr Ala	
	370 375 380	
25	atc gcg ctg gaa cgg ctg agc taa	1176
	Ile Ala Leu Glu Arg Leu Ser	
	385 390	
	<210> 176	
	<211> 391	
	<212> PRT	
	<213> 副球菌属菌种 (Paracoccus sp.)	
30	<400> 176	
	Met Asp Pro Ile Val Ile Thr Gly Ala Met Arg Thr Pro Met Gly Ala	
	1 5 10 15	
35	Phe Gln Gly Asp Leu Ala Ala Met Asp Ala Pro Thr Leu Gly Ala Ala	
	20 25 30	
	Ala Ile Arg Ala Ala Leu Asn Gly Leu Ser Pro Asp Met Val Asp Glu	
	35 40 45	
40	Val Leu Met Gly Cys Val Leu Pro Ala Gly Gln Gly Gln Ala Pro Ala	
	50 55 60	
45	Arg Gln Ala Ala Leu Asp Ala Gly Leu Pro Leu Ser Ala Gly Ala Thr	
	65 70 75 80	
	Thr Ile Asn Lys Met Cys Gly Ser Gly Met Lys Ala Ala Met Leu Gly	
	85 90 95	
50	His Asp Leu Ile Ala Ala Gly Ser Ala Gly Ile Val Val Ala Gly Gly	
	100 105 110	
	Met Glu Ser Met Ser Asn Ala Pro Tyr Leu Leu Pro Lys Ala Arg Ser	
	115 120 125	
55	Gly Met Arg Met Gly His Asp Arg Val Leu Asp His Met Phe Leu Asp	
	130 135 140	
60	Gly Leu Glu Asp Ala Tyr Asp Lys Gly Arg Leu Met Gly Thr Phe Ala	
	145 150 155 160	

```

      Glu Asp Cys Ala Gly Asp His Gly Phe Thr Arg Glu Ala Gln Asp Asp
                165                      170                      175
5  Tyr Ala Leu Thr Ser Leu Ala Arg Ala Gln Asp Ala Ile Ala Ser Gly
                180                      185                      190

      Ala Phe Ala Ala Glu Ile Ala Pro Val Thr Val Thr Ala Arg Lys Val
                195                      200                      205
10 Gln Thr Thr Val Asp Thr Asp Glu Met Pro Gly Lys Ala Arg Pro Glu
                210                      215                      220

      Lys Ile Pro His Leu Lys Pro Ala Phe Arg Asp Gly Gly Thr Val Thr
15 225                      230                      235                      240

      Ala Ala Asn Ser Ser Ser Ile Ser Asp Gly Ala Ala Ala Leu Val Met
                245                      250                      255

20 Met Arg Gln Ser Gln Ala Glu Lys Leu Gly Leu Thr Pro Ile Ala Arg
                260                      265                      270

      Ile Ile Gly His Ala Thr His Ala Asp Arg Pro Gly Leu Phe Pro Thr
25 275                      280                      285

      Ala Pro Ile Gly Ala Met Arg Lys Leu Leu Asp Arg Thr Asp Thr Arg
                290                      295                      300

30 Leu Gly Asp Tyr Asp Leu Phe Glu Val Asn Glu Ala Phe Ala Val Val
305                      310                      315                      320

      Ala Met Ile Ala Met Lys Glu Leu Gly Leu Pro His Asp Ala Thr Asn
35 325                      330                      335

      Ile Asn Gly Gly Ala Cys Ala Leu Gly His Pro Ile Gly Ala Ser Gly
                340                      345                      350

40 Ala Arg Ile Met Val Thr Leu Leu Asn Ala Met Ala Ala Arg Gly Ala
355                      360                      365

      Thr Arg Gly Ala Ala Ser Val Cys Ile Gly Gly Gly Glu Ala Thr Ala
                370                      375                      380

45 Ile Ala Leu Glu Arg Leu Ser
385                      390

<210> 177
<211> 1980
50 <212> DNA
    <213> 副球菌属菌种(Paracoccus sp.)
    <220>
    <221> 基因
    <222> (1)..(1980)
55 <223>
    <400> 177
atgaccaaag ccgtaatcgt atctgccgca cgtacccccg tcggcagctt catggggcgca 60
ttcgccaatg tccccgcaca tgatctgggc gccgccgtcc tgcgcgaggt cgtggcccgcg 120
gccggtgtcg accccgccga ggtcagcgag acgatccttg gccaggtgct gaccgcgcgcg 180
60 cagggccaga accccgcgcg ccaggcgcat atcaatgcgg gcctgcccga ggaatcggcg 240

```


	gcgtggctca	tcaaccaggt	ctgcggctcg	gggctgcgcg	ccgtcgcgct	ggcggcgag	300
	cagggtcatgc	tgggcgatgc	gcagatcggt	ctggcgggg	gccaggagag	catgtcgcgtg	360
	tcgacccatg	cgcctatct	gcgcgcgggc	cagaagatgg	gcgacatgaa	gatgatcgac	420
	accatgatcc	gcgacgggct	gtgggatgcc	ttcaacggct	atcacatggg	tcagaccgcc	480
5	gagaacgtgg	ccgaccagtg	gtcgatcagc	cgcgaccagc	aggacgaatt	cgccctggct	540
	tcgcagaaca	aggccgaggc	cgcgcagaat	gcgggccgct	tcgatgacga	aatcgtcgcc	600
	tataccgtca	agggccgcaa	gggcgacacg	gtcgtcgaca	aggacgaata	catccgccac	660
	ggcgccacga	tcgagggcat	gcagaagctg	cgccccgcct	tcaccaagga	aggctcggtc	720
	acggcgggca	acgcgtcggg	cctgaacgac	ggcgcgggcg	ccgtcatggt	catgtccgag	780
10	gacgaggccg	cacgccgcgg	gctgacgccg	ctggcgcgca	tcgcctccta	tgcgacggcg	840
	ggcctcgacc	cggcgatcat	gggcaccggg	ccgatccct	ccagccgcaa	ggcgtggaa	900
	aaggcgggct	ggtcggtcgg	cgacctggac	ctggctcagg	cgaacgaggc	ctttgccgcg	960
	caggcctgcg	ccgtgaacaa	ggacatgggc	tgggatccgt	ccatcgtgaa	cgtcaacggc	1020
	ggcgcgatcg	ccatcgcca	cccgatcggc	gcctcggggg	cgcgatcct	gaacaccctg	1080
15	ctgttcgaga	tgcagcgccg	cgacgccaa	aagggccttg	cgacgctgtg	catcggcgcc	1140
	ggcatgggcg	tcgccatgtg	cctcgaacgc	tgaacgaccg	gcgtgtgcgc	aatttaattg	1200
	cgcacacgcc	ccctgcaaag	tagcaatgtt	ttacgataac	gaatgaaggg	gggaatcatg	1260
	tccaaggtag	cactggtcac	cgcgcgatcg	cgcgccatcg	gcgccgagat	ctgcaaggcg	1320
	cttcaggccg	caggctatac	cgtcgcgcgcg	aactatgccg	gcaatgacga	cgcgccaaag	1380
20	gccttcaccg	aggaaaccgg	catcaagacc	tacaagtggg	cggtcgccga	ttacgatgcc	1440
	tgcaaggccg	gcacgcgcca	ggtcgaagag	gatctggggc	cgatcgccgt	gctgatcaac	1500
	aatgccggga	tcacccgcga	cgcgcccttc	cacaagatga	cgcccgagaa	gtgcaaggag	1560
	gtcatcgaca	ccaacctgac	cggcaccttc	aacatgacct	atccggtctg	gccgggcatg	1620
	cgcgaaacga	agttcgggacg	cgatcatcaac	atcagctcga	tcaacgggca	gaagggccag	1680
25	ttcgggcagg	cgaactatgc	cgcgcccaag	gcgggcgacc	tgggcttcac	caagtgcgtg	1740
	gcgcaggaag	gcgcgcgcaa	caacatcacc	gtcaacgcga	tctgcccccg	ctatatcgcg	1800
	acggacatgg	tgatggccgt	tcccgaacag	gtccgcgagg	ggatcatcgc	gcagatcccc	1860
	gtcggccgct	tgggcgagcc	gtccgagatc	gcgcgctgcg	tggtgttcct	ggcctccgac	1920
	gatgcgggct	tcgtcacagg	ctcgaccatc	acggcgcaatg	gcggccagta	ctacatctga	1980
30	<210> 178						
	<211> 390						
	<212> PRT						
	<213> 副球菌属菌种 (Paracoccus sp.)						
35	<400> 178						
	Met Thr Lys Ala Val Ile Val Ser Ala Ala Arg Thr Pro Val Gly Ser						
	1 5 10 15						
40	Phe Met Gly Ala Phe Ala Asn Val Pro Ala His Asp Leu Gly Ala Ala						
	20 25 30						
	Val Leu Arg Glu Val Val Ala Arg Ala Gly Val Asp Pro Ala Glu Val						
	35 40 45						
45	Ser Glu Thr Ile Leu Gly Gln Val Leu Thr Ala Ala Gln Gly Gln Asn						
	50 55 60						
	Pro Ala Arg Gln Ala His Ile Asn Ala Gly Leu Pro Lys Glu Ser Ala						
	65 70 75 80						
50	Ala Trp Leu Ile Asn Gln Val Cys Gly Ser Gly Leu Arg Ala Val Ala						
	85 90 95						
	Leu Ala Ala Gln Gln Val Met Leu Gly Asp Ala Gln Ile Val Leu Ala						
	100 105 110						
55	Gly Gly Gln Glu Ser Met Ser Leu Ser Thr His Ala Ala Tyr Leu Arg						
	115 120 125						
60	Ala Gly Gln Lys Met Gly Asp Met Lys Met Ile Asp Thr Met Ile Arg						
	130 135 140						

Asp Gly Leu Trp Asp Ala Phe Asn Gly Tyr His Met Gly Gln Thr Ala
 145 150 155 160
 5 Glu Asn Val Ala Asp Gln Trp Ser Ile Ser Arg Asp Gln Gln Asp Glu
 165 170 175
 Phe Ala Leu Ala Ser Gln Asn Lys Ala Glu Ala Ala Gln Asn Ala Gly
 180 185 190
 10 Arg Phe Asp Asp Glu Ile Val Ala Tyr Thr Val Lys Gly Arg Lys Gly
 195 200 205
 Asp Thr Val Val Asp Lys Asp Glu Tyr Ile Arg His Gly Ala Thr Ile
 210 215 220
 15 Glu Gly Met Gln Lys Leu Arg Pro Ala Phe Thr Lys Glu Gly Ser Val
 225 230 235 240
 20 Thr Ala Gly Asn Ala Ser Gly Leu Asn Asp Gly Ala Ala Ala Val Met
 245 250 255
 Val Met Ser Glu Asp Glu Ala Ala Arg Arg Gly Leu Thr Pro Leu Ala
 260 265 270
 25 Arg Ile Ala Ser Tyr Ala Thr Ala Gly Leu Asp Pro Ala Ile Met Gly
 275 280 285
 30 Thr Gly Pro Ile Pro Ser Ser Arg Lys Ala Leu Glu Lys Ala Gly Trp
 290 295 300
 Ser Val Gly Asp Leu Asp Leu Val Glu Ala Asn Glu Ala Phe Ala Ala
 305 310 315 320
 35 Gln Ala Cys Ala Val Asn Lys Asp Met Gly Trp Asp Pro Ser Ile Val
 325 330 335
 Asn Val Asn Gly Gly Ala Ile Ala Ile Gly His Pro Ile Gly Ala Ser
 340 345 350
 Gly Ala Arg Ile Leu Asn Thr Leu Leu Phe Glu Met Gln Arg Arg Asp
 355 360 365
 45 Ala Lys Lys Gly Leu Ala Thr Leu Cys Ile Gly Gly Gly Met Gly Val
 370 375 380
 Ala Met Cys Leu Glu Arg
 385 390
 50 <210> 179
 <211> 240
 <212> PRT
 <213> 副球菌属菌种(Paracoccus sp.)
 55 <400> 179
 Met Ser Lys Val Ala Leu Val Thr Gly Gly Ser Arg Gly Ile Gly Ala
 1 5 10 15
 Glu Ile Cys Lys Ala Leu Gln Ala Ala Gly Tyr Thr Val Ala Ala Asn
 20 25 30
 60

	Tyr	Ala	Gly	Asn	Asp	Asp	Ala	Ala	Lys	Ala	Phe	Thr	Glu	Glu	Thr	Gly	
			35					40					45				
5	Ile	Lys	Thr	Tyr	Lys	Trp	Ser	Val	Ala	Asp	Tyr	Asp	Ala	Cys	Lys	Ala	
		50					55					60					
	Gly	Ile	Ala	Gln	Val	Glu	Glu	Asp	Leu	Gly	Pro	Ile	Ala	Val	Leu	Ile	
	65					70					75					80	
10	Asn	Asn	Ala	Gly	Ile	Thr	Arg	Asp	Ala	Pro	Phe	His	Lys	Met	Thr	Pro	
					85					90					95		
	Glu	Lys	Trp	Lys	Glu	Val	Ile	Asp	Thr	Asn	Leu	Thr	Gly	Thr	Phe	Asn	
				100					105					110			
15	Met	Thr	His	Pro	Val	Trp	Pro	Gly	Met	Arg	Glu	Arg	Lys	Phe	Gly	Arg	
			115					120					125				
	Val	Ile	Asn	Ile	Ser	Ser	Ile	Asn	Gly	Gln	Lys	Gly	Gln	Phe	Gly	Gln	
20		130					135					140					
	Ala	Asn	Tyr	Ala	Ala	Ala	Lys	Ala	Gly	Asp	Leu	Gly	Phe	Thr	Lys	Ser	
	145					150					155					160	
25	Leu	Ala	Gln	Glu	Gly	Ala	Arg	Asn	Asn	Ile	Thr	Val	Asn	Ala	Ile	Cys	
					165					170					175		
	Pro	Gly	Tyr	Ile	Ala	Thr	Asp	Met	Val	Met	Ala	Val	Pro	Glu	Gln	Val	
				180					185					190			
30	Arg	Glu	Gly	Ile	Ile	Ala	Gln	Ile	Pro	Val	Gly	Arg	Leu	Gly	Glu	Pro	
			195					200					205				
	Ser	Glu	Ile	Ala	Arg	Cys	Val	Val	Phe	Leu	Ala	Ser	Asp	Asp	Ala	Gly	
35		210					215					220					
	Phe	Val	Thr	Gly	Ser	Thr	Ile	Thr	Ala	Asn	Gly	Gly	Gln	Tyr	Tyr	Ile	
	225					230					235					240	
40	<210> 180																
	<211> 729																
	<212> DNA																
	<213> Paracoccus carotinifaciens E-396																
	<220>																
45	<221> CDS																
	<222> (1)..(726)																
	<223>																
	<400> 180																
	atg	agc	gca	cat	gcc	ctg	ccc	aag	gca	gat	ctg	acc	gcc	acc	agt	ttg	48
50	Met	Ser	Ala	His	Ala	Leu	Pro	Lys	Ala	Asp	Leu	Thr	Ala	Thr	Ser	Leu	
	1				5					10					15		
	atc	gtc	tcg	ggc	ggc	atc	atc	gcc	gcg	tgg	ctg	gcc	ctg	cat	gtg	cat	96
	Ile	Val	Ser	Gly	Gly	Ile	Ile	Ala	Ala	Trp	Leu	Ala	Leu	His	Val	His	
55				20					25				30				
	gcg	ctg	tgg	ttt	ctg	gac	gcg	gcg	gcg	cat	ccc	atc	ctg	gcg	gtc	gcg	144
60	Ala	Leu	Trp	Phe	Leu	Asp	Ala	Ala	Ala	His	Pro	Ile	Leu	Ala	Val	Ala	

	35	40	45	
5	aat ttc ctg ggg ctg acc tgg ctg tcg gtc ggt ctg ttc atc atc gcg Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala 50 55 60			192
10	cat gac gcg atg cat ggg tcg gtc gtg ccg ggg cgc ccg cgc gcc aat His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn 65 70 75 80			240
	gcg gcg atg ggc cag ctt gtc ctg tgg ctg tat gcc gga ttt tcc tgg Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp 85 90 95			288
15	cgc aag atg atc gtc aag cac atg gcc cat cat cgc cat gcc gga acc Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr 100 105 110			336
20	gac gac gac cca gat ttc gac cat ggc ggc ccg gtc cgc tgg tac gcc Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala 115 120 125			384
25	cgc ttc atc ggc acc tat ttc ggc tgg cgc gag ggg ctg ctg ctg ccc Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro 130 135 140			432
30	gtc atc gtg acg gtc tat gcg ctg atg ttg ggg gat cgc tgg atg tac Val Ile Val Thr Val Tyr Ala Leu Met Leu Gly Asp Arg Trp Met Tyr 145 150 155 160			480
	gtg gtc ttc tgg ccg ttg ccg tcg atc ctg gcg tcg atc cag ctg ttc Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe 165 170 175			528
35	gtg ttc ggc atc tgg ctg ccg cac cgc ccc ggc cac gac gcg ttc ccg Val Phe Gly Ile Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro 180 185 190			576
40	gac cgc cac aat gcg cgg tcg tcg cgg atc agc gac ccc gtg tcg ctg Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu 195 200 205			624
	ctg acc tgc ttt cac ttt ggc ggt tat cat cac gaa cac cac ctg cac Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His 210 215 220			672
45	ccg acg gtg cct tgg tgg cgc ctg ccc agc acc cgc acc aag ggg gac Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp 225 230 235 240			720
50	acc gca tga Thr Ala			729
55	<210> 181 <211> 242 <212> PRT <213> Paracoccus carotinifaciens E-396 <400> 181			
60	Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu 1 5 10 15			

	Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His	
	20 25 30	
5	Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Val Ala	
	35 40 45	
	Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala	
10	50 55 60	
	His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn	
	65 70 75 80	
15	Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp	
	85 90 95	
	Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr	
	100 105 110	
20	Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala	
	115 120 125	
	Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro	
	130 135 140	
25	Val Ile Val Thr Val Tyr Ala Leu Met Leu Gly Asp Arg Trp Met Tyr	
	145 150 155 160	
	Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe	
	165 170 175	
30	Val Phe Gly Ile Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro	
	180 185 190	
	Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu	
35	195 200 205	
	Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His	
	210 215 220	
40	Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp	
	225 230 235 240	
	Thr Ala	
45	<210> 182	
	<211> 510	
	<212> DNA	
	<213> 副球菌属菌种 (Paracoccus sp.)	
50	<220>	
	<221> CDS	
	<222> (1)..(507)	
	<223>	
	<400> 182	
55	atg agc act tgg gcc gca atc ctg acc gtc atc ctg acc gtc gcc gcg	48
	Met Ser Thr Trp Ala Ala Ile Leu Thr Val Ile Leu Thr Val Ala Ala	
	1 5 10 15	
60	atg gag ctg acg gcc tac tcc gtc cat cgg tgg atc atg cat ggc ccc	96
	Met Glu Leu Thr Ala Tyr Ser Val His Arg Trp Ile Met His Gly Pro	

	20	25	30	
5	ctg ggc tgg ggc tgg cat aaa tcg cac cac gac gag gat cac gac cac Leu Gly Trp Gly Trp His Lys Ser His His Asp Glu Asp His Asp His	144		
10	gcg ctc gag aag aac gac ctc tat ggc gtc atc ttc gcg gta atc tcg Ala Leu Glu Lys Asn Asp Leu Tyr Gly Val Ile Phe Ala Val Ile Ser	192		
15	atc gtg ctg ttc gcg atc ggc gcg atg ggg tcg gat ctg gcc tgg tgg Ile Val Leu Phe Ala Ile Gly Ala Met Gly Ser Asp Leu Ala Trp Trp	240		
20	ctg gcg gtg ggg gtc acc tgc tac ggg ctg atc tac tat ttc ctg cat Leu Ala Val Gly Val Thr Cys Tyr Gly Leu Ile Tyr Tyr Phe Leu His	288		
25	gac ggc ttg gtg cat ggg cgc tgg ccg ttc cgc tat gtc ccc aag cgc Asp Gly Leu Val His Gly Arg Trp Pro Phe Arg Tyr Val Pro Lys Arg	336		
30	ggc tat ctt cgt cgc gtc tac cag gca cac agg atg cat cac gcg gtc Gly Tyr Leu Arg Arg Val Tyr Gln Ala His Arg Met His His Ala Val	384		
35	cat ggc cgc gag aac tgc gtc agc ttc ggt ttc atc tgg gcg ccc tcg His Gly Arg Glu Asn Cys Val Ser Phe Gly Phe Ile Trp Ala Pro Ser	432		
40	gtc gac agc ctc aag gca gag ctg aaa cgc tcg ggc gcg ctg ctg aag Val Asp Ser Leu Lys Ala Glu Leu Lys Arg Ser Gly Ala Leu Leu Lys	480		
45	gac cgc gaa ggg gcg gat cgc aat aca tga Asp Arg Glu Gly Ala Asp Arg Asn Thr	510		
50	<210> 183 <211> 169 <212> PRT <213> 副球菌属菌种(Paracoccus sp.) <400> 183			
55	Met Ser Thr Trp Ala Ala Ile Leu Thr Val Ile Leu Thr Val Ala Ala 1 5 10 15			
60	Met Glu Leu Thr Ala Tyr Ser Val His Arg Trp Ile Met His Gly Pro 20 25 30			
65	Leu Gly Trp Gly Trp His Lys Ser His His Asp Glu Asp His Asp His 35 40 45			
70	Ala Leu Glu Lys Asn Asp Leu Tyr Gly Val Ile Phe Ala Val Ile Ser 50 55 60			
75	Ile Val Leu Phe Ala Ile Gly Ala Met Gly Ser Asp Leu Ala Trp Trp 65 70 75 80			
80	Leu Ala Val Gly Val Thr Cys Tyr Gly Leu Ile Tyr Tyr Phe Leu His			

	85	90	95	
	Asp Gly Leu Val His Gly Arg Trp Pro Phe Arg Tyr Val Pro Lys Arg			
	100	105	110	
5	Gly Tyr Leu Arg Arg Val Tyr Gln Ala His Arg Met His His Ala Val			
	115	120	125	
	His Gly Arg Glu Asn Cys Val Ser Phe Gly Phe Ile Trp Ala Pro Ser			
10	130	135	140	
	Val Asp Ser Leu Lys Ala Glu Leu Lys Arg Ser Gly Ala Leu Leu Lys			
	145	150	155	160
15	Asp Arg Glu Gly Ala Asp Arg Asn Thr			
	165			
	<210> 184			
	<211> 888			
20	<212> DNA			
	<213> 副球菌属菌种(Paracoccus sp.)			
	<220>			
	<221> CDS			
	<222> (1)..(885)			
25	<223>			
	<400> 184			
	atg acg ccc aag cag caa ttc ccc cta cgc gat ctg gtc gag atc agg			48
	Met Thr Pro Lys Gln Gln Phe Pro Leu Arg Asp Leu Val Glu Ile Arg			
	1 5 10 15			
30	ctg gcg cag atc tcg ggc cag ttc ggc gtg gtc tcg gcc ccg ctc ggc			96
	Leu Ala Gln Ile Ser Gly Gln Phe Gly Val Val Ser Ala Pro Leu Gly			
	20 25 30			
	gcg gcc atg agc gat gcc gcc ctg tcc ccc ggc aaa cgc ttt cgc gcc			144
35	Ala Ala Met Ser Asp Ala Ala Leu Ser Pro Gly Lys Arg Phe Arg Ala			
	35 40 45			
	gtg ctg atg ctg atg gtc gcc gaa agc tcg ggc ggg gtc tgc gat gcg			192
40	Val Leu Met Leu Met Val Ala Glu Ser Ser Gly Gly Val Cys Asp Ala			
	50 55 60			
	atg gtc gat gcc gcc tgc gcg gtc gag atg gtc cat gcc gca tcg ctg			240
	Met Val Asp Ala Ala Cys Ala Val Glu Met Val His Ala Ala Ser Leu			
	65 70 75 80			
45	atc ttc gac gac atg ccc tgc atg gac gat gcc agg acc cgt cgc ggt			288
	Ile Phe Asp Asp Met Pro Cys Met Asp Asp Ala Arg Thr Arg Arg Gly			
	85 90 95			
50	cag ccc gcc acc cat gtc gcc cat ggc gag ggg cgc gcg gtg ctt gcg			336
	Gln Pro Ala Thr His Val Ala His Gly Glu Gly Arg Ala Val Leu Ala			
	100 105 110			
	ggc atc gcc ctg atc acc gag gcc atg cgg att ttg ggc gag gcg cgc			384
55	Gly Ile Ala Leu Ile Thr Glu Ala Met Arg Ile Leu Gly Glu Ala Arg			
	115 120 125			
	ggc gcg acg ccg gat cag cgc gca agg ctg gtc gca tcc atg tcg cgc			432
60	Gly Ala Thr Pro Asp Gln Arg Ala Arg Leu Val Ala Ser Met Ser Arg			
	130 135 140			

	gcg atg gga ccg gtg ggg ctg tgc gca ggg cag gat ctg gac ctg cac	480
	Ala Met Gly Pro Val Gly Leu Cys Ala Gly Gln Asp Leu Asp Leu His	
	145 150 155 160	
5	gcc ccc aag gac gcc gcc ggg atc gaa cgt gaa cag gac ctc aag acc	528
	Ala Pro Lys Asp Ala Ala Gly Ile Glu Arg Glu Gln Asp Leu Lys Thr	
	165 170 175	
10	ggc gtg ctg ttc gtc gcg ggc ctc gag atg ctg tcc att att aag ggt	576
	Gly Val Leu Phe Val Ala Gly Leu Glu Met Leu Ser Ile Ile Lys Gly	
	180 185 190	
15	ctg gac aag gcc gag acc gag cag ctc atg gcc ttc ggg cgt cag ctt	624
	Leu Asp Lys Ala Glu Thr Glu Gln Leu Met Ala Phe Gly Arg Gln Leu	
	195 200 205	
	ggt cgg gtc ttc cag tcc tat gac gac ctg ctg gac gtg atc ggc gac	672
	Gly Arg Val Phe Gln Ser Tyr Asp Asp Leu Leu Asp Val Ile Gly Asp	
	210 215 220	
20	aag gcc agc acc ggc aag gat acg ggg cgc gac acc gcc gcc ccc ggc	720
	Lys Ala Ser Thr Gly Lys Asp Thr Gly Arg Asp Thr Ala Ala Pro Gly	
	225 230 235 240	
25	cca aag cgc ggc ctg atg gcg gtc gga cag atg ggc gac gtg gcg cag	768
	Pro Lys Arg Gly Leu Met Ala Val Gly Gln Met Gly Asp Val Ala Gln	
	245 250 255	
30	cat tac cgc gcc agc cgc gcg caa ctg gac gag ctg atg cgc acc cgg	816
	His Tyr Arg Ala Ser Arg Ala Gln Leu Asp Glu Leu Met Arg Thr Arg	
	260 265 270	
35	ctg ttc cgc ggg ggg cag atc gcg gac ctg ctg gcc cgc gtg ctg ccg	864
	Leu Phe Arg Gly Gly Gln Ile Ala Asp Leu Leu Ala Arg Val Leu Pro	
	275 280 285	
	cat gac atc cgc cgc agc gcc tag	888
	His Asp Ile Arg Arg Ser Ala	
	290 295	
40	<210> 185	
	<211> 295	
	<212> PRT	
	<213> 副球菌属菌种(Paracoccus sp.)	
45	<400> 185	
	Met Thr Pro Lys Gln Gln Phe Pro Leu Arg Asp Leu Val Glu Ile Arg	
	1 5 10 15	
50	Leu Ala Gln Ile Ser Gly Gln Phe Gly Val Val Ser Ala Pro Leu Gly	
	20 25 30	
	Ala Ala Met Ser Asp Ala Ala Leu Ser Pro Gly Lys Arg Phe Arg Ala	
	35 40 45	
55	Val Leu Met Leu Met Val Ala Glu Ser Ser Gly Gly Val Cys Asp Ala	
	50 55 60	
	Met Val Asp Ala Ala Cys Ala Val Glu Met Val His Ala Ala Ser Leu	
	65 70 75 80	
60	Ile Phe Asp Asp Met Pro Cys Met Asp Asp Ala Arg Thr Arg Arg Gly	

	85	90	95
	Gln Pro Ala Thr His Val Ala His	Gly Glu Gly Arg Ala Val Leu Ala	
	100	105	110
5	Gly Ile Ala Leu Ile Thr Glu Ala Met Arg Ile Leu Gly Glu Ala Arg		
	115	120	125
	Gly Ala Thr Pro Asp Gln Arg Ala Arg Leu Val Ala Ser Met Ser Arg		
10	130	135	140
	Ala Met Gly Pro Val Gly Leu Cys Ala Gly Gln Asp Leu Asp Leu His		
	145	150	155
15	Ala Pro Lys Asp Ala Ala Gly Ile Glu Arg Glu Gln Asp Leu Lys Thr		
	165	170	175
	Gly Val Leu Phe Val Ala Gly Leu Glu Met Leu Ser Ile Ile Lys Gly		
20	180	185	190
	Leu Asp Lys Ala Glu Thr Glu Gln Leu Met Ala Phe Gly Arg Gln Leu		
	195	200	205
	Gly Arg Val Phe Gln Ser Tyr Asp Asp Leu Leu Asp Val Ile Gly Asp		
25	210	215	220
	Lys Ala Ser Thr Gly Lys Asp Thr Gly Arg Asp Thr Ala Ala Pro Gly		
	225	230	235
30	Pro Lys Arg Gly Leu Met Ala Val Gly Gln Met Gly Asp Val Ala Gln		
	245	250	255
	His Tyr Arg Ala Ser Arg Ala Gln Leu Asp Glu Leu Met Arg Thr Arg		
35	260	265	270
	Leu Phe Arg Gly Gly Gln Ile Ala Asp Leu Leu Ala Arg Val Leu Pro		
	275	280	285
40	His Asp Ile Arg Arg Ser Ala		
	290	295	
	<210> 186		
	<211> 30		
	<212> DNA		
45	<213> 人工的		
	<220>		
	<221> misc_feature		
	<222> (1)..(30)		
	<223> 引物		
50	<400> 186		
	aaggcctcat atgagcgcac atgccctgcc		
	<210> 187		
	<211> 28		
55	<212> DNA		
	<213> 人工的		
	<220>		
	<221> misc_feature		
	<222> (1)..(28)		
60	<223> 引物		

30

	<400> 187		
	cgggatcctc atgcggtgtc ccccttgg	28	
5	<210> 188		
	<211> 30		
	<212> DNA		
	<213> 人工的		
	<220>		
	<221> misc_feature		
10	<222> (1)..(30)		
	<223> 引物		
	<400> 188		
	aaggcctcat atgagcactt gggccgcaat	30	
15	<210> 189		
	<211> 30		
	<212> DNA		
	<213> 人工的		
	<220>		
20	<221> misc_feature		
	<222> (1)..(30)		
	<223> 引物		
	<400> 189		
	aggatcctca tgtattgcga tccgcccctt	30	
25	<210> 190		
	<211> 52		
	<212> DNA		
	<213> 人工的		
	<220>		
30	<221> misc_feature		
	<222> (1)..(52)		
	<223> 多接头		
	<400> 190		
	gtgcagcctc aggtcgacat atgcggccgc atccggatcc ctctctctcc ag	52	
35	<210> 191		
	<211> 52		
	<212> DNA		
	<213> 人工的		
40	<220>		
	<221> misc_feature		
	<222> (1)..(52)		
	<223> 多接头		
	<400> 191		
45	cacgtcggag tccagctgta tacgccggcg taggcctagg gaggaggagg tc	52	
	<210> 192		
	<211> 52		
	<212> DNA		
50	<213> 人工的		
	<220>		
	<221> misc_feature		
	<222> (1)..(52)		
	<223> 多接头		
55	<400> 192		
	gtgcaggagg aggtcgacat atgcggccgc atccggatcc ctgaggctcc ag	52	
	<210> 193		
	<211> 52		
60	<212> DNA		

```

    <213> 人工的
    <220>
    <221> misc_feature
    <222> (1)..(52)
5  <223> 多接头
    <400> 193
    cacgtcctcc tccagctgta tacgccggcg taggcctagg gactccgagg tc          52
    <210> 194
    <211> 52
10 <212> DNA
    <213> 人工的
    <220>
    <221> misc_feature
    <222> (1)..(52)
15 <223> 多接头
    <400> 194
    ctggagcctc aggtcgacat atgcggccgc atccggatcc ctctcctgc ac          52
    <210> 195
20 <211> 52
    <212> DNA
    <213> 人工的
    <220>
    <221> misc_feature
25 <222> (1)..(52)
    <223> 多接头
    <400> 195
    gacctcggag tccagctgta tacgccggcg taggcctagg gaggaggacg tg          52
    <210> 196
30 <211> 52
    <212> DNA
    <213> 人工的
    <220>
35 <221> misc_feature
    <222> (1)..(52)
    <223> 多接头
    <400> 196
    ctggaggagg aggtcgacat atgcggccgc atccggatcc ctgaggctgc ac          52
40 <210> 197
    <211> 52
    <212> DNA
    <213> 人工的
45 <220>
    <221> misc_feature
    <222> (1)..(52)
    <223> 多接头
    <400> 197
50 gacctcctcc tccagctgta tacgccggcg taggcctagg gactccgacg tg          52

```