



# (12) 发明专利申请

(10) 申请公布号 CN 102021206 A

(43) 申请公布日 2011. 04. 20

(21) 申请号 201010530644. 3

苏珊·M·马德里德

(22) 申请日 2004. 01. 15

乔恩·D·米克尔森 乔恩·B·索伊

(30) 优先权数据

(74) 专利代理机构 隆天国际知识产权代理有限公司 72003

0301120. 2 2003. 01. 17 GB

代理人 吴小瑛 吕俊清

0301117. 8 2003. 01. 17 GB

0301121. 0 2003. 01. 17 GB

(51) Int. Cl.

0301118. 6 2003. 01. 17 GB

*C12P 7/62* (2006. 01)

0301122. 8 2003. 01. 17 GB

*C12P 7/64* (2006. 01)

0301119. 4 2003. 01. 17 GB

0330016. 7 2003. 12. 24 GB

60/489, 441 2003. 07. 23 US

(62) 分案原申请数据

200480006382. 3 2004. 01. 15

(83) 生物保藏信息

NCIMB 41205 2003. 12. 22

NCIMB 41204 2003. 12. 22

(71) 申请人 丹尼斯科公司

地址 丹麦哥本哈根

(72) 发明人 阿尔诺·D·克赖杰

权利要求书 3 页 说明书 58 页 序列表 43 页  
附图 26 页

(54) 发明名称

制备碳水化合物酯的方法

(57) 摘要

制备碳水化合物酯的方法,所述方法包括将酰基供体、酰基受体与水混合以产生包含 5-98% 水的高水分环境,其中所述酰基供体是选自磷脂、溶血磷脂、甘油三酯、甘油二酯、糖脂或溶血糖脂组成的组中的一或多种的脂质底物,且所述酰基受体是碳水化合物;以及将所述混合物与脂质酰基转移酶接触,使得所述脂质酰基转移酶催化以下反应中的一或所有两种:醇解或酯基转移作用。本发明还涉及所述脂质酰基转移酶在制备碳水化合物酯中的应用。

1. 制备碳水化合物酯的方法，所述方法包括将酰基供体、酰基受体与水混合以产生含 5-98% 水的高水分环境，其中所述酰基供体是脂质底物，其选自磷脂、溶血磷脂、甘油三酯、甘油二酯、糖脂或溶血糖脂组成的组中的一或多种，且所述酰基受体是碳水化合物；以及将所述混合物与脂质酰基转移酶接触，使得所述脂质酰基转移酶催化以下反应中的一或所有两种：醇解或酯基转移作用。

2. 权利要求 1 的方法，其中所述脂质酰基转移酶是固定化的。

3. 权利要求 1 或 2 的方法，其中所述方法包括纯化所述碳水化合物酯。

4. 权利要求 1-3 之一的方法，其中所述脂质酰基转移酶的特征是，其为具有酰基转移酶活性并包含氨基酸基序 GDSX 的酶，其中 X 是以下氨基酸残基中的一或多种：L、A、V、I、F、Y、H、Q、T、N、M 或 S。

5. 前述权利要求任意一项所述的方法，其中所述脂质酰基转移酶是在使用“经缓冲的底物中的转移酶测定法”检测时具有至少 2% 酰基转移酶活性的脂质酰基转移酶；所述转移酶测定法包括以下步骤：

i) 将 450mg 磷酸卵磷脂和 50mg 胆固醇溶于氯仿，在真空条件下蒸发干燥；

ii) 将 300mg 胆固醇 / 磷脂酰胆碱混合物移到 Wheaton 玻璃杯中，添加 15ml 50mM HEPES 缓冲液 pH 7，并在搅拌过程中使脂质分散到缓冲液中；

iii) 在使用磁力搅拌器混和的同时，将底物加热至 35°C，并添加 0.25ml 酶溶液；

iv) 在反应 0、5、10、15、25、40 和 60 分钟后取出 2ml 样品，并立即添加 25  $\mu$ l 4M HCl 以酸化游离脂肪酸，终止酶反应；

v) 添加 3ml 氯仿，并剧烈摇晃振摇 30 秒，将样品离心，分离 2ml 氯仿相，并用 0.45  $\mu$  m 的滤纸过滤，进入 10ml 涂焦油的 Dram 玻璃杯中；

vi) 在 60°C 氮蒸汽条件下蒸发氯仿，并称量样品；

vii) 通过 GLC 分析提取的脂质。

6. 前述权利要求任意一项所述的方法，其中所述脂质酰基转移酶被分类为 E.C.2.3.1.x。

7. 前述权利要求任意一项所述的方法，其中所述脂质酰基转移酶可以从一或多种下列属的生物体获得：气单胞菌、链霉菌、酵母菌、乳球菌、分枝杆菌、链球菌、乳杆菌、脱亚硫酸菌、芽胞杆菌、弯曲杆菌、弧菌科、木杆菌、硫化叶菌、曲霉、裂殖酵母、李司忒氏菌、奈瑟氏球菌、Mesorhizobium、雷尔氏菌、黄单胞菌和念珠菌。

8. 前述权利要求任意一项所述的方法，其中所述脂质酰基转移酶包含一种或多种以下氨基酸序列：(i) SEQ ID No.2 所示氨基酸序列；(ii) SEQ ID No.3 所示氨基酸序列；(iii) SEQ ID No.4 所示氨基酸序列；(iv) SEQ ID No.5 所示氨基酸序列；(v) SEQ ID No.6 所示氨基酸序列；(vi) SEQ ID No.12 所示氨基酸序列；(vii) SEQ ID No.20 所示氨基酸序列；(viii) SEQ ID No.22 所示氨基酸序列；(ix) SEQ ID No.24 所示氨基酸序列；(x) SEQ ID No.26 所示氨基酸序列；(xi) SEQ ID No.28 所示氨基酸序列；(xii) SEQ ID No.30 所示氨基酸序列；(xiii) SEQ ID No.32 所示氨基酸序列；(xiv) SEQ ID No.34 所示氨基酸序列，或者与 SEQ ID No.2、SEQ ID No.3、SEQ ID No.4、SEQ ID No.5、SEQ ID No.6、SEQ ID No.12、SEQ ID No.20、SEQ ID No.22、SEQ ID No.24、SEQ ID No.26、SEQ ID No.28、SEQ ID No.30、SEQ ID No.32 或 SEQ ID No.34 所示的任意一种序列具有 75% 或

75%以上同一性的氨基酸序列。

9. 脂质酰基转移酶在制备碳水化合物酯中的用途，所述制备通过在酰基供体、酰基受体与水的混合物中催化醇解和 / 或酯基转移作用来实施，所述混合物包含 5-98% 水、其中所述酰基供体是脂质底物，其选自磷脂、溶血磷脂、甘油三酯、甘油二酯、糖脂或溶血糖脂组成的组中的一或多种，所述酰基受体是碳水化合物。

10. 如权利要求 9 所述的用途，其中所述脂质酰基转移酶是固定化的。

11. 如权利要求 9 所述的用途，其中所述碳水化合物酯是经纯化的。

12. 如权利要求 9-11 中任一项所述的用途，其中所述脂质酰基转移酶的特征在于其为这样一种酶，它具有酰基转移酶活性，并包含氨基酸基序 GDSX，其中 X 是以下氨基酸残基中的一种或多种残基：L、A、V、I、F、Y、H、Q、T、N、M 或 S。

13. 如权利要求 9-12 中任一项所述的用途，其中所述脂质酰基转移酶是在使用“经缓冲的底物中的转移酶测定法”检测时具有至少 2% 酰基转移酶活性的脂质酰基转移酶；所述转移酶测定法包括以下步骤：

i) 将 450mg 磷酸卵磷脂和 50mg 胆固醇溶于氯仿，在真空条件下蒸发干燥；

ii) 将 300mg 胆固醇 / 磷脂酰胆碱混合物移到 Wheaton 玻璃杯中，添加 15ml 50mM HEPES 缓冲液 pH 7，并在搅拌过程中使脂质分散到缓冲液中；

iii) 在使用磁力搅拌器混和的同时，将底物加热至 35°C，并添加 0.25ml 酶溶液；

iv) 在反应 0、5、10、15、25、40 和 60 分钟后取出 2ml 样品，并立即添加 25  $\mu$ l 4M HCl 以酸化游离脂肪酸，终止酶反应；

v) 添加 3.00ml 氯仿，并剧烈摇晃振摇 30 秒，将样品离心，分离 2ml 氯仿相，并用 0.45  $\mu$  m 的滤纸过滤，进入 10ml 涂焦油的 Dram 玻璃杯中；

vi) 在 60°C 氮蒸汽条件下蒸发氯仿，并称量样品；

vii) 通过 GLC 分析提取的脂质。

14. 如权利要求 9-13 中任一项所述的用途，其中所述脂质酰基转移酶被分类为 E.C.2.3.1.x。

15. 如权利要求 9-14 中任一项所述的用途，其中所述的脂质酰基转移酶可以从一或多种下列属的生物体获得：气单胞菌、链霉菌、酵母菌、乳球菌、分枝杆菌、链球菌、乳杆菌、脱亚硫酸菌、芽胞杆菌、弯曲杆菌、弧菌科、木杆菌、硫化叶菌、曲霉、裂殖酵母、李司忒氏菌、奈瑟氏球菌、Mesorhizobium、雷尔氏菌、黄单胞菌和念珠菌。

16. 如权利要求 9-15 中任一项所述的用途，其中所述脂质酰基转移酶包含一种或多种以下氨基酸序列：(i) SEQ ID No.2 所示氨基酸序列；(ii) SEQ ID No.3 所示氨基酸序列；(iii) SEQ ID No.4 所示氨基酸序列；(iv) SEQ ID No.5 所示氨基酸序列；(v) SEQ ID No.6 所示氨基酸序列；(vi) SEQ ID No.12 所示氨基酸序列；(vii) SEQ ID No.20 所示氨基酸序列；(viii) SEQ ID No.22 所示氨基酸序列；(ix) SEQ ID No.24 所示氨基酸序列；(x) SEQ ID No.26 所示氨基酸序列；(xi) SEQ ID No.28 所示氨基酸序列；(xii) SEQ ID No.30 所示氨基酸序列；(xiii) SEQ ID No.32 所示氨基酸序列；(xiv) SEQ ID No.34 所示氨基酸序列、或者与 SEQ ID No.2、SEQ ID No.3、SEQ ID No.4、SEQ ID No.5、SEQ ID No.6、SEQ ID No.12、SEQ ID No.20、SEQ ID No.22、SEQ ID No.24、SEQ ID No.26、SEQ ID No.28、SEQ ID No.30、SEQ ID No.32 或 SEQ ID No.34 所示的任意一种序列具有 75% 或 75% 以上

同一性的氨基酸序列。

## 制备碳水化合物酯的方法

[0001] 参考的相关申请

[0002] 本申请是于 2004 年 1 月 15 日提交的发明名称为“方法”，申请号为 200480006382.3 发明专利申请的分案申请。

[0003] 本申请参考了如下相关申请：1999 年 7 月 20 日提交的、申请号为 09/750,990 的美国申请和申请号为 10/409,391 的美国申请。每一篇申请和这些申请的每一篇中引用的每篇文献（“申请引用的文献”），每一篇本发明所参考和引用的每篇文献，无论在正文中或者在那些申请的审查过程中，以及在那些审查过程中提出的所有支持专利性的文献，都包含在本文中作为参考。本文正文也引用了各种文件（“本文引用的文件”）。每一篇“本文引用的文件”和每一篇在“本文引用的文件”中引用或者参考的文献都包含在本文中作为参考。

### 技术领域

[0004] 本发明涉及通过利用脂质酰基转移酶对脂质进行生物转化以制备碳水化合物酯和 / 或蛋白质酯和 / 或蛋白质亚基酯和 / 或羧基酸酯 (hydroxy acid ester) 的方法。

[0005] 本发明还涉及脂质酰基转移酶在对脂质进行生物转化使其成为以下一或多种物质中的用途：碳水化合物酯和 / 或蛋白质和 / 或蛋白质亚基酯和 / 或羧基酸酯。

[0006] 本发明还涉及本文固定化的脂质酰基转移酶的用途，该固定化的脂质酰基转移酶可用于在高水分环境中对脂质进行生物转化以制备一或多种碳水化合物酯和 / 或蛋白质酯和 / 或蛋白质亚基酯和 / 或羧基酸酯。

[0007] 本发明还涉及固定化的脂质酰基转移酶。

### 背景技术

[0008] 脂酶已经广泛用于脂质的生物转化以制备高价值产品，例如糖酯，以便用于广泛的工业中，包括食品和 / 或饲料工业，化妆品和 / 或皮肤护理工业，油化学 (oleochemical) 工业和制药工业。

[0009] 当生物转化过程需要脂质底物水解时，脂解酶可用于高水分环境中。然而，当生物转化过程需要酯交换反应或酯基转移反应诸如通过醇解，由于不必要的水解反应利用酯酶高水分环境可以是有害的，所述不必要的水解反应可导致不需要的生物产物和 / 或使得生物转化产物的产量较低。

[0010] 通常，需要酯交换作用和 / 或酯基转移作用的生物转化过程利用非 - 水环境中的酯酶，所述非 - 水环境诸如油系统和 / 或有机溶剂系统诸如丁醇，甲醇或己烷。这样的系统提供这样的环境，其中极性受体分子和脂质供体分子可以至少部分溶解，且所述酯酶具有足够的酶活性。尽管少量水在任何酶活性所需的，水的量严格保持在较低水平以避免该酶的水解活性。

[0011] 通常，糖酯，蛋白质酯，羧基酸酯已经通过化学合成利用无机催化物产生。常规生物转化法制备糖酯或羧基酸酯利用有机溶剂环境或仅存在少量水的超临界液体中的

酯酶。

[0012] Lecointe et al *Biotechnology Letters*, Vol 18., No.8(August), pp869-874 公开了对多种脂酶的研究, 以及它们在含水介质中的活性, 所述研究根据分别来自甲醇和丁醇的甲基酯或丁基酯。Lecointe et al 教导来自近平滑念珠菌 (*Candida parapsilosis*) 的酯酶 / 酰基转移酶, 其随甲醇或丁醇浓度增加显示降低的水解活性以及增强的产生甲基酯和丁基酯的能力。来自近平滑念珠菌的酯酶 / 酰基转移酶在制备脂肪羟氨基 (hydroxamic) 酸中的用途在 Vaysse et al.*J.of Biotechnology* 53(1997)41-46 中教导。

[0013] 脂酶: 胆固醇酰基转移酶 (cholesterol acyltransferase) 已经是已知的 (参见例如 Buckley-Biochemistry 伯克利生物化学 1983, 22, 5490-5493)。具体地, 已经发现甘油磷脂: 胆固醇酰基转移酶 (GCAT), 它很象植物和 / 或哺乳动物卵磷脂: 胆固醇酰基转移酶 (LCAT), 可以催化脂肪酸在卵磷脂与胆固醇之间的转移。

[0014] 厄普顿 (Upton) 和伯克利 (Buckley) (*TIBS* 20, May 1995 p 178-179) 及布鲁米尔科 (Brumlik) 和伯克利 (Buckley) (*J.of Bacteriology* Apr.1996 p2060-2064) 教导了源自嗜水气单胞菌 (*Aeromona hydrophila*) 的脂酶 / 酰基转移酶能够在水介质中将酰基转移到乙醇受体上。

## 发明内容

[0015] 本发明第一方面提供了制备一或多种碳水化合物酯, 蛋白质酯, 蛋白质亚基酯或羧基酯的方法, 所述方法包括将酰基供体, 酰基受体与水混合以产生包含 5-98% 水的高水分环境, 其中所述酰基供体是脂质底物, 其选自由磷脂, 溶血磷脂 (lysophospholipid), 甘油三酯, 甘油二酯, 糖脂或溶血糖脂 (lysoglycolipid) 组成的组中的一或多种, 且所述酰基受体是选自由碳水化合物, 蛋白质, 蛋白质亚基, 或羧基组成的组中的一或多种物质; 以及将所述混合物与脂质酰基转移酶接触, 使得所述脂质酰基转移酶催化以下反应中的一或所有两种: 醇解或酯基转移作用。

[0016] 本发明另一方面提供脂质酰基转移酶在制备一或多种碳水化合物酯, 蛋白质酯, 蛋白质亚基酯或羧基酯中的用途, 其可通过在酰基供体, 酰基受体与水的混合物中催化醇解和 / 或酯基转移作用进行, 所述混合物包含 5-98% 水, 其中所述酰基供体是脂质底物, 其选自由磷脂, 溶血磷脂, 甘油三酯, 甘油二酯, 糖脂或溶血糖脂组成的组中的一或多种, 且所述酰基受体是选自由碳水化合物, 蛋白质, 蛋白质亚基, 或羧基组成的组中的一或多种物质。

[0017] 根据本发明另一方面, 提供了通过本发明的方法制备的碳水化合物酯, 蛋白质酯, 蛋白质亚基酯或羧基酯。

[0018] 根据本发明另一方面, 提供了通过本发明的方法制备的药物, 化妆品, 食品, 颜料 (paint), 包含碳水化合物酯, 蛋白质酯, 蛋白质亚基酯或羧基酯。

[0019] 根据本发明另一方面, 提供了本发明所述的固定化的脂质酰基转移酶。

[0020] 本发明所涉及内容的小结:

[0021] 1. 制备一或多种碳水化合物酯, 蛋白质酯, 蛋白质亚基酯或羧基酯的方法, 所述方法包括将酰基供体, 酰基受体与水混合以产生含 5-98% 水的高水分环境, 其中所述酰基供体是脂质底物, 其选自由磷脂, 溶血磷脂, 甘油三酯, 甘油二酯, 糖脂或溶血

糖脂组成的组中的一或多种，且所述酰基受体是选自由碳水化合物，蛋白质，蛋白质亚基，或羧基组成的组中的一或多种物质；以及将所述混合物与脂质酰基转移酶接触，使得所述脂质酰基转移酶催化以下反应中的一或所有两种：醇解或酯基转移作用。

[0022] 2. 如 1 的方法，其中所述脂质酰基转移酶是固定化的。

[0023] 3. 如 1 或 2 的方法，其中所述方法包括纯化所述碳水化合物酯，蛋白质酯，蛋白质亚基酯，羧基酯。

[0024] 4. 如 1-3 中任一项的方法，其中所述脂质酰基转移酶的特征是，其为具有酰基转移酶活性并包含氨基酸基序 GDSX 的酶，其中 X 是以下氨基酸残基中的一或多种：L, A, V, I, F, Y, H, Q, T, N, M 或 S。

[0025] 5. 如前述任意一项所述的方法，其中所述脂质酰基转移酶包含 H-309，或包含组氨酸残基，该组氨酸位于与 SEQ ID No.2 或 SEQ ID No.32 所示嗜水气单胞菌脂解酶的氨基酸序列中的 His-309 相对应的位置。

[0026] 6. 如前述所述的方法，其中所述的脂质酰基转移酶可以从一或多种下列属的生物体获得：气单胞菌、链霉菌、酵母菌、乳球菌、分枝杆菌、链球菌、乳杆菌、脱亚硫酸菌、芽胞杆菌、弯曲杆菌、弧菌科、木杆菌、硫化叶菌、曲霉、裂殖酵母、李司忒氏菌、奈瑟氏球菌、Mesorhizobium、雷尔氏菌、黄单胞菌和念珠菌。

[0027] 7. 如前述任意一项所述的方法，其中所述脂质酰基转移酶包含一种或多种以下氨基酸序列：(i) SEQ ID No.2 所示氨基酸序列；(ii) SEQ ID No.3 所示氨基酸序列；(iii) SEQ ID No.4 所示氨基酸序列；(iv) SEQ ID No.5 所示氨基酸序列；(v) SEQ ID No.6 所示氨基酸序列；(vi) SEQ ID No.12 所示氨基酸序列；(vii) SEQ ID No.20 所示氨基酸序列；(viii) SEQ ID No.22 所示氨基酸序列；(ix) SEQ ID No.24 所示氨基酸序列；(x) SEQ ID No.26 所示氨基酸序列；(xi) SEQ ID No.28 所示氨基酸序列；(xii) SEQ ID No.30 所示氨基酸序列；(xiii) SEQ ID No.32 所示氨基酸序列；(xiv) SEQ ID No.34 所示氨基酸序列，或者与 SEQ ID No.2, SEQ ID No.3, SEQ ID No.4, SEQ ID No.5, SEQ ID No.6, SEQ ID No.12, SEQ ID No.20, SEQ ID No.22, SEQ ID No.24, SEQ ID No.26, SEQ ID No.28, SEQ ID No.30, SEQ ID No.32, 或 SEQ ID No.34 所示的任意一种序列有 75% 或者 75% 以上同一性的氨基酸序列。

[0028] 8. 脂质酰基转移酶在制备一或多种碳水化合物酯，蛋白质酯，蛋白质亚基酯或羧基酯中的用途，所述制备通过对酰基供体，酰基受体与水的混合物的醇解和 / 或酯基转移作用来实施，所述混合物包含 5-98% 水，其中所述酰基供体是脂质底物，其选自由磷脂，溶血磷脂，甘油三酯，甘油二酯，糖脂或溶血糖脂组成的组中的一或多种，所述酰基受体是选自由碳水化合物，蛋白质，蛋白质亚基，或羧基组成的组中的一或多种物质。

[0029] 9. 如 8 所述的用途，其中所述脂质酰基转移酶是固定化的。

[0030] 10. 如 8 所述的用途，其中所述碳水化合物酯，蛋白质酯，蛋白质亚基酯或羧基酯是经纯化的。

[0031] 11. 如 8-10 中任一项所述的用途，其中所述脂质酰基转移酶的特征在于其为这样一种酶，它具有酰基转移酶活性，并包含氨基酸基序 GDSX，其中 X 是以下氨基酸残基中的一种或多种残基：L, A, V, I, F, Y, H, Q, T, N, M 或 S。

[0032] 12. 如 8-10 中任一项所述的用途，其中所述脂质酰基转移酶包含 H-309，或包含组氨酸残基，该组氨酸位于与 SEQ ID No.2 或 SEQ ID No.32 所示嗜水气单胞菌脂解酶的氨基酸序列中的 His-309 相对应的位置。

[0033] 13. 如 8-12 中任一项所述的用途，其中所述的脂质酰基转移酶可以从一或多种下列属的生物体获得：气单胞菌、链霉菌、酵母菌、乳球菌、分枝杆菌、链球菌、乳杆菌、脱亚硫酸菌、芽胞杆菌、弯曲杆菌、弧菌科、木杆菌、硫化叶菌、曲霉、裂殖酵母、李司忒氏菌、奈瑟氏球菌、Mesorhizobium、雷尔氏菌、黄单胞菌和念珠菌。

[0034] 14. 如 8-13 中任一项所述的用途，其中所述脂质酰基转移酶包含一种或多种以下氨基酸序列：(i) SEQ ID No.2 所示氨基酸序列；(ii) SEQ ID No.3 所示氨基酸序列；(iii) SEQ ID No.4 所示氨基酸序列；(iv) SEQ ID No.5 所示氨基酸序列；(v) SEQ ID No.6 所示氨基酸序列；(vi) SEQ ID No.12 所示氨基酸序列；(vii) SEQ ID No.20 所示氨基酸序列；(viii) SEQ ID No.22 所示氨基酸序列；(ix) SEQ ID No.24 所示氨基酸序列；(x) SEQ ID No.26 所示氨基酸序列；(xi) SEQ ID No.28 所示氨基酸序列；(xii) SEQ ID No.30 所示氨基酸序列；(xiii) SEQ ID No.32 所示氨基酸序列；(xiv) SEQ ID No.34 所示氨基酸序列，或者与 SEQ ID No.2, SEQ ID No.3, SEQ ID No.4, SEQ ID No.5, SEQ ID No.6, SEQ ID No.12, SEQ ID No.20, SEQ ID No.22, SEQ ID No.24, SEQ ID No.26, SEQ ID No.28, SEQ ID No.30, SEQ ID No.32, 或 SEQ ID No.34 所示的任意一种序列有 75% 或者 75% 以上同一性的氨基酸序列。

[0035] 15. 碳水化合物酯，其通过前述 1-7 任一项所述的方法产生。

[0036] 16. 蛋白质酯，其通过前述 1-7 任一项所述的方法产生。

[0037] 17. 蛋白质亚基酯，其通过前述 1-7 任一项所述的方法产生。

[0038] 18. 羧基酯，其通过前述 1-7 任一项所述的方法产生。

[0039] 19. 固定化的脂质酰基转移酶。

[0040] 20. 如 19 所述的固定化的脂质酰基转移酶，其中所述脂质酰基转移酶的特征在于其为这样一种酶，它具有酰基转移酶活性，并包含氨基酸基序 GDSX，其中 X 是以下氨基酸残基中的一种或多种残基：L, A, V, I, F, Y, H, Q, T, N, M 或 S。

[0041] 21. 如 19 或 20 的固定化的脂质酰基转移酶，其中所述脂质酰基转移酶包含 H-309，或包含组氨酸残基，该组氨酸位于与 SEQ ID No.2 或 SEQ ID No.32 所示嗜水气单胞菌脂解酶的氨基酸序列中的 His-309 相对应的位置。

[0042] 22. 如 19-21 中任一项所述的固定化的脂质酰基转移酶，其中所述的脂质酰基转移酶可以从一或多种下列属的生物体获得：气单胞菌、链霉菌、酵母菌、乳球菌、分枝杆菌、链球菌、乳杆菌、脱亚硫酸菌、芽胞杆菌、弯曲杆菌、弧菌科、木杆菌、硫化叶菌、曲霉、裂殖酵母、李司忒氏菌、奈瑟氏球菌、Mesorhizobium、雷尔氏菌、黄单胞菌和念珠菌。

[0043] 23. 如 19-22 中任一项所述的固定化的脂质酰基转移酶，其中所述脂质酰基转移酶包含一种或多种以下氨基酸序列：(i) SEQ ID No.2 所示氨基酸序列；(ii) SEQ ID No.3 所示氨基酸序列；(iii) SEQ ID No.4 所示氨基酸序列；(iv) SEQ ID No.5 所示氨基酸序列；(v) SEQ ID No.6 所示氨基酸序列；(vi) SEQ ID No.12 所示氨基酸序列；(vii) SEQ ID No.20 所示氨基酸序列；(viii) SEQ ID No.22 所示氨基酸序列；(ix) SEQ ID No.24 所示氨

氨基酸序列；(x) SEQ ID No.26 所示氨基酸序列；(xi) SEQ ID No.28 所示氨基酸序列；(xii) SEQ ID No.30 所示氨基酸序列；(xiii) SEQ ID No.32 所示氨基酸序列；(xiv) SEQ ID No.34 所示氨基酸序列，或者与 SEQ ID No.2, SEQ ID No.3, SEQ ID No.4, SEQ ID No.5, SEQ ID No.6, SEQ ID No.12, SEQ ID No.20, SEQ ID No.22, SEQ ID No.24, SEQ ID No.26, SEQ ID No.28, SEQ ID No.30, SEQ ID No.32, 或 SEQ ID No.34 所示的任意一种序列有 75% 或者 75% 以上同一性的氨基酸序列。

[0044] 发明详述

[0045] 本文应用的术语“脂质酰基转移酶”是指一种也具有脂酶活性的酶（根据国际生物化学分子生物学联合会命名委员会的酶命名法推荐 (Enzyme Nomenclature Recommendations) (1992) 通常分类为 E.C.3.1.1.x），该酶也具有酰基转移酶活性（通常分类为 E.C.2.3.1.x），该酶能够将酰基从一种脂质转移到一种或多种受体底物，例如以下物质的一种或多种：碳水化合物、蛋白质或其亚单位或羟基酸。

[0046] 优选，本发明的“酰基受体”不是水。

[0047] 一方面，优选所述酶能够将酰基从脂质底物转移到碳水化合物。

[0048] 碳水化合物酰基受体可以为以下物质中的一种或多种：单糖、二糖、寡糖或多糖。优选的碳水化合物是以下物质中的一种或多种：葡萄糖、果糖、脱水果糖 (anhydrofructose)、麦芽糖、乳糖、蔗糖、半乳糖、木糖 (xylose)、木寡糖 (xylooligosaccharide)、阿拉伯糖、麦芽寡糖 (maltooligosaccharide)、塔格糖 (tagatose)、microthecin、ascopyrone P、ascopyrone T 或 cortalcerone。

[0049] 碳水化合物酯可作为有价值的乳化剂在例如食品中起作用。

[0050] 一方面，优选所述酶能够将酰基从脂质底物转移到蛋白质和 / 或蛋白质亚基。

[0051] 优选所述蛋白质亚基是以下物质中的一种或多种：氨基酸，蛋白质水解产物，肽，二肽，寡肽，多肽。

[0052] 适宜的蛋白质可以是食品中例如在乳制品和 / 或肉制品中的一种或多种蛋白质。仅作为实例，适宜的蛋白质可能是乳凝块或乳清中的蛋白质，如乳球蛋白。其它适宜的蛋白质包括卵清蛋白（来自蛋类）、麸蛋白 (gliadin)、麦谷蛋白 (glutenin)、puroindoline、来自谷物的脂质转移蛋白和肉中的肌球蛋白，或以下乳蛋白：酪蛋白，乳白蛋白 (lactalbumin) 和乳铁传递蛋白 (lactoferrin)。

[0053] 适宜地，在所述蛋白质或蛋白质亚基中，所述酰基受体可以是以下蛋白质或蛋白质亚基组分中的一种或多种：丝氨酸，苏氨酸，酪氨酸，或半胱氨酸。

[0054] 当所述蛋白质亚基是氨基酸时，适宜地所述氨基酸可以是任何氨基酸，优选所述氨基酸是例如丝氨酸，苏氨酸，酪氨酸或半胱氨酸中的一或多种。

[0055] 一方面，优选所述酶能将酰基从脂质底物转移到羟基酸。

[0056] 适宜地，所述羟基酸是以下酸中的一种或多种：柠檬酸，酒石酸，乳酸，抗坏血酸，羟基乙酸，苹果酸， $\alpha$ -羟基醋酸 ( $\alpha$ -hydroxyethanoic acid)， $\alpha$ -羟基辛酸 ( $\alpha$ -hydroxyoctanoic acid)， $\alpha$ -羟基辛酸 ( $\alpha$ -hydroxycaprylic acid)，羟基辛酸 (hydroxycaprylic acid)，葡糖酸，乳糖酸 (lactobionic acid) 或麦芽糖酸 (maltobionic acid)。

[0057] 适宜地，所述羟基酸可以是水果酸 (fruit acid)，例如苹果酸，乳酸，酒石酸，柠

檬酸或羟基乙酸中的一种或多种。

[0058] 一个实施方案中，优选所述羟基酸是柠檬酸，乳酸，酒石酸或苹果酸中的一种或多种。

[0059] 本文术语“羟基酸”指其中烷基的一或多个氢原子被羟基取代的羧酸。

[0060] 一方面，所述脂质酰基转移酶可将酰基从脂质底物转移到碳水化合物，蛋白质亚基或羟基酸中的一种或多种，所述脂质酰基转移酶还可将酰基从脂质转移到固醇和/或 stanol，具体是 phytosterol 和/或 hytostanol 中的一种或多种。

[0061] 适宜地，当所述脂质底物是磷脂，其可为卵磷脂，例如磷脂酰胆碱。本文术语卵磷脂包括磷脂酰胆碱，磷脂酰乙醇胺，磷脂酰肌醇，磷脂酰丝氨酸和磷脂酰甘油。

[0062] 适宜地，当脂质底物是溶血磷脂，其可为溶血卵磷脂，例如溶血磷脂酰胆碱。本文术语溶血磷脂酰胆碱与术语溶血卵磷脂同义，这些术语可互换使用。

[0063] 适宜地，当脂质底物是糖脂，其可为例如双半乳糖甘油二酯 (DGDG)。

[0064] 所述脂质底物可称为“脂质酰基供体”或“酰基供体”。这些术语可互换使用。

[0065] 一些方面，优选脂质酰基转移酶针对其起作用的脂质底物是磷脂，诸如卵磷脂，例如磷脂酰胆碱。

[0066] 一些方面，优选所述脂质底物是糖脂，诸如 DGDG。

[0067] 一些方面，所述脂质底物可以是食品脂质，即食品的脂质组分。

[0068] 在一些方面，优选地，本发明的脂质酰基转移酶不能或基本不能作用于甘油三酯和/或 1-甘油单酯和/或 2-甘油单酯。

[0069] 脂质底物或脂质酰基供体宜为一种或多种脂质，其存在于一种或多种下列底物中：脂肪，包括猪油 (lard)、牛羊油 (tallow)、乳脂 (butter fat)；油，包括从棕榈油 (palm oil)、葵花子油 (sunflower oil)、大豆油 (soya bean oil)、红花油 (safflower oil)、棉籽油 (cotton seed oil)、花生油 (ground nut oil)、玉米油 (corn oil)、橄榄油 (olive oil)、花生油 (peanut oil)、椰子油 (coconut oil) 和油菜籽油 (rape seed oil) 中提取或由它们衍生的油。来自大豆、油菜籽、或蛋黄的卵磷脂也是适宜的脂质底物。脂质底物也可以是燕麦脂质或其它包含半乳糖脂的基于植物的物质。

[0070] 在本发明的一些方面，脂质可选自脂肪酸链长为 8-22 碳的脂质。

[0071] 本发明的一些方面，脂质可选自脂肪酸链长为 16-22 碳的脂质，更优选为 16-20 碳的脂质。

[0072] 本发明的一些方面，脂质可选自脂肪酸链长不多于 14 碳的脂质，适宜地选自脂肪酸链长 4-14 碳的脂质，适宜地为 4-10 碳的脂质，适宜地为 4-8 碳的脂质。

[0073] 优选所述酰基供体不是游离脂肪酸。

[0074] 优选，所述酰基供体不是碳水化合物（糖）酯。

[0075] 适宜地，本发明中的脂质酰基转移酶可显示出一种或多种下列脂酶的活性：糖脂酶活性 (E.C.3.1.1.26)，甘油三酯酶活性 (E.C.3.1.1.3)，磷脂酶 A2 活性 (E.C.3.1.1.4)，磷脂酶 A1 活性 (E.C.3.1.1.32)。这里应用的术语“糖脂酶活性”包括“半乳糖脂酶活性”。

[0076] 适宜地，本发明中的脂质酰基转移酶具有至少一种或多种下列活性：糖脂酶活

性 (E.C.3.1.1.26) 和 / 或磷脂酶 A1 活性 (E.C.3.1.1.32) 和 / 或磷脂酶 A2 活性 (E.C.3.1.1.4)。

[0077] 对于本发明一些方面, 脂质酰基转移酶至少具有糖脂酶活性 (E.C.3.1.1.26)。

[0078] 适宜地, 对于一些方面, 本发明的脂质酰基转移酶可将酰基从糖脂和 / 或磷脂上转移到以下一种或多种受体底物: 碳水化合物, 蛋白质, 甘油羧基酸。

[0079] 一些方面, 优选本发明的脂质酰基转移酶可将酰基从糖脂和 / 或磷脂转移到碳水化合物, 形成至少碳水化合物酯。

[0080] 一些方面, 优选本发明的脂质酰基转移酶可将酰基从糖脂或磷脂转移到蛋白质或蛋白质亚基, 形成至少蛋白质酯 (或蛋白质脂肪酸缩合物) 或蛋白质亚基酯。

[0081] 本文术语 “蛋白质亚基酯” 酯从任何蛋白质亚基形成的酯, 诸如二肽酯, 寡肽酯, 多肽酯或蛋白质水解产物酯。

[0082] 一些方面, 优选本发明脂质酰基转移酶不表现甘油三酯酶活性 (E.C.3.1.1.3)。

[0083] 优选地, 本发明所述的脂质酰基转移酶可以用下列标准来表征:

[0084] (i) 该酶具有酰基转移活性 (可定义为酯转移活性), 由此将脂质酰基供体的原始酯键上的酰基部分转移到一或多种碳水化合物, 蛋白质, 蛋白质亚基或羧基酸酰基受体形成新酯, 即碳水化合物酯, 蛋白质酯, 蛋白质亚基酯或羧基酸酯; 和

[0085] (ii) 该酶包含氨基酸序列基序 GDSX, 其中 X 是下列氨基酸残基 L, A, V, I, F, Y, H, Q, T, N, M 或 S 中的一种或多种。

[0086] 优选地, GDSX 基序中的 X 是 L, 这样, 本发明所述的酶优选包含氨基酸基序 GSDL。

[0087] GDSX 基序是由四种保守氨基酸构成。 优选地, 所述基序中的丝氨酸是脂质酰基转移酶的催化性丝氨酸。 适宜地, GDSX 基序中的丝氨酸可以在与嗜水气单胞菌脂解酶中与 Ser-16 相应的位置, 如 Brumlik 和 Buckley (Journal of Bacteriology Apr.1996, Vol.1, No.7, p 2060-2064) 的教导。

[0088] 为确定一种蛋白质是否含有本发明所述 GDSX 基序, 优选将该序列与 Pfam 数据库中的隐藏 markov 模型图谱 (model profile) (HMM 图谱) 进行比较。

[0089] Pfam 是蛋白质结构域家族的数据库。 Pfam 包括用于每个家族的辅助 (curated) 多重序列比对以及用于鉴定新序列中的这些结构域的隐藏 Markov 模型图谱 (HMM 图谱)。 对 Pfam 的介绍见于 Bateman A 等. (2002) Nucleic Acids Res.30 ; 276-280。 隐藏 Markov 模型被应用于许多旨在对蛋白质进行分类的数据库中, 综述见 Bateman A 和 Haft DH (2002) Brief Bioinform3 ; 236-245。

[0090] [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list\\_uids=12230032&dopt=Abstract](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=12230032&dopt=Abstract)

[0091] [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list\\_uids=11752314&dopt=Abstract](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=11752314&dopt=Abstract)

[0092] 对隐藏 Markov 模型的详细解释和这些模型如何在 Pfam 数据库中应用参见 Durb 中 R, Eddy S, 和 Krogh A (1998) Biological sequence analysis ; probabilistic models of proteins and nucleic acids. Cambridge University Press, ISBN 0-521-62041-4 (Durb 中 R, Eddy S, 和 Krogh A (1998) 生物序列分析, 蛋白质和核苷酸概率模型, 剑桥大学出版社, ISBN 0-521-62041-4。 Hammer 程序包可以从华盛顿大学, St Louis, USA 获得。

[0093] 可选择地，GDSX 基序可以通过应用 Hammer 软件包识别，说明由 Durb 中 R, Eddy S, 和 Krogh A(1998) Biological sequence 分析; probabilistic models of proteins and nucleic acids. Cambridge University Press, ISBN0-521-62041-4 (Durb 中 R, Eddy S, 和 Krogh A(1998) 生物序列分析, 蛋白质和核苷酸概率模型, 剑桥大学出版社, ISBN 0-521-62041-4) 及其中的参考文献提供, 以及本说明书中关于 HMMER2 的资料。

[0094] PFAM 数据库可以通过, 例如, 目前位于如下网址的几个服务器进行访问:

[0095] <http://www.sanger.ac.uk/Software/Pfam/index.shtml>

[0096] <http://pfam.wustl.edu/>

[0097] <http://pfam.jouy.inra.fr/>

[0098] <http://pfam.cgb.ki.se/>

[0099] 数据库提供了检索工具, 在此可以输入蛋白质序列。应用数据库的默认参数, 对蛋白质序列中 Pfam 结构域的存在进行分析。GDSX 结构域是在数据库中已建立的区域, 因此任何查询序列中存在的该结构域可以得到识别。数据库将返回 Pfam00657 共有序列与查询序列的比对。

[0100] 多重比对, 包括杀鲑气单胞菌或嗜水气单胞菌可以通过下列方法获得:

[0101] a) 手工方法

[0102] 通过上述程序, 可获得目的蛋白与 Pfam00657 保守序列的比对并获得 P10480 序列与 Pfam00657 保守序列的比对; 或

[0103] b) 通过数据库

[0104] 在鉴定 Pfam00657 共有序列之后, 数据库提供选项显示查询序列与 Pfam00657 保守序列的种子比对 (seed alignment), P10480 是该种子比对的一部分, 且表示为 GCAT\_AERHY。查询序列和 P10480 都将在同一窗口显示。

[0105] 嗜水气单胞菌参比序列:

[0106] 嗜水气单胞菌 GDSX 脂酶的残基在 NCBI 文件 P10480 中进行编号, 所述文本中的编号是指由该文件给出的编号, 在本发明中它用来确定具体的氨基酸残基, 所述具体的氨基酸残基在优选具体实施方案中, 存在于本发明的脂质酰基转移酶中。

[0107] 实施 Pfam 比对 (图 33 和 34):

[0108] 下列保守的残基可被识别, 并且在优选的实施方案中存在于本发明的组合物和方法所应用的酶中。

[0109] 1 区 -GDSX 区

[0110] hid hid hid hid Gly Asp Ser hid

[0111] 28 29 30 31 32 33 34 35

[0112] 2 区 -GANDY 区

[0113] hid Gly hid Asn Asp hid

[0114] 130 131 132 133 134 135

[0115] 3 区 -HPT 区

[0116] His

[0117] 309

[0118] 其中 'hid' 是指选自 Met, Ile, Leu, Val, Ala, Gly, Cys, His, Lys, Trp,

Tyr 或 Phe 的一个疏水残基。

[0119] 优选地，应用于本发明的组合物/方法中的脂质酰基转移酶可用 Pfam00657 共有序列进行比对。

[0120] 优选地，与 pfam00657 结构域家族的隐藏 markov 模型图谱 (HMM 图谱) 的阳性匹配表明本发明中的 GDSL 或 GDSX 结构域的存在。

[0121] 优选地，当与 Pfam00657 共有序列比对时，应用于本发明的组合物/方法中的脂质酰基转移酶含有至少一个，优选一个以上，优选两个以上下列区域，GDSx 区、GANDY 区、HPT 区。适宜地，脂质酰基转移酶含有 GDSx 区和 GANDY 区。可选择地，所述酶含有 GDSx 区和 HPT 区。优选地，所述酶含有至少一个 GDSx 区。

[0122] 优选地，当与 Pfam00657 共有序列比对时，应用于本发明的组合物/方法中的酶与参比嗜水气单胞菌多肽序列即 SEQ ID No.32 比较时含有至少一个，优选一个以上，优选两个以上，优选三个以上，优选四个以上，优选五个以上，优选六个以上，优选七个以上，优选八个以上，优选九个以上，优选十个以上，优选十一个以上，优选十二以上，优选十三以上，优选十四个以上下列氨基酸残基：28hid, 29hid, 30hid, 31hid, 32gly, 33Asp, 34Ser, 35hid, 130hid, 131Gly, 132hid, 133Asn, 134Asp, 135hid, 309His。

[0123] pfam00657 GDSX 结构域是将具有该区域蛋白与其它酶区分的独特标识。

[0124] Pfam00657 共有序列，表示为图 1 的 SEQ ID No.1。这是从第 6 版数据库 Pfam00657 家族的鉴定衍生出来的，本文也称其为 pfam00657.6。

[0125] 保守序列可通过新版 Pfam 数据库来更新。

[0126] 例如，图 33 和 34 显示第 11 版数据库中 00657 家族的 pfam 比对，本文也称其为 pfam00657.11。

[0127] 发现两个版本的数据库中 Pfam00657 家族中都存在 GDSx 区、GANDY 区和 HPT 区。其它版本的 (release)pfam 数据库可用来鉴定 pfam00657 家族。

[0128] 优选地，本发明所述的脂质酰基转移酶可以用下列标准来表征：

[0129] (i) 该酶具有转移酰基的活性 (可定义为酯基转移活性)，由此将脂质酰基供体的原始酯键上的酰基部分转移到酰基受体上形成新酯；

[0130] (ii) 该酶包含氨基酸序列基序 GDSX，其中 X 是一种或多种下列氨基酸残基：L, A, V, I, F, Y, H, Q, T, N, M 或 S；

[0131] (iii) 该酶包含 His-309 或包含对应于嗜水气单胞菌脂解酶的 His-309 位的组氨酸残基，所述嗜水气单胞菌脂解酶如图 2 所示 (SEQ ID No.2 或 SEQ ID No.32)。优选地，GDSX 基序的氨基酸残基是 L。

[0132] 在 SEQ ID No.2 或 SEQ ID No.32 中，前 18 个氨基酸残基形成信号序列。全长序列 (包含信号序列的蛋白质) 的 His-309，等同于该蛋白质成熟部分 (即不含有信号序列的序列) 中的 His-291。

[0133] 优选地，本发明所述的脂质酰基转移酶包括下列催化三联体：Ser-34, Asp-134 和 His-309 或者包含位置分别对应于图 2 (SEQ ID No.2) 或图 28 (SEQ ID No.32) 所示的嗜水气单胞菌脂解酶中 Ser-34, Asp-134 和 His-309 的丝氨酸残基、天冬氨酸残基和组氨酸残基。如上所述，在 SEQ ID No.2 或 SEQ ID No.32 所示序列中，前 18 个氨基酸残基形

成信号序列。全长序列（包含信号序列的蛋白质）的 Ser-34, Asp-134 和 His-309, 等同于蛋白质成熟部分（即不含有信号序列的序列）的 Ser-16, Asp-116 和 His-291。在图 1 (SEQ ID No.1) 所示的 Pfam00657 共有序列中, 活性位点残基对应于 Ser-7, Asp-157 和 His-348。

[0134] 优选地, 本发明所述的脂质酰基转移酶可以用下列标准来表征:

[0135] (i) 该酶具有酰基转移活性（可定义为酯转移活性）, 由此将脂质酰基供体的原始酯键上的酰基部分转移到一或多种碳水化合物, 蛋白质, 蛋白质亚基或羧基酰基受体形成新酯, 即碳水化合物酯, 蛋白质酯, 蛋白质亚基酯或羧基酯; 并且

[0136] (ii) 该酶包含氨基酸序列基序 GDSX, 其中 X 是一种或多种下列氨基酸残基: L, A, V, I, F, Y, H, Q, T, N, M 或 S;

[0137] (iii) 该酶包含 His-309 或包含对应于嗜水气单胞菌脂解酶的 His-309 位的组氨酸残基, 所述嗜水气单胞菌脂解酶如图 2 所示 (SEQ ID No.2 或 SEQ ID No.32)。

[0138] 优选地, GDSX 基序的氨基酸残基是 L。

[0139] 在 SEQ ID No.2 或 SEQ ID No.32 中, 前 18 个氨基酸残基形成信号序列。全长序列（包含信号序列的蛋白质）的 His-309, 等同于该蛋白质成熟部分（即不含有信号序列的序列）中的 His-291。

[0140] 优选地, 本发明所述的脂质酰基转移酶包括下列催化三联体: Ser-34, Asp-134 和 His-309 或者包含位置分别对应于图 2 (SEQ ID No.2) 或图 28 (SEQ ID No.32) 所示的嗜水气单胞菌脂解酶中 Ser-34, Asp-134 和 His-309 的丝氨酸残基、天冬氨酸残基和组氨酸残基。如上所述, 在 SEQ ID No.2 或 SEQ ID No.32 所示序列中, 前 18 个氨基酸残基形成信号序列。全长序列（包含信号序列的蛋白质）的 Ser-34, Asp-134 和 His-309, 等同于蛋白质成熟部分（即不含有信号序列的序列）的 Ser-16, Asp-116 和 His-291。在图 1 (SEQ ID No.1) 所示的 Pfam00657 共有序列中, 活性位点残基对应于 Ser-7, Asp-157 和 His-348。

[0141] 优选地, 本发明所述的脂质酰基转移酶可以用下列标准来表征:

[0142] (i) 该酶具有酰基转移活性（可定义为酯转移活性）, 由此将脂质酰基供体的原始酯键上的酰基部分转移到一或多种碳水化合物, 蛋白质, 蛋白质亚基或羧基酰基受体形成新酯, 即碳水化合物酯, 蛋白质酯, 蛋白质亚基酯或羧基酯; 并且

[0143] (ii) 该酶包含至少 Gly-32, Asp-33, Ser-34, Asp-134 和 His-309 或者包含位置分别对应于图 2 (SEQ ID No.2) 或图 28 (SEQ ID No.32) 所示的嗜水气单胞菌脂解酶中 Gly-32, Asp-33, Ser-34, Asp-134 和 His-309 的甘氨酸、天冬氨酸、丝氨酸、天冬氨酸和组氨酸。

[0144] 适宜地, 本发明所述的脂质酰基转移酶可以并优选从下列来自一个或多个属的生物体中获得: 气单胞菌属 (Aeromonas)、链霉菌属 (Streptomyces)、酵母菌属 (Saccharomyces)、乳球菌属 (Lactococcus)、分枝杆菌属 (Mycobacterium)、链球菌属 (Streptococcus)、乳杆菌属 (Lactobacillus)、脱亚硫酸菌属 (Desulfitobacterium)、芽胞杆菌属 (Bacillus)、弯曲杆菌属 (Campylobacter)、弧菌属 (Vibrionaceae)、木杆菌属 (Xylella)、硫叶菌属 (Sulfolobus)、曲霉属 (Aspergillus)、裂殖酵母属 (Schizosaccharomyces)、李斯特氏菌属 (Listeria)、奈瑟氏球菌 (Neisseria)、

Mesorhizobium、雷尔氏菌属 (*Ralstonia*)、黄杆菌属 (*Xanthomonas*) 和假丝酵母属 (*Candida*)。

[0145] 适宜地，本发明所述的脂质酰基转移酶可以并优选从下列一种或多种生物体中获得：嗜水气单胞菌 (*Aeromonas hydrophila*)、杀鲑气单胞菌 (杀鲑气单胞菌)、天蓝色链霉菌 (*Streptomyces coelicolor*)、龟裂链霉菌 (*Streptomyces rimosus*)、分枝杆菌 (*Mycobacterium*)、化脓性链球菌 (*Streptococcus pyogenes*)、乳酸乳球菌 (*Lactococcus lactis*)、嗜热链球菌 (*Streptococcus thermophilus*)、瑞士乳杆菌 (*Lactobacillus helveticus*)、脱卤脱亚硫酸菌 (*Desulfitobacterium dehalogenans*)、芽孢杆菌 (*Bacillus* sp)、空肠弯曲杆菌 (*Campylobacter jejuni*)、弧菌 (*Vibrionaceae*)、苛养木杆菌 (*Xylella fastidiosa*)、硫磺矿硫叶菌 (*Sulfolobus solfataricus*)、酿酒酵母 (*Saccharomyces cerevisiae*)、土曲霉 (*Aspergillus terreus*)、粟酒裂殖酵母 (*Schizosaccharomyces pombe*)、无害李司忒氏菌 (*Listeria innocua*)、单核细胞增多性李司忒氏菌 (*Listeria monocytogenes*)、脑膜炎奈瑟氏球菌 (*Neisseria meningitidis*)、*Mesorhizobium loti*、青枯雷尔氏菌 (*Ralstonia solanacearum*)、野油菜黄单胞菌 (*Xanthomonas campestris*)、地毯草黄单胞菌 (*Xanthomonas axonopodis*)、近平滑假丝酵母 (近平滑念珠菌)。

[0146] 在一方面，优选地，本发明中的脂质酰基转移酶可以并优选从一种或多种嗜水气单胞菌或杀鲑气单胞菌中获得。

[0147] 适宜地，本发明所述的脂质酰基转移酶包含一种或多种下列氨基酸序列。

[0148] (i) SEQ ID No.2 所示的氨基酸序列 (见图 2)

[0149] (ii) SEQ ID No.3 所示的氨基酸序列 (见图 3)

[0150] (iii) SEQ ID No.4 所示的氨基酸序列 (见图 4)

[0151] (iv) SEQ ID No.5 所示的氨基酸序列 (见图 5)

[0152] (v) SEQ ID No.6 所示的氨基酸序列 (见图 6)

[0153] (vi) SEQ ID No.12 所示的氨基酸序列 (见图 14)

[0154] (vii) SEQ ID No.20 所示的氨基酸序列 (见图 16)

[0155] (viii) SEQ ID No.22 所示的氨基酸序列 (见图 18)

[0156] (ix) SEQ ID No.24 所示的氨基酸序列 (见图 20)

[0157] (xi) SEQ ID No.26 所示的氨基酸序列 (见图 22)

[0158] (xii) SEQ ID No.28 所示的氨基酸序列 (见图 24)

[0159] (xiii) SEQ ID No.30 所示的氨基酸序列 (见图 26)

[0160] (ixi) SEQ ID No.32 所示的氨基酸序列 (见图 28)

[0161] (xiv) SEQ ID No.34 所示的氨基酸序列 (见图 30)

[0162] 或者与 SEQ ID No.2, SEQ ID No.3, SEQ ID No.4, SEQ ID No.5, SEQ ID No.6, SEQ ID No.12, SEQ ID No.20, SEQ ID No.22, SEQ ID No.24, SEQ ID No.26, SEQ ID No.28, SEQ ID No.30, SEQ ID No.32, 或 SEQ ID No.34 所示的任意一种序列有 75% 或者更高同一性的氨基酸序列。

[0163] 适宜地，本发明所述的脂质酰基转移酶或者包含 SEQ ID No.2 或 SEQ ID No.3 或 SEQ ID No.32 或 SEQ ID No.34 所示的氨基酸序列，或者包含与 SEQ ID No.2 或 SEQ ID No.3 或 SEQ ID No.32 或 SEQ ID No.34 所示的氨基酸序列有 75% 或 75% 以上，优选 80%

或 80%以上, 优选 85%或 85%以上, 优选 90%或 90%以上, 优选 95%或 95%以上同一性的氨基酸序列。

[0164] 为了达到本发明的目的, 同一性的程度取决于相同序列元素的数目。依照本发明, 同一性程度可通过本领域已知电脑程序适宜地测定, 所述程序诸如 GCG 程序包提供的 GAP(Program Manual for the Wisconsin Package, Version 8, August 1994, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, US53711) (Needleman & Wunsch(1970), J.of Molecular Biology48, 443-45), 利用下列设置进行多肽序列比较: GAP 生成罚分 3.0 和 GAP 延伸罚分 0.1。

[0165] 适宜地, 本发明所述的脂质酰基转移酶包含一段氨基酸序列, 该序列与 SEQ ID No.2, SEQ ID No.3, SEQ ID No.4, SEQ ID No.5, SEQ ID No.6, SEQ ID No.12, SEQ ID No.20, SEQ ID No.22, SEQ ID No.24, SEQ ID No.26, SEQ ID No.28, SEQ ID No.30, SEQ ID No.32 或 SEQ ID No.34 所示的任何一段序列具有 80%或 80%以上, 优选 85%或 85%以上, 更优选 90%或 90%以上, 还更优选 95%或 95%以上的同一性。

[0166] 适宜地, 本发明所述的脂质酰基转移酶包含下列氨基酸序列中的一或多种:

[0167] (a) 如 SEQ ID No.2 或 SEQ ID No.32 的 1-100 氨基酸残基所示的氨基酸序列;

[0168] (b) 如 SEQ ID No.2 或 SEQ ID No.32 的 101-200 氨基酸残基所示的氨基酸序列;

[0169] (c) 如 SEQ ID No.2 或 SEQ ID No.32 的 201-300 氨基酸残基所示的氨基酸序列;

或

[0170] (d) 与上述 (a)-(c) 定义的任意氨基酸序列有 75%或 75%以上, 优选 85%或 85%以上, 更优选 90%或 90%以上, 还更优选 95%或 95%以上同一性的氨基酸序列。

[0171] 适宜地, 本发明所述的脂质酰基转移酶包含下列氨基酸序列中的一或多种:

[0172] (a) 如 SEQ ID No.2 或 SEQ ID No.32 的氨基酸残基 28-39 所示的氨基酸序列;

[0173] (b) 如 SEQ ID No.2 或 SEQ ID No.32 的氨基酸残基 77-88 所示的氨基酸序列;

[0174] (c) 如 SEQ ID No.2 或 SEQ ID No.32 的氨基酸残基 126-136 所示的氨基酸序列;

[0175] (d) 如 SEQ ID No.2 或 SEQ ID No.32 的氨基酸残基 163-175 所示的氨基酸序列;

[0176] (e) 如 SEQ ID No.2 或 SEQ ID No.32 的氨基酸残基 304-311 所示的氨基酸序列;

或

[0177] (f) 与上述 (a)-(e) 定义的任意氨基酸序列有 75%或 75%以上, 优选 85%或 85%以上, 更优选 90%或 90%以上, 还更优选 95%或 95%以上同一性的氨基酸序列。

[0178] 适宜地, 本发明所述的脂质酰基转移酶可以包含下列核苷酸序列中的一个或多个序列表达而产生的氨基酸序列。

[0179] (a) 如 SEQ ID No.7 所示的核苷酸序列 (见图 9);

[0180] (b) 如 SEQ ID No.8 所示的核苷酸序列 (见图 10);

[0181] (c) 如 SEQ ID No.9 所示的核苷酸序列 (见图 11);

[0182] (d) 如 SEQ ID No.10 所示的核苷酸序列 (见图 12);

[0183] (e) 如 SEQ ID No.11 所示的核苷酸序列 (见图 13);

[0184] (f) 如 SEQ ID No.13 所示的核苷酸序列 (见图 15);

[0185] (g) 如 SEQ ID No.21 所示的核苷酸序列 (见图 17);

[0186] (h) 如 SEQ ID No.23 所示的核苷酸序列 (见图 19);

[0187] (i) 如 SEQ ID No.25 所示的核苷酸序列 (见图 21) ;

[0188] (j) 如 SEQ ID No.27 所示的核苷酸序列 (见图 23) ;

[0189] (k) 如 SEQ ID No.29 所示的核苷酸序列 (见图 25) ;

[0190] (l) 如 SEQ ID No.31 所示的核苷酸序列 (见图 27) ;

[0191] (m) 如 SEQ ID No.33 所示的核苷酸序列 (见图 29) ;

[0192] (n) 如 SEQ ID No.35 所示的核苷酸序列 (见图 31) ;

[0193] (o) 或

[0194] 与 SEQ ID No.7, SEQ ID No.8, SEQ ID No.9, SEQ ID No.10, SEQ ID No.11, SEQ ID No.13, SEQ ID No.21, SEQ ID No.23, SEQ ID No.25, SEQ ID No.27, SEQ ID No.29, SEQ ID No.31, SEQ ID No.33 或 SEQ ID No.35 所示的任意一序列具有 75% 或 75% 以上同一性的核苷酸序列。

[0195] 适宜地, 核苷酸序列可以与 SEQ ID No.7, SEQ ID No.8, SEQ ID No.9, SEQ ID No.10, SEQ ID No.11, SEQ ID No.13, SEQ ID No.21, SEQ ID No.23, SEQ ID No.25, SEQ ID No.27, SEQ ID No.29, SEQ ID No.31, SEQ ID No.33 或 SEQ ID No.35 所示的任意一序列具有 80% 或 80% 以上, 优选 85% 或 85% 以上, 更优选 90% 或 90% 以上, 还更优选 95% 或 95% 以上的同一性。

[0196] 一方面, 本发明所述的脂质酰基转移酶可以为卵磷脂: 胆固醇酰基转移酶 (LCAT) 或其变体 (例如分子进化产生的变体)。

[0197] 本领域所知的适宜的 LCAT 可从下列一种或多种生物体中得到, 例如: 哺乳动物、大鼠、小鼠、小鸡、黑腹果蝇 (*Drosophila melanogaster*)、植物, 包括拟南芥 (*Arabidopsis*) 和稻 (*Oryza sativa*)、线虫 (nematodes)、真菌和酵母。

[0198] 在一个具体实施方案中, 本发明所述的脂质酰基转移酶可以是并优选从含有 pPet12aAhydro 和 pPet12aASalmo 的大肠杆菌菌株 E.TOP 10 (*E.coli* strains TOP 10) 中获得的脂质酰基转移酶, 该菌株由丹麦哥本哈根 K, DK-1001, Langebrogade 1 的丹尼斯科公司 (Danisco A/S of Langebrogade 1, DK-1001 Copenhagen K, Denmark), 根据布达佩斯条约关于用于专利程序的国际认可的微生物保藏 (the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the purpose of Patent 方法) 在 2003 年 12 月 22 日保藏于英国苏格兰阿伯丁圣马恰尔街 23 号国立工业和海洋微生物保藏有限公司 (NCIMB) (the National Collection of Industrial, Marine and Food Bacteria (NCIMB) 23 St.Machar Street, Aberdeen Scotland, GB), 保藏号分别为 NCIMB 41204 和 NCIMB 41205。

[0199] 这里使用的术语“转移酶”与术语“脂质酰基转移酶”可以互换。

[0200] 适宜地, 本发明定义的脂质酰基转移酶催化下列一种或两种反应: 酯基转移作用 (transesterification), 醇解作用。

[0201] 根据本发明, 可获得一或多种以下的优越性质: 从脂质到形成一或多种碳水化合物酯, 蛋白质酯, 蛋白质亚基酯或羧基酯的生物转化可在高水分环境中发生, 所述环境不包含有机溶剂或包含于常规生物转化法相比而言较少量的有机溶剂。

[0202] 术语“生物转化”指修饰一种有机化合物以产生另一种有机化合物和 / 或通过酶催化从其它有机化合物合成有机化合物。

[0203] 本文术语“酯基转移作用”指的是酰基从脂质供体（除外游离脂肪酸）上经酶催化转移到酰基受体（除外水）。为避免怀疑，本文所用术语“酯基转移作用”指的是酰基在脂质供体与酰基受体（不是水）之间的酶催化转移，其中所述酰基受体包含适宜地化学基团，其可以为例如 -OH 或 -SH 基团。

[0204] 本发明使用的术语“醇解作用”指的是通过与醇基团 ROH 的反应对酸衍生物的共价键的酶裂解，使得一产物与醇基团的 H 结合而另一产物与醇基团的 OR 基团结合。

[0205] 本发明使用的术语“水解作用”指的是酰基从脂质到水分子 OH 基团的酶催化转移作用。由水解作用引起的酰基转移作用需要水分子的分离。

[0206] 术语“酯交换作用”指的是酰基在脂质供体与酰基受体之间的酶催化转移，其中的脂质供体不是游离的酰基。换言之“酯交换作用”指两个脂质分子之间的脂肪酸互换。

[0207] 一方面，本发明定义的脂质酰基转移酶催化酯交换作用。

[0208] 适宜地，本发明的方法或用途可进一步包含一或多种以下步骤：将酰基受体溶于水；将脂质酰基供体加入溶解的酰基受体，形成双相系统或乳液；对反应混合物进行搅拌或超声处理；加热反应混合物，例如，使得酶变性；通过标准分离技术将水相从脂肪/乳液相分离，所述标准分离技术诸如溶剂提取或水蒸发；通过疏水反应层析使得脂肪相分级，结晶或高真空蒸馏 (high vacuum distillation)。适宜地，一或多种加热，分离或分级步骤可在所述反应达到平衡以后实施。

[0209] 一个实施方案中，本发明方法所用的酯酶酰基转移酶是固定化的。当所述酶被固定的情况下，所述混合物包含酰基供体，酰基受体和流经柱子的水，所述水例如包含固定化的酶。通过将所述酶固定化，可能容易地再利用它。

[0210] 适宜地，所述固定的酶可用于流动反应器 (flow reactor) 中或批量反应器 (batch reactor) 中，所述反应器含有反应混合物，该混合物包含溶于水的酰基受体以及脂质酰基供体作为双相系统或乳液。反应混合物可选被搅拌或经超声处理。一旦反应到达例如平衡，所述反应混合物和固定化的酶可被分离。适宜地，所述反应产物可以例如通过疏水反应层析，结晶或高真空蒸馏来分级。

[0211] 固定化的脂质酰基转移酶可以使用本领域熟知的固定化技术制备。本领域有大量的对本领域技术人员而言清楚明白的制备固定化的酶的方法（例如以下文献中所述的技术：EP 0 746 608；或 Balcao VM, Paiva AL, Malcata FX., *Enzyme MicrobTechnol.*1996 May 1; 18(6): 392-416；或 Reetz MT, Jaeger KE.*Chem Phys Lipids.*1998 Jun; 93(1-2): 3-14；或 Bornscheuer UT, Bessler C, Srinivas R, Krishna SH.*TrendsBiotechnol.*2002 Oct; 20(10): 433-7；Plou et al, *J.Biotechnology* 92(2002)55-66；Warmuth et al., 1992.*Bio Forum* 9, 282-283；Ferrer et al., 2000. *J.Chem.Technol.Biotechnol.*75, 1-8；或 Christensen et al., 1998.*Nachwachsende Rohstoff* 10, 98-105；Petersen and Christenen, 2000, *Applied Biocatalysis.*Harwood Academic Publishers, Amsterdam。（将每篇都包含在本文中作为参照）。本文所用技术包括例如与 Eupergit C 的共价偶联，吸附于聚丙烯和硅颗粒。

[0212] 本文术语“高水分环境 (high water environment)”优选指低或缺乏有机溶剂的环境，优选低或缺乏极性有机溶剂。本文术语有机溶剂作为脂质底物时优选不包括食品

油，并有选不包括例如含高非极性脂质的食品油。适宜地，本发明高水分环境包含不超过 50% 体积的有机溶剂，不超过 30% 体积的有机溶剂，更优选不超过 15% 体积的有机溶剂，更优选不超过 5%，更优选不超过 1%，更优选不超过 0.5% 体积的有机溶剂，更优选 0% 体积的有机溶剂。

[0213] 当根据本发明产生碳水化合物酯时，优选所述碳水化合物酯是寡糖酯、单糖酯或二糖酯。

[0214] 适宜地，当根据本发明产生碳水化合物酯时，碳水化合物酯可以是以下的一种或多种：葡萄糖酯、果糖酯、脱水果糖酯、麦芽糖酯、乳糖酯、半乳糖酯、木糖酯、木寡糖酯 (xylooligosaccharide ester)、阿拉伯糖酯 (arabinose ester)、麦寡糖酯 (maltooligosaccharide ester)、塔格糖酯 (tagatose ester)、蔗糖酯、microthecin 酯、ascopyrone P 酯、ascopyrone T 酯、或 cortalcerone 酯。

[0215] 根据本发明优选产生的碳水化合物酯是以下一种或者多种：碳水化合物单酯 (碳水化合物 mono-ester)、糖单酯 (sugar mono-ester)、寡糖单酯、三糖单酯、二糖单酯、单糖单酯、葡萄糖单酯、果糖单酯、脱水果糖单酯、麦芽糖单酯、乳糖单酯、半乳糖单酯、木糖单酯、木寡糖单酯、阿拉伯糖单酯、麦寡糖单酯、塔格糖单酯、蔗糖单酯、microthecin 酯、ascopyrone P 酯、ascopyrone T 酯、或 cortalcerone 酯。

[0216] 在一个具体实施方案中，microthecin 酯、ascopyrone P 酯、ascopyrone T 酯、和 / 或 cortalcerone 酯可以作为抗微生物剂起作用。可选地或者进一步地，microthecin 酯、ascopyrone P 酯、ascopyrone T 酯、和 / 或 cortalcerone 酯可以作为抗氧化剂和 / 或乳化剂的一个或两个来起作用。

[0217] 优选地，本发明碳水化合物酯的形成 (如果有的话) 不依赖于 UDP- 葡萄糖。

[0218] 优选地，本发明的食品不包含 UDP- 葡萄糖或仅仅包含不显著量的 UDP- 葡萄糖。

[0219] 已发现本发明组合物和方法中应用的脂质酰基转移酶与脂解酶相比具有独特的性质，它们明显优先地将酰基从脂质转移到水以外的受体上，即使在有大量水存在的条件下也是如此。与现有技术中的酶相比发现，本发明中应用的脂质酰基转移酶在存在 6%、54%、73%、89% 和大约 95% 的水的条件下具有高的相对转移酶活性。被试脂解酶在这些水浓度条件下事实上不具有明显的相对转移酶活性。

[0220] % 转移酶活性 (即作为总酶活性百分比的转移酶活性) 可通过以下方案测定：

[0221] 测定转移酶活性百分率的实验设计：

[0222] 酶促反应后，可用  $\text{CHCl}_3$  :  $\text{CH}_3\text{OH}$  2 : 1 提取已添加本发明的脂质酰基转移酶的食品，分离包含脂质物质的有机相，并可根据在下文描述的详细步骤通过 GLC 和 HPLC 分析。通过 GLC 和 HPLC 的分析，确定游离脂肪酸和一种或多种固醇 / stanol 酯类；碳水化合物酯；蛋白质酯；甘油二酯；或单甘油酯的量。不添加本发明的酶的对照食品用同种方法分析。

[0223] 计算：

[0224] 从 GLC 和 HPLC 分析的结果可以计算出游离脂肪酸和碳水化合物酯和 / 或蛋白质酯和 / 或蛋白质亚基酯和 / 或羟基酸的增加量：

[0225]  $\Delta$  % 脂肪酸 = % 脂肪酸 (酶) - % 脂肪酸 (对照)； $M_v$  脂肪酸 = 脂肪酸平均分子

量

[0226]  $A = \Delta \% \text{蛋白质酯} / Mv \text{ 蛋白质酯}$  (其中  $\Delta \% \text{蛋白质酯} = \% \text{蛋白质酯 (酶)} - \% \text{蛋白质酯 (对照)}$ ),  $Mv \text{ 蛋白质酯} = \text{蛋白质酯的平均分子量}$ ), 这适用于酰基受体是蛋白质;

[0227]  $B = \Delta \% \text{碳水化合物酯} / My \text{ 碳水化合物酯}$  (其中  $\Delta \% \text{碳水化合物酯} = \% \text{碳水化合物酯 (酶)} - \% \text{碳水化合物酯 (对照)}$ ),  $Mv \text{ 碳水化合物酯} = \text{碳水化合物酯平均分子量}$ ), 这适用于酰基受体是碳水化合物。

[0228]  $C = \Delta \% \text{蛋白质亚基酯} / Mv \text{ 蛋白质亚基酯}$  (其中  $\Delta \% \text{蛋白质亚基酯} = \% \text{蛋白质亚基酯 (酶)} - \% \text{蛋白质亚基酯 (对照)}$ ),  $Mv \text{ 蛋白质亚基酯} = \text{蛋白质亚基酯平均分子量}$ ), 这适用于酰基受体是蛋白质亚基; 和

[0229]  $D = \Delta \% \text{羧基酯} / Mv \text{ 羧基酯}$  (其中  $\Delta \% \text{羧基酯} = \% \text{羧基酯 (酶)} - \% \text{羧基酯 (对照)}$ ),  $Mv \text{ 羧基酯} = \text{羧基酯的平均分子量}$ ), 这适用于酰基受体是羧基酸。

[0230] 转移酶活性以占总酶活性的百分率计算:

[0231]

$$\% \text{转移酶活性} = \frac{A^* + B^* + C^* + D^* \times 100}{A^* + B^* + C^* + D^* + \Delta \% \text{脂肪酸} / (Mv \text{ 脂肪酸})}$$

[0232] \* 指适当时删除。

[0233] 酶的酯酶和酰基转移酶活性可利用以下实验评估。由此, 具有本发明所述酶特征的脂质酰基转移酶可被获得 / 鉴定。

[0234] 经缓冲的底物中的转移酶测定法 (参见实施例 6)

[0235] 起脂质酰基转移酶作用以用于本发明组合物和方法中的酶可以用实施例 12 中教导的测定法常规鉴定出来。此测定方法将在下文称为“经缓冲的底物中的转移酶测定法”。在实施例 6 中对本发明源于杀鲑气单胞菌的脂质酰基转移酶进行分析并与本发明未涉及的一定范围的脂解酶进行比较。可以看出脂解酶中只有 LIPOPAN<sup>®</sup> F (Novozymes, 丹麦) 具有转移酶活性, 且仅具有较低的水平 (1.3%)。

[0236] 适于本发明的组合物和方法中应用的酶可以利用在经缓冲的底物中的转移酶测定法来常规确定。应用这种测定方法 (其中含水量非常高 - 大约 95%), 本发明应用的脂质酰基转移酶是至少有 2% 酰基转移酶活性 (相对转移酶活性), 优选至少 5% 相对转移酶活性, 优选至少 10% 相对转移酶活性, 优选至少 15%、20%、25%、26%、28%、30%、40%、50%、60% 或 75% 相对转移酶活性的酶。本发明适宜的脂质酰基转移酶可以具有小于 28%、小于 30%、优选小于 40%、50%、60%、70%、80%、90% 或 100% 的酰基转移酶活性。

[0237] 低水分蛋黄中的转移酶测定法 (参见实施例 11)

[0238] 作为“经缓冲底物中的转移酶测定法”的替换 (或附加), 本发明所用脂质酰基转移酶可以用“低水分蛋黄中的转移酶测定法”来鉴定。

[0239] 为了判定一种酶是否是属于根据本发明的脂质酰基转移酶, 可进行“低水分环境中的转移酶测定法”, 即如实施例 9 教导在含水 6% 的油性环境中。这个例子说明在水含量 6% 的油性环境中本发明的脂质酰基转移酶具有高的相对转移酶活性, 而现有技术的

脂解酶具有水解活性。

[0240] 在一具体实施方案中，本发明组合物和 / 或方法中应用的适宜脂质酰基转移酶是用“低水分环境中的转移酶测定法”检测的酶，在 30、20 或 120 分钟后进行测定时，其相对转移酶活性至少 1%，优选至少 2%，优选至少 5%，优选至少 10%，优选至少 20%，优选至少 30%，优选至少 40%，优选至少 50%，优选至少 60%，优选至少 70%，优选至少 75%。适宜地，用“低水分环境中的转移酶测定法”在 10、20、30 或 120 分钟后进行测定时，本发明脂质酰基转移酶，可有少于 30%、40%、50%、60%、70% 或 80% 的活性。

[0241] 如上所述，本发明中的脂酶酰基转移酶可以用“经缓冲的底物中的转移酶测定法”或“低水分环境中的转移酶测定法”以胆固醇为酰基受体进行鉴定。当然，本领域技术人员会容易意识到，通过对分析方法进行明显的修改，“经缓冲的底物中的转移酶测定法”或“低水分环境中的转移酶测定法”可用于确定脂质酰基转移酶对任何脂质酰基供体或任何酰基受体组合的活性。如有必要，技术人员可简单地用可选酰基供体底物（如糖脂，甘油三酯）替换酰基供体底物（如磷脂）和 / 或用可选酰基受体底物（如碳水化合物，蛋白质，蛋白质亚基或羧基酸）替换酰基受体底物（如胆固醇）。（例如见实施例 10-13）。

[0242] 本发明术语“高水分环境”指任何包含 5-98% 水的环境。优选所述环境包含不超过 6% 水含量，优选不超过 7%，8%，9%，10%，20%，30%，40%，50%，60%，70%，80% 或 90% 水含量。适宜地，所述高水分环境包含 20-98%，适宜地包含 50-98% 水，适宜地包含 70-98% 水，适宜地包含 75-98% 水。

[0243] 一个实施方案中，所述混合物中加入的脂质酰基转移酶的量与水的重量比至少 1 : 700，优选 1 : 10,000。

[0244] 本文使用的术语“低水分”意味着任何底物或食品，其水含量小于 6%，优选小于 5%、4%、3%、2%、1% 或 0.5%。

[0245] 优选本发明的方法和 / 或用途可在 15-60 °C 实施，优选在 20-60 °C，优选 20-50 °C，优选 20-45 °C，优选 20-40 °C。

[0246] 适宜地，本发明的方法或用途包含另一步骤或纯化和 / 或分离所述反应产物，即一或多种碳水化合物酯，蛋白质酯，蛋白质亚基酯，或羧基酸酯。因此，优选所述反应产物是纯化的和 / 或分离的形式。

[0247] 纯化酯的多种方法是本领域技术人员已知的。例如，只有通过本发明教导的方法 / 用途产生的酯可利用层析诸如疏水反应，过滤，离心，溶剂提取 / 蒸馏或结晶。适宜的方法见 Ulmann's Encyclopedia of Industrial Chemistry (2002)，Wiley-VCH Verlag GmbH & Co.KgaA。

[0248] 本发明所述脂质酰基转移酶可在任何适宜的表达宿主表达。例如本发明所述脂质酰基转移酶可在枯草芽孢杆菌 (*Bacillus subtilis*) 中表达，并且可以通过超滤法和 / 或乙醇沉淀法和 / 或离心法纯化，接着可以使用淀粉（麦芽糊精）作为酶的载体进行喷雾干燥。经喷雾干燥后的酶通过添加另外的粉末形式载体而标准化为规定的 PLU 活性。有关技术在本领域中是很明确且常规的。

[0249] 一个实施方案中，本发明的方法是体外方法。所述方法适宜地是连续或批量方

法。

[0250] 本发明的酶可以和一或多种其它酶联用。因此，除了本发明的酶以外，将所述混合物与至少一种其它的酶接触也包含在本发明的范围内。所述另外的酶包括淀粉降解酶，诸如内 (endo) 或外 (exo) 淀粉酶，支链淀粉酶 (pullulanases)，去支化 (debranching) 酶，半纤维素酶 (hemicellulases) 包括木聚糖酶 (xylanases)，纤维素酶 (cellulase)，氧化还原酶 (oxidoreductases)，例如葡萄糖氧化酶或碳水化合物氧化酶诸如麦芽糖的氧化酶，例如己糖氧化酶 (HOX)，酯酶，磷脂酶和己糖氧化酶，以及蛋白酶。所述混合物可以和本发明的酶以及至少一种其它的酶同时或依次接触。

[0251] 一个实施方案中例如所述脂质酰基转移酶可以和具有一或多种以下脂酶活性的脂酶联用：糖脂酶活性 (E.C.3.1.1.26)，三酰甘油酯酶活性 (E.C.3.1.1.3)，磷脂酶 A2 活性 (E.C.3.1.1.4) 或磷脂酶 A1 活性 (E.C.3.1.1.32)。适宜的酯酶是本领域已知的，并作为实例包括以下酯酶：**LIPOPAN**<sup>®</sup> F 和 / 或 **LECITASE**<sup>®</sup> ULTRA (Novozymes A/S, Denmark)，磷脂酶 A2 (例如磷脂酶 A2 来自 **LIPOMOD**<sup>™</sup> 22L 来自 Biocatalysts，**LIPOMAX**<sup>™</sup> 来自 Genecor)，**LIPOLASE**<sup>®</sup> (Novozymes A/S, Denmark)，酯酶的教导见 WO03/97835，EP 0977 869 或 EP 1 193 314。

[0252] 用途

[0253] 因此，本发明的方法产生一或多种碳水化合物酯，蛋白质酯，蛋白质亚基酯，羟基酸酯。这些酯中的许多是有用的乳化剂。作为实例只有例如氨基酸酯，肽酯，蛋白质酯，碳水化合物酯和羟基酸酯 (诸如酒石酸酯) 是功能重要的乳化剂。乳化剂可用于多种工业中，诸如食品工业，饲料工业，化妆品工业 (例如在化妆品基础中)，制药工业 (例如在药物合成和配制中) 以及例如颜料工业中。乳化剂可作为湿润剂，食品成分和活性成分起作用。

[0254] 此外，蛋白质脂肪酸缩合物由于它们优越的生理性质适于作在例如化妆品和个人卫生产品中。例如，蛋白质酯可用于沐浴制品以及香波和身体清洁剂中。蛋白质脂肪酸缩合物也可用于药物组合物，例如作为基质。

[0255] 已知蛋白质脂肪酸缩合物可用于化妆品工业中。通常，这些产品通过利用水作为溶剂，使蛋白质水解产物与脂肪酸氯化物在 Schotten-Baumann 条件下反应来制备。

[0256] ([http://www.scf-online.com/english/26\\_e/rawmaterial26\\_e.htm#5](http://www.scf-online.com/english/26_e/rawmaterial26_e.htm#5))。

[0257] 在开发蛋白质 - 脂肪酸缩合物的过程中，可能将可再生来源脂肪酸 (来自植物油) 与蛋白质结合，其可获自动物废弃物 (皮) 以及来自许多植物，来构建具有疏水 (脂肪酸) 和亲水 (蛋白质) 部分的表面活性剂结构。在该过程中，脂肪酸氯化物 (fatty acid chloride) 与氨基酸的氨基反应，并形成蛋白质脂肪酸缩合物 (见图 49)。获得具有很好的皮肤相容性以及另外具有好的清洁效果的产品。

[0258] 甚至加入较少的酰基化蛋白质水解产物对其它表面活性剂的皮肤相容性具有协同作用这一事实从技术配制角度来看是非常重要的。该保护作用可由该产物的两性特性来解释。蛋白质 - 脂肪酸缩合物与皮肤胶原之间存在相互作用。这导致保护性层的形成，其减少了表面活性剂对上皮层的过度攻击，它们强的去油效应以及阴离子表面活性剂与皮肤的直接相互作用。

[0259] 在化妆品工业中，基于蛋白质的表面活性剂主要用于温和的沐浴产品 (shower

and bath product), 温和的香波, 基于表面活性剂的洗面剂, 冷-剃须制品和固定剂或婴儿用表面活性剂制品。

[0260] 蛋白质水解产物脂肪酸缩合物也可用作药物制剂的基质, 例如用作含有用于局部给药皮肤的活性成分的乳膏或油膏。

[0261] 本发明提供无需利用脂肪酸氯化物制备蛋白质脂肪酸缩合物的新方法。本发明的反应的描述见图 50。所述反应可在水或缓冲体系中在低温进行而不形成废物。

[0262] 本文术语“蛋白质脂肪酸缩合物”包括所有以下物质: 蛋白质酯, 多肽酯, 二肽酯, 寡肽酯, 肽酯, 以及氨基酸酯。

[0263] 本领域技术人员已知, 碳水化合物酯(具体是糖酯)在食品工业中具有广泛的应用。其它应用领域包括化妆品, 口腔护理产品, 和医学用品。此外这些化合物可用作抗生素, 抗肿瘤剂, 杀真菌剂和杀虫剂。本发明的脂质酰基转移酶能催化在高水分环境中形成葡萄糖糖酯(图 51)。

[0264] 本发明制备的酯可用于以下领域:

[0265] 化妆品: 包括精华油乳液(essential oil emulsion)(o/w, HLB 16-18)石蜡油乳液, o/w, HLB 10-14; 硬脂酸乳液; 蜡乳液(wax emulsion), o/w, HLB14-16; 羊毛脂乳液, o/w, HLB 12-14; 硅酮乳液; 牙膏, o/w; 泡沫沐浴品(foam bath), o/w, HLB 14-18; 洗发液。

[0266] 药物制剂: 包括药物乳剂; 油膏基质; 栓剂化合物, w/o; 包囊化(encapsulation); 注射制剂。

[0267] 农业: 包括土壤改良; 作为增肥添加剂(fertiliser additive); 作为全能洗涤剂(all-purpose cleaner); 水果和蔬菜的洗涤剂; 搅乳器用清洁物。

[0268] 农作物保护: 包括天然存在的杀虫剂; 氯化烃(chlorinated hydrocarbon), 和 140; 磷酸酯 o/w, HLB 10-14; 杀真菌剂, o/w; 除莠剂, o/w。

[0269] 食品工业: 包括面包和蛋糕; 人造黄油; 巧克力; 脂肪粉霜预防(fat bloom prevention), w/o, HLB 5-10; 糖霜化(sugar frosting), o/w, HLB 14-16; 焦糖(caramel)和口香糖的软化剂, w/o, HLB.2-4; 粘连预防, w/o, HLB2-4; 冰淇淋添加剂 w/o, HLB 4-6; 奶和烘焙粉的湿化, w/o, HLB 9-11; 蛋奶(custard)粉, w/o, HLB 2-4; 在饮料工业中; 在水果和蔬菜中; 在调味剂中, w/o 和 o/w, HLB 10-12; 在肉, 沙拉或其它调味酱中, o/w; 在食品色素中, w/o, HLB 2-4; o/w, HLB 8-18; 在泡沫抑制剂中。

[0270] 利用本发明制备的蛋白质脂肪酸酯, 羧基酸酯和碳水化合物酯作为食品应用中的乳化剂的益处是这些是无害的食品相容性成分, 它们与其它常规使用的乳化剂例如乙氧基化的脂肪酸酯相比更容易发生生物降解。因此, 这些乳化剂在食品工业和非食品工业应用中更为环保。

[0271] 在一个具体实施方案中, microthecin 酯、ascopyrone P 酯、ascopyrone T 酯、和 / 或 cortalcerone 酯可以作为抗微生物剂起作用。可选地或者进一步地, microthecin 酯、ascopyrone P 酯、ascopyrone T 酯、和 / 或 cortalcerone 酯可以作为抗氧化剂和 / 或乳化剂的来起作用。

[0272] 一个实施方案中, 本发明的方法和用途可用于利用脂质酰基-转移酶制备用于

药物配制剂中的乳化剂，具体是用于活性成分的控制释放的产生中，其中所述活性成分是酰基化的。这种慢释放型制剂具体可用于口服给药的药物组合物，其中所述酯在消化道中的逐渐水解可逐渐释放活性成分。所述酰基化的组合物可进一步用于皮下或静脉内制剂。

[0273] 另一实施方案中，本发明的方法或用途可用于制备相转移催化剂用于例如在有机反应中将盐转移到有机溶剂溶液中。例如，将酰基基团转移到适当的阳离子受体，诸如羧酸（柠檬酸），或可选具有阴离子受体基团，诸如羟基-胺可产生用于将盐转移到有机溶剂溶液中的相转移催化剂。

[0274] 另一实施方案中，本发明的方法可用于制备具有低生物利用度和 / 或低溶解度的药物化合物的前药，例如抗病毒剂如阿昔洛韦 (acyclovir) 和更昔洛韦 (ganciclovir)。所述方法可进一步用于其它具有游离羟基的药物化合物，例如伯，仲，叔羟基。

[0275] 优选，本发明制备的酯可用于药物配制。

[0276] 优选，本发明制备的酯可用于化妆品和 / 或个人卫生用品中。

[0277] 优选，本发明制备的酯可用于食品和 / 或饲料中。

[0278] 本发明的方法可以是在生产一或多种药物，化妆品，个人卫生用品，食品或饲料的方法中的一个步骤。

[0279] 优点

[0280] 本发明方法的一个优点是其导致一或多种碳水化合物酯，蛋白质酯，蛋白质亚基酯或羧基酯的生产而无需利用有机溶剂。因此，本发明允许利用较少的有机溶剂或不利用有机溶剂。这具有许多优点，例如降低生产成本，降低人和 / 或环境的有机溶剂暴露，简化生产过程。

[0281] 在用于食品应用的酯的生产中，具体优选利用脂质而不是脂肪酸，因为不必要去除过多的脂质，这是由于它们可来自利用反应产物的食品的一部分。另一方面，多余的游离脂肪酸必须被去除，这是由于它们对于大多数食品是有害的。

[0282] 分离

[0283] 在一方面，优选地，本发明应用的多肽或蛋白质是分离形式的。术语“分离的”意思是序列至少基本上不含有至少一种其它成分，该成分于此序列天然相结合，并在自然界发现。

[0284] 一方面，优选本发明的生物转化产品例如碳水化合物酯和 / 或蛋白质酯和 / 或蛋白质亚基酯和 / 或羧基酯分离自反应混合物。术语“分离的”指所述生物转化产物至少基本不含至少一种其它组分，所述生物转化产物在生物转化反应过程中与该其它组分结合。

[0285] 纯化

[0286] 在一方面，优选地，本发明应用的多肽或蛋白质是纯化的。术语“纯化的”意思是序列处于相对纯化的状态 - 如至少大约 51% 纯度、或至少大约 75% 纯度、至少大约 80% 纯度、至少大约 90% 纯度、至少大约 95% 纯度、至少大约 98% 纯度。

[0287] 一方面，优选本发明制备的生物转化产物，例如碳水化合物酯和 / 或蛋白质酯和 / 或蛋白质亚基酯和 / 或羧基酯自反应混合物纯化，并因此为纯化的形式。术语“纯化的”指所述生物转化产物是相对纯的状态 - 例如至少约 51% 纯，或至少约 75%，或至

少约 80%，或至少约 90% 纯，或至少约 95% 纯或至少约 98% 纯。

[0288] 药物组合物

[0289] 本发明还提供了药物组合物，其包含本发明的产物和可药用的载体，稀释剂或赋形剂（包括其组合）。

[0290] 所述药物组合物可在人或兽医医学中用于人或动物并通常包含任何一或多种可药用稀释剂，载体或赋形剂。治疗用可接受的载体或稀释剂是制药领域已知的，并在例如 Remington's Pharmaceutical Sciences, Mack Publishing Co. (A.R.Gennaro edit.1985) 中描述。药物载体，赋形剂或稀释剂可根据目的给药途径和标准药物操作来选择。所述药物组合物可包含或另外包含，载体，赋形剂或稀释剂，任何适宜的粘合剂，润滑剂，助悬剂，包被剂，助溶剂。

[0291] 防腐剂，稳定剂，染料甚至调味剂可在药物组合物中提供。防腐剂的实例包括苯甲酸钠，山梨酸，以及对羟基苯甲酸的酯。抗氧化剂和助悬剂也可使用。

[0292] 根据不同的递送系统，有不同的组合物 / 配制要求。例如，本发明的药物组合物可配制成利用微型泵或通过粘膜途径给药的形式，例如作为喷鼻剂或吸入气溶胶或可摄入溶液，或经胃肠外给药，其中所述组合物可配制成可注射形式用于例如经静脉内，肌内或皮下途径来递送。可选，所述配制剂可设计为通过多种途径给药。

[0293] 当所述药剂经胃肠道粘膜给药时，其应当能够在通过胃肠道的过程中保持稳定；例如其应抵抗蛋白水解降解，在酸性 pH 稳定并抵抗胆汁的清洁作用。

[0294] 适当的情况下，所述药物组合物可通过吸入给药，以栓剂或阴道栓的形式，以洗液，溶液，乳膏，油膏或粉尘的形式，通过利用皮肤贴片 (patch)，以含有赋形剂诸如淀粉或乳糖的口服片剂，或以单独或与赋形剂混合的胶囊或卵形胶囊 (ovule) 的形式，或以含有调味或着色剂的酏剂，溶液或悬液的形式，或它们可经胃肠外例如经静脉内，肌内或皮下注射。对于经胃肠外给药，所述组合物最好以含有其它物质的无菌含水溶液形式使用，所述其它物质例如足够的盐或单糖使得所述溶液与血液等张。对于经颊或经舌下给药，所述组合物可以以可通过传统方式配制的片剂或锭剂形式给药。

[0295] 克隆编码本发明所述多肽的核苷酸序列

[0296] 编码具有本文定义的具体性质的多肽或适合被修饰的多肽的核苷酸序列，可以从产生上述多肽的细胞或生物体中分离。本领域有多种不同的已知的方法分离核苷酸序列。

[0297] 例如，基因组 DNA 和 / 或 cDNA 库可以应用由源自产生多肽的生物体的染色体 DNA 或信使 RNA 构建。如果多肽氨基酸序列是已知的，可以合成标记的寡核苷酸探针，并用于识别从生物体制备的基因组文库的多肽编码克隆。可替代的，包含与另一个已知的多肽基因同源的氨基酸序列的标记的寡核苷酸探针可以用于识别多肽编码型克隆。在后面的例子中，应用低严谨性的杂交和洗涤条件。

[0298] 可替代的，多肽编码克隆能够通过将基因组 DNA 片段插入表达载体如质粒中进行鉴别，使用得到的基因组 DNA 文库转化不含酶的细菌，然后将转化后的细菌涂布于含有一种被多肽抑制的酶琼脂皿，从而使表达多肽的克隆被鉴定。

[0299] 作为另外的选择，编码多肽的核苷酸序列可以用确定的标准方法合成制备，该标准方法例如 Beucage S.L. 等 (1981) Tetrahedron Letters 22, p1859-1869 描述的亚磷酰胺

(phosphoroamidite) 方法, 或 Matthes 等 (1984) EMBO J.3, p 801-805 描述的方法。在亚磷酰胺方法中, 寡核苷酸被合成 (例如在自动 DNA 合成器中)、纯化、退火、连接并克隆到适宜的载体中。

[0300] 核苷酸序列可以是混合的基因组和合成来源, 混合的合成与 cDNA 来源, 或混合的基因组与 cDNA 来源, 其可通过使用标准技术连接合成 DNA、基因组 DNA (genomic DNA) 或 cDNA 来源 (适当的) 的片段来制备。每一个连接片段对应于整个核苷酸序列中不同的位置。DNA 序列也可通过聚合酶链式反应 (PCR) 使用特异性的引物制备, 例如 US 4,683,202 中所描述的或在 Saiki R K 等 (Science(1988)239, pp 487-491) 中所描述的。

[0301] 核苷酸序列

[0302] 本发明也包括具有这里定义的特殊性质的核苷酸序列编码多肽。这里使用的术语“核苷酸序列”指寡核苷酸序列或多核苷酸序列, 它的变体、同族体、片段和它的衍生物 (例如它的部分)。该核苷酸序列可以属于基因组的、合成的、或重组的来源, 不论是代表同义链或反义链, 它可以是双链的或是单链的。

[0303] 关于本发明使用的术语“核苷酸序列”包含基因组 DNA、cDNA、合成 DNA, 和 RNA。优选的是 DNA, 更优选为编码序列的 cDNA。

[0304] 一个优选的实施方案中, 当结合到也存在它的自然环境中的、与其相结合的序列时, 编码具有本文定义具体特性的多肽的核苷酸序列本身不涵盖自然条件下存在的天然核苷酸序列。为了便于参考, 我们称这种优选的实施方案为“非天然核苷酸序列”。在这种条件下, 术语“天然核苷酸序列”是指在其自身自然环境中的完整核苷酸序列, 并且条件是与其天然所结合的完整启动子可操作地连接, 所述启动子也存在于其自身天然环境中。因此, 本发明的多肽可以通过存在于自然生物体中的核苷酸序列表达, 但是其中的核苷酸序列不受该生物体中与其天然结合的启动子的控制。

[0305] 优选的多肽不是一种天然多肽。这种条件下, 术语“天然多肽”是指一种处于其自身的自然环境中的完整多肽, 且条件是它被其天然核苷酸序列表达。

[0306] 一般地, 编码具有本发明定义的具体性质的多肽的核苷酸序列可以通过重组 DNA 技术 (即重组 DNA) 制备。然而, 在本发明一个可选择的实施方案中, 核苷酸序列的全部或部分可以应用本领域已知的化学方法合成 (参考 Caruthers MH 等 (1980) Nuc Acids Res Symp Ser 215-23 and Horn T et al (1980) Nuc Acids Res Symp Ser 225-232)。

[0307] 分子进化

[0308] 一旦编码酶的核苷酸序列被分离, 或者推定的编码酶的核苷酸序列被鉴定, 修改选定的核苷酸序列是合乎需要的, 例如为了制备本发明所述的酶而使核苷酸序列突变是合乎需要的。

[0309] 突变可以由合成的寡核苷酸引入。这些寡核苷酸包含位于所需的突变位点侧翼的氨基酸序列。

[0310] 适宜的方法被 Morinaga 等在 (Biotechnology(1984)2, p646-649) 中公开。另一种将突变引入编码酶的核苷酸序列的方法记述在 Nelson 和 Long 的 (Analytical Biochemistry(1989), 180, p147-151)。

[0311] 作为定点突变的替代, 例如上面所述, 可以随机引入突变, 例如使用商业试

剂盒诸如 Stratagene 生产的 GeneMorph PCR 突变试剂盒，或 Clontech 生产的多样化 (Diversify) PCR 随机突变试剂盒。EP 0 583 265 涉及对基于 PCRd 诱变的优化，它能与突变 DNA 类似物 (例如 EP 0 866 796 中描述的那些物质) 的应用相结合。易错聚合酶链式反应 (Error prone PCR) 技术适宜于生产具有优选特征的脂质酰基转移酶变体。WO0206457 涉及脂酶分子进化。

[0312] 获得新序列的第三种方法是非同一的氨基酸序列的片段化，或应用任何数量的限制性内切酶，或诸如 DNase I 的酶，并重新组装编码功能蛋白的全部氨基酸序列。或者，可以引入一种或多种不可识别的氨基酸序列，并在重新组装编码功能蛋白的全部氨基酸序列的过程中引入突变。DNA 改组和 DNA 家族改组 (DNA shuffling and family shuffling) 技术适宜于生产具有优选特性的脂质酰基转移酶的突变体。适宜的实现“改组” (‘shuffling’) 的方法被 EPO 752 008、EP1 138 763、EP1 103 606 公开。如 US 6,180,406 和 WO01/34835 所描述的，重排也可以与 DNA 突变的其它形式相结合。

[0313] 因此，能在体内或在体外在核苷酸序列中产生大量的定点突变或随机突变，随后用不同的方法筛选功能改善的编码多肽。应用 *silico* 和 *exo* 介导的重组方法 (见 WO00/58517, US 6,344,328, US 6,361,974)，例如，分子进化能够在产生的突变体保持与已知的酶或蛋白十分低的同源性的情况下完成。这些获得的突变体与已知的转移酶具有显著的结构近似性，但是具有十分低的氨基酸序列同源性。

[0314] 作为一个非限制的例子，另外地，多肽序列的突变或天然突变体能够被整合到或野生型或其它突变或天然突变体中用于生产新的突变体。这种新的突变体也被筛选出来以改善编码多肽的功能。

[0315] 通过上述提到的应用和类似的分子进化方法，可在没有任何关于蛋白质结构或功能的已知知识的情况下，鉴别和选择具有优选的性质本发明所述酶的突变体，同时可产生不可预测的但有益的突变或突变体。本领域有关分子进化的应用中有很多有关酶活力优化或改变的例子。这些例子包括，但不限于以下一种或多种，优化在宿主细胞内或体外优化的表达或活性，增加酶活性，改变酶作用底物和 / 或产品特异性，提高或降低酶或结构的稳定性，在优选的环境条件下 (例如温度、pH、酶作用底物) 酶活性 / 特异性的改变。

[0316] 应用分子进化工具可使酶发生改变以改善酶的功能，这对于本领域的技术人员是显然的。

[0317] 适宜地，本发明中使用的脂质酰基转移酶可以是一种突变体，即与其亲代酶比较，可以包含至少一个氨基酸置换、缺失或添加。突变酶保持与其亲代酶至少 1%、2%、3%、5%、10%、15%、20%、30%、40%、50%、60%、70%、80%、90%、95%、97%、99% 的同源性。适宜的亲代酶可包括任何具有酯酶或脂酶活性的酶。优选地，亲代酶与 Pfam00657 共有序列一致。

[0318] 在一个优选的实施方案中，脂质酰基转移酶突变体在 GDSx、GANDY 和 HPT 区保留或掺入至少一个或多个 Pfam00657 共有序列氨基酸残基。

[0319] 酶，例如在水环境中无或低脂质酰基转移酶活性的脂酶可应用分子进化工具进行突变，引入或增强转移酶活性，因此产生适用于本发明的组合物和方法的具有显著转移酶活性的脂质酰基转移酶。

[0320] 适宜地，用于本发明的脂质酰基转移酶可以是一种突变体，这种突变体与其亲代酶相比，对于极性脂质，优选磷脂和 / 或糖脂具有增强的酶活性。 优选地，这种突变体优选对于溶解 (lyso) 极性脂质表现出低或无活性。 对于极性脂质，磷脂和 / 或糖脂的增强的活性可能是水解作用和 / 或转移酶活性的结果。

[0321] 与其亲代酶相比，用于本发明的脂质酰基转移酶突变体可能减弱了对甘油三酯和 / 或甘油单酯和 / 或甘油二酯的活性。

[0322] 适宜地，酶突变体对甘油三酯和 / 或甘油单酯和 / 或甘油二酯无活性。

[0323] 或者，本发明应用的酶突变体可具有增强的对甘油三酯的活性，和 / 或也增强对下列一种或多种物质的活性，极性脂质、磷脂、卵磷脂、磷脂酰胆碱、糖脂、二半乳糖基甘油单酯、单半乳糖基甘油单酯。

[0324] 已知的脂质酰基转移酶突变体，和一种或多种所述突变体可适用于本发明的方法和用途和 / 或用于本发明的酶组合中。 仅仅作为举例，在下列参考文献中描述的脂质酰基转移酶突变体可以适用于本发明：

[0325] Hilton S, Buckley JT. Studies on the reaction mechanism of a microbial lipase/acyltransferase using chemical modification and site-directed mutagenesis. *J Biol Chem.* 1991 Jan 15 ; 266 (2) : 997-1000。

[0326] Robertson DL, Hilton S, Wong KR, Koepke A, Buckley JT. Influence of active site and tyrosine modification on the secretion and activity of the *Aeromonas hydrophila* lipase/acyltransferase. *J Biol Chem.* 1994 Jan 21 ; 269 (3) : 2146-50。

[0327] Brumlik MJ, Buckley JT. Identification of the catalytic triad of the lipase/acyltransferase from *Aeromonas hydrophila*. *J Bacteriol.* 1996 Apr ; 178 (7) : 2060-4。

[0328] Peelman F, Vinaimont N, Verhee A, Vanloo B, Verschelde JL, Labeur C, Seguret-Mace S, Duverger N, Hutchinson G, Vandekerckhove J, Tavernier J, Rosseneu M. A proposed architecture for lecithin cholesterol acyl transferase (LCAT) : identification of the catalytic triad and molecular modeling. *Protein Sci.* 1998 Mar ; 7 (3) : 587-99。

[0329] 氨基酸序列

[0330] 本发明还包括具有本文定义具体特性的多肽的氨基酸序列。

[0331] 这里使用的术语“氨基酸序列”与术语“多肽”和 / 或术语“蛋白质”同义。在某些情况下，术语“氨基酸序列”与术语“肽”同义。

[0332] 氨基酸序列可从适宜的来源制备 / 分离，或者可以通过合成制备或可通过应用 DNA 重组技术制备。

[0333] 适宜地，在此处提到的氨基酸序列可通过标准技术从分离出的本文教导的多肽中获得。

[0334] 以下是测定分离的多肽中确定氨基酸序列的一种适宜方法：

[0335] 纯化的多肽可为冰冻干燥的，100  $\mu$ g 冰冻干燥的原材料溶于 50  $\mu$ l 由 8M 尿素和 0.4M 碳酸氢铵组成的混合溶液中，pH 8.4。溶解的蛋白质 50 $^{\circ}$ C 变性还原 15 分钟，然后以氮气覆盖，添加 5  $\mu$ l 45mM 的二硫苏糖醇。经冷却至室温后，在室温，黑暗，氮气环境中加入 5  $\mu$ l 100mM 的碘乙酰胺，使半胱氨酸残基衍生 15 分钟。

[0336] 将 135  $\mu$ l 的水与溶于 5  $\mu$ l 水的 5  $\mu$ g 赖氨酸 -C 内蛋白酶加入上述反应混合物，

在氮气环境中 37°C 消化 24 小时。

[0337] 得到的多肽通过反相 HPLC 在 VYDAC 碳 -18 柱 (0.46x15cm ; 10 μ m ; The Separation Group, California, USA) 分离, 使用水中的溶剂 A : 0.1 % TFA 和乙腈中的溶剂 B : 0.1 % TFA。在氨基末端测序之前, 可对选出的多肽利用相同溶液体系通过 Develosil 碳 -18 柱再进行色谱分析。可以根据生产厂家的技术说明书 (Applied Biosystems, California, USA) 利用脉冲液态快速循环, 使用 Applied Biosystems 476A 序列分析仪进行测序。

[0338] 序列同一性或序列同源性

[0339] 本发明还涉及与具有本文定义的具体性质多肽的氨基酸序列或任何编码这种多肽的核苷酸序列的氨基酸序列 (后面称为“同源序列”) 具有一定程度的序列同一性或序列同源性的序列的用途。这里, 术语“同源体”指的是与受试氨基酸序列和受试核苷酸序列具有一定同源性的实体。这里, 术语“同源性”可与“同一性”等同。

[0340] 同源氨基酸序列和 / 或核苷酸序列应提供和 / 或编码保持功能活性和 / 或增强酶活性的多肽。

[0341] 在本文中, 指定同源序列包含与受试序列具有至少 75、85 或 90% 一致性, 优选 95 或 98% 一致性的氨基酸序列。一般地, 同源序列包括与受试氨基酸序列相同的活性位点等。虽然同源性也可被认为是相似性 (即具有相似的化学性质 / 功能的氨基酸残基), 在本发明的上下文中, 优选以序列的同一性表述同源性。

[0342] 在本发明的上下文中, 指定同源序列包括与编码本发明多肽的核苷酸序列 (主体序列) 具有至少 75, 85 或 90% 的同一性, 优选 95 或 98% 的同一性的核苷酸序列。一般地, 同源序列包括与受试序列相同的编码活性位点及其它的序列。虽然同源性也可被认为是相似性 (即具有相似的化学性质 / 功能的氨基酸残基), 在本发明的上下文中, 优选以序列的同一性表述同源性。

[0343] 同源性比较可用肉眼进行, 或更普遍的是, 借助与容易获得的序列比对软件进行比较。这种市场上可买到的电脑软件可计算出两个或更多个序列之间的 % 同源性。

[0344] % 同源性可在相邻的序列中计算, 即一条序列与另一条对齐, 一条序列的每一个氨基酸直接与另一条的氨基酸比较, 每次一个残基。这被称为“无间隙”排列。一般地, 这种无间隙排列只对相对较少数目的残基进行。

[0345] 虽然这是一种非常简单和稳定的方法, 但不能考虑到例如在否则同一的序列对中, 插入和缺失将导致后面的氨基酸残基不能对齐, 因此当总体比对已经完成时, 很可能地导致同源性百分比大幅度下降。因此, 为产生最佳排列而设计的许多序列比较方法考虑到可能存在的插入和缺失, 而不会过分降低整体同源性得分。这是通过在序列中插入“缺口”, 最大化局部同源性来实现的。

[0346] 但是, 这些更为复杂的方法将“缺口罚分” (“ gap penalties”) 赋予每一个在排列上出现的缺口, 使得对于相同数目的同一性氨基酸, 具有尽可能少的缺口数的比对 (反映两条被比较序列之间更高的相关性) 比缺口数多的比对可得到更高的分数。一般应用的“亲和缺口成本” (“ Affine gap costs”) 缺口的存在的罚分较高, 而对缺口后的每个后续的残基罚分较少。这是最通常应用的缺口评分系统。高缺口罚分当然产生具有较少缺口的优化排列。大多数排列程序允许对缺口罚分进行更改。然而, 使用这些软件进

行序列比较时优选使用缺省值。例如，当使用 GCG Wisconsin 最佳拟合 (Bestfit) 程序包时，缺省的氨基酸序列缺口罚分为每个缺口 -12 和每个延伸 -4。

[0347] 计算最大同源性百分比首先需要考虑到缺口罚分，产生最佳排列，实现这样排列的适宜的计算机程序是 GCG Wisconsin 中最佳拟合程序包 (Devereux 等 1984 *Nuc.Acids Research* 12p387)。其它可进行序列对比的软件实例，包括但不限于，BLAST 程序包 (参见 Ausubel et al 1999 *Short Protocols* 中 *Molecular Biology*, 4<sup>th</sup> Ed-Chapter 18)，FASTA (Altschul et al 1990 *J.Mol.Biol.*403-410) 和 GENWORKS 比较工具组。BLAST 和 FASTA 都可以脱机和在线检索 (参考 Ausubel 等 1999, pages 7-58 to 7-60)。然而，在一些应用中优选使用 GCG 最佳拟合程序。名为 BLAST 2 Sequence 的新工具也可以用于比较蛋白质和核苷酸序列 (参见 *FEMS Microbiol Lett* 1999 174 (2) : 247-50 ; *FEMS Microbiol Lett* 1999 177 (1) : 187-8 和 *tatiana&commat ; ncbi.nlm.nih.gov*)

[0348] 尽管最终同源性百分比能够以同一性的方式测定，比对程序通常不是基于全有或全无 (all-or-nothing) 的成对比较。作为替代，成比例的相似性分数矩阵是普遍应用的，该矩阵为基于化学相似性和进化距离的每个配对比较分配分值。平常应用的这样一种矩阵的一个实例就是 BLOSUM62 矩阵 -BLAST 程序组件的缺省矩阵。如果能提供的话 (想得到进一步的纤细信息，参见用户手册)，GCG Wisconsin 程序通常或应用公开的缺省值或应用定制的标记比较表。在一些应用中，优选使用 GCG 程序包公开的缺省值，或者在应用其它软件的情况下，使用缺省的矩阵，如 BLOSUM62。

[0349] 可选择地，同源性百分比可以在 DNASIS™ (日立公司软件) 通过序列特征多重对比来计算，基于一种类似于 CLUSTAL (Higgins DG & Sharp PM (1988), *Gene* 73 (1), 237-244) 的算法。

[0350] 一旦软件产生了一个最佳的排列，就能够计算序列同源性百分比，优选序列同一性百分比。通常软件进行部分序列的比较，产生一个数值型结果。

[0351] 序列也可以进行氨基酸残基的删除、插入或置换，这会产生一个沉默变化并导致生成一个功能上等同的物质。只要物质保持第二结合活性，考虑周全的氨基酸置换可以基于在极性、电荷、溶解度、疏水性、亲水性、和 / 或亲水亲油性方面残基性质的近似性进行。例如，带负电荷的氨基酸包括门冬氨酸和谷氨酸；带正电荷的氨基酸包括赖氨酸和精氨酸；具有相似亲水性并具有不带电荷的极性头基团的氨基酸包括亮氨酸、异亮氨酸、缬氨酸、甘氨酸、丙氨酸、天门冬酰胺、谷氨酰胺、丝氨酸、苏氨酸、苯丙氨酸、和酪氨酸。

[0352] 可以进行保守置换，例如下面表中所述。第二栏中同一区的氨基酸，优选第三栏中同一行的可以互相取代。

[0353]	脂肪族的	非极性的	GAP
			ILV
		极性不带电核的	CSTM
			NQ
		极性带电核的	DE
			KR
芳香族的		HF WY	

[0354] 本发明也包含同源取代的产生，（取代和置换都为此处应用的，即用另外的残基取代现有的氨基酸残基）即相似取代，例如碱性氨基酸取代碱性氨基酸，酸性氨基酸取代酸性氨基酸，极性氨基酸取代极性氨基酸等。也可以发生非同源取代，即从一类残基到另一类或可选择地涉及非天然氨基酸例如鸟氨酸（下文中记为 Z）、二氨基丁酸鸟氨酸（下文中记为 B）、原亮氨酸鸟氨酸（下文中记为 O）、pyriylalanine、噻吩基丙氨酸（thienylalanine）、萘基丙氨酸（naphthylalanine）和苯基甘氨酸。

[0355] 取代也可以通过非天然氨基酸产生。

[0356] 氨基酸序列的突变体可以包含适宜的可插入任意两个氨基酸残基序列间的间隔基团包括烷基基团诸如甲基、乙基、丙基，以及氨基酸间隔物，例如甘氨酸或丙氨酸残基。进一步的突变形式是本领域技术人员容易理解的，该突变形式包含类肽（peptoid）形式的一种或多种氨基酸残基。为了避免产生疑问，“类肽形式”用于指代变体氨基酸残基，其中  $\alpha$ -碳取代基团在该残基的氮原子上而非  $\alpha$ -碳上。制备类肽形式的肽是本领域已知的，例如 Simon RJ 等., PNAS(1992)89(20), 9367-9371 和 Horwell DC, Trends Biotechnol.(1995)13(4), 132-134。

[0357] 本发明中应用的核苷酸序列或编码具有本文定义的具体性质的多肽的核苷酸序列可以包括合成的或修饰的核苷酸。寡核苷酸的许多不同种类的修饰是本领域已知的。这些包括甲基磷酸酯和磷硫酰骨架和 / 或在分子的 3' 和 / 或 5' 端附加的吡啶或聚赖氨酸链。为了本发明所述的目的，本领域任何可得的方法均可以用于修饰本文描述的氨基酸序列，这是可以理解的。这些修饰可以用于提高核苷酸序列在体内的活性或延长氨基酸序列的寿命。

[0358] 本发明也涉及氨基酸序列作为本文讨论的序列，或任何衍生物、片段及其衍生物的互补序列的用途。如果序列为它的片段的互补序列，那么该序列可以用作探针鉴定其它生物体中相似编码序列的存在等。

[0359] 与本发明非 100% 同源但落入发明保护范围的多核苷酸序列，是可以通过多种途径获得的。此处描述的序列的其它突变体也可以获得，例如通过检测来自不同范围个体例如来自不同种群的个体的 DNA 文库。而且，其它病毒的 / 细菌的，或细胞的同源体特别是哺乳动物细胞（例如大鼠、小鼠、牛或灵长类动物细胞）的细胞同源物可以获得，此同源体和它的片段一般而言能够选择性地与本文列表所列的序列杂交。此序列可以通过检测其它动物物种的 cDNA 文库或基因组 DNA 文库来获得，并在中到高严谨条件下用

探针探测这样的文库，所述探针包含序列表中所示任何序列的全部或部分。相似的考虑适用于获得本发明的多肽和核苷酸序列的种属同源物和等位基因突变体。

[0360] 变体和菌株 / 种属同源体也可以应用简并 PCR 获得，其使用设计用来靶向变体和同源体中编码本发明序列中保守氨基酸序列的序列的引物。保守序列可以例如通过比对来自数种变体 / 同源体的氨基酸序列来预知。序列比对可以通过本领域公知的计算机软件进行。例如广泛应用的 GCGWisconsin PileUp 程序。

[0361] 用于简并 PCR 的引物可以包含一或多个简并位置，其可用于严格条件，该严格条件低于用针对已知序列的单个序列引物来克隆序列的那些严格条件。

[0362] 或者，此多核苷酸能够通过特征序列的定点诱变获得。在需要例如沉默密码子序列变化来最优化密码子对其中表达多核苷酸的具体宿主细胞的偏好时，这是有用的。其它序列的变化是为了引入限制多肽识别位点，或改变多核苷酸编码的多肽的性质或功能。

[0363] 本发明的多核苷酸（核苷酸序列）可以的探针用于制备引物例如 PCR 引物、用于替换扩增反应的引物，探针例如通过常规方法用放射性或非放射性标记从而由暴露 (revealing) 标记物来标记，或者多核苷酸可以克隆到载体中。所述引物、探针和其它片段可以为至少 15，优选至少 20，例如至少 25、30 或 40 个核苷酸的长度，也包含在本发明所用的术语多核苷酸中。

[0364] 本发明所述的多核苷酸例如 DNA 多核苷酸和探针可以重组地制备、合成地制备或用本领域技术人员能够获得的任意有效的方法制备。它们也可以通过标准技术克隆。

[0365] 一般而言，引物是用合成方法生产的，包括以一次一种核苷酸的方式所需核酸序列的分步制备方法。使用自动技术完成上述任务的技术是本领域容易获得的技术。

[0366] 较长的多核苷酸通常可以用基因重组技术制备，例如应用 PCR（多聚酶链式反应）克隆技术。这包含制备位于需要克隆的脂质靶向序列区域侧翼的引物对（例如大约 15 到 30 个氨基酸），使该引物与从动物或人细胞获得的 mRNA 或 cDNA 接触，在使所需区域扩增的条件下进行多聚酶链式反应，分离被扩增的片段（例如通过在琼脂糖凝胶中纯化反应混合物）并回收被扩增的 DNA。引物可以被设计为包含适宜的酶识别位点限制性内切酶，以至于被扩增的 DNA 能够被克隆到适宜的克隆载体。

[0367] 杂交

[0368] 本发明也包含本发明序列的互补序列，或能够或杂交到本发明序列中或其互补序列中的序列。

[0369] 本发明应用的术语“杂交”包括“通过碱基配对将核酸链连接到互补链的过程”，也就是使用多聚酶链式反应 (PCR) 技术的扩增过程。

[0370] 本发明也涉及能够与本文讨论的受试序列的互补序列或它的任何衍生物、它的片段或其衍生物杂交的核苷酸序列的用途。

[0371] 本发明也涉及与能够与本文讨论的核苷酸序列杂交的序列的互补序列。

[0372] 杂交条件基于核苷酸结合复合物的解链温度 ( $T_m$ )，如 Berger 和 Kimmel (1987, Guide to Molecular Cloning Techniques, Methods in Enzymology, Vol.152, Academic Press, San Diego CA) 中的教导，并显示如下定义的“严格性”。

[0373] 最大严格性有代表性的发生在  $T_m - 5^\circ\text{C}$ （比探针的  $T_m$  低  $5^\circ\text{C}$ ）条件下；高严格

性发生在大约  $T_m$  下  $5^\circ\text{C}$  -  $10^\circ\text{C}$  ; 中等严格性发生在大约  $T_m$  下  $10^\circ\text{C}$  -  $20^\circ\text{C}$  ; 低严格性发生在大约  $T_m$  下  $20^\circ\text{C}$  -  $25^\circ\text{C}$  。正如本领域技术人员理解的, 最大严格性杂交能够用于鉴定或检测相同的核苷酸序列, 而中等 (或低) 严格性杂交能够用于鉴定或检测类似或相关的多聚核苷酸序列。

[0374] 优选地, 本发明包含在高严格性或中等严格性条件下能够与编码具有本发明定义具体性质的多肽的核苷酸序列杂交的序列之互补序列。

[0375] 更加优选地, 本发明涉及能够在高严格条件 (例如  $65^\circ\text{C}$  和  $0.1\times\text{SSC}$  { $0.15\text{M NaCl}$ ,  $0.015\text{M}$  柠檬酸钠  $\text{pH}$  7.0}) 下与编码具有本文定义的具体性质的多肽的核苷酸序列杂交的序列之互补序列。

[0376] 本发明也涉及能够与本文讨论的核苷酸序列 (包括那些本文讨论的序列的互补序列) 杂交的序列。

[0377] 本发明也涉及能够与本文讨论的核苷酸序列 (包括那些本文讨论的序列的互补序列) 杂交的序列的互补序列。

[0378] 本发明也包括在中等到最大严格性条件下, 能够与本文讨论的核苷酸序列杂交的多聚核苷酸序列。

[0379] 在优选的方面中, 本发明覆盖了能够在严格条件下 (例如  $50^\circ\text{C}$  和  $0.2\times\text{SSC}$ ) 与本文讨论的核苷酸序列, 或它的互补序列杂交的核苷酸序列。

[0380] 在更加优选的方面中, 本发明覆盖了能够在高严格条件下 (例如  $65^\circ\text{C}$  和  $0.1\times\text{SSC}$ ) 与本文讨论的核苷酸序列, 或它的互补序列杂交的核苷酸序列。

[0381] 多肽的表达

[0382] 本发明使用的或编码具有本发明所限定具体性质的多肽的核苷酸序列能够被掺入到重组的可复制载体中。该载体可以在和 / 或从相容的宿主细胞中复制并表达核苷酸序列为多肽的形式。表达可以用控制序列, 包括启动子 / 增强子和其它表达调节信号来控制。可使用原核生物的启动子和在真核细胞中具有功能的启动子。可使用组织特异性启动子或刺激特异性启动子。也可使用嵌合的启动子, 它包含的序列元件来自两种或多种上述不同启动子。

[0383] 由宿主重组细胞产生的核苷酸序列表达的多肽可以依赖于使用的序列和 / 或载体被分泌或被包含在细胞内。编码序列可以设计成与信号序列一起, 该信号序列可以引导所述物质的编码序列穿过具体的原核生物或真核生物细胞膜。

[0384] 表达载体

[0385] 术语 “表达载体” 意思是一种能够在体内或体外表达的构建体。

[0386] 优选地, 表达载体是掺入到生物体基因组中的。术语 “掺入” 优选包括稳定地掺入到基因中。

[0387] 本发明的核苷酸序列或编码具有本发明所限定具体性质的多肽的核苷酸序列可以存在于载体中, 其中所述核苷酸序列可操作地连接到调节序列, 以至于该调节序列能够通过适宜的宿主生物体表达核苷酸序列, 即该载体是表达载体。

[0388] 本发明的载体可以被转化到合适的下面描述的宿主细胞来表达具有本发明所限定的具体性质的多肽。

[0389] 载体的选择, 例如质粒、粘端质粒、病毒或噬菌体载体, 常常依赖于它被引入

的宿主细胞。

[0390] 载体可以包含一种或多种可选标记基因 - 诸如提供抗生素抗性的基因, 如氨苄青霉素、卡那霉素、氯霉素或四环素抗性。或者, 该选择可以通过共转化完成 (如 WO91/17243 中描述的)。

[0391] 载体可被用于体外, 例如产生 RNA 或用于转染或转化宿主细胞。

[0392] 因此, 在一个进一步的具体实施方案中, 本发明提供了一种制备本发明的核苷酸序列或编码具有本发明所限定具体性质的多肽的核苷酸序列的方法, 所述方法是通过引入核苷酸序列到可复制载体中, 引入该载体到适合的宿主细胞中, 并在能够引起所述载体复制的条件下使所述宿主细胞生长。

[0393] 载体可以进一步包含在正讨论的宿主细胞中促进载体复制的核苷酸序列。这些序列的例子是质粒 pUC19、pACYC177、pUB110、pE194、pAMB1 和 pIJ702 的复制起点。

[0394] 调节序列

[0395] 在一些应用中, 本发明使用的核苷酸序列或编码具有本发明所限定具体性质的多肽的核苷酸序列可以可操作地连接到能 (如通过被选宿主细胞) 表达所述核苷酸序列的调节序列上。例如, 本发明包括这样的载体, 其包含可操作连接到这样的调节序列上的本发明核苷酸序列, 即该载体是表达载体。

[0396] 术语“可操作连接”(“operably linked”)指并列(juxtaposition), 其中描述的组分为使它们以意图方式作用的关系。“可操作连接”到编码序列的调节序列以通过在与控制序列相容的条件下完成表达编码序列的方式来连接。

[0397] 术语“调节序列”包括启动子和增强子和其它表达调节信号。

[0398] 这里使用的术语“启动子”是本领域内可常规使用的, 例如 RNA 聚合酶结合位点。

[0399] 编码具有本发明限定的具体性质的酶的核苷酸序列的增强表达也可以通过选择异源的调节序列来完成, 例如启动子、分泌前导序列(secretion leader)和终止区。

[0400] 优选地, 本发明的核苷酸序列可以与至少一个启动子可操作连接。

[0401] 指导细菌、真菌或酵母宿主内核苷酸序列转录的适宜启动子的例子是本领域公知的。

[0402] 构建体

[0403] 术语“构建体”与术语如“偶联物(conjugate)”、“盒(cassette)”和“杂合体(hybrid)”同义, 它包括编码具有本发明所限定具体性质的多肽的核苷酸序列, 其用于根据本发明直接或间接与启动子连接。间接连接的例子是在启动子和本发明核苷酸序列中间提供适宜的间隔组(spacer group), 例如内含子序列, 如 Sh1 内含子或 ADH 内含子。同样, 本发明的术语“融合”包括直接或间接连接。在一些方面, 这些术语不包括编码通常与野生型基因启动子连接的蛋白质的核苷酸序列的天然组合并且条件是所述它们都处于它们的自然环境中。

[0404] 所述构建体甚至可以包含或表达允许选择遗传构建体的标记物。

[0405] 对于一些应用中, 优选所述构建体包含至少本发明的核苷酸序列或编码具有本发明限定的具体性质的多肽的核苷酸, 其可操作连到启动子。

**[0406] 宿主细胞**

[0407] 本发明的术语“宿主细胞”包括任何具有以下性质的细胞：其包含编码具有本发明所限定具体性质的多肽的核苷酸序列，或者上面描述的用于重组制备具有本发明所限定具体性质的多肽的表达载体。

[0408] 因此，本发明进一步的具体实施方案提供了用本发明核苷酸序列或编码具有本发明所限定具体性质的多肽的核苷酸序列转化或转染的宿主细胞。可选择与上述载体相容的所述细胞，可以例如是原核生物（如细菌）、真菌、酵母或植物的细胞。优选的宿主细胞不是人的细胞。

[0409] 适合的细菌宿主生物体的例子是革兰阴性菌或革兰氏阳性菌。

[0410] 依赖于编码具有本发明所限定具体性质的多肽的核苷酸序列的性质，和/或进一步加工所表达蛋白质的需求，真核生物的宿主例如酵母或其它真菌是优选的。一般而言，酵母细胞是相对真菌细胞优选的，因为酵母细胞较容易操纵。然而，一些蛋白质或者从酵母细胞中分泌很少，或者有时没有被恰当加工（例如酵母中的高糖基化）。在这些例子中，应选择不同的真菌宿主生物体。

[0411] 适宜宿主细胞，如酵母、真菌和植物宿主细胞的用途，可以提供翻译后修饰（例如豆蔻酰化、糖基化、截短、投石 (lipidation) 和酪氨酸、丝氨酸或苏氨酸磷酸化），因为可能需要使本发明的重组表达产物具有最佳生物活性。

[0412] 所述宿主细胞可以是蛋白酶缺陷的或蛋白酶阴性的 (minus) 菌株。

**[0413] 生物体**

[0414] 本发明的术语“生物体”包括任何包含本发明所述的核苷酸序列，或包含用于编码具有本发明所限定具体性质的多肽的核苷酸序列和/或由此获得的产物的生物体。

[0415] 适宜的生物体可包括原核生物、真菌，酵母或植物。

[0416] 本发明的术语“转基因生物体”包括任何生物体，所述生物体包含用于编码具有本发明所限定具体性质的多肽的核苷酸序列和/或由此获得的产品的，和/或其中的启动子能够允许在生物体中表达用于编码具有本发明所限定具体性质的多肽的核苷酸序列。优选的核苷酸序列被掺入在生物体基因组中。

[0417] 术语“转基因生物体”不包括在天然环境中的天然核苷酸编码序列，此时所述序列受同样处于天然环境中的它们的天然启动子控制。

[0418] 因此，本发明转基因生物体包括含有选自如下物质之一，或其组合的生物体：用于编码具有本发明所限定的具体性质的多肽的核苷酸序列，本文定义的构建体，本文限定的载体，本文限定的质粒，本文限定的细胞，或它的产物。例如，转基因生物体也包含在异源的启动子控制下用于编码具有本发明所限定的具体性质的多肽的核苷酸序列。

**[0419] 宿主细胞 / 宿主生物体转化**

[0420] 正如先前指出的，宿主生物体可以是原核生物体或真核生物体。适合的原核生物宿主的例子包括大肠杆菌和枯草杆菌。

[0421] 原核生物宿主转化是本领域公知的，例如参见 Sambrook 等 (Molecular Cloning : A Laboratory Manual, 2nd edition, 1989, Cold Spring Harbor Laboratory Press)。如果应用原核生物宿主，那么核苷酸序列在转化之前需要被适当修饰 - 例如去除内含子。

[0422] 另一个具体实施方案中，转基因生物体可以是酵母。

[0423] 丝状真菌细胞可以使用本领域公知的各种方法转化，例如涉及原生质体形成和原生质体转化，之后用公知的方式再生细胞壁的方法。曲霉作为宿主微生物的用途已经被 EP 0 238 023 公开。

[0424] 另一种宿主生物体可以是植物。用于转化植物的一般技术的综述公开在 Potrykus 的文章中 (Annu Rev Plant Physiol Plant Mol Biol[1991]42 : 205-225) 和 Christou 的文章中 (Agro-Food-Industry Hi-Tech March/April 1994:17-27)。进一步有关于植物转化方面的教导见于 EP-A-0449375。

[0425] 对于真菌、酵母和植物转化的广泛教导见下述部分。

[0426] 转化的真菌

[0427] 宿主生物体可以是真菌 - 例如丝状真菌。此类适合的宿主的例子包括属于嗜热酶属、枝顶孢霉属 (Acremonium)、曲霉属、青霉属、毛霉属 (Mucor)、脉孢菌属 (Neurospora)、木霉属 (Trichoderma) 等菌属中的任意一种。

[0428] 转化丝状真菌的教导可以参见 US-A-5741665，它综述了转化丝状真菌和培养真菌本领域公知的标准技术。一个适用于粗糙脉孢菌 (N.crassa) 的范围更广的综述，例如见于 Davis 和 de Serres 的 Methods Enzymes (1971) 17A : 79-143 中。

[0429] 进一步转化丝状真菌的教导被公开在 US-A-5674707 中。

[0430] 一方面，宿主生物体可以是曲霉属，例如黑曲霉。

[0431] 本发明的转基因曲霉也可用以下方法制得，例如 Turner C.1994 (Vectors for genetic manipulation. In : Martinelli S.D., Kinghorn J.R. (编者) Aspergillus : 50 years on. Progress in industrial microbiology vol 29. Elsevier Amsterdam 1994. pp.641-666) 教导的方法。

[0432] 丝状真菌中的基因表达已经被公开在 Punt 等 (2002) Trends Biotechnol 2002 May ; 20 (5) : 200-6, Archer & Peberdy Crit Rev Biotechnol (1997) 17 (4) : 273-306 中。

[0433] 转化的酵母

[0434] 在另一个具体实施方案中，转基因生物体可以是酵母。

[0435] 有关酵母中异源基因表达的的原理的综述见于，例如 Methods Mol Biol (1995), 49 : 341-54, 和 Curr Op 中 Biotechnol (1997) Oct ; 8 (5) : 554-60。

[0436] 从这个角度，酵母 - 例如酿酒酵母或毕赤酵母 (Saccharomyces cerevisi or Pichia pastoris) (参见 FEMS Microbiol Rev (2000) 24 (1) : 45-66) 可以被用作异源基因表达的载体。

[0437] 有关酿酒酵母异源基因表达和基因产物的分泌原理的综述见 E Hinchcliffe E Kenny (1993, " Yeast as a vehicle for the expression of heterologous genes " , Yeasts, Vol 5, Anthony H Rose and J Stuart Harrison, eds, 2nd edition, Academic Press Ltd.)。

[0438] 对于酵母的转化，已开发了几个试验方案。例如，本发明的转基因酵母可以通过以下 Hinnen 等 (1978, Proceedings of the National Academy of Sciences of the USA 75, 1929) ; Beggs, J D (1978, Nature, London, 275, 104) ; and Ito, H et al (1983, J Bacteriology 153, 163-168) 教导的方法制备。

[0439] 适宜的酵母宿主生物体可以选自生物工程相关酵母种类，例如，但不限于，

选自毕赤酵母属、汉逊酵母属 (*Hansenula*)、克鲁维酵母属、*Yarrowinia* spp.、酵母属 (*Saccharomyces* spp.) 包括酿酒酵母、或裂殖酵母 (*Schizosaccharomyce* spp.) 诸如粟酒裂殖酵母 (*Schizosaccharomyce pombe*)。

[0440] 甲基营养 (methylophilic) 酵母种的菌株巴氏毕赤酵母 (*Pichia pastoris*) 可以用作宿主生物体。

[0441] 在一具体实施方案中, 宿主生物体可以是汉逊酵母属, 例如多形汉逊酵母 (*H.polymorpha*) (如在 WO01/39544 中公开的)。

[0442] 转化的酵母细胞可以通过使用不同的选择性标记例如营养缺陷型标记显性的抗菌素抗性标记来选择。

[0443] 转化的植物 / 植物细胞

[0444] 本发明适宜的宿主生物体可以是植物。有关一般技术的综述见于 Potrykus 的文章 (*Annu Rev Plant Physiol Plant Mol Biol*[1991]42 : 205-225) 和 Christou 的文章 (*Agro-Food-Industry Hi-Tech March/April 1994 17-27*), 或在 WO01/16308 中。例如, 转基因植物可产生升高水平的植物固醇酯以及 phytosterol 酯。

[0445] 因此, 本发明也涉及生产具有提高的植物甾醇酯和 phytosterol 酯水平的转基因植物的方法, 包含用本发明定义的脂质酰基转移酶转化植物细胞 (具体是用包含本文定义的脂质酰基转移酶的表达载体或结构体) 和从转化的细胞培育植物的步骤。

[0446] 分泌

[0447] 通常, 使多肽从表型宿主分泌到培养基中是希望得到的, 而酶能够很容易从所述培养基中回收。分泌前导序列可基于所需的表达宿主进行选择。杂交信号序列也在本发明中得到应用。

[0448] 异源分泌前导序列的常见例子是那些源于真菌淀粉葡苷酶 (AG) 基因 (*glaA*- 具有 18 和 24 氨基酸, 例如来自曲霉),  $\alpha$ - 因子基因 (酵母例如酵母属、克鲁维酵母属和汉逊酵母属) 或  $\alpha$ - 淀粉酶基因 (芽孢杆菌属) 的序列。

[0449] 检测

[0450] 本领域有多种检测和测定氨基酸序列表达的公知方案。例如包括酶联免疫吸附剂测定 (ELISA)、放射免疫测定法 (RIA) 和荧光活化的细胞分选法 (FACS)。

[0451] 本领域技术人员掌握已知多种标记和连接技术, 并可以被应用在核酸和氨基酸测定方面中。

[0452] 许多公司, 例如 Pharmacia Biotech (Piscataway, NJ)、Promega (Madison, WI)、和 US Biochemical Corp (Cleveland, OH) 为这些方法提供商业试剂盒和实验方案。

[0453] 适宜的报告分子或标记包括那些放射性核素、酶、荧光素、化学发光物 (chemiluminescent)、显色剂, 也包括底物、辅助因子、抑制因子、磁颗粒等等类似的物质。专利包括 US-A-3,817,837、US-A-3,850,752、US-A-3,939,350、US-A-3,996,345、US-A-4,277,437、US-A-4,275,149 和 US-A-4,366,241 已经教导了这些标记的用途。

[0454] 重组免疫球蛋白也可以如 US-A-4,816,567 中所述生产。

[0455] 融合蛋白

[0456] 具有本发明所述具体性质的多肽可以被当作融合蛋白生产, 例如有助于它的提取和纯化。融合蛋白配偶体 (partner) 的例子包括谷胱甘肽-S- 转移酶 (GST)、6xHis、

GAL4(DNA 结合区域和 / 或转录活化结构域) 和  $\beta$ -半乳糖苷酶。在融合蛋白配偶体和目的蛋白序列之间包含蛋白水解切割位点以便允许去除融合蛋白序列的蛋白质序列也是方便的。优选的融合蛋白不妨碍所述蛋白质序列的活性。

[0457] 大肠杆菌中的基因融合表达系统已经公开在 *Curr.Opin.Biotechnol.* (1995) 6(5) : 501-6 中。

[0458] 本发明的另一个具体实施方案中, 具有本发明所述具体性质的多肽的氨基酸序列可以连接到异源序列以便编码融合蛋白。例如, 为了筛选多肽文库寻找能够影响物质活性的因子, 编码表达可识别可购得的抗体的异源表位的嵌合物质是有用的。

#### 附图说明

[0459] 本发明下面的部分将仅通过实施例并参照下图与实施例进行详细描述。

[0460] 图 1 显示来自第 6 版数据库的 pfam00657 共有序列 (SEQ ID No.1) ;

[0461] 图 2 显示从生物体嗜水气单胞菌中得到的氨基酸序列 (SEQ ID No.2) (P 10480 ; GI : 121051) ;

[0462] 图 3 显示从生物体杀鲑气单胞菌中得到的氨基酸序列 (SEQ ID No.3) (AAG098404 ; GI : 9964017) ;

[0463] 图 4 显示从生物体天蓝色链霉菌 A3(2) 中得到的氨基酸序列 (SEQ ID No.4) (Genebank 登录号 NP631558) ;

[0464] 图 5 显示从生物体天蓝色链霉菌 A3(2) 中得到的氨基酸序列 (SEQ ID No.5) (Genebank 登录号 CAC42140) ;

[0465] 图 6 显示从生物体酿酒酵母中得到的氨基酸序列 (SEQ ID No.6) (Genebank 登录号 P41734) ;

[0466] 图 7 显示所选序列与 pfam00657 共有序列的比对。

[0467] 图 8 显示 SEQ ID No.3 与 SEQ ID No.2 配对比对, 其显示 93% 的氨基酸序列同一性。信号序列加下划线。+ 表示有差别。GDSX 基序包括活性位点丝氨酸 16, 并且活性位点天冬氨酸 116 和组氨酸 291 突出显示 (见于阴影区域)。氨基酸后的数字是负信号序列 (minus the signal sequence) ;

[0468] 图 9 显示编码从生物体嗜水气单胞菌获得的本发明脂质酰基转移酶的核苷酸序列 (SEQ ID No.7) ;

[0469] 图 10 显示编码从生物体杀鲑气单胞菌获得的本发明脂质酰基转移酶的核苷酸序列 (SEQ ID No.8) ;

[0470] 图 11 显示编码从生物体天蓝色链霉菌 A3(2) 获得的本发明脂质酰基转移酶的核苷酸序列 (SEQ ID No.9) (Genebank 登录号 Nec003888.1 : 8327480..8328367) ;

[0471] 图 12 显示编码从生物体天蓝色链霉菌 A3(2) 获得的本发明脂质酰基转移酶的核苷酸序列 (SEQ ID No.10) (Genebank 登录号 AL939131.1 : 265480..266367) ;

[0472] 图 13 显示编码从生物体酿酒酵母获得的本发明脂质酰基转移酶的核苷酸序列 (SEQ ID No.11) (Genebank 登录号 Z75034) ;

[0473] 图 14 显示从生物体雷尔氏菌 (*Ralstonia*) 获得的氨基酸序列 (SEQ ID No.12) (Genebank 登录号 AL646052)。

[0474] 图 15 显示编码从生物体雷尔氏菌获得的本发明脂质酰基转移酶的核苷酸序列 (SEQ ID No.13) ;

[0475] 图 16 显示 SEQ ID No.20。所述序列编码 Scoe1, 其为 NCBI 蛋白质登录号 CAB39707.1 GI : 4539178 保守的公知 (hypothetical) 蛋白质 [天蓝色链霉菌 A3(2)];

[0476] 图 17 显示如 SEQ ID No.21 所示的核苷酸序列, 其编码 NCBI 蛋白质登录号 CAB39707.1 GI : 4539178 保守的公知蛋白质 [天蓝色链霉菌 A3(2)];

[0477] 图 18 显示 SEQ ID No.22 所示氨基酸序列。所述序列编码 Scoe2, 其为 NCBI 蛋白质登录号 CAC01477.1 GI : 9716139 的保守公知蛋白质 [天蓝色链霉菌 A3(2)];

[0478] 图 19 显示如 SEQ ID No.23 所示核苷酸序列, 该核苷酸序列编码 Scoe2, 其为 NCBI 蛋白质登录号 CAC01477.1 GI : 9716139 的保守公知蛋白质 [天蓝色链霉菌 A3(2)];

[0479] 图 20 显示 Scoe3 的氨基酸序列 (SEQ ID No.24), 所述 Scoe3 为 NCBI 蛋白质登录号 CAB88833.1 GI : 7635996 的公知分泌性蛋白 [天蓝色链霉菌 A3(2)];

[0480] 图 21 显示如 SEQ ID No.25 所示核苷酸序列, 该核苷酸序列编码 Scoe3, 其为 NCBI 蛋白登录号为 CAB88833.1 GI : 7635996 的公知分泌性蛋白 [天蓝色链霉菌 A3(2)];

[0481] 图 22 显示 Scoe4 的氨基酸序列 (SEQ ID No.26), 所述 Scoe4 是 NCBI 蛋白登录号为 CAB89450.1 GI : 7672261 的公知分泌性蛋白 [天蓝色链霉菌 A3(2)];

[0482] 图 23 显示如 SEQ ID No.27 所示核苷酸序列, 该核苷酸序列编码 Scoe4, 其为 NCBI 蛋白质登录号 CAB89450.1 GI : 7672261 的公知分泌性蛋白 [天蓝色链霉菌 A3(2)];

[0483] 图 24 显示 Scoe5 的氨基酸序列 (SEQ ID No.28), Scoe5 为 NCBI 蛋白质登录号 CAB62724.1 GI : 6562793 的公知脂蛋白 [天蓝色链霉菌 A3(2)];

[0484] 图 25 显示如 SEQ ID No.29 所示的核苷酸序列, 其编码 Scoe5, Scoe5 为 NCBI 蛋白质登录号 CAB62724.1 GI : 6562793 的公知脂蛋白 [天蓝色链霉菌 A3(2)];

[0485] 图 26 显示 Srim1 的氨基酸序列 (SEQ ID No.30), Srim1 为 NCBI 蛋白质登录号为 AAK84028.1 GI : 15082088 的 GDSL-脂酶 [龟裂链霉菌 (*streptomyces rimosus*)];

[0486] 图 27 显示 SEQ ID No.31 所示的核苷酸序列, 其编码 Srim1, Srim1 为 NCBI 蛋白登录号为 AAK84028.1 GI : 15082088 的 GDSL-脂肪酶 [龟裂链霉菌];

[0487] 图 28 显示从嗜水气单胞菌 (ATCC#7965) 获得的脂质酰基转移酶的氨基酸序列 (SEQ ID No.32) ;

[0488] 图 29 显示编码从嗜水气单胞菌 (ATCC#7965) 获得的脂质酰基转移酶的核苷酸序列 (SEQ ID No.33) ;

[0489] 图 30 显示从杀鲑气单胞菌杀鲑亚种 (*Aeromonas salmonicida* subsp.*Salmonicida*) (ATCC#14174) 获得的脂质酰基转移酶的氨基酸序列 (SEQ ID No.34) ;

[0490] 图 31 显示编码从杀鲑气单胞菌杀鲑亚种 (ATCC#14174) 获得的脂质酰基转移酶的核苷酸序列 (SEQ ID No.35) ;

[0491] 图 32 显示气单胞菌属基因的同源物可使用国家生物技术信息中心 (the National Center for Biotechnology Information), NIH, MD, USA 的基础局部序列比对检索工具

服务 (basic local alignment search tool service) 和完整基因组数据库鉴定。用 GDSX 基序作数据库搜索, 而且鉴定了潜在编码具有脂解活性的酶的许多序列 / 基因。已从链霉菌属, 黄单胞菌属和雷尔氏菌属 (*Ralstonia*) 中鉴定了基因。如下面的例子, 青枯雷尔氏菌 (*Ralstonia solanacearum*) 与杀鲑气单胞菌 (*satA*) 基因作比对。成对比对表现出 23% 的同一性。活性位点丝氨酸出现在氨基端, 且可以鉴定催化性残基组氨酸和天冬氨酸。

[0492] 图 33 显示 Pfam00657.11[ 家族 00657, 数据库版本 11] 共有序列 (此后称作 Pfam 共有序列) 以及不同序列与 Pfam 共有序列的比对。箭头表示活性位点残基, 加下划线的框 (boxes) 表示三种已由 [Upton C and Buckley JT (1995) Trends Biochem Sci 20 ; 179-179] 标明的同源框。Pfam 共有序列中的大写字母表示在许多家族成员中保守的残基。“-” 标记表示预期在 Pfam 共有序列的隐藏 Markov 模型中发现残基而没有发现残基, 因而有缺口被插入的位点。“.” 标记表示在 Pfam 共有序列中没有对应残基的残基。这些序列是在图 16、18、20、22、24、26、28 和 30 中列出的氨基酸序列。

[0493] 图 34 显示 Pfam00657.11[ 家族 00657, 数据库版本 11] 共有序列 (此后称作 Pfam 序列), 以及不同序列与 Pfam 共有序列的比对。箭头表示活性位点残基, 加下划线的框表示三种已由 [Upton C and Buckley JT (1995) Trends Biochem Sci 20 ; 179-179] 标明的同源框。Pfam 共有序列中的大写字母表示在许多家族成员中保守的残基。“-” 标记表示预期在 Pfam 共有序列的隐藏 Markov 模型中发现残基而没有发现残基, 因而有缺口被插入的位点。“.” 标记表示在 Pfam 共有序列中没有对应残基的残基。这些序列是在图 2、16、18、20、26、28 和 30 中列出的氨基酸序列。这些蛋白质都对脂质底物具有活性。

[0494] 图 35 显示包含羧基末端组氨酸标记的杀鲑气单胞菌脂质酰基转移酶基因的表达载体 pet12-AsalGCAT-pSM ;

[0495] 图 36 显示用 NEFA 试剂盒测定法检测细胞提取物的结果, 结果描述了重组杀鲑气单胞菌脂质酰基转移酶对卵磷脂的活性。从左至右的孔表示的是: 阳性对照, 阴性对照 (即空质粒提取物) 和 IPTG 诱导后保温 0、1、2 和 3 小时收集的样本。

[0496] 图 37 显示含表达载体 pet12-AsalGCAT-pSM 的 BL21 (DE3)pLysS 的生长优化, 其显示在 30°C 培养生成了对卵磷脂有高活性的酶。用 NEFA 试剂盒测定法来检测细胞提取物的磷脂酶活性。从左至右的孔表示的是: 阳性对照; 阴性对照; 20°C; 30°C ;

[0497] 图 38 显示与底物卵磷脂一同保温的、表达活性脂质酰基转移酶的 BL21 (DE3) pLysS 的细胞粗提取物, 反应混合物通过薄层色谱法分析, 显示降解产物的存在。泳道: 1. 无酶; 2.+A.sal-10  $\mu$  l 37°C; 3.+A.sal-20  $\mu$  l 37°C; 4.+A.sal-10  $\mu$  l 24°C; 5.+A.sal-20  $\mu$  l 24°C ;

[0498] 图 39 所示为杀鲑气单胞菌酰基转移酶的部分纯化, 其显示了与纯化的组氨酸标记蛋白相关的磷脂酶活性。SE = 超声提取物, His = 用 Qiagen 的 Ni-NTA 离心 - 试剂盒 (spin-kit) 纯化 ;

[0499] 图 40 显示包含 C- 末端组氨酸标记的嗜水气单胞菌甘油酯酰基转移酶 (Glycerolipid Acyl Transferase) (GCAT) 基因的表达载体 pet12-A.h.GCAT = pSMa, 被用于转化大肠杆菌 BL21 (DE3)pLysS ;

[0500] 图 41 显示使用 Non-Esterified Fatty Acid (NEFA) 试剂盒 (Roche, 瑞士), 测定含有重组嗜水气单胞菌 GCAT 酶的粗提物 (5 & 10  $\mu$  l) 对卵磷脂的活性, 显示存在对磷脂,

卵磷脂有活性的酶；

[0501] 图 42 显示含有表达载体 pet12-AsalGCAT-pSM 的 BL21(DE3)pLysS 的生长优化，显示在 30°C 培养生成了对卵磷脂有高活性的酶。用 NEFA 试剂盒测定法来检测细胞提取物的磷脂酶活性。

[0502] 图 43 显示嗜水气单胞菌和杀鲑气单胞菌的酰基转移酶的部分纯化，其显示了与纯化的组氨酸标记蛋白相关的磷脂酶活性。SE = 超声提取物，His = 用 Qiagen 的 Ni-NTA 自旋 - 试剂盒纯化；

[0503] 图 44 显示了气单胞菌属基因在枯草芽孢杆菌 163 中的表达，其显示产生了对卵磷脂和 DGDG 均具有活性的分泌的酶。pUB-AH 为包含嗜水气单胞菌基因的构建体，pUB-AS 为带有杀鲑气单胞菌基因的构建体，培养物滤出液与底物保温 60 分钟。

[0504] 图 45 与图 46 图示了脂肪酸和胆固醇酯作为时间的函数。所述图显示在利用卵磷脂和胆固醇缓冲液作为底物的条件下，食品中酰基转移酶活性测定实验的 GLC 分析所得结果；

[0505] 图 47 显示在例 17 中用于诱变嗜水气单胞菌脂质酰基转移酶基因的融合构建体的氨基酸序列 (SEQ ID No.36)，加下划线的氨基酸是木聚糖酶信号肽；

[0506] 图 48 描述了编码包含木聚糖酶信号肽的编码嗜水气单胞菌酶的核苷酸序列 (SEQ ID No.54)；

[0507] 图 49 显示氨基酸的蛋白质 - 脂肪酸缩合物 (condensate) 的结构；

[0508] 图 50 图示了来自磷脂酰胆碱的脂肪酸在转移到具有可用于酯化的游离羟基的氨基酸例如酪氨酸或丝氨酸的游离羟基时所发生的反应；并且

[0509] 图 51 显示 DGDG 与葡萄糖在脂质酰基转移酶的催化下的反应图示。

## 具体实施方式

[0510] 实施例 1：来源于杀鲑气单胞菌杀鲑亚种 (*Aeromonas salmonicida* subsp. *Salmonicida*) 的转移酶的克隆，序列测定和异源性表达。

[0511] 所用的菌株：

[0512] 杀鲑气单胞菌杀鲑亚种 (ATCC 14174) 获自 ATCC，在 30°C、Luria-Bertani 培养基 (LB) 中保温过夜。对所述细胞进行离心并用 Qiagen Ltd. 的基因组 DNA 分离方法分离基因组 DNA。基因组 DNA 缓冲液组 (cat.19060)，蛋白激酶 K (cat.19131) 和 RNase A (cat.19101) 均获自于 Qiagen.Ltd (Boundary court Gatwick Court, West Sussex, RH 102AX)。

[0513] 宿主菌株 BL21(DE3)pLysS (Novagen) 用于产生重组气单胞菌属酶。BL21(DE3)pLysS 的感受态细胞作为用表达载体 pet12-AsalGCAT-pSM 转化的宿主。包含适宜质粒的转化体在含有 100- $\mu$ g 氨苄西林每毫升的 LB 琼脂培养基中、37°C 生长。

[0514] 表达载体 pet12-AsalGCAT-pSM 的构建：

[0515] 对于来自气单胞菌属的转移酶基因的所有 DNA 扩增，以基因组 DNA (0.2-1  $\mu$ l) 作为模板，pfu DNA 聚合酶 (2.5 单位) 与 10  $\mu$ l 10xpfu 缓冲液，均为 1  $\mu$ l 的每一种引物 (50pmol/ $\mu$ l)，200  $\mu$ M dNTP 在 100  $\mu$ l 的总反应容积中一起使用。PCR 反应在可编程的热循环仪中利用如下条件进行：95°C 维持 30 秒，95°C 30 秒、60°C 1 分钟、68°C 2 分钟循

环 30 次。72℃再延伸 5 分钟。

[0516] 来自杀鲑气单胞菌的转移酶基因的 PCR 扩增通过两个分离的 PCR 反应进行。PCR 反应 1 使用以下引物对实施：as1USNEW(5' AGCATATGAAAAAATGGTTTGT TGT TTTATTG GGG 3' [SEQ ID No.36]) 和 asls950new(5' GTG ATG GTG GGC GAG GAA CTC GTA CTG3' [SEQ ID No.37])。实施第二次 PCR 反应以掺入 C-末端组氨酸标记，所述反应使用第一次反应的 PCR 产物以及以下引物进行：as1USNEW(5' AGCATATGAAAA AATGGTTTGT TTTATTG GGG 3' [SEQ ID No.38]) 和 AHLS1001(5' TTGGATCCGAATTCAT CAATG GTG ATG GTG ATG GTG GGC3' [SEQ ID No.39])。纯化第二次反应的 PCR 产物并利用限制性内切酶 NdeI 和 BamHI 消化。2 μ gpET 12a 载体 DNA 同样经限制性内切酶 NdeI 和 BamHI 消化并用磷酸酶处理。限制性内切酶处理的 pET 12a 与反应 2 的 PCR 产物经纯化，用快速连接试剂盒 (Rapid Ligation Kit) (Roche, Switzerland) 连接。连接混合物用来转化大肠杆菌 TOP10 细胞。转化体铺在含有 100 μ g/ml 氨苄西林的 LB 琼脂培养基上。

[0517] T7 启动子引物 (5' TAATACGACTCACTATAG3' [SEQ ID No.40]) 和 T7 终止子引物 (5' CTAGTTATTGCTCAGCGG3' [SEQ ID No.41]) 用于验证 pET12a 载体中所克隆的转移酶基因的序列和方向。用 ABI Prism® BigDye™ 终止子循环测序试剂盒 (Terminators Cycle sequencing kit)，以 500ng 质粒 DNA 作为模板，以及 3.2pmol T7 启动子和终止子引物实施 DNA 序列测定。

[0518] 图 35 所示构建体用于转化感受态细菌宿主菌株 BL21 (DE3)pLysS (Novagen) 并且可选出氨苄西林抗性转化体用于表达分析。

[0519] 重组杀鲑气单胞菌脂质酰基转移酶的表达

[0520] 用非酯化的脂肪酸 (Non-Esterified Fatty Acid) (NEFA) 试剂盒 (Roche, Switzerland) 在细胞提取物中定量对于卵磷脂的酶活性。

[0521] 在图 36 中，包含表达载体 pet12-AsalGCAT-pSM 的 BL21 (DE3)pLysS 在 LB 培养基 +100 μ g/ml 氨苄西林中生长，37℃振荡保温，直到 OD<sub>600</sub> = 0.6 到 1.0。用 IPTG (0.4mM) 诱导所述培养物，再保温 3 小时。IPTG 诱导 0、1、2 和 3 小时后取样品。以卵磷脂作为底物用 NEFA 试剂盒检测酶活性。

[0522] 用于制备活性更强的酶的生长优化

[0523] 包含表达载体 pet12-AsalGCAT-pSM 的 BL21 (DE3)pLysS 在 LB 培养基 +100 μ g/ml 氨苄西林中生长，在不同生长温度 (37℃、30℃、20℃)、振荡条件下保温。产生活性脂质酰基转移酶的最适条件是当如图 37 所示培养物在 30℃生长时。

[0524] 重组杀鲑气单胞菌转移酶的部分纯化

[0525] 包含表达载体 pet12-AsalGCAT-pSM 的菌株 BL21 (DE3)pLysS 在 37℃生长，通过超声处理 (sonification) 制备细胞粗提物。重组体酶用 Qiagen 的 Ni-NTA 自旋 (spin) 试剂盒从经超声处理的细胞粗提物进一步纯化。使用 NEFA 试剂盒且以卵磷脂作为底物测定磷脂酶活性。来自表达活性转移酶的 BL21 (DE3)pLysS 的细胞粗提物与底物卵磷脂一同保温，用薄层色谱法分析反应混合物，显示降解产物的存在 (参见图 38)。

[0526] 重组杀鲑气单胞菌转移酶的部分纯化

[0527] 包含表达载体 pet12-AsalGCAT-pSM 的菌株 BL21 (DE3)pLysS 在 37℃生长，通

过超声处理制备细胞粗提物。重组酶用 Qiagen 的 Ni-NTA 自旋试剂盒从经超声处理后的细胞粗提物进一步纯化。使用 NEFA 试剂盒且以卵磷脂作为底物测定磷脂酶活性（参见图 39）。

#### [0528] 实施例 2 嗜水气单胞菌转移酶在大肠杆菌中的克隆与表达

[0529] 从 ATCC 获得的嗜水气单胞菌 (ATCC#7965)，30 °C、Luria-Bertani 培养基 (LB) 中保温过夜。离心细胞并用 Qiagen Ltd. 的基因组 DNA 分离方法分离基因组 DNA (genomic DNA)。基因组 DNA 缓冲液组 (cat.19060)，蛋白激酶 K (cat.19131) 和 RNase A (cat.19101) 均来自于 Qiagen Ltd. (Boundary court Gatwick Court, West Sussex, RH10 2AX)。

[0530] 宿主细菌菌株 BL21 (DE3) pLysS (Novagen) 用于制备重组气单胞菌属酶。BL21 (DE3) pLysS 的感受态细胞作为用表达载体 pet12a-A.h.GCAT-pSMa 转化的宿主。包含适宜质粒的转化体在含有 100- $\mu$ g 氨苄西林每毫升的 LB 琼脂培养基中、于 37°C 生长。

[0531] 表达载体 pet12a-A.h.GCAT-pSMa 的构建：

[0532] 对于来自气单胞菌属的转移酶基因的所有 DNA 扩增，以基因组 DNA (0.2-1  $\mu$ l) 作为模板，pfu DNA 聚合酶 (2.5 单位) 与 10  $\mu$ l 10xpfu 缓冲液，均为 1  $\mu$ l 的每一种引物 (50pmol/ $\mu$ l)，200  $\mu$ M dNTP 在 100  $\mu$ l 的总反应容积中一起使用。PCR 反应在可编程的热循环仪中利用如下条件进行：95°C 维持 30 秒，95°C 30 秒、60°C 1 分钟、68°C 2 分钟循环 30 次。72°C 再延伸 5 分钟。

[0533] 来自嗜水气单胞菌 (ATCC#7965) 的转移酶基因的 PCR 扩增通过两个分离的 PCR 反应进行。

[0534] PCR 反应 1 使用以下引物对进行：AHUS1 (5' GTCATATGAAAAAATGGTTTGTGTGTTTATTGGGATTGGTC3'，SEQ ID No.42) 和 ahls950 (5' ATGGTGATGGTGGGCGAGGAACCTCGTACTG3'，SEQ ID No.43)。

[0535] 实施第二次 PCR 反应以掺入 C- 末端组氨酸标记，所述反应使用第一次反应的 PCR 产物以及以下引物进行：AHUS1 (5' GTCATATGAAAAAATGGTTTGTGTGTTTATTGGGATTGGTC3' SEQ ID No.44,) 和 AHLS1001 (5' TTGGATCCGAATTCATCAATGGTGATGGTGATGGTGGGC3' SEQ ID No.45)。

[0536] 纯化第二次反应的 PCR 产物并利用限制性内切酶 NdeI 和 BamHI 消化。2  $\mu$ g pET 12a 载体 DNA 同样经限制性内切酶 NdeI 和 BamHI 消化并用磷酸酶处理。限制性内切酶处理的 pET 12a 与反应 2 的 PCR 产物经纯化，用快速连接试剂盒 (Rapid Ligation Kit) (Roche, Switzerland) 连接。连接混合物用来转化大肠杆菌 TOP10 细胞。转化体铺在含有 100  $\mu$ g/ml 氨苄西林的 LB 琼脂培养基上。

[0537] T7 启动子引物 (5' TAATACGACTCACTATAG3' ) 和 T7 终止子引物 (5' CTAGTTATTGCTCAGCGG3' ) 用于验证 pET12a 载体中所克隆 GCAT 的序列和方向。用 ABI Prism® BigDye™ 终止子循环测序试剂盒 (Terminators Cycle sequencing kit)，以 500ng 质粒 DNA 作为模板，以及 3.2pmol T7 启动子和终止子引物实施 DNA 序列测定。

[0538] 图 40 所示构建体用于转化感受态细菌宿主菌株 BL21 (DE3) pLysS (Novagen) 并且可选出氨苄西林抗性转化体用于表达分析。

[0539] 嗜水气单胞菌转移酶在 BL21 (DE3) pLysS 中的表达

[0540] 包含表达载体 pet12a-A.h.GCAT-pSMa 的大肠杆菌菌株 BL21 (DE3)pLysS 在 LB 培养基 +100  $\mu$ g/ml 氨苄西林中生长, 37 $^{\circ}$ C 振荡保温, 直到  $OD_{600} = 0.6$  到 1.0。用 IPTG (0.4mM) 诱导所述培养物, 再保温 3 小时。IPTG 诱导 0、1、2 和 3 小时后取样品。以卵磷脂作为底物用 NEFA 试剂盒检测酶活性 (图 41)。

[0541] 用于制备活性更强的酶的生长优化

[0542] 包含表达载体 pet12a-A.h.GCAT-pSMa 的 BL21 (DE3)pLysS 在 LB 培养基 +100  $\mu$ g/ml 氨苄西林中生长, 在不同生长温度 (37 $^{\circ}$ C、30 $^{\circ}$ C、20 $^{\circ}$ C)、振荡条件下保温。产生活性脂质酰基转移酶的最适条件是当如图 42 所示培养物在 30 $^{\circ}$ C 生长时。

[0543] 重组嗜水气单胞菌转移酶 (GCAT) 的部分纯化

[0544] 包含表达载体 pet12a-A.h.GCAT-pSMa 的菌株 BL21 (DE3)pLysS 在 37 $^{\circ}$ C 生长, 通过超声处理制备细胞粗提物。重组体酶用 Qiagen 的 Ni-NTA 自旋 (spin) 试剂盒从经超声处理的细胞粗提物进一步纯化。使用 NEFA 试剂盒且以卵磷脂作为底物测定磷脂酶活性 (图 43)。

[0545] 实施例 3: 气单胞菌属转移酶在枯草芽孢杆菌 (枯草芽孢杆菌) 163 中的表达

[0546] 质粒构建

[0547] 两种不同的枯草芽孢杆菌表达载体 (pUB 110&pBE5) 用于气单胞菌属基因在枯草芽孢杆菌的异源表达。pUB 110 载体包含  $\alpha$  淀粉酶启动子而 pBE 载体含有作为融合的气单胞菌属基因的表达调节区的 P32 启动子。在 pUB110 中, 气单胞菌属成熟 GCAT 基因的第二个氨基酸与枯草芽孢杆菌的木聚糖酶信号肽的最后一个氨基酸通过在限制酶切位点 NheI 在框内融合, 在成熟蛋白之前产生两个额外的氨基酸。pBE5 含有在 Nco I 位点融合的 cgtase 信号序列, 使重组蛋白质分泌入培养过滤物。

[0548] 实施 PCR 反应获得与 pUB 110 和 pBE5 载体的信号序列框内融合的气单胞菌属基因。用下列的配对引物对嗜水气单胞菌基因进行 PCR。

[0549] PCR 反应 1: usAHncol (5' ATGCCATGGCCGACAGCCCGTCCCGCC3', SEQ ID No.46) 和 IsAH (5' TTGGATCCGAATTCATCAATGGTGATG3', SEQID No.47)

[0550] PCR 反应 2: US-AhnheI (5' TTGCTAGCGCCGACAGCCCGTCCCGCC3', SEQ ID No.48.) 和 IsAH (5' TTGGATCCGAATTCATCAATGGTGATG3, SEQID No.49)

[0551] 用下列引物对杀鲑气单胞菌基因进行 PCR。

[0552] PCR 反应 3: US-Asncol (5' TTGCCATGGCCGACACTCGCCCCGCC3', SEQ ID No.50) 和 IsAH (5' TTGGATCCGAATTCATCAATGGTGATG3', SEQID No.51)

[0553] PCR 反应 4: US-ASnheI (5' TTGCTAGCGCCGACACTCGCCCCGCC3', SEQ ID No.52) 和 IsAH (5' TTGGATCCGAATTCATCAATGGTGATG3', SEQID No.53)

[0554] 全部 PCR 产物克隆到 PCR blunt II (TOPO 载体) 中并用反向或正向测序引物测序。

[0555] 来自 PCR 反应 1 和 3 的克隆用 NcoI 和 BamHI 切割, 作为插入物与 NcoI/BamHI/磷酸酶切割的 pBE5 载体连接。来自 PCR 反应 2 和 4 的克隆用 NheI 和 BamHI 切割, 作为插入物与 NheI/BamHI/磷酸酶切割的 pUB 载体连接。

[0556] 枯草芽孢杆菌中气单胞菌属转移酶基因的表达与所述酶的活性表征。

[0557] 来自两种气单胞菌种的酰基转移酶已成功地在大肠杆菌中表达 (结果如上)。芽

孢杆菌 pUB 110 和 pBE5 基因融合构建体用来转化枯草芽孢杆菌，并通过铺在卡那霉素板上选择转化株。经分离并在 2xYT 中生长的卡那霉素抗性转化体能在枯草芽孢杆菌中异源表达气单胞菌基因。培养过滤物除具有酰基转移酶与磷脂酶活性外，还具有二半乳糖基二酰甘油 (digalactosyldiacylglycerol) (DGDG) 半乳糖脂酶活性。针对二半乳糖基二酰甘油 (DGDG) 的活性在将上清液与底物保温 60 分钟后测定，来自小麦粉的 DGDG (来自 Sigma) 以及针对卵磷脂的活性在图 44 中显示。芽孢杆菌在培养基中培养过夜 (20-24 小时) 至 48 小时以后作为分泌性蛋白产生。在一些情况中，气单胞菌基因的表达影响芽孢杆菌和大肠杆菌中的细胞活力和生长，因此有必要仔细选择表达菌株并优化生长条件以确保表达。例如，数种芽孢杆菌宿主菌株 (B.s 163, DB104 和 OS 21) 用表达载体转化用于生长比较。B.s163 可由两种气单胞菌基因转化并具有表达活性蛋白的能力。DB104 可用所有构建转化但只有表达杀鲑气单胞菌转移酶的能力。

[0558] 实施例 4：大肠杆菌中产生的气单胞菌脂质酰基转移酶的发酵和纯化

[0559] 大肠杆菌发酵：

[0560] 微生物

[0561] 本研究中使用大肠杆菌的两个菌株，一个包括嗜水气单胞菌 (实施 2) 脂质酰基转移酶而另外的包括杀鲑气单胞菌 (实施例 1) 脂质酰基转移酶。

[0562] 包含嗜水气单胞菌基因的大肠杆菌菌株命名为 DIDK0124，包含杀鲑气单胞菌基因的大肠杆菌菌株命名为 DIDK0125。DIDK0124 的发酵物命名为 HYDRO0303 并且 DIDK0125 的发酵物命名为 SAL0302。从 HYDRO025 纯化的蛋白质命名为 REF#138。从 HYDRO0303 纯化的蛋白命名为 REF#135。

[0563] 生长培养基与培养条件

[0564] LB-琼脂

[0565] 用于维持菌株的 LB-琼脂板包括：10g/L 胰蛋白胨，5g/L 酵母提取液，5g/L NaCl，15g/L 琼脂，100mg/L 氨苄西林和 35mg/L 氯霉素。琼脂板在 30°C 保温。

[0566] LB 摇瓶

[0567] 用于产生生物反应物培养用接种物原料的 LB 培养基 (50mL pr 摇瓶) 包括：10g/L 胰蛋白胨，5g/L 酵母提取物，5g/L NaCl，100mg/L 氨苄西林 (ampicillin) 和 35mg/L 氯霉素 (chloramphenicol)。摇瓶从 LB 琼脂板接种，30°C、200rpm 保温。

[0568] 生物反应物的培养

[0569] 生物反应物培养在 6L 内部建造的 (in-built) 生物反应器中进行，所述生物反应器中充满含有以下物质的 4L 培养基：10g/L 胰蛋白胨、5g/L 酵母提取液、5g/L NaCl、8g/L  $\text{KH}_2\text{PO}_4$ 、0.9g/L  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 、40g/L 葡萄糖一水合物、0.4mL/ADD APT® Foamstop S 中 260 (ADD APT Chemicals AG, Helmond, The Netherlands)、10mg/L  $(\text{NH}_4)_2\text{Fe}(\text{SO}_4)_2 \cdot 6\text{H}_2\text{O}$ 、0.7mg/L  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ 、3mg/L  $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ 、3mg/L  $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ 、10mg/L EDTA、0.1mg/L  $\text{NiSO}_4 \cdot 6\text{H}_2\text{O}$ 、0.1mg/L  $\text{CoCl}_2$ 、0.1mg/L  $\text{H}_3\text{BO}_3$ 、0.1mg/L KI、0.1mg/L  $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ 、1g/L 氨苄西林和 35mg/L 氯霉素。

[0570] 在生物反应器中接种一定量的 LB 培养基确保在大约 20 小时培养后的生长终点 (通过以下参数计算：最大比生长速率  $0.6\text{h}^{-1}$ ，LB 摇瓶的  $\text{OD}_{600}$  和生物反应器终末  $\text{OD}_{600}$ ，其为大约 20)。

[0571] SAL0302 接种了 10mL 的 LB 培养基, 而 HYDR00303 接种了 4mL 的 LB 培养基。

[0572] 生物反应器在下列条件下操作: 温度 30°C, 搅拌 800-1000rpm(依照实验), 通气 5L/min, pH 6.9, pH 控制 8.75% (w/v) NH<sub>3</sub>-水和 2M H<sub>2</sub>SO<sub>4</sub>。添加异丙基 β-D-硫半乳糖苷 (thiogalactoside) 至终末浓度 0.6mM, 诱导完成, 分别产生 0.4 摩尔 (HYDR00303) 和 0.7 摩尔 CO<sub>2</sub>。

[0573] 收获

[0574] 以下操作步骤用于生物质的收获与匀浆:

[0575] 1) 来自发酵的发酵肉汤在 5000×g 离心, 4°C 放置 10 分钟, 弃去上清液。所述生物质用前 20°C 放置保存。融化所述生物质并在 500mL 20mM NaH<sub>2</sub>PO<sub>4</sub>, pH 7.4, 500mM NaCl, 10mM 咪唑和完全 (无 EDTA) 蛋白酶抑制剂 (Roche, Germany) 中重悬。

[0576] 2) 悬浮的生物质在 2kbar, 4°C, 在来自 Constant Systems 公司 (Warwick, UK) 的细胞破裂仪 (cell disrupter) 中匀浆。

[0577] 3) 细胞碎片在 10.000×g 离心去除, 4°C 放置 30 分钟, 收集上清液。

[0578] 4) 使上清液在 13.700×g 离心而澄清化, 4°C 放置 60 分钟, 收集上清。

[0579] 5) 所述上清液通过 0.2 μm Vac Cap 过滤器 (Pall LifeSciences, UK) 过滤, 收集滤出液立即层析纯化。

[0580] 转移酶的层析纯化

[0581] 用 50ml 螯合型琼脂糖 ff. 凝胶填充柱 (2.5×10cm), 并用镍-硫酸盐装料 (charged with) (根据生产商 Amersham Biosciences 所述方法)。以 200ml 20mM NaH<sub>2</sub>PO<sub>4</sub>, pH 7.4, 500mM NaCl, 10mM 咪唑平衡柱, 将 400ml 粗提物以 5ml/min 的流速加样于柱。用 20mM NaH<sub>2</sub>PO<sub>4</sub>, pH 7.4, 500mM NaCl, 10mM 咪唑洗柱直到 UV<sub>280</sub> 达到基线, 然后用 40ml 20mM NaH<sub>2</sub>PO<sub>4</sub>, pH 7.4, 500mM NaCl, 500mM 咪唑洗脱 GCAT。

[0582] 实施例 5: 枯草芽孢杆菌中产生的气单胞菌脂质酰基转移酶的发酵与纯化

[0583] 发酵

[0584] BAC0318-19, BAC0323-24

[0585] 微生物

[0586] 本研究中所用的微生物来源于含有插入 pUB110OIS 载体中、编码杀鲑气单胞菌转移酶的基因的质粒对枯草芽孢杆菌宿主菌株 #163 的转化。该基因的表达受 α 淀粉酶启动子控制, 转移酶的分泌由枯草芽孢杆菌木聚糖酶信号序列介导 (实施例 3)。菌株命名为 DIDK0138 (发酵 BAC0318-19) 和 DIDK0153 (发酵 BAC0323-24)。

[0587] 生长培养基及培养条件

[0588] 预培养基

[0589] 摇瓶 (总容积 500mL, 有隔板) 中加入 100mL 培养基, 所述培养基包括:

[0590] NaCl 5g/L

[0591] K<sub>2</sub>HPO<sub>4</sub> 10g/L

[0592] 大豆粉 20g/L

[0593] 酵母提取物, BioSpringer 106 20g/L

[0594] 消泡剂, SIN260 5mL/L

[0595] 高压灭菌前 pH 调整到 7.0

[0596] 高压灭菌后向每个摇瓶中添加 6mL 50% (w/w) Nutriose。并且在高压灭菌后添加浓度 50mg/L 的卡那霉素。

[0597] 接种

[0598] 预培养摇瓶用直接来自 25% (w/v) 甘油储存液的冰冻培养物接种。摇瓶在 33°C、175rpm 保温大约 16 小时，取 50mL 接种发酵罐。

[0599] 发酵

[0600] 发酵在 6L 内部建造的发酵罐中进行。

[0601] 成批发酵培养基 (batch medium) (3L) 包括：

[0602] 玉米浆 (50% dw) 40g/L

[0603] 酵母提取物 BioSpringer 153 (50% dw) 10g/L

[0604] NaCl 5g/L

[0605]  $\text{CaCl}_2, 2\text{H}_2\text{O}$  0.25g/L

[0606]  $\text{Mn}(\text{NO}_3)_2, \text{H}_2\text{O}$  0.2g/L

[0607] 消泡剂 SIN260 1mL/L

[0608] 卡那霉素 (经滤过灭菌处理后加入高压灭菌的发酵罐) 50mg/L

[0609] 进料 (feed) 包括：

[0610] 葡萄糖一水合物 540g/kg

[0611]  $\text{MgSO}_4, 7\text{H}_2\text{O}$  4.8g/kg

[0612] 消泡剂 SIN260 4mL/kg

[0613] 酵母提取物, BioSpringer 153 (50% dw) 150g/kg

[0614] (分别高压灭菌)

[0615] 根据下列方程式, 发酵 BAC0318 和 BAC0323 中的进料基于蓄积的  $\text{CO}_2$  开始：

[0616] 进料率 - 流率 (flow) [g/h] = 0,  $\text{AcCO}_2 < 0.15$

[0617] 进料率 - 流率 [g/h] =  $2.85 + t \cdot 1.54$ ,  $\text{AcCO}_2 \geq 0.15$  和  $t < 12$

[0618] 进料率 - 流率 [g/h] = 21.3,  $t > 12$

[0619] t: 从蓄积的  $\text{CO}_2$  ( $\text{AcCO}_2$ ) 达到 0.15 摩尔的点开始的时间 (小时)。

[0620] 根据下列方程式, 发酵 BAC0319 和 BAC0324 中的进料基于蓄积的  $\text{CO}_2$  开始：

[0621] 进料率 - 流率 [g/h] = 0,  $\text{AcCO}_2 < 0.15$

[0622] 进料率 - 流率 [g/h] =  $2.0 + t \cdot 1.08$ ,  $\text{AcCO}_2 \geq 0.15$  和  $t < 12$

[0623] 进料率 - 流率 [g/h] = 15,  $t > 12$

[0624] t: 从蓄积的  $\text{CO}_2$  ( $\text{AcCO}_2$ ) 达到 0.15 摩尔的点开始的时间 (小时)。

[0625] 加入 12.5% (w/v) 氨水或 2M 磷酸使 pH 值控制在 7.0。

[0626] 通气量为 3L/min 对应 1vvm。

[0627] 温度 33°C。

[0628] 发酵罐配备有相距 10cm 的两个 8 cm  $\varnothing$  Rushton 推进器 (impeller)。

[0629] 收获

[0630] 在室温 16,000 $\times$ g 离心 10 分钟, 去除生物质。上清液经过滤灭菌, 滤出液用于纯化和应用试验。

[0631] 实施例 6: “经缓冲的底物中的转移酶实验” 用于测定酶的酰基转移酶活性

[0632] 从杀鲑气单胞菌分离脂质酰基转移酶，并在枯草芽孢杆菌中表达。该酶在形成胆固醇酯过程中将脂肪酸从卵磷脂转移到胆固醇非常有效。还显示，该酶具有一些水解活性，这通过游离脂肪酸的形成观察。常规磷酸酯酶 (EC3.1.1.4 和 EC3.1.1.32) 具有在形成游离脂肪酸和溶血卵磷脂过程中水解卵磷脂的能力，对于这些酶没有转移酶反应的报道。

[0633] 我们详细描述了能够测定酶的转移酶和水解活性的实验，并由此鉴定本发明的脂质酰基转移酶，所述实验利用含有卵磷脂和胆固醇的底物。在该研究中，使用基于分散在缓冲液中的磷脂酰胆碱和胆固醇的底物。反应产物的定量通过从所述底物提取脂质并随后对该脂质组分进行 GLC 分析来进行。

[0634] 步骤

[0635] 材料

[0636] L- $\alpha$ - 磷脂酰胆碱 95% (植物) Avanti no.441601

[0637] 胆固醇: Sigma cat.C 8503

[0638] 胆固醇基棕榈酸酯 (Cholesteryl Palmitate), Sigma C 6072

[0639] 胆固醇基硬脂酸酯 (Cholesteryl Stearate), Sigma C 3549

[0640] HEPES 缓冲液 Sigma cat.No.H 3375

[0641] 氯仿, 分析级.

[0642] 酶

[0643] 纯化的 GCAT 来自杀鲑气单胞菌 #178-9

[0644] TLC 分析。

[0645] TLC- 板在加热箱中 (110°C) 活化 1/2h.

[0646] 将 100ml 流动缓冲液倾倒入有盖层析小室。所述小室的壁覆盖有滤纸 (Whatman 2) 以用溶剂蒸气饱和和所述小室。

[0647] 将 TLC- 板置于框架中 (in frame) 并将样品加于 TLC 板 2cm 自底部。将该 TLC 板置于具有流动缓冲液的 TLC 小室中。当流动缓冲液的高度达到距离底部 14cm 时, 将该 TLC 板取出并在烘板上干燥, 然后置于加热箱中 110°C 10 分。

[0648] 随后将该 TLC- 板浸于展开剂中, 然后在加热箱中干燥 110°C 15 分。

[0649] 流动 - 缓冲液:

[0650] Nr.IV: 氯仿: 甲醇: H<sub>2</sub>O (65 : 25 : 4)

[0651] Nr.I: P- 醚: MTBE : 乙酸 (60 : 40 : 1)

[0652] 展开缓冲液 (钒酸盐 - 缓冲液):

[0653] 32g Na<sub>2</sub>CO<sub>3</sub> ad 300ml H<sub>2</sub>O (1M)

[0654] 将 18.2g 五氧化二钒 (V<sub>2</sub>O<sub>5</sub>) 加入, 轻柔搅拌使其溶解。

[0655] 将所述溶液冷却至室温。

[0656] 小心加入 460ml 2.5M H<sub>2</sub>SO<sub>4</sub>. (460ml H<sub>2</sub>O+61ml H<sub>2</sub>SO<sub>4</sub>)

[0657] 加入水至 1000ml。

[0658] GLC 分析

[0659] Perkin Elmer Autosystem 9000 Capillary Gas Chromatograph, 其配备有 WCOT 融合的硅柱 12.5m×0.25mm ID×0.1 $\mu$  薄膜厚度 5% 苯基甲基硅酮 (phenyl-methyl-silicone)

(CP Sil 8CB 来自 Chrompack)。

[0660] 载气：氦

[0661] 注射器 .PSSI 冷裂注射 (最初温度 50°C 加热到 385°C)，体积 1.0  $\mu$ l

[0662] 检测器 FID：395°C

[0663] 烘箱程序： 1 2 3

[0664] 烘箱温度， °C . 90 280 350

[0665] 等温，时间，分钟 . 1 0 10

[0666] 温度比， °C /min. 15 4

[0667] 样品制备：30mg 样品溶于 9ml 2 : 1 庚烷：吡啶，其含有内部标准十七烷，0.5mg/ml。将 300  $\mu$ l 的样品溶液移到曲管形瓶 (crimp vial)s，添加 300  $\mu$ l IMSTFA (N-甲基 N-三甲基甲硅烷基三氟乙酰胺 (N-Methyl-N-trimethylsilyl-trifluoroacetamide))，60°C 反应 20 分钟。

[0668] 计算：甘油单-二-三酯 (mono-di-triglycerides) 和游离脂肪酸的反应因子 (response factor) 通过标准 2 (单-双-甘油三酯) 确定，胆固醇，胆固醇基棕榈酸酯 (Cholesteryl palmitate) 和胆固醇基硬脂酸酯的反应因子通过纯化的参考物质 (称重纯化物质 10mg) 来确定。

[0669] 结果：基于磷脂酰胆碱和胆固醇作为底物的转移酶测定。

[0670] 在下文中，转移酶的转移酶活性在基于磷脂酰胆碱和胆固醇的底物中根据以下方法测定。

[0671] 450mg 磷酸卵磷脂 (> 95% PC Avanti 产品号 441601) 和 50mg 胆固醇溶于氯仿，在真空条件下蒸发干燥。300mg 胆固醇 / 磷脂酰胆碱混合物移到一 Wheaton 玻璃杯中并且添加 15ml 50mM HEPES 缓冲液 pH 7。在搅拌过程中脂质分散到缓冲液中。

[0672] 底物用磁力搅拌器混和同时加热至 35°C，添加 0.25ml 酶溶液。这是含水量大约 95% 的高水分量环境。

[0673] 在反应 0, 5, 10, 15, 25, 40 和 60 分钟后取出 2ml 样品。立即添加 25  $\mu$ l 4M HCl 以酸化游离脂肪酸，终止酶反应。添加 3.00ml 氯仿。将样品在 Whirley 中剧烈摇晃振摇 30 秒。将样品离心，分离 2ml 氯仿相并用 0.45- $\mu$ m 的滤纸过滤，进入 10ml 涂焦油的 Dram 玻璃杯中。在 60°C 氮蒸汽条件下蒸发氯仿。并再次称量样品。提取的脂质通过 GLC 分析。

[0674] GLC 分析结果如图 17 所示。结果以所提取脂质的百分比计算。形成的脂肪酸和胆固醇酯的数量作为时间的函数，如图 45 所示。从图 45 推断酶反应作为时间的函数不是线性的，因为可观察到最初的水解和转移酶活性较高。在大约 10 分钟之后，大约 60 分钟之前，反应显示出脂肪酸和胆固醇酯的形成与时间的线性关系。因此决定以这种时间间隔观察酶反应。

[0675] 表 1

[0676]

分钟	0	5	10	15	25	40	60
胆固醇, %	10.064	8.943	8.577	8.656	8.102	7.856	7.809
胆固醇酯, %	0.000	1.571	2.030	2.058	2.282	2.659	3.081
总 FFA, %	0.260	1.197	1.239	1.466	2.445	2.943	3.940

[0677] 根据已知的反应混合物中的脂质含量与添加的酶量可计算脂肪酸和胆固醇酯的生成量, 用  $\mu\text{mol/ml}$  酶表示 (表 18 与图 58)

[0678] 表 2

[0679]

分钟	10	15	25	40	60
	$\mu\text{mol/ml}$	$\mu\text{mol/ml}$	$\mu\text{mol/ml}$	$\mu\text{mol/ml}$	$\mu\text{mol/ml}$
总 FFA	58.1	68.7	114.6	138.0	184.7
胆固醇酯	88.8	90.0	99.3	115.6	133.8

[0680] 从表 2 的结果和图 46 的曲线斜率, 可以计算出作为时间函数的脂肪酸和胆固醇酯的数量, 用  $\mu\text{mol/min}$  每 ml 酶表示。

[0681] 水解活性与转移酶活性的计算如表 3 所示。 用在上文描述的 % 酰基转移酶活性的测定方案测定相对转移酶活性。

[0682] 表 19

[0683]

水解活性 (脂肪酸)	2.52	$\mu\text{mol/min}$ 每 ml 酶
转移酶活性 (胆固醇酯)	0.94	$\mu\text{mol/min}$ 每 ml 酶
总活性	3.45	$\mu\text{mol/min}$ 每 ml 酶
相对转移酶活性	27.1	%
相对水解活性	72.9	%

[0684] 其它酶的转移酶活性筛选

[0685] 上面提到的方法用来筛选不同脂解酶的转移酶活性和水解活性。 检测的酶如表

4 所示。

[0686] 表 4

[0687]

		1	2	3	4	5
底物	ml	15	15	15	15	15
# 178-9 的转移酶杀鲑气单胞菌 PLU-7/ml	ml	0.25				
5% # 3016, LIOPAN <sup>®</sup> F (尖镰孢)	ml		0.25			
5%细毛嗜热霉	ml			0.25		
5%皱落念珠菌( <i>Candida rugosa</i> ) # 2983	ml				0.25	
5%柱状念珠菌( <i>Candida cylindracea</i> ) # 3076	ml					0.25

[0688] 包含 300mg 磷脂酰胆碱 / 胆固醇的底物分散到 50mM pH 7.0 的 HEPES 缓冲液中，搅拌下加热至 35℃，添加酶溶液，搅拌样品并将其维持在 35℃。以规律的时间间隔取出样品并用氯仿提取。分离的脂质用 GLC 分析，结果如表 21 所示。

[0689] 表 5

[0690]

样品								
1	<b>转移酶 178-9</b>							
	分钟	0	5	10	15	25	40	60
	FFA	1.216	2.516	2.983	2.62	2.894	3.448	3.911
	胆固醇	7.547	6.438	6.365	6.15	6.136	5.936	5.662

[0691]

	胆固醇酯	0	1.835	2.177	2.44	2.58	2.851	3.331
2	尖 镰 孢 (LIPOPAN® F)	0	5	10	15	25	40	60
	FFA	1.216	1.345	1.796	1.95	2.487	2.424	2.977
	胆固醇	7.547	7.309	7.366	7.33	7.429	7.341	7.326
	胆固醇酯	0	0.26	0.386	0.35	0.267	0.36	0.394
3	细毛嗜热霉	0	5	10	15	25	40	60
	FFA	1.216	0.853	0.875	1	0.896	1.105	1.009
	胆固醇	7.547	7.384	7.639	7.63	7.675	7.603	7.529
	胆固醇酯	0	0	0	0	0	0	0
4	皱落念珠菌 (# 2938)	0	5	10	15	25	40	60
	FFA	1.216	0.982	0.987	1.02	1.135	1.131	1.15
4	胆固醇	7.547	7.438	7.656	7.66	7.638	7.575	7.585
	胆固醇酯	0	0	0	0	0	0	0
5	柱状念珠菌 (# 3076)	0	5	10	15	25	40	60
	FFA	1.216	1.032	1.097	1.07	1.203	1.131	1.43
	胆固醇	7.547	7.502	7.425	7.65	7.619	7.502	7.411
	胆固醇酯	0	0	0	0	0	0	0

[0692] 从 GLC 分析可以观察到只有脂质酰基转移酶 (178-9) 产生大量的胆固醇酯和脂肪酸。并观察到来自尖镰孢的磷脂酶使得游离脂肪酸量稳定增加, 仅最初有少量的胆固醇酯形成, 但是并未观察到胆固醇酯作为时间函数的增高。

[0693] 根据已知脂质底物的量与 GLC 分析, 可以基于得自 10 到 60 分钟反应时间的结果计算出相对转移酶活性和相对水解活性。转移酶 178-9 和尖镰孢脂酶的结果如表 6 所示。其它被检验的酶不表现活性。

[0694] 表 6

[0695]

	转移酶 178-9	尖镰孢
水解活性, 微摩尔/分钟 每毫升酶	1.03	0.96
转移酶活性, 微摩尔/分钟 每毫升酶	0.40	0.01
总活性, 微摩尔/分钟 每 毫升酶	1.43	0.98
相对水解活性	71.8	98.7
相对转移酶活性	28.2	1.3

[0696] 表 6 所示结果证实脂质酰基转移酶（样品 178-9）具有显著的转移酶活性。同样也观察到相对转移酶活性与表 3 所示实验十分一致。

[0697] 然而观察到来自尖镰孢的磷脂酶具有很低的转移酶活性。转移酶活性非常低以至于其落在分析的不确定性中。正如意料中一样，尖镰孢的磷脂酶具有显著的水解活性。

[0698] 结论

[0699] 基于纯化的磷脂酰胆碱和胆固醇的人造底物作为测定杀鲑气单胞菌转移酶活性的底物。在反应时间 10 到 60 分钟之间，游离脂肪酸与胆固醇酯形成作为时间的函数呈现几乎线性。基于反应时间 10 到 60 分钟之间的活性，可计算出水解活性与转移酶活性。

[0700] 本测定方法中的底物浓度相对低于蛋黄中而言较低，而含水量相对较高。

[0701] 根据来自杀鲑气单胞菌的脂质酰基转移酶（这种情况中是 GCAT）在缓冲液中的人造底物磷脂酰胆碱 / 胆固醇中的实验的结果，推论这种酶在含水量很高的体系中也具有很好的转移酶活性。

[0702] 可利用缓冲液中的磷脂酰胆碱 / 胆固醇实验来测定酶的转移酶活性和水解活性。缓冲液中的磷脂酰胆碱 / 胆固醇仅在某一时间范围内呈线性。

[0703] 实施例 7：来源于杀鲑气单胞菌的脂质酰基转移酶的固定化

[0704] 来源于杀鲑气单胞菌的脂质酰基转移酶（本例中为 GCAT）通过丙酮的沉淀作用固定在 Celite 535535（来自 Fluka）上。20mM pH 7 的 TEA 缓冲液中的 10ml 酶溶液与 0.1 克 Celite 535 535（来自 Fluka）在室温缓慢搅拌 2 小时。

[0705] 在持续搅拌过程中加入 50ml 冷丙酮。

[0706] 沉淀物通过在 5000g 离心 1 分钟分离。

[0707] 沉淀物用 20ml 冷丙酮洗涤 2 次。

[0708] Celite 在环境温度干燥 1 小时

[0709] 所述酶在含水量高（6-89%）的环境中具有高的活性，所述转移酶和用于本发明的其它转移酶因此也可用于固定化的酶在含水量高的应用中。这就允许在利用转移酶的脂质生物转化中替代目前的固定化的脂酶。

[0710] 实施例 8：来自嗜水气单胞菌 (Ahyd2) 的脂质酰基转移酶的变体 (SEQ ID No.36 (参见图 47))

[0711] 用来自 Stratagene, La Jolla, CA 92037, USA 的 QuikChange<sup>®</sup> Multi-Site

[0712] Directed Mutagenesis 试剂盒引入突变，按 Stratagene 提供说明进行。

[0713] 变异位于酪氨酸 256 的变体表现出对磷脂的活性增加。

[0714] 变异位于酪氨酸 256 和酪氨酸 260 的变体显示对半乳糖脂的活性增加。

[0715] 变异位于酪氨酸 265 的变体显示对作为酰基供体的半乳糖脂的转移酶活性增加。

[0716] 编号指示在下列序列中的位置：来自嗜水气单胞菌的酶，其氨基酸序列如图 471SEQ ID No.36 所示（加下划线的氨基酸表示木聚糖酶信号肽）。核苷酸序列如图 72 中 SEQ ID No 54 所示。

[0717] 实施例 9 “低水分环境中的测定法”

[0718] 脂解酶在低含水量环境中的转移酶反应。

- [0719] 方法
- [0720] 材料
- [0721] 胆固醇 Sigma 目录 .C 8503
- [0722] L- $\alpha$ - 磷脂酰胆碱 95% (植物) Avanti # 441601
- [0723] 大豆油, Aarhus United, DK.
- [0724] 氯仿, 分析级
- [0725] 酶
- [0726] #179, GCAT, 来自杀鲑气单胞菌
- [0727] #2427, 磷脂酶 A1, 来自尖镰孢 (*Fusarium oxysporum*)。 **LIPOPAN® F** 来自 Novozymes, Denmark
- [0728] #1991, 来自胰腺的磷脂酶 A2, LIPOMOD 22L 来自 Biocatalysts, UK
- [0729] #2373, 南极念珠菌 (*Candida antarctica*) 脂酶, Novozyme 525 L 来自 Novozymes Denmark.
- [0730] 酶测定
- [0731] 通过在搅拌过程中加热到 60°C, 使 13.1% 卵磷脂与 6.6% 胆固醇溶于大豆油
- [0732] 用 20ml Wheaton 玻璃杯称量底物并加热到 46°C
- [0733] 添加水与酶溶液并开始计时
- [0734] 以规律间隔将 50mg 样品转移到 10ml Dram 玻璃杯中并冷冻
- [0735] 分离出的脂质通过 GLC 分析
- [0736] GLC 分析
- [0737] GLC 分析方案 - 见实施例 6
- [0738] 结果
- [0739] 实验设计如表 33 所示
- [0740] 将以含有 13.1% 卵磷脂与 6.6% 胆固醇的大豆油为基础的底物加热到 46°C。添加酶溶液并开始计时。
- [0741] 反应 30, 60 和 120 分钟后, 取样品做 GLC 分析。
- [0742] 表 33

		1	2	3	4	5
	底物	克	5	5	5	5
	转移酶 # 179-C72, 56PLU-7/ml	ml		0.3		
	# 2427, 200 PLU-7/ml	ml			0.3	
[0743]	胰腺 PLA 2 # 1991 6300 PLU/ml	ml				0.3
	Novozyme 525 L, # 2373, 200 LIPU/ml	ml				
	水	ml	0.3			
	%水		6	6	6	6

[0744] GLC 结果如表 9 所示。结果以基于总样品组分的百分比表示。以 GLC 结果为基础, 可计算出相对不加酶的对照样品, 通过酶反应产生的脂肪酸与胆固醇酯的量。在这样的实验条件下, 总体酶活性可以作为水解活性 (通过游离脂肪酸形成测定) 与转移酶活性 (通过胆固醇酯形成估计) 来评价。从这些结果以及脂肪酸与胆固醇酯的分子量信息可计算出相对摩尔水解活性与相对摩尔转移酶活性, 如表 10 所示。

[0745] 表 9

	酶	反应时间 分钟	脂肪酸 %	胆固醇 %	胆固醇酯 %
	对照	120	0.533	7.094	0.000
	#179	30	0.770	5.761	2.229
	#179	60	0.852	5.369	2.883
[0746]	#179	120	0.876	4.900	3.667
	#2427	30	3.269	7.094	0.000
	#2427	60	3.420	7.094	0.000
	#2427	120	3.710	7.094	0.000
	#1991	30	2.871	7.094	0.000
	#1991	60	3.578	7.094	0.000
	#1991	120	3.928	7.094	0.000

[0747] 表 10

[0748]

酶	反应时间 分钟	产生的 脂肪酸	消耗的 胆固醇	产生的 胆固醇酯	水解 活性%	转移酶 活性%
#179	30	0.238	1.334	2.229	20	80
#179	60	0.319	1.725	2.883	21	79
#179	120	0.343	2.195	3.667	18	82
#2427	30	2.737	0.000	0.000	100	0
#2427	60	2.887	0.000	0.000	100	0
#2427	120	3.177	0.000	0.000	100	0
#1991	30	2.338	0.000	0.000	100	0
#1991	60	3.046	0.000	0.000	100	0
#1991	120	3.395	0.000	0.000	100	0
#2373	30	0.885	0.000	0.000	100	0
#2373	60	0.888	0.000	0.000	100	0
#2373	120	1.383	0.000	0.000	100	0

## [0749] 结论

[0750] 在这些实验中可观察到所有检测酶都表现有水解活性，这是由于脂肪酸数量增加。但只有来自杀鲑气单胞菌的 GCAT 表现出转移酶活性。因此推论在具有含 6% 水的卵磷脂与胆固醇的油性体系中，来自尖镰孢的磷脂酶 A1，来自胰腺的磷脂酶 A2 和来自南极念珠菌的脂酶只显示水解活性。

## [0751] 实施例 10：利用本发明的脂质酰基转移酶制备蛋白质酯

[0752] 脂肪酸诸如蔗糖酯和葡萄糖酯的碳水化合物酯通常通过使脂肪酸或脂肪酸皂以及碳水化合物在高温反应制备 (Journal of the Americal Oil Chemists' Society (1978) 55 ; 4 ; 398-401)。但该方法的缺点在于形成副反应和有色的副产物。

[0753] 本发明中，脂肪酸的碳水化合物酯通过转移酶反应利用卵磷脂作为脂肪酸供体及碳水化合物如葡萄糖作为受体分子来制备。

[0754] 该反应在流动反应器 (flow reactor) 中利用固定在固体支持物上的脂质酰基转移酶进行。

## [0755] 方法

[0756] 在搅拌下将 100g 葡萄糖溶于 1000ml 水，然后将 200g 磷脂酰胆碱在搅拌下分散到该水相中并加热到 40°C。

[0757] 调节 pH 到 pH 6.5。

[0758] 在流动反应器中充满 100g 固定在固体支持物上的、来自杀鲑气单胞菌的脂质酰基转移酶。

[0759] 将该流动反应器置于加热箱 40°C。

[0760] 将该反应混合物以 2ml/min 泵入该柱。

[0761] 收集反应产物。

[0762] 反应产物中的水通过薄膜真空蒸发 (thin film vacuum evaporation) 去除并分离脂质。

[0763] 葡萄糖酯通过溶剂分级自脂质分离。

[0764] 碳水化合物酯可用于多种用途中，诸如食品和非食品工业中的有效乳化剂。

[0765] 实施例 11- 利用本发明的脂质酰基转移酶制备蛋白质酯

[0766] 本发明中，氨基酸，肽或蛋白质的脂肪酸缩合物通过转移酶反应制备。在该反应中，磷脂酰胆碱用作供体将脂肪酸转移到氨基酸（诸如酪氨酸，丝氨酸或苏氨酸）的游离羟基，所述氨基酸具有可用于酯化的游离羟基。

[0767] 方法 1。

[0768] 50g 1- 酪氨酸（或丝氨酸或苏氨酸）在搅拌下溶于 1000ml 水，在搅拌中将 200g 磷脂酰胆碱分散于水相，并加热到 40°C。

[0769] pH 调节到 pH 7，并用 NaOH 或 HCl 维持该 pH。

[0770] 加入 50ml 来自杀鲑气单胞菌的脂质酰基转移酶，并在搅拌下使该反应在 40°C 继续。

[0771] 间隔规律的时间取样，并通过 TLC 以及 HPLC 分析。

[0772] 20h 反应时间后，该反应到达平衡并终止反应。

[0773] 酪氨酸脂肪酸缩合物，卵磷脂及溶血卵磷脂通过离心自该反应介质分离，所述离心根据标准方法（见“Centrifuges, Filtering”中 Ullmann's Encyclopedia of Industrial Chemistry for example (2002) by Wiley-VCH Verlag GmbH & Co.KGaA) 实施。

[0774] 酪氨酸脂肪酸缩合物进一步通过疏水反应柱层析纯化并且分离含有酪氨酸脂肪酸缩合物的级分，通过蒸发去除溶剂。（见“Basic Principles of Chromatography”中 Ullmann's Encyclopedia of Industrial Chemistry (2002) 通过 Wiley-VCH Verlag GmbH & Co.KGaA.)

[0775] 方法 2

[0776] 下文中，脂质酰基转移酶的转移酶活性在基于磷脂酰胆碱及 1- 酪氨酸的底物中根据以下方法测定。

[0777] 在 Wheaton 玻璃杯中称量 450mg 磷脂酰胆碱 (> 95% PC Avanti 产品号 441601) 及 50mg 1- 酪氨酸，并加入 15ml 50mM HEPES 缓冲液 pH 7。在搅拌下使脂质分散于缓冲液中。

[0778] 磁力搅拌同时将该底物加热到 35°C，并加入 0.25ml 10PLU/ml 的转移酶。

[0779] 0, 5, 10, 15, 25, 40 及 60 分钟的反应时间后取出 2ml 样品。

[0780] 立即加入 25  $\mu$ l 4M HCl 以酸化该游离脂肪酸并终止酶反应。加入 3.00ml 氯仿，在 Whirley 上剧烈振摇所述样品 30 秒。对该样品进行离心，并分离 2ml 氯仿相，通过 0.45- $\mu$  m 滤纸过滤到 10ml tared Dram 玻璃杯中。

[0781] 在氮气流条件下、60°C 蒸发氯仿，再次称量所述样品。提取的脂质通过 TLC 分析。

[0782] 实施例 12- 利用本发明的脂质酰基转移酶制备羧基酯（具体是乳酸酯）

[0783] 脂肪酸的羧基酯通常由反应脂肪酸和羧基酸之间的反应在高温利用无机盐或金属离子作为催化剂制备（见例如 Bailey' s Industrial Oil 及 Fat Products, Fifth edition, Volume 3. Edible Oil 及 Fat Products : Products 及 Application Technology, page 502-511.）。但是该方法的缺点在于会形成副反应以及有色副产物。

[0784] 本发明中，羧基酯脂肪酸利用卵磷脂作为脂肪酸供体及羧基酸（具体是乳酸酯）作为受体分子通过转移酶反应制备。

[0785] 方法 .

[0786] 50g 乳酸在搅拌下溶于 1000ml 水，在搅拌中将 200g 磷脂酰胆碱分散于水相，并加热到 40°C。

[0787] 将 pH 调节到 pH 7，并用 NaOH 或 HCl 维持该 pH。

[0788] 加入 50ml 来自杀鲑气单胞菌的脂质酰基转移酶，并在搅拌下使该反应在 40°C 继续。

[0789] 间隔规律的时间取样，并通过 TLC 以及 HPLC 分析。

[0790] 20h 反应时间后，该反应到达平衡并终止反应。

[0791] 乳酸酯，卵磷脂及溶血卵磷脂通过离心自该反应介质分离，所述离心根据标准方法（见“Centrifuges, Filtering”中 Ullmann' s Encyclopedia of Industrial Chemistry for example (2002) by Wiley-VCH Verlag GmbH & Co.KgaA）实施。乳酸酯进一步通过分子蒸馏来纯化，获得具有高纯度的乳酸酯脂肪酸。

[0792] 实施例 13- 利用本发明的脂质酰基转移酶制备柠檬酸酯

[0793] 基于磷脂酰胆碱柠檬酸作为底物的转移酶测定

[0794] 下文中，来自杀鲑气单胞菌的脂质酰基转移酶的转移酶活性在基于磷脂酰胆碱及 L- 酪氨酸的底物中根据以下方法测定。

[0795] 在 Wheaton 玻璃杯中称量 450mg 磷脂酰胆碱（> 95% PC Avanti 产品号 441601）及 50mg 柠檬酸，并加入 15ml 50mM HEPES 缓冲液 pH 7。在搅拌下使脂质分散于缓冲液中。

[0796] 磁力搅拌同时将该底物加热到 35°C，并加入 0.25ml 10 PLU/ml 来自杀鲑气单胞菌的转移酶。

[0797] 0, 5, 10, 15, 25, 40 及 60 分钟的反应时间后取出 2ml 样品。

[0798] 立即加入 25  $\mu$ l 4M HCl 以酸化该游离脂肪酸并终止酶反应。加入 3.00ml 氯仿，在 Whirley 上剧烈振摇所述样品 30 秒。对该样品进行离心，并分离 2ml 氯仿相，通过 0.45- $\mu$  m 滤纸过滤到 10ml tared Dram 玻璃杯中。

[0799] 在氮气流条件下、60°C 蒸发氯仿，再次称量所述样品。提取的脂质通过 TLC 分析。

[0800] 上面说明书中提及的所有公开的内容包含在本文中作为参考。本发明方法和体系的不同修饰和改变对于本领域普通技术人员来说是清楚的而不会偏离本发明的范围和精。尽管本发明根据优选的具体实施方案进行了描述，但本发明权利要求保护的范围不应该不适当地限于这些特殊的具体实施方案是可以理解的。事实上，对用于实现发明的所述模式的各种修饰对于生物化学领域和生物工程学领域或相关领域的普通技术人员来说是清楚明白的，意图包含在下面的权利要求的保护范围内。

[0801] 布达佩斯条约关于用于专利程序的国际认可的微生物保藏

[0802] 国际局表格

[0803] 丹尼斯科公司 (Danisco A/S)

[0804] Langebrogade 1

[0805] DK-1001 Copenhagen

[0806] 丹麦

[0807] 保藏人的姓名和地址

[0808]

<b>I. 微生物的鉴定</b>	
保藏人给出的鉴定参考 Escherichia coli TOP10pPet12aAhydro	由国际保藏单位给出的登记号码 NCIMB 41204
<b>II. 科学描述和/或建议的分类学名称</b>	
上述项 I 所说明的微生物被给出的 <input type="checkbox"/> 科学描述 <input checked="" type="checkbox"/> 建议的分类学名称 (用打叉表示选定)	
<b>III. 收到和接受</b>	
本国际保藏单位于 2003 年 12 月 22 日(即初始保藏日) <sup>1</sup> 收到并接受上述项 I 所说明的微生物	
<b>IV. 收到转移请求</b>	
本国际保藏单位于 _____(即初始保藏日)接受了上述项 I 所说明的微生物, 并且, 接受了于 _____(即收到转换请求的日期)将该原始保藏转为根据布达佩斯条约的保藏的请求	
<b>IV. 国际保藏单位</b>	
名称: NCIMB 地址: 23 St Machar Drive Aberdeen AB24.3RY Scotland, 英国	能代表国际保藏单位或被授权的官员的人的 签名:  日期: 2004 年 1 月 9 日

[0809] <sup>1</sup> 当适用于细则 6.4(d) 时, 所述的日期为国际保藏单位的资格被认定的日期。

[0810] 表格 sp/4( 只此一页 )

[0811] 布达佩斯条约关于用于专利程序的国际认可的微生物保藏

[0812] 国际局表格

[0813] 丹尼斯科公司 (Danisco A/S) 存活证明

[0814] Langebrogade 1 由下一页所指明的国际保

[0815] DK-1001 Copenhagen 藏单位根据细则 10.2 出具

[0816] 丹麦

[0817] 应接到此存活证明的

[0818] 单位的名称和地址

[0819]

<b>I. 保藏人</b>	<b>II. 微生物的鉴定</b>
名称：如上所述	由国际保藏单位给出的保藏号：
地址：	NCIMB 41204
	保藏日或转保藏日：2003 年 12 月 22 日
<b>II. 存活证明</b>	
对由上述项 II 所说明的微生物的存活性在 检验	
2. 在该日期，所述微生物是	
<input checked="" type="checkbox"/> <sup>3</sup> 存活的	
<input type="checkbox"/> <sup>3</sup> 不存活的	

[0820] 1 指的是初始保藏日，或者，当进行了新的保藏或保藏的转换时，指的是最相关的日期（新保藏日期或是转保藏日期）。

[0821] 2 对于那些法规 10.2(a) (ii) 和 (iii) 所代表的情况，指的是最近的存活性检验。

[0822] 3 用打叉表示选定。

<b>IV. 实施存活性检验的条件<sup>4</sup></b>	
<b>V. 国际保藏单位</b>	
[0823] 名称： NCIMB Ltd.	能代表国际保藏单位或被授权的官员的人的签名：
地址： 23 St Machar Drive Aberdeen AB24 3RY Scotland	日期：2004 年 1 月 9 日

[0824] 4 如果要求该信息并且如果检验为阴性，则填写此项。

[0825] 布达佩斯条约关于用于专利程序的国际认可的微生物保藏

[0826] 国际局表格

[0827] 丹尼斯科公司 (Danisco A/S)

[0828] Langebrogade 1

[0829] DK-1001 Copenhagen

[0830] 丹麦

[0831] 保藏人的姓名和地址

<b>I. 微生物的鉴定</b>	
保藏人给出的鉴定参考 <i>Escherichia coli</i> TOP10pPet12aAsalmo	由国际保藏单位给出的登记号码 NCIMB 41205
<b>II. 科学描述和/或建议的分类学名称</b>	
上述项 I 所说明的微生物被给出的 <input type="checkbox"/> 科学描述 <input checked="" type="checkbox"/> 建议的分类学名称 (用打叉表示选定)	
<b>III. 收到和接受</b>	
[0832]	本国际保藏单位于 2003 年 12 月 22 日(即初始保藏日)收到并接受上述项 I 所说明的微生物
<b>IV. 收到转移请求</b>	
本国际保藏单位于 _____(即初始保藏日)接受了上述项 I 所说明的微生物, 并且, 接受了于 _____(即收到转换请求的日期)将该原始保藏转为根据布达佩斯条约的保藏的请求	
<b>IV. 国际保藏单位</b>	
姓名: NCIMB 地址: 23 St Machar Drive Aberdeen AB24.3RY Scotland, 英国	能代表国际保藏单位或被授权的官员的人的 签名:  日期: 2004 年 1 月 9 日

[0833] 当适用于细则 6.4(d) 时, 所述的日期为国际保藏单位的资格被认定的日期。

[0834] 表格 sp/4( 只此一页 )

[0835] 布达佩斯条约关于用于专利程序的国际认可的微生物保藏

[0836] 国际局表格

[0837] 丹尼斯科公司 (Danisco A/S) 存活证明

[0838] Langebrogade 1 由下一页所指明的国际保

[0839] DK-1001 Copenhagen 藏单位根据细则 10.2 出具

[0840] 丹麦

[0841] 应接到此存活证明的

[0842] 单位的名称和地址

[0843]

<b>I. 保藏人</b>	<b>II. 微生物的鉴定</b>
名称：如上所述	由国际保藏单位给出的保藏号：
地址：	NCIMB 41205
	保藏日或转保藏日：2003 年 12 月 22 日
<b>III. 存活证明</b>	
对由上述项 II 所说明的微生物的存活性在	检验
	<sup>2</sup> . 在该日期，所述微生物是
<input checked="" type="checkbox"/> <sup>3</sup> 存活的	
<input type="checkbox"/> <sup>3</sup> 不存活的	

[0844] 1 指的是初始保藏日，或者，当进行了新的保藏或保藏的转换时，指的是最相关的日期（新保藏日期或是转保藏日期）。

[0845] 2 对于那些法规 10.2(a)(ii) 和 (iii) 所代表的情况，指的是最近的存活性检验。

[0846] 3 用打叉表示选定。

[0847]

<b>IV. 实施存活性检验的条件<sup>4</sup></b>	
<b>V. 国际保藏单位</b>	
姓名： NCIMB Ltd.	能代表国际保藏单位或被授权的官员的人的
地址： 23 St Machar Drive	签名：
Aberdeen	
AB24 3RY	日期：2004 年 1 月 9 日
Scotland	

[0848] 4 如果要求该信息并且如果检验为阴性，则填写此项。

[0001]

## 序列表

&lt;110&gt; 丹尼斯科公司 (Danisco A/S)

&lt;120&gt; 制备碳水化合物酯的方法

&lt;130&gt; p015575WO AAW

&lt;140&gt; PCT/IB2004/000575

&lt;141&gt; 2004-01-15

&lt;150&gt; GB0301121.0

&lt;151&gt; 2003-01-17

&lt;150&gt; GB0301122.8

&lt;151&gt; 2003-01-17

&lt;150&gt; GB0301117.8

&lt;151&gt; 2003-01-17

&lt;150&gt; GB0301120.2

&lt;151&gt; 2003-01-17

&lt;150&gt; GB0301119.4

&lt;151&gt; 2003-01-17

&lt;150&gt; GB0301118.6

&lt;151&gt; 2003-01-17

[0002]

<150> GB030330016.7

<151> 2003-12-24

<150> US60/489441

<151> 2003-07-23

<160> 54

<170> PatentIn version 3.0

<210> 1

<211> 361

<212> PRT

<213> 人工序列

<220>

<223> 共有序列

<400> 1

Ile Val Ala Phe Gly Asp Ser Leu Thr Asp Gly Glu Ala Tyr Tyr Gly  
1 5 10 15

Asp Ser Asp Gly Gly Gly Trp Gly Ala Gly Leu Ala Asp Arg Leu Thr  
20 25 30

Ala Leu Leu Arg Leu Arg Ala Arg Pro Arg Gly Val Asp Val Phe Asn  
35 40 45

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

Leu Val Ala Leu Leu Phe Leu Ala Gln Ser Leu Gly Leu Pro Asn Leu  
65 70 75 80

Pro Pro Tyr Leu Ser Gly Asp Phe Leu Arg Gly Ala Asn Phe Ala Ser  
85 90 95

Ala Gly Ala Thr Ile Leu Pro Thr Ser Gly Pro Phe Leu Ile Gln Val  
100 105 110

[0003]

Gln Phe Lys Asp Phe Lys Ser Gln Val Leu Glu Leu Arg Gln Ala Leu  
 115 120 125  
 Gly Leu Leu Gln Glu Leu Leu Arg Leu Leu Pro Val Leu Asp Ala Lys  
 130 135 140  
 Ser Pro Asp Leu Val Thr Ile Met Ile Gly Thr Asn Asp Leu Ile Thr  
 145 150 155 160  
 Ser Ala Phe Phe Gly Pro Lys Ser Thr Glu Ser Asp Arg Asn Val Ser  
 165 170 175  
 Val Pro Glu Phe Lys Asp Asn Leu Arg Gln Leu Ile Lys Arg Leu Arg  
 180 185 190  
 Ser Asn Asn Gly Ala Arg Ile Ile Val Leu Ile Thr Leu Val Ile Leu  
 195 200 205  
 Asn Leu Gly Pro Leu Gly Cys Leu Pro Leu Lys Leu Ala Leu Ala Leu  
 210 215 220  
 Ala Ser Ser Lys Asn Val Asp Ala Ser Gly Cys Leu Glu Arg Leu Asn  
 225 230 235 240  
 Glu Ala Val Ala Asp Phe Asn Glu Ala Leu Arg Glu Leu Ala Ile Ser  
 245 250 255  
 Lys Leu Glu Asp Gln Leu Arg Lys Asp Gly Leu Pro Asp Val Lys Gly  
 260 265 270  
 Ala Asp Val Pro Tyr Val Asp Leu Tyr Ser Ile Phe Gln Asp Leu Asp  
 275 280 285  
 Gly Ile Gln Asn Pro Ser Ala Tyr Val Tyr Gly Phe Glu Thr Thr Lys  
 290 295 300  
 Ala Cys Cys Gly Tyr Gly Gly Arg Tyr Asn Tyr Asn Arg Val Cys Gly  
 305 310 315 320  
 Asn Ala Gly Leu Cys Asn Val Thr Ala Lys Ala Cys Asn Pro Ser Ser  
 325 330 335  
 Tyr Leu Leu Ser Phe Leu Phe Trp Asp Gly Phe His Pro Ser Glu Lys  
 340 345 350  
 Gly Tyr Lys Ala Val Ala Glu Ala Leu  
 355 360

<210> 2

<211> 335

<212> PRT

<213> 嗜水气单胞菌(Aeromonas hydrophila)

[0004]

&lt;400&gt; 2

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

[0005]

Met Ala Ala Arg Ser Ala Ser Thr Leu Asn Cys Glu Gly Lys Met Phe  
 290 295 300

Trp Asp Gln Val His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu  
 305 310 315 320

Pro Ala Ala Thr Phe Ile Glu Ser Gln Tyr Glu Phe Leu Ala His  
 325 330 335

<210> 3

<211> 336

<212> PRT

<213> 杀鲑气单胞菌(*Aeromonas salmonicida*)

<400> 3

Met Lys Lys Trp Phe Val Cys Leu Leu Gly Leu Ile Ala Leu Thr Val  
 1 5 10 15

Gln Ala Ala Asp Thr Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly  
 20 25 30

Asp Ser Leu Ser Asp Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr  
 35 40 45

Leu Pro Ser Ser Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro  
 50 55 60

Val Trp Leu Glu Gln Leu Thr Lys Gln Phe Pro Gly Leu Thr Ile Ala  
 65 70 75 80

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

Trp Asn Pro Lys Tyr Gln Val Tyr Asn Asn Leu Asp Tyr Glu Val Thr  
 100 105 110

Gln Phe Leu Gln Lys Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu  
 115 120 125

Trp Val Gly Ala Asn Asp Tyr Leu Ala Tyr Gly Trp Asn Thr Glu Gln  
 130 135 140

Asp Ala Lys Arg Val Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met  
 145 150 155 160

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

Gly Gln Asn Pro Ser Ala Arg Ser Gln Lys Val Val Glu Ala Val Ser  
 180 185 190

[0006]

His Val Ser Ala Tyr His Asn Lys Leu Leu Leu Asn Leu Ala Arg Gln  
 195 200 205  
 Leu Ala Pro Thr Gly Met Val Lys Leu Phe Glu Ile Asp Lys Gln Phe  
 210 215 220  
 Ala Glu Met Leu Arg Asp Pro Gln Asn Phe Gly Leu Ser Asp Val Glu  
 225 230 235 240  
 Asn Pro Cys Tyr Asp Gly Gly Tyr Val Trp Lys Pro Phe Ala Thr Arg  
 245 250 255  
 Ser Val Ser Thr Asp Arg Gln Leu Ser Ala Phe Ser Pro Gln Glu Arg  
 260 265 270  
 Leu Ala Ile Ala Gly Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro  
 275 280 285  
 Met Ala Arg Arg Ser Ala Ser Pro Leu Asn Cys Glu Gly Lys Met Phe  
 290 295 300  
 Trp Asp Gln Val His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu  
 305 310 315 320  
 Arg Ala Ala Thr Phe Ile Glu Thr Gln Tyr Glu Phe Leu Ala His Gly  
 325 330 335

<210> 4

<211> 295

<212> PRT

<213> 天蓝色链霉菌 (Streptomyces coelicolor)

<400> 4

Met Pro Lys Pro Ala Leu Arg Arg Val Met Thr Ala Thr Val Ala Ala  
 1 5 10 15  
 Val Gly Thr Leu Ala Leu Gly Leu Thr Asp Ala Thr Ala His Ala Ala  
 20 25 30  
 Pro Ala Gln Ala Thr Pro Thr Leu Asp Tyr Val Ala Leu Gly Asp Ser  
 35 40 45  
 Tyr Ser Ala Gly Ser Gly Val Leu Pro Val Asp Pro Ala Asn Leu Leu  
 50 55 60  
 Cys Leu Arg Ser Thr Ala Asn Tyr Pro His Val Ile Ala Asp Thr Thr  
 65 70 75 80  
 Gly Ala Arg Leu Thr Asp Val Thr Cys Gly Ala Ala Gln Thr Ala Asp  
 85 90 95

[0007]

Phe Thr Arg Ala Gln Tyr Pro Gly Val Ala Pro Gln Leu Asp Ala Leu  
 100 105 110  
 Gly Thr Gly Thr Asp Leu Val Thr Leu Thr Ile Gly Gly Asn Asp Asn  
 115 120 125  
 Ser Thr Phe Ile Asn Ala Ile Thr Ala Cys Gly Thr Ala Gly Val Leu  
 130 135 140  
 Ser Gly Gly Lys Gly Ser Pro Cys Lys Asp Arg His Gly Thr Ser Phe  
 145 150 155 160  
 Asp Asp Glu Ile Glu Ala Asn Thr Tyr Pro Ala Leu Lys Glu Ala Leu  
 165 170 175  
 Leu Gly Val Arg Ala Arg Ala Pro His Ala Arg Val Ala Ala Leu Gly  
 180 185 190  
 Tyr Pro Trp Ile Thr Pro Ala Thr Ala Asp Pro Ser Cys Phe Leu Lys  
 195 200 205  
 Leu Pro Leu Ala Ala Gly Asp Val Pro Tyr Leu Arg Ala Ile Gln Ala  
 210 215 220  
 His Leu Asn Asp Ala Val Arg Arg Ala Ala Glu Glu Thr Gly Ala Thr  
 225 230 235 240  
 Tyr Val Asp Phe Ser Gly Val Ser Asp Gly His Asp Ala Cys Glu Ala  
 245 250 255  
 Pro Gly Thr Arg Trp Ile Glu Pro Leu Leu Phe Gly His Ser Leu Val  
 260 265 270  
 Pro Val His Pro Asn Ala Leu Gly Glu Arg Arg Met Ala Glu His Thr  
 275 280 285  
 Met Asp Val Leu Gly Leu Asp  
 290 295

<210> 5

<211> 295

<212> PRT

<213> 天蓝色链霉菌

<400> 5

Met Pro Lys Pro Ala Leu Arg Arg Val Met Thr Ala Thr Val Ala Ala  
 1 5 10 15  
 Val Gly Thr Leu Ala Leu Gly Leu Thr Asp Ala Thr Ala His Ala Ala  
 20 25 30

[0008]

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

<210> 6

<211> 238

<212> PRT

[0009]

<213> 酿酒酵母 (*Saccharomyces cerevisiae*)

<400> 6

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

<210> 7

<211> 1005

<212> DNA

[0010]

## &lt;213&gt; 嗜水气单胞菌

## &lt;400&gt; 7

atgaaaaaat ggtttgtgtg tttattggga ttggtcgcgc tgacagtcca ggcagccgac 60  
 agccgtcccg cttctcccg gatcgtgatg ttggcgaca gcctctccga taccggcaag 120  
 atgtacagca agatgcgcgg ttacctcccc tccagcccc cctactatga gggccgcttc 180  
 tccaacgggc ccgtctggct ggagcagctg accaacgagt tcccggcct gaccatagcc 240  
 aacgaggcgg aaggcggacc gaccgccgtg gcttacaaca agatctcctg gaatccaag 300  
 tatcaggtea tcaacaacct ggactacgag gtcacccagt tctgcaaaa agacagcttc 360  
 aagccggacg atctgggatg cctctgggtc ggcgccaacg actatctggc ctatggctgg 420  
 aacacagagc aggatgcca gcggtgcgc gacgccatca gcgatgcggc caaccgcatg 480  
 gtgctgaacg gcgccaagga gatactgctg ttcaacctgc cggatctggg ccagaacccc 540  
 tcggcccgca gccagaaggt ggtcggaggc gccagccatg tctccgcta ccacaaccag 600  
 ctgctgctga acctggcacg ccagctggct cccaccggca tggatgaagct gttcgagatc 660  
 gacaagcagt ttgccgagat gctgcgtgat ccgcagaact tcggcctgag cgaccagagg 720  
 aacgctget acggtggcag ctatgtatgg aagccgtttg cctccgcag cgccagcacc 780  
 gacagccagc tctccgctt caaccgcag gagegcctcg ccatgcceg caaccgctg 840  
 ctggcccagg ccgtgccag cccatggct gccgcagcg ccagaccct caactgtgag 900  
 ggcaagatgt tctgggatca ggtccacccc accactgtcg tgcacgcgc cctgagcgag 960  
 cccgccgca cttcatcga gagccagtac gagttcctcg cccac 1005

## &lt;210&gt; 8

## &lt;211&gt; 1011

## &lt;212&gt; DNA

## &lt;213&gt; 杀鲑气单胞菌

## &lt;400&gt; 8

atgaaaaaat ggtttgtttg tttattgggg ttgatcgcgc tgacagtcca ggcagccgac 60  
 actgccccg cttctcccg gatcgtgatg ttggcgaca gcctctccga taccggcaaa 120  
 atgtacagca agatgcgcgg ttacctcccc tccagcccgc cctactatga gggccgttc 180

[0011]

tccaacggac ccgtctgget ggagcagctg accaagcagt tcccgggtct gaccatcgcc	240
aacgaagcgg aaggcgggtgc cactgccgtg gcttacaaca agatctcctg gaatcccaag	300
tatcaggtct acaacaacct ggactacgag gtcacccagt tcttgcaaaa agacagcttc	360
aagccggacg atctggtgat cctctgggtc ggtgccaatg actatctggc atatggctgg	420
aatacggagc aggatgcaa gcgagttcgc gatgccatca gcgatgcggc caaccgcatg	480
gtactgaacg gtgccaagca gatactgctg ttcaacctgc cggatctggg ccagaacctg	540
tcagcccga gtcagaaggt ggtcgaggcg gtcagccatg tctccgcta tcacaacaag	600
ctgctgctga acctggcagc ccagctggcc cccaccggca tggtaaagct gttcgagatc	660
gacaagcaat ttgccgagat gctgcgtgat ccgcagaact tcggcctgag cgacgctgag	720
aacccctgct acgacggcgg ctatgtgtgg aagccgtttg ccaccgcag cgtcagcacc	780
gaccgccagc tctccgctt cagtcgcag gaacgcctcg ccatcgccgg caaccgctg	840
ctggcacagg ccgttgccag tctatggcc cggcgcagcg ccagccccct caactgtgag	900
ggcaagatgt tctgggatca ggtacaccgc accactgtcg tgcacgcagc cctgagcgag	960
cgcgccgcca cttcatga gaccagctac gagttcctcg cccacggatg a	1011

<210> 9

<211> 888

<212> DNA

<213> 天蓝色链霉菌

<400> 9

atgccgaagc ctgcccttcg ccgtgtcatg accgcgacag tcgccgccgt cggcacgctc	60
gccctcggcc tcaccgacgc caccgcccac gccgcgcccg ccaggccac tccgacctg	120
gactacgtcg ccctcggcga cagctacagc gccggctccg gcttctgcc cgtcgacccc	180
gccaacctgc tctgtctgcg ctcgacggcc aactaccccc acgtcatcgc ggacacgacg	240
ggcggcccgc tcacggacgt cacctcggcg gccgcgcaga ccgccgactt cacgcgggcc	300
cagtaccggg gcgtcgcacc ccagttggac gcgctcggca ccggcacgga cctggtaacg	360
ctcaccatcg gcggcaacga caacagcacc ttcatcaacg ccatacggc ctgcggcag	420
gcgggtgtcc tcagcggcgg caagggcagc ccctgcaagg acaggcacgg cacctcttc	480

[0012]

gacgacgaga tcgaggccaa cacgtacccc gcgctcaagg aggcgctgct cggcgctcgc 540  
gccagggctc cccacgccag ggtggcggct ctcggctacc cgtggatcac cccggccacc 600  
gccgacctgt cctgcttctt gaagctcccc ctcgccgccc gtgacgtgcc ctacctgcgg 660  
gccatccagg cacacctcaa cgacgcggtc cggcgggccc cggaggagac cggagccacc 720  
tacgtggact tctccggggg gtccgacggc cacgacgcct gcgaggcccc cggcaccgcg 780  
tggatcgaac cgctgctctt cgggcacagc ctcgttcccc tccaccccaa cgccttgggc 840  
gagcggcgca tggccgagca cacgatggac gtcctcggcc tggactga 888

<210> 10

<211> 888

<212> DNA

<213> 天蓝色链霉菌

<400> 10

tcagtccagg ccgaggacgt ccatcgtgtg ctcggccatg cggcgctcgc ccaggcgctt 60  
gggggtggacg ggaacgaggc tgtgcccga gacgacgggt tcgatccagc ggggtgccggg 120  
ggcctcgcag gcgtcgtggc cgtcggacac cccggagaag tccacgtagg tggctccggt 180  
ctcctcggcg gcccgccgga ccgctcgtt gaggtgtgcc tggatggccc gcaggtaggg 240  
cacgtcaccg gcggcgaggg ggagcttcag gaagcaggac gggtcggcgg tggccggggg 300  
gatccacggg tagccgagag ccgccaccct ggctggggga gccctggcgc ggacgccgag 360  
cagcgcctcc ttgagcgcgg ggtacgtgtt ggctcgtatc tcgtcgtcga aggaggtgcc 420  
gtgcctgtcc ttgcaggggc tgcccttgcc gccgctgagg acaccgccg tgccgcaggc 480  
cgtgatggcg ttgatgaagg tgctgttgtc gttgccgccc atggtgagcg tgaccaggtc 540  
cgtgccggtg ccgagcgcgt ccaactgggg tgcgacgccc gggactggg cccgcgtgaa 600  
gtcggcggtc tgcgcggcgc cgcagggtac gtccgtgagg cgggcgccc tcgtgtccgc 660  
gatgacgtgg gggtagttgg ccgtcagcgc cagacagagc aggttggcgg ggtcgcaggg 720  
caggacccc gagccggcgc ttagctgtc gccgagggcg acgtagtcca gggtcggagt 780  
ggcctgggcg ggccggcgt gggcgggtggc gtcggtgagg ccgagggcga gcgtgccgac 840  
ggcggcgact gtcgcggtca tgacacggcg aagggcaggc ttcggcat 888

[0013]

<210> 11  
 <211> 717  
 <212> DNA  
 <213> 酿酒酵母

<400> 11  
 atggattacg agaagtttct gttatttggg gattccatta ctgaatttgc ttttaatact 60  
 aggcccattg aagatggcaa agatcagtat gctcttggag ccgcattagt caacgaatat 120  
 acgagaaaaa tggatattct tcaaagaggg ttcaaagggt acacttctag atgggcgttg 180  
 aaaatacttc ctgagatttt aaagcatgaa tccaatattg tcatggccac aatatttttg 240  
 ggtgccaacg atgcatgctc agcagggtccc caaagtgtcc ccttccccga atttatcgat 300  
 aatattcgtc aatgggtatc tttgatgaag tcttaccata tccgtcctat tataatagga 360  
 ccggggctag tagatagaga gaagtgggaa aaagaaaaat ctgaagaaat agctctcgga 420  
 tacttccgta ccaacgagaa ctttgccatt tattccgatg ccttagcaaa actagccaat 480  
 gaggaaaaag ttcccttctg ggctttgaat aaggcgtttc aacaggaagg tggatgatgct 540  
 tggcaacaac tgctaacaga tggactgcac ttttccggaa aagggtacaa aatttttcat 600  
 gacgaattat tgaaggtcat tgagacattc taccaccaat atcatcccaa aaacatgcag 660  
 tacaactga aagattggag agatgtgcta gatgatggat ctaacataat gtcttga 717

<210> 12  
 <211> 347  
 <212> PRT  
 <213> 雷尔氏菌(Ralstonia sp.)

<400> 12  
 Met Asn Leu Arg Gln Trp Met Gly Ala Ala Thr Ala Ala Leu Ala Leu  
 1 5 10 15  
 Gly Leu Ala Ala Cys Gly Gly Gly Gly Thr Asp Gln Ser Gly Asn Pro  
 20 25 30  
 Asn Val Ala Lys Val Gln Arg Met Val Val Phe Gly Asp Ser Leu Ser  
 35 40 45

[0014]

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

[0015]

<210> 13

<211> 1044

<212> DNA

<213> 雷尔氏菌

<400> 13

```

atgaacctgc gtcaatggat gggcgccgcc acggctgccc ttgccttggg cttggccgcg      60
tgcgggggcg gtgggaccga ccagagcggc aatcccaatg tcgccaaggt gcagcgcgatg    120
gtggtgttcg gcgacagcct gagcgatata ggcacctaca cccccgtcgc gcaggcgggtg    180
ggcggcggca agttcaccac caaccgggc ccgatctggg ccgagaccgt ggccgcgcaa     240
ctgggcgtga cgctcacgcc ggcggtgatg ggctacgcca cctccgtgca gaattgcccc     300
aaggccggct gcttcgacta tgcgcagggc ggctcgcgcg tgaccgatcc gaacggcatc     360
ggccacaacg gcggcgcggg ggcgctgacc taccggttc agcagcagct cgccaacttc     420
tacgcgccca gcaacaacac attcaacggc aataacgatg tcgtcttcgt gctggccggc     480
agcaacgaca tttcttctg gaccactgcg gcggccacca gcggctccgg cgtgacgccc     540
gccattgcca cggcccaggt gcagcaggcc gcgacggacc tggtcggcta tgtcaaggac     600
atgatcgcca agggtgcgac gcaggctctac gtgttcaacc tgcccagacag cagcctgacg     660
ccggacggcg tggcaagcgg cacgaccggc caggcgtgc tgcaacgcgt ggtgggcaacg     720
ttcaacacga cgctgcaaag cgggctggcc ggcacctcgg cgcgcatcat cgacttcaac     780
gcacaactga ccgcggcgat ccagaatggc gcctcgttcg gcttcgcca caccagcgcc     840
cgggcctgcg acgccaccaa gatcaatgcc ctggtgccga gcgccggcgg cagctcgctg     900
ttctgctcgg ccaacacgct ggtggcttcc ggtgcggacc agagctacct gttcgccgac     960
ggcgtgcacc cgaccacggc cggccatcgc ctgatcgcca gcaacgtgct ggccgcgctg    1020
ctggcggata acgtcgcgca ctga                                             1044

```

<210> 14

<211> 261

<212> PRT

<213> 天蓝色链霉菌

[0016]

&lt;400&gt; 14

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

&lt;210&gt; 15

&lt;211&gt; 786

[0017]

<212> DNA

<213> 天蓝色链霉菌

<400> 15

```

gtgatcgggt cgtacgtggc ggtgggggac agcttcaccg agggcgtcgg cgacccccggc      60
cccgcagggg cgttcgtcgg ctgggcccac cggetcgccg tactgctcgc ggaccggcgc      120
cccgagggcg acttcacgta cacgaacctc gccgtgcgcg gcaggctcct cgaccagatc      180
gtggcggaac aggtcccgcg ggtcgtcggg ctcgcgcccg acctcgtctc gttcgcggcg      240
ggcggcaacg acatcatccg gcccggcacc gatcccgcag aggtcgccga gcggttcgag      300
ctggcggtgg ccgcgctgac cgccgcgcc ggaaccgtcc tggtgaccac cgggttcgac      360
acccgggggg tgcccgtcct caagcacctg cgcggcaaga tcgccacgta caacgggcac      420
gtccgcgcca tcgccgaccg ctacggctgc ccggtgctcg acctgtggte gctgcggagc      480
gtccaggacc gcagggcgtg ggacgccgac cggtgcacc tgtcgccgga ggggcacacc      540
cgggtggcgc tgcgcgcggg gcaggccctg ggctgcgcg tcccggccga ccctgaccag      600
ccctggccgc ccctgccgcc gcgcggcacg ctcgacgtcc ggcgcgacga cgtgcactgg      660
gcgcgcgagt acctggtgcc gtggatcggg cgccggtgce ggggcgagtc gtcgggcgac      720
cacgtgacgg ccaaggggac gctgtcgccg gacgcatca agacgcgat cgccgcggtg      780
gcctga                                             786

```

<210> 16

<211> 260

<212> PRT

<213> 天蓝色链霉菌

<400> 16

```

Met Gln Thr Asn Pro Ala Tyr Thr Ser Leu Val Ala Val Gly Asp Ser
1          5          10          15
Phe Thr Glu Gly Met Ser Asp Leu Leu Pro Asp Gly Ser Tyr Arg Gly
20          25          30
Trp Ala Asp Leu Leu Ala Thr Arg Met Ala Ala Arg Ser Pro Gly Phe
35          40          45

```

[0018]

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

<210> 17

<211> 783

<212> DNA

<213> 天蓝色链霉菌

<400> 17

atgcagacga accccgcgta caccagtctc gtcgccgtcg ggaactcctt caccgagggc 60

atgtcggacc tgctgccca cgctcctac cgtggctggg ccgacctct cgccaccgg 120

[0019]

atggcgcccc gctccccggg ctteccgttac gccaacctgg cgggtgcgcgg gaagctgac 180  
 ggacagatcg tcgacgagca ggtggacgtg gccgccgcca tgggagccga cgtgatcacg 240  
 ctggtcggcg ggctcaacga cacgctgcgg cccaagtgcg acatggcccc ggtgcggggac 300  
 ctgctgacct aggccgtgga acggctcgcc ccgcaactgcg agcagctggt gctgatgcgc 360  
 agtccccgtc gccagggtcc ggtgctggag cgcttccggc cccgcatgga ggccctgttc 420  
 gccgtgatcg acgacctggc cgggcggcac ggcgccgtgg tcgtcgacct gtacggggcc 480  
 cagtcgctgg ccgaccctcg gatgtgggac gtggaccggc tgcacctgac cgccgagggc 540  
 caccgccggg tcgcgagggc ggtgtggcag tcgctcggcc acgagccccga ggacccccgag 600  
 tggcacgcgc cgatccccgc gacgcccgcc ccgggggtggg tgacgcgcag gaccgcggac 660  
 gtccggttcg cccggcagca cctgctgccc tggataggcc gcaggctgac cgggcgctcg 720  
 tccggggacg gctgcccgc caagcgcccc gacctgctgc cctacgagga ccccgcacgg 780  
 tga 783

<210> 18

<211> 454

<212> PRT

<213> 天蓝色链霉菌

<400> 18

Met Thr Arg Gly Arg Asp Gly Gly Ala Gly Ala Pro Pro Thr Lys His  
 1                   5                   10                   15  
 Arg Ala Leu Leu Ala Ala Ile Val Thr Leu Ile Val Ala Ile Ser Ala  
           20                   25                   30  
 Ala Ile Tyr Ala Gly Ala Ser Ala Asp Asp Gly Ser Arg Asp His Ala  
           35                   40                   45  
 Leu Gln Ala Gly Gly Arg Leu Pro Arg Gly Asp Ala Ala Pro Ala Ser  
           50                   55                   60  
 Thr Gly Ala Trp Val Gly Ala Trp Ala Thr Ala Pro Ala Ala Ala Glu  
 65                   70                   75                   80  
 Pro Gly Thr Glu Thr Thr Gly Leu Ala Gly Arg Ser Val Arg Asn Val  
           85                   90                   95  
 Val His Thr Ser Val Gly Gly Thr Gly Ala Arg Ile Thr Leu Ser Asn

[0020]

100					105					110					
Leu	Tyr	Gly	Gln	Ser	Pro	Leu	Thr	Val	Thr	His	Ala	Ser	Ile	Ala	Leu
		115					120					125			
Ala	Ala	Gly	Pro	Asp	Thr	Ala	Ala	Ala	Ile	Ala	Asp	Thr	Met	Arg	Arg
	130					135					140				
Leu	Thr	Phe	Gly	Gly	Ser	Ala	Arg	Val	Ile	Ile	Pro	Ala	Gly	Gly	Gln
145					150					155					160
Val	Met	Ser	Asp	Thr	Ala	Arg	Leu	Ala	Ile	Pro	Tyr	Gly	Ala	Asn	Val
				165					170					175	
Leu	Val	Thr	Thr	Tyr	Ser	Pro	Ile	Pro	Ser	Gly	Pro	Val	Thr	Tyr	His
			180					185					190		
Pro	Gln	Ala	Arg	Gln	Thr	Ser	Tyr	Leu	Ala	Asp	Gly	Asp	Arg	Thr	Ala
		195					200					205			
Asp	Val	Thr	Ala	Val	Ala	Tyr	Thr	Thr	Pro	Thr	Pro	Tyr	Trp	Arg	Tyr
	210					215					220				
Leu	Thr	Ala	Leu	Asp	Val	Leu	Ser	His	Glu	Ala	Asp	Gly	Thr	Val	Val
225					230					235					240
Ala	Phe	Gly	Asp	Ser	Ile	Thr	Asp	Gly	Ala	Arg	Ser	Gln	Ser	Asp	Ala
				245					250					255	
Asn	His	Arg	Trp	Thr	Asp	Val	Leu	Ala	Ala	Arg	Leu	His	Glu	Ala	Ala
			260				265						270		
Gly	Asp	Gly	Arg	Asp	Thr	Pro	Arg	Tyr	Ser	Val	Val	Asn	Glu	Gly	Ile
		275					280					285			
Ser	Gly	Asn	Arg	Leu	Leu	Thr	Ser	Arg	Pro	Gly	Arg	Pro	Ala	Asp	Asn
	290					295					300				
Pro	Ser	Gly	Leu	Ser	Arg	Phe	Gln	Arg	Asp	Val	Leu	Glu	Arg	Thr	Asn
305					310					315					320
Val	Lys	Ala	Val	Val	Val	Val	Leu	Gly	Val	Asn	Asp	Val	Leu	Asn	Ser
				325					330					335	
Pro	Glu	Leu	Ala	Asp	Arg	Asp	Ala	Ile	Leu	Thr	Gly	Leu	Arg	Thr	Leu
			340					345					350		
Val	Asp	Arg	Ala	His	Ala	Arg	Gly	Leu	Arg	Val	Val	Gly	Ala	Thr	Ile
			355				360					365			
Thr	Pro	Phe	Gly	Gly	Tyr	Gly	Gly	Tyr	Thr	Glu	Ala	Arg	Glu	Thr	Met
	370					375					380				
Arg	Gln	Glu	Val	Asn	Glu	Glu	Ile	Arg	Ser	Gly	Arg	Val	Phe	Asp	Thr
385					390					395					400
Val	Val	Asp	Phe	Asp	Lys	Ala	Leu	Arg	Asp	Pro	Tyr	Asp	Pro	Arg	Arg

[0021]

	405	410	415
Met Arg Ser	Asp Tyr Asp Ser Gly	Asp His Leu His Pro	Gly Asp Lys
	420	425	430
Gly Tyr Ala	Arg Met Gly Ala Val Ile Asp Leu Ala Ala	Leu Lys Gly	
	435	440	445
Ala Ala Pro	Val Lys Ala		
	450		

<210> 19

<211> 1365

<212> DNA

<213> 天蓝色链霉菌

<400> 19

```

atgacctggg gtcgtgacgg ggggtgcgggg ggcccccca ccaagcaccg tgccctgctc      60
gcggcgatcg tcacctgat agtggcgatc tccgcggcca tatacgccgg agcgtccgcg      120
gacgacggca gcagggacca cgcgctgcag gccggaggcc gtctcccacg aggagacgcc      180
gccccgcgt ccaccggtgc ctgggtgggc gcctgggcca ccgcaccggc cgcggccgag      240
ccgggcaccg agacgaccgg cctggcgggc cgctccgtgc gcaacgtcgt gcacacctcg      300
gtcggcggca ccggcgcgcg gatcacctc tcgaacctgt acgggcagtc gccgctgacc      360
gtcacacacg cctcgatcgc cctggccgcc gggccccaca ccgccccgc gatcgccgac      420
accatgcgcc ggctcacctt cggcggcagc gcccggtga tcatcccggc gggcggccag      480
gtgatgagcg acaccgccg cctcgccatc ccctacgggg cgaacgtcct ggtcaccacg      540
tactccccca tcccgccgg gccggtgacc taccatcgc aggcccgca gaccagctac      600
ctggccgacg gcgaccgcac ggcggacgtc accgccgtcg cgtacaccac ccccacgccc      660
tactggcgct acctgaccgc cctcgacgtg ctgagccacg aggccgacgg cacggtcgtg      720
gcgttcggcg actccatcac cgacggcgcc cgctcgcaga gcgacgcaa ccaccgctgg      780
accgacgtcc tcgccgcacg cctgcacgag gcggcgggcg acggccggga cacgccccgc      840
tacagcgtcg tcaacgaggg catcagcggc aaccggctcc tgaccagcag gccggggcgg      900
ccggccgaca acccgagcgg actgagccgg ttccagcggg acgtgctgga acgcaccaac      960
gtcaaggccg tcgtcgtcgt cctcgcgctc aacgacgtcc tgaacagccc ggaactcgcc     1020
gaccgcgacg ccatactgac cgccctgcgc accctcgtcg accgggcgca cccccgggga     1080
    
```

[0022]

ctgcgggtcg tcggcgccac gatcacgccg ttcggcggtt acggcggcta caccgaggcc 1140  
 cgcgagacga tgcggcagga ggtcaacgag gagatccgct ccggccgggt cttcgacacg 1200  
 gtcgtcgact tcgacaagge cctgcgcgac ccgtacgacc cgcgccggat gcgctccgac 1260  
 tacgacagcg gcgaccacct gcaccccggc gacaaggggt acgcgcgcat gggcgcggtc 1320  
 atcgacctgg ccgcgctgaa gggcgcggcg ccggtcaagg cgtag 1365

<210> 20

<211> 340

<212> PRT

<213> 天蓝色链霉菌

<400> 20

Met Thr Ser Met Ser Arg Ala Arg Val Ala Arg Arg Ile Ala Ala Gly  
 1 5 10 15  
 Ala Ala Tyr Gly Gly Gly Gly Ile Gly Leu Ala Gly Ala Ala Ala Val  
 20 25 30  
 Gly Leu Val Val Ala Glu Val Gln Leu Ala Arg Arg Arg Val Gly Val  
 35 40 45  
 Gly Thr Pro Thr Arg Val Pro Asn Ala Gln Gly Leu Tyr Gly Gly Thr  
 50 55 60  
 Leu Pro Thr Ala Gly Asp Pro Pro Leu Arg Leu Met Met Leu Gly Asp  
 65 70 75 80  
 Ser Thr Ala Ala Gly Gln Gly Val His Arg Ala Gly Gln Thr Pro Gly  
 85 90 95  
 Ala Leu Leu Ala Ser Gly Leu Ala Ala Val Ala Glu Arg Pro Val Arg  
 100 105 110  
 Leu Gly Ser Val Ala Gln Pro Gly Ala Cys Ser Asp Asp Leu Asp Arg  
 115 120 125  
 Gln Val Ala Leu Val Leu Ala Glu Pro Asp Arg Val Pro Asp Ile Cys  
 130 135 140  
 Val Ile Met Val Gly Ala Asn Asp Val Thr His Arg Met Pro Ala Thr  
 145 150 155 160  
 Arg Ser Val Arg His Leu Ser Ser Ala Val Arg Arg Leu Arg Thr Ala  
 165 170 175

[0023]

Gly Ala Glu Val Val Val Gly Thr Cys Pro Asp Leu Gly Thr Ile Glu  
 180 185 190

Arg Val Arg Gln Pro Leu Arg Trp Leu Ala Arg Arg Ala Ser Arg Gln  
 195 200 205

Leu Ala Ala Ala Gln Thr Ile Gly Ala Val Glu Gln Gly Gly Arg Thr  
 210 215 220

Val Ser Leu Gly Asp Leu Leu Gly Pro Glu Phe Ala Gln Asn Pro Arg  
 225 230 235 240

Glu Leu Phe Gly Pro Asp Asn Tyr His Pro Ser Ala Glu Gly Tyr Ala  
 245 250 255

Thr Ala Ala Met Ala Val Leu Pro Ser Val Cys Ala Ala Leu Gly Leu  
 260 265 270

Trp Pro Ala Asp Glu Glu His Pro Asp Ala Leu Arg Arg Glu Gly Phe  
 275 280 285

Leu Pro Val Ala Arg Ala Ala Ala Glu Ala Ala Ser Glu Ala Gly Thr  
 290 295 300

Glu Val Ala Ala Ala Met Pro Thr Gly Pro Arg Gly Pro Trp Ala Leu  
 305 310 315 320

Leu Lys Arg Arg Arg Arg Arg Arg Val Ser Glu Ala Glu Pro Ser Ser  
 325 330 335

Pro Ser Gly Val  
 340

<210> 21

<211> 1023

<212> DNA

<213> 天蓝色链霉菌

<400> 21

atgacgagca tgtcgagggc gaggggtggcg cggcggatcg cggccggcgc ggcgtacggc 60

ggcggcggca tcggcctggc gggagcggcg gcggtcggtc tgggtggtggc cgaggtgcag 120

ctggccagac gcagggtggg ggtgggcacg ccgaccggg tgccgaacgc gcagggactg 180

tacggcggca ccctgcccac ggccggcgac ccgccgtgc ggctgatgat gctgggcgac 240

tccacggccg ccgggcaggg cgtgcaccgg gccgggcaga cgccgggcgc gctgctggcg 300

tccgggctcg cggcgggtggc ggagcggccg gtgcggctgg ggtcggtcgc ccagccgggg 360

gcgtgctcgg acgacctgga ccggcaggtg gcgctggtgc tcgccgagcc ggaccgggtg 420

[0024]

cccgacatct gcgtgatcat ggtcggcgcc aacgacgtca cccaccggat gccggcgacc 480  
 cgctcgggtgc ggcacctgtc ctcggcggta cggcggctgc gcacggccgg tgcggaggtg 540  
 gtggtcggca cctgtccgga cctgggcacg atcgagcggg tgcggcagcc gctgcgctgg 600  
 ctggccccggc gggcctcacg gcagctcgcg gcggcacaga ccatcggcgc cgtcgagcag 660  
 ggcggggcgca cgggtgtcgt gggcgacctg ctgggtccgg agttcgcgca gaaccgcgg 720  
 gagctcttcg gccccgacaa ctaccacccc tccgccgagg ggtacgccac ggccgcgatg 780  
 gcggtactgc cctcgggtgtg cgccgcgctc ggccctgtggc cggccgacga ggagcacccg 840  
 gacgcgctgc gccgcgaggg ctctctgccg gtggcgcgcg cggcggcgga ggcggcgtcc 900  
 gaggcgggta cggaggtcgc cgccgccatg cctacggggc ctcggggggc ctgggcgctg 960  
 ctgaagcgcc ggagacggcg tcgggtgtcg gaggcggaac cgtccagccc gtccggcgtt 1020  
 tga 1023

<210> 22

<211> 305

<212> PRT

<213> 天蓝色链霉菌

<400> 22

Met Gly Arg Gly Thr Asp Gln Arg Thr Arg Tyr Gly Arg Arg Arg Ala  
 1 5 10 15  
 Arg Val Ala Leu Ala Ala Leu Thr Ala Ala Val Leu Gly Val Gly Val  
 20 25 30  
 Ala Gly Cys Asp Ser Val Gly Gly Asp Ser Pro Ala Pro Ser Gly Ser  
 35 40 45  
 Pro Ser Lys Arg Thr Arg Thr Ala Pro Ala Trp Asp Thr Ser Pro Ala  
 50 55 60  
 Ser Val Ala Ala Val Gly Asp Ser Ile Thr Arg Gly Phe Asp Ala Cys  
 65 70 75 80  
 Ala Val Leu Ser Asp Cys Pro Glu Val Ser Trp Ala Thr Gly Ser Ser  
 85 90 95  
 Ala Lys Val Asp Ser Leu Ala Val Arg Leu Leu Gly Lys Ala Asp Ala  
 100 105 110

[0025]

Ala Glu His Ser Trp Asn Tyr Ala Val Thr Gly Ala Arg Met Ala Asp  
 115 120 125

Leu Thr Ala Gln Val Thr Arg Ala Ala Gln Arg Glu Pro Glu Leu Val  
 130 135 140

Ala Val Met Ala Gly Ala Asn Asp Ala Cys Arg Ser Thr Thr Ser Ala  
 145 150 155 160

Met Thr Pro Val Ala Asp Phe Arg Ala Gln Phe Glu Glu Ala Met Ala  
 165 170 175

Thr Leu Arg Lys Lys Leu Pro Lys Ala Gln Val Tyr Val Ser Ser Ile  
 180 185 190

Pro Asp Leu Lys Arg Leu Trp Ser Gln Gly Arg Thr Asn Pro Leu Gly  
 195 200 205

Lys Gln Val Trp Lys Leu Gly Leu Cys Pro Ser Met Leu Gly Asp Ala  
 210 215 220

Asp Ser Leu Asp Ser Ala Ala Thr Leu Arg Arg Asn Thr Val Arg Asp  
 225 230 235 240

Arg Val Ala Asp Tyr Asn Glu Val Leu Arg Glu Val Cys Ala Lys Asp  
 245 250 255

Arg Arg Cys Arg Ser Asp Asp Gly Ala Val His Glu Phe Arg Phe Gly  
 260 265 270

Thr Asp Gln Leu Ser His Trp Asp Trp Phe His Pro Ser Val Asp Gly  
 275 280 285

Gln Ala Arg Leu Ala Glu Ile Ala Tyr Arg Ala Val Thr Ala Lys Asn  
 290 295 300

Pro  
 305

<210> 23

<211> 918

<212> DNA

<213> 天蓝色链霉菌

<400> 23

atgggtcgag ggacggacca gcggacgcgg tacggccgtc gccgggcgcg tgtcgcgctc 60

gccgcctga ccgccccgt cctgggcgtg ggcgtggcgg gctgcgactc cgtgggcggc 120

gactcaccgc ctcttccgg cagcccgtcg aagcggacga ggacggcgcc cgctgggac 180

accagcccgg cgccgtcgc cgccgtgggc gactccatca cgcgcgctt cgacgcctgt 240

[0026]

gcggtgctgt cggactgcc ggaggtgtcg tgggcgaccg gcagcagcgc gaaggtcgac 300  
 tcgctggccg tacggctgct ggggaaggcg gacgcggccg agcacagctg gaactacgcg 360  
 gtcaccgggg cccggatggc ggacctgacc gctcaggtga cgcgggcggc gcagcgcgag 420  
 ccggagctgg tggcgggtgat ggccggggcg aacgacgcgt gccggtccac gacctcggcg 480  
 atgacgccgg tggcggactt ccgggcgcag ttcgaggagg cgatggccac cctgcgcaag 540  
 aagctcccca aggcgcaggt gtacgtgtcg agcatcccgg acctcaagcg gctctggtcc 600  
 cagggccgca ccaaccgct gggcaagcag gtgtggaagc tcggcctgtg cccgtcgatg 660  
 ctgggcgacg cggactccct ggactcggcg gcgaccctgc ggcgcaacac ggtgcgcgac 720  
 cgggtggcgg actacaacga ggtgctgcgg gaggtctgcg cgaaggaccg gcggtgccgc 780  
 agcgcgacg gcgcggtgca cgagttccgg ttcggcacgg accagttgag cactgggac 840  
 tggttccacc cgagtgtgga cggccaggcc cggctggcgg agatgccta ccgcgcggtc 900  
 accgcaaga atccctga 918

<210> 24

<211> 268

<212> PRT

<213> 龟裂链霉菌(*Streptomyces rimosus*)

<400> 24

Met Arg Leu Ser Arg Arg Ala Ala Thr Ala Ser Ala Leu Leu Leu Thr  
 1 5 10 15  
 Pro Ala Leu Ala Leu Phe Gly Ala Ser Ala Ala Val Ser Ala Pro Arg  
 20 25 30  
 Ile Gln Ala Thr Asp Tyr Val Ala Leu Gly Asp Ser Tyr Ser Ser Gly  
 35 40 45  
 Val Gly Ala Gly Ser Tyr Asp Ser Ser Ser Gly Ser Cys Lys Arg Ser  
 50 55 60  
 Thr Lys Ser Tyr Pro Ala Leu Trp Ala Ala Ser His Thr Gly Thr Arg  
 65 70 75 80  
 Phe Asn Phe Thr Ala Cys Ser Gly Ala Arg Thr Gly Asp Val Leu Ala  
 85 90 95  
 Lys Gln Leu Thr Pro Val Asn Ser Gly Thr Asp Leu Val Ser Ile Thr

[0027]

	100		105		110
Ile Gly Gly Asn Asp Ala Gly Phe Ala Asp Thr Met Thr Thr Cys Asn	115		120		125
Leu Gln Gly Glu Ser Ala Cys Leu Ala Arg Ile Ala Lys Ala Arg Ala	130		135		140
Tyr Ile Gln Gln Thr Leu Pro Ala Gln Leu Asp Gln Val Tyr Asp Ala	145		150		155
Ile Asp Ser Arg Ala Pro Ala Ala Gln Val Val Val Leu Gly Tyr Pro	165		170		175
Arg Phe Tyr Lys Leu Gly Gly Ser Cys Ala Val Gly Leu Ser Glu Lys	180		185		190
Ser Arg Ala Ala Ile Asn Ala Ala Ala Asp Asp Ile Asn Ala Val Thr	195		200		205
Ala Lys Arg Ala Ala Asp His Gly Phe Ala Phe Gly Asp Val Asn Thr	210		215		220
Thr Phe Ala Gly His Glu Leu Cys Ser Gly Ala Pro Trp Leu His Ser	225		230		235
Val Thr Leu Pro Val Glu Asn Ser Tyr His Pro Thr Ala Asn Gly Gln	245		250		255
Ser Lys Gly Tyr Leu Pro Val Leu Asn Ser Ala Thr	260		265		

<210> 25

<211> 1068

<212> DNA

<213> 龟裂链霉菌

<400> 25

```

ttcatcacia cgatgtcaca acaccggcca tccgggtcat cctgatcgt gggaatgggt      60
gacaagcctt cccgtgacga aagggtcctg ctacatcaga aatgacagaa atcctgctca      120
gggaggttcc atgagactgt cccgacgcgc ggccacggcg tccgcgctcc tctcaccctc      180
ggcgctcgcg ctcttcggcg cgagcgccgc cgtgtccgcg ccgcgaatcc aggccaccga      240
ctacgtggcc ctcgcgact cctactcctc gggggtcggc gcgggcagct acgacagcag      300
cagtggtctc tgtaagegca gcaccaagtc ctaccggcc ctgtgggccc cctcgcacac      360
cggtaacgcg ttcaacttca ccgctgttc gggcgcccgc acaggagacg tgctggccaa      420
    
```

[0028]

gcagctgacc ccggtcaact ccggcaccga cctggtcagc attaccatcg gcggcaacga 480  
 cgcgggcttc gccgacacca tgaccacctg caacctccag ggcgagagcg cgtgcctggc 540  
 gcggatcgcc aaggcgcgcg cctacatcca gcagacgctg cccgcccagc tggaccaggt 600  
 ctacgacgcc atcgacagcc gggccccgcg agcccaggtc gtcgtcctgg gctacccgcg 660  
 cttctacaag ctggggcggca gctgcgccgt cggtctctcg gagaagtccc gcgcggccat 720  
 caacgccgcc gccgacgaca tcaacgccgt caccgccaag cgcgccgcg accacggctt 780  
 cgccctcggg gacgtcaaca cgaccttcgc cgggcacgag ctgtgctccg gcgccccctg 840  
 gctgcacagc gtcacccttc ccgtggagaa ctctaccac cccacggcca acggacagtc 900  
 caagggctac ctgcccgtcc tgaactccgc cacctgatct cgcggctact ccgccctga 960  
 cgaagtcccg cccccggcg gggcttcgcc gtaggtgcgc gtaccgccgt cgcccgtcgc 1020  
 gccggtggcc ccgccgtacg tgccgccgc cccggacgcg gtcggttc 1068

<210> 26

<211> 335

<212> PRT

<213> 嗜水气单胞菌

<400> 26

Met Lys Lys Trp Phe Val Cys Leu Leu Gly Leu Val Ala Leu Thr Val  
 1 5 10 15  
 Gln Ala Ala Asp Ser Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly  
 20 25 30  
 Asp Ser Leu Ser Asp Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr  
 35 40 45  
 Leu Pro Ser Ser Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro  
 50 55 60  
 Val Trp Leu Glu Gln Leu Thr Lys Gln Phe Pro Gly Leu Thr Ile Ala  
 65 70 75 80  
 Asn Glu Ala Glu Gly Gly Ala Thr Ala Val Ala Tyr Asn Lys Ile Ser  
 85 90 95  
 Trp Asn Pro Lys Tyr Gln Val Ile Asn Asn Leu Asp Tyr Glu Val Thr  
 100 105 110  
 Gln Phe Leu Gln Lys Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu

[0029]

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

&lt;210&gt; 27

&lt;211&gt; 1008

&lt;212&gt; DNA

&lt;213&gt; 嗜水气单胞菌

&lt;400&gt; 27

atgaaaaaat ggtttgtgtg tttattggga ttggtcgcgc tgacagttca ggcagccgac 60

agtcgccccg ccttttcccg gatcgtgatg ttcggcgaca gcctctccga taccggcaaa 120

atgtacagca agatgcgegg ttacctcccc tccagcccgc cctactatga gggccgtttc 180

[0030]

tccaacggac ccgtctggct ggagcagctg accaaacagt tcccgggtct gaccatcgcc 240  
aacgaagcgg aaggcgggtgc cactgccgtg gcttacaaca agatctcctg gaatcccaag 300  
tatcaggtca tcaacaacct ggactacgag gtcacccagt tcttgcagaa agacagcttc 360  
aagccggacg atctgggtgat cctctgggtc ggtgccaatg actatctggc ctatggctgg 420  
aacacggagc aggatgccaa gegggttcgc gatgccatca gcgatgcggc caaccgcatg 480  
gtactgaacg gtgccaagca gatactgctg ttcaacctgc cggatctggg ccagaacccg 540  
tcagctcgca gtcagaaggt ggtcgaggcg gtcagccatg tctccgccta tcacaaccag 600  
ctgctgctga acctggcacg ccagctggcc cccaccggca tggtaaagct gttcgagatc 660  
gacaagcaat ttgccgagat gctgcgtgat ccgcagaact tcggcctgag cgacgtcgag 720  
aacccttgc acgacggcgg ctatgtgtgg aagccgtttg ccaccgcag cgtcagcacc 780  
gaccgccagc tctccgcctt cagtccgag gaacgcctcg ccatcgccgg caaccgctg 840  
ctggcacagg ccgttgccag tcctatggcc cgccgcagcg ccagcccct caactgtgag 900  
ggcaagatgt tctgggatca ggtacaccgg accactgtcg tgcaagcagc cctgagcgag 960  
cgcgcgccca ccttcacgca gaaccagtac gagttcctcg cccactga 1008

<210> 28

<211> 336

<212> PRT

<213> 杀鲑气单胞菌

<400> 28

Met Lys Lys Trp Phe Val Cys Leu Leu Gly Leu Ile Ala Leu Thr Val  
1 5 10 15  
Gln Ala Ala Asp Thr Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly  
20 25 30  
Asp Ser Leu Ser Asp Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr  
35 40 45  
Leu Pro Ser Ser Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro  
50 55 60  
Val Trp Leu Glu Gln Leu Thr Lys Gln Phe Pro Gly Leu Thr Ile Ala  
65 70 75 80

[0031]

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

<210> 29

<211> 1011

<212> DNA

<213> 杀鲑气单胞菌

[0032]

<400> 29  
 atgaaaaaat ggtttgtttg tttattgggg ttgatcgcgc tgacagttca ggcagccgac 60  
 actcgccccg ccttctcccc gatcgtgatg ttcggcgaca gcctctccga taccggcaaa 120  
 atgtacagca agatgcgcgg ttacctccc tccagcccgc cctactatga gggccgttcc 180  
 tccaacggac ccgtctggct ggagcagctg accaagcagt tcccgggtct gaccatcgcc 240  
 aacgaagcgg aaggcgggtgc cactgccgtg gcttacaaca agatctcctg gaatcccaag 300  
 tatcaggtca tcaacaacct ggactacgag gtcaccctcagt tcttgacaga agacagcttc 360  
 aagccggacg atctggtgat cctctgggtc ggtgccaatg actatctggc atatggctgg 420  
 aatacggagc aggatgccaa gcgagttcgc gatgccatca gcgatgcggc caaccgcatg 480  
 gtactgaacg gtgccaagca gatactgctg ttcaacctgc cggatctggg ccagaacccg 540  
 tcagcccgca gtcagaaggt ggtcgaggcg gtcagccatg tctccgccta tcacaacaag 600  
 ctgctgctga acctggcacg ccagctggcc cccaccggca tggtaaagct gttcgagatc 660  
 gacaagcaat ttgccgagat gctgcgtgat ccgcagaact tcggcctgag cgacgtcgag 720  
 aacctctget acgacggcgg ctatgtgtgg aagccgtttg ccaccgcag cgtcagcacc 780  
 gaccgccagc tctccgcctt cagtccgcag gaacgcctcg ccatcgccgg caaccgctg 840  
 ctggcacagg ccggtgccag tcctatggcc cggcgcagcg ccagccccct caactgtgag 900  
 ggcaagatgt tctgggatca ggtacaccgg accactgtcg tgcacgcagc cctgagcgag 960  
 cgcgcgcgca ccttcatcga gaccagtac gaggctcctcg cccacggatg a 1011

<210> 30

<211> 347

<212> PRT

<213> 嗜水气单胞菌

<400> 30

Met Phe Lys Phe Lys Lys Asn Phe Leu Val Gly Leu Ser Ala Ala Leu  
 1 5 10 15  
 Met Ser Ile Ser Leu Phe Ser Ala Thr Ala Ser Ala Ala Ser Ala Asp  
 20 25 30  
 Ser Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly Asp Ser Leu Ser  
 35 40 45

[0033]

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

[0034]

<210> 31

<211> 27

<212> DNA

<213> 人工序列

<220>

<223> PCR引物

<400> 31

gtgatggtgg gcgaggaact cgtactg

27

<210> 32

<211> 35

<212> DNA

<213> 人工序列

<220>

<223> PCR引物

<400> 32

agcatatgaa aaaatggttt gtttgtttat tgggg

35

<210> 33

<211> 39

<212> DNA

<213> 人工序列

<220>

<223> PCR引物

<400> 33

ttggatccga attcatcaat ggtgatggtg atggtgggc

39

<210> 34

[0035]

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 启动子引物

<400> 34  
taatacgact cactatag

18

<210> 35

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 终止子引物

<400> 35  
ctagttattg ctcagcgg

18

<210> 36

<211> 41

<212> DNA

<213> 人工序列

<220>

<223> PCR引物

<400> 36  
gtcatatgaa aaaatggttt gtgtgtttat tgggattggt c

41

<210> 37

<211> 30

<212> DNA

[0036]

<213> 人工序列

<220>

<223> PCR引物

<400> 37

atggtgatgg tgggcgagga actcgtactg

30

<210> 38

<211> 41

<212> DNA

<213> 人工序列

<220>

<223> PCR引物

<400> 38

gtcatatgaa aaaatggttt gtgtgtttat tgggattggt c

41

<210> 39

<211> 39

<212> DNA

<213> 人工序列

<220>

<223> PCR引物

<400> 39

ttggatccga attcatcaat ggtgatggtg atggtgggc

39

<210> 40

<211> 26

<212> DNA

<213> 人工序列

[0037]

<220>

<223> PCR引物

<400> 40

atgccatggc cgacagccgt cccgcc

26

<210> 41

<211> 27

<212> DNA

<213> 人工序列

<220>

<223> PCR引物

<400> 41

ttggatccga attcatcaat ggtgatg

27

<210> 42

<211> 26

<212> DNA

<213> 人工序列

<220>

<223> PCR引物

<400> 42

ttgctagcgc cgacagccgt cccgcc

26

<210> 43

<211> 27

<212> DNA

<213> 人工序列

<220>

[0038]

<223> PCR引物

<400> 43

ttggatccga attcatcaat ggtgatg

27

<210> 44

<211> 26

<212> DNA

<213> 人工序列

<220>

<223> PCR引物

<400> 44

ttgccatggc cgacactcgc cccgcc

26

<210> 45

<211> 27

<212> DNA

<213> 人工序列

<220>

<223> PCR引物

<400> 45

ttggatccga attcatcaat ggtgatg

27

<210> 46

<211> 26

<212> DNA

<213> 人工序列

<220>

<223> PCR引物

[0039]

<400> 46		
ttgctagcgc cgacactcgc cccgcc		26
<210> 47		
<211> 27		
<212> DNA		
<213> 人工序列		
<220>		
<223> PCR引物		
<400> 47		
ttggatccga attcatcaat ggtgatg		27
<210> 48		
<211> 1047		
<212> DNA		
<213> 嗜水气单胞菌		
<400> 48		
atgtttaagt ttaaaaagaa tttcttagtt ggattatcgg cagctttaat gagtattagc		60
ttgttttcgg caaccgcctc tgcagctagc gccgacagcc gtcccgcctt ttcccggatc		120
gtgatgttcg gcgacagcct ctccgatacc ggcaaaatgt acagcaagat gcgcggttac		180
ctcccctcca gcccgccta ctatgagggc cgtttctcca acggaccgt ctggctggag		240
cagctgacca aacagttccc ggtctgacc atcgccaacg aagcggagg cggtgccact		300
gccgtggett acaacaagat ctctggaat cccaagtatc aggtcatcaa caacctggac		360
tacgaggtca cccagttctt gcagaaagac agcttcaagc cggacgatct ggtgatcctc		420
tgggtcgggtg ccaatgacta tctggcctat ggctggaaca cggagcagga tgccaagcgg		480
gttcgcgatg ccatcagcga tgcggccaac cgcgatgtac tgaacggtgc caagcagata		540
ctgctgttca acctgccgga tctgggccag aaccgctcag ctcgcagtca gaaggtggtc		600
gaggcggta gccatgtctc cgctatac aaccagctgc tgctgaacct ggcacgccag		660
ctggccccca ccggcatggt aaagctgttc gagatcgaca agcaatttgc cgagatgctg		720

[0040]

cgtgatccgc agaacttcgg cctgagcgcac gtcgagaacc cctgctacga cggcggctat 780  
 gtgtggaagc cgtttgccac ccgcagcgtc agcaccgacc gccagctctc cgccttcagt 840  
 ccgcaggaac gcctcgccat cgccggcaac ccgctgctgg cacaggccgt tgccagtctt 900  
 atggcccgcc gcagcgcag cccctcaac tgtgaggga agatgttctg ggatcaggta 960  
 caccgacca ctgtcgtgca cgcagccctg agcgagcgcg ccgccacctt catcgcgaac 1020  
 cagtacgagt tcctcgccca ctgatga 1047

<210> 49

<211> 1007

<212> DNA

<213> 嗜水气单胞菌 (seq 33 底链)

<400> 49

tactttttta ccaaacacac aaataaccct aaccagcgcg actgtcaagt ccgtcggctg 60  
 tcagcggggc ggaaaagggc ctgacactac aagccgctgt cggagaggct atggccgttt 120  
 tacatgtcgt tctacgcgcc aatggagggg aggtcgggcg ggatgatact cccggcaaag 180  
 aggttgctg ggcagaccga cctcgtcgac tggtttgca agggcccaga ctggtagcgg 240  
 ttgcttcgcc ttccgccacg gtgacggcac cgaatgtgt tctagaggac cttagggta 300  
 tagtccagta gttgttgac ctgatgctcc agtgggtcaa gaacgtcttt ctgtcgaagt 360  
 tcggcctgct agaccactag gagaccagc cacggttact gatagaccgg ataccgacct 420  
 tgtgcctcgt cctacggttc gcccaagcgc tacggtagtc gctacgccgg ttggcglacc 480  
 atgacttgcc acggttcgtc tatgacgaca agttggacgg cctagaccgg gtcttgggca 540  
 gtcgagcgtc agtcttcac cagctccgcc agtcggtaca gaggcggata gtgttggtcg 600  
 acgacgactt ggaccgtgcg gtcgaccggg ggtggccgta ccatttcgac aagctctagc 660  
 tgttcgttaa acggtctac gacgactag gcgtcttgaa gccggactcg ctgcagctct 720  
 tggggacgat gctgcccg atacacacct tcggcaaac gtggcgctcg cagtcgtggc 780  
 tggcggtcga gaggcgaag tcaggcgtcc ttgcggagcg gtagcggccg ttggcgacg 840  
 accgtgtccg gcaacggtca ggataccggg cggcgtcgcg gtcgggggag ttgacactcc 900  
 cgttctacaa gaccctagtc catgtgggt ggtgacagca cgtgcgtcgg gactcgtcgc 960  
 cgcggcggtg gaagtagcgc ttggtcatgc tcaaggagcg ggtgact 1007

[0041]

<210> 50

<211> 1011

<212> DNA

<213> 杀鲑气单胞菌 (Seq ID 35 底链)

<400> 50

```
tactttttta ccaaacaac aaataacccc aactagcgcg actgtcaagt ccgtcggctg      60
tgagcggggc ggaagagggc ctagcactac aagccgctgt cggagaggct atggccgttt      120
tacatgtcgt tctacgcgcc aatggagggg aggtcgggcg ggatgatact cccggcaaag      180
aggttgcctg ggcagaccga cctcgtcgac tggttcgtca agggcccaga ctggtagcgg      240
ttgcttcgcc ttccgccacg gtgacggcac cgaatgttgt tctagaggac cttagggttc      300
atagtccagt agttgttga cctgatgctc cagtgggtca agaacgtctt tctgtcgaag      360
ttcggcctgc tagaccacta ggagaccag ccacggttac tgatagaccg tataccgacc      420
ttatgcctcg tcctacggtt cgctcaagcg ctacggtagt cgctacgccg gttggcgtac      480
catgacttgc cacggttcgt ctatgacgac aagttggacg gcctagaccg ggtcttgggc      540
agtcgggcgt cagtcttcca ccagctccgc cagtcggtac agaggcggat agtgttgttc      600
gacgacgact tggaccgtgc ggtcgaccgg ggggtggcct accatttoga caagctctag      660
ctgttcgtta aacggctcta cgacgacta ggcgtcttga agccggactc gctgcagctc      720
ttggggacga tgctccgcc gatacacacc ttcggcaaac ggtgggcgtc gcagtcgtgg      780
ctggcggtcg agaggcggaa gtcaggcgtc cttgcggagc ggtagcggcc gttgggcgac      840
gaccgtgtcc ggcaacggtc aggataccgg gcggcgctgc ggtcggggga gttgacactc      900
ccgttctaca agaccctagt ccatgtgggc tggtgacagc acgtgcgtcg ggactcgctc      960
gcgcggcggg ggaagtagct ctgggtcatg ctcaaggagc ggtgcctac t      1011
```

<210> 51

<211> 35

<212> DNA

<213> 人工序列

[0042]

<220>

<223> PCR引物

<400> 51

agcatatgaa aaaatggttt gtttgtttat tgggg

35

<210> 52

<211> 1047

<212> DNA

<213> 嗜水气单胞菌 (seq ID 54 底链)

<400> 52

tacaaattca aatTTTTtctt aaagaatcaa cctaatagcc gtcgaaatta ctcataatcg 60

aacaaaagcc gttggcggag acgtcgatcg cggetgtcgg cagggcggaa aagggcctag 120

cactacaagc cgctgtcggg gaggctatgg ccgttttaca tgcgtttcta cgcgccaatg 180

gaggggaggt cgggcgggat gatactcccg gcaaagaggt tgcttgggca gaccgacctc 240

gtcgactggt ttgtcaaggc cccagactgg tagcggttgc ttcgccttcc gccacggtga 300

cggcaccgaa tgttgttcta gaggacctta gggttcatag tccagtagtt gttggacctg 360

atgtccagt gggtaagaa cgtctttctg tcgaagtccg gcctgctaga ccactaggag 420

accagccac ggttactgat agaccggata ccgaccttgt gcctcgtcct acggttcgcc 480

caagcgtac ggtagtcgct acgccggtt gcgtaccatg acttgccacg gttcgtctat 540

gacgacaagt tggacggcct agaccggctc ttgggcagtc gagcgtcagt cttccaccag 600

ctccgccagt cggtagagag gcgatagtg ttggtcgacg acgacttga cgtgcccgtc 660

gaccgggggt ggccgtacca ttctgacaag ctctagctgt tcgttaaagc gctctacgac 720

gcactaggcg tcttgaagcc ggactcgtc cagctcttgg ggacgatgct gccgccgata 780

cacaccttcg gcaaacggtg ggcgtcgcag tcgtggctgg cggtcgagag gcggaagtca 840

ggcgtccttg cggagcggta gcggccggtt ggcgacgacc gtgtccgca acggtcagga 900

taccgggcgg cgtcgcggtc gggggagttg aactcccgt tctacaagac cctagtccat 960

gtgggctggt gacagcacgt gcgtcgggac tcgctcgcgc ggcggtggaa gtagcgttg 1020

gtcatgctca aggagcgggt gactact 1047

[0043]

<210> 53

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 区1 GDSX区(说明书第12页)

<220>

<221> NON\_CONS

<222> (1)..(8)

<223> X是选自Met, Ile, Leu, Val, Ala, Gly , Cys, His, Lys, Trp, Tyr或Ph的疏水性残基

<400> 53

Xaa Xaa Xaa Xaa Gly Asp Ser Xaa  
1 5

<210> 54

<211> 6

<212> PRT

<213> 人工序列

<220>

<223> 区2 GANDY区(说明书第28页)

<220>

<221> NON\_CONS

<222> (1)..(6)

<223> X是选自Met, Ile, Leu, Val, Ala, Gly , Cys, His, Lys, Trp, Tyr或Ph的疏水性残基

<400> 54

Xaa Gly Xaa Asn Asp Xaa

[0044]

1 5

## SEQ ID No. 1

```

1 ivaIGD1ld geayygdsgd ggwgagladr Ltallrlrar prgvdvfnrg isGrtsdGr1
61 ivDalvallF laqslqlpnL pPYLsgdflr GANFAsaqAt Ilptsqpfli QvqFkdksq
121 vlelrqalgl lqellrlpv ldakspdlvt imiGtNDlit saffgpkste sdrnsvvpef
181 kdnlrqlikr Lrsnngarii vlitlvilnl gplGClPlkl alalassknv dasgclerln
241 eavadfneal relaiskled qlrkdglp dv kgadvpyvDl ysifqldgi qnpsayvyGF
301 ettkaCCGYG gryNyrvCG naglcnvtak aCnpssylls flfwDgfFps ekGykavAea
361 1

```

图 1

## SEQ ID No. 2

```

1 mkkwfvcllg lvaltvqaad srpafsrivm fgdsldtgk myskmrgylyp sspyyegrif
61 sngpvwleql tnefpgltia neaeggptav aynkiswnpk yqvinnldye vtqflqkdsf
121 kpddlvilw gandylaygw nteqdakrvr daisdaanrm vngakeill fnlpdlgqnp
181 sarsqkvvea ashvsayhnq llnlarqla ptgmvklfei dkqfaemlrd pqnfglsdqr
241 nacyggsyv w kpfasrsast dsqlsafnpq erlaiagnpl laqavaspma arsastlnce
301 gkmfwdqvhv ttvhaalse paatfiesqy eflah

```

图 2

## SEQ ID No. 3

```

1 mkkwfvcllg lialtvqaad trpafsrivm fgdsldtgk myskmrgylyp sspyyegrif
61 sngpvwleql tkqfpgitia neaeggatav aynkiswnpk yqvynnlde vtqflqkdsf
121 kpddlvilw gandylaygw nteqdakrvr daisdaanrm vngakqill fnlpdlgqnp
181 sarsqkvvea vshvsayhnk llnlarqla ptgmvklfei dkqfaemlrd pqnfglsdve
241 npcydggv w kpfatrsvst drqlsafspq erlaiagnpl laqavaspma rrsasplnce
301 gkmfwdqvhv ttvhaalse raatfietqy eflahg

```

图 3

## SEQ ID No. 4

```

1 mpkpalrrvm tatvaavgti algltdatah aapaqatptl dyvalgdsys agsgvlpvdp
61 anlcllrsta nyphviadtt garltdvtcg aaqtadftra qypqvapqid algtgtdlvt
121 ltiggnndst finaitacgt agvlgsgkgs pckdrhgtsf ddeieantyp alkeallgvr
181 arapharvaa lgyptwtpat adpscfklp laagdvpvyr aigahlnlav rraaetgat
241 yvdfsgvsdg hdaceapgtr wlepllfghs lvpvhpnalg ermaehtmd vlgld

```

图 4

## SEQ ID No. 5

```
1 mpkpalrrvm tatvaavgtl algltdatah aapaqatptl dyvalgdsys agsgvlpvdp
61 anllclirsta nyphviadtt garltdvtcg aaqtadftra qygvapqid algtgtdlvt
121 ltiggndnst finaitacgt agvlsggkgs pckdrhgtsf ddeieantyp alkeallgvr
181 arapharvaa lgywitpat adpscflklp laagdvpylr aiqahlndav rraaetgat
241 yvdfsgvsdg hdaceapgtr wiepllfghs lvpvhpnalq errmaehtmd vlgid
```

图 5

## SEQ ID No. 6

```
1 mdyekfillfg dsitefafnt rpiedgkdqy algaalvney trkmdilqrg fkgytsrwal
61 kilpeilkhe snivmatifl gandacsagg qsvplpefid nirquvslmk syhirpiiig
121 pglvdrekwe kekseeialg yfrtnenfai ysdalaklan eekvpfvaln kafqgegga
181 wqqlitdglh fsgkykifh dellkvietf ypqyhpknmq ykldwrdvl ddgsnims
```

图 6

Alignment of pfam00657.6 consensus sequence with P10480  
 \*->ivafGDSLtdg.....eayygdsgggwgagladrL  
 iv+fGDSL+d+++ ++ ++ ++++++ ++s+g w ++l ++  
 P10480 28 IVMFGDLSLSDTgkmyskmrgylpssppYYEGRFSNGPWWLEQLTNEF 74

tall..rlrarprgvdvfnrgisGrtsdGrIivDalvallFlaqlglpn  
 +l + ++++++ +n+ +  
 P10480 75 PGLTIANEAEGGPTAVAYNKISWNPK----- 100

lPpYlsgdfIrgANFAsagAtIiptsgpfliQvqFkdfksqvlelrqal  
 ++ ++  
 P10480 101 -----YQVINN 106

llqellrllpvlakspdlvtimiGtNDlitsaffgpkstesdrnvspe  
 l++e+ ++l +++ k+ dlV+++G+ND+ ++ ++ ++++++  
 P10480 107 LDYEVTQFLQKDSFKPDDLVLWVGANDY-----LAYGWNTQDAKR 148

fkdnlrqlikrLrsnngariivlitlviInlgplGClPlklalalasskn  
 ++d +++++r+ nga+ +++++nl+ lG+ P+  
 P10480 149 VRDAISDAANRMV-LNGAK-----ETLLFNLPDLGQNPS----- 181

vdasgclerlneavadfnealrelaiskledqlrkdgIpdvkgadvpyvD  
 +++ te + ++a++n++l +la +ql+++g++++++d ++++  
 P10480 182 ARSQKVVVEASHVSAHNNQLLNLA-----RQLAPTGMVKLFEIDKQFAE 226

lysifqldldginqpsayv.y....GFe.ttkACCyGgr.yNyn.rv.CG  
 + +q+++ + + ++++++ +++ t++t++++++ +N+++r+ ++  
 P10480 227 MLRDPQNFGLSDQRNACyGsyvWKPfASRSASTDSQLSaFNPQeRLaIA 276

nag.l.c.nvtakaC.npssyll.sflfwDgfHpskGykavAeal<-\*  
 +++ l + +++++a++ +s+ +++++fwD++Hp+ ++a+ e  
 P10480 277 GNPfLaQaVASPMARrSASTLNceGKMfWDQVHPTTVVHAALSEPA 322

Alignment of pfam00657.6 consensus sequence with AAG09804  
 \*->ivafGDSLtdg.....eayygdsgggwgagladrL  
 iv+fGDSL+d+++ ++ ++ ++++++ ++s+g w ++l ++  
 AAG09804 28 IVMFGDLSLSDTgkmyskmrgylpssppYYEGRFSNGPWWLEQLTKQF 74

tallrlraiprgvdvfnrgisGrtsdGrIivDalvallFlaqlglpnlp  
 +g+++ n + +G+t  
 AAG09804 75 -----PGLTIANEAEGGAT----- 88

PYlsgdfIrgANFAsagAtIiptsgpfliQvqFkdfksqvlelrqa...  
 ++++ + +++++ +  
 AAG09804 89 -----AVAYNKISWNpkyq 102

..lglqellrllpvlakspdlvtimiGtNDlitsaffgpkstesdrnv  
 ++l++e+ ++l +++ k+ dlV+++G+ND+ ++ ++ ++  
 AAG09804 103 vyNNLDYEVTQFLQKDSFKPDDLVLWVGANDY-----LAYGWNTQ 144

svpefkdnlrqlikrLrsnngariivlitlviInlgplGClPlklalala  
 +++++d +++++r+ nga+ +++++nl+ lG+ P+  
 AAG09804 145 DAKRVRDAISDAANRMV-LNGAK-----QILLFNLPDLGQNPS----- 181

ssknvdasgclerlneavadfnealrelaiskledqlrkdgIpdvkgadv  
 +++ te + ++a++n++l +la +ql+++g++++++d  
 AAG09804 182 ----ARsqKVVVEAVSHVSAHNNKLLNLA-----RQLAPTGMVKLFEIDK 222

pyvDlysifqldldginqpsayv.y....GFe.ttkACCyGgr.yNyn.r  
 +++++ +q+++ + ++ ++++++ +++ t++ ++l +++ + +++++  
 AAG09804 223 QFAEMLRDPQNFGLSDVENPCyDgyvWKPfATRsvSTRQLSaFSPQeR 272

v.CGnag.l.c.nvtakaC.npssyll.sflfwDgfHpskGykavAeal  
 + +++++ l + +++++a++ +s +++++fwD++Hp+ ++a+ e+  
 AAG09804 273 LaIAGNPfLaQaVASPMARrSASPLNceGKMfWDQVHPTTVVHAALSERa 322

<-\*

AAG09804 - -

Alignment of pfam00657.6 consensus sequence with NP\_631558

```

*->ivaFGDSLIdgeayygdsgggwgagladrLtalrlrarprgvdvf
+va+GDS ++g      +g + +++L  + + + ++ +
NP_631558  42  YVALGDSYSAG-----SGVLPVDPANL----LCLRSTANYPHV 75

nrgisGrtsdGrIivD.a.l.vallFlaqsIglpnlpPYLagdfIrgANF
+ ++G++      D + + +
NP_631558  76  IADTTGAR-----LTDvTcGaAQ----- 93

AsagAtIlptsqpfliQvqFkdfksqvlelrqalglIqellrIipvidak
+++ ++ + ++ +++
NP_631558  94  -----TADfTRAQYPGVAPQLDALGT 114

spdlvtimIGtNDl.....itsaffgpkstesdrnvsvp
+ dlvt+ iG+ND ++ + + ++ + ++ + +k ++ + +++
NP_631558  115 GTDLVTLTIGGNDstffinaitacgtagvLSGGKGSCKDRHGTSFDDEI 164

efkdn..lrqlikrLrs.nngariivlitlviInlg.....plG
e +++ l++++ +r+++ +ar+ +l ++i+++ ++ + + G
NP_631558  165 EANTYpaLKEALLGVRArAPHARVAALGYFWITPATadpscflklplAAG 214

ClPlklalalassknvdaagclerIneavadfnealrelaiskledqlrk
P+          l+ ++a n a+r a
NP_631558  215 DVFPY-----LRAIQaHLNDaVRRaA----- 234

dglpdkvadpvyvDlysiFqldgIqnpSayvyGFettkaCCGyGgryN
++ + tyvD+ ++
NP_631558  235 -----EETGATYVDFSGVSDG----- 250

ynrvCGnaglcNvtakaC.npsyll.sflwDgf...EpsakGykavAe
++aC+ p +++ + lf + + + Ep++ C +++Ae
NP_631558  251 -----HDACeAPGTRWIEFLLFghSLvpvRPNALGERRMAE 286

al<-+
+
NP_631558  287 ET 288

```

Alignment of pfam00657.6 consensus sequence with CAC42140

```

*->ivaFGDSLIdgeayygdsgggwgagladrLtalrlrarprgvdvf
+va+GDS ++g      +g + +++L  + + + ++ +
CAC42140  42  YVALGDSYSAG-----SGVLPVDPANL----LCLRSTANYPHV 75

nrgisGrtsdGrIivD.a.l.vallFlaqsIglpnlpPYLagdfIrgANF
+ ++G++      D + + +
CAC42140  76  IADTTGAR-----LTDvTcGaAQ----- 93

AsagAtIlptsqpfliQvqFkdfksqvlelrqalglIqellrIipvidak
+++ ++ + ++ +++
CAC42140  94  -----TADfTRAQYPGVAPQLDALGT 114

spdlvtimIGtNDl.....itsaffgpkstesdrnvsvp
+ dlvt+ iG+ND ++ + + ++ + ++ + +k ++ + +++
CAC42140  115 GTDLVTLTIGGNDstffinaitacgtagvLSGGKGSCKDRHGTSFDDEI 164

efkdn..lrqlikrLrs.nngariivlitlviInlg.....plG
e +++ l++++ +r+++ +ar+ +l ++i+++ ++ + + G
CAC42140  165 EANTYpaLKEALLGVRArAPHARVAALGYFWITPATadpscflklplAAG 214

ClPlklalalassknvdaagclerIneavadfnealrelaiskledqlrk
P+          l+ ++a n a+r a
CAC42140  215 DVFPY-----LRAIQaHLNDaVRRaA----- 234

dglpdkvadpvyvDlysiFqldgIqnpSayvyGFettkaCCGyGgryN
+ + + tyvD+ ++
CAC42140  235 -----EETGATYVDFSGVSDG----- 250

ynrvCGnaglcNvtakaC.npsyll.sflwDgf...EpsakGykavAe
++aC+ p +++ + lf + + + Ep++ G +++Ae
CAC42140  251 -----HDACeAPGTRWIEFLLFghSLvpvRPNALGERRMAE 286

```

```

al<-*
+
CAC42140 287 HT 288
Alignment of pfam00657.6 consensus sequence with P41734
*->ivafGDSLTDg...eayygdsgggwgagladrltallrlrarprg
++fGDS+T+ +++ + + d+ ga+l + + +tr+
P41734 6 FLfGDSITeFafntRPIEDGKRQYALGAALVNEY-----TRK 43

vdfnrgisGrtsdGrIivDalvallFlaqlslglpnlpPYLsgdflrGAN
+ d+ rg++G+t
P41734 44 MDILQRGFkGYT----- 55

FAsagAtIltptsgpfliQvqFkdfksqvlrlrqalglqlrllpvlida
+ r+al++l+e+l+ +
P41734 56 -----SRWALKILPEILKH-----E 70

kspdlvtimiGtNDlitsaffgpkstesdrnsvpefkdnlrqlikrLrs
+ + ti++G+ND+ ++ +++ v++pef+dn+rq+++++s
P41734 71 SNIVMATIFLGANDA-----CSAGPQSVPLPEFIDNIRQMVSIMKS 111

nngariivlitlviilngplGC1PlklalalassknvdasgcIerlneav
++++ii++++lv ++ ++ k ++ + + r+ne +
P41734 112 YHIRPIIIGPLVDREKW-----EKEKSEEIALGYFRTNENF 148

adfnealrelaiskledqlrkdglpdkvadpvyvDlyisifqldldiqnp
a + al +la ++ +vp+v l+++fq+ +g++++
P41734 149 AIYSDALAKLA-----NEEKVFPFVALNKAFQOEGGDAWQ 182

sayvyGFettkaCCGyGgryNynrvCGnaglcncvtakaCnpssyllsflf
+ l+
P41734 183 Q-----LL 185

wDgfHhpkGykavAeal<-*
Dg+H+s kGyk++++l
P41734 186 TDGLHFSGKGYKIFHDEL 203

```

图 7

```

A.sal 1 MKKWFVCLLGLIALTVQAAADTRPAFSRIVMFGDSLSDTGKMYSKMRGYLPSSPPYYEGRF 60
+ +
A.hyd 1 MKKWFVCLLGLVALTVQAAADSRPAFSRIVMFGDSLSDTGKMYSKMRGYLPSSPPYYEGRF 60

A.sal 61 SNGPVWLEQLTKQFPGLTIANEAEGGATAVAYNKISWNPKYQVINNLDYEVTQFLQKDSF 120
++ +
A.hyd 61 SNGPVWLEQLTNEFPGLTIANEAEGGPTAVAYNKISWNPKYQVINNLDYEVTQFLQKDSF 120

A.sal 121 KPDDLVLWVGANDYLAYGWNTQDAKRVRDAISDAANRMVNLNGAKQILLFNLPDLGQNP 180
+
A.hyd 121 KPDDLVLWVGANDYLAYGWNTQDAKRVRDAISDAANRMVNLNGAKEILLFNLPDLGQNP 180

A.sal 181 SARSQVVEAVSHVSAYHNKLLLNLARQLAPTGMVKLFEIDKQFAEMLRDPQNFGLSDVE 240
+ +
A.hyd 181 SARSQVVEAASHVSAYHNQLLLNLARQLAPTGMVKLFEIDKQFAEMLRDPQNFGLSDQR 240
++

A.sal 241 NPCYDGGYVWKPFASTRSVSTDRQLSAFSPQERLAIAGNPLLAQAVASPMARRSASPLNCE 300
+ + + + + + + +
A.hyd 241 NACYGGSYVWKPFASTRSASTDSQLSALNPFQERLAIAGNPLLAQAVASPMARRSASTLNCE 300

A.sal 301 GKMFWDQVHPTTVVHAALSERAAFFIETQYEF LAH 335
+ +
A.hyd 301 GKMFWDQVHPTTVVHAALSEPAATFIESQYEF LAH 335

```

图 8

```

1 ATGAAAAAAT GGTTTGTGTG TTTATTGGGA TTGGTCGCGC TGACAGTTCA GGCAGCCGAC
61 AGCCGTFCCCG CCTTCTCCCG GATCGTGATG TTTGGCGACA GCCTCTCCGA TACCGGCAAG
121 ATGTACAGCA AGATGCGCGG TTACCTCCCC TCCAGCCCCC CCTACTATGA GGGCCGCTTC
181 TCCAACGGGC CCGTCTGGCT GGAGCAGCTG ACCAACGAGT TCCCGGGCCT GACCATAGCC
241 AACGAGGCGG AAGGCGGACC GACCGCCGTG GCTTACAACA AGATCTCCTG GAATCCCAAG
301 TATCAGGTCA TCAACAACCT GGACTACGAG GTCACCCAGT TCTTGCAAAA AGACAGCTTC
361 AAGCCGGACG ATCTGGTGAT CCTCTGGGTC GGCGCCAAG ACTATCTGGC CTATGGCTGG
421 AACACAGAGC AGGATGCCAA GCGGGTGC GC GCCCATCA GCGATGCGGC CAACCCGATG
481 GTGCTGAACG GCGCCAAGGA GATACTGCTG TTCAACCTGC CGGATCTGGG CCAGAACCCC
541 TCGGCCCGCA GCCAGAAGGT GGTGAGGCG GCCAGCCATG TCTCCGCCA CCACAACCAG
601 CTGCTGCTGA ACCTGGCAGC CCAGCTGGCT CCCACCGCA TGGTGAAGCT GTTCGAGATC
661 GACAAGCAGT TTGCCGAGAT GCTGCGTGAT CCGCAGAACT TCGGCCTGAG CGACCAAGAGG
721 AACGCCTGCT ACGGTGGCAG CTATGTATGG AAGCCGTTT GCTCCCGCAG CGCCAGCACC
781 GACAGCCAGC TCTCCGCTT CAACCCGCGAG GAGCGCCTCG CCATCGCCGG CAACCCGCTG
841 CTGGCCACAG CCGTGGCCAG CCCCATGGCT GCCCGCAGCG CCAGCACCCCT CAACTGTGAG
901 GGCAAGATGT TCTGGGATCA GGTCCACCCC ACCACTGTG TGCACGCCG CCTGAGCGAG
961 CCCGCCGCA CCTTCATCGA GAGCCAGTAC GAGTTCCTCG CCCAC
    
```

图 9

```

1 ATGAAAAAAT GGTTTGTTTG TTTATTGGGG TTGATCGCGC TGACAGTTCA GGCAGCCGAC
61 ACTCGCCCCG CCTTCTCCCG GATCGTGATG TTCGGCGACA GCCTCTCCGA TACCGGCAAA
121 ATGTACAGCA AGATGCGCGG TTACCTCCCC TCCAGCCCCC CCTACTATGA GGGCCGTTTC
181 TCCAACGGAC CCGTCTGGCT GGAGCAGCTG ACCAAGCAGT TCCCGGGTCT GACCATCGCC
241 AACGAGGCGG AAGGCGGTGC CACTGCCGTG GCTTACAACA AGATCTCCTG GAATCCCAAG
301 TATCAGGTCT ACAACAACCT GGACTACGAG GTCACCCAGT TCTTGCAAAA AGACAGCTTC
361 AAGCCGGACG ATCTGGTGAT CCTCTGGGTC GGTGCCAATG ACTATCTGGC ATATGGCTGG
421 AATACGGAGC AGGATGCCAA GCGAGTTCGC GATGCCAATG GCGATGCGGC CAACCCGATG
481 GACTGGAACG GTGCCAAGCA GATACTGCTG TTCAACCTGC CGGATCTGGG CCAGAACCCC
541 TCAGCCCCGA GTCAGAAGGT GGTGAGGCG GTCAGCCATG TCTCCGCCA CCACAACCAG
601 CTGCTGCTGA ACCTGGCAGC CCAGCTGGCC CCCACCGCA TGGTGAAGCT GTTCGAGATC
661 GACAAGCAAT TTGCCGAGAT GCTGCGTGAT CCGCAGAACT TCGGCCTGAG CGACCTCGAG
721 AACCCCTGCT ACGACGGCCG CTATGTGTGG AAGCCGTTT CCACCCGCGC CGTCAGCACC
781 GACCGCCAGC TCTCCGCTT CAGTCCGCG GACCGCCTCG CCATCGCCGG CAACCCGCTG
841 CTGGCACAGG CCGTGGCCAG TCCATGGGCC CGCCGCGAGC CCAGCCCCCT CAACTGTGAG
901 GGCAAGATGT TCTGGGATCA GGTACACCCC ACCACTGTG TGCACGCAGC CTTGAGCGAG
961 GCGGCCGCA CCTTCATCGA GACCCAGTAC GAGTTCCTCG CCCACGGATG A
    
```

图 10

```

1 ATGCCGAAGC CTGCCCTTCG CCGTGTGATG ACCGCGACAG TCGCCGCCGT CCGCACGCTC
61 GCCCTCGGCC TCACCGACGC CACCGCCAC GCCGCGCCCG CCCAGGCCAC TCCGACCCTG
121 GACTACGTCG CCTTCGGCGA CAGCTACAGC GCCGGTCCG GCGTCTGCCC CGTCGACCCC
181 GCCAACCTGC TCTGTCTGCG CTCGACGGCC AACTACCCCC ACGTCATCGC GGACACGAGC
241 GCGGCCCGCC TCACGGACGT CACTGCGGCC GCCGCGCAGA CCGCCGACTT CACGCGGGCC
301 CAGTACCCGG GCGTCGCACC CCAGTTGGAC GCGCTCGGCA CCGGCACGGA CCTGCTCAGC
361 CTCACCATCG GCGGCAACGA CAACACCACC TTCATCAACG CCATCACGGC CTGCGCCACG
421 GCGGGTGTCC TCAGCGGCGG CAAGGGCAGC CCTGCAAGG ACAGGCACGG CACTCTCTTC
481 GACGACGAGA TCGAGGCCAA CACGTACCCC GCGCTCAAGG AGGGCTGCT CCGGCTCCGC
541 GCCAGGGCTC CCCACGCCAG GGTGGCGGCT CTCGGCTACC CGTGGATCAC CCCGGCCACC
601 GCGGACCCGT CCTGCTTCTT GAAGCTCCCC CTCGCGCCCG GTGACGTGCC CTACCTGCGG
661 GCCATCCAGG CACACCTCAA CGACCGGTC CGGCGGGCCG CCGAGGAGAC CGGAGCCACC
721 TACGTGGACT TCTCCGGGGT GTCGACGGC CACGACGCT GCGAGGCCCC CGGCACCCGC
781 TGATCGAAC CGTGTCTCTT CGGCACAGC CTCGTCTCCG TCCACCCCAA CGCCCTGGGC
841 GAGCGGCGCA TGGCCGAGCA CACGATGGAC GTCCTCGGCC TGGACTGA
    
```

图 11

```

1   TCAGTCCAGG CCGAGGACGT CCATCGTGTG CTCGGCCATG CGCCGCTCGC CCAGGGCGTT
61  GGGTGGACG GGAACGAGGC TGTGCCCGAA GAGCAGCGGT TCGATCCAGC GGGTGCCGGG
121 GGCCTCGCAG GCGTCTGGCC CGTCCGACAC CCCGGAGAAG TCCACGTAGG TGGCTCCGGT
181 CTCCTCGGCG GCCCGCCGGA CCGCGTCGTT GAGGTGTGCC TGGATGGCCC GCAGGTAGGG
241 CACGTACCCG GCGGCGAGGG GGAGCTTCAG SAAGCAGGAC GGGTGGCGGG TGGCCGGGGT
301 GATCCACGGG TAGCCGAGAG CCGCCACCCT GCGGTGGGGA GCCCTGGCCG GGACGCCGAG
361 CAGCGCTCC TTGAGCGCGG GGTACGTGTT GGCCTCGATC TCGTUGLGA AGGAGGTGCC
421 GTGCCTGTCC TTGCAGGGGC TGCCCTTGCC GCCCTGAGG ACACCCGCGG TGCCCGAGGC
481 CGTGATGGCG TTGATGAAGG TGCTGTTGTC GTTGCCTGCC ATGGTGAGCG TGACCAGGTC
541 CGTGCCGGTG CCGAGCGCGT CCAACTGGGG TCGCACGCCG GGGTACTGGG CCCGCTGAA
601 GTCGGCGGTC TCGCGGGCGC CGCAGGTGAC GTCCGTGAGG CCGGCGCCCG TCGTGTCCGC
661 GATGACGTGG GGGTAGTTGG CCGTCGAGCG CAGACAGAGC AGGTGGCCGG GGTGACGGG
721 CAGGACGCCG GAGCCGGGCG TGTAGCTGTC GCCGAGGGCG ACGTAGTCCA GGGTCGGAGT
781 GGCCTGGGCG GCGCGGGCGT GGGCGGTGGC GTCGGTGAGG CCGAGGGCGA GCGTGCCGAC
841 GCGCGGACT GTCGCGGTCA TGACACGGCG AAGGGCAGGC TTCGGCAT
    
```

图 12

```

1   ATGGATTACG AGAAGTTTCT GTTATTTGGG GATTCCATTA CTGAATTGCT TTTAATACT
61  AGGCCCATTC AAGATGGCAA AGATCAGTAT GCTCTTGGAG CCGCATTAGT CAACGAATAT
121 ACGAGAAAAA TGGATATTCT TCAAAGAGGG TTCAAAGGGT ACACCTTAGT ATGGGCGTTG
181 AAAATACTTC CTGAGATTTT AAAGCATGAA TCAAATATTG TCGTGGCCAC AATATTTTGG
241 GGTGCCAACG ATGCATGCTC AGCAGGTGCC CAAAGTGTCC CCTTCCCGA ATTTATCGAT
301 AATATTCGTC AAATGGTATC TTTGATGAAG TCTTACCATA TCCGTCCTAT TATAATAGGA
361 CCGGGGCTAG TAGTAGAGA GAGTGGGAA AAGAAAAAT CTGAAGAAAT AGCTCTCGGA
421 TACTTCCGTA CCAACGAGAA CTTTCCATT TATCCGATG CCTTAGCAA ACTAGCCAAT
481 GAGGAAAAAG TTCCCTTCGT GGCTTTGAAT AAGGCGTTC AACAGGAAGG TGGTATGCT
541 TGGCAACAAC TGCTAACAGA TGGACTGCAC TTTCCGGAA AAGGGTACAA AATTTTCAT
601 GACGAATTAT TGAAGTCAT TGAGACATTC TACCCCAAT ATCATCCCAA AAACATGCAG
661 TACAAACTGA AAGATTGGAG AGATGTGCTA GATGATGGAT CTAACATAAT GTCTTGA
    
```

图 13

(SEQ ID No. 12)

```

      10      20      30      40      50      60
      |      |      |      |      |      |
MNLRQWMAA TAALALGLAA CGGGGTDQSG NPNVAKVORM VVFGDSLSDI GTYTPVAQAV

      70      80      90      100     110     120
      |      |      |      |      |      |
GGGKFTTNPQ PIWAETVAAQ LGVTLTPAVM GYATSVQNCPE KAGCFDYAQG GSRVTDPNGI

      130     140     150     160     170     180
      |      |      |      |      |      |
GHNGGAGALT YPVQQQLANF YAASNVTENG NNDVVFVLAG SNDIFFWTTA AATSGSGVTF

      190     200     210     220     230     240
      |      |      |      |      |      |
AIATAQVQQA ATDLVGYVKD MIAKGATQVY VFNLPDSSLT PDGVASGTTG QALLHALVGT

      250     260     270     280     290     300
      |      |      |      |      |      |
FNNTLQSGLA GTSARIIDFN AQLTAAIQNG ASFGFANTSA RACDATKINA LVPSAGGSSL

      310     320     330     340
      |      |      |      |
FCSANTLVAS GADQSYLFDG GVHPTTAGHR LTASNVLRL LADNVAH
    
```

图 14

## (SEQ ID No. 13)

```

atgaacctgc gtcaatggat gggcgccgcc acggtgccc ttgccttggg cttggccgcg      60
tgcgggggcy gtdggaccga ccagagcggc aatcccaatg tcgccaaggt gcagcgcag      120
gtggtgttcg gcgacagcct gagcgatc ggcacctaca ccccgtcgc gcaggcggtg      180
ggcggcgcca agttcaccac caaccgggc ccgatctggg ccgagaccgt ggccgcgcaa      240
ctgggctgta cgtcaccgcc ggcggtgat gctacacca cctccgtca gaattgcccc      300
aaggccggct gcttcgacta tgcgcagggc ggctcgcgcy tgaccgatcc gaacggcacc      360
ggccacaacg gcggcgcggg ggcgctgacc taccgggttc agcagcagct cgccaacttc      420
tacgcgccca gcaacaacac attcaacggc aataacgatg tcgtcttcgt gctggccggc      480
agcaacgaca ttttcttctg gaccactgcy gcggccacca gcggctccgg cgtgacgccc      540
gccattgcca cggcccaggt gcagcaggcc gcgacggacc tggtcggcta tgtcaaggac      600
atgatcgcca agggtgcygc gcaggctac gtgttcaacc tgcccagacag cagcctgacg      660
ccggaccgcy tggcaagcgg cacgaccgqc caqcgctgc tgacgcgct ggtgggcacg      720
ttcaacacga cgtgcaaaag cgggctggcc ggcacctcgg cgcgcacat cgacttcaac      780
gcacaactga ccgcgcgcat ccagaatggc gcctcgttcg gcttcgcca caccagcgcc      840
cgggcctcgc acgccaacaa gatcaatgcc ctggtgcccga gcgcggcggc cagctcgtg      900
ttctgctcgg ccaacacgct ggtggttcc ggtgaggacc agagctacct gttcgcgac      960
ggcgtgcacc cgaccacggc cggccatcgc ctgatcgcca gcaacgtgct ggccgcgctg      1020
ctggcggata acgtcgcgca ctga      1044

```

图 15

## (SEQ ID No. 20)

```

1 migsyvavgd sftcgvdpq pdgafvgwad rlavlladdr pegdftytnl avrgrlldqi
61 vaeqvprvvg lapdlvsfaa gndiirpqt dpdevaerfe lavaalataa gtlvlttgfd
121 trgvplkhl rgkiatyngv vraiadrygc pvldlwslrs vqdrrawdada rhlspcpgt
181 rvalraggal glrvpadpdq pwpplpprgt ldvrrddvhv areylvpwig rrlrgessgd
241 hvtakgtlsp daiktriaav a

```

图 16

## (SEQ ID No. 21)

```

1 gtgatcgggt cgtacgtggc ggtgggggac agcttcaccg agggcgtcgg cgaccccgcc
61 cccgacgggg cgttcgtcgg ctgggcccgc cggctcgcg tactgctcgc ggaccggcgc
121 cccgagggcy acttcacgta cacgaacctc gccgtgcygc gcaggctcct cgaccagatc
181 gtggcggaac aggtcccgcg ggtcgtcggc ctcgcgcccg acctcgtctc gttcggcggc
241 ggcggcaacg acatcatccg gcccggcacc gatcccgcag aggtcggcga gcgggtcggc
301 ctggcggtgg ccgcgctgac cgcgcgggccc ggaaccgtcc tggtgaccac cgggttcgac
361 acccgggggg tgcccgtcct caagcacctg cgcggcaaga tcgccacgta caacgggcac
421 gtcccgcgca tcgcccagcc ctacggctgc ccggtgctgc acctgtggtc gctgcggagc
481 gtccaggacc gcaggggcgtg ggacgcccgc cggctgcacc tgtcggcggg ggggcacacc
541 cgggtggcgc tgcggcggg gcaggccctg ggctgcccgc tcccggcga ccctgaccag
601 ccctggccgc ccctgccgcc gcgcggcagc ctgcagctcc ggcgcgacga cgtgcactgg
661 gcgcgcgagt acctggtgcc gtggatcggg cgcgggctgc ggggcgagtc gtcggggcag
721 cacgtgacgg ccaaggggac gctgtcggc gacgccatca agacgggat cgcccgggtg
781 gcctga

```

图 17

## (SEQ ID No. 22)

```

1  mqtncpyatsl vavgdstfey msdlldpdsy rgwadllatr mearspgfry anlavrghli
61  qqivdeqvdv aaamgadvit lvvgllndtir pkcdmarvrd lltqaverla phceqlvlmr
121 spgrqgppvle rfrprmealf aviddlagrh gavvvdlyga qsladprmwv vdrhlhtaeg
181 hrrvaeavwq slghpedpde whapipatpp pgwvtrrtad vrfarqhlip wigrrltgrs
241 sgdglpakrp dllypedpar

```

图 18

## (SEQ ID No. 23)

```

1  atgcagacga accccgcgta caccagtctc gtcgccgctc gcgactcctt caccgagggc
61  atgtcggacc tgetgcccga cggctectac cgtggctggg cggactcctt cggcaccggg
121 atggcggccc gctccccggg cttccgggtac gccaacctgg cggtgccgcg gaagctgatc
181 ggacagatcg tcgacgagca ggtggacgtg gccgccgcca tgggagccga cgtgatcacg
241 ctggtcggcg ggtccaacga cacgctgcgg cccaagtgcg acatggcccg ggtgcccggc
301 ctgctgacc aggccgtgga acgctcgcc cgcactgcg agcagctggt gctgatgcgc
361 agtcccggtc gccaggggtc ggtgctggag cgcttcggc cccgatgga ggcctgttc
421 gccgtgatcg acgacctggc cggcggcac ggcgccgtg tcgtgcacct gtacggggcc
481 cagtccgtgg ccgaccctcg gatgtggac gtggaccggc tgcacctgac cggcagggc
541 caccgcccgg tcgcccgggc ggtgtggcag tcgctcggc acgagcccga ggaccccag
601 tggcacgcgc cgatcccggc gacgccggc cgggggtggg tgacgcgac gaccgcccg
661 gtccggttcg cccggcagca cctgctgccc tggataggcc gcaggtgac cggcgcctcg
721 tccggggacg gcctgccggc caagcgccc gacctgctc cctacgagga ccccgcacg
781 tga

```

图 19

## (SEQ ID No. 24)

```

1  mtrgrdygag apptkhrall aaivtlivai saaiyagasa ddgsrdhalq eggrlpdyda
61  apastgawvg awatapaaae pgtettglag rsvrnvvhts vgggtgaritl snlygqsplt
121 vthasialaa gpdtaaaiaa tmrrltfggs arviipaggg vmsdtarlai pyganvlvtt
181 yspipsgpvt ynpqarqtsy ladgdrtdv tavaytptp ywryltaldv lsheadgtvv
241 afgdsitdga rsqsdanhrw tdvlaarlhe aagdgrdtp rsvvnegisg nrlltsrpgr
301 padnpsqlsr fgrdvlertn vkavvvlgv ndvlaspela drdailtglr tlvdraharg
361 lrvvgatitp fggygytea retmrqevne eirsgrvfdt vvdfdkalrd pydprmrds
421 ydsqdhlpq dkgyarmgav idlaalkgaa pvka

```

图 20

## (SEQ ID No. 25)

```

1 atgaccggg gtcgtgacgg ggggtcgggg ggcggcccca ccaagcaccg tgccttgcct
61 gcggcgatcg tcacctgat agtggcgatc tccgcggcca tatacgcccg agcgtccggc
121 gacgacggca gcagggacca cgcgctgcag gccggaggcc gtctcccacg aggagacgcc
181 gccccgcgt ccaccggtgc ctgggtgggc gcctgggcca cgcaccggc cgcggccgag
241 ccgggacccg agacgaccgg cctggcgggc cgctccgtgc gcaacgtgt gcacacctcg
301 gtcggcgcca ccggcgcgcg gatcaccctc tcgaacctgt acgggcagtc gccgtgacc
361 gtcaacacag cctcgatcgc cctggccgcc gggcccgaca ccgcccgcg gatcgccgac
421 accatggccc ggctcacctt cggcggcagc gccgggtga tcatcccggc gggcgccag
481 gtgatgagcg acaccggcgg cctcgccatc cctcagggg cgaacgtct ggtcacacg
541 tactccccca tcccgtccgg gccggtgacc taccatccgc aggccggca gaccagctac
601 ctggccgacg gcgaccgcac ggcggacgtc accgcccgtc cgtacaccac cccacgccc
661 tactggcgct acctgaccgc cctcgacgtg ctgagccacg agggccgacg cacggtcgtg
721 gcgttcggcg actccatcac cgacggcgcc cgctcgaga gcgacgcaa ccaccgctg
781 accyacytcc tcyccyccy cctgcaayag ycyccyccy acggccyga cacgcccyt
841 tacagcgtcg tcaacgaggg catcagcggc aaccggctcc tgaccagcag gccggggcgg
901 ccggccgaca acccgagcgg actgagccgg ttccagcggg acgtgctgga accgaccaac
961 gtcaaggccg tcgtcgtcgt cctcggcgtc aacgacgtcc tgaacagccc ggaactcgcc
1021 gaccgagcag ccatcctgac cggcctgcgc accctcgtcg accggcgca cgcgggggga
1081 ctgcccgtcg tcggcgccac gatcacgccc ttccggcggc acggcggcta caccgagcc
1141 cgcgagacga tgcggcagga ggtcaacgag gagatcccct ccggccgggt cttcgacacg
1201 gtcgctgact tcgacaaggc cctgcgcgac ccgtacgacc cgcggcggat gcgctccgac
1261 tacgacagcg gcgaccacct gacccccggc gacaaggggt acgcccgat gggcgcggtc
1321 atcgacctgg ccgctgtaa gggcgggcg cgggtcaagg cgtag

```

图 21

## (SEQ ID No. 26)

```

1 mtsmsrarva rriaagaayg gggiglagaa avglvvaevq larrrvgvgt ptrvpnaqgl
61 yggtlptaqd pplrlmmlgd staagggvhr agqtpgalla sglavaaerp vrlgsvaqpg
121 acsddidrxv alvlaepdrv pdicvimvga ndvthmpat rsvrhlssav rrlrtagaev
181 vvgtcpdlgt iervrqplrw larrasrqla aaqtigaveq ggrtvslgdl lgpefaqnpr
241 elfgpdnyhp saegyataam avlpsvcaal glwpadeehp dalrregflp varaaaaaas
301 eagtevaam ptgprgpwal lkrrrrrvs eaepsspsgv

```

图 22

## (SEQ ID No. 27)

```

1 atgacgagca tgcgagggc gagggtgycg cyccygtatc cggccggcgc ggcgtacggc
61 ggcggcggca tcggcctggc gggagcggcg gcggtcgtc tgggtggtggc cgaggtgacg
121 ctggccagac gcagggtggg ggtgggcacg ccgaccggg tgccgaacgc gcagggactg
181 tacggcgcca ccctgcccac ggcggcgac ccgcccgtgc ggctgatgat gctggggcag
241 tccacggccg ccgggcaggc cgtgcaccgg gccgggcaga cgcggggcgc gctgctggcg
301 tccgggctcg cggcgggtgg ggagcggccg gtgcccgtg ggtcggctcg ccagccgggg
361 gcgtgctcgg acgacctgga ccgacaggtg gcgctgggtc tcgcccagcc ggaccgggtg
421 cccgacatct gcgtgatcat ggtcggcggc aacgacgtca cccaccggat gccggcgacc
481 cgtcgggtgc ggcacctgtc ctcggcggtc cggcggctgc gcacggccgg tcgaggagtg
541 gtagtcggca cctgtecggc cctgggcagc atcgagcggg tcggccagcc gctgcctcag
601 ctggcccggc gggcctcacg gcagctcgc gcggcacaga ccacccggc cgtcgagcag
661 ggcggggcga cgggtgcctg gggcgacctg ctgggtcccg agttcgcgca gaaccggcgg
721 gagctcttcg gcccgcacaa ctaccacccc tccgcccagg ggtacgccac gggcgcgatg
781 gcggtactgc cctcgggtgt gcgcccgtc ggcctgtggc cggccgacga ggaacccccg
841 gacgcgctgc gccgcgaggg ctccctgccc gtygcgcgcg cggcggcgga ggcggcgtcc
901 gagcgggta cggaggtcgc cgcccacatg cctacggggc ctcggggggc ctggggcgtg
961 ctgaagcgcc ggagacggcg tcgggtgtcg gaggcggaac cgtccagccc gtcggcgctt
1021 tga

```

图 23

## (SEQ ID No. 28)

```

1 mgrgtdqrtr ygrrrarval aaltaavlqv gvagcdsvvg dspapsqspk krtrtapawd
61 tspasvaavg dsitrgfdac avlsdcpevs watgssakvd slavrligka daaehswnya
121 vtgarmadlt aqvtraaqre pelvavmaga ndacrsttsa mtpvadfrac feeamatlrk
181 klpkavvyvs sipdlkrlws qgrtnplgkq vwklglcpms lgdadslds atrlrrntvrd
241 rvadynevlr evcakdrrec sddgavhefr fgtdqlshwd wfhpsvdgqa rlaeiayrav
301 taknp

```

图 24

## (SEQ ID No. 29)

```

1 atgggtcgag ggacggacca gcggacgcgg tacggccgct gccgggcccg tgcgcgctc
61 gccgccctga ccgcccccgt cctgggctgt ggcgtggcgg gctgcgactc cgtgggcccg
121 gactcaccgg ctccctccgg cagcccgtcg aagcggacga ggacggcccg ccctggggac
181 accagcccgg cgtccgtcgc cgcctgtggc gactccatca cgcgcccgtt cgacgcctgt
241 gcggtgctgt cggactgccc ggaggtgtcg tgggagaccg gcagcagcgc gaaggtcgac
301 tcgctggccg tacggctgct ggggaaggcg gacgcggccg agcacagctg gaactacgcg
361 gtcaccgggg ccgggatggc ggacctgacc gctcaggtga cgcgggcccg gcagcgcgag
421 ccggagctgg tggcggatgt ggccggggcg aacgacgcgt gccggtccac gacctcggcg
481 atgaccccgg tggcggactt ccggcgcaag ttcgagggag ccatggccac cctgcgcaag
541 aagctcccca aggcgcaggt gtacgtgtcg agcatcccgy acctcaagcy gctctggtcc
601 cagggcccga ccaaccgctt gggcaagcag gtgtggaagc tcggcctgtg cccgtcgatg
661 ctgggcccag cggactccct ggactcggcg gcgaccctgc ggcgcaacac ggtgcgagc
721 cgggtggcgg actacaacga ggtgctgcgg gaggtctgcg cgaaggaccg cgggtgcccg
781 agcgacgacg gcgcccgtga cgagttccgg ttcggcacgg accagttgag ccactgggac
841 tggttccacc cgagtggtga cggccaggcc cggctggcgg agatcgccca ccgcccgtg
901 accgcaaga atccctga

```

图 25

## (SEQ ID No. 30)

```

1 mrlsrraata sallltpala lfgasaavsa priqatdyva lgdsyssvgv agsydsssgs
61 ckrstksypa lwaashtgtr fnftacsgar tgdvllakqlt pvnsgtdlvs itiggnadagf
121 adtmattcnlq gesaclaria karayiqqtl paqldqvyda idsrpaaqv vvlgyprfyk
181 lggscavglc eksraainaa addinavtak raadhgfafg dvnttfaghe lcsgapwlhs
241 vtlpvensyh ptangqskgy lplvlnsat

```

图 26

## (SEQ ID No. 31)

```

1  ttcatcacia cgatgtcaca acaccggcca tccgggtcat ccctgatcgt gggaatgggt
61  gacaagcctt cccgtgacga aagggtcctg ctacatcaga aatgacagaa atcctgctca
121 gggagggttc atgagactgt cccgacgcgc gccacggcg tccgcgctcc tcctcaccct
181 ggcgctcgcg ctcttcggcg cgagcgccgc cgtgtccgcg ccgcgaatcc aggccaccga
241 ctacgtggcc ctggcgact cctactcctc gggggtcggc gcgggcagct acgacagcag
301 cagtggctcc tgtaagcgca gcaccaagtc ctaccggcc ctgtggggcg cctcgcacac
361 cgttacgcgg ttcaacttca ccgctgttc gggcgcccgc acaggagacg tgctggccaa
421 gcagctgacc ccggtcaact ccggcacgca cctggtcagc attaccatcg gcggcaacga
481 cgcgggcttc gccgacacca tgaccacctg caacctccag ggcgagagcg cgtgcctggc
541 gcggatcgcc aaggcgcgcg cctacatcca gcagacgctg cccgcccagc tggaccaggt
601 ctacgacgcc atcgacagcc gggccccgc agcccagtc gtcgtcctgg gctaccggcg
661 cttctacaag ctggggcgca gctgcgcctg cggctctctg gagaagtccc gcgcccct
721 caacgcccgc gccgacgaca tcaacgccgt caccgccaag cgcgcccgcg accacggctt
781 cgccttcggg gacgtcaaca cgacctcgc cgggcacgag ctgtgtctcg gcgcccctg
841 gctgcacagc gtcacccttc ccgtggagaa ctcctaccac cccacggcca acggacagtc
901 caagggctac ctgcccgtcc tgaactcgc cacctgatct cgcggctact ccgcccctga
961 cgaagtcccc cccccgggcg gggcttcgca gtaggtgcgc gtaccgctg cgcgctcg
1021 gccgggtggc ccgcccgtac tgcccgcgc cccggacgcg gtcggttc

```

图 27

## (SEQ ID No. 32)

```

1  MKKWFVCLLG LVALTVQAAD SRPAFSRIVM FGDSLSDTGK MYSKMRGYLP
51  SSPPYYEGRF SNGPVWLEQL TKQFPGLTIA NEAEGGATAV AYNKISWMPK
101 YQVINLDYE VTQFLQKDSF KPDDLVLWV GANDYLAYGW NTEQDAKVRV
151 DAISDAANRM VINGAKQILL FNLFDLQNP SARSQKVVEA VSHVSAYHNQ
201 LLLNLARQLA PTGMVKLFEI DKQFAEMLRD PQNFGLSDVE NPCYDGGYVW
251 KPFATRSVST DRQLSAFSPQ ERLAIAGNPL LAQAVASPMA RRSASPLNCE
301 GKMFWDQVHP TTVVHAALSE RAATFIANQY EFLAH*

```

图 28

## (SEQ ID No. 33)

```

1  ATGAAAAAAT  GCTTCTGTC  TTTATTGGGA  TTGCTGCGGC  TGACAGTTCA
   TACTTTTTTA  CCAAACACAC  AAATAACCCT  AACCAGCGCG  ACTGTCAAGT
51  GGCAGCCGAC  AGTCGCCCG  CCTTTCCCG  GATCGTGAIG  TTCGGCGACA
   CCGTCGGCTG  TCAGCGGGG  GGAAAGGGC  CTAGCACTAC  AAGCCGCTGT
101  GCCTCTCGGA  TACCGGCAAA  ATGTACAGCA  AGATGCGCG  TTACCTCCCG
   CGGAGAGGCT  ATGGCCGTT  TACATGTCT  TCTACGCGCC  AATGGAGGGG
151  TCCAGCCCGC  CCTACTATGA  GGGCCGTTT  TCCAACGGAC  CCGTCTGGCT
   AAGTCGGGGC  GGATGATACT  CCCGGCAAAG  AAGTTGCCCT  GGCAGACCGA
201  GGAGCAGCTG  ACCAACAGT  TCCCGGGTCT  GACCATCGCC  AACGAAGCGG
   CCTCGTCGAC  TGTTTGTCA  AAGGCCCAGA  CTGCTAGCGG  TTGCTTCGCC
251  AAGCGGCTGC  CACTGCCGTG  GCTTACAACA  AGATCTCTG  GAATCCCAAG
   TTCCGCCACG  GTACGGCAC  CGAATGTGT  TCTAGAGGAC  CTTAGGGTTC
301  TATCAGGTCA  TCAACAACCT  GGACTACGAG  GTCACCCAGT  TCTTGCAAG
   ATAGTCCAGT  AATTGTTGGA  CCTGATGCTC  CAGTGGGTCA  AGAACGTCTT
351  AGACAGCTTC  AAGCCGGAC  ATCTGGTGAT  CCTCTGGGTC  GGTGCCAATG
   TCTGTGGAAG  TTCGGCCTGC  TAGACCACTA  GGAGACCCAG  CCACGGTTAC
401  ACTATCTGGC  CTATGGCTGG  AACACGGAG  AAGATGCCAA  GCGGGTTCGC
   TGATAGACCG  GATACCGACC  TTGTGCCTCG  TCCTACGGTT  CGCCCAAGCG
451  GATGCCATCA  GCGATCGGC  CAACCGCATG  GTACTGAAG  GTGCCAAGCA
   CTACGGTAGT  CGCTACGCC  GTTGGCGTAC  CATGACTTGC  CACGGTCTGT
501  GATACTGCTG  TTCACCTGC  CGGATCTGG  CCAGAACCCG  TCAGCTCGCA
   CTATGACGAC  AAGTTGGAC  GCCTAGACCC  GGTCTTGGG  AGTCGAGCGT
551  GTCAGAAGGT  GGTGAGGCG  GTCAGCCATG  TCTCCGCTA  TCACAACCAG
   CAGTCTTCCA  CCAGCTCCG  CAGTCGGTAC  AGAGCGGAT  AGTGTGCTC
601  CTGCTGCTGA  ACCTGGCAG  CCAGCTGGC  CCCACCGCA  TGGTAAAGCT
   GACGACGACT  TGGACCGTC  GGTGACCGG  GGGTGGCGT  ACCATTTCGA
651  GTTCGAGATC  GACAAGCAAT  TTGCCGAGAT  GCTGCGTGT  CCGCAGAACT
   CAAGCTCTAG  CTGTTCGTT  AACGGCTCTA  CGACGACTA  GCGTCTTGA
701  TCGGCCTGAG  CGACGTCAG  AACCCCTGCT  ACGACGGCG  CTATGTGTGG
   AGCCGGACTC  GCTGCACTC  TTGGGRCGA  TGCTGCCGCC  GATACACACC
751  AAGCCGTTG  CCACCCGAG  CGTCAGCACC  GACCGCCAG  TCTCCGCTT
   TTCGGCAAAC  GGTGGCGTC  GCAGTCGTGG  CTGGCGGTC  AGAGCGGAA
801  CAGTCCGAG  GAACCCCTG  CCATCGCCG  CAACCCGTC  CTGGCAGAG
   GTCAGGCGTC  CTTGCGGAG  GGTAGCGCC  GTTGGCGAC  GACCGTGTCC
851  CCGTTGCCAG  TCCTATGCC  CGCCGACCG  CCAGCCCTT  CAACTGTGAG
   GGCAACGGTC  AGGATACGG  GCGGCTCGC  GGTGGGGGA  GTTGACACTC
901  GGCAAGATGT  TCTGGATCA  GGTACACCC  ACCACTGTC  TGCACGACG
   CCGTCTACA  AGACCTAGT  CCATGTGGC  TGGTGACGC  ACGTGGCTG
951  CCTGAGCGAG  CGCGCCGCA  CCTTCATCG  GAACCACTG  GAGTCTCTG
   GGACTCGCTC  GCCCGGCGT  GGAAGTAGC  CTTGGTCAI  CTCAGGAGC
1001  CCCAC TGA
      GGGTG ACT

```

图 29

(SEQ ID No. 34)

```
1 MKKWFVCLLG LIALTVQAAD TRPAFSRIVM FGDSLSDTGK MYSKMRGYLP
51 SSPPYYEGRF SNGPVWLEQL TKQFPGLTIA NEAEGGATAV AYNKISWNEK
101 YQVINLDYE VTQFLQKDSF KPDDLVLWV GANDYLAYGW NTEQDAKRVR
151 DAISDAANRM VLNQAKQILL FNLFDLGQNE SARSQKVVEA VSHVSAYHNK
201 LLLNLARQLA PTGMVKLFEL DKQFAEMLRD PQNFGLSDVE NPCYDGGYVW
251 KPFATRSVST DRQLSAFSPQ ERLAIAGNPL LAQAVASPMA RRSASPLNCE
301 GKMFWDQVHP TTVVHAALSE RAATFIETQY EFLAHG*
```

图 30

## (SEQ ID No. 35)

```

1  ATGAAAAAAT GGTTTGTTTG TTTATTGGGG TTGATCGCGC TGACAGTTCA
   TACTTTTTTA CCAAACAAAC AAATAACCCC AACTAGCGCG ACTGTCAAGT
51  GGCAGCCGAC ACTCGCCCGG CCTTCTCCCG GATCGTGATG TFCGGCGACA
   CCSTCGGCTG TGAGCGGGGC GGAAGAGGGC CTAGCACTAC AAGCCGCTGT
101 GCCTCTCCGA TACCGGCAAA ATGTACAGCA AGATGCGGGG TTAACCTCCC
   CGGAGAGGCT ATGGCCGTTT TACATGTGCT TCTACGCGCC AATGGAGGGG
151 TCCAGCCCGC CCTACTATGA GGGCCGTTT TCCAACGGAC CCGTCTGGCT
   AAGTCGGGCG GGATGATACT CCCGGCAAAG AGGTTCCTG GGCAGACCGA
201 GGAGCAGCTG ACCAAGCACT TCCCAGGCT GACCATCGCC AACGAAGCGG
   CCTCGTCGAC TGGTTCGTCA AGGGCCAGA CTGGTAGCGG TTGCTTCGCC
251 AAGCCGGTGC CACTGCCGTG GCTTACAACA AGATCTCCTG GAATCCCAG
   TTCCGCCAG GTGACGGCAC CGAATGTGT TCTAGAGGAC CTTAGGGTTC
301 TATCAGGTCA TCAACAACCT GGACTACGAG GTCACCCAGT TCTTGAGAA
   ATAGTCCAGT AGTTGTTGGA CCTGATGCTC CAGTGGGTCA AGAACGTCTT
351 AGACAGCTTC AAGCCGGAGC ATCTGGTGAT CCTCTGGGTC GGTGCCAATG
   TCTGTGGAAG TTCGGCCTGC TAGACCACTA GAGACCCAG CCACGGTTAC
401 ACTATCTGGC ATATGGCTGG AATACGGAGC AGGATGCCAA GCGAGTTCGC
   TGATAGACCG TATACCGACC TTATGCCTCG TCCTACGGTT CGCTCAAGCG
451 GATGCCATCA GCGATCGGC CAACCGCATG GACTGACG GTGCCAAGCA
   CTACGGTAGT CGCTACGGCG GTTGGCGTAC CATGACTTGC CACGGTTCGT
501 GATACTGCTG TTCAACCTGC CGGATCTGGG CCAGAACCCG TCAGCCCGCA
   CTATGACGAC AAGTTGGAGC GCCTAGACCC GGTCTTGGGC AGTCGGGCGT
551 GTCAGAAAGT GGTGAGGCG GTCAGCCATG TCTCCGCTA TCACAACAA
   CAGTCTTCCA CCAGCTCCGC CAGTCGGTAC AGAGGCGGAT AGTGTGTGTC
601 CTGCTGCTGA ACCTGGCAGC CCAGCTGGCC CCCACCGCA TGGTAAAGT
   GACGACGACT TGGACCCTGC GGTGACCCGG GGTGGCCGT ACCATTTCGA
651 GTTCGAGATC GACAAGCAAT TTCCCGAGT GCTGCGTAT CCGCAGACT
   CAAGCTCTAG CTGTTGTTA AACGGCTCTA CGACGCACTA GGCGTCTGA
701 TCGCCCTGAG CGACSTCGAG AACCCCTGCT ACGACGGCGG CFATGTGTGG
   AGCCGGACTC GCTGCAGCTC TTGGGGAGA TGCTGCCGCC GATACALACC
751 AAGCCGTTG CCACCCGAGC GTCACACCC GACCGCCAGC TCTCCGCTT
   TTCGGCAAAC GGTGGGCGTC GCAGTCTGTG CTGGCGGTCB AGAGGCGGAA
801 CAGTCCGAG GAACGCTCG CCATCGCCGG CAACCCGCTG CTGGCACAGG
   GTCAGGCGTC CTGCGGAGC GGTAGCGGCC GTTGGCGAC GACCGTGTCC
851 CCGTTCAGC TCCTATGGCC CGCCGAGCG CCAGCCCCCT CAACTGTGAG
   GGCAACGGTC AGGATACCGG GCGGCGTGC GGTGCGGGGA GTTGACACTC
901 GGCAAGATGT TCTGGATCA GGTACACCCG ACCACTGTG TGACCGCAGC
   CCGTCTACA AGACCCTAGT CCATGTGGC TGGTGACAGC ACGTGCCTGC
951 CCTGAGCGAG CGCGCCGCA CCTTCATCA GACCCAGTAC GAGTTCCTCG
   GGACTCGCTC GCGCGGCGT GGAAGTAGCT CTGGTCTATG CTCAGGAGC
1001 CCCACGGATG A
      GGTGCCCTAC T

```

图 31

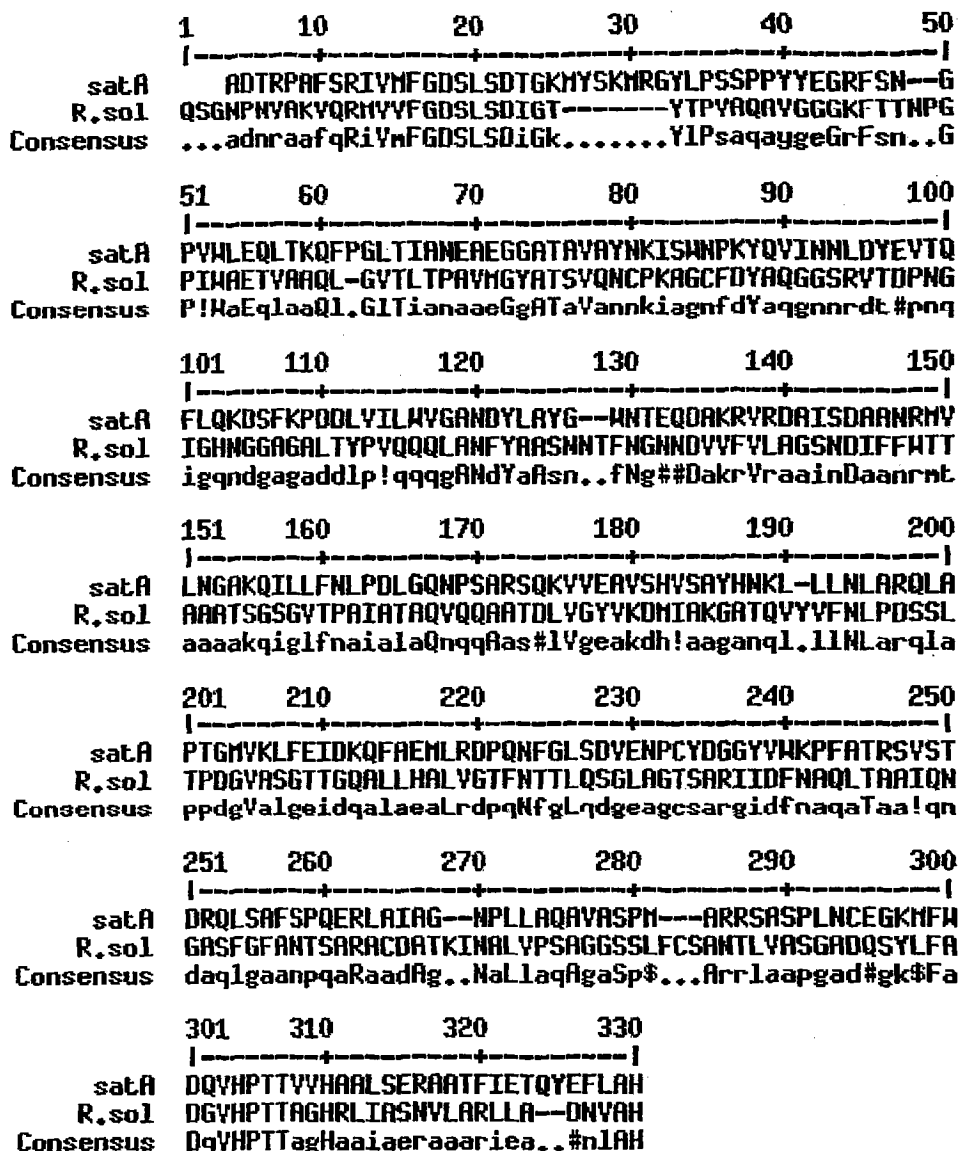


图 32

▼

Pfam	*->ivaFGDSltDggg.....ayygdsgggwgagladrtsla..rlrargrvdv	
Srim1	38 YVALGDSYSSevg.....agSYDSSGCSCKRSTRSYTALWAS..-----HTOTRF	81
Scoe1	5 YVAVGDSFTTEG.....--VGDPGPDGAFVGMADRLAVLL..ADRPEGDFTY	47
Scoe2	10 LVAVGDSFTTEG.....--MSDLLPDSYRGMADLLATRM..--AARSPGFRY	50
Scoe3	239 VVAVGDSITDG.....ARSQSDANHRWTDVLAARLEEA..GDGRDTPRYSV	283
Scoe4	75 LNMIGDSTAG.....--QGVHRAGQTPGALLASG..LAAVAERPVRL	113
Scoe5	66 VAAVGDSITRFD.....acAVLSDCFEVSWSATGSSAKVDSLAVr.LLGAADAEHS	116
Ahyd1	28 IVMFGDSLSDTGkmyskargylpasppyYEGRFSGNPFVWLEQLTNEFFGLTIANEAEGGTAVA	91
Asal1	28 IVMFGDSLSDTGkmyskargylpasppyYEGRFSGNPFVWLEQLTKQF.....PGLTI	79
Ahyd2	40 <u>IVMFGDSLSDTGkmyskargylpasppyYEGRFSGNPFVWLEQLTKQFFGLTIANEAEGGTAVA</u>	103

Pfam	fnrgisGrtsdGrIvvdarIvatllFlaqfIglNlpPYLsgdflrGANFAsagAtIlgtsIipflni	
Srim1	82 MFTACSGAR.....-----	90
Scoe1	48 TMLAVRGRL.....-----	56
Scoe2	51 ANLAVHGKL.....-----	59
Scoe3	284 VNEGISGNR.....-----	292
Scoe4	114 GSVQAQPGAC.....-----	122
Scoe5	117 WNYAVTGAR.....-----	125
Ahyd1	92 YNKISWNPK.....-----	100
Asal1	80 ANEAEGGAT.....-----	88
Ahyd2	104 YNKISWNPK.....-----	112

▼

Pfam	QvqFkdFkskvlrge.....IglIqellrIvvpvldakapdlvtimIGtNDL...itvakfopks	
Srim1	91 -----TGDVLAQKLPVNSGTDLVSIITIGGNDagfaDTMTTCNLOG	131
Scoe1	57 -----LDQIVAEQVERVVG LAPDLVSPAAGNDI...-----I-----	86
Scoe2	60 -----IGQIVDEQVDVAAMGADVITLVGGLNDT...-----	88
Scoe3	293 -----LLTSRPRGPA.....DNPSGLSRFQBDVLERTNVKAVVVVLGVNDV...-----	333
Scoe4	123 -----SDDLDRQVALVLAEPDRVDPICVIMVGANDV...-----	153
Scoe5	126 -----MADETAQVTRAAQREPELVAVMAGANDA...-----CR	155
Ahyd1	101 -----YQVI.....NNLDYEVTFQFLQKDSFKPDDLVLWVGANDY...-----LA	137
Asal1	89 -----AVAYNKISWNPKyqyNNLDYEVTFQFLQKDSFKPDDLVLWVGANDY...-----LA	137
Ahyd2	113 -----YQVI.....NNLDYEVTFQFLQKDSFKPDDLVLWVGANDY...-----LA	149

Pfam	.....tksdrvsvpefrdnrlkIkrLrsangariIiilitvlInlpI.....plGCl	
Srim1	132 esaclarIARARAYIQTLPAQLDQVDAIDSRAPAA-----QVVVLGYF.....	176
Scoe1	87 -----RPGTDPDRVAERPELAVAALT-AAAGTVLVTTGFDTRGVF.....	125
Scoe2	89 -----LRPKCDMAKVRDILLTOAVERLAEHCEQLVIMRSP.....	122
Scoe3	334 -----LNSPELADRDAILTGLRTLVDRAHARGLRVVGATITPFGGYGG.....	376
Scoe4	154 -----THRMPTATRSVRHLSSAVRRLR-TAGAEVVVGTCPDLGTIE.....	192
Scoe5	156 -----SITTSAMTFVADFRQAQFEAMATLR-KKLPKAQVYVSSIPLDKRLwsggrtnplgkQVWKL	214
Ahyd1	138 -----YGMNTEQDAKVRDAISDAANRMV-LNGAK-----QILLFNLP.....	174
Asal1	138 -----YGMNTEQDAKVRDAISDAANRMV-LNGAK-----QILLFNLP.....	174
Ahyd2	150 -----YGMNTEQDAKVRDAISDAANRMV-LNGAK-----QILLFNLP.....	186

Pfam	pq.kIalalaseknvdatgclerIneavadynealrelaeI.ek.l.q.aqlrkdgIpdIksanvpy	
Srim1	177 --.RFYKLGSCAVGLSEKSRATNAAADINAVTAKRA--...-----ADHGFAF	219
Scoe1	126 --.-----VLKHLRGKIATYNGHVRATA--...-----DRYGCPV	152
Scoe2	123 --.-----GRQGFVLEFRFRMEALFAVIDDLA--...-----GRHGAVV	154
Scoe3	377 --.YTEARETRQEVWEEIRSGRVFDTVVDFDKALRDPY--...-----	412
Scoe4	193 --.-----RVRQFLRWLARRASrQLAAAQTIGAVEQGGRTVSL	227
Scoe5	215 GLGFSMLGDADSLDAAATLRRNTVDRVADYMEVLBEVC--...-AkDRRCRSDDGAVHEFRFGT	273
Ahyd1	175 --.-----DLGQNFARSQRVVEAVSHVSAYHNQLLNLA--...-.RQLAPTGMVKLFEIDKQF	224
Asal1	175 --.-----DLGQNFARSQRVVEAVSHVSAYHNQLLNLA--...-.RQLAPTGMVKLFEIDKQF	224
Ahyd2	187 --.-----DLGQNFARSQRVVEAVSHVSAYHNQLLNLA--...-.RQLAPTGMVKLFEIDKQF	236

Pfam	VDlysifqldgIqnpSayv.y....GFeeT.kACCGyGgr.yMyn.rv.CEnag.l.ck.vtakaC	
Srim1	220 GDVNT.....-----TFAGHEICSGAPwL.HS.VT-----	242
Scoe1	153 LDLWLSRSVQDRRA.....-----	166
Scoe2	155 VDLYGASLADFRM.....-----	168
Scoe3	413 .....-----	413
Scoe4	228 GDLLGPEFAQNFREL.....-----	242

```

Scoe5 274 DQL-----.....----- 276
Ahyd1 225 AEMLRDPQNFGLSDQRNACYgGsyvwKPFASrSASTDSQLSaFNPOeRLaIAGNPILaQAvASPMAA 291
Asa11 225 AEMLRDPQNFGLSDVENPCYdGgyvwKPFATrSVSTDRQLSaFSPQeRLaIAGNPILaQAvASPMAR 291
Ahyd2 237 AEMLRDPQNFGLSDVENPCYdGgyvwKPFATrSVSTDRQLSaFSPQeRLaIAGNPILaQAvASPMAR 303

▼
Pfam .dassyll.atlfdGf.HpsekGykavAeal<-*
Srim1 243 .-----LPVENSyHPTANGQSKGYLPV 263
Scoe1 167 .-----WDADRL.HLSFEGHTRVALRA 106
Scoe2 169 .-----WDVDRL.HLTAEGRHRRVAEAV 188
Scoe3 413 .-DPRRMRsDYDSGDHL.HPGDKGYARMGAVI 441
Scoe4 243 .-----FGPDNY.HPSAEGYATAAMAV 262
Scoe5 277 .-----SHWDWF.HPSVDGQARLAEIA 296
Ahyd1 292 rSASTLNCeGKMFWQV.HPTTVVHAALSERA 322
Asa11 292 rSASPLNCeGKMFWQV.HPTTVVHAALSERA 322
Ahyd2 304 rSASPLNCeGKMFWQV.HPTTVVHAALSERA 334

```

图 33

▼	
Pfam	*->ivafGDSltldggg.....ayygdsgggwgagladrltsla..rlrargrgv
Srim1	38 YVALGDSYSSGVG.....agSYDSSSGSCKRSTKSYPALWAAS.....HTGTRF 81
Scoe1	5 YVAVGDSFTEG-----VGDGPDGAFVGVADRLAVLL..ADRRPEGDFTY 47
Scoe2	10 LVAVGDSFTEG-----MSDLLPDGSYRGWADLLATRM...AARSPGFERY 50
Ahyd1	28 IVMPGDSLSDTGKmyskmrgylpssppyYEGRFSSNGPVWLEQLTNEFPGLTiaNEAEGGPTAVA 91
Asal1	28 IVMPGDSLSDTGKmyskmrgylpssppyYEGRFSSNGPVWLEQLTKQF-----PGLTI 79
Ahyd2	40 <u>IVMPGDSLSDTGKmyskmrgylpssppyYEGRFSSNGPVWLEQLTKQFPGLTiaNEAEGGATAVA</u> 103
▼	
Pfam	fnrgisGrtsdGrivvDarlvatl1FlaqfiGlnlpPYLsgdf1rGANFasagAt1lgtslipflni
Srim1	82 NFTAACSGAR-----90
Scoe1	48 TNLAVRGRL-----56
Scoe2	51 ANLAVRGKL-----59
Ahyd1	92 YNKISWNEK-----100
Asal1	80 ANEAEGGAT-----88
Ahyd2	104 YNKISWNEK-----112
▼	
Pfam	QvqFkdfkskvlelrqa.....lgl1qellrlvvpvldakspdlvtimiGtND1...itvakfgpks
Srim1	91 -----TGDVLAQLTPVNSGTDLVSITIGGNDagfaDTMTTCNLQG 131
Scoe1	57 -----LDQIVAEQVPRVVGVLAPDLVSFAAGGNDI.....I-----86
Scoe2	60 -----IGQIVDEQVDVAAAMGADVITLVGGLNDF.....88
Ahyd1	101 -----YQVI.....NNLDYEVTQFLQKDSFKPDDLVLVWVGANDY.....LA 137
Asal1	89 -----AVAYNKISWNPkyqvyNNLDYEVTQFLQKDSFKPDDLVLVWVGANDY.....LA 137
Ahyd2	113 -----YQVI.....NNLDYEVTQFLQKDSFKPDDLVLVWVGANDY.....LA 149
▼	
Pfam	.....tksdnrvsvpefrdnlrk1krLrsangariii1itlv1ln1pl1GCl
Srim1	132 esaclarIAKARAYIQOTLPAQLDQVYDAIDSRAFAA-----QVVVLGYP-----176
Scoe1	87 -----RPGTDFEVAERFELAVAALT-AAAGTVLVTTIGFDTRGVP-----125
Scoe2	89 -----LRPKCDMARVRDLLTQAVERLAPHCEQLVLMRSP-----122
Ahyd1	138 -----YGNWTEQDAKRVRDAISDAANRMV-LNGAK-----EILLFNLP-----174
Asal1	138 -----YGNWTEQDAKRVRDAISDAANRMV-LNGAK-----QILLFNLP-----174
Ahyd2	150 -----YGNWTEQDAKRVRDAISDAANRMV-LNGAK-----QILLFNLP-----186
▼	
Pfam	pqklalalassknvdatgclerlneavadynealrelaeieklqazlrkdglpdlkeanvpy
Srim1	177 --RFYKLGGSACAVGLSEKSRAAINAAADDINAVTAKRA-----ADHGFAF 219
Scoe1	126 -----VLKHLRGKIATYNGHVRAJA-----DRYCCPV 152
Scoe2	123 -----GRQGPVLERFRPRMEALFAVIDDLA-----GRHGAVV 154
Ahyd1	175 -----DLGQNPSARSQKVVEAVSHVSAYHNQLLLNLA-----RQLAPTGMVKLFEDKQF 224
Asal1	175 -----DLGQNPSARSQKVVEAVSHVSAYHNKLLLNLA-----RQLAPTGMVKLFEDKQF 224
Ahyd2	187 -----DLGQNPSARSQKVVEAVSHVSAYHNQLLLNLA-----RQLAPTGMVKLFEDKQF 236
▼	
Pfam	VDlysisfqdlidg1qnp1sayv.y...GFeet.kaCCGyGgr.yNyn.rv.CGnag.1.ck.vtakaC
Srim1	220 GDVNT-----TFAGHEICSGAPwL.HS.VT-----242
Scoe1	153 LDLWLSLRSVQDRRA-----166
Scoe2	155 VDLYGAQSLADPRM-----168
Ahyd1	225 AEMLRDPQNFGLSDQRNACyGgyvwwKPFASrSASTDSQLSaFNPQeRLaIAGNPLLaQAvASPMMA 291
Asal1	225 AEMLRDPQNFGLSDVENPCYdGgyvwwKPFATrSVSTDRLQLSaFSPQeRLaIAGNPLLaQAvASPMAR 291
Ahyd2	237 AEMLRDPQNFGLSDVENPCYdGgyvwwKPFATrSVSTDRLQLSaFSPQeRLaIAGNPLLaQAvASPMAR 303
▼	
Pfam	.dassyl1.at1fwDgf.HpsekGykavAeal<*
Srim1	243 -----LPVENSyHPTANGQSKGYLPV 263
Scoe1	167 -----WDADRL.HLSPEGHTRVALRA 186
Scoe2	169 -----WDVDRL.HLTAEGHRRVAVAV 188
Ahyd1	292 rSASTLNCeGKMFWDQV.HPTTVVHAALSERA 322
Asal1	292 rSASPLNCeGKMFWDQV.HPTTVVHAALSERA 322
Ahyd2	304 rSASPLNCeGKMFWDQV.HPTTVVHAALSERA 334

图 34

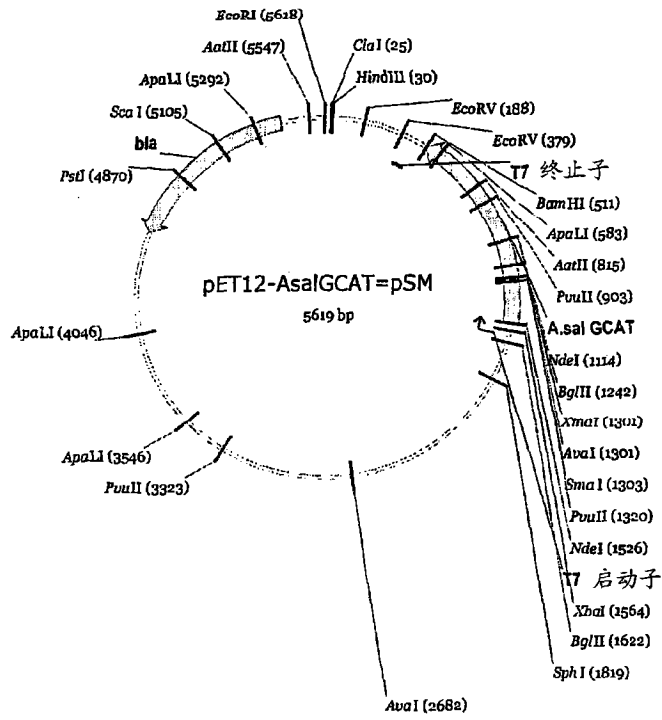


图 35

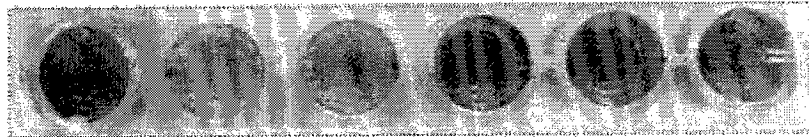


图 36



图 37

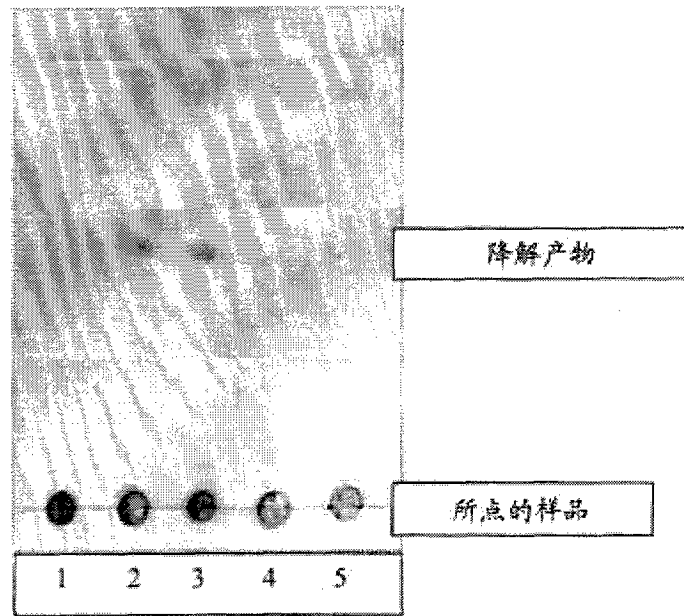


图 38

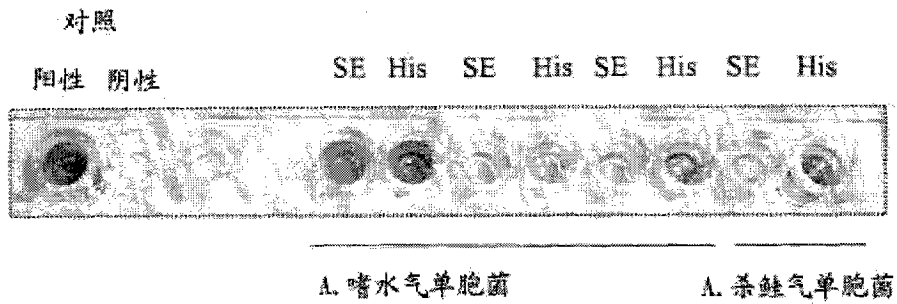


图 39

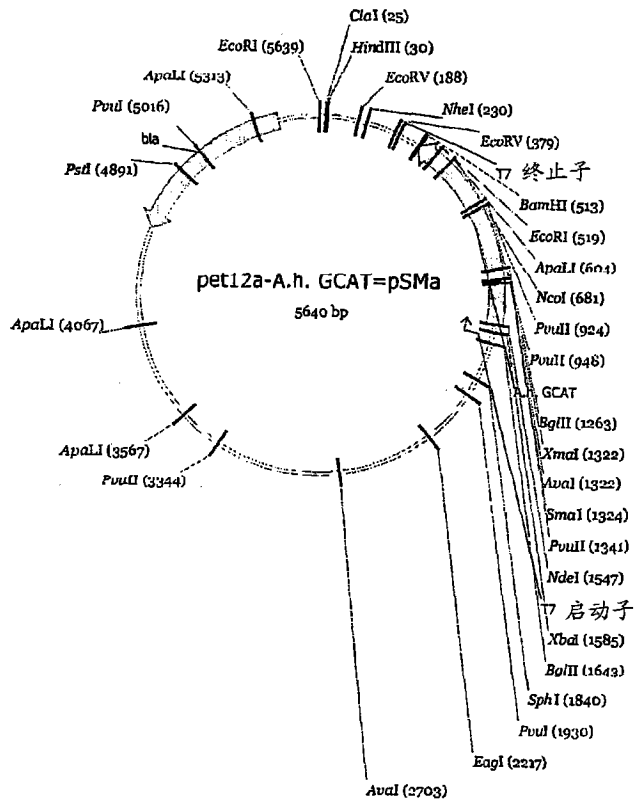


图 40

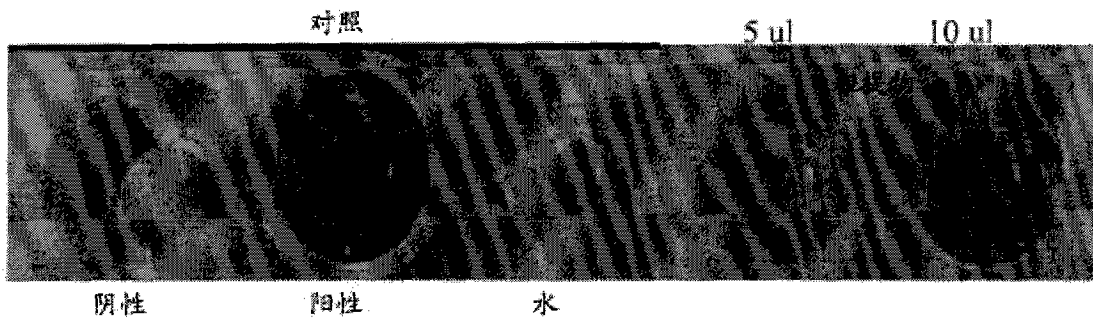


图 41

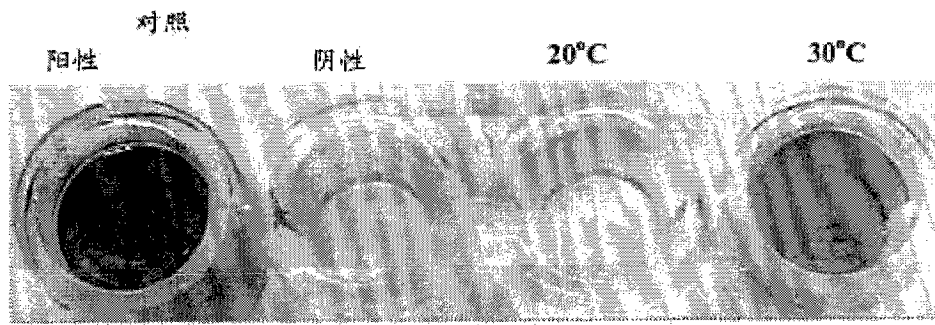


图 42

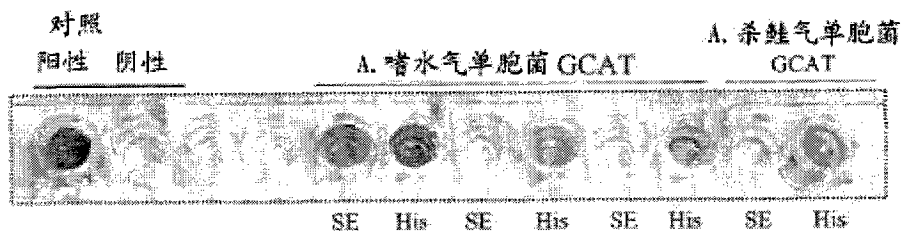


图 43

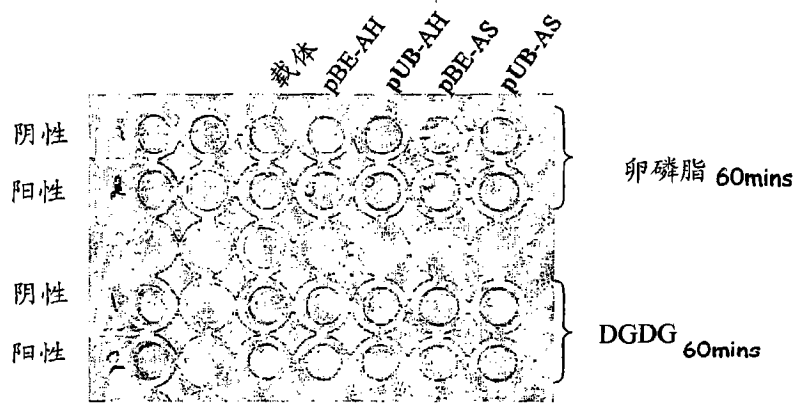


图 44

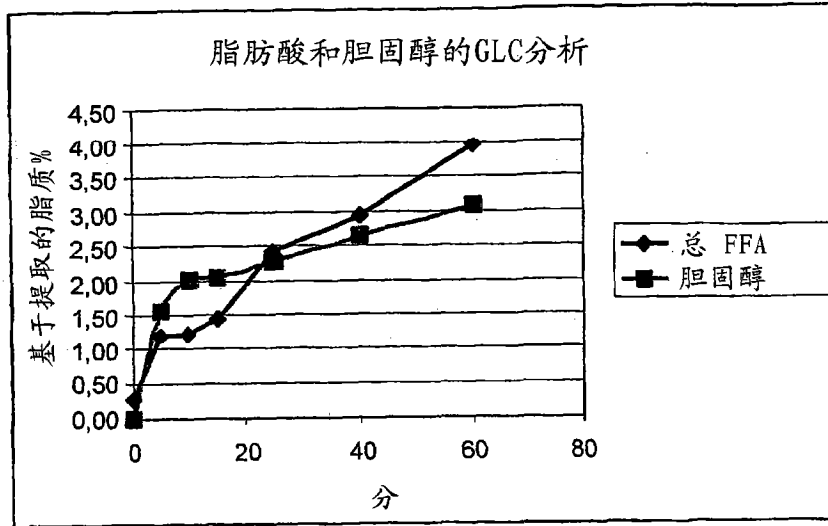


图 45

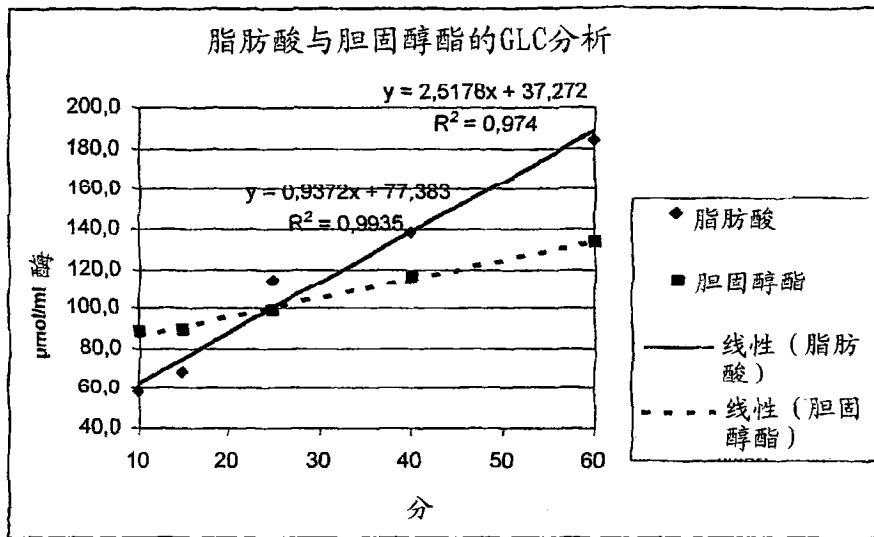


图 46

(SEQ ID No. 36)

```

1  MFKFKKNFLV GLSAALMSIS LFSATASAAS ADSRPAFSRI VMFGDSLSDT
51  GKMYSKMRGY LPSSPPYYEG RFSNGPVWLE QLTKQFPGLT IANEAEAGGAT
101 AVAYNKISWN EKQVINNLD YEVTQFLQKD SFKPDLLVIL WVGANDYLAY
151 GWNTEQDAKR VRDAISDAAN RMVLNGAKQI LLENLPDLGQ NPSARSQKVY
201 EAVSHVSAYH NOLLNLLARO LAPTMGVKLF EIDKQFAEML RDPQNFGLSD
251 VENPCYDGGY VWKPEATRSV STDRQLSAFS POERLAIAGN PLLAQAVASP
301 MARRSASPLN CEGKMFWDQV HPTTVVHAAL SERAATFIAN QYEFLAH**
    
```

图 47

(SEQ ID No. 54)

1 ATGTTAAGT TTAAAAGAA TTTCITAGT GGAATATCG CAGCTTTAAT  
 TACAATTCA AATTTTCTT AAAGATCAA CCTAATAGCC GTCGAAATTA  
 51 GAGTATTAGC TTGTTTCGG CAACCGCCTC TGCAGCTAGC GCCGACAGCC  
 CTCATAATCG AACAAAAGCC GTTGGCGGAG ACGTCGATCG CGGCTGTCGG  
 101 GTCCCGCCTT TTCCCGGATC GTGATGTTCG GCGACAGCCT CTCCGATACC  
 CAGGGCGGAA AAGGGCCTAG CACTACAAGC CGCTGTGCGA GAGGCTATGG  
 151 GGCAAAATGT ACAGCAAGAT GCGCGGTAC CTCCCTCCA GCCCGCCCTA  
 CCGTTTTACA TGTCTTCTA CCGCCCAATG GAGGGGAGGT CGGGCGGGAT  
 201 CTATGAGGC CGTTTCTCCA ACGGACCCGT CTGGCTGGAG CAGCTGACCA  
 GATACTCCC GCAAAGAGGT TGCCTGGCA GACCGACCTC GTCGACTGGT  
 251 AACAGTCCC GGGTCTGACC ATCGCCAACG AAGCGGAAGG CGGTGCCACT  
 TTGTCAAGGG CCCAGACTGG TACGGTTCG TTCGCCCTCC GCCACGGTGA  
 301 GCCGTGCTT ACAACAAGAT CTCTGGAAT CCCAAGTATC AGGTATCAA  
 CCGCACCGAA TGTGTCTA GAGGACCTA GGGTTCATAG TCCAGTAGTT  
 351 CAACCTGGC TACGAGTCA CCCAGTCTT GCAGAAAGC AGCTCAACG  
 GTTGGACCTG ATGCTCCAGT GGGTCAAGAA CGTCTTCTG TCGAAGTTCG  
 401 CGGACGATCT GGTGATCCTC TGGTCCGGT CCAATGACTA TCTGGCCTAT  
 GCCTGTAGA CCACTAGGAG ACCCAGCCAC GGTACTGAT AGACCGGATA  
 451 GGCTGGAACA CCGAGCAGGA TGCCAAGCGG GTTCGCGATG CCATCAGCGA  
 CCGACCTGT GCCTGCTCT ACGGTTCGCC CAAGCGCTAC GGTAGTCCGT  
 501 TCGGCCAAC CGCATGGTAC TGAACGGTGC CAAGCAGATA CTCTGTTC  
 ACGCCGGTTC GCGTACCATG ACTTGCCAGG GTTCGTCTAT GAGCACAAGT  
 551 ACCTGCCGGA TCTGGCCAG AACCCGTGAG CTCGCACTCA GAAGTGGTTC  
 TGGACGGCT AGACCCGGTC TTGGCAGTC GAGCGTCAAT CTCCACCAG  
 601 GAGCGGTC ACCATGTCTC GGCCTATCAC AACCACTGC TGCTGAACCT  
 CTCCCCAGT CGGTACAGAG CGCGATGTC TTGGTCCAGC ACGACTTGG  
 651 GGCAGCCAG CTGGCCCCA CCGCATGGT AAAGCTGTC GAGATCGACA  
 CCGTCCGGTC GACCGGGGT GCGCGTACCA TTTCGACAAG CTCTAGCTGT  
 701 AGCAATTTGC CGAGATGCTG CGTGATCCG AGAAGTTCGG CCTGAGCGAC  
 TCGTTAAACG GCTCTAGAC GCACTAGGCG TCTTGAAGCC GGAATCCGCT  
 751 GTCGAGAACC CCTGCTACA CCGCGGCTAT GTGTGGAGC CGTTGCCAC  
 CAGCTTGG GCAGATGCT GCGCGGATA CACACCTTC GCAAACGGT  
 801 CCGCAGCTC AGCACCACC GCCAGTCTC GCGCTTCAAT CCGCAGGAAC  
 GCGTCCAG TCGTGGCTGG CGGTCCAGAG CCGGAAGTCA GCGTCCCTG  
 851 GCCTCGCAT CGCGGCAAC CCGTGTCTG CACAGGCCGT TGCCAGTCTT  
 CCGAGCGTA GCGCCGTTG GCGACGACC GTGTCCGCA ACGTCCAG  
 901 ATGGCCGCC GCAGGCCAG CCCCCTAAC TGTGAGGCA AGATGTCTG  
 TACCGGGCG CGTCCGGTC GGGGAGTTG AACTCCCGT TCTACAAGC  
 951 GGATCAGTA CACCGACCA CTGTGTGCA GCGAGCCTG AGCGAGCGG  
 CCTAGTCCAT GTGGGCTGT GACAGCAGT GCGTCCGGAC TCGTCCGCG  
 1001 CCGCCACCTT CATCGCAAC CAGTACGAT TCTCGCCA CTGATGA  
 GCGGTGGAA GTAGGCTTG GTCATGCTA AGGAGCGGT GACTACT

图 48

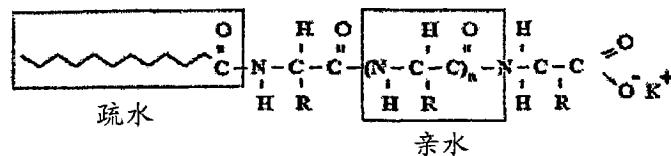


图 49

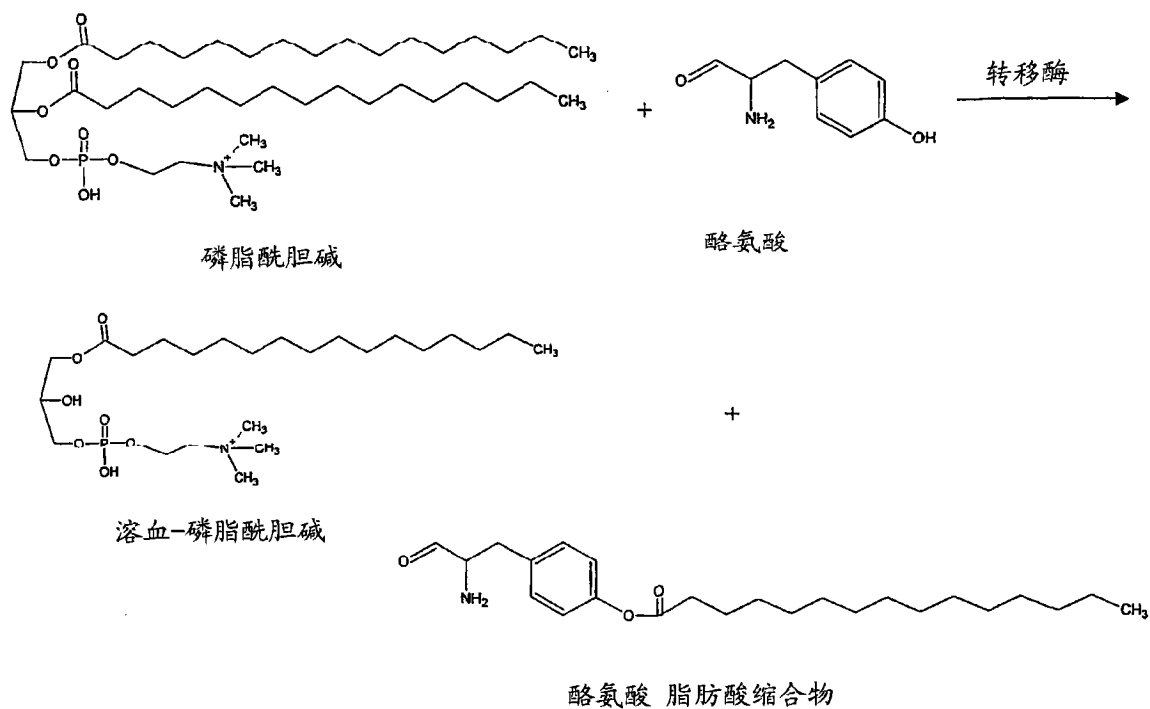


图 50

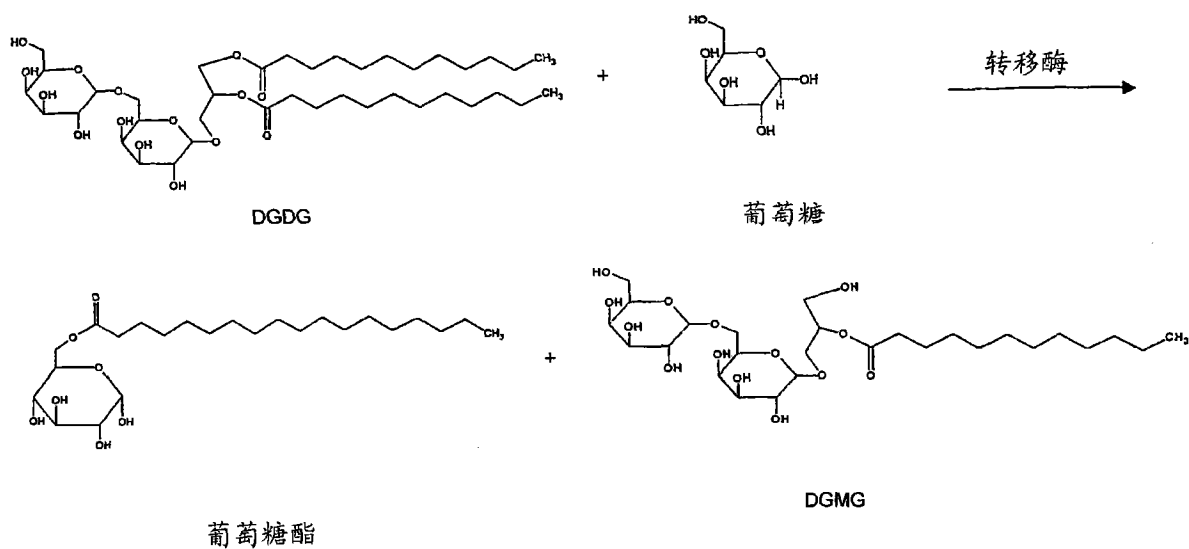


图 51