



(12) 发明专利

(10) 授权公告号 CN 102382805 B

(45) 授权公告日 2015. 03. 25

(21) 申请号 201110259204. 3

C12P 7/38(2006. 01)

(22) 申请日 2007. 12. 07

C12R 1/72(2006. 01)

(30) 优先权数据

A2027/2006 2006. 12. 07 AT

(62) 分案原申请数据

200780045107. 6 2007. 12. 07

(73) 专利权人 IEP 有限责任公司

地址 德国威斯巴登

(72) 发明人 A·古普塔 A·申特谢尔

M·博布科瓦

(74) 专利代理机构 中国专利代理(香港)有限公

司 72001

代理人 柯珂 李炳爱

(56) 对比文件

CN 1091472 A, 1994. 08. 31,

WO 2004111083 A2, 2004. 12. 23,

Vinod Bihari, et al. Biotransformation of 3-Methoxy-8, 14-seco-1, 3, 5(10), 9(11) Estratetraen-14, 17-dione (Secodione) to Its 17p-HydroxyDerivative (Secol) Using Immobilized Yeast Cells (Pichia farinosa Y-118). 《Biotechnology and Bioengineering》. 1985, 第 27 卷 1347-1352.

审查员 王瑶

(51) Int. Cl.

C12N 9/02(2006. 01)

权利要求书1页 说明书11页

序列表21页

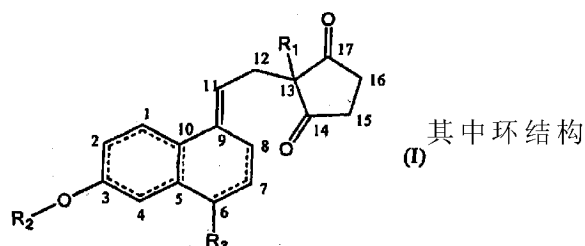
(54) 发明名称

具有氧化还原酶活性的多肽及其用途

(57) 摘要

本发明涉及通式 I 的开环二酮 (Secodion) 衍生物的对映体选择性酶促还原方法

还原酶/脱氢酶还原。根据本发明, 反应批料中使用的开环二酮衍生物浓度 $\geq 10\text{g/l}$, 并且由氧化还原酶/脱氢酶形成的氧化的 NAD 或 NADP 辅因子被连续再生。



不包括杂原子、包括一个杂原子或包括数个杂原子, R_1 是氢或 C_1-C_4 烷基, R_2 是氢、 C_1-C_8 烷基或现有技术已知的对 OH 的保护基, 诸如酯, R_3 是氢、甲

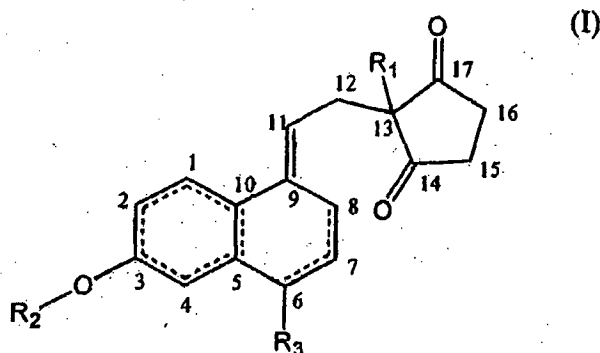
基或卤化物所述结构单元 (A)



(A) 表

示苯环或具有 0、1 或 2 个 C-C 双键的 C_6 环, 任选地在第 6/7 或 7/8 位包括双键, 和第 1、2、4、5、6、7、8、9、11、12 和 16 位的碳独立地被氢、 C_1-C_4 烷基、卤化物或苯基取代, 在所述方法中在 NADH 或 NADPH 作为辅因子存在的条件下开环二酮衍生物被氧化

1. 具有氧化还原酶活性的多肽, 其特征在于所述多肽
 - a) 为氨基酸序列 SEQ ID NO :3 或
 - b) 由核酸序列 SEQ ID NO :8 编码并且具有在辅因子 NADH 或 NADPH 的存在下还原通式 I 的开环二酮衍生物的活性



其中

R_1 是氢或 C_1-C_4 烷基,

R_2 是氢、 C_1-C_8 烷基或现有技术已知的对 OH 的保护基,

R_3 是氢、甲基或卤化物,

所述结构单元



表示苯环或具有 0、1 或 2 个 C-C 双键的 C_6 环,

任选地在第 6/7 或 7/8 位包括双键, 和

第 1、2、4、5、6、7、8、11、12 和 16 位的碳独立地被氢、 C_1-C_4 烷基、卤化物或苯基取代。

2. 根据权利要求 1 的多肽, 其中所述对 OH 的保护基是酯。

3. 根据权利要求 1 或 2 的多肽, 其中所述多肽获自木兰假丝酵母 (*Candida magnoliae*)。

具有氧化还原酶活性的多肽及其用途

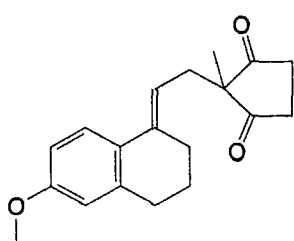
[0001] 本案是申请号为 200780045107.6 的中国专利申请的分案申请。

[0002] 本发明涉及通式 I 的开环二酮 (Secodion) 衍生物的对映体选择性酶促还原方法, 其中在 NADH 或 NADPH 作为辅因子存在的条件下开环二酮衍生物被氧化还原酶 / 脱氢酶还原。

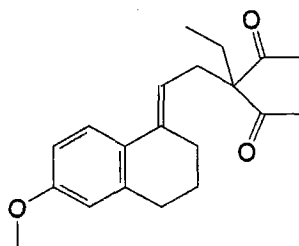
[0003] 甾族激素的工业制备存在两条彼此独立的路线, 即, 一方面, 从天然存在的植物来源的甾族化合物开始, 和另一方面, 从前手性前体通过对映体选择性合成的完全合成方式。在这两条路线中, 甾族化合物全合成是日益重要的, 尤其是因为它还允许引入天然存在的甾族化合物不具有的结构单元。

[0004] 对映体纯的甾族化合物的全合成的关键组分是通式 I 的化合物, 其还被称为开环甾族化合物 (Secosteroides), 8,14- 开环 - 甾 - 四烯 -14,17- 二酮 (8,14-seco-gona-tetraen-14,17-dione) 或开环二酮。这组的具体代表物是诸如化合物甲基开环二酮 (式 II, 13- 甲基 -3- 甲氧基 -8,14- 开环 - 甾 -1,3,5(10),9(11)- 四烯 -14,17- 二酮) 和乙基开环二酮 (式 III, 13- 乙基 -3- 甲氧基 -8,14- 开环 - 甾 -1,3,5(10),9(11)- 四烯 -14,17- 二酮), 从这些化合物例如可以制备具有药理学活性的化合物乙炔基雌甾二醇 (式 IV) 和炔诺孕酮 (式 V)。

[0005]



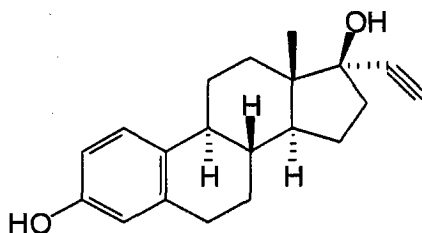
式 II



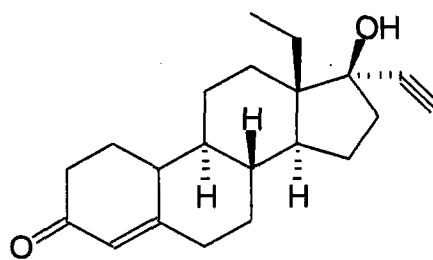
式 III

[0006]

[0007]



式 IV



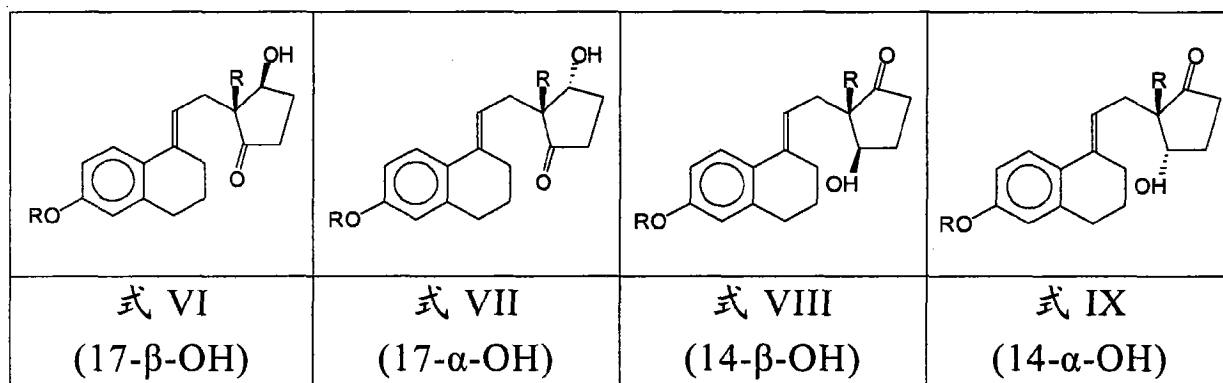
式 V

[0008]

[0009] 制备对映体纯的甾族化合物的关键步骤是通过将酮基团之一对映体选择性还原成羟基, 使式 I (例如, II 和 III) 的化合物转化为具有预先形成的不对称 C-13 的光学活性化合物。随后可以通过环化将获得的光学活性羟基开环甾族化合物 (开环醇 (Secole), 式 VI 到 IX) 进一步加工为手性甾族化合物, 同时保持手性。

[0010] 通过式 I 化合物的酮基团的对映体选择性还原, 理论上可以形成 4 种光学活性化合物 (式 VI 到 IX)。

[0011]



[0012] 式 VI 的化合物（其中在第 17 位羟基显示 β-构型）是尤其具有经济利益的，因为它们产生天然雌酮的衍生物。这种化合物还被称为 17-β-羟基开环甾族化合物。

[0013] 尤其在十九世纪六七十年代充分研究了借助于不同的微生物对通式 I 的开环二酮衍生物的立体选择性还原。由此，能够证明假丝酵母属 (*Candida*)、德巴利酵母属 (*Debaryomyces*)、克勒克酵母属 (*Kloeckera*)、毕赤酵母属 (*Pichia*)、隐球酵母属 (*Cryptococcus*)、红酵母属 (*Rhodotorula*)、球拟酵母属 (*Torulopsis*) 和汉逊氏酵母属 (*Hansenula*) 的不同酵母菌株能够将开环二酮还原成各种羟基化合物 (US 3616226, US 1252524, US 3616225)。

[0014] 尤其是，酵母属 (*Saccharomyces*) 的酵母，诸如葡萄汁酵母 (*S. uvarum*) 可有利的用于制备例如相应的 17-β-羟基开环甾族化合物 (Kosmol 等人; Liebig's Ann. Chem. 701, 199 (1967))。其他酵母菌株诸如 *Saccharomyces drosophilum* 优选地将开环二酮还原成对应的 14-α-羟基开环甾族化合物 (*Acta microbiol. Acad. Sci. hung.* 22, 463-471 (1975))。此外，还描述了通过苏云金芽孢杆菌 (*Bacillus thuringiensis*) 还原开环二酮形成 14-α-羟基开环甾族化合物 (Kosmol 等人; Liebig's Ann. Chem. 701, 199 (1967))。

[0015] 在世界范围内孕激素和雌激素试剂被广泛用作避孕剂并在激素替代疗法中使用。到目前为止雌激素和孕激素衍生物的大部分合成还基于上述反应原理，其中的关键步骤是将开环二酮对映体选择性还原为相应的 17-β-羟基开环甾族化合物。

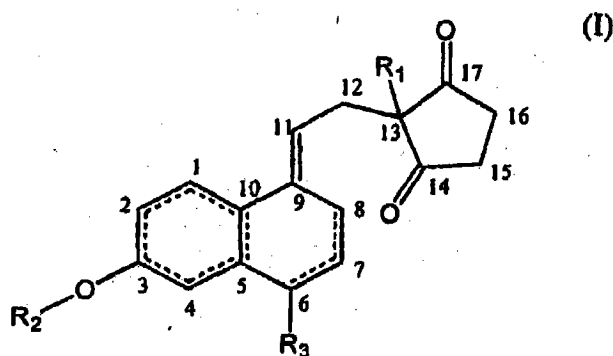
[0016] 在这种情况下，迄今都是利用毕赤酵母属或酵母属的不同酵母菌株以全细胞生物转化形式进行开环二酮衍生物的立体选择性还原。但是，这些方法具有缺点，即只能实现远低于 1%（通常从 1 到 5g/l）的极低底物浓度 (US 3697379; *Current Science*, 1984 年 2 月 5 日, 第 53 卷, 第 3 期, 第 124 页; *Indian Journal of Experimental Biology*, 第 27 卷, 1989 年 8 月, 第 742-743 页)。因此，尤其是大体积反应产物的再处理和分离及大量生物物质的分离变得非常复杂。就本发明人所知，到目前为止还没有分离、鉴定和描述参与还原的酶。并且，还没有鉴定出编码氧化还原酶（通过该氧化还原酶可以还原开环二酮衍生物）的 DNA 序列。

[0017] 因此，本发明的目的是提供一种可以对映体选择性还原通式 I 的开环二酮衍生物（尤其是式 II 和 III 的开环二酮衍生物）的方法。由此尤其还应当可以制备相应的 17-β-羟基开环甾族化合物。

[0018] 第一个方面，根据本发明通过通式 I 的开环二酮衍生物的对映体选择性酶促还原

方法实现所述目的，

[0019]



[0020] 其中环结构不包括杂原子、包括一个杂原子或包括数个杂原子，

[0021] R_1 是氢或 C_1-C_4 烷基，

[0022] R_2 是氢、 C_1-C_8 烷基或现有技术已知的对 OH 的保护基，诸如酯，

[0023] R_3 是氢、甲基或卤化物

[0024] 所述结构单元

[0025]



[0026] 表示苯环或具有 0、1 或 2 个 C-C 双键的 C_6 环，

[0027] 任选地在第 6/7 或 7/8 位包括双键，和

[0028] 第 1、2、4、5、6、7、8、9、11、12 和 16 位的碳独立地被氢、 C_1-C_4 烷基、卤化物或苯基取代，

[0029] 其中在 NADH 或 NADPH 作为辅因子存在的条件下开环二酮衍生物被氧化还原酶 / 脱氢酶还原，

[0030] 该方法的特征在于在反应批料中使用的开环二酮衍生物浓度 $\geq 10g/l$ ，并且由氧化还原酶 / 脱氢酶形成的氧化的辅因子 NAD 或 NADP 被连续再生。

[0031] 该方法代表了开环二酮衍生物的对映体选择性酶促还原相对于现有技术的显著改进。本发明方法允许利用远远超过现有技术描述的浓度范围的游离酶将开环二酮衍生物还原为不同的相应羟基开环甾族化合物。

[0032] 第二个方面，根据本发明通过通式 I 的开环二酮衍生物的对映体选择性酶促还原方法实现上述目的，其中在 NADH 或 NADPH 作为辅因子存在的条件下开环二酮衍生物被氧化还原酶 / 脱氢酶还原，该方法的特征在于所述氧化还原酶 / 脱氢酶

[0033] a) 包括与氨基酸序列 SEQ ID NO :1、SEQ ID NO :2、SEQ ID NO :3、SEQ ID NO :4 或 SEQ ID NO :5 至少 50% 氨基酸相同的氨基酸序列，

[0034] b) 由核酸序列 SEQ ID NO :6、SEQ ID NO :7、SEQ ID NO :8、SEQ ID NO :9 或 SEQ ID NO :10 编码，或

[0035] c) 由严格条件下与 SEQ ID NO :6、SEQ ID NO :7、SEQ ID NO :8、SEQ ID NO :9 或 SEQ ID NO :10 杂交的核酸序列编码。

[0036] 本发明人已经鉴定出能够将开环二酮衍生物还原为羟基开环甾族化合物并且可以在工业规模上重组产生的氧化还原酶。与目前使用的全细胞方法相比,按照本发明的方法可以实现显著更高的底物浓度。

[0037] 在本发明的方法中,具有序列 SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4 或 SEQ ID NO:5 的氧化还原酶或可衍生自这些多肽的多肽可以分别以完全纯化的状态,部分纯化的状态或作为包含多肽 SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4 或 SEQ ID NO:5 的细胞而被使用。由此,所用细胞可以天然、透化或裂解状态提供。优选地,氧化还原酶及可由其衍生的衍生物可分别在合适的宿主生物诸如大肠杆菌 (*Escherichia coli*) 中过表达,并且重组多肽被用于通式 I 的开环二酮衍生物的还原。

[0038] 例如从生物橙色绿屈挠菌 (*Chloroflexus aurantiacus*) DSM 635 的基因组可获得编码具有 SEQ ID NO:1 的多肽的 DNA 序列 SEQ ID NO:6。

[0039] 例如从生物 *Rubrobacter xylophilus* DSM 9941 的基因组可获得编码具有 SEQ ID NO:2 的多肽的 DNA 序列 SEQ ID NO:7。

[0040] 从酵母木兰假丝酵母 (*Candida magnoliae*) CBS 6396 可获得编码具有 SEQ ID NO:3 的多肽的 DNA 序列 SEQ ID NO:8。

[0041] 例如通过同源性筛选从木兰假丝酵母 (*Candida magnoliae*) DSMZ70638 可获得 SEQ ID NO:4 和 SEQ ID NO:5 的氧化还原酶。

[0042] 在严格条件下例如与 SEQ ID NO:6 杂交的核酸序列应被理解为可以通过集落杂交法、噬菌斑杂交法、DNA 杂交法或可比较的方法利用 SEQ ID NO:6 或 SEQ ID NO:6 的部分序列作为 DNA 探针鉴定出的多核苷酸。为了这个目的,在 60°C 0.7-1M NaCl 溶液中,固定在滤器上的多核苷酸与诸如 SEQ ID NO:6 杂交。按照例如分子克隆,实验室手册 (*Molecular Cloning, A Laboratory Manual*), 第二版 (Cold Spring Harbor Laboratory Press, 1989) 或类似出版物的描述进行杂交。随后,在 65°C 用 0.1 到 2 倍 SSC 溶液洗涤滤器,其中 1 倍 SSC 溶液理解为由 150mM NaCl 和 15mM 柠檬酸钠组成的混合物。

[0043] 在上述严格条件下与序列的多核苷酸 SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9 或 SEQ ID NO:10 杂交的多核苷酸应该与多核苷酸序列 SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9 或 SEQ ID NO:10 显示至少 60% 的序列同一性,更好的是同一性为至少 80%,甚至更好的是同一性为 95%。

[0044] 在另一个方面,根据本发明通过通式 I 的开环二酮衍生物的对映体选择性酶促还原方法实现上述目的,其中在 NADH 或 NADPH 作为辅因子存在的条件下开环二酮衍生物被氧化还原酶 / 脱氢酶还原,该方法的特征在于所述氧化还原酶 / 脱氢酶长度为 230 到 260 个氨基酸,并包括选自 [序列 SEQ ID NO:18 到 SEQ ID NO:42] :

[0045] nalvtgasrgig, nalvtggsrgig, nalitggsrgig, nalitgasrgig, nalitggsrgmg, halvtgasrgig,

[0046] gysvtla, gynvtla, gysvtlv, gynvtlv,

[0047] fkgaplpa, fkaaplpa,

[0048] fvsnag, ffsnag, fvcnag, fvanag,

[0049] spialtkal, spvaltkti, spialtktl, spvamtkal, sqialtkal,

[0050] avysask, avysatk,

[0051] pikgwi 和 pisgwi,

[0052] 的一个或数个部分序列。

[0053] 在本发明的方法中, NADH 或 NADPH 被用作辅因子。术语“NADP”理解为烟酰胺腺嘌呤二核苷酸磷酸, “NADPH”理解为还原型烟酰胺腺嘌呤二核苷酸磷酸。术语“NAD”表示烟酰胺腺嘌呤二核苷酸, 术语“NADH”表示还原型烟酰胺腺嘌呤二核苷酸。

[0054] 根据所述方法的优选实施方案, 在反应批料中使用的开环二酮衍生物浓度 $\geq 10\text{g/l}$, 并且由氧化还原酶 / 脱氢酶形成的氧化的辅因子 NAD 或 NADP 被连续再生, 所述氧化还原酶 / 脱氢酶

[0055] a) 包括与氨基酸序列 SEQ ID NO :1、SEQ ID NO :2、SEQ ID NO :3、SEQ ID NO :4 或 SEQ ID NO :5 至少 50% 氨基酸相同的氨基酸序列,

[0056] b) 所述氧化还原酶 / 脱氢酶由核酸序列 SEQ ID NO :6、SEQ ID NO :7、SEQ ID NO :8、SEQ ID NO :9 或 SEQ ID NO :10 编码, 或

[0057] c) 所述氧化还原酶 / 脱氢酶由严格条件下与 SEQ ID NO :6、SEQ ID NO :7、SEQ ID NO :8、SEQ ID NO :9 或 SEQ ID NO :10 杂交的核酸序列编码。

[0058] 根据所述方法另一个优选的实施方案, 反应批料中使用的开环二酮衍生物浓度 $\geq 10\text{g/l}$, 并且由氧化还原酶 / 脱氢酶形成的氧化的辅因子 NAD 或 NADP 被连续再生, 所述氧化还原酶 / 脱氢酶长度为 230 到 260 个氨基酸, 包括选自 [序列 SEQ ID NO :18 到 SEQ ID NO :42] :

[0059] nalvtgasrgig, nalvtggsrgig, nalitggsrgig, nalitgasrgig, nalitggsrgmg, halvtgasrgig, gysvtla, gynvtla, gysvtlv, gynvtlv, fkgaplpa, fkaaplpa, fvsnag, ffsnag, fvcnag, fvanag, spialtkal, spvaltkti, spialtktl, spvamtkal, sqialtkal, avysask, avysatk, pikgwi 和 pisgwi 的一个或数个部分序列。

[0060] 在根据本发明第二和第三个方面的方法中, 由氧化还原酶 / 脱氢酶形成的氧化的辅因子 NAD 或 NADP 优选的被连续再生。

[0061] 根据本发明所有方法的优选实施方案, 通过醇的氧化来再生氧化的辅因子 NAD 或 NADP。

[0062] 在这种情况下, 伯醇和仲醇诸如乙醇, 2- 丙醇, 2- 丁醇, 2- 戊醇, 3- 戊醇, 4- 甲基-2- 戊醇, 2- 己醇, 2- 庚醇, 2- 辛醇或环己醇优选地被用作共底物。以总体积计, 用于再生的共底物比例为 5 到 95% (体积比)。

[0063] 具有通式 $R_xR_y\text{CHOH}$ 的仲醇优选被用于辅因子再生, 其中 R_x 和 R_y 彼此独立的是氢、分枝或未分枝的 C_1 - C_8 烷基并且 $C_{\text{总数}} \geq 3$ 。

[0064] 根据本发明方法另一个优选的实施方案, 还添加氧化还原酶 / 脱氢酶用于辅因子的再生。

[0065] 合适的 NADH 依赖性醇脱氢酶例如可获自面包酵母, 近平滑假丝酵母 (*Candida parapsilosis*, CPCP) (US 5,523,223 和 US 5,763,236, *Enzyme Microb. Technol.*, 1993,15(11):950-8), *Pichia capsulata* (DE10327454.4), 红串红球菌 (*Rhodococcus erythropolis*, RECR) (US5,523,223), 褐色诺卡氏菌 (*Norcardia fusca*) (Biosci. Biotechnol. Biochem., 63(10), 1999, 第 1721-1729 页; *Appl. Microbiol. Biotechnol.*, 2003,62(4):380-6; Epub 2003 年 4 月 26 日) 或赤红球菌 (*Rhodococcus ruber*) (*J. Org.*

Chem., 2003, 68(2): 402-6)。这些醇脱氢酶的合适共底物是, 例如, 已经提及的仲醇诸如 2- 丙醇 (异丙醇), 2- 丁醇, 2- 戊醇, 4- 甲基 -2- 戊醇, 2- 辛醇或环己醇。

[0066] 用于 NADPH 再生的合适仲醇脱氢酶是例如上文描述的及从乳杆菌目生物分离的那些仲醇脱氢酶, 诸如高加索酸奶乳杆菌 (*Lactobacillus kefir*) (US 5, 200, 335), 短乳杆菌 (*Lactobacillus brevis*) (DE 19610984A1; Acta Crystallogr. D. Biol. Crystallogr. 2000 年 12 月; 56Pt 12: 1696-8), 稍小乳杆菌 (*Lactobacillus minor*) (DE 10119274), 肉色明串珠菌 (*Leuconostoc carnosum*) (A 1261/2005, Kl. C12N), 或所述来自布氏热厌氧菌 (*Thermoanaerobium brockii*), *Thermoanaerobium ethanolicus* 或拜氏梭状芽孢杆菌 (*Clostridium beijerinckii*) 的那些仲醇脱氢酶。

[0067] 但是, 原则上其他酶系统也可用于辅因子再生。例如, 可以利用 NAD 或 NADP 依赖性甲酸脱氢酶进行辅因子再生 (Tishkov 等人, J. Biotechnol. Bioeng. [1999] 64, 187-193, Pilot-scale production and isolation of recombinant NAD and NADP specific formate dehydrogenase)。甲酸脱氢酶的合适共底物是, 例如, 甲酸盐诸如甲酸铵、甲酸钠或甲酸钙。

[0068] 本发明方法获得的 TTN (总转换数 = 还原的开环二酮化合物的摩尔数 / 使用的辅因子的摩尔数) 范围通常是 10^2 到 10^5 , 但是优选 $\geq 10^3$ 。

[0069] 根据优选的实施方案, 在含水有机两相系统中实施本发明的方法。

[0070] 因此, 开环二酮衍生物的转化发生在两相系统中, 其包含诸如, 用于辅因子再生的 2- 醇、氧化还原酶、水、辅因子和开环二酮化合物。但是, 还可以包括另外的有机溶剂, 这些有机溶剂不参与辅因子再生, 即, 不包含任何可氧化的羟基。二乙醚、叔丁基甲醚、二异丙醚、二丁醚、醋酸乙酯、醋酸丁酯、庚烷、己烷、甲苯、二氯甲烷、环己烷或其混合物优选被用作另外的有机溶剂。

[0071] 因此, 两相系统中不可与水共混的有机组分的比例相对于反应批料总体积可以为 10% 到 90%, 优选 20% 到 80%。水相的比例相对于反应批料总体积可以为 90% 到 10%, 优选 80% 到 20%。

[0072] 还可以向水中添加缓冲剂, 诸如, 磷酸钾、Tris/ 盐酸、甘氨酸或三乙醇胺缓冲剂, 其 pH 值为 5 到 10, 优选的 6 到 9。此外, 所述缓冲液可以包含稳定或活化上述两种酶的离子, 诸如, 镁离子或锌离子。

[0073] 此外, 在本发明方法中可以使用稳定所用酶的其他添加剂, 例如, 甘油, 山梨糖醇, 1, 4-DL- 二硫苏糖醇 (DTT) 或二甲基亚砷 (DMSO)。

[0074] 根据水相计算, 辅因子 NAD(P)H 的浓度为 0.001mM 到 10mM, 尤其是 0.01mM 到 1.0mM。根据所用酶的特性, 温度可以是 10°C 到 70°C, 优选 20°C 到 35°C。

[0075] 通常, 有待还原的开环二酮衍生物难溶于水。因此, 在反应期间底物可以以完全或不完全溶解状态存在。如果底物不完全溶于反应混合物, 则底物的一部分以固体形式存在, 从而可以形成第三固相。在转化期间反应混合物还可能暂时形成乳状液。

[0076] 在本发明的方法中, 按总体积计算, 反应批料中使用的通式 I 的开环二酮衍生物的量优选为 10g/l 到 500g/l, 更优选为 25g/l 到 300g/l, 尤其优选为 50g/l 到 200g/l。

[0077] 本发明优选实施方案的特征还在于 13- 乙基 -3- 甲氧基 -8, 14- 开环 - 甾 -1, 3, 5(10), 9(11)- 四烯 -14, 17- 二酮 (乙基开环二酮 - 式 III) 或 13- 甲基 -3- 甲氧基 -8,

14-开环-甾-1,3,5(10),9(11)-四烯-14,17-二酮(甲基开环二酮-式 II)被用作开环二酮衍生物。

[0078] 在例如由玻璃或金属制成的反应容器中实施本发明的方法。为了这个目的,将组分分别转移入反应容器,并在例如氮气或空气的气氛中搅拌。根据使用的开环二酮化合物及氧化还原酶,反应时间为1小时到7天,尤其是2小时到48小时。在此期间,至少50%的开环二酮化合物被还原为对应的羟基开环甾族化合物。

[0079] 下面,将通过实施例更详细的阐述本发明。

[0080] 实施例 1

[0081] 从橙色绿屈挠菌(Chloroflexus auratiacus)DSM 635 克隆氧化还原酶

[0082] A) 橙色绿屈挠菌 DSM 635 的培养

[0083] 在 48°C 光照环境的细菌培养箱中,橙色绿屈挠菌 DSM 635 的细胞在下列培养基(pH 8.2)中培养:0.1%酵母抽提物,0.1%甘氨酸甘氨酸,0.01% $\text{Na}_2\text{HPO}_4 \times 2\text{H}_2\text{O}$,0.01% $\text{MgSO}_4 \times 7\text{H}_2\text{O}$,0.01% KNO_3 ,0.05% NaNO_3 ,0.01% NaCl ,0.005% $\text{CaCl}_2 \times 2\text{H}_2\text{O}$,5ml 0.01% Fe(III) 柠檬酸盐溶液,1ml 痕量元素溶液 SL-6[500 μl /l H_2SO_4 ,2.28g/l $\text{MnSO}_4 \times \text{H}_2\text{O}$,500mg/l $\text{ZnSO}_4 \times 7\text{H}_2\text{O}$,500mg H_3BO_3 ,25mg/l $\text{CuSO}_4 \times 5\text{H}_2\text{O}$,25mg/l $\text{Na}_2\text{MoO}_4 \times 2\text{H}_2\text{O}$,45mg/l $\text{CoCl}_2 \times 6\text{H}_2\text{O}$]。在培养的第 12 天,通过离心从培养基分离细胞,并在 -80°C 保存。

[0084] B) 扩增编码选择性氧化还原酶的基因

[0085] 按照 Manniatis & Sambrook 的“Molecular Cloning(分子克隆)”中描述的方法提取基因组 DNA。获得的核酸用作含特异性引物的聚合酶链式反应(PCR)的模板,所述特异性引物源自 NCBI 数据库编号 76258197 公开的基因序列。在这种情况下,引物在其 5' 末端分别配置有核酸内切酶 Nde I 和 Hind III 或 Sph I 的限制性酶切位点(SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13),以便随后克隆入表达载体。

[0086] 在含有 500ng 基因组 DNA 的 PCR 缓冲液[10mM Tris-HCl, (pH8.0);50mM KCl;10mM MgSO_4 ;1mM dNTP Mix;在每一情形下 20pMol 引物和 2.5U Platinum Pfx DNA 聚合酶(Invitrogen)]中和如下温度循环进行扩增:

[0087] 循环 1:94°C,2min

[0088] 循环 2x 30:94°C,30sec

[0089] 56°C,30sec

[0090] 68°C,60sec

[0091] 循环 3:68°C,7min

[0092] 4°C,∞

[0093] 获得的大小约 750bp 的 PCR 产物在 1% 琼脂糖凝胶上纯化后分别用核酸内切酶 NdeI 和 Hind III 或核酸内切酶 Sph I 和 Hind III 限制性酶切,将其分别连接入 pET21a 载体(Novagen)或 pQE70 载体(Qiagen)的骨架(该骨架已经用相同的核酸内切酶处理)。将 2 μl 连接批料转化入大肠杆菌 Top 10F' 细胞(Invitrogen)后,分别通过核酸内切酶 Nde I 和 Hind III 或核酸内切酶 Sph I 和 Hind III 的限制性分析检测氨苄青霉素(或卡那霉素)抗性菌落的质粒 DNAs 中大小为 750bp 的插入物的存在。来自所述片段阳性的克隆的质粒制品被用于序列分析,随后将其分别转化入大肠杆菌 BL21Star(Invitrogen)和大肠杆菌 RB791(genetic stock, Yale)。

[0094] 实施例 2

[0095] 重组绿屈挠菌氧化还原酶在大肠杆菌中的表达

[0096] 用表达构建体转化的大肠杆菌菌株 BL21 Star (Invitrogen, Karlsruhe, 德国) 和 RB791 (E. coli genetic stock, Yale, USA) 分别在含有氨苄青霉素 (50 μ g/ml) 或羧苄青霉素 (50 μ g/ml) 的 200ml LB- 培养基 (1% 胰蛋白胨, 0.5% 酵母抽提物, 1% NaCl) 中培养, 直至 550nm 测量的光密度 (OD) 达到 0.5。通过添加浓度为 0.1mM 的异丙基硫代半乳糖苷 (IPTG) 诱导重组蛋白的表达。在 25℃ 和 220rpm 诱导 8 小时或 16 小时后, 收获细胞并冷冻于 -20℃。为了进行活性试验, 将 10mg 细胞与 500 μ l 100mM TEA 缓冲液 pH 7.0 和 500 μ l 玻璃珠混合, 利用球磨机消化 10 分钟。然后所获裂解物在稀释状态下进行相应的测量。活性试验组成如下: 870 μ l 100mM TEA 缓冲液 pH 7.0, 160 μ g NADH, 10 μ l 稀释的细胞裂解物。通过向反应混合物中添加 100 μ l 100mM 底物溶液开始反应。

[0097] 为了大量回收酶, 将 30g 细胞重悬于 150ml 三乙醇胺缓冲液 (100mM, pH 7, 2mM $MgCl_2$, 10% 甘油), 并使用高压匀浆器消化。随后, 将酶溶液与 150ml 甘油混合, -20℃ 保存。

[0098] 实施例 3

[0099] 生物的培养及乙基开环二酮 (式 III) 还原转化后的筛选

[0100] 为了进行筛选, 酵母菌株粉状毕赤氏酵母 (*Pichia farinosa*) DSM70362, *Candida gropengiesseri* MUCL 29836, *Candida vaccinii* CBS7318, 粉状毕赤氏酵母 (*Pichia farinosa*) DSM 3316, 酿酒酵母 (*Saccharomyces cerevisiae*) CBS 1508 和木兰假丝酵母 (*Candida magnoliae*) CBS 6396 在下列培养基中培养: 酵母抽提物 (5), 蛋白胨 (5) 和葡萄糖 (20) (括号中数字的单位都是 g/l)。培养基在 121℃ 灭菌, 酵母在 25℃ 140 转每分的摇床上培养, 无需进一步的 pH 调节。

[0101] 在下列全细胞生物转化批料中检测式 III 的乙基开环二酮向对应羟基开环甾族化合物的还原转化: 将一批次中 400mg 新收获的细胞与 50mg 葡萄糖、10mg 式 III 的乙基开环二酮和 900 μ l 100mM 三乙醇胺缓冲液 (TEA) pH 7.0 在 28℃ 和 1400rpm 振荡 24 小时。随后, 所述批料用 1ml 二氯甲烷抽提, 离心, 用氮气干燥, 在容纳于乙腈后进行 HPLC 分析。

[0102] 筛选结果概述于表 1。

[0103] 表 1

[0104]

菌株序号	微生物	24 小时后利用野生型菌株 进行乙基开环二酮的转化
		24 h 批料
DSM 70362	粉状毕赤氏酵母	0.7%
MUCL 29836	<i>Candida gropengiesseri</i>	0.2%
CBS 7318	<i>Candida vaccinii</i>	3.2%

[0105]

DSM 3316	粉状毕赤氏酵母	15.8%
CBS 1508	酿酒酵母	0.7%
CBS 6396	木兰假丝酵母	41%

[0106] 菌株 CBS 6396 显示乙基开环二酮的最高转化,因此被挑选为制备 cDNA 文库的起始生物。

[0107] 实施例 4

[0108] 从木兰假丝酵母 CBS 6396 制备 cDNA 文库及克隆氧花还原酶

[0109] A) cDNA 文库的分离 (全部和 mRNA) 及制备

[0110] 将 600mg 新鲜细胞重悬于 2.5ml 冰冷却的 LETS 缓冲液。向所述细胞悬液中添加 5ml (约 20g) 玻璃珠,所述玻璃珠已用硝酸洗涤并用 3ml 苯酚平衡 (pH 7.0)。全体批料交替经 30 秒的涡旋和 30 秒的冰上冷却处理,共 10 分钟。随后,添加 5ml 冰冷却的 LETS 缓冲液,再将其激烈涡旋混合。在 4℃ 将所述细胞悬液用 11000g 离心 5 分钟。回收水相,并用等体积的苯酚:氯仿:异戊醇 (24:24:1) 抽提两次。然后用氯仿抽提。在最后一次抽提后,在 -20℃ 通过添加 1/10 体积的 5M LiCl₂ 沉淀总 RNA 4 小时。

[0111] 将获得的 1mg 总 RNA 通过 Oligo-dT 纤维素 (NEB Biolabs) 以富集 mRNA 分子。在随后的沉淀后,5 μg mRNA 用于 cDNA 合成 (pBluescript IIXR cDNA 文库构建试剂盒,Stratagene)。按照制造商的说明书将构建的文库转化入 XL-10Gold 大肠杆菌并筛选 ADH 的活性。利用 NADPH 或 NADH 分别作为辅因子和乙基开环二酮 (式 III) 作为底物,根据吸光度的降低鉴定并分离出克隆 (cM4)。利用引物 T7 和引物 T3 测序从所述克隆分离的质粒,获得 789bp 的 ORF。所述片段编码大小为 262 个氨基酸的融合蛋白,由 β-半乳糖苷酶的 α-片段和推定的短链醇脱氢酶序列组成。

[0112] B) 通过 PCR 合成编码来自木兰假丝酵母 CBS 6396 的短链 ADH 的全长转录物

[0113] 构建特异性引物用于后续将全长转录物克隆入合适的表达系统。在这种情况下,分别具有 Nde I 和 Sph I 识别序列的 5' 引物和分别具有 XhoI 和 SacI 识别序列的 3' 引物被修饰 (SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17)。从木兰假丝酵母表达文库的克隆 (cM4) 分离的质粒 DNA 用作聚合酶链式反应的模板。在含有 50ng 模板的 PCR 缓冲液 [10mM Tris-HCl, (pH 8.0); 50mM KCl; 10mM MgSO₄; 1mM dNTP Mix; 在每一情形下 20pMol 引物和 2.5U Platinum Pfx DNA 聚合酶 (Invitrogen)] 中和下面温度循环条件下进行扩增:

[0114] 循环 1: 94℃, 2min

[0115] 循环 2x 30: 94℃, 15sec

[0116] 58℃, 30sec

[0117] 68℃, 75sec

[0118] 循环 3: 68℃, 7min

[0119] 4℃, ∞

[0120] 获得的 PCR 产物在 1% 琼脂糖凝胶上纯化后分别用核酸内切酶 Nde I 和 Xho I 或核酸内切酶 Sph I 和 Sac I 限制性酶切,将其分别连接入 pET21a 载体 (Novagen) 或 pQME70 载体的骨架 (该骨架已经用相同的核酸内切酶处理)。将 2 μl 连接批料转化入大肠杆菌

Top 10F' 细胞 (Invitrogen) 后, 分别通过核酸内切酶 Nde I 和 Xho I 或核酸内切酶 Sph I 和 Sac I 的限制性分析检测氨苄青霉素 (或卡那霉素) 抗性菌落的质粒 DNAs 中大小为 750bp 的插入物的存在。对表达构建体 pET21-MgIV 和 pQME70-MgIV 进行测序。编码短链氧化还原酶的木兰假丝酵母基因具有共 729bp (包含于 SEQ ID NO:8) 的开放阅读框, 其对应于 243 个氨基酸的蛋白 (SEQ ID NO:3)。

[0121] 实施例 5

[0122] 重组氧化还原酶在大肠杆菌细胞中的表达

[0123] 用分别编码氧化还原酶的表达构建体 pET21-MgIV 和 pQME70-MgIV 分别转化感受态的大肠杆菌 StarBL21 (De3) 细胞 (Invitrogen) 和 RB791 细胞 (大肠杆菌 genetic stock, Yale, USA)。然后将用表达构建体转化的大肠杆菌菌落分别在 200ml 含有 50 μ g/ml 氨苄青霉素或 40 μ g/ml 卡那霉素的 LB 培养基 (1% 胰蛋白胨, 0.5% 酵母抽提物, 1% NaCl) 中培养, 直至在 550nm 时测量的光密度达到 0.5。通过添加 0.1mM 浓度的异丙基硫代半乳糖苷 (IPTG) 诱导重组蛋白的表达。在 25°C 和 220rpm 诱导 16 小时后, 收获细胞并冷冻于 -20°C。为了进行活性试验, 将 10mg 细胞与 500 μ l 100mM TEA 缓冲液 pH 7.0、1mM $MgCl_2$ 和 500 μ l 玻璃珠混合, 利用球磨机消化 10 分钟。然后所获裂解物在稀释状态下进行相应的测量。

[0124] 活性试验组成如下: 960 μ l 100mM TEA 缓冲液 pH 7.0, 1mM $MgCl_2$, 160 μ g NADH, 10 μ l 稀释的细胞裂解物。通过向反应混合物中添加 10 μ l 100mM 底物溶液 (溶于 70% 甲醇) 开始反应。

[0125] 为了大量回收酶, 将 30g 细胞重悬于 150ml 三乙醇胺缓冲液 (100mM, pH 7, 2mM $MgCl_2$, 10% 甘油), 并用高压匀浆器消化。随后, 将酶溶液与 150ml 甘油混合, 在 -20°C 保存。

[0126] 实施例 6

[0127] 通过氧化还原酶 SEQ ID NO:1 还原乙基开环二酮 (式 III)

[0128] 为了还原乙基开环二酮 (式 III), 室温下 800 μ l 缓冲液 (100mM 磷酸钾, pH = 7, 2mM $MgCl_2$), 1.2ml 2- 丙醇, 0.08mg NAD, 100mg 乙基开环二酮 (式 III) 和 1ml 酶悬液氧化还原酶 SEQ ID NO:1 (参见实施例 3) 的混合物在恒定充分混合条件下在反应容器中温育 24 小时。96 小时后, > 90% 的使用的乙基开环二酮 (式 III) 被还原。

[0129] 在反应完成时, 通过用二氯甲烷抽提对反应混合物进行再处理, 分离包含产物的有机相, 通过蒸发 / 馏出溶剂获得 17- β - 羟基化合物 (乙基开环醇 (Ethylsecol))。

[0130] 然后通过 HPLC 将乙基开环二酮转化为乙基开环醇 (Ethylsecol)。为了这个目的, 使用乙腈和水作为流动相的分离柱 EC 125/4Nucleodur100-5 C18ec (Machery-Nagel, Düren, 德国)。为了进行分析, 使用流动相中乙腈比例从 30% 到 70% 的线性梯度。通过与参照物比较进行反应产物的鉴定。

[0131] 实施例 7

[0132] 通过氧化还原酶 SEQ ID NO:2 还原乙基开环二酮 (式 III)

[0133] 为了还原乙基开环二酮 (式 III), 室温下 250 μ l 缓冲液 (100mM 三乙醇胺, pH = 8, 2mM $MgCl_2$), 250 μ l 4- 甲基 -2- 戊醇, 0.02mg NAD, 25mg 乙基开环二酮 (式 III) 和 25 μ l 酶悬液氧化还原酶 SEQ ID NO:2 (参见实施例 3) 的混合物在恒定充分混合条件下在反应容器中温育 96 小时。96 小时后, > 30% 的使用的乙基开环二酮 (式 III) 被还原为羟基化合物。

[0134] 在反应完成时,通过用二氯甲烷抽提对反应混合物进行再处理,分离包含产物的有机相,通过蒸发/馏出溶剂获得 17- β -羟基化合物(乙基开环醇(Ethylsecol))。

[0135] 实施例 8

[0136] 通过氧化还原酶 SEQ ID NO:3 还原乙基开环二酮(式 III)

[0137] 为了还原乙基开环二酮(式 III),室温下 100 μ l 缓冲液(100mM 三乙醇胺, pH = 7, 2mM $MgCl_2$), 400 μ l 4-甲基-2-戊醇, 0.02mg NADP, 25mg 乙基开环二酮(式 III) 和 100 μ l 酶悬液氧化还原酶 SEQ ID NO:3(参见实施例 3) 的混合物在恒定充分混合条件下在反应容器中温育 72 小时。72 小时后, > 95% 的所用乙基开环二酮(式 III) 被还原为羟基化合物。

[0138] 实施例 9

[0139] 通过氧化还原酶 SEQ ID NO:4 还原乙基开环二酮(式 III)

[0140] 为了还原乙基开环二酮(式 III),室温下 200 μ l 缓冲液(100mM 三乙醇胺, pH = 9, 2mM $MgCl_2$), 300 μ l 2-庚醇, 0.025mg NADP, 100mg 乙基开环二酮(式 III) 和 50 μ l 酶悬液氧化还原酶 SEQ ID NO:4(参见实施例 3) 的混合物在恒定充分混合条件下在反应容器中温育 72 小时。72 小时后, > 80% 的所用乙基开环二酮(式 III) 被还原为羟基化合物。

[0141] 实施例 10

[0142] 通过氧化还原酶 SEQ ID NO:5 还原乙基开环二酮(式 III)

[0143] 为了还原乙基开环二酮(式 III),室温下 300 μ l 缓冲液(100mM 三乙醇胺, pH = 7, 2mM $MgCl_2$), 1.2ml 4-甲基-2-戊醇, 0.12mg NADP, 150mg 乙基开环二酮(式 III) 和 0.6ml 酶悬液氧化还原酶 SEQ ID NO:5(参见实施例 3) 的混合物在恒定充分混合条件下在反应容器中温育 72 小时。72 小时后, > 90% 的所用乙基开环二酮(式 III) 被还原为羟基化合物。

[0001]

序列表

<110> IEP GmbH

<120> 在 NADH 或 NADPH 存在的条件下使用氧化还原酶/脱氢酶通过对映体
选择性酶促还原开环二酮衍生物来制备开环醇衍生物的方法

<130> I 11005

<160> 42

<170> PatentIn version 3.3

<210> 1

<211> 252

<212> PRT

<213> 橙色绿屈挠菌 DSM 635

<400> 1

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

Ala Gly Ile Gly Arg Ala Ser Ala Leu Ala Phe Ala Arg Glu Gly Ala
20 25 30

Lys Val Val Val Ala Asp Val Asn Val Glu Gly Gly Glu Glu Thr Ile
35 40 45

Ala Leu Cys Arg Ala Leu Asn Thr Asp Ala Met Phe Val Arg Cys Asp
50 55 60

Val Ser Gln Arg Asp Glu Val Glu Arg Leu Ile Ala Leu Ala Val Asp
65 70 75 80

Thr Phe Gly Arg Ile Asp Phe Ala His Asn Asn Ala Gly Ile Glu Gly
85 90 95

[0002]

Val Gln Ala Met Leu Ala Asp Tyr Pro Glu Glu Val Trp Asp Arg Val
100 105 110

Ile Glu Ile Asn Leu Lys Gly Val Trp Leu Cys Met Lys Tyr Glu Ile
115 120 125

Arg His Met Leu Lys Gln Gly Gly Gly Ala Ile Val Asn Thr Ser Ser
130 135 140

Val Ala Gly Leu Ala Gly Ser Arg Gly Val Ser Ala Tyr Val Ala Ser
145 150 155 160

Lys His Gly Ile Val Gly Ile Thr Lys Ala Ala Ala Leu Glu Tyr Ala
165 170 175

Arg Asn Gly Ile Arg Val Asn Ala Ile Cys Pro Gly Thr Ile His Thr
180 185 190

Ala Met Ile Asp Arg Phe Thr Gln Gly Asp Pro Gln Leu Leu Ala Gln
195 200 205

Phe Ala Glu Gly Glu Pro Ile Gly Arg Leu Gly Ser Pro Glu Glu Val
210 215 220

Ala Asn Ala Val Ile Trp Leu Cys Ser Asp Lys Ala Ser Phe Val Thr
225 230 235 240

Gly Ala Thr Leu Ala Val Asp Gly Gly Arg Leu Ala
245 250

<210> 2

<211> 249

[0003]

<212> PRT

<213> Rubrobacter xylanophilus DSM 9941

<400> 2

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

Gly Arg Ala Thr Ala Leu Lys Phe Ala Arg Glu Gly Ala Arg Val Val
20 25 30

Ala Ala Glu Leu Asp Glu Arg Gly Gly Glu Gly Val Val Arg Glu Val
35 40 45

Arg Ser Leu Gly Gly Glu Ala Val Phe Val Arg Thr Asp Val Ser Glu
50 55 60

Phe Ala Gln Val Glu Asp Ala Val Glu Arg Ala Val Gly Glu Tyr Gly
65 70 75 80

Thr Leu Asp Val Met Phe Asn Asn Ala Gly Ile Gly His Tyr Ala Pro
85 90 95

Leu Leu Glu His Glu Pro Glu His Tyr Asp Arg Val Val Arg Val Asn
100 105 110

Gln Tyr Gly Val Tyr Tyr Gly Ile Leu Ala Ala Gly Arg Lys Met Val
115 120 125

Ala Leu Lys Asn Pro Gly Leu Ile Ile Asn Thr Ala Ser Val Tyr Ala
130 135 140

Phe Leu Ala Ser Pro Gly Val Ile Gly Tyr His Ala Ala Lys Gly Ala
145 150 155 160

[0004]

Val Lys Met Met Thr Gln Ala Ala Ala Leu Glu Leu Ala Pro His Gly
165 170 175

Ile Arg Val Val Ala Ile Ala Pro Gly Gly Val Asp Thr Pro Ile Ile
180 185 190

Gln Gly Tyr Lys Asp Met Gly Leu Gly Glu Arg Leu Ala Arg Gly Gln
195 200 205

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

Leu Leu Ala Thr Asp Glu Ala Asp Ala Ile Asn Gly Ser Val Val Met
225 230 235 240

Thr Asp Asp Gly Tyr Ala Glu Phe Lys
245

<210> 3

<211> 243

<212> PRT

<213> 木兰假丝酵母 CBS 6396

<400> 3

Met Ser Ala Thr Ser Asn Ala Leu Ile Thr Gly Ala Ser Arg Gly Met
1 5 10 15

Gly Glu Ala Thr Ala Ile Lys Leu Ala Leu Glu Gly Tyr Ser Val Thr
20 25 30

Leu Ala Ser Arg Gly Ile Glu Gln Leu Asn Ala Ile Lys Glu Lys Leu
35 40 45

[0005]

Pro Ile Val Lys Lys Gly Gln Gln His Tyr Val Trp Gln Leu Asp Leu
50 55 60

Ser Asp Ile Glu Ala Ala Ser Thr Phe Lys Gly Ala Pro Leu Pro Ala
65 70 75 80

Ser Ser Tyr Asp Val Phe Phe Ser Asn Ala Gly Val Val Asp Phe Ala
85 90 95

Pro Phe Ala Asp Gln Ser Glu Thr Ala Gln Lys Asp Leu Phe Thr Val
100 105 110

Asn Leu Leu Ser Pro Val Ala Leu Thr Lys Thr Ile Val Lys Ala Ile
115 120 125

Ala Asp Lys Pro Arg Glu Thr Pro Ala His Ile Ile Phe Thr Ser Ser
130 135 140

Ile Val Gly Ile Arg Gly Val Pro Asn Val Ala Val Tyr Ser Ala Thr
145 150 155 160

Lys Gly Ala Ile Asp Ser Phe Ala Arg Ser Leu Ala Arg Glu Phe Gly
165 170 175

Pro Lys Asn Ile His Val Asn Cys Val Asn Pro Gly Thr Thr Arg Thr
180 185 190

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

Lys Gly Trp Ile Glu Val Asp Ala Ile Ala Asp Ala Val Leu Phe Leu
210 215 220

[0006]

Ile Lys Ser Lys Asn Ile Thr Gly Gln Ser Leu Val Val Asp Asn Gly
225 230 235 240

Phe Gly Val

<210> 4

<211> 241

<212> PRT

<213> 木兰假丝酵母 DSM 70638

<400> 4

Met Thr Ser Thr Pro Asn Ala Leu Ile Thr Gly Gly Ser Arg Gly Ile
1 5 10 15

Gly Ala Ser Ala Ala Ile Lys Leu Ala Gln Glu Gly Tyr Ser Val Thr
20 25 30

Leu Ala Ser Arg Asp Leu Glu Lys Leu Thr Glu Val Lys Asp Lys Leu
35 40 45

Pro Ile Val Arg Gly Gly Gln Lys His Tyr Val Trp Gln Leu Asp Leu
50 55 60

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

Ser Ser Tyr Asp Leu Phe Val Ser Asn Ala Gly Ile Ala Gln Phe Ser
85 90 95

Pro Thr Ala Glu His Thr Asn Ser Glu Trp Leu Asn Ile Met Thr Ile
100 105 110

Asn Leu Val Ser Pro Ile Ala Leu Thr Lys Ala Leu Leu Gln Ala Val

[0007]

115	120	125	
Ser Gly Arg Ser Ser Glu Asn Pro Phe Gln Ile Val Phe Ile Ser Ser			
130	135	140	
Val Ala Ala Leu Arg Gly Val Ala Gln Thr Ala Val Tyr Ser Ala Ser			
145	150	155	160
Lys Ala Gly Thr Asp Gly Phe Ala Arg Ser Leu Ala Arg Glu Leu Gly			
	165	170	175
Pro Gln Gly Val His Val Asn Val Val Asn Pro Gly Trp Thr Lys Thr			
	180	185	190
Asp Met Thr Glu Gly Val Glu Thr Pro Lys Asp Met Pro Ile Lys Gly			
	195	200	205
Trp Ile Gln Pro Glu Ala Ile Ala Asp Ala Val Val Phe Leu Ala Arg			
210	215	220	
Ser Lys Asn Ile Thr Gly Ala Asn Ile Val Val Asp Asn Gly Phe Ser			
225	230	235	240
Thr			
<210>	5		
<211>	241		
<212>	PRT		
<213>	木兰假丝酵母 DSM 70638		
<400>	5		
Met Thr Thr Thr Ser Asn Ala Leu Val Thr Gly Gly Ser Arg Gly Ile			
1	5	10	15

[0008]

Gly Ala Ala Ser Ala Ile Lys Leu Ala Gln Glu Gly Tyr Asn Val Thr
20 25 30

Leu Ala Ser Arg Ser Val Asp Lys Leu Asn Glu Val Lys Ala Lys Leu
35 40 45

Pro Ile Val Gln Asp Gly Gln Lys His Tyr Ile Trp Glu Leu Asp Leu
50 55 60

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

Arg Ser Tyr Asp Val Phe Val Ser Asn Ala Gly Val Ala Ala Phe Ser
85 90 95

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

Asn Leu Ser Ser Pro Ile Ala Leu Thr Lys Ala Leu Leu Lys Asp Val
115 120 125

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

Val Ala Gly Leu His Gly Ala Ala Gln Val Ala Val Tyr Ser Ala Ser
145 150 155 160

Lys Ala Gly Leu Asp Gly Phe Met Arg Ser Val Ala Arg Glu Val Gly
165 170 175

Pro Lys Gly Ile His Val Asn Ser Ile Asn Pro Gly Tyr Thr Lys Thr
180 185 190

[0009]

Glu Met Thr Ala Gly Ile Glu Ala Leu Pro Asp Leu Pro Ile Lys Gly
195 200 205

Trp Ile Glu Pro Glu Ala Ile Ala Asp Ala Val Leu Phe Leu Ala Lys
210 215 220

Ser Lys Asn Ile Thr Gly Thr Asn Ile Val Val Asp Asn Gly Leu Ile
225 230 235 240

Ala

<210> 6

<211> 759

<212> DNA

<213> 橙色绿屈挠菌 DSM 635

<400> 6

atggagccac ctttcattgg gaaggttgcg ctggtcaccg gcgcagcagc cggtattggt	60
cgtgcttcag cactggcggt tgcccgtgag ggtgccaaagg ttgtcgttgc tgatgtgaat	120
gtcgagggcg gggaagagac gattgcgctg tgtcgggctt tgaataccga tgcaatgttc	180
gtgcgttgtg atgtttcgca acgcgatgaa gtggagcgat taattgctct ggcagttgac	240
acgttcggtc ggatcgactt tgcgcacaac aacgccggga ttgaaggcgt gcaggcaatg	300
ctggccgatt atcccgaaga ggtctgggat cgggtgatcg agatcaacct caaaggggtc	360
tggttgtgta tgaagtacga aatccggcac atgctcaagc aggggtggcgg tgcgattgtg	420
aatacctcat cggtcgccgg tctggccgga tcacgtggcg tttcggcgta tgtagccagc	480
aagcacggta ttgttggtat taccaaagcg gcagcccttg agtatgcgcg taacggtatt	540
cgtgtcaacg caatctgtcc aggtacgatt catactgcga tgatcgaccg ctttaccag	600

[0010]

ggtgatcccc aactgcttgc ccagttcgct gagggatgaac cgattggctg gctcggctcg 660
 cctgaagagg tcgccaatgc ggtgatctgg ctctgctcag ataaggcttc gtttgtgacc 720
 ggagcgacac tggcggttga tggtagccgc ctggcgtaa 759

<210> 7

<211> 750

<212> DNA

<213> Rubrobacter xylanophilus DSM 9941

<400> 7

atgctcgagg ggaaggtcgc ggtcatcacg ggggcccga ggcgcatagg ccgggcccacc 60
 gcgctcaagt tcgcccgcga gggggcccgg gtcgtcgccg ccgagctcga cgagcgcggc 120
 ggggaggggg tggccggga ggtgcgcagc ctggggggcg aggcgggtctt cgtccggacc 180
 gacgtctcgg agttcgcgca ggtggaggac gccgtcgagc gggcggctcg ggagtacggc 240
 accctcgacg tgatgttcaa caacgccgc atcgggcact acgccccct gctggagcac 300
 gagcccagc actacgaccg ggtggtccgg gtgaaccagt acggcgtcta ctacgggata 360
 ctgcccgcg ggagaaagat ggtcgccctg aagaaccccg gcttgatcat caacaccgcc 420
 tcggtctacg ccttcctcgc ctgcccgggg gtcateggct accacgccgc caagggggcg 480
 gtcaagatga tgaccaggc ggcggcgctg gagctcgccc cgcacggcat aagggtcgtc 540
 gccatcgccc cgggcggggg ggacaccccc atcatecagg gctacaagga catggggctc 600
 ggcgagaggc tggcccgcgg ccagatgcgc cgccggctcc agacccccga gcagatcgcc 660
 ggggcggctg cctgtctcgc caccgacgag gccgaccca taaacggctc ggtggtcatg 720
 accgacgacg gctacgcgga gttcaagtag 750

<210> 8

<211> 732

<212> DNA

<213> 木兰假丝酵母 CBS 6396

[0011]

<400> 8	
atgtctgcta cttcgaacgc tcttatcact ggtgccagcc gcggaatggg cgaggccaca	60
gctattaagc ttgcccttga ggggtacagc gtcacccttg catcacgcgg tattgagcag	120
ctcaatgcca tcaaggaaaa actaccatc gtgaagaagg gccagcagca ctacgtttgg	180
cagctcgatc ttagtgacat cgaggcggct tccaccttca agggggctcc tctgcctgcc	240
agcagctacg acgtgttctt cagcaacgcc ggtgtggtgg actttgctcc gttcgcagac	300
caaagcgaga ctgcgcaaaa ggacctgttc acggttaacc tgctgtcgcc tgttgcggtg	360
accaagacca ttgtaaggc catcgccgac aagccccgcg agacgcctgc tcacattatc	420
ttcacctcgt ccattgtcgg aattcgcggt gtccccaacg tggcggctta cagcgccacc	480
aagggcgcga ttgacagctt tgcgcgctcg ctgtctcgtg agttcgggtcc caagaacatc	540
cacgttaact gcgtgaaccc gggcacgacg cgcaccgaga tgacaaagg cgttgatctc	600
gcggctttcg gcgatgttcc tatcaagggc tggatcgagg tcgatgcgat tgccgacgct	660
gtgctgtttt tgatcaagtc caagaacatc actggccagt cgtcgtttgt tgacaacgga	720
ttcgggtgtt aa	732

<210> 9

<211> 726

<212> DNA

<213> 木兰假丝酵母 DSM 70638

<400> 9

atgacatcta cacctaagc cctcatcacg ggaggcagcc gcggcattgg cgcttcgcc	60
gccatcaaac tggtcaaga aggggtacagc gtcacgctgg cgtcccgca ccttgagaaa	120
cttactgagg tcaaggacaa gctgccaatc gtgagaggtg gacagaaaca ctacgtttgg	180
cagctcgatc ttgccgatgt ggaggctgca tcgtctttca aggcggctcc tctgccggcc	240
agcagctacg atttgtttgt ttcgaacgcc ggaattgccc agttctcgcc tacggcagag	300

[0012]

catactaata gtgagtggct gaacattatg accattaact tagtgtcccc gattgccctg	360
acgaaggctc ttttgcaggc cgtttctggg aggtcgagcg agaaccggtt tcagatcgtc	420
ttcatctcgt cggttgcagc actacgtggc gttgcacaaa cggccgtcta cagtgcgtcg	480
aaggctggta ctgatggatt cgcacgctca cttgctcgcg aactaggtcc tcaaggtgtt	540
catgtgaacg tggatgaacc tggctggact aagacagaca tgacggaagg agtcgaaacc	600
ccaaaggaca tgcccattaa gggctggatc cagcctgagg caattgctga tgctgtagta	660
ttccttgcga ggtcgaaaaa cattaccggc gcgaatattg tagtggacaa tggtttctcg	720
acgtaa	726

<210> 10

<211> 726

<212> DNA

<213> 木兰假丝酵母 DSM 70638

<400> 10

atgacgacta cttcaaacgc gcttgtcact ggaggcagcc gcggcattgg cgctgcctcc	60
gccattaagc tggctcagga gggctacaat gttacgttgg cctctcgag tgttgataaa	120
ctgaatgaag taaaggcgaa actcccaatt gtacaggacg ggcagaagca ctacatttgg	180
gaactcgatc tggctgatgt ggaagctgct tcgtcgttca aggggtgctcc tttgcctgct	240
cgcagctacg acgtctttgt ttcgaacgcg ggctcgtg cgttctcgcc cacagccgac	300
cacgatgata aggagtggca gaacttgctt gccgtgaact tgctcgtgcc cattgccctc	360
acgaaggccc tcttgaagga tgtctccgaa aggcctgtgg acaagccact gcagattatc	420
tacatttcgt cgggtggccgg cttgcatggc gccgcgcagg tcgccgtgta cagtgcattc	480
aaggccggtc ttgatggttt tatgcgtcc gtcgccgtg aggtgggccc gaagggcac	540
catgtgaact ccatcaacc cggatacacg aagactgaaa tgaccgcggg cattgaagcc	600

[0013]

cttcctgatt tgcctatcaa ggggtggatc gagcccgagg caattgctga cgcggttctg 660
tttctggcaa agtccaagaa taccaccggc acaaacattg tggtcgacaa tggcttgatt 720
gcttaa 726

<210> 11
<211> 38
<212> DNA
<213> 人工构建物

<220>
<221> misc_feature
<222> (1)..(38)

<400> 11
ggaattccat atgatggagc cacctttcat tgggaagg 38

<210> 12
<211> 34
<212> DNA
<213> 人工构建物

<220>
<221> misc_feature
<222> (1)..(34)

<400> 12
cccaagctta ttattacgcc aggcggccac catc 34

<210> 13
<211> 34
<212> DNA
<213> 人工构建物

<220>
<221> misc_feature
<222> (1)..(34)

[0014]

<400> 13
cccaagctta ttattacgcc aggcggccac catc 34

<210> 14
<211> 35
<212> DNA
<213> 人工构建物

<220>
<221> misc_feature
<222> (1)..(35)

<400> 14
ggaattccat atgatgtctg ctacttcgaa cgctc 35

<210> 15
<211> 33
<212> DNA
<213> 人工构建物

<220>
<221> misc_feature
<222> (1)..(33)

<400> 15
ccgctcgagt tattaaacac cgaatccggtt gtc 33

<210> 16
<211> 35
<212> DNA
<213> 人工构建物

<400> 16
cacatgcatg cagatgtctg ctacttcgaa cgctc 35

<210> 17
<211> 34

[0015]

<212> DNA

<213> 人工构建物

<220>

<221> misc_feature

<222> (1)..(34)

<400> 17

gcccgagctc ttattaaaca ccgaatccgt tgtc

34

<210> 18

<211> 12

<212> PRT

<213> 人工构建物

<400> 18

Asn Ala Leu Val Thr Gly Ala Ser Arg Gly Ile Gly
1 5 10

<210> 19

<211> 12

<212> PRT

<213> 人工构建物

<400> 19

Asn Ala Leu Val Thr Gly Gly Ser Arg Gly Ile Gly
1 5 10

<210> 20

<211> 12

<212> PRT

<213> 人工构建物

<400> 20

Asn Ala Leu Ile Thr Gly Gly Ser Arg Gly Ile Gly
1 5 10

[0016]

<210> 21
<211> 12
<212> PRT
<213> 人工构建物

<400> 21

Asn Ala Leu Ile Thr Gly Ala Ser Arg Gly Ile Gly
1 5 10

<210> 22
<211> 12
<212> PRT
<213> 人工构建物

<400> 22

Asn Ala Leu Ile Thr Gly Gly Ser Arg Gly Met Gly
1 5 10

<210> 23
<211> 12
<212> PRT
<213> 人工构建物

<400> 23

His Ala Leu Val Thr Gly Ala Ser Arg Gly Ile Gly
1 5 10

<210> 24
<211> 7
<212> PRT
<213> 人工构建物

<400> 24

Gly Tyr Ser Val Thr Leu Ala
1 5

[0017]

<210> 25
<211> 7
<212> PRT
<213> 人工构建物

<400> 25

Gly Tyr Asn Val Thr Leu Ala
1 5

<210> 26
<211> 7
<212> PRT
<213> 人工构建物

<400> 26

Gly Tyr Ser Val Thr Leu Val
1 5

<210> 27
<211> 7
<212> PRT
<213> 人工构建物

<400> 27

Gly Tyr Asn Val Thr Leu Val
1 5

<210> 28
<211> 8
<212> PRT
<213> 人工构建物

<400> 28

Phe Lys Gly Ala Pro Leu Pro Ala
1 5

[0018]

<210> 29
<211> 8
<212> PRT
<213> 人工构建物

<400> 29

Phe Lys Ala Ala Pro Leu Pro Ala
1 5

<210> 30
<211> 6
<212> PRT
<213> 人工构建物

<400> 30

Phe Val Ser Asn Ala Gly
1 5

<210> 31
<211> 6
<212> PRT
<213> 人工构建物

<400> 31

Phe Phe Ser Asn Ala Gly
1 5

<210> 32
<211> 6
<212> PRT
<213> 人工构建物

<400> 32

Phe Val Cys Asn Ala Gly
1 5

[0019]

<210> 33
<211> 6
<212> PRT
<213> 人工构建物

<400> 33

Phe Val Ala Asn Ala Gly
1 5

<210> 34
<211> 9
<212> PRT
<213> 人工构建物

<400> 34

Ser Pro Ile Ala Leu Thr Lys Ala Leu
1 5

<210> 35
<211> 9
<212> PRT
<213> 人工构建物

<400> 35

Ser Pro Val Ala Leu Thr Lys Thr Ile
1 5

<210> 36
<211> 9
<212> PRT
<213> 人工构建物

<400> 36

Ser Pro Ile Ala Leu Thr Lys Thr Leu
1 5

[0020]

<210> 37
<211> 9
<212> PRT
<213> 人工构建物

<400> 37

Ser Pro Val Ala Met Thr Lys Ala Leu
1 5

<210> 38
<211> 9
<212> PRT
<213> 人工构建物

<400> 38

Ser Gln Ile Ala Leu Thr Lys Ala Leu
1 5

<210> 39
<211> 7
<212> PRT
<213> 人工构建物

<400> 39

Ala Val Tyr Ser Ala Ser Lys
1 5

<210> 40
<211> 7
<212> PRT
<213> 人工构建物

<400> 40

Ala Val Tyr Ser Ala Thr Lys
1 5

[0021]

<210> 41
<211> 6
<212> PRT
<213> 人工构建物

<400> 41

Pro Ile Lys Gly Trp Ile
1 5

<210> 42
<211> 6
<212> PRT
<213> 人工构建物

<400> 42

Pro Ile Ser Gly Trp Ile
1 5