

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200680047754.6

[51] Int. Cl.

C12P 7/62 (2006.01)

C12P 41/00 (2006.01)

C12P 13/00 (2006.01)

[43] 公开日 2009 年 5 月 13 日

[11] 公开号 CN 101432435A

[22] 申请日 2006.12.14

[21] 申请号 200680047754.6

[30] 优先权

[32] 2005.12.19 [33] AT [31] A2027/2005

[86] 国际申请 PCT/EP2006/012037 2006.12.14

[87] 国际公布 WO2007/073875 德 2007.7.5

[85] 进入国家阶段日期 2008.6.19

[71] 申请人 IEP 有限责任公司

地址 德国威斯巴登

[72] 发明人 A·古普塔 M·博布科瓦

A·申茨彻

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 刘维升 李连涛

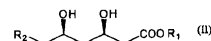
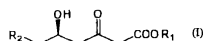
权利要求书 3 页 说明书 15 页 序列表 13 页

[54] 发明名称

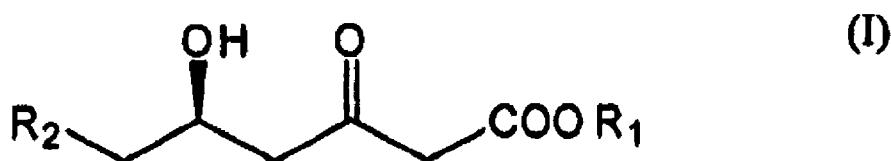
对映体选择性酶还原羟基酮化合物的方法

[57] 摘要

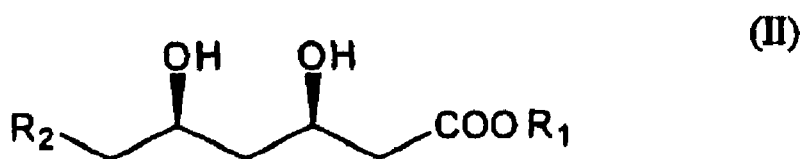
对映体选择性酶还原通式 I 之羟基酮的方法，所述通式 I 中， $R_1 = C_1 - C_6$ 烷基， $R_2 = -Cl, -CN, -OH, -H$ 或者 $C_1 - C_6$ 烷基，以得到通式 II 的手性二醇，其中 R_1 与 R_2 与通式 I 中的含义相同。在 NADH 或者 NADPH 作为辅因子的条件下，用氧化还原酶还原羟基酮，其中 a) 所述羟基酮在反应混合物中的浓度为 $> 50g/l$ ，b) 通过氧化通式 $R_X R_Y CHOH$ 的二级醇持续再生所得的氧化辅因子 NAD 或者 NADP，其中 R_X, R_Y 彼此独立是氢、支链或者无支链 $C_1 - C_8$ 烷基，且 $C_{总数} \geq 3$ ，并且 c) 羟基酮的还原和二级醇的氧化是由相同的氧化还原酶催化的。



1、对映体选择性酶促还原通式 I 羟基酮的方法



所述通式 I 中, $\text{R}_1=\text{C}_1\text{-C}_6$ 烷基, $\text{R}_2=-\text{Cl}$, $-\text{CN}$, $-\text{OH}$, $-\text{H}$ 或者 $\text{C}_1\text{-C}_6$ 烷基, 以得到通式 II 的手性二醇:



其中 R_1 与 R_2 与通式 I 中的含义相同,

其中在 NADH 或者 NADPH 作为辅因子的条件下, 用氧化还原酶还原羟基酮, 其特征在于

a) 羟基酮在反应混合物中的浓度 $\geq 50\text{g/l}$,

b) 通过通式 $\text{R}_\text{X}\text{R}_\text{Y}\text{CHOH}$ 的二级醇氧化持续再生所得的氧化辅因子 NAD 或者 NADP, 其中 R_X , R_Y 彼此独立是氢、支链或者无支链 $\text{C}_1\text{-C}_8$ 烷基, 且 $\text{C}_{\text{总数}} \geq 3$, 和

c) 羟基酮的还原和二级醇的氧化是由相同的氧化还原酶催化。

2、根据权利要求 1 的方法, 其特征在于, 所述氧化还原酶

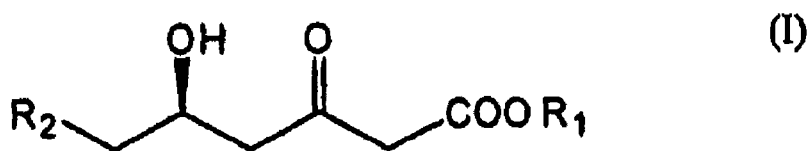
a) 含有氨基酸序列, 其中至少 50% 的氨基酸与氨基酸序列 SEQ ID NO:1、SEQ ID NO:8、SEQ ID NO:11 或者 SEQ ID NO:14 的那些相同,

b) 由 SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:13 或者 SEQ ID NO:15 的核酸序列编码, 或者

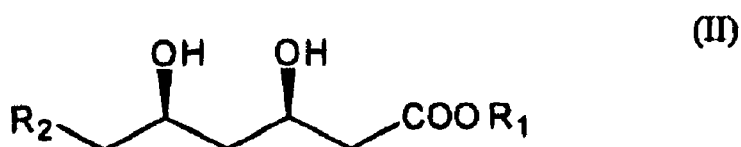
c) 由在严格条件下, 与 SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:13 或者 SEQ ID NO:15 杂交的核酸序列

编码。

3、对映体选择性酶促还原通式 I 羟基酮的方法



所述通式 I 中, $\text{R}_1=\text{C}_1\text{-C}_6$ 烷基, $\text{R}_2=\text{-Cl, -CN, -OH, -H}$ 或者 $\text{C}_1\text{-C}_6$ 烷基, 以得到通式 II 的手性二醇:



其中 R_1 与 R_2 与通式 I 中的含义相同,

其中在辅因子存在的条件下, 用氧化还原酶还原羟基酮, 其特征在于所述氧化还原酶

a) 含有氨基酸序列, 其中至少 50% 的氨基酸与氨基酸序列 SEQ ID NO:1、SEQ ID NO:8、SEQ ID NO:11 或者 SEQ ID NO:14 的那些相同,

b) 由 SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:13 或者 SEQ ID NO:15 的核酸序列编码, 或者

c) 由在严格条件下, 与 SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:13 或者 SEQ ID NO:15 杂交的核酸序列编码。

4、根据权利要求 3 的方法, 其特征在于, 辅因子被共底物持续还原。

5、权利要求 3 或者 4 的方法, 其特征在于用 NAD(P)H 作为辅因子。

6、权利要求 1-5 的任一项的方法, 其特征在于用 2-丙醇、2-丁醇、2-戊醇、4-甲基-2-戊醇、2-庚醇或者 2-辛醇作为共底物或者二级醇。

7、根据权利要求 1-6 的任一项的方法, 其特征在于所述羟基酮占总反应体积的 5-50 重量%, 优选 8-40 重量%, 特别优选 10-25 重量%。

8、根据权利要求 1-7 的任一项的方法, 其特征在于用叔-丁基(5R)-6-氰基-5-

羟基-3-氧代己酸酯、叔-丁基(5S)-6-氯-5-羟基-3-氧代己酸酯或者叔-丁基(5S)-5,6-二羟基-3-氧代己酸酯作为通式I的羟基酮。

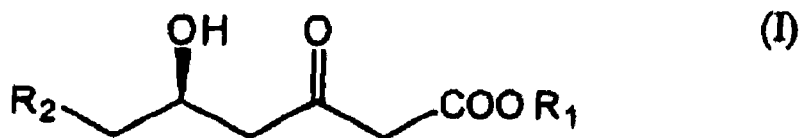
9、根据权利要求 1-8 的任一项的方法，其特征在于 TTN(总转化数= mol 还原的羟基酮/ mol 所用的辅因子) $\geq 10^3$ 。

10、根据权利要求 1-9 的任一项的方法，其特征在于所述方法在含水有机两相系统中进行。

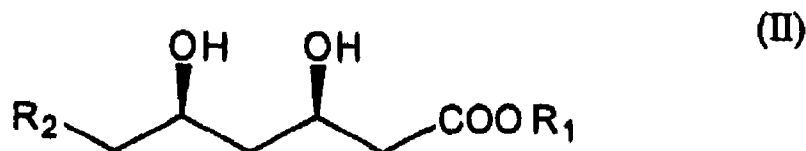
11、根据权利要求 1-10 的任一项的方法，其特征在于别外地使用有机溶剂，如乙醚、叔-丁基甲基醚、二异丙基醚、二丁基醚、乙酸丁酯、庚烷、己烷或者环己醇。

对映体选择性酶还原羟基酮化合物的方法

本发明涉及对映体选择性酶还原通式 I 之羟基酮的方法:



所述通式 I 中, $\text{R}_1 = \text{C}_1\text{--C}_6$ 烷基, $\text{R}_2 = \text{--Cl}$, --CN , --OH , --H 或者 $\text{C}_1\text{--C}_6$ 烷基, 以得到通式 II 的手性二醇:



其中 R_1 与 R_2 与通式 I 中的含义相同, 在辅因子的条件下, 用氧化还原酶还原羟基酮。

通式 II 的手性二醇是药品生产, 特别是 HMG-CoA-还原酶抑制剂生产过程中的重要中间化合物。所述手性二醇是例如叔-丁基(3R,5S)-6-氯-3,5-二羟基己酸酯、叔-丁基(3R,5R)-6-氰基-3,5-二羟基己酸酯或者叔-丁基(5S,3R)-3,5,6-三羟基己酸酯。

这种类型的二醇特别是通过对映体选择性还原相应的通式 I 羟基酮, 例如叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯、叔-丁基(5S)-6-氯-5-羟基-3-氧代己酸酯或者叔-丁基(5S)-5,6-二羟基-3-氧代己酸酯而形成。这种化学催化还原反应的缺点是, 一方面, 苛刻的反应条件会形成副产物, 另一方面, 形成不需要的对映体或者非对映异构体过量, 并且该技术的成本非常高。

由于上述原因, 人们一直在探索生物催化方法, 以使用所述方法对映体选择性还原所述的羟基酮。生物催化方法通常在温和的条件下完成, 期待着在没有进一步的副产物形成的情况下, 还原已经很不稳定的 5-羟基-3-氧代己酸酯衍生物。但是, 迄今一直没有找到适宜的生物催化剂, 用所述催化剂能够有效地酶

促还原所述的 5-羟基-3-氧代己酸酯并且是用分离的酶。

美国专利 6001615 和国际专利申请 WO01/85975 A1 例如描述了用来自白僵菌(*Beauveria*)、*Pichia*、念珠菌、克鲁维酵母菌、*Torulaspora* 的不同酵母还原叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯或者叔-丁基(5S)-5,6-二羟基-3-氧代己酸酯的方法。但是所述转化使用野生菌株的整个细胞,并且在远低于 5%的很小浓度下完成。这是由于迄今为止,尚未鉴定出与转化有关的酶以及其 DNA 序列。

此外,在 EP 0569 998 B1 中描述了微生物转化结构相似的化合物。其中从 *Acinetobacter calcoaceticus* ATCC 33305 中纯化了 NADH 依赖性酶,并以与葡萄糖脱氢酶一起分离的形式,用于辅酶再生。在所述方法中,所述底物的浓度为 1%,NADH 的“总转化数”仅为 10。因此,尚未描述适于工业应用的方法。

在美国专利 6645746 B1 中公开了 *Candida magnoliae* 的一种氨基酸序列,用所述氨基酸序列可以有助于 NADPH 还原叔-丁基(5S)-6-氯-5-羟基-3-氧代己酸酯至叔-丁基(3R,5S)-6-氯-3,5-二羟基己酸酯。在该文献的说明书中,优选将所述酶与共表达的巨大芽孢杆菌的葡萄糖脱氢酶一起使用,葡萄糖脱氢酶有助于辅因子 NADPH 的再生,且葡萄糖作为共底物而产生。

本发明面对该任务,提供了一种方法,该方法可以经济地以高产率和高对映体纯度,且没有副产物而产生通式 II 的纯对映体二醇。

根据本发明,由前述类型的方法而完成了所述任务,即在 NADH 或者 NADPH 作为辅因子的条件下,通过氧化还原酶还原而实现,其特征在于,其中,

a) 羟基酮在反应混合物中的浓度 $\geq 50\text{g/l}$,

b) 通过通式 $\text{R}_\text{X}\text{R}_\text{Y}\text{CHOH}$ 的二级醇氧化持续再生所得的氧化辅因子 NAD 或者 NADP,其中 $\text{R}_\text{X}\text{R}_\text{Y}$ 彼此独立是氢、支链或者无支链 $\text{C}_1\text{-C}_8$ 烷基,且 $\text{C}_{\text{总数}} \geq 3$,和

c) 羟基酮的还原和二级醇的氧化由相同的氧化还原酶催化。

因此,所述方法优选实施方案的特征在于,所述氧化还原酶

a) 含有氨基酸序列,其中至少 50%的氨基酸与氨基酸序列 SEQ ID NO:1、SEQ ID NO:8、SEQ ID NO:11 或者 SEQ ID NO:14 的那些相同,

b) 由 SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:13 或者 SEQ ID NO:15 的核酸序列编码,或者

c) 由在严格条件下,与 SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:9、SEQ ID

NO:10、SEQ ID NO:12、SEQ ID NO:13 或者 SEQ ID NO:15 杂交的核酸序列编码。

根据本发明，上述任务还可以通过前述类型的方法完成，其中，所用氧化还原酶

- a) 含有氨基酸序列，其中至少 50%的氨基酸与氨基酸序列 SEQ ID NO:1、SEQ ID NO:8、SEQ ID NO:11 或者 SEQ ID NO:14 的那些相同，
- b) 由 SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:13 或者 SEQ ID NO:15 的核酸序列编码，或者
- c) 由在严格条件下，与 SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:13 或者 SEQ ID NO:15 杂交的核酸序列编码。

已发现具有 SEQ ID NO:1、SEQ ID NO:8 和 SEQ ID NO:11 氨基酸序列的多肽具有氧化还原酶活性，并可用于以>99%的非对映异构体过量将叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯还原为叔-丁基(3R,5R)-6-氰基-3,5-二羟基己酸酯。用同样的方法，用 SEQ ID NO:1、SEQ ID NO:8 和 SEQ ID NO:11 氨基酸序列可以还原通式 I 的其他羟基酮。

另外，所述氧化还原酶所具有的优点是，能够将在还原反应过程中形成的氧化的辅因子，通过二级醇的还原而再生。即，所述氧化还原酶的一种特别经济的优点也是，与现有技术的方法(US6645746 B 和 EP0569998 B)相比，不必再用其他酶使辅因子再生。

编码 SEQ ID NO:1 之多肽的 DNA 序列 SEQ ID NO:2 例如来自喜木聚糖红色杆菌 (*Rubrobacter xylophilus*) DSM 9941 的基因组。此外，还发现，可以用 DNA 序列 SEQ ID NO:3，以便在大肠杆菌中表达 SEQ ID NO:1 的多肽。

编码 SEQ ID NO:8 之多肽的 DNA 序列 SEQ ID NO:9 例如来自 *Geobacillus thermodenitrificans* DSM 465 的基因组。此外，还发现，可以用 DNA 序列 SEQ ID NO:10，以便在大肠杆菌中表达 SEQ ID NO:8 的多肽。

编码 SEQ ID NO:11 之多肽的 DNA 序列 SEQ ID NO:12 来自橙色绿屈挠菌 DSM 635 的基因组。

编码 SEQ ID NO:14 之多肽的 DNA 序列 SEQ ID NO:15 来自 *Candida magnoliae* CBS 6396。

因此,本发明还涉及用氨基酸序列 SEQ ID NO:1、SEQ ID NO:8、SEQ ID NO:11 或者 SEQ ID NO:14 之多肽,或者含有与氨基酸序列 SEQ ID NO:1、SEQ ID NO:8、SEQ ID NO:11 或者 SEQ ID NO:14 的氨基酸序列有至少 50%的同一性的氨基酸序列的多肽,即通过取代、插入、缺失或增加 SEQ ID NO:1、SEQ ID NO:8、SEQ ID NO:11 或者 SEQ ID NO:14 的至少一个氨基酸而获得的多肽,或者用由 SEQ ID NO:2、SEQ ID NO:9、SEQ ID NO:12 或者 SEQ ID NO:15 的核苷酸序列,或者由与 SEQ ID NO:2、SEQ ID NO:9、SEQ ID NO:12 或者 SEQ ID NO:15 的核苷酸序列在严格条件下杂交之核苷酸序列编码的多肽,将通式 I 的羟基酮还原成通式 II 的二醇的方法。

例如在严格条件下与 SEQ ID NO:2 杂交的核苷酸序列应理解为,是指用 SEQ ID NO:2 作为探针,通过菌落杂交方法、噬菌斑杂交方法、Southern 杂交法或者类似的方法可以鉴定的多核苷酸。

为此目的,在 0.7-1 M NaCl 溶液中, 60°C, 将在过滤器上固定的多核苷酸与 SEQ ID NO:2 杂交。用描述于例如, 分子克隆(Molecular Cloning), 实验室手册, 第二版(Cold Spring Harbor Laboratory Press, 1989)或者类似文献中的方法, 完成所述杂交。最后,用 0.1-2 倍 SSC 溶液,于 65°C 冲洗所述过滤器,其中,1 倍 SSC 溶液是一种混合物,由 150 mM NaCl 和 15mM 柠檬酸钠组成。

根据本发明方法,含 SEQ ID NO:1、SEQ ID NO:8、SEQ ID NO:11 或者 SEQ ID NO:14 序列之多肽,或者由所述多肽衍生的多肽可以以完全纯化、部分纯化或者以含 SEQ ID NO:1、SEQ ID NO:8、SEQ ID NO:11 或者 SEQ ID NO:14 多肽之细胞的形式使用。这里所用的细胞可以是天然的、渗透性的或者溶解形式。优选在一种适宜的宿主生物,例如大肠杆菌中,过表达含有 SEQ ID NO:1、SEQ ID NO:8、SEQ ID NO:11 或者 SEQ ID NO:14 序列的多肽或者其衍生物,然后用重组多肽还原通式 I 的羟基酮。

每公斤待转化的通式 I 的化合物使用 5000 至 10 Mio U 氧化还原酶 SEQ ID NO:1、SEQ ID NO:8、SEQ ID NO:11 或者 SEQ ID NO:14 或其衍生物(如上所公开的)。这里酶单位 1U 是每分钟(min)转化 1 μ mol 式 I 的羟基酮所需要的酶量。

根据本发明的酶促还原反应在温和的条件下进行,使得通常不稳定的羟基酮的降解和因此不需要的副产物的形成可以在很大程度上受到抑制。根据本发明方法,手性二醇形成的时间长,且非对映异构体的纯度通常>99%。

本发明优选实施方案的特征在于,在所述方法中用共底物持续地还原所用的辅因子。

优选用 NAD(P)H 作为辅因子,其中在还原反应中形成的 NAD(P)由共底物再次还原成 NAD(P)H。

这里,优选用二级醇,例如 2-丙醇、2-丁醇、2-戊醇、3-戊醇、4-甲基-2-戊醇、2-庚醇、2-辛醇或者环己醇作为共底物。根据特别优选的实施方案,将 2-丙醇用于辅酶再生。再生所用的共底物含量可占总体积的 5-95 Vol%。

同样例如用含有 SEQ ID NO:1、SEQ ID NO:8、SEQ ID NO:11 或者 SEQ ID NO:14 的多肽完成辅酶再生。

根据本发明方法,优选通式 I 的羟基酮占总体积的 5 重量%-50 重量%(50 g/l-50g/l),优选 8 重量%-40 重量%,更优选 10 重量%-25 重量%。

一个特别优选的实施方案的特征在于,用叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯、叔-丁基(5S)-6-氯-5-羟基-3-氧代己酸酯或者叔-丁基(5S)-5,6-二羟基-3-氧代己酸酯作为通式 I 的羟基酮。

优选本发明方法在水-有机两相系统中完成。

其中进行酶促还原反应的反应混合物的水部分,优选含有 pH 为 5-10,优选 pH6-9 的缓冲液,例如磷酸钙-、Tris/HCl-或者三乙醇胺-缓冲液。为了稳定或者激活酶,所述缓冲液还可以含有离子,例如锌离子或者镁离子。

合适地,本发明方法的反应温度为大约 10°C-70°C,优选 20°C-45°C。

在本发明方法进一步优选的实施方案中,所述酶促转化在与水不混溶或仅与水有限可混溶的有机溶剂存在下进行。所述溶剂是例如对称或者不对称二(C₁-C₆)烷基醚,直链(geradkettiges)或者支链烷或者环烷或者水溶性二级醇,同时,形成共底物。优选的有机溶剂是乙醚、叔-丁基甲基醚、二异丙基醚、二丁基醚、乙酸丁酯、庚烷、己烷、2-辛醇、2-庚醇、4-甲基-2-戊醇和环己醇。后者所述二级醇的情况下溶剂可同时可作为共底物使辅因子再生。

在加入水溶性溶剂或者共底物时,反应混合物由水和有机相组成。所述羟基酮根据其相应的溶解性分散在有机相和水相之间。所述有机相通常占总反应体积的 5-95%,优选 10-90%比例。优选将所述两个液相进行机械混合,这样在两相之间形成一个大上层。例如在本实施方案中,酶促还原反应形成的 NAD 还可利用共底物,如上所述,再还原成 NADH。

辅因子,特别是 NADH 或者 NADPH 在水相中的浓度通常为 0.001 mM-10 mM,特别是 0.01mM-1 mM。

根据本发明方法,还可以使用氧化还原酶/脱氢酶的稳定剂。适宜的稳定剂是例如甘油、山梨醇、1,4-DL-二硫苏糖醇(DTT)或者二甲亚砜(DMSO)。

本发明方法例如在密封的玻璃或者金属反应容器中完成。所述组分分别转移到所述反应容器中,并在气体,例如氮气或空气下搅拌。

根据本发明进一步可能的实施方案,可以将氧化的共底物(例如丙酮)不断除去,和/或不断补充新的共底物(例如 2-丙醇),以便反应向反应产物(通式 II 的二醇)方向进行。

在另一实施方案中,在反应过程中逐渐加入 SEQ ID NO:1、SEQ ID NO:8、SEQ ID NO:11 或者 SEQ ID NO:14 的氧化还原酶和/或共底物。

反应结束后,处理反应混合物。如果需要,将水相与有机相分离,然后过滤含产物的有机相。如果需要还可以提取水相,如同进一步处理有机相一样。然后,蒸发有机相的溶剂,其中含有作为粗产物的通式 II 的二醇。可以进一步纯化所述粗产物,或者直接用于下面产物的合成。

下列实施例将用于进一步说明本发明。

实施例 1

从喜木聚糖红色杆菌 DSM 9941 中克隆氧化还原酶

A) 培养喜木聚糖红色杆菌 DSM 9941

在细菌振荡培养箱中,将喜木聚糖红色杆菌 DSM 9941 细胞于 50°C(pH7.2),以 140rpm,在下列介质中培养:0.1%酵母提取物、0.1% Trypton、0.004% CaSO₄ x2H₂O、0.02% MgCl₂ x6H₂O、0.01%次氨基三乙酸、100 mg 磷酸盐缓冲液[5.44 g/l KH₂PO₄,43 g/l Na₂HPO₄ x12H₂O]、500μl/l 0.01 M 柠檬酸铁、500μl/l 示踪元素 [500μl/l H₂SO₄, 2.28 g/l MnSO₄ x H₂O, 500 mg/l ZnSO₄ x7 H₂O, 500mg H₃BO₃, 25 mg/l CuSO₄ x 5 H₂O, 25mg/l Na₂MoO₄ x2 H₂O, 45 mg/l CoCl₂ x 6 H₂O]。在培养第 6 天后,将培养介质离心分离细胞,然后于-80°C贮存。

B) 扩增编码选择性氧化还原酶的基因

按照 Manniatis & Sambrook 在“分子克隆”中的描述方法提取基因组 DNA。用所得的核酸作为聚合酶链反应(PCR)反应的模板,这里使用 NCBI 数据库中 46106817 公开的基因序列衍生的特异性引物。这里,引物在 5' -末端提供有

用于核酸内切酶 Nde I 和 Hind III 的限制性切割位，以用于随后克隆到表达载体。(SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:7)。

在 PCR 缓冲液[10 mM Tris-HCl(pH8.0); 50 mM KCl; 10 mM MgSO₄; 1mM dNTP Mix; 每 20 pMol 引物和 2.5 U 铂 Pfx DNA 聚合酶(Invitrogen)]中，使用 500ng 基因组 DNA 和下列温度循环进行扩增：

循环 1: 94°C, 2 分钟

循环 2x30: 94°C, 15 秒

54°C, 30 秒

68°C, 60 秒

循环 3: 68°C, 7 分钟

4°C, ∞

所得大小约 750 bp 的 PCR 产物用 1%琼脂糖凝胶纯化后，用核酸内切酶 Nde I 和 Hind III 或者 SphI 和 Hind III 限制切割，然后与用相同核酸内切酶处理过的 pET21a 载体主链(Novagen)或者 pQE70 载体(Qiagen)连接。转化 2μl 连接液至大肠杆菌 Top 10F'细胞(Invitrogen)后，用核酸内切酶 Nde I 和 Hind III 或者 SphI 和 Hind III 限制分析检测质粒-DNA-氨苄青霉素-抗性菌落，检测 750 bp 大小的插入片段的存在情况。对来自所述片段正性的菌落的质粒制备物进行序列分析，然后转化至大肠杆菌 BL21 菌株(Invitrogen)或者大肠杆菌 RB791(genetic stock, Yale)。

实施例 2

在大肠杆菌细胞中有效表达 SEQ ID NO:1 多肽

为了在大肠杆菌细胞中有效表达 SEQ ID NO:1 多肽，为了在表达载体中克隆，用编码 DNA SEQ ID NO:3 作为 PCR 反应中的模板。所述 DNA 序列与前述 DNA 序列(SEQ ID NO:2)的前 160 个碱基对有 51 个碱基不同。所述修饰是保守的，不会改变氨基酸序列。

在含 50 ng DNA SEQ ID NO:3 作为模板的 PCR 缓冲液[10 mM Tris-HCl, (pH8.0); 50 mM KCl; 10 mM MgSO₄; 1mM dNTP Mix; 每 20 pMol 引物(SEQ ID NO:6、SEQ ID NO:5)和 2.5 U 铂 Pfx DNA 聚合酶(Invitrogen)]中，按下列温度循环完成扩增：

循环 1: 94°C, 2 分钟
循环 2x30: 94°C, 40 秒
56°C, 30 秒
68°C, 60 秒
循环 3: 68°C, 7 分钟
4°C, ∞

所得大小约 750 bp 的 PCR 产物用 1%琼脂糖凝胶纯化后,用核酸内切酶 Nde I 和 Hind III 在与用相同核酸内切酶处理过的 pET21a 载体(Novagen)主链中连接。转化 2 μ l 连接液至大肠杆菌 Top 10F'细胞(Invitrogen)后,用核酸内切酶 Nde I 和 Hind III 限制分析检测质粒-DNA-氨苄青霉素-抗性菌落,检测 750 bp 大小的插入片段的存在情况。对来自所述片段正性的菌落的质粒制备物进行序列分析,然后转化至大肠杆菌 BL21 菌株(Invitrogen)

实施例 3

来自喜木聚糖红色杆菌 DSM 9941 的氧化还原酶的制备

将表达构建体转化的大肠杆菌菌株 BL21 Star(Invitrogen, Karlsruhe, 德国)或者 RB791(大肠杆菌 genetic stock, Yale, USA)在含氨苄青霉素(50 μ g/ml)的培养基(1% Tryptone, 0.5%酵母提取物, 1% NaCl)中培养,直到光密度可达到 0.5 在 550nm 下测量。通过加入 0.1 mM 异丙基硫代半乳糖苷(IPTG)诱导重组蛋白的表达。在 25°C, 220rpm 诱导后 16 小时,收集细胞,并冷冻于-20°C。

为了进行酶提取,将 30g 细胞悬浮于 150 ml 三乙醇胺缓冲液(100 mM, pH=7, 2 mM MgCl₂, 10%甘油)中,然后用高压匀浆机分解。然后,将所述酶溶液用 150ml 甘油混合,并于-20°C贮存。

将这样获得的酶溶液用于合成叔-丁基(3R,5R)-6-氰基-3,5-二羟基己酸酯。

在稀释的情况下,所得的酶溶液也用于相应的酶测量。活性检测按如下: 870 μ l 100mM TEA-缓冲液, pH7.0, 160 μ g NAD(P)H,10 μ l 稀释的细胞溶解物。通过向反应混合物中加入 100 μ l 100 mM 叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯溶液而启动反应。

实施例 4

用氧化还原酶 SEQ ID NO:1 将叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯转化为叔-丁基(3R,5R)-6-氰基-3,5-二羟基己酸酯

为了将叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯转化为叔-丁基(3R,5R)-6-氰基-3,5-二羟基己酸酯,按 Eppendorf 反应设备中,将由下列物质组成的混合物在室温恒定混合下培养 24 小时: 900 μ l 缓冲液(100 mM TEA, pH=7, 1mM MgCl₂), 100 μ l 2-丙醇、10 μ l 叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯粗产品(对映异构体纯度>99%)、0.1 mg NAD 和 100 μ l 酶悬浮液(见实施例 3)。24 小时后, 96%所用叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯还原为叔-丁基(3R,5R)-6-氰基-3,5-二羟基己酸酯。剩余非对映异构体过量>99%。

用手性气相色谱跟踪测定转化率以及非对映异构体过量。为此, 使用 Shimadzu 的 Gax-Chromatograph GC-17A, 其带有手性分离柱 CP-Chirasil-DEX CB(Varian Chrompack, Darmstadt, 德国),Flammen-Ionisations-Detektor, 用氦作为载气。

在 0.72 bar, 10 分钟, 50°C, 5°C/分钟 $\rightarrow$ 200°C 10 分钟分离叔-丁基-6-氰基-3,5-二羟基己酸酯。

保留时间为: (R-BCH)4.4 分钟; (R,R-BCH)47.1 分钟和(R,S-BCH)48.2 分钟。

实施例 5

用氧化还原酶 SEQ ID NO:1 从叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯合成为叔-丁基(3R,5R)-6-氰基-3,5-二羟基己酸酯

为了进一步将叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯转化成为叔-丁基(3R,5R)-6-氰基-3,5-二羟基己酸酯,在 Eppendorf 反应设备中,将由下列物质组成的混合物培养: 550 μ l 缓冲液(100 mM TEA, pH=7, 1mM MgCl₂), 150 μ l 2-丙醇、200 μ l 叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯粗产品(对映异构体纯度>99%)、0.1 mg NAD 和 200 μ l 酶悬浮液(见实施例 3)。在反应过程中,产生的丙酮/2-丙醇混合物多次通入氮气蒸发,并加入新鲜的 2-丙醇和 50 μ l 酶。24 小时内重复 2-3 次后, >90%的所用叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯被还原为叔-丁基(3R,5R)-6-氰基-3,5-二羟基己酸酯。非对映异构体过量>99%。

实施例 6

通过氧化还原酶 SEQ ID NO:1 从叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯合成叔-丁基(3R,5R)-6-氰基-3,5-二羟基己酸酯

为了进一步将叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯转化为叔-丁基(3R,5R)-6-氰基-3,5-二羟基己酸酯,在反应容器中,在 45°C 培养下列混合物: 6.7

ml 缓冲液(100 mM TEA, pH=9), 1.7 ml 2-丙醇、2 ml 叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯粗产品(对映异构体纯度>99%)、1.0 mg NAD 和 150 mg 冷冻的大肠杆菌 BL21 菌株细胞, 所述细胞中含氧化还原酶 SEQ ID NO:1(见实施例 3)。24 小时后, >90%的所用叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯被还原为叔-丁基(3R,5R)-6-氰基-3,5-二羟基己酸酯。非对映异构体过量>99%。

实施例 7

通过氧化还原酶 SEQ ID NO:8 从叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯合成叔-丁基(3R,5R)-6-氰基-3,5-二羟基己酸酯

为了进一步将叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯转化为叔-丁基(3R,5R)-6-氰基-3,5-二羟基己酸酯, 在反应容器中, 在 40℃保温下列混合物: 5.0 ml 缓冲液(100 mM TEA, pH=7.5), 2.0 ml 2-丙醇、4.0 ml 叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯粗产品(对映异构体纯度>99%)、1.0 mg NAD 和 250 mg 冷冻大肠杆菌 RB791 菌株细胞, 所述细胞中含氧化还原酶 SEQ ID NO:8(见相应实施例 2 和 3)。24 小时后, >90%的所用叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯被还原为叔-丁基(3R,5R)-6-氰基-3,5-二羟基己酸酯。非对映异构体过量>99%。

实施例 8

通过氧化还原酶 SEQ ID NO:11 从叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯合成叔-丁基(3R,5R)-6-氰基-3,5-二羟基己酸酯

为了将叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯转化为叔-丁基(3R,5R)-6-氰基-3,5-二羟基己酸酯, 在 Eppendorf 反应容器中, 在室温且恒定混合下, 保温下列混合物 48 小时: 350 μ l 缓冲液(100 mM 磷酸钙, pH=7), 150 μ l 2-丙醇、50 μ l 叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯粗产品(对映异构体纯度>99%)、0.025 mg NAD 和 15 μ l 酶悬浮液 SEQ ID NO:11(见实施例 3)。48 小时后, >80%的所用叔-丁基(5R)-6-氰基-5-羟基-3-氧代己酸酯被还原为叔-丁基(3R,5R)-6-氰基-3,5-二羟基己酸酯。非对映异构体过量>99%。

SEQ ID NO:1

1 mlegkvavit gagsgigrat alrfaregar vvaelderr
41 geevvreile sggeavfvt dvsefeqvea averaveeyg
81 tidvmfnag ighyaplleh dpehydrvr vnqygvygi
121 laagrkmæl enpgviinta svyafiaspg vigyhaskga
161 vkmmtqaaal elaphgirvv aiapggvdtg iiqgykdmgl
201 gerlargqmr rrlqtpeqia gavvllatee adaingsvwm
241 tddgyaefk

SEQ ID NO:2

1 atgctcgaggggaaggtcgcggtcatcacgggggccggca
41 gcggcataggccgggccaccgcgctcaggttcgcccgcga
81 agggggcccggtggtcgtggcgagctcgacgagcggagg
121 ggggaggaggtcgtccgggagatcctcgagtcggcgggg
161 aggccgtcttcgtgaggacggacgtctcggagttcgagca
201 ggtgagggccgcccgtcgagcgccgctcgaggagtacggg
241 acgctggacgtcatgttcaacaacgccggcatcgggcact
281 acgccccctgctggagcacgacccggagcactacgaccg
321 ggtggtccgggtgaaccagtacggcgtctactacgggata
361 ctgcccggcggcaggaagatggccgagctggagaaccccg
401 gcgtgatcatcaacaccgcctcgggtctacgtttcctggc
441 cccccgggtgtgatcggtatcacgcttcaagggggcg
481 gtgaagatgatgaccagggccgcagccctggagctcgccc
521 cccacggcatacgggtcgtcgccatgccccggcggggt
561 ggacacccgatcatccagggtacaaggacatgggcctc
601 ggtgagcggctggcccggccagatgcgtcgaggctcc
641 agacccccgagcagatcgccggcggcgtcgtcctgctcg
681 caccgaggaggcagacgccataaacggctcgggtggtgatg
721 accgacgacggctacgcggagttcaagtaa

SEQ ID NO:3

1 atgctggaag gtaaagtggc agtcatcacc ggtgcaggca
41 gcggcattgg gcgtgccact gcgtgcgtt ttgcgcgtga
81 aggcgctcgc gtcgttgtg cagagctgga tgaacgtcgc
121 ggtgaggaag ttgtacgtga gattctgga tctggcgggg
161 gggaggccgt cttcgtgagg acggacgtct cgaggtcga

201 gcaggttag gccgccgtc agcgccgct cgaggagtag
 241 gggacgctgg acgtcatgtt caacaacgcc ggcacgggc
 281 actacgccc cctgctggag cagacccgg agcactacga
 321 ccgggtggc cgggtgaacc agtacggcgt ctactacggg
 361 atactcgccg ccggcaggaa gatggccgag ctggagaacc
 401 ccggcgtgat catcaacacc gcctcggctt acgctttcct
 441 ggctcccc ggtgtgatc gctatcacgc ttccaagggg
 481 gcggtgaaga tgatgacca ggccgcagcc ctggagctcg
 521 ccccccacgg catacgggtc gtcgccatcg ccccgggcgg
 561 ggtggacacc ccgatcatcc agggctacaa ggacatggg
 601 ctggtgagc ggctggccc cggccagatg cgtcgaggc
 641 tccagacccc cgagcagatc gccggcggc tctctctgct
 681 cgccaccgag gaggcagacg ccataaacgg ctggtggtg
 721 atgaccgacg acggctacgc ggagttcaag taa

SEQ ID NO:4

GGGAATTCCATATGATGCTCGAGGGGAAGGTCTG

SEQ ID NO:5

CCCAAGCTTATTACTTGAAGTCCGCGTAGCCGTC

SEQ ID NO:6

CCTAGCTAGCATGCTGGAAGGTAAAGTGGC

SEQ ID NO:7

CACATGCATGCGAATGCTCGAGGGGAAGGTC

SEQ ID NO:8

Geobacillus thermodenitrificans DSM 465 蛋白质序列羧基还原酶

1 mrlkgkaaiv tggasgigra tairfaeega kvavsdinee ggeetvrlir ekggeaifvq
 61 tdvadsqvs rlvqtavdaf gglhilfnna gighsevrst dlseeewdrv invnlkgvfl
 121 gikyavpvmk qcgggaivnt sslgikgkk yesaynaska gvilltknaa leygkfnirv
 181 naiapgvidt niitpwkqde rkwpiskn algrigtpee vanavlflas deasfitgat
 241 lsvdgggltf

SEQ ID NO:9*Geobacillus thermodenitrificans* DSM 465 核酸序列 羧基还原酶

```

1  atgaggctaa aaggaaaagc ggcgattgtc accggcggcg cgagcggcat cggccgggcg
61  acggcgattc gctttgcgga agaaggcgcc aaagtggcgg tgagcgacat caatgaggaa
121 ggaggggaag aaacgggtccg cctgattcgg gaaaaaggag gggaggcgat ttttgtccaa
181 acggacgtag ccgattccaa gcaagtgagc cgccttgtcc aaacggcggg tgatgccttt
241 ggcggcctac atattctctt taacaatgcc ggcacggcc attcggagt gcggagcacc
301 gacttgtctg aagaagagtg ggaccgggtc atcaacgtta atttgaaagg agtgttcctt
361 ggcacaaat acgcggtgcc cgtgatgaag caatgcggtg gcggggccat tgtcaacaca
421 tcgagcctgc ttggaatcaa agggaaaaag tacgaatcgg cctacaacgc ctcaaggcc
481 ggggtgattt tgttgacgaa aaatgcagca ttggaatatg ggaagtttaa cattcgctc
541 aatgccattg caccgggggt cattgatacg aacatcatca cgccgtggaa acaagatgag
601 cgcaaatggc cgatcatttc gaaagcgaac gccctcgcc gcacgggac gccagaggaa
661 gtggcgaacg cgggtgtgtt tttggcgctc gatgaagcgt cgtttatcac cggcgcgaca
721 ttgtcggtcg acggcggcg gctgacgttt tag

```

SEQ ID NO:10*Geobacillus thermodenitrificans* DSM 465 核酸序列 合成基团

```

1  atgcgcctga aagggaagc ggcaattgtg acgggtggcg ccagcggcat cgggcgcgcg
61  actgcgatcc gttttgcaga agaggtgctg aaagttgccg ttagcgacat taacgaggaa
121 ggcggtgagg aaaccgttcg cctgatccgt gaaaaaggcg gtgaggcaat ctctgtgcag
181 acggatgtgg ccgactcaaa acaggatatc cgtctggttc agaccgcggt cgacgcgttt
241 ggtggcctgc acatcctgtt caataacgcc ggcattggcc atagcgaagt gcgtagtact
301 gatctgagcg aggaagagtg ggatcgcgtg attaacgtga acctgaaagg tgtgtttctg
361 ggtattaagt atgcagtccc tgttatgaaa cagtgtggcg gtggtgcgat tgtgaatacc
421 tctagtctgt tgggaatcaa agggaaaaag tatgaatcgg cctacaacgc atcgaaagcc
481 ggcgcatcc tgctgaccaa aaatgcggcc ctggagtatg gcaagttcaa tattcgtgtc
541 aacgcgatcg ctccaggcgt tatcgatacc aacatcatta caccgtggaa gcaagatgaa
601 cgcaagtggc cgattatctc caaagctaag gcgctgggcc gtatcggtac gccggaagaa
661 gtggctaata cggttctgtt tctggcaagc gacgaagcga gctttattac gggtgcaacc
721 ctctccgtag atggggggcg gttaaccttc taa

```

SEQ ID NO:11

橙色绿屈挠菌

DSMZ 635

蛋白质序列羧基还原酶

```

1  meppfigkva lvtgaaagig rasalafare gakkvvadvn veggeetial cralntdamf
61  vrcdvsqrde verlialavd tfgridfahn nagiegvqam ladypeevwd rvieinlkgv
121 wlcmyeirh mlkqgggaiv ntssvaglag srgvsayvas khgivgitka aaleyarngi
181 rvnaicpgti htamidrftg gdpqlaqa egepigrigs peevanaviw lcsdkasfvt
241 gatlavdggr la

```

SEQ ID NO:12

橙色绿屈挠菌

DSMZ 635

核酸序列羧基还原酶

```

1  atggagccac ctttcattgg gaaggttgcg ctggtcaccg gcgcagcagc cggatttggt
61  cgtgcttcag cactggcggt tgcccgtag ggtgccagg ttgtcggtgc tgatgtgaat
121 gtcgagggcg gggaagagac gattgcgctg tgcgggctt tgaataccga tgcaatgttc
181 gtgcgtttgt atgtttcgca acgcgatgaa gtggagcgat taattgctct ggcagttgac
241 acgttcggtc ggatcgactt tgcgcacaac aacgccggga ttgaaggcgt gcaggcaatg
301 ctggccgatt atcccgaaga ggtctgggat cgggtgatcg agatcaacct caaaggggtc
361 tggttgtgta tgaagtacga aatccggcac atgctcaagc aggggtggcg tgcatgtgtg
421 aatacctcat cggtcgccgg tctggccgga tcacgtggcg tttcggcgta ttagaccagc
481 aagcacggta ttgttggtat taccaaagcg gcagcccttg agtatgcgcg taacggtatt
541 cgtgtcaacg caatctgtcc aggtacgatt catactgca tgatcgaccg ctttaccag
601 ggtgatcccc aactgcttgc ccagttcgct gagggatgaac cgattggctg gtcggtctcg
661 cctgaagagg tcgccaatgc ggtgatctgg ctctgctcag ataaggcttc gtttgtgacc
721 ggagcgacac tggcggttga tgggtggcgc ctggcgtaa

```

SEQ ID NO:13

橙色绿屈挠菌

DSMZ 635

核酸序列合成基团

```

1  atggagcccc catttatcgg gaaagttgcg ttagttacgg gggcagcggc agggatcggc
61  agggcgagtg ccctggcggt tgctagagaa gggccaagg tcgttggtgc agacgttaac
121 gtagaggggt gcgaagagac aattgcttta tgcagagctc tcaacactga tgccatgttc
181 gtccgctgtg atgtgtcaca gcgagacgaa gtcgaaagge taatcgccct agcggtagac
241 acattcggcc gtattgactt tgctcataat aacgcgggca tagaggaggt acaagcaatg

```

301 ttggttgact atcctgagga agtatgggat cgagtaattg aaatcaatct caaggggggtt
 361 tggctgtgta tgaagtacga aataaggcac atgctcaagc aaggtggcgg agcgatcgta
 421 aacactagct ctgtcgccgg tctagcagga tctcgggggg tttccgcata cgtcgcctcg
 481 aaacacggca ttgtagggat taccaaagct gcagcccttg agtatgccc aaatggaata
 541 agagtgaatg ctatctgccc aggcacaata catactgcaa tgatagatcg gtttacgcag
 601 ggtgatccgc aacttttggc gcagttcgcc gaaggtgagc ctataggteg ccttggtagc
 661 ccggaagagg tcgctaatagc ggtgatttgg ttgtgttcag acaaagcaag tttcgtgacg
 721 ggagctaccc tggcagtgga tggaggacgt ttagct

SEQ ID NO:14

Candida magnoliae CBS 6396

蛋白质序列羧基还原酶

1 msatsnalit gasrgmgeat aiklalegys vtlasrgieq lnaikeklpi vkkggqhyvw
 61 qldlsdieaa stfkgaplpa ssydvffsna gvvdapfad qsetaqkdlf tvnllspval
 121 tktivkaiad kpretpahii ftssivgiring vpvavysat kgaidsfars larefgpkni
 181 hvncvnpqgt rtentkgvdl aafgdpikg wievdaiada vlflikskni tggslvvdng
 241 fgv

SEQ ID NO:15

Candida magnoliae CBS 6396

核酸序列合成基团

1 atgtctgcta cttcgaacgc tcttatcaact ggtgccagcc gcggaatggg cgaggccaca
 61 gctattaagc ttgcccttga ggggtacagc gtcacccttg catcacgcgg tattgagcag
 121 ctcaatgcc acaaggaaaa actaccatc gtgaagaagg gccagcagca ctacgtttgg
 181 cagctcgatc ttagtgacat cgaggcggct tccacettca agggggctcc tctgcctgcc
 241 agcagctacg acgtgttctt cagcaacgcc ggtgtggtgg actttgctcc gttcgagac
 301 caaagcgaga ctgcgcaaaa ggacctgttc acggttaacc tgctgtcgcc tgttgcttg
 361 accaagacca ttgttaaggc catcgccgac aagccccgcg agacgcctgc tcacattatc
 421 ttcacctcgt ccattgtcgg aattcgcggt gttcccaacg tggcggtcta cagcgccacc
 481 aagggcgaga ttgacagctt tgcgcgctcg cttgctcgtg agttcggtcc caagaacatc
 541 cacgttaact gcgtgaaccc gggcacgacg cgcaccgaga tgacaaaggc cgttgatctc
 601 gcggctttcg gcgatgttcc tatcaagggc tggatcgagg tcgatcgcat tgccgacgct
 661 gtgctgtttt tgatcaagtc caagaacatc actggccagt cgctcgttgt tgacaacgga
 721 ttcggtgttt aa

<110> IEP GmbH

<120> 对映体选择性酶促还原羟基酮化合物的方法

<130> I 9844

<150> AT A 2027/2005

<151> 2005-12-19

<160> 15

<170> PatentIn version 3.3

<210> 1

<211> 249

<212> PRT

<213> 喜木聚糖红色杆菌 DSM 9941

<400> 1

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 Arg Phe Ala Arg Glu Gly Ala Arg Val Val
20 25 30

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

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

Phe Glu Gln Val Glu Ala Ala Val Glu Arg Ala Val Glu 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 Asp 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 Ala
 115 120 125

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

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

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 Val
 210 215 220

Leu Leu Ala Thr Glu 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> 2

<211> 750

<212> DNA

<213> 喜木聚糖红色杆菌 DSM 9941

<400> 2

atgctcgagg ggaaggtcgc ggtcatcacg ggggccggca gcggcatagg ccgggccacc	60
gcgctcaggt tcgcccgca aggggcccg gtggtcgtgg cggagctcga cgagcggagg	120
ggggaggagg tcgtccggga gatcctcgag tccggccggg aggccgtctt cgtgaggacg	180
gacgtctcgg agttcgagca ggttgaggcc gccgtcgagc gcgccgtcga ggagtacggg	240
acgctggacg tcatgttcaa caacgccggc atcgggcact acgccccct gctggagcac	300
gacccgagc actacgaccg ggtggtcgg gtgaaccagt acggcgtcta ctacgggata	360
ctgccgccg gcaggaagat ggccgagctg gagaacccg gcgtgatcat caacaccgcc	420
tcggtctacg ctttcttggc ctccccgggt gtgateggct atcacgttc caagggggcg	480
gtgaagatga tgaccaggc cgcagccctg gagctcgccc cccacggcat acgggtcgtc	540
gccatcgccc cgggcgggggt ggacaccccg atcatccagg gctacaagga catgggcctc	600
ggtgagcggc tggcccgcg ccagatgcgt cgcaggctcc agaccccgga gcagatcgcc	660
ggcgccgtcg tcctgtcgc caccgaggag gcagacgcca taaacggctc ggtggtgatg	720
accgacgacg gctacgcgga gttcaagtaa	750

<210> 3

<211> 753

<212> DNA

<213> 合成的构建体

<400> 3

atgctggaag gtaaagtggc agtcatcacc ggtgcaggca gcggcattgg gcgtgccact	60
gcgctgcgtt ttgcgcgtga aggcgtcgc gtcgttgtgg ccgagctgga tgaacgtcgc	120
ggtgaggaag ttgtacgtga gattctggaa tetggccggg gggaggccgt ctctgtgagg	180
acggacgtct cggagttcga gcaggttgag gccgccgtcg agcgcgccgt cgaggagtac	240

gggacgctgg acgtcatgtt caacaacgcc ggcacgggc actacgccc cctgctggag	300
cacgacccgg agcactacga ccgggtggtc cgggtgaacc agtacggcgt ctactacggg	360
atactcgccg ccggcaggaa gatggccgag ctggagaacc ccggcgtgat catcaacacc	420
gcctcggtct acgctttcct ggctccccc ggtgtgatcg gctatcacgc ttccaagggg	480
gcggtgaaga tgatgacca ggccgcagcc ctggagctcg ccccccacgg catacgggtc	540
gtcgccatcg ccccgggcgg ggtggacacc ccgatcatcc agggctacaa ggacatgggc	600
ctcggtgagc ggctggcccg cgccagatg cgtcgcaggc tccagacccc cgagcagatc	660
gccggcgccg tcgtcctgct cgccaccgag gaggcagacg ccataaacgg ctcggtggtg	720
atgaccgacg acggctacgc ggagttcaag taa	753

- <210> 4
- <211> 33
- <212> DNA
- <213> 合成的构建体

<400> 4	
gggaattcca tatgatgctc gaggggaagg teg	33

- <210> 5
- <211> 34
- <212> DNA
- <213> 合成的构建体

<400> 5	
cccaagctta ttacttgaac tccgcgtage cgtc	34

- <210> 6
- <211> 30
- <212> DNA
- <213> 合成的构建体

<400> 6	
---------	--

cctagctagc atgctggaag gtaaagtggc

30

<210> 7

<211> 31

<212> DNA

<213> 合成的构建体

<400> 7

cacatgcatg cgaatgctcg aggggaaggt c

31

<210> 8

<211> 250

<212> PRT

<213> Geobacillus thermodenitrificans DSM 465

<400> 8

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

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

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

Ile Arg Glu Lys Gly Gly Glu Ala Ile Phe Val Gln Thr Asp Val Ala
50 55 60

Asp Ser Lys Gln Val Ser Arg Leu Val Gln Thr Ala Val Asp Ala Phe
65 70 75 80

Gly Gly Leu His Ile Leu Phe Asn Asn Ala Gly Ile Gly His Ser Glu
85 90 95

Val Arg Ser Thr Asp Leu Ser Glu Glu Glu Trp Asp Arg Val Ile Asn

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

- <210> 9
- <211> 753
- <212> DNA
- <213> geobacillus thermodenitrificans DSM 465

<400> 9

atgaggctaa aaggaaaagc ggcgattgtc accggcggcg cgagcggcat cggccgggcg	60
acggcgattc gctttgcgga agaaggcgcc aaagtggcgg tgagcgacat caatgaggaa	120
ggaggggaag aaacgggtccg cctgattcgg gaaaaaggag gggaggcgat ttttgtccaa	180
acggacgtag ccgattccaa gcaagtgagc cgccttgtcc aaacggcggt tgatgccttt	240
ggcggcctac atattctctt taacaatgcc ggcacggccc attcggaagt gcggagcacc	300
gacttgtctg aagaagagtg ggaccgggtc atcaacgtta atttgaaagg agtggttcctt	360
ggcatcaaat acgcggtgcc cgtgatgaag caatgcggtg gcggggccat tgtcaacaca	420
tcgagcctgc ttggaatcaa agggaaaaag tacgaatcgg cctacaacgc ctgaaaggcc	480
ggggtgattt tgttgacgaa aaatgcagca ttggaatatg ggaagtttaa cattcgcgtc	540
aatgccattg caccgggggt cattgatacg aacatcatca cgccgtggaa acaagatgag	600
cgcaaatggc cgatcaattc gaaagcgaac gccctcggcc gcacggggac gccagaggaa	660
gtggcgaacg cggigtgttt ttggcgctcc gatgaagcgt cgtttatcac cggcgcgaca	720
ttgtcggteg acggcggcgg getgacgttt tag	753

<210> 10

<211> 753

<212> DNA

<213> 合成的构建体

<400> 10

atgcgcctga aagggaaagc ggcaattgtg acgggtggcg ccagcggcat cgggcgcgcg	60
actgcgatcc gtittgcaga agagggtgcg aaagttgccg tiagcgacat taacgaggaa	120
ggcggtgagg aaaccgttcg cctgatccgt gaaaaaggcg gtgaggcaat cttcgtgcag	180
acggatgtgg ccgactcaaa acaggtatct cgtctgttcc agaccgcggt cgacgcgttt	240
ggtggccctgc acatcctgtt caataacgcc ggcattggcc atagcgaagt gcgtagtact	300

gatctgagcg aggaagagtg ggatcgctg attaacgtga acctgaaagg tgtgtttctg 360
ggtattaagt atgcagtcce tgttatgaaa cagtgtggcg gtggtgcgat tgtgaatacc 420
tctagtctgt tgggaatcaa agggaaaaag tatgaatcgg cctacaacgc atcgaaagcc 480
ggcgtcatcc tgctgaccaa aaatgcggcc ctggagtatg gcaagttcaa tattegtgtc 540
aacgcgatcg ctccaggcgt tatcgatacc aacatcatta caccgtggaa gcaagatgaa 600
cgcaagtggc cgattatctc caaagctaat gcgctgggcc gtatcggtac gccggaagaa 660
gtggctaata cggtttctgt tctggcaagc gacgaagcga gctttattac gggtgcaacc 720
ctctccgtag atgggggcgg gttaaccttc taa 753

<210> 11
<211> 252
<212> PRT
<213> 橙色绿屈挠菌 DSM 635

<400> 11

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

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> 12
 <211> 759
 <212> DNA
 <213> 橙色绿屈挠菌 DSM 635

<400> 12
 atggagccac ctttcattgg gaaggttgcg ctggtcaccg gcgcagcagc cggtatttgt 60
 cgtgcttcag cactggcggtt tgcccgtgag ggtgccaaagg ttgtcgttgc tgatgtgaat 120
 gtcgagggcg gggaagagac gattgcgctg tgtcgggctt tgaataccga tgcaatgttc 180
 gtgcgttgtg atgtttcgca acgcgatgaa gtggagcgat taattgctct ggcagttgac 240
 acgttcggtc ggatecgactt tgcgcacaac aacgccggga ttgaaggcgt gcaggcaatg 300
 ctggccgatt atcccgaaaga ggtctgggat cgggtgatcg agatcaacct caaaggggtc 360
 tggttgtgta tgaagtaaga aatccggcac atgctcaagc agggtggcgg tgcgatttgt 420
 aatacctcat cggtcgccgg tctggccgga tcacgtggcg ttteggcgta tgtagccagc 480
 aagcacggta ttgttggtat taccaaagcg gcagcccttg agtatgcgcg taacggtatt 540
 cgtgtcaacg caatctgtcc aggtacgatt cctactgcga tgatcgaccg ctttaccag 600
 ggtgatcccc aactgcttgc ccagttcgct gaggggtgaac cgattggteg gctcggctcg 660
 cctgaagagg tcgccaatgc ggtgatctgg ctctgctcag ataagcttc gtttgtgacc 720
 ggagcgacac tggcggttga tggtggccgc ctggcgtaa 759

<210> 13
 <211> 756
 <212> DNA
 <213> 合成的构建体

<400> 13
 atggagcccc catttatcgg gaaagttgcg ttagttacgg gggcagcggc agggatcggc 60
 agggcgagtg cccctggcggtt tgctagagaa ggggccaaagg tcgttgttggc agacgttaac 120
 gtagagggtg gcgaagagac aattgcttta tgcagagctc tcaacactga tgccatgttc 180

gtccgctgtg atgtgtcaca gcgagacgaa gtcgaaaggc taatcgccct agcggtagac	240
acattcggcc gtattgactt tgctcataat aacgcgggca tagagggagt acaagcaatg	300
ttggctgact atcctgagga agtatgggat cgagtaattg aaatcaatct caagggggtt	360
tggtctgtga tgaagtacga aataaggcac atgtcaagc aagtggcgg agcgatcgta	420
aacactagct ctgtcgccgg tctagcagga tctcgggggg ttccgcata cgtegcctcg	480
aaacacggca ttgtagggat taccaaagct gcagcccttg agtatgcccg aaatggaata	540
agagtgaatg ctatctgccc aggcacata catactgcaa tgatagatcg gtttacgcag	600
ggtgatccgc aacttttggc gcagttcgcc gaaggtgagc ctataggteg cettggtagc	660
ccggaagagg tcgctaatgc ggtgatttgg ttgtgttcag acaaagcaag ttctgtgacg	720
ggagctaccc tggcagtgga tggaggacgt ttagct	756

- <210> 14
<211> 243
<212> PRT
<213> Candida magnoliae CBS 6396

<400> 14

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

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

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

Phe Gly Val

- <210> 15
- <211> 732
- <212> DNA
- <213> Candida magnoliae CBS 6396

<400> 15	
atgtctgcta cttcgaacgc tcttatact ggtgccagcc gcggaatggg cgaggccaca	60
gctattaagc ttgcccttga ggggtacagc gtcaccttg catcacgagg tattgagcag	120
ctcaatgcca tcaaggaaaa actacceatc gtgaagaagg gccagcagca ctacgtttgg	180
cagctcgatc ttagtgacat cgaggcggct tccaccttca agggggctcc tctgcctgcc	240
agcagctacg acgtgttctt cagcaacgcc ggtgtggtgg actttgctcc gttcgcagac	300
caaagcgaga ctgcgcaaaa ggacctgttc acggttaacc tgctgtcgcc tgttgcgttg	360
accaagacca ttgttaaggc catcgccgac aagccccgcg agacgcctgc tcacattatc	420
ttcacctegt ccatigtctg aattcgggt gttcccaacg tggcggctct cagegccacc	480
aagggcgcga ttgacagctt tgcgcgtcg cttgctcgtg agttcggtcc caagaacatc	540
cacgttaact gcgtgaaccc gggcacgacg cgcaccgaga tgacaaaggc cgttgatctc	600
gcggctttcg gcgatgttcc tateagggc tggatcgagg tcgatcgat tgccgacgct	660
gtgctgtttt tgatcaagtc caagaacatc actggccagt cgctcgttgt tgacaacgga	720
ttcggtgttt aa	732