(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 117120616 A
(43)申请公布日 2023.11.24

(21)申请号 202280025099.3
(22)申请日 2022.03.11
(30)优先权数据
2021-057308 2021.03.30 JP
(85)PCT国际申请进入国家阶段日
2023.09.26
(86)PCT国际申请的申请数据
PCT/JP2022/010912 2022.03.11
(87)PCT国际申请的公布数据
WO2022/209763 JA 2022.10.06
(83)生物保藏信息
NITE BP-02919 2019.03.18
FERM BP-11515 2011.10.25
(71)申请人 旭化成株式会社
地址 日本东京都
(72)发明人 井阪光二 宫武令 片山沙绫香
(74)专利代理机构 北京三友知识产权代理有限公司 11127
专利代理师 李洋 褚瑶杨
(51)Int.Cl.
*C12N 15/53*(2006.01)

权利要求书3页 说明书26页
序列表33页 附图10页

(54)发明名称

具有羧酸还原活性的重组多肽

(57)摘要

本发明提供具有良好的羧酸还原活性的重组多肽、以及使用了该重组多肽的脂肪族化合物的制造方法。一种重组多肽,其中,(a)由与序列编号1所示的氨基酸序列的序列一致性为60%以上、65%以上、70%以上、75%以上、80%以上、85%以上、88%以上、90%以上、95%以上、97%以上、98%以上或99%以上的氨基酸序列A构成;(b)上述氨基酸序列A中,与由序列编号1所示的氨基酸序列构成的多肽的底物结合部位相对应的位置的氨基酸中的至少一个被置换;(c)具有羧酸还原活性。

CN 117120616 A

1.一种重组多肽,其中,

(a)由与序列编号1所示的氨基酸序列的序列一致性为60%以上、65%以上、70%以上、75%以上、80%以上、85%以上、88%以上、90%以上、95%以上、97%以上、98%以上或99%以上的氨基酸序列A构成,

(b)所述氨基酸序列A中,与由序列编号1所示的氨基酸序列构成的多肽的底物结合部位相对应的位置的氨基酸中的至少一个被置换,

(c)具有羧酸还原活性。

2.如权利要求1所述的重组多肽,其中,以序列编号1所示的氨基酸序列为基准,所述(b)中的与由序列编号1所示的氨基酸序列构成的多肽的所述底物结合部位相对应的位置是与283位、284位、298位、303位、306位、335位、512位和926位相对应的位置。

3.如权利要求1或2所述的重组多肽,其中,(b)所述氨基酸序列A中,与由序列编号1所示的氨基酸序列构成的多肽的底物结合部位相对应的位置的氨基酸中的至少两个被置换。

4.如权利要求1~3中任一项所述的重组多肽,其中,

(b)所述氨基酸序列A中,以序列编号1所示的氨基酸序列为基准,

- 与283位和303位相对应的位置被置换,
- 与283位和335位相对应的位置被置换,
- 与303位和306位相对应的位置被置换,
- 与303位和335位相对应的位置被置换,
- 与306位和335位相对应的位置被置换,或者
- 与283位和303位相对应的位置被置换。

5.如权利要求1~4中任一项所述的重组多肽,其中,

(b)所述氨基酸序列A中,以序列编号1所示的氨基酸序列为基准,满足下述(i)~(v)中的至少一者:

(i)与283位相对应的位置的氨基酸残基被精氨酸、苏氨酸和赖氨酸中的任一者置换,

(ii)与284位相对应的位置的氨基酸残基被精氨酸、苏氨酸和缬氨酸中的任一者置换,

(iii)与303位相对应的位置的氨基酸残基被半胱氨酸和蛋氨酸中的任一者置换,

(iv)与306位相对应的位置的氨基酸残基被精氨酸和赖氨酸中的任一者置换,

(v)与335位相对应的位置的氨基酸残基被精氨酸、酪氨酸、苯丙氨酸和蛋氨酸中的任一者置换。

6.如权利要求1~5中任一项所述的重组多肽,其中,

(b)所述氨基酸序列A中,以序列编号1所示的氨基酸序列为基准,满足下述(xi)~(xviii)中的任一者:

(xi)与283位相对应的位置的氨基酸残基为精氨酸,与303位相对应的位置的氨基酸为蛋氨酸,

(xii)与283位相对应的位置的氨基酸残基为精氨酸,与303位相对应的位置的氨基酸为半胱氨酸,

(xiii)与283位相对应的位置的氨基酸残基为精氨酸,与335位相对应的位置的氨基酸为苯丙氨酸,

(xiv)与303位相对应的位置的氨基酸残基为蛋氨酸,与306位相对应的位置的氨基酸

为赖氨酸，

(xv)与303位相对应的位置的氨基酸残基为蛋氨酸，与335位相对应的位置的氨基酸为苯丙氨酸，

(xvi)与303位相对应的位置的氨基酸残基为半胱氨酸，与306位相对应的位置的氨基酸为赖氨酸，

(xvii)与303位相对应的位置的氨基酸残基为半胱氨酸，与335位相对应的位置的氨基酸为苯丙氨酸，

(xviii)与306位相对应的位置的氨基酸残基为赖氨酸，与335位相对应的位置的氨基酸为苯丙氨酸。

7.如权利要求1～6中任一项所述的重组多肽，其中，与由序列编号1所示的氨基酸序列构成的多肽的羧酸还原活性相比，所述(c)的羧酸还原活性增高。

8.如权利要求1～7中任一项所述的重组多肽，其来源于选自由分枝杆菌(Mycobacterium)属、诺卡氏菌(Nocardia)属、脉孢菌(Neurospora)属和慢反应脂肪酸菌(Segniliparus)属组成的组中的属的微生物。

9.如权利要求1～8中任一项所述的重组多肽，其来源于分枝杆菌(Mycobacterium)属的微生物。

10.如权利要求1～9中任一项所述的重组多肽，其中，所述羧酸由下述式(I)或式(II)所表示，

式(I)：$R^1-(CH_2)_{n1}-COOH$

式中，$R^1$为甲基、羧基或醛基，$n^1$为1～10的整数；

式(II)：$R^2-(CH_2)_{n2}-COOH$

式中，$R^2$为氢、羟基或氨基，$n^2$为2～11的整数。

11.如权利要求10所述的重组多肽，其中，

式(I)中，$n^1$为1～7的整数，

式(II)中，$n^2$为2～8的整数。

12.如权利要求10或11所述的重组多肽，其中，

式(I)中，$n^1$为2～4的整数，

式(II)中，$n^2$为3～5的整数。

13.一种DNA，其编码权利要求1～12中任一项所述的重组多肽。

14.一种基因重组质粒，其具有权利要求13所述的DNA。

15.一种重组微生物，其导入有权利要求14所述的DNA。

16.如权利要求14所述的重组微生物，其中，所述重组微生物是具有选自由碳原子数4～12的二羧酸组成的组中的至少一者的生产能力的微生物。

17.一种培养物和/或培养物的提取物，其是权利要求15或16所述的重组微生物的培养物和/或培养物的提取物。

18.一种目标化合物的制造方法，其包括以下步骤：

将权利要求1～12中任一项所述的重组多肽或者权利要求17所述的培养物和/或培养物的提取物与底物化合物混合，得到混合液。

19.如权利要求18所述的制造方法，其中，所述底物化合物由式(I)或式(II)所表示，所

述目标化合物为醛，

式(I)：$R^1$-$(CH_2)_{n1}$-COOH

式中，$R^1$为甲基、羧基或醛基，$n^1$为1～10的整数；

式(II)：$R^2$-$(CH_2)_{n2}$-COOH

式中，$R^2$为氢、羟基或氨基，$n^2$为2～11的整数。

20.如权利要求19所述的制造方法，其中，所述醛选自由4-氨基丁烷-1-醛、5-氨基戊烷-1-醛、6-氨基己烷-1-醛、丁烷二醛、戊烷二醛、己烷二醛、4-羟基丁烷-1-醛、5-羟基戊烷-1-醛、6-羟基己烷-1-醛和戊二醛组成的组。

21.如权利要求18所述的制造方法，其中，

所述目标化合物为醇，

该方法包括以下步骤：向所述混合液中添加具有将醛转换成醇的催化剂活性的酶，得到相应的醇。

22.如权利要求18所述的制造方法，其中，

所述目标化合物为胺，

该方法包括以下步骤：向所述混合液中添加具有将醛转换成胺的催化剂活性的酶，得到相应的胺。

# 具有羧酸还原活性的重组多肽

## 技术领域

[0001] 本发明涉及具有羧酸还原活性的重组多肽、以及使用了该多肽的脂肪族化合物的制造方法。

## 背景技术

[0002] 近年来为了应对与地球温室化相伴的气象灾害、气候变化,以与地球环境的共生和环境保全为目标的可持续创新的必要性日渐提高。其中,对于能够利用可再生原料、并且利用了生物体反应的物质生产工序即生物工艺寄予很大期待。截至目前已经开发出了各种化学合成品的发酵生产工艺。例如,对于作为聚合物原料使用的羧酸化合物、例如琥珀酸和己二酸等的制造,已经研究了使用微生物的发酵生产法(专利文献1和2)。

[0003] 羧酸化合物不仅可作为聚合物原料,而且作为可转换成醛、醇和胺的原料也是有吸引力的化合物(专利文献3～7、非专利文献1和2)。通过将这些羧酸化合物的羧基转换成醛、醇和胺等取代基,能够由生物物质来源的原料开发成各种化合物,因此需要该转换技术,但羧基处于热力学稳定的状态,在化学反应中需要较大的能量。从在发酵生产中将羧基转换成其他取代基的方面出发,尝试进行了具有羧酸还原能力的酶(以下也称为“羧酸还原酶”)的探索和应用(专利文献8～10、非专利文献3)。

[0004] 羧酸还原酶中,脓肿分枝杆菌(Mycobacterium abscessus)来源的酶对于比较广泛的底物显示出羧基还原活性,因此报告了其用于获取目标化合物的应用可能性(专利文献11、非专利文献4和5)。但是,对于包括该酶在内的羧酸还原酶整体,其反应速度普遍不高,不足以在工业上利用。尽管通过酶修饰进行了提高催化剂活性和提高酶的热稳定性的研究,但依然未得到足以用于工业用途的催化剂活性(非专利文献6)。特别是越是碳链长度较短的底物,则酶活性越趋于降低,需要进行活性的改善。

[0005] 现有技术文献

[0006] 专利文献

[0007] 专利文献1:美国专利第5573931号说明书

[0008] 专利文献2:国际公开第2012/177721号

[0009] 专利文献3:日本特表2012-525856号公报

[0010] 专利文献4:日本特表2016-538870号公报

[0011] 专利文献5:日本特表2016-501031号公报

[0012] 专利文献6:日本特表2017-533734号公报

[0013] 专利文献7:日本特表2016-533162号公报

[0014] 专利文献8:日本特开2011-254747号公报

[0015] 专利文献9:国际公开第2014/095986号

[0016] 专利文献10:日本特表2015-512645号公报

[0017] 专利文献11:美国专利申请公开第2020/0080064号说明书

[0018] 非专利文献

[0019] 非专利文献1:Yu,et.al.Direct biosynthesis of adipic acid from a synthetic pathway in recombinant Escherichia coli,Biotechinology and Bioengineering,Wiley Periodicals,Inc.,2014,vol.111,pp.2580.

[0020] 非专利文献2:Yu Zhou,et.al.Biosynthesis of adipic acid by a highly efficient induction-free systemin Escherichia Coli,Journal of Biotechnology,Elsevier,2020,8,pp.314-315

[0021] 非专利文献3:Napora-Wijata et al.Biocatalytic reduction of Carboxylic acid,Biothechnology Journal,Biotechvisions,2014,9,pp.822-843

[0022] 非专利文献4:Khusnutdinova et al.Exploring Bacterial Carboxylate Reductase for the Reduction of Bifunctional Carboxylic acids,Biothechnology Journal,Wiley-VCH,2017,12,1600751

[0023] 非专利文献5:Fedrochuk et al.Site-directed mutagenesis and stability of the carboxylic acid reductase MAB4714 from Mycobacterium abscessus,Journal of Biotechnology,Elsevier,2019,303,pp.72-29

[0024] 非专利文献6:Fedrchuk et.al.One-pot Biocatalytic Transformation of Adipic Acid to6-aminocaproic Acid and 1,6-hexamethylenediamine Using Carboxylic Acid Reductases and Transaminases.J.Am.Chem.Soc.,2020,142,2,pp.1038-1048

## 发明内容

[0025] 发明所要解决的课题

[0026] 本发明的课题在于提供具有羧酸还原活性的重组多肽。另外还提供使用了该重组多肽的脂肪族化合物的制造方法。

[0027] 用于解决课题的手段

[0028] 本发明人为了解决上述课题反复进行了深入研究,结果发现,通过在根据脓肿分枝杆菌(Mycobacterium abscessus)来源的酶(由序列编号1所示的氨基酸序列构成的多肽)的立体结构分析推定为底物结合部位的部位中导入突变,可提高该酶活性。即,本发明如下所述:

[0029] [1]一种重组多肽,其中,

[0030] (a)由与序列编号1所示的氨基酸序列的序列一致性为60%以上、70%以上、80%以上、85%以上、90%以上或95%以上的氨基酸序列A构成,

[0031] (b)上述氨基酸序列A中,与由序列编号1所示的氨基酸序列构成的多肽的底物结合部位相对应的位置的氨基酸中的至少一个被置换,

[0032] (c)具有羧酸还原活性;

[0033] [2]如[1]中所述的重组多肽,其中,以序列编号1所示的氨基酸序列为基准,上述(b)中的与由序列编号1所示的氨基酸序列构成的多肽的上述底物结合部位相对应的位置是与283位、284位、298位、303位、306位、335位、512位和926位相对应的位置;

[0034] [3]如[1]或[2]中所述的重组多肽,其中,(b)上述氨基酸序列A中,与由序列编号1所示的氨基酸序列构成的多肽的底物结合部位相对应的位置的氨基酸中的至少两个被置

换;

[0035] [4]如[1]~[3]中任一项所述的重组多肽,其中,

[0036] (b)上述氨基酸序列A中,以序列编号1所示的氨基酸序列为基准,

[0037] •与283位和303位相对应的位置被置换,

[0038] •与283位和335位相对应的位置被置换,

[0039] •与303位和306位相对应的位置被置换,

[0040] •与303位和335位相对应的位置被置换,

[0041] •与306位和335位相对应的位置被置换,或者

[0042] •与283位和303位相对应的位置被置换;

[0043] [5]如[1]~[4]中任一项所述的重组多肽,其中,

[0044] (b)上述氨基酸序列A中,以序列编号1所示的氨基酸序列为基准,满足下述(i)~(v)中的至少一者:

[0045] (i)与283位相对应的位置的氨基酸残基被精氨酸、苏氨酸和赖氨酸中的任一者置换,

[0046] (ii)与284位相对应的位置的氨基酸残基被精氨酸、苏氨酸和缬氨酸中的任一者置换,

[0047] (iii)与303位相对应的位置的氨基酸残基被半胱氨酸和蛋氨酸中的任一者置换,

[0048] (iv)与306位相对应的位置的氨基酸残基被精氨酸和赖氨酸中的任一者置换,

[0049] (v)与335位相对应的位置的氨基酸残基被精氨酸、酪氨酸、苯丙氨酸和蛋氨酸中的任一者置换;

[0050] [6]如[1]~[5]中所述的重组多肽,其中,

[0051] (b)上述氨基酸序列A中,以序列编号1所示的氨基酸序列为基准,满足下述(xi)~(xviii)中的任一者:

[0052] (xi)与283位相对应的位置的氨基酸残基为精氨酸,与303位相对应的位置的氨基酸为蛋氨酸,

[0053] (xii)与283位相对应的位置的氨基酸残基为精氨酸,与303位相对应的位置的氨基酸为半胱氨酸,

[0054] (xiii)与283位相对应的位置的氨基酸残基为精氨酸,与335位相对应的位置的氨基酸为苯丙氨酸,

[0055] (xiv)与303位相对应的位置的氨基酸残基为蛋氨酸,与306位相对应的位置的氨基酸为赖氨酸,

[0056] (xv)与303位相对应的位置的氨基酸残基为蛋氨酸,与335位相对应的位置的氨基酸为苯丙氨酸,

[0057] (xvi)与303位相对应的位置的氨基酸残基为半胱氨酸,与306位相对应的位置的氨基酸为赖氨酸,

[0058] (xvii)与303位相对应的位置的氨基酸残基为半胱氨酸,与335位相对应的位置的氨基酸为苯丙氨酸,

[0059] (xviii)与306位相对应的位置的氨基酸残基为赖氨酸,与335位相对应的位置的氨基酸为苯丙氨酸;

[0060] [7]如[1]～[6]中任一项所述的重组多肽,其中,与由序列编号1所示的氨基酸序列构成的多肽的羧酸还原活性相比,上述(c)的羧酸还原活性增高;

[0061] [8]如[1]～[7]中任一项所述的重组多肽,其来源于选自由分枝杆菌(Mycobacterium)属、诺卡氏菌(Nocardia)属、脉孢菌(Neurospora)属和慢反应脂肪酸菌(Segniliparus)属组成的组中的属的微生物;

[0062] [9]如[1]～[8]中任一项所述的重组多肽,其来源于分枝杆菌(Mycobacterium)属的微生物;

[0063] [10]如[1]～[9]中任一项所述的重组多肽,其中,上述羧酸由下述式(I)或式(II)所表示,

[0064] 式(I):$R^1-(CH_2)_{n1}-COOH$

[0065] (式中,$R^1$为甲基、羧基或醛基,$n^1$为1～10的整数),

[0066] 式(II):$R^2-(CH_2)_{n2}-COOH$

[0067] (式中,$R^2$为氢、羟基或氨基,$n^2$为2～11的整数);

[0068] [11]如[10]中所述的重组多肽,其中,

[0069] 式(I)中,$n^1$为1～7的整数,

[0070] 式(II)中,$n^2$为2～8的整数;

[0071] [12]如[10]或[11]中所述的重组多肽,其中,

[0072] 式(I)中,$n^1$为2～4的整数,

[0073] 式(II)中,$n^2$为3～5的整数;

[0074] [13]一种DNA,其编码[1]～[12]中任一项所述的重组多肽;

[0075] [14]一种基因重组质粒,其具有[13]中所述的DNA;

[0076] [15]一种重组微生物,其导入有[14]中所述的DNA;

[0077] [16]如[15]中所述的重组微生物,其中,上述重组微生物是具有选自由碳原子数4～12的二羧酸组成的组中的至少一者的生产能力的微生物;

[0078] [17][15]或[16]中所述的重组微生物的培养物和/或培养物的提取物;

[0079] [18]一种目标化合物的制造方法,其包括以下步骤:

[0080] 将[1]～[12]中任一项所述的重组多肽或者[17]中所述的培养物和/或培养物的提取物与底物化合物混合,得到混合液;

[0081] [19]如[18]中所述的制造方法,其中,上述底物化合物由式(I)或式(II)所表示,上述目标化合物为醛,

[0082] 式(I):$R^1-(CH_2)_{n1}-COOH$

[0083] (式中,$R^1$为甲基、羧基或醛基,$n^1$为1～10的整数)、

[0084] 式(II):$R^2-(CH_2)_{n2}-COOH$

[0085] (式中,$R^2$为氢、羟基或氨基,$n^2$为2～11的整数);

[0086] [20]如[19]中所述的制造方法,其中,上述醛选自由4-氨基丁烷-1-醛、5-氨基戊烷-1-醛、6-氨基己烷-1-醛、丁烷二醛、戊烷二醛、己烷二醛、4-羟基丁烷-1-醛、5-羟基戊烷-1-醛、6-羟基己烷-1-醛和戊二醛组成的组;

[0087] [21]如[18]中所述的制造方法,其中,

[0088] 上述目标化合物为醇,

[0089] 该方法包括以下步骤:向上述混合液中添加具有将醛转换成醇的催化剂活性的酶,得到相应的醇;

[0090] [22]如[18]中所述的制造方法,其中,

[0091] 上述目标化合物为胺,

[0092] 该方法包括以下步骤:向上述混合液中添加具有将醛转换成胺的催化剂活性的酶,得到相应的胺。

[0093] 发明的效果

[0094] 根据本发明,能够提供具有羧酸还原活性的重组多肽、以及使用了该重组多肽的脂肪族化合物的制造方法。

## 附图说明

[0095] 图1是示出野生型羧酸还原酶的氨基酸序列(序列编号1)的图。

[0096] 图2是示出野生型羧酸还原酶的碱基序列(序列编号2)的图。

[0097] 图3是示出修饰型羧酸还原酶1的氨基酸序列(序列编号3)的图。

[0098] 图4是示出修饰型羧酸还原酶2的氨基酸序列(序列编号4)的图。

[0099] 图5A是示出褶皱慢反应脂肪酸菌(Segniliparus rugosus)来源的羧酸还原酶的氨基酸序列(序列编号72)的图。

[0100] 图5B示出序列编号1的氨基酸序列和与序列编号1具有61%的序列一致性的褶皱慢反应脂肪酸菌(Segniliparus rugosus)来源的羧酸还原酶的氨基酸序列的比对结果的图。

[0101] 图5C是示出序列编号1的氨基酸序列和与序列编号1具有61%的序列类似性的褶皱慢反应脂肪酸菌(Segniliparus rugosus)来源的羧酸还原酶的氨基酸序列的比对结果的一部分的图。

[0102] 图6是pSK000的质粒的概要图。

[0103] 图7A是将各种修饰型羧酸还原酶针对底物(己二酸)的酶活性(以每单位时间、每单位酶量的底物消耗速度(nmol/min/μg酶)的形式表示)与野生型羧酸还原酶进行比较而得到的图。图中,“WT”是指野生型酶。

[0104] 图7B是将各种修饰型羧酸还原酶针对底物(6-氧代己酸)的酶活性(以每单位时间、每单位酶量的底物消耗速度(nmol/min/μg酶)的形式表示)与野生型羧酸还原酶进行比较而得到的图。图中,“WT”是指野生型酶。

[0105] 图7C是将各种修饰型羧酸还原酶针对底物(6-氨基己酸)的酶活性(以每单位时间、每单位酶量的底物消耗速度(nmol/min/μg酶)的形式表示)与野生型羧酸还原酶进行比较而得到的图。图中,“WT”是指野生型酶。

[0106] 图7D是将各种修饰型羧酸还原酶针对底物(6-羟基己酸)的酶活性(以每单位时间、每单位酶量的底物消耗速度(nmol/min/μg酶)的形式表示)与野生型羧酸还原酶进行比较而得到的图。图中,“WT”是指野生型酶。

## 具体实施方式

[0107] 以下对本发明的具体实施方式进行具体说明。需要说明的是,本发明并不限定于

以下的实施方式,可以在其要点的范围内进行各种变形来实施。另外,只要没有特别载明,本说明书中记载的DNA的取得、载体的制备和转化等基因操作可通过Molecular Cloning 4th Edition(Cold Spring Harbor Laboratory Press,2012)、Current Protocols in Molecular Biology(Greene Publishing Associates and Wiley-Interscience)、基因工程实验笔记(遗伝子工学実験ノート)(羊土社田村隆明)等公知的文献中记载的方法来进行。需要说明的是,只要不特别声明,本说明书中,核苷酸序列以从5'朝向3'的方向来记载。

[0108] 本发明的重组多肽是按照所具有的羧酸还原活性增高的方式进行了修饰的多肽。本说明书中,关于多肽,羧酸还原活性是指将羧酸的羧基转换成醛基的活性。此处,"羧酸"中包括包含1个或复数个羧基的化合物。

[0109] 羧酸例如由式:R-COOH(式中,R是由选自由C、H、O、N、S组成的组中的1个以上的原子构成的链状或环状有机基团)所表示。

[0110] 在一个方式中,羧酸由

[0111] 式(I):$R^1-(CH_2)_{n1}-COOH$

[0112] (式中,$R^1$为甲基、羧基或醛基,$n^1$为1~10的整数)

[0113] 所表示。式(I)中,$n^1$优选为1~7的整数,更优选为2~4的整数,进一步优选为3~4的整数。

[0114] 在另一方式中,羧酸由

[0115] 式(II):$R^2-(CH_2)_{n2}-COOH$

[0116] (式中,$R^2$为氢、羟基或氨基,$n^2$为2~11的整数)

[0117] 所表示。式(II)中,$n^2$优选为2~8的整数,更优选为3~5的整数,进一步优选为4~5的整数。

[0118] 本发明的重组多肽具有下述(a)~(c)所示的特性:

[0119] (a)由与野生型羧酸还原酶的氨基酸序列的序列一致性为60%以上的氨基酸序列A构成;

[0120] (b)上述氨基酸序列A中,与野生型羧酸还原酶的底物结合部位相对应的位置的氨基酸中的至少一个被置换;

[0121] (c)具有羧酸还原活性。

[0122] 与由序列编号1所示的氨基酸序列构成的多肽的羧酸还原活性相比,上述(c)的羧酸还原活性优选增高。因此,在优选的方式中,与由序列编号1所示的氨基酸序列构成的多肽的羧酸还原活性相比,本发明的重组多肽所具有的羧酸还原活性增高。

[0123] 本发明的重组多肽由与野生型羧酸还原酶的氨基酸序列的序列一致性为一定比例以上的氨基酸序列A构成。野生型羧酸还原酶多肽被分类在EC 1.2.1.30中。

[0124] 野生型羧酸还原酶可以不受限定地从例如选自由分枝杆菌(Mycobacterium)属、诺卡氏菌(Nocardia)属、脉孢菌(Neurospora)属和慢反应脂肪酸菌(Segniliparus)属组成的组中的微生物中取得。野生型羧酸还原酶优选可以从属于分枝杆菌(Mycobacterium)属和诺卡氏菌(Nocardia)属中的微生物中取得,更优选可以从属于分枝杆菌(Mycobacterium)属的微生物中取得,最优选可以从脓肿分枝杆菌(Mycobacterium abscessus)中取得。

[0125] "野生型羧酸还原酶"并不限定于具有序列编号1所示的氨基酸序列的多肽,也可

以是在序列编号1所示的氨基酸序列中置换、添加、插入或缺失1个或复数个氨基酸而成的氨基酸序列。另外,在该多肽的N末端和C末端中的任意一方或两方可以添加有1个或复数个氨基酸。具体地说,在“野生型羧酸还原酶”中,除了具有序列编号1所示的氨基酸序列的多肽以外,还包括具有与序列编号1所示的氨基酸序列的序列一致性为60%以上的氨基酸序列的多肽。

[0126] 本说明书中,“野生型羧酸还原酶”是具有与序列编号1所示的氨基酸序列的序列一致性为60%以上的氨基酸序列的多肽,并且是在烟酰胺腺嘌呤二核苷酸磷酸(Nicotinamide adenine dinucleotide phosphate;NADPH)和三磷酸腺苷(Adenosine triphosphate;ATP)的存在下能够以羧酸化合物作为底物进行还原反应的多肽。因此,“羧酸还原酶活性”或“羧酸还原活性”更具体地说是指在NADPH和ATP的存在下能够以羧酸化合物作为底物进行还原反应的活性。

[0127] 作为野生型羧酸还原酶的一例,将脓肿分枝杆菌(Mycobacterium abscessus)来源的野生型羧酸还原酶的氨基酸序列示于序列编号1(图1)。作为野生型羧酸还原酶的另一示例,将褶皱慢反应脂肪酸菌(Segniliparus rugosus)来源的野生型羧酸还原酶的氨基酸序列示于序列编号72(图5A)。该酶与序列编号1所示的氨基酸序列具有61%的序列一致性。

[0128] 本发明的重组多肽由与野生型羧酸还原酶的氨基酸序列的序列一致性为一定比例以上的氨基酸序列A构成。氨基酸序列A与序列编号1所示的氨基酸序列的序列一致性为60%以上、65%以上、70%以上、75%以上、80%以上、85%以上、88%以上、90%以上、95%以上、97%以上、98%以上或99%以上。

[0129] 关于氨基酸序列的“序列一致性”可通过公知的方法来确定,例如可使用在BLAST(基本局部比对搜索工具(Basic Local Alignment Search Tool)https://blast.ncbi.nlm.nih.gov/Blast.cgi)中能够使用的比对搜索工具来得到。本领域技术人员可通过参考上述网站、例如参考“帮助(Help)”部分来理解使用该程序确定序列一致性的方法。序列的比较中,使用缺省参数(缺口开放罚分11、每一残基的缺口罚分1),可以使用BLAST(注册商标)(Blastp)程序的Blast2序列函数。对于多核苷酸序列的“序列一致性”也同样地可通过公知的方法确定。

[0130] 关于可在本发明中使用的野生型羧酸还原酶的基因,可以由Genebank gene ID5967171取得序列信息,例如为序列编号2所示的多核苷酸(图2)。可以通过常见的基因工程方法由脓肿分枝杆菌(Mycobacteroides abscessus)ATCC 19977取得该基因。即,可以以基因组DNA作为模板,作为通过PCR扩增得到的基因产物来取得,用于宿主微生物内的表达。另外,也可以通过有机合成制备来取得。另外,可以利用最终翻译成同一氨基酸的替代密码子。代表性方法为置换成考虑了宿主微生物的密码子使用频率的基因序列的方法。

[0131] 在一个方式中,本发明的重组多肽具有下述(a)~(c)所示的特性:

[0132] (a)由与序列编号1所示的氨基酸序列的序列一致性为60%以上的氨基酸序列A构成;

[0133] (b)上述氨基酸序列A中,与由序列编号1所示的氨基酸序列构成的多肽的底物结合部位相对应的位置的氨基酸中的至少一个被置换;

[0134] (c)具有羧酸还原活性。

[0135] 下文中,对于野生型羧酸还原酶为由序列编号1所示的氨基酸序列构成的多肽的

情况下的本发明进行详细说明。

[0136] 氨基酸序列A中,与由序列编号1所示的氨基酸序列构成的多肽的底物结合部位相对应的位置的氨基酸中的至少一个被置换。优选在氨基酸序列A中,与由序列编号1所示的氨基酸序列构成的多肽的底物结合部位相对应的位置的氨基酸中的至少两个被置换。

[0137] 关于氨基酸序列A,“与由序列编号1所示的氨基酸序列构成的多肽的底物结合部位相对应的位置”是指,以序列编号1所示的氨基酸序列为基准,与283位、284位、298位、303位、306位、335位、512位和926位相对应的位置。因此,氨基酸序列A中,以序列编号1所示的氨基酸序列为基准,在与283位、284位、298位、303位、306位、335位、512位和926位相对应的位置的氨基酸中,1个或复数个、优选至少1个、更优选至少2个被置换。通过使氨基酸序列A具有上述置换,针对目的底物可得到良好的酶活性。优选的是,通过使氨基酸序列A具有上述置换,针对目的底物的酶活性比野生型羧酸还原酶增高。关于底物结合部位的确定方法在下文叙述。

[0138] 关于氨基酸序列A,“以序列编号1所示的氨基酸序列为基准,与X位相对应的位置的氨基酸残基(X表示1以上的整数)”这一用语包括下述两种情况:

[0139] • 以序列编号1所示的氨基酸序列为基准,与X位相对应的位置的氨基酸残基(情况1);以及

[0140] • 序列编号1所示的氨基酸序列的X位的氨基酸残基(情况2)。

[0141] 具体地说,“以序列编号1所示的氨基酸序列为基准,与283位、284位、298位、303位、306位、335位、512位和926位相对应的位置的氨基酸残基”这一用语包括下述两种情况:

[0142] • 以序列编号1所示的氨基酸序列为基准,与283位、284位、298位、303位、306位、335位、512位和926位相对应的位置的氨基酸残基;以及

[0143] • 序列编号1所示的氨基酸序列的283位、284位、298位、303位、306位、335位、512位和926位的氨基酸残基。

[0144] 本说明书中,“相对应的位置的氨基酸残基”是指,使用BLAST(https://blast.ncbi.nlm.nih.gov/Blast.cgi)序列分析软件,将作为基准的氨基酸序列(具体地说为序列编号1所示的氨基酸序列。以下也称为“基准序列”)与进行对比的氨基酸序列(具体地说为氨基酸序列A)排列整齐时,与基准序列中的特定位置相对的位置的氨基酸残基。例如,将使用BLAST序列分析软件对于将序列编号1所示的氨基酸序列与褶皱慢反应脂肪酸菌(Segniliparus rugosus)来源的羧酸还原酶的氨基酸序列(序列编号72的氨基酸序列。与序列编号1所示的氨基酸序列具有61%的序列一致性)(图5A)排列整齐时的比对结果及其一部分(61位之后的氨基酸序列)示于图5(图5B和图5C)。如图5所示,在褶皱慢反应脂肪酸菌(Segniliparus rugosus)来源的羧酸还原酶的氨基酸中,与序列编号1所示的氨基酸序列的115位(序列编号1中为谷氨酰胺残基)相对应的位置的氨基酸残基为119位的丙氨酸残基。

[0145] 将由氨基酸序列A构成的本发明的重组多肽(也称为“修饰型羧酸还原酶”)的示例示于图3(修饰型羧酸还原酶1、序列编号3)和图4(修饰型羧酸还原酶2、序列编号4)。图3所示的修饰型羧酸还原酶1(由序列编号3所示的多肽构成)与序列编号1相比包含W283R和A303M的氨基酸置换。图4所示的修饰型羧酸还原酶2(由序列编号4所示的多肽构成)与序列编号1相比包含W283R和L335F的氨基酸置换。即,修饰型羧酸还原酶1中,以序列编号1所示

的氨基酸序列为基准,与283位相对应的位置的氨基酸残基以及与303位相对应的位置的氨基酸残基分别被精氨酸和蛋氨酸置换。修饰型羧酸还原酶2中,以序列编号1所示的氨基酸序列为基准,与283位相对应的位置的氨基酸残基以及与335位相对应的位置的氨基酸残基分别被精氨酸和苯丙氨酸置换。

[0146] 在一个方式中,关于本发明的重组多肽,在氨基酸序列A中,以序列编号1所示的氨基酸序列为基准,满足下述(i)~(v)中的至少一者:

[0147] (i)与283位相对应的位置的氨基酸残基被精氨酸、苏氨酸和赖氨酸中的任一者置换,

[0148] (ii)与284位相对应的位置的氨基酸残基被精氨酸、苏氨酸和缬氨酸中的任一者置换,

[0149] (iii)与303位相对应的位置的氨基酸残基被半胱氨酸和蛋氨酸中的任一者置换,

[0150] (iv)与306位相对应的位置的氨基酸残基被精氨酸和赖氨酸中的任一者置换,

[0151] (v)与335位相对应的位置的氨基酸残基被精氨酸、酪氨酸、苯丙氨酸和蛋氨酸中的任一者置换。

[0152] 关于氨基酸序列中的突变,可以为上述置换中的一者,也可以为二者以上的组合。

[0153] 关于本发明的重组多肽,在氨基酸序列A中,

[0154] •以序列编号1所示的氨基酸序列为基准,与283位和303位相对应的位置被置换,

[0155] •以序列编号1所示的氨基酸序列为基准,与283位和335位相对应的位置被置换,

[0156] •以序列编号1所示的氨基酸序列为基准,与303位和306位相对应的位置被置换,

[0157] •以序列编号1所示的氨基酸序列为基准,与303位和335位相对应的位置被置换,

[0158] •以序列编号1所示的氨基酸序列为基准,与306位和335位相对应的位置被置换,或者

[0159] •以序列编号1所示的氨基酸序列为基准,与283位和303位相对应的位置被置换。

[0160] 在优选的一个方式中,本发明的重组多肽包含2个突变,在氨基酸序列A中,以序列编号1所示的氨基酸序列为基准,满足下述(xi)~(xviii)中的任一者:

[0161] (xi)与283位相对应的位置的氨基酸残基为精氨酸,与303位相对应的位置的氨基酸为蛋氨酸,

[0162] (xii)与283位相对应的位置的氨基酸残基为精氨酸,与303位相对应的位置的氨基酸为半胱氨酸,

[0163] (xiii)与283位相对应的位置的氨基酸残基为精氨酸,与335位相对应的位置的氨基酸为苯丙氨酸,

[0164] (xiv)与303位相对应的位置的氨基酸残基为蛋氨酸,与306位相对应的位置的氨基酸为赖氨酸,

[0165] (xv)与303位相对应的位置的氨基酸残基为蛋氨酸,与335位相对应的位置的氨基酸为苯丙氨酸,

[0166] (xvi)与303位相对应的位置的氨基酸残基为半胱氨酸,与306位相对应的位置的氨基酸为赖氨酸,

[0167] (xvii)与303位相对应的位置的氨基酸残基为半胱氨酸,与335位相对应的位置的氨基酸为苯丙氨酸,

[0168] (xviii)与306位相对应的位置的氨基酸残基为赖氨酸,与335位相对应的位置的氨基酸为苯丙氨酸。

[0169] 例如可以使用现有的蛋白质晶体结构信息(蛋白质数据库,http://www.rcsb.org/)以及计算科学的方法通过酶-底物结合结构预测来推定底物结合部位。

[0170] 在进行酶-底物结合结构预测的情况下,例如可使用以swiss-model软件(https://swissmodel.expasy.org/)为代表的蛋白质三维结构预测软件,取得基于具有类似氨基酸序列的酶的晶体结构的粗模型。通过将本模型供于分子动力学计算,能够取得动态稳定化状态的模型。

[0171] 通过将利用上述方法取得的模型与经预测的现有的酶-底物结合结构重叠,能够推定底物结合部位。蛋白质分子彼此的重叠例如可以使用分子图形工具Pymol(https://pymol.org/2/)软件来实施。

[0172] 或者可以将目标底物应用于上述模型中,供于分子动力学计算,得到酶-底物结合结构的动态稳定化模型。此处,通过量子化学计算来推定目标反应中的底物的迁移状态结构并用于酶-底物结合结构预测中,这样在更准确的结合模型推定中也是有用的。在所得到的酶-底物结合结构预测模型中,例如使用上述Pymol软件,显示出在底物分子周边的任意距离内存在的氨基酸残基,由此能够推定存在于底物分子附近的氨基酸。

[0173] 制作在像这样推定的酶氨基酸(组)中导入有点突变的综合性突变型酶文库,进行筛选,由此可以确定有助于改善酶活性的底物结合部位。

[0174] 本发明的第2方面涉及编码上述重组多肽的DNA。具体地说,是编码氨基酸序列A的DNA。编码重组多肽的DNA的一例为上述序列编号2所示的多核苷酸。

[0175] 本发明的另一方面涉及导入有编码重组多肽的上述DNA(例如序列编号2所示的多核苷酸)的重组微生物。可以通过将具有编码重组多肽的多核苷酸序列的DNA与载体DNA连结而制作重组体DNA,使用该重组体DNA对宿主微生物的菌株进行转化,由此取得该重组微生物。

[0176] 宿主微生物例如为原本具有表达选自二羧酸中的至少一者的能力的微生物。即,宿主微生物例如原本具有选自二羧酸中的至少一者的生产能力。此处,二羧酸优选为碳原子数4～12的二羧酸,更优选为碳原子数4～6的二羧酸。在一个方式中,宿主微生物更优选为原本具有表达选自由琥珀酸、戊二酸和己二酸组成的组中的至少一者的能力的微生物。

[0177] 作为宿主微生物,可以利用能够在发酵工艺中使用的各种微生物,例如选自细菌、酵母和真菌等。宿主微生物例如选自埃希氏菌属、芽孢杆菌属、棒状杆菌属、克雷伯氏菌属、梭菌属、葡糖杆菌属、发酵单胞菌属、乳杆菌属、乳球菌属、链球菌属、假单胞菌属和链霉菌属微生物。从容易进行基因重组的方面出发,宿主微生物优选为大肠杆菌(Escherichia Coli)。

[0178] 本发明的重组多肽的DNA可使用能够在宿主微生物中自体复制的载体导入至宿主微生物中,或者本发明的重组多肽的DNA可以插入到染色体中进行复制。载体例如选自由质粒、噬菌体、转座子、IS元件、噬粒、粘粒或者线状或环状的DNA等组成的组。重组多肽的DNA被插入载体中并被导入到宿主微生物中。载体优选为质粒或噬菌体。

[0179] 宿主微生物为大肠杆菌的情况下,优选的质粒例如为pLG338、pACYC184、pBR322、pUC18、pUC19、pHSG298、pHSG398、pKC30、pRep4、pHS1、pKK223-3、pDHE19.2、pHS2、pPLc236、

pMBL24、pLG200、pUR290、pIN-III113-B1、λgt11和pBdCI。宿主微生物为芽孢杆菌属等杆菌的情况下,优选的质粒例如为pUB110、pC194和pBD214。除这些以外能够使用的质粒记载于“Gene Cloning and DNAnalysis 6th edition”、Wiley-Blackwell 2016中。

[0180] 关于本发明,载体例如可如下制作:在编码基因的多核苷酸序列的上游以可工作的方式连结启动子(控制区域),并且在下游以可工作的方式连结终止子,根据情况以可工作的方式连结基因标记和/或其他控制序列,由此制作该载体。

[0181] 在重组微生物的制作中,可以选择使用适合于重组多肽的表达增强的表达机构、例如启动子和终止子。关于启动子,不论是结构表达型启动子还是诱导表达型启动子,均被定义为使RNA聚合酶与DNA结合、引发RNA合成的DNA序列。强启动子是指以高频率引发mRNA合成的启动子,也适用于本发明中。在大肠杆菌中,能够利用lac系、trp系、tac或trc系、λ噬菌体的主要操纵子和启动子区域、fd外壳蛋白的调节区域、针对糖酵解系酶(例如3-磷酸甘油酸激酶、甘油醛-3-磷酸脱氢酶)、谷氨酸脱羧酶A、丝氨酸羟甲基转移酶的启动子等。作为终止子,例如可以举出rrnBT1T2终止子和lac终止子。

[0182] 除了启动子和终止子以外能够使用的其他调节元件例如为选择标记、扩增信号和复制起点等。调节元件例如记载于“Gene Expression Technology:Methods in Enzymology 185”、Academic Press(1990)中。

[0183] 本发明中使用的宿主微生物可以是该微生物原本具有的编码醇脱氢酶的基因(下文中也称为“醇脱氢酶基因”)受到了破坏的微生物。通过使用这样的宿主微生物,可避免宿主微生物固有的醇脱氢酶的混入,抑制目的外反应的进行。

[0184] 本发明的另一方面涉及上述重组多肽、或者该重组微生物的培养物和/或培养物的提取物。

[0185] 可以使用表达了本发明的修饰型羧酸还原酶的重组微生物的菌体来制造所期望的目标化合物。例如,目标化合物为醇化合物的情况下,可以添加醛还原酶。目标化合物为氨基化合物的情况下,可以通过添加氨基转移酶、或者在宿主内共表达氨基转移酶来制造氨基化合物。即还提供了下述方法:将本发明的编码修饰型羧酸还原酶的DNA导入至宿主微生物中,对所得到的重组微生物进行培养,使用该培养物来制造目标化合物。

[0186] 因此,本发明的又一方面涉及使用了上述重组微生物的目标化合物的制造方法。具体地说,该制造方法包括将培养物和/或培养物的提取物与底物化合物混合来得到混合液的步骤。使培养物和/或培养物的提取物与底物化合物在混合液中进行反应的反应时间是可生成目的物的时间。反应时间例如为15分钟至48小时。

[0187] 作为底物化合物的羧酸例如由下述式(I)或式(II)所表示,

[0188] 式(I):$R^1-(CH_2)_{n1}-COOH$

[0189] (式中,$R^1$为甲基、羧基或醛基,$n^1$为1～10的整数),

[0190] 式(II):$R^2-(CH_2)_{n2}-COOH$

[0191] (式中,$R^2$为氢、羟基或氨基,$n^2$为2～11的整数)。

[0192] 上述式(I)中,$n^1$优选为1～7的整数,更优选为2～4的整数。上述式(II)中,$n^2$优选为2～8的整数,更优选为3～5的整数。

[0193] 目标化合物为醛的情况下,醛例如选自由4-氨基丁醛、5-氨基戊醛、6-氨基己醛、丁烷二醛、戊烷二醛、己烷二醛、4-羟基丁醛、5-羟基戊醛、6-羟基己醛、4-氧代丁酸、5-氧代

戊酸和6-氧代己酸组成的组。醛优选选自由6-氨基己醛、己烷二醛、6-羟基己醛和6-氧代己酸组成的组。

[0194] 在另一方式中,本制造方法中包括以下步骤:在混合液中添加具有将醛转换成醇的催化剂活性的酶,得到相应的醇。本方式中,目标化合物为醇。醇例如选自由4-羟基丁酸、5-羟基戊酸、6-羟基己酸、4-氨基-1-丁醇、5-氨基-1-戊醇、6-氨基-1-己醇、4-羟基丁醛、5-羟基戊醛、6-羟基己醛、1,4-丁二醇、1,5-戊二醇和1,6-己二醇组成的组,优选选自由6-羟基己酸、6-氨基-1-己醇、6-羟基己醛和1,6-己二醇组成的组。

[0195] 在又一方式中,本制造方法包括以下步骤:在混合液中添加具有将醛转换成胺的催化剂活性的酶,得到相应的胺。本方式中,目标化合物为胺。胺例如选自由4-氨基丁酸、5-氨基戊酸、6-氨基己酸、4-氨基丁醛、5-氨基戊醛、6-氨基己醛、4-氨基-1-丁醇、5-氨基-1-戊醇、6-氨基-1-己醇、1,4-二氨基丁烷、1,5-二氨基戊烷和1,6-二氨基己烷组成的组,优选选自由6-氨基己酸、6-氨基己醛、6-氨基-1-己醇和1,6-二氨基己烷组成的组。

[0196] 本发明的制造方法可以包括对上述重组微生物进行培养,使目标化合物中的任一者蓄积在培养液中,进行分离精制的步骤。

[0197] 对于在本发明的羧酸还原酶的存在下进行羧酸还原反应的条件,只要为生成目的物的条件即可,可以利用本领域技术人员通常进行的方法对反应液的组成、pH、反应温度、反应时间等进行调整、设定。例如,作为反应液,可以举出例如pH7～9的缓冲液,更优选可以举出包含pH8～9的HEPES-KOH的缓冲液。

[0198] 在羧酸还原酶活性中心处,由ATP和底物羧酸生成酰基AMP后,酰基暂时转位至磷酸泛酰巯基乙胺基。为了将磷酸泛酰巯基乙胺基供给至羧酸还原酶,向反应液中加入磷酸泛酰巯基乙胺基转移酶也是有效的。本发明中使用的磷酸泛酰巯基乙胺基转移酶多肽和编码其的多核苷酸没有特别限定,可以利用由诺卡氏菌(Nocardia)属得到的多核苷酸(登录号No.DQ904035)和其翻译产物。反应温度通常为20～40℃、更优选为30～37℃。反应时间只要为可生成目的物的时间即可,例如为15分钟至48小时。

[0199] 重组微生物的培养物和/或培养物的提取物可以以含有本发明的重组多肽的含酶物、例如酶液的形式来制备。关于酶液的制备方法,可以对于表达本发明的酶的重组微生物例如进行超声破碎、或珠磨机破碎、或者使用溶菌酶进行溶菌,使用其离心上清。

[0200] 用于进行本发明的重组微生物的培养的培养基组成、培养条件和培养时间可以通过本领域技术人员通常进行的方法适当地选择。培养温度代表性地为20～40℃、优选为30～37℃的范围内。

[0201] 培养基可以是包含1种以上的碳源、氮源、无机盐、维生素、以及必要时的微量元素或维生素等微量成分的天然、半合成或合成培养基。培养基必须恰当地满足所要培养的转化体的营养要求。具体地说,宿主微生物为好氧性的情况下,为了确保发酵中的恰当的氧浓度,需要进行振荡。这些培养条件可由本领域技术人员容易地进行设定。

[0202] 另外,在转化体表达有用的附加特性的情况下、例如具有针对抗生素的抗性标记的情况下,培养基可以包含相对应的抗生素。通过添加抗生素,可使质粒保持稳定。作为抗生素,例如可以举出氨苄青霉素、卡那霉素、氯霉素、四环素、红霉素、链霉素和奇霉素等,但并不限定于这些。

[0203] 根据本发明,能够提供具有良好的羧酸还原活性重组多肽、以及使用了该重组多

肽的脂肪族化合物的制造方法。具体地说,通过在野生型羧酸还原酶中导入突变,与该野生型酶相比,能够得到所具有的酶活性进一步提高的修饰型酶。通过使用该修饰型酶,能够有效地制造目标化合物。另外,通过根据所期望的目标化合物使用追加的酶进行转换,能够得到各种目标化合物。此外,本发明的修饰型酶中,由于酶活性提高,因此期待其能够用于工业规模的目标化合物的生产。

[0204] 接受了上述说明的本领域技术人员能够充分实施本发明。以下为了进一步进行说明而记载了实施例,但本发明并不限定于该实施例。

[0205] 实施例

[0206] 下述实施例所示的全部PCR使用PrimeSTAR Max DNAPolymerase(Takara Bio)来实施。大肠杆菌的转化使用氯化钙法(参照羊土社基因工程实验笔记(遺伝子工学実験ノート)上田村隆明著)来进行。在质粒的构建中,使用LB培养基,添加必要的抗生素来使用。只要没有特别载明,本实施例中使用的试剂使用富士胶片和光纯药公司制造的试剂。热循环仪全部使用Bio-rad公司制造的C1000Touch热循环仪。

[0207] 各实施例中使用的培养基的组成和制作方法如下所述。

[0208] (LB培养基)

[0209] 对于20g/L LB培养基Lenox(Difco公司制造),在120℃进行20分钟蒸气灭菌。根据需要加入最终浓度100mg/L的氨苄青霉素、1.5%的Bacto琼脂(Difco公司制造)。

[0210] (SOC培养基)

[0211] 首先制作1L的20g/L bacto胰蛋白胨(Difco公司制造)、5g/L bacto酵母提取物(Difco公司制造)、1M NaCl、1M KCl的溶液,在120℃进行20分钟蒸气灭菌,冷却到室温。接下来将1M $MgSO_4$、1M $MgCl_2$、2M葡萄糖各10ml通过过滤杀菌,添加到上述混合液中。

[0212] <1>羧酸还原酶活性中心的推定

[0213] 在制作修饰型酶时,按照下述方法推定被认为形成了活性中心的氨基酸残基。作为模板的序列编号1的多肽的蛋白质晶体结构尚未被报告。因此,首先使用swiss-model(https://swissmodel.expasy.org/、Swiss Institute of Bioinformatics)软件,制作序列编号1的多肽的三维结构的粗模型。此时,以作为“5mst”登记的PDB Entry作为模板。将该粗模型使用MOE(分子操作环境,Molecular Operating Environment)平台(加拿大CCG公司制造)进行氢键、范德华力等分子内相互作用的最佳化、以及基于Amber10力场和Born溶剂模型的结构最佳化,构建初期模型。进一步进行分子动力学计算,取得分子整体的稳定化模型。分子动力学计算使用作为力场参数包的Amber14和GB(Generalized Born)溶剂模型来进行。对分子的状态进行以2fs(飞秒)间隔持续50万次的计算,由此模拟1ns(纳米秒)间的分子运动。将该计算从1ns延长至2ns和3ns,由此推定动态稳定的底物-酶复合物的结构以何种方式趋同。此处,用于计算酶内部的底物的动态结构的力场参数是必要的,因此使用半经验性量子力学计算软件MOPAC对于酶反应中的底物(6-氨基己酸或者己烷二酸)的迁移状态结构进行预测。使用哈密顿算符AM1进行反应路径探索和迁移状态最佳化,进一步进行基于基准振动分析的检验。使用将该迁移状态的结构最佳化和电荷计算搭载于Gaussian的Antechamber,利用Hatree Fock 6/31G的哈密顿算符进行计算。

[0214] 结合至酶的底物配体与其周边氨基酸残基的相互作用可使用MOE,根据酶与迁移状态底物的结合数据进行结构最佳化来进行评价。此时使用结构最佳化运算法

Protonate3D进行分子内相互作用的最佳化。进一步利用MOE内的Amber99参数实施复合物整体的最佳化。另外,利用VMD(Visual Molecular Dynamics、伊利诺伊大学)软件根据酶-底物复合物结构对位于底物周边5Å以内的氨基酸残基进行鉴定。

[0215] <2>质粒的构建

[0216] 如下制作基因表达用载体。首先,利用质粒pNFP-A51(于2011年10月25日作为FERM P-22182保藏在独立行政法人制品评价技术基础机构专利生物保藏中心(IPOD)(住所:日本茨城县筑波市东1丁目1番地1中央第6)。国际保藏号:FERM BP-11515)的多克隆位点的BglII和EcoRV位点插入启动子序列,使用XbaI和HindIII位点插入终止子序列,构建pSK000。此时,通过在启动子序列末端添加EcoRV限制酶识别序列的上游3碱基对,使EcoRV限制酶位点残留在启动子和终止子序列间,在后述的酶基因克隆化时进行利用。将启动子序列和终止子序列示于表1。将所使用的引物序列示于表2。将pSK000质粒结构示于图6。所使用的限制酶由Takara公司制造。

[0217] [表1]

[0218]

| | 基因序列 | 序列编号 |
|---|---|---|
| 启动子 | CTCAAGCCCAAAGGAAGAGTGAGGC GAGTCAGTCGCGTAATGCTTAGGCAC AGGATTGATTTGTCGCAATGATTGAC ACGATTCCGCTTGACGCTGCGTAAGG TTTTTGTAATTTTACAGGCAACCTTT TATTCACTAACAAATAGCTGGTGGAA | 5 |
| 终止子 | AAAAAAAAAAAGGTGCCGTGGAAAC ACGGCACCTTTTTTTTTCTTTA | 6 |

[0219] [表2]

[0220]

| 引物编号 | 碱基序列 | 序列编号 |
|---|---|---|
| 1 | ATGACTGAAACGATCTCCACAGCGG | 7 |
| 2 | CTACACCAGGCCCAACAGCTGAATA | 8 |
| 3 | ATGATCGAGACAATTTTGCCTGCTG | 9 |
| 4 | TCAGGCGTACGCGATCGCGG | 10 |
| 5 | ATGATACGCGAGCCTCCGGA | 11 |
| 6 | TTACGCTTCTTCGACACTTACTC | 12 |
| 7 | CTTCTTCTGCCACAAGTTGGCGACC | 13 |
| 8 | TCCGAAGTGCTCTTCGTGCCGC | 14 |
| 9 | ATGGGCAGCAGCCATCACCATC | 15 |
| 10 | GGTATATCTCCTTATTAAAGTTAAAC | 16 |
| 11 | TTTAATAAGGAGATATACCATGATACGCGAGCCTC | 17 |
| 12 | GGTGATGGCTGCTGCCCATTTACGCTTCTTCGACA | 18 |
| 13 | AGCTTGCGGCCGCATAATGCTTAAG | 19 |
| 14 | ATGTATATCTCCTTCTTATACTTAACTAATATAC | 20 |
| 15 | AGAAGGAGATATACATATGACTGAAACGATCTCCA | 21 |
| 16 | GGTGGCAGCAGCCTAGCTACACCAGGCCCAACA | 22 |

[0221] (羧酸还原酶和磷酸泛酰巯基乙胺基转移酶表达质粒的制作)

[0222] 利用Eurofins Genomics株式会社的人工基因合成服务来合成编码羧酸还原酶和

磷酸泛酰巯基乙胺基转移酶的多核苷酸。以合成基因序列作为模板进行各酶的PCR扩增。关于编码羧酸还原酶的基因序列,使用表2中记载的引物1和2,在98℃(10sec)、55℃(5sec)、72℃(20sec)、30次循环的反应条件下进行扩增。关于编码磷酸泛酰巯基乙胺基转移酶的基因序列,使用引物3和4,在98℃(10sec)、55℃(5sec)、72℃(10sec)、30次循环的反应条件下进行扩增。

[0223] 接下来,将所得到的DNA片段利用Mighty Cloning Reagent Set(Takara公司制造)进行磷酸化,用作克隆的DNA片段。另行将pSK000质粒利用EcoRV限制酶进行处理,接下来利用碱性磷酸酯酶(BAP,Takara公司制造)进行处理。将上述DNA片段和载体片段使用Mighty Cloning Reagent Set(Takara公司制造)进行连接(16℃、一夜)。利用1μL的连接反应液转化大肠杆菌JM109株。由出现的菌落提取质粒,进行序列分析,制作出按照表达酶蛋白的方式插入了基因的质粒。

[0224] (氨基转移酶表达质粒的制作)

[0225] 利用Eurofins Genomics株式会社的人工基因合成服务合成出氨基转移酶的多核苷酸。由genebank gene ID7435770取得基因序列。以合成基因序列作为模板,使用表2中记载的引物5和6,在98℃(10sec)、55℃(5sec)、72℃(10sec)、30次循环的反应条件下进行PCR。将所得到的DNA片段利用上述方法连接至pSK000载体的启动子下游(16℃、一夜)。利用1μL连接反应液转化大肠杆菌JM109株。由出现的菌落提取质粒进行序列分析,取得了按照表达酶蛋白的方式插入有基因的质粒。

[0226] (羧酸氧化还原酶的点突变导入酶表达质粒文库的制作)

[0227] 接着进行羧酸还原酶基因序列的点突变导入综合性酶表达文库的制作。首先,利用Eurofins Genomics株式会社的人工基因片段合成服务取得5种将与所期望的氨基酸部位相对应的3个碱基对作为“NNK”的200个碱基对的基因片段。将序列示于表3。

[0228] [表3]

[0229]

| 突变导入部位 | 综合性点突变导入中使用的合成片段的碱基序列 | 序列编号 |
|---|---|---|
| Trp 283 | CTTGTGGCAGAAGAAGNNKCTCACCGACGATGTGATTCCGTCC ATAGGCGTGAACTTCATGCCCATGAGCCACCTGGCGGGTCGCC TCACTCTCATGGGCACCCTTTCCGGTGGCGGAACCGCCTACTAC ATCGCTTCGAGCGATCTTTCGACTTTCTTCGAGGACATCGCGCT CATCCGCCCCTCCGAAGTGCTCTTCG | 23 |
| Leu 284 | CTTGTGGCAGAAGAAGTGGNNKACCGACGATGTGATTCCGTCC ATAGGCGTGAACTTCATGCCCATGAGCCACCTGGCGGGTCGCC TCACTCTCATGGGCACCCTTTCCGGTGGCGGAACCGCCTACTAC ATCGCTTCGAGCGATCTTTCGACTTTCTTCGAGGACATCGCGCT CATCCGCCCCTCCGAAGTGCTCTTCG | 24 |
| Ala 303 | CTTGTGGCAGAAGAAGTGGCTCACCGACGATGTGATTCCGTCCA TAGGCGTGAACTTCATGCCCATGAGCCACCTGNNKGGTCGCCT CACTCTCATGGGCACCCTTTCCGGTGGCGGAACCGCCTACTACA TCGCTTCGAGCGATCTTTCGACTTTCTTCGAGGACATCGCGCTC ATCCGCCCCTCCGAAGTGCTCTTCG | 25 |
| Leu 306 | CTTGTGGCAGAAGAAGTGGCTCACCGACGATGTGATTCCGTCCA TAGGCGTGAACTTCATGCCCATGAGCCACCTGGCGGGTCGCNN KACTCTCATGGGCACCCTTTCCGGTGGCGGAACCGCCTACTACA TCGCTTCGAGCGATCTTTCGACTTTCTTCGAGGACATCGCGCTC ATCCGCCCCTCCGAAGTGCTCTTCG | 26 |
| Leu 335 | CTTGTGGCAGAAGAAGTGGCTCACCGACGATGTGATTCCGTCCA TAGGCGTGAACTTCATGCCCATGAGCCACCTGGCGGGTCGCCT CACTCTCATGGGCACCCTTTCCGGTGGCGGAACCGCCTACTACA TCGCTTCGAGCGATCTTTCGACTTTCTTCGAGGACATCGCGNNK ATCCGCCCCTCCGAAGTGCTCTTCG | 27 |

[0230] 接着,以上述羧酸还原酶表达质粒10ng作为模板,对用于克隆点突变导入酶基因片段的载体片段进行扩增。使用表2中记载的引物7和8,在98℃(10sec)、55℃(5sec)、72℃(15sec)、30次循环的反应条件下进行PCR。将所得到的载体片段使用Takara公司的In-Fusion(注册商标)HD Cloning Kit与200个碱基对的基因片段连结,制作点突变导入酶表达质粒文库。

[0231] <3>点突变导入酶表达质粒文库的筛选

[0232] 作为修饰型羧酸还原酶的一次筛选,进行了导入有上述[2]中制作的点突变的综合性羧酸还原酶文库的筛选。首先,利用点突变导入酶表达质粒文库,转化大肠杆菌JM109株,利用琼脂培养基在30℃进行1天培养。同样地为了进行比较,使用野生型酶表达质粒或者不具有酶基因的空质粒,对大肠杆菌JM109株进行转化,利用琼脂培养基在30℃进行1天培养。将所出现的菌落接种至LB培养基中(深孔板、1mL/孔),在30℃振荡培养一夜(M·BR-1212FP微孔板振荡器、TAITEC公司制造)。将各30μL的培养液转移到微孔板中,进行10倍稀释,利用Tecan Infinite 200酶标仪测定OD600值,进行记录(OD600)。利用Thremo Fisher公司制造的Sorvall ST-8FL离心机(安装有微孔板转子)将培养后的深孔板以2,456xg离心10分钟。弃掉上清,将团块用100mM Tris-HCl缓冲液(pH=8.0)50μL混悬后,加入溶菌酶溶液150μL(总量200μL、最终浓度0.65mg/mL溶菌酶、1mM EDTA)。使用TAITEC公司制造的Deepwell Maximizer台式振荡机在30℃以400rpm培养30分钟,以2546xg离心10分钟。将上清50μL转移至96孔微孔板中,添加以30℃进行了预热的底物溶液(总量250μL、最终浓度0.2mM NADPH、0.5mM ATP、10mM己二酸二钠、10mM $MgCl_2$、100mM HEPES钾缓冲液pH=8.0)。使用Tecan Infinite 200酶标仪,对于340nm的吸光度连续测定15分钟。计算出每单位时间

的340nm吸光度减少量,将所计算出的值除以上述OD600值。在野生型酶与突变型酶之间对于所得到的值进行比较,结果得到了复数个显示出高于野生型酶的活性的突变体。将340nm吸光度的减少速度高于野生型的孔的传代菌落利用LB培养基进行培养,提取质粒,进行序列分析。对于由筛选得到的各修饰型酶,将突变点、以及相对于野生型羧酸还原酶的活性的相对酶活性示于表4。

[0233] [表4]

[0234] 表4:各修饰型酶的比活性

[0235]

| | | 设野生型为1时的各突变体的比活性 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 导入至突变体中的氨基酸 | | | | | | | |
| | | Lys | Arg | Thr | Val | Met | Phe | Tyr | Cys |
| 酶中的突变部位 | Trp 283 | 1.22 | 1.26 | 1.05 | | | | | |
| | Leu 284 | | 1.08 | 1.05 | 1.24 | | | | |
| | Ala 303 | | | | | 1.51 | | | 1.36 |
| | Leu 306 | 1.14 | 1.12 | | | | | | |
| | Leu 335 | | 1.27 | | | 1.22 | 1.20 | 1.25 | |

[0236] <4>双突变导入羧酸还原酶的构建以及针对各种底物的活性评价

[0237] (双突变导入羧酸氧化还原酶表达质粒的制作)

[0238] 基于由上述筛选得到的提高羧酸还原酶活性的氨基酸突变信息,制作将2个突变点组合的突变体表达质粒,进行酶活性比较。在羧酸还原酶点突变基因序列中第二点突变的导入使用InversePCR法在98℃(10sec)、55℃(5sec)、72℃(10sec)、30次循环的反应条件下进行。将所得到的DNA片段利用Mighty Cloning Reagent Set(Takara公司制造)进行磷酸化,进行自连接(16℃、20分钟)。利用1μL连接反应液转化大肠杆菌JM109株。由出现的菌落提取质粒,进行序列分析,确认到导入了目标突变。

[0239] (6-氧代己酸的合成)

[0240] 参照文献(Turk et.al.Metabolic Engineering toward Sustainable Production of Nylon-,ACS Synthetic Biology,ACS Publications,2016,vol.5,pp 65-73.)来合成6-氧代己酸。将Toronto Research Chem.公司制造的6-氧代己酸甲酯1g添加到水中,制成10%w/w%。一边利用磁力搅拌进行搅拌,一边加入8N的NaOH,调整为pH=14。在该状态下在常温下搅拌24小时后,加入5N的HCl,中和至pH=7。将每100μL分注到1.5mL管中,使用东京理化公司制造的CVE-3110离心蒸发器浓缩2小时。在-80℃冷冻3小时后,利用真空干燥机干燥一夜。

[0241] (双突变导入羧酸氧化还原酶的活性评价)

[0242] 利用双突变导入酶表达质粒转化大肠杆菌JM109株。同样地,为了进行比较,使用野生型酶表达质粒或者不具有酶基因的空质粒中的任一者转化大肠杆菌JM109株,利用琼脂培养基在30℃进行1天培养。将出现的菌落接种在LB培养基中(15m管、液量2mL),在30℃振荡培养一夜。将残留的培养液转移到2mL管中,利用Eppendorf公司制造的

Centrifuge5424R台式微量离心机以8,000rpm离心3分钟。弃掉上清,将团块利用100mM Tris-HCl缓冲液(pH＝8.0)200μL混悬后,利用超声波破碎机将菌体破碎。关于超声波破碎,在冰上以存储3的输出使用30秒振荡、30秒休止、30秒振荡的方案来进行(Tomy公司制造UR-21P)。将破碎液以13,200rpm离心15分钟,将该上清用作酶液。将酶液用100mM Tris-HCl缓冲液(pH＝8.0)进行2倍稀释,将50μL转移至96孔微孔板中,添加以30℃进行了预热的底物溶液(总量250μL、最终浓度0.2mM NADPH、0.5mM ATP、10mM $MgCl_2$、100mM HEPES钾缓冲液pH＝8.0)。此时,在底物中使用了1mM己二酸二钠、1mM 6-氧代己酸、10mM 6-氨基己酸、0.1mM 6-羟基己酸中的任一者。6-羟基己酸由ACCUSTANDARD公司购入。使用Tecan Infinite 200酶标仪,对于340nm的吸光度连续测定15分钟。对于反应后的分析液采样5μL,使用Bio-rad公司制造的蛋白质分析浓缩色素试剂,基于Bradford法计算出全蛋白浓度。

[0243] 另行分取5μL的酶液作为SDS-PAGE用样品,与Bio-rad公司制造的2xlaemmuli样品缓冲液5μL(5％2-巯基-1,3-丙二醇)进行混合。在100℃加热20分钟后,使用Bio-rad公司制造的Mini-PROTEANTGX凝胶(Any kD)和Mini-PROTEAN Tetra Cell SDS-PAGE电泳槽,以200V进行30分钟电泳。将电泳后的凝胶利用Bio-Safe公司制造的Coomassie-STAIN染色液进行30分钟染色,利用蒸馏水进行2小时脱色。将脱色后的凝胶利用Bio-rad公司制造的Gel DocEZ系统进行拍摄,使用Image lab软件计算出各泳道的全蛋白条带中的羧酸还原酶条带比例。将羧酸还原酶条带比例乘以全蛋白浓度,将其作为分析液中的羧酸还原酶浓度。进一步由340nm的吸光度变化计算出NADPH的消耗速度,计算出每单位酶蛋白的NADPH消耗速度。值全部为独立的2个培养样品的平均值。将所评价的双突变导入羧酸还原酶、以及针对各底物的活性示于图7A、图7B、图7C和图7D。在第283位的氨基酸突变成精氨酸、第303位的氨基酸突变成蛋氨酸的双突变体中,对任一底物均显示出了增高的活性。

[0244] <5>双突变导入羧酸还原酶的酶学性质的变化

[0245] 对双突变体酶和野生型酶的酶学性质进行比较。使用野生型酶表达质粒、序列编号3所示的突变型酶表达质粒、不具有酶基因的空质粒中的任一者,对大肠杆菌JM109株进行转化。将出现的菌落接种在LB培养基中(15mL管、液量2mL),在30℃振荡培养一夜。将各30μL的培养液稀释10倍,记录OD600值。将余下的培养液转移至2mL管中,利用Eppendorf公司制造的Centrifuge5424R台式微量离心机在8,000rpm离心3分钟。弃掉上清,将团块利用100mM Tris-HCl缓冲液(pH＝8.0)按照OD600＝30进行混悬后,利用超声波破碎机破碎菌体。将破碎液以13,200rpm离心15分钟,将其上清用作酶液。将酶液用100mM Tris-HCl缓冲液(pH＝8.0)进行3倍稀释,将50μL转移至96孔微孔板中,添加在30℃预加热的底物溶液(总量250μL、最终浓度0.2mM NADPH、0.5mM ATP、10mM $MgCl_2$、100mM HEPES钾缓冲液pH＝8.0)。在野生型酶的实验中,以50mM、25mM、12.5mM、6.25mM和3.125mM中的任一浓度使用己二酸二钠。在突变型酶的实验中,以3mM、1.5mM、0.75mM、0.375mM和0.1875mM中的任一浓度使用己二酸二钠。使用Tecan Infinite 200酶标仪,对340nm的吸光度进行60分钟持续测定。对反应后的分析液采样5μL,使用Bio-rad公司制造的蛋白质分析浓缩色素试剂,基于Bradford法计算出全蛋白浓度。

[0246] 另行分取5μL的酶液作为SDS-PAGE用样品,与Bio-rad公司制造的2xlaemmuli样品缓冲液5μL(5％2-巯基-1,3-丙二醇)进行混合。在100℃加热20分钟后,使用Bio-rad公司制造的Mini-PROTEANTGX凝胶(Any kD)和Mini-PROTEAN Tetra Cell SDS-PAGE电泳槽,以

200V进行30分钟电泳。将电泳后的凝胶利用Bio-Safe公司制造的Coomassie-STAIN染色液染色30分钟,利用蒸馏水进行2小时脱色。将脱色后的凝胶利用Bio-rad公司制造的Gel DocEZ系统进行拍摄,使用Image lab软件计算出各泳道的全蛋白条带中的羧酸还原酶条带比例。将羧酸还原酶条带比例乘以全蛋白浓度,将其作为分析液中的羧酸还原酶浓度。进一步由340nm的吸光度变化计算出NADPH的消耗速度,计算出每单位酶蛋白的NADPH消耗速度。将底物浓度的倒数和NADPH消耗速度的倒数作图,根据与x轴的交点以及与y轴的交点分别计算出Km和Vmax。将各数值列于下表。数值为独立的2次研究的平均值。突变型酶针对己二酸的Km为野生型酶的1/20以下,与现有的突变型羧酸还原酶、即L342突变型和G418E突变型(Fedrchuk et.al.One-pot Biocatalytic Transformation of Adipic Acid to6-aminocaproic Acid and 1,6-hexamethylenediamine Using Carboxylic Acid Reductases and Transaminases.J.Am.Chem.Soc.,2020,142,2,pp.1038-1048(非专利文献6))相比为1/17,与野生型酶相比得到了改善。

[0247] [表5]

[0248] 表5:野生型羧酸还原酶、修饰型酶(序列编号3)以及现有的修饰型酶的酶学性质的比较

[0249]

| | Km (mmol/L) | $V_{max}$ (umol/L/min) | Kcat ($sec^{-1}$) |
|---|---|---|---|
| 比较例 (野生型、序列编号1) | 23.3 | 18.1 | 1.68 |
| 实施例 (W283R/A303M突变型、序列编号3) | 1.06 | 24.8 | 3.53 |
| 比较例 (L342E突变型、文献值) | 56.6+/-5.0 | 未记载 | 2.72+/-0.09 |
| 比较例 (G418E突变型、文献值) | 18.7+/-2.5 | 未记载 | 4.25+-/0.26 |

[0250] <6>利用了双突变导入羧酸还原酶的化学品的合成

[0251] 在使用大肠杆菌制作酶液时,为了避免大肠杆菌固有的醇脱氢酶的混入,表达酶的大肠杆菌株使用了醇脱氢酶ADH基因破坏株。以下示出ADH基因破坏质粒的制作方法。

[0252] (ADH基因破坏质粒的制作)

[0253] 关于ADH基因的破坏,使用pHAK1(于2019年3月18日作为NITE P-02919保藏在独立行政法人制品评价技术基础机构生物技术中心专利微生物保藏中心(NPMD)(地址:千叶县木更津市かずさ镰足2-5-8 122号室)。国际保藏号:NITE BP-02919)通过同源重组法进行。pHAK1包含温度敏感性突变型repA基因、卡那霉素抗性基因、枯草杆菌(Bacillus subtilis)来源的左聚糖蔗糖酶基因SacB。左聚糖蔗糖酶基因在蔗糖存在下对宿主微生物产生致死作用。PCR片段的扩增中使用PrimeSTAR Max DNAPolymerase(产品名、Takara Bio制)进行,质粒制备使用大肠杆菌HST08株进行。以大肠杆菌BL21(DE3)株的基因组DNA作为模板,得到包含破坏靶基因的上游区域、编码区域、以及下游区域的PCR产物。将靶基因与引物序列的组合示于表6。

[0254] [表6]

[0255]

| 靶基因 | 正向引物 | | | 反向引物 | | |
|---|---|---|---|---|---|---|
| | 引物编号 | 碱基序列 | 序列编号 | 引物编号 | 碱基序列 | 序列编号 |
| yqhD | 17 | TTGTTGAATAAATCGTCTGTTTGCCGAGAATACGCG | 28 | 18 | TATCGCATGCAGATCTGCGCTAATGGGCTCCAGGA | 29 |
| fucO | 19 | TTGTTGAATAAATCGCGCGCCAGCGCTGGCTGTTT | 30 | 20 | TATCGCATGCAGATCCCGCCAATGCCGGAAGAGTT | 31 |
| adhP | 21 | TTGTTGAATAAATCGCGATCGTGATGCCGCTGTCT | 32 | 22 | TATCGCATGCAGATCGACAACGTAGGCTTTGTTCA | 33 |
| eutG | 23 | TTGTTGAATAAATCGCGCCATCTCGACACTCTTGA | 34 | 24 | TATCGCATGCAGATCGCGAACATCGATGGGTTAGC | 35 |
| ybbO | 25 | TTGTTGAATAAATCGTGGCATTTGCCCTTCCTGTT | 36 | 26 | TATCGCATGCAGATCGTCCTGATCCTGCAACGGAA | 37 |
| ahr | 27 | TTGTTGAATAAATCGCTCATAACGGTACTGCAAAC | 38 | 28 | TATCGCATGCAGATCTCGCAGCAGGTAAGATGATT | 39 |
| yahK | 29 | TTGTTGAATAAATCGATATTCGTCCTAACGAACAG | 40 | 30 | TATCGCATGCAGATCTTTTTGATTTTCAAGTATGT | 41 |

[0256] 接着,使用In-Fusion HD cloning kit(产品名、Clontech公司制造),将该PCR产物插入到使用引物31和32扩增得到的pHAK1质粒片段中,进行环状化。将引物31和32的序列示于表7。

[0257] [表7]

[0258]

| 引物编号 | 碱基序列 | 序列编号 |
|---|---|---|
| 31 | GATCTGCATGCGATATCTCTAGAACGCGTAAGCTT | 42 |
| 32 | TCTCGAGCCGATTTATTCAACAAAGCCGC | 43 |

[0259] 将所得到的插入有破坏靶基因的上游区域、编码区域、以及下游区域的DNA片段的pHAK1质粒作为模板,使用表8中记载的引物进行PCR,得到除去了破坏靶基因的编码区域的一部分区域或全部区域的质粒片段。

[0260] [表8]

[0261]

| 靶基因 | 正向引物 | | | 反向引物 | | |
|---|---|---|---|---|---|---|
| | 引物编号 | 碱基序列 | 序列编号 | 引物编号 | 碱基序列 | 序列编号 |
| yqhD | 33 | ACTTTCGTTTTCGGGCATTTCGTCC | 44 | 34 | CCAATATGAGGGCAGAGAACGATC | 45 |
| fucO | 35 | ATGCGCTGATGTGATAATGC | 46 | 36 | CCTTCTCCTTGTTGCTTTA | 47 |
| adhP | 37 | GAGGCCTTTGCTGCGACTGC | 48 | 38 | AGTTCCTCCTTTTCGGATGAT | 49 |
| eutG | 39 | ATGCCGGATGCGACGCTT | 50 | 40 | TCATTTTGCATATAGCCCCT | 51 |
| ybbO | 41 | TGACCTGGGCAGTAATGGTG | 52 | 42 | CAGGATCTCCGTTGCTTTATGAGTC | 53 |
| ahr | 43 | CGTGGTGTTGAAAGCCGATTATTG | 54 | 44 | CATAAACTTCCAGTTCTCCGCCC | 55 |
| yahK | 45 | TCGCACACTAACAGACTGAA | 56 | 46 | TGTGTTTACTCCTGATTAGC | 57 |

[0262] 对所得到的质粒片段通过末端磷酸化、自连接而进行环状化,得到基因破坏用质粒。

[0263] (ADH基因破坏大肠杆菌株的构建)

[0264] 通过氯化钙法在大肠杆菌BL21(DE3)株中转化用于破坏所期望的基因的质粒后,涂布至含有卡那霉素硫酸盐100mg/L的LB琼脂培养基,在30℃培养一夜,得到转化体。将一白金环的该转化体种到含有卡那霉素硫酸盐100mg/L的LB培养基1mL中,在30℃进行振荡培养。将所得到的培养液涂布至含有卡那霉素硫酸盐100mg/L的LB琼脂培养基,在42℃培养一夜。对于所得到的菌落,通过单交换将质粒插入到基因组中。将一白金环的菌落接种到LB液体培养基1mL中,在30℃进行振荡培养。将所得到的培养液涂布至含有蔗糖10%的LB琼脂培养基,培养一夜。对于所得到的菌落,使用表9所示的引物集通过菌落直接PCR确认所期望的基因被破坏的情况。根据以上的操作构建出ADH7基因被破坏的ADH缺损大肠杆菌。

[0265] [表9]

[0266]

| 靶基因 | 正向引物 | | | 反向引物 | | |
|---|---|---|---|---|---|---|
| | 引物编号 | 碱基序列 | 序列编号 | 引物编号 | 碱基序列 | 序列编号 |
| yqhD | 47 | TATTCTCAATCCGTTTCAGCACGCG | 58 | 48 | TTCGGGATCACCACCAGGCCG | 59 |
| fucO | 49 | CGGAAATGGACGAACAGTGG | 60 | 50 | CGTCATCAGCGTTTACCAGATT | 61 |
| adhP | 51 | GCATAAACACTGTCCGCGTC | 62 | 52 | GAAATCGAGAAGGCAGAAGCGAAA | 63 |
| eutG | 53 | GGTGCGGTCACCATTGTTCG | 64 | 54 | CATATCGCACGCCAGCAGTG | 65 |
| ybbO | 55 | GGTGAGGATGGAGAGTTCATG | 66 | 56 | CAGTTCGATTTGCGCCACCAG | 67 |
| ahr | 57 | TCCGCTAGTGTGATTTCAGG | 68 | 58 | GAAATTATTATGCCGCCAGGCGT | 69 |
| yahK | 59 | TTATGGTCTGGGCGACATGC | 70 | 60 | GCATCATCCTGGTCATATACCC | 71 |

[0267] (酶的制备)

[0268] 使用野生型羧酸还原酶表达质粒、序列编号3所示的修饰型酶表达质粒、氨基转移酶表达质粒、不具有酶基因的空质粒中的任一者,对上述ADH缺损大肠杆菌株进行转化。

[0269] 将菌落接种在LB培养基中(15m管、液量2mL),在30℃振荡培养一夜。将各30μL的培养液稀释10倍,利用Tecan Infinite 200酶标仪测定OD600值,进行记录。将余下的培养液转移到2mL管中,利用Eppendorf公司制造的Centrifuge5424R台式微量离心机以8,000rpm离心3分钟。弃掉上清,将团块利用100mM Tris-HCl缓冲液(pH=8.0)按照OD600=30再混悬后,利用超声波破碎机将菌体破碎。将破碎液以13,200rpm离心15分钟,将其上清用作酶液。酶反应在1.5mL管内进行。在反应液中分别各加入50μL的含有羧酸还原酶和氨基转移酶的酶液。反应液的组成如下。

[0270] 反应条件1;最终浓度2mM的己二酸二钠、5mM的谷氨酸钠、5mM NADPH、5mM ATP、10mM $MgCl_2$、0.05mM磷酸吡哆醛、100mM HEPES钾缓冲液pH=8.0、总量1000μL。

[0271] 反应条件2;最终浓度10mM的6-氨基己酸、20mM的谷氨酸钠、20mM NADPH、20mM ATP、10mM $MgCl_2$、0.05mM磷酸吡哆醛、100mM HEPES钾缓冲液pH=8.0、总量1000μL。

[0272] 将所制备的管利用TAITEC公司制造的台式振荡器(M·BR-022UP)在30℃以1000rpm振荡4小时。将反应后的样品用0.5%甲酸进行5倍稀释,利用表10所示的分析方法进行分析。

[0273] [表10]

[0274]

| 羧酸、醇的分析 | |
|---|---|
| 柱 | Shim-Pack SCR(H) +Shim-pack SCR-101H 7.9x300(岛津公司制造) |
| 柱温度 | 40℃ |
| 洗脱液 | 磷酸水溶液 (pH2.3) 0.1wt% |
| 流速 | 1mL/min |
| 检测器 | RI + UV 210nm |
| 注入量 | 50uL |
| | |
| 氨基醇的分析 | |
| 柱 | Shodex IC YS-G + YS-50 (4.6mm x 125mm) (昭和电工公司制造) |
| 柱温度 | 40℃ |
| 洗脱液 | 6mM MSA |
| 流速 | 1mL/min |
| 检测器 | RI |
| 注入量 | 10uL |
| | |
| 胺的分析 | |
| 柱 | Dionex IonPac CG19+CS19 (2mm) (Thermo Fisher公司制造) |
| 柱温度 | 30℃ |
| 洗脱液 | 9mM MSA→70mM MSA (10min)<br>→保持1min |
| 流速 | 0.35mL/min |
| 检测器 | Suppressed Conductivity (CSRS 300) |
| 注入量 | 5uL |

[0275] 表11-1和表11-2中示出了将底物化合物利用羧酸还原酶和氨基转移酶进行了酶处理的情况下所生成的六亚甲基二胺HMD的浓度(μM)的分析结果。各值为独立的2次分析的平均值。表11-1和表11-2中,所使用的底物分别为己二酸(表11-1)和6-氨基己酸(表11-2)。表中,“CAR”表示羧酸还原酶,“AT”表示氨基转移酶、“空载体”表示利用了不具有酶基因序列的空载体进行了转化的菌体的破碎物。此处使用的“突变型CAR”是由序列编号3所示的氨基酸序列构成的修饰型酶。

[0276] [表11-1]

[0277]

| 反应条件1样品 | HMD(μM) |
|---|---|
| AT+野生型CAR(比较例) | 8.63 |
| AT+突变型CAR(实施例) | 403 |
| 仅AT(比较例) | 0.959 |
| 仅野生型CAR(比较例) | N.D. |
| 仅突变型CAR(比较例) | N.D. |
| 空载体(比较例) | N.D. |

[0278] 如表11-1所示,在既未添加羧酸还原酶也未添加氨基转移酶的样品、以及仅羧酸还原酶的样品中,未观察到从己二酸生成HMD的情况。在共同添加了野生型的羧酸还原酶和氨基转移酶的情况下,确认到微量的HMD生成,与其相比,在使用突变型的羧酸还原酶和氨基转移酶的情况下,确认到了其50倍的HMD的生成。可知通过使用该突变型酶,即使不存在辅酶的再生系统也能够有助于HMD生成。

[0279] [表11-2]

[0280]

| 反应条件2样品 | HMD(μM) |
|---|---|
| AT+野生型CAR(比较例) | 16.1 |

| | |
|---|---|
| AT+突变型CAR(实施例) | 81.4 |
| 仅AT(比较例) | 4.96 |
| 仅野生型CAR(比较例) | N.D. |
| 仅突变型CAR(比较例) | 19.7 |
| 空载体(比较例) | N.D. |

[0281] 如表11-2所示,在既未添加羧酸还原酶也未添加氨基转移酶的样品、以及仅野生型羧酸还原酶的样品中,未观察到从6-氨基己酸生成HMD的情况。在共同添加了野生型的羧酸还原酶和氨基转移酶的情况下,确认到HMD生成。有趣的是,在仅添加了突变型羧酸还原酶的样品中,确认到比共同添加了野生型的羧酸还原酶和氨基转移酶的情况更有效地生成了HMD的情况。关于这一点,可以推定出,通过突变型羧酸还原酶可更有效地对6-氨基己酸进行还原,通过大肠杆菌内源性氨基转移酶可进一步转换成HMD。并且在共同添加了突变型的羧酸还原酶和氨基转移酶的样品中,确认到了最高的HMD浓度。

[0282] <7>利用了双突变导入羧酸还原酶表达菌的化学品的合成

[0283] (羧酸还原酶、氨基转移酶共表达用质粒的制作)

[0284] 共表达用质粒的主干使用pACYCDuet(默克公司制造)。首先,为了在多克隆位点1插入ygjg,以pACYCDuet作为模板,使用引物编号9,10进行PCR,制备载体片段。接下来,以克隆有氨基转移酶的质粒作为模板、使用引物编号11,12进行PCR,得到ygjg基因片段。将所得到的基因片段使用Takara公司的In-Fusion(注册商标)HD Cloning Kit进行连接。接下来,为了在多克隆位点2插入羧酸还原酶,以通过之前的操作得到的含有ygjg的pACYC质粒作为模板、使用引物编号13,14进行PCR,得到构成载体的片段。PCR反应条件使用98℃(10sec)、55℃(5sec)、72℃(30sec)、30次循环。

[0285] 接着,以克隆有野生型羧酸还原酶或双突变型羧酸还原酶的质粒作为模板,使用引物编号15,16进行PCR。PCR反应条件使用98℃(10sec)、55℃(5sec)、72℃(20sec)、30次循环。将载体片段与包含野生型或突变型羧酸还原酶的PCR片段使用Takara公司的In-Fusion(注册商标)HD Cloning Kit进行连接,制作2基因共表达质粒。

[0286] (培养评价)

[0287] 作为表达酶的大肠杆菌株,使用上述<6>中制作的ADH破坏株。利用2基因共表达质粒、或者仅克隆有ygjg的pACYC、或者pACYC质粒对上述株进行转化。将菌落接种在LB培养基中(15m管、液量1mL),在37℃振荡培养3小时。接下来,在包含己二酸10mM的SOC培养基中接种5%,在30℃培养一夜(15ml管、液量2ml)。将各30μL的培养液稀释10倍,使用Tecan Infinite 200酶标仪测定OD600值,进行记录。将余下的培养液转移至2mL管中,利用Eppendorf公司制造的Centrifuge5424R台式微量离心机以8,000rpm离心3分钟。将上清用0.5%甲酸进行5倍稀释,利用上述表10所示的分析方法进行分析。

[0288] 表12中示出了利用含有己二酸的培养基对表达野生型或修饰型的羧酸还原酶、以及氨基转移酶的菌株进行一定时间培养后,在上清中检测出的HMD的浓度(μM)。表中,“修饰型CAR1”是由序列编号3所示的氨基酸序列构成的重组多肽(修饰型羧酸还原酶1),“修饰型CAR2”是由序列编号4所示的氨基酸序列构成的重组多肽(修饰型羧酸还原酶2)。各值为独立的2次分析的平均值。

[0289] [表12]

[0290]

| 所导入的羧酸还原酶 | HMD(μM) |
|---|---|
| 突变型CAR1(实施例) | 116.1 |
| 突变型CAR2(实施例) | 104.4 |
| 野生型(比较例) | 72.1 |
| 空载体(比较例) | N.D. |

[0291] 如表12所示,在未表达羧酸还原酶和氨基转移酶的利用空载体转化后的菌株的培养样品中,未观察到从己二酸生成HMD的情况。在共表达了野生型的羧酸还原酶和氨基转移酶的菌株的培养样品中,确认到HMD生成。与其相比,在使用了突变型的羧酸还原酶和氨基转移酶的情况下,确认到1.4倍至1.6倍的HMD生成。可知通过使用突变型羧酸还原酶,能够促进从己二酸向HMD的菌体内转换。

[0292] 工业实用性

[0293] 本发明能够用于包括羧酸还原反应的有效的发酵生产,可期待应用于目标化合物的工业规模的生产。

[0294]

| | | |
|---|---|---|
| 0-1 | PCT/RO/134 表<br>有关保藏的微生物或其他生物材料的说明(PCT 细则第 13 条之 2) | |
| 0-1-1 | 软件版本 | JPO-PAS<br>i450 |
| 0-2 | 国际申请号 | |
| 0-3 | 申请人或代理人的档案号 | 200888WO01 |

[0295]

| | | |
|---|---|---|
| 1 | 下面的说明与本申请说明书中此处提到的保藏的微生物或其他生物材料相关。 | |
| 1-1 | 段落编号 | 0076 |
| 1-3 | 保藏事项 | |
| 1-3-1 | 保藏单位名称 | IPOD 独立行政法人 制品评价技术基础机构 专利生物保藏中心(IPOD, NITE) |
| 1-3-2 | 保藏单位地址 | 日本〒292-0818 千叶县木更津市かずさ鎌足 2-5-8 120 号室 |
| 1-3-3 | 保藏日期 | 2011 年 10 月 25 日(25.10.2011) |
| 1-3-4 | 保藏号 | IPOD BP-11515 |
| 1-5 | 本说明针对的指定国 | 全部指定国 |
| 2 | 下面的说明与本申请说明书中此处提到的保藏的微生物或其他生物材料相关。 | |
| 2-1 | 段落编号 | 0095 |
| 2-3 | 保藏事项 | |
| 2-3-1 | 保藏单位名称 | IPOD 独立行政法人 制品评价技术基础机构 专利生物保藏中心(NPMD) |
| 2-3-2 | 保藏单位地址 | 日本〒292-0818 千叶县木更津市かずさ鎌足 2-5-8 122 号室 |
| 2-3-3 | 保藏日期 | 2019 年 03 月 18 日(18.03.2019) |
| 2-3-4 | 保藏号 | NPMD BP-02919 |
| 2-5 | 本说明针对的指定国 | 全部指定国 |

[0296] 由受理局填写

[0297]

| | | |
|---|---|---|
| 0-4 | 本页已经和国际申请一起收到(是/否) | √ |
| 0-4-1 | 授权官员 | 永井恒男 |

[0298] 由国际局填写

[0299]

| | | |
|---|---|---|
| 0-5 | 国际局收到本页日期 | |
| 0-5-1 | 授权官员 | |

序列表

<110> 旭化成株式会社

<120> 具有羧酸还原活性的重组多肽

<130> 200888W001

<150> JP2021-057308

<151> 2021-03-30

<160> 72

<170> PatentIn version 3.5

<210> 1

<211> 1178

<212> PRT

<213> 脓肿分枝杆菌

<400> 1

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

195 200 205

Ala Leu Glu Ala Val Arg Ala Arg Leu Glu Ser Ala Gly Arg Thr Ile

210 215 220

Val Val Glu Ala Leu Ala Glu Ala Leu Ala Arg Gly Arg Asp Leu Pro

225 230 235 240

Ala Ala Pro Leu Pro Ser Ala Asp Pro Asp Ala Leu Arg Leu Leu Ile

245 250 255

Tyr Thr Ser Gly Ser Thr Gly Thr Pro Lys Gly Ala Met Tyr Pro Gln

260 265 270

Trp Leu Val Ala Asn Leu Trp Gln Lys Lys Trp Leu Thr Asp Asp Val

275 280 285

Ile Pro Ser Ile Gly Val Asn Phe Met Pro Met Ser His Leu Ala Gly

290 295 300

Arg Leu Thr Leu Met Gly Thr Leu Ser Gly Gly Gly Thr Ala Tyr Tyr

305 310 315 320

Ile Ala Ser Ser Asp Leu Ser Thr Phe Phe Glu Asp Ile Ala Leu Ile

325 330 335

Arg Pro Ser Glu Val Leu Phe Val Pro Arg Val Val Glu Met Val Phe

340 345 350

Gln Arg Phe Gln Ala Glu Leu Asp Arg Ser Leu Ala Pro Gly Glu Ser

355 360 365

Asn Ser Glu Ile Ala Glu Arg Ile Lys Val Arg Ile Arg Glu Gln Asp

370 375 380

Phe Gly Gly Arg Val Leu Ser Ala Gly Ser Gly Ser Ala Pro Leu Ser

385 390 395 400

Pro Glu Met Thr Glu Phe Met Glu Ser Leu Leu Gln Val Pro Leu Arg

405 410 415

Asp Gly Tyr Gly Ser Thr Glu Ala Gly Gly Val Trp Arg Asp Gly Val

420 425 430

Leu Gln Arg Pro Pro Val Thr Asp Tyr Lys Leu Val Asp Val Pro Glu

435 440 445

Leu Gly Tyr Phe Thr Thr Asp Ser Pro His Pro Arg Gly Glu Leu Arg

450 455 460

Leu Lys Ser Glu Thr Met Phe Pro Gly Tyr Tyr Lys Arg Pro Glu Thr

465 470 475 480

Thr Ala Asp Val Phe Asp Asp Glu Gly Tyr Tyr Lys Thr Gly Asp Val

485 490 495

Val Ala Glu Leu Gly Pro Asp His Leu Lys Tyr Leu Asp Arg Val Lys

500 505 510

Asn Val Leu Lys Leu Ala Gln Gly Glu Phe Val Ala Val Ser Lys Leu
515 520 525
Glu Ala Ala Tyr Thr Gly Ser Pro Leu Val Arg Gln Ile Phe Val Tyr
530 535 540
Gly Asn Ser Glu Arg Ser Phe Leu Leu Ala Val Val Val Pro Thr Pro
545 550 555 560
Glu Val Leu Glu Arg Tyr Ala Asp Ser Pro Asp Ala Leu Lys Pro Leu
565 570 575
Ile Gln Asp Ser Leu Gln Gln Val Ala Lys Asp Ala Glu Leu Gln Ser
580 585 590
Tyr Glu Ile Pro Arg Asp Phe Ile Val Glu Thr Val Pro Phe Thr Val
595 600 605
Glu Ser Gly Leu Leu Ser Asp Ala Arg Lys Leu Leu Arg Pro Lys Leu
610 615 620
Lys Asp His Tyr Gly Glu Arg Leu Glu Ala Leu Tyr Ala Glu Leu Ala
625 630 635 640
Glu Ser Gln Asn Glu Arg Leu Arg Gln Leu Ala Arg Glu Ala Ala Thr
645 650 655
Arg Pro Val Leu Glu Thr Val Thr Asp Ala Ala Ala Ala Leu Leu Gly
660 665 670
Ala Ser Ser Ser Asp Leu Ala Pro Asp Val Arg Phe Ile Asp Leu Gly
675 680 685
Gly Asp Ser Leu Ser Ala Leu Ser Tyr Ser Glu Leu Leu Arg Asp Ile
690 695 700
Phe Glu Val Asp Val Pro Val Gly Val Ile Asn Ser Val Ala Asn Asp
705 710 715 720
Leu Ala Ala Ile Ala Arg His Ile Glu Ala Gln Arg Thr Gly Ala Ala
725 730 735
Thr Gln Pro Thr Phe Ala Ser Val His Gly Lys Asp Ala Thr Val Ile
740 745 750
Thr Ala Gly Glu Leu Thr Leu Asp Lys Phe Leu Asp Glu Ser Leu Leu
755 760 765
Lys Ala Ala Lys Asp Val Gln Pro Ala Thr Ala Asp Val Lys Thr Val
770 775 780
Leu Val Thr Gly Gly Asn Gly Trp Leu Gly Arg Trp Leu Val Leu Asp
785 790 795 800
Trp Leu Glu Arg Leu Ala Pro Asn Gly Gly Lys Val Tyr Ala Leu Ile
805 810 815
Arg Gly Ala Asp Ala Glu Ala Ala Arg Ala Arg Leu Asp Ala Val Tyr

820 825 830

Glu Ser Gly Asp Pro Lys Leu Ser Ala His Tyr Arg Gln Leu Ala Gln

835 840 845

Gln Ser Leu Glu Val Ile Ala Gly Asp Phe Gly Asp Gln Asp Leu Gly

850 855 860

Leu Ser Gln Glu Val Trp Gln Lys Leu Ala Lys Asp Val Asp Leu Ile

865 870 875 880

Val His Ser Gly Ala Leu Val Asn His Val Leu Pro Tyr Ser Gln Leu

885 890 895

Phe Gly Pro Asn Val Ala Gly Thr Ala Glu Ile Ile Lys Leu Ala Ile

900 905 910

Ser Glu Arg Leu Lys Pro Val Thr Tyr Leu Ser Thr Val Gly Ile Ala

915 920 925

Asp Gln Ile Pro Val Thr Glu Phe Glu Glu Asp Ser Asp Val Arg Val

930 935 940

Met Ser Ala Glu Arg Gln Ile Asn Asp Gly Tyr Ala Asn Gly Tyr Gly

945 950 955 960

Asn Ser Lys Trp Ala Gly Glu Val Leu Leu Arg Glu Ala His Asp Leu

965 970 975

Ala Gly Leu Pro Val Arg Val Phe Arg Ser Asp Met Ile Leu Ala His

980 985 990

Ser Asp Tyr His Gly Gln Leu Asn Val Thr Asp Val Phe Thr Arg Ser

995 1000 1005

Ile Gln Ser Leu Leu Leu Thr Gly Val Ala Pro Ala Ser Phe Tyr

1010 1015 1020

Glu Leu Asp Ala Asp Gly Asn Arg Gln Arg Ala His Tyr Asp Gly

1025 1030 1035

Val Pro Gly Asp Phe Thr Ala Ala Ser Ile Thr Ala Ile Gly Gly

1040 1045 1050

Val Asn Val Val Asp Gly Tyr Arg Ser Phe Asp Val Phe Asn Pro

1055 1060 1065

His His Asp Gly Val Ser Met Asp Thr Phe Val Asp Trp Leu Ile

1070 1075 1080

Asp Ala Gly Tyr Lys Ile Ala Arg Ile Asp Asp Tyr Asp Gln Trp

1085 1090 1095

Leu Ala Arg Phe Glu Leu Ala Leu Lys Gly Leu Pro Glu Gln Gln

1100 1105 1110

Arg Gln Gln Ser Val Leu Pro Leu Leu Lys Met Tyr Glu Lys Pro

1115 1120 1125

Gln Pro Ala Ile Asp Gly Ser Ala Leu Pro Thr Ala Glu Phe Ser
1130 1135 1140
Arg Ala Val His Glu Ala Lys Val Gly Asp Ser Gly Glu Ile Pro
1145 1150 1155
His Val Thr Lys Glu Leu Ile Leu Lys Tyr Ala Ser Asp Ile Gln
1160 1165 1170
Leu Leu Gly Leu Val
1175

<210> 2
<211> 3534
<212> DNA
<213> 脓肿分枝杆菌
<400> 2

atgactgaaa cgatctccac agcggctgtc cccactacgg atctcgaaga gcaggtgaag 60
cgacgcatcg agcaggtcgt gtccaacgat ccgcagctgg cggcgcttct cccggaagat 120
tcggtcaccg aggcggtcaa cgagcccgat ctaccgctgg tcgaggtgat caggcgactg 180
ctggagggct acggtgaccg cccggcactc ggccagcgcg ccttcgagtt cgtcaccggg 240
gacgacggtg cgaccgtgat cgcgctgaag cccgaataca ccaccgtctc ctaccgcgag 300
ttgtgggaac gtgccgaggc tatcgctgcc gcgtggcacg agcagggcat ccgtgacggc 360
gacttcgtcg ctcagttggg tttcaccagc acggacttcg cgtcgctcga cgtcgcggga 420
ttgcgtctgg gcaccgtctc ggtgcccctg cagacgggcg cgtcgctgca gcagcgcaac 480
gcgattctcg aagagacccg gcccgcagtc tttgccgcga gtatcgaata ccttgatgcc 540
gccgtcgatt cggtgcttgc gacccccctcg gtgcgactcc tctcggtttt cgactatcac 600
gcggaggtcg acagccagcg cgaagcgctg gaggctgtgc gggcccggct tgagagtgcc 660
ggccggacga tcgtcgtcga ggccctggcg gaggctctcg cgcgggggcg ggacctgccc 720
gccgcgccgc tgcccagtgc agatcccgat gccttgcgtc tgctcatcta cacctccggc 780
agcaccggta cccccaaggg cgccatgtat ccgcaatggc tggtcgccaa cttgtggcag 840
aagaagtggc tcaccgacga tgtgattccg tccataggcg tgaacttcat gcccatgagc 900
cacctggcgg gtcgcctcac tctcatgggc acccttttccg gtggcggaac cgcctactac 960
atcgcttcga gcgatctttc gactttcttc gaggacatcg cgctcatccg cccctccgaa 1020
gtgctcttcg tgccgcgtgt ggtggagatg gtgttccagc gttttcaggc agaattggac 1080
cggtcccttg ccccgggtga gagcaactcc gagatcgcgg agcgaatcaa ggtccgcatc 1140
cgggaacagg acttcggcgg gcgtgtgctc agtgctggct ccgggtcggc cccgttgtct 1200
cctgagatga cggagttcat ggagtcgctg ctgcaggtgc cgttgcgcga cgggtatggg 1260
tccaccgagg ccggtggtgt gtggcgtgac ggagtcctgc agcgtccgcc cgtcaccgac 1320
tacaagctgg ttgacgttcc ggaactcgga tacttcacca cagattcgcc gcatccccgt 1380
ggcgagctgc ggttgaagtc ggagacgatg ttccccggct actacaagcg cccggagacc 1440
actgccgatg tcttcgatga cgaggggtac tacaagaccg gtgacgtggt cgccgagctc 1500
gggccggatc acctcaagta cctcgaccgc gtcaagaacg tcctcaagct cgcgcaggga 1560

gagtttgtcg cggtgtcaaa gctggaggcc gcttacaccg gcagcccgct ggtccggcag 1620
atctttgtgt acgggaacag tgaacgctcg ttcctgctgg ctgtcgtggt cccgacaccc 1680
gaagtccttg agcggtacgc agattcgcca gatgcgctca agcccttgat ccaggattcg 1740
ctgcagcagg tcgccaagga cgcggagctg caatcctatg agataccgcg cgacttcatc 1800
gttgagacgg tgccgttcac cgtcgagtcc ggattgctat cggacgcgcg aaagctgctg 1860
cgccccaagc tgaaggatca ctacggagag aggctggagg cgctgtacgc cgaactggcg 1920
gaaagccaga atgagcggct gcgccagttg gccagggagg cagccacgcg cccggtcctg 1980
gagacggtga ccgatgcggc cgccgcgctg ctgggcgcat cgtcctcgga tctggctcct 2040
gatgtgcgat tcatcgacct cggtggcgac tcactgtcgg cgctgtcgta ctccgagctg 2100
ctgcgcgaca tctttgaggt ggacgttccg gtgggcgtca tcaacagcgt cgccaacgac 2160
cttgccgcga tcgcccggca catcgaggcg cagcggaccg gcgccgctac gcagccgacc 2220
tttgcgtcgg tccacggcaa ggacgcgacg gtcatcaccg ccggtgaact caccctcgac 2280
aagttcttgg acgagtcact gttgaaagcg gccaaggacg ttcagccggc aacggccgat 2340
gtcaagaccg ttctagtgac cggcggcaac ggctggttgg gtcgttggct ggtgctcgat 2400
tggctggagc ggttggcacc caatggtggc aaggtctacg ccctcattcg tggcgccgat 2460
gccgaagcag cccgggcacg gttggacgcc gtgtacgaat cgggtgatcc caagctgtcc 2520
gcgcattatc gtcagctggc gcaacagagt ctggaagtta tcgccggcga tttcggcgac 2580
caggatctcg gtctatccca ggaagtttgg cagaagctgg ccaaggacgt ggacctgatc 2640
gtgcactccg gtgccttggt gaaccacgtg ctgccgtaca gccagttgtt cggtccgaat 2700
gtggcgggta ccgccgagat catcaagctg gcaatttcgg agcggctcaa gccggtcacc 2760
tacctgtcga cggtgggcat cgccgaccag attccggtga cggagttcga ggaagactcc 2820
gatgttcgtg tgatgtcggc cgagcgccag atcaatgacg gctacgcgaa cggatacggc 2880
aactcaaaat gggccggcga ggtgctgttg cgggaggctc atgacctagc ggggctgccg 2940
gtgcgtgtgt tccgctccga catgatcctg gcgcacagtg actaccacgg acagctcaac 3000
gtcaccgacg tgttcacccg gagcatccag agtctgctgc tcaccggtgt tgcaccggcc 3060
agcttctatg aattggatgc cgacggcaat cggcagcgcg ctcactatga cggtgtgccc 3120
ggcgatttca ccgccgcatc gatcaccgcc atcggcggtg tgaacgtggt agacggttac 3180
cgcagcttcg acgtgttcaa cccgcaccat gacggtgtct cgatggatac cttcgtcgac 3240
tggctgatcg acgcaggcta caagatcgcg cggatcgacg attacgacca gtggctcgcc 3300
cggttcgagc tggccctcaa gggattgccc gagcagcagc ggcaacagtc ggtgttgcca 3360
cttctcaaga tgtacgagaa gccgcaaccg gcgatcgacg gaagtgcact tccgaccgca 3420
gaattcagtc gcgccgtgca cgaggcgaag gtcggagaca gcggtgagat accgcacgtc 3480
accaaggagc tgatcctcaa gtacgccagc gatattcagc tgttgggcct ggtg 3534

<210> 3
<211> 1178
<212> PRT
<213> 人工序列
<220>
<223> W283R/ A303M

<400> 3

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

Arg Leu Thr Leu Met Gly Thr Leu Ser Gly Gly Gly Thr Ala Tyr Tyr
305 310 315 320
Ile Ala Ser Ser Asp Leu Ser Thr Phe Phe Glu Asp Ile Ala Leu Ile
325 330 335
Arg Pro Ser Glu Val Leu Phe Val Pro Arg Val Val Glu Met Val Phe
340 345 350
Gln Arg Phe Gln Ala Glu Leu Asp Arg Ser Leu Ala Pro Gly Glu Ser
355 360 365
Asn Ser Glu Ile Ala Glu Arg Ile Lys Val Arg Ile Arg Glu Gln Asp
370 375 380
Phe Gly Gly Arg Val Leu Ser Ala Gly Ser Gly Ser Ala Pro Leu Ser
385 390 395 400
Pro Glu Met Thr Glu Phe Met Glu Ser Leu Leu Gln Val Pro Leu Arg
405 410 415
Asp Gly Tyr Gly Ser Thr Glu Ala Gly Gly Val Trp Arg Asp Gly Val
420 425 430
Leu Gln Arg Pro Pro Val Thr Asp Tyr Lys Leu Val Asp Val Pro Glu
435 440 445
Leu Gly Tyr Phe Thr Thr Asp Ser Pro His Pro Arg Gly Glu Leu Arg
450 455 460
Leu Lys Ser Glu Thr Met Phe Pro Gly Tyr Tyr Lys Arg Pro Glu Thr
465 470 475 480
Thr Ala Asp Val Phe Asp Asp Glu Gly Tyr Tyr Lys Thr Gly Asp Val
485 490 495
Val Ala Glu Leu Gly Pro Asp His Leu Lys Tyr Leu Asp Arg Val Lys
500 505 510
Asn Val Leu Lys Leu Ala Gln Gly Glu Phe Val Ala Val Ser Lys Leu
515 520 525
Glu Ala Ala Tyr Thr Gly Ser Pro Leu Val Arg Gln Ile Phe Val Tyr
530 535 540
Gly Asn Ser Glu Arg Ser Phe Leu Leu Ala Val Val Val Pro Thr Pro
545 550 555 560
Glu Val Leu Glu Arg Tyr Ala Asp Ser Pro Asp Ala Leu Lys Pro Leu
565 570 575
Ile Gln Asp Ser Leu Gln Gln Val Ala Lys Asp Ala Glu Leu Gln Ser
580 585 590
Tyr Glu Ile Pro Arg Asp Phe Ile Val Glu Thr Val Pro Phe Thr Val
595 600 605
Glu Ser Gly Leu Leu Ser Asp Ala Arg Lys Leu Leu Arg Pro Lys Leu

610 615 620
Lys Asp His Tyr Gly Glu Arg Leu Glu Ala Leu Tyr Ala Glu Leu Ala
625 630 635 640
Glu Ser Gln Asn Glu Arg Leu Arg Gln Leu Ala Arg Glu Ala Ala Thr
645 650 655
Arg Pro Val Leu Glu Thr Val Thr Asp Ala Ala Ala Ala Leu Leu Gly
660 665 670
Ala Ser Ser Ser Asp Leu Ala Pro Asp Val Arg Phe Ile Asp Leu Gly
675 680 685
Gly Asp Ser Leu Ser Ala Leu Ser Tyr Ser Glu Leu Leu Arg Asp Ile
690 695 700
Phe Glu Val Asp Val Pro Val Gly Val Ile Asn Ser Val Ala Asn Asp
705 710 715 720
Leu Ala Ala Ile Ala Arg His Ile Glu Ala Gln Arg Thr Gly Ala Ala
725 730 735
Thr Gln Pro Thr Phe Ala Ser Val His Gly Lys Asp Ala Thr Val Ile
740 745 750
Thr Ala Gly Glu Leu Thr Leu Asp Lys Phe Leu Asp Glu Ser Leu Leu
755 760 765
Lys Ala Ala Lys Asp Val Gln Pro Ala Thr Ala Asp Val Lys Thr Val
770 775 780
Leu Val Thr Gly Gly Asn Gly Trp Leu Gly Arg Trp Leu Val Leu Asp
785 790 795 800
Trp Leu Glu Arg Leu Ala Pro Asn Gly Gly Lys Val Tyr Ala Leu Ile
805 810 815
Arg Gly Ala Asp Ala Glu Ala Ala Arg Ala Arg Leu Asp Ala Val Tyr
820 825 830
Glu Ser Gly Asp Pro Lys Leu Ser Ala His Tyr Arg Gln Leu Ala Gln
835 840 845
Gln Ser Leu Glu Val Ile Ala Gly Asp Phe Gly Asp Gln Asp Leu Gly
850 855 860
Leu Ser Gln Glu Val Trp Gln Lys Leu Ala Lys Asp Val Asp Leu Ile
865 870 875 880
Val His Ser Gly Ala Leu Val Asn His Val Leu Pro Tyr Ser Gln Leu
885 890 895
Phe Gly Pro Asn Val Ala Gly Thr Ala Glu Ile Ile Lys Leu Ala Ile
900 905 910
Ser Glu Arg Leu Lys Pro Val Thr Tyr Leu Ser Thr Val Gly Ile Ala
915 920 925

Asp Gln Ile Pro Val Thr Glu Phe Glu Glu Asp Ser Asp Val Arg Val
930 935 940
Met Ser Ala Glu Arg Gln Ile Asn Asp Gly Tyr Ala Asn Gly Tyr Gly
945 950 955 960
Asn Ser Lys Trp Ala Gly Glu Val Leu Leu Arg Glu Ala His Asp Leu
965 970 975
Ala Gly Leu Pro Val Arg Val Phe Arg Ser Asp Met Ile Leu Ala His
980 985 990
Ser Asp Tyr His Gly Gln Leu Asn Val Thr Asp Val Phe Thr Arg Ser
995 1000 1005
Ile Gln Ser Leu Leu Leu Thr Gly Val Ala Pro Ala Ser Phe Tyr
1010 1015 1020
Glu Leu Asp Ala Asp Gly Asn Arg Gln Arg Ala His Tyr Asp Gly
1025 1030 1035
Val Pro Gly Asp Phe Thr Ala Ala Ser Ile Thr Ala Ile Gly Gly
1040 1045 1050
Val Asn Val Val Asp Gly Tyr Arg Ser Phe Asp Val Phe Asn Pro
1055 1060 1065
His His Asp Gly Val Ser Met Asp Thr Phe Val Asp Trp Leu Ile
1070 1075 1080
Asp Ala Gly Tyr Lys Ile Ala Arg Ile Asp Asp Tyr Asp Gln Trp
1085 1090 1095
Leu Ala Arg Phe Glu Leu Ala Leu Lys Gly Leu Pro Glu Gln Gln
1100 1105 1110
Arg Gln Gln Ser Val Leu Pro Leu Leu Lys Met Tyr Glu Lys Pro
1115 1120 1125
Gln Pro Ala Ile Asp Gly Ser Ala Leu Pro Thr Ala Glu Phe Ser
1130 1135 1140
Arg Ala Val His Glu Ala Lys Val Gly Asp Ser Gly Glu Ile Pro
1145 1150 1155
His Val Thr Lys Glu Leu Ile Leu Lys Tyr Ala Ser Asp Ile Gln
1160 1165 1170
Leu Leu Gly Leu Val
1175

<210> 4
<211> 1178
<212> PRT
<213> 人工序列
<220>

<223> W283R/ L335F

<400> 4

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

290 295 300
Arg Leu Thr Leu Met Gly Thr Leu Ser Gly Gly Gly Thr Ala Tyr Tyr
305 310 315 320
Ile Ala Ser Ser Asp Leu Ser Thr Phe Phe Glu Asp Ile Ala Phe Ile
325 330 335
Arg Pro Ser Glu Val Leu Phe Val Pro Arg Val Val Glu Met Val Phe
340 345 350
Gln Arg Phe Gln Ala Glu Leu Asp Arg Ser Leu Ala Pro Gly Glu Ser
355 360 365
Asn Ser Glu Ile Ala Glu Arg Ile Lys Val Arg Ile Arg Glu Gln Asp
370 375 380
Phe Gly Gly Arg Val Leu Ser Ala Gly Ser Gly Ser Ala Pro Leu Ser
385 390 395 400
Pro Glu Met Thr Glu Phe Met Glu Ser Leu Leu Gln Val Pro Leu Arg
405 410 415
Asp Gly Tyr Gly Ser Thr Glu Ala Gly Gly Val Trp Arg Asp Gly Val
420 425 430
Leu Gln Arg Pro Pro Val Thr Asp Tyr Lys Leu Val Asp Val Pro Glu
435 440 445
Leu Gly Tyr Phe Thr Thr Asp Ser Pro His Pro Arg Gly Glu Leu Arg
450 455 460
Leu Lys Ser Glu Thr Met Phe Pro Gly Tyr Tyr Lys Arg Pro Glu Thr
465 470 475 480
Thr Ala Asp Val Phe Asp Asp Glu Gly Tyr Tyr Lys Thr Gly Asp Val
485 490 495
Val Ala Glu Leu Gly Pro Asp His Leu Lys Tyr Leu Asp Arg Val Lys
500 505 510
Asn Val Leu Lys Leu Ala Gln Gly Glu Phe Val Ala Val Ser Lys Leu
515 520 525
Glu Ala Ala Tyr Thr Gly Ser Pro Leu Val Arg Gln Ile Phe Val Tyr
530 535 540
Gly Asn Ser Glu Arg Ser Phe Leu Leu Ala Val Val Val Pro Thr Pro
545 550 555 560
Glu Val Leu Glu Arg Tyr Ala Asp Ser Pro Asp Ala Leu Lys Pro Leu
565 570 575
Ile Gln Asp Ser Leu Gln Gln Val Ala Lys Asp Ala Glu Leu Gln Ser
580 585 590
Tyr Glu Ile Pro Arg Asp Phe Ile Val Glu Thr Val Pro Phe Thr Val
595 600 605

Glu Ser Gly Leu Leu Ser Asp Ala Arg Lys Leu Leu Arg Pro Lys Leu
610 615 620
Lys Asp His Tyr Gly Glu Arg Leu Glu Ala Leu Tyr Ala Glu Leu Ala
625 630 635 640
Glu Ser Gln Asn Glu Arg Leu Arg Gln Leu Ala Arg Glu Ala Ala Thr
645 650 655
Arg Pro Val Leu Glu Thr Val Thr Asp Ala Ala Ala Ala Leu Leu Gly
660 665 670
Ala Ser Ser Ser Asp Leu Ala Pro Asp Val Arg Phe Ile Asp Leu Gly
675 680 685
Gly Asp Ser Leu Ser Ala Leu Ser Tyr Ser Glu Leu Leu Arg Asp Ile
690 695 700
Phe Glu Val Asp Val Pro Val Gly Val Ile Asn Ser Val Ala Asn Asp
705 710 715 720
Leu Ala Ala Ile Ala Arg His Ile Glu Ala Gln Arg Thr Gly Ala Ala
725 730 735
Thr Gln Pro Thr Phe Ala Ser Val His Gly Lys Asp Ala Thr Val Ile
740 745 750
Thr Ala Gly Glu Leu Thr Leu Asp Lys Phe Leu Asp Glu Ser Leu Leu
755 760 765
Lys Ala Ala Lys Asp Val Gln Pro Ala Thr Ala Asp Val Lys Thr Val
770 775 780
Leu Val Thr Gly Gly Asn Gly Trp Leu Gly Arg Trp Leu Val Leu Asp
785 790 795 800
Trp Leu Glu Arg Leu Ala Pro Asn Gly Gly Lys Val Tyr Ala Leu Ile
805 810 815
Arg Gly Ala Asp Ala Glu Ala Ala Arg Ala Arg Leu Asp Ala Val Tyr
820 825 830
Glu Ser Gly Asp Pro Lys Leu Ser Ala His Tyr Arg Gln Leu Ala Gln
835 840 845
Gln Ser Leu Glu Val Ile Ala Gly Asp Phe Gly Asp Gln Asp Leu Gly
850 855 860
Leu Ser Gln Glu Val Trp Gln Lys Leu Ala Lys Asp Val Asp Leu Ile
865 870 875 880
Val His Ser Gly Ala Leu Val Asn His Val Leu Pro Tyr Ser Gln Leu
885 890 895
Phe Gly Pro Asn Val Ala Gly Thr Ala Glu Ile Ile Lys Leu Ala Ile
900 905 910
Ser Glu Arg Leu Lys Pro Val Thr Tyr Leu Ser Thr Val Gly Ile Ala

915 920 925

Asp Gln Ile Pro Val Thr Glu Phe Glu Glu Asp Ser Asp Val Arg Val

930 935 940

Met Ser Ala Glu Arg Gln Ile Asn Asp Gly Tyr Ala Asn Gly Tyr Gly

945 950 955 960

Asn Ser Lys Trp Ala Gly Glu Val Leu Leu Arg Glu Ala His Asp Leu

965 970 975

Ala Gly Leu Pro Val Arg Val Phe Arg Ser Asp Met Ile Leu Ala His

980 985 990

Ser Asp Tyr His Gly Gln Leu Asn Val Thr Asp Val Phe Thr Arg Ser

995 1000 1005

Ile Gln Ser Leu Leu Leu Thr Gly Val Ala Pro Ala Ser Phe Tyr

1010 1015 1020

Glu Leu Asp Ala Asp Gly Asn Arg Gln Arg Ala His Tyr Asp Gly

1025 1030 1035

Val Pro Gly Asp Phe Thr Ala Ala Ser Ile Thr Ala Ile Gly Gly

1040 1045 1050

Val Asn Val Val Asp Gly Tyr Arg Ser Phe Asp Val Phe Asn Pro

1055 1060 1065

His His Asp Gly Val Ser Met Asp Thr Phe Val Asp Trp Leu Ile

1070 1075 1080

Asp Ala Gly Tyr Lys Ile Ala Arg Ile Asp Asp Tyr Asp Gln Trp

1085 1090 1095

Leu Ala Arg Phe Glu Leu Ala Leu Lys Gly Leu Pro Glu Gln Gln

1100 1105 1110

Arg Gln Gln Ser Val Leu Pro Leu Leu Lys Met Tyr Glu Lys Pro

1115 1120 1125

Gln Pro Ala Ile Asp Gly Ser Ala Leu Pro Thr Ala Glu Phe Ser

1130 1135 1140

Arg Ala Val His Glu Ala Lys Val Gly Asp Ser Gly Glu Ile Pro

1145 1150 1155

His Val Thr Lys Glu Leu Ile Leu Lys Tyr Ala Ser Asp Ile Gln

1160 1165 1170

Leu Leu Gly Leu Val

1175

<210> 5

<211> 154

<212> DNA

<213> 人工序列

<220>
<223> 启动子
<400> 5
tcaagcccaa aggaagagtg aggcgagtca gtcgcgtaat gcttaggcac aggattgatt 60
tgtcgcaatg attgacacga ttccgcttga cgctgcgtaa ggtttttgta attttacagg 120
caacctttta ttcactaaca aatagctggt ggaa 154
<210> 6
<211> 47
<212> DNA
<213> 人工序列
<220>
<223> 终止子
<400> 6
aaaaaaaaaa aggtgccgtg gaaacacggc acctttttt ttcttta 47
<210> 7
<211> 25
<212> DNA
<213> 人工序列
<220>
<223> 引物1
<400> 7
atgactgaaa cgatctccac agcgg 25
<210> 8
<211> 25
<212> DNA
<213> 人工序列
<220>
<223> 引物2
<400> 8
ctacaccagg cccaacagct gaata 25
<210> 9
<211> 25
<212> DNA
<213> 人工序列
<220>
<223> 引物3
<400> 9
atgatcgaga caattttgcc tgctg 25
<210> 10

<211> 20
<212> DNA
<213> 人工序列
<220>
<223> 引物4
<400> 10
tcaggcgtac gcgatcgcgg 20
<210> 11
<211> 20
<212> DNA
<213> 人工序列
<220>
<223> 引物5
<400> 11
atgatacgcg agcctccgga 20
<210> 12
<211> 23
<212> DNA
<213> 人工序列
<220>
<223> 引物6
<400> 12
ttacgcttct tcgacactta ctc 23
<210> 13
<211> 25
<212> DNA
<213> 人工序列
<220>
<223> 引物7
<400> 13
cttcttctgc cacaagttgg cgacc 25
<210> 14
<211> 22
<212> DNA
<213> 人工序列
<220>
<223> 引物8
<400> 14
tccgaagtgc tcttcgtgcc gc 22

<210> 15
<211> 22
<212> DNA
<213> 人工序列
<220>
<223> 引物9
<400> 15
atgggcagca gccatcacca tc 22
<210> 16
<211> 26
<212> DNA
<213> 人工序列
<220>
<223> 引物10
<400> 16
ggtatatctc cttattaaag ttaaac 26
<210> 17
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 引物11
<400> 17
tttaataagg agatatacca tgatacgcga gcctc 35
<210> 18
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 引物12
<400> 18
ggtgatggct gctgcccatt tacgcttctt cgaca 35
<210> 19
<211> 25
<212> DNA
<213> 人工序列
<220>
<223> 引物13
<400> 19

agcttgcggc cgcataatgc ttaag 25

<210> 20

<211> 34

<212> DNA

<213> 人工序列

<220>

<223> 引物14

<400> 20

atgtatatct ccttcttata cttaactaat atac 34

<210> 21

<211> 35

<212> DNA

<213> 人工序列

<220>

<223> 引物15

<400> 21

agaaggagat atacatatga ctgaaacgat ctcca 35

<210> 22

<211> 33

<212> DNA

<213> 人工序列

<220>

<223> 引物16

<400> 22

ggtggcagca gcctagctac accaggccca aca 33

<210> 23

<211> 200

<212> DNA

<213> 人工序列

<220>

<223> 合成片段/ Trp283

<220>

<221> misc_feature

<222> (17)..(18)

<223> n is a, c, g, t or u

<400> 23

cttgtggcag aagaagnnkc tcaccgacga tgtgattccg tccataggcg tgaacttcat 60

gcccatgagc cacctggcgg gtcgcctcac tctcatgggc acccttccg gtggcggaac 120

cgcctactac atcgcttcga gcgatctttc gactttcttc gaggacatcg cgctcatccg 180

cccctccgaa gtgctcttcg 200

<210> 24

<211> 200

<212> DNA

<213> 人工序列

<220>

<223> 合成片段/ Leu 284

<220>

<221> misc_feature

<222> (20)..(21)

<223> n is a, c, g, t or u

<400> 24

cttgtggcag aagaagtggn nkaccgacga tgtgattccg tccataggcg tgaacttcat 60

gcccatgagc cacctggcgg gtcgcctcac tctcatgggc acccttccg gtggcggaac 120

cgcctactac atcgcttcga gcgatctttc gactttcttc gaggacatcg cgctcatccg 180

cccctccgaa gtgctcttcg 200

<210> 25

<211> 200

<212> DNA

<213> 人工序列

<220>

<223> 合成片段/ Ala 303

<220>

<221> misc_feature

<222> (77)..(78)

<223> n is a, c, g, t or u

<400> 25

cttgtggcag aagaagtggc tcaccgacga tgtgattccg tccataggcg tgaacttcat 60

gcccatgagc cacctgnnkg gtcgcctcac tctcatgggc acccttccg gtggcggaac 120

cgcctactac atcgcttcga gcgatctttc gactttcttc gaggacatcg cgctcatccg 180

cccctccgaa gtgctcttcg 200

<210> 26

<211> 200

<212> DNA

<213> 人工序列

<220>

<223> 合成片段/ Leu 306

<220>

<221> misc_feature

<222> (86)..(87)
<223> n is a, c, g, t or u
<400> 26
cttgtggcag aagaagtggc tcaccgacga tgtgattccg tccataggcg tgaacttcat 60
gcccatgagc cacctggcgg gtcgcnnkac tctcatgggc acccttccg gtggcggaac 120
cgcctactac atcgcttcga gcgatctttc gactttcttc gaggacatcg cgctcatccg 180
cccctccgaa gtgctcttcg 200
<210> 27
<211> 200
<212> DNA
<213> 人工序列
<220>
<223> 合成片段/ Leu 335
<220>
<221> misc_feature
<222> (173)..(174)
<223> n is a, c, g, t or u
<400> 27
cttgtggcag aagaagtggc tcaccgacga tgtgattccg tccataggcg tgaacttcat 60
gcccatgagc cacctggcgg gtcgcctcac tctcatgggc acccttccg gtggcggaac 120
cgcctactac atcgcttcga gcgatctttc gactttcttc gaggacatcg cgnnkatccg 180
cccctccgaa gtgctcttcg 200
<210> 28
<211> 36
<212> DNA
<213> 人工序列
<220>
<223> 引物17
<400> 28
ttgttgaata aatcgtctgt ttgccgagaa tacgcg 36
<210> 29
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 引物18
<400> 29
tatcgcatgc agatctgcgc taatgggctc cagga 35
<210> 30

<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 引物19
<400> 30
ttgttgaata aatcgcgcgc cagcgctggc tgttt 35
<210> 31
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 引物20
<400> 31
tatcgcatgc agatcccgcc aatgccggaa gagtt 35
<210> 32
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 引物21
<400> 32
ttgttgaata aatcgcgatc gtgatgccgc tgtct 35
<210> 33
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 引物22
<400> 33
tatcgcatgc agatcgacaa cgtaggcttt gttca 35
<210> 34
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 引物23
<400> 34
ttgttgaata aatcgcgcca tctcgacact cttga 35

<210> 35
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 引物24
<400> 35
tatcgcatgc agatcgcgaa catcgatggg ttagc 35
<210> 36
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 引物25
<400> 36
ttgttgaata aatcgtggca tttgcccttc ctgtt 35
<210> 37
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 引物26
<400> 37
tatcgcatgc agatcgtcct gatcctgcaa cggaa 35
<210> 38
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 引物27
<400> 38
ttgttgaata aatcgctcat aacggtactg caaac 35
<210> 39
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 引物28
<400> 39

tatcgcatgc agatctcgca gcaggtaaga tgatt 35

<210> 40

<211> 35

<212> DNA

<213> 人工序列

<220>

<223> 引物29

<400> 40

ttgttgaata aatcgatatt cgtcctaacg aacag 35

<210> 41

<211> 35

<212> DNA

<213> 人工序列

<220>

<223> 引物30

<400> 41

tatcgcatgc agatcttttt gattttcaag tatgt 35

<210> 42

<211> 35

<212> DNA

<213> 人工序列

<220>

<223> 引物31

<400> 42

gatctgcatg cgatatctct agaacgcgta agctt 35

<210> 43

<211> 29

<212> DNA

<213> 人工序列

<220>

<223> 引物32

<400> 43

tctcgagccg atttattcaa caaagccgc 29

<210> 44

<211> 25

<212> DNA

<213> 人工序列

<220>

<223> 引物33

<400> 44
actttcgttt tcgggcattt cgtcc 25
<210> 45
<211> 24
<212> DNA
<213> 人工序列
<220>
<223> 引物34
<400> 45
ccaatatgag ggcagagaac gatc 24
<210> 46
<211> 20
<212> DNA
<213> 人工序列
<220>
<223> 引物35
<400> 46
atgcgctgat gtgataatgc 20
<210> 47
<211> 19
<212> DNA
<213> 人工序列
<220>
<223> 引物36
<400> 47
ccttctcctt gttgcttta 19
<210> 48
<211> 20
<212> DNA
<213> 人工序列
<220>
<223> 引物37
<400> 48
gaggcctttg ctgcgactgc 20
<210> 49
<211> 21
<212> DNA
<213> 人工序列
<220>

<223> 引物38
<400> 49
agttcctcct tttcggatga t 21
<210> 50
<211> 18
<212> DNA
<213> 人工序列
<220>
<223> 引物39
<400> 50
atgccggatg cgacgctt 18
<210> 51
<211> 20
<212> DNA
<213> 人工序列
<220>
<223> 引物40
<400> 51
tcattttgca tatagcccct 20
<210> 52
<211> 20
<212> DNA
<213> 人工序列
<220>
<223> 引物41
<400> 52
tgacctgggc agtaatggtg 20
<210> 53
<211> 25
<212> DNA
<213> 人工序列
<220>
<223> 引物42
<400> 53
caggatctcc gttgctttat gagtc 25
<210> 54
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 引物43
<400> 54
cgtggtgttg aaagccgatt attg 24
<210> 55
<211> 23
<212> DNA
<213> 人工序列
<220>
<223> 引物44
<400> 55
cataaacttc cagttctccg ccc 23
<210> 56
<211> 20
<212> DNA
<213> 人工序列
<220>
<223> 引物45
<400> 56
tcgcacacta acagactgaa 20
<210> 57
<211> 20
<212> DNA
<213> 人工序列
<220>
<223> 引物46
<400> 57
tgtgtttact cctgattagc 20
<210> 58
<211> 25
<212> DNA
<213> 人工序列
<220>
<223> 引物47
<400> 58
tattctcaat ccgtttcagc acgcg 25
<210> 59
<211> 21
<212> DNA

<213> 人工序列
<220>
<223> 引物48
<400> 59
ttcgggatca ccaccaggcc g 21
<210> 60
<211> 20
<212> DNA
<213> 人工序列
<220>
<223> 引物49
<400> 60
cggaaatgga cgaacagtgg 20
<210> 61
<211> 22
<212> DNA
<213> 人工序列
<220>
<223> 引物50
<400> 61
cgtcatcagc gtttaccaga tt 22
<210> 62
<211> 20
<212> DNA
<213> 人工序列
<220>
<223> 引物51
<400> 62
gcataaacac tgtccgcgtc 20
<210> 63
<211> 24
<212> DNA
<213> 人工序列
<220>
<223> 引物52
<400> 63
gaaatcgaga aggcagaagc gaaa 24
<210> 64
<211> 20

<212> DNA
<213> 人工序列
<220>
<223> 引物53
<400> 64
ggtgcggtca ccattgttcg 20
<210> 65
<211> 20
<212> DNA
<213> 人工序列
<220>
<223> 引物54
<400> 65
catatcgcac gccagcagtg 20
<210> 66
<211> 21
<212> DNA
<213> 人工序列
<220>
<223> 引物55
<400> 66
ggtgaggatg gagagttcat g 21
<210> 67
<211> 21
<212> DNA
<213> 人工序列
<220>
<223> 引物56
<400> 67
cagttcgatt tgcgccacca g 21
<210> 68
<211> 20
<212> DNA
<213> 人工序列
<220>
<223> 引物57
<400> 68
tccgctagtg tgatttcagg 20
<210> 69

<211> 23
<212> DNA
<213> 人工序列
<220>
<223> 引物58
<400> 69
gaaattatta tgccgccagg cgt 23
<210> 70
<211> 20
<212> DNA
<213> 人工序列
<220>
<223> 引物59
<400> 70
ttatggtctg ggcgacatgc 20
<210> 71
<211> 22
<212> DNA
<213> 人工序列
<220>
<223> 引物60
<400> 71
gcatcatcct ggtcatatac cc 22
<210> 72
<211> 1188
<212> PRT
<213> 褶皱慢反应脂肪酸菌
<400> 72

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

85 90 95

Val Ser Tyr Arg Glu Leu Leu Glu Arg Ser His Ala Ile Ala Ser Glu
100 105 110

Leu Ala Asn His Ala Glu Ala Pro Val Lys Ala Gly Glu Phe Ile Ala
115 120 125

Thr Ile Gly Phe Thr Ser Thr Asp Tyr Thr Ser Leu Asp Ile Ala Gly
130 135 140

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

Asp Thr Leu Lys Ala Ile Ala Glu Glu Thr Ala Pro Ala Val Phe Gly
165 170 175

Ala Ser Val Glu His Leu Asp Asn Ala Val Thr Thr Ala Leu Ala Thr
180 185 190

Pro Ser Val Arg Arg Leu Leu Val Phe Asp Tyr Arg Gln Gly Val Asp
195 200 205

Glu Asp Arg Glu Ala Val Glu Ala Ala Arg Ser Arg Leu Ala Glu Ala
210 215 220

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

Arg Ala Leu Pro Arg Val Ala Leu Pro Pro Ala Thr Asp Ala Gly Asp
245 250 255

Asp Ser Leu Ser Leu Leu Ile Tyr Thr Ser Gly Ser Thr Gly Thr Pro
260 265 270

Lys Gly Ala Met Tyr Pro Glu Arg Asn Val Ala Gln Phe Trp Gly Gly
275 280 285

Ile Trp His Asn Ala Phe Asp Asp Gly Asp Ser Ala Pro Asp Val Pro
290 295 300

Asp Ile Met Val Asn Phe Met Pro Leu Ser His Val Ala Gly Arg Ile
305 310 315 320

Gly Leu Met Gly Thr Leu Ser Ser Gly Gly Thr Thr Tyr Phe Ile Ala
325 330 335

Lys Ser Asp Leu Ser Thr Phe Phe Glu Asp Tyr Ser Leu Ala Arg Pro
340 345 350

Thr Lys Leu Phe Phe Val Pro Arg Ile Cys Glu Met Ile Tyr Gln His
355 360 365

Tyr Gln Ser Glu Leu Asp Arg Ile Gly Ala Ala Asp Gly Ser Pro Gln
370 375 380

Ala Glu Ala Ile Lys Thr Glu Leu Arg Glu Lys Leu Leu Gly Gly Arg
385 390 395 400

Val Leu Thr Ala Gly Ser Gly Ser Ala Pro Met Ser Pro Glu Leu Thr
405 410 415
Ala Phe Ile Glu Ser Val Leu Gln Val His Leu Val Asp Gly Tyr Gly
420 425 430
Ser Thr Glu Ala Gly Pro Val Trp Arg Asp Arg Lys Leu Val Lys Pro
435 440 445
Pro Val Thr Glu His Lys Leu Ile Asp Val Pro Glu Leu Gly Tyr Phe
450 455 460
Ser Thr Asp Ser Pro Tyr Pro Arg Gly Glu Leu Ala Ile Lys Thr Gln
465 470 475 480
Thr Ile Leu Pro Gly Tyr Tyr Lys Arg Pro Glu Thr Thr Ala Glu Val
485 490 495
Phe Asp Glu Asp Gly Phe Tyr Leu Thr Gly Asp Val Val Ala Glu Val
500 505 510
Ala Pro Glu Glu Phe Val Tyr Val Asp Arg Arg Lys Asn Val Leu Lys
515 520 525
Leu Ser Gln Gly Glu Phe Val Ala Leu Ser Lys Leu Glu Ala Ala Tyr
530 535 540
Gly Thr Ser Pro Leu Val Arg Gln Ile Ser Val Tyr Gly Ser Ser Gln
545 550 555 560
Arg Ser Tyr Leu Leu Ala Val Val Val Pro Thr Pro Glu Ala Leu Ala
565 570 575
Lys Tyr Gly Asp Gly Glu Ala Val Lys Ser Ala Leu Gly Asp Ser Leu
580 585 590
Gln Lys Ile Ala Arg Glu Glu Gly Leu Gln Ser Tyr Glu Val Pro Arg
595 600 605
Asp Phe Ile Ile Glu Thr Asp Pro Phe Thr Ile Glu Asn Gly Ile Leu
610 615 620
Ser Asp Ala Gly Lys Thr Leu Arg Pro Lys Val Lys Ala Arg Tyr Gly
625 630 635 640
Glu Arg Leu Glu Ala Leu Tyr Ala Gln Leu Ala Glu Thr Gln Ala Gly
645 650 655
Glu Leu Arg Ser Ile Arg Val Gly Ala Gly Glu Arg Pro Val Ile Glu
660 665 670
Thr Val Gln Arg Ala Ala Ala Ala Leu Leu Gly Ala Ser Ala Ala Glu
675 680 685
Val Asp Pro Glu Ala His Phe Ser Asp Leu Gly Gly Asp Ser Leu Ser
690 695 700
Ala Leu Thr Tyr Ser Asn Phe Leu His Glu Ile Phe Gln Val Glu Val

705 710 715 720
Pro Val Ser Val Ile Val Ser Ala Ala Asn Asn Leu Arg Ser Val Ala
725 730 735
Ala His Ile Glu Lys Glu Arg Ser Ser Gly Ser Asp Arg Pro Thr Phe
740 745 750
Ala Ser Val His Gly Ala Gly Ala Thr Thr Ile Arg Ala Ser Asp Leu
755 760 765
Lys Leu Glu Lys Phe Leu Asp Ala Gln Thr Leu Ala Ala Ala Pro Ser
770 775 780
Leu Pro Arg Pro Ala Ser Glu Val Arg Thr Val Leu Leu Thr Gly Ser
785 790 795 800
Asn Gly Trp Leu Gly Arg Phe Leu Ala Leu Ala Trp Leu Glu Arg Leu
805 810 815
Val Pro Gln Gly Gly Lys Val Val Val Ile Val Arg Gly Lys Asp Asp
820 825 830
Lys Ala Ala Lys Ala Arg Leu Asp Ser Val Phe Glu Ser Gly Asp Pro
835 840 845
Ala Leu Leu Ala His Tyr Glu Asp Leu Ala Asp Lys Gly Leu Glu Val
850 855 860
Leu Ala Gly Asp Phe Ser Asp Ala Asp Leu Gly Leu Arg Lys Ala Asp
865 870 875 880
Trp Asp Arg Leu Ala Asp Glu Val Asp Leu Ile Val His Ser Gly Ala
885 890 895
Leu Val Asn His Val Leu Pro Tyr Ser Gln Leu Phe Gly Pro Asn Val
900 905 910
Val Gly Thr Ala Glu Val Ala Lys Leu Ala Leu Thr Lys Arg Leu Lys
915 920 925
Pro Val Thr Tyr Leu Ser Thr Val Ala Val Ala Val Gly Val Glu Pro
930 935 940
Ser Ala Phe Glu Glu Asp Gly Asp Ile Arg Asp Val Ser Ala Val Arg
945 950 955 960
Ser Ile Asp Glu Gly Tyr Ala Asn Gly Tyr Gly Asn Ser Lys Trp Ala
965 970 975
Gly Glu Val Leu Leu Arg Glu Ala Tyr Glu His Ala Gly Leu Pro Val
980 985 990
Arg Val Phe Arg Ser Asp Met Ile Leu Ala His Arg Lys Tyr Thr Gly
995 1000 1005
Gln Leu Asn Val Pro Asp Gln Phe Thr Arg Leu Ile Leu Ser Leu
1010 1015 1020

Leu Ala Thr Gly Ile Ala Pro Lys Ser Phe Tyr Gln Leu Asp Ala
1025 1030 1035
Thr Gly Gly Arg Gln Arg Ala His Tyr Asp Gly Ile Pro Val Asp
1040 1045 1050
Phe Thr Ala Glu Ala Ile Thr Thr Leu Gly Leu Ala Gly Ser Asp
1055 1060 1065
Gly Tyr His Ser Phe Asp Val Phe Asn Pro His His Asp Gly Val
1070 1075 1080
Gly Leu Asp Glu Phe Val Asp Trp Leu Val Glu Ala Gly His Pro
1085 1090 1095
Ile Ser Arg Val Asp Asp Tyr Ala Glu Trp Leu Ser Arg Phe Glu
1100 1105 1110
Thr Ser Leu Arg Gly Leu Pro Glu Ala Gln Arg Gln His Ser Val
1115 1120 1125
Leu Pro Leu Leu His Ala Phe Ala Gln Pro Ala Pro Ala Ile Asp
1130 1135 1140
Gly Ser Pro Phe Gln Thr Lys Asn Phe Gln Ser Ser Val Gln Glu
1145 1150 1155
Ala Lys Val Gly Ala Glu His Asp Ile Pro His Leu Asp Lys Ala
1160 1165 1170
Leu Ile Val Lys Tyr Ala Glu Asp Ile Lys Gln Leu Gly Leu Leu
1175 1180 1185

(序列编号1)

MTETISTAAVPTTDLEEQVKRRIEQVVSNDPQLAALLPED
SVTEAVNEPDLPLVEVIRRLLEGYGDRPALGQRAFEFVTG
DDGATVIALKPEYTTVSYRELWERAEAIAAAWHEQGIRDG
DFVAQLGFTSTDFASLDVAGLRLGTVSVPLQTGASLQQRN
AILEETRPAVFAASIEYLDAAVDSVLATPSVRLLSVFDYH
AEVDSQREALEAVRARLESAGRTIVVEALAEALARGRDLP
AAPLPSADPDALRLLIYTSGSTGTPKGAMYPQWLVANLWQ
KKWLTDDVIPSIGVNFMPMSHLAGRLTLMGTLSGGGTAYY
IASSDLSTFFEDIALIRPSEVLFVPRVVEMVFQRFQAELD
RSLAPGESNSEIAERIKVRIREQDFGGRVLSAGSGSAPLS
PEMTEFMESLLQVPLRDGYGSTEAGGVWRDGVLQRPPVTD
YKLVDVPELGYFTTDSPHPRGELRLKSETMFPGYYKRPET
TADVFDDEGYYKTGDVVAELGPDHLKYLDRVKNVLKLAQG
EFVAVSKLEAAYTGSPLVRQIFVYGNSERSFLLAVVVPTP
EVLERYADSPDALKPLIQDSLQQVAKDAELQSYEIPRDFI
VETVPFTVESGLLSDARKLLRPKLKDHYGERLEALYAELA
ESQNERLRQLAREAATRPVLETVTDAAAALLGASSSDLAP
DVRFIDLGGDSLSALSYSELLRDIFEVDVPVGVINSVAND
LAAIARHIEAQRTGAATQPTFASVHGKDATVITAGELTLD
KFLDESLLKAAKDVQPATADVKTVLVTGGNGWLGRWLVLD
WLERLAPNGGKVYALIRGADAEAARARLDAVYESGDPKLS
AHYRQLAQQSLEVIAGDFGDQDLGLSQEVWQKLAKDVDLI
VHSGALVNHVLPYSQLFGPNVAGTAEIIKLAISERLKPVT
YLSTVGIADQIPVTEFEEDSDVRVMSAERQINDGYANGYG
NSKWAGEVLLREAHDLAGLPVRVFRSDMILAHSDYHGQLN
VTDVFTRSIQSLLLTGVAPASFYELDADGNRQRAHYDGVP
GDFTAASITAIGGVNVVDGYRSFDVFNPHHDGVSMDTFVD
WLIDAGYKIARIDDYDQWLARFELALKGLPEQQRQQSVLP
LLKMYEKPQPAIDGSALPTAEFSRAVHEAKVGDSGEIPHV
TKELILKYASDIQLLGLV

图1

(序列编号2)

ATGACTGAAACGATCTCCACAGCGGCTGTCCCCACTACGGATCTCGAAGAGCAGGTGAAGCGACGCATCGAGCAGGTCGTGTCCAA
CGATCCGCAGCTGGCGGCGCTTCTCCCGGAAGATTCGGTCACCGAGGCGGTCAACGAGCCCGATCTACCGCTGGTCGAGGTGATCA
GGCGACTGCTGGAGGGCTACGGTGACCGCCCGGCACTCGGCCAGCGCGCCTTCGAGTTCGTCACCGGGGACGACGGTGCGACCGTG
ATCGCGCTGAAGCCCGAATACACCACCGTCTCCTACCGCGAGTTGTGGGAACGTGCCGAGGCTATCGCTGCCGCGTGGCACGAGCA
GGGCATCCGTGACGGCGACTTCGTCGCTCAGTTGGGTTTCACCAGCACGGACTTCGCGTCGCTCGACGTCGCGGGATTGCGTCTGG
GCACCGTCTCGGTGCCCCTGCAGACGGGCGCGTCGCTGCAGCAGCGCAACGCGATTCTCGAAGAGACCCGGCCCGCAGTCTTTGCC
GCGAGTATCGAATACCTTGATGCCGCCGTCGATTCGGTGCTTGCGACCCCCTCGGTGCGACTCCTCTCGGTTTTCGACTATCACGC
GGAGGTCGACAGCCAGCGCGAAGCGCTGGAGGCTGTGCGGGCCCGGCTTGAGAGTGCCGGCCGGACGATCGTCGTCGAGGCCCTGG
CGGAGGCTCTCGCGCGGGGGCGGGACCTGCCCGCCGCGCCGCTGCCCAGTGCAGATCCCGATGCCTTGCGTCTGCTCATCTACACC
TCCGGCAGCACCGGTACCCCCAAGGGCGCCATGTATCCGCAATGGCTGGTCGCCAACTTGTGGCAGAAGAAGTGGCTCACCGACGA
TGTGATTCCGTCCATAGGCGTGAACTTCATGCCCATGAGCCACCTGGCGGGTCGCCTCACTCTCATGGGCACCCTTTCCGGTGGCG
GAACCGCCTACTACATCGCTTCGAGCGATCTTTCGACTTTCTTCGAGGACATCGCGCTCATCCGCCCCTCCGAAGTGCTCTTCGTG
CCGCGTGTGGTGGAGATGGTGTTCCAGCGTTTTCAGGCAGAATTGGACCGGTCCCTTGCCCCGGGTGAGAGCAACTCCGAGATCGC
GGAGCGAATCAAGGTCCGCATCCGGGAACAGGACTTCGGCGGGCGTGTGCTCAGTGCTGGCTCCGGGTCGGCCCCGTTGTCTCCTG
AGATGACGGAGTTCATGGAGTCGCTGCTGCAGGTGCCGTTGCGCGACGGGTATGGGTCCACCGAGGCCGGTGGTGTGTGGCGTGAC
GGAGTCCTGCAGCGTCCGCCCGTCACCGACTACAAGCTGGTTGACGTTCCGGAACTCGGATACTTCACCACAGATTCGCCGCATCC
CCGTGGCGAGCTGCGGTTGAAGTCGGAGACGATGTTCCCCGGCTACTACAAGCGCCCGGAGACCACTGCCGATGTCTTCGATGACG
AGGGGTACTACAAGACCGGTGACGTGGTCGCCGAGCTCGGGCCGGATCACCTCAAGTACCTCGACCGCGTCAAGAACGTCCTCAAG
CTCGCGCAGGGAGAGTTTGTCGCGGTGTCAAAGCTGGAGGCCGCTTACACCGGCAGCCCGCTGGTCCGGCAGATCTTTGTGTACGG
GAACAGTGAACGCTCGTTCCTGCTGGCTGTCGTGGTCCCGACACCCGAAGTCCTTGAGCGGTACGCAGATTCGCCAGATGCGCTCA
AGCCCTTGATCCAGGATTCGCTGCAGCAGGTCGCCAAGGACGCGGAGCTGCAATCCTATGAGATACCGCGCGACTTCATCGTTGAG
ACGGTGCCGTTCACCGTCGAGTCCGGATTGCTATCGGACGCGCGAAAGCTGCTGCGCCCCAAGCTGAAGGATCACTACGGAGAGAG
GCTGGAGGCGCTGTACGCCGAACTGGCGGAAAGCCAGAATGAGCGGCTGCGCCAGTTGGCCAGGGAGGCAGCCACGCGCCCGGTCC
TGGAGACGGTGACCGATGCGGCCGCCGCGCTGCTGGGCGCATCGTCCTCGGATCTGGCTCCTGATGTGCGATTCATCGACCTCGGT
GGCGACTCACTGTCGGCGCTGTCGTACTCCGAGCTGCTGCGCGACATCTTTGAGGTGGACGTTCCGGTGGGCGTCATCAACAGCGT
CGCCAACGACCTTGCCGCGATCGCCCGGCACATCGAGGCGCAGCGGACCGGCGCCGCTACGCAGCCGACCTTTGCGTCGGTCCACG
GCAAGGACGCGACGGTCATCACCGCCGGTGAACTCACCCTCGACAAGTTCTTGGACGAGTCACTGTTGAAAGCGGCCAAGGACGTT
CAGCCGGCAACGGCCGATGTCAAGACCGTTCTAGTGACCGGCGGCAACGGCTGGTTGGGTCGTTGGCTGGTGCTCGATTGGCTGGA
GCGGTTGGCACCCAATGGTGGCAAGGTCTACGCCCTCATTCGTGGCGCCGATGCCGAAGCAGCCCGGGCACGGTTGGACGCCGTGT
ACGAATCGGGTGATCCCAAGCTGTCCGCGCATTATCGTCAGCTGGCGCAACAGAGTCTGGAAGTTATCGCCGGCGATTTCGGCGAC
CAGGATCTCGGTCTATCCCAGGAAGTTTGGCAGAAGCTGGCCAAGGACGTGGACCTGATCGTGCACTCCGGTGCCTTGGTGAACCA
CGTGCTGCCGTACAGCCAGTTGTTCGGTCCGAATGTGGCGGGTACCGCCGAGATCATCAAGCTGGCAATTTCGGAGCGGCTCAAGC
CGGTCACCTACCTGTCGACGGTGGGCATCGCCGACCAGATTCCGGTGACGGAGTTCGAGGAAGACTCCGATGTTCGTGTGATGTCG
GCCGAGCGCCAGATCAATGACGGCTACGCGAACGGATACGGCAACTCAAAATGGGCCGGCGAGGTGCTGTTGCGGGAGGCTCATGA
CCTAGCGGGGCTGCCGGTGCGTGTGTTCCGCTCCGACATGATCCTGGCGCACAGTGACTACCACGGACAGCTCAACGTCACCGACG
TGTTCACCCGGAGCATCCAGAGTCTGCTGCTCACCGGTGTTGCACCGGCCAGCTTCTATGAATTGGATGCCGACGGCAATCGGCAG
CGCGCTCACTATGACGGTGTGCCCGGCGATTTCACCGCCGCATCGATCACCGCCATCGGCGGTGTGAACGTGGTAGACGGTTACCG
CAGCTTCGACGTGTTCAACCCGCACCATGACGGTGTCTCGATGGATACCTTCGTCGACTGGCTGATCGACGCAGGCTACAAGATCG
CGCGGATCGACGATTACGACCAGTGGCTCGCCCGGTTCGAGCTGGCCCTCAAGGGATTGCCCGAGCAGCAGCGGCAACAGTCGGTG
TTGCCACTTCTCAAGATGTACGAGAAGCCGCAACCGGCGATCGACGGAAGTGCACTTCCGACCGCAGAATTCAGTCGCGCCGTGCA
CGAGGCGAAGGTCGGAGACAGCGGTGAGATACCGCACGTCACCAAGGAGCTGATCCTCAAGTACGCCAGCGATATTCAGCTGTTGG
GCCTGGTG

图2

(序列编号3)

MTETISTAAVPTTDLEEQVKRRIEQVVSNDPQLAALLPEDSVTEAVNEPDLPLVEVIRRLLEGYGDRPALGQRAFEFVTGDDGATVIALKPEYTTVSYRE
LWERAEAIAAAWHEQGIRDGDFVAQLGFTSTDFASLDVAGLRLGTVSVPLQTGASLQQRNAILEETRPAVFAASIEYLDAAVDSVLATPSVRLLSVFDYH
AEVDSQREALEAVRARLESAGRTIVVEALAEALARGRDLPAAPLPSADPDALRLLIYTSGSTGTPKGAMYPQWLVANLWQKKRLTDDVIPSIGVNFMPMS
HLMGRLTLMGTLSGGGTAYYIASSDLSTFFEDIALIRPSEVLFVPRVVEMVFQRFQAELDRSLAPGESNSEIAERIKVRIREQDFGGRVLSAGSGSAPLS
PEMTEFMESLLQVPLRDGYGSTEAGGVWRDGVLQRPPVTDYKLVDVPELGYFTTDSPHPRGELRLKSETMFPGYYKRPETTADVFDDEGYYKTGDVVAEL
GPDHLKYLDRVKNVLKLAQGEFVAVSKLEAAYTGSPLVRQIFVYGNSERSFLLAVVVPTPEVLERYADSPDALKPLIQDSLQQVAKDAELQSYEIPRDFI
VETVPFTVESGLLSDARKLLRPKLKDHYGERLEALYAELAESQNERLRQLAREAATRPVLETVTDAAAALLGASSSDLAPDVRFIDLGGDSLSALSYSEL
LRDIFEVDVPVGVINSVANDLAAIARHIEAQRTGAATQPTFASVHGKDATVITAGELTLDKFLDESLLKAAKDVQPATADVKTVLVTGGNGWLGRWLVLD
WLERLAPNGGKVYALIRGADAEAARARLDAVYESGDPKLSAHYRQLAQQSLEVIAGDFGDQDLGLSQEVWQKLAKDVDLIVHSGALVNHVLPYSQLFGPN
VAGTAEIIKLAISERLKPVTYLSTVGIADQIPVTEFEEDSDVRVMSAERQINDGYANGYGNSKWAGEVLLREAHDLAGLPVRVFRSDMILAHSDYHGQLN
VTDVFTRSIQSLLLTGVAPASFYELDADGNRQRAHYDGVPGDFTAASITAIGGVNVVDGYRSFDVFNPHHDGVSMDTFVDWLIDAGYKIARIDDYDQWLA
RFELALKGLPEQQRQQSVLPLLKMYEKPQPAIDGSALPTAEFSRAVHEAKVGDSGEIPHVTKELILKYASDIQLLGLV

图3

(序列编号4)

MTETISTAAVPTTDLEEQVKRRIEQVVSNDPQLAALLPEDSVTEAVNEPDLPLVEVIRRLLEGYGDRPALGQRAFEFVTGDDGATVIALKPEYTTVSYRE
LWERAEAIAAAWHEQGIRDGDFVAQLGFTSTDFASLDVAGLRLGTVSVPLQTGASLQQRNAILEETRPAVFAASIEYLDAAVDSVLATPSVRLLSVFDYH
AEVDSQREALEAVRARLESAGRTIVVEALAEALARGRDLPAAPLPSADPDALRLLIYTSGSTGTPKGAMYPQWLVANLWQKKRLTDDVIPSIGVNFMPMS
HLAGRLTLMGTLSGGGTAYYIASSDLSTFFEDIAFIRPSEVLFVPRVVEMVFQRFQAELDRSLAPGESNSEIAERIKVRIREQDFGGRVLSAGSGSAPLS
PEMTEFMESLLQVPLRDGYGSTEAGGVWRDGVLQRPPVTDYKLVDVPELGYFTTDSPHPRGELRLKSETMFPGYYKRPETTADVFDDEGYYKTGDVVAEL
GPDHLKYLDRVKNVLKLAQGEFVAVSKLEAAYTGSPLVRQIFVYGNSERSFLLAVVVPTPEVLERYADSPDALKPLIQDSLQQVAKDAELQSYEIPRDFI
VETVPFTVESGLLSDARKLLRPKLKDHYGERLEALYAELAESQNERLRQLAREAATRPVLETVTDAAAALLGASSSDLAPDVRFIDLGGDSLSALSYSEL
LRDIFEVDVPVGVINSVANDLAAIARHIEAQRTGAATQPTFASVHGKDATVITAGELTLDKFLDESLLKAAKDVQPATADVKTVLVTGGNGWLGRWLVLD
WLERLAPNGGKVYALIRGADAEAARARLDAVYESGDPKLSAHYRQLAQQSLEVIAGDFGDQDLGLSQEVWQKLAKDVDLIVHSGALVNHVLPYSQLFGPN
VAGTAEIIKLAISERLKPVTYLSTVGIADQIPVTEFEEDSDVRVMSAERQINDGYANGYGNSKWAGEVLLREAHDLAGLPVRVFRSDMILAHSDYHGQLN
VTDVFTRSIQSLLLTGVAPASFYELDADGNRQRAHYDGVPGDFTAASITAIGGVNVVDGYRSFDVFNPHHDGVSMDTFVDWLIDAGYKIARIDDYDQWLA
RFELALKGLPEQQRQQSVLPLLKMYEKPQPAIDGSALPTAEFSRAVHEAKVGDSGEIPHVTKELILKYASDIQLLGLV

图4

(序列编号72)

>WP_007468889.1 羧酸还原酶 [Segniliparus rugosus]

MTESQSYETRQARPAGQSLAERVARLVAIDPQAAAAVPDKAVAERATQQGLRLAQRIEAFLSGYGDRPAL
AQRAFEITKDPITGRAVATLLPKFETVSYRELLERSHAIASELANHAEAPVKAGEFIATIGFTSTDYTSL
DIAGVLLGLTSVPLQTGATTDTLKAIAEETAPAVFGASVEHLDNAVTTALATPSVRRLLVFDYRQGVDED
REAVEAARSRLAEAGSAVLVDTLDEVIARGRALPRVALPPATDAGDDSLSLLIYTSGSTGTPKGAMYPER
NVAQFWGGIWHNAFDDGDSAPDVPDIMVNFMPLSHVAGRIGLMGTLSSGGTTYFIAKSDLSTFFEDYSLA
RPTKLFFVPRICEMIYQHYQSELDRIGAADGSPQAEAIKTELREKLLGGRVLTAGSGSAPMSPELTAFIE
SVLQVHLVDGYGSTEAGPVWRDRKLVKPPVTEHKLIDVPELGYFSTDSPYPRGELAIKTQTILPGYYKRP
ETTAEVFDEDGFYLTGDVVAEVAPEEFVYVDRRKNVLKLSQGEFVALSKLEAAYGTSPLVRQISVYGSSQ
RSYLLAVVVPTPEALAKYGDGEAVKSALGDSLQKIAREEGLQSYEVPRDFIIETDPFTIENGILSDAGKT
LRPKVKARYGERLEALYAQLAETQAGELRSIRVGAGERPVIETVQRAAAALLGASAAEVDPEAHFSDLGG
DSLSALTYSNFLHEIFQVEVPVSVIVSAANNLRSVAAHIEKERSSGSDRPTFASVHGAGATTIRASDLKL
EKFLDAQTLAAAPSLPRPASEVRTVLLTGSNGWLGRFLALAWLERLVPQGGKVVVIVRGKDDKAAKARLD
SVFESGDPALLAHYEDLADKGLEVLAGDFSDADLGLRKADWDRLADEVDLIVHSGALVNHVLPYSQLFGP
NVVGTAEVAKLALTKRLKPVTYLSTVAVAVGVEPSAFEEDGDIRDVSAVRSIDEGYANGYGNSKWAGEVL
LREAYEHAGLPVRVFRSDMILAHRKYTGQLNVPDQFTRLILSLLATGIAPKSFYQLDATGGRQRAHYDGI
PVDFTAEAITTLGLAGSDGYHSFDVFNPHHDGVGLDEFVDWLVEAGHPISRVDDYAEWLSRFETSLRGLP
EAQRQHSVLPLLHAFAQPAPAIDGSPFQTKNFQSSVQEAKVGAEHDIPHLDKALIVKYAEDIKQLGLL

图5A

一致性(61%)

```
序列编号1          1    MTETISTAAVPTTDLEEQVKRRIEQVVSNDPQLAALLPEDSVTEAVNEPDLPLVEVIRRL  60
                        MTE+ S          + +  R+ ++V+ DPQ AA +P+ +V E   +  L L + I
S.S.Rugosus CAR    1    MTESQSYETRQARPAGQSLAERVARLVAIDPQAAAAVPDKAVAERATQQGLRLAQRIEAF  60

Query  61    LEGYGDRPALGQRAFEFVTGD--DGATVIALKPEYTTVSYRELWERAEAIAA--AWH-EQ  115
             L GYGDRPAL QRAFE +T D   G  V  L P++ TVSYREL ER+ AIA+  A H E
Sbjct  61    LSGYGDRPALAQRAFE-ITKDPITGRAVATLLPKFETVSYRELLERSHAIASELANHAEA  119

Query  116   GIRDGDFVAQLGFTSTDFASLDVAGLRLGTVSVPLQTGASLQQRNAILEETRPAVFAASI  175
              ++ G+F+A +GFTSTD+ SLD+AG+ LG  SVPLQTGA+     AI EET PAVF AS+
Sbjct  120   PVKAGEFIATIGFTSTDYTSLDIAGVLLGLTSVPLQTGATTDTLKAIAEETAPAVFGASV  179

Query  176   EYLDAAVDSVLATPSVRLLSVFDYHAEVDSQREALEAVRARLESAGRTIVVEALAEALAR  235
             E+LD AV + LATPSVR L VFDY   VD  REA+EA R+RL  AG  ++V+ L E +AR
Sbjct  180   EHLDNAVTTALATPSVRRLLVFDYRQGVDEDREAVEAARSRLAEAGSAVLVDTLDEVIAR  239

Query  236   GRDLPAAPLP---SADPDALRLLIYTSGSTGTPKGAMYPQWLVANLWQK--KRLTDDV--  288
             GR LP   LP    A  D+L LLIYTSGSTGTPKGAMYP+  VA  W       DD
Sbjct  240   GRALPRVALPPATDAGDDSLSLLIYTSGSTGTPKGAMYPERNVAQFWGGIWHNAFDDGDS  299

Query  289   ---IPSIGVNFMPMSHLMGRLTLMGTLSGGGTAYYIASSDLSTFFEDIALIRPSEVLFVP  345
                +P I VNFMP+SH+ GR+ LMGTLS GGT Y+IA SDLSTFFED +L RP+++ FVP
Sbjct  300   APDVPDIMVNFMPLSHVAGRIGLMGTLSSGGTTYFIAKSDLSTFFEDYSLARPTKLFFVP  359

Query  346   RVVEMVFQRFQAELDRSLAPGESNSEIAERIKVRIREQDFGGRVLSAGSGSAPLSPEMTE  405
             R+ EM++Q +Q+ELDR  A     S  AE IK  +RE+  GGRVL+AGSGSAP+SPE+T
Sbjct  360   RICEMIYQHYQSELDRIGA--ADGSPQAEAIKTELREKLLGGRVLTAGSGSAPMSPELTA  417

Query  406   FMESLLQVPLRDGYGSTEAGGVWRDGVLQRPPVTDYKLVDVPELGYFTTDSPHPRGELRL  465
             F+ES+LQV L DGYGSTEAG VWRD  L +PPVT++KL+DVPELGYF+TDSP+PRGEL +
Sbjct  418   FIESVLQVHLVDGYGSTEAGPVWRDRKLVKPPVTEHKLIDVPELGYFSTDSPYPRGELAI  477

Query  466   KSETMFPGYYKRPETTADVFDDEGYYKTGDVVAELGPDHLKYLDRVKNVLKLAQGEFVAV  525
             K++T+ PGYYKRPETTA+VFD++G+Y TGDVVAE+ P+   Y+DR KNVLKL+QGEFVA+
Sbjct  478   KTQTILPGYYKRPETTAEVFDEDGFYLTGDVVAEVAPEEFVYVDRRKNVLKLSQGEFVAL  537
```

图5B-1

(续)

```
Query  526   SKLEAAYTGSPLVRQIFVYGNSERSFLLAVVVPTPEVLERYADSPDALKPLIQDSLQQVA  585
             SKLEAAY  SPLVRQI VYG+S+RS+LLAVVVPTPE L +Y D  +A+K  + DSLQ++A
Sbjct  538   SKLEAAYGTSPLVRQISVYGSSQRSYLLAVVVPTPEALAKYGDG-EAVKSALGDSLQKIA  596

Query  586   KDAELQSYEIPRDFIVETVPFTVESGLLSDARKLLRPKLKDHYGERLEALYAELAESQNE  645
             ++  LQSYE+PRDFI+ET PFT+E+G+LSDA K LRPK+K  YGERLEALYA+LAE+Q
Sbjct  597   REEGLQSYEVPRDFIIETDPFTIENGILSDAGKTLRPKVKARYGERLEALYAQLAETQAG  656

Query  646   RLRQLAREAATRPVLETVTDAAAALLGASSSDLAPDVRFIDLGGDSLSALSYSELLRDIF  705
              LR +   A  RPV+ETV  AAAALLGAS++++ P+  F DLGGDSLSAL+YS  L +IF
Sbjct  657   ELRSIRVGAGERPVIETVQRAAAALLGASAAEVDPEAHFSDLGGDSLSALTYSNFLHEIF  716

Query  706   EVDVPVGVINSVANDLAAIARHIEAQRTGAATQPTFASVHGKDATVITAGELTLDKFLDE  765
             +V+VPV VI S AN+L ++A HIE +R+  + +PTFASVHG  AT I A +L L+KFLD
Sbjct  717   QVEVPVSVIVSAANNLRSVAAHIEKERSSGSDRPTFASVHGAGATTIRASDLKLEKFLDA  776

Query  766   SLLKAAKDVQPATADVKTVLVTGGNGWLGRWLVLDWLERLAPNGGKVYALIRGADAEAAR  825
               L AA  +    ++V+TVL+TG NGWLGR+L L WLERL P GGKV  ++RG D +AA+
Sbjct  777   QTLAAAPSLPRPASEVRTVLLTGSNGWLGRFLALAWLERLVPQGGKVVVIVRGKDDKAAK  836

Query  826   ARLDAVYESGDPKLSAHYRQLAQQSLEVIAGDFGDQDLGLSQEVWQKLAKDVDLIVHSGA  885
             ARLD+V+ESGDP L AHY  LA + LEV+AGDF D DLGL +  W +LA +VDLIVHSGA
Sbjct  837   ARLDSVFESGDPALLAHYEDLADKGLEVLAGDFSDADLGLRKADWDRLADEVDLIVHSGA  896

Query  886   LVNHVLPYSQLFGPNVAGTAEIIKLAISERLKPVTYLSTVGIADQIPVTEFEEDSDVRVM  945
             LVNHVLPYSQLFGPNV GTAE+ KLA+++RLKPVTYLSTV +A  +  + FEED D+R +
Sbjct  897   LVNHVLPYSQLFGPNVVGTAEVAKLALTKRLKPVTYLSTVAVAVGVEPSAFEEDGDIRDV  956

Query  946   SAERQINDGYANGYGNSKWAGEVLLREAHDLAGLPVRVFRSDMILAHSDYHGQLNVTDVF  1005
             SA R I++GYANGYGNSKWAGEVLLREA++ AGLPVRVFRSDMILAH  Y GQLNV D F
Sbjct  957   SAVRSIDEGYANGYGNSKWAGEVLLREAYEHAGLPVRVFRSDMILAHRKYTGQLNVPDQF  1016

Query  1006  TRSIQSLLLTGVAPASFYELDADGNRQRAHYDGVPGDFTAASITAIGGVNVVDGYRSFDV  1065
             TR I SLL TG+AP SFY+LDA G RQRAHYDG+P DFTA +IT + G+   DGY SFDV
Sbjct  1017  TRLILSLLATGIAPKSFYQLDATGGRQRAHYDGIPVDFTAEAITTL-GLAGSDGYHSFDV  1075

Query  1066  FNPHHDGVSMDTFVDWLIDAGYKIARIDDYDQWLARFELALKGLPEQQRQQSVLPLLKMY  1125
             FNPHHDGV +D FVDWL++AG+ I+R+DDY +WL+RFE +L+GLPE QRQ SVLPLL  +
Sbjct  1076  FNPHHDGVGLDEFVDWLVEAGHPISRVDDYAEWLSRFETSLRGLPEAQRQHSVLPLLHAF  1135

Query  1126  EKPQPAIDGSALPTAEFSRAVHEAKVGDSGEIPHVTKELILKYASDIQLLGLV  1178
              +P PAIDGS   T  F  +V EAKVG   +IPH+ K LI+KYA DI+ LGL+
Sbjct  1136  AQPAPAIDGSPFQTKNFQSSVQEAKVGAEHDIPHLDKALIVKYAEDIKQLGLL  1188
```

图5B-2

| | 氨基酸编号 | | 氨基酸编号 |
|---|---|---|---|
| 序列编号1 | 61 | LEGYGDRPALGQRAFEFVTGD--DGATVIALKPEYTTVSYRELWERAEAIAA--AWH-EQ | 115 |
| | | L GYGDRPAL QRAFE +T D G V L P++ TVSYREL ER+ AIA+ A H E | |
| S.Rugosus CAR | 61 | LSGYGDRPALAQRAFE-ITKDPITGRAVATLLPKFETVSYRELLERSHAIASELANHAEA | 119 |

图5C

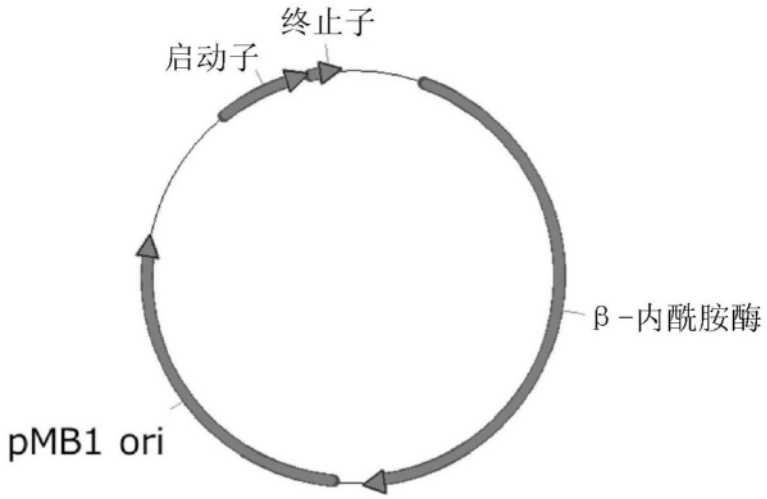



图6

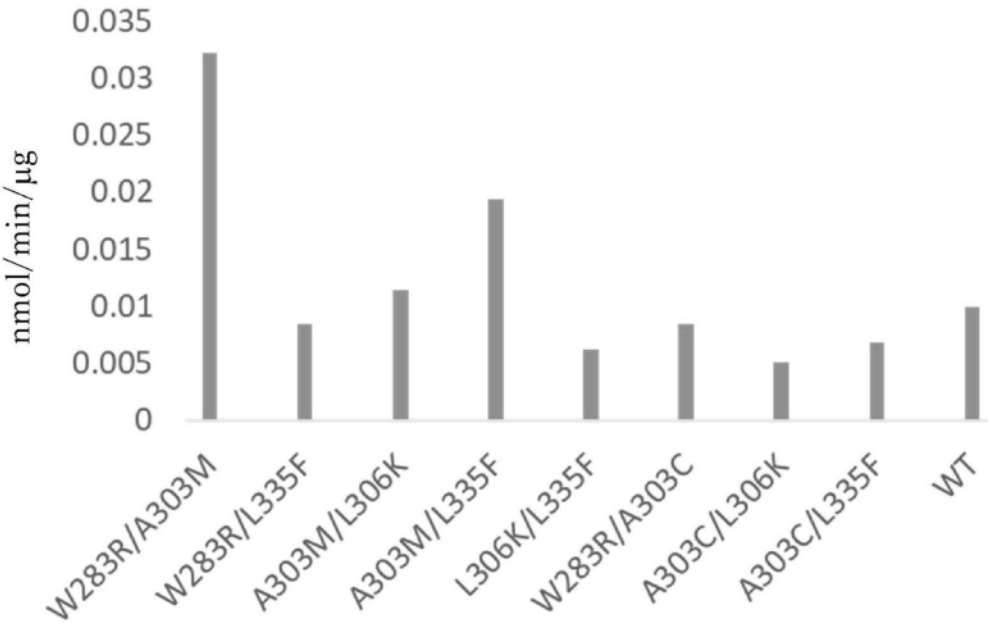



图7A

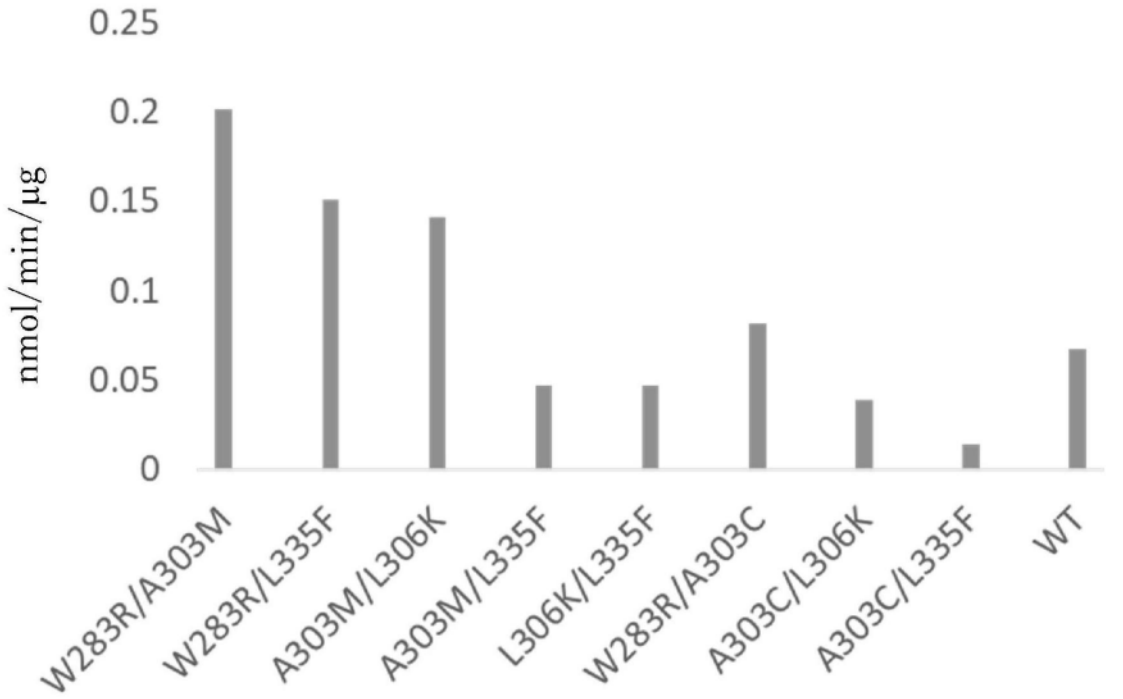

0.25
0.2
0.15
0.1
0.05
0
nmol/min/μg
W283R/A303M
W283R/L335F
A303M/L306K
A303M/L335F
L306K/L335F
W283R/A303C
A303C/L306K
A303C/L335F
WT

图7B

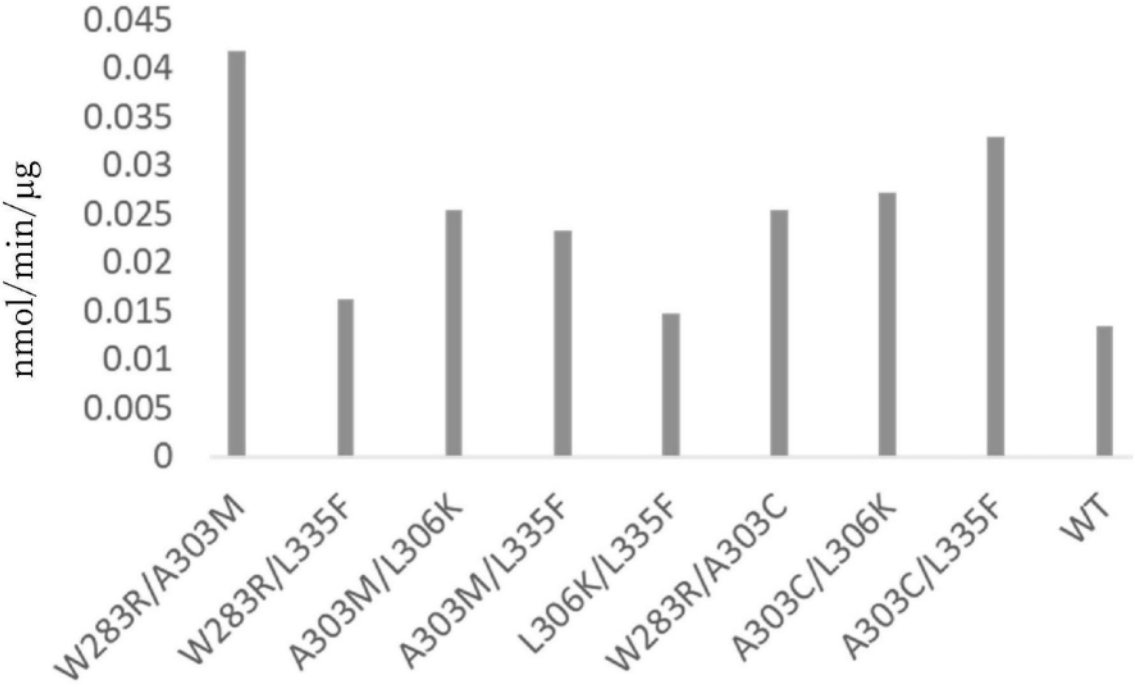

0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
nmol/min/μg
W283R/A303M
W283R/L335F
A303M/L306K
A303M/L335F
L306K/L335F
W283R/A303C
A303C/L306K
A303C/L335F
WT

图7C

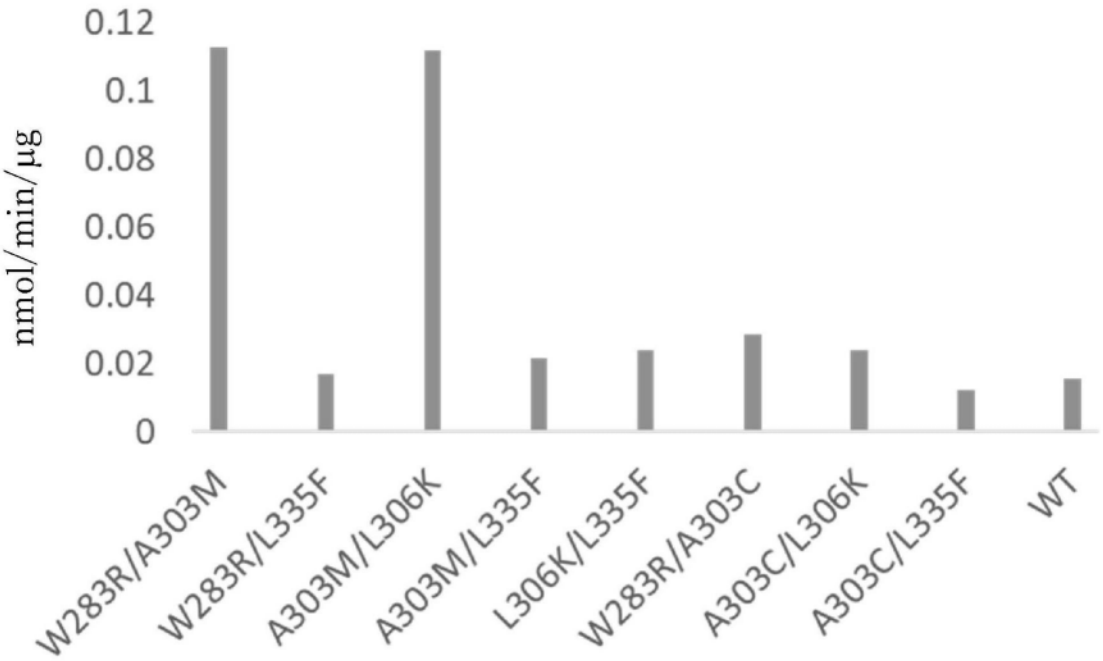

0.12
0.1
0.08
0.06
0.04
0.02
0
nmol/min/μg
W283R/A303M
W283R/L335F
A303M/L306K
A303M/L335F
L306K/L335F
W283R/A303C
A303C/L306K
A303C/L335F
WT

图7D