



(12) 发明专利申请

(10) 申请公布号 CN 114395496 A

(43) 申请公布日 2022.04.26

(21) 申请号 202111260243.5

(51) Int.Cl.

(22) 申请日 2013.03.06

C12N 1/19 (2006.01)

(30) 优先权数据

C12N 1/21 (2006.01)

61/607,479 2012.03.06 US

C12N 15/55 (2006.01)

61/619,112 2012.04.02 US

(62) 分案原申请数据

201380023454.4 2013.03.06

(71) 申请人 利戈斯股份有限公司

地址 美国加利福尼亚

(72) 发明人 J·A·迪特里希 J·L·福特曼

E·J·斯汀

(74) 专利代理机构 中国贸促会专利商标事务所

有限公司 11038

代理人 张小勇

权利要求书1页 说明书61页

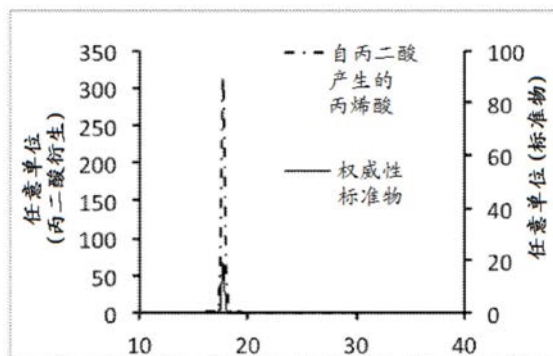
序列表78页 附图1页

(54) 发明名称

用于产生丙二酸的重组宿主细胞

(57) 摘要

本发明涉及用于生物产生丙二酸;筛选丙二酸产生得以改进的丙二酸产生性宿主细胞;纯化生物产生的丙二酸;以及使丙二酸合成转化成工业重要化学品的重组宿主细胞、材料和方法。本发明的重组宿主细胞,其为酵母细胞或细菌细胞,其包含异源丙二酰基-辅酶A水解酶。本发明的重组宿主细胞包含增加乙酰基-辅酶A生物合成、增加丙二酰基-辅酶A生物合成、降低丙二酸分解代谢、增加丙二酸自宿主细胞分泌、增加宿主细胞对丙二酸的耐受性以及增加各种碳源的分解代谢的遗传修饰,从而增加丙二酸产率、效价和/或生产力。



1. 一种重组宿主细胞, 其为酵母细胞或细菌细胞, 其包含异源丙二酰基-辅酶A水解酶, 其中所述异源丙二酰基-辅酶A水解酶选自:

(a) SEQ ID NO: 46、49、51和52, 其中X1选自S、T、N、Q和Y;

(b) SEQ ID NO: 45和48, 其中X1选自R、H、K、S、T、N、Q和Y;

(c) SEQ ID NO: 47, 其中X1是S;

(d) SEQ ID NO: 50, 其中X1选自A、D、K、S、T、N和Y。

2. 根据权利要求1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 45, 其中X1是N。

3. 根据权利要求1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 45, 其中X1是S。

4. 根据权利要求1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 45, 其中X1是Y。

5. 根据权利要求1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 46, 其中X1是N。

6. 根据权利要求1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 46, 其中X1是S。

7. 根据权利要求1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 46, 其中X1是Y。

8. 根据权利要求1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 47, 其中X1是S。

9. 根据权利要求1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 48, 其中X1是N。

10. 根据权利要求1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 48, 其中X1是S。

用于产生丙二酸的重组宿主细胞

[0001] 本申请是申请日为2013年3月6日、申请号为201380023454.4的同名申请的分案申请。

[0002] 政府权益

[0003] 本发明是在由能源部授予的授予号DE-SC0006469下在政府支持下进行。政府享有本发明中的某些权利。

[0004] 发明背景

[0005] 与石油化学工业相关的长期经济和环境顾虑已为增加对用于使碳原料转化成可替代由石油原料产生的化学品的化学品的方法的研究、开发和商业化提供推动力。一种方法是开发用以使可再生原料转化成可替代石油衍生化学品的产品的生物精制方法。改进生物精制方法的两个共同目标包括实现较低生产成本和降低二氧化碳排放。

[0006] 丙二酸(Propanedioic acid/“malonate”,CAS号141-82-2)当前是自非可再生石油原料产生。用醇(例如甲醇或乙醇)使丙二酸的一个或两个羧酸部分单酯化或二酯化分别产生丙二酸单烷基酯和丙二酸二烷基酯。2,2-二甲基-1,3-二噁烷-4,6-二酮(“梅尔德伦氏酸(Meldrum's acid)”CAS号2033-24-1)是使用丙酮在乙酸酐中或使用乙酸异丙烯基酯在酸中自丙二酸产生。

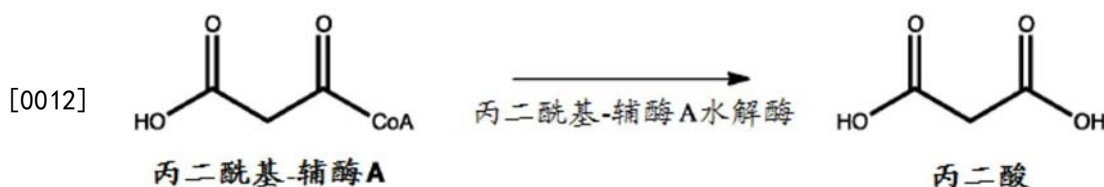
[0007] 化学合成当前是用于合成丙二酸和丙二酸衍生化合物的优选途径。举例来说,丙二酸二烷基酯是通过氰化氢或一氧化碳方法产生。在氰化氢方法中,使氰化钠与氯乙酸钠在高温下反应以产生氰乙酸钠,其随后与醇/矿物酸混合物反应以产生丙二酸二烷基酯。Hildbrand等报道产率75-85%(参见“Malonic acid and Derivatives”,Ullmann's Encyclopedia of Industrial Chemistry,Wiley-VCH,Weinheim,New York(2002))。在一氧化碳方法中,丙二酸二烷基酯(在本文中也称为丙二酸二酯)是通过在醇存在下在高温和高压下,氯乙酸与一氧化碳进行由钴催化的烷氧基羰基化反应来产生。

[0008] 用以获得丙二酸和丙二酸衍生化合物的现存石油化学基产生途径具有低产性、环境损害性、非可再生原料依赖性,且需要对废水和废气进行昂贵处理。因此,仍然需要用于使可再生原料生物催化转化成丙二酸,纯化生物合成丙二酸,以及随后制备下游化学品和产品的材料和方法。

[0009] 发明概述

[0010] 本发明提供用于生物产生丙二酸的重组宿主细胞、材料和方法;用于在丙二酸产生性宿主细胞中检测丙二酸的存在以及测定丙二酸的含量(在本文中称为“感测丙二酸”)的方法;以及用于筛选丙二酸产生得以增加的宿主细胞的方法。此外,本发明提供用于纯化生物产生的丙二酸的方法以及用于使丙二酸转化成其它工业重要化学品的方法。

[0011] 在第一方面,本发明提供包含编码酰基-辅酶A水解酶的异源核酸的重组宿主细胞,所述水解酶催化丙二酰基-辅酶A转化成丙二酸,如此处所说明:



[0013] 这些重组宿主细胞比不包含所述异源水解酶的对细胞产生更多丙二酸。在各种实施方案中,宿主细胞在适当发酵条件下可产生至少10g/L丙二酸,且在各种实施方案中,产生水平可高达50g/L至100g/L或更高。在一些实施方案中,异源核酸编码突变形式的内源表达的酶;因此,本发明提供多种突变酰基-辅酶A水解酶、编码它们的核酸、以及包含那些核酸的重组表达载体。在其它实施方案中,异源核酸编码过表达的内源性酶。此外,在一些实施方案中,异源核酸编码宿主细胞异源性(不在宿主细胞中天然表达)酰基-辅酶A水解酶的野生型或突变酶。在一些实施方案中,宿主细胞是酵母细胞。在其它实施方案中,宿主细胞是细菌细胞。

[0014] 因此,在各种实施方案中,由本发明提供的异源核酸编码野生型或突变形式的酰基-辅酶A水解酶。由本发明提供的核酸编码且适于丙二酰基-辅酶A水解的酰基-辅酶A水解酶的非限制性实例包括选自以下组成的组的野生型和修饰的酶:3-羟基异丁酰基-辅酶A水解酶(EC 3.1.2.4)、3-羟基丙酰基-辅酶A水解酶(EC 3.1.2.4)、乙酰乙酰基-辅酶A水解酶(EC 3.1.2.11)、甲基丙二酰基-辅酶A水解酶(EC 3.1.2.17)、丙酰基-辅酶A水解酶(EC 3.1.2.18)、丁二酰基-辅酶A水解酶(EC 3.1.2.3)和如本文提供的被突变来具有丙二酰基-辅酶A水解酶活性的丙二酰基-辅酶A:ACP转酰基酶(EC 2.3.1.39)。

[0015] 在各种实施方案中,本发明提供一种是3-羟基异丁酰基-辅酶A水解酶(EC 3.1.2.4)的突变体的丙二酰基-辅酶A水解酶。适合的3-羟基异丁酰基-辅酶A水解酶可自真核生物体与原核生物体(包括革兰氏阳性与革兰氏阴性两者)两者获得。在各种实施方案中,3-羟基异丁酰基-辅酶A水解酶是自酵母菌株、芽孢杆菌属种类(*Bacillus species*)和假单胞菌属种类(*Pseudomonas species*)获得。

[0016] 在其它实施方案中,本发明提供一种是如本文提供的被突变来具有丙二酰基-辅酶A水解酶活性的由原核生物编码的丙二酰基-辅酶A:ACP转酰基酶(EC 2.3.1.39)的丙二酰基-辅酶A水解酶。在各种实施方案中,原核生物是革兰氏阴性细菌。在本发明的各种实施方案中,革兰氏阴性细菌是埃希氏菌属(*Escherichia*)。

[0017] 在第二方面,本发明提供编码催化丙二酰基-辅酶A转化成丙二酸的野生型或突变酰基-辅酶A水解酶的重组表达载体。在一些实施方案中,表达载体是酵母表达载体;在其它实施方案中,表达载体是细菌表达载体。在各种实施方案中,细菌表达载体是大肠杆菌(*Escherichia coli*)表达载体。

[0018] 在第三方面,本发明提供适于在使得丙二酸能够分离以及用作化学合成其它有用产品的起始物质的水平下生物合成产生丙二酸的重组宿主细胞。在一些实施方案中,宿主细胞是真核生物。在一些实施方案中,宿主细胞是酵母细胞。在各种实施方案中,酵母是假丝酵母属(*Candida*)、隐球酵母属(*Cryptococcus*)、汉逊酵母属(*Hansenula*)、伊萨酵母属(*Issatchenkia*)、克鲁维酵母属(*Kluyveromyces*)、*Komagataella*、油脂酵母属(*Lipomyces*)、毕赤酵母属(*Pichia*)、红冬孢酵母属(*Rhodospiridium*)、红酵母属(*Rhodotorula*)、酵母属(*Saccharomyces*)或耶罗威亚酵母属(*Yarrowia*)种类。在一些实施

方案中,真核宿主细胞是真菌。在一些实施方案中,细胞宿主是藻类。

[0019] 在其它实施方案中,宿主细胞是细菌细胞。在各种实施方案中,宿主细胞是选自由以下组成的组的细菌细胞:芽孢杆菌属、梭菌属(*Clostridium*)、棒杆菌属(*Corynebacterium*)、埃希氏菌属、假单胞菌属和链霉菌属(*Streptomyces*)。在一些实施方案中,宿主细胞是大肠杆菌细胞。

[0020] 通常,本发明的重组宿主细胞已被遗传修饰以改进丙二酸产率、效价和/或生产力。在各种实施方案中,宿主细胞已通过一种或多种选自由以下修饰组成的组的宿主细胞修饰加以修饰以增加丙二酸生物合成:导致乙酰基-辅酶A生物合成增加、丙二酰基-辅酶A生物合成增加、丙二酰基-辅酶A利用降低、丙二酸分解代谢降低、丙二酸向发酵液中的分泌增加、宿主细胞对发酵液中的丙二酸的耐受性增加、和/或宿主细胞对碳源(例如乙酸盐、海藻酸盐、乙醇、脂肪酸、木质纤维物质、甲醇、戊糖和合成气)的分解代谢增加的修饰。

[0021] 在第四方面,本发明提供用于在重组宿主细胞中产生丙二酸的方法,所述方法大体上包括在使得所述重组宿主细胞能够产生丙二酸的条件下在发酵液中培养所述重组宿主细胞。在一些实施方案中,宿主细胞已被工程化来表达更多或更少的导致比尚未如此工程化的相应细胞产生更多丙二酸的内源性酶。在一些实施方案中,所述方法包括培养表达导致丙二酸的产生得以增加的异源(外来或非天然)酶的重组宿主细胞。在一些实施方案中,用于方法中的宿主细胞包含一种或多种包含异源丙二酰基-辅酶A水解酶的表达载体。在这些方法的一些实施方案中,发酵液补充以促进丙二酸产生的碳源且所述碳源选自由以下组成的组:纤维糊精、5碳糖、6碳糖、二氧化碳、乙醇、甲醇、甘油、乙酸盐和/或脂肪酸。

[0022] 在第五方面,本发明提供包含丙二酸转录因子和对所述转录因子具有反应性的可操作地连接于标记基因的启动子的生物检测器。本发明也提供用于“感测”丙二酸、丙二酸产生和丙二酸产生性宿主细胞的方法以及用于筛选丙二酸产生得以增加的宿主细胞的方法。在各种实施方案中,所述方法包括培养表达丙二酸转录因子且含有对所述转录因子具有反应性并可操作地连接于标记基因的启动子的宿主细胞,以及通过针对标记基因产物的表达进行筛选且选择表达较高水平的标记基因产物的那些宿主细胞来选择丙二酸产生得以改进的宿主细胞。在一些实施方案中,丙二酸是在一种宿主细胞中产生,使来自第一细胞的发酵液与包含丙二酸转录因子和对所述转录因子具有反应性的可操作地连接于标记基因的启动子的第二细胞接触(添加至含有第二细胞的培养基中),且通过鉴定标记基因产物的表达水平最高的细胞来鉴定丙二酸产生得以改进的宿主细胞。在其它实施方案中,丙二酸是在包含丙二酸转录因子和对丙二酸具有反应性的可操作地连接于标记基因的启动子的宿主细胞中产生,且通过筛选并鉴定表达最高水平的标记基因产物的细胞来针对丙二酸产生得以增加筛选丙二酸产生得以增加的宿主细胞。在一些实施方案中,转录因子可结合丙二酸,此导致转录因子结合同源启动子且活化可操作地连接于所述启动子的标记基因。在一些实施方案中,转录因子是MdcY转录因子。在一些实施方案中,实施所述方法以筛选或选择相对于对照细胞,丙二酸产生得以改进的遗传修饰宿主细胞。

[0023] 在第六方面,本发明提供自产生丙二酸的宿主细胞(任选是本发明的宿主细胞)的发酵液分离的纯化丙二酸。本发明也提供用于自产生丙二酸的宿主细胞的发酵液纯化丙二酸的方法,所述方法大体上包括在使得宿主细胞能够产生丙二酸的条件下在发酵液中培养所述宿主细胞,以及自所述发酵液纯化所述丙二酸。在本发明的一些实施方案中,通过在纯

化过程期间使发酵液脱水来增加发酵液中的丙二酸浓度。在本发明的各种实施方案中,脱水是通过反渗透处理、蒸发或两者组合来实现。在各种实施方案中,纯化是通过添加以下一者或多者来实现:二价阳离子、单价阳离子、铵、单取代的胺、双取代的胺、三取代的胺、阳离子纯化树脂或酸。在本发明的各种实施方案中,这些试剂是连同一种或多种有机溶剂一起添加。在本发明的一些实施方案中,疏水性溶剂用于发酵液的液-液提取。在其它实施方案中,丙二酸是通过用酸催化剂和醇进行反应性提取或蒸馏来自发酵液纯化。

[0024] 在第七方面,本发明提供制备衍生自丙二酸的化合物的方法以及由所述方法产生的化合物。所述方法大体上包括使丙二酸与一种或多种基质反应以产生化合物。在这些方法的一些实施方案中,用已建立的合成途径自丙二酸获得的化学品是使用生物源性丙二酸产生。在这些方法的其它实施方案中,提供适于与合成或生物源性丙二酸一起使用的用于产生有用化学品的新合成途径。在一些实施方案中,丙二酸单烷基酯是自生物源性丙二酸合成。在其它实施方案中,丙二酸二烷基酯是自生物源性丙二酸合成。在一些实施方案中,丙烯酸酯是自丙二酸或丙二酸合成。在其它实施方案中,丙烯酸酯是自丙二酸单酯或二酯合成。在其它实施方案中,二羧酸是自丙二酸产生。可根据本发明的方法产生的说明性二羧酸包括选自由以下组成的组的二羧酸:戊二酸、己二酸、庚二酸、辛二酸、壬二酸、癸二酸、十一烷二酸、十二烷二酸、以及各者的相应单烷基和二烷基酯。在本发明的其它实施方案中,二羧酸是自丙二酸衍生化合物产生。在本发明的其它实施方案中, ϵ -己内酰胺是自丙二酸产生。在本发明的其它实施方案中, δ -戊内酰胺是自丙二酸产生。

[0025] 本发明的这些和其它方面和实施方案在附图中加以说明且在以下更详细描述。

[0026] 附图简述

[0027] 图1显示本发明的利用启动子 P_{MdcL} 的大肠杆菌MdcY丙二酸生物感测器的剂量-反应曲线。X轴是添加至发酵液中的外源性丙二酸的浓度;Y轴是在具有25 $\mu\text{g}/\text{ml}$ 四环素的培养基中生长12小时之后的细胞培养物密度(OD_{600})。用包含在 P_{MdcL} 启动子的控制下的MdcY转录因子和tetA基因的质粒S14转化的大肠杆菌在外源性添加丙二酸后产生四环素抗性蛋白质TetA。如实施例21中另外详细所述,生物感测器显示四环素抗性呈丙二酸依赖性增加,如通过 OD_{600} 随外源添加的丙二酸的浓度增加而增加所度量。

[0028] 图2是HPLC色谱图迹线,其显示对根据如实施例28中所述的本发明的方法自丙二酸产生丙烯酸的分离和检测。X轴显示丙烯酸在约17.5分钟洗脱,且Y轴显示由通过监测210nm的UV检测器检测丙烯酸获得的任意单位。根据本发明的方法自丙二酸产生的丙烯酸(虚线)展现的滞留时间与权威性丙烯酸标准物(实线)相同。

[0029] 发明详述

[0030] 本发明提供用于生物产生丙二酸;筛选丙二酸产生得以改进的丙二酸产生性宿主细胞;纯化生物产生的丙二酸;以及使丙二酸合成转化成工业重要化学品的重组宿主细胞、材料和方法。

[0031] 尽管本发明在本文中关于其方面和特定实施方案加以描述,但本领域技术人员将认识到可进行各种变化且可用等效物进行替代而不脱离本发明。本发明不限于特定核酸、表达载体、酶、宿主微生物或方法,因为所述各物可改变。本文中使用的术语是仅出于描述特定方面和实施方案的目的,且不应解释为具有限制性。此外,可根据本发明进行许多修改以适应特定情况、材料、物质组成、方法、一个或多个工艺步骤。所有所述修改都在随附于此

的权利要求的范围内。

[0032] 本文引用的所有专利、专利申请和出版物以引用的方式整体并入本文。

[0033] 章节1:定义

[0034] 在本说明书中和在随后权利要求中,将提及将被定义为具有以下含义的许多术语。

[0035] 如说明书和随附权利要求中所用,除非上下文另外明确规定,否则单数形式“一个(a)”、“一种(an)”和“所述”包括复数个指示物。因此,举例来说,提及“表达载体”包括单个表达载体以及多个相同(例如相同操纵子)或不同的表达载体;提及“细胞”包括单个细胞以及多个细胞;等。

[0036] 蛋白质编码序列中的氨基酸在本文中是由以下缩写和符号来标识。特定氨基酸是由如下单字母缩写来标识:A是丙氨酸,R是精氨酸,N是天冬酰胺,D是天冬氨酸,C是半胱氨酸,Q是谷氨酰胺,E是谷氨酸,G是甘氨酸,H是组氨酸,L是亮氨酸,I是异亮氨酸,K是赖氨酸,M是甲硫氨酸,F是苯丙氨酸,P是脯氨酸,S是丝氨酸,T是苏氨酸,W是色氨酸,Y是酪氨酸,且V是缬氨酸。共有序列中的划线(-)指示在指定位置处不存在氨基酸。共有序列中的加号(+)指示任何氨基酸都可存在于指定位置处。因此,本文共有序列中的加号指示氨基酸在该处通常是非保守的位置;同源酶序列在与共有序列比对时可在指示的“+”位置处具有任何氨基酸。在共有序列中可存在一子组氨基酸之一的位置处,使用以下缩写:B表示氨基酸R、K或H之一存在于指示的位置处;J表示氨基酸D或E之一存在于指示的位置处;O表示氨基酸I、L或V之一存在于指示的位置处;U表示氨基酸S或T之一存在于指示的位置处;且X₁表示氨基酸A、D、R、H、K、S、T、N、Q或Y之一(或那些氨基酸的子组)存在于指示的位置处。X₁的说明性子组包括X₁是A、D、K、S、T、N或Y且X₁是S或N。蛋白质编码序列中的特定氨基酸是通过它们的相应单字母缩写继之以在蛋白质编码序列中的氨基酸位置来标识,其中1对应于在蛋白质的N末端处的氨基酸(通常是甲硫氨酸)。举例来说,酿酒酵母(*S.cerevisiae*)野生型EHD3中的E124是指在自EHD3 N末端甲硫氨酸(即M1)开始的位置124处的谷氨酸。氨基酸取代(即点突变)是通过在亲本蛋白质编码序列中的单字母符号和编号之后标识突变(即子代)氨基酸来指示;举例来说,酿酒酵母EHD3中的E124A是指在EHD3蛋白质编码序列中的位置124处丙氨酸取代谷氨酸。突变亦可标识在括号中,例如EHD3 (E124A)。蛋白质编码序列中的多个点突变由反斜杠(/)分隔;举例来说,EHD3 E124A/Y125A指示突变E124A与Y125A均存在于EHD3蛋白质编码序列中。引入一些实例中的突变的数目已通过括号标识突变之前的划线继之以突变的数目加以注解(例如A5W8H3-1 (E95Q))。带有和不带有划线和数目的Uniprot ID在本文中可互换使用(即A5W8H3-1 (E95Q) = A5W8H3 (E95Q))。

[0037] 如本文所用,术语“表达”在与编码酶的核酸或细胞中的酶本身关联使用时是指可为内源性或外源性(异源)酶的所述酶是在所述细胞中产生。在这些情形下,术语“过表达”是指在内源性酶的情况下,相较于野生型,酶是以较高水平产生,即酶水平增加。本领域技术人员了解的是酶的过表达可通过以下方式来实现:增加用于驱动编码序列的表达的启动子的强度或改变所述启动子的类型,增加核糖体结合位点或Kozak序列的强度,增加mRNA转录物的稳定性,改变密码子使用,增加酶的稳定性等。

[0038] 术语“表达载体”或“载体”是指以下核酸和/或组合物:其包含可例如通过转导、转化或感染引入宿主细胞中,以使所述细胞接着产生(“表达”)不同于为所述细胞天然具有的

核酸和/或蛋白质,或在某种意义上不为所述细胞天然具有的含于如此引入的核酸中或由如此引入的核酸编码的核酸和/或蛋白质的核酸。因此,“表达载体”含有待由宿主细胞表达的核酸(通常是DNA)。任选地,表达载体可含于用以有助于实现核酸进入宿主细胞中的材料中,如与病毒、脂质体、蛋白质包衣等相关的材料。适合用于本发明的各种方面和实施方案中的表达载体包括核酸序列可或已连同任何优选或所需操作元件一起插入其中的表达载体。因此,表达载体可被转移至宿主细胞中且通常在其中复制(但在一些实施方案中,也可采用提供“短暂”表达的非可复制载体)。在一些实施方案中,采用整合至染色体、线粒体或质粒DNA中的表达载体。在其它实施方案中,采用在染色体外复制的表达载体。典型表达载体包括质粒,且表达载体通常含有为转录载体中的核酸所需的操作元件。所述质粒以及其它表达载体在本文中加以描述或为本领域普通技术人员所熟知。

[0039] 术语“发酵(ferment)”、“发酵性(fermentative)”和“发酵(fermentation)”在本文中用于描述在用以产生有用化学品的条件下培养微生物,所述条件包括但不限于发生微生物生长所处的好氧或厌氧条件。

[0040] 如本文所用的术语“异源”是指物质不为细胞天然具有。举例来说,如果以下至少一者是成立的,那么核酸对于细胞而言是异源的,因此就那个细胞而言是“异源核酸”: (a) 所述核酸不天然见于那个细胞中(也就是说它是“外源性”核酸); (b) 所述核酸天然见于给定宿主细胞中(也就是说“内源性”),但所述核酸或由这个核酸的转录和翻译产生的RNA或蛋白质是以非天然(例如大于或小于天然存在)量在所述宿主细胞中产生或存在; (c) 所述核酸包含编码宿主细胞内源性蛋白质,但在序列方面不同于编码相同蛋白质(具有相同或大致上相同氨基酸序列)的内源性核苷酸序列的核苷酸序列,从而通常导致蛋白质在细胞中以较大量产生,或在酶的情况下,产生具有改变的(例如较高或较低或不同)活性的突变形式;和/或(d) 所述核酸包含两个或更多个不以相同彼此关系见于细胞中的核苷酸序列。作为另一实例,如果蛋白质是通过翻译RNA产生或相应RNA是通过转录异源核酸产生,那么所述蛋白质对于宿主细胞而言是异源的;如果蛋白质是内源性蛋白质的突变形式,且突变是通过遗传工程化引入,那么所述蛋白质对于宿主细胞而言也是异源的。

[0041] 术语“宿主细胞”与“宿主微生物”在本文中可互换用于指代可(或已)通过插入表达载体加以转化的活细胞。如本文所述的宿主微生物或细胞可为原核细胞(例如真细菌界的微生物)或真核细胞。如本领域技术人员所将了解,原核细胞缺乏膜束缚的核,而真核细胞具有膜束缚的核。

[0042] 术语“分离的”或“纯净的”是指物质大致上(例如大于50%或大于75%或基本上例如大于90%、95%、98%或99%)不含在它的天然状态下通常伴随它的组分,所述天然状态例如是它天然所见的状态或当它最初被产生时它所存在的状态。

[0043] 视结构、pH和存在的离子而定,如本文所述的羧酸可为盐、酸、碱或衍生物。术语“丙二酸(malonate)”和“丙二酸(malonic acid)”在本文中可互换使用。丙二酸也称为丙二酸($C_3H_4O_4$;CAS号141-82-2)。

[0044] 如本文所用的术语“丙二酸衍生化合物”是指丙二酸单烷基酯,包括例如且不加限制地丙二酸单甲酯(也称为丙二酸单甲酯,CAS号16695-14-0)、丙二酸单乙酯(也称为丙二酸单乙酯,CAS号1071-46-1)、丙二酸单丙酯、丙二酸单丁酯、丙二酸单叔丁酯(CAS号40052-13-9)等;丙二酸二烷基酯,例如且不加限制地丙二酸二甲酯(CAS号108-59-8)、丙二酸二乙

酯(CAS号105-53-3)、丙二酸二丙酯(CAS号1117-19-7)、丙二酸二丁酯(CAS号1190-39-2)等以及梅尔德伦氏酸(CAS号2033-24-1)。丙二酸衍生化合物可自丙二酸合成产生且它们自身是有价值化合物,但也是许多其它有价值化合物的化学合成中的有用基质。

[0045] 如本文所用,术语“核酸”及其变化形式将为聚脱氧核糖核苷酸(含有2-脱氧-D-核糖)和聚核糖核苷酸(含有D-核糖)所通用。“核酸”也可指作为嘌呤或嘧啶碱基的N-糖苷的任何其它类型的聚核苷酸以及含有非核苷酸骨架的其它聚合物,前提是聚合物含有呈允许如DNA和RNA中所见的碱基配对和碱基堆积的构型的核苷碱基。如本文所用,核苷酸和聚核苷酸的符号是由IUPAC-IUB生物化学命名委员会(IUPAC-IUB Commission of Biochemical Nomenclature)(Biochem.9:4022,1970)推荐的符号。“核酸”在本文中也可关于它的序列(即不同核苷酸在核酸中出现的次序)加以提及,因为核酸中的核苷酸序列通常限定它的生物活性,例如如在编码区域的序列中,即基因中主要由启动子和编码区域组成的核酸,所述编码区域编码基因的产物,其可为RNA(例如rRNA、tRNA或mRNA)或蛋白质(当基因编码蛋白质时,所述mRNA与所述蛋白质两者均是那个基因的“基因产物”)。

[0046] 术语“可操作地连接”是指核酸表达控制序列(如启动子、核糖体结合位点和转录终止子)与第二核酸序列、编码序列或编码区域之间的功能性键联,其中所述表达控制序列引导或另外调控所述编码序列的转录和/或翻译。

[0047] 如本文所用的术语“任选的”或“任选地”是指随后描述的特征或结构可或可不存在,或随后描述的事件或事项可或可不发生,且描述包括特定特征或结构存在的情况和所述特征或结构不存在的情况,或事件或事项发生的情况以及它不发生的情况。

[0048] 如本文所用,“重组”是指通过人介入来改变遗传物质。通常,重组是指通过分子生物学(重组DNA技术)方法,包括克隆和重组来操作细胞或病毒或表达载体中的DNA或RNA。重组也可指通过随机或定向突变诱发来操作细胞或病毒中的DNA或RNA。“重组”细胞或核酸可通常关于它如何不同于天然存在的对应物(“野生型”)加以描述。此外,对已“工程化”或“修饰”的细胞或核酸以及那些术语的变化形式的任何提及都意图指代重组细胞或核酸。

[0049] 如本文所用,术语“转录因子生物传感器”是指用以通过活化“标记”或“报道体”基因的表达来检测例如丙二酸的物质的系统,其中报道基因表达是由能够在结合例如丙二酸的那个物质后结合启动子且活化转录的转录因子介导。举例来说,丙二酸可结合转录因子(例如MdcY)且活化自启动子(例如P_{MdcL})的转录。“丙二酸转录因子”是在结合于丙二酸时可活化启动子的转录因子。因此,MdcY是丙二酸转录因子。

[0050] 如本文所用的术语“转导”、“转化”、“转染”及其变化形式是指将一种或多种核酸引入细胞中。出于实际目的,对于待称为“被转导”、“被转化”或“被转染”细胞的细胞而言,核酸必须由细胞稳定维持或复制足以使得它编码的功能或产物能够表达的时期。如将为本领域技术人员所了解,核酸的稳定维持或复制可通过将核酸的序列并入细胞染色体DNA(例如基因组)中(如通过染色体整合而发生),或通过染色体外复制(如用自由复制质粒来发生)来进行。当病毒具有“感染性”时,它可被稳定维持或复制:当它转导宿主微生物时,复制且(不利用任何互补病毒或载体)向其它微生物散布子代表达载体,例如与原始转导表达载体具有相同类型的病毒,其中所述子代表达载体具有相同繁殖能力。

[0051] 章节2:丙二酰基-辅酶A水解酶

[0052] 根据本发明的一个方面,丙二酸是通过丙二酰基-辅酶A水解酶催化丙二酰基-辅

酶A转化成丙二酸的作用产生。迄今为止,尚未鉴定野生型丙二酰基-辅酶A水解酶基因,但在非工程化菌株的发酵培养基中存在少量丙二酸指示可能存在具有这个活性的野生型酶。本发明提供作为野生型基因的对应物的已被突变来赋予丙二酰基-辅酶A水解酶活性的许多基因。制备丙二酰基-辅酶A水解酶的宿主细胞是重组宿主细胞;在许多实施方案中,宿主细胞已被遗传修饰来包含编码催化丙二酰基-辅酶A水解成丙二酸的丙二酰基-辅酶A水解酶的异源核酸。在一些实施方案中,重组宿主细胞是真核生物。在各种实施方案中,真核生物是选自以下非限制性实例属的酵母菌株:假丝酵母属、隐球酵母属、汉逊酵母属、伊萨酵母属、克鲁维酵母属、Komagataella、油脂酵母属、毕赤酵母属、红冬孢酵母属、红酵母属、酵母属或耶罗威亚酵母属。本领域技术人员将认识到这些属广泛涵盖酵母,包括区分为油脂酵母的酵母。在一些实施方案中,宿主细胞是酿酒酵母。在其它实施方案中,宿主细胞是库德里阿兹威毕赤酵母(*Pichia kudriavzevii*)。在本发明的其它实施方案中,真核宿主细胞是真菌或藻类。在其它实施方案中,重组宿主细胞是选自以下非限制实例属的原核生物:芽孢杆菌属、梭菌属、棒杆菌属、埃希氏菌属、假单胞菌属、如红杆菌属(*Rhodobacter*)和链霉菌属。在一些实施方案中,宿主细胞是大肠杆菌。

[0053] 本发明部分地起因于发现各种酰基-辅酶A水解酶和转酰基酶可被工程化来具有丙二酰基-辅酶A水解酶活性且因此用于生物产生丙二酸。适于修饰以进行丙二酰基-辅酶A水解的酰基-辅酶A水解酶的非限制性实例包括来自以下组成的组的那些酰基-辅酶A水解酶的任一者:3-羟基异丁酰基-辅酶A水解酶(EC 3.1.2.4)、3-羟基丙酰基-辅酶A水解酶(EC 3.1.2.4)、乙酰乙酰基-辅酶A水解酶(EC 3.1.2.11)、甲基丙二酰基-辅酶A水解酶(EC 3.1.2.17)、丙酰基-辅酶A水解酶(EC 3.1.2.18)、丁二酰基-辅酶A水解酶(EC 3.1.2.3)和如本文提供的被突变来具有丙二酰基-辅酶A水解酶活性的丙二酰基-辅酶A:ACP转酰基酶(EC 2.3.1.39)。

[0054] 在一些实施方案中,用于根据本发明产生丙二酸的丙二酰基-辅酶A水解酶是突变的酿酒酵母EHD3酰基-辅酶A水解酶(对于野生型EHD3氨基酸序列,参见SEQ ID NO:1)。一种具有改变的底物特异性的所述突变体是E124V突变体(参见Rouhier, "Characterization of YDR036C from *Saccharomyces cerevisiae*." Dissertation, Miami University, Miami University and OhioLINK (2011)), 尽管先前已报道,但未报道其具有丙二酰基-辅酶A水解酶活性。在本发明的一些实施方案中,表达E124V突变体的大肠杆菌宿主细胞用于产生丙二酸,其接着自细胞或发酵液纯化。在本发明的其它实施方案中,表达E124V突变体的酵母细胞用于根据本发明产生丙二酸。在另一实施方案中,表达E124V突变体的油脂酵母细胞用于根据本发明产生丙二酸。

[0055] 用以产生EHD3的E124A突变体的先前努力导致细胞在诱导蛋白质自pET28a表达载体表达后死亡;所述蛋白质不能被纯化(参见Rouhier, 上文)。本发明提供E124A突变体的可用于大肠杆菌宿主细胞中,从而致使它们能够产生丙二酸的表达载体。这些大肠杆菌表达载体的特征在于相对于Rouhier的pET28a载体,E124A突变体是在较低无毒水平下产生。这是例如通过采用启动子的拷贝数低于由Rouhier使用的启动子的拷贝数或启动子弱于由Rouhier使用的启动子的表达载体来实现。本领域技术人员也将了解翻译可通过核糖体结合位点(RBS)或Kozak序列对核糖体的亲和力来调节。因此,较弱RBS或Kozak序列也可用于降低基因表达。拷贝数较低的表达载体的实例包括但不限于pSC101起点表达载体、p15a起

点表达载体和整合至染色体DNA中的表达载体。弱于由Rouhier使用的T7启动子的启动子的实例包括但不限于 P_{LacO1} 、 P_{TRC} 和 P_{BAD} 启动子。在一些实施方案中,载体具有pSC101复制起点。在其它实施方案中,用于表达EHD3E124A突变体编码序列的启动子是 P_{LacO1} 启动子。另外,本发明提供用于酵母宿主细胞的编码E124A突变体的表达的载体。本发明的遗传修饰的酿酒酵母EHD3 E124A表达载体可体内用于在大肠杆菌和酿酒酵母中产生丙二酸,且本发明的方法提供用于随后自这些菌株的发酵液纯化丙二酸以及使丙二酸合成转化成衍生小分子化合物的手段。

[0056] 本发明也提供EHD3的用作丙二酰基-辅酶A水解酶的E124S突变体、用于表达这个突变体的载体、以及表达这个突变体且产生丙二酸的宿主细胞(参见实施例31)。野生型酿酒酵母EHD3催化3-羟基丙酰基-辅酶A (3HPA-CoA) 和3-羟基异丁酰基-辅酶A (3HIBA-CoA) 的水解,且尽管本发明不受理论限制,但预测E124会与3HPA-辅酶A上的末端羟基部分相互作用,从而使底物在EHD3活性位点中稳定(参见Rouhier,上文)。本发明的某些方面起因于发现特定E124点突变使产生丙二酸的丙二酰基-辅酶A的酶水解增加。E124突变成亲核氨基酸(例如S或T)、碱性氨基酸(例如H、K或R)或酰胺氨基酸(例如N或Q)会使丙二酰基-辅酶A在EHD3活性位点中的结合改进超过3-羟基丙酰基-辅酶A且使丙二酸产生增加(相对于未突变对应酶)。由于内源性宿主细胞酰基-辅酶A分子的水解降低,E124S、E124T、E124N、E124Q、E124H、E124K和E124R突变也降低副产物(例如乙酸盐、丙酸盐、异丁酸盐和丁二酸盐)的产生。E124S点突变将羟基部分放置在促进丝氨酸残基与丙二酰基-辅酶A的末端羧酸根基团之间的氢键结的位置中。E124Q点突变将谷氨酰胺酰胺基团放置在接近丙二酰基-辅酶A的末端羧酸根基团的位置中。E124K点突变将赖氨酸胺基团放置在促进赖氨酸残基与丙二酰基-辅酶A的末端羧酸根基团之间的氢键结的位置中。不同于上述亲核、酰胺和碱性E124点突变,突变E124A和E124V移除在位置124处存在的带电荷氨基酸;这些突变消除丙二酸上的末端羧酸根与EHD3 124氨基酸侧链之间的氢键结且打开关于混杂活性的EHD3活性位点,从而增加不合需要副产物形成且降低丙二酸产生。

[0057] 在本发明的一些实施方案中,表达E124S突变体的大肠杆菌宿主细胞用于产生丙二酸。在本发明的其它实施方案中,表达E124S突变体的酵母宿主细胞用于产生丙二酸。在其它实施方案中,表达E124S突变体的油脂酵母宿主细胞用于产生丙二酸。在本发明的一些实施方案中,表达E124Q突变体的大肠杆菌宿主细胞用于产生丙二酸。在本发明的其它实施方案中,表达E124Q突变体的酵母宿主细胞用于产生丙二酸。在其它实施方案中,表达E124Q突变体的油脂酵母宿主细胞用于产生丙二酸。在本发明的一些实施方案中,表达E124K突变体的大肠杆菌宿主细胞用于产生丙二酸。在本发明的其它实施方案中,表达E124K突变体的酵母宿主细胞用于产生丙二酸。在其它实施方案中,表达E124K突变体的油脂酵母宿主细胞用于产生丙二酸。在本发明的一些实施方案中,表达E124H突变体的大肠杆菌宿主细胞用于产生丙二酸。在本发明的其它实施方案中,表达E124H突变体的酵母宿主细胞用于产生丙二酸。在本发明的其它实施方案中,表达E124H突变体的油脂酵母宿主细胞用于产生丙二酸。在本发明的一些实施方案中,表达E124R突变体的大肠杆菌宿主细胞用于产生丙二酸。在本发明的其它实施方案中,表达E124R突变体的酵母宿主细胞用于产生丙二酸。在本发明的其它实施方案中,表达E124R突变体的油脂酵母宿主细胞用于产生丙二酸。在其它实施方案中,表达EHD3 E124亲核氨基酸点突变(即E124S或E124T)的重组宿主细胞用于产生丙二酸。

在其它实施方案中,表达EHD3 E124碱性氨基酸点突变(即E124H、E124K或E124R)的重组宿主细胞用于产生丙二酸。在其它实施方案中,表达EHD3 E124酰胺氨基酸点突变(即E124N或E124Q)的重组宿主细胞用于产生丙二酸。

[0058] 本发明也提供一种包含突变活性位点的突变EHD3、用于表达所述突变体的载体、以及表达所述突变体且产生丙二酸的宿主细胞。本发明的某些方面部分地起因于发现特定氨基酸(即F121和F177)对于酰基-辅酶A底物结合是重要的,且引入特定点突变使丙二酰基-辅酶A水解和丙二酸产生增加。引入突变F121I或F121L会增加丙二酰基-辅酶A进入活性位点。类似地,引入突变F177I或F177L会增加丙二酰基-辅酶A进入活性位点。在氨基酸位置F121或F177处的一个或多个点突变可单独或连同E124点突变一起引入。在各种实施方案中,F121和/或F177点突变是连同E124点突变一起引入。在一些实施方案中,表达EHD3 F121I或F121L突变体的重组宿主细胞用于产生丙二酸。在其它实施方案中,表达EHD3 F177I或F178L突变体的重组宿主细胞用于产生丙二酸。在这些实施方案中,重组宿主细胞可不加限制地为大肠杆菌或酵母(包括但不限于酿酒酵母或其它酵母)宿主细胞。

[0059] 本发明也提供包含突变线粒体靶向性序列的突变EHD3、用于表达所述突变体的载体、以及表达所述突变体且产生丙二酸的宿主细胞。在酿酒酵母宿主中,野生型EHD3定位在线粒体中。丙二酰基-辅酶A见于线粒体与胞质液两者中;EHD3催化胞质液丙二酰基-辅酶A水解需要使EHD3定位于胞质液。本发明的某些方面起因于发现EHD3线粒体靶向性序列的突变可使丙二酸产生增加。对线粒体靶向性重要的EHD3氨基酸包括R3、K7、K14、K18和R22,且使这些碱性氨基酸的一者或多者突变成疏水性氨基酸(即A或V)会消除线粒体靶向性。在一些实施方案中,包含由在选自由R3、K7、K14、K18和R22组成的组的氨基酸处的一个或多个A或V突变组成的EHD3的重组宿主用于产生丙二酸。在一些实施方案中,重组宿主是酵母菌株。在其它实施方案中,宿主是酿酒酵母。在其它实施方案中,重组宿主细胞含有线粒体靶向性序列未改变的EHD3(即野生型)的一个或多个拷贝以及线粒体靶向性序列突变的EHD3的一个或多个拷贝。用于本发明的这个方面的线粒体靶向性序列的其它实例是:WT COX4、SynA1、SynA2、Syn B1和Syn B2,如由Allison和Schatz (Allison和Schatz (1986) PNAS 83: 9011-9015) 对于其它应用所概述。在本发明的其它实施方案中,使最常含有Ser-Lys-Leu基序的过氧化物酶体靶向性信号(例如PTS1、PTS2)融合于丙二酰基-辅酶A水解酶的C末端以使这个蛋白质定位于过氧化物酶体。

[0060] 因此,在本发明的一个方面,重组宿主细胞包含编码突变酿酒酵母EHD3的异源核酸,所述突变酿酒酵母EHD3导致丙二酸产生相对于不包含所述突变EHD3的宿主细胞得以增加。在一些实施方案中,突变EHD3在大肠杆菌中异源表达。在其它实施方案中,突变EHD3在酿酒酵母中异源表达。在其它实施方案中,突变EHD3在油脂酵母细胞中异源表达。在一些实施方案中,突变EHD3含有在位置E124处的点突变。在一些实施方案中,在残基E124处的点突变是E124A或E124V。在一些实施方案中,在E124处的点突变是E124S或E124T。在一些实施方案中,在E124处的点突变是E124S。在一些实施方案中,在E124处的点突变是选自由E124H、E124K和E124R组成的组的碱性氨基酸。在一些实施方案中,在E124处的点突变是E124H。

[0061] 在一些实施方案中,在E124处的点突变是E124K。在一些实施方案中,在E124处的点突变是E124R。在一些实施方案中,在残基E124处的点突变是E124N或E124Q。在一些实施方案中,在残基E124处的点突变是E124Q。在一些实施方案中,使一个或多个选自由F121和

F177组成的组的EHD3氨基酸突变成I或L。在一些实施方案中,使一个或多个选自R3、K7、K14、K18和R22组成的组的EHD3氨基酸突变成A或V。

[0062] 在本发明的另一方面,不同于EHD3或除EHD3之外的酶也用作丙二酰基-辅酶A水解酶以根据本发明产生丙二酸。在一些实施方案中,流感嗜血杆菌 (*Haemophilus influenzae*) YciA在异源宿主中异源表达以根据本发明产生丙二酸(参见Zhuang等 *Biochemistry* 47:2789-2796(2008))。在其它实施方案中,丙二酰基-辅酶A水解酶是褐鼠 (*Rattus norvegicus*) 内源性酰基-辅酶A水解酶(参见Kovachy等, *J. Biol. Chem.* 258: 11415-11421(1983))。在其它实施方案中,丙二酰基-辅酶A水解酶是来自金色大鼠 (*Mesocricetus auratus*) 的棕色脂肪组织线粒体蛋白质部分的酰基-辅酶A水解酶(参见Alexson等, *J. Biol. Chem.* 263:13564-13571(1988))。

[0063] 因此,根据本发明,不同于EHD3或除EHD3之外的酰基-辅酶A水解酶(来自酿酒酵母或来自其它生物体的同源酶)也可用于在重组宿主中生物合成丙二酸。在一些实施方案中,重组宿主是酿酒酵母。在其它实施方案中,重组宿主是大肠杆菌。在其它实施方案中,重组宿主是不同于酿酒酵母的如以下另外详述的酵母。在各种实施方案中,宿主被修饰来表达选自自由以下组成的组的突变酶:白色假丝酵母 (*S. albicans*) EHD3、智人 (*H. sapiens*) HIBCH (UniProt:Q6NVY1)、拟南芥 (*A. thaliana*) CHY1 (UniProt:Q9LKJ1)、褐家鼠 (*R. norvegicus*) HIBC H (UniProt:Q5XIE6)、小家鼠 (*M. musculus*) HIBCH (UniProt:Q8QZS1)、原鸡 (*G. gallus*) HIBCH (UniProt:Q5ZJ60)、家牛 (*B. taurus*) HIBCH (Uni Prot:Q2HJ73)、斑马鱼 (*D. rerio*) HIBCH (UniProt:Q58EB4)、芽孢杆菌 (*B. cereus*) Bch、铜绿假单胞菌 (*P. aeruginosa*) Hich、大肠杆菌YciA、流感嗜血杆菌 (*H. influenzae*) YciA、小家鼠ACOT4、小家鼠ACOT8、肠道沙门菌 (*S. enterica*) SARI_01218、嗜热泉生古细菌 (*A. pernix*) K1、哈氏嗜纤维菌 (*C. hutchinsonii*) Chut02003666、硫磺矿硫化叶菌 (*S. sol fataricus*) P2 SS02287、酸热硫化叶菌 (*S. acidocaldarius*) DSM 639 Saci_0145、喜气热棒菌 (*P. aerophilum*) str. IM2 PAE3404、黑腹果蝇 (*D. melanogaster*) CG1635、甲醇降解菌 (*P. carbinolicus*) DSM 2380 Pcar_1366、厌氧性粘菌 (*A. dehalogenans*) 2CP-C 110、原鸡ACOT9和非洲爪蟾 (*X. laevis*) MGC114623。

[0064] 本领域技术人员将了解一种或多种适合突变的酰基-辅酶A水解酶可根据本发明用于在宿主细胞中使丙二酰基-辅酶A转化成丙二酸。此外,不同于本文明确公开的酰基-辅酶A水解酶的酰基-辅酶A水解酶可以突变或异源表达形式加以利用,且将为本领域技术人员鉴于本公开充分了解的是可如何鉴定、修饰和表达其它适当酶以获得如本文公开的所需丙二酰基-辅酶A水解酶活性。

[0065] 共有序列

[0066] 本发明的丙二酰基-辅酶A水解酶包括与由本发明提供的共有序列同源的丙二酰基-辅酶A水解酶。如上所指示,大致上与本文明确所述的酶同源的任何酶都可用于本发明的宿主细胞中。当一种酶展现相同目标活性时,它与另一酶(“参照酶”)同源,且可用于大致上类似目的。通常,同源酶共有实质性序列同一性。各组同源酶通常具有一个或多个在共有序列蛋白质类别的所有成员之间保守的特定氨基酸。

[0067] 酶相对于共有序列的序列同一性百分比是通过针对共有序列比对酶序列来确定。本领域技术人员将认识到各种序列比对算法适于将酶与共有序列比对。参见例如

Needleman,SB等“A general method applicable to the search for similarities in the amino acid sequence of two proteins.”Journal of Molecular Biology 48(3): 443-53 (1970)。在相对于共有序列比对酶序列之后,酶具有由共有序列中的相同位置描述的某一氨基酸(或划线)所处的位置的百分比决定序列同一性百分比。当简并氨基酸(即B、J、O、U、“+”)存在于共有序列中时,由简并氨基酸描述的任何氨基酸都可存在于酶中的酶比对位置处以与共有序列在比对位置处相同。当划线存在于共有序列中时,酶不必具有氨基酸存在于酶比对位置处以与共有序列在比对位置处相同。

[0068] 本发明提供用于鉴定和构建本发明的丙二酰基-辅酶A水解酶的共有序列。在各种实施方案中,这些丙二酰基-辅酶A水解酶共有序列含有据信为形成氧离子孔所必需(但本发明不受任何作用机制理论限制)的活性位点氨基酸残基,所述氧离子孔负责使由丙二酰基-辅酶A底物产生的烯醇阴离子中间体以及如下所述的对丙二酰基-辅酶A结合重要的氨基酸残基稳定。就水解丙二酰基-辅酶A的能力而言,由本文提供的共有序列涵盖的丙二酰基-辅酶A水解酶具有的酶活性与本文例示的一种酶的酶活性相同、或基本上相同、或至少大致上类似。丙二酰基-辅酶A水解酶可见于自然界中,或更通常,是根据本发明加以修饰来具有丙二酰基-辅酶A水解酶活性的野生型酶的工程化突变体。丙二酰基-辅酶A水解酶可被鉴定或通过使另一酶的序列突变以产生由本文共有序列涵盖的序列来由另一酶构建;如果酶与本文共有序列共有实质性同源性,但具有次优(包括不具有)丙二酰基-辅酶A水解酶活性,那么根据本发明,使它突变来符合本文提供的共有序列以提供本发明的丙二酰基-辅酶A水解酶。

[0069] 本发明提供四个丙二酰基-辅酶A水解酶共有序列:(i)基于EHD3 EC 3.1.2.4的丙二酰基-辅酶A水解酶(SEQ ID NO:7), (ii)基于芽孢杆菌属EC 3.1.2.4的丙二酰基-辅酶A水解酶(SEQ ID NO:8), (iii)基于假单胞菌属EC 3.1.2.4的丙二酰基-辅酶A水解酶(SEQ ID NO:9), 以及(iv)基于芽孢杆菌属EC 3.1.2.4与假单胞菌属EC 3.1.2.4两者的丙二酰基-辅酶A水解酶(SEQ ID NO:10)。共有序列提供各位置标识氨基酸(如果特定氨基酸被鉴定)或最可能见于那个类别的丙二酰基-辅酶A水解酶中的指定位置处的一子组氨基酸(如果位置被鉴定为可变)的氨基酸序列。本领域技术人员将认识到这些共有序列中的固定氨基酸和保守氨基酸与共有序列所基于的野生型序列相同(在固定氨基酸的情况下)或一致(在保守氨基酸的情况下)。共有序列中的划线指示适于根据本发明进行突变的酶可见于自然界中,所述酶在序列中的划线位置处可具有另一氨基酸,但通常无氨基酸存在于划线位置处。

[0070] 基于EHD3 EC 3.1.2.4酶的丙二酰基-辅酶A水解酶共有序列

[0071] 本发明提供一种基于EHD3 EC 3.1.2.4酶的丙二酰基-辅酶A水解酶共有序列(SEQ ID NO:7), 且在各种实施方案中,适合用于本发明的方法中的丙二酰基-辅酶A水解酶与这个丙二酰基-辅酶A水解酶共有序列具有至少63%同一性。在各种实施方案中,适于根据本发明的方法使关键谷氨酸残基突变成X₁以赋予丙二酰基-辅酶A水解酶活性的酶与SEQ ID NO:7具有65%、70%、80%、90%或95%或更高的同一性。与这个共有序列具有显著同源性的蛋白质包括UniProt ID:C5DE94 (63%同一性)、UniProt ID:Q6CJH2 (64%同一性)、UniProt ID:G2WAE2 (66%同一性)、UniProt ID:J8Q6P9 (66%同一性)、UniProt ID:G8C0H0 (68%同一性)、UniProt ID:C5DX08 (68%同一性)、UniProt ID:P28817 (69%同一性)、

UniProt ID:A7TTD5 (69%同一性)、UniProt ID:J7S9J9 (70%同一性)、UniProt ID:Q6FM09 (71%同一性)、UniProt ID:I2H4L2 (71%同一性)、UniProt ID:H2AME2 (73%同一性)、UniProt ID:G8ZTJ4 (77%同一性)、UniProt ID:G0W4I8 (77%同一性)、UniProt ID:G0V818 (78%同一性)和UniProt ID:J5S5X3 (79%同一性)。在一些实施方案中,与共有序列SEQ ID NO:7等同或具有大于63%同一性的丙二酰基-辅酶A水解酶在重组宿主细胞中表达且用于根据本发明产生丙二酸。

[0072] 在与这个共有序列(SEQ ID NO:7)同源的突变和野生型酶中,高度保守的氨基酸是V101、R110、L114、R116、K119、L120、N121、A122、L123、L135、E137、Y138、K140、S141、S151、R156、C159、G161、G162、D163、V164、A168、F185、E188、Y189、S190、N192、A196、T197、K200、M206、G208、I209、T210、M211、G212、G213、G214、V215、G216、H220、P222、F223、R224、T227、E228、T230、M234、P235、E236、D238、I239、G240、F242、P243、D244、V245、F249、P252、Q263、Y267、L268、T271、G272、G277、G284、S287、H288、Y289、L298、R301、L302、E304、E333、F334、L352、V354、I355、F359、L374、F391、L399、K402、S403、S406、N417、D429、L430、T432、A433、E449、F450、K457、L458、K461、W468、L494、T502、Y506、P507、L514、P515和K561。在各种实施方案中,与这个共有序列(SEQ ID NO:7)同源的丙二酰基-辅酶A水解酶含有这些保守氨基酸的至少25%,常为这些保守氨基酸的大多数(大于50%),且有时是这些保守氨基酸的全部。

[0073] 这个共有序列(SEQ ID NO:7)中的一些氨基酸为活性所必需且在所述类别的所有成员之间是保守的。由基于EHD3 EC 3.1.2.4的共有序列涵盖的丙二酰基-辅酶A水解酶含有对水解酶活性重要的六个活性位点残基:(i)共有序列中的三个活性位点氨基酸残基(G161、G162、G213)据信为形成氧离子孔所必需(但本发明不受任何作用机制理论限制),所述氧离子孔负责使由丙二酰基-辅酶A底物产生的烯醇阴离子中间体稳定;(ii)共有序列的两个氨基酸残基(E236、D244)为酰基-辅酶A水解所必需;且(iii)在(SEQ ID NO:7)的位置188处的氨基酸残基据信为丙二酰基-辅酶A底物结合所必需。然后,在这六个残基中,五个存在于共有序列(SEQ ID NO:7)中和由那个序列涵盖的所有丙二酰基-辅酶A水解酶中,且在位置188处的第六个残基(共有序列中的氨基酸X₁)选自由极性或带正电荷氨基酸(R、H、K、S、T、N、Q、Y)以及A和D组成的组以提供本发明的能够在重组宿主细胞中产生丙二酸的丙二酰基-辅酶A水解酶。来自共有序列的六个必需残基(G161、G162、G213、E236、D244、X₁188)分别对应于用于在实施例31中说明本发明的酿酒酵母EHD3中的G99、G100、G149、E172、D180和E124(通常突变成X₁)。

[0074] 基于芽孢杆菌属EC 3.1.2.4酶的丙二酰基-辅酶A水解酶共有序列

[0075] 本发明提供一种基于芽孢杆菌属EC 3.1.2.4酶的丙二酰基-辅酶A水解酶共有序列(SEQ ID NO:8),且在各种实施方案中,适合用于本发明的方法中的丙二酰基-辅酶A水解酶与这个丙二酰基-辅酶A水解酶共有序列具有至少86%同一性。在各种实施方案中,适于根据本发明的方法使关键谷氨酸残基突变成X₁以赋予丙二酰基-辅酶A水解酶活性的酶与SEQ ID NO:8具有90%或95%或更高的同一性。与这个共有序列具有显著同源性的蛋白质包括UniProt ID:C2TX63 (92%同一性)、UniProt ID:C2UV40 (91%同一性)、UniProt ID:C2QBT2 (93%同一性)、UniProt ID:C2XTU0 (93%同一性)、UniProt ID:C2PVQ0 (93%同一性)、UniProt ID:C3A5N3 (93%同一性)、UniProt ID:C2SJV4 (93%同一性)、UniProt ID:

C2Z7U1 (92%同一性)、UniProt ID:C2VTI4 (97%同一性)、UniProt ID:B3Z9Y3 (97%同一性)、UniProt ID:B7JNH7 (97%同一性)、UniProt ID:Q63BK8 (97%同一性)、UniProt ID:B0Q3Q4 (97%同一性)、UniProt ID:B0AQX0 (97%同一性)、UniProt ID:B3YSW2 (97%同一性)、UniProt ID:C2NHG5 (97%同一性)、UniProt ID:B3ZIZ8 (97%同一性)、UniProt ID:C2QSV2 (97%同一性)、UniProt ID:C3C255 (97%同一性)、UniProt ID:B5UZZ1 (96%同一性)、UniProt ID:C2MKL7 (95%同一性)、UniProt ID:B9IZZ9 (95%同一性)、UniProt ID:F0PNG8 (95%同一性)、UniProt ID:Q738L0 (97%同一性)、UniProt ID:C2PEV7 (95%同一性)、UniProt ID:C2YRH7 (96%同一性)、UniProt ID:Q4MU30 (95%同一性)、UniProt ID:Q81DR3 (96%同一性)、UniProt ID:C2W7W8 (89%同一性) 和UniProt ID:A7GPH6 (86%同一性)。在各种实施方案中,与共有序列SEQ ID NO:8等同或具有大于86%同一性的丙二酰基-辅酶A水解酶在重组宿主细胞中表达且用于根据本发明产生丙二酸。在序列表中提供在根据本发明突变以赋予丙二酰基-辅酶A水解酶活性的关键谷氨酸残基的位置处含有X₁的B9IZZ9 (SEQ ID NO:46)、C3ALI3 (SEQ ID NO:47)、F0PNG8 (SEQ ID NO:49)、Q63BK8 (SEQ ID NO:51) 和Q81DR3 (SEQ ID NO:52) 的序列。

[0076] 在与这个共有序列 (SEQ ID NO:8) 同源的突变和野生型酶中,高度保守的氨基酸是M1、T2、E3、V5、L6、F7、S8、G13、V14、A15、I17、T18、L19、N20、R21、P22、K23、A24、L25、N26、S27、L28、S29、Y30、M32、L33、I36、G37、K39、L40、K41、E42、W43、E44、I49、I52、V53、L54、K55、G56、A57、G58、K60、G61、F62、C63、A64、G65、G66、D67、I68、K69、T70、L71、Y72、E73、A74、R75、S76、N77、E78、A80、L81、Q82、A84、E85、F87、F88、E90、E91、Y92、I94、D95、T96、Y99、Y101、K103、P104、I105、I106、A107、C108、L109、D110、G111、I112、V113、M114、G115、G116、G117、V118、G119、L120、T121、N122、G123、A124、R127、I128、V129、T130、T133、K134、W135、A136、M137、P138、E139、M140、N141、I142、G143、F144、F145、P146、D147、V148、G149、A150、A151、Y152、F153、L154、N155、A157、P158、G159、G162、V165、A166、L167、A169、L172、K173、A174、D176、V177、L178、I180、A182、A183、D184、L192、F195、L196、W204、V210、L214、K215、L231、E236、H241、F242、E248、I250、I251、S253、L254、E255、F261、L269、L270、S271、K272、S273、P274、S276、L277、K278、V279、T280、L281、K282、Q283、G287、K290、S291、E293、C295、F296、A297、T298、D299、L300、L302、A303、K304、N305、F306、M307、R308、H309、D311、F312、F313、E314、G315、V316、R317、S318、V320、D322、K323、D324、Q325、N326、P327、Y329、K330、Y331、D336、V337、V342、N343、F345、F346、L348和L349。在各种实施方案中,与这个共有序列 (SEQ ID NO:8) 同源的丙二酰基-辅酶A水解酶含有这些保守氨基酸的至少25%,常为这些保守氨基酸的大多数 (大于50%),且有时是这些保守氨基酸的全部。

[0077] 这个共有序列 (SEQ ID NO:8) 中的一些氨基酸为活性所必需且在所述类别的所有成员之间是保守的。由基于芽孢杆菌属EC 3.1.2.4的共有序列涵盖的丙二酰基-辅酶A水解酶含有对水解酶活性重要的六个活性位点残基:(i) 共有序列中的三个活性位点氨基酸残基 (G65、G66、G116) 据信为形成氧离子孔所必需 (但本发明不受任何作用机制理论限制),所述氧离子孔负责使由丙二酰基-辅酶A底物产生的烯醇阴离子中间体稳定;(ii) 共有序列的两个氨基酸残基 (E139、D147) 为酰基-辅酶A水解所必需;且(iii) (SEQ ID NO:8的) 突变氨基酸 (X₁91) 据信为丙二酰基-辅酶A底物结合所必需。然后,在这六个残基中,五个存在于共有序列 (SEQ ID NO:8) 中和由那个序列涵盖的所有丙二酰基-辅酶A水解酶中,且第六个残

基X₁91为提供本发明的能够在重组宿主细胞中产生丙二酸的丙二酰基-辅酶A水解酶所必需。来自共有序列的六个必需残基(G65、G66、G116、E139、D147、X₁91)分别对应于用于在实施例31中说明本发明的苏云金杆菌幕虫亚种(*Bacillus thuringiensis* subsp.finitimus)(菌株YBT-020)F0PNG8中的G65、G66、G116、E139、D147和E91(通常突变成X₁)(参见含有突变E91S的SEQ ID NO:49)。

[0078] 与共有序列(SEQ ID NO:8)同源且由本发明提供的克隆或合成核酸编码的适于丙二酰基-辅酶A水解的酶的非限制性实例包括含有至少一个通过由以下组成的组的突变酶说明的突变的突变酶:蜡样芽胞杆菌(*Bacillus cereus*)(菌株Q1)B9IZZ9(E91S)、B9IZZ9(E91A)、B9IZZ9(E91H)、B9IZZ9(E91K)、B9IZZ9(E91R)、B9IZZ9(E91Q)、B9IZZ9(E91T)、B9IZZ9(E91N)、B9IZZ9(E91Y)、B9IZZ9(E91D);苏云金杆菌幕虫亚种(菌株YBT-020)F0PNG8(E91S)、F0PNG8(E91A)、F0PNG8(E91H)、F0PNG8(E91K)、F0PNG8(E91R)、F0PNG8(E91Q)、F0PNG8(E91T)、F0PNG8(E91N)、F0PNG8(E91Y)、F0PNG8(E91D);蜡样芽胞杆菌(菌株ATCC 14579/DSM 31)Q81DR3、Q81DR3(E91S)、Q81DR3(E91A)、Q81DR3(E91H)、Q81DR3(E91K)、Q81DR3(E91R)、Q81DR3(E91Q)、Q81DR3(E91T)、Q81DR3(E91N)、Q81DR3(E91Y)、Q81DR3(E91D);蜡样芽胞杆菌(菌株ZK/E33L)Q63BK8、Q63BK8(E91S)、Q63BK8(E91A)、Q63BK8(E91H)、Q63BK8(E91K)、Q63BK8(E91R)、Q63BK8(E91Q)、Q63BK8(E91T)、Q63BK8(E91N)、Q63BK8(E91Y)、Q63BK8(E91D)。

[0079] 基于假单胞菌属EC 3.1.2.4酶的丙二酰基-辅酶A水解酶共有序列

[0080] 本发明提供一种基于假单胞菌属EC 3.1.2.4酶的丙二酰基-辅酶A水解酶共有序列(SEQ ID NO:9),且在各种实施方案中,适合用于本发明的方法中的丙二酰基-辅酶A水解酶与这个丙二酰基-辅酶A水解酶共有序列具有至少75%同一性。在各种实施方案中,适于根据本发明使关键谷氨酸残基突变成X₁以赋予丙二酰基-辅酶A水解酶活性的酶与SEQ ID NO:9具有80%、90%或95%或更高的同一性。与这个共有序列具有显著同源性的蛋白质包括:UniProt ID:F5KBQ4(80%同一性)、UniProt ID:A6VAN3(81%同一性)、UniProt ID:A4XS22(81%同一性)、UniProt ID:F6AA82(75%同一性)、UniProt ID:E2XN63(84%同一性)、UniProt ID:F2KE35(85%同一性)、UniProt ID:C3KDS5(83%同一性)、UniProt ID:F8G3B7(86%同一性)、UniProt ID:G8PYD2(85%同一性)、UniProt ID:Q4KGS1(82%同一性)、UniProt ID:Q3KGL5(85%同一性)、UniProt ID:B0KV51(86%同一性)、UniProt ID:B1J4J2(86%同一性)、UniProt ID:A5W8H3(86%同一性)、UniProt ID:Q88N06(86%同一性)、UniProt ID:Q1I5T5(84%同一性)、UniProt ID:F8H1A4(77%同一性)、UniProt ID:A4VIV7(77%同一性)和UniProt ID:Q9I5I5(81%同一性)。在一些实施方案中,与共有序列SEQ ID NO:9等同或具有大于75%同一性的丙二酰基-辅酶A水解酶在重组宿主细胞中表达且用于根据本发明产生丙二酸。序列表中包括各自在根据本发明突变的关键谷氨酸残基的位置处含有X₁的A4XS22(SEQ ID NO:45)、F6AA82(SEQ ID NO:50)和E2XN63(SEQ ID NO:48)的序列。

[0081] 这个共有序列(SEQ ID NO:9)中的高度保守氨基酸是M1、E6、G13、R15、I16、A19、L21、D22、A23、L27、N28、A29、L30、L32、P33、M34、I35、L38、W45、A46、C53、V54、L56、R57、G58、N59、G60、K62、A63、F64、C65、A66、G67、G68、V70、L73、C77、P81、G82、P85、L87、A88、F91、F92、Y96、R97、L98、H103、P106、K107、P108、C111、W112、H114、G115、V117、G119、G120、G121、M122、

G123、L124、Q126、R131、I132、V133、T134、P135、R138、L139、M141、P142、E143、I146、G147、L148、D151、V152、G153、S155、F157、L158、R160、P162、G163、L165、G166、L167、F168、L171、N177、D180、A181、D183、L184、L186、A187、D188、R189、Q195、Q196、L199、L203、Q205、N207、W208、E210、Q215、L216、S218、L219、A222、P232、L237、R239、R240、D244、L247、D248、A258、D267、L269、G280、P282、V288、W289、Q291、R294、R296、L298、S299、L300、E307、Y308、S311、L312、N313、C314、R316、H317、P318、F320、E322、G323、V324、R325、A326、R327、L328、D330、D332、P335、W337、W339、P346、A352、H353和F354。在各种实施方案中,与这个共有序列(SEQ ID NO:9)同源的丙二酰基-辅酶A水解酶含有这些保守氨基酸的至少25%,常为这些保守氨基酸的大多数(大于50%),且有时是这些保守氨基酸的全部。

[0082] 这个共有序列(SEQ ID NO:9)中的一些氨基酸为活性所必需且在所述类别的所有成员之间是保守的。由基于假单胞菌属EC 3.1.2.4的共有序列涵盖的丙二酰基-辅酶A水解酶含有对水解酶活性必需的六个保守活性位点残基:(i)共有序列中的三个活性位点氨基酸残基(G67、G68、G120)据信为形成氧离子孔所必需(但本发明不受任何作用机制理论限制),所述氧离子孔负责使由酰基-辅酶A底物产生的烯醇阴离子中间体稳定;(ii)共有序列的两个氨基酸残基(E143、D151)据信为酰基-辅酶A水解所必需;且(iii)(SEQ ID NO:9的)氨基酸X₁₉₅据信为丙二酰基-辅酶A底物结合所必需。然后,在这六个残基中,五个存在于共有序列(SEQ ID NO:9)中和由那个序列涵盖的所有丙二酰基-辅酶A水解酶中,且第六个残基X₁₉₅为提供本发明的能够在重组宿主细胞中产生丙二酸的丙二酰基-辅酶A水解酶所必需。在本发明的各种实施方案中,(己)使关键野生型谷氨酸残基(E95)突变成极性带正电荷氨基酸(即R、H、K、S、T、N、Q、Y)或A或D以产生X₁₉₅且提供本发明的能够在重组宿主细胞中产生丙二酸的丙二酰基-辅酶A水解酶。在本发明的一些实施方案中,氨基酸E95是(己突变成)选自由K、S、T、N、Y、A和D组成的组的氨基酸。在本发明的一些实施方案中,氨基酸E95是S或N。来自共有序列的六个必需残基(G67、G68、G120、E143、D151、X₁₉₅)分别对应于用于在实施例31中说明本发明的黄褐假单胞菌(*Pseudomonas fulva*) (12-X) F6AA82-2中的G67、G68、G120、E143、D151和E95(通常突变成X₁₉₅) (参见含有突变E95S/Q348A的SEQ ID NO:50)。

[0083] 与共有序列(SEQ ID NO:9)同源且由本发明提供的克隆或合成核酸编码的适于丙二酰基-辅酶A水解的酶的非限制性实例包括含有至少一个通过由以下组成的组的突变酶说明的突变的突变酶:黄褐假单胞菌(菌株12-X) F6AA82 (E95S)、F6AA82 (E95N)、F6AA82 (E95A)、F6AA82 (E95H)、F6AA82 (E95K)、F6AA82 (E95R)、F6AA82 (E95Q)、F6AA82 (E95D)、F6AA82 (E95T)、F6AA82 (E95Y),如实施例43中所说明;荧光假单胞菌(*Pseudomonas fluorescens*) WH6E2XN63 (E95S)、E2XN63 (E95N)、E2XN63 (E95A)、E2XN63 (E95H)、E2XN63 (E95K)、E2XN63 (E95R)、E2XN63 (E95Q)、E2XN63 (E95D)、E2XN63 (E95T)、E2XN63 (E95Y);门多萨假单胞菌(*Pseudomonas mendocina*) (菌株ymp) A4XS22 (E95S)、A4XS22 (E95N)、A4XS22 (E95A)、A4XS22 (E95H)、A4XS22 (E95K)、A4XS22 (E95R)、A4XS22 (E95Q)、A4XS22 (E95D)、A4XS22 (E95T)、E2XN63 (E95Y)。

[0084] 在本发明的各种实施方案中,丙二酰基-辅酶A水解酶是来自黄褐假单胞菌(菌株12-X)的F6AA82 (E95S)、来自荧光假单胞菌WH6的E2XN63 (E95S)、来自门多萨假单胞菌(菌株ymp)的A4XS22 (E95S),如实施例31中所说明。

[0085] 如实施例43中所说明,F6AA82 E95X₁₉₅突变产生丙二酰基-辅酶A水解酶活性。含有

突变E95S、E95Y、E95T、E95N、E96K、E95A和E95D的F6AA82蛋白产生的丙二酸显著(t检验, $p < 0.05$) 多于野生型F6AA82蛋白。含有这些突变的F6AA82蛋白适于用作丙二酰基-辅酶A水解酶以及产生丙二酸。在含有赋予丙二酰基-辅酶A水解酶活性的 X_1 突变的F6AA82酶中, 突变E95S或E95N是优选的。在含有赋予丙二酰基-辅酶A水解酶活性的 X_1 突变的F6AA82酶中, 突变E95A、E95T、E95K、E95Y和E95D是适合的。在本发明的各种实施方案中, 丙二酰基-辅酶A水解酶是F6AA82 (E95S)。在本发明的其它实施方案中, 丙二酰基-辅酶A水解酶是F6AA82 (E95N)。含有突变E95H、E95Q或E95R的F6AA82蛋白在如实施例中所采用的测试条件下不导致丙二酸产生得以增加。

[0086] 当与本文共有序列具有实质性同源性的酶具有次优或不具有丙二酰基-辅酶A水解酶活性时, 那么根据本发明, 它可被突变来符合本文提供的共有序列以提供本发明的丙二酰基-辅酶A水解酶。举例来说, 蛋白A5W8H3与丙二酰基-辅酶A水解酶共有SEQ ID NO:9显示86%同一性, 但不展现丙二酰基-辅酶A水解酶活性(实施例31)。可使在A5W8H3与共有SEQ ID NO:9之间不同的一个或多个氨基酸突变以引入丙二酰基-辅酶A水解酶活性。具体来说, A5W8H3突变T2N、C5F、V7J、L8B、G10U、D12B、P24J、A26U、N31U、Q41B、A72B、A740、Q75J、S83J、S90B、A94J、F100B、A101B、N129U、A159U、F1690、P175B、G185J、G192B、A198J、A213J、N217B、Q224J、C228B、A229J、W2360、H241Q、E242B、Q245J、A2520、R261A、Q264J、D272B、G274A、Q275B、Y297B、Q302J、Q305B、M3100、N333B、A347J、A355J、A3570和/或G368U可用于赋予丙二酰基-辅酶A水解酶活性。

[0087] 基于细菌EC 3.1.2.4酶的丙二酰基-辅酶A水解酶共有序列

[0088] 尽管芽孢杆菌属和假单胞菌属在进化上是疏远的(即芽孢杆菌属是格兰氏阳性而假单胞菌属是革兰氏阴性), 但在芽孢杆菌属EC 3.1.2.4酶与假单胞菌属EC 3.1.2.4酶之间存在显著序列保守性。本发明提供一种基于这些细菌EC 3.1.2.4酰基-辅酶A水解酶的丙二酰基-辅酶A水解酶共有序列(SEQ ID NO:10)。由这个共有序列涵盖的丙二酰基-辅酶A水解酶通常具有来自这个序列的高度保守氨基酸的至少25% (或大多数或全部), 所述保守氨基酸选自由以下组成的组: L53、L59、N60、L62、M66、L88、F97、C98、A99、G100、G101、F124、F125、Y129、K140、P141、G148、G152、G153、G154、G156、L157、T167、M174、P175、E176、I179、G180、D184、V185、G186、L191、L210、D219、A226、P333、N364、F375、E377、D385和P390。本发明的与这个共有序列同源的适合丙二酰基-辅酶A水解酶包括共有序列的为丙二酰基-辅酶A水解所必需的活性位点氨基酸(G100、G101、G153、E176和D184) 以及 X_1 128, 其中(已) 使关键野生型谷氨酸残基(E128) 突变成极性 or 带电荷氨基酸(即R、H、K、S、T、N、Q、Y) 或D或A, 且能够在重组宿主细胞中产生丙二酸。

[0089] 基于丙二酰基-辅酶A:ACP转酰基酶的丙二酰基-辅酶A水解酶

[0090] 在本发明的其它实施方案中, 丙二酰基-辅酶A水解酶是突变丙二酰基-辅酶A:ACP转酰基酶(EC 2.3.1.39)。本发明提供一种大肠杆菌FabD序列(SEQ ID NO:53), 且在各种实施方案中, 适合用于本发明的方法中的丙二酰基-辅酶A水解酶在相对于SEQ ID NO:53比时具有至少50%、60%、70%、80%、90%、95%或更高的序列同一性且在比对位置处含有一个或多个以下氨基酸突变: S92C、H201N、R117D、R117E、R117N、R117Y、R117G、R117H、Q11D、Q11E、Q11N、Q11Y、Q11G、Q11H、L93A、L93V、L93I、L93F、L93S、L93G。

[0091] 在本发明的一些实施方案中, 丙二酰基-辅酶A水解酶是具有一个或多个选自自由以

下组成的组的突变的突变大肠杆菌FabD丙二酰基-辅酶A:ACP转酰基酶(对于野生型序列,参见SEQ ID NO:53):S92C、H201N、R117D、R117E、R117N、R117Y、R117G、R117H、Q11D、Q11E、Q11N、Q11Y、Q11G、Q11H、L93A、L93V、L93I、L93F、L93S、L93G。实施例35说明表达含有一个以下突变组合的突变大肠杆菌FabD酶且在高于无突变FabD酶的野生型酵母的水平下产生丙二酸的重组酵母细胞:S92C/L91V/R117H、L91I/R117Y/A246E、Q80L/L91S/R117G和L91I/R117Y。

[0092] 章节3:表达载体

[0093] 在本发明的各种方面,重组宿主细胞已通过“遗传工程化”加以修饰以产生重组丙二酰基-辅酶A水解酶和丙二酸。宿主细胞通常是通过重组DNA技术加以工程化以表达编码丙二酰基-辅酶A水解酶的异源核酸,所述丙二酰基-辅酶A水解酶是突变形式的天然存在的酰基-辅酶A水解酶或转酰基酶或根据本文提供的一种共有序列制备的非天然存在的丙二酰基-辅酶A水解酶,或是具有丙二酰基-辅酶A水解酶活性的天然存在的酰基-辅酶A水解酶,所述天然存在的酰基-辅酶A水解酶在它天然存在的细胞中过表达或在它不天然存在的细胞中异源表达。

[0094] 本发明的核酸构建体包括包含编码一种或多种丙二酰基-辅酶A水解酶的核酸的表达载体。使编码酶的核酸可操作地连接于启动子和任选其它控制序列以使当在适合条件下培养时,主题酶在含有表达载体的宿主细胞中表达。采用的启动子和控制序列取决于选择用于产生丙二酸的宿主细胞。因此,本发明不仅提供表达载体,而且也提供用于构建表达载体的核酸构建体。用于设计和制备核酸构建体和表达载体的方法通常为本领域技术人员所熟知且因此在本文中仅简要综述。

[0095] 编码丙二酰基-辅酶A水解酶的核酸可通过为本领域普通技术人员所知的任何适合方法制备,所述方法包括例如直接化学合成和克隆。此外,用于本发明中的核酸序列可自提供从头合成核酸的商业供应商获得。

[0096] 编码所需酶的核酸可通过已知方法并入表达载体中,所述方法包括例如使用限制酶来裂解例如质粒的表达载体中的特定位点,由此产生本发明的表达载体。一些限制酶产生单链末端,其可被退火以获得具有有序列互补于裂解的表达载体的末端的末端的核酸序列,或被合成以具有有序列互补于裂解的表达载体的末端的末端。接着使用例如DNA连接酶的适当酶共价连接末端。DNA接头可用于促进核酸序列连接至表达载体中。

[0097] 也可通过利用为本领域技术人员所知的基于聚合酶链反应(PCR)的方法来组合一组个别核酸序列。举例来说,各所需核酸序列可最初在单独PCR中产生。此后,设计特定引物以使PCR产物的末端含有互补序列。当混合、变性和再退火PCR产物时,在它们的3'末端处具有匹配序列的各链重叠且可充当彼此的引物。通过DNA聚合酶使这个重叠部分延伸会产生原始序列“拼接”在一起的分子。以这个方式,一系列个别核酸序列可被接合且随后同时转导至宿主细胞中。因此,实现多个核酸序列的各者的表达。

[0098] 典型表达载体含有之前是一种或多种控制序列或调控区域以及任选继之以一种或多种控制序列或调控区域的所需核酸序列,所述控制序列或调控区域包括启动子和(当基因产物是蛋白质时)核糖体结合位点,例如长度通常是3-9个核苷酸且通常位于先于编码序列的起始密码子上游3-11个核苷酸处的核苷酸序列,在大肠杆菌或其它原核宿主的情况下,所述编码序列继之以转录终止子。参见Shine等,Nature 254:34(1975)以及Steitz,

Biological Regulation and Development:Gene Expression(R.F.Goldberger编),第1卷,第349页(1979)Plenum Publishing,N.Y.。在如酵母的真核宿主的情况下,典型表达载体含有之前是一种或多种调控区域以及用以起始翻译的Kozak序列且继之以终止子的所需核酸编码序列。参见Kozak,Nature 308:241-246(1984)。

[0099] 调控区域或控制序列包括例如含有启动子和操纵子的那些区域。使启动子可操作地连接于所需核酸编码序列,由此通过RNA聚合酶起始核酸序列的转录。操纵子是邻近于启动子的核酸序列,其含有其中转录因子可进行结合的蛋白质结合结构域。转录因子活化或阻遏自启动子起始转录。以这个方式,对转录的控制是基于所用特定调控区域和存在或不存在相应转录因子所达成。用于原核表达的非限制性实例包括乳糖启动子(当与乳糖接触时,LacI阻遏蛋白改变构象,由此阻止LacI阻遏蛋白结合操纵子)和色氨酸启动子(当与色氨酸复合时,TrpR阻遏蛋白具有结合操纵子的构象;在不存在色氨酸下,TrpR阻遏蛋白具有不结合操纵子的构象)。用于真核表达的启动子的非限制性实例包括pTDH3、pTEF1(如实施例31中所说明)、pTEF2、pRNR2、pRPL18B、pREV1、pGAL1、pGAL10、pGAPDH、pCUP1、pMET3、pPGK1、pPYK1、pHXT7、pPDC1、pFBA1、pTDH2、pPGI1、pPDC1、pTPI1、pEN02、pADH1和pADH2。如实施例44中所述,来自酿酒酵母菌株BY4741的用于基因HSP150、PGK1、PH05、SCT1、PRB1、TPI1、ACH1、HXK2、ACO1、JEN1、MDH2、POX1、CIT1、ALD4、ADH1、TDH3、ADH2和SDH1的启动子被证明用于根据本发明产生丙二酸。如将为本领域普通技术人员所了解,多种表达载体及其组分可用于本发明中。

[0100] 尽管任何适合表达载体都可用于并有所需序列,但可易于获得的表达载体不加限制地包括:质粒,如pESC、pTEF、p414CYC1、p414GALS、pSC101、pBR322、pBBR1MCS-3、pUR、pEX、pMR100、pCR4、pBAD24、pUC19、pRS系列;以及噬菌体,如M13噬菌体和 λ 噬菌体。当然,所述表达载体仅可适于特定宿主细胞或表达特定丙二酰基-辅酶A水解酶。然而,本领域普通技术人员可易于通过常规实验确定任何特定表达载体是否适合于任何给定宿主细胞或蛋白质。举例来说,可将表达载体引入宿主细胞中,接着监测所述宿主细胞的活力和载体中含有的序列的表达。此外,可参照描述表达载体和它们针对任何特定宿主细胞的适合性的相关文本和文献。除使用表达载体之外,也构造表达盒直接整合至宿主基因组中的菌株。

[0101] 例如通过转导、转染或转化将表达载体引入或转移至宿主细胞中。用于将表达载体引入宿主细胞中的所述方法为本领域普通技术人员所熟知。举例来说,一种用于用表达载体转化大肠杆菌的方法涉及氯化钙处理,其中表达载体是通过钙沉淀引入。

[0102] 对于鉴定核酸是否已成功引入或进入宿主细胞中,可用多种方法。举例来说,潜在转化的宿主细胞的培养物可使用适合稀释法分离成个别细胞且此后个别地生长并测试引入的核酸中含有的基因的所需基因产物的表达。举例来说,常使用的规范涉及基于已由表达载体中的抗生素抗性赋予基因赋予的抗生素抗性选择细胞,所述基因如 β 内酰胺酶(amp)、氨基糖苷磷酸转移酶(neo)和潮霉素磷酸转移酶(hyg、hph、hpt)基因。

[0103] 通常,本发明的宿主细胞将已用至少一种表达载体转化。当仅使用单一表达载体时,载体将通常含有丙二酰基-辅酶A水解酶基因。一旦宿主细胞已用表达载体转化,就在含有如糖(例如葡萄糖)的碳源的适合培养基中培养宿主细胞。当培养宿主细胞时,发生用于产生丙二酸的酶的表达。一旦表达,酶就催化丙二酰基-辅酶A的硫酯键水解,由此释放丙二酸和辅酶A。

[0104] 如果本发明的宿主细胞将包括一种以上异源基因,那么多种基因可自一种或多种载体表达。举例来说,单一表达载体可包含一种、两种或更多种编码一种、两种或更多种丙二酰基-辅酶A水解酶和/或提供某一有用功能(例如丙二酸产率、效价和/或生产力改进)的其它蛋白质的基因。异源基因可含于以游离体形式复制的载体中或整合至宿主细胞基因组中的载体中,且当采用一种以上载体时,那么所有载体都可以游离体形式(在染色体外)复制,或所有载体都可整合或一些载体可整合且一些载体可以游离体形式复制。染色体整合通常用于将经受持续增殖的细胞,例如用于产生供工业应用的丙二酸的细胞。实施例45说明由本发明提供的丙二酰基-辅酶A水解酶基因的基因组整合的益处;在那个实施例中,基因的单一整合拷贝显示导致丙二酸效价高于当相同基因自质粒表达时获得的效价。这个实施例也用于证明以每个细胞的丙二酰基-辅酶A水解酶编码基因的拷贝数增加的方式调节由本发明提供的丙二酸产生导致发酵培养基中的丙二酸增加。尽管“基因”通常主要由单一启动子和单一编码序列组成,但在某些宿主细胞中,两种或更多种编码序列可受操纵子中的一种启动子控制。在一些实施方案中,使用两操纵子或三操纵子系统。

[0105] 在一些实施方案中,相对于某一参照序列,采用的编码序列已被修饰来反映所选宿主细胞的密码子偏好。众多生物体的密码子使用表可易于获得且可用于指导序列设计。使用给定宿主生物体的普遍密码子通常会改进靶序列在宿主细胞中的翻译。作为一个非限制性实例,在一些实施方案中,主题核酸序列将针对酵母密码子偏好加以修饰(参见例如 Bennetzen等, *J. Biol. Chem.* 257:3026-3031 (1982))。在一些实施方案中,核苷酸序列将针对大肠杆菌密码子偏好加以修饰(参见例如 Nakamura等, *Nucleic Acids Res.* 28:292 (2000))。在其它实施方案中,核苷酸序列被修饰来包括针对酿酒酵母密码子偏好加以最优化的密码子(参见实施例42)。

[0106] 核酸可通过多种常规重组技术制备。简而言之,主题核酸可自全部可通过包括但不限于PCR和rt-PCR的各种扩增方法直接从细胞提取或重组产生的基因组DNA片段、cDNA和RNA制备。主题核酸也可通过直接化学合成制备。

[0107] 可使用众多技术增加(或降低)宿主微生物中的核酸转录水平。举例来说,可通过使用较高拷贝数的包含核酸序列的表达载体,或通过使所需核酸的多个拷贝整合至宿主微生物的基因组中来增加核酸的拷贝数,如实施例45中所说明。使所需核酸序列整合于宿主染色体上的非限制性实例包括recA介导的重组、 λ 噬菌体重组酶介导的重组和转位子插入。可通过改变多顺反子mRNA上的编码区域的次序,或将多顺反子操纵子分解成多个各自具有它自身启动子的多顺反子或单顺反子操纵子来增加核酸转录物水平。可通过增加(或降低)蛋白质编码区域可操作地连接的启动子的强度来增加(或降低)RNA水平。用于质粒设计和装配以提供丙二酸产生的说明性技术提供于实施例1、3、31和35中。

[0108] 也可以许多方式增加所需多肽序列在宿主微生物中的翻译水平。非限制性实例包括增加mRNA稳定性、修饰核糖体结合位点(或Kozak)序列、改进核糖体结合位点(或Kozak序列)与编码所需多肽的核酸序列的起始密码子之间的距离或序列、修饰位于编码所需多肽的核酸序列的起始密码子的5'的顺反子间区域、使mRNA转录物的3'末端稳定、改进多肽的密码子使用、改变多肽的生物合成中使用的低使用率/稀有密码子tRNA的表达。可基于对源于宿主微生物的基因的序列分析确定优选密码子和低使用率/稀有密码子tRNA。

[0109] 可通过使编码所需多肽的核酸序列突变,从而导致相对于对照多肽序列,所需多

肽序列得以修饰来增加多肽半衰期或稳定性。当修饰的多肽是酶时,酶在宿主中的活性可由于以下因素而改变:在宿主细胞中的溶解性增加;在所需pH下的功能改进;抑制酶活性的结构域的移除;对所需底物的动力学参数改进(较低 K_m 或较高 K_{cat} 值);由细胞内代谢物进行的变构调控的移除等。也可通过使酶随机突变诱发来分离改变/修饰的酶,以使改变/修饰的酶可自游离载体或自整合至宿主微生物的基因组中的重组基因表达。

[0110] 章节4:重组宿主细胞

[0111] 在一个方面,本发明提供适于生物产生丙二酸的重组宿主细胞。任何适合宿主细胞都可用于实施本发明的方法。在一些实施方案中,宿主细胞是重组宿主微生物,其中核酸分子已被插入、缺失或修饰(即突变;例如通过核苷酸的插入、缺失、取代和/或倒位)以产生丙二酸,或以相对于“对照细胞”或“参照细胞”增加丙二酸的产率、效价和/或生产力。“对照细胞”可用于比较目的,且通常是不含有对目标宿主细胞进行的一种或多种修饰的野生型或重组亲本细胞。

[0112] 在一重要实施方案中,本发明提供适于在足以随后进行如本文所述的纯化和使用的水平下产生丙二酸的重组酵母细胞。酵母宿主细胞是用于构建包含催化产生小分子产物的异源酶的重组代谢路径的优异宿主细胞。存在已建立的分子生物学技术和编码为构建酵母表达载体所必需的遗传元件的核酸,所述遗传元件包括但不限于启动子、复制起点、抗生素抗性标记、营养缺陷标记、终止子等。其次,已明确建立用于使核酸整合至酵母染色体中的技术。酵母也提供作为工业发酵宿主的许多优势。酵母可耐受高浓度的有机酸且在低pH下维持细胞活力并可在好氧培养条件与厌氧培养条件两者下生长,且存在已建立的发酵液和发酵方案。菌株能够在低pH下增殖和/或产生所需产物提供关于本发明的许多优势。首先,这个特征提供对由丙二酸的产生所创建的环境的耐受性。其次,就工艺观点来看,维持低pH环境的能力会限制能够污染且破坏批料的生物体的数目。第三,这个特征也消除或至少降低对添加额外酸来有助于通过由本发明提供的一些方法(参见实施例37)纯化丙二酸的需要。

[0113] 在本发明的一些实施方案中,包含编码丙二酰基-辅酶A水解酶的异源核酸的重组宿主细胞是真核生物。在各种实施方案中,真核生物是选自以下各属的非限制性清单的酵母:假丝酵母属、隐球酵母属、汉逊酵母属、伊萨酵母属、克鲁维酵母属、Komagataella、油脂酵母属、毕赤酵母属、红冬孢酵母属、红酵母属、酵母属或耶罗威亚酵母属种类。在各种实施方案中,酵母具有选自由以下组成的组的种类:白色假丝酵母(*Candida albicans*)、嗜酒假丝酵母(*Candida ethanolica*)、克鲁斯假丝酵母(*Candida krusei*)、*Candida methanosorbosa*、*Candida sonorensis*、热带假丝酵母(*Candida tropicalis*)、弯曲隐球酵母(*Cryptococcus curvatus*)、多形汉逊酵母(*Hansenula polymorpha*)、东方伊萨酵母(*Issatchenkia orientalis*)、乳酸克鲁维酵母(*Kluyveromyces lactis*)、马克斯克鲁维酵母(*Kluyveromyces marxianus*)、耐热克鲁维酵母(*Kluyveromyces thermotolerans*)、*Komagataella pastoris*、斯氏油脂酵母(*Lipomyces starkeyi*)、安格斯毕赤酵母(*Pichia angusta*)、*Pichia deserticola*、盔形毕赤酵母(*Pichia galeiformis*)、库德毕赤酵母(*Pichia kodamae*)、库德里阿兹威毕赤酵母、膜醭毕赤氏酵母(*Pichia membranaefaciens*)、甲醇毕赤酵母(*Pichia methanolica*)、巴斯德毕赤酵母(*Pichia pastoris*)、柳毕赤酵母(*Pichia salictaria*)、树干毕赤酵母(*Pichia stipitis*)、耐热毕

赤酵母 (*Pichia thermotolerans*)、喜海藻糖毕赤酵母 (*Pichia trehalophila*)、圆红冬孢酵母 (*Rhodospiridium toruloides*)、粘红酵母 (*Rhodotorula glutinis*)、禾本红酵母 (*Rhodotorula graminis*)、贝酵母 (*Saccharomyces bayanus*)、布拉迪酵母 (*Saccharomyces boulardi*)、酿酒酵母、克鲁维酵母 (*Saccharomyces kluyveri*) 和解脂耶罗威亚酵母 (*Yarrowia lipolytica*)。本领域技术人员将认识到这个清单以广义涵盖酵母,包括油脂菌株与非油脂菌株两者。实施例46和47说明根据本发明使用库德里阿兹威毕赤酵母菌株Y-134。

[0114] 本发明提供用于生物产生丙二酸的替代性重组宿主细胞。说明性实例包括真核、原核和古细菌细胞。真核细胞的说明性实例包括但不限于:黑曲霉 (*Aspergillus niger*)、米曲霉 (*Aspergillus oryzae*)、寇氏隐甲藻 (*Cryptothecodinium cohnii*)、日本小克银汉霉 (*Cunninghamella japonica*)、冠状虫霉 (*Entomophthora coronata*)、高山被孢霉 (*Mortierella alpina*)、卷枝毛霉菌 (*Mucor circinelloides*)、粗糙链孢霉 (*Neurospora crassa*)、终极腐霉 (*Pythium ultimum*)、裂殖壶菌 (*Schizochytrium limacinum*)、破囊壶菌 (*Thraustochytrium aureum*)、里氏木霉 (*Trichoderma reesei*) 和法夫酵母 (*Xanthophyllomyces dendrorhous*)。一般来说,如果使用真核细胞,那么采用非病原菌株。非病原菌株的说明性实例包括但不限于:巴斯德毕赤酵母和酿酒酵母。此外,包括酿酒酵母的某些菌株已由食物与药品管理局 (Food and Drug Administration) 指定为通常认为是安全的 (或GRAS) 且因此可合宜地用于本发明的方法的各种实施方案中。

[0115] 由本发明提供的重组原核宿主细胞的说明性实例包括但不限于枯草芽孢杆菌 (*Bacillus subtilis*)、产氨短杆菌 (*Brevibacterium ammoniagenes*)、拜氏梭菌 (*Clostridium beijerinckii*)、阪崎肠杆菌 (*Enterobacter sakazakii*)、嗜酸乳杆菌 (*Lactobacillus acidophilus*)、乳酸乳球菌 (*Lactococcus lactis*)、百脉根根瘤菌 (*Mesorhizobium loti*)、绿脓杆菌 (*Pseudomonas aeruginosa*)、恶臭假单胞菌 (*Pseudomonas putida*)、荚膜红细菌 (*Rhodobacter capsulatus*)、类球红细菌 (*Rhodobacter sphaeroides*)、肠道沙门氏菌 (*Salmonella enterica*)、伤寒沙门氏菌 (*Salmonella typhi*)、鼠伤寒沙门氏菌 (*Salmonella typhimurium*)、弗累克斯纳氏杆菌 (*Shigella flexneri*)、金黄色葡萄球菌 (*Staphylococcus aureus*)、产二素链霉菌 (*Streptomyces ambofaciens*)、金素链霉菌 (*Streptomyces aureofaciens*)、金色链霉菌 (*Streptomyces aureus*)、杀真菌素链霉菌 (*Streptomyces fungicidicus*)、灰色产色链霉素 (*Streptomyces griseochromogenes*)、灰色链霉菌 (*Streptomyces griseus*)、变铅青链霉菌 (*Streptomyces lividans*)、橄榄灰链霉菌 (*Streptomyces olivogriseus*)、枝链霉菌 (*Streptomyces rameus*)、田无链霉菌 (*Streptomyces tanashiensis*) 和葡萄色链霉菌 (*Streptomyces vinaceus*)。包括枯草芽孢杆菌、嗜酸乳杆菌的某些这些细胞已由食物与药品管理局指定为通常认为是安全的 (或GRAS) 且因此可用于本发明的方法的各种实施方案中。尽管自公共安全和管理的观点来看合乎需要,但GRAS状态不影响宿主菌株用于实施本发明的能力;因此,非GRAS以及甚至病原生物体包括在适合用于实施本发明的说明性宿主菌株的清单中。

[0116] 大肠杆菌也是用于代谢路径构建的优异原核宿主细胞,且大肠杆菌也良好用于小分子产物的工业发酵中。不同于大多数野生型酵母菌株,野生型大肠杆菌可分解代谢作为

碳源的戊糖与己糖两者。相对于酵母,大肠杆菌也具有较短倍增时间,从而使得实验能够更快速进行。本发明提供广泛多种适于如本文所述产生丙二酸的大肠杆菌宿主细胞。在本发明的方法的各种实施方案中,包含编码丙二酰基-辅酶A水解酶的异源核酸的重组宿主细胞是大肠杆菌细胞。

[0117] 章节5:用于达成改进的丙二酸产生的其它修饰和发酵条件

[0118] 在本发明的其它方面,丙二酸产率、效价和/或生产力增加是通过以下方式来实现:采用由本发明提供的已以不同于引入异源丙二酰基-辅酶A水解酶的方式或也以除引入异源丙二酰基-辅酶A水解酶之外的方式遗传修饰的宿主细胞和/或采用由本发明的某些方法提供的发酵条件。简而言之,本发明的重组宿主细胞包含增加乙酰基-辅酶A生物合成、增加丙二酰基-辅酶A生物合成、降低丙二酸分解代谢、增加丙二酸自宿主细胞分泌、增加宿主细胞对丙二酸的耐受性、以及增加各种碳源的分解代谢的遗传修饰。

[0119] 增加乙酰基-辅酶A生物合成的遗传修饰和发酵条件

[0120] 根据本发明,丙二酸效价、产率和/或生产力增加可通过增加乙酰基-辅酶A生物合成的遗传修饰来实现,且本发明提供增加乙酰基-辅酶A生物合成的酶、用于表达增加乙酰基-辅酶A生物合成的酶的载体、表达增加乙酰基-辅酶A生物合成且增加丙二酸效价、产率和/或生产力的酶的宿主细胞、以及与此相关的方法。如上所述,丙二酸是通过水解可自乙酰基-辅酶A产生的丙二酰基-辅酶A产生;因此,乙酰基-辅酶A生物合成增加可改进丙二酸产生。

[0121] 产生乙酰基-辅酶A所通过的一个途径是通过乙酰基-辅酶A合成酶(EC 6.2.1.1)来达成,所述合成酶催化自乙酸和辅酶A(CoA)形成乙酰基-辅酶A。本发明提供适于根据本发明的方法产生丙二酸的重组宿主细胞,相对于不包含异源乙酰基-辅酶A合成酶的宿主细胞,所述重组宿主细胞包含一种或多种增加丙二酸效价、产率和/或生产力的异源乙酰基-辅酶A合成酶(ACS)。适合ACS酶的非限制性实例是酿酒酵母ACS1(GenBank: AAC04979.1)和ACS2(GenBank: CAA97725.1)。在一些实施方案中,包含酿酒酵母乙酰基-辅酶A合成酶ACS1和/或ACS2的重组宿主细胞用于增加丙二酸效价、产率和/或生产力。在其它实施方案中,包含选自由肠道沙门氏菌Acs、大肠杆菌AcsA和枯草芽孢杆菌AcsA组成的组的乙酰基-辅酶A合成酶的重组宿主细胞用于增加丙二酸产率、效价和/或生产力。本领域技术人员了解的是其它乙酰基-辅酶A合成酶可根据本发明在产生丙二酸的重组宿主细胞中表达以增加丙二酸产率、效价和/或生产力。这个修饰说明于实施例9和46中。

[0122] 产生乙酰基-辅酶A所通过的第二途径是通过催化自丙酮酸形成乙酰基-辅酶A的丙酮酸脱氢酶复合物来达成。本发明提供适于根据本发明的方法产生丙二酸的重组宿主细胞,相对于不包含异源丙酮酸脱氢酶复合物酶的宿主细胞,所述重组宿主细胞包含一种或多种增加丙二酸效价、产率和/或生产力的异源丙酮酸脱氢酶复合物酶。适合丙酮酸脱氢酶复合物酶的非限制性实例包括酿酒酵母PDA1、PDB1、LAT1、LPD1和PDX1。在本发明的一些实施方案中,用于产生丙二酸的重组宿主细胞中的丙二酸产率、效价和/或生产力是通过表达一种或多种选自由酿酒酵母PDA1、PDB1、LAT1、LPD1和PDX1组成的组的丙酮酸脱氢酶来增加。本领域技术人员了解的是其它丙酮酸脱氢酶可根据本发明在产生丙二酸的重组宿主细胞中表达以增加丙二酸产率、效价和/或生产力。这个修饰说明于实施例10中。

[0123] 产生乙酰基-辅酶A所通过的第三途径是通过包含催化乙醇转化成乙酰基-辅酶A

的酶的异源乙醇分解代谢路径来达成。相较于丙二酸,乙醇是一种价格比较低廉的化学品,且产生丙二酸且表达乙醇分解代谢路径的宿主细胞可使乙醇转化成丙二酸。醇脱氢酶(EC 1.1.1.1)催化乙醇转化成乙醛。适合醇脱氢酶的非限制性实例包括选自由酿酒酵母ADH2、大肠杆菌AdhP、智人ADH1A、智人ADH1B和智人ADH1C组成的组的醇脱氢酶。除醇脱氢酶之外,乙醇分解代谢路径也包含乙醛脱氢酶(酰化;EC 1.2.1.10)或醛脱氢酶(EC 1.2.1.3)和乙酰基-辅酶A合成酶(EC 6.2.1.1)。乙醛脱氢酶(酰化)催化乙醛转化成乙酰基-辅酶A,醛脱氢酶催化乙醛转化成乙酸,且如上所述的乙酰基-辅酶A合成酶催化自乙酸和辅酶A形成乙酰基-辅酶A。适合乙醛脱氢酶(酰化)的非限制性实例包括选自由大肠杆菌MhpF、大肠杆菌AdhE、假单胞菌属CF600 DmpF和恶臭假单胞菌TodL组成的组的乙醛脱氢酶。醛脱氢酶的非限制性实例包括酿酒酵母ALD2、ALD3、ALD4、ALD5和ALD6;以及智人ALD1、ALD2、ALD4和ALD10。乙酰基-辅酶A合成酶的非限制性实例包括酿酒酵母ACS1、酿酒酵母ACS2和大肠杆菌Acs。

[0124] 本发明提供适于根据本发明的方法产生丙二酸的重组宿主细胞,相对于不包含异源乙醇分解代谢路径酶的宿主细胞,所述重组宿主细胞包含一种或多种增加丙二酸产率、效价和/或生产力的异源乙醇分解代谢路径酶。在一些实施方案中,异源乙醇分解代谢路径酶是乙醇脱氢酶和乙醛脱氢酶(酰化)。在一些实施方案中,异源乙醇分解代谢路径酶是酿酒酵母ADH2乙醇脱氢酶和大肠杆菌MhpF乙醛脱氢酶(酰化)。在一些实施方案中,异源酿酒酵母ADH2和大肠杆菌MhpF在表达异源酿酒酵母EHD3丙二酰基-辅酶A水解酶的重组大肠杆菌中表达。在其它实施方案中,异源酿酒酵母ADH2和大肠杆菌MhpF在表达异源酿酒酵母EHD3丙二酰基-辅酶A水解酶的重组酿酒酵母中表达。在其它实施方案中,异源酿酒酵母ADH2和大肠杆菌MhpF在表达异源酿酒酵母EHD3丙二酰基-辅酶A水解酶的重组油脂酵母中表达。在其它实施方案中,异源乙醇分解代谢路径酶是酿酒酵母ADH2乙醇脱氢酶和假单胞菌属种类CF600 DmpF乙醛脱氢酶(酰化)。在一些实施方案中,异源酿酒酵母ADH2和假单胞菌属种类CF600 DmpF在表达异源酿酒酵母EHD3丙二酰基-辅酶A水解酶的重组大肠杆菌中表达。在一些实施方案中,异源酿酒酵母ADH2和假单胞菌属种类CF600 DmpF在表达酿酒酵母EHD3丙二酰基-辅酶A水解酶的重组酿酒酵母中表达。在一些实施方案中,异源酿酒酵母ADH2和假单胞菌属种类CF600 DmpF在表达异源酿酒酵母EHD3丙二酰基-辅酶A水解酶的重组油脂酵母中表达。在其它实施方案中,异源乙醇分解代谢路径酶是酿酒酵母ADH2乙醇脱氢酶和恶臭假单胞菌TodL乙醛脱氢酶(酰化)。在一些实施方案中,异源酿酒酵母ADH2和恶臭假单胞菌TodL在表达异源酿酒酵母EHD3丙二酰基-辅酶A水解酶的重组大肠杆菌中表达。在一些实施方案中,异源酿酒酵母ADH2和恶臭假单胞菌TodL在表达异源酿酒酵母EHD3丙二酰基-辅酶A水解酶的重组酿酒酵母中表达。在一些实施方案中,异源酿酒酵母ADH2和恶臭假单胞菌TodL在表达异源酿酒酵母EHD3丙二酰基-辅酶A水解酶的重组油脂酵母中表达。在其它实施方案中,异源乙醇分解代谢路径酶是一种或多种选自含有酿酒酵母ADH2、大肠杆菌AdhP、智人ADH1A、智人ADH1B和/或智人ADH1C的组的醇脱氢酶以及一种或多种选自含有大肠杆菌MhpF、大肠杆菌AdhE、假单胞菌属CF600 DmpF和恶臭假单胞菌TodL的组的乙醛脱氢酶(酰化)。本领域技术人员了解的是其它醇脱氢酶和乙醛脱氢酶(酰化)可在适于根据本发明的方法产生丙二酸的重组宿主细胞中表达以增加丙二酸产率、效价和/或生产力。

[0125] 在其它实施方案中,异源乙醇分解代谢路径酶是乙醇脱氢酶、醛脱氢酶和乙酰基-

辅酶A合成酶。在一些实施方案中,异源乙醇分解代谢路径酶是酿酒酵母ALD2醇脱氢酶、酿酒酵母ALD2醛脱氢酶和酿酒酵母ACS1乙酰基-辅酶A合成酶。在其它实施方案中,异源乙醇分解代谢路径酶是酿酒酵母ALD2醇脱氢酶、酿酒酵母ALD2醛脱氢酶和酿酒酵母ACS2乙酰基-辅酶A合成酶。在其它实施方案中,异源乙醇分解代谢路径酶是酿酒酵母ALD2醇脱氢酶、酿酒酵母ALD6醛脱氢酶和酿酒酵母ACS1乙酰基-辅酶A合成酶。在其它实施方案中,异源乙醇分解代谢路径酶是一种或多种选自含有酿酒酵母ADH2、大肠杆菌AdhP、智人ADH1A、智人ADH1B和/或智人ADH1C的组的醇脱氢酶、一种或多种选自含有酿酒酵母ALD2、酿酒酵母ALD3、酿酒酵母ALD4、酿酒酵母ALD5、酿酒酵母ALD6、智人ALD1、智人ALD2、智人ALD4和/或智人ALD10的组的醛脱氢酶、以及一种或多种选自含有酿酒酵母ACS1、酿酒酵母ACS2和/或大肠杆菌Acs的组的乙酰基-辅酶A合成酶。

[0126] 在一些实施方案中,适于根据本发明的方法产生丙二酸的重组宿主细胞包含异源乙醇分解代谢路径酶且使内源产生的乙醇转化成乙酰基-辅酶A并增加丙二酸产率、效价和/或生产力。在其它实施方案中,乙醇是外源添加至发酵液中且适于根据本发明的方法产生丙二酸的重组宿主细胞包含异源乙醇分解代谢路径酶且使外源添加的乙醇转化成乙酰基-辅酶A并增加丙二酸产率、效价和/或生产力。当外源添加至发酵液中时,添加乙醇以获得最小浓度1%乙醇体积/体积,且通常添加至发酵液中以获得在1-15%体积/体积之间的浓度。用于根据本发明改进丙二酸产生的乙醇分解代谢说明于实施例11中。

[0127] 增加乙酰基-辅酶A的胞质液池是增加丙二酸生物合成的第四途径;在众多植物和动物细胞而非酿酒酵母中,ATP柠檬酸裂解酶(EC2.3.3.8)是负责胞质液乙酰基-辅酶A生物合成的主要酶。更详细而言,线粒体中的乙酰基-辅酶A通过柠檬酸合成酶的活性与草酰乙酸缩合以形成柠檬酸。随后,柠檬酸被自线粒体转运至胞质液中,在胞质液中,ATP柠檬酸裂解酶催化形成乙酰基-辅酶A、草酰乙酸和ADP。尽管酿酒酵母不含有天然ATP柠檬酸裂解酶,但已描述油脂酵母菌株中的适合异源ATP柠檬酸裂解酶(参见例如Boulton等, J.Gen.Microbiol.127:169-176(1981))。本发明提供包含一种或多种异源核油脂酵母ATP柠檬酸裂解酶的重组宿主细胞。油脂酵母ATP柠檬酸裂解酶的非限制性实例包括选自由以下组成的油脂酵母的组的柠檬酸裂解酶:弯假丝酵母(*Candida curvata*)、浅白隐球酵母(*Cryptococcus albidus*)、产油油脂酵母(*Lipomyces lipofer*)、圆红冬孢酵母(*Rhodospiridium toruloides*)、粘红酵母(*Rhodotorula glutanis*)、皮状丝孢酵母(*Trichosporon cutaneum*)、解脂耶罗威亚酵母等。在各种实施方案中,重组宿主细胞包含编码ATP柠檬酸裂解酶的异源核酸。在各种实施方案中,ATP柠檬酸裂解酶来自选自由弯假丝酵母、浅白隐球酵母、产油油脂酵母、圆红冬孢酵母、粘红酵母、皮状丝孢酵母、解脂耶罗威亚酵母组成的组的生物体。这个修饰说明于实施例12中。

[0128] 乙酰基-辅酶A生物合成也可根据本发明,通过改变一种或多种编码影响脂肪酸储存或分解代谢的蛋白质的核酸的表达来增加。本发明提供包含对一种或多种编码影响脂肪酸储存和分解代谢的蛋白质的核酸的遗传修饰的宿主细胞。在酿酒酵母中,这些蛋白质包括SNF2、IRA2、PRE9、PHO90、SPT21、POX1、ANT1、FOX3、PAS1、PAS3、ARE1、ARE2、DGA1、LR01、ACL1、MAE1、GLC3、GLG1、GLG2、PAT1和PEX11。这个修饰说明于实施例13中。

[0129] 在本发明的一些实施方案中,宿主细胞包含影响脂肪酸分解代谢中涉及的蛋白质

的表达和/或活性的遗传修饰。举例来说,大多数宿主细胞将通过 β -氧化路径天然降解脂肪酸、羟基脂肪酸和许多二酸。在大多数情况下, β -氧化是通过用酰基-辅酶A连接酶使游离脂肪酸基团活化成辅酶A硫酯来发生。酰基-辅酶A中间体进一步氧化并降解-通过2,3烯酰基-辅酶A、3-羟基酰基-辅酶A和3-酮酰基-辅酶A进行-且随后裂解导致产生乙酰基-辅酶A和相对于初始底物缩短两个碳的酰基-辅酶A。为 β -氧化所需的酶活性是已知的。本发明提供相较于对照宿主细胞,具有的对于中链(C4-C8)和长链(>C8)脂肪酸、羟基脂肪酸和二酸的分解代谢路径活性增加的宿主细胞。举例来说,在酵母(例如酿酒酵母)中, β -氧化发生在过氧化物酶体中;影响过氧化物酶体 β -氧化的示例性非限制性核酸产物是酿酒酵母PAT1和PEX11。在本发明的一些实施方案中,提供用于供产生丙二酸使用的本文方法中的被修饰来达成PAT1和/或PEX11的表达增加的宿主细胞。这个修饰说明于实施例14中。

[0130] 增加丙二酰基-辅酶A生物合成的遗传修饰和发酵条件

[0131] 根据本发明,丙二酸效价、产率和/或生产力增加可通过增加丙二酰基-辅酶A生物合成来实现,且本发明提供与此相关的宿主细胞、载体、酶和方法。丙二酰基-辅酶A是通过乙酰基-辅酶A羧化酶(EC6.4.1.2)催化自乙酰基-辅酶A和二氧化碳形成丙二酰基-辅酶A的活性在宿主细胞中产生。本发明提供用于产生丙二酸的表达异源乙酰基-辅酶A羧化酶(ACC)的重组宿主细胞。在一些实施方案中,宿主细胞是包含异源酿酒酵母乙酰基-辅酶A羧化酶ACC1或与其同源的酶的酿酒酵母细胞。在一些实施方案中,被修饰来异源表达如酿酒酵母ACC1的ACC的宿主细胞被通过对酿酒酵母SNF1蛋白质激酶或与其同源的酶进行遗传修饰来进一步修饰以消除ACC1翻译后调控。本发明也提供一种适于根据本发明产生丙二酸的重组宿主细胞,其为一种包含编码大肠杆菌乙酰基-辅酶A羧化酶复合蛋白AccA、AccB、AccC和AccD或一种或多种与其同源的酶的表达的异源核酸的大肠杆菌细胞。根据本发明,可异源表达其它乙酰基-辅酶A羧化酶以增加丙二酰基-辅酶A和丙二酸生物合成。这个修饰说明于实施例15中。

[0132] 在本发明的各种实施方案中,BirA(即生物素-[乙酰基辅酶A羧化酶]全酶合成酶)的表达是与大肠杆菌乙酰基-辅酶A羧化酶复合蛋白AccA、AccB、AccC和AccD共表达以增强ACC复合物的活性且导致丙二酰基-辅酶A和丙二酸产生得以增加。在本发明的各种实施方案中,通过在以下残基的任一者、全部或任何组合处引入丝氨酸至丙氨酸突变来进一步修饰酿酒酵母ACC1以消除ACC1翻译后调控:S10、S233、S430、S1114、S1145、S1148、S1157、S1159、S1162、S1163、S1169。在本发明的一些实施方案中,所用乙酰基-辅酶A羧化酶来自解脂耶罗威亚酵母CLIB122,在本文中称为Y1ACC。实施例42说明如何在由本发明提供的丙二酸产生性宿主菌株中包括这个酶会相对于在无这个酶下发酵的相同宿主,导致在发酵120小时之后,来自那个宿主的丙二酸的效价加倍。在本发明的其它实施方案中,这个酶是与也源于这个生物体的生物素-[乙酰基-辅酶A羧化酶]全酶合成酶(BPL1)共表达。在本发明的其它实施方案中,乙酰基-辅酶A羧化酶和生物素-[乙酰基辅酶A羧化酶]全酶合成酶编码基因是dtsR1accBC且源于谷氨酸棒杆菌(*Corynebacterium glutamicum*)。在本发明的其它实施方案中,这些基因源于包括但不限于以下各属的酵母菌株的酵母菌株:假丝酵母属、毕赤酵母属或本文任何其它酵母。在本发明的各种实施方案中,产生丙二酸的宿主细胞表达这些乙酰基-辅酶A羧化酶和生物素-[乙酰基-辅酶A羧化酶]全酶合成酶的任何组合。

[0133] 在本发明的一些实施方案中,适于根据本发明的方法产生丙二酸的宿主细胞包含

影响脂肪酸生物合成中涉及的蛋白质的表达和/或活性的遗传修饰。丙二酰基-辅酶A是脂肪酸的生物合成中的天然底物,且使丙二酰基-辅酶A转向脂肪酸产生会降低宿主细胞产生丙二酸的能力。本发明提供用于产生丙二酸的表达异源脂肪酸合成酶(FAS)多酶复合物的重组宿主细胞。酿酒酵母脂肪酸合成酶复合物的温度敏感型突变是已知的(参见Knobling等, *Eur. J. Biochem.*, 59:415-421 (1975))。异源温度敏感型脂肪酸合成酶复合物的表达会使丙二酰基-辅酶A转向将受培养宿主细胞所处的温度控制的脂肪酸生物合成。在一些实施方案中,宿主细胞是包含酿酒酵母脂肪酸合成酶FAS1和FAS2或与其同源的酶的酿酒酵母细胞。在本发明的一些实施方案中,FAS1和FAS2酶是温度敏感型FAS1或FAS2酶。

[0134] 除对宿主细胞的遗传修饰之外,可通过向细胞培养基中添加FAS抑制剂来降低脂肪酸生物合成。举例来说,FAS抑制剂浅蓝菌素(cerulenin)与FAS复合物的酿酒酵母酮酰基合成酶结构域中的活性位点半胱氨酸C1305形成共价键,从而抑制酶活性(Johansson等, *PNAS*, 105:12803-12808 (2008))。浅蓝菌素不仅有效抑制酿酒酵母FAS活性,而且通常是在酮酰基合成酶结构域中含有Cys-His-His或Cys-His-Asn催化三联体的FAS复合物的抑制剂。在一些实施方案中,添加浅蓝菌素至发酵液中以获得在5mg/l与100mg/l之间的最终浓度来抑制脂肪酸生物合成且增加根据本发明的方法在产生丙二酸的重组宿主细胞中产生丙二酸。在本发明的方法的各种实施方案中,添加FAS抑制剂至含有产生丙二酸的重组宿主细胞的发酵液中。在本发明的方法的一些实施方案中,FAS抑制剂是如实施例16中所说明的浅蓝菌素。在本发明的方法的一些实施方案中,浅蓝菌素是在5mg/l与100mg/l之间的浓度下补充于发酵液中。在本发明的方法的其它实施方案中,脂肪酸合成酶复合物抑制剂选自自由平板霉素(platensimycin)、硫乳霉素(thiolactomycin)和三氯生(triclosan)组成的组。

[0135] 乙酰基-辅酶A羧化酶的一种底物是二氧化碳,且增加发酵液中的二氧化碳分压会促进丙二酰基-辅酶A的形成。发酵液应含有最小溶解二氧化碳压力0.01大气压,且增加溶解二氧化碳分压高于这个阈值是合乎需要的。发酵液应通常含有0.1大气压与1大气压之间的溶解二氧化碳分压。发酵液中的溶解二氧化碳分压可增加至高于饱和条件,或高于1大气压溶解二氧化碳。在本发明的方法的一些实施方案中,使发酵液中的溶解二氧化碳分压增加至0.1大气压与1大气压之间。在本发明的方法的一些实施方案中,二氧化碳分压是通过向发酵液中添加碳酸盐或碳酸氢盐来增加。举例来说且不加限制地,可添加碳酸钙至发酵液中以增加溶解二氧化碳分压。在本发明的方法的其它实施方案中,发酵是在含有处于高于大气压下的二氧化碳的加压容器中运行。在本发明的方法的其它实施方案中,向发酵液中喷射二氧化碳气体。如果添加的组分不干扰宿主细胞生长或丙二酸产生,那么所喷射的气体混合物可含有其它气体。可为有利的是共定位二氧化碳气源与丙二酸发酵。举例来说且不加限制地,可将由各种发酵过程(例如乙醇、异丁醇、3-羟基丙酸等)、化学过程(例如下游丙二酸合成化学)或能量产生(例如煤或天然气发电厂)产生的气态二氧化碳泵送至来自丙二酸产生性宿主细胞的发酵液中以增加二氧化碳分压,如实施例17中所说明。

[0136] 降低丙二酸分解代谢的遗传修饰

[0137] 根据本发明,丙二酸效价、产率和/或生产力增加可通过降低丙二酸分解代谢来实现,且本发明提供与此相关的宿主细胞、载体、酶和方法。丙二酸在宿主细胞中被分解代谢所通过的一种代谢路径是通过酰基-辅酶A合成酶催化丙二酸和辅酶A转化成丙二酰基-辅

酶A的活性来达成。在本发明的一些实施方案中,适于根据本发明的方法产生丙二酸的重组宿主细胞包含导致一种或多种编码酰基-辅酶A合成酶的核酸的缺失、减弱或修饰的遗传修饰。在本发明的一些实施方案中,重组宿主细胞是酵母且一种或多种酰基-辅酶A合成酶选自由FAA1、FAA2、FAA3、FAA4、LSC1和LSC2组成的组。在本发明的其它实施方案中,重组宿主细胞是大肠杆菌且一种或多种酰基-辅酶A合成酶选自由FadD、FadK、FadI、SucC、SucD和YahF组成的组。本发明的这个方面说明于实施例18中。

[0138] 增加丙二酸自宿主细胞分泌的遗传修饰

[0139] 根据本发明,丙二酸效价、产率和/或生产力增加可通过增加丙二酸转运至发酵液中来实现,且本发明提供与此相关的宿主细胞、材料和方法。在本发明的一些实施方案中,适合用于本发明的方法中的重组宿主细胞是包含编码酿酒酵母转运蛋白的表达的异源核酸的酿酒酵母细胞,所述转运蛋白选自由PDR5、PDR10、PDR11、PDR12、PDR15和PDR18组成的组。在本发明的一些实施方案中,适于根据本发明的方法产生丙二酸的重组宿主细胞是包含编码大肠杆菌DcuC的表达的异源核酸的大肠杆菌细胞。本发明的这个方面说明于实施例19中。

[0140] 增加宿主细胞对丙二酸的耐受性的遗传修饰

[0141] 根据本发明,丙二酸效价、产率和/或生产力增加可通过增加宿主细胞对丙二酸的耐受性来实现,且本发明提供与此相关的宿主细胞、材料和方法。高浓度的丙二酸可竞争性抑制丁二酸脱氢酶(EC1.3.5.1)活性(参见Slater,Methods Enzymol.10:48-57(1967))。本发明部分地基于发现突变丁二酸脱氢酶展现受丙二酸的竞争性抑制较低。举例来说,酿酒酵母丁二酸脱氢酶SDH1残基E300、R331和R442对底物(例如丁二酸)识别是重要的。增加SDH1活性位点的尺寸会降低由丙二酸进行的竞争性抑制,同时仍然允许酶维持对天然底物丁二酸的活性。具体来说,引入一个或多个选自由E300D、R331K或R331H、以及R442K和R442H组成的组的突变会降低由丙二酸对SDH1的竞争性抑制。在一些实施方案中,表达具有点突变R300D的SDH1的重组宿主细胞用于根据本发明产生丙二酸。在其它实施方案中,表达具有点突变R331K或R331H的SDH1的重组宿主细胞用于根据本发明产生丙二酸。在其它实施方案中,表达具有点突变R442K或R442H的SDH1的重组宿主细胞用于根据本发明产生丙二酸。本发明的这个方面说明于实施例20中。

[0142] 增加各种碳源的分解代谢的遗传修饰

[0143] 在本发明的方法中,碳原料用于产生丙二酸。适合碳源不加限制地包括选自由以下组成的组的碳源:纯化糖(例如右旋糖、蔗糖、木糖、阿拉伯糖、乳糖等);植物源性混合糖(例如甘蔗、甜高粱、糖蜜、玉米淀粉、马铃薯淀粉、甜菜糖、小麦等)、植物油、脂肪酸、甘油、纤维素物质、海藻酸盐、乙醇、二氧化碳、甲醇和合成气体(“合成气”)。给定宿主细胞可高效或无效分解代谢特定原料。如果宿主细胞无效分解代谢原料,那么可修饰宿主细胞以增强或创建针对那个原料的分解代谢路径。举例来说,实施例47显示使用库德里阿兹威毕赤酵母作为宿主菌株,根据本发明自许多不同碳源产生丙二酸。尽管这个生物体在所测试的条件下不分解代谢蔗糖,但本发明提供引入蔗糖转化酶(β -呋喃果糖苷酶,EC 3.2.1.26)以有助于使用这个碳源。实施例30也说明使用多种碳源,根据本发明的方法产生丙二酸。本发明的其它实施方案包括使用甲醇分解代谢性宿主菌株。在一些实施方案中,宿主是酵母菌株。在一些实施方案中,宿主选自*Komagataella pastoris*、甲醇毕赤酵母或巴斯德毕赤酵

母。

[0144] 本发明提供包含通过增加分解代谢非天然碳源的能力来增加丙二酸效价、产率和/或生产力的遗传修饰的宿主细胞。野生型酿酒酵母细胞不能分解代谢戊糖、木质纤维物质或海藻酸盐原料。在一些实施方案中,本发明提供一种包含编码使得能够分解代谢用于如本文所述产生丙二酸的戊糖的酶的异源核酸的酿酒酵母细胞。在其它实施方案中,异源核酸编码使得能够分解代谢木质纤维原料的酶。在本发明的其它实施方案中,异源核酸编码增加海藻酸盐原料的分解代谢的酶。

[0145] 本领域技术人员将认识到用以增加自给定宿主细胞产生丙二酸的个别操作可以实际上无限制组合加以使用且常赋予导致丙二酸输出量比单一操作以及在多种情况下个别操作的总和高得多的促进益处。因此,本发明不仅涵盖本文所述的单一操作,它也由这些扰动的导致宿主菌株产生的丙二酸增加的任何组合加以体现。

[0146] 章节6:检测丙二酸产生性宿主细胞以及筛选丙二酸产生得以改进的宿主细胞

[0147] 本发明也提供一种可用于准确感测液体培养基中的丙二酸的转录因子生物感测器系统。在许多应用中,这个系统用于通过感测含有丙二酸产生性宿主细胞的发酵培养基中的丙二酸的存在来检测宿主细胞内产生的丙二酸。“准确感测”是指检测丙二酸的存在和/或直接或间接测定丙二酸的浓度;因此,通过准确感测丙二酸,本发明的这个方面可应用于菌株改良,即用于增加宿主细胞中的丙二酸产生。在这个系统中,丙二酸结合存在于转录因子上的蛋白质部分。丙二酸结合丙二酸结合部分导致转录因子结合由转录因子活化的启动子,或在一些实施方案中,导致启动子由丙二酸结合的转录因子去阻遏。

[0148] 结合丙二酸,或反应于由丙二酸结合结合部分产生的信号而被活化来结合启动子的许多转录因子适合用于这个系统中。

[0149] 在一些实施方案中,转录因子可结合丙二酸,此导致转录因子结合同源启动子且活化可操作地连接于所述启动子的基因。结合丙二酸的转录因子的非限制性实例包括转录因子乙酸钙不动杆菌(*Acinetobacter calcoaceticus*) MdcY (SEQ ID NO:3)、豆科植物根瘤菌(*Rhizobium leguminosarum*) MatR (SEQ ID NO:4)、肺炎克雷伯氏菌(*Klebsiella pneumoniae*) MauR (SEQ ID NO:5) 及其同源物。

[0150] 在一些实施方案中,本发明中使用的转录因子是MdcY转录因子。MdcY转录因子可直接结合丙二酸且调控由如 P_{MdcL} 的启动子介导的转录。MdcY和MdcY反应性启动子在本领域中是已知的(参见例如Koo等,J Bacteriol.182:6382-6390 (2000))。

[0151] 来自乙酸钙不动杆菌的MdcY多肽序列的实例提供于SEQ ID NO:3中。可根据本发明采用的MdcY多肽包括SEQ ID NO:3中阐述的MdcY多肽序列的变体和同源物。因此,MdcY转录因子多肽可具有优选历经至少100个或更多个氨基酸、或至少200个或更多个氨基酸的区域,或历经整个多肽的长度,与氨基酸序列SEQ ID NO:3具有至少60%同一性、通常至少75%、90%、95%、99%或更高的氨基酸序列同一性的氨基酸序列。本领域技术人员鉴于本公开了解的是也可采用变体,例如使用已知序列作为选择将不导致功能损失的氨基酸取代的指导。

[0152] 在一些实施方案中,用于本发明中的MdcY转录因子天然存在于宿主细胞中。在其它实施方案中,通过将编码丙二酸转录因子的异源核酸引入宿主细胞中来遗传修饰宿主细胞以表达外来转录因子。在一些实施方案中,包含编码丙二酸转录因子的异源核酸的遗传

修饰的宿主细胞是大肠杆菌宿主细胞。

[0153] MdcY转录因子可结合许多启动子且活化可操作地连接于启动子的基因的表达。适于根据本发明使用的MdcY反应性启动子的实例提供于SEQ ID NO:6中。在一些实施方案中,用于本发明中的MdcY反应性启动子通常包含操纵子序列ATTGTATACAAT。在一些实施方案中,启动子与SEQ ID NO:6中所示的启动子序列至少75%相同。在一些实施方案中,启动子包含SEQ ID NO:6的子序列,所述子序列包含SEQ ID NO:6的10、20、25、30、35个或更多个连续核苷酸。

[0154] 因此,在一个方面,本发明提供一种用于准确感测丙二酸的方法。在一些实施方案中,感测步骤包括检测(例如测量)报道基因或标记基因(例如荧光报道基因)的基因产物的量。在一些实施方案中,报道基因的基因产物影响包含本发明的丙二酸转录因子生物感测器的组分的宿主细胞的生长速率。在一些实施方案中,报道基因的基因产物导致修饰的宿主细胞变得对化合物具有抗性或敏感性。举例来说,在一些实施方案中,报道基因是抗生素抗性基因(例如tetA基因),其中在培养基中存在丙二酸会诱导抗生素抗性,以使当抗生素存在时,宿主细胞在丙二酸存在下展现生长改进,如实施例21中所说明。在一些实施方案中,包含本发明的转录因子生物感测器的组分的宿主细胞是能够产生丙二酸的宿主细胞。实施例36说明使用抗生素抗性(tetA)与比色信号(lacZ)两者作为与发酵培养基中根据本发明的其它方面产生的丙二酸浓度相关联的输出来实施本发明的这个方面。

[0155] 通常,接着,本发明提供一种用于筛选或选择产生丙二酸的宿主细胞的方法,其包括:(a)提供本发明的修饰的宿主细胞,(b)培养所述宿主细胞,以及(c)基于由所述宿主细胞对报道基因的表达来筛选或选择所述宿主细胞。

[0156] 在本发明的一些实施方案中,用于筛选或选择产生丙二酸的宿主细胞的方法包括:(a)提供本发明的多种修饰的宿主细胞,其中具有不同修饰的所述修饰的宿主细胞是单独培养的,(b)培养宿主细胞的各单独培养物,(c)基于由所述宿主细胞对报道基因的表达来筛选或选择所述宿主细胞,以及(d)比较所述单独培养物的报道基因的所述表达。在本发明的一些实施方案中,步骤(d)包括鉴定报道基因的基因产物的表达增加的一个或多个培养物和/或相应宿主细胞。将由本领域技术人员认识到的是这些培养物由单一转化体产生且表示克隆性群体。

[0157] 在一些实施方案中,本发明的方法是一种用于选择产生丙二酸的宿主细胞的方法,其中所述选择是阳性选择。在阳性选择中,选择步骤选择使保持活力和倍增的概率增加的报道基因的表达较高,且因此保持活力和倍增的概率较高的宿主细胞。举例来说,产生丙二酸的宿主细胞将保持活力且在快于不产生丙二酸的宿主细胞的速率下增殖。类似地,相较于对照菌株,丙二酸产生得以增加的宿主细胞将在快于对照菌株的速率下增殖。

[0158] 在本发明的一些实施方案中,用于选择产生丙二酸的大肠杆菌宿主细胞的方法包括:(a)提供本发明的多种修饰的大肠杆菌宿主细胞,其中具有不同修饰的所述修饰的宿主细胞是单独培养的,(b)培养宿主细胞的各单独培养物,(c)基于由所述宿主细胞对报道基因的表达来阳性选择所述宿主细胞,以及(d)比较所述单独培养物的报道基因的所述表达。

[0159] 在本发明的其它实施方案中,用于选择产生丙二酸的大肠杆菌宿主细胞的方法包括:(a)提供本发明的多种修饰的大肠杆菌宿主细胞,其中具有不同修饰的所述修饰的宿主细胞是同时培养的,(b)在含有阳性选择剂的生长培养基中培养修饰的宿主细胞的异质混

合物以使展现丙二酸的产生高于多种宿主细胞的宿主细胞将在快于展现丙二酸的产生较低的宿主细胞的速率下增殖,以及(d)分离展现丙二酸的产生最高的宿主细胞。

[0160] 章节7:丙二酸纯化

[0161] 在第六方面,本发明提供用于自发酵液纯化丙二酸的方法,所述方法包括:(a)在适于产生丙二酸的条件下培养宿主细胞,(b)以及自发酵液回收(即纯化)丙二酸。本发明也提供根据本发明的方法产生的纯化丙二酸。生物合成的丙二酸可在细胞内产生和/或分泌至培养基中。细胞内产生的丙二酸通常是使用如上所述的膜转运体分泌至培养基中。未分泌的丙二酸可通过化学、酶促或机械细胞溶解自宿主细胞移除。可自细胞、发酵液或两者回收丙二酸。如果细胞被工程化来分泌丙二酸,那么可选择仅自发酵液回收丙二酸或可选择自培养基与细胞(即通过使细胞溶解)两者将它回收。如果细胞未被工程化来分泌丙二酸,那么可使宿主细胞溶解以分离其中的丙二酸。

[0162] 本发明提供用以分离生物产生的丙二酸的方法。如本文所用,“分离”、“纯化”和“回收”用于指代使丙二酸与存在的其它物质分离。如这个情形下使用的“分离”、“纯化”或“回收”意图表达制备相对于产生丙二酸,但基于重量/重量(w/w)可或可不大致上(即50%以上)纯净的细胞或发酵液,富含丙二酸的丙二酸。根据这些方法分离丙二酸涉及使产生的丙二酸与产生丙二酸的发酵培养基、宿主细胞及其各部分的至少一部分或全部分离。丙二酸可根据本发明自发酵液和/或产生性细胞加以纯化,即基于w/w达到50%以上纯度,其中使产生丙二酸的任何天然存在或重组宿主细胞(例如大肠杆菌、酿酒酵母、油脂酵母等)生长,即宿主细胞不限于本发明的重组宿主细胞。分离的丙二酸可不含或基本上不含来自宿主细胞的杂质。丙二酸是在使存在的任何杂质都不干扰丙二酸的随后使用的程度上被分离或纯化。举例来说,如果随后使用是作为工业化学品(如待用于聚合反应中的化学品),那么丙二酸应基本上不含杂质,此时任何剩余杂质都将不干扰丙二酸在聚合反应中的使用。通常,用于聚合反应的丙二酸具有至少95%w/w或更高的纯度。如果丙二酸将用作燃料(如待用于燃烧反应中的燃料),那么化合物应基本上不含杂质,此时任何剩余杂质都将不干扰丙二酸作为燃料的使用。如果丙二酸用作动物饲料,那么丙二酸应基本上不含杂质,此时任何剩余杂质都将不干扰所述物质作为动物饲料的使用。当丙二酸用作动物饲料时,可选择自发酵液回收含有丙二酸的生物质且使用生物质作为动物饲料。

[0163] 在本发明的纯化方法的一些实施方案中,浓缩发酵液以增加丙二酸的工作浓度且降低需要处理的液体体积。在本发明的纯化方法的各种实施方案中,这个浓缩是通过蒸发(包括真空和加热)、反渗透、“高通”膜脱水和/或薄膜蒸发来实现。

[0164] 在一些实施方案中,本发明的纯化方法包括回收产生的丙二酸的步骤,其中所述回收步骤是与培养步骤同时或在培养步骤之后。在一些实施方案中,丙二酸是自发酵液和宿主细胞纯化。在其它实施方案中,自发酵液分离宿主细胞,溶解,且接着自宿主细胞回收丙二酸。在其它实施方案中,溶解发酵液中的宿主细胞且自溶解的细胞和发酵液回收丙二酸。一种用于自由本发明提供的发酵液回收丙二酸的方法是用阳离子使丙二酸沉淀。在一些实施方案中,这是单价阳离子,在其它实施方案中,它是二阶阳离子。通常,阳离子是以盐形式添加至发酵液(或溶解产物)中。举例来说,根据本发明使丙二酸钙自可为发酵液或细胞溶解产物或两者的混合物的含有丙二酸的水溶液沉淀是通过添加钙盐来达成(Weiner, Org.Synth.18:50(1938);Weiner,Org.Synth.Coll.2:376(1943))。各种钙盐(例如氢氧化

钙、碳酸钙、氯化钙)可根据本发明用于使丙二酸自发酵液沉淀。

[0165] 当两个丙二酸羧酸未质子化且钙存在于发酵液中时,形成丙二酸钙盐。丙二酸钙不溶于例如发酵液的水溶液中,将沉淀,且可接着自发酵液回收。两个丙二酸羧酸部分的 pK_a 值是2.83和5.69;因此,当发酵液pH低于5.7时,不形成丙二酸钙且丙二酸仍然溶解于发酵液中。当发酵液pH高于5.7时,可形成丙二酸钙且将自发酵液沉淀。羧酸是弱酸,且丙二酸在发酵液中累积将降低pH,且如果pH下降至5.7以下,那么添加钙阳离子不导致形成丙二酸钙。本发明的一种用以自发酵液纯化丙二酸的方法在于添加钙盐,其中钙盐阴离子是碱(即碳酸钙和氢氧化钙)且添加钙盐会升高和/或维持发酵液pH高于5.7。在本发明的方法的一些实施方案中,添加氢氧化钙至发酵液中以达到/维持pH在5.69与7.5之间且使丙二酸钙沉淀。在本发明的方法的第二实施方案中,添加碳酸钙至发酵培养基中以达到/维持pH在5.69与7.5之间且使丙二酸钙沉淀。在本发明的第三实施方案中,当发酵液pH在5.69与7.5之间时,添加氯化钙至发酵液中且使丙二酸钙沉淀。根据本发明的方法,其它钙盐可添加至发酵液中以使丙二酸沉淀。添加碳酸钙具有增加发酵液中的二氧化碳分压,从而促进通过上述乙酰基-辅酶A羧化酶的活性形成丙二酰基-辅酶A的额外优势。实施例23通过描述使用钙盐纯化丙二酸来说明本发明的这个方面。

[0166] 可在发酵开始时(即在已累积大量丙二酸之前)、在发酵期间、或在发酵结束时添加钙盐。在一些实施方案中,碳酸钙或氢氧化钙是在宿主细胞已开始产生丙二酸之前添加至发酵液中。在其它实施方案中,碳酸钙或氢氧化钙是在丙二酸浓度超过5、10、15、20、25、30、40或50g/l之前添加至发酵液中。在其它实施方案中,碳酸钙或氢氧化钙是在发酵结束时添加至发酵液中。可选择与发酵同时或在发酵之后自发酵液回收丙二酸钙沉淀。

[0167] 因为丙二酸钙单水合物溶解度随温度增加而降低,所以可通过在回收步骤期间使温度升高至50℃与100℃之间来增加自发酵液回收丙二酸。在本发明的方法的一些实施方案中,添加钙盐(即碳酸钙、氢氧化钙和/或氯化钙)至发酵液中,使温度增加至50℃与100℃之间,且自发酵液回收丙二酸。在本发明的其它实施方案中,使温度增至高于100℃。然而,在约100℃的温度下可发生丙二酸的热分解,因此,使温度增加至100℃或高于100℃可负面影响产率。

[0168] 在本发明的一些实施方案中,丙二酸是通过用单价阳离子沉淀来自产生培养基纯化。通常,单价阳离子是以盐形式添加至发酵液(或溶解产物)中。举例来说,根据本发明使丙二酸钠自可为发酵液或细胞溶解产物或两者的混合物的含有丙二酸的水溶液沉淀是通过添加钠盐来达成。各种钠盐(例如氢氧化钠、碳酸钠、氯化钠)可根据本发明用于使丙二酸自发酵液沉淀。当两个丙二酸羧酸未质子化且钠存在于发酵液中时,形成丙二酸钠盐。丙二酸钠可溶于例如发酵液的水溶液中,但可通过使培养基的pH增至高于丙二酸的较高 pK_a (5.69)来压出溶液;参见实施例32。丙二酸钠的额外沉淀是通过蒸发溶剂(水)以增加离子的浓度来实现。如同本发明的与丙二酸钙沉淀相关的方面一样,添加含有碱的钠盐(例如氢氧化钠或碳酸钠)可用于维持发酵液的pH高于5.7,由此促进钠盐沉淀。根据本发明自发酵液分离沉淀会提供纯化丙二酸。

[0169] 本发明的用于自发酵液分离丙二酸的另一方法是通过用脂族伯胺、仲胺或叔胺或伯醇、仲醇或叔醇进行反应性提取来达成。这个方法部分地起因于发现伯醇与叔胺两者均为高度有效用于自发酵液选择性移除丙二酸的试剂,如实施例24-26、36-41中所说明。在本

发明的一些实施方案中,丙二酸是通过用三辛胺在有机溶剂中进行反应性提取来自发酵液纯化。在本发明的其它实施方案中,丙二酸是使用选自以下组成的组的叔胺自发酵液纯化:三乙胺、三丙胺、三丁胺、三戊胺、三己胺、三庚胺、三壬胺和三癸胺。在本发明的其它实施方案中,胺化合物是氨。在其它实施方案中,它是二烷基胺,包括但不限于二乙胺、二丙胺、二异丙胺、二丁胺、二戊胺、二己胺、二庚胺、二辛胺、二壬胺、二癸胺。实施例33说明使用二乙胺自发酵液分离丙二酸。在本发明的各种实施方案中,氨基化合物是连同包括但不限于以下的有机溶剂或溶剂的混合物一起添加:戊醇、己醇、庚醇、辛醇、壬醇、癸醇、己烷、癸烷、煤油、醚、乙酸乙酯或其任何组合。在各种实施方案中,溶剂是含有4至20个碳的分支或未分支烷、醇、二醇、酯、醚、二醚或内酯。在各种实施方案中,添加溶剂有助于将氨基化合物转移至水相中以及将胺配位的丙二酸转移至有机相中,由此改进纯化。

[0170] 许多参数可根据本发明加以调整以改进丙二酸提取。烷基的长度、发酵液的离子强度、pH、三烷基胺的摩尔当量、温度和共溶剂各自对三烷基胺的提取效率具有影响。烷基长度的影响说明于实施例36中。三丙胺、三己胺和三辛胺各自能够使丙二酸自水相分离且使它根据本发明溶解至1-辛醇有机相中。提取效率三辛胺>三己胺>三丙胺指示较长的这些链长度是在低pH下,在1-辛醇共溶剂下,本发明的优选实施方案(针对高产率而言)。

[0171] 如同盐沉淀一样,当使用三烷基胺时,pH对自水相提取丙二酸具有影响。当两个酸部分均被羧化时,三烷基胺提取具有最高效率。实施例37显示历经一定pH值范围的三辛胺提取效率。最高提取水平发生在pH 1.5下且当pH增加时减弱。当高于丙二酸的第二个pKa (5.69)时,实际上不存在提取。尽管本发明提供使用三烷基胺在不限制pH下自水相提取丙二酸,但在pH 2.0或低于2.0下的实施方案为最大效率所优选。在本发明的一些实施方案中,如由实施例48所说明的在低pH下的宿主细胞生长和丙二酸产生有助于使丙二酸质子化以增强用三烷基胺进行的提取。在本发明的其它实施方案中,添加外源性酸以有助于这个质子化。

[0172] 在本发明的各种实施方案中,使用不同浓度的三烷基胺。如实施例38中所报道,观察到提取效率与三辛胺:丙二酸摩尔分数之间的线性关系;具体来说,线性关系存在于三辛胺:丙二酸摩尔分数0至1之间。在高于1摩尔比下,100%的丙二酸被提取至有机相中。因此,为使提取效率最大化,有机相中的三烷基胺的量必须与水相(例如发酵液)中的丙二酸的量等摩尔。理想地,将添加大于等摩尔量的三烷基胺至有机相中以补偿归因于发酵液中的其它有机酸和阴离子的提取效率降低。

[0173] 如实施例39中所说明,通过增加水相的离子强度(盐浓度),三烷基胺对丙二酸的提取效率受负面影响。可通过许多方法实现各种离子的脱盐,所述方法包括:添加其它试剂以产生不溶性盐、pH调整和/或穿过各种膜。

[0174] 在本发明的各种实施方案中,反应性提取是在酸催化剂存在下,丙二酸在一个或两个羧酸处与醇进行费歇尔(Fischer)酯化以分别导致形成丙二酸单烷基酯和丙二酸二烷基酯。举例来说,丙二酸二乙酯可通过使用乙醇使丙二酸进行费歇尔酯化来产生。丙二酸单烷基酯和丙二酸二烷基酯在水溶液中具有低溶解度和低挥发性,从而使得能够通过蒸馏使丙二酸酯与水及发酵液中的其它挥发性组分分离。在本发明的各种实施方案中,酯的分离是通过向有机溶剂中的相分配来实现。一般来说,伯或仲脂族醇适于与丙二酸进行费歇尔酯化;举例来说且不加限制地,醇可尤其选自以下组成的组:甲醇(CAS号67-56-1)、乙醇

(CAS号64-17-5)、1-丙醇(CAS号71-23-8)、1-丁醇(CAS号71-36-3)、异丙醇(CAS号67-64-0)和异丁醇(CAS号78-83-1)。许多酸催化剂适于醇与丙二酸的费歇尔酯化;举例来说且不加限制地,酸催化剂可选自由以下组成的组:硫酸、对甲苯磺酸、三氟甲磺酸铊(III)、N,N'-二环己基碳化二亚胺和三溴化四丁铵。在本发明的各种实施方案中,丙二酸是通过使用乙醇进行反应性蒸馏以使丙二酸转化成丙二酸二乙酯来自发酵液纯化。在本发明的其它实施方案中,丙二酸是通过使用甲醇进行反应性蒸馏以使丙二酸转化成丙二酸二甲酯来自发酵液纯化。在本发明的各种实施方案中,酸催化剂是硫酸。

[0175] 如实施例41中所说明,使用硫酸作为催化剂,与乙醇进行费歇尔酯化可导致自水相纯化>99%的丙二酸。在本发明的各种实施方案中,这个过程中使用的过量乙醇是在减压下回收且再循环以用于随后发酵批次中。在本发明的各种其它实施方案中,催化剂是基于基质的(例如安伯来特(amberlite)树脂)且被回收、再装载以及再使用于随后批次中。

[0176] 通过费歇尔酯化来纯化丙二酸的经济状况可通过共定位丙二酸生物精制与醇生物精制或化学精制(在本文中用于指代使用合成化学而非发酵过程自可再生或石油原料产生醇)来根据本发明加以改进。在本发明的一些实施方案中,使丙二酸生物精制与醇生物精制或化学精制共定位以降低丙二酸单烷基酯或丙二酸二烷基酯合成的成本。在本发明的一些实施方案中,由生物精制或化学精制产生的醇具有的纯度低于通常贩卖和销售的用于燃料和/或化学应用的醇。

[0177] 在其它实施方案中,用于自发酵液分离丙二酸的本发明的方法涉及使用有机溶剂进行液-液提取。两个丙二酸羧酸部分的 pK_a 值是2.83和5.69;因此,当培养基的pH降低至2.8以下时,大于50%的分子不电离。这些物质对有机溶剂的亲合力高于对水性发酵液的亲合力。使这些分子分离至有机相中用于驱动平衡朝向质子化物质且因此渐进进入有机溶剂中。在本发明的一些实施方案中,发酵液的酸性pH是发酵条件的结果。在本发明的其它实施方案中,通过添加酸使pH降低至适当酸性。在本发明的各种实施方案中,溶剂选自但不限于由以下组成的组:乙酸乙酯、二氯甲烷、二氯乙烷、癸烷、十二烷、己烷、辛醇、戊醇或其混合物。在本发明的一些实施方案中,在低pH下的宿主细胞生长和丙二酸产生有助于使丙二酸质子化以增强丙二酸在有机溶剂中的溶解度。在本发明的其它实施方案中,添加外源性酸以有助于这个质子化。在本发明的一些实施方案中,溶剂是通过真空蒸馏移除(完全或部分)。在本发明的各种实施方案中,如其它地方所述,使浓缩的丙二酸与醇酯化,且通过蒸馏纯化酯。

[0178] 在本发明的其它实施方案中,丙二酸是通过结合离子交换树脂来自发酵液移除。在本发明的各种实施方案中,树脂选自但不限于以下:**Lewatit® VP OC 1065**、**Lewatit® MP-64**(氯离子型)、**Toyoparl®**、**Dowex® 66**(游离碱)、**Amberlite® IRA-67**(游离碱)、**Amberlite® IRA-96**(游离碱)、**Amberjet® 4200**(氯离子型)、**Lewatit® MonoPlus M500**(氯离子型)、**Dowex® 1X8**(氯离子型)、**Amberlyst A26**(氢氧根型)、**Amberlite® IRA958**(氯离子型)。在本发明的各种实施方案中,丙二酸是使用pH或盐梯度自树脂洗脱。在本发明的其它实施方案中,这些树脂的一者或多者用于自浓缩的丙二酸移除杂质。

[0179] 尽管本发明的纯化方法可用作用于自发酵液纯化丙二酸的单步纯化方法,但两种或更多种纯化方法可根据本发明连续使用。举例来说且不加限制地,用氢氧化钙使丙二酸

沉淀可继之以丙二酸与乙醇的费歇尔酯化以及使用蒸馏进一步纯化所得丙二酸二乙酯。

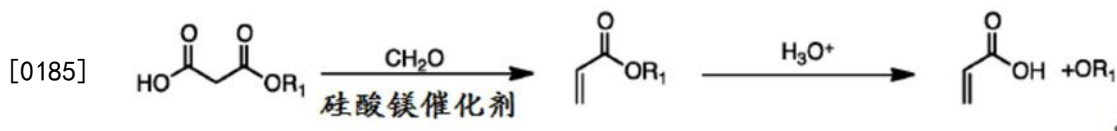
[0180] 章节8:使用丙二酸合成其它化学品的化学

[0181] 丙二酸是许多重要工业反应中的化学前体。本发明首次使得能够在这些反应中使用生物源性丙二酸。在利用丙二酸作为反应物的任何已知反应中,都可使用生物源性丙二酸。在使用丙二酸单烷基酯、丙二酸二烷基酯或梅尔德伦氏酸的任何反应中,都可在转化成适当反应基质之后使用生物源性丙二酸。此外,本发明提供丙二酸用作起始物质的新型化学合成方法。对于这些反应,可使用由任何手段产生的丙二酸,所述手段包括通过丙二酸半醛的酶催化氧化来生物产生丙二酸、生物或合成源性3-羟基丙酸的合成氧化、或生物或合成源性1,3-丙二醇的合成氧化。

[0182] 因此,除提供用于已知化学反应中的生物源性丙二酸之外,本发明提供使用丙二酸和丙二酸衍生化合物来产生其它适用化合物的新方法。

[0183] 本发明部分地归因于以下事实:丙二酸和丙二酸衍生化合物可用作通过与甲醛反应来合成丙烯酸酯(即丙烯酸酯、丙烯酸甲酯、丙烯酸乙酯、丙烯酸丙酯和丙烯酸丁酯)中的基质。丙烯酸酯在销售额方面占据每年十亿美元且用于范围宽泛达尿布至洗发精至薄膜和涂料的产品中。本发明提供用于通过使丙二酸或丙二酸二酯与甲醛反应来产生丙烯酸酯的两种碱性合成。在一种方法中,使丙二酸二酯与甲醛在适当碱(例如二乙胺)存在下反应以产生2-亚甲基丙二酸的二酯。随后酯皂化和加热导致产生1份丙烯酸、1份CO₂和2份醇。所用的丙二酸二酯的酯部分的组成、碱和溶剂可广泛变化。举例来说且不加限制地,可利用二甲酯、二乙酯、二丙酯、二丁酯、二异丙酯或二己酯。类似地,所用的碱可为许多碱中的一者。在各种实施方案中,碱是三取代的胺。适于根据本发明的方法使用的三取代的胺的非限制性实例包括选自含有哌啶、三甲胺、三乙胺、三丙胺、三异丙胺、三丁胺或更长链三烷基胺的组的三取代的胺。在本发明的这个方面的一些实施方案中,在克诺维纳盖尔(Knoevenagel)缩合中使丙二酸二乙酯与甲醛在吡咯烷中反应以形成丙烯酸酯且随后皂化并加热以产生1份丙烯酸、1份CO₂和2份乙醇。在本发明的方法的其它实施方案中,使丙二酸二甲酯与甲醛在二乙胺中反应以形成2-亚甲基丙二酸二甲酯,其随后被转化成丙烯酸、CO₂和甲醇。

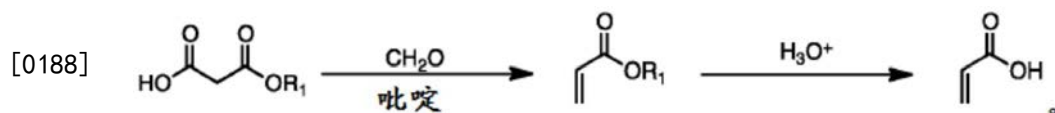
[0184] 在其它实施方案中,镁盐催化丙二酸二酯与甲醛缩合以形成2-亚甲基丙二酸二酯,其接着以所述方式被处理。举例来说且不加限制地,硅酸镁是适于这个反应的镁盐。在一些实施方案中,使丙二酸二乙酯与甲醛在硅酸镁催化剂存在下反应以形成2-亚甲基丙二酸二乙酯。在其它实施方案中,使丙二酸二异丙酯与甲醛在硅酸镁催化剂存在下反应以形成2-亚甲基丙二酸二异丙酯。一种使用丙二酸单酯的一般性流程说明于此:



[0186] 在这个情况下,R₁选自非限制性实例:H、CH₃、CH₃CH₂、CH₃(CH₂)₂、CH₃(CH₂)₃、(CH₃)₂CH₂或(CH₃)₂CH₂CH₂。

[0187] 在另一实施方案中,采用克诺维纳盖尔缩合的多布纳(Doebner)改进形式且使丙二酸与甲醛在吡啶中反应以产生丙烯酸酯。实施例28说明根据本发明的这个实施方案自丙二酸产生丙烯酸酯。在其它实施方案中,使丙二酸单甲酯与甲醛在吡啶中反应以形成丙烯酸甲酯。在此处说明的流程中,丙二酸单酯充当克诺维纳盖尔缩合的多布纳改进形式中的

起始物质



[0189] R_1 选自非限制性实例： H 、 CH_3 、 CH_3CH_2 、 $\text{CH}_3(\text{CH}_2)_2$ 、 $\text{CH}_3(\text{CH}_2)_3$ 、 $(\text{CH}_3)_2\text{CH}_2$ 或 $(\text{CH}_3)_2\text{CH}_2\text{CH}_2$ 。在其它实施方案中，使丙二酸单乙酯与甲醛在吡啶中反应以形成丙烯酸乙酯。

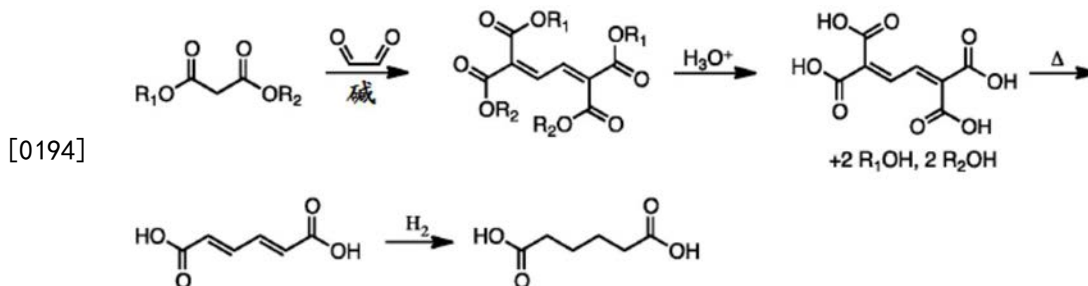
[0190] 在本发明的其它实施方案中，丙二酸和丙二酸衍生化合物可用作通过与各种基质反应来合成不饱和二羧酸中的基质，所述各种基质包括但不限于醛、烷基卤化物、二醛、烷基二卤化物、末端烯烃及以上的组（即双官能基质）。在丙二酸或丙二酸衍生化合物与醛反应的情况下，形成烯，其可随后被氢化以产生饱和产物。通常，丙二酸和丙二酸衍生化合物可用于合成C5-C12直链、饱和或不饱和二酸（即戊二酸（CAS 110-94-1）、己二酸（CAS 124-04-9）、庚二酸（CAS 111-16-0）、辛二酸（CAS 505-48-6）、壬二酸（CAS 123-99-9）、癸二酸（CAS 111-20-6）、十一烷二酸（CAS 1852-04-6）、十二烷二酸（CAS 693-23-2））和它们的相应二烷基酯。在本发明的各种实施方案中，丙二酸用作C5-C12直链饱和二酸的化学合成中的基质。在本发明的各种实施方案中，丙二酸或丙二酸衍生化合物用作戊二酸、己二酸、庚二酸、辛二酸、壬二酸、癸二酸、十一烷二酸或十二烷二酸的化学合成中的基质。

[0191] 在本发明的其它实施方案中，衍生自由或不由本发明提供的丙二酸的化合物用作戊二酸的化学合成中的基质。戊二酸可通过合成源性丙二酸二乙酯和甲醛（参见Ahluwalia等，“Organic reaction mechanisms,”第2版Alpha Science International:Harrow (2005)第340-341页）；合成源性丙二酸二乙酯和二氯甲烷（参见Perkin等,J Am Chem Soc 59:990-995 (1891)）；以及合成源性梅尔德伦氏酸和甲醛（参见Hedge等,J.Org.Chem.26:3166-3170 (1961)）的反应根据本发明来形成。

[0192] 在本发明的其它实施方案中，丙二酸或丙二酸衍生化合物用作己二酸的化学合成中的基质。己烷二酸是用于产生尼龙6,6的两种化合物之一且在每年原料销售额方面占据数十亿美元。尼龙6,6用于广泛范围的耐用消费品中，包括地毯、汽车气囊和绳索。己二酸可通过镁螯合的丙二酸单酯和1,2-二氯乙烷的反应根据本发明来形成。已描述二氯乙烷和丙二酸二甲酯的烷基化（美国专利号6,262,298）；然而，非所要产物环丙烷-1,1-二甲酸二甲酯是在99.3%纯度和83%理论产率下报道的唯一产物。环丙烷-1,1-二甲酸酯的形成由在1,2-二氯乙烷的初始烷基化之后的分子内烷基化所致。相反，本发明提供使用不允许对环丙烷的非所要分子内烷基化的镁螯合的丙二酸单酯的材料和方法。镁螯合的丙二酸单酯与1,2-二氯乙烷通过脱羧促进的烯醇产生加以反应；在脱羧之后不能进行第二分子内反应（即生成环丙烷二甲酸酯）。供根据本发明的这个实施方案使用的示例性催化剂是硅酸镁。本领域技术人员了解的是各种二卤代烷（例如1,3-二卤代丙烷、1,2-二卤代丙烷）也可与镁螯合的丙二酸单酯一起用于形成二羧酸产物且那些卤化物基团可为I、Cl、Br或F。举例来说且不加限制地，可使1,3-二卤代丙烷与镁螯合的丙二酸单酯反应以产生庚二酸。

[0193] 本发明也提供用于通过丙二酸和丙二酸衍生化合物与乙二醛的反应来产生己二酸的方法。通过丙二酸二酯和脂族单醛的克诺维纳盖尔反应形成烯酸是已知的（参见例如Rao等,J Am Oil Chem Soc,70:297-299 (1993)；以及Zhang等,Synth Comm,40:3093-3100 (2010)）。本发明部分地起因于发现可使二醛与丙二酸衍生化合物反应以产生相应1,3-二

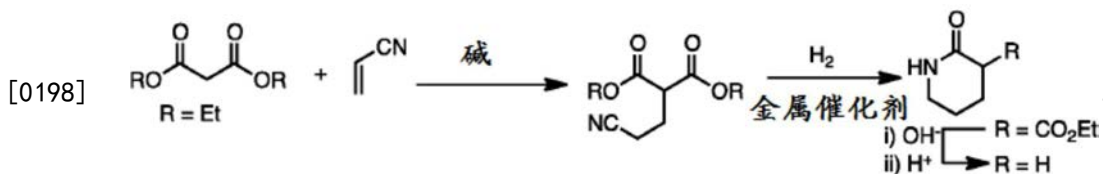
烯-1,1,4,4-四甲酸四酯。举例来说,丙二酸二乙酯与乙二醛反应产生丁-1,3-二烯-1,1,4,4-四甲酸四乙酯;类似地,丙二酸与乙二醛(草醛)反应产生己-2,4-二烯二酸。当需要时,通过向反应混合物中添加酸来进行酯基水解。热诱导脱羧以形成不饱和二羧酸(即己-2,4-二烯二酸),且随后使不饱和二羧酸氢化产生己二酸。在本发明的一些实施方案中,使丙二酸二乙酯和乙二醛缩合以产生二烷基化双丙二酸乙酯;随后皂化以水解酯基,继之以热诱导脱羧产生不饱和二羧酸。氢化不饱和二羧酸产生己二酸。这个方法概述于以下反应流程中,其中烷基(R_1 和 R_2)选自非限制性实例: H 、 CH_3 、 CH_3CH_2 、 $CH_3(CH_2)_2$ 、 $CH_3(CH_2)_3$ 、 $(CH_3)_2CH_2$ 或 $(CH_3)_2CH_2CH_2$ 。



[0195] 本领域技术人员了解的是各种二醛和丙二酸衍生化合物的克诺维纳盖尔缩合都可根据本发明的方法用于合成饱和和不饱和(其中烯在2位和n-2位处)二羧酸。举例来说且不加限制地,丙二醛可用于合成庚烷二甲酸,丁烷-1-4-二醛用于合成辛烷二甲酸,且戊烷-1,5-二醛可用于合成壬烷二甲酸。

[0196] 在本发明的其它实施方案中,丙二酸或丙二酸衍生化合物用作庚二酸的化学合成中的基质。除上述用以获得庚二酸的二卤代烷和二醛途径之外,本发明也提供用于通过可或可不衍生自由由本发明提供的丙二酸的丙-2-烯醛(丙烯醛)和丙二酸的反应来合成庚二酸的方法。丙二酸向丙-2-烯醛的迈克尔加成(Michael addition)产生2-(3-氧代丙基)丙二酸,且随后与具有另一丙二酸分子的醛部分缩合产生己-1-烯-1,1,6,6-四甲酸。热诱导的脱羧继之以氢化产生庚二酸。

[0197] 在本发明的另一方法中,丙二酸用作内酰胺或内酯的化学合成中的基质。在一个实施方案中,通过以下步骤使丙二酸二乙酯合成转化成 δ -戊内酰胺:丙二酸二乙酯是通过丙二酸与乙醇的费歇尔酯化产生。丙二酸二乙酯向丙烯腈的迈克尔加成产生2-(2-氰基乙基)丙二酸二乙酯。使2-(2-氰基乙基)丙二酸二乙酯的腈部分氢化会产生在温和条件下内酰胺化以产生2-氧代哌啶-3-甲酸乙酯的氨基酸。使2-氧代哌啶-3-甲酸乙酯的酯部分皂化,继之以加热会导致脱羧以得到 δ -戊内酰胺。这个非限制性实例说明于此:



[0199] 在本发明的另一方法中,丙二酸单乙酯用作乙酸乙酯的化学合成中的基质。加热丙二酸单乙酯导致热诱导的脱羧以形成乙酸乙酯。在一些实施方案中,丙二酸单乙酯用于合成乙酸乙酯。本领域技术人员将认识到存在许多已建立的用于自由本发明提供的丙二酸制备丙二酸单烷基酯的方法。

[0200] 在本发明的另一方法中,丙二酸单烷基酯(例如丙二酸单甲酯、丙二酸单乙酯、丙

二酸单丙酯、丙二酸单丁酯、丙二酸单异丁酯)用作3-羟基丙酸或它的相关烷基酯的化学合成中的基质。硼烷将化学选择性降低丙二酸单烷基酯的游离羧酸部分且产生相应3-羟基丙酸烷基酯。3-羟基丙酸烷基酯本身是有价值产品。酯也可被水解以产生3-羟基丙酸。3-羟基丙酸可又脱水生成本文更详细描述的一种有价值大宗化学品的丙烯酸。在一些实施方案中,丙二酸单烷基酯用于合成3-羟基丙酸。在一些实施方案中,丙二酸单烷基酯是丙二酸单甲酯。在一些实施方案中,丙二酸单烷基酯是丙二酸单乙酯。

[0201] 在本发明的另一方法中,丙二酸用于合成二烷基化1,3-二羰基化合物。在各种实施方案中,丙二酸用作选自以下组成的组的化合物的化学合成中的基质:2,2-二甲基丙二酸二乙酯、2,2-二乙基丙二酸二乙酯、2,2-二丙基丙二酸二乙酯、2,2-二丁基丙二酸二乙酯、2,2-二甲基丙二酸二甲酯、2,2-二乙基丙二酸二甲酯、2,2-二丙基丙二酸二甲酯和2,2-二丁基丙二酸二甲酯。可自使用烷基卤化物合成获得的丙二酸二烷基酯根据本发明形成合成的二烷基化1,3-二羰基化合物(参见Gatterman等,“The practical methods of organic chemistry”第3版,The Macmillan Company:New York,第189-191页(1916))。

[0202] 在本发明的其它实施方案中,使由本发明提供的丙二酸或丙二酸衍生物与脲反应以形成当前在售的约50种相关医药(包括苯巴比妥(phenobarbital))的核心结构巴比妥酸(barbituric acid)。

[0203] 在本发明的合成方法的一些实施方案中,高度纯化的丙二酸制剂用于合成反应和/或相关下游应用;在其它实施方案中,使用纯度较小的丙二酸。在其它情况下,粗溶解产物或其它相对未纯化丙二酸可用于合成反应和/或下游应用。对于大多数合成化学反应,通常需要丙二酸纯度大于90%w/w。在丙二酸用于聚合反应的情况下,可需要甚至更大纯度(即95%w/w或以上)。然而,在其它应用中,可采用纯度小得多的丙二酸制剂。举例来说,丙二酸用作动物饲料,且出于这个目的,可采用产生性细胞本身或它们的粗溶解产物。

[0204] 在本发明的另一实施方案中,使丙二酸(通过任何手段产生)与戊二酸半醛(其也可自丙二酸产生)反应以产生庚-2-烯二酸。这是克诺维纳盖尔缩合的另一实例且可用广泛系列的与这个反应相关的溶剂和催化试剂进行。接着使用催化氢化还原庚-2-烯二酸以产生庚二酸(heptanedioic acid/pimelic acid)。庚二酸具有广泛多种用途,包括作为尼龙5,7中的单体;作为用于增强生物素生物合成的发酵培养基补充剂;以及作为塑化剂。用于庚二酸合成的先前技术方法是昂贵且低产的。本发明提供一种用于可衍生自由由本发明提供的丙二酸的这个重要化合物的新化学合成法。

[0205] 在本发明的其它实施方案中,蒸汽或液相氢化用于使丙二酸二烷基酯转化成1,3-丙二醇。由于丙二酸和丙二酸酯的2号碳的反应性性质,各种化合物的多取代衍生物可使用由本发明提供的丙二酸产生。在一些实施方案中,新戊二醇(2,2-二甲基丙烷-1,3-二醇)作为一些聚酯的重要组分是通过首先产生丙二酸2,2-二甲基-二烷基酯来产生。这个化合物可使用多种碱作为催化剂以及甲基卤化物作为甲基的来源来产生。随后酯基团氢化导致自酯基团产生新戊二醇和相关醇。这代表一种用以获得这个重要工业化合物的新颖途径。此外,由于丙二酸和丙二酸衍生物化合物的2号碳的高度反应性性质,其它烷基卤化物可用于产生当并入聚合物中时具有新颖性质潜力的新型化合物。

[0206] 已详细描述的本发明是由以下实施例所说明,鉴于本发明的不同方面、实施方案和应用,所述实施例不应解释为限制本发明。

实施例

[0207] 实施例1:构建用于在大肠杆菌中产生丙二酸的编码野生型EHD3和EHD3突变体E124S、E124A、E124A/E308V、和E124V丙二酰基-辅酶A水解酶的重组核酸和表达载体

[0208] 本发明提供用于在大肠杆菌宿主中产生丙二酸的方法、以及产生丙二酸和表达EHD3突变酶的大肠杆菌宿主细胞,所述突变酶包括但不限于E124S、E124A和E124V。这个实施例描述构建用于本发明中的EHD3和突变EHD3蛋白的蛋白质编码序列、含有那些编码序列的表达载体、和包含那些表达载体的宿主细胞。使用引物对A93/A94,通过PCR自面包酵母(Baker's yeast)扩增编码野生型酿酒酵母EHD3的核酸。使用引物对A93/A96和A95/A94引入点突变E124A,使用引物对A93/A98和A97/A94引入突变E124V,且使用引物对A93/A100和A99/A94引入突变E124S。使用标准技术将所得核酸克隆至含有pSC101复制起点、氯霉素抗性盒和P_{Lac01}启动子的大肠杆菌表达载体中。将所得载体转化至大肠杆菌DH10b宿主中且涂铺在含有50μg/ml氯霉素(Cm⁵⁰)和2%w/v葡萄糖的鲁里亚-贝塔尼(Luria-Bertani, LB)琼脂板上。接着将个别菌落接种至48孔板中的3ml LB培养基中;在6小时生长之后,分离质粒且将EHD3蛋白编码区域测序。当对源于EHD3的克隆(E124A)的质粒测序时,发现第二点突变E308V使EHD3(E124A)克隆稳定;不带电荷缬氨酸残基的存在可能使蛋白质折叠稳定。EHD3(E124A、E308V)菌株也用于产生丙二酸(参见实施例2)。引物A93-A100分别是SEQ ID NO: 13-20。

[0209] 实施例2:使用EHD3突变体E124S和E124A/E308V在大肠杆菌中体内产生丙二酸

[0210] 这个实施例描述导致使用异源EHD3丙二酰基-辅酶A水解酶在大肠杆菌宿主细胞中体内产生丙二酸的宿主细胞和培养条件。大肠杆菌菌株K12用含有野生型EHD3、EHD3(E124A)、EHD2(E124V)、EHD3(E124S)、EHD3(E124A、E308V)的载体或空载体阴性对照转化。在LB琼脂板(Cm⁵⁰、2%葡萄糖)上将转化体划线。在37℃下过夜生长之后,将个别菌落接种至48孔板中的3ml LB(Cm⁵⁰、2%葡萄糖)中。在37℃下在板振荡器上孵育培养物6小时,此时将各培养物以1%v/v接种至48孔板中的补充有Cm⁵⁰和混合碳源(0.5%甘油、0.05%葡萄糖、0.2%乳糖)的M9基本培养基中。在30℃下在板振荡器上孵育培养物,且在孵育48小时之后移除500μL发酵液样本以进行分析。

[0211] 离心(x6000g, 1分钟)样本且分析上清液以定量丙二酸。于20μM水中制备化学标准物。在连接于示差折光检测器和UV检测器(210nm监测)的Shimadzu Prominence XR UPLC上对丙二酸进行分离。在Bio-Rad Aminex HPX-87h发酵监测柱上进行产物分离。UPLC被编程来等度运行,使用5mM H₂SO₄作为洗脱剂,流速是每分钟600μL。每个样本注射10μL,且样本板温度保持在4℃下。丙二酸标准物在约19.8分钟开始洗脱。向含有丙二酸的样本中添加标准物显示丙二酸峰面积按比例增加,从而确认产生丙二酸。通过与由权威性丙二酸制备的标准曲线进行比较来计算丙二酸浓度(mg/l)。

[0212] 对于具有空载体对照的培养物,在与丙二酸标准物相同的滞留时间未观察到峰,且积分峰面积低于仪器的检测限(即未观察到丙二酸产生)。观察到具有野生型EHD3、EHD3(E124A)、EHD3(E124V)、EHD3(E124S)和EHD3(E124A、E308V)的样本中产生丙二酸;丙二酸浓度是(平均值±标准偏差;n=3):野生型EHD3, 6.0±0.2mg/l;EHD3(E124A), 7.6±0.7mg/l;EHD3(E124V), 0.28±0.03mg/l;EHD3(E124S), 82.3±7.8mg/l;以及EHD3(E124A/E308V), 8.35±2.5mg/l。相对于野生型蛋白质,EHD3(E124V)导致丙二酸的产生降低,这可能是由于

丙二酰基-辅酶A底物结合性不良与对其它内源性酰基-辅酶A分子具有高混杂活性两者所致。类似地,相对于野生型EHD3,EHD3 (E124A) 仅使丙二酸产生微小增加。EHD3突变体E124S导致丙二酸产生超过野生型EHD3、EHD3 (E124A) 和EHD (E124V) 显著 ($p < 0.05$, t检验) 增加,从而说明E124S在增加丙二酰基-辅酶A底物结合性和丙二酸产生方面具有重要性。

[0213] 实施例3:构建用于在大肠杆菌中体内产生丙二酸的其它EHD3E124突变体和表达载体

[0214] 实施例1描述构建野生型EHD3和一子组E124突变(具体来说E124A、E124V、E124S和E124A/E308V)的大肠杆菌表达载体,且实施例2描述它们用于在大肠杆菌细胞中产生丙二酸。这个实施例描述构建所有EHD3 E124点突变,即E124G、E124T、E124C、E124L、E124I、E124M、E124P、E124Y、E124W、E124D、E124N、E124Q、E124H、E124K、E124R和E124F的大肠杆菌表达载体。使用如实施例1中所述的具有pSC101起点和 P_{Lac01} 复制起点的大肠杆菌载体构建这些EHD3突变体。正向PCR引物包含核酸序列(5'-aatttttttactgatNNNtattcctttgaattttcaaatagc-3'),其中序列“NNN”是编码所需E124氨基酸点突变的三个核苷酸;同样,反向PCR引物包含核酸序列(5'-ttcaaagaataaNNNatcagtaaaaaatttgatggacttg-3'),其中序列“NNN”互补于编码所需E124氨基酸点突变的三个核苷酸。正向PCR引物连同PCR引物A94 (SEQ ID NO:14) 一起使用,且反向PCR引物连同PCR引物A93 (SEQ ID NO:13) 一起使用以产生含有所需点突变的两个重叠EHD3基因片段。用引物A93和A94扩增两个EHD3基因片段产生含有所需点突变的全长EHD3基因。使用标准克隆方案构建表达载体且转化至大肠杆菌DH10b中并随后如实施例1中所述加以分离且通过测序来验证。

[0215] 实施例4:使用EHD3 E124突变体在重组大肠杆菌中体内产生丙二酸

[0216] 这个实施例描述一组大肠杆菌宿主细胞的发酵以及产生的丙二酸水平的测定,所述组的各成员含有19种可能EHD3 E124氨基酸取代之一。相对于野生型,特定E124点突变应理论上改进丙二酰基-辅酶A结合性和丙二酸产生;具体来说,点突变E124T、E124N、E124Q、E124H、E124K和E124R应改进丙二酸产生,理论上是由于引入含有改进与丙二酸的末端羧酸部分的相互作用的官能团的氨基酸。点突变E124S、E124Q和E124K具有位于处于应在理论上最佳配合丙二酸结合的位置中的EHD3结合袋中的侧链。大致上如实施例2中所述进行丙二酸产生且定量。

[0217] 实施例5:构建用于在大肠杆菌中体内产生丙二酸的YciA丙二酰基-辅酶A水解酶表达载体

[0218] 这个实施例描述根据本发明使用大肠杆菌YciA在大肠杆菌中产生丙二酸。使用引物A120 (SEQ ID NO:21) 和A121 (SEQ ID NO:22) 自宿主基因组PCR扩增野生型大肠杆菌酰基-辅酶A YciA。将所得核酸克隆在含有pSC101复制起点和氨苄青霉素抗性基因的大肠杆菌表达质粒上的 P_{Lac01} 启动子之后。对照载体包括未在 P_{Lac01} 启动子之后插入yciA基因的空载体。如实施例1中所述,培养个别菌落,分离它们的质粒,且将YciA基因插入物的编码区域测序。

[0219] 实施例6:使用异源YciA在重组大肠杆菌中体内产生丙二酸

[0220] 这个实施例描述导致使用如实施例5中所述的编码异源YciA丙二酰基-辅酶A水解酶的表达载体在大肠杆菌宿主细胞中体内产生丙二酸的宿主细胞和培养条件。用含有大肠杆菌YciA的载体转化野生型大肠杆菌菌株K12;具有空载体的野生型大肠杆菌充当阴性对

照。如实施例2中所述进行丙二酸产生且定量。

[0221] 实施例7:构建用于在酵母中体内产生丙二酸的工程化EHD3丙二酰基-辅酶A水解酶表达载体

[0222] 本发明也提供用于在酵母细胞中体内产生丙二酸的表达载体、宿主细胞和方法。一般来说,酵母细胞比大肠杆菌可耐受发酵液中更高浓度的有机酸且具有更好建立的工业发酵方案。

[0223] 部分地由使用实施例1和3中所述的大肠杆菌表达载体作为EHD3基因的PCR模板来产生这个实施例中所述的酵母表达载体。酵母表达载体含有2微米复制起点、ura3营养缺陷标记和TEF启动子;载体也含有puc复制起点和氨苄青霉素或氯霉素抗性盒以达成在大肠杆菌中扩增载体。将质粒转化至作为S288C亲本菌株的两种衍生物的酿酒酵母BY4741或BY4742背景中。

[0224] 可将其它突变引入EHD3编码序列中以消除线粒体靶向性。EHD3编码序列中的碱性氨基酸R3、K7、K14、K18和R22对线粒体靶向性是重要的,且使它们中的任何一者或多者突变成A或V都会降低线粒体EHD3表达且增加胞质液EHD3表达。通过用含有所需点突变的正向和反向引物对EHD3模板进行PCR扩增来引入突变。使用匹配于EHD3基因,例外之处是所需氨基酸点突变的位置处的核苷酸序列被改变成“gyt”(其中y是寡核苷酸PCR引物的混合群体中的半胱氨酸或胸腺嘧啶核苷酸)的引物引入突变。首先将基因片段克隆至具有pSC101复制起点、 p_{Lac01} 启动子和氯霉素抗性标记的大肠杆菌表达载体中,且如实施例1中所述确认载体序列;在分离所需突变体之后,扩增EHD3基因且克隆至酵母表达载体中。

[0225] 实施例8:使用工程化EHD3丙二酰基-辅酶A水解酶在重组酵母中体内产生丙二酸

[0226] 使用标准方案,用大致上如实施例7中所述制备的具有异源EHD3的酵母表达载体或空载体阴性对照质粒转化酿酒酵母BY4742宿主细胞。在缺乏尿嘧啶的合成完全省却培养基(SD)琼脂板上将转化体划线且在30℃下培养;使个别菌落在30℃下于3mL SD培养基中生长过夜且随后以1%v/v稀释至3mL缺乏尿嘧啶的SD中。在30℃下培养菌株72小时;在24、48和72小时时间点取样500 μ l等分试样以定量丙二酸产生和OD₆₀₀。TEF启动子是一种组成型启动子。

[0227] 实施例9:通过表达异源乙酰基-辅酶A合成酶增加工程化酵母中的丙二酸生物合成

[0228] 除如实施例1-8中所说明的用于表达异源丙二酰基-辅酶A水解酶和体内产生丙二酸的方法、载体和宿主细胞之外,本发明也提供用于改进丙二酸的效价、产率和/或生产力的方法和宿主细胞。在一个方面,通过增加乙酰基-辅酶A的生物合成来改进丙二酸产生。这个实施例描述在包含丙二酰基-辅酶A水解酶路径的重组酿酒酵母宿主中异源表达乙酰基-辅酶A合成酶以及在丙二酸产生方面的所得改进。说明的5种乙酰基-辅酶A合成酶蛋白是酿酒酵母ACS1和ACS2、大肠杆菌AcsA、肠道沙门氏菌Acs以及枯草芽孢杆菌AcsA。所有基因都自它们的相应宿主加以PCR扩增且克隆至具有2微米复制起点、ura3营养缺陷标记和TEF启动子的酵母表达载体中;载体也含有puc复制起点和氨苄青霉素抗性盒以达成在大肠杆菌中扩增载体。

[0229] 实施例10:通过表达异源丙酮酸脱氢酶增加工程化酵母中的丙二酸生物合成

[0230] 尽管实施例9描述通过表达异源乙酰基-辅酶A合成酶增加乙酰基-辅酶A生物合

成,但这个实施例描述通过表达异源丙酮酸脱氢酶增加乙酰基-辅酶A生物合成。具体来说,在包含丙二酰基-辅酶A水解酶路径的重组酿酒酵母中异源表达酿酒酵母丙酮酸脱氢酶PDA1、PDB1、LAT1、LPD1和PDX1。基因都自酿酒酵母染色体加以PCR扩增且克隆至具有2-微米复制起点、ura3营养缺陷标记和TEF启动子的酵母表达载体中;载体也含有puc复制起点和氨苄青霉素抗性盒以达成在大肠杆菌中扩增载体。

[0231] 实施例11:通过异源表达乙醇分解代谢路径增加大肠杆菌和酿酒酵母中的丙二酸生物合成

[0232] 这个实施例描述用以增加乙酰基-辅酶A生物合成的第三途径:异源表达乙醇分解代谢路径。乙醇分解代谢路径包含两种或三种酶。醇脱氢酶和乙醛脱氢酶(酰化)、或醇脱氢酶、乙醛脱氢酶(非酰化)、以及乙酰基-辅酶A合成酶。将醇脱氢酶酿酒酵母ADH2、大肠杆菌AdhP、以及智人ADH1A、智人ADH1B和智人ADH1C与乙醛脱氢酶(酰化)或醛脱氢酶和乙酰基-辅酶A合成酶组合克隆。乙醛脱氢酶(酰化)大肠杆菌MhpF、大肠杆菌AdhE、假单胞菌属CF600DmpF和恶臭假单胞菌TodL也都组合克隆。在利用乙醛脱氢酶(非酰化)的乙醇分解代谢路径中,使用酿酒酵母ALD2、ALD3、ALD4、ALD5和ALD6。所用乙酰基-辅酶A合成酶是酿酒酵母ACS1和ACS2以及大肠杆菌Acs。所有基因都自基因组DNA加以PCR扩增。

[0233] 对于大肠杆菌宿主细胞,自具有p15a复制起点、氨苄青霉素抗性标记和 P_{lacO1} 启动子的载体骨架表达乙醇分解代谢路径。所有两基因和三基因路径的所有组合都构建成单一操纵子。用EHD3表达质粒和乙醇分解代谢路径质粒共转化大肠杆菌K12且在LB琼脂板(Cm^{50} 、 Cb^{50} 、2%w/v葡萄糖)上划线。对照菌株具有空载体。如实施例2中所述进行产生培养以及对其的分析,其中显著例外之处是添加为维持第二质粒所需的第二抗生素。

[0234] 可使用与大肠杆菌实验相同的乙醇浓度进行酿酒酵母发酵。酵母表达载体具有2微米起点、ura3营养缺陷标记和CUP启动子,且在这个载体骨架上构建乙醇分解代谢路径的所有组合;不存在乙醇分解代谢路径的质粒充当阴性对照。将乙醇分解代谢路径质粒转化至在酵母染色体上包含工程化EHD3丙二酰基-辅酶A水解酶路径的重组酿酒酵母BY4742中。所有发酵都在30℃下于无尿嘧啶的SD培养基中进行。在生长12或24小时之后,用100μM硫酸铜诱导乙醇分解代谢路径表达。在24、48和72小时取样500μl等分试样以定量乙醇、乙醛、乙酸和丙二酸浓度;也记录各时间点的细胞密度 OD_{600} 测量结果。除丙二酸效价之外,也计算乙醇消耗量。

[0235] 实施例12:通过异源表达ATP柠檬酸裂解酶增加酿酒酵母中的丙二酸生物合成

[0236] 这个实施例描述用以在重组酿酒酵母中增加乙酰基-辅酶A生物合成以及改进丙二酸产生的第四方法。ATP柠檬酸裂解酶(EC 2.3.3.8)催化在胞质液中自柠檬酸形成乙酰基-辅酶A、草酰乙酸和ADP。自基因组DNA PCR扩增来自油脂酵母弯假丝酵母、浅白隐球酵母、产油油脂酵母、圆红冬孢酵母、粘红酵母、皮状丝孢酵母、解脂耶罗威亚酵母的ATP柠檬酸裂解酶且克隆至酵母表达载体中的CUP启动子之后;表达载体含有2微米起点和1eu2d营养缺陷标记。不存在ATP柠檬酸裂解酶的质粒充当阴性对照。

[0237] 将ATP柠檬酸裂解酶路径质粒转化至在酵母染色体上包含工程化EHD3丙二酰基-辅酶A水解酶路径的重组酿酒酵母BY4742中。所有实验都在30℃下于无尿嘧啶的SD培养基中进行。在生长12或24小时之后,用100μM硫酸铜诱导ATP柠檬酸裂解酶路径表达。一些培养物也补充以0.5、1、2.5或5g/l柠檬酸盐以提供对路径活性的额外证明。在24、48和72小时取

样500 μ l等分试样以定量柠檬酸(当适用时)和丙二酸浓度;也记录各时间点的细胞密度OD₆₀₀测量结果。除丙二酸效价之外,可计算柠檬酸盐消耗量。

[0238] 实施例13:通过改进宿主细胞脂肪酸储存增加重组酵母中的丙二酸生物合成

[0239] 这个实施例描述用以在重组酿酒酵母中增加乙酰基-辅酶A生物合成以及改进丙二酸产生的第五方法。脂肪酸生物合成路径与丙二酸产生竞争乙酰基-辅酶A和丙二酰基-辅酶A,且改变宿主细胞脂肪酸合成代谢可增加丙二酸产生。本发明提供包含对一种或多种编码影响脂肪酸储存和分解代谢的蛋白质的核酸的遗传修饰的宿主细胞。在酿酒酵母中,蛋白质SNF2、IRA2、PRE9、PH090、SPT21、POX1、ANT1、FOX3、EHD3、PAS1、PAS3、ARE1、ARE2、DGA1、LR01、ACL1、MAE1、GLC3、GLG1、GLG2、PAT1和PEX11被个别地以及组合地敲除且培养所得菌株以产生丙二酸。构建的所有酿酒酵母菌株都包含用于产生丙二酸的工程化EHD3丙二酰基-辅酶A水解酶路径。如实施例8中所述进行发酵,且如实施例2中所述定量丙二酸。

[0240] 实施例14:通过增加 β -氧化酶活性增加重组酵母中的丙二酸生物合成

[0241] 这个实施例描述用以在重组酿酒酵母中增加乙酰基-辅酶A生物合成以及改进丙二酸产生的第六方法。除降低宿主细胞脂肪酸合成代谢之外,增加宿主细胞脂肪酸分解代谢也可增加丙二酸产生。本发明提供被修饰来增加PAT1和/或PEX11的表达的宿主细胞。自基因组酿酒酵母DNA PCR扩增PAT1和PEX11且克隆至酵母表达载体中的CUP启动子之后;表达载体含有2微米起点和1 μ 2d营养缺陷标记。无 β -过氧化物酶的质粒充当阴性对照。

[0242] 将 β -氧化酶路径质粒转化至在酵母染色体上包含工程化EHD3丙二酰基-辅酶A水解酶路径的重组酿酒酵母BY4742中。使用标准重组方法使工程化丙二酰基-辅酶A水解酶整合于染色体上。所得菌株充当用以测试随后修饰和它们对丙二酸产生的影响的基础。所有实验都在30 $^{\circ}$ C下于无尿嘧啶的SD培养基中进行。在生长12或24小时之后,用100 μ M硫酸铜诱导路径表达。一些培养物也补充以0.5、1、2.5或5g/l棕榈酸以提供对路径活性的额外证明。在24、48和72小时取样500 μ l等分试样以定量棕榈酸(当适用时)和丙二酸浓度;也记录各时间点的细胞密度OD₆₀₀测量结果。除丙二酸效价之外,可计算棕榈酸消耗量。

[0243] 实施例15:通过增加乙酰基-辅酶A羧化酶活性改进工程化酵母中的丙二酸生物合成

[0244] 除用于表达异源丙二酰基-辅酶A水解酶和体内产生丙二酸的方法、载体和宿主细胞之外,本发明也提供用于改进丙二酸的效价、产率和/或生产力的方法和宿主细胞。在一个方面,通过增加丙二酰基-辅酶A的生物合成来改进丙二酸产生。丙二酰基-辅酶A是自乙酰基-辅酶A生物合成丙二酸中的倒数第二个中间体,且在酿酒酵母中,这个反应是由乙酰基-辅酶A羧化酶(ACC1)催化。

[0245] 通过过表达酿酒酵母ACC1来增加丙二酰基-辅酶A生物合成。为此,使用标准方法将ACC1基因克隆在含有2微米复制起点和ura3营养缺陷标记的酿酒酵母表达质粒上的CUP启动子之后。对照载体包括空载体。以染色体缺失负责ACC1磷酸调控的ACC1和SNF1蛋白质激酶来使酿酒酵母宿主菌株工程化;以独立与组合两种方式构建染色体缺失。如实施例8中所述使具有表达质粒或对照质粒的宿主细胞生长且如实施例2中所述定量丙二酸产生。

[0246] 实施例16:通过用浅蓝菌素补充发酵液改进宿主细胞中的丙二酸生物合成

[0247] 在这个实施例中,通过用浅蓝菌素补充发酵液来改进表达EHD3源性丙二酰基-辅酶A水解酶的重组宿主细胞中的丙二酸产生。将包含在P_{Lac01}启动子控制之下的酿酒酵母

EHD3 (E124S) 的丙二酸产生质粒A4转化至大肠杆菌K12宿主中。

[0248] 将个别菌落接种至48孔板中的3ml LB培养基 (Cm^{50} 、2%w/v葡萄糖) 中且在37℃下在板振荡器上培养6小时。接着将菌株以1%v/v续代培养至3mL M9基本培养基 (Cm^{50} 、0.5%w/v甘油、0.05%w/v葡萄糖、0.2%w/v乳糖) 中且在30℃下在板振荡器上培养。在生长6小时之后,一半培养物补充以10mg/1浅蓝菌素。在生长48小时之后,如实施例2中所述测量上清液中的丙二酸浓度。

[0249] 实施例17:通过用二氧化碳补充发酵液改进工程化酵母中的丙二酸生物合成

[0250] 在这个实施例中,修改发酵条件以增加丙二酰基-辅酶A的生物合成。由酶乙酰基-辅酶A羧化酶将乙酰基-辅酶A酶促转化成丙二酰基-辅酶A需要化学计量的二氧化碳,且用二氧化碳补充生长培养基会增加丙二酸产生。以固体碳酸钙或气态二氧化碳形式添加二氧化碳至生长培养基中。

[0251] 使具有丙二酰基-辅酶A生物合成路径的重组酵母细胞在补充有0.1-10g/l之间的碳酸钙的确定基本培养基中生长。对照培养物不补充以碳酸钙。历经48小时生长过程,如实施例2中所述定量丙二酸产生。

[0252] 实施例18:通过降低丙二酸分解代谢改进工程化酵母中的丙二酸生物合成

[0253] 在这个实施例中,通过消除宿主细胞中的内源性丙二酸分解代谢来增加丙二酸产生。酿酒酵母含有多种酰基-辅酶A合成酶,包括FAA1、FAA2、FAA3、FAA4、LSC1和LSC2;通过缺失或修饰宿主基因组上编码这些蛋白质的核酸,可降低生长培养基中的丙二酸的分解代谢。

[0254] 构建具有各酵母酰基-辅酶A合成酶的丙二酰基-辅酶A基因敲除菌株。接着在补充有1-5g/l丙二酸钠的确定培养基中培养所得菌株。历经48小时过程监测发酵液中的丙二酸浓度且如实施例2中所述加以定量。可遵循类似程序构建具有多个基因敲除的菌株。

[0255] 实施例19:通过改进丙二酸自宿主细胞的分泌改进工程化酵母中的丙二酸生物合成

[0256] 在这个实施例中,通过增加丙二酸自宿主细胞的分泌来改进丙二酸产生。这是通过过表达各酿酒酵母多效性抗性泵(即PDR5、PDR10、PDR11、PDR12、PDR15和PDR18)的一者或多者来达成。

[0257] 实施例20:通过降低由丙二酸进行的丁二酸脱氢酶竞争性抑制改进工程化酵母中的丙二酸生物合成

[0258] 在这个实施例中,降低对酿酒酵母丁二酸脱氢酶SDH1的竞争性抑制,从而使得能够获得较高丙二酸效价。首先,使用标准方法遗传修饰酿酒酵母以缺失丁二酸脱氢酶(SDH1)的天然染色体拷贝。使所得菌株厌氧生长以促进在不存在SDH1蛋白下的生长。随后用具有遗传修饰的SDH表达盒的载体转化SDH1缺失菌株,所述表达盒含有E300D、R331K、R331H、R442K、R442H突变或这些突变的组合。将这些突变SDH基因克隆在具有2微米起点和ura3营养缺陷标记的酵母骨架上的组成型TEF启动子之后。随后将载体转化至在染色体上编码丙二酰基-辅酶A水解酶的酿酒酵母菌株中且使所得菌株在缺乏尿嘧啶的SD培养基中生长。如实施例8中所述培养转化的菌株以产生丙二酸且如实施例2中所述定量丙二酸以定量这些SDH突变对丙二酸产生的影响。

[0259] 实施例21:使用外源添加的丙二酸的MdcY丙二酸转录因子生物感测器。

[0260] 质粒S14用于说明在大肠杆菌宿主细胞中生物传感器对外源添加的丙二酸的反应。S14采用源于乙酸钙不动杆菌的丙二酸反应性转录因子MdcY (SEQ ID NO:3) 和MdcY反应性启动子 P_{MdcL} (SEQ ID NO:6)。使用具有氨苄青霉素抗性标记和ColE1复制起点的大肠杆菌载体骨架构建本发明的生物传感器; tetA四环素抗性基因被放置在 P_{MdcL} 启动子的控制之下。将质粒S14转化至大肠杆菌宿主中产生在用丙二酸补充发酵液之后表达tetA基因产物的菌株, 且所述菌株在四环素抗性方面展现丙二酸依赖性增加。

[0261] 合成产生编码MdcY和TetA基因产物、 P_{MdcL} 启动子以及大肠杆菌载体骨架的核酸; 接着通过PCR扩增核酸以及随后克隆至大肠杆菌载体骨架中来构建生物传感器载体。将质粒转化至化学感受态大肠杆菌DH10b中且将所得克隆涂铺在含有50 μ g/ml羧苄青霉素(Cb⁵⁰)的LB琼脂板上。使个别菌落在补充有抗生素的3ml LB培养基中生长过夜且验证纯化质粒的序列。

[0262] 用质粒S14共转化大肠杆菌菌株K12且自LB琼脂板(Cb⁵⁰)分离个别菌落。使菌落在25ml LB发酵液(Cb⁵⁰)中生长直至达到在600nm下的光密度(OD_{600})约0.50, 在这个时间点, 制备细胞储备物且储存在-80 $^{\circ}$ C下; 细胞储备物是0.5ml细胞培养物和0.5ml 50% v/v甘油溶液。

[0263] 所有生物传感器说明都用丙二酸进行。解冻生物传感器细胞储备物的等分试样且用于接种250ml有挡板爱伦美烧瓶(Erlenmeyer flask)中的50ml LB培养基(Cb⁵⁰)。在37 $^{\circ}$ C下孵育培养物2小时; 随后, 添加0.6ml生物传感器培养物至以补充有所需浓度的四环素和丙二酸的2.3ml LB培养基(Cb⁵⁰)制备的48孔板中($n=4$)。接着使各板在30 $^{\circ}$ C下在轨道滴定板振荡器上生长。在孵育12小时之后, 获取200 μ l样本以测量 OD_{600} 。

[0264] 具有S14的生物传感器培养物显示丙二酸剂量依赖性反应(图1)。动态范围(充分诱导的样本与不存在丙二酸补充的那些样本之间的最大 OD_{600} 值差异)是1.20 OD_{600} 单位, 从而指示基于MdcY- P_{MdcL} 的生物传感器对外源添加的丙二酸具有高度反应性。在0.5-1mM外源添加的丙二酸之间观察到 OD_{600} 增加, 从而提供可使用这个方法定量丙二酸所历经的适合范围。

[0265] 实施例22. 用以检测发酵液中生物产生的丙二酸的MdcY丙二酸转录因子生物传感器

[0266] 在这个实施例中, 丙二酸转录因子生物传感器用于检测丙二酸自如本发明的其它方面所述加以工程化的酵母菌株的产生。

[0267] 使用遗传工程化酵母菌株如下产生丙二酸。具有用于表达丙二酰基-辅酶A水解酶的包含CYC1终止子、氨苄青霉素抗性盒、PMB1复制起点、CEN/ARS复制起点和URA3选择标记的载体的酿酒酵母BY4741酵母细胞用于发酵。F0PNG8-1丙二酰基-辅酶A水解酶(来自苏云金杆菌幕虫亚种菌株YBT-020; UniProt ID F0PNG8, 具有E91S突变)和F6AA82-2丙二酰基-辅酶A水解酶(来自黄褐假单胞菌菌株12-X; UniProt ID F6AA82, 具有E95S和Q348A突变)各自在TEF1启动子的控制之下自这个质粒表达。实施例30中所述的培养基与作为碳源的20g/L葡萄糖一起使用。如下进行产生。一式四份用产生菌株的20 μ l起子培养物接种48孔板中的2ml培养基。用可透气膜将板覆盖, 在30 $^{\circ}$ C下在板振荡器上孵育, 且在生长142小时之后取样以对累积的产物进行HPLC和生物传感器分析。通过离心和经0.45微米膜过滤来自培养基移除细胞和细胞碎片, 随后由HPLC或生物传感器进行分析。

[0268] 具有在丙二酸反应性 P_{mdcL} 启动子的表达控制下编码四环素抗性基因(tetA)的质粒pS14、或在丙二酸反应性 P_{mdcL} 启动子的表达控制下编码lacZ基因的质粒pS27的大肠杆菌细胞用作生物传感器指示菌株。以与实施例21中对于pS14所述相同的方式构建载体pS27,其中插入编码 β -半乳糖苷酶的lacZ基因而非tetA基因。如实施例21中所述制备生物传感器菌株。

[0269] 添加自96孔产生板获得的酵母消耗培养基至含有120 μ l生物传感器细胞培养物的96孔板中。对于TetA (pS14) 生物传感器,添加10 μ l四环素储备溶液(以提供20-35 μ g/ml的范围)至各孔中。对于TetA (pS14) 生物传感器与lacZ (pS27) 生物传感器两者,各孔的剩余体积用LB培养基(Cb^{50})填充至最终体积600 μ l且如先前所述加以生长。在2小时之后(对于S27培养物)以及在5-8小时之后(对于S14培养物),将样本(200 μ l)收集至96孔板中,且测量OD600。如下对来自S27(β -半乳糖苷酶报道体)生物传感器板的样本进行邻硝基苯基- β -半乳糖苷(ONPG)测定。于25 μ l溶解缓冲液中1:4稀释细胞且随后添加90 μ l ONPG储备溶液(10mg/ml于去离子水中)至各孔中。彻底混合各孔的内含物且在30 $^{\circ}\text{C}$ 下保持4-16小时。在420nm下测量光密度。

[0270] 观察到对丙二酸的剂量依赖性反应。也如实施例2中所述,通过HPLC测量特定丙二酸浓度,且通过与标准曲线进行比较来定量。计算生物传感器的可定量输出结果OD420(对于pS27)或OD600(对于pS14)与通过HPLC测量的特定丙二酸浓度之间的线性回归分析结果以得到决定系数(R^2) 0.88492(绘制37个OD420样本)和0.89755(绘制18个OD600样本)。

[0271] 本领域技术人员将认识到这些极高决定系数指示生物传感器输出结果与培养基中的丙二酸浓度之间具有关联。本发明的这个方面提供关于筛选生物丙二酸产生中的差异性输出结果的成本与时间两方面的巨大优势。稀释用于激发生物传感器的培养基可有助于传感器的动态反应范围自零延伸至丙二酸完全溶液饱和。相较于HPLC分析每个样本需要2-20分钟或更多的时间,使用基于板的筛选使得能够在数分钟内筛选96个样本。通过限制或完全替换HPLC,在资本投资和溶剂用量以及造成的排放方面的节约也是大量的。

[0272] 实施例23.通过添加氢氧化钙、碳酸钙和氯化钙使丙二酸自发酵液沉淀

[0273] 在这个实施例中,通过用二价阳离子(具体来说钙)进行沉淀来自发酵液纯化丙二酸。使用外源添加至发酵液中的合成代谢物说明纯化方法。使酿酒酵母BY4741的酵母培养物在30 $^{\circ}\text{C}$ 、200rpm下在0.5L合成完全培养基中生长72小时。在72小时之后,全细胞发酵液的25ml等分试样用于溶解0、0.5、1、5、10、25、50、75、100g/l当量的丙二酸。将各25ml样本分成五个5ml等分试样,且将各浓度的一个等分试样的pH调整至5.5、6.0、6.5、7.0或7.5。接着离心($\times 6000\text{rcf}$,5分钟,25 $^{\circ}\text{C}$)所有样本,且将上清液转移至单独管中。HPLC分析显示上清液部分中仅存在丙二酸。接着,通过在25 $^{\circ}\text{C}$ 下以与各样本的0.5、1、5、10、25、50、75、100g/l丙二酸等摩尔的当量进行添加,代表性二价阳离子氯化钙用于使丙二酸自澄清上清液部分沉淀。接着如实施例2中所述通过HPLC测量上清液中剩余和沉淀中的丙二酸浓度。在低于5g/L的浓度下,提取效率是10%或以下。一旦pH降低至5.5,所述提取效率也是可忽略的。然而,在较高浓度和6-7.5的pH值下,这个方法十分有效纯化丙二酸。这个方法在100g/L、75g/L、50g/L、25g/L和10g/L下的相应提取效率(自发酵液分离的百分比)如下:pH 7.5=89.8%、88.5%、83.3%、70.1%和66.4%;pH 7.0=88.0%、86.9%、81.8%、71.5%和63.9%;pH 6.5=80.5%、79.0%、75.8%、65.0%和59.0%;pH 6.0=54.4%、55.0%、52.8%、43.0%、

34.9%。

[0274] 这些结果说明本发明的这个方法使丙二酸自发酵液得以纯化,从而使所述丙二酸与酵母细胞与其它溶解化学物质分离。

[0275] 实施例24:通过用乙醇和甲醇进行反应性提取来自发酵液纯化生物源性丙二酸

[0276] 在这个实施例中,通过反应性提取来自发酵液纯化自具有丙二酸生物合成路径的酿酒酵母的培养物内源产生的丙二酸。首先通过离心 ($\times 6000$ rcf, 5分钟) 自50mL发酵液移除宿主细胞。使用自:Gatterman L.和Babsinian VS. “The practical methods of organic chemistry” 第3版, The Macmillan Company: New York, 第161-162 (1916) 页改编的方法来进行发酵液中二乙酯和二甲酯的形成。简而言之, 添加氯化钠至发酵液中; 随后, 添加乙醇和硫酸, 从而导致形成丙二酸二乙酯。

[0277] 实施例25: 使用叔胺自发酵液反应性提取丙二酸

[0278] 如实施例23中所述制备酵母发酵液。外源添加丙二酸至发酵液中以达到最终浓度50g/l。通过添加酸来调整溶液pH至 <4.0 的值。

[0279] 使用1-辛醇作为稀释剂, 在0.25、0.5和0.75mol-TA/kg 1-辛醇下制备三种叔胺(TA)溶液。所用叔胺是三乙胺、三丙胺、三丁胺、三戊胺、三己胺、三庚胺、三壬胺和三癸胺。

[0280] 使相等体积的含有胺和稀释剂的有机溶剂与含有丙二酸的发酵液混合。在1000rpm、25℃下搅拌反应2小时; 随后, 通过离心 ($\times 6000$ rcf, 10分钟) 分离两相。可如实施例2中所述通过HPLC测量各相中的丙二酸浓度。

[0281] 实施例26: 使用甲醇和乙醇自发酵液反应性蒸馏丙二酸

[0282] 这个实施例描述使用甲醇和乙醇自发酵液反应性蒸馏丙二酸; 阳离子交换树脂Amberlyst-15用作固体催化剂。在使用之前, 在70℃下在真空烘箱中干燥树脂6小时。组件由装填有陶瓷的玻璃柱连接于收集容器组成。酯化发生在连接于柱底部的反应器中。将冷凝器放置在柱的顶部以冷凝低挥发性蒸汽。酯化反应器用发酵液进料且离子交换树脂Amberlyst-15 (2%w/w) 作为催化剂加以添加。向反应器施加足够热量以蒸发反应混合物, 且一旦获得所需温度, 就添加甲醇或乙醇至反应器中。自酯化反应器和收集容器的排放阀取出样本。在反应完成之后, 测量来自酯化反应器和收集容器的最终反应混合物的体积和质量。如实施例2中所述测量丙二酸消耗和产物形成。

[0283] 实施例27: 丙二酸合成转化成丙二酸二甲酯和丙二酸二乙酯

[0284] 如下在1L培养烧瓶中自重组酿酒酵母产生生物源性丙二酸: 在缺乏尿嘧啶的合成完全省却培养基(SD)琼脂板上将菌株划线且在30℃下培养; 使个别菌落在30℃下于3mL SD培养基中生长过夜且随后以1%v/v稀释至500ml缺乏尿嘧啶的SD中。在30℃下培养菌株72小时。离心 ($\times 6000g$, 10分钟) 所得发酵液且使上清液与细胞离心块分离。使用氯化钙使丙二酸自发酵液沉淀(参照实施例23)。接着通过添加过量乙醇和等摩尔硫酸以催化费歇尔酯化来使所得丙二酸钙转化成丙二酸二乙酯。

[0285] 实施例28: 通过克诺维纳盖尔冷凝的多布纳改进形式自丙二酸和甲醛产生丙烯酸酯

[0286] 根据本发明通过丙二酸与多聚甲醛在吡啶中缩合来产生丙烯酸酯。在具有磁力搅拌器的3颈圆底烧瓶中进行反应。添加15ml吡啶和15ml甲苯至烧瓶中, 且以5等份添加10g粉末丙二酸; 随后, 历经30分钟时期添加1.1当量(3.2g)多聚甲醛至反应容器中。剧烈搅拌混

合物以促进组分溶解。在0℃下起始反应温度且接着历经数小时的过程增加至50℃直至观察到由形成气泡所证实的二氧化碳形成。在50℃下加热2小时之后,使烧瓶返回至室温且于水中将反应的等分试样稀释100倍并通过使用实施例2中所述的方法进行HPLC加以分析。丙二酸衍生的丙烯酸的样本用权威性标准物在洗脱时间约17.5分钟下共洗脱(参见图2)。

[0287] 尽管克诺维纳盖尔反应的多布纳改进形式已用于产生许多化合物,但它用于丙烯酸酯产生中是新颖和重要的。每年超过十亿千克丙烯酸用于制备自尿布至薄膜和涂料的广泛范围的产品。它当前主要来源于石油且使用本发明中所述的方法进行生产视多聚甲醛的来源而定提供这个大宗化学品的一种部分或任选完全可再生途径。此外,通过本发明的方法产生的丙二酸将实质上比它的石油衍生对应物廉价。这有助于在可与在用石油化学途径竞争的成本下产生可持续性或部分可持续性丙烯酸。

[0288] 实施例29:自丙二酸和甲醛产生戊二酸

[0289] 如实施例8中所述在1L培养烧瓶中自重组酿酒酵母产生生物源性丙二酸。离心($\times 6000g$, 10分钟)所得发酵液且使上清液与细胞离心块分离。使用氯化钙使丙二酸自发酵液沉淀且转化成丙二酸二乙酯。接着根据文献方法使生物源性丙二酸二乙酯与甲醛在吡啶中反应且处理成所需酸(Hedge等(1961) JOC 26:3166-3170)。

[0290] 实施例30:自各种碳源生物催化产生丙二酸。

[0291] 根据本发明,使具有包含CYC1终止子、氨苄青霉素抗性基因、PMB1复制起点、CEN/ARS复制起点和URA3选择标记的丙二酰基-辅酶A水解酶表达载体的酿酒酵母BY4741酵母细胞在酵母发酵培养基中生长,所述培养基包含5g/L硫酸铵、1g/L磷酸二氢钾、0.5g/L硫酸镁、0.1g/L氯化钠、0.1g/L氯化钙、2mg/L肌醇、0.5mg/L硼酸、0.4mg/L泛酸钙、0.4mg/L烟酸、0.4mg/L盐酸吡哆醇、0.4mg/L盐酸硫胺素、0.4mg/L硫酸锌、0.4mg/L硫酸锰、0.2mg/L对氨基苯甲酸、0.2mg/L核黄素、0.2mg/L钼酸钠、0.2mg/L氯化铁、0.1mg/L碘化钾、40 μg /L硫酸铜、2 μg /L叶酸、2 μg /L生物素、10mg/L腺嘌呤、50mg/L L-精氨酸盐酸盐、80mg/L L-天冬氨酸、20mg/L L-组氨酸盐酸盐、50mg/L L-异亮氨酸、100mg/L L-亮氨酸、50mg/L L-赖氨酸盐酸盐、20mg/L甲硫氨酸、50mg/L L-苯丙氨酸、100mg/L L-苏氨酸、50mg/L L-色氨酸、50mg/L L-酪氨酸和140mg/L L-缬氨酸(基础培养基)。在不同发酵中,以下各者用作唯一碳源:20g/L葡萄糖、2% v/v乙醇或2% v/v甘油。在这个实施例中,在TEF启动子的控制之下使用F0PNG8-1丙二酰基-辅酶A水解酶(来自苏云金杆菌幕虫亚种菌株YBT-020;UniProt ID F0PNG8,具有E91S突变)。一式三份用50 μl 培养产生菌株的饱和培养物接种48孔板中的1.5ml补充有2%碳源的基础培养基。用可透气膜覆盖培养板,在30℃下在板振荡器上孵育,且在生长138小时之后取样以对产物累积进行HPLC分析。

[0292] 如实施例2中所述对丙二酸累积进行HPLC分析。结果如下:葡萄糖作为碳源:4.8mM \pm 0.2mM(标准偏差)丙二酸;乙醇作为碳源:7.5mM \pm 0.8mM丙二酸;以及甘油作为碳源:1.7mM \pm 0.1mM丙二酸。这些结果显示测试的碳源都适合用于根据本发明产生丙二酸。

[0293] 实施例31:编码各种丙二酰基-辅酶A水解酶的重组质粒载体的构建和表达以及它们在酵母中产生丙二酸的用途

[0294] 使用如下所列的引物通过PCR自质粒扩增编码由本发明提供的各种丙二酰基-辅酶A水解酶的核酸:EHD3(E124S)引物Y1-11_A13-R(SEQ ID NO:23)/Y1-11_A13-F(SEQ ID NO:24);B9IZZ9-1引物Y0012(SEQ ID NO:25)/Y0013(SEQ ID NO:26);F0PNG8-1引物Y0014

(SEQ ID NO:27)/Y0015 (SEQ ID NO:28);C3ALI3-1引物Y0018 (SEQ ID NO:29)/Y0019 (SEQ ID NO:30);Q81DR3-1引物Y0020 (SEQ ID NO:31)/Y0021 (SEQ ID NO:32);A4XS22-1引物Y0024 (SEQ ID NO:33)/Y0025 (SEQ ID NO:34);E2XN63-1引物Y0026 (SEQ ID NO:35)/Y0027 (SEQ ID NO:36);A5W8H3-1引物Y0028 (SEQ ID NO:37)/Y0029 (SEQ ID NO:38);以及F6AA82-1引物Y0030 (SEQ ID NO:39)/Y0031 (SEQ ID NO:40)。纯化PCR产物被克隆在含有氨苄青霉素抗性盒、PMB1复制起点、CEN/ARS复制起点和URA3选择标记的穿梭载体中TEF1启动子的下游以及CYC1终止子的上游。将所得质粒转化至大肠杆菌感受态宿主细胞中且在含有Cb⁵⁰的LB琼脂板上加以选择。在37℃下过夜孵育之后,将个别菌落接种在48孔板中的2ml LB-Cb⁵⁰中且在37℃下在振荡器上生长5小时,随后分离质粒且通过测序确认。在对含有蛋白质F6AA82 (E95S)的构建体测序后,发现没有预期到的点突变Q348A;这个突变归因于PCR扩增期间的错误。所得蛋白质F6AA82 (E95S/Q348A)在本文中也称为F6AA82-2。Q348A点突变未显示为获得蛋白质活性所必需。

[0295] 酿酒酵母BY4741细胞用作表达各种丙二酰基-辅酶A水解酶的载体的宿主。使用标准程序将质粒载体个别地引入酵母宿主细胞中。在具有实施例30中所述的培养基的琼脂板上选择转化体,所述培养基含有2%葡萄糖作为碳源。

[0296] 如实施例22中所述培养8个水解酶表达性酿酒酵母菌株且如实施例2中所述通过HPLC加以分析。发酵培养基中的相对丙二酸浓度如下(表示为丙二酸峰下积分面积;平均值平均值±S.D.;n=4):酿酒酵母BY4741(阴性对照)48,865±9,345;EHD3 (E124S)94,721±8,115;B9IZZ9 (E91S)261,717±38,012;F0PNG8 (E91S)216,654±31,145;F6AA82 (E95S/Q348A)212,096±29,338;E2XN63 (E95S)198,046±35,084;Q81DR3 (E91S)193,665±37,898;Q63BK8 (E91S)167,477±8,110;以及A5W8H3 (E95S)52,047±9,042。识别符是Uniprot ID(<http://www.uniprot.org/>)继之以由本发明提供的导致丙二酸产生的突变。在由未用酵母细胞接种的培养基组成的样本中未检测到丙二酸。

[0297] 在其它实施例中,分别含有E101S和E95S突变的来自蕈状芽孢杆菌(*Bacillus mycoides*)的C3ALI3和来自门多萨假单胞菌(菌株ymp)的A4XS22用作丙二酰基-辅酶A水解酶。因为培养基条件因缓冲至pH 4.0而略微变化,所以包括F0PNG8-1和F6AA82-1以进行比较;所有其它发酵、取样和分析条件都如上所述。结果如下:C3ALI3 (E101S)10±1mM、A4XS22 (E95S)7±1mM、F0PNG8 (E91S)11±2mM以及F6AA82 (E95S/Q348A)23±2mM丙二酸。在不存在丙二酰基-辅酶A水解酶蛋白质下,酿酒酵母细胞不产生可检测浓度的丙二酸。

[0298] 这个实施例说明根据本发明,丙二酸可在表达含有赋予丙二酰基-辅酶A水解酶活性的X₁突变的酶的酵母宿主细胞中产生。为这个实施例中使用的的所有突变水解酶所共有的E至S活性位点突变可用于这些酶类别的其它成员中以提供类似结果。

[0299] 实施例32:使用单价阳离子使丙二酸自细胞和发酵液沉淀

[0300] 在这个实施例中,通过用单价阳离子进行沉淀来自发酵液纯化丙二酸。单价阳离子是钠。使用外源添加至发酵液中的合成丙二酸说明纯化方法。如实施例23中所述生长、制备以及用丙二酸接种酿酒酵母BY4741的培养物。接着,代表性单价阳离子氯化钠用于使丙二酸自澄清上清液部分沉淀。在25℃下添加2或4摩尔当量(相较于丙二酸浓度)的氯化钠至各样本的上清液部分中。接着如实施例2中所述通过HPLC测量上清液中剩余和沉淀中的丙二酸浓度。用100g/L丙二酸接种且用2摩尔当量的氯化钠处理的培养基中剩余的丙二酸的

浓度根据pH变化如下:pH 5.5=103%、pH 6.0=96%、pH 6.5=71%、pH 7.0=74%、pH 7.5=86%。用100g/L丙二酸接种且用4摩尔当量的氯化钠处理的培养基中剩余的丙二酸的浓度根据pH变化如下:pH 5.5=92%、pH 6.0=86%、pH 6.5=66%、pH 7.0=81%、pH 7.5=86%。

[0301] 实施例33:使用二乙胺自细胞和发酵液纯化丙二酸

[0302] 在这个实施例中,通过添加二乙胺来自发酵液纯化丙二酸。使用外源添加至发酵液中的商购获得的丙二酸说明纯化方法。制备酵母发酵培养基且如先前实施例中所述添加已知量的丙二酸。代表性二取代的胺二乙胺用于自澄清上清液部分纯化丙二酸。在25℃下添加0.5与100g/l之间的二乙胺至各样本的上清液部分中。接着如实施例2中所述通过HPLC如先前所述测量上清液中剩余和沉淀中的代谢物浓度。用100g/L丙二酸接种的用每当量丙二酸4当量的二乙胺处理的培养基中剩余的丙二酸的浓度根据pH变化如下:pH 5.5=100%、pH 6.0=86%、pH 6.5=65%、pH 7.0 67%、pH 7.5=57%。

[0303] 实施例34:合成庚二酸

[0304] 这个实施例描述通过戊二酸半醛与丙二酸缩合以及催化氢化来合成庚二酸。在250ml圆底烧瓶中,将10克戊二酸半醛溶解于40ml吡啶中。添加9克丙二酸,且在80℃下在搅拌下加热混合物5小时。使用旋转蒸发移除吡啶,且将含有中间体庚-2-烯-1-7-二酸的所得物质再溶解于己醇中。添加Pd/C催化剂且在氢气氛围下搅拌所得混合物24小时。通过经硅藻土过滤移除催化剂;在减压下移除溶剂;且使用硅胶快速色谱纯化所得庚二酸。

[0305] 庚二酸是尼龙5,7的关键组分,其在发酵中用于弥补生物素营养缺陷性以及用于产生若干塑料。用于庚二酸合成的先前技术方法是昂贵且低产的。本发明提供一种用于可衍生自由由本发明提供的丙二酸的这个重要化合物的新化学合成法。

[0306] 实施例35:编码其它丙二酰基-辅酶A水解酶的重组载体的构建和表达以及在酵母中产生丙二酸

[0307] 在这个实施例中,使大肠杆菌丙二酰基-辅酶A:ACP转酰基酶FabD (SEQ ID NO:53) 突变以含有在指示位置处的一个或多个以下氨基酸变化:S92C、H201N、R117D、R117E、R117N、R117Y、R117G、R117H、Q11D、Q11E、Q11N、Q11Y、Q11G、Q11H、L93A、L93V、L93I、L93F、L93S、L93G,且用作酿酒酵母中的丙二酰基-辅酶A水解酶。使用引物F1 (5' - ATGACGCAATTTGCATTTGTGTTCCC-3') 和F2 (5' - TTAAAGCTCGAGCGCCGCT-3') 自大肠杆菌菌株K12 PCR扩增编码大肠杆菌FabD的核酸。接着使用标准方法使扩增的基因突变并在TEF1启动子的控制下且在CYC1终止子的上游插入穿梭表达质粒中。这个载体含有氨苄青霉素抗性盒、PMB1复制起点、CEN/ARS复制起点和URA3选择标记。测定的个别突变组合与结果一起列于以下。

[0308] 将个别菌落接种至1ml实施例30中所述的含有2%葡萄糖作为碳源的培养基中。在30℃下在振荡器上孵育培养物24小时,且20μl这些培养物用于接种2ml相同培养基的产生培养物。用可透气膜覆盖产生培养物,在30℃下在振荡下孵育,且在生长96小时和168小时之后取样以对产物累积进行HPLC分析。

[0309] 在由未用酵母细胞接种的培养基组成的样本中未检测到丙二酸。野生型酵母在发酵168小时之后产生小于0.1mM的丙二酸。表达四个FabD变体的任一者所产生的丙二酸的水平都高于未表达这些蛋白质的细胞。在使用在酿酒酵母中表达的各种工程化FabD丙二酰

基-辅酶A-ACP转酰基酶发酵96和168小时之后的丙二酸累积如下:FabD S92C/L93V/R117H 96小时=1.01mM、168小时=2.49mM;FabD L93I/R117Y 96小时=1.47mM、168小时=2.48mM;FabD L93S/R117G 96小时=1.11mM、168小时=2.89mM;FabD L93I/R117Y 96小时=1.64mM、168小时=3.47mM。

[0310] 实施例36:用含三烷基胺的1-辛醇自水反应性提取丙二酸。

[0311] 在这个实施例中,通过用3种三烷基胺(三丙胺、三己胺和三辛胺)进行反应性提取来自水纯化丙二酸。使用添加至蒸馏水中的权威性丙二酸说明纯化方法。添加丙二酸至水中以达到最终浓度100g/l;溶液的pH是约1.5。

[0312] 使250ul丙二酸水溶液与250ul由25%v/v三烷基胺和75%v/v1-辛醇组成的有机相混合。制备一个不添加有机相的样本;这个样本提供各样本中初始丙二酸浓度的测量结果。通过倒置来混合样本18小时,离心($\times 18,000g$)1分钟,且对水相取样以通过HPLC分析丙二酸浓度。

[0313] 对于HPLC分析水相中的丙二酸,我们采用配备有UV检测器的Shimadzu XR HPLC系统。将5 μ l各样本注射至系统中且用Aminex HPX-87h发酵监测柱(Bio-Rad,Hercules,CA)分离。流动相是去离子水(用硫酸调整至pH 1.95),流速是0.6ml/分钟,烘箱温度是50C,且UV检测器在210nm下进行监测。在注射后约10分钟监测样本的丙二酸洗脱。

[0314] $33 \pm 2.6\%$ 丙二酸被三丙胺提取, $73 \pm 4.4\%$ 丙二酸被三己胺提取,且 $89 \pm 11.9\%$ 丙二酸被三辛胺提取($n=3$)。因此,长链长度三烷基胺优于短链长度三烷基胺来增加提取效率。在这个实施例中使用的三烷基胺中,三辛胺优于较短链长度三己胺和三丙胺。

[0315] 实施例37:降低水相pH以增加使用三烷基胺对丙二酸的反应性提取。

[0316] 在这个实施例中,在不同pH值下通过用含三辛胺的1-辛醇进行反应性提取来自水纯化丙二酸。首先于水中制备100g/l丙二酸的储备溶液。接着将储备溶液分成调整至所需pH的工作样本。因为添加碱使在各pH值下的丙二酸浓度稀释,所以获取反应性提取之前与之后两者的丙二酸浓度。计算提取前样本与提取后样本之间的丙二酸浓度差异提供在测试的各pH值下的产率百分比。

[0317] 如下进行反应性提取。使250ul丙二酸水溶液与250ul由25%v/v三辛胺和75%v/v1-辛醇组成的有机相混合。所有样本都通过倒置来混合18小时,离心($\times 18,000g$)1分钟,且对水相取样以如实施例#36中所述通过HPLC分析丙二酸浓度。

[0318] 在各pH下的提取效率如下:pH 1.5,70%;pH 2.26,57%;pH 2.93,45%;pH 4.05,30%;pH 4.62,23%;pH 5.0,15%;pH 5.5,5%;pH 6.0,0%;以及pH 7.0,3%。

[0319] 提取效率随pH增加而降低,且在pH 1.5下获得最高提取效率。高于pH 6.0时的提取效率可忽略。因此,优选的是当使用三烷基胺提取丙二酸时,发酵液的pH低于2.0。

[0320] 实施例38:增加有机相中的三烷基胺浓度来增加自水溶液反应性提取丙二酸。

[0321] 在这个实施例中,使用含有不同浓度三辛胺的三烷基胺/1-辛醇上覆有机相自水纯化丙二酸。使用添加至蒸馏水中以达到最终浓度100g/l的合成丙二酸说明纯化方法;溶液的pH是约1.5。

[0322] 使250ul丙二酸水溶液与250ul由含指示量的三烷基胺(表示为相对于水相中的丙二酸的摩尔分数)的1-辛醇组成的有机相混合。制备一个不添加有机相的样本且用于提供水相中初始丙二酸浓度的测量结果。

[0323] 通过倒置来混合样本18小时,离心($\times 18,000g$)1分钟,且对水相取样以如实施例#36中所述通过HPLC分析丙二酸浓度。

[0324] 观察到提取效率与三辛胺:丙二酸摩尔分数之间存在线性关系;具体来说,所述线性关系存在于三辛胺:丙二酸摩尔分数0至1之间。当摩尔比高于1时,100%丙二酸被提取至有机相中。因此,有机相中的三烷基胺的量必须与水相(例如发酵液)中的丙二酸的量等摩尔以使提取产率最大化。理想地,将添加大于等摩尔量的三烷基胺至有机相中以补偿归因于发酵液中的其它有机酸和阴离子的提取效率降低。

[0325] 实施例39:通过使用等摩尔量的于1-辛醇中的三辛胺增加在不同离子强度下自水溶液反应性提取丙二酸的产率。

[0326] 在这个实施例中,通过用三辛胺进行反应性提取来在不同离子强度下自水纯化丙二酸;通过使有机相中的三辛胺与水相中的丙二酸的摩尔比增加至1:1,我们能够改进丙二酸提取效率。

[0327] 添加丙二酸至水中以达到最终浓度100g/l。通过添加氯化钠来调整离子强度至指示浓度。在添加丙二酸和氯化钠之后,所有样本的最终pH都是约pH 1.5。

[0328] 使250 μ l丙二酸水溶液与250 μ l由含三辛胺的1-辛醇组成的有机相混合。计算三辛胺体积以使三辛胺:丙二酸摩尔比如所指示;其余体积是1-辛醇。制备一个不添加有机相的样本;这个样本提供初始丙二酸浓度的测量结果。通过倒置来在18 $^{\circ}$ C下混合样本18小时,离心($\times 18,000g$)1分钟,且对水相取样以如实施例36中所述通过HPLC分析丙二酸浓度。

[0329] 使用0.59、1.07、1.61和2.14摩尔当量的三辛胺,离子强度(mM浓度)如下影响丙二酸的提取:在离子强度0下,分别回收63.5%、98.6%、99.5%和99.4%的丙二酸;在离子强度75下,分别回收60.3%、91.3%、93.7%和93.6%的丙二酸;在离子强度150下,分别回收55.9%、85.8%、89.5%和89.1%的丙二酸;在离子强度225下,分别回收52.4%、83.5%、86.0%和85.9%的丙二酸;在离子强度300下,分别回收46.5%、79.4%、83.1%和82.6%的丙二酸;在离子强度375下,分别回收44.3%、76.7%、79.9%和80.1%的丙二酸;在离子强度450下,分别回收43.3%、72.8%、77.0%和78.1%的丙二酸;在离子强度500下,分别回收43.1%、72.2%、75.1%和76.6%的丙二酸。

[0330] 在测试的所有离子浓度下,增加三辛胺:丙二酸摩尔比至1:1都使自水溶液提取丙二酸的效率增加。通过使三辛胺:丙二酸摩尔比增加高于1:1,未获得丙二酸提取效率的有意义改进。当自水溶液提取丙二酸时,优选使用三辛胺:丙二酸摩尔比至少1:1,且如果不存在其它有机酸,那么摩尔比恰好1:1是优选的。如果存在污染性有机酸,那么摩尔比高于1:1可优选用以改进丙二酸提取效率。

[0331] 实施例40:通过增加水相的离子强度增加自三烷基胺/1-辛醇有机相反提取丙二酸的产率。

[0332] 在这个实施例中,我们说明用以自由含三辛胺的1-辛醇组成的有机相反提取丙二酸且进入水相中的方法。

[0333] 添加丙二酸至水中以达到最终浓度100g/l且提取至由三辛胺(在与丙二酸的摩尔比1:1下使用)溶解于1-辛醇中组成的有机相中。在18 $^{\circ}$ C下持续18小时将10ml水溶液提取至10ml有机相中。接着通过离心($\times 4000g$,5分钟)分离有机相。也对水相取样以定量非提取的丙二酸浓度。

[0334] 接着使250 μ l有机相与250 μ l水溶液在给定离子强度氯化钠或氢氧化钠下混合。通过倒置来在18℃下混合样本18小时。接着离心(x18,000g,1分钟)样本且对水相取样以如对于实施例36所述通过HPLC分析丙二酸浓度。

[0335] 相对于反提取至蒸馏水中的丙二酸浓度来使以下呈现的数据标准化。添加0.5摩尔的氯化钠或氢氧化钠使反提取反应的产率增加超过25倍。添加高于1.5M的氯化钠使反提取产率降低;然而,其中氢氧化钠反提取产率随进一步添加氢氧化钠而继续改进。通过高温(即高于18℃)下操作反提取反应,反提取产率将进一步增加。NaCl的毫摩尔浓度和改进倍数:0mM,1;0.5mM,27.3;1mM,43.1;1.5mM,74.5;2mM,63.4;2.5mM,66.5;3mM,69.4;3.5mM,63.7;4mM,57.8;4.5mM,53.3;以及5mM,45.2。NaOH的毫摩尔浓度和改进倍数:0mM,1;0.5mM,33.6;1mM,46.5;1.5mM,57.5;2mM,52.7;2.5mM,56.4;3mM,62.6;3.5mM,58.2;4mM,68.5;4.5mM,67;5mM,74.2。

[0336] 实施例41:通过酯化和随后相分离自水溶液纯化丙二酸

[0337] 在这个实施例中,我们说明丙二酸和乙醇酯化以形成丙二酸二乙酯以及随后将丙二酸二乙酯相分离至己烷有机相中。

[0338] 制备500 μ l含有100g/l浓度的丙二酸、指示浓度的乙醇和指示浓度的硫酸的水溶液。丙二酸和乙醇是用于形成丙二酸二乙酯的底物;添加硫酸作为催化剂。添加250 μ l己烷的有机上覆层至各样本中,通过倒置来在25℃和大气压下混合它们18小时,离心(x18,000g)1分钟,且对水相取样以如实施例36中所述通过HPLC分析丙二酸浓度。

[0339] 通过水相中丙二酸浓度的降低来度量丙二酸的消耗。自含有0%v/v乙醇和0%v/v硫酸的样本确定水相中的基线丙二酸浓度(即100%未反应丙二酸)。

[0340] 向反应混合物中添加乙醇与硫酸两者为催化丙二酸酯化所必需。在高乙醇和硫酸浓度下获得如由丙二酸的消耗百分比度量的最优选反应条件。也就是,大于40%v/v乙醇(与丙二酸的摩尔比7.14)和大于10%v/v硫酸(与丙二酸的摩尔比1.94)为催化丙二酸接近完全消耗所必需。这个说明的结果如下:用5、10、20、30、40和50%(V/V)乙醇进行酯化反应。水相中剩余的丙二酸分别如下:5%(V/V)硫酸,78.9%、61.6%、43.2%、43.2%、28.2%和16.6%;10%(V/V)硫酸,83.5%、70.3%、49.5%、28.6%、16.6%和8.5%;5%(V/V)硫酸,71.8%、53.2%、28.1%、11.8%、4.4%和0.6%。无酸对照显示95.1%、93.2%、85.8%、69.1%、63.0%和87.3%丙二酸剩余在具有乙醇的相应溶液中。在无乙醇存在下添加酸导致以下浓度的丙二酸剩余在水相中:对于5、10和20%(V/V)硫酸而言分别是96.1%、92.7%和84.4%。

[0341] 实施例42:通过增加乙酰基-辅酶A羧化酶活性改进工程化酵母中的丙二酸生物合成

[0342] 除用于表达异源丙二酰基-辅酶A水解酶和体内产生丙二酸的方法、载体和宿主细胞之外,本发明也提供用于改进丙二酸的效价、产率和/或生产力的方法和宿主细胞。在一个方面,通过增加丙二酰基-辅酶A的生物合成来改进丙二酸产生。丙二酰基-辅酶A是自乙酰基-辅酶A生物合成丙二酸中的倒数第二个中间体,且在酿酒酵母中,这个反应是由乙酰基-辅酶A羧化酶(ACC1)催化。

[0343] 在这个实施例中,使用针对在酿酒酵母中表达的标准密码子最优化表,将来自酵母解脂耶罗威亚酵母CLIB122(NCBI参照序列:XP_501721)的ACC回翻译成DNA序列。Y1ACC

DNA序列的结果是作为SEQ ID NO:11包括在以下。

[0344] 使用标准方案以单片DNA形式合成依次编码上游酿酒酵母FAA1同源物的50个核苷酸、酿酒酵母TEF1启动子序列的300个碱基对、Y1ACC序列(SEQ ID NO:11)以及最后下游酿酒酵母FAA1的50个核苷酸的DNA且通过PCR加以扩增。使用标准乙酸锂转化程序,将所得侧接Y1ACC PCR产物和含有可选择URA3基因和侧接FAA1同源序列以直接重组至染色体上的FAA1位点中的线性DNA构建体共转化至含有由本发明提供的丙二酰基-辅酶A水解酶F6AA82-2的2个拷贝的BY4741(菌株LYM004)中,且通过涂铺在含有2%葡萄糖且缺乏尿嘧啶(-Ura)的合成确定培养基(SD)板上进行选择。分离克隆和相应基因组DNA且通过PCR验证以含有整合至FAA1中的Y1ACC。

[0345] 对于体内丙二酸产生,将挑选的各转化体的一个个别菌落接种在96孔板中的SD-Ura培养基的50 μ l等分试样中。使LYM004生长以比较缺乏异源Y1ACC的表达的酵母。在30℃下在振荡器上将板孵育约4小时且25 μ l的这些预培养物用于接种发酵板。在这个实施例中,48孔板中的2ml SD培养基用于发酵。用可透气膜覆盖发酵板,在30℃下在板振荡器上孵育,且在发酵72和120小时之后取样以对产物累积进行HPLC分析。如实施例36中所述对发酵液中的丙二酸累积进行HPLC分析。

[0346] 在仅由培养基组成的样本中未检测到丙二酸。表达Y1ACC的LYM004在72和120小时产生的丙二酸分别比单独LYM004多1.94和2.43倍。观察到在发酵120小时之后收集的样本中的发酵培养基中测量的丙二酸浓度最高。

[0347] 实施例43:使用工程化F6AA82丙二酰基-辅酶A水解酶在酿酒酵母中产生丙二酸

[0348] 在这个实施例中,在蛋白质F6AA82(Q348A)中的位置E95处引入全部20个蛋白基因氨基酸点突变。据信氨基酸位置95与丙二酰基-辅酶A的末端相互作用,且E95的突变将把丙二酰基-辅酶A水解酶活性引入蛋白质中。

[0349] 使用标准方法构建突变在位置95处的F6AA82(Q348A)(即全部氨基酸)且克隆至含有尿嘧啶营养缺陷标记、CEN/ARS复制起点、TEF启动子和CYC终止子的酵母质粒中。将F6AA82突变体克隆在TEF启动子之后且直接在CYC终止子的上游。所有质粒都含有氨苄青霉素抗性标记和用于质粒在大肠杆菌中增殖的ColE1起点。

[0350] 使用乙酸锂程序将质粒转化至酿酒酵母中且在30℃下在SD-尿嘧啶琼脂板上选择转化体。将各突变体的12个菌落接种至96孔板中的500 μ l SD-Ura培养基的预培养物中且在30℃下在振荡下孵育约16小时。这些预培养物的5 μ l等分试样用于接种含有500 μ l RD4-Ura(1X YNB、3X SC补充剂、2%葡萄糖、75mM丁二酸盐缓冲剂(pH 4.0),尿嘧啶省却)培养基的96孔产生板。在30℃下在振荡下孵育产生板3天,随后对发酵液取样。通过离心使样本澄清且在0.45 μ m膜上过滤,随后进行HPLC分析。如实施例36中所述对发酵液中的丙二酸累积进行HPLC分析。

[0351] 各F6AA82突变体(除所述E95突变之外也含有突变Q348A)在这些条件下的产量如下:F6AA82(E95N) 8.03 \pm 0.14mM, F6AA82(E95S) 4.18 \pm 0.61mM, F6AA82(E95Y) 3.87 \pm 0.30mM, F6AA82(E95A) 2.33 \pm 0.50mM, F6AA82(E95K) 1.65 \pm 0.23mM, F6AA82(E95T) 1.16 \pm 0.62mM, F6AA82(E95D) 0.75 \pm 0.27mM, F6AA82(E95F) 0.17 \pm 0.13mM, F6AA82(E95V) 0.11 \pm 0.32mM, F6AA82(E95L) 0.11 \pm 0.12mM, F6AA82(E95G) 0.00 \pm 0.01mM, F6AA82(E95P) -0.31 \pm 0.78mM, F6AA82(E95R) -0.32 \pm 0.69mM, F6AA82(E95) -0.32 \pm 0.40mM, F6AA82(E95W) -0.34 \pm

0.73mM, F6AA82 (E95Q) -0.37±0.74mM, F6AA82 (E95H) -0.53±0.70mM, F6AA82 (E95C) -0.57±0.69mM, F6AA82 (E95I) -0.64±0.81mM, 以及F6AA82 (E95M) -0.77±0.79mM。丙二酸的负浓度是由于产生效价下降至HPLC的可检测水平以下且指示不存在丙二酸产生。

[0352] 含有突变E95S、E95Y、E95T、E95N、E96K、E95V和E95D的F6AA82 (Q348A) 蛋白产生的丙二酸显著(t检验, $p < 0.05$) 多于F6AA82 (Q348A) 蛋白。含有这些突变的F6AA82蛋白适于水解丙二酰基-辅酶A以及产生产生丙二酸。F6AA82突变E95S和E95N优选用于水解丙二酰基-辅酶A以及产生丙二酸。

[0353] E95Y、E95T、E95K、E95V和E95D产生的丙二酸显著(t检验, $p < 0.05$) 多于野生型F6AA82蛋白。含有这些突变的F6AA82蛋白适于水解丙二酰基-辅酶A以及产生产生丙二酸。

[0354] 实施例44: 启动子选择对在酿酒酵母中产生丙二酸的影响

[0355] 在这个实施例中, 自质粒Y20扩增包含CEN/ARS复制起点、PMB1复制起点、氨苄青霉素抗性标记、URA3标记和hph标记的表达质粒骨架。通过PCR自质粒Y1-F6AA82-2扩增由CYC1终止子侧接的丙二酰基-辅酶A水解酶F6AA82-2且使用标准技术与骨架一起装配。所得质粒pPLIB0-Q9I用作用于克隆所选不同启动子的骨架。

[0356] 自酿酒酵母菌株BY4741选择96个参照基因。使用酿酒酵母菌株BY4741基因组DNA作为模板, 通过PCR扩增对应于约750个碱基对的核酸片段来产生这些基因的启动子序列, 所述碱基对紧靠所述基因的开放阅读框的上游, 包括目标基因的起始密码子。使用标准克隆技术, 克隆紧靠pPLIB0-Q9I中的F6AA82-2 (起始密码子在SEQ ID NO:12中的3739处) 的上游的这些启动子序列。使所得质粒在大肠杆菌中增殖且通过测序来验证存在所需启动子。

[0357] 上述质粒用于使用标准乙酸锂程序转化酿酒酵母BY4741。在30℃下在无尿嘧啶的CSM培养基的琼脂板上选择转化体。将具有各质粒的一个转化体接种在由96孔板中的300u1 RD-Ura培养基 (2X YNB、3X SC-U、2%葡萄糖) 组成的预培养板中。在30℃下在振荡下孵育预培养板约20小时, 且30u1这些预培养物用于接种由48孔培养板中的1.8ml RD4-Ura培养基 (2X YNB、3X SC-U、2%葡萄糖、75mM丁二酸盐缓冲剂 (pH 4.0)) 组成的产生板 (各预培养物对应1个产生孔)。在30℃下在振荡下孵育各板144小时, 随后对发酵液取样。通过离心使样本澄清且在0.45uM膜上过滤, 随后进行HPLC分析。

[0358] 如实施例36中所述对发酵液中的丙二酸累积进行HPLC分析。在与驱动F6AA82-2的表达的不同启动子一起发酵144小时之后, 发酵液中的丙二酸毫摩尔浓度如下: HSP150 8.8; PGK1 8.6; PHO5 8.5; SCT1 7.8; PRB1 7.6; TPI1 7.1; ACH1 7; HXK2 6.9; AC01 6.9; JEN1 6.9; MDH2 6.8; POX1 6.6; CIT1 6.6; ALD4 6.6; ADH1 6.5; TDH3 6.4; ADH2 6.4; SDH1 6.4; TDH1 6.1; MLS1 6.1; RPN6 6; GLK1 5.9; POT1 5.8; HSP26 5.8; FBA1 5.7; LPD1 5.7; CYC1 5.5; COX5a 5.5; TEF1 5.5; SHH4 5.5; GND2 5.5; TPS1 5.5; MDH1 5.4; PDC1 5.4; HXK1 5.3; TDH2 5.3; IDH2 5.3; DDR2 5.2; SLT2 5.2; ENO2 5.1; COX6 5; CH01 4.9; PHO3 4.9; PFK1 4.9; ACS1 4.9; GUT2 4.8; PHM7 4.8; CIT2 4.7; ACS2 4.7; ALD2 4.5; IDH1 4.5; IDP2 4.5; FBP1 4.3; PHO12 4.3; PDC5 4.2; PFY1 4.1; GDH1 4.1; PEX13 4.1; ICT1 4.1; YSA1 4; KGD2 3.9; GIM5 3.9; GAP1 3.9; DBP5 3.8; STE5 3.7; BIO2 3.7; PDC6 3.6; HXT5 3.6; REG1 3.6; TPS3 3.5; BIO5 3.3; PHO8 3.3; IRC7 3.2; GPM1 3.2; ALA1 3.1; KGD1 3.1; MIG1 3.1; YKT6 2.9; SN04 2.8; ARA1 2.8; PDR10 2.7; YBR139W 2.3; ERR2 2.2; CRC1 1.9; TSL1 1.3; ENO1 1.2; PFK2 1.1; RPL6A 1.1。因为这些启动子中的许多未命名, 所以在各数值前方的字

母数字码指示在酿酒酵母菌株BY4741中紧靠克隆的启动子的下游的基因。

[0359] 这个实施例说明广泛多种启动子可用于控制丙二酰基-辅酶A水解酶的表达以在酵母中体内产生丙二酸。此外,在这个实验中,40个不同启动子产生至少5mM丙二酸的效价,且7个启动子导致产生至少7mM。

[0360] 实施例45:使丙二酰基-辅酶A水解酶编码基因整合至酿酒酵母基因组中以产生丙二酸。

[0361] 在这个实施例中,酿酒酵母菌株BY4741用作用于基因组整合由本发明提供的丙二酰基-辅酶A水解酶的宿主菌株。选择酿酒酵母基因组上的三个位点用于在这个菌株中整合合成核酸构建体且待整合的三个线性核酸盒各自携带独特同源性位点以靶向整合,且独特可选择标记编码丙二酰基-辅酶A水解酶的上游。在全部三例中,编码由TEF1启动子和CYC1终止子侧接的丙二酰基-辅酶A水解酶F6AA82 (E95S/Q348A) 的核酸用作盒的表达部分。

[0362] 在两例中,靶向作为染色体的非编码DNA的单独长末端重复序列YPRC δ 15和YORW δ 22以达成整合。在第三例中,替代BUD21基因来整合由本发明提供的丙二酰基-辅酶A水解酶盒(F6AA82 (E95S/Q348A))。这些线性插入盒的序列作为SEQ ID NO:41、SEQ ID NO:42和SEQ ID NO:43包括在以下。独特整合位点和可选择标记分别是YPRC δ 15 (HIS3)、YORW δ 22 (LEU2) 和BUD21 (URA3)。

[0363] 使用标准方案构建这些核酸盒且转化至酿酒酵母BY4741中。在30℃下在由省却适当氨基酸的SC培养基制得的琼脂板上选择转化体。在选择性琼脂板上将各构建的若干转化体重划线。接着筛选转化体且通过丙二酸的产生以及通过PCR加以验证。以下菌株由这些整合所产生:LYM002含有整合在YPRC δ 15 (HIS3) 处的F6AA82 (E95S/Q348A) 的1个拷贝;LYM004携带F6AA82 (E95S/Q348A) 的2个拷贝,1个在YPRC δ 15 (HIS3) 处,第2个在YORW δ 22 (LEU2) 处;LYM007源于LYM004且具有整合在BUD21基因座处的F6AA82 (E95S/Q348A) 的第三拷贝。

[0364] 对于发酵和丙二酸产生,携带质粒Y1-F6AA82 (E95S/Q348A) 的酿酒酵母BY4741充当阳性对照。在30℃下在振荡下孵育96孔板中的500 μ l RD4培养基的预培养物16-20小时。这些预培养物的5 μ l等分试样用于接种含有500 μ l RD4培养基(1X YNB、3X SC补充剂、2%葡萄糖、75mM丁二酸缓冲剂(pH 4.0))的96孔产生板。在30℃下在振荡下孵育产生板120小时,随后对发酵液取样。如实施例36中所述制备样本且通过HPLC加以分析。

[0365] 用权威性标准物建立的标准曲线用于确定发酵液中的丙二酸浓度。由这些这些发酵获得的丙二酸产生水平如下:酿酒酵母BY4741+Y1-F6AA82 (E95S/Q348A), 2.6 ± 0.4 mM; LYM002, 3.2 ± 0.2 mM; LYM004, 4.6 ± 0.2 mM; LYM007, 7.7 ± 0.3 mM。

[0366] 这个实施例说明由本发明提供的丙二酰基-辅酶A水解酶编码基因的基因组整合的益处,因为基因的单一整合拷贝产生的丙二酸效价高于相同基因自质粒表达所获得的效价。这个实施例也用于说明由本发明通过细胞中存在的丙二酰基-辅酶A水解酶编码基因的数目来调节丙二酸。

[0367] 实施例46:表达乙酰基-辅酶A合成酶以增加自工程化宿主细胞产生丙二酸。

[0368] 乙酰基-辅酶A合成酶(ACS)为许多生物体所共有。这些酶使用乙酸和三磷酸腺苷(ATP)作为底物来产生乙酰基-辅酶A。根据本发明,这些酶可用于使来自任何来源(包括内源代谢、原料水解或馈入)的乙酸转化成作为丙二酰基-辅酶A生物合成中的倒数第二个前体的乙酰基-辅酶A。通过增加乙酰基-辅酶A水平,可增加丙二酰基-辅酶A,且因此增加由本

发明提供的宿主细胞中的丙二酸。

[0369] 适于这个目的的说明性ACS可自细菌肠道沙门氏菌获得。为在酿酒酵母中表达肠道沙门菌ACS (SeACS), 使用密码子最优化表使氨基酸序列回翻译成DNA序列, 且使被确定涉及于降低酶活性的一个残基突变 (L641P)。由本发明提供的所得密码子最优化DNA序列是 SEQ ID NO:44。

[0370] 自寡核苷酸重新合成编码SeACS1 (L641P) 的核酸且使用标准方案插入在来自酿酒酵母的TEF1启动子的控制下的酵母质粒Y1中。这个质粒也含有CEN/ARS起点和赋予尿嘧啶原养型的基因。通过在SD-尿嘧啶板上涂铺来选择含有这个质粒的LYM004的克隆且通过测序验证以含有SeACS1 (L641P) 基因。

[0371] LYM004用作用于表达SeACS1 (L641P) 乙酰基-辅酶A合成酶的宿主。对于体内丙二酸产生, 将挑选的各转化体的一个个别菌落接种在96孔板中的SD-Ura培养基的50 μ l等分试样中。使LYM004生长以比较缺乏异源SeACS1 (L641P) 的表达, 但含有Y1空载体的酵母。在30 $^{\circ}$ C下在振荡器上将板孵育约48小时, 且50 μ l这些预培养物用于接种发酵板。在这个实施例中, 96孔板中的500 μ L RD4U培养基用于发酵。用可透气膜覆盖发酵板, 在30 $^{\circ}$ C下在板振荡器上孵育, 且在发酵120小时之后取样以对产物累积进行HPLC分析。如实施例36中所述进行HPLC分析。

[0372] 在由未用酵母细胞接种的培养基组成的样本中未检测到丙二酸。在所测试的条件下, 表达SeACS1的LYM004在120小时产生的丙二酸比具有单独空Y1对照的LYM004多1.86倍 (5.5 ± 0.62 mM), 由此说明由本发明的这个实施方案提供的丙二酸产生改进。

[0373] 实施例47: 使用修饰的库德里阿兹威毕赤酵母菌株利用不同碳源产生丙二酸。

[0374] 菌株LPK3003是通过基因组整合编码由P_kTEF1启动子驱动的hph潮霉素(hygromycin)磷酸转移酶(赋予对潮霉素B的抗性)、以及由P_kTDH1启动子驱动的F6AA82 (E95S/Q348A) 丙二酰基-辅酶A水解酶的核酸盒, 由库德里阿兹威毕赤酵母菌株Y-134 (自USDA Agricultural Research Services, Peoria, IL获得) 获得。

[0375] 使LPK3003的种子培养物在30 $^{\circ}$ C下在振荡下在YPD培养基中生长16-20小时。这个种子培养物用于接种1.1X YNB, 接着在每孔450 μ l下等分于96孔板中。添加各种碳源(葡萄糖、蔗糖、乙醇、甘油或来自乙酸钠的乙酸)的50 μ l溶液至各孔中(一式三份)以达到最终浓度2% (w/v)。在30 $^{\circ}$ C下在振荡下将板孵育115小时且对发酵液取样。如实施例36中所述制备样本且通过HPLC加以分析。

[0376] 在这个实施例中, 对于蔗糖或乙酸盐碳源, 观察到存在少许生长或不存在生长, 且这些样本中的丙二酸累积可忽略。如上所述, 宿主细胞可通过分别引入蔗糖转化酶和/或乙酰基-辅酶A合成酶(ACS)来被修饰以赋予或增强蔗糖和/或乙酸盐分解代谢。这个实施例中测试的其它碳源导致丙二酸累积(平均平行测定值 \pm S.D.): 4.6 ± 0.5 mM(自葡萄糖)、 5.38 ± 0.05 mM(自甘油)以及 3.7 ± 0.3 mM(自乙醇)。

[0377] 这些结果说明多种碳源可用于自表达丙二酸-辅酶A水解酶的工程化库德里阿兹威毕赤酵母菌株产生丙二酸。值得注意的是, 甘油在这个实施例中提供最高效价。

[0378] 实施例48: 基于生物反应器的丙二酸产生。

[0379] 在这个实施例中, 以分批馈料控制方式使酵母菌株LYM004 (对于构建细节, 参见实施例45) 在0.5L生物反应器中生长。自SC板分离LYM004的单一菌落且在5mL RD4培养基中培

养(对于配方,参见实施例43)。在30℃下,在200rpm下振荡下维持培养物过夜。4mL培养物用于接种250mL非挡板烧瓶中的50mL新鲜RD4培养基且在30℃、200rpm下生长过夜。时间零点OD 600nm吸光度是0.304。在过夜生长(16小时)之后,这个培养物用于接种1L RD4培养基。将这个培养物分成2个单独500mL等分试样且添加至两个单独生物反应器中。两个发酵均维持在30℃下,其中单一叶轮在400rpm下运转,且使用压缩空气达成的喷射速率是每分钟1个容器体积(VVM)。使培养物生长过夜(21小时)以使在起始分批进料阶段之前达成葡萄糖消耗。每980秒递送进料(配方在以下)2秒。每日获取0.5mL样本且分析丙二酸的产生。在4天之后,培养物已累积34mM丙二酸且达到OD 600nm 16.2。在9天之后,培养物已累积116mM丙二酸且达到OD 600nm 52.1。

[0380] 分批进料培养基由17g/L Difco YNB;50g/L硫酸铵;49.8g/L缺乏组氨酸、甲硫氨酸和亮氨酸的合成完全(SC)补充剂;2.57g/L甲硫氨酸;8.85g/L丁二酸和20g/L葡萄糖组成;调整pH至4.0。

[0381] 对表达丙二酰基-辅酶A水解酶的3个拷贝的菌株LYM007(参见实施例45)重复这个方案。在这个实施例中,单一叶轮在700rpm下运转且进料的组成如下:68g/L YNB(Sigma)、16.6g/L SC-his-met-leu(Sunrise Science)、1.284g/L甲硫氨酸、75mM丁二酸盐缓冲剂(pH 4.0)、400g/L葡萄糖。所用进料速率是每980秒循环5秒。每日获取0.2mL样本且通过HPLC分析丙二酸、乙酸、丁二酸、丙酮酸的产生。通过测量600nm下的光密度(OD600)来监测生长。各时间点的OD600和丙二酸浓度如下:18小时,OD=6.9,3.2mM;45小时,OD=13.8,14.0mM;69小时,OD=16.6,24.2mM;88.5小时OD=24.4,31.1mM;111.8小时,OD=36.9,47.8mM。

[0382] 在比较实施例中,以分批进料控制方式使LPK3003(参见实施例47)在0.5L发酵罐中生长。自YPD板分离LPK3003的单一菌落且在5mL YPD培养基中培养。在30℃下,在200rpm下振荡下使培养物生长过夜。5mL培养物用于接种50mL含有6.8g/L YNB(Difco)和2g/L葡萄糖的新鲜基本培养基且在30℃、200rpm下生长过夜。这个培养物接着用于接种500mL 6.8g/L YNB(Difco)培养基、2g/L葡萄糖且在1升发酵罐中生长。温度维持在30℃下,单一叶轮在700rpm下运转,且使用压缩空气,喷射速率被设置成每分钟1个容器体积(VVM)。使培养物生长过夜21小时以使在起始分批进料阶段之前达成葡萄糖消耗。通过将发酵罐进料设置成每980秒循环5秒来起始含有68g/L YNB和400g/L葡萄糖的进料。每日获取200uL样本且分析生长(OD600)和丙二酸的产生(如实施例36中所述)。通过在每个取样间隔下添加200uL消泡剂来控制起泡。在每个取样间隔下测量pH且用10N NaOH调整至4.5。丙二酸继续产生超过111小时,从而累积36.4mM丙二酸。重要的是,这是首次显示在不添加氨基酸、尿嘧啶或腺嘌呤下达成产生。各时间点的OD600、丙二酸浓度和pH如下:18小时,OD=2.1,0.7mM,pH=2.3;45小时,OD=12.2,8.8mM,pH=2.0;69小时,OD=12.8,18.1mM,pH=2.0;88.5小时,OD=27.6,22.9mM,pH=2.7;111.8小时,OD=40.1,36.4mM,pH=2.4;194小时,OD=60.6,94mM,pH=3.3。

[0383] 尽管来自LYM007的丙二酸的效价较高,但可使LPK3003生长所处的条件(基本培养基,低pH)以及它仅含有F6AA82-2丙二酰基-辅酶A水解酶的单一拷贝的事实使得这个菌株成为用于进行其它工程化尝试的优越菌株。

[0384] 本申请还涉及以下实施方案:

- [0385] 1. 一种重组宿主细胞, 其为酵母细胞或细菌细胞, 其包含异源丙二酰基-辅酶A水解酶, 其中所述异源丙二酰基-辅酶A水解酶选自:
- [0386] (a) SEQ ID NO: 46、49、51和52, 其中X1选自S、T、N、Q和Y;
- [0387] (b) SEQ ID NO: 45和48, 其中X1选自R、H、K、S、T、N、Q和Y;
- [0388] (c) SEQ ID NO: 47, 其中X1是S;
- [0389] (d) SEQ ID NO: 50, 其中X1选自A、D、K、S、T、N和Y。
- [0390] 2. 根据实施方案1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 45, 其中X1是N。
- [0391] 3. 根据实施方案1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 45, 其中X1是S。
- [0392] 4. 根据实施方案1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 45, 其中X1是Y。
- [0393] 5. 根据实施方案1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 46, 其中X1是N。
- [0394] 6. 根据实施方案1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 46, 其中X1是S。
- [0395] 7. 根据实施方案1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 46, 其中X1是Y。
- [0396] 8. 根据实施方案1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 47, 其中X1是S。
- [0397] 9. 根据实施方案1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 48, 其中X1是N。
- [0398] 10. 根据实施方案1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 48, 其中X1是S。
- [0399] 11. 根据实施方案1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 48, 其中X1是Y。
- [0400] 12. 根据实施方案1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 49, 其中X1是N。
- [0401] 13. 根据实施方案1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 49, 其中X1是S。
- [0402] 14. 根据实施方案1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 49, 其中X1是Y。
- [0403] 15. 根据实施方案1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 50, 其中X1是A。
- [0404] 16. 根据实施方案1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 50, 其中X1是N。
- [0405] 17. 根据实施方案1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是SEQ ID NO: 50, 其中X1是S。
- [0406] 18. 根据实施方案1所述的重组宿主细胞, 其中所述异源丙二酰基-辅酶A水解酶是

SEQ ID NO:50,其中X1是Y。

[0407] 19.根据实施方案1所述的重组宿主细胞,其选自假丝酵母属(*Candida*)、隐球酵母属(*Cryptococcus*)、*Komagataella*、油脂酵母属(*Lipomyces*)、毕赤酵母属(*Pichia*)、红冬孢酵母属(*Rhodospiridium*)、红酵母属(*Rhodotorula*)、酵母属(*Saccharomyces*)、毛孢子菌属(*Trichosporon*)、耶罗威亚酵母属(*Yarrowia*)、芽孢杆菌属(*Bacillus*)、梭菌属(*Clostridium*)、棒杆菌属(*Corynebacterium*)、埃希氏菌属(*Escherichia*)、假单胞菌属(*Pseudomonas*)或链霉菌属(*Streptomyces*)的种。

[0408] 20.根据实施方案19所述的重组宿主细胞,其为毕赤酵母属、酵母属或大肠埃希菌属的种。

[0409] 21.根据实施方案20所述的重组宿主细胞,其为库德里阿兹威毕赤酵母(*Pichia kudriavzevii*)。

[0410] 22.根据实施方案20所述的重组宿主细胞,其为酿酒酵母(*Saccharomyces cerevisiae*)。

[0411] 23.根据实施方案20所述的重组宿主细胞,其为大肠杆菌(*Escherichia coli*)。

[0412] 24.根据实施方案1所述的重组宿主细胞,其进一步包含以下的一种或多种:

[0413] (a) 异源丙二酰基-辅酶A合成酶,其选自酿酒酵母ACS1、酿酒酵母ACS2、肠道沙门氏菌(*Salmonella enterica*)ACS、大肠杆菌AcsA和枯草芽孢杆菌(*Bacillus subtilis*)AcsA;

[0414] (b) 异源酿酒酵母丙酮酸脱氢酶复合物酶,其选自PDA1、PDB1、LAT1、LPD1和PDX1;

[0415] (c) 异源乙醇分解代谢路径酶;

[0416] (d) 异源ATP柠檬酸裂解酶;

[0417] (e) 异源乙酰基-辅酶A羧化酶;

[0418] (f) 降低丙二酸分解代谢的遗传修饰;

[0419] (g) 异源转运蛋白,其选自PDR5、PDR10、PDR11、PDR12、PDR15和PDR18;

[0420] (h) 增加宿主细胞对丙二酸的耐受性的遗传修饰;

[0421] (i) 改进宿主细胞对碳源的分解代谢的遗传修饰,所述碳源选自乙醇、乙酸盐和蔗糖。

[0422] 25.根据实施方案22所述的重组宿主细胞,其进一步包含乙酰基-辅酶A羧化酶,所述乙酰基-辅酶A羧化酶是解脂耶罗威亚酵母乙酰基-辅酶A羧化酶ACC。

[0423] 26.一种产生丙二酸的方法,其包括在导致产生丙二酸的条件下培养根据实施方案1所述的重组宿主细胞的步骤;以及自宿主细胞发酵液分离丙二酸的步骤,其中所述分离通过以下任一实现:

[0424] (a) 添加钙盐至所述发酵液中以使丙二酸钙沉淀;以及自所述发酵液分离沉淀的丙二酸以提供丙二酸,或

[0425] (b) 调整所述发酵液的pH至2.0或以下,添加脂族胺至所述发酵液中,通过相转移将丙二酸自所述发酵液转移至有机溶剂或有机溶剂的组合中,以及移除脂族胺和有机溶剂以提供纯化丙二酸。

[0426] 27.一种产生丙烯酸酯的方法,其包括以下步骤:

[0427] (a) 在导致产生丙二酸的条件下培养根据实施方案1所述的宿主细胞;

- [0428] (b) 自宿主细胞发酵液回收丙二酸;以及
- [0429] (c) 使丙二酸和多聚甲醛在50℃与90℃之间的温度下在含有碱的有机溶剂中反应。
- [0430] 28. 一种丙二酰基-辅酶A水解酶,其选自:
- [0431] (a) SEQ ID NO:46、49、51和52,其中X1选自S、T、N、Q和Y;
- [0432] (b) SEQ ID NO:45和48,其中X1选自R、H、K、S、T、N、Q和Y;
- [0433] (c) SEQ ID NO:47,其中X1是S;
- [0434] (d) SEQ ID NO:50,其中X1选自A、D、K、S、T、N和Y。
- [0435] 29. 一种重组库德里阿兹威毕赤酵母宿主细胞,其包含以下二者:
- [0436] (a) 与SEQ ID NO:9具有至少75%同一性的异源黄褐假单胞菌丙二酰基-辅酶A水解酶;和
- [0437] (b) 选自ScALD 2至6的异源酿酒酵母(Sc)醛脱氢酶(ALD),
- [0438] 其中重组宿主细胞比不包含异源丙二酰基-辅酶A水解酶的对细胞产生更多丙二酸。
- [0439] 30. 根据实施方案29所述的重组库德里阿兹威毕赤酵母宿主细胞,其包含异源酿酒酵母醛脱氢酶ScALD 6。
- [0440] 31. 一种产生丙二酸或其盐的方法,其包括在导致产生丙二酸或其盐的条件下培养根据实施方案1或29所述的宿主细胞。
- [0441] 32. 根据实施方案31所述的方法,其中所述盐是钠盐、铵盐或其他含一价阳离子的盐。
- [0442] 33. 根据实施方案31所述的方法,其进一步包括使所述丙二酸或其盐酯化以提供丙二酸单烷基酯或二烷基酯。
- [0443] 34. 根据实施方案33所述的方法,其中通过反应性提取分离所述丙二酸单烷基酯或二烷基酯。
- [0444] 35. 根据实施方案33所述的方法,其中所述酯化采用酸催化剂。
- [0445] 36. 根据实施方案33所述的方法,其中所述酯化采用酸催化剂,所述酸催化剂是树脂。
- [0446] 37. 根据实施方案33所述的方法,其中所述丙二酸二烷基酯是丙二酸二乙酯或丙二酸二甲酯。

序列表

<110> 利戈斯股份有限公司

<120> 用于产生丙二酸的重组宿主细胞

<130> IIC212348

<160> 53

<170> PatentIn version 3.5

<210> 1

<211> 500

<212> PRT

<213> 酿酒酵母

<400> 1

```

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

```


210	215	220
Tyr Val Ser Ser Glu Asn Leu Asp Ala Leu Gln Lys Arg Leu Gly Glu		
225	230	235
Ile Ser Pro Pro Phe Asn Asn Asp Pro Gln Ser Ala Tyr Phe Phe Gly		240
	245	250
Met Val Asn Glu Ser Ile Asp Glu Phe Val Ser Pro Leu Pro Lys Asp		255
	260	265
Tyr Val Phe Lys Tyr Ser Asn Glu Lys Leu Asn Val Ile Glu Ala Cys		270
	275	280
Phe Asn Leu Ser Lys Asn Gly Thr Ile Glu Asp Ile Met Asn Asn Leu		285
290	295	300
Arg Gln Tyr Glu Gly Ser Ala Glu Gly Lys Ala Phe Ala Gln Glu Ile		
305	310	315
Lys Thr Lys Leu Leu Thr Lys Ser Pro Ser Ser Leu Gln Ile Ala Leu		320
	325	330
Arg Leu Val Gln Glu Asn Ser Arg Asp His Ile Glu Ser Ala Ile Lys		335
	340	345
Arg Asp Leu Tyr Thr Ala Ala Asn Met Cys Met Asn Gln Asp Ser Leu		350
355	360	365
Val Glu Phe Ser Glu Ala Thr Lys His Lys Leu Ile Asp Lys Gln Arg		
370	375	380
Val Pro Tyr Pro Trp Thr Lys Lys Glu Gln Leu Phe Val Ser Gln Leu		
385	390	395
Thr Ser Ile Thr Ser Pro Lys Pro Ser Leu Pro Met Ser Leu Leu Arg		400
	405	410
Asn Thr Ser Asn Val Thr Trp Thr Gln Tyr Pro Tyr His Ser Lys Tyr		415
	420	425
Gln Leu Pro Thr Glu Gln Glu Ile Ala Ala Tyr Ile Glu Lys Arg Thr		430
	435	440
Asn Asp Asp Thr Gly Ala Lys Val Thr Glu Arg Glu Val Leu Asn His		445
450	455	460
Phe Ala Asn Val Ile Pro Ser Arg Arg Gly Lys Leu Gly Ile Gln Ser		
465	470	475
Leu Cys Lys Ile Val Cys Glu Arg Lys Cys Glu Glu Val Asn Asp Gly		480
	485	490
Leu Arg Trp Lys		495
500		
<210> 2		
<211> 1000		

<212> PRT

<213> 流感嗜血杆菌

<400> 2

Ala	Thr	Gly	Thr	Thr	Thr	Thr	Ala	Cys	Ala	Cys	Thr	Gly	Ala	Ala	Ala
1				5					10					15	
Cys	Thr	Thr	Ala	Thr	Gly	Ala	Thr	Gly	Thr	Gly	Ala	Thr	Thr	Gly	Thr
			20					25						30	
Gly	Ala	Thr	Cys	Gly	Gly	Thr	Gly	Gly	Thr	Gly	Gly	Thr	Cys	Ala	Thr
			35				40						45		
Gly	Cys	Gly	Gly	Gly	Thr	Ala	Cys	Ala	Gly	Ala	Ala	Gly	Cys	Cys	Gly
			50				55					60			
Cys	Ala	Cys	Thr	Thr	Gly	Cys	Ala	Cys	Cys	Ala	Gly	Cys	Thr	Cys	Gly
65					70					75					80
Thr	Ala	Thr	Gly	Gly	Gly	Ala	Thr	Thr	Thr	Ala	Ala	Ala	Ala	Cys	Cys
					85					90					95
Cys	Thr	Thr	Thr	Thr	Ala	Thr	Thr	Ala	Ala	Cys	Ala	Cys	Ala	Thr	Ala
					100					105				110	
Ala	Thr	Gly	Thr	Ala	Gly	Ala	Thr	Ala	Cys	Thr	Thr	Thr	Ala	Gly	Gly
					115					120				125	
Gly	Cys	Ala	Ala	Ala	Thr	Gly	Thr	Cys	Thr	Thr	Gly	Thr	Ala	Ala	Cys
					130					135				140	
Cys	Cys	Thr	Gly	Cys	Ala	Ala	Thr	Thr	Gly	Gly	Thr	Gly	Gly	Gly	Ala
145					150					155					160
Thr	Cys	Gly	Gly	Thr	Ala	Ala	Ala	Gly	Gly	Thr	Cys	Ala	Thr	Thr	Thr
					165					170					175
Ala	Gly	Thr	Ala	Ala	Ala	Ala	Gly	Ala	Ala	Gly	Thr	Ala	Gly	Ala	Thr
					180					185				190	
Gly	Cys	Ala	Ala	Thr	Gly	Gly	Gly	Cys	Gly	Gly	Thr	Thr	Thr	Ala	Ala
					195					200				205	
Thr	Gly	Gly	Cys	Gly	Cys	Ala	Thr	Gly	Cys	Thr	Gly	Cys	Ala	Gly	Ala
					210					215				220	
Thr	Ala	Ala	Ala	Gly	Cys	Ala	Gly	Gly	Gly	Ala	Thr	Cys	Cys	Ala	Ala
225					230					235					240
Thr	Thr	Thr	Cys	Gly	Thr	Ala	Cys	Thr	Thr	Thr	Ala	Ala	Ala	Thr	Ala
					245					250					255
Gly	Cys	Ala	Gly	Thr	Ala	Ala	Ala	Gly	Gly	Cys	Cys	Cys	Ala	Gly	Cys
					260					265				270	
Ala	Gly	Thr	Gly	Cys	Gly	Thr	Gly	Cys	Thr	Ala	Cys	Thr	Cys	Gly	Ala
					275					280					285

Gly Cys Thr Cys Ala Ala Gly Cys Thr Gly Ala Cys Ala Gly Ala Gly		
290	295	300
Thr Thr Cys Thr Ala Thr Ala Thr Cys Gly Thr Cys Ala Ala Gly Cys		
305	310	315
Thr Gly Thr Thr Cys Gly Thr Ala Cys Thr Gly Cys Ala Thr Thr Ala		
	325	330
Gly Ala Ala Ala Ala Thr Cys Ala Ala Cys Cys Thr Ala Ala Thr Thr		
	340	345
Thr Ala Gly Ala Thr Ala Thr Thr Thr Thr Cys Cys Ala Ala Cys Ala		
	355	360
Ala Gly Ala Ala Gly Cys Gly Ala Cys Cys Gly Ala Thr Ala Thr Thr		
	370	375
Cys Thr Gly Ala Thr Thr Ala Ala Gly Cys Ala Ala Gly Ala Thr Cys		
385	390	395
Gly Ala Gly Thr Thr Ala Cys Ala Gly Gly Cys Gly Thr Thr Ala Gly		
	405	410
Cys Ala Cys Ala Ala Ala Ala Thr Gly Gly Gly Ala Thr Thr Ala		
	420	425
Ala Cys Thr Thr Thr Thr Cys Gly Thr Gly Cys Thr Ala Ala Ala Thr		
	435	440
Cys Ala Gly Thr Gly Gly Thr Ala Thr Thr Ala Ala Cys Thr Gly Cys		
450	455	460
Gly Gly Gly Thr Ala Cys Thr Thr Thr Cys Thr Thr Ala Gly Cys Thr		
465	470	475
Gly Gly Thr Ala Ala Ala Ala Thr Thr Cys Ala Thr Ala Thr Thr Gly		
	485	490
Gly Thr Thr Thr Gly Gly Ala Ala Ala Thr Thr Ala Thr Gly Ala		
	500	505
Ala Gly Gly Thr Gly Gly Cys Cys Gly Thr Gly Cys Ala Gly Gly Gly		
	515	520
Gly Ala Thr Cys Cys Thr Gly Cys Thr Thr Cys Thr Gly Thr Ala Ala		
530	535	540
Ala Thr Cys Thr Thr Thr Cys Ala Cys Ala Thr Cys Gly Ala Thr Thr		
545	550	555
Ala Ala Gly Ala Gly Ala Thr Cys Thr Cys Gly Gly Ala Thr Thr Ala		
	565	570
Cys Gly Thr Gly Thr Ala Gly Ala Thr Cys Gly Cys Cys Thr Thr Ala		
	580	585
Ala Ala Ala Cys Ala Gly Gly Thr Ala Cys Ala Cys Cys Gly Cys Cys		
		590

595	600	605
Gly Cys Gly Thr Ala Thr Thr	Gly Ala Thr Gly Cys Ala Cys Gly Thr	
610	615	620
Ala Cys Gly Ala Thr Cys Ala Ala Thr Thr Thr Thr Gly Ala Thr Ala		
625	630	635
Thr Thr Thr Thr Ala Gly Cys Thr Ala Ala Ala Cys Ala Ala Cys Ala		640
	645	650
Cys Gly Gly Thr Gly Ala Thr Gly Cys Thr Gly Thr Thr Thr Thr Ala		655
	660	665
Cys Cys Thr Gly Thr Gly Thr Thr Thr Thr Cys Thr Thr Thr Thr Ala		670
675	680	685
Thr Gly Gly Gly Ala Thr Cys Ala Gly Thr Thr Gly Ala Thr Gly Ala		
690	695	700
Thr Cys Ala Cys Cys Cys Thr Cys Ala Ala Cys Ala Ala Ala Thr Thr		
705	710	715
Cys Cys Thr Thr Gly Thr Thr Ala Thr Ala Thr Ala Ala Cys Thr Cys		
	725	730
Ala Thr Ala Cys Cys Ala Ala Thr Gly Ala Ala Cys Ala Ala Ala Cys		735
	740	745
Cys Cys Ala Thr Gly Ala Ala Gly Thr Gly Ala Thr Cys Cys Gly Thr		
755	760	765
Ala Ala Thr Ala Ala Cys Thr Thr Gly Gly Ala Thr Cys Gly Cys Ala		
770	775	780
Gly Thr Cys Cys Ala Ala Thr Gly Thr Ala Thr Ala Cys Thr Gly Gly		
785	790	795
Thr Gly Thr Gly Ala Thr Thr Gly Ala Ala Gly Gly Gly Ala Thr Cys		
	805	810
Gly Gly Thr Cys Cys Ala Cys Gly Thr Thr Ala Thr Thr Gly Cys Cys		815
	820	825
Cys Ala Thr Cys Cys Ala Thr Thr Gly Ala Ala Gly Ala Thr Ala Ala		
835	840	845
Ala Gly Thr Gly Ala Thr Gly Cys Gly Thr Thr Thr Cys Thr Cys Gly		
850	855	860
Gly Ala Thr Cys Gly Thr Ala Ala Thr Thr Cys Ala Cys Ala Thr Cys		
865	870	875
Ala Ala Ala Thr Thr Thr Ala Thr Thr Thr Ala Gly Ala Ala Cys Cys		880
	885	890
Ala Gly Ala Ala Gly Gly Cys Thr Thr Ala Ala Cys Cys Ala Gly Thr		895
900	905	910

Ala Ala Thr Gly Ala Ala Gly Thr Gly Thr Ala Thr Cys Cys Ala Ala
 915 920 925
 Ala Cys Gly Gly Gly Ala Thr Cys Thr Cys Thr Ala Cys Cys Ala Gly
 930 935 940
 Thr Thr Thr Ala Cys Cys Gly Thr Thr Thr Gly Ala Cys Gly Thr Gly
 945 950 955 960
 Cys Ala Ala Ala Thr Gly Gly Gly Cys Ala Thr Thr Gly Thr Gly Ala
 965 970 975
 Ala Thr Thr Cys Thr Ala Thr Gly Ala Ala Ala Gly Gly Thr Thr Thr
 980 985 990
 Ala Gly Ala Ala Ala Ala Cys Gly
 995 1000

<210> 3

<211> 224

<212> PRT

<213> 乙酸钙不动杆菌

<400> 3

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

Phe Thr Glu His Asn Glu Ile Leu Glu Ala Ile Phe Ala Gly Asn Glu
 180 185 190
 Gln Gln Ala Glu Ala Leu Leu Lys Ala His Val Val Ile Gln Gly Gln
 195 200 205
 Lys Phe Thr Asp Phe Ile Ser Thr Ile Glu Ser Leu Gln Pro Lys Ser
 210 215 220
 <210> 4
 <211> 222
 <212> PRT
 <213> 根瘤菌
 <400> 4
 Met Arg Lys Val Lys Arg Met Ser Glu Asn Val Gly Arg Trp Leu Arg
 1 5 10 15
 Asp Glu Ile Glu Asn Ser Ile Leu Ser Asn Glu Phe Ser Pro Gly Glu
 20 25 30
 Arg Leu Asp Glu Thr Val Leu Ala Thr Arg Phe Gly Val Ser Arg Thr
 35 40 45
 Pro Val Arg Glu Ala Leu Met Gln Leu Asp Ala Ile Gly Leu Ile Glu
 50 55 60
 Ile Arg Pro Arg Arg Gly Ala Ile Val Ile Asp Pro Gly Pro His Arg
 65 70 75 80
 Val Tyr Glu Met Phe Glu Val Met Ala Glu Leu Glu Gly Leu Ala Gly
 85 90 95
 Ser Leu Ala Ala Arg Arg Leu Asp Lys Thr Ser Arg Glu Ala Ile Thr
 100 105 110
 Ala Thr His Gly Arg Cys Glu Lys Ser Ala Ala Ala Gly Asp Ser Asp
 115 120 125
 Ala Tyr Tyr Tyr Asp Asn Glu Glu Phe His Lys Ala Ile Tyr Ala Ala
 130 135 140
 Gly Arg Ser Asp Phe Leu Glu Glu Gln Cys Leu Gln Leu His Arg Arg
 145 150 155 160
 Leu Arg Pro Asp Arg Arg Leu Gln Leu Arg Val Arg Asn Arg Leu Ser
 165 170 175
 Thr Ser Phe Leu Glu His Cys Ala Ile Val Asp Ala Ile Phe Ala Gly
 180 185 190
 Asp Gly Asp Glu Ala Arg Arg Leu Leu Arg Gly His Val Gly Ile Gln
 195 200 205
 Gly Glu Arg Phe Ser Asp Leu Val Ala Ser Met Ala Ala Arg
 210 215 220

<210> 5

<211> 308

<212> PRT

<213> 肺炎克雷伯氏菌

<400> 5

```

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

```


Gln Met Arg Gln Leu Ile Ser Ile Val Tyr Ser His His Arg Glu Arg
275 280 285
Asp Ala Asp Leu Leu Ala Leu Ala Ala Glu Gly Arg Met Tyr Ala Arg
290 295 300
Ser Ile Asn Arg
305

<210> 6

<211> 77

<212> DNA

<213> 乙酸钙不动杆菌

<400> 6

aaaaaaattg tatacaattt atgtttattt gagtacaaag cattgtacac tgaatacaga 60
taggctataa ctatacc 77

<210> 7

<211> 598

<212> PRT

<213> EHD3

<220>

<221> misc_feature

<222> (1) ... (16)

<223> Xaa的定义见说明书和原始序列表文件

<220>

<221> misc_feature

<222> (18) ... (24)

<223> Xaa的定义见说明书和原始序列表文件

<220>

<221> misc_feature

<222> (27) ... (41)

<223> Xaa的定义见说明书和原始序列表文件

<220>

<221> misc_feature

<222> (44) ... (90)

<223> Xaa的定义见说明书和原始序列表文件

<220>

<221> misc_feature

<222> (93) ... (93)

<223> Xaa的定义见说明书和原始序列表文件

<220>

<221> misc_feature

<222> (95) ... (100)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (102) ... (102)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (104) ... (105)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (107) ... (108)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (111) ... (113)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (118) ... (118)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (125) ... (126)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (128) ... (131)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (133) ... (134)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (139) ... (139)
<223> Xaa的定义见说明书和原始序列表文件
<220>

<221> misc_feature
<222> (142) ... (144)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (146) ... (150)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (152) ... (152)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (157) ... (158)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (166) ... (167)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (169) ... (170)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (172) ... (175)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (177) ... (183)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (186) ... (188)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (191) ... (191)
<223> Xaa的定义见说明书和原始序列表文件

<220>
<221> misc_feature
<222> (195) ... (195)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (199) ... (199)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (202) ... (205)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (207) ... (207)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (217) ... (219)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (221) ... (221)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (225) ... (225)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (231) ... (231)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (247) ... (248)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (251) ... (251)

<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (253) ... (257)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (260) ... (262)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (264) ... (264)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (266) ... (266)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (270) ... (270)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (273) ... (276)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (278) ... (279)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (282) ... (283)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (285) ... (285)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature

<222> (290) ... (293)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (295) ... (297)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (299) ... (300)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (305) ... (306)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (308) ... (318)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (320) ... (322)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (325) ... (327)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (329) ... (332)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (335) ... (336)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (338) ... (339)
<223> Xaa的定义见说明书和原始序列表文件
<220>

<221> misc_feature
<222> (341) ... (342)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (344) ... (344)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (346) ... (346)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (348) ... (348)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (350) ... (351)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (353) ... (353)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (356) ... (357)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (360) ... (373)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (375) ... (381)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (383) ... (383)
<223> Xaa的定义见说明书和原始序列表文件

<220>
<221> misc_feature
<222> (385) ... (387)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (389) ... (390)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (393) ... (398)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (400) ... (401)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (405) ... (405)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (407) ... (407)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (409) ... (409)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (411) ... (416)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (418) ... (424)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (426) ... (428)

<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (431) ... (431)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (434) ... (434)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (438) ... (438)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (440) ... (443)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (445) ... (448)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (451) ... (452)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (454) ... (456)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (459) ... (460)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (463) ... (464)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature

<222> (469) ... (478)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (480) ... (485)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (487) ... (487)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (489) ... (490)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (492) ... (493)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (495) ... (496)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (498) ... (499)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (501) ... (501)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (504) ... (505)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (508) ... (511)
<223> Xaa的定义见说明书和原始序列表文件
<220>

<221> misc_feature
<222> (516) ... (521)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (524) ... (528)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (530) ... (530)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (532) ... (535)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (537) ... (550)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (552) ... (552)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (554) ... (554)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (556) ... (560)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (562) ... (562)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (564) ... (575)
<223> Xaa的定义见说明书和原始序列表文件

<220>
 <221> misc_feature
 <222> (577) ... (580)
 <223> Xaa的定义见说明书和原始序列表文件
 <220>
 <221> misc_feature
 <222> (582) ... (582)
 <223> Xaa的定义见说明书和原始序列表文件
 <220>
 <221> misc_feature
 <222> (585) ... (586)
 <223> Xaa的定义见说明书和原始序列表文件
 <220>
 <221> misc_feature
 <222> (588) ... (598)
 <223> Xaa的定义见说明书和原始序列表文件
 <400> 7

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

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Phe	Phe	Xaa	Xaa	Xaa	Tyr	Ser	Xaa	Asn	
			180						185						190	
Phe	Gln	Xaa	Ala	Thr	Tyr	Xaa	Lys	Pro	Xaa	Xaa	Xaa	Xaa	Met	Xaa	Gly	
			195						200						205	
Ile	Thr	Met	Gly	Gly	Gly	Val	Gly	Xaa	Xaa	Xaa	His	Xaa	Pro	Phe	Arg	
			210						215						220	
Xaa	Ala	Thr	Glu	Asn	Thr	Xaa	Trp	Ala	Met	Pro	Glu	Met	Asp	Ile	Gly	
225						230						235			240	
Phe	Phe	Pro	Asp	Val	Gly	Xaa	Xaa	Phe	Ala	Xaa	Pro	Xaa	Xaa	Xaa	Xaa	
			245						250						255	
Xaa	Ala	Asn	Xaa	Xaa	Xaa	Gln	Xaa	Ala	Xaa	Tyr	Leu	Cys	Xaa	Thr	Gly	
			260						265						270	
Xaa	Xaa	Xaa	Xaa	Gly	Xaa	Xaa	Ala	Tyr	Xaa	Xaa	Gly	Xaa	Ala	Ser	His	
			275						280						285	
Tyr	Xaa	Xaa	Xaa	Xaa	Asn	Xaa	Xaa	Xaa	Leu	Xaa	Xaa	Arg	Leu	Gly	Glu	
290						295						300				
Xaa	Xaa	Pro	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Gln	Xaa	
305						310						315			320	
Xaa	Xaa	Phe	Phe	Xaa	Xaa	Xaa	Asn	Xaa	Xaa	Xaa	Xaa	Glu	Phe	Xaa	Xaa	
			325						330						335	
Pro	Xaa	Xaa	Pro	Xaa	Xaa	Tyr	Xaa	Phe	Xaa	Tyr	Xaa	Asn	Xaa	Xaa	Leu	
			340						345						350	
Xaa	Val	Ile	Xaa	Xaa	Cys	Phe	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	
			355						360						365	
Xaa	Xaa	Xaa	Xaa	Xaa	Leu	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Tyr	Xaa	Gly	
370						375						380				
Xaa	Xaa	Xaa	Ala	Xaa	Xaa	Phe	Ala	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Leu	Xaa	
385						390						395			400	
Xaa	Lys	Ser	Pro	Xaa	Ser	Xaa	Gln	Xaa	Ala	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	
			405						410						415	
Asn	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Ala	Xaa	Xaa	Xaa	Asp	Leu	Xaa	Thr	
			420						425						430	
Ala	Xaa	Asn	Met	Cys	Xaa	Asn	Xaa	Xaa	Xaa	Xaa	Gln	Xaa	Xaa	Xaa	Xaa	
			435						440						445	
Glu	Phe	Xaa	Xaa	Ala	Xaa	Xaa	Xaa	Lys	Leu	Xaa	Xaa	Lys	Gln	Xaa	Xaa	
450						455						460				
Pro	Tyr	Pro	Trp	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Gln	Xaa	
465						470						475			480	
Xaa	Xaa	Xaa	Xaa	Xaa	Pro	Xaa	Pro	Xaa	Xaa	Pro	Xaa	Xaa	Leu	Xaa	Xaa	

485	490	495
Asn Xaa Xaa Asn Xaa Thr Trp Xaa Xaa Tyr Pro Xaa Xaa Xaa Xaa Tyr		
500	505	510
Gln Leu Pro Xaa Xaa Xaa Xaa Xaa Xaa Gln Tyr Xaa Xaa Xaa Xaa Xaa		
515	520	525
Asn Xaa Asn Xaa Xaa Xaa Xaa Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa		
530	535	540
Xaa Xaa Xaa Xaa Xaa Xaa Phe Xaa Asn Xaa Asn Xaa Xaa Xaa Xaa Xaa		
545	550	555
Lys Xaa Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys		
565	570	575
Xaa Xaa Xaa Xaa Ala Xaa Gly Gly Xaa Xaa Trp Xaa Xaa Xaa Xaa Xaa		
580	585	590
Xaa Xaa Xaa Xaa Xaa Xaa		
595		

<210> 8

<211> 351

<212> PRT

<213> 芽孢杆菌属

<220>

<221> misc_feature

<222> (4) ... (4)

<223> Xaa的定义见说明书和原始序列表文件

<220>

<221> misc_feature

<222> (9) ... (11)

<223> Xaa的定义见说明书和原始序列表文件

<220>

<221> misc_feature

<222> (16) ... (16)

<223> Xaa的定义见说明书和原始序列表文件

<220>

<221> misc_feature

<222> (31) ... (31)

<223> Xaa的定义见说明书和原始序列表文件

<220>

<221> misc_feature

<222> (45) ... (48)

<223> Xaa的定义见说明书和原始序列表文件

<220>
<221> misc_feature
<222> (51) ... (51)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (59) ... (59)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (79) ... (79)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (83) ... (83)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (86) ... (86)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (89) ... (89)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (91) ... (91)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (93) ... (93)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (98) ... (98)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (102) ... (102)

<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (125) ... (125)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (131) ... (132)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (156) ... (156)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (161) ... (161)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (163) ... (163)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (168) ... (168)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (170) ... (171)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (175) ... (175)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (188) ... (191)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature

<222> (194) ... (194)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (197) ... (202)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (205) ... (209)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (211) ... (213)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (216) ... (220)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (223) ... (230)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (232) ... (235)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (237) ... (238)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (240) ... (240)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (245) ... (247)
<223> Xaa的定义见说明书和原始序列表文件
<220>

<221> misc_feature
<222> (249) ... (249)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (252) ... (252)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (256) ... (257)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (259) ... (260)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (263) ... (268)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (275) ... (275)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (285) ... (286)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (288) ... (289)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (292) ... (292)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (294) ... (294)
<223> Xaa的定义见说明书和原始序列表文件

<220>
 <221> misc_feature
 <222> (301) ... (301)
 <223> Xaa的定义见说明书和原始序列表文件
 <220>
 <221> misc_feature
 <222> (310) ... (310)
 <223> Xaa的定义见说明书和原始序列表文件
 <220>
 <221> misc_feature
 <222> (319) ... (319)
 <223> Xaa的定义见说明书和原始序列表文件
 <220>
 <221> misc_feature
 <222> (321) ... (321)
 <223> Xaa的定义见说明书和原始序列表文件
 <220>
 <221> misc_feature
 <222> (332) ... (332)
 <223> Xaa的定义见说明书和原始序列表文件
 <220>
 <221> misc_feature
 <222> (334) ... (335)
 <223> Xaa的定义见说明书和原始序列表文件
 <220>
 <221> misc_feature
 <222> (338) ... (341)
 <223> Xaa的定义见说明书和原始序列表文件
 <220>
 <221> misc_feature
 <222> (344) ... (344)
 <223> Xaa的定义见说明书和原始序列表文件
 <400> 8

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

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

<210> 9

<211> 370
<212> PRT
<213> 假单胞菌属
<220>
<221> misc_feature
<222> (3) ... (4)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (7) ... (12)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (18) ... (18)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (20) ... (20)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (24) ... (26)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (31) ... (31)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (36) ... (37)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (40) ... (43)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (47) ... (48)
<223> Xaa的定义见说明书和原始序列表文件

<220>
<221> misc_feature
<222> (51) ... (52)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (55) ... (55)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (69) ... (69)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (71) ... (72)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (74) ... (75)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (78) ... (80)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (83) ... (84)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (90) ... (90)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (94) ... (95)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (99) ... (102)

<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (104) ... (104)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (109) ... (110)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (116) ... (116)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (118) ... (118)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (125) ... (125)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (129) ... (130)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (136) ... (137)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (144) ... (145)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (159) ... (159)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature

<222> (161) ... (161)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (164) ... (164)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (169) ... (169)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (172) ... (172)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (175) ... (176)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (179) ... (179)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (182) ... (182)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (185) ... (185)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (191) ... (194)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (197) ... (198)
<223> Xaa的定义见说明书和原始序列表文件
<220>

<221> misc_feature
<222> (200) ... (202)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (204) ... (204)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (206) ... (206)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (212) ... (214)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (217) ... (217)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (220) ... (221)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (223) ... (226)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (228) ... (231)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (233) ... (233)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (236) ... (236)
<223> Xaa的定义见说明书和原始序列表文件

<220>
<221> misc_feature
<222> (242) ... (243)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (245) ... (246)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (249) ... (249)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (251) ... (252)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (254) ... (254)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (257) ... (257)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (259) ... (260)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (262) ... (266)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (270) ... (270)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (272) ... (272)

<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (275) ... (279)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (283) ... (284)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (286) ... (287)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (290) ... (290)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (292) ... (293)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (297) ... (297)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (302) ... (303)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (305) ... (305)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (309) ... (310)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature

<222> (319) ... (319)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (321) ... (321)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (329) ... (329)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (331) ... (331)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (333) ... (333)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (336) ... (336)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (338) ... (338)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (341) ... (342)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (345) ... (345)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (347) ... (347)
<223> Xaa的定义见说明书和原始序列表文件
<220>

<221> misc_feature
 <222> (349) ... (351)
 <223> Xaa的定义见说明书和原始序列表文件
 <220>
 <221> misc_feature
 <222> (355) ... (357)
 <223> Xaa的定义见说明书和原始序列表文件
 <220>
 <221> misc_feature
 <222> (359) ... (359)
 <223> Xaa的定义见说明书和原始序列表文件
 <220>
 <221> misc_feature
 <222> (361) ... (362)
 <223> Xaa的定义见说明书和原始序列表文件
 <220>
 <221> misc_feature
 <222> (364) ... (364)
 <223> Xaa的定义见说明书和原始序列表文件
 <220>
 <221> misc_feature
 <222> (366) ... (370)
 <223> Xaa的定义见说明书和原始序列表文件
 <400> 9

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

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

Xaa Xaa

370

<210> 10

<211> 505

<212> PRT

<213> 一般性细菌

<220>

<221> misc_feature

<222> (1) ... (30)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (32) ... (32)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (34) ... (35)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (39) ... (40)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (42) ... (43)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (48) ... (48)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (52) ... (52)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (57) ... (57)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (63) ... (63)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (67) ... (67)
<223> Xaa的定义见说明书和原始序列表文件
<220>

<221> misc_feature
<222> (70) ... (70)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (73) ... (75)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (80) ... (80)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (83) ... (83)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (86) ... (87)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (89) ... (90)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (95) ... (95)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (102) ... (106)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (108) ... (108)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (116) ... (116)
<223> Xaa的定义见说明书和原始序列表文件

<220>
<221> misc_feature
<222> (127) ... (128)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (131) ... (132)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (135) ... (135)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (142) ... (143)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (150) ... (150)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (164) ... (166)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (170) ... (171)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (193) ... (193)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (195) ... (200)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (208) ... (208)

<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (211) ... (211)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (215) ... (215)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (221) ... (221)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (223) ... (223)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (227) ... (227)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (233) ... (233)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (239) ... (239)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (242) ... (242)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (245) ... (245)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature

<222> (252) ... (253)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (261) ... (263)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (265) ... (265)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (275) ... (275)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (278) ... (278)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (281) ... (284)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (289) ... (289)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (296) ... (298)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (301) ... (301)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (306) ... (306)
<223> Xaa的定义见说明书和原始序列表文件
<220>

<221> misc_feature
<222> (308) ... (308)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (311) ... (311)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (313) ... (314)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (316) ... (316)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (328) ... (328)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (332) ... (332)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (334) ... (335)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (337) ... (338)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (350) ... (351)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (353) ... (353)
<223> Xaa的定义见说明书和原始序列表文件

<220>
<221> misc_feature
<222> (358) ... (358)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (361) ... (361)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (367) ... (372)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (374) ... (374)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (379) ... (380)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (383) ... (384)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (386) ... (387)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (393) ... (393)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (397) ... (397)
<223> Xaa的定义见说明书和原始序列表文件
<220>
<221> misc_feature
<222> (400) ... (400)

<223> Xaa的定义见说明书和原始序列表文件

<220>

<221> misc_feature

<222> (402) ... (402)

<223> Xaa的定义见说明书和原始序列表文件

<220>

<221> misc_feature

<222> (405) ... (420)

<223> Xaa的定义见说明书和原始序列表文件

<220>

<221> misc_feature

<222> (427) ... (427)

<223> Xaa的定义见说明书和原始序列表文件

<220>

<221> misc_feature

<222> (431) ... (431)

<223> Xaa的定义见说明书和原始序列表文件

<220>

<221> misc_feature

<222> (433) ... (434)

<223> Xaa的定义见说明书和原始序列表文件

<220>

<221> misc_feature

<222> (440) ... (505)

<223> Xaa的定义见说明书和原始序列表文件

<400> 10

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

1 5 10 15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Met Xaa

20 25 30

Met Xaa Xaa Thr Glu His Xaa Xaa Phe Xaa Xaa Ser Glu Asn Gly Xaa

35 40 45

Ala Ser Ile Xaa Leu Asn Arg Pro Xaa Ala Leu Asn Ser Leu Xaa Tyr

50 55 60

Asp Met Xaa Gln Pro Xaa Gly Gln Xaa Xaa Xaa Glu Trp Glu Asn Xaa

65 70 75 80

Glu Arg Xaa Ala Leu Xaa Xaa Leu Xaa Xaa Gly Ala Gly Thr Xaa Gly

85 90 95

Phe Cys Ala Gly Gly Xaa Xaa Xaa Xaa Xaa Tyr Xaa Ala Arg Ser Asn

			100				105				110				
Glu	Pro	Gly	Xaa	Ala	Leu	Gln	His	Ala	Glu	Arg	Phe	Phe	Glu	Xaa	Xaa
			115				120				125				
Tyr	Glu	Xaa	Xaa	Thr	Tyr	Xaa	Tyr	Gln	Tyr	Lys	Lys	Pro	Xaa	Xaa	Ala
			130				135				140				
Cys	Leu	Asp	Gly	Ile	Xaa	Met	Gly	Gly	Gly	Val	Gly	Leu	Thr	Asn	Gly
145			150				155				160				
Ala	Lys	Tyr	Xaa	Xaa	Xaa	Thr	Glu	Arg	Xaa	Xaa	Trp	Ala	Met	Pro	Glu
			165				170				175				
Met	Asn	Ile	Gly	Phe	Phe	Pro	Asp	Val	Gly	Ala	Ala	Tyr	Phe	Leu	Asn
			180				185				190				
Xaa	Ala	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Pro	Gly	Tyr	Leu	Gly	Arg	Tyr	Xaa
			195				200				205				
Ala	Leu	Xaa	Ala	Ser	Ile	Xaa	Lys	Ala	Ser	Asp	Val	Xaa	Phe	Xaa	Asn
210			215				220								
Ala	Ala	Xaa	Tyr	Phe	Met	Thr	Ser	Xaa	Ser	Leu	Pro	Ala	Phe	Xaa	Thr
225			230				235				240				
Glu	Xaa	Glu	Ser	Xaa	Asn	Trp	His	Lys	Glu	Asp	Xaa	Xaa	His	Thr	His
			245				250				255				
Leu	Leu	Lys	Glu	Xaa	Xaa	Xaa	Val	Xaa	Arg	Thr	Phe	Ala	Thr	Ala	Pro
			260				265				270				
Asn	Leu	Xaa	Ser	Glu	Xaa	Ala	Pro	Xaa	Xaa	Xaa	Xaa	Ser	Leu	Glu	Glu
			275				280				285				
Xaa	Asn	Ser	His	Phe	Ala	Phe	Xaa	Xaa	Xaa	Asp	Thr	Xaa	Glu	Glu	Ile
290			295				300								
Trp	Xaa	Ala	Xaa	His	Ser	Xaa	Glu	Xaa	Xaa	Lys	Xaa	Gln	Ser	Ser	Phe
305			310				315				320				
Ala	Leu	Lys	Thr	Lys	Glu	Thr	Xaa	Leu	Ser	Lys	Xaa	Pro	Xaa	Xaa	Leu
			325				330				335				
Xaa	Xaa	Thr	Leu	Lys	Gln	Phe	Ile	Asp	Gly	Arg	Asp	Lys	Xaa	Xaa	Glu
			340				345				350				
Xaa	Cys	Phe	Ala	Thr	Xaa	Leu	Val	Xaa	Ala	Lys	Asn	Phe	Met	Xaa	Xaa
			355				360				365				
Xaa	Xaa	Xaa	Xaa	Glu	Xaa	Phe	Phe	Glu	Gly	Xaa	Xaa	Ser	Val	Xaa	Xaa
370			375				380								
Asp	Xaa	Xaa	Gln	Asn	Pro	Asn	Tyr	Xaa	Tyr	Lys	Gln	Xaa	Ser	Asp	Xaa
385			390				395				400				
Ser	Xaa	Glu	Asp	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
			405				410				415				

Xaa Xaa Xaa Xaa Asn Arg Phe Phe Asn Leu Xaa Asn Ala Gly Xaa His		
420	425	430
Xaa Xaa Pro Leu Ala Asp Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa		
435	440	445
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa		
450	455	460
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa		
465	470	475
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa		
485	490	495
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa		
500	505	

<210> 11

<211> 6801

<212> DNA

<213> 人工序列

<400> 11

```

atgagattgc agttgagaac cttgactaga aggtttttct ctatggcttc tggttcttct 60
actccagatg ttgtccatt ggttgatcca aatattcata agggtttggc ctctcacttt 120
ttcggtttga actctgttca tactgctaag ccatctaagg tcaaagaatt tgttgcttct 180
catggtggtc ataccgttat caacaaagt ttgattgcca acaacggtat tgctgccgta 240
aaagaaatta ggtctgttag aaaatgggcc tacgaaactt ttggtgacga aagagctatt 300
tccttactg ttatggctac acctgaagat ttggctgcta atgctgatta cattagaatg 360
gccgatcaat acgttgaagt tccaggtggg acaacaaca acaattacgc taacgttgaa 420
ttgatcggtg atgtcgctga aagatttggg gttgatgctg tttgggctgg ttggggctcat 480
gcttctgaaa atccattatt gccagaatct ttggctgctt ctccaagaaa gatcgttttt 540
attggtccac caggtgctgc tatgagatca ttgggtgata agatttcttc taccatcggt 600
gctcaacatg ctaaggttcc atgtattcca tggcttggtg ctggtgttga tgaagtgtt 660
gttgataagt ccaccaactt ggtttccgtt tctgaagaag ttacactaa gggttgtact 720
actggtccaa aacaagggtt ggaaaaggct aagcaattg gtttcccagt tatgattaag 780
gcttctgaag gtggtggtgg taaaggtatt agaaagggtg aaagggaaga agatttcgaa 840
gctgcttacc atcaagttga aggtgaaatt ccaggttccc caattttcat tatgcaattg 900
gctggtaacg ccagacattt ggaagttcaa ttattggctg atcagtacgg caacaacatt 960
tctttgtttg gtagagattg ctccgttcaa agaaggcatc aaaagattat tgaagaagcc 1020
ccagttactg ttgctggtca acaaactttt actgctatgg aaaaagctgc cgtcagattg 1080
ggtaaattgg ttggttatgt ttctgccggt actgtcgaat acttgactc tcatgaagat 1140
gacaagttct acttcttga attgaacca agattgcaag ttgaacatcc aactactgaa 1200
atggttaccg gtgttaattt gccagctgct caattgcaaa ttgctatggg tattccattg 1260
gacagaatca aggatattag gttgttctac ggtgttaacc cacataccac tactccaatt 1320

```

gatttcgatt tctctggtga agatgctgac aaaactcaaa gaaggccagt tccaagaggt 1380
catacaactg cttgtagaat tacttctgaa gatccaggtg aaggttttta accatctggt 1440
ggtactatgc acgaattgaa ctttaggtcc tcttctaattg tttggggtta cttctctggt 1500
ggtaatcaag gtggtatcca ctctttctct gattctcaat tcggtcatat tttcgccctt 1560
ggtgaaaaca gatccgcctc aagaaaacat atggttggtg ccttgaaaga attgtccatc 1620
agaggtgatt tcagaaccac tgttgaatac ttgatcaaat tattggaaac cccagacttc 1680
gaggataaca ctattactac tggttggttg gacgagttga tctctaacaa attgactgct 1740
gaaagaccag attccttctt ggctggtgtt tgtggtgctg ctacaaaagc tcatagagct 1800
tcagaagatt ctatcgctac ttacatggct tctttggaaa aagggtcaagt tccagccaga 1860
gacattttga aaactttgtt cccagttgac ttcactacg aagggtcaaag atacaagttt 1920
accgctacca gatcatccga agattcttac actttgttca tcaacggttc cagatgcgat 1980
attggtgtta gaccattgtc tgatggtggt attttgtgtt tgggttggtg tagatcccat 2040
aacgtttatt ggaaagaaga agttggcgct accagattgt ctgttgattc taaaacttgc 2100
ttgttggaag tcgaaaacga tccaactcaa ttgagatcac catctccagg taagttggtt 2160
aagttcttgg ttgaaaacgg tgatcacgtt agagctaate aaccatacgc tgaaatcgaa 2220
gtcatgaaga tgtacatgac tttgaccgct caagaagatg gtatcgttca attgatgaag 2280
caaccaggtt ctactattga agccggtgat attttgggta ttttggcttt ggatgatcca 2340
tccaaggtca aacatgctaa accatttgaa ggtcagttgc cagaattggg tccaccaaca 2400
ttgtctggta acaaaccaca tcaaagatac gaacattgcc aaaacgtctt gcacaacatt 2460
ttgttgggtt tcgataacca ggttgtcatg aagtctacat tgcaagaaat ggtcggtttg 2520
ttgagaaatc cagaattgcc atacttgcaa tgggctcatc aagtttcttc attgcataca 2580
agaatgtccg ctaagttgga tgctactttg gctggtttga ttgataaggc taaacaaagg 2640
ggtggtgaat ttccagctaa gcaattattg agagccttgg aaaaagaagc ttcacttggc 2700
gaagttgatg ctttgtttca acaaacattg gcccctttgt ttgatttggc tagagaatat 2760
caagatgggt tggccatcca tgaattgcaa gttgctgctg gtttggtgca agcttattat 2820
gattctgaag ctagattctg cgggtccaaac gttagagatg aagatgttat cttgaagttg 2880
aggggaagaaa acagggactc tttgagaaaa gttgttatgg cccaattgtc ccattcaaga 2940
gttggtgcta aaaacaactt ggttttggct ttgttggtg agtacaaggc tgctgatcaa 3000
gctggtactg attctccagc ttctaattgt catgttgcta aatacttgag gccagtcttg 3060
agaaaagatt tcgaattgga atcaagagct tccgctaagg tttctttgaa ggctagagaa 3120
attttgatcc aatgcgcttt gccatccttg aaagaaagaa ctgatcaatt ggaacacatc 3180
ttgagatcct ctgttggtga atcaagatac ggtgaagttg gtttggaaca tagaactcca 3240
agagctgaca tcttgaaaga agttgttgac tccaagtaca tcgtgttcga tgttttggct 3300
caattcttcg ctcatgatga tccatggata gttttggctg ctttgagatt gtatattaga 3360
agggttgta aggctactc cattttggtt attactacc accaagactc tgatttgcca 3420
ccagttattt cttggagatt cagattgcca actatgtcat ctgccttgta caactctgtt 3480
gtttcttctg gttctaagac tccaacttct ccatctgttt caagagctga ttctgtttcc 3540
gatttctctt acaccgttga aagagattct gctccagcta gaactggtgc tatagttgct 3600
gttccacatt tggatgattt ggaagatgct ttgaccagag tcttgaaaaa tttgcaaaaa 3660

agagggtgctg gtttggctat ttctgttgggt gcttctaaca aatcagctgc tgcttctgct 3720
 agagatgctg ctgctgcagc tgcttcttct gttgatactg gtttgtctaa catctgcaac 3780
 gttatgatcg gtagagttga tgaatccgat gatgatgata ccttgatcgc cagaatttcc 3840
 caagttatcg aagatttcaa agaggacttc gaagcttgct ccttgagaag aattactttc 3900
 tcattcggtg actccagagg tacttaccca aagtacttta cttttagagg tccagcctat 3960
 gaagaagatc caaccattag acatattgaa ccagctttgg cttttcaatt ggaattggct 4020
 agattgtcta acttcgacat caagccagtt cataccgata acagaaacat ccatgtttac 4080
 gaagctactg gtaagaatgc tgcttccgat aagagatttt tcaccagagg tatagttaga 4140
 ccaggtagat tgagagaaaa catcccaaca tccgagtact tgatttctga agctgataga 4200
 ttgatgtccg atatttttga tgcttggaa gttattggta ctaccaactc tgatttgaac 4260
 cacatcttca ttaacttctc cgtgtttttt gctttgaagc cagaagaagt tgaagctgct 4320
 tttggtggtt ttttggaaaag attcggtaga agattgtgga gattgagagt tactgggtgcc 4380
 gaaattagaa tgatggtttc tgatccagaa actggttctg cttttccatt gagagctatg 4440
 atcaacaacg tttccggtta cgttgtccaa tctgaattat acgctgaagc caagaatgat 4500
 aagggtcaat ggatctttaa gtccttgggt aaaccaggtt caatgcatat gagatccatt 4560
 aacactccat accctaccaaa agaattggtt caacctaaaa gatacaaggc ccatttgatg 4620
 ggtactacct actgttatga tttcccagag ttgttcagac agtccattga atctgattgg 4680
 aaaaagtacg atggttaaggc tccagatgat ttgatgactt gcaacgaatt gatcttggac 4740
 gaagattctg gtgaattgca agaagttaat agagaaccag gtgctaacaa cgttgggtatg 4800
 gttgcttggg aatttgaagc taagactcca gaatatccaa ggggtagaag ttttatcggt 4860
 gttgccaacg atattacctt ccagattggt tcttttggtc cagctgaaga tcaattcttc 4920
 ttcaaggtta ctgaattggc cagaaagttg ggtattccaa gaatctactt gtctgctaata 4980
 tccggtgcta gaattgggtat tgctgatgaa ttggtcggta agtataaagt tgcttggaa 5040
 gacgaaactg atccatctaa gggtttcaag tacttgtact tcactccaga atcattggct 5100
 accttgaaac cagatactgt tgttaccacc gaaattgaag aagaaggtcc aaacggcggt 5160
 gaaaagagac atgttattga ttacatcgtc ggtgaaaagg atggtttggg tgttgaatgt 5220
 ttgagagggt ctggtttaat tgctgggtgct acttcaagag cttacaagga tattttcacc 5280
 ttgaccttgg ttacctgtag atcagttggt attggtgctt acttggttag attgggtcaa 5340
 agagccattc aaattgaagg tcagccaatt atcttgactg gtgctccagc tattaacaag 5400
 ttgttgggta gagaagtcta ctctctaac ttgcaattgg gtggtactca aatcatgtac 5460
 aacaacggtg tttctcattt gaccgctaga gatgatttga acggtgttca taagatcatg 5520
 cagtggttgt cttatatctc agcttcaaga ggtttgccag ttccagtttt gccacataag 5580
 actgatgttt gggatagaga tgttacctc caaccagtta gaggtgaaca atatgatgtc 5640
 agatggttga tttctggtag gactttggaa gatggtgctt ttgaatctgg tttgttcgat 5700
 aaggactctt tccaagaaac tttatctggt tgggctaagg gtgttgttgt tggtagagct 5760
 agattgggtg gtattccatt tgggtgttatt ggtgttgaaa ctgccactgt tgataacact 5820
 actccagctg atccagctaa tccagattct attgaaatgt ctacttccga agctgggtcaa 5880
 gtttgggtatc caaattctgc tttcaagacc tccaagcca ttaacgattt taatcatggt 5940
 gaagccttgc cattgatgat tttggctaata tggagaggtt tttccggtgg tcaaagagat 6000

atgtacaacg aagttttgaa gtacggctcc tttattgtcg atgctttggt tgattacaag 6060
 cagccaatta tggtttacat tccaccaact ggtgaattga gaggtggttc ttgggttggt 6120
 gttgacccaa ctattaactc cgatatgatg gaaatgtacg ccgatgttga aagtagaggt 6180
 ggtgttttgg aaccagaagg tatggttggt attaagtaca gaagagacaa gttgttagat 6240
 accatggcca gattagatcc agagtactct tccttgaaaa aacaattgga agagtcccca 6300
 gactccgaag aattgaaaagt taagttgtcc gtcagggaaa agtctttgat gccaatctac 6360
 caacaaatct ccgttcaatt tgctgacttg catgatagag ctggtagaat ggaagctaaa 6420
 ggtgttatta gagaagcctt ggtttggaag gatgctagaa gatttttctt ttggaggatc 6480
 agaagaaggt tggtcgagga atatttgatc accaagatca actccatttt gccatcttgt 6540
 accagattgg aatgtttggc tagaatcaaa tcttggaagc cagctacttt ggatcaaggt 6600
 tctgatagag gtgttgctga atggtttgac gaaaattctg atgctgtttc tgccaggttg 6660
 tctgaattga aaaaagatgc ttctgctcag tccttcgctt cccaattgag aaaagataga 6720
 caaggtacat tgcagggtat gaagcaagct ttggcttctt tgtctgaagc tgaaagagct 6780
 gaattattga agggcttgta a 6801

<210> 12

<211> 7891

<212> DNA

<213> 人工序列

<400> 12

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccc gagacggtca 60
 cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
 ttggcgggtg tcggggctgg cttactatg cggcatacaga gcagattgta ctgagagtgc 180
 accacgcctt tcaattcaat tcatcatttt ttttttattc ttttttttga tttcggtttc 240
 tttgaaattt ttttgattcg gtaatctccg aacagaagga agaacgaagg aaggagcaca 300
 gacttagatt ggtatatata cgcatatgta gtgttgaaaga aacatgaaat tgcccagtat 360
 tcttaaccca actgcacaga acaaaaacct gcaggaaacg aagataaatc atgtcgaaag 420
 ctacatatata ggaacgtgct gctactcatc ctagtcctgt tgctgccaag ctattttaata 480
 tcatgcacga aaagcaaaca aacttggtg cttcattgga tgttcgtacc accaaggaat 540
 tactggagtt agttgaagca ttaggtccca aaatttggtt actaaaaaca catgtggata 600
 tcttgactga tttttccatg gagggcacag ttaagccgct aaaggcatta tccgccaagt 660
 acaatttttt actcttcgaa gacagaaaat ttgctgacat tggtaatata gtcaaattgc 720
 agtactctgc ggggtgtatac agaatagcag aatgggcaga cattacgaat gcacacggtg 780
 tgggtgggcc aggtattggt agcggtttga agcaggcggc agaagaagta acaaaggaac 840
 ctagaggcct tttgatgta gcagaattgt catgcaaggg ctccctatct actggagaat 900
 atactaaggg tactgttgac attgcgaaga gcgacaaaga ttttgttatc ggctttattg 960
 ctcaaagaga catgggtgga agagatgaag gttacgattg gttgattatg acaccgggtg 1020
 tgggttttaga tgacaaggga gacgcattgg gtcaacagta tagaaccgtg gatgatgtgg 1080
 tctctacagg atctgacatt attattgttg gaagaggact atttgcaaag ggaagggatg 1140
 ctaaggtaga ggggtgaacgt tacagaaaag caggctggga agcatatttg agaagatgcg 1200

gccagcaaaa ctaaaaaact gtattataag taaatgcatg tataactaaac tcacaaatta 1260
gagcttcaat ttaattatat cagttattac cctgcggtgt gaaataccgc acagatgcgt 1320
aaggagaaaa taccgcatca ggaaattgta aacgttaata ttttgttaaa attcgcgtta 1380
aatTTTTgtt aaatcagctc atTTTTtaac caataggccg aaatcggcaa aatcccttat 1440
aaatcaaaaag aatagaccga gatagggttg agtggtgttc cagtttgtaa caagagtcca 1500
ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc 1560
ccactacgtg aaccatcacc ctaatcaagt tttttggggt cgaggtgccg taaagcacta 1620
aatcggaacc ctaaaggag ccccgattt agagcttgac ggggaaagcc gggaacgtg 1680
gcgagaaaagg aagggaagaa agcgaaagga gcggcgctta gggcgctggc aagtgtagcg 1740
gtcacgctgc gcgtaaccac cacaccgcc gcgcttaatg cgccgctaca gggcgctcg 1800
cgccattcgc cattcaggt gcgcaactgt tgggaagggc gatcggtcgc ggctcttctg 1860
ctattacgcc agctggcgaa ggggggatgt gctgcaaggc gattaagttg ggtaacgcca 1920
gggttttccc agtcacgacg ttgtaaaacg acggccagt aattgtaata cgactcacta 1980
tagggcgaat tggagctcca ccgcggtggc ggccgcatag gccactagt gatctgatat 2040
catcgatgaa ttcgagctcg ttttcgacac tggatggcgg cgtagtatc gaatcgacag 2100
cagtatagcg accagcattc acatacgatt gacgcatgat attactttct gcgcacttaa 2160
cttcgcatct gggcagatga tgtcgaggcg aaaaaaata taaatcacgc taacatttga 2220
ttaaataaga acaactacaa tataaaaaaa ctatacaaat gacaagttct tgaaaacaag 2280
aatcttttta ttgtcagtac tgattattcc tttgcctcg gacgagtgt ggggcgtcgg 2340
tttccactat cggcgagtac ttctacacag ccatcgttcc agacggccgc gcttctgcgg 2400
gcgatttgtg tacgcccgc agtcccggt ccgcatcgga cgattgcgtc gcacgaccc 2460
tgcgcccaag ctgcatcatc gaaattgccg tcaaccaagc tctgatagag ttggtcaaga 2520
ccaatgcgga gcatatacgc ccggagccgc ggcatcctg caagctccg atgcctccgc 2580
tcgaagtagc gcgtctgtg ctccatacaa gccaacacg gcctccagaa gaagatgttg 2640
gcgacctcgt attgggaatc cccgaacatc gcctcgctcc agtcaatgac cgctgttatg 2700
cggccattgt ccgtcaggac attgttgag ccgaaatccg cgtgcacgag gtgccggact 2760
tcggggcagt cctcggccca aagcatcagc tcatcgagag cctgcgcgac ggacgcactg 2820
acggtgtcgt ccatcacagt ttgccagtga tacacatggg gatcagcaat cgcgcatatg 2880
aaatcacgcc atgtagtgta ttgaccgatt cttgcggtc cgaatgggcc gaaccgctc 2940
gtctggctaa gatcggccgc agcgatcgca tccatggcct ccgcgaccg ctgcagaaca 3000
gcgggcagtt cggtttcagg caggtcttgc aacgtgacac cctgtgcacg gcgggagatg 3060
caataggtca ggctctcgt gaattccca atgtcaagca cttccggaat cgggagcgcg 3120
gccgatgcaa agtgccgata aacataacga tctttgtaga aaccatcggc gcagctatTT 3180
accgcagga catatccacg cctcctaca tcgaagctga aagcacgaga ttcttcgccc 3240
tccgagagct gcatcaggtc ggagacgtg tcgaactttt cgatcagaaa cttctcgaca 3300
gacgtcgcgg tgagttcagg ctttttacc atggttgttt atgttcggat gtgatgtgag 3360
aactgtatcc tagcaagatt ttaaaaggaa gtatatgaaa gaagaacctc agtggcaaat 3420
cctaaccttt tataatttct tacagggcg cggcgtgggg acaattcaac gcgtctgtga 3480
ggggagcgtt tccctgctcg caggtctgca gcgaggagcc gtaatttttg cttcgcgccg 3540

```

tgcggccatc aaaatgtatg gatgcaaattg attatacatg gggatgtatg ggctaaatgt 3600
acgggcgaca gtcacatcat gcccctgagc tgcgcacgtc aagactgtca aggagggtat 3660
tctgggcctc catgtcgtg gccgggtgac ccggcgggga cgaggcaagc taaacagatc 3720
tcccgggtag tttaaacaat gaatgtcacc tttgaagaaa gagccagttt acacggttac 3780
agaatcggta tcgcatcctt agacgcccc gcacacctga acgctttgtc cttaccaatg 3840
attgatgctt tacaagacag attgagagct tgggcagaag atgcagacat agcctgtgtc 3900
ttgttacgtg gtaatggttc taaagcattt tgcgcaggtg gtgacgttgt ccaattagct 3960
aagaaatgtt tggcctcccc aggtgaagcc cctgaattgg ctgaaagatt tttcgctaga 4020
agttacagat tagatcatta tttgcacaca tacccaaagc ctttgatatg ctgggctcat 4080
ggtcacgttt taggtggtgg tatgggtttg ttacaaggtg ctggtattag aatagtcacc 4140
ccatcttcaa gattggcaat gcctgaaatc tctattggtt tatttcaga tgttggtggt 4200
tcccatttct taagtagatt gcctggtaaa ttgggtttat ttttcggtt aaccgcttca 4260
ccattgaatg ccagagatgc tttggactta aacttggtg atagattctt gttagatact 4320
caacaagacg cattgatcga tggtttgatc caattgaact ggagagaaca accagatttg 4380
caattgcatt ccttgtaaaa ggcttttagaa caacaagcaa gaagtgaatt accagctgca 4440
caatggttgc ctagaagaga aagattagac gccttggttag atcaagctac tttaccattg 4500
tcttggaag ccttagcttc attggaaaac gatgaagacg cattgttagc caaagccgct 4560
aagactatgt tgggtggttc acctttaaca ggtcatttgg tctggggtca aattagaaga 4620
gctagacact tatctttggc acaagtattt caaatggaat acggcatgtc attgaattgt 4680
tgcagacatc cagaatttgc agaaggtgta agagccagat taatcgataa agaccacgct 4740
ccacattggc actggcctga cgtaaacc aa gttcctgaag ccgttattgc agccccattt 4800
gcaccattag atgaccaccc tttagctgat ttggcataag tttaaactca tgtaattagt 4860
tatgtcacgc ttacattcac gccctcccc cacatccgct ctaaccgaaa aggaaggagt 4920
tagacaacct gaagtctagg tccctattta tttttttata gttatgttag tattaagaac 4980
gttattttata tttcaaattt ttctttttt tctgtacaga cgcgtgtacg catgtaacat 5040
tatactgaaa accttgcttg agaaggtttt gggacgctcg aaggcttta tttgcggcct 5100
ctagcagctt ttgttccctt tagtgagggt taatttcgag cttggcgtaa tcatggtcat 5160
agctgtttcc tgtgtgaaat tggtatccgc tcacaattcc acacaacata ggagccggaa 5220
gcataaagt taaagcctgg ggtgcctaat gtagtgagga actcacatta attgcgttgc 5280
gctcactgcc cgctttccag tcgggaaacc tgcgtgcc a gctgcattaa tgaatcggcc 5340
aacgcgcggg gagaggcgg ttcggtattg ggcgtcttc cgcttcctcg ctactgact 5400
cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc tctactcaaag gcggtataac 5460
ggttatccac agaattcagg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa 5520
aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc ggccccctg 5580
acgagcatca caaaaaatcga cgtcaagtc agaggtggcg aaaccgcaca ggactataaa 5640
gataaccagg gtccccccct ggaagctccc tcgtgcgctc tctgtttccg acctgcccgc 5700
ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct caatgctcac 5760
gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac 5820
ccccgttca gcccgaccgc tgcgccttat ccggttaact tcgtcttgag tccaaccg 5880

```


taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt 5940
atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagga 6000
cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct 6060
cttgatccgg caaacaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga 6120
ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg gggcttgacg 6180
ctcagtggaa cgaaaactca cgtaaggga ttttggtcat gagattatca aaaaggatct 6240
tcacctagat ctttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt 6300
aaacttggtc tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc 6360
tatttcgttc atccatagtt gcctgactgc ccgtcgtgta gataactacg atacgggagg 6420
gcttaccatc tggccccagt gctgcaatga taccgcgaga cccacgctca ccggctccag 6480
atztatcagc aataaaccag ccagccggaa gggccgagcg cagaagtggc cctgcaactt 6540
tatccgcctc catccagtct attaatgtt gccgggaagc tagagtaagt agttcgccag 6600
ttaatagttt gcgcaacgtt gttgccattg ctacaggcat cgtggtgtca cgctcgtcgt 6660
ttggtatggc ttcatcagc tccggttccc aacgatcaag gcgagttaca tgatccccc 6720
tgttgtgaaa aaaagcgggt agtccttcg gtctccgat cgttgtcaga agtaagttgg 6780
ccgcagtgtt atcactcatg gttatggcag cactgcataa ttctcttact gtcatgcat 6840
ccgtaagatg cttttctgtg actggtgagt actcaacaa gtcattctga gaatagtgt 6900
tgccgagacc gagttgctct tgcccggcgt caatacggga taataccgcg ccacatagca 6960
gaactttaaa agtgctcatc attggaaaac gttcttcggg gcgaaaactc tcaaggatct 7020
taccgctgtt gagatccagt tcgatgtaac ccactcgtgc acccaactga tcttcagcat 7080
cttttacttt caccagcgtt tctgggtgag caaaaacagg aaggcaaat gccgcaaaaa 7140
agggaataag ggcgacacgg aaatgttgaa tactcatact cttccttttt caatattatt 7200
gaagcattta tcagggttat tgtctcatga gcgatacat atttgaatgt atttagaaaa 7260
ataaacaat aggggttccg cgcacatttc cccgaaaagt gccacctggg tccttttcat 7320
cacgtgctat aaaaataatt ataatttaa tttttaata taaatatata aattaaaaat 7380
agaaagtaaa aaaagaaatt aaagaaaaa tagttttgt tttccgaaga tgtaaaagac 7440
tctaggggga tcgccaacaa atactacctt ttatcttgc cttcctgctc tcaggtatta 7500
atgccgaatt gtttcatctt gtctgtgtag aagaccacac acgaaaatcc tgtgatttta 7560
cattttactt atcgtaatac gaatgtatat ctatttaate tgcttttctt gtctaataaa 7620
tatatatgta aagtacgctt tttgttgaaa tttttaaac ctttgtttat tttttttct 7680
tcattccgta actcttctac cttctttatt tactttctaa aatccaaata caaacataa 7740
aaataaataa acacagagta aattcccaaa ttattccatc attaaaagat acgaggcgcg 7800
tgtaagttac aggcaagcga tccgtcctaa gaaaccatta ttatcatgac attaacctat 7860
aaaaataggc gtatcacgag gccctttcgt c 7891

<210> 13

<211> 60

<212> DNA

<213> 人工序列

<400> 13

ccaatatata ataaaaatatg gaggaatgcg atgctcagaa atacgctaaa atgtgcccaa 60

<210> 14

<211> 60

<212> DNA

<213> 人工序列

<400> 14

tgccctggaga tccttactcg agttggatcc ttatttccat ctttaagccat cgттаacttc 60

<210> 15

<211> 37

<212> DNA

<213> 人工序列

<400> 15

ttttactgat gcgtattctt tgaattttca aatagca 37

<210> 16

<211> 34

<212> DNA

<213> 人工序列

<400> 16

tcaaagaata cgcattcagta aaaaatttga tgga 34

<210> 17

<211> 41

<212> DNA

<213> 人工序列

<400> 17

ttttactgat gtttattctt tgaattttca aatagcaact t 41

<210> 18

<211> 39

<212> DNA

<213> 人工序列

<400> 18

tcaaagaata aacattcagta aaaaatttga tggacttgg 39

<210> 19

<211> 39

<212> DNA

<213> 人工序列

<400> 19

ttttactgat tcgtattctt tgaattttca aatagcaac 39

<210> 20

<211> 36

<212> DNA
<213> 人工序列
<400> 20
tcaaagaata cgaatcagta aaaaatttga tggact 36
<210> 21
<211> 55
<212> DNA
<213> 人工序列
<400> 21
ccaatatata ataaaaatatg gaggaatgcg atgtctacaa cacataacgt ccctc 55
<210> 22
<211> 53
<212> DNA
<213> 人工序列
<400> 22
tgcctggaga tccttactcg agttggatcc ttactcaaca ggtaaggcgc gag 53
<210> 23
<211> 61
<212> DNA
<213> 人工序列
<400> 23
ctaattacat gactcgaggt cgacggtatc gttatttcca tcttaagcca tcgttaactt 60
c 61
<210> 24
<211> 67
<212> DNA
<213> 人工序列
<400> 24
cattagaaaag aaagcatagc aatctaattct aagtttataaa caatgactac tcaacccag 60
ctaaatg 67
<210> 25
<211> 60
<212> DNA
<213> 人工序列
<400> 25
gaaagcatag caatctaatc taagtttataa acaatgaccg aacaagtctt attctcagta 60
<210> 26
<211> 59
<212> DNA

<213> 人工序列

<400> 26

ctaattacat gactcgaggt cgacggtatc gttaagcgtt caacaaattg aaaaatctg 59

<210> 27

<211> 58

<212> DNA

<213> 人工序列

<400> 27

gaaagcatag caatctaate taagtttaaa acaatgaccg aacatgtatt attctcag 58

<210> 28

<211> 57

<212> DNA

<213> 人工序列

<400> 28

ctaattacat gactcgaggt cgacggtatc gttaagcgtt taacaaattg aaaaatc 57

<210> 29

<211> 57

<212> DNA

<213> 人工序列

<400> 29

gaaagcatag caatctaate taagtttaaa acaatgagaa gatacatcag aggtggt 57

<210> 30

<211> 57

<212> DNA

<213> 人工序列

<400> 30

ctaattacat gactcgaggt cgacggtatc gttatgcagc gttcaacaaa ttgaaaa 57

<210> 31

<211> 58

<212> DNA

<213> 人工序列

<400> 31

gaaagcatag caatctaate taagtttaaa acaatgaccg aacaagtctt attctcag 58

<210> 32

<211> 57

<212> DNA

<213> 人工序列

<400> 32

ctaattacat gactcgaggt cgacggtatc gttaagcgtt caacaaattg aaaaatc 57

<210> 33
<211> 60
<212> DNA
<213> 人工序列
<400> 33
gaaagcatag caatctaate taagtttaaa acaatgaact tacaatttga agaaagacca 60
<210> 34
<211> 57
<212> DNA
<213> 人工序列
<400> 34
ctaattacat gactcgaggt cgacggtatc gttacaaatc agctaaaggg tgttcac 57
<210> 35
<211> 59
<212> DNA
<213> 人工序列
<400> 35
gaaagcatag caatctaate taagtttaaa acaatgaact tacactttga agaattgac 59
<210> 36
<211> 57
<212> DNA
<213> 人工序列
<400> 36
ctaattacat gactcgaggt cgacggtatc gttagtagtc agacaaatct gctaaag 57
<210> 37
<211> 59
<212> DNA
<213> 人工序列
<400> 37
gaaagcatag caatctaate taagtttaaa acaatgacaa tccactgtga agtattaac 59
<210> 38
<211> 55
<212> DNA
<213> 人工序列
<400> 38
ctaattacat gactcgaggt cgacggtatc gttaccaaac gtcagccaaa ggggtg 55
<210> 39
<211> 58
<212> DNA

<213> 人工序列

<400> 39

gaaagcatag caatctaatac taagtttaaa acaatgaatg tcacctttga agaaagag 58

<210> 40

<211> 55

<212> DNA

<213> 人工序列

<400> 40

ctaattacat gactcgaggt cgacggatc gttatgcaa atcagctaaa ggggtg 55

<210> 41

<211> 4431

<212> DNA

<213> 人工序列

<400> 41

cgaaacccta tgctctgttg ttcggatttg aaatttttaa actacattaa tgtgttagtt 60
 tttctttctt tctttctttg tcttgacgtg atttgactt ctgtcttgca ttcgcgtcca 120
 ttcatctgac ccaatatcc ttttggtttt gttatcctta taaaaagaaa ggaagcttct 180
 tagagggaaa aaaatgatga agagtaatgc caaaatataa ataaataaat aaatatgaaa 240
 atcattttct atttttaata gaataagaag agcatcttaa gattacaatt tcaagaaata 300
 gtttacacag tatatccaat aactccaata aactactttc ctatacaaat ttctatgggtg 360
 ggattaatag taaaacttct gtacttctct aattcaccaa gaaattaagg taaacatctg 420
 gtaagcacta tccagctttt tgctattaca catatggctt ttctgcaatc atttcttccc 480
 attttgtctc aagccgtag tcttgaaacc acaggcggag tagagttact tgatgcggta 540
 ttttacatgc cttttttcac tgcaaaaaaa atgaaatata tatttacacg atttgcagga 600
 cagtttacga tagtgagtat gcagaatagt taacacctt gttttatcct tttgtgtctt 660
 aattatatga tataaaggcg cctggcaa atccgtttta agagcttggg gagcgctagg 720
 agtcaactgcc aggtatcggt tgaacacggc attagtcagg gaagtcataa cacagtcctt 780
 tcccgcatt ttctttttct attactcttg gcctcctcta gtacactcta ttttttttta 840
 tgcctcggtg atgattttca tttttttttt tcccctagcg gatgactctt tttttttctt 900
 agcgattggc attatcacat aatgaattat acattatata aagtaatgtg atttcttcga 960
 agaatatact aaaaaatgag caggcaagat aaacgaaggc aaagatgaca gagcagaaag 1020
 ccctagtaaa gcgtattaca aatgaaacca agattcagat tgcgatctct ttaaagggtg 1080
 gtcccctagc gatagagcac tcgatcttcc cagaaaaaga ggcagaagca gtagcagaac 1140
 aggccacaca atcgcaagtg attaacgtcc acacaggtat agggtttctg gaccatatga 1200
 tacatgctct ggccaagcat tccggctggg cgctaactgt tgagtgcatt ggtgacttac 1260
 acatagacga ccatcacacc actgaagact gcgggattgc tctcggtcaa gcttttaag 1320
 aggccctact ggcgctgga gtaaaaagggt ttggatcagg atttgccct ttggatgagg 1380
 cactttccag agcggtggtg gatctttcga acaggccgta cgcagttgtc gaacttggtt 1440
 tgcaaaggga gaaagtagga gatctctctt gcgagatgat cccgcatttt cttgaaagct 1500

```

ttgcagaggc tagcagaatt accctccacg ttgattgtct gcgaggcaag aatgatcatc 1560
accgtagtga gagtgcgttc aaggctcttg cggttgccat aagagaagcc acctcgccca 1620
atggtaccaa cgatgttccc tccaccaaag gtgttcttat gtagtgacac cgattattta 1680
aagctgcagc atacgatata tatacatgtg tatatatgta tacctatgaa tgtcagtaag 1740
tatgtatacg aacagtatga tactgaagat gacaaggtaa tgcattcattc tatacgtgtc 1800
attctgaacg aggcgcgctt tccttttttc tttttgcttt ttcttttttt ttctcttgaa 1860
ctcgacggat ctatgcggtg tgaaataaccg cacagatgcg taaggagaaa atacgcgcatc 1920
aggaaattgt aaacgttaat attttgtaa aattcgcgtt aaatttttgt taaatcagct 1980
catttttttaa ccaataggcc gatagcttca aaatgtttct actccttttt tactcttcca 2040
gattttctcg gactccgcgc atcgccgtac cacttcaaaa cacccaagca cagcactacta 2100
aatttccctt ctttcttctt ctagggtgtc gttaattacc cgtactaaag gtttgaaaaa 2160
gaaaaaagag accgcctcgt ttctttttct tcgtcgaaaa aggcaataaa aattttttatc 2220
acgtttcttt ttcttgaaaa tttttttttt gatttttttc tctttcgatg acctcccatt 2280
gatatttaag ttaataaacg gtcttcaatt tctcaagttt cagtttcatt tttcttgttc 2340
tattacaact ttttttactt cttgtctcatt agaaagaaag catagcaatc taatctaagt 2400
ttaaaacaat gaatgtcacc tttgaagaaa gagccagttt acacggttac agaatcggtg 2460
tcgcatcctt agacgccccg gcacccctga acgctttgtc cttaccaatg attgatgctt 2520
tacaagacag attgagagct tgggcagaag atgcagacat agcctgtgtc ttgttacgtg 2580
gtaatgggtc taaagcattt tgcgaggtg gtgacgttgt ccaattagct aagaaatgtt 2640
tggcctcccc aggtgaagcc cctgaattgg ctgaaagatt tttcgctaga agttacagat 2700
tagatcatta ttgacacaca tacccaaagc ctttgatatg ctgggctcat ggtcacgttt 2760
taggtgggtg tatgggtttg ttacaagggt ctggtattag aatagtcacc ccatcttcaa 2820
gattggcaat gcctgaaatc tctattgggt tatttccaga tggttggtgg tccattttct 2880
taagtagatt gcctggtaaa ttgggtttat ttttcggttt aaccgcttca ccattgaatg 2940
ccagagatgc tttggactta aacttggttg atagattcct gttagatact caacaagacg 3000
cattgatcga tggtttgatc caattgaact ggagagaaca accagatttg caattgcatt 3060
ccttggttaa ggcttttaga caacaagcaa gaagtgaatt accagctgca caatggttgc 3120
ctagaagaga aagattagac gccttgtag atcaagctac tttaccattg tcttggaag 3180
ccttagcttc atttgaaaac gatgaagacg cattgttagc caaagccgct aagactatgt 3240
tgggtggttc acctttaaca ggtcatttgg tctggggtca aattagaaga gctagacact 3300
tatctttggc acaagtattt caaatggaat acggcatgtc attgaattgt tgcagacatc 3360
cagaatttgc agaaggtgta agagccagat taatcgataa agaccacgct ccacattggc 3420
actggcctga cgtaaaccaa gttcctgaag ccgttattgc agccatttt gcaccattag 3480
atgaccaccc tttagctgat ttggcataac gataccgtcg acctcgagtc atgtaattag 3540
ttatgtcacg cttacattca cgccctcccc ccacatccgc tctaaccgaa aaggaaggag 3600
ttagacaacc tgaagtctag gtccctattt atttttttat agttatgtta gtattaagaa 3660
cgttattttat atttcaaatt tttctttttt ttctgtacag acgctgttac gcatgtaaca 3720
ttatactgaa aaccttgctt gagaaggttt tgggacgctc gaaggcttta atttgcggcc 3780
ataaagcagc cgctacaaaa cagacaagat tcagtatgta aggtaaatac ctttttgcac 3840

```

agttaaacta cccaaactta ttaaagcttg ataaattact gaaattccac ctttcagtta 3900
 gattcaggcc tcatatagat tagatatagg gtacgtaaca ttctgtcaac caagttgttg 3960
 gaatgaaagt ctaaaatgtc atctattcgg tagcactcat gttactagta tactgtcaca 4020
 tgcggtgtaa cgtggggaca taaaacagac atcaaata atggaagctg aaatgcaaag 4080
 atcgataatg taataggaat gaaacataata aaacgaaagg agaagtaatg gtaatattag 4140
 tatgtagaaa taccgattca attttgggga ttcttatatt ctcgagagaa tttctagtat 4200
 aatctgtata cataatatta taggctttac caacaatgga atttcgacaa ttatcatatt 4260
 attcaccaat taatcacaag ttggtaatga gtttgataac aagttacttt cttaacaacg 4320
 ttagtatcgt caaaacactc ggttttactc gagctttag cacaataata ccgtgtagag 4380
 ttctgtattg ttcttcttag tgcttgata tgctcatccc gaccttccat t 4431

<210> 42

<211> 5356

<212> DNA

<213> 人工序列

<400> 42

accggagctt ggatatgata aacgaaatat tcttgaatcg tgagatcgcc tgttttcaaa 60
 accgttggag gcagaaacaa ttttgtcaca agatgggcat tctaccccat cttgtctgta 120
 ttattgtagt ctcgctttct tttatgctgg acaaatgaga ctactgcaca tttttatacg 180
 ttcttggttt tttttaaaagg tgtggtttcg gcattatcct gccgcacgtt tcttgataa 240
 ttcacctga ttctctattt taaacgcttc agcctatcag gatttggttt tgatacatc 300
 tgcaagagtg tatctcgga acagtcattt attccgcaac aaacttaatt gcggaacgcg 360
 ttaggcgatt tctagcatat atcaaatacc gtccgcgatt tcttctgggt tcgtctcttt 420
 tcttttaaat acttattaac gtactcaaac aactacactt cgttgtatct cagaatgaga 480
 tccctcagta tgacaataca tcattctaaa cgttcgtaaa acacatatga aacaacttta 540
 taacaaagcg aacaaaatgg gcaacatgag atgaaactcc gcgtccctta gctgaactac 600
 ccaaacgtac gaatgcctga acaattagtt tagatccgag attccgcgct tccatcattt 660
 agtataatcc atattttata taatatatag gataagtaac agcccgcgaa tcgaggagaa 720
 cttctagtat atccacatac ctaatattat tgccttatta aaaatggaat cccaacaatt 780
 acatcaaaaat ccacattctc ttcaaaaatca attgtcctgt acttccttgt tcatgtgtgt 840
 tcaaaaacgt tatatttata ggataattat actctatttc tcaacaagta attggttggt 900
 tggccgagcg gtctaaggcg cctgattcaa gaaatatctt gaccgcagtt aactgtggga 960
 atactcaggt atcgtaagat gcaagagttc gaatctctta gcaaccatta ttttttctct 1020
 caacataacg agaacacaca ggggcgctat cgcacagaat caaattcgat gactggaaat 1080
 tttttgttaa tttcagaggt cgctgacgc atatacttt ttcaactgaa aaattgggag 1140
 aaaaaggaaa ggtgagaggc cggaaccggc ttttcatata gaatagagaa gcgttcatga 1200
 ctaaatgctt gcatcacaat acttgaagtt gacaatatta ttttaaggacc tattgttttt 1260
 tccaataggt ggtagcaat cgtcttactt tctaactttt cttacctttt acatttcagc 1320
 aatatatata tatatttcaa ggatatacca ttctaagtgc tgcccctatg tctgccctta 1380
 agaagatcgt cgttttgcca ggtgaccacg ttggtcaaga aatcacagcc gaagccatta 1440


```

aggttcttaa agctatttct gatgttcggt ccaatgtcaa gttcgatttc gaaaatcatt 1500
taattggtgg tgctgctatc gatgctacag gtgtcccact tccagatgag gcgctggaag 1560
cctccaagaa gggtgatgcc gttttgttag gtgctgtggc tggtcctaaa tggggtaccg 1620
gtagtgttag acctgaacaa ggtttactaa aaatccgtaa agaacttcaa ttgtacgcca 1680
acttaagacc atgtaacttt gcatccgact ctctttttaga cttatctcca atcaagccac 1740
aatttgctaa aggtactgac ttcgttggtg tcagagaatt agtgggaggt atttactttg 1800
gtaagagaaa ggaagacgat ggtgatgggt tcgcttgga tagtgaacaa tacaccgttc 1860
cagaagtgc aagaatcaca agaatggccg ctttcatggc cctacaacat gagccaccat 1920
tgcctatttg gtccttggt aaagctaate ttttggcctc ttcaagatta tggagaaaaa 1980
ctgtggagga aaccatcaag aacgaattcc ctacattgaa ggttcaacat caattgattg 2040
attctgccgc catgatccta gttaagaacc caaccacct aaatggtatt ataatcacca 2100
gcaacatgtt tggatgatc atctccgatg aagcctccgt tatcccaggt tccttgggtt 2160
tggtgccatc tgcgtccttg gcctctttgc cagacaagaa caccgcattt ggtttgtacg 2220
aaccatgcca cggttctgct ccagatttgc caaagaataa ggttgaccct atcgccacta 2280
tcttgctgc tgcaatgatg ttgaaattgt cattgaactt gcctgaagaa ggtaaggcca 2340
ttgaagatgc agttaaaaag gttttgatg caggtatcag aactggtgat ttaggtggtt 2400
ccaacagtac caccgaagtc ggtgatgctg tcgccgaaga agttaagaaa atccttgctt 2460
aaaaagattc tcttttttta tgatatttgt acataaactt tataaatgaa attcataata 2520
gaaacgacac gaaattacaa aatggaatat gttcataggg tagacgaaac tatatacgca 2580
atctacatac atttatcaag aaggagaaaa aggaggatag taaaggaata caggtaagca 2640
aattgatact aatggctcaa cgtgataagg aaaaagaatt gcactttaac attaataattg 2700
acaaggagga gggcaccaca caaaaagtta ggtgtaacag aaaatcatga aactacgatt 2760
cctaatttga tattggagga ttttctctaa aaaaaaaaaa atacaacaaa taaaaaacac 2820
tcaatgacct gaccatttga tggagttaa gtcaatacct tcttgaagca tttcccataa 2880
tggtgaaagt tccctcaaga attttactct gtcagaaacg gcctatagct tcaaaatggt 2940
tctactcctt ttttactcct ccagatttcc tcggactccg cgcacgcccg taccacttca 3000
aaacacccaa gcacagcata cttaaatttcc cctcttttct cctctagggt gtcgttaatt 3060
acccgtacta aaggtttga aaagaaaaaa gagaccgcct cgtttctttt tcttcgtcga 3120
aaaaggcaat aaaaattttt atcacgttcc ttttcttga aaattttttt tttgattttt 3180
ttctctttcg atgacctccc attgatattt aagttaataa acggtcttca atttctcaag 3240
tttcagtttc atttttcttg ttctattaca acttttttta cttcttgctc attagaaaga 3300
aagcatagca atctaatac agtttaaaac aatgaatgtc accttgaag aaagagccag 3360
tttacacggc tacagaatcg gtatcgcatc cttagacgcc ccagcatcct tgaacgcttt 3420
gtccttacca atgattgatg ctttacaaga cagattgaga gcttgggcag aagatgcaga 3480
catagcctgt gtcttggtac gtggtaatgg ttctaaagca ttttgccgag gtggtgacgt 3540
tgtccaatta gctaagaaat gtttggcctc cccaggtgaa gcccctgaat tggctgaaag 3600
atttttcgct agaagttaca gattagatca ttatttgcac acatacccaa agcctttgat 3660
atgctgggct catggtcacg ttttaggtgg tggtatgggt ttgttacaag gtgctggtat 3720
tagaatagtc accccatctt caagattggc aatgcctgaa atctctattg gtttatttcc 3780

```

agatgttggg ggttcccatt tcttaagtag attgcctggg aaattggggtt tatttttcgg 3840
 ttttaaccgct tcaccattga atgccagaga tgctttggac ttaaacttgg ctgatagatt 3900
 cttgttagat actcaacaag acgcattgat cgatggtttg atccaattga actggagaga 3960
 acaaccagat ttgcaattgc attccttggt aaaggcttta gaacaacaag caagaagtga 4020
 attaccagct gcacaatggg tgcctagaag agaaagatta gacgccttgt tagatcaagc 4080
 tactttacca ttgtcttggc aagccttagc ttcatggaa aacgatgaag acgcattggt 4140
 agccaaagcc gctaagacta tggtgggtgg ttcaccttta acaggtcatt tggctctgggg 4200
 tcaaataga agagctagac acttatcttt ggcacaagta tttcaaatgg aatacggcat 4260
 gtcattgaat tggtgcagac atccagaatt tgcagaagg gtaagagcca gattaatcga 4320
 taaagaccac gctccacatt ggcactggcc tgacgtaaac caagttcctg aagccgttat 4380
 tgcagcccat tttgcacat tagatgacca ccttttagct gatttggcat aacgataccg 4440
 tcgacctga gtcatgtaat tagttatgtc acgttacat tcacgcctc cccccacatc 4500
 cgctctaacc gaaaaggaag gagttagaca acctgaagtc taggtcccta tttatttttt 4560
 tatagttatg ttagtattaa gaacgttatt tatatttcaa atttttcttt tttttctgta 4620
 cagacgcgtg tacgcatgta acattatact gaaaacctg cttgagaagg ttttgggacg 4680
 ctggaaggct ttaatttgcg gccggaccaa ctatcatccg ctaattactg acattaccaa 4740
 atgagatctg tgaatgggca agataaaaaa caaaaattga aatgtttgac gttatgtaaa 4800
 actattaatt ccttcgcttt cggcgggtcac agaatttgcg tgtagctgac tcttgttcaa 4860
 tcaatatcat ttgttacttt atttgaaagt ctgtattact gcgcctattg tcatccgtac 4920
 caaagaacgt caaaaagaaa caagataatt tttgtgctta caccatttat agatcactga 4980
 gccagaata tcgctggagc tcagtgtgaa tggcatgaac acaactctga ctgatcgac 5040
 atattgccgt tatcataaat actagttgta cttgtcaatg cgacgaatgg catcatgcct 5100
 attattacgt tcctcttttt ccgtttcatg tttccagaat gctattgaat ctaacacttc 5160
 aattataaaa aagaataaat ccgcaataat tttaggctaa ttgttgact gtcaagcgaa 5220
 cctaattgggt aaaattcaga ggaaccttcg acgtagtctg atcgctactt ctatatctta 5280
 tgttcccagt caatcaaaag ttgatactat aatagctgcc atttatacct gttagttatg 5340
 gcgatcgttt atcacg 5356

<210> 43

<211> 3492

<212> DNA

<213> 人工序列

<400> 43

aactcaatcg tgctaagcga caaactgaaa aggtatatca accagagatg aaatcggaaa 60
 ttacggggga tagcactctg aaacgtgttc aagtcacctt ccacgggaat ggtcgggtaa 120
 ataaaaagaa aaagaagggt cctccacgga aagatggtat caagttccgt tgattgaaaa 180
 tttttctttt caagcgatga ggtagcgaa gttgttcgaa caagcccaa tatgtttaga 240
 aggagtagc aattgatacc ctgaactgtt actacatacc acagcttttc aattcaattc 300
 atcatTTTTT ttttattctt ttttttgatt tcggtttctt tgaaattttt ttgattcggg 360
 aatctccgaa cagaaggaag aacgaaggaa ggagcacaga cttagattgg tatatatacg 420

catatgtagt gttgaagaaa catgaaattg cccagtattc ttaacccaac tgcacagaac 480
 aaaaacctgc aggaaacgaa gataaatcat gtcgaaagct acatataagg aacgtgctgc 540
 tactcatcct agtcctgttg ctgccaagct atttaatatc atgcacgaaa agcaaacaaa 600
 cttgtgtgct tcattggatg ttcgtaccac caaggaatta ctggagttag ttgaagcatt 660
 aggtcccaaaa atttgtttac taaaaacaca tgtggatatac ttgactgatt tttccatgga 720
 gggcacagtt aagccgctaa aggcatatc cgccaagtac aatTTTTTtac tcttcgaaga 780
 cagaaaaattt gctgacattg gtaatacagt caaattgcag tactctgcgg gtgtatacac 840
 aatagcagaa tgggcagaca ttacgaatgc acacggtgtg gtgggcccag gtattgttag 900
 cggtttgaag caggcggcag aagaagtaac aaaggaacct agaggccttt tgatgttagc 960
 agaattgtca tgcaagggct ccctatctac tggagaatat actaagggtg ctgttgacat 1020
 tgcgaagagc gacaaaagatt ttgttatcgg ctttattgct caaagagaca tgggtggaag 1080
 agatgaaggt tacgattggt tgattatgac acccggtgtg ggttttagatg acaagggaga 1140
 cgcattgggt caacagtata gaaccgtgga tgatgtggtc tctacaggat ctgacattat 1200
 tattgttgga agaggactat ttgcaaaggg aagggatgct aaggtagagg gtgaacgtta 1260
 cagaaaagca ggctgggaag catatttgag aagatgcggc cagcaaaact aaaaaactgt 1320
 attataagta aatgcatgta tactaaactc acaaattaga gcttcaattt aattatatca 1380
 gttattaccc tatgcggtgt gaaataccgc acagatgcgt aaggagaaaa taccgcatca 1440
 ggatagcttc aaaatgtttc tactcctttt ttactcttcc agattttctc ggactccgcg 1500
 catcgccgta ccacttcaaa acaccaagc acagcatact aaatttcccc tctttcttcc 1560
 tctaggggtg cgtaattac ccgtactaaa ggtttggaag agaaaaaga gaccgcctcg 1620
 tttctttttc ttcgtcgaaa aaggcaataa aaatttttat cacgtttctt tttcttgaaa 1680
 attttttttt tgattttttt ctctttcgat gacctcccat tgatatttaa gttaataaac 1740
 ggtcttcaat ttctcaagtt tcagtttcat tttcttggtt ctattacaac tttttttact 1800
 tcttgctcat tagaaagaaa gcatagcaat ctaatctaag tttaaaacaa tgaatgtcac 1860
 ctttgaagaa agagccagtt tacacgggta cagaatcggg atcgcatcct tagacgcccc 1920
 agcatccttg aacgctttgt ccttaccaat gattgatgct ttacaagaca gattgagagc 1980
 ttgggcagaa gatgcagaca tagcctgtgt cttgttacgt ggtaatgggt ctaaagcatt 2040
 ttgcgcaggt ggtgacgttg tccaattagc taagaaatgt ttggcctccc caggtgaagc 2100
 ccctgaattg gctgaaagat ttttcgctag aagttacaga ttagatcatt atttgcacac 2160
 atacccaaaag cctttgatat gctgggctca tgggtcacgtt ttaggtgggtg gtatgggttt 2220
 gttacaaggt gctggtatta gaatagtcac cccatcttca agattggcaa tgcctgaaat 2280
 ctctattggt ttatttccag atgttggtgg ttccatttcc ttaagtagat tgcctggtaa 2340
 attgggttta tttttcggtt taaccgcttc accattgaat gccagagatg ctttggactt 2400
 aaacttggct gatagattct tgtagatac tcaacaagac gcattgatcg atggtttgat 2460
 ccaattgaac tggagagaac aaccagattt gcaattgcat tccttgttaa aggcctttaga 2520
 acaacaagca agaagtgaat taccagctgc acaatggttg cctagaagag aaagattaga 2580
 cgccttggtta gatcaagcta ctttaccatt gtcttggcaa gccttagctt cattggaaaa 2640
 cgatgaagac gcattgttag ccaaagccgc taagactatg ttgggtgggt cacttttaac 2700
 aggtcatttg gtctggggtc aaattagaag agctagacac ttatctttgg cacaagtatt 2760

tcaaatggaa tacggcatgt cattgaattg ttgcagacat ccagaatttg cagaaggtgt 2820
aagagccaga ttaatcgata aagaccacgc tccacattgg cactggcctg acgtaaacca 2880
agttcctgaa gccgttattg cagcccattt tgcaccatta gatgaccacc ctttagctga 2940
tttggcataa cgataccgtc gacctcgagt catgtaatta gttatgtcac gcttacattc 3000
acgccctccc cccacatccg ctctaaccga aaaggaagga gttagacaac ctgaagtcta 3060
ggtccttatt tattttttta tagttatgtt agtattaaga acgttattta tatttcaaata 3120
ttttcttttt tttctgtaca gacgcgtgta cgcattgtaac attatactga aaaccttgct 3180
tgagaagggt ttgggacgct cgaaggcttt aatttgcggc cgtacaatca caaaattaat 3240
cataatgttc atttatctat aatcaaacaa tgagaatata tactgaagtt tggtaggtta 3300
ttcttcagag atgcagtacg atatcaccaa cggtagcgag acacctatca aaacataaac 3360
ggattcatca tgaggtaaaa cgacttctg ttcaacgctg agatcgtgat catgagaatc 3420
tccaccaaca aatgctgtga aagacgcata tagtaaatg ccaccactca taagcagaag 3480
atttccactt at 3492

<210> 44

<211> 1959

<212> DNA

<213> 人工序列

<400> 44

atgtcccaaa ctcataagca cgctattcca gctaatttg ctgatagatg cttgatcaac 60
ccagaacaat acgaaactaa gtacaagcaa tccatcaacg atccagatac tttttggggt 120
gaacaaggta agattttgga ttggattacc ccataccaaa aggtcaagaa tacttctttt 180
gctccaggta acgtttccat caaatgggtat gaagatggta ctttgaactt ggctgctaac 240
tgtttgata gacacttgca agaaaacggt gatagaaccg ctattatttg ggaagggtgat 300
gatgcctctc aatccaaaca tatctcttac agagaattgc acagagatgt ctgtagattc 360
gctaacactt tggttgattt gggtattaag aagggtgacg ttgttgctat ctatatgcc 420
atggttcctg aagctgctgt tgctatgttg gcttggtgta gaattgggtg tggtcattct 480
gttattttcg gtggtttttc accagaagct gttgccggtga gaattatcga ttcttcatcc 540
agattgggtta tcaccgctga tgaagggtt agagctggta gatcaattcc attgaaaaag 600
aacgttgatg acgccttgaa gaacccaaat gttacttctg ttgaacacgt catcgttttg 660
aagagaaccg attctgatat tgactggcaa gaaggtagag atttgtgggtg gagagatttg 720
attgaaaagg cttctccaga acatcaacca gaagcaatga acgctgaaga tcctttgttt 780
atcttgata cttctgggtc tactggtaag ccaaagggtg ttttacacac tactgggtgt 840
tatttggttt acgtgctac tactttcaag tacgttttcg attatcatcc aggtgacatc 900
tattgggtga ctgctgatgt tgggtgggtt actggtcatt cttatttggt gtatgggtcca 960
ttagcttggt gtgctactac cttgatgttc gaaggtgttc caaattggcc aactccagct 1020
agaatgtgtc aagttgttga caaacaccaa gtcaacatct tgtatactgc tccaactgct 1080
attagagctt tgatggctga aggtgataag gctattgaag gtactgatag atcctccttg 1140
agaatcttgg gttctgttgg tgaacctatt aacctgaag cttgggaatg gtactggaag 1200
aaaattggta aagaaaagt cccagttgtt gatacttgggt ggcaactga aacaggtggt 1260

```

tttatgatta ctccattgcc aggtgctatt gaattgaaag ctggttctgc tactagacca 1320
ttttttggtg ttcaaccagc cttggttgat aatgaaggct atccacaaga aggtgctact 1380
gaaggtaatt tggttattac tgattcttgg ccagggtcaag ctagaacttt gtttggtgat 1440
cacgaaagat tcgaacaaac ctacttttcc accttcaaga acatgtactt ttctggtgat 1500
ggtgctagaa gagatgaaga tggttactat tggattaccg gtagagttga tgatgtcttg 1560
aatgtttctg gtcacagatt gggtagagcc gaaattgaat ctgctttggt tgctcatcca 1620
aagattgctg aagcagcagt tgttggtatt ccacatgcta ttaagggtca agcaatctac 1680
gcttacgtta ccttgaatca tggatgaaga ccatccccag aattatacgc tgaagttaga 1740
aactgggtca gaaaagaaat tgggtccattg gctactccag atgttttaca ttggacagat 1800
tctttgccaa agaccagatc aggtgaagatc atgagaagaa tcttgagaaa aattgctgcc 1860
ggtgatactt ctaacttggg tgatacctct actttggctg atccagggtg tgttgaaaag 1920
ccattagaag aaaagcaagc cattgcatg ccactctaa 1959

```

<210> 45

<211> 367

<212> PRT

<213> 门多萨假单胞菌

<220>

<221> misc_feature

<222> (95)...(95)

<223> Xaa的定义见说明书和原始序列表文件

<400> 45

```

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

```

```

Asn Ile Gly Leu Tyr Pro Asp Val Gly Gly Ser Trp Phe Leu Ala Arg
145                150                155                160
Leu Pro Gly Arg Leu Gly Leu Phe Leu Gly Leu Thr Ala Ala Ser Ile
                165                170                175
Asn Ala Arg Asp Ala Leu Asp Leu Asn Leu Ala Asp Arg Phe Leu Arg
                180                185                190
Asp Asp Gln Gln Asp Ala Leu Leu Glu Gly Leu Val Gln Leu Asn Trp
                195                200                205
Arg Glu Gln Pro Ala Ala Gln Leu His Ser Leu Leu Arg Ala Leu Glu
                210                215                220
Asn Glu Ala Arg Gly Glu Leu Pro Ala Ala Gln Trp Leu Pro Arg Arg
225                230                235                240
Glu Arg Ile Asp Glu Leu Leu Asp Val Ala Asp Leu Pro Ala Ala Val
                245                250                255
Gln Ala Ile Ser Ala Leu Gln Gln Asp Asp Asp Ala Leu Leu Ala Arg
                260                265                270
Ala Ala Lys Thr Leu Ala His Gly Cys Pro Leu Thr Ala His Leu Val
                275                280                285
Trp Gln Gln Ile Arg Arg Ala Arg His Leu Ser Leu Ala Glu Val Phe
                290                295                300
Arg Met Glu Tyr Ala Met Ser Leu Asn Cys Cys Arg His Pro Asp Phe
305                310                315                320
Pro Glu Gly Val Arg Ala Arg Leu Ile Asp Lys Asp Gln Thr Pro His
                325                330                335
Trp His Trp Pro Asp Val Ala Ala Ile Pro Glu Ala Val Ile Glu Ala
                340                345                350
His Phe Ala Pro Ala Trp Glu Gly Glu His Pro Leu Ala Asp Leu
                355                360                365

```

<210> 46

<211> 351

<212> PRT

<213> 蜡样芽胞杆菌

<220>

<221> misc_feature

<222> (91) ... (91)

<223> Xaa的定义见说明书和原始序列文件

<400> 46

```

Met Thr Glu His Val Leu Phe Ser Val Ser Glu Asn Gly Val Ala Ser
1                5                10                15

```

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

	325		330		335
Val Ser Glu Glu Asp Val Asn Arg Phe Phe Asn Leu Leu Asn Ala					
	340		345		350
<210> 47					
<211> 360					
<212> PRT					
<213> 蕈状芽孢杆菌					
<220>					
<221> misc_feature					
<222> (101) ... (101)					
<223> Xaa的定义见说明书和原始序列文件					
<400> 47					
Met Arg Arg Tyr Ile Arg Gly Gly Phe Lys Met Thr Glu Asn Val Leu					
1	5		10		15
Phe Ser Ile His Glu Asn Gly Val Ala Ser Ile Thr Leu Asn Arg Pro					
	20		25		30
Lys Ala Leu Asn Ser Leu Ser Tyr Asp Met Leu His Pro Ile Gly Gln					
	35		40		45
Lys Leu Lys Glu Trp Glu Lys Asp Asp Arg Ile Ala Val Val Ile Leu					
	50		55		60
Lys Gly Ala Gly Thr Lys Gly Phe Cys Ala Gly Gly Asp Ile Lys Thr					
65	70		75		80
Leu Tyr Glu Ala Arg Ser Asn Glu Val Ala Leu Gln His Ala Glu His					
	85		90		95
Phe Phe Glu Glu Xaa Tyr Glu Ile Asp Thr Tyr Ile Tyr His Tyr Pro					
	100		105		110
Lys Pro Ile Ile Ala Cys Leu Asp Gly Ile Val Met Gly Gly Gly Val					
	115		120		125
Gly Leu Thr Asn Gly Ala Lys Tyr Arg Ile Val Thr Asp Arg Thr Lys					
	130		135		140
Trp Ala Met Pro Glu Met Asn Ile Gly Phe Phe Pro Asp Val Gly Ala					
145	150		155		160
Ala Tyr Phe Leu Asn Lys Ala Pro Gly Gln Thr Gly Arg Tyr Val Ala					
	165		170		175
Leu Thr Ala Ser Val Leu Lys Ala Ala Asp Val Leu Tyr Ile Lys Ala					
	180		185		190
Ala Asp His Tyr Met Pro Ser Glu Thr Leu Pro Thr Phe Leu Asp Ala					
	195		200		205
Ile Glu Lys Val Asn Trp His Asp Gln Asn Ile His Ile Thr Leu Lys					

210	215	220
Glu Leu Ile Asn Lys Tyr Glu Thr Ala Pro Ser Val Glu Ser Glu Leu		
225	230	235
Val Ser Leu Leu Glu Glu Ile Asp Gln His Phe Ser Phe His Thr Val		240
	245	250
Glu Asp Ile Ile His Ser Leu Asp Asn Ala Asp Gly Ser Phe Ala Ser		255
	260	265
Lys Thr Lys Glu Thr Leu Leu Ser Lys Ser Pro Phe Ser Leu Lys Val		270
	275	280
Thr Leu Lys Gln Leu Ile Asp Gly Lys Glu Lys Ser Ile Glu Glu Cys		285
290	295	300
Phe Ala Thr Asp Leu Val Leu Ala Lys Asn Phe Met Arg His Glu Asp		
305	310	315
Phe Phe Glu Gly Val Arg Ser Val Val Val Asp Lys Asp Gln Asn Pro		320
	325	330
Lys Tyr Lys Tyr Lys Gln Leu Ser Asp Val Ser Asp Glu Asp Val Asn		335
	340	345
Arg Phe Phe Asn Leu Leu Asn Ala		350
355	360	

<210> 48

<211> 370

<212> PRT

<213> 荧光假单胞菌

<220>

<221> misc_feature

<222> (95) ... (95)

<223> Xaa的定义见说明书和原始序列表文件

<400> 48

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

	85		90		95
Arg Leu Asp Tyr Arg Leu His Thr Phe Pro Lys Pro Leu Ile Cys Trp					
100		105		110	
Gly His Gly Tyr Val Leu Gly Gly Gly Met Gly Leu Leu Gln Ser Ala					
115		120		125	
Ala Val Arg Ile Val Thr Pro Ser Ser Arg Leu Ala Met Pro Glu Ile					
130		135		140	
Ser Ile Gly Leu Tyr Pro Asp Val Gly Ala Ser Trp Phe Leu Ser Arg					
145		150		155	160
Leu Pro Gly Lys Leu Gly Leu Phe Leu Gly Leu Thr Gly Ala His Val					
165		170		175	
Asn Gly Arg Asp Ala Leu Asp Leu Gly Leu Ala Asp Arg Phe Leu Arg					
180		185		190	
Asp Asp Gln Gln Asp Glu Leu Ile Glu Gly Leu Leu Gln Leu Asn Trp					
195		200		205	
Gln Glu Gln Thr Ala Met Gln Leu Asn Ser Leu Phe Lys Ala Leu Ala					
210		215		220	
Gln Glu Ala Val Asp Gln Leu Pro Glu Ala Gln Trp Leu Pro Arg Arg					
225		230		235	240
Ala Gln Ile Asp Glu Trp Leu Asp Val Gly Asp Val Arg Ser Ala Trp					
245		250		255	
Arg Ala Leu Ser Gln Leu Arg Asp His Ala Asp Pro Leu Phe Ser Arg					
260		265		270	
Ala Gly Lys Thr Leu Ser Glu Gly Cys Pro Leu Thr Ala His Leu Val					
275		280		285	
Trp Glu Gln Ile Gln Arg Ala Arg His Leu Ser Leu Ala Gln Val Phe					
290		295		300	
Gln Met Glu Tyr Thr Leu Ser Leu Asn Cys Cys Arg His Pro Glu Phe					
305		310		315	320
Ser Glu Gly Val Arg Ala Arg Leu Ile Asp Lys Asp Gln Thr Pro Arg					
325		330		335	
Trp His Trp Pro Asp Val His Thr Leu Pro Asp Ala Val Val Gln Ala					
340		345		350	
His Phe Asn Lys Ala Trp Glu Gly Arg His Pro Leu Ala Asp Leu Ser					
355		360		365	
Asp Tyr					
370					
<210> 49					
<211> 351					

<212> PRT

<213> 苏云金杆菌幕虫亚种

<220>

<221> misc_feature

<222> (91) ... (91)

<223> Xaa的定义见说明书和原始序列表文件

<400> 49

```

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

```

Asp Gln Ser Ser Phe Ala Leu Lys Ala Lys Glu Thr Leu Leu Ser Lys
 260 265 270
 Ser Pro Ile Ser Leu Lys Val Thr Leu Lys Gln Phe Ile Asp Gly Gln
 275 280 285
 Asn Lys Ser Val Glu Glu Cys Phe Ala Thr Asp Leu Val Leu Ala Lys
 290 295 300
 Asn Phe Met Arg His Glu Asp Phe Phe Glu Gly Val Arg Ser Val Val
 305 310 315 320
 Val Asp Lys Asp Gln Asn Pro Asp Tyr Lys Tyr Lys Gln Leu Ser Asp
 325 330 335
 Val Ser Glu Glu Asp Val Asn Arg Phe Phe Asn Leu Leu Asn Ala
 340 345 350

<210> 50

<211> 366

<212> PRT

<213> 黄褐假单胞菌

<220>

<221> misc_feature

<222> (95)...(95)

<223> Xaa的定义见说明书和原始序列表文件

<400> 50

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

```

Ser Ile Gly Leu Phe Pro Asp Val Gly Gly Ser His Phe Leu Ser Arg
145                      150                      155                      160
Leu Pro Gly Lys Leu Gly Leu Phe Phe Gly Leu Thr Ala Ser Pro Leu
                      165                      170                      175
Asn Ala Arg Asp Ala Leu Asp Leu Asn Leu Ala Asp Arg Phe Leu Leu
                      180                      185                      190
Asp Thr Gln Gln Asp Ala Leu Ile Asp Gly Leu Ile Gln Leu Asn Trp
                      195                      200                      205
Arg Glu Gln Pro Asp Leu Gln Leu His Ser Leu Leu Lys Ala Leu Glu
                      210                      215                      220
Gln Gln Ala Arg Ser Glu Leu Pro Ala Ala Gln Trp Leu Pro Arg Arg
225                      230                      235                      240
Glu Arg Leu Asp Ala Leu Leu Asp Gln Ala Thr Leu Pro Leu Ser Trp
                      245                      250                      255
Gln Ala Leu Ala Ser Leu Glu Asn Asp Glu Asp Ala Leu Leu Ala Lys
                      260                      265                      270
Ala Ala Lys Thr Met Leu Gly Gly Ser Pro Leu Thr Gly His Leu Val
                      275                      280                      285
Trp Gly Gln Ile Arg Arg Ala Arg His Leu Ser Leu Ala Gln Val Phe
                      290                      295                      300
Gln Met Glu Tyr Gly Met Ser Leu Asn Cys Cys Arg His Pro Glu Phe
305                      310                      315                      320
Ala Glu Gly Val Arg Ala Arg Leu Ile Asp Lys Asp His Ala Pro His
                      325                      330                      335
Trp His Trp Pro Asp Val Asn Gln Val Pro Glu Gln Val Ile Ala Ala
                      340                      345                      350
His Phe Ala Pro Leu Asp Asp His Pro Leu Ala Asp Leu Ala
                      355                      360                      365

```

<210> 51

<211> 351

<212> PRT

<213> 蜡样芽胞杆菌

<220>

<221> misc_feature

<222> (91) ... (91)

<223> Xaa的定义见说明书和原始序列文件

<400> 51

```

Met Thr Glu His Val Leu Phe Ser Val Ser Glu Asn Gly Val Ala Ser
1                      5                      10                      15

```

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

	325		330		335
Val Ser Asp Glu Asp Val Asn Arg Phe Phe Asn Leu Leu Asn Ala					
	340		345		350
<210> 52					
<211> 351					
<212> PRT					
<213> 蜡样芽胞杆菌					
<220>					
<221> misc_feature					
<222> (91) ... (91)					
<223> Xaa的定义见说明书和原始序列文件					
<400> 52					
Met Thr Glu Gln Val Leu Phe Ser Val Ser Glu Asn Gly Val Ala Thr					
1	5		10		15
Ile Thr Leu Asn Arg Pro Lys Ala Leu Asn Ser Leu Ser Tyr Asp Met					
	20		25		30
Leu Gln Pro Ile Gly Gln Lys Leu Lys Glu Trp Glu His Asp Glu Arg					
	35		40		45
Ile Ala Leu Ile Val Leu Lys Gly Ala Gly Thr Lys Gly Phe Cys Ala					
	50		55		60
Gly Gly Asp Ile Lys Thr Leu Tyr Glu Ala Arg Ser Asn Glu Val Ala					
65	70		75		80
Leu Gln His Ala Glu Arg Phe Phe Glu Glu Xaa Tyr Glu Ile Asp Thr					
	85		90		95
Tyr Ile Tyr Gln Tyr Thr Lys Pro Ile Ile Ala Cys Leu Asp Gly Ile					
	100		105		110
Val Met Gly Gly Gly Val Gly Leu Thr Asn Gly Ala Lys Tyr Arg Ile					
	115		120		125
Val Thr Glu Arg Thr Lys Trp Ala Met Pro Glu Met Asn Ile Gly Phe					
	130		135		140
Phe Pro Asp Val Gly Ala Ala Tyr Phe Leu Asn Lys Ala Pro Gly Tyr					
145	150		155		160
Thr Gly Arg Phe Val Ala Leu Thr Ala Ser Ile Leu Lys Ala Ser Asp					
	165		170		175
Val Leu Phe Ile Asn Ala Ala Asp Tyr Phe Met Thr Ser Asp Ser Leu					
	180		185		190
Pro Glu Phe Leu Thr Glu Leu Glu Ser Val Asn Trp His Lys Glu Asp					
	195		200		205
Asp Val His Thr Asn Leu Lys Glu Val Ile Arg Thr Phe Ala Thr Ala					

210	215	220
Pro Asn Leu Glu Ser Glu Leu Ala Pro Ser Leu Glu Val Ile Asn Ser		
225	230	235
His Phe Ala Phe Asp Thr Ile Glu Glu Ile Ile His Ser Leu Glu Lys		240
	245	250
Asp Glu Ser Ser Phe Ala Leu Lys Thr Lys Glu Ile Leu Leu Ser Lys		255
	260	265
Ser Pro Ile Ser Leu Lys Val Thr Leu Lys Gln Phe Ile Asp Gly Gln		270
	275	280
Asp Lys Ser Val Glu Glu Cys Phe Ala Thr Asp Leu Ile Leu Ala Lys		285
290	295	300
Asn Phe Met Arg His Glu Asp Phe Phe Glu Gly Val Arg Ser Val Val		
305	310	315
Val Asp Lys Asp Gln Asn Pro Asn Tyr Lys Tyr Lys Gln Leu Ser Asp		320
	325	330
Val Ser Glu Glu Asp Val Asn Arg Phe Phe Asn Leu Leu Asn Ala		335
	340	345
		350
<210> 53		
<211> 309		
<212> PRT		
<213> 大肠杆菌		
<400> 53		
Met Thr Gln Phe Ala Phe Val Phe Pro Gly Gln Gly Ser Gln Thr Val		
1	5	10
Gly Met Leu Ala Asp Met Ala Ala Ser Tyr Pro Ile Val Glu Glu Thr		
	20	25
Phe Ala Glu Ala Ser Ala Ala Leu Gly Tyr Asp Leu Trp Ala Leu Thr		30
	35	40
Gln Gln Gly Pro Ala Glu Glu Leu Asn Lys Thr Trp Gln Thr Gln Pro		45
50	55	60
Ala Leu Leu Thr Ala Ser Val Ala Leu Tyr Arg Val Trp Gln Gln Gln		
65	70	75
Gly Gly Lys Ala Pro Ala Met Met Ala Gly His Ser Leu Gly Glu Tyr		80
	85	90
Ser Ala Leu Val Cys Ala Gly Val Ile Asp Phe Ala Asp Ala Val Arg		95
	100	105
Leu Val Glu Met Arg Gly Lys Phe Met Gln Glu Ala Val Pro Glu Gly		110
	115	120
Thr Gly Ala Met Ala Ala Ile Ile Gly Leu Asp Asp Ala Ser Ile Ala		125

130	135	140
Lys Ala Cys Glu Glu Ala Ala Glu Gly Gln Val Val Ser Pro Val Asn		
145	150	155
Phe Asn Ser Pro Gly Gln Val Val Ile Ala Gly His Lys Glu Ala Val		
	165	170
Glu Arg Ala Gly Ala Ala Cys Lys Ala Ala Gly Ala Lys Arg Ala Leu		
	180	185
Pro Leu Pro Val Ser Val Pro Ser His Cys Ala Leu Met Lys Pro Ala		
	195	200
Ala Asp Lys Leu Ala Val Glu Leu Ala Lys Ile Thr Phe Asn Ala Pro		
210	215	220
Thr Val Pro Val Val Asn Asn Val Asp Val Lys Cys Glu Thr Asn Gly		
225	230	235
Asp Ala Ile Arg Asp Ala Leu Val Arg Gln Leu Tyr Asn Pro Val Gln		
	245	250
Trp Thr Lys Ser Val Glu Tyr Met Ala Ala Gln Gly Val Glu His Leu		
	260	265
Tyr Glu Val Gly Pro Gly Lys Val Leu Thr Gly Leu Thr Lys Arg Ile		
	275	280
Val Asp Thr Leu Thr Ala Ser Ala Leu Asn Glu Pro Ser Ala Met Ala		
290	295	300
Ala Ala Leu Glu Leu		
305		

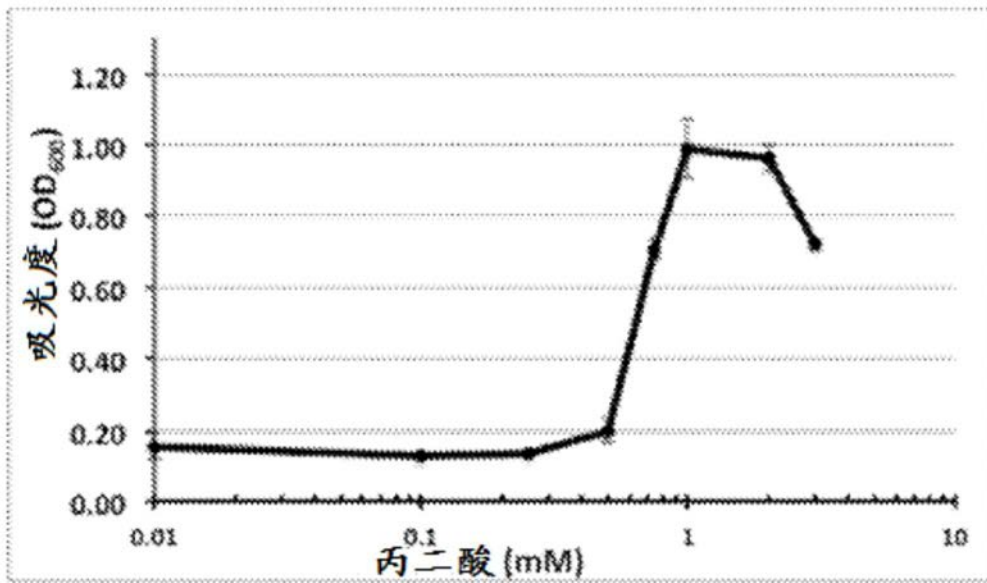


图1

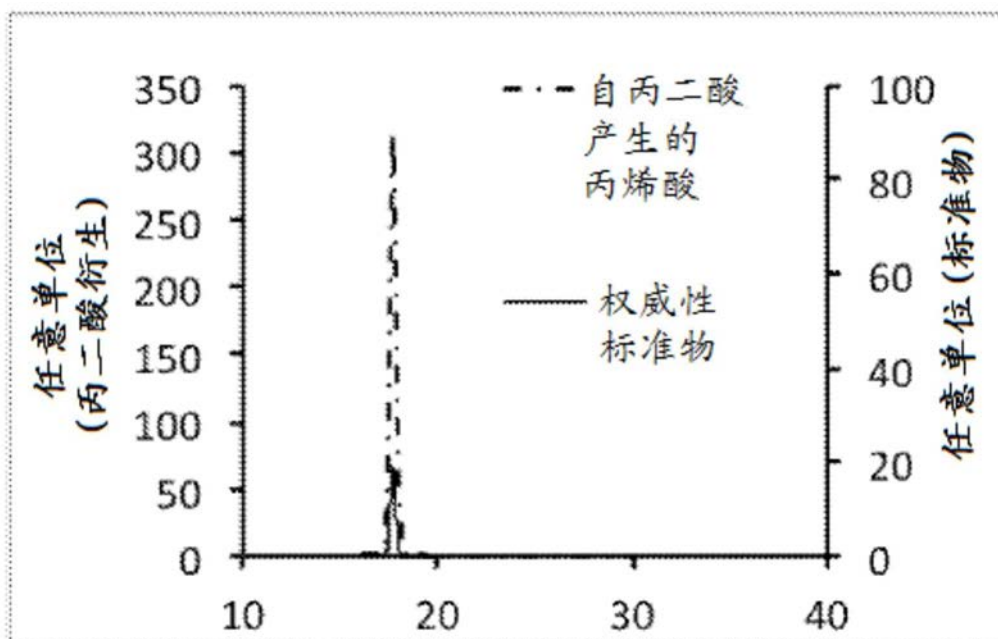


图2