



(12) 发明专利申请

(10) 申请公布号 CN 111748511 A

(43) 申请公布日 2020. 10. 09

(21) 申请号 202010692338.3

(22) 申请日 2015.07.17

(30) 优先权数据

62/026,573 2014.07.18 US

(62) 分案原申请数据

201580044132.7 2015.07.17

(71) 申请人 基因组股份公司

地址 美国加利福尼亚

(72) 发明人 A·W·施尔摩 N·赫尔曼 H·王

Z·胡 V·阿拉加达

A·I·拉默斯-索利斯 E·克拉克

(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 陈晓娜

(51) Int.Cl.

C12N 1/21 (2006.01)

C12N 15/53 (2006.01)

C12P 7/18 (2006.01)

C12R 1/19 (2006.01)

权利要求书1页 说明书50页

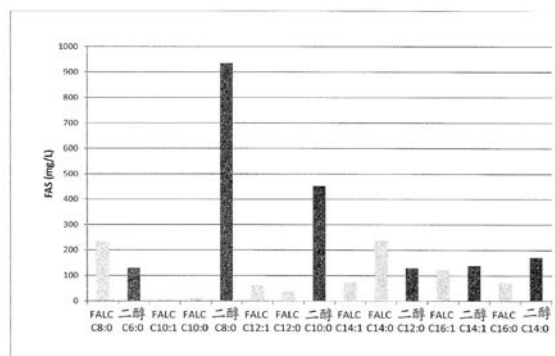
序列表20页 附图11页

(54) 发明名称

微生物的脂肪二醇产生

(57) 摘要

本公开涉及脂肪二醇以及用于产生其的重组微生物。更具体地说,本公开涉及经工程化以经由发酵产生脂肪二醇的重组微生物。进一步涵盖一种使用所述微生物从简单碳源产生脂肪二醇的方法。



1. 一种重组微生物,其在具有简单碳源的发酵肉汤中生长时产生1,3脂肪二醇,所述微生物经工程化以表达编码包含酰基-ACP还原酶(EC 1.2.1.80或EC 1.2.1.42)活性的多肽的核酸序列。

2. 如权利要求1所述的重组微生物,其进一步表达编码包含醇脱氢酶(EC 1.1.1.-)活性的多肽的核酸序列。

3. 如权利要求1所述的重组微生物,其中所述1,3脂肪二醇在体内产生。

4. 如权利要求3所述的重组微生物,其中所述1,3脂肪二醇选自由以下组成的群组:C₅ 1,3脂肪二醇、C₆ 1,3脂肪二醇、C₇ 1,3脂肪二醇、C₈ 1,3脂肪二醇、C₉ 1,3脂肪二醇、C₁₀ 1,3脂肪二醇、C₁₁ 1,3脂肪二醇、C₁₂ 1,3脂肪二醇、C₁₃ 1,3脂肪二醇、C₁₄ 1,3脂肪二醇、C₁₅ 1,3脂肪二醇、C₁₆ 1,3脂肪二醇、C₁₇ 1,3脂肪二醇、C₁₈ 1,3脂肪二醇以及C₁₉ 1,3脂肪二醇。

5. 如权利要求3所述的重组微生物,其中所述1,3脂肪二醇是C₁₂ 1,3脂肪二醇。

6. 如权利要求1所述的重组微生物,其中所述简单碳源来源于可再生原料。

7. 一种细胞培养物,其包含根据权利要求1至6中任一项所述的微生物。

8. 如权利要求7所述的细胞培养物,其中所述细胞培养物产生1,3脂肪二醇。

9. 如权利要求8所述的细胞培养物,其中所述1,3脂肪二醇选自由以下组成的群组:C₅ 1,3脂肪二醇、C₆ 1,3脂肪二醇、C₇ 1,3脂肪二醇、C₈ 1,3脂肪二醇、C₉ 1,3脂肪二醇、C₁₀ 1,3脂肪二醇、C₁₁ 1,3脂肪二醇、C₁₂ 1,3脂肪二醇、C₁₃ 1,3脂肪二醇、C₁₄ 1,3脂肪二醇、C₁₅ 1,3脂肪二醇、C₁₆ 1,3脂肪二醇、C₁₇ 1,3脂肪二醇、C₁₈ 1,3脂肪二醇以及C₁₉ 1,3脂肪二醇。

10. 一种产生1,3脂肪二醇的方法,其包括培养如权利要求1所述的微生物。

11. 一种产生1,3脂肪二醇的方法,其包括:

(a) 在发酵肉汤中培养重组微生物,所述微生物经工程化以表达编码包含酰基-ACP还原酶(EC 1.2.1.80或EC 1.2.1.42)活性的多肽的核酸序列;以及

(b) 从所述发酵肉汤分离1,3脂肪二醇,其中所述发酵肉汤包含简单碳源。

12. 如权利要求11所述的方法,其进一步表达编码包含醇脱氢酶(EC 1.1.1.-)活性的多肽的核酸序列。

13. 如权利要求11所述的方法,其中所述1,3脂肪二醇选自由以下组成的群组:C₅ 1,3脂肪二醇、C₆ 1,3脂肪二醇、C₇ 1,3脂肪二醇、C₈ 1,3脂肪二醇、C₉ 1,3脂肪二醇、C₁₀ 1,3脂肪二醇、C₁₁ 1,3脂肪二醇、C₁₂ 1,3脂肪二醇、C₁₃ 1,3脂肪二醇、C₁₄ 1,3脂肪二醇、C₁₅ 1,3脂肪二醇、C₁₆ 1,3脂肪二醇、C₁₇ 1,3脂肪二醇、C₁₈ 1,3脂肪二醇以及C₁₉ 1,3脂肪二醇。

微生物的脂肪二醇产生

[0001] 本申请是申请号为201580044132.7的发明专利申请的分案申请。原申请的申请日是2015年7月17日,优先权日期是2014年7月18日,发明名称为“微生物的脂肪二醇产生”。

[0002] 相关申请的交叉引用

[0003] 本申请要求2014年7月18日提交的美国临时申请No.62/026,573的权益,所述临时申请的全部公开内容以引用的方式并入本文中。

[0004] 序列列表

[0005] 本申请包含序列列表,其已经呈ASCII格式电子递交并以引用的方式整体并入本文中。所述ASCII拷贝于2015年7月17日创建,命名为LS00052PCT_SL.txt且大小是50,064个字节。

发明领域

[0006] 本公开涉及脂肪二醇及其产生方法。本文中,本公开涉及经工程化以经由发酵产生脂肪二醇的重组微生物。进一步涵盖一种使用微生物从简单碳源产生脂肪二醇的方法。

[0007] 发明背景

[0008] 脂肪醇作为工业试剂和工艺的组分具有许多商业用途,特别是在清洁剂和表面活性剂的生产中。它们用作化妆品和食品中的乳化剂、软化剂和增稠剂,以及作为工业溶剂和增塑剂。脂肪醇可以从石油化学或油脂化学来源的原料产生。石油化学品是来源于石油的化学产品。油脂化学品是来源于例如植物脂肪和动物脂肪等天然来源的精炼油。

[0009] 用于制造脂肪醇的化学途径是能量密集和高环境成本的,并需要使用危险的试剂。举例来说,乙烯可以使用三乙基铝寡聚,接着空气氧化。此方法产生偶数脂肪醇并称为齐格勒法(Ziegler process)。或者,乙烯可以寡聚,得到烯烃混合物,接着对烯烃进行加氢甲酰化,得到奇数醛,随后对奇数醛进行氢化,得到脂肪醇。在另一化学方法中,烯烃产品转变成脂肪醛,然后转变成脂肪醇。所述烯烃产品通过壳牌高碳烯烃法(Shell higher olefin process)制备,此法在1977年由Royal Dutch Shell商业化(例如每年生产大约超过一百万吨烯烃)。

[0010] 用于制造脂肪醇的天然途径虽然被视为是绿色方法,但与化学途径相比仍然成本高。传统上,脂肪醇来源于脂肪酸酯或蜡酯,脂肪酸酯或蜡酯最初是从鲸的鲸油提取,后来从牛脂(例如来自牛肉或羊羔的动物脂肪)提取。蜡酯的一种替代植物来源是西蒙德木植物(jojoba plant)。今天,脂肪醇还可以从例如菜籽油、芥籽油、椰子油或棕榈油等油脂化学来源的原料(例如精炼植物油)产生。此类植物油主要是由含有经三种脂肪酸(FA)酯化的甘油的三酰基甘油(TAG)构成。植物油的多样用途依赖于TAG的FA组成。举例来说,肥皂生产需要高比例的月桂酸(12:0),而推荐将富含油酸的油(18:1)用于烹饪。可以对TAG进行酯交换反应,得到酯,随后将酯氢化成脂肪醇。虽然牛脂通常是C₁₆-C₁₈,但来自植物来源的链长更可变(例如C₆-C₂₄)。长链醇(例如C₂₀-C₂₂)可以从菜籽或芥菜籽获得,而中值脂肪醇(例如C₁₂-C₁₄)可以从椰子或棕榈油获得。椰子和棕榈油富含月桂酸(C₁₂)和肉豆蔻酸(C₁₄)。因为2000年欧洲爆发了牛海绵状脑病(即疯牛病),所以牛脂通常被来源于棕榈油和豆油的植物油脂

肪酸代替。

[0011] 脂肪二醇或脂肪族二醇是脂肪醇的实例,且可以经由化学方法产生。举例来说,1,3-二醇可以由乙烯和羧酰氯化物合成(参见例如Kirchanov等人(1981) Translation from Izvestiya Akademii Nauk SSSR, Seriya Khimicheskaya 4:909-911)。1,3-二醇还可以通过 α,β -不饱和酮和醛的水合来制备,其中所得酮-醇氢化。1,3-二醇的另一化学合成包括环氧化物的加氢甲酰化,接着醛氢化(例如由环氧乙烷制备1,3-丙二醇)。1,3-二醇的更专门途径包括烯烃与甲醛之间的反应以及 β -羟基酮的使用。1,3-二醇已经可用作食品添加剂(参见例如美国专利No.3,806,615)。1,3-二羟基构型使得这些化学实体无毒性。

[0012] 1,3-二醇是双官能的,且可以用作其它分子之间的连接分子,例如在聚合物的产生中。举例来说,1,3-丙二醇用作聚合物产生中的单体。1,3脂肪二醇还可以用作表面活性剂的前驱物,例如“Gemini”表面活性剂,其中两个醇部分经化学改性(例如乙氧基化、糖基化、硫酸化等等)。1,3脂肪二醇的3-羟基部分也是手性的,这使得1,3脂肪二醇适用作产生例如单体、药物、营养品、杀虫剂、除草剂、香精、香料、溶剂等手性重要的化合物的合成子。

[0013] 因为脂肪二醇是工业试剂和工艺的重要组分,所以希望足够大量地生产它们,以满足工业需求,同时维持较低的对环境的影响。本发明解决了此需求。

发明概要

[0014] 本公开的一方面提供一种重组微生物,其在具有简单碳源的发酵肉汤中生长时产生1,3脂肪二醇,所述微生物包括编码具有硫酯酶(EC 3.1.2.-、EC 3.1.1.5或EC 3.1.2.14)活性和羧酸还原酶(EC 6.2.1.3或EC 1.2.1.42)活性的多肽的核酸序列。在一方面,1,3脂肪二醇在体内产生。在另一方面,1,3脂肪二醇包括但不限于C₅ 1,3脂肪二醇、C₆ 1,3脂肪二醇、C₇ 1,3脂肪二醇、C₈ 1,3脂肪二醇、C₉ 1,3脂肪二醇、C₁₀ 1,3脂肪二醇、C₁₁ 1,3脂肪二醇、C₁₂ 1,3脂肪二醇、C₁₃ 1,3脂肪二醇、C₁₄ 1,3脂肪二醇、C₁₅ 1,3脂肪二醇、C₁₆ 1,3脂肪二醇、C₁₇ 1,3脂肪二醇、C₁₈ 1,3脂肪二醇以及C₁₉ 1,3脂肪二醇。在又一方面,简单碳源来源于可再生原料。在一个实施方案中,本公开提供一种重组微生物,其中所述微生物包括编码具有硫酯酶(EC 3.1.2.-、EC 3.1.1.5或EC 3.1.2.14)活性和羧酸还原酶(EC 6.2.1.3或EC 1.2.1.42)活性的多肽的核酸序列,且其中当在具有简单碳源的发酵肉汤中生长时所述微生物产生1,3脂肪二醇。在另一个实施方案中,核酸序列是外源性的。在另一个实施方案中,核酸序列包括一个或多个核酸序列。

[0015] 本公开的另一方面提供一种重组微生物,其在具有简单碳源的发酵肉汤中生长时产生1,3脂肪二醇,所述微生物包括经工程化以表达编码具有硫酯酶(EC 3.1.2.-、EC 3.1.1.5或EC 3.1.2.14)活性和羧酸还原酶(EC 6.2.1.3或EC 1.2.1.42)活性的多肽的核酸序列的途径。在一方面,1,3脂肪二醇在体内产生。在另一方面,1,3脂肪二醇包括但不限于C₅ 1,3脂肪二醇、C₆ 1,3脂肪二醇、C₇ 1,3脂肪二醇、C₈ 1,3脂肪二醇、C₉ 1,3脂肪二醇、C₁₀ 1,3脂肪二醇、C₁₁ 1,3脂肪二醇、C₁₂ 1,3脂肪二醇、C₁₃ 1,3脂肪二醇、C₁₄ 1,3脂肪二醇、C₁₅ 1,3脂肪二醇、C₁₆ 1,3脂肪二醇、C₁₇ 1,3脂肪二醇、C₁₈ 1,3脂肪二醇以及C₁₉ 1,3脂肪二醇。在又一方面,简单碳源来源于可再生原料。在一个实施方案中,本公开提供一种重组微生物,所述微生物具有经工程化以表达编码具有硫酯酶(EC 3.1.2.-、EC 3.1.1.5或EC 3.1.2.14)活性和羧酸还原酶(EC 6.2.1.3或EC 1.2.1.42)活性的多肽的核酸序列的途径,

其中当在具有简单碳源的发酵肉汤中生长时所述微生物产生1,3脂肪二醇。在另一个实施方案中,核酸序列是外源性的。在另一个实施方案中,核酸序列包括一个或多个核酸序列。

[0016] 本公开的另一方面提供一种重组微生物,其在具有简单碳源的发酵肉汤中生长时产生1,3脂肪二醇,所述微生物经工程化以表达编码具有硫酯酶(EC 3.1.2.-、EC 3.1.1.5或EC 3.1.2.14)活性、羧酸还原酶(EC 6.2.1.3或EC 1.2.1.42)活性和任选醇脱氢酶(EC 1.1.1.)活性的多肽的一个或多个核酸序列。在一方面,1,3脂肪二醇在体内产生。在另一方面,1,3脂肪二醇包括但不限于C₅ 1,3脂肪二醇、C₆ 1,3脂肪二醇、C₇ 1,3脂肪二醇、C₈ 1,3脂肪二醇、C₉ 1,3脂肪二醇、C₁₀ 1,3脂肪二醇、C₁₁ 1,3脂肪二醇、C₁₂ 1,3脂肪二醇、C₁₃ 1,3脂肪二醇、C₁₄ 1,3脂肪二醇、C₁₅ 1,3脂肪二醇、C₁₆ 1,3脂肪二醇、C₁₇ 1,3脂肪二醇、C₁₈ 1,3脂肪二醇以及C₁₉ 1,3脂肪二醇。在又一方面,简单碳源来源于可再生原料。在一个实施方案中,本公开提供一种重组微生物,所述微生物经工程化以表达编码具有硫酯酶(EC 3.1.2.-、EC 3.1.1.5或EC 3.1.2.14)活性、羧酸还原酶(EC 6.2.1.3或EC 1.2.1.42)活性和醇脱氢酶(EC 1.1.1.)活性的多肽的一个或多个核酸序列,其中所述微生物当在具有简单碳源的发酵肉汤中生长时产生1,3脂肪二醇。在另一个实施方案中,所述一个或多个核酸序列是外源性的。

[0017] 本公开的另一方面提供一种重组微生物,其在具有简单碳源的发酵肉汤中生长时产生1,3脂肪二醇,所述微生物包括经工程化以表达编码具有硫酯酶(EC 3.1.2.-、EC 3.1.1.5或EC 3.1.2.14)活性、羧酸还原酶(EC 6.2.1.3或EC 1.2.1.42)活性和任选醇脱氢酶(EC 1.1.1.)活性的多肽的一个或多个核酸序列的途径。在一方面,1,3脂肪二醇在体内产生。在另一方面,1,3脂肪二醇包括但不限于C₅ 1,3脂肪二醇、C₆ 1,3脂肪二醇、C₇ 1,3脂肪二醇、C₈ 1,3脂肪二醇、C₉ 1,3脂肪二醇、C₁₀ 1,3脂肪二醇、C₁₁ 1,3脂肪二醇、C₁₂ 1,3脂肪二醇、C₁₃ 1,3脂肪二醇、C₁₄ 1,3脂肪二醇、C₁₅ 1,3脂肪二醇、C₁₆ 1,3脂肪二醇、C₁₇ 1,3脂肪二醇、C₁₈ 1,3脂肪二醇以及C₁₉ 1,3脂肪二醇。在又一方面,简单碳源来源于可再生原料。在一个实施方案中,本公开提供一种重组微生物,所述微生物具有经工程化以表达编码具有硫酯酶(EC 3.1.2.-、EC 3.1.1.5或EC 3.1.2.14)活性、羧酸还原酶(EC 6.2.1.3或EC 1.2.1.42)活性和醇脱氢酶(EC 1.1.1.)活性的多肽的一个或多个核酸序列的途径,其中所述微生物当在具有简单碳源的发酵肉汤中生长时产生1,3脂肪二醇。在另一个实施方案中,所述一个或多个核酸序列是外源性的。

[0018] 本公开的另一方面提供一种重组微生物,其在具有简单碳源的发酵肉汤中生长时产生1,3脂肪二醇,其中所述简单碳源来源于可再生原料。

[0019] 本公开的另一方面提供一种重组微生物,其在具有简单碳源的发酵肉汤中生长时产生1,3脂肪二醇,其中所述微生物表达编码具有硫酯酶(EC 3.1.2.-、EC 3.1.1.5或EC 3.1.2.14)活性和羧酸还原酶(EC 6.2.1.3或EC 1.2.1.42)活性的多肽的一个或多个核酸序列。在一个实施方案中,硫酯酶包括但不限于fatB1、TE_EE182564、TE_CAD63310、phaG和tesA。在另一个实施方案中,羧酸还原酶是carB。在另一个实施方案中,所述一个或多个核酸序列是外源性的。

[0020] 本公开的另一方面提供一种重组微生物,其在具有简单碳源的发酵肉汤中生长时产生1,3脂肪二醇,所述微生物包括经工程化以表达编码具有硫酯酶(EC 3.1.2.-、EC 3.1.1.5或EC 3.1.2.14)活性和羧酸还原酶(EC 6.2.1.3或EC 1.2.1.42)活性的多肽的一

个或多个核酸序列的途径。在一个实施方案中,硫酯酶包括但不限于fatB1、TE_EEI82564、TE_CAD63310、phaG和tesA。在另一个实施方案中,羧酸还原酶是carB。在另一个实施方案中,所述一个或多个核酸序列是外源性的。

[0021] 本公开的另一方面提供一种重组微生物,其在具有简单碳源的发酵肉汤中生长时产生1,3脂肪二醇,其中所述微生物表达编码具有硫酯酶(EC 3.1.2.-、EC 3.1.1.5或EC 3.1.2.14)活性、羧酸还原酶(EC 6.2.1.3或EC 1.2.1.42)活性和醇脱氢酶(EC 1.1.1.)的多肽的一个或多个核酸序列。在一个实施方案中,硫酯酶包括但不限于fatB1、TE_EEI82564、TE_CAD63310、phaG和tesA。在另一个实施方案中,羧酸还原酶是carB。在又一个实施方案中,醇脱氢酶是alrA。在又一个实施方案中,所述一个或多个核酸序列是外源性的。

[0022] 本公开的另一方面提供一种重组微生物,其在具有简单碳源的发酵肉汤中生长时产生1,3脂肪二醇,所述微生物包括经工程化以表达编码具有硫酯酶(EC 3.1.2.-、EC 3.1.1.5或EC 3.1.2.14)活性、羧酸还原酶(EC 6.2.1.3或EC 1.2.1.42)活性和醇脱氢酶(EC 1.1.1.)活性的多肽的一个或多个核酸序列的途径。在一个实施方案中,硫酯酶包括但不限于fatB1、TE_EEI82564、TE_CAD63310、phaG和tesA。在另一个实施方案中,羧酸还原酶是carB。在又一个实施方案中,醇脱氢酶是alrA。在又一个实施方案中,所述一个或多个核酸序列是外源性的。

[0023] 本公开进一步涵盖一种细胞培养物,其包括在具有简单碳源的发酵肉汤中生长时产生1,3脂肪二醇的重组微生物。在一方面,微生物经工程化以表达编码具有硫酯酶(EC 3.1.2.-、EC 3.1.1.5或EC 3.1.2.14)活性和羧酸还原酶(EC 6.2.1.3或EC 1.2.1.42)活性的多肽的核酸序列。在另一方面,微生物经工程化以表达编码具有硫酯酶(EC 3.1.2.-、EC 3.1.1.5或EC 3.1.2.14)活性、羧酸还原酶(EC 6.2.1.3或EC 1.2.1.42)活性和醇脱氢酶(EC 1.1.1.-)活性的多肽的核酸序列。在另一方面,细胞培养物产生1,3脂肪二醇。在另一方面,细胞培养物产生包括C₅ 1,3脂肪二醇、C₆ 1,3脂肪二醇、C₇ 1,3脂肪二醇、C₈ 1,3脂肪二醇、C₉ 1,3脂肪二醇、C₁₀ 1,3脂肪二醇、C₁₁ 1,3脂肪二醇、C₁₂ 1,3脂肪二醇、C₁₃ 1,3脂肪二醇、C₁₄ 1,3脂肪二醇、C₁₅ 1,3脂肪二醇、C₁₆ 1,3脂肪二醇、C₁₇ 1,3脂肪二醇、C₁₈ 1,3脂肪二醇、C₁₉ 1,3脂肪二醇等的1,3脂肪二醇。在一个实施方案中,核酸序列是外源性的。在另一个实施方案中,核酸序列包括一个或多个核酸序列。

[0024] 本公开进一步涵盖一种产生1,3脂肪二醇的方法,包括如上所述的微生物(上文)。

[0025] 本公开的另一方面提供一种产生1,3脂肪二醇的方法,其包括在发酵肉汤中提供重组微生物,所述微生物表达编码具有硫酯酶(EC 3.1.2.-、EC 3.1.1.5或EC 3.1.2.14)活性、羧酸还原酶(EC 6.2.1.3或EC 1.2.1.42)活性和任选醇脱氢酶(EC 1.1.1.)活性的多肽的一个或多个核酸序列;以及从所述发酵肉汤分离1,3脂肪二醇。在一个实施方案中,所述方法进一步包括添加简单碳源至发酵肉汤。在又一个实施方案中,简单碳源来源于可再生原料。在另一方面,本公开提供一种产生1,3脂肪二醇的方法,其包括在发酵肉汤中提供重组微生物,所述微生物经工程化以表达编码具有硫酯酶(EC 3.1.2.-、EC 3.1.1.5或EC 3.1.2.14)活性和羧酸还原酶(EC 6.2.1.3或EC 1.2.1.42)活性的多肽的一个或多个核酸序列;以及从所述发酵肉汤分离1,3脂肪二醇。在一方面,1,3脂肪二醇包括但不限于C₅ 1,3脂肪二醇、C₆ 1,3脂肪二醇、C₇ 1,3脂肪二醇、C₈ 1,3脂肪二醇、C₉ 1,3脂肪二醇、C₁₀ 1,3脂肪二醇、C₁₁ 1,3脂肪二醇、C₁₂ 1,3脂肪二醇、C₁₃ 1,3脂肪二醇、C₁₄ 1,3脂肪二醇、C₁₅ 1,3

脂肪二醇、C₁₆ 1,3脂肪二醇、C₁₇ 1,3脂肪二醇、C₁₈ 1,3脂肪二醇以及C₁₉ 1,3脂肪二醇。在一个实施方案中,所述方法进一步包括添加简单碳源至发酵肉汤。在又一个实施方案中,简单碳源来源于可再生原料。

[0026] 本公开的另一方面提供一种重组微生物,其在具有简单碳源的发酵肉汤中生长时产生1,3脂肪二醇,所述微生物表达编码具有酰基-ACP还原酶(EC 1.2.1.80或EC 1.2.1.42)活性的多肽的核酸序列。在一方面,1,3脂肪二醇在体内产生。在另一方面,脂肪二醇包括但不限于C₅ 1,3脂肪二醇、C₆ 1,3脂肪二醇、C₇ 1,3脂肪二醇、C₈ 1,3脂肪二醇、C₉ 1,3脂肪二醇、C₁₀ 1,3脂肪二醇、C₁₁ 1,3脂肪二醇、C₁₂ 1,3脂肪二醇、C₁₃ 1,3脂肪二醇、C₁₄ 1,3脂肪二醇、C₁₅ 1,3脂肪二醇、C₁₆ 1,3脂肪二醇、C₁₇ 1,3脂肪二醇、C₁₈ 1,3脂肪二醇以及C₁₉ 1,3脂肪二醇。在一个实施方案中,核酸序列是外源性的。

[0027] 本公开的又一方面提供一种重组微生物,其在具有简单碳源的发酵肉汤中生长时产生1,3脂肪二醇,所述微生物表达编码具有酰基-ACP还原酶(EC 1.2.1.80或EC 1.2.1.42)活性和醇脱氢酶(EC 1.1.1.-)活性的多肽的一个或多个核酸序列。在一方面,1,3脂肪二醇在体内产生。在另一方面,脂肪二醇包括但不限于C₅ 1,3脂肪二醇、C₆ 1,3脂肪二醇、C₇ 1,3脂肪二醇、C₈ 1,3脂肪二醇、C₉ 1,3脂肪二醇、C₁₀ 1,3脂肪二醇、C₁₁ 1,3脂肪二醇、C₁₂ 1,3脂肪二醇、C₁₃ 1,3脂肪二醇、C₁₄ 1,3脂肪二醇、C₁₅ 1,3脂肪二醇、C₁₆ 1,3脂肪二醇、C₁₇ 1,3脂肪二醇、C₁₈ 1,3脂肪二醇以及C₁₉ 1,3脂肪二醇。在一个实施方案中,所述一个或多个核酸序列是外源性的。

[0028] 本公开进一步涵盖一种细胞培养物,其包括当在具有简单碳源的发酵肉汤中生长时产生1,3脂肪二醇的重组微生物,所述微生物经工程化以表达编码具有酰基-ACP还原酶(EC 1.2.1.80或EC 1.2.1.42)活性和任选醇脱氢酶(EC 1.1.1.-)活性的多肽的一个或多个核酸序列。在一方面,细胞培养物产生1,3脂肪二醇。在另一方面,脂肪二醇包括但不限于C₅ 1,3脂肪二醇、C₆ 1,3脂肪二醇、C₇ 1,3脂肪二醇、C₈ 1,3脂肪二醇、C₉ 1,3脂肪二醇、C₁₀ 1,3脂肪二醇、C₁₁ 1,3脂肪二醇、C₁₂ 1,3脂肪二醇、C₁₃ 1,3脂肪二醇、C₁₄ 1,3脂肪二醇、C₁₅ 1,3脂肪二醇、C₁₆ 1,3脂肪二醇、C₁₇ 1,3脂肪二醇、C₁₈ 1,3脂肪二醇以及C₁₉ 1,3脂肪二醇。在一个实施方案中,所述一个或多个核酸序列是外源性的。

[0029] 在又一方面,本公开涵盖一种产生1,3脂肪二醇的方法,其包括在发酵肉汤中提供重组微生物,所述微生物经工程化以表达编码具有酰基-ACP还原酶(EC 1.2.1.80或EC 1.2.1.42)活性的多肽的核酸序列;以及从所述发酵肉汤分离1,3脂肪二醇。在一个实施方案中,微生物进一步表达编码具有醇脱氢酶(EC 1.1.1.-)活性的多肽的核酸序列。在另一个实施方案中,所述方法进一步包括添加简单碳源至发酵肉汤。在又一个实施方案中,简单碳源来源于可再生原料。所述方法产生脂肪二醇,包括但不限于C₅ 1,3脂肪二醇、C₆ 1,3脂肪二醇、C₇ 1,3脂肪二醇、C₈ 1,3脂肪二醇、C₉ 1,3脂肪二醇、C₁₀ 1,3脂肪二醇、C₁₁ 1,3脂肪二醇、C₁₂ 1,3脂肪二醇、C₁₃ 1,3脂肪二醇、C₁₄ 1,3脂肪二醇、C₁₅ 1,3脂肪二醇、C₁₆ 1,3脂肪二醇、C₁₇ 1,3脂肪二醇、C₁₈ 1,3脂肪二醇以及C₁₉ 1,3脂肪二醇。

[0030] 在又一方面,本公开进一步涵盖1,3脂肪二醇从以上论述的任何重组微生物(上文)分泌和回收。在一个实施方案中,1,3脂肪二醇分泌至发酵肉汤中。在另一个实施方案中,1,3脂肪二醇经由油水分离,例如经由重力沉降、离心、倾析等等回收。

[0031] 本公开进一步涵盖脂肪二醇组合物。在一方面,所述组合物包括一种或多种脂肪

二醇,包括1,3-二醇。

[0032] 本公开的另一方面更提供脂肪二醇在产生包括乙氧基化物等表面活性剂中的用途。

[0033] 本公开进一步涵盖手性1,3脂肪二醇、其对映异构体和手性混合物。进一步涵盖1,3脂肪二醇、其对映异构体和手性混合物的组合物。

[0034] 附图简述

[0035] 当结合附图阅读时,更好地理解本公开,附图用于说明一些优选实施方案。然而,应了解,本公开不局限于图中公开的特定实施方案。

[0036] 图1描绘一种用于制备1,3-二醇的示例性途径,包括酶功能性。

[0037] 图2描绘一种用于制备1,3-二醇的示例性途径,出于说明的目的,提供酶功能性的实例。

[0038] 图3展示一种用于制备1,3-二醇的替代途径,包括酶功能性。

[0039] 图4展示来自表达TE_EEI82564和CarB的重组大肠埃希氏菌菌株的提取物的GC/MS色谱。所有样品都用BSTFA+1%TMCS衍生化。峰(1)是衍生化的1,3-辛醇且峰(2)是衍生化的1,3-癸醇。

[0040] 图5展示来自图4的衍生化峰1和峰2的质谱,来源于表达TE_EEI82564和CarB的重组大肠埃希氏菌菌株。衍生化剂是BSTFA+1%TMCS。

[0041] 图6展示经BSTFA+1%TMCS衍生化的1,3-癸二醇的离子碎裂图案。

[0042] 图7展示由表达TE_CAD63310和CarB的重组大肠埃希氏菌菌株产生的1,3-二醇(二醇)和脂肪醇(FALC)的组成。

[0043] 以下缩写用于图8-11中:

[0044] FAS-脂肪酸生物合成/脂肪酸合酶

[0045] TE-硫酯酶

[0046] ACS-酰基CoA合酶

[0047] TL-3-酮酰基CoA硫解酶(可逆)

[0048] (S) 3HACS- (S) -3-羟基-酰基CoA脱氢酶(可逆)

[0049] (S) 2ECO- (S) -2-烯酰基CoA水合酶/ (S) -3-羟基酰基CoA脱水酶

[0050] CAR-羧酸还原酶

[0051] FAR-脂肪酰基CoA/ACP还原酶和形成脂肪醇的脂肪酰基CoA/ACP还原酶

[0052] ACR-酰基CoA还原酶

[0053] AAR-酰基ACP/CoA还原酶

[0054] 图8描绘从酰基-ACP产生1,3脂肪二醇的生化途径。途径1使用例如TE、CAR和ADH等酶功能性产生1,3-二醇。途径2使用TE、ACS、ACR和ADH产生1,3-二醇。途径3使用AAR和ADH产生1,3-二醇。途径4使用FAR和ADH产生1,3-二醇。途径5使用FAR产生1,3-二醇。

[0055] 图9描绘从酰基-CoA产生1,3脂肪二醇的生化途径。途径1使用例如TE、CAR和ADH等酶功能性产生1,3-二醇。途径2使用ACR和ADH产生1,3-二醇。途径3使用AAR和ADH产生1,3-二醇。途径4使用FAR和ADH产生1,3-二醇。途径5使用FAR产生1,3-二醇。

[0056] 图10展示(R)-1,3脂肪二醇产生。途径1使用例如TE、CAR和ADH等酶功能性产生右手手性的1,3-二醇。途径2使用TE、ACR和ADH产生右手手性的1,3-二醇。途径3使用AAR和ADH

产生右手手性的1,3-二醇。途径4使用FAR和ADH产生右手手性的1,3-二醇。途径5使用FAR产生右手手性的1,3-二醇。

[0057] 图11展示(S)-1,3脂肪二醇产生。途径1使用例如TE、CAR和ADH等酶功能性产生左手手性的1,3-二醇。途径2使用ACR和ADH产生左手手性的1,3-二醇。途径3使用AAR和ADH产生左手手性的1,3-二醇。途径4使用FAR和ADH产生左手手性的1,3-二醇。途径5使用FAR产生左手手性的1,3-二醇。途径6使用TE、ACS、FadE和(S) 2ECOH产生左手手性的1,3-二醇。途径7使用脂肪酸和ACS产生左手手性的1,3-二醇。途径8使用TE、ACS、TL和(S) 3HACS产生左手手性的1,3-二醇。

[0058] 发明详述

[0059] 综述

[0060] 一种新颖环保的产生脂肪二醇的方法的研发改善了此行业。所述方法容许从来源于可再生原料的简单碳源产生脂肪二醇,包括但不限于来自玉米、藤条、天然气或木质纤维素生物质的碳水化合物;废产物,例如城市固体废物、甘油、烟气、合成气、二氧化碳;或由例如生物质、天然气或其它含碳物质等有机物质的重整所产生的碳流。所述方法进一步容许通过光合生物,例如蓝藻细菌和藻类,从CO₂和光产生脂肪二醇。此方法对于环境来说更佳,因为其不产生石油化学来源工艺所产生的毒性副产物。

[0061] 更具体地说,本公开提供经工程化以将来源于可再生原料的简单碳源转化成脂肪二醇的重组微生物。1,3-二醇是作为无色无味的稳定化学实体的脂肪二醇的实例。预期微生物产生的1,3-二醇有许多工业应用,包括作为清洁剂、表面活性剂、乳化剂、软化剂、溶剂、塑料、香精、香料和生物活性化合物的组分。微生物产生的1,3-二醇还作为天然食物的代替物(或添加剂)用于食品工业,因为其容易代谢、无毒、不挥发,且能量密集,保存期长。

[0062] 本公开的重组微生物用于产生脂肪二醇的发酵过程中。本文中,本公开涵盖微生物脂肪酸代谢及其中间物至1,3-二醇的转变。本公开的一个优点是生产方法更洁净,即采用简单发酵法。可再生原料的使用保护环境,因为其依赖于不消耗自然资源的可再生且可持续的原材料。工业废物(例如甘油)用作原料支持更好的废物管理和再循环。另一个优点是可制造新颖的工业目标产品,即具有选择性链长、手性且呈特定混合物或与衍生物组合的脂肪二醇组合物。

[0063] 定义

[0064] 如本文所用,术语“1,3脂肪二醇”或“1,3-二醇(1,3-diol)”或“1,3-二醇(1,3-dialcohol)”或“3-OH脂肪醇”或“3-羟基脂肪醇”或“1,3-二羟醇”或“1,3-脂肪族二醇”在本文中可互换使用,且是指具有至少5个碳的链长且经由脂肪酰基硫酯中间物从微生物脂肪酸代谢产生,并具有至少两个OH基团,即碳链1位的OH基团和3位的OH基团的化学实体。

[0065] 如本文提及的“1,3-二醇”由重组微生物或重组微生物宿主细胞产生。

[0066] “1,3-二醇组合物”典型地至少包括与另一成分组合的1,3-二醇。

[0067] 术语“酶分类(EC)号”是指表示特定多肽序列或酶的编号。EC号根据酶催化的反应将酶分类。EC号由国际生物化学与分子生物学联合会命名委员会(the nomenclature committee of the international union of biochemistry and molecular biology, IUBMB)建立,其描述可在万维网上IUBMB酶命名网址获得。

[0068] 术语“硫酯酶”是指特征为EC号3.1.2.14.或EC号3.1.1.5或EC号3.1.2.-的酶活

性。

[0069] 术语“羧酸还原酶 (CAR)”是指特征为EC号6.2.1.3或EC号1.2.1.42或EC号1.2.99.6的酶活性。

[0070] 术语“醛还原酶”与“醇脱氢酶”在本文中可互换使用且是指特征为EC号1.1.-.-.的酶活性。

[0071] 术语“酰基-ACP还原酶 (AAR)”是指特征为EC号1.2.1.80或EC号1.2.1.42的酶活性。

[0072] 术语“乙酰基-CoA羧化酶”是指特征为EC号6.4.1.2的酶活性。

[0073] 术语“登录号”和“NCBI登录号”和“GenBank登录号”在本文中可互换使用且是指表示特定核酸序列的编号。此描述中论述的序列登录号是从美国国立卫生研究院 (the National Institutes of Health, U.S.A.) 维护的NCBI (美国国家生物技术信息中心 (National Center for Biotechnology Information)) 提供的数据库获得, 以及从瑞士生物信息学研究所 (the Swiss Institute of Bioinformatics) 提供的the UniProt Knowledgebase (UniProtKB) 和Swiss-Prot数据库 (又称为UniProtKB登录号) 获得。

[0074] 如本文所用, 术语“核苷酸”是指由杂环碱基、糖和一个或多个磷酸酯基组成的多核苷酸的单体单元。天然存在的碱基 (鸟嘌呤 (G)、腺嘌呤 (A)、胞嘧啶 (C)、胸腺嘧啶 (T) 和尿嘧啶 (U)) 典型地是嘌呤或嘧啶的衍生物, 不过应了解也包括天然与非天然存在的碱基类似物。天然存在的糖是戊糖 (五碳糖) 脱氧核糖 (其形成DNA) 或核糖 (其形成RNA), 不过应了解也包括天然与非天然存在的糖类似物。核酸典型地经由磷酸键连接以形成核酸或多核苷酸, 不过本领域中已知许多其它的键 (例如硫代磷酸酯、硼烷磷酸酯等等)。

[0075] 如本文所用, 术语“多核苷酸”是指核苷酸 (RNA) 或脱氧核糖核苷酸 (DNA) 的聚合物, 其可以是单链或双链, 且其可以含有非天然或改变的核苷酸。术语“多核苷酸”、“核酸序列”和“核苷酸序列”在本文中可互换使用, 是指任何长度的核苷酸 (RNA或DNA) 的聚合形式。这些术语是指分子的一级结构, 因此包括双链和单链DNA以及双链和单链RNA。所述术语包括由核苷酸类似物制成的RNA或DNA的类似物以及例如经修饰的多核苷酸作为同等物, 不过不限于甲基化和/或封端多核苷酸。多核苷酸可呈任何形式, 包括但不限于质粒、病毒、染色体、EST、cDNA、mRNA和rRNA。

[0076] 术语“内源性多核苷酸”和“内源性DNA”和“内源性核酸序列”在本文中可互换使用且是指在宿主细胞内产生的DNA。

[0077] 术语“外源性多核苷酸”和“外源性DNA”和“外源性核酸序列”在本文中可互换使用且是指在宿主细胞外产生的DNA。举例来说, 来自宿主细胞A的基因可以插入宿主细胞B中。然而, 源自宿主细胞A的基因可以操纵或修饰 (在宿主细胞A内或外) 并再插入同一宿主细胞A中。

[0078] 术语“经修饰的多核苷酸”和“经修饰的DNA”和“经修饰的核酸序列”在本文中可互换使用且是指相对于原始或天然状态, 已在一定形式上改变的DNA。此改变可影响DNA或其编码基因产物 (例如多肽或蛋白质) 的稳定性、表达、活性或功能。在一个实施方案中, 编码多肽的表达增加。在另一个实施方案中, 编码多肽的表达减少。在另一个实施方案中, 编码多肽的表达缺乏。

[0079] 如本文所用, 术语“多肽”和“蛋白质”和“多肽序列”和“蛋白质序列”在本文中可互

换使用且是指氨基酸残基的聚合物。术语“重组多肽”是指通过重组技术产生的多肽，其中一般编码所表达的蛋白质的DNA或RNA插入适合的表达载体中，随后表达载体用于转化宿主细胞，以产生所述多肽。

[0080] 如本文所用，术语“同源性”和“同源”是指包含与对应多核苷酸或多肽序列至少约50%一致的序列的多核苷酸或多肽。优选地，同源多核苷酸或多肽具有与对应氨基酸序列或多核苷酸序列至少约70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或至少约99%同源的多核苷酸序列或氨基酸序列。如本文所用，术语序列“同源性”与序列“一致性”可互换使用。本领域技术人员知道确定两个或更多个序列之间的同源性。简单地说，可以如下计算两个序列之间的“同源性”。为了最佳比较，将序列比对（例如间隙可以引入用于最佳比对的第一和第二氨基酸或核苷酸序列中的一者或两者中，且为了比较，可以忽略非同源序列）。在一优选实施方案中，为了比较而进行比对的第一序列的长度是第二序列长度的至少约30%、优选至少约40%、更优选至少约50%、甚至更优选至少约60%和甚至更优选至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%或约100%。接着比较第一和第二序列的对应氨基酸位置或核苷酸位置的氨基酸残基或核苷酸。当第一序列中的一个位置被与第二序列中对应位置相同的氨基酸残基或核苷酸占据时，分子在该位置一致。两个序列之间的同源性百分比是所述序列共享的一致位置数目的函数，考虑到为最佳比对两个序列需要引入的间隙数目和每个间隙的长度。两个序列之间的序列比较和同源性百分比的测定可以使用例如BLAST等数学算法完成（Altschul等人（1990）*J. Mol. Biol.* 215 (3) :403-410）。两个氨基酸序列之间的同源性百分比也可以使用Needleman和Wunsch算法测定，该算法已经并入GCG软件包中的GAP程序中，使用Blossum 62矩阵或PAM 250矩阵和16、14、12、10、8、6或4的间隙权重和1、2、3、4、5或6的长度权重（Needleman和Wunsch, (1970) *J. Mol. Biol.* 48:444-453）。两个核苷酸序列之间的同源性百分比也可以使用GCG软件包中的GAP程序测定，使用NWSgapdna.CMP矩阵和40、50、60、70或80的间隙权重和1、2、3、4、5或6的长度权重。本领域技术人员可以执行初始同源性计算并相应地校准算法参数。优选的一组参数（以及在从业者不确定哪些参数应该用来确定分子是否在权利要求书的同源性限制内的情况下应使用的参数）是Blossum 62评分矩阵，间隙罚分为12，间隙延伸罚分为4，且框移间隙罚分为5。生物技术领域中已知序列比对的其它方法（参见例如Rosenberg (2005) *BMC Bioinformatics* 6:278) ;Altschul等人 (2005) *FEBS J.* 272 (20) :5101-5109)。

[0081] “内源性”多肽是指由重组细胞自其中工程化或衍生化的宿主细胞（例如亲本微生物细胞）的基因组编码的多肽。

[0082] “外源性”多肽是指并非由亲本或宿主微生物细胞（例如宿主细胞）的基因组原始编码的多肽。变异（即突变）多肽是外源性多肽的一实例。外源性多肽的另一实例是在天然细胞中存在，但为例如外源性多核苷酸表达等改变表达的结果的蛋白质（例如含有与原生基因一致但经工程化以在宿主细胞中过度表达的基因的载体或质粒；此类基因可任选地插入宿主DNA中）。

[0083] 术语“异源”一般意指来源于不同物种或来源于不同生物体或来源于不同来源。如本文所用，其是指非天然存在于特定生物体中的核苷酸序列或多肽序列。异源表达意指蛋

白质或多肽在通常不表达该蛋白质的细胞中表达。因而,异源意指转移蛋白质初始来源于与接受者不同的细胞类型或不同的物种或不同的来源。举例来说,植物细胞内源性的多核苷酸序列可以通过重组方法引入细菌宿主细胞,那么植物多核苷酸在重组细菌宿主细胞中是异源多核苷酸。外源性多肽的另一实例是在天然细胞中存在,但为例如异源多核苷酸表达等改变表达的结果的蛋白质(例如含有与原生基因一致但经工程化以在宿主细胞中过度表达的基因的载体或质粒;此类基因可任选地插入宿主DNA中)。

[0084] 如本文所用,术语多肽的“片段”是指尺寸在四个氨基酸残基至整个氨基酸序列减去一个氨基酸残基范围内的全长多肽或蛋白质的较短部分。在本公开的某些实施方案中,片段是指多肽或蛋白质的一结构域或整个氨基酸序列(例如底物结合域或催化结构域)。

[0085] 如本文所用,术语“突变诱发”是指使生物体的遗传信息以稳定的方式改变的一种方法。编码核酸序列的蛋白质的突变诱发产生突变蛋白质。突变诱发也指引起蛋白质活性改变的非编码核酸序列的变化。

[0086] 如本文所用,术语“基因”是指编码RNA产物或蛋白质产物的核酸序列,以及影响RNA或蛋白质表达的可操作地连接的核酸序列(例如此类序列包括但不限于启动子或强化子序列)或影响RNA或蛋白质表达的可操作地连接的核酸序列编码序列(例如此类序列包括但不限于核糖体结合位点或翻译控制序列)。

[0087] 本领域中已知表达控制序列且包括例如启动子、强化子、聚腺苷酸化信号、转录终止子、内部核糖体进入位点(IRES)等,其提供多核苷酸序列在宿主细胞中的表达。表达控制序列特异性地与参与转录的细胞蛋白质相互作用(Maniatis等人(1987) *Science* 236: 1237-1245)。示例性表达控制序列描述于例如Goeddel, *Gene Expression Technology: Methods in Enzymology*, 第185卷, Academic Press, San Diego, Calif. (1990) 中。

[0088] 术语“多个”是指数目为至少2个(例如多个多核苷酸序列意指至少两个多核苷酸序列)。

[0089] 在本公开的方法中,表达控制序列可操作地连接于多核苷酸序列。“可操作地连接”意指当适当分子(例如转录活化因子蛋白)结合于表达控制序列时多核苷酸序列和表达控制序列以允许基因表达的方式连接。根据转录和解释的方向,可操作地连接的启动子位于所选多核苷酸序列的上游。可操作地连接的强化子可以位于所选多核苷酸的上游、内部或下游。

[0090] 如本文所用,术语“载体”是指能够输送其连接的另一核酸(即多核苷酸序列)的核酸分子。一种类型适用载体是游离体(即能够染色体外复制的核酸)。适用的载体是能够自主复制和/或表达其连接的核酸的载体。能够指导其可操作地连接的基因的表达的载体在本文中称为“表达载体”。一般说来,重组DNA技术中应用的表达载体经常呈“质粒”形式,其一般是指呈载体形式,不结合于染色体的环状双链DNA环。术语“质粒”与“载体”在本文中可互换使用,因为质粒是载体最常用的形式。然而,还包括功能相等且随后在本领域中已知的表达载体的此类其它形式。在一些实施方案中,重组载体进一步包含可操作地连接于多核苷酸序列的启动子。在一些实施方案中,启动子是发育调节性、细胞器特异性、组织特异性、诱导性、组成性或细胞特异性启动子。重组载体典型地包含至少一种序列,包括(a)可操作地联接于多核苷酸序列的表达控制序列;(b)可操作地联接于多核苷酸序列的选择标记物;(c)可操作地联接于多核苷酸序列的标记物序列;(d)可操作地联接于多核苷酸序列的纯化

部分；(e)可操作地联接于多核苷酸序列的分泌序列；以及(f)可操作地联接于多核苷酸序列的靶向序列。在某些实施方案中，核苷酸序列稳定地并入宿主细胞的基因组DNA中，且核苷酸序列的表达在调节启动子区域的控制下。本文中的表达载体包括本文中描述的多核苷酸序列，其呈适于该多核苷酸序列在宿主细胞中表达的形式。本领域技术人员应了解，表达载体的设计可取决于例如待转化的宿主细胞的选择、所需多肽的表达水平等因素。本文中描述的表达载体可以引入宿主细胞中以产生由如本文中描述的多核苷酸序列编码的多肽，包括融合多肽。例如大肠埃希氏菌等原核生物中基因编码多肽的表达最常用含有指导融合或非融合多肽表达的组成性或诱导性启动子的载体进行。融合载体添加许多氨基酸至在其中编码的多肽，通常添加至重组多肽的氨基或羧基端。此类融合载体典型地用于达成以下三个目的中的一个或多个：(1)增加重组多肽的表达；(2)增加重组多肽的可溶性；以及(3)通过充当亲和力纯化中的配体，帮助纯化重组多肽。经常，在融合表达载体中，蛋白裂解位在融合部分与重组多肽的接合处引入。这能够在纯化融合多肽后将重组多肽与融合部分分离。在某些实施方案中，本公开的多核苷酸序列可操作地连接于来源于噬菌体T5的启动子。在某些实施方案中，宿主细胞是酵母细胞，且表达载体是酵母表达载体。在酵母酿酒酵母(*S. cerevisiae*)中表达的载体的实例包括pYepSec1 (Baldari等人(1987) *EMBO J.* 6:229-234)、pMFa (Kurjan等人(1982) *Cell* 30:933-943)、pJRY88 (Schultz等人(1987) *Gene* 54:113-123)、pYES2 (Invitrogen Corp., San Diego, CA) 和picZ (Invitrogen Corp., San Diego, CA)。在其它实施方案中，宿主细胞是昆虫细胞，且表达载体是杆状病毒表达载体。可用于在培养的昆虫细胞(例如Sf9细胞)中表达蛋白质的杆状病毒载体包括例如pAc系列 (Smith等人(1983) *Mol. Cell Biol.* 3:2156-2165) 和pVL系列 (Lucklow等人(1989) *Virology* 170:31-39)。在又一个实施方案中，本文中描述的多核苷酸序列可使用哺乳动物表达载体在哺乳动物细胞中表达。原核与真核细胞的其它适合表达系统为本领域中所熟知；参见例如Sambrook等人，“*Molecular Cloning: A Laboratory Manual*,” 第二版, Cold Spring Harbor Laboratory, (1989)。

[0091] 如本文所用，“酰基-CoA”是指在烷基链的羰基碳与辅酶A (CoA) 的4'-磷酸泛亚硫酰基部分的巯基之间形成的酰基硫酯，其具有式R-C(O)S-CoA，其中R是具有至少4个碳原子的任何烷基。

[0092] 如本文所用，“酰基-ACP”是指在烷基链的羰基碳与酰基载体蛋白 (ACP) 的磷酸泛酰巯基乙氨基部分的巯基之间形成的酰基硫酯。磷酸泛酰巯基乙氨基部分在全酰基载体蛋白合酶 (ACPS，一种使用辅酶A作为底物和磷酸泛酰巯基乙氨基供体的磷酸泛酰巯基乙氨基转移酶) 的作用下在翻译后附接于ACP上的保守丝氨酸残基。在一些实施方案中，酰基-ACP是完全饱和酰基-ACP的合成中的中间物。在其它实施方案中，酰基-ACP是不饱和酰基-ACP的合成中的中间物。在一些实施方案中，碳链具有约5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25或26个碳。这些酰基-ACP中的每一者是将其转变成脂肪酸衍生物的酶的底物。如技术人员显而易见，全ACP的4'-磷酸泛亚硫酰基部分来源于辅酶A。因此，利用酰基ACP作为底物的酶经常具有一些用于酰基CoA的活性，且利用酰基CoA作为底物的酶具有一些用于酰基ACP的活性。

[0093] 如本文所用，术语“脂肪酸生物合成途径”意指产生脂肪酸、脂肪酸硫酯和/或其衍生物的生物合成途径。脂肪酸生物合成途径可包括具有除本文中论述的酶活性外的酶活性

以产生具有所需特征的脂肪酸衍生物的其它酶或多肽。

[0094] 如本文所用,术语“克隆”典型地是指遗传自单个共同祖先以及基本上与单个共同祖先一致的细胞或一组细胞,例如由单个细菌细胞产生的克隆菌落的细菌。

[0095] 如本文所用,术语“培养物”典型地是指包含活细胞的液体培养基。在一个实施方案中,培养物包含在预先确定的培养基中在控制条件下复制的细胞,例如在包含选择碳源和氮的液体培养基中生长的重组宿主细胞的培养物。“培养(Culturing)”或“培养(cultivation)”是指重组宿主细胞群体在液体或固体培养基中在适合条件下生长。在特定实施方案中,培养是指底物经发酵而生物转变成最终产物。培养基为众所周知的,且此类培养基的个别组分可获自商业来源,例如Difco™和BBL™商标。在一个非限制性实例中,水性培养基是包含氮、盐和碳的复杂来源的“丰富培养基”,例如YP培养基,包含每升此类培养基10g胨和10g酵母提取物。培养物的宿主细胞可另外根据美国专利5,000,000、5,028,539、5,424,202、5,482,846、5,602,030、W02010127318中描述的方法经工程化以有效吸收碳且使用纤维素材料作为碳源。另外,在一些实施方案中,宿主细胞经工程化以表达转化酶以便蔗糖可用作碳源。

[0096] 如本文所用,术语“在有效表达基因工程化的多核苷酸序列的条件下”意指容许宿主细胞表达对应酶功能性以产生例如脂肪二醇等所需脂肪酸衍生物的任何条件。适合条件包括例如发酵条件。

[0097] 术语“重组微生物”是指已经基因修饰或工程化,以使得相对于亲本细胞或原生宿主细胞,宿主细胞内的某些酶活性已经改变、添加和/或缺失的宿主细胞。基因修饰或基因工程化的宿主细胞是重组微生物的一个实例。因而,重组宿主细胞中例如酶等“蛋白质的活性水平变化或改变”是指相对于缺乏相同修饰的亲本或原生宿主细胞,所测定的活性的一个或多个特征的差异。典型地,在具有变化活性的重组宿主细胞与不具有该变化活性的对应野生型宿主细胞之间测定活性差异(例如相对于对应野生型宿主细胞,比较重组宿主细胞的培养物)。变化活性可以是例如以下各者的结果:重组宿主细胞表达的蛋白质的变化量(例如由于编码蛋白质的DNA序列的拷贝数增加或减少、编码蛋白质的mRNA转录物的数目增加或减少和/或蛋白质从mRNA的蛋白质翻译的量增加或减少);蛋白质结构的变化(例如一级结构的变化,例如引起底物特异性的变化的蛋白质编码序列的变化、观测到的动力学参数的变化);以及蛋白质稳定性的变化(例如蛋白质降解增加或减少)。在一些实施方案中,多肽是本文中描述的任何多肽的突变体或变异体。在某些情况下,本文中描述的多肽的编码序列为在特定宿主细胞中表达而密码子最佳化。举例来说,为在大肠埃希氏菌中表达,一个或多个密码子可以最佳化(如例如Grosjean等人(1982)Gene18:199-209中所述)。在一个实施方案中,重组微生物产生例如脂肪酸衍生物(例如脂肪酸、脂肪醛、脂肪醇、脂肪二醇)等所需产物。在一特定实施方案中,重组微生物产生1,3-二醇。

[0098] 如本文所用,术语“调节序列”典型地是指DNA中可操作地连接于编码蛋白质的DNA序列的最终控制蛋白质表达的碱基序列。调节序列的实例包括但不限于RNA启动子序列、转录因子结合序列、转录终止序列、转录调节子(例如强化子元件)、影响RNA稳定性的核苷酸序列以及翻译调节序列(例如核糖体结合位点(例如原核生物中夏因-达加诺尔序列(Shine-Dalgarno sequence)或真核生物中科扎克序列(Kozak sequence))、起始密码子、终止密码子)。

[0099] 术语“表达水平改变”和“表达水平变化”可互换使用且意指在相同条件下与对应野生型细胞中的浓度相比,多核苷酸、多肽、代谢物或产物(例如脂肪酸衍生物)以不同浓度存在于经工程化的宿主细胞。脂肪酸衍生物的实例是脂肪酸、3-羟基脂肪酸、脂肪醛、3-羟基脂肪醛、脂肪醇、1,3-脂肪二醇等等。

[0100] 如本文所用,术语“滴度”是指每单位体积宿主细胞培养物产生的脂肪酸衍生物,例如脂肪二醇(例如1,3-二醇)的量,且一般以质量体积单位报道,例如10g/L。在一个实施方案中,滴度可指既定重组宿主细胞培养物产生的特定1,3-二醇或1,3-二醇组合。在另一个实施方案中,滴度也可指既定重组宿主细胞培养物产生的脂肪二醇组合物(例如1,3-二醇组合物)。

[0101] 如本文所用,“宿主细胞产生的脂肪二醇(例如1,3-二醇)的产率”是指宿主细胞中输入碳源转变为产物(例如脂肪二醇)的效率,且在“质量产率”的情况下,以质量(产物)/质量(碳源)的百分比单位报道,例如30%质量产率将指30g产物由100g碳源产生;20%质量产率将指20g产物由100g碳源产生;10%质量产率将指10g产物由100g碳源产生等等。产率可指既定重组宿主细胞培养物产生的特定1,3-二醇或1,3-二醇组合。

[0102] 如本文所用,术语“生产率”是指每单位体积宿主细胞培养物产生的脂肪二醇(例如1,3-二醇)或衍生物的量(例如以g/L/小时报道)。生产率可指既定重组宿主细胞培养物产生的特定1,3-二醇或1,3-二醇组合。

[0103] 如本文所用,术语“葡萄糖利用率”意指每单位时间培养物使用的葡萄糖的量,以克/升/小时(g/L/小时)报道。

[0104] “原料”是用于制造产品或工业过程的原材料。“可再生原料”是来源于可再生材料,例如生物材料,例如植物物质,且可以由天然方式(例如玉米、藤条、木质纤维素生物质)或废产物(例如城市固体废物、甘油、游离脂肪酸、烟气或合成气;二氧化碳等等)替换的原材料。相比之下,“不可再生原料”是使用会消耗且不能再生的原材料(例如原油、煤、核燃料等等)。

[0105] 如本文所用,术语“简单碳源”是指适用作供原核或简单真核细胞生长的燃料来源的基质或化合物。有资格作为简单碳源的来源可以呈多种形式,包括但不限于聚合物、碳水化合物、酸、醇、醛、酮、氨基酸、肽和气体(例如CO和CO₂)。示例性简单碳源包括但不限于单糖,例如葡萄糖、果糖、甘露糖、半乳糖、木糖和阿拉伯糖;寡糖,例如寡聚果糖和寡聚半乳糖;聚糖,例如淀粉、纤维素、果胶和木聚糖;二糖,例如蔗糖、麦芽糖、纤维二糖和松二糖;纤维素材料和变体,例如半纤维素、甲基纤维素和羧甲基纤维素钠;饱和或不饱和脂肪酸、丁二酸酯、乳酸酯和乙酸酯;醇,例如乙醇、甲醇和丙醇;甘油,或其混合物。在一个实施方案中,简单碳源来源于玉米、甘蔗、高粱、甜菜、柳枝稷、青贮饲料、稻草、木材、纸浆、污水、垃圾、纤维素城市废物、烟气、合成气或二氧化碳。简单碳源还可以为光合作用产物,例如葡萄糖。在一个实施方案中,简单碳源来源于可再生原料。在一特定实施方案中,简单碳源来源于可再生材料,例如来自玉米、藤条或木质纤维素生物质的碳水化合物;或来自废产物,例如甘油、脂肪酸、烟气或合成气;或来自例如生物质等有机材料的重整;或来自光合作用固定的二氧化碳。在另一个实施方案中,简单碳源选自葡萄糖、果糖、甘露糖、半乳糖、木糖、阿拉伯糖、寡聚果糖、寡聚半乳糖、淀粉、纤维素、果胶、木聚糖、蔗糖、麦芽糖、纤维二糖、松二糖、半纤维素、甲基纤维素、羧甲基纤维素钠、丁二酸酯、乳酸酯、乙酸酯、乙醇、甲醇、甘油及

其混合物。在某些实施方案中,简单碳源来源于生物质。生物质的一个示例性来源是植物物质或植被,例如玉米、甘蔗或柳枝稷。生物质的另一个示例性来源是代谢废物,例如动物物质(例如牛粪)。生物质的其它示例性来源包括藻类及其它海生植物。生物质还包括来自工业、农业、林业和家庭的废产物,包括但不限于发酵废物、青贮饲料、稻草、木材、污水、垃圾、纤维素城市废物和吃剩的食物。术语“生物质”还指例如碳水化合物(例如单糖、二糖或聚糖)等碳源。

[0106] 如本文所用,关于产物(例如1,3-二醇或衍生物)的术语“分离”是指产物与细胞组分、细胞培养基或化学或合成前驱物分离。本文中描述的方法产生的脂肪二醇(例如1,3-二醇)和相关组合物可以相对不混合于发酵肉汤以及细胞质中。因此,脂肪二醇组合物可以在细胞内或细胞外集中在有机相中。在一个实施方案中,1,3-二醇组合物在细胞外集中。

[0107] 如本文所用,术语“纯化(purify、purified或purification)”意指通过例如分离(isolation)或分离(separation),将分子自其环境去除或分离。“基本上纯化”分子为至少约60%(例如至少约70%、至少约75%、至少约85%、至少约90%、至少约95%、至少约97%、至少约99%)不含其相关的其它组分。如本文所用,这些术语还指从样品去除污染物。举例来说,污染物的去除可以引起样品中脂肪二醇的百分比增加。举例来说,当1,3-二醇在重组宿主细胞中产生时,1,3-二醇通过去除宿主细胞蛋白质而纯化。纯化后,样品中1,3-二醇的百分比增加。术语“纯化(purify、purified和purification)”是无需绝对纯度的相对术语。因此,举例来说,当1,3-二醇在重组宿主细胞中产生时,纯化的1,3-二醇是基本上与其它细胞组分(例如核酸、多肽、脂质、碳水化合物或其它烃)分离的1,3-二醇。

[0108] 如为了本说明书和权利要求书而用的术语“在体内产生脂肪二醇(例如1,3-二醇)”意指在活和/或重组和/或基因修饰的宿主细胞中从简单碳源产生脂肪二醇,其中简单碳源添加至发酵肉汤中以便宿主细胞在发酵期间可吸收并代谢简单碳源。在一个实施方案中,简单碳源来源于可再生原料。

[0109] 将宿主细胞菌株工程化用于筛选

[0110] 脂肪酸生物合成是细菌生物合成机制的最保守系统之一。脂肪酸合酶(FAS)多酶复合物存在于所有细菌和真核生物中。大多数FAS相关基因是细胞生长和存活所需的。真核生物和细菌FAS驱动基本上相同类型的生物化学转化。在真核生物中,FAS称为FAS I,且其大部分催化结构域由一个多肽链(不可解离)编码。在例如细菌等原核生物中,FAS称为FASII,且其个别酶和载体蛋白由编码离散(可解离)蛋白质的单独基因编码。

[0111] 在FAS途径中酰基载体蛋白(ACP)与酶一起控制原生生物体中产生的脂肪酸的长度、饱和度和分支。此途径中的步骤由脂肪酸生物合成(FAB)和乙酰基-CoA羧化酶(ACC)基因家族的酶催化。举例来说,可以包括在工程化FAS途径的酶包括乙酰基-CoA羧化酶(例如AccABCD、丙二酰基-CoA:ACP酰基转移酶(例如FabD)、3-酮酰基-ACP合酶III(例如FabH)、3-酮酰基-ACP还原酶(例如FabG)、3-羟基酰基-ACP脱水酶/异构酶(例如FabA)、3-羟基酰基-ACP脱水酶(例如FabZ)、反式-2-烯酰基-ACP还原酶(例如FabI或fabL或fabK)、反式-2-烯酰基-ACP异构酶(例如FabM)、3-酮酰基-ACP合酶I(例如FabB)和3-酮酰基-ACP合酶II(例如FabF)。取决于所需产物,可减弱或过度表达这些基因中的一个或多个。因而,宿主细胞经工程化以通过馈入可以来源于可再生原料的简单碳源增加脂肪酸衍生物(例如脂肪二醇、脂肪醇)以及脂肪酸衍生物中间物(例如脂肪醛)的产生。本文中,主要目标是增加调节例如脂

肪二醇等脂肪酸衍生物产生的关键控制酶的活性,以将菌株转变成产生脂肪二醇的微生物工厂。在一个实施方案中,菌株产生脂肪二醇,例如1,3-二醇。在另一个实施方案中,菌株产生脂肪二醇,例如1,3-二醇,以及脂肪醇。在另一个实施方案中,菌株经进一步修饰,使得尤其酰基-ACP还原酶活性增加和/或3-羟基酰基-ACP脱水酶活性降低,使得1,3脂肪二醇产生增加。

[0112] 宿主细胞预先经工程化以增加其它脂肪酸衍生物,包括脂肪酸甲酯 (FAME)、脂肪酸乙酯 (FAEE) 和脂肪醇 (FALC) (参见例如美国专利No.8,283,143,以引用的方式并入本文中)。如本领域技术人员所了解,脂肪酸合成还可以通过使用非酰基ACP依赖性脂肪酸生物合成伸长酰基CoA来进行。负责对应脂肪酸生物合成反应、缩合、还原、脱水、还原等的FAS酶还可以用于合成酰基硫酯,其可以用作产生包括但不限于脂肪酸、脂肪醛、脂肪醇及其3-羟基衍生物(包括1,3脂肪二醇)在内的脂肪酸衍生物的底物。如本领域技术人员所知,负责脂肪酸氧化的生物化学反应(β -氧化循环)可以反过来起作用以支持脂肪酸硫酯的合成。这些酰基硫酯可以用作产生包括但不限于脂肪酸、脂肪醛、脂肪醇及其3-羟基衍生物(包括1,3脂肪二醇)在内的脂肪酸衍生物的底物。此外,在一些生物体中,脂肪酸生物合成可以在无ACP下发生,例如通过合成酰基CoA(参见例如美国专利申请公布No.US2014/0051136A1;美国专利申请公布No.US 2014/0273114A1;以及Dellomonaco等人(2011)Nature 476(7360):355-9)。在一个方面,这些多种且不同FAS系统的组分可以在相同细胞中共同表达,以协同工作,产生脂肪酰基硫酯和衍生物,包括但不限于脂肪酸、脂肪醛、脂肪醇及其3-羟基衍生物(包括1,3脂肪二醇)。

[0113] 手性分子

[0114] 如果分子可以呈彼此不可重叠镜象的立体异构体(即对映异构体)存在,那么称该分子是手性的。此具有重大意义,因为生物体对特定分子的反应经常取决于该分子与生物体中受体分子上特定位点配合的方式。包括手性醇和二醇在内的手性分子是合成例如药物、营养物及其它活性物质等某些化合物的构筑嵌段。在药物和营养物应用中,需要知道哪个对映异构体是活性对映异构体且与预期受体配合。

[0115] 一种获得呈纯活性异构体形式的化合物的方式是通过采用例如微生物等生物体产生该化学物质,因为生物体中生物分子的产生是立体特异性的(即其产生特异性立体异构体)。举例来说,氨基酸、维生素和激素是在糖发酵期间通过酵母天然产生且可以从其中收获。本领域技术人员了解作为手性催化剂的酶的特性,且对对映异构体纯的药物的需求增加燃起了对用于合成精细化学品的酶的兴趣。与经由生物体产生手性分子对比,当通过化学程序制备手性分子时,获得对映异构体的混合物(即外消旋混合物)。

[0116] 对映异构体分析的当前方法包括例如旋光测定法、核磁共振、同位素稀释、量热法和酶技术等非色谱技术。这些技术需要纯样品,且不分离对映异构体。对映异构体的定量(不需要纯样品)和分离可以通过手性色谱法,例如气相色谱法(GC)或高效液相色谱法(HPLC),使用手性柱同时进行(参见Stereochemistry of Organic Compounds, Ernest L. Eliel/Sanuel H. Wilen, 1994, John Wiley&Sons, Inc.)。生物催化剂可用于制备手性化合物且产物手性纯度可以使用例如手性HPLC或LC/MS等手性色谱法鉴别(参见美国专利申请公布No.US2008/0248539A1和US2013/0052699A1)。

[0117] 3-羟基脂肪酸衍生物的手性

[0118] 3-羟基脂肪酸衍生物(例如3-羟基脂肪酸、3-羟基脂肪酸酯、3-羟基脂肪醛、3-羟基脂肪醇等等)的一独特方面是每个分子都是手性的。3-羟基官能团是立体中心,提供每种化合物的手性点。手性可以是界定包括但不限于聚合物性能、生物活性、药物效力等分子应用的一种适用的分子属性。3-羟基脂肪酸衍生物的立体异构体取决于产生其的脂肪酸生物合成(FAS)的选择性。通过操纵哪些FAS酶负责3-羟基脂肪酸衍生物合成,可以控制所得3-羟基脂肪族衍生物的手性。举例来说,采用天然大肠埃希氏菌FAS用于1,3脂肪二醇生物合成将产生(R)-1,3脂肪二醇,其手性中心由大肠埃希氏菌中FabG催化(及其它微生物中的同源物)的形成(R)-3-羟基酰基ACP的3-酮酰基-ACP还原酶的活性建立。(R)-3-羟基酰基ACP是醇生物合成多肽的底物,包括但不限于图10中途径1-5中所示的多肽,将其转变成(R)-1,3脂肪二醇。此外,(S)-3-羟基酰基CoA是通过 β -氧化途径降解脂肪酸中的中间物。游离脂肪酸通过酰基-CoA合酶转变为酰基-CoA,由大肠埃希氏菌中FadD及其它微生物中的同源物催化;通过脂肪酰基-CoA脱氢酶氧化为反式-2-烯酰基-CoA,由大肠埃希氏菌中FadE及其它微生物中的同源物催化;接着通过2-反式-烯酰基-CoA水合酶/(S)-3-羟基-酰基-CoA脱水酶水合为(S)-3-羟基-酰基-CoA,由大肠埃希氏菌中FadB及其它微生物中的同源物催化;接着通过3-酮-酰基-CoA脱氢酶进一步氧化为3-酮-酰基-CoA,也由大肠埃希氏菌中FadB及其它微生物中的同源物催化;最后通过3-酮酰基-CoA硫解酶硫解为酰基-CoA和乙酰基-CoA,由大肠埃希氏菌中FadA及其它微生物中的同源物催化。例如通过大肠埃希氏菌FadB中组氨酸450突变(或不同微生物中或来自不同微生物的功能同源物),在 β -氧化的(S)-3-羟基-酰基-CoA脱氢酶活性中选择性分裂的菌株将在提供游离脂肪酸时积累(S)-3-羟基-酰基CoA(图11中的途径6和7)。组氨酸450是与来自大肠埃希氏菌的脂肪酸氧化的多酶复合物的 α -次单元有关的L-3-羟基酰基辅酶A脱氢酶的催化残基(参见He等人(1996) *Biochemistry* 35 (29):9625-9630)。(S)-3-羟基-酰基CoA接着可以通过形成脂肪醇的多肽(例如图11中途径1-5中描述的多肽)的作用转变为(S)-1,3脂肪二醇。游离脂肪酸可以提供至细胞表面(图11中途径7),或可例如通过酰基ACP由硫酯酶水解在细胞内产生(图11中途径6)。在一个实施方案中,以上反应中的酰基CoA中间物通过3-酮酰基-CoA硫解酶伸长为3-酮酰基-CoA(参见图11中途径8),由大肠埃希氏菌中FadA及其它微生物中的同源物催化;接着由FadB的突变体还原,突变体在其水合酶/脱水酶活性上选择性分裂(例如通过大肠埃希氏菌FadB(或相关酶中其同源物)中Glu 119突变)。此将引起(S)-3-羟基酰基-CoA积累,接着其通过形成脂肪二醇的多肽,例如图11中途径1-5中所示的多肽,转变为(S)-1,3脂肪二醇。大 α -次单元的谷氨酸-119是由来自大肠埃希氏菌的脂肪酸氧化的多酶复合物催化的2-反式-烯酰基-辅酶A的水合中的催化基础(参见He等人(1997) *Biochemistry* 36 (36):11044-11049)。大 α -次单元的谷氨酸-139是由来自大肠埃希氏菌的脂肪酸氧化的多酶复合物催化的D-与L-3-羟基酰基-辅酶A的脱水中而非 δ^3, δ^2 -烯酰基-辅酶A的异构化中的催化基础(参见Yang等人(1995) *Biochemistry* 34 (19):6441-6447)。在另一个实施方案中,以上反应中的酰基CoA中间物通过3-酮酰基-CoA硫解酶伸长为3-酮酰基CoA,由大肠埃希氏菌中FadA及其它微生物中的同源物催化;接着由(S)-3-羟基酰基-CoA脱氢酶(例如来自EC 1.1.1.35)还原(图11途径8)。此将引起(S)-3-羟基酰基CoA积累,接着其通过形成脂肪二醇的多肽,例如图11中途径1-5中所示的多肽,转变为(S)-1,3脂肪二醇。

[0119] 为将宿主细胞工程化以表达某些酶功能性(参见以下表1),可对宿主细胞进行基

因修饰。在一些实施方案中,多核苷酸(或基因)序列借助于包括可操作地连接于多核苷酸序列的启动子的重组载体提供给宿主细胞。在某些实施方案中,启动子是发育调节性、细胞器特异性、组织特异性、诱导性、组成性或细胞特异性启动子。在一些实施方案中,重组载体包括至少一种选自以下的序列:可操作地联接于多核苷酸序列的表达控制序列;可操作地联接于多核苷酸序列的选择标记物;可操作地联接于多核苷酸序列的标记物序列;可操作地联接于多核苷酸序列的纯化部分;可操作地联接于多核苷酸序列的分泌序列;以及可操作地联接于多核苷酸序列的靶向序列。本文中描述的表达载体包括呈适于多核苷酸序列在宿主细胞中表达的形式多核苷酸序列。本领域技术人员应了解,表达载体的设计可取决于例如待转化的宿主细胞的选择、所需多肽的表达水平等因素。本文中描述的表达载体可以引入宿主细胞中以产生由如上所述(上文)的多核苷酸序列编码的多肽,包括融合多肽。例如大肠埃希氏菌等原核生物中基因编码多肽的表达最常用含有指导融合或非融合多肽表达的组成性或诱导性启动子的载体进行。融合载体添加许多氨基酸至在其中编码的多肽,通常添加至重组多肽的氨基或羧基端。此类融合载体典型地用于达成包括以下的三个目的中的一个或多个:增加重组多肽的表达;增加重组多肽的可溶性;以及通过充当亲和力纯化中的配体,帮助纯化重组多肽。经常,在融合表达载体中,蛋白裂解位点在融合部分与重组多肽的接合处引入。这容许在纯化融合多肽后将重组多肽与融合部分分离。此类酶的实例及其同源识别序列包括因子Xa、凝血酶和肠激酶。示例性融合表达载体包括pGEX载体(Pharmacia Biotech, Inc., Piscataway, NJ; Smith等人(1988) Gene 67:31-40)、pMAL载体(New England Biolabs, Beverly, MA)和pRITS载体(Pharmacia Biotech, Inc., Piscataway, N.J.),其分别将谷胱甘肽S-转移酶(GST)、麦芽糖E结合蛋白或蛋白A与标靶重组多肽融合。

[0120] 诱导性非融合大肠埃希氏菌表达载体的实例包括pTrc载体(Amann等人(1988) Gene 69:301-315)和pET 11d载体(Studier等人, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, Calif. (1990) 60-89)。标靶基因从pTrc载体的表达依赖于宿主RNA聚合酶从杂交trp-lac融合启动子的转录。标靶基因从pET 11d载体的表达依赖于由共同表达的病毒RNA聚合酶介导的从T7 gn10-lac融合启动子(T7 gn1)的转录。此病毒聚合酶由例如BL21 (DE3) 或HMS174 (DE3) 等宿主菌株在lacUV 5启动子的转录控制下从具有T7 gn1基因的固有 λ 前噬菌体供应。原核与真核细胞的适合表达系统为本领域中所熟知(参见例如Sambrook等人(1989) Molecular Cloning: A Laboratory Manual, 第二版, Cold Spring Harbor Laboratory)。诱导性非融合大肠埃希氏菌表达载体的实例包括pTrc载体(Amann等人(1988) Gene 69:301-315)和PET 11d载体(Studier等人(1990) Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA, 第60-89页)。在某些实施方案中,本公开的多核苷酸序列可操作地连接于来源于噬菌体T5的启动子。在一个实施方案中,宿主细胞是酵母细胞。在此实施方案中,表达载体是酵母表达载体。载体可以经由多种领域公认的用于将外来核酸(例如DNA)引入宿主细胞中的技术引入原核或真核细胞。适合转化或转染宿主细胞的方法可见于例如Sambrook等人(上文)。为稳定转化细菌细胞,已知(取决于使用的表达载体和转化技术)某一部分的细胞将吸收并复制表达载体。为了鉴别和选择这些转化体,可以将编码选择性标记物(例如对抗生素的抗性)的基因与相关基因一起引入宿主细胞。选择性标记物包括赋予对例如但不限

于氨苄青霉素 (ampicillin)、卡那霉素 (kanamycin)、氯霉素 (chloramphenicol) 或四环素 (tetracycline) 的抗性的标记物。可以将编码选择性标记物的核酸在与编码本文中描述的多肽相同的载体上引入宿主细胞中或可以在分开载体上引入。可以通过在适当的选择药物存在下生长来鉴别经引入的核酸稳定转化的细胞。如本文中描述的工程化或重组宿主细胞是用于产生例如脂肪二酯组合物等脂肪酸衍生物组合物的细胞。在本文中描述的本公开的任何方面中, 宿主细胞可以选自真核生物植物、细菌、藻类、蓝细菌、绿色硫细菌、绿色非硫细菌、紫色硫细菌、紫色非硫细菌、嗜极菌、酵母、真菌、其工程化生物体或合成生物体。在一些实施方案中, 宿主细胞为光依赖性的或固定碳。在一些实施方案中, 宿主细胞具有自给营养活性。如本文中描述, 各种宿主细胞可用于产生脂肪二酯。

[0121] 本公开的宿主细胞或微生物包括可以基因工程化或修饰以含有变化, 从而测试特定酶活性的效率的宿主株系或宿主细胞。各种任选的遗传操纵和变化在宿主细胞之间可互换使用, 取决于何种原生酶途径存在于原始宿主细胞中。宿主菌株可以涵盖许多遗传变化, 以测试特定变量, 包括但不限于培养条件, 包括发酵组分、碳源 (例如原料)、温度、压力、减少的培养污染条件和氧含量。

[0122] 在一个实施方案中, 宿主菌株涵盖参与脂肪酸 β -氧化和/或噬菌体附着位点的一种或多种酶的任选的减弱或缺失。这些遗传修饰被设计成能在细胞内降解脂肪酸并增加对噬菌体的抗性。在一个实施方案中, 宿主菌株是大肠埃希氏菌, 且遗传修饰是 *fadE* 和/或 *fhuA* 的减弱或缺失。酰基-CoA 脱氢酶 (大肠埃希氏菌中的 *FadE*) 是一种对于代谢脂肪酸来说重要的酶。其催化脂肪酸降解中的第二步 (β -氧化), 其为脂肪酸硫酯代谢 (酰基-CoA) 成乙酰基-CoA 分子和 NAD (P) H 的过程。更具体地说, 细菌中脂肪酸降解的 β -氧化循环的第二步是酰基-CoA 氧化成 2-烯酰基-CoA, 由 *FadE* 催化。当大肠埃希氏菌或其它细菌缺乏或 *FadE* 或脂肪酰基 CoA 脱氢酶减弱时, 其几乎不能作为碳源在脂肪酸上生长。不能利用任何链长的脂肪酸符合所报道的 *fadE* 株系表现型, 即其中 *FadE* 功能被破坏的 *fadE* 突变株。*fadE* 基因可以任选地敲除或减弱以确保可作为脂肪酸衍生物途径中的中间物的酰基-CoA 可以积累在细胞中, 使得所有酰基-CoA 可以有效地转变为脂肪酸衍生物。然而, 在非限制性条件下当糖用作碳源时, *fadE* 减弱可为任选的, 因为在此类条件下 *FadE* 的表达可能受抑, 因此 *FadE* 可以仅少量存在, 而无法与酯合酶或用于酰基-CoA 底物的其它酶有效竞争。在这些情况下, 认为 *FadE* 由于分解代谢物抑制而受抑。大肠埃希氏菌和许多其它微生物更喜欢消耗糖而非脂肪酸, 因此当两个来源都可获得时, 由于 *fad* 调节子的抑制, 预期糖将首先被消耗 (参见 D. Clark, *J Bacteriol.* (1981) 148 (2) : 521-6)。此外, 糖的缺乏和脂肪酸的存在诱发 *FadE* 表达。酰基-CoA 中间物可不再属于 β -氧化途径, 因为 *fad* 调节子 (包括 *FadE*) 表达的蛋白质将上调并将有效竞争酰基-CoA。因此, 将 *fadE* 基因敲除或减弱可为有益的。因为碳源可基于糖, 所以任选地将 *FadE* 减弱。

[0123] 举例来说, 在大肠埃希氏菌中, *fadE* 基因 (编码酰基-CoA 脱氢酶) 或 *fadD* 基因 (编码酰基-CoA 合成酶) 可以缺失。此类菌株不能降解脂肪酸或只能很差地降解, 因此, 细胞内脂肪酸的可用性增加。接着此类脂肪酸可用于增加转变成例如脂肪酸衍生物等产物。脂肪酸还可以通过使例如 *fadA* 或 *fadB* 等其它脂肪酸降解酶缺失而可获得。这些基因中任一者的缺失是任选的, 且当游离脂肪酸是外来供应或是产物途径的中间物时可以执行。表1 (下文) 提供代谢途径内酶活性的综合列表, 包括可以减弱以增加宿主菌株中脂肪酸的可用性的各种

脂肪酸降解酶。

[0124] 大肠埃希氏菌中,基因fhuA编码TonA蛋白质,该蛋白质是大肠埃希氏菌外膜中的能量耦合转运体和受体(V.Braun(2009) J Bacteriol.191(11):3431-3436)。其缺失是任选的。fhuA缺失容许细胞变得对噬菌体攻击更具抗性,噬菌体攻击在商业发酵中可为有害的。因此,需要在发酵操作期间可能遭受潜在污染的宿主细胞中使fhuA缺失。类似地,其它生物体以及其它噬菌体附着位点中的同源蛋白质是为提高噬菌体抗性而进行缺失的潜在候选者。

[0125] 在另一个实施方案中,宿主菌株(上文)还涵盖包括fadR、fabA、fabD、fabG、fabH、fabV和/或fabF在内的一个或多个以下基因的任选的过度表达。此类基因的实例是来自大肠埃希氏菌的fadR、来自鼠伤寒沙门氏菌(*Salmonella typhimurium*)的fabA(NP_460041)、来自鼠伤寒沙门氏菌的fabD(NP_460164)、来自鼠伤寒沙门氏菌的fabG(NP_460165)、来自鼠伤寒沙门氏菌的fabH(NP_460163)、来自霍乱弧菌(*Vibrio cholera*)的fabV(YP_001217283)以及来自丙酮丁醇梭菌(*Clostridium acetobutylicum*)的fabF(NP_350156)。编码脂肪酸生物合成中酶和调节子的这些基因中一或多者的过度表达可以足以增加各培养条件下包括脂肪醛在内的脂肪酸衍生物中间物以及例如脂肪二醇等最终产物的滴度。

[0126] 在一个实施方案中,大肠埃希氏菌菌株用作产生脂肪二醇的宿主细胞。这些宿主细胞可以包括一种或多种生物合成基因(即编码脂肪酸生物合成中酶和调节子的基因)的任选的过度表达,该过度表达可以进一步增加或增强包括但不限于fadR、fabA、fabD、fabG、fabH、fabV和/或fabF的各培养条件下脂肪酸衍生物中间物(例如脂肪醛)以及最终产物(例如脂肪二醇、脂肪醇)的滴度。遗传变化的实例包括来自大肠埃希氏菌的fadR、来自鼠伤寒沙门氏菌的fabA(NP_460041)、来自鼠伤寒沙门氏菌的fabD(NP_460164)、来自鼠伤寒沙门氏菌的fabG(NP_460165)、来自鼠伤寒沙门氏菌的fabH(NP_460163)、来自霍乱弧菌的fabV(YP_001217283)以及来自丙酮丁醇梭菌的fabF(NP_350156)。在一些实施方案中,运载这些生物合成基因的合成操纵子可以工程化且在细胞中表达以测试各培养条件下的脂肪酸衍生物中间物过度表达和/或进一步增强脂肪二醇产生。此类合成操纵子含有一种或多种生物合成基因。举例来说,ifab138操纵子是一种经工程化的操纵子,其含有任选的脂肪酸生物合成基因,包括来自霍乱弧菌的fabV、来自鼠伤寒沙门氏菌的fabH、来自鼠伤寒沙门氏菌的fabD、来自鼠伤寒沙门氏菌的fabG、来自鼠伤寒沙门氏菌的fabA和/或来自丙酮丁醇梭菌的fabF,所述基因可用于促进脂肪酸衍生物和中间物的过度表达以测试特定培养条件。此类合成操纵子的一个优点是脂肪酸衍生物产生速率(例如脂肪酸、脂肪醛、脂肪醇、脂肪二醇等)可在含有其的细胞中进一步增加或增强。

[0127] 在一些实施方案中,用于产生酰基硫酯(例如酰基-CoA或酰基-ACP)和生物合成酶(例如TE、CAR、AR、ADH、ACC、AAR、FAR、ACR;又见图1和3以及图8-11)的宿主细胞或微生物将进一步表达包涵可以增强一种或多种特定脂肪酸衍生物(例如脂肪酸、3-羟基脂肪酸、脂肪醇、1,3脂肪二醇、脂肪醛、3-羟基脂肪醛等)产生的某些酶活性的基因。在一个实施方案中,宿主细胞具有用于产生脂肪酸和3-羟基脂肪酸的硫酯酶(TE)活性(EC 3.1.2.-或EC 3.1.2.14或EC3.1.1.5),其可以通过基因过度表达而增加。在另一个实施方案中,宿主细胞具有用于产生脂肪醛和/或3-羟基脂肪醛的硫酯酶(TE)活性(EC 3.1.2.-或EC 3.1.2.14或EC 3.1.1.5)和羧酸还原酶(CAR)(EC6.2.1.3或EC 1.2.1.42或EC 1.2.99.6)活性。在另一

个实施方案中,宿主细胞具有用于产生脂肪醇和/或脂肪二醇的硫酯酶(TE)活性(EC3.1.2.-或EC 3.1.2.14或EC 3.1.1.5)和羧酸还原酶(CAR)活性(EC6.2.1.3或EC 1.2.1.42或EC 1.2.99.6)和醇脱氢酶(ADH)/醛还原酶(AR)活性(EC 1.1.1.-)。在另一个实施方案中,宿主细胞具有用于产生脂肪醛和/或3-羟基-脂肪醛的酰基-ACP还原酶(AAR)活性(EC 1.2.1.80或EC 1.2.1.42)。在另一个实施方案中,宿主细胞具有用于产生脂肪醇和/或脂肪二醇的酰基-ACP还原酶(AAR)活性(EC 1.2.1.80或EC 1.2.1.42)和醇脱氢酶(ADH)/醛还原酶(AR)活性(EC 1.1.1.-)。基因组合可以通过微生物相应地工程化而过度表达或表达不足。在一个实施方案中,一个或多个过度表达基因是内源性的。在另一个实施方案中,一个或多个过度表达基因是外源性的。

[0128] 在替代实施方案中,宿主细胞具有用于产生脂肪醇的酰基-ACP还原酶(AAR)活性(EC 1.2.1.80或EC 1.2.1.42)和/或酰基ACP/酰基CoA还原酶(AAR/ACR)活性(EC 1.2.1.80或EC 1.2.1.42或EC1.2.1.50)和/或醇脱氢酶活性(E.C.1.1.-.-)和/或形成脂肪醇的酰基-CoA/Acy1 ACP还原酶(FAR)活性(EC 1.1.1.-)和/或羧酸还原酶(CAR)活性(EC6.2.1.3或EC 1.2.1.42或EC 1.2.99.6)和/或硫酯酶(TE)活性(EC 3.1.2.-或EC 3.1.2.14或EC 3.1.1.5)。在其它替代实施方案中,宿主细胞具有用于产生脂肪醇的酰基-CoA还原酶活性(EC1.2.1.50)和酰基-CoA合酶(FadD)活性(EC 2.3.1.86)和硫酯酶(TE)活性(EC 3.1.2.-或EC 3.1.2.14或EC 3.1.1.5)。微生物和微生物细胞中这些替代酶活性的表达由以引用的方式并入本文中的美国专利号8,097,439、8,110,093、8,110,670、8,183,028、8,268,599、8,283,143、8,232,924、8,372,610和8,530,221教导。在其它实施方案中,用于产生酰基-ACP和/或酰基-CoA及其它生物合成酶的宿主细胞或微生物将包括某些天然酶活性,所述酶活性上调或过度表达以产生一种或多种特定脂肪酸衍生物,例如脂肪醛和/或脂肪醇和/或脂肪二醇。在一个实施方案中,宿主细胞具有用于产生脂肪酸的天然硫酯酶(TE)活性,其可以通过过度表达硫酯酶基因而增加。

[0129] 本公开包括表达编码生物合成酶(上文)的基因的宿主菌株或微生物。重组宿主细胞产生脂肪酸衍生物中间物,例如脂肪醛,和脂肪酸衍生物最终产物,例如脂肪醇和/或脂肪二醇及其组合物和掺合物。脂肪酸衍生物最终产物典型地从培养基回收和/或从宿主细胞分离。在一个实施方案中,脂肪二醇和/或脂肪醇从培养基(细胞外)回收。在另一个实施方案中,脂肪二醇和/或脂肪醇从宿主细胞(细胞内)分离。在另一个实施方案中,脂肪二醇和/或脂肪醇从培养基回收以及从宿主细胞分离。在另一个实施方案中,脂肪二醇和/或脂肪醇是细胞外的且与宿主细胞缔合并从宿主细胞分离。宿主细胞产生的脂肪二醇组合物可以使用本领域中已知的例如GC-FID等方法分析,以确定特定脂肪二醇的分配以及脂肪二醇组合物的组分的链长和饱和度。

[0130] 用作微生物(例如微生物细胞)的宿主细胞的实例包括但不限于来自以下种类的细胞:埃希氏菌属(*Escherichia*)、芽孢杆菌属(*Bacillus*)、乳酸杆菌(*Lactobacillus*)、单胞发酵菌属(*Zymomonas*)、红球菌属(*Rhodococcus*)、假单胞细菌属(*Pseudomonas*)、曲霉属(*Aspergillus*)、木霉属(*Trichoderma*)、链孢霉属(*Neurospora*)、镰刀霉属(*Fusarium*)、腐质霉属(*Humicola*)、根毛霉属(*Rhizomucor*)、克卢费氏酵母属(*Kluyveromyces*)、毕赤氏酵母属(*Pichia*)、毛霉属(*Mucor*)、毁丝霉属(*Myceliophthora*)、青霉属(*Penicillium*)、显草菌属(*Phanerochaete*)、侧耳属(*Pleurotus*)、栓菌属(*Trametes*)、聚球藻属(*Synechococcus*)、

集胞藻属 (*Synechocystis*)、乳球菌属 (*Lactococcus*)、金孢子菌属 (*Chrysosporium*)、酵母属 (*Saccharomyces*)、窄食单胞菌属 (*Stenotrophomonas*)、裂殖酵母属 (*Schizosaccharomyces*)、耶氏酵母属 (*Yarrowia*) 或链霉菌属 (*Streptomyces*)。在一些实施方案中,宿主细胞是革氏阳性细菌细胞。在一些实施方案中,宿主细胞是革氏阴性细菌细胞。在一些实施方案中,宿主细胞是大肠埃希氏菌细胞。在一些实施方案中,宿主细胞是大肠埃希氏菌B细胞、大肠埃希氏菌C细胞、大肠埃希氏菌K细胞或大肠埃希氏菌W细胞。在其它实施方案中,宿主细胞是迟缓芽孢杆菌 (*Bacillus lentus*) 细胞、短芽孢杆菌 (*Bacillus brevis*) 细胞、嗜热脂肪芽孢杆菌 (*Bacillus stearothermophilus*) 细胞、地衣芽孢杆菌 (*Bacillus licheniformis*) 细胞、嗜碱芽孢杆菌 (*Bacillus alkalophilus*) 细胞、凝结芽孢杆菌 (*Bacillus coagulans*) 细胞、环状芽孢杆菌 (*Bacillus circulans*) 细胞、短小芽孢杆菌 (*Bacillus pumilis*) 细胞、苏云金芽孢杆菌 (*Bacillus thuringiensis*) 细胞、克劳氏芽孢杆菌 (*Bacillus clausii*) 细胞、巨大芽孢杆菌 (*Bacillus megaterium*) 细胞、枯草芽孢杆菌 (*Bacillus subtilis*) 细胞或解淀粉芽孢杆菌 (*Bacillus amyloliquefaciens*) 细胞。在其它实施方案中,宿主细胞是康氏木霉 (*Trichoderma koningii*) 细胞、绿色木霉 (*Trichoderma viride*) 细胞、里氏木霉 (*Trichoderma reesei*) 细胞、长枝木霉 (*Trichoderma longibrachiatum*) 细胞、泡盛曲霉 (*Aspergillus awamori*) 细胞、烟曲霉 (*Aspergillus fumigatus*) 细胞、臭曲霉 (*Aspergillus foetidus*) 细胞、构巢曲霉 (*Aspergillus nidulans*) 细胞、黑曲霉 (*Aspergillus niger*) 细胞、米曲霉 (*Aspergillus oryzae*) 细胞、特异腐质霉 (*Humicola insolens*) 细胞、疏棉状腐质霉 (*Humicola lanuginosa*) 细胞、浑浊红球菌 (*Rhodococcus opacus*) 细胞、米赫根毛霉 (*Rhizomucor miehei*) 细胞或米奇毛霉 (*Mucor miehei*) 细胞。在其它实施方案中,宿主细胞是青紫链霉菌 (*Streptomyces lividans*) 细胞或鼠链霉菌 (*Streptomyces murinus*) 细胞。在其它实施方案中,宿主细胞是放线菌属 (*Actinomycetes*) 细胞。在一些实施方案中,宿主细胞是酿酒酵母 (*Saccharomyces cerevisiae*) 细胞。在其它实施方案中,宿主细胞是来自真核生物植物、藻类、蓝细菌、绿色硫细菌、绿色非硫细菌、紫色硫细菌、紫色非硫细菌、嗜极菌、酵母、真菌、其工程化生物体或合成生物体的细胞。在一些实施方案中,宿主细胞为光依赖性的或固定碳。在一些实施方案中,宿主细胞具有自给营养活性。在一些实施方案中,宿主细胞具有光合自养活性,例如在光存在下。在一些实施方案中,在缺乏光下宿主细胞是异养的或混养的。在某些实施方案中,宿主细胞是来自拟南芥 (*Arabidopsis thaliana*)、柳枝稷 (*Panicum virgatum*)、奇岗 (*Miscanthus giganteus*)、玉蜀黍 (*Zea mays*)、布朗葡萄藻 (*Botryococcus braunii*)、莱茵衣藻 (*Chlamydomonas reinhardtii*)、盐生杜氏藻 (*Dunaliella salina*)、聚球藻属 (*Synechococcus* Sp.) PCC 7002、聚球藻属 PCC 7942、集胞藻属 (*Synechocystis* Sp.) PCC 6803、细长嗜热藻 (*Thermosynechococcus elongatus*) BP-1、绿硫菌 (*Chlorobium tepidum*)、嗜热光合绿曲菌 (*Chloroflexus auranticus*)、酒色着色菌 (*Chromatium vinosum*)、深红红螺菌 (*Rhodospirillum rubrum*)、荚膜红细菌 (*Rhodobacter capsulatus*)、沼泽红假单胞菌 (*Rhodopseudomonas palustris*)、扬氏梭菌 (*Clostridium ljungdahlii*)、热纤梭菌 (*Clostridium thermocellum*)、产黄青霉 (*Penicillium chrysogenum*)、巴斯德毕赤氏酵母 (*Pichia pastoris*)、酿酒酵母、非洲粟酒裂殖酵母 (*Schizosaccharomyces pombe*)、荧光假单胞菌 (*Pseudomonas fluorescens*)、恶

臭假单胞菌 (*Pseudomonas putida*) 或运动发酵单胞菌 (*Zymomonas mobilis*) 的细胞。在一特定实施方案中, 微生物细胞来自蓝藻细菌, 包括但不限于原绿球藻 (*Prochlorococcus*)、聚球藻属、集胞藻属、蓝杆藻 (*Cyanothece*) 和点形念珠藻 (*Nostoc punctiforme*)。在另一个实施方案中, 微生物细胞来自特定蓝藻物种, 包括但不限于细长聚球藻 (*Synechococcus elongatus*) PCC 7942、集胞藻属 PCC 6803 和聚球藻属 PCC 7001。

[0131] 经工程化以产生 1,3-二醇的重组宿主细胞

[0132] 本公开鉴别编码酶功能多肽以修饰产生例如脂肪二醇 (例如 1,3-二醇) 等所需化合物的酶途径的多核苷酸。本文中通过酶登录号 (EC 号, 参见以下表 1) 鉴别的这些多肽可用于工程化产生脂肪二醇的脂肪酸途径。更具体地说, 图 1-3 和 8-11 描绘经工程化以产生 1,3-二醇的途径。如所示, 运载酰基中间物 (酰基-ACP 或 3-羟基酰基-ACP) 的 3' 羟基酰基载体蛋白 (ACP) 可转变成 1,3-二醇, 采用 3' 羟基脂肪酸 (3' OH FA) 和 3' 羟基脂肪醛 (3' OH 脂肪醛) 作为中间物。在一个实施方案中, 图 1-3 和 8-11 中描绘产生 1,3-二醇的工程化途径。本文中, 例如葡萄糖等简单碳源首先通过微生物生物体 (例如埃希氏菌属、芽孢杆菌属、乳酸杆菌、红球菌属、聚球藻属、集胞藻属、假单胞细菌属、曲霉属、木霉属、链孢霉属、镰刀霉、腐质霉属、根毛霉属、克卢费氏酵母属、毕赤氏酵母属、毛霉属、毁丝霉属、青霉属、显革菌属、侧耳属、栓菌属、金孢子菌属、酵母属、窄食单胞菌属、裂殖酵母属、耶氏酵母属或链霉属) 转变为 3' 羟基酰基-ACP。在一些实施方案中, 通用和高度保守的酰基-ACP 或 3' 羟基酰基-ACP 通过微生物生物体的天然途径产生。在一个实施方案中, 3' 羟基酰基-ACP 可用于引发工程化途径。举例来说, 3' 羟基酰基-ACP 可以通过具有硫酯酶 (TE) 活性的酶 (参见下表 2) 转变为中间物, 例如 3' OH FA。接着中间物 3' OH FA 可以通过具有羧酸还原酶 (CAR) 活性的酶 (参见下表 3) 转变为另一中间物 3' OH 醛。接着具有醇脱氢酶 (ADH) 或醛还原酶 (AR) 活性的酶 (参见下表 4) 可以将 3' OH 醛转变成 1,3-二醇。为进一步说明此类途径, 图 2 提供具有硫酯酶活性 (例如 *fatB1*、*tesA*、*phaG*)、CAR 活性 (例如 *carB*) 和 ADH/AR 活性 (例如 *alrA*) 的特定酶的实例。可进行 3' OH 酰基-ACP 转变成 3' OH FA 的反应的硫酯酶 (TE) 的其它实例展示于表 2 中。在一个实施方案中, 编码这些硫酯酶的基因是 *tesA*、*tesB*、*fatB*、*fatB1*、*fatB2*、*fatB3*、*TE_EEI82564*、*TE_CAD63310* 和 *phaG*。在另一个实施方案中, 编码这些硫酯酶的基因是 *TE_EEI82564* 和/或 *TE_CAD63310*, 其先前不与 3' OH 酰基-ACP 转变成 3' OH FA 的能力相关 (参见例如 Jing 等人 (2011) *BMC Biochemistry* 12 (44):1471-2091)。可进行 3' OH FA 转变成 3' OH 醛的反应的 CAR 酶的其它实例展示于表 3 中。在一个实施方案中, 编码 CAR 酶的基因是 *carB*。可进行 3' OH 醛转变成 1,3-二醇的反应的 ADH/AR 酶的其它实例展示于表 4 中。在一个实施方案中, 编码这些 ADH/AR 酶的基因是 *alrA* 和/或 *yqhD*。

[0133] 在另一个实施方案中, 图 3 中描绘也产生 1,3-二醇的工程化途径。类似地, 例如葡萄糖等简单碳源首先通过微生物生物体 (例如埃希氏菌属、芽孢杆菌属、乳酸杆菌、红球菌属、聚球藻属、集胞藻属、假单胞细菌属、曲霉属、木霉属、链孢霉属、镰刀霉、腐殖菌属、根毛霉属、克卢费氏酵母属、毕赤氏酵母属、毛霉属、毁丝霉属、青霉属、显革菌属、侧耳属、栓菌属、金孢子菌属、酵母属、窄食单胞菌属、裂殖酵母属、耶氏酵母属或链霉属) 转变为 3' 羟基酰基-ACP。在一些实施方案中, 通用和高度保守的 3' 羟基酰基-ACP 通过微生物生物体的天然途径产生。如上所述, 3' 羟基酰基-ACP 可用于引发工程化途径。举例来说, 3' 羟基酰基-ACP 通过具有酰基-ACP 还原酶 (AAR) 活性的酶 (参见表 1) 转变为中间物, 例如 3' OH 脂肪醛。脂

肪醇和/或脂肪醛通过AAR的产生可通过编码乙酰基-CoA羧化酶的称为accABCD的基因的异源表达来增强。可进行3' OH酰基-ACP转变成3' OH醛的反应的AAR酶的实例包括但不限于来自细长聚球藻、蓝杆藻属、集胞藻属和海洋原绿球藻 (*Prochlorococcus marinus*) 的酶。接着具有醇脱氢酶 (ADH) 或醛还原酶 (AR) 活性的酶 (参见下表4) 可以将3' OH醛转变成脂肪二醇, 例如1,3-二醇。因此, 本公开提供可以有效且选择性地体内产生包括1,3-二醇在内的脂肪二醇的重组微生物。应该注意大多数细胞天然产生能够还原醛的酶, 因为醛会具有细胞毒性。因此, AR和ADH的异源表达可能并非产生脂肪醇和二醇所需, 但其可以提高脂肪醇和二醇的产生效率。

[0134] 另外, 编码具有脂肪酸降解酶活性的多肽的多核苷酸可以任选在宿主细胞中减弱。此类多肽的非限制性实例是酰基-CoA合成酶 (例如大肠埃希氏菌FadD) 和酰基-CoA脱氢酶 (例如大肠埃希氏菌FadE)。表1提供示例性代谢途径中酶活性的综合列表, 包括可以任选地根据本领域中已知的方法 (参见例如美国专利No. 8,283,143, 上文) 减弱的各种脂肪酸降解酶。举例来说, FadR (参见表1) 是参与大肠埃希氏菌中脂肪酸降解和脂肪酸生物合成途径的关键调节因子 (Cronan等人, *Mol. Microbiol.*, 29 (4) :937-943 (1998))。大肠埃希氏菌酶FadD (参见表1) 和脂肪酸转运蛋白FadL是脂肪酸吸收系统的组分。FadL及其同源物介导脂肪酸转运至细菌细胞, 且FadD及其同源物介导酰基-CoA酯的形成。脂肪酸和脂肪酸衍生物的一种替代异源吸收系统是来自假单胞菌属的外膜蛋白AlkL (Julsing等人 (2012) *Appl. Environ. Microbiol.* 78:5724-5733)。当不可利用其它碳源时, 外源性脂肪酸由细菌吸收且转变为酰基-CoA酯, 酰基-CoA酯可以结合于转录因子FadR且减少编码负责脂肪酸转运 (FadL)、活化 (FadD) 和 β -氧化 (FadA、FadB和FadE) 的蛋白质的fad基因的表达。当可利用替代碳源时, 细菌合成如酰基-ACP等脂肪酸, 用于磷脂合成, 而非 β -氧化的底物。因此, 酰基-CoA和酰基-ACP都是可以产生不同最终产物的脂肪酸的独立来源 (Caviglia等人, *J. Biol. Chem.*, 279 (12) :1163-1169 (2004))。FadR和/或FabB及其功能同源物可以增强宿主细胞 (例如大肠埃希氏菌) 中脂肪酸衍生物的产生, 但是其过度表达是任选的。本文中, 涵盖FabB过度表达可以增加伸长速率 (脂肪酸链合成), 且FadR过度表达可以增加FabA和FabB的表达。后者是可能的, 因为预期FadR是FabA和FabB的正调节剂。

[0135] 表1: 酶活性

[0136]

基因名称	来源生物体	酶名称	登录号	EC号	示例性用途
脂肪酸产生增加					
accA	大肠埃希氏菌 (<i>E. Coli</i>)、乳酸 球菌	乙酰基 -CoA 羧化 酶，次单 元 A (羧基 转移酶 α)	AAC73296、 NP_414727	6.4. 1.2	增加丙 二酰基 -CoA 产生
accB	大肠埃希氏菌、 乳酸球菌	乙酰基 -CoA 羧化 酶，次单 元 B (BCCP: 生 物素羧基 载体蛋白)	NP_417721	6.4. 1.2	增加丙 二酰基 -CoA 产生
accC	大肠埃希氏菌、 乳酸球菌	乙酰基 -CoA 羧化 酶，次单 元 C (生物 素羧化酶)	NP_417722	6.4. 1.2、 6.3. 4.14	增加丙 二酰基 -CoA 产生

[0137]

accD	大肠埃希氏菌、 乳酸球菌	乙酰基 -CoA 羧化 酶, 次单 元 D (羧 基转移酶 β)	NP_416819	6.4. 1.2	增加丙 二酰基 -CoA 产生
fadD	大肠杆菌 W3110	酰基-CoA 合酶	AP_002424	2.3. 1.86 、 6.2. 1.3	增加脂 肪酸产 生
fabA	大肠杆菌 K12	β -羟基癸 酰基硫酯 脱水酶/异 构酶	NP_415474	4.2. 1.60	增加脂 肪酰基 -ACP/C oA 产 生
fabB	大肠埃希氏菌	3-氧代基 酰基-[酰 基-载体- 蛋白]合酶 I	BAA16180	2.3. 1.41	增加脂 肪酰基 -ACP/C oA 产 生
fabD	大肠杆菌 K12	[酰基-载 体-蛋白] S-丙二酰 基转移酶	AAC74176	2.3. 1.39	增加脂 肪酰基 -ACP/C oA 产 生
fabF	大肠杆菌 K12	3-氧代基 酰基-[酰 基-载体- 蛋白]合酶 II	AAC74179	2.3. 1.17 9	增加脂 肪酰基 -ACP/C oA 产 生
fabG	大肠杆菌 K12	3-氧代基 酰基-[酰 基-载体蛋 白]还原酶	AAC74177	1.1. 1.10 0	增加脂 肪酰基 -ACP/C oA 产 生

[0138]

fabH	大肠杆菌 K12	3-氧代基 酰基-[酰 基-载体- 蛋白]合酶 III	AAC74175	2.3. 1.18 0	增加脂 肪酰基 -ACP/C oA 产 生
fabI	大肠杆菌 K12	烯酰基 -[酰基-载 体-蛋白] 还原酶	NP_415804	1.3. 1.9	增加脂 肪酰基 -ACP/C oA 产 生
fabR	大肠杆菌 K12	转录抑制 因子	NP_418398	无	调节不 饱和脂 肪酸产 生
fabV	霍乱弧菌	烯酰基 -[酰基-载 体-蛋白] 还原酶	YP_001217 283	1.3. 1.9	增加脂 肪酰基 -ACP/C oA 产 生
fabZ	大肠杆菌 K12	(3R)-羟基 肉豆蔻酰 基载体蛋 白脱水酶	NP_414722	4.2. 1.-	增加脂 肪酰基 -ACP/C oA 产 生
fadE	大肠杆菌 K13	酰基-CoA 脱氢酶	AAC73325	1.3. 99.3 , 1.3. 99.-	减少脂 肪酸降 解
fadD	大肠杆菌 K12	酰基-CoA 合成酶	NP_416319	6.2. 1.3	减少脂 肪酸降 解
fadA	大肠杆菌 K12	3-酮酰基 -CoA 硫解 酶	YP_02627	2.3. 1.16	减少脂 肪酸降 解

fadB	大肠杆菌 K12	烯酰基-CoA 水合酶、3-OH 酰基-CoA 差向异构酶/脱氢酶	NP_418288	4.2.1.17、5.1.2.3、1.1.1.35	减少脂肪酸降解
fadR	大肠埃希氏菌	转录调节蛋白	NP_415705	无	阻断或逆转脂肪酸降解
链长控制					
tesA (有或无前导序列)	大肠埃希氏菌	硫酯酶 - 前导序列包括氨基酸 1-26	P0ADA1	3.1.2.2.-、3.1.1.5、3.1.2.14	C18 链长
tesA (无前导序列)	大肠埃希氏菌	硫酯酶	AAC73596、NP_415027	3.1.2.2.-、3.1.1.5、3.1.2.14	C18:1 链长
tesA (与辛酸复合的大肠埃希氏菌硫酯酶 I 的突变体)	大肠埃希氏菌	硫酯酶	L109P	3.1.2.2.-、3.1.1.5、3.1.2.14	<C18 链长
fatB1	加州月桂 (<i>Umbellularia californica</i>)	硫酯酶	Q41635	3.1.2.14	C12:0 链长

[0139]

[0140]

fatB2	萼距花(<i>Cuphea hookeriana</i>)	硫酯酶	AAC49269	3.1. 2.14	C8:0 - C10:0 链长
fatB3	萼距花	硫酯酶	AAC72881	3.1. 2.14	C14:0 - C16:0 链长
fatB	樟 树 (<i>Cinnamomum amphora</i>)	硫酯酶	Q39473	3.1. 2.14	C14:0 链长
fatB	拟南芥	硫酯酶	CAA85388	3.1. 2.14	C16:1 链长
fatA1	向日葵 (<i>Helianthus annuus</i>)	硫酯酶	AAL79361	3.1. 2.14	C18:1 链长
fatA	拟南芥	硫酯酶	NP_189147, NP_193041	3.1. 2.14	C18:1 链长
fatA	芥菜(<i>Brassica juncea</i>)	硫酯酶	CAC39106	3.1. 2.14	C18:1 链长
fatA	萼距花	硫酯酶	AAC72883	3.1. 2.14	C18:1 链长
tes	深海发光杆菌 (<i>Photobacterium profundum</i>)	硫酯酶	YP_130990	3.1. 2.2. 、 3.1. 2.14	链长
tesB	大肠埃希氏菌	硫酯酶	NP_414986	3.1. 2.2.	链长
fadM	大肠埃希氏菌	硫酯酶	NP_414977	3.1. 2.2.	链长
yciA	大肠埃希氏菌	硫酯酶	NP_415769	3.1. 2.2.	链长
ybgC	大肠埃希氏菌	硫酯酶	NP_415264	3.1. 2.2.	链长
phaG	恶臭假单胞菌	硫酯酶、 3-羟基酰 基 -CoA- 酰基载体 蛋白转移 酶	AAN67031	3.1. 2.14 、	C6 至 C12 链 长
饱和度控制					

Sfa	大肠埃希氏菌	fabA 遏制因子	AAN79592, AAC44390	无	增加单不饱和脂肪酸
fabA	大肠杆菌 K12	β -羟基癸酰基硫酯脱水酶/异构酶	NP_415474	4.2.1.60	产生不饱和脂肪酸
GnsA	大肠埃希氏菌	secG 无效突变遏制剂	ABD18647.1	无	增加不饱和脂肪酸酯
GnsB	大肠埃希氏菌	secG 无效突变遏制剂	AAC74076.1	无	增加不饱和脂肪酸酯
fabB	大肠埃希氏菌	3-氧代基酰基-[酰基-载体-蛋白]合酶 I	BAA16180	2.3.1.41	调节不饱和脂肪酸产生
des	枯草芽胞杆菌	D5 脂肪酰基去饱和酶	O34653	1.14.19	调节不饱和脂肪酸产生
酯产生					
AT3G51970	拟南芥	长链醇 O-脂肪酰基转移酶	NP_190765	2.3.1.26	酯产生
ELO1	安格斯毕赤氏酵母 (<i>Pichia angusta</i>)	脂肪酸延长酶	BAD98251	2.3.1.-	产生极长链长脂肪酸
plsC	酿酒酵母	酰基转移酶	AAA16514	2.3.1.51	酯产生
DAGAT/DGAT	拟南芥	二酰基甘油酰基转移酶	AAF19262	2.3.1.20	酯产生
hWS	智人	酰基-CoA 蜡醇酰基转移酶	AAX48018	2.3.1.20	酯产生

[0141]

aft1	不动细菌属 (<i>Acinetobacter</i> <i>sp.</i>) ADP1	双官能蜡 酯合酶/酰 基 -CoA: 二酰基甘 油酰基转 移酶	AAO17391	2.3. 1.20	酯产生
ES9	除烃海杆菌 (<i>Marinobacter</i> <i>hydrocarbonocl</i> <i>asticus</i>)	蜡酯合酶	ABO21021	2.3. 1.20	酯产生
mWS	加州希蒙得木 (<i>Simmondsia</i> <i>chinensis</i>)	蜡酯合酶	AAD38041	2.3. 1.-	酯产生
脂肪醇产出					
		硫酯酶(参 见上文)			增加脂 肪酸/ 脂肪醇 产生
BmFA R	家 蚕 (<i>Bombyxmori</i>)	FAR (形 成脂肪醇 的酰基 -CoA还原 酶)	BAC79425	1.2. 1.50 、 1.2. 1.84	酰基 -CoA 转变成 脂肪醇
acr1	不动细菌属 ADP1	酰基-CoA 还原酶	YP_047869	1.2. 1.42 、 1.2. 1.50	脂肪酰 基-CoA 还原成 脂肪醛
yqhD	大肠杆菌 W3110	醇脱氢酶	AP_003562	1.1. 1.-	脂肪醛 还原成 脂肪 醇; 增 加脂肪 醇产生
alrA	不动细菌属 ADP1	醇脱氢酶	CAG70252	1.1. 1.-	脂肪醛 还原成 脂肪醇

[0142]

[0143]

GTNG_1865	嗜热脂肪地芽胞杆菌 (<i>Geobacillus thermodenitrificans</i>) NG80-2	长链醛脱氢酶	YP_001125970	1.2. 1.3	脂肪醛还原成脂肪醇
AAR	细长聚球藻	酰基-ACP还原酶	YP_400611	1.2. 1.42	脂肪酰基-ACP/CoA还原成脂肪醛
carB	耻垢分枝杆菌 (<i>Mycobacterium smegmatis</i>)	羧酸还原酶蛋白	YP_889972	6.2. 1.3. 1.2. 1.42	脂肪酸还原成脂肪醛
FadD	大肠杆菌 K12	酰基-CoA合成酶	NP_416319	6.2. 1.3	脂肪酸活化成脂肪酰基-CoAs
atoB	胡萝卜软腐欧文氏菌 (<i>Erwinia carotovora</i>)	乙酰基-CoA乙酰基转移酶	YP_049388	2.3. 1.9	产生丁醇
hbd	溶纤维丁酸弧菌 (<i>Butyrivibrio fibrisolvens</i>)	β -羟基丁酰基-CoA脱氢酶	BAD51424	1.1. 1.15 7	产生丁醇
CPE0095	产气荚膜梭状芽胞杆菌 (<i>Clostridium perfringens</i>)	巴豆酸酶 丁酰基-CoA脱氢酶	BAB79801	4.2. 1.55	产生丁醇
bcd	贝氏梭状芽胞杆菌 (<i>Clostridium beijerinckii</i>)	丁酰基-CoA脱氢酶	AAM14583	1.3. 99.2	产生丁醇

ALDH	贝氏梭状芽胞杆菌	辅酶 A-酰化醛脱氢酶	AAT66436	1.2. 1.3	产生丁醇
AdhE	大肠杆菌 CFT073	醛-醇脱氢酶	AAN80172	1.1. 1.1、 1.2. 1.10	产生丁醇
脂肪醇乙酰基酯产出					
		硫酯酶(参见上文)			改变产出
acr1	不动细菌属 ADP1	酰基-CoA 还原酶	YP_047869	1.2. 1.42 、 1.2. 1.50	改变产出
yqhD	大肠杆菌 K12	醇脱氢酶	AP_003562	1.1.- .-	改变产出
AAT	草莓(<i>Fragaria x ananassa</i>)	醇 O-乙酰基转移酶	AAG13130	2.3. 1.84	改变产出
端烯产出					
OleT	泡菜球菌属 (<i>Jeotgalicoccus sp.</i>)	脂肪酸脱羧酶	HQ709266	1.11 .2.4	脂肪酸脱羧
产物输出					
AtMRP5	拟南芥	拟南芥多药抗性相关	NP_171908	无	改变产物输出量
AmiS2	红球菌属	ABC 转运体 AmiS2	JC5491	无	改变产物输出量
AtPGP1	拟南芥	拟南芥 p 糖蛋白 1	NP_181228	无	改变产物输出量
AcrA	皮氏候选种 (<i>Candidatus Prot ochlamydia amoebophila</i>) UWE25	推定多药流出转运蛋白, acrA	CAF23274	无	改变产物输出量

[0144]

AcrB	皮氏候选种 <i>UWE25</i>	可能多药 流出转运 蛋白， acrB	CAF23275	无	改变产 物输出 量
TolC	土拉弗朗西斯 菌亚种诺威达 (<i>Francisellatula rensis subsp. novicida</i>)	外膜蛋白 [细胞外被 膜生源，	ABD59001	无	改变产 物输出 量
AcrE	索氏志贺氏菌 (<i>Shigellasonnei</i>) <i>Ss046</i>	跨膜蛋白 影响隔膜 形成和细 胞膜渗透 性	YP_312213	无	改变产 物输出 量
AcrF	大肠埃希氏菌	吡啶黄 (Acridflavi ne)抗性蛋 白 F	P24181	无	改变产 物输出 量
[0145]	细长嗜热聚球 藻 (<i>Thermosynecho coccus elongatus</i>) [BP-1]	多药流出 转运体	NP_682409. 1	无	改变产 物输出 量
tll0139	细长嗜热聚球 藻[BP-1]	多药流出 转运体	NP_680930. 1	无	改变产 物输出 量
AlkL	恶臭假单胞菌	外膜蛋白	YP_009076 010	无	改变产 物输出 量
发酵					
复制检 查点基 因					增加产 出效率
umuD	索氏志贺氏菌 <i>Ss046</i>	DNA 聚合 酶 V，次 单元	YP_310132	3.4. 21.-	增加产 出效率

	umuC	大肠埃希氏菌	DNA 聚合酶 V, 次单元	ABC42261	2.7. 7.7	增加产出效率
	pntA、 pntB	福氏志贺氏菌 (<i>Shigella flexneri</i>)	NADH:NADPH 转氢酶(α 和 β 次单元)	P07001、 P0AB70	1.6. 1.2	增加产出效率
	其它					
[0146]	fabK	肺炎链球菌 (<i>Streptococcus pneumoniae</i>)	反式-2-烯酰基-ACP 还原酶 II	AAF98273	1.3. 1.9	促进脂肪酸生物合成
	fabL	地衣形芽孢杆菌 (<i>Bacillus licheniformis</i>) DSM 13	烯酰基-(酰基载体蛋白)还原酶	AAU39821	1.3. 1.9	促进脂肪酸生物合成
	fabM	变形链球菌 (<i>Streptococcus mutans</i>)	反式-2, 顺式-3-癸烯酰基-ACP 异构酶	DAA05501	4.2. 1.17	促进脂肪酸生物合成

[0147] 表2: 硫酯酶活性

名称/名字	生物体	EC 号
tesA (有或无前导序列)	大肠埃希氏菌	3.1.2.2-或 3.1.2.14 或 3.1.1.5
tesA (无前导序列)	大肠埃希氏菌	3.1.2.2-或 3.1.2.14 或 3.1.1.5
tes	深海发光杆菌	3.1.2.2 或 3.1.2.14
tesB	大肠埃希氏菌	3.1.2.2
[0148] TE_CAD63310	胚芽乳杆菌 (<i>Lactobacillus plantarum</i>)	3.1.2.2 或 3.1.2.14
TE_EEI82564	四联厌氧球菌 (<i>Anaerococcus tetradus</i>)	3.1.2.2 或 3.1.2.14
fatB1	加州月桂	3.1.2.14
fatB2	萼距花	3.1.2.14

名称/名字	生物体	EC 号
fatB3	萼距花	3.1.2.14
fatB	樟树	3.1.2.14
fatB	拟南芥	3.1.2.14
fatA1	向日葵	3.1.2.14
[0149] fatA	芥菜	3.1.2.14
fatA	萼距花	3.1.2.14
fadM	大肠埃希氏菌	3.1.2.2
yciA	大肠埃希氏菌	3.1.2.2
ybgC	大肠埃希氏菌	3.1.2.2
phaG	恶臭假单胞菌	3.1.2.2 或 3.1.2.14

[0150] 表3: 羧酸还原酶 (CAR) 活性

名称/名字	生物体	Genpept 登录	EC 号
carB	耻垢分枝杆菌	YP_889972	1.2.99.6
carA	耻垢分枝杆菌	ABK75684	1.2.99.6
FadD9	耻垢分枝杆菌	WP_003897160	1.2.99.6
[0151] car	日内瓦分枝杆菌 (<i>Mycobacterium genavense</i>)	WP_025734970	1.2.99.6
car	艾奥瓦诺卡氏菌 (<i>Nocardia iowensis</i>)	Q6RKB1	1.2.99.6
car	巴西诺卡氏菌 (<i>Nocardia brasiliensis</i>)	WP_029899937	1.2.99.6

[0152] 表4: 醇脱氢酶 (ADH) 或醛还原酶 (AR) 活性

名称/名字	生物体	EC 号
yqhD	大肠埃希氏菌	1.1.1.1 或 1.1.1.2
alrA	不动细菌属 ADP1	1.1.1.1 或 1.1.1.2
[0153] yjgB	大肠埃希氏菌	1.1.1.1 或 1.1.1.2
yahK	大肠埃希氏菌	1.1.1.1 或 1.1.1.2
adhP	大肠埃希氏菌	1.1.1.1 或 1.1.1.2
ydjL	大肠埃希氏菌	1.1.1.1 或 1.1.1.2
yhdH	大肠埃希氏菌	1.1.1.1 或 1.1.1.2

	名称/名字	生物体	EC 号
[0154]	yafB (dkgB)	大肠埃希氏菌	1.1.1.1 或 1.1.1.2
	yqhE (dkgA)	大肠埃希氏菌	1.1.1.1 或 1.1.1.2
	ybbo	大肠埃希氏菌	1.1.1.1 或 1.1.1.2
	gldA	大肠埃希氏菌	1.1.1.1 或 1.1.1.2

[0155] 本公开鉴别编码可用于重组宿主细胞和产生方法的具有酶活性的多肽的多核苷酸。具有酶活性的多肽促进包括脂肪二酯化合物的组合物的产生。一般公认无需与此类多核苷酸的绝对序列一致性。举例来说,可以使特定多核苷酸序列(例如编码具有酶功能的多肽的多核苷酸)变化且针对活性筛选编码多肽。此类变化典型地包含保守突变和沉默突变(例如密码子最佳化)。可以使用本领域中已知的方法,针对所需功能,包括但不限于催化活性增加、稳定性增加或抑制减少(例如反馈抑制减少),筛选基因工程化或修饰的多核苷酸和编码的变异多肽。

[0156] 另外,本公开根据酶分类(EC)号鉴别与如本文中描述(上文)的参与脂肪二酯产生的工程化途径的各个步骤(即反应)有关的酶活性,并提供按此类EC号分类的示例性多肽(例如酶)和编码此类多肽的示例性多核苷酸。本文中通过登录号和/或序列鉴别编码(SEQ ID NO)鉴别的此类示例性多肽和多核苷酸可用于将在亲本宿主细胞中产生包括1,3脂肪二酯在内的脂肪二酯的脂肪酸途径工程化,以获得本文中描述的重组或基因修饰的宿主细胞。本文中描述的多肽和多核苷酸是示例性和非限制性的。本文中描述的示例性多肽的同源物的序列可由本领域技术人员通过各个数据库(例如美国国家生物技术信息中心(NCBI)提供的Entrez数据库、瑞士生物信息学研究所提供的ExPasy数据库、不伦瑞克工业大学(The Technical University of Braunschweig)提供的BRENDA资料库和京都大学与东京大学生物信息中心(The Bioinformatics Center of Kyoto University and University of Tokyo)提供的KEGG资料库,都可在万维网上获得)获得。

[0157] 发酵和脂肪二酯的产生

[0158] 如本文所用,发酵泛指重组宿主细胞将有机物质转变成目标物质。举例来说,这包括通过在包含碳源的培养基中繁殖重组宿主细胞的培养物,重组宿主细胞将碳源转变成脂肪酸衍生物,例如脂肪二酯。可用于产生例如脂肪二酯和/或脂肪醇等目标物质的条件是容许宿主细胞产生例如脂肪二酯组合物等所需产物的任何条件。类似地,这包括其中在宿主中表达的载体的多核苷酸序列容许宿主细胞合成目标多肽的任何条件。适合条件包括例如典型发酵条件。发酵条件可以包括许多参数,包括但不限于温度范围、pH水平、通风水平、馈送速率和培养基组成。这些条件中的每一者单独和组合下容许宿主细胞生长。发酵可以是好氧、厌氧或其变体(例如微好氧)。示例性培养基包括肉汤(液体)或凝胶(固体)。一般来说,培养基包括可以通过宿主细胞直接代谢的碳源(例如来源于可再生原料的简单碳源)。另外,酶可用于培养基中以促进碳源的流动(例如淀粉或纤维素解聚成发酵性糖)和后面代谢。

[0159] 对于小规模生产,工程化的宿主细胞可以呈例如约100 μ L、200 μ L、300 μ L、400 μ L、500 μ L、1mL、5mL、10mL、15mL、25mL、50mL、75mL、100mL、500mL、1L、2L、5L或10L的批量生长;发酵;以及诱导表达所需多核苷酸序列,例如编码具有特定酶活性的多肽的多核苷酸(例如

TE、CAR、ADH、FAR、ACR、ACC和/或AAR酶活性)。对于大规模生产,工程化的宿主细胞可以在体积批量为约10L、100L、1000L、10,000L、100,000L、1,000,000L或更大的培养物中生长;发酵;以及诱发表达任何所需多核苷酸序列。本文中描述的脂肪二醇组合物可以在重组宿主细胞培养物的细胞外环境中发现且可以容易从培养基分离。例如脂肪二醇和/或脂肪醇等脂肪酸衍生物可以由重组宿主细胞分泌,输送至细胞外环境或被动转移至重组宿主细胞培养物的细胞外环境中。脂肪二醇组合物可以使用本领域中已知的常规方法从重组宿主细胞培养物分离。

[0160] 为了产生脂肪二醇,进行许多修饰以产生宿主细胞(上文)。因此,本公开提供相对于未经工程化或原生宿主细胞(例如用作对照细胞的野生型宿主细胞),经工程化以提供生物合成途径的重组宿主细胞,工程化例如通过特定菌株改良来实现。例如细菌、蓝藻细菌、酵母、藻类或丝状真菌等微生物可以用作生产宿主。可以用作生产宿主的微生物的非限制性实例包括大肠埃希氏菌、酿酒酵母等等。微生物菌株有效地将葡萄糖或其它可再生原料转变成脂肪酸衍生物,包括脂肪醇和脂肪二醇。为了实现此,菌株已经小心地工程化以表达具有特定功能性的关键酶。已经建立了用于产生各种化合物的高密度发酵的方案和程序(参见例如以引用的方式并入本文中的美国专利No.8,372,610;8,323,924;8,313,934;8,283,143;8,268,599;8,183,028;8,110,670;8,110,093;和8,097,439)。

[0161] 值得注意地,直到现在才出现直接并有效地从葡萄糖或其它可再生的原料产生包括1,3-二醇在内的脂肪二醇的方法。然而,这些脂肪二醇可用作清洁剂、表面活性剂、乳化剂、软化剂、溶剂、塑料和食品添加剂的组分。用于产生如本文中呈现的脂肪二醇及其组合物的基于发酵的方法提供了本领域中所用的化学方法的环保替代方法。在一些实施方案中,宿主细胞在包含约20g/L至约900g/L起始浓度的碳源(例如简单碳源)的培养基(例如发酵培养基)中培养。在其它实施方案中,培养基包含约2g/L至约10g/L、约10g/L至约20g/L、约20g/L至约30g/L、约30g/L至约40g/L或约40g/L至约50g/L起始浓度的碳源(例如简单碳源)。在一些实施方案中,培养基中可用碳源的水平可以在发酵进行期间监测。在一些实施方案中,所述方法进一步包括当培养基中初始碳源的水平小于约0.5g/L时添加补充碳源至培养基。在一些实施方案中,当培养基中碳源的水平小于约0.4g/L、小于约0.3g/L、小于约0.2g/L或小于约0.1g/L时添加补充碳源至培养基。在一些实施方案中,添加补充碳源以维持约1g/L至约25g/L的碳源水平。在一些实施方案中,添加补充碳源以维持约2g/L或更多(例如约2g/L或更多、约3g/L或更多、约4g/L或更多)的碳源水平。在一些实施方案中,添加补充碳源以维持约5g/L或更少(例如约5g/L或更少、约4g/L或更少、约3g/L或更少)的碳源水平。在一些实施方案中,添加补充碳源以维持约2g/L至约5g/L、约5g/L至约10g/L或约10g/L至约25g/L的碳源水平。

[0162] 在一个实施方案中,用于发酵的碳源来源于可再生原料。在一些实施方案中,碳源是葡萄糖。在一些实施方案中,碳源是甘油。其它可能的碳源包括但不限于果糖、甘露糖、半乳糖、木糖、阿拉伯糖、淀粉、纤维素、果胶、木聚糖、蔗糖、麦芽糖、纤维二糖和松二糖;纤维素材料和变体,例如半纤维素、甲基纤维素和羧甲基纤维素钠;饱和或不饱和脂肪酸、丁二酸酯、乳酸酯和乙酸酯;醇,例如乙醇、甲醇和甘油或其混合物。在一个实施方案中,碳源来源于玉米、甘蔗、高粱、甜菜、柳枝稷、青贮饲料、稻草、木材、纸浆、污水、垃圾、纤维素城市废物、烟气、合成气或二氧化碳。简单碳源还可以为光合作用产物,例如葡萄糖或蔗糖。在一个

实施方案中,碳源来源于废物,例如甘油、烟气或合成气;或例如生物质等有机材料的重整;或天然气或甲烷;或这些物质重整成合成气;或光合作用固定的二氧化碳,例如1,3-二醇可以通过光合且使用CO₂为碳源生长的重组蓝藻细菌产生。在某些实施方案中,碳源来源于生物质。生物质的一个示例性来源是植物物质或植被,例如玉米、甘蔗或柳枝稷。生物质的另一个示例性来源是代谢废物,例如动物物质(例如牛粪)。生物质的其它示例性来源包括藻类及其它海生植物。生物质还包括来自工业、农业、林业和家庭的废产物,包括但不限于发酵废物、青贮饲料、稻草、木材、污水、垃圾、纤维素城市废物、城市固体废物和吃剩的食物。

[0163] 在一些实施方案中,脂肪二醇(例如1,3-二醇)以约0.5g/L至约40g/L的浓度产生。在一些实施方案中,脂肪二醇以1g/L或更多(例如约1g/L或更多、约10g/L或更多、约20g/L或更多、约50g/L或更多、约100g/L或更多)的浓度产生。在一些实施方案中,脂肪二醇以约1g/L至约170g/L、约1g/L至约10g/L、约40g/L至约170g/L、约100g/L至约170g/L、约10g/L至约100g/L、约1g/L至约40g/L、约40g/L至约100g/L或约1g/L至约100g/L的浓度产生。

[0164] 在一些实施方案中,脂肪二醇以约25mg/L、约50mg/L、约75mg/L、约100mg/L、约125mg/L、约150mg/L、约175mg/L、约200mg/L、约225mg/L、约250mg/L、约275mg/L、约300mg/L、约325mg/L、约350mg/L、约375mg/L、约400mg/L、约425mg/L、约450mg/L、约475mg/L、约500mg/L、约525mg/L、约550mg/L、约575mg/L、约600mg/L、约625mg/L、约650mg/L、约675mg/L、约700mg/L、约725mg/L、约750mg/L、约775mg/L、约800mg/L、约825mg/L、约850mg/L、约875mg/L、约900mg/L、约925mg/L、约950mg/L、约975mg/L、约1000mg/L、约1050mg/L、约1075mg/L、约1100mg/L、约1125mg/L、约1150mg/L、约1175mg/L、约1200mg/L、约1225mg/L、约1250mg/L、约1275mg/L、约1300mg/L、约1325mg/L、约1350mg/L、约1375mg/L、约1400mg/L、约1425mg/L、约1450mg/L、约1475mg/L、约1500mg/L、约1525mg/L、约1550mg/L、约1575mg/L、约1600mg/L、约1625mg/L、约1650mg/L、约1675mg/L、约1700mg/L、约1725mg/L、约1750mg/L、约1775mg/L、约1800mg/L、约1825mg/L、约1850mg/L、约1875mg/L、约1900mg/L、约1925mg/L、约1950mg/L、约1975mg/L、约2000mg/L(2g/L)、3g/L、5g/L、10g/L、20g/L、30g/L、40g/L、50g/L、60g/L、70g/L、80g/L、90g/L、100g/L或由前述值中的任两者限定的范围的滴度产生。在一些实施方案中,脂肪二醇(例如1,3-二醇)以超过100g/L、超过200g/L、超过300g/L或更高,例如500g/L、700g/L、1000g/L、1200g/L、1500g/L或2000g/L的滴度产生。根据本公开的方法由重组宿主细胞产生的例如1,3-二醇等脂肪二醇的优选滴度为5g/L至200g/L、10g/L至150g/L、20g/L至120g/L和30g/L至100g/L、100g/L至150g/L以及120g/L至180g/L。在一个实施方案中,根据本公开的方法由重组宿主细胞产生的例如1,3-二醇等脂肪二醇的滴度为约1g/L至约250g/L且更具体地说,90g/L至约120g/L。滴度可指特定1,3-二醇或给定重组宿主细胞培养物产生的不同链长或不同功能性的1,3-二醇的组合。

[0165] 在其它实施方案中,根据本公开的方法经工程化以产生例如1,3-二醇等脂肪二醇的宿主细胞的产率为至少1%、至少2%、至少3%、至少4%、至少5%、至少6%、至少7%、至少8%、至少9%、至少10%、至少11%、至少12%、至少13%、至少14%、至少15%、至少16%、至少17%、至少18%、至少19%、至少20%、至少21%、至少22%、至少23%、至少24%、至少25%、至少26%、至少27%、至少28%、至少29%、至少30%或至少40%或由前述值中的任两个限定的范围。在其它实施方案中,例如1,3-二醇等脂肪二醇以超过30%、40%、50%、60%、70%、80%、90%或更多的产率产生。或者或另外,产率为约30%或更少、约27%或更

少、约25%或更少或约22%或更少。因此,产率可由前述终点中的任两者限定。举例来说,根据本公开的方法由重组宿主细胞产生的例如1,3-二醇等脂肪二醇的产率可以是5%至15%、10%至25%、10%至22%、15%至27%、18%至22%、20%至28%或20%至30%。在一特定实施方案中,由重组宿主细胞产生的例如1,3-二醇等脂肪二醇的产率为约10%至约40%。在另一特定实施方案中,由重组宿主细胞产生的例如1,3-二醇等脂肪二醇的产率为约25%至约30%。产率可指由给定重组宿主细胞培养物产生的例如1,3-二醇等特定脂肪二醇或1,3-二醇的组合。另外,产率还将依赖于所用原料。

[0166] 在一些实施方案中,重组宿主细胞产生的例如1,3-二醇等脂肪二醇的生产率是至少100mg/L/小时、至少200mg/L/小时、至少300mg/L/小时、至少400mg/L/小时、至少500mg/L/小时、至少600mg/L/小时、至少700mg/L/小时、至少800mg/L/小时、至少900mg/L/小时、至少1000mg/L/小时、至少1100mg/L/小时、至少1200mg/L/小时、至少1300mg/L/小时、至少1400mg/L/小时、至少1500mg/L/小时、至少1600mg/L/小时、至少1700mg/L/小时、至少1800mg/L/小时、至少1900mg/L/小时、至少2000mg/L/小时、至少2100mg/L/小时、至少2200mg/L/小时、至少2300mg/L/小时、至少2400mg/L/小时或至少2500mg/L/小时。举例来说,根据本公开的方法由重组宿主细胞产生的例如1,3-二醇等脂肪二醇的生产率可以是500mg/L/小时至2500mg/L/小时或700mg/L/小时至2000mg/L/小时。在一特定实施方案中,生产率为约0.7mg/L/小时至约3g/L/小时。生产率可指由既定重组宿主细胞产生的例如1,3-二醇等特定脂肪二醇。

[0167] 在一些实施方案中,用于本文中论述的发酵程序(上文)中的宿主细胞是哺乳动物细胞、植物细胞、昆虫细胞、酵母细胞、真菌细胞、丝状真菌细胞、藻类细胞、蓝藻细胞和细菌细胞。在特定实施方案中,宿主细胞选自埃希氏菌属、芽孢杆菌属、假单胞细菌属、乳酸杆菌、红球菌属、聚球藻属、集胞藻属、假单胞细菌属、曲霉属、木霉属、链孢霉属、镰刀霉、腐质霉属、根毛霉属、克卢费氏酵母属、毕赤氏酵母属、毛霉属、毁丝霉属、青霉属、显革菌属、侧耳属、栓菌属、金孢子菌属、酵母属、窄食单胞菌属、裂殖酵母属、耶氏酵母属或链霉属。在其它实施方案中,宿主细胞是迟缓芽孢杆菌细胞、短芽孢杆菌细胞、嗜热脂肪芽孢杆菌细胞、地衣芽孢杆菌细胞、嗜碱芽孢杆菌细胞、凝结芽孢杆菌细胞、环状芽孢杆菌细胞、短小芽孢杆菌细胞、苏云金芽孢杆菌细胞、克劳氏芽孢杆菌细胞、巨大芽孢杆菌细胞、枯草芽孢杆菌细胞或解淀粉芽孢杆菌细胞。在其它实施方案中,宿主细胞是恶臭假单胞菌细胞。在某些实施方案中,宿主细胞是聚球藻属PCC7002、细长聚球藻PCC 7942、集胞藻属PCC 6803、细长聚球藻PCC6301、海洋原绿球藻CCMP1986 (MED4)、多变鱼腥藻(*Anabaena variabilis*) ATCC29413、点形念珠藻ATCC29133 (PCC73102)、无类囊体蓝藻(*Gloeobacter violaceus*) ATCC29082 (PCC7421)、念珠藻属(*Nostoc* sp.) ATCC27893 (PCC7120)、蓝杆藻属(*Cyanothece* sp.) PCC7425 (29141)、蓝杆藻属ATCC51442或聚球藻属ATCC27264 (PCC7002)。在其它实施方案中,宿主细胞是康氏木霉细胞、绿色木霉细胞、里氏木霉细胞、长枝木霉细胞、泡盛曲霉细胞、烟曲霉细胞、臭曲霉细胞、构巢曲霉细胞、黑曲霉细胞、米曲霉细胞、特异腐质霉细胞、疏棉状腐质霉细胞、浑浊红球菌细胞、米赫根毛霉细胞或米奇毛霉细胞。在其它实施方案中,宿主细胞是放线菌属细胞。在其它实施方案中,宿主细胞是青紫链霉菌细胞或鼠链霉菌细胞。在其它实施方案中,宿主细胞是酿酒酵母细胞。

[0168] 在其它实施方案中,宿主细胞是来自真核生物植物、藻类、蓝细菌、绿色硫细菌、绿

色非硫细菌、紫色硫细菌、紫色非硫细菌、嗜极菌、酵母、真菌、其工程化生物体或合成生物体的细胞。在某些实施方案中,宿主细胞是来自拟南芥、柳枝稷、奇岗、玉蜀黍、布朗葡萄藻、莱茵衣藻、盐生杜氏藻、细长嗜热藻、聚球藻属、集胞藻属、绿硫菌、嗜热光合绿曲菌、酒色着色菌、深红红螺菌、荚膜红细菌、沼泽红假单胞菌、扬氏梭菌、热纤梭菌或产黄青霉的细胞。在某些其它实施方案中,宿主细胞是来自巴斯德毕赤氏酵母、酿酒酵母、解脂耶氏酵母(*Yarrowia lipolytica*)、非洲粟酒裂殖酵母、荧光假单胞菌、恶臭假单胞菌或运动发酵单胞菌的细胞。在其它实施方案中,宿主细胞是来自聚球藻属PCC 7002、聚球藻属PCC 7942或集胞藻属PCC 6803的细胞。在一些实施方案中,宿主细胞是CHO细胞、COS细胞、VERO细胞、BHK细胞、海拉细胞(HeLa cell)、Cv1细胞、MDCK细胞、293细胞、3T3细胞或PC12细胞。。在特定实施方案中,宿主细胞是大肠埃希氏菌细胞。在一些实施方案中,大肠埃希氏菌细胞是菌株B、菌株C、菌株K或菌株W大肠埃希氏菌细胞。

[0169] 脂肪二醇的组合物和配方

[0170] 生物产物(例如根据本公开产生的脂肪二醇组合物),包括生物产生的有机化合物和尤其使用本文公开的脂肪酸生物合成途径产生的脂肪二醇组合物,是从可再生来源(例如从来源于可再生原料的简单碳源)产生的,因而是新的物质组合物。根据双碳同位素指纹法或¹⁴C年代测定,可将这些新生物产物与来源于石油化学碳的有机化合物区别。另外,生物来源碳的特定来源(例如葡萄糖对比甘油)可以通过双碳同位素指纹法测定(参见例如美国专利No.7,169,588)。区别例如本公开的脂肪二醇等生物产物与基于石油的有机化合物的能力有益于跟踪贸易中的这些物质。举例来说,包含基于生物和基于石油的碳同位素概况的有机化合物或化学品可以与仅仅由基于石油的物质制成的有机化合物和化学品相区别。因此,本文中产生的生物产物可以根据其独特的碳同位素概况而在贸易追踪或跟踪。可以通过比较每个样品中稳定碳同位素比率(¹³C/¹²C),来区别生物产物与基于石油的有机化合物。既定生物产物中¹³C/¹²C比率是二氧化碳固定时常压二氧化碳中¹³C/¹²C比率的结果。其还反映出准确的代谢途径。也发生区域性变化。石油、C3植物(阔叶树)、C4植物(草)和海相碳酸盐岩都展示¹³C/¹²C和对应 $\delta^{13}\text{C}$ 值的显著差异。C4与C3植物都展现一系列¹³C/¹²C同位素比率,但是C4植物的典型值为约-7至约-13千分位数,且C3植物的典型值为约-19至约-27千分位数(参见例如Stuiver等人,Radiocarbon 19:355(1977))。例如煤和石油等不可再生物质一般在此后者范围内。

[0171] $\delta^{13}\text{C}(\text{‰}) = [({}^{13}\text{C}/{}^{12}\text{C})_{\text{样品}} - ({}^{13}\text{C}/{}^{12}\text{C})_{\text{标准品}}] / ({}^{13}\text{C}/{}^{12}\text{C})_{\text{标准品}} \times 1000$

[0172] 已经发展一系列替代RM并与IAEA、USGS、NIST及其它精选国际同位素实验室合作。与PDB的千分偏差的标记是 $\delta^{13}\text{C}$ 。在CO₂上,通过高精度稳定比率质谱分析(IRMS),在质量44、45和46的分子离子上进行测量。本文中描述的组合物包括通过本文中描述的任何方法产生的脂肪二醇组合物和产物。具体地说,脂肪二醇组合物或产物的 $\delta^{13}\text{C}$ 为约-28或更大、约-27或更大、-20或更大、-18或更大、-15或更大、-13或更大、-10或更大或-8或更大。举例来说,脂肪二醇组合物或产物的 $\delta^{13}\text{C}$ 可以为约-30至约-15、约-27至约-19、约-25至约-21、约-15至约-5、约-13至约-7或约-13至约-10。在其它情况下,脂肪二醇组合物或产物的 $\delta^{13}\text{C}$ 可以为约-10、-11、-12或-12.3。根据本公开产生的脂肪二醇组合物和产物还可以通过比较每种化合物中¹⁴C的量而与基于石油的有机化合物相区别。因为¹⁴C具有5730年的核半衰期,所以含有“较古老”碳的基于石油的燃料可以与含有“较新”碳的脂肪二醇组合物和生物产物相区

别(参见例如Currie,“Source Apportionment of Atmospheric Particles”, Characterization of Environmental Particles, J. Buffle和H.P. van Leeuwen编辑, IUPAC Environmental Analytical Chemistry Series (Lewis Publishers, Inc.) 第I卷第1部分3-74, (1992))。

[0173] 放射性碳年代测定中的基本假设是大气中 ^{14}C 浓度的恒久不变导致活生物中 ^{14}C 的恒久不变。然而,由于二十世纪五十年代以来的大气层核试验和十九世纪五十年代以来的化石燃料的燃烧,所以 ^{14}C 已经获得了第二个地球化学时期特征。其在常压 CO_2 中和由此在活生物层中的浓度大约是二十世纪六十年代中期核试验高峰期时的两倍。此后其逐渐回落至约 1.2×10^{-12} 的稳态宇生(常压)基线同位素比率($^{14}\text{C}/^{12}\text{C}$),其中松弛“半衰期”大约是7-10年。此后者半衰期不能取文字意思;而是必须使用详细的常压核输入/衰变函数描绘常压和生物层 ^{14}C 自核时代以来的变化。证实此生物层 ^{14}C 时期特征提供了近代生物层碳的年代测定的希望。 ^{14}C 可以通过加速质谱分析(AMS)测量,其中结果以现代碳的分数的单位(fM)给出。在这方面,fM具有与国家标准和技术研究所(National Institute of Standards and Technology, NIST)标准参考物质(SRM 4990B和4990C)所定义相同的含义,称为草酸标准品 HOxI 和 HOxII 。基本定义涉及 $0.95 \times 14\text{C}/12\text{C}$ 同位素比率 HOxI (参考AD 1950)。此大致相当于衰变校正的工业革命前的木材。对于当前活生物层(植物物质),fM大约是1.1。本文中描述的脂肪二醇组合物和产物包括可以具有至少约1的fM ^{14}C 的生物产物。举例来说,本公开的生物产物可以具有至少约1.01的fM ^{14}C 、约1至约1.5的fM ^{14}C 、约1.04至约1.18的fM ^{14}C 或约1.111至约1.124的fM ^{14}C 。

[0174] ^{14}C 的另一测量值称为现代碳的百分比(pMC)。对于使用 ^{14}C 年代测定的考古学家或地质学家,AD 1950等于零岁。这也表示100pMC。大气中的爆炸碳在1963年热核武器高峰期时几乎是正常水平的两倍。已经粗略估计自从出现以来其在大气内的分配,显示自AD 1950以来活的植物和动物的值超过100pMC。其随时间逐渐减少,其中如今的值接近107.5pMC。这意指例如玉米等新鲜生物质材料将得到接近107.5pMC的 ^{14}C 特征。基于石油的化合物的pMC值为零。化石碳与现代碳组合将导致现代pMC含量的稀释。通过假定107.5pMC表示现代生物质的 ^{14}C 含量且0pMC表示基于石油的产物的 ^{14}C 含量,针对物质所测量的pMC值将反映两种组分类型的比例。举例来说,从现代大豆100%获得的物质将得到接近107.5pMC的放射性碳特征。如果该物质用基于石油的产物稀释50%,那么将得到大约54pMC的放射性碳特征。基于生物的碳含量通过分配100%等于107.5pMC且0%等于0pMC来获得。举例来说,测得99pMC的样品将得到93%的同等的基于生物的碳含量。此值称为平均基于生物的碳结果,且假定分析物质内的所有组分来源于现代生物物质或基于石油的物质。包含如本文中描述的一种或多种脂肪二醇的生物产物可以具有至少约50、60、70、75、80、85、90、95、96、97、98、99或100的pMC。在其它情况下,本文中描述的脂肪二醇组合物可以具有约50与约100、约60与约100、约70与约100、约80与约100、约85与约100、约87与约98或约90与约95之间的pMC。在其它情况下,本文中描述的脂肪二醇组合物可以具有约90、91、92、93、94或94.2的pMC。

[0175] 例如1,3-二醇等脂肪二醇是在许多工业应用中价值且合乎需要的分子。本公开通过重组微生物产生此类化合物,包括体内,且因此产生一系列适用产物。此类产物包括1,3-二醇及其组合物。1,3-二醇的实例包括但不限于 C_5 1,3-二醇(1,3-戊二醇); C_6 1,3-二醇(1,3-己二醇); C_7 1,3-二醇(1,3-庚二醇); C_8 1,3-二醇(1,3-辛二醇); C_9 1,3-二醇(1,3-壬二

醇);C₁₀ 1,3-二醇(1,3-癸二醇);C₁₁ 1,3-二醇(1,3-十一烷二醇);C₁₂ 1,3-二醇(1,3-十二烷二醇);C₁₃ 1,3-二醇(1,3-十三烷二醇);C₁₄ 1,3-二醇(1,3-十四烷二醇);C₁₅ 1,3-二醇(1,3-十五烷二醇);C₁₆ 1,3-二醇(1,3-十六烷二醇);C₁₇ 1,3-二醇(1,3-十七烷二醇);C₁₈ 1,3-二醇(1,3-十八烷二醇);C₁₉ 1,3-二醇(1,3-十九烷二醇)等等。虽然本文中通常描述偶数链1,3-二醇,但也包括奇数链1,3-二醇,例如具有7-21个碳且更优选5-19个碳的1,3-二醇。

[0176] 本公开的1,3-二醇具有各种链长和/或饱和和/或分支特征。在一些实施方案中,1,3-二醇组合物通常包括一种类型1,3-二醇,例如C₅1,3-二醇(1,3-戊二醇);C₆ 1,3-二醇(1,3-己二醇);C₇ 1,3-二醇(1,3-庚二醇);C₈ 1,3-二醇(1,3-辛二醇);C₉ 1,3-二醇(1,3-壬二醇);C₁₀ 1,3-二醇(1,3-癸二醇);C₁₁ 1,3-二醇(1,3-十一烷二醇);C₁₂ 1,3-二醇(1,3-十二烷二醇);C₁₃ 1,3-二醇(1,3-十三烷二醇);C₁₄ 1,3-二醇(1,3-十四烷二醇);C₁₅ 1,3-二醇(1,3-十五烷二醇);C₁₆ 1,3-二醇(1,3-十六烷二醇);C₁₇ 1,3-二醇(1,3-十七烷二醇);C₁₈ 1,3-二醇(1,3-十八烷二醇);C₁₉ 1,3-二醇(1,3-十九烷二醇)等等。在另一个实施方案中,1,3-二醇组合物通常包括特定链长的特定1,3-二醇呈特定比率的混合物。在又一个实施方案中,1,3-二醇组合物包括特定链长的一种或多种1,3-二醇的组合与其它成分或组分组合以产生清洁剂、表面活性剂、乳化剂、软化剂、溶剂、塑料和食品添加剂。

[0177] 在一个实施方案中,脂肪二醇组合物包括C₁₂ 1,3-二醇于直链脂肪醇的混合物中。在另一个实施方案中,脂肪二醇组合物包括C₁₂ 1,3-二醇于支链脂肪醇的混合物中。在另一个实施方案中,脂肪二醇组合物包括C₁₂ 1,3-二醇于直链与支链脂肪醇的混合物中。在另一个实施方案中,脂肪二醇组合物包括C₁₂ 1,3-二醇与例如清洁剂或表面活性剂成分等其它成分组合。在又一个实施方案中,脂肪二醇组合物包括C₁₂ 1,3-二醇与例如乳化剂或溶剂成分等其它成分组合。在又一个实施方案中,脂肪二醇组合物包括C₁₂ 1,3-二醇与聚合物组合。本文中,脂肪二醇组合物可以用作塑料的组分。在另一个实施方案中,脂肪二醇组合物包括C₁₂ 1,3-二醇于食物成分的混合物中。在另一个实施方案中,脂肪二醇组合物包括1,3-二醇作为表面活性剂或清洁剂(例如糖苷或乙醇盐)的合成中的中间物,或与香料或其它化学品可以从中合成的化学构筑嵌段组合。

[0178] 在另一个实施方案中,脂肪二醇组合物包括一定比率的C₈-、C₁₀-和C₁₂ 1,3-二醇于直链脂肪醇的混合物中。在另一个实施方案中,脂肪二醇组合物包括一定比率的C₈-、C₁₀-和C₁₂ 1,3-二醇于支链脂肪醇的混合物中。在另一个实施方案中,脂肪二醇组合物包括一定比率的C₈-、C₁₀-和C₁₂ 1,3-二醇于直链与支链脂肪醇的混合物中。在另一个实施方案中,脂肪二醇组合物包括一定比率的C₈-、C₁₀-和C₁₂ 1,3-二醇与例如清洁剂或表面活性剂成分等其它成分组合。在另一个实施方案中,脂肪二醇组合物包括一定比率的C₈-、C₁₀-和C₁₂ 1,3-二醇与例如乳化剂或溶剂成分等其它成分组合。在又一个实施方案中,脂肪二醇组合物包括一定比率的C₈-、C₁₀-和C₁₂ 1,3-二醇与聚合物组合。本文中,脂肪二醇组合物可以用作塑料的组分。在另一个实施方案中,脂肪二醇组合物包括一定比率的C₈-、C₁₀-和C₁₂ 1,3-二醇于食物成分的混合物中。

[0179] 本公开进一步涵盖一种脂肪二醇组合物,其包括一定比率的C₅-、C₆-、C₇-、C₈-、C₉-、C₁₀-和/或C₁₁ 1,3-二醇于食品相关成分的混合物中。此类脂肪二醇将适用作食物稳定剂、食品强化剂、食品添加剂或食品替代物。本公开的脂肪二醇的化合物和组合物可以调配成制成所需产品,包括清洁剂、表面活性剂、乳化剂、软化剂、溶剂、塑料、食品添加剂等等。在

一个实施方案中，C₁₀-C₁₈ 1,3-二醇预期用作表面活性剂。在另一个实施方案中，1,3-二醇直接使用或作为合成营养物质、药物、农用化学品及其它生物活性分子中的中间物。

[0180] 讨论手性1,3-二醇。

实施例

[0181] 以下实施例进一步说明本公开，但不应视为以任何方式限制其范围。

[0182] 实施例1：培养重组大肠埃希氏菌菌株用于产生1,3-二醇

[0183] 所有实验都从特定微生物菌株的单一菌落或冷冻储备料开始。96孔板中的高通量(HTP)方案如下针对每种菌株一式四份地进行：40μL卢里亚-贝尔塔尼(Luria-Bertani, LB)培养物(来自在96孔板中生长的LB培养物)用来接种360μL LB培养基，接着在32°C振荡下培育3-4小时。将40μL LB种子用来接种360μL Nlim培养基(见下文)。在30-35°C下在32°C下生长2小时后，培养物用IPTG(最终浓度1mM)诱导。接着如果未另外说明，那么培养物在30-35°C下在振荡下培育20小时，随后根据下文详述的标准提取方案提取。摇瓶方案类似进行，除了培养物体积按比例增大，使得最终的产生培养基体积是15ml而非400μl。摇瓶培养基还含有0.25%(v/v) Triton X100。根据微生物菌株，在所有阶段都将适当的抗生素添加至培养基。

Nlim 培养基配方		
1	x	具有 NH ₄ Cl 的 5x 盐溶液
1	x	1000x 微量维生素
1	mg/L	10 mg/mL 硫胺
1	mM	1M MgSO ₄
[0184]	0.1	mM 1M CaCl ₂
40	g/L	500g/L 葡萄糖
1	x	1000x 微量矿物质
10	mg/L	10g/L 柠檬酸铁
100	μg/mL	100 mg/ml 壮观霉素 (spectinomycin)
100	mM	2M BisTris (pH7.0)

[0185] 生物反应器中的基线工艺如下：将菌株的细胞库小瓶在含有抗生素的LB摇瓶中在32°C下培养，直到培养物的OD读数>1。将5%v/v的此培养物转移至最低种子培养基(含有氯化铵、氯化钠、磷酸二氢钾、硫酸镁、氯化钙、葡萄糖、微量元素溶液、柠檬酸铁一水化物、缓冲剂和抗生素)，并在32°C下培养整夜。接着此种子培养物用于接种准备好的用于生产的生物反应器。

[0186] 用于此工艺的初始生物反应器培养基含有各种浓度的与种子培养基相同的组分，以及微量维生素溶液，和任选少量的复杂培养基组分，例如酪蛋白氨基酸、玉米浆粉或酵母提取物。生物反应器的后无菌添加任选地包括不耐热的维生素或氨基酸、葡萄糖和抗生素。

[0187] 在接种前，将生物反应器参数稳定化，且打开控制器-溶解氧设定：10-50%；温度设定：27-37°C；通风设定：0.25-1vvm；pH设定：6.5-7.5。生物反应器用5%v/v种子

培养物接种并当培养物密度达到所需设定点时用1mM IPTG诱导。

[0188] 由葡萄糖、蔗糖、果糖、木糖或甘油与并入生物反应器基础培养基中的其它可能培养基组分构成的饲料溶液以1-50g/L/小时葡萄糖(基于标称培养物体积)的最大速率馈送至培养物,使用DO或pH触发器,在培养基的碳源耗尽且应添加下一剂饲料溶液时向控制器作出指示。在培养48小时与96小时之间收获生物反应器。

[0189] 实施例2:分析1,3-二醇

[0190] 利用以下程序提取通过重组大肠埃希氏菌菌株产生的发酵肉汤样品:

[0191] 1.在称重前将肉汤以3000rpm涡旋30秒

[0192] 2.在Vortex Genie上涡旋后,立即取500 μ L肉汤样品

[0193] 3.添加5mL 500mg/L(1-十一烷醇)乙酸丁酯作为内标

[0194] 4.在涡旋机(DVX-2500多管涡旋机,VWR)中在2500rpm下提取肉汤,历时20分钟

[0195] 5.在室温下将提取物离心(在4750rpm下)10分钟

[0196] 6.将100 μ L顶层清液经插入物移至GC小瓶中

[0197] 7.在室温下通过添加100 μ L(BSTFA+1%TMCS)至GC小瓶而衍生化

[0198] 8.将提取物和BSTFA试剂混合30秒,接着注射在如下所述的GC/MS上:

[0199] 用于鉴别的仪器条件

[0200] 起始温度:60 $^{\circ}$ C

[0201] 起始时间:5分钟

[0202] 平衡时间:1分钟

[0203] 程序速率:25 $^{\circ}$ C/分钟

[0204] 最终温度:300 $^{\circ}$ C

[0205] 最终时间:1.6分钟

[0206] 检测器:MSD

[0207] 入口温度:300 $^{\circ}$ C

[0208] 转移管线温度:300 $^{\circ}$ C

[0209] MS源:230 $^{\circ}$ C

[0210] MS Quad:150 $^{\circ}$ C

[0211] 分流比:20:1

[0212] 柱流速:1mL/分钟

[0213] 样品大小:1 μ L

[0214] 实施例3:使用具有来自四联厌氧球菌或植物乳杆菌的TE和carB的途径产生1,3-二醇

[0215] 此实施例展示了使用包括来自四联厌氧球菌(TE_EEI82564,genbank登录号WP_004837416)或植物乳杆菌(TE_CAD63310,genbank登录号WP_003640969)的微生物硫酯酶和来自解皂菌状杆菌(genbank登录号YP_889972)的羧酸还原酶的变体CarB的代谢途径,1,3-二醇在重组大肠埃希氏菌中的意外产生。

[0216] 编码carB2(SEQ ID NO:6)和TE_EEI82564的基因克隆至pCL1920-衍生载体(SC101复制子,壮观霉素抗性标记物),使得其转录由IPTG-诱导性Ptrc启动子控制,且其与编码醇脱氢酶alrA(其并非脂肪醇或脂肪二醇产生所需,但提高其产生速率)和3-酮酰基-ACP合酶

(fabB)的变异体和转录调节子(fadR)的基因一起形成操纵子。质粒命名为pVA369(参见表5)。来自植物乳杆菌的TE_CAD63310的基因以一致方式,与carB12(SEQ ID NO:4)的基因一起克隆,所得质粒命名为pJP2(参见下表5)。

[0217] 用于质粒转型的基础菌株是V668和DJ81。简单地说,基础菌株的基因组如下操纵:在V668中,fadE(酰基-CoA脱氢酶)基因缺失,且合成脂肪酸生物合成操纵子和磷酸泛酰巯基乙氨基转移酶(entD)过度表达。简单地说,基础菌株DJ81的基因组如下操纵:酰基-CoA脱氢酶(fadE)基因缺失,且合成脂肪酸生物合成操纵子、磷酸泛酰巯基乙氨基转移酶(entD)和变异硫酯酶(tesA)过度表达。

[0218] 质粒pVA369和pJP2分别转化至基础菌株D848和V668中,产生菌株VA370和JP-11(参见下文表6)。接着如实施例1和2中所述培养菌株并分析其产生脂肪醇的能力。意外地,两种菌株产生若干未知峰。

[0219] 图4展示表达TE_EEI82564的菌株VA370的提取物的GC-MS色谱。GC-MS色谱中RT=8.199min和RT=9.094min的两个峰不匹配例如十二烷醇和十四烷醇或十二烷酸和十四烷酸等预期脂肪醇和脂肪酸的保留时间。峰1在十二烷醇前流出,且峰2在十四烷醇后和十四烷酸前流出。峰1和2的离子碎裂图案(参见图5)表明此两个峰是1,3-三甲基硅氧烷基辛烷和1,3-三甲基硅氧烷基癸烷,是BSTFA(参见实施例2)分别与1,3-辛二醇和1,3-癸二醇的衍生化产物。为了说明,图6展示如在图5中峰1中观测到的1,3三甲基硅氧烷基癸烷的离子片段的示意图。还观测到微量的衍生化1,3-十二烷二醇和1,3-十四烷二醇。

[0220] 类似地,表达TE_CAD63310的菌株JP-11的提取物含有新峰,如上所述,基于其离子碎裂图案和保留时间,鉴别为1,3-辛二醇、1,3-癸二醇、1,3-十四烷醇和1,3-十四碳烯醇。在HTP发酵方案中,JP-11产生总共 1.9 ± 0.05 g/L 1,3-二醇,以及脂肪醇,例如辛醇、癸醇、十二烷醇、十二碳烯醇、十四醇和十四碳烯醇。菌株JP-11的产物分配展示于图7中。

[0221] 1,3-二醇的产生由于以下两个原因而为意外的:(i)其需要3-OH脂肪酸作为中间物,3-OH脂肪酸很可能在硫酯酶的作用下来源于3-OH酰基-ACP(参见图2)。用于此实施例中的两种硫酯酶先前已在大肠埃希氏菌中表达,且仅仅报道其产生脂肪酸(Jing等人BMC Biochemistry 2011,12:44),表明其不适合产生3-OH脂肪酸,因此不适合产生1,3-二醇,(ii)其需要3-OH脂肪酸中间物,通过羧酸还原酶CarB还原成3-OH脂肪醛,接着通过醇脱氢酶(ADH)进一步还原成1,3-二醇。虽然已知ADH相对混杂,但CarB先前未展示将3-OH脂肪酸转变成3-OH脂肪醛。因此,申请人发现某些微生物硫酯酶,例如TE_EEI82564和TE_CAD63310,在与CarB一起在大肠埃希氏菌宿主细胞中共同表达时过度产生1,3-二醇。1,3-二醇的分析可展示其高度富含(R)对映异构体,证实天然大肠埃希氏菌脂肪酸生物合成机制的3-酮酰基ACP还原酶(FabG)的对映异构体选择性。

[0222] 表5:用于产生1,3-二醇的质粒

质粒名称	途径基因
pVA369	<i>pCL-carB2-TE_EEI82564alrA-fabB-fadR</i>
pJP2	<i>pCL-carB12-TE_CAD63310-alrA-fabB-fadR</i>
pNH328	<i>pCL-trc-carB8-phaG-alrA</i>
pHN330	<i>pCL-trc-carB8-fatB1-alrA</i>

[0224]

pNT16	<i>pCL-trc-AAR-alsA</i>
-------	-------------------------

[0225] 表6:用于产生1,3-二醇的菌株

	菌株名称	描述
	VA370	具有质粒 <i>pVA369</i> 的 <i>V668</i>
	JP-11	具有质粒 <i>pJP2</i> 的 <i>DJ81</i>
[0226]	stNH1371	具有质粒 <i>pNH330</i> 的 <i>D178</i>
	stNH1369	具有质粒 <i>pNH328</i> 的 <i>D178</i>
	Becos247	具有质粒 <i>pNT16</i> 的 <i>DV2</i>

[0227] 实施例4A:使用具有来自加州月桂的 *fatB1* 和 *carB* 的途径产生1,3-二醇

[0228] 此实施例展示了使用包括来自加州月桂 (genbank登录号Q41635) 的植物硫酸酯酶 *fatB1* 和来自解皂菌状杆菌的变异羧酸还原酶 *CarB* 的代谢途径,1,3-二醇在重组大肠埃希氏菌中的产生。

[0229] 编码 *carB8* (SEQ ID NO:8) 和 *fatB1* 的基因克隆至 *pCL1920*-衍生载体 (SC101复制子,壮观霉素抗性标记物),使得其转录由 IPTG-诱导性 *P_{trc}* 启动子控制,且其与编码醇脱氢酶 *alsA* 的基因一起形成操纵子。质粒命名为 *pNH330* (参见表5)。

[0230] 用于质粒转化的基础菌株是菌株 *D178*。简单地说,菌株 *D178* 的基因组如下修饰:*fadE* (酰基-CoA脱氢酶) 基因缺失,且磷酸泛酰巯基乙氨基转移酶 (*entD*) 过度表达。质粒 *pNH330* 转化至 *D178* 中,得到菌株 *stNH1371* (参见表6)。接着如实施例1和2中所述培养菌株并分析其产生脂肪醇和1,3-二醇的能力。1,3-二醇峰如实施例2中所述鉴别。

[0231] 在HTP发酵方案中菌株 *stNH1371* 产生 39.5 ± 3.2 mg/L 1,3-二醇。1,3-十二烷二醇是产生的1,3-二醇之一。除1,3-二醇外,还检出例如十二烷醇等脂肪醇。1,3,二二醇的分析可展示其高度富含 (R) 对映异构体,证实天然大肠埃希氏菌脂肪酸生物合成机制的3-酮酰基 ACP 还原酶 (*FabG*) 的对映异构体选择性。

[0232] 实施例4B:使用具有来自恶臭假单胞菌的 *phaG* 和 *carB* 的途径产生1,3-二醇

[0233] 此实施例展示了使用包括来自恶臭假单胞菌 (genbank登录号AAN67031) 的硫酸酯酶/酰基转移酶和来自解皂菌状杆菌的变异羧酸还原酶 *CarB* 的代谢途径,1,3-二醇在重组大肠埃希氏菌中的产生。

[0234] 编码 *carB8* (参见下文与此一同封装的序列列表) 和 *phaG* 的基因克隆至 *pCL1920*-衍生载体 (SC101复制子,壮观霉素抗性标记物),使得其转录由 IPTG-诱导性 *P_{trc}* 启动子控制,且其与编码醇脱氢酶 *alsA* 的基因一起形成操纵子。质粒命名为 *pNH328* (参见表5)。

[0235] 用于质粒转化的基础菌株是菌株 *D178*。简单地说,菌株 *D178* 的基因组如下修饰:*fadE* (酰基-CoA脱氢酶) 基因缺失,且磷酸泛酰巯基乙氨基转移酶 (*entD*) 过度表达。质粒 *pNH328* 转化至 *D178* 中,得到菌株 *stNH1369* (参见上文表6)。接着如实施例1和2中所述培养菌株并分析其产生脂肪醇和1,3-二醇的能力。1,3-二醇峰如实施例2中所述鉴别。

[0236] 在HTP发酵方案中菌株 *stNH1369* 产生 600 ± 27 mg/L 1,3-二醇。所产生的1,3-二醇是1,3-辛二醇、1,3-癸二醇、1,3-十二烷二醇和1,3-十四烷二醇。除1,3-二醇外,仅检出较

少量的脂肪酸。1,3,二醇的分析可展示其高度富含(R)对映异构体,证实天然大肠埃希氏菌脂肪酸生物合成机制的3-酮酰基ACP还原酶(FabG)的对映异构体选择性。

[0237] 实施例5:使用具有来自细长聚球藻的AAR的途径产生1,3-二醇

[0238] 此实施例展示了使用包括来自细长聚球藻(genbank登录号YP_400611;野生型)的变异酰基-ACP还原酶AAR的代谢途径,1,3-二醇在重组大肠埃希氏菌中的产生。关于变异AAR序列参见与此一起封装的序列列表(下文)。

[0239] 编码AAR变异体(SEQ ID NO:2)的基因克隆至pCL1920-衍生载体(SC101复制子,壮观霉素抗性标记物),使得其转录由IPTG诱导性Ptrc启动子控制,且其与编码醇脱氢酶alrA的基因一起形成操纵子。质粒命名为pNT 16(参见表5)。

[0240] 用于质粒转化的基础菌株是DV2。简单地说,通过使fadE(酰基-CoA脱氢酶)基因缺失来操纵菌株DV2的基因组。质粒pNT16转化至基础菌株DV2中,得到菌株Becos 247(参见表6)。接着如实施例1和2中所述培养Becos247并分析其产生脂肪醇的能力。意外地,菌株产生1,3-二醇。1,3-二醇峰如实施例2中所述鉴别。

[0241] 菌株Becos247在5L发酵中产生0.57g/L 1,3-二醇,其占所产生的总脂肪酸物质的9.1%。所产生的1,3-二醇是1,3-十二烷二醇、1,3-十四烯二醇和1,3-十四烷二醇,另外,产生脂肪醇癸醇、十二碳烯醇、十二烷醇、十四碳烯醇、十四烷醇、十六碳烯醇、十六烷醇和十八碳烯醇以及少量的脂肪酸。

[0242] 此实验中经由3-OH脂肪醛作为中间物产生1,3-二醇是意外的(参见图3),因为用于此实施例的酰基-ACP还原酶,来自细长聚球藻的野生型AAR先前已经在大肠埃希氏菌中表达,且仅仅报道从酰基-ACP产生脂肪醇,而非从3-OH酰基-ACP产生1,3-二醇(Schirmer等人(2010)Science 329,559)。1,3,二醇的分析可展示其高度富含®对映异构体,证实天然大肠埃希氏菌脂肪酸生物合成机制的3-酮酰基ACP还原酶(FabG)的对映异构体选择性。

[0243] 实施例6A:使用具有来自加州月桂的fatB1和carB的途径产生1,3-二醇

[0244] 此实施例描述了如何使用包括来自加州月桂的植物硫酯酶fatB1和来自解皂菌状杆菌的羧酸还原酶carB的代谢途径,去证实1,3-二醇在重组大肠埃希氏菌中的产生。

[0245] 编码野生型carB和fatB1的基因克隆至pCL1920-衍生载体(SC101复制子,壮观霉素抗性标记物),使得其转录由IPTG-诱导性Ptrc启动子控制,且其与编码醇脱氢酶alrA的基因一起形成操纵子。质粒转化至例如菌株DV2等基础菌株(参见实施例5)。

[0246] 接着如实施例1和2中所述培养所得菌株并分析其产生脂肪醇和二醇的能力。预期菌株产生1,3-二醇。1,3,二醇的分析展示其高度富含(R)对映异构体,证实天然大肠埃希氏菌脂肪酸生物合成机制的3-酮酰基ACP还原酶(FabG)的对映异构体选择性。

[0247] 实施例6B:使用具有来自加州月桂的fatB1和carB的简化途径产生1,3-二醇

[0248] 此实施例描述了如何使用包括来自加州月桂的植物硫酯酶fatB1和来自解皂菌状杆菌的羧酸还原酶carB,的简化代谢途径,去证实1,3-二醇在重组大肠埃希氏菌中的产生。

[0249] 编码野生型carB和fatB1的基因克隆至pCL1920-衍生载体(SC101复制子,壮观霉素抗性标记物),使得其转录由IPTG-诱导性Ptrc启动子控制。质粒转化至例如菌株DV2等基础菌株(参见实施例5)。

[0250] 接着如实施例1和2中所述培养所得菌株并分析其产生脂肪醇和二醇的能力。预期菌株产生1,3-二醇,这证实硫酯酶和羧酸还原酶足以使微生物细胞能够产生1,3二醇。1,3,

二醇的分析展示其高度富含(R)对映异构体,证实天然大肠埃希氏菌脂肪酸生物合成机制的3-酮酰基ACP还原酶(FabG)的对映异构体选择性。

[0251] 实施例7:使用具有来自大肠埃希氏菌的tesA和carB的途径产生1,3-二醇

[0252] 此实施例描述了如何使用包括硫酯酶tesA和来自解皂菌状杆菌的羧酸还原酶CarB的代谢途径,去证实1,3-二醇在重组大肠埃希氏菌中的产生。

[0253] 编码野生型carB和野生型tesA的基因克隆至pCL1920-衍生载体(SC101复制子,壮观霉素抗性标记物),使得其转录由IPTG-诱导性Ptrc启动子控制,且其与编码醇脱氢酶alrA的基因一起形成操纵子。质粒转化至例如菌株DV2等基础菌株(参见实施例5)。

[0254] 接着如实施例1和2中所述培养所得菌株并分析其产生脂肪醇和二醇的能力。预期菌株产生1,3-二醇,这证实硫酯酶和羧酸还原酶足以使微生物细胞能够产生1,3-二醇。1,3-二醇的分析展示其高度富含(R)对映异构体,证实天然大肠埃希氏菌脂肪酸生物合成机制的3-酮酰基ACP还原酶(FabG)的对映异构体选择性。

[0255] 实施例8:使用具有来自细长聚球藻的野生型AAR的简化途径产生1,3-二醇

[0256] 此实施例描述了如何使用包括来自细长聚球藻的酰基-ACP还原酶AAR的代谢途径,去证实1,3-二醇在重组大肠埃希氏菌中的产生。

[0257] 编码野生型AAR的基因克隆至pCL1920-衍生载体(SC101复制子,壮观霉素抗性标记物),使得其转录由IPTG诱导性Ptrc启动子控制。质粒转化至例如菌株DV2等基础菌株(参见实施例5)。

[0258] 接着如实施例1和2中所述培养所得菌株并分析其产生脂肪醇和二醇的能力。预期菌株产生1,3-二醇,这证实AAR的异源产生足以使微生物细胞能够产生1,3-二醇。1,3-二醇的分析展示其高度富含(R)对映异构体,证实天然大肠埃希氏菌脂肪酸生物合成机制的3-酮酰基ACP还原酶(FabG)的对映异构体选择性。

[0259] 实施例9:使用具有来自樟树的fatB和carB的途径产生1,3-二醇

[0260] 此实施例描述了如何使用包括来自樟树的植物硫酯酶fatB和来自解皂菌状杆菌的羧酸还原酶carB的代谢途径,去证实1,3-二醇在重组大肠埃希氏菌中的产生。

[0261] 编码野生型carB和来自樟树的fatB的基因克隆至pCL1920-衍生载体(SC101复制子,壮观霉素抗性标记物),使得其转录由IPTG-诱导性Ptrc启动子控制,且其与编码醇脱氢酶alrA的基因一起形成操纵子。质粒转化至例如菌株DV2等基础菌株(参见实施例5)。

[0262] 接着如实施例1和2中所述培养所得菌株并分析其产生脂肪醇和二醇的能力。预期菌株产生1,3-二醇,这证实硫酯酶和羧酸还原酶足以使微生物细胞能够产生1,3-二醇。1,3-二醇的分析展示其高度富含(R)对映异构体,证实天然大肠埃希氏菌脂肪酸生物合成机制的3-酮酰基ACP还原酶(FabG)的对映异构体选择性。

[0263] 实施例10:使用具有来自贝氏不动细菌(Acinetobacter baylyi)的acr1的途径产生1,3-二醇

[0264] 此实施例描述了如何使用包括来自贝氏不动细菌(genbank登录号AAC45217)的脂肪酰基-CoA还原酶acr1的代谢途径,去证实1,3-二醇在重组大肠埃希氏菌中的产生。

[0265] 编码acr1的基因克隆至pCL1920-衍生载体(SC101复制子,壮观霉素抗性标记物),使得其转录由IPTG-诱导性Ptrc启动子控制,且其与编码酰基-CoA合成酶(fadD)和硫酯酶的基因一起形成操纵子。质粒转化至例如菌株DV2等基础菌株(参见实施例5)。

[0266] 接着如实施例1和2中所述培养所得菌株并分析其产生脂肪醇和二醇的能力。预期菌株产生1,3-二醇。

[0267] 实施例11:使用具有来自水油海杆菌 (*Marinobacter aquaeolei*) 的FAR的途径产生1,3-二醇

[0268] 此实施例描述了如何使用包括来自水油海杆菌 (genbank登录号YP_959486) 的脂肪酰基-ACP还原酶FAR的代谢途径,去证实1,3-二醇在重组大肠埃希氏菌中的产生。

[0269] 编码野生型FAR的基因克隆至pCL1920-衍生载体 (SC101复制子,壮观霉素抗性标记物),使得其转录由IPTG-诱导性Ptrc启动子控制。质粒转化至例如菌株DV2等基础菌株 (参见实施例5)。

[0270] 接着如实施例1和2中所述培养所得菌株并分析其产生脂肪醇和二醇的能力。预期菌株产生1,3-二醇,这证实异源产生的FAR足以使微生物细胞能够产生1,3-二醇。

[0271] 实施例12:使用具有来自发光杆菌 (*Photobacterium luminescens*) 的FAR复合物的途径产生1,3-二醇

[0272] 此实施例描述了如何使用包括来自发光杆菌 (genbank登录号AHH25015-17) 的脂肪酰基-ACP还原酶FAR复合物 (包括LuxC、LuxD和LuxE) 的代谢途径,去证实1,3-二醇在重组大肠埃希氏菌中的产生。

[0273] 编码LuxC、LuxD和LuxE的基因克隆至pCL1920-衍生载体 (SC101复制子,壮观霉素抗性标记物),使得其转录由IPTG-诱导性Ptrc启动子控制,且其与编码醇脱氢酶alrA的基因一起形成操纵子。质粒转化至例如菌株DV2等基础菌株 (参见实施例5)。

[0274] 接着如实施例1和2中所述培养所得菌株并分析其产生脂肪醇和二醇的能力。预期菌株产生1,3-二醇。

[0275] 实施例13:使用fadB (His450Gln) 产生3-(S)-脂肪二醇

[0276] 此实施例描述了如何使用包括保留烯酰基-CoA水合酶活性但是缺乏脱氢酶活性且表达来自贝氏不动细菌 (genbank登录号AAC45217) 的脂肪酰基-CoA还原酶acr1的3-羟基-酰基-ACP酰基-CoA酰基转移酶或硫酯酶fadB (His450Gln) 的代谢途径,去证实3-(S)-脂肪二醇在重组大肠埃希氏菌中的产生。

[0277] 编码TesA、FadD、FadB (His450Gln) 和Acr1的基因克隆至pCL1920-衍生载体 (SC101复制子,壮观霉素抗性标记物),使得其转录由IPTG-诱导性Ptrc启动子控制,且其实现足以合成脂肪醇的操纵子。质粒转化至例如菌株MG1655等基础菌株 (参见实施例5),其中编码FadE的额外基因已在IPTG诱导性Ptrc启动子控制下引入基因组中。

[0278] 接着如实施例1和2中所述培养所得菌株并分析其产生脂肪醇和二醇的能力。预期菌株产生3-(S)-脂肪二醇。

[0279] 实施例14. 使用fadB (Glu119Gln) 产生3-(S)-脂肪二醇

[0280] 此实施例描述了如何使用包括保留脱氢酶活性但是缺乏脱水酶活性且表达来自贝氏不动细菌 (genbank登录号AAC45217) 的脂肪酰基-CoA还原酶acr1的3-羟基-酰基-ACP酰基-CoA酰基转移酶或硫酯酶fadB (Glu119Gln) 的代谢途径,去证实3-(S)-脂肪二醇在重组大肠埃希氏菌中的产生。

[0281] 编码TesA、FadD、FadB (Glu119Gln) 和Acr1的基因克隆至pCL1920-衍生载体 (SC101复制子,壮观霉素抗性标记物),使得其转录由IPTG-诱导性Ptrc启动子控制,且其实现足以

合成脂肪醇的操纵子。质粒转化至例如菌株MG1655等基础菌株(参见实施例5),其中编码FadA的额外基因已在IPTG诱导性Ptrc启动子控制下引入基因组。

[0282] 接着如实施例1和2中所述培养所得菌株并分析其产生脂肪醇和二醇的能力。预期菌株产生3-(S)-脂肪二醇。

[0283] 如本领域技术人员显而易见,在不脱离本公开的精神和范围下可以对以上方面和实施方案进行各种修改和改变。此类修改和改变在本公开的范围內。

序列表

- <110> 基因组股份公司
 <120> 脂肪二醇的微生物生产
 <130> IIC201867
 <140>
 <141>
 <150> 62/026,573
 <151> 2014-07-18
 <160> 8
 <170> PatentIn 3.5版
 <210> 1
 <211> 342
 <212> PRT
 <213> 人工序列
 <220>
 <221> 来源
 <223> /注释="人工序列的描述:合成AAR变异多肽"
 <400> 1

```

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

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

<210> 2

<211> 1029

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释="人工序列的描述:合成AAR变异多核苷酸"

<400> 2

atggcattcg gtcttatcgg tcatgcaacc agtttggagc aggcccgcga cgtttggcgc 60
 aggctgggct acgacgaata cgccgatcaa ggattggagt tttggagtag cgctcctcct 120
 caaatcgttg atgaaatcac agtcaccagt gccacaggca aggtgattca cggtcgctac 180
 atcgaatcgg ggttcttgcc ggaaatgctg gcggcgcgcc gttcaaac agcaacgcgc 240
 aaagttctca atgcatgtc ccatgccc aaacacggca tcgacatctc ggccttgggg 300
 ggctttacct cgattatctt cgagaatttc gatttggcca agttgcggca agtgcgcgac 360
 actaccttgg agtttgaacg gttcaccacc ggcaatactc acacggccta cgtaatctgt 420

agacaggtgg aagccgctgc taaaacgctg ggcacgaca ttgcgcaagc gacagtagcg 480
 gttgtcggcg cgactggcga tatcggtagc gctgtctgcc gctggctcga cctcaactg 540
 ggtgtcggtg atttgatcct gacggcgcgc aatcaggagc gtttgataa cctgcaggct 600
 gaactcggcc ggggcaagat tctgcccttg gaagccgctc tgccggaagc tgactttatc 660
 gtgtgggtcg ccagtatgcc tcagggcgta gtgatcgacc cagcaaccct gaagcaacc 720
 tgcgtcctaa tcgacggggg ctacccaaa aacttgggca gcaaagtcca aggtgagggc 780
 atctatgtcc tcaatggcgg ggtagttaa cattgcttcg acatcgactg gcagatcatg 840
 tccttggcag agatggcgcg gcccgagcgc cagatgtttg cctgctttgc cgaggcgatg 900
 ctcttggaat ttgaaggctg gcatactaac ttctctggg gccgcaacca aatcacgatc 960
 gagaagatgg aagcgatcgg tgaggcatcg gtgcgccacg gcttccaacc cttggcattg 1020
 gcaatttga 1029

<210> 3

<211> 1174

<212> PRT

<213> 人工序列

<220>

<221> 来源

<223> /注释="人工序列的描述:合成CAR变异多肽"

<400> 3

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

Ala Pro Val Ser Arg Leu Ala Pro Ile Leu Ala Glu Val Glu Pro Arg
 165 170 175
 Ile Leu Thr Val Ser Ala Glu Tyr Leu Asp Leu Ala Val Glu Ser Val
 180 185 190
 Arg Asp Val Asn Ser Val Ser Gln Leu Val Val Phe Asp His His Pro
 195 200 205
 Glu Val Asp Asp His Arg Asp Ala Leu Ala Arg Ala Arg Glu Gln Leu
 210 215 220
 Ala Gly Lys Gly Ile Ala Val Thr Thr Leu Asp Ala Ile Ala Asp Glu
 225 230 235 240
 Gly Ala Gly Leu Pro Ala Glu Pro Ile Tyr Thr Ala Asp His Asp Gln
 245 250 255
 Arg Leu Ala Met Ile Leu Tyr Thr Ser Gly Ser Thr Gly Ala Pro Lys
 260 265 270
 Gly Ala Met Tyr Thr Glu Ala Met Val Ala Arg Leu Trp Thr Met Ser
 275 280 285
 Gly Ile Thr Gly Asp Pro Thr Pro Val Ile Asn Val Asn Phe Met Pro
 290 295 300
 Leu Asn His Leu Gly Gly Arg Ile Pro Ile Ser Thr Ala Val Gln Asn
 305 310 315 320
 Gly Gly Thr Ser Tyr Phe Val Pro Glu Ser Asp Met Ser Thr Leu Phe
 325 330 335
 Glu Asp Leu Ala Leu Val Arg Pro Thr Glu Leu Gly Leu Val Pro Arg
 340 345 350
 Val Ala Asp Met Leu Tyr Gln His His Leu Ala Thr Val Asp Arg Leu
 355 360 365
 Val Thr Gln Gly Ala Asp Glu Leu Thr Ala Glu Lys Gln Ala Gly Ala
 370 375 380
 Glu Leu Arg Glu Gln Val Leu Gly Gly Arg Val Ile Thr Gly Phe Val
 385 390 395 400
 Ser Thr Ala Pro Leu Ala Ala Glu Met Arg Ala Phe Leu Asp Ile Thr
 405 410 415
 Leu Gly Ala His Ile Val Asp Gly Tyr Gly Leu Thr Glu Thr Gly Ala
 420 425 430
 Val Thr Arg Asp Gly Val Ile Val Arg Pro Pro Val Ile Asp Tyr Lys
 435 440 445
 Leu Ile Asp Val Pro Glu Leu Gly Tyr Phe Ser Thr Asp Lys Pro Tyr
 450 455 460
 Pro Arg Gly Glu Leu Leu Val Arg Ser Ile Thr Leu Thr Pro Gly Tyr

465	470	475	480
Tyr Lys Arg Pro Glu Val Thr Ala Ser Val Phe Asp Arg Asp Gly Tyr			
	485	490	495
Tyr His Thr Gly Asp Val Met Ala Glu Thr Ala Pro Asp His Leu Val			
	500	505	510
Tyr Val Asp Arg Arg Asn Asn Val Leu Lys Leu Ala Gln Gly Glu Phe			
	515	520	525
Val Ala Val Ala Asn Leu Glu Ser Val Phe Ser Gly Ala Ala Leu Val			
	530	535	540
Arg Gln Ile Phe Val Tyr Gly Asn Ser Glu Arg Ser Phe Leu Leu Ala			
545	550	555	560
Val Val Val Pro Thr Pro Glu Ala Leu Glu Gln Tyr Asp Pro Ala Ala			
	565	570	575
Leu Lys Ala Ala Leu Ala Asp Ser Leu Gln Arg Thr Ala Arg Asp Ala			
	580	585	590
Glu Leu Gln Ser Tyr Glu Val Pro Ala Asp Phe Ile Val Glu Thr Glu			
	595	600	605
Pro Phe Ser Ala Ala Asn Gly Leu Leu Ser Gly Val Gly Lys Leu Leu			
	610	615	620
Arg Pro Asn Leu Lys Asp Arg Tyr Gly Gln Arg Leu Glu Gln Met Tyr			
625	630	635	640
Ala Asp Ile Ala Ala Thr Gln Ala Asn Gln Leu Arg Glu Leu Arg Arg			
	645	650	655
Ala Ala Ala Thr Gln Pro Val Ile Asp Thr Leu Thr Gln Ala Ala Ala			
	660	665	670
Thr Ile Leu Gly Thr Gly Ser Glu Val Ala Ser Asp Ala His Phe Thr			
	675	680	685
Asp Leu Gly Gly Asp Ser Leu Ser Ala Leu Thr Leu Ser Asn Leu Leu			
	690	695	700
Ser Asp Phe Phe Gly Phe Glu Val Pro Val Gly Thr Ile Val Asn Pro			
705	710	715	720
Ala Thr Asn Leu Ala Gln Leu Ala Gln His Ile Glu Ala Gln Arg Thr			
	725	730	735
Ala Gly Asp Arg Arg Pro Ser Phe Thr Thr Val His Gly Ala Asp Ala			
	740	745	750
Thr Glu Ile Arg Ala Ser Glu Leu Thr Leu Asp Lys Phe Ile Asp Ala			
	755	760	765
Glu Thr Leu Arg Ala Ala Pro Gly Leu Pro Lys Val Thr Thr Glu Pro			
	770	775	780

1085	1090	1095
Phe Glu Thr Ala Leu Thr	Ala Leu Pro Glu Lys	Arg Arg Ala Gln
1100	1105	1110
Thr Val Leu Pro Leu Leu	His Ala Phe Arg Ala	Pro Gln Ala Pro
1115	1120	1125
Leu Arg Gly Ala Pro Glu	Pro Thr Glu Val Phe	His Ala Ala Val
1130	1135	1140
Arg Thr Ala Lys Val Gly	Pro Gly Asp Ile Pro	His Leu Asp Glu
1145	1150	1155
Ala Leu Ile Asp Lys Tyr	Ile Arg Asp Leu Arg	Glu Phe Gly Leu
1160	1165	1170

Ile

<210> 4

<211> 3525

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释="人工序列的描述:合成CAR变异(carB12)多核苷酸"

<400> 4

```

atgggcacga gcgatgttca cgacgcgacc gacggcgta ccgagactgc actgcgtgat 60
cgccagcgca ctcgtcgtat tgcagaactg tacgcaacgg acccagagtt cgcagcagca 120
gctcctctgc cggccgttgt cgatgcggcg cacaaaccgg gcctgcgtct ggcggaaatc 180
ctgcagaccc tgttcaccgg ctacggcgat cgtccggcgc tgggctatcg tgcacgtgag 240
ctggcgacgg acgaaggcgg tcgtacggtc acgcgtctgc tgccgcgctt cgataccctg 300
acctatgcac aggtgtggag ccgtgttcaa gcagtggctg cagcgttgcg tcacaatttc 360
gcacaaccga tttaccggg cgacgcggtc gcgactatcg gctttgcgag cccggactat 420
ttgacgctgg atctggtgtg cgcgtatctg ggcttggtca gcgttccttt gcagcataac 480
gctccggtgt ctcgcctggc cccgattctg gccgaggtg aaccgcgtat tctgacggtg 540
agcgcagaat acctggacct ggcggttgaa tccgtccgtg atgtgaactc cgtcagccag 600
ctggttgttt tcgaccatca tccggaagtg gacgateacc gtgacgact ggctcgcgca 660
cgcgagcagc tggccggcaa agtatcgca gttacgacc tggatgcgat cgcagacgaa 720
ggcgcaggtt tgccggtga gccgatttac acggcggatc acgatcagcg tctggccatg 780
attctgtata ccagcggtc tacgggtgct ccgaaaggcg cgatgtacac cgaagcgatg 840
gtggctcgcc tgtggactat gagcgggatc acgggcgacc cgaccccggt tatcaacgtg 900
aacttcatgc cgctgaacca tctgggcggt cgtatccga ttagcaccgc cgtgcagaat 960
ggcggtagca gctacttctg tccgaaagc gacatgagca cgctgtttga ggatctggcc 1020
ctgggtccgc ctaccgaact gggctctggtg ccgcgtgttg cggacatgct gtaccagcat 1080
catctggcga ccgtggatcg cctggtgacc cagggcgcgg acgaactgac tgcggaaaag 1140

```

caggccggtg cggaactgcg tgaacaggtc ttgggcggtc gtgttatcac cggttttggt 1200
tccaccgcgc cgttggcggc agagatgcgt gcttttctgg atatcacctt gggatgcacac 1260
atcgttgacg gttacggctt gaccgaaacc ggtgcgggtc cccgtgatgg tgtgattggt 1320
cgtcctccgg tcattgatta caagctgac gatgtgccgg agctgggtta cttctccacc 1380
gacaaaccgt acccgcgtgg cgagctgctg gttcgtagca tcacgttgac tccgggttac 1440
tacaagcgcc cagaagtcac cgcgtccgtt ttcgatcgcg acggctatta ccacaccggc 1500
gacgtgatgg cagaaaccgc gccagaccac ctgggtgatg tggaccgccc caacaatggt 1560
ctgaagctgg cgcaaggtga atttgtcgcc gtggctaacc tggagtccgt tttcagcggc 1620
gctgctctgg tccgccagat tttcgtgtat ggtaacacgc agcgcagctt tctgttggct 1680
gttgttgctc ctaccccgga ggcgctggag caatacgacc ctgccgatt gaaagcagcc 1740
ctggcggatt cgctgcagcg tacggcgcgt gatgccgagc tgcagagcta tgaagtgccg 1800
gcggaactca ttgttgagac tgagcctttt agcgtcgca acggtctgct gagcgggtgt 1860
ggcaagttgc tgcgtccgaa tttgaaggat cgctacggtc agcgtttgga gcagatgtac 1920
gcgacatcg cggctacgca ggcaaccaa ttgcgtgagc tgcgtcgcg tgcggctact 1980
caaccggtga tcgacacgct gacgcaagct gcggcgacca tcctgggtac cggcagcgag 2040
gttgcaagcg acgcacactt tactgatttg ggcggtgatt ctctgagcg gctgacgttg 2100
agcaacttgc tgtctgactt ctttggcttt gaagtcccg ttggcacgat tgttaacca 2160
gcgactaatc tggcacagct ggcgcaacat atcagggcg agcgcacggc gggtgaccgc 2220
cgtccatcct ttacgacggt ccacggtgcg gatgctacgg aatccgtgc aagcgaactg 2280
actctggaca aattcatcga cgctgagact ctgcgcgag cacctggtt gccgaaggtt 2340
acgactgagc cgcgtacggt cctgttgagc ggtgccaatg gttggttggg ccgcttctg 2400
accctgcagt ggctggaacg tttggcaccg gttggcggta ccctgatcac cattgtgcg 2460
ggctgtgacg atgcagcggc ctgtgcagc ttgactcagg cttacgatac ggaccagag 2520
ctgtcccgcc gcttcgctga gttggcggat cgccacttgc gtgtggtggc aggtgatatc 2580
ggcgatcaga atctgggctt gaccccgag ctgtggcacc gtctggcagc agaggtcgat 2640
ctggctgctt atccagcggc cctggtcaac cacgtcctgc cgtaccgcca gctgtttggt 2700
ccgaatgttg ttggcaccgc cgaagttatc aagttggctc tgaccgagcg catcaagcct 2760
gttacctacc tgtccacggc gaaggtcgcg atgggtattc ctgattttga ggagacggt 2820
gacattcgta ccgtcagccc ggttcgtccg ctggatggtg gctatgcaa tggctatggc 2880
aacagcaagt gggctggcga ggtgctgctg cgcgaggcac atgacctgtg tggcctgccg 2940
gttgcgacgt ttcgtagcga catgattctg gccaccgc gctaccgtgg ccaagtgaat 3000
gtgccggaca tgttcaccg tctgtgctg teectgctga tcacgggtgt ggcaccgcgt 3060
tccttctaca ttggtgatgg cgagcgtccg cgtgcacact accgggctt gaccgtcgat 3120
tttgttgcg aagcggttac tacctgggt gctcagcaac gtgagggtta tgtctcgat 3180
gacgttatga atccgcacga tgacgttatt agcttgatg tctttgtgga ctggctgatt 3240
cgtgcgggcc acccaattga ccgtgttgac gactatgatg actgggtgcg tcgttttgaa 3300
accgcgttga ccgccttgc ggagaaacgt cgtgcgcaga ccgttctgcc gctgctgcat 3360
gccttctcg cgccacaggc gccgttgcgt ggcgccctg aaccgaccga agtgtttcat 3420
gcagcgggtc gtaccgctaa agtcggtccg ggtgatattc cgcacctgga tgaagccctg 3480

atcgacaagt acatccgtga cctgcgcgag ttcggtctga tttag 3525
 <210> 5
 <211> 1173
 <212> PRT
 <213> 人工序列
 <220>
 <221> 来源
 <223> /注释="人工序列的描述:合成CAR变异(carB2)多肽"
 <400> 5
 Met Thr Ser Asp Val His Asp Ala Thr Asp Gly Val Thr Glu Thr Ala
 1 5 10 15
 Leu Asp Asp Arg Gln Ser Thr Arg Arg Ile Ala Glu Leu Tyr Ala Thr
 20 25 30
 Asp Pro Glu Phe Ala Ala Ala Ala Pro Leu Pro Ala Val Val Asp Ala
 35 40 45
 Ala His Lys Pro Gly Leu Arg Leu Ala Glu Ile Leu Gln Thr Leu Phe
 50 55 60
 Thr Gly Tyr Gly Asp Arg Pro Ala Leu Gly Tyr Arg Ala Arg Glu Leu
 65 70 75 80
 Ala Thr Asp Glu Gly Gly Arg Thr Val Thr Arg Leu Leu Pro Arg Phe
 85 90 95
 Asp Thr Leu Thr Tyr Ala Gln Val Trp Ser Arg Val Gln Ala Val Ala
 100 105 110
 Ala Ala Leu Arg His Asn Phe Ala Gln Pro Ile Tyr Pro Gly Asp Ala
 115 120 125
 Val Ala Thr Ile Gly Phe Ala Ser Pro Asp Tyr Leu Thr Leu Asp Leu
 130 135 140
 Val Cys Ala Tyr Leu Gly Leu Val Ser Val Pro Leu Gln His Asn Ala
 145 150 155 160
 Pro Val Ser Arg Leu Ala Pro Ile Leu Ala Glu Val Glu Pro Arg Ile
 165 170 175
 Leu Thr Val Ser Ala Glu Tyr Leu Asp Leu Ala Val Glu Ser Val Arg
 180 185 190
 Asp Val Asn Ser Val Ser Gln Leu Val Val Phe Asp His His Pro Glu
 195 200 205
 Val Asp Asp His Arg Asp Ala Leu Ala Arg Ala Arg Glu Gln Leu Ala
 210 215 220
 Gly Lys Gly Ile Ala Val Thr Thr Leu Asp Ala Ile Ala Asp Glu Gly
 225 230 235 240

Ala Gly Leu Pro Ala Glu Pro Ile Tyr Thr Ala Asp His Asp Gln Arg
 245 250 255
 Leu Ala Met Ile Leu Tyr Thr Ser Gly Ser Thr Gly Ala Pro Lys Gly
 260 265 270
 Ala Met Tyr Thr Glu Ala Met Val Ala Arg Leu Trp Thr Met Ser Gly
 275 280 285
 Ile Thr Gly Asp Pro Thr Pro Val Ile Asn Val Asn Phe Met Pro Leu
 290 295 300
 Asn His Leu Gly Gly Arg Ile Pro Ile Ser Thr Ala Val Gln Asn Gly
 305 310 315 320
 Gly Thr Ser Tyr Phe Val Pro Glu Ser Asp Met Ser Thr Leu Phe Glu
 325 330 335
 Asp Leu Ala Leu Val Arg Pro Thr Glu Leu Gly Leu Val Pro Arg Val
 340 345 350
 Ala Asp Met Leu Tyr Gln His His Leu Ala Thr Val Asp Arg Leu Val
 355 360 365
 Thr Gln Gly Ala Asp Glu Leu Thr Ala Glu Lys Gln Ala Gly Ala Glu
 370 375 380
 Leu Arg Glu Gln Val Leu Gly Gly Arg Val Ile Thr Gly Phe Val Ser
 385 390 395 400
 Thr Ala Pro Leu Ala Ala Glu Met Arg Ala Phe Leu Asp Ile Thr Leu
 405 410 415
 Gly Ala His Ile Val Asp Gly Tyr Gly Leu Thr Glu Thr Gly Ala Val
 420 425 430
 Thr Arg Asp Gly Val Ile Val Arg Pro Pro Val Ile Asp Tyr Lys Leu
 435 440 445
 Ile Asp Val Pro Glu Leu Gly Tyr Phe Ser Thr Asp Lys Pro Tyr Pro
 450 455 460
 Arg Gly Glu Leu Leu Val Arg Ser Ile Thr Leu Thr Pro Gly Tyr Tyr
 465 470 475 480
 Lys Arg Pro Glu Val Thr Ala Ser Val Phe Asp Arg Asp Gly Tyr Tyr
 485 490 495
 His Thr Gly Asp Val Met Ala Glu Thr Ala Pro Asp His Leu Val Tyr
 500 505 510
 Val Asp Arg Arg Asn Asn Val Leu Lys Leu Ala Gln Gly Glu Phe Val
 515 520 525
 Ala Val Ala Asn Leu Glu Ser Val Phe Ser Gly Ala Ala Leu Val Arg
 530 535 540
 Gln Ile Phe Val Tyr Gly Asn Ser Glu Arg Ser Phe Leu Leu Ala Val

545	550	555	560
Val Val Pro Thr Pro Glu Ala Leu Glu Gln Tyr Asp Pro Ala Ala Leu			
	565	570	575
Lys Ala Ala Leu Ala Asp Ser Leu Gln Arg Thr Ala Arg Asp Ala Glu			
	580	585	590
Leu Gln Ser Tyr Glu Val Pro Ala Asp Phe Ile Val Glu Thr Glu Pro			
	595	600	605
Phe Ser Ala Ala Asn Gly Leu Leu Ser Gly Val Gly Lys Leu Leu Arg			
	610	615	620
Pro Asn Leu Lys Asp Arg Tyr Gly Gln Arg Leu Glu Gln Met Tyr Ala			
625	630	635	640
Asp Ile Ala Ala Thr Gln Ala Asn Gln Leu Arg Glu Leu Arg Arg Ala			
	645	650	655
Ala Ala Thr Gln Pro Val Ile Asp Thr Leu Thr Gln Ala Ala Ala Thr			
	660	665	670
Ile Leu Gly Thr Gly Ser Glu Val Ala Ser Asp Ala His Phe Thr Asp			
	675	680	685
Leu Gly Gly Asp Ser Leu Ser Ala Leu Thr Leu Ser Asn Leu Leu Ser			
	690	695	700
Asp Phe Phe Gly Phe Glu Val Pro Val Gly Thr Ile Val Asn Pro Ala			
705	710	715	720
Thr Asn Leu Ala Gln Leu Ala Gln His Ile Glu Ala Gln Arg Thr Ala			
	725	730	735
Gly Asp Arg Arg Pro Ser Phe Thr Thr Val His Gly Ala Asp Ala Thr			
	740	745	750
Glu Ile Arg Ala Ser Glu Leu Thr Leu Asp Lys Phe Ile Asp Ala Glu			
	755	760	765
Thr Leu Arg Ala Ala Pro Gly Leu Pro Lys Val Thr Thr Glu Pro Arg			
	770	775	780
Thr Val Leu Leu Ser Gly Ala Asn Gly Trp Leu Gly Arg Phe Leu Thr			
785	790	795	800
Leu Gln Trp Leu Glu Arg Leu Ala Pro Val Gly Gly Thr Leu Ile Thr			
	805	810	815
Ile Val Arg Gly Arg Asp Asp Ala Ala Ala Arg Ala Arg Leu Thr Gln			
	820	825	830
Ala Tyr Asp Thr Asp Pro Glu Leu Ser Arg Arg Phe Ala Glu Leu Ala			
	835	840	845
Asp Arg His Leu Arg Val Val Ala Gly Asp Ile Gly Asp Pro Asn Leu			
	850	855	860

gcggacttca ttgttgagac tgagcctttt agcgcctgca acggtctgct gagcgggtgtt 1860
 ggcaagttgc tgcgtccgaa tttgaaggat cgctacggtc agcgtttgga gcagatgtac 1920
 gcggacatcg cggctacgca ggcgaaccaa ttgcgtgagc tgcgtcgcgc tgcggctact 1980
 caaccgggtga tcgacacgct gacgcaagct gcggcgacca tcctgggtac cggcagcgag 2040
 gttgcaagcg acgcacactt tactgatttg ggcggtgatt ctctgagcgc gctgacgttg 2100
 agcaacttgc tgtctgactt ctttggcttt gaagtcccgg ttggcacgat tgttaacca 2160
 gcgactaatc tggcacagct ggcgcaacat atcgaggcgc agcgcacggc gggtgaccgc 2220
 cgtccatcct ttacgacggt ccacgggtgc gatgctacgg aatccgtgc aagcgaactg 2280
 actctggaca aattcatcga cgctgagact ctgcgcgcag cacctggttt gccgaaggtt 2340
 acgactgagc cgcgtacggt cctgttgagc ggtgccaatg gttggttggg ccgcttcctg 2400
 accctgcagt ggctggaacg tttggcaccg gttggcggtg ccctgatcac cattgtgcgc 2460
 ggtcgtgacg atgcagcggc ccgcgcacgc ttgactcagg cttacgatac ggaccagag 2520
 ctgtcccgcc gcttcgctga gttggcggtg cgccacttgc gtgtggtggc aggtgatatac 2580
 ggcgatccga atctgggcct gaccccggag atttggcacc gtctggcagc agaggtcgat 2640
 ctggctgctt atccagcggc cctggtcaac cacgtctgc cgtaccgcca gctgtttggt 2700
 ccgaatgttg ttggcaccgc cgaagttatc aagttggctc tgaccgagcg catcaagcct 2760
 gttacctacc tgtccacggt tagcgtcgcg atgggtattc ctgattttga ggaggacggt 2820
 gacattcgta ccgtcagccc ggttcgtccg ctggatggtg gctatgcaa tggctatggc 2880
 aacagcaagt gggctggcga ggtgctgctg cgcgaggcac atgacctgtg tggcctgccg 2940
 gttgcgacgt ttcgtagcga catgattctg gccaccgcc gctaccgtgg ccaagtgaat 3000
 gtgccggaca tgttcaccg tctgctgctg tcctgctga tcacgggtgt ggcaccgcgt 3060
 tccttctaca ttggtgatgg cgagcgtccg cgtgcacact acccgggcct gaccgtcgat 3120
 tttgttgccg aagcggttac taccctgggt gctcagcaac gtgagggtta tgtctcgtat 3180
 gacgttatga atccgcacga tgacggtatt agcttgatg tctttgtgga ctggctgatt 3240
 cgtgcccggc acccaattga ccgtgttgac gactatgatg actgggtgcg tcgttttgaa 3300
 accgcgttga ccgccttgcc ggagaaacgt cgtgcgcaga ccgttctgcc gctgctgcat 3360
 gcctttcgcg cgccacaggc gccgttgcgt ggcgccctg aaccgaccga agtgtttcat 3420
 gcagcgggtc gtaccgctaa agtcggtccg ggtgatattc cgcacctgga tgaagccctg 3480
 atcgacaagt acatccgtga cctgcgcgag ttcggtctga tttag 3525

<210> 7

<211> 1174

<212> PRT

<213> 人工序列

<220>

<221> 来源

<223> /注释=“人工序列的描述:合成CAR变异(carB8)多肽”

<400> 7

Met Gly Thr Ser Asp Val His Asp Ala Thr Asp Gly Val Thr Glu Thr

1

5

10

15

Ala Leu Asp Asp Arg Gln Arg Thr Arg Arg Ile Ala Glu Leu Tyr Ala
 20 25 30
 Thr Asp Pro Glu Phe Ala Ala Ala Ala Pro Leu Pro Ala Val Val Asp
 35 40 45
 Ala Ala His Lys Pro Gly Leu Arg Leu Ala Glu Ile Leu Gln Thr Leu
 50 55 60
 Phe Thr Gly Tyr Gly Asp Arg Pro Ala Leu Gly Tyr Arg Ala Arg Glu
 65 70 75 80
 Leu Ala Thr Asp Glu Gly Gly Arg Thr Val Thr Arg Leu Leu Pro Arg
 85 90 95
 Phe Asp Thr Leu Thr Tyr Ala Gln Val Trp Ser Arg Val Gln Ala Val
 100 105 110
 Ala Ala Ala Leu Arg His Asn Phe Ala Gln Pro Ile Tyr Pro Gly Asp
 115 120 125
 Ala Val Ala Thr Ile Gly Phe Ala Ser Pro Asp Tyr Leu Thr Leu Asp
 130 135 140
 Leu Val Cys Ala Tyr Leu Gly Leu Val Ser Val Pro Leu Gln His Asn
 145 150 155 160
 Ala Pro Val Ser Arg Leu Ala Pro Ile Leu Ala Glu Val Glu Pro Arg
 165 170 175
 Ile Leu Thr Val Ser Ala Glu Tyr Leu Asp Leu Ala Val Glu Ser Val
 180 185 190
 Arg Asp Val Asn Ser Val Ser Gln Leu Val Val Phe Asp His His Pro
 195 200 205
 Glu Val Asp Asp His Arg Asp Ala Leu Ala Arg Ala Arg Glu Gln Leu
 210 215 220
 Ala Gly Lys Gly Ile Ala Val Thr Thr Leu Asp Ala Ile Ala Asp Glu
 225 230 235 240
 Gly Ala Gly Leu Pro Ala Glu Pro Ile Tyr Thr Ala Asp His Asp Gln
 245 250 255
 Arg Leu Ala Met Ile Leu Tyr Thr Ser Gly Ser Thr Gly Ala Pro Lys
 260 265 270
 Gly Ala Met Tyr Thr Glu Ala Met Val Ala Arg Leu Trp Thr Met Ser
 275 280 285
 Gly Ile Thr Gly Asp Pro Thr Pro Val Ile Asn Val Asn Phe Met Pro
 290 295 300
 Leu Asn His Leu Gly Gly Arg Ile Pro Ile Ser Thr Ala Val Gln Asn
 305 310 315 320
 Gly Gly Thr Ser Tyr Phe Val Pro Glu Ser Asp Met Ser Thr Leu Phe

	325	330	335
Glu Asp Leu Ala Leu Val Arg Pro Thr	Glu Leu Gly Leu Val Pro Arg		
	340	345	350
Val Ala Asp Met Leu Tyr Gln His His	Leu Ala Thr Val Asp Arg Leu		
	355	360	365
Val Thr Gln Gly Ala Asp Glu Leu Thr	Ala Glu Lys Gln Ala Gly Ala		
	370	375	380
Glu Leu Arg Glu Gln Val Leu Gly Gly	Arg Val Ile Thr Gly Phe Val		
385	390	395	400
Ser Thr Ala Pro Leu Ala Ala Glu Met	Arg Ala Phe Leu Asp Ile Thr		
	405	410	415
Leu Gly Ala His Ile Val Asp Gly Tyr	Gly Leu Thr Glu Thr Gly Ala		
	420	425	430
Val Thr Arg Asp Gly Val Ile Val Arg	Pro Pro Val Ile Asp Tyr Lys		
	435	440	445
Leu Ile Asp Val Pro Glu Leu Gly Tyr	Phe Ser Thr Asp Lys Pro Tyr		
	450	455	460
Pro Arg Gly Glu Leu Leu Val Arg Ser	His Thr Leu Thr Pro Gly Tyr		
465	470	475	480
Tyr Lys Arg Pro Glu Val Thr Ala Ser	Val Phe Asp Arg Asp Gly Tyr		
	485	490	495
Tyr His Thr Gly Asp Val Met Ala Glu	Thr Ala Pro Asp His Leu Val		
	500	505	510
Tyr Val Asp Arg Arg Asn Asn Val Leu	Lys Leu Ala Gln Gly Glu Phe		
	515	520	525
Val Ala Val Ala Asn Leu Glu Ser Val	Phe Ser Gly Ala Ala Leu Val		
	530	535	540
Arg Gln Ile Phe Val Tyr Gly Asn Ser	Glu Arg Ser Phe Leu Leu Ala		
545	550	555	560
Val Val Val Pro Thr Pro Glu Ala Leu	Glu Gln Tyr Asp Pro Ala Ala		
	565	570	575
Leu Lys Ala Ala Leu Ala Asp Ser Leu	Gln Arg Thr Ala Arg Asp Ala		
	580	585	590
Glu Leu Gln Ser Tyr Glu Val Pro Ala	Asp Phe Ile Val Glu Thr Glu		
	595	600	605
Pro Phe Ser Ala Ala Asn Gly Leu Leu	Ser Gly Val Gly Lys Leu Leu		
	610	615	620
Arg Pro Asn Leu Lys Asp Arg Tyr Gly	Gln Arg Leu Glu Gln Met Tyr		
625	630	635	640

Ala Asp Ile Ala Ala Thr Gln Ala Asn Gln Leu Arg Glu Leu Arg Arg
645 650 655

Ala Ala Ala Thr Gln Pro Val Ile Asp Thr Leu Thr Gln Ala Ala Ala
660 665 670

Thr Ile Leu Gly Thr Gly Ser Glu Val Ala Ser Asp Ala His Phe Thr
675 680 685

Asp Leu Gly Gly Asp Ser Leu Ser Ala Leu Thr Leu Ser Asn Leu Leu
690 695 700

Ser Asp Phe Phe Gly Phe Glu Val Pro Val Gly Thr Ile Val Asn Pro
705 710 715 720

Ala Thr Asn Leu Ala Gln Leu Ala Gln His Ile Glu Ala Gln Arg Thr
725 730 735

Ala Gly Asp Arg Arg Pro Ser Phe Thr Thr Val His Gly Ala Asp Ala
740 745 750

Thr Glu Ile Arg Ala Ser Glu Leu Thr Leu Asp Lys Phe Ile Asp Ala
755 760 765

Glu Thr Leu Arg Ala Ala Pro Gly Leu Pro Lys Val Thr Thr Glu Pro
770 775 780

Arg Thr Val Leu Leu Ser Gly Ala Asn Gly Trp Leu Gly Arg Phe Leu
785 790 795 800

Thr Leu Gln Trp Leu Glu Arg Leu Ala Pro Val Gly Gly Thr Leu Ile
805 810 815

Thr Ile Val Arg Gly Arg Asp Asp Ala Ala Ala Arg Ala Arg Leu Thr
820 825 830

Gln Ala Tyr Asp Thr Asp Pro Glu Leu Ser Arg Arg Phe Ala Glu Leu
835 840 845

Ala Asp Arg His Leu Arg Val Val Ala Gly Asp Ile Gly Asp Pro Asn
850 855 860

Leu Gly Leu Thr Pro Glu Ile Trp His Ser Leu Ala Ala Glu Val Asp
865 870 875 880

Leu Val Val His Pro Ala Ala Leu Val Asn His Val Leu Pro Tyr Arg
885 890 895

Gln Leu Phe Gly Pro Asn Val Val Gly Thr Ala Glu Val Ile Lys Leu
900 905 910

Ala Leu Thr Glu Arg Ile Lys Pro Val Thr Tyr Leu Ser Thr Val Gly
915 920 925

Val Ala Arg Gly Ile Pro Asp Phe Glu Glu Asp Gly Asp Ile Arg Thr
930 935 940

Val Ser Pro Val Arg Pro Leu Asp Gly Gly Tyr Ala Asn Gly Tyr Gly

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

Ile

<210> 8

<211> 3525

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释="人工序列的描述:合成CAR变异(carB8)多核苷酸"

<400> 8

atgggcacga gcgatgttca cgacgcgacc gacggcgтта ccgagactgc actggatgat 60

cgccagagga ctcgtcgtat tgcagaactg tacgcaacgg acccagagtt cgcagcagca 120
gctcctctgc cggccgttgt cgatgcggcg cacaaaccgg gcctgcgtct ggcggaaatc 180
ctgcagaccc tgttcaccgg ctacggcgat cgtccggcgc tgggctatcg tgcacgtgag 240
ctggcgacgg acgaaggcgg tcgtacggtc acgcgtctgc tgccgcgctt cgataccctg 300
acctatgcac aggtgtggag ccgtgttcaa gcagtggctg cagcgttgcg tcacaatttc 360
gcacaaccga tttaccggg cgacgcggtc gcgactatcg gctttgcgag cccggactat 420
ttgacgctgg atctgggtg cgcgtatctg ggcttggtca gcgttccttt gcagcataac 480
gctccggtgt ctcgcctggc cccgattctg gccgaggtgg aaccgcgtat tctgacgggtg 540
agcgcagaat acctggacct ggcggtttaa tccgtccgtg atgtgaactc cgtcagccag 600
ctggttgttt tcgaccatca tccggaagtg gacgateacc gtgacgact ggctcgcgca 660
cgcgagcagc tggccggcaa aggtatcgca gttacgacc tggatgcgat cgcagacgaa 720
ggcgcaggtt tgccggetga gccgatttac acggcggatc acgatcagcg tctggccatg 780
attctgtata ccagcggtc tacgggtgct ccgaaaggcg cgatgtacac cgaagcgatg 840
gtggctcgc tgtggactat gagcgggatc acgggcgacc cgaccccggt tatcaacgtg 900
aacttcatgc cgctgaacca tctgggcggt cgtatccga ttagcaccgc cgtgcagaat 960
ggcggtagca gctacttctg tccgaaagc gacatgagca cgctgttga ggatctggcc 1020
ctgggtccgc ctaccgaact gggctctggtg ccgcgtgttg cggacatgct gtaccagcat 1080
catctggcga ccgtggatcg cctggtgacc cagggcgcgg acgaactgac tgcggaaaag 1140
caggccggtg cggaactgcg tgaacaggtc ttgggcggtc gtgttatcac cggttttgtt 1200
tccaccgcgc cgttggcggc agagatgcgt gcttttctgg atatcacctt ggggtgcacac 1260
atcgttgacg gttacggtct gaccgaaacc ggtgcggtca cccgtgatgg tgtgattgtt 1320
cgtcctccgg tcattgatta caagctgatc gatgtgccgg agctgggtta cttctccacc 1380
gacaaaccgt acccgcgtgg cgagctgctg gttcgtagcc acacgttgac tccgggttac 1440
tacaagcgc cagaagtcac cgcgtccgtt ttcgatcgcg acggctatta ccacaccggc 1500
gacgtgatgg cagaaaccgc gccagaccac ctggtgtatg tggaccgcc caacaatgtt 1560
ctgaagctgg cgcaaggatg atttgtcgc gtggctaacc tggagtccgt tttcagcggc 1620
gctgctctgg tccgccagat tttcgtgtat ggtaacagcg agcgcagctt tctgttggct 1680
gttgttgtcc ctaccccgga ggcgctggag caatacagcc ctgccgatt gaaagcagcc 1740
ctggcggatt cgctgcagcg tacggcgcgt gatgccgagc tgcagagcta tgaagtgccg 1800
gcggacttca ttgttgagac tgagcctttt agcgtcgcga acggtctgct gagcgggtgtt 1860
ggcaagttgc tgcgtccgaa tttgaaggat cgctacggtc agcgtttgga gcagatgtac 1920
gcggacatcg cggtacgca ggcaaccaa ttgcgtgagc tgcgtcgcgc tgcggctact 1980
caaccggtga tcgacacgct gacgcaagct gcggcgacca tctgggttac cggcagcgag 2040
gttgaagcg acgcacactt tactgatttg ggcggtgatt ctctgagcgc gctgacgttg 2100
agcaacttgc tgtctgactt ctttggcttt gaagtcccgg ttggcacgat tgttaacca 2160
gcgactaatc tggcacagct ggcaacat atcgaggcgc agcgcacggc gggtagccgc 2220
cgtccatcct ttacgacggt ccacggtgcg gatgctacgg aatccgtgc aagcgaactg 2280
actctggaca aattcatcga cgctgagact ctgcgcgag cacctggttt gccgaaggtt 2340
acgactgagc cgcgtacggt cctgttgagc ggtgccaatg gttggttggg ccgcttctg 2400

accctgcagt ggctggaacg tttggcaccg gttggcggta ccctgatcac cattgtgcgc 2460
ggctcgtgacg atgcagcggc ccgcgcacgc ttgactcagg cttacgatac ggaccagag 2520
ctgtcccgcc gcttcgctga gttggcggat cgccacttgc gtgtggtggc aggtgatatc 2580
ggcgatccga atctgggcct gaccccggag atttggcaca gtctggcagc agaggtcgat 2640
ctggctgttc atccagcggc cctgggtcaac cacgtcctgc cgtaccgcca gctgtttggt 2700
ccgaatgttg ttggcaccgc cgaagttatc aagttggctc tgaccgagcg catcaagcct 2760
gttacctacc tgtccacggt tggggtcgcg aggggtattc ctgattttga ggaggacggt 2820
gacattcgta ccgtcagccc ggttcgtccg ctggatggtg gctatgcaaa tggctatggc 2880
aacagcaagt gggctggcga ggtgctgctg cgcgaggcac atgacctgtg tggcctgccg 2940
gttgcgacgt ttcgtagcga catgattctg gcccaccgc gctaccgtgg ccaagtgaat 3000
gtgccggaca tgttacccg tctgctgctg tcctgctga tcacgggtgt ggcaccgcgt 3060
tccttctaca ttggtgatgg cgagcgtccg cgtgcacact acccggcct gaccgtcgat 3120
tttgttgccg aagcggttac taccctgggt gctcagcaac gtgagggtta tgtctcgat 3180
gacgttatga atccgcacga tgacgttatt agcttgatg tctttgtgga ctggctgatt 3240
cgtgcgggcc acccaattga ccgtgttgac gactatgatg actgggtgcg tcgttttgaa 3300
accgcgttga ccgccttgc ggagaaacgt cgtgcgcaga ccgttctgcc gctgctgcat 3360
gcctttcgcg cgccacaggc gccgtggcgt ggcgccctg aaccgaccga agtgtttcat 3420
gcagcgggtc gtaccgctaa agtcggtccg ggtgatattc cgcacctgga tgaagccctg 3480
atcgacaagt acatccgtga cctgcgcgag ttcggtctga tttag 3525

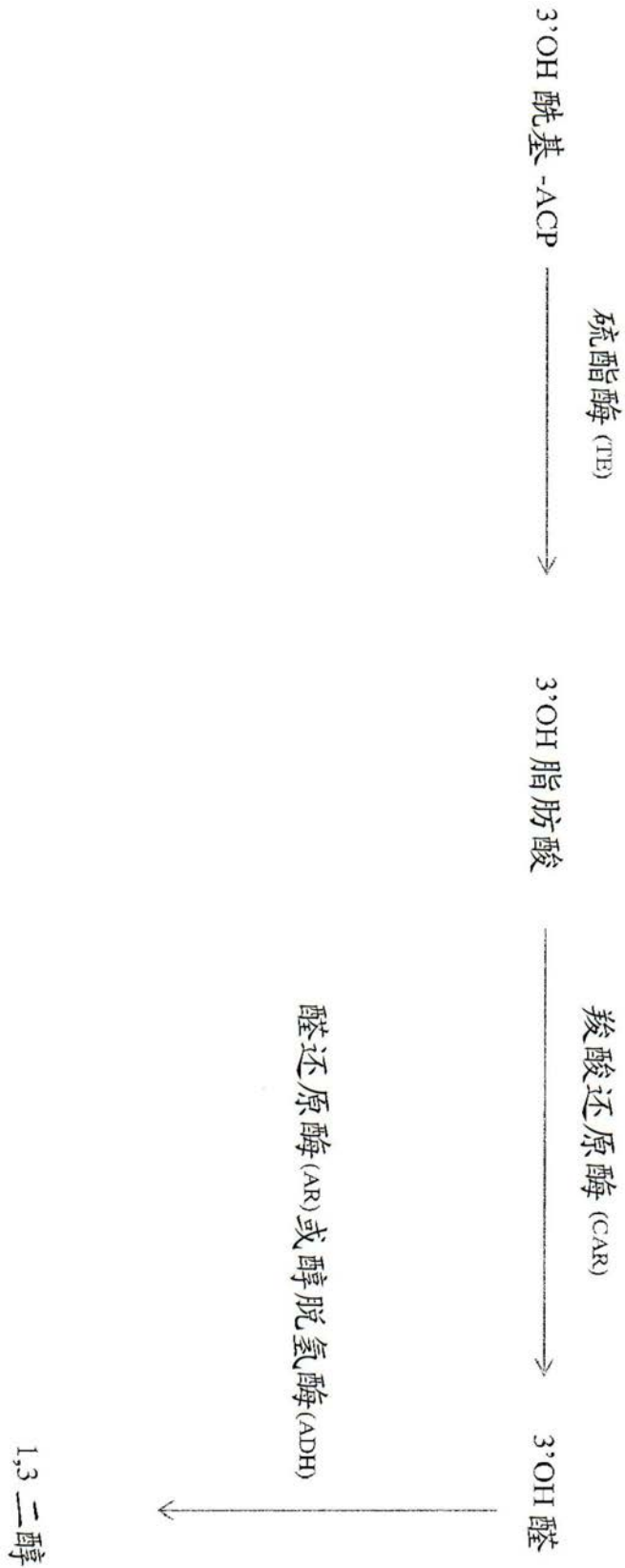


图1

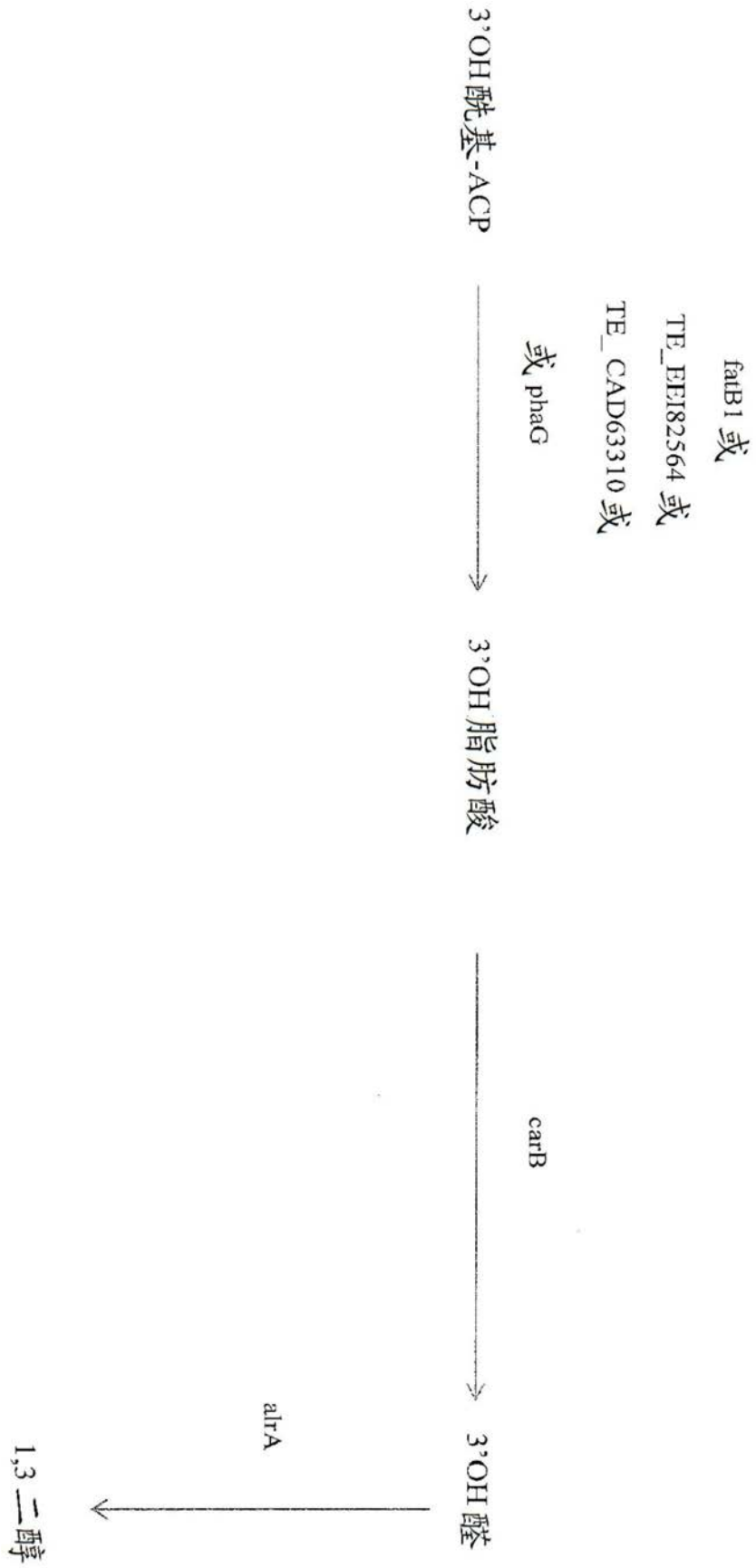


图2

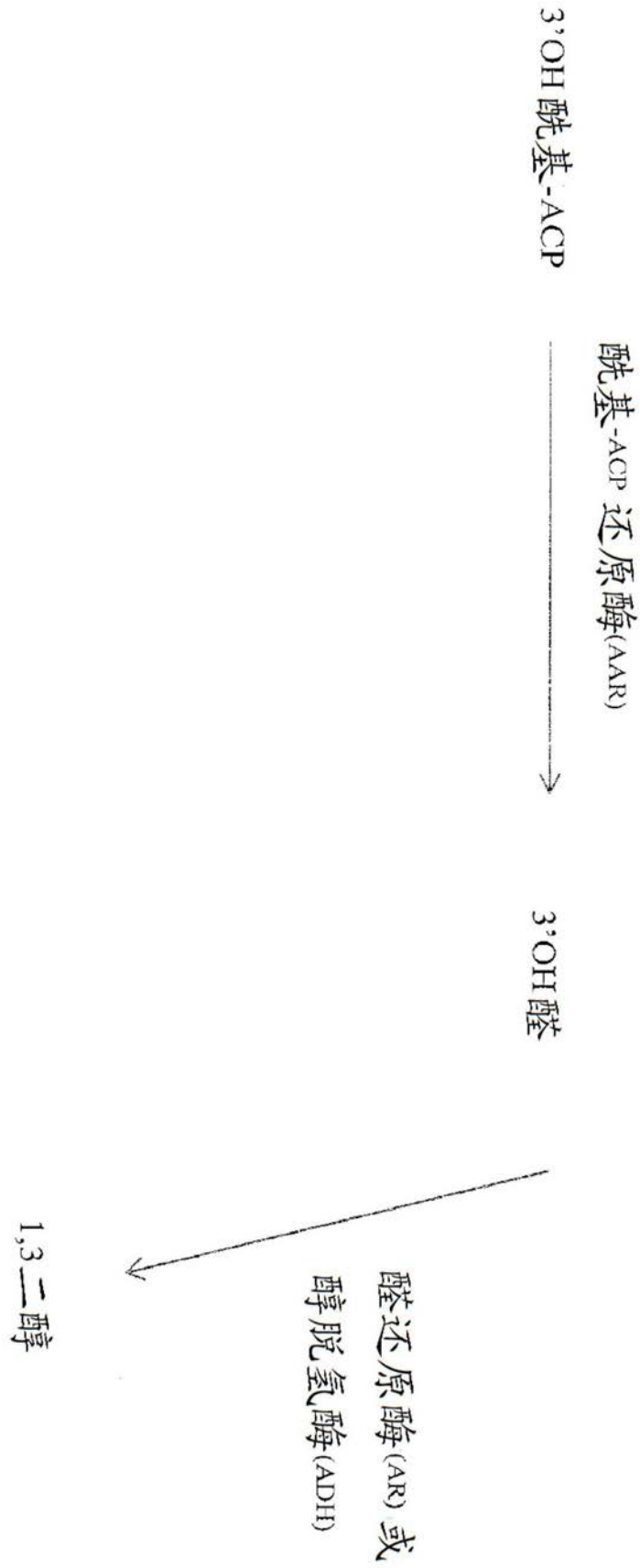


图3

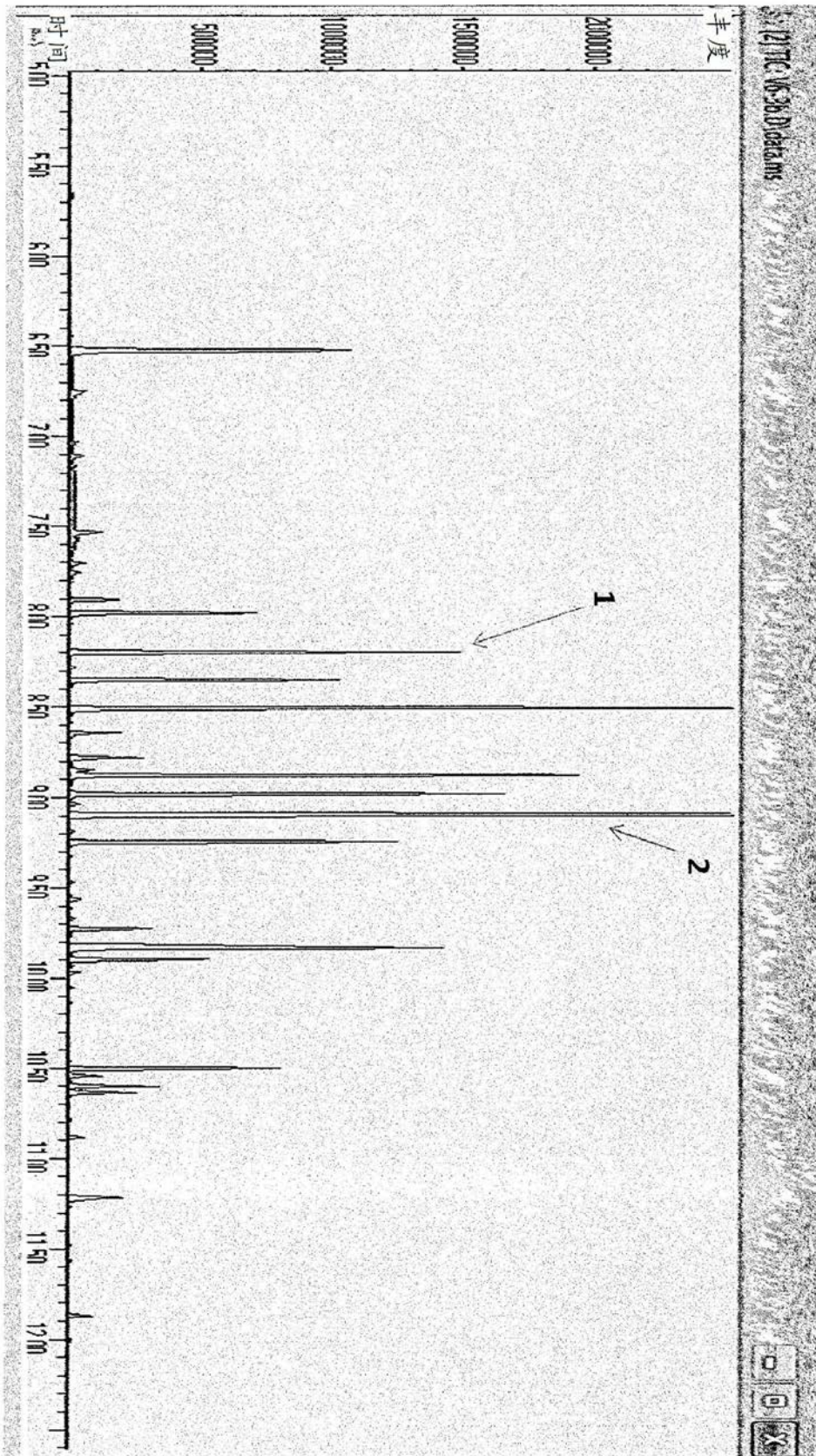


图4

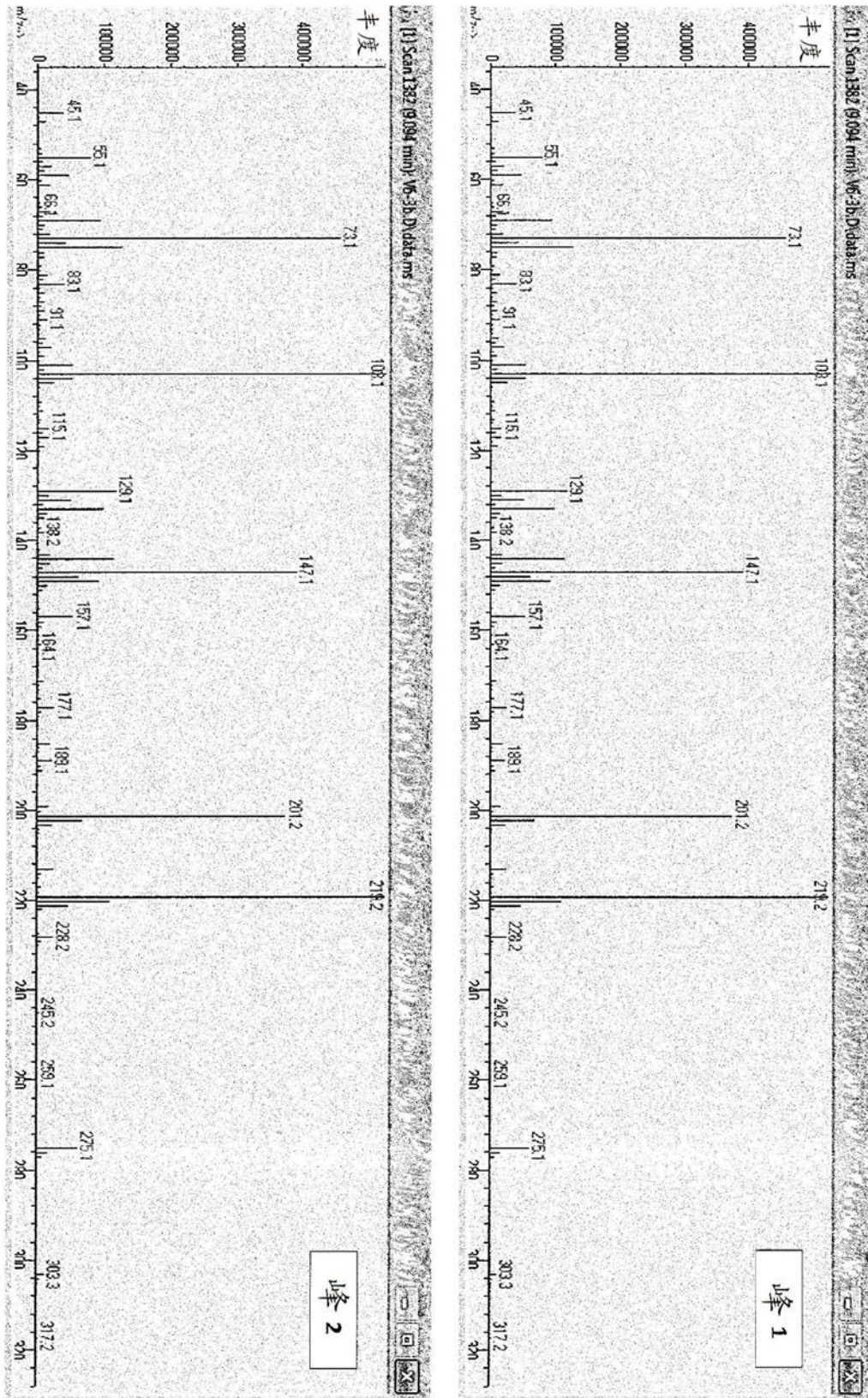


图5

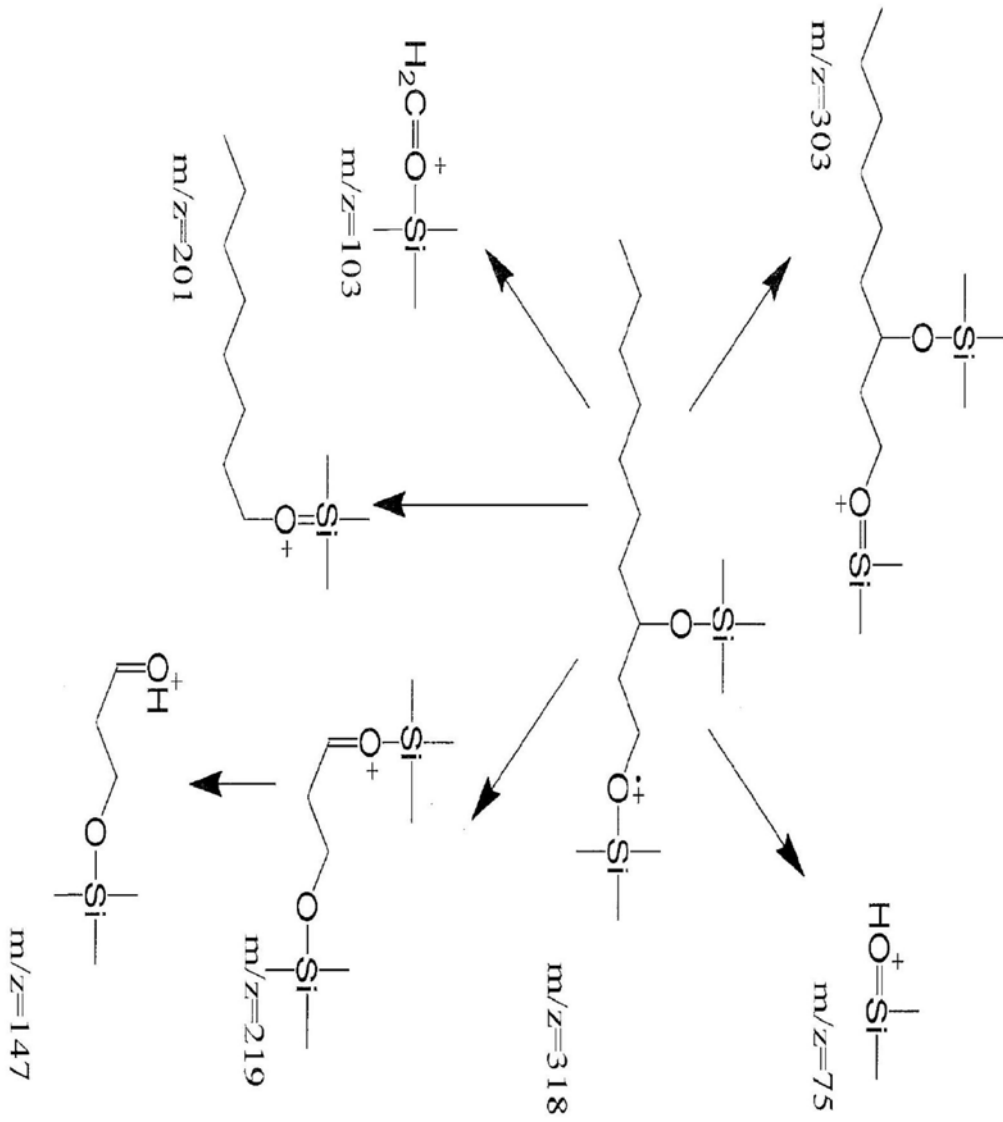


图6

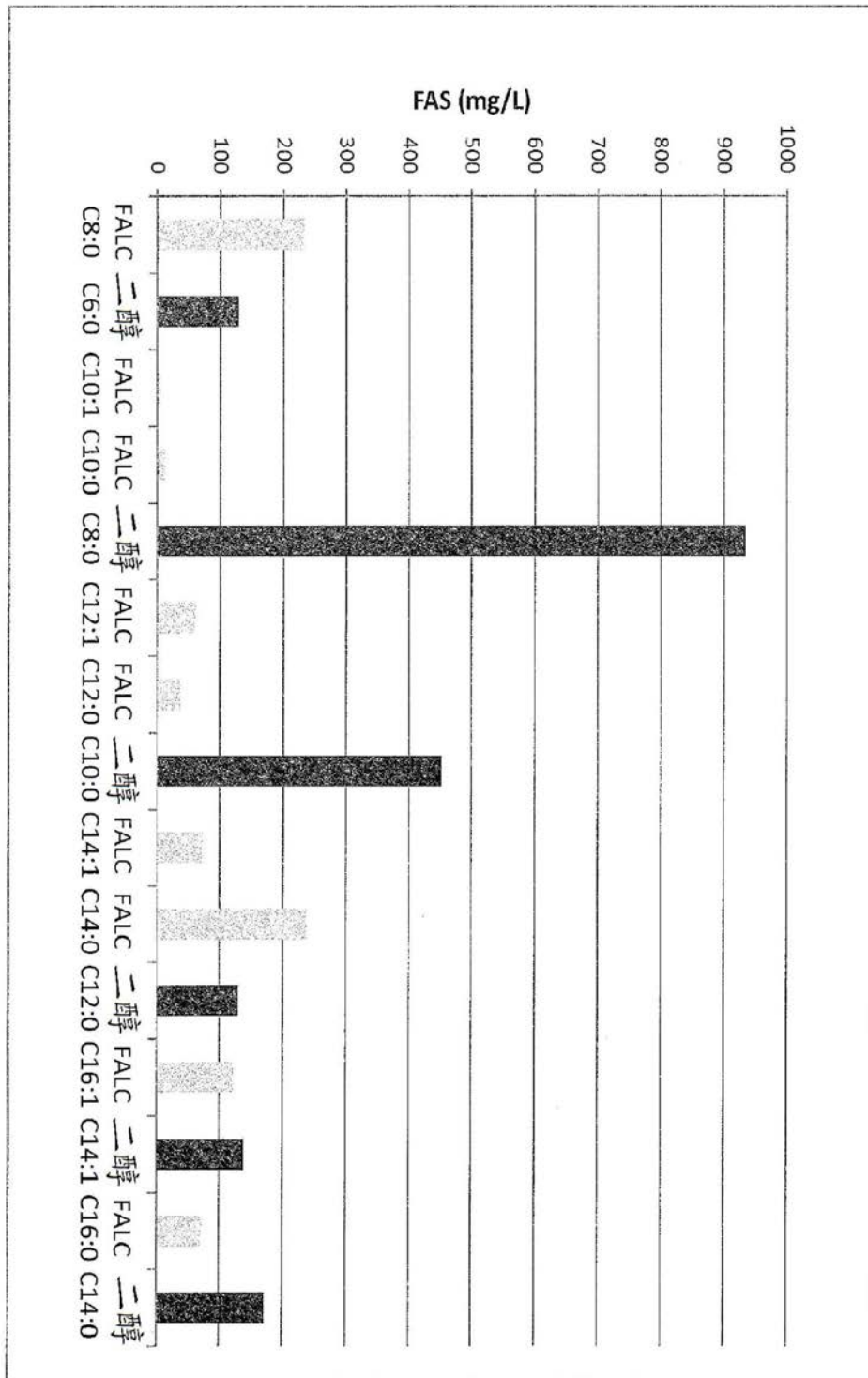


图7

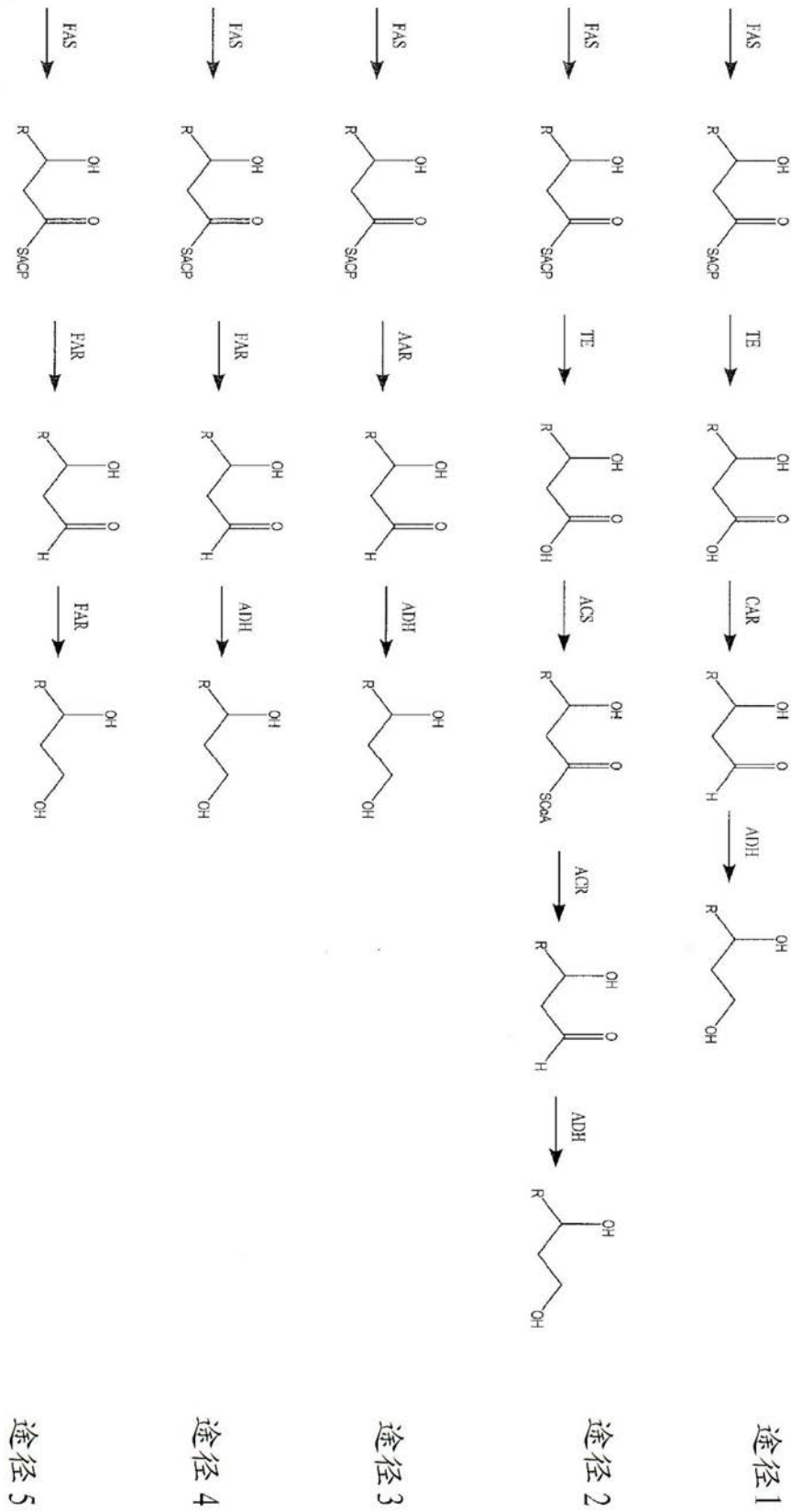


图8

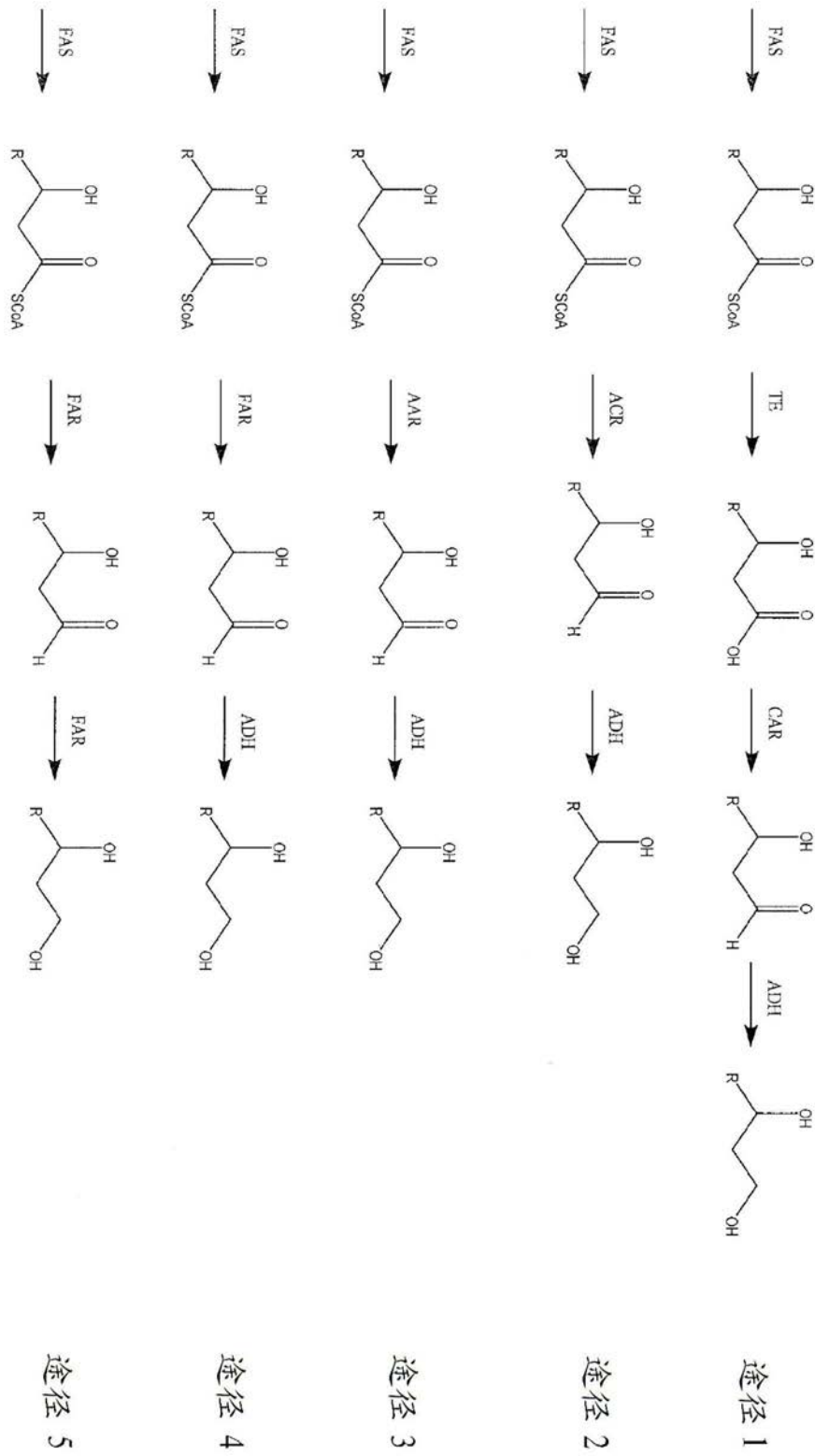


图9

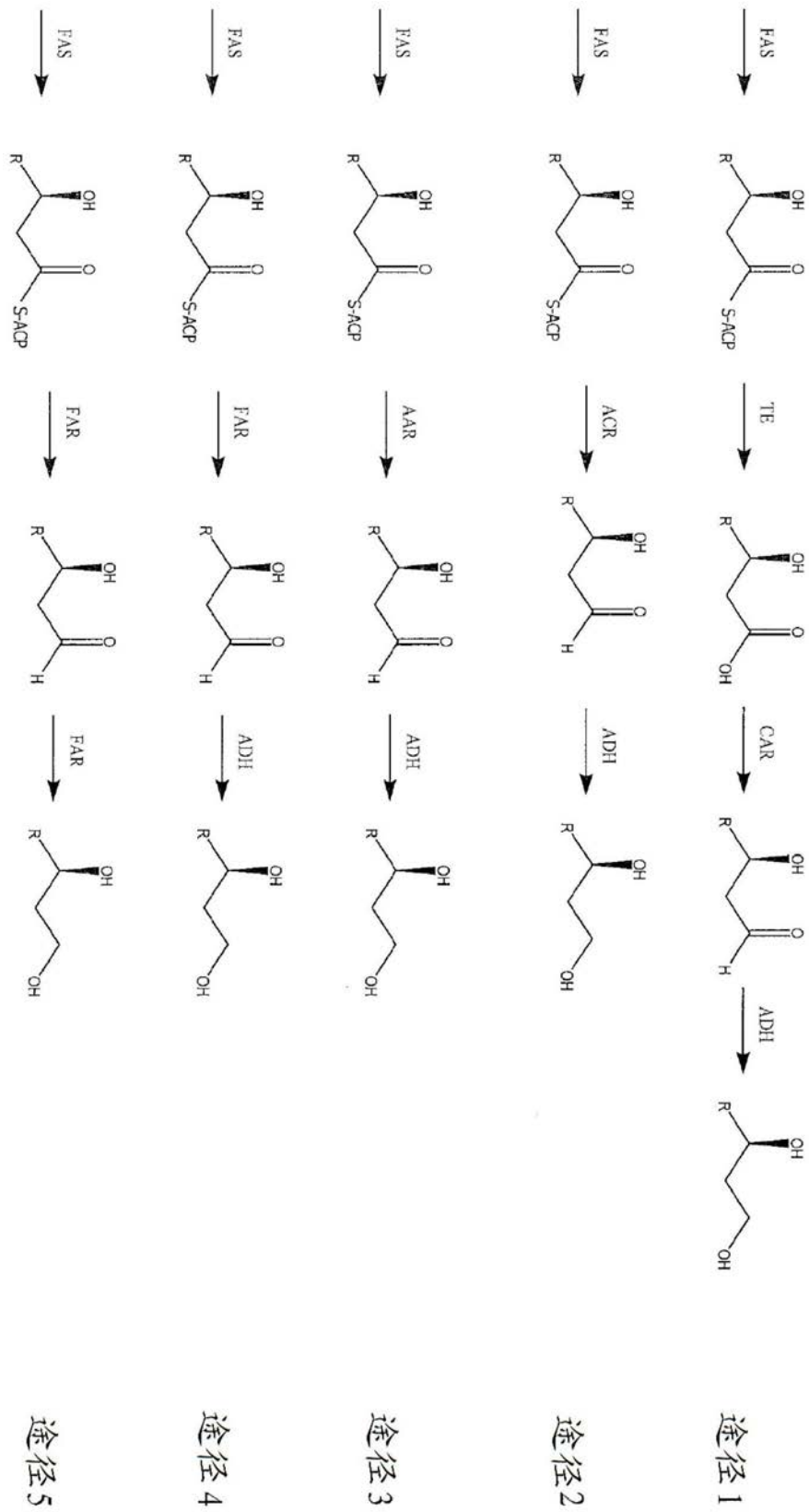


图10

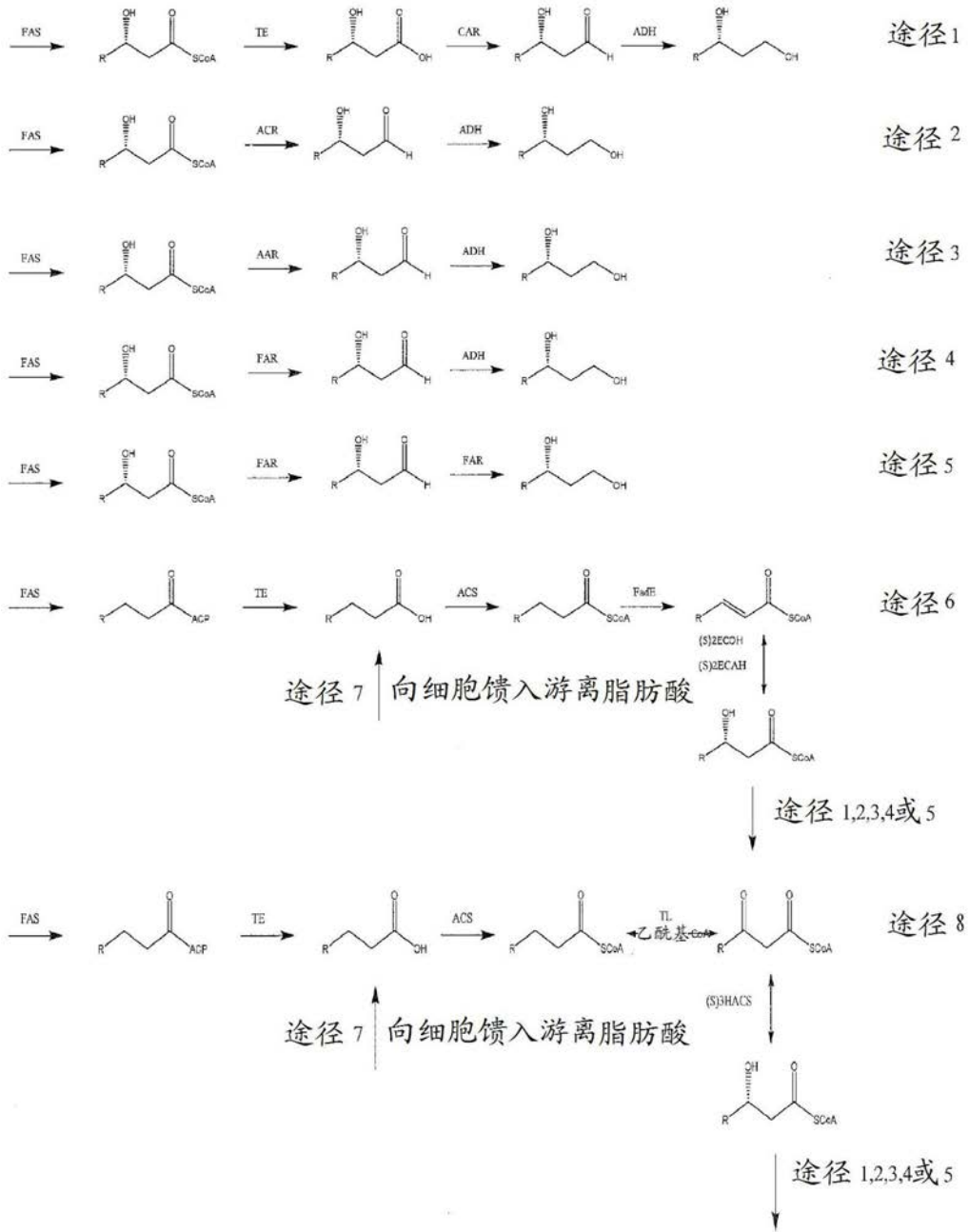


图11