



(12)发明专利

(10)授权公告号 CN 104053771 B

(45)授权公告日 2016.11.09

(21)申请号 201280067329.9

(22)申请日 2012.11.14

(65)同一申请的已公布的文献号
申请公布号 CN 104053771 A

(43)申请公布日 2014.09.17

(30)优先权数据
61/561,665 2011.11.18 US

(85)PCT国际申请进入国家阶段日
2014.07.17

(86)PCT国际申请的申请数据
PCT/US2012/065046 2012.11.14

(87)PCT国际申请的公布数据
W02013/074650 EN 2013.05.23

(73)专利权人 科德克希思公司
地址 美国加利福尼亚州

(72)发明人 法比安·卡比罗尔 陈海滨
阿努潘·苟呵尔
史蒂文·J·科利尔

德里克·史密斯

波吉特·科斯耶克 雅各布·杰恩

(74)专利代理机构 北京安信方达知识产权代理有限公司 11262

代理人 王思琪 郑霞

(51)Int.Cl.

C12N 9/04(2006.01)

C12P 13/00(2006.01)

C12P 7/00(2006.01)

C12P 17/10(2006.01)

C12N 9/00(2006.01)

(56)对比文件

W0 2010025238 A2,2010.03.04,

W0 2010025085 A2,2010.03.04,

CN 101627116 A,2010.01.13,

CN 101784669 A,2010.07.21,

W0 2009046153 A1,2009.04.09,

审查员 毛舒燕

权利要求书8页 说明书67页
序列表132页

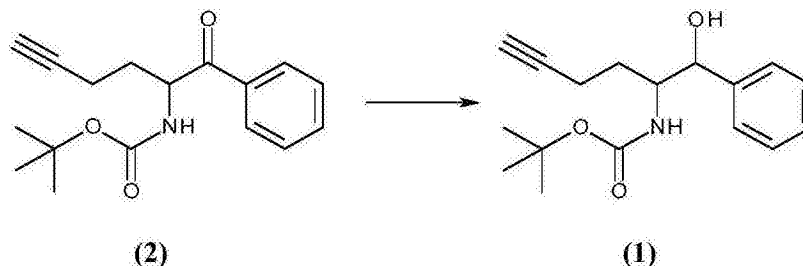
(54)发明名称

用于制备羟基取代的氨基甲酸酯的生物催
化剂

(57)摘要

本公开内容涉及用于制备羟基取代的氨基
甲酸酯化合物的工程化酮还原酶多肽、和多核苷
酸、载体、宿主细胞,以及制备和使用该酮还原酶
多肽的方法。

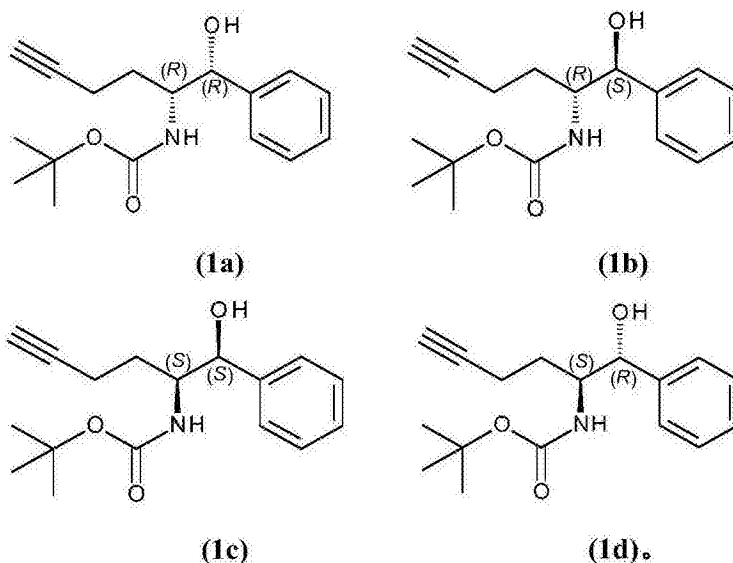
1. 一种工程化酮还原酶多肽, 所述工程化酮还原酶多肽能够将底物化合物(2), (1-氧代-1-苯基己-5-炔-2-基)氨基甲酸叔丁酯, 转化为产物化合物(1), (1-羟基-1-苯基己-5-炔-2-基)氨基甲酸叔丁酯,



所述酮还原酶多肽的氨基酸序列基于SEQ ID NO:4并且与SEQ ID NO:4相比, 所述酮还原酶多肽中的残基差异为H40R、D150Y和M206F, 以及任选地选自以下的一个或多个残基差异: L17M; L17Q; L17R; A64V; T71P; V87L; G94P; V95M; E105G; I144V; M147I; M147Q; N157C; L196M; D198S; Y199H; Q208R; 和W249Y。

2. 如权利要求1所述的工程化酮还原酶多肽, 其中与SEQ ID NO:4相比, 所述一个或多个残基差异选自: L17Q/R/M; A64V; G94P; I144V; M147Q/I; N157C; 和L196M。

3. 如权利要求1所述的工程化酮还原酶多肽, 所述工程化酮还原酶多肽能够形成非对映体过量于反式-(1-羟基-1-苯基己-5-炔-2-基)氨基甲酸叔丁酯化合物(1b)和(1d)的产物顺式-(1-羟基-1-苯基己-5-炔-2-基)氨基甲酸叔丁酯化合物(1a)和(1c)



4. 如权利要求3所述的工程化酮还原酶多肽, 其中与SEQ ID NO:4相比, 所述一个或多个残基差异选自: L17Q; A64V; T71P; V87L; M147I; N157C; 和L196M。

5. 如权利要求3所述的工程化酮还原酶多肽, 所述工程化酮还原酶多肽能够以相比于化合物(1b)和(1d)的至少50的非对映体比形成产物化合物(1a)和(1c)。

6. 如权利要求1所述的工程化酮还原酶多肽, 其中所述酮还原酶多肽的氨基酸序列是选自由以下组成的组的序列: SEQ ID NO:40、44、68、136、138、140、148、152、154和158。

7. 如权利要求1至6中任一项所述的工程化酮还原酶多肽, 其中所述肽固定在固体支持物上。

8. 一种多核苷酸, 所述多核苷酸编码如权利要求1至6中任一项所述的工程化酮还原酶

多肽。

9. 如权利要求8所述的多核苷酸,所述多核苷酸是选自由以下组成的组的核苷酸序列: 39、43、67、135、137、139、147、151、153和157。

10. 一种表达载体,所述表达载体包含如权利要求8所述的多核苷酸。

11. 如权利要求10所述的表达载体,所述表达载体包含控制序列。

12. 如权利要求11所述的表达载体,其中所述控制序列为启动子。

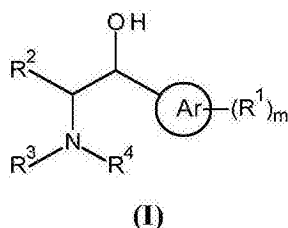
13. 一种宿主细胞,所述宿主细胞包括如权利要求8所述的多核苷酸。

14. 如权利要求13所述的宿主细胞,所述宿主细胞为大肠杆菌(E.coli)。

15. 一种用于制备如权利要求1至6中任一项所述的工程化酮还原酶多肽的方法,包括在适合于表达所述工程化酮还原酶多肽的培养条件下培养权利要求13所述的宿主细胞。

16. 如权利要求15所述的方法,所述方法还包括分离所表达的多肽。

17. 一种用于制备式(I)的产物化合物的方法



其中

Ar是5至7元碳环或杂环;

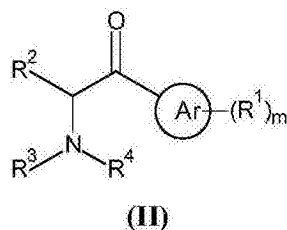
R¹的每次出现是H;

R²是炔基;

R³和R⁴,彼此独立地,选自H和烷氧基羰基,以及

m为0至10;

所述方法包括使式(II)的底物化合物,



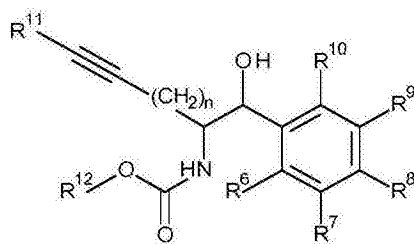
其中R¹、R²、R³、R⁴和m如以上定义,

在适当的反应条件下在辅因子NADPH或NADH的存在下与权利要求1至6中任一项的工程化酮还原酶多肽接触。

18. 如权利要求17所述的方法,其中Ar为5至7元芳基或杂芳基。

19. 如权利要求18所述的方法,其中Ar选自噻吩基、苯基和吡啶基。

20. 如权利要求17所述的方法,其中所述式(I)的产物化合物包括式(Ib)的产物化合物



(Ib)

其中

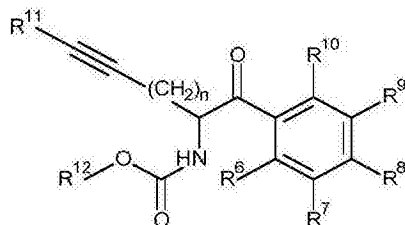
R^6 、 R^7 、 R^8 、 R^9 和 R^{10} 各自独立地是H；

n 为1至4，

R^{11} 是H；以及

R^{12} 是烷基；

所述方法包括使式(IIb)的底物化合物



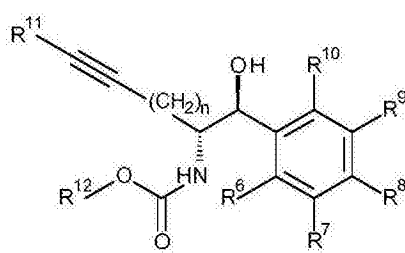
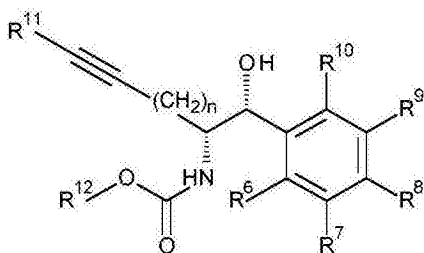
(IIb)

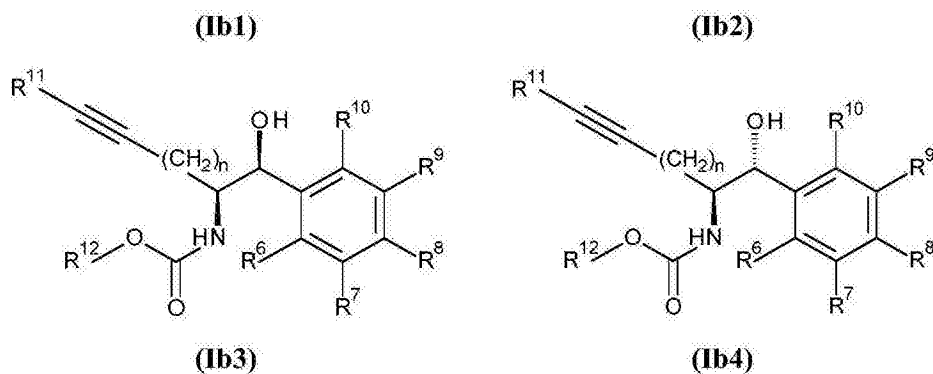
其中

其中 R^6 、 R^7 、 R^8 、 R^9 、 R^{10} 、 R^{11} 、 R^{12} 和 n 如以上定义，

在适当的反应条件下在辅因子NADPH或NADH的存在下与权利要求1至6中任一项的工程化酮还原酶多肽接触。

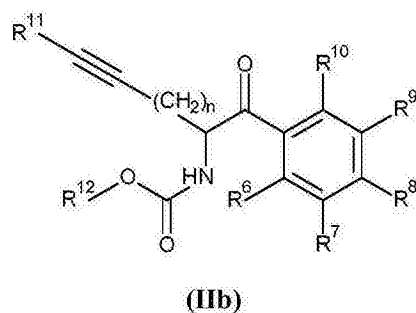
21. 如权利要求20所述的方法，其中所述式(Ib)的产物化合物包括式(Ib1)和(Ib3)的产物化合物，并且所述方法形成非对映体过量于式(Ib2)和(Ib4)的化合物的式(Ib1)和(Ib3)的产物化合物，





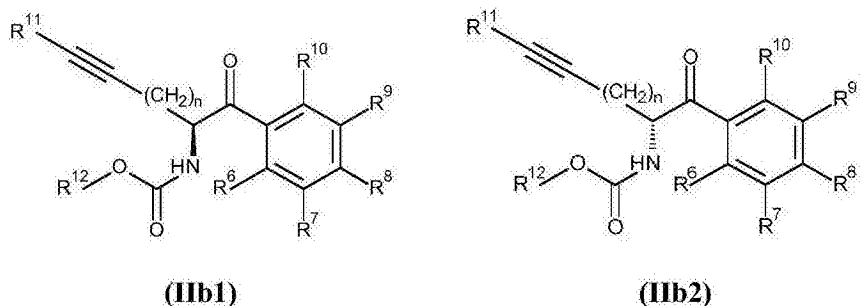
其中

所述方法包括使式(IIb)的底物化合物



在适当的反应条件下在辅因子NADPH或NADH的存在下与权利要求3至5中任一项的工程化酮还原酶多肽接触。

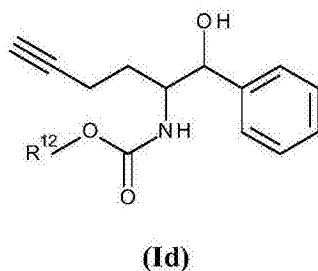
22. 如权利要求20至21中任一项所述的方法, 其中所述式(IIb)的底物化合物包括式(IIb1)和(IIb2)的化合物的混合物



23. 如权利要求22所述的方法, 其中所述混合物包括式(IIb1)和(IIb2)化合物的外消旋混合物。

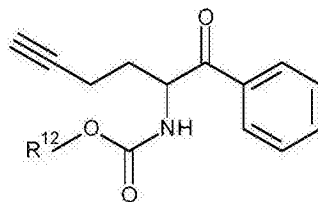
24. 如权利要求20-21和23中任一项所述的方法, 其中n为2。

25. 如权利要求20所述的方法, 其中所述式(Ib)的产物化合物包括式(Id)的化合物



其中

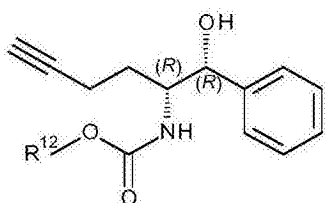
所述方法包括使式(IIId)的底物化合物



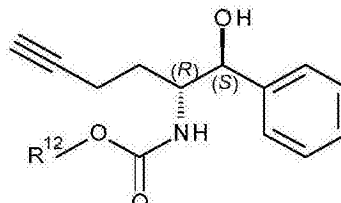
(IIId)

在适当的反应条件下在辅因子NADPH或NADH的存在下与权利要求1至6中任一项的工程化酮还原酶多肽接触。

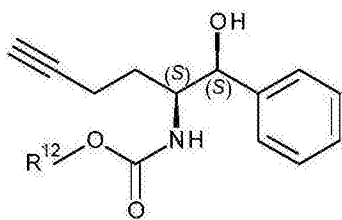
26. 如权利要求20所述的方法,其中所述式(Ib)的产物化合物包括式(Id1)和(Id3)的产物化合物,并且所述方法形成非对映体过量于式(Id2)和(Id4)的化合物的式(Id1)和(Id3)的产物化合物,



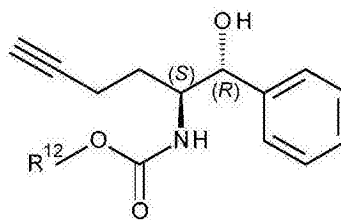
(Id1)



(Id2)



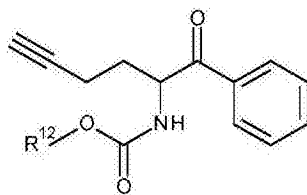
(Id3)



(Id4)

其中

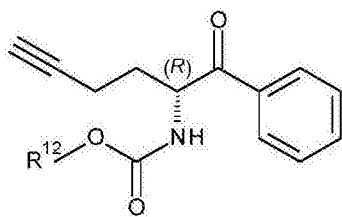
所述方法包括使式(IIId)的底物化合物



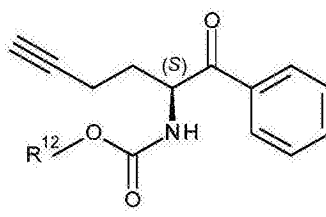
(IIId)

在适当的反应条件下在辅因子NADPH或NADH的存在下与权利要求3至5中任一项的工程化酮还原酶多肽接触。

27. 如权利要求25或26所述的方法,其中所述式(IIId)的底物化合物包括式(IIId1)和(IIId2)的化合物的混合物



(IIId1)



(IIId2)

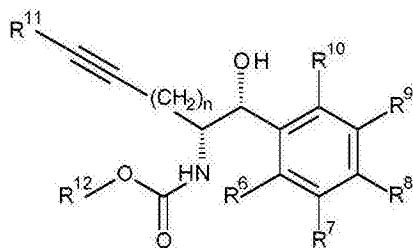
28. 如权利要求27所述的方法, 其中所述混合物是式(IIId1)和(IIId2)的化合物的外消旋混合物。

29. 如权利要求20-21、23、25-26和28中任一项所述的方法, 其中 R^{12} 选自甲基、乙基、正丙基、异丙基、正丁基、异丁基、仲丁基、和叔丁基。

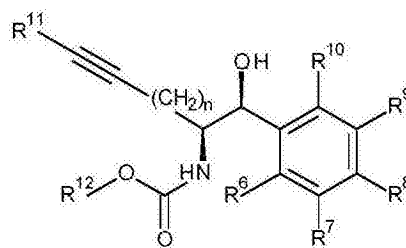
30. 如权利要求29所述的方法, 其中 R^{12} 是叔丁基。

31. 如权利要求26所述的方法, 其中以相比于式(Id2)和(Id4)的化合物的至少50的非对映体比形成所述式(Id1)和(Id3)的产物化合物, 所述方法包括使所述式(IIId)的底物化合物在适当的反应条件下在辅因子NADPH或NADH的存在下与权利要求5的工程化酮还原酶多肽接触。

32. 一种用于制备非对映体过量于式(Ib3)的化合物的式(Ib1)的产物化合物的方法,



(Ib1)



(Ib3)

其中,

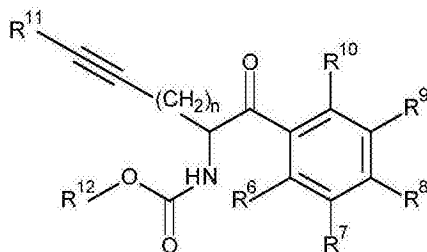
R^6 、 R^7 、 R^8 、 R^9 和 R^{10} 各自独立地是H;

n 为1至4,

R^{11} 是H; 以及

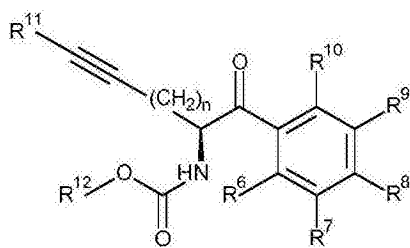
R^{12} 选自烷基;

所述方法包括使式(IIb)的底物化合物

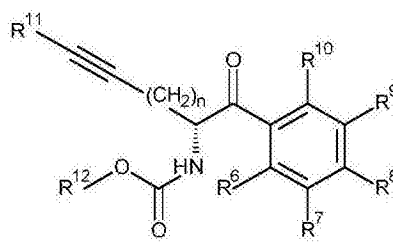


(IIb)

其中 R^6 、 R^7 、 R^8 、 R^9 、 R^{10} 、 R^{11} 、 R^{12} 和 n 如以上定义, 在用于平衡式(IIb1)和(IIb2)的底物化合物的适当的反应条件下在NADPH或NADH的存在下与权利要求3至5中任一项的工程化酮还原酶接触



(IIb1)



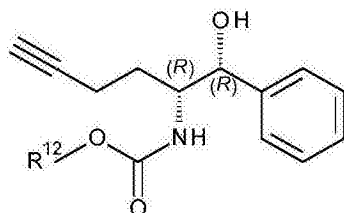
(IIb2)

33. 如权利要求32所述的方法, 其中所述底物化合物(IIb)包括式(IIb1)和(IIb2)的化合物的混合物。

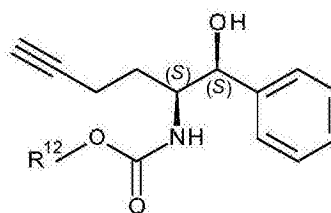
34. 如权利要求33所述的方法, 其中所述混合物包括式(IIb1)和(IIb2)的化合物的外消旋混合物。

35. 如权利要求32至34中任一项所述的方法, 其中n为2。

36. 如权利要求32所述的方法, 其中所述式(Ib1)的产物化合物包括式(Id1)的化合物, 并且形成非对映体过量于式(Id3)的化合物的式(Id1)的化合物,

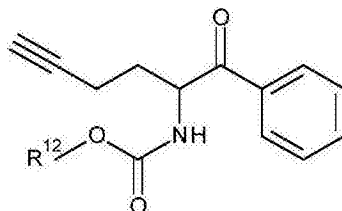


(Id1)



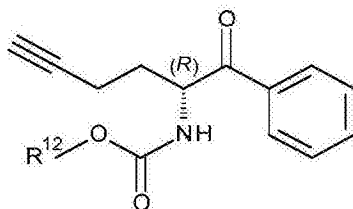
(Id3)

所述方法包括使式(IIId)的底物化合物

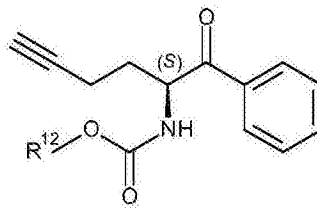


(IIId)

在用于平衡式(IIId1)和(IIId2)的底物化合物的适当的反应条件下在NADPH或NADH的存在下与权利要求3至5中任一项的工程化酮还原酶接触



(IIId1)



(IIId2)

37. 如权利要求36所述的方法, 其中所述式(IIId)的化合物包括式(IIId1)和(IIId2)的底物化合物的混合物。

38. 如权利要求37所述的方法, 其中所述混合物包括式(IIId1)和(IIId2)的底物化合物的外消旋混合物。

39. 如权利要求32-34和36-38中任一项所述的方法,其中用于平衡底物化合物的反应条件包括9至12的pH和45℃至60℃的温度。

40. 如权利要求36所述的方法,其中所述适当的反应条件包括以下底物化合物载量浓度:1g/L、10g/L、20g/L、40g/L、50g/L、75g/L、100g/L、125g/L、150g/L、175g/L、或200g/L。

41. 如权利要求36所述的方法,其中所述适当的反应条件包括以下工程化酮还原酶多肽浓度:20g/L、10g/L、5g/L、4g/L、2g/L、1g/L、0.5g/L、0.2g/L或0.1g/L。

42. 如权利要求17-21、23、25-26、28、30-34、36-38和40-41中任一项所述的方法,所述方法还包括用辅因子再生系统将NAD⁺或NADP⁺转化为NADH或NADPH。

43. 如权利要求42所述的方法,其中所述辅因子再生系统包括酮还原酶和仲醇。

44. 如权利要求43所述的方法,其中所述仲醇为异丙醇(IPA)。

45. 如权利要求44所述的方法,所述方法还包括除去由酮还原酶介导的将所述IPA至丙酮的转化形成的丙酮的步骤。

46. 如权利要求43所述的方法,其中所述辅因子再生系统的所述酮还原酶为权利要求1至6中任一项所述的酮还原酶。

47. 如权利要求43所述的方法,其中所述辅因子再生系统的所述酮还原酶为与权利要求1至6中任一项所述的酮还原酶不同的酮还原酶。

48. 如权利要求36所述的方法,其中所述反应条件包括共溶剂系统,所述共溶剂系统包括水性缓冲溶液和IPA,其中所述IPA浓度为5%(v/v)、10%(v/v)、20%(v/v)、30%(v/v)、40%(v/v)、45%(v/v)、50%(v/v)、55%(v/v)、60%(v/v)、65%(v/v)、70%(v/v)、75%(v/v)、80%(v/v)、85%(v/v)、或90%(v/v)。

49. 如权利要求17所述的方法,其中所述反应条件包括:

(a) 1g/L至10g/L的工程化酮还原酶多肽浓度;

(b) 50g/L至200g/L的载量浓度下的底物化合物;

(c) 0.1g/L至0.5g/L的NADPH浓度;

(d) 水性缓冲液和30%至70%(v/v)的IPA的共溶剂溶液;以及

(e) 35℃至60℃的温度。

50. 如权利要求17-21、23、25-26、28、30-34、36-38、40-41和43-49中任一项所述的方法,所述方法还包括选自以下的一个或多个步骤:提取所述产物化合物;分离所述产物化合物;纯化所述产物化合物;和结晶所述产物化合物。

51. 如权利要求17-21、23、25-26、28、30-34、36-38、40-41和43-49中任一项所述的方法,其中所述式(II)的底物化合物是所述化合物的氘化形式。

用于制备羟基取代的氨基甲酸酯的生物催化剂

技术领域

[0001] 本公开内容涉及用于生产手性化合物的生物催化剂介导的方法和在该方法中使用的生物催化剂。

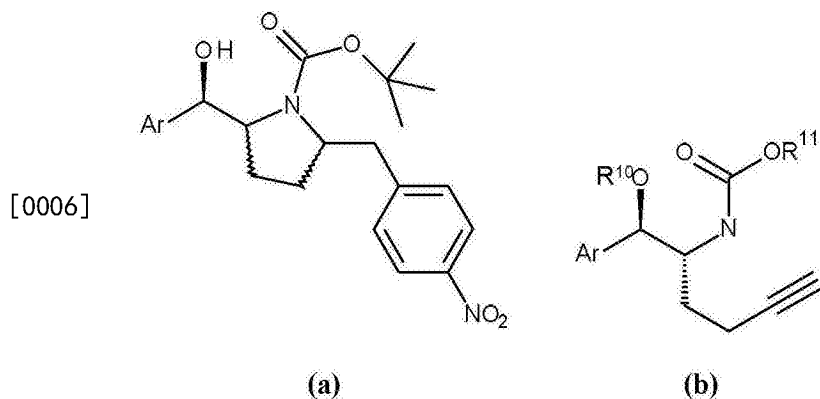
[0002] 参考序列列表、表格或计算机程序

[0003] 序列列表的正式复本作为ASCII格式化文本文件与说明书经EFS-Web同时被提交,文件名为“CX2-103W01_ST25.txt”,创建日期为2012年11月13日,且大小为290,640字节。相对于2011年11月18日随母案美国临时申请61/561,665提交的具有文件名“CX2-103USP1_ST25.txt”的序列列表,该序列列表仅有较少的格式更改。经EFS-Web提交的序列列表为本说明书的一部分,并且据此通过引用整体并入本文。

背景技术

[0004] β -肾上腺素能受体为G蛋白偶联受体超家族的成员,并通过将 G_{as} 的刺激与腺苷酸环化酶的激活偶联,介导对外部配体,特别是肾上腺素和去甲肾上腺素的响应,导致增加的细胞内第二信使cAMP。已知 β -肾上腺素能受体有三个亚型。 β_1 -受体主要表达于心脏、冠状动脉、肾、肌肉和中枢神经系统,并影响心脏刺激、冠状血管扩张以及结肠和食管的松弛。在 β_1 -受体中的多态性与几种形式的高血压和心脏衰竭相关并影响静息心率。 β_2 -肾上腺素能受体发现于肺、胃肠道、肝、子宫、血管平滑肌、心脏和骨骼肌中,在那里其部分地调节平滑肌松弛(例如,支气管扩张)。 β_3 -肾上腺素能受体主要表达于脂肪组织(棕色和白色)、膀胱、胆囊、结肠和心脏。

[0005] 各个肾上腺素能受体亚型的种间特异性表达存在差异,提示不同亚型在不同物种的相同组织中的作用。在人中, β_3 -肾上腺素能受体表达于膀胱逼尿肌,且 β_3 -肾上腺素能受体mRNA是遭受根治性膀胱切除术的患者中的主要mRNA。此外, β_1 和 β_2 选择性激动剂对逼尿肌功能无影响,而 β_3 -选择性激动剂产生浓度依赖性舒张(Igawa Y.等人,1999,Br J Pharmacol.126:819-25)。这些观察结果导致了开发 β_3 -肾上腺素能受体激动剂用于治疗各种形式的膀胱过度活动综合征(参见,例如,Igawa Y.等人,2010,Korean J Urol.51:811-818)。基于羟甲基吡咯烷的 β_3 -肾上腺素能受体特异性激动剂描述于W02009124167和W02009124166中。羟基吡咯烷药物的合成可使用具有结构(a)和(b)的中间体,其中Ar表示芳基基团,且 R^{10} 和 R^{11} 表示保护基团。



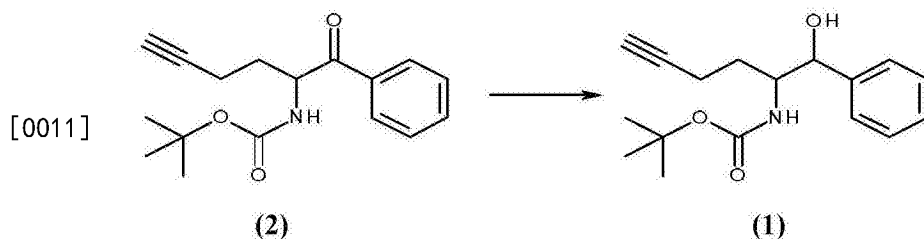
[0007] 中间体(a)作为顺式和反式异构体存在,所述顺式和反式异构体必须进行分离和纯化,以获得用于合成最终药物产物的手性中间体。用于获得中间体(b)的合成途径使用了起始化合物5-己炔酸并与(S)-(-)-4-苄基或(S)-(-)-4-苯基-2-噁唑烷酮反应以使羟基基团的手性位于吡咯烷基的左侧(参见,例如,通过引用并入本文的W02009124167和W02009124166)。

[0008] 鉴于需要分离手性中间体或使用具有手性特异的试剂来合成 β_3 -肾上腺素能受体激动剂,希望发现可选择的合成途径,特别是导致感兴趣的特定立体异构体的高度非对映体过量,提供起始材料至所需产物的高效率的转化率,使用温和条件且避免使用具有手性的特定化学试剂的方法。

发明内容

[0009] 本公开内容提供了非天然存在的具有酮还原酶活性的工程化多肽、编码多肽的多核苷酸、制备多肽的方法以及使用多肽将底物化合物(2)(1-氧代-1-苯基己-5-炔-2-基)氨基甲酸叔丁酯或其结构类似物转化为产物化合物(1),(1-羟基-1-苯基己-5-炔-2-基)氨基甲酸叔丁酯或相应结构类似物的方法,如方案1中显示的。

[0010] 方案1



[0012] 为了方案1的转化率中的改进特性而开发的工程化酮还原酶多肽与来自克菲尔乳杆菌(*Lactobacillus kefir*)的具有SEQ ID NO:2的氨基酸序列的天然存在的酮还原酶多肽或具有SEQ ID NO:4的氨基酸序列的参考工程化酮还原酶相比具有一个或多个残基差异。这些残基差异在影响以下方面的残基位置出现:活性、立体选择性、底物结合、稳定性(热和溶剂)、表达、及其各种组合以及其他。

[0013] 在一个方面,工程化酮还原酶多肽包括以下氨基酸序列:具有与选自以下的参考序列的至少80%序列同一性:SEQ ID NO:4、6、8、10、12、14、16、18、20、22、24、28、30、32、34、36、38、40、42、44、46、48、50、52、54、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、96、98、102、104、106、108、110、114、116、122、124、126、128、130、132、136、138、140、

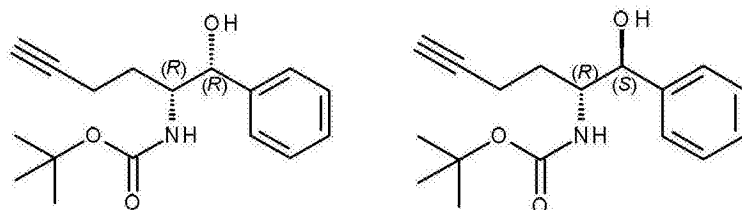
142、144、146、148、150、152、154、156、158、160、162、164、168、170和172,以及与SEQ ID NO: 2或4相比选自以下的一个或多个残基差异:X7S;X17M;X17Q;X17R;X23V;X27L;X29G;X40R;X60I;X64V;X71P;X87L;X94A;X94P;X94S;X95M;X96L;X96Y;X105G;X113I;X122A;X127R;X131S;X144V;X145L;X147I;X147L;X147Q;X150Y;X152G;X153G;X157C;X173L;X195A;X195G;X196M;X198S;X199H;X206F;X206L;X208R;X216R;X221S;X243S;X245I;X249F;X249G;和X249Y。如本文所提供,在一些实施方案中,公开的氨基酸差异可单独使用或以各种组合使用以产生具有改进的酶特性的工程化酮还原酶多肽。

[0014] 在一些实施方案中,酮还原酶多肽包括具有与SEQ ID NO:4的至少80%序列同一性和特征X40R和X206L或F的氨基酸序列,条件是当X206为F时,序列还包括与SEQ ID NO:4相比选自以下的一个或多个残基差异:X7S;X17M;X17Q;X17R;X23V;X27L;X29G;X60I;X64V;X71P;X87L;X94A;X94P;X94S;X95M;X96L;X96Y;X105G;X113I;X122A;X127R;X131S;X144V;X145L;X147I;X147L;X147Q;X150Y;X152G;X153G;X157C;X173L;X195A;X195G;X196M;X198S;X199H;X208R;X216R;X221S;X243S;X245I;X249F;X249G;和X249Y。

[0015] 在一些实施方案中,工程化酮还原酶包括具有特征X40R和X206F或L,以及与SEQ ID NO:4相比选自以下的至少一个或多个残基差异的氨基酸序列:X17Q/R/M;X64V;X94P;X144V;X147Q/I/L;X157C;和X196M。

[0016] 在一些实施方案中,在将底物化合物(2)转化为产物化合物(1)中具有改进特性的工程化酮还原酶多肽包括选自以下的氨基酸序列:SEQ ID NO:4、6、8、10、12、14、16、18、20、22、24、28、30、32、34、36、38、40、42、44、46、48、50、52、54、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、96、98、102、104、106、108、110、114、116、122、124、126、128、130、132、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、168、170和172,以及特别是选自以下的氨基酸序列:SEQ ID NO:18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172。

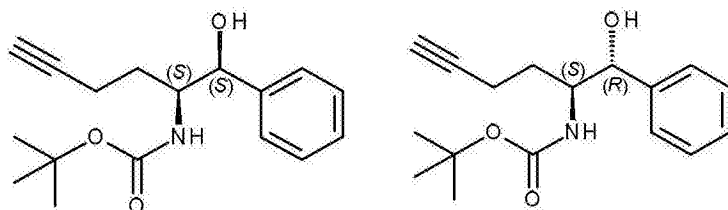
[0017] 在一些实施方案中,工程化酮还原酶多肽能够在适当的反应条件下形成非对映体过量于(in diastereomeric excess of)化合物(1b)和(1d)的产物化合物(1a)和(1c)



(1a)

(1b)

[0018]



(1c)

(1d)

[0019] 在一些实施方案中,工程化酮还原酶多肽能够在适当的反应条件下形成非对映体过量于化合物(1c)的产物化合物(1a)。

[0020] 在一些实施方案中,能够形成非对映体过量于化合物(1b)和(1d)的产物化合物(1a)和(1c),或能够形成非对映体过量于化合物(1c)的产物化合物(1a)的工程化酮还原酶多肽包括,具有特征X40R和206F或L,以及与SEQ ID NO:4相比选自以下的至少一个或多个残基差异的氨基酸序列:X17Q;X64V;X71P;X87L;X147I;X157C;X196M;和X249F。

[0021] 展示相关立体选择性的示例性多肽可选自SEQ ID NO:18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172。

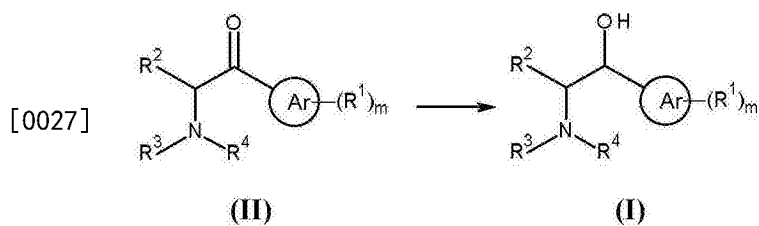
[0022] 在一些实施方案中,工程化酮还原酶能够在适当的反应条件下形成非对映体过量于化合物(1a)和(1c)的化合物(1b)和(1d)。在一些实施方案中,能够形成非对映体过量于化合物(1a)和(1c)的化合物(1b)和(1d)的工程化酮还原酶包括具有特征X195G的氨基酸序列。展示该立体选择性的示例性多肽可选自SEQ ID NO:50和74的多肽。

[0023] 在一些实施方案中,工程化酮还原酶多肽被固定在固体支持物上。在一些实施方案中,固体支持物是包括具有环氧官能团的聚甲基丙烯酸酯、具有氨基环氧官能团的聚甲基丙烯酸酯、具有十八烷基官能团的苯乙烯/DVB共聚物或聚甲基丙烯酸酯的珠或树脂。

[0024] 在另一个方面,本公开内容提供了编码工程化酮还原酶的多核苷酸,以及包括该多核苷酸的表达载体,以及能够表达编码该多肽的多核苷酸的宿主细胞。在一些实施方案中,本公开内容还提供了制备工程化酮还原酶多肽的方法。编码工程化酮还原酶的示例性多核苷酸序列包括以下的序列:SEQ ID NO:5、7、9、11、13、15、17、19、21、23、25、27、29、31、33、35、37、39、41、43、45、47、49、51、53、55、57、59、61、63、65、67、69、71、73、75、77、79、81、83、85、87、89、91、93、95、97、99、101、103、105、107、109、111、113、115、117、119、121、123、125、127、129、131、133、135、137、139、141、143、145、147、149、151、153、155、157、159、161、163、165、167、169和171,特别是以下的序列:SEQ ID NO:17、19、21、23、25、27、29、31、33、35、37、39、41、43、55、57、59、61、63、65、67、69、71、73、75、77、79、81、83、85、87、89、91、93、95、97、99、101、103、105、107、109、111、113、115、117、119、121、123、125、127、129、131、133、135、137、139、141、143、145、147、149、151、153、155、157、159、161、163、165、167、169和171。

[0025] 在另一个方面,工程化酮还原酶多肽可在用于制备式(I)的羟基化合物的方法中被使用,式(I)的羟基化合物可用于合成药物化合物。因此,在一些实施方案中,工程化酮还原酶可用于将式(II)的底物化合物转化为式(I)的产物化合物,如方案2中所示

[0026] 方案2



[0028] 其中

[0029] Ar是5至7元碳环或杂环；

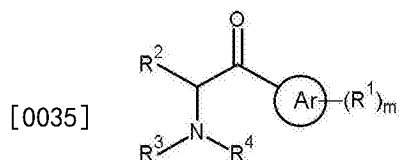
[0030] R^1 的每次出现独立地选自卤素、 $-\text{COOR}^a$ 、 $-\text{C}(\text{O})\text{R}^b$ 、 $-\text{OR}^c$ 、 $-\text{SO}_2$ 、 $-\text{SR}^d$ 、 $-\text{S}(\text{O})\text{R}^e$ 、 $-\text{NR}^f\text{R}^g$ 、 $-\text{C}(\text{O})\text{NR}^h\text{R}^i$ 、 $-\text{NO}_2$ 、 $-\text{CN}$ 、烷基、杂烷基、环烷基、环烷基烷基、杂环烷基、杂环烷基烷基、芳基、芳基烷基、杂芳基和杂芳基烷基，其中 R^a 、 R^b 、 R^c 、 R^d 、 R^e 、 R^f 、 R^g 、 R^h 和 R^i 独立地选自H和任选地取代的烷基、杂烷基、环烷基、环烷基烷基、杂环烷基、杂环烷基烷基、芳基、芳基烷基、杂芳基、和杂芳基烷基；

[0031] R^2 选自任选地取代的烷基、烯基、炔基、杂烷基、环烷基、环烷基烷基、环烷基烯基、环烷基炔基、杂环烷基、杂环烷基烷基、杂环烷基烯基、杂环烷基炔基、芳基、芳基烷基、芳基烯基、芳基炔基、杂芳基、杂芳基烷基、杂芳基烯基、杂芳基炔基、烷基硫代、烷基亚磺酰基和芳基亚磺酰基；

[0032] R^3 和 R^4 ，彼此独立地，选自H、任选地取代的烷基、环烷基、环烷基烷基、杂环烷基、杂环烷基烷基、芳基、芳基烷基、杂芳基、杂芳基烷基、烷氧基羰基、芳基烷氧基羰基、杂芳基烷氧基羰基、和保护基团，或 R^3 和 R^4 中的一个与 R^2 形成任选地取代的5至7元杂环，以及

[0033] m 为0至10。

[0034] 因此，用于制备式(I)的化合物的方法可包括使式(II)的化合物

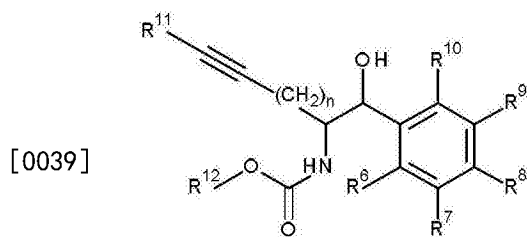


(II)

[0036] 其中 R^1 、 R^2 、 R^3 、 R^4 和 m 如以上定义，

[0037] 在适当的反应条件下在辅因子NADPH或NADH的存在下与工程化酮还原酶多肽接触。

[0038] 在方法的一些实施方案中，式(I)化合物包括式(Ib)的化合物



(Ib)

[0040] 其中

[0041] R^6 、 R^7 、 R^8 、 R^9 、和 R^{10} 各自独立地选自H、卤素、 $-\text{COOR}^a$ 、 $-\text{C}(\text{O})\text{R}^b$ 、 $-\text{OR}^c$ 、 $-\text{SO}_2$ 、 $-\text{SR}^d$ 、 $-\text{S}(\text{O})\text{R}^e$ 、 $-\text{NR}^f\text{R}^g$ 、 $-\text{C}(\text{O})\text{NR}^h\text{R}^i$ 、 $-\text{NO}_2$ 、 $-\text{CN}$ 、烷基、杂烷基、环烷基、环烷基烷基、杂环烷基、杂环烷基烷基、芳基、芳基烷基、杂芳基和杂芳基烷基，其中 R^a 、 R^b 、 R^c 、 R^d 、 R^e 、 R^f 、 R^g 、 R^h 和 R^i 独立地选自H和任选地取代的烷基、杂烷基、环烷基、环烷基烷基、杂环烷基、杂环烷基烷基、芳基、芳基烷基、杂芳基、和杂芳基烷基；

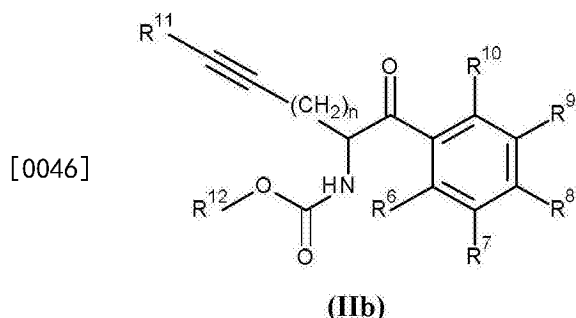
[0042] n 为1至4，

[0043] R^{11} 选自H和任选地取代的烷基、杂烷基、环烷基、环烷基烷基、芳基、芳基烷基、杂芳

基、和杂芳基烷基；以及

[0044] R^{12} 选自H、任选地取代的烷基、环烷基、环烷基烷基、杂环烷基、杂环烷基烷基、芳基、芳基烷基、杂芳基、杂芳基烷基和保护基团。

[0045] 式(Ib)的化合物可通过使式(IIb)的底物化合物

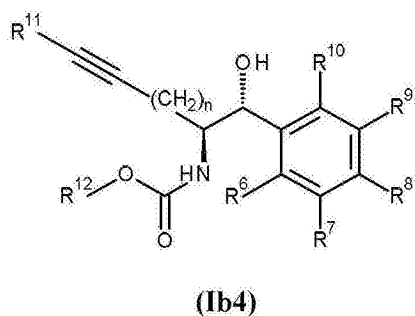
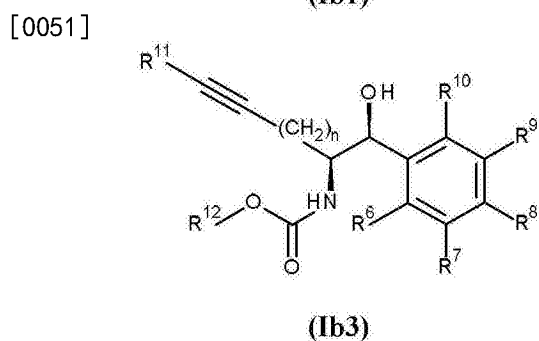
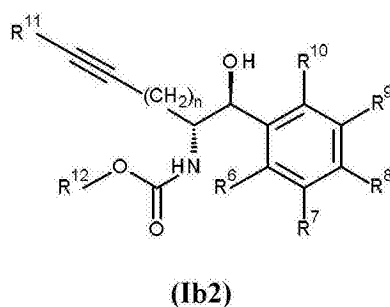
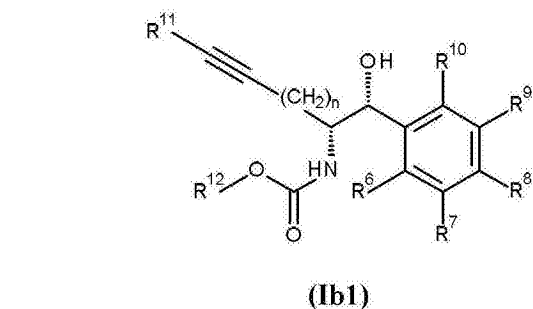


[0047] 其中

[0048] 其中 R^6 、 R^7 、 R^8 、 R^9 、 R^{10} 、 R^{11} 、 R^{12} 和 n 如以上定义，

[0049] 在适当的反应条件下在辅因子NADPH或NADH的存在下与公开内容的工程化酮还原酶多肽接触而制备。

[0050] 在一些实施方案中，式(Ib)的产物化合物包括式(Ib1)和(Ib3)的化合物，并且该方法形成非对映体过量于式(Ib2)和(Ib4)的化合物的式(Ib1)和(Ib3)的产物化合物



[0052] 因此，在一些实施方案中，用于制备非对映体过量于式(Ib2)和(Ib4)的化合物的式(Ib1)和(Ib3)的产物化合物的方法包括，使式(IIb)的底物化合物与相比于对式(Ib2)和(Ib4)的化合物具有对式(Ib1)和(Ib3)的化合物的非对映选择性的工程化酮还原酶接触。

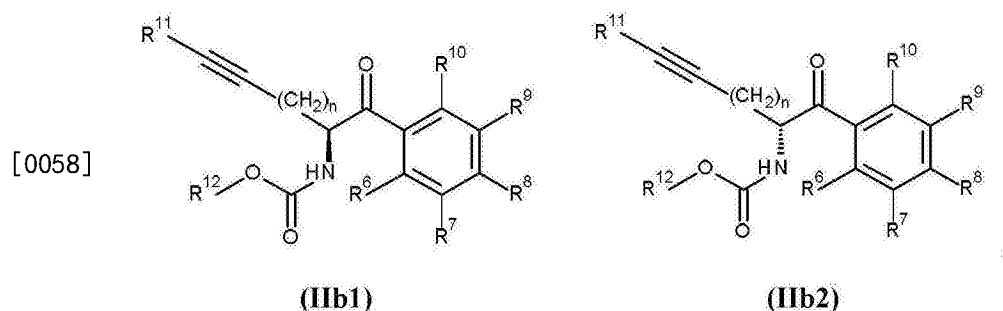
[0053] 在一些实施方案中，式(Ib)的产物化合物包括式(Ib1)的化合物，并且该方法形成非对映体过量于式(Ib3)的化合物的式(Ib1)的产物化合物。因此，在一些实施方案中，用于制备非对映体过量于式(Ib3)的化合物的式(Ib1)的产物化合物的方法包括，使式(IIb)的底物化合物与相比于对式(Ib3)的化合物具有对式(Ib1)的化合物的非对映选择性的工程化酮还原酶接触。

[0054] 相比于对式(Ib2)和(Ib4)的化合物具有对式(Ib1)和(Ib3)的化合物的非对映选择性或相比于对式(Ib3)的化合物具有对式(Ib1)的化合物的非对映选择性的示例性工程化酮还原酶多肽可包括选自以下的氨基酸序列:SEQ ID NO:18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172。

[0055] 在一些实施方案中,式(Ib)的产物化合物包括式(Ib2)和(Ib4)的化合物,并且该方法形成非对映体过量于式(Ib1)和(Ib3)的化合物的式(Ib2)和(Ib4)的产物化合物。在一些实施方案中,用于制备非对映体过量于式(Ib1)和(Ib3)的化合物的式(Ib2)和(Ib4)的产物化合物的方法包括,使式(IIb)的底物化合物与相比于对式(Ib1)和(Ib3)的化合物具有对式(Ib2)和(Ib4)的化合物的非对映选择性的工程化酮还原酶接触。

[0056] 相比于对式(Ib1)和(Ib3)的化合物具有对式(Ib2)和(Ib4)的化合物的非对映选择性的示例性工程化酮还原酶多肽可包括选自SEQ ID NO:50和74的氨基酸序列。

[0057] 在一些实施方案中,用于制备非对映体过量于式(Ib2)和(Ib4)的化合物的式(Ib1)和(Ib3)的化合物、制备非对映体过量于式(Ib3)的化合物的式(Ib1)的化合物、以及制备非对映体过量于式(Ib1)和(Ib3)的化合物的式(Ib2)和(Ib4)的化合物的方法可在促进式(IIb1)和(IIb2)的底物化合物之间的平衡的适当的反应条件下进行



[0059] 在转化为产物的过程中底物化合物(IIb1)和(IIb2)之间的平衡可通过动态动力学拆分来将产物化合物的产量增加至超过存在于反应溶液中的式(IIb1)和(IIb2)的化合物的初始浓度。

[0060] 在本文使用式(IIb)的底物化合物的方法的一些实施方案中,n选自2、3或4。在一些实施方案中,n为2。在一些实施方案中,n为2且R⁶、R⁷、R⁸、R⁹、R¹⁰和R¹¹为H。

[0061] 用于进行生物催化方法的参数,包括:底物化合物载量、酶载量、辅因子载量、溶剂条件(例如,缓冲液、异丙醇等)、pH和温度以及其他,被进一步描述于下面的详述中。

[0062] 在一些实施方案中,适当的反应条件可包括:(a)约1g/L至约10g/L的工程化酮还原酶多肽浓度;(b)约50g/L至约200g/L的载量浓度下的底物化合物;(c)约0.1g/L至约0.5g/L的NADP(H)浓度;(d)水性缓冲液和约30%至约70%(v/v)的IPA的共溶剂溶液;以及(e)35℃至约60℃的温度。在一些实施方案中,适当的反应条件可包括:(a)约0.1g/L至约1g/L的工程化酮还原酶多肽浓度;(b)约5g/L至约50g/L的载量浓度下的底物化合物;(c)约0.01g/L至约0.1g/L的NADP(H)浓度;(d)水性缓冲液和约30%至约70%(v/v)的IPA的共溶剂溶液;以及(e)约30℃至约45℃的温度。

[0063] 在一些实施方案中,用于动态动力学拆分的用于平衡式(IIb1)和(IIb2)的底物化合物的反应条件可包括约9至约12的pH和约45℃至约60℃的温度。

[0064] 各种实施方案以及关于制备和使用工程化多肽、编码多肽的多核苷酸的指导,和用于生物催化转化的方法将在以下详述中提供。

具体实施方式

[0065] 除非上下文另外清楚地指明,否则对于本文中该说明书和所附的权利要求书中的单数形式“一(a)”、“一(an)”和“该(the)”包括复数指示物。因此,例如,提及“一种多肽”包括多于一种多肽,以及提及“一种化合物”是指多于一种化合物。

[0066] 还应理解的是,当各个实施方案的描述使用术语“包括(comprising)”时,本领域技术人员将理解,在一些特定情况下,一种实施方案可替代地利用措辞“基本上由…组成”或“由…组成”来描述。

[0067] 并且,除非另有说明,使用“或”意指“和/或”。类似地,“包括(comprise)”、“包括(comprises)”、“包括(comprising)”、“包括(include)”、“包括(includes)”、“包括(including)”、“具有(has)”、“具有(have)”和“具有(having)”是可互换的,且不意图为限制性的。

[0068] 上述的一般说明和详述都只是示例性和说明性的,而不是限制本公开内容。

[0069] 本文所用的章节标题仅用于组织目的,且不应当被解释为限制所描述的主题。

[0070] 缩写

[0071] 用于遗传编码的氨基酸的缩写是常规的,并如下:

	氨基酸	三个字母的缩写	一个字母的缩写
	丙氨酸	Ala	A
	精氨酸	Arg	R
	天冬酰胺	Asn	N
	天冬氨酸	Asp	D
	半胱氨酸	Cys	C
	谷氨酸	Glu	E
	谷氨酰胺	Gln	Q
	甘氨酸	Gly	G
	组氨酸	His	H
[0072]	异亮氨酸	Ile	I
	亮氨酸	Leu	L
	赖氨酸	Lys	K
	蛋氨酸	Met	M
	苯丙氨酸	Phe	F
	脯氨酸	Pro	P
	丝氨酸	Ser	S
	苏氨酸	Thr	T
	色氨酸	Trp	W
	酪氨酸	Tyr	Y
	缬氨酸	Val	V

[0073] 当使用三个字母缩写时,除非前面明确加有“L”或“D”或从使用缩写的上下文明显,否则氨基酸可为关于 α -碳(C_{α})成L-构型或D-构型。例如,尽管“Ala”表示丙氨酸,而没有规定关于 α -碳的构型,但“D-Ala”与“L-Ala”分别表示D-丙氨酸与L-丙氨酸。当使用单字母缩写时,大写字母表示关于 α -碳成L-构型的氨基酸类,而小写字母表示关于 α -碳成D-构型的氨基酸类。例如,“A”表示L-丙氨酸而“a”表示D-丙氨酸。当多肽序列被呈现为一串单个字母或三个字母缩写(或其混合)时,根据通常惯例,序列以氨基(N)至羧基(C)方向呈现。

[0074] 用于基因编码的核苷的缩写是常规的,并且是如下:腺苷(A);鸟苷(G);胞苷(C);胸苷(T);以及尿苷(U)。除非特别描绘,否则缩写的核苷可为核糖核苷或2'-脱氧核糖核苷。核苷可以单个计或以聚集体计被指定为核糖核苷或2'-脱氧核糖核苷。当核酸序列被呈现为一串一个字母缩写时,根据通常惯例,序列以5'至3'方向呈现,且没有显示磷酸盐。

[0075] 定义

[0076] 关于本公开内容,除非另外具体指明,否则本文的说明书中使用的技术术语和科学术语具有本领域普通技术人员通常理解的含义。因此,以下术语意为具有以下含义:

[0077] “蛋白”、“多肽”和“肽”在本文中可互换使用,表示通过酰胺键共价连接的至少两个氨基酸的聚合物,而不论长度或翻译后修饰(例如,糖基化、磷酸化、脂化、肉豆蔻化(myristilation)、泛素化等等)。包括在这一定义中的是D-氨基酸和L-氨基酸,以及D-氨基酸和L-氨基酸的混合物。

[0078] “多核苷酸”或“核酸”是指共价地连接在一起的两个或更多个核苷。多核苷酸可完全由核糖核苷(即,RNA)组成,完全由2'-脱氧核糖核苷酸(即,DNA)组成或由核糖核苷和2'-脱氧核糖核苷的混合物组成。尽管核苷将通常通过标准磷酸二酯键合连接在一起,但多核苷酸可包括一个或多个非标准的键合。多核苷酸可以是单链的或双链的,或可包括单链区和双链区两者。此外,虽然多核苷酸将通常由天然存在的编码核碱基(即,腺嘌呤、鸟嘌呤、尿嘧啶、胸腺嘧啶和胞嘧啶)组成,但其可包括一个或多个修饰的和/或合成的核碱基,诸如,例如,肌苷、黄嘌呤、次黄嘌呤等。优选地,这种修饰的或合成的核碱基将为编码核碱基。

[0079] “编码序列”是指编码蛋白的氨基酸序列的核酸(例如,基因)的那部分。

[0080] “天然存在的”或“野生型”是指天然发现的形式。例如,天然存在的或野生型多肽或多核苷酸序列是生物体中存在的序列,其可从天然来源分离且未通过人为操纵而被有意识地修改。

[0081] 当关于例如细胞、核酸或多肽在本公开内容使用时,“非天然存在的”或“工程化的”或“重组的”指将以天然不存在其他方式被修饰或与其相同但由合成的材料产生或衍生和/或通过使用重组技术处理产生的材料或与该材料的自然或天然形式相应的材料。非限制性实例包括,表达在细胞的天然(非重组的)形式中未发现的基因或以与天然基因本来表达的水平不同的水平表达天然基因的重组细胞以及其他。

[0082] “序列同一性的百分比”、“百分比同一性”和“百分比同一”在本文使用以指多核苷酸序列或多肽序列之间的对比,并通过在比较窗上比较两个最佳比对序列来确定,其中多核苷酸或多肽序列在比较窗中的部分与用于两个序列的最佳比对的参考序列相比可以包括添加或缺失(即,缺口)。百分比如下计算,通过确定两个序列中出现相同的核酸碱基或氨基酸残基或者核酸碱基或氨基酸残基与缺口对齐的位置的数目,以产生匹配位置的数目,将匹配位置的数目除以比较窗中位置的总数目,并将结果乘以100以得到序列同一性的百

分比。最佳比对和百分比序列同一性的确定使用BLAST和BLAST 2.0算法来进行(参见,例如,Altschul等人,1990,J.Mol.Biol.215:403-410以及Altschul等人,1977,Nucleic Acids Res.3389-3402)。用于进行BLAST分析的软件为通过美国国家生物技术信息中心网站公共可获得的。

[0083] 简言之,BLAST分析涉及首先通过确定查询序列(query sequence)中长度W的短字来确定高分序列对(HSP),当其与数据库序列中相同长度的字比对时,所述短字匹配或满足一些正值阈值评分T。T被称为邻近字评分阈值(Altschul等人,如上)。这些最初的邻近字击中(word hit)用作启动检索以找到更长的包括它们的HSP的种子。然后字击中沿着每个序列的两个方向延伸到累积比对评分不能增加的程度。对于核苷酸序列,累积评分使用参数M(用于一对匹配残基的奖励评分;总是 >0)和N(用于错配残基的惩罚评分;总是 <0)来计算。对于氨基酸序列,使用评分矩阵以计算累积评分。当:累积比对评分从其最大获得的值跌落量X时;由于一个或多个负评分残基比对的累积,累积评分变成零或以下时;或达到任一序列的末端时,字击中在每个方向的延伸被停止。BLAST算法参数W、T和X确定比对的灵敏度和速度。BLASTN程序(对于核苷酸序列)使用11的字长(W)、10的期望值(E)、 $M=5$ 、 $N=-4$ 、以及两个链的比较作为缺省参数。对于氨基酸序列,BLASTP程序使用3的字长(W)、10的期望值(E),和BLOSUM62评分矩阵作为缺省参数(参见Henikoff和Henikoff,1989,Proc Natl Acad Sci USA 89:10915)。

[0084] 在提供两个序列的百分比同一性中与BLAST类似地起作用的许多其他算法是可用的。用于比较的最佳序列比对可如下进行,例如,通过Smith和Waterman,1981,Adv.Appl.Math.2:482的局部同源性算法,通过Needleman和Wunsch,1970,J.Mol.Biol.48:443的同源性比对算法,通过Pearson和Lipman,1988,Proc.Natl.Acad.Sci.USA 85:2444的相似度检索方法,通过这些算法的计算机化执行(在GCG Wisconsin软件包中的GAP、BESTFIT、FASTA和TFASTA),或通过目测(通常参见,Current Protocols in Molecular Biology,F.M.Ausubel等人编著,Current Protocols,Green Publishing Associates, Inc.和John Wiley&Sons, Inc.,(1995增刊)(Ausubel))。另外,序列比对与百分比序列同一性的确定可使用GCG Wisconsin软件包(Accelrys,Madison WI)中的BESTFIT或GAP程序,使用提供的缺省参数。

[0085] “参考序列”是指与另一个序列比较的确定的序列。参考序列可以是更大序列的子集,例如,全长基因或多肽序列的区段。通常,参考序列为至少20个核苷酸或氨基酸残基的长度,至少25个残基的长度,至少50个残基的长度,或核酸或多肽的全长。因为两种多核苷酸或多肽可以各自(1)包括两个序列之间相似的序列(即,完整序列的一部分),且(2)还可包括在两种序列之间不同的序列,所以两种(或更多种)多核苷酸或多肽之间的序列比较通常通过在比较窗内比较两种多核苷酸的序列来进行以确定和比较具有序列相似性的局部区域。术语“参考序列”并非意图限于野生型序列,而可包括工程化或改变的序列。例如,在一些实施方案中,“参考序列”可以是先前工程化或改变的氨基酸序列。

[0086] “比较窗”是指至少约20个连续核苷酸位置或氨基酸残基的概念性片段,其中序列可与至少20个连续的核苷酸或氨基酸的参考序列进行比较,并且其中在比较窗中序列的一部分与用于两个序列的最佳比对的参考序列(其不包括添加或缺失)相比,可以包括20%或更少的添加或缺失(即,缺口)。比较窗可以比20个连续的残基更长,并任选地包括30、40、

50、100或更长的窗。

[0087] 当在给定的氨基酸或多核苷酸序列的编号的上下文中使用时,“相应于”、“参考”或“相对于”指当给定的氨基酸或多核苷酸序列与参考序列相比时,指定的参考序列的残基的编号。换句话说,给定的聚合物的残基编号或残基位置是关于参考序列被指定的,而不是通过给定的氨基酸或多核苷酸序列内的残基的实际数值位置被指定的。例如,给定的氨基酸序列,诸如工程化酮还原酶的氨基酸序列,可以通过引入缺口以优化两个序列之间的残基匹配而与参考序列进行比对。在这些情况中,虽然存在缺口,给定的氨基酸或多核苷酸序列中残基的编号是关于与其比对的参考序列作出的。如本文所用,除非另外特别说明,否则提及残基位置,诸如,如下面进一步描述的“Xn”,将被解释为是指“相应于...的残基”。

[0088] “立体选择性”是指在化学反应或酶促反应中一种立体异构体比另一种立体异构体或另一组立体异构体优先形成。立体选择性可以是部分的,其中一种立体异构体的形成优于另一种立体异构体的形成,或立体选择性可以是完全的,其中只形成一种立体异构体。当立体异构体是对映体时,立体选择性被称为对映体选择性,即一种对映体在两种对映体的总和中的分数(通常被报告为百分比)。它在本领域通常可选择地被报告为(通常为百分比)对映体过量(e.e.),其中根据以下式计算[主要对映体-次要对映体]/[主要对映体+次要对映体]。当立体异构体是非对映异构体时,立体选择性被称为非对映选择性,即一种非对映体在两种非对映体的混合物中的分数(通常被报告为百分比),通常可选择地报告为非对映体过量(d.e.)。对映体过量和非对映体过量是立体异构体过量的类型。还应当理解,立体选择性不局限于单一立体异构体,而是可对立体异构体的组描述,例如相比于对化合物(1b)和(1d)对化合物(1a)和(1c)的立体选择性,如本文进一步描述的。

[0089] “高立体选择性”是指能够以至少约85%立体异构体过量将底物例如,化合物(2)转化为其相应手性醇产物例如化合物(1a)和(1c)的化学或酶促反应。

[0090] “增加的酶活性”或“增加的活性”是指工程化酶的改进特性,其可被表示为与参考酶相比,比活性(例如,产生的产物/时间/重量蛋白)的增加,或底物至产物的百分比转化率的增加(例如,在指定的时间段使用指定量的酮还原酶,起始量的底物至产物的百分比转化率)。确定酶活性的示例性方法被提供于实施例。可以影响与酶活性相关的任何特性,包括经典的酶特性 K_m 、 V_{max} 或 k_{cat} ,它们的改变能够导致增加的酶活性。酮还原酶活性可以通过用于测量酮还原酶的标准测定法,诸如在底物或产物浓度中的变化,或在辅因子的浓度(不存在辅因子再生系统)中的变化的任何一个来测量。酶活性的比较使用确定的酶制剂,一组条件下确定的测定,和一种或多种确定的底物来进行,如在本文中详细地进一步描述的。通常,当比较细胞溶解产物中的酶时,确定细胞的数目和测定的蛋白的量,以及使用相同表达系统和相同宿主细胞以最小化由宿主细胞产生并存在于溶解产物中的酶的量的变化。

[0091] “转化率”是指底物至相应的产物的酶促转化。“百分比转化率”是指在指定条件下一段时间内被转化为产物的底物的百分比。因此,例如,酮还原酶多肽的“酶活性”或“活性”可以表示为底物至产物的“百分比转化率”。

[0092] “热稳定的(thermostable)”或“热稳定的(thermal stable)”可互换使用,是指与未处理的酶相比当暴露于一组温度条件(例如,40-80℃)持续一段时间(例如,0.5-24小时)时耐失活,因此在暴露于高温后保留某些水平的残留活性(例如,诸如高于60%至80%)的多肽。

[0093] “溶剂稳定的”指与未处理的酶相比，多肽在暴露于不同浓度(例如5-99%)的溶剂(例如，异丙醇、四氢呋喃、2-甲基四氢呋喃、丙酮、甲苯、乙酸丁酯、甲基叔丁基醚等)持续一段时间(例如，0.5-24小时)之后维持相似的活性(例如，多于例如，60%至80%)的能力。

[0094] “氨基酸差异”或“残基差异”是指相对于参考序列中在相应位置上的氨基酸残基，在多肽序列的位置上的氨基酸残基中的差异。氨基酸差异的位置通常在本文中称为“Xn”，其中n是指残基差异所基于的参考序列中的相应位置。例如，“与SEQ ID NO:4相比在位置X40上的残基差异”是指相应于SEQ ID NO:4的位置40的多肽位置上的氨基酸残基的差异。因此，如果SEQ ID NO:4的参考多肽在位置40上具有组氨酸，那么“与SEQ ID NO:4相比在位置X40上的残基差异”指相应于SEQ ID NO:4的位置40的多肽位置上除组氨酸之外的任何残基的氨基酸置换。在本文大多数情况下，位置上的特定氨基酸残基差异表示为“XnY”，其中“Xn”指定如上所述的相应位置，且“Y”是工程化多肽中发现的氨基酸的单字母标识符(即，与参考多肽相比的不同残基)。在一些情况下(例如，在表2A、2B、和2C中)，本公开内容还提供了由常规符号“AnB”表示的特定氨基酸差异，其中A是参考序列中残基的单字母标识符，“n”是参考序列中的残基位置的数目，且B是工程化多肽的序列中的残基置换的单字母标识符。在一些情况下，本公开内容的多肽相对于参考序列可以包括一个或多个氨基酸残基差异，这由相对于参考序列存在残基差异的特定位置的列表指示。在一些实施方案中，当多于一个氨基酸可以在多肽的特定残基位置中被使用时，可被使用的各种氨基酸残基由“/”隔开(例如，X192A/G)。本公开内容包括工程化多肽序列，该工程化多肽序列包括一个或多个氨基酸差异，该氨基酸差异包括任一种和/或两种保守和非保守的氨基酸置换。包括在本公开内容的序列表中的特定重组碳酸酐酶多肽的氨基酸序列包括起始蛋氨酸(M)残基(即，M表示残基位置1)。然而，技术人员理解该起始蛋氨酸残基可通过生物学加工机制除去，诸如在宿主细胞内或体外翻译系统中，以产生缺乏起始蛋氨酸残基，但在其他方面保留了酶的性能的成熟蛋白。因此，术语“相对于SEQ ID NO:4在位置Xn的氨基酸残基差异”，如本文使用的，可指位置“Xn”或已被加工以缺乏起始蛋氨酸的参考序列中的相应位置(例如，位置(X-1)n)。

[0095] “保守氨基酸置换”指具有相似侧链的残基的可互换性，并且因此，通常涉及用在相同或相似定义类别的氨基酸内的氨基酸置换多肽中的氨基酸。通过示例的方式而非限制，具有脂族侧链的氨基酸可被另一个脂族氨基酸置换，例如，丙氨酸、缬氨酸、亮氨酸和异亮氨酸；具有羟基侧链的氨基酸被具有羟基侧链的另一个氨基酸置换，例如，丝氨酸和苏氨酸；具有芳香族侧链的氨基酸被具有芳香族侧链的另一个氨基酸置换，例如，苯丙氨酸、酪氨酸、色氨酸和组氨酸；具有碱性侧链的氨基酸被具有碱性侧链的另一个氨基酸置换，例如，赖氨酸和精氨酸；具有酸性侧链的氨基酸被具有酸性侧链的另一个氨基酸置换，例如，天冬氨酸或谷氨酸；且疏水性氨基酸或亲水性氨基酸分别被另一个疏水性氨基酸或亲水性氨基酸置换。示例性保守置换被提供在下面的表1中。

[0096] 表1

[0097]	残基	可能的保守置换
	A、L、V、I	其它脂族残基 (A、L、V、I) 其它非极性残基 (A、L、V、I、G、M)
	G、M	其它非极性残基 (A、L、V、I、G、M)
	D、E	其它酸性残基 (D、E)
	K、R	其它碱性残基 (K、R)
	N、Q、S、T	其它极性残基
	H、Y、W、F	其它芳香族残基 (H、Y、W、F)
	C、P	无

[0098] “非保守置换”是指用具有显著差异侧链性质的氨基酸置换多肽中的氨基酸。非保守置换可以利用限定组之间，而不是它们之内的氨基酸，并影响(a)置换的区域中肽骨架的结构(例如，脯氨酸置换甘氨酸)(b)电荷或疏水性，或(c)侧链的体积。通过示例的方式而非限制，示例性的非保守置换可以是酸性氨基酸被碱性或脂族氨基酸置换；芳香族氨基酸被小氨基酸置换；以及亲水性氨基酸被疏水性氨基酸置换。

[0099] “缺失”是指通过从参考多肽除去一个或多个氨基酸而对多肽进行的修饰。缺失可以包括除去1个或多个氨基酸、2个或更多个氨基酸、5个或更多个氨基酸、10个或更多个氨基酸、15个或更多个氨基酸、或20个或更多个氨基酸、多至组成多肽氨基酸总数的10%、多至氨基酸总数的20%而同时保留酶活性和/或保留工程化酶的改进特性。缺失可以涉及多肽的内部和/或端部。在各个实施方案中，缺失可以包括连续的区段或可以是不连续的。

[0100] “插入”是指通过向参考多肽添加一个或多个氨基酸而对多肽进行的修饰。在一些实施方案中，改进的工程化酮还原酶类包括一个或多个氨基酸插入天然存在的酮还原酶多肽，以及一个或多个氨基酸插入工程化酮还原酶多肽。插入可以是在多肽的内部或到羧基或氨基末端。如本文所用的插入包括如本领域已知的融合蛋白。插入可以是氨基酸的连续区段或由天然存在的多肽中的一个或多个氨基酸分隔。

[0101] 如本文所用的“片段”是指如下多肽：所述多肽具有氨基端和/或羧基端缺失，但剩余的氨基酸序列与该序列中相应位置相同。片段通常可具有全长酮还原酶多肽例如SEQ ID NO:4的多肽的约80%、90%、95%、98%和99%。

[0102] “分离的多肽”是指如下多肽：所述多肽基本上与其天然伴随的其它污染物，例如，蛋白、脂质和多核苷酸分离。术语包括已从它们天然存在环境或表达系统(例如，宿主细胞或体外合成)中除去或纯化的多肽。改进的酮还原酶类可以存在于细胞内，存在于细胞培养基中，或以各种形式制备，诸如溶解产物或分离的制剂。因此，在一些实施方案中，本公开内容的工程化酮还原酶多肽可以是分离的多肽。

[0103] “基本上纯的多肽”是指如下组合物，在所述组合物中多肽物质是存在的优势物质(即，在摩尔基础或重量基础上，它比在该组合物中的任何其它个体大分子物质更丰富)，并且当目标物质构成存在的大分子物质的按摩尔或%重量计至少约50%时，一般是基本上纯化的组合物。一般而言，基本上纯的工程化酮还原酶多肽组合物将构成该组合物中存在的所有大分子物质的按摩尔或%重量计约60%或更多、约70%或更多、约80%或更多、约90%或更多、约95%或更多以及约98%或更多。溶剂物质、小分子(<500道尔顿)、以及元素离子物质不被认为是大分子物质。在一些实施方案中，分离的改进的酮还原酶多肽是基本上纯

的多肽组合物。

[0104] “异源的”多核苷酸是指通过实验室技术引入宿主细胞中的任何多核苷酸,并且包括从宿主细胞中除去,经历实验室操作且然后再引入到宿主细胞中的多核苷酸。

[0105] “密码子优化的”是指编码蛋白的多核苷酸的密码子变为特定生物体中优先使用的那些密码子,以致所编码的蛋白在感兴趣的生物体中被有效表达。在一些实施方案中,可以对编码酮还原酶类的多核苷酸进行密码子优化,以用于从被选择用于表达的宿主生物体优化生产。

[0106] “控制序列”在本文中定义为包括对本公开内容的多核苷酸和/或多肽的表达是必要的或有利的的所有组分。每个控制序列对感兴趣的多核苷酸来说可以是天然的或外源的。此类控制序列包括但不限于,前导序列、聚腺苷酸化序列、前肽序列、启动子、信号肽序列和转录终止子。

[0107] “可操作地连接的”在本文中定义为如下一种配置:在所述配置中控制序列相对于感兴趣的多核苷酸被适当放置(即,以功能关系)的位置使得控制序列指导或调整感兴趣的多核苷酸和/或多肽的表达。

[0108] “辅因子再生系统”或“辅因子再循环系统”是指一组参与还原辅因子的氧化形式的反应(例如,NADP⁺到NADPH)的反应物。由酮基底物的酮还原酶催化的还原而氧化的辅因子由辅因子再生系统的还原形式再生。辅因子再生系统包括化学计量的还原剂,其是还原氢等同物的来源并且能够还原辅因子的氧化形式。辅因子再生系统还可包括催化剂,例如酶催化剂,其催化还原剂对辅因子的氧化形式的还原。分别从NAD⁺或NADP⁺再生NADH或NADPH的辅因子再生系统,是本领域已知的并且可在本文描述的方法中使用。

[0109] “仲醇脱氢酶”在本文中用来指NAD⁺或NADP⁺依赖性酶,其催化仲醇(例如,异丙醇)和NAD⁺或NADP⁺分别转化为酮和NADH或NADPH。

[0110] “适当的反应条件”是指在生物催化反应溶液中的那些条件(例如,酶载量、底物载量、辅因子载量、温度、pH、缓冲液、共溶剂等的范围),在上述条件下本公开内容的酮还原酶多肽能够将底物化合物转化为产物化合物(例如,将化合物(2)转化为化合物(1))。示例性的“适当的反应条件”被提供在本公开内容中,并通过实施例说明。

[0111] “载量”,诸如在“化合物载量”或“酶载量”或“辅因子载量”中,是指在反应开始时反应混合物中组分的浓度或量。

[0112] 在生物催化剂介导的方法的上下文中的“底物”是指生物催化剂所作用于的化合物或分子。例如,在本文所公开的方法中的酮还原酶生物催化剂的示例性底物是化合物(2)。

[0113] 在生物催化剂介导的方法的上下文中的“产物”是指由生物催化剂的作用产生的化合物或分子。例如,在本文所公开的方法中的酮还原酶生物催化剂的示例性产物是化合物(1)。

[0114] 本文所用的“平衡”是指导致化学物质在化学或酶促反应(例如,两种物质A和B的互变)中的稳态浓度的方法,包括立体异构体的互变,如由化学或酶促反应的正向速率常数和逆向速率常数确定的。

[0115] “烷基(alkyl)”是指1个至18个包含1个和18个碳原子,直链的或支链的,更优选地1个至8个包含1个和8个碳原子,且最优选地1个至6个包含1个和6个碳原子的饱和烃基。具

有指定数目的碳原子的烷基被表示在括号中,例如,(C₁-C₄)烷基是指1个至4个碳原子的烷基。

[0116] “烯基”是指含有至少一个双键,但任选地含有多于一个双键的2个至12个包含2个和12个碳原子的直链或支链的基团。

[0117] “炔基”是指含有至少一个三键,但任选地含有多于一个三键,并且另外任选地含有一个或多个双键部分的2个至12个包含2个和12个碳原子的直链或支链的基团。

[0118] “杂烷基”、“杂烯基”和“杂炔基”是指其中一个或多个碳原子各自独立地被相同或不同的杂原子或杂原子基团代替的如本文所定义的烷基、烯基和炔基。可代替碳原子的杂原子和/或杂原子基团包括,但不限于,-O-、-S-、-S-O-、-NR^a-、-PH-、-S(O)-、-S(O)₂-、-S(O)NR^a-、-S(O)₂NR^a-以及类似的,包括其组合,其中每一R^a独立地选自氢、烷基、杂烷基、环烷基、杂环基、芳基和杂芳基。

[0119] “烷氧基”是指基团-OR^b,其中R^b是如上面定义的烷基基团,包括如也在本文中定义的任选地取代的烷基。

[0120] “芳基”是指具有单环(例如,苯基)或多个稠环(例如,萘基或蒽基)的6个至12个包含6个和12个碳原子的不饱和芳香族碳环基团。示例性芳基包括苯基、吡啶基、萘基以及类似的。

[0121] “芳基烷基”是指被芳基取代的烷基,即芳基-烷基-基团,优选地具有在烷基部分的1个至6个包含1个和6个碳原子和在芳基部分的6个至12个包含6个和12个碳原子。此类芳基烷基基团由苯甲基、苯乙基以及类似的举例说明。

[0122] “芳基烯基”是指被芳基取代的烯基,即,芳基-烯基-基团,优选地具有在烯基部分的2个至6个包含2个和6个碳原子和在芳基部分的6个至12个包含6个和12个碳原子。

[0123] “芳基炔基”是指被芳基取代的炔基,即,芳基-炔基-基团,优选具有在炔基部分的2个至6个碳原子包含2个和6个和在芳基部分的6个至12个包含6个和12个碳原子。

[0124] “碳环”和“碳环的”在本文中可互换使用,以指3个至12个原子的环结构,其中所有环原子是碳原子。碳环包括环烷基和芳基基团。示例性的碳环包括,但不限于,环戊基、环己基、环庚基、环辛基、环戊烯基、环己烯基、环己二烯基和苯基。

[0125] “环烷基”是指具有可任选地被1个至3个烷基基团取代的单环或多个稠环的3个至12个包含3个和12个碳原子的环状烷基基团。示例性的环烷基基团包括,但不限于,单环结构诸如环丙基、环丁基、环戊基、环辛基、1-甲基环丙基、2-甲基环戊基、2-甲基环辛基以及类似的,或多环结构,包括桥环系统,诸如金刚烷基以及类似的。

[0126] “环烷基烷基”是指被环烷基取代的烷基,即,环烷基-烷基-基团,优选具有在烷基部分中的1个至6个包含1个和6个碳原子和在环烷基部分中的3个至12个包含3个和12个碳原子。此类环烷基烷基基团由环丙基甲基、环己基乙基以及类似的举例说明。

[0127] “环烷基烯基”是指被环烷基取代的烯基,即,环烷基-烯基-基团,优选具有在烯基部分中的2个至6个包含2个和6个碳原子和在环烷基部分中的3个至12个包含3个和12个碳原子。

[0128] “环烷基炔基”是指被环烷基取代的炔基,即,环烷基-炔基-基团,优选具有在炔基部分中的1个至6个包含1个和6个碳原子和在环烷基部分中的3个至12个包含3个和12个碳原子。

[0129] “氨基”是指基团 -NH_2 。取代的氨基是指基团 -NHR^δ 、 $\text{-NR}^\delta\text{R}^\delta$ 和 $\text{-NR}^\delta\text{R}^\delta\text{R}^\delta$ ，其中各 R^δ 独立地选自取代的或未被取代的烷基、环烷基、环杂烷基、烷氧基、芳基、杂芳基、杂芳基烷基、酰基、烷氧基羰基、硫酰基、亚磺酰基、磺酰基，以及类似的。典型的氨基基团包括但不限于二甲基氨基、二乙基氨基、三甲基胺、三乙基胺、甲基磺酰基氨基(methylsulfonylamino)、呋喃基-氧基-磺氨基以及类似的。

[0130] “烷基氨基”是指 -NHR^δ 基团其中 R^δ 是烷基，N-氧化物衍生物，或其保护的衍生物，例如，甲氨基、乙氨基、正丙氨基、异丙基氨基、正丁基氨基、异丁基氨基、叔丁基氨基、或甲氨基-N-氧化物以及类似的。

[0131] “氨基烷基”是指其中氢原子中的一个或多个被如本文定义的氨基(包括取代的氨基)取代的烷基基团。

[0132] “氧代”是指 $=\text{O}$ 。

[0133] “氧基”是指二价基团 -O- ，其可具有各种取代基以形成不同的氧基基团，包括醚和酯。

[0134] “羧基”是指 -COOH 。“羰基”是指 -C(=O)- ，其可具有多种取代基以形成不同的羰基基团，包括酸、酰基卤、醛、酰胺、酯和酮。

[0135] “羰基”是指 -C(=O)- ，其可具有各种取代基以形成不同的羰基，包括酯和酮。

[0136] “烷氧基羰基”是指 -C(=O)OR^e ，其中 R^e 是如本文所定义的烷基，其可任选地被取代。

[0137] “芳氧基羰基”是指 -C(=O)OR^f ，其中 R^f 是如本文所定义的芳基，其可任选地被取代。

[0138] “芳基烷氧基羰基”是指 -C(=O)OR^p ，其中 R^p 为如本文所定义的芳基-烷基-基团，其可任选地被取代。

[0139] “杂芳氧基羰基”是指 -C(=O)OR^o ，其中 R^o 为如本文所定义的杂芳基基团，其可任选地被取代。

[0140] “杂芳基烷氧基羰基”是指 -C(=O)OR^w ，其中 R^w 为如本文所定义的杂芳基烷基基团，其可任选地被取代。

[0141] “氨基羰基”是指 -C(=O)NH_2 。取代的氨基羰基是指 $\text{-C(=O)NR}^\delta\text{R}^\delta$ ，其中氨基 $\text{NR}^\delta\text{R}^\delta$ 如本文所定义。

[0142] “卤素”或“卤代”是指氟、氯、溴和碘。

[0143] “羟基”是指 -OH 。

[0144] “氰基”是指 -CN 。

[0145] “烷基硫代”是指 -S-R^n 。其中 R^n 是烷基，典型的烷基硫代包括，但不限于，甲基硫代、乙基硫代、正丙基硫代，以及类似的。

[0146] “磺酰基”是指 $\text{-SO}_2\text{-}$ 。取代的磺酰基是指 $\text{-SO}_2\text{R}^k$ ，其中 R^k 是烷基、芳基或如下文所述的其它适合的取代基。

[0147] “烷基磺酰基”是指 $\text{-SO}_2\text{R}^k$ ，其中 R^k 是如本文所定义的烷基基团。

[0148] “芳基磺酰基”是指 $\text{-SO}_2\text{R}^k$ ，其中 R^k 是如本文所定义的芳基基团。

[0149] “亚磺酰基”是指 -S(O)R^λ ，其中 R^λ 选自烷基、环烷基、芳基、杂芳基(通过环碳键合的)和杂环(通过环碳键合的)。

[0150] “烷基亚磺酰基”是指 -S(O)R^λ ，其中 R^λ 为如本文所定义的烷基基团。

[0151] “芳基亚磺酰基”是指 -S(O)R^λ ，其中 R^λ 为如本文所定义的芳基基团。

烷基、三氟甲基、卤代烷基、羟烷基、烷氧基烷基、硫代、烷基硫代、酰基、羧基、烷氧基羰基、甲酰氨基、取代的甲酰氨基、烷基磺酰基、烷基亚磺酰基、烷基磺酰基氨基、磺酰氨基、取代的亚磺酰氨基、氰基、氨基、取代的氨基、烷基氨基、二烷基氨基、氨基烷基、酰氨基、脒基、氨基胍基(amidoximo)、羟基草氨酰基(hydroxamoyl)、苯基、芳基、取代的芳基、芳氧基、芳基烷基、芳基烯基、芳基炔基、吡啶基、咪唑基、杂芳基、取代的杂芳基、杂芳氧基、杂芳基烷基、杂芳基烯基、杂芳基炔基、环丙基、环丁基、环戊基、环己基、环烷基、环烯基、环烷基烷基、取代的环烷基、环烷氧基、吡咯烷基、哌啶基、吗啉代、杂环、(杂环)氧基和(杂环)烷基；且优选的杂原子是氧、氮和硫。应理解，当这些取代基上存在开放化合价时，它们可被烷基、环烷基、芳基、杂芳基和/或杂环基团进一步取代，当碳上存在这些开放化合价时，它们可被卤素和被氧键合取代基、氮键合取代基或硫键合取代基进一步取代，并且当多个这样的开放化合价存在时，这些基团可通过直接形成键或通过新的杂原子(优选氧、氮或硫)形成键而被结合以形成环。还应理解，可以进行上述取代，只要用取代基取代氢没有向本发明的分子引入不可接受的不稳定性，并且在其他方面是化学上合理的。

[0164] “任选的”或“任选地”是指随后描述的事件或情况可能发生或可能不发生，并且该描述包括其中所述事件或情况发生的实例和其中所述事件或情况不发生的实例。本领域普通技术人员将理解，对于描述为含有一个或多个任选的取代基的任何分子，仅意在包括立体上实际的和/或合成上可行的化合物。“任选地取代的”是指在化学基团的术语或系列中的所有其后的修饰词。例如，在术语“任选地取代的芳基烷基”中，分子的“烷基”部分和“芳基”部分可以被取代或可以不被取代，并且对于系列“任选地取代的烷基、环烷基、芳基和杂芳基，”该烷基、环烷基、芳基和杂芳基基团，独立于其它，可以被取代或可以不被取代。

[0165] “保护基团”是指当连接到分子中的反应性官能团时掩蔽、减少或阻止该官能团的反应性的原子团。通常，保护基团可在合成过程中根据需要被选择性地除去。保护基团的实例可见于Wuts和Greene, “Greene’s Protective Groups in Organic Synthesis,” 第4版, Wiley Interscience(2006), 以及Harrison等人, Compendium of Synthetic Organic Methods, 第1-8卷, 1971-1996, John Wiley & Sons, NY。可以具有保护基团的官能团包括但不限于, 羟基、氨基和羧基。代表性的氨基保护基团包括但不限于, 甲酰基、乙酰基、三氟乙酰基、苄基、苄氧羰基(“CBZ”)、叔丁氧羰基(“Boc”)、三甲基甲硅烷基(“TMS”)、2-三甲基甲硅烷基-乙烷磺酰基(“SES”)、三苯甲基和取代的三苯甲基、烯丙氧基羰基、9-苄基甲氧基羰基(“FMOC”)、硝基-芴氧基羰基(“NVOC”)以及类似的。代表性的羟基保护基团包括但不限于, 其中羟基被酰化的那些(例如, 甲基和乙基酯、乙酸酯或丙酸酯基团或乙二醇酯)或其中羟基被烷基化的那些诸如苄基和三苯甲基醚, 以及烷基醚、四氢吡喃醚、三烷基甲硅烷基醚(例如, TMS或TIPPS基团)和烯丙基醚。其它保护基团可以见于本文引用的参考文献。

[0166] 工程化酮还原酶多肽

[0167] 酮还原酶(KRED)或羰基还原酶生物催化剂(EC1.1.1.184)可用于由醛和酮合成醇, 并且由相应的前立体异构的酮底物合成光学活性仲醇。KRED还可催化逆反应, 即, 将醇底物氧化为相应醛/酮产物。通过KRED的醛和酮的还原以及醇的氧化使用辅因子, 最常见的是用于氧化反应的还原型烟酰胺腺嘌呤二核苷酸(NADH)或还原型烟酰胺腺嘌呤二核苷酸磷酸(NADPH)和烟酰胺腺嘌呤二核苷酸(NAD)或烟酰胺腺嘌呤二核苷酸磷酸(NADP+)。NADH和NADPH作为电子供体, 而NAD⁺和NADP⁺作为电子受体。

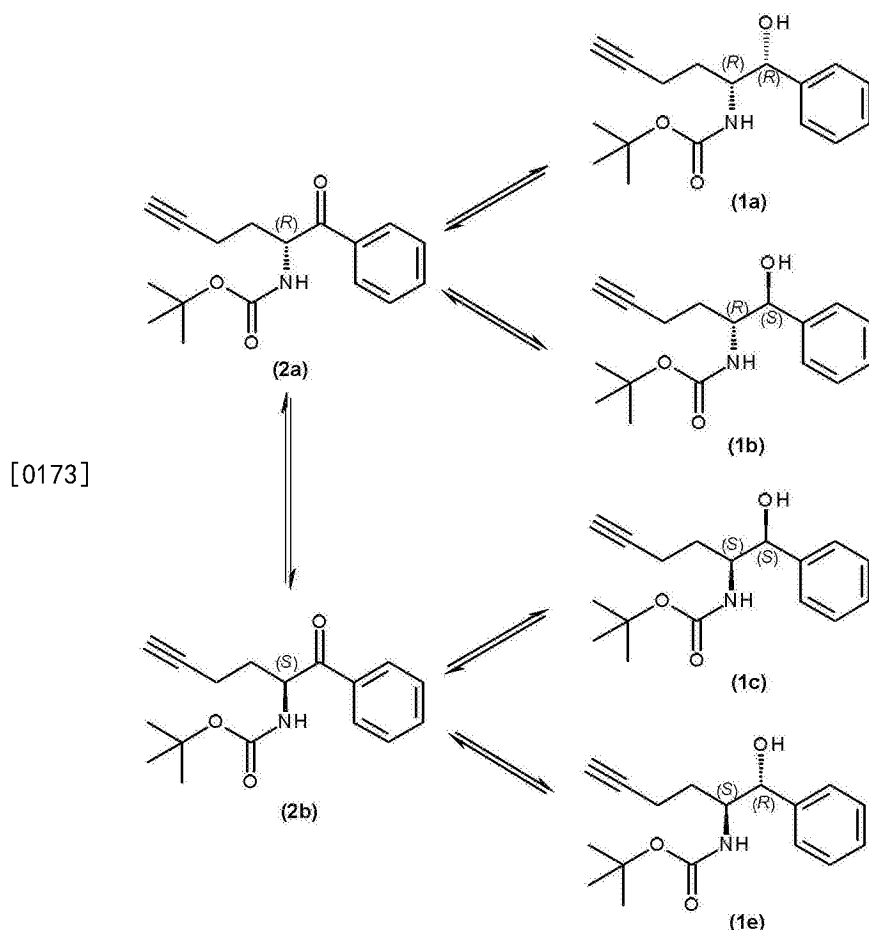
[0168] KRED可发现于各种各样的细菌和酵母中(关于综述参见,例如,Kraus和Waldman, *Enzyme catalysis in organic synthesis* 卷1&2.VCH Weinheim 1995;Faber,K., *Biotransformations in organic chemistry*,第4版,Springer,Berlin Heidelberg New York.2000;Hummel和Kula *Eur.J.Biochem.*1989184:1-13)。已报告了很多KRED基因和酶序列,包括:木兰假丝酵母(*Candida magnoliae*)(Genbank登录号JC7338;GI:11360538);近平滑假丝酵母(*Candida parapsilosis*)(Genbank登录号BAA24528.1;GI:2815409)、赭色掷孢酵母(*Sporobolomyces salmonicolor*)(Genbank登录号AF160799;GI:6539734);克菲尔乳杆菌(Genbank登录号AAP94029.1;GI:33112056);短乳杆菌(*Lactobacillus brevis*)(Genbank登录号1NXQ_A;GI:30749782);和布氏热厌氧菌(*Thermoanaerobium brockii*)(Genbank登录号P14941;GI:1771790)。

[0169] 酮还原酶的立体选择性已应用于制备重要的药物构建块(pharmaceutical building block)(参见例如Broussy等人,2009,Org Lett.11(2):305-308)。天然存在的或工程化的KRED在产生有用化合物的生物催化方法中的特定应用已针对以下被证明:4-氯乙酰乙酸酯的还原(Zhou,J.Am.Chem.Soc.1983105:5925-5926;Santaniello,J.Chem.Res.(S)1984:132-133;美国专利号5,559,030;美国专利号5,700,670和美国专利号5,891,685)、二氧代羧酸的还原(例如,美国专利号6,399,339)、(S)-氯-5-羟基-3-氧代己酸叔丁酯的还原(例如,美国专利号6,645,746和WO 01/40450)、基于吡咯并三嗪的化合物的还原(例如,美国申请号2006/0286646);取代的苯乙酮的还原(例如,美国专利号6,800,477);以及酮基硫代坊的还原(WO 2005/054491)。

[0170] 如本文中讨论的,(1-羟基-1-苯基己-5-炔-2-基)氨基甲酸叔丁酯为在基于羟甲基吡咯烷的 β_3 -肾上腺素能受体激动剂的合成中的中间体(例如,W02009124167和W02009124166,通过引用并入本文)。天然存在的和先前工程化的KRED不能有效地将底物化合物(2),(1-氧代-1-苯基己-5-炔-2-基)氨基甲酸叔丁酯,转化为产物化合物(1),(1-羟基-1-苯基己-5-炔-2-基)氨基甲酸叔丁酯,如在上述方案1所示。然而,本公开内容描述了在进行方案1的转化中具有改进的特性的非天然存在的(即,工程化的)KRED的开发。这些改进的特性包括,增加的活性,对某些非对映异构体的增加的选择性,增加的稳定性(例如,热和溶剂),高百分比底物转化率(特别是,在高底物载量的存在下),增强的多肽表达,及其各种改进组合以及其他。

[0171] 对于本文描述的目的,对本领域技术人员来说化合物(2)具有两个手性中心,并且可以至少两个不同的非对映异构形式(例如,化合物(2a)和(2b))存在将是明显的。因此,通过酮还原酶的还原可导致产物的至少四种不同的立体异构形式(例如,化合物(1a)、(1b)、(1c)和(1d)),如方案3中所示。

[0172] 方案3



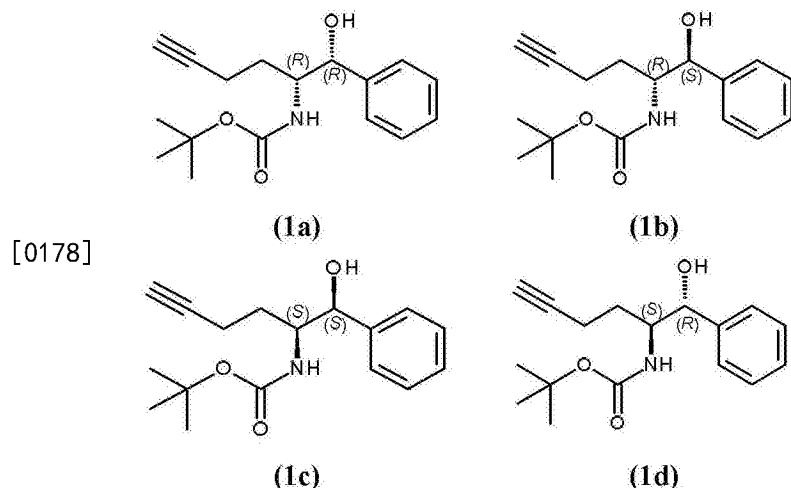
[0174] 因此,如本文所用,提及化合物(2)或其结构类似物而没有提及任何特定的立体异构结构,是指为本文公开的工程化酮还原酶的底物的化合物(例如,化合物(2a)和(2b))的立体异构体形式的任何混合物(包括外消旋混合物)或纯制剂。类似地,提及任何指示的产物化合物(1)或其结构类似物而没有提及特定立体异构结构是指在酮还原酶反应中形成的产物化合物(1)(例如,化合物(1a)、(1b)、(1c)、和(1d))的立体异构形式的任何混合物。此外,本公开内容的工程化酮还原酶还可进行逆反应,以将化合物(1)转化为化合物(2)的相应的酮,并且促进两种底物化合物(2a)和(2b)之间的平衡,如方案2中所示。

[0175] 本公开内容的非天然存在的多肽是与SEQ ID NO:2的天然存在的酮还原酶或SEQ ID NO:4的工程化酮还原酶相比具有改进的特性的工程化酮还原酶。工程化酮还原酶多肽适于将化合物(2)有效转化为化合物(1)且与SEQ ID NO:2的天然存在的酮还原酶或SEQ ID NO:4的参考工程化酮还原酶多肽相比具有一个或多个残基差异,其相对于来自克菲尔乳杆菌的SEQ ID NO:2的天然存在的酮还原酶具有以下11个氨基酸差异:A94G、S96V、E145F、F147M、L153T、Y190P、L195M、V196L、L199Y、I226V、和Y249W。这些残基差异与酶性能中的改进,特别是增加的活性、增加的立体选择性、增加的稳定性、以及对底物和/或产物浓度的增加的耐受性(例如,降低的产物抑制)相关。

[0176] 在一些实施方案中,工程化酮还原酶多肽能够在适当的反应条件下将底物化合物(2)转化为化合物(1),具有相对于SEQ ID NO:4的参考多肽的活性增加至少约1.2倍、1.5倍、2倍、3倍、4倍、5倍、10倍、20倍、30倍、40倍、50倍、100倍、200倍、500倍、1000倍、5000倍、10000倍或更多倍的活性。在一些实施方案中,工程化酮还原酶多肽能够在适当的反应条件

下在约48小时、约36小时、约24小时或甚至更短时间长度的反应时间内将化合物(2)的底物转化为化合物(1),具有至少约40%、至少约50%、至少约60%、至少约70%、至少约80%、或至少约90%、至少约95%、至少约98%、至少约99%的百分比转化率。

[0177] 在一些实施方案中,本文所描述的工程化酮还原酶多肽在适当的反应条件下相比于对反式化合物(1b)和(1d),展现对顺式化合物(1a)和(1c)的非对映选择性。



[0179] 在一些实施方案中,工程化酮还原酶多肽能够以相比于反式化合物(1b)和(1d)(即,[1b+1d])的大于10:1、20:1、30:1、40:1、50:1、60:1、70:1、80:1、90:1、或100:1或更大的非对映体比将化合物(2)转化为顺式化合物(1a)和(1c)(即,[1a+1c])。在一些实施方案中,工程化酮还原酶多肽能够在适当的反应条件下以相比于化合物(1b)和(1d)(即,[1b+1d])的大于50:1的非对映体比将底物化合物(2)转化为产物化合物(1a)和(1c)(即,[1a+1c])。

[0180] 在一些实施方案中,工程化酮还原酶能够以相比于化合物(1c)的非对映体过量(in diastereomeric excess over compound(1c))将底物化合物(2)转化为产物化合物(1a)。在一些实施方案中,工程化酮还原酶能够在适当的反应条件下以相比于化合物(1c)的至少10%、20%、30%、40%、50%、60%、70%、80%、90%或更大的非对映体过量将化合物(2)转化为化合物(1a)。

[0181] 在一些实施方案中,工程化酮还原酶多肽相比于对化合物(1a)和(1c),展现对化合物(1b)和(1d)的非对映选择性。因此,在一些实施方案中,工程化酮还原酶能够在适当的反应条件下将化合物(2)转化为非对映体过量于化合物(1a)和(1c)的化合物(1b)和(1d)。

[0182] 如下面进一步讨论的,工程化酮还原酶的非对映选择性和在底物化合物之间的平衡可用于动态动力学拆分方法以制备过量的某些非对映体,并且在一些情况下,制备非对映体或非对映体的基本上纯的制剂。

[0183] 在一些实施方案中,工程化酮还原酶多肽能够在适当的反应条件下以相对于SEQ ID NO:4的参考多肽对底物存在的增加的耐受性将化合物(2)转化为化合物(1)。因此,在一些实施方案中,工程化酮还原酶多肽能够在适当的反应条件下在约72h、约48h、约36h、约24h或甚至更短的时间长度的反应时间内以至少约40%、至少约50%、至少约60%、至少约70%、至少约80%、至少约90%、至少约95%、至少约98%、或至少约99%的百分比转化率在以下的底物载量浓度的存在下将化合物(2)的底物转化为化合物(1):至少约1g/L、5g/L、

10g/L、20g/L、约30g/L、约40g/L、约50g/L、约70g/L、约100g/L、约125g/L、约150g/L、约175g/L或约200g/L或更多。

[0184] 工程化多肽的上面描述的改进特性在其下进行转化的适当的反应条件可关于多肽、底物、辅因子、缓冲液、共溶剂的浓度或量、pH和/或包括温度及反应时间的条件来确定，如下面和实施例中进一步描述的。

[0185] 用于将化合物(2)转化为化合物(1)的与其改进特性相关的示例性工程化多肽包括与SEQ ID NO:4相比在以下残基位置上的一个或多个残基差异:X7;X17;X23;X27;X29;X40;X60;X64;X71;X87;X94;X95;X96;X105;X113;X122;X127;X131;X144;X145;X147;X150;X152;X153;X157;X173;X195;X196;X198;X199;X206;X208;X216;X221;X243;X245;和X249。在与表2A、2B和2C的示例性多肽的改进特性相关的这些位置的每一个上的特定氨基酸差异包括:X7S;X17M;X17Q;X17R;X23V;X27L;X29G;X40R;X60I;X64V;X71P;X87L;X94A;X94P;X94S;X95M;X96L;X96Y;X105G;X113I;X122A;X127R;X131S;X144V;X145L;X147I;X147L;X147Q;X150Y;X152G;X153G;X157C;X173L;X195A;X195G;X196M;X198S;X199H;X206F;X206L;X208R;X216R;X221S;X243S;X245I;X249F;X249G;和X249Y。

[0186] 本公开内容的示例性工程化酮还原酶多肽的结构与功能信息显示于下面表2A、2B和2C中。奇数序列标识符(即,“SEQ ID NO:”)是指编码由偶数的SEQ ID NO提供的氨基酸序列的核苷酸序列,并且该序列被提供在伴随该公开内容的电子序列文件,该序列文件在此通过引用并入本文。氨基酸残基差异是基于与SEQ ID NO:4的参考多肽序列比较,SEQ ID NO:4的参考多肽序列为相对于克菲尔乳杆菌的天然存在的酮还原酶(SEQ ID NO:2)具有以下11个氨基酸差异的工程化酮还原酶多肽:A94G、S96V、E145F、F147M、L153T、Y190P、L195M、V196L、L199Y、I226V、和Y249W。“顺式:反式非对映体比”(本文中还为“d.r.”)为两种可能的顺式非对映体产物化合物(1a)和化合物(1c)与两种可能的反式非对映体产物化合物(1b)和化合物(1d)的比。非对映体比可由式 $[1a+1c]/[1b+1d]$ 来计算。

[0187] 在本文酮还原酶的筛选中,发现SEQ ID NO:4的工程化多肽以>100的d.r.(在反应条件E下)将化合物(2)转化为化合物(1a)和(1c),但具有比选择的在将底物化合物(2)转化为产物化合物(1a)和(1c)中具有改进特性的工程化酮还原酶更低的活性。因此,SEQ ID NO:4的工程化多肽被用作在将底物化合物(2)转化为产物化合物(1a)和(1c)时具有增加的活性以及具有较高的d.r.的工程化多肽的进一步进化的起点。各工程化多肽的活性使用高通量(HTP)测定(作为初级筛选)来确定。表2A中的HTP测定值,使用大肠杆菌澄清细胞溶解产物以~200μL体积/孔的96孔板格式,根据如表中标注的测定反应条件来确定。在某些情况下,使用次级摇瓶粉(SFP)和/或下游处理(DSP)粉测定来评估工程化酮还原酶的特性。SFP和DSP形式提供工程化多肽的更纯的粉制剂。例如,SFP制剂中的工程化酮还原酶是总蛋白的约30%。表2B中的SFP测定值使用在2mL小瓶格式中的工程化多肽的SFP使用表中标注的反应条件来确定。表2C中的DSP测定值使用在2mL或5mL小瓶格式中的工程化多肽的DSP粉使用表中标注的反应条件来确定。HTP、SFP和DSP制备和测定的更多细节被描述于实施例中。

[0188] 表2A

[0189]

SEQ ID NO: (nt/aa)	氨基酸差异 (相对于 SEQ ID NO: 4)	活性 (相对于 SEQ ID NO: 4)	测定 ¹
-----------------------	-----------------------------	-----------------------------	-----------------

3/4	n/a	1.0	i
5/6	V96L	4.4	i
7/8	M195A	3.5	i
9/10	M206F	4.9	i
17/18	H40R; G94P; M206F	112.7	ii
19/20	H40R; V95M; M206F	81.3	ii
21/22	H40R; I144V; M206F	77.5	ii
23/24	H40R; L196M; M206F	77.0	ii
25/26	H40R; D198S; M206F;	323.8	ii
27/28	H40R; A64V; V87L; M147Q; Y199H; M206F;	137.8	ii
29/30	H40R; A64V; M147Q; Y199H; M206F	121.4	ii
33/34	L17Q; H40R; A64V; T71P; G94P; V95M; T122A; I144V; M147I; N157C; L196M; D198S; M206F; W249Y	25190.1	iii
35/36	L17Q; H40R; A64V; T71P; G94P; I144V; M147I; N157C; L196M; M206F; W249F	23882.6	iii
37/38	L17M; H40R; A64V; V87L; G94P; I144V; N157C; L196M; D198S; Y199H; M206F; W249F	22691.3	iii
39/40	L17R; H40R; A64V; V87L; G94P; I144V; M147I; D150Y; N157C; L196M; D198S; M206F	10801.3	iii
41/42	L17Q; I23V; H40R; A64V; T71P; V87L; G94P; V95M; I144V; M147I; T153G; N157C; L196M; D198S; M206F; W249F	29528.6	iii
43/44	L17R; H40R; A64V; T71P; G94P; V95M; I144V; D150Y; L196M; D198S; M206F	11467.5	iii
45/46	V96Y	1.5	i
47/48	F145L	2.1	i
49/50	M195G	13.8 (相对 d.r.)	i
51/52	M206L	2.3	i
53/54	Y199H	2.4	i
55/56	L17Q; H40R; M206F	346.6	ii
57/58	L17R; H40R; M206F	248.1	ii
59/60	E29G; H40R; T71P; M206F	49.1	ii
61/62	H40R; G94A; M206F	68.9	ii
63/64	H40R; G94S; M206F	107.6	ii
65/66	H40R; N131S; M206F	35.5	ii
67/68	H40R; D150Y; M206F	48.7	ii
69/70	H40R; T152G; M206F	55.7	ii
71/72	H40R; N157C; M206F	313.9	ii
73/74	H40R; M195G; M206F	500.9 (相对 d.r.)	ii
75/76	H40R; V113I; M206F	19.0	ii
77/78	H40R; L196M; M206F	74.2	ii
79/80	H40R; M206F; W249G	27.5	ii
81/82	L17M; H40R; A64V; V87L; V96Y; M147L; Y199H; M206F	154.8	ii
83/84	L17M; H40R; V87L; M147L; D173L; Y199H; M206F	202.6	ii
85/86	L17M; H40R; A64V; V87L; M147Q; D173L; Y199H; M206F	587.4	ii
87/88	L17M; H40R; V87L; V96L; Y199H; M206F	147.9	ii
89/90	L17M; H40R; V96L; M147Q; M206F	90.5	ii
91/92	L17M; H40R; A64V; V87L; V96Y; M147Q; Y199H; M206F	71.4	ii
93/94	L17M; H40R; V87L; V96L; M147Q; Y199H; M206F	100.3	ii

[0190]

[0191]

95/96	H40R; V96L; M147Q; D173L; M206F	38.6	ii
97/98	H40R; Y199H; M206F	63.6	ii
99/100	H40R; V87L; M147Q; D173L; M206F	27.3	ii
101/102	H40R; M147Q; M206F	31.7	ii
103/104	L17M; H40R; V87L; V96Y; D173L; M206F	40.8	ii
105/106	L17M; H40R; A64V; Y199H; M206F; N221S	351.6	ii
107/108	L17M; H40R; M206F	101.0	ii
109/110	L17M; H40R; A64V; V96Y; M147I; D173L; Y199H; M206F	241.2	ii
111/112	L17M; H40R; V96L; M147Q; M206F; W249Y	213.1	ii
113/114	L17M; H40R; A64V; M147Q; D173L; Y199H; M206F	242.0	ii
115/116	L17M; H40R; A64V; V87L; V96Y; M147I; D173L; M206F	257.8	ii
117/118	H40R; A64V; V87L; V96Y; M147Q; Y199H; M206F; W249Y	91.7	ii
119/120	L17M; H40R; A64V; M147Q; Y199H; M206F; W249F	612.5	ii
121/122	L17M; H40R; A64V; M147Q; D173L; M206F	211.7	ii
123/124	H40R; A64V; M147Q; D173L; Y199H; M206F	89.7	ii
125/126	L17M; H40R; A64V; M147L; M206F	267.4	ii
127/128	L17M; H40R; V87L; M147Q; Y199H; M206F	170.2	ii
129/130	L17M; H40R; A64V; V96L; M147I; Y199H; M206F;	444.1	ii
131/132	G7S; H40R; V60I; A64V; Q127R; M206F; H216R; V245I	185.7	ii
133/134	L17R; H40R; A64V; T71P; G94P; I144V; M147I; N157C; L196M; D198S; M206F; F243S	8985.6	iii
135/136	L17R; H40R; A64V; T71P; V87L; G94P; I144V; M147I; D150Y; N157C; L196M; D198S; M206F	11808.6	iii
137/138	L17Q; H40R; A64V; V87L; G94P; V95M; I144V; M147I; D150Y; N157C; L196M; D198S; M206F	11931.9	iii
139/140	L17R; H40R; A64V; V87L; G94P; I144V; M147I; D150Y; N157C; L196M; D198S; M206F; Q208R	11538.1	iii
141/142	L17Q; H40R; G94P; I144V; M147I; N157C; L196M; D198S; M206F; W249Y	12100.3	iii
143/144	L17M; H40R; A64V; T71P; G94P; I144V; M147I; N157C; L196M; D198S; M206F; W249Y	11180.5	iii
145/146	L17Q; H40R; A64V; T71P; G94P; V95M; I144V; M147I; N157C; L196M; M206F	5635.6	iii
147/148	L17R; H40R; A64V; G94P; I144V; M147Q; D150Y; N157C; L196M; M206F	9200.2	iii
149/150	L17M; H40R; A64V; T71P; G94P; I144V; M147I; N157C; L196M; D198S; M206F	8707.5	iii
151/152	L17R; H40R; T71P; V87L; G94P; I144V; M147I; D150Y; N157C; L196M; M206F	7487.5	iii
153/154	L17M; H40R; A64V; T71P; G94P; E105G; I144V; M147I; D150Y; N157C; L196M; Y199H; M206F	6634.1	iii
155/156	L17R; H40R; G94P; I144V; M147I; N157C; L196M; D198S; Y199H; M206F	5827.1	iii
157/158	L17R; H40R; A64V; G94P; V95M; I144V; D150Y; L196M; M206F; W249Y	10168.4	iii
159/160	L17Q; H40R; G94P; V95M; I144V; M147Q; L196M; D198S; Y199H; M206F; W249F	5869.0	iii
161/162	L17Q; H40R; A64V; V87L; G94P; I144V; M147I; N157C; L196M; D198S; Y199H; M206F	6201.2	iii

163/164	L17Q; F27L; H40R; A64V; T71P; G94P; V95M; I144V; M147I; L196M; D198S; M206F; W249Y	5838.0	iii
165/166	L17Q; H40R; A64V; T71P; G94P; I144V; N157C; L196M; M206F; W249Y	5883.6	iii
167/168	L17R; H40R; T71P; V87L; G94P; I144V; M147I; N157C; L196M; M206F; W249F	6007.8	iii
169/170	L17Q; H40R; A64V; T71P; G94P; I144V; M147I; L196M; D198S; M206F	5487.0	iii
171/172	L17R; H40R; A64V; G94P; V95M; I144V; N157C; L196M; Y199H; M206F; W249F	8629.4	iii

¹HTP 测定条件:

[0192]

- i: 通过在 pH 10.5 下在 300 μ L 的含有 0.2 M 硼酸盐、0.5 g/L 溶菌酶以及 0.4 g/L 硫酸多粘菌素 B 的溶解缓冲液中在 250 rpm 以及室温摇动持续 1 h 来溶解细胞。反应条件包括: 20 g/L 的底物化合物 (2)、60 μ L 来自溶解在 300 μ L 溶解缓冲液中的含有酮还原酶多肽的细胞的澄清的细胞溶解产物、30% IPA (v/v)、0.5 g/L NADP、0.2 M 硼酸盐、pH 9.5 且 35 $^{\circ}$ C 下温育 24 h。
- ii: 通过在 pH 10.5 下在 300 μ L 的含有 0.2 M 硼酸盐、0.5 g/L 溶菌酶以及 0.4 g/L 硫酸多粘菌素 B 的溶解缓冲液中在 250 rpm 以及室温摇动持续 1 h 来溶解细胞。反应条件包括: 50 g/L 的底物化合物 (2)、20 μ L 来自溶解在 300 μ L 溶解缓冲液中的含有酮还原酶多肽的细胞的澄清的细胞溶解产物、50% IPA (v/v)、0.1 g/L NADP、0.2 M 硼酸盐、pH 10 且 45 $^{\circ}$ C 下温育 24 h。
- iii: 通过在 pH 10.5 下在 300 μ L 的含有 0.2 M 硼酸盐、0.5 g/L 溶菌酶以及 0.4 g/L 硫酸多粘菌素 B 的溶解缓冲液中在 250 rpm 以及室温摇动持续 1 h 来溶解细胞。反应条件包括: 50 g/L 的底物化合物 (2)、40 μ L 来自溶解在 300 μ L 溶解缓冲液中的含有酮还原酶多肽的细胞的澄清的细胞溶解产物、50% IPA (v/v)、0.1 g/L NADP、0.2 M 硼酸盐、pH 10 且 55 $^{\circ}$ C 下温育 24 h。

[0193] 表2B

[0194]

SEQ ID NO: (nt/aa)	氨基酸差异 (相对于 SEQ ID NO: 4)	活性 (相对于 SEQ ID NO: 4)	d.r.	测定 ²
3/4	n/a	1.0	25	A
5/6	V96L	6.5	15	A
7/8	M195A	4.5	11	A
9/10	M206F	7.0	39	A
11/12	H40R; V96L	33.0	9	B
13/14	H40R; M206F	37.4	15	B
15/16	H40R; M195A	16.0	7	B
13/14	H40R; M206F	3.5	>100	C
17/18	H40R; G94P; M206F	28.3	>100	C
19/20	H40R; V95M; M206F	18.2	83	C
21/22	H40R; I144V; M206F	18.1	55	C
23/24	H40R; L196M; M206F	36.9	>100	C
25/26	H40R; D198S; M206F	51.9	>100	C
27/28	H40R; A64V; V87L; M147Q; Y199H; M206F	27.2	51	C
29/30	H40R; A64V; M147Q; Y199H; M206F	20.0	34	C
31/32	H40R; G94P; I144V; L196M; M206F	45 $^{\circ}$ C 测定时无活性	n/a	D
33/34	L17Q; H40R; A64V; T71P; G94P; V95M; T122A; I144V; M147I; N157C; L196M; D198S; M206F; W249Y	345.6	>100	D
35/36	L17Q; H40R; A64V; T71P; G94P; 144V; M147I;	341.7	>100	D

[0195]

	N157C; L196M; M206F; W249F			
37/38	L17M; H40R; A64V; V87L; G94P; I144V; N157C; L196M; D198S; Y199H; M206F; W249F	343.1	78	D
39/40	L17R; H40R; A64V; V87L; G94P; I144V; M147I; D150Y; N157C; L196M; D198S; M206F	234.4	>100	D
41/42	L17Q; I23V; H40R; A64V; T71P; V87L; G94P; V95M; I144V; M147I; T153G; N157C; L196M; D198S; M206F; W249F	341.3	83	D
43/44	L17R; H40R; A64V; T71P; G94P; V95M; I144V; D150Y; L196M; D198S; M206F	203.4	>100	D
² SFP 测定条件: A: 在 2 ml 小瓶中: 5 g/L 底物化合物(2)、5 g/L 酮还原酶多肽的 SFP、30% IPA (v/v)、0.5 g/L NADP、0.2 M 磷酸盐、pH 9.5 (缓冲液)且 35℃ 下温育 24 h。 B: 在 2 ml 小瓶中: 5 g/L 底物化合物(2)、5 g/L 酮还原酶多肽的 SFP、30% IPA (v/v)、0.5 g/L NADP、0.2 M 磷酸盐、pH 9.5 (缓冲液)且 35℃ 下温育 6 h。 C: 在 2 ml 小瓶中: 40 g/L 底物化合物(2)、5 g/L 酮还原酶多肽的 SFP、50% IPA (v/v)、0.1 g/L NADP、0.2 M 磷酸盐、pH 10.2 (反应)且 35℃ 下温育 24 h。 D: 在 2 ml 小瓶中: 40 g/L 底物化合物(2)、2 g/L 酮还原酶多肽的 SFP、50% IPA (v/v)、0.1 g/L NADP、0.2 M 磷酸盐、pH 10 (缓冲液)且 45℃ 下温育 22 h。				

[0196] 表2C

[0197]

SEQ ID NO: (nt/aa)	氨基酸差异 (相对于 SEQ ID NO: 4)	活性 (相对于 SEQ ID NO: 4)	d.r.	测定 ³
3/4	n/a	1.0	>100	E
9/10	M206F	6.0	>100	E
13/14	H40R; M206F	29.7	30	E
13/14	H40R; M206F	80.7	14	F
17/18	H40R; G94P; M206F	73.5	>100	F
21/22	H40R; I144V; M206F	253.1	42	F
23/24	H40R; L196M; M206F	264.0	>100	F
25/26	H40R; D198S; M206F	74.9	>100	F
31/32	H40R; G94P; I144V; L196M; M206F	470.5 45℃ 测定时无活性	51 n/a	F G
33/34	L17Q; H40R; A64V; T71P; G94P; V95M; T122A; I144V; M147I; N157C; L196M; D198S; M206F; W249Y	1629.1	>100	G
35/36	L17Q; H40R; A64V; T71P; G94P; I144V; M147I; N157C; L196M; M206F; W249F	1620.0	>100	G
37/38	L17M; H40R; A64V; V87L; G94P; I144V; N157C; L196M; D198S; Y199H; M206F; W249F	1660.0	87	G
41/42	L17Q; I23V; H40R; A64V; T71P; V87L; G94P; V95M; I144V; M147I; T153G; N157C; L196M; D198S; M206F; W249F	1669.1	88	G

[0198]

³DSP 测定条件:**E:** 100 g/L 底物化合物(2)、5 g/L 酮还原酶多肽的 DSP 粉、30% IPA (v/v)、0.1 g/L NADP、0.2 M 磷酸盐、pH 9.5 (缓冲液)、且 35℃ 下温育 24 h。**F:** 40 g/L 底物化合物(2)、1 g/L 酮还原酶多肽的 DSP 粉、50% IPA (v/v)、0.1 g/L NADP、0.2 M 磷酸盐、pH 10.3 (反应)、且 35℃ 下温育 22 h。**G:** 100 g/L 底物化合物(2)、1 g/L 酮还原酶多肽的 DSP 粉、50% IPA (v/v)、0.1 g/L NADP、0.2 M 磷酸盐、pH 10 (缓冲液)、且 45℃ 下温育 22 h。

[0199] 根据示例性多肽的特性,酶特性中的改进(例如,将化合物(2)转化为化合物(1a)和(1c)中的活性)与相比于SEQ ID NO:4在以下残基位置上的残基差异相关:X7;X17;X23;X27;X29;X40;X60;X64;X71;X87;X94;X95;X96;X105;X113;X122;X127;X131;X144;X145;X147;X150;X152;X153;X157;X173;X195;X196;X198;X199;X206;X208;X216;X221;X243;X245;和X249。在与改进特性相关的这些位置的每一个上的特定残基差异包括:X7S;X17M;X17Q;X17R;X23V;X27L;X29G;X40R;X60I;X64V;X71P;X87L;X94A;X94P;X94S;X95M;X96L;X96Y;X105G;X113I;X122A;X127R;X131S;X144V;X145L;X147I;X147L;X147Q;X150Y;X152G;X153G;X157C;X173L;X195A;X195G;X196M;X198S;X199H;X206F;X206L;X208R;X216R;X221S;X243S;X245I;X249F;X249G;和X249Y。在酶活性中的增加与以下残基位置上的残基差异相关:X60;X71;X94;X95;X96;X127;X144;X152;X196;X199;X206;X216;和X245。部分地影响酶活性和非对映选择性,特别是相比于化合物(1b)和(1d)形成产物化合物(1a)和(1c)的底物结合功能与在以下残基位置上的残基差异相关:X40;X94;X95;X96;X144;X145;X150;X152;X153;X157;X195;X196;X198;X199;X206;和X249。辅因子NADP结合与残基位置X40上的残基差异相关,并影响酶活性。化合物(1a)和(1c)相比于化合物(1b)和(1d)的非对映体比的增加与以下残基位置上的残基差异相关:X17;X64;X71;X87;X147;X157;X196;X206和X249。非对映体比的少许减少与残基位置X144上的残基差异相关,但d.r.的减少由显著增加的酶活性来补偿。以相比于化合物(1a)和(1c)的非对映体过量形成化合物(1b)和(1d)的非对映选择性中的变化与X195G相关。酶稳定性的增加,特别是在增加的温度下(例如35℃下的活性对45℃下的活性),与以下残基位置上的残基差异以及其他相关:X17;X64;X71;X87;X94;X147;X147;X147;X157;X198;和X249。如本领域技术人员将理解的,前述残基位置和每个残基位置的特定氨基酸残基,可以单独或以各种组合用于合成具有所需改进特性的工程化酮还原酶多肽,所需改进特性包括酶活性、立体选择性和稳定性以及其他。

[0200] 根据本文提供的指导,可进一步设想,以下的示例性工程化多肽的任何一种:SEQ ID NO:4、6、8、10、12、14、16、18、20、22、24、28、30、32、34、36、38、40、42、44、46、48、50、52、54、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、96、98、102、104、106、108、110、114、116、122、124、126、128、130、132、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、168、170和172,并且特别是以下的工程化多肽:SEQ ID NO:18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172可用作用于合成其他工程化酮还原酶多肽的起始氨基酸序列,例如,通过添加来自表2A、2B和2C中的其他多肽以及本文描述的其他残基位置的各种氨基

酸差异的新组合通过随后几轮进化。进一步的改进可以通过包括在贯穿前几轮的进化中已被保持为不变的残基位置上的氨基酸差异而产生。

[0201] 因此,在一些实施方案中,能够以多肽SEQ ID NO:4的活性的至少1.5倍将底物化合物(2)(1-氧代-1-苯基己-5-炔-2-基)氨基甲酸叔丁酯,转化为产物化合物(1)(1-羟基-1-苯基己-5-炔-2-基)氨基甲酸叔丁酯的酮还原酶多肽包括这样的氨基酸序列,该氨基酸序列具有与参考序列SEQ ID NO:2的至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多的序列同一性和特征X40R和X206L或F,条件是当X206为F时,该序列还包括与SEQ ID NO:4相比选自以下的一个或多个残基差异:X7S;X17M;X17Q;X17R;X23V;X27L;X29G;X60I;X64V;X71P;X87L;X94A;X94P;X94S;X95M;X96L;X96Y;X105G;X113I;X122A;X127R;X131S;X144V;X145L;X147I;X147L;X147Q;X150Y;X152G;X153G;X157C;X173L;X195A;X195G;X196M;X198S;X199H;X208R;X216R;X221S;X243S;X245I;X249F;X249G;和X249Y。

[0202] 在一些实施方案中,能够以多肽SEQ ID NO:4的活性的至少1.5倍将底物化合物(2)转化为产物化合物(1)的酮还原酶多肽包括这样的氨基酸序列,该氨基酸序列具有与参考序列SEQ ID NO:2的至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多的同一性和特征X40R和X206L或F,条件是当X206为F时,该序列还包括与SEQ ID NO:4相比选自以下的一个或多个残基差异:X7S;X17M;X17Q;X17R;X23V;X27L;X29G;X64V;X71P;X87L;X94A;X94P;X94S;X95M;X96L;X96Y;X105G;X113I;X122A;X127R;X131S;X147I;X147L;X147Q;X150Y;X152G;X153G;X157C;X173L;X195A;X195G;X196M;X198S;X199H;X208R;X216R;X221S;X243S;X249F;X249G;和X249Y。

[0203] 在一些实施方案中,能够以多肽SEQ ID NO:4的活性的至少1.5倍将底物化合物(2)转化为产物化合物(1)的酮还原酶多肽包括这样的氨基酸序列,该氨基酸序列具有与选自以下的参考序列的至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多的同一性:SEQ ID NO:4、6、8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、38、40、42、44、46、48、50、52、54、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172,并且具有特征X40R和X206L或F,条件是当X206为F时,该序列还包括与SEQ ID NO:4相比选自以下的一个或多个残基差异:X7S;X17M;X17Q;X17R;X23V;X27L;X29G;X60I;X64V;X71P;X87L;X94A;X94P;X94S;X95M;X96L;X96Y;X105G;X113I;X122A;X127R;X131S;X144V;X145L;X147I;X147L;X147Q;X150Y;X152G;X153G;X157C;X173L;X195A;X195G;X196M;X198S;X199H;X208R;X216R;X221S;X243S;X245I;X249F;X249G;和X249Y,且特别是与SEQ ID NO:4相比选自以下的一个或多个残基差异:X7S;X17M;X17Q;X17R;X23V;X27L;X29G;X64V;X71P;X87L;X94A;X94P;X94S;X95M;X96L;X96Y;X105G;X113I;X122A;X127R;X131S;X147I;X147L;X147Q;X150Y;X152G;X153G;X157C;X173L;X195A;X195G;X196M;X198S;X199H;X208R;X216R;X221S;X243S;X249F;X249G;和X249Y。

[0204] 在一些实施方案中,参考序列选自SEQ ID NO:4、14、36、42、和130。在一些实施方案中,参考序列为SEQ ID NO:4。在一些实施方案中,参考序列为SEQ ID NO:36。在一些实施

方案中,参考序列为SEQ ID NO:42。在一些实施方案中,参考序列为SEQ ID NO:130。

[0205] 在一些实施方案中,与如以上描述的参考序列具有氨基酸序列同一性的工程化酮还原酶多肽包括特征X40R和X206L。

[0206] 在一些实施方案中,工程化酮还原酶包括具有特征X40R和X206L以及与SEQ ID NO:4相比选自以下的一个或多个残基差异的氨基酸序列:X7S;X17M;X17Q;X17R;X23V;X27L;X29G;X60I;X64V;X71P;X87L;X94A;X94P;X94S;X95M;X96L;X96Y;X105G;X113I;X122A;X127R;X131S;X144V;X145L;X147I;X147L;X147Q;X150Y;X152G;X153G;X157C;X173L;X195A;X195G;X196M;X198S;X199H;X208R;X216R;X221S;X243S;X245I;X249F;X249G;和X249Y。

[0207] 在一些实施方案中,与如以上描述的参考序列具有氨基酸序列同一性的工程化酮还原酶多肽包括特征X40R和X206L以及与SEQ ID NO:4相比选自以下的一个或多个残基差异:X7S;X17M;X17Q;X17R;X23V;X27L;X29G;X60I;X64V;X71P;X87L;X94A;X94P;X94S;X95M;X96L;X96Y;X105G;X113I;X122A;X127R;X131S;X144V;X145L;X147I;X147L;X147Q;X150Y;X152G;X153G;X157C;X173L;X195A;X195G;X196M;X198S;X199H;X208R;X216R;X221S;X243S;X245I;X249F;X249G;和X249Y。

[0208] 在一些实施方案中,与如以上描述的参考序列具有氨基酸序列同一性的工程化酮还原酶多肽包括特征X40R和X206L以及与SEQ ID NO:4相比选自以下的一个或多个残基差异:X7S;X17M;X17Q;X17R;X23V;X27L;X29G;X64V;X71P;X87L;X94A;X94P;X94S;X95M;X96L;X96Y;X105G;X113I;X122A;X127R;X131S;X147I;X147L;X147Q;X150Y;X152G;X153G;X157C;X173L;X195A;X195G;X196M;X198S;X199H;X208R;X216R;X221S;X243S;X249F;X249G;和X249Y。

[0209] 在一些实施方案中,具有特征X40R和X206F,以及与SEQ ID NO:4相比选自以下的一个或多个残基差异的工程化酮还原酶多肽:X7S;X17M;X17Q;X17R;X23V;X27L;X29G;X64V;X71P;X87L;X94A;X94P;X94S;X95M;X96L;X96Y;X105G;X113I;X122A;X127R;X131S;X147I;X147L;X147Q;X150Y;X152G;X153G;X157C;X173L;X195A;X195G;X196M;X198S;X199H;X208R;X216R;X221S;X243S;X249F;X249G;和X249Y,还可包括与SEQ ID NO:4相比选自X60I,X144V;X145L;和X245I的一个或多个残基差异。

[0210] 在一些实施方案中,具有特征X40R和X206F,以及与SEQ ID NO:4相比选自以下的一个或多个残基差异的工程化酮还原酶:X7S;X17M;X17Q;X17R;X23V;X27L;X29G;X60I;X64V;X71P;X87L;X94A;X94P;X94S;X95M;X96L;X96Y;X105G;X113I;X122A;X127R;X131S;X144V;X145L;X147I;X147L;X147Q;X150Y;X152G;X153G;X157C;X173L;X195A;X195G;X196M;X198S;X199H;X208R;X216R;X221S;X243S;X245I;X249F;X249G;和X249Y包括选自以下的氨基酸序列:SEQ ID NO:18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172。

[0211] 在一些实施方案中,工程化酮还原酶包括具有特征X40R和X206F或L以及与SEQ ID NO:4相比选自以下的至少一个或多个残基差异的氨基酸序列:X17Q/R/M;X64V;X94P;X144V;X147Q/I/L;X157C;和X196M。在一些实施方案中,工程化酮还原酶包括具有特征X40R

和X206F/L以及至少X17Q的氨基酸序列。在一些实施方案中,工程化酮还原酶包括具有特征X40R和X206F/L以及至少X64V的氨基酸序列。在一些实施方案中,工程化酮还原酶包括具有特征X40R和X206F/L以及至少X94P的氨基酸序列。在一些实施方案中,工程化酮还原酶包括具有特征X40R和X206F/L以及至少X144V的氨基酸序列。在一些实施方案中,工程化酮还原酶包括具有特征X40R和X206F/L以及至少X147I的氨基酸序列。在一些实施方案中,工程化酮还原酶包括具有特征X40R和X206F/L以及至少X157C的氨基酸序列。在一些实施方案中,工程化酮还原酶包括具有特征X40R和X206F/L以及至少X196M的氨基酸序列。在前述实施方案中的每一个中,在本文描述的其他残基位置上的另外残基差异可存在于工程化酮还原酶中。

[0212] 在一些实施方案中,工程化酮还原酶包括具有与SEQ ID NO:4相比至少选自以下的残基差异组合的氨基酸序列:(a)X40R、X196M、和X206F/L;(b)X40R、X144V、和X206F/L;(c)X40R、X17H/R/Q、和X206F/L;(d)X40R、X94P、和X206F/L;(e)X40R、X198S、和X206F/L;(f)X40R、X17H/R/Q、X94P和X206F/L;(g)X40R、X71P、X157C和X206F/L;(h)X40R、X94P、X144V、X196M和X206F/L;(i)X17H/R/Q、X40R、X64V、X147I/Q/L、和X206F/L;(g)X17H/R/Q、X40R、X64V、X94P、X144V、X147I/Q/L、X157C、X196M和X206F/L;和(k)X17Q、X40R、X64V、X71P、X94P、X144V、X147I、X157C、X196M、X206F、和X249F。

[0213] 在一些实施方案中,工程化酮还原酶多肽能够将底物化合物(2)(1-氧代-1-苯基己-5-炔-2-基)氨基甲酸叔丁酯,转化为产物化合物(1)(1-羟基-1-苯基己-5-炔-2-基)氨基甲酸叔丁酯,具有相对于SEQ ID NO:4的参考多肽的活性至少2倍、3倍、4倍、5倍、10倍、20倍、30倍、40倍、50倍、100倍、200倍、500倍、1000倍、5000倍、10000倍或更多倍的活性。在一些实施方案中,能够将底物化合物(2)转化为产物化合物(1),具有相对于SEQ ID NO:4的参考多肽的活性至少2倍、3倍、4倍、5倍、10倍、20倍、30倍、40倍、50倍、100倍、200倍、500倍、1000倍、5000倍、10000倍或更多倍的活性的工程化酮还原酶多肽,包括具有特征H40R和X206F或L以及与SEQ ID NO:4相比选自以下的一个或多个残基差异的氨基酸序列:X60I;X71P;X94P;X94A;X95M;X96L;X96Y;X127R;X144V;X145I;X150Y;X152G;X153G;X157C;X195A;X195G;X196M;X198S;X199H;X206F,X216R,X245I,X245F;X249Y;和X249F。

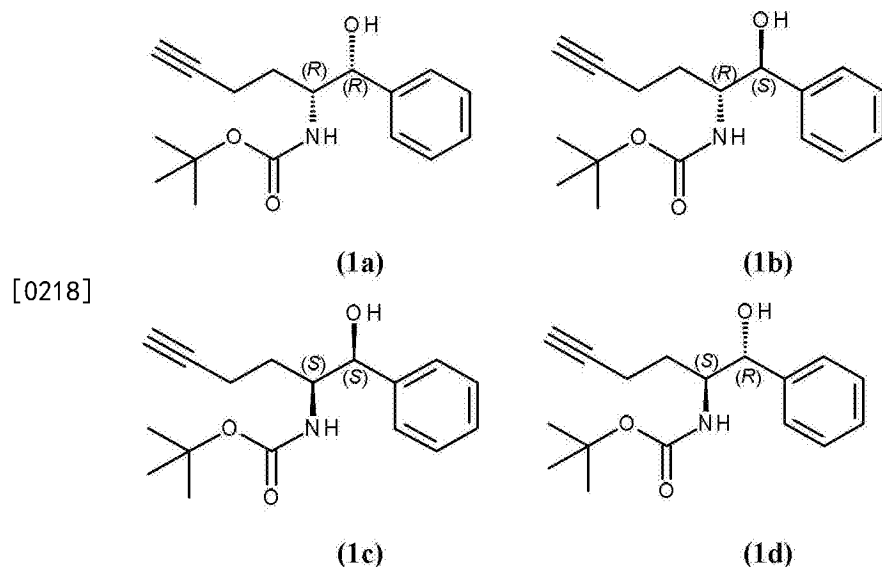
[0214] 在一些实施方案中,能够将底物化合物(2)转化为产物化合物(1),具有SEQ ID NO:4的参考多肽的活性的至少10倍活性的工程化酮还原酶多肽,当在HTP条件下测量时,包括选自以下的氨基酸序列:SEQ ID NO:18、20、22、24、26、28、30、34、36、38、40、42、44、50、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172。

[0215] 在一些实施方案中,当在HTP条件下测量时,工程化酮还原酶能够以SEQ ID NO:4的活性的至少100倍活性将化合物(2)转化为化合物(1)。在一些实施方案中,能够以SEQ ID NO:4的活性的至少100倍活性将化合物(2)转化为化合物(1)的工程化酮还原酶,包括选自以下的氨基酸序列:SEQ ID NO:34、36、38、40、42、44、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172。

[0216] 在一些实施方案中,能够将化合物(2)转化为化合物(1)的工程化酮还原酶与SEQ ID NO:4或32的参考多肽相比具有增加的热稳定性。在一些实施方案中,具有增加的热稳定

性的工程化酮还原酶多肽包括具有与SEQ ID NO:4相比选自以下的一个或多个残基差异的氨基酸序列:X17Q;X17R;X17W;X64V;X71P;X87L;X94S;X94P;X147Q;X147I;X147L;X157C;X198S;X249Y;和X249F。在一些实施方案中,具有增加的热稳定性的工程化酮还原酶包括具有特征H40R和X206F或L,以及与SEQ ID NO:4相比选自以下的一个或多个残基差异的氨基酸序列:X17Q;X17R;X17W;X64V;X71P;X87L;X94S;X94P;X147Q;X147I;X147L;X157C;X198S;X249Y;和X249F。

[0217] 在一些实施方案中,本文公开的工程化酮还原酶多肽能够形成非对映体过量于反式-(1-羟基-1-苯基己-5-炔-2-基)氨基甲酸叔丁酯化合物(1b)和(1d)的产物顺式-(1-羟基-1-苯基己-5-炔-2-基)氨基甲酸叔丁酯化合物(1a)和(1c)



[0219] 在一些实施方案中,相比于对化合物(1b)和(1d)具有对产物化合物(1a)和(1c)的非对映选择性的工程化酮还原酶多肽包括具有特征X40R和X206F或L以及与SEQ ID NO:4相比选自以下的一个或多个残基差异的氨基酸序列:X17Q;X64V;X71P;X87L;X147I;X157C;X196M;和X249F。

[0220] 如上面所提及,示例性工程化酮还原酶指示,X195G的存在改变了非对映选择性以相比于产物化合物(1a)和(1c)有利于形成产物化合物(1b)和(1d)。因此,在相比于对产物化合物(1b)和(1d)具有对产物化合物(1a)和(1c)的非对映选择性的工程化酮还原酶多肽的一些实施方案中,X195G的存在是不受欢迎的。

[0221] 在一些实施方案中,相比于对化合物(1b)和(1d)具有对产物化合物(1a)和(1c)的非对映选择性的工程化酮还原酶多肽包括选自以下的氨基酸序列:SEQ ID NO:18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172。

[0222] 在一些实施方案中,工程化酮还原酶能够以相比于化合物(1b)和(1d)的至少50的非对映体比形成产物化合物(1a)和(1c)。在一些实施方案中,能够以相比于化合物(1b)和(1d)的至少50的非对映体比形成产物化合物(1a)和(1c)的工程化酮还原酶包括选自以下的氨基酸序列:SEQ ID NO:18、20、22、24、26、28、32、34、36、38、40、42和44。

[0223] 在一些实施方案中,工程化酮还原酶能够以相比于产物化合物(1c)的非对映体过量形成产物化合物(1a)。相比于对化合物(1c)具有对化合物(1a)的非对映选择性的酮还原酶的氨基酸序列具有与相比于对化合物(1b)和(1d)具有对化合物(1a)和(1c)的非对映选择性的工程化酮还原酶类似的特征。因此,在一些实施方案中,相比于对化合物(1c)具有对产物化合物(1a)的非对映选择性的酮还原酶多肽包括具有特征X40R和X206F或L以及与SEQ ID NO:4相比选自以下的一个或多个残基差异的氨基酸序列:X17Q、X64V、X71P、X87L、X147I、X157C、X196M、和X249F。

[0224] 在一些实施方案中,能够以相比于化合物(1c)的非对映体过量形成产物化合物(1a)的工程化酮还原酶多肽包括选自以下的氨基酸序列:SEQ ID NO:18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172。

[0225] 在一些实施方案中,工程化酮还原酶多肽能够形成非对映体过量于化合物(1a)和(1c)的产物化合物(1b)和(1d)。在一些实施方案中,能够形成非对映体过量于化合物(1a)和(1c)的产物化合物(1b)和(1d)的工程化酮还原酶包括这样的氨基酸序列,该氨基酸序列具有与SEQ ID NO:4的参考序列至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多的同一性和具有特征X195G。

[0226] 在一些实施方案中,能够形成非对映体过量于化合物(1a)和(1c)的产物化合物(1b)和(1d)的工程化酮还原酶包括具有特征X40R和X206F或L、以及X195G的氨基酸序列。

[0227] 在一些实施方案中,能够形成非对映体过量于化合物(1a)和(1c)的产物化合物(1b)和(1d)的工程化酮还原酶包括选自SEQ ID NO:50和74的氨基酸序列。

[0228] 在一些实施方案中,能够在适当的反应条件下将化合物(2)转化为化合物(1)的工程化酮还原酶包括这样的氨基酸序列,该氨基酸序列具有与以下中的一个至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、或99%同一性:SEQ ID NO:6、8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、38、40、42、44、46、48、50、52、54、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172,和与SEQ ID NO:4相比存在于以下中的任何一个中的氨基酸残基差异:SEQ ID NO:6、8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、38、40、42、44、46、48、50、52、54、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172。在一些实施方案中,存在于选自SEQ ID NO:6、8、10、12、14、16、46、48、50、52、和54的氨基酸序列中的一个或多个的残基差异明确地被排除在前述实施方案之外。

[0229] 因此,在一些实施方案中,能够在适当的反应条件下将化合物(2)转化为化合物(1)的工程化酮还原酶多肽包括这样的氨基酸序列,该氨基酸序列具有与以下之一的至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、或

99%同一性:SEQ ID NO:18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172,和与SEQ ID NO:4相比存在于以下中的任何一个中的氨基酸残基差异:SEQ ID NO:18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172。

[0230] 除上面指定的残基位置之外,本文所公开的任何工程化酮还原酶多肽还可以包括相对于SEQ ID NO:2或4的参考多肽序列,在其它残基位置即,除以下残基位置外的残基位置的残基差异:X7;X17;X23;X27;X29;X60;X64;X71;X87;X94;X95;X96;X105;X113;X122;X127;X131;X144;X145;X147;X150;X152;X153;X157;X173;X195;X196;X198;X199;X208;X216;X221;X243;X245;和X249。在这些其它残基位置上的残基差异可以提供氨基酸序列的另外变体而没有改变多肽将底物化合物(2)转化为产物化合物(1)的能力,特别是关于用于形成化合物(1a)和(1c)的增加的活性;相比于对化合物(1b)和(1d)具有对产物化合物(1a)和(1c)的非对映选择性;相比于对化合物(1c)具有对产物化合物(1a)的非对映选择性;和相比于对化合物(1a)和(1c)具有对产物化合物(1b)和(1d)的非对映选择性。因此,在一些实施方案中,除选自以下的工程化酮还原酶多肽的任何一个的氨基酸残基差异之外:SEQ ID NO:6、8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、38、40、42、44、46、48、50、52、54、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172,序列还可以包括与SEQ ID NO:4相比在其它氨基酸残基位置上的1-2个、1-3个、1-4个、1-5个、1-6个、1-7个、1-8个、1-9个、1-10个、1-11个、1-12个、1-14个、1-15个、1-16个、1-18个、1-20个、1-22个、1-24个、1-26个、1-30个、1-35个、1-40个、1-45个、或1-50个残基差异。在一些实施方案中,与参考序列相比氨基酸残基差异的数目可以是1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个、21个、22个、23个、24个、25个、30个、30个、35个、40个、45个、或50个残基位置。在一些实施方案中,在其它氨基酸残基位置上的残基差异可以包括与SEQ ID NO:2的野生型多肽或SEQ ID NO:4的工程化酮还原酶多肽的参考序列相比的保守置换和/或非保守置换。

[0231] 相对于SEQ ID NO:2的野生型序列在其它位置上的氨基酸残基差异和这些差异对酶功能的影响针对其它工程化酮还原酶多肽被描述,该其它工程化酮还原酶多肽公开于已公布的PCT申请W02008103248、W02009029554、W02009036404、W02009042984、W02010027710、W02010025238、W02010025287、W02010025085、W0/2009/046153、W02011022548;和W02011/140219;以及2011年4月13日提交的美国临时申请号61/475,103;这些申请的每一个据此通过引用并入本文。因此,在一些实施方案中,与SEQ ID NO:2或4的序列相比的氨基酸差异中的一个或多个还可以在选自以下的残基位置上被引入到本公开内容的工程化酮还原酶多肽:X2;X3;X4;X8;X9;X10;X11;X12;X16;X19;X21;X25;X41;X42;X43;X45;X46;X49;X53;X54;X57;X60;X66;X68;X72;X74;X75;X76;X77;X78;X80;X82;X86;

X93;X97;X99;X100;X101;X104;X106;X108;X109;X111;X112;X117;X120;X124;X125;X126;X129;X134;X141;X144;X145;X148;X149;X151;X155;X159;X163;X165;X169;X176;X177;X178;X179;X185;X186;X190;X192;X194;X197;X200;X201;X202;X203;X204;X205;X207;X210;X211;X212;X214;X217;X223;X225;X226;X228;X233;X235;X236;X245;X248;X250;和X251。特别地,在所述位置上的氨基酸残基的选择可以选自以下:X2A/S;X3Y/N/V;X4C;X8R/N;X9G;X10T;X11V/T/F/L;X12I;X16A/G/V/S;X19V;X21F/R;X25N/R/T;X41V/T;X42G;X43A/I/R;X45G;X46R;X49R;X53D/V;X54A;X57V;X60A/I;X66E;X68V;X72R/E/T;X74L;X75N;X76A/I;X77A;X78D;X80T/V;X82S;X86I;X93S/A/T;X97G/E/L/H/I/M/R/T/V;X99L;X100K;X101G/N;X104M;X106G/D;X108H/N/S/D/K;X109R/E;X111M;X112D;X117S/A/G;X120V;X124Q;X125S;X126V;X129T;X134M;X141V;X144V;X145D/S/L/Q/F/Y/A/M/V/K;X148I;X149F;X151A;X155C;X159T;X163I;X165T/N;X169C;X176V;X177R;X178G;X179F;X185S;X186I;X190A/C/P/H/G/F/N/L/E/I/V;X192E/R;X194D/G/N/L/Q/S/R;X197G/E/V/A;X200K/P;X201A/I/L;X202G/I/L/M/W/Y/V/F/N;X203G;X204A/V;X205T/R/V;X207C/I/N/T;X210R;X211R/I/L/T/V;X212S/V;X214T/V;X217F;X223V/G/I;X225V;X226L/T/V;X228A;X233A/G;X235W;X236R;X245I;X248K/R;X250I/Y;和X251T。对在残基位置上的氨基酸残基的选择的指导可以参见所引用的参考文献。

[0232] 在一些实施方案中,本公开内容还提供了包括本文描述的任何工程化酮还原酶多肽的片段的工程化酮还原酶多肽,该片段保留了该工程化酮还原酶的功能活性和/或改进特性。因此,在一些实施方案中,本公开内容提供了能够在适当的反应条件下将化合物(2)转化为化合物(1)的多肽片段,其中该片段包括本公开内容的工程化酮还原酶多肽的全长氨基酸序列的至少约80%、90%、95%、98%、或99%,工程化酮还原酶多肽诸如选自以下的示例性工程化酮还原酶多肽:SEQ ID NO:6、8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、38、40、42、44、46、48、50、52、54、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172,并且特别是选自以下的示例性工程化多肽:SEQ ID NO:18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172。

[0233] 在一些实施方案中,本公开内容的工程化酮还原酶多肽可以具有包括本文描述的工程化酮还原酶多肽中的任一个的缺失的氨基酸序列,诸如以下的示例性工程化多肽:SEQ ID NO:4、6、8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、38、40、42、44、46、48、50、52、54、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172,并且特别是选自以下的示例性工程化多肽:SEQ ID NO:18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、

140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172。因此,对于本公开内容的工程化酮还原酶多肽的每一个实施方案,当保留本文描述的工程化酮还原酶的相关功能活性和/或改进特性时,氨基酸序列可以包括酮还原酶多肽的一个或多个氨基酸、2个或更多个氨基酸、3个或更多个氨基酸、4个或更多个氨基酸、5个或更多个氨基酸、6个或更多个氨基酸、8个或更多个氨基酸、10个或更多个氨基酸、15个或更多个氨基酸、或20个或更多个氨基酸、多至氨基酸的总数的10%、多至氨基酸的总数的10%、多至氨基酸的总数的20%、或多至氨基酸的总数的30%的缺失。在一些实施方案中,缺失可以包括1-2个、1-3个、1-4个、1-5个、1-6个、1-7个、1-8个、1-9个、1-10个、1-15个、1-20个、1-21个、1-22个、1-23个、1-24个、1-25个、1-30个、1-35个、1-40个、1-45个、或1-50个氨基酸残基。在一些实施方案中,缺失的数目可以是1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个、21个、22个、23个、24个、25个、30个、30个、35个、40个、45个、或50个氨基酸残基。在一些实施方案中,缺失可以包括1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、18个、20个、21个、22个、23个、24个、或25个氨基酸残基的缺失。

[0234] 在一些实施方案中,本公开内容的工程化酮还原酶多肽可具有与本文描述的工程化酮还原酶多肽中的任何一个相比包括插入的氨基酸序列,工程化酮还原酶多肽诸如以下的示例性工程化多肽:SEQ ID NO:4、6、8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、38、40、42、44、46、48、50、52、54、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172,并且特别是选自以下的示例性工程化多肽:SEQ ID NO:18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172。因此,对于本公开内容的酮还原酶多肽的每一个实施方案,当保留本文描述的工程化酮还原酶的相关功能活性和/或改进特性时,插入可以包括一个或多个氨基酸、2个或更多个氨基酸、3个或更多个氨基酸、4个或更多个氨基酸、5个或更多个氨基酸、6个或更多个氨基酸、8个或更多个氨基酸、10个或更多个氨基酸、15个或更多个氨基酸、20个或更多个氨基酸、30个或更多个氨基酸、40个或更多个氨基酸、或50个或更多个氨基酸。插入可以是酮还原酶多肽的氨基末端或羧基末端、或内部部分。

[0235] 在一些实施方案中,本公开内容提供能够在适当的反应条件下将化合物(2)转化为化合物(1)的工程化酮还原酶多肽,该工程化酮还原酶多肽包括这样的氨基酸序列,该氨基酸序列具有与SEQ ID NO:6、8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、38、40、42、44、46、48、50、52、54、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、或99%的同一性,并且特别是选自以下的示例性工程化多肽:SEQ ID NO:18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、74、76、78、

80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172,条件是,该氨基酸序列与公开于以下的示例性工程化酮还原酶多肽氨基酸序列不同(也就是,该氨基酸序列不包括公开于以下的示例性工程化酮还原酶多肽氨基酸序列):已公布的PCT申请W02008103248、W02009029554、W02009036404、W02009042984、W02010027710、W02010025238、W02010025287、W02010025085、W02009046153、W02011022548;和W02011/140219;和2011年4月13日提交的美国临时申请号61/475,103;这些申请的每一个据此通过引用并入本文。

[0236] 在上述实施方案中,用于工程化多肽的适当的反应条件可以是表2A、2B、和2C中所描述的那些条件。因此,在一些实施方案中,适当的反应条件是对HTP测定描述的那些条件,该条件包括:20或50g/L的底物化合物(2);60 μ L、20 μ L、或40 μ L含有工程化酮还原酶的细胞溶解产物;30%或50%(v/v)异丙醇;0.5g/L或0.1g/L NADP;0.2M硼酸盐;pH 9.5或pH 10;35 $^{\circ}$ C、45 $^{\circ}$ C或55 $^{\circ}$ C的温育温度;和24h的反应时间。这些反应条件和酮还原酶多肽的使用的指导提供于表2A以及其他中。

[0237] 在一些实施方案中,适当的反应条件是对摇瓶粉(SFP)测定描述的那些条件,该条件包括:5g/L或40g/L底物化合物(2);5g/L或2g/L的酮还原酶多肽;30%或50%(v/v)异丙醇;0.5或0.1g/L NADP;0.2M硼酸盐;pH 9.5、10、10.2;35 $^{\circ}$ C或45 $^{\circ}$ C的温育温度;和6h或24h的反应时间。这些反应条件和酮还原酶多肽的使用的指导提供于表2B以及其他中。

[0238] 在一些实施方案中,适当的反应条件是对下游处理粉(DSP)测定描述的那些条件,该条件包括:40g/L或100g/L底物化合物(2);5g/L或1g/L的酮还原酶多肽;30%或50%(v/v)异丙醇;0.1g/L NADP;0.2M硼酸盐;pH 9.5或10;35 $^{\circ}$ C或45 $^{\circ}$ C的温育温度;和22h至24h的反应时间。这些反应条件和酮还原酶多肽的使用的指导提供于表2C以及其他中。

[0239] 在一些实施方案中,本公开内容的多肽可以是融合多肽的形式,其中工程化多肽与其它多肽融合,诸如通过举例的方式而非限制,抗体标签(例如,myc表位)、纯化序列(例如,用于结合至金属的His标签)和细胞定位信号(例如,分泌信号)。因此,本文描述的工程化多肽可与其它多肽融合或不与其它多肽融合使用。

[0240] 应理解的是,本文描述的工程化酮还原酶多肽不限于遗传编码的氨基酸。因此,除了遗传编码的氨基酸以外,本文描述的多肽可以完全或部分由天然存在的和/或合成的非编码氨基酸组成。本文描述的多肽可包括的某些常见非编码氨基酸包括但不限于:遗传编码的氨基酸的D-立体异构体;2,3-二氨基丙酸(Dpr); α -氨基异丁酸(Aib); ϵ -氨基己酸(Aha); δ -氨基戊酸(Ava);N-甲基甘氨酸或肌氨酸(MeGly或Sar);鸟氨酸(Orn);瓜氨酸(Cit);叔丁基丙氨酸(Bua);叔丁基甘氨酸(Bug);N-甲基异亮氨酸(MeIle);苯基甘氨酸(Phg);环己基丙氨酸(Cha);正亮氨酸(Nle);萘基丙氨酸(Nal);2-氯苯丙氨酸(OCF);3-氯苯丙氨酸(Mcf);4-氯苯丙氨酸(Pcf);2-氟苯丙氨酸(Off);3-氟苯丙氨酸(Mff);4-氟苯丙氨酸(Pff);2-溴苯丙氨酸(Obf);3-溴苯丙氨酸(Mbf);4-溴苯丙氨酸(Pbf);2-甲基苯丙氨酸(OMf);3-甲基苯丙氨酸(Mmf);4-甲基苯丙氨酸(Pmf);2-硝基苯丙氨酸(Onf);3-硝基苯丙氨酸(Mnf);4-硝基苯丙氨酸(Pnf);2-氰基苯丙氨酸(OCf);3-氰基苯丙氨酸(Mcf);4-氰基苯丙氨酸(Pcf);2-三氟甲基苯丙氨酸(OTf);3-三氟甲基苯丙氨酸(Mtf);4-三氟甲基苯丙氨酸(Ptf);4-氨基苯丙氨酸(Paf);4-碘苯丙氨酸(Pif);4-氨甲基苯丙氨酸(Pamf);2,4-

二氯苯丙氨酸(Opbf);3,4-二氯苯丙氨酸(Mpcf);2,4-二氟苯丙氨酸(Opff);3,4-二氟苯丙氨酸(Mpff);吡啶-2-基丙氨酸(2pAla);吡啶-3-基丙氨酸(3pAla);吡啶-4-基丙氨酸(4pAla);萘-1-基丙氨酸(1nAla);萘-2-基丙氨酸(2nAla);噻唑基丙氨酸(taAla);苯并噻吩基丙氨酸(bAla);噻吩基丙氨酸(tAla);呋喃基丙氨酸(fAla);高苯丙氨酸(hPhe);高酪氨酸(hTyr);高色氨酸(hTrp);五氟苯丙氨酸(5ff);苯乙烯基丙氨酸(sAla);苄基丙氨酸(aAla);3,3-二苯丙氨酸(Dfa);3-氨基-5-苯基戊酸(Afp);青霉素(Pen);1,2,3,4-四氢异喹啉-3-羧酸(Tic);E-2-噻吩基丙氨酸(Thi);蛋氨酸亚砷(Mso);N(w)-硝基精氨酸(nArg);高赖氨酸(hLys);磷酸基甲基苯丙氨酸(pmPhe);磷酸丝氨酸(pSer);磷酸苏氨酸(pThr);高天冬氨酸(hAsp);高谷氨酸(hGlu);1-氨基环戊-(2或3)-烯-4-羧酸;哌可酸(PA);氮杂环丁烷-3-羧酸(ACA);1-氨基环戊烷-3-羧酸;烯丙基甘氨酸(aOly);炔丙基甘氨酸(pgGly);高丙氨酸(hAla);正缬氨酸(nVal);高亮氨酸(hLeu);高缬氨酸(hVal);高异亮氨酸(hIle);高精氨酸(hArg);N-乙酰赖氨酸(AcLys);2,4-二氨基丁酸(Dbu);2,3-二氨基丁酸(Dab);N-甲基缬氨酸(MeVal);高半胱氨酸(hCys);高丝氨酸(hSer);羟基脯氨酸(Hyp)和高脯氨酸(hPro)。本文描述的多肽可包括的另外的非编码氨基酸对本领域技术人员将是明显的(参见,例如,在Fasman,1989,CRC Practical Handbook of Biochemistry and Molecular Biology,CRC Press,Boca Raton,FL,在第3-70页及其中引用的参考文献中提供的多种氨基酸,全部参考文献通过引用并入本文)。这些氨基酸可以是以L-构型或D-构型。

[0241] 本领域技术人员将认识到,带有侧链保护基的氨基酸或残基还可以构成本文所描述的多肽。在这种情况下属于芳香族类别的这些受保护的氨基酸的非限制性实例包括(在括号中列出的保护基)但不限于:Arg(tos)、Cys(苄基)、Cys(硝基吡啶亚磺酰基)、Glu(δ -苄基酯)、Gln(咕吨基)、Asn(N- δ -咕吨基)、His(bom)、His(苄基)、His(tos)、Lys(fmoc)、Lys(tos)、Ser(O-苄基)、Thr(O-苄基)和Tyr(O-苄基)。

[0242] 本文所述的多肽可包括的构型上受限制的非编码氨基酸包括但不限于,N-甲基氨基酸(L-构型);1-氨基酸环戊-(2或3)-烯-4-羧酸;哌可酸;氮杂环丁烷-3-羧酸;高脯氨酸(hPro);以及1-氨基环戊烷-3-羧酸。

[0243] 在一些实施方案中,工程化酮还原酶多肽可以被提供在固体支持物上,诸如膜、树脂、固体载体(solid carrier)、或其它固相材料。固体支持物可以由有机聚合物如聚苯乙烯、聚乙烯、聚丙烯、聚氟乙烯、聚乙烯氧和聚丙烯酰胺以及其共聚物和接枝物组成。固体支持物还可以是无机的,诸如玻璃、二氧化硅、可控孔度玻璃(CPG)、反相二氧化硅或金属诸如金或铂。固体支持物的结构可以呈珠、球、粒子、颗粒、凝胶、膜或表面的形式。表面可以是平面的、基本上平面的或非平面的。固体支持物可以是多孔的或非多孔的,并且可以具有溶胀的或非溶胀的性质。固体支持物可以被配置成孔、凹部或其它容器、器皿(vessel)、特征(feature)或位置(location)的形式。

[0244] 在一些实施方案中,本公开内容的具有酮还原酶活性的工程化多肽可以被固定在固体支持物上,使得它们保留它们的相对于SEQ ID NO:4的参考多肽的改进活性、立体选择性、和/或其它改进特性。在这样的实施方案中,固定的多肽可以促进化合物(2)或其结构类似物的底物生物催化转化为化合物(1)或相应结构类似物的产物(例如,如本文描述的方案1、2和3的方法中所示的),并且在反应完全后很容易保留(例如,通过保留在其上固定多肽的珠),并且然后在后续反应中再利用或回收。这样的固定的酶的方法允许更高的效率和成

本降低。因此,进一步设想,使用本公开内容的工程酮还原酶多肽的方法中的任一种,可以使用结合或固定在固体支持物上的相同的工程酮还原酶多肽来进行。

[0245] 酶固定的方法是本领域中熟知的。可非共价地或共价地结合工程化酮还原酶多肽。用于缀合和固定酶至固体支持物(例如,树脂、膜、珠、玻璃等等)的各种方法是本领域熟知的并描述于例如:Yi等人,“Covalent immobilization of ω -transaminase from *Vibrio fluvialis* JS17 on chitosan beads,”*Process Biochemistry* 42(5):895-898 (2007年五月);Martin等人,“Characterization of free and immobilized(S)-aminotransferase for acetophenone production,”*Applied Microbiology and Biotechnology* 76(4):843-851(2007年九月);Koszelewski等人,“Immobilization of ω -transaminases by encapsulation in a sol-gel/celite matrix,”*Journal of Molecular Catalysis B:Enzymatic*,63:39-44(2010年四月);Truppo等人,“Development of an Improved Immobilized CAL-B for the Enzymatic Resolution of a Key Intermediate to Olanzapine,”*Organic Process Research&Development*,在线公布:dx.doi.org/10.1021/op200157c;Hermanson,G.T.,*Bioconjugate Techniques*,第二版,Academic Press(2008);Mateo等人,“Epoxy sephabeads:a novel epoxy support for stabilization of industrial enzymes via very intense multipoint covalent attachment,”*Biotechnology Progress* 18(3):629-34(2002);以及*Bioconjugation Protocols:Strategies and Methods*,In *Methods in Molecular Biology*,C.M.Niemeyer 编著,Humana Press(2004);其每个的公开内容通过引用并入本文。可用于固定本公开内容的工程化酮还原酶的固体支持物包括但不限于,包括以下的珠或树脂:具有环氧官能团的聚甲基丙烯酸酯、具有氨基环氧官能团的聚甲基丙烯酸酯、具有十八烷基官能团的苯乙烯/DVB共聚物或聚甲基丙烯酸酯。可用于固定本公开内容的工程化酮还原酶的示例性固体支持物包括但不限于,壳聚糖珠、Eupergit C和SEPABEAD(Mitsubishi),包括以下不同类型的SEPABEAD:EC-EP、EC-HFA/S、EXA252、EXE119和EXE120。

[0246] 在一些实施方案中,工程化多肽可以是各种形式,例如,诸如分离的制剂,作为基本上纯的酶、用编码酶的基因转化的整个细胞、和/或作为细胞提取物和/或此类细胞的溶解产物。酶可以冻干、喷雾干燥、沉淀、或为粗浆料的形式,如下面进一步讨论的。

[0247] 在一些实施方案中,工程化多肽可以试剂盒的形式被提供。试剂盒中的多肽可以单独地存在或作为多种多肽存在。试剂盒还可以包括用于进行酶促反应的试剂、用于评估多肽活性的底物、以及用于检测产物的试剂。试剂盒还可包括试剂分配器和试剂盒的使用说明书。

[0248] 在一些实施方案中,多肽可以以阵列的形式被提供在固体支持物上,其中多肽被布置在定位上不同的位置中。阵列可用于测试用于由多肽转化的各种底物化合物。多种支持物可以配置在阵列上的不同的位置,这对于试剂的机器人递送或通过检测方法和/或仪器是可寻址的。用于缀合底物,例如膜、珠、玻璃等等的各种方法描述于Hermanson,G.T.,*Bioconjugate Techniques*,第2版,Academic Press;(2008),以及*Bioconjugation Protocols:Strategies and Methods*,In *Methods in Molecular Biology*,C.M.Niemeyer 编著,Humana Press(2004),以及其他;这些的公开内容通过引用并入本文。

[0249] 在一些实施方案中,本公开内容的试剂盒包括以下的阵列,其包括在不同的可寻

址的位置的本文公开的多种不同的工程化酮还原酶多肽,其中不同的多肽是参考序列的不同的变体,所述变体各自具有至少一种不同的改进的酶性质。包括多种工程化多肽的这样的阵列和它们的使用方法被描述在例如W02009008908中。

[0250] 可用于制备工程化酮还原酶多肽的多核苷酸、表达载体、和宿主细胞

[0251] 在另一个方面,本公开内容提供编码本文描述的非天然存在的酮还原酶多肽的多核苷酸。这些多核苷酸可以与控制基因表达的一个或多个异源的调节序列可操作地连接,以产生能够表达酮还原酶多肽的重组的多核苷酸。包括编码工程化酮还原酶多肽的异源的多核苷酸的表达构建体可以被引入适当的宿主细胞以表达相应的多肽。

[0252] 如对本领域技术人员将是明显的,蛋白序列的可用性和相应于各种氨基酸的密码子的知识提供了对能够编码该主题多肽的所有多核苷酸的描述。当相同氨基酸由替代的或同义的密码子编码时,遗传密码的简并性允许极大数目的核酸被制出,所有这些核酸编码改进的酮还原酶。因此,如果具有具体的氨基酸序列的知识,本领域技术人员能够以不改变蛋白的氨基酸序列的方式通过仅仅变更序列的一个或多个密码子来制出许多的不同核酸。在这点上,本公开内容明确涵盖可通过选择基于可能的密码子选择的组合制出的编码本文描述的多肽的多核苷酸的每一种可能的变体,并且所有这些变体将被认为针对本文描述的任何多肽被明确地公开,所述本文描述的任何多肽包括在表2A、2B和2C中呈现的以及作为以下通过引用并入本文的序列表中公开的氨基酸序列:SEQ ID NO:4、6、8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、38、40、42、44、46、48、50、52、54、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172,并且特别是SEQ ID NO:18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172。

[0253] 在一些实施方案中,可以选择和/或工程化多核苷酸以包括优选地选择的密码子以适合在其中产生蛋白的宿主细胞。例如,细菌中使用的优选的密码子被用于在细菌中表达基因,例如大肠杆菌;酵母中使用的优选的密码子被用于酵母中的表达;并且哺乳动物中使用的优选的密码子被用于哺乳动物细胞中的表达。因为不是所有的密码子需要被代替以优化酮还原酶的密码子使用(例如,因为天然序列可以有优选的密码子,并可以不需要对所有的氨基酸残基使用优选的密码子),编码酮还原酶多肽的密码子优化的多核苷酸可在全长编码区的密码子位置的约40%、50%、60%、70%、80%或大于90%处含有优选的密码子。

[0254] 在一些实施方案中,多核苷酸编码能够将化合物(2)或其结构类似物转化为化合物(1)或相应的结构类似物的非天然存在的多肽,其中多肽包括具有与参考序列的至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、或99%同一性的氨基酸序列,参考序列是选自以下的参考序列:SEQ ID NO:4、6、8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、

158、160、162、164、166、168、170和172,并且特别是选自以下的参考序列:18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172。

[0255] 在一些实施方案中,多核苷酸编码能够将化合物(2)或其结构类似物转化为化合物(1)或相应的结构类似物并且具有与本文描述的任何参考多肽的任何指定序列同一性的工程化酮还原酶多肽,并且包括与SEQ ID NO:2或4相比在以下残基位置上的一个或多个残基差异:X7S;X17M;X17Q;X17R;X23V;X27L;X29G;X40R;X60I;X64V;X71P;X87L;X94A;X94P;X94S;X95M;X96L;X96Y;X105G;X113I;X122A;X127R;X131S;X144V;X145L;X147I;X147L;X147Q;X150Y;X152G;X153G;X157C;X173L;X195A;X195G;X196M;X198S;X199H;X206F;X206L;X208R;X216R;X221S;X243S;X245I;X249F;X249G;和X249Y。因此,在一些实施方案中,多核苷酸编码具有与以上描述的参考多肽的任何指定序列同一性并且包括指定残基差异,包括残基差异的组的酮还原酶多肽,如本公开内容中提供的。

[0256] 在一些实施方案中,多核苷酸编码能够以多肽SEQ ID NO:4的活性的至少1.5倍将底物化合物(2)转化为产物化合物(1)的酮还原酶多肽,其中多肽包括这样的氨基酸序列,该氨基酸序列具有与参考序列SEQ ID NO:2或4的至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多的同一性和特征X40R和X206L或F,条件是当X206为F时,该序列还包括与SEQ ID NO:4相比选自以下的一个或多个残基差异:X7S;X17M;X17Q;X17R;X23V;X27L;X29G;X60I;X64V;X71P;X87L;X94A;X94P;X94S;X95M;X96L;X96Y;X105G;X113I;X122A;X127R;X131S;X144V;X145L;X147I;X147L;X147Q;X150Y;X152G;X153G;X157C;X173L;X195A;X195G;X196M;X198S;X199H;X208R;X216R;X221S;X243S;X245I;X249F;X249G;和X249Y;特别是与SEQ ID NO:4相比选自以下的一个或多个残基差异:X7S;X17M;X17Q;X17R;X23V;X27L;X29G;X64V;X71P;X87L;X94A;X94P;X94S;X95M;X96L;X96Y;X105G;X113I;X122A;X127R;X131S;X147I;X147L;X147Q;X150Y;X152G;X153G;X157C;X173L;X195A;X195G;X196M;X198S;X199H;X208R;X216R;X221S;X243S;X249F;X249G;和X249Y。

[0257] 在一些实施方案中,多核苷酸编码包括这样的氨基酸序列的酮还原酶多肽,该氨基酸序列具有至少与SEQ ID NO:4相比选自以下的残基差异组合:(a)X40R、X196M、和X206F/L;(b)X40R、X144V、和X206F/L;(c)X40R、X17H/R/Q、和X206F/L;(d)X40R、X94P、和X206F/L;(e)X40R、X198S、和X206F/L;(f)X40R、X17H/R/Q、X94P和X206F/L;(g)X40R、X71P、X157C和X206F/L;(h)X40R、X94P、X144V、X196M和X206F/L;(i)X17H/R/Q、X40R、X64V、X147I/Q/L、和X206F/L;(j)X17H/R/Q、X40R、X64V、X94P、X144V、X147I/Q/L、X157C、X196M和X206F/L;以及(k)X17Q、X40R、X64V、X71P、X94P、X144V、X147I、X157C、X196M、X206F、和X249F。

[0258] 在一些实施方案中,多核苷酸编码能够形成非对映体过量于化合物(1b)和(1d)的产物化合物(1a)和(1c)的酮还原酶多肽。在一些实施方案中,多核苷酸编码相比于对产物化合物(1b)和(1d)具有对产物化合物(1a)和(1c)的非对映选择性的酮还原酶多肽,其中酮还原酶多肽包括这样的氨基酸序列,该氨基酸序列具有特征X40R和X206F或L,以及与SEQ ID NO:4相比选自以下的一个或多个残基差异:X17Q;X64V;X71P;X87L;X147I;X157C;

X196M;和X249F。

[0259] 在一些实施方案中,多核苷酸编码相比于对化合物(1b)和(1d)具有对产物化合物(1a)和(1c)的非对映选择性的酮还原酶多肽,其中酮还原酶多肽包括选自以下的氨基酸序列:SEQ ID NO:18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172。

[0260] 在一些实施方案中,多核苷酸编码能够将底物化合物(2)转化为产物化合物(1),具有相对于SEQ ID NO:4的参考多肽的活性增加了至少约1.2倍、1.5倍、2倍、3倍、4倍、5倍、10倍、20倍、30倍、40倍、50倍、100倍、200倍、500倍、1000倍、5000倍、10000倍或更多倍的活性的工程化多肽,其中多肽包括这样的氨基酸序列,该氨基酸序列具有与选自以下的任何一个的参考多肽的至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、或99%同一性:SEQ ID NO:4、6、8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172,条件是氨基酸序列包括与SEQ ID NO:4相比在以下的任何一个的多肽序列中:SEQ ID NO:4、6、8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、38、40、42、44、46、48、50、52、54、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172,并且特别是以下的多肽序列中含有的残基差异的组的任何一个:SEQ ID NO:18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172,如在表2A、2B和2C中列出的。如本文中所讨论的,在一些实施方案中,参考多肽选自SEQ ID NO:4、14、36、42、和130。

[0261] 在一些实施方案中,编码酮还原酶多肽的多核苷酸选自SEQ ID NO:3、5、7、9、11、13、15、17、19、21、23、25、27、29、31、33、35、37、39、41、43、55、57、59、61、63、65、67、69、71、73、75、77、79、81、83、85、87、89、91、93、95、97、99、101、103、105、107、109、111、113、115、117、119、121、123、125、127、129、131、133、135、137、139、141、143、145、147、149、151、153、155、157、159、161、163、165、167、169和171,特别是选自以下的多核苷酸:SEQ ID NO:17、19、21、23、25、27、29、31、33、35、37、39、41、43、55、57、59、61、63、65、67、69、71、73、75、77、79、81、83、85、87、89、91、93、95、97、99、101、103、105、107、109、111、113、115、117、119、121、123、125、127、129、131、133、135、137、139、141、143、145、147、149、151、153、155、157、159、161、163、165、167、169和171。

[0262] 在一些实施方案中,多核苷酸能够在高度严格的条件下与选自以下的多核苷酸:3、5、7、9、11、13、15、17、19、21、23、25、27、29、31、33、35、37、39、41、43、55、57、59、61、63、65、67、69、71、73、75、77、79、81、83、85、87、89、91、93、95、97、99、101、103、105、107、109、111、

113、115、117、119、121、123、125、127、129、131、133、135、137、139、141、143、145、147、149、151、153、155、157、159、161、163、165、167、169和171,或其互补物;特别是选自以下的多核苷酸杂交:SEQ ID NO:17、19、21、23、25、27、29、31、33、35、37、39、41、43、55、57、59、61、63、65、67、69、71、73、75、77、79、81、83、85、87、89、91、93、95、97、99、101、103、105、107、109、111、113、115、117、119、121、123、125、127、129、131、133、135、137、139、141、143、145、147、149、151、153、155、157、159、161、163、165、167、169和171,或其互补物,其中高度严格杂交多核苷酸编码能够以相对于SEQ ID NO:4的多肽的活性的增加了至少1.2倍、1.5倍、2倍、3倍、4倍、5倍、10倍、20倍、30倍、40倍、50倍、100倍、200倍、500倍、1000倍、5000倍、10000倍或更多倍的活性,将化合物(2)转化为化合物(1)的非天然存在的多肽。

[0263] 在一些实施方案中,多核苷酸能够在高度严格的条件下与编码包括SEQ ID NO:2或4的氨基酸序列的多肽的多核苷酸杂交,该氨基酸序列具有特征X40R和X206L或F,条件是当X206L为F时,该序列还包括与SEQ ID NO:4相比选自以下的一个或多个残基差异:X7S;X17M;X17Q;X17R;X23V;X27L;X29G;X60I;X64V;X71P;X87L;X94A;X94P;X94S;X95M;X96L;X96Y;X105G;X113I;X122A;X127R;X131S;X144V;X145L;X147I;X147L;X147Q;X150Y;X152G;X153G;X157C;X173L;X195A;X195G;X196M;X198S;X199H;X208R;X216R;X221S;X243S;X245I;X249F;X249G;和X249Y,特别是与SEQ ID NO:4相比选自以下的一个或多个残基差异:X7S;X17M;X17Q;X17R;X23V;X27L;X29G;X64V;X71P;X87L;X94A;X94P;X94S;X95M;X96L;X96Y;X105G;X113I;X122A;X127R;X131S;X147I;X147L;X147Q;X150Y;X152G;X153G;X157C;X173L;X195A;X195G;X196M;X198S;X199H;X208R;X216R;X221S;X243S;X249F;X249G;和X249Y。

[0264] 在一些实施方案中,多核苷酸编码本文描述的多肽,但具有在核苷酸水平上与编码本文描述的工程化酮还原酶多肽的参考多核苷酸的约80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、或99%或更多的序列同一性。在一些实施方案中,参考多核苷酸选自:SEQ ID NO:3、5、7、9、11、13、15、17、19、21、23、25、27、29、31、33、35、37、39、41、43、55、57、59、61、63、65、67、69、71、73、75、77、79、81、83、85、87、89、91、93、95、97、99、101、103、105、107、109、111、113、115、117、119、121、123、125、127、129、131、133、135、137、139、141、143、145、147、149、151、153、155、157、159、161、163、165、167、169和171;特别是选自以下的多核苷酸:SEQ ID NO:17、19、21、23、25、27、29、31、33、35、37、39、41、43、55、57、59、61、63、65、67、69、71、73、75、77、79、81、83、85、87、89、91、93、95、97、99、101、103、105、107、109、111、113、115、117、119、121、123、125、127、129、131、133、135、137、139、141、143、145、147、149、151、153、155、157、159、161、163、165、167、169和171。

[0265] 编码本文公开的非天然存在的多肽的分离的多核苷酸可以多种方式操作以提供多肽的表达。在一些实施方案中,编码多肽的多核苷酸可以被提供为表达载体,其中存在一个或多个控制序列以调控多核苷酸和/或多肽的表达。取决于表达载体,所分离的多核苷酸在其插入载体中之前的操作可能是令人期望的或必要的。利用重组DNA方法修饰多核苷酸和核酸序列的技术是本领域熟知的。在Sambrook等人,2001,Molecular Cloning:A Laboratory Manual,第3版,Cold Spring Harbor Laboratory Press;以及Current Protocols in Molecular Biology,Ausubel.F.编著,Greenlee Pub.Associates,1998,更新至2010中提供了指导。

[0266] 在一些实施方案中,控制序列包括启动子、前导序列、多腺苷酸化序列、前肽序列、信号肽序列、和转录终止子以及其他。适合的启动子可以基于使用的宿主细胞来选择。示例性细菌启动子包括大肠杆菌的乳糖操纵子、大肠杆菌的色氨酸操纵子、噬菌体 λ 、天蓝色链霉菌(*Streptomyces coelicolor*)琼脂酶基因(*dagA*)、枯草芽孢杆菌(*Bacillus subtilis*)果聚糖蔗糖酶基因(*sacB*)、地衣芽孢杆菌(*Bacillus licheniformis*) α -淀粉酶基因(*amyL*)、 β -内酰胺酶基因、和*tac*启动子;用于丝状真菌宿主细胞的示例性启动子,包括获自以下基因的启动子:米曲霉(*Aspergillus oryzae*)TAKA淀粉酶、米黑根毛霉(*Rhizomucor miehei*)天冬氨酸蛋白酶、黑曲霉(*Aspergillus niger*)中性 α -淀粉酶、黑曲霉酸稳定型 α -淀粉酶、黑曲霉或泡盛曲霉(*Aspergillus awamori*)葡糖淀粉酶(*glaA*)、米黑根毛霉脂肪酶、米曲霉碱性蛋白酶、米曲霉磷酸丙糖异构酶、构巢曲霉(*Aspergillus nidulans*)乙酰胺酶和尖孢镰刀菌(*Fusarium oxysporum*)胰蛋白酶样蛋白酶,以及其突变体、截短的、和杂合启动子,以及示例性酵母细胞启动子可以来自以下基因:酿酒酵母(*Saccharomyces cerevisiae*)烯醇化酶(*ENO-1*)、酿酒酵母半乳糖激酶(*GAL1*)、酿酒酵母乙醇脱氢酶/甘油醛-3-磷酸脱氢酶(*ADH2/GAP*)和酿酒酵母3-磷酸甘油酸激酶。用于酵母宿主细胞的其它有用的启动子由Romanos等人,1992,*Yeast* 8:423-488描述。

[0267] 控制序列还可以是适当的转录终止子序列,转录终止子序列是由宿主细胞识别以终止转录的序列。终止子序列可操作地连接到编码多肽的核酸序列的3'末端。在选择的宿主细胞中有功能的任何终止子可以用于本公开内容。例如,用于丝状真菌宿主细胞的示例性转录终止子可以获自以下的基因:米曲霉TAKA淀粉酶、黑曲霉葡萄糖淀粉酶、构巢曲霉邻氨基苯甲酸合酶、黑曲霉 α -葡糖苷酶和尖孢镰刀菌胰蛋白酶样蛋白酶。用于酵母宿主细胞的示例性终止子可以获自以下的基因:酿酒酵母烯醇化酶、酿酒酵母细胞色素C(*CYC1*)和酿酒酵母甘油醛-3-磷酸脱氢酶。用于酵母宿主细胞的其它有用的终止子由上述Romanos等人,1992描述。

[0268] 在一些实施方案中,控制序列还可以是信号肽编码区,其编码与多肽的氨基末端连接的氨基酸序列并将编码的多肽引导入细胞的分泌通路。信号序列通常取决于用于表达多肽的宿主细胞的类型。用于细菌宿主细胞的有效信号肽编码区是获自以下的基因的信号肽编码区:芽孢杆菌(*Bacillus*)NC1B 11837麦芽糖淀粉酶、嗜热脂肪芽孢杆菌 α -淀粉酶、地衣芽孢杆菌枯草蛋白酶、地衣芽孢杆菌 β -内酰胺酶、嗜热脂肪芽孢杆菌中性蛋白酶(*nprT*、*nprS*、*nprM*)和枯草芽孢杆菌*prsa*。用于丝状真菌宿主细胞的示例性信号肽编码区可以为获自以下的基因的信号肽编码区:米曲霉TAKA淀粉酶、黑曲霉中性淀粉酶、黑曲霉葡萄糖淀粉酶、米黑根毛霉天冬氨酸蛋白酶、特异腐质霉(*Humicola insolens*)纤维素酶和柔毛腐质霉(*Humicola lanuginosa*)脂肪酶。用于酵母宿主细胞的有用的信号肽可以来自酿酒酵母 α 因子和酿酒酵母转化酶的基因。

[0269] 加入调节序列还可以是期望的,调节序列允许关于宿主细胞的生长调节多肽的表达。调节系统的实例是引起响应化学或物理刺激物(包括调节化合物的存在)而开启或关闭基因的表达式的那些调节系统。在原核宿主细胞中,适当的调节序列包括*lac*、*tac*和*trp*操纵子系统。在酵母宿主细胞中,适当的调节系统包括,作为实例的*ADH2*系统或*GAL1*系统。在丝状真菌中,适当的调节序列包括TAKA α -淀粉酶启动子、黑曲霉葡萄糖淀粉酶启动子和米曲霉葡萄糖淀粉酶启动子。调节序列的其它实例是允许基因扩增的那些。在真核系统中,这些

包括二氢叶酸还原酶基因,其在氨甲喋呤的存在下被扩增;和金属硫蛋白基因,其用重金属扩增。在这些情况中,编码本公开内容的多肽的核酸序列将与调节序列可操作地连接。

[0270] 其他控制序列,诸如前导序列、聚腺苷酸化序列和转录终止子序列,可以使用本领域中可用的那些(参见Sambrook,同上,和Current Protocols in Molecular Biology,同上)。

[0271] 在另一方面,本公开内容还涉及重组表达载体,取决于它们被引入的宿主的类型,重组表达载体包括编码工程化酮还原酶多肽的多核苷酸,和一种或多种表达调节区诸如启动子和终止子,复制起点等。重组表达载体可以是任何载体(例如,质粒或病毒),其可以方便地经历重组DNA程序并且可以引起多核苷酸序列的表达。载体的选择将通常取决于载体与载体将被引入的宿主细胞的相容性。载体可以是线性的或闭合的环状质粒。

[0272] 表达载体可以是自主复制载体,即,作为染色体外的实体存在的载体,其复制独立于染色体复制,自主复制载体例如质粒、染色体外元件、微型染色体或人工染色体。载体可以包含用于保证自身复制的任何工具(means)。可选地,载体可以是当被引入宿主细胞时,被整合进入基因组并与其被整合进入的染色体一起复制的载体。表达载体优选地含有一个或多个可选择的标记物,其使得容易选择转化细胞。可选择的标记物是其产物提供杀生物剂或病毒抗性,对重金属的抗性,营养缺陷型的原养型对化学试剂的抗性(例如抗生素)以及类似性质的基因。

[0273] 在另一个方面,本公开内容提供了包括编码本公开内容的工程化酮还原酶多肽的多核苷酸的宿主细胞,该多核苷酸与用于在该宿主细胞中表达酮还原酶多肽的一个或多个控制序列可操作地连接。在表达由本发明的表达载体编码的酮还原酶多肽中使用的宿主细胞是本领域熟知的并且包括但不限于:细菌细胞,诸如大肠杆菌、乳杆菌(*Lactobacillus*)、链霉菌属(*Streptomyces*)和鼠伤寒沙门氏菌(*Salmonella typhimurium*)的细胞;真菌细胞,诸如酵母细胞;昆虫细胞诸如果蝇S2细胞和夜蛾(*Spodoptera*)Sf9细胞;动物细胞诸如CHO、COS、BHK、293和Bowes黑色素瘤细胞;以及植物细胞。示例性宿主细胞是大肠杆菌BL21和W3110。

[0274] 用于上面描述的宿主细胞的适当培养基和生长条件是本领域熟知的。可以通过本领域已知的多种方法将用于表达酮还原酶的多核苷酸引入宿主细胞中(例如,电穿孔、生物射弹粒子轰击、脂质体介导的转染、氯化钙转染和原生质体融合)。

[0275] 在本文的实施方案中,非天然存在的工程化酮还原酶多肽和编码此类多肽的核苷酸,可以使用由本领域技术人员通常使用的方法来制备。如上所提及的,天然存在的氨基酸序列和编码在产生工程化酮还原酶中使用的克菲尔乳杆菌的酮还原酶的相应的多核苷酸是作为Genban登录号AAP94029.1;GI:33112056可获得的。在一些实施方案中,亲本多核苷酸序列是密码子优化的以增强特定的宿主细胞中酮还原酶的表达。

[0276] 工程化酮还原酶多肽可以通过使编码天然存在的酮还原酶的多核苷酸经历诱变和/或定向进化方法而获得(参见,例如,Stemmer,1994,Proc Natl Acad Sci USA 91:10747-10751;PCT公布号WO 95/22625,WO 97/0078,WO 97/35966,WO 98/27230,WO 00/42651,和WO 01/75767;美国专利号6,537,746;6,117,679;6,376,246;和6,586,182;和美国专利公布号20080220990A1和20090312196A1;这些中每一个据此通过引用并入本文)。

[0277] 可以使用的其它定向进化程序包括交错延伸过程(StEP)、体外重组(Zhao等人,

1998, Nat. Biotechnol. 16:258-261)、诱变PCR(Caldwell等人, 1994, PCR Methods Appl. 3: S136-S140)、和盒式诱变(Black等人, 1996, Proc Natl Acad Sci USA 93:3525-3529)以及其他。可用于本文目的的诱变和定向进化技术也被描述在以下文献中:例如, Ling等人, 1997, Anal. Biochem. 254(2):157-78; Dale等人, 1996, Methods Mol. Biol. 57:369-74; Smith, 1985, Ann. Rev. Genet. 19:423-462; Botstein等人, 1985, Science 229:1193-1201; Carter, 1986, Biochem. J. 237:1-7; Kramer等人, 1984, Cell, 38:879-887; Wells等人, 1985, Gene 34:315-323; Minshull等人, 1999, Curr Opin Chem Biol 3:284-290; Christians等人, 1999, Nature Biotech 17:259-264; Cramer等人, 1998, Nature 391:288-291; Cramer等人, 1997, Nature Biotech 15:436-438; Zhang等人, 1997, Proc Natl Acad Sci USA 94:45-4-4509; Cramer等人, 1996, Nature Biotech 14:315-319; Stemmer, 1994, Nature 370:389-391。所有出版物通过引用并入本文。

[0278] 在一些实施方案中, 诱变处理后获得的克隆被筛选用于具有期望的酶特性的工程化酮还原酶。使用标准技术, 诸如分离产物(例如, 通过HPLC或GC)和通过测量分离的底物和产物的UV吸光度检测和/或通过使用串联质谱(例如, MS/MS)检测产物, 可以进行测量来自表达文库的酮还原酶的酶活性。然后包括编码期望的工程化多肽的多核苷酸的克隆被分离, 测序以鉴定核苷酸序列变化(如果有的话), 并且用于在宿主细胞中表达酶。示例性测定提供在以下实施例中。

[0279] 当多肽的序列是已知的时, 编码酶的多核苷酸可以根据已知的合成方法, 例如由 Beaucage等人1981, Tet Lett 22:1859-69描述的经典的亚磷酰胺方法, 或由Matthes等人, 1984, EMBO J. 3:801-05描述的方法, 通过标准的固相方法被制备。在一些实施方案中, 多至约100个碱基的片段可以被单独地合成, 然后连接(例如, 通过酶或化学连接方法(chemical ligation method)或聚合酶介导的方法)以形成任何期望的连续序列。

[0280] 在一些实施方案中, 本公开内容还提供用于制备或制造能够将化合物(2)转化为化合物(1)的非天然存在的多肽的方法, 其中所述方法包括在适合表达多肽的培养条件下培养能够表达编码非天然存在的多肽的多核苷酸的宿主细胞。在一些实施方案中, 多肽可以在无细胞表达系统, 例如描述于以下中的那些中表达: Kudlicki等人, Cell Free Expression, 第1版, Landes Biosciences(2007)和Cell Free Protein Synthesis: Methods and Protocols, 第1版, Spirin等人, 编著, Wiley-VCH(2007), 所有这些都通过引用并入本文。在一些实施方案中, 用于制备或制造非天然存在的酮还原酶多肽的方法还包括分离多肽的步骤。非天然存在的多肽可在如以上描述的适当的细胞中表达, 并使用用于蛋白纯化的熟知技术中的任何一种或多种从宿主细胞、培养基、和/或表达培养基分离(或回收), 用于蛋白纯化的熟知技术包括溶菌酶处理、声处理、过滤、盐析、超离心和色谱法以及其他。用于分离酮还原酶多肽的色谱方法包括, 反相色谱法、高效液相色谱法、离子交换色谱法、凝胶电泳和亲和色谱法以及其他。

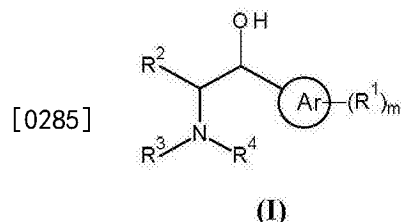
[0281] 在一些实施方案中, 本公开内容的非天然存在的多肽可以以各种形式来制备和使用, 各种形式包括, 但不限于: 粗提取物(例如, 无细胞溶解产物)、粉末(例如, 摇瓶粉)、冻干产物、和基本上纯的制剂(例如, DSP粉), 如在以下实施例中进一步阐述的。

[0282] 在一些实施方案中, 非天然存在的多肽可以以纯化的方式, 例如基本上纯化的方式来制备和使用。通常, 用于纯化具体酶的条件将部分地取决于诸如净电荷、疏水性、亲水

性、分子量、分子形状等的因素,并且对于本领域技术人员将是明显的。为了便于纯化,可以设想,在一些实施方案中,工程化酮还原酶多肽可以表达为具有纯化标签,诸如具有对金属的亲力的His-标签、或结合到抗体上的抗体标签,例如myc表位标签的融合蛋白。

[0283] 使用酮还原酶和产物化合物的方法

[0284] 在另一个方面,本文公开的工程化酮还原酶多肽可以在用于将底物化合物(2)或其结构类似物转化为化合物(1)的产物或相应的结构类似物的方法中使用。通常,化合物(1)的结构类似物包括在结构式(I)内,



[0286] 其中

[0287] Ar是5至7元碳环或杂环;

[0288] R^1 的每次出现独立地选自卤素、 $-\text{COOR}^a$ 、 $-\text{C}(\text{O})\text{R}^b$ 、 $-\text{OR}^b$ 、 $-\text{SO}_2$ 、 $-\text{SR}^c$ 、 $-\text{S}(\text{O})\text{R}^d$ 、 $-\text{NR}^e\text{R}^f$ 、 $-\text{C}(\text{O})\text{NR}^g\text{R}^h$ 、 $-\text{NO}_2$ 、 $-\text{CN}$ 、烷基、杂烷基、环烷基、环烷基烷基、杂环烷基、杂环烷基烷基、芳基、芳基烷基、杂芳基和杂芳基烷基,其中 R^a 、 R^b 、 R^c 、 R^d 、 R^e 、 R^f 、 R^g 和 R^h 独立地选自H和任选地取代的烷基、杂烷基、环烷基、环烷基烷基、杂环烷基、杂环烷基烷基、芳基、芳基烷基、杂芳基、和杂芳基烷基;

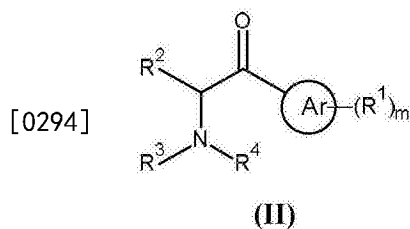
[0289] R^2 选自任选地取代的烷基、烯基、炔基、杂烷基、环烷基、环烷基烷基、环烷基烯基、环烷基炔基、杂环烷基、杂环烷基烷基、杂环烷基烯基、杂环烷基炔基、芳基、芳基烷基、芳基烯基、芳基炔基、杂芳基、杂芳基烷基、杂芳基烯基、杂芳基炔基、烷基硫代、烷基亚磺酰基和芳基亚磺酰基;

[0290] R^3 和 R^4 ,彼此独立地,选自H、任选地取代的烷基、环烷基、环烷基烷基、杂环烷基、杂环烷基烷基、芳基、芳基烷基、杂芳基、杂芳基烷基、烷氧基羰基、芳基烷氧基羰基、杂芳基烷氧基羰基、和保护基团,或 R^3 和 R^4 中的一个与 R^2 形成任选地取代的5至7元杂环,以及

[0291] m 为0至10。

[0292] 在一些实施方案中,可以进行用于将化合物(2)的底物或其结构类似物转化为化合物(1)的产物或相应的结构类似物的方法,其中化合物(2)的底物为化合物(2)的氘化形式(例如,具有与化合物(2)相同的结构但化合物(2)的一个或多个氢原子被氘原子取代的分子)。药物化合物的氘化形式的一些实例描述于例如美国专利号5,846,514、6,503,921、6,613,739、和7,705,036中。类似地,用于本文描述的结构类似物的转化的方法,包括以下描述的将各种式(II)的化合物转化为相应的式(I)的化合物,可以使用这些化合物的氘化形式进行。

[0293] 因此,在一些实施方案中,用于制备式(I)的化合物的方法包括使式(II)的化合物



[0295] 其中

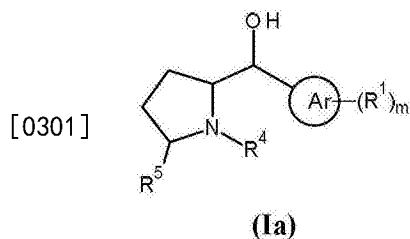
[0296] R^1 、 R^2 、 R^3 、 R^4 和 m 如以上定义，

[0297] 在适当的反应条件下在辅因子NADPH或NADH的存在下与本文公开的工程化酮还原酶多肽接触。

[0298] 在方法的一些实施方案中，Ar为5至7元芳基或杂芳基基团。在一些实施方案中，Ar选自噻吩基、苯基和吡啶基。

[0299] 在一些实施方案中， m 为1、2、3、或4。在一些实施方案中， R^1 为卤素，特别是溴或氟。在一些实施方案中，卤素基团为氟。

[0300] 在方法的一些实施方案中，式(I)的产物化合物包括式(Ia)的化合物，

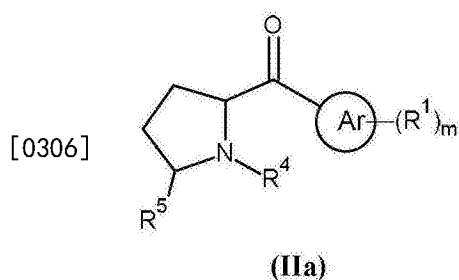


[0302] 其中

[0303] Ar、 R^1 、 R^4 和 m 如以上定义；

[0304] R^5 是任选取代的烷基、烯基、炔基、杂烷基、环烷基、环烷基烷基、环烷基烯基、环烷基炔基、杂环烷基、杂环烷基烷基、芳基、芳基烷基、杂芳基、和杂芳基烷基。

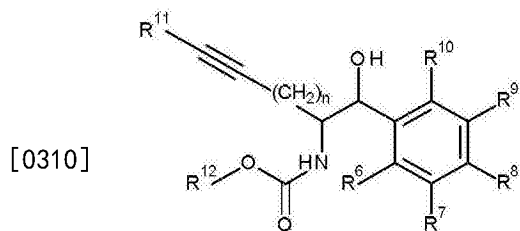
[0305] 因此，在一些实施方案中，用于制备式(Ia)的化合物的方法包括使式(IIa)的底物化合物，



[0307] 其中Ar、 R^1 、 R^4 、 R^5 和 m 如以上定义，

[0308] 在适当的反应条件下在辅因子NADPH或NADH的存在下与本文描述的工程化酮还原酶多肽接触。

[0309] 在方法的一些实施方案中，式(I)的产物化合物包括式(Ib)的化合物，



(Ib)

[0311] 其中

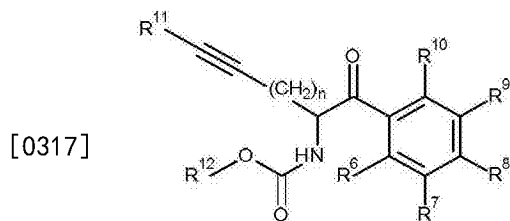
[0312] R^6 、 R^7 、 R^8 、 R^9 和 R^{10} 各自独立地选自H、卤素、 $-\text{COOR}^a$ 、 $-\text{C(O)}R^b$ 、 $-\text{OR}^c$ 、 $-\text{SO}_2$ 、 $-\text{SR}^d$ 、 $-\text{S(O)}R^e$ 、 $-\text{NR}^fR^g$ 、 $-\text{C(O)}\text{NR}^hR^i$ 、 $-\text{NO}_2$ 、 $-\text{CN}$ 、烷基、杂烷基、环烷基、环烷基烷基、杂环烷基、杂环烷基烷基、芳基、芳基烷基、杂芳基和杂芳基烷基，其中 R^a 、 R^b 、 R^c 、 R^d 、 R^e 、 R^f 、 R^g 、 R^h 和 R^i 独立地选自H和任选地取代的烷基、杂烷基、环烷基、环烷基烷基、杂环烷基、杂环烷基烷基、芳基、芳基烷基、杂芳基、和杂芳基烷基；

[0313] n 为1至4，

[0314] R^{11} 选自H和任选地取代的烷基、杂烷基、环烷基、环烷基烷基、芳基、芳基烷基、杂芳基、和杂芳基烷基；以及

[0315] R^{12} 选自H、任选地取代的烷基、环烷基、环烷基烷基、杂环烷基、杂环烷基烷基、芳基、芳基烷基、杂芳基、杂芳基烷基、和保护基团。

[0316] 因此，在一些实施方案中，用于制备式(Ib)的产物化合物的方法包括使式(IIb)的底物化合物，



(IIb)

[0318] 其中

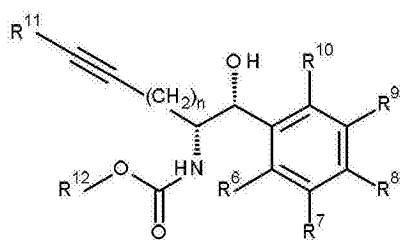
[0319] R^6 、 R^7 、 R^8 、 R^9 、 R^{10} 、 R^{11} 、 R^{12} 和 n 如以上定义，

[0320] 在适当的反应条件下在辅因子NADPH或NADH的存在下与本文描述的工程化酮还原酶多肽接触。

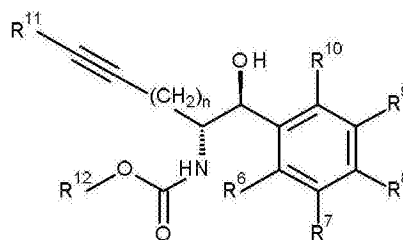
[0321] 在一些实施方案中，式(IIb)的化合物的 R^{11} 包括任选地取代的苯基。

[0322] 在一些实施方案中，式(IIb)的化合物的 n 包括2、3或4。

[0323] 在方法的一些实施方案中，式(Ib)的产物化合物包括式(Ib1)和式(Ib3)的化合物，并且方法形成非对映体过量于式(Ib2)和(Ib4)的化合物的式(Ib1)和(Ib3)的产物化合物，

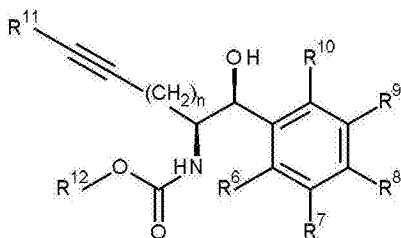


(Ib1)

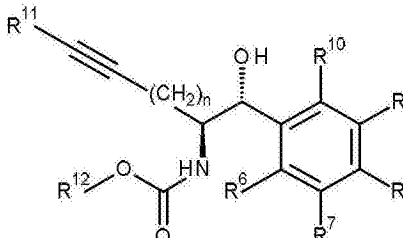


(Ib2)

[0324]



(Ib3)

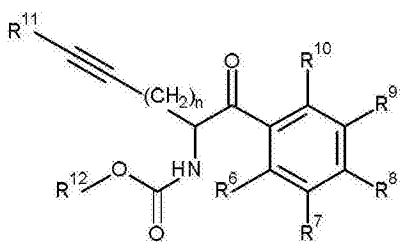


(Ib4)

[0325] 其中

[0326] R^6 、 R^7 、 R^8 、 R^9 、 R^{10} 、 R^{11} 、 R^{12} 和 n 如关于式(Ib)中所定义。

[0327] 因此,在一些实施方案中,用于制备非对映体过量于式(Ib2)和(Ib4)的化合物的式(Ib1)和(Ib3)的化合物的方法包括使式(IIb)的底物化合物



[0328]

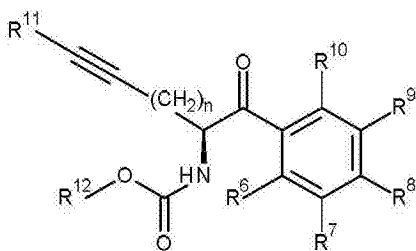
(IIb)

[0329] 其中

[0330] R^6 、 R^7 、 R^8 、 R^9 、 R^{10} 、 R^{11} 、 R^{12} 和 n 如关于式(Ib)中所定义;

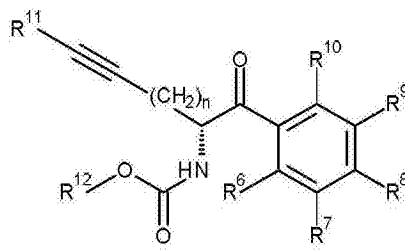
[0331] 在适当的反应条件下在辅因子NADPH或NADH的存在下与相比于对式(Ib2)和(Ib4)的化合物具有对式(Ib1)和(Ib3)的化合物的非对映选择性的工程化酮还原酶接触。

[0332] 在方法的一些实施方案中,式(IIb)的化合物包括式(IIb1)和式(IIb2)的化合物的混合物



[0333]

(IIb1)



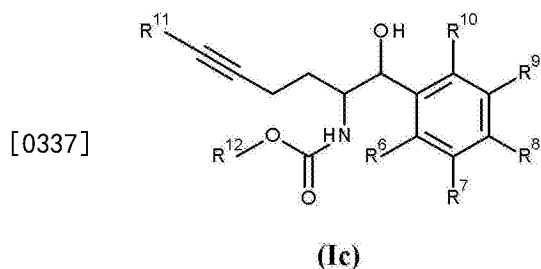
(IIb2)

[0334] 在方法的一些实施方案中,式(IIb)的化合物包括式(IIb1)和式(IIb2)的化合物

的外消旋混合物。

[0335] 在使用式(IIb)的化合物的方法的一些实施方案中,n选自2、3或4。在一些实施方案中,n为2。在一些实施方案中,n为2且R⁶、R⁷、R⁸、R⁹、R¹⁰和R¹¹为H。

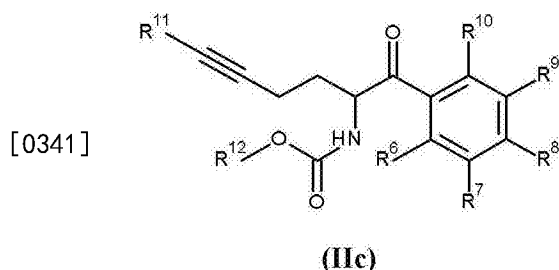
[0336] 在方法的一些实施方案中,式(Ib)的产物化合物包括式(Ic)的化合物



[0338] 其中

[0339] R⁶、R⁷、R⁸、R⁹、R¹⁰、R¹¹和R¹²如关于式(Ib)的化合物所定义。

[0340] 因此,在一些实施方案中,用于制备式(Ic)的产物化合物的方法包括使式(IIc)的底物化合物

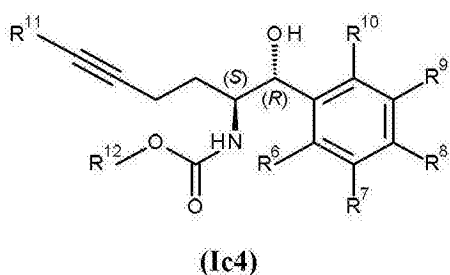
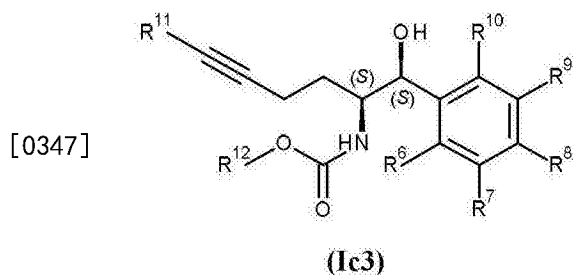
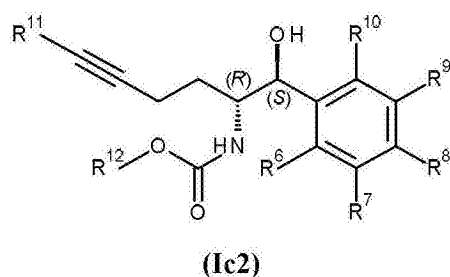
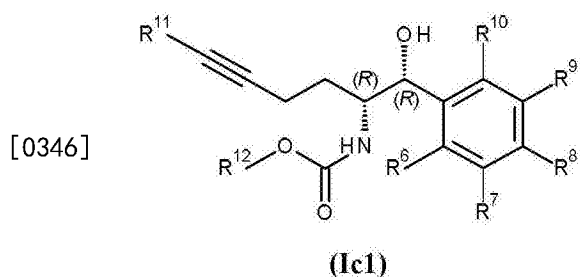


[0342] 其中

[0343] R⁶、R⁷、R⁸、R⁹、R¹⁰、R¹¹和R¹²如关于式(Ib)的化合物所定义

[0344] 在适当的反应条件下在辅因子NADPH或NADH的存在下与本文描述的工程化酮还原酶多肽接触。

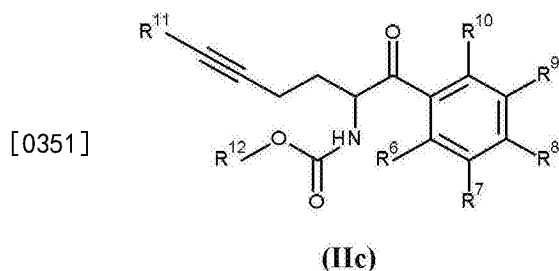
[0345] 在方法的一些实施方案中,式(Ib)的产物化合物包括式(Ic1)和(Ic3)的化合物,并且方法形成非对映体过量于式(Ic2)和(Ic4)的化合物的式(Ic1)和(Ic3)的产物化合物,



[0348] 其中

[0349] R^6 、 R^7 、 R^8 、 R^9 、 R^{10} 、 R^{11} 和 R^{12} 如关于式(Ib)的化合物所定义。

[0350] 因此,在一些实施方案中,用于制备非对映体过量于式(Ic2)和(Ic4)的化合物的式(Ic1)和(Ic3)的化合物的方法包括使式(IIc)的底物化合物

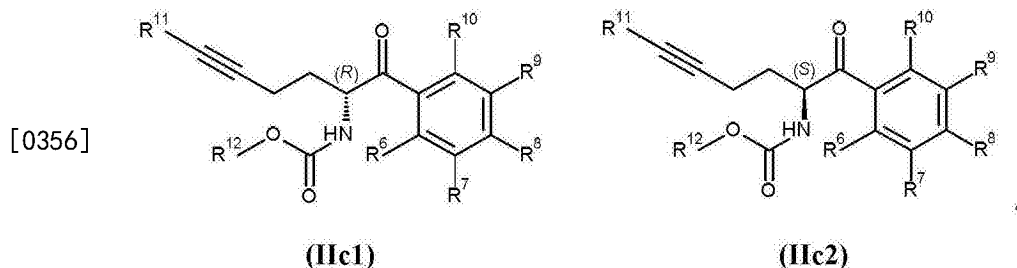


[0352] 其中

[0353] R^6 、 R^7 、 R^8 、 R^9 、 R^{10} 、 R^{11} 和 R^{12} 如关于式(Ib)的化合物所定义;

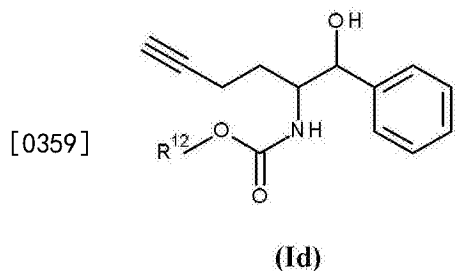
[0354] 在适当的反应条件下在辅因子NADPH或NADH的存在下与相比于对式(Ic2)和(Ic4)的化合物具有对式(Ic1)和(Ic3)的产物化合物的非对映选择性的工程化酮还原酶多肽接触。

[0355] 在方法的一些实施方案中,式(IIc)的化合物包括式(IIc1)和式(IIc2)的化合物的混合物,



[0357] 在方法的一些实施方案中,式(IIb)的化合物包括式(IIb1)和式(IIb2)的化合物的外消旋混合物。

[0358] 在方法的一些实施方案中,式(Ib)的产物化合物包括式(Id)的化合物,

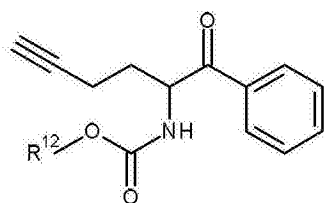


[0360] 其中

[0361] R^{12} 如关于式(Ib)的化合物所定义。

[0362] 因此,在一些实施方案中,用于制备式(Id)的产物化合物的方法包括使式(IIId)的底物化合物,

[0363]



(IIId)

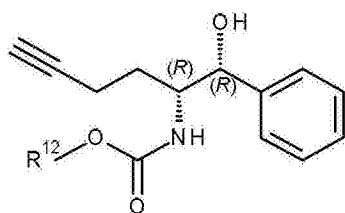
[0364] 其中

[0365] R¹²如关于式(Ib)的化合物所定义；

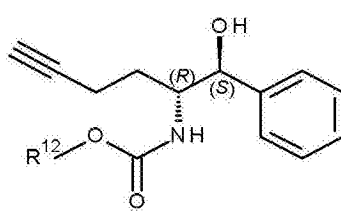
[0366] 在适当的反应条件下在辅因子NADPH或NADH的存在下与本文描述的工程化酮还原酶多肽接触。

[0367] 在方法的一些实施方案中,式(Ib)的产物化合物包括式(Id1)和(Id3)的化合物,并且方法形成非对映体过量于式(Id2)和(Id4)的化合物的式(Id1)和(Id3)的产物化合物,

[0368]

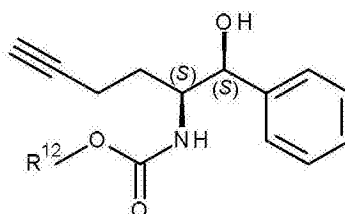


(Id1)

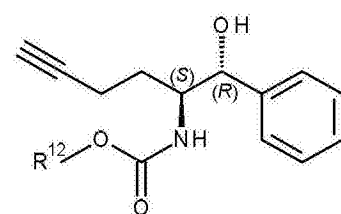


(Id2)

[0369]



(Id3)

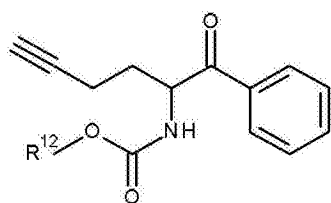


(Id4)

[0370] 其中R¹²如关于式(Ib)的化合物所定义。

[0371] 因此,在一些实施方案中,用于制备非对映体过量于式(Id2)和(Id4)的化合物的式(Id1)和(Id3)的产物化合物的方法包括使式(IIId)的底物化合物

[0372]



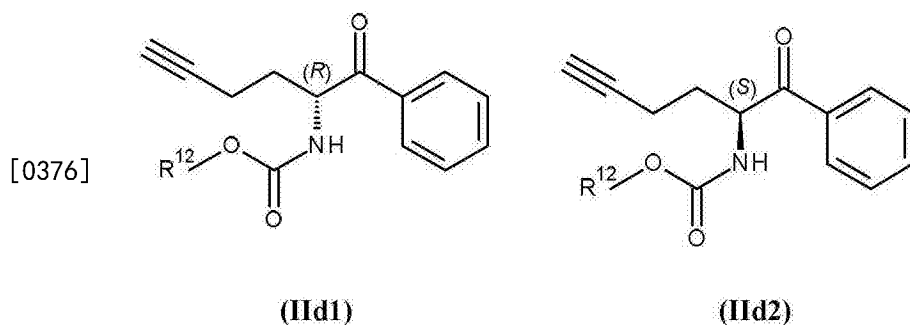
(IIId)

[0373] 其中R¹²如关于式(Ib)的化合物所定义；

[0374] 在适当的反应条件下在辅因子NADPH或NADH的存在下与相比于对式(Id2)和(Id4)的化合物具有对式(Id1)和(Id3)的产物化合物的非对映选择性的工程化酮还原酶多肽接触。

[0375] 在方法的一些实施方案中,式(IIId)的化合物包括式(IIId1)和(IIId2)的化合物的

混合物。



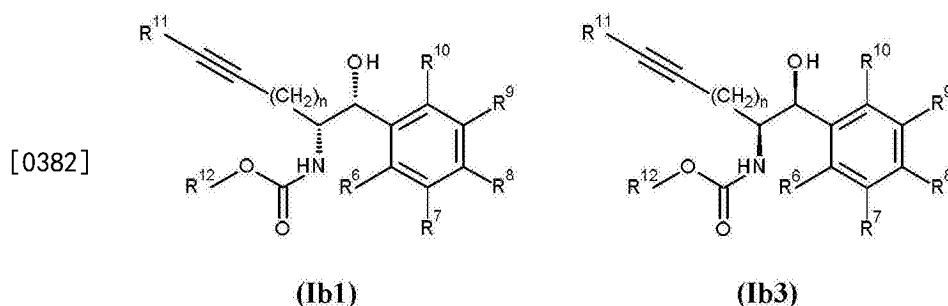
[0377] 在方法的一些实施方案中,式(IIId)的化合物包括式(IIId1)和(IIId2)的化合物的外消旋混合物。

[0378] 在用于以下的方法的一些实施方案中:制备非对映体过量于化合物(Ib2)和(Ib4)的式(Ib1)和(Ib3)的化合物、制备非对映体过量于化合物(Ic2)和(Ic4)的式(Ic1)和(Ic3)的化合物或制备非对映体过量于化合物(Id2)和(Id4)的式(Id1)和(Id3)的化合物,基团 R^{12} 选自甲基、乙基、正丙基、异丙基、丁基、异丁基、仲丁基、和叔丁基。在一些实施方案中, R^{12} 为叔丁基。

[0379] 在一些实施方案中,相比于对式(Ib2)和(Ib4)的产物化合物具有对式(Ib1)和(Ib3)的产物化合物的非对映选择性的示例性工程化酮还原酶包括这样的酮还原酶多肽,该酮还原酶多肽包括选自以下的氨基酸序列:SEQ ID NO:18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172。

[0380] 在以上方法的一些实施方案中,以相比于式(Id2)和(Id4)的化合物的至少50的非对映体比形成式(Id1)和(Id3)的产物化合物。在一些实施方案中,能够以相比于式(Id2)和(Id4)的化合物的至少50的非对映体比形成式(Id1)和(Id3)的产物化合物的示例性工程化酮还原酶包括这样的多肽,该多肽包括选自以下的氨基酸序列:SEQ ID NO:18、20、22、24、26、28、32、34、36、38、40、42和44。

[0381] 在方法的一些实施方案中,式(Ib)的产物化合物包括式(Ib1)的化合物,并且形成非对映体过量于式(Ib3)的化合物的式(Ib1)的产物化合物

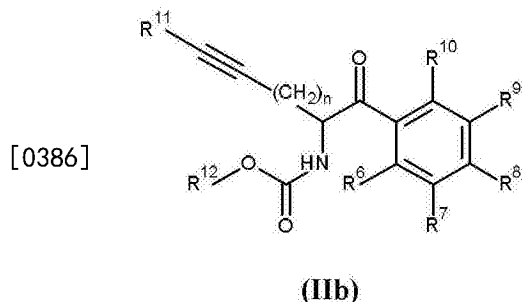


[0383] 其中

[0384] R^6 、 R^7 、 R^8 、 R^9 、 R^{10} 、 R^{11} 、 R^{12} 和 n 如以上提供的。在一些实施方案中, n 选自2、3或4。在一些实施方案中, n 为2。在一些实施方案中, n 为2且 R^6 、 R^7 、 R^8 、 R^9 、 R^{10} 和 R^{11} 为H。

[0385] 因此,在一些实施方案中,用于制备非对映体过量于式(Ib3)的化合物的式(Ib1)

的化合物的方法包括使式(IIb)的底物化合物

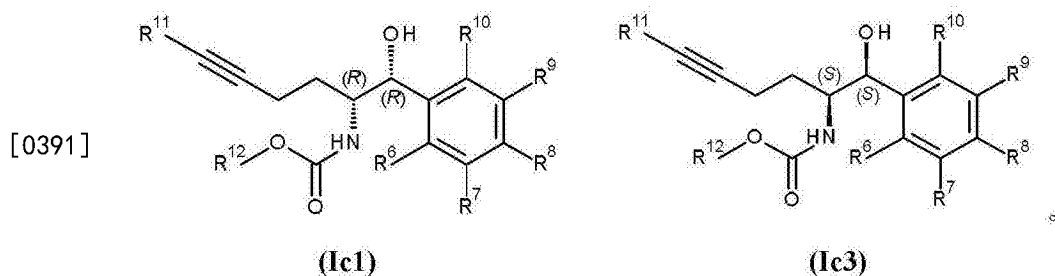


[0387] 在适当的反应条件下在辅因子NADPH或NADH的存在下与相比于对式(Ib3)的化合物具有对式(Ib1)的产物化合物的非对映选择性的工程化酮还原酶接触。

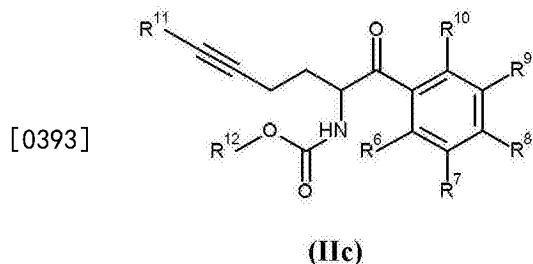
[0388] 在用于制备非对映体过量于式(Ib3)的化合物的式(Ib1)的化合物的方法的一些实施方案中,底物化合物(IIb)包括式(IIb1)和(IIb2)的化合物的混合物。

[0389] 在以上方法的一些实施方案中,底物化合物(IIb)包括式(IIb1)和(IIb2)的化合物的外消旋混合物。

[0390] 在方法的一些实施方案中,式(Ib1)的产物化合物包括式(Ic1)的化合物,并且形成非对映体过量于式(Ic3)的化合物的式(Ic1)的化合物



[0392] 因此,在一些实施方案中,用于制备非对映体过量于式(Ic3)的化合物的式(Ic1)的化合物的方法包括使式(IIc)的底物化合物



[0394] 其中

[0395] R^6 、 R^7 、 R^8 、 R^9 、 R^{10} 、 R^{11} 、和 R^{12} 如关于式(Ib)的化合物所定义;

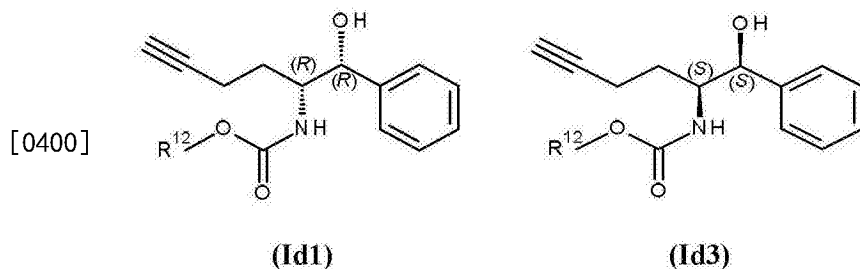
[0396] 在适当的反应条件下在辅因子NADPH或NADH的存在下与相比于对式(Ic3)的化合物具有对式(Ic1)的产物化合物的非对映选择性的工程化酮还原酶接触。

[0397] 在用于制备非对映体过量于式(Ic3)的化合物的式(Ic1)的化合物的方法的一些实施方案中,底物化合物(IIc)包括式(IIc1)和(IIc2)的化合物的混合物。

[0398] 在以上方法的一些实施方案中,底物化合物(IIc)包括式(IIc1)和(IIc2)的化合物的外消旋混合物。

[0399] 在方法的一些实施方案中,式(Ib)的产物化合物包括式(Id1)的化合物,并且形成

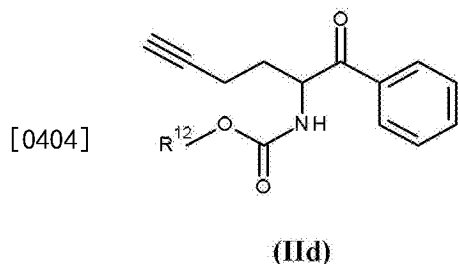
非对映体过量于式(Id3)的化合物的式(Id1)的化合物



[0401] 其中

[0402] R^{12} 如关于式(Ib)的化合物所定义。

[0403] 因此,在一些实施方案中,用于制备非对映体过量于式(Id3)的化合物的式(Id1)的化合物的方法包括使式(IIId)的底物化合物



[0405] 其中

[0406] R^{12} 如关于式(Ib)的化合物所定义;

[0407] 在适当的反应条件下在NADPH或NADH的存在下与相比于对式(Id3)的化合物具有对式(Id1)的产物化合物的非对映选择性的工程化酮还原酶接触。

[0408] 在用于制备非对映体过量于式(Id3)的化合物的式(Id1)的化合物的方法的一些实施方案中,式(IIId)的化合物包括式(IIId1)和(IIId2)的底物化合物的混合物。

[0409] 在方法的一些实施方案中,式(IIId)的化合物包括式(IIId1)和(IIId2)的底物化合物的外消旋混合物。

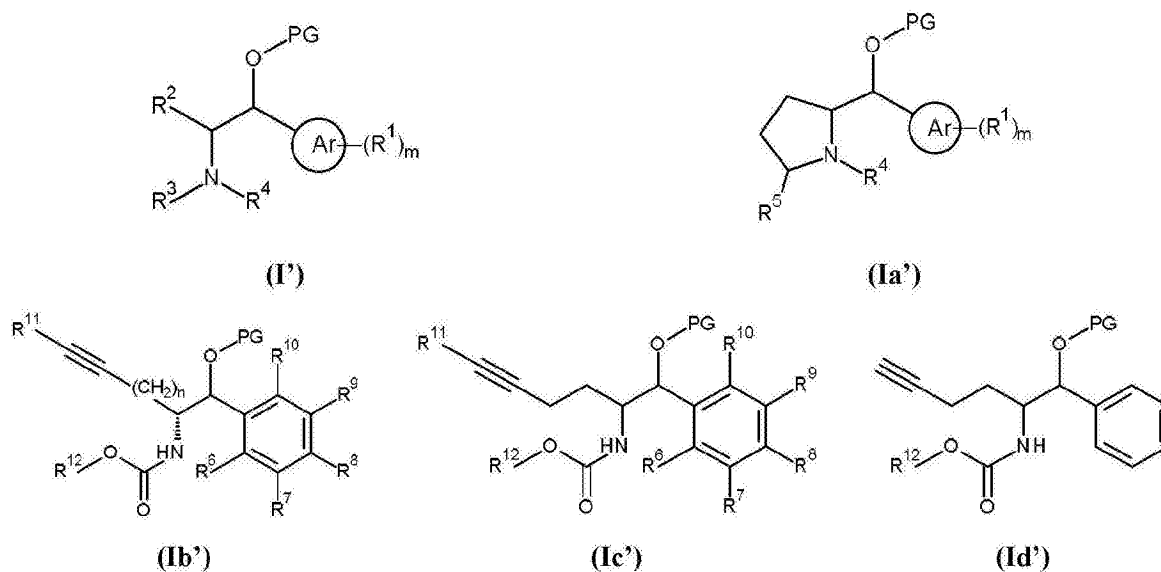
[0410] 在用于以下的方法的一些实施方案中:制备非对映体过量于式(Ib3)的化合物的式(Ib1)的化合物、制备非对映体过量于式(Ic3)的化合物的式(Ic1)的化合物、或制备非对映体过量于式(Id3)的化合物的式(Id1)的化合物,基团 R^{12} 选自甲基、乙基、正丙基、异丙基、丁基、异丁基、仲丁基、和叔丁基。在一些实施方案中, R^{12} 为叔丁基。

[0411] 在一些实施方案中,相比于对式(Ib2)和(Ib4)的化合物具有对式(Ib1)和(Ib3)的产物化合物、相比于对式(Ic2)和(Ic4)的化合物具有对式(Ic1)和(Ic3)的产物化合物、相比于对式(Id2)和(Id4)的化合物具有对式(Id1)和(Id3)的产物化合物的非对映选择性的示例性工程化酮还原酶包括这样的多肽,该多肽包括选自以下的氨基酸序列:18、20、22、24、26、28、30、32、34、36、38、40、42、44、56、58、60、62、64、66、68、70、72、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、136、138、140、142、144、146、148、150、152、154、156、158、160、162、164、166、168、170和172。

[0412] 在一些实施方案中,可以修饰酮还原酶反应的产物化合物以在药物化合物合成中提供其他有用的中间体。因此,在一些实施方案中,可以加工产物化合物的羟基基团以具有

保护基团。此类保护的产物化合物,包括以下式(I')、(Ia')、(Ib')、(Ic')、和(Id')的化合物:

[0413]



[0414] 其中

[0415] Ar、R¹、R²、R³、R⁴、R⁵、R⁶、R⁷、R⁸、R⁹、R¹⁰、R¹¹、和R¹²如先前所定义;且

[0416] PG为保护基团。

[0417] 因此,在一些实施方案中,用于制备化合物(I')、(Ia')、(Ib')、(Ic')、或(Id')的方法可包括分别用保护基团修饰产物化合物(I)、(Ia)、(Ib)、(Ic)、或(Id)的步骤。

[0418] 在以上的实施方案中,其中取代基包括保护基团,该保护基团可以是相关官能团的任何适合的保护基团。例如,胺保护基团可选自叔丁氧羰基(Boc)、9-芴甲氧羰基(Fmoc)、三氯乙基氯甲酸酯基(Troc)、对甲氧基苄基羰基(Moz)、3,4-二甲氧基苄基(DMPM)、对甲氧基苄基(Pmb)、甲苯磺酰基(Ts)和苄氧羰基(Cbz)以及其他。羧基保护基团可选自烷基酯(例如,叔丁基酯)、芳基烷基酯(例如,苄酯)、甲硅烷基酯、和噁唑啉以及其他。羟基保护基团可选自烷氧基(例如,甲基醚)、烷氧基烷基(例如,甲氧基甲基醚:MOM)、甲氧基乙氧基甲基醚(MEM)、甲基硫甲基醚(MTM)、苄氧基甲基醚(BOM)、四氢吡喃基醚(THP)、对甲氧基苄基醚(PMB)甲基、三苯甲基醚、甲氧基三苯甲基醚和甲硅烷基醚(例如,三甲基甲硅烷基醚:TMS)以及其他。其他适合的保护基团描述于Wuts和Greene中,“Greene's Protective Groups in Organic Synthesis,”第4版,Wiley Interscience(2006),通过引用并入本文。

[0419] 在本文和实施例中所示的实施方案中,本公开内容涵盖可以在方法中使用的适当的反应条件的范围,包括但不限于:pH、温度、缓冲液、溶剂系统、底物载量、底物化合物立体异构体的混合物、多肽载量、辅因子载量、压力和反应时间的范围。用于使用本文所述的工程化酮还原酶多肽进行用于将底物化合物生物催化转化为产物化合物的方法的另外的适当的反应条件可易于通过常规实验来优化,常规实验包括但不限于:使工程化酮还原酶多肽与底物化合物在浓度、pH、温度、溶剂条件的实验反应条件下接触,并且例如使用本文提供的实施例中描述的方法检测产物化合物。

[0420] 考虑到例如,产物化合物的需要量、底物浓度对酶活性的影响、酶在反应条件下的稳定性、以及底物至产物的百分比转化率,反应混合物中的底物化合物可以变化。在方法的

一些实施方案中,适当的反应条件包括以下的底物化合物载量浓度:至少约1至约200g/L、约10g/L至约200g/L、约20g/L至约200g/L、约40g/L至约200g/L、约50g/L至约200g/L、约75g/L至约200g/L、约100g/L至约200g/L、约125g/L至约200g/L、约150g/L至约200g/L或约75g/L至约150g/L。在一些实施方案中,适当的反应条件包括以下的底物化合物载量浓度:至少约1g/L、约10g/L、约20g/L、约40g/L、约50g/L、约75g/L、约100g/L、约125g/L、约150g/L、约175g/L、或约200g/L、或甚至更大。本文中提供的底物载量的值是基于化合物(2)的分子量;然而还设想,等摩尔量的化合物(2)的各种水合物和盐还可以在方法中被使用。此外,由式(II)、(IIa)、(IIb)、(IIc)、和(IId)的化合物涵盖的底物化合物,还可以根据针对化合物(2)使用的量以及相关溶剂的条件以适当的量被使用。

[0421] 本公开内容的工程化酮还原酶多肽的改进的活性和/或立体选择性提供其中可以以工程化多肽的较低浓度实现较高百分比转化率的方法。其还降低残留蛋白的量,该残留蛋白可能需要在用于纯化产物化合物(例如,化合物(1))的随后步骤中被除去。在方法的一些实施方案中,适当的反应条件包括以下的工程化酮还原酶多肽浓度:约0.1至约20g/L、约0.1至约10g/L、约0.5至约10g/L、约1.0至约10g/L、约0.1至约5g/L、约0.1至约2g/L、约0.1至约1g/L。在方法的一些实施方案中,适当的反应条件包括以下的工程化酮还原酶多肽浓度:约20g/L或更少、约15g/L或更少、约10g/L或更少、约5g/L或更少、约4g/L或更少、约2g/L或更少、约1g/L或更少、约0.5g/L或更少、约0.2g/L或更少、或约0.1g/L。

[0422] 在酮还原酶反应的过程中,反应混合物的pH可以变化。反应混合物的pH可以被维持在期望的pH或在期望的pH范围内。这可以通过在反应的过程之前和/或期间加入酸或碱来实现。可选地,pH可以通过使用缓冲液来控制。因此,在一些实施方案中,反应条件包括缓冲液。用于维持期望的pH范围的适当的缓冲液是本领域已知的且包括,通过举例的方式而非限制,硼酸盐缓冲液、碳酸盐缓冲液、磷酸盐缓冲液、三乙醇胺缓冲液和类似缓冲液。在一些实施方案中,缓冲液是硼酸盐。在一些实施方案中,适当的反应条件包括硼酸盐的缓冲溶液,其中硼酸盐浓度是从约0.01至约0.4M、约0.05至约0.4M、约0.1至约0.3M、或约0.1至约0.2M。在一些实施方案中,反应条件包括以下的硼酸盐浓度:约0.01、约0.02、约0.03、约0.04、约0.05、约0.07、约0.1、约0.12、约0.14、约0.16、约0.18、约0.2、约0.3、或约0.4M。在一些实施方案中,反应条件包括作为适当的溶剂的水,没有缓冲液存在。

[0423] 在方法的一些实施方案中,反应条件可以包括适当的pH。如上所提及的,期望的pH或期望的pH范围可以通过使用酸或碱、适当的缓冲液、或缓冲液和酸或碱添加的组合来维持。反应混合物的pH可以在反应的过程之前和/或期间被控制。在一些实施方案中,适当的反应条件包括约6至约12的pH、约6至约11的pH、约7至约11的pH、或约7至约10的pH、约8至约10的pH、约9至约10的pH、或约8至约9.5的pH、或约8至约9的pH的溶液pH。在一些实施方案中,反应条件包括约6、约6.5的pH、约7的pH、约7.5的pH、约8的pH、约8.5的pH、约9的pH、约9.5的pH、约10的pH、约10.5的pH、约11的pH、约11.5的pH、或约12的pH的溶液pH。本领域普通技术人员可以通过考虑,通过示例的方式而非限制,酮还原酶的稳定性和活性、底物和产物的稳定性、和辅因子的稳定性,使用适当的pH或pH范围。

[0424] 在本文方法的实施方案中,例如,考虑到:在较高的温度下反应速率的增加、对于反应中的酶的活性的增加、以及如下文进一步描述的,用于动态动力学拆分反应的底物非对映体的平衡(例如,消旋)的增加速率,适当的温度可以用于反应条件。例如,本公开内容

的工程化多肽具有相对于天然存在的酮还原酶多肽例如SEQ ID NO:2的野生型多肽以及在一些实施方案中SEQ ID NO:32的工程化酮还原酶的增加的稳定性,其允许工程化多肽以较高的温度用于反应的增加的转化率和改进的底物可溶性特性。因此,在一些实施方案中,适当的反应条件包括以下的温度:约20℃至约60℃、约25℃至约60℃、约30℃至约60℃、约35℃至约60℃、约40℃至约60℃、约45℃至约60℃、约45℃至约55℃、或约50℃至约55℃。在一些实施方案中,适当的反应条件可包括环境温度(例如,25℃)或以下的温度:约27℃、约30℃、约32℃、约35℃、约37℃、约40℃、约45℃、约50℃、约55℃或约60℃的温度。在一些实施方案中,酶促反应过程中的温度可以在整个反应过程中保持在一个温度。在一些实施方案中,酶促反应过程中的温度可以在反应过程中按温度曲线调节。

[0425] 方法通常在溶剂中进行。适当的溶剂包括水、缓冲水溶液、有机溶剂、和/或共溶剂系统,共溶剂系统通常包括水性溶剂和有机溶剂。适当的溶剂包括水、缓冲水溶液、有机溶剂、和/或共溶剂系统,共溶剂系统通常包括水性溶剂和有机溶剂。水性溶剂(水或水性的共溶剂系统)可以是pH缓冲的或无缓冲的。在一些实施方案中,使用工程化酮还原酶多肽的方法通常在包括有机溶剂(例如,乙醇、异丙醇(IPA)、二甲基亚砜(DMSO)、乙酸乙酯、乙酸丁酯、1-辛醇烷、庚烷、辛烷、甲基叔丁基醚(MTBE)、甲苯以及类似物)、离子液体(例如,1-乙基-4-甲基咪唑鎓四氟硼酸盐、1-丁基-3-甲基咪唑鎓四氟硼酸盐、1-丁基-3-甲基咪唑鎓六氟磷酸盐以及类似物)的水性共溶剂系统中进行。水性共溶剂系统的有机溶剂组分可以是与水性组分易混溶的,提供单一的液相,或可以是与水性组分部分地易混溶的或不易混溶的,提供两个液相。示例性的水性共溶剂系统包括水和一种或多种有机溶剂。大体上,水性共溶剂系统的有机溶剂组分被选择为使得其不使酮还原酶完全失活。适当的共溶剂系统可以通过使用酶活性测定,如本文描述的那些,用确定的感兴趣的底物在候选溶剂系统中测量特定的工程化酮还原酶的酶促活性来容易地确定。

[0426] 在方法的一些实施方案中,适当的反应条件包括水性共溶剂,该水性共溶剂包括以下浓度的DMSO:约1%至约50%(v/v)、约1至约40%(v/v)、约2%至约40%(v/v)、约5%至约30%(v/v)、约10%至约30%(v/v)、或约10%至约20%(v/v)。在方法的一些实施方案中,适当的反应条件包括水性共溶剂,该水性共溶剂包括至少约1%、约5%、约10%、约15%、约20%、约25%、约30%、约35%、约40%、约45%、或约50%(v/v)的浓度的DMSO。

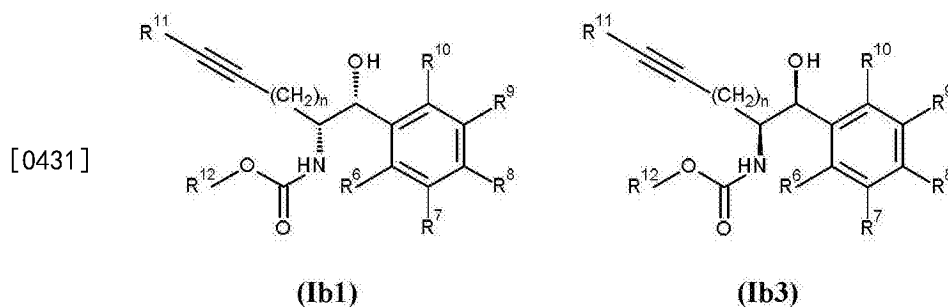
[0427] 在一些实施方案中,方法可在包括缓冲水溶液、有机溶剂、或共溶剂系统的反应条件下进行。在一些实施方案中,缓冲溶液为,例如,约0.025M至约0.25M硼酸盐的硼酸盐。在一些实施方案中,共溶剂系统包括约95%至约5%(v/v)的缓冲水溶液(例如,约0.2M硼酸盐),和约5%(v/v)至约95%(v/v)的有机溶剂溶液,例如,异丙醇。在一些实施方案中,共溶剂系统包含约30%(v/v)至约70%(v/v)的缓冲水溶液(例如,约0.2M硼酸盐),约70%(v/v)至约30%的有机溶剂溶液(例如,IPA)。在一些实施方案中,反应条件包括作为适当的溶剂的水,没有缓冲液存在。

[0428] 在一些实施方案中,共溶剂系统包括缓冲水溶液和IPA,其中IPA浓度为约5%至约95%(v/v)、约10%至约90%(v/v)、约15%至约90%(v/v)、约20%至约90%(v/v)、约25%至约80%(v/v)、约25%至约75%(v/v)、约35%至约75%(v/v)、约45%至约75%(v/v)、约55%至约75%(v/v)、约60%至约70%(v/v)、或约60%至约65%(v/v)。在一些实施方案中,IPA浓度为至少约25%(v/v)、至少约35%(v/v)、至少约45%(v/v)、至少约55%(v/v)、约

60% (v/v)、约65% (v/v)、约70% (v/v)、约75% (v/v)、约80% (v/v)、约85% (v/v)、约90% (v/v)或约95% (v/v)。在一些实施方案中,反应条件包括0.2M硼酸盐缓冲剂和约60% (v/v)至70% (v/v) IPA的共溶剂系统。在一些实施方案中,反应条件包括以下的共溶剂系统:约0.2M硼酸盐缓冲剂和约90% (v/v) IPA、约0.2M硼酸盐缓冲剂和约85% (v/v) IPA、约0.2M硼酸盐缓冲剂和约80% (v/v) IPA、约0.2M硼酸盐缓冲剂和约75% (v/v) IPA、约0.2M硼酸盐缓冲剂和约70% (v/v) IPA、约0.2M硼酸盐缓冲剂和约65% (v/v) IPA、约0.2M硼酸盐缓冲剂和约60% (v/v) IPA、约0.2M硼酸盐缓冲剂和约55% (v/v) IPA、约0.2M硼酸盐缓冲剂和约50% (v/v) IPA。

[0429] 如以上所描述,使用工程酮还原酶多肽将各种底物化合物转化为其相应产物化合物的方法可利用底物化合物,例如化合物(IIc1)和(IIc2)的立体异构体的平衡和本文公开的工程化酮还原酶多肽的高立体选择性以进行动态动力学拆分(DKR)动态,还称为动力学不对称转化(DKAT),由此形成的特定立体异构体或立体异构体的量比存在于反应开始时的相应立体异构体底物或底物,例如,化合物(IIc1)和化合物(IIc2)的量。

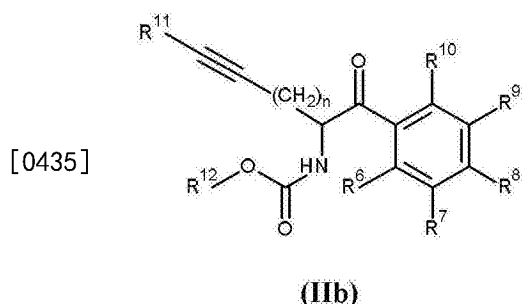
[0430] 因此,在一些实施方案中,用于制备非对映体过量于式(Ib3)的化合物的式(Ib1)的化合物的方法



[0432] 其中

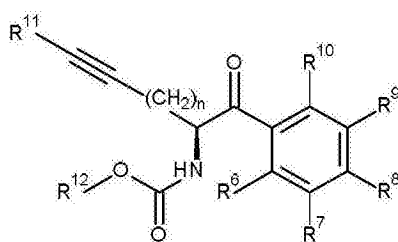
[0433] R^6 、 R^7 、 R^8 、 R^9 、 R^{10} 、 R^{11} 、 R^{12} 和 n 如以上关于式(Ib)的化合物所定义;

[0434] 包括使式(IIb)的底物化合物

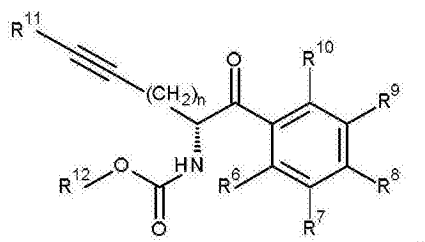


[0436] 在用于平衡式(IIb1)和(IIb2)的底物化合物的适当的反应条件下在辅因子NADPH或NADH的存在下与相比于对式(Ib2)和(Ib4)的化合物具有对式(Ib1)和(Ib3)的产物化合物的非对映选择性的工程化酮还原酶接触

[0437]



(IIb1)



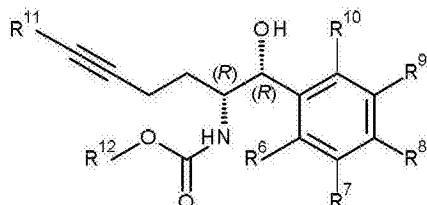
(IIb2)

[0438] 在以上方法的一些实施方案中,底物化合物(IIb)包括式(IIb1)和(IIb2)的化合物的混合物。

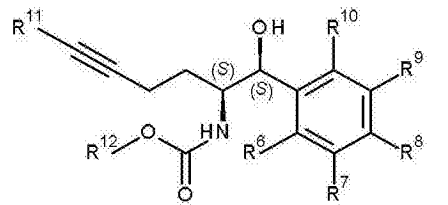
[0439] 在以上方法的一些实施方案中,底物化合物(IIb)包括式(IIb1)和(IIb2)的化合物的外消旋混合物。

[0440] 在方法的一些实施方案中,式(Ib1)的产物化合物包括式(Ic1)的化合物,并且形成非对映体过量于式(Ic3)的化合物的式(Ic1)的产物化合物

[0441]



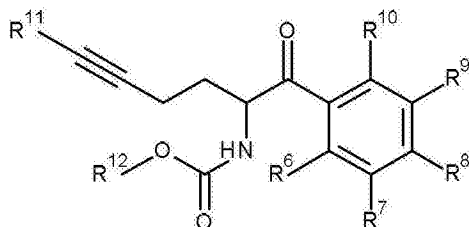
(Ic1)



(Ic3)

[0442] 因此,在一些实施方案中,用于制备非对映体过量于式(Ic3)的化合物的式(Ic1)的化合物的方法包括使式(IIc)的底物化合物

[0443]

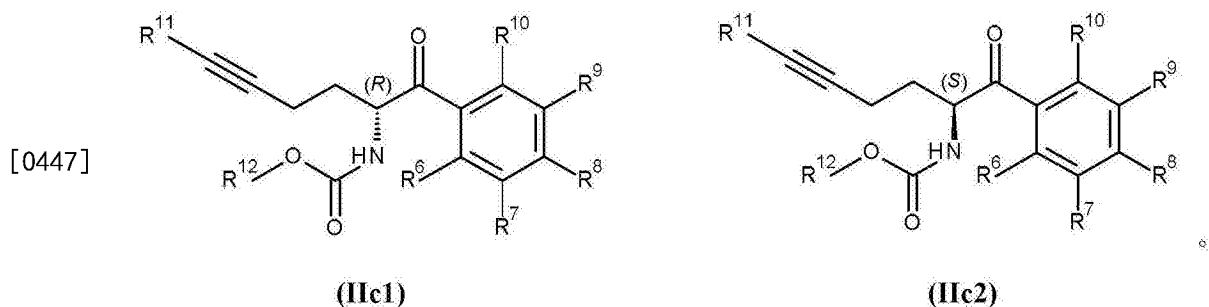


(IIc)

[0444] 其中

[0445] R⁶、R⁷、R⁸、R⁹、R¹⁰、R¹¹和R¹²如关于式(Ib)的化合物所定义;

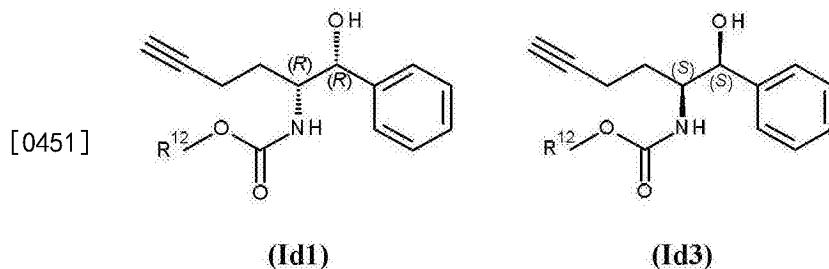
[0446] 在用于平衡式(IIc1)和(IIc2)的底物化合物的适当的反应条件下在辅因子NADPH或NADH的存在下与相比于对式(Ic2)和(Ic4)的化合物具有对式(Ic1)和(Ic3)的化合物的非对映选择性的工程化酮还原酶接触



[0448] 在以上方法的一些实施方案中,底物化合物(IIc)包括式(IIc1)和式(IIc2)的化合物的混合物。

[0449] 在以上方法的一些实施方案中,底物化合物(IIc)包括式(IIc1)和式(IIc2)的化合物的外消旋混合物。

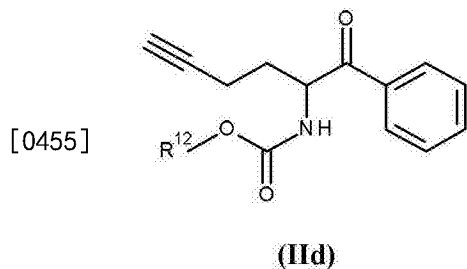
[0450] 在方法的一些实施方案中,式(Ib)的产物化合物包括式(Id1)的化合物,并且形成非对映体过量于式(Id3)的化合物的式(Id1)的产物化合物



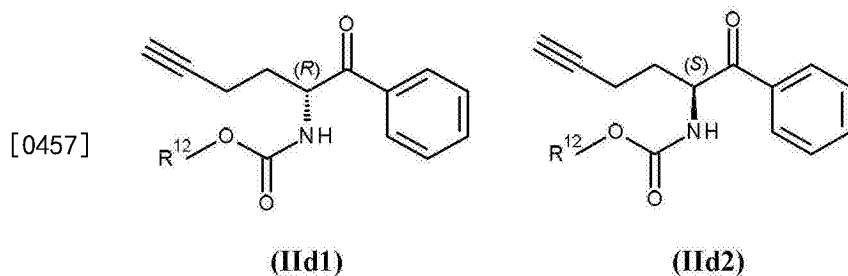
[0452] 其中

[0453] R^{12} 如关于式(Ib)的化合物所定义。

[0454] 因此,在一些实施方案中,用于制备非对映体过量于式(Id3)的化合物的式(Id1)的化合物的方法包括使式(IIId)的底物化合物



[0456] 在用于平衡式(IIId1)和(IIId2)的底物化合物的适当的反应条件下在辅因子NADPH或NADH的存在下与相比于对式(Id2)和(Id4)的化合物具有对式(Id1)和(Id3)的化合物的非对映选择性的工程化酮还原酶接触



[0458] 在用于制备非对映体过量于式(Id3)的化合物的式(Id1)的化合物的方法的一些

实施方案中,式(II_d)的化合物包括式(II_{d1})和(II_{d2})的底物化合物的混合物。

[0459] 在用于制备非对映体过量于式(I_{d3})的化合物的式(I_{d1})的化合物的方法的一些实施方案中,式(II_d)的化合物包括式(II_{d1})和(II_{d2})的底物化合物的外消旋混合物。

[0460] 因为酮还原酶反应中的底物化合物,例如,(II_{b1})和(II_{b2});(II_{c1})和(II_{c2});或(II_{d1})和(II_{d2})的平衡在pH 9或更高的条件下和45℃或更高的温度下是有利的,在一些实施方案中,用于DKR或DKAT的适当的条件可包括至少pH 9、至少pH 9.5、至少pH 10.0、至少pH 10.5、至少pH 11.0、至少pH 11.5的溶液pH,和至少45℃、至少50℃、至少55℃、至少60℃、或至少65℃的溶液温度。

[0461] 在以上方法的一些实施方案中,用于平衡底物化合物的反应条件包括约9至约12的pH和约45℃至约60℃的温度。在以上方法的一些实施方案中,用于平衡底物化合物的反应条件包括约10至约12的pH和约50℃至约65℃的温度。

[0462] 通常,酮还原酶介导的还原使用电子供体。在一些实施方案中,电子供体为辅因子。适合的辅因子包括,但不限于,NADP⁺(烟酰胺腺嘌呤二核苷酸磷酸)、NADPH(NADP⁺的还原的形式)、NAD⁺(烟酰胺腺嘌呤二核苷酸)和NADH(NAD⁺的还原的形式)。通常,将辅因子的还原形式加入至反应混合物中,并因此在一些实施方案中,方法在选自NADPH辅因子或NADH辅因子的电子供体的存在下进行。在一些实施方案中,电子供体为NADPH辅因子。在一些实施方案中,方法可以在其中反应条件包括以下的NADH或NADPH辅因子浓度下进行:约0.01至约1g/L、约0.03至约0.8g/L、约0.05至约0.5g/L、约0.1至约0.3g/L、约0.05至约0.2g/L或约0.1至约0.2g/L。在一些实施方案中,方法在以下的NADH或NADPH辅因子浓度下进行:约1g/L、约0.8g/L、约0.5g/L、约0.3g/L、约0.2g/L、约0.1g/L、约0.05g/L、或约0.01g/L。

[0463] 在方法的一些实施方案中,任选的辅因子再循环系统,还称为辅因子再生系统,可以用于从酶促反应中产生的NADP⁺/NAD⁺再生辅因子NADPH/NADH。辅因子再生系统是指参加还原辅因子的氧化形式的反应(例如,NADP⁺为NADPH)的一组反应物。由多肽还原酮类底物氧化的辅因子通过辅因子再生系统的还原形式再生。辅因子再生系统包括化学计量还原剂,其是还原氢等同物的来源并且能够还原辅因子的氧化形式。辅因子再生系统还可包括催化剂,例如酶催化剂,其通过还原剂还原催化辅因子的氧化形式。用于从NAD⁺或NADP⁺分别再生NADH或NADPH的辅因子再生系统是本领域已知的并且可以在本文描述的方法中使用。

[0464] 可以应用的适合的示例性辅因子再生系统包括,但不限于,葡萄糖和葡萄糖脱氢酶、甲酸和甲酸脱氢酶、葡萄糖-6-磷酸和葡萄糖-6-磷酸脱氢酶、仲(例如,异丙醇)醇和仲醇脱氢酶、亚磷酸和亚磷酸脱氢酶、分子氢和氢化酶等。这些系统可以与或NADP⁺/NADPH或NAD⁺/NADH作为辅因子的组合使用。使用氢化酶的电化学再生也可以用作辅因子再生系统,例如,描述于美国专利号5,538,867和6,495,023中的那些,这两者通过引用并入本文。包含金属催化剂和还原剂(例如,分子氢或甲酸)的化学辅因子再生系统还可以是适合的。参见,例如,PCT公布WO 2000/053731,其通过引用并入本文。

[0465] 在一些实施方案中,辅因子再生系统包括葡萄糖脱氢酶(GDH),其是分别将D-葡萄糖和NAD⁺或NADP⁺催化转化为葡糖酸和NADH或NADPH的NAD⁺或NADP⁺-依赖性酶。适合于在实践本文方法中使用的葡萄糖脱氢酶包括天然存在的葡萄糖脱氢酶以及非天然存在的葡萄糖脱氢酶。编码天然存在的葡萄糖脱氢酶的基因已报告于文献中,例如,枯草芽孢杆菌61297GDH基因、蜡状芽孢杆菌ATCC14579和巨大芽孢杆菌。例如使用诱变、定向进化等产生

的非天然存在的葡萄糖脱氢酶提供于PCT公布W02005/018579,和美国公布号2005/0095619和2005/0153417中。所有这些序列通过引用并入本文。

[0466] 在一些实施方案中,辅因子再生系统包括甲酸脱氢酶,其是分别将甲酸盐和 NAD^+ 或 NADP^+ 催化转化为二氧化碳和 NADH 或 NADPH 的 NAD^+ 或 NADP^+ -依赖性酶。适合于在本文描述的酮还原酶反应中用作辅因子再生系统的甲酸脱氢酶包括天然存在的甲酸脱氢酶以及非天然存在的甲酸脱氢酶。适合的甲酸脱氢酶描述于通过引用并入本文的PCT公布W0 2005/018579中。甲酸盐可以以盐,通常碱盐或铵盐的形式(例如, HCO_2Na 、 KHCO_2NH_4 以及类似的),以甲酸,通常甲酸水溶液或其混合物的形式来提供。碱或缓冲液可以用于提供所需的pH。

[0467] 在一些实施方案中,辅因子再生系统包括仲醇脱氢酶,其是分别将仲醇和 NAD^+ 或 NADP^+ 催化转化为酮和 NADH 或 NADPH 的 NAD^+ 或 NADP^+ -依赖性酶。适合于用作辅因子再生系统的仲醇脱氢酶包括天然存在的酮还原酶以及非天然存在的酮还原酶。天然存在的仲醇脱氢酶包括来自以下的已知的醇脱氢酶,例如并非限制:布氏热厌氧菌(*Thermoanaerobium brockii*)、红平红球菌(*Rhodococcus erythropolis*)、克菲尔乳杆菌、和短乳杆菌,且非天然存在的仲醇脱氢酶包括由此衍生的工程化醇脱氢酶。在一些实施方案中,可以使用为了热稳定性或溶剂稳定性而工程化的非天然存在的酮还原酶。此类酮还原酶描述于本申请和专利申请US 20080318295A1、US 20090093031A1、US 20090155863A1、US 20090162909A1、US 20090191605A1、US 20100055751A1、W0/2010/025238A2、W0/2010/025287A2、和US 20100062499A1;其中每一个通过引用并入本文。

[0468] 如从本公开内容将是明显的,本文描述的工程化酮还原酶能够将IPA转化为丙酮以分别再生辅因子 NADH / NADPH 或 NAD^+ / NADP^+ 。因此,在用于进行将化合物(2)或其结构类似物转化为化合物(1)或其相应的结构类似物的方法的一些实施方案中,辅因子再生系统的酮还原酶是本公开内容的工程化酮还原酶多肽,并且使用仲醇作为还原剂以回收在反应混合物中的 NADPH 或 NADH 辅因子。

[0469] 辅因子再生系统中有用的适合的仲醇,包括低级仲烷醇和芳基烷基甲醇。示例性低级仲醇包括,但不限于,异丙醇、2-丁醇、3-甲基-2-丁醇、2-戊醇、3-戊醇、3,3-二甲基-2-丁醇,以及类似的。在一些实施方案中,仲醇是异丙醇(IPA)。适合的芳基-烷基甲醇包括未取代的和取代的1-芳基乙醇。

[0470] 在一些实施方案中,方法可以在反应过程中没有添加 NADPH 或 NADH 辅因子且没有任何其他酶系统存在(例如,葡萄糖脱氢酶或甲酸脱氢酶)时进行。

[0471] 在一些实施方案中,本公开内容的方法可以在其中除了工程化酮还原酶多肽之外不存在辅因子再循环酶下进行。例如,反应条件可以包括约45-75%(v/v)的IPA浓度、约0.01-0.5g/L的 NADPH 或 NADH 辅因子载量,且其中除了工程化酮还原酶多肽外不存在辅因子再循环酶。

[0472] 在其中辅因子再循环系统产生挥发性产物,诸如来自异丙醇的丙酮的一些实施方案中,可以除去挥发性产物,例如,通过用非反应性气体喷射反应溶液或通过应用真空降低反应压力以除去气相中存在的挥发性产物。非反应性气体是不与反应组分反应的任何气体。各种非反应性气体包括氮气和稀有气体(例如,惰性气体)。在一些实施方案中,非反应性气体是氮气。例如,由异丙醇氧化形成的丙酮可以通过用氮气喷射反应溶液或对反应溶液应用真空且通过丙酮阱,诸如冷凝器或其他冷阱从气相中除去丙酮来除去。

[0473] 在本文的实施方案中,在方法中使用的酮还原酶多肽和任选的辅因子再生系统的任何另外的酶可以以纯化的酶、用编码酶的基因转化的全细胞、和/或细胞提取物和/或此类细胞的溶解产物的形式加入至反应混合物中。

[0474] 在一些实施方案中,编码酮还原酶多肽的基因和任选的辅因子再生酶可以单独转化进入宿主细胞或一起转化进入表达酶的相同宿主细胞。用编码工程化酮还原酶酶的基因转化的全细胞和/或任选的辅因子再生酶,或细胞提取物和/或其溶解产物,可以以各种不同的形式来被采用,包括固体(例如,低压冻干、喷雾干燥以及类似的)或半固体(例如,粗制糊)。在一些实施方案中,可将转化的细胞固定在固体支持物上,并通过使固定的细胞与底物化合物接触进行转化反应。

[0475] 通常,加入反应物(例如,底物、辅因子、多肽等)的顺序不是本公开内容的方法的关键。可以将反应物一起同时加入至溶剂(例如,单相溶剂、二相水性共溶剂系统、以及类似的溶剂),或可选地,可将反应物中的一些单独加入,以及在不同的时间点一起加入一些反应物。

[0476] 在一些实施方案中,可使用以上或本文其他处公开的混合物和反应条件的任何组合进行方法。因此,在一些实施方案中,本文描述的方法包括以下反应条件:(a)约1g/L至约10g/L的工程化酮还原酶多肽浓度;(b)约50g/L至约200g/L的载量浓度下的底物化合物;(c)约0.1g/L至约0.5g/L的NADP(H)浓度;(d)水性缓冲液和约30%至约70%(v/v)的IPA的共溶剂溶液;以及(e)35℃至约60℃的温度。

[0477] 在一些实施方案中,本文描述的方法包括以下的反应条件:(a)约0.1g/L至约1g/L的工程化酮还原酶多肽浓度;(b)约5g/L至约50g/L的载量浓度下的底物化合物;(c)约0.01g/L至约0.1g/L的NADP(H)浓度;(d)水性缓冲液和约30%至约70%(v/v)的IPA的共溶剂溶液;以及(e)约30℃至约45℃的温度。

[0478] 在一些实施方案中,在用于将底物化合物(2)转化为产物化合物(1)的方法中,化合物(2)为约50g/L至约200g/L的载量且方法在24h或更少时间内导致化合物(2)至化合物(1)的至少约91%、92%、93%、94%、95%、96%、97%、98%、99%或更多的转化率。

[0479] 在一些实施方案中,在用于将底物化合物(2)转化为产物化合物(1a)和(1c)的方法中,化合物(2)为约50g/L至约200g/L的载量且方法在24h或更少时间内导致以相比于化合物(1b)和(1d)的至少约91%、92%、93%、94%、95%、96%、97%、98%、99%或更多的非对映体过量形成化合物(1a)和(1c)。

[0480] 在一些实施方案中,还涵盖,包括使用工程化酮还原酶多肽将底物化合物生物催化转化为产物化合物的方法还包括以下的化学步骤:产物处理、提取、分离、纯化和/或结晶,其中的每一个可以在各种条件下进行。

[0481] 用于从通过如上公开的方法产生的生物催化反应混合物中提取、分离产物化合物、形成产物化合物的盐、纯化、和/或结晶产物化合物的方法、技术和试验方案是普通技术人员已知的和/或通过常规实验可获得的。另外,例证性的方法在下面的实施例中提供。

[0482] 本公开内容的各种特征和实施方案在以下代表性实施例中说明,代表性实施例意图是说明性的而非限制性的。

[0483] 实施例

[0484] 本公开内容的各种特征和实施方案在以下代表性实施例中说明,代表性实施例

意图是说明性的而非限制性的。

[0485] 实施例1:野生型酮还原酶基因的获得和表达载体的构建。

[0486] 将来自克菲尔乳杆菌的野生型酮还原酶基因(SEQ ID NO:1)设计为使用标准密码子优化在大肠杆菌中表达。(密码子优化软件参见例如,“OPTIMIZER:a web server for optimizing the codon usage of DNA sequences,”Puigbò等人,Nucleic Acids Res.2007年7月;35(Web Server issue):W126-31.Epub 2007年4月16)。基因使用寡核苷酸来合成并克隆进入表达载体pCK110900(如美国专利申请公布20060195947中的图3描述的载体,其据此通过引用并入本文)中的lac启动子的控制下。表达载体还含有P15a复制起点和氯霉素抗性基因。将产生的质粒使用标准方法转化进入大肠杆菌W3110(fhu⁻)。编码工程化酮还原酶多肽的多核苷酸还被克隆进入在大肠杆菌W3110中表达的载体pCK110900。

[0487] 编码SEQ ID NO:4的工程化酮还原酶多肽的多核苷酸(SEQ ID NO:3)通过编码克菲尔乳杆菌的野生型酮还原酶的密码子优化的基因(Genbank登录号AAP94029.1;GI:33112056)的定向进化来获得。与SEQ ID NO:2的天然存在的酮还原酶相比,SEQ ID NO:4具有11个氨基酸残基差异(A94G、S96V、E145F、F147M、L153T、Y190P、L195M、V196L、L199Y、I226V、和Y249W)。发现SEQ ID NO:4的多肽在SFP反应条件下(表2B)以相比于化合物(1b)和(1d)的约25的非对映体比将化合物(2)转化为化合物(1a)和(1c),且在DSP反应条件下(表2C)以相比于化合物(1b)和(1d)的>100的非对映体比但具有低酶活性将化合物(2)转化为化合物(1a)和(1c)。SEQ ID NO:3的多核苷酸被用作起始骨架用于随后几轮进化以产生具有改进特性的工程化酮还原酶。进行编码SEQ ID NO:4的基因(即,SEQ ID NO:3)的多轮定向进化。每轮使用来自每轮的编码最大改进的工程化多肽的基因作为用于随后轮进化的亲本“骨架”序列。产生的工程化酮还原酶多肽序列和具体突变以及相关活性列于表2A、2B和2C中。

[0488] 实施例2:生产工程化酮还原酶

[0489] 工程化酮还原酶多肽在大肠杆菌W3110中作为在lac启动子的控制下表达的细胞内蛋白产生。多肽主要累积为可溶性的细胞溶质的活性酶(soluble cytosolic active enzyme)。摇瓶程序用于产生工程化多肽粉,其可在本文公开的活性测定或生物催化方法中使用。

[0490] 摇瓶粉的发酵。将含有编码感兴趣的工程化酮还原酶的质粒的大肠杆菌的单个微生物菌落接种到含有30μg/ml氯霉素和1%葡萄糖的50mL Luria Bertani肉汤中。细胞在培养箱(incubator)中在30℃下以250rpm摇动生长过夜(至少16小时)。将培养物稀释到1升烧瓶中含有30μg/ml氯霉素的250mL Terrific Broth(12g/L细菌用胰蛋白胨、24g/L酵母提取物、4mL/L甘油、65mM磷酸钾、pH 7.0、1mM MgSO₄)中,至600nm的光密度(OD₆₀₀)为0.2,并允许在30℃生长。当培养物的OD₆₀₀是0.6至0.8时,通过加入异丙基-β-D-硫代半乳糖苷(“IPTG”)至最终浓度1mM来诱导酮还原酶基因的表达。然后培养持续过夜(至少16小时)。通过离心(5000rpm、15min、4℃)收获细胞,并丢弃上清液。

[0491] 生产酮还原酶摇瓶粉:将细胞团块用等体积的冷(4℃)100mM磷酸盐缓冲液,pH 9.0(任选地含有2mM MgSO₄)再悬浮,并如以上通过离心来收获。洗涤的细胞再悬浮于2体积的冷磷酸盐缓冲液中,并且以12,000psi经过French Press两次,同时保持在4℃。通过离心(9000rpm、45min、4℃)除去细胞碎片。收集澄清的溶解产物上清液,并储存在-20℃。冷冻的

澄清溶解产物的冷冻干燥提供了粗制酮还原酶多肽的干摇瓶粉。可选地,细胞团块(洗涤前或洗涤后)可储存在4℃或-80℃。

[0492] 用于生产下游处理(DSP)粉的发酵。用于生产DSP粉的工程化酮还原酶的较大规模(~100-120g)发酵可以作为短分批进行,随后根据标准生物加工方法通过进料分批法进行。简言之,通过添加IPTG至1mM的最终浓度来诱导酮还原酶表达。在发酵之后,收获细胞,并将其再悬浮在100mM三乙醇胺-H₂SO₄缓冲液中,然后通过匀化而被机械破裂。细胞碎片和核酸用聚乙烯亚胺(PEI)絮凝,并通过离心使悬浮液澄清。使用切向横流超滤膜浓缩所产生的澄清上清液以除去盐和水。然后浓缩的和部分纯化的酶浓缩物可在冷冻干燥器中干燥并被包装(例如,聚乙烯容器中)。

[0493] 高通量生长&表达。细胞被挑选并使用对于W3110的标准KRED方案以定向诱导来生长:(1)主生长-单个菌落通过Q-bot从琼脂Q-托盘中挑选并且在30℃,200rpm,85%湿度在含有1%葡萄糖和30μg/ml氯霉素(CAM)的LB培养基中生长过夜。(2)传代培养-将过夜生长的20μl转移至含有380μl的含有30μg/ml CAM、1mM IPTG、1mM MgSO₄的2xYT生长培养基的深孔板,且在30℃,200rpm,85%湿度下温育持续~18h。传代培养TB培养基由TB培养基(380μl/孔)、30μg/ml CAM、1mM MgSO₄和1mM IPTG组成。细胞培养物以4000rpm、4℃离心10min,并丢弃培养基。细胞团块再悬浮在200-400μl溶解缓冲液(0.1M三乙醇胺(TEA)缓冲液、pH 9.0,含有1mM MgSO₄、400μg/ml PMBS和500μg/ml溶菌酶)中。

[0494] 实施例3:酮还原酶活性测定

[0495] HTP筛选测定。标准HTP反应测定在96-孔深孔板(反应测定块)中以200μl反应体积规模进行。每孔中的反应混合物通常由以下组成:20g/L或50g/L化合物(2);0.05g/L NADP⁺;30%或50%异丙醇(IPA);和10μl、20μl或60μl的澄清溶解产物,如表2A中进一步指定的。

[0496] 反应条件A、B、D、E和G:反应条件G通过将2.5mL的异丙醇加入20mL小瓶中的500mg化合物(2)中,随后加入2mL的pH 10.0含有1mM MgSO₄的0.2M硼酸盐来进行。将所得的混合物加热至反应温度,并搅拌约15min。10g/L酶和1g/L的NADP的储备溶液在含有1mM MgSO₄的pH 10的0.2M硼酸盐中新鲜制备。将储备溶液(500μl)加入反应混合物中以启动反应。将小瓶保持紧密封闭并允许反应在期望的温度下在搅拌(1200rpm)下进行。为了跟踪反应的过程,5μl样品被取出并稀释到730μl的乙腈(MeCN)中。激烈混合后,将悬浮液短暂离心以分离不溶性微粒。澄清上清液的样品通过HPLC使用以下描述的条件来分析。针对如表2B和2C中指定的测定条件A、B、D、E、和G中的每个调整试剂。

[0497] 反应条件C和F:预混合物测定溶液通过将22.5mL异丙醇加入33mL含有1mM MgSO₄的pH 9的0.2M硼酸盐来制备。将预混合物溶液的pH用浓HCl调至10.2。对于每个实验,3.7mL的新鲜制备的储备溶液放置在20mL螺帽玻璃小瓶中,紧密封闭,并在搅拌(1200rpm)下加热至35℃。15min后,将250μl的在pH 9的0.2M硼酸盐中的100g/L酶粉储备溶液和50μl的在无菌水中的10g/L NADP储备溶液在35℃下以搅拌(1200rpm)加入每个小瓶中。加入酶和NADP后,立刻将1mL的在异丙醇中的200g/L化合物(2)储备溶液加入以启动反应。将小瓶紧密封闭并允许反应在35℃下在搅拌(1200rpm)下进行。为了跟踪反应的过程,15μl样品被取出并稀释到730μl的乙腈(MeCN)中。激烈混合后,将悬浮液短暂离心以分离不溶性微粒。澄清上清液的样品通过HPLC来分析。

[0498] HPLC分析:样品通过如下反相HPLC来分析。

[0499]

柱:	Chiralpak IC 柱(4.6 x 50 mm, 5 μ m)
温度:	25 °C
流速	1 mL/min
流动相:	等度 60%溶剂 A: 2 mM 甲酸铵, 用甲酸调 pH 3.5; 40%溶剂 B: 乙腈
运行时间	15 min
检测	210 nm

[0500] 底物和产物化合物具有以下保留时间:

[0501]

化合物(1c)/化合物(1d)	4.8min
化合物(1a)	5.6min
化合物(1b)	8.1min
底物化合物(2a)、(2b)、(2c)、和(2d)	10min和12min

[0502] 在本申请中引用的所有出版物、专利、专利申请和其它文件据此出于所有目的均通过引用以其整体并入,其程度如同分别指出将每个单独的出版物、专利、专利申请或其它文件出于所有目的通过引用并入一样。

[0503] 尽管已经阐释和描述了各种具体实施方案,但应理解可以作出各种改变而不背离本发明的精神和范围。

序列表

- <110> 科德克希思公司
 法比安·卡比罗尔
 陈海滨
 史蒂文·J·科利尔
 阿努潘·苟呵尔
 雅各布·杰恩
 波吉特·科斯耶克
 德里克·史密斯
- <120> 用于制备羟基取代的氨基甲酸酯的生物催化剂
- <130> CX2-103W01
- <150> 61/561,665
- <151> 2011-11-18
- <160> 172
- <170> PatentIn 3.5 版本

[0001]

- <210> 1
- <211> 756
- <212> DNA
- <213> 人工序列
- <220>
- <223> 密码子优化的克菲尔乳杆菌酮还原酶基因

<400> 1

```

atgaccgata gtctgaaggg caaagtagcc atcgtaaccg gcgggactct gggatcgggt    60
ttggcaatcg ccgataaatt ttagaggag ggtgcgaaag tagttattac tggtegtcac    120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc    180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggea    240
ttcggcccggt ttacgaccgt cgtgaacaat gcagggattg cagtttccaa aagcgttgaa    300
gacaactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttt    360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat    420
atgagcagta ttgaggggtt ctaggcgat ccgacgttgg gggcatataa cgcttccaag    480
ggggcggtae gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat    540
gtgcgtgtca acacagtaca tccgggtat atcaagaccg cgctggctga tgatctggaa    600
gggtctgagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg    660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt    720
gcgaatttg tggtcgacgg cgggtatacc gcacag                    756

```

<210> 2
 <211> 252
 <212> PRT
 <213> 克菲尔乳杆菌 (Lactobacillus kefir)

 <400> 2
 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg His Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
 50 55 60
 Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80
 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Ala Val Ser
 85 90 95
 [0002] Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
 100 105 110
 Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg
 115 120 125
 Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
 130 135 140
 Glu Gly Phe Val Gly Asp Pro Thr Leu Gly Ala Tyr Asn Ala Ser Lys
 145 150 155 160
 Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu
 165 170 175
 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Tyr Ile Lys
 180 185 190
 Thr Pro Leu Val Asp Asp Leu Glu Gly Ala Glu Glu Met Met Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Ile Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Tyr Thr Ala Gln
 245 250

<210> 3
 <211> 756
 <212> DNA
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 3
 atgacogac gtctgaaggg caaagtagec atcgtaaccg gegggacact gggatcgggt 60
 ttggcaatcg cegataaatt tglagaggag ggtgcgaaag tagttattac tggcgcgcac 120
 geggatgtag gtgaaaagge cgccaaatca atcggeggea ctgatgttat tegtgttgc 180
 cagcacgatg cgtccgatga agcaggctgg acgaaactgt tegacaccac cgaggaggca 240
 ttcgccccgg ttaegacegt cgtgaacaat gcagggattg gggttgttaa aagegttgaa 300
 gacactacca cggagggaatg gegtaaacig ctgtccgtta atctggatgg tgttttttcc 360
 ggcacccgctc tgggcattca gcgcattgaaa aataaagget tgggcgetag catcataaat 420
 atgagcagta ttttcgggat ggtaggcgat ccgactaccg gggcatacaa cgttcccaag 480
 ggggcgggtac gtatcatgtc gaaaagcgea gcgctggatt gcgcaetgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgcctga tgattatgaa 600
 ggtgctgagg aaatgatgtc acagegtacg aaaaccccta tgggcacacat tggcgaaccg 660
 aatgacatcg catgggtctg tgtgtacctg geatetgaeg aatcgaaatt tgcgaegggt 720
 gcagaatttg tggtegaegg cgggtggacc gcacag 756

[0003]

<210> 4
 <211> 252
 <212> PRT
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 4
 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg His Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
 50 55 60
 Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80
 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Gly Val Val

	85	90	95
Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser			
100	105	110	
Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg			
115	120	125	
Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile			
130	135	140	
Phe Gly Met Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys			
145	150	155	160
Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu			
165	170	175	
Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys			
180	185	190	
Thr Pro Met Leu Asp Asp Tyr Glu Gly Ala Glu Glu Met Met Ser Gln			
195	200	205	
Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala			
210	215	220	
Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly			
225	230	235	240
Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln			
245	250		

[0004]

<210> 5

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 5

atgaccgate gtctgaaggg caaagtagcc atcgtaaccg gegggacact gggatcgggt	60
ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggtcgtcac	120
gcggatgtag gtgaaaaggc cgccaaatca atcgggcgca ctgatgttat tcgctttgtc	180
cagcacgatg cgtccgatga ageaggetgg acgaaactgt tcgacaccac cgaggaggca	240
ttcggcccggt ttacgaccgt cgtgaacaat gcagggattg gggttcttaa aagcgttgaa	300
gacactacca cggagggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttct	360
ggcacccgctc tgggcattca gcgcatgaaa aataaaggct tgggcgetag catcatcaat	420
atgagcagta ttctcgggat ggtaggcgat ccgaactaccg gggcatacaa cgttctcaag	480
ggggcggttac gtatcatgtc gaaaagcgca gcgtggatt gegcactgaa ggactacgat	540
gtgcgtgtca acacagtaca tccgggcccc atcaagacc ccatgctega tgattatgaa	600
ggtgctgagg aatgatgtc acagcgtacg aaaaccccta tgggcccacat tggegaaccg	660
aatgacateg catgggtctg tctgtacctg gcatctgacg aatcgaaatt tgcgaagggt	720

gcagaatttg tggtegacgg cgggtggacc gcacag

756

<210> 6

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 6

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

[0005]

<210> 7
 <211> 756
 <212> DNA
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 7
 atgaccgatac gtctgaaggg caaagtagcc atcgtaaccg ggggacact gggatcggt 60
 ttggcaatcg ccgataaatt ttagaggag ggtgcgaaag tagttattac tggtegtcac 120
 gcggatgtag gtgaaaaggc cgccaaatca atcggeggca ctgatgttat tcgtttgtc 180
 cagcagatg cgtccgatga agcaggettg acgaaactgt tcgacaccac cgaggaggca 240
 ttggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
 gacaactacca cggaggaatg gegttaaactg ctgtccgtta atctggatgg tgttttttc 360
 ggcacccgtc tgggcattca gcgcataaaa aataaaggct tgggcgctag catcatcaat 420
 atgagcagta ttttcgggat ggtaggcgat ccgactaccg gggcatacaa cgttccaag 480
 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tcgggcccc atcaagacce cggctctega tgattatgaa 600
 ggtgctgagg aaatgatgtc acagcgtaac aaaaccccta tgggccacat tggcgaaccg 660
 aatgacatcg catgggtctg tgtgtacctg gcactgaacg aatcgaaatt tgcgacgggt 720
 gcagaatttg tggtcgacgg cgggtggacc gcacag 756

[0006]

<210> 8
 <211> 252
 <212> PRT
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 8
 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg His Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
 50 55 60
 Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala

65	70	75	80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Gly Val Val			
	85	90	95
Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser			
	100	105	110
Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg			
	115	120	125
Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile			
	130	135	140
Phe Gly Met Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys			
	145	150	155
Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu			
	165	170	175
Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys			
	180	185	190
Thr Pro Ala Leu Asp Asp Tyr Glu Gly Ala Glu Glu Met Met Ser Gln			
	195	200	205
Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala			
	210	215	220
Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly			
	225	230	235
Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln			240
	245	250	

[0007]

<210> 9

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 9

atgaccgatac gtctgaaggc caaagtagcc atcgtaaccg ggggacact gggatcggt	60
ttggcaatcg ccgataaatt ttagaggag ggtgcgaaag tagttattac tggtcgtcac	120
gcggatgtag gtgaaaagc cgcgaatca atcgggcgca ctgatgttat tcgtttgtc	180
cagcacgatg cgtccgatga agcaggtgg acgaaactgt tcgacaccac cgaggaggca	240
ttcgccccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa	300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc	360
ggcaccgctc tgggcattca gcgcataaaa aataaagct tgggcgttag catcatcaat	420
atgagcagta ttttcgggat ggtaggcgat ccgactaccg gggcatacaa cgtttccaag	480
ggggcggtac gtatcatgtc gaaaagcgca gcctgggatt gcgcactgaa ggactacgat	540
gtgcgtgtca acacagtaca tccgggcccc atcaagacct egatgetoga tgattatgaa	600

```

ggtgctgagg aaatgttttc acagegtacg aaaacccta tgggccacat tggcgaaccg      660
aatgacatcg catgggtctg tgtgtacctg gcattctgacg aatcgaaatt tgcgacgggt      720
gcagaatttg tggtcgacgg cgggtggacc gcacag                                756

```

<210> 10
 <211> 252
 <212> PRT
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

[0008]

```

<400> 10
Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
1          5          10          15
Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
20          25          30
Lys Val Val Ile Thr Gly Arg His Ala Asp Val Gly Glu Lys Ala Ala
35          40          45
Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
50          55          60
Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
65          70          75          80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Gly Val Val
85          90          95
Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
100         105         110
Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg
115         120         125
Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
130         135         140
Phe Gly Met Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys
145         150         155         160
Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu
165         170         175
Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys
180         185         190
Thr Pro Met Leu Asp Asp Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln
195         200         205
Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
210         215         220
Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
225         230         235         240

```

Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
245 250

<210> 11
<211> 756
<212> DNA
<213> 人工序列

<220>
<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 11
atgaccgatc gtctgaaggc caaagtagec atcgtaaccg geggacact gggatcggt 60
ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggtegtcgt 120
gcggtatgtag gtgaaaagge cgccaaatca atcggcggca ctgatgttat tcgtttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttcttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggegat ccgaactaccg gggcatacaa cgttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgtggatt gcgcactgaa ggaactacgat 540
gtgcgigtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgattatgaa 600
gggtctgagg aatgatgtc acagcgtacg aaaacccta tgggtacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcactcgaag aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacag 756

[0009]

<210> 12
<211> 252
<212> PRT
<213> 人工序列

<220>
<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 12
Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
1 5 10 15
Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
20 25 30
Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
35 40 45
Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala

	50	55	60
	Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala		
	65	70	75
	Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Gly Val Leu		80
		85	90
	Lys Ser Val Glu Asp Thr Thr Thr Glu Trp Arg Lys Leu Ser		95
	100	105	110
	Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg		
	115	120	125
	Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile		
	130	135	140
	Phe Gly Met Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys		
	145	150	155
	Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu		160
		165	170
	Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys		175
	180	185	190
	Thr Pro Met Leu Asp Asp Tyr Glu Gly Ala Glu Glu Met Met Ser Gln		
	195	200	205
	Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala		
	210	215	220
[0010]	Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly		
	225	230	235
	Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln		240
		245	250

<210> 13

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 13

atgaccgatc gtctgaaggg caaagtagcc ategtaaccg gegggacact gggatcgggt	60
ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggtcgtcgt	120
gcggatgtag gtgaaaaggc cgcacaaatca ategggcgca ctgatgttat tegtcttctc	180
cagcacgatg cgtccgatga agcaggetgg acgaaactgt tgcacaccac cgaggaggca	240
ttcgccccgg ttaegaccgt cgtgaacaat gcagggattg gggttgttaa aagecgttgaa	300
gacactacca cggagggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttctc	360
ggcacccgtc tgggcattca gcgcattgaaa aataaagget tgggcgctag catcatcaat	420
atgagcagta ttttcgggat ggtaggcgat ccgaactaccg gggcatacaa cgttccaag	480

```

ggggcgggtac gtatcatgtc gaaaagcgca gcgctggatt ggcactgaa ggactacgat      540
gtgcgtgtca acacagfaca tcgggceccc atcaagaccc cgatgetcga tgattatgaa      600
gggtcgtgagg aaatgttttc acagcgtacg aaaacccta tgggtcacat tggcgaaaccg      660
aatgacatcg catgggtctg tgtgtacctg gcattctgacg aatcgaaatt tgcgacgggt      720
gcagaatttg tggtcgacgg cgggtggacc gcacag                                756

```

<210> 14

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 14

[0011]

```

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

```

<400> 16
Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
1 5 10 15
Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
20 25 30
Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala

	35	40	45
	Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala		
	50	55	60
	Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala		
	65	70	75
	Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Gly Val Val		80
	85	90	95
	Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser		
	100	105	110
	Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg		
	115	120	125
	Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile		
	130	135	140
	Phe Gly Met Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys		
	145	150	155
	Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu		
	165	170	175
	Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys		
	180	185	190
	Thr Pro Ala Leu Asp Asp Tyr Glu Gly Ala Glu Glu Met Met Ser Gln		
	195	200	205
[0013]	Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala		
	210	215	220
	Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly		
	225	230	235
	Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln		240
	245	250	

<210> 17

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 17

atgaccgatac gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggatcgggt	60
ttagcaatcg ccgataaatt ttagaggag ggtgcgaaag tagttattac tggctcgtcgt	120
gcggatgtag gtgaaaaggc cgcgaatca atcgggcgca ctgatgttat tegettgtc	180
cagcacgatg cgtcgatga agcaggtgg acgaaactgt tcgacaccac cgaggaggca	240
ttcggcccggt ttacgaccgt cgtgaacaat gcagggatc ctgtttgttaa aagcgttgaa	300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc	360

```

ggcaccgcgc tgggcattca gcgcattgaaa aataaagget tgggegetag catcateaat 420
atgagcagta ttttcgggat ggtaggcgat cegactaccg gggcatacaa cgetteccaag 480
ggggcgggtac gtatcatgtc gaaaagegca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtea acacagtaca tccgggcccc atcaagacc ccatgctcga tgattatgaa 600
gggtcgtgagg aatgttttc acagcgtacg aaaaccctta tgggtcacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcattctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggctgacgg cgggtggacc gcacag 756

```

<210> 18

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 18

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

[0014]

Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

<210> 19
 <211> 756
 <212> DNA
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

[0015]
 <400> 19
 atgaccgac gtctgaaggg eaaagtagcc atcgtaaccg geggacact gggatcggg 60
 ttggcaatcg cegataaatt tctagaggag ggtgcgaaag tagttattac tggtcgtcgt 120
 gcggatgtag gtgaaaagge cgccaaatca atcggcggea ctgatgttat tgcctttgtc 180
 cagcaogatg cgtccgatga agcaggetgg acgaaactgt tcgacaccac cgaggaggca 240
 ttccggcccg ttaegacegt cgtgaacaat gcagggaattg ggatggttaa aagegttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttcc 360
 ggcaccgctc tgggcattca gcgcattgaaa aataaaggct tgggcgctag catcataaat 420
 atgagcagta ttttcgggat ggtaggcgat cegactaccg gggcatacaa cgettcgaag 480
 gggcgcggtac gtatcatgtc gaaaagcgca gcctgggatt gcgcactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgattatgaa 600
 ggtgctgagg aaatgttttc acagegtacg aaaaccctta tgggtcacaat tggcgaaccg 660
 aatgacatcg catgggtctg tgtgtacctg geatctgaag aatcgaaatt tgcgaegggt 720
 gcagaatttg tggtcgacgg cgggtggacc gcacag 756

<210> 20
 <211> 252
 <212> PRT
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 20
 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala

	20	25	30
	Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala		
	35	40	45
	Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala		
	50	55	60
	Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala		
	65	70	75
	Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Gly Met Val		
	85	90	95
	Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser		
	100	105	110
	Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg		
	115	120	125
	Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile		
	130	135	140
	Phe Gly Met Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys		
	145	150	155
	Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu		
	165	170	175
	Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys		
	180	185	190
[0016]	Thr Pro Met Leu Asp Asp Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln		
	195	200	205
	Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala		
	210	215	220
	Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly		
	225	230	235
	Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln		240
	245	250	

<210> 21

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 21

atgaccgata gtctgaaggc caaagtagcc atcgtaaccg ggggacact gggatcggg	60
ttggcaatcg ccgaiaaatt tctagaggag ggtgcgaaag tagttattac tggctcgcgt	120
gcggatgtag gtgaaaagcc cgcgaatca atcgggcgca ctgatgttat tcgctttgtc	180
cagcagcatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240

```

ttcggcccggt ttaacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa    300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttct    360
ggcacccgctc tgggcattca ggcgatgaaa aataaagget tgggegetag catcatcaat    420
atgagcagtg ttttcgggat ggtaggcgat ccgaactaccg gggcatacaa cgtttccaag    480
ggggcggtac gtatcatgtc gaaaagcgca gcgttgatt gcgcactgaa ggactacgat    540
gtgcgtgtea acacagtaca tccgggcccc atcaagaccc cgatgetcga tgattatgaa    600
gggtgctgagg aaatgttttc acagegtacg aaaacccta tgggtcacat tggcgaaccg    660
aatgacatcg catgggtctg tgtgtacctg gcattctgacg aategaaatt tgcgacgggt    720
gcagaatttg tggtcgacgg cgggtggacc gcacag                                756

```

<210> 22

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 22

[0017]

```

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

```

Thr Pro Met Leu Asp Asp Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

<210> 23

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 23

[0018]

atgaccgate gtctgaaggc caaagtagcc atcgtaaccg ggggacact gggatcgggt 60
 ttggcaatcg ccgataaatt tctagaggag ggtgagaaag tagttattac tggctgctgt 120
 ggggatgtag gtgaaaagcg cgcgaaatca atcgggcgca ctgatgttat tegtctgtgc 180
 cagcagcatg cgtccgatga agcaggtctg acgaaactgt tcgacaccae cgaggaggca 240
 ttgggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
 gacactacca eggaggaaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttgc 360
 ggcaccgctc tgggcattea gcgcattgaa aataaaggct tgggcgctag catcatcaat 420
 atgagcagta ttttcgggat ggtaggcgat ccgactaccg gggcatacaa cgttcccaag 480
 ggggagggtac gtatcatgtc gaaaagcgca gcctggattg gcgcactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgaatgatga tgattatgaa 600
 ggtgtctagg aaatgttttc acagcgtaacg aaaaccctta tgggtcacat tggcgaaccg 660
 aatgacatcg catgggtctg tgtgtacctg gcattctgacg aatcgaaatt tgcgacgggt 720
 gcagaatttg tggctgacgg cgggtggacc gcacag 756

<210> 24

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 24

Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr

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

[0019]

<210> 25

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 25

atgaccgatac gtctgaaggg caaagtagcc atcgtaaccg ggggacact gggatcggt 60
 ttggcaatcg ccgataaatt ttagaggag ggtcgaaag tagttattac tggtcgtcgt 120

```

gcgatgtag gtgaaaagge cgccaaatca atcggeggca ctgatgttat tcgttttgtc 180
cagcaogatg cgtccgatga agcaggetgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggccegg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg cgtccgtta atctggatgg tgtttttttc 360
ggcaccegtc tgggcattca ggcgatgaaa aataaaggct tgggegetag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgactaccg gggcatacaa cgtttccaag 480
ggggcgggtac gtatcatgtc gaaaagcgea gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga ttctgtatgaa 600
ggtgctgagg aaatgttttc acagcgtacg aaaaccoccta tgggtcacat tggcgaaccg 660
aatgacateg catgggtctg tgtgtacctg gcctctgacg aatcgaaatt tgcgacgggt 720
gcgaatttg tggtcgacgg cgggtggacc gcacag 756

```

<210> 26

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

[0020]

<400> 26

```

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

```

Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys
 180 185 190
 Thr Pro Met Leu Asp Ser Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

<210> 27

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 27

[0021] atgaccgac gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggatateggt 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcgt 120
 gcggatgtag gtgaaaaggc cgcgaaatca atcggcgga ctgatgttat tcgctttgtc 180
 cagcacgatg ttccgatga agcaggettg acgaaactgt tegacaccac cgaggaggca 240
 ttcgccccgg ttacgacctt agtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
 gacactacca cggagggaatg gcgtaactg ctgtccgta atctggatgg tgttttttc 360
 ggcaaccgtc tgggcattca gcgatgaaa aataaagget tgggcgctag catcatcaat 420
 atgagcagta ttttcgggca ggtaggcgat ccgactaccg gggeatacaa cgtttccaag 480
 ggggcggta cgtatcatgc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcccc atcaagacce cgatgetcga tgatcatgaa 600
 ggtgtctgag aaatgttttc acagcgtacg aaaaccccta tgggtcacat tggcgaaaccg 660
 aatgacatcg catgggtctg tgtgtacctg gcactctgacg aatcgaaatt tgcgacgggt 720
 gcagaatttg tggtegacgg cgggtggacc gcacag 756

<210> 28

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 28
 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Val
 50 55 60
 Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80
 Phe Gly Pro Val Thr Thr Leu Val Asn Asn Ala Gly Ile Gly Val Val
 85 90 95
 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
 100 105 110
 Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg
 115 120 125
 Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
 130 135 140
 Phe Gly Gln Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys
 145 150 155 160
 Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu
 165 170 175
 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys
 180 185 190
 Thr Pro Met Leu Asp Asp His Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

[0022]

<210> 29

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 29

```

atgaccgate gtctgaaggg caaagtagec atcgtaaccg geggacact gggatcggt      60
ttggcaatcg cggataaatt tctagaggag ggtgcgaaag tagttattac tggtegtcgt      120
gcggatgtag gtgaaaagge cgcctaatca atcggcggca ctgatgttat tcgttttgc      180
cagcaccgatg ttccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggea      240
ttcgcccggt ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa      300
gacactacea cggaggaatg gcgtaaacgt ctgtccgtta atctggatgg tgttttttc      360
ggcaccgctc tgggcattca gcgcataaaa aataaaggct tgggcgctag catcateaat      420
atgagcagta ttttcgggca ggtaggegat ccgaactacc gggcatacaa cgttcccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgcgtggatt gcgcactgaa ggactacgat      540
glcgctgtca acacagtaca tccgggcccc atcaagacc cgaigctoga tgalcatgaa      600
gggtctgagg aaatgttttc acagcgtacg aaaacccta tgggtcacat tggcgaaccg      660
aatgacatcg catgggtctg tgtgtacctg gcattcgacg aatcgaaatt tgcgacgggt      720
gcagaatttg tggtcgacgg cgggtggacc gcacag                                756

```

<210> 30

<211> 252

<212> PRT

<213> 人工序列

<220>

[0023]

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 30

```

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

```

Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu
 165 170 175
 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys
 180 185 190
 Thr Pro Met Leu Asp Asp His Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

<210> 31
 <211> 756
 <212> DNA
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

[0024]

<400> 31
 atgaccgatac gtctgaaggc caaagtagcc atcgtaaccg gcgggacact gggatateggt 60
 ttggcaatcg cegataaatt tctagaggag ggtgcgaaag tagttattac tggctcgtcgt 120
 gcggatgtag gtgaaaagcc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
 cagcacgatg cgtccgatga ageaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 ttcggcccggt ttaagaccgt cgtgaacaat gcagggatcc cggttgttaa aagcgttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
 ggcaccgctc tgggcattca gcgcattgaa aataaaggct tgggcgctag catcatcaat 420
 atgagcagtg tgttcgggat ggtaggcgat ccgactaccg gggcatacaa cgtttccaag 480
 ggggcggtag gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgatgga tgattatgaa 600
 ggtgctgagg aaatgttttc acagcgtaag aaaaccctta tgggtcacat tggcgaaccg 660
 aatgacatcg catgggtctg tgtgtacctg gcattctgac aatcgaaatt tgcgacgggt 720
 gcagaatttg tggctcgacgg cgggtggacc gcacag 756

<210> 32
 <211> 252
 <212> PRT
 <213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 32

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

<210> 33

756 <211>

<212> DNA

〈213〉 人工序列

 $\langle 220 \rangle$

〈223〉 克菲尔乳杆菌酮还原酶的工程化变体

<400> 33
 atgaccgatac gtctgaaggg caaagtagcc atcgtaaccg gcgggacaca aggtateggt 60
 ttggcaateg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggctgctgt 120
 gcgcatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
 cagcacgatg tatecgatga agcagctgg ectaaactgt tegacaccac cgaggaagea 240
 ttgggccggg ttacgaccgt agtgaacaat gcagggattc cgatgggtta aagecgttgaa 300
 gaaactacca cggaggaatg gegttaaactg ctgtccgtta atctggatgg tgttttttc 360
 ggcgccgtc tgggcattca gcgcatgaaa aataaagget tgggcgctag catcataat 420
 atgagcagtg tgttcgggat tgtaggcgac ccgactaccg gggeatactg tgcctccaag 480
 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgatgga tagctatgaa 600
 ggtgctgagg aaatgttttc acagegtacg aaaacccta tgggtcacat tggcgaaccg 660
 aatgacatcg catgggtctg tgtgtacctg gcactgaag aatcgaaatt tgcgacgggt 720
 gcagaatttg tggtegacgg egggtacacc gcacag 756

<210> 34
 <211> 252
 <212> PRT
 <213> 人工序列

[0026] <220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 34
 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15

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

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

<210> 35

<211> 756

<212> DNA

<213> 人工序列

[0027]

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 35

atgaccgatac gtctgaaggg caaagtagcc atcgtaaccg gcgggacaca aggtatcggt	60
ttggcaatcg cegataaatt ttagaggag ggtgcgaaag tagttattac tggtegtcgt	120
gcggatgtag gtgaaaaggc cgccaaatca atcgggcgca ctgatgttat tcgctttgtc	180
cagcaecatg tatccgatga agcaggctgg cctaaactgt tegacaccac cgaggaggca	240
ttcggcccggt ttacgaccgt agtgaacaat gcagggatcc cggtgggtta aagcgttgaa	300
gacactacca cggaggaaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc	360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat	420
atgagcagtg tgttcgggat ttaggcgac ccgactaccg gggcatactg tgettccaag	480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgactgaa ggactacgat	540
gtgcgtgtea acacagtaca tccgggcccc atcaagaccc cgatgatgga tgattatgaa	600
gggtgtgagg aaatgtttc acagcgtacg aaaaccccta tgggtcaca tggcgaccg	660
aatgacateg catgggtctg tgtgtacctg gcactgacg aatcgaaatt tgcgacgggt	720
gcagaatttg tggtcgacgg cgggttcacc gcacag	756

<210> 36

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 36

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

[0028]

<210> 37

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 37

```

atgaccgate gtctgaaggg caaagtagcc atcgtaaccg gegggacaat gggatcgggt      60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtegtcgt      120
gggatgtag gtgaaaaggc cgcgaaatca atcggcggca ctgatgttat tcegtttgtc      180
cagcacgatg tatccgatga agcaggctgg actaaactgt tegacaccac cgaggaagca      240
ttcggeccgg ttacgacett agtgaacaat gcagggatte cggttgttaa aagcgttgaa      300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc      360
ggcacccgtc tgggcattca ggcgatgaaa aataaagget tgggcgctag catcatcaat      420
atgagcagtg tgttcgggat ggtaggcgac ccgactaccg gggcatactg tgettccaag      480
ggggcggtac gttatcgtc gaaaagegca gcgtcggatt ggcgactgaa ggactacgat      540
gtgcgtgtca acacagtaca tccgggcccc atcaagacce cgatgatgga tagccatgaa      600
gggtcgtagg aaatgttttc acagcgtacg aaaaccccta tgggtcacat tggcgaaccg      660
aatgacateg catgggtctg tgtgtacctg gcactcgacg aatcgaaatt tgcgacgggt      720
gcgaatttg tggtcgacgg cgggttcacc gcacag                                     756

```

<210> 38

<211> 252

[0029]

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 38

```

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

```

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

[0030]

<210> 39

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 39

atgacogatc gtctgaaggg caaagtagcc atcgtaaceg geggacacg cggatateggt	60
ttagcaatcg ccgataaatt ttagaggag ggtgcgaaag tagttattac tggctegct	120
gaggatgtag gtgaaaaggc cgcacaaatca atcgccggca ctgatgttat tcgctttgtc	180
cagcagcatg tatccgatga agcaggctgg actaaactgt tgcacaccac cgaggaggca	240
ttcgggcccgg ttacgacctt agtgaacaat gcagggatc cggttgttaa aagcgttgaa	300
gacaactacca cggagggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc	360
ggcacccgte tgggcattca gcgatgaaa aataaaggct tgggcgctag catcataat	420
atgagcagtg tgttcgggat ttaggctac ccgactaccg gggcatactg tgcctccaag	480
ggagcggtae gtaatcgtc gaaaagcga gcgtggatt gcgactgaa ggactacga	540
gtgcgtgtca acacagtaca tccgggccc atcaagacc cgatgatgga tagctatgaa	600
ggtgctgagg aaatgttttc acagegtacg aaaacccta tgggtcacat tggcgaaccg	660
aatgacateg catgggtctg tgtgtacctg gcactgtacg aatcgaaatt tgcgacgggt	720
gcagaatttg tggtcgacgg cgggtggaac gcacag	756

<210> 40

<211> 252
 <212> PRT
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 40

[0031] Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Arg Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Val
 50 55 60
 Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80
 Phe Gly Pro Val Thr Thr Leu Val Asn Asn Ala Gly Ile Pro Val Val
 85 90 95
 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
 100 105 110
 Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg
 115 120 125
 Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Val
 130 135 140
 Phe Gly Ile Val Gly Tyr Pro Thr Thr Gly Ala Tyr Cys Ala Ser Lys
 145 150 155 160
 Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu
 165 170 175
 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys
 180 185 190
 Thr Pro Met Met Asp Ser Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

<210> 41
 <211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 41

```

atgacccgac gtctgaaggc caaagtagcc atcgtaaccg gggggacaca aggtatcggt      60
ttggcagtcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggtcgtcgt      120
gggatgtag gtgaaaagcc cgcgaatca atcgggcgca ctgatgttat tegtcttgtc      180
cagcacgatg tatccgatga agcagcctgg cctaaactgt tcgacaccac cgaggaggca      240
ttcgccccgg ttaagacctt agtgaacaat gcagggatcc cgatgggtta aagcgttgaa      300
gacactacca cggaggaaat gcgtaaactg ctgtccgtta atctggatgg tgttttttgc      360
ggcaccgcgc tgggcattca gcgcattgaa aataaagget tgggccttag cactcatcaat      420
atgagcagtg igtctgggat tctagggcac ccgactggcg gggcactatg tctctccaag      480
ggggcggtag gtatcatgtc gaaaagcgca gcctctggat gcgcactgaa ggactacgat      540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgatgga tagctatgaa      600
ggctctgagg aaatgttttc acagcgtacg aaaaccctta tgggtccatc tggcgaccgc      660
aatgacatcg catgggtctg tctgtacctg gcctctgaag aatcgaaatt tgcgacgggt      720
gcagaatttg tggtcgacgg cgggttcacc gcacag                                     756

```

[0032]

<210> 42

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 42

```

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

```

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

[0033]

<210> 43

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 43

atgaccgatc gtcftgaaggg caaagtagcc atcgtaaccg gegggacacg cggatcgggt	60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtegtcgt	120
gcggatgtag gtgaaaaggc cgcgaaatec atcggcggca ctgatgttat tcgctttgtc	180
cagcacgatg tatccgatga agcaggctgg cctaaactgt tcgacaccac cgaggaggea	240
ttcgggcccg ttacgaccgt agtgaacaat gcagggatc cgatggttaa aagcgttgaa	300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc	360
ggcaccgctc tgggcattca gcgatgaaa aataaagget tgggcgctag catcatcaat	420
atgagcagtg tgttcgggat ggtaggctac ccgactaccg gggcatacaa tgcctccaag	480
ggggcggtac gtatcatgtc gaaaagcgca gcgatggatt gcgcactgaa ggactacgat	540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgatgga tagctatgaa	600
ggtgetgagg aaatgttttc acagcgtaacg aaaaccccta tgggtcacat tggcgaaaccg	660
aatgacatcg catgggtctg tgtgtacctg geatctgaacg aatcgaaatt tgcgacgggt	720
gcgaatttg tggtcgacgg cgggtggacc gcacag	756

<210> 44
 <211> 252
 <212> PRT
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 44

Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Arg Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Val
 50 55 60
 Ser Asp Glu Ala Gly Trp Pro Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80
 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Pro Met Val
 [0034] 85 90 95
 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
 100 105 110
 Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg
 115 120 125
 Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Val
 130 135 140
 Phe Gly Met Val Gly Tyr Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys
 145 150 155 160
 Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu
 165 170 175
 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys
 180 185 190
 Thr Pro Met Met Asp Ser Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

<210> 45
 <211> 756
 <212> DNA
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 45
 atgacogate gtctgaaggg caaagtagec atcgtaaccg gegggacact gggatcgggt 60
 ttggcaatcg cegataaatt tgltagaggag ggtgcgaaag tagttattac tggtcgtcac 120
 geggatgtag gtgaaaagge cgccaaatca atcggeggea ctgatgttat tegtittgtc 180
 cagcacgatg cgtccgatga agcaggctgg acgaaactgt tegacaccac cgaggaggca 240
 ttcgcccggt ttaegacegt cgtgaacaat gcagggattg gggtttataa aagegttgaa 300
 gacactacca cggaggaatg gegtaaacgt ctgtccgtta atctggatgg tgttttttcc 360
 ggcacccggtc tgggcattca gcgcattgaaa aataaagget tgggcgetag catcateaat 420
 atgagcagta ttttcgggat ggtaggcgat cegactaccg gggcatacaa cgttcccaag 480
 ggggcgggtac gtatcatgtc gaaaagcgea gcgtcggatt gcgcactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgattatgaa 600
 ggtgctgagg aaatgatgtc acagegtacg aaaaccccta tgggcacat tggcgaaccg 660
 aatgacatcg catgggtctg tgtgtacctg geatetgaeg aatcgaaatt tgcgaegggt 720
 gcagaatttg tggtegaegg cgggtggacc gcacag 756

[0035]

<210> 46
 <211> 252
 <212> PRT
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 46
 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg His Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
 50 55 60
 Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80
 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Gly Val Tyr

	85	90	95
Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser			
100	105	110	
Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg			
115	120	125	
Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile			
130	135	140	
Phe Gly Met Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys			
145	150	155	160
Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu			
165	170	175	
Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys			
180	185	190	
Thr Pro Met Leu Asp Asp Tyr Glu Gly Ala Glu Glu Met Met Ser Gln			
195	200	205	
Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala			
210	215	220	
Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly			
225	230	235	240
Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln			
245	250		

[0036]

<210> 47

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 47

atgacogate gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggatcggt	60
ttggcaatcg ccgataaatt ttagaggag ggtgcgaaag tagttattac tggtegtac	120
gcggaigttag gtgaaaaggc cgcacaaatca atcggcggea ctgatgttat tgcctttgtc	180
cagcagcatg cgtccgatga agcaggctgg acgaaactgt tegacaccac cgaggaggca	240
ttcggeccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa	300
gacactacca cggaggaatg gcgtaaactg ctgtcogtta atctggatgg tgttttttc	360
ggcacccgte tgggcattca gcgcattgaaa aataaaggct tgggcctag catcatcaat	420
atgagcagta ttttggggat ggtaggcgat ccgactaccg gggcatacaa cgttccaag	480
ggggcggtac gtatcatgtc gaaaagegca gcctggatt gcgcactgaa ggactacgat	540
gtgcgtgtca acacagtaca tccgggcccc atcaagacce cgatgetega tgattatgaa	600
ggtgctgagg aatgatgtc acagcgtacg aaaacccta tgggeacat tggcgaaccg	660

aatgacatcg catgggtctg tgtgtacctg gcattctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacag 756

<210> 48

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 48

[0037]

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

245

250

<210> 49

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 49

atgaccgate gtctgaaggc caaagtagcc atcgtaaccg ggggacact gggatcggt	60
ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggtegtcac	120
gcggatgtag gtgaaaagcc cgcgaatca atcgccggca ctgatgttat tctctttgtc	180
cagcacgatg cgtccgatga agcaggctgg acgaaaactgt tcgacaccac cgaggaggca	240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa	300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttct	360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat	420
atgagcagta ttctcgggat ggtagggat ccgaactacc gggcatacaa cgtctccaag	480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggaactacgat	540
gtgcgigtca acacagtaca tccgggcccc atcaagaccc cggggctcga tgattatgaa	600
ggctctgagg aatgatgtc acagcgtacg aaaaacctta tgggceacat tggcgaaccg	660
aatgacatcg catgggtctg tctgtacctg gcactcgaag aatcgaaatt tgcgacgggt	720
gcagaatttg tggtcgacgg cgggtggacc gcacag	756

[0038]

<210> 50

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 50

Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr	
1 5 10 15	
Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala	
20 25 30	
Lys Val Val Ile Thr Gly Arg His Ala Asp Val Gly Glu Lys Ala Ala	
35 40 45	
Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala	

	50	55	60
	Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala		
	65	70	75
	Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Gly Val Val		80
		85	90
	Lys Ser Val Glu Asp Thr Thr Thr Glu Trp Arg Lys Leu Ser		95
	100	105	110
	Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg		
	115	120	125
	Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile		
	130	135	140
	Phe Gly Met Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys		
	145	150	155
	Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu		160
		165	170
	Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys		
	180	185	190
	Thr Pro Gly Leu Asp Asp Tyr Glu Gly Ala Glu Glu Met Met Ser Gln		
	195	200	205
	Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala		
	210	215	220
[0039]	Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly		
	225	230	235
	Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln		240
		245	250

<210> 51

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 51

atgaccgatc gtctgaaggg caaagtagcc ategtaaccg gegggacact gggtatcggt	60
ttggcaatcg ccgataaatt thtagaggag ggtgcgaaag tagttattac tggtcgtcac	120
gcggatgtag gtgaaaaggc cgcacaaatca ategggcgca ctgatgttat tegtattgtc	180
cagcacgatg cgtccgatga agcaggetgg acgaaactgt tcgacaccac cgaggaggca	240
ttcgccccgg ttaegaccgt cgtgaacaat gcagggattg gggttgttaa aagegttgaa	300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc	360
ggcacccgtc tgggcattca gcgcattgaaa aataaagget tgggcgctag catcatcaat	420
atgagcagta ttttcgggat ggtaggcgat ccgaactaccg gggcatacaa cgttccaag	480

```

ggggcgggtac gtatcatgtc gaaaagcgca gcgctggatt ggcactgaa ggactacgat      540
gtgcgtgtca acacagtaca tcgggceccc atcaagaccc cgatgetcga tgattatgaa      600
gggtcgtgagg aaatgttgtc acagcgtacg aaaacccta tgggceacat tggcgaaccg      660
aatgacatcg catgggtctg tgtgtacctg gcattctgacg aatcgaaatt tgcgacgggt      720
gcagaatttg tggtcgacgg cgggtggacc gcacag                                756

```

<210> 52

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 52

[0040]

```

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

```

<400> 54
Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
1 5 10 15
Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
20 25 30
Lys Val Val Ile Thr Gly Arg His Ala Asp Val Gly Glu Lys Ala Ala

	35	40	45	
	Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala			
	50	55	60	
	Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala			
	65	70	75	80
	Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Gly Val Val			
	85	90	95	
	Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser			
	100	105	110	
	Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg			
	115	120	125	
	Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile			
	130	135	140	
	Phe Gly Met Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys			
	145	150	155	160
	Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu			
	165	170	175	
	Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys			
	180	185	190	
	Thr Pro Met Leu Asp Asp His Glu Gly Ala Glu Glu Met Met Ser Gln			
	195	200	205	
[0042]	Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala			
	210	215	220	
	Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly			
	225	230	235	240
	Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln			
	245	250		

<210> 55

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 55

atgaccgatac gtctgaaggg caaagtagcc atcgtaaccg gcgggacaca gggatcgggt	60
ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggctcgtct	120
gcggatgtag gtgaaaaggc cgcgaatca atcgggcgca ctgatgttat tegtctgtc	180
cagcacgatg cgtccgatga agcaggtgg acgaaactgt tgcacaccac cgaggaggca	240
ttcggcccggt ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa	300
gacactacca cggaggaatg gcgtaaaactg ctgtccgtta atctggatgg tgttttttc	360

```

ggcacccgctc tgggcattca gcgcattgaaa aataaagget tgggegetag catcateaat 420
atgagcagta ttttcgggat ggtaggcgat cegactaccg gggcatacaa cgetteccaag 480
ggggcgggtac gtatcatgtc gaaaagegca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtea acacagtaca tccgggcccc atcaagacc ccatgctcga tgattatgaa 600
gggtcgtgagg aatgttttc acagcgtaacg aaaaccacct tgggtcacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacag 756

```

<210> 56

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 56

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

[0043]

Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

<210> 57

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 57

[0044]

```

atgaccgatac gtctgaaggg eaaagtagcc atcgtaaccg ggggacacg gggatcgggt      60
ttggcaatcg cegataaatt tctagaggag ggtgcgaaag tagttattac tggtcgtcgt      120
gcggatgtag gtgaaaagge cgccaaatca atcggcggea ctgatgttat tgcctttgtc      180
cagcaogatg cgtccgatga agcaggetgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggcccggt ttaegacegt cgtgaacaat gcagggaattg gggttgttaa aagegttgaa      300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttcc      360
ggcaccgctc tgggcattca gcgcattgaaa aataaaggct tgggcgctag catcataaat      420
atgagcagta ttttcgggat ggtaggcgat cegactaccg gggcatacaa cgettcctaa      480
ggggcggtac gtatcatgtc gaaaagcgca gcctgggatt gcgcactgaa ggactacgat      540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgattatgaa      600
ggtgctgagg aaatgttttc acagegtacg aaaaccctta tgggtcactat tggcgaaccg      660
aatgacatcg catgggtctg tgtgtacctg geatctgaag aatcgaaatt tgcgaegggt      720
gcagaatttg tggtcgacgg cgggtggacc gcacag                                756

```

<210> 58

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 58

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

	20	25	30
	Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala		
	35	40	45
	Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala		
	50	55	60
	Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala		
	65	70	75
	Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Gly Val Val		
	85	90	95
	Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser		
	100	105	110
	Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg		
	115	120	125
	Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile		
	130	135	140
	Phe Gly Met Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys		
	145	150	155
	Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu		
	165	170	175
	Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys		
	180	185	190
[0045]	Thr Pro Met Leu Asp Asp Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln		
	195	200	205
	Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala		
	210	215	220
	Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly		
	225	230	235
	Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln		240
	245	250	

<210> 59

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 59

atgaccgata gtctgaaggc caaagtagcc atcgtaaccg ggggacact gggatcggc	60
ttggcaatcg ccgaiaaatt ttaggggag ggtgcgaaag tagttattac tggctcgtcgt	120
gcggatgtag gtgaaaagcc cgcgaatca atcgggcgca ctgatgttat tcgctttgtc	180
cagcagcatg cgccgatga agcaggctgg cctaaactgt tcgacaccac cgaggaggca	240

```

ttcggccccg ttaacaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa    300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttct    360
ggcacccgtc tgggcattca ggcgatgaaa aataaagget tgggegetag catcatcaat    420
atgagcagta ttttcgggat ggtaggcgat ccgaactaccg gggcatacaa cgtttccaag    480
ggggcggtac gtatcatgtc gaaaagcgca gcgttgatt gegcactgaa ggactacgat    540
gtgcgtgtea acacagtaca tccgggcccc atcaagaccc cgatgetcga tgattatgaa    600
gggtgctgagg aaatgttttc acagegtacg aaaaccccta tgggtcacat tggcgaaccg    660
aatgacatcg catgggtctg tgtgtacctg gcattctgacg aategaaatt tgcgacgggt    720
gcagaatttg tggtcgacgg cgggtggacc gcacag                                756

```

<210> 60

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 60

[0046]

```

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

```

Thr Pro Met Leu Asp Asp Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

<210> 61
 <211> 756
 <212> DNA
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

[0047] <400> 61
 atgaccgate gtctgaaggc caaagtagcc atcgtaaccg geggacact gggatcggt 60
 ttggcaatcg ccgataaatt tctagaggag ggtgagaaag tagttattac tggctgctgt 120
 gcgatgtag gtgaaaagcg cgcacaaatc atcggegga ctagatgttat tegtctgtc 180
 cagcagatg cgtccgatga agcaggctgg acgaaactgt tcgacacac caggaggca 240
 ttgggcgcgg ttacgaccgt cgtgaacaat gcagggattg ctgttggtta aagcgttgaa 300
 gacactacca eggaggaaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
 ggcaccgctc tgggcattea gcgcattgaa aataaaggct tgggcgctag catcatcaat 420
 atgagcagta ttttcgggat ggtaggcgat ccgactaccg gggcatacaa cgttccaaag 480
 ggggcggta cgtatcatgt gaaaagcgca gcctggatt gcgcactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgtatctcga tgattatgaa 600
 ggtgctgagg aaatgttttc acagcgtacg aaaaccctta tgggtcacat tggcgaaccg 660
 aatgacatcg catgggtctg tgtgtacctg gcattctgac aatcgaaatt tgcgacgggt 720
 gcagaatttg tggctgacgg cgggtggacc gcacag 756

<210> 62
 <211> 252
 <212> PRT
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 62
 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr

1 5 10 15
 Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
 50 55 60
 Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80
 Phe Gly Pro Val Thr Val Val Asn Asn Ala Gly Ile Ala Val Val
 85 90 95
 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
 100 105 110
 Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg
 115 120 125
 Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
 130 135 140
 Phe Gly Met Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys
 145 150 155 160
 Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu
 165 170 175
 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys
 [0048] 180 185 190
 Thr Pro Met Leu Asp Asp Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

<210> 63

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 63

atgaccgatac gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggatcggt 60
 ttggcaatcg ccgataaatf ttagtaggag ggtgagaaag tagttattac tggctcgctg 120

```

gcggatgtag gtgaaaagge cgccaaatca atcggeggca ctgatgttat tcgttttgtc 180
cagcaogatg cgtccgatga agcaggetgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggccegg ttacgaccgt cgtgaacaat gcagggatta gtgttgtaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcaccegtc tgggcattca ggcgatgaaa aataaaggct tgggegetag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgactaccg gggcatacaa cgtttccaag 480
ggggcgggtac gtatcatgtc gaaaagcgea gcgtcggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgattatgaa 600
gggtcgtgagg aaatgttttc acagcgtacg aaaaccccta tgggtcacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcattcgaag aatcgaaatt tgcgacgggt 720
gcgaattttg tggtcgacgg cgggtggacc gcacag 756

```

<210> 64

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

[0049]

<400> 64

```

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

```

Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys
 180 185 190
 Thr Pro Met Leu Asp Asp Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

<210> 65

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 65

[0050] atgaccgatac gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggatateggt 60
 ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggctgctgt 120
 gcggatgtag gtgaaaaggc cgcgaaatca atcggcgga ctagatgttat tctgtttgtc 180
 cagcacgatg cgtccgatga agcaggettg acgaaactgt tgcacaccac cgaggaggca 240
 ttcgccccgg ttacgaccgt cgtgaacaat gcagggtatt gggttgttaa aagcgttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgta aactggatgg tgttttttc 360
 ggcaaccgtc tgggcattca ggcgatgaaa agtaaagget tgggcgtag catcatcaat 420
 atgagcagta ttttcgggat ggtaggcgat ccgactaccg gggeatacaa cgtttccaag 480
 ggggcggta cgtatcatgtc gaaaagcgca gcgttgatt gcgcactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcccc atcaagacce cgatgctcga tgattatgaa 600
 ggtgctgagg aaatgttttc acagcgtacg aaaaccccta tgggtcacat tggcgaaaccg 660
 aatgacatcg catgggtctg tgtgtacctg gcactgtacg aatcgaaatt tgcgacgggt 720
 gcagaatttg tggtegacgg cgggtggacc gcacag 756

<210> 66

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 66
 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
 50 55 60
 Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80
 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Gly Val Val
 85 90 95
 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
 100 105 110
 Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg
 115 120 125
 Met Lys Ser Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
 130 135 140
 Phe Gly Met Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys
 145 150 155 160
 Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu
 [0051] 165 170 175
 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys
 180 185 190
 Thr Pro Met Leu Asp Asp Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

<210> 67

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 67

```

atgaccgatac gtctgaaggg caaagtagec atcgtaaccg gggggacact gggatcggt      60
ttggcaatcg cggataaatt tctagaggag ggtgcgaaag tagttattac tggtegtcgt      120
gcggatgtag gtgaaaagge cggcaaatca atcggcggca ctgatgttat tcgttttgc      180
cagcaccgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggtcccg ttaacaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa      300
gacactacea cggaggaaatg gcgtaaacgt ctgtccgtta atctggatgg tgttttttc      360
ggcaccgctc tgggcaattca gcgcataaaa aataaaggct tgggcgctag catcataaat      420
atgagcagta ttttcgggat ggtaggetat ccgaactacc gggcatacaa cgtttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat      540
gtgcgtgtca acacagtaca tccgggcccc atcaagacc ccatgctoga tgattatgaa      600
ggtgctgagg aaatgttttc acagcgtacg aaaacccta tgggtcacat tggcgaaccg      660
aatgacatcg catgggtctg tgtgtacctg gcatttgacg aatcgaaatt tgcgacgggt      720
gcagaatttg tggtcgacgg cgggtggacc gcacag      756

```

<210> 68

<211> 252

<212> PRT

<213> 人工序列

<220>

[0052]

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 68

```

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

```

Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu
 165 170 175
 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys
 180 185 190
 Thr Pro Met Leu Asp Asp Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

<210> 69
 <211> 756
 <212> DNA
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

[0053]

<400> 69
 atgaccgatac gctgaaggc caaagtagcc atcgtaaccg gcgggacact gggatateggt 60
 ttggcaatcg cegataaatt ttagaggag ggtgcgaaag tagttattac tggctgctgt 120
 gcggatgtag gtgaaaagc cgcgaatca atcgggcgca ctgatgttat tcgtttgtgc 180
 cagcacgatg cgtccgatga ageaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 ttggcccggt ttaagaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttgc 360
 ggcaccgctc tgggcattca gcgcattgaa aataaaggct tgggcgctag catcatcaat 420
 atgagcagta ttttcgggat ggtaggcgat cgggggaccg gggcatacaa cgtttccaag 480
 ggggcgggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tcggggcccc atcaagaccc cgatgctcga tgattatgaa 600
 ggtgctgagg aatgttttc acagcgtacg aaaaccctta tgggtcaccat tggcgaaccg 660
 aatgacatcg catgggtctg tgtgtacctg gcatttgacg aatcgaaatt tgcgacgggt 720
 gcagaatttg tggtcgacgg cgggtggacc gcacag 756

<210> 70
 <211> 252
 <212> PRT
 <213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 70

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

<210> 71

<211> 756

<212> DNA

〈213〉 人工序列

 $\langle 220 \rangle$

〈223〉 克菲尔乳杆菌酮还原酶的工程化变体

<400> 71
 atgaccgatac gtctgaaggg caaagtagcc atcgtaaccg ggggacact gggatcggt 60
 ttggcaatcg ccgataaatt ttagaggag ggtgcgaaag tagttattac tggctcgctg 120
 gggatgtag gtgaaaaggc cgccaaatec atcggeggca ctgatgttat tcgtttgtc 180
 cagcacgatg cgtccgatga agcaggettg acgaaactgt tcgacaccac cgaggaggca 240
 ttgggccggg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
 ggcaccgctc tgggcattca gcgcataaaa aataaaggct tgggcgctag catcatcaat 420
 atgagcagta ttttcgggat ggtaggegat ccgactaccg gggcatactg tgettccaag 480
 ggggcggtac gtatcatgtc gaaaagcgca gcgtggatt gcgcactgaa ggaetacgat 540
 gtgcgtgtca acacagtaca tccgggcccc atcaagacct cgatgctcga tgattatgaa 600
 ggtgctgagg aaatgttttc acagcgtacg aaaaccccta tgggtcacat tggegaaccg 660
 aatgacatcg catgggtctg tgtgtacctg gcactgacg aatcgaaatt tgcgacgggt 720
 gcagaatttg tggtegacgg cgggtggaac gcacag 756

<210> 72
 <211> 252
 <212> PRT
 <213> 人工序列

[0055] <220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 72
 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
 50 55 60
 Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80
 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Gly Val Val
 85 90 95
 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
 100 105 110
 Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg
 115 120 125
 Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
 130 135 140

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

<210> 73
 <211> 756
 <212> DNA
 <213> 人工序列

[0056]

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 73
 atgaccgatac gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggatcgggt 60
 ttggcaatcg cegataaatt tctagaggag ggtcgaaaag tagttattac tggtcgtcgt 120
 gcggatgtag gtgaaaagge cgccaaatca atcgggcgca ctgatgttat tgcctttgtc 180
 cagcacgatg cgtccgatga agcaggetgg acgaaactgt tcgacaccac cgaggaggea 240
 ttccggcccg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagegttgaa 300
 gacaetacca cggaggaatg gcgtaaaactg ctgtccgtta atctggatgg tgttttttcc 360
 ggcaaccgtc tgggcattca gcgcattgaa aataaaggct tgggcgctag catcateaat 420
 atgagcagta ttttcgggat ggtaggcgat cgaactaccg gggeatacaa cgttccaaag 480
 ggggcgggtac gtatcatgtc gaaaagcgca gcgctggatt gcgactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgggtctcga tgattatgaa 600
 ggtgctgagg aaatgttttc acagcgtacg aaaaccctca tgggteacat tggcgaaccg 660
 aatgacatcg catgggtctg tgtgtacctg geatctgacg aatcgaaatt tgcgacgggt 720
 gcagaatttg tggtegaagg cgggtggacc gcacag 756

<210> 74
 <211> 252
 <212> PRT
 <213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 74

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

[0057]

<210> 75

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 75

```

atgaacgate gletgaaggg caaagiagcc ategtaaccg gegggacact gggatateggt    60
ttggcaateg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtegtcgt    120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc    180
cagcaogatg cgtccgatga agcaggettg acgaaactgt tcgacaccac cgaggaggca    240
ttcgccccgg ttacgacgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa    300
gacactacca cggaggaatg gcgtaaactg ctgtccatta atctggatgg tgttttttc    360
ggcaccgcgc tgggcattca gcgcataaaa aataaaggct tgggcgctag catcateaat    420
atgagcagta ttttcgggat ggtaggegat ccgactaccg gggcatacaa cgttccaaag    480
ggggcgggtac gtatcatgtc gaaaagcgca gcgciggtatt gcgcactgaa ggactacgat    540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgcttga tgattatgaa    600
gggtcgtagg aaatgttttc acagcgtacg aaaaccctta tgggtcacat tggegaaccg    660
aatgacateg catgggtctg tgtgtacctg gcctctgacg aatcgaaatt tgcgaagggt    720
gcagaatttg tggtcgacgg cgggtggacc gcacag                                756

```

<210> 76

<211> 252

<212> PRT

<213> 人工序列

[0058]

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 76

```

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

```

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

<210> 77

<211> 756

<212> DNA

<213> 人工序列

[0059]

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 77

atgaccgatac gtctgaaggc caaagtagcc atcgtaaccg ggggacact gggatcggt 60
 ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggtegtgt 120
 gcggatgtag gtgaaaaggc cgcgaatca atcgccgcca ctgatgttat tgcctttgtc 180
 cagcagcatg cgtccgatga agcaggttg acgaaactgt tcgacaccac cgaggaggca 240
 ttccggcccg itacgaccgt cgtgaacaat gcagggttg gggttgttaa aagcgttgaa 300
 gacactacca cggaggaaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
 ggcaccgctc tgggcattca gcgcattgaa aataaaggt tgggcgctag catcatcaat 420
 atgagcagta ttttcgggat ggtaggcgat ccgaactacc gggcatacaa cgttccaaag 480
 ggggcggta cgtatcatgtc gaaaagcgca gcctgggatt gcgcacfgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggccc atcaagacce cgaatgatga tgattatgaa 600
 ggtgtgtagg aaatgttttc acagcgtacg aaaaccctta tgggtcacat tggcgaaccg 660
 aatgacatcg catgggtctg tgtgtacctg gcattctgac aatcgaaatt tgcgacgggt 720
 gcagaatttg tggtcgacgg cgggtggacc gcacag 756

<210> 78

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 78

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

[0060]

<210> 79

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 79

```

atgaccgatac gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggatcggt      60
ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggtcgtcgt      120
gcggatgtag gtgaaaaggc cgcctaatca atcggcggca ctgatgttat tcgctttgtc      180
cagcagcatg cgtccgatga agcaggetgg acgaaactgt tegacaccac cgaggaggea      240
ttcgcccggt ttacgaccgt cgtgaacaat gcagggtatt gggttgttaa aagcgttgaa      300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttcc      360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcctag catcatcaat      420
atgagcagta ttttcgggat ggtaggegat ccgactaccg gggcatacaa cgttccaag      480
ggggcggtac gtatcatgtc gaaaagegca gcgctggatt gcgcaetgaa ggactacgat      540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgetega tgattatgaa      600
gggtctgagg aaatgttttc acagcgtacg aaaaccccta tgggtcacat tggcgaaccg      660
aatgacatcg catgggtctg tgtgtacctg gcctctgaac aatcgaaatt tgcgacgggt      720
gcagaatttg tggtcgacgg cgggggtacc gcacag      756

```

[0061]

<210> 80

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 80

```

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

```

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

[0062]

<210> 81

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 81

```

atgaccgatc gtcftgaaggg caaagtagcc atcgtaaccg gegggacaat gggatcgggt      60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggctcgtcgt      120
gcggatgtag gtgaaaaggc cgcgaaatec atcggcggca ctgatgttat tcgctttgtc      180
cagcacgatg ttccgatga agcaggctgg acgaaactgt tegacaccac cgaggaggea      240
ttcgccccgg ttacgacctt agtgaacaat gcagggattg gggtttaca aagcgttgaa      300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc      360
ggcaccgctc tgggcattca gcgatgaaa aataaagget tgggcgetag catcatcaat      420
atgagcagta ttttcgggct ggtaggcgat ccgactaccg gggcatacaa cgcctccaag      480
ggggcggtac gtatcatgtc gaaaagcgea gcgatggatt gcgcactgaa ggactacgat      540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgatcatgaa      600
gggtgetgagg aaatgttttc acagcgtaacg aaaaccccta tgggtcacat tggcgaaccg      660
aatgacatcg catgggtctg tgtgtacctg geatctgaac aatcgaaatt tgcgacgggt      720
gcagaatttg tggtcgacgg aggggtggacc gcacag                                756

```

<210> 82

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 82

Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Met Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Val
 50 55 60
 Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80
 Phe Gly Pro Val Thr Thr Leu Val Asn Asn Ala Gly Ile Gly Val Tyr
 [0063] 85 90 95
 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
 100 105 110
 Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg
 115 120 125
 Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
 130 135 140
 Phe Gly Leu Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys
 145 150 155 160
 Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu
 165 170 175
 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys
 180 185 190
 Thr Pro Met Leu Asp Asp His Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

<210> 83
 <211> 756
 <212> DNA
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 83
 atgacogatc gtctgaaggg caaagtagec atcgtaaccg gegggacaat gggatcgggt 60
 ttggcaatcg cegataaatt tglagaggag ggtgcgaaag tagttattac tggtcgtcgt 120
 gcggatgtag gtgaaaagge cgccaaatca atcggeggea ctgatgttat tgcgtttgtc 180
 cagcacgatg cgtccgatga agcaggctgg acgaaactgt tegacaccac cgaggaggca 240
 ttcgcccggt ttaegacett agtgaacaat gcagggaattg gggttgttaa aagegttgaa 300
 gacactacca cggagggaatg gegtaaacgt ctgtccgtta atctggatgg tgttttttcc 360
 ggcaccggtc tgggcattca gcgcattgaaa aataaagget tgggcgctag catcataaat 420
 atgagcagta ttttcgggct ggtaggcgat cegactaccg gggcatacaa cgttccaag 480
 ggggcgggtac gtatcatgtc gaaaagcgea gcgcgttat gcgcaetgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcccc atcaagacce cgatgctcga tgatcatgaa 600
 ggtgctgagg aaatgttttc acagegtacg aaaaccecta tgggteacat tggcgaaccg 660
 aatgacatcg catgggtctg tgtgtacctg geatetgaeg aatcgaaatt tgcgaegggt 720
 gcagaatttg tggtegaagg cgggtggacc gcacag 756

[0064]

<210> 84
 <211> 252
 <212> PRT
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 84
 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Met Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
 50 55 60
 Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80
 Phe Gly Pro Val Thr Thr Leu Val Asn Asn Ala Gly Ile Gly Val Val

	85	90	95
Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser			
100	105	110	
Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg			
115	120	125	
Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile			
130	135	140	
Phe Gly Leu Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys			
145	150	155	160
Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Leu Cys Ala Leu			
165	170	175	
Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys			
180	185	190	
Thr Pro Met Leu Asp Asp His Glu Gly Ala Glu Glu Met Phe Ser Gln			
195	200	205	
Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala			
210	215	220	
Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly			
225	230	235	240
Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln			
245	250		

[0065]

<210> 85

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 85

atgaccgate gtctgaaggg caaagtagcc atcgtaaccg gegggacaat gggatcgggt	60
ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggctcgtcgt	120
gcggatgtag gtgaaaaggc cgccaaatca atcgcgcgca ctgatgttat tcgctttgtc	180
cagcacgatg ttccgatga ageaggetgg acgaaactgt tcgacaccac cgaggaggca	240
ttcggcccggt ttacgacctt agtgaacaat gcagggattg gggttgttaa aagcgttgaa	300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttct	360
ggcacccgctc tgggcattca gcgcattgaaa aataaaggct tgggcgctag catcatcaat	420
atgagcagta ttctcgggca ggtaggcgat ccgaactaccg gggcatacaa cgttctcaag	480
ggggcggttac gtatcatgtc gaaaagcgca gcgctgttat gcgcactgaa ggactacgat	540
gtgcgtgtca acacagtaca tccgggcccc atcaagacct cgatgctega tgatcatgaa	600
ggtgctgagg aaatgttttc acagcgtacg aaaaccccta tgggtcaccat tggcgaaccg	660
aatgacatcg catgggtctg tctgtacctg gcattctgacg aatcgaaatt tgcgacgggt	720

gcagaatttg tggtegacgg cgggtggacc gcacag

756

<210> 86

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 86

Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Met Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Val
 50 55 60

[0066]

Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80
 Phe Gly Pro Val Thr Thr Leu Val Asn Asn Ala Gly Ile Gly Val Val
 85 90 95
 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
 100 105 110
 Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg
 115 120 125
 Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
 130 135 140
 Phe Gly Gln Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys
 145 150 155 160
 Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Leu Cys Ala Leu
 165 170 175
 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys
 180 185 190
 Thr Pro Met Leu Asp Asp His Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln

245

250

<210> 87

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 87

atgaccgatac gtcitgaaggg caaagtagcc atcgtaaccg gcgggacaat gggatcggt	60
ttggcaatcg ccgataaatt ttagaggag ggtgcgaaag tagttattac tggtcgtcgt	120
gcggaigttag gtgaaaagge cgccaaatca atcggcggca ctgatgttat tcgctttgtc	180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240
ttcgggcccg ttaegacctt agtgaacaat gcagggattg gggttctgaa aagegttgaa	300
gacaetacca cggaggaatg gegttaaactg ctgtcgtta atctggatgg tgtttttt	360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat	420
atgagcagta ttttcgggat ggtaggcgat ccgaactaccg gggcatacaa cgttccaaag	480
ggggcggiac gtatcatgtc gaaaagcgca gcgcgggatt gcgcacigaa ggactacgat	540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccg cgatgctoga tgatcatgaa	600
gggtgctgagg aaatgttttc acagcgtaacg aaaaccctta tgggtcacat tggcgaaccg	660
aatgacateg catgggtctg tgtgtacctg gcattctgacg aatcgaaatt tgcgacgggt	720
gcagaatttg tggtcgacgg cgggtggacc gcacag	756

[0067]

<210> 88

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 88

Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr	
1	5 10 15
Met Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala	
	20 25 30
Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala	
	35 40 45
Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala	
50	55 60

[0068]

Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80
 Phe Gly Pro Val Thr Thr Leu Val Asn Asn Ala Gly Ile Gly Val Leu
 85 90 95
 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
 100 105 110
 Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg
 115 120 125
 Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
 130 135 140
 Phe Gly Met Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys
 145 150 155 160
 Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu
 165 170 175
 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys
 180 185 190
 Thr Pro Met Leu Asp Asp His Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

<210> 89

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 89

atgaccgatac gtcgaaggg caaagtagcc atcgtaaccg gegggacaat gggatcggg 60
 ttggcaatcg ccgataaatt igttagggag ggtgcgaaag tagttattac tggctgtgt 120
 gcggatgtag gtgaaaagge ccgcaaatca atcggeggea ctgatgttat tegettgtc 180
 cagcacgatg cgtctgatga agcaggetgg acgaaactgt tcgacaccac cgaggaggea 240
 ttcggeccgg ttaagaccgt cgtgaacaat gcagggattg gggttctgaa aagcgttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
 ggcaccgctc tgggcattca gcgcattgaaa aataaagget tgggegetag catcatcaat 420
 atgagcagta ttttcgggca ggtaggcgat ccgactaccg gggcatacaa cgttccaag 480
 ggggcggtag gtatcatgtc gaaaagcgea gcctgggatt gcgcaetgaa ggactacgat 540

```

gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgattatgaa 600
gggtgctgagg aaatgttttc acagcgtacg aaaaccccta tgggtcacat tggegaaccg 660
aatgacateg catgggtctg tgtgtacctg gcattetgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacag 756

```

<210> 90

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 90

[0069]

```

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

```

225	230	235	240
Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln			
	245	250	

<210> 91
 <211> 756
 <212> DNA
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 91

atgaccgatc gtcgaaggg caaagtagcc atcgtaaccg gcgggacaat gggatcggt	60
ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggcgtcgt	120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgtttgtc	180
cagcaecatg ttccgatga agcaggctgg acgaaactgt tegacaccac cgaggaggea	240
ttcgccccgg ttacgacctt agtgaacaat gcagggattg gggtttaca aagcgttgaa	300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc	360
ggcacccgtc tgggcattca gcgcattgaaa aataaaggct tgggcgtcag catcatcaat	420
atgagcagta ttttcgggca ggtagcgcat ccgactaccg gggcatacaa cgttccaag	480
ggggcggtac gtatcatgtc gaaaagcgca gcctggatt gcgactgaa ggactacgat	540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgetcga tgateatgaa	600
gggtgtgagg aaatgttttc acagcgtacg aaaaccccca tgggtcact tggcgaaccg	660
aatgacatcg catgggtctg tgtgtacctg gcctctgacg aatcgaaatt tgcgaegggt	720
gcagaatttg tggctgacgg cgggtggacc gcacag	756

[0070]

<210> 92
 <211> 252
 <212> PRT
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 92

Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr	
1	5 10 15
Met Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala	
	20 25 30
Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala	
	35 40 45

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

[0071]

<210> 93

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 93

atgaccgata gctctgaagg caaagtagec atcgtaacgg ggggacaat gggatcggt 60
 ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggtegtcgt 120
 ggggatgtag gtgaaaagge cgcacaaatca atcgggcgca ctgatgttat tctgtttgtc 180
 cagcacgatg cgtccgatga agcaggctgg acgaaaactgt tgcacaccac cgaggaggca 240
 ttccggcccg ttaagaccit agtgaacaat gcagggattg gggttctgaa aagcgttgaa 300
 gacaactacca cggagggaatg gcgtaaaactg ctgtccgtta atctggatgg tgttttttcc 360
 ggcaccgctc tgggcattca ggcgatgaaa aataaaggct tgggcgctag catcatcaat 420

```

atgagcagta ttttcgggca ggtaggcgat ccgactaccg gggcatacaa cgtttccaag 480
ggggcggtac gtatcatgtc gaaaagegca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtea acacagtaca tccgggcccc atcaagacct cgatgetega tgatcatgaa 600
ggtgctgagg aaatgttttc acagcgtacg aaaaccccta tgggtcacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcattctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtegacgg cgggtggacc gcacag 756

```

<210> 94

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 94

[0072]

```

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

```

210	215	220
Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly		
225	230	235
Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln		240
	245	250

<210> 95

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 95

[0073]

atgaccgatac gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggatcggt	60
ttggcaatcg ccgataaatt ttagaggag ggtgcgaaag tagttattac tggctgctgt	120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc	180
cagcacgatg cgtccgatga agcaggetgg acgaaactgt tcgacaccac cgaggaggca	240
ttcgccccgg ttacgaccgt cgtgaacaat gcagggattg gggttctgaa aagegttgaa	300
gacactacca eggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc	360
ggcacccgtc tgggcattca gcgcataaaa aataaagget tgggcgctag caccataaat	420
atgagcagta ttttcgggca agtaggcgat ccgactaccg gggcatacaa cgtttccaag	480
ggggcggtac gtatcatgtc gaaaagegca gcgtgttat gcgactgaa ggactacgat	540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgattatgaa	600
gggtgtgagg aaatgttttc acagcgtacg aaaaccctta tgggtcacat tggcgaaccg	660
aatgacatcg catgggtctg tgtgtacctg gcattctgacg aatcgaaatt tgcgacgggt	720
gcagaatttg tggtegacgg cgggtggacc gcacag	756

<210> 96

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 96

Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr	
1	5 10 15
Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala	
	20 25 30

[0074]

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

<210> 97

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 97

atgaccgata gtctgaaggg caaagtagcc atcgtaaccg ggggacact gggatcggt 60
 ttggcaatcg ccgataaatt ttagagaggag ggtgcgaaag tagttattac tggctcgtgt 120
 gggatgtag gtgaaaaggc cgcgaaatca atcggeggca ctgatgttat tegettgtgc 180
 cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 ttcggeccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300

```

gacactacca eggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcaccogtc tgggcattca gcgcatgaaa aataaagget tgggegetag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat cgaactaccg gggcatacaa cgtttccaag 480
ggggcgggtac gtatcatgtc gaaaagegca gcgetggatt gcgcaetgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagacc cgaatgetga tgatcacgaa 600
ggtgctgagg aaatgttttc acagcgtaacg aaaacccta tgggtcacat tggegaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcactgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacag 756

```

<210> 98

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 98

[0075]

```

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

```

<400> 100
Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
1 5 10 15

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

[0077]

<210> 101

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 101

atgaccgacg gctcgaaggg caaagtagcc atcgtaaccg geggacact gggtatcggg 60
 ttggcaatcg cagataaatt tctagaggag ggtcgaaaag tagttattac tggctcgtgt 120
 gcggatgtag gtgaaaaggc cgcacaaatca atcggcggca ctgatgttat tcgctttgtc 180

```

cagcacgatg cgtccgatga agcaggctgg acgaaactgt tegacaccac cgaggaggea 240
ttcggcccggt ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagegttgaa 300
gacaetacea cggaggaatg gegtaaactg ctgtcgttta atctggatgg tgttttttct 360
ggcacccgctc tgggcattca gcgcatgaaa aataaagget tgggcgctag catcatcaat 420
atgagcagta itttcgggca ggtaggcgat ccgactaccg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagegea gcgctggatt gcgcactgaa ggactacgat 540
gtgctgtgta acacagtaca tccgggcccc atcaagaccc cgatgetega tgattatgaa 600
ggctgtgagg aaatgttttc acagcgtacg aaaaccctta tgggtcacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcctctgaeg aategaaatt tgcgacgggt 720
gcagaatttg tggtegacgg cgggtggacc gcacag 756

```

<210> 102

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 102

[0078]

```

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

```

180	185	190	
Thr Pro Met Leu Asp Asp Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln			
195	200	205	
Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala			
210	215	220	
Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly			
225	230	235	240
Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln			
245	250		

<210> 103

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 103

[0079]

atgaccgalc gtcgaaggg caaagtagcc atcglaaccg geggagacaat gggatcggc	60
ttggcaatcg ccgataaatt ttagaggag ggtgcgaaag tagttattac tggcgcgt	120
gcggatgtag gtgaaaaggc cgcacaatca atcgcgcgca ctgatgttat tcgctttgtc	180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggea	240
ttcggcccgg ttacgacctt agtgaacaat gcagggattg gggtttacia aagcgttgaa	300
gacactacca cggaggaatg gcgtaaaactg ctgtccgtta atctggatgg tgttttttc	360
ggcacccgte tgggcattca gcgcattgaaa aataaaggct tgggcgctag catcateaat	420
atgagcagta ttttcgggat ggtaggcgat ccgactaccg gggcatacaa cgttccaag	480
ggggcggtac gtatcatgtc gaaaagegca gcgtgttat gcgcactgaa ggactacgat	540
gtgcgtgtca acacagtaca tccgggcccc atcaagacce cgatgctcga tgattatgaa	600
ggtgctgagg aatgttttc acagcgtacg aaaaccctta tgggtcacat tggcgaaccg	660
aatgacatcg catgggtctg tgtgtacctg gcattctgacg aatcgaaatt tgcgacgggt	720
gcagaatttg tggtcgacgg cgggtggacc gcacag	756

<210> 104

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 104

[0080]

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

<210> 105

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 105

atgaccgatac gtctgaaggg caaagtagec atcgtaaccg geggacaat gggtatcggt

60

```

ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcgt      120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc      180
cagcacgatg ttccgatga agcaggctgg acgaaactgt tegacaccac cgaggaggca      240
ttcgccccgg ttaogaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa      300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc      360
ggcacccgtc tgggcattca ggcgatgaaa aataaagget tgggcgctag catcatcaat      420
atgagcagta ttttcgggat ggtaggcgat ccgactaccg gggcatacaa cgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgtggatt gcgcaetgaa ggactacgat      540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgatcatgaa      600
gggtgctgagg aaatgttttc acagcgtaac aaaaccacct tgggtcacat tggcgaaccg      660
agtgcacatcg catgggtctg tgtgtacctg gcattctgacg aatcgaaatt tgcgacgggt      720
gcagaatttg tggtcgacgg cgggtggacc gcacag                                756

```

<210> 106

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

[0081]

<400> 106

```

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

```

	165	170	175
Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys			
	180	185	190
Thr Pro Met Leu Asp Asp His Glu Gly Ala Glu Glu Met Phe Ser Glu			
	195	200	205
Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Ser Asp Ile Ala			
	210	215	220
Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly			
	225	230	235
Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln			240
	245	250	

<210> 107

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

[0082]

<400> 107

atgacagatc gtctgaaggg caaagtagcc atcgtaaccg ggggacaat gggatcggt	60
ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggctgctgt	120
gggatgtag gtgaaaagge cgcacaatec atcggeggca ctgatgttat tegtctgtc	180
cagcaagatg cgtccgatga agcaggtg acgaaactgt tcgacaccac cgaggaggca	240
ttcgccccg ttacgacgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa	300
gacactacca cggaggaaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc	360
ggcaccgctc tgggcattca gcgcattgaaa aataaaggt tgggcctag catcatcaat	420
atgagcagta ttttcgggat ggtaggat cgcactaccg gggcatacaa cgtttccaag	480
ggggcggtac gtatcatgtc gaaaagcgca gcctggatt gcgcactgaa ggactacgat	540
gtgcgigtca acacagtaca tccgggcccc atcaagacce cgatgctoga tgattatgaa	600
gggtctgagg aaatgttttc acagcgtacg aaaaccccta tgggtcact tggcgaaccg	660
aatgacatcg catgggtctg tctgtacctg gcctctgacg aatcgaaatt tgcgacgggt	720
gcagaatttg tggctgacgg cgggtggacc gcacag	756

<210> 108

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 108
 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Met Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
 50 55 60
 Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80
 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Gly Val Val
 85 90 95
 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
 100 105 110
 Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg
 115 120 125
 Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
 130 135 140
 Phe Gly Met Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys
 145 150 155 160
 Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu
 165 170 175
 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys
 180 185 190
 Thr Pro Met Leu Asp Asp Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

[0083]

<210> 109

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 109
 atgaccgac gtctgaaggg caaagtagec atcgtaaceg gegggacaat gggatcggt 60
 ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggctgctgt 120
 gcggatgtag gtgaaaaggc cgccaaatca atcggcgga ctgatgttat tcgctttgtc 180
 cagcacgatg ttccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 ttcgccccgg ttacgaccgt cgtgaacaat gcagggaattg gggtttaca aagegttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
 ggcaccgctc tgggcattca gcgcattgaaa aataaaggt tgggcctag catcatcaat 420
 atgagcagta ttttcgggat tctaggcgat ccgactaccg gggcatacaa cgttccaag 480
 ggggcggta cgtatcatgt gaaaagcga gcctgttat gcgactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccggcccc atcaagacc cgtatctga tgatcatgaa 600
 ggtgctgagg aaatgttttc acagcgtacg aaaacccta tgggtcact tggcgaaccg 660
 aatgacatcg catgggtctg tctgtacctg gcatttgaag aatcgaaatt tgcgaagggt 720
 gcagaatttg tggtegaagg cgggtggacc gcacag 756

<210> 110
 <211> 252
 <212> PRT
 <213> 人工序列

[0084]

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 110
 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Met Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Val
 50 55 60
 Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80
 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Gly Val Tyr
 85 90 95
 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
 100 105 110
 Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg
 115 120 125
 Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
 130 135 140
 Phe Gly Ile Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys

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

<210> 111
 <211> 756
 <212> DNA
 <213> 人工序列

<220>

[0085] <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 111	
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacaat gggatateggt	60
ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggctcgtcgt	120
gcggatgtag gtgaaaaggc egccaaatca atcggcggca ctgatgttat tcgctttgic	180
cagcacgatg cgtccgatga agcaggettg acgaaactgt tcgacaccac cgaggaggca	240
ttcgccccgg ttacgaccgt cgtgaacaat gcagggatgg gggttctgaa aagcgttgaa	300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc	360
ggcacccgtc tgggcattca gcgcataaaa aataaagget tgggcgctag catcatcaat	420
atgagcagta ttttcgggca ggtaggcgat ccgactaccg gggcatacaa cgtttccaag	480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat	540
gtgcgtgtca acacagtaea tccgggcccc atcaagaccg cgatgctcga tgattatgaa	600
gggtctgagg aaatgttttc acagcgtacg aaaaccctta tgggtcacat tggcgaaccg	660
aatgacatcg catgggtctg tgtgtacctg gcattctgacg aatcgaaatt tgcgacgggt	720
gcagaatttg tggtcgacgg cgggtacacc gcacag	756

<210> 112
 <211> 252
 <212> PRT
 <213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 112

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

[0086]

<210> 113

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 113

```

atgaccgate gtctgaaggc caaagtagcc atcgtaaccg ggggacaat gggatcggt      60
ttggcaatcg cegataaatt ttagaggag ggtgcgaaag tagttattac tggctcgct      120
gcggatgtag gtgaaaaggc cgcacaatca atcggcgcca ctgatgttat tcgctttgtc      180
cagcagcatg ttccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcgccccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa      300
gacactacca cggaggaatg gcgtaaactg ctgtccglla atctggatgg tgttttttc      360
ggcaccogtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat      420
atgagcagta ttttcgggca ggtaggegat ccgactaccg gggcatacaa cgtttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcctgtttat gcgcaetgaa ggactacgat      540
gtgcgtgtca acacagtaca tccgggcccc atcaagacce cgatgctega tgatcatgaa      600
ggtgctgagg aatgttttc acagcgtacg aaaacccta tgggtcacat tggcgaaacc      660
aatgacatcg catgggtctg tgtgtacctg gcacttgacg aatcgaaatt tgcgacgggt      720
gcagaatttg tggtcgacgg cgggtggacc gcacag                                756

```

<210> 114

<211> 252

<212> PRT

<213> 人工序列

[0087]

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 114

```

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

```

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

<210> 115

<211> 756

<212> DNA

<213> 人工序列

[0088]

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 115

atgaccgatac gtctgaaggg caaagtagcc atcgtaaccg gcgggacaat gggatateggt	60
ttggcaatcg ccgataaatt ttagaggag ggtgcgaaag tagttattac tggctegctgt	120
gcggatgtag gtgaaaaggc cgcgaaatca atcgggcgca ctgatgttat tgcctttgtc	180
cagcaecatg ttccgatga agcaggctgg acgaaactgt tgcacaccac cgaggaggca	240
ttcgccccgg ttacgacctt agtgaacaat gcagggattg gggtttaca aagcgttgaa	300
gacaactacca cggaggaaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc	360
ggcacccgtc tgggcattca gcgcataaaa aataaaggct tgggcctag catcatcaat	420
atgagcagta ttttcgggat ttaggcgat ccgactaccg gggcatacaa cgettecaag	480
ggggcggtac gtatcatgtc gaaaagcgca gcctgttat gcgcactgaa ggactacgat	540
gtgcgtgtea acacagtaca tccgggcccc atcaagaccc cgatgctcga tgattatgaa	600
ggtgctgagg aaatgttttc acagcgtacg aaaaccctta tgggtccatc tggcgaccg	660
aatgacatcg catgggtctg tgtgtacctg gcactgtacg aatcgaaatt tgcgacgggt	720
gcagaatttg tggtcgacgg cgggtggacc gcacag	756

<210> 116

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 116

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

[0089]

<210> 117

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 117

```

atgaccgate gtctgaaggg caaagtagcc atcgtaaccg gegggacact gggatcgggt      60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtegtcgt      120
geggaatgtag gtgaaaaggc cgcgaaatca atcggcggca ctgatgttat tcegtttgtc      180
cagcacgatg ttccgatga agcaggctgg acgaaactgt tegacaccac cgaggaggca      240
ttcggeccgg ttacgacctt agtgaacaat gcagggattg gggtttaca aagcgttgaa      300
gacactacca cggagggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc      360
ggcacccgtc tgggcattca gcgcattgaaa aataaagget tgggcgctag catcatcaat      420
atgagcagta ttttcgggca ggtaggcgat ccgactaccg gggcatacaa cgtttccaag      480
ggggcggtac gtatcatgtc gaaaagegca gcgtcggatt gcgcactgaa ggactacgat      540
gtgcgtgtca acacagtaca tccgggcccc atcaagacce cgatgctega tgatcatgaa      600
gggtcgtgagg aaatgttttc acagcgtacg aaaaccctta tgggtcacat tggcgaaccg      660
aatgacateg catgggtctg tgtgtacctg gcactcgacg aatcgaaatt tgcgacgggt      720
gcgaatttg tggtcgacgg cgggtacacc gcacag                                     756

```

<210> 118

<211> 252

[0090]

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 118

```

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

```

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

<210> 119

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 119

atgacogacg	gtctgaaggg	caaagtagcc	atcgtaaccg	geggacaaat	gggtatcggt	60
ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttattac	tggtcgtegt	120
geggatgtag	gtgaaaaggc	cgccaaatca	atcgcgggca	ctgatgttat	tcgctttgtc	180
cagcagcatg	ittccgatga	agcaggctgg	acgaaactgt	tcgacaaccac	cgaggaggca	240
ttcggcccgg	ttacgaccgt	cgtgaacaat	gcagggattg	gggttggtta	aagcgttgaa	300
gacaactacca	eggaggaatg	gcgtaaactg	ctgtccgtta	atctggatgg	tgtttttttc	360
ggcacccgte	tgggcattca	gcgcattgaa	aataaaggct	tgggcgctag	catcataaat	420
atgagcagta	ttttcgggca	ggtaggcgat	ccgactaccg	gggcatacaa	cgtttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gccttggtat	gcgcaetgaa	ggactacgat	540
gtgcgtgtca	acacagtaca	tcggggcccc	atcaagaccc	cgatgetcga	tgatcatgaa	600
ggtgetgagg	aaatgttttc	acagegtacg	aaaaccacct	tgggtcacat	tggcgaccgg	660
aatgacateg	catgggtctg	tgtgtacctg	gcattctgacg	aatcgaaatt	tgcgacgggt	720
gcagaatttg	tggtcgacgg	cgggttcacc	gcacag			756

<210> 120

<211> 252
 <212> PRT
 <213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 120

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

[0092]

<210> 121

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 121

```

atgaccgacg gtctgaaggg caaagtagcc atcgtaaccg gggggacaat gggatcgggt    60
ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggtcgtcgt    120
gggatgtag gtgaaaaggc cgcacaatca atcgggcgca ctgatgttat tegtcttgc    180
cagcacgatg ttccgatga agcaggetgg acgaaactgt tcgacaccac cgaggaggca    240
ttcggeccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa    300
gacaactacca cggaggaaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttt    360
ggcaccgctc tgggcattca gcgcattgaaa aataaagget tgggcctag cateatcaat    420
atgagcagta itttcgggca gctaggggat ccgactaccg gggcatacaa cgettcgaag    480
ggggcggtae gtatcatgtc gaaaagegca gcctgttat gcgcactgaa ggactacgat    540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatctcga tgattatgaa    600
ggctctgagg aaatgttttc acagcgtacg aaaaccctta tgggtcact tggcgaccg    660
aatgacatcg catgggtctg tctgtacctg gcattctgacg aatcgaaatt tgcgacgggt    720
gcagaatttg tggtcgacgg cgggtggacc gcacag                                756

```

[0093]

<210> 122

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 122

```

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

```

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

[0094]

<210> 123

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 123

atgaccgatc gtcftgaaggg caaagtagcc atcgtaaccg gegggacact gggatcgggt	60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggctcgtcgt	120
gcggatgtag gtgaaaaggc cgccaaatec atcggcggca ctgatgttat tcgctttgtc	180
cagcacgatg ttccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggea	240
ttcgccccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa	300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc	360
ggcaccgctc tgggcattca gcgatgaaa aataaagget tgggcgetag catcatcaat	420
atgagcagta ttttcgggca ggtaggcgat ccgactaccg gggcatacaa cgcctccaag	480
ggggcggtac gtateatgtc gaaaagcgea gcgatgttat gcgcactgaa ggactacgat	540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgatcatgaa	600
gggtgetgagg aaatgttttc acagcgtacg aaaaccccta tgggtcacat tggcgaaccg	660
aatgacatcg catgggtctg tgtgtacctg geatctgaag aatcgaaatt tgcgacgggt	720
gcagaatttg tggtcgacgg cgggtggacc gcacag	756

〈210〉	124
〈211〉	252
〈212〉	PRT
〈213〉	人工序列

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400>	124																
Met	Thr	Asp	Arg	Leu	Lys	Gly	Lys	Val	Ala	Ile	Val	Thr	Gly	Gly	Thr		
1				5				10						15			
Leu	Gly	Ile	Gly	Leu	Ala	Ile	Ala	Asp	Lys	Phe	Val	Glu	Glu	Gly	Ala		
				20				25						30			
Lys	Val	Val	Ile	Thr	Gly	Arg	Arg	Ala	Asp	Val	Gly	Glu	Lys	Ala	Ala		
				35				40						45			
Lys	Ser	Ile	Gly	Gly	Thr	Asp	Val	Ile	Arg	Phe	Val	Gln	His	Asp	Val		
				50				55				60					
Ser	Asp	Glu	Ala	Gly	Trp	Thr	Lys	Leu	Phe	Asp	Thr	Thr	Glu	Glu	Ala		
65						70				75					80		
Phe	Gly	Pro	Val	Thr	Thr	Val	Val	Asn	Asn	Ala	Gly	Ile	Gly	Val	Val		
				85						90				95			
Lys	Ser	Val	Glu	Asp	Thr	Thr	Thr	Glu	Glu	Trp	Arg	Lys	Leu	Leu	Ser		
				100				105						110			
Val	Asn	Leu	Asp	Gly	Val	Phe	Phe	Gly	Thr	Arg	Leu	Gly	Ile	Gln	Arg		
				115				120						125			
Met	Lys	Asn	Lys	Gly	Leu	Gly	Ala	Ser	Ile	Ile	Asn	Met	Ser	Ser	Ile		
				130				135				140					
Phe	Gly	Gln	Val	Gly	Asp	Pro	Thr	Thr	Gly	Ala	Tyr	Asn	Ala	Ser	Lys		
145						150				155					160		
Gly	Ala	Val	Arg	Ile	Met	Ser	Lys	Ser	Ala	Ala	Leu	Leu	Cys	Ala	Leu		
				165						170					175		
Lys	Asp	Tyr	Asp	Val	Arg	Val	Asn	Thr	Val	His	Pro	Gly	Pro	Ile	Lys		
				180				185						190			
Thr	Pro	Met	Leu	Asp	Asp	His	Glu	Gly	Ala	Glu	Glu	Met	Phe	Ser	Gln		
				195				200						205			
Arg	Thr	Lys	Thr	Pro	Met	Gly	His	Ile	Gly	Glu	Pro	Asn	Asp	Ile	Ala		
				210				215				220					
Trp	Val	Cys	Val	Tyr	Leu	Ala	Ser	Asp	Glu	Ser	Lys	Phe	Ala	Thr	Gly		
225						230				235					240		
Ala	Glu	Phe	Val	Val	Asp	Gly	Gly	Trp	Thr	Ala	Gln						
				245						250							

<210> 125
 <211> 756
 <212> DNA
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 125
 atgacogate gtctgaaggg caaagtagec ategtaaccg gegggacaat gggatcgggt 60
 ttggcaatcg cegataaatt tglagaggag ggtgcgaaag tagttattac tggcgcgcgt 120
 gcggatgtag gtgaaaagge cgccaaatca atcggeggea ctgatgttat tegtittgtc 180
 cagcacgatg tttccgatga agcaggctgg acgaaactgt tegacaccac cgaggaggca 240
 ttccggcccg ttaegacegt cgtgaacaat gcagggattg gggttgttaa aagegttgaa 300
 gacactacca cggagggaatg gegtaaacig ctgtccgtta atctggatgg tgttttttcc 360
 ggcacccgte tgggcattca gcgcattgaaa aataaagget tgggcgctag catcataaat 420
 atgagcagta ttttcgggct ggtaggcgat ccgactaccg gggcatacaa cgttccaag 480
 ggggcggtag gtatcatgtc gaaaagcgca gcctggatt gcgcaetgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcccc atcaagacce cgatgctcga tgattatgaa 600
 ggtgctgagg aaatgttttc acagegtacg aaaaccecta tgggteacat tggcgaaccg 660
 aatgacateg catgggtctg tgtgtacctg geatetgaeg aatcgaaatt tgcgaegggt 720
 gcagaatttg tggtegaegg cgggtggacc gcacag 756

[0096]

<210> 126
 <211> 252
 <212> PRT
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 126
 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Met Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Val
 50 55 60
 Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80
 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Gly Val Val

	85	90	95
Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser			
100	105	110	
Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg			
115	120	125	
Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile			
130	135	140	
Phe Gly Leu Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys			
145	150	155	160
Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu			
165	170	175	
Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys			
180	185	190	
Thr Pro Met Leu Asp Asp Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln			
195	200	205	
Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala			
210	215	220	
Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly			
225	230	235	240
Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln			
245	250		

[0097]

<210> 127

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 127

atgaccgate gtctgaaggg caaagtagcc atcgtaaccg gegggacaat gggatcgggt	60
ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggtcgtcgt	120
gcggatgtag gtgaaaaggc cgcacaatca atcgggcgca ctgatgttat tcgctttgtc	180
cagcacgatg cgtccgatga ageaggetgg acgaaactgt tcgacaccac cgaggaggca	240
ttcggcccgg ttacgacctt agtgaacaat gcagggattg gggttgttaa aagcgttgaa	300
gacactacca cggagggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttct	360
ggcacccgctc tgggcattca gcgcatgaaa aataaaggct tgggcgetag catcatcaat	420
atgagcagta ttctcgggca ggtaggcgat ccgaactaccg gggcatacaa cgttctcaag	480
ggggcggttac gtatcatgtc gaaaagcgca gcgtggatt gcgcaactgaa ggactacgat	540
gtgcgtgtca acacagtaca tccgggcccc atcaagacct cgatgctega tgatcatgaa	600
ggtgctgagg aaatgttttc acagcgtaacg aaaaccccta tgggtcacaat tggegaaccg	660
aatgacateg catgggtctg tctgtacctg gcatctgacg aatcgaaatt tgcgaagggt	720

gcagaatttg tggtegacgg cgggtggacc gcacag

756

<210> 128

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 128

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

[0098]

<210> 129
 <211> 756
 <212> DNA
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 129
 atgaccgatac gtctgaaggg caaagtagcc atcgtaaccg ggggacaat gggatcggt 60
 ttggcaatcg ccgataaatt ttagaggag ggtgcgaaag tagttattac tggctgctgt 120
 gcggatgtag gtgaaaaggc cgccaaatca atcggeggca ctgatgttat tcgctttgtc 180
 cagcacgatg ttccgatga agcaggettg acgaaactgt tcgacaccac cgaggaggca 240
 ttggccccg ttacgaccgt cgtgaacaat gcagggattg gggttctgaa aagcgttgaa 300
 gacaactacca cggaggaatg gegttaaactg ctgtccgtta atctggatgg tgttttttc 360
 ggcacccgtc tgggcattca gcgcataaaa aataaaggct tgggcgctag catcatcaat 420
 atgagcagta ttttcgggat ttaggcgat ccgactaccg gggcatacaa cgttccaag 480
 ggggcggtac gtatcatgtc gaaaagcga gcgctggatt gcgcactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tcgggcccc atcaagacce cgatgctega tgatcatgaa 600
 ggtgctgagg aaatgttttc acagcgtaacg aaaaccccta tgggtcacat tggcgaaccg 660
 aatgacatcg catgggtctg tgtgtacctg gcactgaacg aatcgaaatt tgcgacgggt 720
 gcagaatttg tggtcgacgg cgggtggacc gcacag 756

[0099]

<210> 130
 <211> 252
 <212> PRT
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 130
 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Met Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Val
 50 55 60
 Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala

65	70	75	80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Gly Val Leu			
	85	90	95
Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser			
	100	105	110
Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg			
	115	120	125
Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile			
	130	135	140
Phe Gly Ile Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys			
	145	150	155
Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu			
	165	170	175
Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys			
	180	185	190
Thr Pro Met Leu Asp Asp His Glu Gly Ala Glu Glu Met Phe Ser Gln			
	195	200	205
Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala			
	210	215	220
Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly			
	225	230	235
Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln			240
	245	250	

[0100]

<210> 131

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 131

atgaccgatac gtctgaagag caaagtagcc atcgtaaccg ggggacact gggatcggt	60
ttggcaatcg ccgataaatt ttagaggag ggtgcgaaag tagttattac tggtcgtcgt	120
gcggatgtag gtgaaaagc cgcgaatca atcgggcgca ctgatgttat tcgtttatc	180
cagcacgatg tgtccgatga agcaggtgg acgaaactgt tcgacaccac cgaggaggca	240
ttcgccccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa	300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc	360
ggcaccgctc tgggcattcg ccgcatgaaa aataaagct tgggcgttag catcatcaat	420
atgagcagta ttttcgggat ggtaggcgat ccgactaccg gggcatacaa cgtttccaag	480
ggggcggtac gtatcatgtc gaaaagcgca gcgtggatt gcgactgaa ggactacgat	540
gtgcgtgtca acacagtaca tccgggccc atcaagacc cgtatgetga tgattatgaa	600

```

ggtgctgagg aaatgttttc acagegtacg aaaaccccta tgggtcgtat tggogaaccg      660
aatgacatcg catgggtctg tgtgtacctg gcattctgacg aatcgaaatt tgcgacgggt      720
gcagaatttg tgatcgacgg cgggtggacc gcacag                                     756

```

<210> 132
 <211> 252
 <212> PRT
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

[0101]

```

<400> 132
Met Thr Asp Arg Leu Lys Ser Lys Val Ala Ile Val Thr Gly Gly Thr
1           5           10          15
Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
20          25          30
Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
35          40          45
Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Ile Gln His Asp Val
50          55          60
Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
65          70          75          80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Gly Val Val
85          90          95
Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
100         105         110
Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Arg Arg
115         120         125
Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
130         135         140
Phe Gly Met Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys
145         150         155         160
Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu
165         170         175
Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys
180         185         190
Thr Pro Met Leu Asp Asp Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln
195         200         205
Arg Thr Lys Thr Pro Met Gly Arg Ile Gly Glu Pro Asn Asp Ile Ala
210         215         220
Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
225         230         235         240

```

Ala Glu Phe Val Ile Asp Gly Gly Trp Thr Ala Gln
245 250

<210> 133
<211> 756
<212> DNA
<213> 人工序列

<220>
<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 133
atgaccgatc gtctgaaggc caaagtagec atcgtaaccg geggacacg cggatcgggt 60
ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggtegtcgt 120
gcggtatgtag gtgaaaagge cgccaaatca atcggcggca ctgatgttat tcgtttgtc 180
cagcacgatg tatccgatga agcaggctgg cctaaactgt tegacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt agtgaacaat gcagggatc cggtggttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgct tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagtg tgttcgggat tctaggegac ccgaetaccg gggcatactg tcttccaag 480
[0102] ggggcggtac gtatcatgtc gaaaagcgca gcgtggatt gcgcactgaa ggactacgat 540
gtgcgigtca acacagtaca tccgggcccc atcaagacce cgatgatgga tagctatgaa 600
gggtctgagg aatgttttc acagcgtacg aaaaacctta tgggtcactt tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcactcgaag aatcgaaatt tgcgacgggt 720
gcagaatctg tggtegacgg cgggtggacc gcacag 756

<210> 134
<211> 252
<212> PRT
<213> 人工序列

<220>
<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 134
Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
1 5 10 15
Arg Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
20 25 30
Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
35 40 45
Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Val

	50	55	60
	Ser Asp Glu Ala Gly Trp Pro Lys Leu Phe Asp Thr Thr Glu Glu Ala		
	65	70	75
	Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Pro Val Val		80
		85	90
	Lys Ser Val Glu Asp Thr Thr Thr Glu Trp Arg Lys Leu Leu Ser		95
	100	105	110
	Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg		
	115	120	125
	Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Val		
	130	135	140
	Phe Gly Ile Val Gly Asp Pro Thr Thr Gly Ala Tyr Cys Ala Ser Lys		
	145	150	155
	Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu		160
		165	170
	Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys		175
	180	185	190
	Thr Pro Met Met Asp Ser Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln		
	195	200	205
	Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala		
	210	215	220
[0103]	Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly		
	225	230	235
	Ala Glu Ser Val Val Asp Gly Gly Trp Thr Ala Gln		240
		245	250

<210> 135

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 135

atgaccgatc gtctgaaggg caaagtagcc ategtaaccg gegggacacg cggatcgggt	60
ttggcaatcg ccgataaatt thtagaggag ggtgcgaaag tagttattac tggtcgtcgt	120
gcggatgtag gtgaaaaggc cgcacaaatca ateggcggca ctgatgttat tegtcttgc	180
cagcacgatg tatecgatga agcaggetgg cctaaactgt tcgacaccac cgaggaggca	240
ttcgccccgg ttaegacett agtgaacaat gcagggatte cgggtggttaa aagecgttgaa	300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc	360
ggcacccgtc tgggcattca gcgcattgaaa aataaagget tgggcgctag catcatcaat	420
atgagcagtg tgttcgggat ttaggctac ccgaactaccg gggcatactg tgcttccaag	480

```

ggagcgggtac gtatcatgtc gaaaagcgca gcgctggatt ggcactgaa ggactacgat      540
gtgcgtgtca acacagtaca tcgggceccc atcaagaccc cgatgatgga tagctatgaa      600
gggtcgtgagg aaatgttttc acagcgtacg aaaacccta tgggtcacat tggcgaaaccg      660
aatgacatcg catgggtctg tgtgtacctg gcattctgacg aatcgaaatt tgcgacgggt      720
gcagaatttg tggtcgacgg cgggtggacc gcacag                                     756

```

<210> 136

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 136

```

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

```

[0104]

<400> 138
Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
1 5 10 15
Gln Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
20 25 30
Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala

	35	40	45
	Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Val		
	50	55	60
	Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala		
	65	70	75
	Phe Gly Pro Val Thr Thr Leu Val Asn Asn Ala Gly Ile Pro Met Val		80
	85	90	95
	Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser		
	100	105	110
	Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg		
	115	120	125
	Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Val		
	130	135	140
	Phe Gly Ile Val Gly Tyr Pro Thr Thr Gly Ala Tyr Cys Ala Ser Lys		
	145	150	155
	Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu		
	165	170	175
	Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys		
	180	185	190
	Thr Pro Met Met Asp Ser Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln		
	195	200	205
[0106]	Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala		
	210	215	220
	Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly		
	225	230	235
	Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln		240
	245	250	

<210> 139

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 139

atgaccgatac gtctgaaggg caaagtagcc atcgtaaccg gcgggacacg cggatctcgt	60
ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggctcgtcgt	120
gcggatgtag gtgaaaaggc cgcgaatca atcgggcgca ctgatgttat tegtctgtc	180
cagcacgatg tatccgatga agcaggtcgg actaaactgt tcgacaccac cgaggaggca	240
ttcggcccggt ttacgacctt agtgaacaat gcagggatcc cggttgttaa aagcgttgaa	300
gacactacca cggaggaatg gcgtaaaactg ctgtccgtta atctggatgg tgttttttc	360

```

ggcacccgctc tgggcattca gcgcattgaaa aataaagget tgggegetag catcateaat 420
atgagcagtg tgttcgggat ttaggctac cegactaccg gggcatactg tgettecaag 480
ggggcgggtac gtatcatgtc gaaaagegca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtea acacagtaca tccgggcccc atcaagacc ccatgatgga tagctatgaa 600
gggtcgtgagg aaatgttttc acggcgtaag aaaacccta tgggtcacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgaagg cgggtggacc gcacag 756

```

<210> 140

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 140

```

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

```

[0107]

Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

<210> 141
 <211> 756
 <212> DNA
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

[0108] <400> 141
 atgaccgatac gtctgaaggg eaaagtagcc atcgtaaccg gegggaacaca aggtatcggg 60
 ttggcaatcg cegataaatt tctagaggag ggtgcgaaag tagttattac tggtcgctcg 120
 gcggatgtag gtgaaaagge cgccaaatca atcggcggea ctgatgttat tgcctttgtc 180
 cagcaogatg cgtccgatga agcaggetgg acgaaactgt tcgacaccac cgaggaggca 240
 ttccggcccg itacgacgt agtgaacaat gcagggaite egggtgttaa aagegttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttcc 360
 ggcaccgctc tgggcattca gcgcattgaaa aataaaggct tgggcgctag catcataaat 420
 atgagcagtg tgttcgggat tctagcgac cgaactaccg gggcatactg tgettecaag 480
 gggcgcggtac gtatcatgtc gaaaagcgca gcctgggatt gcgcactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgatgga tagctatgaa 600
 ggtgctgagg aaatgttttc acagcgtacg aaaaccctta tgggtcacaat tggcgaaccg 660
 aatgacatcg catgggtctg tgtgtacctg geatctgaag aatcgaaatt tgcgaegggt 720
 gcagaatttg tggtcgacgg cgggtacacc gcacag 756

<210> 142
 <211> 252
 <212> PRT
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 142
 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Gln Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala

	20	25	30
	Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala		
	35	40	45
	Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala		
	50	55	60
	Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala		
	65	70	75
	Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Pro Val Val		
	85	90	95
	Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser		
	100	105	110
	Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg		
	115	120	125
	Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Val		
	130	135	140
	Phe Gly Ile Val Gly Asp Pro Thr Thr Gly Ala Tyr Cys Ala Ser Lys		
	145	150	155
	Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu		
	165	170	175
	Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys		
	180	185	190
[0109]	Thr Pro Met Met Asp Ser Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln		
	195	200	205
	Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala		
	210	215	220
	Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly		
	225	230	235
	Ala Glu Phe Val Val Asp Gly Gly Tyr Thr Ala Gln		240
	245	250	

<210> 143

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 143

atgaccgata gtctgaaggc caaagtagcc atcgtaaccg ggggacaat gggatcggg	60
ttggcaatcg ccgaiaaatt tctagaggag ggtgcgaaag tagttattac tggctcgtcg	120
gcggatgtag gtgaaaagcc cgcgaatca atcgggcgca ctgatgttat tgcctttgtc	180
cagcagcatg tatccgatga agcaggctgg cctaaactgt tcgacaccac cgaggaggca	240

```

ttcggcccg ttaacgacgt agtgaacaat gcagggattc cggtggttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttct 360
ggcaccgctc tgggcattca ggcgatgaaa aataaagget tgggegetag catcatcaat 420
atgagcagtg igticgggat tgtaggcgac ccgaactaccg gggcatactg tgcctccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gccttgatt gcgcactgaa ggactacgat 540
gtgcgtgtea acacagtaca tccgggcccc atcaagaccc cgatgatgga tagctatgaa 600
gggtgctgagg aaatgttttc acagegtacg aaaaccccta tgggtcacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcattctgacg aategaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtacacc gcacag 756

```

<210> 144

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 144

[0110]

```

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

```

Thr Pro Met Met Asp Ser Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Tyr Thr Ala Gln
 245 250

<210> 145

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 145

[0111]

atgaccgate gtctgaaggc caaagtagcc atcgtaaccg geggacaca aggtatcggt 60
 ttggcaatcg ccgataaatt tctagaggag ggtgegaag tagttattac tggctgctgt 120
 gcgatgtag gtgaaaagc cgcacaaatc atcggegga ctagatgttat tegettgtgc 180
 cagcaagatg tatecgatga ageaggtgg cctaaactgt tcgacacac caggaggca 240
 ttggcgcgg ttacgaccgt cgtgaacaat gcagggatc egatgggtta aagcgttgaa 300
 gacactacca eggaggatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
 ggcaccgctc tgggcattea gcgcattgaa aataaaggct tgggcgctag catcatcaat 420
 atgagcagtg tgttcgggat tctaggegac ccgactaccg gggcactatg tgettcacag 480
 ggggcggta cgtatcatgc gaaaagcga gcctgggatt gcgcactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggccc atcaagacc cgaatgatga tgactatgaa 600
 ggtgtcagg aaatgtttc acagcgtacg aaaacccta tgggtcacat tggcgaaccg 660
 aatgacatcg catgggtctg tgtgtacctg gcattcgaag aatcgaaatt tgcgacgggt 720
 gcagaatttg tggtcgacgg cgggtggacc gcacag 756

<210> 146

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 146

Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr

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

[0112]

<210> 147

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 147

atgaccgatac gtctgaaggg caaagtagcc atcgtaaccg gcgggacacg cggatateggt 60
 ttggcaatcg ccgataaatt thtagaggag ggtgcgaaag tagttattac tggtcgtcgt 120

```

gcgatgtag gtgaaaagge cgccaaatca atcggeggca ctgatgttat tcgttttgtc 180
cagcaogatg tatccgatga agcaggetgg actaaactgt tcgacaccac cgaggaggca 240
ttcggecegg ttacgaccgt cgtgaacaat gcagggatgc cgggtggttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg cigtccgtta atctggatgg tgttttttgc 360
ggcaccegtc tgggcattca ggcgatgaaa aataaaggct tgggegetag catcatcaat 420
atgagcagtg tgttcgggca agtaggetac ccgactaccg gggcatactg tgetttcaag 480
ggggcggtag gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgatgga tgactatgaa 600
ggtgctgagg aaatgttttc acagcgtacg aaaaccoccta tgggtcacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcattctgacg aatcgaaatt tgcgacgggt 720
gcgaattttg tggtcgacgg cgggtggacc gcacag 756

```

<210> 148

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

[0113]

<400> 148

```

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

```

Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys
 180 185 190
 Thr Pro Met Met Asp Asp Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

<210> 149
 <211> 756
 <212> DNA
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 149
 [0114] atgaccgatac gtctgaaggg caaagtagcc atcgtaaccg gcgggacaat gggatateggt 60
 ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggctcgtcgt 120
 gcggatgtag gtgaaaaggc cgcgaaatca atcggcgga ctgatgttat tctctttgtc 180
 cagcacgatg tatccgatga agcaggettg cctaaactgt tegacaccac cgaggaggca 240
 ttccgcccgg ttacgaccgt agtgaacaat gcagggatc cgggtggtta aagcgttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
 ggcaaccgtc tgggcattca ggcgatgaaa aataaagget tgggcgctag catcatcaat 420
 atgagcagtg tgttcgggat tctagggcgc ccgactaccg gggeatactg tcttccaag 480
 ggggcggtag gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcccc atcaagacce cgatgatgga tagctatgaa 600
 ggtgtctagg aaatgttttc acagcgtacg aaaaccccta tgggtcacat tggcgaaaccg 660
 aatgacatcg catgggtctg tgtgtacctg gcactctgac aatcgaaatt tgcgacgggt 720
 gcagaatttg tggtegacgg cgggtggacc gcacag 756

<210> 150
 <211> 252
 <212> PRT
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 150
 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Met Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Val
 50 55 60
 Ser Asp Glu Ala Gly Trp Pro Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80
 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Pro Val Val
 85 90 95
 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
 100 105 110
 Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg
 115 120 125
 Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Val
 130 135 140
 Phe Gly Ile Val Gly Asp Pro Thr Thr Gly Ala Tyr Cys Ala Ser Lys
 145 150 155 160
 Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu
 165 170 175
 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys
 180 185 190
 Thr Pro Met Met Asp Ser Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

[0115]

<210> 151
 <211> 756
 <212> DNA
 <213> 人工序列

 <220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

 <400> 151

```

atgaccgate gtctgaagg gaaagtagec atcgtaaccg gggggacacg cggatateggt    60
ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggtegtcgt    120
gcggatgtag gtgaaaagge cgcgaaatca atcggcggca ctgatgttat tcgttttgc    180
cagcagcatg cctcgatga agcaggctgg cctaaactgt tcgacaccac cgaggaggea    240
ttcgcccggt ttaegacctt agtgaacaat gcagggattc cggttgttaa aagcgttgaa    300
gacactacea cggaggaatg gcgtaaacgt ctgtccgtta atctggatgg tgttttttc    360
ggcaccgctc tgggcattca gcgcattgaa aataaaggct tgggcgctag catcateaat    420
atgagcagtg tgttcgggat tctaggetac ccgaactacc gggcatactg tgcctccaag    480
ggggcggtac gtatcatgtc gaaaagcgea gcgcctggatt gcgcactgaa ggactacgat    540
glcgctgtca acacagtlaca tccgggcctt atcaagacc cgaatgatga tgactatgaa    600
gggtctgagg aatgttttc acagcgtacg aaaacccta tgggtcacat tggcgaaccg    660
aatgacatcg catgggtctg tctgtacctg gcattctgac aatcgaaatt tgcgacgggt    720
gcagaatttg tggtcgacgg cgggtggacc gcacag                                756

```

<210> 152

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

[0116]

<400> 152

```

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

```

Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu
 165 170 175
 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys
 180 185 190
 Thr Pro Met Met Asp Asp Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

<210> 153
 <211> 756
 <212> DNA
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

[0117]

<400> 153
 atgaccgatac gctgaaggg caaagtagcc atcgtaaccg gcgggacaaat gggatateggt 60
 ttggcaatcg cegataaatt tctagaggag ggtgcgaaag tagttattac tggctcgtcgt 120
 gcggatgtag gtgaaaagge cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
 cagcacgatg tatecgatga ageaggctgg cctaaactgt tcgacaccac cgaggaggca 240
 ttgcgccegg ttaacaccgt agtgaacaat gcggggatgc cgggtggtta aagcgttgaa 300
 gacactacca cgggggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
 ggcaccgcgc tgggcattca gcgcattgaaa aataaaggct tgggcgctag catcatcaat 420
 atgagcagtg tgttcgggat tctaggctac ccgaactaccg gggcactatg tgcctccaag 480
 ggggcgggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgatgga tgaccatgaa 600
 ggtgctgagg aaatgttttc acagcgtaacg aaaaccctta tgggtcacaat tggcgaaccg 660
 aatgacatcg catgggtctg tgtgtacctg gcattctgacg aatcgaaatt tgcgacgggt 720
 gcagaatttg tggtcgacgg cgggtggacc gcacag 756

<210> 154
 <211> 252
 <212> PRT
 <213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 154

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

[0118]

<210> 155

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 155
 atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg ggggacacg cggtatcggt 60
 ttggcaatcg ccgataaatt ttagaggag ggtgcgaaag tagttattac tggctgctgt 120
 ggggatgtag gtgaaaaggc cgccaaatec atcggeggca ctgatgttat tcgtttgtc 180
 cagcacgatg cgtccgatga agcaggettg actaaactgt tcgacaccac cgaggaggca 240
 ttgggcccgg ttacgaccgt cgtgaacaat gcagggatte cgggtggtta aagcgttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
 ggcaccggtt tgggcattca gcgcataaaa aataaaggct tgggcgctag cateatcaat 420
 atgagcagtg tgttcgggat ttaggegac ccgactaccg gggcatactg tgettccaag 480
 ggggcggtac gtatcatgtc gaaaagcgca gcctggatt gcgcactgaa ggaactcgt 540
 gtgcgtgtca acacagtaca tccgggcccc atcaagacc cgaigatgga tagccatgaa 600
 ggtgtctagg aaatgttttc acagcgtacg aaaaccccta tgggtcacat tggegaaccg 660
 aatgacatcg catgggtctg tgtgtacctg gcctctgacg aatcgaaatt tgcgacgggt 720
 gcagaatttg tggtegacgg cgggtggacc gcacag 756

<210> 156
 <211> 252
 <212> PRT
 <213> 人工序列

[0119] <220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 156
 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Arg Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
 50 55 60
 Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80
 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Pro Val Val
 85 90 95
 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
 100 105 110
 Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg
 115 120 125
 Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Val
 130 135 140

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

<210> 157
 <211> 756
 <212> DNA
 <213> 人工序列

[0120]

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 157
 atgaccgatac gtctgaaggg caaagtagcc atcgtaaccg gcgggacacg cggatcgggt 60
 ttggcaatcg cegataaatt tctagaggag ggtcgaaag tagttattac tggtcgtcgt 120
 gcggatgtag gtgaaaagge cgccaaatca atcggcggea ctgatgttat tgcgtttgtc 180
 cagcagcatg tatccgatga agcaggetgg actaaactgt tcgacaccac cgaggaggea 240
 ttccggcccg ttacgaccgt cgtgaacaat gcagggatte cgatgggtta aagegttgaa 300
 gacaetacca cggaggaatg gcgtaaaactg ctgtccgtta atctggatgg tgttttttcc 360
 ggcaaccgtc tgggcattca gcgcattgaa aataaaggct tgggcgctag catcateaat 420
 atgagcagtg tgttcgggat ggtaggctac cgaactaccg gggeatacaa tgettccaag 480
 ggggcgggtac gtatcatgtc gaaaagcgea gcgctggatt gcgactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcccc atcaagacce cgatgatgga tgattatgaa 600
 ggtgctgagg aaatgttttc acagegtacg aaaaccctta tgggteacat tggcgaaccg 660
 aatgacatcg catgggtctg tgtgtacctg geatetgaag aategaaatt tgcgacgggt 720
 gcagaatttg tggtegaagg cgggtacacc gcacag 756

<210> 158
 <211> 252
 <212> PRT
 <213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 158

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

[0121]

<210> 159

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 159

```

atgaacgate gtetgaaggg caaagtagcc ategtaaccg gegggacaca aggtatcggt      60
ttggcaateg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtegtcgt      120
gcggatgtag gtgaaaaggc egccaaatca ateggcggca ctgatgttat tegctttgtc      180
cageacgatg catccgatga agcaggetgg actaaactgt tegacaccac cgaggaggca      240
ttcggeccgg ttaacacgt agtgaacaat gcagggattc cgatggttaa aagcgttgaa      300
gacactacca cggaggaatg gcglaaactg ctgtccgtta atciggatgg tgttttttc      360
ggcacccgtc tgggcattca gcgcataaaa aataaaggct tgggcgctag catcateaat      420
atgagcagtg tgttcgggca agtaggegac ccgactaccg gggcatacaa tgcctccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgciggatt gcgcactgaa ggactacgat      540
gtgcgtgtca acacagtaca tccgggcccc atcaagacc ccatgatgga tagccatgaa      600
ggtgtcaggg aaatgttttc acagcgtacg aaaaccctta tgggtcacat tggegaaccg      660
aatgacatcg catgggtctg tgtgtacctg gcctctgacg aatcgaaatt tgcgaagggt      720
gcagaatttg tggtcgacgg cgggttcacc gcacag                                756

```

<210> 160

<211> 252

<212> PRT

<213> 人工序列

[0122]

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 160

```

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

```

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

<210> 161
 <211> 756
 <212> DNA
 <213> 人工序列

[0123]

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 161
 atgaccgata gtctgaaggc caaagtagcc atcgtaaccg gcgggacaca aggtatcggt 60
 ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggtegtcgt 120
 gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgttttgte 180
 cagcacgatg tatccgatga agcaggtcgg actaaactgt tcgacaccac cgaggaggca 240
 ttggccccgg itacgacctt agtgaacaat gcagggatcc cggttggtta aagcgttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttcc 360
 ggcaccogtc tgggcattca gcgcattgaa aataaagget tgggcgctag catcatcaat 420
 atgagcagtg tgttcgggat tctaggcgac ccgaactacc gggcactatg tgcctccaag 480
 ggggcggtag gtatcatgtc gaaaagcgca gcctgtgatt gcgcactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgaigatgga tagccatgaa 600
 ggtgctgagg aaatgttttc acagcgtaacg aaaaccacct tgggtcacat tggcgaaccg 660
 aatgacatcg catgggtctg tctgtacctg gcattctgac aatcgaaatt tgcgacgggt 720
 gcagaatttg tggtcgacgg cgggtggacc gcacag 756

<210> 162
 <211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 162

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

[0124]

<210> 163

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 163

```

atgaccgate gtctgaaggc caaagtagcc atcgtaacgg gcgggacaca aggtatcggt      60
ttggcaatcg ccgataaact tctagaggag ggtgcgaaag tagttattac tggtcgtcgt      120
gcggatgtag gtgaaaagcg cgccaaatec atcgccggca ctgatgttat tcgctttgtc      180
cagcacgatg tatccgatga agcaggctgg cctaaactgt tcgacaccac cgaggaggca      240
ttcggcccggt ttacgaccgt cgtgaacaat gcagggatcc cgatggttaa aagcgttgaa      300
gacactacca cggaggaaat gcgtaaactg ctgtccgtta atctggatgg tgttttttcc      360
ggcacccgtc tgggcattca gcgcattgaa aataaagget tgggcgctag catcatcaat      420
atgagcagtg tgttcgggat tctaggcgac ccgactaccg gggcatacaa tgcctccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcctggatt gcgcactgaa ggactacgat      540
gtgcgtgtca acacagtaca tccgggcccc atcaagaacc cgatgatgga tagctatgaa      600
gggtctgagg aaatgttttc acagcgtacg aaaaccacct tgggtcacat tggegaaccg      660
aatgacatcg catgggtctg tctgtacctg gcattctgac aatcgaaatt tgcgacgggt      720
gcagaatttg tggtcgacgg cgggtacacc gcacag                                     756

```

[0125]

<210> 164

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 164

```

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

```

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

[0126]

<210> 165

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 165

atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacaca aggtatcggt	60
ttggcaatcg ccgataaatt ttagaggag ggtgcgaaag tagttattac tggctgctgt	120
gaggatgtag gtgaaaagge cgcacaaatca atcggcgga ctgatgttat tegtattgtc	180
cagcagatg tatccgatga agcaggctgg cctaaactgt tcgacaccac cgaggaggca	240
ttcgccccgg ttacgaccgt cgtgaacaat gcagggatc cggttgttaa aagcgttgaa	300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc	360
ggcacccgtc tgggcattca gcgatgaaa aataaaggct tgggcgctag catcatcaat	420
atgagcagtg tgttcgggat gtaggggat ccgactaccg gggcatactg tgcctccaag	480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat	540
gtgcgtgtca acacagtaca tccgggcccc atcaagacc ccatgatgga tgactatgaa	600
ggtgtctgagg aaatgttttc acagcgtacg aaaaccccta tgggtcacat tggcgaaccg	660
aatgacatcg catgggtctg tgtgtacctg gcacttgacg aatcgaaatt tgcgacgggt	720
gcagaatttg tggtcgacgg cgggtacacc gcacag	756

<210> 166
 <211> 252
 <212> PRT
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 166
 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Gln Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Val
 50 55 60
 Ser Asp Glu Ala Gly Trp Pro Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80
 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Pro Val Val
 85 90 95
 [0127] Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
 100 105 110
 Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg
 115 120 125
 Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Val
 130 135 140
 Phe Gly Met Val Gly Asp Pro Thr Thr Gly Ala Tyr Cys Ala Ser Lys
 145 150 155 160
 Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu
 165 170 175
 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys
 180 185 190
 Thr Pro Met Met Asp Asp Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Tyr Thr Ala Gln
 245 250

<210> 167

<211> 756
 <212> DNA
 <213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 167

```

atgacccgac gtctgaaggg caaagtagcc atcgtaaccg gggggacacg cggtatcggt    60
ttggcaatcg ccgataaatt tctagaggag ggtgcgaaag tagttattac tggtcgtcgt    120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc    180
cagcacgatg catccgatga agcaggetgg cctaaactgt tcgacaccac cgaggaggea    240
ttcggcccggt ttacgacett agtgaacaat gcagggattc cgggtggttaa aagcgttgaa    300
gacaactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttcc    360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat    420
atgagcagtg tgttcgggat tctagggcac ccgactaccg gggcatactg tgettcacag    480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggaactacgat    540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgaatgatga tgactatgaa    600
gggtctgagg aaatgttttc acagcgtacg aaaaccoccta tgggtcacat tggcgaaccg    660
aatgacatcg catgggtctg tgtgtacctg gcattcgacg aatcgaaatt tgcgacgggt    720
gcagaatttg tggtegacgg cgggttcacc gcacag                                756
  
```

[0128]

<210> 168
 <211> 252
 <212> PRT
 <213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 168

```

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

```

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

```

[0129]

<210> 169

<211> 756

<212> DNA

<213> 人工序列

<220>

<223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 169

```

atgaccgatac gtctgaaggg caaagtagcc atcgtaaccg gegggacaca aggtatcggc      60
ttggcaatcg cegataaatt thtagaggag ggtgcgaaag tagttattac tggtcgtcgt      120
geggaatgtag gtgaaaaggc cgccaaatea atcgcgcgca ctgatgttat tegetttgtc      180
cagcagcatg tatccgatga agcaggctgg cctaaactgt tcgacaccac cgaggaggca      240
ttcggecccg ttacgaccgt cgtgaacaat gcagggatcc eggttggtta aagcgttgaa      300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc      360
ggcacccgtc tgggcattca gcgcattgaa aataaagget tgggcgctag catecateaat      420
atgagcagtg tgttcgggat ttaggcgac ccgactaccg gggcatataa tgettccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat      540
gtgcgtgtca acacagtaca tcgggcccc atcaagacc ccatgatgga tagctatgaa      600
gggtctgagg aatgttttc acagcgtaac aaaacccta tgggtcact tggcgaccg      660
aatgacatcg catgggtctg tgtgtacct gcattctgac aatcgaaatt tgcgacgggt      720
gcagaatttg tggtcgacgg cgggtggacc gcacag      756

```

<210> 170
 <211> 252
 <212> PRT
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 170

Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Gln Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Val
 50 55 60
 Ser Asp Glu Ala Gly Trp Pro Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80
 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Pro Val Val
 [0130] 85 90 95
 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
 100 105 110
 Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg
 115 120 125
 Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Val
 130 135 140
 Phe Gly Ile Val Gly Asp Pro Thr Thr Gly Ala Tyr Asn Ala Ser Lys
 145 150 155 160
 Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu
 165 170 175
 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys
 180 185 190
 Thr Pro Met Met Asp Ser Tyr Glu Gly Ala Glu Glu Met Phe Ser Gln
 195 200 205
 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
 210 215 220
 Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240
 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
 245 250

<210> 171
 <211> 756
 <212> DNA
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 171
 atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacacg cggatateggt 60
 ttggcaatcg ccgataaatt tgtagaggag ggtagcgaag tagttattac tggtagtcgt 120
 gcggatgtag gtgaaaaggc cgcctaaatca atcgccggca ctgatgttat tegetttgtc 180
 cagcacgatg tatccgatga agcaggettg actaaactgt tegacaccac caggaggcca 240
 ttccgcccgg ttaagaccgt cgtgaacaat gcagggattc cgatgggtta aagcgttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
 ggcaccgcgc tgggcattca gcgcataaaa aataaaggct tgggcctag catcatcaat 420
 atgagcagtg tgttcgggat ggtaggcgac ccgactaccg gggcactactg tgcttccaag 480
 ggggcggta cgtatcatgt gaaaagcgca gcctggatt gcgcactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcccc atcaagacct cgatgatgga tgaccatgaa 600
 ggtgctgagg aatgttttc acagcgtacg aaaaccctta tgggtcacat tggcgaaccg 660
 aatgacatcg catgggtctg tgtgtacctg gcctctgacg aatcgaaatt tgcgacgggt 720
 gcagaatttg tggtagcgg cgggttcacc gcacag 756

[0131]

<210> 172
 <211> 252
 <212> PRT
 <213> 人工序列

<220>
 <223> 克菲尔乳杆菌酮还原酶的工程化变体

<400> 172
 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 Arg Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30
 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
 35 40 45
 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Val
 50 55 60
 Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
 65 70 75 80

[0132]

Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Pro Met Val		
85	90	95
Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser		
100	105	110
Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg		
115	120	125
Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Val		
130	135	140
Phe Gly Met Val Gly Asp Pro Thr Thr Gly Ala Tyr Cys Ala Ser Lys		
145	150	155
Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu		
165	170	175
Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Pro Ile Lys		
180	185	190
Thr Pro Met Met Asp Asp His Glu Gly Ala Glu Glu Met Phe Ser Gln		
195	200	205
Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala		
210	215	220
Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly		
225	230	235
Ala Glu Phe Val Val Asp Gly Gly Phe Thr Ala Gln		
245	250	