(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 101889081 B
(45)授权公告日 2014.06.18

(21)申请号 200880118039.6
(22)申请日 2008.09.28
(30)优先权数据
60/976,345 2007.09.28 US
(85)PCT国际申请进入国家阶段日
2010.05.27
(86)PCT国际申请的申请数据
PCT/US2008/078046 2008.09.28
(87)PCT国际申请的公布数据
WO2009/042984 EN 2009.04.02
(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 杰克·梁 波蒂·博勒普
韦丝娜·米切尔 艾米丽·穆德弗
詹姆士·拉隆德 吉伽特·哈思曼
(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 7/42(2006.01)
C12P 7/66(2006.01)
C12P 41/00(2006.01)

(56)对比文件
US 5225339 A,1993.07.06, 全文.
EP 0796914 A2,1997.09.24, 全文.
CN 1300833 A,2001.06.27, 全文.
WO 2008103248 A1,2008.08.28, 全文.
Philippe Bisel 等.Stereochemical clarification of the enzyme-catalysed reduction of 2-acetylchromen-4-one.《Tetrahedron: Asymmetry》.2007, 第18卷(第9期),1142-1144.

审查员 韦炜

权利要求书5页 说明书85页
序列表139页 附图1页

(54)发明名称

酮还原酶多肽及其用途

(57)摘要

本发明提供与天然存在的野生型酮还原酶相比具有改进性质的人工改造的酮还原酶。还提供了编码人工改造的酮还原酶的多核苷酸、能够表达人工改造的酮还原酶的宿主细胞和使用人工改造的酮还原酶来合成各种手性化合物的方法。

CN 101889081 B

1. 一种能够以SEQ ID NO:4多肽速率的至少1.5倍的速率将底物2-[3-[3-[2-(7-氯-2-喹啉基)乙烯基]苯基]-3-氧代丙基]苯甲酸甲酯转化为产物(S,E)-2-(3-(3-(2-(7-氯喹啉-2-基)乙烯基)苯基)-3-羟丙基)苯甲酸甲酯的酮还原酶多肽,其包括与具有下述特征的基于SEQ ID NO:2、4或114的参考序列至少85%相同的氨基酸序列:对应于X190的残基是脯氨酸,对应于X195的残基是精氨酸,而且对应于X202的残基是色氨酸,并且其中所述酮还原酶多肽的对应于X190的残基是脯氨酸、对应于X195的残基是精氨酸、而且对应于X202的残基是色氨酸,其中所述酮还原酶多肽能够将所述底物转化为具有至少99%的百分比立体异构过量的所述产物,其中所述酮还原酶的氨基酸序列选自SEQ ID NO:6、8、10、12、14、16、18、20、22、26、28、30、32、34、36、50、64、76、78、80、82、84、108和110。

2. 一种编码根据权利要求1的多肽的多核苷酸。

3. 根据权利要求2所述的多核苷酸,其选自对应于SEQ ID NO:5、7、9、11、13、15、17、19、21、23、25、27、29、31、33、35、49、63、75、77、79、81、83、107或109的序列。

4. 一种包括权利要求2所述多核苷酸的表达载体,所述多核苷酸可操作地连接至适合指导在宿主细胞中表达的控制序列。

5. 根据权利要求4所述的表达载体,其中所述控制序列包括启动子。

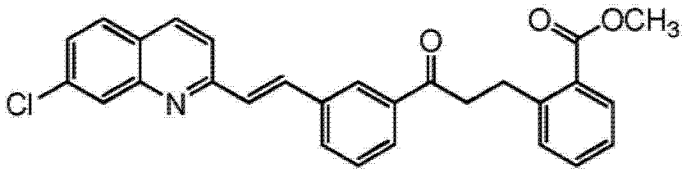
6. 根据权利要求5所述的表达载体,其中所述启动子是大肠杆菌启动子。

7. 一种包括权利要求4所述表达载体的宿主细胞。

8. 根据权利要求7所述的宿主细胞,其为大肠杆菌。

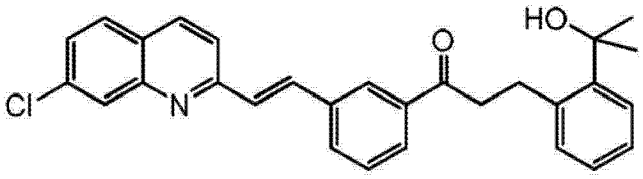
9. 根据权利要求7所述的宿主细胞,其中编码所述多肽的所述表达载体的密码子已被优化以在所述宿主细胞中表达。

10. 一种组合物,其包括权利要求1的酮还原酶和结构式(I)的底物2-[3-[3-[2-(7-氯-2-喹啉基)乙烯基]苯基]-3-氧代丙基]苯甲酸甲酯



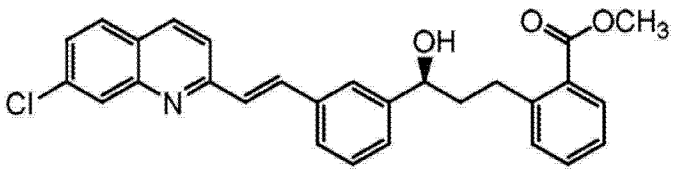
(I);或

结构式(III)的底物2-[2-[3-[3[2-(7-氯-2-喹啉基)乙烯基]苯基]-3-氧代丙基]苯基]-2-丙醇



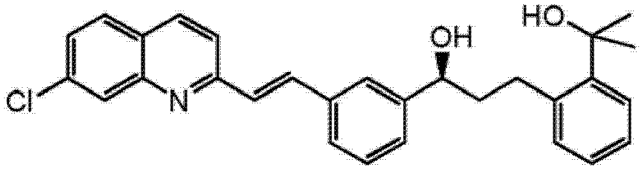
(III)。

11. 根据权利要求10所述的组合物,其中所述底物是结构式(I)的底物而且所述组合物还包括结构式(II)的产物(S,E)-2-(3-(3-(2-(7-氯喹啉-2-基)乙烯基)苯基)-3-羟丙基)苯甲酸甲酯



(II)。

12. 根据权利要求 10 所述的组合物，其中所述底物是结构式（III）的底物而且所述组合物还包括结构式（IV）的产物 (S,E)-1-(3-(2-(7- 氯喹啉 -2- 基）乙烯基）苯基)-3-(2-(2- 羟丙 -2- 基）苯基）丙 -1- 醇

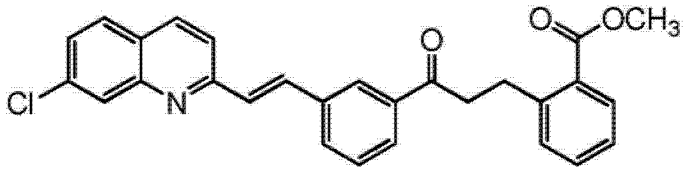


(IV)。

13. 根据权利要求 10 所述的组合物，还包括辅助因子再生系统。

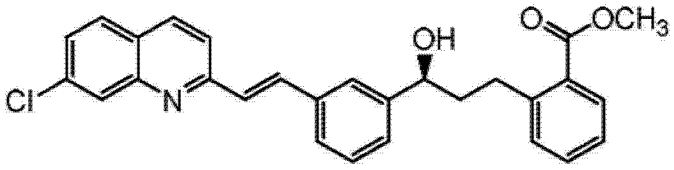
14. 根据权利要求 13 所述的组合物，其中所述辅助因子再生系统包括葡萄糖脱氢酶和葡萄糖；甲酸脱氢酶和甲酸；或者异丙醇和仲醇脱氢酶。

15. 一种将式（I）的底物 2-[3-[3-[2-(7- 氯 -2- 喹啉基）乙烯基］苯基］-3- 氧代丙基］苯甲酸甲酯



(I)

还原为式（II）的产物 (S,E)-2-(3-(3-(2-(7- 氯喹啉 -2- 基）乙烯基）苯基)-3- 羟丙基）苯甲酸甲酯



(II)

的方法，该方法包括将式（I）底物在适于将所述底物还原或转化为式（II）产物的反应条件下与权利要求 1 所述的酮还原酶多肽接触。

16. 根据权利要求 15 所述的方法，其中所述产物以超过 99% 的立体异构过量存在。

17. 根据权利要求 15 所述的方法，该方法使用表达酮还原酶的完整细胞或这种细胞的提取物或裂解产物进行。

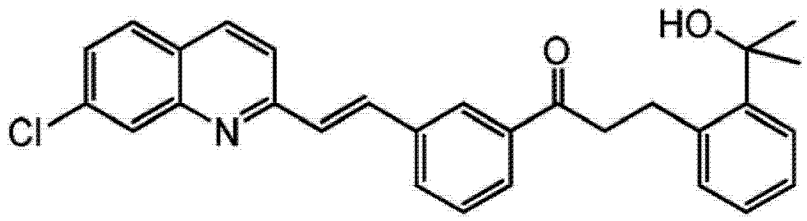
18. 根据权利要求 15 所述的方法，其中所述酮还原酶是分离和 / 或纯化的而且还原反应是在酮还原酶的辅助因子的存在下进行的。

19. 根据权利要求 15 所述的方法，其中所述酮还原酶是分离和 / 或纯化的而且所述还原反应是在酮还原酶的辅助因子以及辅助因子再生系统的存在下进行的。

20. 根据权利要求 18 所述的方法，其中所述辅助因子再生系统包括葡萄糖脱氢酶和葡萄糖；甲酸脱氢酶和甲酸；或者异丙醇和仲醇脱氢酶。

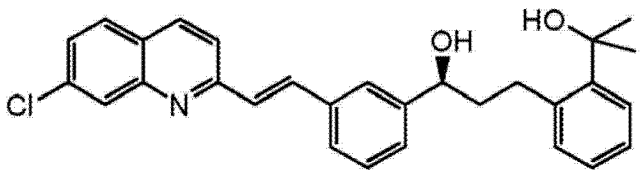
21. 根据权利要求 20 所述的方法，其中所述仲醇脱氢酶是如权利要求 1 所述的酮还原酶。

22. 一种将式(III)底物 2-[2-[3-[3[2-(7- 氯 -2- 喹啉基) 乙烯基] 苯基]-3- 氧代丙基] 苯基]-2- 丙醇



(III)

还原为式(IV)产物 (S,E)-1-(3-(2-(7- 氯喹啉 -2- 基) 乙烯基) 苯基)-3-(2-(2- 羟丙 -2- 基) 苯基) 丙 -1- 醇



(IV)

的方法,该方法包括将式(III)底物在适于将所述底物还原为式(IV)产物的反应条件下与权利要求 1 所述的酮还原酶多肽接触。

23. 根据权利要求 22 所述的方法,其中所述产物以超过 99% 的立体异构过量存在。

24. 根据权利要求 22 所述的方法,其使用表达酮还原酶的完整细胞或这种细胞的提取物或裂解产物进行。

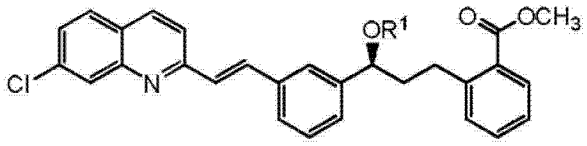
25. 根据权利要求 22 所述的方法,其中所述酮还原酶是分离和 / 或纯化的而且还原反应是在酮还原酶的辅助因子的存在下进行的。

26. 根据权利要求 22 所述的方法,其中所述酮还原酶是分离和 / 或纯化的而且所述还原反应是在酮还原酶的辅助因子以及辅助因子再生系统的存在下进行的。

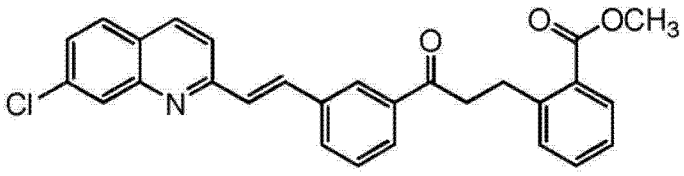
27. 根据权利要求 25 所述的方法,其中所述辅助因子再生系统包括葡萄糖脱氢酶和葡萄糖 ;甲酸脱氢酶和甲酸 ;或者异丙醇和仲醇脱氢酶。

28. 根据权利要求 27 所述的方法,其中所述仲醇脱氢酶是如权利要求 1 所述的酮还原酶。

29. 一种用于合成式(XX)化合物的方法 :

(XX) 

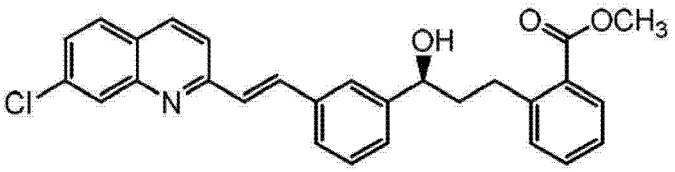
其中 R^1 是烷基磺酰基或芳基磺酰基,所述方法包括 :(a) 通过在适于将底物还原或转化为式(II)产物的反应条件下将式(I)底物与权利要求 1 所述的酮还原酶多肽接触而将式(I)底物 2-[3-[3-[2-(7- 氯 -2- 喹啉基) 乙烯基] 苯基]-3- 氧代丙基] 苯甲酸甲酯



(I)

还原为式(II)产物 (S,E)-2-(3-(3-(2-(7- 氯喹啉 -2- 基) 乙烯基) 苯基)-3- 羟丙

基）苯甲酸甲酯

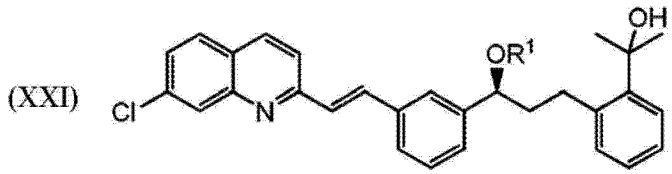


（II）；

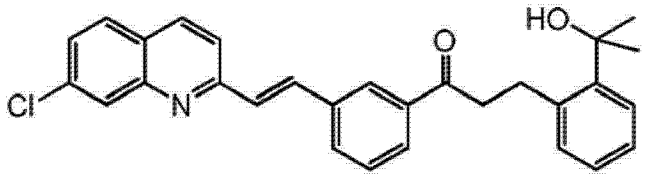
以及（b）将式（II）化合物与烷基磺酰基卤素或芳基磺酰基卤素在适于形成化合物（XX）的条件下反应。

30. 根据权利要求 29 所述的方法，其中所述烷基磺酰基卤素是甲基磺酰氯。

31. 一种用于合成式（XXI）化合物的方法：

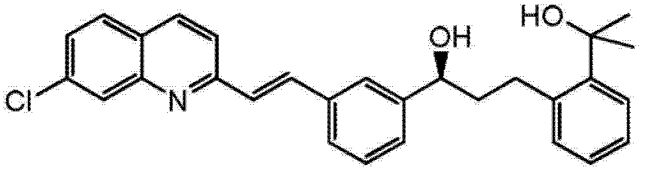


其中 R^1 是烷基磺酰基或芳基磺酰基，所述方法包括：（a）通过在适于将底物还原或转化为式（IV）产物的反应条件下将式（III）底物



（III）

与权利要求 1 所述的酮还原酶多肽接触而将式（III）底物 2-[2-[3-[3[2-(7-氯-2-喹啉基）乙烯基］苯基]-3-氧代丙基］苯基]-2-丙醇还原为式（IV）产物 (S,E)-1-(3-(2-(7-氯喹啉-2-基）乙烯基）苯基)-3-(2-(2-羟丙-2-基）苯基）丙-1-醇

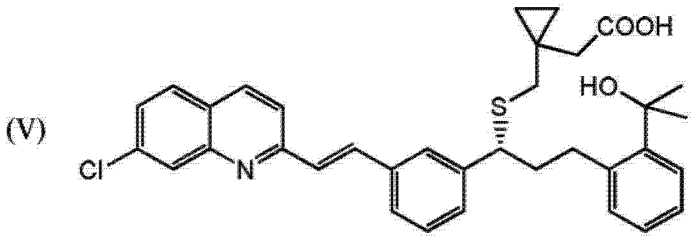


（IV）；

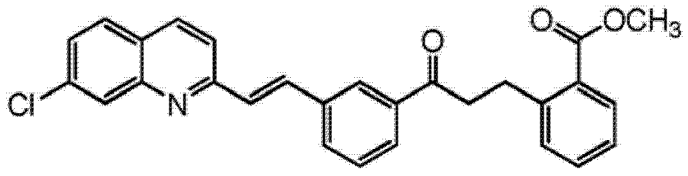
以及（b）将式（IV）化合物与烷基磺酰基卤素或芳基磺酰基卤素在适于形成化合物（XXI）的条件下反应。

32. 根据权利要求 29 所述的方法，其中所述烷基磺酰基卤素是甲基磺酰氯。

33. 一种用于合成化合物（V）或其盐、水合物或溶剂化物的方法：

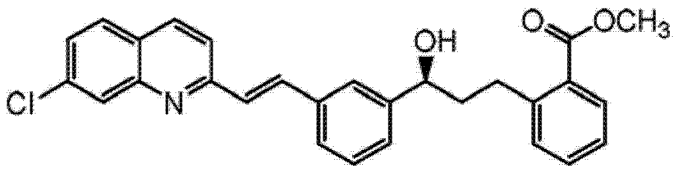


其中所述方法中的步骤包括通过在适于将底物还原或转化为式（II）产物的反应条件下将式（I）底物与权利要求 1 所述的酮还原酶多肽接触而将式（I）底物 2-[3-[3-[2-(7-氯-2-喹啉基）乙烯基］苯基]-3-氧代丙基］苯甲酸甲酯



(I)

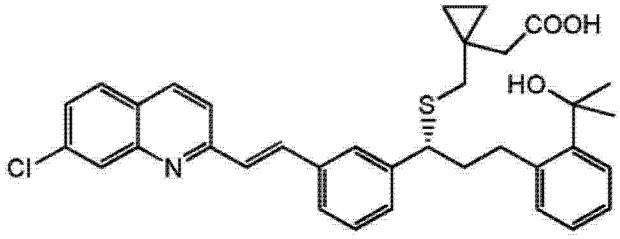
还原为式(II)产物 (S,E)-2-(3-(3-(2-(7- 氯喹啉 -2- 基) 乙烯基) 苯基)-3- 羟丙基) 苯甲酸甲酯



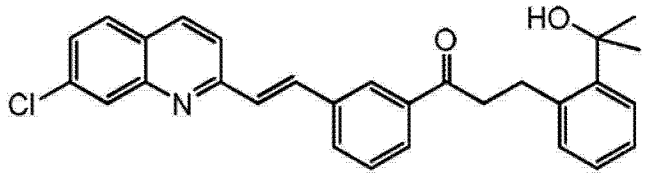
(II)。

34. 一种用于合成化合物(V)或其盐、水合物或溶剂化物的方法 :

(V)

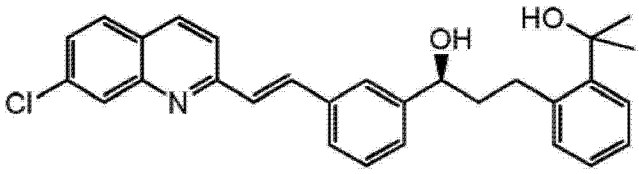


其中所述方法中的步骤包括通过在适于将底物还原为式(IV)产物的反应条件下将式(III)底物与权利要求 1 所述的酮还原酶多肽接触而将式(III)底物 2-[2-[3-[3[2-(7- 氯 -2- 喹啉基) 乙烯基] 苯基]-3- 氧代丙基] 苯基]-2- 丙醇



(III)

还原为式(IV)产物 (S,E)-1-(3-(2-(7- 氯喹啉 -2- 基) 乙烯基) 苯基)-3-(2-(2- 羟丙 -2- 基) 苯基) 丙 -1- 醇



(IV)。

酮还原酶多肽及其用途

[0001] 1. 相关申请的交叉引用

[0002] 本申请根据美国法典第 35 编专利法第 119(e) 条要求在 2007 年 9 月 28 日提交的序号为 60/976,345 的申请的权益，所述申请的内容通过引用被并入本文。

[0003] 2. 对序列表、表格或计算机程序的引用

[0004] 根据 37C.F.R.1.821 以计算机可读形式 (CRF) 通过 EFS-Web 使用文件名 376247-018.txt 随本申请共同提交的序列表通过引用被并入本文。电子版的序列表创建于 2008 年 9 月 28 日，其文件大小为 254 千字节。

[0005] 发明领域

[0006] 本发明涉及酮还原酶多肽及其用途。

[0007] 3. 背景

[0008] 属于酮还原酶 (KRED) 或羰基还原酶类别 (EC1.1.1.184) 的酶用于从相应的前立体异构 (prostereoisomeric) 酮底物或相应的外消旋 (racemic) 醛底物合成旋光的醇。KRED 通常将酮或醛转化为相应的醇产物，但是还可以催化逆反应，即醇底物被氧化为相应的酮 / 醛产物。酶例如 KRED 对酮和醛的还原以及醇的氧化需要辅助因子，其中最常见的为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) 以及用于氧化反应的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 充当电子供体，而 NAD 和 NADP 充当电子受体。经常观察到酮还原酶和醇脱氢酶接受磷酸化的或非磷酸化的辅助因子 (以其氧化态和还原态)。

[0009] 可在大范围的细菌和酵母中发现 KRED 酶 (关于综述，参见 Kraus 和 Waldman，1995，Enzyme catalysis in organic synthesis(有机合成中的酶催化)，卷 1 和 2.VCH Weinheim ;Faber，K.，Biotransformations in organic chemistry(有机化学中的生物转化)，第 4 版 Springer，Berlin Heidelberg New York.2000 ; 以及 Hummel 和 Kula，1989，Eur.J.Biochem.184 :1-13)。已经报道了若干 KRED 基因和酶序列，例如木兰假丝酵母 (Candida magnoliae)(Genbank 登记号 JC7338 ;GI :11360538)、近平滑假丝酵母 (Candidaparapsilosis)(Genbank 登记号 BAA24528.1 ;GI :2815409)、赭色掷孢酵母 (Sporobolomyces salmonicolor)(Genbank 登记号 AF160799 ;GI :6539734)。

[0010] 为了省去产生关键化合物的许多化学合成步骤，渐增地利用酮还原酶将不同的酮和醛底物向手性醇产物进行酶促转化。这些应用可利用表达酮还原酶的完整细胞来进行生物催化的酮还原，或在多种酮还原酶在完整细胞中的存在将不利地影响立体纯度和期望产物收率的情况下使用纯化的酶。对于体外应用，将再生酶例如葡萄糖脱氢酶 (GDH)、甲酸脱氢酶等的辅助因子 (NADH 或 NADPH) 与酮还原酶共同使用。使用酮还原酶来产生有用的化合物的实例包括 4- 氯乙酰基乙酸酯的不对称还原 (Zhou，1983，J.Am.Chem.Soc.105 : 5925-5926 ;Santaniello，J.Chem.Res.(S)1984 :132-133 ; 美国专利第 5,559,030 号 ; 美国专利第 5,700,670 号和美国专利第 5,891,685 号)、二氧代羧酸的还原 (例如美国专利第 6,399,339 号)、(S) 氯 -5- 羟基 -3- 氧代己酸叔丁基酯的还原 (例如美国专利第 6,645,746 号和 WO 01/40450)、基于吡咯并三嗪的化合物的还原 (例如美国申请第 2006/0286646

号）、取代苯乙酮的还原（例如美国专利第6,800,477号）和酮噻吩烷(ketothiolanes)的还原(WO 2005/054491)。

[0011] 期望鉴定可被用来进行从各种酮和醛底物向其对应的手性醇产物转化的其他的酮还原酶。

[0012] 4. 概述

[0013] 在一个方面，本公开提供能够将2-[3-[3-[2-(7-氯-2-喹啉基）乙烯基]苯基]-3-氧代丙基]苯甲酸甲酯（“底物”）还原或转化为(S,E)-2-(3-(3-(2-(7-氯喹啉-2-基）乙烯基）苯基)-3-羟丙基）苯甲酸甲酯（“产物”）的酮还原酶多肽，编码这种多肽的多核苷酸以及使用所述多肽的方法。相比之下，天然存在 的酮还原酶不展示将特定的酮底物还原或转化为相应的醇产物的显著活性。

[0014] 本公开中人工改造的酮还原酶多肽衍生自乳酸杆菌属(Lactobacillus)物种的酮还原酶。照这样，在一些实施方案中，与得自乳酸杆菌例如高加索酸奶乳杆菌(Lactobacillus kifir)(“L.kefir”;SEQ ID NO:4)、短乳杆菌(Lactobacillus brevis)(“L.brevis”;SEQ ID NO:2)和微小乳杆菌(Lactobacillus minor)(“L.minor”,SEQ ID NO:114)的天然存在的野生型酮还原酶相比，本公开中人工改造的酮还原酶多肽在将特定的酮底物转化为相应的醇产物中具有增加的活性。在一些实施方案中，与SEQ ID NO:2、4或114的序列相比，人工改造的酮还原酶至少具有下面的特征：(1)对应于位置190(即X190)的残基是受限(constrained)残基，而且(2)对应于位置202(即X202)的残基是芳香族残基。在一些实施方案中，与SEQ ID NO:2、4或114的序列相比，本公开的酮还原酶至少具有下面的特征：(1)对应于位置190(即X190)的残基是受限残基，(2)对应于位置195(即X195)的残基是碱性的、酸性的、非极性的或极性的残基，而且(3)对应于位置202(即X202)的残基是芳香族残基。

[0015] 除了上述特征之外，与SEQ ID NO:2、4或114的序列相比，酮还原酶在其他残基位置可具有一个或多个残基差异。在一些实施方案中，本文的酮还原酶多肽包括与具有下述特征的基于SEQ ID NO:2、4或114的参考序列至少约80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多地相同的氨基酸序列：对应于X190的残基是受限残基，特别是脯氨酸；而且对应于X202的残基是芳香族残基，特别是色氨酸；条件是酮还原酶多肽至少具有上述特征，即，对应于X190的残基是受限残基，而且对应于X202的残基是芳香族残基。在一些实施方案中，酮还原酶多肽至少具有下面的特征：对应于X190的残基是脯氨酸；而且对应于X202的残基是色氨酸。

[0016] 在一些实施方案中，本文的酮还原酶多肽包括与具有下述特征的基于SEQ ID NO:2、4或114的参考序列至少约80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更 多地相同的氨基酸序列：对应于X190的残基是受限残基，特别是脯氨酸；对应于X195的残基是碱性的、酸性的、非极性的或极性的残基，特别是脯氨酸、精氨酸、赖氨酸、天冬氨酸或天冬酰胺；而且对应于X202的残基是芳香族残基，特别是色氨酸；条件是酮还原酶多肽至少具有上述特征，即，对应于X190的残基是受限残基，对应于X195的残基是碱性的、酸性的、非极性的或极性的残基，而且对应于X202的残基是芳香族残基。在一些实施方案中，酮还原酶多肽至少具有下面的特征：对应于X190的残基是脯氨酸；对应于X195的残基是脯氨酸、精氨酸、赖氨酸、天冬氨酸或天冬酰胺；而且

对应于 X202 的残基是色氨酸。

[0017] 在一些实施方案中，例如当改进的性质是来自单个残基差异或残基差异的特定组合时，人工改造的酮还原酶可任选地包括与参考序列相比在多肽中其他位置一个或多个保守的残基差异。在一些实施方案中，可以引入其他残基位置的另外的残基差异以对酶的性质产生进一步的改进。这些改进可以是针对确定底物的酶活性的进一步增加，但也可包括立体选择性、热稳定性、溶剂稳定性的增加和 / 或产物抑制的减少。能够导致一种或多种改善的酶性质的各种残基差异在详述中被提供。在一些实施方案中，这些人工改造的酮还原酶多肽基于 SEQ ID NO :111、112 和 139 或其区域（例如对应于残基 90 至 211 的残基）中所陈列的序列式。SEQ ID NO :112 基于高加索酸奶乳杆菌酮还原酶的野生型氨基酸序列 (SEQ ID NO :4) ;SEQ IDNO :111 基于短乳杆菌酮还原酶的野生型氨基酸序列 (SEQ ID NO :2) ;而且 SEQ ID NO :139 基于微小乳杆菌酮还原酶的野生型氨基酸序列 (SEQ IDNO :114)。序列式说明，对应于 X190 的残基是受限残基 ;对应于 X195 的残基是碱性的、酸性的、非极性的或极性的残基 ;而且对应于 X202 的残基是芳香族残基。如详述中所提供，序列式还说明在其他残基位置的特征。

[0018] 在一些实施方案中，与 SEQ ID NO :2、4 或 114 相比，在酶活性速率方面改进了本公开的酮还原酶多肽，所述酶活性速率即它们将底物转化为产物的速率。改进量范围可是 1.5 倍（或重）于对应的野生型或参考酮还原酶的酶活性至多达 2 倍、5 倍、10 倍、20 倍、25 倍、50 倍、75 倍、100 倍或更多倍。在一些实施方案中，酮还原酶多肽能够以至少 5 倍、10 倍、 25 倍、50 倍、75 倍、100 倍、150 倍、200 倍、250 倍、300 倍、500 倍或 1000 倍于 SEQ ID NO :4 或 SEQ ID NO :90 的多肽所展现的速率的速率将底物转化为产物。

[0019] 在一些实施方案中，人工改造的酮还原酶多肽能够以比参考多肽改进的速率将底物 2-[3-[3-[2-(7- 氯 -2- 喹啉基）乙烯基] 苯基]-3- 氧代丙基] 苯甲酸甲酯转化为产物 (S,E)-2-(3-(3-(2-(7- 氯喹啉 -2- 基）乙烯基）苯基)-3- 羟丙基）苯甲酸甲酯，其中所述参考多肽具有 SEQ ID NO :90 的氨基酸序列并能够将底物转化为具有至少约 95％的百分比立体异构过量的产物。具有这种性质的示例性多肽包括但不限于含有对应于 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、92、94、96、98、100、102、104、106、108 和 110 的氨基酸序列的多肽。

[0020] 在一些实施方案中，酮还原酶多肽还能够将底物转化为具有至少约 99％的百分比立体异构过量的产物。具有这种性质的示例性多肽包括但不限于含有对应于 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、92、94、96、98、100、102、104、106、108 和 110 的氨基酸序列的多肽。因为具有 SEQ ID NO :90 的氨基酸序列的参考多肽能够以比野生型 (SEQ IDNO :4) 改进的速率（例如，用约 10g/L 的 KRED 将 1g/L 的底物的 10-20％在 24 小时内转化为产物）和立体选择性 (99％的立体异构过量）将底物转化为产物，所以比 SEQ ID NO :90 改进的本文的多肽也是比野生型改进的。

[0021] 在一些实施方案中，酮还原酶多肽能够以至少约 99％的百分比立体异构过量和比参考多肽改进了至少 1-50 倍的速率将底物 2-[3-[3-[2-(7- 氯 -2- 喹啉基）乙烯基] 苯基]-3- 氧代丙基] 苯甲酸甲酯转化为产物 (S, E)-2-(3-(3-(2-(7- 氯喹啉 -2- 基）乙烯

基）苯基)-3-羟丙基）苯甲酸甲酯，其中所述参考多肽具有 SEQ ID NO:90 的氨基酸序列。具有这种性质的示例性多肽包括但不限于含有对应于SEQ ID NO:42、44、46、48、86、88、92、94、96、100 或 102 的氨基酸序列的多肽。

[0022] 在一些实施方案中，酮还原酶多肽能够以至少约 99％的百分比立体异构过量和比参考多肽改进了至少 50-250 倍的速率将底物 2-[3-[3-[2-(7- 氯 -2- 喹啉基）乙烯基] 苯基]-3- 氧代丙基] 苯甲酸甲酯转化为产物 (S, E)-2-(3-(3-(2-(7- 氯喹啉 -2- 基）乙烯基）苯基)-3- 羟丙基）苯甲酸甲酯，其中所述参考多肽具有 SEQ ID NO:90 的氨基酸序列。具有这种性质的示例性多肽包括但不限于含有对应于SEQ ID NO:10、32、34、36、50、64、66、68、70、72、74、76、78、80、82、84、98、104、106 和 108 的氨基酸序列的多肽。

[0023] 在一些实施方案中，酮还原酶多肽能够以至少约 99％的百分比立体异构过量和比参考多肽改进了至少 250-1000 倍的速率将底物 2-[3-[3-[2-(7- 氯 -2- 喹啉基）乙烯基] 苯基]-3- 氧代丙基] 苯甲酸甲酯转化为产物 (S,E)-2-(3-(3-(2-(7- 氯喹啉 -2- 基）乙烯基）苯基)-3- 羟丙基）苯甲酸甲酯，其中所述参考多肽具有 SEQ ID NO:90 的氨基酸序列。具有这种性质的示例性多肽包括但不限于含有对应于SEQ ID NO:12、18、22、24、26、28、30、38、40、52、54、56、58、60、62 和 110 的氨基酸序列的多肽。

[0024] 在一些实施方案中，本发明提供能够以至少约 99％的百分比立体异构过量和比参考多肽改进了至少 1000 倍的速率将底物 2-[3-[3-[2-(7- 氯 -2- 喹啉基）乙烯基] 苯基]-3- 氧代丙基] 苯甲酸甲酯转化为产物 (S, E)-2-(3-(3-(2-(7- 氯喹啉 -2- 基）乙烯基）苯基)-3- 羟丙基）苯甲酸甲酯的酮还原酶多肽，其中所述参考多肽具有 SEQ ID NO:90 的氨基酸序列。具有这种性质的示例性多肽包括但不限于含有对应于 SEQ ID NO:6、8、14、16 和 20 的氨基酸序列的多肽。

[0025] 在一些实施方案中，与具有 SEQ ID NO:90 的氨基酸序列的参考多肽相比，本公开的改进的酮还原酶多肽还具有增加的对丙酮的抵抗性，其中丙酮作为底物向产物转化中的竞争性抑制剂。具有这种增加的丙酮抵抗性的示例性多肽包括但不限于含有对应于SEQ ID NO:6、8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、38、40、50、52、54、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、 98、104、106、108 和 110 的氨基酸序列的多肽。

[0026] 在一些实施方案中，当用超过约 100g/L 的底物和小于约 5g/L 的多肽进行时，酮还原酶多肽能够在小于约 24 小时中将至少约 95％的底物转化为产物。具有这种能力的示例性多肽包括但不限于含有对应于 SEQ IDNO:6、8、14、16 和 20 的氨基酸序列的多肽。

[0027] 在一些实施方案中，本公开的酮还原酶是高度地立体选择性的，并且因此能够将底物还原为具有超过约 99％、99.1％、99.2％、99.3％、99.4％、99.5％、99.6％、99.7％、99.8％或 99.9％立体异构过量的产物。具有这种高度立体选择性的示例性酮还原酶多肽包括但不限于含有对应于SEQ IDNO: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 和 110 的氨基酸序列的多肽。

[0028] 在另一方面，本公开提供编码本文所述人工改造的酮还原酶的多核苷酸或者与这种多核苷酸在高度严格的条件下杂交的多核苷酸。所述多核苷酸可包括启动子和用于表达所编码的人工改造的酮还原酶的其他调节元件，并且可利用为特定的所期望的表达系统而

优化的密码子。示例性多核苷酸包括但不限于 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 或 109。示例性多核苷酸还包括编码对应于 SEQ ID NO:111、112 和 139 的序列式的多肽的多核苷酸。

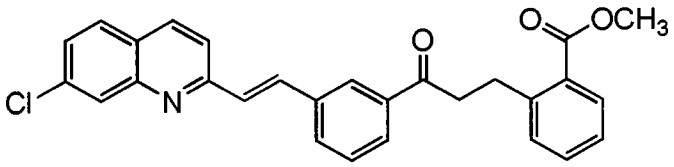
[0029] 在另一方面，本公开提供含有本文所述多核苷酸和 / 或表达载体的宿主细胞。所述宿主细胞可以是高加索酸奶乳杆菌或短乳杆菌或微小乳杆菌，或者它们可以是不同的生物体。所述宿主细胞可被用来表达和分离本文所述人工改造的酮还原酶，或者可选地，它们可被直接用来将底物转化为立体异构产物。

[0030] 无论用完整的细胞、细胞提取物还是纯化的酮还原酶进行该方法，都 可以使用单独的酮还原酶，或者可选地，可以使用两种或更多种酮还原酶的混合物。

[0031] 本文所述的酮还原酶能够催化具有结构式 (I) 的化合物即 2-[3-[3-[2-(7- 氯 -2- 喹啉基）乙烯基］苯基］-3- 氧代丙基］苯甲酸甲酯（“底物”）中的酮基：

[0032]

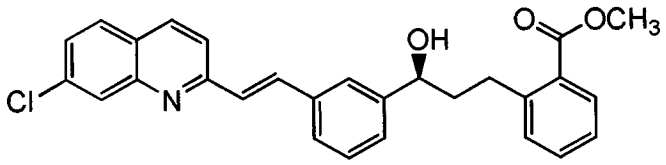
(I)



[0033] 向具有结构式 (II) 的相应的立体异构醇产物即 (S, E)-2-(3-(3-(2-(7- 氯喹啉 -2- 基）乙烯基）苯基）-3- 羟丙基）苯甲酸甲酯（“产物”）还原的反应：

[0034]

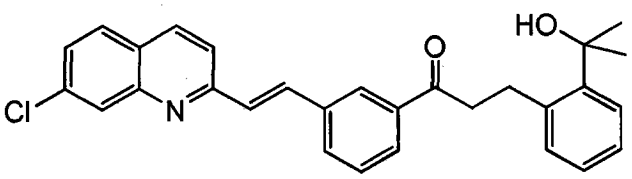
(II)



[0035] 在一些实施方案中，本文所述的酮还原酶能够催化具有结构式 (III) 的化合物即 2-[2-[3-[3[2-(7- 氯 -2- 喹啉基）乙烯基］苯基］-3- 氧代丙基］苯基］-2- 丙醇（也是“底物”）中的酮基，

[0036]

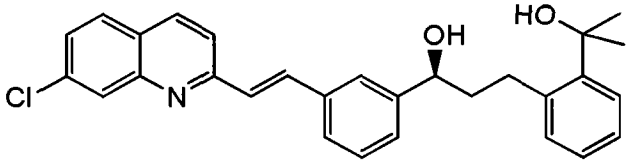
(III)



[0037] 向具有结构式 (IV, (S, E)-1-(3-(2-(7- 氯喹啉 -2- 基）乙烯基）苯基）-3-(2-(2- 羟丙 -2- 基）苯基）丙 -1- 醇（也是有关式 (III) 中底物的产物））的相应的立体异 构醇产物还原的反应：

[0038]

(IV)

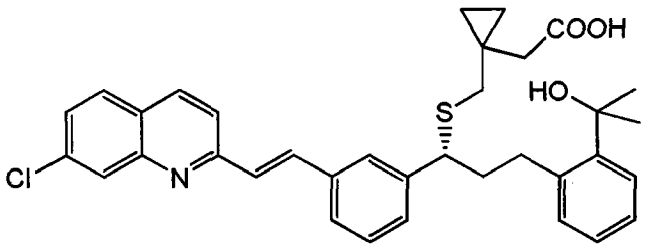


[0039] 因此，在一些实施方案中，本文提供了将结构式 (I) 或结构式 (III) 中底物还原为相应的醇产物的方法，其包括将所述底物在适于将所述底物分别还原或转化为结构式 (II) 或结构式 (IV) 中产物的反应条件下与本公开的酮还原酶多肽接触或孵化。

[0040] 在一些实施方案中，本文所提供的酮还原酶多肽可被用来生产用于合成结构式 (V) 中化合物 2-[1-[[(1R)-1-[3-[(E)-2-(7-氯喹啉-2-基）乙烯基］苯基］-3-[2-(2-羟丙-2-基）苯基］丙基］硫甲基］环丙基］乙酸的中间产物：

[0041]

(V)



[0042] 包括其溶剂化物、水合物及药学上可接受的盐。在一些实施方案中，在用于合成式 (V) 中化合物的方法中，所述方法中的步骤可包括在适于将所述底物还原或转化为式 (II) 中产物的反应条件下使用本公开的酮还原酶多肽将式 (I) 中的底物还原或转化为式 (II) 中的产物。

[0043] 在一些实施方案中，式 (IV) 的中间产物还可被用来合成结构式 (V) 中的化合物。因此，在一些实施方案中，在用于合成式 (V) 中化合物的方法中，所述方法中的步骤可包括通过在适于将所述底物还原或转化为式 (IV) 中产物的反应条件下将底物与本文公开的酮还原酶接触而将式 (III) 中的底物还原或转化为式 (IV) 中的产物。

[0044] 在一些将底物还原为产物的方法的实施方案中，底物被还原为具有超过约 99％立体异构过量的产物，其中所述酮还原酶多肽包括对应于 SEQ IDNO：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、111 或 112 的序列。

[0045] 在一些将底物还原为产物的方法的实施方案中，当用超过约 100g/L 的底物和小于约 5g/L 的多肽进行时，至少约 95％的底物在小于约 24 小时中被转化为产物，其中所述多肽包括对应于 SEQ ID NO：6、8、14、16 或 20 的氨基酸序列。

[0046] 在一些将底物还原为产物的方法的实施方案中，使用约 10g/L 的多肽在小于约 24 小时中将至少约 10-20％的 1g/L 的底物转化为产物，其中所述多肽包括对应于 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、111 或 112 的氨基酸序列。

[0047] 5. 附图简述

[0048] 图 1 说明酮还原酶 (KRED) 在将式 (I) 中底物化合物向式 (II) 中相应产物转化中的作用。该还原使用本文所述的 KRED 和辅助因子例如 NADPH。葡萄糖脱氢酶 (GDH) 被用来将 $NAD(P)^+$ 转化 / 再循环为 NAD(P)H。葡萄糖被转化为葡萄糖酸，而随着氢氧化钠的加入，葡萄糖酸又被转化为它的钠盐（葡萄糖酸钠）。可选地，可使用仲醇脱氢酶（例如 KRED) 来

将异丙醇转化为丙酮并将 $NAD(P)^+$ 转化 / 再循环为 NAD(P)H。

[0049] 6. 详述

[0050] 6.1. 定义

[0051] 如本文所用，下列术语旨在具有下列意义。

[0052] “酮还原酶”和“KRED”在本文中被可互换地使用以指具有将羰基还原为它相应的醇的酶能力的多肽。更具体地说，本发明的酮还原酶多肽能够立体选择性地将上述式 (I) 中的化合物还原为上述式 (II) 中的相应产物。所述多肽通常利用辅助因子即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) 作为还原剂。本文所用的酮还原酶包括天然存在的（野生型的）酮还原酶以及通过人类操作而产生的非天然存在的人工改造的多肽。

[0053] “编码序列”是指编码蛋白的氨基酸序列的那部分核酸（例如基因）。

[0054] “天然存在的”或“野生型的”是指在自然中所发现的形式。例如，天然存在的或野生型的多肽或多核苷酸序列是存在于生物体中的、可从自然中的来源分离的并且未被人类操作有意识地修改的序列。

[0055] “重组的”在提及例如细胞、核酸或多肽而被使用时，是指对应于天然或天生形式的物质的物质，该物质以否则在自然中不存在的方式被修饰，或者与天然或天生形式的物质相同但是是从合成的物质和 / 或通过使用重组技术的操作而产生或衍生出来的。非限制性的实例包括但不限于这样的重组细胞，其表达在天然（非重组）形式的细胞中未被发现的基因或者表达否则以不同水平表达的天然基因。

[0056] “序列同一性百分比”和“百分比同源性”在本文中可被互换使用以提及多核苷酸和多肽中的比较，并且是通过将两个最佳地比对的序列在一个比较窗上进行比较而被确定的，其中比较窗中的多核苷酸或多肽序列的部分与参考序列（其不包括添加或缺失）相比可包括添加或缺失（即，缺口）以将这两个序列最佳地比对。百分比可如下计算：确定两个序列中存在相同核酸碱基或氨基酸残基的位置的数量来得到匹配位置数量，将匹配位置数量除以比较窗中位置的总数量并将该结果乘以 100 以得到序列同一性百分比。可选地，百分比可如下计算：确定两个序列中存在相同核酸碱 基或氨基酸残基或者核酸碱基和氨基酸残基与缺口对齐的位置的数量而产生匹配位置数量，将匹配位置数量除以比较窗中位置的总数量并将该结果乘以 100 以得到序列同一性百分比。本领域技术人员理解，存在许多可用的确立的算法比对两个序列。可通过例如 Smith 和 Waterman，1981，Adv. Appl. Math. 2：482 的局部同源性算法、Needleman 和 Wunsch，1970，J. Mol. Biol. 48：443 的同源性比对算法、Pearson 和 Lipman，1988，Proc. Natl. Acad. Sci. USA 85：2444 的相似性搜索方法、这些算法的计算机化执行 (GCGWisconsin 软件包中的 GAP、BESTFIT、FASTA 和 TFASTA) 或目视检查（通常参见 Current Protocols in Molecular Biology（最新分子生物学实验方法汇编），F. M. Ausubel 等，编辑，Current Protocols（最新实验方法），GreenePublishing Associates，Inc. 和 John Wiley & Sons，Inc. 之间的合资企业，(1995 补编）(Ausubel)) 来进行比较序列的最佳比对。适合于确定百分比序列同一性和序列相似性的算法的实例是 BLAST 和 BLAST 2.0 算法，其分别被描述在 Altschul 等，1990，J. Mol. Biol. 215：403-410 和 Altschul 等，1977，Nucleic Acids Res. 3389-3402 中。用于进行 BLAST 分析的软件可通过美国国家生物技术信息中心网站公开获得。该算法包括首先通过鉴定查询序列中长度 W 的

短字来鉴定高分值序列对 (HSP),其在与数据库序列中相同长度的字比对时匹配或满足一些正值的阈值分 T。T 被称作邻近字分值的阈值 (Altschul 等,如上述)。这些最初的邻近的字符合 (word hit) 被作为启动对更长的含有字符合的 HSP 的搜索的种子 (seed) 而起作用。然后将字符合向双方向沿着每个序列扩展,直到累计比对分值被增加。对于核苷酸序列,使用参数 M(匹配的残基对儿的奖赏分值 ;总是 > 0) 和 N(错配的残基的惩罚分值 ;总是 < 0) 来计算累计分值。对于氨基酸序列,评分矩阵被用来计算累计分值。字符合在各方向的延伸在下面这时停止 :累计比对分值从它所获得的最大值减少了数量 X ;由于一个或多个负分值残基比对的积累导致累计分值降到零或以下 ;或者到达任一个序列的末端。BLAST 算法参数 W、T 和 X 确定比对的灵敏度和速度。BLASTN 程序 (用于核苷酸序列) 默认使用字长度 (W)11、期望值 (E)10、M = 5、N = 4 和两条链的对比。对于氨基酸序列,BLASTP 程序默认使用字长度 (W)3、期望值 (E)10 和 BLOSUM62 评分矩阵 (参见 Henikoff 和 Henikoff, 1989, Proc Natl Acad Sci USA 89 :10915)。序列比对和 % 序列同一性的示例性的确定可利用 GCG Wisconsin 软件包 (Accelrys, Madison WI) 中的 BESTFIT 或 GAP 程序,其使用所提供的默认参数。

[0057] “参考序列”是指被用作序列比较基础的确定序列。参考序列可以是较大序列的亚组 (subset),例如全长基因或多肽序列的一部分。通常,参考序列的长度为至少 20 个核苷酸或氨基酸残基、至少 25 个残基、至少 50 个残基或者全长的核酸或多肽。因为两个多核苷酸或多肽可各自 (1) 包括两个序列之间相似的序列 (即完整序列的一部分),并且 (2) 还包括两个序列之间有分歧的序列,所以通常通过在“比较窗”中比较两个多核苷酸的序列来进行两个 (或更多个) 多核苷酸或多肽之间的序列比较,以便鉴定并比较具有序列相似性的局部区域。

[0058] 在一些实施方案中,“参考序列”可基于一级氨基酸序列,其中所述参考序列是可在一级序列中具有一个或多个改变的序列。例如,“在对应于 X190 的残基处具有脯氨酸的基于 SEQ ID NO :4 的参考序列”是指在其中 SEQ ID NO :4 的 X190 处的相应残基已被改变为脯氨酸的参考序列。

[0059] “比较窗”是指具有至少约 20 个连续的核苷酸位置或氨基酸残基的概念性区段,其中可将序列与具有至少 20 个连续核苷酸或氨基酸的参考序列进行比较,而且其中为了两个序列的最佳比对,比较窗中的序列部分与参考序列 (其不包括添加或缺失) 相比可包括 20% 或更少的添加或缺失 (即缺口)。比较窗可以比 20 个连续残基长,并且包括任选地 30、40、50、100 个或更长的窗。

[0060] “基本同一性”是指与参考序列在具有至少 20 个残基位置的比较窗且常常为至少 30-50 个残基的窗中相比具有至少 80% 序列同一性、至少 85% 序列同一性和 89 至 95% 序列同一性,更通常地至少 99% 序列同一性的多核苷酸或多肽序列,其中通过将参考序列与包括缺失或添加在内的序列进行比较来计算序列同一性百分比,所述缺失或添加在比较窗上总共占参考序列的 20% 或更少。在应用于多肽的特定实施方案中,术语“基本同一性”是指在被任选地比对例如通过程序 GAP 或 BESTFIT 使用默认的缺口权重进行比对时两个多肽序列共享至少 80% 的序列同一性,优选地至少 89% 的序列同一性、至少 95% 的序列同一性或更高 (例如 99% 序列同一性)。优选地,不相同的残基位置的区别在于保守的氨基酸取代。

[0061] “对应于”、“参照”或“相对于”在被用于给定的氨基酸或多核苷酸序列编号的上下文中时是指在将给定的氨基酸或多核苷酸序列与参考序列比较时指定的参考序列的残基编号。换句话说，给定的多聚体的残基编号或残基位置是参照参考序列指定的而非通过在给定的氨基酸或多核苷酸序列中残基的实际编号位置指定的。例如，给定的氨基酸序列例如人工改造的酮还原酶的氨基酸序列，可与参考序列进行比对，其通过引入缺口而将两个序列之间的残基匹配最佳化。在这些情况下，虽然存在缺口，但是给定的氨基酸或多核苷酸序列中残基的编号是参照它与之比对的参考序列而制定的。

[0062] “立体选择性”是指在化学或酶促反应中一种立体异构体与另一种相比优先差异形成。立体选择性可以是部分的，其中对一种立体异构体的形成的偏爱超过另一种；或者它可以是全部的，其中只形成一种立体异构体。当立体异构体是对映体时，立体选择性被称作对映选择性，即一种对映体在两种的总和中的分数（通常被报道为百分数）。它通常可选地在本领域中被报道（通常作为百分数）为对映体过量 (e.e.)，它是根据公式 [主要对映体一次要对映体]/[主要对映体 + 次要对映体] 从中计算而得的。当立体异构体是非对映异构体 (diastereoisomer) 时，立体选择性被称作非对映立体选择性 (diastereoselectivity)，在两种非对映体 (diastereomer) 的混合物中一种非对映体的分数（通常被报道为百分数）通常被可选地报道为非对映体过量 (d.e.)。对映体过量和非对映体过量是立体异构过量的类型。

[0063] “高立体选择性的”是指这样的酮还原酶多肽，其能够将底物转化或还原为具有至少约 99％立体异构过量的相应的 (S)- 产物。

[0064] “立体特异性”是指在化学或酶促反应中一种立体异构体与另一种相比优先差异转化。立体特异性可以是部分的，其中对一种立体异构体的转化的偏爱超过另一种；或者它可以是全部的，其中只转化一种立体异构体。

[0065] “化学选择性”是指在化学或酶促反应中一种产物与另一种相比优先差异形成。

[0066] “改进的酶性质”是指与参考酮还原酶相比展现了任何酶性质改进的酮还原酶多肽。关于本文所述的人工改造的酮还原酶多肽，通常是针对野生型酮还原酶进行比较，虽然在一些实施方案中，参考酮还原酶可以是另一个改进的人工改造的酮还原酶。期望得到改进的酶性质包括但不限于酶活性（其可被表示为底物的转化百分比）、热稳定性、溶剂稳定性、pH 活性概况、辅助因子的要求、对抑制剂的不应性（例如产物抑制）、立体特异性和立体选择性（包括对映体选择性）。

[0067] “增加的酶活性”是指人工改造的酮还原酶多肽的改进的性质，其可被表示为与参考酮还原酶相比之下比活性（例如，所产生的产物 / 时间 / 蛋白重量）的增加或底物向产物的转化百分比（例如，使用特定量的 KRED 在特定时间段内起始量的底物向产物的转化百分比）的增加。确定酶活性的示例性方法在实施例中提供。可影响与酶活性相关的任何特性，包括经典的酶性质 K_m、V_{max} 或 k_{cat}，其改变可导致酶活性的增加。酶活性的改进可以是从约 1.5 倍于相应野生型酮还原酶的酶活性到多达 2 倍、5 倍、10 倍、20 倍、25 倍、50 倍、75 倍、100 倍或更多倍于天然存在的酮还原酶或酮还原酶多肽所衍生自的另一种人工改造的酮还原酶的酶活性。在特定的实施方案中，人工改造的酮还原酶展现改进的酶活性，其范围是超过母体酮还原酶的酶活性 1.5 至 50 倍、1.5 至 100 倍。技术人员所理解的是，任何酶的活性被扩散限制了以至于催化反转速率不能超过底物（其包括任何必需的辅助因子）的扩散速

率。扩散限制的理论最大值或 k_{cat}/K_m 通常是约 10^8 至 $10^9(M^{-1}s^{-1})$。因此，酮还原酶的酶活性的任何改进将具有与所述酮还原酶所作用于其上的底物的扩散速率相关的上限。可通过用于测量酮还原酶的任一个标准测定来测量酮还原酶活性，例如由与酮向醇的还原伴随的其氧化所致的 NADPH 吸光度或荧光的减少，或者通过耦合测定中所产生的产物来测量酮还原酶活性。如本文进一步更详细地描述，使用酶的确定制剂、设定条件下的确定测定和一种或多种确定底物来进行酶活性的比较。通常，当比较裂解产物时，确定了所测定的细胞的数目和蛋白的量，以及使用相同表达系统和相同宿主细胞来将宿主细胞所产生的并存在于裂解产物中的酶量的变化降至最低。

[0068] “转化”是指将底物酶促还原为相应的产物。“转化百分比”是指在特定条件下一段时间内被还原为产物的底物的百分比。因此，酮还原酶多肽的“酶活性”或“活性”可被表达为底物向产物的“转化百分比”。

[0069] “热稳定的”是指与未处理的酶相比在暴露于升高的温度（例如 40-80℃）下一段时间（例如 0.5-24 小时）后保持相似活性（例如超过 60％至 80％）的酮还原酶多肽。

[0070] “溶剂稳定的”是指与未处理的酶相比在暴露于变化的浓度（例如 5-99％）的溶剂（异丙基醇、四氢呋喃、2- 甲基四氢呋喃、丙酮、甲苯、醋酸丁酯、甲基叔丁基醚等）中一段时间（例如 0.5-24 小时）后保持相似活性（超过 60％至 80％）的酮还原酶多肽。

[0071] “pH 稳定的”是指与未处理的酶相比在暴露于高或低 pH（例如 4.5-6 或 8 至 12）中一段时间（例如 0.5-24 小时）后保持相似活性（超过 60％至 80％）的酮还原酶多肽。

[0072] “热和溶剂稳定的”是指既是热稳定的又是溶剂稳定的酮还原酶多肽。

[0073] “衍生自”如本文中在人工改造的酮还原酶的上下文中所用，鉴定所述人工改造所基于的起始酮还原酶和 / 或编码这种酮还原酶的基因。例如，通过人工地使编码 SEQ ID NO：4 中的高加索酸奶乳杆菌酮还原酶的基因经过多代而进化来获得 SEQ ID NO：90 中的人工改造的酮还原酶。因此，该人工改造的酮还原酶“衍生自”SEQ ID NO：4 中的野生型酮还原酶。

[0074] “亲水性氨基酸或残基”是指具有展现小于零的疏水性的侧链的氨基酸或残基，所述疏水性是根据 Eisenberg 等，1984，J.Mol.Biol.179：125-142 中归一化的共有疏水性量表。被遗传编码的亲水性氨基酸包括 L-Thr(T)、L-Ser(S)、L-His(H)、L-Glu(E)、L-Asn(N)、L-Gln(Q)、L-Asp(D)、L-Lys(K) 和 L-Arg(R)。

[0075] “酸性氨基酸或残基”是指当氨基酸被包含在肽或多肽中时具有展现小于约 6 的 pK 值的侧链的亲水性氨基酸或残基。酸性氨基酸在生理 pH 下由于失去了氢离子而通常具有带负电的侧链。被遗传编码的酸性氨基酸包 括 L-Glu(E) 和 L-Asp(D)。

[0076] “碱性氨基酸或残基”是指当氨基酸被包含在肽或多肽中时具有展现大于约 6 的 pK 值的侧链的亲水性氨基酸或残基。碱性氨基酸在生理 pH 下由于缔合了水合氢离子而通常具有带正电的侧链。被遗传编码的碱性氨基酸包括 L-Arg(R) 和 L-Lys(K)。

[0077] “极性氨基酸或残基”是指具有在生理 pH 下不带电的侧链的亲水性氨基酸或残基，但是它具有至少一个这样的键，在其中两个原子所共享的电子对儿与其中一个原子保持得更近。被遗传编码的极性氨基酸包括 L-Asn(N)、L-Gln(Q)、L-Ser(S) 和 L-Thr(T)。

[0078] “疏水性氨基酸或残基”是指具有展现大于零的疏水性的侧链的氨基酸或残基，所述疏水性是根据 Eisenberg 等，1984，J.Mol.Biol.179：125-142 中归一化的共有疏水性量

表。被遗传编码的疏水性氨基酸包括 L-Pro(P)、L-Ile(I)、L-Phe(F)、L-Val(V)、L-Leu(L)、L-Trp(W)、L-Met(M)、L-Ala(A) 和 L-Tyr(Y)。

[0079] “芳香族氨基酸或残基”是指具有包括至少一个芳香族或杂芳香族环的侧链的亲水性或疏水性氨基酸或残基。被遗传编码的芳香族氨基酸包括 L-Phe(F)、L-Tyr(Y) 和 L-Trp(W)。虽然由于它的杂芳香族氮原子的 pKa,L-His(H) 有时候被分类为碱性残基,或者由于它的侧链包括杂芳香族环而被分类为芳香族残基,但是在本文中组氨酸被分类为亲水性残基或“受限残基”(参见如下)。

[0080] “受限氨基酸或残基”是指具有受限的几何形状的氨基酸或残基。在本文中,受限残基包括 L-Pro(P) 和 L-His(H)。组氨酸具有受限的几何形状,因为它具有相对小的咪唑环。脯氨酸具有受限的几何形状,因为它也具有五元环。

[0081] “非极性氨基酸或残基”是指具有在生理 pH 下不带电的侧链的疏水性氨基酸或残基,而且所述侧链具有这样的键,在其中两个原子所共享的电子对几与这两个原子中的每一个通常保持同等距离(即,该侧链不是极性的)。被遗传编码的非极性氨基酸包括 L-Gly(G)、L-Leu(L)、L-Val (V)、L-Ile(I)、L-Met(M) 和 L-Ala(A)。

[0082] “脂肪族氨基酸或残基”是指具有脂肪烃侧链的疏水性氨基酸或残基。被遗传编码的非极性氨基酸包括 L-Ala(A)、L-Val(V)、L-Leu(L) 和 L-Ile(I)。

[0083] “半胱氨酸”或 L-Cys(C) 是不普通的,因为它能够与其他 L-Cys(C) 氨基酸或其他含有硫烷基或硫氢基的氨基酸形成二硫桥。“类半胱氨酸残基”包括半胱氨酸和含有能够形成二硫桥的硫氢基部分的其他氨基酸。L-Cys(C)(和具有含 -SH 的侧链的其他氨基酸)以还原的游离 -SH 或氧化的二硫桥形式存在于肽中的能力影响 L-Cys(C) 是否为肽贡献净疏水或亲水特性。当 L-Cys(C) 展示根据 Eisenberg(Eisenberg 等,1984,见上文)归一化的共有量表为 0.29 的疏水性时,可以理解的是,为了本公开的目的,L-Cys(C) 被分类到它自己的独特的组。

[0084] “小氨基酸或残基”是指具有由总共三个或更少的碳和 / 或杂原子(不包括 α - 碳和氢)所组成的侧链的氨基酸或残基。小氨基酸或残基可根据上述定义被进一步分类为脂肪族、非极性、极性或酸性的小氨基酸或残基。被遗传编码的小氨基酸包括 L-Ala(A)、L-Val(V)、L-Cys(C)、L-Asn(N)、L-Ser(S)、L-Thr(T) 和 L-Asp(D)。

[0085] “含羟基的氨基酸或残基”是指含有羟基 (-OH) 部分的氨基酸。被遗传编码的含羟基的氨基酸包括 L-Ser(S)、L-Thr(T) 和 L-Tyr(Y)。

[0086] “保守的”氨基酸取代或突变是指具有相似侧链的残基的可互换性,并且因此通常包括多肽中的氨基酸被相同或相似的确定分类内的氨基酸所取代。然而,如本文所用,在一些实施方案中,保守的突变不包括从亲水性残基到亲水性残基、疏水性残基到疏水性残基、含羟基的残基到含羟基的残基或者小残基到小残基的取代,如果取而代之地保守的突变可以是从脂肪族残基到脂肪族残基、非极性残基到非极性残基、极性残基到极性残基、酸性残基到酸性残基、碱性残基到碱性残基、芳香族残基到芳香族残基或者受限残基到受限残基的取代。此外,如本文所用,A、V、L 或 I 可被保守地突变为另一种脂肪族残基或另一种非极性残基。下面的表 1 显示示例性的保守取代。

[0087] 表 1 :保守取代

[0088]

残基	可能的保守突变
A,L,V,I	其他脂肪族的 (A,L,V,I) 其他非极性的 (A,L,V,I,G,M)
G,M	其他非极性的 (A,L,V,I,G,M)
D,E	其他酸性的 (D,E)
K,R	其他碱性的 (K,R)
P,H	其他受限的 (P,H)
N,Q,S,T	其他极性的
Y,W,F	其他芳香族的 (Y,W,F)
C	无

[0089] “非保守的取代”是指多肽中的氨基酸向具有显著不同的侧链性质的氨基酸进行的取代或突变。非保守的取代可使用上面所列的确定组之间而非之内的氨基酸。在一个实施方案中，非保守的突变影响 (a) 取代（例如脯氨酸取代甘氨酸）的区域中肽骨架的结构，(b) 电荷或疏水性，或 (c) 侧链的体积。

[0090] “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的修饰。缺失可包括除去 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者 20 个或更多个氨基酸、多达构成多肽的氨基酸总数目的 10％、多达构成多肽的氨基酸总数目的 20％或多达构成多肽的氨基酸总数目的 30％，同时保留酶活性并且 / 或者保留人工改造的酮还原酶的改进的性质。缺失可针对多肽的内部部分和 / 或末端部分。在各种实施方案中，缺失可包括连续的区段或者可以是不连续的。

[0091] “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的修饰。在一些实施方案中，改进的人工改造的酮还原酶包括向天然存在的酮还原酶多肽插入一个或多个氨基酸以及向其他改进的酮还原酶多肽插入一个或多个氨基酸。插入可以是在多肽的内部部分，或者在羧基或氨基末端。本文所用的插入包括本领域已知的融合蛋白。插入可以是连续的氨基酸区段或者是被天然存在的多肽中一个或多个氨基酸所分开的。

[0092] “片段”如本文所用是指具有氨基末端和 / 或羧基末端缺失的多肽，但是其中余下的氨基酸序列与序列中相应位置相同。片段可以是至少 14 个 氨基酸的长度、至少 20 个氨基酸的长度、至少 50 个氨基酸的长度或更长，并且长达 SEQ ID NO :2 或 SEQ ID NO :4 或 SEQ ID NO :114 中天然存在的酮还原酶多肽全长的 70％、80％、90％、95％、98％和 99％。

[0093] “分离的多肽”是指基本上同天然伴随它的其他杂质例如蛋白、脂质和多核苷酸分开的多肽。该术语包括已经从它们天然存在的环境或表达系统（例如宿主细胞或体外合成）中除去或纯化的多肽。改进的酮还原酶可以存在于细胞中，存在于细胞培养基中或者以各种形式例如裂解产物或分离的制剂而被制备。像这样，在一些实施方案中，改进的酮还原酶可以是分离的多肽。

[0094] “基本上纯的多肽”是指这样的组合物，在其中所述多肽种类是所存在的优势种类（即在摩尔或重量基础上，它比组合物中任何其他单独的大分子种类的含量更多），而且当目标种类占存在的大分子种类至少约 50％（以摩尔或％重量）时所述组合物通常是基本上纯化的组合物。通常，基本上纯的酮还原酶组合物将包括约 60％或更多、约 70％或更多、约 80％或更多、约 90％或更多、约 95％或更多和约 98％或更多的（以摩尔或％重量）存在于组合物中的所有大分子种类。在一些实施方案中，目标种类被纯化为大致均一的（即通过常规检测方法不能在组合物中检测到杂质种类），其中组合物基本上由单独的大分子种类所组成。溶剂种类、小分子（＜ 500 道尔顿）和元素离子种类不被视为大分子种类。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。

[0095] “严格杂交”在本文中被用来指这样的条件，在其中核酸杂交体是稳定的。如本领域技术人员所知，杂交体的稳定性反映在杂交体的解链温度（T_m）上。通常，杂交体的稳定性是离子强度、温度、G/C 含量和离液剂的存在的函数。可使用用于预测解链温度的已知方法来计算多核苷酸的 T_m 值（参见例如 Baldino 等，Methods Enzymology 168 :761-777 ; Bolton 等，1962，Proc.Natl.Acad.Sci.USA 48 :1390 ;Bresslauer 等，1986，Proc.Natl.Acad.SciUSA 83 :8893-8897 ;Freier 等，1986，Proc.Natl.Acad.Sci USA 83 :9373-9377 ; Kierzek 等，Biochemistry 25 :7840-7846 ;Rychlik 等，1990，Nucleic Acids Res18 : 6409-6412（勘误表，1991，Nucleic Acids Res 19 :698) ;Sambrook 等，见上 文 ;Suggs 等，1981，在Developmental Biology Using Purified Genes（使用纯化的基因的发育生物学）中 (Brown 等，编辑），第 683-693 页，AcademicPress ;和 Wetmur，1991，Crit Rev Biochem Mol Biol 26 :227-259。所有出版物通过引用并入本文）。在一些实施方案中，多核苷酸编码本文所公开的多肽，并在确定条件（例如适度严格的或高度严格的条件）下与编码本公开的人工改造的酮还原酶的序列的互补体进行杂交。

[0096] “杂交严格性”涉及核酸的杂交中的杂交条件，例如洗涤条件。通常，在较低严格性条件下进行杂交反应，接着在不同的但是更高的严格性下洗涤。术语“适度严格的杂交”是指允许靶 DNA 结合互补核酸的条件，其中所述互补核酸与靶 DNA 具有大约 60％的同一性、优选地大约 75％的同一性、大约 85％的同一性 ;与靶多核苷酸具有超过大约 90％的同一性。示例性的适度严格的条件是这样的条件，其等同于在 42℃下在 50％甲酰胺、5× 登纳特氏溶液、5×SSPE、0.2％ SDS 中杂交，接着在 42℃下用 0.2×SSPE、0.2％ SDS 洗涤。“高度严格的杂交”通常是指与在溶液条件下针对确定的核苷酸序列所确定的热解链温度 T_m 相差约 10℃或更少的条件。在一些实施方案中，高度严格的杂交是指仅允许在 65℃下 0.018M NaCl 中形成稳定杂交体的那些核酸序列进行杂交的条件。（即，如果杂交体在 65℃下 0.018MNaCl 中不稳定，它在如本文所包括的高度严格的条件下将不稳定）。可通过例如在等同于在 42℃下在 50％甲酰胺、5× 登纳特氏溶液、5×SSPE、0.2％ SDS 并接着在 65℃下用 0.1×SSPE 和 0.1％ SDS 洗涤的条件下杂交来提供高度严格的条件。另一种高度严格的条件是在等同于在 65℃下在含有 0.1％ (w:v)SDS 的 5×SSC 中杂交并在 65℃下用含有 0.1％ SDS 的 0.1×SSC 洗涤的条件下杂交。其他高度严格的杂交条件以及适度严格的条件在上面所引用的参考文献中描述。

[0097] “异源性的”多核苷酸是指通过实验室技术被引入宿主细胞的任何多核苷酸，而且包括从宿主细胞中取出，进行实验室操作并且之后被重新引入宿主细胞的多核苷酸。

[0098] “密码子优化的”是指将编码蛋白的多核苷酸的密码子改变为在特定生物体中优先使用的那些，以使被编码的蛋白被有效地表达在感兴趣的生 物体中。虽然遗传密码因为大多数氨基酸由若干密码子（被称作“同义密码子”或“同义的”密码子）而是简并的 (degenerate)，但是众所周知特定生物体的密码子选择是非随机的并且偏向于特定密码子三联体。根据给定基因、具有共同功能或祖先起源的基因、高表达蛋白与低拷贝数蛋白的对比以及生物体基因组的聚集蛋白编码区，密码子选择偏向可能更高。在一些实施方案中，编码酮还原酶的多核苷酸可以是密码子优化的以从被选作表达的宿主生物体最优化地生产。

[0099] “优选的”、“最优化的”或“高度密码子选择偏向的密码子”可互换地指与编码相

同氨基酸的其他密码子相比在蛋白编码区中以更高频率使用的密码子。可根据单个基因、具有共同功能或起源的一组基因、高表达的基因、整个生物体的聚集蛋白编码区中的密码子频率、相关生物体的聚集蛋白编码区中的密码子频率或其组合来确定优选的密码子。其频率与基因表达水平一起增加的密码子通常是表达的最优化的密码子。已经知道若干方法用于确定密码子频率（例如，密码子选择、相对同义的密码子选择）和特定生物体中的密码子偏好，其包括多变量分析，例如使用聚类分析或对应分析和用于基因中的密码子的有效数目（参见 GCG CodonPreference，Genetics Computer Group Wisconsin Package；CodonW，John Peden，University of Nottingham；McInerney，J.O，1998，Bioinformatics 14：372-73；Stenico 等，1994，Nucleic Acids Res. 222437-46；Wright，F.，1990，Gene87：23-29）。对于越来越多的生物体可获得密码子选择表（参见例如 Wada 等，1992，Nucleic Acids Res. 20：2111-2118；Nakamura 等，2000，Nucl. Acids Res. 28：292；Duret，等，见上文；Henaut 和 Danchin，“Escherichia coli andSalmonella，”1996，Neidhardt，等编辑，ASM Press，Washington D.C.，第 2047-2066 页）。获得密码子选择的数据来源可依靠能够编码蛋白的任何可得的核苷酸序列。这些数据集包括实际上已知编码被表达蛋白的核酸序列（例如，完整的编码蛋白的序列 -CDS）、被表达的序列标签 (ESTS) 或预测的基因组序列编码区（参见例如 Mount，D.，Bioinformatics：Sequence andGenome Analysis，第 8 章，Cold Spring Harbor Laboratory Press，Cold SpringHarbor，N.Y.，2001；Uberbacher，E.C.，1996，Methods Enzymol. 266：259-281；Tiwari 等.，1997，Comput. Appl. Biosci. 13：263-270）。

[0100] “控制序列”在本文中被定义为包括这样的所有成分，其对于感兴趣的多核苷酸和 / 或多肽的表达是必要的或有利的。每个控制序列对于编码多肽的核酸序列可以是天生的或外来的。这种控制序列包括但不限于前导序列、多腺苷酸化序列、前肽序列、启动子、信号肽序列和转录终止子。至少，控制序列包括启动子以及转录和翻译终止信号。为了引入特定的限制位点以促进控制序列与编码多肽的核酸序列的编码区连接，可为控制序列提供连接序列。

[0101] “可操作地连接的”在本文中被定义为这样的构型，在其中将控制序列适当置于（即，以功能的关系）相对于感兴趣的多核苷酸的一个位置以使所述控制序列指导或调节感兴趣的多核苷酸和 / 或多肽的表达。

[0102] “启动子序列”是被宿主细胞识别以表达感兴趣的多核苷酸（例如编码序列）的核酸序列。控制序列可包括适当的启动子序列。启动子序列含有转录控制序列，其介导感兴趣的多核苷酸的表达。启动子可以是在所选的宿主细胞中显示转录活性的任何核酸序列，其包括突变、截短的和杂交的启动子，并且可从编码与宿主细胞同源或异源的细胞外或细胞内多肽的基因获得。

[0103] 6.2. 酮还原酶

[0104] 本公开提供人工改造的酮还原酶（“KRED”），其能够立体选择性地将底物 2-[3-[3-[2-(7- 氯 -2- 喹啉基）乙烯基］苯基］-3- 氧代丙基］苯甲酸甲酯还原或转化为产物 (S，E)-2-(3-(3-(2-(7- 氯喹啉 -2- 基）乙烯基）苯基）-3- 羟丙基）苯甲酸甲酯，而且与从高加索酸奶乳杆菌 (SEQ ID NO：4) 或短乳杆菌 (SEQ IDNO：2) 或微小乳杆菌 (SEQ ID NO：114) 获得的天然存在的野生型 KRED 酶相比或者与其他人工改造的酮还

原酶相比，其具有改进的特性。所述产物化合物在孟鲁司特的合成中是有用的，孟鲁司特也被称作 Singulari ®这个名字，即被用作治疗过敏症和控制哮喘的治疗剂的一种白三烯受体拮抗剂。如本文所示，来自高加索酸奶乳杆菌的野生型酶不以任何显著的速率将 2-[3-[3-[2-(7- 氯 -2- 喹啉基) 乙烯基] 苯基]-3- 氧代丙基] 苯甲酸甲酯还原或转化为产物 (S,E)-2-(3-(3-(2-(7- 氯喹啉 -2- 基) 乙烯基) 苯基)-3- 羟丙基) 苯甲酸甲酯。与此相反，本公开的酮还原酶能够将特定底物还原或转化为明显水 平的产物。在一些实施方案中，本公开的酮还原酶多肽具有如下要求 :(1) 对应于 SEQ ID NO :2、4 或 114 的位置 190(即 X190) 的残基是受限残基 ;而且 (2) 对应于 SEQ ID NO :2、4 或 114 的位置 202(即 X202) 的残基是芳香族残基。在一些实施方案中，改进的酮还原酶多肽具有如下要求 :(1) 对应于 SEQ ID NO :2、4 或 114 的位置 190(即 X190) 的残基是受限残基 ;

[0105] (2) 对应于 SEQ ID NO :2、4 或 114 的位置 195(即 X195) 的残基是碱性的、酸性的、非极性的或极性的残基 ;而且 (3) 对应于 SEQ ID NO :2、4 或 114 的位置 202(即 X202) 的残基是芳香族残基。

[0106] 如上所述，可针对高加索酸奶乳杆菌、短乳杆菌或微小乳杆菌的天然存在的酮还原酶 (也被称作“ADH”或“醇脱氢酶”) 或另外人工改造的酮还原酶的氨基酸序列来描述本公开的酮还原酶。像这样，在从起始的甲硫氨酸 (M) 残基 (即 M 代表残基位置 1) 开始的酮还原酶中确定氨基酸残基位置，虽然技术人员将理解可通过生物加工装置在例如宿主细胞或体外翻译系统中除去该起始的甲硫氨酸残基以便产生缺少起始的甲硫氨酸残基的成熟蛋白。特殊氨基酸或氨基酸改变存在于氨基酸序列中的氨基酸残基位置在本文中有时候被描述为术语“Xn”或“位置 n”，其中 n 是指残基位置。当相同残基位置处的氨基酸残基在酮还原酶之间有所不同时，不同的残基可表示为含有“/”的排列“高加索酸奶乳杆菌残基 / 短乳杆菌残基 / 微小乳杆菌残基”。取代突变，即不同氨基酸残基对参考序列例如 SEQID NO :2 和 SEQ ID NO :4 和 SEQ ID NO :114 的野生型酮还原酶中的氨基酸残基的替换，可被表示为符号“→”。在本文中，突变有时候被描述为突变“至一个”类型的氨基酸。例如，SEQ ID NO :4 的残基 96 可被突变“至一个”脂肪族残基。但是短语“至一个”的使用不排除从一类的一个氨基酸突变至同类的另一个氨基酸。例如，SEQ ID NO :4 的残基 96 是极性残基丝氨酸，但是它可被突变为不同的极性残基 ;例如，所述突变可以是“S96T”(96 → T) 突变。可从已知编码酮还原酶活性的多核苷酸来获得高加索酸奶乳杆菌、短乳杆菌或微小乳杆菌的天然存在的酮还原酶 (也被称作“ADH”或“醇脱氢酶”) 的氨基酸序列 (例如，关于高加索酸奶乳杆菌，Genbank 登录号 AAP94029 GI :33112056 或 SEQ ID NO :3 ;关于短乳杆菌，Genbank 登录号 CAD66648 GI :28400789 或 SEQ ID NO :1 ;以及关于微小乳 杆菌，SEQ ID NO :113)。

[0107] 在一些实施方案中，与野生型或另外人工改造的多肽相比，酮还原酶多肽的改进的特性是关于它将式 (I) 中的底物 2-[3-[3-[2-(7- 氯 -2- 喹啉基) 乙烯基] 苯基]-3- 氧代丙基] 苯甲酸甲酯还原或转化为其相应的式 (II) 中的 (S)- 醇产物 (S,E)-2-(3-(3-(2-(7- 氯喹啉 -2- 基) 乙烯基) 苯基)-3- 羟丙基) 苯甲酸甲酯的活性的增加。在一些实施方案中，酶活性的增加可体现为底物向产物转化的速率的增加，或体现为与野生型或另外的参考序列 (例如 SEQ IDNO :90) 相比在还原或转化相同量产物中使用较少的改进的多肽的能力。在一些实施方案中，酮还原酶性质的改进的性质是关于还原底物至产物中立体选择性的增加，即，这里是产物的立体异构过量的增加。在一些实施方案中，酮

还原酶多肽的改进的性质是关于它的稳定性或热稳定性。在一些实施方案中，酮还原酶多肽具有多于一个改进的性质，例如改进的酶活性和改进的立体选择性。

[0108] 在一些实施方案中，本文的酮还原酶多肽可具有对参考序列（例如天然存在的多肽或人工改造的多肽）的若干修饰，其在一些实施方案中可产生改进的酮还原酶性质。如本文所用，“修饰”包括氨基酸取代、缺失和插入。可将任一个或一组合的修饰引入天然存在的或人工改造的多肽以产生人工改造的酶。在这种实施方案中，对氨基酸序列修饰的数目可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 个或更多个氨基酸或者 20 个或更多个氨基酸，参考多肽序列的氨基酸总数目的至多 10％、氨基酸总数目的至多 15％、氨基酸总数目的至多 20％或氨基酸总数目的至多 30％。在一些实施方案中，产生改进的酮还原酶性质的对天然存在的多肽或人工改造的多肽的修饰的数目可包括约 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个参考序列修饰。在一些实施方案中，修饰的数目可以是 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个氨基酸残基。所述修饰可包括插入、缺失、取代或其组合。

[0109] 在一些实施方案中，修饰包括对参考序列的氨基酸取代。可产生改进的酮还原酶性质的取代可以在一个或多个氨基酸、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者 20 个或更多个氨基酸，参考酶序列的氨基酸总数目的至多 10％、氨基酸总数目的至多 20％或氨基酸总数目的至多 30％。在一些实施方案中，产生改进的酮还原酶性质的对天然存在的多肽或人工改造的多肽的取代的数目可包括参考序列的约 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个氨基酸取代。在一些实施方案中，取代的数目可以是 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个氨基酸残基。

[0110] 在一些实施方案中，本文的酮还原酶多肽包括与具有下述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少约 80％、85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或更多地相同的氨基酸序列；对应于位置 X190 的残基是受限残基，特别是脯氨酸；而且对应于 X202 的残基是芳香族残基，特别是色氨酸；条件是酮还原酶多肽具有这样的氨基酸序列，在其中对应于 X190 的残基是受限残基，而且对应于 X202 的残基是芳香族残基。在一些实施方案中，对应于 X190 的残基是脯氨酸。在一些实施方案中，对应于 X202 的残基是色氨酸。在一些实施方案中，酮还原酶多肽具有这样的氨基酸序列，在其中对应于 X190 的残基是脯氨酸，而且对应于 X202 的残基是色氨酸。在一些实施方案中，这些酮还原酶多肽与参考氨基酸序列相比可在其他残基位置上具有一个或多个残基差异。该差异可包括各种修饰，例如取代、缺失和插入。取代可以是非保守取代、保守取代或者非保守取代和保守取代的组合。在一些实施方案中，这些酮还原酶多肽可任选地与参考序列相比在其他氨基酸残基位置上具有约 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35、1-40、 1-45 或约 1-50 个残基差异。在一些实施方案中，与参考序列相比在其他氨基酸残基上的差异的数

目可以是1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35、40、45或50个残基差异。

[0111] 在一些实施方案中，本文的酮还原酶多肽包括与具有下述特征的基于SEQ ID NO：2、4或114的参考序列至少约80％、85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或更多地相同的氨基酸序列；对应于位置X190的残基是受限残基，特别是脯氨酸；对应于X195的残基是碱性的、酸性的、非极性的或极性的残基，特别是脯氨酸、精氨酸、赖氨酸、天冬氨酸或天冬酰胺；而且对应于X202的残基是芳香族残基，特别是色氨酸；条件是酮还原酶多肽包括至少具有上述特征的氨基酸序列，即，对应于X190的残基是受限残基；对应于X195的残基是碱性的、酸性的、非极性的或极性的残基；而且对应于X202的残基是芳香族残基。在一些实施方案中，对应于X190的残基是脯氨酸。在一些实施方案中，对应于X195的残基是精氨酸。在一些实施方案中，对应于X202的残基是色氨酸。在一些实施方案中，酮还原酶多肽具有这样的氨基酸序列，在其中对应于X190的残基是脯氨酸；对应于X195的残基是脯氨酸、精氨酸、赖氨酸、天冬氨酸或天冬酰胺，特别是精氨酸；而且对应于X202的残基是色氨酸。在一些实施方案中，这些酮还原酶多肽与参考氨基酸序列相比可在其他残基位置上具有一个或多个残基差异。该差异可包括各种修饰，例如取代、缺失和插入。取代可以是非保守取代、保守取代或者非保守取代和保守取代的组合。在一些实施方案中，这些酮还原酶多肽可任选地与参考序列相比在其他氨基酸残基位置上具有约1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35、1-40、1-45或约1-50个残基差异。在一些实施方案中，与参考序列相比在其他氨基酸残基位置上的差异的数目可以是1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35、40、45或50个残基差异。

[0112] 在一些实施方案中，改进的酮还原酶多肽包括基于SEQ ID NO：111或SEQ ID NO：112或SEQ ID NO：139或其区域（例如对应于90-211的残基）中所陈列的序列式的氨基酸序列。SEQ ID NO：112基于高加索酸奶乳杆菌酮还原酶的野生型氨基酸序列(SEQ ID NO：4)；SEQ ID NO：111基于短乳杆菌酮还原酶的野生型氨基酸序列(SEQ ID NO：2)；而且SEQ ID NO：139基于微小乳杆菌酮还原酶的野生型氨基酸序列(SEQ ID NO：114)。基于SEQ ID NO：111、112或139中的序列式的酮还原酶说明，对应于X190的残基是受限残基；对应于X195的残基是碱性的、酸性的、非极性的或极性的残基；而且X202是芳香族残基。在SEQ ID NO：111、112或139中的序列式的一些实施方案中，对应于X190的残基是脯氨酸；而且对应于X202的残基是色氨酸。在SEQ ID NO：111、112或139中的序列式的一些实施方案中，对应于X190的残基是脯氨酸；对应于X195的残基是脯氨酸、精氨酸、赖氨酸、天冬氨酸或天冬酰胺，特别是精氨酸；而且对应于X202的残基是色氨酸。如下所述，该序列式还说明了各种残基位置的特征。

[0113] 在一些实施方案中，包括基于SEQ ID NO：111、112或139或其区域（例如残基90-211）中的序列式的氨基酸序列并且具有本文所述对应于X190和X202的残基的特定特征或对应于X190、X195和X202的残基的特定特征的改进的酮还原酶多肽，还可包括如序列式中所提供的一个或多个选自如下的特征：对应于X7的残基是受限的或非极性的残基；对应于X17的残基是非极性的或脂肪族的残基；对应于X25的残基是极性的或酸性的残基；对应于X27的残基是芳香族的残基；对应于X41的残基是极性的、脂肪族的或非极性

的残基；对应于 X53 的残基是极性的、非极性的或酸性的残基；对应于 X57 的残基是非极性的或脂肪族的残基；对应于 X64 的残基是极性的、脂肪族的或非极性的残基；对应于 X66 的残基是酸性的残基；对应于 X68 的残基是极性的、酸性的、脂肪族的或非极性的残基；对应于 X72 的残基是碱性的残基；对应于 X76 的残基是脂肪族的、非极性的或极性的残基；对应于 X80 的残基是脂肪族的、非极性或极性的残基；对应于 X93 的残基是极性的、脂肪族的或非极性的残基；对应于 X94 的残基是半胱氨酸、脂肪族的或非极性的残基；对应于 X95 的残基是脂肪族的、非极性的或酸性的残基；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基；对应于 X97 的残基是碱性的、受限的或极性的残基； 对应于 X100 的残基是碱性的、酸性的、极性的、脂肪族的或非极性的残基；对应于 X108 的残基是碱性的或受限的残基；对应于 X111 的残基是非极性的或脂肪族的残基；对应于 X139 的残基是非极性的或脂肪族的残基；对应于 X145 的残基是酸性的、脂肪族的或非极性的残基；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基；对应于 X148 的残基是非极性的或脂肪族的残基；对应于 X151 的残基是受限的、非极性的或脂肪族的残基；对应于 X152 的残基是极性的、芳香族的、脂肪族的或非极性的残基；对应于 X163 的残基是非极性的或脂肪族的残基；对应于 X165 的残基是极性的、非极性的或脂肪族的残基；对应于 X173 的残基是酸性的、非极性的或脂肪族的残基；对应于 X179 的残基是芳香族的、碱性的、脂肪族的或非极性的残基；对应于 X194 的残基是受限的、极性的、脂肪族的或非极性的残基；对应于 X196 的残基是非极性的或脂肪族的残基；对应于 X197 的残基是酸性的残基；对应于 X198 的残基是酸性的或极性的残基；对应于 X210 的残基是极性的或碱性的残基；对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基；对应于 X212 的残基是极性的、脂肪族的或非极性的残基；对应于 X216 的残基是受限的或碱性的残基；对应于 X223 的残基是非极性的或脂肪族的残基；对应于 X226 的残基是非极性的或极性的残基；对应于 X233 的残基是极性的或酸性的残基；对应于 X235 的残基是极性的或芳香族的残基；对应于 X248 的残基是非极性的或芳香族的残基；而且对应于 X249 的残基是芳香族的残基。在一些实施方案中，氨基酸序列可具有至少 2、3、4、5、6 或更多种特征。在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139（或其区域）中的序列式的氨基酸序列的多肽与 SEQ ID NO：2、4 或 114 中的参考序列相比可另外在未被上面的 X 所指出的残基位置上具有一个或多个残基差异。在一些实施方案中，所述差异可以是在未被上面的 X 所限定的其他氨基酸残基位置上 的 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基位置上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异 包括保守突变。

[0114] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列的多肽与 SEQ ID NO：2、4 或 114 中的氨基酸序列相比可具有一个或多个保守突变。示例性的保守突变包括氨基酸替换，例如但不限于：用另一种芳香族残基例如酪氨酸替换对应于 X27 的残基苯丙氨酸 (F)；用另一种非极性的或脂肪族残基例如缬氨酸替换对应于 X57 的残基异亮氨酸 (I)；用另一种非极性的或脂肪族残基例如缬氨酸替换对应于 X64 的残基丙氨酸 (A)；用另一种酸性的残基例如谷氨酸替换对应于 X66 的残基天冬氨酸 (D)；用另一种碱性残基例如精氨酸替换对应于 X72 的残基赖氨酸 (K)；用

另一种非极性的或脂肪族残基例如缬氨酸替换对应于 X80 的残基丙氨酸 (A) ;用另一种极性的残基例如丝氨酸替换对应于 X93 的残基异亮氨酸 (I) ;用另一种非极性的残基例如甘氨酸替换对应于 X94 的残基丙氨酸 (A) ;用另一种非极性的或脂肪族残基例如亮氨酸替换对应于 X95 的残基缬氨酸 (V) ;用另一种极性的残基例如苏氨酸替换对应于 X96 的残基丝氨酸 (S) ;用另一种非极性的或脂肪族残基例如异亮氨酸替换对应于 X148 的残基缬氨酸 (V) ;用另一种极性的残基例如丝氨酸替换对应于 X152 的残基苏氨酸 ;用另一种非极性的或脂肪族残基例如异亮氨酸替换对应于 X163 的残基缬氨酸 (V) ;用另一种芳香族残基例如苯丙氨酸替换对应于 X179 的残基酪氨酸 (Y) ;用另一种脂肪族残基例如亮氨酸替换对应于 X196 的残基缬氨酸 (V) ;用另一种酸性的残基例如谷氨酸替换对应于 X198 的残基天冬氨酸 (D) ;用另一种非极性的或脂肪族残基例如缬氨酸替换对应于 X223 的残基异亮氨酸 (I) ;以及用另一种芳香族残基例如色氨酸替换对应于 X249 的残基酪氨酸 (Y)。

[0115] 在一些实施方案中,包括基于 SEQ ID NO :111、112 或 139 或其区域 (例如残基 90-211) 中的序列式的氨基酸序列的,具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的酮还原酶多肽,还可包括一个或多个选自如下的特征 :对应于 X7 的残基是组氨酸、脯氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是甘氨酸或组氨酸 ;对应于 X17 的残基是甘氨酸、甲 硫氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸,特别是甲硫氨酸 ;对应于 X25 的残基是谷氨酸、天冬氨酸、丝氨酸、苏氨酸、谷氨酰胺或天冬酰胺,特别是苏氨酸或天冬氨酸 ;对应于 X27 的残基是酪氨酸、苯丙氨酸、色氨酸,特别是酪氨酸 ;对应于 X41 的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是丝氨酸或丙氨酸 ;对应于 X53 的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、天冬氨酸或谷氨酸,特别是苏氨酸、甘氨酸或天冬氨酸 ;对应于 X57 的残基是甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是缬氨酸 ;对应于 X64 的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是丝氨酸、丙氨酸或缬氨酸 ;对应于 X66 的残基是天冬氨酸或谷氨酸 ;对应于 X68 的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、天冬氨酸、谷氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是苏氨酸、天冬氨酸、丙氨酸、缬氨酸或谷氨酸 ;对应于 X72 的残基是赖氨酸或精氨酸,特别是精氨酸 ;对应于 X76 的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、甘氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是丙氨酸或苏氨酸 ;对应于 X80 的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是丙氨酸、苏氨酸或缬氨酸 ;对应于 X93 的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是异亮氨酸、丝氨酸、丙氨酸或甘氨酸 ;对应于 X94 的残基是半胱氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是半胱氨酸、甘氨酸或丙氨酸 ;对应于 X95 的残基是天冬氨酸、谷氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是缬氨酸、亮氨酸或谷氨酸 ;对应于 X96 的残基是半胱氨酸、丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是天冬酰胺、丝氨酸、缬氨酸、亮氨酸、异亮氨酸或苏氨酸 ;对应于 X97 的残基是赖氨酸、精氨酸、组氨酸、脯氨酸、丝氨酸、苏氨酸、谷氨酰胺或天冬酰胺,特别是赖氨酸、组氨酸或丝氨酸 ;对应于 X100 的残基是赖氨酸、精氨酸、 丝

氨酸、苏氨酸、谷氨酰胺、天冬酰胺、甘氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是谷氨酸、赖氨酸、谷氨酰胺或丙氨酸；对应于 X108 的残基是精氨酸、赖氨酸、组氨酸或脯氨酸，特别是精氨酸或组氨酸；对应于 X111 的残基是甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是甲硫氨酸；对应于 X139 的残基是甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是亮氨酸；对应于 X145 的残基是天冬氨酸、谷氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是谷氨酸或亮氨酸；对应于 X147 的残基是酪氨酸、苯丙氨酸、色氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是苯丙氨酸或亮氨酸；对应于 X148 的残基是甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是异亮氨酸；对应于 X151 的残基是脯氨酸、组氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是丙氨酸；对应于 X152 的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、酪氨酸、苯丙氨酸、色氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是苏氨酸、色氨酸、丙氨酸或丝氨酸；对应于 X163 的残基是甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是异亮氨酸；对应于 X165 的残基是丝氨酸、苏氨酸、谷氨酰胺天冬酰胺、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是天冬酰胺；对应于 X173 的残基是天冬氨酸、谷氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸，特别是缬氨酸；对应于 X179 的残基是酪氨酸、苯丙氨酸、色氨酸、精氨酸、赖氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是酪氨酸、苯丙氨酸、赖氨酸或亮氨酸；对应于 X194 的残基是脯氨酸、组氨酸、丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是脯氨酸、谷氨酰胺、丝氨酸或亮氨酸；对应于 X196 的残基是甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是亮氨酸；对应于 X197 的残基是天冬氨酸或谷氨酸，特别是谷氨酸；对应于 X198 的残基是天冬氨酸、谷氨酸、丝氨酸、苏氨酸、谷氨酰胺或天冬酰胺，特别是天冬氨酸、谷氨酸或谷氨酰胺；对应于 X210 的残基是精氨酸、赖氨酸、丝氨酸、苏氨酸、谷氨酰胺或天冬酰胺，特别是苏氨酸或精氨酸；对应于 X211 的残基是赖氨酸、精氨酸、天冬氨酸、谷氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸、苏氨酸、谷氨酰胺或天冬酰胺，特别是赖氨酸、缬氨酸、亮氨酸、精氨酸、苏氨酸、天冬氨酸或异亮氨酸；对应于 X212 的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是苏氨酸、丝氨酸或缬氨酸；对应于 X216 的残基是脯氨酸、组氨酸、精氨酸或赖氨酸，特别是组氨酸或精氨酸；对应于 X223 的残基是甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是缬氨酸；对应于 X226 的残基是甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸、苏氨酸、谷氨酰胺或天冬酰胺，特别是苏氨酸；对应于 X233 的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、天冬氨酸、谷氨酸，特别是天冬酰胺或天冬氨酸；对应于 X235 的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、酪氨酸、苯丙氨酸、色氨酸，特别是丝氨酸或色氨酸；对应于 X248 的残基是甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、酪氨酸、苯丙氨酸或色氨酸，特别是甘氨酸或色氨酸；并且对应于 X249 的残基是酪氨酸、苯丙氨酸或色氨酸，特别是色氨酸。在一些实施方案中，氨基酸序列可具有至少 2、3、4、5、6 或更多种由上述 X 所定义的另外的特征。在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139（或其区域）中的序列式的氨基酸序列的酮还原酶多肽与 SEQ ID NO：2、4 或 114 中的参考序列相比在未被上面的 X 所指出的残基位置上可另外具有一个或多个残基差异。在一些

实施方案中，所述差异可以是在未被上面的 X 所限定的其他氨基酸残基位置上的 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基位置上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或 40 个残基差异。在一些实施方案中，所述差异包括保守突变。

[0116] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特 征的改进的酮还原酶，可另外具有一个或多个或至少所有的下述特征：对应于 X7 的残基是受限的或非极性的残基，特别是组氨酸；对应于 X94 的残基是半胱氨酸、脂肪族的或非极性的残基，特别是半胱氨酸或甘氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是苏氨酸、半胱氨酸、缬氨酸、亮氨酸或异亮氨酸；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基，特别是亮氨酸。在一些实施方案中，酮还原酶多肽除了前述特征之外可具有一种或多种如下特征：对应于 X93 的残基是极性的、脂肪族的或非极性的残基，特别是丝氨酸；对应于 X194 的残基是受限的、极性的、脂肪族的或非极性的残基，特别是亮氨酸、丝氨酸或谷氨酰胺；对应于 X196 的残基是非极性的或脂肪族的残基，特别是亮氨酸；并且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是缬氨酸、亮氨酸、异亮氨酸或苏氨酸。在一些实施方案中，酮还原酶多肽除了前述特征之外可具有一种或多种如下特征：对应于 X17 的残基是非极性的或脂肪族的残基，特别是甲硫氨酸；对应于 X27 的残基是芳香族的残基，特别是酪氨酸；对应于 X64 的残基是极性的、非极性的或脂肪族的残基，特别是缬氨酸；对应于 X80 的残基是脂肪族的、非极性的或极性的残基，特别是苏氨酸；对应于 X100 的残基是碱性的、酸性的、极性的、脂肪族的或非极性的残基，特别是赖氨酸；对应于 X198 的残基是酸性的或极性的残基，特别是谷氨酸；并且对应于 X216 的残基是受限的或碱性的残基，特别是精氨酸。在一些实施方案中，酮还原酶多肽除了前述特征之外可具有一种或多种如下特征：对应于 X25 的残基是极性的或酸性的残基，特别是苏氨酸；对应于 X41 的残基是极性的、脂肪族的或非极性的残基，特别是丝氨酸或丙氨酸；对应于 X53 的残基是极性的、非极性的或酸性的残基，特别是天冬氨酸；对应于 X57 的残基是非极性的或脂肪族的残基，特别是缬氨酸；对应于 X66 的残基是酸性的残基，特别是谷氨酸；对应于 X68 的残基是极性的、酸性的、脂肪族的或非极性的残基，特别是缬氨酸；对应于 X72 的残基是碱性的残基，特别是精氨酸；对应于 X76 的残基是脂肪族的、非极性的或极性的残基，特别是丙氨酸；对应于 X95 的残基是脂肪族的、非极性的或酸性的残基，特别是亮氨酸或谷氨酸；对应于 X97 的残基是碱性的、受限的或极性的残基，特别 是组氨酸；对应于 X108 的残基是碱性的或受限的残基，特别是组氨酸；对应于 X111 的残基是非极性的或脂肪族的残基，特别是甲硫氨酸；对应于 X139 的残基是非极性的或脂肪族的残基，特别是亮氨酸；对应于 X145 的残基是酸性的、脂肪族的或非极性的残基，特别是亮氨酸或谷氨酰胺；对应于 X148 的残基是非极性的或脂肪族的残基，特别是异亮氨酸；对应于 X151 的残基是受限的、非极性的或脂肪族的残基，特别是丙氨酸；对应于 X152 的残基是极性的、芳香族的、脂肪族的或非极性的残基，特别是丝氨酸；对应于 X163 的残基是非极性的或脂肪族的残基，特别是异亮氨酸；对应于 X165 的残基是极性的、非极性的或脂肪族的残基，特别是天冬酰胺；对应于 X173 的残基是

酸性的、非极性的或脂肪族的残基，特别是缬氨酸；对应于 X179 的残基是芳香族的、碱性的、脂肪族的或非极性的残基，特别是苯丙氨酸；对应于 X197 的残基是酸性的残基，特别是谷氨酸；对应于 X210 的残基是极性的或碱性的残基，特别是精氨酸；对应于 X212 的残基是极性的、脂肪族的或非极性的残基，特别是缬氨酸或丝氨酸；对应于 X223 的残基是非极性的或脂肪族的残基，特别是缬氨酸；对应于 X226 的残基是非极性的或极性的残基，特别是苏氨酸；对应于 X233 的残基是极性的或酸性的残基，特别是天冬酰胺或天冬氨酸；对应于 X235 的残基是极性的或芳香族的残基，特别是色氨酸；对应于 X248 的残基是非极性的或芳香族的残基，特别是甘氨酸或色氨酸；并且对应于 X249 的残基是芳香族的残基，特别是色氨酸。对于本领域技术人员将是明显的是，前述组中特征的各种组合可被用来形成本公开的酮还原酶。

[0117] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还包括一个或多个或至少所有的下述另外特征：对应于 X94 的残基是半胱氨酸、脂肪族的或非极性的残基特别是半胱氨酸或甘氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基特别是苏氨酸、半胱氨酸、缬氨酸、亮氨酸或异亮氨酸；并且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基特别是缬氨酸、亮氨酸、异亮氨酸或苏氨酸。在一些实施方案中，酮还原酶多肽与 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0118] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还包括一个或多个或至少所有的下述另外特征：对应于 X94 的残基是半胱氨酸、脂肪族的或非极性的残基，特别是半胱氨酸或甘氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是苏氨酸、半胱氨酸、缬氨酸、亮氨酸或异亮氨酸；并且对应于 X147 的残基是脂肪族的、脂肪族的或非极性的残基，特别是亮氨酸。在一些实施方案中，酮还原酶多肽与 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%相同的氨基酸序列，条件是酮还原酶具有至

少含前述特征的氨基酸序列。

[0119] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还包括一个或多个或至少所有的下述另外特征：对应于 X94 的残基是半胱氨酸、脂肪族的或非极性的残基，特别是半胱氨酸或甘氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是苏氨酸、半胱氨酸、缬氨酸、亮氨酸或异亮氨酸；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基，特别是亮氨酸；并且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是缬氨酸、亮氨酸、异亮氨酸或苏氨酸。在一些实施方案中，酮还原酶多肽与 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具 有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0120] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还包括一个或多个或至少所有的下述另外特征：对应于 X94 的残基是半胱氨酸、脂肪族的或非极性的残基，特别是半胱氨酸或甘氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是苏氨酸、半胱氨酸、缬氨酸、亮氨酸或异亮氨酸；对应于 X194 的残基是受限的、极性的、脂肪族的或非极性的残基，特别是亮氨酸、丝氨酸或谷氨酰胺；并且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是缬氨酸。在一些实施方案中，酮还原酶 多肽与 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具 有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0121] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还包括一个或多个或至少所有的下述另外特征：对应于 X94 的残基是半胱氨酸、脂肪族的或非极性的残基，特别是半胱氨酸或甘氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是苏氨酸、半胱氨酸、缬氨酸、亮氨酸或异亮氨酸；对应于 X147 的残基是芳香族的、脂肪

族的或非极性的残基，特别是亮氨酸；对应于 X194 的残基是受限的、极性的、脂肪族的或非极性的残基，特别是亮氨酸、丝氨酸或谷氨酰胺；并且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是缬氨酸、亮氨酸、异亮氨酸或苏氨酸。在一些实施方案中，酮还原酶多肽与 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、 98％或 99％相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0122] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还包括一个或多个或至少所有的下述另外特征：对应于 X93 的残基是极性的、脂肪族的或非极性的残基，特别是丝氨酸；对应于 X94 的残基是半胱氨酸、脂肪族的或非极性的残基，特别是半胱氨酸或甘氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是苏氨酸或亮氨酸；并且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是苏氨酸、异亮氨酸、亮氨酸或缬氨酸。在一些实施方案中，酮还原酶多肽与 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0123] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还包括一个或多个或至少所有的下述另外特征：对应于 X93 的残基是极性的、脂肪族的或非极性的残基，特别是丝氨酸；对应于 X94 的残基是半胱氨酸、脂肪族的或非极性的残基，特别是半胱氨酸或甘氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的 残基，特别是半胱氨酸、苏氨酸、缬氨酸、亮氨酸或异亮氨酸；对应于 X100 的残基是碱性的、酸性的、极性的、脂肪族的或非极性的残基，特别是赖氨酸；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基，特别是亮氨酸；对应于 X152 的残基是极性的、芳香族的、脂肪族的或非极性的残基，特别是丝氨酸；对应于 X194 的残基是受限的、极性的、脂肪族的或非极性的残基，特别是亮氨酸、丝氨酸或谷氨酰胺；对应于 X196 的残基是非极性的或脂肪族的残基，特别是亮氨酸；并且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是缬氨

酸、亮氨酸、异亮氨酸或苏氨酸。在一些实施方案中，酮还原酶多肽与 SEQ IDNO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0124] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还包括一个或多个或至少所有的下述另外特征：对应于 X7 的残基是受限的或非极性的残基，特别是组氨酸；对应于 X66 的残基是酸性的残基，特别是谷氨酸；对应于 X93 的残基是极性的、脂肪族的或非极性的残基，特别是丝氨酸或甘氨酸；对应于 X94 的残基是半胱氨酸、非极性的或脂肪族的残基，特别是半胱氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是异亮氨酸、亮氨酸或苏氨酸；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基，特别是亮氨酸；对应于 X194 的残基是受限的、极性的、脂肪族的或非极性 的残基，特别是亮氨酸、丝氨酸或谷氨酰胺；对应于 X196 的残基是非极性的或脂肪族的残基，特别是亮氨酸；对应于 X198 的残基是酸性的或极性的残基，特别是谷氨酸；并且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是苏氨酸、缬氨酸、亮氨酸或异亮氨酸。在一些实施方案中，酮还原酶多肽与 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0125] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还包括一个或多个或至少所有的下述另外特征：对应于 X7 的残基是受限的或非极性的残基，特别是组氨酸；对应于 X17 的残基是非极性的或脂肪族的残基，特别是甲硫氨酸；对应于 X27 的残基是芳香族的残基，特别是酪氨酸；对应于 X64 的残基是极性的、非极性的或脂肪族的残基，特别是缬氨酸；对应于 X66 的残基是酸性的残基，特别是谷氨酸；对应于 X80 的残基是脂肪族的、非极性的或极性的残基，特别是苏氨酸；对应于 X93 的残基是极性的、脂肪族的或非极性的残基，特别是丝氨酸或甘氨酸；对应于 X94 的残基是半胱氨酸、非极性的或脂肪族的残基，特

别是半胱氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是异亮氨酸、亮氨酸或苏氨酸；对应于 X100 的残基是碱性的、酸性的、极性的、脂肪族的、非极性的残基或受限的残基，特别是赖氨酸；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残 基，特别是亮氨酸；对应于 X194 的残基是受限的、极性的、脂肪族的或非极性的残基，特别是亮氨酸、丝氨酸或谷氨酰胺；对应于 X196 的残基是非极性的或脂肪族的残基，特别是亮氨酸；对应于 X198 的残基是酸性的或极性的残基，特别是谷氨酸；对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是苏氨酸、缬氨酸、亮氨酸或异亮氨酸；并且对应于 X216 的残基是受限的或碱性的残基，特别是精氨酸。在一些实施方案中，酮还原酶多肽与 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ ID NO：2、4 或 114 的参考序列至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0126] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于 X7 的残基是受限的或非极性的残基，特别是甘氨酸或组氨酸。在一些实施方案中，酮还原酶多肽与 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ ID NO：2、4 或 114 的参考序列至少 85％、86％、87％、88％、89％、 90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0127] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于 X17 的残基是非极性的或脂肪族的残基，特别是甲硫氨酸。在一些实施方案中，酮还原酶多肽与 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具 有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ ID NO：2、4 或 114 的参考序列至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的氨基酸

序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0128] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于 X25 的残基是极性的或酸性的残基，特别是苏氨酸。在一些实施方案中，酮还原酶多肽与 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0129] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于 X27 的残基是芳香族的残基，特别是酪氨酸。在一些实施方案中，酮还原酶多肽与 SEQ IDNO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0130] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于 X53 的残基是极性的、非极性的或酸性的残基，特别是天冬氨酸。在一些实施方案中，酮还原酶多肽与 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异 包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ ID NO：2、4或114的参考序列至少85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0131] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征

或对应于X190、X195和X202的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于X64的残基是非极性的或脂肪族的残基，特别是缬氨酸。在一些实施方案中，酮还原酶多肽与SEQ ID NO:2、4或114中的参考序列相比在其他氨基酸残基位置上可另外具有1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35或约1-40个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35或约40个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ ID NO:2、4或114的参考序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0132] 在一些实施方案中，包括基于SEQ ID NO:111、112或139或其区域（例如残基90-211）中的序列式的氨基酸序列，具有本文所述对应于X190和X202的残基的特定特征或对应于X190、X195和X202的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于X80的残基是脂肪族的、非极性的或极性的残基，特别是缬氨酸或苏氨酸。在一些实施方案中，酮还原酶多肽与SEQ ID NO:2、4或114中的参考序列相比在其他氨基酸残基位置上可另外具有1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35或约1-40个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、 18、20、22、24、26、30、35或约40个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ ID NO:2、4或114的参考序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0133] 在一些实施方案中，包括基于SEQ ID NO:111、112或139或其区域（例如残基90-211）中的序列式的氨基酸序列，具有本文所述对应于X190和X202的残基的特定特征或对应于X190、X195和X202的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于X93的残基是极性的、脂肪族的或非极性的残基，特别是丝氨酸。在一些实施方案中，酮还原酶多肽与SEQ ID NO:2、4或114中的参考序列相比在其他氨基酸残基位置上可另外具有1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35或约1-40个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35或约40个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ ID NO:2、4或114的参考序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0134] 在一些实施方案中，包括基于SEQ ID NO:111、112或139或其区域（例如残基90-211）中的序列式的氨基酸序列，具有本文所述对应于X190和X202的残基的特定特征或对应于X190、X195和X202的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于X94的残基是半胱氨酸、脂肪族的或非极性的残基，特别是甘氨酸或半胱氨酸。在一些实施方案中，酮还原酶多肽与SEQ ID NO:2、4或114中的参考序列相比在其他氨

基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是 在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ ID NO:2、4或114的参考序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0135] 在一些实施方案中，包括基于 SEQ ID NO:111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征 :对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是缬氨酸、亮氨酸、异亮氨酸、苏氨酸或半胱氨酸。在一些实施方案中，酮还原酶多肽与 SEQ IDNO:2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或 约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ ID NO:2、4或114的参考序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0136] 在一些实施方案中，包括基于 SEQ ID NO:111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征 :对应于 X100 的残基是碱性的、酸性的、极性的、脂肪族的或非极性的残基，特别是赖氨酸。在一些实施方案中，酮还原酶多肽与 SEQ ID NO:2、4 或 114 中的参考序列 相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ ID NO:2、4或114的参考序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0137] 在一些实施方案中，包括基于 SEQ ID NO:111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征 :对应于 X108 的残基是碱性的或受限的残基，特别是组氨酸。在一些实施方案中，酮还原酶多肽与 SEQ ID NO:2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其

他氨基酸残基上的1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35或约40个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ IDNO:2、4或114的参考序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0138] 在一些实施方案中，包括基于SEQ ID NO:111、112或139或其区域（例如残基90-211）中的序列式的氨基酸序列，具有本文所述对应于X190和X202的残基的特定特征或对应于X190、X195和X202的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征:对应于X147的残基是 芳香族的、脂肪族的或非极性的残基，特别是亮氨酸。在一些实施方案中，酮还原酶多肽与SEQ ID NO:2、4或114中的参考序列相比在其他氨基酸残基位置上可另外具有1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35或约1-40个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35或约40个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ ID NO:2、4或114的参考序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0139] 在一些实施方案中，包括基于SEQ ID NO:111、112或139或其区域（例如残基90-211）中的序列式的氨基酸序列，具有本文所述对应于X190和X202的残基的特定特征或对应于X190、X195和X202的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征:对应于X151的残基是受限的、非极性的或脂肪族的残基，特别是丙氨酸。在一些实施方案中，酮还原酶多肽与SEQ ID NO:2、4或114中的参考序列相比在其他氨基酸残基位置上可另外具有1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35或约1-40个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35或约40个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ ID NO:2、4或114的参考序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0140] 在一些实施方案中，包括基于SEQ ID NO:111、112或139或其区域（例如残基90-211）中的序列式的氨基酸序列，具有本文所述对应于X190和X202的残基的特定特征或对应于X190、X195和X202的残基的特定特 征的改进的酮还原酶，还至少包括下述另外特征:对应于X165的残基是极性的、非极性的或脂肪族的残基，特别是天冬酰胺。在一些实施方案中，酮还原酶多肽与SEQ ID NO:2、4或114中的参考序列相比在其他氨基酸残基位置上可另外具有1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35或约1-40个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35或约40个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ ID NO:2、4或114的参考序列至少

85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0141] 在一些实施方案中，包括基于SEQ ID NO:111、112或139或其区域（例如残基90-211）中的序列式的氨基酸序列，具有本文所述对应于X190和X202的残基的特定特征或对应于X190、X195和X202的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征:对应于X173的残基是酸性的、非极性的或脂肪族的残基，特别是缬氨酸。在一些实施方案中，酮还原酶多肽与SEQ ID NO:2、4或114中的参考序列相比在其他氨基酸残基位置上可另外具有1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35或约1-40个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35或约40个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ ID NO:2、4或114的参考序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0142] 在一些实施方案中，包括基于SEQ ID NO:111、112或139或其区域（例如残基90-211）中的序列式的氨基酸序列，具有本文所述对应于X190和X202的残基的特定特征或对应于X190、X195和X202的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征:对应于X194的残基是受限的、极性的、脂肪族的或非极性的残基，特别是丝氨酸、谷氨酰胺或亮氨酸。在一些实施方案中，酮还原酶多肽与SEQ ID NO:2、4或114中的参考序列相比在其他氨基酸残基位置上可另外具有1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35或约1-40个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35或约40个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ ID NO:2、4或114的参考序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0143] 在一些实施方案中，包括基于SEQ ID NO:111、112或139或其区域（例如残基90-211）中的序列式的氨基酸序列，具有本文所述对应于X190和X202的残基的特定特征或对应于X190、X195和X202的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征:对应于X196的残基是非极性的或脂肪族的残基，特别是亮氨酸。在一些实施方案中，酮还原酶多肽与SEQ ID NO:2、4或114中的参考序列相比在其他氨基酸残基位置上可另外具有1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35或约1-40个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35或约40个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ ID NO:2、4或114的参考序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0144] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于 X198 的残基是酸性的或极性的残基，特别是谷氨酸或谷氨酰胺。在一些实施方案中，酮还原酶多肽与 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98% 或 99% 相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0145] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于 X210 的残基是极性的或碱性的残基，特别是精氨酸。在一些实施方案中，酮还原酶多肽与 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ IDNO：2、4 或 114 的参考序列至少 85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98% 或 99% 相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0146] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是缬氨酸、亮氨酸、异亮氨酸或苏氨酸。在一些实施方案中，酮还原酶多肽与 SEQ IDNO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98% 或 99% 相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0147] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征

或对应于X190、X195和X202的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于X212的残基是极性的、脂肪族的或非极性的残基，特别是缬氨酸或丝氨酸。在一些实施方案中，酮还原酶多肽与SEQ ID NO:2、4或114中的参考序列相比在其他氨基酸残基位置上可另外具有1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35或约1-40个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35或约40个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ ID NO:2、4或114的参考序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99% 相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0148] 在一些实施方案中，包括基于SEQ ID NO:111、112或139或其区域（例如残基90-211）中的序列式的氨基酸序列，具有本文所述对应于X190和X202的残基的特定特征或对应于X190、X195和X202的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于X216的残基是受限的或碱性的残基，特别是精氨酸。在一些实施方案中，酮还原酶多肽与SEQ ID NO:2、4或114中的参考序列相比在其他氨基酸残基位置上可另外具有1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35或约1-40个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35或约40个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ IDNO:2、4或114的参考序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0149] 在一些实施方案中，包括基于SEQ ID NO:111、112或139或其区域（例如残基90-211）中的序列式的氨基酸序列，具有本文所述对应于X190和X202的残基的特定特征或对应于X190、X195和X202的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于X223的残基是非极性的或脂肪族的残基，特别是缬氨酸。在一些实施方案中，酮还原酶多肽与SEQ ID NO:2、4或114中的参考序列相比在其他氨基酸残基位置上可另外具有1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35或约1-40个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35或约40个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ ID NO:2、4或114的参考序列至少85%、86%、87%、88%、89%、 90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0150] 在一些实施方案中，包括基于SEQ ID NO:111、112或139或其区域（例如残基90-211）中的序列式的氨基酸序列，具有本文所述对应于X190和X202的残基的特定特征或对应于X190、X195和X202的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于X226的残基是非极性的或极性的残基，特别是苏氨酸。在一些实施方案中，酮还原酶多肽与SEQ ID NO:2、4或114中的参考序列相比在其他氨基酸残基位置上可另外具有

1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的氨基酸序列，条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0151] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于 X7 的残基是受限的或非极性的残基，特别是组氨酸；对应于 X94 的残基是半胱氨酸、非极性的或脂肪族的残基，特别是半胱氨酸或甘氨酸；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基，特别是苯丙氨酸或亮氨酸；并且对应于 X196 的残基是非极性的或脂肪族的残基，特别是亮氨酸。在一些实施方案中，酮还原酶多肽与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异（例如 SEQ ID NO：48）。 在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的氨基酸序列（例如 SEQ ID NO：48），条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0152] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于 X7 的残基是受限的或非极性的残基，特别是组氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是天冬酰胺、丝氨酸、缬氨酸、亮氨酸、异亮氨酸或苏氨酸；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基，特别是苯丙氨酸或亮氨酸；并且对应于 X196 的残基是非极性的或脂肪族的残基，特别是亮氨酸。在一些实施方案中，酮还原酶多肽与具有前述特征的 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异（例如 SEQ ID NO：42、44、46 或 90）。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的氨基酸序列（例如 SEQ ID NO：42、44、46 或 90），条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0153] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基

90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于 X7 的残基是受限的或非极性的残基，特别是组氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是天冬酰胺、丝氨酸、缬氨酸、亮氨酸、异亮氨酸或苏氨酸；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基，特别是苯丙氨酸或亮氨酸；对应于 X196 的残基是非极性的或脂肪族的残基，特别是亮氨酸；并且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是缬氨酸、亮氨酸、异亮氨酸或苏氨酸。在一些实施方案中，酮还原酶多肽与具有前述特征的 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异（例如 SEQ ID NO：88、94 或 102）。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的氨基酸序列（例如 SEQID NO：88、94 或 102），条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0154] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于 X7 的残基是受限的或非极性的残基，特别是组氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是天冬酰胺、丝氨酸、缬氨酸、亮氨酸、异亮氨酸或苏氨酸；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基，特别是苯丙氨酸或亮氨酸；对应于 X194 的残基是受限的、极性的、脂肪族的或非极性的残基，特别是亮氨酸、谷氨酰胺或丝氨酸；对应于 X196 的残基是非极性的或脂肪族的残基，特别是亮氨酸；并且对 应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是缬氨酸、亮氨酸、异亮氨酸或苏氨酸。在一些实施方案中，酮还原酶多肽与具有前述特征的 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异（例如 SEQ ID NO：98）。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ IDNO：2、4 或 114 的参考序列至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的氨基酸序列（例如 SEQ ID NO：98），条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0155] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于 X7 的残基是受限的或非极性的残基，特别是组氨酸；对应于 X93 的残基是极性的、脂肪族的或非极性的残基，特别是丝氨酸；对应于 X94 的残基是半胱氨酸、非极性

的或脂肪族的残基，特别是半胱氨酸或甘氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是天冬酰胺、丝氨酸、缬氨酸、亮氨酸、异亮氨酸或苏氨酸；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基，特别是苯丙氨酸或亮氨酸；对应于 X196 的残基是非极性的或脂肪族的残基，特别是亮氨酸；并且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是缬氨酸、亮氨酸、异亮氨酸或苏氨酸。在一些实施方案中，酮还原酶多肽与具有前述特征的 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异（例如 SEQ ID NO：84）。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、 14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%相同的氨基酸序列（例如 SEQID NO：84），条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0156] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于 X7 的残基是受限的或非极性的残基，特别是组氨酸；对应于 X66 的残基是酸性的残基，特别是谷氨酸；对应于 X93 的残基是极性的、脂肪族的或非极性的残基，特别是丝氨酸；对应于 X94 的残基是半胱氨酸、非极性的或脂肪族的残基，特别是半胱氨酸或甘氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是天冬酰胺、丝氨酸、缬氨酸、亮氨酸、异亮氨酸或苏氨酸；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基，特别是苯丙氨酸或亮氨酸；对应于 X196 的残基是非极性的或脂肪族的残基，特别是亮氨酸；并且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是缬氨酸、亮氨酸、异亮氨酸或苏氨酸。在一些实施方案中，酮还原酶多肽与具有前述特征的 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异（例如 SEQ IDNO：64）。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%相同的氨基酸序列（例如 SEQ ID NO：64），条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0157] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于 X7 的残基是受限的或非极性的残基，特别是组氨酸；对应于 X17 的残基是非极性的或脂肪族的残基，特别是甲硫氨酸；对应于 X27 的残基是芳香族的残基，特别是酪氨酸；对应于 X64 的残基是极性的、非极性的或脂肪族的残基，特别是缬氨酸；对应于 X80 的

残基是脂肪族的、非极性的或极性的残基，特别是苏氨酸或缬氨酸；对应于 X93 的残基是极性的、脂肪族的或非极性的残基，特别是丝氨酸；对应于 X94 的残基是半胱氨酸、非极性的或脂肪族的残基，特别是半胱氨酸或甘氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是天冬酰胺、丝氨酸、缬氨酸、亮氨酸、异亮氨酸或苏氨酸；对应于 X100 的残基是碱性的、酸性的、极性的或脂肪族的，特别是赖氨酸；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基，特别是苯丙氨酸或亮氨酸；对应于 X196 的残基是非极性的或脂肪族的残基，特别是亮氨酸；并且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是缬氨酸、亮氨酸、异亮氨酸或苏氨酸。在一些实施方案中，酮还原酶多肽与具有前述特征的 SEQID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异（例如 SEQ ID NO：22 或 24）。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%相同的氨基酸序列（例如 SEQ ID NO：22 或 24），条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0158] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于 X7 的残基是受限的或非极性的残基，特别是组氨酸；对应于 X66 的残基是酸性的残基，特别是谷氨酸；对应于 X93 的残基是极性的、脂肪族的或非极性的残基，特别是丝氨酸；对应于 X94 的残基是半胱氨酸、非极性的或脂肪族的残基，特别是半胱氨酸或甘氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是天冬酰胺、丝氨酸、缬氨酸、亮氨酸、异亮氨酸或苏氨酸；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基，特别是苯丙氨酸或亮氨酸；对应于 X194 的残基是受限的、极性的或非极性的残基，特别是天冬酰胺或亮氨酸；对应于 X196 的残基是非极性的或脂肪族的残基，特别是亮氨酸；对应于 X198 的残基是酸性的或极性的残基，特别是谷氨酸或天冬酰胺；对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是缬氨酸、亮氨酸、异亮氨酸或苏氨酸；并且对应于 X216 的残基是受限的或碱性的残基，特别是精氨酸。在一些实施方案中，酮还原酶多肽与具有前述特征的 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异（例如 SEQ IDNO：58）。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%相同的氨基酸序列（例如 SEQ ID NO：58），条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0159] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基

90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于 X7 的残基是受限的或非极性的残基，特别是组氨酸；对应于 X66 的残基是酸性的残基，特别是谷氨酸；对应于 X93 的残基是极性的、脂肪族的或非极性的残基，特别是丝氨酸；对应于 X94 的残基是半胱氨酸、脂肪族的或非极性的残基，特别是半胱氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是半胱氨酸、苏氨酸、缬氨酸、亮氨酸或异亮氨酸；对应于 X147 的残基是芳香族的或脂肪族的残基，特别是亮氨酸；对应于 X194 的残基是受限的、极性的、脂肪族的或非极性的残基，特别是谷氨酰胺；对应于 X196 的残基是非极性的或脂肪族的残基，特别是亮氨酸；对应于 X198 的残基是酸性的或极性的残基，特别是谷氨酸；并且对应于 X211 的残基是碱性的、脂肪族的或非极性的残基，特别是缬氨酸缬氨酸、亮氨酸、异亮氨酸或苏氨酸。在一些实施方案中，酮还原酶多肽与具有前述特征的 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异（例如 SEQ ID NO：40）。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的氨基酸序列（例如 SEQ ID NO：40），条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0160] 在一些实施方案中，包括基于 SEQ ID NO：111、112 或 139 或其区域（例如残基 90-211）中的序列式的氨基酸序列，具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于 X7 的残基是受限的或非极性的残基，特别是组氨酸；对应于 X17 的残基是非极性的或脂肪族的残基，特别是甲硫氨酸；对应于 X27 的残基是芳香族的残基，特别是酪氨酸；对应于 X64 的残基是非极性的或脂肪族的残基，特别是缬氨酸；对应于 X80 的残基是脂肪族的、非极性的或极性的残基，特别是苏氨酸或缬氨酸；对应于 X93 的残基是极性的、脂肪族的或非极性的残基，特别是丝氨酸；对应于 X94 的残基是半胱氨酸、非极性的或脂肪族的残基，特别 是半胱氨酸或甘氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是天冬酰胺、丝氨酸、缬氨酸、亮氨酸、异亮氨酸或苏氨酸；对应于 X100 的残基是碱性的、酸性的、极性的、非极性的或脂肪族的残基，特别是赖氨酸；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基，特别是苯丙氨酸或亮氨酸；对应于 X194 的残基是受限的、极性的或非极性的残基，特别是天冬酰胺或亮氨酸；对应于 X196 的残基是非极性的或脂肪族的残基，特别是亮氨酸；对应于 X198 的残基是酸性的或极性的残基，特别是谷氨酸或天冬酰胺；对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是缬氨酸、亮氨酸、异亮氨酸或苏氨酸；并且对应于 X216 的残基是受限的或碱性的残基，特别是精氨酸。在一些实施方案中，酮还原酶多肽与具有前述特征的 SEQ ID NO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异（例如 SEQ ID NO：14 或 16）。在一些实施方案中，差异的数目可以是

在其他氨基酸残基上的1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35或约40个残基差异。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ ID NO:2、4或114的参考序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列（例如SEQ ID NO:14或16），条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0161] 在一些实施方案中，包括基于SEQ ID NO:111、112或139或其区域（例如残基90-211）中的序列式的氨基酸序列，具有本文所述对应于X190和X202的残基的特定特征或对应于X190、X195和X202的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于X7的残基是受限的或非极性的残基，特别是组氨酸；对应于X17的残基是非极性的或脂肪族的残基，特别是甲硫氨酸；对应于X25的残基是极性的或酸性的残基，特别是苏氨酸；对应于X27的残基是芳香族的残基，特别是酪氨酸；对应于X53的残基是极性的、非极性的或酸性的残基，特别是天冬氨酸；对应于X64的残基是非极性的或脂肪族的残基，特别是缬氨酸；对应于X80 的残基是脂肪族的、非极性的或极性的残基，特别是苏氨酸或缬氨酸；对应于X93的残基是极性的、脂肪族的或非极性的残基，特别是丝氨酸；对应于X94的残基是半胱氨酸、非极性的或脂肪族的残基，特别是半胱氨酸或甘氨酸；对应于X96的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是天冬酰胺、丝氨酸、缬氨酸、亮氨酸、异亮氨酸或苏氨酸；对应于X100的残基是碱性的、酸性的、极性的、非极性的或脂肪族的残基，特别是赖氨酸；对应于X147的残基是芳香族的、脂肪族的或非极性的残基，特别是苯丙氨酸或亮氨酸；对应于X194的残基是受限的、极性的或非极性的残基，特别是天冬酰胺或亮氨酸；对应于X196的残基是非极性的或脂肪族的残基，特别是亮氨酸；对应于X198的残基是酸性的或极性的残基，特别是谷氨酸或天冬酰胺；对应于X211的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是缬氨酸、亮氨酸、异亮氨酸或苏氨酸；并且对应于X216的残基是受限的或碱性的残基，特别是精氨酸。在一些实施方案中，酮还原酶多肽与具有前述特征的SEQ IDNO:2、4或114中的参考序列相比在其他氨基酸残基位置上可另外具有1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35或约1-40个残基差异（例如SEQ ID NO:6）。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35或约40个残基差异。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于SEQ ID NO:2、4或114的参考序列至少85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列（例如SEQ ID NO:6），条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0162] 在一些实施方案中，包括基于SEQ ID NO:111、112或139或其区域（例如残基90-211）中的序列式的氨基酸序列，具有本文所述对应于X190和X202的残基的特定特征或对应于X190、X195和X202的残基的特定特征的改进的酮还原酶，还至少包括下述另外特征：对应于X7的残基是受限的或非极性的残基，特别是组氨酸；对应于X17的残基是非极性的或脂肪族的残基，特别是甲硫氨酸；对应于X25的残基是极性的或酸性的残基， 特别是苏氨酸；对应于X27的残基是芳香族的残基，特别是酪氨酸；对应于X53的残基是极性的、非极性的或酸性的残基，特别是天冬氨酸；对应于X64的残基是非极性的或脂肪族的残基，特别是缬氨酸；对应于X76的残基是脂肪族的、非极性的或极性的残基，特别是丙氨

酸；对应于 X80 的残基是脂肪族的、非极性的或极性的残基，特别是苏氨酸或缬氨酸；对应于 X93 的残基是极性的、脂肪族的或非极性的残基，特别是丝氨酸；对应于 X94 的残基是半胱氨酸、非极性的或脂肪族的残基，特别是半胱氨酸或甘氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是天冬酰胺、丝氨酸、缬氨酸、亮氨酸、异亮氨酸或苏氨酸；对应于 X100 的残基是碱性的、酸性的、极性的、非极性的或脂肪族的残基，特别是赖氨酸；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基，特别是苯丙氨酸或亮氨酸；对应于 X194 的残基是受限的、极性的或非极性的残基，特别是天冬酰胺或亮氨酸；对应于 X196 的残基是非极性的或脂肪族的残基，特别是亮氨酸；对应于 X198 的残基是酸性的或极性的残基，特别是谷氨酸或天冬酰胺；对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是缬氨酸、亮氨酸、异亮氨酸或苏氨酸；并且对应于 X216 的残基是受限的或碱性的残基，特别是精氨酸。在一些实施方案中，酮还原酶多肽与具有前述特征的 SEQ IDNO：2、4 或 114 中的参考序列相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-26、1-30、1-35 或约 1-40 个残基差异（例如 SEQ ID NO：8）。在一些实施方案中，差异的数目可以是在其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个残基差异。在一些实施方案中，酮还原酶多肽包括与具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列至少 85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%相同的氨基酸序列（例如 SEQ ID NO：8），条件是酮还原酶具有至少含前述特征的氨基酸序列。

[0163] 在一些实施方案中，本公开的改进的酮还原酶包括具有对应于 SEQ IDNO：111、112 或 139 中的序列式的残基 90-211 的区域或结构域的氨基酸序 列，其中对应于 X190 的残基是受限的残基，而对应于 X202 的残基是芳香族的残基。在一些实施方案中，对应于残基 90-211 的结构域或区域包括这样的氨基酸序列，其中对应于 X190 的残基是脯氨酸，而对应于 X202 的残基是色氨酸。在一些实施方案中，对应于残基 90-211 的区域或结构域与基于 SEQ ID NO：2、4 或 114 的参考序列的相应结构域相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18 或 1-20 个残基差异。在一些实施方案中，差异的数目可以是在所述结构域中其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18 或约 20 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括具有与在对应于 X190 和 X202 的残基上具有前述特征的基于 SEQ IDNO：2、4 或 114 的参考序列的相应结构域至少 85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%相同的结构域或区域的氨基酸序列，条件是酮还原酶结构域或区域包括至少含前述特征的氨基酸序列。

[0164] 在一些实施方案中，本公开的改进的酮还原酶包括具有对应于 SEQ IDNO：111、112 或 139 中的序列式的残基 90-211 的区域或结构域的氨基酸序列，其中对应于 X190 的残基是受限的残基；对应于 X195 的残基是碱性的、酸性的、非极性的或极性的残基；而对应于 X202 的残基是芳香族的残基。在一些实施方案中，对应于残基 90-211 的结构域或区域包括这样的氨基酸序列，其中对应于 X190 的残基是脯氨酸；对应于 X195 的残基是脯氨酸、精氨酸、赖氨酸、天冬氨酸或天冬酰胺，特别是精氨酸；而对应于 X202 的残基是色氨酸。在一些实施方案中，对应于残基 90-211 的区域或结构域与基于 SEQ ID NO：2、4 或 114 的参考序列

的相应结构域相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18 或 1-20 个残基差异。在一些实施方案中，差异的数目可以是在所述结构域中其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18 或约 20 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶多肽包括具有与在对应于 X190、X195 和 X202 的残基上具有前述特征的基 于 SEQ ID NO：2、4 或 114 的参考序列的相应结构域至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的结构域或区域的氨基酸序列，条件是酮还原酶结构域或区域包括至少含前述特征的氨基酸序列。

[0165] 在一些实施方案中，具有对应于 SEQ ID NO：111、112 或 139 中的序列式的残基 90-211 的结构域或区域并且具有对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的酮还原酶多肽，还可在所述区域或结构域中包括选自如下的一种或多种特征：对应于 X93 的残基是极性的、脂肪族的或非极性的残基；对应于 X94 的残基是半胱氨酸、脂肪族的或非极性的残基；对应于 X95 的残基是脂肪族的、非极性的或酸性的残基；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基；对应于 X97 的残基是碱性的、受限的或极性的残基；对应于 X100 的残基是碱性的、酸性的、极性的、脂肪族的或非极性的残基；对应于 X108 的残基是碱性的或受限的残基；对应于 X111 的残基是非极性的或脂肪族的残基；对应于 X139 的残基是非极性的或脂肪族的残基；对应于 X145 的残基是酸性的、脂肪族的或非极性的残基；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基；对应于 X148 的残基是非极性的或脂肪族的残基；对应于 X151 的残基是受限的、非极性的或脂肪族的残基；对应于 X152 的残基是极性的、芳香族的、脂肪族的或非极性的残基；对应于 X163 的残基是非极性的或脂肪族的残基；对应于 X165 的残基是极性的、非极性的或脂肪族的残基；对应于 X173 的残基是酸性的、非极性的或脂肪族的残基；对应于 X179 的残基是芳香族的、碱性的、脂肪族的或非极性的残基；对应于 X194 的残基是受限的、极性的、脂肪族的或非极性的残基；对应于 X196 的残基是非极性的或脂肪族的残基；对应于 X197 的残基是酸性的残基；对应于 X198 的残基是酸性的或极性的残基；对应于 X210 的残基是极性的或碱性的残基；而且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基。在一些实施方案中，对应于残基 90-211 的区域或结构域与基于 SEQ ID NO：2、4 或 114 的参考序列的相应结构域相比在其他未被上述 X 所指定的氨基酸残基上可另外具有约 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、 1-15、1-16、1-18 或 1-20 个残基差异。在一些实施方案中，差异的数目可以是在所述结构域中其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18 或约 20 个残基差异。在一些实施方案中，所述差异包括保守突变。

[0166] 在一些实施方案中，如上所述具有包含与 SEQ ID NO：111、112 或 139 中的序列式的残基 90-211 相对应的氨基酸序列的结构域或区域的酮还原酶多肽，与 SEQ ID NO：2、4 或 114 的相应结构域的氨基酸序列相比在所述结构域或区域中可具有一个或多个保守突变。这种保守突变的实例包括氨基酸替换例如但不限于：用另一种极性残基例如丝氨酸替换对应于 X93 的残基异亮氨酸 (I)；用另一种非极性残基例如甘氨酸替换对应于 X94 的残基丙氨酸 (A)；用另一种非极性或脂肪族残基例如亮氨酸替换对应于 X95 的残基缬氨酸 (V)；以及用另一种极性残基例如苏氨酸替换对应于 X96 的残基丝氨酸 (S)；用另一种非极性或脂

肪族残基例如异亮氨酸替换对应于X148的残基缬氨酸(V);用另一种极性残基例如丝氨酸替换对应于X152苏氨酸;用另一种非极性或脂肪族残基例如异亮氨酸替换对应于X163的残基缬氨酸(V);用另一种芳香族残基例如苯丙氨酸替换对应于X179的残基酪氨酸(Y);以及用另一种脂肪族残基例如亮氨酸替换对应于X196的残基缬氨酸(V);用另一种酸性残基例如谷氨酸替换对应于X198的残基天冬氨酸(D)。

[0167] 在一些实施方案中,具有对应于SEQ ID NO:111、112或139中的序列式的残基90-211的结构域或区域并且具有对应于X190和X202的残基的特定特征或对应于X190、X195和X202的残基的特定特征的酮还原酶多肽,还可在所述区域或结构域中包括选自如下的一种或多种特征:对应于X93的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是异亮氨酸、丝氨酸、丙氨酸或甘氨酸;对应于X94的残基是半胱氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是半胱氨酸、甘氨酸或丙氨酸;对应于X95的残基是天冬氨酸、谷氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是缬氨酸、亮氨酸或谷氨酸;对应于X96的 残基是半胱氨酸、丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是天冬酰胺、丝氨酸、缬氨酸、亮氨酸、异亮氨酸或苏氨酸;对应于X97的残基是赖氨酸、精氨酸、组氨酸、脯氨酸、丝氨酸、苏氨酸、谷氨酰胺或天冬酰胺,特别是赖氨酸、组氨酸或丝氨酸;对应于X100的残基是赖氨酸、精氨酸、丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、甘氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是谷氨酸、赖氨酸、谷氨酰胺或丙氨酸;对应于X108的残基是精氨酸、赖氨酸、组氨酸或脯氨酸,特别是精氨酸或组氨酸;对应于X111的残基是甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是甲硫氨酸;对应于X139的残基是甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是亮氨酸;对应于X145的残基是天冬氨酸、谷氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是谷氨酸或亮氨酸;对应于X147的残基是酪氨酸、苯丙氨酸、色氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是苯丙氨酸或亮氨酸;对应于X151的残基是脯氨酸、组氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是丙氨酸;对应于X152的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、酪氨酸、苯丙氨酸、色氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是苏氨酸、色氨酸、丙氨酸或丝氨酸;对应于X163的残基是甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是异亮氨酸;对应于X165的残基是丝氨酸、苏氨酸、谷氨酰胺天冬酰胺、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是天冬酰胺;对应于X173的残基是天冬氨酸、谷氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸,特别是缬氨酸;对应于X179的残基是酪氨酸、苯丙氨酸、色氨酸、精氨酸、赖氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是酪氨酸、苯丙氨酸、赖氨酸或亮氨酸;对应于X194的残基是脯氨酸、组氨酸、丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是脯氨酸、谷氨酰胺、丝氨酸或亮氨酸;对应于X196的残基是甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸,特别是亮氨酸;对应于X197的残基是天冬氨 酸或谷氨酸,特别是谷氨酸;对应于X198的残基是天冬氨酸、谷氨酸、丝氨酸、苏氨酸、谷氨酰胺或天冬酰胺,特别是天冬氨酸、谷氨酸或谷氨酰胺;对应于X210的残基是精氨酸、赖氨酸、丝氨酸、苏氨酸、谷氨酰胺或天冬酰胺,特别是苏氨酸或精氨酸;而且对应于

X211 的残基是赖氨酸、精氨酸、天冬氨酸、谷氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸、苏氨酸、谷氨酰胺或天冬酰胺，特别是缬氨酸、亮氨酸、苏氨酸或异亮氨酸。在一些实施方案中，对应于残基 90-211 的区域或结构域与基于 SEQ ID NO :2、4 或 114 的参考序列的相应结构域相比在其他未被上述 X 所指定的氨基酸残基上可另外具有约 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18 或 1-20 个残基差异。在一些实施方案中，差异的数目可以是在所述结构域中其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18 或约 20 个残基差异。在一些实施方案中，所述差异包括保守突变。

[0168] 在一些实施方案中，具有对应于 SEQ ID NO :111、112 或 139 中的序列式的残基 90-211 的结构域或区域并且具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的酮还原酶多肽，还可在所述区域或结构域中包括选自如下的一种或多种或所有特征 :对应于 X94 的残基是半胱氨酸、脂肪族的或非极性的残基，特别是半胱氨酸或甘氨酸 ;对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是苏氨酸、半胱氨酸、缬氨酸、亮氨酸或异亮氨酸 ;而且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是缬氨酸、亮氨酸、苏氨酸或异亮氨酸。在一些实施方案中，对应于残基 90-211 的区域或结构域与基于 SEQ ID NO :2、4 或 114 的参考序列的结构域相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18 或 1-20 个残基差异。在一些实施方案中，差异的数目可以是在所述结构域中其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18 或约 20 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶区域或结构域包括至少具有前述特征的氨基酸序列，而且其中所述氨基酸序列与对应于具有前述特征的基于 SEQ ID NO :2、4 或 114 的参考序列的残基 90-211 的氨基酸序列相比具有至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％的同一性。

[0169] 在一些实施方案中，具有对应于 SEQ ID NO :111、112 或 139 中的序列式的残基 90-211 的结构域或区域并且具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的酮还原酶多肽，还可在所述区域或结构域中包括选自如下的一种或多种或所有特征 :对应于 X94 的残基是半胱氨酸、脂肪族的或非极性的残基，特别是半胱氨酸或甘氨酸 ;对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是苏氨酸、半胱氨酸、缬氨酸、亮氨酸或异亮氨酸 ;对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基，特别是亮氨酸 ;而且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是缬氨酸、亮氨酸、苏氨酸或异亮氨酸。在一些实施方案中，对应于残基 90-211 的区域或结构域与基于 SEQ ID NO :2、4 或 114 的参考序列的结构域相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18 或 1-20 个残基差异。在一些实施方案中，差异的数目可以是在所述结构域中其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18 或约 20 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶区域或结构域包括至少具有前述特征的氨基酸序列，而且其中所述氨基酸序列与对应于具有前述特征的基于 SEQ ID NO :2、4 或 114 的参考序列的残基 90-211 的氨基酸序列

相比具有至少 85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%的同一性。

[0170] 在一些实施方案中，具有对应于 SEQ ID NO:111、112 或 139 中的序列式的残基 90-211 的结构域或区域并且具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的酮还原酶多肽，还可在所述区域或结构域中包括选自如下的一种或多种或所有特征：对应于 X94 的残基是半胱氨酸、脂肪族的或非极性的残基，特别 是半胱氨酸或甘氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是苏氨酸、半胱氨酸、缬氨酸、亮氨酸或异亮氨酸；对应于 X194 的残基是受限的、极性的、脂肪族的或非极性的残基，特别是丝氨酸或谷氨酰胺；而且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是缬氨酸、亮氨酸、苏氨酸或异亮氨酸。在一些实施方案中，对应于残基 90-211 的区域或结构域与基于 SEQ IDNO:2、4 或 114 的参考序列的结构域相比在其他氨基酸残基上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18 或 1-20 个残基差异。在一些实施方案中，差异的数目可以是在所述结构域中其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18 或约 20 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶区域或结构域包括至少具有前述特征的氨基酸序列，而且其中所述氨基酸序列与对应于具有前述特征的基于 SEQ ID NO:2、4 或 114 的参考序列的残基 90-211 的氨基酸序列相比具有至少 85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%的同一性。

[0171] 在一些实施方案中，具有对应于 SEQ ID NO:111、112 或 139 中的序列式的残基 90-211 的结构域或区域并且具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的酮还原酶多肽，还可在所述区域或结构域中包括选自如下的一种或多种或所有特征：对应于 X94 的残基是半胱氨酸、脂肪族的或非极性的残基，特别是半胱氨酸或甘氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是苏氨酸、半胱氨酸、缬氨酸、亮氨酸或异亮氨酸；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基，特别是亮氨酸；对应于 X194 的残基是受限的、极性的、脂肪族的或非极性的残基，特别是亮氨酸、丝氨酸或谷氨酰胺；而且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是缬氨酸、亮氨酸、异亮氨酸或苏氨酸。在一些实施方案中，对应于残基 90-211 的区域或结构域与基于 SEQ ID NO:2、4 或 114 的参考序列的结构域相比在其他氨基酸残基上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、 1-12、1-14、1-15、1-16、1-18 或 1-20 个残基差异。在一些实施方案中，差异的数目可以是在所述结构域中其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18 或约 20 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶区域或结构域包括至少具有前述特征的氨基酸序列，而且其中所述氨基酸序列与对应于具有前述特征的基于 SEQ ID NO:2、4 或 114 的参考序列的残基 90-211 的氨基酸序列相比具有至少 85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%的同一性。

[0172] 在一些实施方案中，具有对应于 SEQ ID NO:111、112 或 139 中的序列式的残基 90-211 的结构域或区域并且具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于

X190、X195 和 X202 的残基的特定特征的酮还原酶多肽，还可在所述区域或结构域中包括选自如下的一种或多种或所有特征：对应于 X93 的残基是极性的、脂肪族的或非极性的残基，特别是丝氨酸；对应于 X94 的残基是半胱氨酸、脂肪族的或非极性的残基，特别是半胱氨酸或甘氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是苏氨酸或亮氨酸；而且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是苏氨酸、异亮氨酸、亮氨酸或缬氨酸。在一些实施方案中，对应于残基 90-211 的区域或结构域与基于 SEQ ID NO：2、4 或 114 的参考序列的结构域相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18 或 1-20 个残基差异。在一些实施方案中，差异的数目可以是在所述结构域中其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18 或约 20 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶区域或结构域包括至少具有前述特征的氨基酸序列，而且其中所述氨基酸序列与对应于具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列的残基 90-211 的氨基酸序列相比具有至少 85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%的同一性。

[0173] 在一些实施方案中，具有对应于 SEQ ID NO：111、112 或 139 中的序 列式的残基 90-211 的结构域或区域并且具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的酮还原酶多肽，还可在所述区域或结构域中包括选自如下的一种或多种或所有特征：对应于 X93 的残基是极性的、脂肪族的或非极性的残基，特别是丝氨酸；对应于 X94 的残基是半胱氨酸、脂肪族的或非极性的残基，特别是半胱氨酸或甘氨酸；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是半胱氨酸、苏氨酸、缬氨酸、亮氨酸或异亮氨酸；对应于 X100 的残基是碱性的、酸性的、极性的、脂肪族的或非极性的残基，特别是赖氨酸；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基，特别是亮氨酸；对应于 X152 的残基是极性的、芳香族的、脂肪族的或非极性的残基，特别是丝氨酸；对应于 X194 的残基是受限的、极性的、脂肪族的或非极性的残基，特别是丝氨酸或谷氨酰胺；对应于 X196 的残基是非极性的或脂肪族的残基，特别是亮氨酸；而且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是缬氨酸。在一些实施方案中，差异的数目可以是在所述结构域中其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18 或约 20 个残基差异。在一些实施方案中，对应于残基 90-211 的区域或结构域与基于 SEQ ID NO：2、4 或 114 的参考序列的结构域相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18 或 1-20 个残基差异。在一些实施方案中，差异的数目可以是在所述结构域中其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18 或约 20 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶区域或结构域包括至少具有前述特征的氨基酸序列，而且其中所述氨基酸序列与对应于具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列的残基 90-211 的氨基酸序列相比具有至少 85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%的同一性。

[0174] 在一些实施方案中，具有对应于 SEQ ID NO：111、112 或 139 中的序列式的残基 90-211 的结构域或区域并且具有本文所述对应于 X190 和 X202 的残基的特定特征或对应

于 X190、X195 和 X202 的残基的特定特征的酮 还原酶多肽，还可在所述区域或结构域中包括选自如下的一种或多种或所有特征 ：对应于 X93 的残基是极性的、脂肪族的或非极性的残基，特别是丝氨酸或甘氨酸 ；对应于 X94 的残基是半胱氨酸、非极性的或脂肪族的残基，特别是半胱氨酸 ；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是异亮氨酸、亮氨酸或苏氨酸 ；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基，特别是亮氨酸 ；对应于 X194 的残基是受限的、极性的、脂肪族的或非极性的残基，特别是谷氨酰胺或丝氨酸 ；对应于 X196 的残基是非极性的或脂肪族的残基，特别是亮氨酸 ；对应于 X198 的残基是酸性的或极性的残基，特别是谷氨酸 ；而且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是苏氨酸、缬氨酸、亮氨酸或异亮氨酸。在一些实施方案中，对应于残基 90-211 的区域或结构域与基于 SEQ ID NO ：2、4 或 114 的参考序列的结构域相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18 或 1-20 个残基差异。在一些实施方案中，差异的数目可以是在所述结构域中其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18 或约 20 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶区域或结构域包括至少具有前述特征的氨基酸序列，而且其中所述氨基酸序列与对应于具有前述特征的基于 SEQ ID NO ：2、4 或 114 的参考序列的残基 90-211 的氨基酸序列相比具有至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％的同一性。

[0175] 在一些实施方案中，具有对应于 SEQ ID NO ：111、112 或 139 中的序列式的残基 90-211 的结构域或区域并且具有本文所述对应于 X190 和 X202 的残基的特定特征或对应于 X190、X195 和 X202 的残基的特定特征的酮还原酶多肽，还可在所述区域或结构域中包括选自如下的一种或多种或所有特征 ：对应于 X93 的残基是极性的、脂肪族的或非极性的残基，特别是丝氨酸或甘氨酸 ；对应于 X94 的残基是半胱氨酸、非极性的或脂肪族的残基，特别是半胱氨酸 ；对应于 X96 的残基是半胱氨酸、极性的、脂肪族的或非极性的残基，特别是异亮氨酸、亮氨酸或苏氨酸 ；对应于 X100 的残 基是 is a 碱性的、酸性的、极性的、脂肪族的、非极性的残基或受限的残基，特别是赖氨酸 ；对应于 X147 的残基是芳香族的、脂肪族的或非极性的残基，特别是亮氨酸 ；对应于 X194 的残基是受限的、极性的、脂肪族的或非极性的残基，特别是谷氨酰胺或丝氨酸 ；对应于 X196 的残基是非极性的或脂肪族的残基，特别是亮氨酸 ；对应于 X198 的残基是酸性的或极性的残基，特别是谷氨酸 ；而且对应于 X211 的残基是碱性的、脂肪族的、非极性的、酸性的或极性的残基，特别是苏氨酸、缬氨酸、亮氨酸或异亮氨酸。在一些实施方案中，对应于残基 90-211 的区域或结构域与基于 SEQ ID NO ：2、4 或 114 的参考序列的结构域相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18 或 1-20 个残基差异。在一些实施方案中，差异的数目可以是在所述结构域中其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18 或约 20 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶区域或结构域包括至少具有前述特征的氨基酸序列，而且其中所述氨基酸序列与对应于具有前述特征的基于 SEQ ID NO ：2、4 或 114 的参考序列的残基 90-211 的氨基酸序列相比具有至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％的同一性。

[0176] 在一些实施方案中，酮还原酶多肽还可包括对应于 SEQ ID NO：111、112 或 139 中的序列式的残基 1-89 的结构域或区域，其中所述结构域或区域可具有一种或多种下述特征：对应于 X7 的残基是受限的或非极性的残基；对应于 X17 的残基是非极性的或脂肪族的残基；对应于 X25 的残基是极性的或酸性的残基；对应于 X27 的残基是芳香族的残基；对应于 X41 的残基是极性的、脂肪族的或非极性的残基；对应于 X53 的残基是极性的、非极性的或酸性的残基；对应于 X57 的残基是非极性的或脂肪族的残基；对应于 X64 的残基是极性的、脂肪族的或非极性的残基；对应于 X66 的残基是酸性的残基；对应于 X68 的残基是极性的、酸性的、脂肪族的或非极性的残基；对应于 X72 的残基是碱性的残基；对应于 X76 的残基是脂肪族的、非极性的或极性的残基；对应于 X80 的残基是脂肪族的、非极性的或极性的残基。在一些实施方案中，包括对应于 SEQ ID NO：111、112 或 139 中的序列式的残基 1-89 的区域或结构域的多肽与 SEQ ID NO：2、4 或 114 的参考序列相比在未被上述 X 所指定的残基位置上可另外具有一个或多个残基差异。在一些实施方案中，所述差异可以是与 SEQ ID NO：2、4 或 114 的参考序列相比在其他未被上述 X 所界定的氨基酸残基位置上的 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15 或 1-16 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基位置上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15 或 16 个残基差异。在一些实施方案中，所述差异包括保守突变。

[0177] 在一些实施方案中，酮还原酶多肽还可包括对应于 SEQ ID NO：111、112 或 139 中的序列式的残基 1-89 的结构域或区域，其中所述结构域或区域可具有一种或多种下述特征：对应于 X7 的残基是组氨酸、脯氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是甘氨酸或组氨酸；对应于 X17 的残基是甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸，特别是甲硫氨酸；对应于 X25 的残基是谷氨酸、天冬氨酸、丝氨酸、苏氨酸、谷氨酰胺或天冬酰胺，特别是苏氨酸或天冬氨酸；对应于 X27 的残基是酪氨酸、苯丙氨酸、色氨酸，特别是酪氨酸；对应于 X41 的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是丝氨酸或丙氨酸；对应于 X53 的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、天冬氨酸或谷氨酸，特别是苏氨酸、甘氨酸或天冬氨酸；对应于 X57 的残基是甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是缬氨酸；对应于 X64 的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是丝氨酸、丙氨酸或缬氨酸；对应于 X66 的残基是天冬氨酸或谷氨酸；对应于 X68 的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、天冬氨酸、谷氨酸、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是苏氨酸、天冬氨酸、丙氨酸、缬氨酸或谷氨酸；对应于 X72 的残基是赖氨酸或精氨酸，特别是精氨酸；对应于 X76 的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、甘氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是丙氨酸或苏氨酸；而且对应于 X80 的残基是丝氨酸、 苏氨酸、谷氨酰胺、天冬酰胺、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是丙氨酸、苏氨酸或缬氨酸。在一些实施方案中，包括对应于 SEQ ID NO：111、112 或 139 中的序列式的残基 1-89 的区域或结构域的多肽与 SEQ ID NO：2、4 或 114 的参考序列相比在未被上述 X 所指定的残基位置上可另外具有一个或多个残基差异。在一些实施方案中，所述差异可以是与 SEQ ID NO：2、4 或 114 的参考序列相比在其他未被上述 X 所界定的氨基酸残基位置上的 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、

1-14、1-15 或 1-16 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基位置上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15 或 16 个残基差异。在一些实施方案中，所述差异包括保守突变。

[0178] 在一些实施方案中，酮还原酶多肽还可包括对应于 SEQ ID NO：111、112 或 139 中的序列式的残基 1-89 的结构域或区域，其中所述结构域或区域可至少具有下述特征：对应于 X7 的残基是受限的或非极性的残基，特别是组氨酸。在一些实施方案中，包括对应于 SEQ ID NO：111、112 或 139 中的序列式的残基 1-89 的区域或结构域的多肽与 SEQ ID NO：2、4 或 114 的参考序列相比在其他残基位置上可另外具有一个或多个残基差异。在一些实施方案中，对应于残基 1-89 的区域或结构域与基于 SEQ ID NO：2、4 或 114 的参考序列的相应结构域或区域相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15 或 1-16 个残基差异。在一些实施方案中，差异的数目可以是在所述结构域中其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15 或约 16 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶区域或结构域包括至少具有前述特征的氨基酸序列，而且其中所述氨基酸序列与对应于具有前述特征的基于 SEQ IDNO：2、4 或 114 的参考序列的残基 1-89 的氨基酸序列相比具有至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％的同一性。

[0179] 在一些实施方案中，酮还原酶多肽还可包括对应于 SEQ ID NO：111、 112 或 139 中的序列式的残基 1-89 的结构域或区域，其中所述结构域或区域可具有一种或多种或至少所有的下述特征：对应于 X7 的残基是受限的或非极性的残基，特别是组氨酸；而且对应于 X66 的残基是酸性的残基，特别是谷氨酸。在一些实施方案中，包括对应于 SEQ ID NO：111、112 或 139 中的序列式的残基 1-89 的区域或结构域的多肽与 SEQ ID NO：2、4 或 114 的参考序列相比在其他残基位置上可另外具有一个或多个残基差异。在一些实施方案中，对应于残基 1-89 的区域或结构域与基于 SEQ ID NO：2、4 或 114 的参考序列的相应结构域相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15 或 1-16 个残基差异。在一些实施方案中，差异的数目可以是在所述结构域中其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15 或约 16 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶区域或结构域包括至少具有前述特征的氨基酸序列，而且其中所述氨基酸序列与对应于具有前述特征的基于 SEQ IDNO：2、4 或 114 的参考序列的残基 1-89 的氨基酸序列相比具有至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％的同一性。

[0180] 在一些实施方案中，酮还原酶多肽还可包括对应于 SEQ ID NO：111、112 或 139 中的序列式的残基 1-89 的结构域或区域，其中所述结构域或区域可具有一种或多种或至少所有的下述特征：对应于 X7 的残基是受限的或非极性的残基，特别是组氨酸；对应于 X17 的残基是非极性的或脂肪族的残基，特别是甲硫氨酸；对应于 X27 的残基是芳香族的残基，特别是酪氨酸；对应于 X64 的残基是极性的、脂肪族的或非极性的残基，特别是缬氨酸；而且对应于 X80 的残基是脂肪族的、非极性的或极性的残基，特别是苏氨酸或缬氨酸。在一些实施方案中，包括对应于 SEQ ID NO：111、112 或 139 中的序列式的残基 1-89 的区域或结构域的多肽与 SEQ ID NO：2、4 或 114 的参考序列相比在其他残基位置上可另外具有一个或

多个残基差异。在一些实施方案中，对应于残基 1-89 的区域或结构域与基于 SEQ IDNO :2、4 或 114 的参考序列的相应结构域相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、 1-15 或 1-16 个残基差异。在一些实施方案中，差异的数目可以是在所述结构域中其他氨基酸残基上的 1、2、3、4、5、6、7、8、9、10、11、12、14、15 或约 16 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶区域或结构域包括至少具有前述特征的氨基酸序列，而且其中所述氨基酸序列与对应于具有前述特征的基于 SEQ IDNO :2、4 或 114 的参考序列的残基 1-89 的氨基酸序列相比具有至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％的同一性。

[0181] 在一些实施方案中，酮还原酶多肽还可包括对应于 SEQ ID NO :111、112 或 139 中的序列式的残基 212-252 的结构域或区域，其中所述结构域或区域可具有一种或多种下述特征 :对应于 X212 的残基是极性的、脂肪族的或非极性的 ;对应于 X216 的残基是受限的或碱性的残基 ;对应于 X223 的残基是非极性的或脂肪族的残基 ;对应于 X226 的残基是非极性的或极性的残基 ;对应于 X233 的残基是极性的或酸性的 ;对应于 X235 的残基是极性的或芳香族的残基 ;对应于 X248 的残基是非极性的或芳香族的残基 ;而且对应于 X249 的残基是芳香族的残基。在一些实施方案中，所述氨基酸序列可具有由上述 X 所界定的至少 2、3、4、5、6 或更多种另外特征。在一些实施方案中，包括基于 SEQ ID NO :111、112 或 139 中的序列式的氨基酸序列（或其区域）的多肽与 SEQ ID NO :2、4 或 114 的参考序列相比在未被上述 X 所指定的残基位置上可另外具有一个或多个残基差异。在一些实施方案中，所述差异可以是在其他未被上述 X 所界定的氨基酸残基位置上的 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9 或 1-10 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基位置上的 1、2、3、4、5、6、7、8、9 或约 10 个残基差异。在一些实施方案中，所述差异包括保守突变。

[0182] 在一些实施方案中，酮还原酶多肽还可包括对应于 SEQ ID NO :111、112 或 139 中的序列式的残基 212-252 的结构域或区域，其中所述结构域或区域可具有一种或多种下述特征 :对应于 X212 的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸 或异亮氨酸，特别是苏氨酸、丝氨酸或缬氨酸 ;对应于 X216 的残基是脯氨酸、组氨酸、精氨酸或赖氨酸，特别是组氨酸或精氨酸 ;对应于 X223 的残基是甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸或异亮氨酸，特别是缬氨酸 ;对应于 X226 的残基是甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸、苏氨酸、谷氨酰胺或天冬酰胺，特别是苏氨酸 ;对应于 X233 的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、天冬氨酸、谷氨酸，特别是天冬酰胺或天冬氨酸 ;对应于 X235 的残基是丝氨酸、苏氨酸、谷氨酰胺、天冬酰胺、酪氨酸、苯丙氨酸、色氨酸，特别是丝氨酸或色氨酸 ;对应于 X248 的残基是甘氨酸、甲硫氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、酪氨酸、苯丙氨酸或色氨酸，特别是甘氨酸或色氨酸 ;而且对应于 X249 的残基是酪氨酸、苯丙氨酸或色氨酸，特别是色氨酸。在一些实施方案中，所述氨基酸序列可具有由上述 X 所界定的至少 2、3、4、5、6 或更多种另外特征。在一些实施方案中，包括基于 SEQID NO :111、112 或 139 中的序列式的氨基酸序列（或其区域）的多肽与 SEQID NO :2、4 或 114 的参考序列相比在未被上述 X 所指定的残基位置上可另外具有一个或多个残基差异。在一些实施方案中，所述差异可以是在其他未被上述 X 所界定的

氨基酸残基位置上的 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9 或 1-10 个残基差异。在一些实施方案中，差异的数目可以是在其他氨基酸残基位置上的 1、2、3、4、5、6、7、8、9 或约 10 个残基差异。在一些实施方案中，所述差异包括保守突变。

[0183] 在一些实施方案中，酮还原酶多肽还可包括对应于 SEQ ID NO：111、112 或 139 中的序列式的残基 212-252 的结构域或区域，其中所述结构域或区域至少具有下述特征：对应于 X216 的残基是受限的或碱性的残基，特别是精氨酸。在一些实施方案中，包括对应于 SEQ ID NO：111、112 或 139 中的序列式的残基 212-252 的区域或结构域的多肽与具有前述特征的 SEQ ID NO：2、4 或 114 的参考序列相比在其他残基位置上可另外具有一个或多个残基差异。在一些实施方案中，对应于残基 212-252 的区域或结构域与基于 SEQ ID NO：2、4 或 114 的参考序列的相应结构域相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9 或 1-10 个残基差异。在一些实施方案中，差异的数目可以是在所述结构域中其他 氨基酸残基上的 1、2、3、4、5、6、7、8、9 或约 10 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶区域或结构域包括至少具有前述特征的氨基酸序列，而且其中所述氨基酸序列与对应于具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列的残基 212-252 的氨基酸序列相比具有至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％的同一性。

[0184] 在一些实施方案中，酮还原酶多肽还可包括对应于 SEQ ID NO：111、112 或 139 中的序列式的残基 212-252 的结构域或区域，其中所述结构域或区域可具有一种或多种或至少所有的下述特征：对应于 X212 的残基是极性的、脂肪族的或非极性的残基，特别是苏氨酸、丝氨酸或缬氨酸；对应于 X216 的残基是受限的或碱性的残基，特别是组氨酸或精氨酸；对应于 X223 的残基是极性的、脂肪族的或非极性的，特别是缬氨酸；对应于 X233 的残基是极性的或酸性的残基，特别是天冬酰胺或天冬氨酸；对应于 X235 的残基是极性的或芳香族的残基，特别是丝氨酸或色氨酸；对应于 X248 的残基是非极性的或芳香族的残基，特别是甘氨酸或色氨酸；而且对应于 X249 的残基是芳香族的残基，特别是色氨酸。在一些实施方案中，包括对应于 SEQ ID NO：111、112 或 139 中的序列式的残基 212-252 的区域或结构域的多肽与 SEQ ID NO：2、4 或 114 的参考序列相比在其他残基位置上可另外具有一个或多个残基差异。在一些实施方案中，对应于残基 212-252 的区域或结构域与基于 SEQ ID NO：2、4 或 114 的参考序列的相应结构域相比在其他氨基酸残基位置上可另外具有 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9 或 1-10 个残基差异。在一些实施方案中，差异的数目可以是在所述结构域中其他氨基酸残基上的 1、2、3、4、5、6、7、8、9 或约 10 个残基差异。在一些实施方案中，所述差异包括保守突变。在一些实施方案中，酮还原酶区域或结构域包括至少具有前述特征的氨基酸序列，而且其中所述氨基酸序列与对应于具有前述特征的基于 SEQ ID NO：2、4 或 114 的参考序列的残基 212-252 的氨基酸序列相比具有至少 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％的同一性。

[0185] 下表 2 提供本文公开的具有相关活性的特定 SEQ ID NO 的列表。除非另外说明，下面的序列衍生自野生型高加索酸奶乳杆菌酮还原酶序列 (SEQID NO：3 和 4)。在下表 2 中，每行列出两个 SEQ ID NO，其中的奇数指的是编码由偶数所提供的氨基酸序列的核苷酸序列。列出突变数目的列是关于与 SEQ ID NO：4 中的高加索酸奶乳杆菌 KRED 氨基酸序列相比之下氨基酸取代的数目。在活性列中，一个“+”对应于 10g/L 的酶在约 24 小时内转化

10-20％的 1g/L 底物的能力。两个加号“++”表示多肽比 SEQ ID NO :90 的活性高约 1 至 50 倍。三个加号“+++”表示多肽比 SEQ ID NO :90 的活性高约 50 至 250 倍。四个加号“++++”表示多肽比 SEQ ID NO :90 的活性高约 250 至 1000 倍，而五个加号“+++++”表示多肽比 SEQ ID NO :90 的活性高 1000 倍以上。

[0186] 表 2 :序列和相应的活性改进的列表

[0187]

SEQ ID NO	与 SEQ ID NO :4 相比的残基 改变	来自高加索酸奶乳 杆菌 SEQ ID NO :4 的突变数目	活性[a]
89/90	G7H ;S96T ;F147L ;Y190P ; V196L ;A202W	6	+
47/48	G7H ;A94G ;F147L ;Y190P ; V196L ;A202W	6	++
45/46	G7H ;S96L ;F147L ;Y190P ; V196L ;A202W	6	++
43/44	G7H ;S96V ;F147L ;Y190P ; V196L ;A202W	6	++
41/42	G7H ;S96I ;F147L ;Y190P ; V196L ;A202W	6	++
99/100	G7H ;S96T ;E145L ;F147L ; Y190P ;V196L ;A202W	7	++
91/92	G7H ;S96T ;F147L ;Y190P ; V 196L ;A202W ;Y249W	7	++
93/94	G7H ;S96T ;F147L ;Y190P ; V196L ;A202W ;K211V	7	++
101/102	G7H ;S96T ;F147L ;Y190P ; V196L ;A202W ;K211I	7	++
85/86	G7H ;S96T ;F147L ;Y190P ;	7	++

[0188]

SEQ ID NO	与 SEQ ID NO :4 相比的残基 改变	来自高加索酸奶乳 杆菌 SEQ ID NO :4 的突变数目	活性[a]
	L195R ;V196L ;A202W		
95/96	G7H ;S96T ;F147L ;Y190P ; L195D ;V196L ;A202W	7	++
87/88	G7H ;S96T ;F147L ;Y190P ; L195R ;V196L ;A202W ; K211I	8	++
97/98	G7H ;S96T ;F147L ;Y190P ; P194L ;L195R ;V196L ; A202W	8	+++
69/70	G7H ;D66E ;A94G ;S96V ; F147L ;Y190P ;L195R ; V196L ;A202W ;K211V	10	+++
83/84	G7H ;I93S ;A94C ;S96L ; F147L ;Y190P ;L195R ; V196L ;A202W ;K211V	10	+++
75/76	G7H ;D66E ;A80V ;A94C ; S96L ;F147L ;Y190P ;L195R ; V196L ;A202W ;K211L	11	+++
73/74	G7H ;D66E ;A80V ;I93S ; A94C ;S96V ;F147L ;Y190P ; L195R ;V196L ;A202W ;K211T	12	+++
63/64	G7H ;D66E ;I93S ;A94C ; S96L ;E145Q ;F147L ;Y190P ; L195R ;V196L ;A202W ;K211V	12	+++
103/104	G7H ;A94C ;S96I ;F147L ; Y190P ;L195R ;V196L ; A202W ;K211T	9	+++
67/68	G7H ;D66E ;A94G ;S96I ; F147L ;Y190P ;L195R ; V196L ;A202W ;K211L	10	+++
105/106	G7H ;A80V ;S96C ;F 147L ; Y190P ;L195R ;V196L ; A202W ;K211V	9	+++
71/72	G7H ;D66E ;A94G ;S96V ; F147L ;T152S ;Y190P ; L195R ;V196L ;A202W ;K211V ;S235W	12	+++
65/66	G7H ;D66E ;A94G ;V95L ;	11	+++

[0189]

SEQ ID NO	与 SEQ ID NO :4 相比的残基 改变	来自高加索酸奶乳 杆菌 SEQ ID NO :4 的突变数目	活性[a]
	S96V ;F147L ;Y190P ;L195R ; V196L ;A202W ;K211V		
23/24	G7H ;L17M ;F27Y ;A64V ; A80T ;I93S ;A94C ;S96L ; E100K ;F147L ;Y190P ; L195R ;V196L ;A202W ; K211V	15	++++
81/82	G7H ;I93S ;A94C ;S96L ; F147L ;Y190P ;P194Q ; L195R ;V196L ;A202W ; K211V	11	+++
79/80	G7H ;I93S ;A94C ;S96L ; F147L ;Y190P ;P194Q ; L195R ;V196L ;A202W ; K211V ;I226T	12	+++
107/108	G7H ;A68V ;I93S ;A94C ; S96L ;R108H ;F147L ;Y179F ; Y190P ;L195R ;V196L ; D198E ;A202W ;K211V	14	+++
49/50	G7H ;A68V ;I93S ;A94C ; S96L ;R108H ;F147L ;I165N ; Y190P ;P194S ;L195R ; V196L ;A202W ;K211I	14	+++
77/78	G7H ;I93S ;A94C ;S96L ; R108H ;F147L ;I165N ; Y190P ;P194Q ;L195R ; V196L ;A202W ;K211V ;	13	+++
9/10	G7H L17M ;F27Y ;A80T ; I93S ;A94C ;S96L ;F147L ; Y190P ;L195R ;V196L ; A202W ;K211V ;T212S	14	+++
33/34	G7H ;L17M ;I57V ;I93S ; A94C ;S96L ;F147L ;V148I ; D173V ;Y190P ;L195R ; V196L ;A202W ;K211L	14	+++
35/36	G7H ;L17M ;I57V ;I93S ; A94C ;S96L ;F147L ;Y190P ; L195R ;V196L ;A202W ; K211V	12	+++
57/58	G7H ;D66E ;I93S ;A94C ; S96I ;F147L ;Y190P ;P194Q ;	14	++++

[0190]

SEQ ID NO	与 SEQ ID NO :4 相比的残基 改变	来自高加索酸奶乳 杆菌SEQ ID NO :4 的突变数目	活性 a
	L195R ;V196L ;D198E ; A202W ;K211V ;H216R		
39/40	G7H ;D66E ;I93S ;A94C ; S96I ;F147L ;Y190P ;P194Q ; L195R ;V196L ;D198E ; A202W ;K211V	13	++++
31/32	G7H ;L17M ;I57V ;A80T ; I93S ;A94C ;S96L ;F147L ; V163I ;Y190P ;L195R ; V196L ;A202W ;K211V	14	+++
21/22	G7H ;L17M ;F27Y ;A64V ; A80T ;I93S ;A94C ;S96L ; E100K ;F147L ;Y190P ; L195R ;V196L ;A202W ; K211V	15	++++
37/38	G7H ;D66E ;I93S ;A94C ; S96I ;F147L ;Y190P ;P194Q ; L195R ;V196L ;D198E ; A202W ;K211V ;H216R	14	++++
53/54	G7H ;D66E ;I93S ;A94C ; S96I ;F147L ;P151A ;Y190P ; P194Q ;L195R ;V196L ; D198E ;A202W ;K211V ; H216R	15	++++
11/12	G7H ;L17M ;F27Y ;A64V ; A80T ;I93S ;A94C ;S96L ; E100K ;F147L ;Y190P ; L195R ;V196L ;A202W ; T210R ;K211V	16	++++
29/30	G7H ;L17M ;A64V ;A80T ; I93S ;A94C ;S96L ;E100K ; F147L ;Y190P ;L195R ; V196L ;A202W ;K211V ; H216R	15	++++
27/28	G7H ;L17M ;F27Y ;A64V ; I93S ;A94C ;S96L ;F147L ; Y190P ;P194Q ;L195R ; V196L ;A202W ;K211V ; H216R	15	++++
15/16	G7H ;L17M ;F27Y ;A64V ; A80T ;I93S ;A94C ;S96L ;	18	+++++

[0191]

SEQ ID NO	与 SEQ ID NO :4 相比的残基 改变	来自高加索酸奶乳 杆菌 SEQ ID NO :4 的突变数目	活性 a
	E100K ;F147L ;Y190P ; P194Q ;L195R ;V196L ; D198E ;A202W ;K211V ; H216R		
109/110	G7H ;A64V ;A80T ;I93S ; A94C ;S96L ;E100K ;F147L ; Y190P ;P194Q ;L195R ; V196L ;D198E ; A202W ; K211V ;H216R	16	++++
61/62	G7H ;D66E ;I93S ;A94C ; S96I ;K97H ;F147L ;Y190P ; P194Q ;L195R ;V196L ; D198E ;A202W ; K211V ; H216R	15	++++
59/60	G7H ;D66E ;I93S ;A94C ; S96I ;F147L ;T152S ;Y190P ; P194Q ;L195R ;V196L ; D198E ;A202W ; K211V ; H216R	15	++++
51/52	G7H ;D66E ;I93S ;A94C ; V95E ;S96I ;F147L ;Y190P ; P194Q ;L195R ;V196L ; D198E ;A202W ; K211V ; H216R	15	++++
55/56	G7H ;D66E ;I93S ;A94C ; S96I ;F147L ;Y190P ;P194Q ; L195R ;V196L ;D198Q ; A202W ;K211V ; H216R	14	++++
25/26	G7H ;L17M ;F27Y ;A64V ; I93S ;A94C ;S96L ;E 100K ; F147L ;Y190P ;P194Q ; L195R ;V196L ; D198E ; A202W ;K211V ;H216R	17	++++
13/14	G7H ;L17M ;F27Y ;A64V ; A80T ;I93S ;A94C ;S96L ; E100K ;F147L ;Y190P ; P194Q ;L195R ; V196L ; D198Q ;A202W ;K211V ; H216R	18	+++++
19/20	G7H ;L17M ;F27Y ;A64V ; A80T ;I93S ;A94C ;S96L ;	19	+++++

[0192]

SEQ ID NO	与 SEQ ID NO :4 相比的残基 改变	来自高加索酸奶乳 杆菌 SEQ ID NO :4 的突变数目	活性 a
	E100K ;F147L ;T152S ; Y190P ;P194Q ;L195R ; V196L ;D198Q ;A202W ; K211V ;H216R		
5/6	G7H ;L17M ;D25T ;F27Y ; G53D ;A64V ;A80T ;I93S ; A94C ;S96L ;E100K ;F147L ; Y190P ;P194Q ; L195R ; V196L ;D198E ;A202W ; K211V ;H216R	20	+++++
7/8	G7H ;L17M ;D25T ;F27Y ; G53D ;A64V ;T76A ;A80T ; I93S ;A94C ;S96L ;E100K ; F147L ;Y190P ; P194Q ; L195R ;V196L ;D198E ; A202W ;K211V ;H216R	21	+++++
17/18	G7H ;L17M ;F27Y ;A64V ; A80T ;I93S ;A94C ;S96L ; E100K ;F147L ;Y190P ; P194Q ;L195R ; V196L ; D198E ;A202W ;K211V ; T212V ;H216R ;I223V	20	++++

[0193] 在一些实施方案中，所述多肽与基于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、92、94、96、98、100、102、104、106、108或110的参考序列的多肽相比具有至少85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或99％的序列同一性，条件是酮还原酶多肽氨基酸序列与SEQ ID NO:4相比包括列在表2中的任何一组取代组合。在一些实施方案中，酮还原酶多肽与参考序列相比在其他氨基酸残基位置上可另外具有1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-25、1-30、1-35或约1-40个残基差异。在一些实施方案中，差异的数目可以是在其他氨基 酸残基上的1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35或约40个残基差异。在一些实施方案中，所述差异包括保守突变。

[0194] 在一些实施方案中，酮还原酶多肽能够以至少约99％的百分比立体异构过量和比具有SEQ ID NO:90的氨基酸序列的参考多肽改进的速率将底物2-[3-[3-[2-(7-氯-2-喹啉基）乙烯基]苯基]-3-氧代丙基]苯甲酸甲酯转化为产物(S,E)-2-(3-(3-(2-(7-氯喹啉-2-基）乙烯基）苯基)-3-羟丙基）苯甲酸甲酯。具有这种性质的示例性多肽包括但不限于含有对应于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、92、94、96、98、100、102、104、106、108和110的氨基酸序列的多肽。因为具有SEQ ID NO:90的氨基酸序列的参考多肽能够以比野生型(SEQ ID NO:4)改进的速率（例如，用约10g/L的KRED在24小时内将10-20％的1g/L底物转化为产物）和立体选择性(99％的立体异构过量）将底物转化为产物，所以本文中比SEQ IDNO:90改进的多肽也是比野生型改进的。

[0195] 在一些实施方案中，酮还原酶多肽能够以至少约99％的百分比立体异构过量和比具有SEQ ID NO:90的氨基酸序列的参考多肽改进了至少1-50倍的速率将底物2-[3-[3-[2-(7-氯-2-喹啉基）乙烯基]苯基]-3-氧代丙基]苯甲酸甲酯转化为产物(S,E)-2-(3-(3-(2-(7-氯喹啉-2-基）乙烯基）苯基)-3-羟丙基）苯甲酸甲酯。具有这种性质的示例性多肽包括但不限于含有对应于SEQ IDNO:42、44、46、48、86、88、92、94、96、100和102的氨基酸序列的多肽。

[0196] 在一些实施方案中，酮还原酶多肽能够以至少约99％的百分比立体异构过量和比具有SEQ ID NO:90的氨基酸序列的参考多肽改进了至少50-250倍的速率将底物2-[3-[3-[2-(7-氯-2-喹啉基）乙烯基]苯基]-3-氧代丙基]苯甲酸甲酯转化为产物(S,E)-2-(3-(3-(2-(7-氯喹啉-2-基）乙烯基）苯基)-3-羟丙基）苯甲酸甲酯。具有这种性质的示例性多肽包括但不限于含有对应于SEQ ID NO:10、32、34、36、50、64、66、68、70、72、74、76、78、80、 82、84、98、104、106和108的氨基酸序列的多肽。

[0197] 在一些实施方案中，酮还原酶多肽能够以至少约99％的百分比立体异构过量和比具有SEQ ID NO:90的氨基酸序列的参考多肽改进了至少250-1000倍的速率将底物2-[3-[3-[2-(7-氯-2-喹啉基）乙烯基]苯基]-3-氧代丙基]苯甲酸甲酯转化为产物(S,E)-2-(3-(3-(2-(7-氯喹啉-2-基）乙烯基）苯基)-3-羟丙基）苯甲酸甲酯。具有这种性质的示例性多肽包括但不限于含有对应于SEQ ID NO:12、18、22、24、26、28、30、38、40、52、54、56、58、60、62和110的氨基酸序列的多肽。

[0198] 在一些实施方案中，酮还原酶多肽能够以至少约 99％的百分比立体异构过量和比具有 SEQ ID NO：90 的氨基酸序列的参考多肽改进了至少 1000 倍的速率将底物 2-[3-[3-[2-(7- 氯 -2- 喹啉基）乙烯基］苯基］-3- 氧代丙基］苯甲酸甲酯转化为产物 (S, E)-2-(3-(3-(2-(7- 氯喹啉 -2- 基）乙烯基）苯基）-3- 羟丙基）苯甲酸甲酯。具有这种性质的示例性多肽包括但不限于含有对应于 SEQ IDNO：6、8、14、16 和 20 的氨基酸序列的多肽。

[0199] 在一些实施方案中，与具有 SEQ ID NO：90 的氨基酸序列的参考多肽相比，改进的酮还原酶多肽可具有增加的对丙酮的抵抗性，其中丙酮被作为底物向产物转化中的竞争性抑制剂。具有这种增加的抵抗性的示例性多肽包括但不限于含有对应于 SEQ ID NO：6、8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、38、40、50、52、54、56、58、60、62、64、66、68、70、72、74、76、78、80、82、84、86、88、98、104、106、108 和 110 的氨基酸序列的多肽。

[0200] 在一些实施方案中，当用超过约 100g/L 的底物和小于约 5g/L 的多肽进行时，酮还原酶多肽能够在小于约 24 小时中将至少约 95％的底物转化为产物。具有这种能力的示例性多肽包括但不限于含有对应于 SEQ IDNO：6、8、14、16 和 20 的氨基酸序列的多肽。

[0201] 在一些实施方案中，本公开的酮还原酶多肽是高度地立体选择性的，以便它能够将底物还原为具有超过约 99％、99.1％、99.2％、99.3％、99.4％、99.5％、99.6％、99.7％、99.8％或 99.9％立体异构过量的产物。具有这种高度立体选择性的示例性酮还原酶多肽包括但不限于含有对应于 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 和 110 的氨基酸序列的多肽。

[0202] 在一些实施方案中，改进的酮还原酶多肽可包括与 SEQ ID NO：2 或 SEQ ID NO：4 或 SEQ ID NO：114 或者其区域或结构域（例如残基 90-211）至少约 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的氨基酸序列，条件是对应于 SEQ IDNO：2、4 或 114 中的残基 190 的残基是脯氨酸，对应于 SEQ ID NO：2、4 或 114 中的残基 195 的残基是精氨酸，而且对应于 SEQ ID NO：2、4 或 114 中的残基 202 的残基是色氨酸，并且改进的酮还原酶多肽另外具有一种或多种下述取代以使该多肽比野生型高加索酸奶乳杆菌酮还原酶或另一种人工改造的酮还原酶（例如 SEQ ID NO：90）进一步改进（例如，关于立体选择性、酶活性和 / 或热稳定性）：7 → H；17 → M；25 → T；27 → Y；41 → S；53 → D；57 → V；64 → V；66 → E；68 → V，E；72 → R；76 → A；80 → T，V；93 → S，A，G；94 → C，G；95 → L，E；96 → V，L，I，T；97 → H，S；100 → K，Q，A；108 → H；111 → M；139 → L；145 → L；147 → L；152 → W，A，S；163 → I；179 → F，K，L；194 → Q，S，L；196 → L；197 → E；198 → E，Q；210 → R；211 → V，L，R，T，D，I；212 → S，V；216 → R；223 → V；233 → N；235 → W；248 → W 和 249 → W。

[0203] 在一些实施方案中，改进的酮还原酶多肽可包括与 SEQ ID NO：2、4 或 114（或者其区域或结构域，例如残基 90-211）至少约 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的氨基酸序列，条件是对应于 SEQ ID NO：2、4 或 114 中的残基 190 的残基是脯氨酸；对应于 SEQ ID NO：2、4 或 114 中的残基 195 的残基是精氨酸；而且对应于 SEQ ID NO：2、4 或 114 中的残基 202 的残基是色氨酸，并且改进的酮还原酶多肽另外具有至少一种或多种下述取代以使该多肽比野生型高加索酸奶乳杆

菌酮还原酶或另一种人工改造的酮还原酶（例如 SEQ ID NO:90）进一步改进（例如，关于立体选择性、酶活性和/或热稳定性）：17→M；27→Y；64→V；80→T；93→S；94→C；96→L；100→K； 147→L；152→S；194→Q 和 211→V。

[0204] 如本领域技术人员将理解的，除非另外说明，上述界定的氨基酸类别中的一些不互斥。因此，具有展现两种或更多种物理化学性质的侧链的氨基酸可被包括在多个类别中。任何氨基酸或残基的适当分类对于本领域技术人员将是明显的，特别是鉴于本文提供的详细公开。

[0205] 在一些实施方案中，改进的人工改造的酮还原酶包括天然存在的酮还原酶多肽的缺失或其他人工改造的酮还原酶多肽的缺失。在一些实施方案中，本文所述的每种改进的人工改造的酮还原酶可包括本文所述的多肽的缺失。因此，对于本公开的酮还原酶多肽的各个和每个实施方案来说，缺失可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 个或更多个氨基酸或者 20 个或更多个氨基酸，酮还原酶多肽的氨基酸总数目的至多 10%、氨基酸总数目的至多 20%或氨基酸总数目的至多 30%，只要能维持酮还原酶活性的功能性活性。在一些实施方案中，缺失可包括 1-2、1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、1-11、1-12、1-14、1-15、1-16、1-18、1-20、1-22、1-24、1-25、1-30、1-35 或约 1-40 个氨基酸残基。在一些实施方案中，缺失的数目可以是 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18、20、22、24、26、30、35 或约 40 个氨基酸。在一些实施方案中，缺失可包括 1、2、3、4、5、6、7、8、9、10、11、12、14、15、16、18 或 20 个氨基酸残基的缺失。

[0206] 如本文所述，本公开的酮还原酶多肽可以是融合多肽的形式，其中所述酮还原酶多肽被融合至其他多肽，例如但不限于抗体标签（例如 myc 表位）、纯化序列（例如 His 标签）和细胞定位信号（例如分泌信号）。因此，酮还原酶多肽可以与或者不与其他多肽融合而使用。

[0207] 本文所述多肽不限于被遗传编码的氨基酸。除了被遗传编码的氨基酸之外，本文所述的多肽可以全部或部分地包括天然存在的和/或合成的非编码氨基酸。可能包含本文所述多肽的某些常见的非编码氨基酸包括但不限于：被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)；t- 丁基丙氨酸 (Bua)；t- 丁基甘氨酸 (Bug)；N- 甲基异亮氨酸 (MeIle)；苯甘氨酸 (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- 氟代苯丙氨酸 (Off)；3- 氟代苯丙氨酸 (Mff)；4- 氟代苯丙氨酸 (Pff)；2- 溴代苯丙氨酸 (Obf)；3- 溴代苯丙氨酸 (Mbf)；4- 溴代苯丙氨酸 (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- 三氟代甲基苯丙氨酸 (Otf)；3- 三氟代甲基苯丙氨酸 (Mtf)；4- 三氟代甲基苯丙氨酸 (Ptf)；4- 氨基苯丙氨酸 (Paf)；4- 碘代苯丙氨酸 (Pif)；4- 氨基甲基苯丙氨酸 (Pamf)；2,4- 二氯苯丙氨酸 (Opef)；3,4- 二氯苯丙氨酸 (Mpcf)；2,4- 二氟代苯丙氨酸 (Opff)；3,4- 二氟代苯丙氨酸 (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)；戊氟代苯丙氨酸(5ff)；苯乙烯基丙氨酸(sAla)；蒽基丙氨酸(aAla)；3,3-二苯丙氨酸(Dfa)；3-氨基-5-苯基戊酸(Afp)；青霉胺(Pen)；1,2,3,4-四氢异喹啉-3-羧酸(Tic)；β-2-噻吩基丙氨酸(Thi)；甲硫氨酸亚砜(Mso)；N(w)-硝基精氨酸(nArg)；高赖氨酸(hLys)；膦酰基甲基苯丙氨酸(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；哌啶酸(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)；羟脯氨酸(Hyp)和高脯氨酸(hPro)。可能包含本文所述多肽的另外的非编码氨基酸对于本领域技术人员将是 明显的(参见例如Fasman,1989,CRC Practical Handbook of Biochemistryand Molecular Biology,CRC Press,Boca Raton,FL,第3-70页和本文所引用的参考文献中提供的各种氨基酸,其全部通过引用被并入)。这些氨基酸可以是L-或D-构型的。

[0208] 本领域技术人员将发现,含有侧链保护基团的氨基酸或残基还可包括本文所述的多肽。这种受保护的氨基酸(在这种情况下其属于芳香族类别)的非限制性实例包括(括号中列出保护基团)但不限于:Arg(tos)、Cys(甲基苄基)、Cys(硝基吡啶亚氧硫基)、Glu(δ-苄酯)、Gln(氧杂蒽基)、Asn(N-δ-氧杂蒽基)、His(bom)、His(苄基)、His(tos)、Lys(fmoc)、Lys(tos)、Ser(O-苄基)、Thr(O-苄基)和Tyr(O-苄基)。

[0209] 可能包含本文所述多肽的、构象受限的非编码氨基酸包括但不限于N-甲基氨基酸(L-构型)；1-氨基环戊-(2或3)-烯-4-羧酸；哌啶酸；氮杂环丁烷-3-羧酸；高脯氨酸(hPro)和1-氨基环戊烷-3-羧酸。

[0210] 如上所述,被引入天然存在的多肽以产生人工改造的酮还原酶的各种修饰可针对该酶的特殊性质。

[0211] 6.3编码人工改造的酮还原酶的多核苷酸

[0212] 在另一方面,本公开提供编码人工改造的酮还原酶的多核苷酸。所述多核苷酸可以被可操作地连接至一个或多个控制基因表达的异源性调节序列以产生能够表达该多肽的重组多核苷酸。含有编码人工改造的酮还原酶的异源性多核苷酸的表达构建体可被引入适当的宿主细胞以表达相应的酮还原酶多肽。

[0213] 因为对应于各种氨基酸的密码子的知识,蛋白序列的可用性提供对所有能够编码所述蛋白序列的多核苷酸的描述。遗传密码的简并性(在其中相同的氨基酸被可选的或同义的密码子所编码)允许产生极大量的核酸,其所有都编码本文公开的改进的酮还原酶。因此,在已经鉴定了特定氨基酸序列之后,本领域技术人员可以以不改变蛋白氨基酸序列的方式通过简单地修饰一个或多个密码子的序列来制备任何数量的不同核酸。在这点上,本公开特别预期能够通过选择基于可能的密码子选择的组合来进行多 核苷酸的各个和每个可能的改变,而且所有这种改变都将被视为为了本文所公开的任何多肽而被特别公开的,其中所述多肽包括呈现于表2中的氨基酸序列。在一些实施方案中,优选地选择密码子以适应产生蛋白的宿主细胞。例如,用于细菌中的优选的密码子被用来表达细菌中的基因；

用于酵母中的优选的密码子被用于酵母中的表达；而且用于哺乳动物中的优选的密码子被用于哺乳动物细胞中的表达。举例来说，SEQ ID NO:3 的多核苷酸是为了大肠杆菌中的表达而被优化的密码子，但是另外还编码高加索酸奶乳杆菌的天然存在的酮还原酶。

[0214] 在一些实施方案中，不需要替换所有密码子以将酮还原酶的密码子选择最佳化，因为天然序列将包括优选的密码子并且因为优选的密码子的使用对于所有氨基酸残基可能不是必需的。因此，编码酮还原酶的密码子优化的多核苷酸可在全长编码区域中约 40%、50%、60%、70%、80%或超过 90%的密码子位置上含有优选的密码子。

[0215] 在一些实施方案中，多核苷酸包括编码酮还原酶多肽的核苷酸序列，所述酮还原酶多肽具有与本文所述任何参考人工改造的酮还原酶多肽相比具有至少约 80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%或更高序列同一性的氨基酸序列。因此，在一些实施方案中，多核苷酸编码与具有下述特征的基于 SEQID NO:2、4 或 114 的参考序列至少约 80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多地相同的氨基酸序列：对应于位置 X190 的残基是受限残基，特别是脯氨酸；而且对应于 X202 的残基是芳香族残基，特别是色氨酸；条件是被编码的酮还原酶多肽具有这样的氨基酸序列，在其中对应于 X190 的残基是受限残基，而且对应于 X202 的残基是芳香族残基。在一些实施方案中，多核苷酸编码具有一种氨基酸序列的酮还原酶，所述氨基酸序列中对应于 X190 的残基是脯氨酸；而且对应于 X202 的残基是色氨酸。

[0216] 在一些实施方案中，多核苷酸编码包括与具有下述特征的基于 SEQ IDNO:2、4 或 114 的参考序列至少约 80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多地相 同的氨基酸序列的酮还原酶多肽：对应于位置 X190 的残基是受限残基，特别是脯氨酸；对应于 X195 的残基是碱性的、酸性的、非极性的或极性的残基，特别是脯氨酸、精氨酸、赖氨酸、天冬氨酸或天冬酰胺；而且对应于 X202 的残基是芳香族残基，特别是色氨酸；条件是被编码的酮还原酶多肽包括至少具有前述特征的氨基酸序列（即，对应于位置 X190 的残基是受限残基；对应于 X195 的残基是碱性的、酸性的、非极性的或极性的残基；而且对应于 X202 的残基是芳香族残基。在一些实施方案中，被编码的酮还原酶具有这样的氨基酸序列，其中对应于 X190 的残基是脯氨酸；对应于 X195 的残基是脯氨酸、精氨酸、赖氨酸、天冬氨酸或天冬酰胺；而且对应于 X202 的残基是色氨酸。在一些实施方案中，多核苷酸编码包括选自 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、92、94、96、98、100、102、104、106、108 和 110 的氨基酸序列的人工改造的酮还原酶多肽。

[0217] 在一些实施方案中，多核苷酸编码包括与选自 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、92、94、96、98、100、102、104、106、108 和 110 的参考序列至少约 80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多地相同的氨基酸序列的酮还原酶多肽，其中被编码的多肽具有本文中对应于 X190 和 X202 的残基的特定特征，或者具有本文中的残基 X190、X195 和 X202 的特定特征。

[0218] 在一些实施方案中，编码人工改造的酮还原酶的多核苷酸选自 SEQ IDNO: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 和 109。在一些实施方案中，多核苷酸能够在高度严 格的条件下与包括 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 或 109 的多核苷酸杂交，其中在高度严格的条件下杂交的多核苷酸编码能够将式 (I) 的底物还原或转化为式 (II) 的产物的酮还原酶多肽。在一些实施方案中，由严格地杂交的多核苷酸所编码的多肽能够将式 (III) 的底物还原或转化为式 (IV) 的产物。在一些实施方案中，由严格地杂交的多核苷酸所编码的多肽将具有特定的酶活性，以及对应于 X190 和 X202 的残基的特定特征或者对应于 X190、X195 和 X202 的残基的特定特征。

[0219] 在一些实施方案中，多核苷酸编码本文所述的多肽，但在核苷酸水平上与编码人工改造的酮还原酶的参考多核苷酸相比具有约 80％或更高的序列同一性，约 85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％或更高的序列同一性。在一些实施方案中，所述参考多核苷酸选自 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 和 109。

[0220] 可以以各种方式对编码改进的酮还原酶多肽的分离的多核苷酸进行操作以提供多肽的表达。取决于表达载体，在插入载体之前对分离的多核苷酸的操作可以是期望的或必要的。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众所周知的。指导被提供在 Sambrook 等.，2001，Molecular Cloning:A Laboratory Manual，第 3 版，Cold Spring HarborLaboratory Press;和 Current Protocols in Molecular Biology，Ausubel.F. 编辑，Greene Pub.Associates，1998，更新至 2006 中。

[0221] 关于细菌宿主细胞，用于指导本公开核酸构建体的转录的适合的启动子包括从大肠杆菌 lac 操纵子、大肠杆菌 trp 操纵子、天蓝色链霉菌 (Streptomyces coelicolor) 琼脂水解酶基因 (dagA)、枯草芽孢杆菌果聚糖蔗糖酶基因 (sacB)、地衣型芽胞杆菌 (Bacillus licheniformis) α-淀粉酶 基因 (amyL)、嗜热脂肪芽胞杆菌 (Bacillus stearothermophilus) 麦芽糖淀粉酶基因 (amyM)、解淀粉芽胞杆菌 (Bacillus amyloliquefaciens) α-淀粉酶基因 (amyQ)、地衣型芽胞杆菌青霉素酶基因 (penP)、枯草芽孢杆菌 xylA 和 xylB 基因和原核 β-内酰胺酶基因获得的启动子 (Villa-Kamaroff 等.，1978，Proc.Natl Acad.Sci.USA 75:3727-3731)，以及 tac 启动子 (DeBoer 等.，1983，Proc.Natl Acad.Sci.USA 80:21-25)。

[0222] 关于丝状真菌宿主细胞，用于指导本公开核酸构建体的转录的适合的启动子包括从米曲霉 (Aspergillus oryzae) TAKA 淀粉酶、米赫根毛霉 (Rhizomucor miehei) 天冬氨酸蛋白酶、黑曲霉 (Aspergillus niger) 中性 α-淀粉酶、黑曲霉酸稳定性 α-淀粉酶、黑曲霉或泡盛曲霉 (Aspergillusawamori) 葡萄糖淀粉酶 (glaA)、米赫根毛霉脂肪酶、米曲霉碱性蛋白酶、米曲霉丙糖磷酸异构酶、构巢曲霉 (Aspergillus nidulans) 乙酰胺酶和尖孢镰刀菌 (Fusarium oxysporum) 胰蛋白酶样蛋白酶的基因获得的启动子 (WO96/00787)，以及 NA2-tpi 启动子 (来自黑曲霉中性 α-淀粉酶和米曲霉丙糖磷酸异构酶基因的启动子的杂

交体）及其突变的、截短的和杂交的启动子。

[0223] 在酵母宿主中，有用的启动子可来自酿酒酵母烯醇化酶 (ENO-1)、酿酒酵母半乳糖激酶 (GAL1)、酿酒酵母醇脱氢酶 / 甘油醛 -3- 磷酸脱氢酶 (ADH2/GAP) 和酿酒酵母 3- 磷酸甘油酸酯激酶的基因。酵母宿主细胞的其他有用的启动子被描述于 Romanos 等 .，1992，Yeast 8 :423-488。

[0224] 控制序列还可以是适合的转录终止子序列，即被宿主细胞识别以终止转录的序列。终止子序列被可操作地连接至编码多肽的核酸序列的 3’末端。在所选宿主细胞中有功能性的任何终止子可被用于本发明。

[0225] 例如，丝状真菌宿主细胞的示例性转录终止子可获得自米曲霉 TAKA 淀粉酶、黑曲霉葡萄糖淀粉酶、构巢曲霉邻氨基苯甲酸合成酶、黑曲霉 α - 葡萄糖苷酶和尖孢镰刀菌胰蛋白酶样蛋白酶的基因。

[0226] 酵母宿主细胞的示例性终止子可获得自酿酒酵母烯醇化酶、酿酒酵母细胞色素 C(CYC1) 和酿酒酵母甘油醛 -3- 磷酸脱氢酶的基因。酵母宿主细胞的其他有用的终止子被描述于 Romanos 等 .，1992，见上文。

[0227] 控制序列还可以是适合的前导序列，即对于宿主细胞的翻译重要的 mRNA 的非翻译区。前导序列被可操作地连接至编码多肽的核酸序列的 5’末端。在所选宿主细胞中有功能性的任何前导序列可被使用。示例性的前导序列包括 M13 噬菌体 pIII 前导序列、大肠杆菌麦芽糖结合蛋白前导序列和大肠杆菌 pelB 前导序列。丝状真菌宿主细胞的前导序列可获得自米曲霉 TAKA 淀粉酶和构巢曲霉丙糖磷酸异构酶的基因。酵母宿主细胞的适合的前导序列获得自酿酒酵母烯醇化酶 (ENO-1)、酿酒酵母 3- 磷酸甘油酸酯激酶、酿酒酵母 α - 因子和酿酒酵母醇脱氢酶 / 甘油醛 -3- 磷酸脱氢酶 (ADH2/GAP) 的基因。

[0228] 控制序列还可以是多腺苷酸化序列，即被可操作地连接至核酸序列的 3’末端的序列，而且其在被转录时作为向被转录的 mRNA 添加多腺苷酸残基的信号而被宿主细胞识别。在所选宿主细胞中有功能性的任何多腺苷酸化序列可被用于本发明。丝状真菌宿主细胞的示例性多腺苷酸化序列可来自米曲霉 TAKA 淀粉酶、黑曲霉葡萄糖淀粉酶、构巢曲霉邻氨基苯甲酸合成酶、尖孢镰刀菌胰蛋白酶样蛋白酶和黑曲霉 α - 葡萄糖苷酶。酵母宿主细胞的有用的多腺苷酸化序列被描述于 Guo 和 Sherman，1995，Mol Cell Bio15 :5983-5990。

[0229] 控制序列还可以是信号肽编码区，其编码被连接至多肽氨基末端并将被编码的多肽引导进入细胞分泌途径的氨基酸序列。核酸序列的编码序列的 5’端可内在地含有在翻译阅读框架中被天然地连接编码分泌多肽的编码区片段的信号肽编码区。可选地，编码序列的 5’端可含有对于所述编码序列来说是外来的单个肽编码区。当编码序列不天然地含有信号肽编码区时，外来的信号肽编码区可能是必需的。

[0230] 可选地，外来的信号肽编码区可简单地替换天然的信号肽编码区以便增强多肽的分泌。然而，将表达多肽引导进入所选宿主细胞分泌途径的任何信号肽编码区可被用于本发明。

[0231] 细菌宿主细胞的有效的信号肽编码区是从杆菌属 (Bacillus)NClB11837 麦芽糖淀粉酶、嗜热脂肪芽胞杆菌 α - 淀粉酶、地衣型芽胞杆菌枯草蛋白酶、地衣型芽胞杆菌 β - 内酰胺酶、嗜热脂肪芽胞杆菌中性蛋白酶 (nprT、nprS、nprM) 和枯草芽孢杆菌 prsA 的基因获得的信号肽编码区。进一步的信号肽被描述于 Simonen 和 Palva，1993，Microbiol

Rev 57 :109-137。

[0232] 丝状真菌宿主细胞的有效的信号肽编码区可以是从米曲霉 TAKA 淀粉酶、黑曲霉中性淀粉酶、黑曲霉葡萄糖淀粉酶、米赫根毛霉天冬氨酸蛋白酶、特异腐质霉 (Humicola insolens) 纤维素酶和嗜热羊毛状腐质霉 (Humicola lanuginosa) 脂肪酶的基因获得的信号肽编码区。

[0233] 酵母宿主细胞的有用的信号肽可来自酿酒酵母 α - 因子和酿酒酵母转化酶的基因。其他有用的信号肽编码区被描述于 Romanos 等 .,1992,见上文。

[0234] 控制序列还可以是前肽编码区,其编码被置于多肽氨基末端的氨基酸序列。所得的多肽被称作前酶或前多肽 (或者在一些情况下的酶原)。前多肽通常是无活性的,并且可通过从所述前多肽中将前肽催化或自催化裂解而被转化为成熟的活性多肽。前肽编码区可获得自枯草芽孢杆菌碱性蛋白酶 (aprE)、枯草芽孢杆菌中性蛋白酶 (nprT)、酿酒酵母 α - 因子、米赫根毛霉天冬氨酸蛋白酶和嗜热毁丝霉 (Myceliophthora thermophila) 乳糖酶的基因 (WO 95/33836)。

[0235] 当信号肽区域和前肽区域二者都存在于多肽的氨基末端时,所述前肽区域被置于多肽氨基末端邻近处,而且所述信号肽区域被置于所述前肽区域的氨基末端邻近处。

[0236] 还可能期望加入调节序列,其允许对与宿主细胞生长相关的多肽表达进行调节。调节系统的实例是导致基因表达根据化学或物理刺激而被打开或关闭的那些,所述刺激包括调节化合物的存在。在原核宿主细胞中,适合的调节序列包括 lac、tac 和 trp 操纵子系统。在酵母宿主细胞中,适合的调节系统包括例如 ADH2 系统或 GAL1 系统。在丝状真菌中,适合的调节序列包括 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。

[0237] 调节序列的其他实例是允许基因扩增的那些。在真核系统中,这些包括在氨甲蝶呤存在下被扩增的二氢叶酸还原酶基因以及与重金属一起的被扩增的金属硫蛋白基因。在这些情况下,编码本发明的 KRED 多肽的核 酸序列将被可操作地连接至调节序列。

[0238] 因此,在另一个实施方案中,本公开还针对包括编码人工改造的酮还原酶多肽或其变体的多核苷酸以及一个或多个表达调节区例如启动子和终止子、复制起点等等的重组表达载体,这取决于它们将被引入的宿主类型。上述各种核酸和控制序列可被连接在一起以产生重组表达载体,其可包括一个或多个方便的限制位点以允许在此类位点插入或取代编码多肽的核酸序列。可选地,可通过将包括所述序列的核酸序列或核酸构建体插入用于表达的适合的载体来表达本公开的核酸序列。在产生表达载体时,编码序列位于载体中以便将编码序列可操作地连接至用于表达的适合的控制序列。

[0239] 重组表达载体可以是任何载体 (例如质粒或病毒),其可被方便地进行重组 DNA 程序并且可导致多核苷酸序列的表达。载体的选择通常将取决于载体与其将被引入的宿主细胞之间的相容性。载体可以是线状或闭环质粒。

[0240] 表达载体可以是其复制独立于染色体复制的自主复制载体,即作为染色体外的实体存在的载体,例如质粒、染色体外元件、微型染色体或人工染色体。该载体可含有用于保证自复制的任何工具。可选地,该载体可以是在被引入宿主细胞时被整合进基因组并与其被整合进入的染色体一起复制的载体。此外,可以使用待引入宿主细胞基因组的单个载体或质粒或者共同含有总 DNA 的两个或更多个载体或质粒,或者转座子。

[0241] 本发明的表达载体优选地含有一个或多个可选择标志物，其允许对被转化的细胞的便利选取。可选择标志物是这样的基因，其产物提供杀生物剂或病毒抗性、重金属抗性、营养缺陷体的原营养以及诸如此类。细菌的可选择标志物的实例是来自枯草芽孢杆菌或地衣型芽胞杆菌的 dal 基因，或者赋予抗生素抗性例如氨苄西林、卡那霉素、氯霉素（实施例 1）或四环素抗性的标志物。酵母宿主细胞的适合的标志物是 ADE2、HIS3、LEU2、LYS2、MET3、TRP1 和 URA3。

[0242] 用于丝状真菌宿主细胞的可选择标志物包括但不限于 amdS（乙酰胺酶）、argB（鸟氨酸氨甲酰转移酶）、bar（草丁膦乙酰转移酶）、hph（潮霉 素磷酸转移酶）、niaD（硝酸盐还原酶）、pyrG（乳清酸核苷 -5′ - 磷酸脱羧酶）、sC（硫酸腺苷基转移酶）和 trpC（邻氨基苯甲酸合成酶）及其等价物。用于曲霉属细胞的实施方案包括构巢曲霉或米曲霉的 amdS 和 pyrG 基因和吸水链霉菌 (Streptomyces hygroscopicus) 的 bar 基因。

[0243] 本发明的表达载体优选地含有这样的元件，其允许载体向宿主细胞基因组中的整合或细胞中载体独立于基因组的自主复制。关于向宿主细胞基因组中的整合，该载体可依赖于编码多肽的核酸序列或用于通过同源或非同源的重组将载体整合进入基因组的任何其他载体元件。

[0244] 可选地，表达载体可含有另外的核酸序列，其用于指导通过同源重组向宿主细胞基因组中的整合。所述另外的核酸序列使得载体被整合进宿主细胞基因组中染色体的精确位置处。为了增加在精确位置处整合的相似性，整合元件应该优选地含有足够数目的核酸，例如 100 至 10,000 个碱基对儿，优选地 400 至 10,000 个碱基对儿，而且最优选地 800 至 10,000 个碱基对儿，其中所述核酸与相应的靶序列是高度同源的以提高同源重组的可能性。整合元件可以是与宿主细胞基因组中的靶序列同源的任何序列。此外，整合元件可以是非编码的或编码的核酸序列。在另一方面，可通过非同源重组将载体整合进入宿主细胞的基因组。

[0245] 关于自主复制，载体还可包括复制起点，其使得载体能够在所讨论的宿主细胞中自主地复制。细菌复制起点的实例是 P15A ori（如 FIG. 5 中的质粒所示）或者质粒 pBR322、pUC19、pACYC177（其中质粒具有 P15A ori）的复制起点，或者允许大肠杆菌中的复制的 pACYC184 和允许芽胞杆菌中的复制的 pUB110、pE194、pTA1060 或 pAMβ1。用于酵母宿主细胞的复制起点的实例是 2 微米的复制起点，即 ARS1、ARS4、ARS1 和 CEN3 的组合、以及 ARS4 和 CEN6 的组合。复制起点可以是具有一种突变的复制起点，所述突变使其在宿主细胞中的功能是温度敏感的（参见例如 Ehrlich，1978，Proc Natl Acad Sci.USA 75：1433）。

[0246] 可将本发明多于一个拷贝的核酸序列插入宿主细胞中以增加基因产物的生产。可通过将至少一个另外拷贝的序列整合进入宿主细胞基因组或者通过包括具有核酸序列的可扩增的可选择标志物基因来获得核酸序列 拷贝数的增加，其中含有扩增的拷贝数的可选择标志物基因以及因此额外拷贝的核酸序列的细胞可通过在适当的可选择的剂的存在下培养细胞来进行选择。

[0247] 用于本发明的许多表达载体是商业上可得的。适合的商品化的表达载体包括来自 Sigma-Aldrich Chemicals，St.Louis MO. 的 p3xFLAGTM™ 表达载体，其包括用于哺乳动物宿主细胞中表达的 CMV 启动子和 hGH 多腺苷酸化位点以及用于大肠杆菌中扩增的 pBR322 复制起点和氨苄西林抗性标志物。其他适合的表达载体是 pBluescriptII SK(-)

和 pBK-CMV(其可通过商业从 Stratagene，LaJolla CA 获得)，以及衍生自 pBR322(Gibco BRL)、pUC(Gibco BRL)、pREP4、pCEP4(Invitrogen) 或 pPoly(Lathe 等.，1987，Gene 57：193-201) 的质粒。

[0248] 6.4 用于表达酮还原酶多肽的宿主细胞

[0249] 在另一方面，本公开提供包括编码本公开改进的酮还原酶多肽的多核苷酸的宿主细胞，所述多核苷酸被可操作地连接至用于表达宿主细胞中酮还原酶的一个或多个控制序列。用于通过本发明的表达载体来表达 KRED 多肽的宿主细胞是本领域众所周知的，并且包括但不限于细菌细胞例如大肠杆菌、高加索酸奶乳杆菌、短乳杆菌、微小乳杆菌、链霉菌和鼠伤寒沙门氏菌 (Salmonella typhimurium) 细胞；真菌细胞例如酵母细胞 (例如酿酒酵母或巴斯德毕赤酵母 (Pichia pastoris)(ATCC 登录号 201178))；昆虫细胞例如果蝇 S2 和夜蛾 Sf9 细胞；动物细胞例如 CHO、COS、BHK、293 和 Bowes 黑素瘤细胞；以及植物细胞。上述宿主细胞的适合的培养基和生长条件是本领域众所周知的。

[0250] 用于表达酮还原酶的多核苷酸可通过本领域已知的各种方法被引入细胞。技术包括但不限于电穿孔、基因枪粒子轰击、脂质体介导的转染、氯化钙转染和原生质体融合。用于将多核苷酸引入细胞的各种方法对于技术人员将是明显的。

[0251] 示例性的宿主细胞是大肠杆菌 W3110。通过将编码改进的酮还原酶的多核苷酸连接至质粒 pCK110900 来产生表达载体，其中所述质粒被可操作地连接至受 lacI 抑制子 (repressor) 控制的 lac 启动子。所述表达载体还含 有 P15a 复制起点和氯霉素抗性基因。通过将细胞经过氯霉素选择来分离含有大肠杆菌 W3110 中目标多核苷酸的细胞。

[0252] 6.5 产生人工改造的酮还原酶多肽的方法

[0253] 在一些实施方案中，为了制备本公开中改进的 KRED 多核苷酸和多肽，从高加索酸奶乳杆菌或短乳杆菌或微小乳杆菌获得 (或衍生) 催化还原反应的天然存在的酮还原酶。在一些实施方案中，亲本多核苷酸序列是密码子优化的以便增强酮还原酶在特定宿主细胞中的表达。作为说明，从低聚核苷酸构建了编码高加索酸奶乳杆菌野生型 KRED 多肽的亲本多核苷酸序列，其中所述低聚核苷酸是基于 Genbank 数据库中可得的高加索酸奶乳杆菌 KRED 序列 (Genbank 登录号 AAP94029 GI：33112056) 的已知多肽序列来制备的。被指定为 SEQ ID NO：1 的亲本多核苷酸序列是密码子优化的以用于大肠杆菌中的表达，而且密码子优化的多核苷酸被克隆进表达载体，将酮还原酶基因的表达置于 lac 启动子和 lacI 抑制子基因的控制之下。表达大肠杆菌中活性酮还原酶的克隆被鉴定，而且基因被测序以确认它们的身份。指定的序列 (SEQ ID NO：1) 是被用作大多数实验起始点的亲本序列，而且人工改造的酮还原酶的文库构建是从高加索酸奶乳杆菌酮还原酶发展而来的。

[0254] 如上面讨论的，可通过将编码天然存在的酮还原酶的多核苷酸经过诱变和 / 或定向进化方法来获得人工改造的酮还原酶。示例性的定向进化技术是诱变和 / 或 DNA 重排，如 Stemmer，1994，Proc Natl Acad Sci USA91：10747-10751；WO 95/22625；WO 97/0078；WO 97/35966；WO 98/27230；WO 00/42651；WO 01/75767 和美国专利 6,537,746 中所述。其他可用的定向进化程序包括但不限于交错延伸过程 (StEP)、体外重组 (Zhao 等.，1998，Nat.Biotechnol.16：258-261)、诱变 PCR(Caldwell 等.，1994，PCR MethodsAppl.3：S136-S140) 以及盒式诱变 (Black 等.，1996，Proc Natl Acad Sci USA93：3525-3529)。可在下述参考文献中发现对于本文目的有用的另外的诱变和定向进化技术：Ling，

等.,1997,“Approaches to DNA mutagenesis:anoverview(DNA 诱变方法:综述)”,Anal.Biochem.254(2):157-78;Dale 等.,1996,“Oligonucleotide-directed random mutagenesis using the phosphorothioate method(使用磷硫酰方法的低聚核苷酸指导的随机诱变)”,Methods Mol.Biol.57:369-74;Smith,1985,“In vitro mutagenesis(体外诱变)”,Ann.Rev.Genet.19:423-462;Botstein等.,1985,“Strategies andapplications of in vitro mutagenesis(体外诱变的策略和应用)”,Science229:1193-1201;Carter,1986,“Site-directed mutagenesis(位点定向诱变)”,Biochem.J.237:1-7;Kramer等.,1984,“Point Mismatch Repair(点错配修复)”,Cell,38:879-887;Wells 等.,1985,“Cassette mutagenesis:an efficientmethod for generation of multiple mutations at defined sites(盒式诱变:在确定位点产生多重突变的有效方法)”,Gene 34:315-323;Minshull 等.,1999,“Protein evolution by molecular breeding(通过分子育种进行的蛋白进化)”,Curr Opin Chem Biol 3:284-290;Christians 等.,1999,“Directed evolution ofthymidine kinase for AZT phosphorylation using DNA family shuffling(使用 DNA 家族重排进行的用于 AZT 磷酸化的胸腺嘧啶核苷激酶定向进化)”,Nature Biotech 17:259-264;Crameri 等.,1998,“DNA shuffling of a family ofgenes from diverse species accelerates directed evolution(来自各种物种的基因家族的 DNA 重排加速定向进化)”,Nature 391:288-291;Crameri 等.,1997,“Molecular evolution of an arsenate detoxification pathway by DNA shuffling(通过 DNA 重排进行的砷酸盐解毒途径分子进化)”,Nature Biotech15:436-438;Zhang 等.,1997,“Directed evolution of an effective fructosidasefrom a galactos idase by DNA shuffling and screening(通过DNA重排和筛选进行的从半乳糖苷酶向有效的果糖苷酶的定向进化)”,Proc Natl Acad SciUSA 94:45-4-4509;Crameri等.,1996,“Improved green fluorescent protein bymolecular evolution using DNA shuffling(使用 DNA 重排通过分子进化来改进绿色荧光蛋白)”,Nature Biotech 14:315-319;以及 Stemmer,1994,“Rapidevolution of a protein in vitro by DNA shuffling(通过 DNA 重排进行的体外蛋白快速进化)”,Nature 370:389-391。所有出版物通过引用被并入本文。

[0255] 筛选在诱变处理之后获得的克隆中具有期望的改进的酶性质的人工改造的酮还原酶。当被转化为 NAD^+ 或 $NADP^+$ 时,可使用监测 NADH 或 NADPH 浓度的减少率的标准生化技术(通过吸光度或荧光的减少)来进行表达文库中的酶活性测量。在该反应中,随着酮还原酶将酮底物还原为 相应的羟基,NADH 或 NADPH 被酮还原酶消耗(氧化)。如吸光度或荧光的减少所测量的,每单位时间中 NADH 或 NADPH 浓度的减少率表示固定量的裂解产物(或从中制备的冻干粉末)中 KRED 多肽的相对(酶促)活性。当期望改进的酶性质是热稳定性时,可在将酶制剂经过确定的温度并测量热处理之后残留的酶活性的量之后测量酶活性。含有编码酮还原酶的多核苷酸的克隆然后被分离,被测序以鉴定核苷酸序列改变(如果有任何的话),并被用来在宿主细胞中表达酶。

[0256] 当已知人工改造的多肽的序列时,可根据已知的合成方法通过标准的固相方法来制备编码酶的多核苷酸。在一些实施方案中,多达约 100 个碱基的片段可被分别合成,然后连接(例如通过酶促或化学连接方法,或聚合酶介导的方法)以形成任何期望的连续序列。例如,可使用例如 Beaucage 等.,1981,Tet Lett 22:1859-69 中描述的经典的亚磷酰胺方

法或Matthes等.,1984,EMBO J.3:801-05中描述的方法(例如,如它通常在自动化的合成方法中实践的那样)通过化学合成来制备本发明的多核苷酸和低聚核苷酸。根据亚磷酰胺方法,将低聚核苷酸在例如自动化DNA合成器中合成、纯化、退火、连接并克隆进适合的载体。另外,可从各种商业来源中的任一种获得基本上任何核酸,其中所述商业来源例如The Midland CertifiedReagent Company,Midland,TX,The Great American Gene Company,Ramona,CA,ExpressGen Inc.Chicago,IL,Operon Technologies Inc.,Alameda,CA,以及许多其他的。

[0257] 可使用蛋白纯化的熟知技术中的任一种或多种从细胞和/或培养基中回收宿主细胞中表达的人工改造的酮还原酶,其中所述技术包括但不限于溶菌酶处理、超声裂解、过滤、盐析、超离心和色谱法。用于细胞裂解和从细菌如大肠杆菌中高效提取蛋白的适合溶液是可以以CelLytic B™的商标名从St.Louis MO的Sigma-Aldrich通过商业获得的。

[0258] 用于分离酮还原酶多肽的色谱技术包括但不限于反相色谱高效液相色谱、离子交换色谱、凝胶电泳和亲和色谱。用于纯化特定酶的条件将部分取决于例如净电荷、疏水性、亲水性、分子量、分子形状等因素,并且对于本领域技术人员将是明显的。

[0259] 在一些实施方案中,亲和技术可被用来分离改进的酮还原酶。对于亲和色谱纯化,可以使用特异性结合酮还原酶多肽的任何抗体。为了产生抗体,可通过注射酮还原酶多肽来对各种宿主动物进行免疫接种,所述宿主动物包括但不限于兔、小鼠、大鼠等。可通过侧链功能基团或连接至侧链功能基团的连接体将多肽连接至适合的载体例如BSA。可根据宿主物种而使用各种佐剂来增强免疫应答,所述佐剂包括但不限于弗氏佐剂(完全和不完全的)、矿物凝胶例如氢氧化铝、表面活性物质例如溶血卵磷脂、普卢兰尼克多元醇类、多聚阴离子、肽、油乳剂、钥孔血蓝蛋白、二硝基苯酚和潜在有用的人类佐剂例如BCG(卡介苗)和短小棒状杆菌(Corynebacterium parvum)。

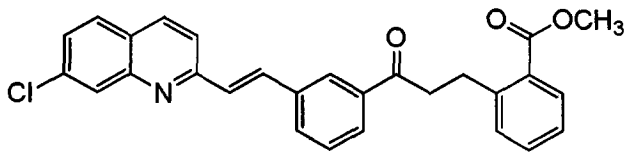
[0260] 在一些实施方案中,可以制备酮还原酶并以表达酶的细胞的形式将其用作粗提取物或者分离的或纯化的制剂。可将酮还原酶制备成冻干产物、粉末形式(例如丙酮粉末)或制备为酶溶液。在一些实施方案中,酮还原酶可以是基本上纯的制剂的形式。

[0261] 在一些实施方案中,酮还原酶多肽可被连接至固体基质。所述基质可以是固相、表面和/或膜。固体支持体可由有机多聚体(例如聚苯乙烯、聚乙烯、聚丙烯、聚氟乙烯、聚乙烯氧基和聚丙烯酰胺)及其共聚物和移植物所组成。固体支持体还可以是无机的,例如玻璃、二氧化硅、可控孔度玻璃(CPG)、反相二氧化硅或者金属例如金或铂。底物的构型可以是珠子、球形、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。固体支持体可以是多孔的或无孔的,并且可具有膨胀的或不膨胀的特征。可以以孔、凹陷或其他容器、器皿、特征或位置的形式形成固体支持体。可在阵列上对于试剂的机器人递送可寻址的各种位置上或者通过检测方法和/或仪器形成许多支持体。

[0262] 6.6 人工改造的酮还原酶和与之一起制备的化合物的使用方法

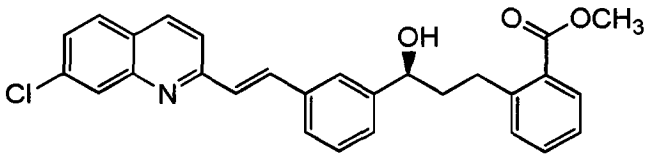
[0263] 本文所述的酮还原酶可催化结构式(I)中的底物即2-[3-[3-[2-(7-氯-2-喹啉基)乙烯基]苯基]-3-氧代丙基]苯甲酸甲酯:

[0264]

(I) 

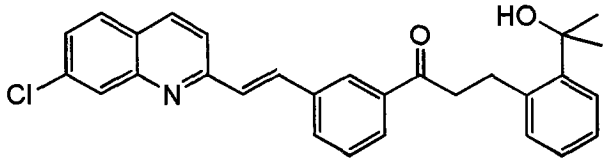
[0265] 向结构式 (II) 中相应的立体异构产物即 (S,E)-2-(3-(3-(2-(7- 氯喹啉 -2- 基) 乙烯基) 苯基)-3- 羟丙基) 苯甲酸甲酯 (“产物”) 的还原 :

[0266]

(II) 

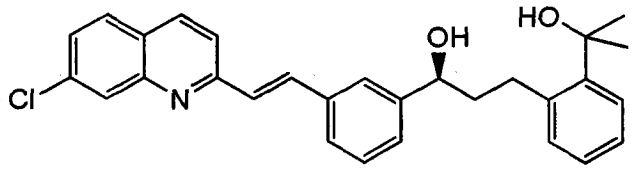
[0267] 在另一个实施方案中,本文所述的酮还原酶能够催化结构式 (III) 中的化合物即 2-[2-[3-[3[2-(7- 氯 -2- 喹啉基) 乙烯基] 苯基]-3- 氧代丙基] 苯基]-2- 丙醇 (也是 “底物”),

[0268]

(III) 

[0269] 向结构式 (IV, (S,E)-1-(3-(2-(7- 氯喹啉 -2- 基) 乙烯基) 苯基)-3-(2-(2- 羟丙 -2- 基) 苯基) 丙 -1- 醇 (也可被称作有关式 (III) 中底物的 “产物”)) 中相应的立体异构产物的还原反应 :

[0270]

(IV) 

[0271] 在一些实施方案中,本公开的酮还原酶可被用于将具有化学式 (I) 或 (III) 的底物还原为相应产物的方法中,其中所述方法包括将底物在适于将所述底物分别还原或转化为式 (II) 或式 (IV) 中产物的反应条件下与本文所述的酮还原酶多肽中任一种 (或多种) 接触或孵化。

[0272] 在一些将底物还原为产物的方法的实施方案中,与 SEQ ID NO :4、2 和 114 的野生型高加索酸奶乳杆菌或短乳杆菌或微小乳杆菌 KRED 序列相比,酮还原酶多肽至少具有下面的特征 :(1) 对应于 X190 的残基是受限残基而且 (2) 对应于 X202 的残基是芳香族残基。在一些实施方案中,与 SEQ ID NO :4、2 和 114 的野生型高加索酸奶乳杆菌或短乳杆菌或微小乳杆菌 KRED 序列相比,酮还原酶多肽至少具有下面的特征 :(1) 对应于 X190 的残基是受限残基 ;(2) 对应于 195 的残基是碱性的、酸性的或极性的残基 ;而且 (3) 对应于 X202 的残基是芳香族残基。

[0273] 在一些将底物还原为产物的方法的实施方案中,底物被还原为具有超过约 99% 立体异构过量的产物,其中所述酮还原酶多肽包括对应于 SEQ IDNO :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、111 或 112 的序列。

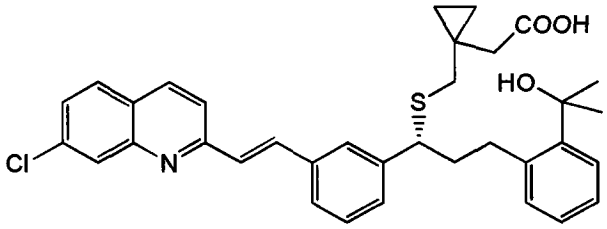
[0274] 在一些将底物还原为产物的方法的实施方案中，当用超过约 100g/L 的底物和小于约 5g/L 的多肽进行时，至少约 95％的底物在小于约 24 小时中被转化为产物，其中所述多肽包括对应于 SEQ ID NO：6、8、14、16 和 20 的氨基酸序列。

[0275] 在一些将底物还原为产物的方法的实施方案中，使用约 10g/L 的多肽在小于约 24 小时中将至少约 10-20％的 1g/L 的底物转化为产物，其中所述多肽包括对应于 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、111 或 112 的氨基酸序列。

[0276] 在一些实施方案中，本文所提供的酮还原酶多肽可被用来生产用于合成结构式 (V) 中化合物的中间产物：

[0277]

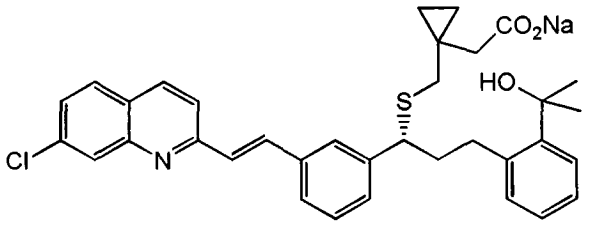
(V)



[0278] 其包括溶剂化物、水合物及其药学上可接受的盐。式 (V) 中的化合物即 2-[1-[[(1R)-1-[3-[(E)-2-(7-氯喹啉-2-基）乙烯基］苯基］-3-[2-(2-羟丙-2-基）苯基］丙基］硫甲基］环丙基］乙酸（也被称作孟鲁司特）是可以以式 (Va) 中所示的钠盐形式（即 (R,E)-2-(1-((1-(3-(2(7-氯喹啉-2-基）乙烯基）苯基)-3-(2-(2-羟丙-2-基）苯基）丙基硫）甲基）环丙基）乙酸钠）通过商业获得的：

[0279]

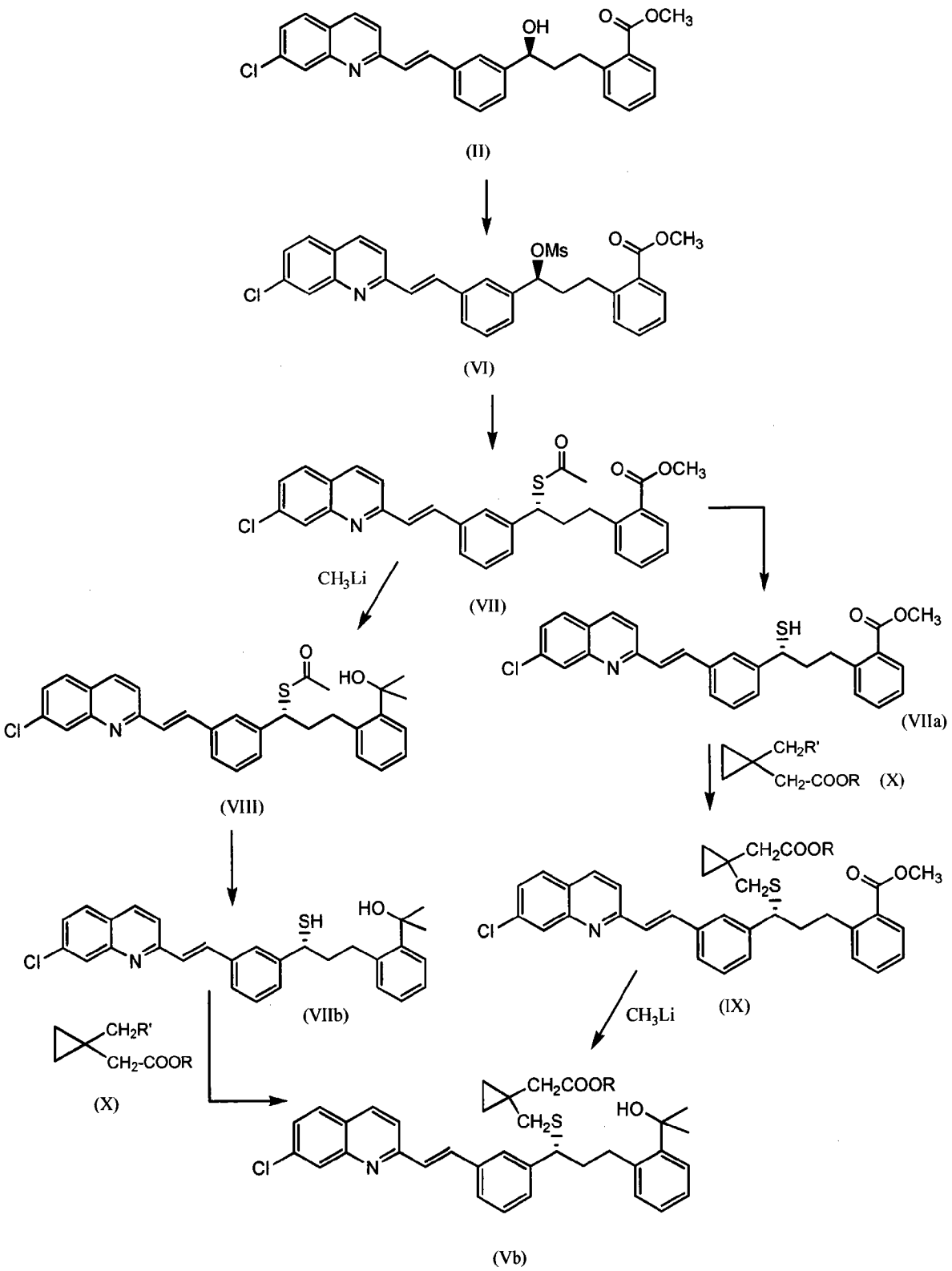
(Va)



[0280] 如上所述而不被理论所约束，孟鲁司特是白三烯受体拮抗剂 (LTRA) 并且看来似乎阻断白三烯 D4 对半胱氨酰白三烯受体 CysLT1 的作用。为了过敏症的治疗和哮喘的维持治疗而将孟鲁司特开处方（参见例如美国专利第 5,565,473 号）。在一些合成孟鲁司特的实施方案中，重要的中间产物是式 (II) 的化合物。因此，在用于合成化合物 (V) 的方法中，该方法的一个步骤可包括在适于将所述底物还原或转化为式 (II) 中产物的条件下将式 (I) 的底物与本公开的酮还原酶多肽反应或接触。

[0281] 在一些实施方案中，可以从式 (II) 中的化合物合成式 (V) 中的化合物，如方案 I 中所说明以及如美国申请公布 20070135480 和 20050245568 中所述。

[0282]



方案 1

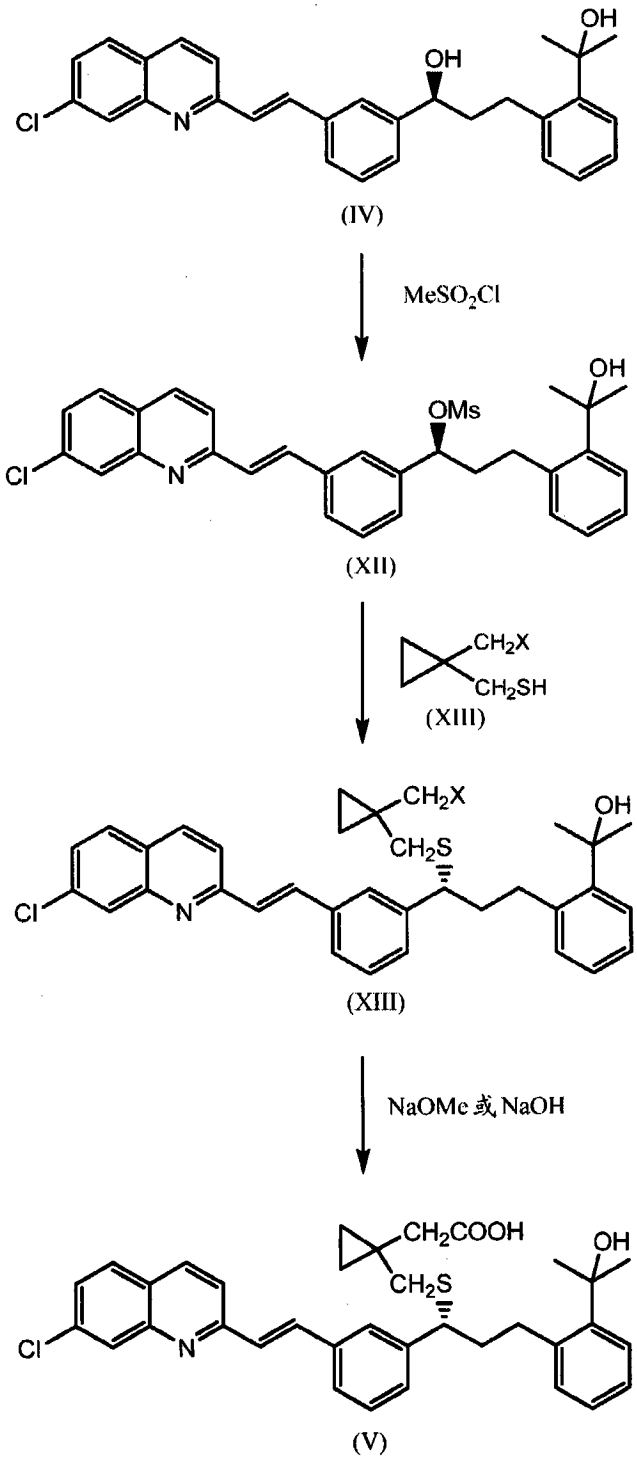
[0283] 因此,在一些实施方案中,用于合成化合物 (V) 的方法可包括 :(a) 用本公开的酮还原酶将化合物 (I) 还原或转化为化合物 (II) ;(b) 将化合物 (II) 与甲基磺酰氯反应以形成式 (VI) 中的化合物 ;(c) 将化合物 (VI) 转化成式 (VII) 中的硫代 (thio) 化合物,接着转化为式 (VIIa) 中的化合物 ;(d) 将硫代化合物 (VIIa) 与式 (X) 的化合物反应以形成式 (IX) 中的化合物,其中 R' 是离去基团而 R 是 H 或 C1-C4 烷基 ;以及 (e) 将化合物 (IX)

转化为式 (Vb) 中的化合物或其盐例如式 (Va)。在一些实施方案中，其他的烷基磺酰卤素或芳基磺酰卤素可替代甲基磺酰氯来使用。式 (IX) 中的化合物向式 (Vb) 中的化合物的转化可使用 CH_3Li 但该步骤还可使用格氏试剂例如甲基卤化镁来进行（参见美国申请公布 20070135480)。在一些实施方案中，上述步骤 (c) 之后的可选过程可包括：(d) 将式 (VII) 中的化合物转化成式 (VIII) 中的化合物，接着转化为式 (VIIb) 中的化合物，以及 (e) 将硫代化合物 (X) 与化合物 (VIIb) 反应以形成式 (Vb) 中的化合物或其盐。各步骤的细节和条件被描述在美国申请公布 20070135480 和 20050245568 中。

[0284] 在一些实施方案中，孟鲁司特的合成可使用式 (IV) 的中间产物。因此，在一些实施方案中，在用于合成式 (V) 的化合物的方法中，该方法的一个步骤可包括在适于将所述底物还原或转化为式 (VI) 中产物的条件下通过将底物与本文所公开的酮还原酶接触而将式 (III) 中的底物还原或转化为式 (IV) 中的产物。

[0285] 在一些实施方案中，可由式 (IV) 中的产物来合成孟鲁司特化合物，如方案 II 中所说明以及如美国申请公布第 2005/0234241 号和美国专利第 7,189,853 号中所述：

[0286]



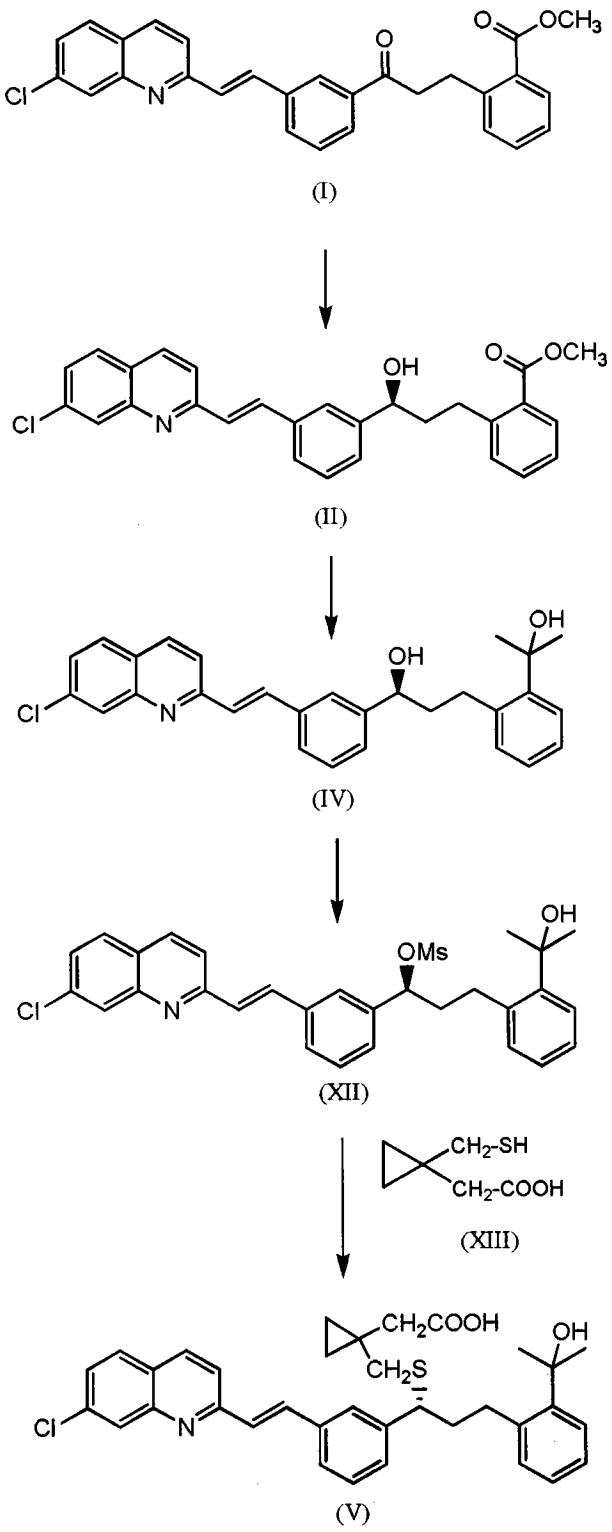
方案 2

[0287] 因此，在一些实施方案中，用于合成化合物 (V) 的方法可包括：(a) 用本公开的酮还原酶将化合物 (I) 还原或转化为化合物 (II)；(b) 将式 (IV) 中的化合物与甲基磺酰氯反应以形成式 (VI) 中的化合物；(c) 将式 (XIII) 中的巯基化合物与式 (IV) 中的化合物偶合以形成式 (VII) 中的化合物，其中 X 是 -CN 或 $CONH_2$；以及 (d) 将式 (VII) 中的化合物转化为式 (V) 中的化合物或其盐。

[0288] 在一些实施方案中，还可由式 (II) 中的产物来合成孟鲁司特化合物，如下面的方

案 III 中所说明以及如美国专利第 5,614632 号中所述，其公开通过引用被并入本文（参见美国专利 5,614,632 的实施例 5-8）。在方案 III 中，使用本公开的酮还原酶将式 (I) 中的底物还原为式 (II) 中的产物；然后通过例如使用格氏试剂（甲基氯化镁）转化为式 (IV) 中的化合物。将式 (IV) 中的化合物与甲基磺酰氯反应以形成甲磺酸盐化合物 (XIII)，虽然可以使用其他烷基磺酰卤素或芳基磺酰卤素。化合物 (XIII) 与式 (XIII) 中的巯基化合物的反应得到化合物 (V)，其可被转化为钠盐形式。

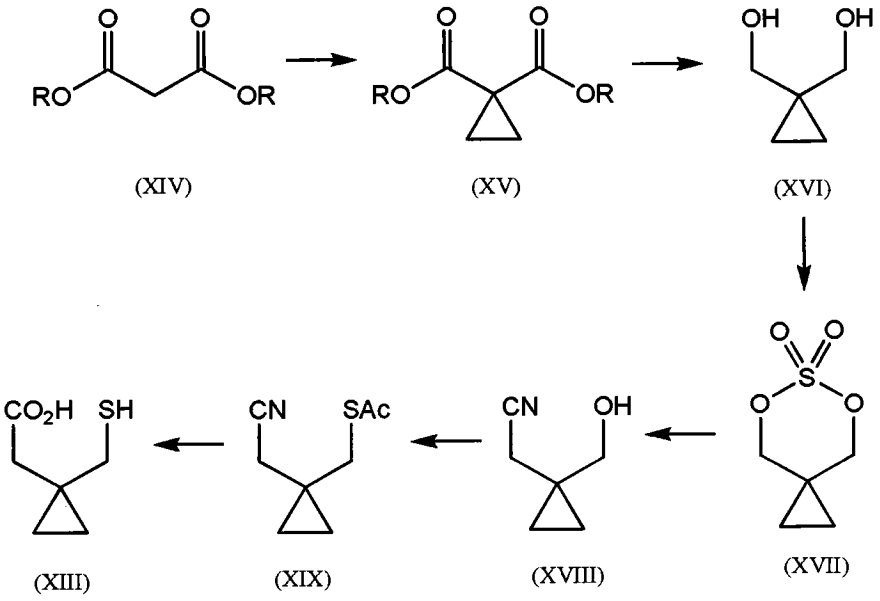
[0289]



方案 3

[0290] 可根据方案 4 来合成式 (XIII) 中的巯基化合物 (参见例如美国专利第 5,614,632 号) :

[0291]

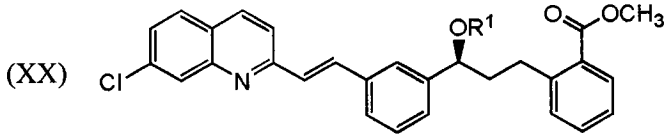


方案 4

[0292] 方案 4 中的 R 基团可以是低级烷基，优选地为 C1-C4 烷基。从式 (II) 或式 (IV) 中的化合物合成孟鲁司特的各种变化也被描述在美国专利第 5,565,473 号；美国公布 20050107612；美国公布 20050245569；美国公布 20060004204；和美国公布 20060194839 中；所有参考文献通过引用并入本文。

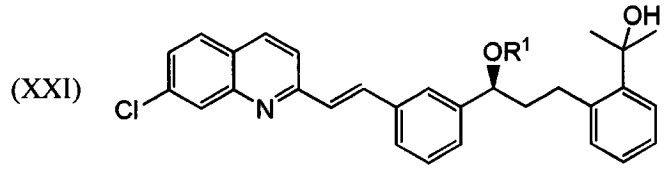
[0293] 鉴于式 (VI) 和式 (XII) 中的甲磺酸盐化合物的重要性，本公开的酮还原酶可被用于合成式 (XX) 中的化合物：

[0294]



[0295] 或者式 (XXI) 中的化合物：

[0296]



[0297] 其中 R^1 是烷基磺酰基、(全) 氟代烷基磺酰基或芳基磺酰基。在一些实施方案中，所述烷基是低级烷基例如 C1-C4 烷基，而芳基是苄基。在一些实施方案中，R^1 是甲苯磺酰基或甲磺酰基。

[0298] 在一些实施方案中，化合物 (XX) 的合成方法可包括：(a) 在适于将化合物 (I) 还原为化合物 (II) 的反应条件下将化合物 (I) 与本文所述的酮还原酶接触；以及 (b) 将化合物 (II) 与烷基磺酰基、(全) 氟代烷基磺酰基、芳基磺酰基卤素或酐反应以形成化合物 (XX)。在一些实施方案中，用于 (b) 中的烷基磺酰卤素是甲基磺酰氯，使得化合物 (XX) 中的 R^1 是甲磺酰基。

[0299] 在一些实施方案中，式 (XXI) 中的化合物的合成方法可包括：(a) 在适于将化合物 (III) 还原为化合物 (IV) 的反应条件下将化合物 (III) 与本文所述的酮还原酶接触；以及

(b) 将化合物 (VI) 与烷基磺酰基或芳基磺酰基卤素或酐反应以形成化合物 (XXI)。在一些实施方案中，用于 (b) 中的烷基磺酰卤素是甲基磺酰氯，使得化合物 (XXI) 中的 R^1 是甲磺酰基。

[0300] 可选地，在一些实施方案中，化合物 (XXI) 的合成方法可使用方案 3 中所说明的过程，其中所述方法包括：(a) 在适于将化合物 (I) 还原为化合物 (II) 的反应条件下将化合物 (I) 与本文所述的酮还原酶接触；(b) 将化合物 (II) 转化为化合物 (IV)；以及 (c) 将化合物 (IV) 与烷基磺酰基、(全) 氟代烷基磺酰基、芳基磺酰基卤素反应以形成化合物 (XXI)。如上所述，化合物 (II) 向 (b) 中的化合物 (IV) 的转化可使用格氏试剂或 CH_3Li。

[0301] 在一些实施方案中，本公开的酮还原酶可以与酮还原酶所作用的底物和 / 或由酮还原酶反应所产生的产物一起作为组合物例如反应组合物存在。因此，在一些实施方案中，组合物可包括本公开的酮还原酶以及结构式 (I) 中的底物或结构式 (III) 中的底物。在一些实施方案中，组合物可包括本公开的酮还原酶以及结构式 (II) 中的产物或结构式 (IV) 中的产物。在一些实施方案中，所述组合物可包括本公开的酮还原酶、式 (I) 中的底物和式 (II) 中的产物。在一些实施方案中，所述组合物可包括本公开的酮 还原酶、式 (III) 中的底物和式 (IV) 中的产物。

[0302] 因为酮还原酶反应可在辅助因子 (NADH 或 NADPH) 再生系统的存在下进行，所以反应条件还可包括辅助因子再生系统的元件，其在下面更详细地描述。因此，在一些实施方案中，前述酮还原酶组合物还可包括辅助因子再生系统，其包括葡萄糖脱氢酶和葡萄糖；甲酸脱氢酶和甲酸；或者异丙醇和仲醇脱氢酶。在一些实施方案中，仲醇脱氢酶是人工改造的酮还原酶。可被用于辅助因子循环的其他酶和底物将是技术人员众所周知的。

[0303] 如本领域技术人员所知，酮还原酶催化的还原反应通常需要辅助因子。由本文所述的人工改造的酮还原酶所催化的还原反应通常也需要辅助因子，虽然人工改造的酮还原酶的许多实施方案需要比野生型酮还原酶所催化的反应少得多的辅助因子。如本文所用，术语“辅助因子”是指与酮还原酶共同起作用的非蛋白化合物。适合与本文所述的人工改造的酮还原酶共同使用的辅助因子包括但不限于 $NADP^+$ (烟酰胺腺嘌呤二核苷酸磷酸)、NADPH($NADP^+$ 的还原型)、NAD^+ (烟酰胺腺嘌呤二核苷酸) 和 NADH(NAD^+ 的还原型)。通常，将辅助因子的还原型加入反应混合物。任选地，可以使用辅助因子再生系统从氧化的 $NAD(P)^+$ 形式再生还原的 NAD(P)H 形式。

[0304] 术语“辅助因子再生系统”是指参加将氧化型的辅助因子还原 (例如 $NADP^+$ 至 NADPH) 的反应的一组反应物。被酮还原酶催化的酮底物还原所氧化的辅助因子被辅助因子再生系统再生为还原型。辅助因子再生系统包括化学计量还原剂，其为还原氢等价物的来源并且能够将辅助因子的氧化型还原。辅助因子再生系统还可包括催化剂，例如为还原剂对辅助因子氧化型的还原进行催化的酶催化剂。用于从 NAD^+ 或 $NADP^+$ 分别再生 NADH 或 NADPH 的辅助因子再生系统是本领域已知的并且可被用于本文所述的方法。

[0305] 可被使用的适合的示例性辅助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、仲醇 (例如异丙醇) 和仲醇脱氢酶、亚磷酸盐和亚磷酸盐脱氢酶、分子氢 和氢化酶以及诸如此类。这些系统可与 NADP+/NADPH 或 NAD+/NADH 辅助因子共同使用。使用氢化酶的电化学再生还可被用作辅助因子再生系统。参见例如美国专利第 5,538,867 和 6,495,023 号，其二者都通过引用被并

入本文。包括金属催化剂和还原剂（例如分子氢或甲酸盐）的化学的辅助因子再生系统也是适合的。参见例如 PCT 公布 WO 2000/053731，其通过引用被并入本文。

[0306] 术语“葡萄糖脱氢酶”和“GDH”在本文中被互换地使用以指 NAD^+ 或 $NADP^+$ 依赖性酶，所述酶催化 D- 葡萄糖和 NAD^+ 或 $NADP^+$ 分别向葡萄糖酸和 NADH 或 NADPH 的转化。下面的反应式 (1) 描述葡萄糖脱氢酶催化的葡萄糖对 NAD^+ 或 $NADP^+$ 的还原。

[0307]

$$(1)\qquad \text{葡萄糖} + NAD(P)^+ + H_2O \xrightarrow{GDH} \text{葡萄糖酸} + NAD(P)H + H^+$$

[0308] 适合用于本文所述方法的实践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二者。编码天然存在的葡萄糖脱氢酶的基因已在文献中报道。例如，枯草芽孢杆菌 61297GDH 基因被表达在大肠杆菌中，并且据报道其展示了与在其天然宿主中所产生的酶相同的物理化学性质 (Vasantha 等 .，1983，Proc. Natl. Acad. Sci. USA 80 :785)。对应于 Genbank 登录号 M12276 的枯草芽孢杆菌 GDH 基因的基因序列被报道在 Lampel 等 .，1986，J. Bacteriol. 166 :238-243 中，而且其作为 Genbank 登录号 D50453 的修正形式被报道在 Yamane 等 .，1996，Microbiology142 :3047-3056 中。天然存在的 GDH 基因还包括编码蜡样芽胞杆菌 (B. cereus) ATCC 14579 (Nature，2003，423 :87-91 ；Genbank 登录号 AE017013) 和巨大芽胞杆菌 (B. megaterium) (Eur. J. Biochem.，1988，174 :485-490，Genbank 登录号 X12370 ；J. Ferment. Bioeng.，1990，70 :363-369，Genbank 登录号 GI216270) 中 GDH 的那些。来自芽孢杆菌属的葡萄糖脱氢酶作为 SEQID NO :10 和 12（分别由对应于该 PCT 公布中的 SEQ ID NO :9 和 11 多核苷酸序列所编码）被提供在 PCT 公布 WO 2005/018579 中，其公开通过引用被并入本文。

[0309] 可使用已知方法例如诱变、定向进化以及诸如此类来产生非天然存在的葡萄糖脱氢酶。可使用 PCT 公布 WO 2005/018579 中实施例 4 所描述的 测定容易地鉴定具有适当活性的 GDH 酶，不论它是天然存在的还是非天然存在的，其公开通过引用被并入本文。示例性的非天然存在的葡萄糖脱氢酶作为 SEQ ID NO :62、64、66、68、122、124 和 126 被提供在 PCT 公布 WO 2005/018579 中。编码它们的多核苷酸序列作为 SEQ ID NO :61、63、65、67、121、123 和 125 被分别提供在 PCT 公布 WO 2005/018579 中。所有这些序列通过引用被并入本文。本文所公开的适合用于酮还原酶催化的还原反应的另外的非天然存在的葡萄糖脱氢酶被提供在美国申请公布第 2005/0095619 和 2005/0153417 号中，其公开通过引用被并入本文。

[0310] 本文所述的用于酮还原酶催化的还原反应的葡萄糖脱氢酶可在 PCT 公布 WO 2005/018579 的实施例 4 中所述的测定中展示至少约 10 μmol/min/mg 而且有时候至少约 10^2 μmol/min/mg 或约 10^3 μmol/min/mg、高达约 10^4 μmol/min/mg 或更高的活性。

[0311] 本文所述的酮还原酶催化的还原反应通常在溶剂中进行。适合的溶剂包括水、有机溶剂（例如乙酸乙酯、乙酸丁酯、1- 辛醇、庚烷、辛烷、甲基叔丁基醚 (MTBE)、甲苯以及诸如此类）、离子液体（例如 1- 乙基 4- 甲基咪唑四氟硼酸盐、1- 丁基 -3- 甲基咪唑四氟硼酸盐、1- 丁基 -3- 甲基咪唑六氟磷酸盐以及诸如此类）。在一些实施方案中，使用水溶剂，其包括水和水共溶剂系统。

[0312] 示例性的水共溶剂系统具有水和一种或多种有机溶剂。通常，选择水共溶剂系统的有机溶剂组分以便它不完全地使酮还原酶失活。可利用酶活性测定例如本文所述的那

些，通过用候选溶剂系统中感兴趣的确定底物来测量指定的人工改造的酮还原酶的酶活性来容易地鉴定适当的共溶剂系统。

[0313] 水共溶剂系统的有机溶剂组分可与水组分混溶以提供单液相，或者可与水组分部分混溶或不可混溶以提供两个液相。通常，当使用水共溶剂系统时，它被选择为双相性的，水被分散在有机溶剂中或者相反。通常，当使用水共溶剂系统时，期望选择可从水相容易地分离的有机溶剂。通常，共溶剂系统中水和有机溶剂的比例通常是在有机溶剂对水约 90 ： 10 至约 10 ： 90(v/v)，以及有机溶剂对水约 80 ： 20 至约 20 ： 80(v/v) 的范围内。可 在加入反应混合物之前预先形成共溶剂系统，或者可将其在反应容器中就地 (in situ) 形成。

[0314] 水溶剂（水或水共溶剂系统）可以是 pH 缓冲的或未缓冲的。一般地，可在约 10 或以下，通常在约 5 至约 10 的范围内的 pH 中进行还原。在一些实施方案中，在约 9 或以下，通常在约 5 至约 9 的范围内的 pH 中进行还原。在一些实施方案中，在约 8 或以下，通常在约 5 至约 8 的范围内并且通常在约 6 至约 8 的范围内的 pH 中进行还原。还可在约 7.8 或以下，或者 7.5 或以下的 pH 中进行还原。可选地，可在中性 pH 即约 7 中进行还原。

[0315] 在还原反应的过程中，可改变反应混合物的 pH。可在反应过程中通过加入酸或碱而将反应混合物的 pH 维持在期望的 pH 或在期望的 pH 范围内。可选地，可使用含有缓冲液的水溶剂来控制 pH。维持期望 pH 范围的适合的缓冲液是本领域已知的并且包括例如磷酸盐缓冲液、三乙醇胺缓冲液以及诸如此类。还可使用缓冲和酸或碱添加的组合。

[0316] 当使用葡萄糖 / 葡萄糖脱氢酶辅助因子再生系统时，如果所得的葡萄糖酸水溶液不被另外中和，那么反应式 (1) 中所表示的葡萄糖酸 (pKa ＝ 3.6) 的共同产生导致反应混合物的 pH 下降。可通过标准的缓冲技术（其中所述缓冲将高达所提供的缓冲能力的葡萄糖酸中和），或者通过与转化过程同时发生的碱的加入而将反应混合物的 pH 维持在期望水平。还可使用缓冲和碱添加的组合。上面描述了维持期望 pH 范围的适合的缓冲液。用于葡萄糖酸中和的适合的碱是有机碱例如胺、醇化物和诸如此类，以及无机碱例如氢氧化物盐（例如 NaOH）、碳酸氢盐（例如 $NaHCO_3$）、碳酸盐（例如 K_2CO_3）、碱式磷酸盐（例如 K_2HPO_4、Na_3PO_4）以及诸如此类。可以在监测反应混合物 pH 时手动地，或者更方便地通过使用被用作 pH 恒定器的自动滴定器来完成与转化过程同时发生的碱的加入。还可将部分缓冲能力和碱添加的组合用于过程控制。

[0317] 当碱添加被用来中和在酮还原酶催化的还原反应过程中释放的葡萄糖酸时，可通过被添加以维持 pH 的碱的量来监测转化过程。通常，在还原过程中被加入未缓冲的或部分缓冲的反应混合物的碱被加入水溶液。

[0318] 在一些实施方案中，辅助因子再生系统可包括甲酸脱氢酶。术语“甲 酸脱氢酶”和“FDH”在本文中被互换地使用以指 NAD^+ 或 $NADP^+$ 依赖性酶，所述酶催化甲酸和 NAD^+ 或 $NADP^+$ 分别向二氧化碳和 NADH 或 NADPH 的转化。适合在本文所述的酮还原酶催化的还原反应中用作辅助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二者。甲酸脱氢酶包括对应于 PCT 公布 WO 2005/018579 中 SEQ ID NOS：70（假单胞菌属）和 72（博伊丁念珠菌 Candida boidinii）的那些，其被分别对应于 PCT 公布 2005/018579 中 SEQ ID NOS：69 和 71 的多核苷酸序列所编码，其公开通过引用被并入本文。用于本文所述方法的甲酸脱氢酶（不论是天然存在的还是非天然存在的）可展示

至少约 1 μ mol/min/mg、有时候至少约 10 μ mol/min/mg 或至少约 10^2 μ mol/min/mg、高达约 10^3 μ mol/min/mg 或更高的活性，并且可在 PCT 公布 WO2005/018579 的实施例 4 中所述的测定中容易地筛选活性。

[0319] 如本文所用，术语“甲酸 (formate)”是指甲酸阴离子 (HCO_2^-)、甲酸 (HCO_2H) 及其混合物。可以以盐（通常为碱或铵盐（例如 HCO_2Na、$KHCO_2NH_4$ 以及诸如此类））、甲酸（通常为甲酸水溶液）或其混合物的形式提供甲酸。甲酸是温和酸。在它的 pKa（水中 pKa = 3.7）的几个 pH 单位以内的水溶液中，甲酸以平衡浓度的 HCO_2^- 和 HCO_2H 二者存在。在高于约 pH 4 的 pH 值下，甲酸主要以 HCO_2^- 存在。当以甲酸 (formic acid) 的形式提供甲酸 (formate) 时，通常通过加入碱而对反应混合物进行缓冲或使其酸性减少以提供期望的 pH，其通常为约 pH 5 或以上。适用于中和甲酸的碱包括但不限于有机碱例如胺、醇化物和诸如此类，以及无机碱例如氢氧化物盐（例如 NaOH）、碳酸氢盐（例如 $NaHCO_3$）、碳酸盐（例如 K_2CO_3）、碱式磷酸盐（例如 K_2HPO_4、Na_3PO_4）以及诸如此类。

[0320] 对于高于约 pH 5 的 pH 值（在此条件下甲酸主要以 HCO_2^- 存在），下面的反应式 (2) 描述甲酸脱氢酶催化的甲酸对 NAD^+ 或 $NADP^+$ 的还原。

[0321]

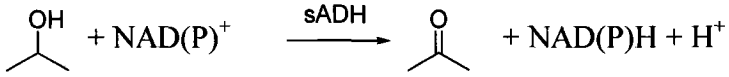
$$(2)\quad HCO_2^- + NAD(P)^+ \xrightarrow{FDH} CO_2 + NAD(P)H$$

[0322] 当使用甲酸和甲酸脱氢酶作为辅助因子再生系统时，可通过标准的缓冲技术（其中所述缓冲释放高达所提供的缓冲能力的质子），或者通过与转化过程同时发生的酸的加入而将反应混合物的 pH 维持在期望水平。在 反应过程中被加入以维持 pH 的适合的酸包括有机酸例如羧酸、磺酸、膦酸和诸如此类，无机酸例如氢卤酸（例如氢氯酸）、硫酸、磷酸和诸如此类，酸式盐例如磷酸二氢盐（例如 KH_2PO_4）、硫酸氢盐（例如 $NaHSO_4$）和诸如此类。一些实施方案利用甲酸 (formic acid)，从而甲酸 (formate) 浓度和溶液 pH 二者都得以维持。

[0323] 当利用酸添加来维持使用甲酸 / 甲酸脱氢酶辅助因子再生系统的还原反应中的 pH 时，可通过被添加以维持 pH 的酸的量来监测转化过程。通常，在转化过程中被加入未缓冲的或部分缓冲的反应混合物的酸被加入水溶液。

[0324] 术语“仲醇脱氢酶”和“sADH”在本文中被互换地使用以指 NAD^+ 或 $NADP^+$ 依赖性酶，所述酶催化仲醇和 NAD^+ 或 $NADP^+$ 分别向酮和 NADH 或 NADPH 的转化。下面的反应式 (3) 描述仲醇（用异丙醇进行说明）对 NAD^+ 或 $NADP^+$ 的还原。

[0325]

(3) 

[0326] 适合用作本文所述酮还原酶催化的还原反应中辅助因子再生系统的仲醇脱氢酶包括天然存在的仲醇脱氢酶以及非天然存在的仲醇脱氢酶二者。天然存在的仲醇脱氢酶包括来自布氏热厌氧菌 (Thermoanaerobiumbrockii)、红串红球菌 (Rhodococcus erythropolis)、高加索酸奶乳杆菌、短乳杆菌和微小乳杆菌的已知的醇脱氢酶，而非天然存在的仲醇脱氢酶包括从其衍生的人工改造的醇脱氢酶。用于本文所述方法的仲醇脱氢酶（不论是天然存在的还是非天然存在的）可展示至少约 1 μ mol/min/mg、有时候至少约 10 μ mol/min/mg 或至少约 10^2 μ mol/min/mg、高达约 10^3 μ mol/min/mg 或更高的活性。

[0327] 适合的仲醇包括低级仲链烷醇和芳基-烷基甲醇。低级仲醇的实例包括异丙醇、2-丁醇、3-甲基-2-丁醇、2-戊醇、3-戊醇、3,3-二甲基-2-丁醇以及诸如此类。在一个实施方案中，仲醇是异丙醇。适合的芳基-烷基甲醇包括未取代的和取代的1-芳基乙醇。

[0328] 当仲醇和仲醇脱氢酶被用作辅助因子再生系统时，所得的 NAD^+ 或 $NADP^+$ 被仲醇的偶合氧化所还原，其中所述偶合氧化是仲醇脱氢酶将仲醇氧化为酮。一些人工改造的酮还原酶还具有对仲醇还原剂脱氢的活性。在一些将仲醇用作还原剂的实施方案中，人工改造的酮还原酶和仲醇脱氢酶是相同的酶。

[0329] 在进行利用辅助因子再生系统的本文所述酮还原酶催化的还原反应的实施方案中，可最初提供氧化或还原型的辅助因子。如上所述，辅助因子再生系统将氧化的辅助因子转化为其还原型，其然后被用于酮还原酶底物的还原中。

[0330] 在一些实施方案中，辅助因子再生系统未被使用。对于不使用辅助因子再生系统而进行的还原反应，将还原型的辅助因子加入反应混合物。

[0331] 在一些实施方案中，当使用宿主生物体的整个细胞进行该过程时，整个细胞可天生地提供该辅助因子。可选地或共同地，该细胞可天生地或重组地提供葡萄糖脱氢酶。

[0332] 在进行本文所述立体选择性还原反应时，人工改造的酮还原酶和包括任选的辅助因子再生系统的任何酶可以以纯化的酶、用编码该酶的基因转化的整个细胞和/或这种细胞的细胞提取物和/或裂解产物的形式被加入反应混合物。编码人工改造的酮还原酶的基因和任选的辅助因子再生系统可被分别转化进宿主细胞或一起转化进相同宿主细胞。例如，在一些实施方案中，可用编码人工改造的酮还原酶的基因来转化一组宿主细胞，并用编码辅助因子再生酶的基因来转化另一组。两组转化细胞可以完整细胞形式或从其衍生的裂解产物或提取物形式一起用于反应混合物。在其他实施方案中，可用编码人工改造的酮还原酶和辅助因子再生酶的基因来转化宿主细胞。

[0333] 用编码人工改造的酮还原酶和/或任选的辅助因子再生酶的基因所转化的整个细胞或者细胞提取物和/或其裂解产物可以以各种不同形式被使用，所述形式包括固体（例如冻干的、喷雾干燥的以及诸如此类）或半固体（例如粗制的糊状物）。

[0334] 可通过沉淀作用（硫酸铵、聚乙烯亚胺、热处理或诸如此类，接着在 冻干之前进行脱盐程序（例如超滤、透析以及诸如此类））将细胞提取物或细胞裂解产物部分地纯化。可通过使用已知的交联剂例如戊二醛的交联或固定于固相（例如 Eupergit C 以及诸如此类）来稳定任何细胞制剂。

[0335] 可以以各种不同形式将固体反应物（例如酶、盐等）提供给反应，所述不同形式包括粉末（例如冻干的、喷雾干燥的以及诸如此类）、溶液、乳浊液、悬液以及诸如此类。可使用本领域普通技术人员已知的方法和设备将反应物容易地冻干或喷雾干燥。例如，可以在-80℃以小等份将蛋白溶液冷冻，然后加入预先冷却的冻干容器中，接着施加真空。在从样品除去水后，在释放真空并回收冻干的样品之前，通常将温度提高到4℃保持两小时。

[0336] 取决于期望产物的量并随着所使用的酮还原酶底物的量，用于还原反应的反应物的量通常将变化。可使用下述准则来确定待使用的酮还原酶、辅助因子和任选的辅助因子再生系统的量。通常，可以使用约50mg至约5g的酮还原酶和约10mg至约150mg的辅助因子来利用浓度为约20至300克/升的酮底物。本领域普通技术人员将容易理解如何改变这些量以使它们达到期望的产率水平和生产规模。可通过基于所用辅助因子和/或酮还原

酶的量的常规实验来容易地确定任选的辅助因子再生系统的适当的量。通常，以超过等摩尔水平的酮还原酶底物的水平来利用还原剂（例如葡萄糖、甲酸、异丙醇）以获得基本完全的或接近完全的酮还原酶底物转化。

[0337] 反应物的加入顺序不重要。可同时将反应物一起加入溶剂（例如单相溶剂、双相水共溶剂系统以及诸如此类），或可选地可将一些反应物分开加入，以及将一些在不同时间点一同加入。例如，辅助因子再生系统、辅助因子、酮还原酶和酮还原酶底物可被首先加入溶剂。

[0338] 为了在使用水共溶剂系统时增加混合效率，可首先将辅助因子再生系统、酮还原酶和辅助因子加入并混合进水相。然后可将有机相加入并混合，接着加入酮还原酶底物。可选地，在加入水相之前，可将酮还原酶底物在有机相中预先混合。

[0339] 进行本文所述的酮还原酶催化的还原反应的适合条件包括各种各样的可被常规实验容易地优化的条件，所述常规实验包括但不限于在实验 pH 和温度下将人工改造的酮还原酶和底物相接触并且使用例如本文所提供的实施例中所述的方法检测产物。

[0340] 通常在约 15℃至约 75℃范围内的温度下进行酮还原酶催化的还原。对于一些实施方案，在约 20℃至约 55℃范围内的温度下进行反应。还在其他的实施方案中，在约 20℃至约 45℃范围内的温度下进行。还可在环境条件下进行反应。

[0341] 通常允许还原反应进行到基本上完成或者接近完成，得到底物的还原。可使用已知方法通过检测底物和 / 或产物来监测底物向产物的还原。适合的方法包括气相色谱法、HPLC 以及诸如此类。反应混合物中产生的醇还原产物的转化率通常超过约 50％，也可超过约 60％，也可超过约 70％，也可超过约 80％，也可超过约 90％，并且往往超过约 97％。

7. 实施例

[0342] 在下面的代表性实例中说明本公开的各种特征和实施方案，其旨在是说明性的而不是限制性的。

[0343] 在下面的描述中，不管是否使用葡萄糖脱氢酶 (GDH)，它是可从 JulichChiral Solutions，Jülich，Germany 获得的 GDH CDX901。

[0344] 实施例 1 ：野生型酮还原酶基因的获得和表达载体的构建。

[0345] 根据报道的酮还原酶氨基酸序列和美国临时申请第 60/848,950 号（其通过引用被并入本文）的实施例 1 中所述的密码子优化算法，为了在大肠杆菌中的表达而设计了编码酮还原酶 (KRED) 的基因。使用由 42 个核苷酸所组成的低聚核苷酸合成了基因，并将其克隆进受 lac 启动子控制的表达载体 pCK110900（其被描述为美国专利申请公布 20060195947 中的图 3）。所述表达载体还含有 P15a 复制起点和氯霉素抗性基因。使用标准方法将所得的质粒转化进大肠杆菌 W3110。密码子优化的基因以及编码多肽被列在表 3 中。如美国临时申请第 60/848,950 号中所述确认了野生型酮还原酶的活性。

[0346] 表 3

[0347]

表 3：所用的酮还原酶的缩写、来源和参考					
KRED	最初从其中鉴定出	Genbank 登	GI 号	多核苷酸	多肽 SEQ

[0348]

	酶的微生物	录号		SEQ ID NO：	ID NO或来 源
ADH-CM	木兰假丝酵母	AB036927.1	12657576	美国专利申 请公布 2006019594 7中的SEQ ID No 1	美国专利申 请公布2006019594 7中的SEQ ID No 2
YDL	酿酒酵母	NP_010159. 1	6320079	SEQ ID NO:121	SEQ ID NO:122
ADH-LB	短乳杆菌	1NXQ_A	30749782	SEQ ID NO:1	SEQ ID N:2
ADH-RE	红串红球菌	AAN73270.1	34776951	SEQ ID NO:123	SEQ ID NO:124
YGL	酿酒酵母	NP_011476	6321399	SEQ ID NO:125	SEQ ID NO:126
YPR	酿酒酵母	NP_010656. 1	6320576	SEQ ID NO:127	SEQ ID NO:128
GRE	酿酒酵母	NP_014490. 1	6324421	SEQ ID NO:129	SEQ ID NO:130
ADH-LK	高加索酸奶乳杆菌	AAP94029.1	33112056	SEQ ID NO:3	SEQ ID NO:4
ADH-SB	赭色掷孢酵母	Q9UUN9	30315955	SEQ ID NO:131	SEQ ID NO:132
ADH-SC	天蓝色链霉菌	NP_631415. 1	21225636	SEQ ID NO:133	SEQ ID NO:134
ADH-TB	布氏热厌氧菌	X64841.1	1771790	SEQ ID NO:135	SEQ ID NO:136
ADH-CP	近平滑假丝酵母	BAA24528	2815409		Julich Chiral Solutions目录号03.11
DR-LB	短乳杆菌 二乙酰还原酶	ABJ63353.1	116098204		Julich Chiral Solutions目录号8.1
ADH-HE	马肝	DEHOAL	625197	SEQ ID NO:137	SEQ ID NO:138
ADH-CB	博伊丁念珠菌	CAD66648	28400789		Julich Chiral Solutions目录号02.10
LDH-LL	莱希曼氏乳杆菌(Lactobacillus leichmannii)				Fluka目录 号61306
ADH-AF	黄曲霉(Aspergillus flavus)	P41747	1168346	SEQ ID NO 115	SEQ ID NO 116
ADH-001	酒类酒球菌 (Oenococcus oeni)	ZP_00318704.1	48864831	SEQ ID NO 117	SEQ ID NO 118
ADH-RU	富养罗尔斯通氏菌(Ralstoniaeutropha)	ZP_00202558.1	46131317	SEQ ID NO 119	SEQ ID NO 120

[0349] 编码本发明中人工改造的酮还原酶的多核苷酸同样被克隆进载体pCK110900以在大肠杆菌W3110中表达。

[0350] 实施例2：酮还原酶粉末的产生；摇瓶程序。

[0351] 包括含有感兴趣的酮还原酶基因的质粒的大肠杆菌的一个单独微生物菌落被接种在50ml的含有30μg/ml氯霉素和1％葡萄糖的Terrific Broth肉汤中。在30℃培养箱中以250rpm摇动培养细胞过夜（至少16小时）。在1升的烧瓶中的250ml的Terrific Broth肉汤(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的$MgSO_4$、30μg/ml氯霉素）中将培养物稀释到在600nm(OD600)下0.2的光密度，并允许其在30℃下生长。当培养物的OD600是0.6至0.8时，用1mM的IPTG诱发酮还原酶基因的表达并培养过夜（至少16小时）。通过离心(5000rpm、15min、4℃）收集细胞并丢弃上清液。用等体积冷的(4℃)100mM三乙醇胺（氯）缓冲液pH 7.0(在ADH-LK和ADH-LB以及从其衍生的人工改造的酮还原酶的情况下包括2mM的$MgSO_4$)将细胞团块重悬浮，并如上通

过离心收集。在两体积的冷的三乙醇胺（氯）缓冲液中将被冲洗的细胞重悬浮，并在维持在4℃时在12000psi下将其通过弗氏压碎器(French Press)两次。通过离心(9000rpm、45min、4℃）除去细胞碎片。收集澄清的裂解产物上清液并储存在-20℃。对冷冻的澄清裂解产物的冻干提供了粗制酮还原酶的干燥粉末。

[0352] 实施例3：酮还原酶的产生；发酵步骤。

[0353] 在通风搅拌的15L发酵器中，含有0.88g/L硫酸铵、0.98g/L柠檬酸钠、12.5g/L三水合磷酸氢二钾、6.25g/L磷酸二氢钾、6.2g/L的Tastone-154酵母提取物、0.083g/L柠檬酸铁铵和8.3ml/L的微量元素溶液（含2g/L二水合氯化钙、2.2g/L七水合硫酸锌、0.5g/L一水合硫酸锰、1g/L七水合硫酸亚铜、0.1g/L四水合钼酸铵和0.02g/L十水合四硼酸钠）的6.0L的生长培养基被升到30℃。用含有带感兴趣的酮还原酶基因的质粒的大肠杆菌W3110的晚期指数培养物在发酵器中接种，并在实施例3中所述的摇瓶中生长至0.5至2.0的起始OD600。以500-1500rpm搅拌发酵器，并以 1.0-15.0L/min向发酵容器中提供空气以维持30％饱和或更高的溶氧水平。通过加入20％ v/v的氢氧化铵来将培养物的pH控制在7.0。通过加入含有500g/L结晶葡萄糖(cerelose)、12g/L氯化铵和10.4g/L七水合硫酸镁的进料溶液来维持培养物的生长。在培养物达到50的OD600之后，通过将异丙基-b-D-硫代半乳糖苷(IPTG)加到1mM的终浓度来诱导酮还原酶的表达。将培养物再生长14小时。然后将培养物冷却到4℃并维持在4℃直到被收集。在4℃下Sorval RC12BP离心机中以5000G离心40分钟来收集细胞。将收集的细胞直接用于接下来的下游回收过程并储存在4℃下直到如此使用。

[0354] 在4℃下在2体积的100mM三乙醇胺（氯）缓冲液(pH6.8)中将细胞团块重悬浮至每体积的湿细胞糊状物。通过将悬浮液穿过装有二阶段均质阀门组件的均质器（其使用12000psig的压力）来从细胞释放细胞内的酮还原酶。破裂之后马上将细胞匀浆冷却到4℃。10％ w/v的聚乙烯亚胺pH 7.2的溶液被加入裂解产物至终浓度0.5％ w/v，并且搅拌30分钟。通过在标准的实验室离心机上以5000G离心30分钟而使所得的悬浮液变澄清。将澄清的上清液倒出并使用具有30Kd的分子量截留(cut off)的纤维素超滤膜将其浓缩十倍。将最终的浓缩物分配到浅容器中，在-20℃下冷冻并冻干为粉末。将酮还原酶粉末储存在-80℃。

[0355] 实施例4：将2-[3-[3-[2-(7-氯-2-喹啉基）乙烯基]苯基]-3-酮丙基]苯甲酸甲酯转化为(S)-2-[3-[3-[2-(7-氯-2-喹啉基）乙烯基]苯基]-3-羟丙基]苯甲酸甲酯的分析方法。

[0356] 2-[3-[3-[2-(7-氯-2-喹啉基）乙烯基]苯基]-3-酮丙基]苯甲酸甲酯的还原和醇产物的手性纯度由如下所确定：反相HPLC(具有保护筒的2.1×50mmZorbax XDB柱子；在30℃下，0.6mL/min，H_2O中的90 ： 10的ACN/0.25％ AcOH；使用下面的保留时间以287nm进行检测：对映体醇0.55min；酮0.72min)或者正相的手性HPLC(具有保护筒的2.1×150mm ChiralPakAS-H柱子，在40℃下，0.5mL/min，10 ： 90的异丙醇/庚烷；使用下面的保留时间以287nm进行检测：(S)-醇7.1min；(R)-醇5.7min；酮4.5min)。

[0357] 实施例5：将2-[3-[3-[2-(7-氯-2-喹啉基）乙烯基]苯基]-3-酮丙基]-苯基]-1-甲基-1-乙醇转化为(S)-2-[3-[3-[2-(7-氯-2-喹啉基）乙烯基]苯基]-3-羟丙基]-苯基]-1-甲基-1-乙醇的分析方法。

[0358] 2-[3-[3-[2-(7- 氯 -2- 喹啉基) 乙烯基] 苯基]-3- 酮丙基]- 苯基]-1- 甲基 -1- 乙醇 (通过将相应的醇氧化而制备的) 的还原和醇产物的纯度由 HPLC(具有下面的洗脱图的 4.6×50mm Phenomix Gemini C18,110A 柱子 ;从 40% 至 80% 在水 /0.1% TFA 梯度持续 4min,保持 1.5min,以 0.5mL/min 的流速在 1min 内回到 40% 的乙腈 /0.1% TFA ;环境温度 ;使用下面的保留时间以 254nm 进行检测 :醇 4.4min,酮 5.7min) 所确定。

[0359] 通过 HPLC(2.1×150mm ChiralPak AD-H 柱子 ;溶剂 :40 ∶ 60 异丙醇 ∶ 庚烷 ;在 40℃下,流速 = 0.5mL/min ;使用下面的保留时间以 280nm 进行检测 :(S)- 醇 12.0min ;(R)- 醇 11.2min ;酮 10.3min) 来评估手性纯度。

[0360] 实施例 6 :对用来将 2-[3-[3-[2-(7- 氯 -2- 喹啉基) 乙烯基] 苯基]-3- 酮丙基] 苯甲酸甲酯还原的野生型酮还原酶的评估

[0361] 使用作为辅助因子的 NADH 和 NADPH 以及作为辅助因子再生系统的葡萄糖脱氢酶 / 葡萄糖或异丙醇来筛选实施例 1 的表 3 中所述的 KRED。向 96 深孔板的孔中加入 250μL 的含有 1.0g/L NAD(P)、30-50g/L 的 KRED 变体的 100mM 三乙醇胺 (氯) 缓冲液 pH7.0、200μL 的 IPA 和 50μL 的 10-20g/L 酮底物的 THF 溶液。可选地,向 96 深孔板的每孔中预先加入 10-20mg 的 $CaCO_3$、250μL 的含有 1.0g/L 的 NAD(P)、30-50g/L 的 KRED、10-20g/L 的 GDH 和 10% 葡萄糖的 100mM 三乙醇胺 (氯) 缓冲液 pH7.0。用铝 / 聚丙烯分层热密封胶带 (Velocity 11(Menlo Park, CA), 目录号 06643-001) 对板进行密封,并在室温下摇动过夜。通过加入 1000μL 的 EtOAc 而将反应猝灭。离心 (在 22℃下以 4000rpm 进行 10min) 后,使用实施例 4 中的正相 HPLC 方法来测定有机相的转化。在被测试的任何条件下这些酶中没有一个显示了任何活性 (检测极限为～ 0.5% 转化)。

[0362] 该实施例证明,野生型酮还原酶对 2-[3-[3-[2-(7- 氯 -2- 喹啉基) 乙烯基] 苯基]-3- 酮丙基] 苯甲酸甲酯具有非常少 (如果有的话) 的活性。

[0363] 实施例 7 :对用来将 2-[3-[3-[2-(7- 氯 -2- 喹啉基) 乙烯基] 苯基]-3- 酮丙基] 苯甲酸甲酯还原的 ADH-LK 变体的评估

[0364] 对已经产生的几种 ADH-LK 变体进行了评估并且发现当在实施例 6 中所述的和表 4 中所列的条件下进行评估时具有 SEQ ID NO :90 的 ADH-LK 变体将底物转化为手性的 (S)-2-[3-[3-[2-(7- 氯 -2- 喹啉基) 乙烯基] 苯基]-3- 羟丙基] 苯甲酸甲酯。

[0365] 表 4

[0366]

表 4: ADH-LK 变体的活性		
SEQ ID NO	来自 ADH-LK 的突变的数目	活性
ADH-LK	0	0
90	6	<0.03 g/L.g 酶.天

[0367] 该实施例显示,含有 G7H、S96T、F147L、Y190P、V196L、A202W 突变的 ADH-LK 变体以高度的立体选择性 (＞ 99% 的立体异构过量) 将 2-[3-[3-[2-(7- 氯 -2- 喹啉基) 乙烯基] 苯基]-3- 酮丙基] 苯甲酸甲酯转化为 (S)-2-[3-[3-[2-(7- 氯 -2- 喹啉基) 乙烯基] 苯基]-3- 羟丙基] 苯甲酸甲酯。

[0368] 实施例 8 :利用异丙醇来再生辅助因子的用来测定酮还原酶对 2-[3-[3-[2-(7- 氯 -2- 喹啉基) 乙烯基] 苯基]-3- 酮丙基] 苯甲酸甲酯的活性的高通量

HPLC 测定

[0369] 通过定向进化获得并含有进化的酮还原酶基因的质粒文库被转化进大肠杆菌中并被铺板在含有 1％葡萄糖和 30μg/mL 氯霉素 (CAM) 的 Luria-Bertani(LB) 肉汤上。在 30℃下孵化至少 16 小时后，使用 Q-bot ®机器人菌落采集器 (Genetix USA，Inc.，Beaverton，OR) 将菌落采集进含有 180μL 的 Terrific 肉汤 (TB)、1％葡萄糖、30μg/mL 氯霉素 (CAM) 和 2mM 的 $MgSO_4$ 的 96 孔浅孔微量滴定板中。在 30℃下以 200rpm 摇动使细胞生长过夜。然后将 20μL 的该培养物转移进含有 350μL 的 Terrific 肉汤 (TB)、2mM 的 $MgSO_4$ 和 30μg/mL 的 CAM 的 96 深孔板中。在 30℃下深孔板中以 250rpm 摇动孵化 2.5 至 3 小时后 (OD_{600} 0.6-0.8)，细胞培养物对重组基因的表达被异丙基硫代半乳糖苷 (IPTG) 诱导至 1mM 的终浓度。然后在 30℃下以 250rpm 摇动板孵化 15-23 小时。

[0370] 向 96 孔微量滴定板的孔中加入 100μL 的含有 0.2mM Na-NADP 和 2 mM 的 $MgSO_4$ 的 100mM 三乙醇胺（氯）缓冲液 pH7.0，以及足够体积的细胞培养物以使最熟知的酶将在任何时间点上（例如在第一轮进化中的 SEQ ID：90）给出酮向 (S)- 醇的 10-20％的转化。IPA 和含有 1.67 至 16.7g/L 的 2-[3-[3-[2-(7- 氯 -2- 喹啉基）乙烯基］苯基］-3- 酮丙基］苯甲酸甲酯的 THF 溶液被加入孔中以得到体积比为 4 ： 1 ： 5 的 100mM 三乙醇胺（氯）缓冲液 pH7.0、THF 和 IPA 的最终组合物（最终的底物浓度为 1 至 10g/L）。随意地将酮加到终浓度为 0 至 2％（通常作为 IPA 或 THF 中的溶液）。如实施例 6 中所述将板热密封并在室温或 40℃下摇动孵化 4 或 24 小时。在反应的最后，用 1000μL 的 1 ： 4 的 THF ：乙腈将底物和产物猝灭。在室温下离心之后，使用实施例 4 中所述的反相 HPLC 方法分析上清液的转化。对于待测定对映选择性的反应，用 1mL 的 EtOAc 将反应猝灭，离心并通过实施例 4 中所述的正相 HPLC 方法对有机相进行手性分析。

[0371] 该实施例描述被用来鉴定针对 2-[3-[3-[2-(7- 氯 -2- 喹啉基）乙烯基］苯基］-3- 酮丙基］苯甲酸甲酯的还原而被改进的 KRED 变体的方法。

[0373] 实施例 9：衍生自 ADH-LK 的人工改造的酮还原酶对 2-[3-[3-[2-(7- 氯 -2- 喹啉基）乙烯基］苯基］-3- 酮丙基］苯甲酸甲酯的还原。

[0374] 在空气中随着搅拌棒的搅拌向琥珀色小瓶中按顺序加入 50-500mg 的底物、1.5mL 的含有 0.33g/L 的 NADP-Na 和 10.0-33.3g/L 的 KRED 变体、2mM 的 $MgSO_4$ 的 100mM 三乙醇胺（氯）缓冲液 pH 7.0 或 8.0、2.5mL 的 IPA 和 0.5mL 的共溶剂 (THF 或甲苯)。小瓶被封闭并被浸入被控制在期望温度下的油浴中。在 1、2、4 和 8 小时处取出两个 50μL 等份，并在 t ＝ 24h 处取出四个 50μL 等份。用 1000μL 的 THF 稀释所述等份，并离心 (22℃下，4000rpm，10 分钟）以除去碎屑。用乙腈将上清液稀释 9×，并如实施例 4 中所述确定转化。在第 24 小时的样品上进行手性分析：加入 0.5 体积的盐水和 2 体积的 THF 以获得清晰的相分离 (phase break)，并用 $MgSO_4$ 将有机相干燥。用 EtOAc 将有机相稀释 10×，并如实施例 4 中所述来分析。表 5 给出对应于酮还原酶的 SEQ ID NO.、来自野生型 ADH-LK 的氨基酸突变的数目和 2-[3-[3-[2-(7- 氯 -2- 喹啉基）乙烯基］苯基］-3- 酮丙基］苯甲酸甲酯向 (S)-2-[3-[3-[2-(7- 氯 -2- 喹啉基）乙烯基］苯基］-3- 羟丙基］苯甲酸 甲酯的转化。被测的所有变体给出＞ 99.2％立体异构过量的 (S)-2-[3-[3-[2-(7- 氯 -2- 喹啉基）乙烯基］苯基］-3- 羟丙基］苯甲酸甲酯（参考标准是 99.2％ ee；不期望的立体异构体是不可检测的）。

[0375]

表 5: 在标明的温度下 ADH-LK 变体的活性			
SEQ ID NO	ADH-LK 突变的数目	活性	温度（°C）
90	6	+	室温
48	6	++	室温
46	6	++	室温
44	6	++	室温
42	6	++	室温
100	7	++	室温
92	7	++	室温
94	7	++	室温
102	7	++	室温
86	7	++	室温
96	7	++	室温
88	8	++	室温
98	8	+++	室温
70	10	+++	室温
84	10	+++	室温
76	11	+++	室温
74	12	+++	室温
64	12	+++	室温
104	9	+++	室温
68	10	+++	室温
106	9	+++	室温
72	12	+++	室温
66	11	+++	室温
24	15	++++	室温
82	11	+++	35
80	12	+++	室温
108	14	+++	室温
50	14	+++	室温
78	13	+++	35
10	14	+++	室温
34	14	+++	室温
36	12	+++	室温
58	14	++++	35
40	13	++++	40
32	14	+++	35
22	15	++++	40
38	14	++++	40

[0372]

54	15	++++	40
12	16	++++	40
30	15	++++	40
28	15	++++	40
16	18	+++++	45

110	16	++++	40
62	15	++++	40
60	15	++++	40
52	15	++++	40
56	14	++++	40
26	17	++++	40
14	18	+++++	40
20	19	+++++	45
6	20	+++++	40
8	21	+++++	45
18	20	++++	40

[0373] +：＜ 0.03g 产物 / 升 /g 酶 / 天 ;++ :0.03 至 1g 产物 / 升 /g 酶 / 天 ;+++ :1 至 5g 产物 / 升 /g 酶 / 天 ;++++ :5 至 20g 产物 / 升 /g 酶 / 天 ;+++++ :＞ 20g 产物 / 升 /g 酶 / 天。

[0374] 实施例 10 :衍生自 ADH-LK 的人工改造的酮还原酶对 2-[2-[3-[3[2-(7- 氯 -2- 喹啉基) 乙烯基] 苯基]-3- 氧代丙基] 苯基]-2- 丙醇的还原。

[0375] 在空气中随着搅拌棒的搅拌向琥珀色小瓶中按顺序加入 500mg 的 2-[2-[3-[3[2-(7- 氯 -2- 喹啉基) 乙烯基] 苯基]-3- 氧代丙基] 苯基]-2- 丙醇、1.5mL 的含有 0.89g/L 的 NADP-Na 和 13.3g/L 的 KRED 变体 SEQ ID NO :20、2mM 的 $MgSO_4$ 的 100mM 三乙醇胺 (氯) 缓冲液 pH 7.0、2.5mL 的 IPA 和 0.5mL 的共溶剂 (THF 或甲苯)。小瓶被封闭并被浸入被控制在 40℃下的油浴中。在 6、24 和 29 小时处取出样品 (30μL), 并用 1500μL 的 THF 稀释 ;30μL 稀释的样品被浓缩干燥, 在如实施例 5 中所述的化学转化的 HPLC 分析之前溶解在 1000μL 乙腈中。

[0376] 在 29h 处建立整个反应混合物之后进行手性分析 :用 2mL 盐水和 6mL 的 THF 稀释样品而得到清晰的相分离。用乙酸乙酯 (3×10mL) 萃取水相, 并在 $MgSO_4$ 上将合并的有机层干燥。将有机溶液滤过西莱特 (Celite) 床并浓缩干燥。通过将 1mg 的残基溶解在 1ml 的乙腈中来制备分析样品, 并如实施例 5 中所述进行分析。以＞ 99％的立体异构过量获得 (S)-1-(3-(2-(7- 氯喹啉 -2- 基) 乙烯基) 苯基)-3-(2-(2- 羟丙 -2- 基) 苯基) 丙 -1- 醇。转化超过约 67％。

[0377] 实施例 11 :(S)-2-[3-[3-[2-(7- 氯 -2- 喹啉基) 乙烯基] 苯基]-3- 羟丙基] 苯甲酸甲酯的制备性规模生产。

[0378] 在氮气中向通过外部热交换器将套温度设置为 51℃的装备有内部温度计和机械搅拌子 (250rpm) 的带套的 2L 三颈烧瓶中加入 300mL 的 100mM 的 pH 8.0 的三乙醇胺 *HCl、0.6mL 的 1M 的 $MgSO_4$、如实施例 3 中所制备的 3.0g 的 SEQ ID NO :20 的催化剂、0.1g 的 $NADP\text{-}Na_2$ 以得到淡黄色溶液。搅拌 10 分钟之后, 在～ 5 分钟内分批加入 100g 的 2-[3-[3-[2-(7- 氯 -2- 喹啉基) 乙烯基] 苯基]-3- 酮丙基] 苯甲酸甲酯。向浆体加入 500mL(390g) 的 IPA 和 100mL(85g) 的甲苯。如实施例 4 和 6 中所述通过手性 HPLC 监测反应过程。实施例 9 中所述的过程被用作取样和样品制备。

[0379] 在 45℃下搅拌 24 小时后, 根据手性 HPLC 分析来表明反应完成。通过空气驱动的真空泵将反应混合物滤过三层的 150mm 直径的 Whatman#2 滤纸。需要大约 10 分钟来完成过滤以得到褐色糊状物。用 2×500mL 水冲洗褐色糊状物。将褐色糊状物过夜风干以得到 104g 的褐色固体 (S)-2-[3-[3-[2-(7- 氯 -2- 喹啉基) 乙烯基] 苯基]-3- 羟丙基] 苯甲酸

甲酯一水合物（100％产率；产物是一水合物），其化学纯度是98.5HPLC面积百分比（起始物质的主要杂质是0.3-0.4HPLC面积百分比）。(R)-立体异构体是未检测到的。

[0380] 本申请中所引用的所有出版物、专利、专利申请和其他文件为所有目的通过引用以其整体并入本文，达到如同单个出版物、专利、专利申请或其他文件单独表明为了所有目的而通过引用被并入的程度。

序列表

<110> 杰克·梁
波蒂·博勒普
韦丝娜·米切尔
艾米丽·穆德弗
詹姆士·拉隆德
吉伽特·哈思曼

<120> 酮还原酶多肽及其用途

<130>376247-018

<141>2008-09-26

<150>60/976,345
<150>2007-09-28

<160>139

<210>1
<211>762
<212>DNA
<213> 人工序列

<220>
<223> 密码子优化的短乳杆菌 (L.Brevis) 序列

<400>1
atgtctaacc gtctggatgg caaagtagcc atcattaccg gcgggactct gggtatcggt 60
ttggcaatcg ccacgaaatt tgtagaggag ggtgcgaaag taatgattac tggtcgtcac 120
tccgatgtag gtgaaaaggc cgccaaatca gtaggcactc cggatcagat tcagtttttt 180
cagcacgatt catccgatga agatggctgg acgaaactgt tcgacgccac cgagaaagca 240
ttcggcccgg ttagcacctt agtgaacaat gcagggattg cagttaacaa aagcgttgaa 300
gaaactacca cggccgaatg gcgtaaactg ctggccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggtt cgtaggcgat ccgagcctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggctat atcaagaccc cgctggtcga tgatctgccg 600

ggtgctgagg aagcgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catatatctg tgtgtacctg gcatctaatg aatcgaaatt tgcgacgggt 720
tccgaatttg tggtcgacgg cgggtatacc gcacagtaat ga 762

<210>2
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 密码子优化的短乳杆菌序列的翻译

<400>2

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

Tyr Ile Cys Val Tyr Leu Ala Ser Asn Glu Ser Lys Phe Ala Thr Gly
225 230 235 240

Ser Glu Phe Val Val Asp Gly Gly Tyr Thr Ala Gln
245 250

<210>3
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 密码子优化的高加索酸奶乳杆菌 (L.Kefir) 序列

<400>3
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggactct gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cagtttccaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggtt cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggctat atcaagaccc cgctggtcga tgatctggaa 600
ggtgctgagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>4
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 密码子优化的高加索酸奶乳杆菌序列的翻译

<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 Ala Val Ser
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
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>5
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>5

atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacaat gggtatcggt 60
ttggcaatcg ccaccaaata cgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcgaca ctgatgttat tcgctttgtc 180
cagcacgatg tctccgatga agcaggctgg acgaaactgt tcgacaccac cgaggagacc 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttaaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgcc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc agcggctcga tgagctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccgcat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg tgggtatacc gcacagtga 759

<210>6
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>6

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

180 185 190

Thr Gln Arg Leu Asp Glu Leu Glu Gly Trp Glu Glu Met Met Ser Gln

195 200 205

Arg Thr Val Thr Pro Met Gly Arg 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>7

<211>759

<212>DNA

<213> 人工序列

<220>

<223> 高加索酸奶乳杆菌的变体

<400>7

atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacaat gggtatcggt 60

ttggcaatcg ccaccaaata cgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120

gcggatgtag gtgaaaaggc cgccaaatca atcggcgaca ctgatgttat tcgctttgtc 180

cagcacgatg tctccgatga agcaggctgg acgaaactgt tcgacgccac cgaggagacc 240

ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttaaa 300

gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360

ggcacccgcc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420

atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480

ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540

gtgcgtgtca acacagtaca tccgggcccc atcaagaccc agcggctcga tgagctggaa 600

ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccgcat tggcgaaccg 660

aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720

gcagaatttg tggtcgacgg tgggtatacc gcacagtga 759

<210>8
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>8

```
Met Thr Asp Arg Leu Lys His Lys Val Ala Ile Val Thr Gly Gly Thr
1               5                  10                  15
Met Gly Ile Gly Leu Ala Ile Ala Thr Lys Tyr 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 Asp 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 Ala Thr Glu Glu Thr
65                  70                  75                  80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ser Cys Val Leu
                85                  90                  95
Lys Ser Val Lys 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 Leu 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 Pro Ile Lys
            180                 185                 190
Thr Gln Arg Leu Asp Glu Leu Glu Gly Trp Glu Glu Met Met Ser Gln
        195                 200                 205
Arg Thr Val Thr Pro Met Gly Arg 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>9
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>9
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacaat gggtatcggt 60
ttggcaatcg ccgataaata cgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggagacc 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtctttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggtt ggttggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtatctccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>10
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

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

Met Gly Ile Gly Leu Ala Ile Ala Asp Lys Tyr 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 Thr
65 70 75 80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ser Cys 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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Pro Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val Ser 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>11
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>11

atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacaat gggtatcggt 60
ttggcaatcg ccgataaata cgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg tctccgatga agcaggctgg acgaaactgt tcgacaccac cgaggagacc 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttaaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg cgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtagg gtaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>12
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体
<400>12

Met Thr Asp Arg Leu Lys His Lys Val Ala Ile Val Thr Gly Gly Thr
1 5 10 15
Met Gly Ile Gly Leu Ala Ile Ala Asp Lys Tyr 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 Val
50 55 60
Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Thr
65 70 75 80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ser Cys Val Leu
85 90 95
Lys Ser Val Lys 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 Leu 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 Pro Ile Lys
180 185 190
Thr Pro Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Arg Val 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>13
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体
<400>13
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacaat gggtatcggt 60
ttggcaatcg ccgataaata cgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg tctccgatga agcaggctgg acgaaactgt tcgacaccac cgaggagacc 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttaaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgcc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc agcggctcga tcagctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccgcat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg tgggtatacc gcacagtga 759

<210>14

<211>252

<212>PRT

<213> 人工序列

<220>

<223> 高加索酸奶乳杆菌的变体

<400>14

Met Thr Asp Arg Leu Lys His Lys Val Ala Ile Val Thr Gly Gly Thr
1 5 10 15
Met Gly Ile Gly Leu Ala Ile Ala Asp Lys Tyr 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 Val
50 55 60
Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Thr
65 70 75 80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ser Cys Val Leu
85 90 95
Lys Ser Val Lys 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 Leu 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 Pro Ile Lys
180 185 190
Thr Gln Arg Leu Asp Gln Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val Thr Pro Met Gly Arg 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>15
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>15
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacaat gggtatcggt 60
ttggcaatcg ccgataaata cgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg tctccgatga agcaggctgg acgaaactgt tcgacaccac cgaggagacc 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttaaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgcc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc agcggctcga tgagctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccgcat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg tgggtatacc gcacagtga 759

<210>16
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>16
Met Thr Asp Arg Leu Lys His Lys Val Ala Ile Val Thr Gly Gly Thr
1 5 10 15
Met Gly Ile Gly Leu Ala Ile Ala Asp Lys Tyr 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 Val
50 55 60
Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Thr
65 70 75 80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ser Cys Val Leu
85 90 95
Lys Ser Val Lys 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 Leu 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 Pro Ile Lys
180 185 190
Thr Gln Arg Leu Asp Glu Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val Thr Pro Met Gly Arg 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>17
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>17
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacaat gggtatcggt 60

ttggcaatcg ccgataaata cgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg tctccgatga agcaggctgg acgaaactgt tcgacaccac cgaggagacc 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttaaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgcc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc agcggctcga tgagctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtagtgccta tgggccgcat tggcgaaccg 660
aatgacgtcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg tgggtatacc gcacagtga 759

<210>18
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>18
Met Thr Asp Arg Leu Lys His Lys Val Ala Ile Val Thr Gly Gly Thr
1 5 10 15
Met Gly Ile Gly Leu Ala Ile Ala Asp Lys Tyr 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 IleArg Phe Val Gln His Asp Val
50 55 60
Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Thr
65 70 75 80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ser Cys Val Leu
85 90 95
Lys Ser Val Lys 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 Leu 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 Pro Ile Lys
180 185 190
Thr Gln Arg Leu Asp Glu Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val Val Pro Met Gly Arg Ile Gly Glu Pro Asn Asp Val 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>19
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>19
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacaat gggtatcggt 60
ttggcaatcg ccgataaata cgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg tctccgatga agcaggctgg acgaaactgt tcgacaccac cgaggagacc 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttaaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgcc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgtccctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc agcggctcga tcagctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccgcat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg tgggtatacc gcacagtga 759

<210>20
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>20

```
Met Thr Asp Arg Leu Lys His Lys Val Ala Ile Val Thr Gly Gly Thr
1               5                  10                  15
Met Gly Ile Gly Leu Ala Ile Ala Asp Lys Tyr 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 Val
    50                  55                  60
Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Thr
65                  70                  75                  80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ser Cys Val Leu
                85                  90                  95
Lys Ser Val Lys 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 Leu Val Gly Asp Pro Ser 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 Pro Ile Lys
            180                 185                 190
Thr Gln Arg Leu Asp Gln Leu Glu Gly Trp Glu Glu Met Met Ser Gln
        195                 200                 205
Arg Thr Val Thr Pro Met Gly Arg 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>21
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>21
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacaat gggtatcggt 60
ttggcaatcg ccgataaata cgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg tctccgatga agcaggctgg acgaaactgt tcgacaccac cgaggagacc 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttaaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcataca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>22
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>22
Met Thr Asp Arg Leu Lys His Lys Val Ala Ile Val Thr Gly Gly Thr
1 5 10 15
Met Gly Ile Gly Leu Ala Ile Ala Asp Lys Tyr 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 Val
50 55 60
Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Thr
65 70 75 80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ser Cys Val Leu
85 90 95
Lys Ser Val Lys 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 Leu 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 Pro Ile Lys
180 185 190
Thr Pro Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val 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>23
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>23
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacaat gggtatcggt 60

ttggcaatcg ccgataaata cgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg tctccgatga agcaggctgg acgaaactgt tcgacaccac cgaggagacc 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttaaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>24
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>24
Met Thr Asp Arg Leu Lys His Lys Val Ala Ile Val Thr Gly Gly Thr
1 5 10 15
Met Gly Ile Gly Leu Ala Ile Ala Asp Lys Tyr 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 Val
50 55 60
Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Thr
65 70 75 80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ser Cys Val Leu
85 90 95
Lys Ser Val Lys Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
100 105 110
ValAsn 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 Leu 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 Pro Ile Lys
180 185 190
Thr Pro Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val 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>25
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>25
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacaat gggtatcggt 60
ttggcaatcg ccgataaata cgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg tctccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttaaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc agcggctcga tgagctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccgcat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg tgggtatacc gcacagtga 759

<210>26
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>26

```
Met Thr Asp Arg Leu Lys His Lys Val Ala Ile Val Thr Gly Gly Thr
1               5                  10                  15
Met Gly Ile Gly Leu Ala Ile Ala Asp Lys Tyr 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 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 Ser Cys Val Leu
                85                  90                  95
Lys Ser Val Lys 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 Leu 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 Pro Ile Lys
            180                 185                 190
Thr Gln Arg Leu Asp Glu Leu Glu Gly Trp Glu Glu Met Met Ser Gln
        195                 200                 205
Arg Thr Val Thr Pro Met Gly Arg 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>27
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>27
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacaat gggtatcggt 60
ttggcaatcg ccgataaata cgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg tctccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgcc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc agcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccgcat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>28
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>28
Met Thr Asp Arg Leu Lys His Lys Val Ala Ile Val Thr Gly Gly Thr
1 5 10 15
Met Gly Ile Gly Leu Ala Ile Ala Asp Lys Tyr 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 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 Ser Cys 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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Gln Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val Thr Pro Met Gly Arg 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>29
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>29
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacaat gggtatcggt 60

ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg tctccgatga agcaggctgg acgaaactgt tcgacaccac cgaggagacc 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttaaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccgcat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg tgggtatacc gcacagtga 759

<210>30
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体
<400>30

Met Thr Asp Arg Leu Lys His 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 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 Val
50 55 60
Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Thr
65 70 75 80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ser Cys Val Leu
85 90 95
Lys Ser Val Lys 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 Leu 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 Pro Ile Lys
180 185 190
Thr Pro Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val Thr Pro Met Gly Arg 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>31
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体
<400>31
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacaat gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttgt gcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggagacc 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcgattc gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaattcg tggtcgacgg cgggtatacc gcacagtga 759

<210>32
<211>252

<212>PRT

<213> 人工序列

<220>

<223> 高加索酸奶乳杆菌的变体

<400>32

Met Thr Asp Arg Leu Lys His 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 His Ala Asp Val Gly Glu Lys Ala Ala
35 40 45
Lys Ser Ile Gly Gly Thr Asp Val Val 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 Thr
65 70 75 80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ser Cys 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
Glu Gly Leu Val Gly Asp Pro Thr Leu Gly Ala Tyr Asn Ala Ser Lys
145 150 155 160
Gly Ala Ile 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 Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val 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>33
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>33
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacaat gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttgt gcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggtt gattggcgat ccgacgctgg gggcgtacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggttt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg ttaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>34
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>34
Met Thr Asp Arg Leu Lys His 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 His Ala Asp Val Gly Glu Lys Ala Ala
35 40 45

Lys Ser Ile Gly Gly Thr Asp Val Val 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 Ser Cys 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
Glu Gly Leu Ile 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 Val 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 Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Leu 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>35
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>35
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacaat gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttgt gcgctttgtc 180

cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggtt ggttggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>36
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>36

Met	Thr	Asp	Arg	Leu	Lys	His	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	His	Ala	Asp	Val	Gly	Glu	Lys	Ala	Ala
		35					40					45			
Lys	Ser	Ile	Gly	Gly	Thr	Asp	Val	Val	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	Ser	Cys	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				
Glu	Gly	Leu	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 Pro Ile Lys
180 185 190
Thr Pro Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val 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>37
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>37
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgaaga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttattaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaacta ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgcc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc agcggctcga tgagctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccgcat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg tgggtatacc gcacagtga 759

<210>38
<211>252

<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>38

```
Met Thr Asp Arg Leu Lys His 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 Glu 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 Ser Cys Val Ile
                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
Glu Gly Leu 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 Pro Ile Lys
            180                 185                 190
Thr Gln Arg Leu Asp Glu Leu Glu Gly Trp Glu Glu Met Met Ser Gln
        195                 200                 205
Arg Thr Val Thr Pro Met Gly Arg 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>39
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>39
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgtcat tcgctttgtc 180
cagcacgatg cgtccgaaga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttattaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggtccc atcaagaccc agcggctcga tgagctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcaacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>40
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>40
Met Thr Asp Arg Leu Lys His 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 Glu 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 Ser Cys Val Ile
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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Gln Arg Leu Asp Glu Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val 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>41
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>41
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180

cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cagttattaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>42
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>42
Met Thr Asp Arg Leu Lys His 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 4 045
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 Ile
85 90 95
Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
100 105 110
Val A5n 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 Leu 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 Pro Ile Lys
180 185 190
Thr Pro Leu Leu Asp Asp Leu Glu Gly Trp 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>43
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>43
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cagttgtgaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>44
<211>252

<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>44

```
Met Thr Asp Arg Leu Lys His 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 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
Glu Gly Leu 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 Pro Ile Lys
            180                 185                 190
Thr Pro Leu Leu Asp Asp Leu Glu Gly Trp 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>45
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>45
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cagttttaaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>46
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>46
Met Thr Asp Arg Leu Lys His 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 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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Pro Leu Leu Asp Asp Leu Glu Gly Trp 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 ValTyr 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>47
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>47
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180

cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gagtttccaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>48
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体
<400>48

Met Thr Asp Arg Leu Lys His 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 Ser
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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Pro Leu Leu Asp Asp Leu Glu Gly Trp 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>49
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体
<400>49
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg catccgatga agtgggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttgaa 300
gacactacca cggaggaatg gcataaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtaacatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagacct ctcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg atcaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaattcg tggtcgacgg cgggtatacc gcacagtga 759

<210>50
<211>252
<212>PRT
<213> 人工序列

<220>

<223> 高加索酸奶乳杆菌的变体

<400>50

Met Thr Asp Arg Leu Lys His 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 LysAla 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 Val 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 Ser Cys Val Leu
85 90 95
Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp His 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 Leu Val Gly Asp Pro Thr Leu Gly Ala Tyr Asn Ala Ser Lys
145 150 155 160
Gly Ala Val Arg Asn 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 Ser Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Ile 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>51

<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>51
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgaaga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgagattaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgcc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc agcggctcga tgagctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccgcat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg tgggtatacc gcacagtga 759

<210>52
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>52
Met Thr Asp Arg Leu Lys His 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 Glu 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 Ser Cys Glu Ile
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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Gln Arg Leu Asp Glu Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val Thr Pro Met Gly Arg 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>53
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>53
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgaaga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttattaa aagcgttgaa 300

gacactacca cggaggaatg gcgtaaactg ctatccgtta atctggatgg tgtttttttc 360
ggcacccgcc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat gcgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc agcggctcga tgagctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccgcat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg tgggtatacc gcacagtga 759

<210>54
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>54
Met Thr Asp Arg Leu Lys His 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 Glu 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 Ser Cys Val Ile
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
Glu Gly Leu Val Gly Asp Ala 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 ValArg Val Asn Thr Val His Pro Gly Pro Ile Lys
180 185 190
Thr Gln Arg Leu Asp Glu Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val Thr Pro Met Gly Arg 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>55
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>55
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgaaga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttattaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgcc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc agcggctcga tcagctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccgcat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg tgggtatacc gcacagtga 759

<210>56
<211>252
<212>PRT
<213> 人工序列

<220>

<223> 高加索酸奶乳杆菌的变体

<400>56

Met Thr Asp Arg Leu Lys His 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 Glu 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 Ser Cys Val Ile
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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Gln Arg Leu Asp Gln Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val Thr Pro Met Gly Arg 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>57

<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>57
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgaaga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttattaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgcc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc agcggctcga tgagctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccgcat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg tgggtatacc gcacagtga 759

<210>58
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>58
Met Thr Asp Arg Leu Lys His 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 Glu 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 Ser Cys Val Ile
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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Gln Arg Leu Asp Glu Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val Thr Pro Met Gly Arg 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>59
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>59
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgaaga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttattaa aagcgttgaa 300

gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgcc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgtctctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc agcggctcga tgagctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccgcat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg tgggtatacc gcacagtga 759

<210>60
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>60

Met	Thr	Asp	Arg	Leu	Lys	His	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	Glu	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	Ser	Cys	Val	Ile
				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				
Glu	Gly	Leu	Val	Gly	Asp	Pro	Ser	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 Pro Ile Lys
180 185 190
Thr Gln Arg Leu Asp Glu Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val Thr Pro Met Gly Arg 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>61
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>61
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgaaga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttattca tagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgcc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc agcggctcga tgagctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccgcat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg tgggtatacc gcacagtga 759

<210>62
<211>252
<212>PRT
<213> 人工序列

<220>

<223> 高加索酸奶乳杆菌的变体

<400>62

Met Thr AspArg Leu Lys His 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 Glu 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 Ser Cys Val Ile
85 90 95
His 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 Leu 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 Pro Ile Lys
180 185 190
Thr Gln Arg Leu Asp Glu Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val Thr Pro Met Gly Arg 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>63

<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>63
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgaaga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttcaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>64
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>64
Met Thr Asp Arg Leu Lys His 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 Glu 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 Ser Cys 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
Gln Gly Leu 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 Pro Ile Lys
180 185 190
Thr Pro Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val 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>65
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>65
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgaaga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg ggcttgtgaa aagcgttgaa 300

gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag tatcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>66
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体
<400>66

Met Thr Asp Arg Leu Lys His 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 Glu 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 Leu 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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Pro Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val 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>67
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体
<400>67
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgaaga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttattaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg ttaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>68
<211>252
<212>PRT
<213> 人工序列

<220>

<223> 高加索酸奶乳杆菌的变体

<400>68

Met Thr Asp Arg Leu Lys His 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 Glu 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 Ile
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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Pro Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Leu 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>69

<211>759

<212>DNA

<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>69
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgaaga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgtgaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>70
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>70
Met Thr Asp Arg Leu Lys His 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 Glu 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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Pro Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val 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>71
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>71
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgaaga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgtgaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420

atgagcagta ttgaggggct cgtaggcgat ccgtcgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatggaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>72
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>72
Met Thr Asp Arg Leu Lys His 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 Glu 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
Glu Gly Leu Val Gly Asp Pro Ser 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 Pro Ile Lys

180 185 190

Thr Pro Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205

Arg Thr Val 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 Trp Lys Phe Ala Thr Gly
225 230 235 240

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

<210>73
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>73
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgaaga agcaggctgg acgaaactgt tcgacaccac cgaggaggtt 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttgtgaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg acaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>74
<211>252
<212>PRT
<213> 人工序列

<220>

<223> 高加索酸奶乳杆菌的变体

<400>74

Met Thr Asp Arg Leu Lys His 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 Glu Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Val
65 70 75 80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ser Cys 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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Pro Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Thr 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>75

<211>759

<212>DNA

<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>75
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgaaga agcaggctgg acgaaactgt tcgacaccac cgaggaggtt 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattt gcgttttgaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg ttaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>76
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>76
Met Thr Asp Arg Leu Lys His 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 Glu Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Val
65 70 75 80

Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Cys 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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Pro Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Leu 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>77
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>77
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttgaa 300
gacactacca cggaggaatg gcataaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420

atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtaacatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc agcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc gcagcgtacg gtaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>78
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>78
Met Thr Asp Arg Leu Lys His 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 GluLys 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 Ser Cys Val Leu
85 90 95
Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp His 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 Leu Val Gly Asp Pro Thr Leu Gly Ala Tyr Asn Ala Ser Lys
145 150 155 160
Gly Ala Val Arg Asn 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 Gln Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205

Arg Thr Val 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>79
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>79
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc agcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggacctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>80
<211>252
<212>PRT
<213> 人工序列

<220>

<223> 高加索酸奶乳杆菌的变体

<400>80

Met Thr Asp Arg Leu Lys His 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 Ser Cys 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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Gln Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
210 215 220
Trp Thr 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>81

<211>759

<212>DNA

<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>81
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc agcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>82
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>82
Met Thr Asp Arg Leu Lys His 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 Ser Cys 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 LeuGly Ile Gln Arg
115 120 125
Met Lys Asn Lys Gly Leu Gly AlaSer Ile Ile Asn Met Ser Ser Ile
130 135 140
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Gln Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val 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>83
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>83
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420

atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>84
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体
<400>84

Met Thr Asp Arg Leu Lys His 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 Ser Cys 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
Glu Gly Leu 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 Pro Ile Lys
180 185 190

Thr Pro Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val 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>85
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体
<400>85
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cagttacaaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>86
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>86

Met Thr Asp Arg Leu Lys His 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 Thr
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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Pro Arg Leu Asp Asp Leu Glu Gly Trp 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>87

<211>759

<212>DNA

<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>87
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cagttacaaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg ataaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>88
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>88
Met Thr Asp Arg Leu Lys His 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 Thr
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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Pro Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Ile 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>89
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>89
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cagttacaaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540

gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>90
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>90
Met Thr Asp Arg Leu Lys His 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 Thr
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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Pro Leu Leu Asp Asp Leu Glu Gly Trp 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>91
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>91
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cagttacaaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga 759

<210>92
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>92

Met Thr Asp Arg Leu Lys His Lys Val Ala Ile ValThr 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 ValVal Asn Asn Ala Gly Ile Ala Val Thr
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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Pro Leu Leu Asp Asp Leu Glu Gly Trp 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 Trp Thr Ala Gln
245 250

<210>93

<211>759

<212>DNA

<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>93
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cagttacaaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gttaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>94
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>94
Met Thr Asp Arg Leu Lys His 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 Thr
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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Pro Leu Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val 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>95
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>95
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cagttacaaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540

gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cggatctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>96
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>96
Met Thr Asp Arg Leu Lys His 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 Thr
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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Pro Asp Leu Asp Asp Leu Glu Gly Trp 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>97
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>97
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cagttacaaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc tgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>98
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>98

Met Thr Asp Arg Leu Lys His 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 Thr
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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Leu Arg Leu Asp Asp Leu Glu Gly Trp 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>99

<211>759

<212>DNA

<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>99
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cagttacaaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>100
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>100
Met Thr Asp Arg Leu Lys His 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 Thr
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 Leu 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 Pro Ile Lys
180 185 190
Thr Pro Leu Leu Asp Asp Leu Glu Gly Trp 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>101
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>101
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cagttacaaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggtgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540

gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg attaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>102
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体
<400>102

Met Thr Asp Arg Leu Lys His 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 Thr
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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Pro Leu Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205

Arg Thr Ile 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>103
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体
<400>103
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattt gcgttattaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg acaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>104
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>104
Met Thr Asp Arg Leu Lys His Lys Val Ala Ile Val Thr Gly Gly Thr

1 5 10 15
Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe ValGlu 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 Cys Val Ile
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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Pro Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Thr 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>105
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>105

atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggtt 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cggtgtgtaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcgtctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>106
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>106

Met Thr Asp Arg Leu Lys His 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 Val
65 70 75 80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Ala Val Cys
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
Glu Gly Leu 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 Pro Ile Lys
180 185 190
Thr Pro Arg Leu Asp Asp Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val 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>107
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>107
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agtgggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttaaa aagcgttgaa 300
gacactacca cggaggaatg gcataaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct tgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactttgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgagctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccacat tggcgaaccg 660

aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210>108
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>108

Met	Thr	Asp	Arg	Leu	Lys	His	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	Val	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	Ser	Cys	Val	Leu
				85					90					95	
Lys	Ser	Val	Glu	Asp	Thr	Thr	Thr	Glu	Glu	Trp	His	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	Leu	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	Phe	Asp	Val	Arg	Val	Asn	Thr	Val	His	Pro	Gly	Pro	Ile	Lys
			180					185					190		
Thr	Pro	Arg	Leu	Asp	Glu	Leu	Glu	Gly	Trp	Glu	Glu	Met	Met	Ser	Gln
		195					200					205			
Arg	Thr	Val	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>109
<211>759
<212>DNA
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>109
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg tctccgatga agcaggctgg acgaaactgt tcgacaccac cgaggagacc 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggagtt gcgttttgaa aagcgttaaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc agcggctcga tgagctggaa 600
ggttgggagg aaatgatgtc acagcgtacg gtaaccccta tgggccgcat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg tgggtatacc gcacagtga 759

<210>110
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 高加索酸奶乳杆菌的变体

<400>110
Met Thr Asp Arg Leu Lys His 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 Val
50 55 60
Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Thr
65 70 75 80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ser Cys Val Leu
85 90 95
Lys Ser Val Lys 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 Leu 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 Pro Ile Lys
180 185 190
Thr Gln Arg Leu Asp Glu Leu Glu Gly Trp Glu Glu Met Met Ser Gln
195 200 205
Arg Thr Val Thr Pro Met Gly Arg 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>111
<211>252
<212>PRT
<213> 人工序列

<220>

<223> 具有短乳杆菌主链的人工改造的酮还原酶序列式
<223> 合成的构建体

<220>
<221>MISC_FEATURE
<222>(7)..(7)
<223>Xaa 是受限的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(17)..(17)
<223>Xaa 是非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(25)..(25)
<223>Xaa 是极性的或酸性的残基

<220>
<221>MISC_FEATURE
<222>(27)..(27)
<223>Xaa 是芳香族的残基

<220>
<221>MISC_FEATURE
<222>(41)..(41)
<223>Xaa 是极性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(53)..(53)
<223>Xaa 是极性的、非极性的或酸性的残基

<220>
<221>MISC_FEATURE
<222>(57)..(57)
<223>Xaa 是非极性的或脂肪族的残基

<220>

<221>MISC_FEATURE
<222>(64)..(64)
<223>Xaa 是极性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(66)..(66)
<223>Xaa 是酸性的残基

<220>
<221>MISC_FEATURE
<222>(68)..(68)
<223>Xaa 是极性的、酸性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(72)..(72)
<223>Xaa 是碱性的残基

<220>
<221>MISC_FEATURE
<222>(76)..(76)
<223>Xaa 是脂肪族的、非极性的或极性的残基

<220>
<221>MISC_FEATURE
<222>(80)..(80)
<223>Xaa 是脂肪族的、非极性的或极性的残基

<220>
<221>MISC_FEATURE
<222>(93)..(93)
<223>Xaa 是极性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(94)..(94)
<223>Xaa 是半胱氨酸、脂肪族的或非极性的残基
<220>

<221>MISC_FEATURE
<222>(95)..(95)
<223>Xaa 是脂肪族的、非极性的或酸性的残基

<220>
<221>MISC_FEATURE
<222>(96)..(96)
<223>Xaa 是半胱氨酸、极性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(97)..(97)
<223>Xaa 是受限的或极性的残基

<220>
<221>MISC_FEATURE
<222>(100)..(100)
<223>Xaa 是碱性的、酸性的、极性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(108)..(108)
<223>Xaa 是碱性的或受限的残基

<220>
<221>MISC_FEATURE
<222>(111)..(111)
<223>Xaa 是非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(139)..(139)
<223>Xaa 是非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(145)..(145)
<223>Xaa 是酸性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(147)..(147)
<223>Xaa 是芳香族的、脂肪族的或非极性的残基
<220>
<221>MISC_FEATURE
<222>(151)..(151)
<223>Xaa 是受限的、非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(152)..(152)
<223>Xaa 是极性的、芳香族的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(163)..(163)
<223>Xaa 是非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(165)..(165)
<223>Xaa 是极性的、非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(173)..(173)
<223>Xaa 是酸性的、非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(179)..(179)
<223>Xaa 是芳香族的、碱性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(190)..(190)
<223>Xaa 是受限的残基

<220>
<221>MISC_FEATURE
<222>(194)..(194)
<223>Xaa 是受限的、极性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(195)..(195)
<223>Xaa 是碱性的、酸性的、非极性的或极性的残基

<220>
<221>MISC_FEATURE
<222>(196)..(196)
<223>Xaa 是非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(197)..(197)
<223>Xaa 是酸性的残基

<220>
<221>MISC_FEATURE
<222>(198)..(198)
<223>Xaa 是酸性的或极性的残基

<220>
<221>MISC_FEATURE
<222>(202)..(202)
<223>Xaa 是芳香族的残基

<220>
<221>MISC_FEATURE
<222>(210)..(210)
<223>Xaa 是极性的或碱性的残基

<220>
<221>MISC_FEATURE
<222>(211)..(211)
<223>Xaa 是碱性的、脂肪族的、非极性的、酸性的或极性的残基

<220>
<221>MISC_FEATURE
<222>(212)..(212)
<223>Xaa 是极性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(216)..(216)
<223>Xaa 是受限的或碱性的残基

<220>
<221>MISC_FEATURE
<222>(223)..(223)
<223>Xaa 是非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(226)..(226)
<223>Xaa 是非极性的或极性的残基

<220>
<221>MISC_FEATURE
<222>(233)..(233)
<223>Xaa 是极性的或酸性的残基

<220>
<221>MISC_FEATURE
<222>(235)..(235)
<223>Xaa 是极性的或芳香族的残基

<220>
<221>MISC_FEATURE
<222>(248)..(248)
<223>Xaa 是非极性的或芳香族的残基

<220>
<221>MISC_FEATURE
<222>(249)..(249)

<223>Xaa 是芳香族的残基

<400>111

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

<210>112

<211>252

<212>PRT

<213> 人工序列

<220>
<223> 具有高加索酸奶乳杆菌主链的人工改造的酮还原酶序列式
<223> 合成的构建体

<220>
<221>MISC_FEATURE
<222>(7)..(7)
<223>Xaa 是受限的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(17)..(17)
<223>Xaa 是非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(25)..(25)
<223>Xaa 是极性的或酸性的残基

<220>
<221>MISC_FEATURE
<222>(27)..(27)
<223>Xaa 是芳香族的残基

<220>
<221>MISC_FEATURE
<222>(41)..(41)
<223>Xaa 是极性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(53)..(53)
<223>Xaa 是极性的、非极性的或酸性的残基

<220>
<221>MISC_FEATURE
<222>(57)..(57)

<223>Xaa 是非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(64)..(64)
<223>Xaa 是极性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(66)..(66)
<223>Xaa 是酸性的残基

<220>
<221>MISC_FEATURE
<222>(68)..(68)
<223>Xaa 是极性的、酸性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(72)..(72)
<223>Xaa 是碱性的残基

<220>
<221>MISC_FEATURE
<222>(76)..(76)
<223>Xaa 是脂肪族的、非极性的或极性的残基

<220>
<221>MISC_FEATURE
<222>(80)..(80)
<223>Xaa 是脂肪族的、非极性的或极性的残基

<220>
<221>MISC_FEATURE
<222>(93)..(93)
<223>Xaa 是极性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE

<222>(94)..(94)
<223>Xaa 是半胱氨酸、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(95)..(95)
<223>Xaa 是脂肪族的、非极性的或酸性的残基

<220>
<221>MISC_FEATURE
<222>(96)..(96)
<223>Xaa 是半胱氨酸、极性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(97)..(97)
<223>Xaa 是受限的或极性的残基

<220>
<221>MISC_FEATURE
<222>(100)..(100)
<223>Xaa 是碱性的、酸性的、极性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(108)..(108)
<223>Xaa 是碱性的或受限的残基

<220>
<221>MISC_FEATURE
<222>(111)..(111)
<223>Xaa 是非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(139)..(139)
<223>Xaa 是非极性的或脂肪族的残基

<220>

<221>MISC_FEATURE
<222>(145)..(145)
<223>Xaa 是酸性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(147)..(147)
<223>Xaa 是芳香族的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(151)..(151)
<223>Xaa 是受限的、非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(152)..(152)
<223>Xaa 是极性的、芳香族的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(163)..(163)
<223>Xaa 是非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(165)..(165)
<223>Xaa 是极性的、非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(173)..(173)
<223>Xaa 是酸性的、非极性的或脂肪族的残基
<220>
<221>MISC_FEATURE
<222>(179)..(179)
<223>Xaa 是芳香族的、碱性的、脂肪族的或非极性的残基

<220>

<221>MISC_FEATURE
<222>(190)..(190)
<223>Xaa 是受限的残基

<220>
<221>MISC_FEATURE
<222>(194)..(194)
<223>Xaa 是受限的、极性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(195)..(195)
<223>Xaa 是碱性的、酸性的、非极性的或极性的残基

<220>
<221>MISC_FEATURE
<222>(196)..(196)
<223>Xaa 是非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(197)..(197)
<223>Xaa 是酸性的残基

<220>
<221>MISC_FEATURE
<222>(198)..(198)
<223>Xaa 是酸性的或极性的残基

<220>
<221>MISC_FEATURE
<222>(202)..(202)
<223>Xaa 是芳香族的残基

<220>
<221>MISC_FEATURE
<222>(210)..(210)
<223>Xaa 是极性的或碱性的残基
<220>

<221>MISC_FEATURE
<222>(211)..(211)
<223>Xaa 是碱性的、脂肪族的、非极性的、酸性的或极性的残基

<220>
<221>MISC_FEATURE
<222>(212)..(212)
<223>Xaa 是极性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(216)..(216)
<223>Xaa 是受限的或碱性的残基

<220>
<221>MISC_FEATURE
<222>(223)..(223)
<223>Xaa 是非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(226)..(226)
<223>Xaa 是非极性的或极性的残基

<220>
<221>MISC_FEATURE
<222>(233)..(233)
<223>Xaa 是极性的或酸性的残基

<220>
<221>MISC_FEATURE
<222>(235)..(235)
<223>Xaa 是极性的或芳香族的残基

<220>
<221>MISC_FEATURE
<222>(248)..(248)
<223>Xaa 是非极性的或芳香族的残基

<220>
<221>MISC_FEATURE
<222>(249)..(249)
<223>Xaa 是芳香族的残基

<400>112

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

<210>113
<211>759
<212>DNA
<213> 微小乳杆菌 (Lactobacillus Minor)
<400>113
atgaccgatc ggttgaaggg gaaagtagca attgtaactg gcggtacctt gggaattggc 60
ttggcaatcg ctgataagtt tgttgaagaa ggcgcaaagg ttgttattac cggccgtcac 120
gctgatgtag gtgaaaaagc tgccagatca atcggcggca cagacgttat ccgttttgtc 180
caacacgatg cttctgatga aaccggctgg actaagttgt ttgatacgac tgaagaagca 240
tttggcccag ttaccacggt tgtcaacaat gccggaattg cggtcagcaa gagtgttgaa 300
gataccacaa ctgaagaatg gcgcaagctg ctctcagtta acttggatgg tgtcttcttc 360
ggtacccgtc ttggaatcca acgtatgaag aataaaggac tcggagcatc aatcatcaat 420
atgtcatcta tcgaaggttt tgttggtgat ccagctctgg gtgcatacaa cgcttcaaaa 480
ggtgctgtca gaattatgtc taaatcagct gccttggatt gcgctttgaa ggactacgat 540
gttcgggtta acactgttca tccaggttat atcaagacac cattggttga cgatcttgaa 600
ggggcagaag aaatgatgtc acagcggacc aagacaccaa tgggtcatat cggtgaacct 660
aacgatatcg cttggatctg tgtttacctg gcatctgacg aatctaaatt tgccactggt 720
gcagaattcg ttgtcgacgg agggtacacc gcccaatag 759

<210>114
<211>252
<212>PRT
<213> 微小乳杆菌

<400>114
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
Arg Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
50 55 60
Ser Asp Glu Thr 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
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 AsnMet Ser Ser Ile
130 135 140
Glu Gly Phe Val Gly Asp Pro Ala 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>115
<211>1050
<212>DNA
<213> 人工序列

<220>
<223> 密码子优化型的黄曲霉 (Aspergillus flavus)adh 基因

<400>115
atgtctattc cggaaatgca atgggctcaa gtagcggagc agaaaggcgg tccgctgatt 60
tacaaacaga ttccggttcc gaaaccgggc ccggacgaaa tcctggtaaa agtgcgttac 120
tctggtgtgt gccacacgga tctgcacgcg ctgaaaggtg actggccact gcctgtgaaa 180
atgccactgg tgggtggcca tgagggcgcg ggtgttgtcg tcgctcgtgg cgatctggtt 240
actgagttcg aaattggtga ccatgctggc ctgaaatggc tgaacggtag ctgcctggca 300
tgcgagttct gtaaacaggc tgacgaaccg ctgtgcccta acgcgagcct gtctggttat 360
actgtagatg gtactttcca gcagtatgcc attggcaaag ccacccatgc gagcaagctg 420
ccgaaaaacg tgcctctgga tgccgtggca ccggttctgt gcgcgggcat taccgtatac 480
aagggcctga aagaaagcgg tgttcgtcct ggccaaaccg ttgcgatcgt aggtgcgggt 540
ggtggtctgg gctccctggc gctgcagtac gcgaaagcta tgggtattcg cgtggtggcg 600
attgatggcg gtgaagagaa acaagcgatg tgtgaacagc tgggcgcaga agcctacgtt 660

gactttacca aaacgcaaga tctggtagca gatgttaagg ccgccacccc tgaaggtctg 720
ggcgcgcatg cggtaattct gctggcggtc gccgaaaaac catttcagca ggcggccgaa 780
tatgtcagcc gtggtacggt ggttgccatt ggtctgccgg cgggcgcatt cctgcgcgcg 840
ccggtgttta acactgtggt tcgtatgatt aacattaagg gtagctacgt tggcaaccgc 900
caggacggcg ttgaggcagt ggacttcttc gcgcgcggcc tgatcaaagc gccgttcaaa 960
accgctcctc tgcaagatct gccgaaaatc ttcgaactga tggaacaagg taagattgca 1020
ggtcgctacg ttctggaaat cccggaatga 1050

<210>116
<211>349
<212>PRT
<213> 人工序列

<220>
<223> 黄曲霉的 adh

<400>116

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

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

<210>117
<211>729
<212>DNA
<213> 人工序列

<220>
<223> 密码子优化型的酒类酒球菌 (Oenococcus oeni)adh 基因

<400>117
atgtttgatc gcctgaaggg taaggtggcg atcgtcaccg gcggtaatag cggtattggt 60
aaggcaatcg cggcagactt catcgccgag ggcgcgaaag tagtgattac cggtcgcaac 120
caagagaagg gtcgccagac tgcgctggaa atcgccggtg acattctgtt cattcagcag 180
gacgtgagcc aggaggcgga ctggcagaaa gttatctcca aaaccatcga gaagttcggt 240
aagtttgata ttctggttaa caacgctggc gtcggcggcg tgggtaagcc actggcagaa 300
atgtccctgt ctgaattcaa ttggacccaa tccattaatc tgtccggtaa cttcctgggt 360
attcacttcg cgctgaacaa aatgaccgaa cctggtagca ttattgacgt aagcagcgcg 420

gctggcctgc gtggttttcc gggcgcagcc gattatagcg cttccaaagg tggtacgcgt 480
ctgctgacgc gcgccgcggc gctggaagcg ctgcaaatgg gtaaaaagat ccgtgtaaac 540
tctattcacc caggttggat tgataccgat atcgtaccaa aagatatgcg tgagcaggtt 600
atcgcgatga cccctatgca tcacctgggc caaccgaaag atattgccaa gctggcgacg 660
tatctggcat ctgacgagtc cgagtatact acgggctccg aactggcagc cgacggcggt 720
ctgatggca 729

<210>118
<211>243
<212>PRT
<213> 人工序列

<220>
<223> 酒类酒球菌的 adh

<400>118
Met Phe Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Asn
1 5 10 15
Ser Gly Ile Gly Lys Ala Ile Ala Ala Asp Phe Ile Ala Glu Gly Ala
20 2 530
Lys Val Val Ile Thr Gly Arg Asn Gln Glu Lys Gly Arg Gln Thr Ala
35 40 45
Leu Glu Ile Ala Gly Asp Ile Leu Phe Ile Gln Gln Asp Val Ser Gln
50 55 60
Glu Ala Asp Trp Gln Lys Val Ile Ser Lys Thr Ile Glu Lys Phe Gly
65 70 75 80
Lys Phe Asp Ile Leu Val Asn Asn Ala Gly Val Gly Gly Val Gly Lys
85 90 95
Pro Leu Ala Glu Met Ser Leu Ser Glu Phe Asn Trp Thr Gln Ser Ile
100 105 110
Asn Leu Ser Gly Ash Phe Leu Gly Ile His Phe Ala Leu Asn Lys Met
115 120 125
Thr Glu Pro Gly Ser Ile Ile Asp Val Ser Ser Ala Ala Gly Leu Arg
130 135 140
Gly Phe Pro Gly Ala Ala Asp Tyr Ser Ala Ser Lys Gly Gly Thr Arg
145 150 155 160
Leu Leu Thr Arg Ala Ala Ala Leu Glu Ala Leu Gln Met Gly Lys Lys
165 170 175
Ile Arg Val Asn Ser Ile His Pro Gly Trp Ile Asp Thr Asp Ile Val

180 185 190

Pro Lys Asp Met Arg Glu Gln Val Ile Ala Met Thr Pro Met His His

195 200 205

Leu Gly Gln Pro Lys Asp Ile Ala Lys Leu Ala Thr Tyr Leu Ala Ser

210 215 220

Asp Glu Ser Glu Tyr Thr Thr Gly Ser Glu Leu Ala Ala Asp Gly Gly

225 230 235 240

Leu Met Ala

<210>119
<211>768
<212>DNA
<213> 人工序列

<220>
<223> 密码子优化型的富养罗尔斯通氏菌 (Ralstonia eutropha)adh 基因

<400>119
atgcagggca aaattgcact ggttacgggc ggtgcgtctg gtatgggtct ggccttcacc 60
cgtcgcctgg cccaagaagg cgcacgtgtt tacttcaccg acattaacgg cgccgcgggt 120
cacgcgacgg aaagcgagct gcagggtgct ggtctggcgg tcgtattcct gcgtcatgac 180
gtgacccagg aaaacgactg gctgcgcgtc ctggaacata ttggcgctgt tgacggtcgc 240
ctggatgtac tggtgaataa tgccggcatt gcgatttctc acaacattga gacttgtact 300
actgaggatt ttgatcgcac gctgaatgtt aacctgaagt ccgtgtttct gggctgcaaa 360
cacggtctgg ctctgatgaa acagaagggt ggcagcatca ttaacgttag ctctatcacc 420
gcgatttgcg gcgaacctgt ggcgctggcg tatagcgcta gcaaagccgg cgtccgtttt 480
ctgagcaagt ctgttgccct gcattgtgcg gagaagggct acgccatccg tgttaacagc 540
ctgcacccgg gttacattga taccccgctg ctggcaggca gcaacgcggg tggctccctg 600
accgctgacg aggtaaccgc gagccgtgta cgtattggcg gtgaaattcc gctgaaacgt 660
cgcggcaccc cggacgaagt tgcaggcgcg gttctgtacc tggctagcga tgaatccacc 720
tatgttactg gtactgagct ggtaattgat ggcggctacg catgtcat 768

<210>120
<211>256
<212>PRT
<213> 人工序列

<220>
<223> 富养罗尔斯通氏菌的 adh

<400>120

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

<210>121

<211>939

<212>DNA

<213> 人工序列

<220>
<223> 酿酒酵母 (Saccharomyces cervisiae) 的 YDL ADH

<400>121
atgtcttttc accaacagtt tttcacgctg aacaacggca ataaaatccc ggcgattgcc 60
atcatcggca ctggtacacg ttggtataaa aatgaagaaa ctgacgcgac cttctccaat 120
agtctggttg aacaaatcgt gtatgcgttg aaactgccgg ggattatcca catcgacgcc 180
gcggagattt atcgcaccta cccggaagtg ggtaaagcac tgtccctgac cgaaaagcct 240
cgtaacgcga tttttctgac ggataaatat tctccgcaga ttaaaatgag tgactcccct 300
gcggacggtc tggatttagc attgaagaaa atgggtacag attatgttga tttatatctg 360
ttacattccc cgtttgtttc gaaggaagtg aatggcttaa gcttagaaga ggcttggaaa 420
gatatggagc agttatacaa aagtggtaaa gctaaaaaca tcggggtttc caatttcgca 480
gtggaagacc tgcaacgtat cctgaaagtc gctgaagtta aacctcaggt caaccagatt 540
gagttctctc cgttcctgca aaaccaaaca ccaggcattt ataaattctg tcaggagcac 600
gatatcctgg tggaagcata ttctccgctg ggcccgctgc agaagaaaac cgcgcaggat 660
gacagccaac cattttttga gtacgtcaaa gaattgagcg aaaaatacat caaatccgag 720
gcccagatca tcctgcgctg ggtcactaaa cgcggtgtgc tgccagttac cacctcttca 780
aagcctcagc gcattagcga tgctcagaac ctgttttcct tcgacctgac agcggaagag 840
gttgataaaa tcacggagct gggtctggaa catgaaccgc tgcgcctgta ctggaataaa 900
ttgtatggca aatataacta cgccgcccag aaagtgtaa 939

<210>122
<211>312
<212>PRT
<213> 人工序列
<220>
<223> 酿酒酵母的 YDL ADH

<400>122
Met Ser Phe His Gln Gln Phe Phe Thr Leu Asn Asn Gly Asn Lys Ile
1 5 10 15
Pro Ala Ile Ala Ile Ile Gly Thr Gly Thr Arg Trp Tyr Lys Asn Glu
20 25 30
Glu Thr Asp Ala Thr Phe Ser Asn Ser Leu Val Glu Gln Ile Val Tyr
35 40 45
Ala Leu Lys Leu Pro Gly Ile Ile His Ile Asp Ala Ala Glu Ile Tyr
50 55 60
Arg Thr Tyr Pro Glu Val Gly Lys Ala Leu Ser Leu Thr Glu Lys Pro

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

<210>123
<211>1047
<212>DNA
<213> 人工序列

<220>
<223> 红串红球菌 (Rhodococcus erythropolis) 的 ADH

<400>123

atgaaagcca ttcagtacac tcgtatcggt gcggaaccag aactgactga aatcccgaag 60
ccggaaccgg gcccgggcga agtactgctg gaagtcacgg cagctggcgt gtgccattcc 120
gatgatttca ttatgtctct gccggaagaa cagtacacct acggcctgcc gctgaccctg 180
ggtcatgaag gtgctggtaa agttgccgca gttggcgaag gtgttgaagg gttggatatt 240
ggcaccaatg tggttgtgta cggcccatgg ggttgtggca actgttggca ttgcagtcag 300
ggcctggaga actattgctc ccgtgcgcag gaactgggta ttaacccgcc tggtctgggt 360
gctccggggg ctttggcaga atttatgatt gtcgactcac cacgtcattt ggtcccgatt 420
ggcgatttag accctgttaa aactgttccg ttgactgatg cgggcctgac cccataccat 480
gcaatcaaac gctccctgcc gaaactgcgc ggcggctctt atgcagtagt gatcggtacg 540
ggtggcctgg gccacgtggc tatccaactg ctgcgtcatt tatctgctgc aacggtgatc 600
gccttggacg tttctgccga taaactggaa ctggctacca aagtcggcgc acatgaagta 660
gtcctgtctg ataaagatgc agcggagaat gtgcgtaaaa ttactggtag ccaaggtgca 720
gctttggtgt tggattttgt gggctatcag cctaccattg acaccgccat ggcagtggcg 780
ggcgtgggct ctgacgtcac cattgttggt atcggtgatg gccaggcaca tgcgaaagtt 840
ggtttcttcc agagtcctta tgaggcatcg gttacggtac cttattgggg cgctcgtaat 900
gaactgatcg aattgatcga tctggcgcat gctggtattt tcgacattgc cgttgagacc 960
ttctctttgg ataatggtgc agaggcctat cgtcgcctgg ctgcgggcac actgtcaggc 1020
cgtgcggtag tcgtcccggg cctgtaa 1047

<210>124
<211>348
<212>PRT
<213> 人工序列

<220>
<223> 红串红球菌的 ADH

<400>124

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

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

<210>125
<211>1047
<212>DNA
<213> 人工序列

<220>
<223> 酿酒酵母的 YGL ADH

<400>125
atgaccacgg aaaaaaccgt agtgttcgtg tcaggcgcga ccggttttat tgctctgcac 60
gtggtagatg acctgctgaa aactggttac aaagtaattg gttccggtcg ttctcaggaa 120
aaaaatgacg gtttgctgaa gaagttcaag tccaacccga atctgagcat ggaaattgtg 180
gaagatattg cggcaccaaa cgccttcgat aaagtattcc agaaacatgg taaagaaatt 240
aaagtggtcc tgcatatcgc gtccccggtc catttcaaca ctaccgattt cgaaaaagac 300
ttactgatcc cggcggtaaa cggtaccaaa tctattttgg aagcaattaa gaactatgcc 360
gcagacaccg tggaaaaagt ggttattact tcatctgttg ccgcgttggc ctctccgggt 420
gatatgaaag ataccagctt cgtggttaac gaagaatcct ggaataaaga cacctgggaa 480
tcgtgtcagg cgaatgctgt gtccgcttat tgcggttcta aaaaattcgc agagaaaacg 540
gcgtgggact tcttggaaga aaaccagagc agcattaaat ttactctgtc cacgattaac 600
ccaggcttcg tttttggtcc gcagctgttc gccgactcct tgcgcaatgg tattaactct 660
agcagtgcga ttattgcgaa cctggtgtcg tataaattag gggataactt ctacaattat 720
agcggcccgt ttatcgacgt ccgtgacgtt tccaaagctc atctgctggc atttgagaaa 780
cctgaatgcg ccggtcagcg cctgtttctg tgcgaggata tgttctgttc ccaggaagcc 840
ctggacattc tgaacgaaga atttccacag ctgaagggca agatcgcaac gggcgaacct 900
ggcagcggct cgaccttcct gactaaaaat tgttgcaaat gcgacaatcg taaaactaaa 960
aacttgctgg gcttccagtt caacaaattc cgcgactgca ttgtcgatac tgcgtcccag 1020
ttgctggaag tgcaaagcaa aagctaa 1047

<210>126
<211>348
<212>PRT
<213> 人工序列

<220>
<223> 酿酒酵母的 YGL ADH

<400>126
Met Thr Thr Glu Lys Thr Val Val Phe Val Ser Gly Ala Thr Gly Phe
1 5 10 15
Ile Ala Leu His Val Val Asp Asp Leu Leu Lys Thr Gly Tyr Lys Val
20 25 30
Ile Gly Ser Gly Arg Ser Gln Glu Lys Asn Asp Gly Leu Leu Lys Lys
35 40 45

Phe Lys Ser Asn Pro Asn Leu Ser Met Glu Ile Val Glu Asp Ile Ala
50 55 60
Ala Pro Asn Ala Phe Asp Lys Val Phe Gln Lys His Gly Lys Glu Ile
65 70 75 80
Lys Val Val Leu His Ile Ala Ser Pro Val His Phe Asn Thr Thr Asp
85 90 95
Phe Glu Lys Asp Leu Leu Ile Pro Ala Val Asn Gly Thr Lys Ser Ile
100 105 110
Leu Glu Ala Ile Lys Asn Tyr Ala Ala Asp Thr Val Glu Lys Val Val
115 120 125
Ile Thr Ser Ser Val Ala Ala Leu Ala Ser Pro Gly Asp Met Lys Asp
130 135 140
Thr Ser Phe Val Val Asn Glu Glu Ser Trp Asn Lys Asp Thr Trp Glu
145 150 155 160
Ser Cys Gln Ala Asn Ala Val Ser Ala Tyr Cys Gly Ser Lys Lys Phe
165 170 175
Ala Glu Lys Thr Ala Trp Asp Phe Leu Glu Glu Asn Gln Ser Ser Ile
180 185 190
Lys Phe Thr Leu Ser Thr Ile Asn Pro Gly Phe Val Phe Gly Pro Gln
195 200 205
Leu Phe Ala Asp Ser Leu Arg Asn Gly Ile Asn Ser Ser Ser Ala Ile
210 215 220
Ile Ala Asn Leu Val Ser Tyr Lys Leu Gly Asp Asn Phe Tyr Asn Tyr
225 230 235 240
Ser Gly Pro Phe Ile Asp Val Arg Asp Val Ser Lys Ala His Leu Leu
245 250 255
Ala Phe Glu Lys Pro Glu Cys Ala Gly Gln Arg Leu Phe Leu Cys Glu
260 265 270
Asp Met Phe Cys Ser Gln Glu Ala Leu Asp Ile Leu Asn Glu Glu Phe
275 280 285
Pro Gln Leu Lys Gly Lys Ile Ala Thr Gly Glu Pro Gly Ser Gly Ser
290 295 300
Thr Phe Leu Thr Lys Asn Cys Cys Lys Cys Asp Asn Arg Lys Thr Lys
305 310 315 320
Asn Leu Leu Gly Phe Gln Phe Asn Lys Phe Arg Asp Cys Ile Val Asp
325 330 335
Thr Ala Ser Gln Leu Leu Glu Val Gln Ser Lys Ser
340 345

<210>127
<211>939
<212>DNA
<213> 人工序列

<220>
<223> 酿酒酵母的 YPR ADH

<400>127
atgccggcaa cgttaaaaaa cagcagtgct accttaaaat taaacacagg tgcgagcatt 60
cctgtcctgg ggttcggcac ctggcgctct gtcgataaca acggctatca tagtgtaatt 120
gcggcgctga aagcggggta ccgtcatatc gatgctgcgg ccatctatct gaatgaagaa 180
gaagtcggcc gtgcgatcaa ggactccggt gttcctcgtg aagaaatttt tattaccacc 240
aaactgtggg gcaccgaaca acgcgatcca gaagcagccc tgaacaaatc tctgaaacgt 300
ctgggtctgg actatgtgga cctgtatctg atgcactggc cggtccctct gaaaacagac 360
cgtgtaactg acggtaacgt cctgtgcatc ccgaccctgg aagatggcac cgtggacatc 420
gataccaaag agtggaattt tattaaaacc tgggaactga tgcaggaatt gccgaaaact 480
ggtaagacca aagccgtcgg tgtgtccaat ttttccatca acaatatcaa agaactgctg 540
gaatcgccaa ataacaaggt cgttccagca accaatcaga tcgagattca tccgttgctg 600
ccgcaggatg aattaatcgc cttttgtaaa gaaaaaggca ttgtggtcga agcatatagc 660
ccattcggct ccgctaacgc cccgctgctg aaagaacagg cgattatcga tatggccaaa 720
aagcacggcg tcgaaccggc gcaactgatt atcagctggt cgattcagcg cggttatgtg 780
gtattggcca agtccgtaaa tccggagcgt atcgtgtcga actttaagat ttttaccctg 840
ccagaggatg atttcaaaac catctctaac ctgagcaaag tgcacggtac caaacgtgtc 900
gttgacatga aatggggctc atttccgatt tttcaataa 939

<210>128
<211>312
<212>PRT
<213> 人工序列

<220>
<223> 酿酒酵母的 YPR ADH

<400>128
Met Pro Ala Thr Leu Lys Asn Ser Ser Ala Thr Leu Lys Leu Asn Thr
1 5 10 15
Gly Ala Ser Ile Pro Val Leu Gly Phe Gly Thr Trp Arg Ser Val Asp
20 25 30

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

<210>129
<211>1029

<212>DNA
<213> 人工序列

<220>
<223> 酿酒酵母的 GRE ADH

<400>129
atgtctgtgt tcgtgtcagg cgcgaatggt tttattgctc agcacatcgt agatctgctg 60
ctgaaagaag attacaaagt aattggttcc gcacgttctc aggaaaaagc tgaaaatttg 120
accgaagcct tcggtaacaa cccgaaattt agcatggaag tggtgcctga tattagcaaa 180
ctggatgcct tcgatcatgt attccagaaa catggtaaag atattaaaat cgtcctgcat 240
accgcgtccc cgttttgttt cgatattacc gattccgaac gtgacttact gatcccggcg 300
gtaaacggtg tcaaaggtat tttgcacagt attaagaaat atgccgcaga cagcgtggaa 360
cgcgtggttc tgacttcatc ttacgccgcg gtatttgata tggcgaagga aaacgataag 420
agcctgacct tcaacgaaga atcctggaat ccggcgacct gggaatcgtg tcagagtgat 480
ccggtgaacg cttattgcgg ttctaaaaaa ttcgcagaga aagcagcgtg ggaattcttg 540
gaagaaaacc gtgatagcgt gaaatttgag ctgacagcgg tcaacccagt ttacgttttt 600
ggtccgcaga tgttcgataa agatgttaaa aaacacttga acaccagctg cgaactggtg 660
aactctctga tgcatctgag ccctgaagat aaaattccgg aactgtttgg cggttacatc 720
gacgtccgtg acgttgcgaa agctcatctg gttgcatttc agaaacgtga aacaatcggt 780
cagcgcctga tcgtgtcgga ggcacgtttc acgatgcagg acgttctgga cattctgaac 840
gaagactttc cagtactgaa gggcaatatc ccggtcggca agcctggcag cggcgccacc 900
cataatactc tgggcgccac cctggacaat aaaaaaagca aaaaattgct gggcttcaaa 960
ttccgtaatc tgaaagagac tattgacgat actgcgtccc agatcctgaa attcgaaggt 1020
cgcatttaa 1029

<210>130
<211>342
<212>PRT
<213> 人工序列

<220>
<223> 酿酒酵母的 GRE ADH

<400>130
Met Ser Val Phe Val Ser Gly Ala Asn Gly Phe Ile Ala Gln His Ile
1 5 10 15
Val Asp Leu Leu Leu Lys Glu Asp Tyr Lys Val Ile Gly Ser Ala Arg
20 25 30

Ser Gln Glu Lys Ala Glu Asn Leu Thr Glu Ala Phe Gly Asn Asn Pro
35 40 45
Lys Phe Ser Met Glu Val Val Pro Asp Ile Ser Lys Leu Asp Ala Phe
50 55 60
Asp His Val Phe Gln Lys His Gly Lys Asp Ile Lys Ile Val Leu His
65 70 75 80
Thr Ala Ser Pro Phe Cys Phe Asp Ile Thr Asp Ser Glu Arg Asp Leu
85 90 95
Leu Ile Pro Ala Val Asn Gly Val Lys Gly Ile Leu His Ser Ile Lys
100 105 110
Lys Tyr Ala Ala Asp Ser Val Glu Arg Val Val Leu Thr Ser Ser Tyr
115 120 125
Ala Ala Val Phe Asp Met Ala Lys Glu Asn Asp Lys Ser Leu Thr Phe
130 135 140
Asn Glu Glu Ser Trp Asn Pro Ala Thr Trp Glu Ser Cys Gln Ser Asp
145 150 155 160
Pro Val Asn Ala Tyr Cys Gly Ser Lys Lys Phe Ala Glu Lys Ala Ala
165 170 175
Trp Glu Phe Leu Glu Glu Asn Arg Asp Ser Val Lys Phe Glu Leu Thr
180 185 190
Ala Val Asn Pro Val Tyr Val Phe Gly Pro Gln Met Phe Asp Lys Asp
195 200 205
Val Lys Lys His Leu Asn Thr Ser Cys Glu Leu Val Asn Ser Leu Met
210 215 220
His Leu Ser Pro Glu Asp Lys Ile Pro Glu Leu Phe Gly Gly Tyr Ile
225 230 235 240
Asp Val Arg Asp Val Ala Lys Ala His Leu Val Ala Phe Gln Lys Arg
245 250 255
Glu Thr Ile Gly Gln Arg Leu Ile Val Ser Glu Ala Arg Phe Thr Met
260 265 270
Gln Asp Val Leu Asp Ile Leu Asn Glu Asp Phe Pro Val Leu Lys Gly
275 280 285
Asn Ile Pro Val Gly Lys Pro Gly Ser Gly Ala Thr His Asn Thr Leu
290 295 300
Gly Ala Thr Leu Asp Asn Lys Lys Ser Lys Lys Leu Leu Gly Phe Lys
305 310 315 320
Phe Arg Asn Leu Lys Glu Thr Ile Asp Asp Thr Ala Ser Gln Ile Leu
325 330 335
Lys Phe Glu Gly Arg Ile

340

<210>131
<211>1032
<212>DNA
<213> 人工序列

<220>
<223> 赭色掷孢酵母 (S.salmonicolor) 的密码子优化的 ADH

<400>131
atggcaaaga tcgacaacgc agttctgccg gagggttctc tggtgctggt caccggcgcg 60
aacggctttg tcgctagcca tgtggtcgaa caactgctgg aacacggcta taaggtgcgc 120
ggcactgctc gctctgcctc caaactggcg aacctgcaga aacgttggga cgccaaatac 180
cctggtcgtt tcgagactgc cgttgttgaa gacatgctga agcagggtgc atatgatgaa 240
gttattaaag gcgcggcagg tgtcgcccac atcgcgtccg tggtcagctt ttctaacaaa 300
tatgatgagg tggtaactcc tgcgatcggt ggcacgctga atgccctgcg tgccgcagct 360
gctacgcctt ccgtgaaacg ttttgtgctg accagcagca ctgtttctgc actgattcca 420
aaacctaacg tcgaaggtat ttatctggat gagaagagct ggaacctgga aagcattgat 480
aaggctaaaa ccctgcctga atctgatccg cagaaaagcc tgtgggtcta cgccgcaagc 540
aaaacggaag cggaactggc tgcctggaaa ttcatggacg aaaacaaacc gcactttact 600
ctgaatgccg ttctgccaaa ctacactatc ggtaccattt ttgacccaga aacccaatcc 660
ggttccactt ccggctggat gatgtctctg ttcaatggcg aagtatctcc ggcactggcg 720
ctgatgccgc cgcagtacta tgtctctgca gttgatatcg gtctgctgca cctgggttgt 780
ctggttctgc cgcaaatcga acgccgtcgt gtttacggca ccgcaggcac ctttgattgg 840
aacaccgttc tggcgacctt ccgtaaactg tatccgtcca agacgttccc ggctgacttt 900
ccggatcagg gccaggatct gtccaaattt gataccgccc cgagcctgga gattctgaaa 960
tccctgggcc gccctggctg gcgtagcatc gaggaatcta tcaaagatct ggtgggttcc 1020
gagaccgcct aa 1032

<210>132
<211>343
<212>PRT
<213> 人工序列

<220>
<223> 赭色掷孢酵母的 ADH

<400>132

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

305 310 315 320
Ser Leu Gly Arg Pro Gly Trp Arg Ser Ile Glu Glu Ser Ile Lys Asp
325 330 335
Leu Val Gly Ser Glu Thr Ala
340

<210>133
<211>1041
<212>DNA
<213> 人工序列

<220>
<223> 天蓝色链霉菌 (S.coelicolor) 的密码子优化的 ADH

<400>133
atgaaggcac tgcagtaccg caccatcggc gccccgcccg aggtcgtcac cgtcccggac 60
ccggagccgg gccccggcca ggtgctgttg aaggtgaccg cggccggagt ctgccactcc 120
gacatcgcgg tgatgagctg gcccgccgag ggcttcccgt acgagctgcc gctcaccctc 180
ggccacgagg gcgtcggcac cgtggccgcg ctcggcgccg gggtgacggg gctcgccgag 240
ggcgacgcgg tcgccgtgta cgggccctgg ggctgcggca cctgcgccaa gtgcgcggag 300
ggcaaggaga actactgcct gcgcgccgac gagctgggca tccgtccgcc ggggctcggg 360
cgtccggggt ccatggccga gtacctgctg atcgacgacc cccggcacct ggtcccgctg 420
gacgggctcg acccggtcgc ggcggtgccg ctcaccgacg ccggactgac gccgtaccac 480
gcgatcaagc ggtcgctgcc caagctggtc cccggctcca ccgcggtggt catcggcacc 540
ggtggtctcg gccacgtcgc catccagctg ctgcgcgccc tgacgtccgc ccgggtggtc 600
gccctggacg tcagcgagga gaagctgcgc ctcgcccgtg cggtgggcgc gcacgaggcg 660
gtgctgtcgg acgcgaaggc cgcggacgcg gtgcgcgaga tcaccggcgg tctcggtgcc 720
gaggccgtgt tcgacttcgt cggcgtggcg cccaccgtgc agaccgccgg agccgtcgcg 780
gccgtcgagg gcgatgtcac cctggtcggc atcggcggcg gatcgctgcc cgtcggcttc 840
ggcatgctgc cgttcgaggt gtcggtcaac gccccctact ggggcagccg cagcgagctg 900
accgaggtgc tgaacctggc ccgctccggt gccgtgtcgg tgcacaccga gacgtactcc 960
ctggacgacg ccccgctcgc ctacgagcgg ctgcacgagg gcagggtcaa cggccgcgcg 1020
gtgatcctgc cccacggctg a 1041

<210>134
<211>346
<212>PRT
<213> 人工序列

<220>

<223> 天蓝色链霉菌的 ADH

<400>134

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

275 280 285

Val Asn Ala Pro Tyr Trp Gly Ser Arg Ser Glu Leu Thr Glu Val Leu

290 295 300

Asn Leu Ala Arg Ser Gly Ala Val Ser Val His Thr Glu Thr Tyr Ser

305 310 315 320

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

325 330 335

Asn Gly Arg Ala Val Ile Leu Pro His Gly

340 345

<210>135

<211>1056

<212>DNA

<213> 人工序列

<220>

<223> 布氏热厌氧菌 (T brockii) 的密码子优化的 ADH-TB

<400>135

atgaaaggct tcgccatgct gagcatcggc aaagtgggtt ggattgaaaa agaaaaaccg 60

gcgccaggcc cgttcgatgc aattgtgcgc cctctggcag tagcgccgtg taccagcgat 120

attcatactg tgtttgaagg tgccattggc gagcgtcaca atatgattct gggccatgaa 180

gccgttggtg aagttgttga ggttggcagc gaagtgaagg atttcaaacc gggcgatcgc 240

gttgtcgttc cagcgattac cccggattgg cgcaccagcg aagtccagcg cggctaccat 300

cagcactctg gcggcatgct ggccggctgg aaattcagca atgtaaagga tggtgtgttc 360

ggtgaatttt ttcacgttaa cgacgcagac atgaatctgg cgcacctgcc gaaagaaatc 420

ccgctggaag cagcggttat gattccggat atgatgacca cgggttttca cggcgcagag 480

ctggcggaca ttgaactggg cgctacggta gccgtactgg gcatcggtcc ggtgggcctg 540

atggcagttg caggcgctaa gctgcgcggc gcaggtcgta ttattgccgt tggttctcgc 600

ccggtgtgtg tggacgccgc taagtattat ggtgcaacgg acattgtcaa ttacaaggac 660

ggcccaattg aatctcagat catgaacctg acggaaggta aaggcgttga cgccgcgatt 720

atcgctggcg gcaacgccga catcatggcg accgcagtta aaatcgtcaa gccaggtggt 780

actattgcta acgtcaacta cttcggcgaa ggtgaggtcc tgcctgtccc acgtctggaa 840

tggggttgcg gtatggcaca taaaaccatt aaaggtggcc tgtgcccagg cggccgtctg 900

cgtatggaac gcctgatcga tctggtcttc tacaaacgcg tggatcctag caaactggtg 960

actcacgttt tccgcggctt tgataacatc gaaaaagctt ttatgctgat gaaagataaa 1020

ccgaaagatc tgattaaacc ggttgtcatc ctggct 1056

<210>136

<211>352

<212>PRT
<213> 人工序列

<220>
<223> 布氏热厌氧菌的 ADH-TB

<400>136

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

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

<210>137
<211>1125
<212>DNA
<213> 人工序列

<220>
<223> 马肝 (Horse Liver) 的密码子优化的 ADH,E 型

<400>137
atgagcacag cagggaaagt aattaaatgc aaagcagctg tgctgtggga ggaaaagaaa 60
ccattttcca ttgaggaggt ggaggtcgca cctccgaagg cccatgaagt ccgtattaag 120
atggtggcca cagggatctg tcgctcagat gaccacgtgg ttagtgggac cctggtcaca 180
cctctgcctg tgatcgcagg ccatgaggca gctggcattg tggagagcat cggggaaggc 240
gtcactacag tacgtccagg tgataaagtc atcccactgt ttactcctca gtgtgggaaa 300
tgccgtgttt gtaaacaccc tgaaggcaac ttctgcttga aaaatgatct gagcatgcct 360
cgtgggacca tgcaggatgg taccagccgt ttcacctgcc gtgggaagcc tatccaccac 420
ttcctgggca ccagcacctt ctcccagtac accgtggtgg acgagatctc agtggccaag 480
atcgatgcgg cctcaccgct ggagaaagtc tgtctgattg gctgtgggtt ttctactggt 540
tatgggtctg cagtcaaggt tgccaaggtc acccagggct ccacctgtgc cgtgtttggc 600
ctgggggggg tgggcctgtc tgttatcatg ggctgtaaag cagccggggc ggcccgtatc 660
attggggtgg acatcaacaa agacaagttt gcaaaggcca aagaagtggg tgccactgag 720
tgtgtcaacc ctcaggacta caagaaacct atccaggagg tgctgacaga aatgagcaat 780
gggggcgtgg atttttcatt tgaagtcatt ggtcgtctgg acaccatggt gactgccttg 840
tcatgctgtc aagaagcata tggtgtgagc gtcatcgtag gggtacctcc tgattcccaa 900
aatctgtcta tgaatcctat gttgctgctg agtgggcgta cctggaaagg ggctattttt 960
ggtggtttta agagtaaaga ttctgtccct aaactggtgg ccgattttat ggctaaaaag 1020

tttgcactgg atcctttaat cacccatgtt ttacctttg aaaaaattaa tgaagggttt 1080
gacctgctgc gctctgggga gagtatccgt accatcctga cgttt 1125

<210>138
<211>375
<212>PRT
<213> 人工序列

<220>
<223> 马肝的 ADH, E 型

<400>138

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

210 215 220

Ile Asn Lys Asp Lys Phe Ala Lys Ala Lys Glu Val Gly Ala Thr Glu
225 230 235 240
Cys Val Asn Pro Gln Asp Tyr Lys Lys Pro Ile Gln Glu Val Leu Thr
245 250 255
Glu Met Ser Asn Gly Gly Val Asp Phe Ser Phe Glu Val Ile Gly Arg
260 265 270
Leu Asp Thr Met Val Thr Ala Leu Ser Cys Cys Gln Glu Ala Tyr Gly
275 280 285
Val Ser Val Ile Val Gly Val Pro Pro Asp Ser Gln Asn Leu Ser Met
290 295 300
Asn Pro Met Leu Leu Leu Ser Gly Arg Thr Trp Lys Gly Ala Ile Phe
305 310 315 320
Gly Gly Phe Lys Ser Lys Asp Ser Val Pro Lys Leu Val Ala Asp Phe
325 330 335
Met Ala Lys Lys Phe Ala Leu Asp Pro Leu Ile Thr His Val Leu Pro
340 345 350
Phe Glu Lys Ile Asn Glu Gly Phe Asp Leu Leu Arg Ser Gly Glu Ser
355 360 365
Ile Arg Thr Ile Leu Thr Phe
370 375

<210>139
<211>252
<212>PRT
<213> 人工序列

<220>
<223> 具有微小乳杆菌主链的人工改造的酮还原酶序列式
<223> 合成的构建体

<220>
<221>MISC_FEATURE
<222>(7)..(7)
<223>Xaa 是受限的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(17)..(17)

<223>Xaa 是非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(25)..(25)
<223>Xaa 是极性的或酸性的残基

<220>
<221>MISC_FEATURE
<222>(27)..(27)
<223>Xaa 是芳香族的残基

<220>
<221>MISC_FEATURE
<222>(41)..(41)
<223>Xaa 是极性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(53)..(53)
<223>Xaa 是极性的、非极性的或酸性的残基

<220>
<221>MISC_FEATURE
<222>(57)..(57)
<223>Xaa 是非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(64)..(64)
<223>Xaa 是极性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(66)..(66)
<223>Xaa 是酸性的残基

<220>
<221>MISC_FEATURE

<222>(68)..(68)
<223>Xaa 是极性的、酸性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(72)..(72)
<223>Xaa 是碱性的残基

<220>
<221>MISC_FEATURE
<222>(76)..(76)
<223>Xaa 是脂肪族的、非极性的或极性的残基

<220>
<221>MISC_FEATURE
<222>(80)..(80)
<223>Xaa 是脂肪族的、非极性的或极性的残基

<220>
<221>MISC_FEATURE
<222>(93)..(93)
<223>Xaa 是极性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(94)..(94)
<223>Xaa 是半胱氨酸、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(95)..(95)
<223>Xaa 是脂肪族的、非极性的或酸性的残基

<220>
<221>MISC_FEATURE
<222>(96)..(96)
<223>Xaa 是半胱氨酸、极性的、脂肪族的或非极性的残基

<220>

<221>MISC_FEATURE
<222>(97)..(97)
<223>Xaa 是受限的或极性的残基

<220>
<221>MISC_FEATURE
<222>(100)..(100)
<223>Xaa 是碱性的、酸性的、极性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(108)..(108)
<223>Xaa 是碱性的或受限的残基

<220>
<221>MISC_FEATURE
<222>(111)..(111)
<223>Xaa 是非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(139)..(139)
<223>Xaa 是非极性的或脂肪族的残基
<220>
<221>MISC_FEATURE
<222>(145)..(145)
<223>Xaa 是酸性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(147)..(147)
<223>Xaa 是芳香族的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(151)..(151)
<223>Xaa 是受限的、非极性的或脂肪族的残基

<220>

<221>MISC_FEATURE
<222>(152)..(152)
<223>Xaa 是极性的、芳香族的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(163)..(163)
<223>Xaa 是非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(165)..(165)
<223>Xaa 是极性的、非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(173)..(173)
<223>Xaa 是酸性的、非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(179)..(179)
<223>Xaa 是芳香族的、碱性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(190)..(190)
<223>Xaa 是受限的残基
<220>
<221>MISC_FEATURE
<222>(194)..(194)
<223>Xaa 是受限的、极性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(195)..(195)
<223>Xaa 是碱性的、酸性的、非极性的或极性的残基

<220>

<221>MISC_FEATURE
<222>(196)..(196)
<223>Xaa 是非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(197)..(197)
<223>Xaa 是酸性的残基

<220>
<221>MISC_FEATURE
<222>(198)..(198)
<223>Xaa 是酸性的或极性的残基

<220>
<221>MISC_FEATURE
<222>(202)..(202)
<223>Xaa 是芳香族的残基

<220>
<221>MISC_FEATURE
<222>(210)..(210)
<223>Xaa 是极性的或碱性的残基

<220>
<221>MISC_FEATURE
<222>(211)..(211)
<223>Xaa 是碱性的、脂肪族的、非极性的、酸性的或极性的残基

<220>
<221>MISC_FEATURE
<222>(212)..(212)
<223>Xaa 是极性的、脂肪族的或非极性的残基

<220>
<221>MISC_FEATURE
<222>(216)..(216)
<223>Xaa 是受限的或碱性的残基

<220>
<221>MISC_FEATURE
<222>(223)..(223)
<223>Xaa 是非极性的或脂肪族的残基

<220>
<221>MISC_FEATURE
<222>(226)..(226)
<223>Xaa 是非极性的或极性的残基

<220>
<221>MISC_FEATURE
<222>(233)..(233)
<223>Xaa 是极性的或酸性的残基

<220>
<221>MISC_FEATURE
<222>(235)..(235)
<223>Xaa 是极性的或芳香族的残基

<220>
<221>MISC_FEATURE
<222>(248)..(248)
<223>Xaa 是非极性的或芳香族的残基

<220>
<221>MISC_FEATURE
<222>(249)..(249)
<223>Xaa 是芳香族的残基

<400>139

```
Met Thr Asp Arg Leu Lys Xaa 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 Xaa Asp Val Gly Glu Lys Ala Ala
        35                  40                  45
Arg Ser Ile Gly Xaa Thr Asp Val Ile Arg Phe Val Gln His Asp Xaa
    50                  55                  60
```

Ser Asp Glu Xaa Gly Trp Thr Lys Leu Phe Asp Xaa Thr Glu Glu Xaa
65 70 75 80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Xaa Xaa Xaa Xaa
85 90 95
Xaa Ser Val Xaa Asp Thr Thr Thr Glu Glu Trp Xaa 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
Xaa Gly Xaa Val Gly Asp Pro Xaa 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 Xaa Asp Val Arg Val Asn Thr Val His Pro Gly Xaa Ile Lys
180 185 190
Thr Xaa Xaa Val Asp Xaa Leu Glu Gly Xaa Glu Glu Met Met Ser Gln
195 200 205
Arg Xaa Xaa Xaa Pro Met Gly Xaa Ile Gly Glu Pro Asn Asp Ile Ala
210 215 220
Trp Ile Cys Val Tyr Leu Ala Ser Xaa Glu Xaa Lys Phe Ala Thr Gly
225 230 235 240
Ala Glu Phe Val Val Asp Gly Xaa Tyr Thr Ala Gln
245 250

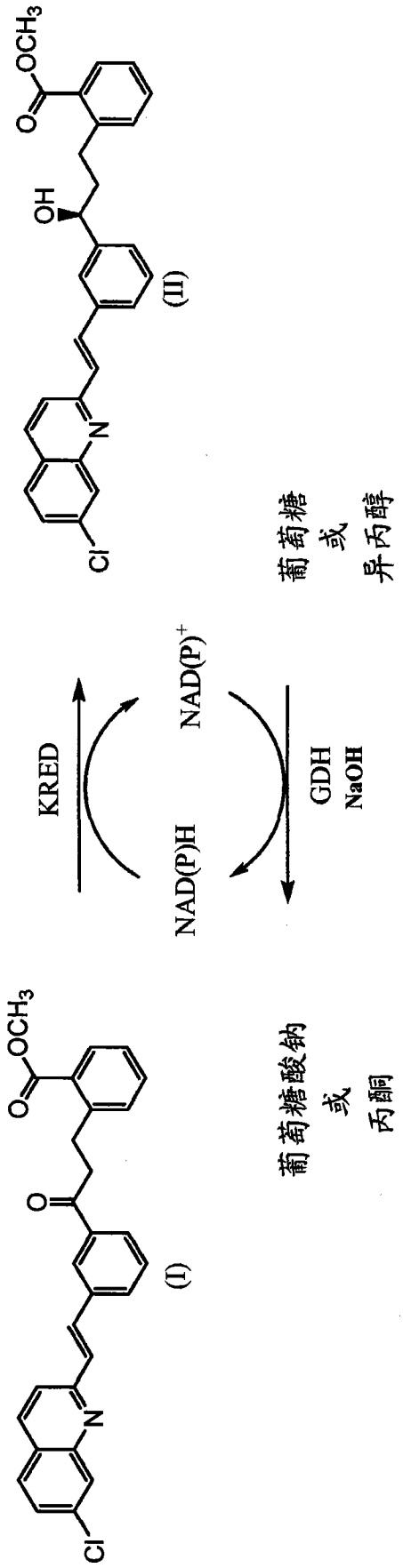

OCH3
O
OH
(II)
N
Cl
葡萄糖
或
异丙醇
KRED
NAD(P)+
NAD(P)H
GDH
NaOH
OCH3
O
O
(I)
N
Cl
葡萄糖酸钠
或
丙酮

图 1