(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 115052982 A
(43)申请公布日 2022.09.13

(21)申请号 202080094930.1
(22)申请日 2020.12.03
(30)优先权数据
1917688.2 2019.12.04 GB
(85)PCT国际申请进入国家阶段日
2022.07.28
(86)PCT国际申请的申请数据
PCT/EP2020/084491 2020.12.03
(87)PCT国际申请的公布数据
WO2021/110848 EN 2021.06.10
(71)申请人 奇华顿股份有限公司
地址 瑞士韦尔涅
(72)发明人 E·艾克霍恩 C·阿尔曼
(74)专利代理机构 中国贸促会专利商标事务所有限公司 11038
专利代理师 李程达
(51)Int.Cl.
C12N 9/90(2006.01)
C12P 17/04(2006.01)

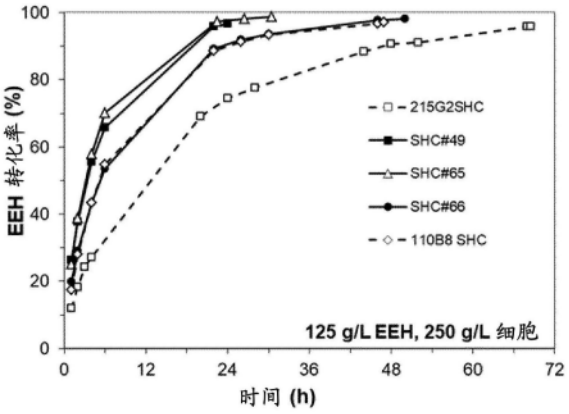
权利要求书3页 说明书117页
序列表58页 附图13页

(54)发明名称

角鲨烯蒎烯环化酶(SHC)变体

(57)摘要

角鲨烯蒎烯环化酶(SHC)酶及其变体及其用于从均法呢醇制备(-)-降龙涎香醚和从双均法呢醇制备龙涎香氧化物的用途。



CN 115052982 A

1.一种制备(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物的方法,所述方法包括使用SHC/HAC酶变体将(3E,7E)-均法呢醇(EEH)或包含EEH的均法呢醇异构体的混合物酶促转化为(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物,

其中SHC/HAC酶变体具有与SEQ ID NO:1具有至少约70.0%同一性的氨基酸序列,

其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置132、224和432的位置处具有相对于SEQ ID NO:1的氨基酸改变,其分别为M132R、A224V和I432T,和

其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557的位置处和对应于SEQ ID NO:1的位置81、431或613的至少一个位置处具有相对于SEQ ID NO:1的氨基酸改变。

2.一种制备龙涎香氧化物或包含龙涎香氧化物的混合物的方法,所述方法包括使用SHC/HAC酶变体将(E,E)-双均法呢醇(BisEEH)或包含BisEEH的双均法呢醇异构体的混合物酶促转化为龙涎香氧化物或包含龙涎香氧化物的混合物,

其中SHC/HAC酶变体具有与SEQ ID NO:1具有至少约70.0%同一性的氨基酸序列,

其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置132、224和432的位置处具有相对于SEQ ID NO:1的氨基酸改变,其分别为M132R、A224V和I432T,和

其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557的位置处和对应于SEQ ID NO:1的位置81、431或613的至少一个位置处具有相对于SEQ ID NO:1的氨基酸改变。

3.一种SHC/HAC酶变体,其具有与SEQ ID NO:1具有至少约70.0%同一性的氨基酸序列,

其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置132、224和432的位置处具有相对于SEQ ID NO:1的氨基酸改变,其分别为M132R、A224V和I432T,和

其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557的位置处和对应于SEQ ID NO:1的位置81、431或613的至少一个位置处具有相对于SEQ ID NO:1的氨基酸改变。

4.权利要求1或2的方法或权利要求3的SHC/HAC酶变体,其中SHC/HAC酶变体具有与SEQ ID NO:1具有至少约90.0%同一性的氨基酸序列。

5.前述权利要求中任一项的方法或SHC/HAC酶变体,其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557和431的位置处具有相对于SEQ ID NO:1的氨基酸改变。

6.前述权利要求中任一项的方法或SHC/HAC酶变体,其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557和613的位置处具有相对于SEQ ID NO:1的氨基酸改变。

7.权利要求5或6的方法或SHC/HAC酶变体,其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置81的位置处具有相对于SEQ ID NO:1的进一步氨基酸改变。

8.前述权利要求中任一项的方法或SHC/HAC酶变体,其中在位置81、431、557或613处的氨基酸改变中的一个或多个是取代,例如非保守取代。

9.前述权利要求中任一项的方法或SHC/HAC酶变体,其中:

对应于SEQ ID NO:1的位置557的位置处的氨基酸改变将SEQ ID NO:1的氨基酸取代为中性亲水性氨基酸,例如苏氨酸;和/或

对应于SEQ ID NO:1的位置81的位置处的氨基酸改变将SEQ ID NO:1的氨基酸取代为

碱性氨基酸,例如组氨酸;和/或

对应于SEQ ID NO:1的位置431的位置处的氨基酸改变将SEQ ID NO:1的氨基酸取代为疏水性氨基酸,例如亮氨酸;和/或

对应于SEQ ID NO:1的位置613的位置处的氨基酸改变将SEQ ID NO:1的氨基酸取代为中性亲水性氨基酸,例如丝氨酸。

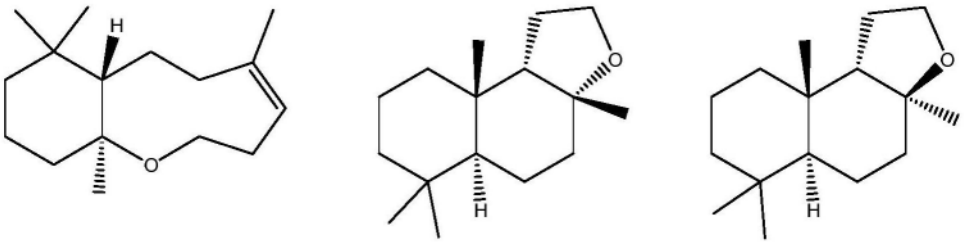
10.前述权利要求中任一项的方法或SHC/HAC酶变体,其中SHC/HAC酶变体具有选自SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4和SEQ ID NO:5的氨基酸序列,例如选自SEQ ID NO:2、SEQ ID NO:4和SEQ ID NO:5的氨基酸序列。

11.前述权利要求中任一项的方法,其中所述酶促转化在约30℃至约50℃,例如约40℃至约50℃的温度和/或约5至约6的pH下发生。

12.前述权利要求中任一项的方法,其中所述方法包括培养产生SHC/HAC酶变体的重组宿主细胞。

13.权利要求12的方法,其中所述重组宿主细胞包含选自SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8或SEQ ID NO:9的核酸序列,例如选自SEQ ID NO:6、SEQ ID NO:8或SEQ ID NO:9的核酸序列。

14.前述权利要求中任一项的方法,其中(-)-降龙涎香醚与副产物(II)、(III)和(IV)中的至少一种或多种混合产生,



(II) (III) (IV)

15.前述权利要求中任一项的方法或SHC/HAC酶变体,其中SHC/HAC酶变体相对于SEQ ID NO:1具有增加的酶活性。

16.前述权利要求中任一项的方法或SHC/HAC酶变体,其中SHC/HAC酶变体在反应的前6小时或前8小时或前12小时内提供与SEQ ID NO:1和/或SEQ ID NO:10相比增加的EEH转化率。

17.通过前述权利要求中任一项的方法获得或可通过前述权利要求中任一项的方法获得的(-)-降龙涎香醚,例如以无定形或结晶形式获得。

18.权利要求17的(-)-降龙涎香醚作为香料或化妆品或消费品例如织物护理、洗护用品、美容护理、清洁产品、洗涤剂产品和/或肥皂产品的一部分的用途。

19.包含权利要求17的(-)-降龙涎香醚的香料或化妆品或消费品。

20.编码前述权利要求中任一项的SHC/HAC酶变体的核酸序列。

21.权利要求20的核酸序列,其中所述核酸序列选自SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8和SEQ ID NO:9,例如选自SEQ ID NO:6、SEQ ID NO:8和SEQ ID NO:9的核酸序列。

22.包含权利要求20或21的核酸序列的构建体。

23.包含权利要求22的构建体的载体。

24.包含权利要求20或21的核酸序列、权利要求22的构建体或权利要求23的载体的重

组宿主细胞。

25.权利要求24的重组宿主细胞，其中所述重组宿主细胞是具有选自埃希氏菌属(Escherichia)、链霉菌属(Streptomyces)、芽孢杆菌属(Bacillus)、假单胞菌属(Pseudomonas)、乳杆菌属(Lactobacillus)和乳球菌属(Lactococcus)的属的细菌，例如其中所述重组宿主细胞是大肠杆菌(E.coli.)。

26.权利要求1-16中任一项的方法，其中包含EEH的异构体的混合物是EE:EZ异构体混合物。

27.权利要求26的方法，其中均法呢醇EE:EZ异构体混合物的重量比选自EE:EZ 90:10;EE:EZ 80:20;EE:EZ 86:14;EE:EZ 70:30;EE:EZ 69:31;和EE:EZ 66:34，或均法呢醇EE:EZ异构体混合物的重量比为EE:EZ 90:10;EE:EZ 80:20;EE:EZ 86:14;EE:EZ 70:30;EE:EZ 69:31;或EE:EZ 66:34。

28.权利要求27的方法，其中EE:EZ异构体混合物的重量比为80:20。

29.权利要求1-16或26-28中任一项的方法，其中SHC/HAC生物催化剂与EEH或包含EEH的均法呢醇异构体的混合物的重量比为约0.5-2:1或约0.25-2:1或约0.1-2:1或约1:1或约0.5:1。

30.权利要求29的方法，其中SHC/HAC生物催化剂与EEH或包含EEH的均法呢醇异构体的混合物的重量比为约1:1或约0.5:1或约0.1:1，所述混合物为80:20重量比的EE:EZ异构体混合物。

角鲨烯藿烯环化酶(SHC)变体

技术领域

[0001] 本发明总体上涉及SHC/HAC酶及其变体。本发明还涉及SHC/HAC酶及其变体的各种用途，例如将(3E,7E)-均法呢醇(homofarnesol)(EEH)酶促转化为(-)-降龙涎香醚(Ambrox)或将E,E-双均法呢醇(bishomofarnesol)(BisEEH)酶促转化为龙涎香氧化物(Ambra oxide)。本发明还涉及酶促反应的产物，例如使用SHC/HAC酶及其变体制备的(-)-降龙涎香醚或龙涎香氧化物，以及所述产物的各种用途。

背景技术

[0002] 角鲨烯藿烯环化酶(SHC)是膜结合酶，其作为生物催化剂将线性三萜类角鲨烯环化为藿烯和藿醇。

[0003] 来自多种细菌的许多野生型和变体SHC酶已被证明可用于将(3E,7E)-均法呢醇转化为(-)-降龙涎香醚(参见，例如WO 2016/170099;WO 2018/157021;Neumann&Simon 1986,Biol Chem Hoppe-Seyler367,723-729;JP2009060799;Seckler&Poralla 1986,Biochem Biophys Act 356-363;Ochs等人1990,J Bacteriol 174,298-302;WO 2010/139719;US 8759043;WO 2012/066059;Seitz等人2012,J Molecular Catalysis B:Enzymatic 84,72-77;和Seitz 2012博士论文(http://elib.uni-stuttgart.de/handle/1 1682/1400)，其内容通过引用并入本文)。需要提供新的和改进的制备(-)-降龙涎香醚的方法，例如使用新的SHC酶或酶变体。还希望提供用于环化其他底物的新的和改进的方法，例如以形成可用于或用作香料的化合物。

发明内容

[0004] 根据本发明的第一方面，提供了制备(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物的方法，所述方法包括使用SHC/HAC酶变体将(3E,7E)-均法呢醇(EEH)或包含EEH的均法呢醇异构体的混合物酶促转化为(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物，

[0005] 其中SHC/HAC酶变体具有与SEQ ID NO:1具有至少约70.0％同一性的氨基酸序列，

[0006] 其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置132、224和432的位置处具有相对于SEQ ID NO:1的氨基酸改变，分别为M132R、A224V和I432T，和

[0007] 其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557的位置处和对应于SEQ ID NO:1的位置81、431或613的至少一个位置处具有相对于SEQ ID NO:1的氨基酸改变。

[0008] 根据本发明的第二方面，提供了通过本发明第一方面的方法获得或可通过本发明第一方面的方法获得的(-)-降龙涎香醚，包括其任何实施方案。

[0009] 根据本发明的第三方面，提供了本发明第二方面的(-)-降龙涎香醚(包括其任何实施方案)作为香料或化妆品或消费品的一部分的用途。

[0010] 根据本发明的第四方面，提供了包含本发明第二方面的(-)-降龙涎香醚(包括其任何实施方案)的香料或化妆品或消费品。

[0011] 根据本发明的第五方面，提供了一种SHC/HAC酶变体，其氨基酸序列与SEQ ID NO:1具有至少约70.0％的同一性，

[0012] 其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置132、224和432的位置处具有相对于SEQ ID NO:1的氨基酸改变，分别为M132R、A224V和I432T，和

[0013] 其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557的位置处和对应于SEQ ID NO:1的位置81、431或613的至少一个位置处具有相对于SEQ ID NO:1的氨基酸改变。

[0014] 根据本发明的第六方面，提供了编码本发明的第五方面的SHC/HAC酶变体的核酸序列，包括其任何实施方案。

[0015] 根据本发明的第七方面，提供了包含本发明的第六方面的核酸序列的构建体，包括其任何实施方案。

[0016] 根据本发明的第八方面，提供了包含本发明的第七方面的构建体的载体，包括其任何实施方案。

[0017] 根据本发明的第九方面，提供了包含本发明的第六方面的核酸序列、本发明的第七方面的构建体或本发明的第八方面的载体的重组宿主细胞，包括其任何实施方案。

[0018] 根据本发明的第十方面，提供了制备(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物的方法，所述方法包括使用SHC/HAC酶变体将(3E,7E)-均法呢醇(EEH)或包含EEH的均法呢醇异构体的混合物酶促转化为(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物，其中SHC/HAC酶变体具有与野生型SHC/HAC酶氨基酸序列具有至少约70.0％同一性的氨基酸序列，并且其中SHC/HAC酶变体氨基酸序列在选自对应于SEQ ID NO:1的位置557、81、431和613的位置的位置处具有相对于野生型SHC/HAC酶的一个或多个氨基酸改变。

[0019] 根据本发明的第十一方面，提供了一种SHC/HAC酶变体，其中该SHC/HAC酶变体具有与野生型SHC/HAC酶氨基酸序列具有至少约70.0％同一性的氨基酸序列，并且其中SHC/HAC酶变体氨基酸序列在选自对应于SEQ ID NO:1的位置557、81、431和613的位置的位置处具有相对于野生型SHC/HAC酶的一个或多个氨基酸改变。

[0020] 根据本发明的第十二方面，提供了制备(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物的方法，所述方法包括使用SHC/HAC酶或SHC/HAC酶变体将(3E,7E)-均法呢醇(EEH)或包含EEH的均法呢醇异构体的混合物酶促转化为(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物，其中SHC/HAC酶或SHC/HAC酶变体具有与野生型SHC/HAC酶氨基酸序列具有至少约70.0％同一性的氨基酸序列。

[0021] 根据本发明的第十三方面，提供了一种SHC/HAC酶或SHC/HAC酶变体，其中SHC/HAC酶或SHC/HAC酶变体具有与野生型SHC/HAC酶氨基酸序列具有至少约70.0％同一性的氨基酸序列。

[0022] 根据本发明的第十四方面，提供了一种制备龙涎香氧化物的方法，所述方法包括使用SHC/HAC酶或SHC/HAC酶变体将E,E-双均法呢醇或包含E,E-双均法呢醇的双均法呢醇异构体的混合物酶促转化为龙涎香氧化物或包含龙涎香氧化物的混合物。例如，SHC/HAC酶或SHC/HAC酶变体可以根据本发明的任何方面。

[0023] 根据本发明的第十五方面，提供了通过本发明的第十四方面的方法获得或可通过本发明的第十四方面的方法获得的龙涎香氧化物，包括其任何实施方案。

[0024] 根据本发明的第十六方面，提供了本发明的第十五方面的龙涎香氧化物(包括其任何实施方案)作为香料或化妆品或消费品的一部分的用途。

[0025] 根据本发明的第十七方面，提供了包含本发明的第十五方面的龙涎香氧化物的香料或化妆品或消费品，包括其任何实施方案。

[0026] 在本发明任何方面的某些实施方案中，SHC/HAC酶变体具有选自SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:17、SEQ ID NO:18、SEQ ID NO:24、SEQ ID NO:25、SEQ ID NO:26、SEQ ID NO:27、SEQ ID NO:28、SEQ ID NO:29和/或SEQ ID NO:30的氨基酸序列。

[0027] 在本发明任何方面的某些实施方案中，SHC/HAC酶变体由具有选自SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:22和SEQ ID NO:23的序列的核酸编码。

[0028] 本发明的某些实施方案可以提供以下优点中的一个或多个：

[0029] • 提高的EEH转化率，特别是在前12小时或前6小时；

[0030] • 新底物的环化和/或新产品的鉴定；

[0031] • 提高的生产率(g/l/小时或g/l/小时/g生物催化剂)；

[0032] • 用于制备最终产品例如(-)-降龙涎香醚和龙涎香氧化物和/或其异构体混合物的工业规模方法；

[0033] • 与其他均法呢醇异构体相比，针对EEH的提高的选择性。

[0034] 提供的与本发明的任何特定一个或多个所述方面相关的细节、实例和偏好将在本文中进一步描述并且同样适用于本发明的所有方面。除非本文另有说明或与上下文明显矛盾，否则本文所述的实施方案、实例和偏好的所有可能变化形式的任何组合都包含在本发明中。

附图说明

[0035] 图1显示了SHC变体在初始和优化反应条件下的相对活性。在初始条件(35℃，pH5.4，0.070％SDS)或在设置为针对每个变体定义为最佳的条件(优化条件)的T、pH和[SDS]下，在4g/l EEH下进行反应，并将生物催化剂加载到10.0的OD_{650nm}。

[0036] 图2显示了SHC变体在优化条件下与亲本215G2 SHC酶相比的相对活性。在4g/l EEH下进行反应，并将生物催化剂加载到10.0的OD_{650nm}，T、pH和[SDS]设置为针对每个变体定义为最佳的条件。

[0037] 图3显示了与215G2亲本SHC/HAC酶相比，在最佳条件下通过SHC/HAC酶变体将EEH转化为(-)-降龙涎香醚。反应在125g/l EEH和250g/l生物催化剂下、在1.3％SDS的存在下和在针对每个变体定义为最佳的T和pH条件下进行。

[0038] 图4显示了与参考215G2 SHC/HAC亲本酶相比，使用最佳SHC/HAC酶变体的EEH转化率的相对改善。反应在125g/l EEH和250g/l生物催化剂下、在1.3％SDS的存在下和在针对每个变体定义为最佳的T和pH条件下进行。在所有时间点，使用215G2 SHC获得的EEH转化率设置为100％(参考)。

[0039] 图5显示了与亲本(参考)215G2 SHC/HAC酶相比，SHC/HAC变体在最佳条件下的相对活性。反应在4g/l BisEEH下进行，并将生物催化剂加载到10.0的OD_{650nm}，T、pH和[SDS]设

置为针对每个变体定义为最佳的条件。

[0040] 图6显示了使用亲本(参考)215G2 SHC/HAC酶和新变体SHC/HAC酶的双均法呢醇的环化。反应在125g/l BisEEH和250g/l生物催化剂下、在1.3％SDS的存在下和在35℃和pH5.4(所有SHC酶)下进行。

[0041] 图7显示了215G2 SHC/HAC(SEQ ID NO:10)和SHC/HAC酶变体#49(SEQ ID NO:2)、#65(SEQ ID NO:3)、#66(SEQ ID NO:4)、#90C7(SEQ ID NO:17)、#110B8(SEQ ID NO:5)和#115A7(SEQ ID NO:18)之间的氨基酸序列比对。

[0042] 图8显示了215G2(SEQ ID NO:10)和衍生自215G2的具有V174I、F601Y和L37Q取代中的一个或多个的SHC/HAC酶变体之间的氨基酸序列比对。

[0043] 图9A显示了使用CLUSTAL O(1.2.4)在野生型AacSHC(SEQ ID NO:1)和野生型TelSHC(SEQ ID NO:19)、野生型ApaSHC1(SEQ ID NO:20)、野生型ZmoSHC1(SEQ ID NO:11)、野生型ZmoSHC2(SEQ ID NO:12)、野生型BjpSHC(SEQ ID NO:13)和野生型GmoSHC(SEQ ID NO:14)之间的氨基酸序列比对。野生型AacSHC中的氨基酸位置81、90、132、224、172、277、431、432、557和613用黑色背景上的白色字母突出显示。

[0044] 图9B显示了使用CLUSTAL O(1.2.4)在野生型AacSHC(SEQ ID NO:1)和野生型TelSHC(SEQ ID NO:19)、野生型ApaSHC1(SEQ ID NO:20)、野生型ZmoSHC1(SEQ ID NO:11)、野生型ZmoSHC2(SEQ ID NO:12)、野生型BjpSHC(SEQ ID NO:13)、野生型GmoSHC(SEQ ID NO:14)、野生型BmeSHC(SEQ ID NO:28)、野生型SalSHC(SEQ ID NO:29)和野生型ApaSHCA(SEQ ID NO:30)之间的氨基酸序列比对。野生型AacSHC中的氨基酸位置81、90、132、224、172、277、431、432、557和613用黑色背景上的白色字母突出显示。

[0045] 图10是用于产生E,E/Z-双均法呢醇的起始混合物的反应方案。

[0046] 图11是显示使用SHC酶的双均法呢醇的环化的反应示意图。所用的双均法呢醇由异构体的混合物组成。

[0047] 图12显示了与215G2亲本SHC/HAC酶相比,在最佳条件下通过SHC/HAC酶变体将EEH转化为(-)-降龙涎香醚。反应在125g/l EEH和250g/l生物催化剂下、在1.3％SDS的存在下和在针对每个变体定义为最佳的T和pH条件下进行。

[0048] 图13显示了与215G2亲本SHC/HAC酶相比,在最佳条件下通过SHC/HAC酶变体将EEH转化为(-)-降龙涎香醚。反应在125g/l EEH和125g/l生物催化剂下、在0.65％SDS的存在下和在针对每个变体定义为最佳的T和pH条件下进行。

[0049] 图14显示了与215G2亲本SHC/HAC酶相比,在最佳条件下通过SHC/HAC酶变体SHC#65将EEH转化为(-)-降龙涎香醚。反应在125g/l至300g/l EEH和250g/l生物催化剂或1的[EEH]:[细胞]比率下进行。SDS以恒定的0.052的[SDS]:[细胞]比率提供。反应在针对SHC变体SHC#65定义为最佳的T和pH条件下进行。

[0050] 图15显示通过野生型AacSHC(SEQ ID NO:1)、野生型BjpSHC(SEQ ID NO:13)、野生型BmeSHC(SEQ ID NO:28)、野生型GmoSHC(SEQ ID NO:14)、野生型ApaSHCA(SEQ ID NO:30)、野生型ApaSHC1(SEQ ID NO:20)、野生型SalSHC(SEQ ID NO:29)、野生型TelSHC(SEQ ID NO:19)、野生型ZmoSHC1(SEQ ID NO:11)和野生型ZmoSHC2(SEQ ID NO:12)将EEH转化至为(-)-降龙涎香醚。所有列出的野生型SHC酶在4g/l EEH下进行反应,以及除野生型ApaSHCA(SEQ ID NO:30)、野生型SalSHC(SEQ ID NO:29)、野生型TelSHC(SEQ ID NO:19)和

野生型ZmoSHC2(SEQ ID NO:12)之外的所有列出的野生型SHC酶在125g/l EEH下进行反应。反应在针对每种列出的野生型SHC酶定义为最佳的T和pH条件和SDS浓度下进行。在125g/l EEH生物转化中调整SDS浓度以允许最佳生物催化剂活性。

[0051] 图16显示了通过SHC/HAC将EEH转化为(-)-降龙涎香醚。使用SHC#65生物催化剂进行生物转化。反应在使用150g/l细胞的125g/l EEH下、在45℃和pH5.6下和在对应于0.040至0.060的[SDS]:[细胞]比率的0.060至0.090％SDS的存在下进行。显示了在反应的72小时时的EEH转化率值。

[0052] 序列概述

[0053] SEQ ID NO:1是野生型酸热脂环酸杆菌(Alicyclobacillus acidocaldarius)(Aac)SHC氨基酸序列。

[0054] SEQ ID NO:2对应于具有取代M132R、A224V、I432T、A557T和H431L的SEQ ID NO:1，并且在本文中可以称为SHC/HAC酶变体#49。

[0055] SEQ ID NO:3对应于具有取代M132R、A224V、I432T、A557T和R613S的SEQ ID NO:1，并且在本文中可以称为SHC/HAC酶变体#65。

[0056] SEQ ID NO:4对应于具有取代M132R、A224V、I432T、Y81H、A557T和R613S的SEQ ID NO:1,并且在本文中可以称为SHC/HAC酶变体#66。

[0057] SEQ ID NO:5对应于具有取代M132R、A224V、I432T、Y81H、H431L和A557T的SEQ ID NO:1,并且在本文中可以称为SHC/HAC酶变体#110B8。

[0058] SEQ ID NO:6是编码SEQ ID NO:2的多肽(SHC/HAC酶变体#49)的核酸序列。

[0059] SEQ ID NO:7是编码SEQ ID NO:3的多肽(SHC/HAC酶变体#65)的核酸序列。

[0060] SEQ ID NO:8是编码SEQ ID NO:4的多肽(SHC/HAC酶变体#66)的核酸序列。

[0061] SEQ ID NO:9是编码SEQ ID NO:5的多肽(SHC/HAC酶变体#110B8)的核酸序列。

[0062] SEQ ID NO:10可称为215G2,并且对应于具有突变M132R、A224V和I432T的野生型AacSHC氨基酸序列。

[0063] SEQ ID NO:11是ZmoSHC1的野生型氨基酸序列。

[0064] SEQ ID NO:12是ZmoSHC2的野生型氨基酸序列。

[0065] SEQ ID NO:13是BjpSHC/BjaSHC的野生型氨基酸序列。

[0066] SEQ ID NO:14是GmoSHC的野生型氨基酸序列。

[0067] SEQ ID NO:15是编码野生型AacSHC的核苷酸序列。

[0068] SEQ ID NO:16是编码215G2 SHC的核苷酸序列。

[0069] SEQ ID NO:17对应于具有取代M132R、A224V、I432T、T90A和R613S的SEQ ID NO:1，并且在本文中可以称为SHC/HAC酶变体#90C7。

[0070] SEQ ID NO:18对应于具有取代M132R、A224V、I432T、A172T和M277K的SEQ ID NO:1,并且在本文中可以称为SHC/HAC酶变体#115A7。

[0071] SEQ ID NO:19是TelSHC的野生型氨基酸序列。

[0072] SEQ ID NO:20是ApaSHC1的野生型氨基酸序列。

[0073] SEQ ID NO:21是GmoSHC变体。

[0074] SEQ ID NO:22是编码SEQ ID NO:17的多肽(SHC/HAC酶变体#90C7)的核苷酸序列。

[0075] SEQ ID NO:23是编码SEQ ID NO:18的多肽(SHC/HAC酶变体#115A7)的核苷酸序

列。

[0076] SEQ ID NO:24是具有额外突变L37Q的SHC/HAC变体215G2 SHC的氨基酸序列。

[0077] SEQ ID NO:25是具有额外突变V174I的SHC/HAC变体215G2 SHC的氨基酸序列。

[0078] SEQ ID NO:26是具有额外突变V174I和F601Y的SHC/HAC变体215G2 SHC的氨基酸序列。

[0079] SEQ ID NO:27是具有额外突变L37Q、V174I和F601Y的SHC/HAC变体215G2 SHC的氨基酸序列。

[0080] SEQ ID NO:28是BmeSHC的野生型氨基酸序列。

[0081] SEQ ID NO:29是SalSHC的野生型氨基酸序列。

[0082] SEQ ID NO:30是ApaSHCA的野生型氨基酸序列。

[0083] 发明内容

[0084] SHC/HAC酶及其变体

[0085] 如本文所用,术语“SHC酶”是指天然存在于例如嗜热细菌例如酸热脂环酸杆菌中的野生型(WT)角鲨烯霍烯环化酶。在均法呢醇环化为降龙涎香醚中起作用的SHC也可称为均法呢醇降龙涎香醚环化酶(HAC)。因此,本文可以使用术语“SHC/HAC酶”。

[0086] 如本文所用,术语“变体”应理解为与衍生其的多肽相比通过氨基酸序列中的一个或多个变化而不同的多肽。衍生变体的多肽也称为亲本或参考多肽。通常,变体是人工构建的,优选是通过基因技术手段。通常,衍生变体的多肽是野生型蛋白质或野生型蛋白质结构域。然而,可用于本公开内容的变体也可以衍生自亲本多肽的同源物、直系同源物或旁系同源物或衍生自人工构建的变体,条件是该变体表现出亲本多肽的至少一种生物活性。氨基酸序列的变化可以是氨基酸交换(取代)、插入、缺失、N末端截短或C末端截短,或这些变化的任何组合,其可以发生在一个或几个位点。

[0087] 如本文所用,术语“SHC/HAC酶变体”是指衍生自野生型SHC酶但与野生型SHC酶相比具有一个或多个氨基酸改变并且因此不是天然存在于原核生物中的酶。一个或多个氨基酸改变可以例如改变(例如增加)底物(例如EEH)的酶活性。

[0088] 用于确定和定量SHC/HAC酶和/或SHC/HAC酶变体活性的测定法在本文中描述并且是本领域已知的。举例来说,SHC/HAC酶和/或SHC/HAC酶变体活性可以通过将纯化的SHC/HAC酶或酶变体或来自产生SHC/HAC酶或酶变体的宿主细胞或完整重组宿主生物体的提取物与适当的底物在适当的条件下孵育,并对反应产物进行分析(例如通过气相色谱法(GC)或HPLC分析)来确定。实施例中提供了关于SHC/HAC酶和/或SHC/HAC酶变体活性测定和反应产物分析的进一步细节。这些测定包括在重组宿主细胞(例如大肠杆菌(E.coli.))中产生SHC/HAC酶变体。

[0089] 如本文所用,术语“活性”是指酶与底物反应以提供所需产物的能力。活性可以在所谓的用于监测所需产物的形成的活性测试中确定。本公开内容的SHC/HAC酶衍生物的特征可以在于它们将均法呢醇(例如EEH)环化成(-)-降龙涎香醚的能力并表现出生物活性,例如HAC活性。本公开内容的SHC/HAC酶衍生物的特征可以在于它们将双均法呢醇(例如E,E-双均法呢醇)环化成龙涎香氧化物的能力。

[0090] 如本文所用,“生物活性”是指多肽可表现出的任何活性,包括但不限于:酶活性;与另一种化合物的结合活性(例如与另一种多肽的结合,特别是与受体的结合,或与核酸的

结合)；抑制活性(例如酶抑制活性)；活化活性(例如酶活化活性)；或毒性作用。不要求变体表现出与亲本或野生型多肽相同程度的这种活性。如果变体表现出亲本多肽活性的至少10％的程度的相关活性，则变体被视为本申请的上下文中的变体。类似地，如果变体表现出亲本多肽活性的至少10％的程度的相关生物活性，则变体被视为本申请的上下文中的变体(因为术语衍生物和变体在整个本公开内容中可互换使用)。在其他实施方案中，本公开内容的SHC/HAC酶变体显示出比参考SHC蛋白(例如野生型SHC/HAC酶或已知的SHC/HAC酶变体)更好的产率。术语“产率”是指每克原料的可回收产物克数(其可以计算为摩尔转化率百分比)。在另外的实施方案中，本公开内容的SHC/HAC酶变体相对于参考SHC蛋白(例如野生型AacSHC或215G2 AacSHC)显示出改变的(例如增加的)生产率。术语“生产率”是指每小时生物转化时间(即添加底物后的时间)的每升反应容量的以克计的可回收产物的量。术语“生产率”还指在反应中使用的每克生物催化剂的每小时生物转化时间(即添加底物后的时间)的每升反应容量的以克计的可回收产物的量。

[0091] 在进一步的实施方案中，本公开内容的SHC/HAC酶变体与参考SHC蛋白(例如野生型AacSHC(SEQ ID NO:1)或215G2AacSHC(SEQ ID NO:10)或野生型ZmoSHC1(SEQ ID NO:11)或野生型ZmoSHC2(SEQ ID NO:12)或野生型BjpSHC(SEQ ID NO:13)或野生型GmoSHC(SEQ ID NO:14)或野生型TelSHC(SEQ ID NO:19)或野生型ApaSHC1(SEQ ID NO:20)或野生型BmeSHC(SEQ ID NO:28)或野生型SalSHC(SEQ ID NO:29)或野生型ApaSHCA(SEQ ID NO:30))相比显示改良的产率。术语“目标产率因子”是指在反应介质中获得的产物浓度与SHC/HAC变体酶(例如纯化的SHC/HAC酶变体或来自产生SHC/HAC酶变体的重组宿主细胞的提取物)浓度之间的比率。在各种实施方案中，本公开内容的SHC/HAC酶变体相对于参考SHC蛋白(例如SEQ ID NO:1或SEQ ID NO:10或SEQ ID NO:11或SEQ ID NO:12或SEQ ID NO:13或SEQ ID NO:14或SEQ ID NO:19或SEQ ID NO:20或SEQ ID NO:28或SEQ ID NO:29或SEQ ID NO:30)显示酶活性(例如改良/增加的均法呢醇降龙涎香醚环化酶(HAC)活性)的改良(例如增加)的倍数增加。这种活性的增加可以是至少2、3、4、6、8、10、12、14、16、18、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95和/或100倍。

[0092] 如本文所用，术语“氨基酸改变”是指在两个氨基酸之间一个或多个氨基酸的插入、一个或多个氨基酸的缺失或相对于参考氨基酸序列的氨基酸序列的一个或多个不同氨基酸对一个或多个氨基酸的取代(可以是保守的或非保守的)。取代将参考序列的氨基酸替换为变体序列中相同数量的氨基酸。例如，参考氨基酸序列可以是野生型(WT)氨基酸序列(例如SEQ ID NO:1或SEQ ID NO:11或SEQ ID NO:12或SEQ ID NO:13或SEQ ID NO:14或SEQ ID NO:19或SEQ ID NO:20或SEQ ID NO:28或SEQ ID NO:29或SEQ ID NO:30)或例如其本身可以是SHC/HAC酶变体序列(例如Aac 215G2变体-SEQ ID NO:10)。

[0093] 通过将SHC/HAC酶变体的氨基酸序列与参考氨基酸序列的氨基酸序列进行比较，可以容易地识别氨基酸改变。

[0094] 例如，可以基于所涉及的氨基酸残基的极性、电荷、大小、溶解度、疏水性、亲水性和/或两亲性性质的相似性进行保守氨基酸取代。上述20种天然存在的氨基酸可分为以下六个标准氨基酸组：

[0095] (1)疏水性:Met、Ala、Val、Leu、Ile；

[0096] (2)中性亲水性:Cys、Ser、Thr;Asn,Gln;

[0097] (3)酸性:Asp、Glu;

[0098] (4)碱性:His、Lys、Arg;

[0099] (5)影响链取向的残基:Gly、Pro;和

[0100] (6)芳香族:Trp、Tyr、Phe。

[0101] 因此,如本文所用,术语“保守取代”是指氨基酸被上面所示的六个标准氨基酸组的同一组中列出的另一氨基酸取代。例如,Glu对Asp的交换在如此修饰的多肽中保留了一个负电荷。此外,甘氨酸和脯氨酸可以根据它们破坏α-螺旋的能力相互替代。上述六组中的一些优选的保守取代是以下子组中的交换:(i)Ala、Val、Leu和Ile;(ii)Ser和Thr;(ii)Asn和Gln:(iv)Lys和Arg;和(v)Tyr和Phe。鉴于已知的遗传密码以及重组和合成DNA技术,熟练的科学家可以容易地构建编码保守氨基酸变体的DNA。

[0102] 如本文所用,“非保守取代”或“非保守氨基酸交换”定义为氨基酸被如上所示的六个标准氨基酸组(1)至(6)的不同组中列出的另一氨基酸的交换。通常,本文所述的SHC/HAC酶变体使用改变所公开的SHC/HAC酶变体的生物学功能(例如HAC活性)的非保守取代来制备。为便于参考,IUPAC-IUB生化命名委员会推荐的单字母氨基酸符号如下所示。还提供了三字母代码以供参考。

[0103]

一字母代码	三字母代码	氨基酸名称
A	Ala	丙氨酸
C	Cys	半胱氨酸
D	Asp	天冬氨酸
E	Glu	谷氨酸
F	Phe	苯丙氨酸
G	Gly	甘氨酸
H	His	组氨酸
I	Ile	异亮氨酸
K	Lys	赖氨酸
L	Leu	亮氨酸
M	Met	甲硫氨酸
N	Asn	天冬酰胺
P	Pro	脯氨酸
Q	Gln	谷氨酰胺

[0104]

R	Arg	精氨酸
S	Ser	丝氨酸
T	Thr	苏氨酸
V	Val	缬氨酸
W	Trp	色氨酸
Y	Tyr	酪氨酸

[0105] 氨基酸改变可以使用重组基因技术的已知方案引入，例如氨基酸取代，包括PCR、基因克隆、cDNA的定点诱变、宿主细胞的转染和体外转录，其可用于将这些改变引入到参考序列中，产生SHC/HAC酶变体。然后可以筛选酶变体的SHC/HAC功能活性。

[0106] SHC/HAC酶的合适来源包括，例如，酸热脂环酸杆菌(Aac)、运动发酵单胞菌(Zymomonas mobilis)(Zmo)、日本慢生根瘤菌(Bradyrhizobium japonicum)(Bjp)、莫比氏葡萄糖杆菌(Gluconobacter morbifer)(Gmo)、双向伯克霍尔德菌(Burkholderia ambifaria)、炭疽杆菌(Bacillus anthracis)、荚膜甲基球菌(Methylococcus capsulatus)、法兰克氏菌(Frankia alni)、巴氏醋杆菌(Acetobacter pasteurianus)(Apa)、细长嗜热聚球藻菌(Thermosynechococcus elongatus)(Tel)、天蓝色链霉菌(Streptomyces coelicolor)(Sco)、沼泽红假单胞菌(Rhodopseudomonas palustris)(Rpa)、特雷迪氏菌(Teredinibacter turnerae)(Ttu)、焦炭疽杆菌(Pelobacter carbinolicus)(Pca)、巨大芽孢杆菌(Bacillus megaterium)(Bme)、白色链霉菌(Streptomyces albolongus)(Sal)和梨形四膜虫(Tetrahymena pyriformis)(参见，例如WO2010/139719、US2012/01345477、WO2012/066059，其内容通过引用并入本文)。

[0107] 特别地，SHC/HAC酶(例如，可从其衍生SHC/HAC酶变体)可以是酸热脂环酸杆菌(Aac)SHC/HAC酶、运动发酵单胞菌SHC/HAC(ZmoSHC1)酶、日本慢生根瘤菌(Bjp或Bja)SHC/HAC酶或莫比氏葡萄糖杆菌(Gmo)SHC/HAC酶或巴氏醋杆菌SHC/HAC(ApaSHC1)酶或巨大芽孢杆菌(Bme)SHC/HAC酶。特别地，SHC/HAC酶(例如，可从其衍生SHC/HAC酶变体)可以是酸热脂环酸杆菌(Aac)SHC/HAC酶。

[0108] 为了便于参考，名称“AacSHC”可用于指酸热脂环酸杆菌(Aac)SHC/HAC酶，“ZmoSHC”可用于指运动发酵单胞菌(Zmo)SHC/HAC酶，“BjpSHC”或“BjaSHC”可用于指日本慢生根瘤菌(Bjp)SHC/HAC酶，“ApaSHC”可用于指巴氏醋杆菌(Apa)SHC/HAC酶，“BmeSHC”可用于指巨大芽孢杆菌(Bme)SHC/HAC酶，“SalSHC”可用于指白色链霉菌(Sal)SHC/HAC酶，“GmoSHC”可用于指莫比氏葡萄糖杆菌(Gmo)SHC/HAC酶。

[0109] AacSHC、ZmoSHC和BjpSHC酶序列公开于BASF WO2010/139719、US2012/01345477A1、Seitz等人(如上文所引用)和Seitz(如上文所引用的2012博士论文)。对于ZmoSHC公开了两种不同的序列，称为ZmoSHC1和ZmoSHC2。Gmo SHC/HAC酶序列公开于WO2018/157021。SalSHC酶公开于Liu等人(2020):A Novel Soluble Squalene-Hopene Cyclase and Its Application in Efficient Synthesis of Hopene,Frontiers in Bioengineering and Biotechnology,vol 8,article 426,https://doi.org/10.3389/fbioe.2020.00426)。

[0110] 表1公开了野生型SHC酶的来源和登录号。

[0111] 表1.野生型(WT)SHC酶的来源和登录号。

[0112]

SHC 来源菌株 (SHC 名称)	参考资料 (并通过引用并入本文)	登录号	根据 WO2010139719 、US2012/0135477(*)或本公开内容(**)的 SEQ ID NO.
酸热脂环酸杆菌 (WT AacSHC)	JP2009-060799 (Kao) Neumann 等人 Biol Chem (1986) 367; 723-729	NBRC15652	SEQ ID NO: 1 **

[0113]

运动发酵单胞菌 (WT ZmoSHC1)	WO2010139719 US20120135477	ATCC31821 PF62207_2 Genpept 登录号 AAV90172 AAF12829.1 AVZ42714.1	SEQ ID NO: 1 * SEQ ID NO: 2 *
运动发酵单胞菌 (WT ZmoSHC2)	Reipen 等人 (1995) Microbiology 141:155-161	EMBL/Genbank 登录号 X80766	SEQ ID NO: 12 **
日本慢生根瘤菌 (WT BjpSHC)	WO2010139719 US2012/0135477	PF62207_5 ABQ33590.1	SEQ ID NO: 5 *
双向伯克霍尔德菌	WO2010139719 US2012/0135477		SEQ ID NO: 6 *
双向伯克霍尔德菌	WO2010139719 US2012/0135477		SEQ ID NO: 7 *
炭疽杆菌	WO2010139719 US2012/0135477		SEQ ID NO: 8 *
法兰克氏菌	WO2010139719 US2012/0135477		SEQ ID NO: 9 *
沼泽红假单胞菌	WO2010139719 US2012/0135477		SEQ ID NO: 10 *
莫比氏葡萄糖杆菌 (WT GmoSHC)	WO2018157021	EHH69691.1	SEQ ID NO: 14 **
细长嗜热聚球藻菌(TelSHC)		BAC09861.1	SEQ ID NO: 19 **
巴氏醋杆菌 (ApaSHC1)		ASC07046.1	SEQ ID NO: 20 **

[0114]

巨大芽孢杆菌 (WT BmeSHC)		WP_016763969	SEQ ID NO: 28 **
白色链霉菌 (WT SalSHC)		AZN28579	SEQ ID NO: 29 **
巴氏醋杆菌 (WT ApaSHCA)		WP_003625617	SEQ ID NO: 30 **

[0115] 野生型AacSHC、野生型ZmoSHC1、野生型ZmoSHC2、野生型BjpSHC、野生型GmoSHC、野生型TelSHC和野生型ApaSHC1、野生型BmeSHC、野生型SalSHC和野生型ApaSHCA的序列也在本文中公开(分别为SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29和SEQ ID NO:30)。

[0116] 由Hoshino和Sato(2002,如上所述)制备的WT SHC序列的比对表明,在所有四个序列中都检测到多个基序,并且其由在运动发酵单胞菌和酸热脂环酸杆菌的SHC序列中发现六次的核心序列Gln-X-X-X-Gly-X-Trp组成(参见Reipen等人1995,Microbiology 141,155-161的图3)。Hoshino和Sato(2002,如上所述)报道了芳香族氨基酸在SHC中异常丰富,并且在SHC中注意到两个特征基序:一个是由特定氨基酸基序[(K/R)(G/A)X2-3(F/Y/W)(L/IV)3X3QX2-5GXW]表示的QW基序,另一个是DXDDTA基序。Wendt等人(1997,Science277,1811-1815和1999,J Mol Biol 286,175-187)报道了酸热脂环酸杆菌的X射线结构分析。DXDDTA基序似乎与SHC活性位点相关。

[0117] 如本文所用的参考AacSHC蛋白可以指如SEQ ID NO:1中公开的野生型AacSHC蛋白。AacSHC具有可用于通过SHC与均法呢醇底物的生物催化反应产生降龙涎香醚衍生物的均法呢醇降龙涎香醚环化酶(HAC)的活性。AacSHC的主要反应是将线性或非线性底物(例如均法呢醇)环化以产生降龙涎香醚。

[0118] 本文所述的野生型SHC/HAC酶或SHC/HAC酶变体的功能同源物也适用于环化反应,例如用于产生(-)-降龙涎香醚,例如在重组宿主中。因此,重组宿主可以包括编码上述多肽的功能同源物的一种或多种异源核酸和/或编码如本文所述的SHC/HAC衍生酶的异源核酸。

[0119] 功能同源物是与参考多肽具有序列相似性并且执行参考多肽的一种或多种生化或生理功能的多肽。功能同源物和参考多肽可以是天然存在的多肽,并且序列相似性可以是由于趋同或趋异的进化事件。因此,功能同源物有时在文献中被命名为同源物、直系同源物或旁系同源物。天然存在的功能同源物的变体,例如由野生型编码序列的突变体编码的多肽,本身可以是功能同源物。功能同源物也可以通过多肽的编码序列的定点诱变或通过组合来自不同天然存在的多肽的编码序列的结构域(“结构域交换”)来产生。用于修饰编码本文所述的功能同源物的基因的技术是已知的并且尤其包括定向进化技术、定点诱变技术和随机诱变技术,并且可用于增加多肽的比活性、改变底物特异性、改变表达水平、改变亚细胞位置或以所需方式修饰多肽:多肽相互作用。这种修饰的多肽被认为是功能同源物。术语“功能同源物”有时应用于编码功能同源多肽的核酸。

[0120] 可以通过分析核苷酸和多肽序列比对来鉴定功能同源物。例如,对核苷酸或多肽序列的数据库进行查询可以鉴定编码SHC衍生多肽等的核酸序列的同源物。

[0121] 杂交还可用于鉴定功能同源物和/或作为两个核酸序列之间同源性的量度。根据标准杂交技术，编码本文公开的任何蛋白质或其部分的核酸序列可用作杂交探针。探针与来自测试来源(例如哺乳动物细胞)的DNA或RNA的杂交表明测试来源中存在相关DNA或RNA。杂交条件是本领域技术人员已知的并且可以在Current Protocols in Molecular Biology,John Wiley&Sons,N.Y.,6.3.1-6.3.6,1991中找到。中等杂交条件被定义为等同于在2x氯化钠/柠檬酸钠(SSC)中在30℃下杂交，然后在1xSSC、0.1％SDS中在50℃下洗涤。高度严格的条件定义为等同于在45℃下在6x氯化钠/柠檬酸钠(SSC)中杂交，然后在65℃下在0.2xSSC、0.1％SDS中洗涤。鉴定功能同源物的序列分析还可以涉及使用相关氨基酸序列作为参考序列的非冗余数据库的BLAST、Reciprocal BLAST或PSI-BLAST分析。在某些情况下，氨基酸序列是从核苷酸序列推导出来的。数据库中具有大于40％序列同一性的那些多肽是用于进一步评估用于SHC/HAC生物转化反应的适用性的候选物。氨基酸序列相似性允许保守氨基酸取代，例如一个疏水残基取代为另一个疏水残基或一个极性残基取代为另一个极性残基。如果需要，可以对此类候选物进行手动检查，以缩小要进一步评估的候选物的数量。手动检查可以通过选择看起来有例如保守的功能结构域的那些候选物来执行。

[0122] 通常，表现出至少约30％氨基酸序列同一性的多肽可用于鉴定保守区。相关多肽的保守区表现出至少30％、40％、41％、42％、43％、44％、45％、46％、47％、48％、49％、50％、51％、52％、53％、54％、55％、56％、57％、58％、59％、60％、61％、62％、63％、64％、65％、66％、67％、68％、69％的氨基酸序列同一性。在一些实施方案中，保守区表现出至少70％、71％、72％、73％、74％、75％、76％、77％、78％、79％、80％、至少81％、至少82％、至少83％、至少84％、至少85％、至少86％、至少87％、至少88％、至少89％、至少90％、至少91％、至少92％、至少93％、至少94％、至少95％、至少96％、至少97％、至少98％或至少99％的氨基酸序列同一性。可以如上文和下文所述确定序列同一性。

[0123] 本文所述和用于本文所述方法中的SHC/HAC酶或酶变体可例如基于SEQ ID NO:1、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29、SEQ ID NO:30，或其变体、同源物、突变体、衍生物或片段。SHC/HAC酶或酶变体可例如具有与SEQ ID NO:1、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30具有至少30％、40％、41％、42％、43％、44％、45％、46％、47％、48％、49％、50％、51％、52％、53％、54％、55％、56％、57％、58％、59％、60％、61％、62％、63％、64％、65％、66％、67％、68％、69％、70％、71％、72％、73％、74％、75％、76％、77％、78％、79％、80％、至少81％、至少82％、至少83％、至少84％、至少85％、至少86％、至少87％、至少88％、至少89％、至少90％、至少91％、至少92％、至少93％、至少94％、至少95％、至少96％、至少97％、至少98％或至少99％的同一性的氨基酸序列。

[0124] 此外，产生的参考SHC可以基于从大肠杆菌产生的氨基酸序列。

[0125] 关于基因的核苷酸序列的“百分比(％)同一性”定义为在比对序列并在必要时引入空位以达到最大序列同一性百分比之后，并且不考虑将任何保守取代作为序列同一性的一部分时，候选DNA序列中与DNA序列中的核苷酸相同的核苷酸的百分比。用于确定百分比核苷酸序列同一性目的的比对可以以本领域技术范围内的各种方式实现，例如，使用可公开获得的计算机软件。本领域技术人员可以确定用于测量比对的适当参数，包括在被比较

的序列的全长上实现最大比对所需的任何算法。术语“多肽”和“蛋白质”在本文中可互换使用，并且是指氨基酸的任何肽连接的链，无论长度或翻译后修饰。

[0126] 如本文所用，术语“衍生物”包括但不限于变体。术语“衍生物”和“变体”在本文中可互换使用。

[0127] 在优选的实施方案中，可用于本公开内容的变体酶表现出氨基酸序列中多达200(多达1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、110、120、130、140、150、160、170、180、190或200)个变化(改变)(即交换、插入、缺失、N末端截短和/或C末端截短)。氨基酸交换可以是保守的和/或非保守的。在优选的实施方案中，可用于本公开内容的变体与衍生其的蛋白质或结构域相差多达1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95或100个氨基酸交换，优选保守氨基酸改变。变体可以另外或替代地包括氨基酸缺失，其可以是N末端截短、C末端截短或内部缺失或这些的任何组合。此类包含N末端截短、C末端截短和/或内部缺失的变体在本申请的上下文中称为“缺失变体”或“片段”。术语“缺失变体”和“片段”在本文中可互换使用。缺失变体可以是天然存在的(例如剪接变体)，或其可以是人工构建的，优选通过基因技术手段构建。通常，产生缺失变体的蛋白质或蛋白质结构域是野生型蛋白质。然而，本公开内容的缺失变体还可以衍生自亲本多肽的同源物、直系同源物或旁系同源物或来自人工构建的变体，条件是缺失变体表现出亲本多肽的至少一种生物活性。优选地，缺失变体(或片段)与亲本多肽相比在其N末端和/或在其C末端和/或内部具有多达1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95或100个氨基酸的缺失。

[0128] 在某些实施方案中，本文所述的SHC/HAC酶变体仅包括取代而不包括任何缺失或插入。

[0129] 如本文所用，“变体”可以替代地或另外地特征在于与衍生其的亲本多肽的一定程度的序列同一性。本公开内容的WT/参考SHC/HAC或SHC/HAC衍生物的变体可以与各自的参考多肽或各自的参考多核苷酸具有至少40％、41％、42％、43％、44％、45％、46％、47％、48％、49％、50％、51％、52％、53％、54％、55％、56％、57％、58％、59％、60％、61％、62％、63％、64％、65％、66％、67％、68％、69％、70％、71％、72％、73％、74％、75％、76％、77％、78％、79％、80％、至少81％、至少82％、至少83％、至少84％、至少85％、至少86％、至少87％、至少88％、至少89％、至少90％、至少91％、至少92％、至少93％、至少94％、至少95％、至少96％、至少97％、至少98％或至少99％同一性的序列同一性。

[0130] 表述“至少30％、40％、41％、42％、43％、44％、45％、46％、47％、48％、49％、50％、51％、52％、53％、54％、55％、56％、57％、58％、59％、60％、61％、62％、63％、64％、65％、66％、67％、68％、69％、70％、71％、72％、73％、74％、75％、76％、77％、78％、79％、80％、至少81％、至少82％、至少83％、至少84％、至少85％、至少86％、至少87％、至少88％、至少89％、至少90％、至少91％、至少92％、至少93％、至少94％、至少95％、至少96％、至少97％、至少98％或至少99％的序列同一性”在整个说明书中用于多肽和多核苷酸序列比较。可以分别基于其与相关基因或蛋白质的相似性来鉴定属于本文公开的任何酶家族的多核苷酸或蛋白质。例如，鉴定可以基于序列同一性。在某些优选的实施方案中，本公开内容的特征在于分离的核酸分子，其与(a)编码本文公开的野生型SHC/HAC酶的多肽(例如SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:

20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30)的核酸分子、(b)核苷酸序列SEQ ID NO:15和(c)包括SEQ ID NO:15的至少30(例如,至少30、40、50、60、80、100、125、150、175、200、250、300、400、500、600、700、800、850、900、950、1000或1010)个核苷酸的区段的核酸分子具有至少30%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、至少81%、至少82%、至少83%、至少84%、至少85%、至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%的同一性。

[0131] 优选地,所讨论的多肽和参考多肽在20、30、40、45、50、60、70、80、90、100或更多个氨基酸的连续区段上表现出指示的序列同一性。优选地,所讨论的多核苷酸和参考多核苷酸在60、90、120、135、150、180、210、240、270、300个或更多个核苷酸的连续区段上表现出指示的序列同一性。在比较两个序列并且未指定要与之比较以计算序列同一性百分比的参考序列的情况下,如果没有具体说明另有说明,序列同一性将参考要比较的两个序列中的较长者来计算。如果指明了参考序列,则根据参考序列(例如SEQ ID NO:1、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30)的全长确定序列同一性,除非另有说明。

[0132] 例如,与具有631个氨基酸残基的全长野生型AacSHC的氨基酸相比,由130个氨基酸组成的肽序列可能表现出20.6%(130/631x100)的最大序列同一性百分比,而具有300个氨基酸的长度的序列可能表现出47.5%(300/631x100)的最大序列同一性百分比。

[0133] 核苷酸和氨基酸序列的相似性,即序列同一性的百分比,可以通过序列比对来确定。这种比对可以用几种本领域已知的算法进行,优选使用Karlin和Altschul的数学算法(Karlin&Altschul(1993)Proc.Natl.Acad.Sci.USA 90:5873-5877),使用hmmalign(HMMER软件包,http://hmmer.wustl.edu/),或使用例如可获自https://www.ebi.ac.uk/Tools/msa/clustalo/的CLUSTAL算法(Thompson,J.D.,Higgins,D.G.&Gibson,T.J.(1994)Nucleic Acids Res.22,4673-80),或GAP程序(爱荷华大学的数学算法)或Myers和Miller(1989-Cabios 4:11-17)的数学算法或Clone Manager9。使用的优选参数是在https://www.ebi.ac.uk/Tools/msa/clustalo/上设置的默认参数。

[0134] 序列同一性(序列匹配)的等级可以使用例如BLAST、BLAT或BlastZ(或BlastX)计算。Altschul等人(1990)J.Mol.Biol.215,403-410的BLASTN和BLASTP程序中并入了类似的算法。使用BLASTN程序(得分=100,字长=12)进行BLAST多核苷酸搜索以获得与编码相关蛋白质的那些核酸同源的多核苷酸序列。

[0135] 使用BLASTP程序(得分=50,字长=3)进行BLAST蛋白质搜索以获得与SHC多肽同源的氨基酸序列。为了获得用于比较目的的空位比对,如Altschul等人(1997)Nucleic Acids Res.25,3389-3402中所述,使用空位BLAST。当使用BLAST和空位BLAST程序时,使用各自程序的默认参数。序列匹配分析可以通过已建立的同源映射技术如Shuffle-LAGAN(Brudno M.,Bioinformatics 2003b,19Suppl 1:154-162)或Markov随机场来补充。当在本申请中提及序列同一性的百分比时,如果没有特别指出,这些百分比是相对于较长序列的全长计算的。

[0136] 在特定实施方案中,使用CLUSTAL O(版本1.2.4)确定两个序列之间的%同一性。

[0137] 下文进一步描述了可用于本文所述方法的特定SHC/HAC酶和酶变体。

[0138] 具有新突变的Aac 215G2 SHC/HAC变体

[0139] 令人惊讶地发现,衍生自WO2016/170099中公开的Aac SHC/HAC酶变体(215G2 SHC/HAC酶变体)的SHC/HAC酶变体为将EEH转化为(-)-降龙涎香醚提供了改善的酶活性。进一步令人惊讶地发现,衍生自215G2的SHC/HAC酶变体为将E,E-双均法呢醇转化为龙涎香氧化物提供了改善的酶活性。

[0140] 除了215G2 SHC/HAC酶变体中已经存在的氨基酸取代外,新的SHC/HAC酶变体具有两个或三个氨基酸改变。

[0141] 因此,本文提供了一种通过将EEH酶促转化为(-)-降龙涎香醚来制备(-)-降龙涎香醚的方法。本文还提供了一种通过将E,E-双均法呢醇酶促转化为龙涎香氧化物来制备龙涎香氧化物的方法。这些过程可以使用本文所述的任何野生型SHC/HAC酶或酶变体,特别是本文所述的Aac 215G2 SHC/HAC变体。

[0142] Aac 215G2 SHC/HAC的SHC/HAC酶变体具有与SEQ ID NO:1具有至少约70.0%同一性的氨基酸序列。例如,SHC/HAC酶变体可与SEQ ID NO:1具有至少约75.0%或至少约80.0%或至少约85.0%或至少约90.0%或至少约95.0%或至少约95.5%或至少约96.0%或至少约96.5%或至少约97.0%或至少约97.5%或至少约98.0%或至少约98.5%或至少约99.0%的同一性。

[0143] Aac 215G2 SHC/HAC的酶变体与SEQ ID NO:1具有小于100%的同一性。例如,SHC/HAC酶变体可以与SEQ ID NO:1具有等于或小于约99.5%或等于或小于约99.0%的同一性。

[0144] 例如,Aac 215G2 SHC/HAC的酶变体可与SEQ ID NO:1具有约70.0%至约99.5%或约80.0%至约99.0%或约85.0%至约98.5%或约90.0%至约98.0%的同一性。

[0145] 关于多肽或核苷酸序列的“百分比(%)同一性”分别定义为在比对序列并在必要时引入空位以达到最大序列同一性百分比之后,并且不考虑将任何保守取代作为序列同一性的一部分时,候选序列中与参考序列中的氨基酸或核苷酸相同的氨基酸或核苷酸的百分比。用于确定序列同一性百分比目的的比对可以以本领域技术范围内的各种方式实现,例如,使用可公开获得的计算机软件。本领域技术人员可以确定用于测量比对的适当参数,包括在被比较的序列的全长上实现最大比对所需的任何算法。术语“多肽”和“蛋白质”在本文中可互换使用,并且是指氨基酸的任何肽连接的链,无论长度或翻译后修饰。

[0146] 核苷酸和氨基酸序列的相似性,即序列同一性的百分比,可以通过序列比对来确定。这样的比对可以用几种本领域已知的算法进行,优选使用Karlin和Altschul的数学算法(Karlin&Altschul(1993)Proc.Natl.Acad.Sci.USA 90:5873-5877),使用hmmalign(HMMER软件包,http://hmmer.wustl.edu/),或使用例如可获自https://www.ebi.ac.uk/Tools/msa/clustalo/的CLUSTAL算法(Thompson,J.D.,Higgins,D.G.&Gibson,T.J.(1994)Nucleic Acids Res.22,4673-80),或GAP程序(爱荷华大学的数学算法)或Myers和Miller(1989-Cabios 4:11-17)的数学算法。使用的优选参数是在https://www.ebi.ac.uk/Tools/msa/clustalo/上设置的默认参数。

[0147] 可以使用例如BLAST、BLAT或BlastZ(或BlastX)计算序列同一性百分比。Altschul等人(1990)J.Mol.Biol.215,403-410的BLASTN和BLASTP程序中并入了类似的算法。可以使

用BLASTN程序(得分＝100,字长＝12)进行BLAST多核苷酸搜索以获得与编码相关蛋白质的那些核酸同源的多核苷酸序列。可以使用BLASTP程序(得分＝50,字长＝3)进行BLAST蛋白质搜索以获得与多肽同源的氨基酸序列。

[0148] 为了获得用于比较目的的空位比对,如Altschul等人(1997)Nucleic Acids Res.25,3389-3402中所述,使用空位BLAST。当使用BLAST和空位BLAST程序时,使用各自程序的默认参数。序列匹配分析可以通过已建立的同源映射技术如Shuffle-LAGAN(Brudno M.,Bioinformatics 2003b,19Suppl 1:154-162)或Markov随机场来补充。当在本申请中提及序列同一性的百分比时,如果没有特别指出,这些百分比是相对于较长序列的全长计算的。

[0149] 在特定实施方案中,使用CLUSTAL O(版本1.2.4)确定两个序列之间的%同一性。

[0150] 在某些实施方案中,与SEQ ID NO:1相比,SHC/HAC酶变体可以具有等于或小于约30个氨基酸改变。例如,与SEQ ID NO:1相比,SHC/HAC酶变体可以具有等于或小于约25或等于或小于约20或等于或小于约15或等于或小于约10或等于或小于约9或等于或小于约8或等于或小于约7或等于或小于约6个氨基酸改变。例如,与SEQ ID NO:1相比,SHC/HAC酶变体可以具有至少约5个或至少约6个氨基酸改变。氨基酸改变例如可以是如上所述的插入、缺失和/或取代。

[0151] 在某些实施方案中,与SEQ ID NO:1相比,SHC/HAC酶变体中唯一的氨基酸改变是取代(即,没有插入或缺失)。

[0152] 氨基酸改变是相对于参考序列定义的。相对于参考序列的氨基酸改变是指变体序列的氨基酸序列与参考序列不同。

[0153] 参考序列和变体序列中的氨基酸可以被分配编号,其中编号从多肽的N末端的氨基酸开始(即多肽N末端的氨基酸被编号为1,下一个氨基酸被编号为2等)。参考序列的“位置”是指参考序列中存在的特定氨基酸残基,如通过参考序列中氨基酸的特定编号所识别的。变体序列的“位置”是指存在于变体序列中的特定氨基酸残基,如通过变体序列中氨基酸的特定编号所识别的。

[0154] 由于与参考序列相比,变体序列可以包括缺失或插入,因此变体序列中的氨基酸的编号可能与参考序列中的相同氨基酸不同。例如,如果在SEQ ID NO:1的氨基酸131和132之间插入氨基酸,则插入的后面的氨基酸将在变体序列中具有编号133,而其在参考序列中保留编号132。在本示例中,与参考序列的位置132对应的变体序列的位置为位置133。因此,变体序列中从参考序列中保留下来的氨基酸可以参照参考序列的“对应位置”来定义。换言之,变体序列中的“位置”可以通过参照参考序列中的“对应位置”来定义。特别地,与参考序列相比变体序列中的取代可以通过参照参考序列的“对应位置”来定义,尽管参考序列中存在任何插入和/或缺失。在参考序列的氨基酸被缺失的情况下,变体序列中没有“对应位置”。如果与参考序列相比没有插入或缺失(即只有取代),则参考序列的“对应位置”将与变体序列中的位置相同。

[0155] 215G2 SHC/HAC酶变体(与野生型Aac SHC相比)中的原始氨基酸改变保留在本文所述的新Aac 215G2 SHC/HAC酶变体中。与野生型Aac SHC相比215G2 SHC/HAC酶变体中的原始氨基酸改变是取代M132R、A224V和I432T(即位置132的M残基取代为R残基,位置224的A残基取代为V残基,位置432的I残基取代为T残基)。

[0156] 因此,新的Aac 215G2 SHC/HAC酶变体在对应于SEQ ID NO:1的位置132、224和432的位置处具有相对于SEQ ID NO:1的氨基酸改变,分别为M132R、A224V和I432T。“M132R”中的数字“132”、“A224V”中的“224”和“I432T”中的“432”指的是SEQ ID NO:1的编号,并且如上所述由于可选的其他插入和/或缺失而不一定对应于变体序列的编号。

[0157] 新的Aac 215G2 SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557的位置和对应于SEQ ID NO:1的位置81、431或613的至少一个位置处也具有相对于SEQ ID NO:1的氨基酸改变。例如,这些氨基酸改变可以是取代,例如非保守取代。

[0158] 例如,SHC/HAC酶变体氨基酸序列可以在对应于SEQ ID NO:1的位置557的位置和对应于SEQ ID NO:1的位置81、431或613的一个位置处具有相对于SEQ ID NO:1的氨基酸改变。这些氨基酸改变可以例如是取代,例如非保守取代。

[0159] 例如,SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557的位置和选自对应于SEQ ID NO:1的位置81、431和613的位置的两个位置处可以具有相对于SEQ ID NO:1的氨基酸改变。这些氨基酸改变可以例如是取代,例如非保守取代。

[0160] 例如,SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557的位置和对应于SEQ ID NO:1的位置81、431和613的所有位置处可以具有相对于SEQ ID NO:1的氨基酸改变。例如,这些氨基酸改变可以是取代,例如非保守取代。

[0161] 例如,SHC/HAC酶变体可以在对应于SEQ ID NO:1的位置431和557的位置处具有相对于SEQ ID NO:1的氨基酸改变。这些氨基酸改变可以例如是取代,例如非保守取代。在某些实施方案中,SHC/HAC酶变体具有SEQ ID NO:2的序列。

[0162] 例如,SHC/HAC酶变体可以在对应于SEQ ID NO:1的位置557和613的位置处具有相对于SEQ ID NO:1的氨基酸改变。这些氨基酸改变可以例如是取代,例如非保守取代。在某些实施方案中,SHC/HAC酶变体具有SEQ ID NO:3的序列。

[0163] 例如,SHC/HAC酶变体可以在对应于SEQ ID NO:1的位置81、557和613的位置处具有相对于SEQ ID NO:1的氨基酸改变。这些氨基酸改变可以例如是取代,例如非保守取代。在某些实施方案中,SHC/HAC酶变体具有SEQ ID NO:4的序列。

[0164] 例如,SHC/HAC酶变体可以在对应于SEQ ID NO:1的位置81、431和557的位置处具有相对于SEQ ID NO:1的氨基酸改变。这些氨基酸改变可以例如是取代,例如非保守取代。在某些实施方案中,SHC/HAC酶变体具有SEQ ID NO:5的序列。

[0165] 对应于SEQ ID NO:1的位置557的位置处的氨基酸改变可以是例如A557X。这是指SEQ ID NO:1的位置557处的氨基酸A被任何不同的氨基酸(X)取代。如上所述,由于SHC/HAC酶变体可能另外包含插入和/或缺失,因此新的SHC/HAC酶变体中新氨基酸(X)的编号可能不是557。

[0166] 对应于SEQ ID NO:1的位置557的位置处的新氨基酸(X)可以例如是Met、Val、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如,在对应于SEQ ID NO:1的位置557的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即A)取代为中性亲水性氨基酸(即Cys、Ser、Thr、Asn或Gln)。例如,对应于SEQ ID NO:1的位置557的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即A)取代为苏氨酸(即对应于SEQ ID NO:1的位置557的位置处的氨基酸改变是A557T)。

[0167] 对应于SEQ ID NO:1的位置81的位置处的氨基酸改变可以是例如Y81X。这是指SEQ

ID NO:1的位置81处的氨基酸Y被任何不同的氨基酸(X)取代。如上所述,由于SHC/HAC酶变体可能另外包含插入和/或缺失,因此新的SHC/HAC酶变体中新氨基酸(X)的编号可能不是81。

[0168] 对应于SEQ ID NO:1的位置81的位置处的新氨基酸(X)可以例如是Met、Ala、Val、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp或Phe。例如,对应于SEQ ID NO:1的位置81的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即Y)取代为碱性氨基酸(即His、Lys或Arg)。例如,对应于SEQ ID NO:1的位置81的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即Y)取代为组氨酸(即对应于SEQ ID NO:1的位置81的位置处的氨基酸改变是Y81H)。

[0169] 对应于SEQ ID NO:1的位置431的位置处的氨基酸改变可以是例如H431X。这是指SEQ ID NO:1的位置431处的氨基酸H被任何不同的氨基酸(X)取代。如上所述,由于SHC/HAC酶变体可能另外包含插入和/或缺失,因此新的SHC/HAC酶变体中新氨基酸(X)的编号可能不是431。

[0170] 对应于SEQ ID NO:1的位置431的位置处的新氨基酸(X)可以是例如Met、Ala、Val、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如,对应于SEQ ID NO:1的位置431的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即H)取代为疏水性氨基酸(即Met、Ala、Val、Leu或Ile)。例如,对应于SEQ ID NO:1的位置431的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即H)取代为亮氨酸(即对应于SEQ ID NO:1的位置431的位置处的氨基酸改变是H431L)。

[0171] 对应于SEQ ID NO:1的位置613的位置处的氨基酸改变可以是例如R613X。这是指SEQ ID NO:1的位置613处的氨基酸R被任何不同的氨基酸(X)取代。如上所述,由于SHC/HAC酶变体可能另外包含插入和/或缺失,因此新的SHC/HAC酶变体中新氨基酸(X)的编号可能不是613。

[0172] 对应于SEQ ID NO:1的位置631的位置处的新氨基酸(X)可以例如是Met、Ala、Val、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Gly、Pro、Trp、Tyr或Phe。例如,在对应于SEQ ID NO:1的位置613的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即R)取代为中性亲水性氨基酸(即Cys、Ser、Thr、Asn或Gln)。例如,对应于SEQ ID NO:1的位置613的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即R)取代为丝氨酸(即对应于SEQ ID NO:1的位置613的位置处的氨基酸改变是R613S)。

[0173] 在某些实施方案中,除了以下氨基酸取代外,新的Aac 215G2SHC变体与SEQ ID NO:1相同:

[0174] (i)M132R、A224V、I432T、A557T和H431L(SEQ ID NO:2);或者

[0175] (ii)M132R、A224V、I432T、A557T和R613S(SEQ ID NO:3);或者

[0176] (iii)M132R、A224V、I432T、A557T、Y81H和R613S(SEQ ID NO:4);或者

[0177] (iv)M132R、A224V、I432T、A557T、Y81H和H431L(SEQ ID NO:5)。

[0178] SHC/HAC酶变体可以例如在对应于SEQ ID NO:1的位置90、172和/或277的位置处具有一个或多个进一步的氨基酸改变。

[0179] 对应于SEQ ID NO:1的位置90的位置处的氨基酸改变可以是例如T90X。这是指SEQ ID NO:1的位置90处的氨基酸T被任何不同的氨基酸(X)取代。如上所述,由于SHC/HAC酶变

体可能另外包含插入和/或缺失,因此新的SHC/HAC酶变体中新氨基酸(X)的编号可能不是90。

[0180] 对应于SEQ ID NO:1的位置90的位置处的新氨基酸(X)可以例如是Met、Ala、Val、Leu、Ile、Cys、Ser、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如,在对应于SEQ ID NO:1的位置90的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即T)替代为疏水性氨基酸(即Met、Ala、Val、Leu、Ile)。例如,对应于SEQ ID NO:1的位置90的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即T)取代为丙氨酸(即对应于SEQ ID NO:1的位置90的位置处的氨基酸改变是T90A)。

[0181] 对应于SEQ ID NO:1的位置172的位置处的氨基酸改变可以是例如A172X。这是指SEQ ID NO:1的位置172处的氨基酸T被任何不同的氨基酸(X)取代。如上所述,由于SHC/HAC酶变体可能另外包含插入和/或缺失,因此新的SHC/HAC酶变体中新氨基酸(X)的编号可能不是172。

[0182] 对应于SEQ ID NO:1的位置172的位置处的新氨基酸(X)可以例如是Met、Val、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如,在对应于SEQ ID NO:1的位置172的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即A)取代为中性亲水性氨基酸(即Cys、Ser、Thr、Asn、Gln)。例如,对应于SEQ ID NO:1的位置172的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即A)取代为苏氨酸(即对应于SEQ ID NO:1的位置172的位置处的氨基酸改变是A172T)。

[0183] 对应于SEQ ID NO:1的位置277的位置处的氨基酸改变可以是例如M277X。这是指SEQ ID NO:1的位置277处的氨基酸M被任何不同的氨基酸(X)取代。如上所述,由于SHC/HAC酶变体可能另外包含插入和/或缺失,因此新的SHC/HAC酶变体中新氨基酸(X)的编号可能不是277。

[0184] 对应于SEQ ID NO:1的位置277的位置的新氨基酸(X)可以是例如Ala、Val、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如,在对应于SEQ ID NO:1的位置277的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即M)取代为碱性氨基酸(即His、Lys、Arg)。例如,对应于SEQ ID NO:1的位置277的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即M)取代为赖氨酸(即对应于SEQ ID NO:1的位置277的位置处的氨基酸改变是M277K)。

[0185] 例如,SHC/HAC酶变体可以在对应于SEQ ID NO:1的位置37、174和/或601的位置处具有一个或多个进一步的氨基酸改变。例如,SHC/HAC酶变体可以在对应于SEQ ID NO:1的位置37、174和/或601的位置处具有一个或多个氨基酸取代(例如保守或非保守取代)。

[0186] 对应于SEQ ID NO:1的位置37的位置处的氨基酸改变可以是例如L37X。这是指SEQ ID NO:1的位置37处的氨基酸L被任何不同的氨基酸(X)取代。如上所述,由于SHC/HAC酶变体可能另外包含插入和/或缺失,因此新的SHC/HAC酶变体中新氨基酸(X)的编号可能不是37。

[0187] 对应于SEQ ID NO:1的位置37的位置的新氨基酸(X)可以是例如Met、Ala、Val、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如,对应于SEQ ID NO:1的位置37的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即L)取代为中性亲水性氨基酸(即Cys、Ser、Thr、Asn或Gln)。例如,对应于SEQ ID NO:1的位置37的位置处

的氨基酸改变可以将SEQ ID NO:1的氨基酸(即L)取代为谷氨酰胺(即对应于SEQ ID NO:1的位置37的位置处的氨基酸改变是L37Q)。

[0188] 对应于SEQ ID NO:1的位置174的位置处的氨基酸改变可以是例如V174X。这是指SEQ ID NO:1的位置174处的氨基酸V被任何不同的氨基酸(X)取代。如上所述,由于SHC/HAC酶变体可能另外包含插入和/或缺失,因此新的SHC/HAC酶变体中新氨基酸(X)的编号可能不是174。

[0189] 对应于SEQ ID NO:1的位置174的位置处的新氨基酸(X)可以例如是Met、Ala、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。对应于SEQ ID NO:1的位置174的位置处的新氨基酸(X)可以是例如疏水性氨基酸(即Met、Ala、Leu或Ile)。例如,对应于SEQ ID NO:1的位置174的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即V)取代为异亮氨酸(即对应于SEQ ID NO:1的位置174的位置处的氨基酸改变是V174I)。

[0190] 对应于SEQ ID NO:1的位置601的位置处的氨基酸改变可以是例如F601X。这是指SEQ ID NO:1的位置601处的氨基酸F被任何不同的氨基酸(X)取代。如上所述,由于SHC/HAC酶变体可能另外包含插入和/或缺失,因此新的SHC/HAC酶变体中新氨基酸(X)的编号可能不是601。

[0191] 对应于SEQ ID NO:1的位置601的位置处的新氨基酸(X)可以例如是Met、Ala、Val、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp或Tyr。对应于SEQ ID NO:1的位置601的位置处的新氨基酸(X)可以是例如芳族酸(即Trp、Tyr、Phe)。例如,对应于SEQ ID NO:1的位置601的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即F)取代为酪氨酸(即对应于SEQ ID NO:1的位置601的位置处的氨基酸改变是F601Y)。

[0192] SHC/HAC酶变体可以例如在对应于SEQ ID NO:1的位置77、92、129、579、601和/或605的位置处具有一个或多个进一步的氨基酸改变。例如,SHC/HAC酶变体可以在对应于SEQ ID NO:1的位置77、92、129、579、601和/或605的位置处具有一个或多个氨基酸取代(例如保守或非保守取代)。

[0193] 对应于SEQ ID NO:1的位置77的位置处的氨基酸改变可以是例如T77X。这是指SEQ ID NO:1的位置77处的氨基酸T被任何不同的氨基酸(X)取代。如上所述,由于SHC/HAC酶变体可能另外包含插入和/或缺失,因此新的SHC/HAC酶变体中新氨基酸(X)的编号可能不是77。

[0194] 对应于SEQ ID NO:1的位置77的位置处的新氨基酸(X)可以是例如Met、Ala、Val、Leu、Ile、Cys、Ser、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如,对应于SEQ ID NO:1的位置77的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即T)取代为疏水性氨基酸(即Met、Ala、Val、Leu或Ile)。例如,对应于SEQ ID NO:1的位置77的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即T)取代为丙氨酸(即对应于SEQ ID NO:1的位置77的位置处的氨基酸改变是T77A)。

[0195] 对应于SEQ ID NO:1的位置92的位置处的氨基酸改变可以是例如I92X。这是指SEQ ID NO:1的位置92处的氨基酸I被任何不同的氨基酸(X)取代。如上所述,由于SHC/HAC酶变体可能另外包含插入和/或缺失,因此新的SHC/HAC酶变体中新氨基酸(X)的编号可能不是92。

[0196] 对应于SEQ ID NO:1的位置92的位置处的新氨基酸(X)可以是例如Met、Ala、Val、

Leu、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如，在对应于SEQ ID NO:1的位置92的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即I)取代为疏水性氨基酸(即Met、Ala、Val、Leu或Ile)。例如，对应于SEQ ID NO:1的位置92的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即I)取代为缬氨酸(即对应于SEQ ID NO:1的位置92的位置处的氨基酸改变是I92V)。

[0197] 对应于SEQ ID NO:1的位置129的位置处的氨基酸改变可以是例如F129X。这是指SEQ ID NO:1的位置129处的氨基酸F被任何不同的氨基酸(X)取代。如上所述，由于SHC/HAC酶变体可能另外包含插入和/或缺失，因此新的SHC/HAC酶变体中新氨基酸(X)的编号可能不是129。

[0198] 对应于SEQ ID NO:1的位置129的位置处的新氨基酸(X)可以是例如Met、Ala、Val、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp或Tyr。例如，对应于SEQ ID NO:1的位置129的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即F)取代为疏水性氨基酸(即Met、Ala、Val、Leu或Ile)。例如，对应于SEQ ID NO:1的位置129的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即F)取代为亮氨酸(即对应于SEQ ID NO:1的位置129的位置处的氨基酸改变是F129L)。

[0199] 对应于SEQ ID NO:1的位置579的位置处的氨基酸改变可以是例如Q579X。这是指SEQ ID NO:1的位置579处的氨基酸Q被任何不同的氨基酸(X)取代。如上所述，由于SHC/HAC酶变体可能另外包含插入和/或缺失，因此新的SHC/HAC酶变体中新氨基酸(X)的编号可能不是579。

[0200] 对应于SEQ ID NO:1的位置579的新氨基酸(X)可以是例如Met、Ala、Val、Leu、Ile、Cys、Ser、Thr、Asn、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如，对应于SEQ ID NO:1的位置579的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即Q)取代为碱性氨基酸(即His、Lys或Arg)。例如，对应于SEQ ID NO:1的位置579的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即Q)取代为组氨酸(即对应于SEQ ID NO:1的位置579的位置处的氨基酸改变是Q579H)。

[0201] 对应于SEQ ID NO:1的位置601的位置处的氨基酸改变可以是例如F601X。这是指SEQ ID NO:1的位置601处的氨基酸F被任何不同的氨基酸(X)取代。如上所述，由于SHC/HAC酶变体可能另外包含插入和/或缺失，因此新的SHC/HAC酶变体中新氨基酸(X)的编号可能不是601。

[0202] 对应于SEQ ID NO:1的位置601的位置处的新氨基酸(X)可以例如是Met、Ala、Val、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp或Tyr。对应于SEQ ID NO:1的位置601的位置处的新氨基酸(X)可以是例如芳族酸(即Trp、Tyr、Phe)。例如，对应于SEQ ID NO:1的位置601的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即F)取代为组氨酸(即对应于SEQ ID NO:1的位置601的位置处的氨基酸改变是F601Y)。

[0203] 对应于SEQ ID NO:1的位置605的位置处的氨基酸改变可以是例如F605X。这是指SEQ ID NO:1的位置605处的氨基酸F被任何不同的氨基酸(X)取代。如上所述，由于SHC/HAC酶变体可能另外包含插入和/或缺失，因此新的SHC/HAC酶变体中的新氨基酸(X)的编号可能不是605。

[0204] 对应于SEQ ID NO:1的位置605的位置处的新氨基酸(X)可以例如是Met、Ala、Val、

Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp或Tyr。对应于SEQ ID NO:1的位置605的位置处的新氨基酸(X)可以是例如芳族酸(即Trp、Tyr、Phe)。例如,对应于SEQ ID NO:1的位置605的位置处的氨基酸改变可以将SEQ ID NO:1的氨基酸(即F)取代为色氨酸(即对应于SEQ ID NO:1的位置601的位置处的氨基酸改变是F605W)。

[0205] 例如,SHC/HAC酶变体可以在对应于SEQ ID NO:1的位置132和432的位置处具有氨基酸改变(例如取代)。

[0206] 例如,SHC/HAC酶变体可以在对应于SEQ ID NO:1的位置601的位置处具有氨基酸改变(例如取代)。

[0207] 例如,SHC/HAC酶变体可以在对应于SEQ ID NO:1的位置77、92和129的位置处具有氨基酸改变(例如取代)。

[0208] 例如,SHC/HAC酶变体可以在对应于SEQ ID NO:1的位置579和601的位置处具有氨基酸改变(例如取代)。

[0209] 例如,SHC/HAC酶变体可以在对应于SEQ ID NO:1的位置129、132和432的位置处具有氨基酸改变(例如取代)。

[0210] 例如,SHC/HAC酶变体可以在对应于SEQ ID NO:1的位置132、432和601的位置处具有氨基酸改变(例如取代)。

[0211] 例如,SHC/HAC酶变体可以在对应于SEQ ID NO:1的位置129、132、432和601的位置处具有氨基酸改变(例如取代)。

[0212] 与SEQ ID NO:1的SHC/HAC酶和/或SEQ ID NO:10的SHC/HAC酶变体相比,新的SHC/HAC酶变体可以例如具有用于将EEH转化为(-)-降龙涎香醚或将BisEEH转化为龙涎香氧化物的增加的酶活性。增加的酶活性可以指EEH到(-)-降龙涎香醚的酶促转化或BisEEH到龙涎香氧化物的酶促转化的任何方面,包括例如EEH或BisEEH的增加的总转化,EEH或BisEEH的增加的转化率(例如在反应的前6小时或前12小时内),(-)-降龙涎香醚或龙涎香氧化物的增加的产量,和副产物的减少的产量。增加的酶活性通常可以通过增加的生产率来定义,其可以根据每克生物催化剂、每小时和每升反应产生的(-)-降龙涎香醚或龙涎香氧化物来定义。

[0213] 与SEQ ID NO:1的SHC/HAC野生型酶和/或SEQ ID NO:10的Aac 215G2 SHC/HAC酶变体相比,新的SHC/HAC酶变体可以例如提供增加的EEH或E,E-双均法呢醇(BisEEH)转化。因此,与使用SEQ ID NO:1的SHC/HAC野生型酶和/或SEQ ID NO:10的SHC/HAC酶变体的方法相比,本文描述的方法可以具有增加的EEH或BisEEH转化水平。新的SHC/HAC酶变体可以例如与SEQ ID NO:1的SHC/HAC野生型酶和/或SEQ ID NO:10的SHC/HAC酶变体相比提供增加的EEH或BisEEH转化率。因此,与SEQ ID NO:1的SHC/HAC野生型酶和/或SEQ ID NO:10的SHC/HAC酶变体相比,本文描述的方法可以具有增加的EEH或BisEEH转化率。新的SHC/HAC酶变体可以例如与SEQ ID NO:1的SHC/HAC野生型酶和/或SEQ ID NO:10的SHC/HAC酶变体相比在反应的前4小时或前6小时或前8小时或前12小时或前24小时内提供增加的EEH或BisEEH转化率。因此,与SEQ ID NO:1的SHC/HAC野生型酶和/或SEQ ID NO:10的SHC/HAC酶变体相比,本文所述的方法在反应的前4小时或前6小时或前8小时或前12小时或前24小时内可以具有增加的EEH或BisEEH转化率。这可以是在与在相同的反应条件(例如相同的pH和温度)下使用两种酶(即新的Aac 215G2 SHC/HAC酶变体和SEQ ID NO:1的野生型酶或SEQ

ID NO:10的215G2酶)进行比较时或在与在其各自优化的反应条件(例如优化的pH和温度)(其可能彼此不同)下使用每种酶进行比较时。

[0214] 例如,新的SHC/HAC酶变体可在反应的前12小时内提供或该方法可在反应的前12小时内具有至少约40%的EEH或BisEEH转化。例如,新的SHC/HAC酶变体可在反应的前12小时内提供或该方法可在反应的前12小时内具有至少约45%或至少约50%或至少约55%或至少约60%的EEH或BisEEH转化。例如,新的SHC/HAC酶变体可在反应的前6小时内提供或该方法可在反应的前6小时内具有至少约30%的EEH或BisEEH转化。例如,新的SHC/HAC酶变体可在反应的前12小时内提供或该方法可在反应的前12小时内具有至少约35%或至少约45%或至少约50%或至少约55%的EEH或BisEEH转化。这可以是在与在相同的反应条件(例如相同的pH和温度)下使用两种酶(即新的SHC/HAC酶变体和SEQ ID NO:1或SEQ ID NO:10的酶)进行比较时或在与在其各自优化的反应条件(例如优化的pH和温度)(其可能彼此不同)下使用每种酶进行比较时。

[0215] EEH到(-)-降龙涎香醚的转化或BisEEH到龙涎香氧化物的转化可以例如使用如上所述的活性测定来确定,并且可以计算为每克原料的可回收产物克数(其可以计算作为摩尔转化率百分比)。

[0216] 如本文所用,本文中提及的均法呢醇底物到(-)-降龙涎香醚的99%/100%转化是指能够使用SHC/HAC酶或酶变体转化为(-)-降龙涎香醚的均法呢醇异构体(即EEH)的99%/100%转化。

[0217] 如本文所用,本文中提及的双均法呢醇底物到龙涎香氧化物的99%/100%转化是指能够使用SHC/HAC酶或酶变体转化为龙涎香氧化物的双均法呢醇异构体(即BisEEH)的99%/100%转化。

[0218] Aac 215G2 SHC/HAC的SHC/HAC酶变体的最适温度可以例如等于或大于约35℃。例如,Aac 215G2 SHC/HAC的SHC/HAC酶变体的最适温度范围可以是约40℃至约50℃,例如约42℃至约48℃或约44℃至约46℃。例如,Aac 215G2 SHC/HAC的SHC/HAC酶变体的最适温度可为约45℃。本文公开的(-)-降龙涎香醚或龙涎香氧化物的制备方法可以在SHC/HAC酶变体的最适温度下进行。

[0219] Aac 215G2 SHC/HAC的SHC/HAC酶变体的最适pH可以例如等于或大于约5.4。例如,Aac 215G2 SHC/HAC的SHC/HAC酶变体的最适pH范围可以为约5.2至约6.0,例如约5.4至约5.8,例如约5.6至约5.8。例如,Aac 215G2 SHC/HAC的SHC/HAC酶变体的最适pH可为约5.6或约5.8。可在SHC/HAC酶变体的最适pH下进行本文公开的用于制备(-)-降龙涎香醚或龙涎香氧化物的方法。

[0220] 本文公开的用于制备(-)-降龙涎香醚或龙涎香氧化物的方法的反应介质中十二烷基硫酸钠(SDS)的最佳浓度可以例如为约0.010w/w%至约0.10w/w%。例如,当底物(例如EEH或BisEEH)以4g/l使用、细胞达到10的OD_{650nm}时,SDS的最佳浓度可为约0.040w/w%至约0.080w/w%,例如约0.050w/w%。当以125g/l与250g/l细胞一起使用底物(例如EEH或BisEEH)时,本文公开的用于制备(-)-降龙涎香醚或龙涎香氧化物的方法的反应介质中十二烷基硫酸钠(SDS)的最佳浓度可以例如为约1.0w/w%至约1.5w/w%。例如,当以125g/l与250g/l细胞一起使用底物(例如EEH或BisEEH)时,SDS的最佳浓度可为约1.2w/w%至约1.4w/w%。

[0221] 本文公开的用于制备(-)-降龙涎香醚或龙涎香氧化物的方法可以在使用的特定酶的最佳温度范围或最佳温度和/或最佳pH范围或最佳pH和/或SDS最佳浓度范围或最佳SDS浓度下进行,如下面实施例中的表7或9或11中所列。

[0222] 在对应于SEQ ID NO:1的位置81、90、172、277、431、557和/或613的位置处具有新突变的其他变体

[0223] 如上所讨论的,令人惊讶地发现衍生自WO2016/170099中公开的Aac SHC/HAC酶变体的SHC/HAC酶变体(215G2 SHC/HAC酶变体)提供了用于将EEH转化为(-)-降龙涎香醚和将BisEEH转化为龙涎香氧化物的改善的酶活性。除了215G2 SHC/HAC酶变体中已经存在的氨基酸取代外,新的SHC/HAC酶变体还具有两个或三个氨基酸改变。

[0224] 预期具有在SEQ ID NO:1的位置81、90、172、277、431、557和613处鉴定的一个或多个新突变的其他SHC/HAC酶变体也将提供用于将EEH转化为(-)-降龙涎香醚或将BisEEH转化为龙涎香氧化物的酶活性(例如改善的酶活性)。

[0225] 特别地,预期来源于其他非Aac物种但具有在SEQ ID NO:1的位置81、90、172、277、431、557和613处鉴定的一个或多个新氨基酸改变的SHC/HAC酶变体也将提供用于将EEH转化为(-)-降龙涎香醚或将BisEEH转化为龙涎香氧化物的酶活性。特别地,预期衍生自ZmoSHC1、ZmoSHC2、BjpSHC、GmoSHC、TelSHC、ApaSHC1、BmeSHC、SalSHC或ApaSHCA具有在SEQ ID NO:1的位置81、90、172、277、431、557和613处鉴定的一个或多个新氨基酸改变的酶变体也将提供用于将EEH转化为(-)-降龙涎香醚或将BisEEH转化为龙涎香氧化物的酶活性。

[0226] 因此,本文提供了一种通过将EEH酶促转化为(-)-降龙涎香醚来制备(-)-降龙涎香醚的方法。本文还提供了一种通过将E,E-双均法呢醇酶促转化为龙涎香氧化物来制备龙涎香氧化物的方法。这些过程可以使用本文所述的任何野生型SHC/HAC酶或酶变体。

[0227] 此外,本文提供了与野生型SHC/HAC酶氨基酸序列具有至少约70.0%同一性的SHC/HAC酶变体,其中SHC/HAC酶变体氨基酸序列在选自对应于SEQ ID NO:1的位置81、90、172、277、431、557和613的位置的位置处具有相对于野生型SHC/HAC酶的一个或多个氨基酸改变。

[0228] 具体地,本文提供了制备(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物的方法,所述方法包括使用与野生型SHC/HAC酶氨基酸序列具有至少约70.0%同一性的SHC/HAC变体将EEH或包含EEH的混合物酶促转化为(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物,其中SHC/HAC酶变体氨基酸序列在选自对应于SEQ ID NO:1的位置81、90、172、277、431、557和613的位置的位置处具有相对于野生型SHC/HAC酶的一个或多个氨基酸改变。

[0229] 具体地,本文提供了制备龙涎香氧化物或包含龙涎香氧化物的混合物的方法,所述方法包括使用具有与野生型SHC/HAC酶氨基酸序列具有至少约70.0%同一性的SHC/HAC变体将BisEEH或包含BisEEH的混合物酶促转化为龙涎香氧化物或包含龙涎香氧化物的混合物,其中SHC/HAC酶变体氨基酸序列在选自对应于SEQ ID NO:1的位置81、90、172、277、431、557和613的位置的位置处具有相对于野生型SHC/HAC酶的一个或多个氨基酸改变。

[0230] SHC/HAC酶变体可以例如具有与野生型SHC/HAC酶氨基酸序列具有至少约70.0%同一性的氨基酸序列。例如,SHC/HAC酶变体具有与野生型SHC/HAC酶氨基酸序列具有至少约75.0%或至少约80.0%或至少约85.0%或至少约90.0%或至少约95.0%或至少约95.5%或至少约96.5%或至少约97.0%或至少约97.5%或至少约98.0%或至少约98.5%

或至少约99.0%同一性的氨基酸序列。

[0231] 例如,SHC/HAC酶变体可以例如具有与野生型SHC/HAC酶氨基酸序列具有小于100%同一性,例如等于或小于约99.5%或等于或小于约99.0%同一性的氨基酸序列。

[0232] 例如,SHC/HAC酶变体与野生型SHC/HAC酶氨基酸序列可以具有约70.0%至约99.5%或约80.0%至约99.0%或约85.0%至约98.5%或约90.0%至约98.0%的同一性。

[0233] 野生型SHC/HAC酶氨基酸序列可以是例如AacSHC(SEQ ID NO:1)、ZmoSHC1(SEQ ID NO:11)、ZmoSHC2(SEQ ID NO:12)、BjpSHC(SEQ ID NO:13)、GmoSHC(SEQ ID NO:14)、TelSHC(SEQ ID NO:19)或ApaSHC1(SEQ ID NO:20)、BmeSHC(SEQ ID NO:28)、SalSHC(SEQ ID NO:29)、或ApaSHCA(SEQ ID NO:30)的氨基酸序列。特别地,野生型SHC/HAC酶氨基酸序列可以是AacSHC(SEQ ID NO:1)的氨基酸序列。

[0234] 因此,在某些实施方案中,SHC/HAC酶变体可以具有与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19或SEQ ID NO:20具有至少约70.0%同一性的氨基酸序列。例如,SHC/HAC酶变体的氨基酸序列与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30具有至少约75.0%或至少约80.0%或至少约85.0%或至少约90.0%或至少约95.0%或至少约95.5%或至少约96.5%或至少约97.0%或至少约97.5%或至少约98.0%或至少约98.5%或至少约99.0%的同一性。

[0235] 例如,SHC/HAC酶变体可以例如具有与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30具有小于100%同一性,例如等于或小于约99.5%或等于或小于约99.0%同一性的氨基酸序列。

[0236] 例如,SHC/HAC酶变体可与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30具有约70.0%至约99.5%或约80.0%至约99.0%或约85.0%至约98.5%或约90.0%至约98.0%的同一性。

[0237] 关于多肽或核苷酸序列的“百分比(%)同一性”分别定义为在比对序列并在必要时引入空位以达到最大序列同一性百分比之后,并且不考虑将任何保守取代作为序列同一性的一部分时,候选序列中与参考序列中的氨基酸或核苷酸相同的氨基酸或核苷酸的百分比。用于确定序列同一性百分比目的的比对可以以本领域技术范围内的各种方式实现,例如,使用可公开获得的计算机软件。本领域技术人员可以确定用于测量比对的适当参数,包括在被比较的序列的全长上实现最大比对所需的任何算法。术语“多肽”和“蛋白质”在本文中可互换使用,并且是指氨基酸的任何肽连接的链,无论长度或翻译后修饰。

[0238] 核苷酸和氨基酸序列的相似性,即序列同一性的百分比,可以通过序列比对来确定。这样的比对可以用几种本领域已知的算法进行,优选使用Karlin和Altschul的数学算法(Karlin&Altschul(1993)Proc.Natl.Acad.Sci.USA 90:5873-5877),使用hmmalign(HMMER软件包,http://hmmer.wustl.edu/),或使用例如可获自https://www.ebi.ac.uk/Tools/msa/clustalo/的CLUSTAL算法(Thompson,J.D.,Higgins,D.G.&Gibson,T.J.(1994)Nucleic Acids Res.22,4673-80),或GAP程序(爱荷华大学的数学算法)或Myers和Miller(1989-Cabios 4:11-17)的数学算法。使用的优选参数是在https://www.ebi.ac.uk/

Tools/msa/clustalo/上设置的默认参数。

[0239] 可以使用例如BLAST、BLAT或BlastZ(或BlastX)计算序列同一性百分比。Altschul等人(1990)J.Mol.Biol.215,403-410的BLASTN和BLASTP程序中并入了类似的算法。可以使用BLASTN程序(得分＝100,字长＝12)进行BLAST多核苷酸搜索以获得与编码相关蛋白质的那些核酸同源的多核苷酸序列。可以使用BLASTP程序(得分＝50,字长＝3)进行BLAST蛋白质搜索以获得与多肽同源的氨基酸序列。

[0240] 为了获得用于比较目的的空位比对,如Altschul等人(1997)Nucleic Acids Res.25,3389-3402中所述,可以使用空位BLAST。当使用BLAST和空位BLAST程序时,使用各自程序的默认参数。序列匹配分析可以通过已建立的同源映射技术如Shuffle-LAGAN(Brudno M.,Bioinformatics 2003b,19Suppl 1:154-162)或Markov随机场来补充。当在本申请中提及序列同一性的百分比时,如果没有特别指出,这些百分比是相对于较长序列的全长计算的。

[0241] 在特定实施方案中,使用CLUSTAL O(版本1.2.4)确定两个序列之间的%同一性。

[0242] 在某些实施方案中,SHC/HAC酶变体与野生型SHC/HAC酶相比,例如与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30相比,可以具有等于或小于约200个氨基酸改变。例如,与野生型SHC/HAC酶相比,例如与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30相比,SHC/HAC酶变体可以具有等于或小于约150或等于或小于约120或等于或小于约100或等于或小于约95或等于或小于约90或等于或小于约85或等于或小于约80或等于或小于约75或等于或小于约70或等于或小于约65或等于或小于约60或等于或小于约55或等于或小于约50或等于或小于约45或等于或小于约40或等于或小于约35或等于或小于约30或等于或小于约25或等于或小于约20或等于或小于约15或等于或小于约10个氨基酸改变。

[0243] 例如,与野生型SHC/HAC酶相比,例如与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30相比,SHC/HAC酶变体可以具有至少约1个或至少约2个或至少约3个或至少约4个或至少约5个或至少约6个氨基酸改变。

[0244] 例如,与野生型SHC/HAC酶相比,例如与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20相比,SHC/HAC酶变体可以具有约1个至约30个氨基酸改变。例如,与野生型SHC/HAC酶相比,例如与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20或SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30相比,SHC/HAC酶变体可以具有约2个至约25个氨基酸改变。例如,与野生型SHC/HAC酶相比,例如与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30相比,SHC/HAC酶变体可以具有约3个至约20个氨基酸改变。例如,与野生型SHC/HAC酶相比,例如与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30相比,SHC/HAC酶变体可以具有约4个至约15个氨基酸改变。例如,与野生型SHC/HAC酶相

比,例如与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30相比,SHC/HAC酶变体可以具有约5个至约10个氨基酸改变。

[0245] 例如,氨基酸改变可以是如上所述的插入、缺失和/或取代。例如,氨基酸改变可以是取代,例如非保守取代。

[0246] 在某些实施方案中,与野生型SHC/HAC酶相比(例如与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30相比),唯一的氨基酸改变是取代(即没有插入或缺失)。

[0247] 氨基酸改变是相对于参考序列定义的。相对于参考序列的氨基酸改变是指变体序列的氨基酸序列与参考序列不同。

[0248] 参考序列和变体序列中的氨基酸可以被分配编号,其中编号从多肽的N末端的氨基酸开始(即多肽N末端的氨基酸被编号为1,下一个氨基酸被编号为2等)。参考序列的“位置”是指参考序列中存在的特定氨基酸残基,如通过参考序列中氨基酸的特定编号所识别的。变体序列的“位置”是指存在于变体序列中的特定氨基酸残基,如通过变体序列中氨基酸的特定编号所识别的。

[0249] 由于与参考序列相比,变体序列可以包括缺失或插入,因此变体序列中的氨基酸的编号可能与参考序列中的相同氨基酸不同。例如,如果在SEQ ID NO:1的氨基酸131和132之间插入氨基酸,则插入的后面的氨基酸将在变体序列中具有编号133,而其在参考序列中保留编号132。在本示例中,与参考序列的位置132对应的变体序列的位置为位置133。因此,变体序列中从参考序列中保留下来的氨基酸可以参照参考序列的“对应位置”来定义。换言之,变体序列中的“位置”可以通过参照参考序列中的“对应位置”来定义。特别地,与参考序列相比变体序列中的取代可以通过参照参考序列的“对应位置”来定义,尽管参考序列中存在任何插入和/或缺失。在参考序列的氨基酸被缺失的情况下,变体序列中没有“对应位置”。如果与参考序列相比没有插入或缺失(即只有取代),则参考序列的“对应位置”将与变体序列中的位置相同。

[0250] 来自不同物种的野生型SHC/HAC酶具有不同的多肽长度。可以使用如上所述的算法比对野生型序列,以便鉴定两种不同野生型SHC/HAC酶中的“对应位置”。因此,与参考序列中的位置相对应的变体序列的位置处的氨基酸可以是例如不同的氨基酸残基和/或可以具有与参考序列的编号不同的编号。例如,AacSHC(SEQ ID NO:1)的位置132的氨基酸M可以对应于ZmoSHC1(SEQ ID NO:11)的位置185的氨基酸Y。

[0251] 因此可以相对于两个不同的参考序列来定义氨基酸改变。例如,氨基酸改变可以是与第一参考序列(例如衍生变体的野生型SHC/HAC酶序列)相比的改变,并且可以通过参照第二参考序列(例如AacSHC(SEQ ID NO:1))来定义变体序列中氨基酸改变的位置。因此,SHC/HAC酶变体中的氨基酸改变可以相对于第一种野生型SHC/HAC酶位于通过参照第二种野生型SHC/HAC酶定义的位置。

[0252] SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC酶氨基酸序列在选自对应于SEQ ID NO:1的位置81、90、172、277、431、557和613的位置的位置处具有一个或多个氨基酸改变。例如,氨基酸改变可以在选自对应于SEQ ID NO:1的位置81、431、557和613的位置的一

个或多个位置处。氨基酸改变可以例如是取代，例如非保守取代。野生型SHC/HAC酶氨基酸序列可以是例如SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30。例如，野生型序列可以是SEQ ID NO:1。

[0253] 在某些实施方案中，相对于野生型SHC/HAC酶氨基酸序列，SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557和对应于SEQ ID NO:1的位置81、431或613的至少一个位置处具有氨基酸改变。氨基酸改变可以例如是取代，例如非保守取代。野生型SHC/HAC酶氨基酸序列可以是例如SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30。例如，野生型序列可以是SEQ ID NO:1。

[0254] 例如，相对于野生型SHC/HAC酶氨基酸序列，SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557的位置和对应于SEQ ID NO:1的位置81、431或613的一个位置处可以具有氨基酸改变。氨基酸改变可以例如是取代，例如非保守取代。野生型SHC/HAC酶氨基酸序列可以是例如SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30。例如，野生型序列可以是SEQ ID NO:1。

[0255] 例如，相对于野生型SHC/HAC酶氨基酸序列，SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557的位置和选自对应于SEQ ID NO:1的位置81、431和613的位置的两个位置处可以具有氨基酸改变。氨基酸改变可以例如是取代，例如非保守取代。野生型SHC/HAC酶氨基酸序列可以是例如SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30。例如，野生型序列可以是SEQ ID NO:1。

[0256] 例如，相对于野生型SHC/HAC酶氨基酸序列，SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557的位置和对应于SEQ ID NO:1的位置81、431和613的所有位置处可以具有氨基酸改变。氨基酸改变可以例如是取代，例如非保守取代。野生型SHC/HAC酶氨基酸序列可以是例如SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30。例如，野生型序列可以是SEQ ID NO:1。

[0257] 例如，相对于野生型SHC/HAC酶氨基酸序列，SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置90和613的位置处具有氨基酸改变。氨基酸改变可以例如是取代，例如非保守取代。野生型SHC/HAC酶氨基酸序列可以是例如SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30。例如，野生型序列可以是SEQ ID NO:1。

[0258] 例如，相对于野生型SHC/HAC酶氨基酸序列，SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置172和277的位置处具有氨基酸改变。氨基酸改变可以例如是取代，例如非保守取代。野生型SHC/HAC酶氨基酸序列可以是例如SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30。例如，野生型序列可以是SEQ ID NO:1。

[0259] 例如，相对于野生型SHC/HAC酶氨基酸序列，SHC/HAC酶变体氨基酸序列在对应于

SEQ ID NO:1的位置557和431的位置处具有氨基酸改变。氨基酸改变可以例如是取代,例如非保守取代。野生型SHC/HAC酶氨基酸序列可以是例如SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30。例如,野生型序列可以是SEQ ID NO:1。

[0260] 例如,相对于野生型SHC/HAC酶氨基酸序列,SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557和613的位置处具有氨基酸改变。氨基酸改变可以例如是取代,例如非保守取代。野生型SHC/HAC酶氨基酸序列可以是例如SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30。例如,野生型序列可以是SEQ ID NO:1。

[0261] 例如,相对于野生型SHC/HAC酶氨基酸序列,SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置81、557和613的位置处具有氨基酸改变。氨基酸改变可以例如是取代,例如非保守取代。野生型SHC/HAC酶氨基酸序列可以是例如SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30。例如,野生型序列可以是SEQ ID NO:1。

[0262] 例如,相对于野生型SHC/HAC酶氨基酸序列,SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置81、431和557的位置处具有氨基酸改变。氨基酸改变可以例如是取代,例如非保守取代。野生型SHC/HAC酶氨基酸序列可以是例如SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30。例如,野生型序列可以是SEQ ID NO:1。

[0263] 在对应于SEQ ID NO:1的位置557的位置处相对于野生型SHC/HAC酶氨基酸序列的氨基酸改变可以是例如野生型SHC/HAC酶的氨基酸取代为不同的氨基酸(X)。如上所述,由于野生型序列可能具有与SEQ ID NO:1不同的长度,并且由于变体可能另外包含插入和/或缺失,因此变体序列中新氨基酸(X)的编号可能不是557。

[0264] SHC/HAC酶变体氨基酸序列中对应于SEQ ID NO:1的位置557的位置处的新氨基酸(X)可以是例如Met、Val、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如,SHC/HAC酶变体中的新氨基酸(X)可以是中性亲水性氨基酸(即Cys、Ser、Thr、Asn或Gln)。例如,SHC/HAC酶变体中的新氨基酸可以是苏氨酸。

[0265] 在对应于SEQ ID NO:1的位置81的位置处相对于野生型SHC/HAC酶氨基酸序列的氨基酸改变可以是例如野生型SHC/HAC酶的氨基酸取代为不同的氨基酸(X)。如上所述,由于野生型序列可能具有与SEQ ID NO:1不同的长度,并且由于变体可能另外包含插入和/或缺失,因此新氨基酸(X)的编号可能不是81。

[0266] SHC/HAC酶变体氨基酸序列中对应于SEQ ID NO:1的位置81的位置处的新氨基酸(X)可以是例如Met、Ala、Val、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp或Phe。例如,SHC/HAC酶变体中的新氨基酸(X)可以是碱性氨基酸(即His、Lys或Arg)。例如,SHC/HAC酶变体中的新氨基酸可以是组氨酸。

[0267] 在对应于SEQ ID NO:1的位置90的位置处相对于野生型SHC/HAC酶氨基酸序列的氨基酸改变可以是例如野生型SHC/HAC酶的氨基酸取代为不同的氨基酸(X)。如上所述,由于野生型序列可能具有与SEQ ID NO:1不同的长度,并且由于变体可能另外包含插入和/或缺失,因此变体序列中新氨基酸(X)的编号可能不是90。

[0268] SHC/HAC酶变体氨基酸序列中对应于SEQ ID NO:1的位置90的位置处的新氨基酸(X)可以例如是Met、Ala、Val、Leu、Ile、Cys、Ser、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如,SHC/HAC酶变体中的新氨基酸(X)可以是疏水性氨基酸(即Met、Ala、Val、Leu、Ile)。例如,SHC/HAC酶变体中的新氨基酸可以是丙氨酸。

[0269] 在对应于SEQ ID NO:1的位置172的位置处相对于野生型SHC/HAC酶氨基酸序列的氨基酸改变可以是例如野生型SHC/HAC酶的氨基酸取代为不同的氨基酸(X)。如上所述,由于野生型序列可能具有与SEQ ID NO:1不同的长度,并且由于变体可能另外包含插入和/或缺失,因此变体序列中新氨基酸(X)的编号可能不是172。

[0270] SHC/HAC酶变体氨基酸序列中对应于SEQ ID NO:1的位置172的位置处的新氨基酸(X)可以例如是Met、Val、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如,SHC/HAC酶变体中的新氨基酸(X)可以是中性亲水性氨基酸(即Cys、Ser、Thr、Asn、Gln)。例如,SHC/HAC酶变体中的新氨基酸可以是苏氨酸。

[0271] 在对应于SEQ ID NO:1的位置277的位置处相对于野生型SHC/HAC酶氨基酸序列的氨基酸改变可以是例如野生型SHC/HAC酶的氨基酸取代为不同的氨基酸(X)。如上所述,由于野生型序列可能具有与SEQ ID NO:1不同的长度,并且由于变体可能另外包含插入和/或缺失,因此变体序列中新氨基酸(X)的编号可能不是277。

[0272] SHC/HAC酶变体氨基酸序列中对应于SEQ ID NO:1的位置277的位置处的新氨基酸(X)可以是例如Ala、Val、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如,SHC/HAC酶变体中的新氨基酸(X)可以是碱性氨基酸(即His、Lys、Arg)。例如,SHC/HAC酶变体中的新氨基酸可以是赖氨酸。

[0273] 在对应于SEQ ID NO:1的位置431的位置处相对于野生型SHC/HAC酶氨基酸序列的氨基酸改变可以例如是野生型SHC/HAC酶的氨基酸取代为不同的氨基酸(X)。如上所述,由于野生型序列可能具有与SEQ ID NO:1不同的长度,并且由于变体可能另外包含插入和/或缺失,因此新氨基酸(X)的编号可能不是431。

[0274] SHC/HAC酶变体氨基酸序列中对应于SEQ ID NO:1的位置431的位置处的新氨基酸(X)可以例如是Met、Ala、Val、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如,SHC/HAC酶变体中的新氨基酸(X)可以是疏水性氨基酸(即Met、Ala、Val、Leu或Ile)。例如,SHC/HAC酶变体中的新氨基酸可以是亮氨酸。

[0275] 在对应于SEQ ID NO:1的位置613的位置处相对于野生型SHC/HAC酶氨基酸序列的氨基酸改变可以是例如野生型SHC/HAC酶的氨基酸取代为不同的氨基酸(X)。如上所述,由于野生型序列可能具有与SEQ ID NO:1不同的长度并且由于变体可能另外包含插入和/或缺失,所以新氨基酸(X)的编号可能不是613。

[0276] SHC/HAC酶变体氨基酸序列中对应于SEQ ID NO:1的位置613的位置处的新氨基酸(X)可以例如是Met、Ala、Val、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Gly、Pro、Trp、Tyr或Phe。例如,SHC/HAC酶变体中的新氨基酸(X)可以是中性亲水性氨基酸(即Cys、Ser、Thr、Asn或Gln)。例如,SHC/HAC酶变体中的新氨基酸可以是丝氨酸。

[0277] 野生型ZmoSHC1、ZmoSHC2、BjpSHC、GmoSHC、TelSHC和ApaSHC1序列(分别为SEQ ID NO:11、12、13、14、19和20)中对应于AacSHC(SEQ ID NO:1)的氨基酸(例如AacSHC的位置81、431、557和613的氨基酸)的氨基酸和位置如图9A所示。野生型ZmoSHC1、ZmoSHC2、BjpSHC、

GmoSHC、TelSHC、ApaSHC1、BmeSHC、SalSHC和ApaSHCA序列(分别为SEQ ID NO:11、12、13、14、19、20、28、29和30)中对应于AacSHC(SEQ ID NO:1)的氨基酸(例如AacSHC的位置81、431、557和613的氨基酸)的氨基酸和位置如图9B所示。

[0278] 野生型AacSHC中的氨基酸位置81、90、132、224、172、277、431、432、557和613用黑色背景上的白色字母突出显示。因此,突出显示的氨基酸正上方或下方的氨基酸是ZmoSHC2、BjaSHC、GmoSHC、ApaSHC1、ApaSHC1、ZmoSHC1和TelSHC中对应于AacSHC(SEQ ID NO:1)的位置81、90、132、224、172、277、431、432、557和613的氨基酸和位置。例如,BjaSHC、GmoSHC、ApaSHC1和ZmoSHC1中与AacSHC(SEQ ID NO:1)的位置81的氨基酸Y相对应的氨基酸是Y。ZmoSHC2和TelSHC中与AacSHC(SEQ ID NO:1)的位置81的氨基酸Y相对应的氨基酸是F。TelSHC的位置84是TelSHC的对应于AacSHC(SEQ ID NO:1)的位置81的位置。

[0279] 在某些实施方案中,215G2 SHC/HAC酶变体中的一个或多个原始氨基酸改变(与野生型AacSHC(SEQ ID NO:1)相比)可以保留在本文所述的新的SHC/HAC酶变体中。与野生型AacSHC相比,215G2SHC/HAC酶变体中的氨基酸改变是取代M132R、A224V和I432T(即,位置132的M残基取代为R残基,位置224的A残基取代为V残基和位置432的I残基取代为T残基)。

[0280] 因此,相对于野生型SHC/HAC酶氨基酸序列,SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置132、224和432的位置可能具有一个或多个氨基酸改变。例如,氨基酸改变可以是取代。

[0281] 例如,相对于野生型SHC/HAC酶氨基酸序列,SHC/HAC酶变体氨基酸序列在选择为对应于SEQ ID NO:1的位置132、224和432的位置处可以具有一个、两个或三个氨基酸改变。

[0282] 例如,相对于野生型SHC/HAC酶氨基酸序列,SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置132和432的位置处可以具有氨基酸改变。

[0283] 例如,相对于野生型SHC/HAC酶氨基酸序列,SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置132、224和432的位置处可以具有氨基酸改变。

[0284] 在对应于SEQ ID NO:1的位置132的位置处相对于野生型SHC/HAC酶氨基酸序列的氨基酸改变可以是例如野生型SHC/HAC酶的氨基酸取代为不同的氨基酸(X)。如上所述,由于野生型序列可能具有与SEQ ID NO:1不同的长度,并且由于变体可能另外包含插入和/或缺失,因此变体序列中新氨基酸(X)的编号可能不是132。

[0285] SHC/HAC酶变体氨基酸序列中对应于SEQ ID NO:1的位置132的位置处的新氨基酸(X)可以是例如Ala、Val、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如,SHC/HAC酶变体中的新氨基酸(X)可以是碱性氨基酸(即His、Lys或Arg)。例如,SHC/HAC酶变体中的新氨基酸可以是精氨酸。

[0286] 在对应于SEQ ID NO:1的位置224的位置处相对于野生型SHC/HAC酶氨基酸序列的氨基酸改变可以是例如野生型SHC/HAC酶的氨基酸取代为不同的氨基酸(X)。如上所述,由于野生型序列可能具有与SEQ ID NO:1不同的长度,并且由于变体可能另外包含插入和/或缺失,因此变体序列中新氨基酸(X)的编号可能不是224。

[0287] SHC/HAC酶变体氨基酸序列中对应于SEQ ID NO:1的位置224的位置处的新氨基酸(X)可以是例如Met、Val、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如,SHC/HAC酶变体中的新氨基酸(X)可以是疏水性氨基酸(即Met、Val、Leu或Ile)。例如,SHC/HAC酶变体中的新氨基酸可以是缬氨酸。

[0288] 在对应于SEQ ID NO:1的位置432的位置处相对于野生型SHC/HAC酶氨基酸序列的氨基酸改变可以是例如野生型SHC/HAC酶的氨基酸取代为不同的氨基酸(X)。如上所述,由于野生型序列可能具有与SEQ ID NO:1不同的长度,并且由于变体可能另外包含插入和/或缺失,因此变体序列中新氨基酸(X)的编号可能不是432。

[0289] SHC/HAC酶变体氨基酸序列中对应于SEQ ID NO:1的位置432的位置处的新氨基酸(X)可以是例如Met、Ala、Val、Leu、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如,SHC/HAC酶变体中的新氨基酸(X)可以是中性亲水性氨基酸(即Cys、Ser、Thr、Asn或Gln)。例如,SHC/HAC酶变体中的新氨基酸可以是苏氨酸。

[0290] 野生型ZmoSHC1、ZmoSHC2、BjpSHC、GmoSHC、TelSHC和ApaSHC1序列(分别为SEQ ID NO:11、12、13、14、19和20)中对应于AacSHC(SEQ ID NO:1)的氨基酸(例如在AacSHC的位置132、224和432处)的氨基酸和位置如图9所示。

[0291] 例如,本文所述的SHC/HAC酶变体可以在WO2016/170099中鉴定的AacSHC的其他位置处另外具有一个或多个其他氨基酸改变(例如取代)。

[0292] 因此,与野生型SHC/HAC酶相比,本文所述的SHC/HAC酶变体在对应于SEQ ID NO:1的位置77、92、129、579、601和/或605的位置处可能具有一个或多个进一步的氨基酸改变。例如,相对于野生型SHC/HAC酶,SHC/HAC酶变体可以在对应于SEQ ID NO:1的位置77、92、129、579、601和/或605的位置处具有一个或多个氨基酸取代(例如保守或非保守取代)。

[0293] 相对于野生型SHC/HAC酶氨基酸序列在对应于SEQ ID NO:1的位置77的位置处的氨基酸改变可以是例如野生型SHC/HAC酶的氨基酸取代为不同的氨基酸(X)。如上所述,由于野生型序列可能具有与SEQ ID NO:1不同的长度,并且由于变体可能另外包含插入和/或缺失,因此变体序列中新氨基酸(X)的编号可能不是77。

[0294] SHC/HAC酶变体氨基酸序列中对应于SEQ ID NO:1的位置77的位置处的新氨基酸(X)可以例如是Met、Ala、Val、Leu、Ile、Cys、Ser、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如,SHC/HAC酶变体中的新氨基酸(X)可以是疏水性氨基酸(即Met、Ala、Val、Leu或Ile)。例如,SHC/HAC酶变体中的新氨基酸可以是丙氨酸。

[0295] 相对于野生型SHC/HAC酶氨基酸序列在对应于SEQ ID NO:1的位置92的位置处的氨基酸改变可以例如是野生型SHC/HAC酶的氨基酸取代为不同的氨基酸(X)。如上所述,由于野生型序列可能具有与SEQ ID NO:1不同的长度,并且由于变体可能另外包含插入和/或缺失,因此变体序列中新氨基酸(X)的编号可能不是92。

[0296] SHC/HAC酶变体氨基酸序列中对应于SEQ ID NO:1的位置92的位置处的新氨基酸(X)可以例如是Met、Ala、Val、Leu、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如,SHC/HAC酶变体中的新氨基酸(X)可以是疏水性氨基酸(即Met、Ala、Val、Leu或Ile)。例如,SHC/HAC酶变体中的新氨基酸可以是缬氨酸。

[0297] 相对于野生型SHC/HAC酶氨基酸序列在对应于SEQ ID NO:1的位置129的位置处的氨基酸改变可以是例如野生型SHC/HAC酶的氨基酸取代为不同的氨基酸(X)。如上所述,由于野生型序列可能具有与SEQ ID NO:1不同的长度,并且由于变体可能另外包含插入和/或缺失,因此变体序列中新氨基酸(X)的编号可能不是129。

[0298] SHC/HAC酶变体氨基酸序列中对应于SEQ ID NO:1的位置129的位置处的新氨基酸(X)可以是例如Met、Ala、Val、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、

Pro、Trp或Tyr。例如，SHC/HAC酶变体中的新氨基酸(X)可以是疏水性氨基酸(即Met、Ala、Val、Leu或Ile)。例如，SHC/HAC酶变体中的新氨基酸可以是亮氨酸。

[0299] 相对于野生型SHC/HAC酶氨基酸序列在对应于SEQ ID NO:1的位置579的位置处的氨基酸改变可以例如是野生型SHC/HAC酶的氨基酸取代为不同的氨基酸(X)。如上所述，由于野生型序列可能具有与SEQ ID NO:1不同的长度，并且由于变体可能另外包含插入和/或缺失，因此变体序列中新氨基酸(X)的编号可能不是579。

[0300] SHC/HAC酶变体氨基酸序列中对应于SEQ ID NO:1的位置579的位置处的新氨基酸(X)可以是例如Met、Ala、Val、Leu、Ile、Cys、Ser、Thr、Asn、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如，SHC/HAC酶变体中的新氨基酸(X)可以是碱性氨基酸(即His、Lys、Arg)。例如，SHC/HAC酶变体中的新氨基酸可以是组氨酸。

[0301] 相对于野生型SHC/HAC酶氨基酸序列在对应于SEQ ID NO:1的位置601的位置处的氨基酸改变可以是例如野生型SHC/HAC酶的氨基酸取代为不同的氨基酸(X)。如上所述，由于野生型序列可能具有与SEQ ID NO:1不同的长度，并且由于变体可能另外包含插入和/或缺失，因此变体序列中新氨基酸(X)的编号可能不是601。

[0302] SHC/HAC酶变体氨基酸序列中对应于SEQ ID NO:1的位置601的位置处的新氨基酸(X)可以例如是Met、Ala、Val、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp或Tyr。例如，SHC/HAC酶变体中的新氨基酸(X)可以是芳香族氨基酸(即Trp、Tyr、Phe)。例如，SHC/HAC酶变体中的新氨基酸可以是酪氨酸。

[0303] 相对于野生型SHC/HAC酶氨基酸序列在对应于SEQ ID NO:1的位置605的位置处的氨基酸改变可以例如是野生型SHC/HAC酶的氨基酸取代为不同的氨基酸(X)。如上所述，由于野生型序列可能具有与SEQ ID NO:1不同的长度，并且由于变体可能另外包含插入和/或缺失，因此变体序列中新氨基酸(X)的编号可能不是605。

[0304] SHC/HAC酶变体氨基酸序列中对应于SEQ ID NO:1的位置605的位置处的新氨基酸(X)可以例如是Met、Ala、Val、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp或Tyr。例如，SHC/HAC酶变体中的新氨基酸(X)可以是芳香族氨基酸(即Trp、Tyr、Phe)。例如，SHC/HAC酶变体中的新氨基酸可以是色氨酸。

[0305] 野生型TelSHC(SEQ ID NO:19)、野生型ApaSHC1(SEQ ID NO:20)、野生型ZmoSHC1(SEQ ID NO:11)、野生型ZmoSHC2(SEQ ID NO:12)、野生型BjpSHC(SEQ ID NO:13)和野生型GmoSHC(SEQ ID NO:14)序列中对应于野生型AacSHC中的氨基酸位置81、90、132、224、172、277、431、432、557和613的氨基酸位置在图9A中在野生型AacSHC(SEQ ID NO:1)中用黑色背景上的白色字母突出显示。

[0306] 野生型TelSHC(SEQ ID NO:19)、野生型ApaSHC1(SEQ ID NO:20)、野生型ZmoSHC1(SEQ ID NO:11)、野生型ZmoSHC2(SEQ ID NO:12)、野生型BjpSHC(SEQ ID NO:13)、野生型GmoSHC(SEQ ID NO:14)、野生型BmeSHC(SEQ ID NO:28)、野生型SalSHC(SEQ ID NO:29)和野生型ApaSHCA(SEQ ID NO:30)中对应于野生型AacSHC中氨基酸位置81、90、132、224、172、277、431、432、557和613的氨基酸位置在图9B中在野生型AacSHC(SEQ ID NO:1)中用黑色背景上的白色字母突出显示。

[0307] SHC/HAC酶变体可以例如相对于野生型SHC/HAC酶氨基酸序列在对应于SEQ ID NO:1的位置37、174和/或601的位置处具有一个或多个进一步的氨基酸改变。例如，相对于

野生型SHC/HAC酶氨基酸序列，SHC/HAC酶变体在对应于SEQ ID NO:1的位置37、174和/或601的位置处具有一个或多个氨基酸取代(例如保守或非保守取代)。

[0308] 相对于野生型SHC/HAC酶氨基酸序列在对应于SEQ ID NO:1的位置37的位置处的氨基酸改变可以是例如野生型SHC/HAC酶的氨基酸取代为不同的氨基酸(X)。如上所述，由于野生型序列可能具有与SEQ ID NO:1不同的长度，并且由于变体可能另外包含插入和/或缺失，因此新氨基酸(X)的编号可能不是37。

[0309] SHC/HAC酶变体氨基酸序列中对应于SEQ ID NO:1的位置37的位置处的新氨基酸(X)可以例如是Met、Ala、Val、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如，SHC/HAC酶变体中的新氨基酸(X)可以是中性亲水性氨基酸(即Cys、Ser、Thr、Asn或Gln)，例如谷氨酰胺。

[0310] 相对于野生型SHC/HAC酶氨基酸序列在对应于SEQ ID NO:1的位置174的位置处的氨基酸改变可以例如是野生型SHC/HAC酶的氨基酸取代为不同的氨基酸(X)。如上所述，由于野生型序列可能具有与SEQ ID NO:1不同的长度并且由于变体可能另外包含插入和/或缺失，因此新氨基酸(X)的编号可能不是174。

[0311] SHC/HAC酶变体氨基酸序列中对应于SEQ ID NO:1的位置174的位置处的新氨基酸(X)可以是例如Met、Ala、Leu、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp、Tyr或Phe。例如，SHC/HAC酶变体中的新氨基酸(X)可以是疏水性氨基酸(即Met、Ala、Val、Leu或Ile)，例如异亮氨酸。

[0312] 相对于野生型SHC/HAC酶氨基酸序列在对应于SEQ ID NO:1的位置601的位置处的氨基酸改变可以是例如野生型SHC/HAC酶的氨基酸取代为不同的氨基酸(X)。如上所述，由于野生型序列可能具有与SEQ ID NO:1不同的长度并且由于变体可能另外包含插入和/或缺失，所以新氨基酸(X)的编号可能不是601。

[0313] SHC/HAC酶变体氨基酸序列中对应于SEQ ID NO:1的位置601的位置处的新氨基酸(X)可以例如是Met、Ala、Leu、Val、Ile、Cys、Ser、Thr、Asn、Gln、Asp、Glu、His、Lys、Arg、Gly、Pro、Trp或Tyr。例如，SHC/HAC酶变体中的新氨基酸(X)可以是芳香族氨基酸(即Trp、Tyr、Phe)，例如酪氨酸。

[0314] 野生型TelSHC(SEQ ID NO:19)、野生型ApaSHC1(SEQ ID NO:20)、野生型ZmoSHC1(SEQ ID NO:11)、野生型ZmoSHC2(SEQ ID NO:12)、野生型BjpSHC(SEQ ID NO:13)和野生型GmoSHC(SEQ ID NO:14)序列中对应于野生型AacSHC中的氨基酸位置81、90、132、224、172、277、431、432、557和613的氨基酸位置在图9A中在野生型AacSHC(SEQ ID NO:1)中用黑色背景上的白色字母突出显示。

[0315] 野生型TelSHC(SEQ ID NO:19)、野生型ApaSHC1(SEQ ID NO:20)、野生型ZmoSHC1(SEQ ID NO:11)、野生型ZmoSHC2(SEQ ID NO:12)、野生型BjpSHC(SEQ ID NO:13)、野生型GmoSHC(SEQ ID NO:14)、野生型BmeSHC(SEQ ID NO:28)、野生型SalSHC(SEQ ID NO:29)和野生型ApaSHCA(SEQ ID NO:30)中对应于野生型AacSHC中氨基酸位置81、90、132、224、172、277、431、432、557和613的氨基酸位置在图9B中在野生型AacSHC(SEQ ID NO:1)中用黑色背景上的白色字母突出显示。

[0316] 设想了本文所述的氨基酸改变的任何组合。具体而言，设想了在对应于本文的AacSHC和WO2016/17009中鉴定的氨基酸改变组合的位置处的氨基酸改变组合。

[0317] 在某些实施方案中,除了以下氨基酸取代,SHC/HAC酶变体与SEQ ID NO:1相同:

[0318] (i)M132R、A224V、I432T、A557T和H4331L(SEQ ID NO:2);或者

[0319] (ii)M132R、A224V、I432T、A557T和R613S(SEQ ID NO:3);或者

[0320] (iii)M132R、A224V、I432T、A557T、Y81H和R613S(SEQ ID NO:4);或者

[0321] (iv)M132R、A224V、I432T、A557T、Y81H和H431L(SEQ ID NO:5);或者

[0322] (v)M132R、A224V、I432T、T90A和R613S(SEQ ID NO:17);或者

[0323] (vi)M132R、A224V、I432T、A172T和M277K(SEQ ID NO:18)。

[0324] 例如,与215G2 SHC酶相比,新的SHC/HAC酶变体可以具有用于将EEH转化为(-)-降龙涎香醚或将BisEEH转化为龙涎香氧化物的增加的酶活性。增加的酶活性可以指EEH到(-)-降龙涎香醚的酶促转化或BisEEH到龙涎香氧化物的酶促转化的任何方面,包括例如EEH或BisEEH的增加的总转化、EEH或BisEEH的增加的转化率(例如,在反应的前6小时或前12小时内),(-)-降龙涎香醚或龙涎香氧化物的增加的产量,和副产物的减少的产量。增加的酶活性通常可以转化为增加的生产率,这可以根据每升反应容量和每小时生物转化时间产生的(-)-降龙涎香醚或龙涎香氧化物,或每小时反应时间(即加入底物后的时间)的每升反应容量和反应中使用的每克生物催化剂产生的(-)-降龙涎香醚或龙涎香氧化物来定义。

[0325] 例如,与215G2 SHC酶相比,新的SHC/HAC酶变体可以提供增加的EEH或BisEEH转化。因此,与使用215G2 SHC酶的方法相比,本文所述的方法可以具有增加的EEH或BisEEH转化水平。例如,与215G2 SHC酶相比,新的SHC/HAC酶变体可以提供增加的EEH或BisEEH转化率。因此,与215G2 SHC酶相比,本文所述的方法可以具有增加的EEH或BisEEH转化率。

[0326] 例如,新的SHC/HAC酶变体可与215G2 SHC酶相比在反应的前4小时或前6小时或前8小时或前12小时或前24小时内提供增加的EEH或BisEEH转化率。因此,与215G2 SHC酶相比,本文所述的方法在反应的前4小时或前6小时或前8小时或前12小时或前24小时内可以具有增加的EEH或BisEEH转化率。这可以是在与在相同的反应条件(例如相同的pH和温度)下使用两种酶(即新的SHC/HAC酶变体和215G2 SHC酶)进行比较时或在与在其各自优化的反应条件(例如优化的pH和温度)(其可能彼此不同)下使用每种酶进行比较时。

[0327] 例如,新的SHC/HAC酶变体可以在反应的前12小时内转化至少约40%EEH或BisEEH,或者该方法可以允许在反应的前12小时内至少约40%EEH或BisEEH转化。例如,新的SHC/HAC酶变体可以在反应的前12小时内转化至少约45%或至少约50%或至少约55%或至少约60%EEH或BisEEH,或者该方法可以允许在反应的前12小时内至少约45%或至少约50%或至少约55%或至少约60%EEH或BisEEH转化。例如,新的SHC/HAC酶变体可以在反应的前6小时内转化至少约30%EEH或BisEEH,或者该方法可以允许在反应的前6小时内至少约30%EEH或BisEEH转化。例如,新的SHC/HAC酶变体可以在反应的前12小时内转化至少约35%或至少约45%或至少约50%或至少约55%EEH或BisEEH,或者该方法可以允许在反应的前12小时内至少约35%或至少约45%或至少约50%或至少约55%EEH或BisEEH转化。这可以是在与在相同的反应条件(例如相同的pH和温度)下使用两种酶(即新的SHC/HAC酶变体和215G2SHC酶)进行比较时或在与在其各自优化的反应条件(例如优化的pH和温度)(其可能彼此不同)下使用每种酶进行比较时。

[0328] EEH至(-)-降龙涎香醚或BisEEH至龙涎香氧化物的转化可以例如使用如上所述的活性测定来确定,并且可以计算为每克原料的可回收产物克数(其可以计算为百分比摩尔

转化率)。

[0329] 如本文所用,本文提及的均法呢醇底物至(-)-降龙涎香醚或双均法呢醇底物至龙涎香氧化物的99%/100%转化是指能够使用SHC/HAC酶或酶变体转化为(-)-降龙涎香醚或龙涎香氧化物的异构体的99%/100%转化。

[0330] SHC/HAC酶变体的最适温度可以例如等于或大于约35℃。例如,SHC/HAC酶变体的最适温度范围可以是约40℃至约50℃,例如约42℃至约48℃或约44℃至约46℃。例如,SHC/HAC酶变体的最适温度可以为约45℃。本文公开的(-)-降龙涎香醚或龙涎香氧化物的制备方法可以在SHC/HAC酶变体的最适温度下进行。

[0331] SHC/HAC酶变体的最适pH可以例如等于或大于约5.4。例如,SHC/HAC酶变体的最适pH可以为约5.2至约6.0,例如约5.4至约5.8,例如约5.6至约5.8。例如,SHC/HAC酶变体的最适pH可以是约5.6或约5.8。可在SHC/HAC酶变体的最适pH下进行本文公开的制备(-)-降龙涎香醚或龙涎香氧化物的方法。

[0332] 本文公开的用于制备(-)-降龙涎香醚或龙涎香氧化物的方法的反应介质中十二烷基硫酸钠(SDS)的最佳浓度在以4g/l与10的OD_{650nm}的细胞一起使用底物(例如EEH或BisEEH)时可以例如为约0.010w/w%至约0.10w/w%。例如,SDS的最佳浓度在以4g/l与10的OD_{650nm}的细胞一起使用底物(例如EEH或BisEEH)时可为约0.040w/w%至约0.080w/w%,例如约0.050w/w%。本文公开的用于制备(-)-降龙涎香醚或龙涎香氧化物的方法的反应介质中十二烷基硫酸钠(SDS)的最佳浓度在以125g/l与250g/l细胞一起使用底物(例如EEH或BisEEH)时可以例如为约1.0w/w%至约1.5w/w%。本文公开的用于制备(-)-降龙涎香醚或龙涎香氧化物的方法的反应介质中十二烷基硫酸钠(SDS)的最佳浓度在以125g/l与125g/l细胞一起使用底物(例如EEH或BisEEH)时可以例如为约0.45w/w%至约0.85w/w%,例如,约0.65w/w%。例如,在以125g/l与250g/l细胞一起使用底物(例如EEH或BisEEH)时,SDS的最佳浓度可以为约1.2w/w%至约1.4w/w%,例如约1.3w/w%。

[0333] 本文公开的制备(-)-降龙涎香醚或龙涎香氧化物的方法可以在使用的特定酶的最佳温度范围或最佳温度和/或最佳pH范围或最佳pH和/或SDS最佳浓度范围或最佳SDS浓度下进行,如下面实施例中的表7或9或11所示。

[0334] 以下编号的段落定义了本公开内容的进一步方面。

[0335] 1.一种制备(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物的方法,所述方法包括使用SHC/HAC酶变体将(3E,7E)-均法呢醇(EEH)或包含EEH的均法呢醇异构体的混合物酶促转化为(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物,

[0336] 其中SHC/HAC酶变体具有与野生型SHC/HAC酶氨基酸序列具有至少约70.0%同一性的氨基酸序列,和

[0337] 其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC酶在选自对应于SEQ ID NO:1的位置81、90、172、277、431、557和613的位置的位置处具有一个或多个氨基酸改变。

[0338] 2.一种制备龙涎香氧化物或包含龙涎香氧化物的混合物的方法,所述方法包括使用SHC/HAC酶变体将E,E-双均法呢醇(BisEEH)或包含BisEEH的双均法呢醇异构体的混合物酶促转化为龙涎香氧化物或包含龙涎香氧化物的混合物,

[0339] 其中SHC/HAC酶变体具有与野生型SHC/HAC酶氨基酸序列具有至少约70.0%同一性的氨基酸序列,和

[0340] 其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC酶在选自对应于SEQ ID NO:1的位置81、90、172、277、431、557和613的位置的位置处具有一个或多个氨基酸改变。

[0341] 3.任何前述段落的方法,其中所述野生型SHC/HAC酶是SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19或SEQ ID NO:20。

[0342] 4.前述段落中任一项的方法,其中SHC/HAC酶变体具有与野生型SHC/HAC酶氨基酸序列具有至少约90.0%或至少约95.0%同一性的氨基酸序列。

[0343] 5.段落1至4中任一项的方法,其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置90和613的位置处具有相对于野生型SHC/HAC酶的氨基酸改变。

[0344] 6.段落1至4中任一项的方法,其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置172和277的位置处具有相对于野生型SHC/HAC酶的氨基酸改变。

[0345] 7.段落1至4中任一项的方法,其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557的位置以及对应于SEQ ID NO:1的位置81、431或613的至少一个位置处具有相对于野生型SHC/HAC酶的氨基酸改变。

[0346] 8.段落7的方法,其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557和431的位置处具有相对于野生型SHC/HAC酶的氨基酸改变。

[0347] 9.段落7的方法,其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557和613的位置处具有相对于野生型SHC/HAC酶的氨基酸改变。

[0348] 10.段落8或9的方法,其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置81的位置处具有相对于野生型SHC/HAC酶的氨基酸改变。

[0349] 11.前述段落中任一项的方法,其中在位置81、90、172、277、431、557和613处的一个或多个例如所有氨基酸改变是取代,例如非保守取代。

[0350] 12.任何前述段落的方法,其中:

[0351] 对应于SEQ ID NO:1的位置81的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为碱性氨基酸,例如组氨酸;和/或

[0352] 对应于SEQ ID NO:1的位置90的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为疏水性氨基酸,例如丙氨酸;和/或

[0353] 对应于SEQ ID NO:1的位置172的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为中性亲水性氨基酸,例如苏氨酸;和/或

[0354] 对应于SEQ ID NO:1的位置277的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为碱性氨基酸,例如赖氨酸;和/或

[0355] 对应于SEQ ID NO:1的位置431的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为疏水性氨基酸,例如亮氨酸;和/或

[0356] 对应于SEQ ID NO:1的位置557的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为中性亲水性氨基酸,例如苏氨酸;和/或

[0357] 对应于SEQ ID NO:1的位置613的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为中性亲水性氨基酸,例如丝氨酸。

[0358] 13.任何前述段落的方法,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC氨基酸序列具有选自以下的一个或多个进一步的氨基酸改变:

[0359] 对应于SEQ ID NO:1的位置132的位置处的氨基酸改变;

[0360] 对应于SEQ ID NO:1的位置224的位置处的氨基酸改变;和

[0361] 对应于SEQ ID NO:1的位置432的位置处的氨基酸改变。

[0362] 14.段落13的方法,其中:

[0363] 对应于SEQ ID NO:1的位置132的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为碱性氨基酸,例如精氨酸;和/或

[0364] 对应于SEQ ID NO:1的位置224的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为疏水性氨基酸,例如缬氨酸;和/或

[0365] 对应于SEQ ID NO:1的位置432的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为中性亲水性氨基酸,例如苏氨酸。

[0366] 15.任何前述段落的方法,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC氨基酸序列具有选自以下的一个或多个进一步的氨基酸改变:

[0367] 对应于SEQ ID NO:1的位置77的位置处的氨基酸改变;和/或

[0368] 对应于SEQ ID NO:1的位置92的位置处的氨基酸改变;和/或

[0369] 对应于SEQ ID NO:1的位置129的位置处的氨基酸改变;和/或

[0370] 对应于SEQ ID NO:1的位置579的位置处的氨基酸改变;和/或

[0371] 对应于SEQ ID NO:1的位置601的位置处的氨基酸改变;和/或

[0372] 对应于SEQ ID NO:1的位置605的位置处的氨基酸改变。

[0373] 16.段落15的方法,其中:

[0374] 对应于SEQ ID NO:1的位置77的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为疏水性氨基酸,例如丙氨酸;和/或

[0375] 对应于SEQ ID NO:1的位置92的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为疏水性氨基酸,例如缬氨酸;和/或

[0376] 对应于SEQ ID NO:1的位置129的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为疏水性氨基酸,例如亮氨酸;和/或

[0377] 对应于SEQ ID NO:1的位置579的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为碱性氨基酸,例如组氨酸;和/或

[0378] 对应于SEQ ID NO:1的位置601的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为芳香族氨基酸,例如酪氨酸;和/或

[0379] 对应于SEQ ID NO:1的位置605的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为芳香族氨基酸,例如色氨酸。

[0380] 17.任何前述段落的方法,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC氨基酸序列具有选自以下的一个或多个进一步的氨基酸改变:

[0381] 对应于SEQ ID NO:1的位置37的位置处的氨基酸改变;和/或

[0382] 对应于SEQ ID NO:1的位置174的位置处的氨基酸改变;和/或

[0383] 对应于SEQ ID NO:1的位置601的位置处的氨基酸改变。

[0384] 18.段落17的方法,其中:

[0385] 对应于SEQ ID NO:1的位置37的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为中性亲水性氨基酸,例如谷氨酰胺;和/或

[0386] 对应于SEQ ID NO:1的位置174的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸

序列的氨基酸取代为疏水性氨基酸,例如异亮氨酸;和/或

[0387] 对应于SEQ ID NO:1的位置601的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为芳香族氨基酸,例如酪氨酸。

[0388] 19.任何前述段落的方法,其中SHC/HAC酶变体具有选自SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:17或SEQ ID NO:18的氨基酸序列。

[0389] 20.任何前述段落的方法,其中所述方法包括培养产生SHC/HAC酶变体的重组宿主细胞。

[0390] 21.段落20的方法,其中所述重组宿主细胞包含编码SHC/HAC酶的核酸序列,例如选自SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:22和SEQ ID NO:23。

[0391] 22.前述段落中任一项的方法,其中(-)-降龙涎香醚与副产物(II)、(III)或(IV)中的至少一种或多种混合产生。

[0392] 23.根据前述权利要求中任一项的制备(-)-降龙涎香醚或包含(-)-降龙涎香醚的反应混合物的方法,所述方法包括使用SHC/HAC酶变体将(3E,7E)-均法呢醇(EEH)或包含EEH的均法呢醇的异构体混合物酶促转化为(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物,

[0393] 其中SHC/HAC酶变体具有与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19或SEQ ID NO:20具有至少约70.0%同一性的氨基酸序列,和

[0394] 其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC酶在选自对应于SEQ ID NO:1的位置81、90、172、277、431、557和613的位置的位置处具有一个或多个氨基酸改变

[0395] 其中包含EEH的异构体混合物选自以下组中的一种或多种:[(3E,7E)和[(3Z,7E)]和/或[(3E,7E)和(3E,7Z)]和/或[(3Z,7E)、(3E,7E)和(3E,7Z)],也分别命名为[EE:EZ]、[EE:ZE]和[EE:EZ:ZE]。

[0396] 24.根据前述权利要求中任一项的制备龙涎香氧化物或包含龙涎香氧化物的混合物的方法,所述方法包括使用SHC/HAC酶变体将(E,E)-双均法呢醇(EEH)或包含BisEEH的双均法呢醇的异构体混合物酶促转化为龙涎香氧化物或包含龙涎香氧化物的混合物,

[0397] 其中SHC/HAC酶变体具有与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19或SEQ ID NO:20具有至少约70.0%同一性的氨基酸序列,和

[0398] 其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC酶在选自对应于SEQ ID NO:1的位置81、90、172、277、431、557和613的位置的位置处具有一个或多个氨基酸改变

[0399] 其中包含BisEEH的异构体混合物选自以下组中的一种或多种:[(E,E)和[(Z,E)]和/或[(E,E)和(E,Z)]和/或[(Z,E)、(E,E)和(E,Z)],也分别命名为[EE:EZ]、[EE:ZE]和[EE:EZ:ZE]。

[0400] 25.一种(-)-降龙涎香醚,其通过任何前述段落的方法获得或可通过任何前述段落的方法获得,例如以无定形或结晶形式获得。

[0401] 26.一种龙涎香氧化物,其通过任何前述段落的方法获得或可通过任何前述段落的方法获得,例如以无定形或结晶形式获得。

[0402] 27.段落25的(-)-降龙涎香醚和/或段落26的龙涎香氧化物作为香料或化妆品或消费品的一部分的用途。

[0403] 28.包含段落25的(-)-降龙涎香醚和/或段落26的龙涎香氧化物的香料或化妆品或消费品。

[0404] 29.一种SHC/HAC酶变体,其具有与野生型SHC/HAC酶氨基酸序列具有至少约70.0%的同一性的氨基酸序列,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC酶在选自对应于SEQ ID NO:1的位置81、90、172、277、431、557和613的位置的位置处具有一个或多个氨基酸改变。

[0405] 30.段落29的SHC/HAC酶变体,其中野生型SHC/HAC酶是SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19或SEQ ID NO:20。

[0406] 31.段落29或30的SHC/HAC酶变体,其中SHC/HAC酶变体具有与野生型SHC/HAC酶氨基酸序列具有至少约90.0%或至少约95.0%同一性的氨基酸序列。

[0407] 32.段落29至31中任一项的SHC/HAC酶变体,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC酶在对应于SEQ ID NO:1的位置90和613的位置处具有氨基酸改变。

[0408] 33.段落29至31中任一项的SHC/HAC酶变体,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC酶在对应于SEQ ID NO:1的位置172和277的位置处具有氨基酸改变。

[0409] 34.段落29至31中任一项的SHC/HAC酶变体,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC酶在对应于SEQ ID NO:1的位置557的位置和对应于SEQ ID NO:1的位置81、431或613的至少一个位置处具有氨基酸改变。

[0410] 35.段落29至31中任一项的SHC/HAC酶变体,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC酶在对应于SEQ ID NO:1的位置557和431的位置处具有氨基酸改变。

[0411] 36.段落29至31中任一项的SHC/HAC酶变体,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC酶在对应于SEQ ID NO:1的位置557和613的位置处具有氨基酸改变。

[0412] 37.段落35或36的SHC/HAC酶变体,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC酶在对应于SEQ ID NO:1的位置81的位置处具有氨基酸改变。

[0413] 38.段落29至37中任一项的SHC/HAC酶变体,其中位置81、90、172、277、431、557或613处的氨基酸改变中的一个或多个例如全部是取代,例如非保守取代。

[0414] 39.段落29至38中任一项的SHC/HAC酶变体,其中:

[0415] 对应于SEQ ID NO:1的位置81的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为碱性氨基酸,例如组氨酸;和/或

[0416] 对应于SEQ ID NO:1的位置90的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为疏水性氨基酸,例如丙氨酸;和/或

[0417] 对应于SEQ ID NO:1的位置172的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为中性亲水性氨基酸,例如苏氨酸;和/或

[0418] 在对应于SEQ ID NO:1的位置277的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为碱性氨基酸,例如赖氨酸;和/或

[0419] 在对应于SEQ ID NO:1的位置431的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为疏水性氨基酸,例如亮氨酸;和/或

[0420] 对应于SEQ ID NO:1的位置557的位置处的氨基酸改变将野生型SHC/HAC氨基酸序

列的氨基酸取代为中性亲水性氨基酸,例如苏氨酸;和/或

[0421] 对应于SEQ ID NO:1的位置613的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为中性亲水性氨基酸,例如丝氨酸。

[0422] 40.段落29至39中任一项的SHC/HAC酶变体,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC氨基酸序列具有选自以下的一个或多个进一步的氨基酸改变:

[0423] 对应于SEQ ID NO:1的位置132的位置处的氨基酸改变;

[0424] 对应于SEQ ID NO:1的位置224的位置处的氨基酸改变;和

[0425] 对应于SEQ ID NO:1的位置432的位置处的氨基酸改变。

[0426] 41.段落40的SHC/HAC酶变体,其中:

[0427] 对应于SEQ ID NO:1的位置132的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为碱性氨基酸,例如精氨酸;和/或

[0428] 对应于SEQ ID NO:1的位置224的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为疏水性氨基酸,例如缬氨酸;和/或

[0429] 对应于SEQ ID NO:1的位置432的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为中性亲水性氨基酸,例如苏氨酸。

[0430] 42.段落29至41中任一项的SHC/HAC酶变体,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC氨基酸序列具有选自以下的一个或多个进一步的氨基酸改变:

[0431] 对应于SEQ ID NO:1的位置77的位置处的氨基酸改变;和/或

[0432] 对应于SEQ ID NO:1的位置92的位置处的氨基酸改变;和/或

[0433] 对应于SEQ ID NO:1的位置129的位置处的氨基酸改变;和/或

[0434] 对应于SEQ ID NO:1的位置579的位置处的氨基酸改变;和/或

[0435] 对应于SEQ ID NO:1的位置601的位置处的氨基酸改变;和/或

[0436] 对应于SEQ ID NO:1的位置605的位置处的氨基酸改变。

[0437] 43.段落42的SHC/HAC酶变体,其中:

[0438] 对应于SEQ ID NO:1的位置77的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为疏水性氨基酸,例如丙氨酸;和/或

[0439] 对应于SEQ ID NO:1的位置92的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为疏水性氨基酸,例如缬氨酸;和/或

[0440] 对应于SEQ ID NO:1的位置129的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为疏水性氨基酸,例如亮氨酸;和/或

[0441] 对应于SEQ ID NO:1的位置579的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为碱性氨基酸,例如组氨酸;和/或

[0442] 对应于SEQ ID NO:1的位置601的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为芳香族氨基酸,例如酪氨酸;和/或

[0443] 对应于SEQ ID NO:1的位置605的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为芳香族氨基酸,例如色氨酸。

[0444] 44.段落29至43中任一项的SHC/HAC酶变体,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC氨基酸序列具有选自以下的一个或多个进一步的氨基酸改变:

[0445] 对应于SEQ ID NO:1的位置37的位置处的氨基酸改变;和/或

[0446] 对应于SEQ ID NO:1的位置174的位置处的氨基酸改变;和/或

[0447] 对应于SEQ ID NO:1的位置601的位置处的氨基酸改变。

[0448] 45.段落44的SHC/HAC酶变体,其中:

[0449] 对应于SEQ ID NO:1的位置37的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为中性亲水性氨基酸,例如谷氨酰胺;和/或

[0450] 对应于SEQ ID NO:1的位置174的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为疏水性氨基酸,例如异亮氨酸;和/或

[0451] 对应于SEQ ID NO:1的位置601的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为芳香族氨基酸,例如酪氨酸。

[0452] 46.段落29至45中任一项的SHC/HAC酶变体,其中SHC/HAC酶变体具有选自SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:17和SEQ ID NO:18的氨基酸序列。

[0453] 47.编码段落29至46中任一项的SHC/HAC酶变体的核酸序列。

[0454] 48.段落47的核酸序列,其中该核酸序列选自SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:22和SEQ ID NO:23。

[0455] 49.包含段落47或48的核酸序列的构建体。

[0456] 50.包含段落49的构建体的载体。

[0457] 51.一种重组宿主细胞,其包含段落47或48的核酸序列、段落49的构建体或段落50的载体。

[0458] 52.段落51的重组宿主细胞,其中所述构建体被整合到所述宿主细胞的基因组中。

[0459] 53.段落51或52的重组宿主细胞,其中所述重组宿主细胞选自原核、酵母、植物和/或昆虫宿主细胞。

[0460] 54.段落51至53中任一项的重组宿主细胞,其中所述重组宿主细胞是具有选自埃希氏菌属(Escherichia)、链霉菌属(Streptomyces)、芽孢杆菌属(Bacillus)、假单胞菌属(Pseudomonas)、乳杆菌属(Lactobacillus)和乳球菌属(Lactococcus)的属的细菌,例如其中所述重组宿主细胞是大肠杆菌(E.coli.)。

[0461] 55.段落1-24中任一项的方法,其中包含EEH的均法呢醇异构体的混合物包括EE:EZ异构体混合物。

[0462] 56.段落55的方法,其中EE:EZ异构体混合物的重量比为:EE:EZ 90:10;EE:EZ 80:20;EE:EZ 86:14;EE:EZ 70:30;EE:EZ 69:31;或EE:EZ 66:34或EE:EZ异构体混合物选自:EE:EZ 90:10;EE:EZ 80:20;EE:EZ 86:14;EE:EZ 70:30;EE:EZ 69:31;和EE:EZ 66:34。

[0463] 57.段落56的方法,其中EE:EZ异构体混合物的重量比为80:20。

[0464] 58.段落1-24中任一项或段落55-57中任一项的方法,其中SHC/HAC生物催化剂与EEH或包含EEH的均法呢醇异构体的混合物(优选80:20重量比的EE:EZ异构体混合物)的重量比在约0.5-2:1或约0.25-2:1或约0.1-2:1或约1:1或约0.5:1的范围内。

[0465] 59.段落58的方法,其中SHC/HAC生物催化剂与EEH或包含EEH的均法呢醇异构体的混合物(优选80:20重量比的EE:EZ异构体混合物)的重量比在约1:1或约0.5:1或约0.1:1的范围内。

[0466] 其他SHC/HAC酶和酶变体

[0467] 进一步令人惊讶地发现,SHC/HAC酶变体也作用于其他底物,例如E,E-双均法呢醇,以制造产品,例如龙涎香氧化物。

[0468] 此外,令人惊讶地发现,与WT AacSHC相比,某些野生型SHC酶提供相对于均法呢醇的其他异构体改善(即更高)的针对EEH的选择性。

[0469] 因此,预期其他野生型SHC/HAC酶和野生型SHC/HAC酶的其他变体也将提供用于将EEH转化为(-)-降龙涎香醚和/或将BisEEH转化为龙涎香氧化物的酶活性(例如改善的酶活性)。

[0470] 因此,本文提供了一种通过将EEH酶促转化为(-)-降龙涎香醚来制备(-)-降龙涎香醚的方法。本文还提供了一种通过将E,E-双均法呢醇酶促转化为龙涎香氧化物来制备龙涎香氧化物的方法。这些方法可以使用本文描述的任何野生型SHC/HAC酶或酶变体。

[0471] 此外,本文提供了与野生型SHC/HAC酶氨基酸序列具有至少约70.0%同一性的SHC/HAC酶或SHC/HAC酶变体。

[0472] 特别地,本文提供了制备(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物的方法,所述方法包括使用与野生型SHC/HAC酶氨基酸序列具有至少约70.0%同一性的SHC/HAC酶或SHC/HAC酶变体将EEH或包含EEH的均法呢醇异构体的混合物酶促转化为(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物。

[0473] 特别地,本文提供了制备龙涎香氧化物或包含龙涎香氧化物的混合物的方法,所述方法包括使用与野生型SHC/HAC酶氨基酸序列具有至少约70.0%同一性的SHC/HAC酶或SHC/HAC酶变体将BisEEH或包含BisEEH的双均法呢醇异构体的混合物酶促转化为龙涎香氧化物或包含龙涎香氧化物的混合物。

[0474] 酶或酶变体的“选择性”是指与另一种底物相比,酶或酶变体与特定底物反应的能力。例如,相对于均法呢醇的其他异构体对EEH具有选择性或相对于双均法呢醇的其他异构体对BisEEH具有选择性的WT SHC酶或WT SHC的酶变体意指WT SHC酶或WT SHC的酶变体相较于均法呢醇的其他异构体更有可能转化EEH或相较于双均法呢醇的其他异构体更有可能转化BisEEH。

[0475] 例如,由于WT SHC酶或WT SHC的酶变体与EEH反应而形成的总产物的wt%可以比由于WT AacSHC与EEH反应而形成的总产物的wt%高至少约1个百分点。例如,由于WT SHC酶或WT SHC的酶变体与EEH的反应而形成的总产物的wt%可以比由于WT AacSHC与EEH反应而形成的总产物的wt%高至少约2或至少约3或至少约4个百分点。例如,由于WT SHC酶或WT SHC的酶变体与EEH的反应而形成的总产物的wt%可以比由于WT AacSHC与EEH的反应而形成的总产物的wt%高多达约40或多达约30或多达约20或多达约15或多达约10个百分点。例如,由于WT SHC酶或WT SHC的酶变体与EEH的反应而形成的总产物的wt%可以比由于WT AacSHC与EEH的反应而形成的总产物的wt%高约1至约40或约2至约30或约3至约20或约4至约10个百分点。由于WT SHC酶或WT SHC的酶变体或WT AacSHC或WT AacSHC的变体的反应而形成的总产物在EEH用作底物时可以例如包含本文描述的式(I)((-)-降龙涎香醚)和式(IV)的化合物、基本上由其组成或由其组成。由于WT SHC酶或WT SHC的酶变体或WT AacSHC或WT AacSHC的变体的反应而形成的总产物在bisEEH用作底物时可以例如包含本文描述的式(X)和/或式(XII)的化合物、基本上由其组成或由其组成。

[0476] 例如,使用WT SHC酶或WT SHC的酶变体形成的(-)-降龙涎香醚的wt%可以例如比

由于WT AacSHC与EEH反应而形成的(-)-降龙涎香醚的wt%高至少约1个百分点。例如,由于WT SHC酶或WT SHC的酶变体与EEH的反应而形成的(-)-降龙涎香醚的wt%可以比由于WT AacSHC与EEH反应而形成的(-)-降龙涎香醚的wt%高至少约2或至少约3或至少约4个百分点。例如,由于WT SHC酶或WT SHC的酶变体与EEH的反应而形成的(-)-降龙涎香醚的wt%可以比由于WT AacSHC与EEH的反应而形成的(-)-降龙涎香醚的wt%高多达约40或多达约30或多达约20或多达约15或多达约10个百分点。例如,由于WT SHC酶或WT SHC的酶变体与EEH的反应而形成的(-)-降龙涎香醚的wt%可以比由于WT AacSHC与EEH的反应而形成的(-)-降龙涎香醚的wt%高约1至约40或约2至约30或约3至约20或约4至约10个百分点。

[0477] WT SHC酶或WT SHC的酶变体的选择性也可以通过比较使用每种酶的反应的EEH:EZH转化率(即EEH的%转化率:EZH的%转化率)或bisEEH:bisEZH转化率(即bisEEH的%转化率:bisEZH的%转化率)来与WT AacSHC或WT AacSHC的变体的选择性进行比较。这可以通过在反应完成时测量剩余在反应混合物中的EEH和EZH或bisEEH和bisEZH的量来确定。可替代地,WT SHC酶或WT SHC的酶变体的选择性也可以通过比较分别由EEH(式I和IV的化合物)和EZH(式II和III的化合物)的转化或分别由bisEEH(式X和XII的化合物)和bisEZH(XI和XIII)的转化产生的产物的比率来与WT AacSHC或WT AacSHC的变体的选择性进行比较。

[0478] SHC酶或WT SHC的酶变体可以例如在由包含EEH和EZH的混合物制备(-)-降龙涎香醚的过程中提供至少约2.0的EEH:EZH转化率。例如,WT SHC酶或WT SHC的酶变体可以在由包含EEH和EZH的混合物制备(-)-降龙涎香醚的过程中提供至少约2.5或至少约3.0或至少约3.5的EEH:EZH转化率。例如,WT SHC酶或WT SHC的酶变体可以在由包含EEH和EZH的混合物制备(-)-降龙涎香醚的过程中提供高达约5.0或高达约4.5或高达约4.0的EEH:EZH转化率。例如,WT SHC酶或WT SHC的酶变体可以在由包含EEH和EZH的混合物制备(-)-降龙涎香醚的过程中提供约2.0至约5.0或约2.5至约4.5或约3.0至约4.0范围内的EEH:EZH转化率。例如,这可能与由AacSHC在由包含EEH和EZH的混合物制备(-)-降龙涎香醚的过程中提供的转化率(其可能例如小于约2.0)形成对比。

[0479] 野生型SHC/HAC酶(例如,可从其衍生SHC/HAC酶变体)可以是例如从酸热脂环酸杆菌(Aac)、运动发酵单胞菌(Zmo)、日本慢生根瘤菌(Bjp)、莫比氏葡萄糖杆菌(Gmo)、双向伯克霍尔德菌、炭疽杆菌、荚膜甲基球菌、法兰克氏菌、巴氏醋杆菌(Apa)、细长嗜热聚球藻菌(Tel)、天蓝色链霉菌(Sco)、沼泽红假单胞菌(Rpa)、特雷迪氏菌(Ttu)、焦炭疽杆菌(Pca)、梨形四膜虫(Tpy)、巨大芽孢杆菌(Bme)或白色链霉菌(Sal)获得的SHC(参见例如WO2010/139719、US2012/01345477、WO2012/066059,其内容通过引用并入本文)。

[0480] 特别地,野生型SHC/HAC酶(例如可从其衍生SHC/HAC酶变体)可以是酸热脂环酸杆菌(Aac)SHC/HAC酶、运动发酵单胞菌(Zmo)SHC/HAC酶、日本慢生根瘤菌(Bjp/Bja)SHC/HAC酶、巴氏醋杆菌(Apa)SHC/HAC酶、巨大芽孢杆菌(Bme)SHC/HAC酶或莫比氏葡萄糖杆菌(Gmo)SHC/HAC酶。特别地,野生型SHC/HAC酶(例如,可以从其衍生SHC/HAC酶变体)可以是酸热脂环酸杆菌(Aac)SHC/HAC酶。

[0481] 当与AacSHC和/或AacSHC的变体相比,WT SHC酶或WT SHC的酶变体相对于均法呢醇的其他异构体具有对EEH的更高选择性时,野生型SHC/HAC酶(例如,可从其衍生SHC/HAC酶变体)不是酸热脂环酸杆菌(Aac)SHC/HAC酶。特别地,当与WT AacSHC和/或WT AacSHC的变体相比,WT SHC酶或WT SHC的酶变体相对于均法呢醇的其他异构体具有对EEH的更高选

择性时，野生型SHC/HAC酶(例如，可从其衍生SHC/HAC酶变体)可选自TelSHC1、ApaSHC1、ZmoSHC1、ZmoSHC2、BjaSHC、GmoSHC、BmeSHC、SalSHC、ApaSHCA。例如，当与WT AacSHC和/或WT AacSHC的变体相比，WT SHC酶或WT SHC的酶变体相对于均法呢醇的其他异构体具有对EEH的更高选择性时，野生型SHC/HAC酶(例如，可从其衍生SHC/HAC酶变体)可以选自ZmoSHC1、BjaSHC、GmoSHC、ApaSHC1和BmeSHC。

[0482] 为了便于参考，名称“AacSHC”可用于指酸热脂环酸杆菌(Aac)SHC/HAC酶，“ZmoSHC”可用于指运动发酵单胞菌(Zmo)SHC/HAC酶，“BjpSHC”或“BjaSHC”可用于指日本慢生根瘤菌(Bjp)SHC/HAC酶，“GmoSHC”可用于指莫比氏葡萄糖杆菌(Gmo)SHC/HAC酶，“BmeSHC”可用于指巨大芽孢杆菌SHC/HAC酶。

[0483] SHC/HAC酶或SHC/HAC酶变体可以例如具有与野生型SHC/HAC酶氨基酸序列具有至少约70.0％同一性的氨基酸序列。例如，SHC/HAC酶或SHC/HAC酶变体可以具有与野生型SHC/HAC酶氨基酸序列具有至少约75.0％或至少约80.0％或至少约85.0％或至少约90.0％或至少约95.0％或至少约95.5％或至少约96.5％或至少约97.0％或至少约97.5％或至少约98.0％或至少约98.5％或至少约99.0％的同一性的氨基酸序列。

[0484] SHC/HAC酶具有与野生型SHC/HAC酶具有100％同一性的氨基酸序列。

[0485] SHC/HAC酶变体具有与野生型SHC/HAC酶氨基酸序列具有小于100％同一性，例如等于或小于约99.5％或等于或小于约99.0％同一性的氨基酸序列。

[0486] 例如，SHC/HAC酶变体与野生型SHC/HAC酶氨基酸序列可以具有约70.0％至约99.5％或约80.0％至约99.0％或约85.0％至约98.5％或约90.0％至约98.0％的同一性。

[0487] 野生型SHC/HAC酶氨基酸序列可以是例如AacSHC(SEQ ID NO:1)、ZmoSHC1(SEQ ID NO:11)、ZmoSHC2(SEQ ID NO:12)、BjpSHC(SEQ ID NO:13)、GmoSHC(SEQ ID NO:14)、TelSHC(SEQ ID NO:19)、ApaSHC1(SEQ ID NO:20)、BmeSHC(SEQ ID NO:28)、SalSHC(SEQ ID NO:29)或ApaSHCA(SEQ ID NO:30)。例如，野生型SHC/HAC酶可以是AacSHC(SEQ ID NO:1)。

[0488] 因此，在某些实施方案中，SHC/HAC酶或SHC/HAC酶变体可以具有与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30具有至少约70.0％同一性的氨基酸序列。例如，SHC/HAC酶或SHC/HAC酶变体具有与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30具有至少约75.0％或至少约80.0％或至少约85.0％或至少约90.0％或至少约95.0％或至少约95.5％或至少约96.5％或至少约97.0％或至少约97.5％或至少约98.0％或至少约98.5％或至少约99.0％同一性的氨基酸序列。

[0489] 例如，SHC/HAC酶可以具有与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30具有100％同一性的氨基酸序列。

[0490] 例如，SHC/HAC酶变体可以例如具有与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30具有小于100％同一性，例如等于或小于约99.5％或等于或小于约99.0％同一性的氨基酸序列。

[0491] 例如，SHC/HAC酶变体可与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID

NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30具有约70.0％至约99.5％或约80.0％至约99.0％或约85.0％至约98.5％或约90.0％至约98.0％的同一性。

[0492] 关于多肽或核苷酸序列的“百分比(％)同一性”分别定义为在比对序列并在必要时引入空位以达到最大序列同一性百分比之后,并且不考虑将任何保守取代作为序列同一性的一部分时,候选序列中与参考序列中的氨基酸或核苷酸相同的氨基酸或核苷酸的百分比。用于确定序列同一性百分比目的的比对可以以本领域技术范围内的各种方式实现,例如,使用可公开获得的计算机软件。本领域技术人员可以确定用于测量比对的适当参数,包括在被比较的序列的全长上实现最大比对所需的任何算法。术语“多肽”和“蛋白质”在本文中可互换使用,并且是指氨基酸的任何肽连接的链,无论长度或翻译后修饰。

[0493] 核苷酸和氨基酸序列的相似性,即序列同一性的百分比,可以通过序列比对来确定。这样的比对可以用几种本领域已知的算法进行,优选使用Karlin和Altschul的数学算法(Karlin&Altschul(1993)Proc.Natl.Acad.Sci.USA 90:5873-5877),使用hmmalign(HMMER软件包,http://hmmer.wustl.edu/),或使用例如可获自https://www.ebi.ac.uk/Tools/msa/clustalo/的CLUSTAL算法(Thompson,J.D.,Higgins,D.G.&Gibson,T.J.(1994)Nucleic Acids Res.22,4673-80),或GAP程序(爱荷华大学的数学算法)或Myers和Miller(1989-Cabios 4:11-17)的数学算法。使用的优选参数是在https://www.ebi.ac.uk/Tools/msa/clustalo/上设置的默认参数。

[0494] 可以使用例如BLAST、BLAT或BlastZ(或BlastX)计算序列同一性百分比。Altschul等人(1990)J.Mol.Biol.215,403-410的BLASTN和BLASTP程序中并入了类似的算法。可以使用BLASTN程序(得分＝100,字长＝12)进行BLAST多核苷酸搜索以获得与编码相关蛋白质的那些核酸同源的多核苷酸序列。可以使用BLASTP程序(得分＝50,字长＝3)进行BLAST蛋白质搜索以获得与多肽同源的氨基酸序列。

[0495] 为了获得用于比较目的的空位比对,如Altschul等人(1997)Nucleic Acids Res.25,3389-3402中所述,使用空位BLAST。当使用BLAST和空位BLAST程序时,使用各自程序的默认参数。序列匹配分析可以通过已建立的同源映射技术如Shuffle-LAGAN(Brudno M.,Bioinformatics 2003b,19Suppl 1:154-162)或Markov随机场来补充。当在本申请中提及序列同一性的百分比时,如果没有特别指出,这些百分比是相对于较长序列的全长计算的。

[0496] 在特定实施方案中,使用CLUSTAL O(版本1.2.4)确定两个序列之间的％同一性。

[0497] 在某些实施方案中,SHC/HAC酶变体与野生型SHC/HAC酶相比,例如与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30相比,可以具有等于或小于约200个氨基酸改变。例如,与野生型SHC/HAC酶相比,例如与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30相比,SHC/HAC酶变体可以具有等于或小于约150或等于或小于约120或等于或小于约100或等于或小于约95或等于或小于约90或等于或小于约85或等于或小于约80或等于或小于约75或等于或小于约70或等于或小于约65或等于或小于约60或等于或小于约55或等于或小于约50或等于或小于约45或等于或小于约40或等于或小于约35或等

于或小于约30或等于或小于约25或等于或小于约20或等于或小于约15或等于或小于约10个氨基酸改变。

[0498] 例如,与野生型SHC/HAC酶相比,例如与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30相比,SHC/HAC酶变体可以具有至少约1个或至少约2个或至少约3个或至少约4个或至少约5个或至少约6个氨基酸改变。

[0499] 例如,与野生型SHC/HAC酶相比,例如与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30相比,SHC/HAC酶变体可以具有约1至约30个氨基酸改变。例如,与野生型SHC/HAC酶相比,例如与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30相比,SHC/HAC酶变体可以具有约2至约25个氨基酸改变。例如,与野生型SHC/HAC酶相比,例如与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30相比,SHC/HAC酶变体可以具有约3至约20个氨基酸改变。例如,与野生型SHC/HAC酶相比,例如与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30相比,SHC/HAC酶变体可以具有约4至约15个氨基酸改变。例如,与野生型SHC/HAC酶相比,例如与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30相比,SHC/HAC酶变体可以具有约5至约10个氨基酸改变。

[0500] 例如,氨基酸改变可以是如上所述的插入、缺失和/或取代。例如,氨基酸改变可以是取代,例如非保守取代。

[0501] 在某些实施方案中,与野生型SHC/HAC酶相比(例如与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30相比),唯一的氨基酸改变是取代(即没有插入或缺失)。

[0502] 氨基酸改变是相对于参考序列定义的。相对于参考序列的氨基酸改变是指变体序列的氨基酸序列与参考序列不同。

[0503] 参考序列和变体序列中的氨基酸可以被分配编号,其中编号从多肽的N末端的氨基酸开始(即多肽N末端的氨基酸被编号为1,下一个氨基酸被编号为2等)。参考序列的“位置”是指参考序列中存在的特定氨基酸残基,如通过参考序列中氨基酸的特定编号所识别的。变体序列的“位置”是指存在于变体序列中的特定氨基酸残基,如通过变体序列中氨基酸的特定编号所识别的。

[0504] 由于与参考序列相比,变体序列可以包括缺失或插入,因此变体序列中的氨基酸的编号可能与参考序列中的相同氨基酸不同。例如,如果在SEQ ID NO:1的氨基酸131和132之间插入氨基酸,则插入的后面的氨基酸将在变体序列中具有编号133,而其在参考序列中保留编号132。在本示例中,与参考序列的位置132对应的变体序列的位置为位置133。因此,变体序列中从参考序列中保留下来的氨基酸可以参照参考序列的“对应位置”来定义。换言之,变体序列中的“位置”可以通过参照参考序列中的“对应位置”来定义。特别地,与参考序

列相比变体序列中的取代可以通过参照参考序列的“对应位置”来定义，尽管参考序列中存在任何插入和/或缺失。在参考序列的氨基酸被缺失的情况下，变体序列中没有“对应位置”。如果与参考序列相比没有插入或缺失(即只有取代)，则参考序列的“对应位置”将与变体序列中的位置相同。

[0505] 来自不同物种的野生型SHC/HAC酶具有不同的多肽长度。可以使用如上所述的算法比对野生型序列，以便鉴定两种不同野生型SHC/HAC酶中的“对应位置”。因此，与参考序列中的位置相对应的变体序列的位置处的氨基酸可以是例如不同的氨基酸残基和/或可以具有与参考序列的编号不同的编号。例如，AacSHC(SEQ ID NO:1)的位置132的氨基酸M可以对应于ZmoSHC1(SEQ ID NO:11)的位置185的氨基酸Y。

[0506] 因此可以相对于两个不同的参考序列来定义氨基酸改变。例如，氨基酸改变可以是与第一参考序列(例如衍生变体的野生型SHC/HAC酶序列)相比的改变，并且可以通过参考第二参考序列(例如AacSHC(SEQ ID NO:1))定义变体序列中氨基酸改变的位置。因此，SHC/HAC酶变体中的氨基酸改变可以相对于第一种野生型SHC/HAC酶位于通过参考第二种野生型SHC/HAC酶定义的位置。

[0507] SHC/HAC酶变体可以例如具有上文参考SEQ ID NO:1定义的一个或多个特定取代或取代的组合。

[0508] 特别地，SHC/HAC酶变体可以在对应于SEQ ID NO:1的位置77、81、90、92、129、132、172、224、277、431、432、557、579、601、605和613的一个或多个位置处具有一个或多个特定取代或取代的组合。

[0509] 例如，这些取代中的每一个可以独立地如上文在标题为“Aac215G2变体”和“对应于SEQ ID NO:1的位置81、90、172、277、431、557和613的位置处具有新突变的其他变体”的小节中所定义。

[0510] 特别地，SHC/HAC酶变体可以具有以下取代组合中的一种或多种：

[0511] (i)对应于SEQ ID NO:1的位置132、224和432的位置处的取代；

[0512] (ii)对应于SEQ ID NO:1的位置132、224、432、557和431的位置处的取代；

[0513] (iii)对应于SEQ ID NO:1的位置132、224、432、557和613的位置处的取代；

[0514] (iv)对应于SEQ ID NO:1的位置132、224、432、557、81和613的位置上的取代；

[0515] (v)对应于SEQ ID NO:1的位置132、224、432、557、81和431的位置处的取代；

[0516] (vi)对应于SEQ ID NO:1的位置132、224、432、90和613的位置处的取代；

[0517] (vii)对应于SEQ ID NO:1的位置132、224、432、172和277的位置处的取代；

[0518] (viii)对应于SEQ ID NO:1的位置132、224、432和37的位置上的取代；

[0519] (ix)对应于SEQ ID NO:1的位置132、224、432和174的位置处的取代；

[0520] (x)对应于SEQ ID NO:1的位置132、224、432、174和601的位置处的取代；

[0521] (xi)对应于SEQ ID NO:1的位置132、224、432、37、174和601的位置处的取代。

[0522] 在某些实施方案中，除了以下氨基酸取代以外，SHC/HAC酶变体与SEQ ID NO:1相同：

[0523] (i)M132R、A224V、I432T、A557T和H431L(SEQ ID NO:2)；或者

[0524] (ii)M132R、A224V、I432T、A557T和R613S(SEQ ID NO:3)；或者

[0525] (iii)M132R、A224V、I432T、A557T、Y81H和R613S(SEQ ID NO:4)；或者

[0526] (iv)M132R、A224V、I432T、A557T、Y81H和H431L(SEQ ID NO:5);或者

[0527] (v)M132R、A224V、I432T、T90A和R613S(SEQ ID NO:17);或者

[0528] (vi)M132R、A224V、I432T、A172T和M277K(SEQ ID NO:18);或者

[0529] (vii)M132R、A224V、I432T和L37Q(SEQ ID NO:24);或者

[0530] (viii)M132R、A224V、I432T、V174I(SEQ ID NO:25);或者

[0531] (ix)M132R、A224V、I432T、V174I和F601Y(SEQ ID NO:26);或者

[0532] (x)M132R、A224V、I432T、L37Q、V174I和F601Y(SEQ ID NO:27)。

[0533] 以下编号的段落定义了本公开内容的其他方面:

[0534] 1.一种制备(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物的方法,所述方法包括使用SHC/HAC酶或SHC/HAC酶变体将(3E,7E)-均法呢醇(EEH)或包含EEH的均法呢醇异构体的混合物酶促转化为(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物,

[0535] 其中SHC/HAC酶或SHC/HAC酶变体具有与野生型SHC/HAC酶氨基酸序列具有至少约70.0%同一性的氨基酸序列,

[0536] 其中WT SHC/HAC酶具有相对于均法呢醇的其他异构体针对EEH的更高的选择性。

[0537] 2.一种制备龙涎香氧化物或包含龙涎香氧化物的混合物的方法,所述方法包括使用SHC/HAC酶或SHC/HAC酶变体将E,E-双均法呢醇(BisEEH)或包含BisEEH的双均法呢醇异构体的混合物酶促转化为龙涎香氧化物或包含龙涎香氧化物的混合物,

[0538] 其中SHC/HAC酶或SHC/HAC酶变体具有与野生型SHC/HAC酶氨基酸序列具有至少约70.0%同一性的氨基酸序列。

[0539] 3.一种制备(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物的方法,所述方法包括使用SHC/HAC酶或SHC/HAC酶变体将3E,7E-均法呢醇(EEH)或包含EEH的均法呢醇异构体的混合物酶促转化为(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物,

[0540] 其中SHC/HAC酶或SHC/HAC酶变体具有与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30具有至少约70.0%同一性的氨基酸序列;和

[0541] 其中包含EEH的异构体混合物选自以下组中的一个或多个:[(3E,7E)和[(3Z,7E)]和/或[(3E,7E)和(3E,7Z)]和/或[(3Z,7E)、(3E,7E)和(3E,7Z)],也分别命名为[EE:EZ]、[EE:ZE]和[EE:EZ:ZE]。

[0542] 4.一种制备龙涎香氧化物或包含龙涎香氧化物的混合物的方法,所述方法包括使用SHC/HAC酶或SHC/HAC酶变体将E,E-双均法呢醇(BisEEH)或包含BisEEH的双均法呢醇异构体的混合物酶促转化为龙涎香氧化物或包含龙涎香氧化物的混合物,

[0543] 其中SHC/HAC酶或SHC/HAC酶变体具有与SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30具有至少约70.0%同一性的氨基酸序列;和

[0544] 其中包含EEH的异构体混合物选自以下组中的一个或多个:[(E,E)和[(Z,E)]和/或[(E,E)和(E,Z)]和/或[(Z,E)、(E,E)和(E,Z)],也分别命名为[EE:EZ]、[EE:ZE]和[EE:EZ:ZE]。

[0545] 5.段落1或3的方法,其中与WT AacSHC相比,SHC/HAC酶具有相对于均法呢醇的其他异构体针对EEH的更高的选择性。

[0546] 6.段落1、3或5中任一项的方法，其中由于SHC/HAC酶或酶变体与EEH反应而形成的总产物的wt％比由于AacSHC与EEH反应而形成的总产物的wt％高至少约1个百分点，例如至少约2个百分点或至少约3个百分点。

[0547] 7.段落1、3、5或6中任一项的方法，其中EEH:EZH转化率为至少约2.0，例如至少约2.5或至少约3.0。

[0548] 8.段落1至7中任一项的方法，其中野生型SHC/HAC酶获自酸热脂环酸杆菌(Aac)、运动发酵单胞菌(Zmo)、日本慢生根瘤菌(Bjp)、莫比氏葡萄糖杆菌(Gmo)、双向伯克霍尔德菌、炭疽杆菌、荚膜甲基球菌、法兰克氏菌、巴氏醋杆菌(Apa)、细长嗜热聚球藻菌(Tel)、天蓝色链霉菌(Sco)、沼泽红假单胞菌(Rpa)、特雷迪氏菌(Ttu)、焦炭疽杆菌(Pca)、梨形四膜虫、巨大芽孢杆菌或白色链霉菌。

[0549] 9.段落1至8中任一项的方法，其中所述野生型SHC/HAC酶是SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:1、SEQ ID NO:14，SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30，例如其中野生型SHC/HAC酶是SEQ ID NO:1。

[0550] 10.段落1的方法，其中均法呢醇异构体的混合物选自以下混合物中的一种或多种:[(3Z,7Z)、(3E,7Z)、(3Z,7E)和(3E,7E)]、[(3Z,7E)、(3E/7E)和(3E,7Z)]、[(3Z,7E)和(3E,7E)]、[(3Z,7E)、(3E,7Z)]和/或[(3E,7E)和(3E,7Z)]。

[0551] 11.任何前述段落的方法，其中该方法使用选自Triton X-100、Tween 80、牛磺脱氧胆酸盐、牛磺脱氧胆酸钠、十二烷基硫酸钠(SDS)和/或月桂基硫酸钠(SLS)的增溶剂。

[0552] 12.段落1至11中任一项的方法，其中SHC/HAC酶变体具有与野生型SHC/HAC酶氨基酸序列具有至少约90.0％或至少约95.0％同一性的氨基酸序列。

[0553] 13.前述段落中任一项的方法，其中SHC/HAC酶变体氨基酸序列在选自对应于SEQ ID NO:1的位置81、90、172、277、431、557和613的位置的一个或多个位置处具有相对于野生型SHC/HAC酶的氨基酸改变。

[0554] 14.段落1至13中任一项的方法，其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置90和613的位置处具有氨基酸改变。

[0555] 15.段落1至13中任一项的方法，其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置172和277的位置处具有氨基酸改变。

[0556] 16.段落1至13中任一项的方法，其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC酶在对应于SEQ ID NO:1的位置557的位置以及对应于SEQ ID NO:1的位置81、431或613的至少一个位置处具有氨基酸改变。

[0557] 17.段落16的方法，其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557和431的位置处具有相对于野生型SHC/HAC酶的氨基酸改变。

[0558] 18.段落16的方法，其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557和613的位置处具有相对于野生型SHC/HAC酶的氨基酸改变。

[0559] 19.段落17或18的方法，其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置81的位置处具有相对于野生型SHC/HAC酶的氨基酸改变。

[0560] 20.前述段落中任一项的方法，其中位置81、90、172、277、431、557或613的氨基酸改变中的一个或多个例如全部是取代，例如非保守取代。

[0561] 21.任何前述段落的方法，其中：

[0562] 对应于SEQ ID NO:1的位置81的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为碱性氨基酸,例如组氨酸;和/或

[0563] 对应于SEQ ID NO:1的位置90的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为疏水性氨基酸,例如丙氨酸;和/或

[0564] 对应于SEQ ID NO:1的位置172的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为中性亲水性氨基酸,例如苏氨酸;和/或

[0565] 在对应于SEQ ID NO:1的位置277的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为碱性氨基酸,例如赖氨酸;和/或

[0566] 在对应于SEQ ID NO:1的位置431的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为疏水性氨基酸,例如亮氨酸;和/或

[0567] 对应于SEQ ID NO:1的位置557的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为中性亲水性氨基酸,例如苏氨酸;和/或

[0568] 对应于SEQ ID NO:1的位置613的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为中性亲水性氨基酸,例如丝氨酸。

[0569] 22.任何前述段落的方法,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC氨基酸序列具有选自以下的一个或多个氨基酸改变:

[0570] 对应于SEQ ID NO:1的位置132的位置处的氨基酸改变;

[0571] 对应于SEQ ID NO:1的位置224的位置处的氨基酸改变;和

[0572] 对应于SEQ ID NO:1的位置432的位置处的氨基酸改变。

[0573] 23.段落22的方法,其中:

[0574] 对应于SEQ ID NO:1的位置132的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为碱性氨基酸,例如精氨酸;和/或

[0575] 对应于SEQ ID NO:1的位置224的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为疏水性氨基酸,例如缬氨酸;和/或

[0576] 对应于SEQ ID NO:1的位置432的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为中性亲水性氨基酸,例如苏氨酸。

[0577] 24.任何前述段落的方法,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC氨基酸序列具有选自以下的一个或多个氨基酸改变:

[0578] 对应于SEQ ID NO:1的位置77的位置处的氨基酸改变;和/或

[0579] 对应于SEQ ID NO:1的位置92的位置处的氨基酸改变;和/或

[0580] 对应于SEQ ID NO:1的位置129的位置处的氨基酸改变;和/或

[0581] 对应于SEQ ID NO:1的位置579的位置处的氨基酸改变;和/或

[0582] 对应于SEQ ID NO:1的位置601的位置处的氨基酸改变;和/或

[0583] 对应于SEQ ID NO:1的位置605的位置处的氨基酸改变。

[0584] 25.段落24的方法,其中:

[0585] 对应于SEQ ID NO:1的位置77的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为疏水性氨基酸,例如丙氨酸;和/或

[0586] 对应于SEQ ID NO:1的位置92的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为疏水性氨基酸,例如缬氨酸;和/或

[0587] 对应于SEQ ID NO:1的位置129的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为疏水性氨基酸,例如亮氨酸;和/或

[0588] 对应于SEQ ID NO:1的位置579的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为碱性氨基酸,例如组氨酸;和/或

[0589] 对应于SEQ ID NO:1的位置601的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为芳香族氨基酸,例如酪氨酸;和/或

[0590] 对应于SEQ ID NO:1的位置605的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为芳香族氨基酸,例如色氨酸。

[0591] 26.任何前述段落的方法,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC氨基酸序列具有选自以下的一个或多个进一步的氨基酸改变:

[0592] 对应于SEQ ID NO:1的位置37的位置处的氨基酸改变;和/或

[0593] 对应于SEQ ID NO:1的位置174的位置处的氨基酸改变;和/或

[0594] 在对应于SEQ ID NO:1的位置601的位置处的氨基酸改变。

[0595] 27.段落26的方法,其中:

[0596] 对应于SEQ ID NO:1的位置37的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为中性亲水性氨基酸,例如谷氨酰胺;和/或

[0597] 对应于SEQ ID NO:1的位置174的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为疏水性氨基酸,例如异亮氨酸;和/或

[0598] 对应于SEQ ID NO:1的位置601的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为芳香族氨基酸,例如酪氨酸。

[0599] 28.任何前述段落的方法,其中SHC/HAC酶变体具有选自SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:17、SEQ ID NO:18、SEQ ID NO:24、SEQ ID NO:25、SEQ ID NO:26和SEQ ID NO:27的氨基酸序列。

[0600] 29.任何前述段落的方法,其中所述方法包括培养产生SHC/HAC酶变体的重组宿主细胞。

[0601] 30.段落29的方法,其中所述重组宿主细胞包含编码SHC/HAC酶的核酸序列,例如选自SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:22或SEQ ID NO:23。

[0602] 31.前述段落中任一项的方法,其中(-)-降龙涎香醚与副产物(II)、(III)或(IV)中的至少一种或多种混合产生。

[0603] 32.一种(-)-降龙涎香醚,其通过任何前述段落的方法获得或可通过任何前述段落的方法获得,例如以无定形或结晶形式获得。

[0604] 33.一种龙涎香氧化物,其通过任何前述段落的方法获得或可通过任何前述段落的方法获得,例如以无定形或结晶形式获得。

[0605] 34.段落32的(-)-降龙涎香醚和/或段落33的龙涎香氧化物作为香料或化妆品或消费品的一部分的用途。

[0606] 35.包含段落32的(-)-降龙涎香醚和/或段落33的龙涎香氧化物的香料或化妆品或消费品。

[0607] 36.一种SHC/HAC酶或SHC/HAC酶变体,其具有与野生型SHC/HAC酶氨基酸序列具有

至少约70.0%同一性的氨基酸序列。

[0608] 37.段落36的SHC/HAC酶或酶变体,其中所述野生型SHC/HAC酶是SEQ ID NO:1、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:28、SEQ ID NO:29或SEQ ID NO:30,例如SEQ ID NO:1。

[0609] 38.段落36或37的SHC/HAC酶或酶变体,其中SHC/HAC酶变体具有与野生型SHC/HAC酶氨基酸序列具有至少约90.0%或至少约95.0%同一性的氨基酸序列。

[0610] 39.段落36至38中任一项的SHC/HAC酶或酶变体,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC酶在选自对应于SEQ ID NO:1的位置81、90、172、277、431、557和613的位置的一个或多个位置处具有氨基酸改变。

[0611] 40.段落36至39中任一项的SHC/HAC酶或酶变体,其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置90和613的位置处具有氨基酸改变。

[0612] 41.段落36至39中任一项的SHC/HAC酶或酶变体,其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置172和277的位置处具有氨基酸改变。

[0613] 42.段落36至39中任一项的SHC/HAC酶或酶变体,其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557的位置和对应于SEQ ID NO:1的位置81、431或613的至少一个位置处具有相对于野生型SHC/HAC酶的氨基酸改变。

[0614] 43.段落36至39中任一项的SHC/HAC酶或酶变体,其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557和431的位置处具有相对于野生型SHC/HAC酶的氨基酸改变。

[0615] 44.段落36至39中任一项的SHC/HAC酶或酶变体,其中SHC/HAC酶变体氨基酸序列在对应于SEQ ID NO:1的位置557和613的位置处具有相对于野生型SHC/HAC酶的氨基酸改变。

[0616] 45.段落43或44的SHC/HAC酶或酶变体,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC酶在对应于SEQ ID NO:1的位置81的位置处具有氨基酸改变。

[0617] 46.段落36至45中任一项的SHC/HAC酶或酶变体,其中在位置81、90、172、277、431、557或613处的一个或多个例如所有的氨基酸改变是取代,例如非保守取代。

[0618] 47.段落36至46中任一项的SHC/HAC酶或酶变体,其中:

[0619] 对应于SEQ ID NO:1的位置81的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为碱性氨基酸,例如组氨酸;和/或

[0620] 对应于SEQ ID NO:1的位置90的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为疏水性氨基酸,例如丙氨酸;和/或

[0621] 对应于SEQ ID NO:1的位置172的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为中性亲水性氨基酸,例如苏氨酸;和/或

[0622] 对应于SEQ ID NO:1的位置277的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为碱性氨基酸,例如赖氨酸;和/或

[0623] 对应于SEQ ID NO:1的位置431的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为疏水性氨基酸,例如亮氨酸;和/或

[0624] 对应于SEQ ID NO:1的位置557的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为中性亲水性氨基酸,例如苏氨酸;和/或

[0625] 对应于SEQ ID NO:1的位置613的位置处的氨基酸改变将野生型SHC/HAC氨基酸序列的氨基酸取代为中性亲水性氨基酸,例如丝氨酸。

[0626] 48.段落36至47中任一项的SHC/HAC酶或酶变体,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC氨基酸序列具有一个或多个氨基酸改变,其选自:

[0627] 对应于SEQ ID NO:1的位置132的位置处的氨基酸改变;

[0628] 对应于SEQ ID NO:1的位置224的位置处的氨基酸改变;和

[0629] 对应于SEQ ID NO:1的位置432的位置处的氨基酸改变。

[0630] 49.段落48的SHC/HAC酶或酶变体,其中:

[0631] 对应于SEQ ID NO:1的位置132的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为碱性氨基酸,例如精氨酸;和/或

[0632] 对应于SEQ ID NO:1的位置224的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为疏水性氨基酸,例如缬氨酸;和/或

[0633] 对应于SEQ ID NO:1的位置432的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为中性亲水性氨基酸,例如苏氨酸。

[0634] 50.段落36至49中任一项的SHC/HAC酶或酶变体,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC氨基酸序列具有一个或多个氨基酸改变,其选自:

[0635] 对应于SEQ ID NO:1的位置77的位置处的氨基酸改变;和/或

[0636] 对应于SEQ ID NO:1的位置92的位置处的氨基酸改变;和/或

[0637] 对应于SEQ ID NO:1的位置129的位置处的氨基酸改变;和/或

[0638] 对应于SEQ ID NO:1的位置579的位置处的氨基酸改变;和/或

[0639] 对应于SEQ ID NO:1的位置601的位置处的氨基酸改变;和/或

[0640] 对应于SEQ ID NO:1的位置605的位置处的氨基酸改变。

[0641] 51.段落50的SHC/HAC酶或酶变体,其中:

[0642] 对应于SEQ ID NO:1的位置77的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为疏水性氨基酸,例如丙氨酸;和/或

[0643] 对应于SEQ ID NO:1的位置92的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为疏水性氨基酸,例如缬氨酸;和/或

[0644] 对应于SEQ ID NO:1的位置129的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为疏水性氨基酸,例如亮氨酸;和/或

[0645] 对应于SEQ ID NO:1的位置579的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为碱性氨基酸,例如组氨酸;和/或

[0646] 对应于SEQ ID NO:1的位置601的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为芳香族氨基酸,例如酪氨酸;和/或

[0647] 对应于SEQ ID NO:1的位置605的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为芳香族氨基酸,例如色氨酸。

[0648] 52.段落36至51中任一项的SHC/HAC酶或酶变体,其中SHC/HAC酶变体氨基酸序列相对于野生型SHC/HAC氨基酸序列具有一个或多个氨基酸改变,其选自:

[0649] 对应于SEQ ID NO:1的位置37的位置处的氨基酸改变;和/或

[0650] 对应于SEQ ID NO:1的位置174的位置处的氨基酸改变;和/或

[0651] 在对应于SEQ ID NO:1的位置601的位置处的氨基酸改变。

[0652] 53.段落52的SHC/HAC酶或酶变体,其中:

[0653] 对应于SEQ ID NO:1的位置37的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为中性亲水性氨基酸,例如谷氨酰胺;和/或

[0654] 对应于SEQ ID NO:1的位置174的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为疏水性氨基酸,例如异亮氨酸;和/或

[0655] 对应于SEQ ID NO:1的位置601的位置处的氨基酸改变将野生型SHC/HAC酶氨基酸序列的氨基酸取代为芳香族氨基酸,例如酪氨酸。

[0656] 54.段落36至53中任一项的SHC/HAC酶或酶变体,其中SHC/HAC酶变体具有选自SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:17、SEQ ID NO:18、SEQ ID NO:24、SEQ ID NO:25、SEQ ID NO:26和SEQ ID NO:27的氨基酸序列。

[0657] 55.编码段落36至54中任一项的SHC/HAC酶或酶变体的核酸序列。

[0658] 56.段落55的核酸序列,其中所述核酸序列选自SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:22或SEQ ID NO:23。

[0659] 57.包含段落55或56的核酸序列的构建体。

[0660] 58.包含段落57的构建体的载体。

[0661] 59.一种重组宿主细胞,其包含段落52或53的核酸序列、段落57的构建体或段落58的载体。

[0662] 60.段落59的重组宿主细胞,其中该构建体被整合到该宿主细胞的基因组中。

[0663] 61.段落59或60的重组宿主细胞,其中所述重组宿主细胞选自原核、酵母、植物和/或昆虫宿主细胞。

[0664] 62.段落59至61中任一项的重组宿主细胞,其中所述重组宿主细胞是具有选自埃希氏菌属、链霉菌属、芽孢杆菌属、假单胞菌属、乳杆菌属和乳球菌属的属的细菌,例如其中所述重组宿主细胞是大肠杆菌。

[0665] 63.段落1-31中任一项的方法,其中包含EEH的均法呢醇异构体的混合物包括EE:EZ异构体混合物。

[0666] 64.段落63的方法,其中EE:EZ异构体混合物选自:EE:EZ 90:10;EE:EZ 80:20;EE:EZ 86:14;EE:EZ 70:30;EE:EZ 69:31;和EE:EZ66:34或其中EE:EZ异构体混合物是EE:EZ 90:10;EE:EZ 80:20;EE:EZ 86:14;EE:EZ 70:30;EE:EZ 69:31;或EE:EZ 66:34。

[0667] 65.段落64的方法,其中EE:EZ异构体混合物的重量比为80:20。

[0668] 66.段落1-31或段落63-65中任一项的方法,其中SHC/HAC酶与EEH或包含EEH的均法呢醇异构体的混合物(优选80:20重量比的EE:EZ异构体混合物)的重量比在约0.5-2:1或约0.25-2:1或约0.1-2:1或约1:1或约0.5:1的范围内。

[0669] 67.段落66的方法,其中SHC/HAC酶与EEH或包含EEH的均法呢醇异构体的混合物(优选80:20重量比的EE:EZ异构体混合物)的重量比在约1:1或约0.5:1或约0.1:1的范围内。

[0670] 以下编号的段落定义了本公开内容的附加方面

[0671] 1.一种制备(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物的方法,其中EE-均法呢醇(EEH)或包含EE-均法呢醇(EEH)的异构体混合物被酶促转化为(-)-降龙涎香醚或包

含(-)-降龙涎香醚的混合物,其中酶促转化使用具有与SEQ ID NO:20(ApaSHC1)、SEQ ID NO:19(TelSHC)、SEQ ID NO:14(GmoSHC)、SEQ ID NO:28(BmeSHC)、SEQ ID NO:29(SalSHC)和/或SEQ ID NO:30(ApaSHCA)具有至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%同一性的多肽序列的角鲨烯藿烯环化酶/均法呢醇降龙涎香醚环化酶(SHC/HAC)生物催化剂在适合于产生(-)-降龙涎香醚的反应条件下进行,其中包含EEH的异构体混合物选自以下组中的一个或多个:[(3E,7E)和[(3Z,7E)]和/或[(3E,7E)和(3E,7Z)]和/或[(3Z,7E)、(3E,7E)和(3E,7Z)],也分别命名为[EE:EZ]和/或[EE:ZE]和/或[EE:EZ:ZE]。

[0672] 2.段落1的方法,其中该方法使用与SEQ ID NO:20(ApaSHC1)或SEQ ID NO:28(BmeSHC)具有至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%同一性的SHC/HAC生物催化剂进行。

[0673] 3.段落2的方法,其中EEH或包含EEH的异构体混合物到(-)-降龙涎香醚的转化在约30℃至约50℃,例如约35℃至约50℃的温度下在约5至约7的pH下进行。

[0674] 4.段落1或段落2的方法,其中该方法在增溶剂例如SDS的存在下进行。

[0675] 5.段落1-4中任一项的方法,其中该方法包括(a)在将EEH或包含EEH的异构体混合物转化至(-)-降龙涎香醚或包含(-)-降龙涎香醚的混合物之前在允许产生WT SHC/HAC生物催化剂的条件下培养表达SHC/HAC酶的一种或多种重组宿主细胞。

[0676] 6.根据段落5的方法,其中培养步骤和随后的转化步骤任选地在不同反应条件下在相同反应容器中进行。

[0677] 7.段落6的方法,其中培养步骤的pH范围为6-7,并且EEH至(-)-降龙涎香醚的步骤的pH范围为4.8-5.5。

[0678] 8.段落1-7中任一项的方法,其中包含EEH的异构体混合物包括EE:EZ异构体混合物。

[0679] 9.段落8的方法,其中EE:EZ异构体混合物的重量比选自:EE:EZ90:10;EE:EZ 80:20;EE:EZ 86:14;EE:EZ 70:30;EE:EZ 69:31;和EE:EZ 66:34或EE:EZ异构体混合物的重量比为:EE:EZ 90:10;EE:EZ 80:20;EE:EZ 86:14;EE:EZ 70:30;EE:EZ 69:31;和/或EE:EZ66:34。

[0680] 10.段落9的方法,其中EE:EZ异构体混合物的重量比为80:20。

[0681] 11.段落1-10中任一项的方法,其中SHC/HAC生物催化剂与EEH或包含EEH的异构体混合物(优选80:20重量比的EE:EZ异构体混合物)的重量比为约0.5-2:1或约0.25-2:1或约0.1-2:1或约1:1或约0.5:1。

[0682] 12.段落11的方法,其中SHC/HAC酶与EEH或包含EEH的异构体混合物(优选80:20重量比的EE:EZ异构体混合物)的重量比为约1:1或约0.5:1或约0.1:1。

[0683] 13.段落1-12中任一项的方法,其中(-)-降龙涎香醚与副产物(II)、(IV)和/或(III)中的一种或多种混合产生。

[0684] 14.段落13的方法,其中从反应介质中分离(-)-降龙涎香醚。

[0685] 15.段落14的方法,其中使用过滤步骤、倾析步骤或过滤和倾析步骤的组合从反应介质中分离(-)-降龙涎香醚。

[0686] 16.段落15的方法,其中过滤步骤是带式过滤或旋转过滤步骤。

[0687] 17.段落13-16中任一项的方法,其中在分离步骤之前,将(-)-降龙涎香醚加热至55℃的最高温度。

[0688] 18.段落17的方法,其中将回收的(-)-降龙涎香醚溶解在溶剂中。

[0689] 19.段落18的方法,其中将溶解的(-)-降龙涎香醚过滤。

[0690] 20.段落18或19的方法,其中通过蒸发除去溶剂来以固体形式回收(-)-降龙涎香醚。

[0691] 21.段落14-20中任一项的方法,其中(-)-降龙涎香醚基本上不含副产物(II)、(IV)和/或(III)。

[0692] 22.一种反应产物,包含可通过段落1-21中任一项的方法获得的(-)-降龙涎香醚。

[0693] 23.段落22的反应产物,其中(-)-降龙涎香醚为固体形式。

[0694] 24.段落23的反应产物,其中(-)-降龙涎香醚为无定形或结晶形式。

[0695] 25.一种制备含有(-)-降龙涎香醚的产品的方法,包括将段落22-24中任一项的反应产物掺入到产品中。

[0696] 26.段落25的方法,其中所述产品是香味产品、化妆品、清洁产品、洗涤剂产品和/或肥皂产品。

[0697] 27.包含段落22-24中任一项的反应产物的香料或化妆品或消费者护理产品。

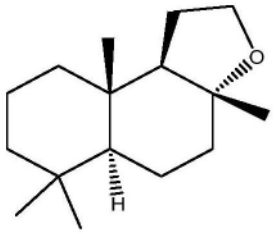
[0698] 28.一种香料或化妆品或消费者护理组合物,其包含段落22-24中任一项的反应产物和附加组分。

[0699] 29.段落22-24中任一项的反应产物作为香料或化妆品消费者护理产品的一部分的用途。

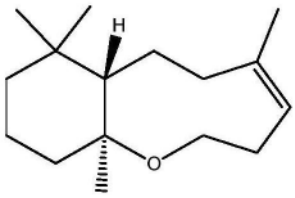
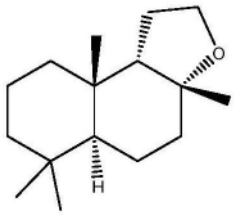
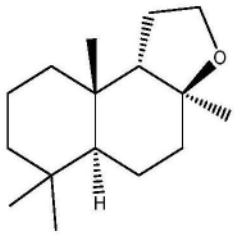
[0700] 降龙涎香醚及其用途

[0701] 本文进一步提供了通过本文所述方法制备的反应产物。例如,反应产物可以包含(-)-降龙涎香醚和一种或多种其他化合物(例如式(II)的化合物、式(III)的化合物和式(IV)的化合物中的一种或多种)、基本上由其组成或由其组成。

[0702] 如本文所用,术语“降龙涎香醚”包括以下式(I)的(-)-降龙涎香醚以及异构体纯形式的或与式(II)、(III)和/或(IV)的以下分子中的一种或多种混合的(-)-降龙涎香醚,

[0703] 

(I)

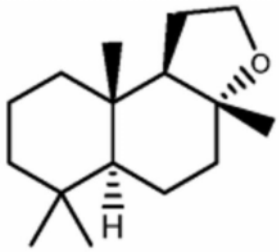
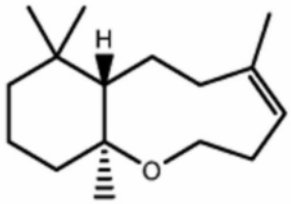
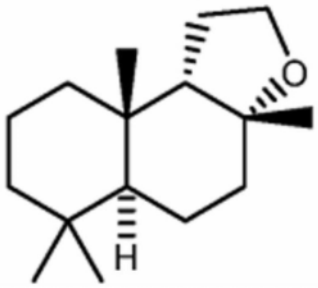
[0704]   

(II) (III) (IV)。

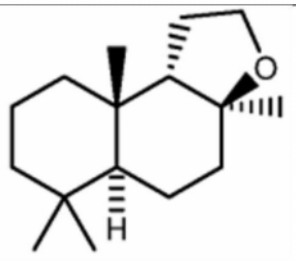
[0705] 式(I)、(II)、(III)和(IV)的反应产物的命名如下所示。

[0706] 表2.式(I)、(II)、(III)和(IV)的反应产物的命名。

[0707]

化合物	描述	名称和结构
(I)	(-)-降龙涎香醚	 (3aR,5aS,9aS,9bR)-3a,6,6,9a-四甲基十二氢萘酚[2,1-b]呋喃
(II)	超分子大环	 (7aS,11aS,Z)-5,8,8,11a-四甲基-2,3,6,7,7a,8,9,10,11,11a-十氢苯并[b]氧杂环辛三烯(oxonine)
(III)	9b-epi-降龙涎香醚	 (3aR,5aS,9aS,9bS)-3a,6,6,9a-四甲基十二氢萘酚[2,1-b]呋喃

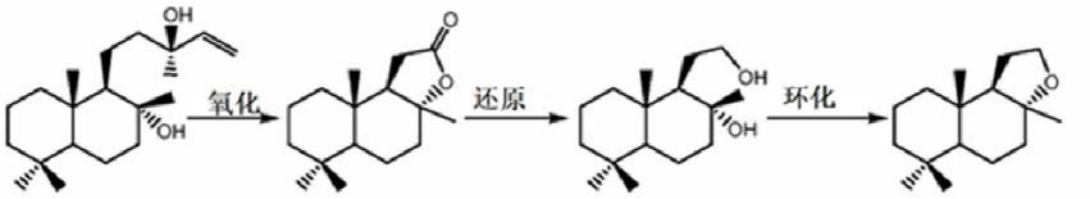
[0708]

(IV)	Escher 等人(1990)	 (3aS,5aS,9aS,9bS)-3a,6,6,9a-四甲基十二氢萘酚[2,1-b]呋喃

[0709] (-)-降龙涎香醚在商业上被称为Ambrox(Firmenich)、Ambroxan(Henkel)、Ambrofix(Givaudan)、Amberlyn(Quest)、Cetalox Laevo(Firmenich)、Ambermor(Aromor)和/或Norambrenolide Ether(Pacific)。

[0710] (-)-降龙涎香醚是一种工业上重要的芳香化合物,并长期以来一直用于香料工业。(-)-降龙涎香醚的特殊期望感官益处来自(-)异构体,而不是(+)异构体。(-)异构体的气味被描述为麝香样、木质、温暖或琥珀香,而(+)-降龙涎香醚对映异构体具有相对较弱的气味。降龙涎香醚类产品的气味和气味阈值也不同。虽然各种富含(-)-降龙涎香醚的材料可商购获得,但希望产生高度浓缩的(-)-降龙涎香醚材料,理想地是纯(-)-降龙涎香醚。本文描述的方法可单独制备或与副产物例如上述式(II)、(III)和/或(IV)所示的化合物混合制备式(I)的(-)-降龙涎香醚。

[0711] (-)-降龙涎香醚可以根据如下的产生工艺从香紫苏内酯(sclareolide)产生。香紫苏醇是从天然植物香紫苏中提取的产品。然而,由于该工艺使用天然起始材料,存在潜在的问题,因为其涉及多级反应,其操作是迂回的,起始材料的供应量和稳定性未必总是令人满意,以及反应可能不是环境友好的,因为在(+)-香紫苏醇氧化降解步骤中使用了氧化剂,例如铬酸或高锰酸盐。

[0712] 

(+)-香紫苏醇 (+)-香紫苏内酯 十氢-2-羟基-2,5,5,8a-十甲基-萘-乙醇 (-)-降龙涎香醚

[0713] (-)-降龙涎香醚也可以使用不同的路线从均法呢醇合成。例如,均法呢醇可以通过溴化、氰化和水解橙花叔醇以产生均法呢醇然后进行还原获得。可替代地,均法呢醇可以从法呢醇、法呢基氯化物、β-法呢烯或其他底物获得。

[0714] 本文描述的方法可单独制备或与副产物例如式(II)、(III)和/或(IV)所示的化合物混合制备式(I)的(-)-降龙涎香醚。例如,式(I)的其他立体异构体可以通过本文描述的方法制备。

[0715] 因此,本文提供了式(I)化合物或包含式(I)化合物的组合物,其通过本文描述的方法获得或可通过本文描述的方法获得,包括其所有实施方案。

[0716] 在某些实施方案中，并非所有的均法呢醇(例如EEH)都转化为(-)-降龙涎香醚或反应的副产物。因此，本文描述的组合物，例如通过本文描述的方法获得的组合物或可通过本文描述的方法获得的组合物可以包含均法呢醇(例如EEH，例如除了式(I)的化合物和/或式(II)、(III)和/或(IV)的化合物以外)。可以将任何剩余的均法呢醇与其他反应产物分离，使得(-)-降龙涎香醚产品不包含均法呢醇。在其他实施方案中，所有均法呢醇起始材料通过本文描述的方法转化为式(I)的(-)-降龙涎香醚或反应的副产物。

[0717] 因此，本文描述的组合物可包含式(I)的化合物、式(II)的化合物、式(III)的化合物、式(IV)的化合物、均法呢醇起始材料(例如EEH)和式(I)的化合物的其他立体异构体中的一种或多种、基本上由其组成或由其组成。例如，本文描述的组合物可包含式(I)的化合物和式(II)的化合物、式(III)的化合物和式(IV)的化合物中的一种或多种、基本上由其组成或由其组成。

[0718] 因此，基于式(I)的化合物、式(II)的化合物、式(III)的化合物和式(IV)的化合物的总重量，本文描述的组合物可包含等于或大于约等于或大于约50wt％的式(I)的化合物。例如，基于式(I)的化合物、式(II)的化合物、式(III)的化合物和式(IV)的化合物的总重量，本文描述的组合物可包含等于或大于约55wt％或等于或大于约60wt％或等于或大于约65wt％或等于或大于约70wt％或等于或大于约75wt％或等于或大于约80wt％或等于或大于约85wt％或等于或大于约90wt％或等于或大于约95wt％的式(I)的化合物。基于式(I)的化合物、式(II)的化合物、式(III)的化合物和式(IV)的化合物的总重量，本文描述的组合物可以例如包含等于或小于约100wt％的式(I)的化合物。例如，基于式(I)的化合物、式(II)的化合物、式(III)的化合物和式(IV)的化合物的总重量，该混合物可以包含等于或小于约99wt％或等于或小于约98wt％或等于或小于约97wt％的式(I)的化合物。例如，基于式(I)的化合物、式(II)的化合物、式(III)的化合物和式(IV)的化合物的总重量，本文描述的组合物可包含约50wt％至约100wt％或约60wt％至约99wt％或约70wt％至约98wt％或约80wt％至约97wt％或约90wt％至约97wt％的式(I)的化合物。

[0719] 在本文描述的组合物中，式(I)的化合物与式(II)的化合物、式(III)的化合物和式(IV)的化合物的总重量的重量比可以例如为大约60:40至大约99:1。例如，式(I)的化合物与式(II)的化合物、式(III)的化合物和式(IV)的化合物的总重量的重量比可以为约65:35至约99:1或约70:30至约99:1或约75:25至约99:1或约80:20至约99:1或约85:15至约99:1或约90:10至约99:1或约95:5至约99:1。例如，式(I)的化合物与式(II)的化合物、式(III)的化合物和式(IV)的化合物的总重量的重量比可以为约65:35至约98:2或约70:30至约97:3或约75:25至约96:4或约80:20至约95:5或约85:15至约90:10。

[0720] 在本文描述的组合物中，式(I)的化合物与均法呢醇(例如EEH)的重量比可以例如为约90:10至约100:0。例如，本文描述的组合物中式(I)的化合物与均法呢醇(例如EEH)的重量比可以为约92:8至约100:0或约94:6至约100:0或约95:5至约100:0或约96:4至约99.5:0.5或约97:3至约99.0:1.0或约98:2至约99.0:1.0。

[0721] 立体异构体的混合物中的式(I)的化合物、式(II)的化合物、式(III)的化合物和式(IV)的化合物的量可以例如通过气相色谱法定量和/或通过NMR光谱鉴定。

[0722] 通过本文描述的方法(例如，使用SHC/HAC酶或其变体和任选地重组宿主细胞)合成的(-)-降龙涎香醚可以是例如无定形形式或结晶形式。

[0723] 通过本文描述的方法(例如使用SHC/HAC酶或其变体和任选地重组宿主细胞)产生的(-)-降龙涎香醚可以通过蒸汽提取/蒸馏或使用非水混溶性溶剂(以分离反应产物和来自留在水相中的生物催化剂的未反应底物)随后蒸发溶剂以获得如通过气相色谱(GC)分析确定的粗反应产物的有机溶剂提取来分离。蒸汽提取/蒸馏和有机溶剂提取方法是本领域技术人员已知的。

[0724] 例如,可以使用有机溶剂例如非水混溶性溶剂(例如甲苯)从整个反应混合物中提取所得的(-)-降龙涎香醚。可替代地,可以使用水混溶性溶剂(例如乙醇)或非水混溶性溶剂(例如甲苯)从反应混合物的固相中提取(通过例如离心或过滤获得)所得的(-)-降龙涎香醚。可替代地,可以使用溶剂混合物从反应混合物的固相中提取所得的(-)-降龙涎香醚。作为进一步的例子,(-)-降龙涎香醚以晶体或无定形形式存在于固相中,并且还可以通过过滤从剩余的固相(细胞材料或其碎片)和液相中分离出来。作为进一步的例子,在高于(-)-降龙涎香醚的熔点(约75℃)的温度下,(-)-降龙涎香醚可以在水相顶部形成油层,该油层可以被移除和收集。为了确保在移除油层后完全回收(-)-降龙涎香醚,可以将有机溶剂添加到含有生物质的水相中以提取包含在生物质中或生物质之上或生物质周围的任何残留的(-)-降龙涎香醚。有机层可以与油层合并,然后整体进行进一步加工以分离和纯化(-)-降龙涎香醚。(-)-降龙涎香醚可以进一步选择性结晶以从最终的(-)-降龙涎香醚产品中去除副产物(II)、(IV)和(III)以及任何未反应的均法呢醇底物。术语“选择性结晶”是指使(-)-降龙涎香醚从溶剂中结晶而化合物(II)、(III)和(IV)保持溶解在结晶溶剂中达到分离的结晶材料仅包含(-)-降龙涎香醚产品或者如果它包含其他化合物(II)、(III)或(IV)中的任何一种则它们仅以嗅觉可接受的量存在的程度的工艺步骤。(-)-降龙涎香醚可以例如不含或基本上不含副产物(II)、(III)和(IV)。选择性结晶步骤可以使用与水混溶的溶剂,例如乙醇等。(-)-降龙涎香醚的选择性结晶可能受未反应的均法呢醇底物的存在以及(-)-降龙涎香醚与其他可检测的副产物(II)、(III)和/或(IV)的比例的影响。即使只获得均法呢醇底物到(-)-降龙涎香醚的10%的转化,(-)-降龙涎香醚的选择性结晶仍然是可能的。

[0725] 最终(-)-降龙涎香醚产品的嗅觉纯度可以使用水中的10%乙醇提取物或通过测试结晶材料来确定。最终(-)-降龙涎香醚产品针对(-)-降龙涎香醚产品的市售参考测试了其嗅觉纯度、质量和其感官特性。(-)-降龙涎香醚材料还由专家在应用研究中进行测试,以确定该材料是否满足关于其感官特性的规格。

[0726] 适用于(-)-降龙涎香醚的提取和/或选择性结晶的合适的水混溶性和非水混溶性有机溶剂的实例包括但不限于脂肪烃,优选具有5至8个碳原子的那些,例如戊烷、环戊烷、己烷、环己烷、庚烷、辛烷或环辛烷,卤代脂肪烃,优选具有一个或两个碳原子的那些,例如二氯甲烷、氯仿、四氯化碳、二氯乙烷或四氯乙烷,芳香烃,例如苯、甲苯、二甲苯、氯苯或二氯苯,脂肪族无环和环状醚或醇,优选具有4至8个碳原子的那些,例如乙醇、异丙醇、二乙醚、甲基叔丁基醚、乙基叔丁基醚、二丙醚、二异丙醚、二丁醚、四氢呋喃或酯例如乙酸乙酯或乙酸正丁酯或酮例如甲基异丁基酮或二氧己环或这些的混合物。特别优选使用的溶剂是上述的庚烷、甲基叔丁基醚(也称为MTBE、叔丁基甲基醚、叔丁基甲基醚和tBME)、二异丙醚、四氢呋喃、乙酸乙酯和/或它们的混合物。优选地,水混溶性溶剂例如乙醇用于从反应混合物的固相中提取(-)-降龙涎香醚。使用乙醇是有利的,因为它易于处理、无毒且对环境友

好。

[0727] 如本文所用,术语“分离的”是指已从伴随它的组分中分离或纯化的生物转化产物,例如(-)-降龙涎香醚。在不同于其天然源自的来源的细胞系统中产生的实体是“分离的”,因为它将必然不含天然伴随它的成分。分离度或纯度可以通过任何合适的方法例如气相色谱(GC)、HPLC或NMR分析来测量。

[0728] 在一些实施方案中,最终产物((-)-降龙涎香醚)被分离和纯化至均质(例如至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%或89.5%纯的或90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或99.5%纯的)。

[0729] 理想地,产生的(-)-降龙涎香醚的量可以为约1mg/l至约20,000mg/l(20g/l)或更高,例如约20g/l至约200g/l或100-200g/l,优选约125g/l或150g/l或约188g/l。

[0730] 使用产生SHC/HAC酶或酶变体的重组大肠杆菌宿主细胞在约2天内在生物转化反应中可以产生至少125g/l(-)-降龙涎香醚。可以在188g/l EEH或更高浓度下进行生物转化,只要实现有效混合,因为搅拌效率可以是系统的唯一限制。此外,具有改善的活性(例如,就具有进一步改进的活性的SHC变体而言或就增加的SHC酶产生而言)的生物催化剂可以使用更少的生物质或增加的底物浓度来改进或维持生产率,这对于混合效率是有利的。

[0731] 例如产生约1至约100mg/l、约30至约100mg/l、约50至约200mg/l、约100至约500mg/l、约100至约1,000mg/l、约250至约5,000mg/l、约1,000(1g/l)至约15,000mg/l(15g/l)或约2,000(2g/l)至约10,000mg/l(10g/l)或约2,000(2g/l)至约25,000mg/l(25g/l)或约2,000(2g/l)至约25,000mg/l(25g/l)、26,000mg/l(26g/l)、27,000mg/l(27g/l)、28,000mg/l(28g/l)、29,000mg/l(29g/l)、30,000mg/l(30g/l)、40g/l、50g/l、60g/l、70g/l、80g/l、90g/l、100g/l、110g/l、120g/l、125g/l、130g/l、140g/l、150g/l、160g/l、170g/l、180g/l、190g/l或200g/l或300g/l或400g/l或500g/l的(-)-降龙涎香醚。

[0732] 优选地,浓度至少为100g/l的(-)-降龙涎香醚在48至72小时的时间段内产生。

[0733] 优选地,浓度为约150g/l的(-)-降龙涎香醚在约48至72小时的时间段内产生。优选地,浓度为约200g/l的(-)-降龙涎香醚在约48至72小时的时间段内产生。

[0734] 优选地,浓度为约250g/l的(-)-降龙涎香醚在约48至72小时的时间段内产生。

[0735] 根据本公开内容的均法呢醇到(-)-降龙涎香醚的生物转化产生(-)-降龙涎香醚作为主要化合物,但也可产生(-)-降龙涎香醚以外的化合物,其可能会或可能不会赋予生物转化混合物令人愉快的嗅觉气味并因此可能对(-)-降龙涎香醚最终产品的感官特性产生积极或消极的影响。因此,使用由受过培训的专家(例如调香师)使用的完善的感官测试进行感官分析,以便测试可以帮助确定化学相关产品相对于参考产品是否也是嗅觉相关的最终产品。从(-)-降龙涎香醚去除一种或多种副产物化合物可以改善包含(-)-降龙涎香醚的反应产物混合物的气味,即使去除的化合物实际上本身是无味的化合物。也就是说,在没有化合物II、III和IV的情况下,可以观察到(-)-降龙涎香醚气味增强。

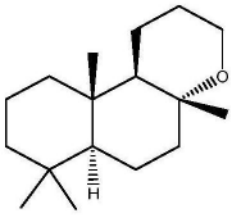
[0736] (-)-降龙涎香醚的各种应用包括但不限于精制香料或消费品,例如织物护理、洗护用品、美容护理和清洁产品、洗涤剂产品和肥皂产品,基本上包括目前可用的降龙涎香醚成分在商业上使用的所有产品,包括但不限于:Ambrox(Firmenich)、Ambroxan(Henkel)、Ambrofix(Givaudan)、Amberlyn(Quest)、Cetalox Laevo(Firmenich)、Ambermor(Aromor)和Norambrenolide Ether(Pacific)产品。

[0737] 因此,本文进一步提供了通过本文描述的方法获得或可通过本文描述的方法获得的(-)-降龙涎香醚作为香料或化妆品或消费品的一部分的用途。本文还提供了一种产品,其包含通过本文描述的方法获得或可通过本文描述的方法获得的(-)-降龙涎香醚。例如,该产品可以是香料或化妆品或消费品。

[0738] 龙涎香氧化物及其用途

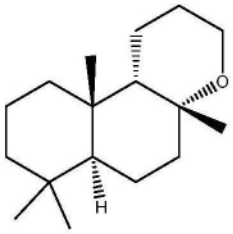
[0739] 本文进一步提供了通过本文描述的方法制备的反应产物。反应产物可以例如包含龙涎香氧化物和一种或多种其他化合物(例如式(XI)的化合物、式(XII)的化合物和式(XIII)的化合物中的一种或多种)、基本上由其组成或由其组成。

[0740] 如本文所用,术语“龙涎香氧化物”包括以下式(X)的龙涎香氧化物以及异构体纯形式的或与式(XI)、(XII)和/或(XIII)的以下分子中的一种或多种混合的龙涎香氧化物,

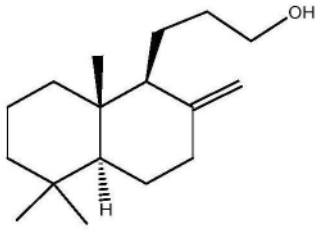


(X)

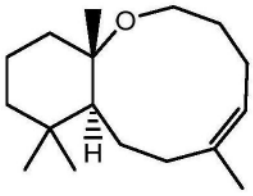
[0741]



(XI)



(XII)

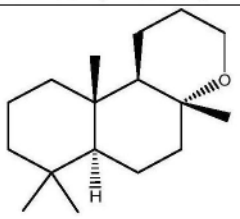
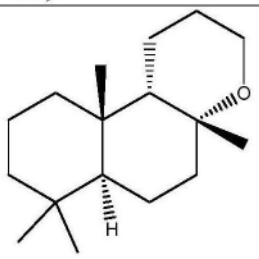


(XIII)

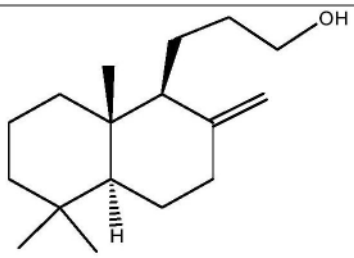
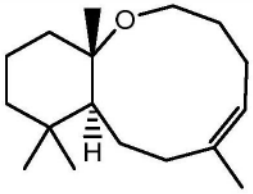
[0742] 式(X)、(XI)、(XII)和(XIII)的反应产物的命名如下所示。

[0743] 表3.式(X)、(XI)、(XII)和(XIII)的反应产物的命名。

[0744]

化合物	名称和结构
(X)	 (4aR,6aS,10aS,10bR)-4a,7,7,10a-四甲基十二氢-1H-苯并[f]色烯](龙涎香氧化物)
(XI)	 [(4aR,6aS,10aS,10bS)-4a,7,7,10a-四甲基十二氢-1H-苯并[f]色烯]

[0745]

(XII)	 [3-((1S,4aS,8aS)-5,5,8a-三甲基-2-亚甲基十二氢萘-1-基)丙烷-1-醇]
(XIII)	 [(8aS,12aS,Z)-6,9,9,12a-四甲基-3,4,7,8,8a,9,10,11,12,12a-十氢-2H-苯并[b]氧杂环癸四烯(oxecine)]

[0746] 如Bolster等人,Tetrahedron,2002,58(26)第5275-5285页中所述,龙涎香氧化物可由(+)-落叶松醇(Larixol)产生。然而,期望提供用于产生龙涎香氧化物的替代或改进方法。

[0747] 本文描述的方法可制备单独的或与副产物例如式(XI)、(XII)和/或(XIII)所示的

化合物混合的式(X)的龙涎香氧化物。例如,式(X)的其他立体异构体可以通过本文描述的方法制备。

[0748] 因此,本文提供了式(X)的化合物或包含式(X)的化合物的组合物,其通过本文描述的方法获得或可通过本文描述的方法获得,包括其所有实施方案。

[0749] 在某些实施方案中,并非所有的双均法呢醇(例如BisEEH)都转化为龙涎香氧化物或反应的副产物。因此,本文描述的组合物,例如通过本文描述的方法获得或可通过本文描述的方法获得的组合物可以包含双均法呢醇(例如BisEEH,例如除了式(X)的化合物和/或式(XI)、(XII)和/或(XIII)的化合物以外)。可以将任何剩余的均法呢醇与其他反应产物分离,使得龙涎香氧化物产物不包含双均法呢醇。在其他实施方案中,通过本文描述的方法将所有的双均法呢醇起始材料转化为式(X)的龙涎香氧化物或反应的副产物。

[0750] 因此,本文描述的组合物可包含式(X)的化合物、式(XI)的化合物、式(XII)的化合物、式(XIII)的化合物、双均法呢醇起始材料(例如bisEEH)和式(X)化合物的其他立体异构体中的一种或多种、基本上由其组成或由其组成。例如,本文描述的组合物可包含式(X)的化合物和式(XI)的化合物、式(XII)的化合物和式(XIII)的化合物中的一种或多种、基本上由其组成或由其组成。

[0751] 因此,基于式(X)的化合物、式(XI)的化合物、式(XII)的化合物和式(XIII)的化合物的总重量,本文描述的组合物可包含等于或大于约等于或大于约50wt%的式(X)的化合物。例如,基于式(X)的化合物、式(XI)的化合物、式(XII)的化合物和式(XIII)的化合物的总重量,本文描述的组合物可包含等于或大于约55wt%或等于或大于约60wt%或等于或大于约65wt%或等于或大于约70wt%或等于或大于约75wt%或等于或大于约80wt%或等于或大于约85wt%或等于或大于约90wt%或等于或大于约95wt%的式(X)化合物。基于式(X)的化合物、式(XI)的化合物、式(XII)的化合物和式(XIII)的化合物的总重量,本文描述的组合物可以例如包含等于或小于约100wt%的式(X)的化合物。例如,基于式(X)的化合物、式(XI)的化合物、式(XII)的化合物和式(XIII)的化合物的总重量,该混合物可以包含等于或小于约99wt%或等于或小于约98wt%或等于或小于约97wt%的式(X)的化合物。例如,基于式(X)的化合物、式(XI)的化合物、式(XII)的化合物和式(XIII)的化合物的总重量,本文描述的组合物可包含约50wt%至约100wt%或约60wt%至约99wt%或约70wt%至约98wt%或约80wt%至约97wt%或约90wt%至约97wt%的式(X)的化合物。

[0752] 在本文描述的组合物中,式(X)的化合物与式(XI)的化合物、式(XII)的化合物和式(XIII)的化合物的总重量的重量比可以例如为大约60:40到大约99:1。例如,式(X)的化合物与式(XI)的化合物、式(XII)的化合物和式(XIII)的化合物的总重量的重量比可以为约65:35至约99:1或约70:30至约99:1或约75:25至约99:1或约80:20至约99:1或约85:15至约99:1或约90:10至约99:1或约95:5至约99:1。例如,式(X)的化合物与式(XI)的化合物、式(XII)的化合物和式(XIII)的化合物的总重量的重量比可以为约65:35至约98:2或约70:30至约97:3或约75:25至约96:4或约80:20至约95:5或约85:15至约90:10。

[0753] 在本文描述的组合物中,式(X)的化合物与双均法呢醇(例如bisEEH)的重量比可以例如为约90:10至约100:0。例如,本文描述的组合物中式(X)的化合物与双均法呢醇(例如bisEEH)的重量比可以为约92:8至约100:0或约94:6至约100:0或约95:5至约100:0或约96:4至约99.5:0.5或约97:3至约99.0:1.0或约98:2至约99.0:1.0。

[0754] 立体异构体混合物中的式(X)的化合物、式(XI)的化合物、式(XII)的化合物和式(XIII)的化合物的量可以例如通过气相色谱法定量和/或通过NMR光谱鉴定。

[0755] 通过本文描述的方法(例如使用SHC/HAC酶或其变体和任选地重组宿主细胞)合成的龙涎香氧化物可以是例如无定形形式或结晶形式。

[0756] 通过本文描述的方法(例如,使用SHC/HAC酶或其变体和重组宿主细胞)产生的龙涎香氧化物可以通过蒸汽提取/蒸馏或使用非水混溶性溶剂(以分离反应产物和来自留在水相中的生物催化剂的未反应底物)随后蒸发溶剂以获得如通过气相色谱(GC)分析确定的粗反应产物的有机溶剂提取来分离。蒸汽提取/蒸馏和有机溶剂提取方法是本领域技术人员已知的。

[0757] 例如,可以使用有机溶剂例如非水混溶性溶剂(例如甲苯)从整个反应混合物中提取所得的龙涎香氧化物。可替代地,可以使用水混溶性溶剂(例如乙醇)或非水混溶性溶剂(例如甲苯)从反应混合物的固相中提取(通过例如离心或过滤获得)所得的龙涎香氧化物。作为进一步的例子,龙涎香氧化物以晶体或无定形形式存在于固相中,并且还可以通过过滤从剩余的固相(细胞材料或其碎片)和液相中分离出来。

[0758] 作为进一步的例子,在高于龙涎香氧化物熔点的温度下,龙涎香氧化物可以在水相顶部形成油层,该油层可以被移除和收集。为了确保在移除油层后完全回收龙涎香氧化物,可以将有机溶剂添加到含有生物质的水相中以提取包含在生物质中或生物质之上或生物质周围的任何残留的龙涎香氧化物。有机层可以与油层合并,然后整体进行进一步加工以分离和纯化龙涎香氧化物。龙涎香氧化物可以进一步选择性结晶以从最终龙涎香氧化物产物中去除副产物(XI)、(XII)和(XIII)以及任何未反应的双均法呢醇底物。术语"选择性结晶"是指使龙涎香氧化物从溶剂中结晶而化合物(XI)、(XII)和(XIII)保持溶解在结晶溶剂中达到分离的结晶材料仅包含龙涎香氧化物产品或者如果它包含其他化合物(X)、(XII)或(XIII)中的任何一种则它们仅以嗅觉可接受的量存在的程度的工艺步骤。例如,龙涎香氧化物可以不含或基本上不含副产物(XI)、(XII)和(XIII)。选择性结晶步骤可以使用与水混溶的溶剂,例如乙醇等。龙涎香氧化物的选择性结晶可能受到未反应的均法呢醇底物的存在以及龙涎香氧化物与其他可检测副产物(XI)、(XII)和/或(XIII)的比例的影响。即使只获得均法呢醇底物到龙涎香氧化物的10%的转化,龙涎香氧化物的选择性结晶仍然是可能的。

[0759] 最终龙涎香氧化物产品的嗅觉纯度可以使用水中的10%乙醇提取物或通过测试结晶材料来确定。最终的龙涎香氧化物产品针对龙涎香氧化物的市售参考测试了其嗅觉纯度、质量和其感官特性。龙涎香氧化物材料还由专家在应用研究中进行测试,以确定该材料是否满足关于其感官特性的规格。

[0760] 适用于龙涎香氧化物的提取和/或选择性结晶的合适的水混溶性和非水混溶性有机溶剂的实例包括但不限于脂肪烃,优选具有5至8个碳原子的那些,例如戊烷、环戊烷、己烷、环己烷、庚烷、辛烷或环辛烷,卤代脂肪烃,优选具有一个或两个碳原子的那些,例如二氯甲烷、氯仿、四氯化碳、二氯乙烷或四氯乙烷,芳香烃,例如苯、甲苯、二甲苯、氯苯或二氯苯,脂肪族无环和环状醚或醇,优选具有4至8个碳原子的那些,例如乙醇、异丙醇、二乙醚、甲基叔丁基醚、乙基叔丁基醚、二丙醚、二异丙醚、二丁醚、四氢呋喃或酯例如乙酸乙酯或乙酸正丁酯或酮例如甲基异丁基酮或二氧己环或这些的混合物。特别优选使用的溶剂是上述

的庚烷、甲基叔丁基醚(也称为MTBE、叔丁基甲基醚、叔丁基甲基醚和tBME)、二异丙醚、四氢呋喃、乙酸乙酯和/或它们的混合物。优选地，水混溶性溶剂例如乙醇用于从反应混合物的固相中提取龙涎香氧化物。使用乙醇是有利的，因为它易于处理、无毒且对环境友好。

[0761] 如本文所用，术语“分离的”是指已从伴随它的组分中分离或纯化的生物转化产物，例如龙涎香氧化物。在不同于其天然源自的来源的细胞系统中产生的实体是“分离的”，因为它将必然不含天然伴随它的成分。分离度或纯度可以通过任何合适的方法例如气相色谱(GC)、HPLC或NMR分析来测量。

[0762] 在一些实施方案中，最终产物(龙涎香氧化物)被分离和纯化至均质(例如至少80％、81％、82％、83％、84％、85％、86％、87％、88％、89％或89.5％纯的或90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或99.5％纯的)。

[0763] 理想地，产生的龙涎香氧化物的量可以为约1mg/l至约20,000mg/l(20g/l)或更高，例如约20g/l至约200g/l或100-200g/l，优选约125g/l或150g/l或约188g/l。

[0764] 例如产生约1至约100mg/l、约30至约100mg/l、约50至约200mg/l、约100至约500mg/l、约100至约1,000mg/l、约250至约5,000mg/l、约1,000(1g/l)至约15,000mg/l(15g/l)或约2,000(2g/l)至约10,000mg/l(10g/l)或约2,000(2g/l)至约25,000mg/l(25g/l)或约2,000(2g/l)至约25,000mg/l(25g/l)、26,000mg/l(26g/l)、27,000mg/l(27g/l)、28,000mg/l(28g/l)、29,000mg/l(29g/l)、30,000mg/l(30g/l)、40g/l、50g/l、60g/l、70g/l、80g/l、90g/l、100g/l、110g/l、120g/l、125g/l、130g/l、140g/l、150g/l、160g/l、170g/l、180g/l、190g/l或200g/l或300g/l或400g/l或500g/l的龙涎香氧化物。

[0765] 优选地，在48至72小时的时间段内产生浓度至少为100g/l的龙涎香氧化物。

[0766] 优选地，浓度为约150g/l的龙涎香氧化物在约48至72小时的时间段内产生。优选地，浓度为约200g/l的龙涎香氧化物在约48至72小时的时间段内产生。

[0767] 优选地，浓度为约250g/l的龙涎香氧化物在约48至72小时的时间段内产生。

[0768] 根据本公开内容的双均法呢醇到龙涎香氧化物的生物转化产生龙涎香氧化物作为主要化合物，但也可能产生除龙涎香氧化物之外的化合物，其可能会或可能不会赋予生物转化混合物令人愉快的嗅觉气味并因此可能对龙涎香氧化物最终产品的感官特性产生积极或消极的影响。因此，使用由受过培训的专家(例如调香师)使用的完善的感官测试进行感官分析，以便测试可以帮助确定化学相关产品相对于参考产品是否也是嗅觉相关的最终产品。从龙涎香氧化物去除一种或多种副产物化合物可以改善剩余的化合物(龙涎香氧化物)的气味，即使去除的化合物实际上本身是无味的化合物。也就是说，在没有化合物XI、XII和XIII的情况下，可以观察到龙涎香氧化物气味增强。

[0769] 龙涎香氧化物的各种应用包括但不限于精制香料或消费品，例如织物护理、洗护用品、美容护理和清洁产品、洗涤剂产品和肥皂产品，基本上包括目前可用的龙涎香氧化物成分在商业上使用的所有产品。

[0770] 因此，本文进一步提供了通过本文描述的方法获得或可通过本文描述的方法获得的龙涎香氧化物作为香料或化妆品或消费品的一部分的用途。本文还提供了包含通过本文描述的方法获得或可通过本文描述的方法获得的龙涎香氧化物的产品。例如，该产品可以是香料或化妆品或消费品。

[0771] 香料组合物

[0772] 本文进一步提供了本文描述的化合物和组合物作为香料组合物或在香料组合物中的用途。

[0773] 因此，本文还提供了包含一种或多种式(I)或(X)的化合物的香料组合物。例如，“香料组合物”可以是包含一种或多种式(I)或(X)的化合物和基础材料的任何组合物。

[0774] 如本文所用，“基础材料”包括选自范围广泛的天然产品和目前可用的合成分子(例如精油、醇、醛和酮、醚和缩醛、酯和内酯、大环化合物和杂环化合物)和/或与通常与香料组合物中的气味剂结合使用的一种或多种成分或赋形剂(例如承载体材料、稀释剂和本领域常用的其他助剂)混合的所有已知香料成分。

[0775] 本领域已知的香料成分可容易地从主要香料制造商处商购获得。此类成分的非限制性例子包括：

[0776] -精油和提取物，例如海狸香、木槿油、橡苔净油、天竺葵油、树苔净油、罗勒油、水果油例如佛手柑油和柑橘油、桃金娘油、棕榈糖油、广藿香油、苦橙叶油、茉莉油、玫瑰油、檀香木油、艾草油、薰衣草油和/或依兰油；

[0777] -醇，例如肉桂醇((E)-3-苯丙基-2-烯-1-醇)；顺式-3-己烯醇((Z)-六-3-烯-1-醇)；香茅醇(3,7-二甲基辛-6-烯-1-醇)；二氢月桂醇(2,6-二甲基辛-7-烯-2-醇)；Ebanol™((E)-3-甲基-5-(2,2,3-三甲基环戊-3-烯-1-基)戊-4-烯-2-醇)；丁香酚(4-烯丙基-2-甲氧基苯酚)；乙基芳樟醇((E)-3,7-二甲基壬-1,6-二烯-3-醇)；法呢醇((2E,6Z)-3,7,11-三甲基十二-2,6,10-三烯-1-醇)；香叶醇((E)-3,7-二甲基辛-2,6-二烯-1-醇)；Super Muguet™((E)-6-乙基-3-甲基辛-6-烯-1-醇)；芳樟醇(3,7-二甲基辛-1,6-二烯-3-醇)；薄荷醇(2-异丙基-5-甲基环己醇)；橙花叔醇(3,7-二甲基-2,6-辛二烯-1-醇)；苯乙醇(2-苯乙醇)；Rhodinol™(3,7-二甲基辛-6-烯-1-醇)；Sandalore™(3-甲基-5-(2,2,3-三甲基环戊-3-烯-1-基)戊-2-醇)；松油醇(2-(4-甲基环己-3-烯-1-基)丙-2-醇)；或Timberol™(1-(2,2,6-三甲基环己)己-3-醇)；2,4,7-三甲基辛-2,6-二烯-1-醇，和/或[1-甲基-2(5-甲基己-4-烯-2-基)环丙基]-甲醇；

[0778] -醛和酮，例如茴香醛(4-甲氧基苯甲醛)；α戊基肉桂醛(2-亚苄基庚醛)；Georgywood™(1-(1,2,8,8-四甲基-1,2,3,4,5,6,7,8-八氢萘酚-2-基)乙烷酮)；羟基香茅醛(7-羟基-3,7-二甲基辛醛)；Iso E**Super®**(1-(2,3,8,8-四甲基-1,2,3,4,5,6,7,8-八氢萘酚-2-基)乙烷酮)；**Isoraldeine®**((E)-3-甲基-4-(2,6,6-三甲基环己-2-烯-1-基)丁-3-烯-2-酮)；3-(4-异丁基-2-甲基苯基)丙醛；麦芽酚；甲基柏木酮；甲基紫罗兰酮；马鞭草酮；和/或香草醛；

[0779] -乙醚和缩醛，例如**降龙涎香醚®**(3a,6,6,9a-四甲基-2,4,5,5a,7,8,9,9b-八氢-1H-苯并[e][1]苯并呋喃)；香叶基甲基醚((2E)-1-甲氧基-3,7-二甲基辛-2,6-二烯)；玫瑰氧化物(4-甲基-2-(2-甲基丙-1-烯-1-基)四氢-2H-吡喃)；和/或**Spirambrene®**(2',2',3,7,7-五甲基螺[双环[4.1.0]庚烷-2,5'-[1,3]二氧己环])；

[0780] -大环化合物，例如Ambretolide((Z)-氧杂环十七烷-10-烯-2-酮)；巴亚基酸次乙酯(1,4-二氧杂环十七烷-5,17-二酮)；和/或**Exaltolide®**(16-氧杂环十六烷-1-酮)；和

[0781] -杂环化合物，例如异丁基喹啉(2-异丁基喹啉)。

[0782] 如本文所用,“承载体材料”是指从气味剂的观点来看实际上是中性的材料,即不会显著改变气味剂的感官特性的材料。

[0783] “稀释剂”是指通常与气味剂结合使用的任何稀释剂,例如邻苯二甲酸二乙酯(DEP)、二丙二醇(DPG)、肉豆蔻酸异丙酯(IPM)、柠檬酸三乙酯(TEC)和醇(例如乙醇)。

[0784] 术语“助剂”是指由于与香料组合物的嗅觉性能没有特别相关的原因而可能用于香料组合物中的成分。例如,助剂可以是辅助用于加工一种或多种香料成分的成分或含有所述一种或多种成分的组合物,或者它可以改善芳香成分或含有其的组合物的处理或储存,例如抗氧化助剂。所述抗氧化剂可以选自例如**Tinogard®** TT(BASF)、**Tinogard®** Q(BASF),生育酚(包括其异构体,CAS 59-02-9;364-49-8;18920-62-2;121854-78-2),2,6-双(1,1-二甲基乙基)-4-甲基苯酚(BHT,CAS 128-37-0)和相关的苯酚、氢醌(CAS 121-31-9)。

[0785] 它也可以是提供额外益处的成分,例如赋予颜色或质地。它也可以是赋予香料组合物中包含的一种或多种成分抗光性或化学稳定性的成分。

[0786] 含有助剂的香料组合物中常用的助剂的性质和类型的详细描述不能是详尽的,但必须提及的是,所述成分是本领域技术人员熟知的。

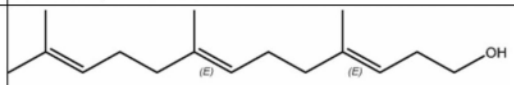
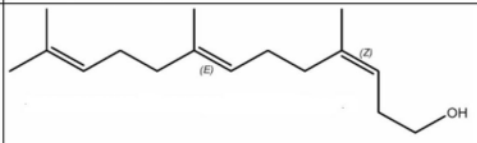
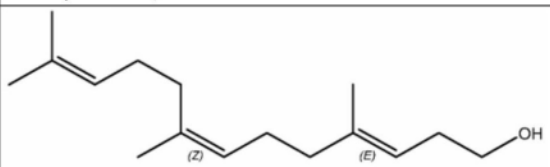
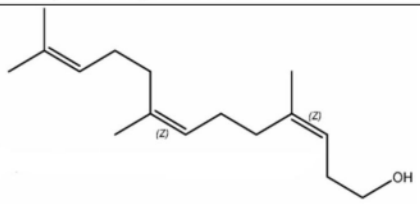
[0787] 本文还提供了包含如本文所述的化合物或组合物或香料组合物的消费品,包括其任何实施方案。例如,消费品可以是化妆品(例如淡香水或淡香精)、清洁产品、去污剂产品或肥皂产品。

[0788] 均法呢醇

[0789] 均法呢醇可能具有如下所示的异构性。

[0790] 表4.均法呢醇异构体。

[0791]

化合物	缩写	名称和结构
E,E-均法呢醇	EEH	 (*3E,7E*)-4,8,12-三甲基十三烷-3,7,11-三烯-1-醇
E,Z-均法呢醇	EZH	 (*3Z,7E*)-4,8,12-三甲基十三烷-3,7,11-三烯-1-醇
Z,E-均法呢醇	ZEH	 (*3E,7Z*)-4,8,12-三甲基十三烷-3,7,11-三烯-1-醇
Z,Z-均法呢醇	ZZH	 (*3Z,7Z*)-4,8,12-三甲基十三烷-3,7,11-三烯-1-醇

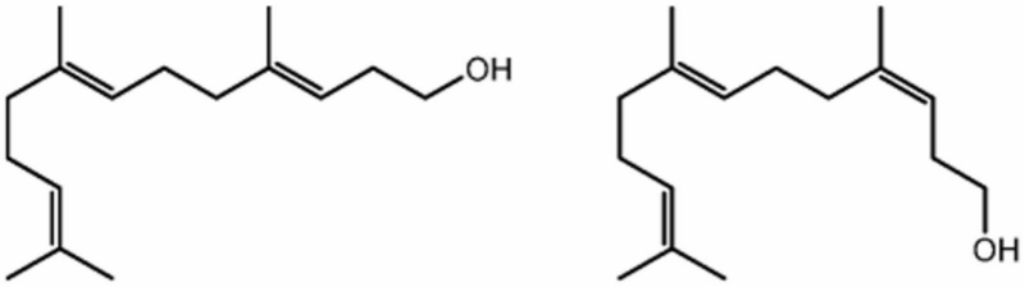
[0792] β-法呢烯可以直接转化为E,E-均法呢醇(EEH)或通过E,E-均法呢酯间接转化为EEH,然后再转化为EEH。关于从不同底物产生(-)-降龙涎香醚的综述可见于US2012/0135477A1、WO2010/139719、US2013.0273619A1、WO2013/156398A1和Seitz博士论文(2012,如上文引用的)和Schaefer 2011(Chemie Unserer Zeit 45,374-388),其内容通过引用并入本文。

[0793] US2012/0135477A1报道了使用ZmoSHC将(3Z,7E)转化为(-)-降龙涎香醚,但根据Schaefer(2011)(如上所述)中的公开内容,(7E,3Z)仅转化为9b-epi-降龙涎香醚(即上文概述的化合物(III))而不是(-)-降龙涎香醚。如本文所用,提及(3Z,7E)均法呢醇是指E,Z-均法呢醇,也称为EZH。

[0794] 虽然均法呢醇可以是四种异构体(即(3Z,7Z)、(3E,7Z)、(3Z,7E)和(3E,7E)异构体)的混合物,但根据文献似乎只能从(3E,7E)均法呢醇获得(-)-降龙涎香醚(参见上面引用的Neumann和Simon(1986))。如本文所用,提及(3E,7E)-均法呢醇是指E,E-均法呢醇,其也称为EEH。

[0795] 用于本文描述的制备(-)-降龙涎香醚的方法的起始材料可以是例如(3E,7E)-均法呢醇或包含(3E,7E)-均法呢醇的混合物,例如包含(3E,7E)-均法呢醇的均法呢醇的异构体混合物。

[0796] 优选地,均法呢醇起始材料包含(3E,7E)和(3Z,7E)的混合物,本文称为EE:EZ异构体混合物。均法呢醇的EE:EZ异构混合物的CAS编号为35826-67-6。

[0797] 

[0798] (3E,7E)-均法呢醇 (3Z,7E)-均法呢醇

[0799] CAS:459-89-2 CAS:138152-06-4

[0800] E/Z-混合物CAS:35826-67-6

[0801] 均法呢醇原液/起始材料可以是异构体的混合物。因此,均法呢醇起始材料还可以包含四种异构体的混合物EE:EZ:ZZ:ZE,其对应于(3E,7E)、(3Z,7E)、(3Z,7Z)和(3E,7Z)。优选地,均法呢醇起始材料选自以下混合物中的一种或多种:[(3Z,7Z)、(3E,7Z)、(3Z,7E)和(3E,7E)]、[(3Z,7E)和(3E,7E)]、[(3Z,7E)、(3E,7Z)]和/或[(3E,7E)和(3E,7Z)]。

[0802] 优选地,均法呢醇起始材料选自以下混合物中的一种或多种:[(3E,7E)、(3Z,7E)]和/或[(3Z,7E)、(3E/7E)和(3E,7Z)],也分别命名为[EE:EZ]和[EE:EZ:ZE]。

[0803] 因此,在某些实施方案中,EEH:EZH的比率为约100:00;99:01;98:02;97:03;96:04;95:05;94:06;93:07;92:08;91:09;90:10;89:11;88:12;87:13;86:14;85:15;84:16;83:17;82:18;81:19;80:20;79:21;78:22;77:23;76:24;75:25;74:26;73:27;72:28;71:29;70:30;69:31;68:32;67:33;66:34;65:35;64:36;63:37;62:38;61:39;60:40;59:41;58:42;57:43;56:44;55:45;54:46;53:47;52:48;51:49;或约50:50。例如,EEH:EZH的比例可以为约50:50至约100:00或约50:50至约99:01或约60:40至约99:1或约70:30至约95:5或约80:20至约95:5。

[0804] 在一些实施方案中,优选地,均法呢醇起始材料包含>90%的E,E-均法呢醇(EEH)。

[0805] 在其他实施方案中,均法呢醇起始材料包含86:14的EE:EZ重量比。

[0806] 在某些实施方案中,均法呢醇起始材料包含80:20的EE:EZ重量比。

[0807] 在某些实施方案中,均法呢醇起始材料包含70:30的EE:EZ重量比。

[0808] 在进一步的实施方案中,均法呢醇起始材料包含69:31的EE:EZ重量比。

[0809] 存在的均法呢醇异构体的数量可能会影响反应速度。SHC/HAC酶或酶变体可能能够从均法呢醇异构体的复杂混合物(例如EE:EZ:ZE:ZZ)将E,E-均法呢醇转化为(-)-降龙涎香醚。然而,可能会观察到较低的转化率,这与EEH以外的均法呢醇异构体可能与EEH竞争接近SHC/HAC衍生酶并因此可能充当EEH转化为(-)-降龙涎香醚的竞争抑制剂和/或还充当替代底物的观点一致(参见例如,Eichhorn等人(2018)Adv.Synth.Catal.360:2339-2351,其内容通过引用并入本文)。因此,均法呢醇底物可以包含2-4种异构体优选两种异构体的异构体混合物。

[0810] 因此，均法呢醇底物可以由2-4种异构体优选两种异构体的异构体混合物组成或基本上由其组成。

[0811] 优选地，均法呢醇底物包含EE:EZ异构混合物。

[0812] 优选地，均法呢醇底物由EE:EZ异构混合物组成或基本上由其组成。

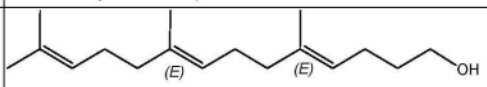
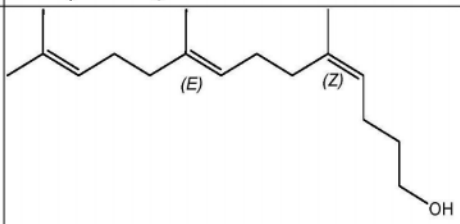
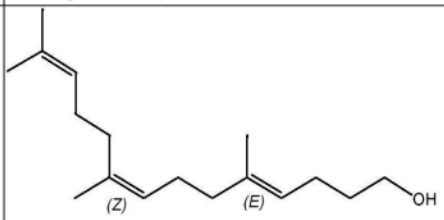
[0813] 例如，如果使用EE:EZ异构体混合物，则(-)-降龙涎香醚以外的化合物(例如表2中列出的化合物II、III和IV)为“油状”形式(而不是固体的形式)，其促进反应混合物的搅拌和由此产生的生物转化过程。

[0814] 双均法呢醇

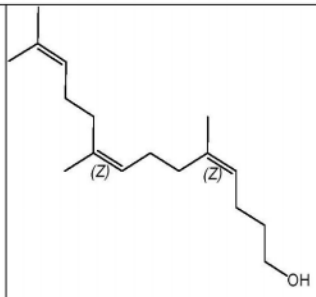
[0815] 双均法呢醇可能具有如下表5所示的异构性。

[0816] 表5.双均法呢醇异构体。

[0817]

化合物	缩写	名称和结构
E,E-双均法呢醇	BisEEH	 (4E,8E)-5,9,13-三甲基十四烷-4,8,12-三烯-1-醇
E,Z-双均法呢醇	BisEZH	 (4Z,8E)-5,9,13-三甲基十四烷-4,8,12-三烯-1-醇
Z,E-双均法呢醇	BisZEH	 (4E,8Z)-5,9,13-三甲基十四烷-4,8,12-三烯-1-醇

[0818]

Z,Z-双均法呢醇	BisZZH	 (4Z,8Z)-5,9,13-三甲基十四烷-4,8,12-三烯-1-醇

[0819] 双均法呢醇可以由E-橙花叔醇产生，如下面的实施例中所述。例如，可将双均法呢醇产生为两种或更多种异构体的混合物(例如，E,E-双均法呢醇和E,Z-双均法呢醇的混合物)。

[0820] 虽然双均法呢醇可能以四种异构体((Z,Z)、(E,Z)、(Z,E)和(E,E)异构体)的混合物形式存在，但似乎龙涎香氧化物只能从E,E-双均法呢醇获得。

[0821] 用于本文描述的用于制备龙涎香氧化物的方法的起始材料可以是例如E,E-双均法呢醇或包含E,E-双均法呢醇的混合物，例如包含E,E-双均法呢醇的双均法呢醇异构体的混合物。

[0822] 优选地，双均法呢醇起始材料包含(BisEEH)和(BisEZH)的混合物，本文称为EE:EZ异构体混合物。

[0823] 双均法呢醇原液/起始材料可以是异构体的混合物。因此，双均法呢醇起始材料还可以包含四种异构体的混合物EE:EZ:ZZ:ZE。

[0824] 因此，在某些实施方案中，BisEEH:BisEZH的比率为约100:00;99:01;98:02;97:03;96:04;95:05;94:06;93:07;92:08;91:09;90:10;89:11;88:12;87:13;86:14;85:15;84:16;83:17;82:18;81:19;80:20;79:21;78:22;77:23;76:24;75:25;74:26;73:27;72:28;71:29;70:30;69:31;68:32;67:33;66:34;65:35;64:36;63:37;62:38;61:39;60:40;59:41;58:42;57:43;56:44;55:45:54:46;53:47:52:48;51:49;或大约50:50。例如，BisEEH:BisEZH的比例可以为约50:50至约100:00或约50:50至约99:01或约60:40至约99:1或约70:30至约95:5或约80:20至约95:5。

[0825] 在一些实施方案中，优选地，双均法呢醇起始材料包含>90％的E,E-双均法呢醇(EEH)。

[0826] 在其他实施方案中，双均法呢醇起始材料包含86:14的BisEEH:BisEZH重量比。在某些实施方案中，双均法呢醇起始材料包含80:20的BisEEH:BisEZH重量比。

[0827] 在某些实施方案中，双均法呢醇起始材料包含70:30的BisEEH:BisEZH重量比。

[0828] 在进一步的实施方案中，双均法呢醇起始材料包含69:31的BisEEH:BisEZH重量比。

[0829] 存在的双均法呢醇异构体的数量可能会影响反应速度。SHC/HAC酶或酶变体可能能够从双均法呢醇异构体的复杂混合物(例如EE:EZ:ZE:ZZ)将E,E-双均法呢醇转化为龙涎香氧化物。然而，可能会观察到较低的转化率，这与BisEEH以外的双均法呢醇异构体可能与BisEEH竞争接近SHC/HAC酶或酶变体并因此可能充当BisEEH转化为龙涎香氧化物的竞争抑

制剂和/或还充当替代底物的观点一致。因此，双均法呢醇底物可以包含2-4种异构体优选两种异构体的异构体混合物。

[0830] 因此，双均法呢醇底物可以由2-4种异构体优选两种异构体的异构体混合物组成或基本上由其组成。

[0831] 优选地，双均法呢醇底物包含EE:EZ异构体混合物。

[0832] 优选地，双均法呢醇底物由EE:EZ异构体混合物组成或基本上由其组成。

[0833] 核酸和制备核酸的方法

[0834] 本文进一步提供了编码如本文所述的SHC/HAC酶或SHC/HAC酶变体的核酸。例如，核酸可以是分离的核酸。

[0835] 特别地，本文提供了包含编码如本文所述的SHC/HAC酶或酶变体的核酸序列的构建体。如本文所用，“构建体”是待转染到靶细胞中的核酸的人工产生的区段。构建体可以包含编码SHC/HAC酶或酶变体的核酸和表达控制器(例如启动子)。

[0836] 本文进一步提供了包含本文所述构建体的载体。如本文所用，“载体”是用作媒介物以将外源遗传物质人工携带到其可以在其中复制和/或表达的细胞中的DNA分子。例如，载体可以是质粒、病毒载体、粘粒或人工染色体。

[0837] 术语“构建体”和“载体”可以重叠，例如在构建体是质粒的情况下。

[0838] 特别地，本文提供了编码SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:17、SEQ ID NO:18、SEQ ID NO:24、SEQ ID NO:25、SEQ ID NO:26和SEQ ID NO:27中任一个的氨基酸序列的核酸。

[0839] 具体而言，本文提供了具有SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:22和SEQ ID NO:23的序列的核酸，其例如可以包含在如本文所述的构建体或载体中。

[0840] 如本文所用，术语“核酸”或“核酸分子”应具体指本公开内容的多核苷酸，其可以是DNA、cDNA、基因组DNA、合成DNA或RNA，并且可以是双链或单链的，有义链和/或反义链。术语“核酸”或“核酸分子”应特别适用于如本文所用的一种或多种多核苷酸，例如作为全长核苷酸序列或其片段或部分，其分别编码具有酶活性(例如代谢途径的酶)的多肽或其片段或部分。

[0841] 该术语还包括单独的分子，例如cDNA，其中相应的基因组DNA具有内含子，因此具有不同的序列；缺少至少一个侧翼基因的基因组片段；由聚合酶链式反应(PCR)产生并且缺少至少一个侧翼基因的cDNA或基因组DNA片段；缺少至少一个侧翼基因的限制性片段；编码非天然存在的蛋白质(例如融合蛋白(例如His标签)、突变蛋白或给定蛋白质的片段)的DNA；以及作为天然存在的核酸或cDNA的简并变体的核酸。此外，它包括作为杂合基因的一部分的重组核苷酸序列，即编码非天然存在的融合蛋白的基因。融合蛋白可以将一个或多个氨基酸(例如但不限于组氨酸(His))添加到蛋白质中，通常在蛋白质的N末端但也可以在C末端或融合在蛋白质的区域内。此类融合蛋白或编码此类蛋白的融合载体通常用于三个目的：(i)增加重组蛋白的产量；(ii)增加重组蛋白的溶解度；和(iii)通过提供用于亲和纯化的配体来帮助纯化重组蛋白。

[0842] 术语“核酸”或“核酸分子”还包括适合在特定微生物宿主细胞(例如大肠杆菌宿主细胞)中表达的密码子优化的序列。如本文所用，术语“密码子优化的”是指核酸蛋白质编码

序列,其已通过用在细菌(例如大肠杆菌)宿主细胞基因中更频繁使用的密码子取代一个或多个或优选相当数量的多个密码子来适应于在原核或真核宿主细胞(特别是细菌宿主细胞例如大肠杆菌宿主细胞)中表达。

[0843] 在这点上,编码参考氨基酸序列(SEQ ID NO:1或SEQ ID NO:10)及其变体/衍生物的核苷酸序列可以是在来源中发现的原始序列(例如在AacSHC中发现的SEQ ID NO:1)或基因可以针对选择的宿主生物例如大肠杆菌进行密码子优化。

[0844] 核糖核酸(RNA)分子可以通过体外转录产生。DNA分子的区段也被认为在本公开内容的范围内,并且可以通过例如聚合酶链式反应(PCR)产生或通过用一种或多种限制性内切核酸酶处理产生。核酸分子的片区段可被称为基因的DNA片段,特别是部分基因的片段。片段还可以包含几个开放阅读框(ORF),其是相同ORF的重复或者是不同的ORF。该术语应具体指编码核苷酸序列,但也应包括非编码核苷酸序列,例如全部或部分未转录或未翻译的序列,或编码多肽。如本文所用的基因(例如用于组装、多样化或重组)可以是非编码序列或编码多肽的序列或编码蛋白质的序列或其部分或片段,其具有用于成功的重组事件的足够的序列长度。更具体地,所述基因具有3bp的最小长度,优选至少100bp,更优选至少300bp。从前述内容将明显的是,提及分离的DNA并不意指存在于例如cDNA或基因组DNA文库或例如限制性消化反应混合物或电泳凝胶切片中的基因组DNA限制性消化物中的数百至数百万个其他DNA分子中的DNA。本公开内容的分离的核酸分子包括在天然状态中未发现的区段。

[0845] 如本文所用,术语“分离的DNA”可以指(1)含有与任何天然存在的序列不同的序列的DNA、非天然存在的多核苷酸或核酸(例如通过人工组合(例如,通过例如基因工程技术对分离的核酸区段进行人工操作)产生的)或(2)在具有天然存在的序列的DNA(例如cDNA或基因组DNA)的情况下,不含在含有感兴趣的DNA的基因天然存在的生物体的基因组中在含有感兴趣的DNA的基因的侧翼的至少一个基因的DNA。

[0846] 如本文所用的术语“分离的DNA”(特别是关于核酸序列)还可以指通过重组DNA技术产生的核酸或多核苷酸,例如包含与宿主细胞异源的多核苷酸的DNA构建体,其任选地掺入宿主细胞中。嵌合核苷酸序列可以特异性地作为重组分子产生。术语“重组”应特别适用于组装多核苷酸,将此类多核苷酸或其部分连接在一起,进行或不进行重组以实现交叉互换或基因镶嵌。例如,将所需功能的核酸区段连接在一起以产生所需的功能组合。编码本文所述多肽的重组基因可以包括该多肽的编码序列,以有义方向可操作地连接到一个或多个适合于表达该多肽的调节区。因为许多微生物能够从多顺反子mRNA表达多个基因产物,如果需要,可以在这些微生物的单个调节区的控制下表达多个多肽。当调节区和编码序列定位成使得调节区对调节序列的转录或翻译有效时,编码序列和调节区被认为是可操作连接的。

[0847] 如本文所用的术语“重组”(特别是关于酶)应指通过重组DNA技术产生的酶,即由编码所需酶的外源DNA构建体转化的细胞产生。“合成”酶是通过化学合成制备的那些。嵌合酶可以特异性地作为重组分子产生。因此,术语“重组DNA”包括被插入载体中、整合至自主复制的质粒或病毒中,或整合至原核生物或真核生物的基因组DNA(或同源细胞的基因组,在除天然染色体外的位置处)中的重组DNA。

[0848] 在另一方面,本公开内容的一种或多种核酸分子与允许在原核和/或真核宿主细胞中表达的表达控制序列可操作地连接。如本文中所用,“可操作地连接”是指整合至遗传

构建体中，使得表达控制序列有效地控制目标编码序列的表达。上述转录/翻译调控元件包括但不限于诱导型和非诱导型、组成型、受细胞周期调节的、代谢调节的启动子、增强子、操纵子、沉默子、阻遏物和本领域技术人员已知的并且驱动或以其它方式调控基因表达的其它元件。此类调控元件包括但不限于指导组成型表达或允许诱导型表达等的调控元件，例如CUP-1启动子、所采用的tet-阻遏物(例如在tet-on或tet-off系统中使用的)、lac系统、trp系统调控元件。作为实例，异丙基β-D-1-硫代吡喃半乳糖苷(IPTG)在100μM至1.0mM浓度范围内是基因表达的有效诱导剂。该化合物是异乳糖(allolactose，一种触发lac操纵子转录的乳糖代谢产物)的分子模拟物，因此其用于在基因处于lac操纵子的控制之下时诱导基因表达。诱导基因表达的调节元件的另一个例子是乳糖。类似地，本公开内容的一种或多种核酸分子可形成编码另外的多肽序列(例如，用作标志物或报告物的序列)的杂合基因的一部分。标志物和报告基因的实例包括β-内酰胺酶、氯霉素乙酰转移酶(CAT)、腺苷脱氨酶(ADA)、氨基糖苷磷酸转移酶、二氢叶酸还原酶(DHFR)、潮霉素-B-磷酸转移酶(HPH)、胸苷激酶(TK)、lacZ(编码β-半乳糖苷酶)和黄嘌呤鸟嘌呤磷酸核糖基转移酶(XGPRT)。与本公开实践相关的许多标准程序一样，本领域技术人员将了解另外的有用试剂，例如可充当标志物或报告物的功能的其它序列。

[0849] 在一些实施方案中，本公开内容提供了编码SHC/HAC酶或其变体的重组多核苷酸，其可被插入用于表达和任选纯化的载体中。一种类型的载体是代表连接另外DNA区段的环状双链DNA环的质粒。某些载体可控制与它们功能连接的基因的表达。这些载体称为“表达载体”。通常适于DNA重组技术的表达载体是质粒类型。通常，表达载体包含基因，例如本文所述的SHC/HAC酶或其变体。在本说明书中，术语“质粒”和“载体”可互换使用，因为质粒是最常用的载体。

[0850] 此类载体可包括DNA序列，其包括但不限于并不天然存在于宿主细胞中的DNA序列，通常不被转录成RNA或翻译成蛋白质(“表达”)的DNA序列以及期望引入至非重组宿主的其它基因或DNA序列。应当理解，通常通过稳定引入一个或多个重组基因来增强本文的重组宿主的基因组。然而，也可在本公开的范围内使用自主或复制型质粒或载体。此外，本公开还可使用低拷贝数(例如单拷贝)或高拷贝数(如本文所例示的)质粒或载体来实施。

[0851] 在一个优选实施方案中，本公开内容的载体包含质粒、噬菌粒、噬菌体、粘粒、人造细菌和人造酵母染色体、敲除或敲入构建体、合成核酸序列或盒，并且可以以线性多核苷酸、质粒、巨大质粒、合成或人造染色体(例如植物、细菌、哺乳动物或酵母人造染色体)的形式产生亚组。

[0852] 优选地，在引入载体后，由引入的多核苷酸编码的蛋白质在细胞内表达。可将多种基因底物掺入质粒。质粒通常是标准克隆载体，例如细菌多拷贝质粒。可将底物掺入相同或不同的质粒中。通常使用具有不同类型的选择标记的至少两种不同类型的质粒来允许选择含有至少两种类型载体的细胞。

[0853] 通常，细菌或酵母细胞可用本领域公知的任一种或多种核苷酸序列来进行转化。对于体内重组，要与基因组或其它基因重组的基因用于使用标准转化技术转化宿主。在一个合适的实施方案中，提供复制起点的DNA包括在构建体中。复制起点可以由本领域技术人员来适当选择。根据基因的性质，如果基因或基因组中已经存在本身可作为复制起点起作用的序列，则可能不需要补充的复制起点。

[0854] 宿主细胞、制备宿主细胞的方法以及使用宿主细胞制备降龙涎香醚和龙涎香氧化物的方法

[0855] 本文进一步提供了包含如本文所述的核酸序列或构建体或载体的重组宿主细胞。本文进一步提供了产生如本文所述的SHC/HAC酶或酶变体的重组宿主细胞。

[0856] 本文的用于产生(-)-降龙涎香醚或龙涎香氧化物的方法可以例如包括培养如本文所述的重组宿主细胞。如本文所用,术语“培养”是指产生活细胞的过程,使得它们产生如本文所述的SHC/HAC酶或酶变体,其可用于如本文所述的用于产生(-)-降龙涎香醚或龙涎香氧化物的方法中。细胞没有必要自身分裂和复制,尽管不排除这种情况。

[0857] 当此类DNA已经被引入细胞内部时,细菌或酵母细胞可被外源或异源DNA转化。转化的DNA可以整合或不整合,即共价连接到细胞的基因组中。例如,在原核生物和酵母中,可在附加型元件例如质粒上保持转化DNA。对于真核细胞,稳定转染的细胞是其中转染的DNA已经整合至染色体中以使其通过染色体复制由子细胞遗传的细胞。这种稳定性通过真核细胞建立由含有转化DNA的子细胞的群体组成的细胞系或克隆的能力来证明。

[0858] 一般来说,引入的DNA最初不存在于作为DNA的受体的宿主中,但从给定的宿主分离DNA区段、随后将该DNA的一个或多个额外拷贝引入同一宿主以例如增加基因产物的产生或改变基因的表达模式也在本公开内容的范围之内。在一些情况下,引入的DNA将通过例如同源重组或定点诱变来修饰或甚至替代内源基因或DNA序列。合适的重组宿主包括微生物、植物细胞和植物。

[0859] 本公开内容还表征重组宿主。术语“重组宿主”(也称为“遗传修饰的宿主细胞”或“转基因细胞”)表示包含异源核酸或其基因组已由至少一个整合的DNA序列增强的宿主细胞。本公开内容的宿主细胞可被如上的多核苷酸或载体进行遗传工程改造。

[0860] 可用于本公开内容目的的宿主细胞包括但不限于原核细胞例如细菌(例如,大肠杆菌和枯草芽孢杆菌(B.subtilis)),其可以例如用例如含有本公开内容的多核苷酸分子的重组噬菌体DNA、质粒DNA、细菌人造染色体或粘粒DNA表达载体转化;简单的真核细胞例如酵母(例如,酵母属和毕赤酵母属(Pichia)),其可以例如用例如含有本公开内容的多核苷酸分子的重组酵母表达载体转化。取决于宿主细胞和用于引入本公开内容的多核苷酸的相应载体,多核苷酸可以整合至例如染色体或线粒体DNA中,或者可以染色体外(例如通过附加体)维持,或者可以仅短暂地包含在细胞中。

[0861] 如本文所用,术语“细胞”,特别是提及遗传工程和将一个或多个基因或组装的基因簇引入细胞时,或产生细胞,被理解为指任何原核或真核细胞。预期原核和真核宿主细胞均根据本公开内容来使用,包括细菌宿主细胞例如大肠杆菌或芽孢杆菌属的种、酵母宿主细胞例如酿酒酵母(S.cerevisiae)、昆虫宿主细胞例如草地夜蛾或人宿主细胞,例如HeLa和Jurkat细胞。

[0862] 具体地,细胞是真核细胞,优选真菌、哺乳动物或植物细胞或原核细胞。合适的真核细胞包括例如但不限于哺乳动物细胞、酵母细胞或昆虫细胞(包括Sf9)、两栖动物细胞(包括载黑素细胞)或包括新杆状线虫(包括秀丽隐杆线虫(Caenorhabditis elegans))的细胞的蠕虫细胞。合适的哺乳动物细胞包括例如但不限于COS细胞(包括Cos-1和Cos-7)、CHO细胞、HEK293细胞、HEK293T细胞、HEK293T-RexTM细胞或其它可转染的真核细胞系。合适的细菌细胞包括但不限于大肠杆菌。

[0863] 优选地，可以使用原核生物，例如大肠杆菌、芽孢杆菌属、链霉菌属，或者哺乳动物细胞，例如HeLa细胞或Jurkat细胞，或者植物细胞，如拟南芥。

[0864] 例如，细胞可以选自原核、酵母、植物和/或昆虫宿主细胞。

[0865] 优选地，细胞是曲霉菌属的种(Aspergillus sp)或真菌细胞，优选地，其可选自酵母属、假丝酵母属(Candida)、克鲁维酵母属(Kluyveromyces)、汉逊酵母属、裂殖酵母属、耶氏酵母属(Yarrowia)、毕赤酵母属和曲霉菌属。

[0866] 优选地，细胞是细菌细胞，例如，具有选自埃希氏菌属、链霉菌属、芽孢杆菌属、假单胞菌属、乳杆菌属和乳球菌属的属。例如，细菌可以是大肠杆菌。

[0867] 优选地，大肠杆菌宿主细胞是被工业和监管机构认可的大肠杆菌宿主细胞(包括但不限于大肠杆菌K12宿主细胞或大肠杆菌BL21宿主细胞)。

[0868] 与本公开内容一起使用的一种优选宿主细胞是大肠杆菌，其可以如本文所述重组制备。因此，重组宿主可以是重组大肠杆菌宿主细胞。存在可用于大肠杆菌的突变体文库、质粒、详细的代谢计算机模型和其它信息，从而允许合理设计各种模块来增加产物产率。可使用与上文针对酵母菌属所描述的方法类似的方法来制备重组大肠杆菌微生物。

[0869] 在一个实施方案中，重组大肠杆菌微生物包含编码SHC/HAC酶或酶变体基因的核苷酸序列。

[0870] 优选地，重组大肠杆菌微生物包含如本文所述的载体构建体。在另一个优选实施方案中，重组大肠杆菌微生物包含编码本文公开的SHC/HAC酶和酶变体的核苷酸序列。

[0871] 与本公开内容一起使用的另一种优选宿主细胞是酿酒酵母，其是合成生物学中广泛使用的基础生物。因此，重组宿主可以是酿酒酵母。存在可用于酿酒酵母的突变体文库、质粒、详细的代谢计算机模型和其它信息，从而允许合理设计各种模块来增加产物产率。用于制备重组酿酒酵母微生物的方法是已知的。

[0872] 可以以常规方式进行细胞的培养。培养基可含有碳源、至少一种氮源和无机盐，并向其中添加维生素。该培养基的成分可以是通常用于培养目标微生物物种的成分。本方法中使用的碳源包括可由重组宿主细胞代谢以促进生长和/或产生(-)-降龙涎香醚或龙涎香氧化物的任何分子。合适的碳源的实例包括但不限于蔗糖(例如，如在糖蜜中发现的)、果糖、木糖、甘油、葡萄糖、纤维素、淀粉、纤维二糖或其它含葡萄糖的聚合物。

[0873] 在使用酵母作为宿主的实施方案中，例如碳源例如蔗糖、果糖、木糖、乙醇、甘油和葡萄糖是合适的。可在整个培养期间给宿主生物提供碳源，可替代地，可使生物体在另一种能量来源例如蛋白质存在的情况下生长一段时间，然后仅在补料分批阶段提供碳源。

[0874] 用于本公开内容的方法的重组宿主细胞微生物的适用性可使用熟知方法通过简单测试程序来确定。例如，可将待测试的微生物在通常用于增殖微生物的pH、温度和反应条件下在丰富培养基(例如，LB-培养基、Bacto-胰蛋白胨酵母提取物培养基、营养培养基等)中进行增殖。一旦选择了产生所需生物转化产物的重组微生物(即重组宿主细胞)，则产物通常通过合适的表达系统和发酵(例如通过细胞培养中的微生物产生)由产生宿主细胞系大规模产生。在本公开内容的一个实施方案中，将成分明确的基本培养基例如M9A用于细胞培养。

[0875] M9A培养基的组分包括：14g/l KH_2PO_4、16g/l K_2HPO_4、1g/l $Na_3Citrate.2H_2O$、7.5g/l $(NH_4)_2SO_4$、0.25g/l $MgSO_4.7H_2O$、0.015g/l $CaCl_2.2H_2O$、5g/l葡萄糖和1.25g/l酵母

提取物)。

[0876] 在本公开内容的另一个实施方案中,使用富含营养物的培养基例如LB。LB培养基的组分包括:10g/l胰蛋白胨,5g/l酵母提取物,5g/l NaCl。矿物培养基和M9矿物培养基的其它实例公开于例如US6524831B2和US 2003/0092143A1中。

[0877] 基本培养基的另一个例子可以如下制备:对于350ml培养物:向35ml柠檬酸/磷酸盐原液(133g/l KH_2PO_4、40g/l $(NH_4)_2HPO_4$、17g/l柠檬酸.H_2O,pH调节至6.3)加入307ml H_2O,根据需要用32%NaOH调节pH至6.8。高压灭菌后,加入0.850ml 50%$MgSO_4$、0.035ml微量元素溶液(见下文)、0.035ml硫胺素溶液和7ml 20%葡萄糖。

[0878] 微量元素溶液:在去离子水中的50g/l $Na_2EDTA.2H_2O$、20g/l $FeSO_4.7H_2O$、3g/l H_3BO_3、0.9g/l $MnSO_4.2H_2O$、1.1g/l $CoCl_2$、80g/L $CuCl_2$、240g/l $NiSO_4.7H_2O$、100g/l KI、1.4g/l $(NH_4)_6Mo_7O_{24}.4H_2O$、1g/l $ZnSO_4.7H_2O$,

[0879] 硫胺素溶液:在去离子水中的2.25g/l硫胺素.HCl

[0880] $MgSO_4$溶液:在去离子水中的50%(w/v)$MgSO_4.7H_2O$

[0881] 重组微生物可以在分批、补料分批或连续过程或其组合中生长。典型地,重组微生物在发酵罐中在确定的温度下在合适的营养源例如碳源的存在下生长所需时间段内以产生足够的酶以将均法呢醇转化为降龙涎香醚或将双均法呢醇转化为龙涎香氧化物并产生所需量的降龙涎香醚包括(-)-降龙涎香醚或所需量的式(X)的龙涎香氧化物。可以以任何合适的方式培养重组宿主细胞,例如通过分批培养或补料分批培养。

[0882] 如本文所用,术语“分批培养”是一种培养方法,其中在培养过程中既不添加也不取出培养基和/或营养物。

[0883] 如本文所用,术语“补料分批”是指一种培养方法,其中在培养过程中添加培养基和/或营养物但不取出培养基。

[0884] 本公开内容的一个实施方案提供一种在细胞系统中产生降龙涎香醚或龙涎香氧化物的方法,其包括在细胞系统中在合适的条件下产生SHC/HAC酶或酶变体,向细胞系统中添加均法呢醇或双均法呢醇,使用利用细胞系统产生的SHC/HAC酶或酶变体将均法呢醇转化为降龙涎香醚或将双均法呢醇转化为龙涎香氧化物,从细胞系统收集降龙涎香醚或龙涎香氧化物,并任选地从系统中分离(-)-降龙涎香醚材料或式(X)的龙涎香氧化物。其他核苷酸序列的表达可用于增强该方法。生物转化方法可以包括在细胞系统中额外表达其他核苷酸序列。其他核苷酸序列的表达可以增强用于制备(-)-降龙涎香醚或龙涎香氧化物的生物转化途径。本公开内容的另一个实施方案是制备(-)-降龙涎香醚或龙涎香氧化物的生物转化方法,包括培养包含SHC/HAC酶或酶变体基因的宿主细胞,在宿主细胞中产生SHC/HAC酶或酶变体,进料均法呢醇(例如EEH)或双均法呢醇(例如bisEEH)到宿主细胞中,在适合促进均法呢醇转化为降龙涎香醚或双均法呢醇转化为龙涎香氧化物的pH、温度和增溶剂条件下孵育宿主细胞并收集(-)-降龙涎香醚或式(X)的龙涎香氧化物。SHC/HAC酶或酶变体在宿主细胞中的产生提供了当在合适的反应条件下将均法呢醇或双均法呢醇添加到宿主细胞中时制备(-)-降龙涎香醚或式(X)的龙涎香氧化物的方法。通过向反应混合物中添加更多的生物催化剂和SDS可以提高实现的转化率。

[0885] 可以以多种方式培养重组宿主细胞微生物,以便为随后的生物转化步骤提供合适量的产生SHC/HAC酶或酶变体的细胞。由于适用于生物转化步骤的微生物差异很大(例如酵

母、细菌和真菌)，因此培养条件当然要根据每个物种的具体要求进行调整，并且这些条件是众所周知的并且有记录的。用于培养重组宿主细胞微生物的细胞的任何本领域已知方法可用于产生可用于本公开内容的后续生物转化步骤的细胞。通常，细胞生长至特定密度(可测量为光密度(OD))以产生足够的生物质用于生物转化反应。

[0886] 选择的培养条件不仅影响获得的细胞的量(生物质)，而且培养条件的质量也影响生物质如何成为生物催化剂。表达SHC/HAC酶或酶变体基因并产生SHC/HAC酶或酶变体的重组宿主细胞微生物被称为适用于生物转化反应的生物催化剂。在一些实施方案中，生物催化剂是产生SHC/HAC酶或酶变体的重组全细胞，或者它可以是悬浮形式或固定形式。在其他实施方案中，生物催化剂是从产生SHC/HAC酶或酶变体的重组全细胞制备的膜级分或液体级分(如例如在如上文所引用的Seitz等人2012中所公开的)。产生SHC/HAC酶或酶变体的重组生物催化剂包括从发酵罐收集的全细胞(用于生物转化反应)或发酵罐中的细胞(然后将其用于一锅反应)。产生SHC/HAC酶或酶变体的重组生物催化剂可以包括完整的重组全细胞和/或细胞碎片。无论哪种方式，SHC/HAC酶或酶变体以某种方式与膜(例如细胞膜)缔合，以便接收底物(例如均法呢醇或双均法呢醇)和/或与底物(例如均法呢醇或双均法呢醇)相互作用，该膜(例如细胞膜)可以是全细胞(例如重组全细胞)的一部分或包含全细胞。SHC/HAC酶或酶变体也可以是允许SHC/HAC酶或酶变体与底物(例如均法呢醇或双均法呢醇)相互作用的固定形式(例如与酶承载体缔合)。SHC/HAC酶或酶变体也可以以可溶形式使用。

[0887] 在一个实施方案中，在开始生物转化步骤之前，以足够的量产生、收获和洗涤(并且任选地储存(例如冷藏、冷冻或冻干))生物催化剂(以产生足够的生物质)。

[0888] 在进一步的实施方案中，以足够的量产生细胞(以产生足够的生物催化剂)，然后调节反应条件，而不需要收获和洗涤用于生物转化反应的生物催化剂。这个一步(或“一锅”)方法是有利的，因为它简化了过程，同时可能降低了成本。用于培养细胞的培养基也适用于生物转化反应，条件是反应条件被调节以促进生物转化反应。

[0889] 用于生长细胞的最佳pH值在6.0-7.0范围内。用于生物转化反应的最佳pH值取决于生物转化反应中使用的SHC/HAC酶或酶变体的类型。使用技术人员熟知的技术调节pH。

[0890] 虽然术语“混合物”或“反应混合物”可以与本公开内容中的术语“培养基”互换使用(尤其是当它涉及“一锅”反应时)，但应注意培养细胞以产生足够的生物质需要细胞培养/发酵培养基，但生物转化步骤不需要培养基，因为反应缓冲液在合适的pH值下就足够了。

[0891] 本公开内容的生物转化方法在时间、温度、pH和增溶剂的条件下进行以提供均法呢醇或双均法呢醇原液(例如包含EEH或bisEEH)到(-)-降龙涎香醚或龙涎香氧化物的转化。

[0892] 对于SHC/HAC酶变体，反应混合物的pH值可以在4-8的范围内，优选5至6.5，更优选4.8-6.0，对于野生型SHC/HAC酶，可以在约pH5.0至约pH7.0的范围内，并且可以通过向反应混合物中添加缓冲液来维持。用于此目的的示例性缓冲剂包括但不限于柠檬酸缓冲剂、磷酸盐缓冲剂、乙酸缓冲剂和/或琥珀酸缓冲剂。

[0893] 优选的温度为约15℃至约60℃，例如约15℃至约50℃或约15℃至约45℃或约30℃至约60℃或约40℃至约50℃。温度可以保持恒定，也可以在生物转化过程中改变。

[0894] 申请人已经证明在生物转化反应中包括增溶剂(例如表面活性剂、去污剂、增溶

剂、水混溶性有机溶剂等)可以是有用的。

[0895] 如本文所用,术语“表面活性剂”是指降低两种液体之间或液体和固体之间的表面张力(或界面张力)的组分。表面活性剂可用作去污剂、润湿剂、乳化剂、发泡剂和分散剂。表面活性剂的实例包括但不限于Triton X-100、Tween 80、牛磺脱氧胆酸盐、牛磺脱氧胆酸钠、十二烷基硫酸钠(SDS)和/或月桂基硫酸钠(SLS)。

[0896] 虽然Triton X-100可用于部分纯化SHC/HAC酶或酶变体(以可溶或膜级分/悬浮液形式),但它也可用于生物转化反应(参见例如Seitz(如上文所引用的2012博士论文)以及Neumann和Simon(1986,如上文所引用的)和JP2009060799中的公开内容)。然而,令人惊讶的是,申请人从一长串其他不太有用的增溶剂中选择并鉴定了SDS作为特别有用的增溶剂。特别地,就均法呢醇到(-)-降龙涎香醚的生物转化反应的反应速度和产率而言(当EEH以4g/l和125g/l使用时),申请人鉴定SDS为比例如Triton X-100显著更好的增溶剂。对于至少一种SHC/HAC衍生酶,在4g/l EEH和10的OD_{650nm}的细胞情况下,反应中使用Triton X-100(浓度范围为约0.005%至0.48%)的均法呢醇到(-)-降龙涎香醚的最大生物转化活性可能仅为使用SDS(浓度为约0.07%)获得的活性的约20%。

[0897] 不希望受理论束缚,将SDS与重组微生物宿主细胞一起使用可以是有利的,因为SDS可以有利地与宿主细胞膜相互作用以使得SHC/HAC酶或酶变体(其是膜结合酶)更容易接近均法呢醇底物。此外,在反应混合物中以合适的水平包含SDS可以改善乳液(水中的均法呢醇)的性质和/或改善宿主细胞内均法呢醇底物对SHC酶的可及性,同时防止破坏(例如SHC(WT或SHC/HAC变体)酶的变性)。生物转化反应中使用的增溶剂(例如SDS)的浓度受生物质的量和底物(EEH)浓度的影响。也就是说,增溶剂(例如SDS)浓度、生物质的量和底物(EEH)浓度之间存在一定程度的相互依赖性。例如,随着均法呢醇底物浓度的增加,需要足够量的生物催化剂和增溶剂(例如SDS)来发生有效的生物转化反应。例如,如果增溶剂(例如SDS)浓度过低,可能会观察到次优的均法呢醇转化。另一方面,例如,如果增溶剂(例如SDS)浓度太高,则可能存在通过破坏完整微生物细胞和/或SHC/HAC酶或酶变体的变性/失活影响生物催化剂的风险。在生物质的量和底物(EEH)浓度的背景下选择合适的SDS浓度在本领域技术人员的知识范围内。例如,技术人员可以开发预测模型以确定合适的SDS、底物(EEH)和生物质浓度。当使用4g/l EEH和10.0的OD(650nm)的生物催化剂时,0.010-0.075%范围内的SDS可以是合适的。当使用125g/l EEH和2倍生物质湿重时,经调整的SDS浓度(1.3%)可以是合适的。然而,使用不同SDS/细胞比率值对EEH转化为(-)-降龙涎香醚的百分比的研究表明,正确选择生物催化剂、均法呢醇底物和增溶剂(例如SDS)的比率有助于开发稳健的生物转化反应系统,其对一定范围的SDS浓度和pH范围表现出一定程度的容忍度。

[0898] WT SHC酶(例如AacSHC)的生物转化反应的温度可以为约30℃至约60℃,例如约45℃至约60℃,例如约50℃至约60℃,例如约55℃。

[0899] WT SHC酶(例如AacSHC)的生物转化反应的pH范围可为约5.0至7.0,更优选约5.6至约6.2,甚至更优选约6.0。

[0900] SHC/HAC酶变体的生物转化反应的温度可以是约30℃至约55℃,例如约40℃至约50℃,例如约45℃。

[0901] SHC/HAC酶变体的生物转化反应的pH可为约4.8-6.4,优选约5.2-6.0。

[0902] 优选地,用于生物转化反应的增溶剂是SDS。当使用约4g/l的EEH和10的OD_{650nm}的细

胞时，用于WT SHC酶(例如AacSHC)的生物转化反应中的SDS浓度可以在约0.010-0.075％的范围内，优选约0.030％。

[0903] 当使用约4g/l的EEH和10的OD_{650nm}的细胞时，用于SHC/HAC酶变体的生物转化反应中的SDS浓度可以在约0.0025-0.090％的范围内，优选约0.050％。

[0904] 当反应加载有约4g/l EEH的EEH浓度的均法呢醇时，可将生物催化剂加载到反应中至约10.0的OD(650nm)。

[0905] 当在生物转化反应中生物催化剂与EEH均法呢醇的比率为大约2:1时，[SDS]/[细胞]比率可以在大约10:1-20:1的范围内，优选大约15:1-18:1，优选大约16:1。

[0906] 在均法呢醇浓度为约125g/l EEH和生物催化剂浓度为250g/l(对应约175的OD(650nm))时，SHC变体酶的生物转化反应中的SDS浓度可以在约1-2％的范围内，优选在约1.4-1.7％的范围内，甚至更优选为约1.5％。

[0907] 生物催化剂与EEH均法呢醇底物的比例可以在约0.5:1-2:1的范围内，在一些实施方案中为2:1，优选约1:1或0.5:1。

[0908] 例如，SHC/HAC酶变体的最佳温度可以等于或大于约35℃。例如，SHC/HAC酶变体的最佳温度可以为约40℃至约50℃，例如约42℃至约48℃或约44℃至约46℃。例如，SHC/HAC酶变体的最佳温度可以为约45℃。本文公开的制备(-)-降龙涎香醚或龙涎香氧化物的方法可以在SHC/HAC酶变体的最佳温度下进行。

[0909] SHC/HAC酶变体的最佳pH可以例如等于或大于约5.4。例如，SHC/HAC酶变体的最佳pH可以为约5.2至约6.0，例如约5.4至约5.8，例如约5.6至约5.8。例如，SHC/HAC酶变体的最佳pH可以是约5.6或约5.8。本文公开的制备(-)-降龙涎香醚或龙涎香氧化物的方法可以在SHC/HAC酶变体的最佳pH下进行。

[0910] 本文公开的制备(-)-降龙涎香醚或龙涎香氧化物的方法的反应介质中十二烷基硫酸钠(SDS)的最佳浓度可以例如为约0.010w/w％至约0.10w/w％。例如，当在10的OD_{650nm}的细胞的情况下以4g/l使用底物(例如EEH或BisEEH)时，SDS的最佳浓度可为约0.040w/w％至约0.080w/w％，例如约0.050w/w％。本文公开的制备(-)-降龙涎香醚或龙涎香氧化物的方法可以使用本文描述的SDS的最佳浓度进行。在本文公开的制备(-)-降龙涎香醚或龙涎香氧化物的方法的反应介质中的十二烷基硫酸钠(SDS)的最佳浓度可以例如在底物(例如EEH或BisEEH)以125g/l与250g/l细胞一起使用时为约1.0w/w％至约1.5w/w％。例如，当底物(例如EEH或BisEEH)以125g/l与250g/l细胞一起使用时，SDS的最佳浓度可以为约1.2w/w％至约1.4w/w％，例如约1.3w/w％或当底物(例如EEH或BisEEH)以125g/l与125g/l细胞一起使用时，为约0.65w/w％。

[0911] 本文公开的制备(-)-降龙涎香醚或龙涎香氧化物的方法可以在针对所用的特定酶的最佳温度范围内或最佳温度下和/或最佳pH范围内或最佳pH下和/或SDS最佳浓度范围内或最佳SDS浓度下进行，如下面实施例中的表7或9或11所示。

[0912] 在一些实施方案中，使用添加了均法呢醇底物的生物催化剂产生降龙涎香醚。

[0913] 在一些实施方案中，使用添加了双均法呢醇底物的生物催化剂产生龙涎香氧化物。

[0914] 可以通过使用已知方式(例如蠕动泵、输液注射器等)进料来添加底物。均法呢醇和双均法呢醇可以是油溶性化合物并以油形式提供。鉴于生物催化剂(微生物细胞，例如完

整的重组全细胞和/或细胞碎片和/或固定化的酶)存在于水相中,当向生物转化反应混合物中加入均法呢醇或双均法呢醇时,生物转化反应可被视为三相系统(包括水相、固相和油相)。即使存在SDS时也是如此。为了澄清的目的,当可溶性WT SHC或SHC/HAC酶变体用作生物催化剂时,这被认为是两相系统。

[0915] 发酵罐可用于将表达SHC/HAC酶或酶变体基因并产生活性SHC/HAC酶或酶变体的重组宿主细胞培养至适合在同一发酵罐容器中用作生物催化剂的足够生物质浓度,其用于将均法呢醇源转化为(-)-降龙涎香醚(例如与副产物(II)、(IV)和/或(III)中的一种或多种混合),或将双均法呢醇源转化为龙涎香氧化物(例如与副产物(XI)、(XII)和/或(XIII)中的一种或多种混合)。

[0916] 本领域技术人员将理解更高的累积产生效价可以通过实施连续过程(例如产品去除、底物进料和生物质添加或(部分)取代)来实现。优选地,在包含SHC/HAC酶或酶变体的重组宿主细胞的存在下,EEH向(-)-降龙涎香醚或BisEEH向龙涎香氧化物的生物转化产生5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100的降龙涎香醚或龙涎香氧化物产率,以摩尔百分比给出并且基于所使用的EEH或BisEEH的摩尔数;特别优选地,产率为5至100、10至100、和100、25至100、30至100、35至100,特别是40至100、45至100、50至100、60至100、70至100摩尔百分比。

[0917] SHC/HAC酶或酶变体的活性通过反应速率(产物量/(产物量+剩余起始材料的量))x100)以摩尔百分比定义。优选地,在存在SHC/HAC酶或酶变体的情况下,EEH到(-)-降龙涎香醚或BisEEH到龙涎香氧化物的生物转化产生5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100的(-)-降龙涎香醚或龙涎香氧化物产率,以摩尔百分比给出并且基于所使用的EEH或BisEEH的摩尔数;特别优选地,产率为5至100、10至100、20至100、25至100、30至100、35至100,特别是40至100、45至100、50至100、60至100、70至100。

[0918] 在本发明的一个优选实施方案中,产率和/或反应速率在例如4、6、8、10、12、16、20、24、36或48小时的定义时间段内测定,在此期间通过包含编码SHC/HAC酶或酶变体的核苷酸序列的重组宿主细胞将EEH转化为(-)-降龙涎香醚或将BisEEH转化为龙涎香氧化物。在进一步的变体中,反应在例如25℃、30℃、40℃、50℃或60℃的精确定义的条件下进行。具体地,产率和/或反应速率通过在24-72小时的时间段内在35℃下进行通过根据本发明的SHC/HAC酶或酶变体将EEH转化为(-)-降龙涎香醚或将BisEEH转化为龙涎香氧化物的反应来确定。

[0919] 在本发明的另一个实施方案中,包含编码SHC/HAC酶变体的核苷酸序列的重组宿主细胞的特征在于它在均法呢醇的反应以产生(-)-降龙涎香醚中和在双均法呢醇的反应以产生龙涎香氧化物中显示出与在相同条件下的WT SHC或SHC/HAC衍生酶相比2-、3-、4-、5-、6-、7-、8-、9-、10-、11-、12-、13-、14-、15-、16-、17-、18-、19-、20-、21-、22-、23-、24-、

25-、26-、27-、28-、29-、30-、31-、32-、33-、34-、35-、36-、37-、38-、39-、40-、41-、42-、43-、44-、45-、46-、47-、48-、49-、50-、51-、52-、53-、54-、55-、56-、57-、58-、59-、60-、61-、62-、63-、64-、65-、66-、67-、68-、69-、70-、71-、72-、73-、74-、75-、76-、77-、78-、79-、80-、81-、82-、83-、84-、85-、86-、87-、88-、89-、90-、91-、92-、93-、94-、95-、96-、97-、98-、99-、100-、200-、500-、1000倍或更高的产率和/或反应速率。这里，术语条件涉及反应条件，例如底物浓度、酶浓度、反应时间和/或温度。

[0920] 在包含编码WT/参考SHC或SHC/HAC衍生物的核苷酸序列的大肠杆菌重组菌株中成功开发用于从均法呢醇制备(-)-降龙涎香醚的生物转化方法可以提供用于(-)-降龙涎香醚产生的低成本和工业经济的方法。

[0921] 在包含编码WT/参考SHC或SHC/HAC衍生物的核苷酸序列的大肠杆菌重组菌株中成功开发用于从双均法呢醇制备龙涎香氧化物的生物转化方法可以提供用于龙涎香氧化物产生的低成本和工业经济的方法。

[0922] 在整个说明书和所附的权利要求书中，除非上下文另有要求，否则词语“包含”以及变型诸如“含有”和“包括”将被理解为暗示包含所陈述的整数或步骤或整数或步骤组，但不排除任何其他整数或步骤或整数或步骤组。术语“包含”也意指“包括”以及“组成”，例如“包含”X的组合物可以仅由X组成，或者可以包括一些额外的东西，例如:X+Y。还必须注意的是，如在本说明书和所附权利要求中使用的，单数形式“一个”、“一种”和“该”包括复数指示物，除非内容另有明确规定。例如，提及“一个/种基因”或“一个/种酶”是提及“一个或多个/一种或多种基因”或“一个或多个/一种或多种酶”。

[0923] 应当理解，本公开内容不限于本文所述的特定方法、方案和试剂，因为这些可能会有所不同。还应理解，本文使用的术语仅出于描述特定实施方案的目的，并不旨在限制本公开内容的范围，本公开内容的范围将仅由所附权利要求限制。除非另有定义，本文使用的所有技术和科学术语具有与本领域技术人员通常理解的相同含义。根据本公开内容，可以使用在本领域技术范围内的常规分子生物学、微生物学和重组DNA技术。

[0924] 本公开内容在其应用中不限于在以下描述中阐述或在附图中示出的构造细节和部件布置。本公开内容能够实现其他实施方案且能够以各种方式实践或执行。此外，本文使用的措辞和术语是出于描述的目的，不应被视为限制。优选地，本文使用的术语如"A multilingual glossary of biotechnological terms:(IUPAC Recommendations)", Leuenberger,H.G.W,Nagel,B.和Kolbl,H.eds.(1995),Helvetica Chimica Acta,CH-4010Basel,Switzerland)中所描述的定义。

[0925] 在本说明书的全文中引用了几个文件。本文引用的每个文件(包括所有专利、专利申请、科学出版物、制造商的说明书、说明书、GenBank登录号序列提交等)(无论是上文还是下文)均通过引用整体并入本文。

[0926] 以g/l(例如生物催化剂悬浮液)或w/w%(例如SDS)浓度制备用于生物转化反应(125g/l底物或更高)的材料的原液。如例如实施例4中所述，通过按重量添加所需量的个体组分(例如底物、生物催化剂悬浮液、SDS溶液、反应缓冲液的原液)来建立生物转化反应。为此，反应混合物的实际体积以质量单位而不是体积单位表示(参见例如实施例4)。因为使用的底物是油(密度<1)，所以例如150g的总反应质量(例如在实施例4中)不会转化为150ml的反应体积。尽管如此，生物转化反应混合物中个体组分的浓度以每体积的质量(例如g/l或

(w/v)%)表示,并因此不考虑例如150g的总反应质量(如实施例4中的示例)不精确对应于150ml的反应体积(如实施例4中的示例)。

[0927] 为避免疑义,实施例、附图和附图说明中所有提到的EEH都是指均法呢醇的EE异构体。所用的EEH取自包含重量比为80:20的EE:EZ异构体混合物的均法呢醇原液。

[0928] 本文描述的实施例对本公开内容是说明性的并且不旨在对其进行限制。已经根据本公开内容描述了本公开内容的不同实施方案。在不背离本公开内容的精神和范围的情况下,可以对本文描述和说明的技术进行许多修改和变化。因此,应当理解,这些实施例仅是说明性的,并不限制本公开内容的范围。

实施例

[0929] 实施例1-新SHC/HAC酶变体的产生和测试

[0930] SHC质粒制备

[0931] 将编码野生型或变体角鲨烯藿烯环化酶(SHC)酶的基因插入到质粒pET-28a(+)中,在其中它在IPTG诱导型T7启动子的控制下用于在大肠杆菌中产生蛋白质。使用标准热激转化方案将质粒转化到大肠杆菌菌株BL21(DE3)中。

[0932] 培养基制备

[0933] 如下制备为350ml培养物默认选择的基本培养基:向35ml柠檬酸/磷酸盐原液(133g/l KH_2PO_4,40g/l $(NH_4)_2HPO_4$,17g/l柠檬酸.H_2O,pH调节至6.3)中加入307ml H_2O,根据需要用32%NaOH将pH调节至6.8。高压灭菌后,加入0.850ml 50%$MgSO_4$、0.035ml微量元素溶液(见下文)、0.035ml硫胺素溶液和7ml 20%葡萄糖。

[0934] 微量元素溶液:在去离子水中的50g/l $Na_2EDTA.2H_2O$、20g/l $FeSO_4.7H_2O$、3g/l H_3BO_3、0.9g/l $MnSO_4.2H_2O$、1.1g/l $CoCl_2$、80g/L $CuCl_2$、240g/l $NiSO_4.7H_2O$、100g/l KI、1.4g/l $(NH_4)_6Mo_7O_{24}.4H_2O$、1g/l $ZnSO_4.7H_2O$,

[0935] 硫胺素溶液:在去离子水中的2.25g/l硫胺素.HCl

[0936] $MgSO_4$溶液:在去离子水中的50%(w/v)$MgSO_4.7H_2O$

[0937] 小规模生物催化剂产生(SHC/HAC野生型酶或酶变体)

[0938] 从含有SHC生产质粒的大肠杆菌菌株BL21(DE3)的预培养物中接种350ml培养物(补充有50μg/ml卡那霉素的培养基)。细胞在37℃下以恒定搅拌(250rpm)生长至光密度为约0.5(OD_{650nm})。

[0939] 然后通过将IPTG添加到300μM的浓度诱导蛋白质产生,然后在不断振荡的情况下再孵育5-6小时。最终通过离心收集所得生物质并用例如50mM Tris-HCl缓冲液pH7.5洗涤。细胞在4℃或-20℃下以沉淀形式储存,直至进一步使用。通常从1升培养物中获得2.5至4克细胞(湿重),与使用的培养基无关。

[0940] 发酵罐中的生物催化剂生产

[0941] 在750ml InforsHT反应器中制备和运行发酵。向发酵容器中加入168ml去离子水。反应容器配备了所有需要的探头(pO_2、pH、取样、消泡剂)、C+N进料和氢氧化钠瓶并进行高压灭菌。高压灭菌后向反应器加入:

[0942] 20ml 10x磷酸盐/柠檬酸缓冲液

[0943] 14ml 50%葡萄糖

[0944] 0.53ml $MgSO_4$溶液

[0945] 2ml $(NH_4)_2SO_4$溶液

[0946] 0.020ml微量元素溶液

[0947] 0.400ml硫胺素溶液

[0948] 0.200ml卡那霉素原液

[0949] 运行参数设置如下:pH=6.95,pO_2=40%,T=30℃,以300rpm搅拌。级联:rpm设定点为300,最小300,最大1000,流量(l/min)设定点0.1,最小0,最大0.6。消泡剂控制:1:9。

[0950] 从种子培养物接种发酵罐到0.4-0.5的OD_{650nm}。该种子培养物在LB培养基(+卡那霉素)中在37℃、220rpm下生长8小时。发酵首先以分批模式运行11.5小时,然后使用进料溶液(灭菌的葡萄糖溶液(143ml H_2O+35g葡萄糖),灭菌后向其添加了:17.5ml $(NH_4)_2SO_4$溶液、1.8ml $MgSO_4$溶液、0.018ml微量元素溶液、0.360ml硫胺素溶液、0.180ml卡那霉素原液)开始C+N进料。进料以约4.2ml/h的恒定流速运行。葡萄糖和NH_4^+测量在外部进行以评估培养物中的C源和N源的可用性。通常葡萄糖水平保持在非常低的水平。

[0951] 培养物总共生长大约25小时,其中它们通常达到40-45的OD_{650nm}。然后通过在发酵罐中将IPTG添加到大约1mM的浓度开始SHC产生(作为IPTG脉冲或使用输注注射器在3-4小时内),将温度设置为40℃,将pO_2设置为20%。SHC产生的诱导在40℃持续16小时。在诱导结束时,通过离心收集细胞,用0.1M柠檬酸/柠檬酸钠缓冲液pH5.4洗涤,并在4℃或-20℃下以沉淀形式储存直至进一步使用。

[0952] 新的SHC/HAC酶变体的产生和选择

[0953] 使用先前在WO2016/170099中描述的酸热脂环酸杆菌215G2SHC变体(215G2 SHC)基因作为模板进行酶进化程序。产生了大约9950个SHC变体的文库,并筛选了显示E,E-均法呢醇(EEH)环化能力增加的变体。在含有34g/l EEH和0.075%十二烷基硫酸钠(SDS)的50mM琥珀酸/NaOH缓冲液pH5.2(0.200ml)中进行筛选(微量滴定板设置)。反应在35℃下在恒定搅拌下进行24小时(轨道振荡,210rpm)。

[0954] 在相同的反应设置中选择了大约495个变体用于确认。从这495个变体中,选择了48个用于在0.500ml体积反应(相同条件)中使用新鲜产生的生物催化剂(2ml培养物)进行确认。

[0955] 最后,在含有EEH底物、0.1%SDS和40.0的OD_{600nm}的已产生SHC变体的细胞的50mM琥珀酸/NaOH缓冲液pH5.2中,使用显示对EEH的提高活性的先前48种变体中的16种进行在34、68或136g/l下的EEH转化。最终体积为1ml,反应在35℃下孵育并在磁力搅拌器上剧烈搅拌。随时间对反应进行取样、提取溶剂并通过气相色谱分析以确定EEH转化为(-)-降龙涎香醚)(参见下面的分析方法)。

[0956] 由此,选择了3个具有改进的EEH环化活性的变体(90C7、110B8和115A7)。在这3个变体上共鉴定了7个突变。从这3种变体产生生物催化剂,并且EEH生物转化在0.975%SDS的存在下在100mM乙酸/乙酸钠缓冲液pH4.7中以100g/l EEH和250g/l细胞运行。将反应(2ml体积)在恒定搅拌(磁力搅拌)下在35℃孵育72小时,并随时间取样用于GC分析。

[0957] 然后进行突变研究以确定哪些已鉴定的单一突变或突变组合有利于EEH环化为(-)-降龙涎香醚。构建了15个额外的变体并产生了它们的生物催化剂。在50mM琥珀酸/NaOH缓冲液pH4.7-5.2中在含有34g/l EEH、77g/l细胞和0.4%SDS或34g/l EEH、155g/l细胞和

0.7％SDS或34g/l EEH、250g/l细胞和1.0％SDS的反应中测试了这些变体的活性。反应(1ml体积)在35℃下在恒定搅拌(磁力搅拌)下孵育，随时间取样并提取用于GC分析。

[0958] GC分析

[0959] 用适当体积的叔丁基甲基醚(MBTE/tBME)提取样品，以量化它们在底物和反应产物方面的含量。在用气相色谱分析之前，通过离心将溶剂部分与水相分离。将1μl溶剂相(分流比10)注入30m x 0.32mm x0.25μm Zebron ZB-5柱。柱以恒定流速(4ml/min H_2)冲洗，使用温度梯度：100℃，15℃/min至200℃，120℃/min至240℃，在240℃ 4分钟。入口温度：250℃，检测器温度：250℃。这导致EEH和EZH的分离以及对应于(-)-降龙涎香醚和由均法呢醇环化产生的3种其他反应产物的峰的分离。它还导致Bis-EEH、Bis-EZH的分离和对应于龙涎香氧化物和由双均法呢醇环化产生的3种其他反应产物的峰的分离。

[0960] EEH转化率由EEH环化和EEH产生的峰面积计算，公式如下：

[0961] EEH转化率(％)＝100x(面积$_{来自EEH的产物的峰}$/(面积$_{来自EEH的产物的峰}$+面积$_{EEH峰}$))

[0962] BisEEH转化率以与EEH转化率相同的方式由BisEEH环化和BisEEH产生的峰面积以相同方式计算。

[0963] 结果

[0964] 在评估的15个变体中，下表6中列出的6个变体显示出最佳活性。表6列出了在这些变体中鉴定的新突变。

[0965] 表6.选择的新SHC变体酶中的突变。

[0966]

SHC 变体	Y81H	T90A	A172T	M277K	H431L	A557T	R613S
110B8 (SEQ ID NO: 5)	+				+	+	
90C7 (SEQ ID NO: 17)		+					+
115A7 (SEQ ID NO: 18)			+	+			
SHC 49 (SEQ ID NO: 2)					+	+	
SHC 65 (SEQ ID NO: 3)						+	+
SHC 66 (SEQ ID NO: 4)	+					+	+

[0967]

[0968] 注意：除了存在于215G2 SHC中的突变之外还出现这些突变：M132R、A224V和I432T

[0969] 如先前在将野生型酸热脂环酸杆菌SHC进化为215G2 SHC时观察到的，在所产生的SHC变体中鉴定的新突变不在酶的活性位点附近。大多数已鉴定的新突变再次位于酶晶体

结构的结构域2中(T90A、A172T、M277K和H431L)。两个位于结构域1(A557T和R613S)中,1个位于2个结构域之间的接口处(Y81H)。

[0970] 实施例2-新SHC/HAC酶变体的反应条件的优化

[0971] 研究的反应参数:温度、SDS浓度和pH

[0972] 与实施例1相关的上表6中列出的SHC变体的反应条件在温度、pH和SDS浓度方面分别进行了优化。

[0973] 使用用相应质粒转化的大肠杆菌细胞如上所述通过发酵从不同变体制备生物催化剂。通过离心收集细胞,并储存在-20℃直至进一步使用。产生的生物催化剂显示出非常相似的SHC含量。因此可以得出结论,观察到的活性差异是由于插入的突变造成的。

[0974] 结果

[0975] 在0.0125-0.125％SDS的存在下,在28至50℃范围内的温度和在恒定搅拌下(Heidolph synthesis 1Liquid装置,900rpm),在0.1M柠檬酸/磷酸钠缓冲液pH5.0-6.8中进行具有4g/l EEH和在10.0的OD_{650nm}下加载的生物催化剂的2-5ml体积的反应。

[0976] 下表7中列出的条件似乎是个体的最佳条件。它们在早先定义为最佳的pH附近的pH下在0.1M琥珀酸/NaOH缓冲液中运行的反应中得到证实。

[0977] 表7.SHC变体酶的最佳反应条件。

[0978]

SHC 变体	最佳温度 (°C)	最佳温度范围(°C)	最佳 pH	最佳 pH 范围	最佳 [SDS]	最佳[SDS] 范围(w/w %)

[0979]

					(w/w %)	
215G2 (SEQ ID NO: 10)	35	32-50	5.4	5.0-6.2	0.060	0.010-0.10
110B8 (SEQ ID NO: 5)	45	40-50	5.8	5.6-6.0	0.050	0.040-0.070
SHC 49 (SEQ ID NO: 2)	45	40-50	5.6	5.4-5.8	0.050	0.040-0.080
SHC 65 (SEQ ID NO: 3)	45	40-50	5.6	5.2-5.8	0.050	0.040-0.10
SHC 66 (SEQ ID NO: 4)	45	40-46	5.8	5.6-6.0	0.060	0.040-0.080
90C7 (SEQ ID NO: 17)	45	40-46	5.6	5.2-6.0	0.040	0.025-0.050
115A7 (SEQ ID NO: 18)	35	32-36	5.4	5.2-5.8	0.050	0.040-0.075

[0980] 注意到与215G2 SHC亲本酶的一些偏差。除了仅一个例外，新突变的引入将最佳温度移动了10℃，从35℃至约45℃。

[0981] 实施例3-在初始和优化的反应条件下测试新的SHC/HAC酶变体活性

[0982] 为了比较SHC变体的活性，通过从用适当质粒转化的大肠杆菌细胞发酵产生生物催化剂。所得细胞用于在初始条件(35℃，pH5.4，0.070％SDS)下或在设置为针对每个变体定义为最佳的条件的T、pH和[SDS]下测试在4g/l EEH下和使用加载至10.0的OD_{650nm}的生物催化剂的EEH转化率(参见上面实施例2中的表7)。将反应混合物的样品加载到SDS-PAGE凝胶上用于分析反应的SHC含量。该分析证实所有反应都含有相同量的SHC酶。

[0983] 图1显示了SHC变体在初始条件(pH5.4、35℃、0.065％SDS、10的OD_{650nm}的细胞)(其先前被确定为215G2 SHC活性的最佳反应条件)以及单独优化的反应条件(可变pH、T和[SDS])下的相对活性。

[0984] 图2显示了SHC变体在优化的条件下与亲本215G2 SHC酶相比的相对活性。得出的结论是,在215G2 SHC中仅引入少数突变导致具有改善或显著改善的活性(EEH环化为(-)-降龙涎香醚)的变体。

[0985] 实施例4-SHC变体在125g/l EEH生物转化中的相对性能

[0986] 通过从用携带编码所选新SHC变体或亲本215G2 SHC的基因的质粒转化的大肠杆菌菌株发酵产生的生物催化剂用于125g/l EEH生物转化。

[0987] 如下在0.75升Infors发酵罐中设置典型的反应(150g总体积)。反应容器加载有适量的均法呢醇,对应于18.75g EEH。从在去离子水中制备的31％(w/w)溶液加入1.95g SDS。通过将细胞悬浮在0.1M琥珀酸/NaOH缓冲液pH5.1中,从已产生感兴趣的SHC变体的大肠杆菌细胞制备细胞悬浮液。通过在10℃和17210g下离心10分钟确定该细胞悬浮液的细胞湿重浓度后,将适当体积的细胞添加到反应容器中以将37.5g细胞引入反应中。用所需量的反应缓冲液pH5.1完成反应体积至150g。在确定为对于其活性最佳的温度下,并在700rpm的恒定搅拌下,用每种SHC变体进行反应。使用85％H_3PO_4将pH设置为对于每个SHC变体或亲本215G2 SHC确定为最佳的值。根据需要使用85％磷酸手动完成pH调节。随时间对反应进行取样(1ml),用5个体积的MTBE/tBME(5ml)提取。在通过离心(台式离心机,13000rpm,2分钟)澄清溶剂相并10至20倍稀释到MTBE/tBME中后,通过GC分析测定反应中的均法呢醇和(-)-降龙涎香醚含量。

[0988] 图3显示了SHC变体和215G2 SHC(亲本酶)在125g/l EEH和250g/l细胞下和在最佳pH和温度条件下使用每种SHC酶运行的生物转化中将EEH转化为(-)-降龙涎香醚。

[0989] 图4显示了4种最佳SHC变体的相对改善,作为与使用亲本酶215G2 SHC所获得的相比较的EEH转化率的％增加。观察到初始反应速度显著增加,即在反应的前6到8小时内实现了EEH转化。预计在使用SHC#65的125g/l EEH转化中应用的反应条件的进一步调整将允许该变体的活性相对于亲本SHC 215G2 SHC的活性的进一步增加,并与其他SHC变体一样达到完全转化。

[0990] 数据表明,使用215G2 SHC作为亲本酶的酶进化产生了具有改进的EEH生物转化特性的SHC酶。仅引入少数氨基酸突变可显著增加EEH生物转化。以这种方式,EEH的转化可以在时间上减少大约1/3或1/2,取决于使用的SHC变体,在大约24小时而不是72小时内获得完成。

[0991] 实施例5-E,E-双均法呢醇环化为龙涎香氧化物

[0992] 双均法呢醇的合成

[0993] 双均法呢醇是作为两种异构体的混合物产生的,如图10所示。

[0994] 步骤1:

[0995] 将图10的化合物1(E-橙花叔醇)(700.0g,3.148mol,1.0当量)、乙基乙烯基醚(703.7g,9.758mol,3.1当量)和苯基膦酸(2.1g,0.0132mol,0.004当量)装入高压釜中,将混合物加热至180℃持续2.5小时。(注意:内部压力升至14bar)。将反应混合物冷却至室温,从高压釜中取出并减压浓缩以除去过量的乙基乙烯基醚。粗产物用MTBE(1400ml)稀释,用

2.0M HCl(2×700ml)、水(500ml)和盐水(500ml)洗涤。有机层用硫酸钠干燥,过滤并减压浓缩。粗产物通过二氧化硅塞,然后蒸馏,得到图10的化合物2(黄色液体,483.0g,61.7%)。

[0996] 步骤2:

[0997] 将在THF(5Vol)中的化合物2(420.0g,1.6907mol,1.0当量)在0℃下在2.0小时内添加到Red-Al的搅拌溶液(283.26g,0.980mol,0.58当量,70%甲苯)中。将反应混合物搅拌过夜并通过TLC监测。完成后,通过使用在室温下搅拌1小时的0℃的10%NaOH溶液(670.0ml)缓慢淬灭反应。然后将反应混合物用冰冷的水(2L)稀释并用MTBE(2x3L)提取。合并的有机层用水(1x2L)和盐水(1x2L)洗涤并用硫酸钠干燥。将有机层真空浓缩得到粗化合物,将其通过真空蒸馏纯化以得到E,E/Z-双均法呢醇(淡黄色液体,280.0g,66.1%)。

[0998] 双均法呢醇是作为约80:20的E,E:E,Z比例的异构体混合物产生的。

[0999] E,E-双均法呢醇环化为龙涎香氧化物

[1000] 产生的新SHC变体用于将双均法呢醇(E,E-和E,Z-双均法呢醇的混合物)生物转化为龙涎香氧化物。鉴定了图11中的反应产物:

[1001] 在4g/l E,E-双均法呢醇下的环化

[1002] 变体的活性在单独定义为最佳和如上表7中所列的反应条件(温度、pH、SDS浓度)下进行测试。反应(5ml体积)含有4g/l E,E-双均法呢醇作为底物。将产生SHC变体的细胞添加至10.0的OD_{650nm}以开始反应。将反应在Heidolph Synthesis 1Liquid 16装置上在恒定搅拌(900rpm)下孵育。

[1003] 125g/l E,E-双均法呢醇下的生物转化率

[1004] 如实施例4所示,在0.75升Infors发酵罐中建立典型反应(150g总体积),但用125g/l E,E-双均法呢醇而不是E,E-均法呢醇加载反应。反应在35℃下在恒定搅拌(700rpm)和pH5.4下在具有所有SHC变体的0.1M琥珀酸/NaOH缓冲液中进行。

[1005] 结果-4g/l E,E-双均法呢醇环化

[1006] 与亲本215G2 SHC/HAC酶变体相比,对于SHC/HAC酶变体#65、#66和#110B8观察到活性的显著增加。结果如图5所示。

[1007] 结果-125g/l E,E-双均法呢醇的生物转化率

[1008] E,E-双均法呢醇转化率如图6所示。对于每个测试的变体都观察到转化率增加。在反应的72小时,所有SHC变体的E,E-BisHF的转化率为95%或更高。对于变体SHC#66或110B8 SHC,在反应的约48小时获得95%的转化率,即仅为用亲本酶215G2 SHC进行生物转化所需时间的2/3。

[1009] 用100ml MTBE通过剧烈摇动然后通过离心(6000g,10分钟,室温)进行相分离将反应提取5次。从回收的溶剂相的CG分析判断,这几乎可以完全提取反应产物。

[1010] 用E,E-双均法呢醇测试的变体显示出比亲本215G2 SHC酶改善的活性,尽管反应条件未设定为针对新SHC变体定义为最佳的pH和温度条件。预计当pH和温度设置为确定为对上述实施例中列出的各个变体最佳的值时会观察到活性的进一步增加。

[1011] 结果-反应产物的鉴定

[1012] 化合物X,龙涎香氧化物

[1013] ^1H NMR(400MHz,$CDCl_3$)δ3.71-3.59(m,2H),1.76(dt,J=11.7,3.8Hz,1H),1.71-1.57(m,6H),1.46-1.24(m,6H),1.26(s,3H),1.15(td,J=13.2,4.2Hz,1H),0.97(dd,J=

13.2,4.2Hz,1H),0.89-0.85(m,1H),0.88(s,3H),0.81(s,3H),0.76(s,3H).^{13}C NMR(101MHz,$CDCl_3$)δppm 74.67,60.89,57.81,56.43,42.10,41.96,38.97,36.88,33.36,33.29,27.68,21.36,19.94,19.91,18.60,18.15,15.60

[1014] 化合物XI

[1015] ^{1}H NMR(400MHz,$CDCl_3$)δppm 3.61-3.57(m,2H),2.09(td,J=12.84,4.40Hz,1H),1.82(dt,J=12.23,3.91Hz,1H),1.70-1.62(m,2H),1.60-1.33(m,7H),1.39(br s,3H),1.31-1.19(m,2H),1.15-0.99(m,3H),1.12(br s,3H),0.87(s,3H),0.80(s,3H).^{13}C NMR(101MHz,$CDCl_3$)δppm 75.56,60.27,54.60,46.99,42.51,38.32,36.62,33.45,33.34,31.74,30.92,26.64,24.99,21.91,21.78,20.96,18.91

[1016] 化合物XII

[1017] ^{1}H NMR(600MHz,BENZENE-d_6)δppm 4.94(s,1H),4.64(s,1H),3.38(br t,J=5.6Hz,2H),2.37(br d,J=12.4Hz,1H),1.94(td,J=12.8,4.1Hz,1H),1.72-1.57(m,3H),1.56-1.46(m,3H),1.44-1.33(m,3H),1.33-1.24(m,2H),1.13(td,J=13.4,3.4Hz,1H),0.96(br d,J=12.8Hz,1H),0.89(td,J=13.4,2.6Hz,1H),0.85(s,3H),0.80(s,3H),0.72(s,3H),0.62(br s,OH).^{13}C NMR(151MHz,BENZENE-d_6)δppm 148.70,106.52,62.78,56.57,55.25,42.11,39.90,38.92,38.34,33.70,33.44,32.03,24.38,21.58,19.64,19.55,14.27

[1018] 化合物XIII

[1019] ^{1}H NMR(600MHz,BENZENE-d_6)δppm 5.15(br dd,J=10.5,4.9Hz,1H),3.32(td,J=13.2,3.4Hz,1H),3.25-3.21(m,1H),3.14-3.09(m,1H),2.79(qd,J=12.3,4.9Hz,1H),1.87-1.80(m,2H),1.72(s,3H),1.72-1.62(m,6H),1.51-1.44(m,1H),1.39-1.22(m,5H),1.27(s,3H),1.03(td,J=13.2,3.0Hz,1H),0.93(s,3H),0.77(s,3H).^{13}C NMR(151MHz,BENZENE-d_6)δppm 137.10,123.6,78.0,56.79,43.05,42.11,37.74,35.0 33.46,29.20,27.29,24.57,24.29,23.48,22.59,22.33,20.27

[1020] 实施例6-使用野生型SHC酶将EEH转化为(-)-降龙涎香醚

[1021] 选择下表8中列出的野生型SHC酶用于测试它们将E,E-均法呢醇转化为(-)-降龙涎香醚的能力。

[1022] 表8.野生型角鲨烯藿烯环化酶。

[1023]

角鲨烯藿烯环化酶	来源(微生物)	欧洲核苷酸档案编号
AacSHC (SEQ ID NO: 1)	酸热脂环酸杆菌	ACV59449.1
ApaSHC1 (SEQ ID NO: 20)	巴氏醋杆菌	ASC07046.1
BjaSHC (SEQ ID NO: 13)	日本慢生根瘤菌	ABQ33590.1
TelSHC (SEQ ID NO: 19)	细长嗜热聚球藻菌	BAC09861.1
ZmoSHC1 (SEQ ID NO: 11)	运动发酵单胞菌	AVZ42714.1
ZmoSHC2 (SEQ ID NO: 12)	运动发酵单胞菌	AAF12829.1

[1024]

GmoSHC (SEQ ID NO: 14)	莫比氏葡萄糖杆菌	EHH69691.1

[1025] 生物催化剂是从用携带如上文实施例1所述的用于产生野生型酶的基因的pET-28a(+)转化的大肠杆菌菌株产生的(小规模生物催化剂产生)。

[1026] 对于产生的每种生物催化剂,粗略地研究了pH、温度和表面活性剂浓度条件的影响。第一步,进行pH优化,然后进行[SDS]和温度优化。使用定义为最佳的pH值附近的0.1M琥珀酸/NaOH缓冲液在另一组反应(可变pH、温度和SDS浓度)中确认优化的反应条件:

[1027] *pH优化:4g/l EEH与0.075%SDS,10的OD_{650nm}的细胞,在pH5.0至7.0的柠檬酸/磷酸钠缓冲液(1ml总体积,40℃,900rpm,Heidolph Synthesis 1)中。

[1028] *[SDS]和温度优化:4g/l EEH,10的OD_{650nm}的细胞,在柠檬酸/磷酸钠缓冲液(1ml总体积,900rpm,Heidolph Synthesis 1)中,最佳pH值下不同[SDS]和温度。

[1029] *反应条件的确认:4g/L EEH,10的OD_{650nm}的细胞,柠檬酸/磷酸钠缓冲液(1ml总体积,900rpm,Heidolph Synthesis 1)中,根据需要的[SDS]、温度和pH值。

[1030] 反应进行24小时,随时间取样(t=0,t=3小时和t=22小时),提取并分析它们的底物和产物含量,并如上所述计算EEH的转化率。这允许大致评估使用每种野生型SHC酶的EEH转化的初始速度以及反应22小时后获得的转化率。使用4g/l EEH和加载到10.0的OD_{650nm}的生物催化剂下,不同的pH、温度和SDS浓度下进行反应。

[1031] 结果

[1032] 下面的表9总结了对于每种测试的野生型SHC酶确定为最佳的条件以及在应用初始(35℃,pH5.4,0.070％SDS)或单独优化的反应条件时在反应的22小时获得的EEH转化率。在所有这些酶中都观察到EEH转化为(-)-降龙涎香醚,尽管水平显著不同。当大肠杆菌细胞产生日本慢生根瘤菌SHC和运动发酵单胞菌SHC1时,EEH转化率最高。非常有趣的是,运动发酵单胞菌SHC2也获得了EEH转化,据其他地方报道,这种酶不会使EEH环化。

[1033] 表9.野生型角鲨烯藿烯环化酶的优化反应条件。

[1034]

角鲨烯藿烯环化酶	优化的反应条件			22h 时的 EEH 转化率
	[SDS] (%)	T (°C)	pH	(%)
AacSHC (SEQ ID NO: 1)	0.030	55	6.0	40.0
ApaSHC1 (SEQ ID NO: 20)	0.0075	45	5.2	9.2
BjaSHC (SEQ ID NO: 13)	0.0050	50	5.8	16.5
TelSHC (SEQ ID NO: 19)	0.0075	45	6.2-6.6	1.5
ZmoSHC1 (SEQ ID NO: 11)	0.005	40	5.0-5.4	70.3
ZmoSHC2 (SEQ ID NO: 12)	0.005	35	6.2-6.4	3.2
GmoSHC (SEQ ID NO: 14)	0.0075	40	5.6	12.2

[1035] 应用单独优化的(T,pH,[SDS])反应条件,在4g/l EEH和在10.0的OD_{650nm}下加载的生物催化剂下运行的反应中获得的EEH转化率。

[1036] 实施例7-野生型SHC的序列比对

[1037] 使用下面指定的参数(使用Clone Manager 9软件)比对各种野生型SHC的序列。结果显示在下表10和图8中。

[1038] 比对:针对参考分子的全局蛋白质比对

[1039] 参数:评分矩阵:BLOSUM 62

[1040] 参考分子:Aac SHC,区域1至631

[1041] 有缺口的比对序列总长度:838aa

[1042] 表10.相对于WT AacSHC的SHC序列同一性百分比

[1043]

序列	开始	结束	匹配	不匹配	%匹配
AacSHC (SEQ ID NO: 1)	1	631	-	-	-
TelSHC (SEQ ID NO: 19)	1	642	285	360	44
ApaSHC1 (SEQ ID NO: 20)	1	720	296	430	40
ZmoSHC1 (SEQ ID NO: 11)	1	725	299	434	40
ZmoSHC2 (SEQ ID NO: 12)	1	658	251	410	37
BjaSHC (SEQ ID NO: 13)	1	684	296	397	42
GmoSHC (SEQ ID NO: 14)	1	685	291	396	42

[1044] -:不适用

[1045] 实施例8-SHC变体的最佳反应条件

[1046] 方法

[1047] 将额外的突变引入产生215G2 SHC的变体[215G2+V174I]、[215G2+V174I+F601]、[215G2+L37Q]和[215G2+L37Q+V174I]。这些变体的序列比对如图8所示。

[1048] 如以上实施例中所述研究这些变体的最佳反应条件。然后使用以10.0的OD_{650nm}加载到反应中的细胞在4g/l底物下测试它们在EEH环化中的性能。

[1049] 结果

[1050] 下表11中列出的反应条件(温度、pH、SDS浓度)似乎是上述变体的环化活性的个体最佳反应条件。具有新突变的变体显示出与215G2 SHC大致相同的活性,或增加的活性,其可能与增加的SHC产生水平相关。

[1051] 表11.SHC变体酶的最佳反应条件。

[1052]

SHC/HAC 酶变体	最佳温度(°C)	最佳 pH	最佳[SDS] (重量/重量%)
215G2 + L37Q (SEQ ID NO: 24)	40	5.8	0.070
215G2 + V174I (SEQ ID NO: 25)	50	5.2-5.4	0.040
215G2 + V174I + F601Y (SEQ ID NO: 26)	40-50	5.2-5.6	0.030-0.070
215G2 + L37Q + V174I + F601Y (SEQ ID NO: 27)	35	5.4-5.8	0.070

[1053] 实施例9-反应产物

[1054] 方法

[1055] 5ml体积的反应用4g/l EEH和以10.0的OD_{650nm}加载的生物催化剂在对于每种酶定义为最佳的pH、温度和SDS浓度下和恒定搅拌(Heidolph synthesis 1Liquid装置,900rpm)下进行。

[1056] 结果

[1057] 表12说明了每种酶在20小时后的EEH转化率,以及所得的反应产物分布。式(I)((-)-降龙涎香醚)和(IV)的化合物由EEH环化产生,式(II)(大环化合物)和(III)(9b-epi-降龙涎香醚)的化合物由EZH环化产生。使用的均法呢醇是异构体的混合物(EEH:EZH 80:20)。

[1058] 观察到与WT AacSHC相比,所有WT SHC酶和特别是转化水平为约9%或更高的WT ZmoSHC1、BjaSHC、GmoSHC、ApaSHC1显示对EEH的更高选择性。与化合物(II)和(III)相比,这导致反应产物中化合物(I)和(IV)的更高比例,并且还转化为更高的EEH:EZH转化率。

[1059] 表12.使用野生型(WT)SHC酶的EEH/EZH混合物的转化。

[1060]

SHC 酶	EEH 转化 (%)	反应产物 (总反应产物的%)					
		I	II	III	IV	I + IV	II + III
AacSHC	40.4	93.5	1.0	4.8	0.7	94.2	5.8
TelSHC1	1.6	94.8	2.0	1.7	1.4	96.2	3.8
ApaSHC1	9.3	98.6	0.3	0.9	0.2	98.8	1.2
ZmoSHC1	70.3	98.1	0.6	1.1	0.2	98.3	1.7
ZmoSHC2	3.2	98.5	0.3	0.9	0.4	98.8	1.2
BjaSHC	16.2	98.0	0.5	1.0	0.5	98.5	1.5
GmoSHC	11.9	98.3	0.5	0.8	0.4	98.7	1.3

[1061] 实施例10-容积生产率

[1062] 新的酸热脂环酸杆菌SHC变体用于在如实施例4所述以125g/l E,E-均法呢醇和250g/l细胞运行的生物转化中改善E,E-均法呢醇到(-)-降龙涎香醚的环化。解决了在125g/l下使用250g/l细胞进行的生物转化中完全EEH转化的反应时间的减少。

[1063] 方法

[1064] 如实施例1(发酵)中所述产生生物催化剂。生物转化的运行如实施例4中所述,应用针对每个变体单独优化的温度和pH条件,并如实施例2中所述,唯一的例外是对于所有变体将SDS浓度设置为1.3w/v%([SDS]:[细胞]比率为0.052)。

[1065] 结果

[1066] 取决于所考虑的SHC/HAC变体,在250g/l细胞的情况下将125g/l E,E-均法呢醇完全转化为(-)-降龙涎香醚所需的时间可减少1/3或2/3,大约在48或24小时中实现了完全转化,而不是使用参考SHC生物催化剂215G2 SHC的大约72小时(图12)。

[1067] 与参考生物催化剂215G2 SHC相比,新SHC变体的初始反应速度提高了1.4-1.8倍。

[1068] 定义为每升和每小时产生的降龙涎香醚或每升、每小时和每克生物催化剂产生的降龙涎香醚的容积生产率与使用215G2 SHC生物催化剂的参考生物转化运行相比增加了1.4-2.6倍(表13)。

[1069] 表13.容积生产率。

[1070]

SHC 变体	[EEH] (g/l)	[细胞] (g/l)	时间 (h)	转化率 (%)	生产率 ($g.l^{-1}.h^{-1}$)	生产率 ($mg.l^{-1}.h^{-1}.g^{-1}$)	增加 (-倍)
215G2	125	250	68.5	95.9	1.75	7.0	-
#49	125	250	26.5	98.2	4.65	18.5	2.6
#65	125	250	30.5	98.7	4.10	16.2	2.3
#66	125	250	50	98.1	2.45	9.8	1.4
110B8	125	250	46	96.8	2.65	10.5	1.5

[1071] 转化率:EEH转化率。

[1072] 实施例11-容积生产率

[1073] 新的酸热脂环酸杆菌SHC变体用于在如实施例4所述以125g/l E,E-均法呢醇和250g/l细胞运行的生物转化中改善E,E-均法呢醇到(-)-降龙涎香醚的环化。解决了在125g/l底物下运行的生物转化中完全EEH转化的生物催化剂加载量的减少。

[1074] 方法

[1075] 如实施例1(发酵)中所述产生生物催化剂。生物转化的运行如实施例4中所述,应用针对每个变体单独优化的温度和pH条件,并如实施例2中所述,唯一的例外是对于所有变体将细胞浓度设置为125g/l和SDS浓度设置为0.65w/v%([SDS]:[细胞]比率为0.052)。

[1076] 结果

[1077] 使用产生新的SHC/HAC变体SHC#49和SHC#65的生物催化剂,对于在72小时内的完全EEH转化,可以将细胞浓度降低2倍,从250到125g/l细胞(图13)。在相同时间,使用参考生物催化剂215G2 SHC的EEH转化率仅为约67%。

[1078] 定义为每升和每小时产生的(-)-降龙涎香醚或每升、每小时和每克生物催化剂产生的(-)-降龙涎香醚的容积生产率与使用215G2 SHC生物催化剂作为参考的相同生物转化运行相比增加了约1.5倍(表14)。对于使用参考生物催化剂215G2 SHC的使用125g/l和250g/l细胞的生物转化,容积生产率提高了约2倍(表14)。

[1079] 表14.容积生产率。

[1080]

SHC 变体	[EEH] (g/l)	[细胞] (g/l)	时间 (h)	转化率 (%)	生产率 ($g.l^{-1}.h^{-1}$)	生产率 ($mg.l^{-1}.h^{-1}.g^{-1}$)	增加 (-倍)	
215G2	125	125	72	66.7	1.16	9.3	-	n.a.
215G2	125	250	68.5	95.9	1.75	7.0	n.a.	-
#49	125	125	72	95.0	1.65	13.2	1.4	1.9
#65	125	125	72	98.5	1.70	13.7	1.5	2.0

[1081] 转化率:EEH转化率。n.a:不适用。

[1082] 实施例12-容积生产率

[1083] 新的酸热脂环酸杆菌SHC变体用于在如实施例4所述以125g/l E,E-均法呢醇和250g/l细胞运行的生物转化中改善E,E-均法呢醇到(-)-降龙涎香醚的环化。解决了生物转化中的EEH浓度的增加。

[1084] 方法

[1085] 如实施例1(发酵)中所述产生生物催化剂。生物转化的运行如实施例4中所述,但EEH与生物催化剂的比例为1:1,并且恒定[SDS]:[细胞]比例为0.052。优化的温度和pH条件定义为对于SHC#65生物催化剂是最佳的(实施例2)。

[1086] 结果

[1087] 使用新的SHC/HAC变体SHC#65,可以在48小时内用300g/l生物催化剂完全转化高达300g/l的EEH(图14)。

[1088] 定义为每升和每小时产生的降龙涎香醚的容积生产率增加了2.3-3.4倍。定义为每升、每小时和每克生物催化剂产生的降龙涎香醚的容积生产率增加了2.8-2.9倍。倍数增加是指使用215G2 SHC生物催化剂的在125g/l EEH、使用250g/l细胞运行的反应(表15)。

[1089] 表15.容积生产率。

[1090]

SHC 变体	[EEH] (g/l)	[细胞] (g/l)	时间 (h)	转化率 (%)	生产率		增加 (-倍)
					($g.l^{-1}.h^{-1}$)	($mg.l^{-1}.h^{-1}.g^{-1}$)	
215G2	125	250	68.5	95.9	1.75	7.0	-/-
#65	125	250	25.5	99.2	4.9	19.5	2.8/2.8
#65	200	200	48	97.7	4.1	20.5	2.3/2.9
#65	250	250	48	93.8	4.9	19.5	2.8/2.8
#65	300	300	48	94.4	5.9	20.0	3.4/2.8

[1091] 转化率:EEH转化率

[1092] 实施例13-使用wt SHC酶的EEH环化

[1093] 方法

[1094] 野生型SHC酶AacSHC(酸热脂环酸杆菌SHC,SEQ ID NO:1)、GmoSHC(莫比氏葡萄糖杆菌SHC,SEQ ID NO:14)、BjpSHC(日本慢生根瘤菌SHC,SEQ ID NO:13)、ApaSHC1和ApaSHCA(巴氏醋杆菌SHC1和SHCA,分别为SEQ ID NO:20和30)、ZmoSHC1和ZmoSHC2(运动发酵单胞菌SHC1和SHC2,分别为SEQ ID NO:11和12)、SalSHC(白色链霉菌SHC,SEQ ID NO:29)、TelSHC(细长嗜热聚球藻菌,SEQ ID NO:19)和BmeSHC巨大芽孢杆菌SHC,SEQ ID NO:28)的生物催化剂在如实施例1中概述的发酵中产生。

[1095] 选择这些野生型SHC酶(也在下表16中列出)用于测试它们将E,E-均法呢醇转化为(-)-降龙涎香醚的能力。

[1096] 表16.野生型角鲨烯藿烯环化酶。

[1097]

角鲨烯藿烯环化酶	来源(微生物)	欧洲核苷酸档案号或 Genebank 登录号

[1098]

AacSHC (SEQ ID NO: 1)	酸热脂环酸杆菌	ACV59449.1
ApaSHC1 (SEQ ID NO: 20)	巴氏醋杆菌	ASC07046.1
BjaSHC (SEQ ID NO: 13)	日本慢生根瘤菌	ABQ33590.1
TelSHC (SEQ ID NO: 19)	细长嗜热聚球藻菌	BAC09861.1
ZmoSHC1 (SEQ ID NO: 11)	运动发酵单胞菌	AVZ42714.1
ZmoSHC2 (SEQ ID NO: 12)	运动发酵单胞菌	AAF12829.1
GmoSHC (SEQ ID NO: 14)	莫比氏葡萄糖杆菌	EHH69691.1
BmeSHC (SEQ ID NO: 28)	巨大芽孢杆菌	WP_016763969
Sal SHC (SEQ ID NO: 29)	白色链霉菌	AZN28579
ApaSHCA (SEQ ID NO: 30)	巴氏醋杆菌	WP_003625617

[1099] 用如实施例6中所述的BmeSHC生物催化剂研究了其他SHC酶的最佳反应条件。最佳反应条件定义为对于反应中10的OD_{650nm}的细胞加载为45℃、pH5.6和0.0025w/v%SDS。

[1100] 使用SalSHC的最佳反应条件为30℃,pH7.0,反应中不添加表面活性剂(Liu等人(2020):A Novel Soluble Squalene-Hopene Cyclase and Its Application in Efficient Synthesis of Hopene,Frontiers in Bioengineering and Biotechnology, vol 8,article 426,https://doi.org/10.3389/fbioe.2020.00426)。

[1101] 在4g/l EEH和125g/l EEH下测试所得生物催化剂的活性,唯一的例外是SalSHC,其中使用1g/l EEH而不是4g/l EEH。

[1102] 使用4g/l EEH(SalSHC:1g/l EEH)的活性测试包含10的OD_{650nm}的细胞(SalSHC生物催化剂:OD_{650nm}为40)并且应用针对每种SHC酶单独优化的温度和pH条件在其他方面如实施例6所述运行。

[1103] 125g/l EEH的生物转化使用250g/l细胞应用针对每种SHC酶单独优化的温度和pH条件如实施例4中所述并如实施例6中所述运行。反应在AacSHC的情况下含有1.3w/v%SDS,在BjpSHC、ZmoSHC1、ApaSHC1的情况下为0.16w/v%,在BmeSHC的情况下为0.10w/v%,和在GmoSHC的情况下为0.14w/v%。没有生物转化在125g/l EEH下使用ApaSHCA、TelSHC、SalSHC和ZmoSHC2生物催化剂运行。

[1104] 结果

[1105] 在低底物浓度(1或4g/l)下测试的所有SHC生物催化剂均观察到EEH环化为(-)-降龙涎香醚。野生型SHC生物催化剂在125g/l EEH下测试,EEH转化率为约15%至约75%(图15)。

[1106] 实施例14-野生型SHC酶的序列比对

[1107] SHC酶相对于WT AacSHC的序列同一性是使用下面指定的参数的Clone Manager 9软件或者使用默认参数的blastp确定的。结果分别总结在表17A和17B中。

[1108] 比对:针对参考分子的全局蛋白质比对

[1109] 参数:评分矩阵:BLOSUM 62

[1110] 参考分子:Aac SHC,区域1至631

[1111] 有缺口的比对序列总长度:838aa

[1112] 表17A.相对于WT AacSHC的SHC序列同一性百分比。

[1113]

序列	开始	结束	匹配	不匹配	%匹配
AacSHC (SEQ ID NO: 1)	1	631	-	-	-
TelSHC (SEQ ID NO: 19)	1	642	285	360	44
ApaSHC1 (SEQ ID	1	720	296	430	40

[1114]

NO: 20)					
ZmoSHC1 (SEQ ID NO: 11)	1	725	299	434	40
ZmoSHC2 (SEQ ID NO: 12)	1	658	251	410	37
BjaSHC (SEQ ID NO: 13)	1	684	296	397	42
GmoSHC (SEQ ID NO: 14)	1	685	291	396	42
BmeSHC (SEQ ID NO: 28)	1	625	194	443	30
SalSHC (SEQ ID NO: 29)	1	691	327	367	47
ApaSHCA (SEQ ID NO: 30)	1	656	259	403	39

[1115] -:不适用。使用Clone Manager 9软件确定序列同一性。

[1116] 表17B.相对于WT AacSHC的SHC序列同一性百分比。

[1117]

SHC 酶/SEQ ID	%同一性
AacSHC (SEQ ID NO: 1)	-
TelSHC (SEQ ID NO: 19)	45.3
ApaSHC1 (SEQ ID NO: 20)	43.5
ZmoSHC1 (SEQ ID NO: 11)	45.3
ZmoSHC2 (SEQ ID NO: 12)	39.2
BjaSHC (SEQ ID NO: 13)	44.3
GmoSHC (SEQ ID NO: 14)	43.9
BmeSHC (SEQ ID NO: 28)	31.4
SalSHC (SEQ ID NO: 29)	51.8
ApaSHCA (SEQ ID NO: 30)	39.9

[1118]

[1119] -:不适用。用blastp确定的序列同一性。

[1120] 实施例15-野生型SHC的序列比对

[1121] 使用Clustal O(1.2.4)比对表16-18中公开的野生型SHC序列。图9A和9B中提供了多序列比对,表19中提供了序列之间的同一性百分比(如使用blastp确定的)。

[1122] 表18.用blastp计算的序列同一性矩阵。

[1123]

% 同一性	AacSHC	BjpSHC	BmeSHC	GmoSHC	ApaSHCA	ApaSHC1	SalSHC	TelSHC	ZmoSHC1	ZmoSHC2
AacSHC	100									
BjpSHC	44.3	100								
BmeSHC	31.4	28.1	100							
GmoSHC	43.9	71.6	30.5	100						
ApaSHCA	39.9	36.8	30.7	37.1	100					
ApaSHC1	43.5	69.4	29.7	75.8	37.2	100				
SalSHC	51.8	43.4	31.3	42.6	39.3	42.4	100			
TelSHC	45.3	39.3	30.4	39.1	40.5	39.2	44.4	100		
ZmoSHC1	45.3	69.6	30.8	76.7	36.8	76.8	45	39.5	100	
ZmoSHC2	39.2	36.8	32.1	35.3	55.1	35.2	42	41.4	36.8	100

[1124]

[1125] 实施例16-培养条件对生物催化剂比活性的影响。

[1126] 方法

[1127] 作为实施例1中概述的生物催化剂产生方法(摇瓶培养和发酵)的替代方案,测试了使用自诱导培养基(AIM)产生生物催化剂。使用包含微量元素的AIM-TB肉汤基质(Formedium)制备培养基。在100ml带挡板的摇瓶中的25ml AIM-TB富含1%甘油。以2E-07的OD_{650nm}接种培养物;在持续摇动(220rpm)下,在37℃下培养持续18±1小时。

[1128] 收集的生物催化剂的活性在标准活性测试中在4g/l E,E-均法呢醇、10的OD_{650nm}的细胞下进行测试,如实施例2所述,并取决于所考虑的SHC酶,应用在其他方面如实施例2和6所述的最佳温度、pH和[SDS]。

[1129] 结果

[1130] 所考虑的SHC生物催化剂的活性取决于培养基和用于生物催化剂产生的条件而有很大差异(表19)。以同样的方式,预期使用替代或改进的发酵方案的生物催化剂生产将产生比应用实施例1中概述的发酵方案产生时具有更低或更高比活性的生物催化剂。

[1131] 表19.培养方法对EEH转化率的影响。

[1132]

SHC 生物催化剂	24 小时的 EEH 转化率(%)		
	AIM-TB 摇瓶	基本培养基 摇瓶	基本培养基 发酵罐
SHC 65	41-70	n.d.	100
ZmoSHC1	47-57	60-70	91
BjpSHC	11-26	14-17	90
ApaSHC1	24-35	4-9	94

[1133] n.d.:未确定

[1134] 实施例17-[SDS]:[细胞]比率对EEH转化的影响

[1135] 方法

[1136] 如实施例1(发酵)中所述从产生SHC变体酶SHC#65的大肠杆菌菌株产生生物催化剂。EEH生物转化如实施例4所述进行,但使用125g/l EEH和150g/l细胞。反应在针对SHC#65定义为最佳的pH5.6和45℃下进行。反应中的SDS在0.060、0.066、0.072、0.078、0.084和0.090%之间变化,导致分别为0.040、0.044、0.048、0.052、0.056和0.060的[SDS]:[细胞]比率。

[1137] 结果

[1138] EEH转化率在所有反应中>85%,在[SDS]:[细胞]比率范围为0.044-0.056时获得最大转化率。在[SDS]:[细胞]比率为0.040和0.060时,转化率略低(图16)。

[1139] 前述内容广泛地描述了而非限制本发明的某些实施方案。对本领域技术人员来说明显的变化和修改旨在落入如所附权利要求中限定和由所附权利要求限定的本发明的范围内。

[1140] 序列表

[1141] SEQ ID NO:1(野生型AacSHC的氨基酸序列)

[1142] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHIPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLGYTMYRHVFPTLALGRYKQAIERR

[1143] SEQ ID NO:2(SHC/HAC酶变体#49的氨基酸序列)

[1144] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDR

VLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNLTPFCDFGEVTDPPSEDVTAHVLE
CFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTTWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLG
YTMYRHVFPTLALGRYKQAIERR

[1145] SEQ ID NO:3(SHC/HAC酶变体#65的氨基酸序列)

[1146] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMPLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDR
VLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLE
CFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTTWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLG
YTMYSHVFPTLALGRYKQAIERR

[1147] SEQ ID NO:4(SHC/HAC酶变体#66的氨基酸序列)

[1148] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALHPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMPLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDR
VLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLE
CFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTTWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLG
YTMYSHVFPTLALGRYKQAIERR

[1149] SEQ ID NO:5(SHC/HAC酶变体#110B8的氨基酸序列)

[1150] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALHPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMPLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDR
VLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNLTPFCDFGEVTDPPSEDVTAHVLE
CFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTTWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLG
YTMYRHVFPTLALGRYKQAIERR

[1151] SEQ ID NO:6(SHC/HAC酶变体#49的核苷酸序列)

[1152] ATGGCTGAGCAGTTGGTGGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT

GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTCGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTCAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGTTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGA
TCATGTTCCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTTCCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGTGCTGCACGGGTATC
AGAAGCTGTCGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTTGGGGCGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCTCACCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGTGTTTCGGCAGCTTCG
GGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATCGACAC
GCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGCTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGACCTGGGCGCTGATGGCG
CTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGCTGGGATGAGCCGTACTACACCGGCACGGGCTTCCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

[1153] SEQ ID NO:7(SHC/HAC酶变体#65的核苷酸序列)

[1154] ATGGCTGAGCAGTTGGTGGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTCGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTCAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGTTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGA
TCATGTTCCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTTCCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGTGCTGCACGGGTATC
AGAAGCTGTCGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT

GGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTTGGGGCGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACACCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGTGTTTCGGCAGCTTCG
GGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATCGACAC
GCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGCTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGACCTGGGCGCTGATGGCG
CTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGCTGGGATGAGCCGTACTACACCGGCACGGGCTTCCCAGGGGATTTCTACCTCGGCTACACCATGTACAGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

[1155] SEQ ID NO:8(SHC/HAC酶变体#66的核苷酸序列)ATGGCTGAGCAGTTGGTGGAAGCTCCG
GCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTCTCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCC
GCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTTGTGCCACATTCTCGATCGCGTCGATCGGGATCGCA
TGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGGACGGCACGTGGGCCCTGCACCCGGGTGGGCCGCCG
GACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTATATCGGCATGTCGCGCGACGAGGAGCCGATGCAGAA
GGCGCTCCGGTTCATTCAGAGCCAGGGCGGGATCGAGTCGTCGCGCGTGTTCACGCGGAGGTGGCTGGCGCTGGTGG
GAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGATCATGTTCCTCGGCAAGCGCATGCCGCTCAACATC
TACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTCTCGATTGTGATGAGCCGCCAGCCGGTGTTCCCGCT
GCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGAGACCGACGTGCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTG
GGTGGATCTTCGACGCGCTCGACCGGGTGCTGCACGGGTATCAGAAGCTGTCGGTGCACCCGTTCCGCCGCGCGGCC
GAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGAGACGGCAGCTGGGGCGGGATTCAGCCGCCTTGGTT
TTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCATCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGC
TGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGGCTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCC
GTGCTCGCGCTGCGCGCTGCGGGGCTTCCGGCCGATCACGACCGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCG
GCAGATCACGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAACCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACA
ACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGTGGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGC
CGCAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTCGGCATGCAGAGCTCGAACGGCGGTTGGGGCGCCTA
CGACGTCGACAACACGAGCGATCTCCCGAACCACACCCCGTTCTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAG
AGGACGTCACCGCCCACGTGCTCGAGTGTTTCGGCAGCTTCGGGTACGATGACGCCTGGAAGGTCATCCGGCGCGCG
GTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGGTTCGGTCGTTGGGGCGTCAATTACCTCTACGGCAC
GGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATCGACACGCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGG
TCGAGCAGCATCAGAACCCGGACGGCGGCTGGGGCGAGGACTGCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAG
GGCGCGAGCACCCCGTCGCAGACGACCTGGGCGCTGATGGCGCTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGC
GCGCCGCGGCGTGCAATACCTCGTGGAGACGCAGCGCCCGGACGGCGGCTGGGATGAGCCGTACTACACCGGCACGG
GCTTCCCAGGGGATTTCTACCTCGGCTACACCATGTACAGCCACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAG
CAAGCCATCGAGCGCAGGTGA

[1156] SEQ ID NO:9(SHC/HAC酶变体#110B8的核苷酸序列)

[1157] ATGGCTGAGCAGTTGGTGGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT

GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGCACCCGGGTGGGCCGCCGGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTCGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTCAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGTTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGA
TCATGTTCCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTTCCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGTGCTGCACGGGTATC
AGAAGCTGTCGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTTGGGGCGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCTCACCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGTGTTTCGGCAGCTTCG
GGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATCGACAC
GCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGCTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGACCTGGGCGCTGATGGCG
CTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGCTGGGATGAGCCGTACTACACCGGCACGGGCTTCCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

[1158] SEQ ID NO:10(Aac215G2SHC/HAC酶变体的氨基酸序列)

[1159] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMPLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDR
VLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLE
CFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLG
YTMYRHVFPTLALGRYKQAIERR

[1160] SEQ ID NO:11(野生型ZmoSHC1的氨基酸序列)

[1161] MGIDRMNSLSRLLMKKIFGAEKTSYKPASDTIIGTDTLKRPNRRPEPTAKVDKTIFKTMGNSLNNTLVS
ACDWLIGQQKPDGHWVGAVESNASMEAEWCLALWFLGLEDHPLRPRLGNALLEMQREDGSWGVYFGAGNGDINATVE
AYAALRSLGYSADNPVLKKAAAWIAEKGGLKNIRVFTRYWLALIGEWPWEKTPNLPPEIIWFPDNFVFSIYNFAQWA
RATMVPIAILSARRPSRPLRPQDRLDELFPEGRARFDYELPKKEGIDLWSQFFRTTDRGLHWVQSNLLKRNSLREAA
IRHVLEWIIRHQDADGGWGGIQPPWVYGLMALHGEGYQLYHPVMAKALSALDDPGWRHDRGESSWIQATNSPVWDTM

LALMALKDAKAEDRFTPEMDKAADWLLARQVKVKGDWSIKLPDVEPGGWAFEYANDRYPDTDDTAVALIALSSYRDK
EEWQKKGVEDAITRGVNWLIAMQSECGGWGAFDKDNNRSILSKIPFCDFGESIDPPSVDVTAHVLEAFGTLGLSRDM
PVIQKAIDYVRSEQEAEGAWFGRWGVNYIYGTGAVLPALAAIGEDMTQPYITKACDWLVAHQQEDGGWGESCSSYME

[1162] SEQ ID NO:12(野生型ZmoSHC2的氨基酸序列)

[1163] MTVSTSSAFHHSPLSDDVEPIIQKATRALLEKQQQDGHWVFELEADATIPAEYILLKHYLGEPEDLEIE
AKIGRYLRRIQGEHGGWSLFYGGDLDLSATVKAYFALKMIGDSPDAPHMLRARNEILARGGAMRANVFTRIQLALFG
AMSWEHVPQMPVELMLMPEWFPVHINKMAYWARTVLVPLLVLQALKPVARNRRGILVDELFVPDVLPTLQESGDPIW
RRFFSALDKVLHKVEPYWPKNMRAKAIHSCVHFVTERLNGEDGLGAIYPAIANSVMMYDALGYPENHPERAIARRAV
EKLMVLDGTEDQGDKEVYCQPCLSPIWDTALVAHAMLEVGGDEAEKSAISALSWLKPQQILDVKGDWAWRRPDLRPG
GWAFQYRNDYYPDVDDTAVVTMAMDRAAKLSDLHDDFEESKARAMEWTIGMQSDNGGWGAFDANNSYTYLNNIPFAD
HGALLDPPTVDVSARCVSMMAQAGISITDPKMKAAVDYLLKEQEEDGSWFGRWGVNYIYGTWSALCALNVAALPHDH
LAVQKAVAWLKTIQNEDGGWGENCDSYALDYSGYEPMDSTASQTAWALLGLMAVGEANSEAVTKGINWLAQNQDEEG
LWKEDYYSGGGFPRVFYLRYHGYSKYFPLWALARYRNLKKANQPIVHYGM

[1164] SEQ ID NO:13(野生型BjpSHC的氨基酸序列)

[1165] MTVTSSASARATRDPGNYQTALQSTVRAAADWLIANQKPDGHWVGRAESNACMEAQWCLALWFMGLEDH
PLRKRLGQSLLDSQRPDGAWQVYFGAPNGDINATVEAYAALRSLGFRDDEPAVRRAREWIEAKGGLRNIRVFTRYWL
ALIGEWPWEKTPNIPPEVIWFPLWFPFSIYNFAQWARATLMPIAVLSARRPSRPLPPENRLDALFPHGRKAFDYELP
VKAGAGGWDRFFRGADKVLHKLQNLGNRLNLGLFRPAATSRVLEWMIRHQDFDGAWGGIQPPWIYGLMALYAEGYPL
NHPVLAKGLDALNDPGWRVDVGDATYIQATNSPVWDTILTLLAFDDAGVLGDYPEAVDKAVDWVLQRQVRVPGDWSM
KLPHVKPGGWAFEYANNYYPDTDDTAVALIALAPLRHDPKWKAKGIDEAIQLGVDWLIGMQSQGGGWGAFDKDNNQK
ILTKIPFCDYGEALDPPSVDVTAHIIEAFGKLGISRNHPSMVQALDYIRREQEPSGPWFGRWGVNYVYGTGAVLPAL
AAIGEDMTQPYIGRACDWLVAHQQADGGWGESCASYMDVSAVGRGTTTASQTAWALMALLAANRPQDKDAIERGCMW
LVERQSAGTWDEPEFTGTGFPGYGVGQTIKLNDPALSQRLMQGPELSRAFMLRYGMYRHYFPLMALGRALRPQSHS

[1166] SEQ ID NO:14(野生型GmoSHC的氨基酸序列)

[1167] MSPADISTKSSSFQRLDNMLPEAVSSACDWLIDQQKPDGHWVGPVESNACMEAQWCLALWFLGQEDHPL
RPRLAQALLEMQREDGSWGIYVGADHGDINTTVEAYAALRSMGYAADMPIMAKSAAWIQQKGGLRNVRVFTRYWLAL
IGEWPWDKTPNLPPEIIWLPDNFIFSIYNFAQWARATMMPLTILSARRPSRPLLPENRLDGLFPEGRENFDYELPVK
GEEDLWGRFFRAADKGLHSLQSFPVRRFVPREAAIRHVIEWIIRHQDADGGWGGIQPPWIYGLMALSVEGYPLHHPV
LAKAMDALNDPGWRRDKGDASWIQATNSPVWDTMLAVLALHDAGAEDRYSPQMDKAIGWLLDRQVRVKGDWSIKLPD
TEPGGWAFEYANDKYPDTDDTAVALIALAGCRHRPEWRERDIEGAISRGVNWLLAMQSSSGGWGAFDKDNNRSILTK
IPFCDFGEALDPPSVDVTAHVLEAFGLLGISRNHPSVQKALAYIRSEQERNGAWFGRWGVNYVYGTGAVLPALAAIG
EDMTQPYIVRACDWLMSVQQENGGWGESCASYMDINAVGHGVATASQTAWALIGLLAAKRPKDREAIARGCQFLIER
QEDGSWTEEEYTGTGFPGYGVGQAIKLDDPSLPDRLLQGAELSRAFMLRYDLYRQYFPVMALSRARRMMKEDASAAA

[1168] SEQ ID NO:15(编码野生型AacSHC的多肽的核苷酸序列)

[1169] ATGGCTGAGCAGTTGGTGGAAGCGCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTCGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTCAGAGCCAGGGCGGGATCGAGTCGTC

GCGCGTGTTCACGCGGATGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGA
TCATGTTCCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTTCCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGCGCTGCACGGGTATC
AGAAGCTGTCGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTTGGGGCGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACATCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGTGTTTCGGCAGCTTCG
GGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATCGACAC
GCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGCTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGGCCTGGGCGCTGATGGCG
CTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGCTGGGATGAGCCGTACTACACCGGCACGGGCTTCCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

[1170] SEQ ID NO:16(编码Aac215G2SHC/HAC酶变体的核苷酸序列)

[1171] ATGGCTGAGCAGTTGGTGGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTCGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTCAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGTTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGA
TCATGTTCCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTTCCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGTGCTGCACGGGTATC
AGAAGCTGTCGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTTGGGGCGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACACCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGTGTTTCGGCAGCTTCG

GGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATCGACAC
GCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGCTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGGCCTGGGCGCTGATGGCG
CTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGCTGGGATGAGCCGTACTACACCGGCACGGGCTTCCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

[1172] SEQ ID NO:17(SHC/HAC酶变体#90C7的氨基酸序列)

[1173] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDATIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMPLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDR
VLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLE
CFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLG
YTMYSHVFPTLALGRYKQAIERR

[1174] SEQ ID NO:18(SHC/HAC酶变体#115A7的氨基酸序列)

[1175] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMPLNIYEFGSWARTTVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDR
VLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDKTQHPAFIKGWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLE
CFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLG
YTMYRHVFPTLALGRYKQAIERR

[1176] SEQ ID NO:19(野生型TelSHC的氨基酸序列)

[1177] MPTSLATAIDPKQLQQAIRASQDFLFSQQYAEGYWWAELESNVTMTAEVILLHKIWGTEQRLPLAKAEQ
YLRNHQRDHGGWELFYGDGGDLSTSVEAYMGLRLLGVPETDPALVKARQFILARGGISKTRIFTKLHLALIGCYDWR
GIPSLPPWIMLLPEGSPFTIYEMSSWARSSTVPLLIVMDRKPVYGMDPPITLDELYSEGRANVVWELPRQGDWRDVF
IGLDRVFKLFETLNIHPLREQGLKAAEEWVLERQEASGDWGGIIPAMLNSLLALRALDYAVDDPIVQRGMAAVDRFA
IETETEYRVQPCVSPVWDTALVMRAMVDSGVAPDHPALVKAGEWLLSKQILDYGDWHIKNKKGRPGGWAFEFENRFY
PDVDDTAVVVMALHAVTLPNENLKRRAIERAVAWIASMQCRPGGWAAFDVDNDQDWLNGIPYGDLKAMIDPNTADVT
ARVLEMVGRCQLAFDRVALDRALAYLRNEQEPEGCWFGRWGVNYLYGTSGVLTALSLVAPRYDRWRIRRAAEWLMQC
QNADGGWGETCWSYHDPSLKGKGDSTASQTAWAIIGLLAAGDATGDYATEAIERGIAYLLETQRPDGTWHEDYFTGT
GFPCHFYLKYHYYQQHFPLTALGRYARWRNLLAT

[1178] SEQ ID NO:20(野生型ApaSHC1的氨基酸序列)

[1179] MNMASRFSLKKILRSGSDTQGTNVNTLIQSGTSDIVRQKPAPQEPADLSALKAMGNSLTHTLSSACEWL

MKQQKPDGHWVGSVGSNASMEAEWCLALWFLGLEDHPLRPRLGKALLEMQRPDGSWGTYYGAGSGDINATVESYAAL
RSLGYAEDDPAVSKAAAWIISKGGLKNVRVFTRYWLALIGEWPWEKTPNLPPEIIWFPDNFVFSIYNFAQWARATMM
PLAILSARRPSRPLRPQDRLDALFPGGRANFDYELPTKEGRDVIADFFRLADKGLHWLQSSFLKRAPSREAAIKYVL
EWIIWHQDADGGWGGIQPPWVYGLMALHGEGYQFHHPVMAKALDALNDPGWRHDKGDASWIQATNSPVWDTMLSLMA
LHDANAEERFTPEMDKALDWLLSRQVRVKGDWSVKLPNTEPGGWAFEYANDRYPDTDDTAVALIAIASCRNRPEWQA
KGVEEAIGRGVRWLVAMQSSCGGWGAFDKDNNKSILAKIPFCDFGEALDPPSVDVTAHVLEAFGLLGLPRDLPCIQR
GLAYIRKEQDPTGPWFGRWGVNYLYGTGAVLPALAALGEDMTQPYISKACDWLINCQQENGGWGESCASYMEVSSIG
HGATTPSQTAWALMGLIAANRPQDYEAIAKGCRYLIDLQEEDGSWNEEEFTGTGFPGYGVGQTIKLDDPAISKRLMQ
GAELSRAFMLRYDLYRQLFPIIALSRASRLIKLGN

[1180] SEQ ID NO:21(GmoSHC变体的氨基酸序列)

[1181] MSPADISTKSSSFQRLDNMLPEAVSSACDWLIDQQKPDGHWVGPLESNACMEAEWCLALWFLGQEDHPL
RPRLAQALLEMQREDGSWGIYVGADHGDINTTVEAYAALRSMGYAADMPIMAKSAAWIQQKGGLRNVRVFTRYWLAL
IGEWPWDKTPNLPPEIIWLPDNFIFSIYNFAQWARATIMPLTILSARRPSRPLLPENRLDGLFPEGRENFDYELPVK
GEEDLWGRFFRAADKGLHSLQSFPVRRFVPREAAIRHVIEWIIRHQDADGGWGGIQPPWIYGLMALSVEGYPLHHPV
LAKAMDALNDPGWRRDKGDASWIQASNSPVWDTMLAVLALHDAGAEDRYSPQMDKAIGWLLDRQVRVKGDWSIKLPD
TEPGGWAFEYANDKYPDTDDTAVALIALAGCRHRPEWRERDIEGAISRGVNWLLAMQSSSGGWGAFDKDNNRSILTK
IPFCDFGEALDPPSVDVTAHVLEAFGLLGISRNHPSVQKALAYIRSEQERNGAWFGRWGVNYVYGTGAVLPALAAIG
EDMTQPYIVRACDWLMSVQQENGGWGESCASYMDINAVGHGVATASQTAWALIGLLAAKRPKDREAIARGCQFLIER
QEDGSWTEEEYTGTGYPGYGVGQAIKLDDPSLPDRLLQGAELSRAFMLRYDLYRQYFPVMALSRARRMMKEDASAAA

[1182] SEQ ID NO:22(SHC/HAC变体#90C7的核苷酸序列)

[1183] ATGGCTGAGCAGTTGGTGGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACGCGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTCGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTCAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGTTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGA
TCATGTTCCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTTCCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGTGCTGCACGGGTATC
AGAAGCTGTCGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTTGGGGCGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACACCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGTGTTTCGGCAGCTTCG
GGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG

TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATCGACAC
GCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGCTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGGCCTGGGCGCTGATGGCG
CTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGCTGGGATGAGCCGTACTACACCGGCACGGGCTTCCCAGGGGATTTCTACCTCGGCTACACCATGTACAGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

[1184] SEQ ID NO:23(SHC/HAC变体#115A7的核苷酸序列)

[1185] ATGGCTGAGCAGTTGGTGGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTCGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTCAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGTTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGA
TCATGTTCCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGACGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTTCCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGTGCTGCACGGGTATC
AGAAGCTGTCGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACAAGACGCAGCA
TCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTTGGGGCGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACACCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGTGTTTCGGCAGCTTCG
GGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATCGACAC
GCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGCTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGGCCTGGGCGCTGATGGCG
CTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGCTGGGATGAGCCGTACTACACCGGCACGGGCTTCCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

[1186] SEQ ID NO:24(SHC/HAC变体215G2+L37Q的氨基酸序列)

[1187] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLQSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMPLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDR
VLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLE

CFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLG
YTMYRHVFPTLALGRYKQAIERR

[1188] SEQ ID NO:25(SHC/HAC变体215G2+V174I的氨基酸序列)

[1189] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMPLNIYEFGSWARATIVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDR
VLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLE
CFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLG
YTMYRHVFPTLALGRYKQAIERR

[1190] SEQ ID NO:26(SHC/HAC变体215G2+V174I+F601Y的氨基酸序列)

[1191] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMPLNIYEFGSWARATIVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDR
VLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLE
CFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGYPGDFYLG
YTMYRHVFPTLALGRYKQAIERR

[1192] SEQ ID NO:27(SHC/HAC变体215G2+L37Q+V174I+F601Y的氨基酸序列)

[1193] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLQSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMPLNIYEFGSWARATIVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDR
VLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLE
CFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGYPGDFYLG
YTMYRHVFPTLALGRYKQAIERR

[1194] SEQ ID NO:28(野生型BmeSHC的氨基酸序列)

[1195] MIILLKEVQLEIQRRIAYLRPTQKNDGSFRYCFETGVMPDAFLIMLLRTFDLDKEVLIKQLTERIVSLQ
NEDGLWTLFDDEEHNLSATIQAYTALLYSGYYQKNDRILRKAERYIIDSGGISRAHFLTRWMLSVNGLYEWPKLFYL
PLSLLLVPTYVPLNFYELSTYARIHFVPMMVAGNKKFSLTSRHTPSLSHLDVREQKQESEETTQESRASIFLVDHLK
QLASLPSYIHKLGYQAAERYMLERIEKDGTLYSYATSTFFMIYGLLALGYKKDSFVIQKAIDGICSLLSTCSGHVHV
ENSTSTVWDTALLSYALQEAGVPQQDPMIKGTTRYLKKRQHTKLGDWQFHNPNTAPGGWGFSDINTNNPDLDDTSAA

IRALSRRAQTDTDYLESWQRGINWLLSMQNKDGGFAAFEKNTDSILFTYLPLENAKDAATDPATADLTGRVLECLGN
FAGMNKSHPSIKAAVKWLFDHQLDNGSWYGRWGVCYIYGTWAAITGLRAVGVSASDPRIIKAINWLKSIQQEDGGFG
ESCYSASLKKYVPLSFSTPSQTAWALDALMTICPLKDQSVEKGIKFLLNPNLTEQQTHYPTGIGLPGQFYIQYHSYN
DIFPLLALAHYAKKHSS

[1196] SEQ ID NO:29(野生型SalSHC的氨基酸序列)

[1197] VTATADGRLDPEYDPEPAAVGERPPVTDRLTGRQTTVATAPAPGAGRRQAEGPERSGPLDPAQALARAT
AELLSRQSPDGWWKGDLETNVTMDAEDLLLRQFLGIREPEQTAATAAWIRSQQREDGTWSTFYGGPPELSTTVEAYV
ALKLAGDDPGAPHMAAAARYVRERGGIAASRVFTRIWLALFGWWPWERLPEMPPEIIFLPRWLPLNIYAFGCWARQT
IVPLTVVSAHRPVRPAPFDLTELHTDPADPYPLRPLAPPTGWDGVFERLDLVLHAYHKRALRPLRRAALAQAGRWIV
ERQEADGCWGGIQPPAVYSLIALHLLGYDLEHPVMRAGLAAFDRFTVHTEDGRRWLEACQSPVWDTCLATIALRDAG
LPADHPALVSAADWMLAEEIRRPGDWSVQRPRLAPGGWAFEFENDNYPDIDDTAEVVLALKRVAHPDRARIDGAVRR
GVEWNLGMQSRNGAWGAFDVDNTSTLPNKLPFCDFGEVVDPPSADVTAHVVEMLAETGLAGDRRTRRGIDWLLKNQE
PDGSWFGRWGTNYIYGTGSVLPALVAAGIPGSHPAVRRAVDWLADRQNPDGGWGEDMRSYEDPVRWSGRGDSTASQT
AWALMALLAAGEGPDGARSEVVERGVQWLCRTQLPSGSWDEPQFTGTGFPWDFSINYHLYRLVFPVTALGRYLHGSP
LTGGGA

[1198] SEQ ID NO:30(野生型ApaSHCA的氨基酸序列)

[1199] MAADGSALSESRLSSEALDRAVLSAHTALSQAQQDDGHWVYELEADATIPAEYILLEHFMDRIDDALEQ
KIAIYLRRIQSEEHGGWPLYHNGKFDLSATVKAYFALKAVGDDINAPHMQRAREAILDHGGAERSNVFTRSQLALFG
EVPWRATPVMPVELMLLPAKAFFSVWNMSYWSRTVIAPLLVLAALRPVAANPRQVHVRELFVTPPEKVQDWIRGPYR
SAWGYVFKGLDSVLRPVVPFIPEKTHKKAIQAALDFIEPRLNGKDGLGAIYPAMANVVMMYRAMGVPDEDPRAKTAW
EAVQALIVEKDDEAYCQPCVSPIWDTGLSGHAMIEAASGPNGIAPEKTVAELKKASAWLRSKQILNVKGDWAVRNPN
LAPGGWAFQYGNDYYPDVDDTAVVGMLLHREGDPTNAEAIERARTWIVGMQSTDGGWGAFDIDNNKDVLNHIPFADH
GALLDPPTADVTARCISFLAQLRNPEDEPVIQRGLEYLRKEQEKDGSWFGRWGTNYIYGTWSALCALNAAGVSHDDP
AVVKAVEWLRSVQRADGGWGEGCESYEGGPHGTYGESLPSQTAWAVLGLMAAGRRDDPAVTRGIAWLADQQDANGEW
HEDPYNAVGFPKVFYLRYHGYKQFFPLMALARYRNLESSNTRRVSFGF

序列表

<110> 奇华顿股份有限公司

<120> 角鲨烯藿烯环化酶(SHC)变体

<130> 31050

<160> 30

<170> PatentIn version 3.5

<210> 1

<211> 631

<212> PRT

<213> 酸热脂环酸杆菌(Alicyclobacillus acidocaldarius)

<400> 1

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

210 215 220

Leu His Gly Tyr Gln Lys Leu Ser Val His Pro Phe Arg Arg Ala Ala
225 230 235 240

Glu Ile Arg Ala Leu Asp Trp Leu Leu Glu Arg Gln Ala Gly Asp Gly
245 250 255

Ser Trp Gly Gly Ile Gln Pro Pro Trp Phe Tyr Ala Leu Ile Ala Leu
260 265 270

Lys Ile Leu Asp Met Thr Gln His Pro Ala Phe Ile Lys Gly Trp Glu
275 280 285

Gly Leu Glu Leu Tyr Gly Val Glu Leu Asp Tyr Gly Gly Trp Met Phe
290 295 300

Gln Ala Ser Ile Ser Pro Val Trp Asp Thr Gly Leu Ala Val Leu Ala
305 310 315 320

Leu Arg Ala Ala Gly Leu Pro Ala Asp His Asp Arg Leu Val Lys Ala
325 330 335

Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala
340 345 350

Val Lys Arg Pro Asn Leu Lys Pro Gly Gly Phe Ala Phe Gln Phe Asp
355 360 365

Asn Val Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Val Trp Ala
370 375 380

Leu Asn Thr Leu Arg Leu Pro Asp Glu Arg Arg Arg Arg Asp Ala Met
385 390 395 400

Thr Lys Gly Phe Arg Trp Ile Val Gly Met Gln Ser Ser Asn Gly Gly
405 410 415

Trp Gly Ala Tyr Asp Val Asp Asn Thr Ser Asp Leu Pro Asn His Ile
420 425 430

Pro Phe Cys Asp Phe Gly Glu Val Thr Asp Pro Pro Ser Glu Asp Val
435 440 445

Thr Ala His Val Leu Glu Cys Phe Gly Ser Phe Gly Tyr Asp Asp Ala
450 455 460

Trp Lys Val Ile Arg Arg Ala Val Glu Tyr Leu Lys Arg Glu Gln Lys
465 470 475 480

Pro Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Leu Tyr Gly
485 490 495

Thr Gly Ala Val Val Ser Ala Leu Lys Ala Val Gly Ile Asp Thr Arg
500 505 510

Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn
515 520 525

Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala
530 535 540
Tyr Ala Gly Lys Gly Ala Ser Thr Pro Ser Gln Thr Ala Trp Ala Leu
545 550 555 560
Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg
565 570 575
Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp
580 585 590
Glu Pro Tyr Tyr Thr Gly Thr Gly Phe Pro Gly Asp Phe Tyr Leu Gly
595 600 605
Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr
610 615 620
Lys Gln Ala Ile Glu Arg Arg
625 630

<210> 2
<211> 631
<212> PRT
<213> 人工序列
<220>
<223> SHC/HAC酶变体#49
<400> 2

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

Val Pro Met Val Pro Pro Glu Ile Met Phe Leu Gly Lys Arg Met Pro
145 150 155 160
Leu Asn Ile Tyr Glu Phe Gly Ser Trp Ala Arg Ala Thr Val Val Ala
165 170 175
Leu Ser Ile Val Met Ser Arg Gln Pro Val Phe Pro Leu Pro Glu Arg
180 185 190
Ala Arg Val Pro Glu Leu Tyr Glu Thr Asp Val Pro Pro Arg Arg Arg
195 200 205
Gly Ala Lys Gly Gly Gly Gly Trp Ile Phe Asp Ala Leu Asp Arg Val
210 215 220
Leu His Gly Tyr Gln Lys Leu Ser Val His Pro Phe Arg Arg Ala Ala
225 230 235 240
Glu Ile Arg Ala Leu Asp Trp Leu Leu Glu Arg Gln Ala Gly Asp Gly
245 250 255
Ser Trp Gly Gly Ile Gln Pro Pro Trp Phe Tyr Ala Leu Ile Ala Leu
260 265 270
Lys Ile Leu Asp Met Thr Gln His Pro Ala Phe Ile Lys Gly Trp Glu
275 280 285
Gly Leu Glu Leu Tyr Gly Val Glu Leu Asp Tyr Gly Gly Trp Met Phe
290 295 300
Gln Ala Ser Ile Ser Pro Val Trp Asp Thr Gly Leu Ala Val Leu Ala
305 310 315 320
Leu Arg Ala Ala Gly Leu Pro Ala Asp His Asp Arg Leu Val Lys Ala
325 330 335
Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala
340 345 350
Val Lys Arg Pro Asn Leu Lys Pro Gly Gly Phe Ala Phe Gln Phe Asp
355 360 365
Asn Val Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Val Trp Ala
370 375 380
Leu Asn Thr Leu Arg Leu Pro Asp Glu Arg Arg Arg Arg Asp Ala Met
385 390 395 400
Thr Lys Gly Phe Arg Trp Ile Val Gly Met Gln Ser Ser Asn Gly Gly
405 410 415
Trp Gly Ala Tyr Asp Val Asp Asn Thr Ser Asp Leu Pro Asn Leu Thr
420 425 430
Pro Phe Cys Asp Phe Gly Glu Val Thr Asp Pro Pro Ser Glu Asp Val
435 440 445
Thr Ala His Val Leu Glu Cys Phe Gly Ser Phe Gly Tyr Asp Asp Ala

450 455 460
Trp Lys Val Ile Arg Arg Ala Val Glu Tyr Leu Lys Arg Glu Gln Lys
465 470 475 480
Pro Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Leu Tyr Gly
485 490 495
Thr Gly Ala Val Val Ser Ala Leu Lys Ala Val Gly Ile Asp Thr Arg
500 505 510
Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn
515 520 525
Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala
530 535 540
Tyr Ala Gly Lys Gly Ala Ser Thr Pro Ser Gln Thr Thr Trp Ala Leu
545 550 555 560
Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg
565 570 575
Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp
580 585 590
Glu Pro Tyr Tyr Thr Gly Thr Gly Phe Pro Gly Asp Phe Tyr Leu Gly
595 600 605
Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr
610 615 620
Lys Gln Ala Ile Glu Arg Arg
625 630

<210> 3
<211> 631
<212> PRT
<213> 人工序列
<220>
<223> SHC/HAC酶变体#65
<400> 3

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

65 70 75 80
Tyr Pro Gly Gly Pro Pro Asp Leu Asp Thr Thr Ile Glu Ala Tyr Val
85 90 95
Ala Leu Lys Tyr Ile Gly Met Ser Arg Asp Glu Glu Pro Met Gln Lys
100 105 110
Ala Leu Arg Phe Ile Gln Ser Gln Gly Gly Ile Glu Ser Ser Arg Val
115 120 125
Phe Thr Arg Arg Trp Leu Ala Leu Val Gly Glu Tyr Pro Trp Glu Lys
130 135 140
Val Pro Met Val Pro Pro Glu Ile Met Phe Leu Gly Lys Arg Met Pro
145 150 155 160
Leu Asn Ile Tyr Glu Phe Gly Ser Trp Ala Arg Ala Thr Val Val Ala
165 170 175
Leu Ser Ile Val Met Ser Arg Gln Pro Val Phe Pro Leu Pro Glu Arg
180 185 190
Ala Arg Val Pro Glu Leu Tyr Glu Thr Asp Val Pro Pro Arg Arg Arg
195 200 205
Gly Ala Lys Gly Gly Gly Gly Trp Ile Phe Asp Ala Leu Asp Arg Val
210 215 220
Leu His Gly Tyr Gln Lys Leu Ser Val His Pro Phe Arg Arg Ala Ala
225 230 235 240
Glu Ile Arg Ala Leu Asp Trp Leu Leu Glu Arg Gln Ala Gly Asp Gly
245 250 255
Ser Trp Gly Gly Ile Gln Pro Pro Trp Phe Tyr Ala Leu Ile Ala Leu
260 265 270
Lys Ile Leu Asp Met Thr Gln His Pro Ala Phe Ile Lys Gly Trp Glu
275 280 285
Gly Leu Glu Leu Tyr Gly Val Glu Leu Asp Tyr Gly Gly Trp Met Phe
290 295 300
Gln Ala Ser Ile Ser Pro Val Trp Asp Thr Gly Leu Ala Val Leu Ala
305 310 315 320
Leu Arg Ala Ala Gly Leu Pro Ala Asp His Asp Arg Leu Val Lys Ala
325 330 335
Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala
340 345 350
Val Lys Arg Pro Asn Leu Lys Pro Gly Gly Phe Ala Phe Gln Phe Asp
355 360 365
Asn Val Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Val Trp Ala
370 375 380

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

<210> 4
<211> 631
<212> PRT
<213> 人工序列
<220>
<223> SHC/HAC酶变体#66
<400> 4

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

305 310 315 320
Leu Arg Ala Ala Gly Leu Pro Ala Asp His Asp Arg Leu Val Lys Ala
325 330 335
Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala
340 345 350
Val Lys Arg Pro Asn Leu Lys Pro Gly Gly Phe Ala Phe Gln Phe Asp
355 360 365
Asn Val Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Val Trp Ala
370 375 380
Leu Asn Thr Leu Arg Leu Pro Asp Glu Arg Arg Arg Arg Asp Ala Met
385 390 395 400
Thr Lys Gly Phe Arg Trp Ile Val Gly Met Gln Ser Ser Asn Gly Gly
405 410 415
Trp Gly Ala Tyr Asp Val Asp Asn Thr Ser Asp Leu Pro Asn His Thr
420 425 430
Pro Phe Cys Asp Phe Gly Glu Val Thr Asp Pro Pro Ser Glu Asp Val
435 440 445
Thr Ala His Val Leu Glu Cys Phe Gly Ser Phe Gly Tyr Asp Asp Ala
450 455 460
Trp Lys Val Ile Arg Arg Ala Val Glu Tyr Leu Lys Arg Glu Gln Lys
465 470 475 480
Pro Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Leu Tyr Gly
485 490 495
Thr Gly Ala Val Val Ser Ala Leu Lys Ala Val Gly Ile Asp Thr Arg
500 505 510
Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn
515 520 525
Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala
530 535 540
Tyr Ala Gly Lys Gly Ala Ser Thr Pro Ser Gln Thr Thr Trp Ala Leu
545 550 555 560
Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg
565 570 575
Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp
580 585 590
Glu Pro Tyr Tyr Thr Gly Thr Gly Phe Pro Gly Asp Phe Tyr Leu Gly
595 600 605
Tyr Thr Met Tyr Ser His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr
610 615 620

Lys Gln Ala Ile Glu Arg Arg
625 630

<210> 5
<211> 631
<212> PRT
<213> 人工序列
<220>
<223> SHC/HAC酶变体#110B8
<400> 5

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

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

545 550 555 560
Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg
565 570 575
Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp
580 585 590
Glu Pro Tyr Tyr Thr Gly Thr Gly Phe Pro Gly Asp Phe Tyr Leu Gly
595 600 605
Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr
610 615 620
Lys Gln Ala Ile Glu Arg Arg
625 630

<210> 6
<211> 1896
<212> DNA
<213> 人工序列
<220>
<223> SHC/HAC酶变体#49
<400> 6

atggctgagc agttggtgga gctccggcc tacgcgcgga cgctggatcg cgcggtggag 60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggaggtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg tgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260

gacgtcgaca acacgagcga tctcccgaac ctcaccccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacgac ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 7
<211> 1896
<212> DNA
<213> 人工序列
<220>
<223> SHC/HAC酶变体#65
<400> 7

atggctgagc agttggtgga agctccggcc tacgcgcgga cgctggatcg cgcggtggag 60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggaggtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg tgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260

gacgtcgaca acacgagcga tctcccgaac cacaccccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacgac ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtacagcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 8
<211> 1896
<212> DNA
<213> 人工序列
<220>
<223> SHC/HAC酶变体#66
<400> 8

atggctgagc agttggtgga agctccggcc tacgcgcgga cgctggatcg cgcggtggag 60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
cacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggaggtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg tgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260

gacgtcgaca acacgagcga tctcccgaac cacaccccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacgac ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtacagcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 9
<211> 1896
<212> DNA
<213> 人工序列
<220>
<223> SHC/HAC酶变体#110B8
<400> 9

atggctgagc agttggtgga agctccggcc tacgcgcgga cgctggatcg cgcggtggag 60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
cacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggaggtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg tgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260

gacgtcgaca acacgagcga tctcccgaac ctcaccccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacgac ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 10
<211> 631
<212> PRT
<213> 人工序列
<220>
<223> 215G2SHC/HAC酶变体
<400> 10

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

165 170 175

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

180 185 190

Ala Arg Val Pro Glu Leu Tyr Glu Thr Asp Val Pro Pro Arg Arg Arg

195 200 205

Gly Ala Lys Gly Gly Gly Gly Trp Ile Phe Asp Ala Leu Asp Arg Val

210 215 220

Leu His Gly Tyr Gln Lys Leu Ser Val His Pro Phe Arg Arg Ala Ala

225 230 235 240

Glu Ile Arg Ala Leu Asp Trp Leu Leu Glu Arg Gln Ala Gly Asp Gly

245 250 255

Ser Trp Gly Gly Ile Gln Pro Pro Trp Phe Tyr Ala Leu Ile Ala Leu

260 265 270

Lys Ile Leu Asp Met Thr Gln His Pro Ala Phe Ile Lys Gly Trp Glu

275 280 285

Gly Leu Glu Leu Tyr Gly Val Glu Leu Asp Tyr Gly Gly Trp Met Phe

290 295 300

Gln Ala Ser Ile Ser Pro Val Trp Asp Thr Gly Leu Ala Val Leu Ala

305 310 315 320

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

325 330 335

Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala

340 345 350

Val Lys Arg Pro Asn Leu Lys Pro Gly Gly Phe Ala Phe Gln Phe Asp

355 360 365

Asn Val Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Val Trp Ala

370 375 380

Leu Asn Thr Leu Arg Leu Pro Asp Glu Arg Arg Arg Arg Asp Ala Met

385 390 395 400

Thr Lys Gly Phe Arg Trp Ile Val Gly Met Gln Ser Ser Asn Gly Gly

405 410 415

Trp Gly Ala Tyr Asp Val Asp Asn Thr Ser Asp Leu Pro Asn His Thr

420 425 430

Pro Phe Cys Asp Phe Gly Glu Val Thr Asp Pro Pro Ser Glu Asp Val

435 440 445

Thr Ala His Val Leu Glu Cys Phe Gly Ser Phe Gly Tyr Asp Asp Ala

450 455 460

Trp Lys Val Ile Arg Arg Ala Val Glu Tyr Leu Lys Arg Glu Gln Lys

465 470 475 480

Pro Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Leu Tyr Gly
485 490 495
Thr Gly Ala Val Val Ser Ala Leu Lys Ala Val Gly Ile Asp Thr Arg
500 505 510
Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn
515 520 525
Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala
530 535 540
Tyr Ala Gly Lys Gly Ala Ser Thr Pro Ser Gln Thr Ala Trp Ala Leu
545 550 555 560
Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg
565 570 575
Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp
580 585 590
Glu Pro Tyr Tyr Thr Gly Thr Gly Phe Pro Gly Asp Phe Tyr Leu Gly
595 600 605
Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr
610 615 620
Lys Gln Ala Ile Glu Arg Arg
625 630

<210> 11
<211> 608
<212> PRT
<213> 运动发酵单胞菌(Zymomonas mobilis)
<400> 11

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

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

420 425 430

Asp Arg Tyr Pro Asp Thr Asp Asp Thr Ala Val Ala Leu Ile Ala Leu

435 440 445

Ser Ser Tyr Arg Asp Lys Glu Glu Trp Gln Lys Lys Gly Val Glu Asp

450 455 460

Ala Ile Thr Arg Gly Val Asn Trp Leu Ile Ala Met Gln Ser Glu Cys

465 470 475 480

Gly Gly Trp Gly Ala Phe Asp Lys Asp Asn Asn Arg Ser Ile Leu Ser

485 490 495

Lys Ile Pro Phe Cys Asp Phe Gly Glu Ser Ile Asp Pro Pro Ser Val

500 505 510

Asp Val Thr Ala His Val Leu Glu Ala Phe Gly Thr Leu Gly Leu Ser

515 520 525

Arg Asp Met Pro Val Ile Gln Lys Ala Ile Asp Tyr Val Arg Ser Glu

530 535 540

Gln Glu Ala Glu Gly Ala Trp Phe Gly Arg Trp Gly Val Asn Tyr Ile

545 550 555 560

Tyr Gly Thr Gly Ala Val Leu Pro Ala Leu Ala Ala Ile Gly Glu Asp

565 570 575

Met Thr Gln Pro Tyr Ile Thr Lys Ala Cys Asp Trp Leu Val Ala His

580 585 590

Gln Gln Glu Asp Gly Gly Trp Gly Glu Ser Cys Ser Ser Tyr Met Glu

595 600 605

<210> 12

<211> 658

<212> PRT

<213> 运动发酵单胞菌

<400> 12

Met Thr Val Ser Thr Ser Ser Ala Phe His His Ser Pro Leu Ser Asp

1 5 10 15

Asp Val Glu Pro Ile Ile Gln Lys Ala Thr Arg Ala Leu Leu Glu Lys

20 25 30

Gln Gln Gln Asp Gly His Trp Val Phe Glu Leu Glu Ala Asp Ala Thr

35 40 45

Ile Pro Ala Glu Tyr Ile Leu Leu Lys His Tyr Leu Gly Glu Pro Glu

50 55 60

Asp Leu Glu Ile Glu Ala Lys Ile Gly Arg Tyr Leu Arg Arg Ile Gln

65 70 75 80

Gly Glu His Gly Gly Trp Ser Leu Phe Tyr Gly Gly Asp Leu Asp Leu

85 90 95

Ser Ala Thr Val Lys Ala Tyr Phe Ala Leu Lys Met Ile Gly Asp Ser
100 105 110

Pro Asp Ala Pro His Met Leu Arg Ala Arg Asn Glu Ile Leu Ala Arg
115 120 125

Gly Gly Ala Met Arg Ala Asn Val Phe Thr Arg Ile Gln Leu Ala Leu
130 135 140

Phe Gly Ala Met Ser Trp Glu His Val Pro Gln Met Pro Val Glu Leu
145 150 155 160

Met Leu Met Pro Glu Trp Phe Pro Val His Ile Asn Lys Met Ala Tyr
165 170 175

Trp Ala Arg Thr Val Leu Val Pro Leu Leu Val Leu Gln Ala Leu Lys
180 185 190

Pro Val Ala Arg Asn Arg Arg Gly Ile Leu Val Asp Glu Leu Phe Val
195 200 205

Pro Asp Val Leu Pro Thr Leu Gln Glu Ser Gly Asp Pro Ile Trp Arg
210 215 220

Arg Phe Phe Ser Ala Leu Asp Lys Val Leu His Lys Val Glu Pro Tyr
225 230 235 240

Trp Pro Lys Asn Met Arg Ala Lys Ala Ile His Ser Cys Val His Phe
245 250 255

Val Thr Glu Arg Leu Asn Gly Glu Asp Gly Leu Gly Ala Ile Tyr Pro
260 265 270

Ala Ile Ala Asn Ser Val Met Met Tyr Asp Ala Leu Gly Tyr Pro Glu
275 280 285

Asn His Pro Glu Arg Ala Ile Ala Arg Arg Ala Val Glu Lys Leu Met
290 295 300

Val Leu Asp Gly Thr Glu Asp Gln Gly Asp Lys Glu Val Tyr Cys Gln
305 310 315 320

Pro Cys Leu Ser Pro Ile Trp Asp Thr Ala Leu Val Ala His Ala Met
325 330 335

Leu Glu Val Gly Gly Asp Glu Ala Glu Lys Ser Ala Ile Ser Ala Leu
340 345 350

Ser Trp Leu Lys Pro Gln Gln Ile Leu Asp Val Lys Gly Asp Trp Ala
355 360 365

Trp Arg Arg Pro Asp Leu Arg Pro Gly Gly Trp Ala Phe Gln Tyr Arg
370 375 380

Asn Asp Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Thr Met Ala
385 390 395 400

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

<210> 13
<211> 684
<212> PRT
<213> 日本慢生根瘤菌(Bradyrhizobium japonicum)
<400> 13
Met Thr Val Thr Ser Ser Ala Ser Ala Arg Ala Thr Arg Asp Pro Gly

1 5 10 15
Asn Tyr Gln Thr Ala Leu Gln Ser Thr Val Arg Ala Ala Ala Asp Trp
20 25 30
Leu Ile Ala Asn Gln Lys Pro Asp Gly His Trp Val Gly Arg Ala Glu
35 40 45
Ser Asn Ala Cys Met Glu Ala Gln Trp Cys Leu Ala Leu Trp Phe Met
50 55 60
Gly Leu Glu Asp His Pro Leu Arg Lys Arg Leu Gly Gln Ser Leu Leu
65 70 75 80
Asp Ser Gln Arg Pro Asp Gly Ala Trp Gln Val Tyr Phe Gly Ala Pro
85 90 95
Asn Gly Asp Ile Asn Ala Thr Val Glu Ala Tyr Ala Ala Leu Arg Ser
100 105 110
Leu Gly Phe Arg Asp Asp Glu Pro Ala Val Arg Arg Ala Arg Glu Trp
115 120 125
Ile Glu Ala Lys Gly Gly Leu Arg Asn Ile Arg Val Phe Thr Arg Tyr
130 135 140
Trp Leu Ala Leu Ile Gly Glu Trp Pro Trp Glu Lys Thr Pro Asn Ile
145 150 155 160
Pro Pro Glu Val Ile Trp Phe Pro Leu Trp Phe Pro Phe Ser Ile Tyr
165 170 175
Asn Phe Ala Gln Trp Ala Arg Ala Thr Leu Met Pro Ile Ala Val Leu
180 185 190
Ser Ala Arg Arg Pro Ser Arg Pro Leu Pro Pro Glu Asn Arg Leu Asp
195 200 205
Ala Leu Phe Pro His Gly Arg Lys Ala Phe Asp Tyr Glu Leu Pro Val
210 215 220
Lys Ala Gly Ala Gly Gly Trp Asp Arg Phe Phe Arg Gly Ala Asp Lys
225 230 235 240
Val Leu His Lys Leu Gln Asn Leu Gly Asn Arg Leu Asn Leu Gly Leu
245 250 255
Phe Arg Pro Ala Ala Thr Ser Arg Val Leu Glu Trp Met Ile Arg His
260 265 270
Gln Asp Phe Asp Gly Ala Trp Gly Gly Ile Gln Pro Pro Trp Ile Tyr
275 280 285
Gly Leu Met Ala Leu Tyr Ala Glu Gly Tyr Pro Leu Asn His Pro Val
290 295 300
Leu Ala Lys Gly Leu Asp Ala Leu Asn Asp Pro Gly Trp Arg Val Asp
305 310 315 320

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

625 630 635 640
Asp Pro Ala Leu Ser Gln Arg Leu Met Gln Gly Pro Glu Leu Ser Arg
645 650 655
Ala Phe Met Leu Arg Tyr Gly Met Tyr Arg His Tyr Phe Pro Leu Met
660 665 670
Ala Leu Gly Arg Ala Leu Arg Pro Gln Ser His Ser
675 680

<210> 14
<211> 685
<212> PRT
<213> 莫比氏葡萄糖杆菌(Gluconobacter morbifer)
<400> 14

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

210 215 220

Glu Glu Asp Leu Trp Gly Arg Phe Phe Arg Ala Ala Asp Lys Gly Leu
225 230 235 240

His Ser Leu Gln Ser Phe Pro Val Arg Arg Phe Val Pro Arg Glu Ala
245 250 255

Ala Ile Arg His Val Ile Glu Trp Ile Ile Arg His Gln Asp Ala Asp
260 265 270

Gly Gly Trp Gly Gly Ile Gln Pro Pro Trp Ile Tyr Gly Leu Met Ala
275 280 285

Leu Ser Val Glu Gly Tyr Pro Leu His His Pro Val Leu Ala Lys Ala
290 295 300

Met Asp Ala Leu Asn Asp Pro Gly Trp Arg Arg Asp Lys Gly Asp Ala
305 310 315 320

Ser Trp Ile Gln Ala Thr Asn Ser Pro Val Trp Asp Thr Met Leu Ala
325 330 335

Val Leu Ala Leu His Asp Ala Gly Ala Glu Asp Arg Tyr Ser Pro Gln
340 345 350

Met Asp Lys Ala Ile Gly Trp Leu Leu Asp Arg Gln Val Arg Val Lys
355 360 365

Gly Asp Trp Ser Ile Lys Leu Pro Asp Thr Glu Pro Gly Gly Trp Ala
370 375 380

Phe Glu Tyr Ala Asn Asp Lys Tyr Pro Asp Thr Asp Asp Thr Ala Val
385 390 395 400

Ala Leu Ile Ala Leu Ala Gly Cys Arg His Arg Pro Glu Trp Arg Glu
405 410 415

Arg Asp Ile Glu Gly Ala Ile Ser Arg Gly Val Asn Trp Leu Leu Ala
420 425 430

Met Gln Ser Ser Ser Gly Gly Trp Gly Ala Phe Asp Lys Asp Asn Asn
435 440 445

Arg Ser Ile Leu Thr Lys Ile Pro Phe Cys Asp Phe Gly Glu Ala Leu
450 455 460

Asp Pro Pro Ser Val Asp Val Thr Ala His Val Leu Glu Ala Phe Gly
465 470 475 480

Leu Leu Gly Ile Ser Arg Asn His Pro Ser Val Gln Lys Ala Leu Ala
485 490 495

Tyr Ile Arg Ser Glu Gln Glu Arg Asn Gly Ala Trp Phe Gly Arg Trp
500 505 510

Gly Val Asn Tyr Val Tyr Gly Thr Gly Ala Val Leu Pro Ala Leu Ala
515 520 525

Ala Ile Gly Glu Asp Met Thr Gln Pro Tyr Ile Val Arg Ala Cys Asp
530 535 540
Trp Leu Met Ser Val Gln Gln Glu Asn Gly Gly Trp Gly Glu Ser Cys
545 550 555 560
Ala Ser Tyr Met Asp Ile Asn Ala Val Gly His Gly Val Ala Thr Ala
565 570 575
Ser Gln Thr Ala Trp Ala Leu Ile Gly Leu Leu Ala Ala Lys Arg Pro
580 585 590
Lys Asp Arg Glu Ala Ile Ala Arg Gly Cys Gln Phe Leu Ile Glu Arg
595 600 605
Gln Glu Asp Gly Ser Trp Thr Glu Glu Glu Tyr Thr Gly Thr Gly Phe
610 615 620
Pro Gly Tyr Gly Val Gly Gln Ala Ile Lys Leu Asp Asp Pro Ser Leu
625 630 635 640
Pro Asp Arg Leu Leu Gln Gly Ala Glu Leu Ser Arg Ala Phe Met Leu
645 650 655
Arg Tyr Asp Leu Tyr Arg Gln Tyr Phe Pro Val Met Ala Leu Ser Arg
660 665 670
Ala Arg Arg Met Met Lys Glu Asp Ala Ser Ala Ala Ala
675 680 685

<210> 15
<211> 1896
<212> DNA
<213> 酸热脂环酸杆菌
<400> 15

atggctgagc agttggtgga agcgccggcc tacgcgcgga cgctggatcg cgcggtggag 60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggatgtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840

ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacatcccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 16
<211> 1896
<212> DNA
<213> 人工序列
<220>
<223> 215G2SHC/HAC酶变体
<400> 16

atggctgagc agttggtgga agctccggcc tacgcgcgga cgctggatcg cgcggtggag 60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggaggtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg tgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840

ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacaccccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 17
<211> 631
<212> PRT
<213> 人工序列
<220>
<223> SHC/HAC酶变体#90C7
<400> 17

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

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

420 425 430

Pro Phe Cys Asp Phe Gly Glu Val Thr Asp Pro Pro Ser Glu Asp Val
435 440 445

Thr Ala His Val Leu Glu Cys Phe Gly Ser Phe Gly Tyr Asp Asp Ala
450 455 460

Trp Lys Val Ile Arg Arg Ala Val Glu Tyr Leu Lys Arg Glu Gln Lys
465 470 475 480

Pro Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Leu Tyr Gly
485 490 495

Thr Gly Ala Val Val Ser Ala Leu Lys Ala Val Gly Ile Asp Thr Arg
500 505 510

Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn
515 520 525

Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala
530 535 540

Tyr Ala Gly Lys Gly Ala Ser Thr Pro Ser Gln Thr Ala Trp Ala Leu
545 550 555 560

Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg
565 570 575

Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp
580 585 590

Glu Pro Tyr Tyr Thr Gly Thr Gly Phe Pro Gly Asp Phe Tyr Leu Gly
595 600 605

Tyr Thr Met Tyr Ser His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr
610 615 620

Lys Gln Ala Ile Glu Arg Arg
625 630

<210> 18

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> SHC/HAC酶变体#115A7

<400> 18

Met Ala Glu Gln Leu Val Glu Ala Pro Ala Tyr Ala Arg Thr Leu Asp
1 5 10 15

Arg Ala Val Glu Tyr Leu Leu Ser Cys Gln Lys Asp Glu Gly Tyr Trp
20 25 30

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

35 40 45

Leu Cys His Ile Leu Asp Arg Val Asp Arg Asp Arg Met Glu Lys Ile
50 55 60

Arg Arg Tyr Leu Leu His Glu Gln Arg Glu Asp Gly Thr Trp Ala Leu
65 70 75 80

Tyr Pro Gly Gly Pro Pro Asp Leu Asp Thr Thr Ile Glu Ala Tyr Val
85 90 95

Ala Leu Lys Tyr Ile Gly Met Ser Arg Asp Glu Glu Pro Met Gln Lys
100 105 110

Ala Leu Arg Phe Ile Gln Ser Gln Gly Gly Ile Glu Ser Ser Arg Val
115 120 125

Phe Thr Arg Arg Trp Leu Ala Leu Val Gly Glu Tyr Pro Trp Glu Lys
130 135 140

Val Pro Met Val Pro Pro Glu Ile Met Phe Leu Gly Lys Arg Met Pro
145 150 155 160

Leu Asn Ile Tyr Glu Phe Gly Ser Trp Ala Arg Thr Thr Val Val Ala
165 170 175

Leu Ser Ile Val Met Ser Arg Gln Pro Val Phe Pro Leu Pro Glu Arg
180 185 190

Ala Arg Val Pro Glu Leu Tyr Glu Thr Asp Val Pro Pro Arg Arg Arg
195 200 205

Gly Ala Lys Gly Gly Gly Gly Trp Ile Phe Asp Ala Leu Asp Arg Val
210 215 220

Leu His Gly Tyr Gln Lys Leu Ser Val His Pro Phe Arg Arg Ala Ala
225 230 235 240

Glu Ile Arg Ala Leu Asp Trp Leu Leu Glu Arg Gln Ala Gly Asp Gly
245 250 255

Ser Trp Gly Gly Ile Gln Pro Pro Trp Phe Tyr Ala Leu Ile Ala Leu
260 265 270

Lys Ile Leu Asp Lys Thr Gln His Pro Ala Phe Ile Lys Gly Trp Glu
275 280 285

Gly Leu Glu Leu Tyr Gly Val Glu Leu Asp Tyr Gly Gly Trp Met Phe
290 295 300

Gln Ala Ser Ile Ser Pro Val Trp Asp Thr Gly Leu Ala Val Leu Ala
305 310 315 320

Leu Arg Ala Ala Gly Leu Pro Ala Asp His Asp Arg Leu Val Lys Ala
325 330 335

Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala
340 345 350

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

<210> 19
<211> 642
<212> PRT

<213> 细长嗜热聚球藻菌(Thermosynechococcus elongatus)

<400> 19

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

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

Gly Phe Pro Cys His Phe Tyr Leu Lys Tyr His Tyr Tyr Gln Gln His
610 615 620
Phe Pro Leu Thr Ala Leu Gly Arg Tyr Ala Arg Trp Arg Asn Leu Leu
625 630 635 640
Ala Thr

<210> 20
<211> 720
<212> PRT
<213> 巴氏醋杆菌(Acetobacter pasteurianus)
<400> 20

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

225 230 235 240
Asp Arg Leu Asp Ala Leu Phe Pro Gly Gly Arg Ala Asn Phe Asp Tyr
245 250 255
Glu Leu Pro Thr Lys Glu Gly Arg Asp Val Ile Ala Asp Phe Phe Arg
260 265 270
Leu Ala Asp Lys Gly Leu His Trp Leu Gln Ser Ser Phe Leu Lys Arg
275 280 285
Ala Pro Ser Arg Glu Ala Ala Ile Lys Tyr Val Leu Glu Trp Ile Ile
290 295 300
Trp His Gln Asp Ala Asp Gly Gly Trp Gly Gly Ile Gln Pro Pro Trp
305 310 315 320
Val Tyr Gly Leu Met Ala Leu His Gly Glu Gly Tyr Gln Phe His His
325 330 335
Pro Val Met Ala Lys Ala Leu Asp Ala Leu Asn Asp Pro Gly Trp Arg
340 345 350
His Asp Lys Gly Asp Ala Ser Trp Ile Gln Ala Thr Asn Ser Pro Val
355 360 365
Trp Asp Thr Met Leu Ser Leu Met Ala Leu His Asp Ala Asn Ala Glu
370 375 380
Glu Arg Phe Thr Pro Glu Met Asp Lys Ala Leu Asp Trp Leu Leu Ser
385 390 395 400
Arg Gln Val Arg Val Lys Gly Asp Trp Ser Val Lys Leu Pro Asn Thr
405 410 415
Glu Pro Gly Gly Trp Ala Phe Glu Tyr Ala Asn Asp Arg Tyr Pro Asp
420 425 430
Thr Asp Asp Thr Ala Val Ala Leu Ile Ala Ile Ala Ser Cys Arg Asn
435 440 445
Arg Pro Glu Trp Gln Ala Lys Gly Val Glu Glu Ala Ile Gly Arg Gly
450 455 460
Val Arg Trp Leu Val Ala Met Gln Ser Ser Cys Gly Gly Trp Gly Ala
465 470 475 480
Phe Asp Lys Asp Asn Asn Lys Ser Ile Leu Ala Lys Ile Pro Phe Cys
485 490 495
Asp Phe Gly Glu Ala Leu Asp Pro Pro Ser Val Asp Val Thr Ala His
500 505 510
Val Leu Glu Ala Phe Gly Leu Leu Gly Leu Pro Arg Asp Leu Pro Cys
515 520 525
Ile Gln Arg Gly Leu Ala Tyr Ile Arg Lys Glu Gln Asp Pro Thr Gly
530 535 540

Pro Trp Phe Gly Arg Trp Gly Val Asn Tyr Leu Tyr Gly Thr Gly Ala
545 550 555 560
Val Leu Pro Ala Leu Ala Ala Leu Gly Glu Asp Met Thr Gln Pro Tyr
565 570 575
Ile Ser Lys Ala Cys Asp Trp Leu Ile Asn Cys Gln Gln Glu Asn Gly
580 585 590
Gly Trp Gly Glu Ser Cys Ala Ser Tyr Met Glu Val Ser Ser Ile Gly
595 600 605
His Gly Ala Thr Thr Pro Ser Gln Thr Ala Trp Ala Leu Met Gly Leu
610 615 620
Ile Ala Ala Asn Arg Pro Gln Asp Tyr Glu Ala Ile Ala Lys Gly Cys
625 630 635 640
Arg Tyr Leu Ile Asp Leu Gln Glu Glu Asp Gly Ser Trp Asn Glu Glu
645 650 655
Glu Phe Thr Gly Thr Gly Phe Pro Gly Tyr Gly Val Gly Gln Thr Ile
660 665 670
Lys Leu Asp Asp Pro Ala Ile Ser Lys Arg Leu Met Gln Gly Ala Glu
675 680 685
Leu Ser Arg Ala Phe Met Leu Arg Tyr Asp Leu Tyr Arg Gln Leu Phe
690 695 700
Pro Ile Ile Ala Leu Ser Arg Ala Ser Arg Leu Ile Lys Leu Gly Asn
705 710 715 720

<210> 21
<211> 685
<212> PRT
<213> 人工序列
<220>
<223> GmoSHC变体
<400> 21

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

Gln Arg Glu Asp Gly Ser Trp Gly Ile Tyr Val Gly Ala Asp His Gly
85 90 95
Asp Ile Asn Thr Thr Val Glu Ala Tyr Ala Ala Leu Arg Ser Met Gly
100 105 110
Tyr Ala Ala Asp Met Pro Ile Met Ala Lys Ser Ala Ala Trp Ile Gln
115 120 125
Gln Lys Gly Gly Leu Arg Asn Val Arg Val Phe Thr Arg Tyr Trp Leu
130 135 140
Ala Leu Ile Gly Glu Trp Pro Trp Asp Lys Thr Pro Asn Leu Pro Pro
145 150 155 160
Glu Ile Ile Trp Leu Pro Asp Asn Phe Ile Phe Ser Ile Tyr Asn Phe
165 170 175
Ala Gln Trp Ala Arg Ala Thr Ile Met Pro Leu Thr Ile Leu Ser Ala
180 185 190
Arg Arg Pro Ser Arg Pro Leu Leu Pro Glu Asn Arg Leu Asp Gly Leu
195 200 205
Phe Pro Glu Gly Arg Glu Asn Phe Asp Tyr Glu Leu Pro Val Lys Gly
210 215 220
Glu Glu Asp Leu Trp Gly Arg Phe Phe Arg Ala Ala Asp Lys Gly Leu
225 230 235 240
His Ser Leu Gln Ser Phe Pro Val Arg Arg Phe Val Pro Arg Glu Ala
245 250 255
Ala Ile Arg His Val Ile Glu Trp Ile Ile Arg His Gln Asp Ala Asp
260 265 270
Gly Gly Trp Gly Gly Ile Gln Pro Pro Trp Ile Tyr Gly Leu Met Ala
275 280 285
Leu Ser Val Glu Gly Tyr Pro Leu His His Pro Val Leu Ala Lys Ala
290 295 300
Met Asp Ala Leu Asn Asp Pro Gly Trp Arg Arg Asp Lys Gly Asp Ala
305 310 315 320
Ser Trp Ile Gln Ala Ser Asn Ser Pro Val Trp Asp Thr Met Leu Ala
325 330 335
Val Leu Ala Leu His Asp Ala Gly Ala Glu Asp Arg Tyr Ser Pro Gln
340 345 350
Met Asp Lys Ala Ile Gly Trp Leu Leu Asp Arg Gln Val Arg Val Lys
355 360 365
Gly Asp Trp Ser Ile Lys Leu Pro Asp Thr Glu Pro Gly Gly Trp Ala
370 375 380
Phe Glu Tyr Ala Asn Asp Lys Tyr Pro Asp Thr Asp Asp Thr Ala Val

385 390 395 400
Ala Leu Ile Ala Leu Ala Gly Cys Arg His Arg Pro Glu Trp Arg Glu
405 410 415
Arg Asp Ile Glu Gly Ala Ile Ser Arg Gly Val Asn Trp Leu Leu Ala
420 425 430
Met Gln Ser Ser Ser Gly Gly Trp Gly Ala Phe Asp Lys Asp Asn Asn
435 440 445
Arg Ser Ile Leu Thr Lys Ile Pro Phe Cys Asp Phe Gly Glu Ala Leu
450 455 460
Asp Pro Pro Ser Val Asp Val Thr Ala His Val Leu Glu Ala Phe Gly
465 470 475 480
Leu Leu Gly Ile Ser Arg Asn His Pro Ser Val Gln Lys Ala Leu Ala
485 490 495
Tyr Ile Arg Ser Glu Gln Glu Arg Asn Gly Ala Trp Phe Gly Arg Trp
500 505 510
Gly Val Asn Tyr Val Tyr Gly Thr Gly Ala Val Leu Pro Ala Leu Ala
515 520 525
Ala Ile Gly Glu Asp Met Thr Gln Pro Tyr Ile Val Arg Ala Cys Asp
530 535 540
Trp Leu Met Ser Val Gln Gln Glu Asn Gly Gly Trp Gly Glu Ser Cys
545 550 555 560
Ala Ser Tyr Met Asp Ile Asn Ala Val Gly His Gly Val Ala Thr Ala
565 570 575
Ser Gln Thr Ala Trp Ala Leu Ile Gly Leu Leu Ala Ala Lys Arg Pro
580 585 590
Lys Asp Arg Glu Ala Ile Ala Arg Gly Cys Gln Phe Leu Ile Glu Arg
595 600 605
Gln Glu Asp Gly Ser Trp Thr Glu Glu Glu Tyr Thr Gly Thr Gly Tyr
610 615 620
Pro Gly Tyr Gly Val Gly Gln Ala Ile Lys Leu Asp Asp Pro Ser Leu
625 630 635 640
Pro Asp Arg Leu Leu Gln Gly Ala Glu Leu Ser Arg Ala Phe Met Leu
645 650 655
Arg Tyr Asp Leu Tyr Arg Gln Tyr Phe Pro Val Met Ala Leu Ser Arg
660 665 670
Ala Arg Arg Met Met Lys Glu Asp Ala Ser Ala Ala Ala
675 680 685

<210> 22

<211> 1896

<212> DNA

<213> 人工序列

<220>

<223> SHC/HAC酶变体#90C7

<400> 22

atggctgagc agttggtgga agctccggcc tacgcgcgga cgctggatcg cgcggtggag 60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacgcg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggaggtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg tgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacaccccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtacagcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 23

<211> 1896

<212> DNA

<213> 人工序列

<220>

<223> SHC/HAC酶变体#115A7

<400> 23

atggctgagc agttggtgga gctccggcc tacgcgcgga cgctggatcg cgcggtggag 60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggaggtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cggacgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg tgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacaa gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacaccccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 24

<211> 631

<212> PRT
<213> 人工序列
<220>
<223> 215G2 SHC/HAC酶变体+ L37Q突变
<400> 24

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

Lys Ile Leu Asp Met Thr Gln His Pro Ala Phe Ile Lys Gly Trp Glu
275 280 285
Gly Leu Glu Leu Tyr Gly Val Glu Leu Asp Tyr Gly Gly Trp Met Phe
290 295 300
Gln Ala Ser Ile Ser Pro Val Trp Asp Thr Gly Leu Ala Val Leu Ala
305 310 315 320
Leu Arg Ala Ala Gly Leu Pro Ala Asp His Asp Arg Leu Val Lys Ala
325 330 335
Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala
340 345 350
Val Lys Arg Pro Asn Leu Lys Pro Gly Gly Phe Ala Phe Gln Phe Asp
355 360 365
Asn Val Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Val Trp Ala
370 375 380
Leu Asn Thr Leu Arg Leu Pro Asp Glu Arg Arg Arg Arg Asp Ala Met
385 390 395 400
Thr Lys Gly Phe Arg Trp Ile Val Gly Met Gln Ser Ser Asn Gly Gly
405 410 415
Trp Gly Ala Tyr Asp Val Asp Asn Thr Ser Asp Leu Pro Asn His Thr
420 425 430
Pro Phe Cys Asp Phe Gly Glu Val Thr Asp Pro Pro Ser Glu Asp Val
435 440 445
Thr Ala His Val Leu Glu Cys Phe Gly Ser Phe Gly Tyr Asp Asp Ala
450 455 460
Trp Lys Val Ile Arg Arg Ala Val Glu Tyr Leu Lys Arg Glu Gln Lys
465 470 475 480
Pro Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Leu Tyr Gly
485 490 495
Thr Gly Ala Val Val Ser Ala Leu Lys Ala Val Gly Ile Asp Thr Arg
500 505 510
Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn
515 520 525
Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala
530 535 540
Tyr Ala Gly Lys Gly Ala Ser Thr Pro Ser Gln Thr Ala Trp Ala Leu
545 550 555 560
Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg
565 570 575
Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp

580 585 590
Glu Pro Tyr Tyr Thr Gly Thr Gly Phe Pro Gly Asp Phe Tyr Leu Gly
595 600 605
Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr
610 615 620
Lys Gln Ala Ile Glu Arg Arg
625 630

<210> 25
<211> 631
<212> PRT
<213> 人工序列
<220>
<223> 215G2 SHC/HAC酶变体+ V174I突变
<400> 25

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

195 200 205

Gly Ala Lys Gly Gly Gly Gly Trp Ile Phe Asp Ala Leu Asp Arg Val

210 215 220

Leu His Gly Tyr Gln Lys Leu Ser Val His Pro Phe Arg Arg Ala Ala

225 230 235 240

Glu Ile Arg Ala Leu Asp Trp Leu Leu Glu Arg Gln Ala Gly Asp Gly

245 250 255

Ser Trp Gly Gly Ile Gln Pro Pro Trp Phe Tyr Ala Leu Ile Ala Leu

260 265 270

Lys Ile Leu Asp Met Thr Gln His Pro Ala Phe Ile Lys Gly Trp Glu

275 280 285

Gly Leu Glu Leu Tyr Gly Val Glu Leu Asp Tyr Gly Gly Trp Met Phe

290 295 300

Gln Ala Ser Ile Ser Pro Val Trp Asp Thr Gly Leu Ala Val Leu Ala

305 310 315 320

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

325 330 335

Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala

340 345 350

Val Lys Arg Pro Asn Leu Lys Pro Gly Gly Phe Ala Phe Gln Phe Asp

355 360 365

Asn Val Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Val Trp Ala

370 375 380

Leu Asn Thr Leu Arg Leu Pro Asp Glu Arg Arg Arg Arg Asp Ala Met

385 390 395 400

Thr Lys Gly Phe Arg Trp Ile Val Gly Met Gln Ser Ser Asn Gly Gly

405 410 415

Trp Gly Ala Tyr Asp Val Asp Asn Thr Ser Asp Leu Pro Asn His Thr

420 425 430

Pro Phe Cys Asp Phe Gly Glu Val Thr Asp Pro Pro Ser Glu Asp Val

435 440 445

Thr Ala His Val Leu Glu Cys Phe Gly Ser Phe Gly Tyr Asp Asp Ala

450 455 460

Trp Lys Val Ile Arg Arg Ala Val Glu Tyr Leu Lys Arg Glu Gln Lys

465 470 475 480

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

485 490 495

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

500 505 510

Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn
515 520 525
Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala
530 535 540
Tyr Ala Gly Lys Gly Ala Ser Thr Pro Ser Gln Thr Ala Trp Ala Leu
545 550 555 560
Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg
565 570 575
Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp
580 585 590
Glu Pro Tyr Tyr Thr Gly Thr Gly Phe Pro Gly Asp Phe Tyr Leu Gly
595 600 605
Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr
610 615 620
Lys Gln Ala Ile Glu Arg Arg
625 630

<210> 26
<211> 631
<212> PRT
<213> 人工序列
<220>
<223> 215G2 SHC/HAC酶变体+ V174I + F601Y突变
<400> 26

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

Phe Thr Arg Arg Trp Leu Ala Leu Val Gly Glu Tyr Pro Trp Glu Lys
130 135 140
Val Pro Met Val Pro Pro Glu Ile Met Phe Leu Gly Lys Arg Met Pro
145 150 155 160
Leu Asn Ile Tyr Glu Phe Gly Ser Trp Ala Arg Ala Thr Ile Val Ala
165 170 175
Leu Ser Ile Val Met Ser Arg Gln Pro Val Phe Pro Leu Pro Glu Arg
180 185 190
Ala Arg Val Pro Glu Leu Tyr Glu Thr Asp Val Pro Pro Arg Arg Arg
195 200 205
Gly Ala Lys Gly Gly Gly Gly Trp Ile Phe Asp Ala Leu Asp Arg Val
210 215 220
Leu His Gly Tyr Gln Lys Leu Ser Val His Pro Phe Arg Arg Ala Ala
225 230 235 240
Glu Ile Arg Ala Leu Asp Trp Leu Leu Glu Arg Gln Ala Gly Asp Gly
245 250 255
Ser Trp Gly Gly Ile Gln Pro Pro Trp Phe Tyr Ala Leu Ile Ala Leu
260 265 270
Lys Ile Leu Asp Met Thr Gln His Pro Ala Phe Ile Lys Gly Trp Glu
275 280 285
Gly Leu Glu Leu Tyr Gly Val Glu Leu Asp Tyr Gly Gly Trp Met Phe
290 295 300
Gln Ala Ser Ile Ser Pro Val Trp Asp Thr Gly Leu Ala Val Leu Ala
305 310 315 320
Leu Arg Ala Ala Gly Leu Pro Ala Asp His Asp Arg Leu Val Lys Ala
325 330 335
Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala
340 345 350
Val Lys Arg Pro Asn Leu Lys Pro Gly Gly Phe Ala Phe Gln Phe Asp
355 360 365
Asn Val Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Val Trp Ala
370 375 380
Leu Asn Thr Leu Arg Leu Pro Asp Glu Arg Arg Arg Arg Asp Ala Met
385 390 395 400
Thr Lys Gly Phe Arg Trp Ile Val Gly Met Gln Ser Ser Asn Gly Gly
405 410 415
Trp Gly Ala Tyr Asp Val Asp Asn Thr Ser Asp Leu Pro Asn His Thr
420 425 430
Pro Phe Cys Asp Phe Gly Glu Val Thr Asp Pro Pro Ser Glu Asp Val

435 440 445
Thr Ala His Val Leu Glu Cys Phe Gly Ser Phe Gly Tyr Asp Asp Ala
450 455 460
Trp Lys Val Ile Arg Arg Ala Val Glu Tyr Leu Lys Arg Glu Gln Lys
465 470 475 480
Pro Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Leu Tyr Gly
485 490 495
Thr Gly Ala Val Val Ser Ala Leu Lys Ala Val Gly Ile Asp Thr Arg
500 505 510
Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn
515 520 525
Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala
530 535 540
Tyr Ala Gly Lys Gly Ala Ser Thr Pro Ser Gln Thr Ala Trp Ala Leu
545 550 555 560
Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg
565 570 575
Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp
580 585 590
Glu Pro Tyr Tyr Thr Gly Thr Gly Tyr Pro Gly Asp Phe Tyr Leu Gly
595 600 605
Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr
610 615 620
Lys Gln Ala Ile Glu Arg Arg
625 630

<210> 27
<211> 631
<212> PRT
<213> 人工序列
<220>
<223> 215G2 SHC/HAC酶变体+ L37Q + V174I + F601Y突变
<400> 27

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

50 55 60

Arg Arg Tyr Leu Leu His Glu Gln Arg Glu Asp Gly Thr Trp Ala Leu
65 70 75 80
Tyr Pro Gly Gly Pro Pro Asp Leu Asp Thr Thr Ile Glu Ala Tyr Val
85 90 95
Ala Leu Lys Tyr Ile Gly Met Ser Arg Asp Glu Glu Pro Met Gln Lys
100 105 110
Ala Leu Arg Phe Ile Gln Ser Gln Gly Gly Ile Glu Ser Ser Arg Val
115 120 125
Phe Thr Arg Arg Trp Leu Ala Leu Val Gly Glu Tyr Pro Trp Glu Lys
130 135 140
Val Pro Met Val Pro Pro Glu Ile Met Phe Leu Gly Lys Arg Met Pro
145 150 155 160
Leu Asn Ile Tyr Glu Phe Gly Ser Trp Ala Arg Ala Thr Ile Val Ala
165 170 175
Leu Ser Ile Val Met Ser Arg Gln Pro Val Phe Pro Leu Pro Glu Arg
180 185 190
Ala Arg Val Pro Glu Leu Tyr Glu Thr Asp Val Pro Pro Arg Arg Arg
195 200 205
Gly Ala Lys Gly Gly Gly Gly Trp Ile Phe Asp Ala Leu Asp Arg Val
210 215 220
Leu His Gly Tyr Gln Lys Leu Ser Val His Pro Phe Arg Arg Ala Ala
225 230 235 240
Glu Ile Arg Ala Leu Asp Trp Leu Leu Glu Arg Gln Ala Gly Asp Gly
245 250 255
Ser Trp Gly Gly Ile Gln Pro Pro Trp Phe Tyr Ala Leu Ile Ala Leu
260 265 270
Lys Ile Leu Asp Met Thr Gln His Pro Ala Phe Ile Lys Gly Trp Glu
275 280 285
Gly Leu Glu Leu Tyr Gly Val Glu Leu Asp Tyr Gly Gly Trp Met Phe
290 295 300
Gln Ala Ser Ile Ser Pro Val Trp Asp Thr Gly Leu Ala Val Leu Ala
305 310 315 320
Leu Arg Ala Ala Gly Leu Pro Ala Asp His Asp Arg Leu Val Lys Ala
325 330 335
Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala
340 345 350
Val Lys Arg Pro Asn Leu Lys Pro Gly Gly Phe Ala Phe Gln Phe Asp
355 360 365

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

<210> 28
<211> 625
<212> PRT
<213> 人工序列
<220>

<223> 巨大芽孢杆菌(Bacillus megaterium)

<400> 28

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

290 295 300

Ser Thr Val Trp Asp Thr Ala Leu Leu Ser Tyr Ala Leu Gln Glu Ala
305 310 315 320

Gly Val Pro Gln Gln Asp Pro Met Ile Lys Gly Thr Thr Arg Tyr Leu
325 330 335

Lys Lys Arg Gln His Thr Lys Leu Gly Asp Trp Gln Phe His Asn Pro
340 345 350

Asn Thr Ala Pro Gly Gly Trp Gly Phe Ser Asp Ile Asn Thr Asn Asn
355 360 365

Pro Asp Leu Asp Asp Thr Ser Ala Ala Ile Arg Ala Leu Ser Arg Arg
370 375 380

Ala Gln Thr Asp Thr Asp Tyr Leu Glu Ser Trp Gln Arg Gly Ile Asn
385 390 395 400

Trp Leu Leu Ser Met Gln Asn Lys Asp Gly Gly Phe Ala Ala Phe Glu
405 410 415

Lys Asn Thr Asp Ser Ile Leu Phe Thr Tyr Leu Pro Leu Glu Asn Ala
420 425 430

Lys Asp Ala Ala Thr Asp Pro Ala Thr Ala Asp Leu Thr Gly Arg Val
435 440 445

Leu Glu Cys Leu Gly Asn Phe Ala Gly Met Asn Lys Ser His Pro Ser
450 455 460

Ile Lys Ala Ala Val Lys Trp Leu Phe Asp His Gln Leu Asp Asn Gly
465 470 475 480

Ser Trp Tyr Gly Arg Trp Gly Val Cys Tyr Ile Tyr Gly Thr Trp Ala
485 490 495

Ala Ile Thr Gly Leu Arg Ala Val Gly Val Ser Ala Ser Asp Pro Arg
500 505 510

Ile Ile Lys Ala Ile Asn Trp Leu Lys Ser Ile Gln Gln Glu Asp Gly
515 520 525

Gly Phe Gly Glu Ser Cys Tyr Ser Ala Ser Leu Lys Lys Tyr Val Pro
530 535 540

Leu Ser Phe Ser Thr Pro Ser Gln Thr Ala Trp Ala Leu Asp Ala Leu
545 550 555 560

Met Thr Ile Cys Pro Leu Lys Asp Gln Ser Val Glu Lys Gly Ile Lys
565 570 575

Phe Leu Leu Asn Pro Asn Leu Thr Glu Gln Gln Thr His Tyr Pro Thr
580 585 590

Gly Ile Gly Leu Pro Gly Gln Phe Tyr Ile Gln Tyr His Ser Tyr Asn
595 600 605

Asp Ile Phe Pro Leu Leu Ala Leu Ala His Tyr Ala Lys Lys His Ser
610 615 620
Ser
625

<210> 29
<211> 691
<212> PRT
<213> 人工序列
<220>
<223> 白色链霉菌(Streptomyces albolongus)
<400> 29

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

Val Pro Leu Thr Val Val Ser Ala His Arg Pro Val Arg Pro Ala Pro
225 230 235 240
Phe Asp Leu Thr Glu Leu His Thr Asp Pro Ala Asp Pro Tyr Pro Leu
245 250 255
Arg Pro Leu Ala Pro Pro Thr Gly Trp Asp Gly Val Phe Glu Arg Leu
260 265 270
Asp Leu Val Leu His Ala Tyr His Lys Arg Ala Leu Arg Pro Leu Arg
275 280 285
Arg Ala Ala Leu Ala Gln Ala Gly Arg Trp Ile Val Glu Arg Gln Glu
290 295 300
Ala Asp Gly Cys Trp Gly Gly Ile Gln Pro Pro Ala Val Tyr Ser Leu
305 310 315 320
Ile Ala Leu His Leu Leu Gly Tyr Asp Leu Glu His Pro Val Met Arg
325 330 335
Ala Gly Leu Ala Ala Phe Asp Arg Phe Thr Val His Thr Glu Asp Gly
340 345 350
Arg Arg Trp Leu Glu Ala Cys Gln Ser Pro Val Trp Asp Thr Cys Leu
355 360 365
Ala Thr Ile Ala Leu Arg Asp Ala Gly Leu Pro Ala Asp His Pro Ala
370 375 380
Leu Val Ser Ala Ala Asp Trp Met Leu Ala Glu Glu Ile Arg Arg Pro
385 390 395 400
Gly Asp Trp Ser Val Gln Arg Pro Arg Leu Ala Pro Gly Gly Trp Ala
405 410 415
Phe Glu Phe Glu Asn Asp Asn Tyr Pro Asp Ile Asp Asp Thr Ala Glu
420 425 430
Val Val Leu Ala Leu Lys Arg Val Ala His Pro Asp Arg Ala Arg Ile
435 440 445
Asp Gly Ala Val Arg Arg Gly Val Glu Trp Asn Leu Gly Met Gln Ser
450 455 460
Arg Asn Gly Ala Trp Gly Ala Phe Asp Val Asp Asn Thr Ser Thr Leu
465 470 475 480
Pro Asn Lys Leu Pro Phe Cys Asp Phe Gly Glu Val Val Asp Pro Pro
485 490 495
Ser Ala Asp Val Thr Ala His Val Val Glu Met Leu Ala Glu Thr Gly
500 505 510
Leu Ala Gly Asp Arg Arg Thr Arg Arg Gly Ile Asp Trp Leu Leu Lys
515 520 525
Asn Gln Glu Pro Asp Gly Ser Trp Phe Gly Arg Trp Gly Thr Asn Tyr

530 535 540

Ile Tyr Gly Thr Gly Ser Val Leu Pro Ala Leu Val Ala Ala Gly Ile
545 550 555 560

Pro Gly Ser His Pro Ala Val Arg Arg Ala Val Asp Trp Leu Ala Asp
565 570 575

Arg Gln Asn Pro Asp Gly Gly Trp Gly Glu Asp Met Arg Ser Tyr Glu
580 585 590

Asp Pro Val Arg Trp Ser Gly Arg Gly Asp Ser Thr Ala Ser Gln Thr
595 600 605

Ala Trp Ala Leu Met Ala Leu Leu Ala Ala Gly Glu Gly Pro Asp Gly
610 615 620

Ala Arg Ser Glu Val Val Glu Arg Gly Val Gln Trp Leu Cys Arg Thr
625 630 635 640

Gln Leu Pro Ser Gly Ser Trp Asp Glu Pro Gln Phe Thr Gly Thr Gly
645 650 655

Phe Pro Trp Asp Phe Ser Ile Asn Tyr His Leu Tyr Arg Leu Val Phe
660 665 670

Pro Val Thr Ala Leu Gly Arg Tyr Leu His Gly Ser Pro Leu Thr Gly
675 680 685

Gly Gly Ala
690

<210> 30
<211> 656
<212> PRT
<213> 人工序列
<220>
<223> 巴氏醋杆菌
<400> 30

Met Ala Ala Asp Gly Ser Ala Leu Ser Glu Ser Arg Leu Ser Ser Glu
1 5 10 15

Ala Leu Asp Arg Ala Val Leu Ser Ala His Thr Ala Leu Ser Gln Ala
20 25 30

Gln Gln Asp Asp Gly His Trp Val Tyr Glu Leu Glu Ala Asp Ala Thr
35 40 45

Ile Pro Ala Glu Tyr Ile Leu Leu Glu His Phe Met Asp Arg Ile Asp
50 55 60

Asp Ala Leu Glu Gln Lys Ile Ala Ile Tyr Leu Arg Arg Ile Gln Ser
65 70 75 80

Glu Glu His Gly Gly Trp Pro Leu Tyr His Asn Gly Lys Phe Asp Leu

85 90 95

Ser Ala Thr Val Lys Ala Tyr Phe Ala Leu Lys Ala Val Gly Asp Asp
100 105 110

Ile Asn Ala Pro His Met Gln Arg Ala Arg Glu Ala Ile Leu Asp His
115 120 125

Gly Gly Ala Glu Arg Ser Asn Val Phe Thr Arg Ser Gln Leu Ala Leu
130 135 140

Phe Gly Glu Val Pro Trp Arg Ala Thr Pro Val Met Pro Val Glu Leu
145 150 155 160

Met Leu Leu Pro Ala Lys Ala Phe Phe Ser Val Trp Asn Met Ser Tyr
165 170 175

Trp Ser Arg Thr Val Ile Ala Pro Leu Leu Val Leu Ala Ala Leu Arg
180 185 190

Pro Val Ala Ala Asn Pro Arg Gln Val His Val Arg Glu Leu Phe Val
195 200 205

Thr Pro Pro Glu Lys Val Gln Asp Trp Ile Arg Gly Pro Tyr Arg Ser
210 215 220

Ala Trp Gly Tyr Val Phe Lys Gly Leu Asp Ser Val Leu Arg Pro Val
225 230 235 240

Val Pro Phe Ile Pro Glu Lys Thr His Lys Lys Ala Ile Gln Ala Ala
245 250 255

Leu Asp Phe Ile Glu Pro Arg Leu Asn Gly Lys Asp Gly Leu Gly Ala
260 265 270

Ile Tyr Pro Ala Met Ala Asn Val Val Met Met Tyr Arg Ala Met Gly
275 280 285

Val Pro Asp Glu Asp Pro Arg Ala Lys Thr Ala Trp Glu Ala Val Gln
290 295 300

Ala Leu Ile Val Glu Lys Asp Asp Glu Ala Tyr Cys Gln Pro Cys Val
305 310 315 320

Ser Pro Ile Trp Asp Thr Gly Leu Ser Gly His Ala Met Ile Glu Ala
325 330 335

Ala Ser Gly Pro Asn Gly Ile Ala Pro Glu Lys Thr Val Ala Glu Leu
340 345 350

Lys Lys Ala Ser Ala Trp Leu Arg Ser Lys Gln Ile Leu Asn Val Lys
355 360 365

Gly Asp Trp Ala Val Arg Asn Pro Asn Leu Ala Pro Gly Gly Trp Ala
370 375 380

Phe Gln Tyr Gly Asn Asp Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val
385 390 395 400

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

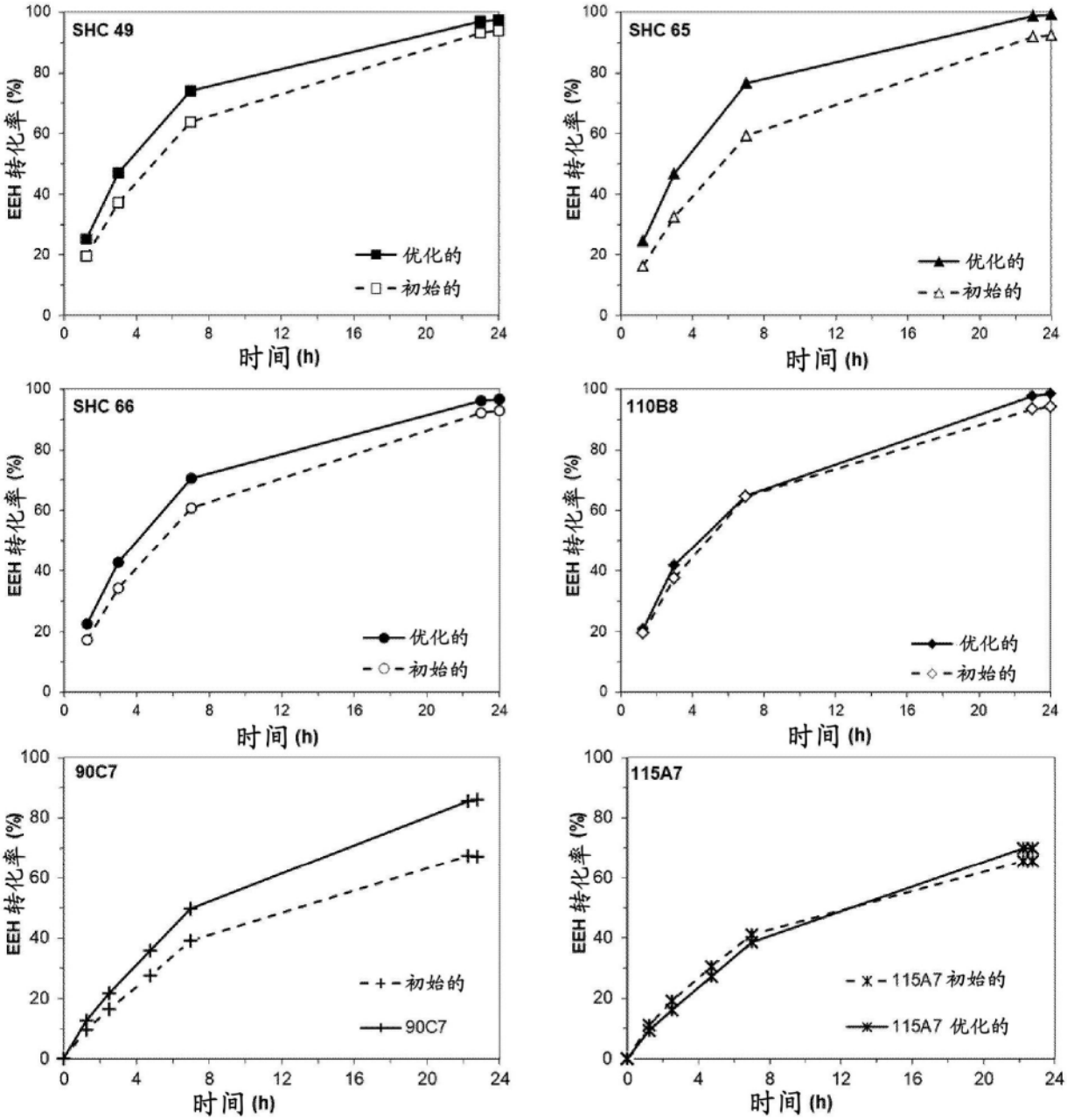

SHC 49
EEH 转化率 (%)
时间 (h)
优化的
初始的
SHC 65
EEH 转化率 (%)
时间 (h)
优化的
初始的
SHC 66
EEH 转化率 (%)
时间 (h)
优化的
初始的
110B8
EEH 转化率 (%)
时间 (h)
优化的
初始的
90C7
EEH 转化率 (%)
时间 (h)
初始的
90C7
115A7
EEH 转化率 (%)
时间 (h)
115A7 初始的
115A7 优化的

图1

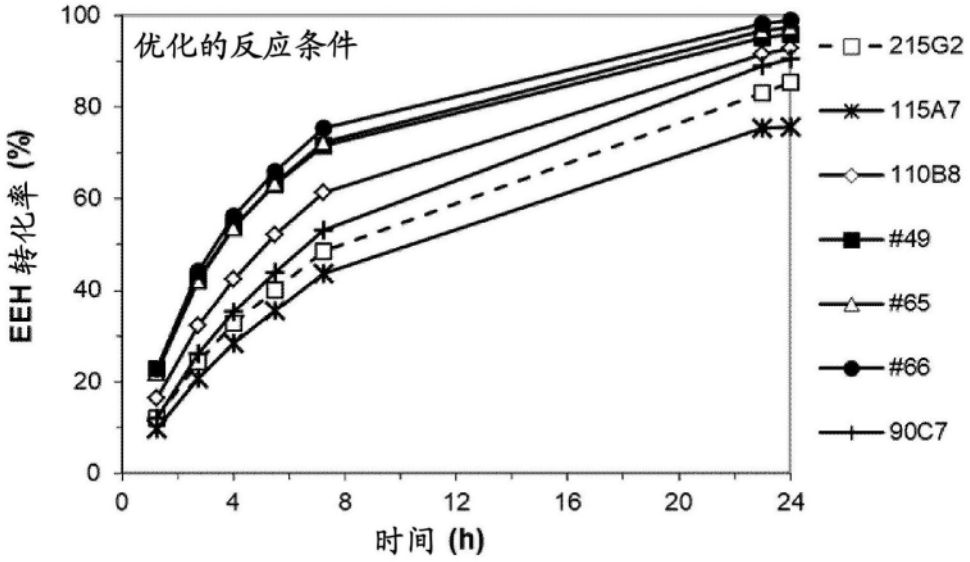

优化的反应条件
EEH 转化率 (%)
时间 (h)
215G2
115A7
110B8
#49
#65
#66
90C7

图2

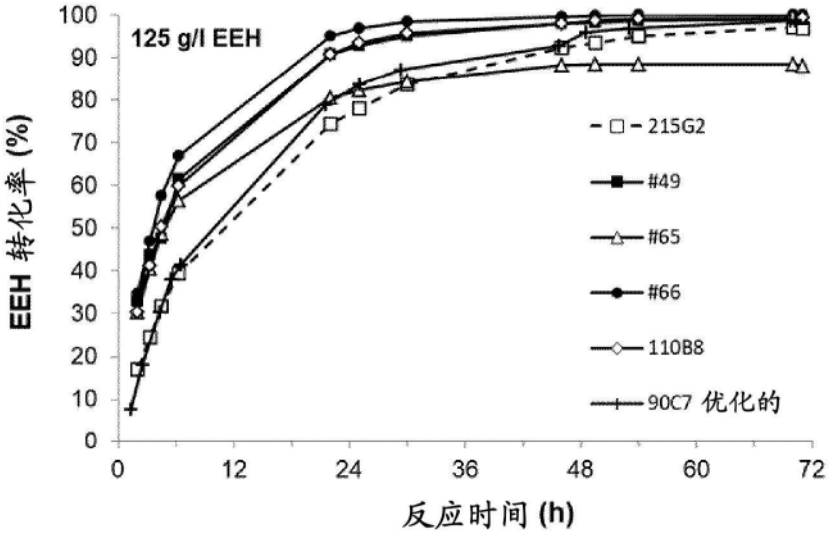

125 g/l EEH
EEH 转化率 (%)
反应时间 (h)
215G2
#49
#65
#66
110B8
90C7 优化的

图3

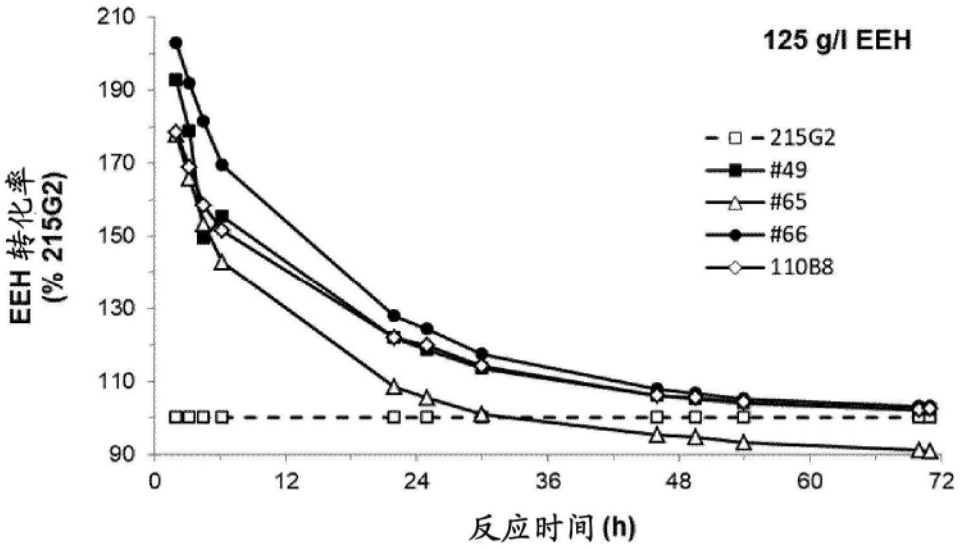

125 g/l EEH
215G2
#49
#65
#66
110B8
EEH 转化率
(% 215G2)
210
190
170
150
130
110
90
0
12
24
36
48
60
72
反应时间 (h)

图4

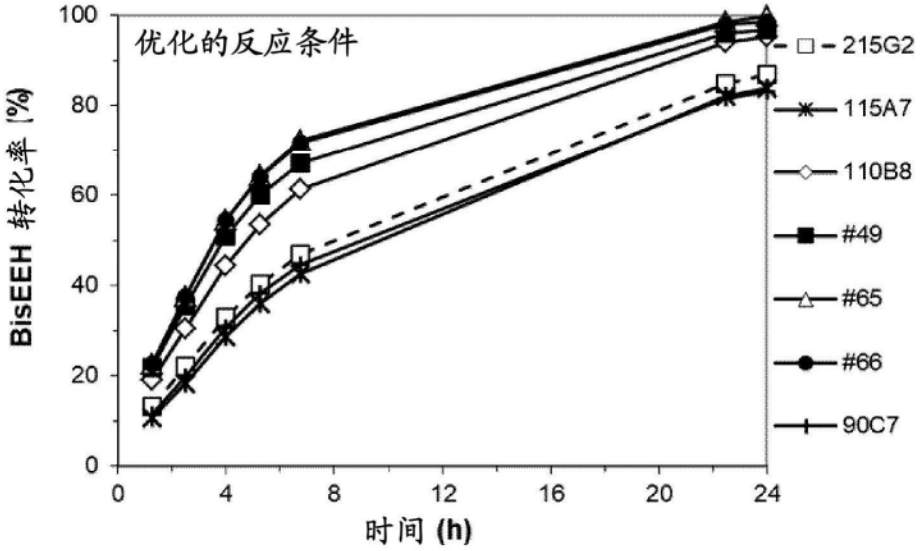

优化的反应条件
215G2
115A7
110B8
#49
#65
#66
90C7
BisEEH 转化率 (%)
100
80
60
40
20
0
0
4
8
12
16
20
24
时间 (h)

图5

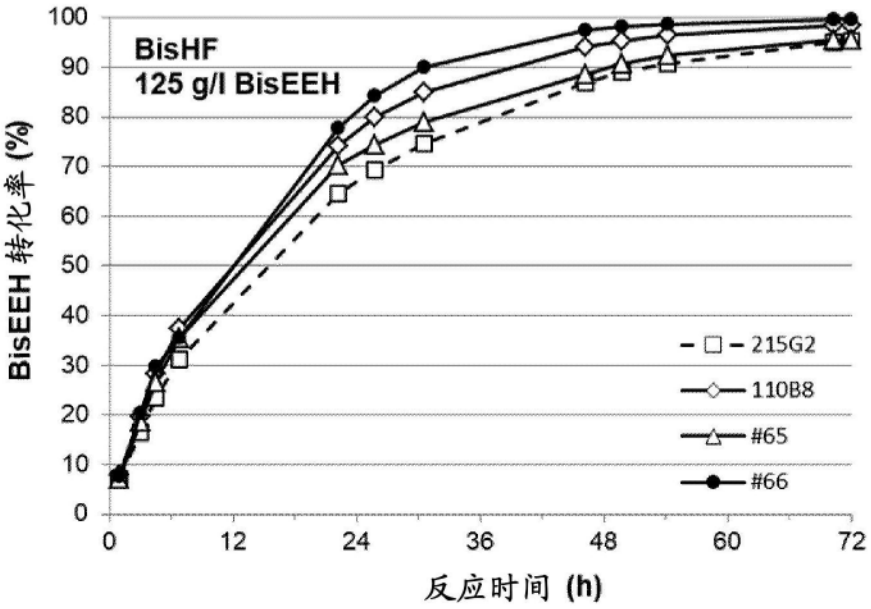

BisHF
125 g/l BisEEH
BisEEH 转化率 (%)
反应时间 (h)
215G2
110B8
#65
#66
0
10
20
30
40
50
60
70
80
90
100
0
12
24
36
48
60
72

图6

```
215G2 SHC       MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDR
SHC #49         MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDR
SHC #65         MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDR
SHC #66         MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDR
90C7 SHC        MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDR
110B8 SHC       MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDR
115A7 SHC       MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDR

215G2 SHC       MEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQ
SHC #49         MEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQ
SHC #65         MEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQ
SHC #66         MEKIRRYLLHEQREDGTWALHPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQ
90C7 SHC        MEKIRRYLLHEQREDGTWALYPGGPPDLDATIEAYVALKYIGMSRDEEPMQKALRFIQSQ
110B8 SHC       MEKIRRYLLHEQREDGTWALHPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQ
115A7 SHC       MEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQ

215G2 SHC       GGIESSRVFTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSIV
SHC #49         GGIESSRVFTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSIV
SHC #65         GGIESSRVFTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSIV
SHC #66         GGIESSRVFTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSIV
90C7 SHC        GGIESSRVFTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSIV
110B8 SHC       GGIESSRVFTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSIV
115A7 SHC       GGIESSRVFTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARTTVVALSIV

215G2 SHC       MSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRVLHGYQKLSVHPFRRAA
SHC #49         MSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRVLHGYQKLSVHPFRRAA
SHC #65         MSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRVLHGYQKLSVHPFRRAA
SHC #66         MSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRVLHGYQKLSVHPFRRAA
90C7 SHC        MSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRVLHGYQKLSVHPFRRAA
110B8 SHC       MSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRVLHGYQKLSVHPFRRAA
115A7 SHC       MSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRVLHGYQKLSVHPFRRAA

215G2 SHC       EIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
SHC #49         EIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
SHC #65         EIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
SHC #66         EIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
90C7 SHC        EIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
110B8 SHC       EIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
115A7 SHC       EIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDKTQHPAFIKGWEGLELYGVELDYG

215G2 SHC       GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKP
SHC #49         GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKP
SHC #65         GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKP
SHC #66         GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKP
90C7 SHC        GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKP
110B8 SHC       GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKP
115A7 SHC       GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKP

215G2 SHC       GGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAY
SHC #49         GGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAY
SHC #65         GGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAY
SHC #66         GGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAY
90C7 SHC        GGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAY
110B8 SHC       GGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAY
115A7 SHC       GGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAY

215G2 SHC       DVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQK
SHC #49         DVDNTSDLPNLTPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQK
SHC #65         DVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQK
SHC #66         DVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQK
90C7 SHC        DVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQK
110B8 SHC       DVDNTSDLPNLTPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQK
115A7 SHC       DVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQK

215G2 SHC       PDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSY
SHC #49         PDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSY
SHC #65         PDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSY
SHC #66         PDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSY
90C7 SHC        PDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSY
110B8 SHC       PDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSY
115A7 SHC       PDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSY

215G2 SHC       EDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTG
SHC #49         EDPAYAGKGASTPSQTTWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTG
SHC #65         EDPAYAGKGASTPSQTTWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTG
SHC #66         EDPAYAGKGASTPSQTTWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTG
90C7 SHC        EDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTG
110B8 SHC       EDPAYAGKGASTPSQTTWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTG
115A7 SHC       EDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTG

215G2 SHC       FPGDFYLGYTMYRHVFPTLALGRYKQAIERR
SHC #49         FPGDFYLGYTMYRHVFPTLALGRYKQAIERR
SHC #65         FPGDFYLGYTMYSHVFPTLALGRYKQAIERR
SHC #66         FPGDFYLGYTMYSHVFPTLALGRYKQAIERR
90C7 SHC        FPGDFYLGYTMYSHVFPTLALGRYKQAIERR
110B8 SHC       FPGDFYLGYTMYRHVFPTLALGRYKQAIERR
115A7 SHC       FPGDFYLGYTMYRHVFPTLALGRYKQAIERR
```

图7

```
215G2 SHC              MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDR
215G2 V174I            MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDR
215G2 V174I F601       MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDR
215G2 L37Q             MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLQSNVTMEAEYVLLCHILDRVDRDR
215G2 L37Q V174I       MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLQSNVTMEAEYVLLCHILDRVDRDR

215G2 SHC              MEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQ
215G2 V174I            MEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQ
215G2 V174I F601       MEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQ
215G2 L37Q             MEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQ
215G2 L37Q V174I       MEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQ

215G2 SHC              GGIESSRVFTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSIV
215G2 V174I            GGIESSRVFTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATIVALSIV
215G2 V174I F601       GGIESSRVFTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATIVALSIV
215G2 L37Q             GGIESSRVFTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSIV
215G2 L37Q V174I       GGIESSRVFTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATIVALSIV

215G2 SHC              MSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRVLHGYQKLSVHPFRRAA
215G2 V174I            MSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRVLHGYQKLSVHPFRRAA
215G2 V174I F601       MSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRVLHGYQKLSVHPFRRAA
215G2 L37Q             MSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRVLHGYQKLSVHPFRRAA
215G2 L37Q V174I       MSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRVLHGYQKLSVHPFRRAA

215G2 SHC              EIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
215G2 V174I            EIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
215G2 V174I F601       EIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
215G2 L37Q             EIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
215G2 L37Q V174I       EIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG

215G2 SHC              GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKP
215G2 V174I            GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKP
215G2 V174I F601       GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKP
215G2 L37Q             GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKP
215G2 L37Q V174I       GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKP

215G2 SHC              GGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAY
215G2 V174I            GGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAY
215G2 V174I F601       GGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAY
215G2 L37Q             GGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAY
215G2 L37Q V174I       GGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAY

215G2 SHC              DVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQK
215G2 V174I            DVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQK
215G2 V174I F601       DVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQK
215G2 L37Q             DVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQK
215G2 L37Q V174I       DVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQK

215G2 SHC              PDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSY
215G2 V174I            PDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSY
215G2 V174I F601       PDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSY
215G2 L37Q             PDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSY
215G2 L37Q V174I       PDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSY

215G2 SHC              EDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTG
215G2 V174I            EDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTG
215G2 V174I F601       EDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTG
215G2 L37Q             EDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTG
215G2 L37Q V174I       EDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTG

215G2 SHC              FPGDFYLGYTMYRHVFPTLALGRYKQAIERR
215G2 V174I            FPGDFYLGYTMYRHVFPTLALGRYKQAIERR
215G2 V174I F601       YPGDFYLGYTMYRHVFPTLALGRYKQAIERR
215G2 L37Q             FPGDFYLGYTMYRHVFPTLALGRYKQAIERR
215G2 L37Q V174I       YPGDFYLGYTMYRHVFPTLALGRYKQAIERR
```

图8

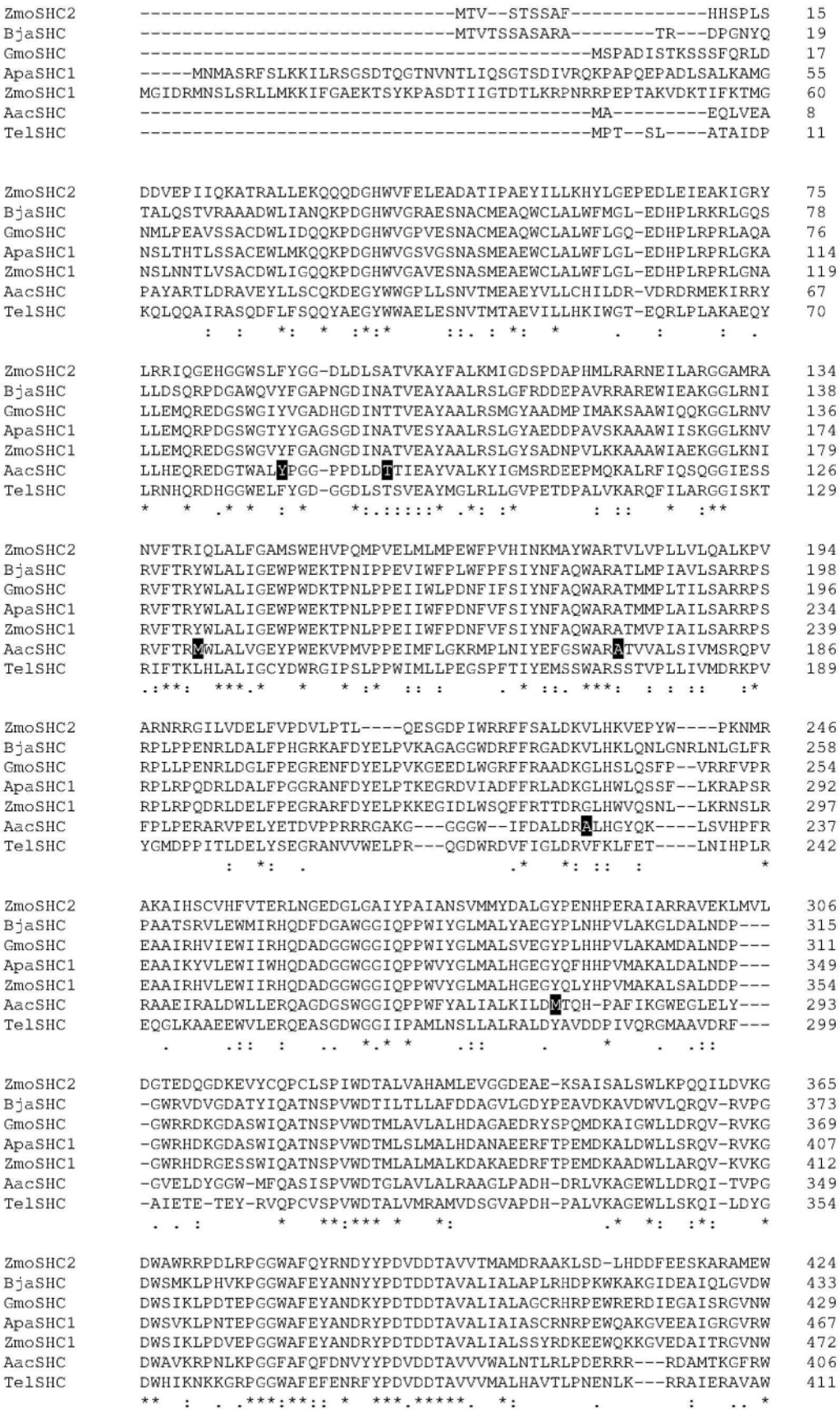

ZmoSHC2 ----------MTV--STSSAF----------HHSPLS 15
BjaSHC ----------MTVTSSASARA--------TR---DPGNYQ 19
GmoSHC ----------MSPADISTKSSSFQRLD 17
ApaSHC1 -----MNMASRFSLKKILRSGSDTQGTNVNTLIQSGTSDIVRQKPAPQEPADLSALKAMG 55
ZmoSHC1 MGIDRMNSLSRLLMKKIFGAEKTSYKPASDTIIGTDTLKRPNRRPEPTAKVDKTIFKTMG 60
AacSHC ----------MA---------EQLVEA 8
TelSHC ----------MPT--SL----ATAIDP 11
ZmoSHC2 DDVEPIIQKATRALLEKQQQDGHWVFELEADATIPAEYILLKHYLGEPEDLEIEAKIGRY 75
BjaSHC TALQSTVRAAADWLIANQKPDGHWVGRAESNACMEAQWCLALWFMGL-EDHPLRKRLGQS 78
GmoSHC NMLPEAVSSACDWLIDQQKPDGHWVGPVESNACMEAQWCLALWFLGQ-EDHPLRPRLAQA 76
ApaSHC1 NSLTHTLSSACEWLMKQQKPDGHWVGSVGSNASMEAEWCLALWFLGL-EDHPLRPRLGKA 114
ZmoSHC1 NSLNNTLVSACDWLIGQQKPDGHWVGAVESNASMEAEWCLALWFLGL-EDHPLRPRLGNA 119
AacSHC PAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDR-VDRDRMEKIRRY 67
TelSHC KQLQQAIRASQDFLFSQQYAEGYWWAELESNVTMTAEVILLHKIWGT-EQRLPLAKAEQY 70
ZmoSHC2 LRRIQGEHGGWSLFYGG-DLDLSATVKAYFALKMIGDSPDAPHMLRARNEILARGGAMRA 134
BjaSHC LLDSQRPDGAWQVYFGAPNGDINATVEAYAALRSLGFRDDEPAVRRAREWIEAKGGLRNI 138
GmoSHC LLEMQREDGSWGIYVGADHGDINTTVEAYAALRSMGYAADMPIMAKSAAWIQQKGGLRNV 136
ApaSHC1 LLEMQRPDGSWGTYYGAGSGDINATVESYAALRSLGYAEDDPAVSKAAAWIISKGGLKNV 174
ZmoSHC1 LLEMQREDGSWGVYFGAGNGDINATVEAYAALRSLGYSADNPVLKKAAAWIAEKGGLKNI 179
AacSHC LLHEQREDGTWALYPGG-PPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESS 126
TelSHC LRNHQRDHGGWELFYGD-GGDLSTSVEAYMGLRLLGVPETDPALVKARQFILARGGISKT 129
ZmoSHC2 NVFTRIQLALFGAMSWEHVPQMPVELMLMPEWFPVHINKMAYWARTVLVPLLVLQALKPV 194
BjaSHC RVFTRYWLALIGEWPWEKTPNIPPEVIWFPLWFPFSIYNFAQWARATLMPIAVLSARRPS 198
GmoSHC RVFTRYWLALIGEWPWDKTPNLPPEIIWLPDNFIFSIYNFAQWARATMMPLTILSARRPS 196
ApaSHC1 RVFTRYWLALIGEWPWEKTPNLPPEIIWFPDNFVFSIYNFAQWARATMMPLAILSARRPS 234
ZmoSHC1 RVFTRYWLALIGEWPWEKTPNLPPEIIWFPDNFVFSIYNFAQWARATMVPIAILSARRPS 239
AacSHC RVFTRMWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSIVMSRQPV 186
TelSHC RIFTKLHLALIGCYDWRGIPSLPPWIMLLPEGSPFTIYEMSSWARSSTVPLLIVMDRKPV 189
ZmoSHC2 ARNRRGILVDELFVPDVLPTL----QESGDPIWRRFFSALDKVLHKVEPYW----PKNMR 246
BjaSHC RPLPPENRLDALFPHGRKAFDYELPVKAGAGGWDRFFRGADKVLHKLQNLGNRLNLGLFR 258
GmoSHC RPLLPENRLDGLFPEGRENFDYELPVKGEEDLWGRFFRAADKGLHSLQSFP--VRRFVPR 254
ApaSHC1 RPLRPQDRLDALFPGGRANFDYELPTKEGRDVIADFFRLADKGLHWLQSSF--LKRAPSR 292
ZmoSHC1 RPLRPQDRLDELFPEGRARFDYELPKKEGIDLWSQFFRTTDRGLHWVQSNL--LKRNSLR 297
AacSHC FPLPERARVPELYETDVPPRRRGAKG---GGGW--IFDALDRALHGYQK----LSVHPFR 237
TelSHC YGMDPPITLDELYSEGRANVVWELPR---QGDWRDVFIGLDRVFKLFET----LNIHPLR 242
ZmoSHC2 AKAIHSCVHFVTERLNGEDGLGAIYPAIANSVMMYDALGYPENHPERAIARRAVEKLMVL 306
BjaSHC PAATSRVLEWMIRHQDFDGAWGGIQPPWIYGLMALYAEGYPLNHPVLAKGLDALNDP--- 315
GmoSHC EAAIRHVIEWIIRHQDADGGWGGIQPPWIYGLMALSVEGYPLHHPVLAKAMDALNDP--- 311
ApaSHC1 EAAIKYVLEWIIWHQDADGGWGGIQPPWVYGLMALHGEGYQFHHPVMAKALDALNDP--- 349
ZmoSHC1 EAAIRHVLEWIIRHQDADGGWGGIQPPWVYGLMALHGEGYQLYHPVMAKALSALDDP--- 354
AacSHC RAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQH-PAFIKGWEGLELY--- 293
TelSHC EQGLKAAEEWVLERQEASGDWGGIIPAMLNSLLALRALDYAVDDPIVQRGMAAVDRF--- 299
ZmoSHC2 DGTEDQGDKEVYCQPCLSPIWDTALVAHAMLEVGGDEAE-KSAISALSWLKPQQILDVKG 365
BjaSHC -GWRVDVGDATYIQATNSPVWDTILTLLAFDDAGVLGDYPEAVDKAVDWVLQRQV-RVPG 373
GmoSHC -GWRRDKGDASWIQATNSPVWDTMLAVLALHDAGAEDRYSPQMDKAIGWLLDRQV-RVKG 369
ApaSHC1 -GWRHDKGDASWIQATNSPVWDTMLSLMALHDANAEERFTPEMDKALDWLLSRQV-RVKG 407
ZmoSHC1 -GWRHDRGESSWIQATNSPVWDTMLALMALKDAKAEDRFTPEMDKAADWLLARQV-KVKG 412
AacSHC -GVELDYGGW-MFQASISPVWDTGLAVLALRAAGLPADH-DRLVKAGEWLLDRQI-TVPG 349
TelSHC -AIETE-TEY-RVQPCVSPVWDTALVMRAMVDSGVAPDH-PALVKAGEWLLSKQI-LDYG 354
ZmoSHC2 DWAWRRPDLRPGGWAFQYRNDYYPDVDDTAVVTMAMDRAAKLSD-LHDDFEESKARAMEW 424
BjaSHC DWSMKLPHVKPGGWAFEYANNYYPDTDDTAVALIALAPLRHDPKWKAKGIDEAIQLGVDW 433
GmoSHC DWSIKLPDTEPGGWAFEYANDKYPDTDDTAVALIALAGCRHRPEWRERDIEGAISRGVNW 429
ApaSHC1 DWSVKLPNTEPGGWAFEYANDRYPDTDDTAVALIAIASCRNRPEWQAKGVEEAIGRGVRW 467
ZmoSHC1 DWSIKLPDVEPGGWAFEYANDRYPDTDDTAVALIALSSYRDKEEWQKKGVEDAITRGVNW 472
AacSHC DWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVWALNTLRLPDERRR---RDAMTKGFRW 406
TelSHC DWHIKNKKGRPGGWAFEFENRFYPDVDDTAVVVMALHAVTLPNENLK---RRAIERAVAW 411

图9A

```
ZmoSHC2      TIGMQSDNGGWGAFDANNSYTYLNNIPFADHGALLDPPTVDVSARCVSMMAQAGISITDP   484
BjaSHC       LIGMQSQGGGWGAFDKDNNQKILTKIPFCDYGEALDPPSVDVTAHIIEAFGKLGISRNHP   493
GmoSHC       LLAMQSSSGGWGAFDKDNNRSILTKIPFCDFGEALDPPSVDVTAHVLEAFGLLGISRNHP   489
ApaSHC1      LVAMQSSCGGWGAFDKDNNKSILAKIPFCDFGEALDPPSVDVTAHVLEAFGLLGLPRDLP   527
ZmoSHC1      LIAMQSECGGWGAFDKDNNRSILSKIPFCDFGESIDPPSVDVTAHVLEAFGTLGLSRDMP   532
AacSHC       IVGMQSSNGGWGAYDVDNTSDLPNHIPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWK   466
TelSHC       IASMQCRPGGWAAFDVDNDQDWLNGIPYGDLKAMIDPNTADVTARVLEMVGRCQLAFDRV   471
               .**.  ***.*:* :*       **: *     ** : **:*: :. ..

ZmoSHC2      KMKAAVDYLLKEQEEDGSWFGRWGVNYIYGTWSALCALNVAALPHDHLAVQKAVAWLKTI   544
BjaSHC       SMVQALDYIRREQEPSGPWFGRWGVNYVYGTGAVLPALAAIGEDMTQPYIGRACDWLVAH   553
GmoSHC       SVQKALAYIRSEQERNGAWFGRWGVNYVYGTGAVLPALAAIGEDMTQPYIVRACDWLMSV   549
ApaSHC1      CIQRGLAYIRKEQDPTGPWFGRWGVNYLYGTGAVLPALAALGEDMTQPYISKACDWLINC   587
ZmoSHC1      VIQKAIDYVRSEQEAEGAWFGRWGVNYIYGTGAVLPALAAIGEDMTQPYITKACDWLVAH   592
AacSHC       VIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQH   526
TelSHC       ALDRALAYLRNEQEPEGCWFGRWGVNYLYGTSGVLTALSLVAPRYDRWRIRRAAEWLMQC   531
              :  .: *:  **.  * *********:*** ..: **   .    .  : :*  *:

ZmoSHC2      QNEDGGWGENCDSYALDYSGYEPMDSTASQTAWALLGLMAVGEAN----SEAVTKGINWL   600
BjaSHC       QQADGGWGESCASYM-DVSAVGRGTTTASQTAWALMALLAANRPQD---KDAIERGCMWL   609
GmoSHC       QQENGGWGESCASYM-DINAVGHGVATASQTAWALIGLLAAKRPKD---REAIARGCQFL   605
ApaSHC1      QQENGGWGESCASYM-EVSSIGHGATTPSQTAWALMGLIAANRPQD---YEAIAKGCRYL   643
ZmoSHC1      QQEDGGWGESCSSYM-EIDSIGKGPTTPSQTAWALMGLIAANRPED---YEAIAKGCHYL   648
AacSHC       QNPDGGWGEDCRSYE-DPAYAGKGASTPSQTAWALMALIAGGRAE----SEAARRGVQYL   581
TelSHC       QNADGGWGETCWSYH-DPSLKGKGDSTASQTAWAIIGLLAAGDATGDYATEAIERGIAYL   590
             *: :***** * **  :        :* ******::.*:*          :*  :*  :*

ZmoSHC2      AQNQDEEGLWKEDYYSGGGFPR-------------------------VFYLRYHGYSKY   634
BjaSHC       VERQ-SAGTWDEPEFTGTGFPGYGVGQTIKLNDPALSQRLMQGPELSRAFMLRYGMYRHY   668
GmoSHC       IERQ-EDGSWTEEEYTGTGFPGYGVGQAIKLDDPSLPDRLLQGAELSRAFMLRYDLYRQY   664
ApaSHC1      IDLQEEDGSWNEEEFTGTGFPGYGVGQTIKLDDPAISKRLMQGAELSRAFMLRYDLYRQL   703
ZmoSHC1      IDRQEQDGSWKEEEFTGTGFPGYGVGQTIKLDDPALSKRLLQGAELSRAFMLRYDFYRQF   708
AacSHC       VETQRPDGGWDEPYYTGTGFPG-------------------------DFYLGYTMYRHV   615
TelSHC       LETQRPDGTWHEDYFTGTGFPC-------------------------HFYLKYHYYQQH   624
              : *   * * *  ::* ***                            * * *  * :

ZmoSHC2      FPLWALARYRNLKKANQPIVHYGM 658
BjaSHC       FPLMALGRALRPQSHS-------- 684
GmoSHC       FPVMALSRARRMMKEDASAAA--- 685
ApaSHC1      FPIIALSRASRLIKLGN------- 720
ZmoSHC1      FPIMALSRAERLIDLNN------- 725
AacSHC       FPTLALGRYKQAIERR-------- 631
TelSHC       FPLTALGRYARWRNLLAT------ 642
             **  **.*  .  .
```

图9A续

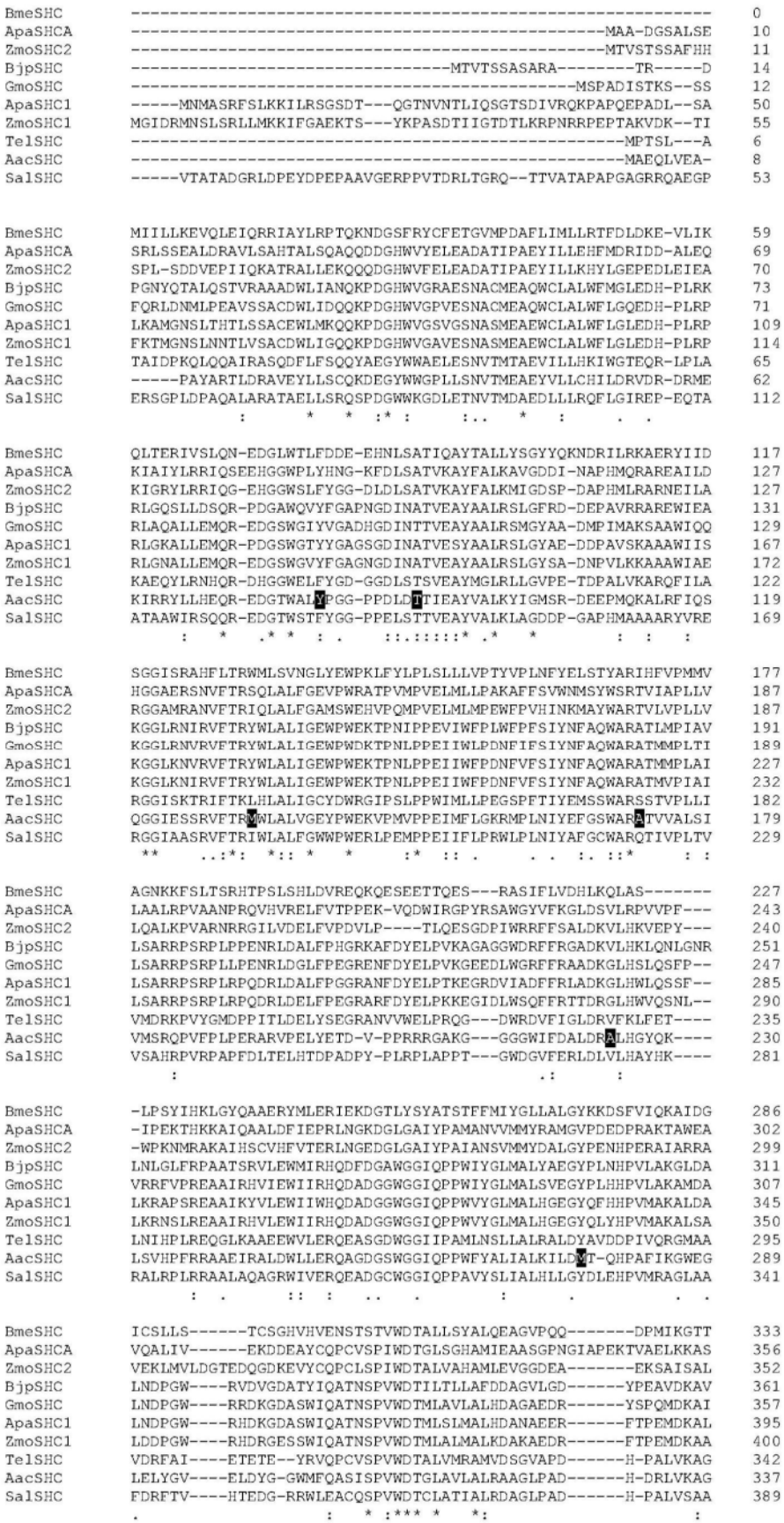

BmeSHC ---------- 0
ApaSHCA ----------MAA-DGSALSE 10
ZmoSHC2 ----------MTVSTSSAFHH 11
BjpSHC ----------MTVTSSASARA--------TR-----D 14
GmoSHC ----------MSPADISTKS--SS 12
ApaSHC1 -----MNMASRFSLKKILRSGSDT---QGTNVNTLIQSGTSDIVRQKPAPQEPADL--SA 50
ZmoSHC1 MGIDRMNSLSRLLMKKIFGAEKTS---YKPASDTIIGTDTLKRPNRRPEPTAKVDK--TI 55
TelSHC ----------MPTSL---A 6
AacSHC ----------MAEQLVEA- 8
SalSHC -----VTATADGRLDPEYDPEPAAVGERPPVTDRLTGRQ--TTVATAPAPGAGRRQAEGP 53
BmeSHC MIILLKEVQLEIQRRIAYLRPTQKNDGSFRYCFETGVMPDAFLIMLLRTFDLDKE-VLIK 59
ApaSHCA SRLSSEALDRAVLSAHTALSQAQQDDGHWVYELEADATIPAEYILLEHFMDRIDD-ALEQ 69
ZmoSHC2 SPL-SDDVEPIIQKATRALLEKQQQDGHWVFELEADATIPAEYILLKHYLGEPEDLEIEA 70
BjpSHC PGNYQTALQSTVRAAADWLIANQKPDGHWVGRAESNACMEAQWCLALWFMGLEDH-PLRK 73
GmoSHC FQRLDNMLPEAVSSACDWLIDQQKPDGHWVGPVESNACMEAQWCLALWFLGQEDH-PLRP 71
ApaSHC1 LKAMGNSLTHTLSSACEWLMKQQKPDGHWVGSVGSNASMEAEWCLALWFLGLEDH-PLRP 109
ZmoSHC1 FKTMGNSLNNTLVSACDWLIGQQKPDGHWVGAVESNASMEAEWCLALWFLGLEDH-PLRP 114
TelSHC TAIDPKQLQQAIRASQDFLFSQQYAEGYWWAELESNVTMTAEVILLHKIWGTEQR-LPLA 65
AacSHC -----PAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDR-DRME 62
SalSHC ERSGPLDPAQALARATAELLSRQSPDGWWKGDLETNVTMDAEDLLLRQFLGIREP-EQTA 112
BmeSHC QLTERIVSLQN-EDGLWTLFDDE-EHNLSATIQAYTALLYSGYYQKNDRILRKAERYIID 117
ApaSHCA KIAIYLRRIQSEEHGGWPLYHNG-KFDLSATVKAYFALKAVGDDI-NAPHMQRAREAILD 127
ZmoSHC2 KIGRYLRRIQG-EHGGWSLFYGG-DLDLSATVKAYFALKMIGDSP-DAPHMLRARNEILA 127
BjpSHC RLGQSLLDSQR-PDGAWQVYFGAPNGDINATVEAYAALRSLGFRD-DEPAVRRAREWIEA 131
GmoSHC RLAQALLEMQR-EDGSWGIYVGADHGDINTTVEAYAALRSMGYAA-DMPIMAKSAAWIQQ 129
ApaSHC1 RLGKALLEMQR-PDGSWGTYYGAGSGDINATVESYAALRSLGYAE-DDPAVSKAAAWIIS 167
ZmoSHC1 RLGNALLEMQR-EDGSWGVYFGAGNGDINATVEAYAALRSLGYSA-DNPVLKKAAAWIAE 172
TelSHC KAEQYLRNHQR-DHGGWELFYGD-GGDLSTSVEAYMGLRLLGVPE-TDPALVKARQFILA 122
AacSHC KIRRYLLHEQR-EDGTWAIYPGG-PPDLDTTIEAYVALKYIGMSR-DEEPMQKALRFIQS 119
SalSHC ATAAWIRSQQR-EDGTWSTFYGG-PPELSTTVEAYVALKLAGDDP-GAPHMAAAARYVRE 169
BmeSHC SGGISRAHFLTRWMLSVNGLYEWPKLFYLPLSLLLVPTYVPLNFYELSTYARIHFVPMMV 177
ApaSHCA HGGAERSNVFTRSQLALFGEVPWRATPVMPVELMLLPAKAFFSVWNMSYWSRTVIAPLLV 187
ZmoSHC2 RGGAMRANVFTRIQLALFGAMSWEHVPQMPVELMLMPEWFPVHINKMAYWARTVLVPLLV 187
BjpSHC KGGLRNIRVFTRYWLALIGEWPWEKTPNIPPEVIWFPLWFPFSIYNFAQWARATLMPIAV 191
GmoSHC KGGLRNVRVFTRYWLALIGEWPWDKTPNLPPEIIWLPDNFIFSIYNFAQWARATMMPLTI 189
ApaSHC1 KGGLKNVRVFTRYWLALIGEWPWEKTPNLPPEIIWFPDNFVFSIYNFAQWARATMMPLAI 227
ZmoSHC1 KGGLKNIRVFTRYWLALIGEWPWEKTPNLPPEIIWFPDNFVFSIYNFAQWARATMVPIAI 232
TelSHC RGGISKTRIFTKLHLALIGCYDWRGIPSLPPWIMLLPEGSPFTIYEMSSWARSSTVPLLI 182
AacSHC QGGIESSRVFTRMWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARATVVALSI 179
SalSHC RGGIAASRVFTRIWLALFGWWPWERLPEMPPEIIFLPRWLPLNIYAFGCWARQTIVPLTV 229
BmeSHC AGNKKFSLTSRHTPSLSHLDVREQKQESEETTQES---RASIFLVDHLKQLAS------- 227
ApaSHCA LAALRPVAANPRQVHVRELFVTPPEK-VQDWIRGPYRSAWGYVFKGLDSVLRPVVPF--- 243
ZmoSHC2 LQALKPVARNRRGILVDELFVPDVLP----TLQESGDPIWRRFFSALDKVLHKVEPY--- 240
BjpSHC LSARRPSRPLPPENRLDALFPHGRKAFDYELPVKAGAGGWDRFFRGADKVLHKLQNLGNR 251
GmoSHC LSARRPSRPLLPENRLDGLFPEGRENFDYELPVKGEEDLWGRFFRAADKGLHSLQSFP-- 247
ApaSHC1 LSARRPSRPLRPQDRLDALFPGGRANFDYELPTKEGRDVIADFFRLADKGLHWLQSSF-- 285
ZmoSHC1 LSARRPSRPLRPQDRLDELFPEGRARFDYELPKKEGIDLWSQFFRTTDRGLHWVQSNL-- 290
TelSHC VMDRKPVYGMDPPITLDELYSEGRANVVWELPRQG---DWRDVFIGLDRVFKLFET---- 235
AacSHC VMSRQPVFPLPERARVPELYETD-V-PPRRRGAKG---GGGWIFDALDRALHGYQK---- 230
SalSHC VSAHRPVRPAPFDLTELHTDPADPY-PLRPLAPPT---GWDGVFERLDLVLHAYHK---- 281
BmeSHC -LPSYIHKLGYQAAERYMLERIEKDGTLYSYATSTFFMIYGLLALGYKKDSFVIQKAIDG 286
ApaSHCA -IPEKTHKKAIQAALDFIEPRLNGKDGLGAIYPAMANVVMMYRAMGVPDEDPRAKTAWEA 302
ZmoSHC2 -WPKNMRAKAIHSCVHFVTERLNGEDGLGAIYPAIANSVMMYDALGYPENHPERAIARRA 299
BjpSHC LNLGLFRPAATSRVLEWMIRHQDFDGAWGGIQPPWIYGLMALYAEGYPLNHPVLAKGLDA 311
GmoSHC VRRFVPREAAIRHVIEWIIRHQDADGGWGGIQPPWIYGLMALSVEGYPLHHPVLAKAMDA 307
ApaSHC1 LKRAPSREAAIKYVLEWIIWHQDADGGWGGIQPPWVYGLMALHGEGYQFHHPVMAKALDA 345
ZmoSHC1 LKRNSLREAAIRHVLEWIIRHQDADGGWGGIQPPWVYGLMALHGEGYQLYHPVMAKALSA 350
TelSHC LNIHPLREQGLKAAEEWVLERQEASGDWGGIIPAMLNSLLALRALDYAVDDPIVQRGMAA 295
AacSHC LSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDYT-QHPAFIKGWEG 289
SalSHC RALRPLRRAALAQAGRWIVERQEADGCWGGIQPPAVYSLIALHLLGYDLEHPVMRAGLAA 341
BmeSHC ICSLLS------TCSGHVHVENSTSTVWDTALLSYALQEAGVPQQ-------DPMIKGTT 333
ApaSHCA VQALIV------EKDDEAYCQPCVSPIWDTGLSGHAMIEAASGPNGIAPEKTVAELKKAS 356
ZmoSHC2 VEKLMVLDGTEDQGDKEVYCQPCLSPIWDTALVAHAMLEVGGDEA-------EKSAISAL 352
BjpSHC LNDPGW----RVDVGDATYIQATNSPVWDTILTLLAFDDAGVLGD-------YPEAVDKAV 361
GmoSHC LNDPGW----RRDKGDASWIQATNSPVWDTMLAVLALHDAGAEDR------YSPQMDKAI 357
ApaSHC1 LNDPGW----RHDKGDASWIQATNSPVWDTMLSLMALHDANAEER------FTPEMDKAL 395
ZmoSHC1 LDDPGW----RHDRGESSWIQATNSPVWDTMLALMALKDAKAEDR------FTPEMDKAA 400
TelSHC VDRFAI----ETETE--YRVQPCVSPVWDTALVMRAMVDSGVAPD------H-PALVKAG 342
AacSHC LELYGV----ELDYG-GWMFQASISPVWDTGLAVLALRAAGLPAD------H-DRLVKAG 337
SalSHC FDRFTV----HTEDG-RRWLEACQSPVWDTCLATIALRDAGLPAD------H-PALVSAA 389

图9B

```
BmeSHC      RYLKKRQHT-KLGDWQFHNPNTAPGGWGFSDINTNNPDLDDTSAAIRALSRRAQTDT---    389
ApaSHCA     AWLRSKQILNVKGDWAVRNPNLAPGGWAFQYGNDYYPDVDDTAVVGMLLHREGDPTN---    413
ZmoSHC2     SWLKPQQILDVKGDWAWRRPDLRPGGWAFQYRNDYYPDVDDTAVVTMAMDRAAKLSDL-H    411
BjpSHC      DWVLQRQVR-VPGDWSMKLPHVKPGGWAFEYANNYYPDTDDTAVALIALAPLRHDPKWKA    420
GmoSHC      GWLLDRQVR-VKGDWSIKLPDTEPGGWAFEYANDKYPDTDDTAVALIALAGCRHRPEWRE    416
ApaSHC1     DWLLSRQVR-VKGDWSVKLPNTEPGGWAFEYANDRYPDTDDTAVALIAIASCRNRPEWQA    454
ZmoSHC1     DWLLARQVK-VKGDWSIKLPDVEPGGWAFEYANDRYPDTDDTAVALIALSSYRDKEEWQK    459
TelSHC      EWLLSKQIL-DYGDWHIKNKKGRPGGWAFEFENRFYPDVDDTAVVVMALHAVTLPNEN--    399
AacSHC      EWLLDRQIT-VPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDER--    394
SalSHC      DWMLAEEIR-RPGDWSVQRPRLAPGGWAFEFENDNYPDIDDTAEVVLALKRVAHPDRA--    446
             ::  .:     ***  :     ***:.*.  *   ** ***: .   :

BmeSHC      -DYLESWQRGINWLLSMQNKDGGFAAFEKNTDSILFTYLPLENAKDAATDPATADLTGRV    448
ApaSHCA     ---AEAIERARTWIVGMQSTDGGWGAFDIDNNKDVLNHIPFADHG-ALLDPPTADVTARC    469
ZmoSHC2     DDFEESKARAMEWTIGMQSDNGGWGAFDANNSYTYLNNIPFADHG-ALLDPPTVDVSARC    470
BjpSHC      KGIDEAIQLGVDWLIGMQSQGGGWGAFDKDNNQKILTKIPFCDYG-EALDPPSVDVTAHI    479
GmoSHC      RDIEGAISRGVNWLLAMQSSSGGWGAFDKDNNRSILTKIPFCDFG-EALDPPSVDVTAHV    475
ApaSHC1     KGVEEAIGRGVRWLVAMQSSCGGWGAFDKDNNKSILAKIPFCDFG-EALDPPSVDVTAHV    513
ZmoSHC1     KGVEDAITRGVNWLIAMQSECGGWGAFDKDNNRSILSKIPFCDFG-ESIDPPSVDVTAHV    518
TelSHC      -LKRRAIERAVAWIASMQCRPGGWAAFDVDNDQDWLNGIPYGDLK-AMIDPNTADVTARV    457
AacSHC      -RRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHIPFCDFG-EVTDPPSEDVTAHV    452
SalSHC      -RIDGAVRRGVEWNLGMQSRNGAWGAFDVDNTSTLPNKLPFCDFG-EVVDPPSADVTAHV    504
                  :   .  *  .**   *.:.*:: :.       :*  :      ** : *::.:

BmeSHC      LECLGNFAGMNKSHPSIKAAVKWLFDHQLDNGSWYGRWGVCYIYGTWAAITGLRAVGVSA    508
ApaSHCA     ISFLAQL-RNPEDEPVIQRGLEYLRKEQEKDGSWFGRWGTNYIYGTWSALCALNAAGVSH    528
ZmoSHC2     VSMMAQA-GISITDPKMKAAVDYLLKEQEEDGSWFGRWGVNYIYGTWSALCALNVAALPH    529
BjpSHC      IEAFGKL-GISRNHPSMVQALDYIRREQEPSGPWFGRWGVNYVYGTGAVLPALAAIGEDM    538
GmoSHC      LEAFGLL-GISRNHPSVQKALAYIRSEQERNGAWFGRWGVNYVYGTGAVLPALAAIGEDM    534
ApaSHC1     LEAFGLL-GLPRDLPCIQRGLAYIRKEQDPTGPWFGRWGVNYLYGTGAVLPALAALGEDM    572
ZmoSHC1     LEAFGTL-GLSRDMPVIQKAIDYVRSEQEAEGAWFGRWGVNYIYGTGAVLPALAAIGEDM    577
TelSHC      LEMVGRC-QLAFDRVALDRALAYLRNEQEPEGCWFGRWGVNYLYGTSGVLTALSLVAPRY    516
AacSHC      LECFGSF-GYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDT    511
SalSHC      VEMLAET-GLAGDR-RTRRGIDWLLKNQEPDGSWFGRWGTNYIYGTGSVLPALVAAGIPG    562
            :. ..              .: ::  .*   * *:****. *:*** ..: .*   .

BmeSHC      SDPRIIKAINWLKSIQQEDGGFGESCYSASLKK-YVPLSFSTPSQTAWALDALMTICPL-    566
ApaSHCA     DDPAVVKAVEWLRSVQRADGGWGEGCESYEGGPHGTY-GESLPSQTAWAVLGLMAAGRR-    586
ZmoSHC2     DHLAVQKAVAWLKTIQNEDGGWGENCDSYALDYSGYFPMDSTASQTAWALLGLMAVGFA-    588
BjpSHC      TQPYIGRACDWLVAHQQADGGWGESCASYMDVS-AVGRGTTTASQTAWALMALLAANRPQ    597
GmoSHC      TQPYIVRACDWLMSVQQENGGWGESCASYMDIN-AVGHGVATASQTAWALIGLLAAKRPK    593
ApaSHC1     TQPYISKACDWLINCQQENGGWGESCASYMEVS-SIGHGATTPSQTAWALMGLIAANRPQ    631
ZmoSHC1     TQPYITKACDWLVAHQQEDGGWGESCSSYME----------------------------    608
TelSHC      DRWRIRRAAEWLMQCQNADGGWGETCWSYHDPS-LKGKGDSTASQTAWAIIGLLAAGDAT    575
AacSHC      REPYIQKALDWVEQHQNPDGGWGEDCRSYEDPA-YAGKGASTPSQTAWALMALIAGGRA-    569
SalSHC      SHPAVRRAVDWLADRQNPDGGWGEDMRSYEDPVRWSGRGDSTASQTAWALMALLAAGEGP    622
                : :*  *:   *. :**:**   *

BmeSHC      ---KDQSVEKGIKFLLNPNLTEQ--QTHYPTGIGLPG-----------------------    598
ApaSHCA     ---DDPAVTRGIAWLADQQDANGEWHEDPYNAVGFPK-----------------------    620
ZmoSHC2     ---NSEAVTKGINWLAQNQDEEGLWKEDYYSGGGFPR-----------------------    622
BjpSHC      ---DKDAIERGCMWLVERQ-SAGTWDEPEFTGTGFPGYGVGQTIKLNDPALSQRLMQGPE    653
GmoSHC      ---DREAIARGCQFLIERQ-EDGSWTEEEYTGTGFPGYGVGQAIKLDDPSLPDRLLQGAE    649
ApaSHC1     ---DYEAIAKGCRYLIDLQEEDGSWNEEEFTGTGFPGYGVGQTIKLDDPAISKRLMQGAE    688
ZmoSHC1     ------------------------------------------------------------    608
TelSHC      GDYATEAIERGIAYLLETQRPDGTWHEDYFTGTGFPC-----------------------    612
AacSHC      ---ESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPG-----------------------    603
SalSHC      DGARSEVVERGVQWLCRTQLPSGSWDEPQFTGTGFPW-----------------------    659

BmeSHC      ---QFYIQYHSYNDIFPLLALAHYAKKHSS---------    625
ApaSHCA     ---VFYLRYHGYKQFFPLMALARYRNLESSNTRRVSFGF    656
ZmoSHC2     ---VFYLRYHGYSKYFPLWALARYRNLKKANQPIVHYGM    658
BjpSHC      LSRAFMLRYGMYRHYFPLMALGRALRPQSHS---------    684
GmoSHC      LSRAFMLRYDLYRQYFPVMALSRARRMMKEDASAAA---    685
ApaSHC1     LSRAFMLRYDLYRQLFPIIALSRASRLIKLGN-------    720
ZmoSHC1     ----------------------------------------    608
TelSHC      ---HFYLKYHYYQQHFPLTALGRYARWRNLLAT------    642
AacSHC      ---DFYLGYTMYRHVFPTLALGRYKQAIERR---------    631
SalSHC      ---DFSINYHLYRLVFPVTALGRYLHGSPLTGGGA----    691
```

图9B续

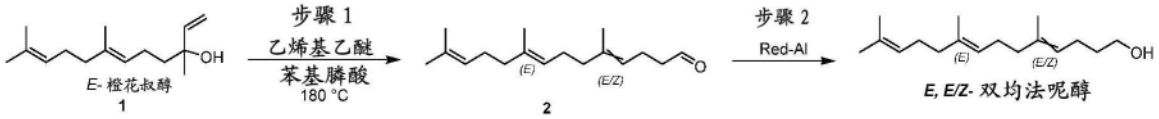

步骤 1
乙烯基乙醚
苯基膦酸
180 °C
步骤 2
Red-Al
OH
E- 橙花叔醇
1
(E)
(E/Z)
O
2
(E)
(E/Z)
OH
E, E/Z- 双均法呢醇

图10

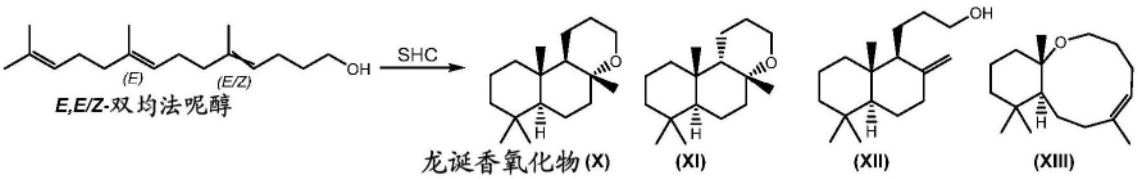

(E)
(E/Z)
OH
E,E/Z-双均法呢醇
SHC
龙诞香氧化物 (X)
(XI)
(XII)
(XIII)

图11

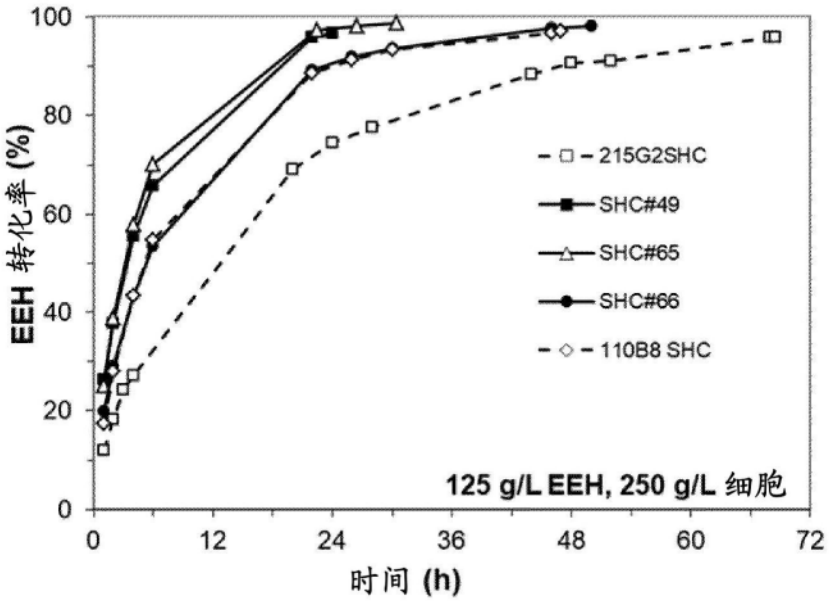

EEH 转化率 (%)
时间 (h)
215G2SHC
SHC#49
SHC#65
SHC#66
110B8 SHC
125 g/L EEH, 250 g/L 细胞

图12

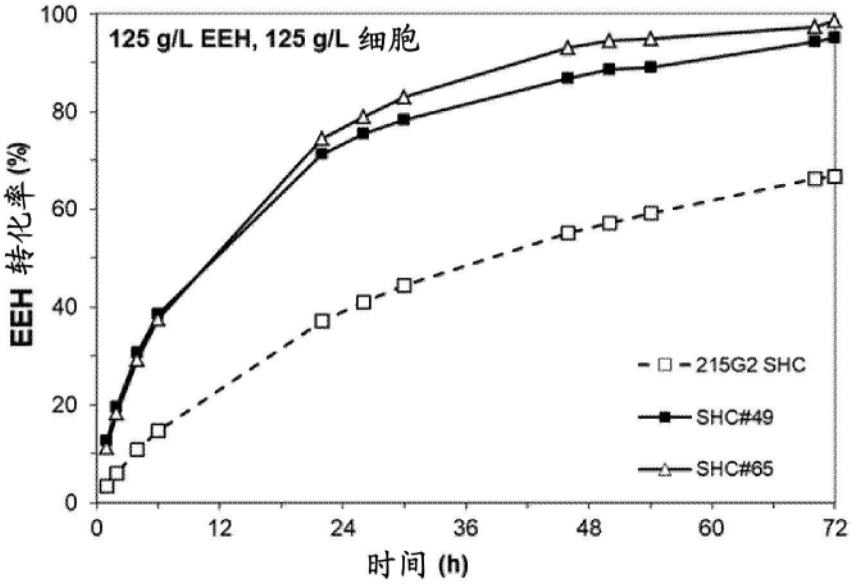

125 g/L EEH, 125 g/L 细胞
EEH 转化率 (%)
时间 (h)
215G2 SHC
SHC#49
SHC#65

图13

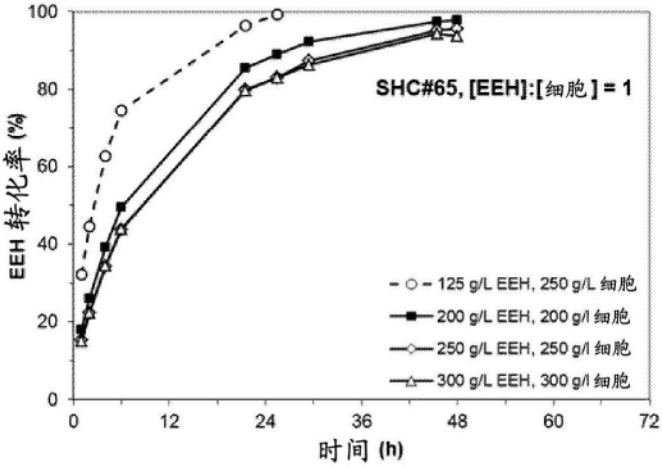

SHC#65, [EEH]:[细胞] = 1
EEH 转化率 (%)
时间 (h)
125 g/L EEH, 250 g/L 细胞
200 g/L EEH, 200 g/l 细胞
250 g/L EEH, 250 g/l 细胞
300 g/L EEH, 300 g/l 细胞

图14

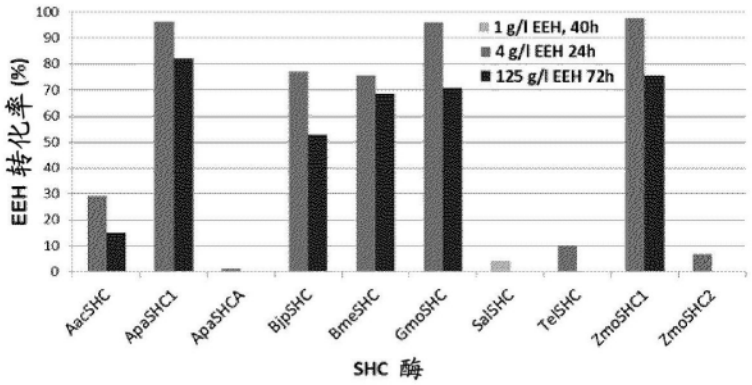

1 g/l EEH, 40h
4 g/l EEH 24h
125 g/l EEH 72h
EEH 转化率 (%)
100
90
80
70
60
50
40
30
20
10
0
AacSHC
ApaSHC1
ApaSHCA
BjpSHC
BmeSHC
GmoSHC
SalSHC
TelSHC
ZmoSHC1
ZmoSHC2
SHC 酶

图15

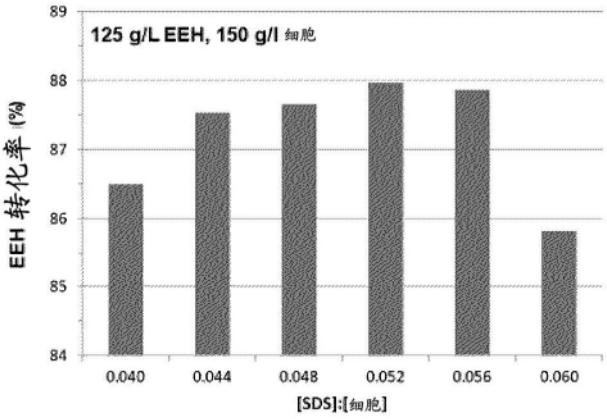

125 g/L EEH, 150 g/l 细胞
EEH 转化率 (%)
89
88
87
86
85
84
0.040
0.044
0.048
0.052
0.056
0.060
[SDS]:[细胞]

图16