



(12) 发明专利申请

(10) 申请公布号 CN 115667538 A

(43) 申请公布日 2023. 01. 31

(21) 申请号 202180038965.8

(22) 申请日 2021.04.14

(30) 优先权数据

2005468.0 2020.04.15 GB

(85) PCT国际申请进入国家阶段日

2022.11.29

(86) PCT国际申请的申请数据

PCT/EP2021/059618 2021.04.14

(87) PCT国际申请的公布数据

W02021/209482 EN 2021.10.21

(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 F·弗拉克斯曼

A·格克

(74) 专利代理机构 中国贸促会专利商标事务所

有限公司 11038

专利代理师 王贵杰

(51) Int.Cl.

C12P 17/18 (2006.01)

C12P 17/02 (2006.01)

C07D 313/08 (2006.01)

C07D 493/08 (2006.01)

C11B 9/00 (2006.01)

权利要求书2页 说明书40页

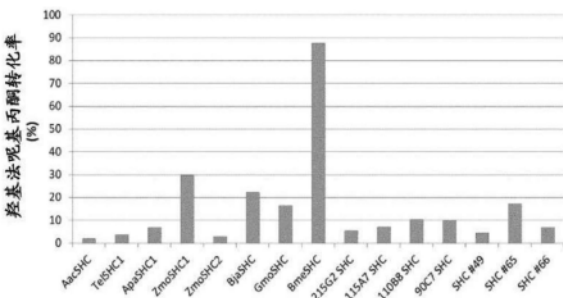
序列表41页 附图1页

(54) 发明名称

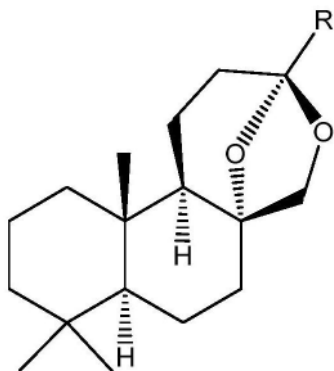
用于制备龙涎缩醛或龙涎缩醛同系物的酶介导的方法

(57) 摘要

一种用于生产龙涎缩醛和龙涎缩醛同系物的酶介导的方法、该方法的产物和所述产物的用途。

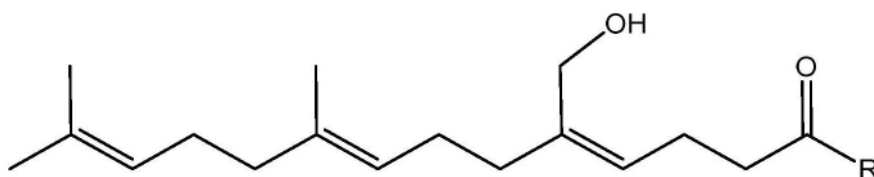


1. 制备式 (I) 的化合物的方法，



式(I)

其中该方法包括使式 (II) 的化合物与角鲨烯-何帕烯环化酶 (SHC) 或酶变体接触，

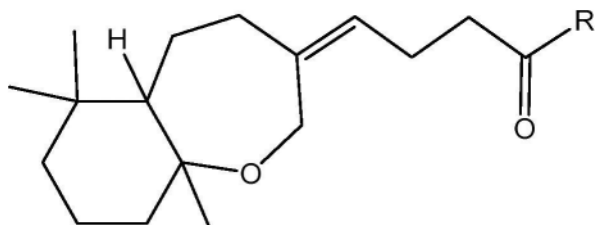


式(II)

其中R为H、甲基或乙基。

2. 权利要求1的方法，其中该方法包括使其中C-8与C-9之间的双键为E-构型且C-4与C-5之间的双键为Z-构型的式 (II) 的化合物与SHC酶或酶变体接触。

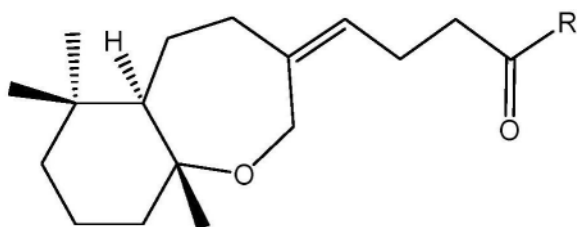
3. 任一前述权利要求的方法，其中式 (III) 的化合物作为副产物制备，



式(III)

其中R为H、甲基或乙基。

4. 权利要求3的方法，其中具有式 (IIIa) 中所示的相对构型的化合物作为副产物制备，



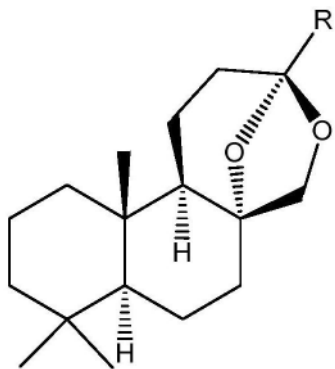
式(IIIa)

其中R为H、甲基或乙基。

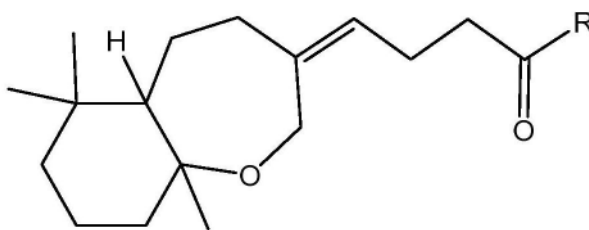
5. 任一前述权利要求的方法，其中R为甲基。

6. 任一前述权利要求的方法，其中该方法还包括纯化式 (I) 的化合物。

7. 组合物,其包含式(I)的化合物和式(III)的化合物:



式(I)



式(III)

其中R为H、甲基或乙基。

8. 权利要求7的组合物,其中R为甲基。

9. 通过或可通过权利要求1至6中任一项的方法得到的组合物,例如其中该组合物为根据权利要求7至8中任一项的组合物。

10. 权利要求7至8中任一项的组合物作为香料组合物的用途或在香料组合物中的用途。

11. 消费品,其包含权利要求7至8中任一项的组合物。

用于制备龙涎缩醛或龙涎缩醛同系物的酶介导的方法

技术领域

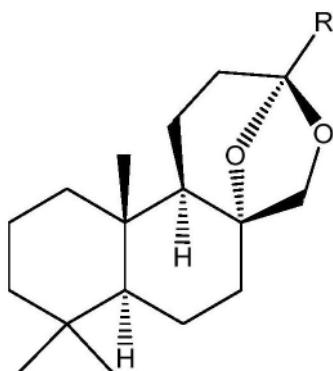
本发明一般涉及使用角鲨烯-何帕烯环化酶 (SHC) 或酶变体制备龙涎缩醛 (Amberketal) 和龙涎缩醛同系物的方法。本发明还涉及通过该方法制备的组合物、所述组合物的不同用途和包含所述组合物的消费品。

背景

龙涎缩醛提供了强烈和持久的琥珀和木香香气,其可单独或与其他木香或琥珀香成分组合用于香料组合物中。龙涎缩醛传统上由泪柏醇通过许多化学转化制备。然而,天然泪柏醇的供应是有限的。因此,期望提供一种新的有效且具有成本效益的合成路线来获得龙涎缩醛和龙涎缩醛同系物。

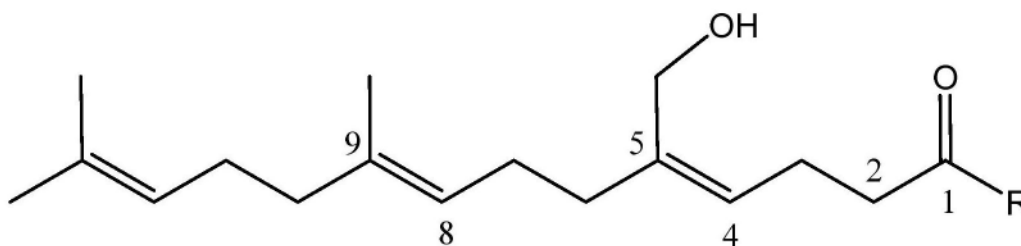
概述

根据本发明的第一个方面,提供了一种用于制备式 (I) 的化合物的方法,



式(I)

其中该方法包括使式 (II) 的化合物与角鲨烯-何帕烯环化酶 (SHC) 或酶变体接触,



式(II)

其中R为H、甲基或乙基。

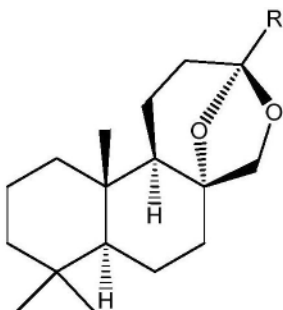
在某些实施方案中,该方法包括使其中C-8与C-9之间的双键为E-构型且C-4与C-5之间的双键为Z-构型的式 (I) 的化合物(式 (II) 的E,Z-化合物)与角鲨烯-何帕烯环化酶 (SHC) (野生型或变体酶) 接触。

在某些实施方案中,该方法包括使包含其中两个双键均为E-构型的式 (II) 的化合物(E,E-化合物)和其中C-8与C-9之间的双键为E-构型且C-4与C-5之间的双键为Z-构型的式 (II) 的化合物(E,Z-化合物)的混合物与角鲨烯-何帕烯环化酶 (SHC) 或酶变体接触。

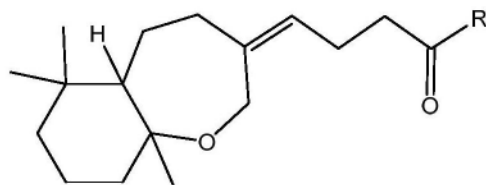
在某些实施方案中,E,Z-化合物与E,E-化合物的重量比在约99:1至约10:90的范

围。例如,式(II)的E,Z化合物与式(II)的E,E-化合物的重量比可以在约95:5至约50:50,或约80:20至约50:50,或约80:20至约60:40的范围。

根据本发明的第二个方面,提供了一种组合物,其包含式(I)的化合物和式(III)的化合物、主要由式(I)的化合物和式(III)的化合物组成、或由式(I)的化合物和式(III)的化合物组成,



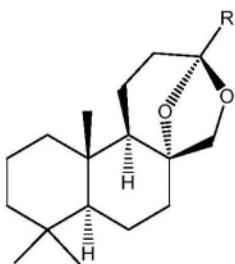
式(I)



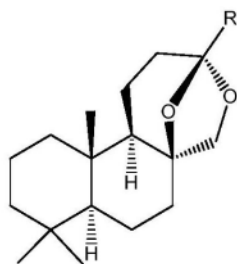
式(III)

其中R为H、甲基或乙基。

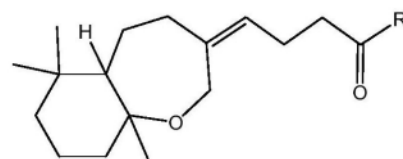
根据本发明的第三个方面,提供了一种组合物,其包含式(I)的化合物、式(IV)的化合物和式(III)的化合物、主要由式(I)的化合物、式(IV)的化合物和式(III)的化合物组成、或由式(I)的化合物、式(IV)的化合物和式(III)的化合物组成,



式(I)



式(IV)



式(III)

其中R为H、甲基或乙基。

根据本发明的第四个方面,提供了通过或可通过本发明的第一个方面的方法得到的化合物或组合物。该组合物可以例如如本发明的第二个方面中所定义,包括其任一实施方案。

根据本发明的第五个方面,提供了本发明的第二个或第三个方面的组合物作为香料组合物的用途或在香料组合物中的用途。

根据本发明的第六个方面,提供了包含本发明的第二个或第三个方面的组合物的消费品。

在本发明的任一方面的某些实施方案中,R为甲基。当R为甲基时,式(I)的化合物可以称作(+)-龙涎缩醛。

本发明的某些实施方案可以提供如下优点中的一个或多个:

- 生产(+)-龙涎缩醛和(+)-龙涎缩醛同系物的生物催化路线;
- 更温和的反应条件(例如较低的温度);
- 高选择性;
- 使用可替代原料(例如用于化学合成的可替代原料)。

关于本发明的任何一个或多个所述方面提供的细节、实施例和优选方案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。除非本文另有指示或与上下文明显矛盾,否则本发明涵盖本文所述的实施方案、实施例和优选方案的所有可能变型的任意组合。

附图简述

图1显示了制备式(II)的化合物的反应方案。对于化合物,R为H、甲基或乙基。

图2显示了用野生型和变体SHC酶将羟基法呢基丙酮转化成(+)-龙涎缩醛的转化率[%],并给出了测试的SHC酶在如表2所示的其最佳反应条件下的性能的概述。

序列概述

SEQ ID NO:1为野生型酸热脂环酸芽孢杆菌(*Alicyclobacillus acidocaldarius*)AacSHC的氨基酸序列。

SEQ ID NO:2为野生型酸热脂环酸芽孢杆菌AacSHC的核苷酸序列。

SEQ ID NO:3为具有取代M132R、A224V、I432T、A557T和R613S的SEQ ID NO:1并且本文可以称作SHC酶变体#65。

SEQ ID NO:4为SHC变体#65的核苷酸序列。

SEQ ID NO:5相应于具有取代M132R、A224V、I432T、Y81H、A557T和R613S的SEQ ID NO:1并且本文可以称作SHC酶变体#66。

SEQ ID NO:6为SHC变体SHC#66的核苷酸序列。

SEQ ID NO:7相应于具有取代M132R、A224V、I432T、T90A和R613S的SEQ ID NO:1并且本文可以称作SHC酶变体#90C7。

SEQ ID NO:8为SHC变体#90C7的核苷酸序列。

SEQ ID NO:9相应于具有取代M132R、A224V、I432T、Y81H、H431L和A557T的SEQ ID NO:1并且本文可以称作SHC酶变体#110B8。

SEQ ID NO:10为SHC变体#110B8的核苷酸序列。

SEQ ID NO:11相应于具有取代M132R、A224V、I432T、A172T和M277K的SEQ ID NO:1并且本文可以称作SHC酶变体#115A7。

SEQ ID NO:12为SHC变体#115A7的核苷酸序列。

SEQ ID NO:13相应于具有突变M132R、A224V和I432T的SEQ ID NO:1并且本文可以称作SHC酶变体215G2。

SEQ ID NO:14为SHC变体#215G2的核苷酸序列。

SEQ ID NO:15为野生型运动发酵单胞菌(*Zymomonas mobilis*)ZmoSHC1的氨基酸序列。

SEQ ID NO:16为野生型运动发酵单胞菌ZmoSHC2的氨基酸序列。

SEQ ID NO:17为野生型日本慢生根瘤菌(*Bradyrhizobium japonicum*)BjaSHC的氨基酸序列。

SEQ ID NO:18为野生型细长嗜热聚球藻(*Thermosynechococcus elongatus*)TelSHC的氨基酸序列。

SEQ ID NO:19为野生型巴氏醋杆菌(*Acetobacter pasteurianus*)ApaSHC1的氨基酸序列。

SEQ ID NO:20为野生型致病葡糖杆菌(*Gluconobacter morbifer*)GmoSHC的氨基

酸序列。

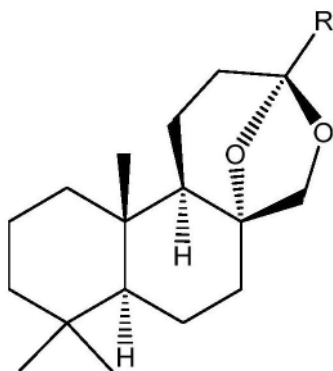
SEQ ID NO:21为野生型巨大芽孢杆菌 (*Bacillus megaterium*) BmeSHC的氨基酸序列。

SEQ ID NO:22相应于具有取代M132R、A224V、I432T、A557T和H431L的SEQ ID NO:1并且本文可以称作SHC酶变体#49。

详细描述

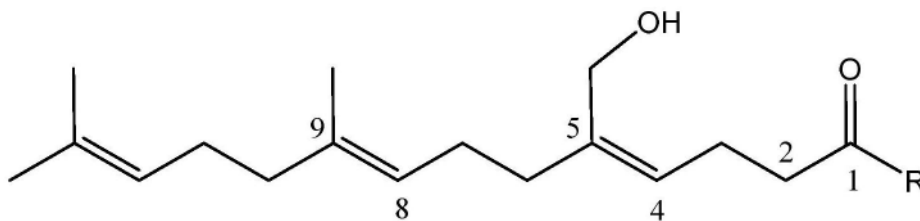
本发明至少部分地基于以下令人惊奇的发现：角鲨烯-何帕烯环化酶 (SHC) 和酶变体可以用于由式 (II) 的多不饱和醇制备 (+)-龙涎缩醛和龙涎缩醛同系物。特别令人惊讶的是，其中如本文式 (II) 所定义的烯基链被羟基甲基取代的底物进行了由内部缩酮化终止的酶促多环化反应。

因此，本文在第一个方面提供了制备式 (I) 的化合物的方法，



式(I)

其中该方法包括使式 (II) 的化合物与角鲨烯-何帕烯环化酶 (SHC) 或酶变体接触，



式(II)

其中R为H、甲基或乙基。

在某些实施方案中，两个双键均为E-构型 (式 (II) 的E,E-化合物)。

在某些实施方案中，C-8与C-9之间的双键为E-构型且C-4与C-5之间的双键为Z-构型 (式 (II) 的E,Z-化合物)。

在具体的实施方案中，R为甲基。

本文提供的方法以酶促方式使用SHC酶或酶变体将式 (II) 的化合物转化成式 (I) 的化合物 (生物转化反应)。

式 (II) 的化合物

式 (II) 的化合物以四种不同立体异构体的形式存在，例如，作为具有E,E-或E,Z-构型的式 (II) 的化合物。

在某些实施方案中，该方法包括使式 (II) 的E,Z-化合物在不存在任何其他式 (II)

的立体异构体的情况下与角鲨烯-何帕烯环化酶(SHC)或酶变体接触。

在其他实施方案中,式(II)的化合物可以例如为立体异构体的混合物。在某些实施方案中,该混合物包含式(II)的E,E-化合物和式(II)的一种或多种其他立体异构体。在某些实施方案中,该混合物包含式(II)的E,Z-化合物和式(II)的一种或多种其他立体异构体。

在某些实施方案中,该方法包括使包含式(II)的E,E-化合物和式(II)的E,Z-化合物的混合物、主要由式(II)的E,E-化合物和式(II)的E,Z-化合物组成的混合物或由式(II)的E,E-化合物和式(II)的E,Z-化合物组成的混合物与SHC酶或酶变体接触。在某些实施方案中,该组合物不含式(II)的任意其他立体异构体。

式(II)的E,Z-化合物与式(II)的全部其他立体异构体的重量比可以例如等于或大于约10:90。例如,式(II)的E,Z-化合物与式(II)的全部其他立体异构体的重量比可以等于或大于约20:80,或等于或大于约30:70,或等于或大于约40:60,或等于或大于约50:50,或等于或大于约60:40,或等于或大于约70:30,或等于或大于约80:20,或等于或大于约90:10,或等于或大于约95:5,或等于或大于约99:1。

式(II)的E,Z-化合物与式(II)的全部其他立体异构体的重量比可以例如等于或小于约99:1。例如,式(II)的E,Z-化合物与式(II)的其他立体异构体的重量比可以等于或小于约95:5,或等于或小于约90:10,或等于或小于约85:15,或等于或小于约80:20,或等于或小于约60:40。

例如,式(II)的E,Z-化合物与式(II)的全部其他立体异构体的重量比可以在约10:90至约99:1,或约10:90至约90:10,或约20:80至约80:20,或约50:50至约80:20,或约60:40至约80:20的范围。

式(II)的E,Z-化合物与式(II)的E,E-化合物的重量比可以例如等于或大于约10:90。例如,式(II)的E,Z-化合物与式(II)的E,E-化合物的重量比可以等于或大于约20:80,或等于或大于约30:70,或等于或大于约40:60,或等于或大于约50:50,或等于或大于约60:40,或等于或大于约70:30,或等于或大于约80:20,或等于或大于约90:10,或等于或大于约95:5,或等于或大于约99:1。

式(II)的E,Z-化合物与式(II)的E,E-化合物的重量比可以例如等于或小于约99:1。例如,式(II)的E,Z-化合物与式(II)的E,E-化合物的重量比可以等于或小于约95:5,或等于或小于约90:10,或等于或小于约85:15,或等于或小于约80:20,或等于或小于约70:30,或等于或小于约60:40。

例如,式(II)的E,Z-化合物与式(II)的E,E-化合物的重量比可以在约10:90至约99:1,或约10:90至约90:10,或约20:80至约80:20,或约50:50至约80:20,或约60:40至约80:20的范围。

立体异构体的混合物中每种立体异构体的量可以例如通过气相色谱法或NMR光谱分析鉴定。

当R为甲基时,式(II)的化合物可以称作羟基法呢基丙酮,包括E,E-羟基法呢基丙酮、E,Z-羟基法呢基丙酮、Z,E-羟基法呢基丙酮和Z,Z-羟基法呢基丙酮及其混合物。

在某些实施方案中,并非所有的式(II)的化合物均转换成式(I)的化合物或反应的副产物。因此,在某些实施方案中,本文所述的组合物,例如通过本文所述的方法得到或

可通过本文所述的方法得到的组合物可以包含式(I)的化合物和式(II)的化合物。

在某些实施方案中,通过本文所述的方法制备的混合物中的任何未转化的式(II)的化合物可以与其他反应产物分离,使得该组合物不含任何式(II)的化合物。

在可替代的实施方案中,通过本文所述的方法将所有式(II)的化合物转化成式(I)的化合物或反应的副产物。

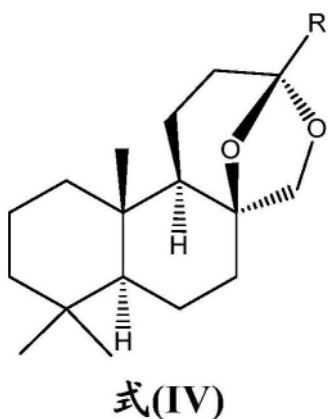
存在的式(II)的化合物的立体异构体的数量可以影响反应速率。SHC酶或酶变体能够将式(II)的E,Z-化合物从式(II)的化合物的立体异构体的复杂混合物(该混合物可以仅包含立体异构体的两种,例如,式(II)的E,Z-化合物和E,E-化合物,或它可以包含三种(即式(II)的E,Z-和E,E-和Z,E-或Z,Z-化合物,乃至所有四种立体异构体)转化成式(I)的化合物。然而,可以观察到较低的转化率,这与其他立体异构体可以与式(II)的E,Z-化合物竞争以接近SHC酶或酶变体的观点一致,并因此可以充当用于将式(II)的E,Z-化合物转化成式(I)的化合物的竞争性抑制剂和/或还充当可替代的底物。因此,式(II)的化合物底物可以包含2-4种异构体、优选两种异构体的异构体混合物。优选地,式(II)的化合物底物包含式(II)的E,Z-和E,E-化合物的异构体混合物、主要由式(II)的E,Z-和E,E-化合物的异构体混合物组成或由式(II)的E,Z-和E,E-化合物的异构体混合物组成。

式(II)的化合物可以按照Fujiwara等人(Tetrahedron Letters,1995Vol 36 (46),8435-8438)描述的一般方法合成。或者,式(II)的化合物可以如图1中简要说明的得到,其中R为H、甲基或乙基。

式(I)的化合物

式(I)的化合物包含多个手性碳原子,且由此也可以存在式(I)的化合物的一种或多种立体异构体,包括对映异构体和非对映异构体。除式(I)的化合物外,通过本文所述的方法制备的产品可以包括式(I)的化合物的立体异构体的一种或多种。得到的立体异构体可以取决于所使用的式(II)的化合物的立体异构体。

例如,除式(I)的化合物外,还可以制备式(IV)的化合物,



其中R为H、甲基或乙基。

其中R为甲基的式(IV)的化合物也称作(-)-表-8-龙涎缩醛。

在某些实施方案中,式(I)的化合物的其他立体异构体没有通过该方法制备或不存在于该方法的产物中,例如在某些实施方案中,式(IV)的化合物没有通过该方法制备或不存在于该方法的产物中。

本文所述的方法可以例如制备式(I)的化合物和式(I)的化合物的一种或多种其

他立体异构体(例如式(IV)的化合物)。因此,本文所述的组合物,例如通过本文所述的方法得到或可通过本文所述的方法得到的组合物可以包含式(I)的化合物和式(I)的化合物的一种或多种立体异构体(例如式(IV)的化合物)。

式(I)的化合物与式(I)的全部其他立体异构体的重量比可以例如等于或大于约50:50。例如,式(I)的化合物与式(I)的全部其他立体异构体的重量比可以等于或大于约55:45,或等于或大于约60:40,或等于或大于约65:35,或等于或大于约70:30,或等于或大于约75:25,或等于或大于约80:20,或等于或大于约85:15,或等于或大于约90:10,或等于或大于约95:5,或等于或大于约99:1。

例如,式(I)的化合物可以是在组合物中制备或存在的唯一的式(I)的立体异构形式。换句话说,组合物基于式(I)的化合物的所有立体异构形式的总重量可以包含100wt%的式(I)的化合物,或者具有的式(I)的化合物与式(I)的全部其他立体异构体的重量比为100:0。或者,式(I)的化合物与式(I)的全部其他立体异构体的重量比可以小于约100wt%。例如,式(I)的化合物与式(I)的全部其他立体异构体的重量比可以等于或小于约99:1,或等于或小于约98:2,或等于或小于约97:3。

例如,式(I)的化合物与式(I)的全部其他立体异构体的重量比可以为约50:50至约100:0,或约60:40至约99:1,或约70:30至约98:2,或约80:20至约97:3,或约90:10至约97:3。

立体异构体的混合物中式(I)的化合物的每种立体异构体的量可以例如通过手性柱上的气相色谱或在位移试剂存在下的NMR光谱来鉴定。

通过本文所述的方法获得的式(I)的化合物可以例如是无定形形式或结晶形式。

通过本文所述的方法制备的式(I)的化合物可以通过蒸汽提取/蒸馏或使用非水混溶性溶剂的有机溶剂萃取(以将反应产物和未反应的底物与停留在水相中的生物催化剂分离),随后蒸发溶剂来分离,以便得到如通过气相色谱(GC)分析确定的粗反应产物。蒸汽提取/蒸馏和有机溶剂萃取方法是本领域技术人员已知的。

作为实例,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萃取所得的式(I)的化合物。或者,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(通过例如离心或过滤获得)中萃取所得的式(I)的化合物。作为另一个实例,式(I)的化合物可以作为晶体或以无定形形式存在于固相中,并且还可以通过过滤与剩余的固相(细胞材料或其碎片)和液相分离。作为另外的实例,在高于式(I)的化合物的熔点的温度下,式(I)的化合物可以在水相的顶部上形成油层,该油层可以被去除并收集。为了确保在除去油层后完全回收式(I)的化合物,可以将有机溶剂加入到含有生物质的水相中,以提取生物质中或生物质上或生物质周围含有的任何残留的式(I)的化合物(例如(+)-龙涎缩醛)。可以将有机层与油层合并,然后进一步处理整体以分离和纯化式(I)的化合物。可以进一步使式(I)的化合物选择性结晶以从终产物中除去副产物和任何未反应的式(II)的化合物。术语“选择性结晶”是指一种方法步骤,其中使式(I)的化合物从溶剂中结晶,同时副产物保持溶解在结晶溶剂中,其溶解程度使得分离的结晶材料仅含有式(I)的化合物,或者如果其含有任何副产物,则它们仅以嗅觉可接受的量存在。例如,式(I)的化合物不含或基本上不含副产物如式(III)或(IIIa)的化合物。选择性结晶步骤可以使用水混溶性溶剂,例如乙醇等。式(I)的化合物的选择性结晶可能受到未反应的式(II)的

化合物的存在以及式(I)的化合物与其他可检测副产物的比例的影响。即使仅10%的式(II)的化合物转化成式(I)的化合物,式(I)的化合物的选择性结晶仍然是可能的。

最终的式(I)的化合物产物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。相对于市售的式(I)的化合物的参考物,测试最终的式(I)的化合物产物的嗅觉纯度、质量及其感官特征。还由专家在应用研究中测试了式(I)的化合物材料,以确定该材料是否满足关于其感官特征的规格。

适用于萃取和/或选择性结晶式(I)的化合物的适合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族烃,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,卤代脂族烃,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,芳族烃,例如苯、甲苯、二甲苯、氯苯或二氯苯,脂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、乙醚、甲基叔丁基醚、乙基叔丁基醚、二丙基醚、二异丙基醚、二丁基醚、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,或二噁烷,或这些的混合物。特别优选使用的溶剂为上述的庚烷、甲基叔丁基醚(也称作MTBE、叔丁基甲基醚、叔丁基甲基醚和tBME)、二异丙基醚、四氢呋喃、乙酸乙酯和/或其混合物。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萃取式(I)的化合物。使用乙醇是有利的,因为它易于处理,它是无毒的并且它对环境是友好的。

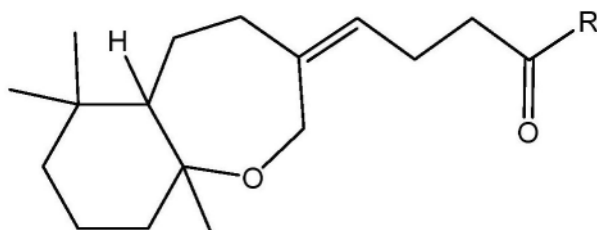
如本文所用的术语“分离的”是指生物转化产物,例如式(I)的化合物,其已经从伴随它的组分中分离或纯化。在不同于其天然起源的来源的细胞系统中产生的实体是“分离的”,因为其必然不含天然伴随其的组分。分离程度或纯度可以通过任何适合的方法测量,例如气相色谱(GC)、HPLC或NMR分析。

在一些实施方案中,从获得的粗产物中分离和纯化式(I)的化合物(例如(+)-龙涎缩醛)(例如至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的纯度)。

理想地,通过本文所述的方法获得的反应液中式(I)的化合物的浓度可以为约1mg/l至约20,000mg/l(20g/l)或更高,例如约20g/l至约200g/l或100-500g/l(包括150g/l、250g/l、300g/l、350g/l、400g/l或450g/l)。

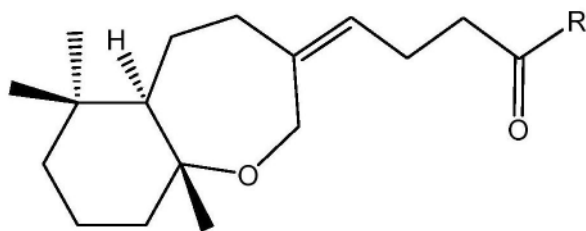
式(III)的化合物

本文所述的方法可以例如制备作为副产物的式(III)的化合物,其中R为H、甲基或乙基:



式(III)

式(III)的化合物可以例如为具有式(IIIa)的相对构型的化合物,其中R为H、甲基或乙基:



式(IIIa)

本文所述的方法可以例如制备式(III)的化合物的一种或多种立体异构体。本文所述的组合物可以包含式(III)的化合物的一种或多种立体异构体。本文所述的方法可以例如制备具有式(IIIa)中所示的相对构型的式(III)的化合物。因此,本文所述的组合物可以包括具有式(IIIa)中所示的相对构型的式(III)的化合物。在某些实施方案中,通过本文所述的方法制备并因此存在于本文所述的组合物中的唯一式(III)的化合物为具有式(IIIa)中所示的相对构型的化合物。

在某些实施方案中,至少约50wt%的全部式(III)的化合物具有式(IIIa)中所示的相对构型。例如,至少约60wt%,或至少约70wt%,或至少约80wt%,或至少约90wt%的全部式(III)的化合物可以具有式(IIIa)中所示的相对构型。

在某些实施方案中,具有式(IIIa)中所示构型的化合物为式(III)的唯一立体异构体。例如,100wt%的全部式(III)的化合物具有式(IIIa)中所示的相对构型。在某些实施方案中,等于或小于约99wt%,或等于或小于约95wt%,或等于或小于约90wt%,或等于或小于约85wt%,或等于或小于约80wt%,或等于或小于约75wt%的全部式(III)的化合物具有式(IIIa)中所示的相对构型。

例如,约50wt%至约100wt%,或约60wt%至约99wt%,或约70wt%至约95wt%的全部式(III)的化合物具有式(IIIa)中所示的相对构型。

立体异构体的混合物中式(III)的化合物的不同异构体的量可以例如通过手性柱上的气相色谱或在位移试剂存在下的NMR光谱来鉴定。

通过本文所述的方法得到的产品

本文还提供了本文所述的方法的产品。因此,本文还提供了通过本文所述的方法(包括其所有实施方案)得到或可通过本文所述的方法得到的组合物。

本文提供了一种组合物,其包含式(I)的化合物和式(III)的化合物、主要由式(I)的化合物和式(III)的化合物组成或由式(I)的化合物和式(III)的化合物组成。该组合物可以例如进一步包含一种或多种其他式(I)的立体异构体,例如式(IV)的化合物。该组合物可以例如包含一种或多种其他式(III)的立体异构体。例如,该组合物可以包含具有式(IIIa)的相对构型的化合物。该组合物可以例如进一步包含任何未反应的式(II)的化合物。在某些实施方案中,R为甲基。

在一个具体的实施方案中,提供了一种组合物,其包含式(I)的化合物、式(IV)的化合物和式(III)的化合物、主要由式(I)的化合物、式(IV)的化合物和式(III)的化合物组成或由式(I)的化合物、式(IV)的化合物和式(III)的化合物组成。

本文还提供了组合物,其包含式(I)的化合物和式(I)的化合物的一种或多种立体异构体,例如式(IV)的化合物。该组合物可以例如进一步包含式(III)的化合物。该组合物

可以进一步包含任何未反应的式 (II) 的化合物。在某些实施方案中, R为甲基。

本文所述的组合物中式 (I) 的化合物与式 (III) 的化合物的重量比可以例如在约 60:40 至约 99:1 的范围。例如, 式 (I) 的化合物与式 (III) 的化合物的重量比可以在约 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 的范围。例如, 式 (I) 的化合物与式 (III) 的化合物的重量比可以在约 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 的范围。

本文所述的粗反应产物中式 (I) 的化合物与式 (II) 的化合物的重量比可以例如在约 90:10 至约 100:0 的范围。例如, 本文所述的组合物中式 (I) 的化合物与式 (II) 的化合物的重量比可以在约 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 的范围。

香料组合物

本文还提供了本文所述的反应产物在香料组合物中或作为香料组合物的用途。

因此, 本文还提供了包含一种或多种式 (I) 的化合物的香料组合物。所谓“香料组合物”是指包含一种或多种式 (I) 的化合物和基质材料的任意组合物。

如本文所用, “基质材料”包括所有已知的香料成分, 其选自广泛的天然产物和目前可获得的合成分子, 例如精油、醇类、醛类和酮类、醚类和缩醛类、酯类和内酯类、大环类和杂环类, 和/或与一种或多种通常与香料组合物中的增香剂结合使用的成分或赋形剂(例如载体材料、稀释剂和本领域常用的其他助剂) 的混合物。

本领域已知的香料成分可容易地从主要的香料制造商商购获得。这类成分的非限制性实例包括:

- 精油和提取物, 例如海狸香、广木香根油、橡苔净油、香叶油、树苔净油、罗勒油、果油(例如香柠檬油和桔油)、香桃木油、玫瑰草油、广藿香油、橙叶油、茉莉油、玫瑰油、檀香油、苦艾油、薰衣草油和/或依兰-依兰油;

- 醇类, 例如肉桂醇 (E)-3-苯基丙-2-烯-1-醇; 顺式-3-己烯醇 (Z)-己-3-烯-1-醇; 香茅醇 (3,7-二甲基辛-6-烯-1-醇); 二氢月桂烯醇 (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)-3,7,11-三甲基十二碳-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); 橙花醇 (3,7-二甲基-2,6-辛二烯-1-醇); 苯基乙醇 (2-苯基乙醇); 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-(2,2,6-三甲基环己基) 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;

- 醛类和酮类, 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-2-基) 乙酮); 羟基香茅醛 (7-羟基-3,7-二甲基辛醛); Iso ESuper® (1-(2,3,8,8-四甲基-1,2,3,4,5,6,7,8-八氢萘-

2-基)乙酮); **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; 马鞭草烯酮; 和/或香草醛;

-醚类和缩醛类,例如**Ambrox**[®] (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); 玫瑰醚(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[二环[4.1.0]庚烷-2,5'-[1,3]二噁烷]);

-大环类,例如黄葵内酯((Z)-氧杂环十七碳-10-烯-2-酮); 十三烷二酸亚乙基酯(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

-杂环类,例如异丁基喹啉(2-异丁基喹啉)。

如本文所用,“载体材料”是指从增香剂的角度来看实际上是中性的材料,即不会显著改变增香剂的感官特性的材料。

所谓“稀释剂”是指通常与增香剂结合使用的任何稀释剂,例如邻苯二甲酸二乙酯(DEP)、一缩二丙二醇(DPG)、肉豆蔻酸异丙酯(IPM)、柠檬酸三乙酯(TEC)和醇(例如乙醇)。

术语“助剂”是指由于与所述组合物的嗅觉性能不特别相关的原因而可以用于香料组合物中的成分,例如,助剂可以是作为加工香料成分或含有所述成分的组合物的助剂的成分,或者它可以改善香料成分或含有香料成分的组合物的处理或储存,例如抗氧化剂助剂,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)和相关酚类、氢醌(CAS 121-31-9)。

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

通常用于包含助剂的香料组合物中的助剂的性质和类型的详细描述不能穷尽,但必须提及的是,所述成分是本领域技术人员众所周知的。

式(I)的化合物的各种应用包括但不限于精细香料或消费品,例如织物护理、盥洗用品、美容护理和清洁产品、洗涤剂产品和皂产品,包括基本上所有其中在商业上使用当前可用的(+)-龙涎缩醛成分的产品。

本文还提供了消费品,其包含如本文所述的组合物或香料组合物,包括其任何实施方案。消费品可以为例如化妆品(例如香水(eau de parfum)或淡香水(eau de toilette)、清洁产品、洗涤剂产品或皂产品。

中间体和原料

本文还提供了用于本文所述的方法的中间体和原料。

本文还提供了包含式(II)的化合物、主要由式(II)的化合物组成或由式(II)的化合物组成的混合物。例如,混合物可以包含其中两个双键均为E-构型的式(II)的化合物(E,E-化合物)和其中C-8与C-9之间的双键为E-构型且C-4与C-5之间的双键为Z-构型的式(II)的化合物(E,Z-化合物)、主要由它们组成或由它们其组成。在一个实施方案中,所述混合物可以包含三种式(II)的立体异构体(即式(II)的E,Z-、E,E-和Z,E-或Z,Z-化合物)乃至所有

四种式 (II) 的立体异构体、主要由它们组成或由它们组成。

式 (II) 的 E,Z-化合物与 E,E-化合物的重量比可以例如等于或大于约 10:90。例如, 式 (II) 的 E,Z-化合物与 E,E-化合物的重量比可以等于或大于约 20:80, 或等于或大于约 30:70, 或等于或大于约 40:60, 或等于或大于约 50:50, 或等于或大于约 60:40, 或等于或大于约 70:30, 或等于或大于约 80:20, 或等于或大于约 90:10, 或等于或大于约 95:5, 或等于或大于 99:1。

式 (II) 的 E,Z-化合物与 E,E-化合物的重量比可以例如等于或小于约 99:1。例如, 式 (II) 的 E,Z-化合物与 E,E-化合物的重量比可以等于或小于约 95:5, 或等于或小于约 90:10, 或等于或小于约 85:15, 或等于或小于约 80:20, 或等于或小于约 70:30, 或等于或小于 60:40。

例如, 式 (II) 的 E,Z-化合物与式 (II) 的 E,E-化合物的重量比可以在约 10:90 至约 99:1, 或约 10:90 至约 90:10, 或约 20:80 至约 80:20, 或约 50:50 至约 80:20, 或约 60:40 至约 80:20 的范围。

SHC 酶或酶变体

本文所述的方法使用 SHC 酶或变体酶以便以酶促方式将式 (II) 的化合物转化成式 (I) 的化合物。

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芽孢杆菌中的野生型角鲨烯何帕烯环化酶。

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作亲本或参考多肽。通常, 变体是人工产生的, 优选通过基因技术方式产生。通常, 衍生变体的多肽为野生型酶或野生型酶。然而, 可用于本公开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。氨基酸序列的变化可以为氨基酸交换 (取代)、插入、缺失、N-端截短或 C-端截短, 或这些变化的任意组合, 其可以在一个或几个位点处发生。

如本文所用, 术语“SHC 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的酶。所述一个或多个氨基酸改变可以例如改变 (例如增加) 底物 (例如式 (II) 的化合物) 的酶活性。或者, 变体 SHC 酶可以衍生自已经存在的 SHC 酶变体。

用于确定和定量 SHC 酶和/或 SHC 酶变体活性的测定法在本文中描述并且是本领域已知的。作为实例, SHC 酶和/或 SHC 酶变体活性可以通过将纯化的 SHC 酶或酶变体或来自已经产生 SHC 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, 并进行底物和反应产物的分析 (例如通过气相色谱 (GC) 或 HPLC 分析) 来确定。关于 SHC 酶和/或 SHC 酶变体活性测定和反应产物分析的进一步细节在实施例提供。这些测定可以包括在重组宿主细胞 (例如大肠杆菌 (*E. coli*)) 中产生 SHC 酶变体。

如本文所用, 术语“活性”是指酶与底物反应以提供目标产物的能力。所述活性可以在所谓的活性测试中通过目标产物的增加、底物 (或原料) 的减少或通过这些参数的组合作为时间的函数来确定。本公开的 SHC 酶的特征可以在于它们将式 (II) 的化合物 (例如羟基法呢基丙酮) 转化成式 (I) 的化合物 (例如 (+)-龙涎缩醛) 的能力。

如本文所用, “生物活性”是指多肽可以表现出的任何活性, 包括但不限于: 酶活

性;与另一种化合物的结合活性(例如与另一种多肽结合,特别是与受体结合,或与核酸结合);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。在其他实施方案中,本文所用的SHC酶变体显示出比参考SHC酶(例如野生型SHC酶或已知的SHC酶变体)更好的底物转化收率。在另外的实施方案中,本文使用的SHC酶变体相对于参考SHC酶(例如215G2 AacSHC相对于野生型AacSHC)可以显示出改变的(例如增加的)生产率。术语“生产率”是指每小时生物转化时间(即加入底物后的时间)每升反应的可回收产物的量(克)。

如本文所用,术语“氨基酸改变”是指相对于参考氨基酸序列的氨基酸序列,在两个氨基酸之间插入一个或多个氨基酸、缺失一个或多个氨基酸或用一個或多个不同的氨基酸取代(其可以是保守的或非保守的)一个或多个氨基酸。取代用变体序列中相同数目的氨基酸替代参考序列的氨基酸。参考氨基酸序列可以例如为野生型(WT)氨基酸序列(例如SEQ ID NO:1或SEQ ID NO:15或SEQ ID NO:16或SEQ ID NO:17或SEQ ID NO:18或SEQ ID NO:19或SEQ ID NO:20或SEQ ID NO:21),或可以例如自身为SHC酶变体序列(例如AacSHC变体215G2-SEQ ID NO:13)。

通过将SHC酶变体的氨基酸序列与参考氨基酸序列(例如野生型或变体SHC氨基酸序列)的氨基酸序列进行比较,可以容易地鉴定氨基酸改变。

SHC酶的适合的来源包括,例如,酸热脂环酸芽孢杆菌(Aac)、运动发酵单胞菌(Zmo)、日本慢生根瘤菌(Bja)、致病葡萄糖杆菌(Gmo)、双向伯克霍尔德氏菌(*Burkholderia ambifaria*)、炭疽芽孢杆菌(*Bacillus anthracis*)、巨大芽孢杆菌(Bme)、荚膜甲基球菌(*Methylococcus capsulatus*)、桤木弗兰克氏菌(*Frankia alni*)、巴氏醋杆菌(Apa)、细长嗜热聚球藻(Tel)、天蓝色链霉菌(*Streptomyces coelicolor*) (Sco)、沼泽红假单胞菌(*Rhodopseudomonas palustris*) (Rpa)、涂氏船蛆杆菌(*Teredinibacter urnerae*) (Ttu)、甲醇黏土杆菌(*Pelobacter carbinolicus*) (Pca)和梨形四膜虫(*Tetrahymena pyriformis*) (参见,例如W02010/139719、US2012/01345477、W02012/066059、W02016/170099;W02018/157021和JP2009060799,其内容通过引用并入本文)。适合的酶还描述在例如Neumann&Simon 1986,Biol Chem Hoppe-Seyler 367,723-729;Seckler&Poralla 1986,Biochem Biophys Act 356-363;Ochs等人1990,J Bacteriol 174,298-302;Seitz等人2012,J Molecular Catalysis B:Enzymatic 84,72-77;和Seitz2012PhD论文(<http://elib.uni-stuttgart.de/handle/11682/1400>)中,其内容通过引用并入本文。这些SHC酶和变体可以用于本文所述的方法。

特别地,SHC酶(例如SHC酶变体可以从其衍生)可以为酸热脂环酸芽孢杆菌(Aac) SHC酶、运动发酵单胞菌(Zmo) SHC酶、日本慢生根瘤菌(Bja) SHC酶、巨大芽孢杆菌(Bme) SHC酶或致病葡萄糖杆菌(Gmo) SHC酶。在某一实施方案中,SHC酶(例如SHC酶变体可以从其衍生)可以为酸热脂环酸芽孢杆菌(Aac) SHC酶。在某一实施方案中,SHC酶(例如SHC酶变体可以从其衍生)可以为巨大芽孢杆菌(Bme) SHC酶。

为了便于参考,名称“AacSHC”可以用于指酸热脂环酸芽孢杆菌(Aac) SHC酶,“ZmoSHC”可以用于指运动发酵单胞菌(Zmo) SHC酶,“BmeSHC”可以用于指巨大芽孢杆菌(Bme) SHC酶,“BjaSHC”可以用于指日本慢生根瘤菌(Bja) SHC酶,并且“GmoSHC”可以用于指致病葡萄糖杆菌(Gmo) SHC酶。

AacSHC、ZmoSHC和BjaSHC酶序列公开在BASF的W02010/139719、US2012/01345477A1、Seitz等人(如上文所引述的)和Seitz(如上文所引述的2012PhD论文)中。公开了ZmoSHC的两种不同的序列,称作ZmoSHC1和ZmoSHC2。GmoSHC酶序列公开在W02018/157021中。表1公开了野生型SHC酶的来源和登录号。

表1.野生型SHC酶的来源和登录号。

SHC 来源菌株(SHC 名称)	参考文献	登录号
酸热脂环酸芽孢杆菌 (WT AacSHC)	JP2009-060799 Neumann 等人 Biol Chem (1986) 367; 723-729	NBRC15652
运动发酵单胞菌 (WT ZmoSHC)	WO2010139719 US20120135477	ATCC31821 PF62207_2 Genpept 登录号 AAV90172
运动发酵单胞菌 (WT ZmoSHC)	Reipen 等人(1995) Microbiology 141:155-161	EMBL/Genbank 登录号 X80766
日本慢生根瘤菌 (WT BjaSHC)	WO2010139719 US2012/0135477	PF62207_5
巨大芽孢杆菌 (WT BmeSHC)	WO2017/150695 WO2019/045058 WO2015/033746	WP_016763969.1
双向伯克霍尔德氏菌 炭疽芽孢杆菌 桤木弗兰克氏菌 沼泽红假单胞菌	WO2010139719 US2012/0135477	
致病葡糖杆菌	WO2018/157021	

野生型AacSHC、野生型ZmoSHC1、野生型ZmoSHC2、野生型BjaSHC、野生型GmoSHC、野生型TelSHC、野生型ApaSHC1和野生型BmeSHC的氨基酸序列也在本文中公开(分别为SEQ ID NO:1、SEQ ID NO:15、SEQ ID NO:16、SEQ ID NO:17、SEQ ID NO:18、SEQ ID NO:19、SEQ ID NO:20和SEQ ID NO:21)。

本文所述并用于本文所述的方法中的SHC酶或酶变体可以例如基于野生型氨基酸序列(SEQ ID NO 1、13、15、16、17、18、19、20、21)之一或其变体、同源物、突变体、衍生物或片段。所述SHC酶或酶变体可以例如具有与本文公开的野生型氨基酸序列之一具有至少

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。

本文所述并且用于本文所述的方法的SHC酶或酶变体可以例如具有等于或大于约75%的选择性。例如，所述SHC酶或酶变体可以例如具有等于或大于约80%，或等于或大于约85%，或等于或大于约90%，或等于或大于约95%的选择性。例如，所述SHC酶或酶变体可以例如具有至多100%、例如小于100%、例如等于或小于约99.5%，或等于或小于约99.0%，或等于或小于约98.0%，或等于或小于约97.0%的选择性。

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引入空位(如果需要)以实现最大百分比序列同一性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。出于确定百分比核苷酸序列同一性的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，并且意指任何氨基酸的肽连接链，与长度或翻译后修饰无关。

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。

下文进一步描述可用于本文所述的方法中的特异性SHC酶和酶变体。

野生型SHC酶

本文所述的方法可以例如使用与野生型SHC酶具有100%序列同一性的SHC酶。野生型SHC酶不必直接从其天然生物体获得，并且可以在实验室中合成，例如使用重组DNA技术。

野生型SHC酶可以例如来自酸热脂环酸芽孢杆菌(Aac)、运动发酵单胞菌(Zmo)、日本慢生根瘤菌(Bja)、致病葡糖杆菌(Gmo)、双向伯克霍尔德氏菌、炭疽芽孢杆菌、巨大芽孢杆菌(Bme)、荚膜甲基球菌、桤木弗兰克氏菌、巴氏醋杆菌(Apa)、细长嗜热聚球藻(Tel)、天蓝色链霉菌(Sco)、沼泽红假单胞菌(Rpa)、涂氏船蛆杆菌(Teredinibacter turnerae)(Ttu)、甲醇黏土杆菌(Pca)和梨形四膜虫(参见，例如W02010/139719、US2012/01345477、W02012/066059，其内容通过引用并入本文)。

特别地，野生型SHC酶可以为酸热脂环酸芽孢杆菌(Aac) SHC酶、运动发酵单胞菌(Zmo) SHC酶、巨大芽孢杆菌(Bme) SHC酶、日本慢生根瘤菌(Bja) SHC酶或致病葡糖杆菌(Gmo) SHC酶。特别地，野生型SHC可以为酸热脂环酸芽孢杆菌(Aac) SHC酶或巨大芽孢杆菌(Bme) SHC酶或运动发酵单胞菌(Zmo) SHC酶(例如ZmoSHC1)。

为了便于参考，名称“AacSHC”可以用于指酸热脂环酸芽孢杆菌SHC酶、“ZmoSHC”可以用于指运动发酵单胞菌SHC酶ZmoSHC1和ZmoSHC2，“BjaSHC”可以用于指日本慢生根瘤菌SHC酶，“BmeSHC”可以用于指巨大芽孢杆菌SHC酶，“GmoSHC”可以用于指致病葡糖杆菌SHC酶，“TelSHC”可以用于指细长嗜热聚球藻SHC酶，且“ApaSHC1”可以用于指巴氏醋杆菌SHC

酶。

野生型SHC酶氨基酸序列可以例如为AacSHC (SEQ ID NO:1)、ZmoSHC1 (SEQ ID NO:15)、ZmoSHC2 (SEQ ID NO:16)、BjaSHC (SEQ ID NO:17)、GmoSHC (SEQ ID NO:20)、BmeSHC (SEQ ID NO:21)、TelSHC (SEQ ID NO:18)或ApaSHC1 (SEQ ID NO:19)。例如,野生型SHC酶可以为BmeSHC (SEQ ID NO:21)、ZmoSHC1 (SEQ ID NO:15)或AacSHC (SEQ ID NO:1)。

SHC变体酶

本文所述的方法可以例如使用SHC酶变体(即与野生型SHC酶的序列同一性小于100%的SHC酶)。

本文所述的方法可以例如使用如W02016/170099或W02018/157021中所述的SHC酶变体,其内容通过引用并入本文。例如,用于本文所述的方法的SHC酶变体可以为W02016/170099中所述的SHC酶变体215G2。

SHC酶变体可以例如具有氨基酸序列,该氨基酸序列具有与野生型SHC酶氨基酸序列至少约70.0%的同一性。例如,SHC酶变体可以具有氨基酸序列,该氨基酸序列具有与野生型SHC酶氨基酸序列至少约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。

SHC酶变体具有的氨基酸序列与野生型SHC酶氨基酸序列的同一性小于100%,例如等于或小于约99.5%,或等于或小于约99.0%。

例如,SHC酶变体可以与野生型SHC酶氨基酸序列具有约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。

野生型SHC酶可以例如来自酸热脂环酸芽孢杆菌(Aac)、运动发酵单胞菌(Zmo)、日本慢生根瘤菌(Bja)、致病葡萄糖杆菌(Gmo)、双向伯克霍尔德氏菌、炭疽芽孢杆菌、巨大芽孢杆菌(Bme)、荚膜甲基球菌、桉木弗兰克氏菌、巴氏醋杆菌(Apa)、细长嗜热聚球藻(Tel)、天蓝色链霉菌(Sco)、沼泽红假单胞菌(Rpa)、涂氏船蛆杆菌(Ttu)、甲醇黏土杆菌(Pca)和梨形四膜虫(参见,例如W02010/139719、US2012/01345477、W02012/066059,其内容通过引用并入本文)。

野生型SHC酶氨基酸序列可以例如为AacSHC (SEQ ID NO:1)、ZmoSHC1 (SEQ ID NO:15)、ZmoSHC2 (SEQ ID NO:16)、BjaSHC (SEQ ID NO:17)、GmoSHC (SEQ ID NO:20)、BmeSHC (SEQ ID NO:21)、TelSHC (SEQ ID NO:18)或ApaSHC1 (SEQ ID NO:19)。例如,野生型SHC酶可以为AacSHC (SEQ ID NO:1)。

因此,在某些实施方案中,SHC酶或SHC酶变体可以具有氨基酸序列,该氨基酸序列具有与SEQ ID NO:1、SEQ ID NO:15、SEQ ID NO:16、SEQ ID NO:17、SEQ ID NO:20、SEQ ID NO:18、SEQ ID NO:19或SEQ ID NO:21至少约70.0%的同一性。例如,SHC酶或SHC酶变体具有氨基酸序列,该氨基酸序列具有SEQ ID NO:1、SEQ ID NO:15、SEQ ID NO:16、SEQ ID NO:17、SEQ ID NO:20、SEQ ID NO:18、SEQ ID NO:19或SEQ ID NO:21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。

例如,SHC酶变体可以例如具有氨基酸序列,该氨基酸序列具有的与SEQ ID NO:1、

SEQ ID NO:15、SEQ ID NO:16、SEQ ID NO:17、SEQ ID NO:20、SEQ ID NO:18、SEQ ID NO:19或SEQ ID NO:21的同一性小于100%，例如等于或小于约99.5%，或等于或小于约99.0%。

例如，SHC酶变体可以与SEQ ID NO:1、SEQ ID NO:15、SEQ ID NO:16、SEQ ID NO:17、SEQ ID NO:20、SEQ ID NO:18、SEQ ID NO:19或SEQ ID NO:21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。

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引入空位(如果需要)以实现最大百分比序列同一性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。出于确定百分比序列同一性的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，并且意指任何氨基酸的肽连接链，与长度或翻译后修饰无关。

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。这样的比对可以使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包，<http://hmm.wustl.edu/>)、或使用例如在<https://www.ebi.ac.uk/Tools/msa/clustalo/>上可得到的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(University of Iowa)的数学算法)或Myers和Miller的数学算法(1989-Cabios 4:11-17)。所用的优选参数为默认参数，如它们在<https://www.ebi.ac.uk/Tools/msa/clustalo/>上所设置的。

可以使用例如BLAST、BLAT或BlastZ (或BlastX) 计算百分比序列同一性。在Altschul等人(1990) J. Mol. Biol. 215, 403-410的BLASTN和BLASTP程序中引入了类似的算法。BLAST多核苷酸检索可以用BLASTN程序进行，得分=100，字长=12，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。可以用BLASTP程序进行BLAST蛋白质检索，得分=50，字长=3，以获得与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，可以如Altschul等人(1997) Nucleic Acids Res. 25, 3389-3402中所述利用空位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(homology mapping)技术如Shuffle-LAGAN(Brudno M., Bioinformatics 2003b, 19 Suppl 1:154-162)或马尔可夫(Markov)随机场来补充。当在本申请中涉及序列同一性的百分比时，如果没有另外特别指出，则相对于较长序列的全长计算这些百分比。

在具体实施方案中，使用CLUSTAL O (1.2.4版) 确定两个序列之间的%同一性。

与亲本SHC酶相比，SHC酶变体可例如具有增加的将式(II)的化合物转化成式(I)的化合物的酶活性。增加的酶活性可以指式(II)的化合物向式(I)的化合物的酶促转化的任何方面，包括例如式(II)的化合物向式(I)的化合物的总转化率增加、式(II)的化合物的转化速率增加(例如在反应的前4小时、或前6小时、或前12小时中)、式(I)的化合物的产量增加和/或副产物的产量减少。通常，增加的酶活性可以通过增加的生产率来定义。其可以根据每小时、每克生物催化剂和每升反应产生的式(I)的化合物来定义。

与亲本SHC酶相比,SHC酶变体可以例如提供增加的式(II)的化合物的转化率。因此,与使用亲本SHC酶的方法相比,本文所述的方法可以具有增加的式(II)的化合物转化水平。与亲本SHC酶相比,SHC酶变体可以例如提供增加的式(II)的化合物的转化速率。因此,与亲本SHC酶相比,本文所述的方法可以具有增加的式(II)的化合物的转化速率。与亲本SHC酶相比,SHC酶变体可以例如在反应的前4小时或前2小时、或前4小时、或前6小时、或前8小时或前12小时或前24小时内提供增加的式(II)的化合物的转化速率。因此,与亲本SHC酶相比,本文所述的方法可以在反应的前2小时、或前4小时、或前6小时、或前8小时、或前12小时、或前24小时内具有增加的式(II)的化合物的转化速率。这可以是当与在相同反应条件(例如相同pH和温度)下使用两种酶(即SHC酶变体和亲本SHC酶)相比时,或者当与在已经单独定义为对于其活性最佳的条件(例如优化的pH和温度)下使用每种酶相比时,并且其可以彼此不同。

式(II)的化合物向式(I)的化合物的转化率可以例如使用如上所述的活性测定法来确定,并且可以计算为每克原料的可回收产物的克数(如果期望,其可以计算为百分比摩尔转化率)。

用于制备本文公开的式(I)的化合物的方法可以在所用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,如下文实施例中所述。

核酸和制备核酸的方法

本文所述的SHC酶和酶变体可以由核酸序列编码。包含编码序列的核酸可以例如分离的核酸。

因此,本文提供了包含编码如本文所述的SHC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人工产生的核酸区段。构建体可以包含编码SHC酶或酶变体的核酸序列和基因表达控制器(例如启动子)。

本文进一步提供了包含如本文所述的构建体的载体。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。载体可以例如质粒、病毒载体、粘粒或人工染色体。

术语“构建体”和“载体”可以重叠,例如当构建体为质粒时。

如本文所用的术语“核酸”或“核酸分子”应具体指本公开的多核苷酸,其可以为DNA、cDNA、基因组DNA、合成DNA或RNA,并且可以为双链或单链的、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的多肽(例如代谢途径的酶)或其片段或部分。

该术语还包括单独的分子,例如cDNA,其中相应的基因组DNA具有内含子并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;通过聚合酶链反应(PCR)产生的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标记)、突变蛋白或给定蛋白质的片段的DNA;和作为cDNA的简并变体的核酸或天然存在的核酸。此外,它包括作为杂合基因(即编码非天然存在的融合蛋白的基因)的一部分的重组核苷酸序列。融合蛋白可以向蛋白质添加一个或多个氨基酸(例如但不限于组氨酸(His)),通常在蛋白质的N-末端,但也在C-末端或融合在蛋白质的区域内。这类融合蛋白或编码这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。

术语“核酸序列”还包括适于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序列。如本文所用,术语“密码子优化的”是指考虑到其特定密码子使用,通过例如用细菌宿主细胞基因(例如大肠杆菌基因)中更频繁使用的密码子取代一个或多个或优选大量密码子,已经适于在特定原核或真核宿主细胞,特别是细菌宿主细胞如大肠杆菌宿主细胞中表达的核酸蛋白质编码序列。

在这方面,编码参考氨基酸序列(例如SEQ ID NO:1或SEQ ID NO:13)及其变体/衍生物的核苷酸序列或基因可以是来源中发现的原始核苷酸序列或基因,或者该基因可以针对所选宿主生物体(例如大肠杆菌)进行密码子优化。

在另一个方面,本公开的核酸序列可操作地连接至允许在原核和/或真核宿主细胞中表达的表达控制序列。如本文所用,“可操作地连接”是指掺入到遗传构建体中,使得表达控制序列有效地控制编码序列或关注的基因的表达。上文涉及的转录/翻译调控元件包括但不限于诱导型和非诱导型、组成型、细胞周期调控型、代谢调控型启动子、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他方式调控基因表达的其他元件。这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的调控元件,例如CUP-1启动子、如例如在tet-on或tet-off系统中使用的tet-阻遏物、lac系统、trp系统调控元件。作为实例,异丙基 β -D-1-半乳糖硫代吡喃糖苷(IPTG)在例如100 μ M至1.0mM的浓度范围内是基因表达的有效诱导物。该化合物为异乳糖的分子模拟物,异乳糖是触发lac操纵子转录的乳糖代谢物,因此当基因在lac操纵子的控制下时,它用于诱导基因表达。诱导基因表达的调控元件的另一个实例为乳糖。类似地,本公开的核酸分子可以形成编码另外的多肽序列的杂合基因的一部分,所述另外的多肽序列例如作为标志物或报道基因起作用的序列。标志物或报道基因的实例包括 β -内酰胺酶、氯霉素乙酰转移酶(CAT)、腺苷脱氨酶(ADA)、氨基葡萄糖苷磷酸转移酶(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转移酶(XGPRT)。与本公开的实施相关的许多标准方法一样,本领域技术人员可以想到另外的有用试剂,例如可以起到标记物或报道基因功能的另外的序列。

本文还提供了编码SHC酶或其变体的重组多核苷酸,其可以插入到载体中用于基因表达和任选的酶纯化。一种类型的载体为代表环状双链DNA环的质粒,另外的DNA区段连接到所述环状双链DNA环中。某些载体可以控制与其功能性连接的基因的表达。这些载体称作“表达载体”。通常适用于DNA重组技术的表达载体具有质粒类型。通常,表达载体包含用于产生SHC野生型或变体酶或如本文所述的基因。在本说明书中,术语“质粒”和“载体”可以互换使用,因为质粒为最常用的载体类型。

这类载体可以包括DNA序列,其包括但不限于非宿主细胞中天然存在的DNA序列、通常不转录成RNA或翻译成蛋白质(“表达”)的DNA序列和期望导入非重组宿主的其他基因或DNA序列。应当理解,通常通过稳定导入一种或多种重组基因来增强本文所述的重组宿主的基因组。然而,自主或复制质粒或载体也可以在本公开的范围内使用。此外,可以使用低拷贝数(例如单拷贝)或高拷贝数(如本文所示例的)质粒或载体来实施本公开。

本公开的载体包括质粒、噬菌粒、噬菌体、粘粒、人工细菌和人工酵母染色体、敲除

或敲入构建体、合成核酸序列或弹夹盒,并且亚组可以以线性多核苷酸、质粒、巨大质粒、合成或人工染色体(例如植物、细菌、哺乳动物或酵母人工染色体)的形式产生。

优选地,在导入载体时,由导入的多核苷酸编码的蛋白质在细胞内表达。质粒通常是标准克隆载体,例如细菌多拷贝质粒。底物可以掺入相同或不同的质粒中。通常使用具有不同类型选择标记的至少两种不同类型的质粒来允许选择包含至少两种类型的载体的细胞。

通常,细菌或酵母细胞可以用本领域众所周知的任何一种或多种核苷酸序列转化。对于体内重组,使用标准转化技术将待与染色体或其他基因重组的基因用于转化宿主。在一个适合的实施方案中,构建体中包括提供复制起点的DNA。本领域技术人员可以适当地选择复制起点。根据基因的性质不同,如果序列已经与作为复制起点本身可操作的基因一起存在,则可能不需要补充的复制起点。

宿主细胞、制备宿主细胞的方法和使用宿主细胞制备式(I)的化合物的方法
重组宿主细胞可以用于本文所述的方法中。

本文进一步提供了一种重组宿主细胞,其包含如本文所述的核酸序列或构建体或载体。本文进一步提供了产生如本文所述的SHC酶或酶变体的重组宿主细胞。

本文所述的用于产生式(I)的化合物的方法可以例如包含培养如本文所述的重组宿主细胞。如本文所用,术语“培养”是指使活细胞增殖以使得它们产生如本文所述的SHC酶或酶变体的过程,所述SHC酶或酶变体可用于产生如本文所述的式(I)的化合物的方法中。

当外源或异源DNA已被导入细胞内时,细菌或酵母细胞可被这类DNA转化。转化DNA可以或可以不整合,即共价连接到细胞的染色体中。例如,在原核生物和酵母中,转化DNA可以维持在附加型元件如质粒上。关于真核细胞,稳定转染的细胞是其中转染的DNA已经整合到染色体中使得其通过染色体复制被子代细胞遗传的细胞。这种稳定性通过真核细胞建立由包含转化DNA的子代细胞群组成的细胞系或克隆的能力来证明。

通常,导入的DNA最初不存在于作为DNA受体的宿主中,但是在本公开范围内从给定宿主分离DNA区段,并且随后将该DNA的一个或多个另外的拷贝导入同一宿主中,例如以增强基因产物的产生或改变基因的表达模式,这在本公开的范围内。在一些情况下,导入的DNA将修饰乃至替换内源基因或DNA序列,例如通过同源重组或定点诱变。适合的重组宿主包括微生物、植物细胞和植物。

本公开的特征还在于重组宿主。术语“重组宿主”(也称作“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经通过至少一个掺入的DNA序列扩增的宿主细胞。可以用如上所述的多核苷酸或载体基因工程改造本发明的宿主细胞。

可以用于本公开的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*)),其可以例如用例如包含本公开的多核苷酸分子的重组噬菌体DNA、质粒DNA、细菌人工染色体或粘粒DNA表达载体转化;简单的真核细胞如酵母(例如,酵母属(*Saccharomyces*)和毕赤酵母属(*Pichia*)),其可以例如用例如包含本公开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的多核苷酸的相应载体的不同,多核苷酸可以整合到例如染色体或线粒体DNA中,或者可以维持在染色体外,例如以附加的形式,或者可以仅短暂地包含在细胞中。

本文所用的术语“细胞”,特别是关于基因工程 and 将一个或多个基因或组装的基因

簇引入细胞或生产细胞时,应理解为是指任何原核或真核细胞。关注用于本公开用途的原核和真核宿主细胞,包括细菌宿主细胞如大肠杆菌或芽孢杆菌属(*Bacillus* sp.)、酵母宿主细胞例如啤酒酵母(*S.cerevisiae*)、昆虫宿主细胞例如草地贪夜蛾(*Spodoptera frugiperda*)或人宿主细胞例如HeLa和Jurkat。

具体地,细胞为真核细胞,优选真菌、哺乳动物或植物细胞,或原核细胞。适合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)或蠕虫细胞,包括隐杆线虫属(*Caenorhabditis*) (包括秀丽隐杆线虫(*Caenorhabditis elegans*))的细胞。适合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、HEK293 T-RexTM细胞或其他可转染的真核细胞系。适合的细菌细胞包括但不限于大肠杆菌。

优选地,可以使用原核细胞,例如大肠杆菌,芽孢杆菌属,链霉菌属,或哺乳动物细胞,如HeLa细胞或Jurkat细胞,或植物细胞,如拟南芥属(*Arabidopsis*)。

细胞可以例如选自原核、酵母、植物和/或昆虫宿主细胞。

优选地,所述细胞为曲霉菌属(*Aspergillus* sp.)或真菌细胞,优选地,它可以选自酵母属、假丝酵母属(*Candida*)、克鲁维酵母属(*Kluyveromyces*)、汉逊酵母属(*Hansenula*)、裂殖酵母属(*Schizosaccharomyces*)、耶氏酵母属(*Yarrowia*)、毕赤酵母属和曲霉菌属。

优选地,所述细胞为细菌细胞,例如选自埃希氏杆菌属(*Escherichia*)、链霉菌属、芽孢杆菌属、假单胞菌属(*Pseudomonas*)、乳杆菌属(*Lactobacillus*)和乳球菌属(*Lactococcus*)的属。例如,所述细菌可以为大肠杆菌。

优选地,大肠杆菌宿主细胞为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)。

与本公开一起使用的一种优选宿主细胞为大肠杆菌,其可以如本文所述重组制备。因此,重组宿主可以为重组大肠杆菌宿主细胞。存在可用于大肠杆菌的突变体、质粒、代谢的详细计算机模型和其他信息的文库,允许合理设计各种模块以提高产物收率。

在一个实施方案中,重组大肠杆菌微生物包含编码SHC酶或酶变体的核苷酸序列(例如,微生物包含选自SEQ ID NO:2、4、6、8、10、12和14的核苷酸序列)。

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。在另一个优选的实施方案中,重组大肠杆菌微生物包含编码本文公开的SHC酶和/或酶变体的核苷酸序列。

与本公开一起使用的另一种优选的宿主细胞为啤酒酵母。存在可用于啤酒酵母的突变体、质粒和详细的代谢计算机模型和其他信息的文库,允许合理设计各种模块以提高产物收率。制备重组啤酒酵母微生物的方法是已知的。

细胞的培养可以以常规方式进行。培养基可以包含碳源、至少一种氮源和无机盐,并且向其中加入维生素。该培养基的成分可以是常规用于培养所讨论的微生物物种的成分。在本发明方法中使用的碳源包括可被重组宿主细胞代谢以促进关注的SHC酶生长和/或产生的任何分子,所述SHC酶用于将式(II)的化合物转化成式(I)的化合物。适合的碳源的实例包括但不限于蔗糖(例如在糖蜜中发现的)、果糖、木糖、甘油、葡萄糖、纤维素、淀粉、纤维二糖或其他含葡萄糖的聚合物。

在采用酵母作为宿主的实施方案中,例如,碳源如蔗糖、果糖、木糖、乙醇、甘油和

葡萄糖是适合的。碳源可以在整个培养期间以分批或补料分批提供给宿主生物体,或者可替代地,可以使用另一种能量源,例如蛋白质或蛋白质水解物。

用于本公开的方法中的重组宿主细胞微生物可以在通常用于微生物增殖的pH、温度和反应条件下在丰富培养基(例如LB培养基、细菌用胰胨酵母提取物培养基、营养培养基等)中增殖。在本公开的一个实施方案中,使用确定的基本培养基例如M9A进行培养。

M9A培养基的组分包含:14g/l KH_2PO_4 、16g/l K_2HPO_4 、1g/l Na_3 柠檬酸盐 $\cdot 2\text{H}_2\text{O}$ 、7.5g/l $(\text{NH}_4)_2\text{SO}_4$ 、0.25g/l $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 、0.015g/l $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ 、5g/l葡萄糖和1.25g/l酵母提取物)。

在本公开的另一个实施方案中,使用富含营养物的培养基例如LB。LB培养基的组分包含:10g/l胰胨、5g/l酵母提取物、5g/l NaCl 。矿物培养基和M9矿物培养基的其他实例公开于例如US6524831B2和US2003/0092143A1中。

基本培养基的另一个实例可以如下制备:对于350ml培养物:向35ml柠檬酸/磷酸盐储备溶液(133g/l KH_2PO_4 、40g/l $(\text{NH}_4)_2\text{HPO}_4$ 、17g/l柠檬酸 $\cdot \text{H}_2\text{O}$, 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(参见下文)、0.035ml硫胺素溶液和7ml 20%葡萄糖。

微量元素溶液:50g/l $\text{Na}_2\text{EDTA} \cdot 2\text{H}_2\text{O}$ 、20g/l $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ 、3g/l H_3BO_3 、0.9g/l $\text{MnSO}_4 \cdot 2\text{H}_2\text{O}$ 、1.1g/l CoCl_2 、80g/L CuCl_2 、240g/l $\text{NiSO}_4 \cdot 7\text{H}_2\text{O}$ 、100g/l KI 、1.4g/l $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24} \cdot 4\text{H}_2\text{O}$ 、1g/l $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ 的去离子水溶液。

硫胺素溶液:2.25g/l硫胺素 $\cdot \text{HCl}$ 的去离子水溶液。

MgSO_4 溶液:50% (w/v) $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 的去离子水溶液。

重组微生物可以以分批、补料分批或连续过程或其组合的方式生长。通常,使重组微生物在发酵罐中在限定的温度下在适合的营养源(例如碳源)存在下生长期望的时间期限,以产生足够的SHC酶,从而能够将式(II)的化合物转化成式(I)的化合物并产生期望的量的式(I)的化合物。重组宿主细胞可以以任意适合的方式培养,例如通过分批培养或补料分批培养。

如本文所用,术语“分批培养”是在培养期间既不添加也不取出培养基的培养方法。

如本文所用,术语“补料分批”是指其中在培养期间添加培养基但不取出培养基的培养方法。

本公开的一个实施方案提供了一种在适合的细胞系统中生产式(I)的化合物的方法,所述方法包括在适合的条件下在细胞系统中产生SHC酶或酶变体,将式(II)的化合物进料至细胞系统,使用细胞系统产生的SHC酶或酶变体将式(II)的化合物转化成式(I)的化合物,从细胞系统收集式(I)的化合物并任选地从该系统中分离式(I)的化合物。其他核苷酸序列的表达可以用于增强用于制备式(I)的化合物的生物转化中使用的细胞系统的活性。

本公开的另一个实施方案是制备式(I)的化合物的生物转化方法,所述方法包括使包含编码SHC酶或酶变体的基因的宿主细胞生长,在宿主细胞中产生SHC酶或酶变体,将式(II)的化合物进料至宿主细胞,在适于促进式(II)的化合物转化成式(I)的化合物的pH、温度和增溶剂的条件下孵育宿主细胞,并收集式(I)的化合物。宿主细胞中SHC酶或酶变体的产生提供了当在适合的反应条件下将式(II)的化合物添加至宿主细胞时制备式(I)的化

合物的方法。通过向反应混合物中加入更多的生物催化剂和SDS,可以提高获得的转化率。

可以以多种方式培养重组宿主细胞微生物,以便提供已产生SHC酶或酶变体的适量的细胞用于随后的生物转化步骤。由于适用于生物转化步骤的微生物变化很大(例如酵母、细菌和真菌),因此培养条件当然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有文献记载。任何本领域已知的用于培养重组宿主细胞微生物的细胞的方法可用于产生可用于本公开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。

所选择的培养条件不仅影响所获得的细胞(生物质)的量,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酶或酶变体基因并产生SHC酶或酶变体的重组宿主细胞微生物称作生物催化剂,其适用于生物转化反应。在一些实施方案中,生物催化剂是产生SHC酶或酶变体的重组全细胞,或者其可以是悬浮液或固定化形式。在其他实施方案中,生物催化剂是由产生SHC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Seitz等人2012中所公开的一如上文所引用)。产生SHC酶或酶变体的重组全细胞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(其然后用于单罐反应)。产生SHC酶或酶变体的重组全细胞可以包括完整重组全细胞和/或细胞碎片。无论哪种方式,SHC酶或酶变体以某种方式与膜(诸如细胞膜)缔合,以便接收底物(例如式(II)的化合物)和/或与底物相互作用,该膜(例如细胞膜)可以为全细胞(例如重组全细胞)的一部分。SHC酶或酶变体也可以是固定化形式(例如与酶载体缔合),其允许SHC酶或酶变体与底物(例如式(II)的化合物)相互作用。SHC酶或酶变体也可以以可溶形式使用。

在一个实施方案中,在生物转化步骤之前,生物催化剂以足够的量产生(以产生足够的生物质)、收获并洗涤(和任选地储存(例如冷冻或冻干))。

在另一个实施方案中,细胞以足够的量产生(以产生足够的生物催化剂),然后调节反应条件,无需收获和洗涤生物催化剂用于生物转化反应。这种一步(或“单罐”)方法是有利的,因为它简化了过程。用于生长细胞的培养基也适用于生物转化反应,条件是调节反应条件以促进生物转化反应。

使细胞生长的最佳pH在6.0-8.0的范围内。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酶或酶变体的类型。使用本领域技术人员众所周知的技术调节pH。

本公开的生物转化方法在时间、温度、pH和增溶剂的条件下进行,以提供式(II)的化合物向式(I)的化合物的转化。

对于所考虑的SHC野生型酶或SHC酶变体,反应混合物的pH可以在4-8,优选4.5-6.5,更优选4.5-6.5的范围内,并且可以通过向反应混合物中加入缓冲液来维持。用于此目的的示例性缓冲液为柠檬酸缓冲液或琥珀酸缓冲液。

对于所考虑的SHC野生型酶或SHC酶变体,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35℃至约55℃。在生物转化过程中,温度可以保持恒定或可以改变。

当生物催化剂与式(II)的化合物之比为约2:1时,[SDS]/[细胞]之比可以在约10:1-20:1,优选约15:1-18:1的范围内,优选约16:1。

用于制备本文公开的式(I)的化合物的方法可以在最佳温度、pH和表面活性剂浓

度下进行,所述最佳温度、pH和表面活性剂浓度使得所考虑的每种单独SHC酶(野生型或变体)具有最佳活性。

在生物转化反应中包括增溶剂(例如表面活性剂、洗涤剂、溶解度增强剂、水混溶性有机溶剂等)可能是有用的。

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)的成分。表面活性剂可用作洗涤剂、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)。

虽然Triton X-100可用于部分纯化SHC酶或酶变体(以可溶性或膜级分/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2012PhD论文)中的公开内容以及Neumann和Simon(1986—如上文所引用)和JP2009060799中的公开内容。SDS可用作增溶剂。

不希望受理论束缚,SDS与重组微生物宿主细胞一起使用可能是有利的,因为SDS可以有利地与宿主细胞膜相互作用,以使SHC酶或酶变体(其为膜结合酶)更容易接近式(II)的化合物底物。此外,在反应混合物中包含适合水平的SDS可以改善乳液的性质(例如,在水中的式(II)的化合物)和/或改善式(II)的化合物底物对宿主内的SHC酶的接近。生物转化反应中使用的增溶剂(例如SDS)的浓度受生物质量和底物浓度的影响。也就是说,增溶剂(例如SDS)浓度、生物质量和底物浓度之间存在一定程度的相互依赖性。作为实例,随着式(II)的化合物底物的浓度增加,需要足量的生物催化剂和增溶剂(例如SDS)来进行有效的生物转化反应。例如,如果增溶剂(例如SDS)浓度太低,则可能观察到式(II)的化合物的次优转化。另一方面,例如,如果增溶剂(例如SDS)浓度过高,则可能存在生物催化剂通过破坏完整微生物细胞和/或SHC酶或酶变体的变性/失活而受到影响的危险。必须仔细研究在生物质量和底物的背景下选择合适浓度的SDS。

在一些实施方案中,使用向其中加入式(II)的化合物底物的生物催化剂生产式(I)的化合物。

能够通过使用已知装置(例如蠕动泵、输注注射器等)进料来添加底物。式(II)的化合物可以是油溶性的并且以油形式提供。鉴于生物催化剂(微生物细胞,例如完整重组全细胞和/或细胞碎片和/或固定化酶)存在于水相中,当将式(II)的化合物加入到生物转化反应混合物中时,生物转化反应可以被认为是三相体系(包含水相、固相和油相)。即使当存在SDS时也是这种情况。

发酵罐可以用于使产生SHC酶或酶变体基因并产生活性SHC酶或酶变体的重组宿主细胞在用于将式(II)的化合物转化成式(I)的化合物(例如与一种或多种式(III)的副产物混合)的相同发酵罐容器中生长至适合用作生物催化剂的足够生物质浓度。

本领域技术人员将理解,通过实施连续过程,例如产物去除、底物进料和生物质添加或(部分)替换,可以实现更高的累积生产效价(production titers)。优选地,在包含SHC酶或酶变体的重组宿主细胞存在下将式(II)的化合物生物转化成式(I)的化合物产生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的式(I)的化合物收率,以摩尔百分比表示并且以所用的式(II)的化合物的摩尔数为基准;特别优选地,收率为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100mol%。

SHC酶或酶变体的活性通过以百分比计的反应速率((产物的量/(产物的量+剩余原料的量)) $\times$ 100)来定义。优选地,在SHC酶或酶变体存在下将式(II)的化合物生物转化成式(I)的化合物产生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的式(I)的化合物收率,以摩尔百分比表示并且以所用的式(II)的化合物的摩尔数为基准;特别优选地,收率为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。

在本发明的一个优选实施方案中,在例如4、6、8、10、12、16、20、24、36、48或72小时的限定时间期限内测定收率和/或反应速率,在该限定时间期限内式(II)的化合物通过包含编码SHC酶或酶变体的核苷酸序列并且已产生SHC酶或酶变体的重组宿主细胞转化成式(I)的化合物。

在另一个实施方案中,反应在例如25℃、30℃、40℃、50℃或60℃的精确确定的条件下进行。特别地,收率和/或反应速率通过在约35℃至约55℃的温度范围内在24-72小时的时间期限内通过本发明的SHC酶或酶变体进行将式(II)的化合物转化成式(I)的化合物的反应来确定。

在本发明的另一个实施方案中,包含编码SHC酶变体的核苷酸序列的重组宿主细胞的特征在于在相同条件下,优选地在已经单独定义为对于所考虑的SHC酶的活性最佳的条件下,与亲本野生型或变体SHC酶相比,其显示在式(II)的化合物的反应中得到式(I)的化合物的收率和/或反应速率为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。本文的术语条件涉及但不限于反应条件,例如pH、温度和增溶剂(例如SDS)的浓度。

用于在包含编码野生型/参考SHC或SHC变体的核苷酸序列的重组大肠杆菌菌株中由式(II)的化合物制备式(I)的化合物的生物转化方法的成功开发可以提供用于式(I)的化合物生产的低成本和工业上经济的方法。

在整个说明书和随后的权利要求书中,除非上下文另有要求,否则词语“包括(comprise)”以及例如“包含(comprises)”和“含有(comprising)”的变化形式将被理解为暗示包含所述整数或步骤或者整数或步骤的组,但不排除任何其他整数或步骤或者整数或步骤的组。术语“包含”还表示“包括”以及“由……组成”,例如“包含”X的组合物可以仅由X组成,或者可以包括另外的成分,例如X+Y。还必须注意,正如在本说明书和所附权利要求

书中所使用的,单数形式“一个”、“一种”和“该”包括复数指示物,内容中另有明确规定的除外。作为实例,涉及的“基因”或“酶”是指“一种或多种基因”或“一种或多种酶”。

应当理解,本公开不限于本文所述的特定方法、方案和试剂,因为这些可以改变。还应理解,本文使用的术语仅用于描述特定实施方案的目的,并不旨在限制本公开的范围,本公开的范围将仅由所附权利要求限定。除非另有定义,否则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含义相同的含义。根据本公开,可以采用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。

本公开的应用不限于以下描述中阐述的或附图中示例的构造细节和部件布置。本公开能够具有其他实施方案并且能够以各种方式实施或进行。此外,本文使用的措辞和术语是为了描述的目的,而不应被认为是限制性的。优选地,本文所用的术语如“A multilingual glossary of biotechnological terms: (IUPAC Recommendations)”, Leuenberger, H.G.W., Nagel, B. 和 Kolbl, H. eds. (1995), Helvetica Chimica Acta, CH-4010 Basel, Switzerland) 中所述定义。

在本说明书的整个文本中引用了若干对比文件。本文引用的每篇对比文件(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 登录号序列提交文件等),无论是上文还是下文,均通过引用整体并入本文。

本文描述的实施例是对本公开的示例,而不旨在对其进行限制。已经根据本公开描述了本公开的不同实施方案。在不脱离本公开的精神和范围的情况下,可以对本文描述和示例的技术进行许多变型和改变。因此,应当理解,这些实施例仅是示例明性的,并不限制本公开的范围。

实施例

实施例1-SHC酶的生产

SHC质粒制备:

将编码野生型或变体角鲨烯何帕烯环化酶(SHC)的基因插入质粒pET-28a(+)中,其中其处于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。

培养基制备:

[0232] 对于350ml培养物,如下制备默认选择的基本培养基:向35ml柠檬酸/磷酸盐储备溶液(133g/l KH_2PO_4 、40g/l $(\text{NH}_4)_2\text{HPO}_4$ 、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(参见下文)、0.035ml硫胺素溶液和7ml 20%葡萄糖。

微量元素溶液:50g/l $\text{Na}_2\text{EDTA} \cdot 2\text{H}_2\text{O}$, 20g/l $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 3g/l H_3BO_3 , 0.9g/l $\text{MnSO}_4 \cdot 2\text{H}_2\text{O}$, 1.1g/l CoCl_2 , 80g/L CuCl_2 , 240g/l $\text{NiSO}_4 \cdot 7\text{H}_2\text{O}$, 100g/l KI , 1.4g/l $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24} \cdot 4\text{H}_2\text{O}$, 1g/l $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ 的去离子水溶液。

硫胺素溶液:2.25g/l硫胺素·HCl的去离子水溶液。

MgSO_4 溶液:50% (w/v) $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 的去离子水溶液。

SHC酶或酶变体生产(生物催化剂生产)。

小规模生物催化剂生产(野生型SHC或SHC变体)

从包含SHC生产质粒的大肠杆菌菌株BL21 (DE3) 的预培养物接种350ml培养物(补充有50 μ g/ml卡那霉素的培养基)。在恒定搅拌(250rpm)下使细胞生长至在37 $^{\circ}$ C下光密度为约0.5 (OD_{650nm})。

然后通过加入IPTG至300 μ M的浓度,随后在恒定振荡下再孵育5-6小时来诱导蛋白质产生。最后通过离心收集所得生物质,并用例如50mM Tris-HCl缓冲液pH 7.5洗涤。将细胞作为沉淀储存在4 $^{\circ}$ C或-20 $^{\circ}$ C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),与所用的培养基无关。

发酵罐中的生物催化剂生产

在750ml InforsHT反应器中准备和运行发酵。向发酵容器中加入168ml去离子水。反应容器配备有所有需要的探头(pO_2 , pH, 采样, 消泡剂), C+N进料和氢氧化钠瓶, 高压灭菌。高压灭菌后, 将如下成分加入到反应器中:

20ml 10x磷酸盐/柠檬酸缓冲液

14ml 50%葡萄糖

0.53ml $MgSO_4$ 溶液

2ml $(NH_4)_2SO_4$ 溶液

0.020ml微量元素溶液

0.400ml硫胺素溶液

0.200ml卡那霉素储备溶液

运行参数设置如下: $pH=6.95$, $pO_2=40\%$, $T=30^{\circ}C$, 以300rpm搅拌。级联: rpm设定点为300, min 300, max 1000, 流速(L/min)设定点0.1, min 0, max 0.6。消泡剂对照: 1:9。

将发酵罐从种子培养物接种至 OD_{650nm} 0.4-0.5。使该种子培养物在LB培养基(+卡那霉素)中在37 $^{\circ}$ C、220rpm下生长8h。发酵首先以分批模式运行11.5h, 之后开始用补料溶液(无菌葡萄糖溶液(143ml H_2O +35g葡萄糖)进行C+N补料, 灭菌后向其中加入: 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卡那霉素储备溶液。进料以约4.2ml/h的恒定流速运行。葡萄糖和 NH_4^+ 在外部进行测量以评估培养物中C-源和N-源的可利用性。葡萄糖水平通常极低。

培养物生长总计约25小时, 其中它们通常达到 OD_{650nm} 40-45。然后通过加入IPTG至在发酵罐中1mM的浓度启动SHC生产(作为IPTG脉冲或使用输注注射器在3-4小时的时间期限内), 将温度设定至40 $^{\circ}$ C和 pO_2 设定至20%。在40 $^{\circ}$ C诱导SHC生产持续16h。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pH 5.4洗涤, 并在4 $^{\circ}$ C或-20 $^{\circ}$ C下作为沉淀储存直至进一步使用。

实施例2-GC分析

用适当体积的叔丁基甲基醚(MBTE/tBME)提取样品以定量其在底物和反应产物中的含量。在用气相色谱分析之前, 通过离心使溶剂级分与水相分离。将1 μ L溶剂相注射(分流比10)到30m x 0.32mm x 0.25 μ m Zebron ZB-Wax柱上。柱在恒定流速(4ml/min H_2)与温度梯度下展开: 100 $^{\circ}$ C, 15 $^{\circ}$ C/min至200 $^{\circ}$ C, 120 $^{\circ}$ C/min至240 $^{\circ}$ C, 4min, 240 $^{\circ}$ C。入口温度: 250 $^{\circ}$ C, 检测器温度: 250 $^{\circ}$ C。这导致底物和产物峰分离。

羟基法呢基丙酮转化率由相应于底物和反应产物的峰的面积用下式计算:

$$\text{转化率}(\%) = 100 \times (\text{面积}_{\text{产物峰}} / (\text{面积}_{\text{产物峰}} + \text{面积}_{\text{底物峰}}))$$

实施例3-使用SHC酶和酶变体筛选羟基法呢基丙酮环化

如实施例1所述在大肠杆菌中产生用于羟基法呢基丙酮环化反应的SHC酶集合。该反应体系包含1g/1羟基法呢基丙酮和细胞至OD_{650nm} 10。应用如表2中所列的反应条件。在表2中列出的最佳反应条件下测试的SHC酶的性能概述显示在图2中。

表2. SHC酶和反应条件。

SHC 酶	SEQ ID NO	温度(°C)	pH*	[SDS](w/w%)
AacSHC	1	55	6.0	0.05
TelSHC	18	45	6.4	0.075
ApaSHC1	19	45	5.2	0.075
ZmoSHC1	15	40	5.2	0.005
ZmoSHC2	16	35	6.4	0.005
BjaSHC	17	50	5.8	0.005
BmeSHC	21	45	5.6	0.0025
GmoSHC	20	40	5.6	0.0075
215G2 SHC	13	35	5.4	0.07
115A7 SHC	11	35	5.4	0.05
110B8 SHC	9	35	5.8	0.05
90C7 SHC	7	45	5.6	0.04
SHC#49	22	45	5.6	0.05
SHC#65	3	45	5.6	0.07
SHC#66	5	45	5.6	0.06

*在柠檬酸/磷酸钠缓冲液中

所有测试的酶都能够将羟基法呢基丙酮环化成(+) -龙涎缩醛。野生型酶的转化率在2%至约90%之间,并且在BmeSHC的情况下最高,如图2所示。如使用在图2中概括的SHC变体215G2 SHC、115A7SHC、110B8 SHC、90C7 SHC、SHC#49、SHC#65、SHC#66中观察到的,当将突变导入野生型AacSHC时,羟基法呢基丙酮环化增加。

实施例4-使用SHC酶的羟基法呢基丙酮环化

分别在包含2和8g/1羟基法呢基丙酮的反应中用产生ZmoSHC1或BmeSHC酶的细胞测试反应中的羟基法呢基丙酮环化。反应包含细胞至OD_{600nm} 80。反应在50mM琥珀酸/NaOH缓冲液pH 5.2中进行,并在35°C下孵育24小时。溶剂萃取反应的GC-FID分析表明,在2g/1底物下运行的反应中,对于ZmoSHC1和BmeSHC,羟基法呢基丙酮转化率分别为52%和90%。在8g/1下,使用ZmoSHC1和BmeSHC的羟基法呢基丙酮转化率分别为17%和79%。

实施例5-使用角鲨烯何帕烯环化酶变体酶的羟基法呢基丙酮的环化

如实施例1所概括的生产酶SHC变体215G2并用于羟基法呢基丙酮的环化反应。

典型的反应(150g总体积)如下在0.75升Infors发酵罐中设置。向反应容器中装入

羟基法呢基丙酮(0.75g, 2.7mmol), 从在去离子水中制备的31% (w/w) 溶液中加入1.95g SDS。通过将细胞悬浮于0.1M琥珀酸/NaOH缓冲液pH 5.1中, 由已产生关注的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10min测定该细胞悬浮液的细胞湿重浓度后, 将适当体积的细胞加入反应容器中以将37.5g细胞导入反应中。用所需量的反应缓冲液pH 5.1将反应体积补足至150g。反应在35℃和pH 5.4下在恒定搅拌(700rpm)下进行。使用85% H_3PO_4 将pH设定为5.4。根据需要使用85%磷酸手动进行pH调节。将反应随时间取样(1ml), 用5体积的MTBE/tBME(5ml)萃取。在通过离心(台式离心机, 13000rpm, 2min)澄清溶剂相后, 通过GC分析测定反应的底物和产物含量。

大约在72小时的反应之后, 通过剧烈振荡用100ml MTBE萃取反应体系5次, 然后通过离心(6000g, 10min, 室温)进行相分离。将粗萃取物经硅胶过滤并将滤液浓缩。将残余物(2.4g)通过硅胶柱色谱法(用在正庚烷中的7-70%乙酸乙酯的梯度洗脱)进行纯化。从其中分离出(+)-龙涎缩醛(式(I)的化合物, 173mg, 23%, 淡黄色固体)和(Z)-5-((5aR, 9aR)-6, 6, 9a-三甲基八氢苯并[b]噁庚因-3(2H)-亚基)戊-2-酮(式(III)的化合物, 65mg, 9%, 灰白色固体)。

(+)-龙涎缩醛的表征。 $[\alpha]_D^{23} = +16.4^\circ$ ($c = 0.42$, $CHCl_3$)。TLC(硅胶, 庚烷/EtOAc 3:2): $R_f = 0.54$ 。

^1H -NMR(400MHz, $CDCl_3$) 4.32(d, $J = 6.8$ Hz, 1H), 3.37(dd, $J = 7.1, 1.2$ Hz, 1H), 1.90(dt, $J = 12.7, 3.4$ Hz, 1H), 1.82(dd, $J = 13.7, 4.6$ Hz, 1H), 1.49-1.78(m, 8H), 1.43-1.47(m, 1H), 1.42(s, 3H), 1.10-1.24(m, 3H), 0.99(dd, $J = 12.2, 2.0$ Hz, 1H), 0.90(s, 6H), 0.83-0.95(m, 1H), 0.82(s, 3H)。 ^{13}C -NMR(101MHz, $CDCl_3$): 106.0(s), 82.6(s), 73.5(t), 55.7(d), 53.3(d), 41.8(t), 38.7(t), 37.3(s), 36.2(t), 35.9(t), 33.6(q), 33.1(s), 24.3(q), 21.7(q), 20.0(t), 18.3(t), 17.4(t), 14.6(q)。

EI-MS(70eV): 278(M^+ , <1), 263(<1), 248(2), 236(4), 218(36), 203(19), 190(55), 175(43), 162(11), 147(24), 137(34), 121(42), 109(47), 95(34), 79(38), 69(36), 55(42), 43(100)。

(Z)-5-((5aR, 9aR)-6, 6, 9a-三甲基八氢苯并[b]噁庚因-3(2H)-亚基)戊-2-酮的表征

TLC(硅胶, 庚烷/EtOAc 3:2): $R_f = 0.29$ 。

^1H -NMR(600MHz, 苯- d_6) δ ppm 5.01(br t, $J = 7.3$ Hz, 1H), 4.44(d, $J = 16.6$ Hz, 1H), 4.29(br d, $J = 16.2$ Hz, 1H), 2.47(ddd, $J = 12.4, 7.9, 3.8$ Hz, 1H), 2.22(ddd, $J = 12.7, 8.6, 4.3$ Hz, 1H), 2.08(q, $J = 7.2$ Hz, 2H), 1.94(t, $J = 1.0$ Hz, 2H), 1.73-1.62(m, 3H), 1.61(s, 3H), 1.49-1.22(m, 4H), 1.21-1.14(m, 2H), 1.19(s, 3H), 0.87(s, 3H), 0.72(s, 3H)。 ^{13}C -NMR(151MHz, 苯- d_6) δ ppm 205.59, 143.75, 120.82, 78.59, 62.81, 55.39, 42.97, 42.28, 41.65, 35.76, 35.38, 33.45, 29.42, 26.21, 21.56, 21.56, 20.89, 19.64。

EI-MS(70eV): 278(M^+ , <1), 260(<1), 245(<1), 220(1), 141(16), 135(15), 123(24), 109(32), 95(25), 81(19), 69(12), 55(13), 43(100)。

实施例5A-使用野生型BmeSHC的羟基法呢基丙酮的环化

如实施例1中所概括的生产野生型BmeSHC, 并将其用于羟基法呢基丙酮的环化反应。

反应(4ml体积)包含在0.2M乙酸/乙酸钠缓冲液pH 5.2中的135g/1羟基法呢基丙酮、221g/1细胞(细胞湿重)、0.09%SDS。将该反应体系在45℃下在恒定搅拌(650rpm, Radleys Carousel)下温育。一个反应用作对照,随时间取样,用MTBE萃取样品,通过离心(台式离心机,13000rpm,2min)澄清溶剂相,并通过GC分析其底物和产物含量。开始后50h,羟基法呢基丙酮转化完全或几乎完全(100%),预期约920mg的(+)-龙涎缩醛。

将合并的反应物离心(4500g,4℃,15min)。回收沉淀并用20ml去离子水洗涤3次(剧烈振荡+离心),弃去水相。最终将沉淀重悬于15ml去离子水中,并用15-20ml MTBE萃取3次(剧烈振荡+离心)。收集有机相并分析其(+)-龙涎缩醛含量。将合并的有机相过滤(硅胶),在氮气流下蒸发溶剂,得到约720mg干燥的结晶粉末。将残余物(720mg)通过硅胶柱色谱法纯化,用在正庚烷中的6-50%MTBE的梯度洗脱。合并包含纯(+)-龙涎缩醛的级分,蒸发溶剂。分离出560mg(+)-龙涎缩醛(式(I)的化合物(收率约61%))。

(+)-龙涎缩醛的表征。 $[\alpha]_D^{25} = +22.6^\circ$ ($c = 0.94$, CHCl_3)。TLC(硅胶,正庚烷/MTBE 3:1): $R_f = 0.55$ 。NMR:光谱数据与实施例5中所示的数据一致。

实施例6-羟基法呢基丙酮(也参见图1)

1a) (E)-6,10-二甲基-1-((四氢-2H-吡喃-2-基)氧基)十一碳-5,9-二烯-2-酮(4)的制备:在氮气气氛中向在1L高压釜中的4-(苄基氧基)-3-氧代丁酸乙酯(1,60g,0.25mol,1当量)在乙醇(600mL)中的溶液中加入10%披钨碳(6.0g,10%w/w),并将该混合物在3atm氢气中搅拌12h。通过C盐(celite)床过滤该反应混合物,并用二氯甲烷/乙醇的1:1混合物洗涤。真空蒸发滤液,得到淡黄色残余物,将其溶于二氯甲烷(600mL)。向该溶液中加入3,4-二氢-2H-吡喃(42.7g,0.51mol,2当量)和PPTS(6.3g,0.025mol;0.1当量),并将该混合物在室温下搅拌16h,然后加入水(500mL),随后用二氯甲烷(2x 200mL)萃取。将合并的有机层用盐水洗涤,用 Na_2SO_4 干燥并真空浓缩,得到粗产物,将其通过硅胶柱色谱法纯化,用乙酸乙酯的石油醚溶液(5-8%)洗脱,得到(E)-5,9-二甲基-2-(2-((四氢-2H-吡喃-2-基)氧基)乙酰基)癸-4,8-二烯酸乙酯(2),为淡黄色液体(45g,77%)。将该产物(44g,0.19mol,1.0当量)溶于THF(440mL),并加入碳酸钾(31.6g,0.23mol,1.2当量)。将该悬浮液在室温下搅拌1h,然后在0℃下加入(E)-1-溴-3,7-二甲基辛-2,6-二烯(香叶基溴,37.3g,0.172mol,0.9当量),并将该混合物在室温下搅拌16h。将该混合物过滤并将滤饼用二氯甲烷洗涤。真空浓缩滤液,得到(E)-5,9-二甲基-2-(2-((四氢-2H-吡喃-2-基)氧基)乙酰基)癸-4,8-二烯酸乙酯(3,70g),将其溶于乙醇(500mL)。加入KOH(35g)在水(100mL)中的溶液,并将该混合物加热至80℃,历时2h。真空浓缩该溶液,得到残余物,将其溶于二氯甲烷(1L),并用水(2x 200mL)和盐水洗涤。将有机层用 Na_2SO_4 干燥并真空浓缩,得到粗油状物,将其通过硅胶柱色谱法、使用5-8%的乙酸乙酯的石油醚溶液纯化,得到(E)-6,10-二甲基-1-((四氢-2H-吡喃-2-基)氧基)十一碳-5,9-二烯-2-酮(4,20g,36%),为淡黄色液体。

1b) (3-(2-甲基-1,3-二氧戊环-2-基)丙基)三苯基磷碘化物(8)的制备:向5-碘戊-2-酮(45g,0.21mol,1当量)在(200mL)甲苯中的溶液中加入三苯基磷(66.8g,0.25mol,1.2当量)。将该混合物在120℃下搅拌16h,然后冷却至室温,此时固体沉淀,将其过滤,用乙醚(100mL)洗涤并真空干燥,得到(4-氧代戊基)三苯基磷碘化物(90g,90%),为浅棕色固体。将该产物(40g,0.10mol,1当量)悬浮于甲苯(400mL)中,并加入乙二醇(80mL)和对甲苯磺酸(1g,0.0167mol,0.16当量),并给装置配备迪安-斯达克冷凝器。将该混合物加热至130

℃, 历时12h, 然后冷却至室温并弃去甲苯层。将剩余的橡胶状残余物溶于二氯甲烷(500mL), 并用水(500mL)和盐水(200mL)洗涤。

将有机层用 Na_2SO_4 干燥并真空浓缩, 得到(3-(2-甲基-1,3-二氧戊环-2-基)丙基)三苯基膦碘化物(8.43g, 80%), 为浅棕色固体。

1c) 羟基法呢基丙酮的制备: 在-78℃下冷却(3-(2-甲基-1,3-二氧戊环-2-基)丙基)三苯基膦碘化物(52.8g, 0.10mol, 2当量)在THF(150mL)中的悬浮液, 然后添加1.6M n-BuLi的己烷溶液(63.8mL, 0.10mol, 2当量)。将该混合物在室温下搅拌30min, 此时形成橙黄色悬浮液, 再将其冷却至-78℃。滴加(E)-6,10-二甲基-1-((四氢-2H-吡喃-2-基)氧基)十一碳-5,9-二烯-2-酮(4.15g, 0.05mol, 1当量)在THF(20mL)中的溶液。并在室温下持续搅拌2h。在该期限过程中, 反应混合物变成黄色悬浮液。将该反应混合物用冰水(100mL)淬灭, 并用乙酸乙酯(2x 100mL)萃取。将有机层用 Na_2SO_4 干燥并真空浓缩, 得到粗产物, 将其与己烷(4x 50mL)一起研磨, 并将沉淀的三苯基膦氧化物通过过滤除去。真空除去溶剂后, 得到2-(((2Z,5E)-6,10-二甲基-2-(3-(2-甲基-1,3-二氧戊环-2-基)亚丙基)十一碳-5,9-二烯-1-基)氧基)四氢-2H-吡喃(9.22g), 为淡黄色液体, 根据 $^1\text{H-NMR}$, 其仍然包含痕量的三苯基膦氧化物。将产物(22g, 0.05mol, 1当量)溶于丙酮(200mL), 并在0℃下加入1.5N HCl-水溶液(220mL)。将该溶液在室温下搅拌16h, 然后加入水(50mL), 并将混合物用乙醚(2x 100mL)萃取。将合并的有机层用盐水洗涤, 用 Na_2SO_4 干燥并真空浓缩, 得到粗油状物, 将其通过硅胶柱色谱法纯化, 用10-15%乙酸乙酯的石油醚溶液洗脱, 得到(5Z,9E)-6-(羟基甲基)-10,14-二甲基十五碳-5,9,13-三烯-2-酮(羟基法呢基丙酮, 7.5g, GC/MS纯度为81%), 通过在85℃/1mm/Hg下蒸馏从其中除去挥发性杂质。使残余物(5.8g)进行第二次硅胶柱色谱, 用10-15%乙酸乙酯的石油醚溶液洗脱, 得到羟基法呢基丙酮(2.8g, 19%), 为淡黄色液体。

$^{13}\text{C-NMR}$ (d_6 -DMSO, 100MHz): 208.4 (s), 140.2 (s), 134.7 (s), 131.1 (s), 125.0 (d), 124.6 (d), 124.6 (d), 58.5 (t), 43.6 (t), 39.7 (2t), 35.0 (t), 30.1 (q), 26.7 (t), 25.9 (q), 21.9 (t), 18.0 (q), 16.2 (q)。

序列表

SEQ ID NO:1 (野生型酸热脂环酸芽孢杆菌SHC (AacSHC) 的氨基酸序列)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVP
MVPPEIMFLGKRMP LNIYEF GSWARATVVALSIVMSRQPVFPLPERARVPELYETDVP PRRRGAKGGGGWIFDALDR
ALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLYGVLDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDNDNTSDLPNHIPFCDFGEVTDPPSEDVTAHVLE
CFGSGFYDDAWKVIRRAVEY LKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAAARRGVQYL VETQRPDGGWDEPYTGTGFP GDFYLG
YTMYRHVFPTLALGRYKQAIERR

SEQ ID NO:2 (编码野生型AacSHC的核苷酸序列)

ATGGCTGAGCAGTTGGTGAAGCGCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG

ACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGTTACGCGGATGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGA
TCATGTTCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTTCCTCGCTGCCCCGAGCGGGCGCGCTGCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGCGCTGCACGGGTATC
AGAAGCTGTGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTACGCCGCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTTATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCGGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTGGGGCGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACATCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCACGTGCTCGAGTGTTTCGGCAGCTTCG
GGTACGATGACGCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTGCGCGCTGAAGGCGGTGGGATCGACAC
GCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGACCCCGTCGAGACGGCCTGGGCGCTGATGGCG
CTCATCGCGGGCGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGTGGGATGAGCCGTACTACACCGGCACGGGCTTCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

SEQ ID NO:3 (AacSHC酶变体#65的氨基酸序列)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDR
VLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQQPWFYALIALKILDMTQHPAFIKWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFQFDNVYYPDVDD
TAVVVWALNTRLRPDERRRRDAMTKGFRWIVGMQSSNGGWGAYVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLE
CFGFSGYDDAWKVIRRAVEYLLKREQKPDGSWFGWGVNLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTTWALMALIAGGRAESEAAARRGVQYL VETQRPDGGWDEPYTGTGFPDGYLG
YTMYSHVFTLALGRYKQAIERR

SEQ ID NO:4 (编码SHC酶变体#65的核苷酸序列)

ATGGCTGAGCAGTTGGTGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGTTACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGA

TCATGTTCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTTCCTCGCTGCCCCGAGCGGGCGCGCTGCCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGTGGGTGGATCTTCGACGCGCTCGACCGGTGCTGCACGGGTATC
AGAAGCTGTGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTACGCCGCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCGGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTTCGCTTCCAGTTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGACGAGCGCCGAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTGGGGCGCCTACGACGTGACAACACGAGCGATCTCCGAACCACACCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCACGTGCTCGAGTGTTCGGCAGCTTCG
GGTACGATGACGCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTGCGCGCTGAAGGCGGTGGGATCGACAC
GCGCGAGCCGTACATTCAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGCACCCGTCGCAGACGACCTGGGCGCTGATGGCG
CTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGTGGGATGAGCCGTACTACACCGGCACGGGCTTCCAGGGGATTTCTACCTCGGCTACACCATGTACAGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

SEQ ID NO:5 (AacSHC酶变体#66的氨基酸序列)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALHPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPVRRRGAAGGGGWIFDALDR
VLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQQPWFYALIALKILDMTQHPAFIKWEGLELYGVLDYD
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDNDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLE
CFGSFYDDAWKVIRRAVEYLKREQKPDGSWFGWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTTWALMALIAGGRAESEAAARRGVQYL VETQRPDGGWDEPYTGTGFPDGYL
YTMYSHVFTLALGRYKQAIERR

SEQ ID NO:6 (编码SHC酶变体#66的核苷酸序列)

ATGGCTGAGCAGTTGGTGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGCACCCGGGTGGGCCCGGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGTTTATTAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGTTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGAGGA
TCATGTTCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTTCCTCGCTGCCCCGAGCGGGCGCGCTGCCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGTGGGTGGATCTTCGACGCGCTCGACCGGTGCTGCACGGGTATC

AGAAGCTGTCGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGA CTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTACGCCGCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCGGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT
GGGCGCTCAACACCCTGCGTTGCCGACGAGCGCCGAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTGGGGCGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACACCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCACGTGCTCGAGTGTTTCGGCAGCTTCG
GGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTGCGCGCTGAAGGCGGTGGGATCGACAC
GCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGGAGACCCCGTCGACAGACCTGGGCGCTGATGGCG
CTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGCGGTGCAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGTGGGATGAGCCGTACTACACCGGCACGGGCTCCCAGGGATTCTACCTCGGCTACACCATGTACAGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

SEQ ID NO:7 (AacSHC酶变体#90C7的氨基酸序列)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLATIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPIRRRGAKGGGGWIFDALDR
VLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQQPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLE
CFGSGFYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPIYQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAAARRGVQYL VETQRPDGGWDEPYTGTGFPDGYLG
YTMYSHVFTLALGRYKQAIERR

SEQ ID NO:8 (编码SHC变体#90C7的核苷酸序列)

ATGGCTGAGCAGTTGGTGGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCCGCGACCTCGACGCGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTACAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGTTACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGA
TCATGTTCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGTGTTCCCGCTGCCCCAGCGGGCGCGCTGCCCCAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGTGGGTGGATCTTCGACGCGCTCGACCGGTGCTGCACGGGTATC
AGAAGCTGTCGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGA CTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTACGCCGCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG

CTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCGGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTTGGGGCGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACACCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTACAGAGGACGTCACCGCCACGTGCTCGAGTGTTCGGCAGCTTCG
GGTACGATGACGCTGGAAGGTCATCCGGCGCGGGTGGAAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTGCGCGCTGAAGGCGGTGGGATCGACAC
GCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTGCGAGCAGCATCAGAACCCGGACGGCGGTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGCACCCCGTCGAGACGGCCTGGGCGCTGATGGCG
CTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGAATACCTCGTGGAGACGAGCGCCCGGA
CGGCGGTGGGATGAGCCGTACTACACCGGCACGGGCTTCCAGGGGATTTCTACCTCGGCTACACCATGTACAGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGTACAAGCAAGCCATCGAGCGCAGGTGA

SEQ ID NO:9 (AacSHC酶变体#110B8的氨基酸序列)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALHPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMPNLIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPRRRGAAGGGGWIFDALDR
VLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQQPWFYALIALKILDMTQHPAFIKWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNLTPFCDFGEVTDPPSEDVTAHVLE
CFGSGFYDDAWKVIRRAVEYLKREQKPDGSWFRGWVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTTWALMALIAGGRAESEAAARRGVQYL VETQRPDGGWDEPYTGTGFPDFYLG
YTMYRHVFTLALGRYKQAIERR

SEQ ID NO:10 (编码SHC酶变体#110B8的核苷酸序列)

ATGGCTGAGCAGTTGGTGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCAGG
ACGGCACGTGGGCCCTGCACCCGGGTGGGCCCGCGGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGTTACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGGA
TCATGTTCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTTCGCGTCCCCGAGCGGGCGCGCTGCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGTGGGTGGATCTTCGACGCGCTCGACCGGTGCTGCACGGGTATC
AGAAGCTGTGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTACGCCGCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCGGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT

GGGCGCTCAACACCCTGCGCTTGCCGACGAGCGCCGAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTTGGGGCGCTACGACGTCGACAACACGAGCGATCTCCCGAACCTCACCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCACGTGCTCGAGTGTTTCGGCAGCTTCG
GGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTGCGGCGCTGAAGGCGGTGCGGATCGACAC
GCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTGCGAGCAGCATCAGAACCCGGACGGCGGTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGACCCCGTCGAGACGACCTGGGGCGTGTATGGCG
CTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGGTGCAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGTGGGATGAGCCGTACTACACGGCACGGGCTTCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

SEQ ID NO:11 (AacSHC酶变体#115A7的氨基酸序列)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMLNIYEFGSWARTTVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDR
VLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQQPFYALIALKILDKTQHPAFIKGWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDNDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLE
CFGSGFYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAAARRGVQYL VETQRPDGGWDEPYTGTGFPDGYLG
YTMVRHVFTLALGRYKQAIERR

SEQ ID NO:12 (编码SHC变体#115A7的核苷酸序列)

ATGGCTGAGCAGTTGGTGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGTTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGA
TCATGTTCTCGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGACGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTCCCGCTGCCCCGAGCGGGCGCGCTGCCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGTGGGTGGATCTTCGACGCGCTCGACCGGTGCTGCACGGGTATC
AGAAGCTGTGCGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTACGCCGCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACAAGACGCAGCA
TCCGGCGTTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGTTGTGGGACACGGGCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCGGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTTCGCGTTCCAGTTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTGCTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGACGAGCGCCGAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTTGGGGCGCTACGACGTCGACAACACGAGCGATCTCCCGAACACACCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCACGTGCTCGAGTGTTTCGGCAGCTTCG

GGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTGCGGCGCTGAAGGCGGTCTGGGATCGACAC
GCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGGAGACCCCGTCGCAGACGGCCTGGGCGCTGATGGCG
CTCATCGCGGGCGGAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTCGTGAGACGCAGCGCCCGGA
CGGCGGTGGGATGAGCCGTACTACACGGCACGGGCTTCCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGTACAAGCAAGCCATCGAGCGCAGGTGA

SEQ ID NO:13 (AacSHC酶变体215G2的氨基酸序列)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPIRRRGAAGGGWIFDALDR
VLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQQPWFYALIALKILDMTQHPAFIKWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLE
CFGSFYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAAARRGVQYLVETQRPDGGWDEPYTGTGFPDGYLG
YTMVRHVFTLALGRYKQAIERR

SEQ ID NO:14 (编码Aac 215G2 SHC酶变体的核苷酸序列)

ATGGCTGAGCAGTTGGTGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGTTACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGAGA
TCATGTTCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGTGTTCCCGCTGCCCGAGCGGGCGCGGTGCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGTGGGTGGATCTTCGACGCGCTCGACCGGTGCTGCACGGGTATC
AGAAGCTGTGCGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTACGCCGCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTCCGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCTGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGACGAGCGCCGAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTGGGGCGCCTACGACGTCGACAACACGAGCGATCTCCGAACCACACCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCACGTGCTCGAGTGTTTCGGCAGCTTCG
GGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTGCGGCGCTGAAGGCGGTCTGGGATCGACAC
GCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGTGGGGCGAGGACT

GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGACCCCGTCGCAGACGGCCTGGGCGCTGATGGCG
CTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGCTGGGATGAGCCGTACTACACCGGCACGGGCTTCCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGTACAAGCAAGCCATCGAGCGCAGGTGA

SEQ ID NO:15 (野生型ZmoSHC1的氨基酸序列)

MGIDRMNSLSRLLMKKIFGAEKTSYKPASDTIIGTDTLKRPNRRPEPTAKVDKTIIFKTMGNSLNNTLVS
ACDWLIGQQKPDGHWVGAVESNASMEAWECLALWFLGLEDHPLRPRLGNALLEMQREDGSWGVYFGAGNGDINATVE
AYAALRSLGYSADNPVLKKAAWIAEKGGKLNIRVFTRYWLALIGEWPEKTPNLPPEIIWFDPDNFVFSIYNFAQWA
RATMVPIAILLSARRPSRPLRPQDRLELDFPEGRARFDYELPKKEGIDLWSQFFRTTDRGLHWVQSNLLKRNSLREAA
IRHVLEWIIIRHQDADGGWGGIQQPWVYGLMALHGEQYQLYHPVMAKALSALDDPGWRHDRGESSWIQATNSPVWDTM
LALMALKDAKAEDRFTPEMDKAADWLLARQVKVKGDSIKLPDVEPGGWAFEYANDRYPD TDDTAVALIALSSYRDK
EEWQKKGVEDAITRGVNWLIA MQSECGGWGAFDKDNNRSILSKIPFCDFGESIDPPSVDVTAHVLEAFGTLGLSRDM
PVIQKAIDYVRSEQEAEGAWFGRWGVNYIYGTGAVLPALAAIGEDMTQPYITKACDWLVAHQEDGGWGESCSSYME

SEQ ID NO:16 (野生型ZmoSHC2的氨基酸序列)

MTVSTSSAFHHSPLSDDVEPIIQKATRALLEKQQQDGHVFELEADATIPA EYILLKHYLGEPEDLEIE
AKIGRYLRRIQGEHGGWSLFYGGDLDSATVKAYFALKMIGDSPDAPHMLRARNEILARGGAMRANVFTRIQLALFG
AMSWEHVPQMPVELMLMPEWFPVHINKMAYWARTVLVPLLVLQALKPVARNRGILVDELFPVDVLP TLQESGDP IW
RRFFSALDKVLHKVEPYWPKNMRKA IHSCVHFVTERLNGEDGLGAIYPAIANSVMMYDALGYPENHPERA IARRAV
EKLMLVDGTEDQGDKVEYVCQPCLSPIWDTALVAHAMLEVGGDEAEKSAISALSWLKPQQILDVKGDWAWRRPD LRP
GWAFQYRNDYYPDVDDTAVVTMAMDRAAKLSDLHDDFEESKARAMEWTIGMQSDNGGWGAFDANN SYTYLNNIPFAD
HGALLDPPTVDVSARCVSMMQAQAGISITDPKMKAAVDYLLKEQEEDGSWFGRWGVNYIYGTWSALCALNVAALPHDH
LAVQKAWAWLKTIQNEDGGWGENCDSYALDYSGYEPMSTASQTAWALLGLMAVG EANSEAVTKGINWLAQNQDEEG
LWKEDYYSGGGFPRVFYLRHYGSKYFPLWALARYRN LKANQPIVHYGM

SEQ ID NO:17 (野生型BjaSHC的氨基酸序列)

MTVTSSASARATRDPGNYQTALQSTVRAAADWLIANQKPDGHWVGRAESNACMEAQWCLALWFMGLEDH
PLRKRLGQSLLDSQRPDGAWQVYFGAPNGDINATVEAYAALRSLGFRDDEPAVRRAREWIEAKGGLRNIRVFTRYWL
ALIGEWPEKTPNIPPEVIWFPLWFPFSIYNFAQWARATLMP IAVLSARRPSRPLPPENRLDALFPHGRKAFDYELP
VKAGAGGWDRFFRGADKVLHKLQNLGNRLNLGLFRPAATSRVLEWMIRHQDFDGA WGGIQQPWIYGLMALYAEGYPL
NHPVLAKGLDALNDPGWRVDVGDATYIQATNSPVWDTIL TLLAFDDAGVLGDYPEAVDKAVDWVLQRQVRVPGDWSM
KLPHVKPGGWAFEYANNYPDTDDTAVALIALAPLRHDPKWKAKGIDEA IQLGVDWLIGMQSQGGGWGAFDKDNNQK
ILTKIPFCDYGEALDPPSVDVTAHII EAFGKLGISRNHPSMVQALDYIRREQEPSGPWFGRWGVNYVYGTGAVLPAL
AAIGEDMTQPYIGRACDWLVAHQADGGWGESCASYMDVSAVGRGTTTASQTAWALMALLAANRPQDKDAIERGCMW
LVERQSAGTWDEPEFTGTGFGYGVGQTIKLN DPALSQRMLMQPELSRAFMLRYGMYRHYFPLMALGRALRPQSHS

SEQ ID NO:18 (野生型TelSHC的氨基酸序列)

MPTSLATAIDPKQLQQAIRASQDFLFSQQYAEGYWWAELESNTMTAEVILLHKI WGT EQRLPLAKAEQ
YLRNHQRDHGGWELFYGDGDLSTSVEAYMGLRLLGVPETDPALVKARQFILARGGISKTRIFTKLHLALIGCYDWR
GIPSLPPWIMLLPEGSPFTIYEMSSWARSSSTVPLLIVMDRKPVYGMPPITLDELYSEGRANVVWELPRQGDWRDVF
IGLDRVFKLFETLNIHPLREQGLKAAEEWVLERQEASGDWGGIIPAM LNSLLALRALDYAVDDPIVQRGMAAVDRFA
IETETERYRVQPCVSPVWDTALVMRAMVDSGVAPDHPALVKAGEWLLSKQILDYGDWHIKNKKGRPGGWAFEFENRFY

PDVDDTAVVVMALHAVTLPNENLKRRRAIERAVAWIASMQCRPGGWAAFDVDNDQDWLNGIPYGLKAMIDPNTADVT
ARVLEMVGRCLAFDRVALDRALAYLRNEQEPEGCWFGRWGVNYLYGTSGVLTALSLVAPRYDRWRIRRAEWLMQC
QNADGGWGETCWSYHDP SLKKG DSTASQTAWAI IGLLAAGDATGDYATEA IERGIAYLLETQRPDGTWHEDYFTGT
GFPCHFYLKYHYQQHFPLTALGRYARWRNLLAT

SEQ ID NO:19 (野生型ApaSHC1的氨基酸序列)

MNMA SRFS LKKIL RSGSDTQGTNVNTLIQSGTSDIVRQKPAPQEPADLSALKAMGNSLTHTLSSACEWL
MKQKQPDGHWVGSVGSNASMEAEWCLALWFLGLEDHPLRPRLGKALLEMQRPDGSWGTYYGAGSGDINATVESYAAL
RSLGYAEDDPAVSKAAAWIISKGGLKNVRVFTRYWLALIGEWPWEKTPNLPPEIIWFPDNFVFSIYNFAQWARATMM
PLAILSARRPSRPLRPQDRLDALFPGGRANFDYELPTKEGRDVIADFFRLADKGLHWLQSSFLKRAPSREAAIKYVL
EWIIWHQDADGGWGGIQQPPWVYGLMALHGEQYQFHHVPMKALDALNDPGWRHDKGDASWIIQATNSPVWDTMLSLMA
LHDANAEEFTPEMDKALDWLLSRQVRVKGDWSVKLPNTEPGGWAFEYANDRYPD TDDTAVALIAIASCRNRPEWQA
KGVEEAIGRGVRWL VAMQSSCGGWGAFDKDNKSILAKIPFCDFGEALDPPSVDVTAHVLEAFGLLGLPRDLPCIQR
GLAYIRKEQDPTGPWFGWGVNYLYGTGAVLPALAAALGEDMTQPYISKACDWLINCCQENGWGESCASEMEVSSIG
HGATTPSQTAWALMGLIAANRPQDYEAIAKGCRYLIDLQEEDGSWNEEEFTGTGFPYGVGQTIKLDDPAISKRLMQ
GAELSRFMLRYDLYRQLFPIIALSRASRLIKLGN

SEQ ID NO:20 (野生型GmoSHC的氨基酸序列)

MSPADISTKSSSFQRLDNMLPEAVSSACDWLIDQKQPDGHWVGPVESNACMEAQWCLALWFLGQEDHPL
RPRLAQALLEMQREDGSWGIYVGADHGDINTTVEAYAALRSMGYAADMPIAKSAAWIQQKGGLRNVRVFTRYWLAL
IGEWPWDKTPNLPPEIIWLPDNFIFSIYNFAQWARATMMPLTILSARRPSRPLLPENRLDGLFPEGRENFDYELPVK
GEEDLWGRFFRAADKGLHSLQSFPVRRFVPREAAIRHVIEWIIIRHQDADGGWGGIQQPWIYGLMALSVEGYPLHHPV
LAKAMDALNDPGWRRDKGDASWIIQATNSPVWDTMLAVLALHDAGAEDRYSPQMDKAIGWLLDRQVRVKGDWSIKLPD
TEPGGWAFEYANDKYPDTDDTAVALIALAGCRHRPEWRERDIEGAISRGVNWLLAMQSSSGGWGAFDKDNRSILTK
IPFCDFGEALDPPSVDVTAHVLEAFGLLGISRNHPSVQKALAYIRSEQERNGAWFGWGVNYVYGTGAVLPALAAIG
EDMTQPYIVRACDWLMSVQQENGWGESCASEMDINAVGHGVATASQTAWALIGLLAAKRPKDREAIARGCQFLIER
QEDGSWTEEEYTGTFPGYGVGQAIKLDDPSLPDRLLQGAELSRFMLRYDLYRQYFPMALSRRARRMMKEDASAAA

SEQ ID NO:21 (野生型BmeSHC的氨基酸序列)

MIILLKEVQLEIQRR IAYLRPTQKNDGSFRYCFETGVMPDAFLIMLLRTFDLDKEVLIKQLTERIVSLQ
NEDGLWTLFDDEEHNLSATIQAYTALLYSGYYQKNDRILRKAERYIIDSGGISRAHFLTRWMLSVNGLYEWPCLFY
PLSLLLVPYVPLNFYELSTYARIHFVPMVAGNKKFSLTSRHTPSLSHLDVREQKQSEETTQESRASIFLVDHLK
QLASLPSYIHKLGYYAAERYMLERIEKDGTLYSYATSTFFMIYGLLALGYKKDSFVIQKAIDGICSLSTCSGHVHV
ENSTSTVWDTALLSYALQEAGVPQQDPMIKGTTRYLKKRQHTKLGDWQFHNPNTPAGGWGFSINTNNPDLDTSAA
IRALSRAQTDTDYLESWQRGINWLLSMQNKDGGFAAFEKNTDSILFTYLPENAKDAATDPATADLTGRVLECLGN
FAGMNKSHPSIKA AVKWLFDHQLDNGSWYGRWVCYIYGTWAAITGLRAVGVSASDPRIIKAINWLKSIQQEDGGFG
ESCYSASLKKYVPLSFSTPSQTAWALDALMTICPLKQSVKGIKFLNPNLTEQQTHYPTGIGLPQGFYIQYHSYN
DIFPLLALAHYAKKHSS

SEQ ID NO:22 (AacSHC酶变体#49的氨基酸序列)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMP LNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDR

VLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDNDNTSDLPNLTPFCDFGEVTDPPSEDVTAHVLE
CFGSGFYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPIYIQALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTTWALMALIAGGRAESEAAARRGVQYLVETQRPDGGWDEPYTGTGFPDGYLG
YTMYRHVFPTLALGRYKQAIERR

序列表

- <110> Givaudan SA
 <120> 酶介导的方法
 <130> 31049 PCT
 <160> 22
 <170> PatentIn version 3.5
 <210> 1
 <211> 631
 <212> PRT
 <213> 酸热脂环酸芽孢杆菌 (*Alicyclobacillus acidocaldarius*) 野生型SHC酶

<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
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
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 Asp	Ala Met
385	390	395
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
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> 1896

<212> DNA

<213> 酸热脂环酸芽孢杆菌野生型SHC酶

<400> 2

atggctgagc agttggtgga agcgccggcc tacgcgcgga cgctggatcg cgcggtggag 60
 tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
 acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
 atggagaaga tccggcggta cctgttgac gagcagcgcg aggacggcac gtggggccctg 240
 taccggggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
 atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
 ggccgggatcg agtcgtcgcg cgtgttcacg cggatgtggc tggcgctggt gggagaatat 420
 ccgtgggaga aggtgccccat ggtcccgcg gagatcatgt tcctcggcaa gcgcatgccg 480
 ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
 atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
 accgacgtgc ctccgcgccg gcgcggtgcc aaggaggagg gtgggtggat cttcgacgcg 660
 ctcgaccggg cgctgcacgg gtatcagaag ctgtcgggtc accggttccg ccgcgcggcc 720
 gagatccgcg ctttggaactg gttgctcgag gccaggccg gagacggcag ctggggcggg 780
 attcagccgc cttggtttta cgcgtcatc gcgtcaaga ttctcgacat gacgcagcat 840
 ccggcgttca tcaagggtg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
 ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
 ctgcgcgctg cggggcttcc ggccgatcac gaccgcttg tcaaggcggg cgagtggctg 1020
 ttggaccggc agatcacggt tccgggcgac tgggcggtga agcggccgaa cctcaagccg 1080
 ggccgggttcg cgttccagtt cgacaacgtg tactaccggg acgtggacga cacggccgtc 1140
 gtgggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200

acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgccctac 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 tgaaggcggg cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggtggtg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taaggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgcccgcg cgtgcaatac 1740
ctcgtggaga cgcagcgcgg ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggtacacc atgtaccgc acgtgtttcc gacgtcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 3

<211> 630

<212> PRT

<213> 人工序列

<220>

<223> SHC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
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
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
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

625 630

<210> 4

<211> 1896

<212> DNA

<213> 人工序列

<220>

<223> SHC酶变体#65

<400> 4

atggctgagc agttggtgga agctccggcc tacgcgcgga cgctggatcg cgcggtggag	60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc	120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc	180
atggagaaga tccggcggta cctgttgac gagcagcgcg aggacggcac gtgggcccctg	240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat	300
atcgcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag	360
ggcgggatcg agtcgtcgcg cgtgttcacg cggaggtggc tggcgctggt gggagaatat	420
ccgtgggaga aggtgcccac ggtccccgcg gagatcatgt tcctcggcaa gcgcatgccg	480
ctcaacatct acgagtttgg ctctgtgggct cgggcgaccg tcgtggcgct ctcgattgtg	540
atgagccgcc agccggtgtt cccgctgcc gagcgggcgc gcgtgcccga gctgtacgag	600
accgacgtgc ctccgcgccg gcgcggtgcc aaggaggagg gtgggtggat cttcgacgcg	660

```

ctcgaccggg tgctgcacgg gtatcagaag ctgtcgggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcggttca tcaagggtg ggaagggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcgggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactaccgag acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgccctac 1260
gacgtcgaca acacgagcga tctcccgaac cacacccgt 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 tgaaggcggg cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggtggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcacccgt cgcagacgac ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcgagg tccgaggccg cgcgcccgcg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtacagcc acgtgtttcc gacgctcgcg 1860
ctcgcccgct acaagcaagc catcgagcgc aggtga 1896

```

<210> 5

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> SHC酶变体#66

<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 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> 6

<211> 1896

<212> DNA

<213> 人工序列

<220>

<223> SHC酶变体#66

<400> 6

atggctgagc agttggtgga agctccggcc tacgcgcgga cgctggatcg cgcggtggag 60
 tatctcctct cctgcaaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120

acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
 atggagaaga tccggcggta cctgttgac gagcagcgcg aggacggcac gtgggccctg 240
 caccgggtg ggccgccga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
 atcggcgatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
 ggccgggatcg agtcgtcgcg cgtgttcacg cggaggtggc tggcgctggt gggagaatat 420
 ccgtgggaga aggtgccccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
 ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
 atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
 accgacgtgc ctccgcgccg gcgcggtgcc aaggaggagg gtgggtggat cttcgacgcg 660
 ctcgaccggg tgctgcacgg gtatcagaag ctgtcggtgc accggttccg ccgcgcggcc 720
 gagatccgcg ccttggaactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
 attcagccgc cttggtttta cgcgtcctc gcgtcaaga ttctcgacat gacgcagcat 840
 ccggcggttca tcaagggtg ggaaggctta gagctgtacg gcgtggagct ggattacgga 900
 ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
 ctgcgcgctg cggggcttcc ggccgatcac gaccgcttg tcaaggcggg cgagtggctg 1020
 ttggaccggc agatcacggt tccgggcgac tgggcggtga agcggccgaa cctcaagccg 1080
 ggccgggttcg cgttccagtt cgacaacgtg tactaccggg acgtggacga cacggccgtc 1140
 gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
 acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgctac 1260
 gacgtcgaca acacgagcga tctcccgaac cacacccgt tctgcgactt cggcgaagtg 1320
 accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
 tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
 ccggacggca gctgggttcg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
 gtgtcggcgc tgaaggcggc cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
 gactgggtcg agcagcatca gaaccgggac ggccgctggg gcgaggactg ccgctcgtac 1620
 gaggatccgg cgtacgcggg taaggcgcg agcaccctgt cgcagacgac ctgggcgctg 1680
 atggcgctca tcgcgggcgg caggcgggag tccgaggccg cgcgcccgcg cgtgcaatac 1740
 ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
 ttcccagggg atttctacct cggtacacc atgtacagcc acgtgtttcc gacgtcgcg 1860
 ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 7

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> SHC酶变体#90C7

<400> 7

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>	8	
<211>	1896	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	SHC酶变体#90C7	
<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 cctgttgac gagcagcgcg aggacggcac gtgggccctg	240	
tacccgggtg ggccgccgga cctcgacgcg accatcgagg cgtacgtcgc gctcaagtat	300	
atcggcgatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag	360	
ggcgggatcg agtcgtcgcg cgtgttcacg cggaggtggc tggcgtggt gggagaatat	420	
ccgtgggaga aggtgcccac ggtcccgcg gagatcatgt tcctcggcaa gcgcgtgccg	480	
ctcaacatct acgagtttgg ctctgtgggt cgggcgaccg tcgtggcgct ctcgattgtg	540	
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag	600	
accgacgtgc ctccgcgccg gcgcggtgcc aaggaggagg gtgggtggat cttcgacgcg	660	
ctcgaccggg tgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc	720	
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg	780	
attcagccgc ctltggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat	840	
ccggcggttca tcaagggtg ggaagggtcta gagctgtacg gcgtggagct ggattacgga	900	
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg	960	
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg	1020	
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcgccga cctcaagccg	1080	
ggcgggttcg cgttccagtt cgacaacgtg tactaccggg acgtggacga cacggccgtc	1140	
gtggtgtggg cgctcaacac cctgcgctt ccggacgagc gccgcaggcg ggacgccatg	1200	
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggtt gggcgcctac	1260	
gacgtcgaca acacgagcga tctcccgaac cacacccgt 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 gaaccgggac ggcggttggg gcgaggactg ccgctcgtac	1620	
gaggatccgg cgtacgcggg taagggcgcg agcaccctgt cgcagacggc ctgggcgctg	1680	
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgcccgcg cgtgcaatac	1740	
ctcgtggaga cgcagcgcgc ggacggcggc tgggatgagc cgtactacac cggcacgggc	1800	
ttcccagggg atttctacct cggctacacc atgtacagcc acgtgtttcc gacgtcgcg	1860	
ctcggccgct acaagcaagc catcgagcgc aggtga	1896	

<210> 9
 <211> 631
 <212> PRT
 <213> 人工序列
 <220>
 <223> SHC酶变体#110B8
 <400> 9
 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
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 Asp Ala Met		
385	390	395
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
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
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>	10					
<211>	1896					
<212>	DNA					
<213>	人工序列					
<220>						
<223>	SHC酶变体#110B8					
<400>	10					
atggctgagc agttggtgga agctccggcc tacgcgcgga cgctggatcg cgcggtggag						60
tatctcctct cctgcaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc						120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc						180
atggagaaga tccggcggta cctgttgac 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 aggtgcccac ggtcccgcg gagatcatgt tcctcggcaa gcgcatgccg						480
ctcaacatct acgagtttgg ctctgtgggct cgggcgaccg tcgtggcgct ctcgattgtg						540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag						600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg						660
ctcgaccggg tgctgcacgg gtatcagaag ctgtcgggtgc acccgttccg ccgcgcggcc						720
gagatccgcg ccttggaactg gttgctcgag cgccaggccg gagacggcag ctggggcggg						780
attcagccgc cttggtttta cgcgctcatc gcgtcaaga ttctcgacat gacgcagcat						840
ccggcggttca tcaagggtg ggaaggtcta gagctgtacg gcgtggagct ggattacgga						900
ggatggatgt ttcaggttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg						960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg						1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcggccgaa cctcaagccg						1080
ggcgggttcg cgttccagtt cgacaacgtg tactaccgg acgtggacga cacggccgtc						1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg						1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac						1260
gacgtcgaca acacgagcga tctcccgaac ctacccccgt 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 tgaaggcggc cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggtctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccg cgcagacgac ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcgcccgct acaagcaagc catcgagcgc aggtga 1896

```

<210> 11

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> SHC酶变体#115A7

<400> 11

```

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	
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	
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 Asp Ala Met					
385		390		395	
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	
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> 12

<211> 1896

<212> DNA

<213> 人工序列

<220>

<223> SHC酶变体#115A7

<400> 12

atggctgagc agttggtgga agctccggcc tacgcgcgga cgctggatcg cgcggtggag 60
 tatctcctct cctgcaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
 acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
 atggagaaga tccggcggta cctgttgac gagcagcgcg aggacggcac gtgggcccctg 240
 taccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
 atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
 ggcgggatcg agtcgtcgcg cgtgttcacg cggaggtggc tggcgctggt gggagaatat 420
 ccgtgggaga aggtgcccac ggtcccgcg gagatcatgt tcctcggcaa gcgcatgccg 480
 ctcaacatct acgagtttgg ctctgtgggt cggacgaccg tcgtggcgct ctcgattgtg 540
 atgagccgcc agccggtgtt cccgtgccc gagcgggcgc gcgtgcccga gctgtacgag 600
 accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttegacgcg 660
 ctcgaccggg tgctgcacgg gtatcagaag ctgtcgtgac acccgttccg ccgcgcggcc 720
 gagatccgcg ccttgactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
 attcagccgc cttggtttta cgcgtcatc gcgtcaaga ttctcgacaa gacgcagcat 840

```

ccggcggttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga      900
ggatggatgtt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg      960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg     1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg     1080
ggcgggttcg cgttccagtt cgacaacgtg tactaccggg acgtggacga cacggccgtc     1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg     1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgccctac     1260
gacgtcgaca acacgagcga tctcccgaac cacaccccg tctgcgactt cggcgaagtg     1320
accgatccgc cgtcagagga cgtcacgcc cactgctcg agtgtttcgg cagcttcggg     1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag     1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg     1500
gtgtcggcgc tgaaggcggc cgggatcgac acgcgcgagc cgtacattca aaaggcgctc     1560
gactgggtcg agcagcatca gaacccggac ggcggtcggg gcgaggactg ccgctcgtac     1620
gaggatccgg cgtacgcggg taagggcgcg agcaccctgt cgcagacggc ctgggcgctg     1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgcccgcg cgtgcaatac     1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc     1800
ttcccagggg atttctacct cggtacacc atgtaccgcc acgtgtttcc gacgtcgcg     1860
ctcggccgct acaagcaagc catcgagcgc aggtga                                  1896

```

<210> 13

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> SHC酶变体#215G2

<400> 13

```

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
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
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>	14	
<211>	1896	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	SHC酶变体#215G2	
<400>	14	
atggctgagc agttggtgga agctccggcc tacgcgcgga cgctggatcg cgcggtggag		60
tatctcctct cctgccaaaa ggacgaagc tactggtggg ggccgcttct gagcaacgtc		120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc		180
atggagaaga tccggcggta cctgttgac gagcagcgcg aggacggcac gtgggccctg		240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat		300

```

atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag      360
ggcgggatcg agtcgtcgcg cgtgttcacg cggaggtggc tggcgctggg gggagaatat      420
ccgtgggaga aggtgcccac ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg      480
ctcaacatct acgagtttgg ctctgtgggct cgggcgaccg tcgtggcgct ctcgattgtg      540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag      600
accgacgtgc ctccgcgccc gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg      660
ctcgaccggg tgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc      720
gagatccgcg ccttggaactg gttgctcgag cgccaggccc gagacggcag ctggggcggg      780
attcagccgc cttggtttta cgcgctcatc gcgtcaaga ttctcgacat gacgcagcat      840
ccggcggttca tcaagggtg ggaaggtcta gagctgtacg gcgtggagct ggattacgga      900
ggatggatgt ttcaggttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg      960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg     1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcggccgaa cctcaagccg     1080
ggcgggttcg cgttccagtt cgacaacgtg tactaccggg acgtggacga cacggccgtc     1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg     1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgccctac     1260
gacgtcgaca acacgagcga tctcccgaa caccacccgt 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 gaaccgggac ggcggtggtg gcgaggactg ccgctcgtac     1620
gaggatccgg cgtacgcggg taaggcgcg agcaccacgt cgcagacggc ctgggcgctg     1680
atggcgctca tcgcgggcgg cagggcgggag tccgaggccg cgcgcccgcg cgtgcaatac     1740
ctcgtggaga cgcagcgcgc ggacggcggc tgggatgagc cgtactacac cggcacgggc     1800
ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg     1860
ctcgcccgct acaagcaagc catcgagcgc aggtga                                1896

```

<210> 15

<211> 608

<212> PRT

<213> 运动发酵单胞菌 (*Zymomonas mobilis*) 野生型SHC酶

<400> 15

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
Asp Gly His Trp Val Gly Ala Val Glu Ser Asn Ala Ser Met Glu Ala		80
	85	90
Glu Trp Cys Leu Ala Leu Trp Phe Leu Gly Leu Glu Asp His Pro Leu		95
	100	105
Arg Pro Arg Leu Gly Asn Ala Leu Leu Glu Met Gln Arg Glu Asp Gly		110
	115	120
Ser Trp Gly Val Tyr Phe Gly Ala Gly Asn Gly Asp Ile Asn Ala Thr		125
	130	135
Val Glu Ala Tyr Ala Ala Leu Arg Ser Leu Gly Tyr Ser Ala Asp Asn		140
145	150	155
Pro Val Leu Lys Lys Ala Ala Ala Trp Ile Ala Glu Lys Gly Gly Leu		160
	165	170
Lys Asn Ile Arg Val Phe Thr Arg Tyr Trp Leu Ala Leu Ile Gly Glu		175
	180	185
Trp Pro Trp Glu Lys Thr Pro Asn Leu Pro Pro Glu Ile Ile Trp Phe		190
	195	200
Pro Asp Asn Phe Val Phe Ser Ile Tyr Asn Phe Ala Gln Trp Ala Arg		205
	210	215
Ala Thr Met Val Pro Ile Ala Ile Leu Ser Ala Arg Arg Pro Ser Arg		220
225	230	235
Pro Leu Arg Pro Gln Asp Arg Leu Asp Glu Leu Phe Pro Glu Gly Arg		240
	245	250
Ala Arg Phe Asp Tyr Glu Leu Pro Lys Lys Glu Gly Ile Asp Leu Trp		255
	260	265
Ser Gln Phe Phe Arg Thr Thr Asp Arg Gly Leu His Trp Val Gln Ser		270
	275	280
Asn Leu Leu Lys Arg Asn Ser Leu Arg Glu Ala Ala Ile Arg His Val		285
	290	295
Leu Glu Trp Ile Ile Arg His Gln Asp Ala Asp Gly Gly Trp Gly Gly		300
305	310	315
Ile Gln Pro Pro Trp Val Tyr Gly Leu Met Ala Leu His Gly Glu Gly		320
	325	330
Tyr Gln Leu Tyr His Pro Val Met Ala Lys Ala Leu Ser Ala Leu Asp		335
	340	345
Asp Pro Gly Trp Arg His Asp Arg Gly Glu Ser Ser Trp Ile Gln Ala		350
	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> 16			
<211> 658			
<212> PRT			
<213> 运动发酵单胞菌野生型SHC酶			
<400> 16			
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> 17
 <211> 684
 <212> PRT
 <213> 日本慢生根瘤菌 (*Bradyrhizobium japonicum*) 野生型SHC酶
 <400> 17
 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
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
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
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
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		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>	18			
<211>	642			
<212>	PRT			
<213>	细长嗜热聚球藻 (Thermosynechococcus elongatus) 野生型SHC酶			
<400>	18			
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
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								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>	19		
<211>	720		
<212>	PRT		
<213>	巴氏醋杆菌 (Acetobacter pasteurianus) 野生型SHC酶		
<400>	19		
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> 20
 <211> 685
 <212> PRT
 <213> 致病葡萄糖杆菌 (*Gluconobacter morbifer*) 野生型SHC酶
 <400> 20
 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>	21														
<211>	625														
<212>	PRT														
<213>	巨大芽孢杆菌 (Bacillus megaterium) 野生型SHC酶														
<400>	21														
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> 22
 <211> 631
 <212> PRT
 <213> SHC酶变体#49
 <400> 22
 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
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
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	

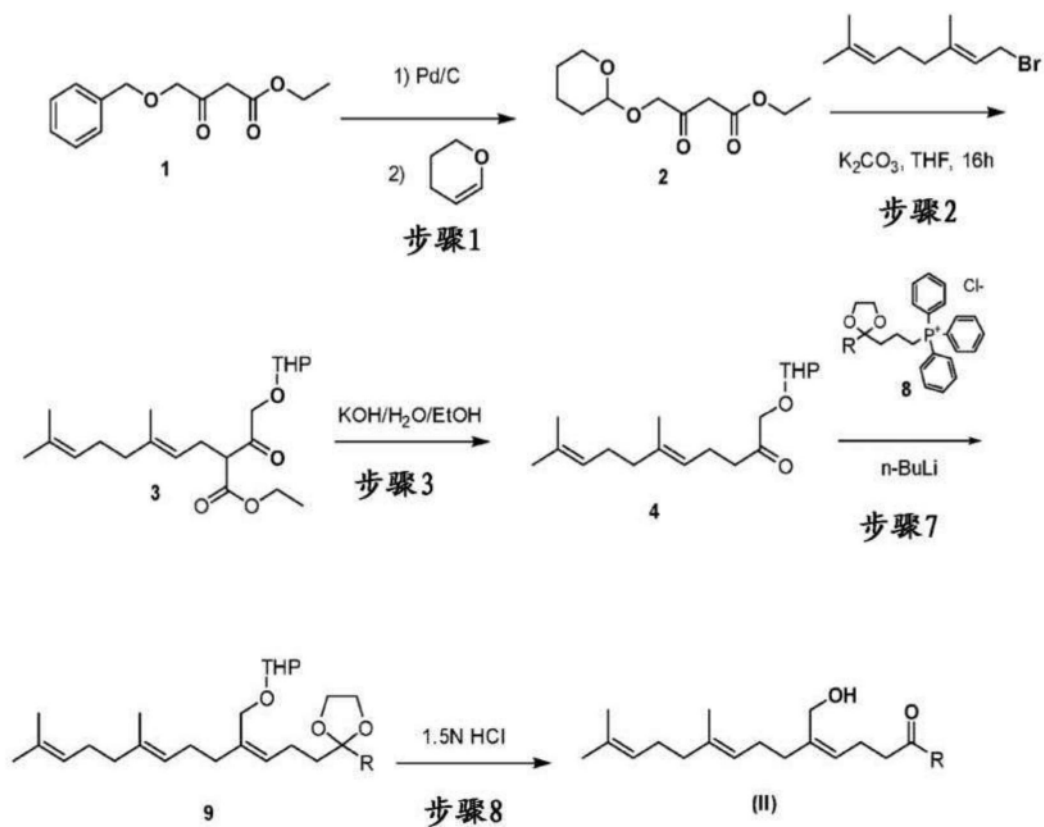


图1

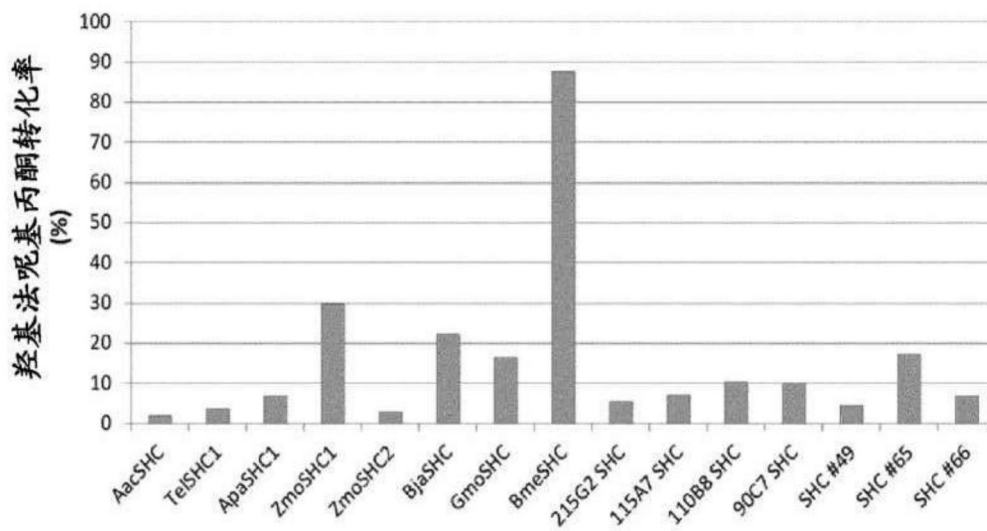


图2