



(12) 发明专利申请

(10) 申请公布号 CN 114438147 A

(43) 申请公布日 2022. 05. 06

(21) 申请号 202210270202.2

(51) Int.Cl.

(22) 申请日 2016.04.22

C12P 17/04 (2006.01)

(30) 优先权数据

G07D 307/92 (2006.01)

1507207.7 2015.04.24 GB

C12N 9/90 (2006.01)

(62) 分案原申请数据

201680023646.9 2016.04.22

(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 B·席林

D·瓦勒尔 L·富拉热

E·洛赫尔

(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 刘晓东

权利要求书2页 说明书107页

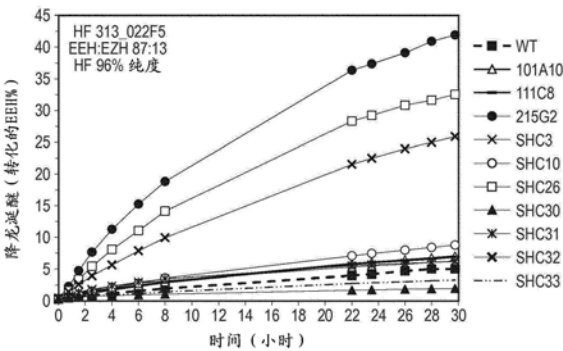
序列表322页 附图21页

(54) 发明名称

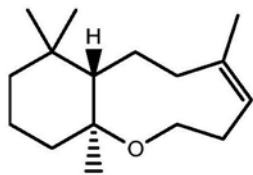
酶及其应用

(57) 摘要

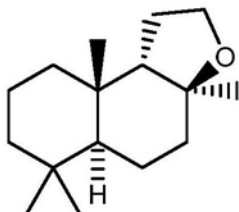
本发明提供了SHC/HAC衍生物、构成所述SHC/HAC衍生物的氨基酸序列、编码所述SHC/HAC衍生物的核苷酸序列、包含编码所述SHC/HAC衍生物的核苷酸序列的载体、包含编码所述SHC/HAC衍生物的核苷酸序列的重组宿主细胞以及包含SHC/HAC衍生物或WT SHC/HAC酶的重组宿主细胞在制备(-)-降龙涎醚和SHC/HAC酶的方法中、以及SHC/HAC酶在制备(-)-降龙涎醚的方法中的应用。



1. 一种制备 (-)-降龙涎醚的异构体混合物的方法, 其中 (3E, 7E)-高法呢醇的异构体混合物被转化为包含副产物 (II) 或 (IV) 和 (III) 中的一种或多种的 (-)-降龙涎醚异构体混合物,

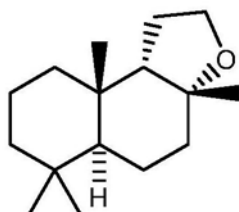


(7a*S*, 11a*S*, *Z*)-5, 8, 8, 11a-四甲基-2, 3, 6, 7, 7a, 8, 9, 10, 11, 11a-十氢苯并[*b*]oxonine (II)



(3a*S*, 5a*S*, 9a*S*, 9b*S*)-3a, 6, 6, 9a-四甲基十二氢苯并[2, 1-*b*]呋喃 (IV)

和 (III)



(3a*R*, 5a*S*, 9a*S*, 9b*S*)-3a, 6, 6, 9a-四甲基十二氢苯并[2, 1-*b*]呋喃 (III)

其中使用SHC/HAC酶的酶促转化在适于产生 (-)-降龙涎醚的反应条件下进行, 并且其中如果所述反应在存在增溶剂的情况下进行, 则不将Triton X-100与选自酸热脂环酸芽孢杆菌 (*Alicyclobacillus acidocaldarius*) SHC (AacSHC)、运动发酵单胞菌 (*Zymomonas mobilis*) (Zmo) SHC (ZmoSHC)、ZmoSHC2和慢生型大豆根瘤菌 (*Bradyrhizobium japonicum*) (Bjp) SHC (BjpSHC) 的野生型角鲨烯何帕烯环化酶/高法呢醇降龙涎香醚环化酶 (SHC/HAC) 酶结合使用, 其中高法呢醇向 (-)-降龙涎醚的转化在4-8的pH下于30℃至60℃的温度下发生。

2. 根据权利要求1所述的方法, 其中所述方法使用选自AacSHC (SEQ ID No. 1)、ZmoSHC1 (SEQ ID No. 2)、ZmoSHC2 (SEQ ID No. 3)、BjpSHC (SEQ ID No. 4) 的SHC/HAC酶或选自表1、表5、表2、表6、表3、表7、表4和/或表8的SHC/HAC衍生物或与SEQ ID No. 1、SEQ ID No. 2、SEQ ID No. 3和/或SEQ ID No. 4具有至少30%同一性、至少40%同一性、至少50%同一性、或至少60%同一性、或至少70%同一性、或至少80%同一性、或至少90%同一性、或至少95%同一性、或至少96%同一性、或至少97%同一性、或至少98%同一性、或至少99%同一性的序列进行。

3. 根据权利要求2所述的方法, 其中高法呢醇向(-)-降龙涎醚的转化在SDS增溶剂的存在下发生。

4. 根据权利要求3所述的方法, 其中使用如表24或表24a中所列的用于所述野生型SHC/HAC或每种SHC/HAC衍生生物的反应条件。

5. 根据权利要求1-4中任一项所述的方法, 其中所述方法包括(a) 在E,E-高法呢醇向(-)-降龙涎醚的转化之前, 在允许产生所述野生型SHC或SHC/HAC衍生物多肽的条件下培养表达野生型SHC或SHC衍生物酶的一种或多种重组宿主细胞。

6. 根据权利要求5所述的方法, 其中所述培养步骤和随后的转化步骤任选地在一个反应容器中在不同的反应条件下进行, 所述反应条件分别适于在细胞培养/发酵培养基中产生生物催化剂和在反应缓冲液中进行生物转化。

7. 根据权利要求6所述的方法, 其中所述培养步骤是在pH6-7的范围, 并且高法呢醇至(-)-降龙涎醚的步骤在pH4.8-5.5的范围。

8. 根据权利要求1-7中任一项所述的方法, 其中所述高法呢醇底物包含EE:EZ异构体。

9. 根据权利要求8所述的方法, 其中所述高法呢醇包含重量比选自以下的EE:EZ异构体混合物: EE:EZ 90:10、EE:EZ 80:20、EE:EZ 86:14、EE:EZ 70:30、EE:EZ 69:31和EE:EZ 66:34。

10. 根据权利要求9所述的方法, 其中所述高法呢醇包含重量比为80:20的EE:EZ。

11. 根据权利要求1-10中任一项所述的方法, 其中使用有机溶剂或汽提/蒸馏步骤或过滤回收从所述反应混合物萃取(-)-降龙涎醚。

12. 根据权利要求11所述的方法, 其中使用有机溶剂从所述反应混合物分离(-)-降龙涎醚。

13. 根据权利要求12所述的方法, 其中使用有机溶剂使所述(-)-降龙涎醚选择性结晶。

14. 根据权利要求12或13所述的方法, 其中所述(-)-降龙涎醚基本上不含副产物(II)、(IV)和/或(III)。

15. 通过根据权利要求11-14中任一项所述的方法能够获得的包含(-)-降龙涎醚的反应产物。

16. 根据权利要求15所述的反应产物, 其为固体形式。

17. 根据权利要求16所述的反应产物, 其为无定形式或晶体形式。

18. 一种制备含有(-)-降龙涎醚的产品的的方法, 包括将根据权利要求16-18中任一项所述的反应产物掺入所述产品中。

19. 根据权利要求18所述的方法, 其中所述产品是香料产品、化妆品、清洁产品、洗涤剂产品或皂产品。

20. 一种香料或化妆品或消费者护理产品, 其中包含根据权利要求15-17中任一项所述的反应产物。

21. 一种香料或化妆品或消费者护理组合物, 其中包含根据权利要求15-17中任一项所述的反应产物和另外的组分。

22. 根据权利要求15-17中任一项所述的反应产物的用途, 用作香料或化妆品或消费者护理产品的一部分。

酶及其应用

[0001] 本申请是申请日为2016年4月22日、发明名称为“酶及其应用”的中国发明专利申请No.201680023646.9的分案申请。

技术领域

[0002] 本发明涉及相对于参照SHC/HAC蛋白经过了修饰的角鲨烯何帕烯环化酶/高法呢醇降龙涎香醚环化酶 (Squalene Hopene Cyclase/Homofarnesol Ambrox Cyclase, SHC/HAC) 衍生酶、构成该SHC/HAC衍生酶的氨基酸序列、编码该SHC/HAC衍生生物的核苷酸序列、包含编码该SHC/HAC衍生生物的核苷酸序列的载体以及包含编码该SHC/HAC衍生生物的核苷酸序列的重组宿主细胞。本发明还涉及功能性表达编码SHC/HAC衍生生物的核苷酸序列的手段以及利用包含编码SHC/HAC衍生物和野生型SHC/HAC的核苷酸序列的重组微生物来制备降龙涎醚 (Ambrox)、优选(-)-降龙涎醚的方法。

背景技术

[0003] 角鲨烯何帕烯环化酶 (Squalene Hopene Cyclases, SHC, EC5.4.99.17) 是膜结合型原核生物酶,其充当使线性三萜角鲨烯环化为何帕烯和何帕醇的生物催化剂。较早的SHC工作集中于嗜热嗜酸性细菌——酸热脂环酸芽孢杆菌 (*Alicyclobacillus acidocaldarius*)) (以前称为酸热芽孢杆菌 (*Bacillus acidocaldarius*)) 的SHC的表征 (参见Neumann&Simon 1986, *Biol Chem Hoppe-Seyler* 367,723-729; Seckler&Poralla 1986, *Biochem Biophys Act* 356-363以及Ochs等人1990, *J Bacteriol* 174,298-302)。然而,更近些时候,已经对来自运动发酵单胞菌 (*Zymomonas mobilis*) 和慢生型大豆根瘤菌 (*Bradyrhizobium japonicum*) 的其它SHC进行了纯化并就它们的天然底物 (如角鲨烯) 和非天然底物 (如高法呢醇和柠檬醛) 进行了表征 (参见例如,WO 2010/139710、WO 2012/066059以及Seitz等人2012, *J. Molecular Catalysis B:Enzymatic* 84,72-77)。

[0004] Neumann和Simon进行的较早工作 (1986-如上所述) 公开了高法呢醇是酸热脂环酸芽孢杆菌SHC (AacSHC) 的另一种底物。然而,Neumann和Simon (1986) 教导的纯化AacSHC的非天然底物高法呢醇的环化速率据报道仅为天然底物角鲨烯的环化速率的3%。在0.25mM至2.0mM的高法呢醇 (产物1b) 浓度下降龙涎醚 (产物2b) 的形成速率增加,在存在4mM产物1b的情况下稍微下降。环化速率的差别可能部分归因于天然SHC底物角鲨烯 (C30碳化合物) 的大小是非天然底物高法呢醇 (其为C16碳化合物) 的两倍这样一个事实。

[0005] (JP2009060799-花王株式会社 (Kao)) 也公开了使用来自酸热脂环酸芽孢杆菌的SHC由高法呢醇制备降龙涎醚的方法。而JP2009060799教导了使用包含SHC的微生物来合成降龙涎醚的可能性,JP2009060799仅公开了使用从表达SHC基因的重组微生物制备的SHC液体萃取物而不是通过表达SHC基因的整个重组微生物细胞的手段由高法呢醇制备降龙涎醚。使用SHC液体萃取物时高法呢醇向降龙涎醚转化的百分比据报道当在pH 5.2-6.0下于60℃的温度下进行14小时时为17.5%,但在6.6的pH下进行仅为6.8%。当使用0.2%高法呢醇 (2g/l) 底物浓度时,利用SHC液体萃取物在pH 5.6下于60℃进行64小时,3E,7E-高法呢

醇向降龙涎醚转化的百分比据报道为63%。

[0006] WO 2010/139719A2及其美国同族(US2012/0135477A1)描述了具有高法呢醇-降龙涎醚环化酶活性的至少三种SHC酶萃取物。当使用10mM (2.36g/l) 高法呢醇浓度时,运动发酵单胞菌(Zmo) SHC和慢生型大豆根瘤菌(Bjp) SHC酶据报道在16小时的反应中分别显示出41%和22%的高法呢醇转化率,而对于AacSHC,其转化率据报道仅为1.2%(假定以相同的高法呢醇浓度),但未提供实验细节。ZmoSHC和BjpSHC酶萃取物是通过使产生SHC酶的大肠杆菌宿主细胞破裂并分离可溶性的SHC级分而由表达SHC基因的重组微生物而制备的。

[0007] Seitz等人(2012-如上所述)报道了三种SHC酶的功能性表达和生化表征,其中两种来自运动发酵单胞菌(ZmoSHC1和ZmoSHC2),一种来自慢生型大豆根瘤菌。据报道,当使用10mM (2.36g/l) 的高法呢醇浓度时,使用野生型ZmoSHC1时观察到高法呢醇向降龙涎醚的“有效”转化(22.95%),使用野生型ZmoSHC2时没有高法呢醇向降龙涎醚的转化,而对于AacSHC,发现高法呢醇向降龙涎醚的转化相对较低(3.4%)。对于AacSHC观察到的高法呢醇向降龙涎醚转化的转化率相对较低的趋势与Neumann和Simon(1986-如上文所引用的)的结果一致,并且如WO 2010/139719A2(也在上面论述)中所公开的。这三种SHC酶以细胞悬浮液形式(通过采用冻-融循环使宿主大肠杆菌细胞部分破裂)并且作为部分纯化的膜结合型级分使用。

[0008] WO2012/066059公开了具有环化酶活性的突变体以及其在以下方法中的用途:用于萜烯的生物催化环化,例如,具体的讲,用于通过使香茅醛环化而制备异蒲勒醇的方法;用于制备薄荷醇的方法以及用于具有萜烯类结构基序的其它化合物的生物催化转化的方法。对各种SHC进行序列比对将苯丙氨酸-486(F486)鉴别为强保守性氨基酸残基,并在运动发酵单胞菌SHC酶中产生了一系列取代变体。这些取代中的一些导致了活性丧失,而其它则导致了从萜烯底物诸如香茅醛形成新的萜类产物(异蒲勒醇)。

[0009] 2012年Seitz的博士论文(<http://elib.uni-stuttgart.de/handle/11682/1400>)中的报道表明,ZmoSHC1中的F486Y突变使高法呢醇生物转化率降低约1.5倍,从34.8%(野生型ZmoSHC1)降低至23.9%(突变体ZmoSHC1 F486Y)。当测试AacSHC中的突变等价物(Y420C)时,推测对较大底物的酶活性将降低,而对较小底物的活性将增加。当在与野生型相同的条件下测试该突变体并比较酶活性时,观察到该突变体根本未显示出高法呢醇底物的任何转化。因此其结论是Y420氨基酸残基对于AacSHC对所有底物的活性是关键的。

[0010] 本领域的其它SHC定点诱变(如Hoshino和Sato 2002,Chem Commun 291-301)聚焦于高度保守区中的突变(如F601)的效果以及它们对天然底物(即角鲨烯或角鲨烯类似物)而不是对非天然底物诸如高法呢醇的效果。

[0011] 概括地说,本领域中涉及用于将高法呢醇成功转化为降龙涎醚的生物转化工艺的有限公开内容仅涉及相对较低浓度/体积的高法呢醇底物(以0.25mM至2mM到10mM或者约0.06g/l至2.36g/l的浓度范围),使用具有高法呢醇-降龙涎醚环化酶(HAC)活性的野生型SHC多肽。具有HAC活性的SHC酶是:(i)通过破坏包含SHC酶的大肠杆菌宿主细胞并分离可溶和不可溶的SHC液体级分而制备的液体萃取物;(ii)部分纯化的膜级分;或者(iii)表达野生型SHC基因并产生SHC酶的重组全细胞,该SHC酶供用于反应中利用增溶剂将高法呢醇生物转化为降龙涎醚,所述增溶剂包括:(i)反应混合物中的Triton X-100(参见如上文所引用的Neumann和Simon 1986,如上文所引用的Seitz等人2012,JP2009060799);或(ii)牛黄

脱氧胆酸盐(如US2012/0135477A1中所公开的)。

[0012] 利用这些野生型SHC萃取物和/或表达SHC基因的整个重组微生物细胞,据发现所获得的高法呢醇向降龙涎醚转化的转化率依SHC酶的来源、高法呢醇原料的量和所采用的反应条件而变化。迄今为止,在所报道的所涉及浓度(0.06-2.36g/l)下,尚未实现利用野生型SHC酶将高法呢醇100%转化为降龙涎醚。此外,采用利用定点诱变实验制备的SHC衍生物进行的初步研究仅提供了阴性的结果(即降低的高法呢醇转化率)而不是阳性的结果(即提高的转化率)。此外,在公开的研究中仅使用纯化的SHC酶萃取物或SHC膜结合级分,或者在使用增溶剂诸如Triton X-100或牛黄脱氧胆酸盐的特定反应条件下使用表达野生型SHC基因的完整重组微生物细胞。尚未证明包含野生型SHC或突变体SHC的重组微生物可在使用最佳反应条件下提供更有效的或高性价比的高法呢醇向降龙涎醚的生物转化。因此,期望通过至少改善反应速率、特异性、产率、生产力以及降低费用(通过例如,利用重组全微生物细胞或通过采用将生物催化剂生产步骤和生物转化步骤合并的“一锅法”工艺来简化工艺)来改善所引用的由高法呢醇制备降龙涎醚的已知方法。

发明内容

[0013] 本发明在多个方面提供SHC/HAC衍生物、构成所述SHC/HAC衍生酶的氨基酸序列、编码SHC/HAC衍生酶的核苷酸序列、包含编码SHC/HAC衍生酶的核苷酸序列的载体、包含具有编码SHC/HAC衍生酶的核苷酸序列的载体的重组宿主细胞以及包含SHC/HAC衍生酶或野生型SHC/HAC酶的重组宿主细胞当在特定反应条件下使用时在制备降龙涎醚材料的方法中的应用,所述降龙涎醚材料包含命名为(-)-降龙涎醚的降龙涎醚异构体和降龙涎醚样分子(作为副产物)。与本领域中涉及AacSHC的公开内容不同,本申请人第一次证明了,表达SHC衍生物基因的整个重组微生物可用于将高法呢醇生物转化为降龙涎醚。此外,表达野生型SHC基因和/或产生SHC酶的完整重组微生物可用于在本领域未公开的特定反应条件下将高法呢醇生物转化为降龙涎醚。

[0014] 还已出乎意料地发现,在本文所公开的野生型SHC/HAC参照序列的氨基酸序列中引入最多五个氨基酸改变得得到这样的SHC/HAC衍生酶,该衍生酶与本文所公开的未经修饰的SHC参照酶相比,具有显著改善的高法呢醇向降龙涎醚转化的转化率。这些新型的SHC/HAC衍生酶单独时以及组合时可用于由高法呢醇底物制备降龙涎醚材料,尤其是(-)-降龙涎醚。

[0015] 另一个出乎意料的发现是,除了一种突变体(F601Y)之外,本文所公开的SHC衍生酶通常在参照SHC多肽序列的非保守部分中的氨基酸残基位置处包含非保守取代。这是出乎意料的发现,因为酶保守区中的改变比蛋白质非保守区中的改改变有可能破坏酶的功能(至少就其天然底物而言)。

[0016] 另一出乎意料的发现是,本公开的经表征的SHC衍生酶在约35℃而不是约60℃下表现最佳(对非天然底物诸如高法呢醇),60℃是嗜热微生物诸如AacSHC的通常的反应温度。本公开的SHC衍生物在较低反应温度下应用于由高法呢醇制备降龙涎醚的方法中对工业规模的降龙涎醚生产周期具有显著的成本优势。

[0017] 本发明的另一个优势是,本公开的SHC衍生酶催化有效的生物转化过程,该生物转化过程在用与本领域先前描述的浓度(如125g/l的EEH)比较相对较高(如,约50倍)的高法

呢醇底物浓度优化时可导致高法呢醇底物100%转化,而参照WT SHC蛋白仅转化约10%的相同底物,即使是在高的酶/细胞浓度下时也是如此。所引用的现有技术中的公开内容全部涉及使用包含SHC的纯化膜萃取物或纯化SHC萃取物(从表达SHC基因的微生物制备)或在特定的生物转化反应条件下(如使用特定的增溶剂)使用表达野生型SHC基因的重组微生物。即使那样,在低很多的EEH浓度下也未见报道100%的高法呢醇转化率。另外,尚未报道“一锅法”反应,其中在第一步骤中,在同一容器中使重组细胞生长并产生SHC酶,随后将EEH转化为(-)-降龙涎醚。本发明的另一个优点是,产生SHC衍生酶的重组宿主细胞显示出高的初始反应速率,这使得能在仅使用相对较低量的生物催化剂的同时在相对较短的时间周期内产生大量的产物。简而言之,包含野生型SHC/HAC或特定SHC/HAC衍生酶的重组微生物在特定生物转化反应条件下的选择和有效表达以及应用导致更有效的生物转化过程。可分离最终产物((-)-降龙涎醚)并容易地纯化。不像所引用的现有技术,所述SHC/HAC衍生酶不作为纯的酶使用而是在完全细胞情形下使用(作为生物催化剂),这是更加节省成本且更加用户友好和环境友好的方法,因为不需要另外的酶纯化和分离步骤。

[0018] 概括地说,本公开提供了用于在重组微生物菌株中制备降龙涎醚的生物转化/生物转换方法,其中该方法:(i)在经济上有吸引力,(ii)环境友好,以及(iii)导致选择性地产生(-)-降龙涎醚作为优势化合物,该化合物在选择性结晶条件下有效地与其它副产物分离,这些副产物不会促成最终产物的嗅觉品质。

具体实施方式

[0019] 如本文所用,术语“SHC”意指来自表10-12中所列来源任一者的角鲨烯何帕烯环化酶。在优选的实施方案中,术语SHC包括如BASF WO 2010/139719、US2012/01345477A1、Seitz等人(2012,如上文所引用的)和Seitz(2012博士论文,如上文所引用的)中所公开的运动发酵单胞菌SHC酶和酸热脂环酸芽孢杆菌SHC酶。为了便于参考,名称“AacSHC”用于酸热脂环酸芽孢杆菌SHC,而名称“ZmoSHC”用于运动发酵单胞菌SHC,名称“BjpSHC”用于慢生型大豆根瘤菌SHC。这些序列相对于野生型AacSHC以及彼此之间的序列同一性百分比(可根据所用的算法而变动)在表18和表19中列出。

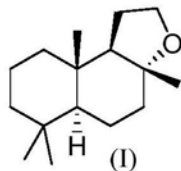
[0020] 由Hoshino和Sato(2002,如上文所引用的)进行的野生型SHC序列的比对表明,在全部四种序列中检测到多个基序并且这些基序由核心序列Gln-X-X-X-Gly-X-Trp组成,该核心序列在运动发酵单胞菌和酸热脂环酸芽孢杆菌二者的SHC序列中出现六次(参见Reipen等人1995, Microbiology 141, 155-161的图3)。Hoshino和Sato(2002,如上文所引用的)报道说,SHC中通常富含芳族氨基酸并且在SHC中发现两个特征性的基序:一个为由特定氨基酸基序[(K/R)(G/A)X2-3(F/Y/W)(L/IV)3X3QX2-5GXW]代表的QW基序,另一个为DXDDTA基序。Wendt等人(1997, Science 277, 1811-1815以及1999, J Mol Biol 286, 175-187)报道了对酸热脂环酸芽孢杆菌SHC的X射线结构分析。DXDDTA基序看起来与SHC活性位点相关。来自现有技术的示例性序列比对显示了反复出现的多个基序,如本文中的图2(来自Hoshino和Sato(2002,如上文所引用的))和图3(来自Seitz博士论文(2012))中所提供的。

[0021] 本文所用的参照(或野生型)AacSHC蛋白是指SEQ ID No.1中所公开的AacSHC蛋白。本公开的参照AacSHC酶具有高法呢醇降龙涎香醚环化酶(HAC)的活性,该活性可用于通过SHC与高法呢醇底物的生物催化反应制备降龙涎醚衍生物。参照AacSHC的主要反应是使

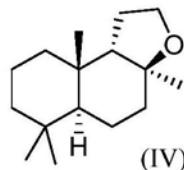
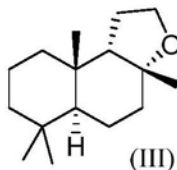
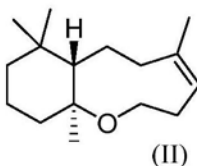
线性或非线性底物诸如高法呢醇环化而产生降龙涎醚。

[0022] 降龙涎醚

[0023] 如本文所用,术语“降龙涎醚”包括式(I)的(-)-降龙涎醚以及立体异构体纯形式的或与下面式(II)、(IV)和/或(III)的分子中的至少一种或多种的混合物形式的(-)-降龙涎醚。



[0024]



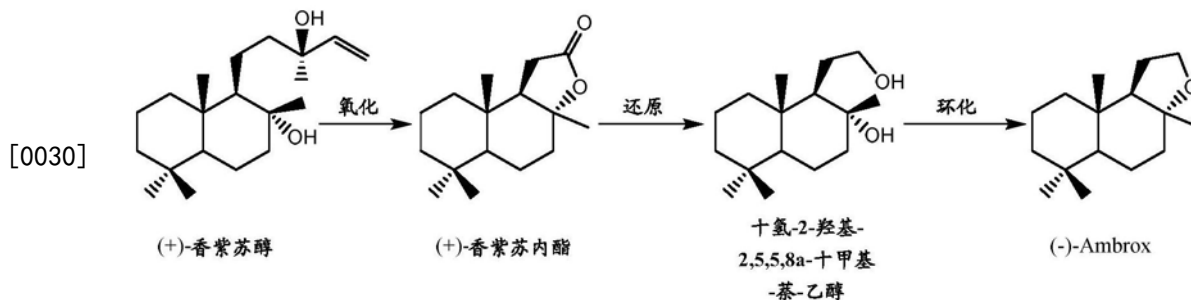
[0025] (-)-降龙涎醚

[0026] (-)-降龙涎醚在商业上称为降龙涎醚(芬美意公司(Firmenich))、降龙涎醚an(汉高公司(Henkel))、Ambrofix(奇华顿公司(Givaudan))、Amberlyn(奎斯特公司(Quest))、CetaloxLaevo(芬美意公司)、Ambermor(埃姆公司(Aromor))和/或Norambrenolide Ether(太平洋公司(Pacific))。

[0027] (-)-降龙涎醚是一种在工业上重要的芳香化合物,已被用于香料行业很久了。(-)-降龙涎醚的特殊受欢迎的感官有益效果来自该(-)立体异构体而不是(+)立体异构体。(-)立体异构体的气味被描述为类似麝香、木香、温暖或琥珀味,而(+) -降龙涎醚对映体具有相对较弱的气味。降龙涎醚样产品的气味和气味阈值也是不同的。虽然各种富含(-)-降龙涎醚的材料可商购获得,但希望制备高度富含(-)-降龙涎醚的材料,理想的是纯的(-)-降龙涎醚。

[0028] (-)-降龙涎醚的制备

[0029] (-)-降龙涎醚可根据下面所描述的制备工艺由香紫苏内酯制备。香紫苏醇是从天然植物香紫苏萃取的产物。然而,由于该工艺中使用天然原料,所以存在潜在的问题,问题在于其涉及多级反应,其操作是迂回的,原料供应的质量及稳定性可能并不会总令人满意,并且该反应可能不是环境友好的,因为在(+)-香紫苏醇氧化降解步骤中使用氧化剂诸如铬酸或高锰酸盐。



[0031] (-)-降龙涎醚还使用不同的路线由高法呢醇合成。举个例子,可通过溴化、氰化以及水解橙花叔醇而得到高法呢酸,然后还原而获得高法呢醇。作为另一种选择,可由法呢

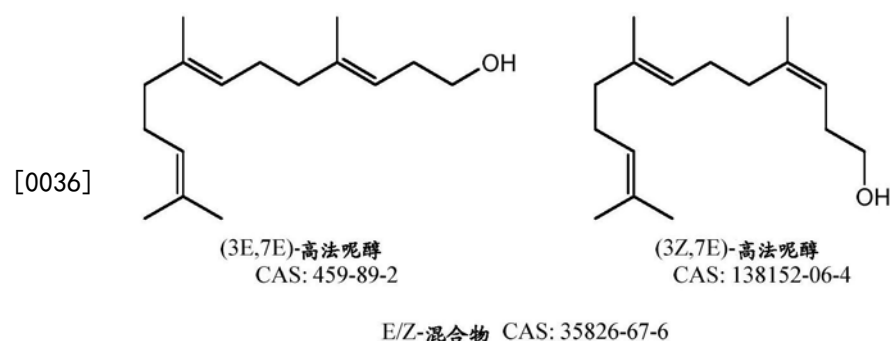
醇、氯化法呢酯、 β -法呢烯或其它底物获得高法呢醇。 β -法呢烯可直接转换为E,E-高法呢醇(EEH)或经由E,E-高法呢酯(其然后转换为EEH)间接转换为EEH。关于由不同底物制备(-)-降龙涎醚的综述可见于US2012/0135477A1、WO 2010/139719、US2013.0273619A1、WO 2013/156398A1和Seitz博士论文(2012,如上文所引用的)以及Schaefer 2011(Chemie Unserer Zeit 45,374-388)。

[0032] 虽然高法呢醇可能作为四种异构体((3Z,7Z)、(3E,7Z)、(3Z,7E)和(3E,7E)异构体)的混合物存在,但根据文献似乎(-)-降龙涎醚仅从(3E,7E)高法呢醇获得(参见Neumann和Simon(1986),如上文所引用的)。如本文所用,提及(3E,7E)高法呢醇即提及E,E-高法呢醇,其也命名为EEH。

[0033] US2012/0135477A1报道了使用ZmoSHC(SEQ ID No.2)使(3Z,7E)转化为(-)-降龙涎醚(参见实施例2-4),但是根据Schaefer(2011)(如上文所引用的)中的公开内容,(7E,3Z)仅转化为如上文所概述的9b-表-降龙涎醚(即化合物III)而不转化为(-)-降龙涎醚。如本文所用,提及(3Z,7E)高法呢醇即提及E,Z-高法呢醇,其也命名为EZH。

[0034] 在一些实施方案中,优选该高法呢醇原料包含(3E,7E)和(3Z,7E)的混合物,该混合物在本文称为EE:EZ立体异构体混合物(具体参见实施例和表20)。

[0035] 高法呢醇的EE:EZ立体异构体混合物的CAS编号为35826-67-6。



[0037] 如实施例所展示的(如参见实施例5、7、9、10、11、18、19和20),在某些实施方案中,该高法呢醇给料/原料是异构体的混合物。

[0038] 因此,在一些实施方案中,该高法呢醇原料还可包含四种异构体EE:EZ:ZZ:ZE的混合物,这四种异构体对应于(3E,7E)和(3Z,7E)、(3Z,7Z)和(3E,7Z)。

[0039] 在一些实施方案中,优选该高法呢醇原料选自以下组中的一者或多者:[(3Z,7Z)、(3E,7Z)、(3Z,7E)和(3E,7E)]、[(3Z,7E)和(3E,7E)]、[(3Z,7E)、(3E,7Z)]和/或[(3E,7E)和(3E,7Z)]。

[0040] 优选地,该高法呢醇原料选自以下组中的一者或多者:[(3E,7E)、(3Z,7E)]和/或[(3Z,7E)、(3E,7E)和(3E,7Z)]。

[0041] 因此,在某些实施方案中,EEH:EZH的比率为约100:00;99:01;98:02;97:03;96:04;95:05;94:06;93:07;92:08;91:09;90:10;89:11;88:12;87:13;86:14;85:15;84:16;83:17;82:18;81:19;80:20;79:21;78:22;77:23;76:24;75:25;74:26;73:27;72:28;71:29;70:30;69:31;68:32;67:33;66:34;65:35;64:36;63:37;62:38;61:39;60:40;59:41;58:42;57:43;56:44;55:45;54:46;53:47;52:48;51:49;或约50:50。

[0042] 在一些实施方案中,优选该高法呢醇原料包含>90%的E,E-高法呢醇(EEH)。

[0043] 在其它实施方案中,该高法呢醇原料包含86:14重量比的EE:EZ。

[0044] 在某些实施方案中,该高法呢醇原料包含80:20重量比的EE:EZ。

[0045] 在某些实施方案中,该高法呢醇原料包含70:30重量比的EE:EZ。

[0046] 在另外的实施方案中,该高法呢醇原料包含69:31重量比的EE:EZ。

[0047] 在一些实施方案中,该高法呢醇原料由四种异构体EE:EZ:ZZ:ZE的混合物组成或基本上由所述混合物组成,这四种异构体对应于(3E,7E)和(3Z,7E)、(3Z,7Z)和(3E,7Z)。

[0048] 在一些实施方案中,优选该高法呢醇原料由选自以下组中的一者或多者的异构体混合物组成或基本上由所述混合物组成:[(3Z,7Z)、(3E,7Z)、(3Z,7E)和(3E,7E)]、[(3Z,7E)和(3E,7E)]、[(3Z,7E)、(3E,7Z)]和/或[(3E,7E)和(3E,7Z)]。

[0049] 优选地,该高法呢醇原料由选自以下组中的一者或多者的异构体混合物组成或基本上由其组成:[(3E,7E)、(3Z,7E)]和/或[(3Z,7E)、(3E,7E)和(3E,7Z)]。

[0050] 因此,在某些实施方案中,EEH:EZH异构体的比率由以下EEH:EZH比率组成或基本上由以下EEH:EZH比率组成:约100:00;99:01;98:02;97:03;96:04;95:05;94:06;93:07;92:08;91:09;90:10;89:11;88:12;87:13;86:14;85:15;84:16;83:17;82:18;81:19;80:20;79:21;78:22;77:23;76:24;75:25;74:26;73:27;72:28;71:29;70:30;69:31;68:32;67:33;66:34;65:35;64:36;63:37;62:38;61:39;60:40;59:41;58:42;57:43;56:44;55:45;54:46;53:47;52:48;51:49;或约50:50。

[0051] 在一些实施方案中,优选高法呢醇原料由>90%的E,E-高法呢醇(EEH)组成或基本上由其组成。

[0052] 在其它实施方案中,该高法呢醇原料由86:14重量比的EE:EZ组成或基本上由其组成。

[0053] 在某些实施方案中,该高法呢醇原料由80:20重量比的EE:EZ组成或基本上由其组成。

[0054] 在某些实施方案中,该高法呢醇原料由70:30重量比的EE:EZ组成或基本上由其组成。

[0055] 在另外的实施方案中,该高法呢醇原料由69:31重量比的EE:EZ组成或基本上由其组成。

[0056] 在本公开的实施方案中,使用SHC/HAC衍生酶制备降龙涎醚。

[0057] SHC/HAC衍生物

[0058] 如本文所用,术语“SHC/HAC衍生物”意指SHC/HAC衍生物的氨基酸序列是经修饰的氨基酸序列或变体氨基酸序列,该序列与根据至少SEQ ID No.1或SEQ ID No.2或SEQ ID No.3或SEQ ID No.4的参照(或野生型)SHC序列的氨基酸序列相比发生改变。一般而言,SHC/HAC衍生物包含具有至少一个改变的SHC改变形式,该改变修饰(如增加)该酶对其底物(如EEH)的活性。

[0059] 对本公开的SHC/HAC衍生物测试它们的高法呢醇降龙涎香醚环化酶活性。因此,使高法呢醇转化为降龙涎醚的这些SHC/HAC衍生物在本文称为HAC衍生物以及SHC衍生物。虽然对源于酸热脂环酸芽孢杆菌、运动发酵单胞菌、慢生型大豆根瘤菌微生物菌株来源的酶已提供了示例性的SHC/HAC衍生物,本公开还涵盖来自其它微生物菌株来源的等同SHC/HAC衍生物,所述来自其它微生物菌株来源的等同SHC/HAC衍生物包括但不限于来自荚膜甲基球菌(*Methylococcus capsulatus*)、桉木弗兰克氏菌(*Frankia alni*)、巴氏醋酸杆菌

(*Acetobacter pasteurianum*) 和梨形四膜虫 (*Tetrahymena pyriformis*) 的SHC/HAC酶 (参见例如, WO 2010/139719、US2012/01345477、WO 2012/066059以及表10-12)。

[0060] 如本文所用, 术语“氨基酸改变”意指相对于参照氨基酸序列 (例如, SEQ ID No.1 或SEQ ID No.2或SEQ ID No.3或SEQ ID No.4的野生型 (WT) 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个或多个不同的氨基酸取代 (其可以是保守取代或非保守取代) 一个或多个氨基酸。可通过将SHC/HAC衍生物氨基酸序列的氨基酸序列与参照氨基酸序列 (例如, SEQ ID No.1或SEQ ID No.2或SEQ ID No.3或SEQ ID No.4的野生型 (WT) 氨基酸序列) 的氨基酸序列相比较而容易地鉴别氨基酸改变。示例性的WT SHC氨基酸序列比对在图1-4以及表18和表19中提供。

[0061] 可例如基于所涉及的氨基酸残基的极性、电荷、大小、溶解性、疏水性、亲水性和/或两亲性性质的相似性作出保守性氨基酸取代。上面概述的20种天然存在的氨基酸可分成以下六个标准氨基酸组:

[0062] (1) 疏水性氨基酸: Met、Ala、Val、Leu、Ile;

[0063] (2) 中性亲水性氨基酸: Cys、Ser、Thr、Asn、Gln;

[0064] (3) 酸性氨基酸: Asp、Glu;

[0065] (4) 碱性氨基酸: His、Lys、Arg;

[0066] (5) 影响链取向的残基: Gly、Pro; 以及

[0067] (6) 芳族氨基酸: Trp、Tyr、Phe。

[0068] 因此, 如本文所用, 术语“保守取代”意指氨基酸替换成上面所示的六个标准氨基酸组的同一组内所列的另一氨基酸。例如, Asp替换为Glu可在经过这样修饰的多肽中保留负电荷。此外, 甘氨酸和脯氨酸根据它们破坏 α -螺旋的能力而可以互相取代。上面六个组内的某些优选的保守取代是以下亚组内的替换: (i) Ala、Val、Leu和Ile; (ii) Ser和Thr; (iii) Asn和Gln; (iv) Lys和Arg; 以及 (v) Tyr和Phe。鉴于已知的遗传密码以及重组和合成DNA技术, 熟练的科学家可容易构建编码保守氨基酸变体的DNA。

[0069] 如本文所用, “非保守取代”或“非保守氨基酸替换”定义为氨基酸替换为如上所示的六个标准氨基酸组 (1) 至 (6) 的不同组中所列的另一氨基酸。通常本公开的SHC/HAC衍生物利用非保守取代制备, 所述非保守取代使所公开的SHC/HAC衍生物的生物功能 (如HAC活性) 发生改变。

[0070] 为了易于参照, IUPAC-IUB生物化学命名委员会推荐的单字母氨基酸符号给出如下。还提供了三字母代码以供参照。

[0071]

单字母代码	三字母代码	氨基酸名称
A	Ala	丙氨酸
C	Cys	半胱氨酸
D	Asp	天冬氨酸
E	Glu	谷氨酸
F	Phe	苯丙氨酸
G	Gly	甘氨酸
H	His	组氨酸
I	Ile	异亮氨酸
K	Lys	赖氨酸
L	Leu	亮氨酸
M	Met	甲硫氨酸

[0072]

N	Asn	天冬酰胺
P	Pro	脯氨酸
Q	Gln	谷氨酰胺
R	Arg	精氨酸
S	Ser	丝氨酸
T	Thr	苏氨酸
V	Val	缬氨酸
W	Trp	色氨酸
Y	Tyr	酪氨酸

[0073] 氨基酸改变诸如氨基酸取代可利用已知的重组基因技术方案引入,所述重组基因技术包括PCR、基因克隆、cDNA的定点诱变、宿主细胞的转染以及体外转录,其可用于将这类改变引入WT SHC序列中从而得到SHC/HAC衍生酶。然后可针对SHC/HAC功能活性对衍生物进行筛选。

[0074] SHC/HAC衍生酶

[0075] 本发明提供SHC/HAC衍生物并且描述具有高法呢醇降龙涎香醚环化酶(HAC)活性的酶,其包含的氨基酸序列相对于根据至少SEQ ID No.1或SEQ ID No.2或SEQ ID No.3或SEQ ID No.4的参照(或野生型)SHC序列的氨基酸序列具有约1至约50个独立选自取代、缺失或插入的突变。

[0076] 在多个实施方案中,与未显示这种缺失/添加的参照SHC酶相比,突变或突变组合增强了所述SHC/HAC衍生物将高法呢醇转换为降龙涎醚的活性。如本文所述的蛋白质建模可用于引导在SHC参照序列中的这种取代、缺失或插入。例如,可利用AacSHC的坐标生成SHC氨基酸序列的结构模型(例如如在图19和20中所示)。如本文所展示的,这种同源建模可用于指导SHC酶将高法呢醇转化为(-)-降龙涎醚的改善。

[0077] 因而,在多个实施方案中,相对于根据至少SEQ ID No.1或SEQ ID No.2或SEQ ID No.3或SEQ ID No.4的参照(或野生型)SHC序列的氨基酸序列,SHC/HAC衍生物可具有约1至约45个突变、约1至约40个突变、约1至约35个突变、约1至约30个突变、约1至约25个突变、约1至约20个突变、约1至约15个突变、约1至约10个突变、或约1至约5个突变。

[0078] 在多个实施方案中,相对于根据至少SEQ ID No.1或SEQ ID No.2或SEQ ID No.3或SEQ ID No.4的参照(或野生型)SHC序列的氨基酸序列,SHC/HAC衍生物包含具有至少5或至少10个突变但不超过约20或30个突变的序列。在多个实施方案中,相对于参照SHC(例如,SEQ ID No.1或2或3或4),SHC衍生物可具有约1个突变、约2个突变、约3个突变、约4个突变、约5个突变、约6个突变、约7个突变、约8个突变、约9个突变、约10个突变、约11个突变、约12个突变、约13个突变、约14个突变、约15个突变、约16个突变、约17个突变、约18个突变、约19个突变、约20个突变、约21个突变、约22个突变、约23个突变、约24个突变、约25个突变、约26个突变、约27个突变、约28个突变、约29个突变、约30个突变、约31个突变、约32个突变、约33个突变、约34个突变、约35个突变、约36个突变、约37个突变、约38个突变、约39个突变、约40个突变、约41个突变、约42个突变、约43个突变、约44个突变、约45个突变、约46个突变、约47个突变、约48个突变、约49个突变、或约50个突变。

[0079] 在这些或其它实施方案中,SHC/HAC衍生物可包含这样的氨基酸序列,该序列与WT SHC(例如,SEQ ID No.1或SEQ ID No.2或SEQ ID No.3或SEQ ID No.4)或者在参照序列(参见例如表18和表19,其中示出了AacSHC(SEQ ID No.1)与其它SHC序列(如W0 2010/139719的ZmoSHC)之间有至少34-52%的同一性)具有至少约50%序列同一性、至少约55%序列同一性、至少约60%序列同一性、至少约65%序列同一性、至少约70%序列同一性、至少约75%序列同一性、至少约80%序列同一性、至少约85%序列同一性、或至少90%序列同一性、或至少91%序列同一性、或至少92%序列同一性、或至少93%序列同一性、或至少94%序列同一性、或至少95%序列同一性、或至少96%序列同一性、或至少97%序列同一性、或至少98%序列同一性、或至少99%序列同一性。

[0080] 在多个实施方案中,与野生型酶相比SHC变体具有较高的将高法呢醇转换为降龙涎醚的活性,诸如与参照野生型酶(例如,SEQ ID No.1或SEQ ID No.2或SEQ ID No.3或SEQ ID No.4)相比,在与高法呢醇底物接触后(-)-降龙涎醚的产量更高。

[0081] 例如,SHC/HAC衍生物可包含这样的氨基酸序列,该氨基酸序列与参照SHC(例如,SEQ ID No.1或2或3或4)或者在参照序列之间(参见例如表18和表19,其中示出了AacSHC(SEQ ID No.1)与其它SHC序列(如W0 2010/139719的ZmoSHC)有至少34-52%的同一性)具有:约50%同一性、约51%同一性、约52%同一性、约53%同一性、约54%同一性、约55%同一性、约56%同一性、约57%同一性、约58%同一性、约59%同一性、约60%同一性、约61%同一性、约62%同一性、约63%同一性、约64%同一性、约65%同一性、约66%同一性、约67%同一性、约68%同一性、约69%同一性、约70%同一性、约71%同一性、约72%同一性、

约73%同一性、约74%同一性、约75%同一性、约76%同一性、约77%同一性、约78%同一性、约79%同一性、约80%同一性、约81%同一性、约82%同一性、约83%同一性、约84%同一性、约85%同一性、约86%同一性、约87%同一性、约88%同一性、约89%同一性、约90%同一性、约91%序列同一性、约92%序列同一性、约93%序列同一性、约94%序列同一性、约95%序列同一性、约96%序列同一性、约97%序列同一性、约98%序列同一性、或约99%序列同一性。

[0082] 已经测试了SHC酶活性的各种SHC/HAC衍生物在表1-9中的一者或多者中列出。因而,在多个实施方案中,SHC/HAC衍生物可具有至少约1、至少约2、至少约3、至少约4、至少约5、至少约6、至少约7、至少约8、至少约9或至少约10个选自表1-9中的一者或多者的突变。在一些实施方案中,SHC/HAC衍生物是经修饰的SHC多肽,该多肽包含相比于根据SEQ ID No.1的野生型/参照氨基酸序列具有最多4个突变且相对于SEQ ID No.1至少包含与F129L和/或I432T中的至少任一者或多者相组合的取代F601Y或M132R的氨基酸序列,并且任选包含支持在大肠杆菌中的表达和活性的前导序列。

[0083] 在其它实施方案中,SHC/HAC衍生物是经修饰的SHC多肽,该多肽包含与根据SEQ ID No.1的野生型/参照氨基酸序列(或其例如经修饰以在大肠杆菌中表达的对应物)相比具有最多8个突变且相对于SEQ ID No.1在选自位置77、92、129、132、224、432、579、601和605的位置中包含一个或多个氨基酸改变的氨基酸序列,其中该SHC/HAC衍生物相对于SEQ ID No.1具有经修饰的(如增加的)酶活性。

[0084] 在一个实施方案中,SHC衍生物相对于SEQ ID No.1包含一个或多个取代,所述取代选自由以下组成的突变体群组:T77X、I92X、F129X、M132X、A224X、I432X、Q579X、F601Y和F605W,其中:

[0085] T77X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0086] I92X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0087] F129X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0088] M132X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0089] A224X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0090] I432X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0091] Q579X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0092] F601X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0093] F605X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0094] 在一个实施方案中,SHC衍生物相对于SEQ ID No.1包含一个或多个取代,所述取代选自由以下组成的突变体群组:T77A、I92V、F129L、M132R、A224V、I432T、Q579H、F601Y和F605W。

[0095] 在另一个实施方案中,SHC衍生物相对于SEQ ID No.2包含一个或多个取代,所述取代选自由以下组成的突变体群组:S129X、V145X、F182X、Y185X、G282X、I498X、H646X和F698X,其中:

[0096] S129X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0097] V145X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0098] F182X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0099] Y185X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0100] G282X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0101] I498X具有选自以下的X:A、B、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0102] H646X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0103] F668X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0104] F698X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0105] 在一个实施方案中,如表2中所示,SHC衍生物相对于SEQ ID No.2包含一个或多个取代,所述取代选自由以下组成的突变体群组:S129A、V145V、F182L、Y185R、G282V、I498T、H646H、F668Y和F698X。

[0106] 在一另外的实施方案中,SHC衍生物相对于SEQ ID No.3包含一个或多个取代,所述取代选自由以下组成的突变体群组:G85X、V100X、F137X、I140X、V233X、I450X、N598X、F620X和F624X,其中:

[0107] G85X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0108] V100X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0109] F137X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0110] I140X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0111] V233X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0112] I450X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0113] N598X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0114] F620X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0115] F624X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0116] 在一个实施方案中,如表3和表3a中所示,SHC衍生物相对于SEQ ID No.3包含一个或多个取代,所述取代选自由以下组成的突变体群组:G85A、V100V、F137L、I140R、V233V、I450T、N598H、F620Y和F624W。

[0117] 在一另外的实施方案中,SHC衍生物相对于SEQ ID No.4包含一个或多个取代,所述取代选自由以下组成的突变体群组:A88X、V104X、F141X、Y144X、V241X、I459X、M607X、F628X和F658X,其中:

[0118] A88X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0119] V104X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0120] F141X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0121] Y144X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0122] V241X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0123] I459X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0124] M607X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0125] F628X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0126] F658X具有选自以下的X:A、C、D、E、F、G、H、I、K、L、M、N、P、Q、R、S、T、V、W或Y。

[0127] 在一另外的实施方案中,如表4中所示,SHC衍生物相对于SEQ ID No.4包含一个或多个取代,所述取代选自由以下组成的组:A88A、V104V、F141L、Y144R、V241V、I459T、M607H、F628Y和F658W。

[0128] SHC衍生物组合

[0129] 在一个实施方案中,如表5中所示,SHC衍生物相对于SEQ ID No.1包含一个或多个取代,所述取代选自由以下组成的突变体群组:T77A、F129L、M132R、I92V、A224V、I432T、Q579H和F601Y。

[0130] 在一个实施方案中,如表6中所示,SHC衍生物包含相对于SEQ ID No.2的一个或多个取代,所述取代选自由以下组成的突变体群组:S129A、V145V F182L、Y185R、G282V、I498T、H646H和F668Y。

[0131] 在一个实施方案中,如表7中所示,SHC衍生物相对于SEQ ID No.3包含一个或多个取代,所述取代选自由以下组成的突变体群组:G85A、V100V F137L、I140R、V233V、I450T、N598H和F620Y。

[0132] 在一另外的实施方案中,如表8中所示,SHC衍生物相对于SEQ ID No.4包含一个或多个取代,所述取代选自由以下组成的组:A88A、V104V、F141L、Y144R、V241V、I459T、M607H和F628Y。

[0133] 表1:相对于野生型AacSHC (SEQ ID No.1) 和来自Hoshino和Sato (2002,如上文所引用的)的图1进行编号的SHC突变的汇总

[0134]

No.	在野生型 AacSHC 中的位置	定位	氨基酸 (AacSHC) (SEQ ID No. 1)	氨基酸 (ZmoSHC) (Hoshino 和 Sato 2002)	氨基酸 智人 (<i>H. sapiens</i>) (Hoshino 和 Sato 2002)	突变	在 SEQ ID No. 中的氨 基酸 (表 14)
1	77	QW5c-QW5b	T	G	G	A	5
8	92	QW5c-QW5b	I	V	A	V	7
2	129	QW5c-QW5b	F	F	W	L	9
3	132	QW5c-QW5b	M	I	F	R	11
4	224	QW5c-QW5b	A	V	V	V	13
5	432	QW4-QW3	I	I	N	T	15
6	579	QW1	Q	N	R	H	17
7	601	QW1	F	F	F	Y	19
9	605		F	F	F	W	171

[0135] 表2:相对于野生型AacSHC (SEQ ID No.1) 进行编号的SHC突变以及AacSHC与ZmoSHC1 (Seitz等人2012,如上文所引用的) 序列比对的汇总,补充数据表

[0136]

No.	在 WT AacSHC 中的位置 和氨基酸 (SEQ ID No. 1)		在 ZmoSHC1 中的相当 氨基酸和位置 (SEQ ID No. 2)		突变	在 SEQ ID No. 中的氨基酸 (参见表 15)
1	77	T	S	129	A	41
8	92	I	V	145	V	43
2	129	F	F	182	L	45
3	132	M	Y	185	R	47
4	224	A	G	282	V	49

[0137]

5	432	I	I	498	T	51
6	579	Q	H	647	H	53
7	601	F	F	668	Y	55
9	605	F	F	698	W	173

[0138] 表3: 相对于野生型AacSHC (SEQ ID No.1) 进行编号的SHC突变以及AacSHC与ZmoSHC2 (Seitz等人2012, 如上文所引用的) 比对的汇总, 补充数据表

[0139]

No.	在 WT AacSHC 中的位置 和氨基酸 (SEQ ID No. 1)		在 ZmoSHC2 中的 氨基酸和位置 (SEQ ID No. 3)		突变	在 SEQ ID No. 中的氨基酸 (参见表 16)
1	77	T	G	85	A	77
8	92	I	V	100	V	79
2	129	F	F	137	L	81
3	132	M	I	140	R	83
4	224	A	V	233	V	85
5	432	I	I	450	T	87
6	579	Q	N	598	H	89
7	601	F	F	620	Y	91
9	605	F	F	624	W	175

[0140] 表3a: 相对于野生型AacSHC (SEQ ID No.1) 编号的SHC突变以及AacSHC与Merkofer博士论文(2004) (参见<http://elib.uni-stuttgart.de/handle/11682/1400>) 的SHC比对图

中的序列No.20的序列比对的汇总

[0141]

No.	在 WT AacSHC 中的 位置和氨基酸 (SEQ ID No. 1)		在 ZmoSHC 中的 氨基酸	突变
1	77	T	G	A
8	92	I	V	V
2	129	F	F	L
3	132	M	I	R
4	224	A	V	V
5	432	I	I	T
6	579	Q	N	H
7	601	F	F	Y
9	605	F	F	W

[0142] 表4:相对于野生型AacSHC (SEQ ID No.1) 编号的SHC突变以及AacSHC与BjpSHC (WO 2010/139719中的SEQ ID No.5) 的序列比对的汇总

[0143]

No.	在 WT AacSHC 中的 位置和氨基酸 (SEQ ID No. 1)		在 WT BjpSHC 中的 氨基酸和位置 (SEQ ID No. 4)		突变	在 SEQ ID No. 中 的氨基酸 (参见表 17)
1	77	T	A	88	A	113
8	92	I	V	104	V	115
2	129	F	F	141	L	117
3	132	M	Y	144	R	119
4	224	A	V	241	V	121
5	432	I	I	459	T	123
6	579	Q	M	607	H	125
7	601	F	F	628	Y	127
9	605	F	F	658	W	177

[0144] 表4a:利用表21a-21j中提供的序列比对(其中将WTAacSHC (SEQ ID No.1) 与SEQ ID No.149、151、153、155、157、159(如下文所确定的) 中的任一者比对),可鉴别对应于WT

AacSHC (SEQ ID No.1) 中的T77、I92、F129、M132、A224、I432、Q579和F601的氨基酸残基和位置并测试SHC/HAC活性

氨基酸 SEQ ID No.	菌株	WO 2010/0139719 中的	核苷酸/氨基酸
		SEQ ID No. (核苷酸/氨基酸序列)	SEQ ID No.
SEQ ID No. 149	双向伯克霍尔德氏菌 (<i>Burkholderia ambifaria</i>) 6	SEQ ID No. 6	150
SEQ ID No. 151	双向伯克霍尔德氏菌	SEQ ID No. 7	152
SEQ ID No. 153	炭疽芽孢杆菌 (<i>Bacillus anthracis</i>)	SEQ ID No. 8	154
SEQ ID No. 155	桉木弗兰克氏菌	SEQ ID No. 9	156
SEQ ID No. 157	<i>Rhodopseudomonas palent</i>	SEQ ID No. 10	158
SEQ ID No. 159	天蓝色链霉菌 (<i>Streptomyces coelicolor</i>)	SEQ ID No. 11	160
SEQ ID No. 161	运动发酵单胞菌2	SEQ ID No. 2	162
SEQ ID No. 163	运动发酵单胞菌	SEQ ID No. 1	164
SEQ ID No. 4	慢生型大豆根瘤菌5	SEQ ID No. 5	168

[0146] 表5:根据野生型AacSHC (SEQ ID No.1) 编号的SHC突变组合的汇总

AacSHC (Seq ID No. 1) 中的突变组合 (Seq ID No. 1)	SHC 衍生物 ID	突变的数目	氨基酸 SEQ ID No.	核苷酸 SEQ ID No.
M132R + A224V + I432T	215G2	3	21	22
M132R + I432T	SHC26	2	23	24
F601Y	SHC3	1	25	26
T77A + I92V+F129L	111C8	3	27	28
Q579H +F601Y	101A10	2	29	30

F129L	SHC10	1	31	32
F129L + F601Y	SHC30	2	33	34
F129L+ M132R + I432T	SHC31	3	35	36
M132R + I432T + F601Y	SHC32	3	37	38
F129L+ M132R + I432T + F601Y	SHC33	4	39	40

[0149] 表6:根据野生型AZmoSHC1序列 (SEQ ID No.2) 编号的SHC突变组合的汇总

[0150]

ZmoSHC1 (Seq ID No. 2) 中的突变组合 (SEQ ID No. 2)	SHC 衍生物 ID	突变的数目	氨基酸 SEQ ID No.	核苷酸 SEQ ID No.
Y185R + G282V + I498T	215G2ZM1	3	57	58
Y185R + I498T	SHC26 ZM1	2	59	60
F668Y	SHC3 ZM1	1	61	62
S129A + V145V+F182L	111C8 ZM1	3	63	64
H646H +F668Y	101A10 ZM1	2	65	66
F182L	SHC10 ZM1	1	67	68
F182L + F668Y	SHC30 ZM1	2	69	70
F182L+ Y185R + I498T	SHC31 ZM1	3	71	72
Y185R + I498T + F668Y	SHC32 ZM1	3	73	74
F182L+ Y185R + I498T + F668Y	SHC33 ZM1	4	75	76

[0151] 表7:根据野生型ZmoSHC2序列 (SEQ ID No.3) 编号的SHC突变组合的汇总

[0152]

ZmoSHC2 (Seq ID No. 3) 中的突变组合 (SEQ ID No. 3)	SHC 衍生物 ID	突变的数目	氨基酸 SEQ ID No.	核苷酸 SEQ ID No.
I140R + V233V + I450T	215G2 ZM2	3	93	94
I140R + I450T	SHC26 ZM2	2	95	96
F620Y	SHC3 ZM2	1	97	98
G85A + V100V+F137L	111C8 ZM2	3	99	100
N598H +F620Y	101A10 ZM2	2	101	102
F137L	SHC10 ZM2	1	103	104
F137L + F620Y	SHC30 ZM2	2	105	106
F137L+ I140R + I450T	SHC31 ZM2	3	107	108
I140R + I450T + F620Y	SHC32 ZM2	3	109	110
F137L+ I140R + I450T + F620Y	SHC33 ZM2	4	111	112

[0153] 表8:根据野生型BjpSHC (SEQ ID No.4) 编号的SHC突变组合的汇总

[0154]

BjpSHC (Seq ID No. 4) 中的突变组合 (SEQ ID No. 4)	SHC 衍生物 ID	突变的数目	氨基酸 SEQ ID No.	核苷酸 SEQ ID No.
Y144R + V241V + I459T	215G2 Bjp	3	129	130
Y144R + I459T	SHC26 Bjp	2	131	132
F628Y	SHC3 Bjp	1	133	134
A88A + V104V + F141L	111C8 Bjp	3	135	136
M607H + F628Y	101A10 Bjp	2	137	138
F141L	SHC10 Bjp	1	139	140
F141L + F628Y	SHC30 Bjp	2	141	142
F141L + Y144R + I459T	SHC31 Bjp	3	143	144

[0155]

M144R + I459T + F628Y	SHC32 Bjp	3	145	146
F141L + Y144R + I459T + F628Y	SHC33 Bjp	4	147	148

[0156] 表9: 示出了相对于WT AacSHC (SEQ ID No.1)、WT ZmoSHC1 (SEQ ID No.2)、WT ZmoSHC2 (SEQ ID No.3) 和BjpSHC (SEQ ID No.4) 的共同SHC突变

[0157]

突变组合				
AacSHC (SEQ ID No. 1) 中 (SEQ ID No. 1)	相对于 ZmoSHC1 (SEQ ID No. 2)	相对于 ZmoSHC2 (SEQ ID No. 3)	相对于 BjpSHC (SEQ ID No. 4)	SHC 衍生物
F601Y	F668Y	F620Y	F628Y	SHC3
F129L	F182L	F137L	F141L	SHC10
F601Y + F129L	F668Y + F182L	F620Y + F137L	F628Y + F141L	SHC30
F601Y + M132R + I432T	F668Y + Y185R + I498T	F620Y + I140R + I450T	F628Y + Y144R + I459T	SHC32
F601Y + F129L + M132R + I432T	F668Y + F182L + Y185R + I498T	F620Y + F137L + I140R + I450T	F628Y + F141L + Y144R + I459T	SHC33
M132R + I432T	Y185R + I498T	I140R + I450T	Y144R + I459T	SHC26
M132R + I432T + A224V	Y185R + I498T + G282V	I140R + I450T + V233V	Y144R + I459T + V241V	215G2
M132R + I432T + F129L	Y185R + I498T + F182L	I140R + I450T + F137L	Y144R + I459T + F141L	SHC31

[0158] 在一个优选的实施方案中,相对于SEQ ID No.1,SHC衍生物至少包含与F129L和/或I432T中的至少任一者或多者相组合的取代F601Y或M132R。

[0159] 与参照SHC蛋白(SEQ ID No.1)相比,本公开中提供的称为SHC3的SHC衍生物包含以下取代:F601Y。

[0160] Hoshino和Sato(2002,如上文所引用的)将F601鉴别为原核和真核生物物种中高度保守的氨基酸残基。据报道,对环氧角鲨烯底物(不是角鲨烯),SHC衍生物F601Y显示出大大增加的 V_{max} 。然而当使用角鲨烯时,相对于野生型AacSHC,F601Y显示出亲和力降低(即更高的 K_M)和催化效率/活性(K_{cat}/K_M)降低。Hoshino和Sato(2002,如上文所引用的)中未提供当将高法呢醇作为酶底物用于F601Y突变体时AacSHC效力的数据。

[0161] 与参照SHC蛋白(SEQ ID No.1)相比,本公开中提供的称为SHC10的SHC衍生物包含以下取代:F129L。

[0162] 与参照SHC蛋白(SEQ ID No.1)相比,本公开中提供的称为SHC30的SHC衍生物包含以下取代:F601Y和F129L。

[0163] 与参照SHC蛋白(SEQ ID No.1)相比,本公开中提供的称为SHC26的SHC衍生物包含以下取代:M132R和I432T。

[0164] 与参照SHC蛋白(SEQ ID No.1)相比,本公开中提供的称为215G2的SHC衍生物包含以下取代:M132R、I432T和A224V。

[0165] 与参照SHC蛋白 (SEQ ID No.1) 相比,本公开中提供的称为SHC32的SHC衍生物包含以下取代:F601Y、M132R和I432T。

[0166] 与参照SHC蛋白 (SEQ ID No.1) 相比,本公开中提供的称为SHC31的SHC衍生物包含以下取代:F129L、M132R和I432T。

[0167] 与参照SHC蛋白 (SEQ ID No.1) 相比,本公开中提供的称为SHC33的SHC衍生物包含以下取代:F601Y、F129L、M132R和I432T。

[0168] 与参照SHC蛋白 (SEQ ID No.1) 相比,本公开中提供的称为101A10的SHC衍生物包含以下取代:F601Y和Q579H。

[0169] 与参照SHC蛋白 (SEQ ID No.1) 相比,本公开中提供的称为111C8的SHC衍生物包含以下取代:T77A+I92V和F129L。

[0170] 在一个优选的实施方案中,相对于SEQ ID No.2,SHC衍生物至少包含与F182L和/或I498T中的至少任一者或多者相组合的取代F668Y或Y185R。

[0171] 与参照SHC蛋白 (SEQ ID No.2) 相比,本公开中提供的称为SHC3ZM1的SHC衍生物包含以下取代:F668Y。

[0172] Hoshino和Sato (2002,如上文所引用的) 将F601鉴别为原核和真核生物物种中高度保守的氨基酸残基。据报道,对环氧角鲨烯底物(不是角鲨烯),SHC衍生物F601Y显示出大大增加的 V_{max} 。然而,当使用角鲨烯时,相对于野生型AacSHC,F601Y显示出亲和力降低(即更高的 K_M)和催化效率/活性(K_{cat}/K_M)降低。Hoshino和Sato中未提供当将高法呢醇作为酶底物用于F601Y突变体时AacSHC效力的数据。相当于ZmoSHC1中的F601Y的SHC衍生物是F668Y。

[0173] 与参照SHC蛋白 (SEQ ID No.2) 相比,本公开中提供的称为SHC10ZM1的SHC衍生物包含以下取代:F182L。

[0174] 与参照SHC蛋白 (SEQ ID No.2) 相比,本公开中提供的称为SHC30ZM1的SHC衍生物包含以下取代:F668Y和F182L。

[0175] 与参照SHC蛋白 (SEQ ID No.2) 相比,本公开中提供的称为SHC26ZM1的SHC衍生物包含以下取代:Y185R和I498T。

[0176] 与参照SHC蛋白 (SEQ ID No.2) 相比,本公开中提供的称为215G2ZM1的SHC衍生物包含以下取代:Y185R、I498T和G282V。

[0177] 与参照SHC蛋白 (SEQ ID No.2) 相比,本公开中提供的称为SHC32ZM1的SHC衍生物包含以下取代:F668Y、Y185R和I498T。

[0178] 与参照SHC蛋白 (SEQ ID No.2) 相比,本公开中提供的称为SHC31ZM1的SHC衍生物包含以下取代:F182L、Y185R和I498T。

[0179] 与参照SHC蛋白 (SEQ ID No.2) 相比,本公开中提供的称为SHC33ZM1的SHC衍生物包含以下取代:F668Y、F182L、Y185R和I498T。

[0180] 与参照SHC蛋白 (SEQ ID No.2) 相比,本公开中提供的称为101A10ZM1的SHC衍生物包含以下取代:F668Y和H646H。

[0181] 与参照SHC蛋白 (SEQ ID No.2) 相比,本公开中提供的称为111C8ZM1的SHC衍生物包含以下取代:S129A+V145V和F182L。

[0182] 在一个优选的实施方案中,相对于SEQ ID No.3,SHC衍生物至少包含与F137L和/

或I450T中的至少任一者或多者相组合的取代F620Y或I140R。

[0183] 与参照SHC蛋白 (SEQ ID No.3) 相比,本公开中提供的称为SHC3ZM2的SHC衍生物包含以下取代:F620Y。

[0184] Hoshino和Sato (2002,如上文所引用的) 将F601鉴别为原核和真核生物SHC种类中高度保守的氨基酸残基。据报道,对环氧角鲨烯底物(不是角鲨烯),AacSHC衍生物F601Y显示出大大增加的 V_{max} 。然而,当使用角鲨烯时,相对于野生型AacSHC,F601Y显示出亲和力降低(即更高的 K_M)和催化效率/活性(K_{cat}/K_M)降低。Hoshino和Sato (2002) 中未提供当将高法呢醇作为酶底物用于F601Y突变体时AacSHC效力的数据。相当于ZmoSHC2中的F601Y的SHC衍生物是F620Y。

[0185] 与参照SHC蛋白 (SEQ ID No.3) 相比,本公开中提供的称为SHC10ZM2的SHC衍生物包含以下取代:F137L。

[0186] 与参照SHC蛋白 (SEQ ID No.3) 相比,本公开中提供的称为SHC30ZM2的SHC衍生物包含以下取代:F620Y和F137L。

[0187] 与参照SHC蛋白 (SEQ ID No.3) 相比,本公开中提供的称为SHC26ZM2的SHC衍生物包含以下取代:I140R和I450T。

[0188] 与参照SHC蛋白 (SEQ ID No.3) 相比,本公开中提供的称为215G2ZM2的SHC衍生物包含以下取代:I140R、I450T和V233V。

[0189] 与参照SHC蛋白 (SEQ ID No.3) 相比,本公开中提供的称为SHC32ZM2的SHC衍生物包含以下取代:F620Y、I140R和I450T。

[0190] 与参照SHC蛋白 (SEQ ID No.3) 相比,本公开中提供的称为SHC31ZM2的SHC衍生物包含以下取代:F137L、I140R和I450T。

[0191] 与参照SHC蛋白 (SEQ ID No.3) 相比,本公开中提供的称为SHC33ZM2的SHC衍生物包含以下取代:F620Y、F137L、I140R和I450T。

[0192] 与参照SHC蛋白 (SEQ ID No.3) 相比,本公开中提供的称为101A10ZM2的SHC衍生物包含以下取代:F620Y和N598H。

[0193] 与参照SHC蛋白 (SEQ ID No.3) 相比,本公开中提供的称为111C8ZM2的SHC衍生物包含以下取代:G85A+V100V和F137L。

[0194] 在一个优选的实施方案中,相对于SEQ ID No.4,SHC衍生物至少包含与F141L和/或I459T中的至少任一者或多者相组合的取代F628Y或Y144R。

[0195] 与参照SHC蛋白 (SEQ ID No.4) 相比,本公开中提供的称为SHC3Bjp的SHC衍生物包含以下取代:F628Y。

[0196] Hoshino和Sato (2002,如上文所引用的) 将F601鉴别为原核和真核生物种类中高度保守的氨基酸残基。据报道,对环氧角鲨烯底物(不是角鲨烯),SHC衍生物F601Y显示出大大增加的 V_{max} 。然而,当使用角鲨烯时,相对于野生型AacSHC,F601Y显示出亲和力降低(即更高的 K_M)和催化效率/活性(K_{cat}/K_M)降低。Hoshino和Sato中未提供当将高法呢醇作为酶底物用于F601Y突变体时AacSHC效力的数据。相当于BjpSHC中的F601Y的SHC衍生物是F628Y。

[0197] 与参照SHC蛋白 (SEQ ID No.4) 相比,本公开中提供的称为SHC10Bjp的SHC衍生物包含以下取代:F141L。

[0198] 与参照SHC蛋白 (SEQ ID No.4) 相比,本公开中提供的称为SHC30Bjp的SHC衍生物包含以下取代:F628Y和F141L。

[0199] 与参照SHC蛋白 (SEQ ID No.4) 相比,本公开中提供的称为SHC26Bjp的SHC衍生物包含以下取代:Y144R和I459T。

[0200] 与参照SHC蛋白 (SEQ ID No.4) 相比,本公开中提供的称为215G2Bjp的SHC衍生物包含以下取代:Y144R、I459T和V241V。

[0201] 与参照SHC蛋白 (SEQ ID No.4) 相比,本公开中提供的称为SHC32Bjp的SHC衍生物包含以下取代:F628Y、Y144R和I459T。

[0202] 与参照SHC蛋白 (SEQ ID No.4) 相比,本公开中提供的称为SHC31Bjp的SHC衍生物包含以下取代:F141L、Y144R和I459T。

[0203] 与参照SHC蛋白 (SEQ ID No.4) 相比,本公开中提供的称为SHC33Bjp的SHC衍生物包含以下取代:F628Y、F141L、Y144R和I459T。

[0204] 与参照SHC蛋白 (SEQ ID No.4) 相比,本公开中提供的称为101A10Bjp的SHC衍生物包含以下取代:F628Y和M607H。

[0205] 与参照SHC蛋白 (SEQ ID No.4) 相比,本公开中提供的称为111C8Bjp的SHC衍生物包含以下取代:A88A+V104V和F141L。

[0206] 氨基酸序列

[0207] 在一些实施方案中,AacSHC/HAC衍生物包含SEQ ID No.5、7、9、11、13、15、17、19、21、23、25、27、29、31、33、35、37、39和/或171中的一者或多者中所列出的多肽中的一者或多者。

[0208] 优选地,本公开的AacSHC/HAC衍生物具有选自以下的氨基酸序列:SEQ ID No.21、SEQ ID No.23、SEQ ID No.25、SEQ ID No.27、SEQ ID No.29、SEQ ID No.31、SEQ ID No.33、SEQ ID No.35、SEQ ID No.37、SEQ ID No.39和/或SEQ ID No.171。

[0209] 在其它实施方案中,ZmoSHC1/HAC衍生物包含SEQ ID No.41、43、45、47、49、51、53、55、57、59、61、63、65、67、69、71、73、75和/或173中的一者或多者中所列出的多肽中的一者或多者。

[0210] 优选地,本公开的ZmoSHC1/HAC衍生物具有选自以下的氨基酸序列:SEQ ID No.57、SEQ ID No.59、SEQ ID No.61、SEQ ID No.63、SEQ ID No.65、SEQ ID No.67、SEQ ID No.69、SEQ ID No.71、SEQ ID No.73、SEQ ID No.75和/或SEQ ID No.173。

[0211] 在另外的实施方案中,ZmoSHC2/HAC衍生物包含SEQ ID No.77、SEQ ID No.79、SEQ ID No.81、SEQ ID No.83、SEQ ID No.85、SEQ ID No.87、SEQ ID No.89、SEQ ID No.91、SEQ ID No.93、SEQ ID No.95、SEQ ID No.97、SEQ ID No.99、SEQ ID No.101、SEQ ID No.103、SEQ ID No.105、SEQ ID No.107、SEQ ID No.109、SEQ ID No.111和/或SEQ ID No.175中的一者或多者中所列出的多肽中的一者或多者。

[0212] 在另外的实施方案中,BjpSHC/HAC衍生物包含以下中的一者或多者中所列出的多肽中的一者或多者:SEQ ID No.113、SEQ ID No.115、SEQ ID No.117、SEQ ID No.119、SEQ ID No.121、SEQ ID No.123、SEQ ID No.125、SEQ ID No.127、SEQ ID No.129、SEQ ID No.131、SEQ ID No.133、SEQ ID No.135、SEQ ID No.137、SEQ ID No.139、SEQ ID No.141、SEQ ID No.143、SEQ ID No.145、SEQ ID No.147和/或SEQ ID No.177。

[0213] 序列比对

[0214] 由于SHC参照序列(例如,AacSHC、ZmoSHC1、ZmoSHC2和BjpSHC多肽序列)长度不同,参照AacSHC序列(SEQ ID No.1)的位置X处的氨基酸残基对应ZmoSHC1参照序列(SEQ ID No.2)上的不同氨基酸位置B、ZmoSHC2参照序列(SEQ ID No.3)上的不同氨基酸位置J以及BjpSHC参照序列(SEQ ID No.4)上的不同氨基酸位置Z。此外,SHC参照序列的改变也可相对该参照SHC序列而言修饰SHC衍生物序列。

[0215] 术语“位置”是指参照SHC蛋白中存在的通过特定氨基酸编号所确定的具体氨基酸残基。通过插入或缺失氨基酸来使SHC参照蛋白改变可导致该参照SHC氨基酸序列与SHC衍生物氨基酸序列之间的编号不一样。举个例子,如果在参照SHC蛋白的氨基酸509与氨基酸510之间插入一个氨基酸,则该插入后面的氨基酸将在SHC衍生蛋白中具有编号511,而在SHC参照蛋白中其将保持编号510。

[0216] 用于测定WT SHC/HAC和SHC/HAC衍生物活性的测定法

[0217] 本文描述了用于测定和定量WT SHC/HAC和/或SHC/HAC衍生酶活性的测定法,它们是本领域已知的。举个例子,WT SHC/HAC和/或SHC/HAC衍生物活性可通过这样来测定:将纯化的SHC/HAC酶或来自产生SHC/HAC酶的宿主细胞的萃取物或完整的重组宿主生物体与适当的底物在适当的条件下温育并对反应产物进行分析(如通过气相色谱(GC)或HPLC分析)。关于SHC/HAC和/或SHC/HAC酶活性测定法以及反应产物分析的进一步细节在实施例中提供。这些测定法包括在重组宿主细胞(如大肠杆菌)中产生SHC衍生物。

[0218] 如本文所用,术语“活性”意指酶与底物反应而提供靶产物的能力。可以在通过靶产物随时间增加、底物(或原料)随时间减少或通过这些参数的组合随时间变化来测定的活性测定法中测定活性。本公开的SHC/HAC衍生物表征为它们将高法呢醇生物转化为(-)-降龙涎醚并且展示出生物活性诸如HAC活性的能力。

[0219] 本文所用的“生物活性”是指多肽可能展现的任何活性,包括但不限于:酶活性;结合另一化合物的活性(如结合至另一多肽,尤其是结合至受体,或者结合至核酸);抑制活性(例如酶抑制活性);激活活性(如酶激活活性);或毒性效应。并非要求该变体或衍生物展现出的这种活性的程度与亲本多肽一样。如果变体展现出的相关活性的程度为亲本多肽的活性的至少10%,则该变体被认为是本申请范围内的变体。同样,如果衍生物展现出的相关生物活性的程度为亲本多肽的活性的至少10%,则该衍生物被认为是本申请范围内的衍生物(因为术语衍生物和变体在整个本公开中可互换使用)。

[0220] 在其它实施方案中,本公开的SHC/HAC衍生物显示出比参照SHC蛋白更好的靶产率。术语“靶产率”是指每克原料的可回收产物克数(其可计算为摩尔转化率百分数)。

[0221] 在另外的实施方案中,相对于参照SHC蛋白,本公开的SHC/HAC衍生物显示出修饰的(如增加的)靶生产力。术语“靶生产力”是指每小时生物转化时间(即添加底物后的时间)每升发酵容量的可回收靶产物的量(克数)。

[0222] 在另外的实施方案中,与参照SHC蛋白相比,本公开的SHC/HAC衍生物显示出改变的靶产率因子。术语“靶产率因子”是指反应介质中所获得的产物浓度与SHC衍生物(例如纯化的SHC酶或来自表达SHC酶的重组宿主细胞的萃取物)浓度之间的比率。

[0223] 在多个实施方案中,相对于参照SHC蛋白(如SEQ ID No.1或SEQ ID No.2或SEQ ID No.3或SEQ ID No.4),本公开的SHC衍生物显示出改变的(如增加的)酶活性增加倍数(如改

变的/增加的高法呢醇降龙涎香醚环化酶(HAC)活性)。该活性增加为至少增加:2、3、4、6、8、10、12、14、16、18、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95和/或100倍。

[0224] 核苷酸序列

[0225] 本公开还涉及分离的核酸分子,该核酸分子包含编码如本文所述的SHC衍生物的核苷酸序列。

[0226] 本文所用的术语“核酸分子”应该特定地指本公开的多核苷酸,其可以为DNA、cDNA、基因组DNA、合成DNA或RNA,并且可以是双链的或单链的、有义链和/或反义链。术语“核酸分子”应该特别适用于本文所用的多核苷酸,如作为全长核苷酸序列或者其片段或部分,其分别编码具有酶活性的多肽(如代谢途径的酶)、或者其片段或部分。

[0227] 该术语还可包括:单独的分子诸如cDNA,其中对应的基因组DNA基因具有内含子,因此是不同的序列;缺少旁侧基因中的至少一者的基因组片段;通过聚合酶链反应(PCR)制备并且缺少旁侧基因中的至少一者的cDNA或基因组DNA的片段;缺少旁侧基因中的至少一者的限制性酶切片断;编码非天然存在的蛋白质诸如融合蛋白(如His标签)、突变蛋白质或给定蛋白质的片段的DNA;以及作为cDNA或天然存在的核酸的简并变体的核酸。此外,其包括作为杂合基因(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。融合蛋白可给蛋白质添加一个或多个氨基酸(例如但不限于组氨酸(His)),通常是在蛋白质的N端,但也有在C端的或者融合于蛋白质区域内。这种融合蛋白或编码这种蛋白质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性;以及(iii)通过提供亲和纯化的配体而帮助对重组蛋白的纯化。术语“核酸分子”还包括适于在特定微生物宿主细胞(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”意指已通过用在细菌(如大肠杆菌)宿主细胞基因中更频繁使用的密码子取代一个密码子或多个密码子或优选大量的密码子而使其适于在原核或真核宿主细胞、尤其是细菌宿主细胞(诸如大肠杆菌宿主细胞)中表达的核酸蛋白质编码序列。就这一点而言,编码参照序列Seq ID No.1、2、3和/或4以及其所有的变体/衍生物的核苷酸序列可以是存在于来源中的原始序列(如分别是AacSHC、ZmoSHC1、ZmoSHC2或BjpSHC)或者该基因可以是针对所选择的宿主生物体(如大肠杆菌)进行了密码子优化的。

[0228] 可以通过体外转录取制备核糖核酸(RNA)分子。DNA分子的区段也可视为在本公开的范围,它们可通过例如聚合酶链反应(PCR)制备或通过用一种或多种限制性内切酶处理来生成。核酸分子的区段可称为基因(尤其是作为部分基因的那些)的DNA片段。片段也可含有若干开放阅读框(ORF),可以是相同ORF的重复或者是不同的ORF。该术语应该具体地指编码核苷酸序列,但还应该包括非编码性的核苷酸序列(如非转录序列或非翻译序列)或编码多肽(整个多肽或部分多肽)的核苷酸序列。本文所用的基因(如用于组装、多样化或重组)可以是非编码序列,或者编码多肽的序列或蛋白质编码序列,或者其具有足以进行成功的重组事件的序列长度的部分或片段。更具体而言,所述基因最低长度为3bp、优选至少100bp、更优选至少300bp。

[0229] 根据上述内容将显而易见的是,提及分离的DNA时,并不是指存在于例如cDNA或基因组DNA文库内或者基因组DNA限制酶切消化物(在例如限制酶切消化反应混合物或电泳凝胶切片中)内的数百到数百万其它DNA分子之中的DNA。本公开的分离的核酸分子涵盖存在方式与天然状态不一样的区段。

[0230] 如本文所用,术语“分离的DNA”可以指:(1)含有与任何天然存在的序列不同的序列的DNA、不是天然存在的多核苷酸或核酸(如,通过经由人为干预将两个原本分离的序列片段人工重组(如人工操作分离的核酸片段,如通过遗传工程技术来进行)而制备,或者(2),在具有天然存在的序列的DNA(如,cDNA或基因组DNA)的背景下,不含有在天然存在含有所关注DNA的基因的生物体基因组中处于含有所关注DNA的基因旁侧的至少一个基因的DNA。

[0231] 本文所用的术语“分离的DNA”,特别是相对于核酸序列而言,也可指通过重组DNA技术制备的核酸或多核苷酸,例如包含宿主细胞异源的多核苷酸的DNA构建体,其任选整合进宿主细胞中。可具体地以重组分子形式制备嵌合核苷酸序列。术语“重组”应该特定地适用于多核苷酸的组装,将这样的多核苷酸或其部分连接在一起,存在或不存在重组来实现交换或基因嵌合。例如,将具有所需功能的核酸区段连接在一起以产生所需的功能组合。编码本文所述的多肽的重组基因包括以有义取向有效连接至适于表达所述多肽的一个或多个调控区的该多肽的编码序列。因为许多微生物能够从多顺反子mRNA表达多种基因产物,因此可根据需要在那些微生物的单个调控区的控制下表达多种多肽。当调控区和编码序列所处位置可使得调控区可有效调控编码序列的转录或翻译时,则所述编码序列和调控区被视为有效连接。

[0232] 本文所用的术语“重组”(具体针对酶时),应指由重组DNA技术制备的酶,即从用编码所需酶的外源DNA构建体转化的细胞产生的酶。“合成的”酶是通过化学合成制备的那些酶。可具体地作为重组分子制备嵌合酶。术语“重组DNA”因此包括掺入进载体中、掺入进自主复制质粒或病毒中、或掺入进原核生物或真核生物的基因组DNA中(或在同源细胞基因组中处于不同于天然染色体位置的位置处)的重组DNA。

[0233] 在另一个方面,本公开的核酸分子有效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,“有效连接的”意指掺入进遗传构建体中从而表达控制序列有效控制所关注的编码序列的表达。上文提及的转录/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、代谢调控型的启动子、增强子、操纵子、沉默子、阻遏子以及本领域技术人员已知的并且驱动或以别的方式调控基因表达的其它元件。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,像例如,CUP-1启动子、如例如在tet-on或tet-off系统中所采用的tet阻遏子、lac系统、trp系统调控元件。举个例子,异丙基 β -D-硫代半乳糖苷(IPTG)在100 μ M至1.0mM的浓度范围内是基因表达的有效诱导物。该化合物是异乳糖的分子模拟物,异乳糖是一种乳糖代谢物,其引发乳糖操纵子的转录,因而在基因处于乳糖操纵子的控制下时其用于诱导基因表达。诱导基因表达的调控元件的另一个实例是乳糖。

[0234] 相似地,本公开的核酸分子可形成编码额外多肽序列(例如,充当标记或报道物的序列)的杂合基因的一部分。标记基因和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)、氨基糖苷磷酸转移酶二氢叶酸还原酶(DHFR)、潮霉素B磷酸转移酶(HPH)、胸苷激酶(TK)、lacZ(编码 β -半乳糖苷酶)以及黄嘌呤鸟嘌呤磷酸核糖转移酶(XGPRT)。如果实践本公开相关的许多标准程序一样,技术人员将会意识到额外的有用试剂,例如可发挥标记或报道物功能的额外序列。

[0235] 在一些实施方案中,本公开提供编码上述WT SHC或SHC/HAC衍生物的重组多核苷

酸,其可插入载体中供表达以及任选的纯化。一种类型的载体是呈现环状双链DNA环的质粒,另外的DNA区段连接在其中。某些载体可控制与它们功能性连接的基因的表达。这些载体称为“表达载体”。通常适于DNA重组技术的表达载体是质粒类型。典型地,表达载体包含诸如本文所述的WT SHC或SHC/HAC变体之类的基因。在本发明描述中,术语“质粒”和“载体”可互换使用,因为质粒是最常使用的载体类型。

[0236] 这种载体可包含DNA序列,所述DNA序列包括但不限于非天然存在于宿主细胞中的DNA序列、通常不转录成RNA或翻译成蛋白质(“表达”)的DNA序列以及期望引入进非重组宿主中的其它基因或DNA序列。应当理解,通常本文所述的重组宿主的基因组通过稳定引入一种或多种重组基因而得到增强。然而,在本公开的范围内也可使用自主型或复制型的质粒或载体。此外,本公开可使用低拷贝数(如单拷贝)或高拷贝数(如本文所示例的)质粒或载体来实践。

[0237] 在一个优选的实施方案中,本公开的载体包含质粒、噬菌粒、噬菌体、粘粒、人工细菌或人工酵母染色体、敲除或敲入构建体、合成的核酸序列或盒,并且亚组可以以线性多核苷酸、质粒、大质粒、合成的或人工的染色体诸如植物、细菌、哺乳动物或酵母人工染色体的形式制备。

[0238] 优选的是引入所述载体后,在细胞内表达由所引入的多核苷酸编码的蛋白质。各式各样的基因底物可掺入质粒中。质粒通常是标准的克隆载体,如细菌多拷贝质粒。所述底物可掺入相同或不同质粒中。通常使用具有不同类型的选择标记的至少两种不同类型的质粒以允许选择含有至少两种类型的载体的细胞。

[0239] 通常,细菌细胞或酵母细胞可用以下核苷酸序列中的任一种或多种转化,如本领域众所周知的。对于体内重组,使用标准的转化技术将待与基因组或其它基因重组的基因用于转化宿主。在一个适当的实施方案中,将提供复制起点的DNA包括在构建体中。复制起点可由技术人员适当地选择。取决于基因的性质,如果序列已经与它们自身可作复制起点起作用的基因或基因组一起存在时,则可能不需要补充的复制起点。

[0240] 当已将外源或异源DNA引入细胞内时细菌或酵母细胞可能会被这类DNA转化。转化的DNA可能会或可能不会整合,即共价连接进细胞的基因组中。例如,在原核细胞以及酵母中,转化的DNA可能保持在附加型元件(诸如质粒)上。对于真核细胞,稳定转染的细胞是其中转染的DNA已经变得整合进染色体中从而其通过染色体复制被子代细胞遗传的细胞。该稳定性通过真核细胞能够确立由含有所述转化DNA的子代细胞的群体组成的细胞系或克隆的能力来证明。

[0241] 一般而言,所引入的DNA最初不存在于接受该DNA的宿主中,但以下情况在本公开的范围内:从给定宿主分离DNA区段,并随后将该DNA的一个或多个额外的拷贝引入进同一宿主中,例如用以增强基因产物的产量或改变基因的表达模式。在某些情况下,引入的DNA将会改变或甚至替代内源基因或DNA序列,例如通过同源重组或定点诱变。合适的重组宿主包括微生物、植物细胞和植物。

[0242] 本公开还描述重组宿主。术语“重组宿主”,也称为“经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通过至少一个引入的DNA序列增强的宿主细胞。本公开的宿主细胞可以用上述多核苷酸或载体进行遗传工程改造。

[0243] 可用于本公开目的的宿主细胞包括但不限于:原核细胞诸如细菌(例如,大肠杆菌

和枯草芽孢杆菌 (*B. subtilis*)), 其可用例如含有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; 简单真核细胞像酵母 (例如酵母属 (*Saccharomyces*) 和毕赤酵母属 (*Pichia*)), 其可用例如含有本公开的多核苷酸分子的重组酵母表达载体转化。取决于用于引入本公开的多核苷酸的宿主细胞和相应载体, 所述多核苷酸可以整合, 例如, 进入染色体或线粒体 DNA 中, 或者可以维持在染色体外, 像例如附加型基因, 或者可以仅暂时包含在细胞内。

[0244] 如本文尤其是关于遗传工程以及将一种或多种基因或组装的基因簇引入进细胞或生产细胞中所用的术语“细胞”应理解为指任何原核细胞或真核细胞。原核和真核宿主细胞二者均考虑根据本公开使用, 包括细菌宿主细胞像大肠杆菌或芽孢杆菌, 酵母宿主细胞诸如酿酒酵母、昆虫宿主细胞诸如草地夜蛾 (*Spodoptora frugiperda*), 或人宿主细胞诸如 HeLa 和 Jurkat。

[0245] 特别地, 所述细胞为真核细胞 (优选真菌细胞、哺乳动物细胞或植物细胞), 或原核细胞。合适的真核细胞包括例如但不限于哺乳动物细胞、酵母细胞或昆虫细胞 (包括 Sf9)、两栖动物细胞 (包括黑素细胞) 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。合适的哺乳动物细胞包括例如但不限 COS 细胞 (包括 Cos-1 和 Cos-7)、CHO 细胞、HEK293 细胞、HEK293T 细胞、HEK293 T-RexTM 细胞或其它可转染的真核细胞系。合适的细菌细胞包括但不限于大肠杆菌。

[0246] 优选地, 可以使用原核生物, 诸如大肠杆菌、芽孢杆菌属 (*Bacillus*)、链霉菌属 (*Streptomyces*), 或哺乳动物细胞, 像 HeLa 细胞或 Jurkat 细胞, 或者植物细胞, 像拟南芥属 (*Arabidopsis*)。

[0247] 优选地, 所述细胞是曲霉属物种 (*Aspergillus* sp.) 或真菌细胞, 优选地, 其可选自以下属: 酵母属、假丝酵母属 (*Candida*)、克鲁维酵母属 (*Kluyveromyces*)、汉森酵母属 (*Hansenula*)、裂殖酵母属 (*Schizosaccharomyces*)、耶氏酵母 (*Yarrowia*)、毕赤酵母属 (*Pichia*) 和曲霉属 (*Aspergillus*)。

[0248] 优选地, 大肠杆菌宿主细胞是本行业及监管部门所认可的大肠杆菌宿主细胞 (包括但不限于大肠杆菌 K12 宿主细胞或如实施例中所展示的, 大肠杆菌 BL21 宿主细胞)。

[0249] 供用于本公开的一种优选的宿主细胞是可如本文所述以重组方法制备的大肠杆菌。因而, 重组宿主可以是重组大肠杆菌宿主细胞。对于大肠杆菌, 存在突变体、质粒、代谢的详细计算机模型以及其它信息的文库供使用, 从而允许对各种模块进行合理设计以增加产物产率。可将类似于上文针对酵母属所述的那些方法的方法用于制备重组大肠杆菌微生物。

[0250] 在一个实施方案中, 重组大肠杆菌微生物包含编码 SHC 基因的核苷酸序列 (如例如在本文表 10、11 和 12 中的任一者或多者中所公开的, 或它们的功能等价物/同源物, 包括但不限于它们的变体、同源突变体、衍生物或片段)。

[0251] 优选地, 重组大肠杆菌微生物包含如图 5 和 21 中所提供的载体构建体。

[0252] 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 WT SHC/HAC 和 WT SHC/HAC 衍生物基因或它们的功能等价物/同源物的核苷酸序列, 所述它们的功能等价物/同源物包括但不限于如表 13、表 14、表 15、表 16、表 17 和/或表 4a 中的任一者或多者中所列出的它们的变体、同源物、突变体、衍生物或片段。

[0253] 用于本公开的另一优选宿主细胞是酿酒酵母 (*S.cerevisiae*)，其为在合成生物学中广泛使用的一类生物体。因而，重组宿主可以是酿酒酵母。对于酿酒酵母，存在突变体、质粒、代谢的详细计算机模型以及其它信息的文库供使用，从而允许对各种模块进行合理设计以增加产物产率。制备重组酿酒酵母微生物的方法是已知的。

[0254] 细胞培养以常规方式进行。培养基含有碳源、至少一种氮源以及无机盐，并向其添加维生素。该培养基的组分可以是常规用于培养所考虑的微生物物种的组分。

[0255] 用于本方法的碳源包括可被重组宿主细胞代谢以利于生长和/或产生(-)-降龙涎醚的任何分子。合适碳源的实例包括但不限于蔗糖(例如，如糖蜜中所存在的)、果糖、木糖、甘油、葡萄糖、纤维素、淀粉、纤维二糖或其它含葡萄糖的聚合物。

[0256] 在采用酵母作为宿主的实施方案中，例如，诸如蔗糖、果糖、木糖、乙醇、甘油和葡萄糖之类的碳源是合适的。可在整个培养周期将碳源提供给宿主生物体，或者备选地，可让该生物体在存在另一种能量来源(如蛋白质)的情况下生长一段时间，然后仅在补料分批培养阶段提供碳源。

[0257] 重组宿主细胞微生物在本公开的方法中的适用性可通过使用众所周知的方法进行简单的测试程序来确定。例如，可将待测试的微生物在丰富培养基(如LB培养基、细菌用胰蛋白胨酵母浸膏培养基、营养培养基等)中在通常用于该微生物繁殖的pH、温度以及通气条件下繁殖。一旦选择了产生所需的生物转化产物的重组微生物(即重组宿主细胞)，则通常通过合适的表达系统和发酵由生产宿主细胞系以大规模(如通过细胞培养中的微生物生产)产生产物。

[0258] 在本公开的一个实施方案中，将已知成份的基本培养基如M9A用于细胞培养。

[0259] M9A培养基的组分包含：14g/l KH_2PO_4 、16g/l K_2HPO_4 、1g/l柠檬酸三钠、 $2\text{H}_2\text{O}$ 、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酵母萃取物。

[0260] 在本公开的另一个实施方案中，使用富营养培养基如LB。LB培养基的组分包含：10g/l胰蛋白胨、5g/l酵母萃取物、5g/l NaCl 。

[0261] 矿质培养基和M9矿质培养基的其它实例例如公开于US 6524831B2和US 2003/0092143A1中。

[0262] 可以以分批工艺、补料分批工艺或连续工艺或它们的组合来培养重组微生物。通常，在发酵罐中在存在合适营养源(如碳源)的情况下于确定温度下培养重组微生物所需的一段时间以产生足够的酶来将高法呢醇生物转化为降龙涎醚以及产生所需量的降龙涎醚，包括(-)-降龙涎醚。

[0263] 可以以任何合适的方式，例如通过分批培养或补料分批培养来培养重组宿主细胞。

[0264] 如本文所用，术语“分批培养”是其中在培养期间既不添加也不移出培养基的培养方法。

[0265] 如本文所用，术语“补料分批”意指其中在培养期间添加培养基但不移出培养基的培养方法。

[0266] 本公开的一个实施方案提供了在细胞系统中制备降龙涎醚的方法，该方法包括在细胞系统中在合适的条件下表达WT SHC或SHC/HAC衍生物，将高法呢醇供料给该细胞系统，

利用用该细胞系统产生的SHC或SHC/HAC衍生物使高法呢醇转化为降龙涎醚,从细胞系统收集降龙涎醚以及任选从该系统分离(-)-降龙涎醚材料。其它核苷酸序列的表达可用来增强该方法。该生物转化方法可包括在细胞系统中额外表达其它核苷酸序列。其它核苷酸序列的表达可增强制备(-)-降龙涎醚的生物转化途径。

[0267] 本公开的另一个实施方案是制备(-)-降龙涎醚的生物转化方法,该方法包括培养包含WT SHC/HAC或SHC/HAC衍生物基因的宿主细胞,在该宿主细胞中产生WT SHC/HAC或SHC/HAC衍生物,将高法呢醇(如EEH) 供料给该宿主细胞,在适于促进高法呢醇转化为降龙涎醚的pH、温度和增溶剂条件下温育所述宿主细胞,以及收集(-)-降龙涎醚。在宿主细胞中产生WT SHC/HAC和/或SHC/HAC衍生物提供了当在合适反应条件下将高法呢醇添加至所述宿主细胞时制备(-)-降龙涎醚的方法。所实现的转化可通过将更多生物催化剂和SDS添加至反应混合物来增强。

[0268] 可将重组宿主细胞微生物以多种方式培养以便提供合适量的表达WT SHC或SHC/HAC衍生物的细胞供后续的生物转化步骤。因为适用于该生物转化步骤的微生物多种多样(如酵母、细菌和真菌),所以当然要根据每种物种的具体要求调整培养条件,这些条件是众所周知的并且有文献记载的。可将培养重组宿主细胞微生物的细胞的本领域已知方法的任何一种用于制备可用于本公开的后续生物转化步骤的细胞。通常,将细胞培养至特定密度(可作为光密度(OD) 测量)以产生足够的生物物质供所述生物转化反应。

[0269] 所选择的培养条件不仅影响所获得的细胞(生物物质) 的量,而且培养条件的质量还影响生物物质变成生物催化剂的情况。表达WT SHC或SHC/HAC衍生物基因并产生WT SHC或SHC/HAC衍生酶的重组宿主细胞微生物称为生物催化剂,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WT SHC或SHC/HAC衍生物的重组完整细胞,或者其可以为悬浮液或固定化的形式。在其它实施方案中,生物催化剂是从产生WT SHC或SHC/HAC衍生物的重组完整细胞制备的膜级分或液体级分(如例如上文所引用的Seitz等人2012中所公开的)。

[0270] 产生WT SHC或SHC/HAC衍生物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(则其用于一锅法反应)。产生WT SHC或SHC/HAC衍生物的重组全细胞可包括完整的重组全细胞和/或细胞碎片。无论哪种方式,WT SHC或SHC/HAC衍生物都以某种方式与膜(诸如细胞膜) 结合,以便接受底物(如高法呢醇) 和/或与底物(如高法呢醇) 反应,该膜(诸如细胞膜) 可以是全细胞(如重组全细胞) 的一部分或构成全细胞(如重组全细胞)。WT SHC或SHC/HAC衍生物还是以是固定化形式(如与酶载体关联),该形式允许WT SHC或SHC/HAC衍生物与底物(如高法呢醇) 反应。WT SHC或SHC/HAC衍生物也可以以可溶形式使用。

[0271] 在一个实施方案中,以足够的量产生生物催化剂(生成足够的生物物质),收获并洗涤(任选保存(如冷冻或冻干保存)),然后用于生物转化步骤。

[0272] 在一另外的实施方案中,以足够的量产生细胞(生成足够的生物催化剂),然后调节反应条件而无需收获和洗涤该生物催化剂供生物转化反应。该一步法(或“一锅法”) 是有利的,因为其简化了工艺同时降低了成本。用于培养细胞的培养基也适合用于该生物转化反应,前提条件是调节反应条件以有利于该生物转化反应。

[0273] 用于培养细胞的最佳pH在6.0-7.0的范围内。用于生物转化反应的最佳pH取决于生物转化反应中所用的SHC/HAC酶的类型。采用技术人员众所周知的技术调节pH。

[0274] 如实施例9中所展示的,将“一锅法”用于将高法呢醇生物转化为(-)-降龙涎醚,转化率为100%。如实施例18中所展示的,将“一锅法”用于将高法呢醇生物转化为(-)-降龙涎醚,转化率为99%。

[0275] 如本文所用,本文中任何提及“高法呢醇底物转化为(-)-降龙涎醚的转化率为99%/100%”,均是指“使用WT SHC/HAC或SHC/HAC衍生酶使能够转化的高法呢醇异构体(即EEH) 99%/100%转化为(-)-降龙涎醚”。

[0276] 虽然在本公开中术语“混合物”或“反应混合物”可与术语“培养基”互换使用(尤其是在涉及“一锅法”反应时),应该指出的是,培养细胞以产生足够的生物质需要细胞培养基/发酵培养基,但是对于生物转化步骤不需要培养基,因为在合适的pH下反应缓冲液将足够。

[0277] 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高法呢醇原料向(-)-降龙涎醚的转化。反应混合物的pH对于SHC/HAC衍生酶可以在4-8、优选5至6.5、更优选4.8-6.0的范围内,对于WT SHC酶可以在约pH 5.0至约pH 7.0的范围内,并且可通过添加缓冲液至反应混合物来维持。用于该目的的示例性缓冲液是柠檬酸缓冲液。优选的温度介于约15℃至约45℃之间、优选介于约20℃至约40℃之间,但对于嗜热生物体、尤其是如果使用来自嗜热微生物的野生型酶(如WT SHC/HAC)的话,温度可以更高,可高达55℃。在生物转化过程期间温度可保持恒定或者可改变。

[0278] 本申请人已展示在生物转化反应物中包括增溶剂(如表面活性剂、去垢剂、溶解性增强剂、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”意指降低两种液体之间或液体与固体之间的表面张力(或界面张力)的组分。表面活性剂可充当去垢剂、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、牛黄脱氧胆酸盐、牛黄脱氧胆酸钠、十二烷基硫酸钠(SDS)和/或月桂基硫酸钠(SLS)。

[0279] 虽然Triton X-100可用于部分纯化WT SHC/HAC或SHC/HAC衍生酶(以可溶形式或膜级分/悬浮液形式),其还可用于生物转化反应(参见例如Seitz (2012博士论文,如上文所引用的)中的公开内容,以及Neumann和Simon (1986,如上文所引用的)中的公开内容,以及JP2009060799)

[0280] 然而,出人意料的是,如实施例14所展示的,申请人从众多其它没那么有用的增溶剂中选择了SDS并将其鉴定为特别有用的增溶剂。具体地讲,比起例如Triton X-100,依据高法呢醇转化为(-)-降龙涎醚的生物转化反应的反应速率和产率(当以4g/l和125g/l使用EEH时),申请人将SDS鉴定为明显更好的增溶剂。如实施例12中的对比数据所展示的,申请人已证明,对于至少一种SHC/HAC衍生酶,在反应中使用Triton X-100(以约0.005%至0.48%的浓度范围)时高法呢醇转化为(-)-降龙涎醚的最大生物转化活性仅为使用SDS(以约0.07%的浓度)时所获得的活性的约20%。

[0281] 尽管不希望受理论束缚,将SDS用于重组微生物宿主细胞可能是有利的,因为SDS可能会有利地与宿主细胞膜反应以使SHC酶(其为结合膜的酶)更容易被高法呢醇底物触及。此外,在反应混合物中包含适当水平的SDS可改善乳液(水中的高法呢醇)的性质和/或改善高法呢醇底物向宿主细胞内的SHC酶的接近而同时防止破坏(如SHC (WT或SHC/HAC衍生生物)酶的变性)。

[0282] 生物转化反应中所用的增溶剂(如SDS)的浓度受生物质的量以及底物(EEH)浓度的影响。也就是说,在增溶剂(如SDS)浓度、生物质的量以及底物(EEH)浓度之间存在一定程度的相互依赖性。举个例子,当高法呢醇底物浓度增加时,发生有效生物转化反应需要足够量的生物催化剂和增溶剂(如SDS)。例如,如果增溶剂(如SDS)浓度过低,则可能会观察到次优的高法呢醇转化。另一方面,例如,如果增溶剂(如SDS)浓度过高,则存在生物催化剂因完整微生物细胞的破坏和/或SHC/HAC酶的变性/失活而受到影响的危险。

[0283] 根据生物质的量以及底物(EEH)浓度选择合适的SDS浓度在技术人员知识内。举个例子,有预测模型供技术人员使用来确定合适的SDS、底物(EEH)和生物质的浓度。另外举个例子,实施例3证明,当使用4g/l的EEH以及达到OD=10.0(650nm)的生物催化剂时,0.010-0.075%范围内的SDS是恰当的。实施例7证明,当125g/l的EEH与2倍湿重的生物质一起使用时,调节SDS浓度(1.55%)是恰当的。然而,使用不同的SDS/细胞比值研究EEH转化为(-)-降龙涎醚的转化百分比表明,正确选择生物催化剂、高法呢醇底物和增溶剂(如SDS)的比率有利于开发稳健的生物转化反应系统,该系统展示出一定程度的对一系列SDS浓度(参见例如图17)和pH范围(参见实施例15、图18)的耐受性。

[0284] WT SHC酶(如AacSHC)的生物转化反应的温度为约45-60℃、优选55℃。

[0285] WT SHC酶(如AacSHC)的生物转化反应的pH范围为约5.0至7.0、更优选约5.6至约6.2、甚至更优选约6.0。

[0286] SHC/HAC衍生酶的生物转化反应的温度为约34℃至约50℃、优选约35℃。

[0287] SHC/HAC衍生酶的生物转化反应的pH为约4.8-6.4、优选约5.2-6.0。

[0288] 优选地,用于生物转化反应中的增溶剂是SDS。

[0289] 当使用约4g/l的EEH时,WT SHC酶(如AacSHC)的生物转化反应中所用的SDS浓度在约0.010-0.075%的范围内,优选为约0.030%。

[0290] 当使用约4g/l的EEH时,SHC/HAC衍生酶的生物转化反应中所用的SDS浓度在约0.0025-0.090%的范围内,优选为约0.050%。

[0291] 当反应物中加载有EEH浓度为约4g/l的EEH的高法呢醇时,将生物催化剂加载到反应物中至OD为10.0(650nm)。

[0292] 当生物催化剂与EEH高法呢醇的比率为约2:1时,[SDS]/[细胞]比率在约10:1-20:1的范围内、优选约15:1-18:1、优选约16:1。

[0293] 当高法呢醇浓度约为125g/l的EEH并且生物催化剂浓度为250g/l(对应于OD约为175(650nm))时,SHC变体酶的生物转化反应中的SDS浓度在约1-2%的范围内、优选在约1.4-1.7%的范围内、甚至更优选约1.5%。

[0294] 生物催化剂与EEH高法呢醇底物的比率在约0.5:1-2:1的范围内,在一些实施方案中为2:1、优选为约1:1或0.5:1。

[0295] 在一些实施方案中,利用添加了高法呢醇底物的生物催化剂制备降龙涎醚。通过采用已知的手段(如蠕动泵、输注注射器等)给料来添加底物是可能的。高法呢醇是油溶性化合物并以油形式提供。鉴于生物催化剂(微生物细胞,诸如完整的重组全细胞和/或细胞碎片和/或固定化酶)存在于水相中,当将高法呢醇添加至生物转化反应混合物时,该生物转化反应可视为三相系统(包含水相、固相和油相)。即使当存在SDS时也是如此。通过澄清,当可溶性WT SHC或SHC/HAC衍生物用作生物催化剂时,这被视为两相系统。

[0296] 所存在的高法呢醇异构体的数量可影响反应的速率。如实施例11所展示的,SHC/HAC衍生酶能够从高法呢醇异构体的复杂混合物(如EE:EZ:ZE:ZZ)将E,E-高法呢醇生物转化为(-)-降龙涎醚。然而,通常观察到较低的转化率,这与以下观点一致:除了EEH之外的高法呢醇异构体可与EEH竞争对SHC/HAC衍生酶的靠近,从而可能充当EEH向(-)-降龙涎醚转化的竞争性抑制剂和/或还充当备选的底物。

[0297] 因此,优选地,高法呢醇底物包含2-4种异构体、优选两种异构体的立体异构体混合物。

[0298] 因此,优选地,高法呢醇底物由2-4种异构体、优选两种异构体的立体异构体混合物组成或基本上由其组成。

[0299] 优选地,高法呢醇底物包含EE:EZ立体异构体混合物。

[0300] 优选地,高法呢醇底物由EE:EZ立体异构体混合物组成或基本上由其组成。

[0301] 如实施例9所展示的,在22.5天内进行的“一锅法”发酵和生物转化反应中,观察到重量比为87:13的EE:EZ 100%转化。在该时间段内约10g EEH发生转化。

[0302] 如实施例7中详细描述,在一个优选的实施方案中,将发酵罐用于将表达SHC/HAC衍生基因并且产生活性SHC/HAC衍生酶的重组宿主细胞培养至适合在同一发酵容器中用作生物催化剂的足够生物物质浓度,该生物催化剂用于将高法呢醇源转化为与副产物(II)、(IV)和/或(III)中的一者或多者混合的(-)-降龙涎醚,如例如图12中所公开的。可通过以下方式分离(-)-降龙涎醚:汽提/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保持在水相中的生物催化剂分离),然后后续蒸发掉溶剂,以获得粗制反应产物,这可通过气相色谱(GC)分析来确定。汽提/蒸馏和有机溶剂萃取方法是本领域技术人员已知的。

[0303] 举个例子,可使用有机溶剂诸如非水混溶性溶剂(例如甲苯)从整个反应混合物萃取所得的(-)-降龙涎醚。备选地,可使用水混溶性溶剂(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的)萃取所产生的(-)-降龙涎醚。举另一个例子,(-)-降龙涎醚作为晶体或以无定形式存在于固相中,可通过过滤与其余的固相(细胞材料或其碎片)和液相分离。再举个例子,在(-)-降龙涎醚的熔点(约75℃)以上的温度下,(-)-降龙涎醚可在水相顶部形成油层,可移出该油层并收集。为了确保在移出油层后完全回收(-)-降龙涎醚,可添加有机溶剂至含有生物物质的水相以便萃取生物物质中或生物物质上或生物物质周围所含的任何残留的(-)-降龙涎醚。可将该有机层与该油层合并,然后进一步处理以分离并纯化(-)-降龙涎醚。

[0304] 可使(-)-降龙涎醚进一步选择性结晶以从最终的(-)-降龙涎醚产物移除副产物(II)、(IV)和(III)以及任何未反应的高法呢醇底物。术语“选择性结晶”是指这样的处理步骤:通过该处理步骤引起(-)-降龙涎醚从溶剂结晶,而化合物(II)、(III)和(IV)保持溶解在结晶溶剂中,其程度使得分离的晶体材料仅含有(-)-降龙涎醚产品,或者如果其含有任何其它化合物(II)、(III)或(IV),则它们仅以嗅觉上可接受的量存在。

[0305] 选择性结晶步骤可使用水混溶性溶剂诸如乙醇等。使用水中的10%乙醇萃取物或通过测试晶体材料来测定最终(-)-降龙涎醚产品的嗅觉纯度。相对于(-)-降龙涎醚产品的市售参照物测试最终(-)-降龙涎醚产品的嗅觉纯度、品质及其感官特性(sensory profile)。还在应用研究中由专家测试该(-)-降龙涎醚材料以确定该材料是否在其感官特

性方面符合规格。(-)-降龙涎醚的各种应用包括但不限于精品香料或消费品诸如织物护理产品、化妆用具、美容护理产品和清洁产品,包括其中商业使用目前可用的降龙涎醚成分的基本上所有的产品,包括但不限于:Ambrox产品(芬美意公司(Firmenich))、Ambroxan产品(汉高公司(Henkel))、Ambrofix产品(奇华顿公司(Givaudan))、Amberlyn产品(奎斯特公司(Quest))、CetaloxLaevo产品(芬美意公司)、Ambermor产品(埃姆公司(Aromor))和Norambrenolide Ether产品(太平洋公司(Pacific))。

[0306] (-)-降龙涎醚的选择性结晶可受未反应的高法呢醇底物还有(-)-降龙涎醚与其它可检测副产物(II)、(III)和/或(IV)之间比率的影响。即使仅获得10%的高法呢醇底物向(-)-降龙涎醚的转化(如实施例7中使用WT SHC/HAC酶所展示的),(-)-降龙涎醚的选择性结晶仍是可能的。

[0307] 适用于(-)-降龙涎醚的萃取和/或选择性结晶的合适水混溶性或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个碳原子的脂族烃,诸如戊烷、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,诸如二氯甲烷、氯仿、四氯化碳、二氯乙烷或四氯乙烷;芳族烃,诸如苯、甲苯、二甲苯、氯苯或二氯苯;脂族的无环或环状的醚或醇,优选具有4至8个碳原子的那些,诸如乙醇、异丙醇、二乙醚、甲基叔丁基醚、乙基叔丁基醚、二丙醚、二异丙醚、二丁醚、四氢呋喃;或酯,诸如乙酸乙酯或乙酸正丁酯;或酮,诸如甲基异丁基酮;或二噁烷;或这些物质的混合物。尤其优选使用的溶剂是上述的庚烷、甲基叔丁基醚(也称为MTB、叔丁基甲醚、叔丁基甲基醚和tBME)、二异丙醚、四氢呋喃、乙酸乙酯和/或它们的混合物。

[0308] 优选地,将水混溶性溶剂诸如乙醇用于从反应混合物的固相萃取(-)-降龙涎醚。乙醇的使用是有利的,因为其易于处理,其是无毒的并且其是环境友好的。

[0309] 本文所用的术语“分离的”是指已从与之相伴的组分分离或纯化的生物转化产物诸如(-)-降龙涎醚。细胞系统中产生的与其在自然界中最初起源的来源不同的实体是“分离的”,因为其将必然不含天然状态下与之相伴的组分。分离或纯化的程度可通过任何恰当的方法,如气相色谱(GS)、HPLC或NMR分析来测量。

[0310] 在一些实施方案中,将最终产物((-)-降龙涎醚)分离并纯化至均质(如纯度为至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%或89.5%或者纯度为90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或99.5%)。

[0311] 理想的是,所产生的(-)-降龙涎醚的量可以为约1mg/l至约20,000mg/l(20g/l)或更高,诸如约20g/l至约200g/l或100-200g/l,优选约125g/l或150g/l或约188g/l。

[0312] 如实施例7所展示的,使用产生SHC/HAC衍生酶的重组大肠杆菌宿主细胞,经过约2天在生物转化反应中产生了至少125g/l(-)-降龙涎醚。

[0313] 如实施例19所展示的,倘若实现充分混合,则以188g/l的EEH或更高浓度运行生物转化是可能的,因为搅拌效率看起来是该系统的唯一限制。此外,具有改善的活性的生物催化剂(如,就具有进一步改善的活性的SHC变体而言,或就SHC酶产量增加而言)可在使用较少生物物质的情况下改善或保持生产力,就混合效率而言使用较少生物物质是有利的。

[0314] 例如,产生了约1至约100mg/l、约30至约100mg/l、约50至约200mg/l、约100至约500mg/l、约100至约1,000mg/l、约250至约5,000mg/l、约1,000(1g/l)至约15,000mg/l(15g/l)、或约2,000(2g/l)至约10,000mg/l(10g/l)或约2,000(2g/l)至约25,000mg/l

(25g/l) 或约2,000 (2g/l) 至约25,000mg/l (25g/l)、26,000mg/l (26g/l)、27,000mg/l (27g/l)、28,000mg/l (28g/l)、29,000mg/l (29g/l)、30,000mg/l (30g/l)、40g/l、50g/l、60g/l、70g/l、80g/l、90g/l、100g/l、110g/l、120g/l、125g/l、130g/l、140g/l、150g/l、160g/l、170g/l、180g/l、190g/l或200g/l或300g/l或400g/l或500g/l的(-)-降龙涎醚。

[0315] 优选地,在48小时至72小时的时间周期内产生了浓度为至少100g/l的(-)-降龙涎醚。

[0316] 优选地,在约48小时至72小时的时间周期内产生了浓度为约150g/l的(-)-降龙涎醚。

[0317] 优选地,在约48小时至72小时的时间周期内产生了浓度为约200g/l的(-)-降龙涎醚。

[0318] 优选地,在约48小时至72小时的时间周期内产生了浓度为约250g/l的(-)-降龙涎醚。

[0319] 技术人员将理解,更高的累积产量滴度可通过实施连续工艺(诸如产物移出、底物供料以及生物质的添加或(部分)取代)来实现。

[0320] 优选地,在包含WT SHC/HAC或SHC/HAC衍生生物的重组宿主细胞的存在下EEH向(-)-降龙涎醚的生物转化产生的降龙涎醚产率为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(单位为摩尔百分比,并且是基于所采用的EEH的摩尔数);特别优选地,该产率介于5至100摩尔%、10至100摩尔%、20至100摩尔%、25至100摩尔%、30至100摩尔%、35至100摩尔%之间,尤其是介于40至100摩尔%、45至100摩尔%、50至100摩尔%、60至100摩尔%、70至100摩尔%之间。

[0321] SHC/HAC酶的活性通过反应速率定义(产物的量/(产物的量+剩余原料的量)) $\times$ 100,单位为摩尔百分比。优选地,在WT SHC或SHC/HAC衍生酶的存在下EEH向(-)-降龙涎醚的生物转化产生的(-)-降龙涎醚产率为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(单位为摩尔百分比,并且是基于所采用的EEH的摩尔数);特别优选地,该产率介于5至100摩尔%、10至100摩尔%、20至100摩尔%、25至100摩尔%、30至100摩尔%、35至100摩尔%之间,尤其是介于40至100摩尔%、45至100摩尔%、50至100摩尔%、60至100摩尔%、70至100摩尔%之间。

[0322] 在本发明的一个优选的实施方案中,在例如4、6、8、10、12、16、20、24、36或48个小时的确定时间段内测定产率和/或反应速率,在该时间段内包含编码根据本公开的WTSHC或SHC/HAC衍生酶的核苷酸序列的重组宿主细胞将EEH转化为(-)-降龙涎醚。在一个另外的变型形式中,在例如25℃、30℃、40℃、50℃或60℃的精确确定的条件下进行反应。具体地讲,通过在35℃下进行由根据本发明的SHC/HAC衍生酶催化的EEH转化成(-)-降龙涎醚的反应持续24-72小时来测定产率和/或反应速率。

[0323] 在本发明的另一个实施方案中,包含编码SHC/HAC衍生物的核苷酸序列的重组宿主细胞表征为,在相同条件下与WT SHC或SHC/HAC衍生酶相比,在高法呢醇生成(-)-降龙涎醚的反应中,其显示出2倍、3倍、4倍、5倍、6倍、7倍、8倍、9倍、10倍、11倍、12倍、13倍、14倍、15倍、16倍、17倍、18倍、19倍、20倍、21倍、22倍、23倍、24倍、25倍、26倍、27倍、28倍、29倍、30倍、31倍、32倍、33倍、34倍、35倍、36倍、37倍、38倍、39倍、40倍、41倍、42倍、43倍、44倍、45倍、46倍、47倍、48倍、49倍、50倍、51倍、52倍、53倍、54倍、55倍、56倍、57倍、58倍、59倍、60倍、61倍、62倍、63倍、64倍、65倍、66倍、67倍、68倍、69倍、70倍、71倍、72倍、73倍、74倍、75倍、76倍、77倍、78倍、79倍、80倍、81倍、82倍、83倍、84倍、85倍、86倍、87倍、88倍、89倍、90倍、91倍、92倍、93倍、94倍、95倍、96倍、97倍、98倍、99倍、100倍、200倍、500倍、1000倍或更高的产率和/或反应速率。这里,术语“条件”涉及反应条件,诸如底物浓度、酶浓度、反应期和/或温度。

[0324] 成功开发一种在包含编码WT/参照SHC或SHC/HAC衍生物的核苷酸序列的大肠杆菌重组菌株中由高法呢醇制备(-)-降龙涎醚的生物转化工艺可提供低成本且在工业上经济的(-)-降龙涎醚生产工艺。

[0325] 如实施例7所展示的,在与经优化的SHC/HAC衍生物温育48小时后,本公开提供了E,E-高法呢醇(125g/l)向(-)-降龙涎醚的100%转化,与WT AacSHC酶相比,当使用AacSHC衍生物时产率提高8倍(见图11)。

[0326] WT参照SHC/HAC或本文所述的SHC/HAC衍生物多肽的功能同源物也适合用于在重组宿主中制备降龙涎醚。因而,重组宿主可包含一种或多种编码上述多肽的功能同源物的异源核酸和/或编码如上所述的SHC/HAC衍生酶的异源核酸。

[0327] 功能同系物是与参照多肽具有序列相似性,以及执行参照多肽的一种或多种生化功能或生理功能的多肽。功能同系物和参照多肽可能是天然存在的多肽,并且序列相似性可能是由于趋同进化或趋异进化事件造成。照此,功能同系物有时候在文献中称为同系物,或直系同源物或旁系同源物。天然存在的功能同系物的变体,诸如由野生型编码序列的突变体编码的多肽,可能它们自身是功能同系物。功能同系物还可以经由多肽的编码序列的定点诱变,或通过组合来自不同天然存在的多肽的编码序列的结构域(“结构域交换”)来生成。用于修饰编码本文所述的功能同系物的基因的技术是已知的,并且尤其包括定向进化技术、定点诱变技术和随机诱变技术,并且可用于增加多肽的比活性、改变底物特异性、改变表达水平、改变亚细胞定位或以所需方式修饰多肽:多肽相互作用。这种经修饰的多肽被视为功能同系物。术语“功能同系物”有时候适用于编码在功能上同源的多肽的核酸。

[0328] 功能同系物可通过对核苷酸和多肽序列比对进行分析而鉴定。例如,对核苷酸或多肽序列的数据库进行查询可鉴定编码SHC衍生物多肽等的核酸序列的同系物。

[0329] 还可将杂交用于鉴定功能同系物和/或用作两条核酸序列之间的同源性的量度。编码任何本文所公开的蛋白质的核酸序列或其部分可根据标准的杂交技术用作杂交探针。探针杂交至来自测试来源(如哺乳动物细胞)的DNA或RNA是该测试来源中存在相关DNA或RNA的指示。杂交条件是本领域技术人员已知的并且可见于Current Protocols in Molecular Biology, John Wiley&Sons, N.Y., 6.3.1-6.3.6, 1991。中等杂交条件定义为相当于在30℃下于2x氯化钠/柠檬酸钠(SSC)中杂交,然后在50℃下于1x SSC、0.1%的SDS中洗涤。高严格条件定义为相当于在45℃下于6x氯化钠/柠檬酸钠(SSC)中杂交,然后在65℃

下于0.2x SSC、0.1%的SDS中洗涤。

[0330] 鉴定功能同系物的序列分析也可涉及使用一相关氨基酸序列作为参照序列对非冗余数据库进行BLAST、交互BLAST或PSI-BLAST分析。在某些情况下,氨基酸序列从核苷酸序列推导而来。数据库中具有大于40%序列同一性的那些多肽是用于进一步评价在SHC/HAC生物转化反应中使用的适用性的候选者。氨基酸序列相似性使得能进行保守氨基酸取代,诸如将一个疏水残基取代为另一个疏水残基或者将一个极性残基取代为另一个极性残基。如果需要,可对这些候选者进行人工检测以便缩窄待进一步评价的候选者的数目。通过选择看起来具有例如保守功能结构域的那些候选者来进行人工检测。

[0331] 通常,表现出至少30%氨基酸序列同一性的多肽可用于鉴定保守区。相关多肽的保守区表现出至少30%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%的氨基酸序列同一性。在一些实施方案中,保守区表现出至少70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、至少81%、至少82%、至少83%、至少84%、至少85%、至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、或至少99%的氨基酸序列同一性。可如上文和下文所述测定序列同一性。

[0332] 所产生的WTSHC和/或SHC/HAC衍生物是基于氨基酸SEQ ID No.1或SEQ ID No.2或SEQ ID No.3或SEQ ID No.4或它们的变体、同源物、突变体、衍生物或片段。

[0333] 所产生的SHC是基于与SEQ ID No.1或SEQ ID No.2或SEQ ID No.3或SEQ ID No.4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%同一性的氨基酸序列。

[0334] 此外,所产生的参照SHC是基于由大肠杆菌产生的氨基酸序列。

[0335] 关于基因的核苷酸序列的“同一性百分比(%)”定义为:在比对序列和引入缺口(必要时)以实现最大的序列同一性百分比、并且不将任何保守取代视作序列同一性的一部分之后,候选DNA序列中与所述DNA序列中的核苷酸相同的核苷酸的百分比。为了测定核苷酸序列同一性百分比的比对可以各种方式来实现,这些方式在本领域技术人员的范围内,例如,使用可公开获得的计算机软件。本领域内的技术人员可确定用于测量比对的恰当参数,包括在进行比较的序列的整个长度上实现最大比对所需的任何算法。

[0336] 术语“多肽”和“蛋白质”在本文中可互换使用,意指任何氨基酸的肽键连接的链,而不管长度或翻译后修饰。

[0337] 如本文所用的,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。

[0338] 如本文所用,术语“变体”应该理解为这样一种多肽:与衍生该多肽的多肽相比,在氨基酸序列中有一个或多个改变。衍生出变体的多肽也称为亲本多肽或参照多肽。通常以人工方式构建变体,优选通过基因技术手段。通常,衍生出变体的多肽是野生型蛋白质或野

生型蛋白质结构域。然而,可用于本公开的变体也可以衍生自亲本多肽的同系物、直系同源物或旁系同源物,或者源于人工构建的变体,前提条件是该变体表现出亲本多肽的至少一种生物活性。氨基酸序列中的改变可以是氨基酸交换、插入、缺失、N端截短或C端截短、或这些改变的任何组合,这些改变可在一个或数个位点发生。

[0339] 在优选的实施方案中,可用于本公开的变体在氨基酸序列中表现出总数高达200(高达1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、110、120、130、140、150、160、170、180、190或200)个改变(即交换、插入、缺失、N端截短和/或C端截短)。氨基酸交换可以是保守性的和/或非保守性的。在优选的实施方案中,可用于本公开的变体与衍生该变体的蛋白质或结构域相差最多1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95或100个氨基酸交换,优选保守氨基酸改变。变体可额外地或备选地包含氨基酸缺失,其可以是N端截短、C端截短或内部缺失或这些情况的任何组合。包含N端截短、C端截短和/或内部缺失的这些变体在本申请的上下文中称为“缺失变体”或“片段”。术语“缺失变体”和“片段”在本文中可互换使用。缺失变体可以是天然存在的(如剪接变体)或其可以人工构建,优选通过基因技术手段。通常,衍生缺失变体的蛋白质或蛋白质结构域是野生型蛋白质。然而,本公开的缺失变体也可以衍生自亲本多肽的同系物、直系同源物或旁系同源物,或者源于人工构建的变体,前提条件是该缺失变体表现出亲本多肽的至少一种生物活性。优选地,与亲本多肽相比,缺失变体(或片段)在其N端和/或在其C端和/或在内部具有最多1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95或100个氨基酸的缺失。

[0340] 作为另外一种选择或额外地,本文所用的“变体”可表征为与衍生该变体的亲本多肽具有一定程度的序列同一性。本公开的WT/参照SHC/HAC或SHC/HAC衍生物的变体可以与各自的参照多肽或与各自的参照多核苷酸具有至少40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、至少81%、至少82%、至少83%、至少84%、至少85%、至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、或至少99%的序列同一性。

[0341] 表述“至少30%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、至少81%、至少82%、至少83%、至少84%、至少85%、至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、或至少99%序列同一性”就多肽和多核苷酸序列比较而言在整篇说明书中使用。

[0342] 属于本文所公开的任何酶的多核苷酸或蛋白质可分别根据它们与相关基因或蛋白质的相似性来鉴别。例如,该鉴别可以是基于序列同一性。在某些优选的实施方案中,本公开描述了分离的核酸分子,其与以下核酸分子具有至少30%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、至少81%、至少82%、至少83%、至少84%、

至少85%、至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、或至少99%的同一性：(a) 编码 SEQ ID No.5-163 (参见本文提供的表14-17和表4a) 的多肽的核酸分子，(b) SEQ ID No.6-168、169、170、172、174和176 (参见本文所提供的表14-17和表4a) 的核苷酸序列，以及(c) 包括SEQ ID No.6-168、169、170、172、174和176 (参见本文所提供的表14-17和表4a) 的至少30 (如至少30、40、50、60、80、100、125、150、175、200、250、300、400、500、600、700、800、850、900、950、1000或1010) 个核苷酸的区段的核酸分子。

[0343] 优选地，所考虑的多肽与参照多肽在20、30、40、45、50、60、70、80、90、100或更多个氨基酸的连续链段上表现出所指示的序列同一性。优选地，所考虑的多核苷酸与参照多核苷酸在60、90、120、135、150、180、210、240、270、300或更多个核苷酸的连续链段上表现出所指示的序列同一性。在其中对两条序列进行比较并且在比较中要计算与其他的序列同一性的参照序列未规定的情况中，如果未另外明确指明的话，则序列同一性应参照待比较的两条序列中的较长者进行计算。如果指明了参照序列，则如果未另外明确指明的话，基于SEQ ID No.1、2、3和/或4指示的参照序列的全长测定序列同一性。

[0344] 例如，由130个氨基酸组成的肽序列与具有631个氨基酸残基的参照SHC的全长氨基酸相比较，可展现出20.6%的最大序列同一性百分比 ($130/631 \times 100$)，而具有300个氨基酸长度的序列可展现出47.5%的最大序列同一性百分比 ($300/631 \times 100$)。核苷酸序列和氨基酸序列的相似性(即序列同一性百分比)可通过序列比对测定。这种比对可用数种本领域已知的算法，优选用Karlin和Altschul (Karlin&Altschul (1993) Proc.Natl.Acad.Sci.USA 90:5873-5877) 的数学算法、用hmmalign (HMMER软件包，<http://hmm.wustl.edu/>) 或用CLUSTAL算法 (Thompson, J.D., Higgins, D.G.&Gibson, T.J. (1994) Nucleic Acids Res.22,4673-80) (在例如<http://www.ebi.ac.uk/Tools/clustalw/>上或在

[0345] <http://www.ebi.ac.uk/Tools/clustalw2/index.html>上或在

[0346] http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_clustalw.html上可供使用) 或

[0347] 用GAP程序 (爱荷华大学的数学算法) (如与WT SHC的序列比对中所用的，如本文表18中所提供的) 或Myers和Miller (1989-Cabios 4:11-17) 的数学算法 (如本文所提供的表19中的WTSHC序列比对中所公开的并且如其中所用的) 来进行。

[0348] 所用的优选参数是如它们在<http://www.ebi.ac.uk/Tools/clustalw/>或<http://www.ebi.ac.uk/Tools/clustalw2/index.html>上设定的默认参数。

[0349] 序列同一性 (序列匹配) 等级可使用例如BLAST、BLAT或BlastZ (或BlastX) 计算。类似的算法并入了Altschul等人 (1990) J.Mol.Biol.215,403-410的BLASTN和BLASTP程序中。BLAST多核苷酸搜索用BLASTN程序，得分 (score) = 100，字长 (word length) = 12来进行，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。

[0350] BLAST蛋白质搜索用BLASTP程序，得分 = 50，字长 = 3来进行，以获得与SrK0多肽同源的氨基酸序列。为了获得带缺口的比对以达到比较目的，如Altschul等人 (1997) Nucleic Acids Res.25,3389-3402中所述使用Gapped BLAST。当使用BLAST和Gapped BLAST程序时，使用各程序的默认参数。序列匹配分析可通过确立的同源性作图技术如Shuffle-LAGAN

(BrudnoM.,Bioinformatics 2003b,19,增刊1:154-162)或马尔可夫随机场来补充。当在本申请中提及序列同一性百分比时,如果未另外明确指明,则这些百分比是相对于较长序列的全长计算的。

[0351] 感官方面

[0352] 根据本公开的高法呢醇向(-)-降龙涎醚的生物转化产生(-)-降龙涎醚作为优势化合物,但是也可能产生除了(-)-降龙涎醚之外的化合物,这些化合物可能会或可能不会给该生物转化混合物赋予令人愉悦的嗅觉香型,从而可能以正面或负面的方式促成(-)-降龙涎醚终产物的感官品质。因此,感官分析使用由经训练的专家(香水师)进行的已确立的感官测试进行,从而这种测试可帮助确定化学相关的靶产物相对于参照产品是否也是嗅觉相关的终产物。如实施例22中的感官分析所展示的,从(-)-降龙涎醚移除一种或多种副产物化合物可改善剩余化合物((-)-降龙涎醚)的气味,即使所移除的化合物实际上本身是无臭化合物。也就是说,在缺少化合物II、III和IV的情况下观察到(-)-降龙涎醚气味增强。

[0353] 本发明的方面

[0354] 1.一种制备(-)-降龙涎醚或包含(-)-降龙涎醚的混合物的方法,其中(3E,7E)-高法呢醇(EEH)或包含EEH的立体异构体混合物被酶促转化为(-)-降龙涎醚或包含(-)-降龙涎醚的混合物,其中所述酶促转化使用SHC/HAC酶在适于产生(-)-降龙涎醚的反应条件下进行,并且其中包含EEH的立体异构体混合物基本上由选自以下的高法呢醇异构体组成:[(3E,7E)和[(3Z,7E)]和/或[(3E,7E)和(3E,7Z)]和/或[(3Z,7E)、(3E,7E)和(3E,7Z)],它们也分别被命名为[EE:EZ]、[EE:ZE]和[EE:EZ:ZE]。

[0355] 2.一种制备(-)-降龙涎醚或包含(-)-降龙涎醚的混合物的方法,其中(3E,7E)-高法呢醇(EEH)或包含EEH的立体异构体混合物经酶促转化而得到(-)-降龙涎醚或包含(-)-降龙涎醚的混合物,其中所述酶促转化使用SHC/HAC酶在适于产生(-)-降龙涎醚的反应条件下进行,并且其中如果所述反应在存在增溶剂的情况下进行,则不将Triton X-100或牛黄脱氧胆酸盐与野生型SHC/HAC酶结合使用。

[0356] 3.一种制备(-)-降龙涎醚或包含(-)-降龙涎醚的混合物的方法,其中(3E,7E)-高法呢醇(EEH)或包含EEH的立体异构体混合物被酶促转化为(-)-降龙涎醚或包含(-)-降龙涎醚的混合物,其中所述酶促转化使用SHC/HAC酶在适于产生(-)-降龙涎醚的反应条件下进行,并且其中包含EEH的立体异构体混合物基本上由选自以下的高法呢醇异构体组成:[(3E,7E)和[(3Z,7E)]和/或[(3E,7E)和(3E,7Z)]和/或[(3Z,7E)、(3E,7E)和(3E,7Z)],它们也分别被命名为[EE:EZ]、[EE:ZE]和[EE:EZ:ZE],并且其中所述反应在包含水相、固相和油相的三相系统中发生。

[0357] 4.根据段落1或段落2或段落3的方法,其中所述方法使用选自SEQ ID No.1、SEQ ID No.2、SEQ ID No.3、SEQ ID No.4的SHC/HAC酶多肽序列或选自表1、表5、表2、表6、表3、表7、表4、表8或表13、表14,或选自SEQ ID No.5、7、9、11、13、15、17、19、21、23、25、27、29、31、33、35、37、39、171、173、175、177和/或178的SHC/HAC衍生物或与SEQ ID No.1、SEQ ID No.2、SEQ ID No.3或SEQ ID No.4具有至少30%同一性、至少40%同一性、至少50%同一性、或至少60%同一性、或至少70%同一性、或至少80%同一性、或至少90%同一性、或至少95%同一性、或至少96%同一性、或至少97%同一性、或至少98%同一性、或至少99%同一性的序列进行。

- [0358] 5.段落1-4中任一项的方法,其中所述方法使用产生SHC/HAC酶的重组宿主细胞。
- [0359] 6.根据段落4或段落5的方法,其中所述编码SHC/HAC酶的核苷酸序列选自SEQ ID No.165、166、167、168、169或SEQ ID No.6、8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、38、40和/或170、172、174和/或176。
- [0360] 7.根据段落1-6中任一项的方法,其中高法呢醇向(-)-降龙涎醚的转化在约4-8的pH下于30℃至60℃的温度下发生。
- [0361] 8.根据段落1-7中任一项的方法,其中高法呢醇向(-)-降龙涎醚的转化使用如表24或表24a中所列的用于野生型SHC/HAC或SHC/HAC衍生酶的反应条件中的一者或多者,优选在5.0至6.2的pH范围下、优选在35℃的温度下进行。
- [0362] 9.根据段落3-8中任一项的方法,其中当生物催化剂与EEH的比率约为2:1时,SDS/细胞比率在10:1至20:1的范围内,优选为16:1。
- [0363] 10.根据段落3-9中任一项的方法,其中生物催化剂与高法呢醇的重量比在约0.5-2:1的范围内,优选约1:1或0.5:1。
- [0364] 11.根据段落3-10中任一项的方法,其中细胞培养步骤和生物转化反应步骤在同一反应容器中进行。
- [0365] 12.根据段落2的方法,其中所述高法呢醇底物包含一种或多种高法呢醇立体异构体。
- [0366] 13.根据段落12的方法,其中所述高法呢醇底物包含两种高法呢醇立体异构体或基本上由两种高法呢醇立体异构体组成。
- [0367] 14.根据段落13的方法,其中所述高法呢醇底物包含EE:EZ立体异构体或基本上由EE:EZ立体异构体组成。
- [0368] 15.根据段落14中任一项的方法,其中所述高法呢醇包含EE:EZ立体异构体混合物或基本上由EE:EZ立体异构体混合物组成,所述EE:EZ立体异构体混合物的重量比选自:100:00;99:01;98:02;97:03;96:04;95:05;94:06;93:07;92:08;91:09;90:10;89:11;88:12;87:13;86:14;85:15;84:16;83:17;82:18;81:19;80:20;79:21;78:22;77:23;76:24;75:25;74:26;73:27;72:28;71:29 70:30;69:31;68:32;67:33;66:34;65:35;64:36;63:37;62:38;61:39;以及60:40。
- [0369] 16.根据段落15的方法,其中所述高法呢醇包含EE:EZ立体异构体混合物或基本上由EE:EZ立体异构体混合物组成,所述EE:EZ立体异构体混合物的重量比选自:EE:EZ 92:8;EE:EZ 90:10;EE:EZ 80:20;EE:EZ 86:14;EE:EZ 70:30;EE:EZ 69:31;以及EE:EZ 66:34。
- [0370] 17.根据段落15或段落16的方法,其中所述高法呢醇包含EE:EZ立体异构体混合物或基本上由EE:EZ立体异构体混合物组成,所述EE:EZ立体异构体混合物的重量比为80:20。
- [0371] 18.根据段落1-17中任一项的方法,其中(-)-降龙涎醚以与副产物(II)、(IV)和/或(III)中的至少一种或多种的混合物产生。
- [0372] 19.根据段落1-18中任一项的方法,其中使用有机溶剂或汽提/蒸馏步骤或过滤从所述生物转化反应混合物分离(-)-降龙涎醚。
- [0373] 20.根据段落1-19中任一项的方法,其中使用有机溶剂或汽提/蒸馏步骤从所述生物转化反应混合物的固相分离(-)-降龙涎醚。
- [0374] 21.根据段落19或段落20的方法,其中使用有机溶剂从所述反应混合物分离(-)-

降龙涎醚。

[0375] 22. 根据段落21的方法, 其中使用乙醇或甲苯从所述反应混合物分离(-)-降龙涎醚。

[0376] 23. 根据段落19-22中任一项的方法, 其中使用有机溶剂使(-)-降龙涎醚选择性结晶。

[0377] 24. 根据段落23的方法, 其中所述(-)-降龙涎醚基本上不含副产物(II)、(IV)和/或(III)。

[0378] 25. 根据段落1-24中任一项的方法, 其中产生了浓度范围为约125-200g/l的(-)-降龙涎醚。

[0379] 26. 通过根据段落1-25中任一项的方法能够获得的(-)-降龙涎醚, 其中所述(-)-降龙涎醚具有约0.1至约0.5ng/l的嗅觉阈值。

[0380] 27. 段落26的(-)-降龙涎醚, 其为固体形式, 优选无定形形式或晶体形式。

[0381] 28. 一种制备含有(-)-降龙涎醚的产品的的方法, 包括将段落26或段落27中任一项的(-)-降龙涎醚掺入所述产品中。

[0382] 29. 段落28的方法, 其中所述产品是香料产品、化妆品、清洁产品、洗涤剂产品或皂产品。

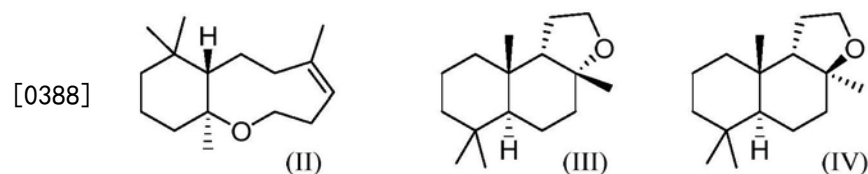
[0383] 30. 一种香料或化妆品或消费者护理产品, 其中包含段落26或段落27中任一项的(-)-降龙涎醚。

[0384] 31. 一种香料或化妆品或消费者护理组合物, 其中包含段落26或段落27的(-)-降龙涎醚和一种或多种另外的组分。

[0385] 32. 段落26或段落27的(-)-降龙涎醚的用途, 其用作香料或化妆品或消费品诸如织物护理、化妆用具、美容护理和/或清洁产品的一部分。

[0386] 33. 用于在香料组合物中或为香料组合物增加、增强或赋予芳香的方法, 包括使所述芳香组合物与芳香增加或增强产品混合的步骤, 所述芳香增加或增强产品根据包括以下步骤的方法制备:

[0387] (a) 制备反应混合物, 所述反应混合物包含与副产物化合物(II)、(III)或(IV)中一种或多种相混合的(-)-降龙涎醚。



[0389] (b) 萃取与副产物化合物(II)、(III)或(IV)中的一者或多者相混合的(-)-降龙涎醚; 以及

[0390] (c) 从所述萃取混合物中选择性结晶(-)-降龙涎醚;

[0391] 其中通过使用SHC/HAC酶在适于产生(-)-降龙涎醚的反应条件下酶促转化(3E, 7E)-高法呢醇(EEH)或包含EEH的立体异构体混合物来制备(-)-降龙涎醚, 并且其中所述包含EEH的立体异构体混合物基本上由选自以下的高法呢醇异构体组成: [(3E, 7E)和[(3Z, 7E)]和/或[(3E, 7E)和(3E, 7Z)]和/或[(3Z, 7E)、(3E, 7E)和(3E, 7Z)], 它们也分别命名为[EE:EZ]、[EE:ZE]和[EE:EZ:ZE]。

[0392] 34.段落33的方法,其中所述反应在包含水相、固相和油相的三相系统中发生。

[0393] 35.根据段落33或段落34的方法,其中所述方法使用选自SEQ ID No.1、SEQ ID No.2、SEQ ID No.3、SEQ ID No.4的SHC/HAC酶多肽序列或选自表1、表5、表2、表6、表3、表7、表4、表8或表14,或选自SEQ ID No.5、7、9、11、13、15、17、19、21、23、25、27、29、31、33、35、37、39、171、173、175和/或177的SHC/HAC衍生物或与SEQ ID No.1、SEQ ID No.2、SEQ ID No.3或SEQ ID No.4具有至少30%同一性、至少40%同一性、至少50%同一性、或至少60%同一性、或至少70%同一性、或至少80%同一性、或至少90%同一性、或至少95%同一性、或至少96%同一性、或至少97%同一性、或至少98%同一性、或至少99%同一性的序列进行。

[0394] 36.段落33-35中任一项的方法,其中所述方法使用产生SHC/HAC酶的重组宿主细胞。

[0395] 37.根据段落35或段落36的方法,其中所述编码SHC/HAC酶的核苷酸序列选自SEQ ID No.165、166、167、168、169或SEQ ID No.6、8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、38、40和/或170、172、174和/或176。

[0396] 38.根据段落33-37中任一项的方法,其中高法呢醇向(-)-降龙涎醚的转化在约4-8的pH下于30℃至60℃的温度下发生。

[0397] 39.根据段落33-38中任一项的方法,其中高法呢醇向(-)-降龙涎醚的转化使用如表24或表24a中所列的用于野生型SHC/HAC或SHC/HAC衍生酶的反应条件中的一者或多者,优选在5.0至6.2的pH范围下、优选在35℃的温度下发生。

[0398] 40.根据段落34-39中任一项的方法,其中当生物催化剂与EEH的比率约为2:1时,SDS/细胞比率在10:1至20:1的范围内,优选为16:1。

[0399] 41.根据段落34-40中任一项的方法,其中生物催化剂与高法呢醇的重量比在约0.5-2:1的范围内,优选约1:1或0.5:1。

[0400] 42.根据段落34-41中任一项的方法,其中细胞培养步骤和生物转化反应步骤在同一反应容器中进行。

[0401] 43.根据段落33-42中任一项的方法,其中所述高法呢醇包含EE:EZ立体异构体混合物或基本上由EE:EZ立体异构体混合物组成,所述EE:EZ立体异构体混合物的重量比选自:100:00;99:01;98:02;97:03;96:04;95:05;94:06;93:07;92:08;91:09;90:10;89:11;88:12;87:13;86:14;85:15;84:16;83:17;82:18;81:19;80:20;79:21;78:22;77:23;76:24;75:25;74:26;73:27;72:28;71:29;70:30;69:31;68:32;67:33;66:34;65:35;64:36;63:37;62:38;61:39;以及60:40。

[0402] 44.根据段落43的方法,其中所述高法呢醇包含EE:EZ立体异构体混合物或基本上由EE:EZ立体异构体混合物组成,所述EE:EZ立体异构体混合物的重量比选自:EE:EZ 92:08;EE:EZ 90:10;EE:EZ 80:20;EE:EZ 86:14;EE:EZ 70:30;EE:EZ 69:31;以及EE:EZ 66:34。

[0403] 45.根据段落43或段落44的方法,其中所述高法呢醇包含EE:EZ立体异构体混合物或基本上由EE:EZ立体异构体混合物组成,所述EE:EZ立体异构体混合物的重量比为80:20。

[0404] 46.根据段落33-45中任一项的方法,其中(-)-降龙涎醚以与副产物(II)、(IV)和/或(III)中的至少一种或多种的混合物产生。

[0405] 47.根据段落33-46中任一项的方法,其中使用有机溶剂或汽提/蒸馏步骤或过滤

从所述生物转化反应混合物分离(-)-降龙涎醚。

[0406] 48. 根据段落33-47中任一项的方法,其中使用有机溶剂或汽提/蒸馏步骤从所述生物转化反应混合物的固相分离(-)-降龙涎醚。

[0407] 49. 根据段落47或段落48的方法,其中使用有机溶剂从所述反应混合物分离(-)-降龙涎醚。

[0408] 50. 根据段落49的方法,其中使用乙醇或甲苯从所述反应混合物分离(-)-降龙涎醚。

[0409] 51. 根据段落47-49中任一项的方法,其中使用有机溶剂使(-)-降龙涎醚选择性结晶。

[0410] 52. 根据段落51的方法,其中所述(-)-降龙涎醚基本上不含副产物(II)、(IV)和/或(III)。

[0411] 53. 根据段落33-52中任一项的方法,其中产生了浓度范围为约125-200g/l的(-)-降龙涎醚。

[0412] 54. 根据段落33-53中任一项的方法,其中所述(-)-降龙涎醚具有约0.1至约0.5ng/l的嗅觉阈值。

[0413] 本发明的另外的方面

[0414] 1. 一种角鲨烯何帕烯环化酶(SHC)/高法呢醇降龙涎香醚环化酶(HAC)衍生物,所述衍生物包含相对于SEQ ID No.1具有独立选自取代、缺失或插入的1-50个突变的氨基酸序列。

[0415] 2. 根据段落1的SHC/HAC衍生物,其中所述SHC衍生物包含相对于SEQ ID No.1具有1至40个突变、1-30个突变、1-20个突变、1-10个突变或1-6个突变的氨基酸序列。

[0416] 3. 根据段落1或段落2的SHC/HAC衍生物,其中所述SHC/HAC衍生物包含相对于SEQ ID No.1具有至少40%同一性、至少50%同一性、或至少60%同一性、或至少70%同一性、或至少80%同一性、或至少90%同一性、或至少95%同一性、或至少96%同一性、或至少97%同一性、或至少98%同一性、或至少99%同一性的氨基酸序列。

[0417] 4. 根据段落3的SHC/HAC衍生物,其中所述SHC变体包含与SEQ ID No.1具有至少95%同一性的氨基酸序列。

[0418] 5. 一种SHC/HAC衍生物,其相对于SEQ ID No.1而言包含独立选自取代、缺失或插入的1-10个突变,其中除了SHC活性位点突变之外的所述一个或多个突变位于SHC酶的结构域2中(图19和/或20)。

[0419] 6. 根据段落1-5中任一项的SHC/HAC衍生物,其中相对于SEQ ID No.1的所述一个或多个突变选自表1,其中如果仅选择一个突变,则其不是F601Y。

[0420] 7. 段落6的SHC/HAC衍生物,其中至少2、3、4、5、6、7、8、9或10个突变选自表1或表5。

[0421] 8. 段落2的SHC/HAC衍生物,所述衍生物包含相对于SEQ ID No.1而言具有最多6个突变并且包含与F129L和/或I432T中的至少任一者或多者组合的至少取代F601Y或M132R的氨基酸序列。

[0422] 9. 段落7的SHC/HAC衍生物,所述衍生物包含相对于SEQ ID No.1而言具有最多8个氨基酸改变并且相对于SEQ ID No.1在选自位置77、129、132、192、224、432、579、601和605的位置中包含一个或不止一个氨基酸改变的氨基酸序列,其中相对于SEQ ID No.1而言所

述SHC/HAC衍生物具有增加的HAC酶促活性。

[0423] 10.段落9的SHC/HAC衍生物,所述衍生物相对于SEQ ID No.1而言包含一个或多个取代,所述取代选自:T77A、F129L、M132R、I92V、A224V、I432T、Q579H、F601Y和/或F605W。

[0424] 11.根据段落10的SHC/HAC衍生物,所述衍生物包含F601Y。

[0425] 12.根据段落10的SHC/HAC衍生物,所述衍生物包含F129L。

[0426] 13.根据段落10的SHC/HAC衍生物,所述衍生物包含F601Y和F129L。

[0427] 14.根据段落10的SHC/HAC衍生物,所述衍生物包含M132R和I432T。

[0428] 15.根据段落14的SHC/HAC衍生物,所述衍生物还包含氨基酸取代A224V。

[0429] 16.根据段落14的SHC/HAC衍生物,所述衍生物还包含F601Y。

[0430] 17.根据段落14的SHC/HAC衍生物,所述衍生物还包含F129L。

[0431] 18.根据段落17的SHC/HAC衍生物,所述衍生物还包含F601Y。

[0432] 19.根据段落11的SHC/HAC衍生物,所述衍生物还包含Q579H。

[0433] 20.根据段落10的SHC/HAC衍生物,所述衍生物包含T77A、I92V和F129L。

[0434] 21.根据前述段落任一项的SHC/HAC衍生物,所述衍生物具有选自SEQ ID No.5、7、9、11、13、15、17、19、21、23、25、27、29、31、33、35、37、39和/或171的氨基酸序列。

[0435] 22.一种分离的核苷酸序列,所述核苷酸序列编码根据段落1-21中任一项的SHC衍生物。

[0436] 23.根据段落22的分离的核苷酸序列,其中所述核苷酸序列选自SEQ ID No.6、8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、38、40和/或170。

[0437] 24.一种构建体,所述构建体包含段落22或段落23的核苷酸序列。

[0438] 25.根据段落24的构建体,所述构建体包含与段落22或23的核苷酸序列功能性连接的启动子。

[0439] 26.段落25的构建体,其中所述启动子是诱导型启动子或组成型启动子。

[0440] 27.一种载体,所述载体包含根据段落24-26中任一项的构建体。

[0441] 28.段落27的载体,其中所述载体是质粒。

[0442] 29.根据段落28的载体,所述载体能够在选自原核生物、酵母、植物和昆虫宿主细胞的宿主细胞中指导表达。

[0443] 30.段落24-26中任一项的构建体或根据段落27-29中任一项的载体,其中所述构建体或所述载体能够整合进选自原核生物、酵母、植物和昆虫宿主细胞的宿主细胞的基因组中。

[0444] 31.一种重组宿主细胞,所述重组宿主细胞包含根据段落22或23的核苷酸序列或根据段落24-26或30中任一项的构建体或根据段落27-30中任一项的载体。

[0445] 32.根据段落31的重组宿主细胞,其中所述宿主细胞选自原核宿主细胞,所述原核宿主细胞由埃希杆菌属(*Escherichia*)、链霉菌属(*Streptomyces*)、芽孢杆菌属(*Bacillus*)、假单胞菌属(*Pseudomonas*)、乳杆菌属(*Lactobacillus*)和乳球菌属(*Lactococcus*)的细菌组成。

[0446] 33.段落32的重组宿主细胞,其中所述宿主细胞是大肠杆菌宿主细胞。

[0447] 34.段落33的重组宿主细胞,其中所述宿主细胞过表达编码所述SHC/HAC衍生物的基因。

[0448] 35.一种制备根据段落1-21中任一项的SHC/HAC衍生物的方法,所述方法包括在允许产生所述SHC/HAC衍生酶的条件下培养一种或多种根据段落31-34中任一项的重组宿主细胞的步骤。

[0449] 36.段落35的方法,其中所述细胞培养在适合生物催化剂产生的条件下发生。

[0450] 37.一种制备(-)-降龙涎醚的方法,所述方法包括使用根据段落31-34中任一项的重组宿主细胞或通过使用包含编码WT SHC/HAC的SEQ ID No.169或SEQ ID No.165的重组宿主细胞将高法呢醇转化为(-)-降龙涎醚,其中如果使用WT SHC/HAC,则高法呢醇向(-)-降龙涎醚的生物转化用除了Triton X-100或牛黄脱氧胆酸盐的增溶剂进行。

[0451] 38.根据段落37的方法,其中高法呢醇向(-)-降龙涎醚的转化是在适用于WT SHC/HAC或SHC/HAC衍生酶的生物转化反应条件下进行。

[0452] 39.根据段落37或38的方法,其中高法呢醇向(-)-降龙涎醚的转化在适用于WT SHC/HAC或SHC/HAC衍生酶的pH、温度、增溶剂浓度下发生。

[0453] 40.根据段落39的方法,其中高法呢醇向(-)-降龙涎醚的转化在范围为30℃至60℃的温度下、在范围为约4-8的pH下并且对于WT SHC/HAC酶在存在除了Triton X-100或牛黄脱氧胆酸盐之外的增溶剂的情况下进行。

[0454] 41.根据段落37-40中任一项的方法,其中高法呢醇向(-)-降龙涎醚的转化使用如表24或表24a中所列的用于WT SHC/HAC或SHC/HAC衍生酶的反应条件中的一者或多者进行。

[0455] 42.根据段落37-41中任一项的方法,其中生物催化剂与高法呢醇的重量比在约0.5:1至2:1的范围内,优选约1:1或0.5:1。

[0456] 43.根据段落37-42中任一项的方法,其中细胞培养步骤和生物转化反应步骤在同一反应容器中进行。

[0457] 44.根据段落37-43中任一项的方法,其中所述高法呢醇底物包含一种或多种高法呢醇立体异构体。

[0458] 45.根据段落44的方法,其中所述高法呢醇底物包含两种高法呢醇立体异构体。

[0459] 46.根据段落45的方法,其中所述高法呢醇底物包含EE:EZ立体异构体。

[0460] 47.根据段落44-46中任一项的方法,其中所述高法呢醇包含EE:EZ立体异构体混合物,所述EE:EZ立体异构体混合物的重量比选自:100:00;99:01;98:02;97:03;96:04;95:05;94:06;93:07;92:08;91:09;90:10;89:11;88:12;87:13;86:14;85:15;84:16;83:17;82:18;81:19;80:20;79:21;78:22;77:23;76:24;75:25;74:26;73:27;72:28;71:29以及70:30。

[0461] 48.段落47的方法,其中所述高法呢醇包含EE:EZ立体异构体混合物,所述EE:EZ立体异构体混合物的重量比选自:EE:EZ 90:10;EE:EZ 80:20;EE:EZ 86:14;EE:EZ 70:30;EE:EZ 69:31;以及EE:EZ 66:34。

[0462] 49.段落35或36的方法,其中所述高法呢醇包含EE:EZ立体异构体混合物,所述EE:EZ立体异构体混合物的重量比为80:20。

[0463] 50.段落37-49中任一项的方法,其中(-)-降龙涎醚以与副产物(II)、(IV)和/或(III)中的至少一种或多种的混合物产生。

[0464] 51.根据段落37-50中任一项的方法,其中使用有机溶剂或汽提/蒸馏步骤从所述生物转化反应混合物分离(-)-降龙涎醚,或者通过过滤直接从生物转化反应混合物分离

(-)-降龙涎醚晶体。

[0465] 52. 根据段落51的方法, 其中使用有机溶剂从所述反应混合物分离(-)-降龙涎醚。

[0466] 53. 段落52的方法, 其中使用有机溶剂使(-)-降龙涎醚选择性结晶。

[0467] 54. 段落52或53的方法, 其中所述(-)-降龙涎醚基本上不含副产物(II)、(IV)和/或(III)。

[0468] 55. 通过段落51-54中任一项的方法能够获得的(-)-降龙涎醚。

[0469] 56. 段落55的(-)-降龙涎醚, 其为固体形式, 优选为无定形形式或晶体形式。

[0470] 57. 一种制备含有(-)-降龙涎醚的产品的的方法, 所述方法包括将段落55或56的(-)-降龙涎醚掺入所述产品, 优选香料产品、美容产品、清洁产品、洗涤剂产品或皂产品。

[0471] 58. 一种香料或化妆品或消费者护理产品, 其中包含段落55或56的(-)-降龙涎醚。

[0472] 59. 一种香料或化妆品或消费者护理组合物, 其中包含段落55或56的(-)-降龙涎醚和一种或多种另外的组分。

[0473] 60. 段落55或56的(-)-降龙涎醚的用途, 其用作香料或化妆品或消费品诸如织物护理、化妆用具、美容护理和/或清洁产品的一部分。

[0474] 61. 根据段落1-21中任一项的SHC/HAC衍生酶、根据段落22或23的核苷酸序列、根据段落24-26或30中任一项的构建体、根据段落27-30中任一项的载体或根据段落31-34中任一项的重组宿主细胞或表达WT SHC/HAC酶的重组宿主细胞的用途, 用于将高法呢醇生物转化为(-)-降龙涎醚, 其中所述WT SHC/HAC酶与除了Triton X-100以外的增溶剂一起用于所述生物转化反应。

[0475] 62. 一种制备(-)-降龙涎醚或(-)-降龙涎醚立体异构体混合物的方法, 其中(3E, 7E)-高法呢醇或(3E, 7E)-高法呢醇立体异构体混合物经酶促转化而得到(-)-降龙涎醚或(-)-降龙涎醚立体异构体混合物, 其中所述酶促转化使用SHC/HAC酶在适于产生(-)-降龙涎醚的反应条件下进行, 并且其中如果所述反应在存在增溶剂的情况下进行, 则不将Triton X-100与WT SHC/HAC酶结合使用。

[0476] 63. 根据段落62的方法, 其中所述方法使用选自AacSHC (SEQ ID No.1)、Zmo SHC1 (SEQ ID No.2)、ZmoSHC2 (SEQ ID No.3)、BjpSHC (SEQ ID No.4)的SHC/HAC、选自表1、表5、表2、表6、表3、表7、表4和/或表8的SHC/HAC衍生物或相对于SEQ ID No.1、SEQ ID No.2、SEQ ID No.3和/或SEQ ID No.4具有至少30%同一性、至少40%同一性、至少50%同一性、或至少60%同一性、或至少70%同一性、或至少80%同一性、或至少90%同一性、或至少95%同一性、或至少96%同一性、或至少97%同一性、或至少98%同一性、或至少99%同一性的序列进行。

[0477] 64. 根据段落63的方法, 其中高法呢醇向(-)-降龙涎醚的转化在范围为30℃至60℃的温度下、在范围为4-8的pH下并且对于WT SHC在存在除了Triton X-100的增溶剂的情况下进行。

[0478] 65. 根据段落64的方法, 其中使用表24或表24a中所列出的用于WT SHC/HAC或每种SHC/HAC衍生物的反应条件。

[0479] 66. 根据段落62-65中任一项的方法, 其中所述方法包括(a)在E,E-高法呢醇转化为(-)-降龙涎醚之前, 在允许表达WT SHC或SHC/HAC衍生多肽的条件下培养表达WT SHC或SHC衍生酶的一个或多个重组宿主细胞。

[0480] 67.根据段落66的方法,其中所述培养步骤和后续的转化步骤在同一反应容器中于不同反应条件下进行。

[0481] 68.根据段落67的方法,其中所述培养步骤是在约6至约7的pH范围下进行,并且高法呢醇转化为(-)-降龙涎醚的步骤是在约4.8-5.5的pH范围下进行。

[0482] 69.根据段落62-68中任一项的方法,其中所述高法呢醇底物包含EE:EZ立体异构体。

[0483] 70.根据段落69的方法,其中所述高法呢醇包含EE:EZ立体异构体混合物,所述EE:EZ立体异构体混合物的重量比选自:EE:EZ 90:10;EE:EZ 80:20;EE:EZ 86:14;EE:EZ 70:30;EE:EZ 69:31;以及EE:EZ 66:34。

[0484] 71.段落70的方法,其中所述高法呢醇包含重量比为80:20的EE:EZ。

[0485] 72.段落62-71中任一项的方法,其中(-)-降龙涎醚以与副产物(II)、(IV)和/或(III)中的至少一种或多种的混合物产生。

[0486] 73.段落62-72中任一项的方法,其中使用有机溶剂或汽提/蒸馏步骤或过滤从所述反应混合物分离(-)-降龙涎醚。

[0487] 74.根据段落73的方法,其中使用有机溶剂从所述反应混合物分离(-)-降龙涎醚。

[0488] 75.段落74的方法,其中使用有机溶剂使(-)-降龙涎醚选择性结晶。

[0489] 76.段落74或75的方法,其中所述(-)-降龙涎醚基本上不含副产物(II)、(IV)和/或(III)。

[0490] 77.通过段落72-76中任一项的方法能够获得的(-)-降龙涎醚。

[0491] 78.段落26的(-)-降龙涎醚,其为固体形式,优选为无定形形式或晶体形式。

[0492] 79.一种制备产品的方法,所述方法包括将段落77或78的(-)-降龙涎醚掺入所述产品中。

[0493] 80.段落79的方法,其中所述产品是香料产品、化妆品、清洁产品、洗涤剂产品或皂产品。

[0494] 81.一种香料或化妆品或消费者护理产品,其中包含段落77或78的(-)-降龙涎醚。

[0495] 82.一种香料或化妆品或消费者护理组合物,其中包含段落77或78的(-)-降龙涎醚和另外的组分。

[0496] 83.段落77或78的(-)-降龙涎醚的用途,其用作香料或化妆品、消费者护理产品的一部分。

[0497] 本发明的另外的方面(ZmoSHC1)

[0498] 1.一种角鲨烯何帕烯环化酶(SHC)/高法呢醇降龙涎香醚环化酶(HAC)衍生物,所述衍生物包含相对于SEQ ID No.2具有独立选自取代、缺失或插入的1-50个突变的氨基酸序列。

[0499] 2.根据段落1的SHC/HAC衍生物,其中所述SHC衍生物包含相对于SEQ ID No.2具有1至40个突变、1-30个突变、1-20个突变、1-10个突变或1-6个突变的氨基酸序列。

[0500] 3.根据段落1或段落2的SHC/HAC衍生物,其中所述SHC/HAC衍生物包含相对于SEQ ID No.2具有至少40%同一性、至少50%同一性、或至少60%同一性、或至少70%同一性、或至少80%同一性、或至少90%同一性、或至少95%同一性、或至少96%同一性、或至少97%同一性、或至少98%同一性、或至少99%同一性的氨基酸序列。

- [0501] 4. 根据段落3的SHC/HAC衍生物,其中所述SHC变体包含与SEQ ID No.2具有至少95%同一性的氨基酸序列。
- [0502] 5. 一种SHC/HAC衍生物,其相对于SEQ ID No.2而言包含独立选自取代、缺失或插入的1-10个突变,其中除了SHC活性位点突变之外的所述一个或多个突变位于SHC酶的结构域2中(图19和/或20)。
- [0503] 6. 根据段落1-5中任一项的SHC/HAC衍生物,其中相对于SEQ ID No.2的所述一个或多个突变选自表2,其中如果仅选择一个突变,则其不是F668Y。
- [0504] 7. 段落6的SHC/HAC衍生物,其中至少2、3、4、5、6、7、8、9或10个突变选自表2和/或表6。
- [0505] 8. 段落2的SHC/HAC衍生物,所述衍生物包含相对于SEQ ID No.2而言具有最多6个突变并且包含与F182L和/或I498T中的至少任一者或多者组合的至少取代F668Y或Y185R的氨基酸序列。
- [0506] 9. 段落7的SHC/HAC衍生物,所述衍生物包含相对于SEQ ID No.2而言具有最多8个氨基酸改变并且相对于SEQ ID No.2在选自位置129、145、182、185、282、498、647和668的位置中包含一个或多个氨基酸改变的氨基酸序列,其中相对于SEQ ID No.2而言所述SHC/HAC衍生物具有增加的HAC酶促活性。
- [0507] 10. 段落9的SHC/HAC衍生物,所述衍生物相对于SEQ ID No.2而言包含一个或多个取代,所述取代选自:S129A、V145V、F182L、Y185R、G282V、I498T、H646H和F668Y。
- [0508] 11. 根据段落10的SHC/HAC衍生物,所述衍生物包含F668Y。
- [0509] 12. 根据段落10的SHC/HAC衍生物,所述衍生物包含F182L。
- [0510] 13. 根据段落10的SHC/HAC衍生物,所述衍生物包含F668Y和F182L。
- [0511] 14. 根据段落10的SHC/HAC衍生物,所述衍生物包含Y185R和I498T。
- [0512] 15. 根据段落14的SHC/HAC衍生物,所述衍生物还包含G282V。
- [0513] 16. 根据段落14的SHC/HAC衍生物,所述衍生物还包含F668Y。
- [0514] 17. 根据段落14的SHC/HAC衍生物,所述衍生物还包含F182L。
- [0515] 18. 根据段落17的SHC/HAC衍生物,所述衍生物还包含F668Y。
- [0516] 19. 根据段落11的SHC/HAC衍生物,所述衍生物还包含H646H。
- [0517] 20. 根据段落10的SHC/HAC衍生物,所述衍生物包含S129A和V145V和F182L。
- [0518] 21. 根据前述段落任一项的SHC/HAC衍生物,所述衍生物具有选自SEQ ID No.41、43、45、47、49、51、53、55、57、59、61、63、65、67、69、71、73和/或75的氨基酸序列。
- [0519] 22. 一种分离的核苷酸序列,所述核苷酸序列编码根据段落1-21中任一项的SHC衍生物。
- [0520] 23. 根据段落22的分离的核苷酸序列,其中所述核苷酸序列选自SEQ ID No.42、44、46、48、50、52、54、56、58、60、62、64、66、68、70、72、74和/或76。
- [0521] 24. 一种构建体,所述构建体包含段落22或段落23的核苷酸序列。
- [0522] 25. 根据段落24的构建体,所述构建体包含与段落22或23的核苷酸序列功能性连接的启动子。
- [0523] 26. 段落25的构建体,其中所述启动子是诱导型启动子或组成型启动子。
- [0524] 27. 一种载体,所述载体包含根据段落24-26中任一项的构建体。

- [0525] 28.段落27的载体,其中所述载体是质粒。
- [0526] 29.根据段落28的载体,所述载体能够在选自原核生物、酵母、植物和昆虫宿主细胞的宿主细胞中指导表达。
- [0527] 30.段落24-26中任一项的构建体或根据段落27-29中任一项的载体,其中所述构建体或所述载体能够整合进选自原核生物、酵母、植物和昆虫宿主细胞的宿主细胞的基因组中。
- [0528] 31.一种重组宿主细胞,所述重组宿主细胞包含根据段落22或23的核苷酸序列或根据段落24-26或30中任一项的构建体或根据段落27-30中任一项的载体。
- [0529] 32.根据段落31的重组宿主细胞,其中所述宿主细胞选自原核宿主细胞,所述原核宿主细胞由埃希杆菌属(*Escherichia*)、链霉菌属(*Streptomyces*)、芽孢杆菌属(*Bacillus*)、假单胞菌属(*Pseudomonas*)、乳杆菌属(*Lactobacillus*)和乳球菌属(*Lactococcus*)的细菌组成。
- [0530] 33.段落32的重组宿主细胞,其中所述宿主细胞是大肠杆菌宿主细胞。
- [0531] 34.段落33的重组宿主细胞,其中所述宿主细胞过表达编码所述SHC/HAC衍生物的基因。
- [0532] 35.一种制备根据段落1-21中任一项的SHC/HAC衍生物的方法,所述方法包括在允许产生所述SHC/HAC衍生酶的条件下培养一种或多种根据段落31-34中任一项的重组宿主细胞的步骤。
- [0533] 36.段落35的方法,其中所述细胞培养在适合生物催化剂产生的条件下发生。
- [0534] 37.一种制备(-)-降龙涎醚的方法,所述方法包括使用根据段落31-34中任一项的重组宿主细胞或通过使用包含编码WT SHC/HAC的SEQ ID No.166的重组宿主细胞将高法呢醇转化为(-)-降龙涎醚,其中如果使用WT SHC/HAC,则高法呢醇向(-)-降龙涎醚的生物转化用除了Triton X-100的增溶剂进行。
- [0535] 38.根据段落37的方法,其中高法呢醇向(-)-降龙涎醚的转化是在适用于WT SHC/HAC或SHC/HAC衍生酶的生物转化反应条件下进行。
- [0536] 39.根据段落37或38的方法,其中高法呢醇向(-)-降龙涎醚的转化在适用于WT SHC/HAC或SHC/HAC衍生酶的pH、温度、增溶剂浓度下发生。
- [0537] 40.根据段落39的方法,其中高法呢醇向(-)-降龙涎醚的转化在范围为30℃至60℃的温度下、在范围为约4-8的pH下并且对于WT SHC/HAC酶在存在除了Triton X-100之外的增溶剂的情况下进行。
- [0538] 41.根据段落37-40中任一项的方法,其中高法呢醇向(-)-降龙涎醚的转化使用如表24或表24a中所列的用于WT SHC/HAC或SHC/HAC衍生酶的反应条件中的一者或多者进行。
- [0539] 42.根据段落37-41中任一项的方法,其中生物催化剂与高法呢醇的重量比在约0.5:1至2:1的范围内,优选约1:1或0.5:1。
- [0540] 43.根据段落37-42中任一项的方法,其中细胞培养步骤和生物转化反应步骤在同一反应容器中进行。
- [0541] 44.根据段落37-43中任一项的方法,其中所述高法呢醇底物包含一种或多种高法呢醇立体异构体。
- [0542] 45.根据段落44的方法,其中所述高法呢醇底物包含两种高法呢醇立体异构体。

- [0543] 46. 根据段落45的方法,其中所述高法呢醇底物包含EE:EZ立体异构体。
- [0544] 47. 根据段落44-46中任一项的方法,其中所述高法呢醇包含EE:EZ立体异构体混合物,所述EE:EZ立体异构体混合物的重量比选自:100:00;99:01;98:02;97:03;96:04;95:05;94:06;93:07;92:08;91:09;90:10;89:11;88:12;87:13;86:14;85:15;84:16;83:17;82:18;81:19;80:20;79:21;78:22;77:23;76:24;75:25;74:26;73:27;72:28;71:29以及70:30。
- [0545] 48. 段落47的方法,其中所述高法呢醇包含EE:EZ立体异构体混合物,所述EE:EZ立体异构体混合物的重量比选自:EE:EZ 90:10;EE:EZ 80:20;EE:EZ 86:14;EE:EZ 70:30;EE:EZ 69:31;以及EE:EZ 66:34。
- [0546] 49. 段落35或36的方法,其中所述高法呢醇包含重量比为80:20的EE:EZ立体异构体混合物。
- [0547] 50. 段落37-49中任一项的方法,其中(-)-降龙涎醚以与副产物(II)、(IV)和/或(III)中的至少一种或多种的混合物产生。
- [0548] 51. 段落37-50中任一项的方法,其中使用有机溶剂或汽提/蒸馏步骤或过滤从所述生物转化反应混合物分离(-)-降龙涎醚。
- [0549] 52. 根据段落51的方法,其中使用有机溶剂从所述反应混合物分离(-)-降龙涎醚。
- [0550] 53. 段落52的方法,其中使用有机溶剂使(-)-降龙涎醚选择性结晶。
- [0551] 54. 段落52或53的方法,其中所述(-)-降龙涎醚基本上不含副产物(II)、(IV)和/或(III)。
- [0552] 55. 通过段落51-54中任一项的方法能够获得的(-)-降龙涎醚。
- [0553] 56. 段落55的(-)-降龙涎醚,其为固体形式,优选为无定形形式或晶体形式。
- [0554] 57. 一种制备含有(-)-降龙涎醚的产品的的方法,所述方法包括将段落55或56的(-)-降龙涎醚掺入所述产品,优选香料产品、美容产品、清洁产品、洗涤剂产品或皂产品。
- [0555] 58. 一种香料或化妆品或消费者护理产品,其中包含段落55或56的(-)-降龙涎醚。
- [0556] 59. 一种香料或化妆品或消费者护理组合物,其中包含段落55或56的(-)-降龙涎醚和一种或多种另外的组分。
- [0557] 60. 段落55或56的(-)-降龙涎醚的用途,其用作香料或化妆品或消费品诸如织物护理、化妆用具、美容护理和/或清洁产品的一部分。
- [0558] 61. 根据段落1-21中任一项的SHC/HAC衍生酶、根据段落22或23的核苷酸序列、根据段落24-26或30中任一项的构建体、根据段落27-30中任一项的载体或根据段落31-34中任一项的重组宿主细胞或表达WT SHC/HAC酶的重组宿主细胞的用途,用于将高法呢醇生物转化为(-)-降龙涎醚,其中所述WT SHC/HAC酶与除了Triton X-100以外的增溶剂一起用于所述生物转化反应。
- [0559] 本发明的另外的方面(ZmoSHC2)
- [0560] 1. 一种角鲨烯何帕烯环化酶(SHC)/高法呢醇降龙涎香醚环化酶(HAC)衍生物,所述衍生物包含相对于SEQ ID No.3具有独立选自取代、缺失或插入的1-50个突变的氨基酸序列。
- [0561] 2. 根据段落1的SHC/HAC衍生物,其中所述SHC衍生物包含相对于SEQ ID No.3具有1至40个突变、1-30个突变、1-20个突变、1-10个突变或1-6个突变的氨基酸序列。

[0562] 3. 根据段落1或段落2的SHC/HAC衍生物,其中所述SHC/HAC衍生物包含相对于SEQ ID No.3具有至少40%同一性、至少50%同一性、或至少60%同一性、或至少70%同一性、或至少80%同一性、或至少90%同一性、或至少95%同一性、或至少96%同一性、或至少97%同一性、或至少98%同一性、或至少99%同一性的氨基酸序列。

[0563] 4. 根据段落3的SHC/HAC衍生物,其中所述SHC变体包含与SEQ ID No.3具有至少95%同一性的氨基酸序列。

[0564] 5. 一种SHC/HAC衍生物,其相对于SEQ ID No.3而言包含独立选自取代、缺失或插入的1-10个突变,其中除了SHC活性位点突变之外的所述一个或多个突变位于SHC酶的结构域2中(图19和/或20)。

[0565] 6. 根据段落1-5中任一项的SHC/HAC衍生物,其中相对于SEQ ID No.3的所述一个或多个突变选自表3,其中如果仅选择一个突变,则其不是F620Y。

[0566] 7. 段落6的SHC/HAC衍生物,其中至少2、3、4、5、6、7、8、9或10个突变选自表3和/或表7。

[0567] 8. 段落2的SHC/HAC衍生物,所述衍生物包含相对于SEQ ID No.3而言具有最多6个突变并且包含与F137L和/或I450T中的至少任一者或多者组合的至少取代F620Y或I140R的氨基酸序列。

[0568] 9. 段落7的SHC/HAC衍生物,所述衍生物包含相对于SEQ ID No.3而言具有最多8个氨基酸改变并且相对于SEQ ID No.3在选自位置85、100、137、140、233、450、598和620的位置中包含一个或多个氨基酸改变的氨基酸序列,其中相对于SEQ ID No.3而言所述SHC/HAC衍生物具有增加的HAC酶促活性。

[0569] 10. 段落9的SHC/HAC衍生物,所述衍生物相对于SEQ ID No.3而言包含一个或多个取代,所述取代选自:G85A、V100V、F137L、I140R、V233V、I450T、N598H和F620Y。

[0570] 11. 根据段落10的SHC/HAC衍生物,所述衍生物包含F620Y。

[0571] 12. 根据段落10的SHC/HAC衍生物,所述衍生物包含F137L。

[0572] 13. 根据段落10的SHC/HAC衍生物,所述衍生物包含F620Y和F137L。

[0573] 14. 根据段落10的SHC/HAC衍生物,所述衍生物包含I140R和I450T。

[0574] 15. 根据段落14的SHC/HAC衍生物,所述衍生物还包含V233V。

[0575] 16. 根据段落14的SHC/HAC衍生物,所述衍生物还包含F620Y。

[0576] 17. 根据段落14的SHC/HAC衍生物,所述衍生物还包含F137L。

[0577] 18. 根据段落17的SHC/HAC衍生物,所述衍生物还包含F620Y。

[0578] 19. 根据段落11的SHC/HAC衍生物,所述衍生物还包含N598H。

[0579] 20. 根据段落10的SHC/HAC衍生物,所述衍生物包含G85A和V100V和F137L。

[0580] 21. 根据前述段落任一项的SHC/HAC衍生物,所述衍生物具有选自SEQ ID No.77、79、81、83、85、87、89、91、93、95、97、99、101、103、105、107、109和/或111的氨基酸序列。

[0581] 22. 一种分离的核苷酸序列,所述核苷酸序列编码根据段落1-21中任一项的SHC衍生物。

[0582] 23. 根据段落22的分离的核苷酸序列,其中所述核苷酸序列选自SEQ ID No.78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110和/或112。

[0583] 24. 一种构建体,所述构建体包含段落22或段落23的核苷酸序列。

- [0584] 25. 根据段落24的构建体,所述构建体包含与段落22或23的核苷酸序列功能性连接的启动子。
- [0585] 26. 段落25的构建体,其中所述启动子是诱导型启动子或组成型启动子。
- [0586] 27. 一种载体,所述载体包含根据段落24-26中任一项的构建体。
- [0587] 28. 段落27的载体,其中所述载体是质粒。
- [0588] 29. 根据段落28的载体,所述载体能够在选自原核生物、酵母、植物和昆虫宿主细胞的宿主细胞中指导表达。
- [0589] 30. 段落24-26中任一项的构建体或根据段落27-29中任一项的载体,其中所述构建体或所述载体能够整合进选自原核生物、酵母、植物和昆虫宿主细胞的宿主细胞的基因组中。
- [0590] 31. 一种重组宿主细胞,所述重组宿主细胞包含根据段落22或23的核苷酸序列或根据段落24-26或30中任一项的构建体或根据段落27-30中任一项的载体。
- [0591] 32. 根据段落31的重组宿主细胞,其中所述宿主细胞选自原核宿主细胞,所述原核宿主细胞由埃希杆菌属 (*Escherichia*)、链霉菌属 (*Streptomyces*)、芽孢杆菌属 (*Bacillus*)、假单胞菌属 (*Pseudomonas*)、乳杆菌属 (*Lactobacillus*) 和乳球菌属 (*Lactococcus*) 的细菌组成。
- [0592] 33. 段落32的重组宿主细胞,其中所述宿主细胞是大肠杆菌宿主细胞。
- [0593] 34. 段落33的重组宿主细胞,其中所述宿主细胞过表达编码所述SHC/HAC衍生物的基因。
- [0594] 35. 一种制备根据段落1-21中任一项的SHC/HAC衍生物的方法,所述方法包括以下步骤: (a) 在允许产生SHC/HAC衍生酶的条件下培养一种或多种根据段落31-34中任一项的重组宿主细胞。
- [0595] 36. 段落35的方法,其中所述细胞培养在适合生物催化剂产生的条件下发生。
- [0596] 37. 一种制备(-)-降龙涎醚的方法,所述方法包括使用根据段落31-34中任一项的重组宿主细胞或通过使用包含编码WT SHC/HAC的SEQ ID No.167的重组宿主细胞将高法呢醇转化为(-)-降龙涎醚,其中如果使用WT SHC/HAC,则高法呢醇向(-)-降龙涎醚的生物转化用除了Triton X-100的增溶剂进行。
- [0597] 38. 根据段落37的方法,其中高法呢醇向(-)-降龙涎醚的转化是在适用于WT SHC/HAC或SHC/HAC衍生酶的生物转化反应条件下进行。
- [0598] 39. 根据段落37或38的方法,其中高法呢醇向(-)-降龙涎醚的转化在适用于WT SHC/HAC或SHC/HAC衍生酶的pH、温度、增溶剂浓度下进行。
- [0599] 40. 根据段落39的方法,其中高法呢醇向(-)-降龙涎醚的转化在范围为30℃至60℃的温度下、在范围为约4-8的pH下并且对于WT SHC/HAC酶在存在除了Triton X-100之外的增溶剂的情况下进行。
- [0600] 41. 根据段落37-40中任一项的方法,其中高法呢醇向(-)-降龙涎醚的转化使用如表24或表24a中所列的用于WT SHC/HAC或SHC/HAC衍生酶的反应条件中的一者或多者进行。
- [0601] 42. 根据段落37-41中任一项的方法,其中生物催化剂与高法呢醇的重量比在约0.5:1至2:1的范围内,优选约1:1或0.5:1。
- [0602] 43. 根据段落37-42中任一项的方法,其中细胞培养步骤和生物转化反应步骤在同

一反应容器中进行。

[0603] 44. 根据段落37-43中任一项的方法, 其中所述高法呢醇底物包含一种或多种高法呢醇立体异构体。

[0604] 45. 根据段落44的方法, 其中所述高法呢醇底物包含两种高法呢醇立体异构体。

[0605] 46. 根据段落45的方法, 其中所述高法呢醇底物包含EE:EZ立体异构体。

[0606] 47. 根据段落44-46中任一项的方法, 其中所述高法呢醇包含EE:EZ立体异构体混合物, 所述EE:EZ立体异构体混合物的重量比选自: 100:00; 99:01; 98:02; 97:03; 96:04; 95:05; 94:06; 93:07; 92:08; 91:09; 90:10; 89:11; 88:12; 87:13; 86:14; 85:15; 84:16; 83:17; 82:18; 81:19; 80:20; 79:21; 78:22; 77:23; 76:24; 75:25; 74:26; 73:27; 72:28; 71:29以及70:30。

[0607] 48. 段落47的方法, 其中所述高法呢醇包含EE:EZ立体异构体混合物, 所述EE:EZ立体异构体混合物的重量比选自: EE:EZ 90:10; EE:EZ 80:20; EE:EZ 86:14; EE:EZ 70:30; EE:EZ 69:31; 以及EE:EZ 66:34。

[0608] 49. 段落35或36的方法, 其中所述高法呢醇包含重量比为80:20的EE:EZ立体异构体混合物。

[0609] 50. 段落37-49中任一项的方法, 其中(-)-降龙涎醚以与副产物(II)、(IV)和/或(III)中的至少一种或多种的混合物产生。

[0610] 51. 段落37-50中任一项的方法, 其中使用有机溶剂或汽提/蒸馏步骤或过滤从所述生物转化反应混合物分离(-)-降龙涎醚。

[0611] 52. 根据段落51的方法, 其中使用有机溶剂从所述反应混合物分离(-)-降龙涎醚。

[0612] 53. 段落52的方法, 其中使用有机溶剂使(-)-降龙涎醚选择性结晶。

[0613] 54. 段落52或53的方法, 其中所述(-)-降龙涎醚基本上不含副产物(II)、(IV)和/或(III)。

[0614] 55. 通过段落51-54中任一项的方法能够获得的(-)-降龙涎醚。

[0615] 56. 段落55的(-)-降龙涎醚, 其为固体形式, 优选为无定形形式或晶体形式。

[0616] 57. 一种制备含有(-)-降龙涎醚的产品的的方法, 所述方法包括将段落55或56的(-)-降龙涎醚掺入所述产品, 优选香料产品、美容产品、清洁产品、洗涤剂产品或皂产品。

[0617] 58. 一种香料或化妆品或消费者护理产品, 其中包含段落55或56的(-)-降龙涎醚。

[0618] 59. 一种香料或化妆品或消费者护理组合物, 其中包含段落55或56的(-)-降龙涎醚和一种或多种另外的组分。

[0619] 60. 段落55或56的(-)-降龙涎醚的用途, 其用作香料或化妆品或消费品诸如织物护理、化妆用具、美容护理和/或清洁产品的一部分。

[0620] 61. 根据段落1-21中任一项的SHC/HAC衍生酶、根据段落22或23的核苷酸序列、根据段落24-26或30中任一项的构建体、根据段落27-30中任一项的载体或根据段落31-34中任一项的重组宿主细胞或表达WT SHC/HAC酶的重组宿主细胞的用途, 用于将高法呢醇生物转化为(-)-降龙涎醚, 其中所述WT SHC/HAC酶与除了Triton X-100以外的增溶剂一起用于所述生物转化反应。

[0621] 本发明的另外的方面(BJpSHC)

[0622] 1. 一种角鲨烯何帕烯环化酶(SHC)/高法呢醇降龙涎香醚环化酶(HAC)衍生物, 所

述衍生物包含相对于SEQ ID No.4具有独立选自取代、缺失或插入的1-50个突变的氨基酸序列。

[0623] 2.根据段落1的SHC/HAC衍生物,其中所述SHC衍生物包含相对于SEQ ID No.4具有1至40个突变、1-30个突变、1-20个突变、1-10个突变或1-6个突变的氨基酸序列。

[0624] 3.根据段落1或段落2的SHC/HAC衍生物,其中所述SHC/HAC衍生物包含相对于SEQ ID No.4具有至少40%同一性、至少50%同一性、或至少60%同一性、或至少70%同一性、或至少80%同一性、或至少90%同一性、或至少95%同一性、或至少96%同一性、或至少97%同一性、或至少98%同一性、或至少99%同一性的氨基酸序列。

[0625] 4.根据段落3的SHC/HAC衍生物,其中所述SHC变体包含与SEQ ID No.4具有至少95%同一性的氨基酸序列。

[0626] 5.一种SHC/HAC衍生物,其相对于SEQ ID No.4而言包含4-10个独立选自取代、缺失或插入的1-10个突变,其中除了SHC活性位点突变之外的所述一个或多个突变位于SHC酶的结构域2中(图19和/或20)。

[0627] 6.根据段落1-5中任一项的SHC/HAC衍生物,其中相对于SEQ ID No.4的所述一个或多个突变选自表4,其中如果仅选择一个突变,则其不是F628Y。

[0628] 7.段落6的SHC/HAC衍生物,其中至少2、3、4、5、6、7、8、9或10个突变选自表4和/或表8。

[0629] 8.段落2的SHC/HAC衍生物,所述衍生物包含相对于SEQ ID No.4而言具有最多6个突变并且包含与F137L和/或I450T中的至少任一者或多者组合的至少取代F628Y或I140R的氨基酸序列。

[0630] 9.段落7的SHC/HAC衍生物,所述衍生物包含相对于SEQ ID No.4而言具有最多8个氨基酸改变并且相对于SEQ ID No.4在选自位置88、104、141、144、241、459、607和628的位置中包含一个或多个氨基酸改变的氨基酸序列,其中相对于SEQ ID No.4而言所述SHC/HAC衍生物具有增加的HAC酶促活性。

[0631] 10.段落9的SHC/HAC衍生物,所述衍生物相对于SEQ ID No.4而言包含一个或多个取代,所述取代选自:A88A、V104V、F141L、Y144R、V241V、I459T、M607H和F628Y。

[0632] 11.根据段落10的SHC/HAC衍生物,所述衍生物包含F628Y。

[0633] 12.根据段落10的SHC/HAC衍生物,所述衍生物包含F141L。

[0634] 13.根据段落10的SHC/HAC衍生物,所述衍生物包含F628Y和F141L。

[0635] 14.根据段落10的SHC/HAC衍生物,所述衍生物包含Y144R和I459T。

[0636] 15.根据段落14的SHC/HAC衍生物,所述衍生物还包含V241V。

[0637] 16.根据段落14的SHC/HAC衍生物,所述衍生物还包含F628Y。

[0638] 17.根据段落14的SHC/HAC衍生物,所述衍生物还包含F141L。

[0639] 18.根据段落17的SHC/HAC衍生物,所述衍生物还包含F628Y。

[0640] 19.根据段落11的SHC/HAC衍生物,所述衍生物还包含M607H。

[0641] 20.根据段落10的SHC/HAC衍生物,所述衍生物包含S129A和V145V和F182L。

[0642] 21.根据前述段落任一项的SHC/HAC衍生物,所述衍生物具有选自SEQ ID No.113、115、117、119、121、123、125、127、129、131、133、135、137、139、141、143、145和/或147的氨基酸序列。

- [0643] 22.一种分离的核苷酸序列,所述核苷酸序列编码根据段落1-21中任一项的SHC衍生物。
- [0644] 23.根据段落22的分离的核苷酸序列,其中所述核苷酸序列选自SEQ ID No.114、116、118、120、124、126、128、130、132、134、136、138、140、142、144、146和/或148。
- [0645] 24.一种构建体,所述构建体包含段落22或段落23的核苷酸序列。
- [0646] 25.根据段落24的构建体,所述构建体包含与段落22或23的核苷酸序列功能性连接的启动子。
- [0647] 26.段落25的构建体,其中所述启动子是诱导型启动子或组成型启动子。
- [0648] 27.一种载体,所述载体包含根据段落24-26中任一项的构建体。
- [0649] 28.段落27的载体,其中所述载体是质粒。
- [0650] 29.根据段落28的载体,所述载体能够在选自原核生物、酵母、植物和昆虫宿主细胞的宿主细胞中指导表达。
- [0651] 30.段落24-26中任一项的构建体或根据段落27-29中任一项的载体,其中所述构建体或所述载体能够整合进选自原核生物、酵母、植物和昆虫宿主细胞的宿主细胞的基因组中。
- [0652] 31.一种重组宿主细胞,所述重组宿主细胞包含根据段落22或23的核苷酸序列或根据段落24-26或30中任一项的构建体或根据段落27-30中任一项的载体。
- [0653] 32.根据段落31的重组宿主细胞,其中所述宿主细胞选自原核宿主细胞,所述原核宿主细胞由埃希杆菌属(*Escherichia*)、链霉菌属(*Streptomyces*)、芽孢杆菌属(*Bacillus*)、假单胞菌属(*Pseudomonas*)、乳杆菌属(*Lactobacillus*)和乳球菌属(*Lactococcus*)的细菌组成。
- [0654] 33.段落32的重组宿主细胞,其中所述宿主细胞是大肠杆菌宿主细胞。
- [0655] 34.段落33的重组宿主细胞,其中所述宿主细胞过表达编码所述SHC/HAC衍生物基因。
- [0656] 35.一种制备根据段落1-21中任一项的SHC/HAC衍生物的方法,所述方法包括以下步骤:(a)在允许产生SHC/HAC衍生酶的条件下培养一种或多种根据段落31-34中任一项的重组宿主细胞。
- [0657] 36.段落35的方法,其中所述细胞培养在适合生物催化剂产生的条件下发生。
- [0658] 37.一种制备(-)-降龙涎醚的方法,所述方法包括使用根据段落31-34中任一项的重组宿主细胞或通过使用包含编码WT SHC/HAC的SEQ ID No.168的重组宿主细胞将高法呢醇转化为(-)-降龙涎醚,其中如果使用WT SHC/HAC,则高法呢醇向(-)-降龙涎醚的生物转化用除了Triton X-100的增溶剂进行。
- [0659] 38.根据段落37的方法,其中高法呢醇向(-)-降龙涎醚的转化是在适用于WT SHC/HAC或SHC/HAC衍生酶的生物转化反应条件下进行。
- [0660] 39.根据段落37或38的方法,其中高法呢醇向(-)-降龙涎醚的转化在适用于WT SHC/HAC或SHC/HAC衍生酶的pH、温度、增溶剂浓度下发生。
- [0661] 40.根据段落39的方法,其中高法呢醇向(-)-降龙涎醚的转化在范围为30℃至60℃的温度下、在范围为约4-8的pH下并且对于WT SHC/HAC酶在存在除了Triton X-100之外的增溶剂的情况下进行。

- [0662] 41. 根据段落37-40中任一项的方法,其中高法呢醇向(-)-降龙涎醚的转化使用如表24或表24a中所列的用于WT SHC/HAC或SHC/HAC衍生酶的反应条件中的一者或多者进行。
- [0663] 42. 根据段落37-41中任一项的方法,其中生物催化剂与高法呢醇的重量比在约0.5:1至2:1的范围内,优选约1:1或0.5:1。
- [0664] 43. 根据段落37-42中任一项的方法,其中细胞培养步骤和生物转化反应步骤在同一反应容器中进行。
- [0665] 44. 根据段落27-31中任一项的方法,其中所述高法呢醇底物包含一种或多种高法呢醇立体异构体。
- [0666] 45. 根据段落44的方法,其中所述高法呢醇底物包含两种高法呢醇立体异构体。
- [0667] 46. 根据段落45的方法,其中所述高法呢醇底物包含EE:EZ立体异构体。
- [0668] 47. 根据段落44-46中任一项的方法,其中所述高法呢醇包含EE:EZ立体异构体混合物,所述EE:EZ立体异构体混合物的重量比选自:100:00;99:01;98:02;97:03;96:04;95:05;94:06;93:07;92:08;91:09;90:10;89:11;88:12;87:13;86:14;85:15;84:16;83:17;82:18;81:19;80:20;79:21;78:22;77:23;76:24;75:25;74:26;73:27;72:28;71:29以及70:30。
- [0669] 48. 段落47的方法,其中所述高法呢醇包含EE:EZ立体异构体混合物,所述EE:EZ立体异构体混合物的重量比选自:EE:EZ 90:10;EE:EZ 80:20;EE:EZ 86:14;EE:EZ 70:30;EE:EZ 69:31;以及EE:EZ 66:34。
- [0670] 49. 段落35或36的方法,其中所述高法呢醇包含重量比为80:20的EE:EZ立体异构体混合物。
- [0671] 50. 段落37-49中任一项的方法,其中(-)-降龙涎醚以与副产物(II)、(IV)和/或(III)中的至少一种或多种的混合物产生。
- [0672] 51. 段落37-50中任一项的方法,其中使用有机溶剂或汽提/蒸馏步骤或过滤从所述生物转化反应混合物分离(-)-降龙涎醚。
- [0673] 52. 根据段落51的方法,其中使用有机溶剂从所述反应混合物分离(-)-降龙涎醚。
- [0674] 53. 段落52的方法,其中使用有机溶剂使(-)-降龙涎醚选择性结晶。
- [0675] 54. 段落52或53的方法,其中所述(-)-降龙涎醚基本上不含副产物(II)、(IV)和/或(III)。
- [0676] 55. 通过段落51-54中任一项的方法能够获得的(-)-降龙涎醚。
- [0677] 56. 段落55的(-)-降龙涎醚,其为固体形式,优选为无定形形式或晶体形式。
- [0678] 57. 一种制备含有(-)-降龙涎醚的产品的的方法,所述方法包括将段落55或56的(-)-降龙涎醚掺入所述产品,优选香料产品、美容产品、清洁产品、洗涤剂产品或皂产品。
- [0679] 58. 一种香料或化妆品或消费者护理产品,其中包含段落55或56的(-)-降龙涎醚。
- [0680] 59. 一种香料或化妆品或消费者护理组合物,其中包含段落55或56的(-)-降龙涎醚和一种或多种另外的组分。
- [0681] 60. 段落55或56的(-)-降龙涎醚的用途,其用作香料或化妆品或消费品诸如织物护理、化妆用具、美容护理和/或清洁产品的一部分。
- [0682] 61. 根据段落1-21中任一项的SHC/HAC衍生酶、根据段落22或23的核苷酸序列、根据段落24-26或30中任一项的构建体、根据段落27-30中任一项的载体或根据段落31-34中

任一项的重组宿主细胞或表达WT SHC/HAC酶的重组宿主细胞的用途,用于将高法呢醇生物转化为(-)-降龙涎醚,其中所述WT SHC/HAC酶与除了Triton X-100以外的增溶剂一起用于所述生物转化反应。

[0683] 在另一方面,提供了基于具有本文所述的SHC或衍生物的氨基酸序列的SHC的结构坐标的SHC晶体模型结构(CMS)。SHC CMS包括角鲨烯/高法呢醇结合口袋结构域(SHBD),该结构域包括角鲨烯/高法呢醇结合口袋(SHBP)和结合至SBD的角鲨烯/高法呢醇底物(如参见图19和20)。该SHC晶体模型结构(CMS)有利于潜在的候选SHC/HAC衍生酶候选者的计算机测试(in-silico testing)。

[0684] 因而在又一些其它实施方案中,本公开提供了筛选能够结合至SHBD的酶(如SHC/HAC衍生物)的方法,其中该方法包括使用SHC/HAC CMS。在另一个方面,本公开提供了用于筛选能够结合至SHBP的酶(如参照SHC或SHC/HAC衍生物)的方法,并且该方法包括使SHBP与测试化合物(如SHC衍生物)接触并确定所述测试化合物是否结合至所述SHBP。在一些实施方案中,该方法是筛选可用于调节SHC衍生酶的活性的测试化合物(如调节剂)。

[0685] 在另一个方面,本公开提供了对以下进行预测、模拟或建模的方法:参照SHC和/或SHC/HAC衍生物的分子特征和/或参照SHC和/或SHC/HAC衍生物与角鲨烯/高法呢醇结合结构域(SHBD)的分子相互作用,该方法包括使用计算机模型,所述计算机模型包括使用或描述如上文所限定的角鲨烯/高法呢醇结合结构域的结构坐标以提供所述配体结合结构域的图像以及任选地显示所述图像。

[0686] 在本说明书及后面的权利要求书通篇中,除非上下文另行要求,否则词语“包含”及其变型形式如“包括”和“具有”将理解为暗示包括所述整数或步骤或者整数或步骤的组,但不排除任何其它整数或步骤或者整数或步骤的组。术语“包含”还意指“含有”以及“由...组成”,例如组合物“包含”X可排他性地由X组成或者可含有额外的某些东西,如X+Y。还必须注意的是,如本文中以及随附的权利要求书中所用的,单数形式“一个”、“一种”、“所述”包括多个指代物,除非上下文另外明确规定。举个例子,提及“基因”或“酶”是提及“一种或多种基因”或“一种或多种酶”。

[0687] 应当理解,本公开不限于本文中所述的特定方法、规程和试剂,因为这些可以变化。还应当理解,本文所用的术语仅用于描述具体的实施方案的目的,并非旨在限制本公开的范围,本公开的范围将仅由随附的权利要求书限制。除非另外限定,否则本文使用的所有科技术语的含义与本领域技术人员通常理解的相同。根据本公开,可能会采用本领域技术内的常规分子生物学、微生物学和重组DNA技术。

[0688] 本公开在其应用中不限于下面描述中所述的或附图中所图示的构建的细节以及各组分的布置方式。本公开还可以有其它实施方案并且能够以多种方式实施或能够以多种方式进行。另外,本文所用的措辞和术语是为了描述的目的而不应该视为限制。

[0689] 优选地,本文所用的术语如以下文献中所述进行定义:“*A multilingual glossary of biotechnological terms: (IUPAC Recommendations)*”,Leuenberger, H.G.W, Nagel, B. 和 Kolbl, H. 编辑(1995), *Helvetica Chimica Acta*, CH-4010 Basel, Switzerland。

[0690] 在本说明书通篇中引用了若干文献。本文中引用的每篇文献(包括所有专利、专利申请、科学出版物、制造商说明书、指南、GenBank登录号序列提交等),无论上文或下文,特

此以引用方式全文并入本文。

[0691] 本文所述的实施例示例说明本公开并且无意于对其施加限制。已经根据本公开描述了本公开的不同实施方案。可以对本文所描述和示例说明的技术作出多种修改和改变而不脱离本公开的精神和范围。因此,应当理解实施例仅是示例性的并且不限制本公开的范围。

[0692] 表10:提供了AacSHC的登录号

[0693]	SHC 来源菌株 (SHC 名称)	参照	登录号
	酸热脂环酸 芽孢杆菌 (WT AacSHC)	JP2009-060799 (花王公司) Neumann 等人 Biol Chem (1986) 367; 723-729	NBRC15652

[0694] 表11:提供了ZmoSHC的登录号

[0695]	SHC 来源菌株 (SHC 名称)	参照	菌株及 登录号	根据 WO 2010139719 US2012/0135477 的 SEQ ID No.
	运动发酵单 胞菌 (WT ZmoSHC)	W02010139719 US2012/0135477	ATCC31821 PF62207_2 Genpept 登录 号 AAV90172	SEQ ID No. 1 SEQ ID No. 2
[0696]		Reipen 等人(1995) Microbiology 141: 155-161.	EMBL/Genbank 登录号 X80766	

[0697] 表12:示出了来自W02010139719的其它SHC酶的来源

[0698]

SHC 来源菌株 (SHC 名称)	参照	登录号	根据 WO 2010139719 US2012/0135477 的 SEQ ID No.
慢生型大豆根瘤菌 (WT BjpSHC)	W02010139719 US2012/0135477	PF62207-5	SEQ ID No. 5
双向伯克霍尔德氏 菌	W02010139719 US2012/0135477		SEQ ID No. 6
双向伯克霍尔德氏 菌	W02010139719 US2012/0135477		SEQ ID No. 7
炭疽芽孢杆菌	W02010139719 US2012/0135477		SEQ ID No. 8
桉木弗兰克氏菌	W02010139719 US2012/0135477		SEQ ID No. 9
<i>Rhodopseudomonas</i> <i>palent</i>	W02010139719 US2012/0135477		SEQ ID No. 10

[0699] 表13:WT AacSHC氨基酸和核苷酸SEQ ID No.

[0700]

氨基酸 SEQ ID No.	菌株	核苷酸 SEQ ID No.

[0701]

SEQ ID No 1	WT AacSHC 酸热脂环酸芽孢杆菌 AB007002.1	169
SEQ ID No 1	(GI: 218288697) 酸热脂环酸芽孢杆菌 ZP-03492960.1	165

[0702] 表14:AacSHC衍生物氨基酸和核苷酸SEQ ID No.

[0703]

氨基酸 SEQ ID No.	突变	SHC 衍生物 名称	核苷酸 SEQ ID No.
SEQ ID No. 5	T77A		6
SEQ ID No. 7	I92V		8
SEQ ID No. 9	F129L		10
SEQ ID No. 11	M132R		12
SEQ ID No. 13	A224V		14
SEQ ID No. 15	I432T		16
SEQ ID No. 17	Q579H		18
SEQ ID No. 19	F601Y		20
SEQ ID No. 171	F605W		170
SEQ ID No. 21	M132R + A224V + I432T	215G2	22
SEQ ID No. 23	M132R + I432T	SHC26	24
SEQ ID No. 25	F601Y	SHC3	26
SEQ ID No. 27	T77A + I92V + F129L	111C8	28
SEQ ID No. 29	Q579H + F601Y	101A10	30
SEQ ID No. 31	F129L	SHC10	32
SEQ ID No. 33	F129L + F601Y	SHC30	34

[0704]

SEQ ID No. 35	F129L + M132R + I432T	SHC31	36
SEQ ID No. 37	M132R + I432T + F601Y	SHC32	38
SEQ ID No. 39	F129L + M132R + I432T + F601Y	SHC33	40

[0705] 表15:WT ZmoSHC1和ZmoSHC1衍生物的氨基酸和核苷酸SEQ ID No.

[0706]

氨基酸 SEQ ID No.	菌株/突变	SHC 衍生物 名称	核苷酸 SEQ ID No.
SEQ ID No. 2	WT ZmoSHC1 ZmoSHC1 (GI: 56552444) ATCC 31821]	参照 ZmoSHC1 序列 WO 2010/139719 中的 SEQ ID No. 2	166
SEQ ID No. 41	S129A		42
SEQ ID No. 43	V145V		44
SEQ ID No. 45	F182L		46
SEQ ID No. 47	Y185R		48
SEQ ID No. 49	G282V		50
SEQ ID No. 51	I498T		52
SEQ ID No. 53	H646H		54
SEQ ID No. 55	F668Y		56
SEQ ID No. 57	Y185R + G282V + I498T	215G2 ZM1	58
SEQ ID No. 59	Y185R + I498T	SHC26 ZM1	60
SEQ ID No. 61	F668Y	SHC3 ZM1	62
SEQ ID No. 63	S129A + V145V+ F182L	111C8 ZM1	64
SEQ ID No. 65	H646H + F668Y	101A10 ZM1	66
SEQ ID No. 67	F182L	SHC10 ZM1	68
SEQ ID No. 69	F182L + F668Y	SHC30 ZM1	70

[0707]

SEQ ID No. 71	F182L + Y185R + I498T	SHC31 ZM1	72
SEQ ID No. 73	Y185R + I498T + F668Y	SHC32 ZM1	74
SEQ ID No. 75	F182L + Y185R + I498T + F668Y	SHC33 ZM1	76
SEQ ID No. 173	F698W		172

[0708] 表16:WT ZmoSHC2和ZmoSHC2衍生物的氨基酸和核苷酸SEQ ID No.

[0709]

氨基酸 SEQ ID No.	菌株/突变	SHC 衍生物 名称	核苷酸 SEQ ID No.
SEQ ID No. 3	WT ZmoSHC2 ZmoSHC2 (GI: 677871)	参照 ZmoSHC1 序 列	167
SEQ ID No. 77	G85A		78
SEQ ID No. 79	V100V		80
SEQ ID No. 81	F137L		82
SEQ ID No. 83	I140R		84
SEQ ID No. 85	V233V		86
SEQ ID No. 87	I450T		88
SEQ ID No. 89	N598H		90
SEQ ID No. 91	F620Y		92
SEQ ID No. 93	I140R + V233V + I450T	215G2ZM2	94
SEQ ID No. 95	I140R + I450T	SHC26 ZM2	96
SEQ ID No. 97	F620Y	SHC3 ZM2	98
SEQ ID No. 99	G85A + V100V + F137L	111C8 ZM2	100
SEQ ID No. 101	N598H + F620Y	101A10 ZM2	102
SEQ ID No. 103	F137L	SHC10 ZM2	104
SEQ ID No. 105	F137L + F620Y	SHC30 ZM2	106

[0710]

SEQ ID No. 107	F137L + I140R, + I450T	SHC31 ZM2	108
SEQ ID No. 109	I140R + I450T + F620Y	SHC32 ZM2	110
SEQ ID No. 111	F137L + I140R + I450T + F620Y	SHC33 ZM2	112
SEQ ID No. 175	F624W		174

[0711] 表17:WT BjpSHC1和BjpSHC1衍生物的氨基酸和核苷酸SEQ ID No.

[0712]

氨基酸 SEQ ID No.	菌株/突变	SHC 衍生物 名称	核苷酸 SEQ ID No.
SEQ ID No. 4	野生型慢生型大豆根瘤菌 SHC (GI: ABQ33590.1)		168
SEQ ID No. 113	A88A		114
SEQ ID No. 115	V104V		116
SEQ ID No. 117	F141L		118
SEQ ID No. 119	Y144R		120
SEQ ID No. 121	V241V		122
SEQ ID No. 123	I459T		124
SEQ ID No. 125	M607H		126
SEQ ID No. 127	F628Y		128
SEQ ID No. 129	Y144R + V241V + I459T	215G2 Bjp	130
SEQ ID No. 131	Y144R + I459T	SHC26 Bjp	132
SEQ ID No. 133	F628Y	SHC3 Bjp	134
SEQ ID No. 135	A88A + V104V + F141L	111C8 Bjp	136
SEQ ID No. 137	M607H + F628Y	101A10 Bjp	138
SEQ ID No. 139	F141L	SHC10 Bjp	140
SEQ ID No. 141	F141L + F628Y	SHC30 Bjp	142

[0713]	SEQ ID No. 143	F141L + Y144R + I459T	SHC31 Bjp	144
	SEQ ID No. 145	M144R + I459T + F628Y	SHC32 Bjp	146
	SEQ ID No. 147	F141L + Y144R + I459T + F628Y	SHC33 Bjp	148
	SEQ ID No. 177	F658W		176

[0714] SEQ ID No.1 (酸热脂环酸芽孢杆菌), AacSHC

[0715] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVP
MVPPEIMFLGKRMLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPIRRRGAKGGGWIFDALDR
ALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQQPWFYALIALKILDMTQHPAFIKWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFQFDNVYYPDVDD
TAVVVWALNTLRPLDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVNTSDLPNHIPFCDFGEVTDPPSEDVTAHVLE
CFGSGFYDDAWKVIRRAVEYLKREQKPDGSGWFRGWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAAARRGVQYL VETQRPDGGWDEPYTGTGFPDGYLG
YTMVRHVFTLALGRYKQAIERR

[0716] SEQ ID No.2 (运动发酵单胞菌), ZmoSHC1

[0717] MGIDRMNSLSRLLMKKIFGAEKTSYKPASDTIIGTDTLKRPNRRPEPTAKVDKTIKTMGNSLNNTLVS
ACDWLIGQQKPDGHWVGAVESNASMEAEWCLALWFLGLEDHPLRPRLGNALLEMQREDGSWGVYFGAGNGDINATVE
AYAALRSLGYSADNPVLKKAIAEWIAEKGLKNIRVFTRYWLALIGEPWEKTPNLPEI IWFDPDNFVFSIYNFAQWA
RATMVPIA ILSARRPSRPLRPQDRLELDFPEGRARFDYELPKKEGIDLWSQFFRTTDRGLHWVQSNLLKRNSLREAA
IRHVLEWIIIRHQDADGGWGGIQQPWWYGLMALHGEYQLYHPVMAKALSALDDPGWRHHRGESSWIQATNSPVWDTM
LALMALKDAKAEDRFTPEMDKAADWLLARQVKVKGDWISIKLPDVEPGGWAFEYANDRYPD TDDTAVALIALSSYRDK
EEWQKKGVEDAITRGVNWLIAMQSECGGWGAFDKDNNRSILSKIPFCDFGESIDPPSVDVTAHVLEAFGLGLSRDM
PVIQKAIDYVRSEQAEGAWFGRWGVNYIYGTGAVLPALAAIGEDMTQPYITKACDWLVAHQQEDGGWGESCSSYME
IDSIGKGPTTSPQTAWALMGLIAANRPEDYEAIKGCCHYLIDRQE QDGSWKEEFTGTGFPGYGVGQTIKLDDPALS
KRLLQGAELSRAFMLRYDFYRQFFPIMALSRAERLIDLNN

[0718] SEQ ID No.3 (运动发酵单胞菌), ZmoSHC2

[0719] MTVSTSSAFHHSPLSDDVEPIIQKATRALLEKQQQDGHVFELEADATIPA EYILLKHYLGEPEDELEIE
AKIGRYLRRIQGEHGGWSLFYGGDLDSATVKAYFALKMIGDSPDAPHMLRARNEILARGGAMRANVFTRIQLALFG
AMSWEHVPQMPVELMLMPEWFPVHINKMAYWARTVLVPLLVLQALKPVARNRRGILVDELFPDVLPTLQESGDPIW
RRFFSALDKVLHKVEPYWPKNMRAKAIHSCVHFVTERLNGEDGLGAIYPAIANSVMMYDALGYPENHPERAIARRAV
EKLMLVDGTEDQGDKVEYQCPLSPIWDTALVAHAMLEVGGDEAEKSAISALSWLKPQQI LDVKGDWAWRRPDLRPG
GWAFQYRNDYYPDVDDTAVVTMAMDRAAKLSDLHDDFEESKARAMEWTIGMQSDNGGWGAFDANN SYTYLNNIPFAD
HGALLDPPTVDVSARCVSMMQAQAGISITDPKMKAAVDYLLKEQEEDGSWFRGWGVNYIYGTWSALCALNVAALPHDH
LAVQKAWWLKTIQNE DGGWGENCDSYALDYSGYEPMSTASQTAWALLGLMAVGEANSEAVTKGINWLAQNQDEEG
LWKEDYSSGGGFPRVFYLYRHGYSKYFPLWALARYRNLLKANQPIVHYGM

[0720] SEQ ID No.4 (慢生型大豆根瘤菌), BjpSHC

[0721] MVTVTSSASARATRDPGNYQTALQSTVRAAADWLIANQKPDGHWVGRAESNACMEAQWCLALWFMGLEDH

PLRKRLGQSLLDSQRPDGAWQVYFGAPNGDINATVEAYAALRSLGFRDDEPAVRRAREWIEAKGGLRNIRVFTRYWL
ALIGEWPWEKTPNIPPEVIWFPLWFPFSIYNFAQWARATLMPIAVLSARRPSRPLPPENRLDALFPHGRKAFDYELP
VKAGAGGWDRFFRGADKVLHKLQNLGNRLNLGLFRPAATSRVLEWMIRHQDFDGAWGGIQPPWIYGLMALYAEGYPL
NHPVLAKGLDALNDPGWRVDVGDATYIQATNSPVWDTILTLLAFDDAGVLGDYPEAVDKAVDWVLQRQVRVPGDWSM
KLPHVKPGGWAFHEYANNYPDTDDTAVALIALAPLRHDPKWKAKGIDEAIIQLGVDWLIGMQSQGGGWGAFDKDNNQK
ILTKIPFCDYGEALDPPSVDVTAHIIEAFGKLGISRNHPSMVQALDYIRREQEPSGPWFGRWGVNYVYGTGAVLPAL
AAIGEDMTQPYIGRACDWLVAHQADGGWGESCASYMDVSAVGRGTTTASQTAWALMALLAANRPQDKDAIERGCMW
LVERQSAGTWDEPEFTGTGFPYGVGQTIKLNDPALSQRMLMQPELSRAFMLRYGMYRHYFPLMALGRALRPQSHS

[0722] SEQ ID No.149 (双向伯克霍尔德氏菌)

[0723] MNDLTEMATLSAGTVPAGLDAVASATDALLAAQNADGHWVYELEADSTIPAHEYVLLVHYLGETPNLEL
EQKIGRYLRRVQQADGGWPLFTDGAPNISASVKAYFALKVIGDDENAEHMQRRARAIQAMGGAEMS NVFTRIQLALY
GAIPWRAVPMPVEIMLLPQWFPFHL SKVSYWARTVIVPLLVLNAKRPIAKNPRGVRIDELFVDPPVNAGLLPRQGH
QSPGWFAFFRVVDHALRAADGLFPNYTRERAIRQAVSFVDERLNGEDGLGAIYPAMANAVMMYDVLGYAEDHPNRAI
ARKSIEKLLVVQEDEAYCQPCLSPVWDTSLAAHALLETGDARAEAVIRGLEWLRPLQILDVRGDWISRRPHVRPGG
WAFQYANPHYPDVDDTAVVAVAMDRVQKLKHNDAFRDSIARAREWVVGMSDGGWGAFEPENTQYYLNNIPFSDHG
ALLDPPTADVSGRCLSMLAQLGETPLNSEPARRALDYMLKEQEPDGSWYGRWGMNYVYGTWTALCALNAAGLTPDDP
RVKRGAAQWLLSIQNKDGGWGEGDGSYKLNRYRGFEQAPSTASQTAWALLGLMAAGEVNNPAVARGVEYLIAEQKEHGL
WDETRFTATGFPRVFYLRHYGRKFFPLWALARYRNLKRNNATRVTFGL

[0724] SEQ ID No.151 (双向伯克霍尔德氏菌)

[0725] MIRRMNKS GPSPWSALDAAIARGDALMRLQQPDGSWCFELES DATITA EYILMMHFMDKID DARQEKM
ARYLRAIQRDLTHGGWDLYVDGDPDVSCSVKAYFALKAAGDSEHAPHMVRARDAILELGAARS NVFTRILLATFGQ
VPWRATPFMPIEFVLPKWVPISMYKVAYWARTTMVPLLVLCSLKARARNPRNIAIPELFVTPPDQERQYFPPARGM
RRAFLALDRVVRHVEPLLKRLRQRAIRHAQAWCAERMNGEDGLGIFPPIVYSYQMMDV LGYPDDHPLRRDCENAL
EKLLVTRPDGSMY CQPCLSPVWDTAWSTMALEQARGVAVPEAGAPASALDEL DARIARAYDWLAERQVNDLRGDWIE
NAPADTQPGGWAFQYANPYPDIDDSAVVTAMLDRRGRTHR NADGSHPYAARVARALDWMRGLQSRNGGFAAFDADC
DRLYLNAIPFADHGALLDPPTEDVSGRVLLCFGVTKRADDRASLARAIDYVKRTQQPDGSWWGRWGTYNYLGTWSVL
AGLALAGEDPSQPYIARALAWLRARQHADGGWGETNDSYIDPALAGTNAGESTS NCTAWALLAQMAFGDGESESVR
GIAYLQSVQQDDGFWWHRSHNAPGFPRIFYLKYHGYTAYFPLWALARYRRLAGGVSAAGAHAVPASTGADAALA

[0726] SEQ ID No.153 (炭疽芽孢杆菌)

[0727] MLLYEKAHEEIVRRATALQTMQWQDGTWRFCFEGAPLTDCHMIFLLKLLGRDKEIEPFVERVASLQTNE
GTWKLHEDEVGGNLSATI QSYAALLASKKYTKEDANMKRAENFIQERGGVARAHFMTKFL LAIHGEY EYPSLFHLPT
PIMFLQNDSPFSIFELSSSARIHLIPMMLCLNKRFRVGGKLLPNLNHIAGGGGEWFREDRSPVFQTLSDVKQIISY
PLSLHHKGYEEIERFMKERIDENGLYSYATASFYMIYALLALGHSLQSSMIQKAIAGITSYIWKMERGNHLQNSPS
TVWDTALLSYALQEAQVSKDNKMIQNATAYLLKKQHTKKADWSVHAPALTPGGWGFSDVNTTIPDIDDTTAVLRALA
RSRGNKNIDNAWKKGGNWIKGLQNDGGWGAFEGVTSKLLAKLPIENASDMITDPSTPDITGRVLEFFGTYAQNEL
PEKQIQRAINWLMNVQEENG SWYGKWGICYLYGTWAVMTGLRSLGIPSSNPSLTRAASWLEHIQHEDGGWGESHSS
VEKRFVTLFPSTPSQTAWALDALISYYDTETPAIRKGVSYLLSNPYVNERYP TGTGLPGAFYIRYHSYAHYIPLTL
AHYIKKYRK

[0728] SEQ ID No.155 (桤木弗兰克氏菌) MPAGVGVLVWLDQRLRAMGRPDLVTTTGGA EIPFVLVAA

TASTVGVALALRRPRHPVGWFLALGGVLLLSGGTQGYAAYGAVARPGRLPAADLVAIYADAGFIPWLVVALILHL
TPTGRPLSARWGRIALATAVAGGLWLLVGLVTTETMQPPFQSVTNPLLIGGPLGPLLVARVVLGLATGAGVVLAAVS
LIVRFRRSVDVERRQLLWVAVAAVPLPVLMAASFAASYAGNNTAAGLAAATLIGLLAIGAGLAIGQYHLYDVEEILS
RAVTYLLVSGLLAASYATVVIVVQSLAGRTGRSQISAVLATLAAVAVTAPAYRKIQEGVDRRFSRRRFFETLQVIRR
YLRDPDPDVAVEEVLRRALGDPTLAVAYLVDDRQWVSADGQPANPGNSFMAAVEVYRRGRPIARVTFDRGRAQPGL
VRAAATAATAELDNAGLRAAVALQLVEVRQSRTRIAAAQFAERRTIERNLHDGAQQRLLALALQLRAVQLGGDEASL
RQAISTGIDQLQAAVVELRELANGLHPAVLADGGALAAALDDVAARTPVPIKISAPDRRYPPDLEAAAWFIACEAMAN
AVKHAHPTTIAVDVSAPDGQLIVEVRDDGIGGAQPSGPGLRGIADRAEAFGGSLTVHTDPGTGTTIRALLHRRSPLS
SGRRSVMIEGCVDDVAVRRFRCSRSGSGSRRRRSSWRCGGICGSRCTGMSRSCSRNAASKLIT

[0729] SEQ ID No.157 (Rhodopseudomonas palent)

[0730] MDSILAPRADAPRNIDGALRESVQQAADWLVANQKPDGHWVGRAETNATMEAQWCLALWFLGLEDHPLR
VRLGRALLDTQRPDGAWHVFYGA PNGDINATVEAYAALRSLGHRDDEEPLRKARDWILSKGGLANIRVFTRYWLALI
GEWPWEKTPNILPEVIWLPTWFPFSIYNFAQWARATLMPIAVLSAHRPSRPLAPQDRDLALFPQGRDSFNIDLPARL
GAGVWDVIFRKIDTILHRLQDWGARRGPHGIMRRGAIDHVLQWII RHQDYDGSWGGI QPPWIYGLMALHTEGYAMTH
PVMKALDALNEPGWRIDIGDATFIQATNSPVWDTMLSLLAFDDAGLGERYPEQVERAVRWLKRQVLVPGDWSVKL
PDVKPGGWAFEYANNFYPTDDTSVALMALAPFRHDPKWQAEGIEDAIQRGIDWL VAMQCKEGGWGAFDKDNDKKIL
AKIPFCDFGEALDPPSADVTAHII EAFKVLDRNHPSIVRALDYLKREQEPEGPWFGRWGVNYVYGTGAVLPALAA
IGEDMRQPYIARACDWLIARQQANGGWGESCVSYMDAKQAGEGTATASQTAWALMALIAADRPQDRDAIERGCLYLT
ETQRDGTWQEVHYTGTGFPYGVGQTIKLNPLLSKRLMQPELSRSFMLRYDLRYHYFPMMAIGRVLRRQGRDSGH

[0731] SEQ ID No.159 (天蓝色链霉菌)

[0732] MTATTDGSTGASLRPLAASASDITITIPAAAAGVPEAAARATRRATDFLLAKQDAEGWWKGDLETNVTM
DAEDLLLRLQFLGIQDEETTRAAALFIRGEQREDGTWATFYGGPGELSTTIEAYVALRLAGDSPEAPHMARAAEWIRS
RGGIASARVFTRIWLALFGWWKDDLPPELPELIYFPTWVPLNIYDFGCWARQTIVPLTIVSAKRVP RPAPFPLDEL
HTDPARPNNPRPLAPVASWDGAFQRIDKALHAYRKVAPRRLRRAAMNSAARWI IERQENDGCWGGI QPPAVYSVIAL
YLLGYDLEHPVMRAGLES LDRFAVWREDGARMIEACQSPVWDTCLATIALADAGVPEDHPQLVKASDWMLGEQIVRP
GDWSVKRPGPPGGWAFEFHNDNYPDIDDTAEVVLALRRVRHHDPERVEKAIGRGVRWNLMQSKNGAWGAFDNDNTS
AFPNRLPFCDFGEVIDPPSADVTAHVVEMLAVEGLAHDPRTRRGIQWLLDAQETDGSWFGWGVNYVYGTGSVIPAL
TAAGLPTSHPAIRRAVRWLESVQNEGGWGEDLSYRYVREWSGRGASTASQTGWALMALLAAGERDSKAVERGVAV
LAATQREDGSWDEPYFTGTGFPWDFSINYNLYRQVFPLTALGRYVHGEPFAKKPRAADAPAEAAPAEVKGS

[0733] SEQ ID No.169

[0734] ATGGCTGAGCAGTTGGTGGAAGCGCCGCCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCCGCGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTGCGCGACGAGGAGCCGATGCAGAAAGCGCTCCGGTTCATTACAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGTTCACGCGGATGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGAGAG
TCATGTTCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTTCGCGTCCCCGAGCGGGCGCGGTGCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGCGCTGCACGGGTATC

AGAAGCTGTCGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGA CTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTACGCCGCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTGGGGCGCCTACGACGTGACAAACACGAGCGATCTCCCGAACCACATCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGTGTTTCGGCAGCTTCG
GGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTGCGCGCTGAAGGCGGTGGGATCGACAC
GCGCGAGCCGTACATTCAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGGAGACCCCGTCGACAGGCGCTGGGCGCTGATGGCG
CTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGTGGGATGAGCCGTACTACACCGGCACGGGCTTCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

[0735] 变体101A10 (SEQ ID No. 30)

[0736] ATGGCTGAGCAGTTGGTGAAGCACCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCCGCGACCTCGACACGACCATCGAGGCGTACGTGCGGCTCAAGTAT
ATCGGCATGTGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTACAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGTTCACGCGGATGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGAGA
TCATGTTCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGTGTTCCCGCTGCCCCAGCGGGCGCGGTGCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGCGCTGCACGGGTATC
AGAAGCTGTCGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGA CTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTACGCCGCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTGGGGCGCCTACGACGTGACAAACACGAGCGATCTCCCGAACCACATCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGTGTTTCGGCAGCTTCG
GGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTGCGCGCTGAAGGCGGTGGGATCGACAC
GCGCGAGCCGTACATTCAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGGAGACCCCGTCGACAGGCGCTGGGCGCTGATGGCG

CTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGCGTGCATTACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGCTGGGATGAGCCGTACTACACCGGCACGGGCTACCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

[0737] 变体101A10 (SEQ ID No.29)

[0738] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVP
MVPPEIMFLGKRMLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDR
ALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQQPWFYALIALKILDMTQHPAFIKWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFQFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYVDNTSDLPNHIPFCDFGEVTDPPSEDVTAHVLE
CFGSFQYDDAWKVIRRAVEYLKREQKPDGSWFRGWVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAAARRGVHYLVETQRPDGGWDEPYTGTGYPGDFYLG
YTMYRHVFPTLALGRYKQAIERR

[0739] 变体111C8 (SEQ ID No.28)

[0740] ATGGCTGAGCAGTTGGTGGGAAGCGCCGCCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCGCGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGACCGTCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTGCGCGCAGGAGGACCGATGCAGAAGGCGCTCCGGTTCATTAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGCTCACGCGGATGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGA
TCATGTTCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTCCCGCTGCCCCAGCGGGCGCGCTGCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGCGCTGCACGGGTATC
AGAAGCTGTGCGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTACGCCGCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCGGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTGGGGCGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACATCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGTGTTTCGGCAGCTTCG
GGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTGCGCGCTGAAGGCGGTGGGATCGACAC
GCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGGAGACCCCGTCGACAGGCCTGGGCGCTGATGGCG
CTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGCGGTGCAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGCTGGGATGAGCCGTACTACACCGGCACGGGCTTCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

[0741] 变体111C8 (SEQ ID No.27)

[0742] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGAWALYPGGPPDLDTTVEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVLTRMWLALVGEYPWEKVP
MVPPEIMFLGKRMPNLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDR
ALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQQPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVNTSDLPNHIPFCDFGEVTDPPSEDVTAHVLE
CFGSGFYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAAARRGVQYL VETQRPDGGWDEPYTGTGFPGDFYLG
YTMYRHVFPTLALGRYKQAIERR

[0743] 变体SHC215G2 (SEQ ID No.22)

[0744] ATGGCTGAGCAGTTGGTGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCCGCGGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGTTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGAGAG
TCATGTTCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTTCGCCGCTGCCCCGAGCGGGCGCGCTGCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTTCGACGCGCTCGACCGGTGCTGCACGGGTATC
AGAAGCTGTGCGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGA
GACGGCAGCTGGGGCGGGATTACGCCGCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGTTCCGGGCGACTGGGCGGTGAAGCGCCCCGA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTGCTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGACGAGCGCCGAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTGGGGCGCTACGACGTCGACAACACGAGCGATCTCCGAACACACCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCACGTGCTCGAGTGTTCGGCAGCTTCG
GGTACGATGACGCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTGCGCGCTGAAGGCGGTGCGGATCGACAC
GCGCGAGCCGTACATTCAAAGGCGCTCGACTGGGTGCGAGCAGCATCAGAACCCGGACGGCGGTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGTAAGGGCGCGAGCACCCCGTCGACAGGCCTGGGCGCTGATGGCG
CTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGTGGGATGAGCCGTACTACACCGGCACGGGCTCCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

[0745] 变体SHC215G2 (SEQ ID No.21)

[0746] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVP

MVPPEIMFLGKRMP LNIYEF GSWARATVVALSIVMSRQPVFPLPERARVPELYETDVP PRRRGAKGGGGWIFDALDR
VLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDNDTS DLPNHTPFCDFGEVTDPPSEDVTAHVLE
CFGSGFYDDAWKVIRRAVEY LKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREP YIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAAARRGVQYL VETQRPDGGWDEPYTGTGFP GDFYLG
YTMYRHVFPTLALGRYKQAIERR

[0747] 变体SHC3 (SEQ ID No. 26)

[0748] ATGGCTGAGCAGTTGGTGGAAGCGCCGCCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGTTCACGCGGATGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGAGA
TCATGTTCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTTCCTCGCTGCCGAGCGGGCGCGCTGCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCGGGCGCGGTGCCAAGGGAGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGCGCTGCACGGGTATC
AGAAGCTGTGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTACGCCGCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGACGTGGACGACACGGCCGTCGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGACGAGCGCCGAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTGGGGCGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACATCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCACGTGCTCGAGTGTTTCGGCAGCTTCG
GGTACGATGACGCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGACGGCAGCTGG
TTCGGTCTTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTGCGCGCTGAAGGCGGTGGGATCGACAC
GCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGACGGCGGTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGTAAGGGCGGAGACCCCGTCGAGACGGCCTGGGCGCTGATGGCG
CTCATCGCGGGCGGAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGCTGGGATGAGCCGTACTACACGGCACGGGCTACCCAGGGATTTCTACCTCGGCTACACCATGTACCGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

[0749] 变体SHC3 (SEQ ID No. 25)

[0750] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVP
MVPPEIMFLGKRMP LNIYEF GSWARATVVALSIVMSRQPVFPLPERARVPELYETDVP PRRRGAKGGGGWIFDALDR
ALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD

TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHIPFCDFGEVTDPPSEDVTAHVLE
CFGSFYDDAWKVIRRAVEYLLKREQKPDGWSFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAAARRGVQYL VETQRPDGGWDEPYTGTGYPGDFYLG
YTMYRHVFPTLALGRYKQAIERR

[0751] 变体SHC10 (SEQ ID No.32)

[0752] ATGGCTGAGCAGTTGGTGAAGCGCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCCGGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTGCGCGACGAGGAGCCGATGCAGAAAGCGCTCCGGTTCATTACAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGCTCACGCGGATGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGA
TCATGTTCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTTCCTCGCTGCCCGAGCGGGCGCGCTGCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGCGCTGCACGGGTATC
AGAAGCTGTGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTACGCCGCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTCCGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGAGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTTGGGGCGCTACGACGTCGACAACACGAGCGATCTCCCGAACCACATCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCACGTGCTCGAGTGTTTCGGCAGCTTCG
GGTACGATGACGCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTGCGCGCTGAAGGCGGTGGGATCGACAC
GCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTGCGAGCAGCATCAGAACCCGGACGGCGGTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGGAGACCCCGTCGACAGGCGCTGGGCGCTGATGGCG
CTCATCGCGGGCGGACGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTCGTGGAGACGCGCGCCCGGA
CGGCGGCTGGGATGAGCCGTACTACACGGCACGGGCTTCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCC
ACGTGTTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

[0753] 变体SHC10 (SEQ ID No.31)

[0754] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVLTRMWLALVGEYPWEKVP
MVPPEIMFLGKRMLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAAGGGGWIFDALDR
ALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQQPWFYALIALKILDMTQHPAFIKWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHIPFCDFGEVTDPPSEDVTAHVLE
CFGSFYDDAWKVIRRAVEYLLKREQKPDGWSFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAAARRGVQYL VETQRPDGGWDEPYTGTGYPGDFYLG

YTMYRHVFPTLALGRYKQAIERR

[0755] 变体SHC26 (SEQ ID No.24)

[0756] ATGGCTGAGCAGTTGGTGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTGCGCGCAGGAGGACCGATGCAGAAGGCGCTCCGGTTCATTAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGTTACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGA
TCATGTTCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTCCCGCTGCCCAGCGGGCGCGCTGCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGCGCTGCACGGGTATC
AGAAGCTGTGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTACGCCGCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCGGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTTCGCTTCCAGTTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGACGAGCGCCGAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTTGGGGCGCTACGACGTCGACAACACGAGCGATCTCCGAACACACCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCACGTGCTCGAGTGTTTCGGCAGCTTCG
GGTACGATGACGCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTGCGCGCTGAAGGCGGTGGGATCGACAC
GCGCGAGCCGTACATTAAAAGGCGCTCGACTGGGTGAGCAGCATCAGAACCCGGACGGCGGTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGACCCCGTCGAGACGGCCTGGGCGCTGATGGCG
CTCATCGCGGGCGGACGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGCTGGGATGAGCCGTACTACACCGGCACGGGCTTCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

[0757] 变体SHC26 (SEQ ID No.23)

[0758] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMPLNIEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDR
ALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQQPWFYALIALKILDMTQHPAFIKWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDNDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLE
CFGSFYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAAARRGVQYL VETQRPDGGWDEPYTGTGFPGDFYLG
YTMYRHVFPTLALGRYKQAIERR

[0759] 变体SHC30 (SEQ ID No.34)

[0760] ATGGCTGAGCAGTTGGTGAAGCGCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC

TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGCTACGCGGATGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGA
TCATGTTCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTTCCTCGCTGCCCCGAGCGGCGCGCTGCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGCGCTGCACGGGTATC
AGAAGCTGTGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTACGCCGCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCGGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTTGGGGCGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACATCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGTGTTTCGGCAGCTTCG
GGTACGATGACGCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTGCGCGCTGAAGGCGGTGGGATCGACAC
GCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTGAGCAGCATCAGAACCCGGACGGCGGTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGACCCCGTCGAGACGGCCTGGGCGCTGATGGCG
CTCATCGCGGGCGGACGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGCTGGGATGAGCCGTACTACACCGGCACGGGCTACCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

[0761] 变体SHC30 (SEQ ID No.33)

[0762] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVLTRMWLALVGEYPWEKVP
MVPPEIMFLGKRMLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDR
ALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQQPWFYALIALKILDMTQHPAFIKWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYVDNTSDLPNHIPFCDFGEVTDPPSEDVTAHVLE
CFGSFQYDDAWKVIRRAVEYLLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAAARRGVQYL VETQRPDGGWDEPYTGTGYPGDFYLG
YTMYRHVFPTLALGRYKQAIERR

[0763] 变体SHC31 (SEQ ID No.36)

[0764] ATGGCTGAGCAGTTGGTGAAGCGCCGCCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT

ATCGGCATGTGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGCTACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGAGA
TCATGTTCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTTCGCGTCCCCGAGCGGGCGCGCTGCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGCGCTGCACGGGTATC
AGAAGCTGTGCGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGA
GACGGCAGCTGGGGCGGGATTACGCCGCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGTGTGGGACACGGGCTCGCCGTGCTCGCGCTGCGCGCTGCGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCGGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTTCGCTTCCAGTTTCACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGACGAGCGCCGAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTGGGGCGCTACGACGTCGACAACAGAGCGATCTCCGAACACACCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCACGTGCTCGAGTGTTCGGCAGCTTCG
GGTACGATGACGCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTGCGCGCTGAAGGCGGTGGGATCGACAC
GCGCGAGCCGTACATTCAAAGGCGCTCGACTGGGTGAGCAGCATCAGAACCCGGACGGCGGTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGCACCCGTCGACAGCGCCGTTGGGCGCTGATGGCG
CTCATCGCGGGCGGACGGGCGGAGTCCGAGGCCGCGCGCCGCGCGGTGCAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGTGGGATGAGCCGTACTACACCGGCACGGGCTTCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

[0765] 变体SHC31 (SEQ ID No.35)

[0766] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVLTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMPNLIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPRRRGAAGGGGWIFDALDR
ALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQQPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTRLRLPDERRRRDAMTKGFRWIVGMQSSNGWGAYDVNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLE
CFGSFYDDAWKVIRRAVEYLKREQKPDGSWFGWGVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAAARRGVQYL VETQRPDGGWDEPYTGTGFPDGYLG
YTMYRHVFPTLALGRYKQAIERR

[0767] 变体SHC32 (SEQ ID No.38)

[0768] ATGGCTGAGCAGTTGGTGAAGCGCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCCGCGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGTTACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGAGA
TCATGTTCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC

TCGATTGTGATGAGCCGCCAGCCGGTGTTCCTCGCTGCCCCGAGCGGGCGCGCTGCCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGCGCTGCACGGGTATC
AGAAGCTGTCGGTGACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTACGCCGCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCGGGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGACGAGCGCCGAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTGGGGCGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACACCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCACGTGCTCGAGTGTTCGGCAGCTTCG
GGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTGCGCGCTGAAGGCGGTGGGATCGACAC
GCGCGAGCCGTACATTCAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGGAGACCCCGTCGAGACGGCCTGGGCGCTGATGGCG
CTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGCGGTGCAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGTGGGATGAGCCGTACTACACCGGCACGGGTACCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

[0769] 变体SHC32 (SEQ ID No. 37)

[0770] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDR
ALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLE
CFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPIYQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAAARRGVQYL VETQRPDGGWDEPYTGTGYPGDFYLG
YTMYRHVFPTLALGRYKQAIERR

[0771] 变体SHC33 (SEQ ID No. 40)

[0772] ATGGCTGAGCAGTTGGTGAAGCGCCGGCCTACGCGGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGCTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGGAGA
TCATGTTCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTTCCTCGCTGCCCCGAGCGGGCGCGCTGCCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGCGCTGCACGGGTATC
AGAAGCTGTCGGTGACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGACTGGTTGCTCGAGCGCCAGGCCGGA

GACGGCAGCTGGGGCGGGATTTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGGCCGATCACGAC
CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCGGGGCGACTGGGCGGTGAAGCGCCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCTGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTGGGGCGCTACGACGTGACACAACAGAGCGATCTCCCGAACCACACCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTACAGAGGACGTCACCGCCACGTGCTCGAGTGTTCGGCAGCTTCG
GGTACGATGACGCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCGTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTGCGCGCTGAAGGCGGTGCGGATCGACAC
GCGCGAGCCGTACATTCAAAGGGCGCTCGACTGGGTGCGAGCAGCATCAGAACCCGGACGGCGGTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGCGAGCACCCCGTCGACAGGCCCTGGGCGCTGATGGCG
CTCATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGCGGTGCAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGTGGGATGAGCCGTACTACACCGGCACGGGCTACCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

[0773] 变体SHC33 (SEQ ID No. 39)

[0774] MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLL
HEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVLTTRRWLALVGEYPWEKVP
MVPPEIMFLGKRMPNLIYEFGSWARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDR
ALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYG
GWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDD
TAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHTPFCDFGEVTDPPSEDVTAHVLE
CFGSGFYDDAWKVIRRAVEYLKREQKPDGSWFRGWVNYLYGTGAVVSALKAVGIDTREPYIQKALDWVEQHQNPDG
GWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAAARRGVQYL VETQRPDGGWDEPYTGTGYPGDFYLG
YTMRYHVFPTLALGRYKQAIERR

[0775] 变体F605W (SEQ ID No. 170)

[0776] ATGGCTGAGCAGTTGGTGAAGCGCCGGCCTACGCGCGGACGCTGGATCGCGCGGTGGAGTATCTCCTC
TCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTCTGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTT
GTGCCACATTCTCGATCGCGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCGCGAGG
ACGGCACGTGGGCCCTGTACCCGGGTGGGCCCGCGACCTCGACACGACCATCGAGGCGTACGTCGCGCTCAAGTAT
ATCGGCATGTGCGCGACGAGGAGCCGATGCAGAAGGCGCTCCGGTTCATTACAGAGCCAGGGCGGGATCGAGTCGTC
GCGCGTGTTACGCGGATGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGTGCCCATGGTCCCGCCGAGGA
TCATGTTCTCGGCAAGCGCATGCCGCTCAACATCTACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTC
TCGATTGTGATGAGCCGCCAGCCGGTGTTCGCGTCCCCGAGCGGGCGCGCTGCCCGAGCTGTACGAGACCGACGT
GCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTTCGACGCGCTCGACCGGGCGCTGCACGGGTATC
AGAAGCTGTGCGGTGCACCCGTTCCGCCGCGCGGCCGAGATCCGCGCCTTGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCAAGATTCTCGACATGACGCAGCA
TCCGGCGTTTCATCAAGGGCTGGGAAGGTCTAGAGCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGG
CTTCCATCTCGCCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGGCCGATCACGAC

CGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCACGGTTCGCGGCGACTGGGCGGTGAAGCGCCCGAA
CCTCAAGCCGGGCGGGTTCGCGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGTGT
GGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGAGGCGGGACGCCATGACGAAGGGATTCCGCTGGATTGTC
GGCATGCAGAGCTCGAACGGCGGTTGGGGCGCCTACGACGTGACAAACACGAGCGATCTCCCGAACCACATCCCGTT
CTGCGACTTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGTGTTCGGCAGCTTCG
GGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGAATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGG
TTCGGTCTTTGGGGCGTCAATTACCTCTACGGCACGGGCGCGGTGGTGTGCGGCGCTGAAGGCGGTGCGGATCGACAC
GCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAGAACCCGGACGGCGGTGGGGCGAGGACT
GCCGCTCGTACGAGGATCCGGCGTACGCGGGTAAGGGCGGAGACCCCGTCGACAGCGCCTGGGGCGCTGATGGCG
CTCATCGCGGGCGGACGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTCGTGGAGACGCAGCGCCCGGA
CGGCGGTGGGATGAGCCGTACTACACCGGCACGGGCTTCCAGGGGATTGGTACCTCGGCTACACCATGTACCGCC
ACGTGTTTCCGACGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

[0777] 变体F605W (SEQ ID No.171) MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEA
EYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQG
GIESSRVFTRMWLALVGEYPWEKVPMPPEIMFLGKRMPPLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVPEL
YETDVPPRRRGAKGGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQQPWFYALIALKIL
DMTQHPAFIKGWEGLLEYGVLDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQITVPGDWA
VKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLP
NHIPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFRWGVNLYGTGAVVSALKA
VGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVE
TQRPDGGWDEPYTGTGFPDGYLGYTMYRHVFPTLALGRYKQAIERR

[0778] SEQ ID No.166 (ZmoSHC1)

[0779] ATGGGTATTGACAGAATGAATAGCTTAAGTCGCTTGTTAATGAAGAAGATTTTCGGGGCTGAAAAAACC
TCGTATAAACCGGCTTCCGATACCATAATCGGAACGGATACCCTGAAAAGACCGAACCGGCGGCCTGAACCGACGGC
AAAAGTCGACAAAACGATATTCAAGACTATGGGGAATAGTCTGAATAATACCCTTGTTTCAGCCTGTGACTGGTTGA
TCGGACAACAAAAGCCCGATGGTCATTGGGTCGGTGCCGTGGAATCCAATGCTTCGATGGAAGCAGAATGGTGTCTG
GCCTTGTTGGTTTTTTGGGTCTGGAAGATCATCCGCTTCGTCCAAGATTGGGCAATGCTCTTTTGAAATGCAGCGGGA
AGATGGCTCTTGGGGAGTCTATTTTCGGCGCTGGAAATGGCGATATCAATGCCACGGTTGAAGCCTATGCGGCCTTGC
GGTCTTTGGGGTATTCTGCCGATAATCCTGTTTTGAAAAAGCGGCAGCATGGATTGCTGAAAAAGGCGGATTAAAA
AATATCCGTGTCTTTACCGTTATTGGCTGGCGTTGATCGGGGAATGGCCTTGGGAAAAGACCCCTAACCTTCCCCC
TGAAATTATCTGGTTCCCTGATAATTTTGTCTTTTCGATTTATAATTTTGCCCAATGGGCGCGGGCAACCATGGTGC
CGATTGCTATTCTGTCCGCGAGACGACCAAGCCGCCCGTGTGCGCCCTCAAGACCGATTGGATGAACTGTTTCCAGAA
GGCCGCGCTCGCTTTGATTATGAATTGCCGAAAAAAGAAGGCATCGATCTTTGGTCGCAATTTTCCGAACCACTGA
CCGTGGATTACATTGGGTTCAGTCCAATCTGTAAAGCGCAATAGCTTGCGTGAAGCCGCTATCCGTCATGTTTTGG
AATGGATTATCCGGCATCAGGATGCCGATGGCGGTTGGGGTGAATTCAGCCACCTTGGGTCTATGGTTTGATGGCG
TTACATGGTGAAGGCTATCAGCTTTATCATCCGGTGATGGCCAAGGCTTTGTGCGCTTTGGATGATCCCGGTTGGCG
ACATGACAGAGGCGAGTCTTCTTGATACAGGCCACCAATAGTCCGGTATGGGATACAATGTTGGCCTTGATGGCGT
TAAAGACGCCAAGGCCGAGGATCGTTTTACGCCGAAAATGGATAAGGCCGCCGATTGGCTTTTGGCTCGACAGGTC
AAAGTCAAAGGCGATTGGTCAATCAAACCTGCCCGATGTTGAACCCGGTGGATGGGCATTTGAATATGCCAATGATCG

CTATCCCGATACCGATGATACCGCCGTCGCTTTGATCGCCCTTTCCTCTTATCGTGATAAGGAGGAGTGGCAAAAGA
AAGGCGTTGAGGACGCCATTACCCGTGGGGTTAATTGGTTGATCGCCATGCAAAGCGAATGTGGCGGTTGGGGAGCC
TTTGATAAGGATAATAACAGAAGTATCCTTTCCAAAATTCCTTTTTGTGATTTCCGAGAATCTATTGATCCGCCTTC
AGTCGATGTAACGGCGCATGTTTTAGAGGCCCTTTGGCACCTTGGGACTGTCCCGCGATATGCCGGTCATCCAAAAG
CGATCGACTATGTCCGTTCCGAACAGGAAGCCGAAGGCGCGTGGTTTGGTCGTTGGGGCGTTAATTATATCTATGGC
ACCGGTGCGGTTCTGCCTGCTTTGGCGGCGATCGGTGAAGATATGACCCAGCCTTACATCACCAAGGCTTGCGATTG
GCTGGTCGCACATCAGCAGGAAGACGGCGGTTGGGGCGAAAGCTGCTCTTCCTATATGGAGATTGATTCCATTGGGA
AGGGCCCAACCACGCCGTCCCAGACTGCTTGGGCTTTGATGGGGTTGATCGCGGCCAATCGTCCCGAAGATTATGAA
GCCATTGCCAAGGGATGCCATTATCTGATTGATCGCCAAGAGCAGGATGGTAGCTGGAAAGAAGAAGAATTCACCGG
CACCGGATTCCCCGGTTATGGCGTGGGTCAGACGATCAAGTTGGATGATCCGGCTTTATCGAAACGATTGCTTCAAG
GCGCTGAACTGTCACGGGCGTTTATGCTGCGTTATGATTTTTATCGGCAATTCTTCCCGATTATGGCGTTAAGTCGG
GCAGAGAGACTGATTGATTTGAATAATTGA

[0780] 表18:使用Blast和GAP程序算法进行的WT AacSHC酶相对于W02010/0139719 (巴斯夫公司 (BASF)) 中公开的WT SHC酶的序列同一性百分比计算.

[0781]

WT AacSHC (631AA, SEQ ID No. 1) 与以下序列的序列 比对	% 同一性 Blast 图 谱	% 同一 性 GAP	WT SHC (HAC) 的来源	长度 (AA)
WO 2010139719 SEQ ID No. 1	44%	40%	运动发酵单胞菌	725
WO 2010139719 SEQ ID No. 2	44%	40%	运动发酵单胞菌	725
WO 2010139719 SEQ ID No. 5	44%	42%	慢生型大豆根瘤菌	684
WO 2010139719 SEQ ID No. 6	41%	38%	双向伯克霍尔德氏菌 6	657
WO 2010139719 SEQ ID No. 7	40%	37%	双向伯克霍尔德氏菌	682
WO 2010139719 SEQ ID No. 8	34%	32%	炭疽芽孢杆菌	617
WO 2010139719 SEQ ID No. 9	37%	未测 定	桉木弗兰克氏菌	720
WO 2010139719 SEQ ID No. 10	45%	42%	Rhodopseudomonas palent	685
WO 2010139719 SEQ ID No. 11	52%	46%	天蓝色链霉菌	679

[0782] 表19:使用Blast算法以及Huang和Miller的算法进行的WT AacSHC酶相对于Seitz (2012) 中公开的WT ZmoSHC1和WT ZmoSHC2酶的序列同一性百分比计算.

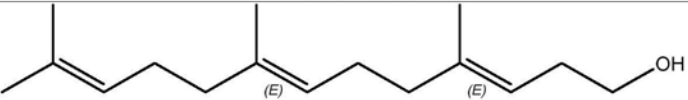
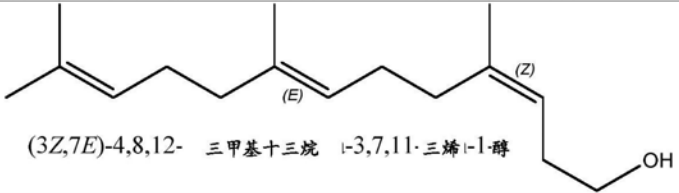
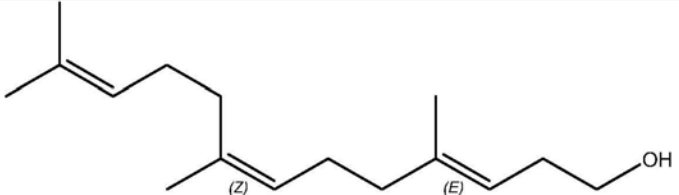
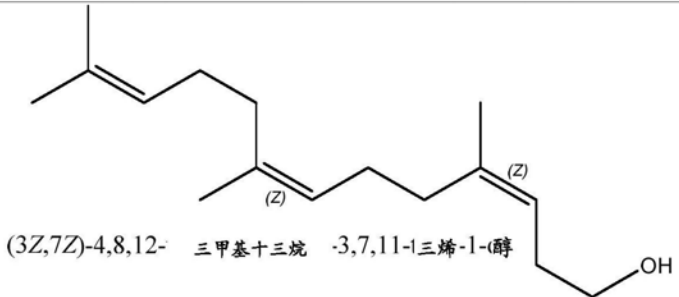
[0783]

比较的 SHC 序列	% 同一性 Blast 图谱	Seitz (2012, 论文) 中 公开的 同一性百分比
AacSHC/HAC 对 ZmoSHC1	44%	41%
AacSHC/HAC 对 ZmoSHC2	39%	37%
AacSHC/HAC 对 AacSHC	98%	未测定
ZmoSHC1 对 ZmoSHC2	36%	34%
AacSHC 对 ZmoSHC2	39%	37%

[0784] Miriam Seitz论文可得自<http://elib.uni-stuttgart.de/handle/11682/1400>

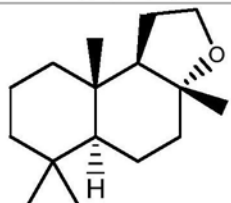
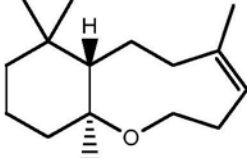
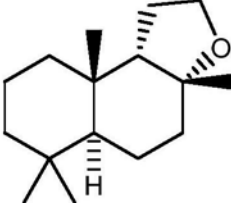
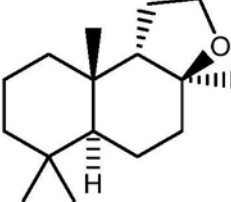
[0785] 表20: 高法呢醇的命名

[0786]

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

[0787] 表21:反应产物的命名

[0788]

化合物	描述	名称和结构
(I)	(-)-降龙涎醚	 $(3aR,5aS,9aS,9bR)$ -3a,6,6,9a-四甲基十二氢萘并[2,1-b]呋喃
(II)	大环	 $(7aS,11aS,Z)$ -5,8,8,11a-四甲基-2,3,6,7,7a,8,9,10,11,11a-十氢萘并[b]oxonine
(III)	9b-表-降龙涎醚	 $(3aR,5aS,9aS,9bS)$ -3a,6,6,9a-四甲基十二氢萘并-[2,1-b]呋喃
(IV)	Escher 等人 (1990)	 $(3aS,5aS,9aS,9bS)$ -3a,6,6,9a-四甲基十二氢萘并-[2,1-b]呋喃

[0789] Escher S, Giersch W., Niclass Y, Bernardinello G 和 Ohloff G (1990). Configuration-odor relationships in 5 β -降龙涎醚. *Helv. Chim. Acta* 73, 1935-1947.

附图说明

[0790] 为了更好地理解本公开, 参考了附图, 其中:

[0791] 图1-4示出了所选择的AacSHC衍生物相对于AacSHC SEQ ID No.1的序列比对。以降序出现, 图1的SEQ ID No.是: SEQ ID No.1、SEQ ID No.29、SEQ ID No.27、SEQ ID No.21、SEQ ID No.19、SEQ ID No.9、SEQ ID No.23、SEQ ID No.33、SEQ ID No.35、SEQ ID No.37和SEQ ID No.39;

[0792] 图5示出了质粒图谱;

[0793] 图6示出了表24中所列出的野生型AacSHC和AacSHC衍生物在标准条件 (pH6.0, 55 $^{\circ}$ C, 0.050% SDS, 细胞达到OD_{650nm}=10) 下的相对HAC活性;

[0794] 图7a示出了采用高法呢醇质量EEH:EZH 87:13以及在96%纯度的高法呢醇(用NMR测定)下,AacSHC衍生物相对于WTAacSHC的HAC活性概况;

[0795] 图7b示出了采用高法呢醇质量EEH:EZH 87:13以及在96%纯度的高法呢醇(用NMR测定)下,AacSHC衍生物相对于WT SHC的相对改善(4小时时(初始速度)和22小时时的产率);

[0796] 图8a示出了采用高法呢醇质量EEH:EZH 92:08以及在100%纯度的高法呢醇(用NMR测定)下,AacSHC衍生物相对于WTAacSHC的HAC活性概况;

[0797] 图8b示出了采用高法呢醇质量EEH:EZH 92:08以及在100%纯度的高法呢醇(用NMR测定)下,AacSHC衍生物相对于WT SHC的相对改善(4小时时(初始速度)和22小时时的产率);

[0798] 图9a示出了采用高法呢醇质量EEH:EZH 66:33以及在76%纯度的高法呢醇(通过NMR测定)下,表24中所列出的AacSHC衍生物相对于WTAacSHC的HAC活性概况;

[0799] 图9b示出了采用高法呢醇质量EEH:EZH 66:33以及在76%纯度的高法呢醇(通过NMR测定)下,AacSHC衍生物相对于WT SHC的相对改善(4小时时(初始速度)和22小时时的产率);

[0800] 图10示出了三种SHC衍生物的HAC活性结果,这些衍生物较于野生型AacSHC/HAC酶显示了大约10倍(215G2)、7倍(SHC26)和6倍(SHC32)的改善。

[0801] 图11示出了由SHC/HAC衍生物(215G2 SHC)和WT AacSHC催化的E,E-高法呢醇向降龙涎醚转化的观察结果。在反应7小时时(初始反应速度的估计),使用变体215G2 SHC时的转化率比使用野生型SHC时实现的转化率高13倍。在反应48小时时,用该变体时的转化率是野生型酶的转化率的8倍。

[0802] 图12示出了当EEH用作原料(以供WT SHC和/或SHC/HAC衍生物生物转化)时所产生的反应产物(降龙涎醚和产物(IV));以及当EE:EZ用作原料时所产生的反应产物((-)-降龙涎醚(I)和产物(II)、(IV)和(III)(见表21);为了易于参考,化合物I-IV可鉴定如下:

[0803] I: (3aR,5aS,9aS,9bR)-3a,6,6,9a-四甲基十二氢萘并[2,1-b]呋喃

[0804] (-)-Ambrox

[0805] II: (7aS,11aS,Z)-5,8,8,11a-四甲基-2,3,6,7,7a,8,9,10,11,11a-十氢苯并[b]oxonine

[0806] IV: (3aS,5aS,9aS,9bS)-3a,6,6,9a-四甲基十二氢萘并[2,1-b]呋喃

[0807] III: (3aRS,5aSR,9aSR,9bSR)-3a,6,6,9a-四甲基十二氢萘并[2,1-b]呋喃

[0808] 9-表-Ambrox

[0809] 图13示出了表25中的降龙涎醚及产物(II)、(IV)和(III)的反应产物GC分析;

[0810] 图14示出了表25中的降龙涎醚及产物(II)、(IV)和(III)的反应产物GC分析;

[0811] 图15提供了在存在Triton X-100或SDS的情况下全细胞生物转化测定法中215G2SHC变体活性的对比数据;

[0812] 图16示出了不同SDS/细胞比率的EEH转化百分比;

[0813] 图17示出了三种不同SDS浓度下标准生物转化反应(如实施例7中所描述的)中的%EEH转化率;

[0814] 图18示出了三种不同pH值下标准生物转化反应(如实施例7中所描述的)中的%

EEH转化率;

[0815] 图19示出了SHC/HAC变体101A10、111C8和215G2中鉴定的突变在SHC晶体结构上的位置(以彩色显示):红色代表变体215G2;紫色(酒红色)代表变体101A10,绿色代表变体111C8。对于鉴定为负责活性增加的氨基酸,侧链在共结晶的底物类似物中以黄色突出显示。所鉴定的变体中不具有改善的活性的其它突变以蓝色标记。应该注意,蓝色突变在酶的两个结构域上约一半对一半(即50:50)散布,而鉴定的有利AacSHC突变大部分位于结构域2中。唯一例外是突变F601Y,其在活性位点附近;

[0816] 图20示出了以下突变(黑色和白色):对SHC/HAC活性不具有有利影响的突变以黑色显示,它们散布在SHC酶的两个结构域上。灰色示出了SHC变体中鉴定的显示出改善的SHC/HAC活性的突变(101A10、111C8和215G2),它们位于SHC酶的结构域2中,仅有一个例外。突出显示了引起变体改善活性的突变的侧链;

[0817] 图21示出了质粒pET-28a(+)的克隆和表达区;图21中的序列的SEQ ID No如下:

[0818] pET 28a(核苷酸序列):SEQ ID No.179;

[0819] pET 28a(氨基酸序列):SEQ ID No.180;

[0820] pET 28b(核苷酸序列):SEQ ID No.181;

[0821] pET 28b(氨基酸序列):SEQ ID No.182;

[0822] pET 28c(核苷酸序列):SEQ ID No.183;以及

[0823] pET 28c(氨基酸序列):SEQ ID No.184。

[0824] 图22示出了含有375g/l的细胞、188g/l的EEH、2.33%的SDS的1.5倍浓缩EEH生物转化反应物的体积生产力以及与之相比的普通生物转化的体积生产力,该常规生物转化以125g/l的EEH、250g/l的细胞、1.55%的SDS平行运行(实施例7);

[0825] 图23示出了常规的生物转化(125g/l的EEH、250g/l的细胞、1.55%的SDS),其如实施例7中所述运行,但是用0.5%或0.9%NaCl替代柠檬酸缓冲液pH 5.4,所有其它反应参数不变。在柠檬酸缓冲液中的生物转化平行运行作为对照;

[0826] 图24示出了(-)-降龙涎醚的固相萃取随甲苯洗涤的演变,以初始存在于200ml整个反应液中的(-)-降龙涎醚量的百分比表示(由于反应液/甲苯的体积比,第一次萃取中的百分比超过100%);以及

[0827] 图25示出了(-)-降龙涎醚的固相萃取随乙醇洗涤的演变,以最初存在的(-)-降龙涎醚量的百分比表示。在4次洗涤后(总共640ml EtOH,即3.2倍初始整个反应液体积或8倍固相体积),约99%的初始存在于反应液中的(-)-降龙涎醚得以回收。

[0828] 实施例

[0829] 为了避免疑问,所有提及WT SHC和SHC变体是提及WT AacSHC(SEQ ID No.1)及其变体(如表23和/或表24中所列出的)。

[0830] 实施例1

[0831] 生物催化剂制备

[0832] 方法1

[0833] SHC质粒制备

[0834] 将编码酸热脂环酸芽孢杆菌角鲨烯何帕烯环化酶(AacSHC)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(见图

5和21)。使用标准的热休克转化方案将该质粒转化进大肠杆菌BL21 (DE3) 菌株中。

[0835] 锥形瓶培养

[0836] 为了产生蛋白质,使用复合培养基(LB)或基本培养基。M9是基本培养基的一个例子,成功地使用了这种培养基。

[0837] 培养基制备

[0838] 对于350ml培养物默认选择的基本培养基制备如下:向35ml柠檬酸/磷酸盐母液(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)添加307ml H_2O ,根据需要用32%NaOH将pH调节至6.8。高压灭菌后添加0.850ml 50% MgSO_4 、0.035ml痕量元素溶液(组成见下一节中)溶液、0.035ml硫胺素溶液和7ml 20%葡萄糖。

[0839] SHC生物催化剂生产(生物催化剂生产)

[0840] 对于小规模生物催化剂生产(野生型SHC或SHC变体),用含有SHC生产质粒的大肠杆菌BL21 (DE3) 菌株的预培养物接种350ml培养物(补充有50 $\mu\text{g}/\text{ml}$ 卡那霉素的培养基)。在37℃下伴随连续搅拌(250转/分钟)将细胞培养至大约0.5的光密度($\text{OD}_{650\text{nm}}$)。

[0841] 然后通过将IPTG添加至300 μM 的浓度,然后伴随连续搅拌另外温育5-6小时来诱导蛋白质产生。最后通过离心收集所得的生物物质,用50mM Tris-HCl缓冲液pH 7.5洗涤。将细胞作为沉淀物在4℃或-20℃下保存直到进一步使用。通常从1升培养物中获得2.5至4克细胞(湿重),而与所用的培养基无关。

[0842] 制备发酵物并在750ml InforsHT反应器中运行。向该发酵容器添加168ml去离子水。该反应容器配备有所有需要的探针(pO_2 探针、pH探针、采样探针、防泡探针)、C+N进料瓶和氢氧化钠瓶并高压灭菌。将以下物质高压灭菌后添加至所述反应器

[0843] 20ml 10x磷酸盐/柠檬酸缓冲液

[0844] 14ml 50%葡萄糖

[0845] 0.53ml MgSO_4 溶液

[0846] 2ml $(\text{NH}_4)_2\text{SO}_4$ 溶液

[0847] 0.020ml痕量元素溶液

[0848] 0.400ml硫胺素溶液

[0849] 0.200ml卡那霉素母液

[0850] 运行参数设定如下:pH=6.95, $\text{pO}_2=40\%$, $T=30^\circ\text{C}$,以300转/分钟搅拌。关联控制参数:转速rpm设定值300,最小300,最大1000,流量l/min设定值0.1,最小0,最大0.6。消泡控制:1:9。

[0851] 用种菌培养物将发酵罐接种至 $\text{OD}_{650\text{nm}}=0.4-0.5$ 。将该种菌培养物在LB培养基(+卡那霉素)中于37℃、220rpm下培养8小时。首先以分批模式运行发酵11.5小时,然后开始用进料溶液(无菌葡萄糖溶液(143ml H_2O +35g葡萄糖),添加灭菌后的:17.5ml $(\text{NH}_4)_2\text{SO}_4$ 溶液、1.8ml MgSO_4 溶液、0.018ml痕量元素溶液、0.360ml硫胺素溶液、0.180ml卡那霉素母液)进行C+N进料。以大约4.2ml/h的恒定流量运行该进料。在外部进行葡萄糖和 NH_4^+ 测量以评价培养物中的C源和N源的可用度。通常葡萄糖水平保持十分低。

[0852] 将培养物培养总共大约25小时,这时它们通常达到 $\text{OD}_{650\text{nm}}=40-45$ 。然后通过向发酵罐中添加IPTG至大约1mM的浓度(使用输注注射器作为IPTG脉冲或在3-4小时的周期内添加),将温度设定至40℃, pO_2 设定至20%来开始SHC生产。在40℃下诱导SHC产生持续16小

时。在诱导结束时,通过离心收集细胞,用0.1M柠檬酸/柠檬酸钠缓冲液pH 5.4洗涤并以沉淀物在4℃或-20℃下保存直至进一步使用。

[0853] 结果1a

[0854] 通常,在所有其它条件不变的情况下,与复合培养基相比,在使用基本培养基时所产生的生物催化剂的比活性更高。在30或37℃下成功进行了诱导。注意到当在40-43℃下进行诱导时,获得更高比活力的生物催化剂。

[0855] 结果1b

[0856] 下面的表22为2个实施例示出了诱导开始时和诱导结束时细胞发酵液的培养物体积、光学密度和细胞量以及所收集的生物质的量(湿重)。

[0857] 表22

[0858]		诱导开始 时 的体积 (ml)	诱导开始 时 的 OD _{650nm}	细胞计 算量 (g)	诱导结束时 的体积 (ml)	诱导结束 时 的 OD _{650nm}	所收集 的细胞 (g)
	实施例 1	273	40	10.9	342	55	28
	实施例 2	272	44	12.0	341	57	23

[0859] 接种时的OD_{650nm}:0.45(实施例1)和0.40(实施例2)。起始体积:205ml。

[0860] 实施例2

[0861] SHC变体的制备及活性筛选

[0862] 方法2

[0863] 为了避免疑问,EE对应于(3E,7E);EZ混合物对应于(3Z,7E);ZE对应于(7Z,3E);ZZ对应于(7Z,3Z);EEH对应于(3E,7E)。

[0864] 使用野生型(WT)酸热脂环酸芽孢杆菌SHC(AacSHC)基因作为模板(GenBank M73834,Swissprot P33247)进行酶进化程序。制备了约10500种SHC变体的文库并筛选显示出增加的EEH环化能力的变体。在反应中运行筛选,该反应在含有4g/l EEH和0.050%SDS的柠檬酸缓冲液(pH 6.0)(0.150ml)中、于55℃下并且在连续搅拌下进行。

[0865] 在选取命中进行验证的情况下,在含有4g/l EEH、0.050%SDS、达到OD_{600nm}=10.0的已表达所述SHC变体的细胞的柠檬酸缓冲液pH 6.0中运行标准的测试。最终的体积为1ml,将反应物在55℃下温育并且在磁力搅拌器上剧烈搅拌。随时间推移对反应进行采样使得能考察活性概况(EEH转化为(-)-降龙涎醚)并且通过气相色谱分析测定(见下面的分析方法)。

[0866] 根据该验证轮次,获得了具有改善的EEH环化活性的3种变体(101A10、111C8和215G2)并在这3种变体上鉴定了总共8个突变。然后运行突变研究以鉴定这些突变中哪些突变对EEH环化成降龙涎醚是有利的。除了该AacSHC衍生物外,还构建了另一种AacSHC变体,其含有所有鉴定的有利突变(如下面表23中概述的SHC33)。筛选条件为:4g/l的EEH,细胞达到OD_{650nm}=10.0,SDS达到0.05%和0.1%(2种浓度),并且在55℃下在连续搅拌下运行反应。

[0867] 结果2a

[0868] 表23:所评价的AacSHC衍生酶中的突变

[0869]	SHC	T77A	I92V	F129L	M132R	A224V	I432T	Q579H	F601Y
	101A10							X	X
	111C8	X	X	X					
	215G2				X	X	X		
	SHC3								X
	SHC10			X					
	SHC26				X		X		
	SHC30			X					X
	SHC31			X	X		X		
	SHC32				X		X		X
	SHC33			X	X		X		X

[0870] 结果2b

[0871] 在三种选择的突变体(101A10、111C8和215G2)中,215G2显示出最佳的活性。

[0872] 实施例3

[0873] 用于SHC变体的最佳反应条件

[0874] 所考察的反应参数:温度、SDS浓度和pH

[0875] 方法3

[0876] 就温度、pH和SDS浓度分别优化表23中确定的SHC变体衍生物的反应条件。为此,用于产生各变体的质粒转化大肠杆菌细胞,然后如上所述将该大肠杆菌细胞在锥形瓶中培养并诱导SHC产生。以这种方式确保了所有的培养物均含有相同或十分相似的SHC量。通过离心收集细胞,用0.1M柠檬酸缓冲液(pH 6.0)洗涤并在-20℃下保存直至进一步使用。

[0877] 结果3

[0878] 该优化研究的结果汇总于下面的表中。还用野生型SHC进行了优化轮次。

[0879] 下面的表24示出了用于野生型以及每种变体的最佳反应条件,这些条件考虑用于表征每种SHC/HAC衍生酶。

[0880] 表24:SHC衍生酶的最佳反应条件

[0881]	SHC	温度(℃)	pH	[SDS] (重量/重量%)
	WT	55 (45-60)	6.0 (5.6-6.2)	0.030 (0.010-0.075)
	101A10	40 (36-50)	6.4 (5.4-7.0)	0.050 (0.010-0.10)
	111C8	40 (36-50)	6.0 (5.6-6.6)	0.070 (0.010-0.090)
	215G2	35 (32-50)	5.4 (5.0-6.2)	0.060 (0.010-0.10)
	SHC3	37 (34-50)	5.8 (5.4-6.4)	0.020 (0.010-0.060)
	SHC10	42 (34-55)	6.0 (5.4-6.4)	0.060 (0.030-0.10)
	SHC26	32 (30-50)	5.4 (5.4-6.2)	0.060 (0.020-0.10)
	SHC30	35 (34-50)	6.2 (5.4-7.0)	0.0050 (0.0025-0.070)
	SHC31	35 (30-50)	5.6 (5.4-6.4)	0.050 (0.010-0.10)
	SHC32	35 (34-50)	5.6 (5.4-6.4)	0.050 (0.010-0.10)
	SHC33	35 (32-50)	5.2 (4.8-6.4)	0.030 (0.0050-0.10)

[0882] 讨论3

[0883] 实施例3示出了SHC衍生物相较于WT SHC的反应条件中值得注意的差异。对于SHC变体观察到在最佳温度、pH和SDS浓度方面与野生型SHC有显著的偏差。仅少数突变对最佳生物转化反应条件具有显著影响。为了确定用于所选择的SHC变体的各反应条件,在底物载荷为4g/l的EEH并且已产生野生型或SHC衍生物的细胞的光密度 $OD_{650nm}=10.0$ 的条件下运行反应。

[0884] 温度

[0885] 表24中的数据展示了惊人的发现,WT SHC酶在55℃(范围为45-60℃)下具有最佳活性时,而许多SHC衍生物在35℃(34-50℃)下具有最佳活性。本公开的SHC衍生物在较低反应温度下应用于由E,E-高法呢醇制备(-)-降龙涎醚的方法中对于以工业规模生产降龙涎醚具有显著的成本优势。

[0886] 增溶剂

[0887] 从大量不可用于该生物转化反应的可能增溶剂中选择并鉴定了SDS(更多信息请参见实施例14)。

[0888] 在反应速度以及产率方面SDS优于例如Triton X-100(在以4g/l的EEH进行的测试中以及如实施例7中所提供的在使用125g/l的EEH时均是如此)。

[0889] 实施例4

[0890] 在标准条件下与WT SHC酶相比测试SHC变体活性

[0891] 方法4

[0892] 为了比较生物催化剂的相对活性,如下描述变体(如表24中所列出的)的产生。用于产生所述SHC变体之一的质粒转化大肠杆菌细胞,然后将该大肠杆菌细胞于37℃和280转/分钟下在LB培养基中培养,让其生长至 $OD_{650nm}=0.50$,然后通过添加IPTG诱导酶产生。在37℃、280转/分钟下持续诱导5.5小时。通过离心收集细胞,用0.1M柠檬酸缓冲液(pH 6.0)洗涤并在-20℃下保存直至进一步使用。当比较SHC变体活性时(见图6),将反应混合物的样品加载至SDS-PAGE凝胶上供分析反应物的SHC含量。该分析确认所有反应物均含有相同量的SHC酶。

[0893] 结果4a

[0894] 图6示出了野生型和SHC变体在标准条件(pH6.0, 55℃, 0.050% SDS, 细胞达到 $OD_{650nm}=10$)下的相对活性;还注意到,野生型SHC和至少根据本公开实施例的测试SHC变体是溶剂耐受的。这意味着可将所选择的水不混溶性溶剂(最高至100%)添加至该生物转化反应。

[0895] 结果4b

[0896] 使用215G2 SHC变体,当将NaCl添加至反应物(测试的浓度为5至100mM(唯一))时,未观察到对该变体活性的明显影响。此外,NaCl添加量高达100mM或高达154mM(0.9% NaCl)对变体215G2的SHC活性未显示出负面影响。这些发现表明,如果在NaCl的生理溶液(0.9%)或类似溶液中进行该生物转化反应并且将pH维持在恰当的值(如约5.4(5.2-5.6)),则可在不存在缓冲液但存在生理NaCl溶液或类似溶液的情况下进行该生物转化反应。

[0897] 讨论4

[0898] 图6图示了所选择的变体及野生型SHC酶在EEH转化为(-)-降龙涎醚方面的活性的

等级。

[0899] 实施例5

[0900] WT SHC和SHC衍生物活性概况

[0901] 方法5

[0902] 在Heidolph Synthesis 1设备上在5ml体积的0.1M柠檬酸缓冲液中于900转/分钟的连续摇动下运行活性测试。该反应中所使用的缓冲液的pH、反应运行的温度以及SDS(十二烷基硫酸钠)的浓度取决于使用的SHC变体(野生型或变体)。所测变体每一者的最佳条件汇总在上面的表24中。

[0903] 使用的高法呢醇原料为96%纯度,高法呢醇底物EEH:EZH比率为87:13。

[0904] 为了避免疑问,EE:EZ混合物是(3E,7E)和(3Z,7E)异构体的混合物。

[0905] 结果5

[0906] 所使用的高法呢醇:EEH:EZH 87:13,纯度(NMR):96%。

[0907] 在最佳条件下运行的标准测试的结果在图7B(SHC衍生物相对于WT SHC的活性概况)和图7A中示出,图7A示出了SHC衍生物相对于WT SHC的相对活性改善(4小时(初始速度)和22小时时的产率)。

[0908] 所使用的高法呢醇:EEH:EZH 92:08,纯度(NMR):100%

[0909] 在最佳条件下运行的标准测试的结果在图8B(AacSHC衍生物相对于WT AacSHC的活性概况)和图8A中示出,图8A示出了AacSHC衍生物相对于WT SHC的相对改善(4小时(初始速度)和22小时时的产率)。

[0910] 所使用的高法呢醇:EEH:EZH 66:33,纯度(NMR):76%

[0911] 在最佳条件下运行的标准测试的结果在图9B(如表24中列出的AacSHC衍生物相对于WT AacSHC的活性概况)和图9A中示出,图9A示出了AacSHC衍生物相对于WT SHC的相对改善(4小时(初始速度)和22小时时的产率)。

[0912] 讨论5

[0913] 主要结论是与所用的高法呢醇底物的质量无关,四种最好的SHC衍生酶以如下次序排序:215GSHC、SHC26、SHC32和SHC3。

[0914] 实施例6

[0915] 从用溶剂完全萃取的反应物测定质量平衡

[0916] 方法6

[0917] 所有条件不变,对每种变体运行2次反应。将高法呢醇用作底物。在温育4小时和22小时后,对于每种变体,用等体积的叔丁基甲基醚(MTBE/tBME)总共洗涤6次完全萃取反应产物和未反应的底物。通过GC分析测定每次洗液的高法呢醇和降龙涎醚含量。从校准曲线计算所形成的降龙涎醚和剩余的高法呢醇的总量,校准曲线使用由认证的降龙涎醚和高法呢醇制成的溶液制得。

[0918] 结果6

[0919] 图10中的结果显示,在使用所述底物的情况下,确认了3种最好的变体,它们较于野生型SHC酶显示出大约10倍(215G2)、7倍(SHC26)和6倍(SHC32)的提高。

[0920] 实施例7

[0921] 125g/l E,E-高法呢醇(EEH)下的生物转化性能

[0922] 方法7

[0923] 使用215G2SHC变体,解决了增加体积生产力的目标。运行系列实验设计(DOE)研究以优化包括参数pH、细胞浓度和SDS浓度在内的测试反应条件。反应条件为:125g/l的EEH(来自EE:EZ 86:14的高法呢醇)、250g/l的细胞、1.55%的SDS,反应在0.1M柠檬酸缓冲液(pH5.4)于35℃下运行。

[0924] 典型的反应(150g总量)设定如下:在0.75升Infors发酵罐中。将该反应容器加载适量的对应18.75g EEH的高法呢醇。从在0.1M柠檬酸缓冲液(pH 5.4)中制备的15.5% (w/w) 溶液添加2.33g SDS。通过将已产生215G2 SHC变体的大肠杆菌细胞悬浮于0.1M柠檬酸缓冲液(pH 5.4)中而由该细胞制备细胞悬液。在通过于10℃和17210g下离心10分钟来测定该悬液的细胞湿重后,将适当体积的细胞添加至反应容器以将37.5g的细胞引入该反应中。用所需量的反应缓冲液将反应物的量补足至150g。在900转/分钟的连续搅拌下于37℃运行该反应。使用水中的40%柠檬酸完成pH调节。随时间推移对反应进行采样(1ml),用5体积的MTBE/tBME(5ml)萃取。在通过离心(台式离心机,13000转/分钟,2分钟)使溶剂相澄清、10倍稀释于MTBE/tBME中后,通过GC分析测定反应物的高法呢醇和降龙涎醚含量。

[0925] 用已产生野生型SHC酶的大肠杆菌细胞进行相同的反应。在该情况中,在0.1M柠檬酸缓冲液(pH 6.0)中于55℃下运行该反应。下面的表24a的第2行提供了用于该实施例的反应条件概要。表24a第1行中给出的反应条件取自前面的实施例(如实施例3-5)。

[0926] 表24a:第2行示出了用于实施例7的反应条件

[0927]	SHC	温度 (℃)	pH	[SDS] (%, w/w)	[EEH] (g/l)	生物催化剂 (g/l) (细胞湿重)
	215G2	35℃	5.4 (5.0-6.2)	0.06	4	1.45 细胞
	215G2	35℃	5.4 (5.0-6.2)	1.55	125	250 细胞

[0928] 结果7

[0929] 图11示出了所观察到的由这2种酶催化的EEH向降龙涎醚的转化。在反应7小时时(初始反应速度的估计),使用变体215G2 SHC时的转化率比使用野生型SHC时实现的转化率高13倍。在反应48小时时,用该变体时的转化率是野生型酶的转化率的约8倍。

[0930] 总体评论7

[0931] 细胞浓度

[0932] 该实施例中所述的反应中的所有细胞浓度(g/l)均以细胞湿重表示。在将该细胞悬液样品在17210g和4℃下离心10分钟后测定细胞悬液的浓度,以细胞湿重(g/l)表示。

[0933] 细胞(g/l)与OD_{650nm}之间的相关性

[0934] 使用125g/l的EEH,用215G2 SHC或WT SHC进行生物转化的情况下,该反应中的250g/l细胞对应于该反应中约172的OD_{650nm}。当测试不同生物催化剂制备物时,观察到OD_{650nm}与生物催化剂量的比率的变化。当生物催化剂用于4g/lEEH下的标准测试中但施加细胞至OD_{650nm}=10.0时,据估计OD_{650nm}=10.0相当于1.45g/l的细胞。

[0935] 讨论7

[0936] 数据证明,与本领域的公开内容相比,已使用相对较高的EEH底物浓度(125g/l)发展了优化且有效的HAC生物转化工艺,其中在所述本领域中仅公开了约0.2g/l(参见JP2009060799)至约2.36g/l(10mM)的高法呢醇底物浓度(参见W02010/139719A2、US2012/0135477A1,以及Seitz等人(2012)(如上文所引用的))。

[0937] 实施例8

[0938] GC分析

[0939] 方法8

[0940] 用适当体积的叔丁基甲基醚(MBTE/tBME)萃取样品供定量它们的EEH和降龙涎醚含量。通过离心从水相分离溶剂级分,然后用气相色谱分析。将1 μ l的溶剂相注入(分流比为3)于30m x 0.32mm x 0.25 μ m Zebron ZB-5柱上。使用温度梯度:100 $^{\circ}$ C,15 $^{\circ}$ C/min至200 $^{\circ}$ C,120 $^{\circ}$ C/min至240 $^{\circ}$ C,240 $^{\circ}$ C 4分钟,以恒定流量(4ml/min H₂)展开柱,这导致降龙涎醚、EEH和EZH分离。进样口温度为200 $^{\circ}$ C,检测器温度:300 $^{\circ}$ C。

[0941] 用以下公式由对应于降龙涎醚和EEH的峰的面积计算EEH转化率:

[0942] 转化率(%) = 100x (面积_{降龙涎醚峰} / (面积_{降龙涎醚峰} + 面积_{E,E-高法呢醇峰}))

[0943] 通过GC-MS确认反应产物降龙涎醚(记录的值和强度:m/z 221(100%),m/z 97(40%),m/z 137(3.3%),m/z 43(2.6%),m/z 41(2.5%),m/z 55(2.4%),m/z 95(1.9%),m/z 67(1.8%),m/z 81(138%),m/z 222(1.7%))。

[0944] 讨论8

[0945] 通过溶剂萃取或汽提进行产品回收。所用的溶剂例如MTBE或己烷:异丙醇(3:2)。用等体积的溶剂反复萃取反应物并对溶剂级分进行GC分析直至不再检测到底物或产物。一般而言,5到6次洗涤就足够了。备选地,通过汽提进行反应产物的萃取。

[0946] 实施例9

[0947] 一锅法反应

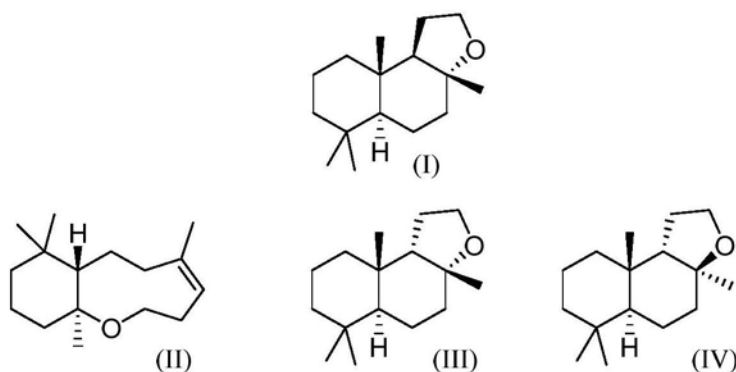
[0948] 方法9

[0949] 使用上述标准的培养和诱导方案,用转化了pET28a(+) 215G2 SHC质粒的大肠杆菌BL21(DE3)运行200ml的发酵,该质粒用于产生带有N末端His标签的215G2 SHC。在诱导阶段结束时,关闭通气,将温度设定为35 $^{\circ}$ C,用柠檬酸将pH调节至5.5,将搅拌器速度设定为500转/分钟。根据培养物生长期间所做的所有添加(进料和基本消耗)估计培养物的体积。根据该体积以及根据培养物的OD,向发酵罐添加适量的SDS。EEH添加至4g/l。随时间推移对反应进行采样,用700 μ l MTBE萃取样品(150-300 μ l)供GC分析。在培养发酵液中EEH直接被转化为降龙涎醚。该反应总共运行22.5天,在这期间反复添加EEH。

[0950] 结果9

[0951] 当达到完成时,10.6g的EEH已环化成降龙涎醚。通过汽提萃取反应产物(下面提供结构)并从反应混合物定量回收。

[0952]



[0953] 反应产物的注释

[0954] 当由SHC转化高法呢醇EE:EZ 87:13时,产生了如图12中所列出的反应产物降龙涎醚、(II)、(IV)和(III),并且反映该原料的EE:EZ比率。

[0955] 当EEH用作原料时,仅生成(-)-降龙涎醚(I)和产物(IV)。

[0956] 当EZH(3Z,7E)用作原料时,仅生成产物(II)和(III)。

[0957] 然而,当使用EEH和EZH的混合物时,生成降龙涎醚(I)以及产物(II)、(IV)和(III)。

[0958] 如果发生EE:EZ 66:34的100%转化,则这将提供66%:34%的((降龙涎醚+(IV)):(II)+(III))。

[0959] 当进行汽提时,其萃取全部4种产物-降龙涎醚以及产物(II)、(IV)和(III),并且结晶步骤产生纯度为99%(GC)的降龙涎醚,产率至少为70%。

[0960] 讨论9

[0961] 数据证明,在生物转化反应或“一锅法”反应系统中生产(-)-降龙涎醚是可能的,并且在汽提和结晶后实现降龙涎醚的选择性富集。

[0962] 如果高法呢醇原料是EE和EZ(如86:14)异构体的混合物,则从这些异构体每一者产生2种产物(总共4种),其中(-)-降龙涎醚目前是粗产物中的主要成分,并且是结晶材料中的优势成分(纯度为99.1%)。未检测到(+)-降龙涎醚。

[0963] 实施例10

[0964] EE:EZ高法呢醇混合物的转化

[0965] 为了避免疑问,

[0966] EE对应于(3E,7E);EZ混合物对应于(3Z,7E);ZE对应于(7Z,3E);ZZ对应于(7Z,3Z);EEH对应于(3E,7E);EZH对应于(3Z,7E)。

[0967] 方法10

[0968] 在以下反应条件下将EE:EZ混合物生物转化:146g/l总高法呢醇,使用250g/l的细胞和1.55%的SDS,采用以下高法呢醇底物(EE:EZ高法呢醇混合物):

[0969] EE:EZ 86:14(该实施例的最高EEH含量)

[0970] EE:EZ 69:31(该实施例的最低EEH含量)

[0971] EE:EZ 80:20

[0972] EEH:EZH 70:30

[0973] 7E,3E/7E,3Z高法呢醇混合物的生物转化

[0974] 使用以下反应条件进行生物转化:

[0975] 使用高法呢醇底物(其为7E,3E:7E,3Z 86:14的混合物)、250g/l的细胞(根据实施例1的方法制备)和1.55%的SDS,在容纳有146g/l总高法呢醇的InforsHT 750ml发酵罐中于0.1M柠檬酸/柠檬酸钠缓冲液(pH 5.4)中进行反应(150.1g总量)。伴随连续搅拌(800转/分钟)在35℃下运行该反应,用水中的10至40%柠檬酸进行pH控制。随时间推移对反应混合物进行采样,对样品进行溶剂萃取供GC分析。注意到,使用2种质量的高法呢醇(EE:EZ 86:14和EE:EZ 69:31)时高法呢醇转化同样地快。

[0976] 结果10

[0977] 当使用WT SHC和一种特异性的SHC衍生物(215G2 SHC)进行125g/l来自EEH:EZH 86:14材料的E,E-高法呢醇的生物转化时,观察到E,E--高法呢醇和E,Z-高法呢醇二者的转化。也就是说,来自酸热脂环酸芽孢杆菌的野生型SHC酶由EEH:EZH 86:14材料产生的反应产物与来自表23的SHC变体由EEH:EZH混合物产生的反应产物一样(即降龙涎醚、产物(II)、(IV)和(III))。图13和图14提供了降龙涎醚及产物(II)、(IV)和(III)的反应产物GC分析。

[0978] 讨论10

[0979] 根据本公开的高法呢醇向降龙涎醚的生物转化产生(-)-降龙涎醚作为优势化合物,但也可能会产生如上文所鉴定的非(-)-降龙涎醚的化合物(如化合物(II)、(IV)和(III)),这些化合物可能会或可能不会给(-)-降龙涎醚产品赋予令人愉悦的嗅觉香型。如上文所展示的,在选择性结晶条件下,降龙涎醚可与其它副产物((II)、(IV)和(III))分离。因此,如果产物对降龙涎醚终产物的感官特性产生负面效果,则从(-)-降龙涎醚终产物选择性分离产物(II)、(IV)和(III)可增加其作为香料或风味剂或化妆品或消费者护理产品的价值。使用由经训练的香水师进行的已确定的感官测试进行感官分析。如果(-)-降龙涎醚终产物自身主要负责所需的感官特性的话,则其纯度可以是该产物嗅觉品质的指标。

[0980] 实施例11

[0981] 从EE:EZ:ZE:ZZ-高法呢醇混合物进行的EEH转化

[0982] 方法11

[0983] 将EE:EZ:ZE:ZZ-高法呢醇40:26:20:14用作底物,用于215G2 SHC进行的EEH转化。为了进行比较,还使用了EE:EZ 2:1或93:07的其它高法呢醇。

[0984] 使用215G2 SHC变体但不在最佳条件下,考察了EE:EZ:ZE:ZZ-高法呢醇混合物的转化。反应条件为pH 5.8、在100mM柠檬酸盐缓冲液中、0.10%的SDS、40℃。观察到以下的EEH转化率,所有的反应均用恒定的2g/l的EEH运行(因此总高法呢醇浓度有变化)。

[0985] 结果11

[0986] 观察到以下高法呢醇异构体混合物转化率:

[0987] EE:EZ 2:1 50-55%

[0988] EE:EZ 93:7 78%

[0989] EE:EZ:ZE:ZZ 40:26:20:14 6%

[0990] 讨论11

[0991] 除了观察到的产率,数据还证明,215G2 SHC变体能够从复杂的EE:EZ:ZE:ZZ高法呢醇混合物将EEH转化为降龙涎醚。正如预期的,观察到较低的转化率,从而导致较低的降龙涎醚产率。该结果与以下观点一致:除了EEH之外的高法呢醇异构体可与EEH竞争对SHC/HAC衍生酶的接近,从而可能充当EEH向(-)-降龙涎醚转化的竞争性抑制剂和/或备选的底

物。

[0992] 实施例12

[0993] 使用Triton X-100和SDS时全细胞生物转化的对比数据

[0994] 方法12

[0995] 根据实施例4的方法4中的方案培养大肠杆菌宿主细胞。根据实施例4中的标准测试进行使用215G2SHC变体的生物转化反应。选择以下条件作为215G2 SHC变体的最适反应条件:4g/l的高法呢醇底物,细胞达到 $OD_{650nm}=10.0$,在0.1M柠檬酸/磷酸钠缓冲液pH5.4中,35℃,以及0.07%的SDS。

[0996] 结果12

[0997] 图15提供了在使用浓度范围为0.005%至0.48%的Triton X-100和浓度为0.07%的SDS时全细胞生物转化测定法中215G2SHC变体活性的比较。

[0998] 讨论12

[0999] 数据证明,使用Triton X-100时最大的活性仅为使用SDS时获得的活性的约20%。

[1000] 实施例13

[1001] SDS/细胞比率

[1002] 方法13

[1003] 使用4g/l的EEH底物、 $OD_{650nm}=5.0$ 的已产生215G2 SHC衍生酶的细胞,根据实施例4中的方法4建立生物转化反应。

[1004] 结果13

[1005] 结果在图16中给出,该图示出了对于不同SDS/细胞比率的EEH转化百分比。

[1006] 图16证明,使用不同的SDS/细胞比值,EEH向(-)-降龙涎醚转化的百分比取决于SDS/细胞比率。必须小心设定该比率以实现最大转化。

[1007] 例如,如果SDS浓度过低,则可能会观察到次优的高法呢醇转化。另一方面,例如,如果SDS浓度过高,则存在通过破坏完整的微生物细胞和/或SHC/HAC酶变性/失活而影响生物催化剂的风险。当使用125g/l的EEH和250g/l生物催化剂根据实施例7中的方法7进行生物转化反应时,最好的生物转化方案显示出16:1的[SDS]/[细胞]比率。

[1008] 讨论13

[1009] 该结果证明,在增溶剂(如SDS)浓度、生物质的量以及底物(EEH)浓度之间存在一定程度的相互依赖性。举个例子,当高法呢醇底物浓度增加时,发生有效生物转化反应需要足够量的生物催化剂和增溶剂(SDS)。

[1010] 实施例14

[1011] 供用于生物转化反应的可能增溶剂的测试

[1012] 方法14

[1013] 在使用与标准测试中相同的条件(4g/l的EEH,细胞达到 $OD_{650nm}=10.0$)的215G2 SHC EEH环化反应中测试各种增溶剂(如下面表26中所列出的)作为SDS的可能替代物。还使用标准测试对通过将处于其最佳浓度(0.060-0.070%)的SDS与所用的其它增溶剂(使用的浓度是根据用这些化合物进行的筛选分别测定为最佳的浓度(见下面的表26))组合来增强活性(累加效果)的可能性进行了测试。此外,还测试了已知可帮助增溶水不溶性化合物的某些“深共熔溶剂”(Deep eutectic solvent)和离子液体。

[1014] 结果14

[1015] 下面的表26汇总了目前在215G2 SHC EEH环化反应中测试的增溶剂(如:表面活性剂、去垢剂、溶解增强剂等)。与使用浓度范围为0.060-0.070%的SDS进行的对照反应相比,均没有改善的活性。在单独以定义为最佳的浓度使用这些化合物时观察到的活性仅为用SDS进行的对照反应中所获得的活性的约20%。注意到当完全不添加增溶剂时,实现了20%的EEH转化。当使用SDS并添加另外的增溶剂(以在测试中定义为最佳的浓度使用)时,没有观察到协同效应。相反,观察到EEH转化百分率降低。从该研究可以得出结论,在测试条件下所述化合物根本不改善EEH转化;相反获得对环化的不利影响,并且SDS是所研究的增溶剂中最有用的。此外,从使用已知可帮助增溶水不溶性化合物的“深共熔溶剂”和离子液体的测试未获得积极的结果。

[1016] 表26:提供了在生物转化反应中测试的增溶剂的列表

增溶剂	所测试的浓度范围
辛基磺基甜菜碱	0.19 - 3.0 %
CHAPS	0.020 - 0.18 %
聚氧乙烯蓖麻油	0.0063 - 0.5 %
二甲基亚砷	0.00032 - 0.2 %
十六烷基氯化吡啶鎓一水合物	0.013 - 0.5 %
十四烷基磺基甜菜碱	0.0009 - 0.03 %
Nonidet P40	0.005 - 0.16 %
[1017]	辛基-β-D-吡喃葡萄糖苷
十六烷基磺基甜菜碱	0.00000003-0.03 %
Pluronic P-105	0.000074 - 0.018 %
季铵盐(如四甲基溴化铵)	20-160 mM
牛磺酸脱氧胆酸钠水合物	0.05 - 0.4 %
Stepan®	0.01 - 0.6 %
Thesit®	0.05 - 0.8 %
Triton X-100	0.005 - 0.32 %
Tween 20	0.05 - 0.8 %
Tween 80	0.01 - 0.4 %

[1018] 讨论14

[1019] 本申请人从大量不可用于本公开的高法呢醇向(-)-降龙涎醚的生物转化反应中

的其它增溶剂中选择了SDS并将其鉴定为可用的增溶剂。

[1020] 实施例15

[1021] 对生物转化反应中的SDS浓度的敏感性

[1022] 方法15

[1023] 所应用的条件是在125g/l的EEH与250g/l生物催化剂以及1.55%的SDS下的标准生物转化(如实施例7中所述)的条件。还测试了两种其它SDS浓度(1.40%和1.70%的SDS)。所有的SDS浓度均为重量/重量%。

[1024] 还将125g/l的EEH与250g/l生物催化剂以及1.55%的SDS下的标准生物转化反应条件(如实施例7中所述)用于测试不同的pH值。

[1025] 在0.1M柠檬酸缓冲液中于pH 5.4下运行对照。用0.1M乙酸缓冲液运行在更低pH下运行的反应。

[1026] 结果15

[1027] 图17中的数据证明,该生物转化反应看起来对SDS浓度改变的敏感性低于当在标准测试(4g/l的EEH并且细胞达到 $OD_{650nm}=10.0$)中测试HAC活性时。

[1028] 图18中的数据证明,当应用该生物转化反应时,系统看起来对pH变化的敏感性低于当在标准测试(4g/l的EEH并且细胞达到 $OD_{650nm}=10.0$)中测试HAC活性时。

[1029] 讨论15

[1030] 数据证明了在125g/l的EEH和250g/l的细胞下的生物转化反应相对于所测试的SDS浓度范围和pH范围的稳健性。

[1031] 实施例16

[1032] 所鉴定的SHC/HAC突变在晶体结构上的位置

[1033] 在AacSHC/HAC变体中鉴定的突变的位置在图19中标记如下:红色代表变体215G2;紫色(酒红色)代表变体101A10,绿色代表变体111C8。对于鉴定为负责活性增加的氨基酸,侧链在共结晶的底物类似物中以黄色突出显示。所鉴定的变体中不具有改善活性的其它突变以蓝色标记。应该注意,蓝色突变在酶的两个结构域上约一半对一半(即50:50)散布,而鉴定的有利AacSHC突变大部分位于结构域2中。唯一例外是突变F601Y,其在活性位点附近。如果仅考虑SHC/HAC衍生酶215G2和111C8二者,则所有的突变体均位于结构域2中。图20提供了以黑色和白色显示的相同信息。

[1034] 结果16

[1035] 对应于215G2、111C8和101A10的所有有利的突变体(红色/绿色/紫色)大部分(除了一个突变体F601Y外)位于SHC晶体结构(如图19中所提供的)的结构域2(Wendt等人(1997) Science 277:1811)。根据野生型AacSHC(SEQ ID No.1)对SHC有利突变组合进行编号。

[1036] 讨论16

[1037] 该晶体结构可用于鉴定具有所需结构/活性关系(尤其是与高法呢醇向(-)-降龙涎醚转化相关的结构/活性关系)的SHC/HAC衍生物。有用的预选步骤可能要将选择限制在位于SHC/HAC晶体结构的结构域2中的氨基酸残基(见图19和20)。

[1038] 实例17

[1039] 高法呢醇的制备

[1040] 方法17

[1041] 一般分析条件

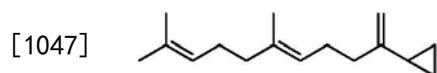
[1042] 非极性GC/MS:50℃/2min,20℃/min 200℃,35℃/min 270℃。带有HP 7890A系列GC系统的GC/MS Agilent 5975C MSD。非极性柱:购自SGE的BPX5,5%苯基95%二甲基聚硅氧烷,0.22mm x 0.25mm x 12m。载气:氮。进样口温度:230℃。分流比:1:50。流量:1.0ml/min。传输线:250℃。MS四极杆:106℃。MS源:230℃。

[1043] A) THF中的MNU的制备

[1044] 将脲(175g,2.9mol)和盐酸甲胺(198g,2.9mol)在水(400ml)中的溶液在搅拌下回流加热(105℃)3.5小时。在40℃下添加溶解于水(200ml)中的NaNO₂(101g,1.45mol)。15分钟后添加THF(1000ml),得到透明的两相混合物。在1.5小时内于0-5℃及搅拌下添加浓H₂SO₄(110g,1.1mol)。于0-5℃下另外0.5小时后,于25℃分离两个透明相。将有机相(A)(1065ml,理论上为1.35M)在0-5℃下保存数天或立刻加入环丙烷化反应器中。

[1045] 相分离后用THF(2x 1:1)萃取水相两次。这得到1100ml的相B和1075的相C。其中相A在后续的环丙烷化反应中有51%的末端链烯转化为环丙烷,相B产生<0.5%的环丙烷,相C不产生可检测的转化。我们得出结论,在第一次相分离后萃取了>99%MNU。因而水相通常在第一次相分离(与有机相A相分离)后在用浓KOH水溶液和乙酸处理后弃去。

[1046] B) 使用THF中的MNU制备E-△-法呢烯



[1048] 在0℃下将THF中的1.35M N-甲基-N-亚硝基脲(136ml,184mmol)滴加至0-5℃的快速搅拌的E-β-法呢烯(CAS 18794-84-8)(25g,122mmol)和KOH水溶液(50ml,40%)混合物。在添加4ml的MNU溶液后,添加预先溶解于0.5ml二氯甲烷中的Pd(acac)₂(7.4mg,0.024mmol,0.02%)。

[1049] 在0-5℃下用4小时添加剩余的MNU溶液。该阶段的GC显示了28%的未转化的E-β-法呢烯、65%的所需单环丙烷(如上所示)以及3%的双环丙烷化合物。25℃下16小时后在0-5℃下添加乙酸(100ml),然后添加叔丁基甲基醚(250ml)。在相分离后用2M HCl(250ml)洗涤有机相并用叔丁基甲基醚(250ml)萃取水相。用水(2x 100ml)、10%NaOH水溶液(2x 100ml)和水(2x 100ml)洗涤合并的有机层,用MgSO₄干燥,过滤并浓缩而得到26.9g微黄的液体,该液体含有9%的E-β-法呢烯、82%的所需单环丙烷化合物和6%的双环丙烷化副产物。

[1050] 可通过蒸馏纯化进一步分离所需化合物。添加1g K₂CO₃(1g)并在135-145℃下于40-60毫巴下在30cm钢卷曲柱上蒸馏得到147g单环丙烷化合物(68%corr)。汇集级分而得到92g 100%纯度的单环丙烷化合物。

[1051] E-△法呢烯的分析数据:

[1052] ¹H-NMR(CDC13,400MHz):5.1(2m,2H),4.6(2H),2.2(2H),2.1(4H),2.0(2H),1.7(s,3H),1.6(2s,6H),1.3(1H),0.6(2H),0.45(2H)ppm.¹³C-NMR(CDC13,400MHz):150.9(s),135.1(s),131.2(s),124.4(d),124.1(d),106.0(t),39.7(t),35.9(t),26.7(t),25.7(q),17.7(q),16.0(d),6.0(t)ppm.GC/MS:218(2%、M⁺),203(5%、[M-15]⁺),175(11%),147(31%),134(15%),133(20%),121(12%),107(55%),95(16%),93(30%),91(20%),82

(11%), 81 (33%), 79 (42%), 69 (100%), 67 (22%), 55 (20%), 53 (21%), 41 (75%)。IR (薄膜): 3081 (w), 2967 (m), 2915 (m), 2854 (m), 1642 (m), 1439 (m), 1377 (m), 1107 (w), 1047 (w), 1018 (m), 875 (s), 819 (m), 629 (w)。C₁₆H₂₆的分析计算值: C, 88.00; H, 12.00。实测值: C, 87.80; H, 12.01。

[1053] C (7E)-4,8,12-三甲基十三烷-3,7,11-三烯-1-醇 ((7E)-高-法呢醇)的制备

[1054] 在150℃下搅拌加热(E)-(6,10-二甲基十一烷-1,5,9-三烯-2-基)环丙烷(E-△法呢烯)(1g, 4.6mmol)、十二烷(0.2g, 1.15mmol, 内标)和L-(+)-酒石酸(1g, 6.9mmol)在压力管中的混合物。在18小时且完全转化(根据GC)后,将该混合物倾注于水(50ml)和甲苯(50ml)上。

[1055] 分离各相并用甲苯(50ml)萃取水相。用浓Na₂CO₃水溶液(50ml)和浓NaCl溶液(2x 50ml)洗涤合并的有机层,用MgSO₄干燥,过滤并减压蒸发而得到棕色树脂(1.35g),将其与30% KOH水溶液(4.3ml)混合并在25℃下搅拌2小时。根据内标,GC分析表明形成了96%的(7E)-4,8,12-三甲基十三烷-3,7,11-三烯-1-醇。E/Z比率为68:22。E异构体的分析数据与来自文献的分析数据一致,参见例如P.Kocienski, S.Wadman J.Org.Chem. 54, 1215 (1989)。

[1056] 结果17

[1057] 数据证明了高法呢醇的制备,该高法呢醇适于生物转化为(-)-降龙涎醚。

[1058] 讨论17

[1059] 这种制备高法呢醇的方法也在两件共同未决的专利申请PCT/EP2014/072882 (W02015/059290)和PCT/EP2014/072891 (W02015/059293)中详细的描述,将这两篇专利申请的全部内容以引用的方式并入本文。

[1060] 实例18

[1061] 一锅法反应

[1062] 在该实验中:(i)用产生215G2 SHC变体的大肠杆菌菌株发酵(如实施例1中所述),然后(ii)直接在发酵液中进行EEH转化。因为3个参数[细胞]、[EEH]和[SDS](g/l)相关联,因此需要根据在发酵结束时获得的细胞浓度(g/l)来调节可用体积发酵液中参数[EEH]和[SDS]。目标是在1.55%的SDS浓度下用250g/l的细胞转化125g/l的EEH。为了进行适当的生物转化,细胞必须在葡萄糖耗竭状况中处于休眠状态。关闭通气。

[1063] 方法18

[1064] 发酵:

[1065] 为了允许相当精确地测定发酵结束时反应器中的发酵液体积,对取出样品的体积以及添加进反应器的所有添加物(进料、碱、酸...)的体积进行记录。

[1066] 发酵液中细胞浓度的测定:

[1067] 在连续搅拌下取出发酵液样品(5-10ml)供细胞湿重(g/l)测定并放进离心管中。记录样品质量。将样品在17210g和4℃下离心10分钟(如12000转/分钟,SS-34转子,Sorvall RC3B离心机)。通过小心的吸移取出上清液并记录离心沉淀物的质量。细胞湿重浓度测定为 $\frac{g_{\text{细胞}}}{l_{\text{发酵液}}}$ 或 $\frac{g_{\text{细胞}}}{g_{\text{发酵液}}}$ 。

[1068] 根据所有的添加物和取出物确定发酵罐中的发酵液体积。在发酵罐处于天平上的情况中,通过称重确定发酵液的质量,如果不是这种情况则假定1ml=1g。

[1069] 高法呢醇和SDS的所需量的确定:

[1070] 根据所测定的细胞浓度和发酵液的体积,确定要添加至反应器的E,E-高法呢醇和SDS的量以便保持相同的如实施例9中所述生物转化中列出的3者之间的比率:125g/l的EEH、250g/l的细胞、1.55%的SDS。

[1071] 建立生物转化:

[1072] 1.将温度设为35℃。关闭通气。

[1073] 2.向发酵液添加计算量的高法呢醇。

[1074] 3.由含水的15.5%SDS母液小心添加所需量的SDS。

[1075] 4.将反应物以800转/分钟充分混合大约15分钟。

[1076] 5.记录该反应物的pH(内部pH电极)。

[1077] 6.吸取样品(大约1ml)至15ml的Falcon管。添加大约5ml去离子水,在彻底混合后记录经外部校准的电极处的pH。

[1078] 7.使用85% H_3PO_4 将反应器中的pH逐步设定至5.4(在经外部校准的电极处测量的值),同时如上所述定期控制外部电极处的pH(6.)。

[1079] 8.在生物转化期间用例如10-25% H_3PO_4 和32%NaOH调节pH。

[1080] 9.反应物采样:将大约1ml反应混合物放入15ml falcon管中。添加大约5ml MTBE。在剧烈振荡下萃取样品。

[1081] 10.将等分试样在台式离心机中以最大速度离心1分钟(埃彭道夫管)。将100 μ l溶剂相添加至装有900 μ l MTBE的GC管中。在生物转化的第一天中每1-1.5小时取样。在接下来的日子里仅每天取3份样品。

[1082] 11.如实施例8中所述分析1 μ l溶剂相的降龙涎醚和EEH含量。

[1083] 12.EEH转化率(%)计算为 $100 \times (\text{降龙涎醚}_{\text{面积}} / (\text{降龙涎醚}_{\text{面积}} + \text{EEH}_{\text{面积}}))$ 。

[1084] 结果18

[1085] 结果证明,以1.9升的规模在KLF2000反应器(生物工程公司(Bioengineering))中进行了一锅法发酵+EEH转化。251g/l的细胞使得在47小时内238g EEH(251g/l的细胞)转化达到 $\geq 93\%$ 。当在开始后93小时测量时,转化率为99%。

[1086] 在Infors HT 0.75l反应器中进行了类似的一锅法实验。在根据标准方案(实施例1)的发酵后,添加从用相同方案平行运行的其它发酵收集而来的反应器细胞。所得的发酵液量为479g。细胞浓度测定为313.7g/l,其为常规生物转化中的细胞浓度(250g/l的细胞)的1.25倍。相应地将EEH和SDS添加至该反应器。在少于90小时内,75.1g EEH(在该实施例中等价于157g/l的EEH)转化达到98%。该结果证明,有可能在 ≥ 125 g/l的EEH下运行一锅法发酵+EEH转化,只要发酵轮次以足够高的细胞密度提供已产生215G2 SHC变体的细胞。

[1087] 讨论18

[1088] 有利的是,获得了99%的底物转化率,当使用昂贵的原料(如EEH)时这在商业上是非常有用的。

[1089] 实施例19

[1090] 增加体积生产力.

[1091] 方法19

[1092] 为了进一步增加体积生产力,运行了1.5倍浓缩的生物转化,其含有375g/l的细胞、188g/l的EEH、2.33%的SDS。以125g/l的EEH、250g/l的细胞、1.55%的SDS平行运行常规

的生物转化(实施例7)。在Infors HT 0.7501反应器中运行这两个反应,所有其它参数不变。

[1093] 结果19

[1094] 图22中的结果证明,在开始后75小时所述1.5倍生物转化中的转化率百分比为88%,相比之下在常规生物转化中为95%。在开始后96小时所述1.5倍生物转化中的转化率百分比为93%的EEH转化,相比之下在常规生物转化中为97%。所述1.5倍生物转化中的转化率百分比为常规生物转化中获得的转化率百分比的96%。注意到,在1.5倍生物转化中搅拌随时间推移变得更困难,因为油状高法呢醇消失,而代之以固体反应产物。这可能解释1.5倍生物转化中转化水平稍微较低的原因。使用配备有更好混合装置的反应器可能会改善1.5倍生物转化中的EEH转化率。结果表明,倘若实现充分混合,则以188g/l的EEH或更高浓度运行生物转化是可能的,搅拌效率看起来是该系统的唯一限制。

[1095] (-)-降龙涎醚生产力

[1096] “(-)-降龙涎醚生产力”是指每小时生物转化时间(即添加底物后的时间)每升生物转化物中的可回收(-)-降龙涎醚的量(克数)。就这一点而言并且参照图22,(-)-降龙涎醚生产力计算如下:

[1097] 125g/l的EEH生物转化(250g/l的细胞)

[1098] 1.25小时时的生产力:每小时每升10.3克

[1099] 8.25小时时的生产力:每小时每升6.3克

[1100] 21.25小时时的生产力:每小时每升4.1克

[1101] 187.5g/l的EEH生物转化(375g/l的细胞)

[1102] 1.25小时时的生产力:每小时每升12.2克

[1103] 8.25小时时的生产力:每小时每升8.2克

[1104] 21.25小时时的生产力:每小时每升5.5克

[1105] 可以认为,在开始后约6-8小时计算的生产力代表反应的初始速度,该初始速度可最好地描述系统的最大转化速率。

[1106] 采用125g/l的EEH和250g/l的细胞的典型生物转化显示,在约6-8小时后降龙涎醚生产力介于6.3至8.5g每小时每升之间(代表反应的初始速度)。

[1107] 实施例20

[1108] 用NaCl溶液替代反应缓冲液

[1109] 方法20

[1110] 如实施例7中所述运行常规的生物转化(125g/l的EEH、250g/l的细胞、1.55%的SDS),但是用0.5%或0.9%NaCl替代柠檬酸缓冲液pH 5.4,所有其它反应参数不变。平行运行在柠檬酸缓冲液中的生物转化作为对照。

[1111] 结果20

[1112] 图23中的结果证明,在缓冲液和0.9%NaCl中运行的反应中EEH转化率相同。当反应在仅0.5%的NaCl中运行时转化率稍低。结果证明只要保证精确的pH调节和足够的离子强度,则在不存在缓冲液的情况下运行生物转化是可能的。

[1113] 实施例21

[1114] 反应液固相的萃取

[1115] 考虑到(-)-降龙涎醚在水中是不可溶的并且在低于大约75℃的温度下不是液体,尽可能利用这些特性使用水混溶性溶剂(如乙醇)和水不混溶性溶剂(如甲苯)来从该生物转化的固相萃取产物。

[1116] 方法21

[1117] 将200ml反应液离心以从液(水)相分离固体(Sorvall GS3,5000转/分钟,10分钟,10℃)。这使大约80ml固体沉淀物与大约120ml液相分离。在MTBE萃取后对水相进行的分析(气相色谱,实施例8)显示,其含有不超过最初存在于200ml反应液体中的(-)-降龙涎醚的大约0.3%。将甲苯和99%乙醇用于从固相萃取降龙涎醚。

[1118] 结果21

[1119] 甲苯萃取:

[1120] 将80ml固相用以下条件萃取6次:使用45ml甲苯(大约固相体积的一半),剧烈振荡30秒,离心(Sorvall GS3,5000转/分钟,10分钟,10℃)。用GC分析溶剂相的(-)-降龙涎醚含量。通过6次萃取超过99.5%的最初存在于反应液中的(-)-降龙涎醚被萃取,6次萃取代表总甲苯体积为最初整个反应液体积(200ml)的1.35倍或者是固相体积的3.4倍。图24中的曲线图示出了萃取随甲苯洗涤的演变,以初始存在于200ml整个反应液中的(-)-降龙涎醚量的百分比表示(由于反应液/甲苯体积比,第一次萃取中的百分比超过100%)。

[1121] 乙醇萃取:

[1122] 用大约160ml(2倍体积)99%的乙醇萃取80ml固相(Infors Multifors HT,35℃,1000转/分钟,30分钟),然后离心。在萃取过程中降龙涎醚不结晶。图25中的曲线图显示,在4次洗涤后(总共640ml EtOH,即3.2倍初始的整个反应液体积或8倍固相体积),约99%的初始存在于反应液中的降龙涎醚得以回收。在第一萃取步骤中需要足量的乙醇来防止降龙涎醚结晶(在乙醇中的溶解性)。当在第一萃取步骤中使用仅1体积或1/2体积的固相时,获得粘性糊状物,其难以处理并且在离心过程中(-)-降龙涎醚以针状结晶物结晶于离心沉淀上。温度看起来不是造成该结晶的因素(萃取和离心在室温以及大约35℃-40℃下测试)。

[1123] EtOH相中的(-)-降龙涎醚浓度以及液相的EtOH/水比率(固相的残留水分)看起来是结晶形成的原因。然而注意到,将乙醇体积减少至1固相体积是可能的。

[1124] 评论21

[1125] 因为在室温下(-)-降龙涎醚不在液相中,其与生物物质分离并且可用有机溶剂(如水混溶性溶剂(如乙醇)或非水混溶性溶剂(如甲苯))萃取。将(-)-降龙涎醚分离进反应混合物固相中的离心步骤是有利的,因为其减少了萃取(-)-降龙涎醚所需的溶剂量。

[1126] 实施例22

[1127] 感官分析

[1128] 目的:为了进行对“粗”萃取物和“结晶”萃取物中所形成的(-)-降龙涎醚和副产物(化合物II、III和IV)进行感官分析。

[1129] 结果22(a)

[1130] EEH转化得到(-)-降龙涎醚(化合物I)和(-)-降龙涎醚异构体(化合物IV)。

[1131] 结果22(b)

[1132] EZH生物转化得到大环醚(化合物II)和9b-表-降龙涎醚(化合物III)。

[1133] 结果22(c)

[1134] 粗(-)-降龙涎醚的组成包含化合物I、II、III和IV,各化合物分别以87.1%、2.8%、2.5%和7.6%的量存在。

[1135] 结果22(d)

[1136] 选择性结晶材料(实验室规模)的组成具有相同组分,分别以99.1%、0.1%、0.1%和0.7%的量存在。

[1137] 感官分析结果如下:

[1138] (-)-降龙涎醚:OTH 0.2ng/l (OTH是嗅觉阈值)。

[1139] 来自EEH的化合物IV:弱,IsoE、木香、GC-TH5-10ng。

[1140] 来自EZH的化合物II:“无臭”(GC-TH>500ng) (GC-TH是检测阈值)。

[1141] 来自EZH的化合物III:GC-TH比降龙涎醚高约10倍(大约2ng)。

[1142] 结论

[1143] “粗制”萃取物中3种副产物(化合物II、III和IV)每一者的总百分比约为3%。

[1144] “结晶”萃取物中3种副产物(化合物II、III和IV)每一者的总百分比约为1%(实验室规模)。

[1145] 对3种副产物(化合物II、III和IV)的感官分析表明气味弱于来自(-)-降龙涎醚的气味。

[1146] 实际上,9b-表-降龙涎醚(化合物III)气味比(-)-降龙涎醚弱约10倍,从而表明其基本上是无臭的。

[1147] 感官分析证明,从(-)-降龙涎醚移除一种或多种副产物化合物可改善剩余化合物((-)-降龙涎醚)的气味,即使所移除的化合物实际上本身是无臭化合物。

[1148] 也就是说,在不存在化合物II、III和IV的情况下观察到降龙涎醚气味增强。

[1149] 实施例23

[1150] 通过汽提回收降龙涎醚

[1151] 方法23

[1152] 所得的粗制(汽提)和结晶(-)-降龙涎醚的纯度

[1153] 汽提EE:EZ 86:14生物转化反应混合物并如下对反应产物进行结晶:收集蒸气溜出液为两相混合物。保留有机相,弃去水相。通过GC分析有机相的组成,结果在下面的表25中示出(见“粗制”)。然后浓缩有机相至干燥。然后将乙醇加至该粗制干燥产品并使该混合物升温直至产品溶解。室温下缓慢添加水,(-)-降龙涎醚在偶尔搅拌和冰浴中冷却下结晶。

[1154] 结果23

[1155] 下面的表25还示出了在汽提/蒸馏步骤(“粗制”)后获得的产品及结晶产品((-)-降龙涎醚)的GC分析结果。表25中提及“EZH”和“EEH”分别指(3Z,7E)高法呢醇和7E,3E高法呢醇。

[1156] 表25表明,使用WTSHC或SHC衍生物,特定的原料(EEH:EZH 86:14)产生了所需的终产物(-)-降龙涎醚和十分特别的副产物混合物(II、IV和III)。选择性结晶的数据显示(-)-降龙涎醚(I)的强烈富集,在结晶样品中几乎不存在副产物(II)、(IV)或(III)。因此,这种EE:EZ混合物提供了在嗅觉上纯的(-)-降龙涎醚产品,该产品以相对直接且成本低的方式选择性结晶。

[1157] 表25:示出了结晶产品的GC分析结果。

[1158]	峰面积 (GC)						降龙 涎醚 (%)
	(II)	(IV)	(III)	降龙涎醚	EZH	EEH	
粗制	215073	190376	588769	6751605	13429	14184	86.9
结晶	10088	8951	64625	9032941	0	0	99.1

[1159] 讨论23

[1160] 汽提/过滤是分离降龙涎醚的环境友好的方法,因为其提供了对降龙涎醚的便捷无溶剂分离,没有相关的生物催化剂失活。

[1161] 总结23

[1162] 采用生物转化反应制备的(-)-降龙涎醚可利用本领域技术人员已知的方法,用溶剂从整个反应混合物萃取(如使用非水混溶性溶剂或通过汽提/蒸馏或通过过滤)或从固相萃取(如使用水混溶性溶剂)。

[1163] 本发明涉及以下实施方案:

[1164] 1.一种制备(-)-降龙涎醚或包含(-)-降龙涎醚的混合物的方法,其中(3E,7E)-高法呢醇(EEH)或包含EEH的立体异构体混合物经酶促转化为(-)-降龙涎醚或包含(-)-降龙涎醚的混合物,其中所述酶促转化使用SHC/HAC酶在适于产生(-)-降龙涎醚的反应条件下进行,并且其中包含EEH的立体异构体混合物基本上由选自由以下组成的组中的一者或多者的高法呢醇异构体组成:[(3E,7E)和[(3Z,7E)]和/或[(3E,7E)和(3E,7Z)]和/或[(3Z,7E)、(3E,7E)和(3E,7Z)],它们也分别命名为[EE:EZ]、[EE:ZE]和[EE:EZ:ZE]。

[1165] 2.一种制备(-)-降龙涎醚或包含(-)-降龙涎醚的混合物的方法,其中(3E,7E)-高法呢醇(EEH)或包含EEH的立体异构体混合物经酶促转化而得到(-)-降龙涎醚或包含(-)-降龙涎醚的混合物,其中所述酶促转化使用SHC/HAC酶在适于产生(-)-降龙涎醚的反应条件下进行,并且其中如果所述反应在存在增溶剂的情况下进行,则不将Triton X-100或牛黄脱氧胆酸盐与野生型SHC/HAC酶结合使用。

[1166] 3.一种制备(-)-降龙涎醚或包含(-)-降龙涎醚的混合物的方法,其中(3E,7E)-高法呢醇(EEH)或包含EEH的立体异构体混合物经酶促转化为(-)-降龙涎醚或包含(-)-降龙涎醚的混合物,其中所述酶促转化使用SHC/HAC酶在适于产生(-)-降龙涎醚的反应条件下进行,并且其中包含EEH的立体异构体混合物基本上由选自由以下组成的组中的一者或多者的高法呢醇异构体组成:[(3E,7E)和[(3Z,7E)]和/或[(3E,7E)和(3E,7Z)]和/或[(3Z,7E)、(3E,7E)和(3E,7Z)],它们也分别命名为[EE:EZ]、[EE:ZE]和[EE:EZ:ZE],并且其中所述反应在包含水相、固相和油相的三相系统中发生。

[1167] 4.根据实施方案1或实施方案2或实施方案3所述的方法,其中所述方法使用选自SEQ ID No.1、SEQ ID No.2、SEQ ID No.3、SEQ ID No.4的SHC/HAC酶多肽序列或选自表1、表5、表2、表6、表3、表7、表4、表8或表13或表14,或选自SEQ ID No.5、7、9、11、13、15、17、19、21、23、25、27、29、31、33、35、37、39、171、173、175、177和/或178的SHC/HAC衍生物或与SEQ ID No.1、SEQ ID No.2、SEQ ID No.3或SEQ ID No.4具有至少30%同一性、至少40%同一性、至少50%同一性、或至少60%同一性、或至少70%同一性、或至少80%同一性、或至少

90%同一性、或至少95%同一性、或至少96%同一性、或至少97%同一性、或至少98%同一性、或至少99%同一性的序列进行。

[1168] 5. 根据实施方案1-4中任一项所述的方法,其中所述方法使用产生所述SHC/HAC酶的重组宿主细胞。

[1169] 6. 根据实施方案4或实施方案5所述的方法,其中所述编码SHC/HAC酶的核苷酸序列选自SEQ ID No. 165、166、167、168、169或SEQ ID No. 6、8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、38、40和/或170、172、174和/或176。

[1170] 7. 根据实施方案1-6中任一项所述的方法,其中高法呢醇向(-)-降龙涎醚的转化在约4-8的pH下于30℃至60℃的温度下发生。

[1171] 8. 根据实施方案1-7中任一项所述的方法,其中高法呢醇向(-)-降龙涎醚的转化使用如表24或表24a中所列的用于所述野生型SHC/HAC或SHC/HAC衍生酶的反应条件中的一者或多者,优选在5.0至6.2的pH范围下、优选在35℃的温度下发生。

[1172] 9. 根据实施方案3-8中任一项所述的方法,其中当生物催化剂与EEH的比率约为2:1时,SDS/细胞比率在10:1至20:1的范围内,优选为16:1。

[1173] 10. 根据实施方案3-9中任一项所述的方法,其中生物催化剂与高法呢醇的重量比在约0.5-2:1的范围内,优选约1:1或0.5:1。

[1174] 11. 根据实施方案3-10中任一项所述的方法,其中所述细胞生长和生物转化反应步骤在同一反应容器中进行。

[1175] 12. 根据实施方案2所述的方法,其中所述高法呢醇底物包含一种或多种高法呢醇立体异构体。

[1176] 13. 根据实施方案12所述的方法,其中所述高法呢醇底物包含两种高法呢醇立体异构体或基本上由两种高法呢醇立体异构体组成。

[1177] 14. 根据实施方案13所述的方法,其中所述高法呢醇底物包含EE:EZ立体异构体或基本上由EE:EZ立体异构体组成。

[1178] 15. 根据实施方案14中任一项所述的方法,其中所述高法呢醇包含EE:EZ立体异构体混合物或基本上由EE:EZ立体异构体混合物组成,所述EE:EZ立体异构体混合物的重量比选自:100:00;99:01;98:02;97:03;96:04;95:05;94:06;93:07;92:08;91:09;90:10;89:11;88:12;87:13;86:14;85:15;84:16;83:17;82:18;81:19;80:20;79:21;78:22;77:23;76:24;75:25;74:26;73:27;72:28;71:29;70:30;69:31;68:32;67:33;66:34;65:35;64:36;63:37;62:38;61:39;和/或60:40。

[1179] 16. 根据实施方案15所述的方法,其中所述高法呢醇包含EE:EZ立体异构体混合物或基本上由EE:EZ立体异构体混合物组成,所述EE:EZ立体异构体混合物的重量比选自:EE:EZ 92:08;EE:EZ 90:10;EE:EZ 80:20;EE:EZ 86:14;EE:EZ 70:30;EE:EZ 69:31;和/或EE:EZ 66:34。

[1180] 17. 根据实施方案15或实施方案16所述的方法,其中所述高法呢醇包含EE:EZ立体异构体混合物或基本上由EE:EZ立体异构体混合物组成,所述EE:EZ立体异构体混合物的重量比为80:20。

[1181] 18. 根据实施方案1-17中任一项所述的方法,其中(-)-降龙涎醚以与副产物(II)、(IV)和/或(III)中的至少一种或多种的混合物产生。

[1182] 19. 根据实施方案1-18中任一项所述的方法, 其中使用有机溶剂或汽提/蒸馏步骤或过滤从所述生物转化反应混合物分离(-)-降龙涎醚。

[1183] 20. 根据实施方案1-19中任一项所述的方法, 其中使用有机溶剂或汽提/蒸馏步骤从所述生物转化反应混合物的固相分离(-)-降龙涎醚。

[1184] 21. 根据实施方案19或实施方案20所述的方法, 其中使用有机溶剂从所述反应混合物分离(-)-降龙涎醚。

[1185] 22. 根据实施方案21所述的方法, 其中使用乙醇或甲苯从所述反应混合物分离(-)-降龙涎醚。

[1186] 23. 根据实施方案19-22中任一项所述的方法, 其中使用有机溶剂使所述(-)-降龙涎醚选择性结晶。

[1187] 24. 根据实施方案23所述的方法, 其中所述(-)-降龙涎醚基本上不含副产物(II)、(IV)和/或(III)。

[1188] 25. 根据实施方案1-24中任一项所述的方法, 其中产生浓度范围为约125-200g/l的(-)-降龙涎醚。

[1189] 26. 通过根据实施方案1-25中任一项所述的方法能够获得的(-)-降龙涎醚, 其中所述(-)-降龙涎醚具有约0.1至约0.5ng/l的嗅觉阈值。

[1190] 27. 根据实施方案26所述的(-)-降龙涎醚, 其为固体形式, 优选为无定形式或晶体形式。

[1191] 28. 一种制备含有(-)-降龙涎醚的产品的的方法, 包括将根据实施方案26或实施方案27中任一项所述的(-)-降龙涎醚掺入所述产品中。

[1192] 29. 根据实施方案28所述的方法, 其中所述产品是香料产品、化妆品、清洁产品、洗涤剂产品或皂产品。

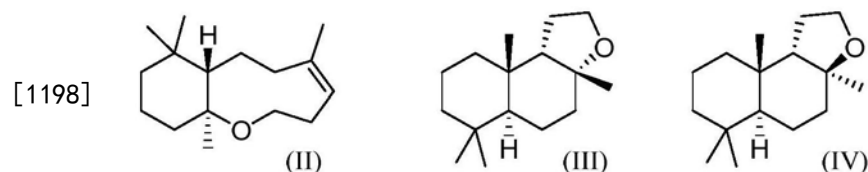
[1193] 30. 一种香料或化妆品或消费者护理产品, 其中包含根据实施方案26或实施方案27中任一项所述的(-)-降龙涎醚。

[1194] 31. 一种香料或化妆品或消费者护理组合物, 其中包含根据实施方案26或实施方案27所述的(-)-降龙涎醚和一种或多种另外的组分。

[1195] 32. 根据实施方案26或实施方案27所述的(-)-降龙涎醚的用途, 其用作香料或化妆品或消费品诸如织物护理、化妆用具、美容护理和/或清洁产品的一部分。

[1196] 33. 一种用于在香料组合物中或为香料组合物增加、增强或赋予芳香的方法, 包括使所述香料组合物与芳香增加或增强产品混合的步骤, 所述芳香增加或增强产品根据包括以下步骤的方法制备:

[1197] (a) 制备反应混合物, 所述反应混合物包含与副产物化合物(II)、(III)或(IV)中的一种或多种混合的(-)-降龙涎醚。



[1199] (b) 萃取与副产物化合物(II)、(III)或(IV)中的一者或多者混合的(-)-降龙涎醚; 以及

[1200] (c) 从所述萃取混合物中选择性结晶(-)-降龙涎醚;

[1201] 其中通过使用SHC/HAC酶在适于产生(-)-降龙涎醚的反应条件下酶促转化(3E, 7E)-高法呢醇(EEH)或包含EEH的立体异构体混合物来制备所述(-)-降龙涎醚,并且其中所述包含EEH的立体异构体混合物基本上由选自以下组成的组中的一者或多者的高法呢醇异构体组成:[(3E, 7E)和[(3Z, 7E)]和/或[(3E, 7E)和(3E, 7Z)]和/或[(3Z, 7E)、(3E, 7E)和(3E, 7Z)],它们也分别命名为[EE:EZ]、[EE:ZE]和[EE:EZ:ZE]。

[1202] 34. 根据实施方案33所述的方法,其中所述反应在包含水相、固相和油相的三相系统中发生。

[1203] 35. 根据实施方案33或实施方案34所述的方法,其中所述方法使用选自SEQ ID No.1、SEQ ID No.2、SEQ ID No.3、SEQ ID No.4的SHC/HAC酶多肽序列或选自表1、表5、表2、表6、表3、表7、表4、表8或表13、表14,或选自SEQ ID No.5、7、9、11、13、15、17、19、21、23、25、27、29、31、33、35、37、39、171、173、175、177和/或178的SHC/HAC衍生物或与SEQ ID No.1、SEQ ID No.2、SEQ ID No.3或SEQ ID No.4具有至少30%同一性、至少40%同一性、至少50%同一性、或至少60%同一性、或至少70%同一性、或至少80%同一性、或至少90%同一性、或至少95%同一性、或至少96%同一性、或至少97%同一性、或至少98%同一性、或至少99%同一性的序列进行。

[1204] 36. 根据实施方案33-35中任一项所述的方法,其中所述方法使用产生所述SHC/HAC酶的重组宿主细胞。

[1205] 37. 根据实施方案35或实施方案36所述的方法,其中所述编码SHC/HAC酶的核苷酸序列选自SEQ ID No.165、166、167、168、169或SEQ ID No.6、8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、38、40和/或170、172、174和/或176。

[1206] 38. 根据实施方案33-37中任一项所述的方法,其中高法呢醇向(-)-降龙涎醚的转化在约4-8的pH下于30℃至60℃的温度下发生。

[1207] 39. 根据实施方案33-38中任一项所述的方法,其中高法呢醇向(-)-降龙涎醚的转化使用如表24或表24a中所列的用于所述野生型SHC/HAC或SHC/HAC衍生酶的反应条件中的一者或多者,优选在5.0至6.2的pH范围下、优选在35℃的温度下发生。

[1208] 40. 根据实施方案34-39中任一项所述的方法,其中当生物催化剂与EEH的比率约为2:1时,SDS/细胞比率在10:1至20:1的范围内,优选为16:1。

[1209] 41. 根据实施方案34-40中任一项所述的方法,其中生物催化剂与高法呢醇的重量比在约0.5-2:1的范围内,优选约1:1或0.5:1。

[1210] 42. 根据实施方案34-41中任一项所述的方法,其中所述细胞生长步骤和生物转化反应步骤在同一反应容器中进行。

[1211] 43. 根据实施方案33-42中任一项所述的方法,其中所述高法呢醇包含EE:EZ立体异构体混合物或基本上由EE:EZ立体异构体混合物组成,所述EE:EZ立体异构体混合物的重量比选自:100:00;99:01;98:02;97:03;96:04;95:05;94:06;93:07;92:08;91:09;90:10;89:11;88:12;87:13;86:14;85:15;84:16;83:17;82:18;81:19;80:20;79:21;78:22;77:23;76:24;75:25;74:26;73:27;72:28;71:29 70:30;69:31;68:32;67:33;66:34;65:35;64:36;63:37;62:38;61:39;和/或60:40。

[1212] 44. 根据实施方案43所述的方法,其中所述高法呢醇包含EE:EZ立体异构体混合物

或基本上由EE:EZ立体异构体混合物组成,所述EE:EZ立体异构体混合物的重量比选自:EE:EZ 92:08;EE:EZ 90:10;EE:EZ 80:20;EE:EZ 86:14;EE:EZ 70:30;EE:EZ 69:31;和/或EE:EZ 66:34。

[1213] 45.根据实施方案43或实施方案44所述的方法,其中所述高法呢醇包含EE:EZ立体异构体混合物或基本上由EE:EZ立体异构体混合物组成,所述EE:EZ立体异构体混合物的重量比为80:20。

[1214] 46.根据实施方案33-45中任一项所述的方法,其中(-)-降龙涎醚以与所述副产物(II)、(IV)和/或(III)中的至少一种或多种的混合物产生。

[1215] 47.根据实施方案33-46中任一项所述的方法,其中使用有机溶剂或汽提/蒸馏步骤或过滤从所述生物转化反应混合物分离(-)-降龙涎醚。

[1216] 48.根据实施方案33-47中任一项所述的方法,其中使用有机溶剂或汽提/蒸馏步骤从所述生物转化反应混合物的固相分离(-)-降龙涎醚。

[1217] 49.根据实施方案47或实施方案48所述的方法,其中使用有机溶剂从所述反应混合物分离(-)-降龙涎醚。

[1218] 50.根据实施方案49所述的方法,其中使用乙醇或甲苯从所述反应混合物分离(-)-降龙涎醚。

[1219] 51.根据实施方案47-49中任一项所述的方法,其中使用有机溶剂使所述(-)-降龙涎醚选择性结晶。

[1220] 52.根据实施方案51所述的方法,其中所述(-)-降龙涎醚基本上不含副产物(II)、(IV)和/或(III)。

[1221] 53.根据实施方案33-52中任一项所述的方法,其中产生了浓度范围为约125-200g/l的(-)-降龙涎醚。

[1222] 54.根据实施方案33-53中任一项所述的方法,其中所述(-)-降龙涎醚具有约0.1至约0.5ng/l的嗅觉阈值。

序列表

<110> Givaudan SA
 <120> 酶及其应用
 <130> 30578
 <150> GB 1507207.7
 <151> 2015-04-24
 <160> 188
 <170> PatentIn version 3.3
 <210> 1
 <211> 631
 <212> PRT
 <213> 酸热脂环酸芽孢杆菌 (*Alicyclobacillus acidocaldarius*)
 <400> 1

```

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

```

195	200	205
Gly Ala Lys Gly Gly Gly Gly Trp Ile Phe Asp Ala Leu Asp Arg Ala		
210	215	220
Leu His Gly Tyr Gln Lys Leu Ser Val His Pro Phe Arg Arg Ala Ala		
225	230	235
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> 725
 <212> PRT
 <213> 运动发酵单胞菌(*Zymomonas mobilis*)
 <400> 2
 Met Gly Ile Asp Arg Met Asn Ser Leu Ser Arg Leu Leu Met Lys Lys
 1 5 10 15
 Ile Phe Gly Ala Glu Lys Thr Ser Tyr Lys Pro Ala Ser Asp Thr Ile
 20 25 30
 Ile Gly Thr Asp Thr Leu Lys Arg Pro Asn Arg Arg Pro Glu Pro Thr
 35 40 45
 Ala Lys Val Asp Lys Thr Ile Phe Lys Thr Met Gly Asn Ser Leu Asn
 50 55 60
 Asn Thr Leu Val Ser Ala Cys Asp Trp Leu Ile Gly Gln Gln Lys Pro
 65 70 75 80
 Asp Gly His Trp Val Gly Ala Val Glu Ser Asn Ala Ser Met Glu Ala
 85 90 95
 Glu Trp Cys Leu Ala Leu Trp Phe Leu Gly Leu Glu Asp His Pro Leu
 100 105 110
 Arg Pro Arg Leu Gly Asn Ala Leu Leu Glu Met Gln Arg Glu Asp Gly
 115 120 125
 Ser Trp Gly Val Tyr Phe Gly Ala Gly Asn Gly Asp Ile Asn Ala Thr
 130 135 140

Val	Glu	Ala	Tyr	Ala	Ala	Leu	Arg	Ser	Leu	Gly	Tyr	Ser	Ala	Asp	Asn
145						150				155					160
Pro	Val	Leu	Lys	Lys	Ala	Ala	Ala	Trp	Ile	Ala	Glu	Lys	Gly	Gly	Leu
						165				170					175
Lys	Asn	Ile	Arg	Val	Phe	Thr	Arg	Tyr	Trp	Leu	Ala	Leu	Ile	Gly	Glu
						180				185					190
Trp	Pro	Trp	Glu	Lys	Thr	Pro	Asn	Leu	Pro	Pro	Glu	Ile	Ile	Trp	Phe
						195				200					205
Pro	Asp	Asn	Phe	Val	Phe	Ser	Ile	Tyr	Asn	Phe	Ala	Gln	Trp	Ala	Arg
						210				215					220
Ala	Thr	Met	Val	Pro	Ile	Ala	Ile	Leu	Ser	Ala	Arg	Arg	Pro	Ser	Arg
225						230				235					240
Pro	Leu	Arg	Pro	Gln	Asp	Arg	Leu	Asp	Glu	Leu	Phe	Pro	Glu	Gly	Arg
						245				250					255
Ala	Arg	Phe	Asp	Tyr	Glu	Leu	Pro	Lys	Lys	Glu	Gly	Ile	Asp	Leu	Trp
						260				265					270
Ser	Gln	Phe	Phe	Arg	Thr	Thr	Asp	Arg	Gly	Leu	His	Trp	Val	Gln	Ser
						275				280					285
Asn	Leu	Leu	Lys	Arg	Asn	Ser	Leu	Arg	Glu	Ala	Ala	Ile	Arg	His	Val
						290				295					300
Leu	Glu	Trp	Ile	Ile	Arg	His	Gln	Asp	Ala	Asp	Gly	Gly	Trp	Gly	Gly
305						310				315					320
Ile	Gln	Pro	Pro	Trp	Val	Tyr	Gly	Leu	Met	Ala	Leu	His	Gly	Glu	Gly
						325				330					335
Tyr	Gln	Leu	Tyr	His	Pro	Val	Met	Ala	Lys	Ala	Leu	Ser	Ala	Leu	Asp
						340				345					350
Asp	Pro	Gly	Trp	Arg	His	Asp	Arg	Gly	Glu	Ser	Ser	Trp	Ile	Gln	Ala
						355				360					365
Thr	Asn	Ser	Pro	Val	Trp	Asp	Thr	Met	Leu	Ala	Leu	Met	Ala	Leu	Lys
						370				375					380
Asp	Ala	Lys	Ala	Glu	Asp	Arg	Phe	Thr	Pro	Glu	Met	Asp	Lys	Ala	Ala
385						390				395					400
Asp	Trp	Leu	Leu	Ala	Arg	Gln	Val	Lys	Val	Lys	Gly	Asp	Trp	Ser	Ile
						405				410					415
Lys	Leu	Pro	Asp	Val	Glu	Pro	Gly	Gly	Trp	Ala	Phe	Glu	Tyr	Ala	Asn
						420				425					430
Asp	Arg	Tyr	Pro	Asp	Thr	Asp	Asp	Thr	Ala	Val	Ala	Leu	Ile	Ala	Leu
						435				440					445
Ser	Ser	Tyr	Arg	Asp	Lys	Glu	Glu	Trp	Gln	Lys	Lys	Gly	Val	Glu	Asp

450	455	460
Ala Ile Thr Arg Gly Val	Asn Trp Leu Ile	Ala Met Gln Ser Glu Cys
465	470	475
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
Tyr Gly Thr Gly Ala Val	Leu Pro Ala Leu 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
Ile Asp Ser Ile Gly Lys	Gly Pro Thr Thr Pro	Ser Gln Thr Ala Trp
610	615	620
Ala Leu Met Gly Leu Ile	Ala Ala Asn Arg Pro	Glu Asp Tyr Glu Ala
625	630	635
Ile Ala Lys Gly Cys His	Tyr Leu Ile Asp Arg	Gln Glu Gln Asp Gly
645	650	655
Ser Trp Lys Glu Glu Glu	Phe Thr Gly Thr Gly	Phe Pro Gly Tyr Gly
660	665	670
Val Gly Gln Thr Ile Lys	Leu Asp Asp Pro Ala	Leu Ser Lys Arg Leu
675	680	685
Leu Gln Gly Ala Glu Leu	Ser Arg Ala Phe Met	Leu Arg Tyr Asp Phe
690	695	700
Tyr Arg Gln Phe Phe Pro	Ile Met Ala Leu Ser	Arg Ala Glu Arg Leu
705	710	715
Ile Asp Leu Asn Asn		
725		

<210> 3

<211> 658

<212> PRT

<213> 运动发酵单胞菌(*Zymomonas mobilis*)

<400> 3

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
Ala Arg Tyr Arg Asn Leu Lys Lys Ala Asn Gln Pro Ile Val His Tyr		640
	645	650
		655
Gly Met		
<210> 4		
<211> 684		
<212> PRT		
<213> 慢生大豆根瘤菌 (<i>Bradyrhizobium japonicum</i>)		
<400> 4		
Met Thr Val Thr Ser Ser Ala Ser Ala Arg Ala Thr Arg Asp Pro Gly		
1	5	10
Asn Tyr Gln Thr Ala Leu Gln Ser Thr Val Arg Ala Ala Ala Asp Trp		
	20	25
		30
Leu Ile Ala Asn Gln Lys Pro Asp Gly His Trp Val Gly Arg Ala Glu		
	35	40
		45
Ser Asn Ala Cys Met Glu Ala Gln Trp Cys Leu Ala Leu Trp Phe Met		
	50	55
		60
Gly Leu Glu Asp His Pro Leu Arg Lys Arg Leu Gly Gln Ser Leu Leu		
65	70	75
		80
Asp Ser Gln Arg Pro Asp Gly Ala Trp Gln Val Tyr Phe Gly Ala Pro		
	85	90
		95
Asn Gly Asp Ile Asn Ala Thr Val Glu Ala Tyr Ala Ala Leu Arg Ser		
	100	105
		110
Leu Gly Phe Arg Asp Asp Glu Pro Ala Val Arg Arg Ala Arg Glu Trp		
	115	120
		125
Ile Glu Ala Lys Gly Gly Leu Arg Asn Ile Arg Val Phe Thr Arg Tyr		
	130	135
		140
Trp Leu Ala Leu Ile Gly Glu Trp Pro Trp Glu Lys Thr Pro Asn Ile		
145	150	155
		160
Pro Pro Glu Val Ile Trp Phe Pro Leu Trp Phe Pro Phe Ser Ile Tyr		
	165	170
		175
Asn Phe Ala Gln Trp Ala Arg Ala Thr Leu Met Pro Ile Ala Val Leu		
	180	185
		190
Ser Ala Arg Arg Pro Ser Arg Pro Leu Pro Pro Glu Asn Arg Leu Asp		
	195	200
		205
Ala Leu Phe Pro His Gly Arg Lys Ala Phe Asp Tyr Glu Leu Pro Val		
210	215	220

Lys	Ala	Gly	Ala	Gly	Gly	Trp	Asp	Arg	Phe	Phe	Arg	Gly	Ala	Asp	Lys
225						230					235				240
Val	Leu	His	Lys	Leu	Gln	Asn	Leu	Gly	Asn	Arg	Leu	Asn	Leu	Gly	Leu
				245						250					255
Phe	Arg	Pro	Ala	Ala	Thr	Ser	Arg	Val	Leu	Glu	Trp	Met	Ile	Arg	His
			260						265					270	
Gln	Asp	Phe	Asp	Gly	Ala	Trp	Gly	Gly	Ile	Gln	Pro	Pro	Trp	Ile	Tyr
		275					280						285		
Gly	Leu	Met	Ala	Leu	Tyr	Ala	Glu	Gly	Tyr	Pro	Leu	Asn	His	Pro	Val
		290				295						300			
Leu	Ala	Lys	Gly	Leu	Asp	Ala	Leu	Asn	Asp	Pro	Gly	Trp	Arg	Val	Asp
305					310					315					320
Val	Gly	Asp	Ala	Thr	Tyr	Ile	Gln	Ala	Thr	Asn	Ser	Pro	Val	Trp	Asp
				325						330					335
Thr	Ile	Leu	Thr	Leu	Leu	Ala	Phe	Asp	Asp	Ala	Gly	Val	Leu	Gly	Asp
			340						345					350	
Tyr	Pro	Glu	Ala	Val	Asp	Lys	Ala	Val	Asp	Trp	Val	Leu	Gln	Arg	Gln
		355					360						365		
Val	Arg	Val	Pro	Gly	Asp	Trp	Ser	Met	Lys	Leu	Pro	His	Val	Lys	Pro
		370					375					380			
Gly	Gly	Trp	Ala	Phe	Glu	Tyr	Ala	Asn	Asn	Tyr	Tyr	Pro	Asp	Thr	Asp
385					390					395					400
Asp	Thr	Ala	Val	Ala	Leu	Ile	Ala	Leu	Ala	Pro	Leu	Arg	His	Asp	Pro
				405						410				415	
Lys	Trp	Lys	Ala	Lys	Gly	Ile	Asp	Glu	Ala	Ile	Gln	Leu	Gly	Val	Asp
			420						425				430		
Trp	Leu	Ile	Gly	Met	Gln	Ser	Gln	Gly	Gly	Gly	Trp	Gly	Ala	Phe	Asp
		435					440						445		
Lys	Asp	Asn	Asn	Gln	Lys	Ile	Leu	Thr	Lys	Ile	Pro	Phe	Cys	Asp	Tyr
		450				455					460				
Gly	Glu	Ala	Leu	Asp	Pro	Pro	Ser	Val	Asp	Val	Thr	Ala	His	Ile	Ile
465					470					475					480
Glu	Ala	Phe	Gly	Lys	Leu	Gly	Ile	Ser	Arg	Asn	His	Pro	Ser	Met	Val
				485						490				495	
Gln	Ala	Leu	Asp	Tyr	Ile	Arg	Arg	Glu	Gln	Glu	Pro	Ser	Gly	Pro	Trp
			500						505				510		
Phe	Gly	Arg	Trp	Gly	Val	Asn	Tyr	Val	Tyr	Gly	Thr	Gly	Ala	Val	Leu
		515					520					525			
Pro	Ala	Leu	Ala	Ala	Ile	Gly	Glu	Asp	Met	Thr	Gln	Pro	Tyr	Ile	Gly

530	535	540
Arg Ala Cys Asp Trp Leu Val Ala His Gln Gln Ala Asp Gly Gly Trp		
545	550	555
Gly Glu Ser Cys Ala Ser Tyr Met Asp Val Ser Ala Val Gly Arg Gly		560
565	570	575
Thr Thr Thr Ala Ser Gln Thr Ala Trp Ala Leu Met Ala Leu Leu Ala		
580	585	590
Ala Asn Arg Pro Gln Asp Lys Asp Ala Ile Glu Arg Gly Cys Met Trp		
595	600	605
Leu Val Glu Arg Gln Ser Ala Gly Thr Trp Asp Glu Pro Glu Phe Thr		
610	615	620
Gly Thr Gly Phe Pro Gly Tyr Gly Val Gly Gln Thr Ile Lys Leu Asn		
625	630	635
Asp Pro Ala Leu Ser Gln Arg Leu Met Gln Gly Pro Glu Leu Ser Arg		640
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> 5		
<211> 631		
<212> PRT		
<213> 人工序列		
<220>		
<223> 合成AacSHC衍生物		
<400> 5		
Met Ala Glu Gln Leu Val Glu Ala Pro Ala Tyr Ala Arg Thr Leu Asp		
1	5	10
Arg Ala Val Glu Tyr Leu Leu Ser Cys Gln Lys Asp Glu Gly Tyr Trp		15
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 Ala Trp Ala Leu		
65	70	75
Tyr Pro Gly Gly Pro Pro Asp Leu Asp Thr Thr Ile Glu Ala Tyr Val		80
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
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	470 475 480
Pro Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Leu Tyr Gly	
	485 490 495
Thr Gly Ala Val Val Ser Ala Leu Lys Ala Val Gly Ile Asp Thr Arg	
	500 505 510
Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn	
	515 520 525
Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala	
	530 535 540
Tyr Ala Gly Lys Gly Ala Ser Thr Pro Ser Gln Thr Ala Trp Ala Leu	
545	550 555 560
Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg	
	565 570 575
Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp	
	580 585 590
Glu Pro Tyr Tyr Thr Gly Thr Gly Phe Pro Gly Asp Phe Tyr Leu Gly	
	595 600 605
Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr	
	610 615 620
Lys Gln Ala Ile Glu Arg Arg	
625	630

<210> 6

<211> 1896

<212> DNA

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 6

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 aggacggcgc gtgggccctg	240

```

taccggtg ggccgccga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcgcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggatgtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccac 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 cgctgcacgg gtatcagaag ctgtcggtgc accggttccg ccgcgcggcc 720
gagatccgcg ctttgactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgtcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggtg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttg 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 tgtcgcatg cagagctcga acggcggttg gggcgctac 1260
gacgtcgaca acacgagcga tctccgaac cacatccgt 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 tgaaggcgt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaaccgggac ggcggtggtg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taaggcgcg agcaccctgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg caggcgggag tccgaggccg cgcgcccgcg cgtgcaatac 1740
ctcgtggaga cgcagcgcgc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggtacacc atgtaccgcc acgtgtttcc gacgtcgcg 1860
ctcgccgct acaagcaagc catcgagcgc aggtga 1896

```

<210> 7

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<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
Tyr Pro Gly Gly Pro Pro Asp Leu Asp Thr Thr Val 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
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
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> 8		
<211> 1896		

<212> DNA

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 8

```

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 gtgggcccctg     240
taccgggtg ggccgccgga cctcgacacg accgtcgagg cgtacgtcgc gctcaagtat     300
atcgcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag     360
ggcgggatcg agtcgtcgcg cgtgttcacg cggatgtggc tggcgtggt gggagaatat     420
ccgtgggaga aggtgcccac ggtcccgcg gagatcatgt tcctcgcaa gcgcatgccg     480
ctcaacatct acgagtttgg ctctgtgggt cgggcgaccg tcgtggcgct ctcgattgtg     540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag     600
accgacgtgc ctccgcgccg gcgcggtgcc aaggaggagg gtgggtggat cttcgacgcg     660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc     720
gagatccgcg ccttggaactg gttgctcgag cgccaggccg gagacggcag ctggggcggg     780
attcagccgc ctgtgtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat     840
ccggcgttca tcaagggtg ggaaggctta 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 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 gaaccgggac ggcggctggg gcgaggactg ccgctcgtac    1620
gaggatccgg cgtacgcggg taaggcgcg agcaccctgt cgcagacggc ctgggcgctg    1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgcccgcg cgtgcaatac    1740
ctcgtggaga cgcagcgcgc ggacggcggc tgggatgagc cgtactacac cggcacgggc    1800
ttcccagggg atttctacct cggtacacc atgtaccgcc acgtgtttcc gacgtcgcg     1860
ctcggccgct acaagcaagc catcgagcgc aggtga                                1896

```

<210> 9

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<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
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
Leu Thr Arg Met Trp Leu Ala Leu Val Gly Glu Tyr Pro Trp Glu Lys
           130          135          140
Val Pro Met Val Pro Pro Glu Ile Met Phe Leu Gly Lys Arg Met Pro
145          150          155          160
Leu Asn Ile Tyr Glu Phe Gly Ser Trp Ala Arg Ala Thr Val Val Ala
           165          170          175
Leu Ser Ile Val Met Ser Arg Gln Pro Val Phe Pro Leu Pro Glu Arg
           180          185          190
Ala Arg Val Pro Glu Leu Tyr Glu Thr Asp Val Pro Pro Arg Arg Arg
           195          200          205
Gly Ala Lys Gly Gly Gly Gly Trp Ile Phe Asp Ala Leu Asp Arg Ala
           210          215          220
Leu His Gly Tyr Gln Lys Leu Ser Val His Pro Phe Arg Arg Ala Ala
225          230          235          240
Glu Ile Arg Ala Leu Asp Trp Leu Leu Glu Arg Gln Ala Gly Asp Gly
           245          250          255
Ser Trp Gly Gly Ile Gln Pro Pro Trp Phe Tyr Ala Leu Ile Ala Leu
           260          265          270

```

Lys Ile Leu Asp Met Thr Gln His Pro Ala Phe Ile Lys Gly Trp Glu		
275	280	285
Gly Leu Glu Leu Tyr Gly Val Glu Leu Asp Tyr Gly Gly Trp Met Phe		
290	295	300
Gln Ala Ser Ile Ser Pro Val Trp Asp Thr Gly Leu Ala Val Leu Ala		
305	310	315
320		
Leu Arg Ala Ala Gly Leu Pro Ala Asp His Asp Arg Leu Val Lys Ala		
325	330	335
Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala		
340	345	350
Val Lys Arg Pro Asn Leu Lys Pro Gly Gly Phe Ala Phe Gln Phe Asp		
355	360	365
Asn Val Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Val Trp Ala		
370	375	380
Leu Asn Thr Leu Arg Leu Pro Asp Glu Arg Arg Arg Arg Asp Ala Met		
385	390	395
400		
Thr Lys Gly Phe Arg Trp Ile Val Gly Met Gln Ser Ser Asn Gly Gly		
405	410	415
Trp Gly Ala Tyr Asp Val Asp Asn Thr Ser Asp Leu Pro Asn His Ile		
420	425	430
Pro Phe Cys Asp Phe Gly Glu Val Thr Asp Pro Pro Ser Glu Asp Val		
435	440	445
Thr Ala His Val Leu Glu Cys Phe Gly Ser Phe Gly Tyr Asp Asp Ala		
450	455	460
Trp Lys Val Ile Arg Arg Ala Val Glu Tyr Leu Lys Arg Glu Gln Lys		
465	470	475
480		
Pro Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Leu Tyr Gly		
485	490	495
Thr Gly Ala Val Val Ser Ala Leu Lys Ala Val Gly Ile Asp Thr Arg		
500	505	510
Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn		
515	520	525
Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala		
530	535	540
Tyr Ala Gly Lys Gly Ala Ser Thr Pro Ser Gln Thr Ala Trp Ala Leu		
545	550	555
560		
Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg		
565	570	575
Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp		

	580		585		590	
Glu Pro Tyr Tyr Thr Gly Thr Gly Phe Pro Gly Asp Phe Tyr Leu Gly						
595			600		605	
Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr						
610			615		620	
Lys Gln Ala Ile Glu Arg Arg						
625			630			
<210>	10					
<211>	1896					
<212>	DNA					
<213>	人工序列					
<220>						
<223>	合成AacSHC衍生物					
<400>	10					
atggctgagc agttggtgga agcgccggcc 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
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat						300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag						360
ggcgggatcg agtcgtcgcg cgtgctcacg cggatgtggc 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 aaggaggagg gtgggtggat cttcgacgcg						660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcgggtgc acccgttccg ccgcgcggcc						720
gagatccgcg ccttggaactg gttgctcgag cgccaggccg gagacggcag ctggggcggg						780
attcagccgc cttggtttta cgcgtcatc gcgtcaaga ttctcgacat gacgcagcat						840
ccggcggttca tcaagggtg ggaaggctta gagctgtacg gcgtggagct ggattacgga						900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg						960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg						1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcggccgaa cctcaagccg						1080
ggcgggttcg cgttccagtt cgacaacgtg tactaccgga acgtggacga cacggccgtc						1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg						1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgccctac						1260
gacgtcgaca acacgagcga tctcccgaa ccatcccggt 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 taaggcgcg agcaccctgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgcccgcg 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> 合成AacSHC衍生物

<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 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> 12

<211> 1896

<212> DNA

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 12

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 gtgggccctg	240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat	300
atcggcgatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag	360
ggcgggatcg agtcgtcgcg cgtgttcacg cggaggtggc tggcgctggt gggagaatat	420
ccgtgggaga aggtgccccat ggtcccgcg gagatcatgt tcctcgcaa gcgcatgccg	480
ctcaacatct acgagtttgg ctctgtgggct cgggcgaccg tcgtggcgct ctcgattgtg	540
atgagccgcc agccggtgtt cccgtgccc gagcgggcgc gcgtgcccga gctgtacgag	600
accgacgtgc ctccgcgccg gcgcggtgcc aaggaggagg gtgggtggat cttcgacgcg	660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcgtgac acccgttccg ccgcgcggcc	720
gagatccgcg ccttgactg gttgctcgag cgccaggccg gagacggcag ctggggcggg	780
attcagccgc cttggtttta cgcgtcatc gcgtcaaga ttctcgacat gacgcagcat	840
ccggcggttca tcaagggtg ggaaggtcta gagctgtacg gcgtggagct ggattacgga	900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg	960


```

ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cggtccagtt cgacaacgtg tactaccggg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaa cacatcccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggtctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taaggcgcg agcacccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgcccgcg cgtgcaatac 1740
ctcgtggaga cgcagcggc 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> 合成AacSHC衍生物

<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 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
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 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
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> 合成AacSHC衍生物

<400> 14

atggctgagc agttggtgga agcgccggcc tacgcgcgga cgctggatcg cgcggtggag	60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc	120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc	180
atggagaaga tccggcggtta cctgttgcaac gagcagcgcg aggacggcac gtgggccctg	240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat	300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag	360
ggcgggatcg agtcgtcgcg cgtgttcacg cggatgtggc tggcgctggt gggagaatat	420

```

ccgtgggaga aggtgcccac ggtcccgcgc gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctctgtgggct cgggcgcaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgcgc gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg tgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggaactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcggttca tcaagggtcg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgcac tgggcggtga agcgcgcgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactaccgag acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcacg cagagctcga acggcggttg gggcgccctac 1260
gacgtcgaca acacgagcga tctcccgaac cacatcccg 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 gaaccggac ggcggtggtg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taaggcgcg agcaccgccg cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgcccgcg cgtgcaatac 1740
ctcgtggaga cgcagcgcgc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

```

<210> 15

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 15

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
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
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	400
Thr Lys Gly Phe Arg Trp Ile Val Gly Met Gln Ser Ser Asn Gly Gly			
	405	410	415
Trp Gly Ala Tyr Asp Val Asp Asn Thr Ser Asp Leu Pro Asn His Thr			
	420	425	430
Pro Phe Cys Asp Phe Gly Glu Val Thr Asp Pro Pro Ser Glu Asp Val			
	435	440	445
Thr Ala His Val Leu Glu Cys Phe Gly Ser Phe Gly Tyr Asp Asp Ala			
	450	455	460
Trp Lys Val Ile Arg Arg Ala Val Glu Tyr Leu Lys Arg Glu Gln Lys			
465	470	475	480
Pro Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Leu Tyr Gly			
	485	490	495
Thr Gly Ala Val Val Ser Ala Leu Lys Ala Val Gly Ile Asp Thr Arg			
	500	505	510
Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn			
	515	520	525
Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala			
	530	535	540
Tyr Ala Gly Lys Gly Ala Ser Thr Pro Ser Gln Thr Ala Trp Ala Leu			
545	550	555	560
Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg			
	565	570	575
Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp			
	580	585	590
Glu Pro Tyr Tyr Thr Gly Thr Gly Phe Pro Gly Asp Phe Tyr Leu Gly			
	595	600	605
Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr			
	610	615	620
Lys Gln Ala Ile Glu Arg Arg			
625	630		
<210>	16		
<211>	1896		
<212>	DNA		
<213>	人工序列		
<220>			

<223> 合成AacSHC衍生物

<400> 16

```

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 gtgggacctg      240
taccggggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat      300
atcggcgatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag      360
ggcgggatcg agtcgtcgcg cgtgttcacg cggatgtggc tggcgctggt gggagaatat      420
ccgtgggaga aggtgccccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg      480
ctcaacatct acgagtttgg ctctgtgggt cgggcgaccg tcgtggcgct ctcgattgtg      540
atgagccgcc agccggtgtt cccgtgccc gagcgggcgc gcgtgcccga gctgtacgag      600
accgacgtgc ctccgcgcgg gcgcggtgcc aaggaggagg gtgggtggat cttcgacgcg      660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc      720
gagatccgcg ccttgactg gttgctcgag cgccaggccg gagacggcag ctggggcggg      780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat      840
ccggcggttca tcaagggtg ggaaggtcta gagctgtacg gcgtggagct ggattacgga      900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg      960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg     1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgccgaa cctcaagccg     1080
ggcgggttcg cgttccagtt cgacaacgtg tactaccgag acgtggacga cacggccgtc     1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg     1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgacctac     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 ggcggtgagg gcgaggactg ccgctcgtac     1620
gaggatccgg cgtacgcggg taaggcgcg agcacccgt cgagacggc ctgggcgctg     1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccggcg cgtgcaatac     1740
ctcgtggaga cgacgcgcc ggacggcggc tgggatgagc cgtactacac cggcacgggc     1800
ttcccagggg atttctacct cggtacacc atgtaccgcc acgtgtttcc gacgtcgcg     1860
ctcgccgct acaagcaagc catcgagcgc aggtga                                1896

```

<210> 17

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 17

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

290	295	300
Gln Ala Ser Ile Ser Pro Val Trp Asp Thr Gly Leu Ala Val Leu Ala		
305	310	315
Leu Arg Ala Ala Gly Leu Pro Ala Asp His Asp Arg Leu Val Lys Ala		320
	325	330
Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala		335
	340	345
Val Lys Arg Pro Asn Leu Lys Pro Gly Gly Phe Ala Phe Gln Phe Asp		350
	355	360
Asn Val Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Val Trp Ala		365
	370	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		400
	405	410
Trp Gly Ala Tyr Asp Val Asp Asn Thr Ser Asp Leu Pro Asn His Ile		415
	420	425
Pro Phe Cys Asp Phe Gly Glu Val Thr Asp Pro Pro Ser Glu Asp Val		430
	435	440
Thr Ala His Val Leu Glu Cys Phe Gly Ser Phe Gly Tyr Asp Asp Ala		445
	450	455
Trp Lys Val Ile Arg Arg Ala Val Glu Tyr Leu Lys Arg Glu Gln Lys		460
465	470	475
Pro Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Leu Tyr Gly		480
	485	490
Thr Gly Ala Val Val Ser Ala Leu Lys Ala Val Gly Ile Asp Thr Arg		495
	500	505
Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn		510
	515	520
Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala		525
	530	535
Tyr Ala Gly Lys Gly Ala Ser Thr Pro Ser Gln Thr Ala Trp Ala Leu		540
545	550	555
Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg		560
	565	570
Gly Val His Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp		575
	580	585
Glu Pro Tyr Tyr Thr Gly Thr Gly Phe Pro Gly Asp Phe Tyr Leu Gly		590
	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> 18

<211> 1896

<212> DNA

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 18

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	gtgggcccctg	240
tacccgggtg	ggccgccgga	cctcgacacg	accatcgagg	cgtagctcgc	gctcaagtat	300
atcggcattg	cgcgcgacga	ggagccgatg	cagaaggcgc	tccggttcat	tcagagccag	360
ggcgggatcg	agtcgtcgcg	cgtgttcacg	cggatgtggc	tggcgctggt	gggagaatat	420
ccgtgggaga	aggtgcccac	ggtcccgcgc	gagatcatgt	tcctcggcaa	gcgcattgccg	480
ctcaacatct	acgagtttgg	ctcgtgggct	cgggcgaccg	tcgtggcgct	ctcgattgtg	540
atgagccgcc	agccggtgtt	cccgtgccc	gagcgggcgc	gcgtgcccga	gctgtacgag	600
accgacgtgc	ctccgcgcgc	gcgcggtgcc	aaggaggggg	gtgggtggat	cttcgacgcg	660
ctcgaccggg	cgctgcacgg	gtatcagaag	ctgtcggtgc	acccgttccg	ccgcgcggcc	720
gagatccgcg	ccttgactg	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	agatcacggg	tccgggcgac	tgggcgggtga	agcggccgaa	cctcaagccg	1080
ggcgggttcg	cggtccagtt	cgacaacgtg	tactaccggg	acgtggacga	cacggccgtc	1140
gtggtgtggg	cgctcaaac	cctgcgcttg	ccggacgagc	gccgcaggcg	ggacgccatg	1200
acgaagggat	tccgctggat	tgtcggcatg	cagagctcga	acggcggttg	gggcgcctac	1260
gacgtcgaca	acacgagcga	tctcccgaac	cacatcccg	tctgcgactt	cggcgaagtg	1320
accgatccgc	cgtcagagga	cgtaaccgcc	cacgtgctcg	agtgtttcgg	cagcttcggg	1380
tacgatgacg	cctggaaggt	catccggcgc	gcggtggaat	atctcaagcg	ggagcagaag	1440
ccggacggca	gctggttcgg	tcgttggggc	gtcaattacc	tctacggcac	gggcgcggtg	1500
gtgtcggcgc	tgaaggcggt	cgggatcgac	acgcgcgagc	cgtacattca	aaaggcgctc	1560
gactgggtcg	agcagcatca	gaacccggac	ggcggctggg	gcgaggactg	ccgctcgtac	1620
gaggatccgg	cgtagcggg	taagggcgcg	agcaccctgt	cgcagacggc	ctgggcgctg	1680

atggcgctca tcgcgggctgg cagggcggag tccgaggccg cgcgccgcgg cgtgcattac 1740
 ctcgtggaga cgcagcgcgc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
 ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
 ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 19

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 19

Met Ala Glu Gln Leu Val Glu Ala Pro Ala Tyr Ala Arg Thr Leu Asp
 1 5 10 15

Arg Ala Val Glu Tyr Leu Leu Ser Cys Gln Lys Asp Glu Gly Tyr Trp
 20 25 30

Trp Gly Pro Leu Leu Ser Asn Val Thr Met Glu Ala Glu Tyr Val Leu
 35 40 45

Leu Cys His Ile Leu Asp Arg Val Asp Arg Asp Arg Met Glu Lys Ile
 50 55 60

Arg Arg Tyr Leu Leu His Glu Gln Arg Glu Asp Gly Thr Trp Ala Leu
 65 70 75 80

Tyr Pro Gly Gly Pro Pro Asp Leu Asp Thr Thr Ile Glu Ala Tyr Val
 85 90 95

Ala Leu Lys Tyr Ile Gly Met Ser Arg Asp Glu Glu Pro Met Gln Lys
 100 105 110

Ala Leu Arg Phe Ile Gln Ser Gln Gly Gly Ile Glu Ser Ser Arg Val
 115 120 125

Phe Thr Arg Met Trp Leu Ala Leu Val Gly Glu Tyr Pro Trp Glu Lys
 130 135 140

Val Pro Met Val Pro Pro Glu Ile Met Phe Leu Gly Lys Arg Met Pro
 145 150 155 160

Leu Asn Ile Tyr Glu Phe Gly Ser Trp Ala Arg Ala Thr Val Val Ala
 165 170 175

Leu Ser Ile Val Met Ser Arg Gln Pro Val Phe Pro Leu Pro Glu Arg
 180 185 190

Ala Arg Val Pro Glu Leu Tyr Glu Thr Asp Val Pro Pro Arg Arg Arg
 195 200 205

Gly Ala Lys Gly Gly Gly Gly Trp Ile Phe Asp Ala Leu Asp Arg Ala
 210 215 220

Leu His Gly Tyr Gln Lys Leu Ser Val His Pro Phe Arg Arg Ala Ala			
225	230	235	240
Glu Ile Arg Ala Leu Asp Trp Leu Leu Glu Arg Gln Ala Gly Asp Gly			
	245	250	255
Ser Trp Gly Gly Ile Gln Pro Pro Trp Phe Tyr Ala Leu Ile Ala Leu			
	260	265	270
Lys Ile Leu Asp Met Thr Gln His Pro Ala Phe Ile Lys Gly Trp Glu			
	275	280	285
Gly Leu Glu Leu Tyr Gly Val Glu Leu Asp Tyr Gly Gly Trp Met Phe			
	290	295	300
Gln Ala Ser Ile Ser Pro Val Trp Asp Thr Gly Leu Ala Val Leu Ala			
305	310	315	320
Leu Arg Ala Ala Gly Leu Pro Ala Asp His Asp Arg Leu Val Lys Ala			
	325	330	335
Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala			
	340	345	350
Val Lys Arg Pro Asn Leu Lys Pro Gly Gly Phe Ala Phe Gln Phe Asp			
	355	360	365
Asn Val Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Val Trp Ala			
	370	375	380
Leu Asn Thr Leu Arg Leu Pro Asp Glu Arg Arg Arg Arg Asp Ala Met			
385	390	395	400
Thr Lys Gly Phe Arg Trp Ile Val Gly Met Gln Ser Ser Asn Gly Gly			
	405	410	415
Trp Gly Ala Tyr Asp Val Asp Asn Thr Ser Asp Leu Pro Asn His Ile			
	420	425	430
Pro Phe Cys Asp Phe Gly Glu Val Thr Asp Pro Pro Ser Glu Asp Val			
	435	440	445
Thr Ala His Val Leu Glu Cys Phe Gly Ser Phe Gly Tyr Asp Asp Ala			
	450	455	460
Trp Lys Val Ile Arg Arg Ala Val Glu Tyr Leu Lys Arg Glu Gln Lys			
465	470	475	480
Pro Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Leu Tyr Gly			
	485	490	495
Thr Gly Ala Val Val Ser Ala Leu Lys Ala Val Gly Ile Asp Thr Arg			
	500	505	510
Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn			
	515	520	525
Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala			

530	535	540
Tyr Ala Gly Lys Gly Ala Ser Thr Pro Ser Gln Thr Ala Trp Ala Leu		
545	550	555 560
Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg		
565	570	575
Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp		
580	585	590
Glu Pro Tyr Tyr Thr Gly Thr Gly Tyr Pro Gly Asp Phe Tyr Leu Gly		
595	600	605
Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr		
610	615	620
Lys Gln Ala Ile Glu Arg Arg		
625	630	
<210> 20		
<211> 1896		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成AacSHC衍生物		
<400> 20		
atggctgagc agttggtgga agcgccggcc 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	
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat	300	
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag	360	
ggcgggatcg agtcgtcgcg cgtgttcacg cggatgtggc tggcgctggt gggagaatat	420	
ccgtgggaga aggtgccc atgtcccgcc gagatcatgt tcctcggcaa gcgcatgccg	480	
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg	540	
atgagccgcc agccggtgtt cccgctgcc gagcgggcgc gcgtgcccga gctgtacgag	600	
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg	660	
ctcgaccggg cgctgcacgg gtatcagaag ctgtcgggtc acccgttcg ccgcgcggcc	720	
gagatccgcg ccttggaactg gttgctcgag cgccaggccg 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 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 gggcgccctac 1260
gacgtcgaca acacgagcga tctcccgaac cacatcccggt 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 tgaaggcgggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggtctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taaggcgcg agcaccctgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgcccgcg cgtgcaatac 1740
ctcgtggaga cgcagcggc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
taccagggg atttctacct cggtacacc atgtaccgcc acgtgtttcc gacgtcgcg 1860
ctcgcccgct acaagcaagc catcgagcgc aggtga 1896

<210> 21

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 21

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 Asp Ala Met			
385	390	395	400
Thr Lys Gly Phe Arg Trp Ile Val Gly Met Gln Ser Ser Asn Gly Gly			
	405	410	415
Trp Gly Ala Tyr Asp Val Asp Asn Thr Ser Asp Leu Pro Asn His Thr			
	420	425	430
Pro Phe Cys Asp Phe Gly Glu Val Thr Asp Pro Pro Ser Glu Asp Val			
	435	440	445
Thr Ala His Val Leu Glu Cys Phe Gly Ser Phe Gly Tyr Asp Asp Ala			
450	455	460	

Trp Lys Val Ile Arg Arg Ala Val Glu Tyr Leu Lys Arg Glu Gln Lys
 465 470 475 480
 Pro Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Leu Tyr Gly
 485 490 495
 Thr Gly Ala Val Val Ser Ala Leu Lys Ala Val Gly Ile Asp Thr Arg
 500 505 510
 Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn
 515 520 525
 Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala
 530 535 540
 Tyr Ala Gly Lys Gly Ala Ser Thr Pro Ser Gln Thr Ala Trp Ala Leu
 545 550 555 560
 Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg
 565 570 575
 Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp
 580 585 590
 Glu Pro Tyr Tyr Thr Gly Thr Gly Phe Pro Gly Asp Phe Tyr Leu Gly
 595 600 605
 Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr
 610 615 620
 Lys Gln Ala Ile Glu Arg Arg
 625 630

<210> 22

<211> 1896

<212> DNA

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 22

atggctgagc agttggtgga agctccggcc tacgcgcgga cgctggatcg cgcggtggag 60
 tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
 acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
 atggagaaga tccggcggta cctgttgac gagcagcgcg aggacggcac gtgggccctg 240
 taccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
 atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
 ggcgggatcg agtcgtcgcg cgtgttcacg cggaggtggc tggcgctggt gggagaatat 420
 ccgtgggaga aggtgccccat ggtccccgcg gagatcatgt tcctcggcaa gcgcatgccg 480
 ctcaacatct acgagtttgg ctctgtgggct 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 cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
 ccggcggttca tcaagggtg ggaagggtcta gagctgtacg gcgtggagct ggattacgga 900
 ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
 ctgcgcgctg cggggcttcc ggccgatcac gaccgcttg tcaaggcggg cgagtggctg 1020
 ttggaccggc agatcacggg tccgggcgac tgggcgggtga agcgcccgaa cctcaagccg 1080
 ggcggggttcg cgttccagtt cgacaacgtg tactaccggg acgtggacga cacggccgtc 1140
 gtgggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
 acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgccctac 1260
 gacgtcgaca acacgagcga tctcccgaac cacacccgt 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 tgaaggcggg cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
 gactgggtcg agcagcatca gaacccggac ggcggtgagg gcgaggactg ccgctcgtac 1620
 gaggatccgg cgtacgcggg taaggcgcg agcaccctgt cgcagacggc ctgggcgctg 1680
 atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgcccgcg cgtgcaatac 1740
 ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
 ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
 ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 23

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 23

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	Ala
				210					215						220
Leu	His	Gly	Tyr	Gln	Lys	Leu	Ser	Val	His	Pro	Phe	Arg	Arg	Ala	Ala
225						230					235				240
Glu	Ile	Arg	Ala	Leu	Asp	Trp	Leu	Leu	Glu	Arg	Gln	Ala	Gly	Asp	Gly
				245					250						255
Ser	Trp	Gly	Gly	Ile	Gln	Pro	Pro	Trp	Phe	Tyr	Ala	Leu	Ile	Ala	Leu
				260					265						270
Lys	Ile	Leu	Asp	Met	Thr	Gln	His	Pro	Ala	Phe	Ile	Lys	Gly	Trp	Glu
				275					280						285
Gly	Leu	Glu	Leu	Tyr	Gly	Val	Glu	Leu	Asp	Tyr	Gly	Gly	Trp	Met	Phe
				290					295						300
Gln	Ala	Ser	Ile	Ser	Pro	Val	Trp	Asp	Thr	Gly	Leu	Ala	Val	Leu	Ala
305						310					315				320
Leu	Arg	Ala	Ala	Gly	Leu	Pro	Ala	Asp	His	Asp	Arg	Leu	Val	Lys	Ala
				325					330						335
Gly	Glu	Trp	Leu	Leu	Asp	Arg	Gln	Ile	Thr	Val	Pro	Gly	Asp	Trp	Ala
				340					345						350
Val	Lys	Arg	Pro	Asn	Leu	Lys	Pro	Gly	Gly	Phe	Ala	Phe	Gln	Phe	Asp
				355					360						365
Asn	Val	Tyr	Tyr	Pro	Asp	Val	Asp	Asp	Thr	Ala	Val	Val	Val	Trp	Ala
				370					375						380
Leu	Asn	Thr	Leu	Arg	Leu	Pro	Asp	Glu	Arg	Arg	Arg	Arg	Asp	Ala	Met

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

<210> 24

<211> 1896

<212> DNA

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 24

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	cctcgacacg	accatcgagg	cgtacgtcgc	gctcaagtat	300
atcggcattgt	cgcgcgacga	ggagccgatg	cagaaggcgc	tccggttcat	tcagagccag	360
ggcgggatcg	agtcgtcgcg	cgtgttcacg	cggaggtggc	tggcgctggt	gggagaatat	420
ccgtgggaga	aggtgccccat	ggtccccg	gagatcatgt	tcctcggcaa	gcgcattgccg	480
ctcaacatct	acgagtttgg	ctcgtgggct	cggcgacacg	tcgtggcgct	ctcgattgtg	540
atgagccgcc	agccggtgtt	cccgtgccc	gagcggcgcg	gcgtgcccga	gctgtacgag	600
accgacgtgc	ctccgcgcg	gcgcggtgcc	aagggagggg	gtgggtggat	cttcgacgcg	660
ctcgaccggg	cgctgcacgg	gtatcagaag	ctgtcggtgc	accggttcg	ccgcgcggcc	720
gagatccgcg	ccttggaactg	gttgtctgag	cgccaggccg	gagacggcag	ctggggcggg	780
attcagccgc	cttggtttta	cgcgtcctc	gcgtcaaga	ttctcgacat	gacgcagcat	840
ccggcggttca	tcaagggtg	ggaaggtcta	gagctgtacg	gcgtggagct	ggattacgga	900
ggatggatgt	ttcaggcttc	catctcgccg	gtgtgggaca	cgggcctcgc	cgtgtctcgcg	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	gggcgcctac	1260
gacgtcgaca	acacgagcga	tctcccgaa	cacacccgt	tctgcgactt	cggcgaagtg	1320
accgatccgc	cgtcagagga	cgtcaccgcc	cacgtgtctg	agtgtttcgg	cagcttcggg	1380
tacgatgacg	cctggaaggt	catccggcgc	gcggtggaat	atctcaagcg	ggagcagaag	1440
ccggacggca	gctgggttcg	tcgttggggc	gtcaattacc	tctacggcac	gggcgcggtg	1500
gtgtcggcgc	tgaaggcggt	cgggatcgac	acgcgcgagc	cgtacattca	aaaggcgctc	1560
gactgggtcg	agcagcatca	gaacccggac	ggcggctggg	gcgaggactg	ccgctcgtac	1620
gaggatccgg	cgtacgcggg	taaggcgcg	agcaccctgt	cgcagacggc	ctgggcgctg	1680
atggcgctca	tcgcgggcgg	cagggcggag	tccgaggccg	cgcgcccgcg	cgtgcaatac	1740
ctcgtggaga	cgcagcgcgc	ggacggcggc	tgggatgagc	cgtactacac	cggcacgggc	1800
ttcccagggg	atttctacct	cggctacacc	atgtaccgcc	acgtgtttcc	gacgtcgcg	1860
ctcggccgct	acaagcaagc	catcgagcgc	aggtga			1896

<210> 25

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 25

Met Ala Glu Gln Leu Val Glu Ala Pro Ala Tyr Ala Arg Thr Leu Asp

1	5	10	15
Arg Ala Val Glu Tyr Leu Leu Ser Cys Gln Lys Asp Glu Gly Tyr Trp			
20	25	30	
Trp Gly Pro Leu Leu Ser Asn Val Thr Met Glu Ala Glu Tyr Val Leu			
35	40	45	
Leu Cys His Ile Leu Asp Arg Val Asp Arg Asp Arg Met Glu Lys Ile			
50	55	60	
Arg Arg Tyr Leu Leu His Glu Gln Arg Glu Asp Gly Thr Trp Ala Leu			
65	70	75	80
Tyr Pro Gly Gly Pro Pro Asp Leu Asp Thr Thr Ile Glu Ala Tyr Val			
85	90	95	
Ala Leu Lys Tyr Ile Gly Met Ser Arg Asp Glu Glu Pro Met Gln Lys			
100	105	110	
Ala Leu Arg Phe Ile Gln Ser Gln Gly Gly Ile Glu Ser Ser Arg Val			
115	120	125	
Phe Thr Arg Met Trp Leu Ala Leu Val Gly Glu Tyr Pro Trp Glu Lys			
130	135	140	
Val Pro Met Val Pro Pro Glu Ile Met Phe Leu Gly Lys Arg Met Pro			
145	150	155	160
Leu Asn Ile Tyr Glu Phe Gly Ser Trp Ala Arg Ala Thr Val Val Ala			
165	170	175	
Leu Ser Ile Val Met Ser Arg Gln Pro Val Phe Pro Leu Pro Glu Arg			
180	185	190	
Ala Arg Val Pro Glu Leu Tyr Glu Thr Asp Val Pro Pro Arg Arg Arg			
195	200	205	
Gly Ala Lys Gly Gly Gly Gly Trp Ile Phe Asp Ala Leu Asp Arg Ala			
210	215	220	
Leu His Gly Tyr Gln Lys Leu Ser Val His Pro Phe Arg Arg Ala Ala			
225	230	235	240
Glu Ile Arg Ala Leu Asp Trp Leu Leu Glu Arg Gln Ala Gly Asp Gly			
245	250	255	
Ser Trp Gly Gly Ile Gln Pro Pro Trp Phe Tyr Ala Leu Ile Ala Leu			
260	265	270	
Lys Ile Leu Asp Met Thr Gln His Pro Ala Phe Ile Lys Gly Trp Glu			
275	280	285	
Gly Leu Glu Leu Tyr Gly Val Glu Leu Asp Tyr Gly Gly Trp Met Phe			
290	295	300	
Gln Ala Ser Ile Ser Pro Val Trp Asp Thr Gly Leu Ala Val Leu Ala			
305	310	315	320

Leu Arg Ala Ala Gly Leu Pro Ala Asp His Asp Arg Leu Val Lys Ala		
325	330	335
Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala		
340	345	350
Val Lys Arg Pro Asn Leu Lys Pro Gly Gly Phe Ala Phe Gln Phe Asp		
355	360	365
Asn Val Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Val Trp Ala		
370	375	380
Leu Asn Thr Leu Arg Leu Pro Asp Glu Arg Arg Arg Arg Asp Ala Met		
385	390	400
Thr Lys Gly Phe Arg Trp Ile Val Gly Met Gln Ser Ser Asn Gly Gly		
405	410	415
Trp Gly Ala Tyr Asp Val Asp Asn Thr Ser Asp Leu Pro Asn His Ile		
420	425	430
Pro Phe Cys Asp Phe Gly Glu Val Thr Asp Pro Pro Ser Glu Asp Val		
435	440	445
Thr Ala His Val Leu Glu Cys Phe Gly Ser Phe Gly Tyr Asp Asp Ala		
450	455	460
Trp Lys Val Ile Arg Arg Ala Val Glu Tyr Leu Lys Arg Glu Gln Lys		
465	470	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	560
Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg		
565	570	575
Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp		
580	585	590
Glu Pro Tyr Tyr Thr Gly Thr Gly Tyr Pro Gly Asp Phe Tyr Leu Gly		
595	600	605
Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr		
610	615	620
Lys Gln Ala Ile Glu Arg Arg		

625	630	
<210>	26	
<211>	1896	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成AacSHC衍生物	
<400>	26	
atggctgagc agttggtgga agcgccggcc tacgcgcgga cgctggatcg cgcggtggag		60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc		120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc		180
atggagaaga tccggcggtg cctgttgac gagcagcgcg aggacggcac gtgggcccctg		240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat		300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag		360
ggcgggatcg agtcgtcgcg cgtgttcacg cggatgtggc 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 ctccgcgcg gcgcggtgcc aaggaggagg gtgggtggat cttcgacgcg		660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcgtgac acccgttccg ccgcgcggcc		720
gagatccgcg ccttggaactg gttgctcgag cgccaggccg gagacggcag ctggggcggg		780
attcagccgc cttggtttta cgcgctcatc gcgtcaaga ttctcgacat gacgcagcat		840
ccggcggttca tcaagggtg ggaaggctta gagctgtacg gcgtggagct ggattacgga		900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg		960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg		1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg		1080
ggcgggttcg cgttccagtt cgacaacgtg tactaccgga acgtggacga cacggccgtc		1140
gtggtgtggg cgctcaacac cctgcgctt ccggacgagc gccgcaggcg ggacgccatg		1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgccctac		1260
gacgtcgaca acacgagcga tctcccgaac cacatccgt tctgcgactt cggcgaagtg		1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg		1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag		1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg		1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc		1560
gactgggtcg agcagcatca gaacccggac ggcggttggg gcgaggactg ccgctcgtac		1620
gaggatccgg cgtacgcggg taaggcgcg agcaccctgt cgcagacggc ctgggcgctg		1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgcccgcg cgtgcaatac		1740
ctcgtggaga cgcagcgcgc ggacggcggc tgggatgagc cgtactacac cggcacgggc		1800
taccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg		1860

ctcggccgct acaagcaagc catcgagcgc aggtga

1896

<210> 27

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 27

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

	245		250		255
Ser Trp Gly Gly Ile Gln Pro Pro Trp Phe Tyr Ala Leu Ile Ala Leu					
	260		265		270
Lys Ile Leu Asp Met Thr Gln His Pro Ala Phe Ile Lys Gly Trp Glu					
	275		280		285
Gly Leu Glu Leu Tyr Gly Val Glu Leu Asp Tyr Gly Gly Trp Met Phe					
	290		295		300
Gln Ala Ser Ile Ser Pro Val Trp Asp Thr Gly Leu Ala Val Leu Ala					
305		310		315	320
Leu Arg Ala Ala Gly Leu Pro Ala Asp His Asp Arg Leu Val Lys Ala					
	325		330		335
Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala					
	340		345		350
Val Lys Arg Pro Asn Leu Lys Pro Gly Gly Phe Ala Phe Gln Phe Asp					
	355		360		365
Asn Val Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Val Trp Ala					
	370		375		380
Leu Asn Thr Leu Arg Leu Pro Asp Glu Arg Arg Arg Arg Asp Ala Met					
385		390		395	400
Thr Lys Gly Phe Arg Trp Ile Val Gly Met Gln Ser Ser Asn Gly Gly					
	405		410		415
Trp Gly Ala Tyr Asp Val Asp Asn Thr Ser Asp Leu Pro Asn His Ile					
	420		425		430
Pro Phe Cys Asp Phe Gly Glu Val Thr Asp Pro Pro Ser Glu Asp Val					
	435		440		445
Thr Ala His Val Leu Glu Cys Phe Gly Ser Phe Gly Tyr Asp Asp Ala					
	450		455		460
Trp Lys Val Ile Arg Arg Ala Val Glu Tyr Leu Lys Arg Glu Gln Lys					
465		470		475	480
Pro Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Leu Tyr Gly					
	485		490		495
Thr Gly Ala Val Val Ser Ala Leu Lys Ala Val Gly Ile Asp Thr Arg					
	500		505		510
Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn					
	515		520		525
Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala					
	530		535		540
Tyr Ala Gly Lys Gly Ala Ser Thr Pro Ser Gln Thr Ala Trp Ala Leu					
545		550		555	560

Met	Ala	Leu	Ile	Ala	Gly	Gly	Arg	Ala	Glu	Ser	Glu	Ala	Ala	Arg	Arg		
					565						570						575
Gly	Val	Gln	Tyr	Leu	Val	Glu	Thr	Gln	Arg	Pro	Asp	Gly	Gly	Trp	Asp		
					580						585						590
Glu	Pro	Tyr	Tyr	Thr	Gly	Thr	Gly	Phe	Pro	Gly	Asp	Phe	Tyr	Leu	Gly		
					595						600						605
Tyr	Thr	Met	Tyr	Arg	His	Val	Phe	Pro	Thr	Leu	Ala	Leu	Gly	Arg	Tyr		
					610						615						620
Lys	Gln	Ala	Ile	Glu	Arg	Arg											
							625									630	

<210> 28

<211> 1896

<212> DNA

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 28

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	aggacggcgc	gtgggcccctg	240
tacccgggtg	ggccgccgga	cctcgacacg	accgtcgagg	cgtacgtcgc	gctcaagtat	300
atcggcgatgt	cgcgcgacga	ggagccgatg	cagaaggcgc	tccggttcat	tcagagccag	360
ggcgggatcg	agtcgtcgcg	cgtgctcacg	cggtatgtggc	tggcgctggt	gggagaatat	420
ccgtgggaga	aggtgcccct	ggtcccgcgc	gagatcatgt	tcctcggcaa	gcgcatgccg	480
ctcaacatct	acgagtttgg	ctcgtgggct	cgggcgaccg	tcgtggcgct	ctcgattgtg	540
atgagccgcc	agccggtgtt	cccgtgccc	gagcgggcgc	gcgtgcccga	gctgtacgag	600
accgacgtgc	ctccgcgcgc	gcgcggtgcc	aagggagggg	gtgggtggat	cttcgacgcg	660
ctcgaccggg	cgctgcacgg	gtatcagaag	ctgtcggtgc	accggttccg	ccgcgcggcc	720
gagatccgcg	ccttggaactg	gttgctcgag	cgccaggccg	gagacggcag	ctggggcggg	780
attcagccgc	cttggtttta	cgcgctcatc	gcgctcaaga	ttctcgacat	gacgcagcat	840
ccggcggttca	tcaagggctg	ggaaggtcta	gagctgtacg	gcgtggagct	ggattacgga	900
ggatggatgt	ttcaggttcc	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	gggcgcctac	1260
gacgtcgaca	acacgagcga	tctcccgaac	cacatcccgt	tctgcgactt	cggcgaagtg	1320

accgatccgc cgtcagagga cgtcacgcc cactgtctcg agtgtttcgg cagcttcggg 1380
 tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
 ccggacggca gctgggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
 gtgtcggcgc tgaaggcggg cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
 gactgggtcg agcagcatca gaacccggac ggcggtgagg gcgaggactg ccgctcgtac 1620
 gaggatccgg cgtacgcggg taagggcgcg agcaccccg cgcagacggc ctgggcgctg 1680
 atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgcccgcg cgtgcaatac 1740
 ctcgtggaga cgcagcgcgc ggacggcgcg tgggatgagc cgtactacac cggcacgggc 1800
 ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
 ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 29

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 29

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 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
Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg		
565	570	575
Gly Val His Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp		
580	585	590
Glu Pro Tyr Tyr Thr Gly Thr Gly Tyr Pro Gly Asp Phe Tyr Leu Gly		
595	600	605
Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr		
610	615	620
Lys Gln Ala Ile Glu Arg Arg		
625	630	

<210> 30

<211> 1896

<212> DNA

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 30

atggctgagc agttggtgga agcaccggcc 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
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat	300
atcggcattgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag	360
ggcgggatcg agtcgtcgcg cgtgttcacg cggatgtggc tggcgctggt gggagaatat	420
ccgtgggaga aggtgcccac ggtcccgcg gagatcatgt tcctcggcaa gcgcatgccg	480
ctcaacatct acgagtttgg ctctgtgggt cgggcgaccg tcgtggcgct ctcgattgtg	540
atgagccgcc agccggtgtt cccgtgccc gagcgggcgc gcgtgcccga gctgtacgag	600
accgacgtgc ctccgcgccg gcgcggtgcc aaggaggagg gtgggtggat cttcgacgcg	660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc	720
gagatccgcg ccttggaactg gttgctcgag cgccaggccg gagacggcag ctggggcggg	780

```

attcagccgc cttggtttta cgcgtcatc gcgtcaaga ttctcgacat gacgcagcat      840
ccggcggttca tcaagggtg ggaaggctta gagctgtacg gcgtggagct ggattacgga      900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg      960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg     1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcggccgaa 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 cacatccgtt tctgcgactt cggcgaagtg     1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg     1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag     1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg     1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc     1560
gactgggtcg agcagcatca gaacccggac ggcggttggg gcgaggactg ccgctcgtac     1620
gaggatccgg cgtacgcggg taagggcgcg agcaccctgt cgcagacggc ctgggcgctg     1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgcccgcg cgtgcattac     1740
ctcgtggaga cgcagcggc ggacggcggc tgggatgagc cgtactacac cggcacgggc     1800
taccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgtcgcg      1860
ctcggccgct acaagcaagc catcgagcgc aggtga                                1896

```

<210> 31

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 31

```

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
Leu 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
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	470 475 480
Pro Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Leu Tyr Gly	
	485 490 495
Thr Gly Ala Val Val Ser Ala Leu Lys Ala Val Gly Ile Asp Thr Arg	
	500 505 510
Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn	
	515 520 525
Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala	
	530 535 540
Tyr Ala Gly Lys Gly Ala Ser Thr Pro Ser Gln Thr Ala Trp Ala Leu	
545	550 555 560
Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg	
	565 570 575
Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp	
	580 585 590
Glu Pro Tyr Tyr Thr Gly Thr Gly Phe Pro Gly Asp Phe Tyr Leu Gly	
	595 600 605
Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr	
	610 615 620
Lys Gln Ala Ile Glu Arg Arg	
625	630

<210> 32

<211> 1896

<212> DNA

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 32

atggctgagc agttggtgga agcgccggcc 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


```

taccggtg ggccgccga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcgcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgctcacg cggatgtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccac 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 cgctgcacgg gtatcagaag ctgtcggtgc accggttccg ccgcgcggcc 720
gagatccgcg ctttgactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgtcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggtg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttg 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 tgtcgcatg cagagctcga acggcggttg gggcgctac 1260
gacgtcgaca acacgagcga tctcccgaac cacatccgt 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 tgaaggcgtt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaaccgggac ggcggtggtg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taaggcgcg agcaccctgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg caggcgggag tccgaggccg cgcgcccgcg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggtacacc atgtaccgcc acgtgtttcc gacgtcgcg 1860
ctcgccgct acaagcaagc catcgagcgc aggtga 1896

```

<210> 33

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 33

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
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
Leu 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
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
Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg		
565	570	575
Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp		
580	585	590
Glu Pro Tyr Tyr Thr Gly Thr Gly Tyr Pro Gly Asp Phe Tyr Leu Gly		
595	600	605
Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr		
610	615	620
Lys Gln Ala Ile Glu Arg Arg		
625	630	
<210> 34		
<211> 1896		

<212> DNA

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 34

```

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 gtgggcccctg     240
taccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat     300
atcgcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag     360
ggcgggatcg agtcgtcgcg cgtgctcacg cggatgtggc tggcgtggt gggagaatat     420
ccgtgggaga aggtgcccac ggtcccgcg gagatcatgt tcctcggcaa gcgcatgccg     480
ctcaacatct acgagtttgg ctctgtgggt cgggcgaccg tcgtggcgct ctcgattgtg     540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag     600
accgacgtgc ctccgcgccg gcgcggtgcc aaggaggagg gtgggtggat cttcgacgcg     660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc     720
gagatccgcg ccttggaactg gttgctcgag cgccaggccg gagacggcag ctggggcggg     780
attcagccgc ctgtgtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat     840
ccggcggttca tcaagggtg ggaaggctta 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 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 gaaccgggac ggcggctggg gcgaggactg ccgctcgtac    1620
gaggatccgg cgtacgcggg taaggcgcg agcaccctgt cgcagacggc ctgggcgctg    1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgcccgcg cgtgcaatac    1740
ctcgtggaga cgcagcgcgc ggacggcggc tgggatgagc cgtactacac cggcacgggc    1800
taccaggggg atttctacct cggtacacc atgtaccgcc acgtgtttcc gacgtcgcg     1860
ctcggccgct acaagcaagc catcgagcgc aggtga                                1896

```

<210> 35

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 35

```

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
Leu 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 Ala
           210          215          220
Leu His Gly Tyr Gln Lys Leu Ser Val His Pro Phe Arg Arg Ala Ala
225          230          235          240
Glu Ile Arg Ala Leu Asp Trp Leu Leu Glu Arg Gln Ala Gly Asp Gly
           245          250          255
Ser Trp Gly Gly Ile Gln Pro Pro Trp Phe Tyr Ala Leu Ile Ala Leu
           260          265          270

```

Lys Ile Leu Asp Met Thr Gln His Pro Ala Phe Ile Lys Gly Trp Glu		
275	280	285
Gly Leu Glu Leu Tyr Gly Val Glu Leu Asp Tyr Gly Gly Trp Met Phe		
290	295	300
Gln Ala Ser Ile Ser Pro Val Trp Asp Thr Gly Leu Ala Val Leu Ala		
305	310	315
Leu Arg Ala Ala Gly Leu Pro Ala Asp His Asp Arg Leu Val Lys Ala		
	325	330
Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala		
	340	345
Val Lys Arg Pro Asn Leu Lys Pro Gly Gly Phe Ala Phe Gln Phe Asp		
	355	360
Asn Val Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Val Trp Ala		
	370	375
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
Trp Gly Ala Tyr Asp Val Asp Asn Thr Ser Asp Leu Pro Asn His Thr		
	420	425
Pro Phe Cys Asp Phe Gly Glu Val Thr Asp Pro Pro Ser Glu Asp Val		
	435	440
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
Thr Gly Ala Val Val Ser Ala Leu Lys Ala Val Gly Ile Asp Thr Arg		
	500	505
Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn		
	515	520
Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala		
	530	535
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
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>	36		
<211>	1896		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	合成AacSHC衍生物		
<400>	36		
atggctgagc agttggtgga agcgccggcc 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
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat			300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag			360
ggcgggatcg agtcgtcgcg cgtgctcacg 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 aaggaggagg gtgggtggat cttcgacgcg			660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcgggtgc acccgttccg ccgcgcggcc			720
gagatccgcg ccttggaactg gttgctcgag cgccaggccg gagacggcag ctggggcggg			780
attcagccgc cttggtttta cgcgtcatc gcgtcaaga ttctcgacat gacgcagcat			840
ccggcggttca tcaagggtg ggaaggctta gagctgtacg gcgtggagct ggattacgga			900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg			960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg			1020
ttggaccggc agatcacggt tccgggcgac tgggcgggtga 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 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 tgaaggcggc cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
 gactgggtcg agcagcatca gaacccggac ggcggtctggg gcgaggactg ccgctcgtac 1620
 gaggatccgg cgtacgcggg taaggcgcg agcaccctgt cgcagacggc ctgggcgctg 1680
 atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgcccgcg cgtgcaatac 1740
 ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
 ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
 ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 37

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 37

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 Ala		
210	215	220
Leu His Gly Tyr Gln Lys Leu Ser Val His Pro Phe Arg Arg Ala Ala		
225	230	235
240		
Glu Ile Arg Ala Leu Asp Trp Leu Leu Glu Arg Gln Ala Gly Asp Gly		
245	250	255
Ser Trp Gly Gly Ile Gln Pro Pro Trp Phe Tyr Ala Leu Ile Ala Leu		
260	265	270
Lys Ile Leu Asp Met Thr Gln His Pro Ala Phe Ile Lys Gly Trp Glu		
275	280	285
Gly Leu Glu Leu Tyr Gly Val Glu Leu Asp Tyr Gly Gly Trp Met Phe		
290	295	300
Gln Ala Ser Ile Ser Pro Val Trp Asp Thr Gly Leu Ala Val Leu Ala		
305	310	315
320		
Leu Arg Ala Ala Gly Leu Pro Ala Asp His Asp Arg Leu Val Lys Ala		
325	330	335
Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala		
340	345	350
Val Lys Arg Pro Asn Leu Lys Pro Gly Gly Phe Ala Phe Gln Phe Asp		
355	360	365
Asn Val Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Val Trp Ala		
370	375	380
Leu Asn Thr Leu Arg Leu Pro Asp Glu Arg Arg Arg Asp Ala Met		
385	390	395
400		
Thr Lys Gly Phe Arg Trp Ile Val Gly Met Gln Ser Ser Asn Gly Gly		
405	410	415
Trp Gly Ala Tyr Asp Val Asp Asn Thr Ser Asp Leu Pro Asn His Thr		
420	425	430
Pro Phe Cys Asp Phe Gly Glu Val Thr Asp Pro Pro Ser Glu Asp Val		
435	440	445
Thr Ala His Val Leu Glu Cys Phe Gly Ser Phe Gly Tyr Asp Asp Ala		
450	455	460
Trp Lys Val Ile Arg Arg Ala Val Glu Tyr Leu Lys Arg Glu Gln Lys		
465	470	475
480		
Pro Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Leu Tyr Gly		
485	490	495
Thr Gly Ala Val Val Ser Ala Leu Lys Ala Val Gly Ile Asp Thr Arg		
500	505	510

Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn	
515	520 525
Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala	
530	535 540
Tyr Ala Gly Lys Gly Ala Ser Thr Pro Ser Gln Thr Ala Trp Ala Leu	
545	550 555 560
Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg	
565	570 575
Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp	
580	585 590
Glu Pro Tyr Tyr Thr Gly Thr Gly Tyr Pro Gly Asp Phe Tyr Leu Gly	
595	600 605
Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr	
610	615 620
Lys Gln Ala Ile Glu Arg Arg	
625	630

<210> 38

<211> 1896

<212> DNA

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 38

atggctgagc agttggtgga agcgccggcc tacgcgcgga cgctggatcg cgcggtggag	60
tatctcctct cctgcaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc	120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc	180
atggagaaga tccggcggta cctgttgac gagcagcgcg aggacggcac gtgggacctg	240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat	300
atcggcgatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag	360
ggcgggatcg agtcgtcgcg cgtgttcacg cggaggtggc tggcgctggt gggagaatat	420
ccgtgggaga aggtgccccat ggtcccgcg gagatcatgt tcctcgcaa gcgcatgccg	480
ctcaacatct acgagtttgg ctctgtgggct cgggcgaccg tcgtggcgct ctcgattgtg	540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag	600
accgacgtgc ctccgcgccg gcgcggtgcc aaggaggagg gtgggtggat cttcgacgcg	660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc	720
gagatccgcg ccttgactg gttgctcgag cgccaggccg gagacggcag ctggggcggg	780
attcagccgc cttggtttta cgcgtcatc gcgtcaaga ttctcgacat gacgcagcat	840
ccggcggttca tcaagggtg ggaaggtcta gagctgtacg gcgtggagct ggattacgga	900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg	960

ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
 ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
 ggcggggttcg cggtccagtt cgacaacgtg tactaccggg acgtggacga cacggccgtc 1140
 gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
 acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
 gacgtcgaca acacgagcga tctcccgaaac cacaccccggt tctgcgactt cggcgaagtg 1320
 accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
 tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
 ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
 gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
 gactgggtcg agcagcatca gaacccggac ggcggtctggg gcgaggactg ccgctcgtac 1620
 gaggatccgg cgtacgcggg taaggcgcg agcaccccggt cgcagacggc ctgggcgctg 1680
 atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgcccgcg cgtgcaatac 1740
 ctcgtggaga cgcagcgcgc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
 taccagggg atttctacct cggtacacc atgtaccgcc acgtgtttcc gacgtcgcg 1860
 ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 39

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 39

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

Leu 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
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 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
Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg		
565	570	575
Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp		
580	585	590
Glu Pro Tyr Tyr Thr Gly Thr Gly Tyr Pro Gly Asp Phe Tyr Leu Gly		
595	600	605
Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr		
610	615	620
Lys Gln Ala Ile Glu Arg Arg		
625	630	

<210> 40

<211> 1896

<212> DNA

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 40

atggctgagc agttggtgga agcgccggcc tacgcgcgga cgctggatcg cgcggtggag	60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc	120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc	180
atggagaaga tccggcggtg cctgttgca cagcagcgcg aggcggcac gtgggccctg	240
taccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat	300
atcggcgatg cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag	360
ggcgggatcg agtcgtcgcg cgtgctcacg cggaggtggc tggcgctggt gggagaatat	420

```

ccgtgggaga aggtgcccac ggtcccgcgc gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctctgtgggct cgggcgcaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccgggtgtt cccgctgccc gagcggggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgcgc gcgcgggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcgggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggaactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcggttca tcaagggtcg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgcac tgggcgggtga agcggccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactaccgga acgtggacga cacggccgtc 1140
gtgggtgtgg 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 tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaaccggac ggcggtggtg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taaggcgcg agcacccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgcccgcg cgtgcaatac 1740
ctcgtggaga cgcagcgcgc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
taccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

```

<210> 41

<211> 725

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 41

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
Ala 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
Ile Asp Ser Ile Gly Lys Gly Pro Thr Thr Pro Ser Gln Thr Ala Trp			
	610	615	620
Ala Leu Met Gly Leu Ile Ala Ala Asn Arg Pro Glu Asp Tyr Glu Ala			
625	630	635	640
Ile Ala Lys Gly Cys His Tyr Leu Ile Asp Arg Gln Glu Gln Asp Gly			
	645	650	655
Ser Trp Lys Glu Glu Glu Phe Thr Gly Thr Gly Phe Pro Gly Tyr Gly			
	660	665	670
Val Gly Gln Thr Ile Lys Leu Asp Asp Pro Ala Leu Ser Lys Arg Leu			

675	680	685
Leu Gln Gly Ala Glu Leu Ser Arg Ala Phe Met	Leu Arg Tyr Asp Phe	
690	695	700
Tyr Arg Gln Phe Phe Pro Ile Met Ala Leu Ser	Arg Ala Glu Arg Leu	
705	710	715
Ile Asp Leu Asn Asn		

725

<210> 42

<211> 2178

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<220>

<221> misc_feature

<222> (387) .. (387)

<223> n是a, c, g或t

<400> 42

atgggtattg acagaatgaa tagcttaagt cgcttggttaa tgaagaagat tttcggggct	60
gaaaaaacct cgtataaacc ggcttccgat accataatcg gaacggatac cctgaaaaga	120
ccgaaccggc ggccatgaacc gacggcaaaa gtgcacaaaa cgatattcaa gactatgggg	180
aatagtctga ataataccct tgtttcagcc tgtgactggt tgatcggaca acaaaagccc	240
gatggtcatt gggtcgggtgc cgtggaatcc aatgcttcga tggaagcaga atgggtgtctg	300
gccttgtggg ttttgggtct ggaagatcat ccgcttcgtc caagattggg caatgctctt	360
ttggaaatgc agcgggaaga tggcgcntgg ggagtctatt tcggcgctgg aaatggcgat	420
atcaatgcca cggttgaagc ctatgcggcc ttgcgggtctt tgggggtattc tgccgataat	480
cctgttttga aaaaagcggc agcatggatt gctgaaaaag gcggattaaa aaatatccgt	540
gtctttaccc gttattggct ggcgttgatc ggggaatggc cttgggaaaa gaccctaac	600
cttccccctg aaattatctg gttccctgat aattttgtct tttcgattta taattttgcc	660
caatgggctg gggcaaccat ggtgccgatt gctattctgt ccgcgagacg accaagccgc	720
ccgctgcgcc ctcaagaccg attggatgaa ctgtttccag aaggccgcgc tcgctttgat	780
tatgaattgc cgaaaaaaga aggcacgat ctttggtcgc aatttttccg aaccactgac	840
cgtggattac attgggttca gtccaatctg ttaaagcgca atagcttgcg tgaagccgct	900
atccgtcatg ttttggaatg gattatccgg catcaggatg ccgatggcgg ttgggggtgga	960
attcagccac cttgggtcta tggtttgatg gcgttacatg gtgaaggcta tcagctttat	1020
catccggtga tggccaaggc tttgtcggct ttggatgatc ccggttggcg acatgacaga	1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg	1140
atggcggttaa aagacgcaa ggccgaggat cgttttacgc cggaatgga taaggccgcc	1200
gattggcttt tggctcgaca ggtcaaagtc aaaggcgatt ggtcaatcaa actgcccgat	1260

```

gttgaacccg gtggatgggc atttgaatat gccaatgata gctatcccga taccgatgat 1320
accgccgtcg ctttgcgcgc cctttcctct tatcgtgata aggaggagtg gcaaaagaaa 1380
ggcggttgagg acgccattac ccgtgggggtt aattgggtga tcgcatgca aagcgaatgt 1440
ggcggttgagg gagcctttga taaggataat aacagaagta tcctttccaa aattcctttt 1500
tgtgatttcg gagaatctat tgatccgcct tcagtcgatg taacggcgca tgttttagag 1560
gcctttggca ccttgggact gtcccgcgat atgccgggtca tccaaaaagc gatcgactat 1620
gtccgttccg aacaggaagc cgaaggcgcg tggtttggtc gttggggcgt taattatata 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gaccagcct 1740
tacatcacca aggcttgcca ttggctggtc gcacatcagc aggaagacgg cggttggggc 1800
gaaagctgct cttcctatat ggagattgat tccattggga agggcccaac cacgccgtcc 1860
cagactgctt gggctttgat ggggttgatc gcggccaatc gtcccgaaga ttatgaagcc 1920
attgccaagg gatgccatta tctgattgat cgccaagagc aggatggtag ctggaaagaa 1980
gaagaattca ccggcaccgg attccccggt tatggcgtgg gtcagacgat caagttagat 2040
gatccggctt tatcgaaacg attgcttcaa ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

```

<210> 43

<211> 725

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 43

```

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

```

Ser Trp Gly Val Tyr Phe Gly Ala Gly Asn Gly Asp Ile Asn Ala Thr			
130	135	140	
Val Glu Ala Tyr Ala Ala Leu Arg Ser Leu Gly Tyr Ser Ala Asp Asn			
145	150	155	160
Pro Val Leu Lys Lys Ala Ala Ala Trp Ile Ala Glu Lys Gly Gly Leu			
	165	170	175
Lys Asn Ile Arg Val Phe Thr Arg Tyr Trp Leu Ala Leu Ile Gly Glu			
	180	185	190
Trp Pro Trp Glu Lys Thr Pro Asn Leu Pro Pro Glu Ile Ile Trp Phe			
	195	200	205
Pro Asp Asn Phe Val Phe Ser Ile Tyr Asn Phe Ala Gln Trp Ala Arg			
	210	215	220
Ala Thr Met Val Pro Ile Ala Ile Leu Ser Ala Arg Arg Pro Ser Arg			
225	230	235	240
Pro Leu Arg Pro Gln Asp Arg Leu Asp Glu Leu Phe Pro Glu Gly Arg			
	245	250	255
Ala Arg Phe Asp Tyr Glu Leu Pro Lys Lys Glu Gly Ile Asp Leu Trp			
	260	265	270
Ser Gln Phe Phe Arg Thr Thr Asp Arg Gly Leu His Trp Val Gln Ser			
	275	280	285
Asn Leu Leu Lys Arg Asn Ser Leu Arg Glu Ala Ala Ile Arg His Val			
	290	295	300
Leu Glu Trp Ile Ile Arg His Gln Asp Ala Asp Gly Gly Trp Gly Gly			
305	310	315	320
Ile Gln Pro Pro Trp Val Tyr Gly Leu Met Ala Leu His Gly Glu Gly			
	325	330	335
Tyr Gln Leu Tyr His Pro Val Met Ala Lys Ala Leu Ser Ala Leu Asp			
	340	345	350
Asp Pro Gly Trp Arg His Asp Arg Gly Glu Ser Ser Trp Ile Gln Ala			
	355	360	365
Thr Asn Ser Pro Val Trp Asp Thr Met Leu Ala Leu Met Ala Leu Lys			
	370	375	380
Asp Ala Lys Ala Glu Asp Arg Phe Thr Pro Glu Met Asp Lys Ala Ala			
385	390	395	400
Asp Trp Leu Leu Ala Arg Gln Val Lys Val Lys Gly Asp Trp Ser Ile			
	405	410	415
Lys Leu Pro Asp Val Glu Pro Gly Gly Trp Ala Phe Glu Tyr Ala Asn			
	420	425	430
Asp Arg Tyr Pro Asp Thr Asp Asp Thr Ala Val Ala Leu Ile Ala Leu			

435	440	445
Ser Ser Tyr Arg Asp Lys Glu Glu Trp Gln Lys Lys Gly Val Glu Asp		
450	455	460
Ala Ile Thr Arg Gly Val Asn Trp Leu Ile Ala Met Gln Ser Glu Cys		
465	470	475
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
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
Ile Asp Ser Ile Gly Lys Gly Pro Thr Thr Pro Ser Gln Thr Ala Trp		
610	615	620
Ala Leu Met Gly Leu Ile Ala Ala Asn Arg Pro Glu Asp Tyr Glu Ala		
625	630	635
Ile Ala Lys Gly Cys His Tyr Leu Ile Asp Arg Gln Glu Gln Asp Gly		
645	650	655
Ser Trp Lys Glu Glu Glu Phe Thr Gly Thr Gly Phe Pro Gly Tyr Gly		
660	665	670
Val Gly Gln Thr Ile Lys Leu Asp Asp Pro Ala Leu Ser Lys Arg Leu		
675	680	685
Leu Gln Gly Ala Glu Leu Ser Arg Ala Phe Met Leu Arg Tyr Asp Phe		
690	695	700
Tyr Arg Gln Phe Phe Pro Ile Met Ala Leu Ser Arg Ala Glu Arg Leu		
705	710	715
Ile Asp Leu Asn Asn		
725		
<210> 44		
<211> 2178		

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 44

```

atgggtattg acagaatgaa tagcttaagt cgcttgtaa tgaagaagat tttcggggct      60
gaaaaaacct cgtataaacc ggcttccgat accataatcg gaacggatac cctgaaaaga      120
ccgaaccggc ggcctgaacc gacggcaaaa gtcgacaaaa cgatattcaa gactatgggg      180
aatagtctga ataataccct tgtttcagcc tgtgactggt tgatcggaca acaaaagccc      240
gatggtcatt gggtcggtgc cgtggaatcc aatgcttcga tggaagcaga atggtgtctg      300
gccttggtgt ttttgggtct ggaagatcat ccgcttcgtc caagattggg caatgctctt      360
ttggaaatgc agcgggaaga tggtctcttg ggagtctatt tcggcgctgg aaatggcgat      420
atcaatgcca cggttgaagc ctatgcggcc ttgcggtctt tgggggtattc tgccgataat      480
cctgttttga aaaaagcggc agcatggatt gctgaaaaag gcgattaaa aaatatccgt      540
gtctttaccc gttattggct ggcttgatc ggggaatggc cttgggaaaa gaccctaac      600
cttccccctg aaattatctg gttccctgat aattttgtct tttcgattta taattttgcc      660
caatgggctc gggcaaccat ggtgccgatt gctattctgt ccgcgagacg accaagccgc      720
ccgctgcgcc ctcaagaccg attggatgaa ctgtttccag aaggccgcgc tcgctttgat      780
tatgaattgc cgaaaaaaga aggcatcgat ctttggtcgc aatttttccg aaccactgac      840
cgtggattac attgggttca gtccaatctg ttaaagcgca atagcttgcg tgaagccgct      900
atccgtcatg ttttggaatg gattatccgg catcaggatg ccgatggcgg ttgggggtgga      960
attcagccac cttgggtcta tggtttgatg gcgttacatg gtgaaggcta tcagctttat     1020
catccggtga tggccaaggc tttgtcggct ttggatgatc ccggttggcg acatgacaga     1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg     1140
atggcggttaa aagacgcaa ggccgaggat cgttttacgc cggaaatgga taaggccgcc     1200
gattggcttt tggctcgaca ggtcaaagtc aaaggcgatt ggtcaatcaa actgcccgat     1260
gttgaacccg gtggatgggc atttgaatat gccaatgatc gctatcccga taccgatgat     1320
accgccgtcg ctttgatcgc cttttcctct tatcgtgata aggaggagtg gcaaaagaaa     1380
ggcggtgagg acgccattac ccgtgggggtt aattggttga tcgcatgca aagcgaatgt     1440
ggcggttggg gagcctttga taaggataat aacagaagta tcctttccaa aattcctttt     1500
tgtgatttcg gagaatctat tgatccgcct tcagtcgatg taacggcgca tgtttttagag     1560
gcctttggca ccttgggact gtcccgcgat atgccggtca tccaaaaagc gatcgactat     1620
gtccgttccg aacaggaagc cgaaggcgcg tggtttggtc gttggggcgt taattatatc     1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gaccagcct     1740
tacatcacca aggcttgca ttggctggtc gcacatcagc aggaagacgg cggttggggc     1800
gaaagctgct ctccctatat ggagattgat tccattggga agggcccaac cacgccgtcc     1860
cagactgctt gggctttgat ggggttgatc gcggccaatc gtcccgaaga ttatgaagcc     1920
attgccaagg gatgccatta tctgattgat cgccaagagc aggatggtag ctggaaagaa     1980
gaagaattca ccggcaccgg attccccggt tatggcgtgg gtcagacgat caagttggat     2040

```

gatccggcctt tatcgaaacg attgcttcaa ggcgctgaac tgtcacgggc gtttatgctg 2100
 cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
 attgatttga ataattga 2178

<210> 45

<211> 725

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 45

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

225	230	235	240
Pro Leu Arg Pro Gln Asp Arg Leu Asp Glu Leu Phe Pro Glu Gly Arg			
	245	250	255
Ala Arg Phe Asp Tyr Glu Leu Pro Lys Lys Glu Gly Ile Asp Leu Trp			
	260	265	270
Ser Gln Phe Phe Arg Thr Thr Asp Arg Gly Leu His Trp Val Gln Ser			
	275	280	285
Asn Leu Leu Lys Arg Asn Ser Leu Arg Glu Ala Ala Ile Arg His Val			
	290	295	300
Leu Glu Trp Ile Ile Arg His Gln Asp Ala Asp Gly Gly Trp Gly Gly			
305	310	315	320
Ile Gln Pro Pro Trp Val Tyr Gly Leu Met Ala Leu His Gly Glu Gly			
	325	330	335
Tyr Gln Leu Tyr His Pro Val Met Ala Lys Ala Leu Ser Ala Leu Asp			
	340	345	350
Asp Pro Gly Trp Arg His Asp Arg Gly Glu Ser Ser Trp Ile Gln Ala			
	355	360	365
Thr Asn Ser Pro Val Trp Asp Thr Met Leu Ala Leu Met Ala Leu Lys			
	370	375	380
Asp Ala Lys Ala Glu Asp Arg Phe Thr Pro Glu Met Asp Lys Ala Ala			
385	390	395	400
Asp Trp Leu Leu Ala Arg Gln Val Lys Val Lys Gly Asp Trp Ser Ile			
	405	410	415
Lys Leu Pro Asp Val Glu Pro Gly Gly Trp Ala Phe Glu Tyr Ala Asn			
	420	425	430
Asp Arg Tyr Pro Asp Thr Asp Asp Thr Ala Val Ala Leu Ile Ala Leu			
	435	440	445
Ser Ser Tyr Arg Asp Lys Glu Glu Trp Gln Lys Lys Gly Val Glu Asp			
	450	455	460
Ala Ile Thr Arg Gly Val Asn Trp Leu Ile Ala Met Gln Ser Glu Cys			
465	470	475	480
Gly Gly Trp Gly Ala Phe Asp Lys Asp Asn Asn Arg Ser Ile Leu Ser			
	485	490	495
Lys Ile Pro Phe Cys Asp Phe Gly Glu Ser Ile Asp Pro Pro Ser Val			
	500	505	510
Asp Val Thr Ala His Val Leu Glu Ala Phe Gly Thr Leu Gly Leu Ser			
	515	520	525
Arg Asp Met Pro Val Ile Gln Lys Ala Ile Asp Tyr Val Arg Ser Glu			
	530	535	540

atgggtattg	acagaatgaa	tagcttaagt	cgcttgtaa	tgaagaagat	tttcggggct	60
gaaaaaacct	cgtataaac	ggcttcgat	accataatcg	gaacggatac	cctgaaaaga	120
ccgaaccggc	ggcctgaacc	gacggcaaaa	gtcgacaaaa	cgatattcaa	gactatgggg	180
aatagtctga	ataataccct	tgtttcagcc	tgtgactgg	tgatcggaca	acaaaagccc	240

gatggtcatt	gggtcgggtgc	cgtggaatcc	aatgcttcga	tggaagcaga	atgggtgtctg	300
gccttggtgt	ttttgggtct	ggaagatcat	ccgcttcgtc	caagattggg	caatgctctt	360
ttggaaatgc	agcgggaaga	tggctcttgg	ggagtctatt	tcggcgctgg	aaatggcgat	420
atcaatgcca	cgggtgaagc	ctatgcggcc	ttgcggctctt	tgggggtattc	tgccgataat	480
cctgttttga	aaaaagcggc	agcatggatt	gctgaaaaag	gcggattaaa	aaatatccgt	540
gtcntnacc	gttattggct	ggcgttgatc	ggggaatggc	cttgggaaaa	gaccacctaac	600
cttccccctg	aaattatctg	gttccctgat	aattttgtct	tttcgattta	taattttgcc	660
caatgggcgc	gggcaaccat	ggtgccgatt	gctattctgt	ccgcgagacg	accaagccgc	720
ccgctgcgcc	ctcaagaccg	attggatgaa	ctgtttccag	aaggccgcgc	tcgctttgat	780
tatgaattgc	cgaaaaaaga	aggcatcgat	ctttggctgc	aatttttccg	aaccactgac	840
cgtggattac	attgggttca	gtccaatctg	ttaaagcgca	atagcttgcg	tgaagccgct	900
atccgtcatg	ttttggaatg	gattatccgg	catcaggatg	ccgatggcgg	ttgggggtgga	960
attcagccac	cttgggtcta	tggtttgatg	gcgttacatg	gtgaaggcta	tcagctttat	1020
catccggtga	tggccaaggc	tttgtcggct	ttggatgatc	ccggttggcg	acatgacaga	1080
ggcgagtctt	cttggataca	ggccaccaat	agtccggtat	gggatacaat	gttggccttg	1140
atggcgtaa	aagacgcaa	ggccgaggat	cgttttacgc	cggaaatgga	taaggccgcc	1200
gattggcttt	tggctcgaca	ggtcaaagtc	aaaggcgatt	ggtcaatcaa	actgcccgat	1260
gttgaacccg	gtggatgggc	atttgaatat	gccaatgatc	gctatcccga	taccgatgat	1320
accgccgtcg	ctttgatcgc	cctttcctct	tatcgtgata	aggaggagtg	gcaaaagaaa	1380
ggcgttgagg	acgccattac	ccgtgggggt	aattggttga	tcgcatgca	aagcgaatgt	1440
ggcggttggg	gagcctttga	taaggataat	aacagaagta	tcctttccaa	aattcctttt	1500
tgtgatttcg	gagaatctat	tgatccgcct	tcagtcgatg	taacggcgca	tgtttttagag	1560
gcctttggca	ccttgggact	gtcccgcgat	atgccggtca	tccaaaaagc	gatcgactat	1620
gtccgttccg	aacaggaagc	cgaaggcgcg	tggtttggtc	gttggggcgt	taattatatac	1680
tatggcaccg	gtgcggttct	gcctgctttg	gcggcgatcg	gtgaagatat	gaccagcct	1740
tacatcacca	aggcttgcca	ttggctggtc	gcacatcagc	aggaagacgg	cggttggggc	1800
gaaagctgct	cttcctatat	ggagattgat	tccattggga	agggcccaac	cacgccgtcc	1860
cagactgctt	gggctttgat	ggggttgatc	gcggccaatc	gtcccgaaga	ttatgaagcc	1920
attgccaagg	gatgccatta	tctgattgat	cgccaagagc	aggatggtag	ctggaaagaa	1980
gaagaattca	ccggcaccgg	attccccggt	tatggcgtgg	gtcagacgat	caagttggat	2040
gatccggctt	tatcgaaaacg	attgcttcaa	ggcgtgaac	tgtcacgggc	gtttatgctg	2100
cgttatgatt	tttatcggca	attcttcccg	attatggcgt	taagtcgggc	agagagactg	2160
attgatattga	ataattga					2178
<210>	47					
<211>	725					
<212>	PRT					
<213>	人工序列					
<220>						
<223>	合成ZmoSHC1衍生物					

<400> 47

```

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

```

Leu	Glu	Trp	Ile	Ile	Arg	His	Gln	Asp	Ala	Asp	Gly	Gly	Trp	Gly	Gly				
305							310				315				320				
Ile	Gln	Pro	Pro	Trp	Val	Tyr	Gly	Leu	Met	Ala	Leu	His	Gly	Glu	Gly				
				325					330						335				
Tyr	Gln	Leu	Tyr	His	Pro	Val	Met	Ala	Lys	Ala	Leu	Ser	Ala	Leu	Asp				
			340					345						350					
Asp	Pro	Gly	Trp	Arg	His	Asp	Arg	Gly	Glu	Ser	Ser	Trp	Ile	Gln	Ala				
		355					360						365						
Thr	Asn	Ser	Pro	Val	Trp	Asp	Thr	Met	Leu	Ala	Leu	Met	Ala	Leu	Lys				
370						375						380							
Asp	Ala	Lys	Ala	Glu	Asp	Arg	Phe	Thr	Pro	Glu	Met	Asp	Lys	Ala	Ala				
385					390					395					400				
Asp	Trp	Leu	Leu	Ala	Arg	Gln	Val	Lys	Val	Lys	Gly	Asp	Trp	Ser	Ile				
				405					410						415				
Lys	Leu	Pro	Asp	Val	Glu	Pro	Gly	Gly	Trp	Ala	Phe	Glu	Tyr	Ala	Asn				
			420					425						430					
Asp	Arg	Tyr	Pro	Asp	Thr	Asp	Asp	Thr	Ala	Val	Ala	Leu	Ile	Ala	Leu				
		435					440						445						
Ser	Ser	Tyr	Arg	Asp	Lys	Glu	Glu	Trp	Gln	Lys	Lys	Gly	Val	Glu	Asp				
450					455							460							
Ala	Ile	Thr	Arg	Gly	Val	Asn	Trp	Leu	Ile	Ala	Met	Gln	Ser	Glu	Cys				
465					470					475					480				
Gly	Gly	Trp	Gly	Ala	Phe	Asp	Lys	Asp	Asn	Asn	Arg	Ser	Ile	Leu	Ser				
				485					490						495				
Lys	Ile	Pro	Phe	Cys	Asp	Phe	Gly	Glu	Ser	Ile	Asp	Pro	Pro	Ser	Val				
			500					505						510					
Asp	Val	Thr	Ala	His	Val	Leu	Glu	Ala	Phe	Gly	Thr	Leu	Gly	Leu	Ser				
		515						520						525					
Arg	Asp	Met	Pro	Val	Ile	Gln	Lys	Ala	Ile	Asp	Tyr	Val	Arg	Ser	Glu				
530						535					540								
Gln	Glu	Ala	Glu	Gly	Ala	Trp	Phe	Gly	Arg	Trp	Gly	Val	Asn	Tyr	Ile				
545					550					555					560				
Tyr	Gly	Thr	Gly	Ala	Val	Leu	Pro	Ala	Leu	Ala	Ala	Ile	Gly	Glu	Asp				
				565					570					575					
Met	Thr	Gln	Pro	Tyr	Ile	Thr	Lys	Ala	Cys	Asp	Trp	Leu	Val	Ala	His				
			580					585						590					
Gln	Gln	Glu	Asp	Gly	Gly	Trp	Gly	Glu	Ser	Cys	Ser	Ser	Tyr	Met	Glu				
		595					600						605						
Ile	Asp	Ser	Ile	Gly	Lys	Gly	Pro	Thr	Thr	Pro	Ser	Gln	Thr	Ala	Trp				

610	615	620
Ala Leu Met Gly Leu Ile	Ala Ala Asn Arg Pro	Glu Asp Tyr Glu Ala
625	630	635
Ile Ala Lys Gly Cys His	Tyr Leu Ile Asp Arg	Gln Glu Gln Asp Gly
645	650	655
Ser Trp Lys Glu Glu Glu	Phe Thr Gly Thr Gly	Phe Pro Gly Tyr Gly
660	665	670
Val Gly Gln Thr Ile Lys	Leu Asp Asp Pro Ala	Leu Ser Lys Arg Leu
675	680	685
Leu Gln Gly Ala Glu Leu	Ser Arg Ala Phe Met	Leu Arg Tyr Asp Phe
690	695	700
Tyr Arg Gln Phe Phe Pro	Ile Met Ala Leu Ser	Arg Ala Glu Arg Leu
705	710	715
Ile Asp Leu Asn Asn		
725		

<210> 48

<211> 2178

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<220>

<221> misc_feature

<222> (553) .. (555)

<223> ngn是agr或cgn

<400> 48

atgggtattg acagaatgaa tagcttaagt cgcttgtaa tgaagaagat tttcggggct	60
gaaaaaacct cgtataaacc ggcttccgat accataatcg gaacggatac cctgaaaaga	120
ccgaaccggc ggcctgaacc gacggcaaaa gtcgacaaaa cgatattcaa gactatgggg	180
aatagtctga ataataccct tgtttcagcc tgtgactggg tgatcggaca acaaaagccc	240
gatggtcatt gggtcggtgc cgtggaatcc aatgcttcga tggaagcaga atgggtgtctg	300
gccttggtgt ttttgggtct ggaagatcat ccgcttcgtc caagattggg caatgctctt	360
ttggaaatgc agcgggaaga tggtcttgg ggagtctatt tcggcgctgg aaatggcgat	420
atcaatgcca cggttgaagc ctatgcggcc ttgcggtctt tgggggtattc tgccgataat	480
cctgttttga aaaaagcggc agcatggatt gctgaaaaag gcggattaaa aaatatccgt	540
gtcttttacc gtngntggct ggcgttgatc ggggaatggc cttgggaaaa gaccctaac	600
cttccccctg aaattatctg gttccctgat aattttgtct tttcgattta taattttgcc	660
caatgggcgc gggcaaccat ggtgccgatt gctattctgt ccgcgagacg accaagccgc	720
ccgctgcgcc ctcaagaccg attggatgaa ctgtttccag aaggccgcgc tcgctttgat	780

```

tatgaattgc cgaaaaaaga aggcattgat ctttggtcgc aatttttccg aaccactgac      840
cgtggattac attgggttca gtccaatctg ttaaagcgca atagcttgcg tgaagccgct      900
atccgtcatg ttttggaatg gattatccgg catcaggatg ccgatggcgg ttgggggtgga      960
attcagccac cttgggtcta tggtttgatg gcgttacatg gtgaaggcta tcagctttat     1020
catccgggtga tggccaaggc tttgtcggct ttggatgata ccggttggcg acatgacaga     1080
ggcgagtctt cttggataca ggccaccaat agtccgggtat gggatacaat gttggccttg     1140
atggcggttaa aagacgccaa ggccaggatg cgttttacgc cggaaatgga taaggccgcc     1200
gattggcttt tggctcgaca ggtcaaagtc aaaggcgatt ggtcaatcaa actgcccgat     1260
gttgaacccg gtggatgggc atttgaatat gccaatgata gctatcccga taccgatgat     1320
accgccgtcg ctttgatcgc cttttcctct tatcgtgata aggaggagtg gcaaaagaaa     1380
ggcgttgagg acgccattac ccgtgggggtt aattggttga tcgccatgca aagcgaatgt     1440
ggcggttggg gagcctttga taaggataat aacagaagta tcctttccaa aattcctttt     1500
tgtgatttcg gagaatctat tgatccgcct tcagtcgatg taacggcgca tgttttagag     1560
gcctttggca ccttgggact gtcccgcgat atgccggtca tccaaaaagc gatcgactat     1620
gtccgttccg aacaggaagc cgaaggcgcg tggtttggtc gttggggcgt taattatata     1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gaccagcct     1740
tacatcacca aggcttgcca ttggctggtc gcacatcagc aggaagacgg cggttggggc     1800
gaaagctgct cttcctatat ggagattgat tccattggga agggcccaac cacgccgtcc     1860
cagactgctt gggctttgat ggggttgatc gcggccaatc gtcccgaaga ttatgaagcc     1920
attgccaagg gatgccatta tctgattgat cgccaagagc aggatggtag ctggaagaa     1980
gaagaattca ccggcaccgg attccccggt tatggcgtgg gtcagacgat caagttggat     2040
gatccggctt tatcgaaacg attgcttcaa ggcgctgaac tgtcacgggc gtttatgctg     2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg     2160
attgatttga ataattga                                     2178

```

<210> 49

<211> 725

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 49

Met Gly Ile Asp Arg Met Asn Ser Leu Ser Arg Leu Leu Met Lys Lys

1 5 10 15

Ile Phe Gly Ala Glu Lys Thr Ser Tyr Lys Pro Ala Ser Asp Thr Ile

20 25 30

Ile Gly Thr Asp Thr Leu Lys Arg Pro Asn Arg Arg Pro Glu Pro Thr

35 40 45

Ala Lys Val Asp Lys Thr Ile Phe Lys Thr Met Gly Asn Ser Leu Asn

50 55 60

Asn Thr Leu Val Ser Ala Cys Asp Trp Leu Ile Gly Gln Gln Lys Pro			
65	70	75	80
Asp Gly His Trp Val Gly Ala Val Glu Ser Asn Ala Ser Met Glu Ala			
	85	90	95
Glu Trp Cys Leu Ala Leu Trp Phe Leu Gly Leu Glu Asp His Pro Leu			
	100	105	110
Arg Pro Arg Leu Gly Asn Ala Leu Leu Glu Met Gln Arg Glu Asp Gly			
	115	120	125
Ser Trp Gly Val Tyr Phe Gly Ala Gly Asn Gly Asp Ile Asn Ala Thr			
	130	135	140
Val Glu Ala Tyr Ala Ala Leu Arg Ser Leu Gly Tyr Ser Ala Asp Asn			
145	150	155	160
Pro Val Leu Lys Lys Ala Ala Ala Trp Ile Ala Glu Lys Gly Gly Leu			
	165	170	175
Lys Asn Ile Arg Val Phe Thr Arg Tyr Trp Leu Ala Leu Ile Gly Glu			
	180	185	190
Trp Pro Trp Glu Lys Thr Pro Asn Leu Pro Pro Glu Ile Ile Trp Phe			
	195	200	205
Pro Asp Asn Phe Val Phe Ser Ile Tyr Asn Phe Ala Gln Trp Ala Arg			
	210	215	220
Ala Thr Met Val Pro Ile Ala Ile Leu Ser Ala Arg Arg Pro Ser Arg			
225	230	235	240
Pro Leu Arg Pro Gln Asp Arg Leu Asp Glu Leu Phe Pro Glu Gly Arg			
	245	250	255
Ala Arg Phe Asp Tyr Glu Leu Pro Lys Lys Glu Gly Ile Asp Leu Trp			
	260	265	270
Ser Gln Phe Phe Arg Thr Thr Asp Arg Val Leu His Trp Val Gln Ser			
	275	280	285
Asn Leu Leu Lys Arg Asn Ser Leu Arg Glu Ala Ala Ile Arg His Val			
	290	295	300
Leu Glu Trp Ile Ile Arg His Gln Asp Ala Asp Gly Gly Trp Gly Gly			
305	310	315	320
Ile Gln Pro Pro Trp Val Tyr Gly Leu Met Ala Leu His Gly Glu Gly			
	325	330	335
Tyr Gln Leu Tyr His Pro Val Met Ala Lys Ala Leu Ser Ala Leu Asp			
	340	345	350
Asp Pro Gly Trp Arg His Asp Arg Gly Glu Ser Ser Trp Ile Gln Ala			
	355	360	365
Thr Asn Ser Pro Val Trp Asp Thr Met Leu Ala Leu Met Ala Leu Lys			

370	375	380
Asp Ala Lys Ala Glu Asp Arg Phe Thr Pro Glu Met Asp Lys Ala Ala		
385	390	395
Asp Trp Leu Leu Ala Arg Gln Val Lys Val Lys Gly Asp Trp Ser Ile		400
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
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
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
Ile Asp Ser Ile Gly Lys Gly Pro Thr Thr Pro Ser Gln Thr Ala Trp		
610	615	620
Ala Leu Met Gly Leu Ile Ala Ala Asn Arg Pro Glu Asp Tyr Glu Ala		
625	630	635
Ile Ala Lys Gly Cys His Tyr Leu Ile Asp Arg Gln Glu Gln Asp Gly		
645	650	655
Ser Trp Lys Glu Glu Glu Phe Thr Gly Thr Gly Phe Pro Gly Tyr Gly		
660	665	670
Val Gly Gln Thr Ile Lys Leu Asp Asp Pro Ala Leu Ser Lys Arg Leu		
675	680	685

Leu Gln Gly Ala Glu Leu Ser Arg Ala Phe Met Leu Arg Tyr Asp Phe
 690 695 700
 Tyr Arg Gln Phe Phe Pro Ile Met Ala Leu Ser Arg Ala Glu Arg Leu
 705 710 715 720
 Ile Asp Leu Asn Asn
 725

<210> 50

<211> 2178

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<220>

<221> misc_feature

<222> (846) .. (846)

<223> n是a, c, g,或t

<400> 50

```

atgggtattg acagaatgaa tagcttaagt cgcttggttaa tgaagaagat tttcggggct      60
gaaaaaacct cgtataaacc ggcttccgat accataatcg gaacggatac cctgaaaaga      120
ccgaaccggc ggccctgaacc gacggcaaaa gtcgacaaaa cgatattcaa gactatgggg      180
aatagtctga ataataccct tgtttcagcc tgtgactggt tgatcggaca acaaaagccc      240
gatggtcatt gggctcggtc cgtggaatcc aatgcttcga tggaagcaga atggtgtctg      300
gccttggtgt ttttggtgtc ggaagatcat ccgcttcgtc caagattggg caatgctctt      360
ttggaaatgc agcgggaaga tggctcttgg ggagtctatt tcggcgctgg aaatggcgat      420
atcaatgcca cggttgaagc ctatgcggcc ttgcggtctt tgggggtattc tgccgataat      480
cctgttttga aaaaagcggc agcatggatt gctgaaaaag gcgattaaa aaatatccgt      540
gtcttttacc gttattggct ggcgttgatc ggggaatggc cttgggaaaa gaccctaac      600
cttccccctg aaattatctg gttccctgat aattttgtct tttcgattta taattttgcc      660
caatgggcgc gggcaaccat ggtgccgatt gctattctgt ccgcgagacg accaagccgc      720
ccgctgcgcc ctcaagaccg attggatgaa ctgtttccag aaggccgcgc tcgctttgat      780
tatgaattgc cgaaaaaaga aggcacgat ctttggtcgc aatttttccg aaccactgac      840
cgtgtnttac attgggttca gtccaatctg ttaaagcgca atagcttgcg tgaagccgct      900
atccgtcatg ttttggaatg gattatccgg catcaggatg ccgatggcgg ttgggggtgga      960
attcagccac cttgggtcta tggtttgatg gcgttacatg gtgaaggcta tcagctttat     1020
catccggtga tggccaaggc tttgtcggct ttggatgac ccggttggcg acatgacaga     1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg     1140
atggcggttaa aagacgccaa ggccgaggat cgttttacgc cggaaatgga taaggccgcc     1200
gattggcttt tggctcgaca ggtcaaagtc aaaggcgatt ggtcaatcaa actgcccgat     1260
gttgaacccg gtggatgggc atttgaatat gccaatgatc gctatcccga taccgatgat     1320

```



```

accgccgtcg ctttgatcgc cttttcctct tatcgtgata aggaggagtg gcaaaagaaa 1380
ggcgttgagg acgccattac cctgggggtt aattggttga tcgccatgca aagcgaatgt 1440
ggcggttggg gagcctttga taaggataat aacagaagta tcctttccaa aattcctttt 1500
tgtgatttcg gagaatctat tgatccgcct tcagtcgatg taacggcgca tgttttagag 1560
gcctttggca ccttgggact gtcccgcgat atgccgggtca tccaaaaagc gatcgactat 1620
gtccgttccg aacaggaagc cgaaggcgcg tggtttggtc gttggggcgt taattatata 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gaccagcct 1740
tacatcacca aggcttgcca ttggctggtc gcacatcagc aggaagacgg cggttggggc 1800
gaaagctgct cttcctatat ggagattgat tccattggga agggcccaac cacgccgtcc 1860
cagactgctt gggctttgat ggggttgatc gcggccaatc gtcccgaaga ttatgaagcc 1920
attgccaagg gatgccatta tctgattgat cgccaagagc aggatggtag ctggaaagaa 1980
gaagaattca ccggcaccgg attccccggt tatggcgtgg gtcagacgat caagttggat 2040
gatccggtt tatcgaaacg attgcttcaa ggcgtgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

```

<210> 51

<211> 725

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 51

```

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

```

130	135	140
Val Glu Ala Tyr Ala Ala Leu Arg Ser Leu Gly Tyr Ser Ala Asp Asn		
145	150	155
Pro Val Leu Lys Lys Ala Ala Ala Trp Ile Ala Glu Lys Gly Gly Leu		
165	170	175
Lys Asn Ile Arg Val Phe Thr Arg Tyr Trp Leu Ala Leu Ile Gly Glu		
180	185	190
Trp Pro Trp Glu Lys Thr Pro Asn Leu Pro Pro Glu Ile Ile Trp Phe		
195	200	205
Pro Asp Asn Phe Val Phe Ser Ile Tyr Asn Phe Ala Gln Trp Ala Arg		
210	215	220
Ala Thr Met Val Pro Ile Ala Ile Leu Ser Ala Arg Arg Pro Ser Arg		
225	230	235
Pro Leu Arg Pro Gln Asp Arg Leu Asp Glu Leu Phe Pro Glu Gly Arg		
245	250	255
Ala Arg Phe Asp Tyr Glu Leu Pro Lys Lys Glu Gly Ile Asp Leu Trp		
260	265	270
Ser Gln Phe Phe Arg Thr Thr Asp Arg Gly Leu His Trp Val Gln Ser		
275	280	285
Asn Leu Leu Lys Arg Asn Ser Leu Arg Glu Ala Ala Ile Arg His Val		
290	295	300
Leu Glu Trp Ile Ile Arg His Gln Asp Ala Asp Gly Gly Trp Gly Gly		
305	310	315
Ile Gln Pro Pro Trp Val Tyr Gly Leu Met Ala Leu His Gly Glu Gly		
325	330	335
Tyr Gln Leu Tyr His Pro Val Met Ala Lys Ala Leu Ser Ala Leu Asp		
340	345	350
Asp Pro Gly Trp Arg His Asp Arg Gly Glu Ser Ser Trp Ile Gln Ala		
355	360	365
Thr Asn Ser Pro Val Trp Asp Thr Met Leu Ala Leu Met Ala Leu Lys		
370	375	380
Asp Ala Lys Ala Glu Asp Arg Phe Thr Pro Glu Met Asp Lys Ala Ala		
385	390	395
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
Gly Gly Trp Gly Ala Phe Asp Lys Asp Asn Asn Arg Ser Ile Leu Ser		
	485	490
Lys Thr Pro Phe Cys Asp Phe Gly Glu Ser Ile Asp Pro Pro Ser Val		
	500	505
Asp Val Thr Ala His Val Leu Glu Ala Phe Gly Thr Leu Gly Leu Ser		
	515	520
Arg Asp Met Pro Val Ile Gln Lys Ala Ile Asp Tyr Val Arg Ser Glu		
	530	535
Gln Glu Ala Glu Gly Ala Trp Phe Gly Arg Trp Gly Val Asn Tyr Ile		
545	550	555
Tyr Gly Thr Gly Ala Val Leu Pro Ala Leu Ala Ala Ile Gly Glu Asp		
	565	570
Met Thr Gln Pro Tyr Ile Thr Lys Ala Cys Asp Trp Leu Val Ala His		
	580	585
Gln Gln Glu Asp Gly Gly Trp Gly Glu Ser Cys Ser Ser Tyr Met Glu		
	595	600
Ile Asp Ser Ile Gly Lys Gly Pro Thr Thr Pro Ser Gln Thr Ala Trp		
610	615	620
Ala Leu Met Gly Leu Ile Ala Ala Asn Arg Pro Glu Asp Tyr Glu Ala		
625	630	635
Ile Ala Lys Gly Cys His Tyr Leu Ile Asp Arg Gln Glu Gln Asp Gly		
	645	650
Ser Trp Lys Glu Glu Glu Phe Thr Gly Thr Gly Phe Pro Gly Tyr Gly		
	660	665
Val Gly Gln Thr Ile Lys Leu Asp Asp Pro Ala Leu Ser Lys Arg Leu		
	675	680
Leu Gln Gly Ala Glu Leu Ser Arg Ala Phe Met Leu Arg Tyr Asp Phe		
	690	695
Tyr Arg Gln Phe Phe Pro Ile Met Ala Leu Ser Arg Ala Glu Arg Leu		
705	710	715
Ile Asp Leu Asn Asn		
	725	

<210> 52

<211> 2178

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<220>

<221> misc_feature

<222> (1494) .. (1494)

<223> n是a, c, g,或t

<400> 52

atgggtattg	acagaatgaa	tagcttaagt	cgcttggttaa	tgaagaagat	tttcggggcct	60
gaaaaaacct	cgtataaacc	ggcttccgat	accataatcg	gaacggatac	cctgaaaaga	120
ccgaaccggc	ggcctgaacc	gacggcaaaa	gtcgacaaaa	cgatattcaa	gactatgggg	180
aatagtctga	ataataccct	tgtttcagcc	tgtgactggt	tgatcggaca	acaaaagccc	240
gatggtcatt	gggtcgggtc	cgtggaatcc	aatgcttcga	tggaaagcaga	atggtgtctg	300
gccttgtggt	ttttgggtct	ggaagatcat	ccgcttcgtc	caagattggg	caatgctctt	360
ttggaaatgc	agcgggaaga	tggctcttgg	ggagtctatt	tcggcgctgg	aaatggcgat	420
atcaatgcca	cggttgaagc	ctatgcggcc	ttgcggtctt	tggggtattc	tgccgataat	480
cctgttttga	aaaaagcggc	agcatggatt	gctgaaaaag	gcggattaaa	aaatatccgt	540
gtctttaccc	gttattggct	ggcgttgatc	ggggaatggc	cttgggaaaa	gaccctaac	600
cttccccctg	aaattatctg	gttccctgat	aattttgtct	tttcgattta	taattttgcc	660
caatgggcgc	gggcaaccat	ggtgccgatt	gctattctgt	ccgcgagacg	accaagccgc	720
ccgctgcgcc	ctcaagaccg	attggatgaa	ctgtttccag	aaggccgcgc	tcgctttgat	780
tatgaattgc	cgaaaaaaga	aggcatcgat	ctttggtcgc	aatttttccg	aaccactgac	840
cgtggattac	attgggttca	gtccaatctg	ttaaagcgca	atagcttgcg	tgaagccgct	900
atccgtcatg	ttttggaatg	gattatccgg	catcaggatg	ccgatggcgg	ttggggtgga	960
attcagccac	cttgggtcta	tggtttgatg	gcgttacatg	gtgaaggcta	tcagctttat	1020
catccggtga	tggccaaggc	ttgtcggct	ttggatgatc	ccggttggcg	acatgacaga	1080
ggcgagtctt	cttggataca	ggccaccaat	agtcgggtat	gggatacaat	gttggccttg	1140
atggcggttaa	aagacgccaa	ggccgaggat	cgttttacgc	cggaaatgga	taaggccgcc	1200
gattggcttt	tggctcgaca	ggtcaaagtc	aaaggcgatt	ggtcaatcaa	actgcccgat	1260
gttgaaccgc	gtggatgggc	atttgaatat	gccaatgatc	gctatcccga	taccgatgat	1320
accgccgtcg	ctttgatcgc	cctttcctct	tatcgtgata	aggaggagtg	gcaaaagaaa	1380
ggcgttgagg	acgccattac	ccgtgggggt	aattggttga	tcgcatgca	aagcgaatgt	1440
ggcggttggg	gagcctttga	taaggataat	aacagaagta	tcctttccaa	aacncctttt	1500
tgtgatttcg	gagaatctat	tgatccgcct	tcagtcgatg	taacggcgca	tgttttagag	1560
gcctttggca	ccttgggact	gtcccgcgat	atgccggtca	tccaaaaagc	gatcgactat	1620
gtccgttccg	aacaggaagc	cgaaggcgcg	tggtttggtc	gttggggcgt	taattatatc	1680
tatggcaccg	gtgcggttct	gcctgctttg	gcggcgatcg	gtgaagatat	gaccagcct	1740
tacatcacca	aggcttgcca	ttggctggtc	gcacatcagc	aggaagacgg	cggttggggc	1800
gaaagctgct	cttcctatat	ggagattgat	tccattggga	agggcccaac	cacgccgtcc	1860

cagactgctt gggctttgat ggggttgatc gcggccaatc gtcccgaaga ttatgaagcc 1920
 attgccaagg gatgccatta tctgattgat cgccaagagc aggatggtag ctggaaagaa 1980
 gaagaattca cccgcaccgg attccccggt tatggcgtgg gtcagacgat caagttggat 2040
 gatccggctt tatcgaaacg attgcttcaa ggcgctgaac tgtcacgggc gtttatgctg 2100
 cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
 attgatttga ataattga 2178

<210> 53

<211> 725

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 53

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

Pro Asp Asn Phe Val Phe Ser Ile Tyr Asn Phe Ala Gln Trp Ala Arg			
210	215	220	
Ala Thr Met Val Pro Ile Ala Ile Leu Ser Ala Arg Arg Pro Ser Arg			
225	230	235	240
Pro Leu Arg Pro Gln Asp Arg Leu Asp Glu Leu Phe Pro Glu Gly Arg			
	245	250	255
Ala Arg Phe Asp Tyr Glu Leu Pro Lys Lys Glu Gly Ile Asp Leu Trp			
	260	265	270
Ser Gln Phe Phe Arg Thr Thr Asp Arg Gly Leu His Trp Val Gln Ser			
	275	280	285
Asn Leu Leu Lys Arg Asn Ser Leu Arg Glu Ala Ala Ile Arg His Val			
290	295	300	
Leu Glu Trp Ile Ile Arg His Gln Asp Ala Asp Gly Gly Trp Gly Gly			
305	310	315	320
Ile Gln Pro Pro Trp Val Tyr Gly Leu Met Ala Leu His Gly Glu Gly			
	325	330	335
Tyr Gln Leu Tyr His Pro Val Met Ala Lys Ala Leu Ser Ala Leu Asp			
	340	345	350
Asp Pro Gly Trp Arg His Asp Arg Gly Glu Ser Ser Trp Ile Gln Ala			
	355	360	365
Thr Asn Ser Pro Val Trp Asp Thr Met Leu Ala Leu Met Ala Leu Lys			
370	375	380	
Asp Ala Lys Ala Glu Asp Arg Phe Thr Pro Glu Met Asp Lys Ala Ala			
385	390	395	400
Asp Trp Leu Leu Ala Arg Gln Val Lys Val Lys Gly Asp Trp Ser Ile			
	405	410	415
Lys Leu Pro Asp Val Glu Pro Gly Gly Trp Ala Phe Glu Tyr Ala Asn			
	420	425	430
Asp Arg Tyr Pro Asp Thr Asp Asp Thr Ala Val Ala Leu Ile Ala Leu			
	435	440	445
Ser Ser Tyr Arg Asp Lys Glu Glu Trp Gln Lys Lys Gly Val Glu Asp			
450	455	460	
Ala Ile Thr Arg Gly Val Asn Trp Leu Ile Ala Met Gln Ser Glu Cys			
465	470	475	480
Gly Gly Trp Gly Ala Phe Asp Lys Asp Asn Asn Arg Ser Ile Leu Ser			
	485	490	495
Lys Ile Pro Phe Cys Asp Phe Gly Glu Ser Ile Asp Pro Pro Ser Val			
	500	505	510
Asp Val Thr Ala His Val Leu Glu Ala Phe Gly Thr Leu Gly Leu Ser			

515	520	525
Arg Asp Met Pro Val Ile Gln Lys Ala Ile Asp Tyr Val Arg Ser Glu		
530	535	540
Gln Glu Ala Glu Gly Ala Trp Phe Gly Arg Trp Gly Val Asn Tyr Ile		
545	550	555
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
Ile Asp Ser Ile Gly Lys Gly Pro Thr Thr Pro Ser Gln Thr Ala Trp		
610	615	620
Ala Leu Met Gly Leu Ile Ala Ala Asn Arg Pro Glu Asp Tyr Glu Ala		
625	630	635
Ile Ala Lys Gly Cys His Tyr Leu Ile Asp Arg Gln Glu Gln Asp Gly		
645	650	655
Ser Trp Lys Glu Glu Glu Phe Thr Gly Thr Gly Phe Pro Gly Tyr Gly		
660	665	670
Val Gly Gln Thr Ile Lys Leu Asp Asp Pro Ala Leu Ser Lys Arg Leu		
675	680	685
Leu Gln Gly Ala Glu Leu Ser Arg Ala Phe Met Leu Arg Tyr Asp Phe		
690	695	700
Tyr Arg Gln Phe Phe Pro Ile Met Ala Leu Ser Arg Ala Glu Arg Leu		
705	710	715
Ile Asp Leu Asn Asn		
725		

<210> 54

<211> 2178

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 54

atgggtattg acagaatgaa tagcttaagt cgcttggttaa tgaagaagat tttcggggct	60
gaaaaaacct cgtataaacc ggcttccgat accataatcg gaacggatac cctgaaaaga	120
ccgaaccggc ggccctgaacc gacggcaaaa gtcgacaaaa cgatattcaa gactatgggg	180
aatagtctga ataataccct tgtttcagcc tgtgactggt tgatcggaca acaaaagccc	240
gatggtcatt gggtcggtgc cgtggaatcc aatgcttcga tggaagcaga atggtgtctg	300

gccttgtggt ttttgggtct ggaagatcat ccgcttcgtc caagattggg caatgctctt	360
ttggaaatgc agcgggaaga tggctcttgg ggagtctatt tcggcgctgg aaatggcgat	420
atcaatgcca cggttgaagc ctatgcggcc ttgcggtctt tggggatttc tgccgataat	480
cctgttttga aaaaagcggc agcatggatt gctgaaaaag gcggattaaa aaatatccgt	540
gtctttaccc gttattggct ggcgttgatc ggggaatggc cttgggaaaa gaccctaac	600
cttccccctg aaattatctg gttccctgat aattttgtct tttcgattta taattttgcc	660
caatgggcgc gggcaaccat ggtgccgatt gctattctgt ccgcgagacg accaagccgc	720
ccgctgcgcc ctcaagaccg attggatgaa ctgtttccag aaggccgcgc tcgctttgat	780
tatgaattgc cgaaaaaaga aggcacgat ctttggctgc aatttttccg aaccactgac	840
cgtggattac attgggttca gtccaatctg ttaaagcgca atagcttgcg tgaagccgct	900
atccgtcatg ttttggaatg gattatccgg catcaggatg ccgatggcgg ttgggggtgga	960
attcagccac cttgggtcta tggtttgatg gcgttacatg gtgaaggcta tcagctttat	1020
catccggtga tggccaaggc tttgtcggct ttggatgac ccggttggcg acatgacaga	1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg	1140
atggcggttaa aagacgcaa ggccgaggat cgttttacgc cggaaatgga taaggccgcc	1200
gattggcttt tggctcgaca ggtcaaagtc aaaggcgatt ggtcaatcaa actgcccgat	1260
gttgaacccg gtggatgggc atttgaatat gccaatgatc gctatcccga taccgatgat	1320
accgccgtcg ctttgatcgc cttttcctct tatcgtgata aggaggagtg gcaaaagaaa	1380
ggcgttgagg acgccattac ccgtggggtt aattggttga tcgccatgca aagcgaatgt	1440
ggcggttggg gagcctttga taaggataat aacagaagta tcctttccaa aattcctttt	1500
tgtgatttcg gagaatctat tgatccgcct tcagtcgatg taacggcgca tgttttagag	1560
gcctttggca ccttgggact gtcccgcgat atgccggtca tccaaaaagc gatcgactat	1620
gtccgttccg aacaggaagc cgaaggcgcg tggtttggtc gttggggcgt taattatate	1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gaccagcct	1740
tacatcacca aggcttgcga ttggctggtc gcacatcagc aggaagacgg cggttggggc	1800
gaaagctgct cttcctatat ggagattgat tccattggga agggcccaac cacgccgtcc	1860
cagactgctt gggctttgat ggggttgatc gcggccaatc gtcccgaaga ttatgaagcc	1920
attgccaagg gatgccatta tctgattgat cgccaagagc aggatggtag ctggaaagaa	1980
gaagaattca ccggcaccgg attccccggt tatggcgtgg gtcagacgat caagttggat	2040
gatccggctt tatcgaaacg attgcttcaa ggcgtgaac tgtcacgggc gtttatgctg	2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg	2160
attgatttga ataattga	2178
<210> 55	
<211> 725	
<212> PRT	
<213> 人工序列	
<220>	
<223> 合成ZmoSHC1衍生物	
<400> 55	

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

305	310	315	320
Ile Gln Pro Pro Trp	Val Tyr Gly Leu Met	Ala Leu His Gly	Glu Gly
325	330	335	
Tyr Gln Leu Tyr His	Pro Val Met Ala Lys	Ala Leu Ser Ala	Leu Asp
340	345	350	
Asp Pro Gly Trp Arg	His Asp Arg Gly Glu	Ser Ser Trp Ile	Gln Ala
355	360	365	
Thr Asn Ser Pro Val	Trp Asp Thr Met Leu	Ala Leu Met Ala	Leu Lys
370	375	380	
Asp Ala Lys Ala Glu	Asp Arg Phe Thr Pro	Glu Met Asp Lys	Ala Ala
385	390	395	400
Asp Trp Leu Leu Ala	Arg Gln Val Lys Val	Lys Gly Asp Trp	Ser Ile
405	410	415	
Lys Leu Pro Asp Val	Glu Pro Gly Gly Trp	Ala Phe Glu Tyr	Ala Asn
420	425	430	
Asp Arg Tyr Pro Asp	Thr Asp Asp Thr Ala	Val Ala Leu Ile	Ala Leu
435	440	445	
Ser Ser Tyr Arg Asp	Lys Glu Glu Trp Gln	Lys Lys Gly Val	Glu Asp
450	455	460	
Ala Ile Thr Arg Gly	Val Asn Trp Leu Ile	Ala Met Gln Ser	Glu Cys
465	470	475	480
Gly Gly Trp Gly Ala	Phe Asp Lys Asp Asn	Asn Arg Ser Ile	Leu Ser
485	490	495	
Lys Ile Pro Phe Cys	Asp Phe Gly Glu Ser	Ile Asp Pro Pro	Ser Val
500	505	510	
Asp Val Thr Ala His	Val Leu Glu Ala Phe	Gly Thr Leu Gly	Leu Ser
515	520	525	
Arg Asp Met Pro Val	Ile Gln Lys Ala Ile	Asp Tyr Val Arg	Ser Glu
530	535	540	
Gln Glu Ala Glu Gly	Ala Trp Phe Gly Arg	Trp Gly Val Asn	Tyr Ile
545	550	555	560
Tyr Gly Thr Gly Ala	Val Leu Pro Ala Leu	Ala Ala Ile Gly	Glu Asp
565	570	575	
Met Thr Gln Pro Tyr	Ile Thr Lys Ala Cys	Asp Trp Leu Val	Ala His
580	585	590	
Gln Gln Glu Asp Gly	Gly Trp Gly Glu Ser	Cys Ser Ser Tyr	Met Glu
595	600	605	
Ile Asp Ser Ile Gly	Lys Gly Pro Thr Thr	Pro Ser Gln Thr	Ala Trp
610	615	620	

Ala Leu Met Gly Leu Ile Ala Ala Asn Arg Pro Glu Asp Tyr Glu Ala
 625 630 635 640
 Ile Ala Lys Gly Cys His Tyr Leu Ile Asp Arg Gln Glu Gln Asp Gly
 645 650 655
 Ser Trp Lys Glu Glu Glu Phe Thr Gly Thr Gly Tyr Pro Gly Tyr Gly
 660 665 670
 Val Gly Gln Thr Ile Lys Leu Asp Asp Pro Ala Leu Ser Lys Arg Leu
 675 680 685
 Leu Gln Gly Ala Glu Leu Ser Arg Ala Phe Met Leu Arg Tyr Asp Phe
 690 695 700
 Tyr Arg Gln Phe Phe Pro Ile Met Ala Leu Ser Arg Ala Glu Arg Leu
 705 710 715 720
 Ile Asp Leu Asn Asn
 725

<210> 56

<211> 2178

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 56

```

atgggtattg acagaatgaa tagcttaagt cgcttgtaa tgaagaagat tttcggggct      60
gaaaaaacct cgtataaacc ggcttccgat accataatcg gaacggatac cctgaaaaga      120
ccgaaccggc ggcctgaacc gacggcaaaa gtcgacaaaa cgatattcaa gactatgggg      180
aatagtctga ataataccct tgtttcagcc tgtgactggt tgatcggaca acaaaagccc      240
gatggtcatt gggtcggtgc cgtggaatcc aatgcttcca tggaagcaga atggtgtctg      300
gccttgtggt ttttgggtct ggaagatcat ccgcttcgtc caagattggg caatgctctt      360
ttggaaatgc agcgggaaga tggctcttgg ggagtctatt tcggcgctgg aaatggcgat      420
atcaatgcca cggttgaagc ctatgcggcc ttgcggtctt tggggatttc tgccgataat      480
cctgttttga aaaaagcggc agcatggatt gctgaaaaag gcggattaaa aaatatccgt      540
gtcttttacc gttattggct ggcgttgatc ggggaatggc cttgggaaaa gaccctaac      600
cttccccctg aaattatctg gttccctgat aattttgtct tttcgattta taattttgcc      660
caatgggcgc gggcaaccat ggtgccgatt gctattctgt ccgcgagacg accaagccgc      720
ccgctgcgcc ctcaagaccg attggatgaa ctgtttccag aaggccgcgc tcgctttgat      780
tatgaattgc cgaaaaaaga aggcatcgat ctttggtcgc aatttttccg aaccactgac      840
cgtggattac attgggttca gtccaatctg ttaaagcgca atagcttgcg tgaagccgct      900
atccgtcatg ttttggaatg gattatccgg catcaggatg ccgatggcgg ttgggggtga      960
attcagccac cttgggtcta tggtttgatg gcgttacatg gtgaaggcta tcagctttat     1020
catccggtga tggccaaggc tttgtcggct ttggatgatc ccggttggcg acatgacaga     1080

```

```

ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcggttaa aagacgccaa ggccgaggat cgttttacgc cggaaatgga taaggccgcc 1200
gattggcttt tggctcgaca ggtcaaagtc aaaggcgatt ggtcaatcaa actgcccgat 1260
gttgaacccg gtggatgggc atttgaatat gccaatgata gctatcccga taccgatgat 1320
accgccgtcg ctttgatcgc cttttcctct tatcgtgata aggaggagtg gcaaaagaaa 1380
ggcggtgagg acgccattac ccgtgggggtt aattggttga tcgccatgca aagcgaatgt 1440
ggcggttggg gagcctttga taaggataat aacagaagta tcctttccaa aattcctttt 1500
tgtgatttcg gagaatctat tgatccgcct tcagtcgatg taacggcgca tgtttttagag 1560
gccttttgga ccttgggact gtcccgcgat atgccggtca tccaaaaagc gatcgactat 1620
gtccgttccg aacaggaagc cgaaggcgcg tggtttggtc gttggggcgt taattatata 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gaccagcct 1740
tacatcacca aggcttgcca ttggttggtc gcacatcagc aggaagacgg cggttggggc 1800
gaaagctgct cttcctatat ggagattgat tccattggga agggcccaac cacgccgtcc 1860
cagactgctt gggctttgat ggggttgatc gcggccaatc gtcccgaaga ttatgaagcc 1920
attgccaagg gatgccatta tctgattgat cgccaagagc aggatggtag ctggaaagaa 1980
gaagaattca ccggcaccgg ataycccggg tatggcgtgg gtcagacgat caagttggat 2040
gatccggctt tatcgaacg attgcttcaa ggcgtgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

```

<210> 57

<211> 725

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 57

```

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

```

100	105	110
Arg Pro Arg Leu Gly Asn Ala Leu Leu Glu Met Gln Arg Glu Asp Gly		
115	120	125
Ser Trp Gly Val Tyr Phe Gly Ala Gly Asn Gly Asp Ile Asn Ala Thr		
130	135	140
Val Glu Ala Tyr Ala Ala Leu Arg Ser Leu Gly Tyr Ser Ala Asp Asn		
145	150	155
Pro Val Leu Lys Lys Ala Ala Ala Trp Ile Ala Glu Lys Gly Gly Leu		
165	170	175
Lys Asn Ile Arg Val Phe Thr Arg Arg Trp Leu Ala Leu Ile Gly Glu		
180	185	190
Trp Pro Trp Glu Lys Thr Pro Asn Leu Pro Pro Glu Ile Ile Trp Phe		
195	200	205
Pro Asp Asn Phe Val Phe Ser Ile Tyr Asn Phe Ala Gln Trp Ala Arg		
210	215	220
Ala Thr Met Val Pro Ile Ala Ile Leu Ser Ala Arg Arg Pro Ser Arg		
225	230	235
Pro Leu Arg Pro Gln Asp Arg Leu Asp Glu Leu Phe Pro Glu Gly Arg		
245	250	255
Ala Arg Phe Asp Tyr Glu Leu Pro Lys Lys Glu Gly Ile Asp Leu Trp		
260	265	270
Ser Gln Phe Phe Arg Thr Thr Asp Arg Val Leu His Trp Val Gln Ser		
275	280	285
Asn Leu Leu Lys Arg Asn Ser Leu Arg Glu Ala Ala Ile Arg His Val		
290	295	300
Leu Glu Trp Ile Ile Arg His Gln Asp Ala Asp Gly Gly Trp Gly Gly		
305	310	315
Ile Gln Pro Pro Trp Val Tyr Gly Leu Met Ala Leu His Gly Glu Gly		
325	330	335
Tyr Gln Leu Tyr His Pro Val Met Ala Lys Ala Leu Ser Ala Leu Asp		
340	345	350
Asp Pro Gly Trp Arg His Asp Arg Gly Glu Ser Ser Trp Ile Gln Ala		
355	360	365
Thr Asn Ser Pro Val Trp Asp Thr Met Leu Ala Leu Met Ala Leu Lys		
370	375	380
Asp Ala Lys Ala Glu Asp Arg Phe Thr Pro Glu Met Asp Lys Ala Ala		
385	390	395
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
Asp	Arg	Tyr	Pro	Asp	Thr	Asp	Asp	Thr	Ala	Val	Ala	Leu	Ile	Ala	Leu
Ser	Ser	Tyr	Arg	Asp	Lys	Glu	Glu	Trp	Gln	Lys	Lys	Gly	Val	Glu	Asp
Ala	Ile	Thr	Arg	Gly	Val	Asn	Trp	Leu	Ile	Ala	Met	Gln	Ser	Glu	Cys
Gly	Gly	Trp	Gly	Ala	Phe	Asp	Lys	Asp	Asn	Asn	Arg	Ser	Ile	Leu	Ser
Lys	Thr	Pro	Phe	Cys	Asp	Phe	Gly	Glu	Ser	Ile	Asp	Pro	Pro	Ser	Val
Asp	Val	Thr	Ala	His	Val	Leu	Glu	Ala	Phe	Gly	Thr	Leu	Gly	Leu	Ser
Arg	Asp	Met	Pro	Val	Ile	Gln	Lys	Ala	Ile	Asp	Tyr	Val	Arg	Ser	Glu
Gln	Glu	Ala	Glu	Gly	Ala	Trp	Phe	Gly	Arg	Trp	Gly	Val	Asn	Tyr	Ile
Tyr	Gly	Thr	Gly	Ala	Val	Leu	Pro	Ala	Leu	Ala	Ala	Ile	Gly	Glu	Asp
Met	Thr	Gln	Pro	Tyr	Ile	Thr	Lys	Ala	Cys	Asp	Trp	Leu	Val	Ala	His
Gln	Gln	Glu	Asp	Gly	Gly	Trp	Gly	Glu	Ser	Cys	Ser	Ser	Tyr	Met	Glu
Ile	Asp	Ser	Ile	Gly	Lys	Gly	Pro	Thr	Thr	Pro	Ser	Gln	Thr	Ala	Trp
Ala	Leu	Met	Gly	Leu	Ile	Ala	Ala	Asn	Arg	Pro	Glu	Asp	Tyr	Glu	Ala
Ile	Ala	Lys	Gly	Cys	His	Tyr	Leu	Ile	Asp	Arg	Gln	Glu	Gln	Asp	Gly
Ser	Trp	Lys	Glu	Glu	Glu	Phe	Thr	Gly	Thr	Gly	Phe	Pro	Gly	Tyr	Gly
Val	Gly	Gln	Thr	Ile	Lys	Leu	Asp	Asp	Pro	Ala	Leu	Ser	Lys	Arg	Leu
Leu	Gln	Gly	Ala	Glu	Leu	Ser	Arg	Ala	Phe	Met	Leu	Arg	Tyr	Asp	Phe
Tyr	Arg	Gln	Phe	Phe	Pro	Ile	Met	Ala	Leu	Ser	Arg	Ala	Glu	Arg	Leu
Ile	Asp	Leu	Asn	Asn											

725

<210>	58	
<211>	2178	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成ZmoSHC1衍生物	
<220>		
<221>	misc_feature	
<222>	(553) .. (555)	
<223>	ngn是agr或cgn	
<220>		
<221>	misc_feature	
<222>	(846) .. (846)	
<223>	n是a, c, g,或t	
<220>		
<221>	misc_feature	
<222>	(1494) .. (1494)	
<223>	n是a, c, g,或t	
<400>	58	
atgggtattg acagaatgaa tagcttaagt cgcttggttaa tgaagaagat tttcggggct	60	
gaaaaaacct cgtataaacc ggcttccgat accataatcg gaacggatac cctgaaaaga	120	
ccgaaccggc ggccatgaacc gacggcaaaa gtgcacaaaa cgatattcaa gactatgggg	180	
aatagtctga ataataccct tgtttcagcc tgtgactggt tgatcggaca acaaaagccc	240	
gatggtcatt gggtcggtgc cgtggaatcc aatgcttcga tggaagcaga atggtgtctg	300	
gccttgtggt ttttgggtct ggaagatcat ccgcttcgtc caagattggg caatgctctt	360	
ttggaaatgc agcgggaaga tggctcttgg ggagtctatt tcggcgctgg aaatggcgat	420	
atcaatgcca cggttgaagc ctatgcggcc ttgcggtctt tggggtattc tgccgataat	480	
cctgttttga aaaaagcggc agcatggatt gctgaaaaag gcggattaaa aaatatccgt	540	
gtcttttacc gtngntggct ggcttgcgc ggggaatggc cttgggaaaa gaccctaac	600	
cttccccctg aaattatctg gttccctgat aattttgtct tttcgattta taattttgcc	660	
caatgggcgc gggcaaccat ggtgccgatt gctattctgt ccgcgagacg accaagccgc	720	
ccgctgcgcc ctcaagaccg attggatgaa ctgtttccag aaggccgcgc tcgctttgat	780	
tatgaattgc cgaaaaaaga aggcacgat ctttggtcgc aatttttccg aaccactgac	840	
cgtgtnttac attgggttca gtccaatctg ttaaagcgca atagcttgcg tgaagccgct	900	
atccgtcatg ttttggaatg gattatccgg catcaggatg ccgatggcgg ttgggggtgga	960	
attcagccac cttgggtcta tggtttgatg gcgttacatg gtgaaggcta tcagctttat	1020	
catccggtga tggccaaggc tttgtcggct ttggatgatc ccggttggcg acatgacaga	1080	
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg	1140	

```

atggcggttaa aagacgccaa ggccgaggat cgttttacgc cggaaatgga taaggccgcc 1200
gattggctttt tggctcgaca ggtcaaagtc aaaggcgatt ggtcaatcaa actgccccgat 1260
gttgaaccccg gtggatgggc atttgaatat gccaatgata gctatcccga taccgatgat 1320
accgccgtcg ctttgatcgc cttttcctct tatcgtgata aggaggagtg gcaaaagaaa 1380
ggcgttgagg acgccattac ccgtgggggtt aattgggtga tcgccatgca aagcgaatgt 1440
ggcggttgagg gagcctttga taaggataat aacagaagta tcctttccaa aacncctttt 1500
tgtgatttcg gagaatctat tgatccgcct tcagtcgatg taacggcgca tgtttttagag 1560
gcctttggca ccttgggact gtcccgcgat atgccgggtca tccaaaaagc gatcgactat 1620
gtccgttccg aacaggaagc cgaaggcgcg tggtttggtc gttggggcgt taattatatc 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gaccagcct 1740
tacatcacca aggcttgcca ttggctggtc gcacatcagc aggaagacgg cggttggggc 1800
gaaagctgct cttcctatat ggagattgat tccattggga agggcccaac cacgccgtcc 1860
cagactgctt gggctttgat ggggttgatc gcggccaatc gtcccgaaga ttatgaagcc 1920
attgccaagg gatgccatta tctgattgat cgccaagagc aggatggtag ctggaaagaa 1980
gaagaattca ccggcaccgg attccccggt tatggcgtgg gtcagacgat caagttggat 2040
gatccggctt tatcgaaacg attgcttcaa ggcgtgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcgga attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

```

<210> 59

<211> 725

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 59

Met Gly Ile Asp Arg Met Asn Ser Leu Ser Arg Leu Leu Met Lys Lys

1 5 10 15

Ile Phe Gly Ala Glu Lys Thr Ser Tyr Lys Pro Ala Ser Asp Thr Ile

20 25 30

Ile Gly Thr Asp Thr Leu Lys Arg Pro Asn Arg Arg Pro Glu Pro Thr

35 40 45

Ala Lys Val Asp Lys Thr Ile Phe Lys Thr Met Gly Asn Ser Leu Asn

50 55 60

Asn Thr Leu Val Ser Ala Cys Asp Trp Leu Ile Gly Gln Gln Lys Pro

65 70 75 80

Asp Gly His Trp Val Gly Ala Val Glu Ser Asn Ala Ser Met Glu Ala

85 90 95

Glu Trp Cys Leu Ala Leu Trp Phe Leu Gly Leu Glu Asp His Pro Leu

100 105 110

Arg	Pro	Arg	Leu	Gly	Asn	Ala	Leu	Leu	Glu	Met	Gln	Arg	Glu	Asp	Gly
115				120				125							
Ser	Trp	Gly	Val	Tyr	Phe	Gly	Ala	Gly	Asn	Gly	Asp	Ile	Asn	Ala	Thr
130				135				140							
Val	Glu	Ala	Tyr	Ala	Ala	Leu	Arg	Ser	Leu	Gly	Tyr	Ser	Ala	Asp	Asn
145				150				155				160			
Pro	Val	Leu	Lys	Lys	Ala	Ala	Ala	Trp	Ile	Ala	Glu	Lys	Gly	Gly	Leu
165				170				175							
Lys	Asn	Ile	Arg	Val	Phe	Thr	Arg	Arg	Trp	Leu	Ala	Leu	Ile	Gly	Glu
180				185				190							
Trp	Pro	Trp	Glu	Lys	Thr	Pro	Asn	Leu	Pro	Pro	Glu	Ile	Ile	Trp	Phe
195				200				205							
Pro	Asp	Asn	Phe	Val	Phe	Ser	Ile	Tyr	Asn	Phe	Ala	Gln	Trp	Ala	Arg
210				215				220							
Ala	Thr	Met	Val	Pro	Ile	Ala	Ile	Leu	Ser	Ala	Arg	Arg	Pro	Ser	Arg
225				230				235				240			
Pro	Leu	Arg	Pro	Gln	Asp	Arg	Leu	Asp	Glu	Leu	Phe	Pro	Glu	Gly	Arg
245				250				255							
Ala	Arg	Phe	Asp	Tyr	Glu	Leu	Pro	Lys	Lys	Glu	Gly	Ile	Asp	Leu	Trp
260				265				270							
Ser	Gln	Phe	Phe	Arg	Thr	Thr	Asp	Arg	Gly	Leu	His	Trp	Val	Gln	Ser
275				280				285							
Asn	Leu	Leu	Lys	Arg	Asn	Ser	Leu	Arg	Glu	Ala	Ala	Ile	Arg	His	Val
290				295				300							
Leu	Glu	Trp	Ile	Ile	Arg	His	Gln	Asp	Ala	Asp	Gly	Gly	Trp	Gly	Gly
305				310				315				320			
Ile	Gln	Pro	Pro	Trp	Val	Tyr	Gly	Leu	Met	Ala	Leu	His	Gly	Glu	Gly
325				330				335							
Tyr	Gln	Leu	Tyr	His	Pro	Val	Met	Ala	Lys	Ala	Leu	Ser	Ala	Leu	Asp
340				345				350							
Asp	Pro	Gly	Trp	Arg	His	Asp	Arg	Gly	Glu	Ser	Ser	Trp	Ile	Gln	Ala
355				360				365							
Thr	Asn	Ser	Pro	Val	Trp	Asp	Thr	Met	Leu	Ala	Leu	Met	Ala	Leu	Lys
370				375				380							
Asp	Ala	Lys	Ala	Glu	Asp	Arg	Phe	Thr	Pro	Glu	Met	Asp	Lys	Ala	Ala
385				390				395				400			
Asp	Trp	Leu	Leu	Ala	Arg	Gln	Val	Lys	Val	Lys	Gly	Asp	Trp	Ser	Ile
405				410				415							
Lys	Leu	Pro	Asp	Val	Glu	Pro	Gly	Gly	Trp	Ala	Phe	Glu	Tyr	Ala	Asn

420	425	430
Asp Arg Tyr Pro Asp Thr Asp Asp Thr Ala Val Ala Leu Ile Ala Leu		
435	440	445
Ser Ser Tyr Arg Asp Lys Glu Glu Trp Gln Lys Lys Gly Val Glu Asp		
450	455	460
Ala Ile Thr Arg Gly Val Asn Trp Leu Ile Ala Met Gln Ser Glu Cys		
465	470	475
Gly Gly Trp Gly Ala Phe Asp Lys Asp Asn Asn Arg Ser Ile Leu Ser		
485	490	495
Lys Thr 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
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
Ile Asp Ser Ile Gly Lys Gly Pro Thr Thr Pro Ser Gln Thr Ala Trp		
610	615	620
Ala Leu Met Gly Leu Ile Ala Ala Asn Arg Pro Glu Asp Tyr Glu Ala		
625	630	635
Ile Ala Lys Gly Cys His Tyr Leu Ile Asp Arg Gln Glu Gln Asp Gly		
645	650	655
Ser Trp Lys Glu Glu Glu Phe Thr Gly Thr Gly Phe Pro Gly Tyr Gly		
660	665	670
Val Gly Gln Thr Ile Lys Leu Asp Asp Pro Ala Leu Ser Lys Arg Leu		
675	680	685
Leu Gln Gly Ala Glu Leu Ser Arg Ala Phe Met Leu Arg Tyr Asp Phe		
690	695	700
Tyr Arg Gln Phe Phe Pro Ile Met Ala Leu Ser Arg Ala Glu Arg Leu		
705	710	715
Ile Asp Leu Asn Asn		720
725		

<210>	60	
<211>	2178	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成ZmoSHC1衍生物	
<220>		
<221>	misc_feature	
<222>	(553) .. (555)	
<223>	ngn是agr或cgn	
<220>		
<221>	misc_feature	
<222>	(1494) .. (1494)	
<223>	n是a, c, g,或t	
<400>	60	
atgggtattg acagaatgaa tagcttaagt cgcttggttaa tgaagaagat tttcggggct	60	
gaaaaaacct cgtataaacc ggcttccgat accataatcg gaacggatac cctgaaaaga	120	
ccgaaccggc ggccctgaacc gacggcaaaa gtcgacaaaa cgatattcaa gactatgggg	180	
aatagtctga ataataccct tgtttcagcc tgtgactggt tgatcggaca acaaaagccc	240	
gatggtcatt gggtcggcgc cgtggaatcc aatgcttcga tggaagcaga atgggtgtctg	300	
gccttgtggg ttttgggtct ggaagatcat ccgcttcgtc caagattggg caatgctctt	360	
ttggaaatgc agcgggaaga tggctcttgg ggagtctatt tcggcgctgg aaatggcgat	420	
atcaatgcca cggftgaagc ctatgcggcc ttgcggctctt tgggggtattc tgccgataat	480	
cctgttttga aaaaagcggc agcatggatt gctgaaaaag gcggattaaa aaatatccgt	540	
gtcttttacc gtngntggct ggcgttgatc ggggaatggc cttgggaaaa gaccctaac	600	
cttccccctg aaattatctg gttccctgat aattttgtct tttcgattta taattttgcc	660	
caatgggcgc gggcaaccat ggtgccgatt gctattctgt ccgcgagacg accaagccgc	720	
ccgctgcgcc ctcaagaccg attggatgaa ctgtttccag aaggccgcgc tcgctttgat	780	
tatgaattgc cgaaaaaaga aggcatcgat ctttggtcgc aatttttccg aaccactgac	840	
cgtggattac attgggttca gtccaatctg ttaaagcgca atagcttgcg tgaagccgct	900	
atccgtcatg ttttggaatg gattatccgg catcaggatg ccgatggcgg ttgggggtgga	960	
attcagccac cttgggtcta tggtttgatg gcgttacatg gtgaaggcta tcagctttat	1020	
catccggtga tggccaaggc tttgtcggct ttggatgac cgggttggcg acatgacaga	1080	
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg	1140	
atggcggttaa aagacgcaa ggccgaggat cgttttacgc cggaaatgga taaggccgcc	1200	
gattggcttt tggctcgaca ggtcaaagtc aaaggcgatt ggtcaatcaa actgcccgat	1260	
gttgaacccg gtggatgggc atttgaatat gccaatgac gctatcccga taccgatgat	1320	
accgccgtcg ctttgatcgc ctttctctct tatcgtgata aggaggagtg gcaaaagaaa	1380	
ggcgttgagg acgccattac ccgtgggggtt aattggttga tcgcatgca aagcgaatgt	1440	

```

ggcggttggg gagcctttga taaggataat aacagaagta tcctttccaa aacncctttt 1500
tgtgatttcg gagaatctat tgatccgcct tcagtcgatg taacggcgca tgttttagag 1560
gcctttggca ccttgggact gtcccgcgat atgccgggtca tccaaaaagc gatcgactat 1620
gtccgttccg aacaggaagc cgaaggcgcg tggtttggtc gttggggcgt taattatata 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gaccagacct 1740
tacatcacca aggcttgcga ttggctggtc gcacatcagc aggaagacgg cggttggggc 1800
gaaagctgct cttcctatat ggagattgat tccattggga agggcccaac cacgccgtcc 1860
cagactgctt gggctttgat ggggttgatc gcggccaatc gtcccgaaga ttatgaagcc 1920
attgccaagg gatgccatta tctgattgat cgccaagagc aggatggtag ctggaaagaa 1980
gaagaattca ccggcaccgg attccccggt tatggcggtg gtcagacgat caagttggat 2040
gatccggctt tatcgaaacg attgcttcaa ggcgtgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

```

<210> 61

<211> 725

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 61

```

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

```

145	150	155	160
Pro Val Leu Lys Lys	Ala Ala Ala Trp	Ile Ala Glu Lys Gly	Gly Leu
165	170	175	
Lys Asn Ile Arg Val	Phe Thr Arg Tyr	Trp Leu Ala Leu	Ile Gly Glu
180	185	190	
Trp Pro Trp Glu Lys	Thr Pro Asn Leu	Pro Pro Glu Ile	Ile Trp Phe
195	200	205	
Pro Asp Asn Phe Val	Phe Ser Ile Tyr	Asn Phe Ala Gln	Trp Ala Arg
210	215	220	
Ala Thr Met Val Pro	Ile Ala Ile Leu	Ser Ala Arg Arg	Pro Ser Arg
225	230	235	240
Pro Leu Arg Pro Gln	Asp Arg Leu Asp	Glu Leu Phe Pro	Glu Gly Arg
245	250	255	
Ala Arg Phe Asp Tyr	Glu Leu Pro Lys	Lys Glu Gly Ile	Asp Leu Trp
260	265	270	
Ser Gln Phe Phe Arg	Thr Thr Asp Arg	Gly Leu His Trp	Val Gln Ser
275	280	285	
Asn Leu Leu Lys Arg	Asn Ser Leu Arg	Glu Ala Ala Ile	Arg His Val
290	295	300	
Leu Glu Trp Ile Ile	Arg His Gln Asp	Ala Asp Gly Gly	Trp Gly Gly
305	310	315	320
Ile Gln Pro Pro Trp	Val Tyr Gly Leu	Met Ala Leu His	Gly Glu Gly
325	330	335	
Tyr Gln Leu Tyr His	Pro Val Met Ala	Lys Ala Leu Ser	Ala Leu Asp
340	345	350	
Asp Pro Gly Trp Arg	His Asp Arg Gly	Glu Ser Ser Trp	Ile Gln Ala
355	360	365	
Thr Asn Ser Pro Val	Trp Asp Thr Met	Leu Ala Leu Met	Ala Leu Lys
370	375	380	
Asp Ala Lys Ala Glu	Asp Arg Phe Thr	Pro Glu Met Asp	Lys Ala Ala
385	390	395	400
Asp Trp Leu Leu Ala	Arg Gln Val Lys	Val Lys Gly Asp	Trp Ser Ile
405	410	415	
Lys Leu Pro Asp Val	Glu Pro Gly Gly	Trp Ala Phe Glu	Tyr Ala Asn
420	425	430	
Asp Arg Tyr Pro Asp	Thr Asp Asp Thr	Ala Val Ala Leu	Ile Ala Leu
435	440	445	
Ser Ser Tyr Arg Asp	Lys Glu Glu Trp	Gln Lys Lys Gly	Val Glu Asp
450	455	460	

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

<210> 62

<211> 2178

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 62

```

atgggtattg acagaatgaa tagcttaagt cgcttgtaa tgaagaagat tttcggggct      60
gaaaaaacct cgtataaacc ggcttccgat accataatcg gaacggatac cctgaaaaga      120
ccgaaccggc ggccatgaacc gacggcaaaa gtcgacaaaa cgatattcaa gactatgggg      180
aatagtctga ataataccct tgtttcagcc tgtgactggg tgatcggaca acaaaagccc      240
gatggtcatt gggctcggcgc cgtggaatcc aatgcttcga tggaagcaga atgggtgtctg      300
gccttggtgg ttttgggtct ggaagatcat ccgcttcgct caagattggg caatgctctt      360
ttggaaatgc agcgggaaga tggctcttgg ggagtctatt tcggcgctgg aaatggcgat      420
atcaatgcca cggttgaagc ctatgcggcc ttgcggctctt tgggggtattc tgccgataat      480
cctgttttga aaaaagcggc agcatggatt gctgaaaaag gcggattaaa aaatatccgt      540
gtctttaccc gttattggct ggcttgatc ggggaatggc cttgggaaaa gaccctaac      600
cttccccctg aaattatctg gttccctgat aattttgtct tttcgattta taattttgcc      660
caatgggcgc gggcaaccat ggtgccgatt gctattctgt ccgcgagacg accaagccgc      720
ccgctgcgcc ctcaagaccg attggatgaa ctgtttccag aaggccgcgc tcgctttgat      780
tatgaattgc cgaaaaaaga aggcacgat ctttggtcgc aatttttccg aaccactgac      840
cgtggattac attgggttca gtccaatctg ttaaagcgca atagcttgcg tgaagccgct      900
atccgtcatg ttttggaatg gattatccgg catcaggatg ccgatggcgg ttgggggtgga      960
attcagccac cttgggtcta tggtttgatg gcgttacatg gtgaaggcta tcagctttat      1020
catccggtga tggccaaggc tttgtcggct ttggatgatc ccggttggcg acatgacaga      1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg      1140
atggcggttaa aagacgcaa ggccaggat cgttttacgc cggaaatgga taaggccgcc      1200
gattggcttt tggctcgaca ggtcaaagtc aaaggcgatt ggtcaatcaa actgcccgat      1260
gttgaacccg gtggatgggc atttgaatat gccaatgatc gctatcccga taccgatgat      1320
accgccgtcg ctttgatcgc ctttctctt tatcgtgata aggaggagtg gcaaaagaaa      1380
ggcgttgagg acgccattac ccgtgggggtt aattggttga tcgccatgca aagcgaatgt      1440
ggcggttggg gagcctttga taaggataat aacagaagta tcctttccaa aattcctttt      1500
tgtgatttcg gagaatctat tgatccgcct tcagtcgatg taacggcgca tgttttagag      1560
gcctttggca ccttgggact gtcccgcgat atgccggtca tccaaaaagc gatcgactat      1620
gtccgttccg aacaggaagc cgaaggcgcg tggtttggtc gttggggcgt taattatatc      1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gaccagcct      1740
tacatcacca aggcttgcca ttggctggtc gcacatcagc aggaagacgg cggttggggc      1800
gaaagctgct cttcctatat ggagattgat tccattggga agggcccaac cacgccgtcc      1860
cagactgctt gggctttgat ggggttgatc gcggccaatc gtcccgaaga ttatgaagcc      1920
attgccaagg gatgccatta tctgattgat cgccaagagc aggatggtag ctggaaagaa      1980
gaagaattca ccggcaccgg atayccccgt tatggcgtgg gtcagacgat caagttggat      2040
gatccggctt tatcgaaacg attgcttcaa ggcgtgaac tgtcacgggc gtttatgctg      2100
cgttatgatt tttatcgga attcttccc attatggcgt taagtcgggc agagagactg      2160
attgatttga ataattga

```

<210> 63

<211> 725

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 63

```

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

```


Ala Arg Phe Asp Tyr Glu Leu Pro Lys Lys Glu Gly Ile Asp Leu Trp		
260	265	270
Ser Gln Phe Phe Arg Thr Thr Asp Arg Gly Leu His Trp Val Gln Ser		
275	280	285
Asn Leu Leu Lys Arg Asn Ser Leu Arg Glu Ala Ala Ile Arg His Val		
290	295	300
Leu Glu Trp Ile Ile Arg His Gln Asp Ala Asp Gly Gly Trp Gly Gly		
305	310	315
Ile Gln Pro Pro Trp Val Tyr Gly Leu Met Ala Leu His Gly Glu Gly		
325	330	335
Tyr Gln Leu Tyr His Pro Val Met Ala Lys Ala Leu Ser Ala Leu Asp		
340	345	350
Asp Pro Gly Trp Arg His Asp Arg Gly Glu Ser Ser Trp Ile Gln Ala		
355	360	365
Thr Asn Ser Pro Val Trp Asp Thr Met Leu Ala Leu Met Ala Leu Lys		
370	375	380
Asp Ala Lys Ala Glu Asp Arg Phe Thr Pro Glu Met Asp Lys Ala Ala		
385	390	395
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
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
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
Ile Asp Ser Ile Gly Lys Gly Pro Thr Thr Pro Ser Gln Thr Ala Trp		
610	615	620
Ala Leu Met Gly Leu Ile Ala Ala Asn Arg Pro Glu Asp Tyr Glu Ala		
625	630	635
Ile Ala Lys Gly Cys His Tyr Leu Ile Asp Arg Gln Glu Gln Asp Gly		
645	650	655
Ser Trp Lys Glu Glu Glu Phe Thr Gly Thr Gly Phe Pro Gly Tyr Gly		
660	665	670
Val Gly Gln Thr Ile Lys Leu Asp Asp Pro Ala Leu Ser Lys Arg Leu		
675	680	685
Leu Gln Gly Ala Glu Leu Ser Arg Ala Phe Met Leu Arg Tyr Asp Phe		
690	695	700
Tyr Arg Gln Phe Phe Pro Ile Met Ala Leu Ser Arg Ala Glu Arg Leu		
705	710	715
Ile Asp Leu Asn Asn		

725

<210> 64

<211> 2178

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<220>

<221> misc_feature

<222> (387) .. (387)

<223> n是a, c, g,或t

<220>

<221> misc_feature

<222> (544) .. (546)

<223> ntn是ttr或ctn

<400> 64

atgggtattg acagaatgaa tagcttaagt cgcttggttaa tgaagaagat tttcggggct	60
gaaaaaacct cgtataaacc ggcttccgat accataatcg gaacggatac cctgaaaaga	120
ccgaaccggc ggctgaacc gacggcaaaa gtcgacaaaa cgatattcaa gactatgggg	180

aatagtctga ataataccct tgtttcagcc tgtgactggt tgatcggaca acaaaagccc	240
gatggtcatt gggtcggtgc cgtggaatcc aatgcttcga tggaagcaga atgggtgtctg	300
gccttgtgggt ttttgggtct ggaagatcat ccgcttcgtc caagattggg caatgctctt	360
ttggaaatgc agcgggaaga tggcgcntgg ggagtctatt tcggcgctgg aaatggcgat	420
atcaatgcca cggttgaagc ctatgcggcc ttgcggtctt tggggatttc tgccgataat	480
cctgttttga aaaaagcggc agcatggatt gctgaaaaag gcggattaaa aaatatccgt	540
gtcntnacc cgttattggct ggcgttgatc ggggaatggc cttgggaaaa gaccctaac	600
cttccccctg aaattatctg gttccctgat aattttgtct tttcgattta taattttgcc	660
caatgggctg gggcaaccat ggtgccgatt gctattctgt ccgcgagacg accaagccgc	720
ccgctgcgcc ctcaagaccg attggatgaa ctgtttccag aaggccgcgc tcgctttgat	780
tatgaattgc cgaaaaaaga aggcacgat ctttggtcgc aatttttccg aaccactgac	840
cgtggattac attgggttca gtccaatctg ttaaagcgca atagcttgcg tgaagccgct	900
atccgtcatg ttttggaatg gattatccgg catcaggatg ccgatggcgg ttgggggtgga	960
attcagccac cttgggtcta tggtttgatg gcgttacatg gtgaaggcta tcagctttat	1020
catccggtga tggccaaggc tttgtcggct ttggatgatc ccggttggcg acatgacaga	1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg	1140
atggcggtta aagacgcaa ggccgaggat cgttttacgc cggaaatgga taaggccgcc	1200
gattggcttt tggctcgaca ggtcaaagtc aaaggcgatt ggtcaatcaa actgcccgat	1260
gttgaacccg gtggatgggc atttgaatat gccaatgatc gctatcccga taccgatgat	1320
accgccgtcg ctttgatcgc cttttcctct tatcgtgata aggaggagtg gcaaaagaaa	1380
ggcgttgagg acgccattac ccgtggggtt aattggttga tcgccatgca aagcgaatgt	1440
ggcggttggg gagcctttga taaggataat aacagaagta tcctttccaa aattcctttt	1500
tgtgatttcg gagaatctat tgatccgcct tcagtcgatg taacggcgca tgttttagag	1560
gcctttggca ccttgggact gtcccgcgat atgccggtca tccaaaaagc gatcgactat	1620
gtccgttccg aacaggaagc cgaaggcgcg tggtttggtc gttggggcgt taattatate	1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gaccagcct	1740
tacatcacca aggcttgcca ttggctggtc gcacatcagc aggaagacgg cggttggggc	1800
gaaagctgct cttcctatat ggagattgat tccattggga agggcccaac cacgccgtcc	1860
cagactgctt gggctttgat ggggttgatc gcggccaatc gtcccgaaga ttatgaagcc	1920
attgccaagg gatgccatta tctgattgat cgccaagagc aggatggtag ctggaaagaa	1980
gaagaattca ccggcaccgg attccccggt tatggcgtgg gtcagacgat caagttggat	2040
gatccggctt tatcgaaacg attgttcaa ggcgtgaac tgtcacgggc gtttatgctg	2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg	2160
attgatttga ataattga	2178
<210> 65	
<211> 725	
<212> PRT	
<213> 人工序列	
<220>	

<223> 合成ZmoSHC1衍生物

<400> 65

Met	Gly	Ile	Asp	Arg	Met	Asn	Ser	Leu	Ser	Arg	Leu	Leu	Met	Lys	Lys
1				5				10					15		
Ile	Phe	Gly	Ala	Glu	Lys	Thr	Ser	Tyr	Lys	Pro	Ala	Ser	Asp	Thr	Ile
		20						25					30		
Ile	Gly	Thr	Asp	Thr	Leu	Lys	Arg	Pro	Asn	Arg	Arg	Pro	Glu	Pro	Thr
		35						40					45		
Ala	Lys	Val	Asp	Lys	Thr	Ile	Phe	Lys	Thr	Met	Gly	Asn	Ser	Leu	Asn
		50						55				60			
Asn	Thr	Leu	Val	Ser	Ala	Cys	Asp	Trp	Leu	Ile	Gly	Gln	Gln	Lys	Pro
65				70					75					80	
Asp	Gly	His	Trp	Val	Gly	Ala	Val	Glu	Ser	Asn	Ala	Ser	Met	Glu	Ala
				85					90					95	
Glu	Trp	Cys	Leu	Ala	Leu	Trp	Phe	Leu	Gly	Leu	Glu	Asp	His	Pro	Leu
				100					105					110	
Arg	Pro	Arg	Leu	Gly	Asn	Ala	Leu	Leu	Glu	Met	Gln	Arg	Glu	Asp	Gly
				115					120					125	
Ser	Trp	Gly	Val	Tyr	Phe	Gly	Ala	Gly	Asn	Gly	Asp	Ile	Asn	Ala	Thr
									135					140	
Val	Glu	Ala	Tyr	Ala	Ala	Leu	Arg	Ser	Leu	Gly	Tyr	Ser	Ala	Asp	Asn
145														160	
Pro	Val	Leu	Lys	Lys	Ala	Ala	Ala	Trp	Ile	Ala	Glu	Lys	Gly	Gly	Leu
														175	
Lys	Asn	Ile	Arg	Val	Phe	Thr	Arg	Tyr	Trp	Leu	Ala	Leu	Ile	Gly	Glu
				180					185					190	
Trp	Pro	Trp	Glu	Lys	Thr	Pro	Asn	Leu	Pro	Pro	Glu	Ile	Ile	Trp	Phe
									200					205	
Pro	Asp	Asn	Phe	Val	Phe	Ser	Ile	Tyr	Asn	Phe	Ala	Gln	Trp	Ala	Arg
														220	
Ala	Thr	Met	Val	Pro	Ile	Ala	Ile	Leu	Ser	Ala	Arg	Arg	Pro	Ser	Arg
225														240	
Pro	Leu	Arg	Pro	Gln	Asp	Arg	Leu	Asp	Glu	Leu	Phe	Pro	Glu	Gly	Arg
														255	
Ala	Arg	Phe	Asp	Tyr	Glu	Leu	Pro	Lys	Lys	Glu	Gly	Ile	Asp	Leu	Trp
														270	
Ser	Gln	Phe	Phe	Arg	Thr	Thr	Asp	Arg	Gly	Leu	His	Trp	Val	Gln	Ser
														285	
Asn	Leu	Leu	Lys	Arg	Asn	Ser	Leu	Arg	Glu	Ala	Ala	Ile	Arg	His	Val

290	295	300
Leu Glu Trp Ile Ile Arg His Gln Asp Ala Asp Gly Gly Trp Gly Gly		
305	310	315
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
Thr Asn Ser Pro Val Trp Asp Thr Met Leu Ala Leu Met Ala Leu Lys		365
	370	375
Asp Ala Lys Ala Glu Asp Arg Phe Thr Pro Glu Met Asp Lys Ala Ala		380
385	390	395
Asp Trp Leu Leu Ala Arg Gln Val Lys Val Lys Gly Asp Trp Ser Ile		400
	405	410
Lys Leu Pro Asp Val Glu Pro Gly Gly Trp Ala Phe Glu Tyr Ala Asn		415
	420	425
Asp Arg Tyr Pro Asp Thr Asp Asp Thr Ala Val Ala Leu Ile Ala Leu		430
	435	440
Ser Ser Tyr Arg Asp Lys Glu Glu Trp Gln Lys Lys Gly Val Glu Asp		445
	450	455
Ala Ile Thr Arg Gly Val Asn Trp Leu Ile Ala Met Gln Ser Glu Cys		460
465	470	475
Gly Gly Trp Gly Ala Phe Asp Lys Asp Asn Asn Arg Ser Ile Leu Ser		480
	485	490
Lys Ile Pro Phe Cys Asp Phe Gly Glu Ser Ile Asp Pro Pro Ser Val		495
	500	505
Asp Val Thr Ala His Val Leu Glu Ala Phe Gly Thr Leu Gly Leu Ser		510
	515	520
Arg Asp Met Pro Val Ile Gln Lys Ala Ile Asp Tyr Val Arg Ser Glu		525
	530	535
Gln Glu Ala Glu Gly Ala Trp Phe Gly Arg Trp Gly Val Asn Tyr Ile		540
545	550	555
Tyr Gly Thr Gly Ala Val Leu Pro Ala Leu Ala Ala Ile Gly Glu Asp		560
	565	570
Met Thr Gln Pro Tyr Ile Thr Lys Ala Cys Asp Trp Leu Val Ala His		575
	580	585
Gln Gln Glu Asp Gly Gly Trp Gly Glu Ser Cys Ser Ser Tyr Met Glu		590
	595	600
		605

Ile Asp Ser Ile Gly Lys Gly Pro Thr Thr Pro Ser Gln Thr Ala Trp	
610	615 620
Ala Leu Met Gly Leu Ile Ala Ala Asn Arg Pro Glu Asp Tyr Glu Ala	
625	630 635 640
Ile Ala Lys Gly Cys His Tyr Leu Ile Asp Arg Gln Glu Gln Asp Gly	
	645 650 655
Ser Trp Lys Glu Glu Glu Phe Thr Gly Thr Gly Tyr Pro Gly Tyr Gly	
	660 665 670
Val Gly Gln Thr Ile Lys Leu Asp Asp Pro Ala Leu Ser Lys Arg Leu	
	675 680 685
Leu Gln Gly Ala Glu Leu Ser Arg Ala Phe Met Leu Arg Tyr Asp Phe	
	690 695 700
Tyr Arg Gln Phe Phe Pro Ile Met Ala Leu Ser Arg Ala Glu Arg Leu	
705	710 715 720
Ile Asp Leu Asn Asn	
	725

<210> 66

<211> 2178

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 66

atgggtattg acagaatgaa tagcttaagt cgcttggttaa tgaagaagat tttcggggct	60
gaaaaaacct cgtataaacc ggcttccgat accataatcg gaacggatac cctgaaaaga	120
ccgaaccggc ggcctgaacc gacggcaaaa gtcgacaaaa cgatattcaa gactatgggg	180
aatagtctga ataataccct tgtttcagcc tgtgactggg tgatcggaca acaaaagccc	240
gatggtcatt gggctcggcgc cgtggaatcc aatgcttcga tggaagcaga atgggtgtctg	300
gccttggtgg ttttggtgct ggaagatcat ccgcttcgct caagattggg caatgctctt	360
ttggaaatgc agcgggaaga tggctcttgg ggagtctatt tcggcgctgg aaatggcgat	420
atcaatgcca cggttgaagc ctatgcggcc ttgcggtctt tgggggtattc tgccgataat	480
cctgttttga aaaaagcggc agcatggatt gctgaaaaag gcggattaaa aaatatccgt	540
gtctttaccc gttattggct ggcgttgatc ggggaatggc cttgggaaaa gaccctaac	600
cttccccctg aaattatctg gttccctgat aattttgtct tttcgattta taattttgcc	660
caatgggctg gggcaaccat ggtgccgatt gctattctgt ccgcgagacg accaagccgc	720
ccgctgcgcc ctcaagaccg attggatgaa ctgtttccag aaggccgcgc tcgctttgat	780
tatgaattgc cgaaaaaaga aggcacgat ctttggtcgc aatttttccg aaccactgac	840
cgtggattac attgggttca gtccaatctg ttaaagcgca atagcttgcg tgaagccgct	900
atccgtcatg ttttggaatg gattatccgg catcaggatg ccgatggcgg ttgggggtgga	960

```

attcagccac cttgggtcta tggtttgatg gcgttacatg gtgaaggcta tcagctttat 1020
catccggtga tggccaaggc tttgtcggct ttggatgac ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcggttaa aagacgccaa ggccgaggat cgttttacgc cggaaatgga taaggccgcc 1200
gattggcttt tggctcgaca ggtcaaagtc aaaggcgatt ggtcaatcaa actgcccgat 1260
gttgaacccg gtggatgggc atttgaatat gccaatgac gctatcccga taccgatgat 1320
accgccgtcg ctttgatcgc cttttcctct tatcgtgata aggaggagt gcaaaagaaa 1380
ggcggtgagg acgccattac ccgtgggggtt aattggttga tcgccatgca aagcgaatgt 1440
ggcggttggg gagcctttga taaggataat aacagaagta tcctttccaa aattcctttt 1500
tgtgatttcg gagaatctat tgatccgcct tcagtcgatg taacggcgca tgttttagag 1560
gccttttgga ccttgggact gtcccgcgat atgccggtca tccaaaaagc gatcgactat 1620
gtccgttccg aacaggaagc cgaaggcgcg tggtttggtc gttggggcgt taattatata 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gaccagcct 1740
tacatcacca aggcttgca ttggctggtc gcacatcagc aggaagacgg cggttggggc 1800
gaaagctgct cttcctatat ggagattgat tccattggga agggcccaac cacgccgtcc 1860
cagactgctt gggctttgat ggggttgatc gcggccaatc gtcccgaaga ttatgaagcc 1920
attgccaagg gatgccatta tctgattgat cgccaagagc aggatggtag ctgaaagaa 1980
gaagaattca ccggcaccgg ataycccggg tatggcgtgg gtcagacgat caagttggat 2040
gatccggctt tatcgaacg attgcttcaa ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttccc attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

```

<210> 67

<211> 725

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 67

```

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

```

				85					90					95	
Glu	Trp	Cys	Leu	Ala	Leu	Trp	Phe	Leu	Gly	Leu	Glu	Asp	His	Pro	Leu
				100					105					110	
Arg	Pro	Arg	Leu	Gly	Asn	Ala	Leu	Leu	Glu	Met	Gln	Arg	Glu	Asp	Gly
				115					120					125	
Ser	Trp	Gly	Val	Tyr	Phe	Gly	Ala	Gly	Asn	Gly	Asp	Ile	Asn	Ala	Thr
				130					135					140	
Val	Glu	Ala	Tyr	Ala	Ala	Leu	Arg	Ser	Leu	Gly	Tyr	Ser	Ala	Asp	Asn
				145					150					155	
Pro	Val	Leu	Lys	Lys	Ala	Ala	Ala	Trp	Ile	Ala	Glu	Lys	Gly	Gly	Leu
				165					170					175	
Lys	Asn	Ile	Arg	Val	Leu	Thr	Arg	Tyr	Trp	Leu	Ala	Leu	Ile	Gly	Glu
				180					185					190	
Trp	Pro	Trp	Glu	Lys	Thr	Pro	Asn	Leu	Pro	Pro	Glu	Ile	Ile	Trp	Phe
				195					200					205	
Pro	Asp	Asn	Phe	Val	Phe	Ser	Ile	Tyr	Asn	Phe	Ala	Gln	Trp	Ala	Arg
				210					215					220	
Ala	Thr	Met	Val	Pro	Ile	Ala	Ile	Leu	Ser	Ala	Arg	Arg	Pro	Ser	Arg
				225					230					235	
Pro	Leu	Arg	Pro	Gln	Asp	Arg	Leu	Asp	Glu	Leu	Phe	Pro	Glu	Gly	Arg
				245					250					255	
Ala	Arg	Phe	Asp	Tyr	Glu	Leu	Pro	Lys	Lys	Glu	Gly	Ile	Asp	Leu	Trp
				260					265					270	
Ser	Gln	Phe	Phe	Arg	Thr	Thr	Asp	Arg	Gly	Leu	His	Trp	Val	Gln	Ser
				275					280					285	
Asn	Leu	Leu	Lys	Arg	Asn	Ser	Leu	Arg	Glu	Ala	Ala	Ile	Arg	His	Val
				290					295					300	
Leu	Glu	Trp	Ile	Ile	Arg	His	Gln	Asp	Ala	Asp	Gly	Gly	Trp	Gly	Gly
				305					310					315	
Ile	Gln	Pro	Pro	Trp	Val	Tyr	Gly	Leu	Met	Ala	Leu	His	Gly	Glu	Gly
				325					330					335	
Tyr	Gln	Leu	Tyr	His	Pro	Val	Met	Ala	Lys	Ala	Leu	Ser	Ala	Leu	Asp
				340					345					350	
Asp	Pro	Gly	Trp	Arg	His	Asp	Arg	Gly	Glu	Ser	Ser	Trp	Ile	Gln	Ala
				355					360					365	
Thr	Asn	Ser	Pro	Val	Trp	Asp	Thr	Met	Leu	Ala	Leu	Met	Ala	Leu	Lys
				370					375					380	
Asp	Ala	Lys	Ala	Glu	Asp	Arg	Phe	Thr	Pro	Glu	Met	Asp	Lys	Ala	Ala
				385					390					395	
															400

Asp Trp Leu Leu Ala Arg Gln Val Lys Val Lys Gly Asp Trp Ser Ile		
405	410	415
Lys Leu Pro Asp Val Glu Pro Gly Gly Trp Ala Phe Glu Tyr Ala Asn		
420	425	430
Asp Arg Tyr Pro Asp Thr Asp Asp Thr Ala Val Ala Leu Ile Ala Leu		
435	440	445
Ser Ser Tyr Arg Asp Lys Glu Glu Trp Gln Lys Lys Gly Val Glu Asp		
450	455	460
Ala Ile Thr Arg Gly Val Asn Trp Leu Ile Ala Met Gln Ser Glu Cys		
465	470	475
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
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
Ile Asp Ser Ile Gly Lys Gly Pro Thr Thr Pro Ser Gln Thr Ala Trp		
610	615	620
Ala Leu Met Gly Leu Ile Ala Ala Asn Arg Pro Glu Asp Tyr Glu Ala		
625	630	635
Ile Ala Lys Gly Cys His Tyr Leu Ile Asp Arg Gln Glu Gln Asp Gly		
645	650	655
Ser Trp Lys Glu Glu Glu Phe Thr Gly Thr Gly Phe Pro Gly Tyr Gly		
660	665	670
Val Gly Gln Thr Ile Lys Leu Asp Asp Pro Ala Leu Ser Lys Arg Leu		
675	680	685
Leu Gln Gly Ala Glu Leu Ser Arg Ala Phe Met Leu Arg Tyr Asp Phe		
690	695	700
Tyr Arg Gln Phe Phe Pro Ile Met Ala Leu Ser Arg Ala Glu Arg Leu		

705	710	715	720
Ile Asp Leu Asn Asn			
	725		
<210> 68			
<211> 2178			
<212> DNA			
<213> 人工序列			
<220>			
<223> 合成ZmoSHC1衍生物			
<220>			
<221> misc_feature			
<222> (544) .. (546)			
<223> ntn是ttr或ctn			
<400> 68			
atgggtattg acagaatgaa tagcttaagt cgcttggttaa tgaagaagat tttcggggct			60
gaaaaaacct cgtataaacc ggcttccgat accataatcg gaacggatac cctgaaaaga			120
ccgaaccggc ggcttgaacc gacggcaaaa gtcgacaaaa cgatattcaa gactatgggg			180
aatagtctga ataataccct tgtttcagcc tgtgactggt tgatcggaca acaaaagccc			240
gatggtcatt gggctcggtc cgtggaatcc aatgcttcga tggaagcaga atggtgtctg			300
gccttgtggt ttttgggtct ggaagatcat ccgcttcgtc caagattggg caatgctctt			360
ttggaaatgc agcgggaaga tggctcttgg ggagtctatt tcggcgctgg aaatggcgat			420
atcaatgcca cggttgaagc ctatgcggcc ttgcggtctt tggggtattc tgccgataat			480
cctgttttga aaaaagcggc agcatggatt gctgaaaaag gcggattaaa aaatatccgt			540
gtcntnaccg gttattggct ggcgttgatc ggggaatggc cttgggaaaa gaccctaac			600
cttccccctg aaattatctg gttccctgat aattttgtct tttcgattta taattttgcc			660
caatgggcgc gggcaaccat ggtgccgatt gctattctgt ccgcgagacg accaagccgc			720
ccgctgcgcc ctcaagaccg attggatgaa ctgtttccag aaggccgcgc tcgctttgat			780
tatgaattgc cgaaaaaaga aggcacgat ctttggtcgc aatttttccg aaccactgac			840
cgtggattac attgggttca gtccaatctg ttaaagcgca atagcttgcg tgaagccgct			900
atccgtcatg ttttggaatg gattatccgg catcaggatg ccgatggcgg ttgggggtgga			960
attcagccac cttgggtcta tggtttgatg gcgttacatg gtgaaggcta tcagctttat			1020
catccggtga tggccaaggc tttgtcggct ttggatgac cggtttggcg acatgacaga			1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg			1140
atggcggttaa aagacgccaa ggccgaggat cgttttacgc cggaatgga taaggccgcc			1200
gattggcttt tggctcgaca ggtcaaagtc aaaggcgatt ggtcaatcaa actgcccgat			1260
gttgaaccgc gtggatgggc atttgaatat gccaatgatc gctatcccga taccgatgat			1320
accgccgtcg ctttgatcgc cttttcctct tatcgtgata aggaggagtg gcaaaagaaa			1380
ggcgttgagg acgccattac ccgtggggtt aattggttga tcgccatgca aagcgaatgt			1440
ggcggttggg gagccttga taaggataat aacagaagta tcctttccaa aattcctttt			1500

```

tgtgatttcg gagaatctat tgatccgcct tcagtcgatg taacggcgca tgtttttagag 1560
gcctttggca ccttgggact gtcccgcgat atgccgggtca tccaaaaagc gatcgactat 1620
gtccgttccg aacaggaagc cgaaggcgcg tggtttggtc gttggggcgt taattatatac 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gaccagcct 1740
tacatcacca aggcttgcga ttggctggtc gcacatcagc aggaagacgg cggttggggc 1800
gaaagctgct cttcctatat ggagattgat tccattggga agggcccaac cacgccgtcc 1860
cagactgctt gggctttgat ggggttgatc gcggccaatc gtcccgaaga ttatgaagcc 1920
attgccaagg gatgccatta tctgattgat cgccaagagc aggatggtag ctggaagaa 1980
gaagaattca ccggcaccgg attccccggt tatggcgtgg gtcagacgat caagttggat 2040
gatccggctt tatcgaaacg attgcttcaa ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcgga attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

```

<210> 69

<211> 725

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 69

```

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

```

Pro Val Leu Lys Lys Ala Ala Ala Trp Ile Ala Glu Lys Gly Gly Leu		
165	170	175
Lys Asn Ile Arg Val Leu Thr Arg Tyr Trp Leu Ala Leu Ile Gly Glu		
180	185	190
Trp Pro Trp Glu Lys Thr Pro Asn Leu Pro Pro Glu Ile Ile Trp Phe		
195	200	205
Pro Asp Asn Phe Val Phe Ser Ile Tyr Asn Phe Ala Gln Trp Ala Arg		
210	215	220
Ala Thr Met Val Pro Ile Ala Ile Leu Ser Ala Arg Arg Pro Ser Arg		
225	230	235
Pro Leu Arg Pro Gln Asp Arg Leu Asp Glu Leu Phe Pro Glu Gly Arg		
245	250	255
Ala Arg Phe Asp Tyr Glu Leu Pro Lys Lys Glu Gly Ile Asp Leu Trp		
260	265	270
Ser Gln Phe Phe Arg Thr Thr Asp Arg Gly Leu His Trp Val Gln Ser		
275	280	285
Asn Leu Leu Lys Arg Asn Ser Leu Arg Glu Ala Ala Ile Arg His Val		
290	295	300
Leu Glu Trp Ile Ile Arg His Gln Asp Ala Asp Gly Gly Trp Gly Gly		
305	310	315
Ile Gln Pro Pro Trp Val Tyr Gly Leu Met Ala Leu His Gly Glu Gly		
325	330	335
Tyr Gln Leu Tyr His Pro Val Met Ala Lys Ala Leu Ser Ala Leu Asp		
340	345	350
Asp Pro Gly Trp Arg His Asp Arg Gly Glu Ser Ser Trp Ile Gln Ala		
355	360	365
Thr Asn Ser Pro Val Trp Asp Thr Met Leu Ala Leu Met Ala Leu Lys		
370	375	380
Asp Ala Lys Ala Glu Asp Arg Phe Thr Pro Glu Met Asp Lys Ala Ala		
385	390	395
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
Ile Asp Ser Ile Gly Lys Gly Pro Thr Thr Pro Ser Gln Thr Ala Trp			
	610	615	620
Ala Leu Met Gly Leu Ile Ala Ala Asn Arg Pro Glu Asp Tyr Glu Ala			
625	630	635	640
Ile Ala Lys Gly Cys His Tyr Leu Ile Asp Arg Gln Glu Gln Asp Gly			
	645	650	655
Ser Trp Lys Glu Glu Glu Phe Thr Gly Thr Gly Tyr Pro Gly Tyr Gly			
	660	665	670
Val Gly Gln Thr Ile Lys Leu Asp Asp Pro Ala Leu Ser Lys Arg Leu			
	675	680	685
Leu Gln Gly Ala Glu Leu Ser Arg Ala Phe Met Leu Arg Tyr Asp Phe			
	690	695	700
Tyr Arg Gln Phe Phe Pro Ile Met Ala Leu Ser Arg Ala Glu Arg Leu			
705	710	715	720
Ile Asp Leu Asn Asn			
	725		

<210> 70

<211> 2178

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<220>

<221> misc_feature

<222> (544) .. (546)

<223> ntn是ttr或ctn

<400> 70

```

atgggtattg acagaatgaa tagcttaagt cgcttgtaa tgaagaagat tttcggggct      60
gaaaaaacct cgtataaacc ggcttccgat accataatcg gaacggatac cctgaaaaga      120
ccgaaccggc ggcctgaacc gacggcaaaa gtcgacaaaa cgatattcaa gactatgggg      180
aatagtctga ataataccct tgtttcagcc tgtgactggt tgatcggaca acaaaagccc      240
gatggtcatt gggtcggtgc cgtggaatcc aatgcttcga tggaagcaga atggtgtctg      300
gccttggtgt ttttgggtct ggaagatcat ccgcttcgtc caagattggg caatgctctt      360
ttggaaatgc agcgggaaga tggtctcttg ggagtctatt tcggcgctgg aaatggcgat      420
atcaatgcca cggttgaagc ctatgcggcc ttgcggtctt tgggggtattc tgccgataat      480
cctgttttga aaaaagcggc agcatggatt gctgaaaaag gcgattaaa aaatatccgt      540
gtcntnaccg gttattggct ggcttgatc ggggaatggc cttgggaaaa gaccctaac      600
cttccccctg aaattatctg gttccctgat aattttgtct tttcgattta taattttgcc      660
caatggggcg gggcaaccat ggtgccgatt gctattctgt ccgcgagacg accaagccgc      720
ccgctgcgcc ctcaagaccg attggatgaa ctgtttccag aaggccgcgc tcgctttgat      780
tatgaattgc cgaaaaaaga aggcatcgat ctttggtcgc aatttttccg aaccactgac      840
cgtggattac attgggttca gtccaatctg ttaaagcgca atagcttgcg tgaagccgct      900
atccgtcatg ttttggaatg gattatccgg catcaggatg ccgatggcgg ttgggggtgga      960
attcagccac cttgggtcta tggtttgatg gcgttacatg gtgaaggcta tcagctttat     1020
catccggtga tggccaaggc tttgtcggct ttggatgatc ccggttggcg acatgacaga     1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg     1140
atggcggttaa aagacgcaa ggccgaggat cgttttacgc cggaaatgga taaggccgcc     1200
gattggcttt tggctcgaca ggtcaaagtc aaaggcgatt ggtcaatcaa actgcccgat     1260
gttgaacccg gtggatgggc atttgaatat gccaatgatc gctatcccga taccgatgat     1320
accgccgtcg ctttgatcgc cttttcctct tatcgtgata aggaggagtg gcaaaagaaa     1380
ggcggttgagg acgccattac ccgtgggggtt aattggttga tcgcatgca aagcgaatgt     1440
ggcggttggg gagcctttga taaggataat aacagaagta tcctttccaa aattcctttt     1500
tgtgatttcg gagaatctat tgatccgcct tcagtcgatg taacggcgca tgtttttagag     1560
gcctttggca ccttgggact gtcccgcat atgccggtca tccaaaaagc gatcgactat     1620
gtccgttccg aacaggaagc cgaaggcgcg tggtttggtc gttggggcgt taattatata     1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gaccagcct     1740
tacatcacca aggcttgca ttggctggtc gcacatcagc aggaagacgg cggttggggc     1800
gaaagctgct ctccctatat ggagattgat tccattggga agggcccaac cacgccgtcc     1860
cagactgctt gggctttgat ggggttgatc gcggccaatc gtcccgaaga ttatgaagcc     1920
attgccaagg gatgccatta tctgattgat cgccaagagc aggatggtag ctggaaagaa     1980
gaagaattca ccggcaccgg ataycccggg tatggcgtgg gtcagacgat caagttggat     2040

```

gatccggcctt tatcgaaacg attgcttcaa ggcgctgaac tgtcacgggc gtttatgctg 2100
 cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
 attgatttga ataattga 2178

<210> 71

<211> 725

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 71

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

225	230	235	240
Pro Leu Arg Pro Gln Asp Arg Leu Asp Glu Leu Phe Pro Glu Gly Arg			
	245	250	255
Ala Arg Phe Asp Tyr Glu Leu Pro Lys Lys Glu Gly Ile Asp Leu Trp			
	260	265	270
Ser Gln Phe Phe Arg Thr Thr Asp Arg Gly Leu His Trp Val Gln Ser			
	275	280	285
Asn Leu Leu Lys Arg Asn Ser Leu Arg Glu Ala Ala Ile Arg His Val			
	290	295	300
Leu Glu Trp Ile Ile Arg His Gln Asp Ala Asp Gly Gly Trp Gly Gly			
305	310	315	320
Ile Gln Pro Pro Trp Val Tyr Gly Leu Met Ala Leu His Gly Glu Gly			
	325	330	335
Tyr Gln Leu Tyr His Pro Val Met Ala Lys Ala Leu Ser Ala Leu Asp			
	340	345	350
Asp Pro Gly Trp Arg His Asp Arg Gly Glu Ser Ser Trp Ile Gln Ala			
	355	360	365
Thr Asn Ser Pro Val Trp Asp Thr Met Leu Ala Leu Met Ala Leu Lys			
	370	375	380
Asp Ala Lys Ala Glu Asp Arg Phe Thr Pro Glu Met Asp Lys Ala Ala			
385	390	395	400
Asp Trp Leu Leu Ala Arg Gln Val Lys Val Lys Gly Asp Trp Ser Ile			
	405	410	415
Lys Leu Pro Asp Val Glu Pro Gly Gly Trp Ala Phe Glu Tyr Ala Asn			
	420	425	430
Asp Arg Tyr Pro Asp Thr Asp Asp Thr Ala Val Ala Leu Ile Ala Leu			
	435	440	445
Ser Ser Tyr Arg Asp Lys Glu Glu Trp Gln Lys Lys Gly Val Glu Asp			
	450	455	460
Ala Ile Thr Arg Gly Val Asn Trp Leu Ile Ala Met Gln Ser Glu Cys			
465	470	475	480
Gly Gly Trp Gly Ala Phe Asp Lys Asp Asn Asn Arg Ser Ile Leu Ser			
	485	490	495
Lys Thr 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		
Ile	Asp	Ser	Ile	Gly	Lys	Gly	Pro	Thr	Thr	Pro	Ser	Gln	Thr	Ala	Trp
	610					615					620				
Ala	Leu	Met	Gly	Leu	Ile	Ala	Ala	Asn	Arg	Pro	Glu	Asp	Tyr	Glu	Ala
625					630					635					640
Ile	Ala	Lys	Gly	Cys	His	Tyr	Leu	Ile	Asp	Arg	Gln	Glu	Gln	Asp	Gly
			645						650					655	
Ser	Trp	Lys	Glu	Glu	Glu	Phe	Thr	Gly	Thr	Gly	Phe	Pro	Gly	Tyr	Gly
			660						665					670	
Val	Gly	Gln	Thr	Ile	Lys	Leu	Asp	Asp	Pro	Ala	Leu	Ser	Lys	Arg	Leu
		675					680							685	
Leu	Gln	Gly	Ala	Glu	Leu	Ser	Arg	Ala	Phe	Met	Leu	Arg	Tyr	Asp	Phe
		690					695				700				
Tyr	Arg	Gln	Phe	Phe	Pro	Ile	Met	Ala	Leu	Ser	Arg	Ala	Glu	Arg	Leu
705					710					715					720
Ile	Asp	Leu	Asn	Asn											
			725												

<210> 72

<211> 2178

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<220>

<221> misc_feature

<222> (544) .. (546)

<223> ntn是ttr或ctn

<220>

<221> misc_feature

<222> (553) .. (555)

<223> ngn是agr或cgn

<220>

<221> misc_feature
 <222> (1494) .. (1494)
 <223> n是a, c, g,或t
 <400> 72

```

atgggtattg acagaatgaa tagcttaagt cgcttggttaa tgaagaagat tttcggggct      60
gaaaaaacct cgtataaaccc ggcttccgat accataatcg gaacggatac cctgaaaaga      120
ccgaaccggc ggccctgaacc gacggcaaaa gtcgacaaaa cgatattcaa gactatgggg      180
aatagtctga ataataccct tgtttcagcc tgtgactggg tgatcggaca acaaaagccc      240
gatggtcatt gggtcgggtc cgtggaatcc aatgcttcga tggaagcaga atgggtgtctg      300
gccttggtgt ttttgggtct ggaagatcat ccgcttcgtc caagattggg caatgctctt      360
ttggaaatgc agcgggaaga tggtctcttg ggagtctatt tcggcgctgg aaatggcgat      420
atcaatgcca cggttgaagc ctatgcggcc ttgcggtctt tggggtatte tgccgataat      480
cctgttttga aaaaagcggc agcatggatt gctgaaaaag gcgattaaa aaatatccgt      540
gtcntnaccg gtngntggct ggcttgatc ggggaatggc cttgggaaaa gaccacctaac      600
cttccccctg aaattatctg gttccctgat aattttgtct tttcgattta taattttgcc      660
caatgggcgc gggcaaccat ggtgccgatt gctattctgt ccgcgagacg accaagccgc      720
ccgctgcgcc ctcaagaccg attggatgaa ctgtttccag aaggccgcgc tcgctttgat      780
tatgaattgc cgaaaaaaga aggcacgat ctttggtcgc aatttttccg aaccactgac      840
cgtggattac attgggttca gtccaatctg ttaaagcgca atagcttgcg tgaagccgct      900
atccgtcatg ttttggaatg gattatccgg catcaggatg ccgatggcgg ttgggggtgga      960
attcagccac cttgggtcta tggtttgatg gcgttacatg gtgaaggcta tcagctttat     1020
catccggtga tggccaaggc tttgtcggct ttggatgatc ccggttggcg acatgacaga     1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg     1140
atggcggttaa aagacgcca ggccaggat cgttttacgc cggaaatgga taaggccgcc     1200
gattggcttt tggctcgaca ggtcaaagtc aaaggcgatt ggtcaatcaa actgcccgat     1260
gttgaacccg gtggatgggc atttgaatat gccaatgatc gctatcccga taccgatgat     1320
accgccgtcg ctttgatcgc cttttcctct tatcgtgata aggaggagtg gcaaaagaaa     1380
ggcgttgagg acgccattac ccgtgggggtt aattggttga tcgccatgca aagcgaatgt     1440
ggcggttggg gagcctttga taaggataat aacagaagta tcctttccaa aacncctttt     1500
tgtgatttcg gagaatctat tgatccgcct tcagtcgatg taacggcgca tgttttagag     1560
gcctttggca ccttgggact gtcccgcgat atgccggtca tccaaaaagc gatcgactat     1620
gtccgttccg aacaggaagc cgaaggcgcg tggtttggtc gttggggcgt taattatatc     1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gaccagcctt     1740
tacatcacca aggcttgcca ttggtgggtc gcacatcagc aggaagacgg cggttggggc     1800
gaaagctgct cttcctatat ggagattgat tccattggga agggcccaac cacgccgtcc     1860
cagactgctt gggctttgat ggggttgatc gcggccaatc gtcccgaaga ttatgaagcc     1920
attgccaagg gatgccatta tctgattgat cgccaagagc aggatggtag ctggaagaa     1980
gaagaattca ccggcaccgg attccccggt tatggcgtgg gtcagacgat caagttggat     2040
gatccggctt tatcgaaacg attgcttcaa ggcgtgaac tgtcacgggc gtttatgctg     2100

```

```

cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178
<210> 73
<211> 725
<212> PRT
<213> 人工序列
<220>
<223> 合成ZmoSHC1衍生物
<400> 73
Met Gly Ile Asp Arg Met Asn Ser Leu Ser Arg Leu Leu Met Lys Lys
1 5 10 15
Ile Phe Gly Ala Glu Lys Thr Ser Tyr Lys Pro Ala Ser Asp Thr Ile
20 25 30
Ile Gly Thr Asp Thr Leu Lys Arg Pro Asn Arg Arg Pro Glu Pro Thr
35 40 45
Ala Lys Val Asp Lys Thr Ile Phe Lys Thr Met Gly Asn Ser Leu Asn
50 55 60
Asn Thr Leu Val Ser Ala Cys Asp Trp Leu Ile Gly Gln Gln Lys Pro
65 70 75 80
Asp Gly His Trp Val Gly Ala Val Glu Ser Asn Ala Ser Met Glu Ala
85 90 95
Glu Trp Cys Leu Ala Leu Trp Phe Leu Gly Leu Glu Asp His Pro Leu
100 105 110
Arg Pro Arg Leu Gly Asn Ala Leu Leu Glu Met Gln Arg Glu Asp Gly
115 120 125
Ser Trp Gly Val Tyr Phe Gly Ala Gly Asn Gly Asp Ile Asn Ala Thr
130 135 140
Val Glu Ala Tyr Ala Ala Leu Arg Ser Leu Gly Tyr Ser Ala Asp Asn
145 150 155 160
Pro Val Leu Lys Lys Ala Ala Ala Trp Ile Ala Glu Lys Gly Gly Leu
165 170 175
Lys Asn Ile Arg Val Phe Thr Arg Arg Trp Leu Ala Leu Ile Gly Glu
180 185 190
Trp Pro Trp Glu Lys Thr Pro Asn Leu Pro Pro Glu Ile Ile Trp Phe
195 200 205
Pro Asp Asn Phe Val Phe Ser Ile Tyr Asn Phe Ala Gln Trp Ala Arg
210 215 220
Ala Thr Met Val Pro Ile Ala Ile Leu Ser Ala Arg Arg Pro Ser Arg
225 230 235 240

```

Pro	Leu	Arg	Pro	Gln	Asp	Arg	Leu	Asp	Glu	Leu	Phe	Pro	Glu	Gly	Arg			
																245	250	255
Ala	Arg	Phe	Asp	Tyr	Glu	Leu	Pro	Lys	Lys	Glu	Gly	Ile	Asp	Leu	Trp			
																260	265	270
Ser	Gln	Phe	Phe	Arg	Thr	Thr	Asp	Arg	Gly	Leu	His	Trp	Val	Gln	Ser			
																275	280	285
Asn	Leu	Leu	Lys	Arg	Asn	Ser	Leu	Arg	Glu	Ala	Ala	Ile	Arg	His	Val			
																290	295	300
Leu	Glu	Trp	Ile	Ile	Arg	His	Gln	Asp	Ala	Asp	Gly	Gly	Trp	Gly	Gly			
305																310	315	320
Ile	Gln	Pro	Pro	Trp	Val	Tyr	Gly	Leu	Met	Ala	Leu	His	Gly	Glu	Gly			
																325	330	335
Tyr	Gln	Leu	Tyr	His	Pro	Val	Met	Ala	Lys	Ala	Leu	Ser	Ala	Leu	Asp			
																340	345	350
Asp	Pro	Gly	Trp	Arg	His	Asp	Arg	Gly	Glu	Ser	Ser	Trp	Ile	Gln	Ala			
																355	360	365
Thr	Asn	Ser	Pro	Val	Trp	Asp	Thr	Met	Leu	Ala	Leu	Met	Ala	Leu	Lys			
																370	375	380
Asp	Ala	Lys	Ala	Glu	Asp	Arg	Phe	Thr	Pro	Glu	Met	Asp	Lys	Ala	Ala			
385																390	395	400
Asp	Trp	Leu	Leu	Ala	Arg	Gln	Val	Lys	Val	Lys	Gly	Asp	Trp	Ser	Ile			
																405	410	415
Lys	Leu	Pro	Asp	Val	Glu	Pro	Gly	Gly	Trp	Ala	Phe	Glu	Tyr	Ala	Asn			
																420	425	430
Asp	Arg	Tyr	Pro	Asp	Thr	Asp	Asp	Thr	Ala	Val	Ala	Leu	Ile	Ala	Leu			
																435	440	445
Ser	Ser	Tyr	Arg	Asp	Lys	Glu	Glu	Trp	Gln	Lys	Lys	Gly	Val	Glu	Asp			
																450	455	460
Ala	Ile	Thr	Arg	Gly	Val	Asn	Trp	Leu	Ile	Ala	Met	Gln	Ser	Glu	Cys			
465																470	475	480
Gly	Gly	Trp	Gly	Ala	Phe	Asp	Lys	Asp	Asn	Asn	Arg	Ser	Ile	Leu	Ser			
																485	490	495
Lys	Thr	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
Ile Asp Ser Ile Gly Lys Gly Pro Thr Thr Pro Ser Gln Thr Ala Trp			
	610	615	620
Ala Leu Met Gly Leu Ile Ala Ala Asn Arg Pro Glu Asp Tyr Glu Ala			
625	630	635	640
Ile Ala Lys Gly Cys His Tyr Leu Ile Asp Arg Gln Glu Gln Asp Gly			
	645	650	655
Ser Trp Lys Glu Glu Glu Phe Thr Gly Thr Gly Tyr Pro Gly Tyr Gly			
	660	665	670
Val Gly Gln Thr Ile Lys Leu Asp Asp Pro Ala Leu Ser Lys Arg Leu			
	675	680	685
Leu Gln Gly Ala Glu Leu Ser Arg Ala Phe Met Leu Arg Tyr Asp Phe			
	690	695	700
Tyr Arg Gln Phe Phe Pro Ile Met Ala Leu Ser Arg Ala Glu Arg Leu			
705	710	715	720
Ile Asp Leu Asn Asn			
	725		

<210> 74

<211> 2178

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<220>

<221> misc_feature

<222> (553) .. (555)

<223> ngn是agr或cgn

<220>

<221> misc_feature

<222> (1494) .. (1494)

<223> n是a, c, g,或t

<400> 74

atgggtattg acagaatgaa tagcttaagt cgcttgtaa tgaagaagat tttcggggct

60

gaaaaaacct	cgtataaacc	ggcttccgat	accataatcg	gaacggatac	cctgaaaaga	120
ccgaaccggc	ggcctgaacc	gacggcaaaa	gtcgacaaaa	cgatattcaa	gactatgggg	180
aatagtctga	ataataccct	tgtttcagcc	tgtgactggt	tgatcggaca	acaaaagccc	240
gatggtcatt	gggtcgggtc	cgtggaatcc	aatgcttcga	tggaagcaga	atggtgtctg	300
gccttgtggt	ttttgggtct	ggaagatcat	ccgcttcgtc	caagattggg	caatgctctt	360
ttggaaatgc	agcgggaaga	tggctcttgg	ggagtctatt	tcggcgctgg	aaatggcgat	420
atcaatgcca	cggttgaagc	ctatgcggcc	ttgcggtctt	tggggtattc	tgccgataat	480
cctgttttga	aaaaagcggc	agcatggatt	gctgaaaaag	gcggtattaa	aaatatccgt	540
gtctttaccc	gtngntggct	ggcgttgatc	ggggaatggc	cttgggaaaa	gaccctaac	600
cttccccctg	aaattatctg	gttcctgat	aattttgtct	tttcgattta	taattttgcc	660
caatgggcgc	gggcaaccat	ggtgccgatt	gctattctgt	ccgcgagacg	accaagccgc	720
ccgctgcgcc	ctcaagaccg	attggatgaa	ctgtttccag	aaggccgcgc	tcgctttgat	780
tatgaattgc	cgaaaaaaga	aggcatcgat	ctttggtcgc	aatttttccg	aaccactgac	840
cgtggattac	attgggttca	gtccaatctg	ttaaagcgca	atagcttgcg	tgaagccgct	900
atccgtcatg	ttttggaatg	gattatccgg	catcaggatg	ccgatggcgg	ttgggggtgga	960
attcagccac	cttgggtcta	tggtttgatg	gcgttacatg	gtgaaggcta	tcagctttat	1020
catccggtga	tggccaaggc	tttgtcggct	ttggatgatc	ccggttggcg	acatgacaga	1080
ggcgagtctt	cttggataca	ggccaccaat	agtccggtat	gggatacaat	gttggccttg	1140
atggcggttaa	aagacgccaa	ggccgaggat	cgttttacgc	cggaaatgga	taaggccgcc	1200
gattggcttt	tggctcgaca	ggtcaaagtc	aaaggcgatt	ggtcaatcaa	actgcccgat	1260
gttgaaccgc	gtggatgggc	atttgaatat	gccaatgatc	gctatcccga	taccgatgat	1320
accgccgtcg	ctttgatcgc	cctttcctct	tatcgtgata	aggaggagtg	gcaaaagaaa	1380
ggcgttgagg	acgccattac	ccgtgggggt	aattggttga	tcgccatgca	aagcgaatgt	1440
ggcggttggg	gagcctttga	taaggataat	aacagaagta	tcctttccaa	aacncctttt	1500
tgtgatttgc	gagaatctat	tgatccgcct	tcagtcgatg	taacggcgca	tgttttagag	1560
gcctttggca	ccttgggact	gtcccgcgat	atgccggtca	tccaaaaagc	gatcgactat	1620
gtccgttccg	aacaggaagc	cgaaggcgcg	tggtttggtc	gttggggcgt	taattatatc	1680
tatggcaccg	gtgcggttct	gcctgctttg	gcggcgatcg	gtgaagatat	gaccagcct	1740
tacatcacca	aggcttgcca	ttggctggtc	gcacatcagc	aggaagacgg	cggttggggc	1800
gaaagctgct	cttcctatat	ggagattgat	tccattggga	agggcccaac	cacgccgtcc	1860
cagactgctt	gggctttgat	ggggttgatc	gcggccaatc	gtcccgaaga	ttatgaagcc	1920
attgccaagg	gatgccatta	tctgattgat	cgccaagagc	aggatggtag	ctggaaagaa	1980
gaagaattca	ccggcaccgg	ataycccgg	tatggcgtgg	gtcagacgat	caagttggat	2040
gatccggtct	tatcgaaacg	attgcttcaa	ggcgtgaac	tgtcacgggc	gtttatgctg	2100
cgttatgatt	tttatcggca	attcttcccg	attatggcgt	taagtcgggc	agagagactg	2160
attgatttga	ataattga					2178
<210>	75					
<211>	725					
<212>	PRT					

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 75

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

275	280	285
Asn Leu Leu Lys Arg Asn Ser	Leu Arg Glu Ala Ala	Ile Arg His Val
290	295	300
Leu Glu Trp Ile Ile Arg His	Gln Asp Ala Asp Gly Gly	Trp Gly Gly
305	310	315
Ile Gln Pro Pro Trp Val Tyr	Gly Leu Met Ala Leu His	Gly Glu Gly
325	330	335
Tyr Gln Leu Tyr His Pro Val	Met Ala Lys Ala Leu Ser	Ala Leu Asp
340	345	350
Asp Pro Gly Trp Arg His Asp	Arg Gly Glu Ser Ser Trp	Ile Gln Ala
355	360	365
Thr Asn Ser Pro Val Trp Asp	Thr Met Leu Ala Leu Met	Ala Leu Lys
370	375	380
Asp Ala Lys Ala Glu Asp Arg	Phe Thr Pro Glu Met Asp	Lys Ala Ala
385	390	395
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
Gly Gly Trp Gly Ala Phe Asp	Lys Asp Asn Asn Arg Ser	Ile Leu Ser
485	490	495
Lys Thr 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
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
 Ile Asp Ser Ile Gly Lys Gly Pro Thr Thr Pro Ser Gln Thr Ala Trp
 610 615 620
 Ala Leu Met Gly Leu Ile Ala Ala Asn Arg Pro Glu Asp Tyr Glu Ala
 625 630 635 640
 Ile Ala Lys Gly Cys His Tyr Leu Ile Asp Arg Gln Glu Gln Asp Gly
 645 650 655
 Ser Trp Lys Glu Glu Glu Phe Thr Gly Thr Gly Tyr Pro Gly Tyr Gly
 660 665 670
 Val Gly Gln Thr Ile Lys Leu Asp Asp Pro Ala Leu Ser Lys Arg Leu
 675 680 685
 Leu Gln Gly Ala Glu Leu Ser Arg Ala Phe Met Leu Arg Tyr Asp Phe
 690 695 700
 Tyr Arg Gln Phe Phe Pro Ile Met Ala Leu Ser Arg Ala Glu Arg Leu
 705 710 715 720
 Ile Asp Leu Asn Asn
 725

<210> 76

<211> 2178

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<220>

<221> misc_feature

<222> (544) .. (546)

<223> ntn是ttr或ctn

<220>

<221> misc_feature

<222> (553) .. (555)

<223> ngn是agr或cgn

<220>

<221> misc_feature

<222> (1494) .. (1494)

<223> n是a, c, g,或t

<400> 76

atgggtattg acagaatgaa tagcttaagt cgcttggttaa tgaagaagat tttcggggct 60
 gaaaaaacct cgtataaacc ggcttccgat accataatcg gaacgatac cctgaaaaga 120

ccgaaccggc	ggcctgaacc	gacggcaaaa	gtcgacaaaa	cgatattcaa	gactatgggg	180
aatagtctga	ataataccct	tgtttcagcc	tgtgactggt	tgatcggaca	acaaaagccc	240
gatggtcatt	gggtcggcgc	cgtggaatcc	aatgcttcga	tggaagcaga	atggtgtctg	300
gccttgtggc	ttttgggtct	ggaagatcat	ccgcttcgtc	caagattggg	caatgctctt	360
ttggaaatgc	agcgggaaga	tggctcttgg	ggagtctatt	tcggcgctgg	aaatggcgat	420
atcaatgcca	cggttgaagc	ctatgcggcc	ttgcggctct	tggggatttc	tgccgataat	480
cctgttttga	aaaaagcggc	agcatggatt	gctgaaaaag	gcggattaaa	aaatatccgt	540
gtcntnacc	gtngntggct	ggcgttgatc	ggggaatggc	cttgggaaaa	gaccacctaac	600
cttccccctg	aaattatctg	gttcctgat	aattttgtct	tttcgattta	taattttgcc	660
caatgggcgc	gggcaaccat	ggtgccgatt	gctattctgt	ccgcgagacg	accaagccgc	720
ccgctgcgcc	ctcaagaccg	attggatgaa	ctgtttccag	aaggccgcgc	tcgctttgat	780
tatgaattgc	cgaaaaaaga	aggcatcgat	ctttggctgc	aatttttccg	aaccactgac	840
cgtggattac	attgggttca	gtccaatctg	ttaaagcgca	atagcttgcg	tgaagccgct	900
atccgtcatg	ttttggaatg	gattatccgg	catcaggatg	ccgatggcgg	ttgggggtgga	960
attcagccac	cttgggtcta	tggtttgatg	gcgttacatg	gtgaaggcta	tcagctttat	1020
catccggtga	tggccaaggc	tttgtcggct	ttggatgatc	ccggttggcg	acatgacaga	1080
ggcgagtctt	cttggataca	ggccaccaat	agtcgggtat	gggatacaat	gttggccttg	1140
atggcgtaa	aagacgcaa	ggccgaggat	cgttttacgc	cggaaatgga	taaggccgcc	1200
gattggcttt	tggctcgaca	ggtcaaagtc	aaaggcgatt	ggtcaatcaa	actgcccgat	1260
gttgaacccg	gtggatgggc	atttgaatat	gccaatgatc	gctatcccga	taccgatgat	1320
accgccgtcg	ctttgatcgc	cctttcctct	tatcgtgata	aggaggagtg	gcaaaaagaaa	1380
ggcgttgagg	acgccattac	ccgtgggggt	aattggttga	tcgcatgca	aagcgaatgt	1440
ggcggttggg	gagcctttga	taaggataat	aacagaagta	tcctttccaa	aacncctttt	1500
tgtgatttcg	gagaatctat	tgatccgcct	tcagtcgatg	taacggcgca	tgtttttagag	1560
gcctttggca	ccttgggact	gtcccgcgat	atgccggatc	tccaaaaagc	gatcgactat	1620
gtccgttccg	aacaggaagc	cgaaggcgcg	tggtttggtc	gttggggcgt	taattatata	1680
tatggcaccg	gtgcggttct	gcctgctttg	gcggcgatcg	gtgaagatat	gaccagcctt	1740
tacatcacca	aggcttgcga	ttggctggtc	gcacatcagc	aggaagacgg	cggttggggc	1800
gaaagctgct	cttcctatat	ggagattgat	tccattggga	agggcccaac	cacgccgtcc	1860
cagactgctt	gggctttgat	ggggttgatc	gcggccaatc	gtcccgaaga	ttatgaagcc	1920
attgccaagg	gatgccatta	tctgattgat	cgccaagagc	aggatggtag	ctggaaagaa	1980
gaagaattca	ccggcaccgg	ataycccggg	tatggcgctg	gtcagacgat	caagttggat	2040
gatccggctt	tatcgaaacg	attgcttcaa	ggcgtgaac	tgtcacgggc	gtttatgctg	2100
cgttatgatt	tttatcgcca	attcttcccg	attatggcgt	taagtcgggc	agagagactg	2160
attgatttga	ataattga					2178
<210>	77					
<211>	658					
<212>	PRT					
<213>	人工序列					

<220>

<223> 合成ZmoSHC2衍生物

<400> 77

```

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 Ala 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
Ala Arg Tyr Arg Asn Leu Lys Lys Ala Asn Gln Pro Ile Val His Tyr		
645	650	655

Gly Met

<210> 78

<211> 1977

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<220>

<221> misc_feature

<222> (255) .. (255)

<223> n是a, c, g,或t

<400> 78

atgactgtat	cgacttcctc	ggcttttcat	catagcccgt	tgtctgatga	tgttgagccg	60
attatccaaa	aggccaccg	tgccttgctt	gagaagcagc	agcaggatgg	ccattggggtt	120
tttgaattgg	aagccgatgc	aaccattccc	gctgaataca	tcctgttaaa	gcattatttg	180
ggtgaacccg	aagatttaga	aatagaggcc	aagataggtc	gctatttgcg	tcgtattcag	240
ggcgagcatg	gcgcntggtc	tttgttttat	ggtggtgatc	ttgatttgag	cgccacggtc	300
aaagcctatt	ttgccttgaa	aatgatcgga	gattctcctg	atgcgcctca	tatgcttcga	360
gccagaaatg	aaattttggc	acgggggtggg	gcgatgcgtg	ccaatgtctt	tacacgtatt	420
caattagctc	tgttcggggc	aatgtcatgg	gagcatgtcc	ctcaaagcc	cgtagagttg	480
atgttgatgc	cggaatgggt	tccggttcac	atcaataaaa	tggcctattg	ggcaagaacc	540
gttttagtcc	cgttattggt	tttacaggcg	ttaaagcctg	tcgcccgtaa	tcggcgcggt	600
atcttggttg	atgaattatt	tgtgccggat	gttttaccga	cccttcagga	aagcggtgac	660
cctatatggc	gtcgtttttt	ttcggcactt	gataaggtat	tgcataaagt	agaaccttat	720
tggccgaaaa	atatgcgcgc	gaaggtatt	catagctgtg	tccattttgt	gaccgagcgt	780
ttgaatgggtg	aagacggggt	gggtgctatt	tatccggcga	ttgccaatag	cgtcatgatg	840
tatgatgcct	tgggatatcc	cgaaaacccat	ccagaaagag	ccattgcccg	tcgggctgtc	900
gaaaaattga	tggtgttaga	tggcacggaa	gatcagggtg	ataaagaagt	ctactgtcag	960
ccttgtttat	ccccgatttg	ggataccgct	ttggttgccc	atgccatgtt	ggaagtcgga	1020
ggcgatgagg	ctgaaaaatc	ggctatttct	gccttgagct	ggttaaagcc	gcaacaaatt	1080
ttggatgtaa	agggcgattg	ggcatggcgg	cggcctgac	tcagaccggg	gggatgggcc	1140
tttcaatata	gaaatgacta	ttatccccgat	gtcgatgata	cggctgttgt	gactatggcg	1200

atggatcgag ccgcaaaatt gtcggatctt cacgatgatt ttgaggaatc taaagcgcgt 1260
gccatggaat ggaccattgg gatgcaaagc gataatggcg gttggggcgc tttcgatgcc 1320
aataacagct atacttatct gaataatatt ccctttgctg atcatggcgc gttacttgat 1380
ccgccaacgg tcgatgtctc ggcacgctgc gtttcaatga tggcgcaagc cggtatctcg 1440
attacagatc ccaaaatgaa agcggcagtt gattatcttc tgaaagagca agaagaggat 1500
ggtagctggg tcgggcgttg ggggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggccttg 1620
ctgaaaacta ttcaaaatga agatgggtggg tggggtgaaa attgcgatag ctatgccctt 1680
gattatagcg gatacgagcc gatggattcg acggcttccc aaacagcatg ggctttattg 1740
ggcttgatgg ctgttgggga agctaattcc gaggccgtga caaagggtat aaactggttg 1800
gcacaaaatc aggatgaaga aggattgtgg aaagaagatt attatagtgg cgggtggtttt 1860
ccccgtgttt tttatcttcg gtatcacggg tattccaaat attttctctt ttgggccttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 79

<211> 658

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 79

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> 80

<211> 1977

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 80

atgactgtat cgacttcctc ggtttttcat catagcccg	tgtctgatga tgttgagccg	60
attatccaaa aggccaccg	tgcttgcctt gagaagcagc agcaggatgg ccattgggtt	120
tttgaattgg aagccgatgc aaccattccc	gctgaataca tcctgttaaa gcattatttg	180
ggtgaaccg aagatttaga aatagaggcc aagataggtc	gctatttgcg tcgtattcag	240
ggcgagcatg gcggatggc tttgttttat ggtggtgatc	ttgatttgag cgccacggc	300
aaagcctatt ttgccttgaa aatgatcgga gattctcctg	atgcgcctca tatgcttcga	360
gccagaaatg aaattttggc acggggtggg	gcgatgcgtg ccaatgtctt tacacgtatt	420


```

caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaatgcc cgtagagttg      480
atgttgatgc cggaatggtt tccggttcac atcaataaaa tggcctattg ggcaagaacc      540
gttttagtcc cgttattggt ttacaggcgc ttaaagcctg tcgcccgtaa tcggcgcggt      600
atcttggttg atgaattatt tgtgccggat gttttaccga cccttcagga aagcggtgac      660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat      720
tgcccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt      780
ttgaatggtg aagacggggt gggtgctatt tatccggcga ttgccaatag cgtcatgatg      840
tatgatgcct tgggatatcc cgaaaacat ccagaaagag ccattgcccg tcgggctgtc      900
gaaaaattga tgggtgtaga tggcacggaa gatcagggtg ataaagaagt ctactgtcag      960
ccttgtttat ccccgatttg ggataccgct ttggttgccc atgcatggtt ggaagtcgga     1020
ggcgtatgagg ctgaaaaatc ggctatttct gccttgagct ggttaaagcc gcaacaaatt     1080
ttggatgtaa agggcgattg ggcatggcgg cggcctgacg tcagaccggg gggatgggccc     1140
tttcaatata gaaatgacta ttatcccgat gtcgatgata cggctgttgt gactatggcg     1200
atggatcgag ccgcaaaatt gtcggatctt cacgatgatt ttgaggaatc taaagcgcgt     1260
gccatggaat ggaccattgg gatgcaaagc gataatggcg gttggggcgc tttcgatgcc     1320
aataacagct atacttatct gaataatatt ccctttgctg atcatggcgc gttacttgat     1380
ccgccaacgg tcgatgtctc ggcacgctgc gtttcaatga tggcgcaagc cggtatctcg     1440
attacagatc ccaaaatgaa agcggcagtt gattatcttc tgaaagagca agaagaggat     1500
ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt     1560
gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggccttg     1620
ctgaaaacta ttcaaaatga agatggtggt tggggtgaaa attgcgatag ctatgccctt     1680
gattatagcg gatacgagcc gatggattcg acggcttccc aaacagcatg ggctttattg     1740
ggcttgatgg ctgttgggga agctaattcc gaggccgtga caaagggtat aaactggttg     1800
gcacaaaatc aggatgaaga aggattgtgg aaagaagatt attatagtgg cggtggtttt     1860
ccccgtgttt ttatcttcg gtatcacggg tattccaaat attttctctt ttgggcttta     1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa      1977

```

<210> 81

<211> 658

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 81

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	Leu	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
Met Asp Arg Ala Ala Lys Leu Ser Asp Leu His Asp Asp Phe Glu Glu		400
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
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
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
Ala Arg Tyr Arg Asn Leu Lys Lys Ala Asn Gln Pro Ile Val His Tyr		
645	650	655
Gly Met		

<210> 82

- <211> 1977
 <212> DNA
 <213> 人工序列
 <220>
 <223> 合成ZmoSHC2衍生物
 <220>
 <221> misc_feature
 <222> (409) .. (411)
 <223> ntn是ttr或ctn
 <400> 82

```

atgactgtat cgacttcctc ggcttttcat catagcccg tgtctgatga tgttgagccg      60
attatccaaa aggccacccg tgccttgctt gagaagcagc agcaggatgg ccattggggtt    120
tttgaattgg aagccgatgc aaccattccc gctgaataca tcctgtttaa gcattatttg    180
ggtgaacccg aagatttaga aatagaggcc aagataggtc gctatttgcg tcgtattcag    240
ggcgagcatg gcggatggtc tttgttttat ggtggtgatc ttgatttgag cgccacggtc    300
aaagcctatt ttgccttgaa aatgatcgga gattctcctg atgcgcctca tatgcttcga    360
gccagaaatg aaattttggc acgggggtggg gcgatgcgtg ccaatgtcnt nacacgtatt    420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaagcc cgtagagttg    480
atgttgatgc cggaatgggt tccggttcac atcaataaaa tggcctattg ggcaagaacc    540
gttttagtcc cgttattggg tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt    600
atcttggttg atgaattatt tgtgccgatg gttttaccga cccttcagga aagcgggtgac    660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat    720
tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt    780
ttgaatgggt aagacggggt gggtgctatt tatccggcga ttgccaatag cgtcatgatg    840
tatgatgcct tgggatatcc cgaaaacatc ccagaaagag ccattgcccg tcgggctgtc    900
gaaaaattga tgggtgttaga tggcacggaa gatcagggtg ataaagaagt ctactgtcag    960
ccttgtttat ccccgatttg ggataccgct ttggttgccc atgccatgtt ggaagtgcga   1020
ggcgatgagg ctgaaaaatc ggctatttct gccttgagct ggttaaagcc gcaacaaatt   1080
ttggatgtaa agggcgattg ggcatggcgg cggcctgacg tcagaccggg gggatgggccc   1140
tttcaatata gaaatgacta ttatcccgat gtcgatgata cggctgttgt gactatggcg   1200
atggatcgag ccgcaaaaatt gtcggatctt cacgatgatt ttgaggaatc taaagcgcgt   1260
gccatggaat ggaccattgg gatgcaaagc gataatggcg gttggggcgc tttcgatgcc   1320
aataacagct atacttatct gaataatatt ccttttgctg atcatggcgc gttacttgat   1380
ccgccaacgg tcgatgtctc ggcacgtctc gtttcaatga tggcgcaagc cggatatctcg   1440
attacagatc ccaaaaatga agcggcagtt gattatcttc tgaaagagca agaagaggat   1500
ggtagctggg tcgggcgttg ggggtgtcaat tacatatatg gcacatggtc ggccttatgt   1560
gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggccttg   1620
ctgaaaacta ttcaaaaatga agatggtggg tggggtgaaa attgcgatag ctatgccctt   1680
gattatagcg gatacgagcc gatggattcg acggcttccc aaacagcatg ggctttattg   1740

```

ggcttgatgg ctgttgggga agctaattcc gaggccgtga caaagggtat aaactggttg 1800
gcacaaaatc aggatgaaga aggattgtgg aaagaagatt attatagtgg cggtaggtttt 1860
ccccgtgttt tttatcttcg gtatcacggg tattccaaat attttcctct ttgggcttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 83

<211> 658

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 83

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 Arg 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
Asp Tyr Ser Gly Tyr Glu Pro Met Asp Ser Thr Ala Ser Gln Thr Ala		560
	565	570
Trp Ala Leu Leu Gly Leu Met Ala Val Gly Glu Ala Asn Ser Glu Ala		575
	580	585
Val Thr Lys Gly Ile Asn Trp Leu Ala Gln Asn Gln Asp Glu Glu Gly		590
	595	600
Leu Trp Lys Glu Asp Tyr Tyr Ser Gly Gly Gly Phe Pro Arg Val Phe		605
	610	615
Tyr Leu Arg Tyr His Gly Tyr Ser Lys Tyr Phe Pro Leu Trp Ala Leu		620
625	630	635
Ala Arg Tyr Arg Asn Leu Lys Lys Ala Asn Gln Pro Ile Val His Tyr		640
	645	650
		655

Gly Met

<210> 84

<211> 1977

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<220>

<221> misc_feature

<222> (418) .. (420)

<223> ngn是agr或cgn

<400> 84

```

atgactgtat cgacttcctc ggcttttcat catagcccgt tgtctgatga tgttgagccg      60
attatccaaa aggccacccg tgccttgctt gagaagcagc agcaggatgg ccattggggtt      120
tttgaattgg aagccgatgc aaccattccc gctgaataca tcctgttaaa gcattatttg      180
ggtgaacccg aagatttaga aatagaggcc aagataggct gctatttgcg tcgtattcag      240
ggcgagcatg gcggatggtc tttgttttat ggtggtgata ttgatttgag cgccacggtc      300
aaagcctatt ttgccttgaa aatgatcgga gattctctg atgcgcctca tatgcttcga      360
gccagaaatg aaattttggc acgggggtggg gcgatgcgtg ccaatgtctt tacacgtngn      420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaatgcc cgtagagttg      480
atgttgatgc cggaatgggt tccggttcac atcaataaaa tggcctattg ggcaagaacc      540
gttttagtcc cgttatttgt ttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt      600
atcttggttg atgaattatt tgtgccgat gttttaccga cccttcagga aagcgggtgac      660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat      720

```

```

tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt      780
ttgaatggtg aagacggggtt gggtgctatt tatccggcga ttgccaatag cgatcatgat      840
tatgatgcct tgggatatcc cgaaaaccat ccagaaagag ccattgcccc tcgggctgtc      900
gaaaaattga tgggtgftaga tggcacggaa gatcagggtg ataaagaagt ctactgtcag      960
ccttgtttat ccccgatttg ggataccgct ttggttgccc atgccatgtt ggaagtcgga     1020
ggcgatgagg ctgaaaaatc ggctatttct gccttgagct ggttaaagcc gcaacaaatt     1080
ttggatgtaa agggcgattg ggcatggcgg cggcctgata tcagaccggg gggatgggcc     1140
tttcaatata gaaatgacta ttatcccgat gtcgatgata cggctgttgt gactatggcg     1200
atggatcgag ccgcaaaaatt gtcggatctt cacgatgatt ttgaggaatc taaagcgcgt     1260
gccatggaat ggaccatttg gatgcaaagc gataatggcg gttggggcgc tttcgatgcc     1320
aataacagct atacttatct gaataatatt ccctttgctg atcatggcgc gttacttgat     1380
ccgccaacgg tcgatgtctc ggcacgctgc gtttcaatga tggcgcaagc cggatatctcg     1440
attacagatc ccaaaaatgaa agcggcagtt gattatcttc tgaaagagca agaagaggat     1500
ggtagctggg tcgggcgttg ggggtgtcaat tacatatatg gcacatggtc ggccttatgt     1560
gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggccttg     1620
ctgaaaacta ttcaaaatga agatggtggg tggggtgaaa attgcatag ctatgccctt     1680
gattatagcg gatacgagcc gatggattcg acggcttccc aaacagcatg ggctttattg     1740
ggcttgatgg ctgttgggga agctaattcc gaggccgtga caaagggtat aaactggttg     1800
gcacaaaatc aggatgaaga aggattgtgg aaagaagatt attatagtgg cgggtggttt     1860
ccccgtgttt tttatcttcg gtatcacggg tattccaaat attttctctt ttgggcttta     1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa      1977

```

<210> 85

<211> 658

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 85

```

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
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
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
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

Gly Met

〈223〉 合成ZmoSHC2衍生物

<400> 86

```

atgactgtat cgacttcctc ggcttttcat catagcccggt tgtctgatga tgttgagccg      60
attatccaaa aggccacccg tgccttgctt gagaagcagc agcaggatgg ccattggggtt      120
tttgaattgg aagccgatgc aaccattccc gctgaataca tcctgttaaa gcattatttg      180
ggtgaacccg aagatttaga aatagaggcc aagataggct gctatttgcg tcgtattcag      240
ggcgagcatg gcggatggtc tttgttttat ggtgggtgatc ttgatttgag cgccacggtc      300
aaagcctatt ttgccttgaa aatgatcgga gattctcctg atgcgcctca tatgcttcga      360
gccagaaatg aaattttggc acgggggtggg gcgatgcgtg ccaatgtctt tacacgtatt      420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaagcc cgtagagtgg      480
atgttgatgc cggaatgggt tccggttcac atcaataaaa tggcctattg ggcaagaacc      540
gttttagtcc cgttattggg tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt      600
atcttggttg atgaattatt tgtgccggat gttttaccga cccttcagga aagcgggtgac      660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat      720
tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt      780
ttgaatgggtg aagacggggt gggtgctatt tatccggcga ttgccaatag cgtcatgatg      840
tatgatgcct tgggatatcc cgaaaacat ccagaaagag ccattgcccg tcgggctgtc      900
gaaaaattga tgggtgttaga tggcacggaa gatcagggtg ataaagaagt ctactgtcag      960
ccttgtttat ccccgatttg ggataccgct ttggttgccc atgcatgtt ggaagtcgga      1020
ggcgatgagg ctgaaaaatc ggctatttct gccttgagct ggttaaagcc gcaacaaatt      1080
ttggatgtaa agggcgattg ggcatggcgg cggcctgatc tcagaccggg gggatgggcc      1140
tttcaatata gaaatgacta ttatcccgat gtcgatgata cggctgttgt gactatggcg      1200
atggatcgag ccgcaaaatt gtcggatctt cacgatgatt ttgaggaatc taaagcgcgt      1260
gccatggaat ggaccattgg gatgcaaagc gataatggcg gttggggcgc tttcgatgcc      1320
aataacagct atacttatct gaataatatt ccctttgctg atcatggcgc gttacttgat      1380
ccgccaacgg tcgatgtctc ggcacgctgc gtttcaatga tggcgcaagc cggtatctcg      1440
attacagatc ccaaaatgaa agcggcagtt gattatcttc tgaaagagca agaagaggat      1500
ggtagctggg tcgggcgttg ggggtgtcaat tacatatatg gcacatggtc ggccttatgt      1560
gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggccttg      1620
ctgaaaacta ttcaaaatga agatggtggg tggggtgaaa attgcgatag ctatgccctt      1680
gattatagcg gatacgagcc gatggattcg acggcttccc aaacagcatg ggctttattg      1740
ggcttgatgg ctgttgggga agctaattcc gaggccgtga caaagggtat aaactggttg      1800
gcacaaaatc aggatgaaga aggattgtgg aaagaagatt attatagtgg cgggtggttt      1860
ccccgtgttt tttatcttcg gtatcacggg tattccaaat attttctctt ttgggcttta      1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa      1977

```

<210> 87

<211> 658

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 87

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
Pro Cys Leu Ser Pro Ile Trp Asp Thr Ala Leu Val Ala His Ala Met		320
	325	330
Leu Glu Val Gly Gly Asp Glu Ala Glu Lys Ser Ala Ile Ser Ala Leu		335
	340	345
Ser Trp Leu Lys Pro Gln Gln Ile Leu Asp Val Lys Gly Asp Trp Ala		350
	355	360
Trp Arg Arg Pro Asp Leu Arg Pro Gly Gly Trp Ala Phe Gln Tyr Arg		365
370	375	380
Asn Asp Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Thr Met Ala		
385	390	395
Met Asp Arg Ala Ala Lys Leu Ser Asp Leu His Asp Asp Phe Glu Glu		400
	405	410
Ser Lys Ala Arg Ala Met Glu Trp Thr Ile Gly Met Gln Ser Asp Asn		415
	420	425
Gly Gly Trp Gly Ala Phe Asp Ala Asn Asn Ser Tyr Thr Tyr Leu Asn		430
435	440	445
Asn Thr 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
Ile Thr Asp Pro Lys Met Lys Ala Ala Val Asp Tyr Leu Leu Lys Glu		480
	485	490
Gln Glu Glu Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Ile		495
	500	505
Tyr Gly Thr Trp Ser Ala Leu Cys Ala Leu Asn Val Ala Ala Leu Pro		510
	515	520
His Asp His Leu Ala Val Gln Lys Ala Val Ala Trp Leu Lys Thr Ile		525
	530	535
Gln Asn Glu Asp Gly Gly Trp Gly Glu Asn Cys Asp Ser Tyr Ala Leu		540
545	550	555
Asp Tyr Ser Gly Tyr Glu Pro Met Asp Ser Thr Ala Ser Gln Thr Ala		560
	565	570
Trp Ala Leu Leu Gly Leu Met Ala Val Gly Glu Ala Asn Ser Glu Ala		575
	580	585
Val Thr Lys Gly Ile Asn Trp Leu Ala Gln Asn Gln Asp Glu Glu Gly		590
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> 88

<211> 1977

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<220>

<221> misc_feature

<222> (1350) .. (1350)

<223> n是a, c, g,或t

<400> 88

```

atgactgtat cgacttcctc ggcttttcat catagcccggt tgtctgatga tgttgagccg      60
attatccaaa aggccacccg tgccttgctt gagaagcagc agcaggatgg ccattggggtt      120
tttgaattgg aagccgatgc aaccattccc gctgaataca tcctgttaaa gcattatttg      180
ggtgaacccg aagatttaga aatagaggcc aagataggct gctatttgcg tcgtattcag      240
ggcgagcatg gcggatggct tttgttttat ggtggtgatc ttgatttgag cgccacggctc      300
aaagcctatt ttgccttgaa aatgatcgga gattctcctg atgcgcctca tatgcttcga      360
gccagaaatg aaattttggc acgggggtggg gcgatgcgtg ccaatgtctt tacacgtatt      420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaagcc cgtagagttag      480
atgttgatgc cggaatgggt tccggttcac atcaataaaa tggcctattg ggcaagaacc      540
gttttagtcc cgttattggt tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt      600
atcttggttg atgaattatt tgtgccggat gttttaccga cccttcagga aagcgggtgac      660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat      720
tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt      780
ttgaatgggt aagacggggt ggggtgctatt tatccggcga ttgccaatag cgtcatgatg      840
tatgatgcct tgggatatcc cgaaaacat ccagaaagag ccattgcccc tcgggctgtc      900
gaaaaattga tgggtgtaga tggcacggaa gatcagggtg ataaagaagt ctactgtcag      960
ccttgtttat ccccgatttg ggataccgct ttggttgccc atgcatgtt ggaagtcgga      1020
ggcgatgagg ctgaaaaatc ggctatttct gccttgagct ggttaaagcc gcaacaaatt      1080
ttggatgtaa agggcgattg ggcatggcgg cggcctgatc tcagaccgga gggatgggcc      1140
tttcaatata gaaatgacta ttatcccgat gtcgatgata cggctgttgt gactatggcg      1200
atggatcgag ccgcaaaatt gtcggatctt cacgatgatt ttgaggaatc taaagcgcgt      1260

```

gccatggaat ggaccattgg gatgcaaagc gataatggcg gttggggcgc tttcgatgcc 1320
 aataacagct atacttatct gaataatacn ccctttgctg atcatggcgc gttacttgat 1380
 ccgccaacgg tcgatgtctc ggcacgctgc gtttcaatga tggcgcaagc cggtatctcg 1440
 attacagatc ccaaaatgaa agcggcagtt gattatcttc tgaaagagca agaagaggat 1500
 ggtagctggg tcgggcgttg ggggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
 gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggccttg 1620
 ctgaaaacta ttcaaaatga agatgggtgg tggggtgaaa attgcgatag ctatgccctt 1680
 gattatagcg gatacgagcc gatggattcg acggcttccc aaacagcatg ggctttattg 1740
 ggcttgatgg ctgttgggga agctaattcc gaggccgtga caaagggtat aaactggttg 1800
 gcacaaaatc aggatgaaga aggattgtgg aaagaagatt attatagtgg cggtggtttt 1860
 ccccggtttt tttatcttcg gtatcacggg tattccaaat attttctctt ttgggcctta 1920
 gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 89

<211> 658

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 89

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 His 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> 90

<211> 1977

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 90

atgactgtat cgacttcctc ggcttttcat catagcccg tgtctgatga tgttgagccg	60
attatccaaa aggccacccg tgccttgctt gagaagcagc agcaggatgg ccattggggtt	120
tttgaattgg aagccgatgc aaccattccc gctgaataca tcctgttaaa gcattatttg	180
ggtgaacccg aagatttaga aatagaggcc aagataggtc gctatttgcg tcgtattcag	240
ggcgagcatg gcggatggtc tttgttttat ggtggtgatc ttgatttgag cgccacggtc	300
aaagcctatt ttgccttgaa aatgatcgga gattctctg atgcgcctca tatgcttcga	360
gccagaaatg aaattttggc acgggggtggg gcgatgcgtg ccaatgtctt tacacgtatt	420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaatgcc cgtagagttg	480

atgttgatgc cggaatgggt tccggttcac atcaataaaa tggcctattg ggcaagaacc 540
 gtttttagtcc cggtatttgt ttacaggcgc ttaaagcctg tcgcccgtaa tcggcgcggg 600
 atcttggttg atgaattatt tgtgccggat gttttaccga cccttcagga aagcgggtgac 660
 cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat 720
 tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt 780
 ttgaatgggt aagacgggtt ggggtgctatt tatccggcga ttgccaatag cgtcatgatg 840
 tatgatgcct tgggatatcc cgaaaacat ccagaaagag ccattgcccg tcgggctgtc 900
 gaaaaattga tgggtgttaga tggcacggaa gatcagggtg ataaagaagt ctactgtcag 960
 ccttgtttat ccccgatttg ggataccgct ttggttgccc atgcatgtt ggaagtgcga 1020
 ggcgatgagg ctgaaaaatc ggctatttct gccttgagct ggttaaagcc gcaacaaatt 1080
 ttggatgtaa agggcgattg ggcatggcgg cggcctgac tcagaccggg gggatgggccc 1140
 tttcaatata gaaatgacta ttatcccgat gtcgatgata cggctgttgt gactatggcg 1200
 atggatcgag ccgcaaaaatt gtcggatctt cacgatgatt ttgaggaatc taaagcgcgt 1260
 gccatggaat ggaccatttg gatgcaaagc gataatggcg gttggggcgc tttcgatgcc 1320
 aataacagct atacttatct gaataatatt ccctttgctg atcatggcgc gttacttgat 1380
 ccgccaacgg tcgatgtctc ggcacgctgc gtttcaatga tggcgcaagc cggtatctcg 1440
 attacagatc ccaaaaatgaa agcggcagtt gattatcttc tgaaagagca agaagaggat 1500
 ggtagctggg tcgggcgttg ggggtgtcaat tacatatatg gcacatggc ggccttatgt 1560
 gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggccttg 1620
 ctgaaaacta ttcaaaatga agatgggtgg tggggtgaaa attgcgatag ctatgccctt 1680
 gattatagcg gatacgagcc gatggattcg acggcttccc aaacagcatg ggctttattg 1740
 ggcttgatgg ctgttgggga agctaattcc gaggccgtga caaagggtat acaytggttg 1800
 gcacaaaatc aggatgaaga aggattgtgg aaagaagatt attatagtgg cggtggtttt 1860
 ccccggtgtt tttatcttcg gtatcacggg tattccaaat attttcctct ttgggcttta 1920
 gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 91

<211> 658

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 91

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
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
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
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
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 Tyr 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>	92	
<211>	1977	

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 92

```

atgactgtat cgacttcctc ggcttttcat catagcccgt tgtctgatga tgttgagccg      60
attatccaaa aggccacccg tgccttgctt gagaagcagc agcaggatgg ccattggggtt    120
tttgaattgg aagccgatgc aaccattccc gctgaataca tcctgttaaa gcattatttg    180
ggtgaacccg aagatttaga aatagaggcc aagataggtc gctatttgcg tcgtattcag    240
ggcgagcatg gcggatggtc tttgttttat ggtggtgatc ttgatttgag cgccacggtc    300
aaagcctatt ttgccttgaa aatgatcgga gattctctg atgcgcctca tatgcttcga    360
gccagaaatg aaattttggc acgggggtggg gcgatgcgtg ccaatgtctt tacacgtatt    420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaatgcc cgtagagttg    480
atgttgatgc cggaatgggt tccggttcac atcaataaaa tggcctattg ggcaagaacc    540
gttttagtcc cgttatttgt tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt    600
atcttggttg atgaattatt tgtgccgatg gttttaccga cccttcagga aagcggtgac    660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat    720
tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt    780
ttgaatgggt aagacggggt gggtgctatt tatccggcga ttgccaatag cgtcatgatg    840
tatgatgcct tgggatatcc cgaaaacat ccagaaagag ccattgcccc tcgggctgtc    900
gaaaaattga tgggtgttaga tggcacggaa gatcagggtg ataaagaagt ctactgtcag    960
ccttgtttat ccccgatttg ggataccgct ttggttgccc atgcatgtt ggaagtcgga   1020
ggcgatgagg ctgaaaaatc ggctatttct gccttgagct ggttaaagcc gcaacaaatt   1080
ttggatgtaa agggcgattg ggcatggcgg cggcctgac tcagaccggg gggatgggcc   1140
tttcaatata gaaatgacta ttatcccgat gtcgatgata cggctgttgt gactatggcg   1200
atggatcgag ccgcaaaatt gtcggatctt cacgatgatt ttgaggaatc taaagcgcgt   1260
gccatggaat ggaccattgg gatgcaaagc gataatggcg gttggggcgc tttcgatgcc   1320
aataacagct atacttatct gaataatatt ccttttgctg atcatggcgc gttacttgat   1380
ccgccaacgg tcgatgtctc ggcacgctgc gtttcaatga tggcgcaagc cggatatctc   1440
attacagatc ccaaaatgaa agcggcagtt gattatcttc tgaaagagca agaagaggat   1500
ggtagctggg tcgggcgttg ggggtgtcaat tacatatatg gcacatggtc ggccttatgt   1560
gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggccttg   1620
ctgaaaacta ttcaaaatga agatggtggg tggggtgaaa attgcgatag ctatgccctt   1680
gattatagcg gatacgagcc gatggattcg acggttccc aaacagcatg ggctttattg   1740
ggcttgatgg ctgttgggga agctaattcc gaggccgtga caaagggtat aaactggttg   1800
gcacaaaatc aggatgaaga aggattgtgg aaagaagatt attatagtgg cgggtggttay   1860
ccccgtgttt tttatcttcg gtatcacggg tattccaaat attttctctt ttgggcttta   1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa     1977

```

<210> 93

<211> 658

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 93

```

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 Arg 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
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
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 Thr 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
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
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
Ala Arg Tyr Arg Asn Leu Lys Lys Ala Asn Gln Pro Ile Val His Tyr		
645	650	655

Gly Met

<210> 94

<211> 1977

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<220>

<221> misc_feature

<222> (418) .. (420)

<223> ngn是agr或cgn

<220>

<221> misc_feature

<222> (1350) .. (1350)

<223> n是a, c, g, 或t

<400> 94

atgactgtat cgacttcctc ggcttttcat catagcccgt tgtctgatga tgttgagccg	60
attatccaaa aggccacccg tgccttgctt gagaagcagc agcaggatgg ccattggggtt	120
tttgaattgg aagccgatgc aaccattccc gctgaataca tcctgttaaa gcattatttg	180
ggtgaacccg aagatttaga aatagaggcc aagataggct gctatttgcg tcgtattcag	240
ggcgagcatg gcggatggct tttgttttat ggtggtgatc ttgatttgag cgccacggctc	300
aaagcctatt ttgccttgaa aatgatcgga gattctctg atgcgcctca tatgcttcga	360
gccagaaatg aaattttggc acggggtggg gcgatgcgtg ccaatgtctt tacacgtngn	420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaagcc cgtagagttg	480
atgttgatgc cggaatggtt tccggttcac atcaataaaa tggcctattg ggcaagaacc	540
gttttagtcc cgttattggt tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt	600
atcttggttg atgaattatt tgtgccgat gttttaccga cccttcagga aagcggtgac	660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat	720
tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt	780


```

ttgaatggtg aagacgggtt ggggtgctatt tatccggcga ttgccaatag cgatcatgatg      840
tatgatgcct tgggatatcc cgaaaacat ccagaaagag ccattgcccg tgggctgtc      900
gaaaaattga tgggtgttaga tggcacggaa gatcagggtg ataaagaagt ctactgtcag      960
ccttgtttat ccccgatttg ggataccgct ttggttgccc atgcatgtt ggaagtcgga     1020
ggcgaatgagg ctgaaaaatc ggctatttct gccttgagct ggttaaagcc gcaacaaatt     1080
ttggatgtaa agggcgattg ggcattggcg cgccctgac tcagaccgg gggatgggc     1140
tttcaatata gaaatgacta ttatcccgat gtcgatgata cggctgttgt gactatggcg     1200
atggatcgag ccgcaaaaatt gtcggatctt cacgatgatt ttgaggaatc taaagcgcgt     1260
gccatggaat ggaccattgg gatgcaaagc gataatggcg gttggggcgc tttcgatgcc     1320
aataacagct atacttatct gaataatacn cctttgctg atcatggcg gttacttgat     1380
ccgccaacgg tcgatgtctc ggcacgtgc gtttcaatga tggcgcaagc cggtatctcg     1440
attacagatc ccaaaatgaa agcggcagtt gattatcttc tgaaagagca agaagaggat     1500
ggtagctggt tcgggcgttg ggggtgtcaat tacatatatg gcacatggc ggccttatgt     1560
gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggccttg     1620
ctgaaaacta ttcaaaatga agatggtggt tggggtgaaa attgcgatag ctatgccctt     1680
gattatagcg gatacgagcc gatggattcg acggcttccc aaacagcatg ggctttattg     1740
ggcttgatgg ctgttgggga agctaattcc gaggccgtga caaagggtat aaactggttg     1800
gcacaaaatc aggatgaaga aggattgtgg aaagaagatt attatagtgg cggtggtttt     1860
ccccgtgttt tttatcttcg gtatcacggt tattccaaat attttctctt ttgggcttta     1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa      1977

```

<210> 95

<211> 658

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 95

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	Arg	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	Thr	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> 96															
<211> 1977															
<212> DNA															
<213> 人工序列															
<220>															
<223> 合成ZmoSHC2衍生物															
<220>															

<221>	misc_feature					
<222>	(418) .. (420)					
<223>	ngn是agr或cgn					
<220>						
<221>	misc_feature					
<222>	(1350) .. (1350)					
<223>	n是a, c, g,或t					
<400>	96					
atgactgtat	cgacttcctc	ggcttttcat	catagcccgt	tgtctgatga	tgttgagccg	60
attatccaaa	aggccacccg	tgccttgctt	gagaagcagc	agcaggatgg	ccattggggtt	120
tttgaattgg	aagccgatgc	aaccattccc	gctgaataca	tcctgtttaa	gcattatttg	180
ggtgaacccg	aagatttaga	aatagaggcc	aagataggtc	gctatttgcg	tcgtattcag	240
ggcgagcatg	gcggatggtc	tttgttttat	ggtggtgatc	ttgatttgag	cgccacggtc	300
aaagcctatt	ttgccttgaa	aatgatcgga	gattctcctg	atgcgcctca	tatgcttcga	360
gccagaaatg	aaattttggc	acggggtggg	gcgatgcgtg	ccaatgtctt	tacacgtngn	420
caattagctc	tgttcggggc	aatgtcatgg	gagcatgtcc	ctcaaatgcc	cgtagagttg	480
atgttgatgc	cggaatgggt	tccggttcac	atcaataaaa	tggcctattg	ggcaagaacc	540
gttttagtcc	cgttattggg	tttacaggcg	ttaaagcctg	tcgcccgtaa	tcggcgcggt	600
atcttggttg	atgaattatt	tgtgccggat	gttttaccga	cccttcagga	aagcgggtgac	660
cctatatggc	gtcgtttttt	ttcggcactt	gataaggtat	tgcataaagt	agaaccttat	720
tggccgaaaa	atatgcgcgc	gaaggctatt	catagctgtg	tccattttgt	gaccgagcgt	780
ttgaatgggtg	aagacggggt	gggtgctatt	tatccggcga	ttgccaatag	cgatcatgatg	840
tatgatgcct	tgggatatcc	cgaaaaccat	ccagaaagag	ccattgcccc	tcgggctgtc	900
gaaaaattga	tgggtgttaga	tggcacggaa	gatcagggtg	ataaagaagt	ctactgtcag	960
ccttgtttat	ccccgatttg	ggataccgct	ttggttgccc	atgccatggt	ggaagtgcga	1020
ggcgatgagg	ctgaaaaatc	ggctatttct	gccttgagct	ggttaaagcc	gcaacaaatt	1080
ttggatgtaa	agggcgattg	ggcatggcgg	cggcctgata	tcagaccggg	gggatgggcc	1140
tttcaatata	gaaatgacta	ttatcccgat	gtcgatgata	cggctgttgt	gactatggcg	1200
atggatcgag	ccgcaaaaatt	gtcggatctt	cacgatgatt	ttgaggaatc	taaagcgcgt	1260
gccatggaat	ggaccattgg	gatgcaaagc	gataatggcg	gttggggcgc	tttcgatgcc	1320
aataacagct	atacttatct	gaataatacn	ccctttgctg	atcatggcgc	gttacttgat	1380
ccgccaacgg	tcgatgtctc	ggcacgtgcg	gtttcaatga	tggcgcaagc	cggtatctcg	1440
attacagatc	ccaaaaatga	agcggcagtt	gattatcttc	tgaaagagca	agaagaggat	1500
ggtagctggg	tcgggcgttg	gggtgtcaat	tacatatatg	gcacatggtc	ggccttatgt	1560
gcattgaatg	tggccgcttt	accccatgat	catttagctg	ttcagaaagc	tgtggccttg	1620
ctgaaaaacta	ttcaaaaatga	agatggtggg	tggggtgaaa	attgcgatag	ctatgccctt	1680
gattatagcg	gatacgagcc	gatggattcg	acggcttccc	aaacagcatg	ggctttattg	1740
ggcttgatgg	ctgttgggga	agctaattcc	gaggccgtga	caaagggtat	aaactggttg	1800
gcacaaaatc	aggatgaaga	aggattgtgg	aaagaagatt	attatagtgg	cggtggtttt	1860

```

ccccgtgttt tttatcttcg gtatcacggt tattccaaat attttcctct ttgggcttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977
<210> 97
<211> 658
<212> PRT
<213> 人工序列
<220>
<223> 合成ZmoSHC2衍生物
<400> 97
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 Tyr 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> 98

<211> 1977

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 98

```

atgactgtat cgacttcctc ggcttttcat catagcccgt tgtctgatga tgttgagccg      60
attatccaaa aggccaccg tgccttgctt gagaagcagc agcaggatgg ccattggggtt      120
tttgaattgg aagccgatgc aaccattccc gctgaataca tcctgttaaa gcattatttg      180
ggatgaacccg aagatttaga aatagaggcc aagataggtc gctatttgcg tcgtattcag      240
ggcgagcatg gcggatggtc tttgttttat ggtggatgat ttgatttgag cgccacggtc      300
aaagcctatt ttgccttgaa aatgatcgga gattctcctg atgcgcctca tatgcttcga      360
gccagaaatg aaattttggc acgggggtggg gcgatgcgtg ccaatgtctt tacacgtatt      420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaagcc cgtagagttg      480
atgttgatgc cggaatgggt tccggttcac atcaataaaa tggcctattg ggcaagaacc      540
gttttagtcc cgttattgggt tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt      600
atcttggttg atgaattatt tgtgccgat gttttaccga cccttcagga aagcggtgac      660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat      720
tggccgaaaa atatgcgcgc gaaggtatt catagctgtg tccattttgt gaccgagcgt      780
ttgaatgggt aagacgggtt ggggtctatt tatccggcga ttgccaatag cgtcatgatg      840
tatgatgcct tgggatatcc cgaaaacccat ccagaaagag ccattgcccg tcgggctgtc      900
gaaaaattga tgggtgtaga tggcacggaa gatcagggtg ataaagaagt ctactgtcag      960
ccttgtttat ccccgatttg ggataccgct ttggttgccc atgcatgtt ggaagtcgga     1020
ggcgatgagg ctgaaaaatc ggctatttct gccttgagct ggttaaagcc gcaacaaatt     1080

```

```

ttggatgtaa agggcgattg ggcattggcgg cggcctgata tcagacccgg gggatgggcc 1140
tttcaatata gaaatgacta ttatcccgat gtcatgata cggctgttgt gactatggcg 1200
atggatcgag ccgcaaaatt gtcggatctt cacgatgatt ttgaggaatc taaagcgcgt 1260
gccatggaat ggaccattgg gatgcaaagc gataatggcg gttggggcgc tttcatgcc 1320
aataacagct atacttatct gaataatatt ccctttgctg atcatggcgc gttacttgat 1380
ccgccaacgg tcgatgtctc ggcacgtgc gtttcaatga tggcgcaagc cggtatctcg 1440
attacagatc ccaaaaatgaa agcggcagtt gattatcttc tgaaagagca agaagaggat 1500
ggtagctggg tcgggcgttg ggggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggccttg 1620
ctgaaaacta ttcaaaatga agatggtggg tggggtgaaa attgcgatag ctatgccctt 1680
gattatagcg gatacgagcc gatggattcg acggttccc aaacagcatg ggctttattg 1740
ggcttgatgg ctgttgggga agctaattcc gaggccgtga caaagggtat aaactggttg 1800
gcacaaaatc aggatgaaga aggattgtgg aaagaagatt attatagtgg cggtggttay 1860
ccccgtgttt tttatcttcg gtatcacggg tattccaaat attttctctt ttgggcttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

```

<210> 99

<211> 658

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 99

```

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 Ala 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 Leu 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
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
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
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
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> 100

<211> 1977

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<220>

<221> misc_feature

<222> (255) .. (255)

<223> n是a, c, g,或t

<220>

<221> misc_feature

<222> (409) .. (411)

<223> ntn是ttr或ctn

<400> 100

```

atgactgtat cgacttcctc ggcttttcat catagcccgt tgtctgatga tgttgagccg      60
attatccaaa aggccaccgc tgccttgctt gagaagcagc agcaggatgg ccattggggtt    120
tttgaattgg aagccgatgc aaccattccc gctgaataca tcctgttaaa gcattatttg    180
ggtgaacccg aagatttaga aatagaggcc aagataggct gctatttgcg tcgtattcag    240
ggcgagcatg gcgcntggtc tttgttttat ggtgggtgatc ttgatttgag cgccacggtc    300
aaagcctatt ttgccttgaa aatgatcgga gattctcctg atgcgcctca tatgcttcga    360
gccagaaatg aaattttggc acggggtggg gcgatgcgtg ccaatgtcnt nacacgtatt    420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaagcc cgtagagtgg    480
atgttgatgc cggaatgggt tccggttcac atcaataaaa tggcctattg ggcaagaacc    540
gttttagtcc cgttattggg tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt    600
atcttggttg atgaattatt tgtgccgatg gttttaccga cccttcagga aagcggtgac    660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat    720
tgGCCGAAAA atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt    780
ttgaatgggt aagacggggt ggggtgctatt tatccggcga ttgccaatag cgtcatgatg    840
tatgatgcct tgggatatcc cgaaaacatc ccagaaagag ccattgcccg tcgggctgtc    900
gaaaaattga tgggtgttaga tggcacggaa gatcagggtg ataaagaagt ctactgtcag    960
ccttgtttat ccccgatttg ggataccgct ttggttgccc atgcatgtt ggaagtcgga   1020
ggcgatgagg ctgaaaaatc ggctatttct gccttgagct ggttaaagcc gcaacaaatt   1080
ttggatgtaa agggcgattg ggcatggcgg cggcctgatc tcagaccggg gggatgggcc   1140
tttcaatata gaaatgacta ttatcccgat gtcgatgata cggctgttgt gactatggcg   1200
atggatcgag ccgcaaaatt gtcggatctt cacgatgatt ttgaggaatc taaagcgcgt   1260
gccatggaat ggaccattgg gatgcaaagc gataatggcg gttggggcgc tttcgatgcc   1320
aataacagct atacttatct gaataatatt ccctttgctg atcatggcgc gttacttgat   1380
ccgccaacgg tcgatgtctc ggcacgtctc gtttcaatga tggcgcaagc cggatatctc   1440
attacagatc ccaaaaatga agcggcagtt gattatcttc tgaaagagca agaagaggat   1500
ggtagctggg tcgggcgttg ggggtgtcaat tacatatatg gcacatggtc ggccttatgt   1560
gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggccttg   1620
ctgaaaacta ttcaaaatga agatggtggg tggggtgaaa attgcgatag ctatgccctt   1680
gattatagcg gatacgagcc gatggattcg acggcttccc aaacagcatg ggctttattg   1740
ggcttgatgg ctgttgggga agctaattcc gagggcgtga caaagggtat aaactggttg   1800
gcacaaaatc aggatgaaga aggattgtgg aaagaagatt attatagtgg cgggtggttt   1860
ccccgtgttt tttatcttcg gtatcacggg tattccaaat attttctctt ttgggcttta   1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa     1977

```

<210> 101

<211> 658

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 101

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
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
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
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
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 His Trp Leu Ala Gln Asn Gln Asp Glu Glu Gly
595 600 605
Leu Trp Lys Glu Asp Tyr Tyr Ser Gly Gly Gly Tyr 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> 102

<211> 1977

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 102

atgactgtat cgacttcctc ggcttttcat catagcccgt tgtctgatga tgttgagccg	60
attatccaaa aggccacccg tgccttgctt gagaagcagc agcaggatgg ccattggggtt	120
tttgaattgg aagccgatgc aaccattccc gctgaataca tcctgttaaa gcattatttg	180
ggtgaacccg aagatttaga aatagaggcc aagataggtc gctatttgcg tcgtattcag	240
ggcgagcatg gcggatggtc tttgttttat ggtggtgatc ttgatttgag cgccacggtc	300
aaagcctatt ttgccttgaa aatgatcgga gattctcctg atgcgcctca tatgcttcga	360
gccagaaatg aaattttggc acgggggtggg gcgatgcgtg ccaatgtctt tacacgtatt	420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaatgcc cgtagagttag	480
atgttgatgc cggaatgggt tccggttcac atcaataaaa tggcctattg ggcaagaacc	540
gttttagtcc cgttattggt tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt	600
atcttggttg atgaattatt tgtgccggat gttttaccga cccttcagga aagcgggtgac	660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat	720
tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt	780
ttgaatgggtg aagacgggtt ggggtgctatt tatccggcga ttgccaatag cgtcatgatg	840
tatgatgcct tgggatatcc cgaaaacat ccagaaagag ccattgcccg tcgggctgtc	900
gaaaaattga tgggtgtaga tggcacggaa gatcagggtg ataaagaagt ctactgtcag	960
ccctgtttat ccccgatttg ggataccgct ttggttgccc atgcatgtt ggaagtcgga	1020
ggcgatgagg ctgaaaaatc ggctatttct gccttgagct ggttaaagcc gcaacaaatt	1080
ttggatgtaa agggcgattg ggcatggcgg cggcctgatc tcagaccggg gggatgggcc	1140
tttcaatata gaaatgacta ttatcccgat gtcgatgata cggctgttgt gactatggcg	1200
atggatcgag ccgcaaaatt gtcggatctt cacgatgatt ttgaggaatc taaagcgcgt	1260
gccatggaat ggaccattgg gatgcaaagc gataatggcg gttggggcgc tttcgatgcc	1320
aataacagct atacttatct gaataatatt ccctttgctg atcatggcgc gttacttgat	1380

ccgccaacgg tcgatgtctc ggcacgctgc gtttcaatga tggcgcaagc cggtatctcg 1440
 attacagatc ccaaaatgaa agcggcagtt gattatcttc tgaaagagca agaagaggat 1500
 ggtagctggg tcgggcgttg ggggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
 gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggccttg 1620
 ctgaaaacta ttcaaaatga agatgggtggg tggggtgaaa attgcgatag ctatgccctt 1680
 gattatagcg gatacgagcc gatggattcg acggcttccc aaacagcatg ggctttattg 1740
 ggcttgatgg ctgttgggga agctaattcc gagggcgtga caaagggtat acaytggttg 1800
 gcacaaaatc aggatgaaga aggattgtgg aaagaagatt attatagtgg cgggtggttay 1860
 ccccggtgttt ttatcttcg gtatcacggg tattccaaat attttcctct ttgggcttta 1920
 gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 103

<211> 658

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 103

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 Leu 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
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
Ala Arg Tyr Arg Asn Leu Lys Lys	Ala Asn Gln Pro Ile Val His Tyr	
645	650	655

Gly Met

<210> 104

<211> 1977

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<220>

<221> misc_feature

<222> (409) .. (411)

<223> ntn是ttr或ctn

<400> 104

atgactgtat cgacttcctc ggettttcat catagcccg	tgtctgatga tgttgagccg	60
attatccaaa aggccacccg tgccttgctt gagaagcagc	agcaggatgg ccattgggtt	120
tttgaattgg aagccgatgc aaccattccc gctgaataca	tcctgttaaa gcattatttg	180
ggtgaacccg aagatttaga aatagaggcc aagataggtc	gctatttgcg tcgtattcag	240
ggcgagcatg gcggatggtc tttgttttat ggtggtgatc	ttgatttgag cgccacggtc	300
aaagcctatt ttgccttgaa aatgatcgga gattctctg	atgcgcctca tatgcttcga	360

```

gccagaaatg aaattttggc acggggtggg gcgatgcgtg ccaatgtcnt nacacgtatt 420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaagcc cgtagagttg 480
atgttgatgc cggaatggtt tccggttcac atcaataaaa tggcctattg ggcaagaacc 540
gttttagtcc cgttattggt ttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt 600
atcttggttg atgaattatt tgtgccgat gttttaccga cccttcagga aagcggtagac 660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat 720
tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt 780
ttgaatggtg aagacgggtt ggggtgctatt tatccggcga ttgccaatag cgtcatgatg 840
tatgatgcct tgggatatcc cgaaaacat ccagaaagag ccattgcccg tcgggctgtc 900
gaaaaattga tgggtgtaga tggcacggaa gatcagggtg ataaagaagt ctactgtcag 960
ccttgtttat ccccgatttg ggataccgct ttggttgccc atgcatgtt ggaagtcgga 1020
ggcgtagagg ctgaaaaatc ggcattttct gccttgagct ggttaaagcc gcaacaaatt 1080
ttggatgtaa agggcgattg ggcatggcgg cggcctgac tcagaccgg gggatgggcc 1140
tttcaatata gaaatgacta ttatcccgat gtcgatgata cggctgttgt gactatggcg 1200
atggatcgag ccgcaaaatt gtcggatctt cacgatgatt ttgaggaatc taaagcgcgt 1260
gccatggaat ggaccattgg gatgcaaagc gataatggcg gttggggcgc tttcgatgcc 1320
aataacagct atacttatct gaataatatt ccctttgctg atcatggcgc gttacttgat 1380
ccgccaacgg tcgatgtctc ggacgctgc gtttcaatga tggcgcaagc cggtatctcg 1440
attacagatc ccaaaatgaa agcggcagtt gattatcttc tgaaagagca agaagaggat 1500
ggtagctggt tcgggcgttg ggggtgcaat tacatatatg gcacatggc ggccttatgt 1560
gcattgaatg tggccgcttt acccatgat catttagctg ttcagaaagc tgtggcttgg 1620
ctgaaaacta ttcaaaatga agatggtggt tggggtgaaa attgcgatag ctatgccctt 1680
gattatagcg gatacgagcc gatggattcg acggcttccc aaacagcatg ggctttattg 1740
ggcttgatgg ctgttgggga agctaattcc gaggccgtga caaagggtat aaactggttg 1800
gcacaaaatc aggatgaaga aggattgtgg aaagaagatt attatagtgg cgggtggttt 1860
ccccgtgttt ttatcttcg gtatcacggt tattccaaat attttctct ttgggcttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

```

<210> 105

<211> 658

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 105

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
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 Leu 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
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
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
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
Met Asp Arg Ala Ala Lys Leu Ser Asp Leu His Asp Asp Phe Glu Glu		
	405	410
Ser Lys Ala Arg Ala Met Glu Trp Thr Ile Gly Met Gln Ser Asp Asn		
	420	425
Gly Gly Trp Gly Ala Phe Asp Ala Asn Asn Ser Tyr Thr Tyr Leu Asn		
	435	440
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
Ile Thr Asp Pro Lys Met Lys Ala Ala Val Asp Tyr Leu Leu Lys Glu		
	485	490
Gln Glu Glu Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Ile		
	500	505
Tyr Gly Thr Trp Ser Ala Leu Cys Ala Leu Asn Val Ala Ala Leu Pro		
	515	520
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
Asp Tyr Ser Gly Tyr Glu Pro Met Asp Ser Thr Ala Ser Gln Thr Ala		
	565	570
Trp Ala Leu Leu Gly Leu Met Ala Val Gly Glu Ala Asn Ser Glu Ala		
	580	585
Val Thr Lys Gly Ile Asn Trp Leu Ala Gln Asn Gln Asp Glu Glu Gly		
	595	600
Leu Trp Lys Glu Asp Tyr Tyr Ser Gly Gly Gly Tyr 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
Ala Arg Tyr Arg Asn Leu Lys Lys Ala Asn Gln Pro Ile Val His Tyr		
	645	650
Gly Met		655

<210>	106	
<211>	1977	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成ZmoSHC2衍生物	
<220>		
<221>	misc_feature	
<222>	(409) .. (411)	
<223>	ntn是ttr或ctn	
<400>	106	
atgactgtat	cgacttcctc	ggcttttcat
attatccaaa	aggccaccgc	tgccttgctt
tttgaattgg	aagccgatgc	aaccattccc
ggtgaacccg	aagatttaga	aatagaggcc
ggcgagcatg	gcggatggtc	tttgttttat
aaagcctatt	ttgccttgaa	aatgatcgga
gccagaaatg	aaattttggc	acgggggtggg
caattagctc	tggtcggggc	aatgtcatgg
atgttgatgc	cggaatgggt	tccggttcac
gttttagtcc	cgttattggg	tttacaggcg
atcttggttg	atgaattatt	tgtgccggat
cctatatggc	gtcgtttttt	ttcggcactt
tggccgaaaa	atatgcgcgc	gaaggctatt
ttgaatgggtg	aagacggggt	gggtgctatt
tatgatgcct	tgggatatcc	cgaaaaccat
gaaaaattga	tgggtgtaga	tggcacggaa
ccttgtttat	ccccgatttg	ggataccgct
ggcgatgagg	ctgaaaaatc	ggctatttct
ttggatgtaa	agggcgattg	ggcatggcgg
tttcaatata	gaaatgacta	ttatcccgat
atggatcgag	ccgcaaaaatt	gtcggatctt
gcatggaat	ggaccattgg	gatgcaaagc
aataacagct	atacttatct	gaataatatt
ccgccaacgg	tcgatgtctc	ggcacgctgc
attacagatc	ccaaaaatgaa	agcggcagtt
ggtagctggg	tcgggcgttg	gggtgtcaat
gcattgaatg	tggccgcttt	accccatgat
ctgaaaacta	ttcaaaaatga	agatggtggg
		tggggtgaaa
		attgcgatag
		ctatgccctt
		60
		120
		180
		240
		300
		360
		420
		480
		540
		600
		660
		720
		780
		840
		900
		960
		1020
		1080
		1140
		1200
		1260
		1320
		1380
		1440
		1500
		1560
		1620
		1680

gattatagcgc gatacgagcc gatggattcg acggcttccc aaacagcatg ggctttattg 1740
 ggcttgatgg ctgttgggga agctaattcc gaggccgtga caaagggtat aaactggttg 1800
 gcacaaaatc aggatgaaga aggattgtgg aaagaagatt attatagtgg cgggtggttay 1860
 ccccggtgttt tttatcttcg gtatcacggc tattccaaat attttcctct ttgggcttta 1920
 gcgcgcctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977
 <210> 107
 <211> 658
 <212> PRT
 <213> 人工序列
 <220>
 <223> 合成ZmoSHC2衍生物
 <400> 107
 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 Leu Thr Arg Arg 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
Trp Pro Lys Asn Met Arg Ala Lys Ala Ile His Ser Cys Val His Phe		240
	245	250
Val Thr Glu Arg Leu Asn Gly Glu Asp Gly Leu Gly Ala Ile Tyr Pro		255
	260	265
Ala Ile Ala Asn Ser Val Met Met Tyr Asp Ala Leu Gly Tyr Pro Glu		270
	275	280
Asn His Pro Glu Arg Ala Ile Ala Arg Arg Ala Val Glu Lys Leu Met		285
290	295	300
Val Leu Asp Gly Thr Glu Asp Gln Gly Asp Lys Glu Val Tyr Cys Gln		
305	310	315
Pro Cys Leu Ser Pro Ile Trp Asp Thr Ala Leu Val Ala His Ala Met		320
	325	330
Leu Glu Val Gly Gly Asp Glu Ala Glu Lys Ser Ala Ile Ser Ala Leu		335
	340	345
Ser Trp Leu Lys Pro Gln Gln Ile Leu Asp Val Lys Gly Asp Trp Ala		350
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
Met Asp Arg Ala Ala Lys Leu Ser Asp Leu His Asp Asp Phe Glu Glu		400
	405	410
Ser Lys Ala Arg Ala Met Glu Trp Thr Ile Gly Met Gln Ser Asp Asn		415
	420	425
Gly Gly Trp Gly Ala Phe Asp Ala Asn Asn Ser Tyr Thr Tyr Leu Asn		430
	435	440
Asn Thr Pro Phe Ala Asp His Gly Ala Leu Leu Asp Pro Pro Thr Val		445
	450	455
Asp Val Ser Ala Arg Cys Val Ser Met Met Ala Gln Ala Gly Ile Ser		
465	470	475
Ile Thr Asp Pro Lys Met Lys Ala Ala Val Asp Tyr Leu Leu Lys Glu		480
	485	490
Gln Glu Glu Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Ile		495
	500	505
Tyr Gly Thr Trp Ser Ala Leu Cys Ala Leu Asn Val Ala Ala Leu Pro		510
	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> 108

<211> 1977

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<220>

<221> misc_feature

<222> (409) .. (411)

<223> ntn是ttr或ctn

<220>

<221> misc_feature

<222> (418) .. (420)

<223> ngn是agr或cgn

<220>

<221> misc_feature

<222> (1350) .. (1350)

<223> n是a, c, g,或t

<400> 108

atgactgtat cgacttcctc ggcttttcat catagcccggt tgtctgatga tgttgagccg 60
 attatccaaa aggccaccg tgccttgctt gagaagcagc agcaggatgg ccattgggtt 120
 tttgaattgg aagccgatgc aaccattccc gctgaatata tcctgttaaa gcattatttg 180


```

ggtgaacccg aagatttaga aatagaggcc aagataggtc gctatttgcg tcgtattcag      240
ggcgagcatg gcggatggtc tttgttttat ggtggatgatc ttgatttgag cgccacggtc      300
aaagcctatt ttgccttgaa aatgatcgga gattctcctg atgcgcctca tatgcttcga      360
gccagaaatg aaattttggc acgggggtggg gcgatgcgtg ccaatgtcnt nacacgtngn      420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaagcc cgtagagttg      480
atgttgatgc cggaatggtt tccggttcac atcaataaaa tggcctattg ggcaagaacc      540
gttttagtcc cgttattggt tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt      600
atcttggttg atgaattatt tgtgccggat gttttaccga cccttcagga aagcgggtgac      660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat      720
tgcccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt      780
ttgaatggtg aagacggggtt ggggtgctatt tatccggcga ttgccaatag cgtcatgatg      840
tatgatgcct tgggatatcc cgaaaacat ccagaaagag ccattgcccg tcgggctgtc      900
gaaaaattga tgggtgtaga tggcacggaa gatcagggtg ataaagaagt ctactgtcag      960
ccttgtttat ccccgatttg ggataccgct ttggttgccc atgcatgtt ggaagtcgga     1020
ggcgatgagg ctgaaaaatc ggctatttct gccttgagct ggttaaagcc gcaacaaatt     1080
ttggatgtaa agggcgattg ggcatggcgg cggcctgatc tcagaccggg ggcatgggcc     1140
tttcaatata gaaatgacta ttatcccgat gtcgatgata cggctgttgt gactatggcg     1200
atggatcgag ccgcaaaatt gtcggatctt cacgatgatt ttgaggaatc taaagcgcgt     1260
gccatggaat ggaccattgg gatgcaaagc gataatggcg gttggggcgc tttcgatgcc     1320
aataacagct atacttatct gaataatacn ccctttgctg atcatggcgc gttacttgat     1380
ccgccaacgg tcgatgtctc ggacgctgc gtttcaatga tggcgcaagc cggtatctcg     1440
attacagatc ccaaaatgaa agcggcagtt gattatcttc tgaaagagca agaagaggat     1500
ggtagctggt tcgggcgttg ggggtgtcaat tacatatatg gcacatggtc ggccttatgt     1560
gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggccttg     1620
ctgaaaacta ttcaaaatga agatgggtgg tggggtgaaa attgcgatag ctatgccctt     1680
gattatagcg gatacgagcc gatggattcg acggcttccc aaacagcatg ggctttattg     1740
ggcttgatgg ctgttgggga agctaattcc gaggccgtga caaagggtat aaactggttg     1800
gcacaaaatc aggatgaaga aggattgtgg aaagaagatt attatagtgg cggtggtttt     1860
ccccgtgttt tttatcttcg gtatcacggg tattccaaat attttcctct ttgggcttta     1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa      1977

```

<210> 109

<211> 658

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 109

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
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 Arg 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
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
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
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 Thr 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 Tyr 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> 110

<211> 1977

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<220>

<221> misc_feature

<222> (418) .. (420)

<223> ngn是agr或cgn

<220>

<221> misc_feature

<222> (1350) .. (1350)

<223> n是a, c, g,或t

<400> 110

```

atgactgtat cgacttcctc ggcttttcat catagcccggt tgtctgatga tgttgagccg      60
attatccaaa aggccacccg tgccttgctt gagaagcagc agcaggatgg ccattggggtt      120
tttgaattgg aagccgatgc aaccattccc gctgaataca tcctgttaaa gcattatttg      180
ggatgaacccg aagatttaga aatagaggcc aagataggtc gctatttgcg tcgtattcag      240
ggcgcagcatg gcggatggtc tttgttttat ggtgggtgatc ttgatttgag cgccacggtc      300
aaagcctatt ttgccttgaa aatgatcgga gattctcctg atgcgcctca tatgcttcga      360
gccagaaatg aaattttggc acgggggtggg gcgatgcgtg ccaatgtctt tacacgtngn      420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaagcc cgtagagttg      480
atgttgatgc cggaatgggt tccggttcac atcaataaaa tggcctattg ggcaagaacc      540
gttttagtcc cgttattggt tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt      600
atcttggttg atgaattatt tgtgccggat gttttaccga cccttcagga aagcgggtgac      660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat      720
tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt      780
ttgaatgggt aagacggggt ggggtgctatt tatccggcga ttgccaatag cgtcatgatg      840
tatgatgcct tgggatatcc cgaaaacat ccagaaagag ccattgcccc tcgggctgtc      900
gaaaaattga tgggtgtaga tggcacggaa gatcagggtg ataaagaagt ctactgtcag      960
ccttgtttat ccccgatttg ggataccgct ttggttgccc atgcatgtt ggaagtcgga     1020
ggcgcgatgagg ctgaaaaatc ggctatttct gccttgagct ggttaaagcc gcaacaaatt     1080
ttggatgtaa agggcgattg ggcatggcgg cggcctgatc tcagaccgga gggatgggcc     1140
tttcaatata gaaatgacta ttatcccgat gtcgatgata cggctgttgt gactatggcg     1200
atggatcgag ccgcaaaatt gtcggatctt cacgatgatt ttgaggaatc taaagcgcgt     1260

```

gccatggaat ggaccattgg gatgcaaagc gataatggcg gttggggcgc tttcgatgcc 1320
 aataacagct atacttatct gaataatacn ccctttgctg atcatggcgc gttacttgat 1380
 ccgccaacgg tcgatgtctc ggcacgctgc gtttcaatga tggcgcaagc cggtatctcg 1440
 attacagatc ccaaaatgaa agcggcagtt gattatcttc tgaaagagca agaagaggat 1500
 ggtagctggg tcgggcgttg ggggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
 gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggccttg 1620
 ctgaaaacta ttcaaaatga agatgggtgg tggggtgaaa attgcgatag ctatgccctt 1680
 gattatagcg gatacgagcc gatggattcg acggcttccc aaacagcatg ggctttattg 1740
 ggcttgatgg ctgttgggga agctaattcc gaggccgtga caaagggtat aaactggttg 1800
 gcacaaaatc aggatgaaga aggattgtgg aaagaagatt attatagtgg cgggtggttay 1860
 ccccggtttt tttatcttcg gtatcacggg tattccaaat attttctctt ttgggcctta 1920
 gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 111

<211> 658

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 111

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 Leu Thr Arg Arg 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	Thr	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 Tyr 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> 112

<211> 1977

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<220>

<221> misc_feature

<222> (409) .. (411)

<223> ntn是ttr或ctn

<220>

<221> misc_feature

<222> (418) .. (420)

<223> ngn是agr或cgn

<220>

<221> misc_feature

<222> (1350) .. (1350)

<223> n是a, c, g,或t

<400> 112

```

atgactgtat cgacttcctc ggcttttcat catagcccgt tgtctgatga tgttgagccg      60
attatccaaa aggccacccg tgccttgctt gagaagcagc agcaggatgg ccattggggtt      120
tttgaattgg aagccgatgc aaccattccc gctgaataca tcctgttaaa gcattatttg      180
ggtgaacccg aagatttaga aatagaggcc aagataggtc gctatttgcg tcgtattcag      240
ggcgagcatg gcggatggtc tttgttttat ggtggtgatc ttgatttgag cgccacggtc      300
aaagcctatt ttgccttgaa aatgatcgga gattctcctg atgcgcctca tatgcttcga      360
gccagaaatg aaattttggc acgggggtggg gcgatgcgtg ccaatgtcnt nacacgtngn      420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaagcc cgtagagttg      480
atgttgatgc cggaatgggt tccggttcac atcaataaaa tggcctattg ggcaagaacc      540
gttttagtcc cgttattggg tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggg      600
atcttggttg atgaattatt tgtgccgat gttttaccga cccttcagga aagcgggtgac      660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat      720
tgGCCGAAAA atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt      780
ttgaatgggt aagacggggt gggtgctatt tatccggcga ttgccaatag cgtcatgatg      840
tatgatgcct tgggatatcc cgaaaaccat ccagaaagag ccattgcccg tcgggctgtc      900
gaaaaattga tgggtgttaga tggcacggaa gatcagggtg ataaagaagt ctactgtcag      960
ccttgtttat ccccgatttg ggataccgct ttggttgccc atgccatgtt ggaagtcgga      1020
ggcgatgagg ctgaaaaatc ggctatttct gccttgagct ggttaaagcc gcaacaaatt      1080
ttggatgtaa agggcgattg ggcatggcgg cggcctgacg tcagaccggg gggatgggcc      1140
tttcaatata gaaatgacta ttatcccgat gtcgatgata cggctgttgt gactatggcg      1200
atggatcgag ccgcaaaatt gtcggatctt cacgatgatt ttgaggaatc taaagcgcgt      1260
gccatggaat ggaccatttg gatgcaaagc gataatggcg gttggggcgc tttcgatgcc      1320
aataacagct atacttatct gaataatacn ccctttgctg atcatggcgc gttacttgat      1380
ccgccaacgg tcgatgtctc ggcacgtgc gtttcaatga tggcgcaagc cggtatctcg      1440
attacagatc ccaaaatgaa agcggcagtt gattatcttc tgaaagagca agaagaggat      1500
ggtagctggg tcgggcgttg ggggtgtcaat tacatatatg gcacatggtc ggccttatgt      1560
gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggccttg      1620
ctgaaaaacta ttcaaaatga agatggtggg tggggtgaaa attgcgatag ctatgccctt      1680
gattatagcg gatacgagcc gatggattcg acggcttccc aaacagcatg ggctttattg      1740
ggcttgatgg ctgttgggga agctaattcc gaggccgtga caaagggtat aaactggttg      1800
gcacaaaatc aggatgaaga aggattgtgg aaagaagatt attatagtgg cgggtggttay      1860
ccccgtgttt tttatcttcg gtatcacggg tattccaaat attttctctt ttgggcttta      1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa      1977

```

<210> 113

<211> 684

<212> PRT

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 113

```

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
Thr Ile Leu Thr Leu Leu Ala Phe Asp Asp Ala Gly Val Leu Gly Asp		
	340	345
Tyr Pro Glu Ala Val Asp Lys Ala Val Asp Trp Val Leu Gln Arg Gln		
	355	360
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
Lys Trp Lys Ala Lys Gly Ile Asp Glu Ala Ile Gln Leu Gly Val Asp		
	420	425
Trp Leu Ile Gly Met Gln Ser Gln Gly Gly Gly Trp Gly Ala Phe Asp		
	435	440
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
Gln Ala Leu Asp Tyr Ile Arg Arg Glu Gln Glu Pro Ser Gly Pro Trp		
	500	505
Phe Gly Arg Trp Gly Val Asn Tyr Val Tyr Gly Thr Gly Ala Val Leu		
	515	520
Pro Ala Leu Ala Ala Ile Gly Glu Asp Met Thr Gln Pro Tyr Ile Gly		
	530	535
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
Thr Thr Thr Ala Ser Gln Thr Ala Trp Ala Leu Met Ala Leu Leu Ala		

580	585	590
Ala Asn Arg Pro Gln Asp Lys Asp Ala Ile Glu Arg Gly Cys Met Trp		
595	600	605
Leu Val Glu Arg Gln Ser Ala Gly Thr Trp Asp Glu Pro Glu Phe Thr		
610	615	620
Gly Thr Gly Phe Pro Gly Tyr Gly Val Gly Gln Thr Ile Lys Leu Asn		
625	630	635
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> 114		
<211> 2055		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成BjapSHC1衍生物		
<400> 114		
atgactgtga ccagctcggc ctccgcgcgt gcgacgcgcg acccgggaaa ttatcagact	60	
gccctgcaat cgacgggtgcg cgcggcggcg gattggctga tcgccaacca gaagccggac	120	
ggccattggg tcggccgcgc cgagtccaat gcctgcatgg aggcgcaatg gtgcctcgcg	180	
ctgtgggttca tggggctcga ggaccatccg ctgcgcaagc gcctgggcca gtcgtgctc	240	
gacagccagc gcccggacgg cgcctggcag gtctatttgc gcgcccccaa tggcgacatc	300	
aacgcgactg tcgaggccta tgccgcgctc cgctcgctgg gcttccgcga cgacgagccg	360	
gcggtgcgcc gggcgcggga atggatcgag gccaaaggcg gcctgcgcaa catccgcgtc	420	
ttcacccgct actggctggc actgatcggc gaatggccgt gggagaagac accgaacatc	480	
ccgccggagg tgatctggtt tccgctctgg tttccgttct cgatctacaa tttcgcgcaa	540	
tgggcccgcg ccaccttgat gccgatcgcc gtgctgtcgg cgcggcggcc gagccggccg	600	
ctgccgccgg agaaccgcct cgatgcgctg tttccgcatg gacggaaggc gttcgactac	660	
gaactgccgg tcaaggccgg cgccggcggc tgggacaggt tcttccgcgg cgccgacaag	720	
gttctgcaca agctgcagaa cctcggcaac cgtctcaatc tcggcctgtt ccgcccggcg	780	
gccaccagcc gcgtgctgga atggatgatc cgccatcagg atttcgacgg cgccctggggc	840	
ggcatccagc cgccctggat ctacgggctg atggcgctct atgccgaagg ctatccgctc	900	
aatcatcccg tgctcgcaaa gggcctcgac gcgctgaacg atcccggctg gcgcgtcgat	960	
gtcggtgacg ccacctacat ccaggccacc aacagcccgg tctgggacac gatcctgacc	1020	
ttgctcgctt tcgacgatgc cggcgtgctc ggcgactatc ccgaggccgt cgacaaggcg	1080	
gtcgactggg tgctgcagcg gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg	1140	

```

catgtcaagc cccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac 1200
gacaccgcgg tcgcgtgat cgcgctggcg ccactgcgcc acgatccgaa atggaaggcc 1260
aaagggatcg acgaggctat ccagctcggg gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcggt cgacaaggac aacaaccaga agatcctgac caagatcccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcacgc ccacatcatc 1440
gaggcggttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggg ccgtgggttcg gccgctgggg cgtcaattac 1560
gtctacggca cccgcgcggt gctgccggcg ctggccgcga tcggcgagga catgaccagg 1620
ccctatatcg gccgcgcctg cgactggctg gttgcccatc agcaggccga tggcggctgg 1680
ggcgagagct ggcctccta catggatgtc agcgcggtcg gccgcggcac cacaacggcc 1740
tcgcagaccg cctgggcgct gatggcgtg ctgccgcga atcgcccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttcccgggc tacggcgctg gccagaccat caagctgaac 1920
gatcccgcg tgcgcagcg gctgatgcag ggcccgaat tgtcccgcgc ctteatgctc 1980
cgctacggca tgtaaccgca ctacttcccg ctgatggcg tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

```

<210> 115

<211> 684

<212> PRT

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 115

```

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
Pro Pro Glu Val Ile Trp Phe Pro Leu Trp Phe Pro Phe Ser Ile Tyr		
	165	170
Asn Phe Ala Gln Trp Ala Arg Ala Thr Leu Met Pro Ile Ala Val Leu		
	180	185
Ser Ala Arg Arg Pro Ser Arg Pro Leu Pro Pro Glu Asn Arg Leu Asp		
	195	200
Ala Leu Phe Pro His Gly Arg Lys Ala Phe Asp Tyr Glu Leu Pro Val		
	210	215
Lys Ala Gly Ala Gly Gly Trp Asp Arg Phe Phe Arg Gly Ala Asp Lys		
225	230	235
Val Leu His Lys Leu Gln Asn Leu Gly Asn Arg Leu Asn Leu Gly Leu		
	245	250
Phe Arg Pro Ala Ala Thr Ser Arg Val Leu Glu Trp Met Ile Arg His		
	260	265
Gln Asp Phe Asp Gly Ala Trp Gly Gly Ile Gln Pro Pro Trp Ile Tyr		
	275	280
Gly Leu Met Ala Leu Tyr Ala Glu Gly Tyr Pro Leu Asn His Pro Val		
	290	295
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
Thr Ile Leu Thr Leu Leu Ala Phe Asp Asp Ala Gly Val Leu Gly Asp		
	340	345
Tyr Pro Glu Ala Val Asp Lys Ala Val Asp Trp Val Leu Gln Arg Gln		
	355	360
Val Arg Val Pro Gly Asp Trp Ser Met Lys Leu Pro His Val Lys Pro		
	370	375
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
Lys Trp Lys Ala Lys Gly Ile Asp Glu Ala Ile Gln Leu Gly Val Asp		
	420	425
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	635
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> 116		
<211> 2055		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成BjapSHC1衍生物		
<400> 116		
atgactgtga ccagctcggc ctccgcgcgt gcgacgcgcg acccgggaaa ttatcagact		

gccctgcaat	cgacggtgcg	cgcggcggcg	gattggctga	tcgccaacca	gaagccggac	120
ggccattggg	tcggccgcgc	cgagtccaat	gcctgcatgg	aggcgcaatg	gtgcctcgcg	180
ctgtggttca	tggggctcga	ggaccatccg	ctgcgcaagc	gcctgggcca	gtcgctgctc	240
gacagccagc	gcccggacgg	cgcttggcag	gtctatttct	gcgcccccaa	tggcgacatc	300
aacgcgactg	tcgaggccta	tgccgcgctc	cgctcgctgg	gcttccgcga	cgacgagccg	360
gcggtgcgcc	gggcgcggga	atggatcgag	gccaaaggcg	gcctgcgcaa	catccgcgtc	420
ttcaccgcgt	actggctggc	actgatcggc	gaatggccgt	gggagaagac	accgaacatc	480
ccgccggagg	tgatctggtt	tccgctctgg	tttccgttct	cgatctacaa	tttcgcgcaa	540
tgggcccgcg	ccaccttgat	gccgatcgcc	gtgctgtcgg	cgcggcggcc	gagccggccg	600
ctgccgccgg	agaaccgcct	cgatgcgctg	tttccgcatg	gacggaaggc	gttcgactac	660
gaactgccgg	tcaaggccgg	cgccggcggc	tgggacaggt	tcttccgcgg	cgccgacaag	720
gttctgcaca	agctgcagaa	cctcggcaac	cgtctcaatc	tcggcctgtt	ccgcccggcg	780
gccaccagcc	gcgtgctgga	atggatgate	cgccatcagg	atttcgacgg	cgcttggggc	840
ggcatccagc	cgccctggat	ctacgggctg	atggcgctct	atgccgaagg	ctatccgctc	900
aatcatcccc	tgctcgcaaa	gggcctcgac	gcgctgaacg	atcccggctg	gcgcgtcgat	960
gtcggtgacg	ccacctacat	ccaggccacc	aacagcccgg	tctgggacac	gatcctgacc	1020
ttgctcgctt	tcgacgatgc	cggcgtgctc	ggcgactatc	ccgaggccgt	cgacaaggcg	1080
gtcgactggg	tgctgcagcg	gcaggtgcg	gtgcccggcg	actggtcgat	gaagctgccg	1140
catgtcaagc	ccggcggtcg	ggcgttcgaa	tacgccaaca	actactatcc	cgacacggac	1200
gacaccgcgg	tcgcgctgat	cgcgctggcg	ccactgcgcc	acgatccgaa	atggaaggcc	1260
aaagggatcg	acgaggctat	ccagctcggt	gtcgactggc	tgatcgcat	gcagagccag	1320
ggcggcggtc	ggggcgcggt	cgacaaggac	aacaaccaga	agatcctgac	caagatcccc	1380
ttctgcgatt	atggcgaggc	gctcgatccg	ccctcggtcg	acgtcacccg	ccacatcatc	1440
gaggcggttc	gcaagctcgg	catctcgcg	aaccatccgt	cgatggtgca	ggcgctggac	1500
tatatctgcc	gtgagcagga	gccgagcggt	ccgtggttcg	gccgctgggg	cgtcaattac	1560
gtctacggca	ccggcgcggt	gctgccggcg	ctggccgcga	tcggcgagga	catgaccag	1620
ccctatatcg	gccgcgcctg	cgactggctg	gttgcccatc	agcaggccga	tggcggtcgg	1680
ggcgagagct	gcgcctccta	catggatgtc	agcgcggtcg	gccgcggcac	cacaacggcc	1740
tcgcagaccg	cctgggcgct	gatggcgctg	ctcgccgcca	atcgccccca	ggacaaggac	1800
gcgatcgagc	gtggctgcat	gtggctggtc	gagcgccagt	cggccggcac	ctgggacgag	1860
ccggaattca	ccggcaccgg	tttcccgggc	tacggcgctg	gccagaccat	caagctgaac	1920
gatcccgcgc	tgtcgcagcg	gctgatgcag	ggcccggaat	tgtcccgcgc	cttcatgctc	1980
cgctacggca	tgtaccgcca	ctacttccc	ctgatggcgc	tcggccgcgc	cctacgcccc	2040
cagagtcata	gctag					2055
<210>	117					
<211>	684					
<212>	PRT					
<213>	人工序列					
<220>						

<223> 合成BjapSHC1衍生物

<400> 117

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	Leu	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		320
	325	330
Thr Ile Leu Thr Leu Leu Ala Phe Asp Asp Ala Gly Val Leu Gly Asp		335
	340	345
Tyr Pro Glu Ala Val Asp Lys Ala Val Asp Trp Val Leu Gln Arg Gln		350
	355	360
Val Arg Val Pro Gly Asp Trp Ser Met Lys Leu Pro His Val Lys Pro		365
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		400
	405	410
Lys Trp Lys Ala Lys Gly Ile Asp Glu Ala Ile Gln Leu Gly Val Asp		415
	420	425
Trp Leu Ile Gly Met Gln Ser Gln Gly Gly Gly Trp Gly Ala Phe Asp		430
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		480
	485	490
Gln Ala Leu Asp Tyr Ile Arg Arg Glu Gln Glu Pro Ser Gly Pro Trp		495
	500	505
Phe Gly Arg Trp Gly Val Asn Tyr Val Tyr Gly Thr Gly Ala Val Leu		510
	515	520
Pro Ala Leu Ala Ala Ile Gly Glu Asp Met Thr Gln Pro Tyr Ile Gly		525
	530	535
Arg Ala Cys Asp Trp Leu Val Ala His Gln Gln Ala Asp Gly Gly Trp		540
545	550	555
Gly Glu Ser Cys Ala Ser Tyr Met Asp Val Ser Ala Val Gly Arg Gly		560
	565	570
Thr Thr Thr Ala Ser Gln Thr Ala Trp Ala Leu Met Ala Leu Leu Ala		575
	580	585
Ala Asn Arg Pro Gln Asp Lys Asp Ala Ile Glu Arg Gly Cys Met Trp		590
595	600	605

Leu Val Glu Arg Gln Ser Ala Gly Thr Trp Asp Glu Pro Glu Phe Thr
 610 615 620
 Gly Thr Gly Phe Pro Gly Tyr Gly Val Gly Gln Thr Ile Lys Leu Asn
 625 630 635 640
 Asp Pro Ala Leu Ser Gln Arg Leu Met Gln Gly Pro Glu Leu Ser Arg
 645 650 655
 Ala Phe Met Leu Arg Tyr Gly Met Tyr Arg His Tyr Phe Pro Leu Met
 660 665 670
 Ala Leu Gly Arg Ala Leu Arg Pro Gln Ser His Ser
 675 680

<210> 118

<211> 2055

<212> DNA

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<220>

<221> misc_feature

<222> (421) .. (423)

<223> ntn是ttr或ctn

<400> 118

atgactgtga ccagctcggc ctccgcgcgt gcgacgcgcg acccgggaaa ttatcagact 60
 gccctgcaat cgacggtgcg cgcggcggcg gattggctga tcgccaacca gaagccggac 120
 ggccattggg tcggccgcgc cgagtccaat gcctgcatgg aggcgcaatg gtgcctcgcg 180
 ctgtggttca tggggctcga ggaccatccg ctgcgcaagc gcctgggcca gtcgctgctc 240
 gacagccagc gcccggacgg cgcctggcag gtctatttcg gcgcccccaa tggcgacatc 300
 aacgcgactg tcgaggccta tgccgcgctc cgctcgcgtg gcttcgcga cgacgagccg 360
 gcggtgcgcc gggcgcggga atggatcgag gccaaaggcg gcctgcgcaa catccgcgtc 420
 ntnaccgcgt actggctggc actgatcggc gaatggccgt gggagaagac accgaacatc 480
 ccgccggagg tgatctggtt tccgctctgg ttccggttct cgatctacaa ttctgcgcaa 540
 tgggcccgcg ccaccttgat gccgatcgcc gtgctgtcgg cgcggcggcc gagccggccg 600
 ctgccgccgg agaaccgcct cgatgcgctg ttccgcgatg gacggaaggc gttcgactac 660
 gaactgccgg tcaaggccgg cgcggcggc tgggacaggt tcttcgcgg cgccgacaag 720
 gttctgcaca agctgcagaa cctcggcaac cgtctcaatc tcggcctgtt ccgcccgcg 780
 gccaccagcc gcgtgctgga atggatgata cgccatcagg atttcgacgg cgcctggggc 840
 ggcatccagc cgccctggat ctacgggctg atggcgctct atgccgaagg ctatccgctc 900
 aatcatcccg tgctcgcaaa gggcctcgac gcgctgaacg atcccggctg gcgcgtcgat 960
 gtcggtgacg ccacctacat ccagccacc aacagcccgg tctgggacac gatcctgacc 1020
 ttgctcgctt tcgacgatgc cggcgtgctc ggcgactatc ccgaggccgt cgacaaggcg 1080

```

gtcgactggg tgctgcagcg gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
catgtcaagc cgggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac 1200
gacaccgcgg tcgcgctgat cgcgctggcg cactgcgcc acgatccgaa atggaaggcc 1260
aaagggatcg acgaggctat ccagctcggg gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcggt cgacaaggac aacaaccaga agatcctgac caagatcccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggg ccgtggttcg gccgctgggg cgtcaattac 1560
gtctacggca ccggcgcggt gctgccggcg ctggccgcga tcggcgagga catgaccag 1620
ccctatatcg gccgcgctg cgactggctg gttgcccatc agcaggccga tggcggctgg 1680
ggcgagagct gcgcctccta catggatgtc agcgcggtcg gccgcggcac cacaacggcc 1740
tcgcagaccg cctgggcgct gatggcgctg ctgcgccca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggtggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttcccgggc tacggcgctg gccagaccat caagctgaac 1920
gatcccgcgc tgtcgagcg gctgatgcag ggcccgaat tgtcccgcgc ctcatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcg ctcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

```

<210> 119

<211> 684

<212> PRT

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 119

```

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 Arg		
130	135	140
Trp Leu Ala Leu Ile Gly Glu Trp Pro Trp Glu Lys Thr Pro Asn Ile		
145	150	155
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
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	635
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>	120	
<211>	2055	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成BjapSHC1衍生物	
<220>		

<221> misc_feature

<222> (430) .. (432)

<223> ngn是agr或cgn

<400> 120

```

atgactgtga ccagctcggc ctccgcgcgt gcgacgcgcg acccgggaaa ttatcagact    60
gccctgcaat cgacggtgcg cgcggcggcg gattggctga tcgccaacca gaagccggac    120
ggccattggg tcggccgcgc cgagtccaat gcctgcatgg aggcgcaatg gtgcctcgcg    180
ctgtggttca tggggctcga ggaccatccg ctgcgcaagc gcctgggcca gtgcgtgctc    240
gacagccagc gcccggacgg cgcttggcag gtctatttcg gcgcccccaa tggcgacatc    300
aacgcgactg tcgaggccta tgccgcgctc cgctcgtcgg gcttccgcga cgacgagccg    360
gcggtgcgcc gggcgcggga atggatcgag gccaaaggcg gcctgcgcaa catccgcgtc    420
ttcaccgcgn gntggctggc actgatcggc gaatggccgt gggagaagac accgaacatc    480
ccgccggagg tgatctggtt tccgctctgg tttccgttct cgatctacaa tttcgcgcaa    540
tgggcccgcg ccacctgat gccgatcgcc gtgctgtcgg cgcggcggcc gagccggccg    600
ctgccgccgg agaaccgcct cgatgcgctg tttccgcatg gacggaaggc gttcgactac    660
gaactgccgg tcaaggccgg cgcggcggc tgggacaggt tcttccgcgg cgccgacaag    720
gttctgcaca agctgcagaa cctcggcaac cgtctcaatc tcggcctgtt ccgcccggcg    780
gccaccagcc gcgtgctgga atggatgatc cgccatcagg atttcgacgg cgcttggggc    840
ggcatccagc cgccctggat ctacgggctg atggcgctct atgccgaagg ctatccgctc    900
aatcatcccg tgctcgcaaa gggcctcgac gcgtgaacg atcccggctg gcgcgtcgat    960
gtcggtgacg ccacctacat ccaggccacc aacagcccgg tctgggacac gatcctgacc   1020
ttgctcgctt tcgacgatgc cggcgtgctc ggcgactatc ccgaggccgt cgacaaggcg   1080
gtcgactggg tgctgcagcg gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg   1140
catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac   1200
gacaccgcgg tcgcgtgat cgcgctggcg cactgcgcc acgatccgaa atggaaggcc   1260
aaagggatcg acgaggctat ccagctcggg gtcgactggc tgatcggcat gcagagccag   1320
ggcggcggct ggggcgcggt cgacaaggac aacaaccaga agatcctgac caagatcccc   1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcacgcg ccacatcatc   1440
gaggcggttc gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac   1500
tatattcgcc gtgagcagga gccgagcggg ccgtgggttc gccgctgggg cgtcaattac   1560
gtctacggca ccggcgcggt gctgccggcg ctggccgcga tcggcgagga catgaccagg   1620
ccctatatcg gccgcgcctg cgactggctg gttgcccatc agcaggccga tggcggctgg   1680
ggcgagagct gcgcctccta catggatgtc agcgcggtcg gccgcggcac cacaacggcc   1740
tcgcagaccg cctgggcgct gatggcgtg ctcgccgcca atcgccccca ggacaaggac   1800
gcgatcgagc gtggctgcat gtggtggtc gagcgccagt cggccggcac ctgggacgag   1860
ccggaattca ccggcaccgg tttcccgggc tacggcgtcg gccagaccat caagctgaac   1920
gatcccgcg cgtcgcagcg gctgatgcag ggcccgaat tgtcccgcgc cttcatgctc   1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg   2040
cagagtcata gctag                                     2055

```

<210> 121

<211> 684

<212> PRT

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 121

```

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		635	640
Asp Pro Ala Leu Ser Gln Arg Leu Met Gln Gly Pro Glu Leu Ser Arg					
	645		650		655
Ala Phe Met Leu Arg Tyr Gly Met Tyr Arg His Tyr Phe Pro Leu Met					
	660		665		670
Ala Leu Gly Arg Ala Leu Arg Pro Gln Ser His Ser					
	675		680		

<210> 122

<211> 2055

<212> DNA

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 122

```

atgactgtga ccagctcggc ctccgcgcgt gcgacgcgcg acccgggaaa ttatcagact      60
gccctgcaat cgacggtgcg cgcggcggcg gattggctga tcgccaacca gaagccggac      120
ggccattggg tcggccgcgc cgagtccaat gcctgcatgg aggcgcaatg gtgcctcgcg      180
ctgtggttca tggggctcga ggaccatccg ctgcgcaagc gcctgggcca gtcgtgctc      240
gacagccagc gcccggacgg cgcctggcag gtctatttcg gcgcccccaa tggcgacatc      300
aacgcgactg tcgaggccta tgccgcgctc cgctcgctgg gcttccgcga cgacgagccg      360
gcggtgcgcc gggcgcggga atggatcgag gccaaaggcg gcctgcgcaa catccgcgtc      420
ttcaccgcgt actggctggc actgatcggc gaatggccgt gggagaagac accgaacatc      480
ccgccggagg tgatctggtt tccgctctgg tttccgttct cgatctacaa tttcgcgcaa      540
tgggcccgcg ccaccttgat gccgatcgcc gtgctgtcgg cgcggcggcc gagccggccg      600
ctgccgccgg agaaccgcct cgatgcgctg tttccgcatg gacggaaggc gttcgactac      660
gaactgccgg tcaaggccgg cgccggcggc tgggacaggt tcttccgcgg cgccgacaag      720
gttctgcaca agctgcagaa cctcggaac cgtctcaatc tcggcctgtt ccgcccggcg      780
gccaccagcc gcgtgctgga atggatgatc cgccatcagg atttcgacgg cgctggggc      840
ggcatccagc cgccctggat ctacgggctg atggcgctct atgccgaagg ctatccgctc      900
aatcatcccg tgctcgcaaa gggcctcgac gcgctgaacg atcccggctg gcgcgtcgat      960
gtcggtgacg ccacctacat ccaggccacc aacagcccgg tctgggacac gatcctgacc     1020

```

```

ttgctcgcct tcgacgatgc cggcgtgctc ggcgactatc ccgaggccgt cgacaaggcg 1080
gtcgactggg tgctgcagcg gcagggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac 1200
gacaccgcgg tcgcgtgat cgcgctggcg ccactgcgcc acgatccgaa atggaaggcc 1260
aaagggatcg acgaggctat ccagctcggg gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcggt cgacaaggac aacaaccaga agatcctgac caagatcccc 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcacgcg ccacatcatc 1440
gaggcggttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggg ccgtggttcg gccgctgggg cgtcaattac 1560
gtctacggca ccggcgcggt gctgccggcg ctggccgcga tcggcgagga catgacccag 1620
ccctatatcg gccgcgcctg cgactggctg gttgcccatc agcaggccga tggcggctgg 1680
ggcgagagct gcgcctccta catggatgtc agcgcggtcg gccgcggcac cacaacggcc 1740
tcgcagaccg cctgggcgct gatggcgctg ctgcgccca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttcccgggc tacggcgctg gccagaccat caagctgaac 1920
gatcccgcg cgtcgcagcg gctgatgcag ggcccgaat tgtcccgcgc cttcatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcg ctcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

```

<210> 123

<211> 684

<212> PRT

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 123

```

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
Pro Pro Glu Val Ile Trp Phe Pro Leu Trp Phe Pro Phe Ser Ile Tyr		
	165	170
Asn Phe Ala Gln Trp Ala Arg Ala Thr Leu Met Pro Ile Ala Val Leu		
	180	185
Ser Ala Arg Arg Pro Ser Arg Pro Leu Pro Pro Glu Asn Arg Leu Asp		
	195	200
Ala Leu Phe Pro His Gly Arg Lys Ala Phe Asp Tyr Glu Leu Pro Val		
	210	215
Lys Ala Gly Ala Gly Gly Trp Asp Arg Phe Phe Arg Gly Ala Asp Lys		
225	230	235
Val Leu His Lys Leu Gln Asn Leu Gly Asn Arg Leu Asn Leu Gly Leu		
	245	250
Phe Arg Pro Ala Ala Thr Ser Arg Val Leu Glu Trp Met Ile Arg His		
	260	265
Gln Asp Phe Asp Gly Ala Trp Gly Gly Ile Gln Pro Pro Trp Ile Tyr		
	275	280
Gly Leu Met Ala Leu Tyr Ala Glu Gly Tyr Pro Leu Asn His Pro Val		
	290	295
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
Thr Ile Leu Thr Leu Leu Ala Phe Asp Asp Ala Gly Val Leu Gly Asp		
	340	345
Tyr Pro Glu Ala Val Asp Lys Ala Val Asp Trp Val Leu Gln Arg Gln		
	355	360
Val Arg Val Pro Gly Asp Trp Ser Met Lys Leu Pro His Val Lys Pro		
	370	375
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
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 Thr 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	635
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> 124		
<211> 2055		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成BjapSHC1衍生物		

<220>

<221> misc_feature

<222> (1377) .. (1377)

<223> n是a, c, g,或t

<400> 124

```

atgactgtga ccagctcggc ctccgcgcgt gcgacgcgcg acccgggaaa ttatcagact      60
gccctgcaat cgacgggtgcg cgcggcggcg gattggctga tcgccaacca gaagccggac      120
ggccattggg tcggccgcgc cgagtccaat gcctgcatgg aggcgcaatg gtgcctcgcg      180
ctgtggttca tggggctcga ggaccatccg ctgcgcaagc gcctgggcca gtcgtgctc      240
gacagccagc gcccggacgg cgcttggcag gtctatttcg gcgcccccaa tggcgacatc      300
aacgcgactg tcgaggccta tgccgcgtc cgctcgttgg gcttcgcga cgacgagccg      360
gcggtgcgcc gggcgcggga atggatcgag gccaaggcg gcctgcgcaa catccgcgtc      420
ttcaccgcgt actggctggc actgatcggc gaatggcgt gggagaagac accgaacatc      480
ccgccggagg tgatctggtt tccgctctgg tttccgttct cgatctacaa tttcgcgcaa      540
tgggcccgcg ccaccttgat gccgatcgcc gtgctgtcgg cgcggcggcc gagccggccg      600
ctgccgccgg agaaccgcct cgatgcgtg tttccgatg gacggaaggc gttcgactac      660
gaactgccgg tcaaggccgg cgcggcggc tgggacaggt tcttcgcgg cgccgacaag      720
gttctgcaca agctgcagaa cctcggcaac cgtctcaatc tcggcctgtt ccgcccggcg      780
gccaccagcc gcgtgctgga atggatgat cgccatcagg atttcgacgg cgcctggggc      840
ggcatccagc cgccctggat ctacgggctg atggcgtct atgccgaagg ctatccgctc      900
aatcatcccg tgctcgcaaa gggcctcgac gcgctgaacg atcccggctg gcgcgtcgat      960
gtcggtgacg ccacctacat ccaggccacc aacagcccgg tctgggacac gatcctgacc     1020
ttgctcgctt tcgacgatgc cggcgtgctc ggcgactatc ccgaggccgt cgacaaggcg     1080
gtcgactggg tgctgcagcg gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg     1140
catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac     1200
gacaccgcgg tcgcgtgat cgcgtggcg ccaactgcgc acgatccgaa atggaaggcc     1260
aaagggatcg acgaggctat ccagctcggg gtcgactggc tgatcggcat gcagagccag     1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagacnccg     1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggctc acgtcacgc ccacatcatc     1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgtggac     1500
tatattcgcc gtgagcagga gccgagcggg ccgtggttcg gccgctgggg cgtcaattac     1560
gtctacggca ccggcgcggg gctgccggcg ctggccgcga tcggcgagga catgaccag     1620
ccctatatcg gccgcgcctg cgactggctg gttgcccatc agcaggccga tggcggctgg     1680
ggcgagagct gcgcctccta catggatgtc agcgcggtcg gccgcggcac cacaacggcc     1740
tcgcagaccg cctgggcgct gatggcgtg ctgcgccca atcgccccca ggacaaggac     1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag     1860
ccggaattca ccggcaccgg tttcccgggc tacggcgtcg gccagaccat caagctgaac     1920
gatcccgcg tgctgcagcg gctgatgcag ggcccgaat tgtcccgcgc cttcatgctc     1980
cgctacggca tgtaccgcca ctacttccc ctgatggcg tcggccgcgc cctacgcccg     2040

```

cagagtcata gctag

2055

<210> 125

<211> 684

<212> PRT

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 125

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

				245					250					255			
Phe	Arg	Pro	Ala	Ala	Thr	Ser	Arg	Val	Leu	Glu	Trp	Met	Ile	Arg	His		
				260					265					270			
Gln	Asp	Phe	Asp	Gly	Ala	Trp	Gly	Gly	Ile	Gln	Pro	Pro	Trp	Ile	Tyr		
				275					280					285			
Gly	Leu	Met	Ala	Leu	Tyr	Ala	Glu	Gly	Tyr	Pro	Leu	Asn	His	Pro	Val		
				290					295					300			
Leu	Ala	Lys	Gly	Leu	Asp	Ala	Leu	Asn	Asp	Pro	Gly	Trp	Arg	Val	Asp		
305						310				315				320			
Val	Gly	Asp	Ala	Thr	Tyr	Ile	Gln	Ala	Thr	Asn	Ser	Pro	Val	Trp	Asp		
				325					330					335			
Thr	Ile	Leu	Thr	Leu	Leu	Ala	Phe	Asp	Asp	Ala	Gly	Val	Leu	Gly	Asp		
				340					345					350			
Tyr	Pro	Glu	Ala	Val	Asp	Lys	Ala	Val	Asp	Trp	Val	Leu	Gln	Arg	Gln		
				355					360					365			
Val	Arg	Val	Pro	Gly	Asp	Trp	Ser	Met	Lys	Leu	Pro	His	Val	Lys	Pro		
				370					375					380			
Gly	Gly	Trp	Ala	Phe	Glu	Tyr	Ala	Asn	Asn	Tyr	Tyr	Pro	Asp	Thr	Asp		
385						390				395				400			
Asp	Thr	Ala	Val	Ala	Leu	Ile	Ala	Leu	Ala	Pro	Leu	Arg	His	Asp	Pro		
				405					410					415			
Lys	Trp	Lys	Ala	Lys	Gly	Ile	Asp	Glu	Ala	Ile	Gln	Leu	Gly	Val	Asp		
				420					425					430			
Trp	Leu	Ile	Gly	Met	Gln	Ser	Gln	Gly	Gly	Gly	Trp	Gly	Ala	Phe	Asp		
				435					440					445			
Lys	Asp	Asn	Asn	Gln	Lys	Ile	Leu	Thr	Lys	Ile	Pro	Phe	Cys	Asp	Tyr		
				450					455					460			
Gly	Glu	Ala	Leu	Asp	Pro	Pro	Ser	Val	Asp	Val	Thr	Ala	His	Ile	Ile		
465						470				475				480			
Glu	Ala	Phe	Gly	Lys	Leu	Gly	Ile	Ser	Arg	Asn	His	Pro	Ser	Met	Val		
				485					490					495			
Gln	Ala	Leu	Asp	Tyr	Ile	Arg	Arg	Glu	Gln	Glu	Pro	Ser	Gly	Pro	Trp		
				500					505					510			
Phe	Gly	Arg	Trp	Gly	Val	Asn	Tyr	Val	Tyr	Gly	Thr	Gly	Ala	Val	Leu		
				515					520					525			
Pro	Ala	Leu	Ala	Ala	Ile	Gly	Glu	Asp	Met	Thr	Gln	Pro	Tyr	Ile	Gly		
				530					535					540			
Arg	Ala	Cys	Asp	Trp	Leu	Val	Ala	His	Gln	Gln	Ala	Asp	Gly	Gly	Trp		
545						550				555				560			

Gly	Glu	Ser	Cys	Ala	Ser	Tyr	Met	Asp	Val	Ser	Ala	Val	Gly	Arg	Gly	
				565					570					575		
Thr	Thr	Thr	Ala	Ser	Gln	Thr	Ala	Trp	Ala	Leu	Met	Ala	Leu	Leu	Ala	
				580					585					590		
Ala	Asn	Arg	Pro	Gln	Asp	Lys	Asp	Ala	Ile	Glu	Arg	Gly	Cys	His	Trp	
				595					600					605		
Leu	Val	Glu	Arg	Gln	Ser	Ala	Gly	Thr	Trp	Asp	Glu	Pro	Glu	Phe	Thr	
				610					615					620		
Gly	Thr	Gly	Phe	Pro	Gly	Tyr	Gly	Val	Gly	Gln	Thr	Ile	Lys	Leu	Asn	
				625					630					635		
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> 126

<211> 2055

<212> DNA

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 126

atgactgtga ccagctcggc ctccgcgcgt gcgacgcgcg acccgggaaa ttatcagact	60
gccctgcaat cgacggtgcg cgcgggcgcg gattggctga tcgccaacca gaagccggac	120
ggccattggg tcggccgcgc cgagtccaat gcctgcatgg aggcgcaatg gtgcctcgcg	180
ctgtggttca tggggctcga ggaccatccg ctgcgcaagc gcctgggcca gtgcgtgctc	240
gacagccagc gcccggacgg cgcttggcag gtctatttcg gcgcccccaa tggcgacatc	300
aacgcgactg tcgaggccta tgccgcgctc cgctcgtcgg gcttcgcga cgacgagccg	360
gcggtgcgcc gggcgcgga atggatcgag gccaaaggcg gcctgcgcaa catccgcgtc	420
ttcacccgct actggctggc actgatcggc gaatggccgt gggagaagac accgaacatc	480
ccgccggagg tgatctggtt tccgctctgg ttccggttct cgatctacaa ttctgcgcaa	540
tgggcccgcg ccaccttgat gccgatcgcc gtgctgtcgg cgcgggcgcc gagccggccg	600
ctgccgccgg agaaccgcct cgatgcgctg ttccgcgatg gacggaaggc gttcgactac	660
gaactgccgg tcaaggccgg cgccggcggc tgggacaggt tcttcgcgg cgccgacaag	720
gttctgcaca agctgcagaa cctcggaac cgtctcaatc tcggcctgtt ccgccggcg	780
gccaccagcc gcgtgctgga atggatgatc cgccatcagg atttcgacgg cgcttggggc	840
ggcatccagc cgccctggat ctacgggctg atggcgctct atgccgaagg ctatccgctc	900
aatcatcccg tgctcgcaaa gggcctcgac gcgctgaacg atcccggtg gcgcgtcgat	960


```

gtcggtgacg ccacctacat ccaggccacc aacagcccgg tctgggacac gatcctgacc 1020
ttgctcgcct tcgacgatgc cggcgtgctc ggcgactatc ccgaggccgt cgacaaggcg 1080
gtcgactggg tgctgcagcg gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
catgtcaagc ccggcggctg ggcgttcgaa tacccaaca actactatcc cgacacggac 1200
gacaccgcgg tcgcgctgat cgcgctggcg cactgcgcc acgatccgaa atggaaggcc 1260
aaagggatcg acgaggctat ccagctcggg gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgct cgacaaggac aacaaccaga agatcctgac caagatcccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggctc acgtcacgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggg ccgtggttcg gccgctgggg cgtcaattac 1560
gtctacggca ccggcgcggt gctgccggcg ctggccgcga tcggcgagga catgaccag 1620
ccctatatcg gccgcgcctg cgactggctg gttgcccatc agcaggccga tggcggctgg 1680
ggcgagagct gcgcctccta catggatgtc agcgcggtcg gccgcggcac cacaacggcc 1740
tcgcagaccg cctgggcgct gatggcgtg ctgccgcga atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcc aytggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttcccgggc tacggcgtcg gccagaccat caagctgaac 1920
gatcccgcg tgctgcagcg gctgatgcag ggcccgaat tgtcccgcgc cttcatgtc 1980
cgctacggca tgtaccgcca ctacttccc ctgatggcg tcggccgcgc cctacggccg 2040
cagagtcata gctag 2055

```

<210> 127

<211> 684

<212> PRT

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 127

```

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
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
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 Tyr Pro Gly Tyr Gly Val Gly Gln Thr Ile Lys Leu Asn		
625	630	635
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>	128	
<211>	2055	
<212>	DNA	
<213>	人工序列	
<220>		

<223> 合成BjapSHC1衍生物

<400> 128

```

atgactgtga ccagctcggc ctccgcgcgt gcgacgcgcg acccgggaaa ttatcagact    60
gccctgcaat cgacggtgcg cgcggcggcg gattggctga tcgccaacca gaagccggac    120
ggccattggg tcggccgcgc cgagtccaat gcctgcatgg aggcgcaatg gtgcctcgcg    180
ctgtggttca tggggctcga ggaccatccg ctgcgcaagc gcctgggcca gtgcgtgctc    240
gacagccagc gcccggacgg cgcctggcag gtctatttcg gcgcccccaa tggcgacatc    300
aacgcgactg tcgaggccta tgccgcgctc cgctcgctgg gcttcgcga cgacgagccg    360
gcggtgcgcc gggcgcggga atggatcgag gccaaaggcg gcctgcgcaa catccgcgtc    420
ttcaccgcgt actggctggc actgatcggc gaatggccgt gggagaagac accgaacatc    480
ccgccggagg tgatctggtt tccgctctgg ttccggttct cgatctacaa ttctgcgcaa    540
tgggcccgcg ccacctgat gccgatcgcc gtgctgtcgg cgcggcggcc gagccggccg    600
ctgccgccgg agaaccgcct cgatgcgctg ttccgcatg gacggaaggc gttcgactac    660
gaactgccgg tcaaggccgg cgcggcggc tgggacaggt tcttcgcgg cgcgcacaag    720
gttctgcaca agctgcagaa cctcggcaac cgtctcaatc tcggcctgtt ccgcccggcg    780
gccaccagcc gcgtgctgga atggatgata cgccatcagg atttcgacgg cgcctggggc    840
ggcatccagc cgccttgat ctacgggctg atggcgctct atgccgaagg ctatccgctc    900
aatcatcccg tgctcgcaaa gggcctcgac gcgtgaacg atcccggctg gcgcgtcgat    960
gtcggtgacg ccacctacat ccaggccacc aacagcccgg tctgggacac gatcctgacc   1020
ttgctgcct tcgacgatgc cggcgtgctc ggcgactatc ccgaggccgt cgacaaggcg   1080
gtcgactggg tgctgcagcg gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg   1140
catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac   1200
gacaccgcgg tcgcgtgat cgcgtggcg ccactgcgcc acgatccgaa atggaaggcc   1260
aaagggatcg acgaggctat ccagctcggg gtcgactggc tgatcggcat gcagagccag   1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagatcccg   1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc   1440
gaggcggttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgtggac   1500
tatattcgcc gtgagcagga gccgagcggg ccgtggttcg gccgctgggg cgtcaattac   1560
gtctacggca ccggcgcggg gctgccggcg ctggccgcga tcggcgagga catgaccagc   1620
ccctatatcg gccgcgcctg cgactggctg gttgccatc agcaggccga tggcggctgg   1680
ggcgagagct gcgcctccta catggatgtc agcgcggctc gccgcggcac cacaacggcc   1740
tcgcagaccg cctgggcgct gatggcgctg ctgcgcgcca atcgccccca ggacaaggac   1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag   1860
ccggaattca ccggcaccgg ttayccgggc tacggcgtcg gccagaccat caagctgaac   1920
gatcccgcg cgtcgagcg gctgatgcag ggcccgaat tgtcccgcgc ctteatgctc   1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg   2040
cagagtcata gctag                                     2055

```

<210> 129

<211> 684

<212> PRT

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 129

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	Arg
	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
Thr Ile Leu Thr Leu Leu Ala Phe Asp Asp Ala Gly Val Leu Gly Asp		
	340	345
Tyr Pro Glu Ala Val Asp Lys Ala Val Asp Trp Val Leu Gln Arg Gln		
	355	360
Val Arg Val Pro Gly Asp Trp Ser Met Lys Leu Pro His Val Lys Pro		
	370	375
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
Lys Trp Lys Ala Lys Gly Ile Asp Glu Ala Ile Gln Leu Gly Val Asp		
	420	425
Trp Leu Ile Gly Met Gln Ser Gln Gly Gly Gly Trp Gly Ala Phe Asp		
	435	440
Lys Asp Asn Asn Gln Lys Ile Leu Thr Lys Thr Pro Phe Cys Asp Tyr		
	450	455
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
Gln Ala Leu Asp Tyr Ile Arg Arg Glu Gln Glu Pro Ser Gly Pro Trp		
	500	505
Phe Gly Arg Trp Gly Val Asn Tyr Val Tyr Gly Thr Gly Ala Val Leu		
	515	520
Pro Ala Leu Ala Ala Ile Gly Glu Asp Met Thr Gln Pro Tyr Ile Gly		
	530	535
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
Thr Thr Thr Ala Ser Gln Thr Ala Trp Ala Leu Met Ala Leu Leu Ala		

580	585	590
Ala Asn Arg Pro Gln Asp Lys Asp Ala Ile Glu Arg Gly Cys Met Trp		
595	600	605
Leu Val Glu Arg Gln Ser Ala Gly Thr Trp Asp Glu Pro Glu Phe Thr		
610	615	620
Gly Thr Gly Phe Pro Gly Tyr Gly Val Gly Gln Thr Ile Lys Leu Asn		
625	630	635
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> 130

<211> 2055

<212> DNA

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<220>

<221> misc_feature

<222> (430) .. (432)

<223> ngn是agr或cgn

<220>

<221> misc_feature

<222> (1377) .. (1377)

<223> n是a, c, g,或t

<400> 130

atgactgtga ccagctcggc ctccgcgcgt gcgacgcgcg acccgggaaa ttatcagact	60
gccctgcaat cgacggtgcg cgcggcggcg gattggctga tcgccaacca gaagccggac	120
ggccattggg tcggccgcgc cgagtccaat gcctgcatgg aggcgcaatg gtgcctcgcg	180
ctgtggttca tggggctcga ggaccatccg ctgcgcaagc gcctgggcca gtcgtgctc	240
gacagccagc gcccggacgg cgcttggcag gtctatttcg gcgccccaa tggcgacatc	300
aacgcgactg tcgaggccta tgccgcgtc cgctcgttgg gcttccgcga cgacgagccg	360
gcggtgcgcc gggcgcgga atggatcgag gccaaaggcg gcctgcgcaa catccgcgtc	420
ttcaccgcgn gntggctggc actgatcggc gaatggcgt gggagaagac accgaacatc	480
ccgccggagg tgatctggtt tccgctctgg tttccgttct cgatctacaa tttcgcgcaa	540
tgggcccgcg ccacctgat gccgatcgcc gtgctgtcgg cgcggcggcc gagccggccg	600
ctgccgccgg agaaccgcct cgatgcgctg tttccgcatg gacggaaggc gttcgactac	660

```

gaactgccgg tcaaggccgg cgccggcggc tgggacaggt tcttccgcgg cgccgacaag 720
gttctgcaca agctgcagaa cctcggcaac cgtctcaatc tcggcctggt cgcgccggcg 780
gccaccagcc gcgtgctgga atggatgata cgccatcagg atttcgacgg cgcctggggc 840
ggcatccagc cgccctggat ctacgggctg atggcgctct atgccgaagg ctatccgctc 900
aatcatcccc tgctcgcaaa gggcctcgac gcgctgaacg atccccggctg gcgctcgat 960
gtcggtgacg ccacctacat ccaggccacc aacagcccgg tctgggacac gatcctgacc 1020
ttgctcgctt tcgacgatgc cggcgtgctc ggcgactatc ccgaggccgt cgacaaggcg 1080
gtcgaactgg tgctgcagcg gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac 1200
gacaccgcgg tcgctcgatc cgcgtggcg ccaactgcgc acgatccgaa atggaaggcc 1260
aaagggatcg acgaggttat ccagctcggg gtgcaactgg tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagacnccg 1380
ttctgcgatt atggcgagcg gctcgatccg cctcggctcg acgtaccgc ccacatcctc 1440
gaggcgcttc gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgtggac 1500
tatattcgcc gtgagcagga gccgagcggg ccgtggttcg gccgctgggg cgtcaattac 1560
gtctacggca ccggcgcggg gctgccggcg ctggccgcga tcggcgagga catgaccag 1620
ccctatatcg gccgcgcctg cgactggctg gttgcccatc agcaggccga tggcggtgg 1680
ggcgagagct gcgcctccta catggatgtc agcgcggtcg gccgcggcac cacaacggcc 1740
tcgcagaccg cctgggcgct gatggcgctg ctgcgccga atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttccggggc tacggcgctg gccagaccat caagctgaac 1920
gatcccgcg cgtcgcagcg gctgatgcag ggcccgaat tgtcccgcg cttcatgctc 1980
cgctacggca tgtaccgca ctacttcccg ctgatggcg tcggccgcgc cctacgccc 2040
cagagtcata gctag 2055

```

<210> 131

<211> 684

<212> PRT

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 131

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 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 Arg			
	130	135	140
Trp Leu Ala Leu Ile Gly Glu Trp Pro Trp Glu Lys Thr Pro Asn Ile			
	145	150	155
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
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		400
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 Thr 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	635
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>	132	
<211>	2055	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成BjapSHC1衍生物	
<220>		
<221>	misc_feature	
<222>	(430) .. (432)	
<223>	ngn是agr或cgn	
<220>		
<221>	misc_feature	
<222>	(1377) .. (1377)	
<223>	n是a, c, g,或t	
<400>	132	
atgactgtga	ccagctcggc	ctccgcgcgt
gccctgcaat	cgacggtgcg	cgcggcggcg
ggccattggg	tcggccgcgc	cgagtccaat
ctgtggttca	tggggctcga	ggaccatccg
gacagccagc	gcccggacgg	cgcttggcag
aacgcgactg	tcgaggccta	tgccgcgctc
gcggtgcgcc	gggcgcggga	atggatcgag
ttcaccgcgn	gntggctggc	actgatcggc
ccgccggagg	tgatctgggt	tccgctctgg
tgggcccgcg	ccacctgat	gccgatcgcc
ctgccgccgg	agaaccgcct	cgatgcgctg
gaactgccgg	tcaaggccgg	cgccggcggc
gttctgcaca	agctgcagaa	cctcggcaac
gccaccagcc	gcgtgctgga	atggatgata
ggcatccagc	cgccctggat	ctacgggctg
aatcatcccc	tgctcgcaaa	gggcctcgac
gtcggtgacg	ccacctacat	ccagccacc
ttgctcgctt	tcgacgatgc	cggcgtgctc
gtcgactggg	tgctgcagcg	gcaggtgcgc
catgtcaagc	ccggcggtg	ggcgttcgaa
gacaccgcgg	tcgcgctgat	cgcgctggcg
aaagggatcg	acgaggctat	ccagctcggt
ggcggcggct	ggggcgcggt	cgacaaggac
ttctgcgatt	atggcgaggc	gctcgatccg

gaggcggttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac	1500
tatatctgcc gtgagcagga gccgagcggg ccgtgggttcg gccgctgggg cgtcaattac	1560
gtctacggca ccggcgcggt gctgccggcg ctggccgcga tcggcgagga catgaccag	1620
ccctatatcg gccgcgcctg cgactggctg gttgcccatc agcaggccga tggcggctgg	1680
ggcgagagct ggcctccta catggatgtc agcgcggtcg gccgcggcac cacaacggcc	1740
tcgcagaccg cctgggcgct gatggcgctg ctgcgccca atcgccccca ggacaaggac	1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag	1860
ccggaattca ccggcaccgg tttccgggc tacggcgctc gccagaccat caagctgaac	1920
gatcccgcg cgtgcagcg gctgatgcag ggcccgaat tgtcccgcg cttcatgtc	1980
cgctacggca tgtaccgca ctacttccg ctgatggcg tcggccgcgc cctacgccc	2040
cagagtcata gctag	2055
<210> 133	
<211> 684	
<212> PRT	
<213> 人工序列	
<220>	
<223> 合成BjapSHC1衍生物	
<400> 133	
Met Thr Val Thr Ser Ser Ala Ser Ala Arg Ala Thr Arg Asp Pro Gly	
1 5 10 15	
Asn Tyr Gln Thr Ala Leu Gln Ser Thr Val Arg Ala Ala Ala Asp Trp	
20 25 30	
Leu Ile Ala Asn Gln Lys Pro Asp Gly His Trp Val Gly Arg Ala Glu	
35 40 45	
Ser Asn Ala Cys Met Glu Ala Gln Trp Cys Leu Ala Leu Trp Phe Met	
50 55 60	
Gly Leu Glu Asp His Pro Leu Arg Lys Arg Leu Gly Gln Ser Leu Leu	
65 70 75 80	
Asp Ser Gln Arg Pro Asp Gly Ala Trp Gln Val Tyr Phe Gly Ala Pro	
85 90 95	
Asn Gly Asp Ile Asn Ala Thr Val Glu Ala Tyr Ala Ala Leu Arg Ser	
100 105 110	
Leu Gly Phe Arg Asp Asp Glu Pro Ala Val Arg Arg Ala Arg Glu Trp	
115 120 125	
Ile Glu Ala Lys Gly Gly Leu Arg Asn Ile Arg Val Phe Thr Arg Tyr	
130 135 140	
Trp Leu Ala Leu Ile Gly Glu Trp Pro Trp Glu Lys Thr Pro Asn Ile	
145 150 155 160	
Pro Pro Glu Val Ile Trp Phe Pro Leu Trp Phe Pro Phe Ser Ile Tyr	

	165		170		175
Asn Phe Ala Gln Trp Ala Arg Ala Thr Leu Met Pro Ile Ala Val Leu					
	180		185		190
Ser Ala Arg Arg Pro Ser Arg Pro Leu Pro Pro Glu Asn Arg Leu Asp					
	195		200		205
Ala Leu Phe Pro His Gly Arg Lys Ala Phe Asp Tyr Glu Leu Pro Val					
	210		215		220
Lys Ala Gly Ala Gly Gly Trp Asp Arg Phe Phe Arg Gly Ala Asp Lys					
225		230		235	240
Val Leu His Lys Leu Gln Asn Leu Gly Asn Arg Leu Asn Leu Gly Leu					
	245		250		255
Phe Arg Pro Ala Ala Thr Ser Arg Val Leu Glu Trp Met Ile Arg His					
	260		265		270
Gln Asp Phe Asp Gly Ala Trp Gly Gly Ile Gln Pro Pro Trp Ile Tyr					
	275		280		285
Gly Leu Met Ala Leu Tyr Ala Glu Gly Tyr Pro Leu Asn His Pro Val					
	290		295		300
Leu Ala Lys Gly Leu Asp Ala Leu Asn Asp Pro Gly Trp Arg Val Asp					
305		310		315	320
Val Gly Asp Ala Thr Tyr Ile Gln Ala Thr Asn Ser Pro Val Trp Asp					
	325		330		335
Thr Ile Leu Thr Leu Leu Ala Phe Asp Asp Ala Gly Val Leu Gly Asp					
	340		345		350
Tyr Pro Glu Ala Val Asp Lys Ala Val Asp Trp Val Leu Gln Arg Gln					
	355		360		365
Val Arg Val Pro Gly Asp Trp Ser Met Lys Leu Pro His Val Lys Pro					
	370		375		380
Gly Gly Trp Ala Phe Glu Tyr Ala Asn Asn Tyr Tyr Pro Asp Thr Asp					
385		390		395	400
Asp Thr Ala Val Ala Leu Ile Ala Leu Ala Pro Leu Arg His Asp Pro					
	405		410		415
Lys Trp Lys Ala Lys Gly Ile Asp Glu Ala Ile Gln Leu Gly Val Asp					
	420		425		430
Trp Leu Ile Gly Met Gln Ser Gln Gly Gly Gly Trp Gly Ala Phe Asp					
	435		440		445
Lys Asp Asn Asn Gln Lys Ile Leu Thr Lys Ile Pro Phe Cys Asp Tyr					
	450		455		460
Gly Glu Ala Leu Asp Pro Pro Ser Val Asp Val Thr Ala His Ile Ile					
465		470		475	480

Glu Ala Phe Gly Lys Leu Gly Ile Ser Arg Asn His Pro Ser Met Val	
485	490 495
Gln Ala Leu Asp Tyr Ile Arg Arg Glu Gln Glu Pro Ser Gly Pro Trp	
500	505 510
Phe Gly Arg Trp Gly Val Asn Tyr Val Tyr Gly Thr Gly Ala Val Leu	
515	520 525
Pro Ala Leu Ala Ala Ile Gly Glu Asp Met Thr Gln Pro Tyr Ile Gly	
530	535 540
Arg Ala Cys Asp Trp Leu Val Ala His Gln Gln Ala Asp Gly Gly Trp	
545	550 555 560
Gly Glu Ser Cys Ala Ser Tyr Met Asp Val Ser Ala Val Gly Arg Gly	
565	570 575
Thr Thr Thr Ala Ser Gln Thr Ala Trp Ala Leu Met Ala Leu Leu Ala	
580	585 590
Ala Asn Arg Pro Gln Asp Lys Asp Ala Ile Glu Arg Gly Cys Met Trp	
595	600 605
Leu Val Glu Arg Gln Ser Ala Gly Thr Trp Asp Glu Pro Glu Phe Thr	
610	615 620
Gly Thr Gly Tyr Pro Gly Tyr Gly Val Gly Gln Thr Ile Lys Leu Asn	
625	630 635 640
Asp Pro Ala Leu Ser Gln Arg Leu Met Gln Gly Pro Glu Leu Ser Arg	
645	650 655
Ala Phe Met Leu Arg Tyr Gly Met Tyr Arg His Tyr Phe Pro Leu Met	
660	665 670
Ala Leu Gly Arg Ala Leu Arg Pro Gln Ser His Ser	
675	680

<210> 134

<211> 2055

<212> DNA

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 134

atgactgtga ccagctcggc ctccgcgcgt gcgacgcgcg acccgggaaa ttatcagact	60
gccctgcaat cgacggtgcg cgcggcggcg gattggctga tcgccaacca gaagccggac	120
ggccattggg tcggccgcgc cgagtccaat gcctgcatgg aggcgcaatg gtgcctcgcg	180
ctgtggttca tggggctcga ggaccatccg ctgcgcaagc gcctgggcca gtcgctgctc	240
gacagccagc gcccggacgg cgcttggcag gtctatttcg gcgcccccaa tggcgacatc	300
aacgcgactg tcgaggccta tgccgcgctc cgctcgctgg gcttccgcga cgacgagccg	360

gcggtgcgcc gggcgcgggga atggatcgag gccaaaggcg gcctgcgcaa catccgcgtc 420
 ttcacccgct actggctggc actgatcggc gaatggccgt gggagaagac accgaacatc 480
 ccgccggagg tgatctggtt tccgctctgg tttccgttct cgatctacaa tttcgcgcaa 540
 tgggcccgcg ccaccttgat gccgatcgcc gtgctgtcgg cgcggcggcc gagccggccg 600
 ctgccgccgg agaaccgcct cgatgcgctg tttccgcatg gacggaaggc gttcgactac 660
 gaactgccgg tcaaggccgg cgccggcggc tgggacaggt tcttccgagg cgccgacaag 720
 gttctgcaca agctgcagaa cctcggcaac cgtctcaatc tcggcctggt ccgcccggcg 780
 gccaccagcc gcgtgctgga atggatgatc cgccatcagg atttcgacgg cgcctggggc 840
 ggcatccagc cgccctggat ctacgggctg atggcgctct atgccgaagg ctatccgctc 900
 aatcatcccc tgctcgcaaa gggcctcgac gcgtgaacg atcccggctg gcgcgtcgat 960
 gtcggtgacg ccacctacat ccaggccacc aacagcccgg tctgggacac gatcctgacc 1020
 ttgctcgctt tcgacgatgc cggcgtgctc ggcgactatc ccgaggccgt cgacaaggcg 1080
 gtcgactggg tgctgcagcg gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
 catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac 1200
 gacaccgcgg tcgcgtgat cgcgtggtcg ccaactgcgc acgatccgaa atggaaggcc 1260
 aaagggatcg acgaggctat ccagctcggg gtcgactggc tgatcggcat gcagagccag 1320
 ggcggcggct ggggcgcggt cgacaaggac aacaaccaga agatcctgac caagatcccg 1380
 ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
 gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgtggac 1500
 tatattcgcc gtgagcagga gccgagcggg ccgtggttcg gccgctgggg cgtcaattac 1560
 gtctacggca ccggcgcggt gctgccggcg ctggccgcga tcggcgagga catgaccag 1620
 ccctatatcg gccgcgcctg cgactggctg gttgccatc agcaggccga tggcggctgg 1680
 ggcgagagct gcgcctccta catggatgtc agcgcggtcg gccgcggcac cacaacggcc 1740
 tcgcagaccg cctgggcgct gatggcgctg ctgcgcgcca atcgccccca ggacaaggac 1800
 gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
 ccggaattca ccggcaccgg ttayccgggc tacggcgtcg gccagaccat caagctgaac 1920
 gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaa tgtcccgcgc cttcatgctc 1980
 cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
 cagagtcata gctag 2055

<210> 135

<211> 684

<212> PRT

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 135

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
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 Leu 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
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
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	635
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>	136			
<211>	2055			
<212>	DNA			
<213>	人工序列			
<220>				
<223>	合成BjapSHC1衍生物			
<220>				
<221>	misc_feature			
<222>	(421) .. (423)			
<223>	ntn是ttr或ctn			
<400>	136			
	atgactgtga ccagctcggc ctccgcgcgt gcgacgcgcg acccgggaaa ttatcagact		60	
	gccctgcaat cgacggtgcg cgcggcggcg gattggctga tcgccaacca gaagccggac		120	
	ggccattggg tcggccgcgc cgagtccaat gcctgcatgg aggcgcaatg gtgcctcgcg		180	
	ctgtggttca tggggctcga ggaccatccg ctgcgcaagc gcctgggcca gtcgtgctc		240	
	gacagccagc gcccggacgg cgcctggcag gtctatttgc gcgcccccaa tggcgacatc		300	
	aacgcgactg tcgaggccta tgccgcgctc cgctcgttgg gcttccgcga cgacgagccg		360	
	gcggtgcgcc gggcgcgggga atggatcgag gccaaaggcg gcctgcgcaa catccgcgtc		420	
	ntnaccgcgt actggctggc actgatcggc gaatggccgt gggagaagac accgaacatc		480	
	ccgccggagg tgatctggtt tccgctctgg tttccgttct cgatctacaa tttcgcgcaa		540	
	tgggcccgcg ccaccttgat gccgatcgcc gtgctgtcgg cgcggcggcc gagccggccg		600	
	ctgccgccgg agaaccgcct cgatgcgctg tttccgcatg gacggaaggc gttcgactac		660	
	gaactgccgg tcaaggccgg cgccggcggc tgggacaggt tcttccgcgg cgccgacaag		720	
	gttctgcaca agctgcagaa cctcggcaac cgtctcaatc tcggcctgtt ccgcccggcg		780	
	gccaccagcc gcgtgctgga atggatgata cgccatcagg atttcgacgg cgcctggggc		840	
	ggcatccagc cgccctggat ctacgggctg atggcgtctt atgccgaagg ctatccgctc		900	
	aatcatcccg tgctcgcaaa gggcctcgac gcgtgaacg atcccggtg gcgcgtcgat		960	
	gtcggtgacg ccacctacat ccaggccacc aacagcccgg tctgggacac gatcctgacc		1020	
	ttgctcgctt tcgacgatgc cggcgtgctc ggcgactatc ccgaggccgt cgacaaggcg		1080	
	gtcgactggg tgctgcagcg gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg		1140	
	catgtcaagc ccggcggtg ggcgttcgaa tacgccaaca actactatcc cgacacggac		1200	
	gacaccgcgg tcgcgtgat cgcgtggcg ccaactgcgc acgatccgaa atggaaggcc		1260	
	aaagggatcg acgaggctat ccagctcggg gtcgactggc tgatcggcat gcagagccag		1320	
	ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagatcccc		1380	

```

ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcggttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggg ccgtgggttcg gccgctgggg cgtcaattac 1560
gtctacggca ccggcgcggt gctgccggcg ctggccgcga tcggcgagga catgaccag 1620
ccctatatcg gccgcgcctg cgactggctg gttgcccatc agcaggccga tggcggctgg 1680
ggcgagagct ggcctccta catggatgtc agcgcggtcg gccgcggcac cacaacggcc 1740
tcgcagaccg cctgggcgct gatggcgctg ctgccgcga atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttccgggc tacggcgctg gccagaccat caagctgaac 1920
gatcccgcg tgctgcagcg gctgatgcag ggcccgaat tgtcccgcg cttcatgctc 1980
cgctacggca tgaccgcca ctacttccg ctgatggcg tcggccgcgc cctacgccc 2040
cagagtcata gctag 2055

```

<210> 137

<211> 684

<212> PRT

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 137

```

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

```

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

465	470	475	480
Glu Ala Phe Gly Lys Leu Gly Ile Ser Arg Asn His Pro Ser Met Val			
	485	490	495
Gln Ala Leu Asp Tyr Ile Arg Arg Glu Gln Glu Pro Ser Gly Pro Trp			
	500	505	510
Phe Gly Arg Trp Gly Val Asn Tyr Val Tyr Gly Thr Gly Ala Val Leu			
	515	520	525
Pro Ala Leu Ala Ala Ile Gly Glu Asp Met Thr Gln Pro Tyr Ile Gly			
	530	535	540
Arg Ala Cys Asp Trp Leu Val Ala His Gln Gln Ala Asp Gly Gly Trp			
545	550	555	560
Gly Glu Ser Cys Ala Ser Tyr Met Asp Val Ser Ala Val Gly Arg Gly			
	565	570	575
Thr Thr Thr Ala Ser Gln Thr Ala Trp Ala Leu Met Ala Leu Leu Ala			
	580	585	590
Ala Asn Arg Pro Gln Asp Lys Asp Ala Ile Glu Arg Gly Cys His 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 Tyr Pro Gly Tyr Gly Val Gly Gln Thr Ile Lys Leu Asn			
625	630	635	640
Asp Pro Ala Leu Ser Gln Arg Leu Met Gln Gly Pro Glu Leu Ser Arg			
	645	650	655
Ala Phe Met Leu Arg Tyr Gly Met Tyr Arg His Tyr Phe Pro Leu Met			
	660	665	670
Ala Leu Gly Arg Ala Leu Arg Pro Gln Ser His Ser			
	675	680	

<210> 138

<211> 2055

<212> DNA

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 138

atgactgtga ccagctcggc ctccgcgcgt gcgacgcgcg acccgggaaa ttatcagact	60
gccctgcaat cgacggtgcg cgcggcggcg gattggctga tcgccaacca gaagccggac	120
ggccattggg tcggccgcgc cgagtccaat gcctgcatgg aggcgcaatg gtgcctcgcg	180
ctgtggttca tggggctcga ggaccatccg ctgcgcaagc gcctgggcca gtcgctgctc	240
gacagccagc gcccggacgg cgcctggcag gtctatttcg gcgcccccaa tggcgacatc	300

```

aacgcgactg tcgaggccta tgccgcgctc cgctcgctgg gcttccgca cgacgagccg 360
gcggtgcgcc gggcgcgga atggatcgag gccaaaggcg gcctgcgcaa catccgcgtc 420
ttcaccgcgt actggctggc actgatcggc gaatggccgt gggagaagac accgaacatc 480
ccgccggagg tgatctggtt tccgctctgg tttccgttct cgatctacaa tttcgcgcaa 540
tgggcccgcg ccaccttgat gccgatcgcc gtgctgtcgg cgcggcggcc gagccggccg 600
ctgccgccgg agaaccgcct cgatgcgctg tttccgcatg gacggaaggc gttcgactac 660
gaactgccgg tcaaggccgg cgccggcggc tgggacaggt tcttccgagg cgccgacaag 720
gttctgcaca agctgcagaa cctcggcaac cgtctcaatc tcggcctggt ccgcccgggc 780
gccaccagcc gcgtgctgga atggatgatc cgccatcagg atttcgacgg cgcctggggc 840
ggcatccagc cgccctggat ctacgggctg atggcgctct atgccgaagg ctatccgctc 900
aatcatcccg tgctcgcaaa gggcctcgac gcgtgaacg atcccgctg gcgcgtcgat 960
gtcggtagcg ccacctacat ccaggccacc aacagcccgg tctgggacac gatcctgacc 1020
ttgctgcct tcgacgatgc cggcgtgctc ggcgactatc ccgaggccgt cgacaaggcg 1080
gtcgactggg tgctgcagcg gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
catgtcaagc ccggcggtg ggcgttcgaa tacccaaca actactatcc cgacacggac 1200
gacaccgcgg tcgcgtgat cgcgtggcg ccatgcgcc acgatccgaa atggaaggcc 1260
aaagggatcg acgaggctat ccagctcggt gtcgactggc tgatcgcat gcagagccag 1320
ggcggcggtt ggggcggtt cgacaaggac aacaaccaga agatcctgac caagatcccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtaccgc ccacatcatc 1440
gaggcggttc gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgtggac 1500
tatattcgcc gtgagcagga gccgagcggc ccgtggttcg gccgctgggg cgtcaattac 1560
gtctacggca ccggcgcggt gctgccggcg ctggccgca tcggcgagga catgaccag 1620
ccctatatcg gccgcgcctg cgactggctg gttgccatc agcaggccga tggcggtg 1680
ggcgagagct gcgcctccta catggatgtc agcgcggtcg gccgcggcac cacaacggcc 1740
tcgcagaccg cctgggcgct gatggcgctg ctgcgcgcca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcca ytggtggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg ttayccgggc tacggcgctc gccagaccat caagctgaac 1920
gatcccgcgc tgtcgagcg gctgatgcag ggcccgaat tgtcccgcg cttcatgctc 1980
cgctacggca tgtaccgcca ctacttccc ctgatggcg tcggccgcgc cctacgccc 2040
cagagtcata gctag 2055

```

<210> 139

<211> 684

<212> PRT

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 139

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
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 Leu 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
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
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	400
Asp Thr Ala Val Ala Leu Ile Ala Leu Ala Pro Leu Arg His Asp Pro					
	405		410		415
Lys Trp Lys Ala Lys Gly Ile Asp Glu Ala Ile Gln Leu Gly Val Asp					
	420		425		430
Trp Leu Ile Gly Met Gln Ser Gln Gly Gly Gly Trp Gly Ala Phe Asp					
	435		440		445
Lys Asp Asn Asn Gln Lys Ile Leu Thr Lys Ile Pro Phe Cys Asp Tyr					
	450		455		460
Gly Glu Ala Leu Asp Pro Pro Ser Val Asp Val Thr Ala His Ile Ile					
465		470		475	480
Glu Ala Phe Gly Lys Leu Gly Ile Ser Arg Asn His Pro Ser Met Val					
	485		490		495
Gln Ala Leu Asp Tyr Ile Arg Arg Glu Gln Glu Pro Ser Gly Pro Trp					
	500		505		510
Phe Gly Arg Trp Gly Val Asn Tyr Val Tyr Gly Thr Gly Ala Val Leu					
	515		520		525
Pro Ala Leu Ala Ala Ile Gly Glu Asp Met Thr Gln Pro Tyr Ile Gly					
	530		535		540
Arg Ala Cys Asp Trp Leu Val Ala His Gln Gln Ala Asp Gly Gly Trp					
545		550		555	560
Gly Glu Ser Cys Ala Ser Tyr Met Asp Val Ser Ala Val Gly Arg Gly					
	565		570		575
Thr Thr Thr Ala Ser Gln Thr Ala Trp Ala Leu Met Ala Leu Leu Ala					
	580		585		590
Ala Asn Arg Pro Gln Asp Lys Asp Ala Ile Glu Arg Gly Cys Met Trp					
	595		600		605
Leu Val Glu Arg Gln Ser Ala Gly Thr Trp Asp Glu Pro Glu Phe Thr					
	610		615		620
Gly Thr Gly Phe Pro Gly Tyr Gly Val Gly Gln Thr Ile Lys Leu Asn					
625		630		635	640

Asp Pro Ala Leu Ser Gln Arg Leu Met Gln Gly Pro Glu Leu Ser Arg
 645 650 655
 Ala Phe Met Leu Arg Tyr Gly Met Tyr Arg His Tyr Phe Pro Leu Met
 660 665 670
 Ala Leu Gly Arg Ala Leu Arg Pro Gln Ser His Ser
 675 680

<210> 140

<211> 2055

<212> DNA

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<220>

<221> misc_feature

<222> (421) .. (423)

<223> ntn是ttr或ctn

<400> 140

```

atgactgtga ccagctcggc ctccgcgcgt gcgacgcgcg acccgggaaa ttatcagact      60
gccctgcaat cgacggtgcg cgcggcggcg gattggctga tcgccaacca gaagccggac      120
ggccattggg tcggccgcgc cgagtccaat gcctgcatgg aggcgcaatg gtgcctcgcg      180
ctgtggttca tggggctcga ggaccatccg ctgcgcaagc gcctgggcca gtcgctgctc      240
gacagccagc gcccggacgg cgcttggcag gtctatttcg gcgcccccaa tggcgacatc      300
aacgcgactg tcgaggccta tgccgcgctc cgctcgcgtg gcttccgcga cgacgagccg      360
gcggtgcgcc gggcgcgggg atggatcgag gccaagggcg gcctgcgcaa catccgcgtc      420
ntnaccgcgt actggctggc actgatcggc gaatggccgt gggagaagac accgaacatc      480
ccgccggagg tgatctggtt tccgctctgg tttccgttct cgatctacaa tttcgcgcaa      540
tgggcccgcg ccaccttgat gccgatcgcc gtgctgtcgg cgcggcggcc gagccggccc      600
ctgccgccgg agaaccgcct cgatgcgctg tttccgcatg gacggaaggc gttcgactac      660
gaactgccgg tcaaggccgg cgccggcggc tgggacaggt tcttccgcgg cgccgacaag      720
gttctgcaca agctgcagaa cctcggcaac cgtctcaatc tcggcctgtt ccgcccggcg      780
gccaccagcc gcgtgctgga atggatgata cgccatcagg atttcgacgg cgcctggggc      840
ggcatccagc cgccctggat ctacgggctg atggcgctct atgccgaagg ctatccgctc      900
aatcatcccc tgctcgcaaa gggcctcgac gcgctgaacg atccccgctg gcgcgtcgat      960
gtcggtgacg ccacctacat ccaggccacc aacagcccgg tctgggacac gatcctgacc     1020
ttgctcgctt tcgacgatgc cggcgtgctc ggcgactatc ccgaggccgt cgacaaggcg     1080
gtcgactggg tgctgcagcg gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg     1140
catgtcaagc ccggcggttg ggcgttcgaa tacgccaaca actactatcc cgacacggac     1200
gacaccgcgg tcgcgctgat cgcgctggcg ccaactgcgc acgatccgaa atggaaggcc     1260
aaagggatcg acgaggctat ccagctcggg gtcgactggc tgatcggcat gcagagccag     1320

```

```

ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagatcccc 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcggttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggg ccgtgggttcg gccgctgggg cgtcaattac 1560
gtctacggca ccggcgcggg gctgccggcg ctggccgcga tcggcgagga catgaccag 1620
ccctatatcg gccgcgcctg cgactggctg gttgcccatc agcaggccga tggcggctgg 1680
ggcgagagct ggcctccta catggatgtc agcgcggtcg gccgcggcac cacaacggcc 1740
tcgcagaccg cctgggcgct gatggcgctg ctgcgccca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttccgggc tacggcgctg gccagaccat caagctgaac 1920
gatcccgcg cgtcgagcg gctgatgcag ggcccgaat tgtccgcgc cttcatgctc 1980
cgctacggca tgtaccgcca ctacttccg ctgatggcg tcggccgcgc cctacgccc 2040
cagagtcata gctag 2055

```

<210> 141

<211> 684

<212> PRT

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 141

```

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 Leu Thr Arg Tyr
          130          135          140
Trp Leu Ala Leu Ile Gly Glu Trp Pro Trp Glu Lys Thr Pro Asn Ile

```

145	150	155	160
Pro Pro Glu Val Ile Trp Phe Pro Leu Trp Phe Pro Phe Ser Ile Tyr			
	165	170	175
Asn Phe Ala Gln Trp Ala Arg Ala Thr Leu Met Pro Ile Ala Val Leu			
	180	185	190
Ser Ala Arg Arg Pro Ser Arg Pro Leu Pro Pro Glu Asn Arg Leu Asp			
	195	200	205
Ala Leu Phe Pro His Gly Arg Lys Ala Phe Asp Tyr Glu Leu Pro Val			
	210	215	220
Lys Ala Gly Ala Gly Gly Trp Asp Arg Phe Phe Arg Gly Ala Asp Lys			
225	230	235	240
Val Leu His Lys Leu Gln Asn Leu Gly Asn Arg Leu Asn Leu Gly Leu			
	245	250	255
Phe Arg Pro Ala Ala Thr Ser Arg Val Leu Glu Trp Met Ile Arg His			
	260	265	270
Gln Asp Phe Asp Gly Ala Trp Gly Gly Ile Gln Pro Pro Trp Ile Tyr			
	275	280	285
Gly Leu Met Ala Leu Tyr Ala Glu Gly Tyr Pro Leu Asn His Pro Val			
	290	295	300
Leu Ala Lys Gly Leu Asp Ala Leu Asn Asp Pro Gly Trp Arg Val Asp			
305	310	315	320
Val Gly Asp Ala Thr Tyr Ile Gln Ala Thr Asn Ser Pro Val Trp Asp			
	325	330	335
Thr Ile Leu Thr Leu Leu Ala Phe Asp Asp Ala Gly Val Leu Gly Asp			
	340	345	350
Tyr Pro Glu Ala Val Asp Lys Ala Val Asp Trp Val Leu Gln Arg Gln			
	355	360	365
Val Arg Val Pro Gly Asp Trp Ser Met Lys Leu Pro His Val Lys Pro			
	370	375	380
Gly Gly Trp Ala Phe Glu Tyr Ala Asn Asn Tyr Tyr Pro Asp Thr Asp			
385	390	395	400
Asp Thr Ala Val Ala Leu Ile Ala Leu Ala Pro Leu Arg His Asp Pro			
	405	410	415
Lys Trp Lys Ala Lys Gly Ile Asp Glu Ala Ile Gln Leu Gly Val Asp			
	420	425	430
Trp Leu Ile Gly Met Gln Ser Gln Gly Gly Gly Trp Gly Ala Phe Asp			
	435	440	445
Lys Asp Asn Asn Gln Lys Ile Leu Thr Lys Ile Pro Phe Cys Asp Tyr			
450	455	460	

Gly	Glu	Ala	Leu	Asp	Pro	Pro	Ser	Val	Asp	Val	Thr	Ala	His	Ile	Ile
465					470					475					480
Glu	Ala	Phe	Gly	Lys	Leu	Gly	Ile	Ser	Arg	Asn	His	Pro	Ser	Met	Val
				485						490				495	
Gln	Ala	Leu	Asp	Tyr	Ile	Arg	Arg	Glu	Gln	Glu	Pro	Ser	Gly	Pro	Trp
			500					505					510		
Phe	Gly	Arg	Trp	Gly	Val	Asn	Tyr	Val	Tyr	Gly	Thr	Gly	Ala	Val	Leu
		515					520					525			
Pro	Ala	Leu	Ala	Ala	Ile	Gly	Glu	Asp	Met	Thr	Gln	Pro	Tyr	Ile	Gly
		530				535					540				
Arg	Ala	Cys	Asp	Trp	Leu	Val	Ala	His	Gln	Gln	Ala	Asp	Gly	Gly	Trp
545					550					555					560
Gly	Glu	Ser	Cys	Ala	Ser	Tyr	Met	Asp	Val	Ser	Ala	Val	Gly	Arg	Gly
				565					570					575	
Thr	Thr	Thr	Ala	Ser	Gln	Thr	Ala	Trp	Ala	Leu	Met	Ala	Leu	Leu	Ala
			580					585					590		
Ala	Asn	Arg	Pro	Gln	Asp	Lys	Asp	Ala	Ile	Glu	Arg	Gly	Cys	Met	Trp
		595					600					605			
Leu	Val	Glu	Arg	Gln	Ser	Ala	Gly	Thr	Trp	Asp	Glu	Pro	Glu	Phe	Thr
		610					615					620			
Gly	Thr	Gly	Tyr	Pro	Gly	Tyr	Gly	Val	Gly	Gln	Thr	Ile	Lys	Leu	Asn
625					630					635					640
Asp	Pro	Ala	Leu	Ser	Gln	Arg	Leu	Met	Gln	Gly	Pro	Glu	Leu	Ser	Arg
				645					650				655		
Ala	Phe	Met	Leu	Arg	Tyr	Gly	Met	Tyr	Arg	His	Tyr	Phe	Pro	Leu	Met
			660					665					670		
Ala	Leu	Gly	Arg	Ala	Leu	Arg	Pro	Gln	Ser	His	Ser				
		675					680								

<210> 142

<211> 2055

<212> DNA

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<220>

<221> misc_feature

<222> (421) .. (423)

<223> ntn是ttr或ctn

<400> 142

atgactgtga ccagctcggc ctccgcgcgt gcgacgcgcg acccgggaaa ttatcagact	60
gccctgcaat cgacggtgcg cgcggcggcg gattggctga tcgccaacca gaagccggac	120
ggccattggg tcggccgcgc cgagtccaat gcctgcatgg aggcgcaatg gtgcctcgcg	180
ctgtggttca tggggctcga ggaccatccg ctgcgcaagc gcctgggcca gtcgctgctc	240
gacagccagc gcccggacgg cgcctggcag gtctatttcg gcgcccccaa tggcgacatc	300
aacgcgactg tcgaggccta tgccgcgctc cgctcgctgg gcttccgcga cgacgagccg	360
gcggtgcgcc gggcgcggga atggatcgag gccaaaggcg gcctgcgcaa catccgcgtc	420
ntnaccgcgt actggctggc actgatcggc gaatggccgt gggagaagac accgaacatc	480
ccgccggagg tgatctggtt tccgctctgg tttccgttct cgatctacaa tttcgcgcaa	540
tgggcccgcg ccaccttgat gccgatcgcc gtgctgtcgg cgcggcggcc gagccggccg	600
ctgccgccgg agaaccgcct cgatgcgctg tttccgcatg gacggaaggc gttcgactac	660
gaactgccgg tcaaggccgg cgcggcggc tgggacaggt tcttccgcgg cgccgacaag	720
gttctgcaca agctgcagaa cctcggcaac cgtctcaatc tcggcctgtt ccgcccgcg	780
gccaccagcc gcgtgctgga atggatgata cgccatcagg atttcgacgg cgcctggggc	840
ggcatccagc cgccctggat ctacgggctg atggcgctct atgccgaagg ctatccgctc	900
aatcatcccg tgctcgcaaa gggcctcgac gcgctgaacg atcccggctg gcgcgtcgat	960
gtcggtgacg ccacctacat ccaggccacc aacagcccgg tctgggacac gatcctgacc	1020
ttgctcgctt tcgacgatgc cggcgtgctc ggcgactatc ccgaggccgt cgacaaggcg	1080
gtcgactggg tgctgcagcg gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg	1140
catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac	1200
gacaccgcgg tcgcgtgat cgcgctggcg ccaactgcgc acgatccgaa atggaaggcc	1260
aaagggatcg acgaggctat ccagctcggg gtcgactggc tgatcggcat gcagagccag	1320
ggcggcggct ggggcgcggt cgacaaggac aacaaccaga agatcctgac caagatcccg	1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggctg acgtcaccgc ccacatcatc	1440
gaggcggttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac	1500
tatatctgcc gtgagcagga gccgagcggg ccgtgggttcg gccgctgggg cgtcaattac	1560
gtctacggca ccggcgcggt gctgccggcg ctggccgcga tcggcgagga catgaccag	1620
ccctatatcg gccgcgcctg cgactggctg gttgcccatc agcaggccga tggcggctgg	1680
ggcgagagct gcgcctccta catggatgtc agcgcggctc gccgcggcac cacaacggcc	1740
tcgcagaccg cctgggcgct gatggcgctg ctgcgccca atcgccccca ggacaaggac	1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag	1860
ccggaattca ccggcaccgg ttayccgggc tacggcgctc gccagaccat caagctgaac	1920
gatcccgcgc tgtcgcagcg gctgatgcag ggcccgaat tgtcccgcgc cttcatgctc	1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg	2040
cagagtcata gctag	2055
<210> 143	
<211> 684	
<212> PRT	
<213> 人工序列	

<220>

<223> 合成BjapSHC1衍生物

<400> 143

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	Leu	Thr	Arg	Arg
		130					135						140		
Trp	Leu	Ala	Leu	Ile	Gly	Glu	Trp	Pro	Trp	Glu	Lys	Thr	Pro	Asn	Ile
145					150					155				160	
Pro	Pro	Glu	Val	Ile	Trp	Phe	Pro	Leu	Trp	Phe	Pro	Phe	Ser	Ile	Tyr
			165						170					175	
Asn	Phe	Ala	Gln	Trp	Ala	Arg	Ala	Thr	Leu	Met	Pro	Ile	Ala	Val	Leu
			180					185					190		
Ser	Ala	Arg	Arg	Pro	Ser	Arg	Pro	Leu	Pro	Pro	Glu	Asn	Arg	Leu	Asp
		195					200					205			
Ala	Leu	Phe	Pro	His	Gly	Arg	Lys	Ala	Phe	Asp	Tyr	Glu	Leu	Pro	Val
		210				215					220				
Lys	Ala	Gly	Ala	Gly	Gly	Trp	Asp	Arg	Phe	Phe	Arg	Gly	Ala	Asp	Lys
225					230					235				240	
Val	Leu	His	Lys	Leu	Gln	Asn	Leu	Gly	Asn	Arg	Leu	Asn	Leu	Gly	Leu
			245						250					255	
Phe	Arg	Pro	Ala	Ala	Thr	Ser	Arg	Val	Leu	Glu	Trp	Met	Ile	Arg	His
			260					265					270		
Gln	Asp	Phe	Asp	Gly	Ala	Trp	Gly	Gly	Ile	Gln	Pro	Pro	Trp	Ile	Tyr
		275						280					285		

Gly Leu Met Ala Leu Tyr Ala Glu Gly Tyr Pro Leu Asn His Pro Val		
290	295	300
Leu Ala Lys Gly Leu Asp Ala Leu Asn Asp Pro Gly Trp Arg Val Asp		
305	310	315 320
Val Gly Asp Ala Thr Tyr Ile Gln Ala Thr Asn Ser Pro Val Trp Asp		
	325	330 335
Thr Ile Leu Thr Leu Leu Ala Phe Asp Asp Ala Gly Val Leu Gly Asp		
	340	345 350
Tyr Pro Glu Ala Val Asp Lys Ala Val Asp Trp Val Leu Gln Arg Gln		
	355	360 365
Val Arg Val Pro Gly Asp Trp Ser Met Lys Leu Pro His Val Lys Pro		
	370	375 380
Gly Gly Trp Ala Phe Glu Tyr Ala Asn Asn Tyr Tyr Pro Asp Thr Asp		
385	390	395 400
Asp Thr Ala Val Ala Leu Ile Ala Leu Ala Pro Leu Arg His Asp Pro		
	405	410 415
Lys Trp Lys Ala Lys Gly Ile Asp Glu Ala Ile Gln Leu Gly Val Asp		
	420	425 430
Trp Leu Ile Gly Met Gln Ser Gln Gly Gly Gly Trp Gly Ala Phe Asp		
	435	440 445
Lys Asp Asn Asn Gln Lys Ile Leu Thr Lys Thr Pro Phe Cys Asp Tyr		
	450	455 460
Gly Glu Ala Leu Asp Pro Pro Ser Val Asp Val Thr Ala His Ile Ile		
465	470	475 480
Glu Ala Phe Gly Lys Leu Gly Ile Ser Arg Asn His Pro Ser Met Val		
	485	490 495
Gln Ala Leu Asp Tyr Ile Arg Arg Glu Gln Glu Pro Ser Gly Pro Trp		
	500	505 510
Phe Gly Arg Trp Gly Val Asn Tyr Val Tyr Gly Thr Gly Ala Val Leu		
	515	520 525
Pro Ala Leu Ala Ala Ile Gly Glu Asp Met Thr Gln Pro Tyr Ile Gly		
	530	535 540
Arg Ala Cys Asp Trp Leu Val Ala His Gln Gln Ala Asp Gly Gly Trp		
545	550	555 560
Gly Glu Ser Cys Ala Ser Tyr Met Asp Val Ser Ala Val Gly Arg Gly		
	565	570 575
Thr Thr Thr Ala Ser Gln Thr Ala Trp Ala Leu Met Ala Leu Leu Ala		
	580	585 590
Ala Asn Arg Pro Gln Asp Lys Asp Ala Ile Glu Arg Gly Cys Met Trp		

595	600	605
Leu Val Glu Arg Gln Ser Ala Gly Thr Trp Asp Glu Pro Glu Phe Thr		
610	615	620
Gly Thr Gly Phe Pro Gly Tyr Gly Val Gly Gln Thr Ile Lys Leu Asn		
625	630	635
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> 144

<211> 2055

<212> DNA

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<220>

<221> misc_feature

<222> (421) .. (423)

<223> ntn是ttr或ctn

<220>

<221> misc_feature

<222> (430) .. (432)

<223> ngn是agr或cgn

<220>

<221> misc_feature

<222> (1377) .. (1377)

<223> n是a, c, g,或t

<400> 144

atgactgtga ccagctcggc ctccgcgcgt gcgacgcgcg acccgggaaa ttatcagact	60
gccctgcaat cgacggtgcg cgcggcggcg gatttggtga tcgccaacca gaagccggac	120
ggccattggg tcggccgcgc cgagtccaat gcctgcatgg aggcgcaatg gtgcctcgcg	180
ctgtggttca tggggctcga ggaccatccg ctgcgcaagc gcctgggcca gtcgctgctc	240
gacagccagc gcccggacgg cgcttggcag gtctatttcg gcgcccccaa tggcgacatc	300
aacgcgactg tcgaggccta tgccgcgctc cgctcgctgg gcttccgcga cgacgagccg	360
gcggtgcgcc gggcgcggga atggatcgag gccaaaggcg gcctgcgcaa catccgcgtc	420
ntnaccgcgn gntggctggc actgatcggc gaatggccgt gggagaagac accgaacatc	480
ccgccggagg tgatctggtt tccgctctgg tttccgttct cgatctacaa tttcgcgcaa	540


```

tgggccccgcg ccaccttgat gccgatcgcc gtgctgtcgg cgcgggcgcc gagccggccg      600
ctgccgccgg agaaccgcct cgatgcgctg tttccgcatg gacggaaggc gttcgactac      660
gaactgccgg tcaaggccgg cgccggcggc tgggacaggt tcttccgcgg cgccgacaag      720
gttctgcaca agctgcagaa cctcggcaac cgtctcaatc tcggcctgtt ccgcccggcg      780
gccaccagcc gcgtgctgga atggatgata cgccatcagg atttcgacgg cgcctggggc      840
ggcatccagc cgccctggat ctacgggctg atggcgctct atgccgaagg ctatccgctc      900
aatcatccccg tgctcgcaaa gggcctcgac gcgtgaacg atcccggctg gcgcgtcgat      960
gtcggtgacg ccacctacat ccaggccacc aacagcccgg tctgggacac gatcctgacc     1020
ttgctcgctt tcgacgatgc cggcgtgctc ggcgactatc ccgaggccgt cgacaaggcg     1080
gtcgactggg tgctgcagcg gcaggtgctc gtgcccggcg actggtcgat gaagctgccg     1140
catgtcaagc ccggcggtg ggcgttcgaa tacccaaca actactatcc cgacacggac     1200
gacaccgcgg tcgcgtgat cgcgtggcg ccaactgcgc acgatccgaa atggaaggcc     1260
aaagggatcg acgaggctat ccagctcggt gtcgactggc tgatcgcat gcagagccag     1320
ggcgggcggt ggggcgctt cgacaaggac aacaaccaga agatcctgac caagacnccg     1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtaccgc ccacatcatc     1440
gaggcggttcg gcaagctcgg catctcgcg aaccatccgt cgatggtgca ggcgtggac     1500
tatattcgcc gtgagcagga gccgagcggc ccgtggttcg gccgctgggg cgtcaattac     1560
gtctacggca ccggcgcggt gctgccggcg ctggccgcga tcggcgagga catgaccag     1620
ccctatatcg gccgcgctg cgactggctg gttgcccatc agcaggccga tggcggttg     1680
ggcgagagct gcgcctccta catggatgtc agcgcggtcg gccgcggcac cacaacggcc     1740
tcgcagaccg cctgggcgct gatggcgctg ctcgccgcca atcgcccca ggacaaggac     1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag     1860
ccggaattca ccggcaccgg tttccggggc tacggcgctc gccagaccat caagctgaac     1920
gatcccgcg cgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc cttcatgctc     1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcg cggccgcgc cctacgcccg     2040
cagagtcata gctag                                         2055

```

<210> 145

<211> 684

<212> PRT

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 145

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 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 Arg		
130	135	140
Trp Leu Ala Leu Ile Gly Glu Trp Pro Trp Glu Lys Thr Pro Asn Ile		
145	150	155 160
Pro Pro Glu Val Ile Trp Phe Pro Leu Trp Phe Pro Phe Ser Ile Tyr		
165	170	175
Asn Phe Ala Gln Trp Ala Arg Ala Thr Leu Met Pro Ile Ala Val Leu		
180	185	190
Ser Ala Arg Arg Pro Ser Arg Pro Leu Pro Pro Glu Asn Arg Leu Asp		
195	200	205
Ala Leu Phe Pro His Gly Arg Lys Ala Phe Asp Tyr Glu Leu Pro Val		
210	215	220
Lys Ala Gly Ala Gly Gly Trp Asp Arg Phe Phe Arg Gly Ala Asp Lys		
225	230	235 240
Val Leu His Lys Leu Gln Asn Leu Gly Asn Arg Leu Asn Leu Gly Leu		
245	250	255
Phe Arg Pro Ala Ala Thr Ser Arg Val Leu Glu Trp Met Ile Arg His		
260	265	270
Gln Asp Phe Asp Gly Ala Trp Gly Gly Ile Gln Pro Pro Trp Ile Tyr		
275	280	285
Gly Leu Met Ala Leu Tyr Ala Glu Gly Tyr Pro Leu Asn His Pro Val		
290	295	300
Leu Ala Lys Gly Leu Asp Ala Leu Asn Asp Pro Gly Trp Arg Val Asp		
305	310	315 320
Val Gly Asp Ala Thr Tyr Ile Gln Ala Thr Asn Ser Pro Val Trp Asp		
325	330	335
Thr Ile Leu Thr Leu Leu Ala Phe Asp Asp Ala Gly Val Leu Gly Asp		
340	345	350
Tyr Pro Glu Ala Val Asp Lys Ala Val Asp Trp Val Leu Gln Arg Gln		

355	360	365
Val Arg Val Pro Gly Asp Trp	Ser Met Lys Leu Pro	His Val Lys Pro
370	375	380
Gly Gly Trp Ala Phe Glu Tyr Ala	Asn Asn Tyr Tyr Pro Asp Thr Asp	
385	390	395
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 Thr 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 Tyr Pro Gly Tyr Gly	Val Gly Gln Thr Ile Lys Leu Asn	
625	630	635
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> 146

<211> 2055

<212> DNA

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<220>

<221> misc_feature

<222> (430) .. (432)

<223> ngn是agr或cgn

<220>

<221> misc_feature

<222> (1377) .. (1377)

<223> n是a, c, g,或t

<400> 146

atgactgtga ccagctcggc ctccgcgcgt gcgacgcgcg acccgggaaa ttatcagact	60
gccctgcaat cgacggtgcg cgcgggcgcg gattggctga tcgccaacca gaagccggac	120
ggccattggg tcggccgcgc cgagtccaat gcctgcatgg aggcgcaatg gtgcctcgcg	180
ctgtggttca tggggctcga ggaccatccg ctgcgcaagc gcctgggcca gtcgctgctc	240
gacagccagc gcccggacgg cgcctggcag gtctatttcg gcgcccccaa tggcgacatc	300
aacgcgactg tcgaggccta tgccgcgctc cgctcgcgtg gcttccgcga cgacgagccg	360
gcggtgcgcc gggcgcgggg atggatcgag gccaaaggcg gcctgcgcaa catccgcgtc	420
ttcaccgcgn gntggctggc actgatcggc gaatggccgt gggagaagac accgaacatc	480
ccgccggagg tgatctggtt tccgctctgg tttccgttct cgatctacaa tttcgcgcaa	540
tgggcccgcg ccaccttgat gccgatcgcc gtgctgtcgg cgcgggcgcc gagccggccg	600
ctgccgccgg agaaccgcct cgatgcgctg tttccgcatg gacggaaggc gttcgactac	660
gaactgccgg tcaaggccgg cgccggcggc tgggacaggt tcttccgcgg cgccgacaag	720
gttctgcaca agctgcagaa cctcggcaac cgtctcaatc tcggcctgtt ccgccggcg	780
gccaccagcc gcgtgctgga atggatgata cgccatcagg atttcgacgg cgcctggggc	840
ggcatccagc cgccctggat ctacgggctg atggcgctct atgccgaagg ctatccgctc	900
aatcatcccc tgctcgcaaa gggcctcgac gcgctgaacg atccccgctg gcgcgtcgat	960
gtcggtgacg ccacctacat ccaggccacc aacagcccgg tctgggacac gatcctgacc	1020
ttgctcgctt tcgacgatgc cggcgtgctc ggcgactatc ccgaggccgt cgacaaggcg	1080
gtcgactggg tgctgcagcg gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg	1140
catgtcaagc ccggcggtg ggcgttcgaa tacgccaaca actactatcc cgacacggac	1200
gacaccgcgg tcgcgctgat cgcgctggcg ccaactgcgc acgatccgaa atggaaggcc	1260
aaagggatcg acgaggctat ccagctcggg gtcgactggc tgatcggcat gcagagccag	1320

```

ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagacnccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcggttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggg ccgtgggttcg gccgctgggg cgtcaattac 1560
gtctacggca ccggcgcggg gctgccggcg ctggccgcga tcggcgagga catgaccag 1620
ccctatatcg gccgcgcctg cgactggctg gttgcccatc agcaggccga tggcggctgg 1680
ggcgagagct gcgcctccta catggatgtc agcgcggtcg gccgcggcac cacaacggcc 1740
tcgcagaccg cctgggcgct gatggcgctg ctgcgccca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg ttayccgggc tacggcgctg gccagaccat caagctgaac 1920
gatccccgcg tgtcgcagcg gctgatgcag ggcccgaat tgtcccgcgc cttcatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

```

<210> 147

<211> 684

<212> PRT

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 147

```

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 Leu Thr Arg Arg
          130          135          140
Trp Leu Ala Leu Ile Gly Glu Trp Pro Trp Glu Lys Thr Pro Asn Ile

```

145	150	155	160
Pro Pro Glu Val Ile Trp Phe Pro Leu Trp Phe Pro Phe Ser Ile Tyr			
	165	170	175
Asn Phe Ala Gln Trp Ala Arg Ala Thr Leu Met Pro Ile Ala Val Leu			
	180	185	190
Ser Ala Arg Arg Pro Ser Arg Pro Leu Pro Pro Glu Asn Arg Leu Asp			
	195	200	205
Ala Leu Phe Pro His Gly Arg Lys Ala Phe Asp Tyr Glu Leu Pro Val			
	210	215	220
Lys Ala Gly Ala Gly Gly Trp Asp Arg Phe Phe Arg Gly Ala Asp Lys			
225	230	235	240
Val Leu His Lys Leu Gln Asn Leu Gly Asn Arg Leu Asn Leu Gly Leu			
	245	250	255
Phe Arg Pro Ala Ala Thr Ser Arg Val Leu Glu Trp Met Ile Arg His			
	260	265	270
Gln Asp Phe Asp Gly Ala Trp Gly Gly Ile Gln Pro Pro Trp Ile Tyr			
	275	280	285
Gly Leu Met Ala Leu Tyr Ala Glu Gly Tyr Pro Leu Asn His Pro Val			
	290	295	300
Leu Ala Lys Gly Leu Asp Ala Leu Asn Asp Pro Gly Trp Arg Val Asp			
305	310	315	320
Val Gly Asp Ala Thr Tyr Ile Gln Ala Thr Asn Ser Pro Val Trp Asp			
	325	330	335
Thr Ile Leu Thr Leu Leu Ala Phe Asp Asp Ala Gly Val Leu Gly Asp			
	340	345	350
Tyr Pro Glu Ala Val Asp Lys Ala Val Asp Trp Val Leu Gln Arg Gln			
	355	360	365
Val Arg Val Pro Gly Asp Trp Ser Met Lys Leu Pro His Val Lys Pro			
	370	375	380
Gly Gly Trp Ala Phe Glu Tyr Ala Asn Asn Tyr Tyr Pro Asp Thr Asp			
385	390	395	400
Asp Thr Ala Val Ala Leu Ile Ala Leu Ala Pro Leu Arg His Asp Pro			
	405	410	415
Lys Trp Lys Ala Lys Gly Ile Asp Glu Ala Ile Gln Leu Gly Val Asp			
	420	425	430
Trp Leu Ile Gly Met Gln Ser Gln Gly Gly Gly Trp Gly Ala Phe Asp			
	435	440	445
Lys Asp Asn Asn Gln Lys Ile Leu Thr Lys Thr Pro Phe Cys Asp Tyr			
450	455	460	

Gly	Glu	Ala	Leu	Asp	Pro	Pro	Ser	Val	Asp	Val	Thr	Ala	His	Ile	Ile
465					470					475					480
Glu	Ala	Phe	Gly	Lys	Leu	Gly	Ile	Ser	Arg	Asn	His	Pro	Ser	Met	Val
				485						490				495	
Gln	Ala	Leu	Asp	Tyr	Ile	Arg	Arg	Glu	Gln	Glu	Pro	Ser	Gly	Pro	Trp
			500					505					510		
Phe	Gly	Arg	Trp	Gly	Val	Asn	Tyr	Val	Tyr	Gly	Thr	Gly	Ala	Val	Leu
		515					520					525			
Pro	Ala	Leu	Ala	Ala	Ile	Gly	Glu	Asp	Met	Thr	Gln	Pro	Tyr	Ile	Gly
		530				535					540				
Arg	Ala	Cys	Asp	Trp	Leu	Val	Ala	His	Gln	Gln	Ala	Asp	Gly	Gly	Trp
545					550					555					560
Gly	Glu	Ser	Cys	Ala	Ser	Tyr	Met	Asp	Val	Ser	Ala	Val	Gly	Arg	Gly
				565					570					575	
Thr	Thr	Thr	Ala	Ser	Gln	Thr	Ala	Trp	Ala	Leu	Met	Ala	Leu	Leu	Ala
			580					585					590		
Ala	Asn	Arg	Pro	Gln	Asp	Lys	Asp	Ala	Ile	Glu	Arg	Gly	Cys	Met	Trp
		595					600					605			
Leu	Val	Glu	Arg	Gln	Ser	Ala	Gly	Thr	Trp	Asp	Glu	Pro	Glu	Phe	Thr
		610					615					620			
Gly	Thr	Gly	Tyr	Pro	Gly	Tyr	Gly	Val	Gly	Gln	Thr	Ile	Lys	Leu	Asn
625					630					635					640
Asp	Pro	Ala	Leu	Ser	Gln	Arg	Leu	Met	Gln	Gly	Pro	Glu	Leu	Ser	Arg
				645					650				655		
Ala	Phe	Met	Leu	Arg	Tyr	Gly	Met	Tyr	Arg	His	Tyr	Phe	Pro	Leu	Met
			660					665					670		
Ala	Leu	Gly	Arg	Ala	Leu	Arg	Pro	Gln	Ser	His	Ser				
		675						680							

<210> 148

<211> 2055

<212> DNA

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<220>

<221> misc_feature

<222> (421) .. (423)

<223> ntn是ttr或ctn

<220>

<221> misc_feature

<222> (430) .. (432)

<223> ngn是agr或cgn

<220>

<221> misc_feature

<222> (1377) .. (1377)

<223> n是a, c, g,或t

<400> 148

atgactgtga	ccagctcggc	ctccgcgcgt	gcgacgcgcg	acccgggaaa	ttatcagact	60
gccctgcaat	cgacggtgcg	cgcggcggcg	gattggctga	tcgccaacca	gaagccggac	120
ggccattggg	tcggccgcgc	cgagtccaat	gcctgcatgg	aggcgcaatg	gtgcctcgcg	180
ctgtggttca	tggggctcga	ggaccatccg	ctgcgcaagc	gcctgggcca	gtcgtgtctc	240
gacagccagc	gccccggacgg	cgcttggcag	gtctattttcg	gcgcccccaa	tggcgacatc	300
aacgcgactg	tcgaggccta	tgccgcgctc	cgctcgtctg	gcttccgcga	cgacgagccg	360
gcggtgcgcc	gggcgcggga	atggatcgag	gccaaaggcg	gcctgcgcaa	catccgcgtc	420
ntnaccgcgn	gntggctggc	actgatcggc	gaatggccgt	gggagaagac	accgaacatc	480
ccgccggagg	tgatctggtt	tccgctctgg	tttccgttct	cgatctacaa	tttcgcgcaa	540
tgggcccgcg	ccaccttgat	gccgatcgcc	gtgctgtcgg	cgcggcggcc	gagccggccg	600
ctgccgccgg	agaaccgcct	cgatgcgctg	tttccgcatg	gacggaaggc	gttcgactac	660
gaactgccgg	tcaaggccgg	cgccggcggc	tgggacaggt	tcttccgcgg	cgccgacaag	720
gttctgcaca	agctgcagaa	cctcggcaac	cgtctcaatc	tcggcctgtt	ccgcccggcg	780
gccaccagcc	gcgtgctgga	atggatgata	cgccatcagg	atttcgacgg	cgcttggggc	840
ggcatccagc	cgccctggat	ctacgggctg	atggcgctct	atgccgaagg	ctatccgctc	900
aatcatcccc	tgctcgcaaa	gggcctcgac	gcgctgaacg	atccccgctg	gcgcgtcgat	960
gtcggtgacg	ccacctacat	ccaggccacc	aacagcccgg	tctgggacac	gatacctgacc	1020
ttgctcgcct	tcgacgatgc	cggcgtgctc	ggcgactatc	ccgaggccgt	cgacaaggcg	1080
gtcgactggg	tgctgcagcg	gcaggtgcgc	gtgcccggcg	actggtcgat	gaagctgccg	1140
catgtcaagc	ccggcggctg	ggcgttcgaa	tacgccaaca	actactatcc	cgacacggac	1200
gacaccgcgg	tcgcgctgat	cgcgctggcg	ccactgcgcc	acgatccgaa	atggaaggcc	1260
aaagggatcg	acgaggctat	ccagctcggg	gtcgactggc	tgatcggcat	gcagagccag	1320
ggcggcggct	ggggcgcggt	cgacaaggac	aacaaccaga	agatcctgac	caagacnccg	1380
ttctgcgatt	atggcgaggc	gctcgatccg	ccctcggtcg	acgtcaccgc	ccacatcatc	1440
gaggcggttc	gcaagctcgg	catctcgcgc	aaccatccgt	cgatggtgca	ggcgctggac	1500
tatatctgcc	gtgagcagga	gccgagcggt	ccgtggttcg	gccgctgggg	cgtcaattac	1560
gtctacggca	ccggcgcggt	gctgccggcg	ctggccgcga	tcggcgagga	catgaccag	1620
ccctatatcg	gccgcgcctg	cgactggctg	gttgcctatc	agcaggccga	tggcggctgg	1680
ggcgagagct	gcgcctccta	catggatgtc	agcgcggtcg	gccgcggcac	cacaacggcc	1740
tcgcagaccg	cctgggcgct	gatggcgctg	ctcgccgccca	atcgccccca	ggacaaggac	1800
gcgatcgagc	gtggctgcat	gtggctggtc	gagcgccagt	cggccggcac	ctgggacgag	1860

ccggaattca ccggcaccgg ttayccgggc tacggcgctcg gccagaccat caagctgaac 1920
 gatcccgcg cgtcgagcg gctgatgcag ggcccgaat tgtcccgcg cttcatgctc 1980
 cgctacggca tgtaccgcca ctacttcccg ctgatggcg tggcccgcg cctacgcccg 2040
 cagagtcata gctag 2055

<210> 149

<211> 657

<212> PRT

<213> 双向伯克霍尔德氏菌(*Burkholderia ambifaria*)

<400> 149

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

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

545	550	555	560
Tyr Arg Gly Phe Glu Gln Ala Pro Ser Thr Ala Ser Gln Thr Ala Trp			
	565	570	575
Ala Leu Leu Gly Leu Met Ala Ala Gly Glu Val Asn Asn Pro Ala Val			
	580	585	590
Ala Arg Gly Val Glu Tyr Leu Ile Ala Glu Gln Lys Glu His Gly Leu			
	595	600	605
Trp Asp Glu Thr Arg Phe Thr Ala Thr Gly Phe Pro Arg Val Phe Tyr			
	610	615	620
Leu Arg Tyr His Gly Tyr Arg Lys Phe Phe Pro Leu Trp Ala Leu Ala			
625	630	635	640
Arg Tyr Arg Asn Leu Lys Arg Asn Asn Ala Thr Arg Val Thr Phe Gly			
	645	650	655
Leu			
<210> 150			
<400> 150			
000			
<210> 151			
<211> 682			
<212> PRT			
<213> 双向伯克霍尔德氏菌(Burkholderia ambifaria)			
<400> 151			
Met Ile Arg Arg Met Asn Lys Ser Gly Pro Ser Pro Trp Ser Ala Leu			
1	5	10	15
Asp Ala Ala Ile Ala Arg Gly Arg Asp Ala Leu Met Arg Leu Gln Gln			
	20	25	30
Pro Asp Gly Ser Trp Cys Phe Glu Leu Glu Ser Asp Ala Thr Ile Thr			
	35	40	45
Ala Glu Tyr Ile Leu Met Met His Phe Met Asp Lys Ile Asp Asp Ala			
	50	55	60
Arg Gln Glu Lys Met Ala Arg Tyr Leu Arg Ala Ile Gln Arg Leu Asp			
65	70	75	80
Thr His Gly Gly Trp Asp Leu Tyr Val Asp Gly Asp Pro Asp Val Ser			
	85	90	95
Cys Ser Val Lys Ala Tyr Phe Ala Leu Lys Ala Ala Gly Asp Ser Glu			
	100	105	110
His Ala Pro His Met Val Arg Ala Arg Asp Ala Ile Leu Glu Leu Gly			
	115	120	125
Gly Ala Ala Arg Ser Asn Val Phe Thr Arg Ile Leu Leu Ala Thr Phe			

130	135	140
Gly Gln Val Pro Trp Arg Ala Thr Pro Phe Met Pro Ile Glu Phe Val		
145	150	155
Leu Phe Pro Lys Trp Val Pro Ile Ser Met Tyr Lys Val Ala Tyr Trp		
165	170	175
Ala Arg Thr Thr Met Val Pro Leu Leu Val Leu Cys Ser Leu Lys Ala		
180	185	190
Arg Ala Arg Asn Pro Arg Asn Ile Ala Ile Pro Glu Leu Phe Val Thr		
195	200	205
Pro Pro Asp Gln Glu Arg Gln Tyr Phe Pro Pro Ala Arg Gly Met Arg		
210	215	220
Arg Ala Phe Leu Ala Leu Asp Arg Val Val Arg His Val Glu Pro Leu		
225	230	235
Leu Pro Lys Arg Leu Arg Gln Arg Ala Ile Arg His Ala Gln Ala Trp		
245	250	255
Cys Ala Glu Arg Met Asn Gly Glu Asp Gly Leu Gly Gly Ile Phe Pro		
260	265	270
Pro Ile Val Tyr Ser Tyr Gln Met Met Asp Val Leu Gly Tyr Pro Asp		
275	280	285
Asp His Pro Leu Arg Arg Asp Cys Glu Asn Ala Leu Glu Lys Leu Leu		
290	295	300
Val Thr Arg Pro Asp Gly Ser Met Tyr Cys Gln Pro Cys Leu Ser Pro		
305	310	315
Val Trp Asp Thr Ala Trp Ser Thr Met Ala Leu Glu Gln Ala Arg Gly		
325	330	335
Val Ala Val Pro Glu Ala Gly Ala Pro Ala Ser Ala Leu Asp Glu Leu		
340	345	350
Asp Ala Arg Ile Ala Arg Ala Tyr Asp Trp Leu Ala Glu Arg Gln Val		
355	360	365
Asn Asp Leu Arg Gly Asp Trp Ile Glu Asn Ala Pro Ala Asp Thr Gln		
370	375	380
Pro Gly Gly Trp Ala Phe Gln Tyr Ala Asn Pro Tyr Tyr Pro Asp Ile		
385	390	395
Asp Asp Ser Ala Val Val Thr Ala Met Leu Asp Arg Arg Gly Arg Thr		
405	410	415
His Arg Asn Ala Asp Gly Ser His Pro Tyr Ala Ala Arg Val Ala Arg		
420	425	430
Ala Leu Asp Trp Met Arg Gly Leu Gln Ser Arg Asn Gly Gly Phe Ala		
435	440	445

Ala Phe Asp Ala Asp Cys Asp Arg Leu Tyr Leu Asn Ala Ile Pro Phe
 450 455 460
 Ala Asp His Gly Ala Leu Leu Asp Pro Pro Thr Glu Asp Val Ser Gly
 465 470 475 480
 Arg Val Leu Leu Cys Phe Gly Val Thr Lys Arg Ala Asp Asp Arg Ala
 485 490 495
 Ser Leu Ala Arg Ala Ile Asp Tyr Val Lys Arg Thr Gln Gln Pro Asp
 500 505 510
 Gly Ser Trp Trp Gly Arg Trp Gly Thr Asn Tyr Leu Tyr Gly Thr Trp
 515 520 525
 Ser Val Leu Ala Gly Leu Ala Leu Ala Gly Glu Asp Pro Ser Gln Pro
 530 535 540
 Tyr Ile Ala Arg Ala Leu Ala Trp Leu Arg Ala Arg Gln His Ala Asp
 545 550 555 560
 Gly Gly Trp Gly Glu Thr Asn Asp Ser Tyr Ile Asp Pro Ala Leu Ala
 565 570 575
 Gly Thr Asn Ala Gly Glu Ser Thr Ser Asn Cys Thr Ala Trp Ala Leu
 580 585 590
 Leu Ala Gln Met Ala Phe Gly Asp Gly Glu Ser Glu Ser Val Arg Arg
 595 600 605
 Gly Ile Ala Tyr Leu Gln Ser Val Gln Gln Asp Asp Gly Phe Trp Trp
 610 615 620
 His Arg Ser His Asn Ala Pro Gly Phe Pro Arg Ile Phe Tyr Leu Lys
 625 630 635 640
 Tyr His Gly Tyr Thr Ala Tyr Phe Pro Leu Trp Ala Leu Ala Arg Tyr
 645 650 655
 Arg Arg Leu Ala Gly Gly Val Ser Ala Ala Gly Ala His Ala Val Pro
 660 665 670
 Ala Ser Thr Gly Ala Asp Ala Ala Leu Ala
 675 680
 <210> 152
 <400> 152
 000
 <210> 153
 <211> 617
 <212> PRT
 <213> 炭疽芽孢杆菌(Bacillus anthracis)
 <400> 153
 Met Leu Leu Tyr Glu Lys Ala His Glu Glu Ile Val Arg Arg Ala Thr

1	5	10	15
Ala Leu Gln Thr Met Gln Trp Gln Asp Gly Thr Trp Arg Phe Cys Phe			
20	25	30	
Glu Gly Ala Pro Leu Thr Asp Cys His Met Ile Phe Leu Leu Lys Leu			
35	40	45	
Leu Gly Arg Asp Lys Glu Ile Glu Pro Phe Val Glu Arg Val Ala Ser			
50	55	60	
Leu Gln Thr Asn Glu Gly Thr Trp Lys Leu His Glu Asp Glu Val Gly			
65	70	75	80
Gly Asn Leu Ser Ala Thr Ile Gln Ser Tyr Ala Ala Leu Leu Ala Ser			
85	90	95	
Lys Lys Tyr Thr Lys Glu Asp Ala Asn Met Lys Arg Ala Glu Asn Phe			
100	105	110	
Ile Gln Glu Arg Gly Gly Val Ala Arg Ala His Phe Met Thr Lys Phe			
115	120	125	
Leu Leu Ala Ile His Gly Glu Tyr Glu Tyr Pro Ser Leu Phe His Leu			
130	135	140	
Pro Thr Pro Ile Met Phe Leu Gln Asn Asp Ser Pro Phe Ser Ile Phe			
145	150	155	160
Glu Leu Ser Ser Ser Ala Arg Ile His Leu Ile Pro Met Met Leu Cys			
165	170	175	
Leu Asn Lys Arg Phe Arg Val Gly Lys Lys Leu Leu Pro Asn Leu Asn			
180	185	190	
His Ile Ala Gly Gly Gly Gly Glu Trp Phe Arg Glu Asp Arg Ser Pro			
195	200	205	
Val Phe Gln Thr Leu Leu Ser Asp Val Lys Gln Ile Ile Ser Tyr Pro			
210	215	220	
Leu Ser Leu His His Lys Gly Tyr Glu Glu Ile Glu Arg Phe Met Lys			
225	230	235	240
Glu Arg Ile Asp Glu Asn Gly Thr Leu Tyr Ser Tyr Ala Thr Ala Ser			
245	250	255	
Phe Tyr Met Ile Tyr Ala Leu Leu Ala Leu Gly His Ser Leu Gln Ser			
260	265	270	
Ser Met Ile Gln Lys Ala Ile Ala Gly Ile Thr Ser Tyr Ile Trp Lys			
275	280	285	
Met Glu Arg Gly Asn His Leu Gln Asn Ser Pro Ser Thr Val Trp Asp			
290	295	300	
Thr Ala Leu Leu Ser Tyr Ala Leu Gln Glu Ala Gln Val Ser Lys Asp			
305	310	315	320

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

<210> 154

<211> 1854

<212> DNA

<213> 炭疽芽孢杆菌(*Bacillus anthracis*)

<400> 154

```

atgttattat acgaaaaagc gcatgaagaa atagtgagaa gagcaacagc acttcaaaca      60
atgcaatggc aagatggtac gtggcgattt tgttttgaag gagctccatt aacagattgc      120
catatgattt ttttattaaa attattaggt agagataaag agatagaacc gttcgtagaa      180
agagtagcat cactccaaac aaatgaagga acatggaaat tgcacgaaga tgaagtagga      240
ggtaatttat cagctacaat tcaatcttat gccgccttac ttgcatcgaa aaaatataca      300
aaagaagatg cgaatatgaa acgagcagaa aattttattc aggaacgcgg tgggtgtggcg      360
cgtgctcatt ttatgacgaa gtttttatta gcaattcatg gagaatatga atatccttca      420
ctcttttcatt taccaacacc aatcatgttt ttacagaatg attccccctt tagtatattt      480
gaattaagta gctcagcacg tattcattta attccgatga tgctatgttt aaataaaaga      540
tttcgagtag ggaaaaagtt attaccaaatt ttaaatacaca ttgcgggcgg aggcggagaa      600
tgggtttcggg aggatcggtc tccagttttt caaacgttat taagtgatgt aaaacaaatt      660
atatcgtatc cactttcgtt acatcataaa ggatatgagg aaatagaacg ttttatgaaa      720
gagcgtattg atgaaaatgg aacgttatat agttacgcaa ctgcctcgtt ttatatgatt      780
tatgctttac ttgcgttagg gcattctctt caatcatcaa tgattcaaaa ggctatagct      840
gggataacat cttatatatg gaagatggaa agagggaatc atttgcaaaa ctctccttca      900
accgtgtggg atacagcttt attaagctat gcgttacaag aggctcaagt ttcaaaggat      960
aataagatga ttcaaaatgc aacagcgtat ttattaaaaa aacagcatac aaaaaaagct     1020
gattggagcg tacatgctcc ggcgcttact cctggcggtt ggggtttttc ggatgtgaat     1080
acgacaattc cagatataga tgatacaaca gctgtgctaa gggcattggc acgaagtaga     1140
ggaaacaaaa atatagataa tgcttggaag aaagggggca attggattaa aggattacaa     1200
aataatgatg gtggctgggg agcatttgaa aaaggtgtga cgagcaaatt attagcaaaa     1260
ttaccaatcg aaaacgcaag tgatatgatt acagatcctt ctacgccaga tattacgggg     1320
agagtgttag agtttttcgg gacgtatgca caaaacgaat tgcctgagaa acagatacaa     1380
agggcaataa attggttaat gaatgtacaa gaggaaaatg gatcatggta tgggaaatgg     1440
gggatttggt atctatatgg tacgtgggct gttatgactg gtttacggtc actcggaatt     1500
ccgtctagca atccttcatt gacacgagca gttcatggc ttgaacatat acagcatgaa     1560
gatggtggtt ggggagaatc atgccacagt agtgtggaga aaaggttcgt tactttacca     1620
tttagtacac catcccaaac tgcattggcg ttagatgctc tcatttctta ctatgataca     1680
gaaacgccag ctattcgaaa aggtgtttca tatttgcttt cgaatcctta tgtgaatgaa     1740
agatatccta ctggaacagg tttaccaggt gcgttttata ttaggtatca tagctatgcc     1800
catatatatc cactacttac tttggcacat tatataaaaa aatatagaaa ataa           1854

```

<210> 155

<211> 720

<212> PRT

<213> 桤木弗兰克氏菌(*Frankia alni*)

<400> 155

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

Gln Ile Ser Ala Val Leu Ala Thr Leu Ala Ala Val Ala Val Thr Ala			
305	310	315	320
Pro Ala Tyr Arg Lys Ile Gln Glu Gly Val Asp Arg Arg Phe Ser Arg			
	325	330	335
Arg Arg Phe Glu Thr Leu Gln Val Ile Arg Arg Tyr Leu Arg Asp Pro			
	340	345	350
Asp Pro Asp Val Ala Val Glu Glu Val Leu Arg Arg Ala Leu Gly Asp			
	355	360	365
Pro Thr Leu Ala Val Ala Tyr Leu Val Asp Asp Arg Arg Gln Trp Val			
	370	375	380
Ser Ala Asp Gly Gln Pro Ala Asn Pro Gly Asn Ser Phe Met Ala Ala			
385	390	395	400
Val Glu Val Tyr Arg Arg Gly Arg Pro Ile Ala Arg Val Thr Phe Asp			
	405	410	415
Arg Gly Arg Ala Gln Pro Gly Leu Val Arg Ala Ala Ala Thr Ala Ala			
	420	425	430
Thr Ala Glu Leu Asp Asn Ala Gly Leu Arg Ala Ala Val Ala Leu Gln			
	435	440	445
Leu Val Glu Val Arg Gln Ser Arg Thr Arg Ile Ala Ala Ala Gln Phe			
	450	455	460
Ala Glu Arg Arg Thr Ile Glu Arg Asn Leu His Asp Gly Ala Gln Gln			
465	470	475	480
Arg Leu Leu Ala Leu Ala Leu Gln Leu Arg Ala Val Gln Leu Gly Gly			
	485	490	495
Asp Glu Ala Ser Leu Arg Gln Ala Ile Ser Thr Gly Ile Asp Gln Leu			
	500	505	510
Gln Ala Ala Val Val Glu Leu Arg Glu Leu Ala Asn Gly Leu His Pro			
	515	520	525
Ala Val Leu Ala Asp Gly Gly Leu Ala Ala Ala Leu Asp Asp Val Ala			
	530	535	540
Ala Arg Thr Pro Val Pro Ile Lys Ile Ser Ala Pro Asp Arg Arg Tyr			
545	550	555	560
Pro Pro Asp Leu Glu Ala Ala Ala Trp Phe Ile Ala Cys Glu Ala Met			
	565	570	575
Ala Asn Ala Val Lys His Ala His Pro Thr Thr Ile Ala Val Asp Val			
	580	585	590
Ser Ala Pro Asp Gly Gln Leu Ile Val Glu Val Arg Asp Asp Gly Ile			
	595	600	605
Gly Gly Ala Gln Pro Ser Gly Pro Gly Leu Arg Gly Ile Ala Asp Arg			

610	615	620
Ala Glu Ala Phe Gly Gly Ser Leu Thr Val His Thr Asp Pro Gly Thr		
625	630	635
Gly Thr Thr Ile Arg Ala Leu Leu His Arg Arg Ser Pro Leu Ser Ser		640
	645	650
Gly Arg Arg Ser Val Met Ile Glu Gly Cys Val Asp Val Val Ala Val		655
	660	665
Arg Arg Phe Arg Cys Arg Ser Ser Arg Gly Ser Gly Ser Arg Arg Arg		670
	675	680
Arg Ser Ser Trp Arg Cys Gly Gly Ile Cys Gly Ser Arg Cys Arg Thr		685
690	695	700
Gly Met Ser Arg Ser Cys Ser Arg Asn Ala Ala Ser Lys Leu Ile Thr		
705	710	715
		720
<210> 156		
<400> 156		
000		
<210> 157		
<211> 685		
<212> PRT		
<213> Rhodopseudomonas palent		
<400> 157		
Met Asp Ser Ile Leu Ala Pro Arg Ala Asp Ala Pro Arg Asn Ile Asp		
1	5	10
Gly Ala Leu Arg Glu Ser Val Gln Gln Ala Ala Asp Trp Leu Val Ala		15
	20	25
Asn Gln Lys Pro Asp Gly His Trp Val Gly Arg Ala Glu Thr Asn Ala		30
	35	40
Thr Met Glu Ala Gln Trp Cys Leu Ala Leu Trp Phe Leu Gly Leu Glu		45
	50	55
Asp His Pro Leu Arg Val Arg Leu Gly Arg Ala Leu Leu Asp Thr Gln		60
65	70	75
Arg Pro Asp Gly Ala Trp His Val Phe Tyr Gly Ala Pro Asn Gly Asp		80
	85	90
Ile Asn Ala Thr Val Glu Ala Tyr Ala Ala Leu Arg Ser Leu Gly His		95
	100	105
Arg Asp Asp Glu Glu Pro Leu Arg Lys Ala Arg Asp Trp Ile Leu Ser		110
	115	120
Lys Gly Gly Leu Ala Asn Ile Arg Val Phe Thr Arg Tyr Trp Leu Ala		125
130	135	140

Leu	Ile	Gly	Glu	Trp	Pro	Trp	Glu	Lys	Thr	Pro	Asn	Ile	Leu	Pro	Glu
145					150					155					160
Val	Ile	Trp	Leu	Pro	Thr	Trp	Phe	Pro	Phe	Ser	Ile	Tyr	Asn	Phe	Ala
				165						170					175
Gln	Trp	Ala	Arg	Ala	Thr	Leu	Met	Pro	Ile	Ala	Val	Leu	Ser	Ala	His
				180						185					190
Arg	Pro	Ser	Arg	Pro	Leu	Ala	Pro	Gln	Asp	Arg	Leu	Asp	Ala	Leu	Phe
				195						200					205
Pro	Gln	Gly	Arg	Asp	Ser	Phe	Asn	Tyr	Asp	Leu	Pro	Ala	Arg	Leu	Gly
Ala	Gly	Val	Trp	Asp	Val	Ile	Phe	Arg	Lys	Ile	Asp	Thr	Ile	Leu	His
225															
				230											240
Arg	Leu	Gln	Asp	Trp	Gly	Ala	Arg	Arg	Gly	Pro	His	Gly	Ile	Met	Arg
				245											255
Arg	Gly	Ala	Ile	Asp	His	Val	Leu	Gln	Trp	Ile	Ile	Arg	His	Gln	Asp
				260											270
Tyr	Asp	Gly	Ser	Trp	Gly	Gly	Ile	Gln	Pro	Pro	Trp	Ile	Tyr	Gly	Leu
				275											285
Met	Ala	Leu	His	Thr	Glu	Gly	Tyr	Ala	Met	Thr	His	Pro	Val	Met	Ala
				290											300
Lys	Ala	Leu	Asp	Ala	Leu	Asn	Glu	Pro	Gly	Trp	Arg	Ile	Asp	Ile	Gly
305															
				310											320
Asp	Ala	Thr	Phe	Ile	Gln	Ala	Thr	Asn	Ser	Pro	Val	Trp	Asp	Thr	Met
				325											335
Leu	Ser	Leu	Leu	Ala	Phe	Asp	Asp	Ala	Gly	Leu	Gly	Glu	Arg	Tyr	Pro
				340											350
Glu	Gln	Val	Glu	Arg	Ala	Val	Arg	Trp	Val	Leu	Lys	Arg	Gln	Val	Leu
				355											365
Val	Pro	Gly	Asp	Trp	Ser	Val	Lys	Leu	Pro	Asp	Val	Lys	Pro	Gly	Gly
				370											380
Trp	Ala	Phe	Glu	Tyr	Ala	Asn	Asn	Phe	Tyr	Pro	Asp	Thr	Asp	Asp	Thr
385															
				390											400
Ser	Val	Ala	Leu	Met	Ala	Leu	Ala	Pro	Phe	Arg	His	Asp	Pro	Lys	Trp
				405											415
Gln	Ala	Glu	Gly	Ile	Glu	Asp	Ala	Ile	Gln	Arg	Gly	Ile	Asp	Trp	Leu
				420											430
Val	Ala	Met	Gln	Cys	Lys	Glu	Gly	Gly	Trp	Gly	Ala	Phe	Asp	Lys	Asp
				435											445
Asn	Asp	Lys	Lys	Ile	Leu	Ala	Lys	Ile	Pro	Phe	Cys	Asp	Phe	Gly	Glu

450	455	460
Ala Leu Asp Pro Pro Ser	Ala Asp Val Thr	Ala His Ile Ile Glu Ala
465	470	475
Phe Ala Lys Val Gly Leu	Asp Arg Asn His	Pro Ser Ile Val Arg Ala
485	490	495
Leu Asp Tyr Leu Lys Arg	Glu Gln Glu Pro	Glu Gly Pro Trp Phe Gly
500	505	510
Arg Trp Gly Val Asn Tyr	Val Tyr Gly Thr	Gly Ala Val Leu Pro Ala
515	520	525
Leu Ala Ala Ile Gly Glu	Asp Met Arg Gln	Pro Tyr Ile Ala Arg Ala
530	535	540
Cys Asp Trp Leu Ile Ala	Arg Gln Gln Ala	Asn Gly Gly Trp Gly Glu
545	550	555
Ser Cys Val Ser Tyr Met	Asp Ala Lys Gln	Ala Gly Glu Gly Thr Ala
565	570	575
Thr Ala Ser Gln Thr Ala	Trp Ala Leu Met	Ala Leu Ile Ala Ala Asp
580	585	590
Arg Pro Gln Asp Arg Asp	Ala Ile Glu Arg	Gly Cys Leu Tyr Leu Thr
595	600	605
Glu Thr Gln Arg Asp Gly	Thr Trp Gln Glu	Val His Tyr Thr Gly Thr
610	615	620
Gly Phe Pro Gly Tyr Gly	Val Gly Gln Thr	Ile Lys Leu Asn Asp Pro
625	630	635
Leu Leu Ser Lys Arg Leu	Met Gln Gly Pro	Glu Leu Ser Arg Ser Phe
645	650	655
Met Leu Arg Tyr Asp Leu	Tyr Arg His Tyr	Phe Pro Met Met Ala Ile
660	665	670
Gly Arg Val Leu Arg Gln	Arg Gly Asp Arg	Ser Gly His
675	680	685
<210> 158		
<400> 158		
000		
<210> 159		
<211> 679		
<212> PRT		
<213> 天蓝色链霉菌(Streptomyces coelicolor)		
<400> 159		
Met Thr Ala Thr Thr Asp	Gly Ser Thr Gly	Ala Ser Leu Arg Pro Leu
1	5	10
		15

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

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

Arg Gln Val Phe Pro Leu Thr Ala Leu Gly Arg Tyr Val His Gly Glu
 645 650 655
 Pro Phe Ala Lys Lys Pro Arg Ala Ala Asp Ala Pro Ala Glu Ala Ala
 660 665 670
 Pro Ala Glu Val Lys Gly Ser
 675
 <210> 160
 <400> 160
 000
 <210> 161
 <211> 725
 <212> PRT
 <213> 运动发酵单胞菌(*Zymomonas mobilis*)
 <400> 161
 Met Gly Ile Asp Arg Met Asn Ser Leu Ser Arg Leu Leu Met Lys Lys
 1 5 10 15
 Ile Phe Gly Ala Glu Lys Thr Ser Tyr Lys Pro Ala Ser Asp Thr Ile
 20 25 30
 Ile Gly Thr Asp Thr Leu Lys Arg Pro Asn Arg Arg Pro Glu Pro Thr
 35 40 45
 Ala Lys Val Asp Lys Thr Ile Phe Lys Thr Met Gly Asn Ser Leu Asn
 50 55 60
 Asn Thr Leu Val Ser Ala Cys Asp Trp Leu Ile Gly Gln Gln Lys Pro
 65 70 75 80
 Asp Gly His Trp Val Gly Ala Val Glu Ser Asn Ala Ser Met Glu Ala
 85 90 95
 Glu Trp Cys Leu Ala Leu Trp Phe Leu Gly Leu Glu Asp His Pro Leu
 100 105 110
 Arg Pro Arg Leu Gly Asn Ala Leu Leu Glu Met Gln Arg Glu Asp Gly
 115 120 125
 Ser Trp Gly Val Tyr Phe Gly Ala Gly Asn Gly Asp Ile Asn Ala Thr
 130 135 140
 Val Glu Ala Tyr Ala Ala Leu Arg Ser Leu Gly Tyr Ser Ala Asp Asn
 145 150 155 160
 Pro Val Leu Lys Lys Ala Ala Ala Trp Ile Ala Glu Lys Gly Gly Leu
 165 170 175
 Lys Asn Ile Arg Val Phe Thr Arg Tyr Trp Leu Ala Leu Ile Gly Glu
 180 185 190
 Trp Pro Trp Glu Lys Thr Pro Asn Leu Pro Pro Glu Ile Ile Trp Phe

195	200	205
Pro Asp Asn Phe Val Phe Ser Ile Tyr Asn Phe Ala Gln Trp Ala Arg		
210	215	220
Ala Thr Met Val Pro Ile Ala Ile Leu Ser Ala Arg Arg Pro Ser Arg		
225	230	235
Pro Leu Arg Pro Gln Asp Arg Leu Asp Glu Leu Phe Pro Glu Gly Arg		
245	250	255
Ala Arg Phe Asp Tyr Glu Leu Pro Lys Lys Glu Gly Ile Asp Leu Trp		
260	265	270
Ser Gln Phe Phe Arg Thr Thr Asp Arg Gly Leu His Trp Val Gln Ser		
275	280	285
Asn Leu Leu Lys Arg Asn Ser Leu Arg Glu Ala Ala Ile Arg His Val		
290	295	300
Leu Glu Trp Ile Ile Arg His Gln Asp Ala Asp Gly Gly Trp Gly Gly		
305	310	315
Ile Gln Pro Pro Trp Val Tyr Gly Leu Met Ala Leu His Gly Glu Gly		
325	330	335
Tyr Gln Leu Tyr His Pro Val Met Ala Lys Ala Leu Ser Ala Leu Asp		
340	345	350
Asp Pro Gly Trp Arg His Asp Arg Gly Glu Ser Ser Trp Ile Gln Ala		
355	360	365
Thr Asn Ser Pro Val Trp Asp Thr Met Leu Ala Leu Met Ala Leu Lys		
370	375	380
Asp Ala Lys Ala Glu Asp Arg Phe Thr Pro Glu Met Asp Lys Ala Ala		
385	390	395
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
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
 Ile Asp Ser Ile Gly Lys Gly Pro Thr Thr Pro Ser Gln Thr Ala Trp
 610 615 620
 Ala Leu Met Gly Leu Ile Ala Ala Asn Arg Pro Glu Asp Tyr Glu Ala
 625 630 635 640
 Ile Ala Lys Gly Cys His Tyr Leu Ile Asp Arg Gln Glu Gln Asp Gly
 645 650 655
 Ser Trp Lys Glu Glu Glu Phe Thr Gly Thr Gly Phe Pro Gly Tyr Gly
 660 665 670
 Val Gly Gln Thr Ile Lys Leu Asp Asp Pro Ala Leu Ser Lys Arg Leu
 675 680 685
 Leu Gln Gly Ala Glu Leu Ser Arg Ala Phe Met Leu Arg Tyr Asp Phe
 690 695 700
 Tyr Arg Gln Phe Phe Pro Ile Met Ala Leu Ser Arg Ala Glu Arg Leu
 705 710 715 720
 Ile Asp Leu Asn Asn
 725

<210> 162

<211> 2178

<212> DNA

<213> 运动发酵单胞菌 (*Zymomonas mobilis*)

<400> 162

atgggtattg acagaatgaa tagcttaagt cgcttggttaa tgaagaagat tttcggggct 60
 gaaaaaacct cgtataaacc ggcttccgat accataatcg gaacggatac cctgaaaaga 120
 ccgaaccggc ggctgaacc gacggcaaaa gtcgacaaaa cgatattcaa gactatgggg 180
 aatagtctga ataataccct tgtttcagcc tgtgactggt tgatcggaca acaaaagccc 240
 gatggtcatt gggtcggtgc cgtggaatcc aatgcttcga tggaagcaga atggtgtctg 300
 gccttggtgtg ttttgggtct ggaagatcat ccgcttcgtc caagattggg caatgctctt 360

```

ttggaaatgc agcgggaaga tggctcttgg ggagtctatt tcggcgctgg aaatggcgat 420
atcaatgcca cggftgaagc ctatgcggcc ttgcggtctt tgggggtattc tgccgataat 480
cctgttttga aaaaagcggc agcatggatt gctgaaaaag gcggattaaa aaatatccgt 540
gtcttttacc gttattggct ggcgttgatc ggggaatggc cttgggaaaa gaccctaac 600
cttccccctg aaattatctg gttccctgat aattttgtct tttcgattta taattttgcc 660
caatgggcgc gggcaaccat ggtgccgatt gctattctgt ccgcgagacg accaagccgc 720
ccgctgcgcc ctcaagaccg attggatgaa ctgtttccag aaggccgcgc tcgctttgat 780
tatgaattgc cgaaaaaaga aggcatcgat ctttggtcgc aatttttccg aaccactgac 840
cgtggattac attgggttca gtccaatctg ttaaagcgca atagcttgcg tgaagccgct 900
atccgtcatg ttttggaatg gattatccgg catcaggatg ccgatggcgg ttgggggtga 960
attcagccac cttgggtcta tggtttgatg gcgttacatg gtgaaggcta tcagctttat 1020
catccggtga tggccaaggc tttgtcggct ttggatgac cgggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgtaa aagacgcaa ggccaggat cgttttacgc cggaaatgga taaggccgcc 1200
gattggcttt tggctcgaca ggtcaaagtc aaaggcgatt ggtcaatcaa actgcccgat 1260
gttgaacccg gtggatgggc atttgaatat gccaatgac gctatcccga taccgatgat 1320
accgccgtcg ctttgatcgc cttttcctct tatcgtgata aggaggagtg gcaaaagaaa 1380
ggcgttgagg acgccattac ccgtggggtt aattggttga tcgcatgca aagcgaatgt 1440
ggcggttggg gagcctttga taaggataat aacagaagta tcctttccaa aattcctttt 1500
tgtgatttcg gagaatctat tgatccgcct tcagtcgatg taacggcgca tgtttttagag 1560
gcctttggca ccttgggact gtcccgcgat atgccggtca tccaaaaagc gatcgactat 1620
gtccgttccg aacaggaagc cgaaggcgcg tggtttggtc gttggggcgt taattatata 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gaccagcct 1740
tacatcacca aggcttgcca ttggctggtc gcacatcagc aggaagacgg cggttggggc 1800
gaaagctgct ctctctatat ggagattgat tccattggga agggcccaac cacgccgtcc 1860
cagactgctt gggctttgat ggggttgatc gcggccaatc gtcccgaaga ttatgaagcc 1920
attgccaagg gatgccatta tctgattgat cgccaagagc aggatggtag ctggaaagaa 1980
gaagaattca ccggcaccgg attccccggt tatggcgtgg gtcagacgat caagttagat 2040
gatccggctt tatcgaaacg attgcttcaa ggcgtgaac tgtcacgggc gtttatgctg 2100
cgttatgatt ttatcggca attcttccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga

```

<210> 163

<211> 725

<212> PRT

<213> 运动发酵单胞菌

<400> 163

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

Tyr Gln Leu Tyr His Pro Val Met Ala Lys Ala Leu Ser Ala Leu Asp		
340	345	350
Asp Pro Gly Trp Arg His Asp Arg Gly Glu Ser Ser Trp Ile Gln Ala		
355	360	365
Thr Asn Ser Pro Val Trp Asp Thr Met Leu Ala Leu Met Ala Leu Lys		
370	375	380
Asp Ala Lys Ala Glu Asp Arg Phe Thr Pro Glu Met Asp Lys Ala Ala		
385	390	395
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
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
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
Ile Asp Ser Ile Gly Lys Gly Pro Thr Thr Pro Ser Gln Thr Ala Trp		
610	615	620
Ala Leu Met Gly Leu Ile Ala Ala Asn Arg Pro Glu Asp Tyr Glu Ala		
625	630	635
Ile Ala Lys Gly Cys His Tyr Leu Ile Asp Arg Gln Glu Gln Asp Gly		

	645	650	655	
Ser Trp Lys Glu Glu Glu Phe Thr Gly Thr Gly Phe Pro Gly Tyr Gly				
660	665	670		
Val Gly Gln Thr Ile Lys Leu Asp Asp Pro Ala Leu Ser Lys Arg Leu				
675	680	685		
Leu Gln Gly Ala Glu Leu Ser Arg Ala Phe Met Leu Arg Tyr Asp Phe				
690	695	700		
Tyr Arg Gln Phe Phe Pro Ile Met Ala Leu Ser Arg Ala Glu Arg Leu				
705	710	715	720	
Ile Asp Leu Asn Asn				
725				
<210> 164				
<211> 2178				
<212> DNA				
<213> 运动发酵单胞菌				
<400> 164				
atgggtattg acagaatgaa tagcttaagt cgcttggttaa tgaagaagat tttcggggct				60
gaaaaaacct cgtataaacc ggcttccgat accataatcg gaacggatac cctgaaaaga				120
ccgaaccggc ggctgaacc gacggcaaaa gtcgacaaaa cgatattcaa gactatgggg				180
aatagtctga ataataccct tgtttcagcc tgtgactggt tgatcggaca acaaaagccc				240
gatggtcatt gggtcggtgc cgtggaatcc aatgcttcga tggaagcaga atggtgtctg				300
gccttgtggt ttttgggtct ggaagatcat ccgcttcgtc caagattggg caatgctctt				360
ttggaaatgc agcgggaaga tggctcttgg ggagtctatt tcggcgctgg aaatggcgat				420
atcaatgcca cggttgaagc ctatgcggcc ttgcggtctt tggggtattc tgccgataat				480
cctgttttga aaaaagcggc agcatggatt gctgaaaaag gcggattaaa aaatatccgt				540
gtctttaccc gttattggct ggcgttgatc ggggaatggc cttgggaaaa gaccctaac				600
cttccccctg aaattatctg gttccctgat aattttgtct tttcgattta taattttgcc				660
caatgggcgc gggcaaccat ggtgccgatt gctattctgt ccgcgagacg accaagccgc				720
ccgctgcgcc ctcaagaccg attggatgaa ctgtttccag aaggccgcgc tcgctttgat				780
tatgaattgc cgaaaaaaga aggcacgat ctttggtcgc aatttttccg aaccactgac				840
cgtggattac attgggttca gtccaatctg ttaaagcgca atagcttgcg tgaagccgct				900
atccgtcatg ttttggaatg gattatccgg catcaggatg ccgatggcgg ttgggggtgga				960
attcagccac cttgggtcta tggtttgatg gcgttacatg gtgaaggcta tcagctttat				1020
catccggtga tggccaaggc tttgtcggtt ttggatgac cgggttggcg acatgacaga				1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg				1140
atggcggttaa aagacgccaa ggccgaggat cgttttacgc cggaaatgga taaggccgcc				1200
gattggcttt tggctcgaca ggtcaaagtc aaaggcgatt ggtcaatcaa actgcccgat				1260
gttgaaccgc gtggatgggc atttgaatat gccaatgatc gctatcccga taccgatgat				1320
accgccgtcg ctttgatcgc ctttcctct tatcgtgata aggaggagtg gcaaaagaaa				1380

ggcggttgagg acgccattac ccgtgggggtt aattggttga tcgcatgca aagcgaatgt 1440
 ggcggttgagg gagcctttga taaggataat aacagaagta tcctttccaa aattcctttt 1500
 tgtgatttcg gagaatctat tgatccgcct tcagtcgatg taacggcgca tgttttagag 1560
 gcctttggca ccttgggact gtcccgcgat atgccggtca tccaaaaagc gatcgactat 1620
 gtccgttccg aacaggaagc cgaaggcgcg tggtttggtc gttggggcgt taattatata 1680
 tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gaccagcct 1740
 tacatcacca aggcttgcga ttggctggtc gcacatcagc aggaagacgg cggttggggc 1800
 gaaagctgct cttcctatat ggagattgat tccattggga agggcccaac cacgccgtcc 1860
 cagactgctt gggctttgat ggggttgatc gcggccaatc gtcccgaaga ttatgaagcc 1920
 attgccaagg gatgccatta tctgattgat cgccaagagc aggatggtag ctggaaagaa 1980
 gaagaattca ccggcaccgg attccccggt tatggcggtg gtcagacgat caagttggat 2040
 gatccggctt tatcgaaacg attgcttcaa ggcgctgaac tgtcacgggc gtttatgctg 2100
 cgttatgatt tttatcgga attcttcccg attatggcgt taagtcgggc agagagactg 2160
 attgatttga ataattga 2178

<210> 165

<211> 1896

<212> DNA

<213> 酸热脂环酸芽孢杆菌 (*Alicyclobacillus acidocaldarius*)

<400> 165

atggctgagc agttggtgga agcgccggcc tacgcgcgga cgctggatcg cgcggtggag 60
 tatctcctct cctgcaaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
 acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
 atggagaaga tccggcggta cctgttgac gagcagcgcg aggacggcac gtggggcctg 240
 taccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
 atcggcgatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcac tcagagccag 360
 ggcgggatcg agtcgtcgcg cgtgttcacg cggtatgtggc tggcgctggt gggagaatat 420
 ccatgggaga aggtgccccat ggtcccgcg gagatcatgt tcctcgcaa gcgtatgccg 480
 ctcaacattt acgagtttgg ctcgtgggcc cggcgaccg tcgtggcgat ctcaattgtc 540
 atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgat 600
 accgacgtgc ctccgcgccg gcgcggcgcc aaggaggcg gcgggcgaat cttcgacgcg 660
 ctggatcgcg ccctgcacgg gtatcagaag ctgtcggtgc accggttccg ccgcgcggcc 720
 gagatccggg ccttgactg gctgctcgag cgccaggccg gagacggcag ttggggcggg 780
 attcagccgc cctggtttta tacgtcatc gcgtcaaga ttctcgacat gacgcagcat 840
 ccggcgttca tcaagggtt ggagggcctc gagctgtacg gactggacct cgactacggc 900
 ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggtcttgc cgtgctcgcg 960
 ctgcgcgccg cggggcttcc ggccgatcac gaccggttg tcaaggcggg cgagtggctt 1020
 ttggaccggc agatcaccgt gccgggagac tgggcggtga agcggccgaa cctcaaaccg 1080
 ggcggttctg cgttccagtt cgacaacgtg tattaccggc acgtcgacga cacggccgtc 1140
 gtggtctggg cgctgaacag ccttcgctt ccggacgagc gccgcaggcg ggacgtgatg 1200

acgaaggggt	tccgctggat	cgctcggtatg	cagagttcca	acggcggctg	gggcgcgtac	1260
gacgtcgaca	acacgagcga	tctgccaaac	cacatcccgt	tctgcgactt	cggcgaagtg	1320
accgatccgc	cgctcggagga	cgtcaccgcc	cacgtgctcg	agtgtttcgg	cagcttcggc	1380
tacgacgacg	cctggaaggt	gatccggcgc	gcggtcgagt	acctcaagcg	cgaacagcgc	1440
ccggatggca	gttggtttgg	ccgctggggc	gtcaactacc	tgtacggcac	gggagcggtc	1500
gtgcccgcgc	tgaaggccgt	cgggatcgac	gtgcgcgagc	cgttcattca	gaaggcgctc	1560
gattgggtcg	agcagcatca	gaacccggac	ggtggctggg	gcgaggactg	ccgctcgtac	1620
gaggatccgg	cgtacgcggg	caaggcgcg	agcaccgccg	cgcagacggc	ttgggcgctg	1680
atggcactca	tcgcgggcgg	cagggcggag	tcggattccg	tgcgccggcg	cgtgcaatat	1740
ctcgtggaga	cgcagcgcgc	ggacggcggc	tgggatgagc	cgtactacac	cggcacgggc	1800
ttcccgggcg	atttctacct	cggtacacc	atgtaccgcc	acgtgtttcc	gacgctcgcg	1860
ctcggccgct	acaagcaagc	catcgagcgc	aggtga			1896
<210>	166					
<211>	2178					
<212>	DNA					
<213>	运动发酵单胞菌					
<400>	166					
atgggtattg	acagaatgaa	tagcttaagt	cgcttgtaa	tgaagaagat	tttcggggct	60
gaaaaaacct	cgtataaacc	ggcttccgat	accataatcg	gaacggatac	cctgaaaaga	120
ccgaaccggc	ggcctgaacc	gacggcaaaa	gtcgacaaaa	cgatattcaa	gactatgggg	180
aatagtctga	ataataccct	tgtttcagcc	tgtgactggt	tgatcggaca	acaaaagccc	240
gatggtcatt	gggtcggcgc	cgtggaatcc	aatgcttcga	tggaagcaga	atgggtgtctg	300
gccttggtgt	ttttgggtct	ggaagatcat	ccgcttcgtc	caagattggg	caatgctctt	360
ttggaaatgc	agcgggaaga	tggctcttgg	ggagtctatt	tcggcgctgg	aaatggcgat	420
atcaatgcca	cggftgaagc	ctatgcggcc	ttgcggctct	tgggggtattc	tgccgataat	480
cctgttttga	aaaaagcggc	agcatggatt	gctgaaaaag	gcggattaaa	aaatatccgt	540
gtcttttacc	gttattggct	ggcgttgatc	ggggaatggc	cttgggaaaa	gaccacctaac	600
cttccccctg	aaattatctg	gttcctgat	aattttgtct	tttcgattta	taattttgcc	660
caatgggcgc	gggcaaccat	ggtgccgatt	gctattctgt	ccgcgagacg	accaagccgc	720
ccgctgcgcc	ctcaagaccg	attggatgaa	ctgtttccag	aaggccgcgc	tcgctttgat	780
tatgaattgc	cgaaaaaaga	aggcatcgat	ctttggctgc	aatttttccg	aaccactgac	840
cgtggattac	attgggttca	gtccaatctg	ttaaagcgca	atagcttgcg	tgaagccgct	900
atccgtcatg	ttttggaatg	gattatccgg	catcaggatg	ccgatggcgg	ttgggggtgga	960
attcagccac	cttgggtcta	tggtttgatg	gcgttacatg	gtgaaggcta	tcagctttat	1020
catccggtga	tggccaaggc	ttgtcggct	ttggatgatc	ccggttggcg	acatgacaga	1080
ggcgagtctt	cttggataca	ggccaccaat	agtccggtat	gggatacaat	gttggccttg	1140
atggcgtaa	aagacgcaa	ggccgaggat	cgttttacgc	cggaaatgga	taaggccgcc	1200
gattggcttt	tggtcgcaca	ggtcaaagtc	aaaggcgatt	ggtcaatcaa	actgcccgat	1260
gttgaacccg	gtggatgggc	atttgaatat	gccaatgatc	gctatcccga	taccgatgat	1320

accgccgtcg	ctttgatcgc	cctttcctct	tatcgtgata	aggaggagtg	gcaaaagaaa	1380
ggcgttgagg	acgccattac	ccgtgggggtt	aattggttga	tcgccatgca	aagcgaatgt	1440
ggcggttggg	gagcctttga	taaggataat	aacagaagta	tcctttccaa	aattcctttt	1500
tgtgatttcg	gagaatctat	tgatccgcct	tcagtcgatg	taacggcgca	tgttttagag	1560
gcctttggca	ccttgggact	gtcccgcgat	atgccgggtca	tccaaaaagc	gatcgactat	1620
gtccgttccg	aacaggaagc	cgaaggcgcg	tggtttggtc	gttggggcgt	taattatata	1680
tatggcaccg	gtgcggttct	gcctgctttg	gcggcgatcg	gtgaagatat	gaccagcct	1740
tacatcacca	aggcttgcca	ttggctggtc	gcacatcagc	aggaagacgg	cggttggggc	1800
gaaagctgct	cttcctatat	ggagattgat	tccattggga	agggcccaac	cacgccgtcc	1860
cagactgctt	gggctttgat	ggggttgatc	gcggccaatc	gtcccgaaga	ttatgaagcc	1920
attgccaagg	gatgccatta	tctgattgat	cgccaagagc	aggatggtag	ctggaaagaa	1980
gaagaattca	ccggcaccgg	attccccggt	tatggcgtgg	gtcagacgat	caagttggat	2040
gatccggctt	tatcgaaaacg	attgcttcaa	ggcgtgaac	tgtcacgggc	gtttatgctg	2100
cgttatgatt	tttatcggca	attcttcccc	attatggcgt	taagtcgggc	agagagactg	2160
attgatttga	ataattga					2178
<210>	167					
<211>	1977					
<212>	DNA					
<213>	运动发酵单胞菌					
<400>	167					
atgactgtat	cgacttcctc	ggcttttcat	catagcccgt	tgtctgatga	tgttgagccg	60
attatccaaa	aggccacccg	tgctttgctt	gagaagcagc	agcaggatgg	ccattggggtt	120
tttgaattgg	aagccgatgc	aaccattccc	gctgaataca	tcctgttaaa	gcattatttg	180
ggtgaacccg	aagatttaga	aatagaggcc	aagataggtc	gctatttgcg	tcgtattcag	240
ggcgagcatg	gcggatggtc	tttgttttat	ggtgggtgatc	ttgatttgag	cgccacggtc	300
aaagcctatt	ttgccttgaa	aatgatcgga	gattctcctg	atgcgcctca	tatgcttcga	360
gccagaaatg	aaattttggc	acgggggtggg	gcgatgcgtg	ccaatgtctt	tacacgtatt	420
caattagctc	tgttcggggc	aatgtcatgg	gagcatgtcc	ctcaaagtcc	cgtagagttg	480
atgttgatgc	cggaatgggt	tccggttcac	atcaataaaa	tggcctattg	ggcaagaacc	540
gttttagtcc	cgttattgggt	tttacaggcg	ttaaagcctg	tcgcccgtaa	tcggcgcggt	600
atcttggttg	atgaattatt	tgtgccggat	gttttaccga	cccttcagga	aagcgggtgac	660
cctatatggc	gtcgtttttt	ttcggcactt	gataaggatg	tgcataaagt	agaaccttat	720
tggccgaaaa	atatgcgcgc	gaaggtctatt	catagctgtg	tccattttgt	gaccgagcgt	780
ttgaatgggtg	aagacggggt	gggtgctatt	tatccggcga	ttgccaatag	cgtcatgatg	840
tatgatgcct	tgggatatcc	cgaaaacat	ccagaaagag	ccattgcccc	tcgggctgtc	900
gaaaaattga	tgggtgtaga	tggcacggaa	gatcagggtg	ataaagaagt	ctactgtcag	960
ccttgtttat	ccccgatttg	ggataccgct	ttggttgccc	atgcatgtt	ggaagtcgga	1020
ggcgtatagg	ctgaaaaatc	ggctatttct	gccttgagct	ggttaaagcc	gcaacaaatt	1080
ttggatgtaa	agggcgattg	ggcatggcgg	cggcctgatc	tcagaccg	gggatgggcc	1140

tttcaatata	gaaatgacta	ttatcccgat	gtcgaatgata	cggctgttgt	gactatggcg	1200
atggatcgag	ccgcaaaatt	gtcggatctt	cacgatgatt	ttgaggaatc	taaagcgcgt	1260
gccatggaat	ggaccattgg	gatgcaaagc	gataatggcg	gttggggcgc	tttcgatgcc	1320
aataacagct	atacttatct	gaataatatt	ccctttgctg	atcatggcgc	gttacttgat	1380
ccgccaacgg	tcgatgtctc	ggcacgctgc	gtttcaatga	tggcgcaagc	cggtatctcg	1440
attacagatc	ccaaaaatgaa	agcggcagtt	gattatcttc	tgaaagagca	agaagaggat	1500
ggtagctggg	tcgggcgttg	gggtgtcaat	tacatatatg	gcacatggtc	ggccttatgt	1560
gcattgaatg	tggccgcttt	accccatgat	catttagctg	ttcagaaagc	tgtggccttg	1620
ctgaaaacta	ttcaaaatga	agatggtggg	tggggtgaaa	attgcgatag	ctatgccctt	1680
gattatagcg	gatacgagcc	gatggattcg	acggcttccc	aaacagcatg	ggctttattg	1740
ggcttgatgg	ctgttgggga	agctaattcc	gaggccgtga	caaagggtat	aaactggttg	1800
gcacaaaatc	aggatgaaga	aggattgtgg	aaagaagatt	attatagtgg	cggtggtttt	1860
ccccgtgttt	tttatcttcg	gtatcacggg	tattccaaat	attttcctct	ttgggccttta	1920
gcgcgctatc	gcaatttgaa	aaaagccaat	cagccgattg	ttcattatgg	gatgtaa	1977
<210> 168						
<211> 2055						
<212> DNA						
<213> 慢生大豆根瘤菌(<i>Bradyrhizobium japonicum</i>)						
<400> 168						
atgactgtga	ccagctcggc	ctccgcgcgt	gcgacgcgcg	acccgggaaa	ttatcagact	60
gccctgcaat	cgacggtgcg	cgcggcggcg	gatttggtga	tcgccaacca	gaagccggac	120
ggccattggg	tcggccgcgc	cgagtccaat	gcctgcatgg	aggcgcaatg	gtgcctcgcg	180
ctgtgggttca	tggggctcga	ggaccatccg	ctgcgcaagc	gcctgggcca	gtcgtgctc	240
gacagccagc	gcccggacgg	cgcttggcag	gtctatttcg	gcgcccccaa	tggcgacatc	300
aacgcgactg	tcgaggccta	tgccgcgctc	cgctcgcgtg	gcttcgcga	cgacgagccg	360
gcggtgcgcc	gggcgcggga	atggatcgag	gccaaggcg	gcctgcgcaa	catccgcgtc	420
ttcacccgct	actggctggc	actgatcggc	gaatggccgt	gggagaagac	accgaacatc	480
ccgccggagg	tgatctggtt	tccgctctgg	tttccgttct	cgatctacaa	tttcgcgcaa	540
tgggcccgcg	ccaccttgat	gccgatcgcc	gtgctgtcgg	cgcggcggcc	gagccggccg	600
ctgccgccgg	agaaccgcct	cgatgcgctg	tttccgcatg	gacggaaggc	gttcgactac	660
gaactgccgg	tcaaggccgg	cgccggcggc	tgggacaggt	tcttccgcgg	cgccgacaag	720
gttctgcaca	agctgcagaa	cctcggcaac	cgtctcaate	tcggcctggt	ccgcccggcg	780
gccaccagcc	gcgtgctgga	atggatgata	cgccatcagg	atttcgacgg	cgcctggggc	840
ggcatccagc	cgccttgat	ctacgggctg	atggcgtctt	atgccgaagg	ctatccgctc	900
aatcatcccc	tgctcgcaaa	gggcctcgac	gcgctgaacg	atccccggtg	gcgcgtcgat	960
gtcggtgacg	ccacctacat	ccaggccacc	aacagcccgg	tctgggacac	gacctgacc	1020
ttgctgcct	tcgacgatgc	cggcgtgctc	ggcgactatc	ccgaggccgt	cgacaaggcg	1080
gtcgactggg	tgctgcagcg	gcaggtgcgc	gtgcccggcg	actggtcgat	gaagctgccg	1140
catgtcaagc	ccggcggctg	ggcgttcgaa	tacgccaaca	actactatcc	cgacacggac	1200

gacaccgcgg	tcgcgctgat	cgcgctggcg	ccactgcgcc	acgatccgaa	atggaaggcc	1260
aaagggatcg	acgaggctat	ccagctcggt	gtcgactggc	tgatcggc	gcagagccag	1320
ggcggcggt	ggggcggtt	cgacaaggac	aacaaccaga	agatcctgac	caagatccc	1380
ttctgcgatt	atggcgaggc	gctcgatccg	ccctcggtcg	acgtcacccg	ccacatcatc	1440
gaggcggttcg	gcaagctcgg	catctcgcg	aaccatccgt	cgatggtgca	ggcgctggac	1500
tatatctgcc	gtgagcagga	gccgagcggt	ccgtgggttcg	gccgctgggg	cgtaattac	1560
gtctacggca	ccggcgcggt	gctgccggcg	ctggccgcga	tcggcgagga	catgaccag	1620
ccctatatcg	gccgcgcctg	cgactggctg	gttgcccatc	agcaggccga	tggcggtgg	1680
ggcgagagct	gcgccctccta	catggatgtc	agcgcggtcg	gccgcggcac	cacaacggcc	1740
tcgcagaccg	cctggggcgct	gatggcgctg	ctcgccgcc	atcgccccca	ggacaaggac	1800
gcgatcgagc	gtggctgcat	gtggctggtc	gagcgccagt	cggccggcac	ctgggacgag	1860
ccggaattca	ccggcaccgg	tttcccgggc	tacggcgctc	gccagaccat	caagctgaac	1920
gatcccgcgc	tgtcgcagcg	gctgatgcag	ggcccgaat	tgtcccgcgc	cttcatgtc	1980
cgctacggca	tgtaccgcca	ctacttccc	ctgatggcg	tcggccgcgc	cctacgccc	2040
cagagtcata	gctag					2055
<210>	169					
<211>	1896					
<212>	DNA					
<213>	酸热脂环酸芽孢杆菌 (<i>Alicyclobacillus acidocaldarius</i>)					
<400>	169					
atggctgagc	agttggtgga	agcgccggcc	tacgcgcgga	cgctggatcg	cgcggtggag	60
tatctcctct	cctgcaaaa	ggacgaaggc	tactggtggg	ggccgcttct	gagcaacgtc	120
acgatggaag	cggagtacgt	cctcttgtgc	cacattctcg	atcgcgtcga	tcgggatcgc	180
atggagaaga	tccggcggt	cctgttgac	gagcagcgcg	aggacggcac	gtgggccctg	240
tacccgggtg	ggccgcccga	cctcgacacg	accatcgagg	cgtacgtcgc	gctcaagtat	300
atcggcgatgt	cgcgcgacga	ggagccgatg	cagaaggcg	tccggttcat	tcagagccag	360
ggcgggatcg	agtcgtcgcg	cgtgttcacg	cggatgtggc	tggcgctggt	gggagaatat	420
ccgtgggaga	aggtgccc	ggtcccgcg	gagatcatgt	tcctcggcaa	gcgcgtgccg	480
ctcaacatct	acgagtttgg	ctcgtgggct	cggcgacccg	tcgtggcgct	ctcgattgtg	540
atgagccgcc	agccggtgtt	cccgtgccc	gagcggcgcg	gcgtgcccga	gctgtacgag	600
accgacgtgc	ctccgcgccg	gcgcggtgcc	aaggaggggg	gtgggtggat	cttcgacgcg	660
ctcgaccggg	cgctgcacgg	gtatcagaag	ctgtcggtgc	accggttccg	ccgcgcggcc	720
gagatccgcg	ccttggaactg	gttgctcgag	cgccaggccg	gagacggcag	ctggggcggg	780
attcagccgc	cttggtttta	cgcgctcatc	gcgtcaaga	ttctcgacat	gacgcagcat	840
ccggcgttca	tcaagggtg	ggaaggtcta	gagctgtacg	gcgtggagct	ggattacgga	900
ggatggatgt	ttcaggcttc	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	gggcgcctac	1260
gacgtcgaca	acacgagcga	tctcccgaac	cacatcccgt	tctgcgactt	cggcgaagtg	1320
accgatccgc	cgtcagagga	cgtcaccgcc	cacgtgctcg	agtgtttcgg	cagcttcggg	1380
tacgatgacg	cctggaaggt	catccggcgc	gcggtggaat	atctcaagcg	ggagcagaag	1440
ccggacggca	gctggttcgg	tcgttggggc	gtcaattacc	tctacggcac	gggcgcggtg	1500
gtgtcggcgc	tgaaggcggt	cgggatcgac	acgcgcgagc	cgtacattca	aaaggcgctc	1560
gactgggtcg	agcagcatca	gaacccggac	ggcggtctgg	gcgaggactg	ccgctcgtac	1620
gaggatccgg	cgtacgcggg	taagggcgcg	agcaccccg	cgcagacggc	ctgggcgctg	1680
atggcgctca	tcgcgggcgg	cagggcggag	tccgaggccg	cgcgcccgcg	cgtgcaatac	1740
ctcgtggaga	cgcagcggcc	ggacggcggc	tgggatgagc	cgtactacac	cggcacgggc	1800
ttcccagggg	atttctacct	cggtacacc	atgtaccgcc	acgtgtttcc	gacgctcgcg	1860
ctcggccgct	acaagcaagc	catcgagcgc	aggtga			1896
<210>	170					
<211>	1896					
<212>	DNA					
<213>	人工序列					
<220>						
<223>	合成AacSHC衍生物					
<400>	170					
atggctgagc	agttggtgga	agcgccggcc	tacgcgcgga	cgctggatcg	cgcggtggag	60
tatctcctct	cctgccaaaa	ggacgaaggc	tactggtggg	ggccgcttct	gagcaacgtc	120
acgatggaag	cggagtacgt	cctcttgtgc	cacattctcg	atcgctcga	tcgggatcgc	180
atggagaaga	tccggcggta	cctgttgac	gagcagcgcg	aggacggcac	gtgggccctg	240
tacccgggtg	ggccgccgga	cctcgacacg	accatcgagg	cgtacgtcgc	gctcaagtat	300
atcggcattg	cgcgcgacga	ggagccgatg	cagaaggcgc	tccggttcac	tcagagccag	360
ggcgggatcg	agtcgtcgcg	cgtgttcacg	cggatgtggc	tggcgctggt	gggagaatat	420
ccgtgggaga	aggtgcccac	ggtcccgcgc	gagatcatgt	tcctcggcaa	gcgcattgcc	480
ctcaacatct	acgagtttgg	ctcgtgggct	cgggcgaccg	tcgtggcgct	ctcgattgtg	540
atgagccgcc	agccggtggt	cccgtgccc	gagcgggcgc	gcgtgcccga	gctgtacgag	600
accgacgtgc	ctccgcgcgc	gcgcggtgcc	aagggagggg	gtgggtggat	cttcgacgcg	660
ctcgaccggg	cgctgcacgg	gtatcagaag	ctgtcggtgc	acccgttccg	ccgcgcggcc	720
gagatccgcg	ccttggaactg	gttgctcgag	cgccaggccg	gagacggcag	ctggggcggg	780
attcagccgc	cttggtttta	cgcgtcctc	gcgtcaaga	ttctcgacat	gacgcagcat	840
ccggcggtca	tcaagggtg	ggaaggtcta	gagctgtacg	gcgtggagct	ggattacgga	900
ggatggatgt	ttcaggcttc	catctcgccg	gtgtgggaca	cgggcctcgc	cgtgctcgcg	960
ctgcgcgctg	cggggcttcc	ggccgatcac	gaccgcttgg	tcaaggcggg	cgagtggctg	1020
ttggaccggc	agatcacggt	tccgggcgac	tgggcggtga	agcggccgaa	cctcaagccg	1080
ggcggggttcg	cgttccagtt	cgacaacgtg	tactaccggg	acgtggacga	cacggccgtc	1140

gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
 acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgccctac 1260
 gacgtcgaca acacgagcga tctcccgaac cacatcccggt 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 tgaaggcgggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
 gactgggtcg agcagcatca gaacccggac ggcggtctggg gcgaggactg ccgctcgtac 1620
 gaggatccgg cgtacgcggg taaggcgcg agcaccctgt cgcagacggc ctgggcgctg 1680
 atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgcccgcg cgtgcaatac 1740
 ctcgtggaga cgcagcggc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
 ttcccagggg attggtacct cggtacacc atgtaccgcc acgtgtttcc gacgtcgcg 1860
 ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 171

<211> 631

<212> PRT

<213> 人工序列

<220>

<223> 合成AacSHC衍生物

<400> 171

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

145	150	155	160
Leu Asn Ile Tyr Glu Phe Gly Ser Trp Ala Arg Ala Thr Val Val Ala			
	165	170	175
Leu Ser Ile Val Met Ser Arg Gln Pro Val Phe Pro Leu Pro Glu Arg			
	180	185	190
Ala Arg Val Pro Glu Leu Tyr Glu Thr Asp Val Pro Pro Arg Arg Arg			
	195	200	205
Gly Ala Lys Gly Gly Gly Gly Trp Ile Phe Asp Ala Leu Asp Arg Ala			
	210	215	220
Leu His Gly Tyr Gln Lys Leu Ser Val His Pro Phe Arg Arg Ala Ala			
225	230	235	240
Glu Ile Arg Ala Leu Asp Trp Leu Leu Glu Arg Gln Ala Gly Asp Gly			
	245	250	255
Ser Trp Gly Gly Ile Gln Pro Pro Trp Phe Tyr Ala Leu Ile Ala Leu			
	260	265	270
Lys Ile Leu Asp Met Thr Gln His Pro Ala Phe Ile Lys Gly Trp Glu			
	275	280	285
Gly Leu Glu Leu Tyr Gly Val Glu Leu Asp Tyr Gly Gly Trp Met Phe			
	290	295	300
Gln Ala Ser Ile Ser Pro Val Trp Asp Thr Gly Leu Ala Val Leu Ala			
305	310	315	320
Leu Arg Ala Ala Gly Leu Pro Ala Asp His Asp Arg Leu Val Lys Ala			
	325	330	335
Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala			
	340	345	350
Val Lys Arg Pro Asn Leu Lys Pro Gly Gly Phe Ala Phe Gln Phe Asp			
	355	360	365
Asn Val Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Val Trp Ala			
	370	375	380
Leu Asn Thr Leu Arg Leu Pro Asp Glu Arg Arg Arg Asp Ala Met			
385	390	395	400
Thr Lys Gly Phe Arg Trp Ile Val Gly Met Gln Ser Ser Asn Gly Gly			
	405	410	415
Trp Gly Ala Tyr Asp Val Asp Asn Thr Ser Asp Leu Pro Asn His Ile			
	420	425	430
Pro Phe Cys Asp Phe Gly Glu Val Thr Asp Pro Pro Ser Glu Asp Val			
	435	440	445
Thr Ala His Val Leu Glu Cys Phe Gly Ser Phe Gly Tyr Asp Asp Ala			
450	455	460	

Trp Lys Val Ile Arg Arg Ala Val Glu Tyr Leu Lys Arg Glu Gln Lys
 465 470 475 480
 Pro Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Leu Tyr Gly
 485 490 495
 Thr Gly Ala Val Val Ser Ala Leu Lys Ala Val Gly Ile Asp Thr Arg
 500 505 510
 Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn
 515 520 525
 Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala
 530 535 540
 Tyr Ala Gly Lys Gly Ala Ser Thr Pro Ser Gln Thr Ala Trp Ala Leu
 545 550 555 560
 Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg
 565 570 575
 Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp
 580 585 590
 Glu Pro Tyr Tyr Thr Gly Thr Gly Phe Pro Gly Asp Trp 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> 172

<211> 2178

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 172

atgggtattg acagaatgaa tagcttaagt cgcttggttaa tgaagaagat tttcggggct 60
 gaaaaaacct cgtataaaccc ggcttccgat accataatcg gaacggatac cctgaaaaga 120
 ccgaaccggc ggcttgaacc gacggcaaaa gtcgacaaaa cgatattcaa gactatgggg 180
 aatagtctga ataataccct tgtttcagcc tgtgactggt tgatcggaca acaaaagccc 240
 gatggtcatt gggctcgggtc cgtggaatcc aatgcttcga tggaagcaga atggtgtctg 300
 gccttgtggt ttttgggtct ggaagatcat ccgcttcgtc caagattggg caatgctctt 360
 ttggaaatgc agcgggaaga tggtctcttg ggagtctatt tcggcgctgg aaatggcgat 420
 atcaatgcca cggttgaagc ctatgcggcc ttgcggtctt tggggtattc tgccgataat 480
 cctgttttga aaaaagcggc agcatggatt gctgaaaaag gcgattaaa aaatatccgt 540
 gtctttaccc gttattggct ggcgttgatc ggggaatggc cttgggaaaa gaccctaac 600

```

cttccccctg aaattatctg gttccctgat aattttgtct tttcgattta taattttgcc 660
caatgggcgc gggcaaccat ggtgccgatt gctattctgt ccgcgagacg accaagccgc 720
ccgctgcgcc ctcaagaccg attggatgaa ctgtttccag aaggccgcgc tcgctttgat 780
tatgaattgc cgaaaaaaga aggcatcgat ctttggtcgc aatttttccg aaccactgac 840
cgtggattac attgggttca gtccaatctg ttaaagcgca atagcttgcg tgaagccgct 900
atccgtcatg ttttggaatg gattatccgg catcaggatg ccgatggcgg ttgggggtgga 960
attcagccac cttgggtcta tggtttgatg gcgttacatg gtgaaggcta tcagctttat 1020
catccggtga tggccaaggc tttgtcggct ttggatgacg ccggttggcg acatgacaga 1080
ggcagagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcggttaa aagacgccaa ggccgaggat cgttttacgc cggaaatgga taaggccgcc 1200
gattggcttt tggctcgaca ggtcaaagtc aaaggcgatt ggtcaatcaa actgcccgat 1260
gttgaacccg gtggatgggc atttgaatat gccaatgacg gctatcccga taccgatgat 1320
accgccgtcg ctttgatcgc ctttctctct tatcgtgata aggaggagtg gcaaaagaaa 1380
ggcgttgagg acgccattac ccgtgggggtt aattggttga tcgccatgca aagcgaatgt 1440
ggcggttggg gagcctttga taaggataat aacagaagta tcctttccaa aattcctttt 1500
tgtgatttcg gagaatctat tgatccgcct tcagtcgatg taacggcgca tgtttttagag 1560
gcctttggca cttgggact gtcccgcgat atgccggtca tccaaaaagc gatcgactat 1620
gtccgttccg aacaggaagc cgaaggcgcg tggtttggtc gttggggcgt taattatata 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gaccagcct 1740
tacatcacca aggcttgca ttggctggtc gcacatcagc aggaagacgg cggttggggc 1800
gaaagctgct ctctctatat ggagattgat tccattggga agggcccaac cacgccgtcc 1860
cagactgctt gggctttgat ggggttgatc gcggccaatc gtcccgaaga ttatgaagcc 1920
attgccaagg gatgccatta tctgattgat cgccaagagc aggatggtag ctggaaagaa 1980
gaagaattca ccggcaccgg attccccggt tatggcgtgg gtcagacgat caagttggat 2040
gatccggctt tatcgaaacg attgcttcaa ggcgctgaac tgtcacgggc gtggatgctg 2100
cgttatgatt ttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

```

<210> 173

<211> 725

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC1衍生物

<400> 173

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		
85	90	95
Glu Trp Cys Leu Ala Leu Trp Phe Leu Gly Leu Glu Asp His Pro Leu		
100	105	110
Arg Pro Arg Leu Gly Asn Ala Leu Leu Glu Met Gln Arg Glu Asp Gly		
115	120	125
Ser Trp Gly Val Tyr Phe Gly Ala Gly Asn Gly Asp Ile Asn Ala Thr		
130	135	140
Val Glu Ala Tyr Ala Ala Leu Arg Ser Leu Gly Tyr Ser Ala Asp Asn		
145	150	155
Pro Val Leu Lys Lys Ala Ala Ala Trp Ile Ala Glu Lys Gly Gly Leu		
165	170	175
Lys Asn Ile Arg Val Phe Thr Arg Tyr Trp Leu Ala Leu Ile Gly Glu		
180	185	190
Trp Pro Trp Glu Lys Thr Pro Asn Leu Pro Pro Glu Ile Ile Trp Phe		
195	200	205
Pro Asp Asn Phe Val Phe Ser Ile Tyr Asn Phe Ala Gln Trp Ala Arg		
210	215	220
Ala Thr Met Val Pro Ile Ala Ile Leu Ser Ala Arg Arg Pro Ser Arg		
225	230	235
Pro Leu Arg Pro Gln Asp Arg Leu Asp Glu Leu Phe Pro Glu Gly Arg		
245	250	255
Ala Arg Phe Asp Tyr Glu Leu Pro Lys Lys Glu Gly Ile Asp Leu Trp		
260	265	270
Ser Gln Phe Phe Arg Thr Thr Asp Arg Gly Leu His Trp Val Gln Ser		
275	280	285
Asn Leu Leu Lys Arg Asn Ser Leu Arg Glu Ala Ala Ile Arg His Val		
290	295	300
Leu Glu Trp Ile Ile Arg His Gln Asp Ala Asp Gly Gly Trp Gly Gly		
305	310	315
Ile Gln Pro Pro Trp Val Tyr Gly Leu Met Ala Leu His Gly Glu Gly		
325	330	335
Tyr Gln Leu Tyr His Pro Val Met Ala Lys Ala Leu Ser Ala Leu Asp		
340	345	350

Asp Pro Gly Trp Arg His Asp Arg Gly Glu Ser Ser Trp Ile Gln Ala		
355	360	365
Thr Asn Ser Pro Val Trp Asp Thr Met Leu Ala Leu Met Ala Leu Lys		
370	375	380
Asp Ala Lys Ala Glu Asp Arg Phe Thr Pro Glu Met Asp Lys Ala Ala		
385	390	395 400
Asp Trp Leu Leu Ala Arg Gln Val Lys Val Lys Gly Asp Trp Ser Ile		
	405	410 415
Lys Leu Pro Asp Val Glu Pro Gly Gly Trp Ala Phe Glu Tyr Ala Asn		
	420	425 430
Asp Arg Tyr Pro Asp Thr Asp Asp Thr Ala Val Ala Leu Ile Ala Leu		
	435	440 445
Ser Ser Tyr Arg Asp Lys Glu Glu Trp Gln Lys Lys Gly Val Glu Asp		
	450	455 460
Ala Ile Thr Arg Gly Val Asn Trp Leu Ile Ala Met Gln Ser Glu Cys		
465	470	475 480
Gly Gly Trp Gly Ala Phe Asp Lys Asp Asn Asn Arg Ser Ile Leu Ser		
	485	490 495
Lys Ile Pro Phe Cys Asp Phe Gly Glu Ser Ile Asp Pro Pro Ser Val		
	500	505 510
Asp Val Thr Ala His Val Leu Glu Ala Phe Gly Thr Leu Gly Leu Ser		
	515	520 525
Arg Asp Met Pro Val Ile Gln Lys Ala Ile Asp Tyr Val Arg Ser Glu		
	530	535 540
Gln Glu Ala Glu Gly Ala Trp Phe Gly Arg Trp Gly Val Asn Tyr Ile		
545	550	555 560
Tyr Gly Thr Gly Ala Val Leu Pro Ala Leu Ala Ala Ile Gly Glu Asp		
	565	570 575
Met Thr Gln Pro Tyr Ile Thr Lys Ala Cys Asp Trp Leu Val Ala His		
	580	585 590
Gln Gln Glu Asp Gly Gly Trp Gly Glu Ser Cys Ser Ser Tyr Met Glu		
	595	600 605
Ile Asp Ser Ile Gly Lys Gly Pro Thr Thr Pro Ser Gln Thr Ala Trp		
	610	615 620
Ala Leu Met Gly Leu Ile Ala Ala Asn Arg Pro Glu Asp Tyr Glu Ala		
625	630	635 640
Ile Ala Lys Gly Cys His Tyr Leu Ile Asp Arg Gln Glu Gln Asp Gly		
	645	650 655
Ser Trp Lys Glu Glu Glu Phe Thr Gly Thr Gly Phe Pro Gly Tyr Gly		

660	665	670
Val Gly Gln Thr Ile Lys Leu Asp Asp Pro Ala Leu Ser Lys Arg Leu		
675	680	685
Leu Gln Gly Ala Glu Leu Ser Arg Ala Trp Met Leu Arg Tyr Asp Phe		
690	695	700
Tyr Arg Gln Phe Phe Pro Ile Met Ala Leu Ser Arg Ala Glu Arg Leu		
705	710	715
Ile Asp Leu Asn Asn		720
	725	

<210> 174

<211> 1977

<212> DNA

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 174

```

atgactgtat cgacttcctc ggcttttcat catagcccg tgtctgatga tgttgagccg      60
attatccaaa aggccacccg tgccttgctt gagaagcagc agcaggatgg ccattggggtt    120
tttgaattgg aagccgatgc aaccattccc gctgaataca tcctgttaaa gcattatttg    180
ggatgaacccg aagatttaga aatagaggcc aagataggtc gctatttgcg tcgtattcag    240
ggcgagcatg gcggatggtc tttgttttat ggtgggtgatc ttgatttgag cgccacggtc    300
aaagcctatt ttgccttgaa aatgatcgga gattctcctg atgcgcctca tatgcttcga    360
gccagaaatg aaattttggc acgggggtggg gcgatgcgtg ccaatgtctt tacacgtatt    420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaatgcc cgtagagttag    480
atgttgatgc cggaatgggt tccggttcac atcaataaaa tggcctattg ggcaagaacc    540
gttttagtcc cgttattggt tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt    600
atcttggttg atgaattatt tgtgccggat gttttaccga cccttcagga aagcgggtgac    660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat    720
tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt    780
ttgaatgggtg aagacggggt ggggtgctatt tatccggcga ttgccaatag cgtcatgatg    840
tatgatgcct tgggatatcc cgaanaacat ccagaaagag ccattgcccc tcgggctgtc    900
gaaaaattga tgggtgtaga tggcacggaa gatcagggtg ataaagaagt ctactgtcag    960
ccttgtttat ccccgatttg ggataccgct ttggttgccc atgcatgtt ggaagtcgga   1020
ggcgatgagg ctgaaaaatc ggctatttct gccttgagct ggttaaagcc gcaacaaatt   1080
ttggatgtaa agggcgattg ggcatggcgg cggcctgatc tcagaccgga gggatgggccc   1140
tttcaatata gaaatgacta ttatcccgat gtcgatgata cggctgttgt gactatggcg   1200
atggatcgag ccgcaaaaatt gtcggatctt cacgatgatt ttgaggaatc taaagcgcgt   1260
gccatggaat ggaccattgg gatgcaaagc gataatggcg gttggggcgc tttcgatgcc   1320
aataacagct atacttatct gaataatatt ccctttgctg atcatggcgc gttacttgat   1380

```

ccgccaacgg tcgatgtctc ggcacgctgc gtttcaatga tggcgcaagc cggtatctcg 1440
 attacagatc ccaaaatgaa agcggcagtt gattatcttc tgaaagagca agaagaggat 1500
 ggtagctggg tcgggcgttg ggggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
 gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggccttg 1620
 ctgaaaacta ttcaaaatga agatgggtggg tggggtgaaa attgcgatag ctatgccctt 1680
 gattatagcg gatacgagcc gatggattcg acggcttccc aaacagcatg ggctttattg 1740
 ggcttgatgg ctgttgggga agctaattcc gaggccgtga caaagggtat aaactggttg 1800
 gcacaaaatc aggatgaaga aggattgtgg aaagaagatt attatagtgg cgggtggtttt 1860
 ccccggtgtt ggtagcttcg gtatcacggg tattccaaat attttcctct ttgggcttta 1920
 gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 175

<211> 658

<212> PRT

<213> 人工序列

<220>

<223> 合成ZmoSHC2衍生物

<400> 175

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		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	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	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
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
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 Trp	
610	615	620
Tyr Leu Arg Tyr His Gly Tyr Ser	Lys Tyr Phe Pro Leu Trp Ala Leu	
625	630	635
Ala Arg Tyr Arg Asn Leu Lys Lys	Ala Asn Gln Pro Ile Val His Tyr	
645	650	655

Gly Met

<210> 176

<211> 2055

<212> DNA

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 176

```

atgactgtga ccagctcggc ctccgcgcgt gcgacgcgcg acccgggaaa ttatcagact      60
gccctgcaat cgacggtgcg cgcggcggcg gattggctga tcgccaacca gaagccggac      120
ggccattggg tcggccgcgc cgagtccaat gcctgcatgg aggcgcaatg gtgcctcgcg      180
ctgtggttca tggggctcga ggaccatccg ctgcgcaagc gcctgggcca gtcgctgctc      240
gacagccagc gcccggacgg cgcttggcag gtctatttcg gcgcccccaa tggcgacatc      300
aacgcgactg tcgaggccta tgccgcgctc cgctcgttgg gcttccgcga cgacgagccg      360
gcggtgcgcc gggcgcggga atggatcgag gccaaaggcg gcctgcgcaa catccgcgtc      420
ttcaccgcgt actggctggc actgatcggc gaatggccgt gggagaagac accgaacatc      480
ccgccggagg tgatctggtt tccgctctgg ttccggttct cgatctacaa tttcgcgcaa      540
tgggcccgcg ccacctgat gccgatcgcc gtgctgtcgg cgcggcggcc gagccggccc      600

```

ctgccgccgg agaaccgcct cgatgcgctg tttccgcatg gacggaaggc gttcgactac 660
 gaactgccgg tcaaggccgg cgccggcggc tgggacaggt tcttccgcgg cgccgacaag 720
 gttctgcaca agctgcagaa cctcggcaac cgtctcaatc tcggcctgtt ccgcccggcg 780
 gccaccagcc gcgtgctgga atggatgata cgccatcagg atttcgacgg cgcctggggc 840
 ggcatccagc cgccctggat ctacgggctg atggcgctct atgccgaagg ctatccgctc 900
 aatcatcccc tgctcgcaaa gggcctcgac gcgctgaacg atcccggctg gcgctcgat 960
 gtcggtgacg ccacctacat ccaggccacc aacagcccgg tctgggacac gatcctgacc 1020
 ttgctcgctt tcgacgatgc cggcgtgctc ggcgactatc ccgaggccgt cgacaaggcg 1080
 gtcgactggg tgctgcagcg gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
 catgtcaagc ccggcggtg ggcgttcgaa tacgccaaca actactatcc cgacacggac 1200
 gacaccgcgg tcgctgctgat cgcgtggcg ccaactgcgc acgatccgaa atggaaggcc 1260
 aaagggatcg acgaggttat ccagctcggg gtcgactggc tgatcggtat gcagagccag 1320
 ggcgcggtt ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagatcccc 1380
 ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcctc 1440
 gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgtggac 1500
 tatattcgcc gtgagcagga gccgagcggg ccgtggttcg gccgctgggg cgtcaattac 1560
 gtctacggca ccggcgcggt gctgccggcg ctggccgcga tcggcgagga catgaccag 1620
 ccctatatcg gccgcgcctg cgactggctg gttgccatc agcaggccga tggcggttg 1680
 ggcgagagct gcgcctccta catggatgtc agcgcggtcg gccgcggcac cacaacggcc 1740
 tcgcagaccg cctgggcgct gatggcgctg ctgccgccga atcgccccca ggacaaggac 1800
 gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
 ccggaattca ccggcaccgg tttcccgggc tacggcgctc gccagaccat caagctgaac 1920
 gatcccgcg cgtcgcagcg gctgatgcag ggcccgaat tgtcccgcgc ctggatgctc 1980
 cgctacggca tgtaccgcca ctacttcccg ctgatggcg ctcggccgcgc cctacgcccg 2040
 cagagtcata gctag 2055

<210> 177

<211> 684

<212> PRT

<213> 人工序列

<220>

<223> 合成BjapSHC1衍生物

<400> 177

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
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
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
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	395 400
Asp Thr Ala Val Ala Leu Ile Ala Leu Ala Pro Leu Arg His Asp Pro		
405	410	415
Lys Trp Lys Ala Lys Gly Ile Asp Glu Ala Ile Gln Leu Gly Val Asp		
420	425	430
Trp Leu Ile Gly Met Gln Ser Gln Gly Gly Gly Trp Gly Ala Phe Asp		
435	440	445
Lys Asp Asn Asn Gln Lys Ile Leu Thr Lys Ile Pro Phe Cys Asp Tyr		
450	455	460
Gly Glu Ala Leu Asp Pro Pro Ser Val Asp Val Thr Ala His Ile Ile		
465	470	475 480
Glu Ala Phe Gly Lys Leu Gly Ile Ser Arg Asn His Pro Ser Met Val		
485	490	495
Gln Ala Leu Asp Tyr Ile Arg Arg Glu Gln Glu Pro Ser Gly Pro Trp		
500	505	510
Phe Gly Arg Trp Gly Val Asn Tyr Val Tyr Gly Thr Gly Ala Val Leu		
515	520	525
Pro Ala Leu Ala Ala Ile Gly Glu Asp Met Thr Gln Pro Tyr Ile Gly		
530	535	540
Arg Ala Cys Asp Trp Leu Val Ala His Gln Gln Ala Asp Gly Gly Trp		
545	550	555 560
Gly Glu Ser Cys Ala Ser Tyr Met Asp Val Ser Ala Val Gly Arg Gly		
565	570	575
Thr Thr Thr Ala Ser Gln Thr Ala Trp Ala Leu Met Ala Leu Leu Ala		
580	585	590
Ala Asn Arg Pro Gln Asp Lys Asp Ala Ile Glu Arg Gly Cys Met Trp		
595	600	605
Leu Val Glu Arg Gln Ser Ala Gly Thr Trp Asp Glu Pro Glu Phe Thr		
610	615	620
Gly Thr Gly Phe Pro Gly Tyr Gly Val Gly Gln Thr Ile Lys Leu Asn		
625	630	635 640
Asp Pro Ala Leu Ser Gln Arg Leu Met Gln Gly Pro Glu Leu Ser Arg		
645	650	655
Ala Trp 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> 178	
<211> 631	
<212> PRT	
<213> 酸热脂环酸芽孢杆菌(<i>Alicyclobacillus acidocaldarius</i>)	
<400> 178	
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	
Ile 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 Asp Thr Asp Val Pro Pro Arg Arg Arg	
195 200 205	
Gly Ala Lys Gly Gly Gly Gly Arg Ile Phe Asp Ala Leu Asp Arg Ala	
210 215 220	
Leu His Gly Tyr Gln Lys Leu Ser Val His Pro Phe Arg Arg Ala Ala	
225 230 235 240	
Glu Ile Arg Ala Leu Asp Trp Leu Leu Glu Arg Gln Ala Gly Asp Gly	
245 250 255	
Ser Trp Gly Gly Ile Gln Pro Pro Trp Phe Tyr Thr 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 Asp 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 Ser Leu Arg Leu Pro Asp Glu Arg Arg Arg Arg Asp Val 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 Arg		
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 Pro Ala Leu Lys Ala Val Gly Ile Asp Val Arg		
500	505	510
Glu Pro Phe 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 Asp Ser Val Arg Arg		
565	570	575

Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp	
580	585
Glu Pro Tyr Tyr Thr Gly Thr Gly Phe Pro Gly Asp Phe Tyr Leu Gly	
595	600
Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr	
610	615
Lys Gln Ala Ile Glu Arg Arg	
625	630
<210>	179
<211>	382
<212>	DNA
<213>	人工序列
<220>	
<223>	合成质粒pET-28a (+) 克隆/表达区
<220>	
<221>	CDS
<222>	(109) .. (285)
<400>	179
agatctcgat cccgcgaaat taatagcact cactataggg gaattgtgag cggataacaa	60
ttccccctcta gaaataattt ttgtttaact ttaagaagga gatatacc atg ggc agc	117
Met Gly Ser	
1	
agc cat cat cat cat cat cac agc agc ggc ctg gtg ccg cgc ggc agc	165
Ser His His His His His His Ser Ser Gly Leu Val Pro Arg Gly Ser	
5	10
cat atg gct agc atg act ggt gga cag caa atg ggt cgc gga tcc gaa	213
His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Gly Ser Glu	
20	25
ttc gag ctc cgt cga caa gct tgc ggc cgc act cga gca cca cca cca	261
Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Thr Arg Ala Pro Pro Pro	
40	45
cca cca ctg aga tcc ggc tgc taa caaagcccga aaggaagctg agttggctgc	315
Pro Pro Leu Arg Ser Gly Cys	
55	
tgccaccgct gagcaataac tagcataacc ccttggggcc tctaaacggg tcttgagggg	375
tttttttg	382
<210>	180
<211>	58
<212>	PRT

<213> 人工序列

<220>

<223> 合成构建体

<400> 180

```
Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1           5           10           15
Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
          20          25          30
Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Thr Arg Ala
          35          40          45
Pro Pro Pro Pro Pro Leu Arg Ser Gly Cys
          50          55
```

<210> 181

<211> 94

<212> DNA

<213> 人工序列

<220>

<223> 合成质粒pET-28b(+) 克隆/表达区

<220>

<221> CDS

<222> (1) .. (72)

<400> 181

```
ggt cgg gat ccg aat tcg agc tcc gtc gac aag ctt gcg gcc gca ctc      48
Gly Arg Asp Pro Asn Ser Ser Ser Val Asp Lys Leu Ala Ala Ala Leu
1           5           10           15
gag cac cac cac cac cac cac tga gatccggctg ctaacaaagc cc          94
Glu His His His His His His
          20
```

<210> 182

<211> 23

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<400> 182

```
Gly Arg Asp Pro Asn Ser Ser Ser Val Asp Lys Leu Ala Ala Ala Leu
1           5           10           15
Glu His His His His His His
          20
```

<210>	183	
<211>	93	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成质粒pET-28c (+) 克隆/表达区	
<220>		
<221>	CDS	
<222>	(1) .. (93)	
<400>	183	
ggt cgg atc cga att cga gct ccg tcg aca agc ttg cgg ccg cac tcg		48
Gly Arg Ile Arg Ile Arg Ala Pro Ser Thr Ser Leu Arg Pro His Ser		
1 5 10 15		
agc acc acc acc acc acc act gag atc cgg ctg cta aca aag ccc		93
Ser Thr Thr Thr Thr Thr Thr Glu Ile Arg Leu Leu Thr Lys Pro		
20 25 30		
<210>	184	
<211>	31	
<212>	PRT	
<213>	人工序列	
<220>		
<223>	合成构建体	
<400>	184	
Gly Arg Ile Arg Ile Arg Ala Pro Ser Thr Ser Leu Arg Pro His Ser		
1 5 10 15		
Ser Thr Thr Thr Thr Thr Thr Glu Ile Arg Leu Leu Thr Lys Pro		
20 25 30		
<210>	185	
<211>	631	
<212>	PRT	
<213>	人工序列	
<220>		
<223>	合成SHC衍生物	
<220>		
<221>	变体	
<222>	(77) .. (77)	
<223>	T77X中的X 选自: A, C, D, E, F, G, H, I, K, L, M, N, P,	
	Q, R, S, T, V, W或Y	
<220>		

<221> 变体
 <222> (92) .. (92)
 <223> I92X中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y
 <220>
 <221> 变体
 <222> (129) .. (129)
 <223> F129X中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.
 <220>
 <221> 变体
 <222> (132) .. (132)
 <223> M132X中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.
 <220>
 <221> 变体
 <222> (224) .. (224)
 <223> A224X中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.
 <220>
 <221> 变体
 <222> (432) .. (432)
 <223> I432X中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.
 <220>
 <221> 变体
 <222> (579) .. (579)
 <223> Q579X中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.
 <220>
 <221> 变体
 <222> (601) .. (601)
 <223> F601X中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.
 <400> 185
 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 Xaa Trp Ala Leu		
65	70	75
80		
Tyr Pro Gly Gly Pro Pro Asp Leu Asp Thr Thr Xaa 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
Xaa Thr Arg Xaa 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 Xaa		
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 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 Xaa		
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 Xaa Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp		
580	585	590
Glu Pro Tyr Tyr Thr Gly Thr Gly Xaa 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>	186	
<211>	725	

- <212> PRT
- <213> 人工序列
- <220>
- <223> 合成SHC衍生物
- <220>
- <221> 变体
- <222> (129) .. (129)
- <223> S129X中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.
- <220>
- <221> 变体
- <222> (145) .. (145)
- <223> V145X中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.
- <220>
- <221> 变体
- <222> (182) .. (182)
- <223> F182X中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.
- <220>
- <221> 变体
- <222> (185) .. (185)
- <223> Y185X中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.
- <220>
- <221> 变体
- <222> (282) .. (282)
- <223> G282X中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.
- <220>
- <221> 变体
- <222> (498) .. (498)
- <223> I498X中的X选自: A, B, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.
- <220>
- <221> 变体
- <222> (646) .. (646)
- <223> H646X中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.

<220>

<221> 变体

<222> (668) .. (668)

<223> F668X中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.

<400> 186

```

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

```

	260						265						270					
Ser	Gln	Phe	Phe	Arg	Thr	Thr	Asp	Arg	Xaa	Leu	His	Trp	Val	Gln	Ser			
	275						280						285					
Asn	Leu	Leu	Lys	Arg	Asn	Ser	Leu	Arg	Glu	Ala	Ala	Ile	Arg	His	Val			
	290						295						300					
Leu	Glu	Trp	Ile	Ile	Arg	His	Gln	Asp	Ala	Asp	Gly	Gly	Trp	Gly	Gly			
305					310					315					320			
Ile	Gln	Pro	Pro	Trp	Val	Tyr	Gly	Leu	Met	Ala	Leu	His	Gly	Glu	Gly			
				325					330					335				
Tyr	Gln	Leu	Tyr	His	Pro	Val	Met	Ala	Lys	Ala	Leu	Ser	Ala	Leu	Asp			
	340						345						350					
Asp	Pro	Gly	Trp	Arg	His	Asp	Arg	Gly	Glu	Ser	Ser	Trp	Ile	Gln	Ala			
	355						360						365					
Thr	Asn	Ser	Pro	Val	Trp	Asp	Thr	Met	Leu	Ala	Leu	Met	Ala	Leu	Lys			
	370						375						380					
Asp	Ala	Lys	Ala	Glu	Asp	Arg	Phe	Thr	Pro	Glu	Met	Asp	Lys	Ala	Ala			
385					390					395					400			
Asp	Trp	Leu	Leu	Ala	Arg	Gln	Val	Lys	Val	Lys	Gly	Asp	Trp	Ser	Ile			
				405					410					415				
Lys	Leu	Pro	Asp	Val	Glu	Pro	Gly	Gly	Trp	Ala	Phe	Glu	Tyr	Ala	Asn			
	420						425						430					
Asp	Arg	Tyr	Pro	Asp	Thr	Asp	Asp	Thr	Ala	Val	Ala	Leu	Ile	Ala	Leu			
	435						440						445					
Ser	Ser	Tyr	Arg	Asp	Lys	Glu	Glu	Trp	Gln	Lys	Lys	Gly	Val	Glu	Asp			
	450						455						460					
Ala	Ile	Thr	Arg	Gly	Val	Asn	Trp	Leu	Ile	Ala	Met	Gln	Ser	Glu	Cys			
465					470					475					480			
Gly	Gly	Trp	Gly	Ala	Phe	Asp	Lys	Asp	Asn	Asn	Arg	Ser	Ile	Leu	Ser			
				485					490					495				
Lys	Xaa	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
Ile Asp Ser Ile Gly Lys Gly Pro Thr Thr Pro Ser Gln Thr Ala Trp		
610	615	620
Ala Leu Met Gly Leu Ile Ala Ala Asn Arg Pro Glu Asp Tyr Glu Ala		
625	630	635
Ile Ala Lys Gly Cys Xaa Tyr Leu Ile Asp Arg Gln Glu Gln Asp Gly		
645	650	655
Ser Trp Lys Glu Glu Glu Phe Thr Gly Thr Gly Xaa Pro Gly Tyr Gly		
660	665	670
Val Gly Gln Thr Ile Lys Leu Asp Asp Pro Ala Leu Ser Lys Arg Leu		
675	680	685
Leu Gln Gly Ala Glu Leu Ser Arg Ala Phe Met Leu Arg Tyr Asp Phe		
690	695	700
Tyr Arg Gln Phe Phe Pro Ile Met Ala Leu Ser Arg Ala Glu Arg Leu		
705	710	715
Ile Asp Leu Asn Asn		
725		

<210> 187

<211> 658

<212> PRT

<213> 人工序列

<220>

<223> 合成SHC衍生物

<220>

<221> 变体

<222> (85) .. (85)

<223> G85X 中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.

<220>

<221> 变体

<222> (100) .. (100)

<223> V100X 中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.

<220>

<221> 变体

<222> (137) .. (137)

<223> F137X 中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.

<220>

<221> 变体

<222> (140) .. (140)

<223> I140X 中的X选自: A,, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.

<220>

<221> 变体

<222> (233) .. (233)

<223> V233X 中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.

<220>

<221> 变体

<222> (450) .. (450)

<223> I450X 中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.

<220>

<221> 变体

<222> (598) .. (598)

<223> N598X 中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y

<220>

<221> 变体

<222> (620) .. (620)

<223> F620X 中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.

<400> 187

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 Xaa Trp Ser Leu Phe Tyr Gly Gly Asp Leu Asp Leu

				85				90				95			
Ser	Ala	Thr	Xaa	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	Xaa	Thr	Arg	Xaa	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	Xaa	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 Xaa 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	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	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 Xaa Trp Leu Ala Gln Asn Gln Asp Glu Glu Gly		
595	600	605
Leu Trp Lys Glu Asp Tyr Tyr Ser Gly Gly Gly Xaa 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	640
Ala Arg Tyr Arg Asn Leu Lys Lys Ala Asn Gln Pro Ile Val His Tyr		
645	650	655

Gly Met

<210> 188

<211> 684

<212> PRT

<213> 人工序列

<220>

<223> 合成SHC衍生物

- <220>
<221> 变体
<222> (88) .. (88)
<223> A88X 中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.
<220>
<221> 变体
<222> (104) .. (104)
<223> V104X 中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.
<220>
<221> 变体
<222> (141) .. (141)
<223> F141X, 中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.
<220>
<221> 变体
<222> (144) .. (144)
<223> Y144X 中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.
<220>
<221> 变体
<222> (241) .. (241)
<223> V241X 中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.
<220>
<221> 变体
<222> (459) .. (459)
<223> I459X 中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.
<220>
<221> 变体
<222> (607) .. (607)
<223> M607X 中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W或Y.
<220>
<221> 变体
<222> (628) .. (628)
<223> F628X 中的X选自: A, C, D, E, F, G, H, I, K, L, M, N, P,

Q, R, S, T, V, W或Y.

<400> 188

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 Xaa Trp Gln Val Tyr Phe Gly Ala Pro
85 90 95

Asn Gly Asp Ile Asn Ala Thr Xaa 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 Xaa Thr Arg Xaa
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

Xaa 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		320
	325	330
Thr Ile Leu Thr Leu Leu Ala Phe Asp Asp Ala Gly Val Leu Gly Asp		335
	340	345
Tyr Pro Glu Ala Val Asp Lys Ala Val Asp Trp Val Leu Gln Arg Gln		350
	355	360
Val Arg Val Pro Gly Asp Trp Ser Met Lys Leu Pro His Val Lys Pro		365
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		400
	405	410
Lys Trp Lys Ala Lys Gly Ile Asp Glu Ala Ile Gln Leu Gly Val Asp		415
	420	425
Trp Leu Ile Gly Met Gln Ser Gln Gly Gly Gly Trp Gly Ala Phe Asp		430
435	440	445
Lys Asp Asn Asn Gln Lys Ile Leu Thr Lys Xaa 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
Gln Ala Leu Asp Tyr Ile Arg Arg Glu Gln Glu Pro Ser Gly Pro Trp		495
	500	505
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		560
	565	570
Thr Thr Thr Ala Ser Gln Thr Ala Trp Ala Leu Met Ala Leu Leu Ala		575
	580	585
Ala Asn Arg Pro Gln Asp Lys Asp Ala Ile Glu Arg Gly Cys Xaa Trp		590
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	Xaa	Pro	Gly	Tyr	Gly	Val	Gly	Gln	Thr	Ile	Lys	Leu	Asn
625							630						635		640
Asp	Pro	Ala	Leu	Ser	Gln	Arg	Leu	Met	Gln	Gly	Pro	Glu	Leu	Ser	Arg
							645						650		655
Ala	Phe	Met	Leu	Arg	Tyr	Gly	Met	Tyr	Arg	His	Tyr	Phe	Pro	Leu	Met
							660						665		670
Ala	Leu	Gly	Arg	Ala	Leu	Arg	Pro	Gln	Ser	His	Ser				
							675								680

SHCwt	MAEQVVEAPAYARTLDRAVEYLLSCQKDEGYWVGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP	86
SHC1	MAEQVVEAPAYARTLDRAVEYLLSCQKDEGYWVGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP	86
SHC2	MAEQVVEAPAYARTLDRAVEYLLSCQKDEGYWVGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP	86
SHC3	MAEQVVEAPAYARTLDRAVEYLLSCQKDEGYWVGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP	86
SHC4	MAEQVVEAPAYARTLDRAVEYLLSCQKDEGYWVGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP	86
SHC5	MAEQVVEAPAYARTLDRAVEYLLSCQKDEGYWVGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP	86
SHC6	MAEQVVEAPAYARTLDRAVEYLLSCQKDEGYWVGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP	86
SHC7	MAEQVVEAPAYARTLDRAVEYLLSCQKDEGYWVGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP	86
SHC8	MAEQVVEAPAYARTLDRAVEYLLSCQKDEGYWVGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP	86
SHC9	MAEQVVEAPAYARTLDRAVEYLLSCQKDEGYWVGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP	86
SHC10	MAEQVVEAPAYARTLDRAVEYLLSCQKDEGYWVGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP	86
SHCwt	DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMPVPEIMFLGKRMPPLNIYEFGSWARA	172
SHC1	DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMPVPEIMFLGKRMPPLNIYEFGSWARA	172
SHC2	DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMPVPEIMFLGKRMPPLNIYEFGSWARA	172
SHC3	DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMPVPEIMFLGKRMPPLNIYEFGSWARA	172
SHC4	DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMPVPEIMFLGKRMPPLNIYEFGSWARA	172
SHC5	DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMPVPEIMFLGKRMPPLNIYEFGSWARA	172
SHC6	DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMPVPEIMFLGKRMPPLNIYEFGSWARA	172
SHC7	DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMPVPEIMFLGKRMPPLNIYEFGSWARA	172
SHC8	DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMPVPEIMFLGKRMPPLNIYEFGSWARA	172
SHC9	DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMPVPEIMFLGKRMPPLNIYEFGSWARA	172
SHC10	DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMPVPEIMFLGKRMPPLNIYEFGSWARA	172

图1

SHCwt	TVVALSIVMSRQPVFPPLPERARVPPELYETDVPPRRRGAKGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW	258
SHC1	TVVALSIVMSRQPVFPPLPERARVPPELYETDVPPRRRGAKGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW	258
SHC2	TVVALSIVMSRQPVFPPLPERARVPPELYETDVPPRRRGAKGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW	258
SHC3	TVVALSIVMSRQPVFPPLPERARVPPELYETDVPPRRRGAKGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW	258
SHC4	TVVALSIVMSRQPVFPPLPERARVPPELYETDVPPRRRGAKGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW	258
SHC5	TVVALSIVMSRQPVFPPLPERARVPPELYETDVPPRRRGAKGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW	258
SHC6	TVVALSIVMSRQPVFPPLPERARVPPELYETDVPPRRRGAKGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW	258
SHC7	TVVALSIVMSRQPVFPPLPERARVPPELYETDVPPRRRGAKGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW	258
SHC8	TVVALSIVMSRQPVFPPLPERARVPPELYETDVPPRRRGAKGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW	258
SHC9	TVVALSIVMSRQPVFPPLPERARVPPELYETDVPPRRRGAKGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW	258
SHC10	TVVALSIVMSRQPVFPPLPERARVPPELYETDVPPRRRGAKGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW	258
SHCwt	GGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQ	344
SHC1	GGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQ	344
SHC2	GGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQ	344
SHC3	GGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQ	344
SHC4	GGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQ	344
SHC5	GGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQ	344
SHC6	GGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQ	344
SHC7	GGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQ	344
SHC8	GGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQ	344
SHC9	GGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQ	344
SHC10	GGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAGEWLLDRQ	344

图2

SHCwt	ITVPGDWAVKRPNLKPGGEAFQFDNVYYPDVDDTAVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVNTSDLPN	430
SHC1	ITVPGDWAVKRPNLKPGGEAFQFDNVYYPDVDDTAVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVNTSDLPN	430
SHC2	ITVPGDWAVKRPNLKPGGEAFQFDNVYYPDVDDTAVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVNTSDLPN	430
SHC3	ITVPGDWAVKRPNLKPGGEAFQFDNVYYPDVDDTAVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVNTSDLPN	430
SHC4	ITVPGDWAVKRPNLKPGGEAFQFDNVYYPDVDDTAVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVNTSDLPN	430
SHC5	ITVPGDWAVKRPNLKPGGEAFQFDNVYYPDVDDTAVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVNTSDLPN	430
SHC6	ITVPGDWAVKRPNLKPGGEAFQFDNVYYPDVDDTAVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVNTSDLPN	430
SHC7	ITVPGDWAVKRPNLKPGGEAFQFDNVYYPDVDDTAVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVNTSDLPN	430
SHC8	ITVPGDWAVKRPNLKPGGEAFQFDNVYYPDVDDTAVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVNTSDLPN	430
SHC9	ITVPGDWAVKRPNLKPGGEAFQFDNVYYPDVDDTAVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVNTSDLPN	430
SHC10	ITVPGDWAVKRPNLKPGGEAFQFDNVYYPDVDDTAVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVNTSDLPN	430
SHCwt	HIPFCDFGEVTDPPSEdVTAHVLECFGSFGYDDAWKVIIRRAVEYLKREQKPDGSGFGRWGVNLYGTGAVVSALKAVGIDTREPVI	516
SHC1	HIPFCDFGEVTDPPSEdVTAHVLECFGSFGYDDAWKVIIRRAVEYLKREQKPDGSGFGRWGVNLYGTGAVVSALKAVGIDTREPVI	516
SHC2	HIPFCDFGEVTDPPSEdVTAHVLECFGSFGYDDAWKVIIRRAVEYLKREQKPDGSGFGRWGVNLYGTGAVVSALKAVGIDTREPVI	516
SHC3	HIPFCDFGEVTDPPSEdVTAHVLECFGSFGYDDAWKVIIRRAVEYLKREQKPDGSGFGRWGVNLYGTGAVVSALKAVGIDTREPVI	516
SHC4	HIPFCDFGEVTDPPSEdVTAHVLECFGSFGYDDAWKVIIRRAVEYLKREQKPDGSGFGRWGVNLYGTGAVVSALKAVGIDTREPVI	516
SHC5	HIPFCDFGEVTDPPSEdVTAHVLECFGSFGYDDAWKVIIRRAVEYLKREQKPDGSGFGRWGVNLYGTGAVVSALKAVGIDTREPVI	516
SHC6	HIPFCDFGEVTDPPSEdVTAHVLECFGSFGYDDAWKVIIRRAVEYLKREQKPDGSGFGRWGVNLYGTGAVVSALKAVGIDTREPVI	516
SHC7	HIPFCDFGEVTDPPSEdVTAHVLECFGSFGYDDAWKVIIRRAVEYLKREQKPDGSGFGRWGVNLYGTGAVVSALKAVGIDTREPVI	516
SHC8	HIPFCDFGEVTDPPSEdVTAHVLECFGSFGYDDAWKVIIRRAVEYLKREQKPDGSGFGRWGVNLYGTGAVVSALKAVGIDTREPVI	516
SHC9	HIPFCDFGEVTDPPSEdVTAHVLECFGSFGYDDAWKVIIRRAVEYLKREQKPDGSGFGRWGVNLYGTGAVVSALKAVGIDTREPVI	516
SHC10	HIPFCDFGEVTDPPSEdVTAHVLECFGSFGYDDAWKVIIRRAVEYLKREQKPDGSGFGRWGVNLYGTGAVVSALKAVGIDTREPVI	516

图3

SHCwt	QKALDWVEQHONPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQLVETQRPDGGWDEPYTGTGFP	602
SHC1	QKALDWVEQHONPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQLVETQRPDGGWDEPYTGTGFP	602
SHC2	QKALDWVEQHONPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQLVETQRPDGGWDEPYTGTGFP	602
SHC3	QKALDWVEQHONPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQLVETQRPDGGWDEPYTGTGFP	602
SHC4	QKALDWVEQHONPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQLVETQRPDGGWDEPYTGTGFP	602
SHC5	QKALDWVEQHONPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQLVETQRPDGGWDEPYTGTGFP	602
SHC6	QKALDWVEQHONPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQLVETQRPDGGWDEPYTGTGFP	602
SHC7	QKALDWVEQHONPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQLVETQRPDGGWDEPYTGTGFP	602
SHC8	QKALDWVEQHONPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQLVETQRPDGGWDEPYTGTGFP	602
SHC9	QKALDWVEQHONPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQLVETQRPDGGWDEPYTGTGFP	602
SHC10	QKALDWVEQHONPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQLVETQRPDGGWDEPYTGTGFP	602
SHCwt	GDFYLGTYMYRHVFPTLALGRYKQAIERR	631
SHC1	GDFYLGTYMYRHVFPTLALGRYKQAIERR	631
SHC2	GDFYLGTYMYRHVFPTLALGRYKQAIERR	631
SHC3	GDFYLGTYMYRHVFPTLALGRYKQAIERR	631
SHC4	GDFYLGTYMYRHVFPTLALGRYKQAIERR	631
SHC5	GDFYLGTYMYRHVFPTLALGRYKQAIERR	631
SHC6	GDFYLGTYMYRHVFPTLALGRYKQAIERR	631
SHC7	GDFYLGTYMYRHVFPTLALGRYKQAIERR	631
SHC8	GDFYLGTYMYRHVFPTLALGRYKQAIERR	631
SHC9	GDFYLGTYMYRHVFPTLALGRYKQAIERR	631

图4

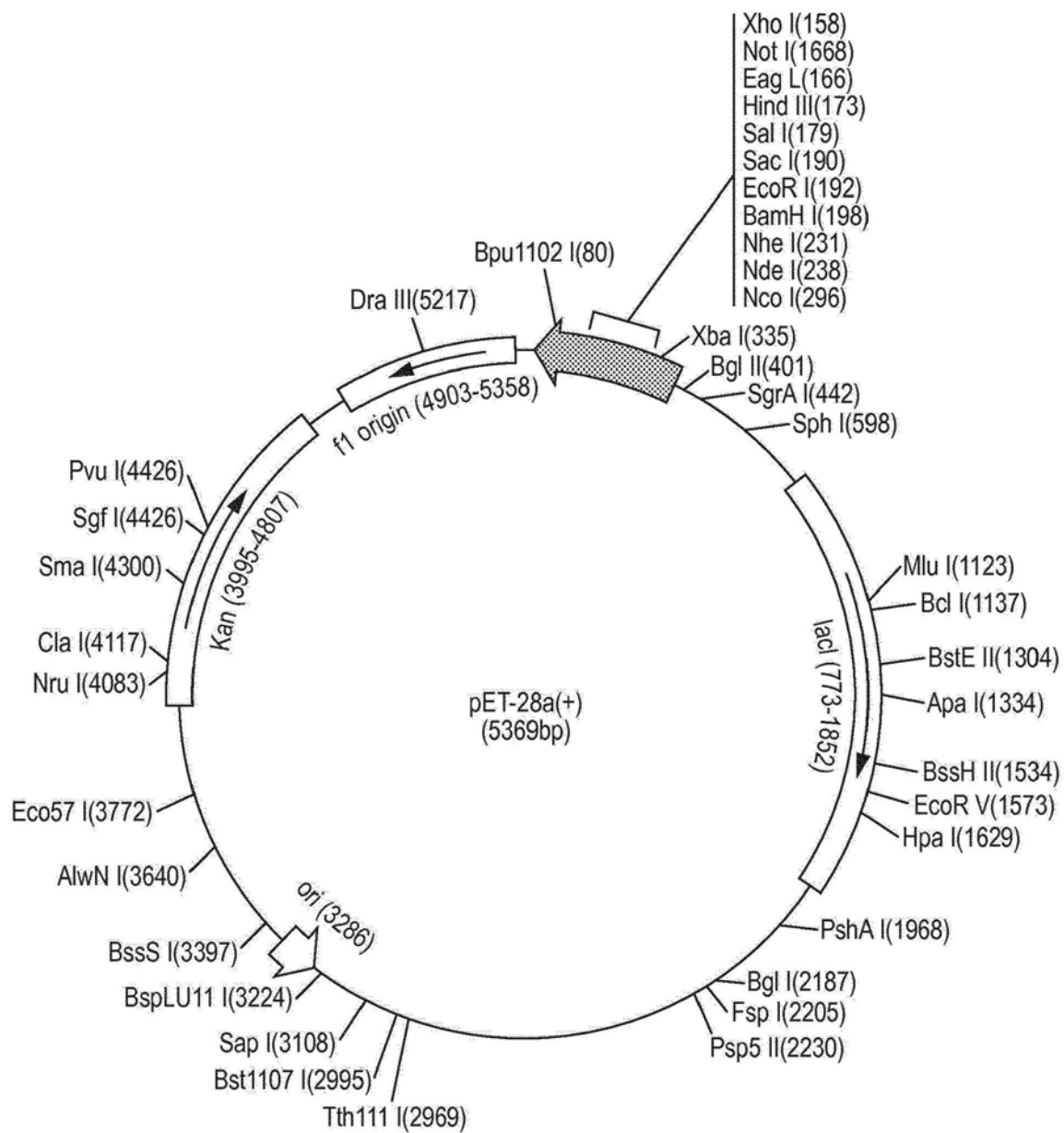


图5

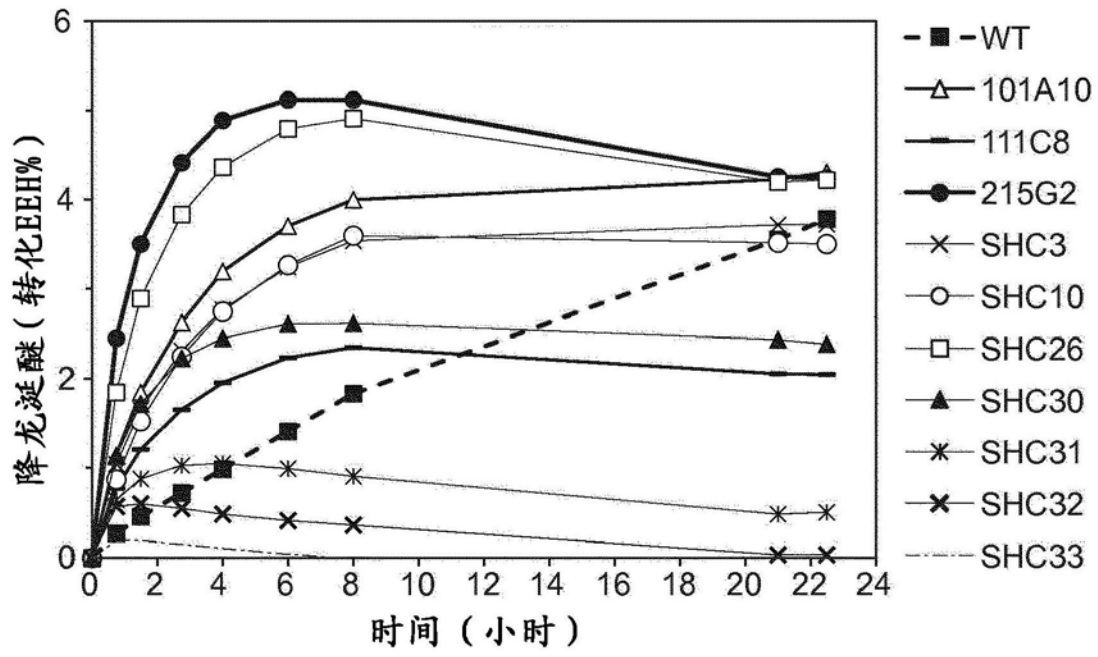


图6

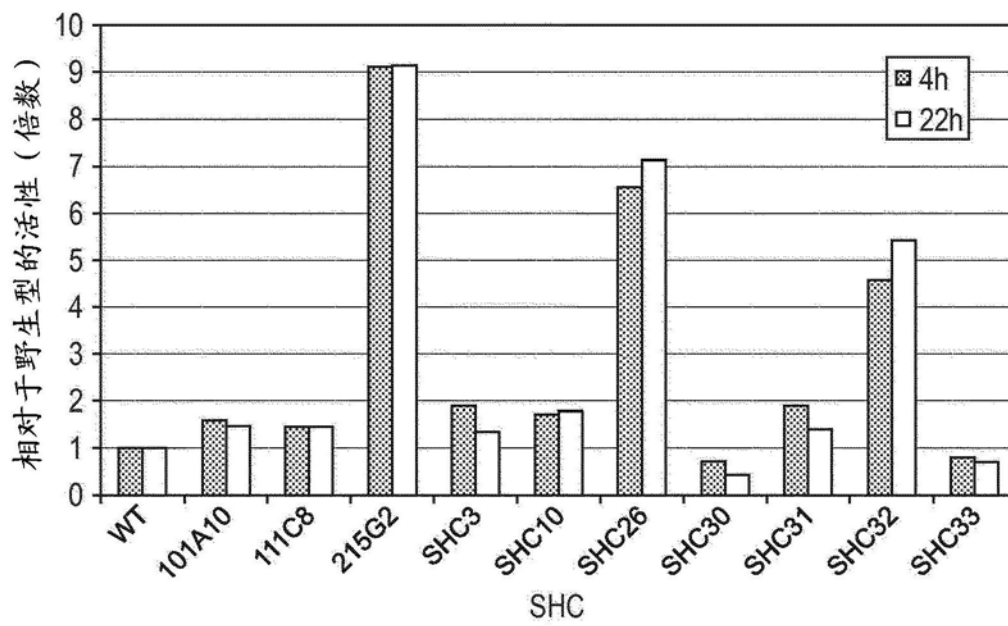


图7a

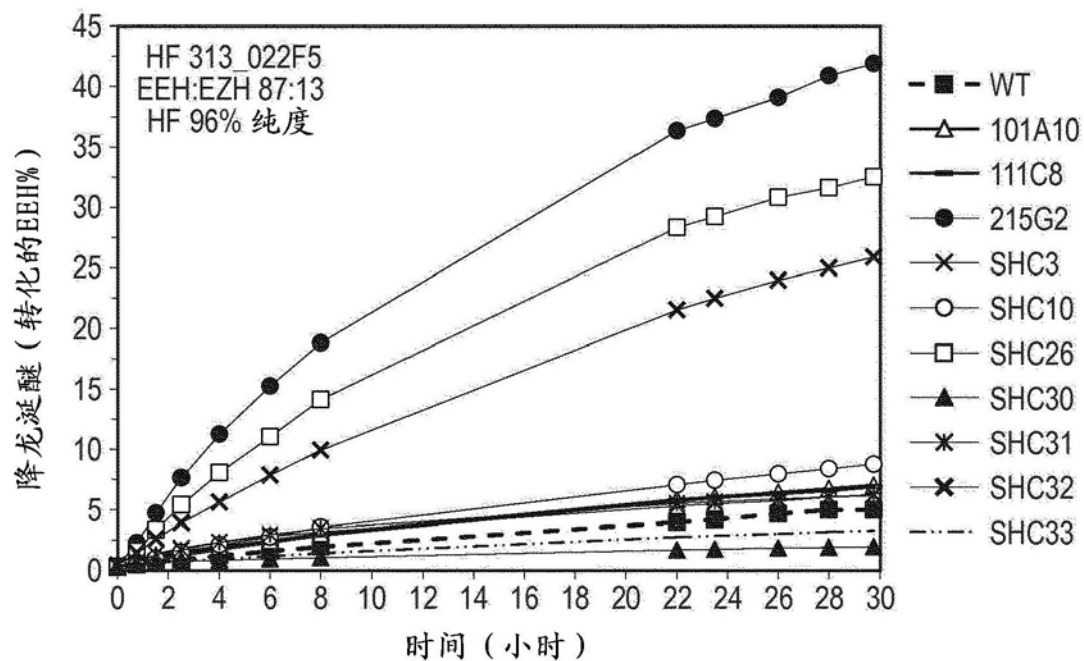


图7b

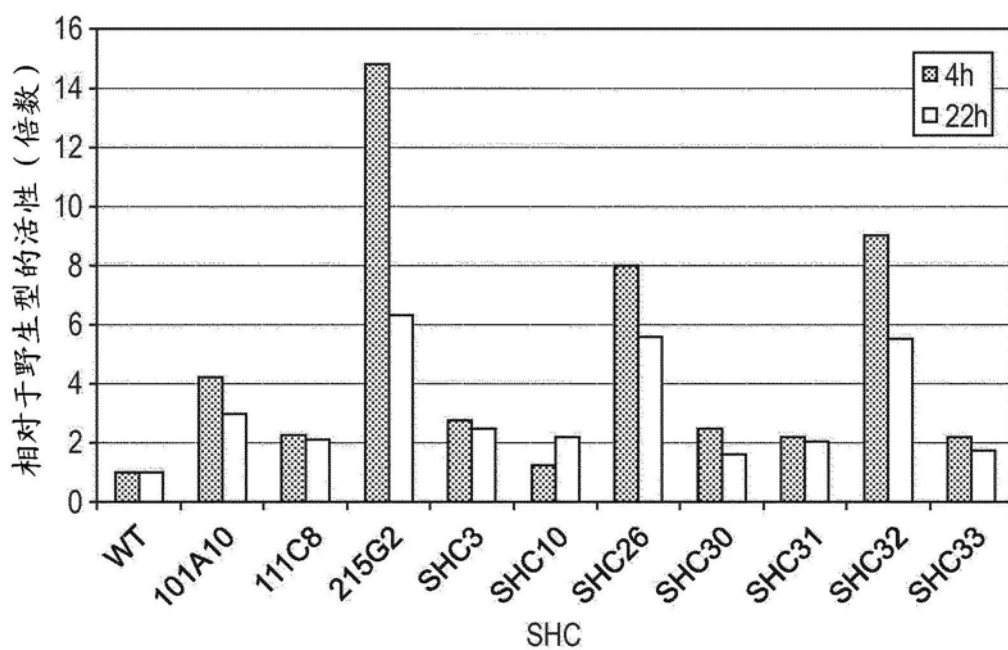


图8a

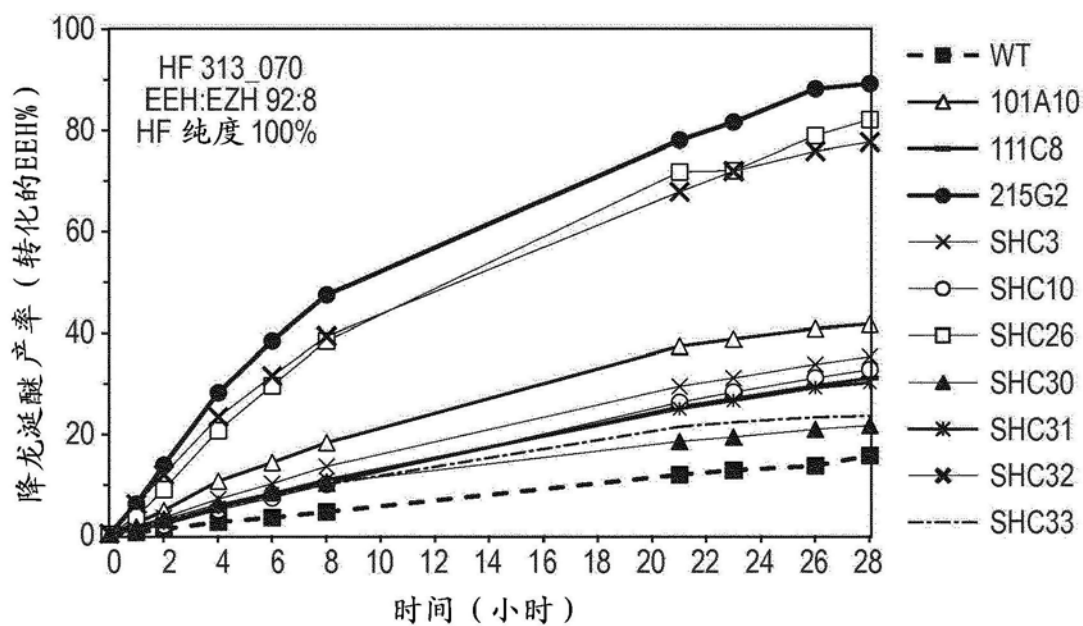


图8b

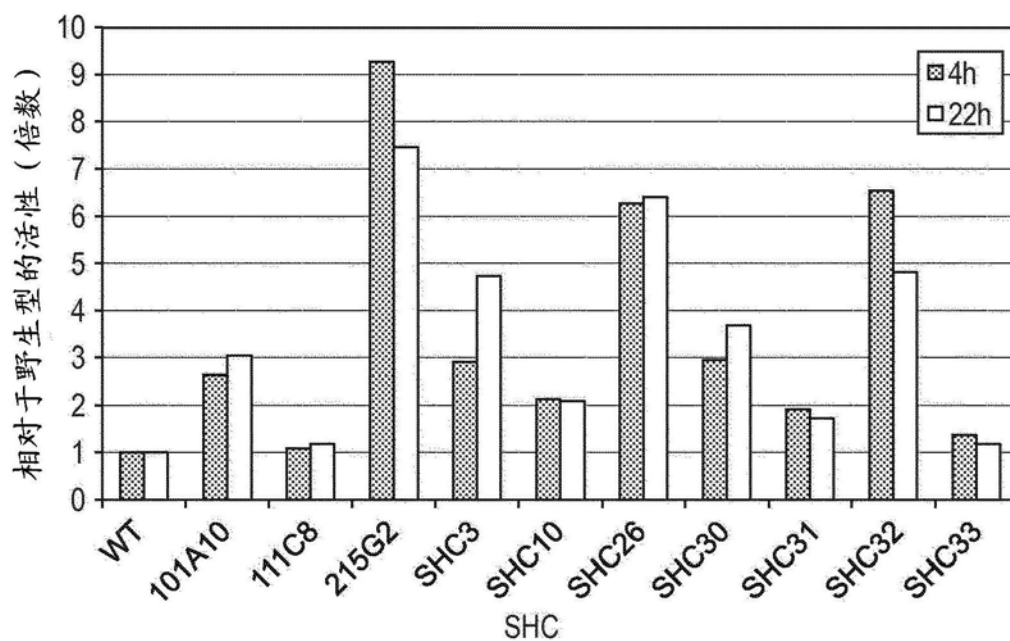


图9a

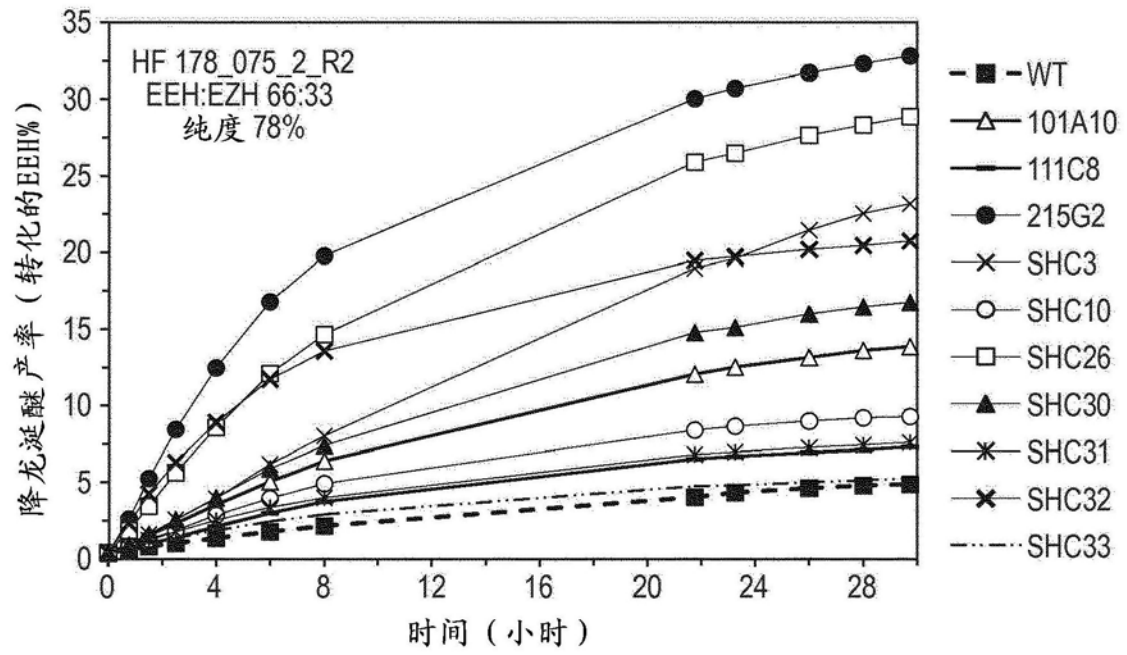


图9b

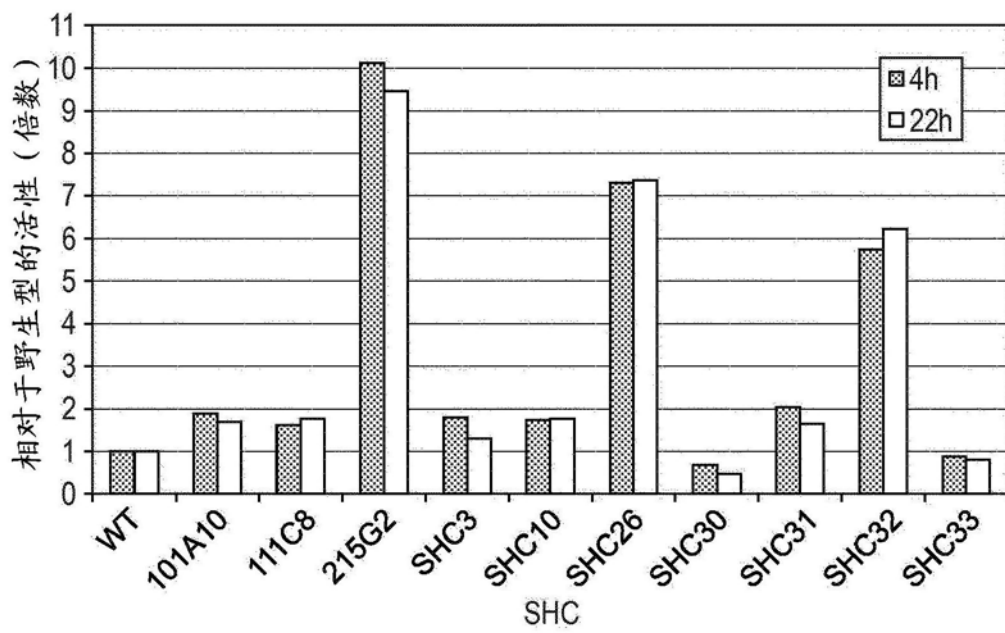


图10

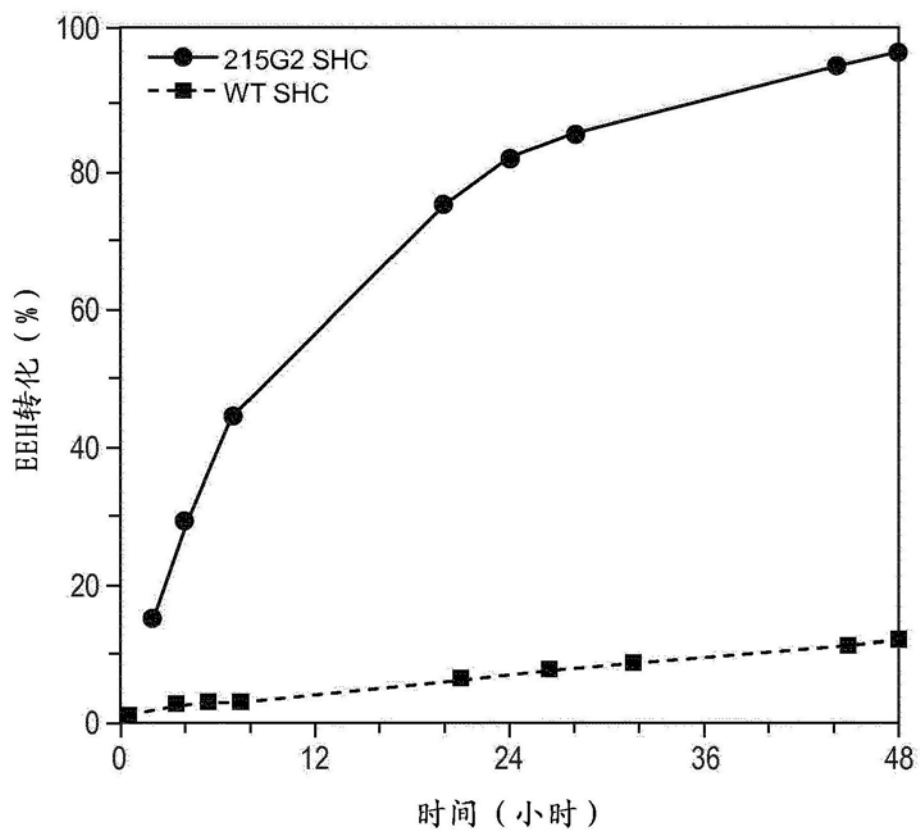


图11

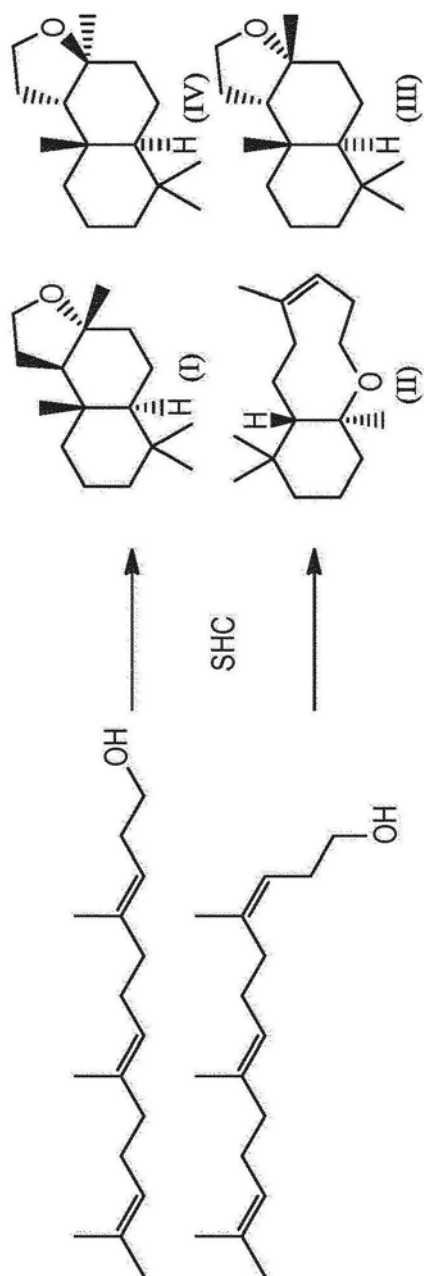


图12

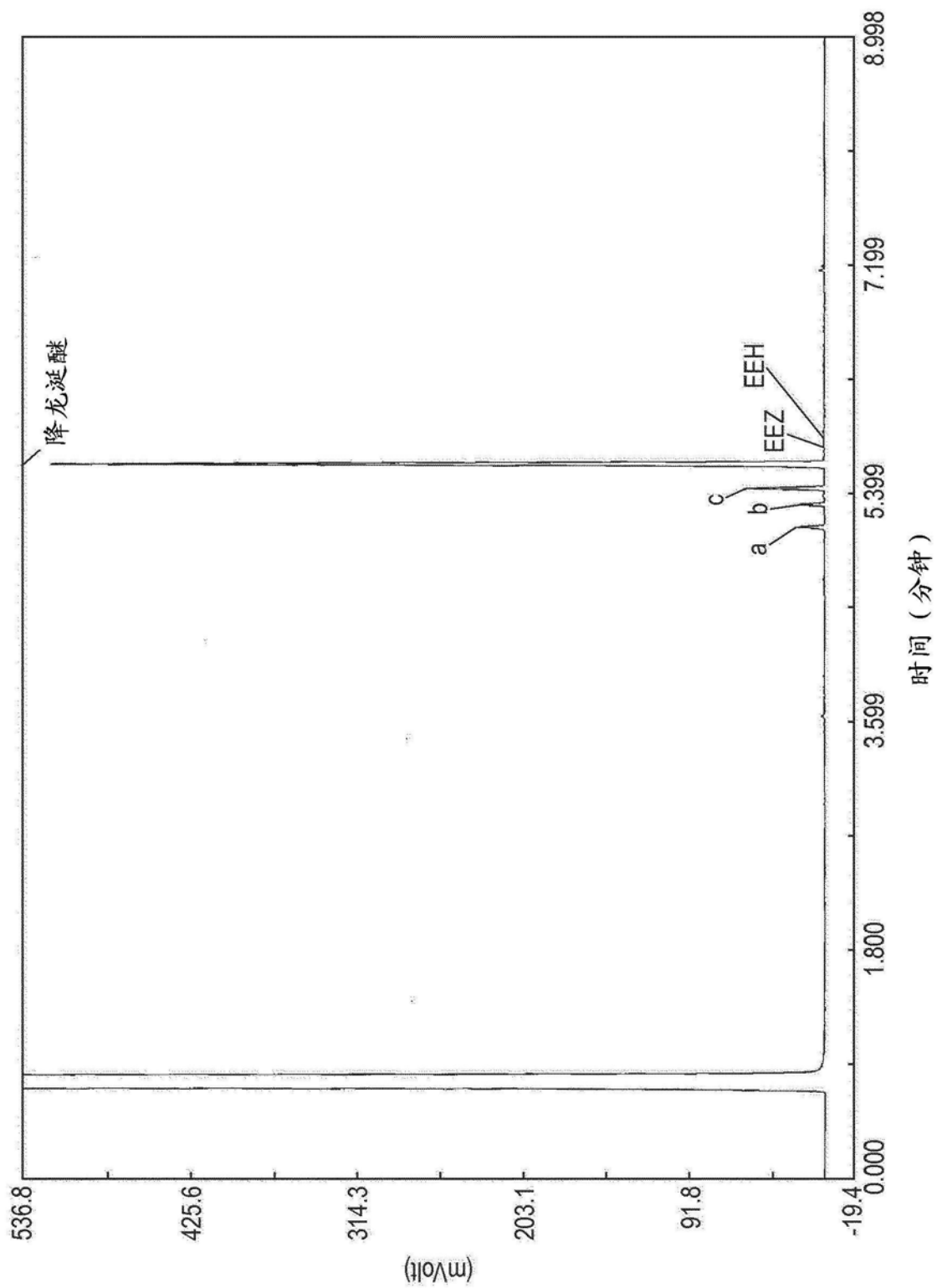


图13

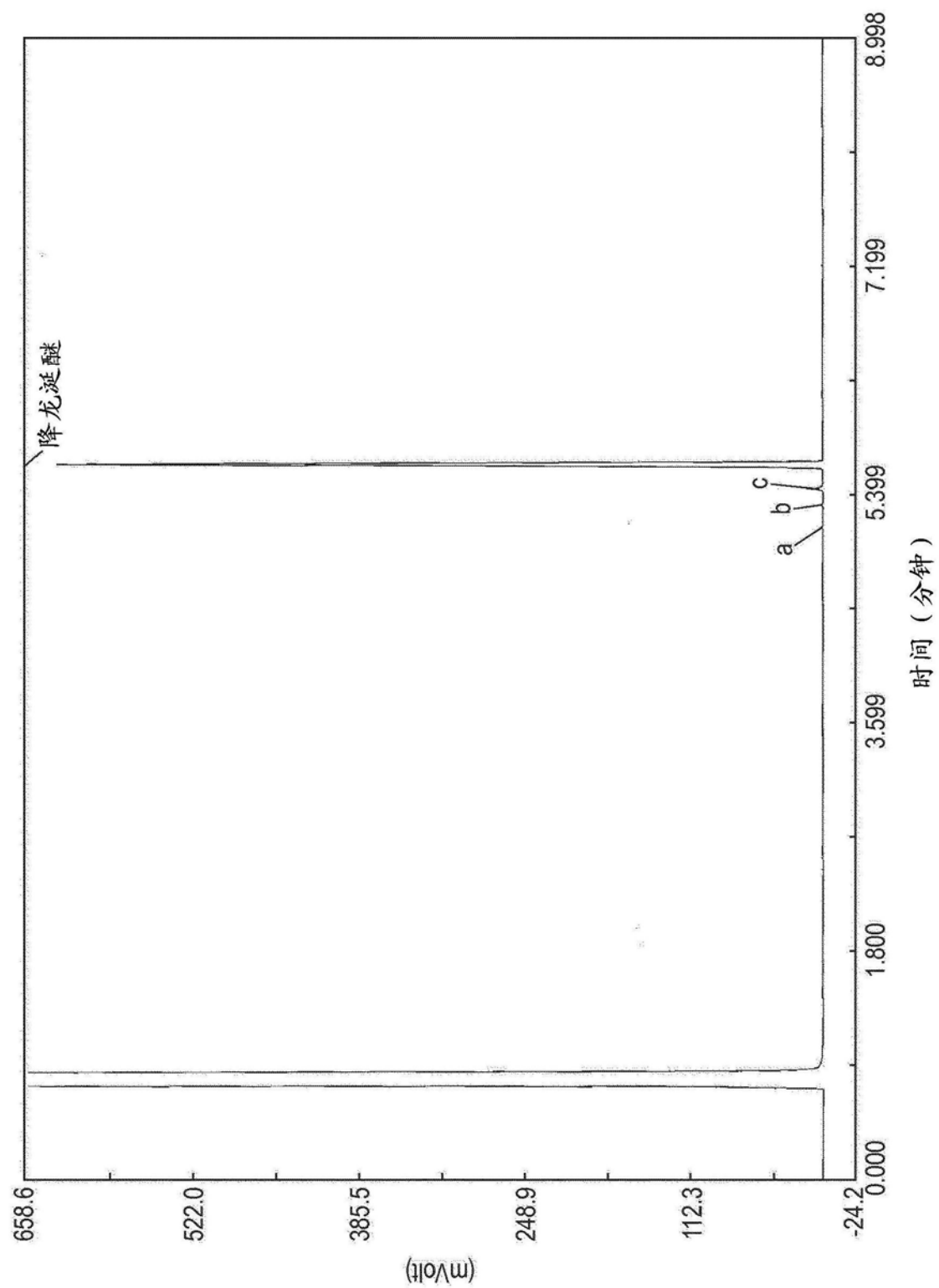


图14

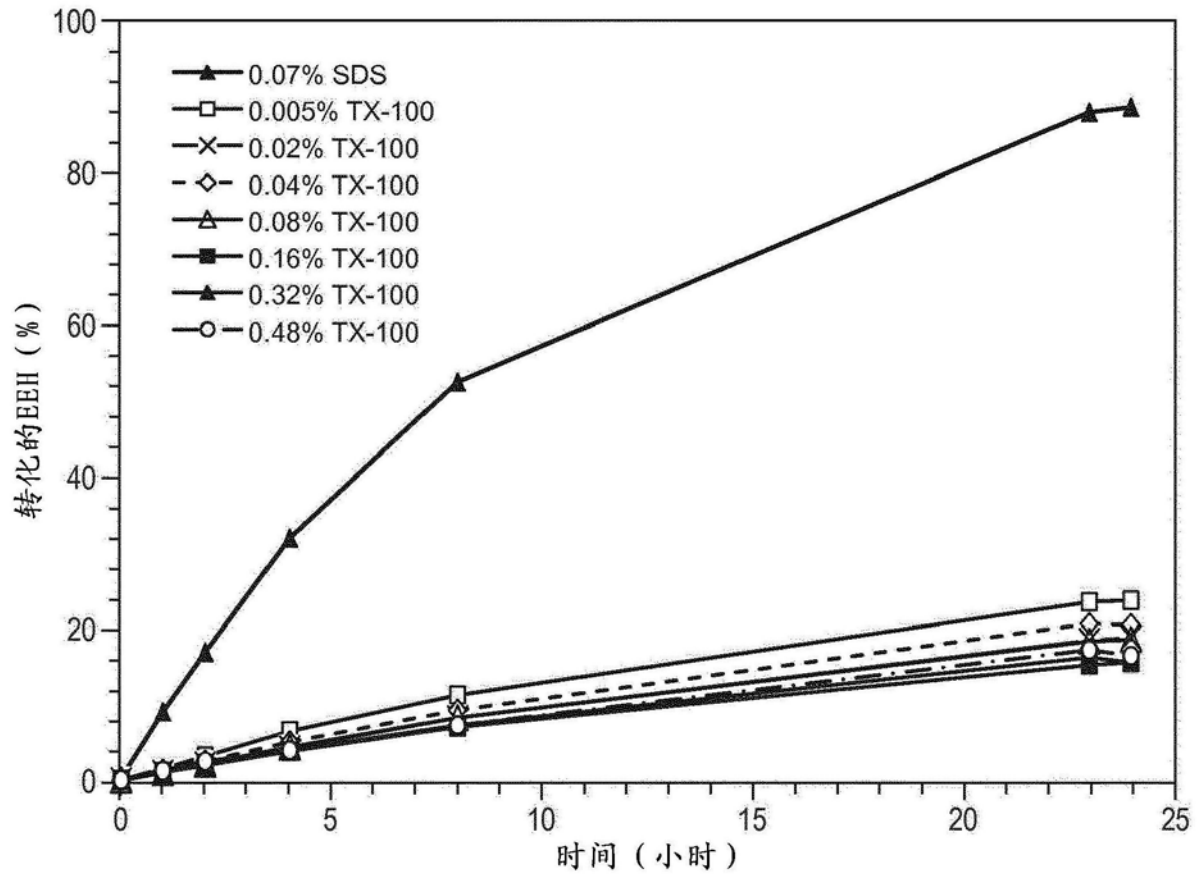


图15

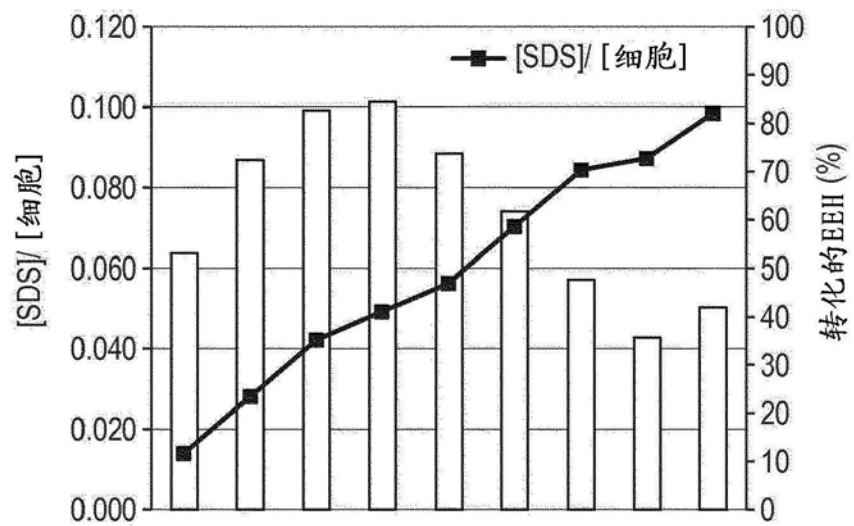


图16

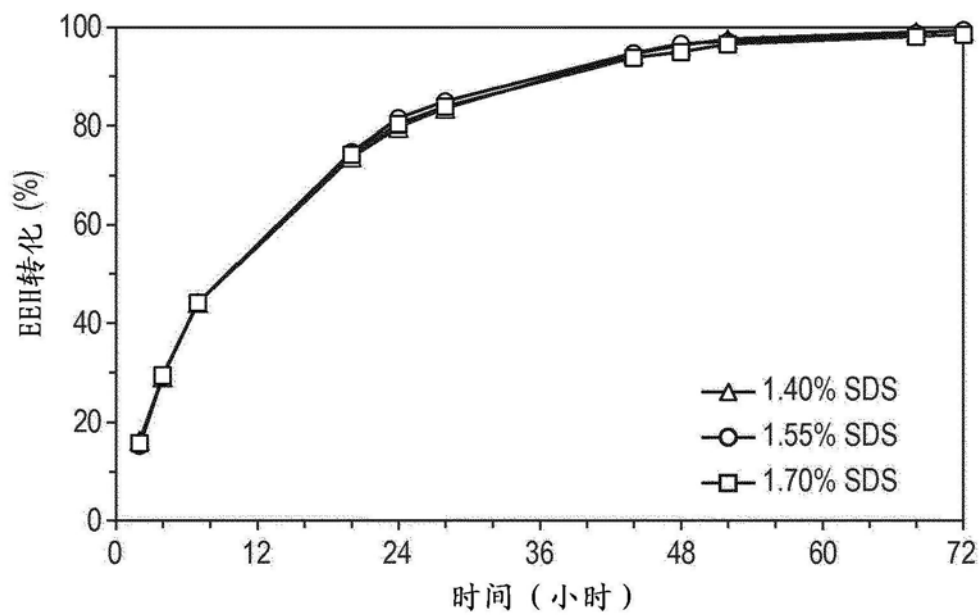


图17

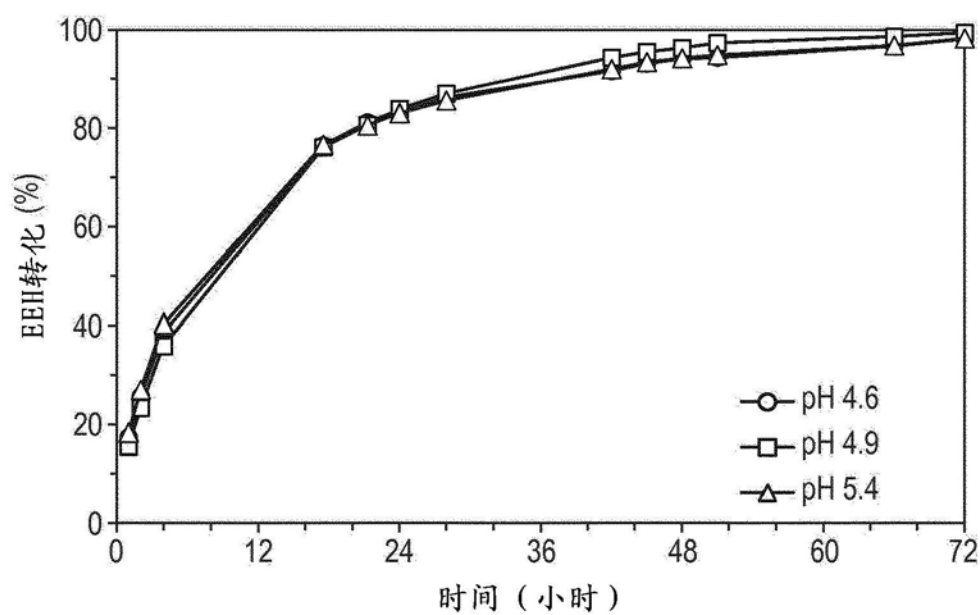


图18

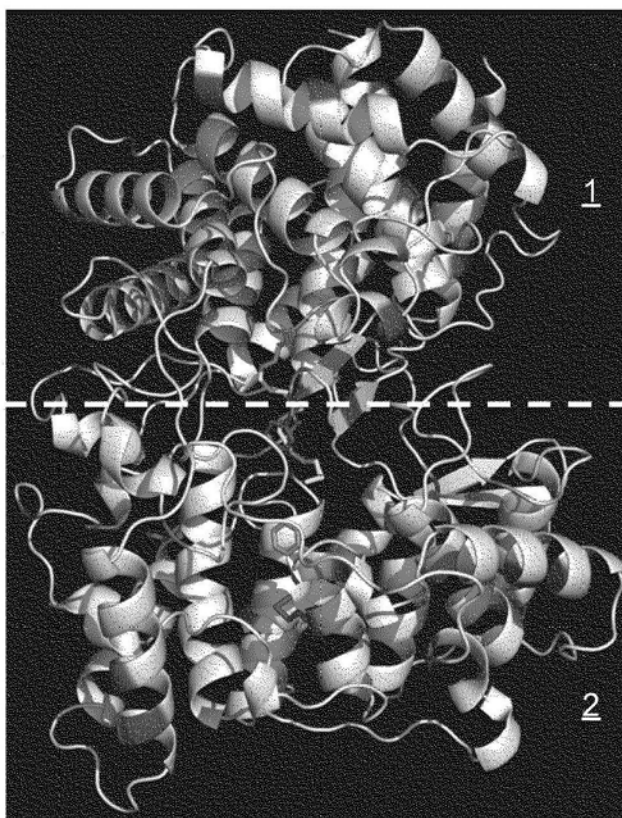


图19

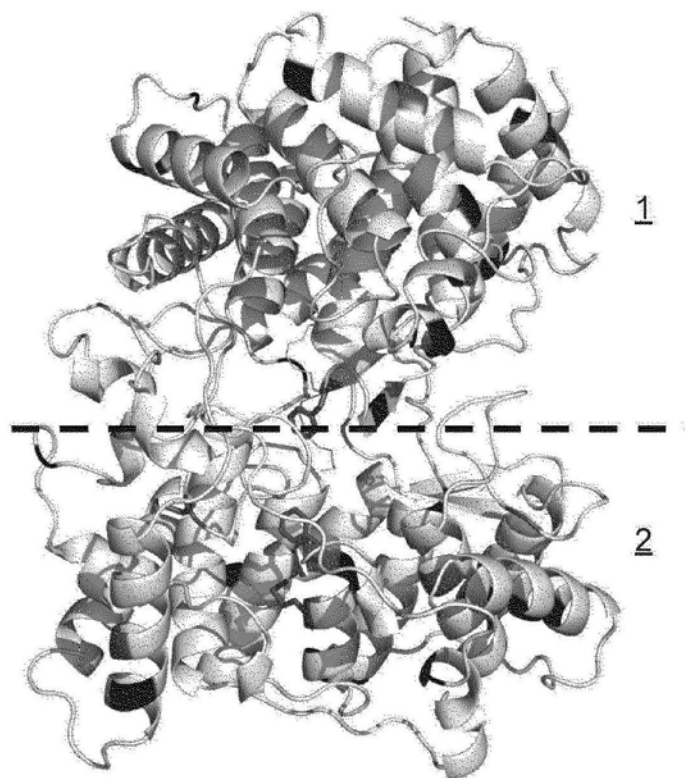


图20

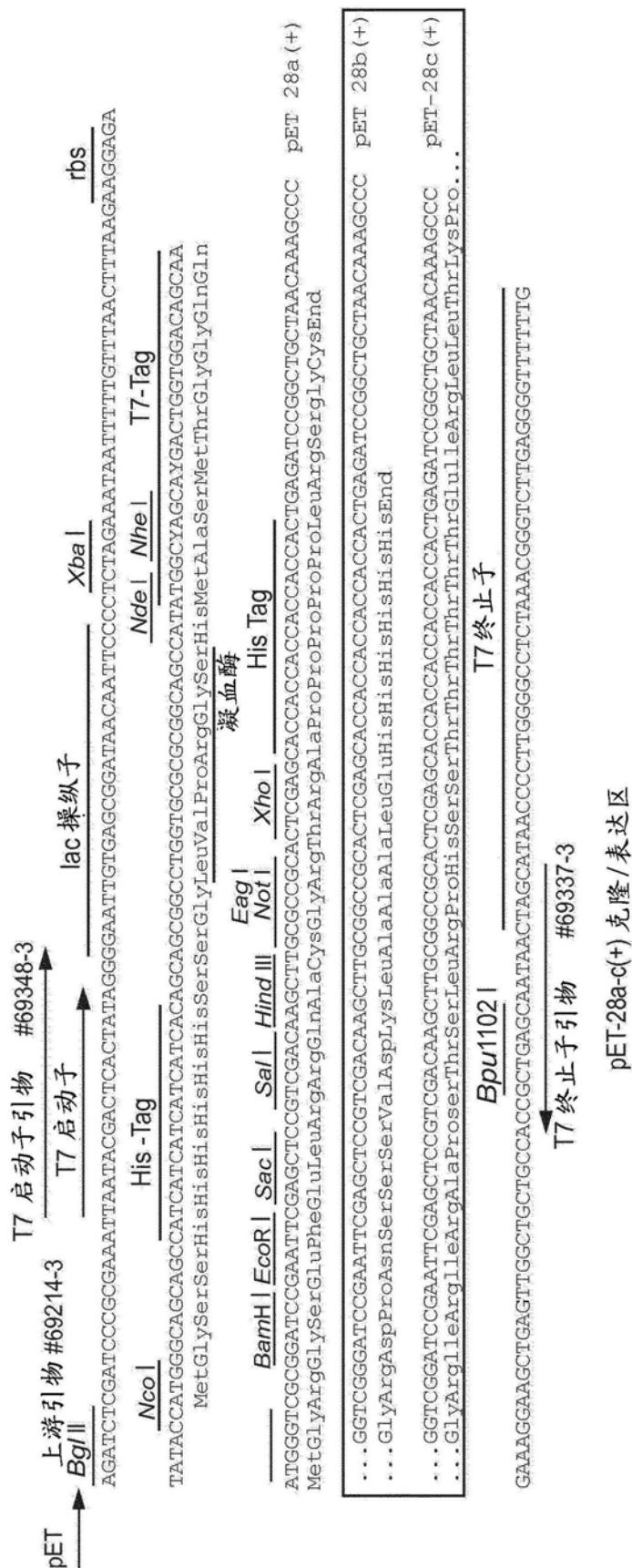


图21

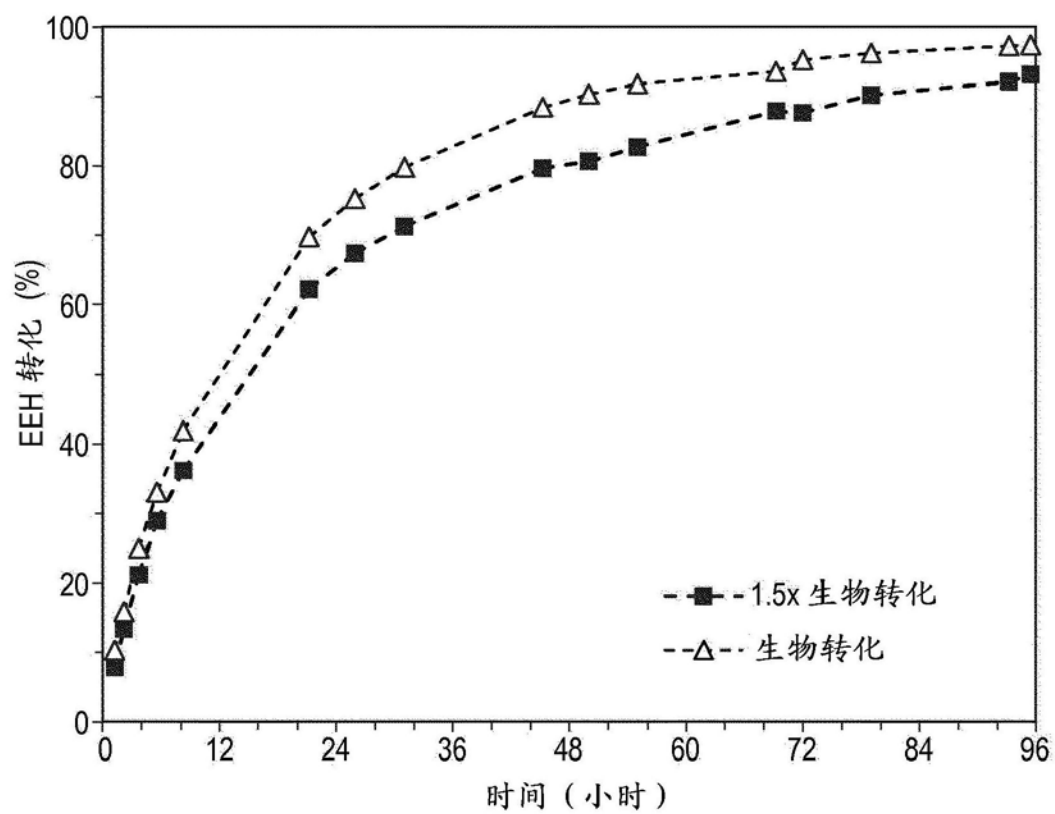


图22

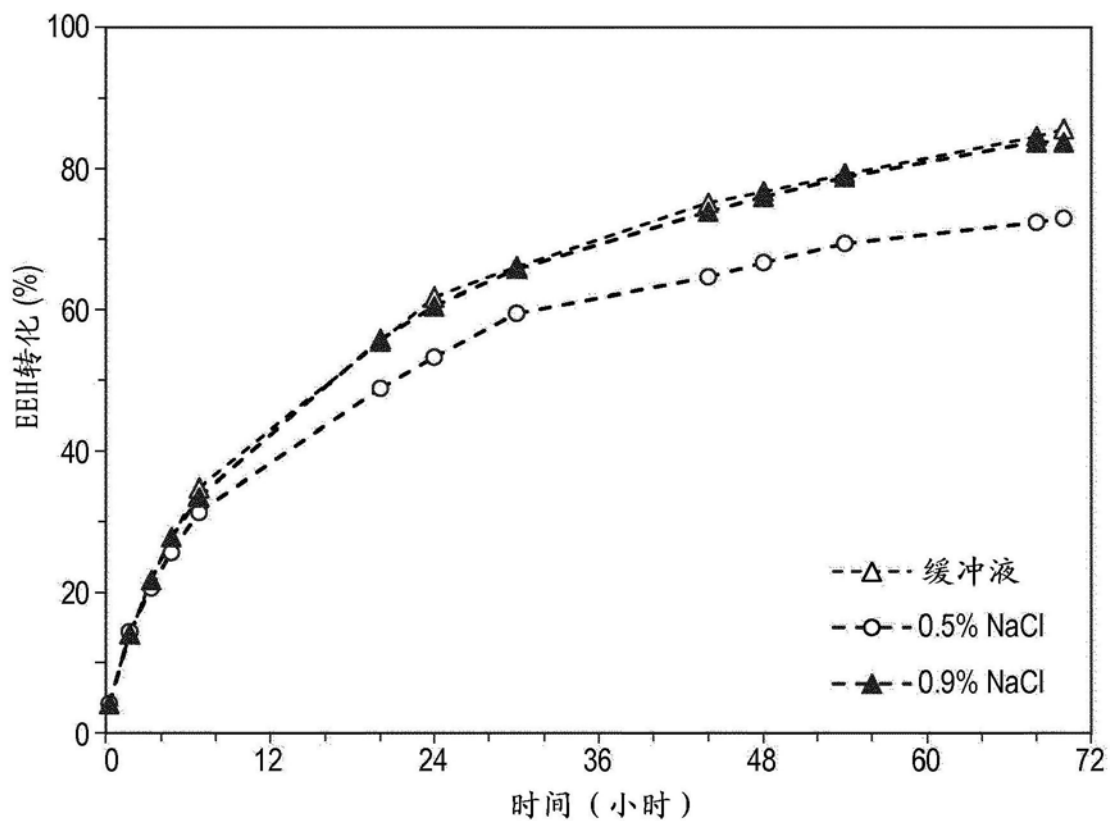


图23

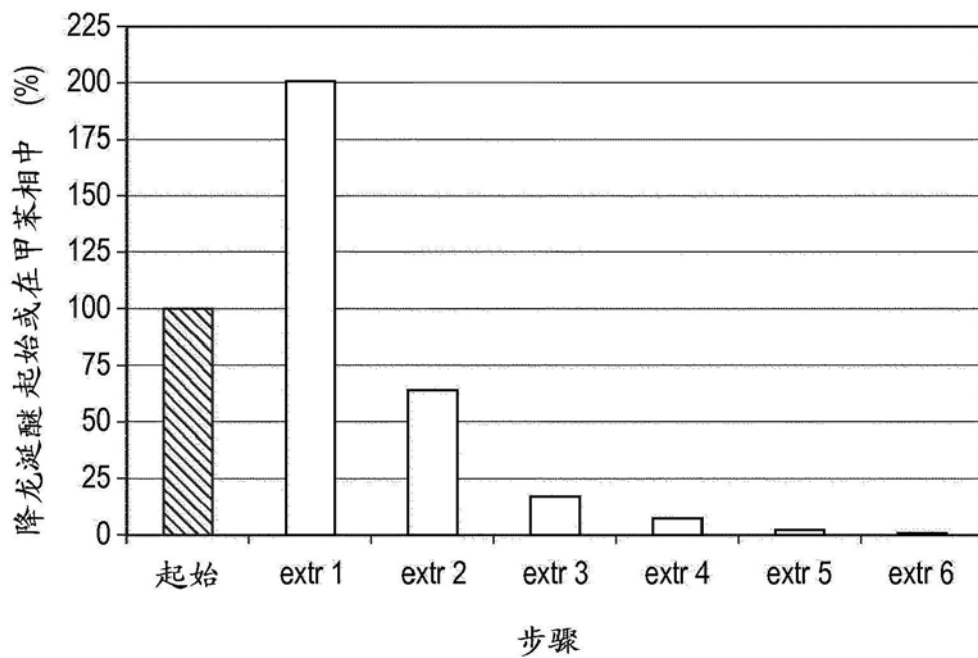


图24

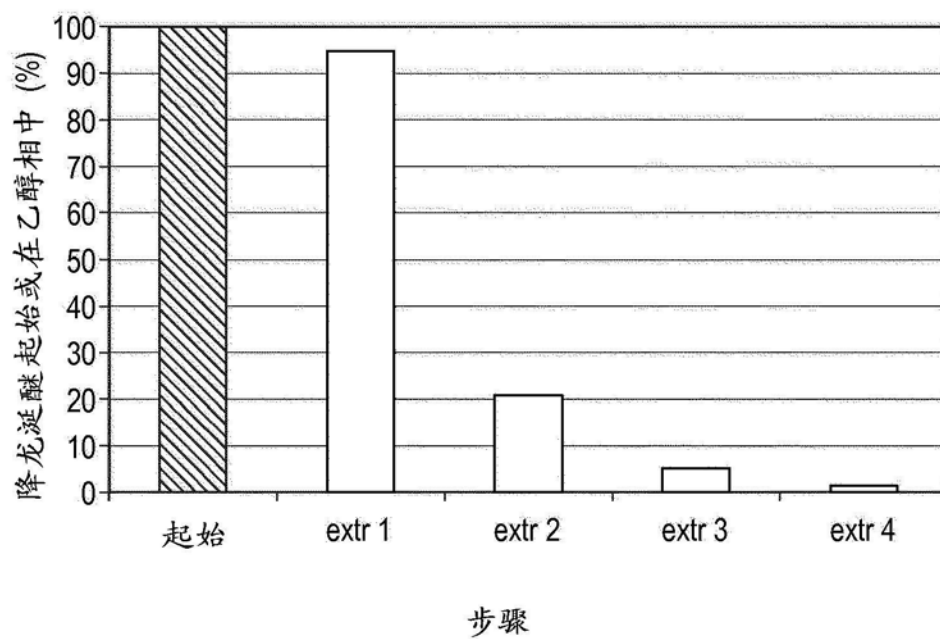


图25