РОССИЙСКАЯ ФЕДЕРАЦИЯ

ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(19) **RU** (11) **2 836 734**(13) **C1**

(51) МПК
*C12P 17/04* (2006.01)
*C12N 9/90* (2006.01)
*C07D 307/92* (2006.01)
*C11D 3/50* (2006.01)
*A61Q 13/00* (2006.01)
*A61K 8/49* (2006.01)

# (12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК
***C12P 17/04** (2025.01); **C12N 9/90** (2025.01); **C07D 307/92** (2025.01); **C11D 3/50** (2025.01); **C12Y 504/99017** (2025.01)*

(21)(22) Заявка: **2020123094, 22.04.2016**

(24) Дата начала отсчета срока действия патента:
**22.04.2016**

Дата регистрации:
**20.03.2025**

Приоритет(ы):
(30) Конвенционный приоритет:
**24.04.2015 GB 1507207.7**
Номер и дата приоритета первоначальной заявки, из которой данная заявка выделена:
**2017134449 24.04.2015**

(45) Опубликовано: **20.03.2025** Бюл. № **8**

Адрес для переписки:
**191036, г. Санкт-Петербург а/я 24 "НЕВИНПАТ", Поликарпов Александр Викторович**

(72) Автор(ы):
**ЭЙХГОРН Эрик (CH),**
**ШИЛЛИНГ Борис (CH),**
**ВАЛЕР Денис (FR),**
**ФУРАЖ Лорэн (FR),**
**ЛОШЭР Эстер (CH)**

(73) Патентообладатель(и):
**ЖИВОДАН СА (CH)**

(56) Список документов, цитированных в отчете о поиске: **US 2012135477 A1, 31.05.2012. JP 2009060799 A, 26.03.2009. US 5616737 A1, 01.04.1997. US 2013273619 A1, 17.10.2013. EP 0165458 A2, 27.12.1985. SU 910561 A1, 07.03.1982. SU 1498767 A1, 07.08.1989.**

## (54) ФЕРМЕНТЫ И ИХ ПРИМЕНЕНИЕ

(57) Реферат:

Группа изобретений относится к биотехнологии, в частности к селективному получению (-)-амброксида. Предложен способ получения (-)-амброксида или смеси, содержащей (-)-амброксид, где смесь изомеров, содержащую (3E,7E)-гомофарнезол (EEH), ферментативно конвертируют в (-)-амброксид или в смесь, содержащую (-)-амброксид, с использованием фермента сквален-гопен-циклазы/гомофарнезол-амброксид-циклазы (SHC/HAC), содержащего аминокислотную последовательность, имеющую по меньшей мере 30% идентичность с SEQ ID NO: 1, в подходящих реакционных условиях. Смесь изомеров, содержащая EEH, выбрана из по меньшей мере одной из [EE:EZ], [EE:ZE] и [EE:EZ:ZE]. (-)-Амброксид присутствует в твердой фазе в виде кристаллов или в аморфной форме. Способ приводит к получению продукта, содержащего 80-97% (-)-амброксида в твердой форме, который применяют в составе ароматизирующего продукта, или косметического продукта, или потребительского продукта. 5 н. и 19 з.п. ф-лы, 25 ил., 26 табл., 23 пр.

RU 2836734 C1

RU 2836734 C1

RUSSIAN FEDERATION

FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(19) **RU** (11) **2 836 734**(13) **C1**

(51) Int. Cl.
*C12P 17/04* (2006.01)
*C12N 9/90* (2006.01)
*C07D 307/92* (2006.01)
*C11D 3/50* (2006.01)
*A61Q 13/00* (2006.01)
*A61K 8/49* (2006.01)

## (12) ABSTRACT OF INVENTION

(52) CPC
***C12P 17/04** (2025.01); **C12N 9/90** (2025.01); **C07D 307/92** (2025.01); **C11D 3/50** (2025.01); **C12Y 504/99017** (2025.01)*

(21)(22) Application: **2020123094, 22.04.2016**

(24) Effective date for property rights:
**22.04.2016**

Registration date:
**20.03.2025**

Priority:
(30) Convention priority:
**24.04.2015 GB 1507207.7**
Number and date of priority of the initial application, from which the given application is allocated:
**2017134449 24.04.2015**

(45) Date of publication: **20.03.2025** Bull. № **8**

Mail address:
**191036, g. Sankt-Peterburg a/ya 24 "NEVINPAT", Polikarpov Aleksandr Viktorovich**

(72) Inventor(s):
**EICHHORN Eric (CH),**
**SCHILLING Boris (CH),**
**WAHLER Denis (FR),**
**FOURAGE Laurent (FR),**
**LOCHER Esther (CH)**

(73) Proprietor(s):
**GIVAUDAN SA (CH)**

RU 2836734 C1

## (54) ENZYMES AND USE THEREOF

(57) Abstract:

FIELD: biotechnology.

SUBSTANCE: group of inventions relates to biotechnology, in particular to selective production of (-)-ambroxide. Disclosed is a method of producing (-)-ambroxide or a mixture containing (-)-ambroxide, where the mixture of isomers containing (3E,7E)-homoparnesol (EEN) is enzymatically converted into (-)-ambroxide or into a mixture containing (-)-ambroxide, using the enzyme squalene-hopene-cyclase/homopharanesol-ambroxide-cyclase (SHC/HAC), containing an amino acid sequence having at least 30% identity with SEQ ID NO: 1, in suitable reaction conditions. Mixture of isomers containing EEH is selected from at least one of [EE:EZ], [EE:ZE] and [EE:EZ:ZE]. (-)-Ambroxide is present in solid phase in form of crystals or in amorphous form.

EFFECT: method leads to obtaining product containing 80–97% (-)-ambroxide in solid form, which is used in a flavoring product, or cosmetic product, or consumer product.

24 cl, 25 dwg, 26 tbl, 23 ex

RU 2836734 C1

## ОБЛАСТЬ ИЗОБРЕТЕНИЯ

Настоящее изобретение относится к производным фермента сквален-гопен-циклазы (Squalene Hopene Cyclase)/гомофарнезол-амброксид-циклазы (Homofarnesol Ambrox Cyclase) (SHC/HAC), модифицированным по отношению к исходному белку SHC/HAC, к аминокислотным последовательностям, составляющим производные ферментов SHC/HAC, нуклеотидным последовательностям, кодирующим производные SHC/HAC, векторам, содержащим нуклеотидные последовательности, кодирующие производные SHC/HAC, и к рекомбинантным клеткам-хозяевам, содержащим нуклеотидные последовательности, кодирующие производные SHC/HAC. Настоящее изобретение также относится к средствам для функциональной экспрессии нуклеотидных последовательностей, кодирующих производные SHC/HAC, и к способам применения рекомбинантных микроорганизмов, содержащих нуклеотидные последовательности, кодирующие производные SHC/HAC и SHC/HAC дикого типа (WT), для получения амброксида, предпочтительно (-)-амброксида.

## ПРЕДШЕСТВУЮЩИЙ УРОВЕНЬ ТЕХНИКИ

Сквален-гопен-циклазы (SHC, EC 5.4.99.17) представляют собой мембраносвязанные прокариотические ферменты, действующие в качестве биокатализаторов реакции циклизации линейного тритерпеноида, представляющего собой сквален, с образованием гопена и гопанола. Ранние работы по SHC были сосредоточены на характеризации SHC термофильной и ацидофильной бактерии Alicyclobacillus acidocaldarius (ранее Bacillus acidocaldarius) (см. Neumann & Simon 1986, Biol Chem Hoppe-Seyler 367, 723-729; Seckler & Poralla 1986, Biochem Biophys Act 356-363 и Ochs et al 1990, J Bacteriol 174, 298-302). Однако позднее другие SHC были очищены из Zymomonas mobilis и Bradyrhizobium japonicum и охарактеризованы в отношении их природных (например, сквалена) и неприродных субстратов (например, гомофарнезола и цитраля) (см., например, WO 2010/139710, WO 2012/066059 и Seitz et al 2012, J. Molecular Catalysis B: Enzymatic 84, 72-77).

В более ранней работе авторов Neumann и Simon (1986 цитируется выше) раскрыто, что гомофарнезол является дополнительным субстратом для SHC Alicyclobacillus acidocaldarius (AacSHC). Однако, согласно публикациям, доля циклизации неприродного гомофарнезола очищенным AacSHC, как сообщили Neumann и Simon (1986), составляла лишь 3% от доли циклизации природного субстрата сквалена. Доля образования амброксида (продукт 2b) увеличивалась с повышением концентрации гомофарнезола (продукт 1b) с 0,25 мМ до 2,0 мМ и несколько уменьшалась в присутствии 4 мМ продукта 1b. Это различие в долях циклизации может отчасти объясняться тем, что природный субстрат SHC сквален в два раза превышает по размеру (C30-углеродное соединение) неприродный гомофарнезол, который представляет собой C16-углеродное соединение.

В (JP 2009060799 Kao) также раскрыт способ получения амброксида из гомофарнезола с использованием SHC из A. acidocadarius. В то время как в JP 2009060799 сообщается о возможности использования микроорганизмов, содержащих SHC, для синтеза амброксида, в JP 2009060799 раскрыто получение амброксида из гомофарнезола только с использованием жидкого экстракта SHC. приготовленного из рекомбинантного микроорганизма, экспрессирующего ген SHC, но не с помощью целых клеток рекомбинантных микроорганизмов, экспрессирующих ген SHC. Процентная доля конверсии гомофарнезола в амброксид с использованием жидкого экстракта SHC была указана как 17,5% при проведении реакции при температуре 60°C в течение 14 часов при pH 5,2-6,0, но лишь 6,8% при проведении реакции при pH 6,6. Сообщалось, что процентная доля конверсии 3E,7E-гомофарнезола в амброксид с использованием

жидкого экстракта SHC при 60°C при pH 5,6 в течение 64 часов составляет 63% при использовании концентрации субстрата гомофарнезола 0,2% (2 г/л).

В WO 2010/139719 A2 и ее эквивалентной заявке на патент США (US 2012/0135477 A1) описаны по меньшей мере три экстракта фермента SHC, обладающих активностью циклазы гомофарнезола до амброксида. Сообщается, что доли конверсии гомофарнезола ферментами SHC Zymomonas mobilis (Zmo) и SHC Bradyrhizobium japonicum (Bjp) составляли 41% при 16 ч реакции и 22% при использовании концентрации гомофарнезола 10 мМ (2,36 г/л), соответственно, при этом доля конверсии ферментом AacSHC согласно сообщениям составляла лишь 1,2% (предположительно при той же концентрации гомофарнезола), однако подробное описание эксперимента не приведено. Экстракты ферментов ZmoSHC и BjpSHC были приготовлены из рекомбинантного микроорганизма, экспрессирующего ген SHC, путем разрушения клеток-хозяев E. coli, продуцирующих ферменты SHC, и отделения растворимых фракций SHC.

В публикации Seitz и соавт. (2012 цитируется выше) описана функциональная экспрессия и биохимическая характеризация трех ферментов SHC: двух из Z. mobilis (ZmoSHC1 и ZmoSHC2) и одного из A. acidocaldarius. Согласно публикации, «эффективную» конверсию (22,95%) гомофарнезола в амброксид наблюдали при использовании ZmoSHC1 дикого типа, при этом при использовании WT ZmoSHC2 конверсия гомофарнезола в амброксид не происходила, а при использовании AacSHC была обнаружена относительно низкая конверсия (3,4%) гомофарнезола в амброксид при использовании концентрации гомофарнезола 10 мМ (2,36 г/л). Склонность к относительно низкой конверсии гомофарнезола в амброксид, наблюдаемая для AacSHC, согласовывалась с результатами, полученными авторами Neumann и Simon (1986 цитируется выше) и обсуждаемыми в WO 2010/139719 A2, что также обсуждается выше. Три фермента SHC использовали в формате клеточной суспензии (посредством частичного разрушения клеток-хозяев E. coli с использованием циклов замораживания и оттаивания) и в виде частично очищенных мембраносвязанных фракций.

В WO 2012/066059 раскрыты мутанты с циклазной активностью и их применение в способе биокаталитической конверсии терпенов, например, в частности, для получения изопулегола путем реакции циклизации цитронеллаля; в способе получения ментола и в способах биокаталитической конверсии других соединений со структурными мотивами терпенового типа. При выравнивании последовательностей различных SHC идентифицировано, что фенилаланин-486 (F486) является сильно консервативным аминокислотным остатком, и создана серия вариантов замен в ферменте SHC Zymomonas mobilis. Некоторые из этих замен приводили к утрате активности, тогда как другие приводили в результате к образованию нового терпеноидного продукта (изопулеголов) из терпеновых субстратов, таких как цитронеллаль.

В диссертации на соискание степени PhD Seitz (http://elib.uni-stuttgart.de/handle/11682/1400) в 2012 г. указано, что мутация F486Y в ZmoSHCl обеспечила снижение показателя биотрансформации гомофарнезола приблизительно в 1,5 раза, с 34,8% (WT ZmoSHCl) до 23,9% (мутантный ZmoSHCl F486Y). При тестировании мутантного эквивалента (Y420C) в AacSHC полагали, что ферментативная активность по отношению к субстратам большего размера будет снижаться, а активность, направленная на субстраты большего размера, будет повышаться. При тестировании этого мутанта в таких же условиях, как и дикого типа, и сравнении их ферментативной активности обнаружили, что мутант вообще не осуществлял какую-либо конверсию субстрата гомофарнезола. Таким образом, был сделан вывод о том, что аминокислотный остаток Y420 является критическим для активности AacSHC по отношению ко всем субстратам.

Другие исследования сайт-направленного мутагенеза SHC в данной области техники (например, Hoshino and Sato 2002, Chem Commun 291-301) были сосредоточены на действии мутаций в высоко консервативных участках (например, F601), и в большей степени на их действии на природные субстраты (т.е. на сквален или аналоги сквалена), чем на неприродные субстраты, такие как гомофарнезол.

Подводя итог вышесказанному, ограниченные сведения из уровня техники, относящиеся к способам биоконверсии для успешной конверсии гомофарнезола в амброксид, относятся только к относительно низким концентрациям/объемам субстрата гомофарнезола (в диапазоне концентраций от 0,25 мМ до 2 мМ до 10 мМ или около 0,06 г/л - 2,36 г/л) с использованием полипептида SHC дикого типа с активностью гомофарнезол-амброксид-циклазы (HAC). Источниками ферментов SHC с активностью HAC были либо: (i) жидкие экстракты, полученные путем разрушения клеток-хозяев E. coli, содержащих ферменты SHC, и отделения нерастворимой фракции от растворимой жидкой фракции SHC; либо (ii) частично очищенные мембранные фракции; либо (iii) рекомбинантные целые клетки, экспрессирующие ген WT SHC и продуцирующие фермент SHC, для применения в реакциях биоконверсии гомофарнезола в амброксид с использованием солюбилизирующих агентов, включающих либо: (i) Тритон X-100 в реакционной смеси (см. Neumann и Simon, 1986, цитируемый выше Seitz et al 2012, цитируемый выше JP2009060799); либо (ii) тауродезоксихолат (как описано в US 2012/0135477 A1).

При использовании этих экстрактов WT SHC и/или целых рекомбинантных клеток микроорганизмов, экспрессирующих ген SHC, было обнаружено, что получаемые доли конверсии изменяются в зависимости от источника фермента SHC, количества исходного вещества гомофарнезола и используемых реакционных условий. При использовании фермента SHC дикого типа 100% конверсия гомофарнезола в амброксид при указанных задействованных концентрациях (0,06-2,36 г/л) до сих пор не достигнута. Кроме того, в предварительных исследованиях с использованием производных SHC, полученных путем сайт-направленного мутагенеза, были только получены скорее отрицательные (т.е. сниженные доли конверсии гомофарнезола), чем положительные (т.е. улучшенные доли конверсии) результаты. Кроме того, в опубликованных исследованиях использовали только очищенные экстракты фермента SHC, или мембраносвязанные фракции фермента SHC, или целые рекомбинантные клетки микроорганизмов, экспрессирующих ген WT SHC в определенных реакционных условиях, при которых используются солюбилизирующие агенты, такие как Тритон X-100 или тауродезоксихолат. До сих пор не продемонстрировано, что рекомбинантный микроорганизм, содержащим либо WT, либо мутантный SHC, может обеспечить более эффективную и экономически результативную биоконверсию гомофарнезола в амброксид при использовании оптимизированных реакционных условий. Соответственно, желательно усовершенствовать цитируемые выше известные способы получения амброксида из гомофарнезола за счет по меньшей мере улучшения скорости, специфичности, выхода, производительности реакции и снижения затрат (за счет, например, упрощения способа при использовании либо рекомбинантных целых клеток микроорганизмов или при осуществлении метода «реакции, протекающей в одном реакционном сосуде», объединяющего стадии биокаталитического получения и биоконверсии).

## КРАТКОЕ ИЗЛОЖЕНИЕ СУЩНОСТИ ИЗОБРЕТЕНИЯ

В различных аспектах настоящего изобретения предложены производные SHC/HAC, аминокислотные последовательности, составляющие производные фермента SHC/HAC,

нуклеотидные последовательности, кодирующие производные фермента SHC/HAC, рекомбинантные клетки-хозяева, содержащие векторы с нуклеотидными последовательностями, кодирующими производные фермента SHC/HAC и применение рекомбинантных клеток-хозяев, содержащих либо производные фермента SHC/HAC, либо WT SHC/HAC, используемые в определенных реакционных условиях, в способах получения материалов амброксида, которые содержат изомер амброксида, обозначаемый как (-)-амброксид, и подобные амброксиду молекулы (в качестве побочных продуктов). В отличие от сведений предшествующего уровня техники, относящихся к AacSHC, автор изобретения впервые продемонстрировал, что цельноклеточный рекомбинантный микроорганизм, экспрессирующий ген производного SHC, можно использовать для биоконверсии гомофарнезола в амброксид. Кроме того, цельноклеточный рекомбинантный микроорганизм, экспрессирующий ген WT SHC и/или продуцирующий ферменты SHC, можно использовать для биоконверсии гомофарнезола в амброксид в определенных реакционных условиях, не раскрытых в предшествующем уровне техники.

Также было неожиданно обнаружено, что включение в аминокислотную последовательность эталонной последовательности WT SHC дикого типа вплоть до пяти аминокислотных изменений, раскрытых в настоящем описании, приводит к получению производных фермента SHC/HAC, которые обладают значительно лучшей долей конверсии гомофарнезола в амброксид по сравнению с немодифицированными эталонными ферментами SHC/HAC, как раскрыто в настоящем описании. Эти новые производные ферментов SHC/HAC полезны сами по себе и в комбинации для получения материалов амброксида, в частности (-)-амброксида, из субстратов гомофарнезола.

Дополнительное неожиданное открытие состоит в том, что, кроме одного мутанта (F601Y), производные фермента SHC, раскрытые в настоящем описании, как правило, содержат неконсервативные замены в положениях аминокислот в неконсервативном участке эталонной полипептидной последовательности SHC. Это открытие является неожиданным, поскольку изменения в консервативной области фермента более вероятно нарушают функцию фермента (по меньшей мере в отношении его природного субстрата), чем изменения в неконсервативной области белка.

Следующее неожиданное открытие состоит в том, что охарактеризованные производные ферменты SHC по настоящему изобретению оптимальнее работают (на неприродном субстрате, таком как гомофарнезол) при температуре около 35°C, чем при температуре около 60°C, которая является обычной реакционной температурой для термофильных микроорганизмов, таких как AacSHC. Применение производных SHC по настоящему изобретению в способах получения амброксида из гомофарнезола при пониженных реакционных температурах обладает значимыми экономическими преимуществами для цикла получения амброксида в промышленном масштабе.

Другое преимущество настоящего изобретения состоит в том, что производные фермента SHC по настоящему изобретению катализируют эффективный процесс биоконверсии, который при оптимизации с относительно высокой (например, около 50-кратной) концентрацией субстрата гомофарнезола по сравнению с концентрациями, ранее описанными в предшествующем уровне техники (например, EEH при 125 г/л) может приводить к 100% конверсии субстрата гомофарнезола, тогда как эталонный белок WT SHC конвертирует только приблизительно 10% того же субстрата даже при высокой концентрации фермента/клеток. Цитированные документы предшествующего уровня техники относятся к применению очищенных мембранных экстрактов, содержащих SHC, или очищенных экстрактов SHC (полученных из микроорганизмов,

экспрессирующих гены SHC), или к применению рекомбинантного микроорганизма, экспрессирующего ген WT SHC, в определенных реакционных условиях биоконверсии (например, с использованием определенных солюбилизирующих агентов). Но и теперь нигде не описана 100% конверсия гомофарнезола при значительно более низкой концентрации EEH. Также нигде не описана «реакция, протекающая в одном реакционном сосуде», где на первой стадии рекомбинантные клетки растут и продуцируют фермент SHC и впоследствии конвертируют EEH в (-)-амброксид в одном и том же реакторе. Дополнительное преимущество настоящего изобретения состоит в том, что, рекомбинантные клетки-хозяева, продуцирующие производные фермента SHC, демонстрируют высокие начальные скорости реакции, что позволяет получить большое количество продукта за относительно короткий период времени при использовании лишь относительно небольших количеств биокатализатора. Вкратце, отбор и эффективная экспрессия и применение рекомбинантных микроорганизмов, содержащих либо фермент WT SHC/HAC, либо определенные производные фермента SHC/HAC, в определенных реакционных условиях биоконверсии приводит к более эффективному процессу биоконверсии. Конечный продукт ((-)-амбр оксид) можно отделить и легко очистить. В отличие от цитируемых сведений из предшествующего уровня техники, производные фермента SHC/HAC применяют не в виде чистого фермента, а в цельноклеточном контексте (в качестве биокатализатора), что представляет собой более экономически эффективный, более удобный для использования и благоприятный для окружающей среды подход, поскольку он не требует дополнительных стадий очистки и выделения фермента.

Подводя итог, в настоящем изобретении предложен способ биоконверсии/биотрансформации для получения амброксида в рекомбинантном штамме микроорганизма, который (i) привлекателен с экономической точки зрения, (ii) благоприятен для окружающей среды и (iii) приводит к селективному получению (-)-амброксида в качестве преобладающего соединения, которое в селективных условиях кристаллизации эффективно отделяется от других побочных продуктов, не вносящих вклад в ольфакторное качество конечного продукта.

## ПОДРОБНОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ

Используемый в настоящем описании термин «SHC» означает фермент сквален-гопен-циклазу из любого источника из перечисленных в таблицах 10-12. В предпочтительных воплощениях изобретения термин «SHC» включает в себя ферменты SHC из Zymomonas mobilis и ферменты SHC из Alicyclobacillus acidocaldarius, как раскрыто в заявке на патент компании BASF WO 2010/139719, заявке на патент США US 2012/01345477 A1, публикациях Seitz et al (2012, цитируется выше) и Seitz (2012, диссертация PhD, цитируется выше). Для простоты ссылки в настоящем описании используют обозначение «AacSHC» для SHC из Alicyclobacillus acidocaldarius, обозначение «ZmoSHC» для SHC из Zymomonas mobilis и обозначение «BjpSHC» для SHC из Bradyrhizobium japonicum. Процентное значение идентичности этих последовательностей относительно последовательности WT AacSHC и друг друга (которое может изменяться в зависимости от используемого алгоритма) приведено в таблицах 18 и 19.

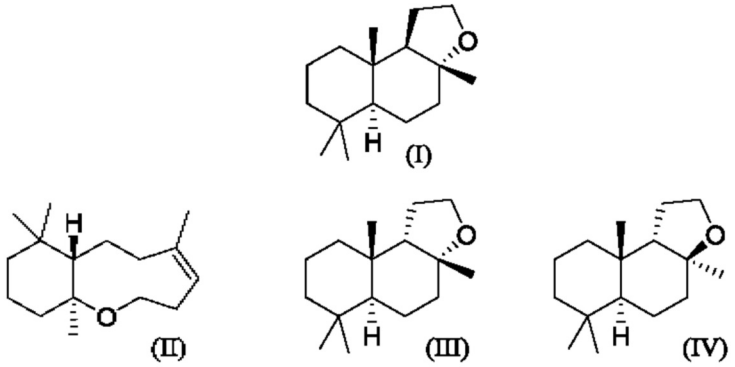
Выравнивание последовательностей WT SHC, полученное Hoshino и Sato (2002, цитируется выше), указывает на то, что во всех четырех последовательностях было обнаружено множество мотивов, состоящих из коровой последовательности Gln-X-X-X-G1y-X-Trp, которая обнаружена шесть раз в последовательностях SHC обоих микроорганизмов, Z. mobilis и A. acidocaldarius (см. Фиг. 3 публикации Reipen et al 1995, Microbiology 141, 155-161). В Hoshino и Sato (2002, цитируется выше) сообщается, что

ароматические аминокислоты с необычно высокой частотой встречаются в SHC, и в SHC было отмечено два характеристических мотива: один из которых представляет собой мотив QW, представленный специфическими аминокислотными мотивами [(K/R)(G/A)X2-3(F/Y/W)(L/IV)3X3QX2-5GXW], и альтернативный мотив, представляющий собой DXDDTA. B Wendt et al (1997, Science 277, 1811-1815 and 1999, J Mol Biol 286, 175-187) описан рентгеноструктурный анализ SHC из A. acidocaldarius. Мотив DXDDTA, по-видимому, соответствует активному центру SHC. Выравнивание иллюстративных последовательностей из предшествующего уровня техники демонстрирует множественные повторяющиеся мотивы, как показано на Фиг. 2 (из публикации Hoshino и Sato (2002, цитируется выше)) и на Фиг. 3 (из диссертации Seitz PhD (2012)) в настоящем описании.

Используемый в настоящем описании эталонный (или дикого типа) белок AacSHC относится к белку AacSHC, как изложено в SEQ ID NO: 1. Эталонный фермент AacSHC согласно настоящему описанию обладает активностью гомофарнезол-амброксид-циклазы (HAC), полезной при получении производных амброксида посредством биокаталитического взаимодействия SHC с субстратом гомофарнезолом. Основная реакция эталонного AacSHC состоит в циклизации линейного или нелинейного субстрата, такого как гомофарнезол, с получением амброксида.

Амброксид

При использовании в настоящем описании термин амброксид включает в себя (-)-амброксид формулы (I) либо в стереоизомерно чистой форме, либо в смеси с по меньшей мере одной или более чем одной из следующих молекул формулы (II), (IV) и/или (III).




(-)-Амброксид

(-)-Амброксид известен под торговыми наименованиями Ambrox (Firmenich), Ambroxan (Henkel), Ambrofix (Givaudan), Amberlyn (Quest), Cetalox Laevo (Firmenich), Ambermor (Aromor) и/или Norambrenolide Ether (Pacific).

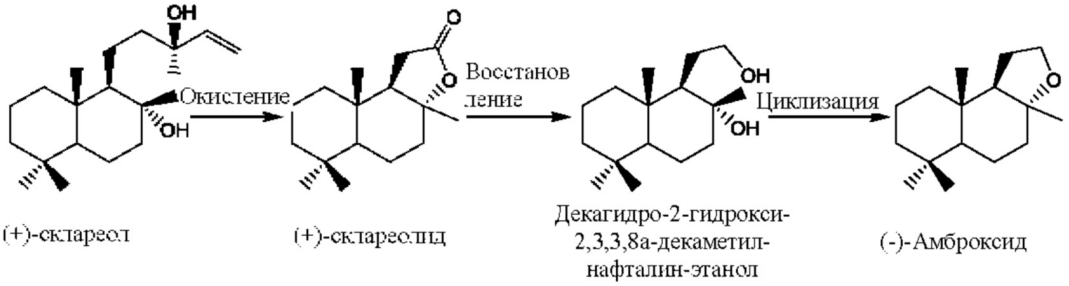
(-)-Амброксид представляет собой промышленно значимое ароматизирующее соединение и длительное время используется в производстве ароматизирующих веществ. Особые желаемые сенсорные преимущества (-)-амброксида происходят от (-)-стереоизомера скорее, чем от (+)-стереоизомера. Запах (-)-стереоизомера описывают как подобный мускусному, древесному, теплому или янтарному, тогда как энантиомер (+)-амброксид имеет относительно слабую ароматную ноту. Запах и пороги восприятия запаха для продуктов, подобных амброксиду, также различаются. Хотя в продаже имеются различные материалы, обогащенные

(-)-амброксидом, желательно получить высоко обогащенные (-)-амброксидом материалы, в идеале - чистый (-)-амбр оксид.

Получение (-)-амброксида

(-)-Амброксид может быть получен из склареолида в соответствии с описанным

ниже способом получения. Склареол представляет собой продукт, экстрагируемый из природного растения шалфея мускатного. Однако, поскольку в этом процессе используется натуральный исходное вещество, существующие потенциальные проблемы заключаются в том, что в нем задействована многостадийная реакция, которая является непрямой, количество и стабильность подачи исходного вещества может не всегда быть удовлетворительной, и реакция может быть неблагоприятной для окружающей среды в связи с использованием на стадии окислительного разложения (+)-склареола окисляющего агента, такого как хромовая кислота или перманганат.



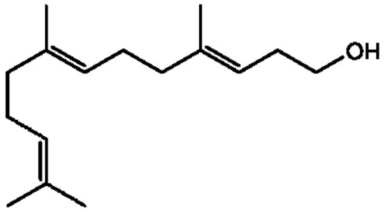

(-)-Амброксид также синтезируют из гомофарнезола, используя различные пути. В качестве примера гомофарнезол может быть получен путем бромирования, цианирования и гидролиза неролидола с получением гомофарнезиловой кислоты с последующим восстановлением. Альтернативно, гомофарнезол может быть получен из фарнезола, фарнезил хлорид а, бета-фарнезена или других субстратов. Бета-фарнезен может быть конвертирован непосредственно в E,E-гомофарнезол (EEH) или опосредованно в EEH через E,E-гомофарнезат, который впоследствии конвертируют в EEH. Обзор по получению (-)-амброксида из различных субстратов можно найти в US 2012/0135477 A1, WO 2010/139719, US 2013/0273619 A1, WO 2013/156398 A1, диссертации PhD Seitz (2012, цитируется выше) и в публикации Schaefer 2011 (Chemie Unserer Zeit 45, 374-388).

Хотя гомофарнезол может присутствовать в виде смеси четырех изомеров, (3Z,7Z), (3E,7Z), (3Z,7E) и (3E,7E) изомеров, исходя из литературы кажется, что (-)-амброксид получается только из (3E,7E) гомофарнезола (см. Neumann и Simon (1986), цитируется выше). Используемая в настоящем описании ссылка на (3E,7E)-гомофарнезол представляет собой ссылку на E,E-гомофарнезол, который также обозначается как EEH.

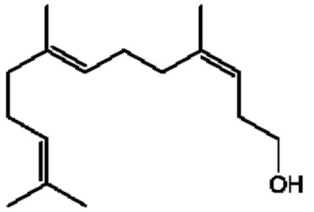
В US 2012/0135477 A1 описана конверсия (3Z,7E) в (-)-амброксид с использованием ZmoSHC (SEQ ID NO: 2) (см. примеры 2-4), однако согласно публикации Schaefer (2011) (цитируется выше), (7E, 3Z) конвертируется только в 9b-эпи-амброксид (т.е. соединение III), как описано выше, но не конвертируется в (-)-амброксид. Используемая в настоящем описании ссылка на (3E,7E)-гомофарнезол представляет собой ссылку на E,Z-гомофарнезол, который также обозначен как EZH.

В некоторых воплощениях изобретения гомофарнезол предпочтительно содержит смесь (3E,7E) и (3Z,7E), обозначаемую в настоящем описании как стереоизомерная смесь EE:EZ (в частности, при ссылки на примеры и таблицу 20).

Стереоизомерная смесь EE:EZ гомофарнезола имеет номер по химическому реферативному журналу (Chemical Abstracts, CAS) 35826-67-6.




(3E,7E)-Гомофарнезол
CAS: 459-89-2




(3Z,7E)-Гомофарнезол
CAS: 138152-06-4

Смесь E/Z CAS: 35826-67-6

Как показано в примерах (например, см. примеры 5, 7, 9, 10, 11, 18, 19 и 20), в некоторых воплощениях изобретения исходное сырье/исходный вещество гомофарнезола представляет собой смесь изомеров.

Соответственно, в некоторых воплощениях изобретения исходное вещество гомофарнезола может также содержать смесь четырех изомеров EE:EZ:ZZ:ZE, которая соответствует (3E,7E) и (3Z,7E), (3Z,7Z и 3E,7Z).

В некоторых воплощениях изобретения исходное вещество гомофарнезола предпочтительно выбрано из одной или более из следующих групп: [(3Z,7Z), (3E,7Z), (3Z,7E) и (3E,7E)], [(3Z,7E) и (3E,7E)], [(3Z,7E), (3E,7Z)] и/или [(3E,7E) и (3E,7Z)].

Предпочтительно исходное вещество гомофарнезола выбрано из одной или более из следующих групп: [(3E,7E), (3Z,7E)] и/или [(3Z,7E), (3E,7E) и (3E,7Z)].

Соответственно, в некоторых воплощениях изобретения соотношение EEH:EZH составляет приблизительно 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; или приблизительно 50:50.

В некоторых воплощениях изобретения исходное вещество гомофарнезола предпочтительно содержит >90% E,E-гомофарнезола (EEH).

В других воплощениях изобретения исходное вещество гомофарнезола содержит EE:EZ в массовом соотношении 86:14.

В некоторых воплощениях изобретения исходное вещество гомофарнезола содержит EE:EZ в массовом соотношении 80:20.

В некоторых воплощениях изобретения исходное вещество гомофарнезола содержит EE:EZ в массовом соотношении 70:30.

В следующих воплощениях изобретения исходное вещество гомофарнезола содержит EE:EZ в массовом соотношении 69:31.

В некоторых воплощениях изобретения исходное вещество гомофарнезола состоит или по существу состоит из смеси четырех изомеров EE:EZ:ZZ:ZE, которая соответствует (3E,7E) и (3Z,7E), (3Z,7Z) и (3E,7Z).

В некоторых воплощениях изобретения исходное вещество гомофарнезола состоит или по существу состоит из смеси изомеров, выбранных из одной или более из следующих групп: [(3Z,7Z), (3E,7Z), (3Z,7E) и (3E,7E)], [(3Z,7E) и (3E,7E)], [(3Z,7E), (3E,7Z)] и/или [(3E,7E) и (3E,7Z)].

Предпочтительно исходное вещество гомофарнезола состоит или по существу состоит из смеси изомеров, выбранных из одной или более из следующих групп: [(3E,7E), (3Z,7E)] и/или [(3Z,7E), (3E,7E) и (3E,7Z)].

Соответственно, в некоторых воплощениях изобретения соотношение изомеров EEH:EZH состоит или по существу состоит из соотношения EEH:EZH приблизительно

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; или приблизительно 50:50.

В некоторых воплощениях изобретения исходное вещество гомофарнезола предпочтительно состоит или по существу состоит из >90% E,E-гомофарнезола (EEH).

В других воплощениях изобретения исходное вещество гомофарнезола состоит или по существу состоит из EE:EZ в массовом соотношении 86:14.

В некоторых воплощениях изобретения исходное вещество гомофарнезола состоит или по существу состоит из EE:EZ в массовом соотношении 80:20.

В некоторых воплощениях изобретения исходное вещество гомофарнезола состоит или по существу состоит из EE:EZ в массовом соотношении 70:30.

В следующих воплощениях изобретения исходное вещество гомофарнезола состоит или по существу состоит из EE:EZ в массовом соотношении 69:31.

В воплощениях настоящего изобретения амброксид получают с использованием производного фермента SHC/HAC.

Производное SHC/HAC

Используемый в настоящем описании термин «производное SHC/HAC» означает, что аминокислотная последовательность производного SHC/HAC представляет собой модифицированную аминокислотную последовательность или ее вариант, измененные по сравнению с аминокислотной последовательностью эталонной (или дикого типа) последовательности SHC, которая соответствует по меньшей мере SEQ ID NO: 1, или SEQ ID NO: 2, или SEQ ID NO: 3, или SEQ ID NO: 4. Как правило, производные SHC/HAC содержат измененные формы SHC, имеющие по меньшей мере одно изменение, которое модифицирует (например, увеличивает) активность фермента по отношению к его субстрату (например, EEH).

Производные SHC/HAC по настоящему изобретению тестируют на их активность гомофарнезол-амброксид-циклазы. Следовательно, эти производные SHC/HAC, которые конвертируют гомофарнезол в амброксид, в настоящем описании называют как производными HAC, так и производные SHC. Хотя в качестве примеров производных SHC/HAC были предложены ферменты, имеющие происхождение из источников, представляющих собой штаммы микроорганизмов Alicyclobacillus acidocaldarius, Zymomonas mobilis, Bradyrhizobium japonicum, настоящее изобретение также охватывает эквивалентные производные SHC/HAC из источников, представляющих собой другие штаммы микроорганизмов, включающих в себя, без ограничений, ферменты SHC/HAC из Methylococcus capsulatus, Frankia alni, Acetobacter pasteurianum и Tetrahymena pyriformis (см., например, WO 2010/139719, US 2012/01345477, WO 2012/066059 и таблицы 10-12).

Используемый в настоящем описании термин «изменение аминокислоты» означает инсерцию одной или более чем одной аминокислоты между двумя аминокислотами, делецию одной или более чем одной аминокислоты или замену (которая может быть консервативной или неконсервативной) одной или более чем одной аминокислоты на одну или более чем одну из других аминокислот относительно аминокислотной последовательности эталонной аминокислотной последовательности (такой как, например, аминокислотная последовательность дикого типа (WT) SEQ ID NO: 1, или SEQ ID NO: 2, или SEQ ID NO: 3, или SEQ ID NO: 4). Изменения аминокислот могут быть легко идентифицированы путем сравнения аминокислотных последовательностей производного SHC/HAC и эталонной аминокислотной последовательности (такой как, например, аминокислотная последовательность дикого типа (WT) SEQ ID NO: 1, или

SEQ ID NO: 2, или SEQ ID NO: 3, или SEQ ID NO: 4). Примеры выравнивания аминокислотной последовательности WT SHC приведены на Фиг. 14 и в таблицах 18 и 19.

Консервативные аминокислотные замены могут быть выполнены, например, на основании подобия полярности, заряда, размера, растворимости, гидрофобности, гидрофильности и/или амфипатической природы задействованных аминокислотных остатков. 20 встречающихся в природе аминокислот, как описано выше, могут быть сгруппированы в шесть следующих стандартных групп аминокислот:

(1) гидрофобные: Met, Ala, Val, Leu, Ile;

(2) нейтральные гидрофильные: Cys, Ser, Thr; Asn, Gln;

(3) кислые: Asp, Glu;

(4) основные: His, Lys, Arg;

(5) остатки, влияющие на ориентацию цепи: Gly, Pro; и

(6) ароматические: Trp, Tyr, Phe.

Соответственно, используемый в настоящем описании термин «консервативные замены» означают замену аминокислоты на другую аминокислоту, включенную в перечень в пределах той же группы из шести стандартных групп аминокислот, представленных выше. Например, в модифицированном таким путем полипептиде замена Asp на Glu сохраняет один отрицательный заряд. Кроме того, глицин и пролин могут быть заменены друг на друга за счет их способности к прерыванию альфа-спиралей. Некоторые предпочтительные консервативные замены в пределах описанных выше шести групп представляют собой обмены внутри следующих подгрупп: (i) Ala, Val, Leu и Ile; (ii) Ser и Thr; (ii) Asn и Gln; (iv) Lys и Arg; и (v) Tyr и Phe. С учетом известного генетического кода и методов рекомбинантных ДНК и синтеза ДНК ученый-специалист сможет легко сконструировать ДНК, кодирующие консервативные варианты аминокислот.

Используемый в настоящем описании термин «неконсервативные замены» или «обмены неконсервативных аминокислот» определяют как обмены аминокислоты с другой аминокислотой, включенной в перечень другой группы из шести стандартных групп аминокислот (1)-(6), представленных выше. Как правило, производные SHC/HAC по настоящему изобретению получают с использованием неконсервативных замен, изменяющих биологическую функцию (например, активность HAC) раскрытых производных SHC/HAC.

Для простоты ссылки ниже указаны однобуквенные символы аминокислот, рекомендованные Комиссией по биохимической номенклатуре IUPAC-IUB. В целях ссылки также приведены трехбуквенные коды.

| Однобуквенный код | Трехбуквенный код | Название аминокислоты |
|---|---|---|
| A | Ala | Аланин |
| C | Cys | Цистеин |
| D | Asp | Аспарагиновая кислота |
| E | Glu | Глутаминовая кислота |
| F | Phe | Фенилаланин |
| G | Gly | Глицин |
| H | His | Гистидин |
| I | Ile | Изолейцин |
| K | Lys | Лизин |
| L | Leu | Лейцин |
| M | Met | Метионин |
| N | Asn | Аспарагин |
| P | Pro | Пролин |
| Q | Gln | Глутамин |
| R | Arg | Аргинин |
| S | Ser | Серин |
| T | Thr | Треонин |
| V | Val | Валин |
| W | Trp | Триптофан |
| Y | Tyr | Тирозин |

Изменения аминокислот, такие как аминокислотные замены, можно вводить, используя известные протоколы технологии рекомбинантных генов, включающие полимеразную цепную реакцию (ПЦР), клонирование генов, сайт-направленный мутагенез кДНК, трансфекция клеток-хозяев и транскрипция in-vitro, которые можно использовать для введения таких изменений в последовательность WT SHC с получением в результате производного фермента SHC/HAC. Впоследствии эти производные можно подвергать скринингу на функциональную активность SHC/HAC.

Производные фермента SHC/HAC

В настоящем изобретении предложено производное SHC/HAC и описан фермент, обладающий активностью гомофарнезол-амброксид-циклазы (HAC) и содержащий аминокислотную последовательность, которая имеет от 1 до приблизительно 50 мутаций, независимо выбранных из замен, делеций или инсерций, относительно аминокислотной последовательности эталонной (или дикого типа) последовательности SHC в

соответствии с по меньшей мере одной из SEQ ID NO: 1, или SEQ ID NO: 2, или SEQ ID NO: 3, или SEQ ID NO: 4.

В различных воплощениях изобретения мутация или комбинация мутаций усиливает активность производного SHC/HAC для конвертирования гомофарнезола в амброксид в сравнении с эталонными ферментами SHC, которые не демонстрируют эту делецию/добавление. Для направленного получения таких замен, делеций или инсерций в эталонной последовательности SHC можно использовать описанное в настоящем описании моделирование белков. Например, можно создать структурную модель аминокислотной последовательности SHC с использованием координат для AacSHC (как показано, например, на Фиг.19 и 20). Как продемонстрировано в настоящем описании, такая модель гомологии полезна для направленного усовершенствования конвертирования гомофарнезола в (-)-амброксид ферментами SHC.

Таким образом, в различных воплощениях изобретения производное SHC/HAC может иметь от приблизительно 1 до приблизительно 45 мутаций, от приблизительно 1 до приблизительно 40 мутаций, от приблизительно 1 до приблизительно 35 мутаций, от приблизительно 1 до приблизительно 30 мутаций, от приблизительно 1 до приблизительно 25 мутаций, от приблизительно 1 до приблизительно 20 мутаций, от приблизительно 1 до приблизительно 15 мутаций, от приблизительно 1 до приблизительно 10 мутаций или от приблизительно 1 до приблизительно 5 мутаций относительно аминокислотной последовательности эталонной (или дикого типа) последовательности SHC в соответствии с по меньшей мере одной из SEQ ID NO: 1, или SEQ ID NO: 2, или SEQ ID NO: 3, или SEQ ID NO: 4.

В различных воплощениях изобретения производное SHC/HAC содержит последовательность, имеющую по меньшей мере 5 или по меньшей мере 10 мутаций относительно аминокислотной последовательности эталонной (или дикого типа) последовательности SHC в соответствии с по меньшей мере одной из SEQ ID NO: 1, или SEQ ID NO: 2, или SEQ ID NO: 3, или SEQ ID NO: 4, но не более 20 или 30 мутаций. В различных воплощениях изобретения производное SHC может иметь приблизительно 1 мутацию, приблизительно 2 мутации, приблизительно 3 мутации, приблизительно 4 мутации, приблизительно 5 мутаций, приблизительно 6 мутаций, приблизительно 7 мутаций, приблизительно 8 мутаций, приблизительно 9 мутаций, приблизительно 10 мутаций, приблизительно 11 мутаций, приблизительно 12 мутаций, приблизительно 13 мутаций, приблизительно 14 мутаций, приблизительно 15 мутаций, приблизительно 16 мутаций, приблизительно 17 мутаций, приблизительно 18 мутаций, приблизительно 19 мутаций, приблизительно 20 мутаций, приблизительно 21 мутацию, приблизительно 22 мутации, приблизительно 23 мутации, приблизительно 24 мутации, приблизительно 25 мутаций, приблизительно 26 мутаций, приблизительно 27 мутаций, приблизительно 28 мутаций, приблизительно 29 мутаций, приблизительно 30 мутаций, приблизительно 31 мутацию, приблизительно 32 мутации, приблизительно 33 мутации, приблизительно 34 мутации, приблизительно 35 мутаций, приблизительно 36 мутаций, приблизительно 37 мутаций, приблизительно 38 мутаций, приблизительно 39 мутаций, приблизительно 40 мутаций, приблизительно 41 мутацию, приблизительно 42 мутации, приблизительно 43 мутации, приблизительно 44 мутации, приблизительно 45 мутаций, приблизительно 46 мутаций, приблизительно 47 мутаций, приблизительно 48 мутаций, приблизительно 49 мутаций или приблизительно 50 мутаций относительно эталонного SHC (такого как, например, SEQ ID NO: 1, или 2, или 3, или 4.

В этих или в других воплощениях изобретения производное SHC/HAC может содержать аминокислотную последовательность, обладающую по меньшей мере

приблизительно 50% идентичностью последовательности, по меньшей мере приблизительно 55% идентичностью последовательности, по меньшей мере приблизительно 60% идентичностью последовательности, по меньшей мере приблизительно 65% идентичностью последовательности, по меньшей мере приблизительно 70% идентичностью последовательности, по меньшей мере приблизительно 75% идентичностью последовательности, по меньшей мере приблизительно 80% идентичностью последовательности, по меньшей мере приблизительно 85% идентичностью последовательности, по меньшей мере приблизительно 90% идентичностью последовательности, по меньшей мере приблизительно 91% идентичностью последовательности, по меньшей мере приблизительно 92% идентичностью последовательности, по меньшей мере приблизительно 93% идентичностью последовательности, по меньшей мере приблизительно 94% идентичностью последовательности, по меньшей мере приблизительно 95% идентичностью последовательности, по меньшей мере приблизительно 96% идентичностью последовательности, по меньшей мере приблизительно 97% идентичностью последовательности, по меньшей мере приблизительно 98% идентичностью последовательности или по меньшей мере приблизительно 99% идентичностью последовательности WT SHC (такой как, например, SEQ ID NO: 1, или SEQ ID NO: 2, или SEQ ID NO: 3, или SEQ ID NO: 4) или между эталонными последовательностями (см., например, таблицы 18 и 19, где показана идентичность 34-52% между AacSHC (SEQ ID NO: 1) и другими последовательностями SHC (например, ZmoSHC по WO 2010/139719).

В различных воплощениях изобретения вариант SHC обладает более высокой активностью конвертирования гомофарнезола в амброксид, чем фермент дикого типа, например более высокой продуктивностью в отношении (-)-амброксида при приведении в контакт с субстратом гомофарнезолом по сравнению с эталонным ферментом дикого типа (таким как, например, SEQ ID NO: 1, или SEQ ID NO: 2, или SEQ ID NO: 3, или SEQ ID NO: 4).

Например, производное SHC/HAC может содержать аминокислотную последовательность, обладающую по меньшей мере: приблизительно 50% идентичностью, приблизительно 51% идентичностью, приблизительно 52% идентичностью, приблизительно 53% идентичностью, приблизительно 54% идентичностью, приблизительно 55% идентичностью, приблизительно 56% идентичностью, приблизительно 57% идентичностью, приблизительно 58% идентичностью, приблизительно 59% идентичностью, приблизительно 60% идентичностью, приблизительно 61% идентичностью, приблизительно 62% идентичностью, приблизительно 63% идентичностью, приблизительно 64% идентичностью, приблизительно 65% идентичностью, приблизительно 66% идентичностью, приблизительно 67% идентичностью, приблизительно 69% идентичностью, приблизительно 70% идентичностью, приблизительно 71% идентичностью, приблизительно 72% идентичностью, приблизительно 73% идентичностью, приблизительно 74% идентичностью, приблизительно 75% идентичностью, приблизительно 76% идентичностью, приблизительно 77% идентичностью, приблизительно 78% идентичностью, приблизительно 79% идентичностью, приблизительно 80% идентичностью, приблизительно 81% идентичностью, приблизительно 82% идентичностью, приблизительно 83% идентичностью, приблизительно 84% идентичностью, приблизительно 85% идентичностью, приблизительно 86% идентичностью, приблизительно 87%

идентичностью, приблизительно 88% идентичностью, приблизительно 89% идентичностью, приблизительно 90% идентичностью, приблизительно 91% идентичностью, приблизительно 92% идентичностью, приблизительно 93% идентичностью, приблизительно 94% идентичностью, приблизительно 95% идентичностью, приблизительно 96% идентичностью, приблизительно 97% идентичностью, приблизительно 98% идентичностью или приблизительно 99% идентичностью последовательности эталонного SHC (такого как, например, SEQ ID NO: 1, или 2, или 3, или 4) или между эталонными последовательностями (см., например, таблицы 18 и 19, где показана идентичность 34-52% между AacSHC (SEQ ID NO: 1) и другими последовательностями SHC (например, ZmoSHC по WO 2010/139719).

Различные производные SHC/HAC, протестированные на активность фермента SHC, включены в перечень одной или более из таблиц 1-9. Таким образом, в различных воплощениях изобретения производное SHC/HAC может иметь по меньшей мере приблизительно 1, по меньшей мере приблизительно 2, по меньшей мере приблизительно 3, по меньшей мере приблизительно 4, по меньшей мере приблизительно 5, по меньшей мере приблизительно 6, по меньшей мере приблизительно 7, по меньшей мере приблизительно 8, по меньшей мере приблизительно 9 или по меньшей мере приблизительно 10 мутаций, выбранных из одной или более мутаций в таблицах 1-9. В некоторых воплощениях изобретения производное SHC/HAC представляет собой модифицированный полипептид SHC, содержащий аминокислотную последовательность, имеющую вплоть до 4 мутаций по сравнению с аминокислотной последовательностью дикого типа/эталонной аминокислотной последовательностью, соответствующей SEQ ID NO: 1, и содержит по меньшей мере замены F601Y или M132R в комбинации с по меньшей мере одной или более чем одной из замен F129L и/или I432T относительно SEQ ID NO: 1, а также необязательно содержит лидерную последовательность, поддерживающую экспрессию и активность в E. coli.

В других воплощениях изобретения производное SHC/HAC представляет собой модифицированный полипептид SHC, содержащий аминокислотную последовательность, имеющую вплоть до 8 мутаций по сравнению с аминокислотной последовательностью дикого типа/эталонной аминокислотной последовательностью, соответствующей SEQ ID NO: 1 (или ее копией, модифицированной для экспрессии в E. coli) и содержит одно или более изменений аминокислот, выбранных из группы, состоящей из положений 77, 92, 129, 132, 224, 432, 579, 601 и 605 относительно SEQ ID NO: 1, где производное SHC/HAC обладает модифицированной (например, повышенной) ферментативной активностью относительно SEQ ID NO: 1.

В одном воплощении изобретения производное SHC содержит одну или более чем одну замену, выбранную из группы мутантов, состоящей из: T77X, I92X, F129X, M132X, A224X, I432X, Q579X, F601Y и F605W относительно SEQ ID NO: 1, где:

В T77X аминокислота X выбрана из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В I92X аминокислота X выбрана из: A C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В F129X аминокислота X выбрана из: A C, D, E, F, G, H, I, K, L, M, N, P, Q, R S, T, V, W или Y.

В M132X аминокислота X выбрана из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В A224X аминокислота X выбрана из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R S, T, V, W или Y.

В I432X аминокислота X выбрана из: A C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В Q579X аминокислота X выбрана из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R S, T, V, W или Y.

В F601X аминокислота X выбрана из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В F605X аминокислота X выбрана из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В одном воплощении изобретения производное SHC содержит одну или более чем одну замену, выбранную из группы мутантов, состоящей из: T77A, 192V, F129L, M132R, A224V, I432T, Q579H, F601Y и F605W относительно SEQ ID NO: 1.

В другом воплощении изобретения производное SHC содержит одну или более чем одну замену, выбранную из группы мутантов, состоящей из: S129X, V145X, F182X, Y185X, G282X, I498X, H646X, и F698X относительно SEQ ID NO: 2, где:

В S129X аминокислота X выбрана из: A C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В V145X аминокислота X выбрана из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В F182X аминокислота X выбрана из: A C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В Y185X аминокислота X выбрана из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В G282X аминокислота X выбрана из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В I498X аминокислота X выбрана из: A, B, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В H646X аминокислота X выбрана из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В F668X аминокислота X выбрана из: A C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В F698X аминокислота X выбрана из: A C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В одном воплощении изобретения производное SHC содержит одну или более чем одну замен, выбранных из группы мутантов, состоящей из: S129A, V145V, F182L, Y185R, G282V, I498T, H646H, F668Y и F698X относительно SEQ ID NO: 2, как показано в таблице 2.

В дополнительном воплощении изобретения производное SHC содержит одну или более чем одну замену, выбранную из группы мутантов, состоящей из: G85X, V100X, F137X, I140X, V233X, I450X, N598X, F620X и F624X относительно SEQ ID NO: 3, где:

В G85X аминокислота X выбрана из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В V100X аминокислота X выбрана из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В F137X аминокислота X выбрана из: A C, D, E, F, G, H, I, K, L, M, N, P, Q, R S, T, V, W или Y.

В I140X аминокислота X выбрана из: A C, D, E, F, G, H, I, K, L, M, N, P, Q, R S, T, V, W или Y.

В V233X аминокислота X выбрана из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T,

V, W или Y.

В I450X аминокислота X выбрана из: A C, D, E, F, G, H, I, K, L, M, N, P, Q, R S, T, V, W или Y.

В N598X аминокислота X выбрана из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В F620X аминокислота X выбрана из: A C, D, E, F, G, H, I, K, L, M, N, P, Q, R S, T, V, W или Y.

В F624X аминокислота X выбрана из: A C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В одном воплощении изобретения производное SHC содержит одну или более чем одну замену, выбранную из группы мутантов, состоящей из: G85A, V100V, F137L, I140R, V233V, I450T, N598H, F620Y и F624W относительно SEQ ID NO: 3, как показано в таблице 3 и в таблице 3a.

В дополнительном воплощении изобретения производное SHC содержит одну или более чем одну замену, выбранную из группы мутантов, состоящей из: A88X, V104X, F141X, Y144X, V241X, I459X, M607X, F628X и F658X относительно SEQ ID NO: 4, где:

В A88X аминокислота X выбрана из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В V104X аминокислота X выбрана из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R S, T, V, W или Y.

В F141X аминокислота X выбрана из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В Y144X аминокислота X выбрана из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В V241X аминокислота X выбрана из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В I459X аминокислота X выбрана из: A C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В M607X аминокислота X выбрана из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В F628X аминокислота X выбрана из: A C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

В F658X аминокислота X выбрана из: A Q D, E, F, G, H, L K, L, M, N, P, Q, R, S, T, V, W или Y.

В дополнительном воплощении изобретения производное SHC содержит одну или более чем одну замену, выбранную из группы, состоящей из: A88A, V104V, F141L, Y144R, V241V, I459T, M607H, F628Y и F658W относительно SEQ ID NO: 4, как показано в таблице 4.

Комбинации производных SHC

В одном воплощении изобретения производное SHC содержит одну или более чем одну замену, выбранную из группы мутантов, состоящей из: T77A, F129L, M132R, I92V, A224V, I432T, Q579H и F601Y относительно SEQ ID NO: 1, как показано в таблице 5.

В одном воплощении изобретения производные SHC содержат одну или более чем одну замену, выбранную из группы мутантов, состоящей из: S129A, V145V, F182L, Y185R, G282V, I498T, H646H и F668Y относительно SEQ ID NO: 2, как показано в таблице 6.

В одном воплощении изобретения производные SHC содержат одну или более чем одну замену, выбранную из группы мутантов, состоящей из: G85A, V100V, F137L, I140R,

V233V, I450T, N598H и F620Y относительно SEQ ID NO: 3, как показано в таблице 7.

В дополнительном воплощении изобретения производное SHC содержит одну или более чем одну замену, выбранную из группы, состоящей из: A88A, V104V, F141L, Y144R, V241V, I459T, M607H, и F628Y относительно SEQ ID NO: 4, как показано в таблице 8.

**Таблица 1: Перечень мутаций SHC, пронумерованных относительно AacSHC дикого типа (SEQ ID NO: 1) и на фигуре 1 из работы Hoshino и Sato (2002, цитируется выше)**

| № | Положение в WT AacSHC | Локализация | Аминокислота (AacSHC) (SEQ ID NO: 1) | Аминокислота (ZmoSHC) (Hoshino и Sato 2002) | Аминокислота (*H. sapiens*) (Hoshino и Sato 2002) | Мутация | Аминокислота в SEQ ID NO: (Таблица 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 |

**Таблица 2: Дополнительный перечень мутаций SHC, пронумерованных относительно AacSHC дикого типа (SEQ ID NO: 1), и выравнивание последовательностей AacSHC и ZmoSHC1 (Seitz *et al* 2012 – цитируется выше)**

| № | Положение и АК в WT AacSHC (SEQ ID NO: 1) | | Эквивалентная АК и положение в ZmoSHC1 (SEQ ID NO: 2) | | Мутация | АК в SEQ ID NO: (см. таблицу 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 |
| 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 |

**Таблица 3: Дополнительный перечень мутаций SHC, пронумерованных относительно AacSHC дикого типа (SEQ ID NO: 1), и выравнивание последовательностей AacSHC и ZmoSHC2 (Seitz *et al* 2012 – цитируется выше)**

| № | Положение и АК в WT AacSHC (SEQ ID NO: 1) | | Эквивалентная АК и положение в ZmoSHC2 (SEQ ID NO: 3) | | Мутация | АК в SEQ ID NO: (см. таблицу 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 |

**Таблица 3a: Перечень мутаций SHC, пронумерованных относительно AacSHC дикого типа (SEQ ID NO: 1), и выравнивание последовательности AacSHC с последовательностью No. 20 на фигуре выравнивания SHC из диссертации PhD Merkofer (2004) (см. http://elib.uni-stuttgart.de/handle/11682/1400)**

| № | Положение и АК в WT AacSHC (SEQ ID NO: 1) | | АК в ZmoSHC | Мутация |
|---|---|---|---|---|
| 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 |

**Таблица 4: Перечень мутаций SHC, пронумерованных относительно AacSHC дикого типа (SEQ ID NO: 1), и выравнивание последовательностей AacSHC и BjpSHC (SEQ ID NO: 5 в WO 2010/139719)**

| № | Положение и АК в WT AacSHC (SEQ ID NO: 1) | | АК и положение в WT BjpSHC (SEQ ID NO: 4) | | Мутация | АК в SEQ ID NO: (см. таблицу 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 |

**Таблица 4a: Используя выравнивания последовательностей, приведенные в таблицах 21a-21j, где WTAacSHC (SEQ ID NO: 1) выравнивают с любой из SEQ ID NO: 149, 151, 153, 155, 157, 159 (как идентифицировано ниже), можно идентифицировать АК остаток и положения, соответствующие T77, I92, F129, M132, A224, I432, Q579 и F601 в WT AacSHC (SEQ ID NO: 1), и тестировать на активность SHC/HAC**

| Аминокислота SEQ ID NO: | Штамм | SEQ ID NO: в WO 2010/0139719 (нуклеотидные/аминокислотные последовательности) | Нуклеотид/аминокислота SEQ ID NO: |
|---|---|---|---|
| SEQ ID NO: 149 | *Burkholderia ambifaria*6 | SEQ ID NO: 6 | 150 |
| SEQ ID NO: 151 | *Burkholderia ambifaria* | SEQ ID NO: 7 | 152 |
| SEQ ID NO: 153 | *Bacillus anthracis* | SEQ ID NO: 8 | 154 |
| SEQ ID NO: 155 | *Frankia alni* | SEQ ID NO: 9 | 156 |
| SEQ ID NO: 157 | *Rhodopseudomonas palent* | SEQ ID NO: 10 | 158 |
| SEQ ID NO: 159 | *Streptomyces coelicolor* | SEQ ID NO: 11 | 160 |
| SEQ ID NO: 161 | *Zymomonas mobilis*2 | SEQ ID NO: 2 | 162 |
| SEQ ID NO: 163 | *Zymomonas mobilis* | SEQ ID NO: 1 | 164 |
| SEQ ID NO: 4 | *Bradyrhizobium japonicum*5 | SEQ ID NO: 5 | 168 |

**Таблица 5: Перечень комбинаций мутаций SHC, пронумерованных в соответствии с AacSHC дикого типа (SEQ ID NO: 1)**

| Комбинации мутаций в AacSHC (SEQ ID NO: 1) | Производное SHC ID | Номер мутаций | Аминокислота SEQ ID NO: | Нуклеотид 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 |

**Таблица 6: Перечень комбинаций мутаций SHC, пронумерованных в соответствии с последовательностью ZmoSHC1 дикого типа (SEQ ID NO: 2)**

| Комбинации мутаций в ZmoSHC1 (SEQ ID NO: 2) | Производное SHC ID | номер мутаций | Аминокислота SEQ ID NO: | Нуклеотид 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 |

**Таблица 7: Перечень комбинаций мутаций SHC, пронумерованных в соответствии с последовательностью ZmoSHC2 дикого типа (SEQ ID NO: 3)**

| Комбинации мутаций в ZmoSHC2 (SEQ ID NO: 3) | Производное SHC ID | Номер мутаций | Аминокислота SEQ ID NO: | Нуклеотид 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 |

**Таблица 8: Перечень комбинаций мутаций SHC, пронумерованных в соответствии с BjpSHC дикого типа (SEQ ID NO: 4)**

| Комбинации мутаций в BjpSHC (SEQ ID NO: 4) | Производное SHC ID | Номер мутаций | Аминокислота SEQ ID NO: | Нуклеотид 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 |
| M144R + I459T + F628Y | SHC32 Bjp | 3 | 145 | 146 |
| F141L + Y144R + I459T + F628Y | SHC33 Bjp | 4 | 147 | 148 |

**Таблица 9: Показаны общие мутации SHC относительно WT AacSHC (SEQ ID NO: 1), WT ZmoSHC1 (SEQ ID NO: 2), WT ZmoSHC2 (SEQ ID NO: 3) и BjpSHC (SEQ ID NO: 4)**

| Комбинации мутаций | | | | |
|---|---|---|---|---|
| в AacSHC (SEQ ID NO: 1) | Относительно ZmoSHC1 (SEQ ID NO: 2) | Относительно ZmoSHC2 (SEQ ID NO: 3) | Относительно BjpSHC (SEQ ID NO: 4) | Производное SHC |
| 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 |

В предпочтительном воплощении изобретения производное SHC содержит по меньшей мере замены F601Y или M132R в комбинации с по меньшей мере любой одной или более из F129L и/или I432T относительно SEQ ID NO: 1.

Производное SHC, обозначенное как SHC3, предложенное в настоящем изобретении, содержит следующую замену F601Y по сравнению с эталонным белком SHC (SEQ ID NO: 1).

Hoshino и Sato (2002, цитируется выше) идентифицировали F601 как высококонсервативный аминокислотный остаток среди прокариотических и эукариотических биологических видов. Сообщали, что производное SHC F601Y характеризуется значительно повышенной Vmax для субстрата оксидосквалена (но не сквалена). Однако F601Y характеризуется снижением аффинности (т.е. повышенной $K_M$) и снижением каталитической эффективности/активности ($Kcat/K_M$) относительно WT AacSHC при использовании сквалена. В работе Hoshino и Sato (2002, цитируется выше) не приведены данные по эффективности AacSHC при использовании гомофарнезола в качестве субстрата фермента с мутацией F601Y.

Производное SHC, обозначенное как SHC10, предложенное в настоящем изобретении, содержит следующую замену F129L по сравнению с эталонным белком SHC (SEQ ID

NO: 1).

Производное SHC, обозначенное как SHC30, предложенное в настоящем изобретении, содержит следующую замену F601Y и F129L по сравнению с эталонным белком SHC (SEQ ID NO: 1).

Производное SHC, обозначенное как SHC26, предложенное в настоящем изобретении, содержит следующую замену M132R и I432T по сравнению с эталонным белком SHC (SEQ ID NO: 1).

Производное SHC, обозначенное как 215G2, предложенное в настоящем изобретении, содержит следующую замену M132R, I432T и A224V по сравнению с эталонным белком SHC (SEQ ID NO: 1).

Производное SHC, обозначенное как SHC32, предложенное в настоящем изобретении, содержит следующую замену F601Y, M132R и I432T по сравнению с эталонным белком SHC (SEQ ID NO: 1).

Производное SHC, обозначенное как SHC31, предложенное в настоящем изобретении, содержит следующую замену F129L, M132R и I432T по сравнению с эталонным белком SHC (SEQ ID NO: 1).

Производное SHC, обозначенное как SHC33, предложенное в настоящем изобретении, содержит следующую замену F601Y, F129L, M132R и I432T по сравнению с эталонным белком SHC (SEQ ID NO: 1).

Производное SHC, обозначенное как 101A10, предложенное в настоящем изобретении, содержит следующую замену F601Y и Q579H по сравнению с эталонным белком SHC (SEQ ID NO: 1).

Производное SHC, обозначенное как 111C8, предложенное в настоящем изобретении, содержит следующую замену T77A+I92V и F129L по сравнению с эталонным белком SHC (SEQ ID NO: 1).

В предпочтительном воплощении изобретения производное SHC содержит по меньшей мере замены F668Y или Y185R в комбинации с по меньшей мере любой одной или более чем одной из F182L и/или I498T относительно SEQ ID NO: 2.

Производное SHC, обозначенное как SHC3ZM1, предложенное в настоящем изобретении, содержит следующую замену F668Y по сравнению с эталонным белком SHC (SEQ ID NO: 2).

Hoshino и Sato (2002, цитируется выше) идентифицировали F601 как высококонсервативный аминокислотный остаток среди прокариотических и эукариотических биологических видов. Сообщалось, что производное SHC F601Y характеризуется значительно повышенной Vmax для субстрата оксидосквалена (но не сквалена). Однако F601Y характеризуется снижением аффинности (т.е. повышенной $K_M$) и снижением каталитической эффективности/активности ($K_{cat}/K_M$) относительно WT AacSHC при использовании сквалена. В работе Hoshino и Sato (2002, цитируется выше) не приведены данные по эффективности AacSHC при использовании гомофарнезола в качестве субстрата фермента с мутацией F601Y. Производное SHC, эквивалентное F601Y в ZmoSHC1, представляет собой F668Y.

Производное SHC, обозначенное как SHC10ZM1, предложенное в настоящем изобретении, содержит следующую замену F182L по сравнению с эталонным белком SHC (SEQ ID NO: 2).

Производное SHC, обозначенное как SHC30ZM1, предложенное в настоящем изобретении, содержит следующую замену F668Y и F182L по сравнению с эталонным белком SHC (SEQ ID NO: 2).

Производное SHC, обозначенное как SHC26ZM1, предложенное в настоящем

изобретении, содержит следующую замену Y185R и I498T по сравнению с эталонным белком SHC (SEQ ID NO: 2).

Производное SHC, обозначенное как 215G2ZM1, предложенное в настоящем изобретении, содержит следующую замену Y185R, I498T и G282V по сравнению с эталонным белком SHC (SEQ ID NO: 2).

Производное SHC, обозначенное как SHC32ZM1, предложенное в настоящем изобретении, содержит следующую замену F668Y, Y185R и I498T по сравнению с эталонным белком SHC (SEQ ID NO: 2).

Производное SHC, обозначенное как SHC31ZM1, предложенное в настоящем изобретении, содержит следующую замену F182L, Y185R и I498T по сравнению с эталонным белком SHC (SEQ ID NO: 2).

Производное SHC, обозначенное как SHC33ZM1, предложенное в настоящем изобретении, содержит следующую замену F668Y, F182L, Y185R и I498T по сравнению с эталонным белком SHC (SEQ ID NO: 2).

Производное SHC, обозначенное как 101A10ZM1, предложенное в настоящем изобретении, содержит следующую замену F668Y и H646H по сравнению с эталонным белком SHC (SEQ ID NO: 2).

Производное SHC, обозначенное как 111C8ZM1, предложенное в настоящем изобретении, содержит следующую замену S129A+V145V и F182L по сравнению с эталонным белком SHC (SEQ ID NO: 2).

В предпочтительном воплощении изобретения производное SHC содержит по меньшей мере замены F620Y или I140R в комбинации с по меньшей мере любой одной или более чем одной из F137L и/или I450T относительно SEQ ID NO: 3.

Производное SHC, обозначенное как SHC3ZM2, предложенное в настоящем изобретении, содержит следующую замену F620Y по сравнению с эталонным белком SHC (SEQ ID NO: 3).

Hoshino и Sato (2002, цитируется выше) идентифицировали F601 как высококонсервативный аминокислотный остаток среди прокариотических и эукариотических биологических видов. Сообщалось, что производное SHC F601Y характеризуется значительно повышенной Vmax для субстрата оксидосквалена (но не сквалена). Однако F601Y характеризуется снижением аффинности (т.е. повышенной $K_M$) и снижением каталитической эффективности/активности ($Kcat/K_M$) относительно WT AacSHC при использовании сквалена. В работе Hoshino и Sato (2002, цитируется выше) не приведены данные по эффективности AacSHC при использовании гомофарнезола в качестве субстрата фермента с мутацией F601Y. Производное SHC, эквивалентное F601Y в ZmoSHC2, представляет собой F620Y.

Производное SHC, обозначенное как SHC10ZM2, предложенное в настоящем изобретении, содержит следующую замену F137L по сравнению с эталонным белком SHC (SEQ Ш NO: 3).

Производное SHC, обозначенное как SHC30ZM2, предложенное в настоящем изобретении, содержит следующую замену F620Y и F137L по сравнению с эталонным белком SHC (SEQ ID NO: 3).

Производное SHC, обозначенное как SHC26ZM2, предложенное в настоящем изобретении, содержит следующую замену I140R и I450T по сравнению с эталонным белком SHC (SEQ ID NO: 3).

Производное SHC, обозначенное как 215G2ZM2, предложенное в настоящем изобретении, содержит следующую замену I140R, I450T и V233V по сравнению с эталонным белком SHC (SEQ ID NO: 3).

Производное SHC, обозначенное как SHC32ZM2, предложенное в настоящем изобретении, содержит следующую замену F620Y, I140R и I450T по сравнению с эталонным белком SHC (SEQ ID NO: 3).

Производное SHC, обозначенное как SHC31ZM2, предложенное в настоящем изобретении, содержит следующую замену F137L, I140R и I450T по сравнению с эталонным белком SHC (SEQ ID NO: 3).

Производное SHC, обозначенное как SHC33ZM2, предложенное в настоящем изобретении, содержит следующую замену F620Y, F137L, I140R и I450T по сравнению с эталонным белком SHC (SEQ ID NO: 3).

Производное SHC, обозначенное как 101A10ZM2, предложенное в настоящем изобретении, содержит следующую замену F620Y и N598H по сравнению с эталонным белком SHC (SEQ ID NO: 3).

Производное SHC, обозначенное как 111C8ZM2, предложенное в настоящем изобретении, содержит следующую замену G85A+V100V и F137L по сравнению с эталонным белком SHC (SEQ ID NO: 3).

В предпочтительном воплощении изобретения производное SHC содержит по меньшей мере замены F628Y или Y144R в комбинации с по меньшей мере любой одной или более чем одной из F141L и/или I459T относительно SEQ ID NO: 4.

Производное SHC, обозначенное как SHC3Bjp, предложенное в настоящем изобретении, содержит следующую замену F628Y по сравнению с эталонным белком SHC (SEQ ID NO: 4).

Hoshino и Sato (2002, цитируется выше) идентифицировали F601 как высококонсервативный аминокислотный остаток среди прокариотических и эукариотических биологических видов. Сообщалось, что производное SHC F601Y характеризуется значительно повышенной Vmax для субстрата оксидосквалена (но не сквалена). Однако F601Y характеризуется снижением аффинности (т.е. повышенной $K_M$) и снижением каталитической эффективности/активности ($Kcat/K_M$) относительно WT AacSHC при использовании сквалена. В работе Hoshino и Sato (2002, цитируется выше) не приведены данные по эффективности AacSHC при использовании гомофарнезола в качестве субстрата фермента с мутацией F601Y. Производное SHC, эквивалентное F601Y в BjpSHC, представляет собой F628Y.

Производное SHC, обозначенное как SHC10Bjp, предложенное в настоящем описании, содержит следующую замену F141L по сравнению с эталонным белком SHC (SEQ ID NO: 4).

Производное SHC, обозначенное как SHC30Bjp, предложенное в настоящем описании, содержит следующую замену F628Y и F141L по сравнению с эталонным белком SHC (SEQ ID NO: 4).

Производное SHC, обозначенное как SHC26Bjp, предложенное в настоящем описании, содержит следующую замену Y144R и I459T по сравнению с эталонным белком SHC (SEQ ID NO: 4).

Производное SHC, обозначенное как 215G2Bjp, предложенное в настоящем описании, содержит следующую замену Y144R, I459T и V241V по сравнению с эталонным белком SHC (SEQ ID NO: 4).

Производное SHC, обозначенное как SHC32Bjp, предложенное в настоящем описании, содержит следующую замену F628Y, Y144R и I459T по сравнению с эталонным белком SHC (SEQ ID NO: 4).

Производное SHC, обозначенное как SHC31Bjp, предложенное в настоящем описании, содержит следующую замену F141L, Y144R и I459T по сравнению с эталонным белком

SHC (SEQ ID NO: 4).

Производное SHC, обозначенное как SHC33Bjp, предложенное в настоящем описании, содержит следующую замену F628Y, F141L, Y144R и I459 по сравнению с эталонным белком SHC (SEQ ID NO: 4).

Производное SHC, обозначенное как 101A10Bjp, предложенное в настоящем описании, содержит следующую замену F628Y и M607H по сравнению с эталонным белком SHC (SEQ ID NO: 4).

Производное SHC, обозначенное как 111C8FBjp, предложенное в настоящем описании, содержит следующую замену A88A+V104V и F141L по сравнению с эталонным белком SHC (SEQ ID NO: 4).

Аминокислотные последовательности

В некоторых воплощениях изобретения производное AacSHC/HAC содержит один или более чем один полипептид, представленный в одной или более чем одной из SEQIDNO: 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33,35, 37, 39 и/или 171.

Предпочтительно производные AacSHC/HAC согласно настоящему описанию имеют аминокислотную последовательность, выбранную из группы, состоящей из 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 и/или SEQ ID NO: 171.

В других воплощениях изобретения производные ZmoSHC1/HAC содержат один или более чем один полипептид, представленный в одной или более чем одной из SEQ ID NO: 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75 и/или 173.

Предпочтительно производные ZmoSHC1/HAC согласно настоящему описанию имеют аминокислотную последовательность, выбранную из группы, состоящей из 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 и/или SEQ ID NO: 173.

В следующих воплощениях изобретения производные ZmoSHC2/HAC содержат один или более чем один полипептид, представленный в одной или более чем одной из 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 и/или SEQ ID NO: 175.

В дополнительных воплощениях изобретения производные BjpSHC/HAC содержат один или более чем один полипептид, представленный в одной или более чем одной из: 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 и/или SEQ ID NO: 177.

Выравнивание последовательностей

В связи с тем, что эталонные последовательности SHC, такие как, например, последовательности полипептидов AacSHC, ZmoSHC1, ZmoSHC2 и BjpSHC, имеют различную длину, аминокислотный остаток в положении X в эталонной последовательности AacSHC (SEQ ID NO: 1) соответствует другому положению аминокислоты B в эталонной последовательности ZmoSHC1 (SEQ ID NO: 2), другому положению аминокислоты J в эталонной последовательности ZmoSHC2 (SEQ ID NO: 3) и другому положению аминокислоты Z в эталонной последовательности BjpSHC (SEQ ID NO: 4). Кроме того, изменение эталонной последовательности SHC может также модифицировать последовательность производного SHC относительно этой эталонной последовательности SHC.

Термин «положение» относится к определенному остатку аминокислоты, находящемуся в эталонном белке SHC, идентифицированному определенной нумерацией аминокислот. Изменение эталонного белка SHC в результате инсерции или делеции аминокислоты приводит к различной нумерации между эталонной аминокислотной последовательностью SHC и аминокислотной последовательностью производного SHC. В качестве примера, в случае инсерции аминокислоты между аминокислотами 509 и 510 эталонного белка SHC, эта аминокислота после инсерции будет иметь номер 511 в белке-производном SHC, тогда как в эталонном белке SHC сохраняется номер 510.

Методы обнаружения активности WT SHC/HAC и производного SHC/HAC

Методы обнаружения и количественного определения активности WT SHC/HAC и/или производного фермента SHC/HAC описаны в настоящем описании и известны на уровне техники. В качестве примера активность WT SHC/HAC и/или производного фермента SHC/HAC можно определить путем инкубирования очищенного фермента SHC/HAC или экстрактов из клеток-хозяев, либо целостного рекомбинантного организма-хозяина, продуцирующего фермент SHC/HAC, с соответствующим субстратом в подходящих условиях и проведения анализа продуктов реакции (например, методом газовой хроматографии (GC) или анализа HPLC (высокоэффективной жидкостной хроматографией)). Дополнительные подробные сведения о количественном определении SHC/HAC и/или ферментативной активности SHC/HAC и анализе продуктов реакции приведены в примерах. Эти методы включают в себя продуцирование производного SHC в рекомбинантных клетках-хозяевах (например, E coli).

Используемый в настоящем описании термин «активность» означает способность фермента к взаимодействию с субстратом с получением целевого продукта. Активность может быть определена в тесте, известном как тест на активность, через увеличение количества целевого продукта, уменьшения количества субстрата (или исходных веществ) или через комбинации этих параметров в зависимости от времени. Производные SHC/HAC согласно настоящему описанию характеризуют по их способности к биоконверсии гомофарнезола в (-)-амброксид и демонстрируют биологическую активность, такую как активность HAC.

Используемый в настоящем описании термин «биологическая активность» относится к любой активности, которую может проявлять полипептид, включающей, без ограничений: ферментативную активность; активность связывания с другим соединением (например, связывание с другим полипептидом, в частности связывание с рецептором или связывание с нуклеиновой кислотой), ингибиторную активность (например, ингибиторную активность по отношению к ферменту), активирующую активность (например, активность, активирующую фермент) или токсические действия. Для варианта или производного не требуется, чтобы они проявляли такую активность в той же степени, что и родительский полипептид. Вариант рассматривают как вариант в контексте настоящей заявки, если он проявляет релевантную активность в степени, составляющей по меньшей мере 10% от активности родительского полипептида. Схожим образом производное рассматривают как вариант в контексте настоящей заявки, если он проявляет релевантную биологическую активность в степени, составляющей по меньшей мере 10% от активности родительского полипептида (поскольку термины производное и вариант на протяжении настоящего описания используют взаимозаменяемо).

В других воплощениях изобретения производные SHC/HAC согласно настоящему описанию демонстрируют лучший выход целевого продукта, чем эталонный белок SHC. Термин «выход целевого продукта» относится к количеству граммов извлекаемого

продукта на один грамм исходного сырья (которое можно рассчитать как скорость конверсии в молярных процентах).

В дополнительных воплощениях изобретения производные SHC/HAC согласно настоящему описанию проявляют модифицированную (например, повышенную) целевую производительность относительно эталонного белка SHC. Термин «целевая производительность» относится к количеству извлекаемого целевого продукта в граммах на литр ферментационной емкости в час времени биоконверсии (т.е. времени после добавления субстрата).

В следующих воплощениях изобретения производные SHC/HAC согласно настоящему описанию проявляют модифицированный коэффициент выхода целевого продукта, чем эталонный белок SHC. Термин «коэффициент выхода целевого продукта» относится к соотношению между концентрацией полученного продукта и концентрацией производного SHC (например, очищенного фермента SHC или экстракта из рекомбинантных клеток-хозяев, экспрессирующих фермент SHC) в реакционной смеси.

В различных воплощениях изобретения производные SHC согласно настоящему описанию демонстрируют модифицированную (например, повышенную) кратность повышения ферментативной активности (например, модифицированную/повышенную активность гомофарнезол-амброксид-циклазы (HAC)) относительно эталонного белка SHC (например, SEQ ID NO: 1, или SEQ ID NO: 2, или SEQ ID NO: 3, или SEQ ID NO: 4). Кратность этого повышения активности составляет по меньшей мере: 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 и/или 100.

Нуклеотидные последовательности

Настоящее изобретение также относится к изолированным молекулам нуклеиновой кислоты, содержащим нуклеотидную последовательность, кодирующую производное SHC, как описано в настоящем описании.

Используемый в настоящем описании термин «молекула нуклеиновой кислоты» следует конкретно относить к полинуклеотидам согласно описанию, которые могут представлять собой ДНК, кДНК, геномную ДНК, синтетическую ДНК или РНК и могут быть двунитевыми или однонитевыми, смысловой и/или антисмысловой нитью.

Термин «молекула нуклеиновой кислоты» следует конкретно применять к описанному (-ым) в настоящем описании полинуклеотиду(-ам), например, в виде полноразмерной нуклеотидной последовательности или ее фрагментов или частей, которые кодируют полипептид, обладающий ферментативной активностью, например фермент метаболического пути или его фрагменты или части, соответственно.

Этот термин также включает в себя отдельную молекулу, такую как кДНК, где соответствующая геномная ДНК имеет интроны и, следовательно, отличающуюся последовательность; геномный фрагмент, в котором отсутствует по меньшей мере один из фланкирующих генов; фрагмент кДНК или геномной ДНК, образующийся в результате полимеразной цепной реакции (ПЦР), и в котором отсутствует по меньшей мере один из фланкирующих генов; фрагмент рестрикции, в котором отсутствует по меньшей мере один из фланкирующих генов; ДНК, кодирующую неприродный белок, такой как слитый белок (например, гистидиновую метку), мутеин или фрагмент данного белка; и нуклеиновую кислоту, представляющую собой вырожденный вариант кДНК или встречающейся в природе нуклеиновой кислоты. Кроме того, он включает в себя рекомбинантную нуклеотидную последовательность, составляющую участок гибридного гена, т.е. гена, кодирующего неприродный слитый белок. В слитых белках к белку может быть присоединена одна или более чем одна аминокислота (такая как, без ограничений, гистидин (His)), обычно на N-конце белка, но также на C-конце белка или

в слитом виде внутри участков белка. Такие слитые белки или слитые векторы, кодирующие такие белки, как правило, предназначены для следующих целей: (i) повысить продуцирование рекомбинантных белков; (ii) повысить растворимость рекомбинантного белка; и (iii) способствовать очистке рекомбинантного белка путем обеспечения лиганда для аффинной очистки. Термин «молекула нуклеиновой кислоты» также включает в себя оптимизированные по кодонам последовательности, подходящие для экспрессии в конкретной клетке микроорганизма-хозяина (например, в клетке-хозяине E. coli). Используемый в настоящем описании термин «оптимизированный по кодонам» означает кодирующую белок последовательность нуклеиновой кислоты, адаптированную для экспрессии в прокариотической или эукариотической клетке-хозяине, в частности в бактериальных клетках-хозяевах, таких как клетки-хозяева E. coli, путем замены одного или более чем одного, или, предпочтительно, значительного числа кодонов кодонами, который чаще используются в генах бактериальных клеток-хозяев (например, E. coli). В связи с этим нуклеотидная последовательность, кодирующая эталонные последовательности SEQ ID NO: 1, 2, 3 и/или 4 и все их варианты/производные, может представлять собой исходную последовательность, обнаруживаемую в источнике (например, AacSHC, ZmoSHC1, ZmoSHC2 или BjpSHC, соответственно), или ген, который может быть оптимизирован по кодонам для выбранных организмов-хозяев, таких как, например, E. coli.

Молекула рибонуклеиновой кислоты (РНК) может быть получена путем транскрипции in vitro. Сегменты молекул ДНК также рассмотрены в пределах объема описания и могут быть получены, например, полимеразной цепной реакцией (ПЦР) или созданы путем обработки одной или более чем одной эндонуклеазой рестрикции. Сегменты молекулы нуклеиновой кислоты могут относиться к ДНК-фрагментам гена, в частности к фрагментам, представляющим собой частичные гены. Фрагмент может также содержать несколько открытых рамок считывания (ОРС), либо повторов одной и той же ОРС, либо различных ОРС. Этот термин следует конкретно относить к кодирующим нуклеотидным последовательностям, но в него также следует включать некодирующие нуклеотидные последовательности, например нетранскрибируемые или нетранслируемые последовательности, или последовательности, кодирующие полноразмерные или частичные полипептиды. Гены, используемые в настоящем описании, например, для сборки, диверсификации или рекомбинации, могут представлять собой либо некодирующие последовательности, либо кодирующие полипептиды последовательности, либо кодирующие белки последовательности или их участки или фрагменты, имеющие достаточную длину последовательности для успешного осуществления событий рекомбинации. Более конкретно, указанные гены имеют минимальную длину 3 пары оснований (п.о.), предпочтительно по меньшей мере 100 п.о., более предпочтительно по меньшей мере 300 п.о.

На основании вышеописанного будет ясно, что ссылка на изолированную ДНК не означает ДНК, находящуюся среди других молекул ДНК в количествах от сотен до миллионов, например, в составе библиотек кДНК, или геномных ДНК, или рестриктных фрагментов геномной ДНК, например в рестрикционной смеси или в пластинке геля на электрофорезе. Изолированная молекула нуклеиновой кислоты согласно настоящему описанию охватывает сегменты, которые сами по себе не обнаруживаются в естественном состоянии.

Используемый в настоящем описании термин «изолированная ДНК» может относиться к (1) ДНК, содержащей последовательность, не идентичную какой-либо встречающейся в природе последовательности, полинуклеотиду или нуклеиновой

кислоте, не встречающимся в природе (например, полученным путем искусственного комбинирования (например, искусственных манипуляций с изолированными сегментами нуклеиновых кислот, например, методами генной инженерии) двух обычно разделенных сегментов последовательностей посредством вмешательства человека) или (2) в контексте ДНК, встречающейся в природе последовательности (например, кДНК или геномной ДНК), ДНК, не содержащей по меньшей мере один из генов, фланкирующих ген, содержащий интересующую ДНК, в геноме организма, в котором в природе встречается ген, содержащий интересующую ДНК.

Используемый в настоящем описании термин «изолированная ДНК», особенно по отношению к последовательности нуклеиновой кислоты, может также относиться к нуклеиновым кислотам или полинуклеотидам, полученным методами рекомбинантных ДНК, например, к конструкции ДНК, содержащей полинуклеотид, гетерологичный для клетки-хозяина, который, возможно, встраивают в клетку-хозяина. Химерная нуклеотидная последовательность может быть, в частности, получена в виде рекомбинантной молекулы. Термин «рекомбинация» конкретно применяют к сборке полинуклеотидов, соединению таких полинуклеотидов или их участков вместе, с рекомбинацией или без рекомбинации, с получением кроссоверного или мозаичного гена. Например, ее осуществляют, чтобы соединить вместе сегменты нуклеиновой кислоты, обладающие требуемыми функциями, для создания требуемой комбинации функций. Рекомбинантный ген, кодирующий описанный в настоящем описании полипептид, включает в себя кодирующую последовательность для этого полипептида, функционально связанную в смысловой ориентации с одной или более регуляторных областей, подходящих для экспрессии этого полипептида. В связи с тем, что многие микроорганизмы способны экспрессировать множественные генные продукты с полицистронной мРНК, при желании множество полипептидов можно экспрессировать под контролем одной регуляторной области для этих микроорганизмов. Кодирующую последовательность и регуляторную область считают функционально связанными, если регуляторная область и кодирующая последовательность расположены так, что регуляторная область эффективно осуществляет регуляцию транскрипции или трансляции этой последовательности.

Используемый в настоящем описании термин «рекомбинантный», особенно по отношению к ферментам, следует относить к ферментам, полученным методами рекомбинантных ДНК, т.е. полученным из клеток, трансформированных экзогенной конструкцией ДНК, кодирующей требуемый фермент. Термин «синтетические ферменты» означает ферменты, полученные путем химического синтеза. Химерный фермент может быть конкретно получен в виде рекомбинантной молекулы. Таким образом, термин «рекомбинантная ДНК» включает в себя рекомбинантную ДНК, встроенную либо в вектор в виде автономно реплицирующейся плазмиды или вируса, либо в геномную ДНК прокариотического или эукариотического организма (или в геном гомологичной клетки в положении, отличающемся от природной хромосомной локализации).

В следующем аспекте молекула (-ы) нуклеиновой кислоты согласно настоящему описанию функционально связана (-ы) с контрольными последовательностями экспрессии, что дает возможность экспрессии в прокариотических и/или эукариотических клетках-хозяевах. Используемый в настоящем описании термин «функционально связанный» означает встраивание в генетическую конструкцию так, чтобы контрольные последовательности экспрессии эффективно контролировали экспрессию интересующей кодирующей последовательности.

Транскрипционные/трансляционные регуляторные элементы, относящиеся к

описанному выше, включают, без ограничений, индуцибельные и неиндуцибельные, конститутивные, регулируемые клеточным циклом, метаболически регулируемые промоторы, энхансеры, операторы, сайленсеры, репрессоры и другие элементы, известные специалистам в данной области техники, которые стимулируют или иным образом регулируют экспрессию гена. Такие регуляторные элементы включают, без ограничений, регуляторные элементы, направляющие конститутивную экспрессию или обеспечивающие возможность индуцибельной экспрессии, такие как, например, промотор

CUP-1, tet-репрессор, используемый, например, в системах tet-on или tet-off, регуляторные элементы lac-системы и trp-системы. В качестве примера, изопропил-β-D-1-тиогалактопиранозид (IPTG) является эффективным индуктором экспрессии генов в диапазоне концентраций от 100 мкМ до 1,0 мМ. Это соединение является молекулярным имитатором аллолактозы, являющейся метаболитом лактозы, который запускает транскрипцию lac оперона и, следовательно, используется для индукции экспрессии гена при нахождении гена под контролем lac оператора. Другим примером регуляторного элемента, который индуцирует экспрессию гена, является лактоза.

Аналогичным образом молекула (-ы) нуклеиновой кислоты согласно настоящему описанию может(гут) образовать участок гибридного гена, кодирующего дополнительные полипептидные последовательности, например последовательность, функционирующую в качестве маркера или репортера. Примеры маркерных и репортерных генов включают гены бета-лактамазы, хлорамфениколацетилтрансферазы (CAT), аденозиндезаминазы (ADA), аминогликозидфосфотрансферазы дигидрофолатредуктазы (DHFR), гигромицин-B-фосфотрансферазы (HPH), тимидинкиназы (TK), lacZ (кодирующий бета-галактозидазу) и ксантингуанинфосфорибозилтрансферазы (XGPRT). Как и для многих стандартных процедур, связанных с осуществлением на практике изобретения согласно описанию, специалистам в данной области техники известны дополнительные полезные реагенты, например, дополнительные последовательности, которые могут выполнять функцию маркера или репортера.

В некотором воплощении в настоящем описании предложен рекомбинантный полипептид, кодирующий WT SHC или его описанное выше производное, который может быть встроен в вектор для экспрессии и дополнительной очистки. Одним из типов вектора является плазмида, представляющая собой кольцевую двунитевую петлю ДНК, в которой лигируют дополнительные сегменты ДНК. Определенные векторы могут контролировать экспрессию генов, с которыми они функционально связаны. Эти векторы называют «экспрессионными векторами». Обычно экспрессионные векторы, подходящие для методов рекомбинантных ДНК, представляют собой векторы плазмидного типа. Как правило, экспрессионный вектор содержит ген, такой как WT SHC или вариант SHC/HAC, как описано в настоящем описании. В настоящем описании термины «плазмида» и «вектор» используют взаимозаменяемо, поскольку плазмида является наиболее часто используемым типом вектора.

Такие векторы могут включать последовательности ДНК, включающие, без ограничений, последовательности ДНК, которые в природе отсутствуют в клетке-хозяине, последовательности ДНК, которые в нормальных условиях не транскрибируются в РНК или не транслируются в белок (не «экспрессируются»), и другие гены или последовательности ДНК, требуемые для введения в нерекомбинантную клетку-хозяина. Должно быть понятно, что, как правило, геном описанной в настоящем описании рекомбинантной клетки-хозяина дополняют в результате стабильного введения

одного или более рекомбинантных генов. Тем не менее, в пределах объема настоящего описания можно также использовать репликативные плазмиды или векторы. Кроме того, настоящее изобретение можно осуществлять на практике с использованием низкокопийной, например однокопийной, или высококопийной плазмиды или вектора (пример, которых приведен в настоящем описании).

В предпочтительном воплощении изобретения вектор согласно настоящему описанию включает в себя плазмиды, фагмиды, фаги, космиды, искусственные бактериальные и искусственные дрожжевые хромосомы, конструкции типа нокаут или нок-ин, синтетические нуклеиново-кислотные последовательности или кассеты, и эти подгруппы могут быть получены в форме линейных полинуклеотидов, плазмид, мегаплазмид, синтетических или искусственных хромосом, таких как искусственные хромосомы растений, бактерий, млекопитающих или дрожжей.

Предпочтительно, чтобы белки, кодируемые встроенным полипептидом, после встраивания в вектор экспрессируются внутри клетки. В плазмиды могут быть встроены гены разнообразных субстратов. Плазмиды часто представляют собой стандартные векторы для клонирования, например бактериальные многокопийные плазмиды. Субстраты могут быть встроены в те же или другие плазмиды. Часто используют по меньшей мере два различных типа плазмид, имеющих различные типы селективных маркеров, что дает возможность отобрать клетки, содержащие по меньшей мере два типа векторов.

Как правило, бактериальные или дрожжевые клетки можно трансформировать любой одной или более чем одной из следующих нуклеотидных последовательностей, как хорошо известно в данной области техники. Для рекомбинации in vivo для трансформации клетки-хозяина используют ген, который должен рекомбинировать с геномом или с другими генами, используя стандартные методы трансформации. В соответствующем воплощении изобретения в конструкцию включают ДНК, обеспечивающую точку начала репликации. Точка начала репликации может быть целесообразно выбрана специалистом в данной области техники. В зависимости от природы генов в том случае, если в генах или в геноме уже присутствуют последовательности, которые могут сами по себе действовать в качестве точек начала репликации, дополнительная точка начала репликации может не требоваться.

Бактериальную или дрожжевую клетку можно трансформировать экзогенной или гетерологичной ДНК, если такая ДНК введена внутрь клетки. Трансформирующая ДНК может быть интегрирована, т.е. ковалентно связана с геномом клетки, или не интегрирована. У прокариот и дрожжей трансформирующая ДНК может сохраняться на эписомном элементе, таком как плазмида. Что касается эукариотических клеток, стабильно трансфицированной клеткой является клетка, в которой произошла интеграция трансфицированной ДНК в хромосому таким образом, что она наследуется дочерними клетками посредством репликации хромосомы. Эта стабильность демонстрируется на основании способности эукариотической клетки основывать клеточные линии или клоны, содержащие популяции дочерних клеток, содержащих трансформирующую ДНК.

Как правило, введенная ДНК исходно не является резидентной для хозяина, который является реципиентом ДНК, но в объем настоящего изобретения входит выделение сегмента ДНК из данного хозяина и последующее введение одной или более дополнительных копий этой ДНК в ту же клетку-хозяина, например, чтобы усилить продуцирование продукта гена или изменить паттерн экспрессии гена. В некоторых случаях введенная ДНК приведет к модификации или даже замене эндогенного гена

или последовательности ДНК посредством, например, гомологической рекомбинации или сайт-направленного мутагенеза. Подходящие рекомбинантные клетки-хозяева включают микроорганизмы, растительные клетки и растения.

Признаками настоящего изобретения также являются рекомбинантные клетки-хозяева. Термином «рекомбинантная клетка-хозяин», которую также называют «генетически модифицированной клеткой-хозяином» или «трансгенной клеткой», обозначают клетку-хозяина, содержащую гетерологичную нуклеиновую кислоту, или клетку, геном которой дополнен по меньшей мере одной встроенной последовательностью ДНК. Клетка-хозяин по настоящему изобретению может быть сконструирована методами генной инженерии с полинуклеотидом или вектором, как описано выше.

Клетки-хозяева, которые можно использовать для целей изобретения, включают без ограничений прокариотические клетки, такие как бактерии (например, E. coli и B. subtilis), которые могут быть трансформированы, например, рекомбинантной ДНК бактериофага, плазмидной ДНК, бактериальной искусственной хромосомой или космидными векторами экспрессии ДНК, содержащими полинуклеотидные молекулы по изобретению; простые эукариотические клетки, такие как дрожжи (например, Saccharomyces и Pichia), которые могут быть трансформированы рекомбинантными дрожжевыми экспрессионными векторами, содержащими полинуклеотидную молекулу по изобретению. В зависимости от клетки-хозяина и соответствующего вектора, используемого для введения полинуклеотида по изобретению, полинуклеотид может быть интегрирован, например, в хромосому или митохондриальную ДНК, либо может поддерживаться экстрахромосомально, например эписомально, либо может только временно содержаться в клетках.

Используемый в настоящем описании термин «клетка», в частности, со ссылкой на генетическое конструирование и включение в клетку одного или более генов или собранного кластера генов или со ссылкой на клетку-продуцент понимают как относящийся к любой прокариотической или эукариотической клетке. Как прокариотические, так и эукариотические клетки-хозяева рассматривают для применения в соответствии с изобретением, включая бактериальные клетки-хозяева, такие как E. coli или Bacillus sp, дрожжевые клетки-хозяева, такие как S. cerevisiae, клетки-хозяева насекомых, таких как Spodoptora frugiperda, или человеческие клетки-хозяева, такие как клетки HeLa и Jurkat.

Более конкретно, клетка представляет собой эукариотическую клетку, предпочтительно клетку грибов, млекопитающих или растений, или прокариотическую клетку. Подходящие эукариотические клетки включают, например, без ограничений, клетки млекопитающих, дрожжевые клетки или клетки насекомых (включая Sf9), клетки амфибий (включая клетки-меланофоры) или клетки червей, включая клетки Caenorhabditis (включая Caenorhabditis elegans). Подходящие клетки млекопитающих включают, например, без ограничений клетки COS (включая Cos-1 и Cos-7), клетки CHO, клетки HEK293, клетки HEK293T, клетки HEK293 T-RexTM или другие способные к трансфекции эукариотические клеточные линии. Подходящие клетки бактерий включают, без ограничений, E. coli.

Предпочтительно можно использовать прокариоты, такие как E. coli, Bacillus, Streptomyces, или клетки млекопитающих, такие как клетки HeLa или клетки Jurkat, или клетки растений, такие как Arabidopsis.

Предпочтительно клетка представляет собой клетку Aspergillus sp. или клетку грибов, предпочтительно она может быть выбрана из группы, состоящей из родов Saccharomyces,

Candida, Kluyveromyces, Hansenula, Schizosaccharomyces, Yarrowia, Pichia и Aspergillus.

Предпочтительно клетка-хозяин E. coli представляет собой клетку-хозяина E. coli, признанную в промышленности и регулирующими органами (включая, без ограничений, клетку-хозяина E. coli K12 или, как продемонстрировано в примерах, клетку-хозяина E. coli BL21).

Одна из предпочтительных клеток-хозяев для использования в настоящем изобретении представляет собой клетку E. coli, которая может быть получена рекомбинантными методами, как описано в настоящем описании. Таким образом, рекомбинантная клетка-хозяин может представлять собой рекомбинантную клетку-хозяина E. coli. Для E. coli существуют библиотеки мутантов, плазмид, подробных компьютерных моделей метаболизма и другая информация, позволяющая осуществлять рациональный дизайн различных модулей для повышения выхода продукта. Для получения рекомбинантных микроорганизмов E. coli можно использовать способы, аналогичные описанным выше для Saccharomyces.

В одном воплощении изобретения рекомбинантный микроорганизм E. coli содержит нуклеотидные последовательности, кодирующие гены SHC (как раскрыто, например, в любой одной или более чем одной из таблиц 10, 11 и 12 в настоящем описании, или их функциональные эквиваленты/гомологи, включая, без ограничений, их варианты, мутанты гомологов, производные или фрагменты).

Предпочтительно рекомбинантный микроорганизм E. coli содержит векторную конструкцию, представленную на Фиг. 5 и 21.

В другом предпочтительном воплощении изобретения рекомбинантный микроорганизм E. coli содержит нуклеотидные последовательности, кодирующие WT SHC/HAC и производные WT SHC/HAC или их функциональные эквиваленты/гомологи, включая без ограничений их варианты, мутанты гомологов, производные или фрагменты, представленные в любой одной или более чем одной из таблицы 13, таблицы 14, таблицы 15, таблицы 16, таблицы 17 и/или таблицы 4a.

Другой предпочтительной клеткой-хозяином для использования в настоящем изобретении является S. cerevisiae, который является широко распространенным организмом-шасси в синтетической биологии. Таким образом, рекомбинантная клетка-хозяин может представлять собой S. cerevisiae. Для S. cerevisiae существуют библиотеки мутантов, плазмид, подробных компьютерных моделей метаболизма и другая информация, позволяющая осуществлять рациональный дизайн различных модулей для повышения выхода продукта. Известны способы создания рекомбинантных микроорганизмов S. cerevisiae.

Культивирование клеток осуществляют традиционным способом. Питательная среда для культивирования содержит источник углерода, по меньшей мере один источник азота и неорганические соли, а также к ней добавляют витамины. Компоненты этой питательной среды могут представлять собой компоненты, традиционно используемые для культивирования рассматриваемого вида микроорганизмов.

Источники углерода, используемые в настоящем способе, включают любую молекулу, которая может метаболизироваться рекомбинантной клеткой-хозяином, содействуя росту и/или продуцированию (-)-амброксида. Примеры приемлемых источников углерода включают, без ограничений, сахарозу (например, находящуюся в мелассах), фруктозу, ксилозу, глицерин, глюкозу, целлюлозу, крахмал, целлобиозу или другой содержащий глюкозу полимер.

В воплощениях изобретения, в которых используют дрожжи в качестве хозяина, приемлемы, например, источники углерода, такие как сахароза, фруктоза, ксилоза,

этанол, глицерин и глюкоза. Организм-хозяин может обеспечиваться источником углерода на протяжении всего периода культивирования или, альтернативно, организм можно выращивать в течение некоторого периода времени в присутствии другого источника энергии, например белка, а затем обеспечивать его источником углерода только во время фазы подпитки.

Приемлемость рекомбинантного микроорганизма-хозяина для использования в способах по настоящему изобретению можно определить с помощью простых методик тестирования, используя хорошо известные методы. Например, тестируемый микроорганизм можно размножать в богатой питательной среде (например, в среде LB, среде с бакто-триптоном и дрожжевым экстрактом, питательной среде и т.п.) в условиях pH, температуры и аэрации, обычно используемых для размножения этого микроорганизма. После отбора рекомбинантных микроорганизмов (т.е. рекомбинантных клеток-хозяев), продуцирующих желаемые продукты биоконверсии, эти продукты, как правило, получают с использованием линии-продуцента клеток-хозяев в крупном масштабе с помощью подходящих экспрессионных систем и ферментации, например, путем продуцирования микроорганизмами в клеточной культуре.

В одном воплощении настоящего изобретения для культивирования клеток используют определенную минимальную питательную среду, такую как M9A.

Компоненты питательной среды M9A включают: 14 г/л $KH_2PO_4$, 16 г/л $K_2HPO_4$, 1 г/л $Na_3$цитрат$\cdot 2H_2O$, 7,5 г/л $(NH_4)_2SO_4$, 0,25 г/л $MgSO_4\cdot 7H_2O$, 0,015 г/л $CaCl_2\cdot 2H_2O$, 5 г/л глюкозы и 1,25 г/л дрожжевого экстракта).

В другом воплощении настоящего изобретения использовали питательную среду, обогащенную питательными веществами, такую как среду LB. Компоненты питательной среды LB включают: 10 г/л триптона, 5 г/л дрожжевого экстракта, 5 г/л NaCl.

Другие примеры минеральной питательной среды и минеральная питательная среда M9 раскрыты, например, в патенте США US 6524831 B2 и в заявке на патент США US 2003/0092143 A1.

Рекомбинантный микроорганизм можно выращивать в периодической культуре, периодической культуре с подпиткой или в непрерывном процессе, либо использовать их комбинацию. Как правило, рекомбинантный микроорганизм выращивают в ферментере при определенной(-ых) температуре(-ах) в присутствии приемлемого источника питательных веществ, например источника углерода, в течение требуемого периода времени для получения достаточного количества фермента для биоконверсии гомофарнезола в амброксид и для получения желаемого количества амброксида, включая (-)-амброксид.

Рекомбинантные клетки-хозяева можно культивировать любым приемлемым способом, например, путем периодического культивирования или периодического культивирования с подпиткой.

Используемый в настоящем описании термин «периодическое культивирование» представляет собой способ культивирования, при котором не добавляют питательную среду для культивирования и не удаляют ее в процессе культивирования.

Используемый в настоящем описании термин «периодическое культивирование с подпиткой» представляет собой способ культивирования, при котором питательную среду для культивирования добавляют в процессе культивирования, но среду для культивирования не удаляют.

В одном воплощении настоящего изобретения предложен способ получения амброксида в клеточной системе, включающий экспрессию WT SHC или производных

SHC/HAC в приемлемых условиях в клеточной системе, подачу в клеточную систему гомофарнезола, конвертирование гомофарнезола в амброксид с использованием WT SHC или производных SHC/HAC, полученных с использованием клеточной системы и, возможно, выделение вещества (-)-амброксида из системы. Для улучшения этого способа можно использовать экспрессию других нуклеотидных последовательностей. Способ биоконверсии может включать дополнительную экспрессию других нуклеотидных последовательностей в клеточной системе. Экспрессия других нуклеотидных последовательностей может усиливать биохимический путь биоконверсии для получения (-)-амброксида.

Следующее воплощение настоящего изобретения представляет собой способ биоконверсии для получения (-)-амброксида, включающий выращивание клеток-хозяев, содержащих гены WT SHC/HAC или производных SHC/HAC, продуцирование WT SHC/HAC или производных SHC/HAC в клетках-хозяевах, подачу гомофарнезола (например, EEH) к клеткам-хозяевам, инкубацию клеток-хозяев в условиях pH, температуры и в присутствии солюбилизирующего агента, подходящих для стимуляции конвертирования гомофарнезола в амброксид и сбор (-)-амброксида. Продуцирование WT SHC/HAC и/или производных SHC/HAC в клетках-хозяевах обеспечивает способ получения (-)-амброксида при добавлении гомофарнезола к клеткам-хозяевам при подходящих реакционных условиях. Достижение конвертирования можно усилить путем добавления к реакционной смеси дополнительного количества биокатализатора и додецилсульфата натрия (SDS).

Рекомбинантную клетку-хозяина микроорганизма можно культивировать различными путями с целью обеспечения клеток, экспрессирующих WT SHC или производные SHC/HAC в приемлемых количествах для последующей стадии биоконверсии. Поскольку микроорганизмы, применяемые для стадии биоконверсии, широко варьируют (например, дрожжи, бактерии и грибы), условия культивирования, которые хорошо известны и документированы, несомненно, регулируют в соответствии со специфичными потребностями каждого биологического вида. Для получения клеток, полезных для использования на последующей стадии биоконверсии по настоящему изобретению, можно использовать любой из известных в данной области техники способов выращивания клеток рекомбинантного микроорганизма-хозяина. Как правило, клетки выращивают до определенной плотности (измеряемой как оптическая плотность (OD)), чтобы получить достаточное количество биомассы для реакции биоконверсии.

На количество получаемых клеток (биомассу) влияют не только выбранные условия культивирования, но также качество условий культивирования, влияющее на превращение биомассы в биокатализатор. Рекомбинантную клетку-хозяина микроорганизма, экспрессирующую ген WT SHC или производного SHC/HAC и продуцирующую ферменты WT SHC или производных SHC/HAC, называют биокатализатором, приемлемым для использования в реакции биоконверсии. В некоторых воплощениях изобретения биокатализатор представляет собой рекомбинантную целую клетку, продуцирующую WT SHC или производные SHC/HAC, либо он может находиться в суспензии или в иммобилизованном формате. В других воплощениях изобретения биокатализатор представляет собой мембранную фракцию или жидкую фракцию, приготовленную из рекомбинантных целых клеток, продуцирующих WT SHC или производное SHC/HAC (как раскрыто, например, в публикации Seitz et al 2012, цитируемой выше).

Рекомбинантная целая клетка, продуцирующая WT SHC или производные SHC/HAC, включает целые клетки, собранные из ферментера (для реакции биоконверсии), или

клетки в ферментере (которые впоследствии используют в реакции, протекающей в одном реакционном сосуде). Рекомбинантная целая клетка, продуцирующая WT SHC или производные SHC/HAC, может включать интактную рекомбинантную целую клетку и/или клеточный дебрис. В любом случае WT SHC или производное SHC/HAC некоторым образом ассоциированы с мембраной (такой как клеточная мембрана) с целью приема субстрата (например, гомофарнезола) и/или взаимодействия с ним, где мембрана (такая как клеточная мембрана) может составлять часть или содержать целую клетку (например, рекомбинантную целую клетку). WT SHC или производные SHC/HAC могут также находиться в иммобилизованной форме (например, ассоциированы с носителем фермента), которая обеспечивает возможность взаимодействия WT SHC или производных SHC/HAC с субстратом (например, гомофарнезолом). WT SHC или производные SHC/HAC можно также использовать в растворимой форме.

В одном воплощении изобретения биокатализатор получают в достаточных количествах (для создания достаточной биомассы), собирают и промывают (и, возможно, хранят (например, в замороженном или лиофилизированном виде)) до стадии биоконверсии.

В следующем воплощении изобретения клетки получают в достаточных количествах (для создания достаточной биомассы), а затем регулируют реакционные условия без необходимости в сборе и промывании биокатализатора для реакции биоконверсии. Этот одностадийный способ (или способ реакции, протекающей в одном реакционном сосуде) обладает преимуществом, поскольку он упрощает процесс, при этом снижая затраты. Питательная среда для культивирования, используемая для выращивания клеток, также приемлема для использования в реакции биоконверсии при условии регулирования реакционных условий таким образом, чтобы способствовать реакции биоконверсии.

Оптимум pH для выращивания клеток находится в диапазоне 6,0-7,0. Оптимум pH для реакции биоконверсии зависит от типа используемого фермента SHC/HAC в реакции биоконверсии. pH регулируют, используя методы, хорошо известные специалистам в данной области техники.

Как показано в примере 9, для биоконверсии гомофарнезола в (-)-амброксид использовали способ «реакции, протекающей в одном реакционном сосуде», с коэффициентом конверсии 100%. Как показано в примере 18, для биоконверсии гомофарнезола в (-)-амброксид использовали способ «реакции, протекающей в одном реакционном сосуде», с коэффициентом конверсии 99%.

При использовании в настоящем описании любая ссылка на коэффициент конверсии 99%/100% для субстрата гомофарнезола в (-)-амброксид представляет собой ссылку на 99%/100% коэффициент конверсии 99%/100% для изомера гомофарнезола (т.е. EEH), способного к конверсии в (-)-амброксид с использованием фермента, представляющего собой WT SHC/HAC или производное SHC/HAC.

Хотя в настоящем описании термины «смесь» или «реакционная смесь» можно использовать взаимозаменяемо с термином «питательная среда» (особенно по отношению к «реакции, протекающей в одном реакционном сосуде»), следует отметить, что для выращивания клеток для создания достаточного количества биомассы требуется питательная/ферментационная среда для культивирования клеток, а для стадии биоконверсии питательная среда не требуется, поскольку для реакции достаточно буферного раствора при подходящем pH.

Способы биоконверсии по настоящему изобретению осуществляют в условиях времени, температуры, pH и в присутствии солюбилизирующего агента, что обеспечивает

превращение исходного сырья гомофарнезола в (-)-амброксид. pH реакционной смеси может находиться в диапазоне 4-8, предпочтительно от 5 до 6,5, более предпочтительно 4,8-6,0 для производных фермента SHC/HAC и в диапазоне для pH от приблизительно 5,0 до pH приблизительно 7,0 для фермента WT SHC, и его можно поддерживать путем добавления к реакционной смеси буферных растворов. Примером буферного раствора для этой цели является буферный раствор лимонной кислоты. Предпочтительная температура составляет от приблизительно 15°C до приблизительно 45°C, предпочтительно от приблизительно 20°C до приблизительно 40°C, хотя для термофильных организмов она может быть выше, вплоть до 55°C, особенно при использовании фермента WT (например, WT SHC/HAC) из термофильного микроорганизма. Температуру можно поддерживать постоянной или изменять в процессе биоконверсии.

Автор изобретения продемонстрировал, что может быть полезным включение в реакцию биоконверсии солюбилизирующего агента (например, поверхностно-активного вещества (ПАВ), детергента, усилителя растворимости, смешиваемого с водой органического растворителя и т.п.). Используемый в настоящем описании термин «поверхностно-активное вещество» означает компонент, который снижает поверхностное натяжение (или межфазное натяжение) между двумя жидкостями или между жидкостью и твердым веществом. Поверхностно-активные вещества могут действовать как детергенты, смачивающие вещества, эмульгаторы, пенообразующие средства и диспергирующие средства. Примеры поверхностно-активных веществ включают, без ограничений, Тритон X-100, Твин 80, тауродезоксихолат, тауродезоксихолат натрия, додецилсульфат натрия (SDS) и/или лаурилсульфат натрия (SLS).

Хотя Тритон X-100 можно использовать для частичной очистки фермента WT SHC/HAC или производного SHC/HAC (в растворимой или мембранной фракции/суспензионной форме), его также можно использовать в реакции биоконверсии (см., например, описание Seitz (2012, диссертация PhD, цитируется выше), а также описание в публикации Neumann и Simon (1986, цитируется выше) и JP 2009060799.

Однако неожиданно, как показано в примере 14, автор изобретения выбрал и идентифицировал SDS как особенно предпочтительное солюбилизирующее средство из большого перечня других солюбилизирующих средств, используемых реже. В частности, автор изобретения идентифицировал SDS как значительно лучшее солюбилизирующее средство, чем, например, Тритон X-100, в отношении скорости реакции и выхода реакции биоконверсии гомофарнезола в (-)-амброксид (при использовании EEH как при 4 г/л, так и при 125 г/л). Как показывают сравнительные данные в примере 12, автор изобретения продемонстрировал, что максимальная активность по меньшей мере одного производного SHC/HAC для биоконверсии гомофарнезола в (-)-амброксид с Тритоном X-100 (в диапазоне концентрации от около 0,005% до 0,48%) в реакции составляла лишь около 20% активности, полученной с SDS (при концентрации около 0,07%).

Не желая ограничиваться теорией, считают, что применение SDS с рекомбинантными клетками-хозяевами микроорганизма может обладать преимуществом в связи с тем, что SDS может взаимодействовать преимущественно с мембраной клетки-хозяина, что делает фермент SHC (который является мембраносвязанным ферментом) более доступным для субстрата гомофарнезола. Кроме того, включение SDS в реакционную смесь в подходящей концентрации может улучшить свойства эмульсии (гомофарнезол в воде) и/или улучшить доступ субстрата гомофарнезола для фермента SHC внутри

клетки-хозяина, в то же время предотвращая разрушение, например денатурацию фермента SHC (WT или производного SHC/HAC).

На концентрацию солюбилизирующего агента (например, SDS), используемого в реакции биоконверсии, влияет количество биомассы и концентрация субстрата (EEH). Таким образом, существует некоторая степень взаимозависимости между концентрацией солюбилизирующего агента (например, SDS), количеством биомассы и концентрацией субстрата (EEH). В качестве примера, при повышении концентрации субстрата гомофарнезола для эффективного протекания реакции биоконверсии требуются достаточные количества биокатализатора и солюбилизирующего агента (например, SDS). Например, в случае слишком низкой концентрации солюбилизирующего агента (например, SDS) может наблюдаться субоптимальная конверсия гомофарнезола. С другом стороны, например, при слишком высокой концентрации солюбилизирующего агента (например, SDS) может быть риск повреждения биокатализатора за счет разрушения интактной клетки микроорганизма и/или денатурации/инактивации фермента SHC/HAC.

Выбор приемлемой концентрации SDS в контексте количества биомассы и концентрации субстрата (EEH) находится в пределах компетенции специалистов в данной области техники. В качестве примера, специалистам в данной области техники доступна прогностическая модель определения приемлемой концентраций SDS, субстрата (EEH) и биомассы. В качестве дополнительного примера, в примере 3 показано, что концентрация SDS в диапазоне 0,010-0,075% является подходящей при использовании 4 г/л EEH и биокатализатора при OD 10,0 (650 нм). В примере 7 показано, что при использовании 125 г/л EEH с 2-кратной массой сырой биомассы приемлема скорректированная концентрация SDS (1,55%). Тем не менее исследование процентной доли конверсии EEH в (-)-амброксид с использованием различных значений отношения SDS/клетки показало, что правильный выбор соотношения биокатализатора, субстрата гомофарнезола и солюбилизирующего агента (например, SDS) способствует разработке надежной системы реакции биоконверсии, которая демонстрирует определенную степень толерантности к диапазону концентраций SDS (см., например, Фиг. 17) и диапазонам pH (см. пример 15, Фиг. 18).

Температура реакции биоконверсии для фермента WT SHC (например, AacSHC) составляет приблизительно 45-60°C, предпочтительно 55°C.

Диапазон pH реакции биоконверсии для фермента WT SHC (например, AacSHC) составляет от приблизительно 5,0 до 7,0, более предпочтительно от приблизительно 5,6 до приблизительно 6,2, даже более предпочтительно приблизительно 6,0.

Температура реакции биоконверсии для производного фермента SHC/HAC составляет от приблизительно 34°C до приблизительно 50°C, предпочтительно приблизительно 35°C.

pH реакции биоконверсии для производного фермента SHC/HAC составляет приблизительно 4,8-6,4, предпочтительно приблизительно 5,2-6,0.

Предпочтительно солюбилизирующий агент, используемый в реакции биоконверсии, представляет собой SDS.

Концентрация SDS, используемая в реакции биоконверсии для фермента WT SHC (например, AacSHC), находится в диапазоне приблизительно 0,010-0,075%, предпочтительно приблизительно 0,030% при использовании приблизительно 4 г/л EEH.

Концентрация SDS, используемая в реакции биоконверсии для производного фермента SHC/HAC, находится в диапазоне приблизительно 0,0025-0,090%, предпочтительно

приблизительно 0,050% при использовании приблизительно 4 г/л ЕЕН.

При загрузке в реакционную смесь гомофарнезола при концентрации ЕЕН приблизительно 4 г/л ЕЕН биокатализатор загружают в реакционную смесь при OD 10,0 (650 нм)

Отношение [SDS]/[клетки] находится в диапазоне приблизительно 10:1-20:1, предпочтительно приблизительно 15:1-18:1, предпочтительно приблизительно 16:1, когда отношение количества биокатализатора к количеству гомофарнезола ЕЕН составляет приблизительно 2:1

Концентрация SDS в реакции биоконверсии для варианта фермента SHC находится в диапазоне приблизительно 1-2%, предпочтительно в диапазоне приблизительно 1,4-1,7%, еще более предпочтительно приблизительно 1,5% при концентрации гомофарнезола, составляющей приблизительно 125 г/л ЕЕН, и концентрации биокатализатор а, составляющей 250 г/л (соответствующей OD приблизительно 175 (650 нм)).

Отношение биокатализатора к субстрату гомофарнезолу ЕЕН находится в диапазоне приблизительно 0,5:1-2:1, в некоторых воплощениях изобретения 2:1, предпочтительно приблизительно 1:1 или 0,5:1.

В некоторых воплощениях изобретения амброксид получают с помощью биокатализатора, к которому добавляют субстрат гомофарнезол. Субстрат можно добавлять путем подачи с использованием известных средств (например, перистальтического насоса, инфузионного шприца и т.п.). Гомофарнезол является маслорастворимым соединением и предоставляется в масляной форме. С учетом того, что биокатализатор (клетки микроорганизмов, такие как интактная рекомбинантная целая клетка и/или клеточный дебрис и/или иммобилизованный фермент) находится в водной фазе, при добавлении гомофарнезола к реакционной смеси биоконверсии реакцию биоконверсии можно рассматривать как трехфазную систему (содержащую водную фазу, твердую фазу и масляную фазу). Это происходит даже в случае присутствия SDS. В качестве пояснения, при использовании в качестве биокатализатора растворимого WT SHC или SHC/HAC эту систему рассматривают как двухфазную.

Число присутствующих изомеров гомофарнезола может влиять на скорость реакции. Как показано в примере 11, производное фермента SHC/HAC способно к биоконверсии E,E-гомофарнезола в (-)-амброксид из сложной смеси изомеров гомофарнезола (например, EE:EZ:ZE:ZZ). Тем не менее, как правило, наблюдается снижение коэффициента конверсии, что согласуется с точкой зрения, что изомеры гомофарнезола, отличающиеся от ЕЕН, могут конкурировать с ЕЕН за доступ к производным фермента SHC/HAC и, следовательно, могут действовать как конкурентные ингибиторы конверсии ЕЕН в (-)-амброксид и/или также действовать в качестве альтернативных субстратов.

Соответственно, предпочтительно субстрат гомофарнезола содержит стереоизомерную смесь из 2-4 изомеров, предпочтительно из двух изомеров.

Соответственно, предпочтительно субстрат гомофарнезола состоит или по существу состоит из стереоизомерной смеси из 2-4 изомеров, предпочтительно из двух изомеров.

Предпочтительно субстрат гомофарнезола содержит стереоизомерную смесь EE:EZ.

Предпочтительно субстрат гомофарнезола состоит или по существу состоит из стереоизомерной смеси EE:EZ.

Как показано в примере 9, 100% конверсию смеси EE:EZ в массовом отношении 87:13 наблюдали в реакции ферментации и биоконверсии, «протекающей в одном реакционном сосуде», которую проводят в течение периода 22,5 дня. За этот период времени приблизительно 10 г ЕЕН претерпевало превращение.

Как подробно описано в примере 7, в предпочтительном воплощении изобретения используют ферментер для выращивания рекомбинантной клетки-хозяина, экспрессирующей ген производного SHC/HAC и продуцирующей активные производные ферментов SHC/HAC, до достаточной концентрации биомассы, приемлемой для использования в качестве биокатализатора, который используют в качестве биокатализатора в том же ферментере, который используют для конверсии источника гомофарнезола в (-)-амброксид в смеси с одним или более чем одним из побочных продуктов (II), (IV) и/или III (как показано, например, на Фиг. 12). Полученный в результате (-)-амброксид можно выделить путем экстракции/дистилляции паром или экстракции органическим растворителем, используя растворитель, не смешиваемый с водой (чтобы отделить продукты реакции и не прореагировавший субстрат от биокатализатора, который остается в водной фазе) с последующим выпариванием растворителя с получением неочищенного продукта реакции, который определяют с помощью анализа методом газовой хроматографии (GC). Способы экстракции/дистилляции паром и экстракции органическим растворителем известны специалистам в данной области техники.

В качестве примера, полученный в результате (-)-амброксид можно экстрагировать из всей реакционной смеси, используя органический растворитель, такой как не смешиваемый с водой растворитель (например, толуол). Альтернативно, полученный в результате (-)-амброксид можно выделить из твердой фазы реакционной смеси (полученной, например, центрифугированием или фильтрованием), используя смешиваемый с водой растворитель (например, этанол) или не смешиваемый с водой растворитель (например, толуол). В качестве следующего примера, (-)-амброксид присутствует в твердой фазе в виде кристаллов или в аморфной форме, и его можно отделить от остальной твердой фазы (клеточного материала или его дебриса) и жидкой фазы также с помощью фильтрования. В качестве следующего примера, при температуре выше температуры плавления (-)-амброксида (около 75°C) (-)-амброксид может образовать масляный слой на поверхности водной фазы, при этом масляный слой можно удалить и собрать. Чтобы обеспечить полное извлечение (-)-амброксида после удаления масляного слоя к водной фазе, содержащей биомассу, для экстракции любого остаточного (-)-амброксида, содержащегося в биомассе или вокруг биомассы, можно добавлять органический растворитель. Органический слой можно объединять с масляным слоем, после чего все вместе подвергают дополнительной обработке для выделения и очистки (-)-амброксида.

(-)-Амброксид можно дополнительно подвергать селективной кристаллизации для удаления побочных продуктов (II), (IV) и (III) и любого непрореагировавшего субстрата гомофарнезола из конечного продукта (-)-амброксида. Термин «селективная кристаллизация» относится к стадии процесса, посредством которого вызывают кристаллизацию (-)-амброксида из растворителя, при этом соединения (II), (III) и (IV) остаются в растворе кристаллизационного растворителя в такой степени, чтобы изолированный кристаллический материал содержал только продукт (-)-амброксид, или в случае содержания в нем каких-либо других соединений (II), (III) или (IV) они присутствовали только в ольфакторно приемлемых количествах.

На стадии селективной кристаллизации можно использовать смешиваемый с водой растворитель, такой как этанол или подобный растворитель. Ольфакторную чистоту конечного продукта (-)-амброксида определяют, используя 10% этанольный экстракт в воде, или путем тестирования кристаллического вещества. Конечный продукт (-)-амброксид тестируют в сравнении с имеющимся в продаже эталонным продуктом (-)-

амброксидом в отношении его ольфакторной чистоты, качества и сенсорного профиля. Также проводят тестирование материала (-)-амброксида экспертами в исследованиях применения с целью определения соответствия материала спецификациям в отношении органолептического профиля. Различные применения (-)-амброксида включают, без ограничений, высококачественные ароматизирующие или потребительские продукты, такие как продукты по уходу за тканями, гигиено-косметические продукты, продукты для заботы о красоте и чистящие продукты, включая по существу все продукты, где используют доступные в настоящее время коммерчески применяемые ингредиенты, содержащие амброксид, включающие без ограничений: Ambrox (Firmenich), Ambroxan (Henkel), Ambrofix (Givaudan), Amberlyn (Quest), Cetalox Laevo (Firmenich), Ambermor (Aromor) и/или Norambrebolide Ether (Pacific).

На селективную кристаллизацию (-)-амброксида может влиять присутствие непрореагировавшего субстрата гомофарнезола, а также соотношение количеств (-)-амброксида и других обнаружимых побочных продуктов (II), (III) и/или (IV). Селективная кристаллизация все еще возможна даже в случае получения только 10% конверсии субстрата гомофарнезола в (-)-амброксид (как показано в примере 7 с использованием фермента WT SHC/HAC).

Примеры приемлемых смешиваемых с водой и не смешиваемых с водой органических растворителей, приемлемых для использования в экстракции и/или селективной кристаллизации (-)-амброксида, включают, без ограничений, алифатические углеводороды, предпочтительно имеющие от 5 до 8 атомов углерода, такие как пентан, циклопентан, гексан, циклогексан, гептан, октан или циклооктан, галогенированные алифатические углеводороды, предпочтительно имеющие один или два атома углерода, такие как дихлорметан, хлороформ, тетрахлорид углерода, дихлорэтан или тетрахлорэтан, ароматические углеводороды, такие как бензол, толуол и ксилолы, хлорбензол или дихлорбензол, алифатические ациклические и циклические простые эфиры или спирты, предпочтительно имеющие от 4 до 8 атомов углерода, такие как этанол, изопропанол, диэтиловый эфир, метил-трет-бутиловый эфир, этил-трет-бутиловый эфир, дипропиловый эфир, диизопропиловый эфир, дибутиловый эфир, тетрагидрофуран, или сложные эфиры, такие как этилацетат или н-бутилацетат, или кетоны, такие как метилизобутилкетон или диоксан, или их смеси. Особенно предпочтительно применяемыми растворителями являются упомянутые выше гептан, метил трет-бутиловый эфир (также известный как MTBE, трет-бутилметиловый эфир, третичный бутилметиловый эфир и третBME), диизопропиловый эфир, тетрагидрофуран, этилацетат и/или их смеси.

Предпочтительно смешиваемый с водой растворитель, такой как этанол, используют для экстракции (-)-амброксида из твердой фазы реакционной смеси. Применение этанола обладает преимуществом, поскольку он прост в обращении, нетоксичен и благоприятен для окружающей среды.

Используемый в настоящем описании термин «изолированный» относится к продукту биоконверсии, такому как (-)-амброксид, который был отделен или очищен от сопутствующих ему компонентов. Объект, продуцируемый клеточной системой, отличающейся от источника, из которого она имеет происхождение в природе, является изолированным, поскольку он обязательно будет свободен от компонентов, сопутствующих ему в природе. Степень изоляции или чистоты можно измерить любым подходящим способом, например, с помощью анализа методом газовой хроматографии (GC), высокоэффективной жидкостной хроматографии (HPLC) или ядерного магнитного резонанса (NMR).

В некоторых воплощениях изобретения конечный продукт ((-)-амброксид) выделяют и очищают до однородности (например, по меньшей мере до 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% или 89,5% чистоты или 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 99,5% чистоты).

Желательно, количество продуцируемого (-)-амброксида может составлять от приблизительно 1 мг/л до приблизительно 20000 мг/л (20 г/л) или более, например от приблизительно 20 г/л до приблизительно 200 г/л или 100-200 г/л, предпочтительно приблизительно 125 г/л или 150 г/л или приблизительно 188 г/л.

Как показано в примере 7, по меньшей мере 125 г/л (-)-амброксида получают в реакции биоконверсии с использованием рекомбинантной клетки-хозяина E. coli, продуцирующей производное фермента SHC/HAC, приблизительно за 2 дня.

Как показано в примере 19, можно проводить реакции биоконверсии при 188 г/л ЕЕН или более при условии достижения эффективного перемешивания, поскольку эффективность перемешивания, по-видимому, является единственным ограничением системы. Кроме того, биокатализатор с улучшенной активностью (например, в отношении вариантов SHC с дополнительно улучшенной активностью или в отношении повышенной производительности фермента SHC) может улучшить или сохранить производительность при использовании меньшего количества биомассы, что обладает преимуществом в отношении эффективности перемешивания.

Например, получают от приблизительно 1 до приблизительно 100 мг/л, от приблизительно 30 до приблизительно 100 мг/л, от приблизительно 50 до приблизительно 200 мг/л, от приблизительно 100 до приблизительно 500 мг/л, от приблизительно 100 до приблизительно 1000 мг/л, от приблизительно 250 до приблизительно 5000 мг/л, от приблизительно 1000 (1 г/л) до приблизительно 15000 мг/л (15 г/л) или от приблизительно 2000 (2 г/л) до приблизительно 10000 мг/л (10 г/л) или от приблизительно 2000 (2 г/л) до приблизительно 25000 мг/л (25 г/л) или от приблизительно 2000 (2 г/л) до приблизительно 25000 мг/л (25 г/л), 26000 мг/л (26 г/л), 27000 мг/л (27 г/л), 28000 мг/л (28 г/л), 29000 мг/л (29 г/л), 30000 мг/л (30 г/л), 40 г/л, 50 г/л, 60 г/л, 70 г/л, 80 г/л, 90 г/л, 100 г/л, 110 г/л, 120 г/л, 125 г/л, 130 г/л, 140 г/л, 150 г/л, 160 г/л, 170 г/л, 180 г/л, 190 г/л, или 200 г/л, или 300 г/л, или 400 г/л, или 500 г/л (-)-амброксида.

Предпочтительно за период времени от 48 до 72 часов получают (-)-амброксид в концентрации по меньшей мере 100 г/л.

Предпочтительно за период времени от приблизительно 48 до 72 часов получают (-)-амброксид в концентрации приблизительно 150 г/л.

Предпочтительно за период времени от приблизительно 48 до 72 часов получают (-)-амброксид в концентрации приблизительно 200 г/л.

Предпочтительно за период времени от приблизительно 48 до 72 часов получают (-)-амброксид в концентрации приблизительно 250 г/л.

Специалисту в данной области техники должно быть понятно, что более высокие совокупные титры продуктивности могут быть достигнуты путем применения непрерывного процесса, такого как извлечение продукта, подача субстрата и добавление или (частичная) замена биомассы.

Предпочтительно биоконверсия ЕЕН в (-)-амброксид в присутствии рекомбинантной клетки-хозяина, содержащей WT SHC/HAC или производное SHC/HAC приводит к выходу амброксида, составляющему 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 молярных процентов от используемого числа моль EEH; особенно предпочтительно выход составляет от 5 до 100, от 10 до 100, 100, от 25 до 100, от 30 до 100, от 35 до 100, в частности, от 40 до 100, от 45 до 100, от 50 до 100, от 60 до 100, от 70 до 100 молярных процентов.

Активность фермента SHC/HAC определяют на основании коэффициента скорости реакции (количество продукта/(количество продукта + количество остаточного исходного материала)) × 100) в молярных процентах. Предпочтительно биоконверсия EEH в (-)-амброксид в присутствии WT SHC/HAC или производного фермента SHC/HAC приводит к выходу (-)-амброксида, составляющему 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 молярных процентов от используемого числа моль EEH; особенно предпочтительно выход составляет от 5 до 100, от 10 до 100, от 20 до 100, от 25 до 100, от 30 до 100, от 35 до 100, в частности, от 40 до 100, от 45 до 100, от 50 до 100, от 60 до 100, от 70 до 100.

В предпочтительном воплощении изобретения выход и/или коэффициент скорости реакции определяют за определенный период времени, например, 4, 6, 8, 10, 12, 16, 20, 24, 36 или 48 часов, в течение которого EEH конвертируется в (-)-амброксид рекомбинантной клеткой-хозяином, содержащей нуклеотидную последовательность, кодирующую WTSHC или производное фермента SHC/HAC согласно настоящему изобретению. В следующем варианте реакцию проводят в точно определенных условиях, например, при 25°C, 30°C, 40°C, 50°C или 60°C. В частности, выход и/или коэффициент скорости реакции определяют путем проведения реакции конверсии EEH в (-)-амброксид производными фермента SHC/HAC согласно изобретению при 35°C за период 24-72 часа.

В следующем воплощении настоящего изобретения рекомбинантная клетка-хозяин, содержащая нуклеотидную последовательность, кодирующую производное SHC/HAC, характеризуется тем, что она проявляет 2-, 3-, 4-, 5-, 6-, 7-, 8-, 9-, 10-, 11-, 12-, 13-, 14-, 15-, 16-, 17-, 18-, 19-, 20-, 21-, 22-, 23-, 24-, 25-, 26-, 27-, 28-, 29-, 30-, 31-, 32-, 33-, 34-, 35-, 36-, 37-, 38-, 39-, 40-, 41-, 42-, 43-, 44-, 45-, 46-, 47-, 48-, 49-, 50-, 51-, 52-, 53-, 54-, 55-, 56-, 57-, 58-, 59-, 60-, 61-, 62-, 63-, 64-, 65-, 66-, 67-, 68-, 69-, 70-, 71-, 72-, 73-, 74-, 75-, 76-, 77-, 78-, 79-, 80-, 81-, 82-, 83-, 84-, 85-, 86-, 87-, 88-, 89-, 90-, 91-, 92-, 93-, 94-, 95-, 96-, 97-, 98-, 99-, 100-, 200-, 500-, 1000-кратный или более высокий выход и/или коэффициент скорости реакции в реакции гомофарнезола с получением (-)-амброксида по сравнению с WT SHC или производным фермента SHC/HAC в одинаковых условиях. В данном случае термин «условие» относится к условиям реакции, таким как концентрация субстрата, концентрация фермента, период и/или температура реакции.

Успешная разработка способа биоконверсии для получения (-)-амброксида из гомофарнезола в рекомбинантном штамме E. coli, содержащем нуклеотидную последовательность, кодирующую WT/эталонный SHC или производное SHC/HAC, может обеспечить снижение затрат и экономичный с промышленной точки зрения процесс производства (-)-амброксида.

Как показано в примере 7, настоящее изобретение обеспечивает 100% конверсию E,E-гомофарнезола (125 г/л) в (-)-амброксид после 48 часов инкубации с оптимизированным производным SHC/HAC с 8-кратным улучшением выхода при использовании производного AacSHC по сравнению с ферментом WT AacSHC (см. Фиг. 11).

Для использования в получении амброксида в рекомбинантной клетке-хозяине также приемлемы функциональные гомологи полипептидов эталонного WT SHC/HAC или производного SHC/HAC, описанных в настоящем описании. Таким образом, рекомбинантная клетка-хозяин может включать в себя одну или более чем одну гетерологичную нуклеиновую кислоту, кодирующую функциональные гомологи описанных выше полипептидов, и/или гетерологичную нуклеиновую кислоту, кодирующую производное фермента SHC/HAC, как описано в настоящем описании.

Функциональный гомолог представляет собой полипептид, обладающий подобием последовательности с эталонным полипептидом и осуществляющий одну или более биохимических или физиологических функций эталонного полипептида. Функциональный гомолог и эталонный полипептид могут представлять собой встречающиеся в природе полипептиды, и подобие последовательностей может быть результатом конвергентных или дивергентных эволюционных событий. Как таковые, функциональные гомологи иногда обозначают в литературе как гомологи, ортологи или паралоги. Сами варианты встречающихся в природе функциональных гомологов, такие как полипептиды, кодируемые мутантами кодирующей последовательности дикого типа, могут быть функциональными гомологами. Функциональные гомологи можно также создать посредством сайт-направленного мутагенеза кодирующей последовательности для полипептида или комбинирования доменов из кодирующих последовательностей для различных встречающихся в природе полипептидов («замены доменов»). Методы модификации генов, кодирующих описанные выше функциональные гомологи, известны и включают среди прочего методы направленного развития, методы сайт-направленного мутагенеза и методы случайного мутагенеза и могут быть полезны для повышения специфической активности полипептида, изменения субстратной специфичности, изменения уровней экспрессии, изменения субклеточной локализации или модификации взаимодействий между полипептидами требуемым образом. Такие модифицированные полипептиды считают функциональными гомологами. Термин «функциональный гомолог» иногда применяют к нуклеиновой кислоте, кодирующей функционально гомологичный полипептид.

Функциональные гомологи можно идентифицировать с помощью анализа выравнивания нуклеотидных и аминокислотных последовательностей. Например, при проведении поиска нуклеотидных или полипептидных последовательностей в базе данных можно идентифицировать гомологи нуклеиново-кислотных последовательностей, кодирующих полипептиды производных SHC и т.п.

Для идентификации функциональных гомологов и/или в качестве меры гомологии между двумя нуклеотидными последовательностями можно также использовать гибридизацию. В качестве зонда для гибридизации можно использовать нуклеиново-кислотную последовательность, кодирующую любой из раскрытых в настоящем описании белков в соответствии со стандартными методами гибридизации. Гибридизация зонда с ДНК или РНК из тестируемого источника (например, из клетки млекопитающего) является показателем присутствия релевантной ДНК или РНК в тестируемом источнике. Условия гибридизации известны специалистам в данной области техники, и их можно найти в руководстве Current Protocols in Molecular Biology, John Wiley & Sons, N.Y., 6.3.1-6.3.6, 1991. Умеренные условия гибридизации определяют как эквивалентные гибридизации в 2x растворе хлорида натрия/цитрата натрия (SSC) при 30°C с последующей отмывкой в 1x SSC, 0,1% SDS при 50°C. Условия высокой степени жесткости определяют как эквивалентные гибридизации в 6x растворе хлорида натрия/цитрата натрия (SSC) при 45°C с последующей отмывкой в 0,2x SSC, 0,1% SDS при 65°C.

Анализ последовательностей для идентификации функциональных гомологов может также включать в себя анализ неизбыточных баз данных BLAST, Reciprocal BLAST или PSI-BLAST с использованием релевантной аминокислотной последовательности в качестве эталонной последовательности. В некоторых случаях аминокислотную последовательность выводят из нуклеотидной последовательности. Полипептиды в базе данных, имеющие более 40% идентичности последовательности, являются кандидатами для дополнительной оценки на пригодность для применения в реакции биоконверсии SHC/HAC. Подобие аминокислотных последовательностей дает возможность проведения консервативных аминокислотных замен, таких как замена одного гидрофобного остатка другим или замена одного полярного остатка другим. При желании можно проводить проверку таких кандидатов вручную, чтобы сузить число кандидатов для дальнейшей оценки. Проверку вручную можно проводить путем отбора тех кандидатов, в которых выявляются, например, консервативные функциональные домены.

Как правило, полипептиды, обладающие по меньшей мере приблизительно 30% идентичностью аминокислотной последовательности, полезны для идентификации консервативных областей. Консервативные области родственных полипептидов обладают по меньшей мере 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% идентичностью аминокислотной последовательности. В некоторых воплощениях изобретения консервативная область обладает по меньшей мере 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, по меньшей мере 81%, по меньшей мере 82%, по меньшей мере 83%, по меньшей мере 84%, по меньшей мере 85%, по меньшей мере 86%, по меньшей мере 87%, по меньшей мере 88%, по меньшей мере 89%, по меньшей мере 90%, по меньшей мере 91%, по меньшей мере 92%, по меньшей мере 93%, по меньшей мере 94%, по меньшей мере 95%, по меньшей мере 96%, по меньшей мере 97%, по меньшей мере 98% или по меньшей мере 99% идентичностью аминокислотной последовательности. Идентичность аминокислотной последовательности можно определять, как описано выше и ниже.

Продуцируемый фермент WTSHC и/или производное фермента SHC/HAC основан на аминокислотной последовательности SEQ ID NO: 1, или SEQ ID NO: 2, или SEQ ID NO: 3, или SEQ ID NO: 4, или их варианте, гомологе, мутанте, производном или фрагменте.

Продуцируемый фермент SHC основан на аминокислотной последовательности, обладающей по меньшей мере 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%, по меньшей мере 81%, по меньшей мере 82%, по меньшей мере 83%, по меньшей мере 84%, по меньшей мере 85%, по меньшей мере 86%, по меньшей мере 87%, по меньшей мере 88%, по меньшей мере 89%, по меньшей мере 90%, по меньшей мере 91%, по меньшей мере 92%, по меньшей мере 93%, по меньшей мере 94%, по меньшей мере 95%, по меньшей мере 96%, по меньшей мере 97%, по меньшей мере 98% или по меньшей мере 99% идентичностью последовательности SEQ ID NO: 1, или SEQ ID NO: 2, или SEQ ID NO: 3, или SEQ ID NO: 4.

Кроме того, продуцируемый эталонный SHC основан на аминокислотной последовательности, продуцируемой в E. coli.

Процент (%) идентичности в отношении нуклеотидной последовательности гена определяют как процентную долю нуклеотидов в последовательности-кандидате ДНК,

идентичных нуклеотидам в последовательности ДНК после выравнивания последовательности и при необходимости введения гэпов для достижения максимального процента идентичности последовательности, и не считают какие-либо консервативные замены как часть идентичности последовательности. Выравнивание в целях определения процента идентичности нуклеотидных последовательностей может быть достигнуто различными путями, находящимися в пределах компетенции специалистов в данной области техники, например, при использовании общедоступного компьютерного программного обеспечения. Специалисты в данной области техники могут определить подходящие параметры для измерения выравнивания, включая любые алгоритмы, требующиеся для достижения максимального выравнивания на протяжении всей длины сравниваемых последовательностей.

Термины «полипептид» и «белок» в настоящем описании используют взаимозаменяемо, и они означают любую цепь связанных пептидной связью аминокислот независимо от ее длины и посттрансляционных модификаций.

Используемый в настоящем описании термин «производное» включает, без ограничений, вариант. Термины «производное» и «вариант» в настоящем описании используют взаимозаменяемо

Используемый в настоящем описании термин «вариант» следует понимать как полипептид, который отличается по сравнению с полипептидом, от которого он имеет происхождение, одним или более чем одним изменением в аминокислотной последовательности. Полипептид, от которого имеет происхождение вариант, также известен как исходный или эталонный полипептид. Как правило, вариант конструируют искусственным путем, предпочтительно с использованием генных технологий. Как правило, полипептид, от которого имеет происхождение вариант, представляет собой белок дикого типа или домен белка дикого типа. Однако варианты, которые можно использовать в настоящем изобретении, могут также иметь происхождение от гомологов, ортологов или паралогов родительского полипептида или от искусственно сконструированных вариантов при условии, что этот вариант проявляет по меньшей мере одну биологическую активность родительского полипептида. Изменения в аминокислотной последовательности могут представлять собой замены аминокислот, инсерции, делеции, N-концевые укорочения или C-концевые укорочения или любую комбинацию этих изменений, которые могут встречаться в одном или нескольких сайтах.

В предпочтительных воплощениях изобретения вариант, который можно использовать в настоящем изобретении, характеризуется общим числом изменений в аминокислотной последовательность вплоть до 200 (вплоть до 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 или 200) (т.е. замен, инсерций, делеций, N-концевых укорочений и/или C-концевых укорочений). Аминокислотные замены могут быть консервативными и/или неконсервативными. В предпочтительных воплощениях изобретения вариант, который можно использовать в настоящем изобретении, отличается от белка или домена, от которого он имеет происхождение, заменами аминокислот, предпочтительно консервативными заменами аминокислот, в количестве вплоть до 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 или 100. Варианты могут дополнительно или альтернативно содержать делеции аминокислот, которые могут представлять собой N-концевые укорочения, C-концевые укорочения или внутренние делеции, либо любую их комбинацию. Такие варианты, содержащие N-концевые укорочения, C-концевые укорочения и/или внутренние делеции, в контексте настоящей заявки называют «делеционными вариантами» или «фрагментами». Термины

«делеционный вариант» и «фрагмент» в настоящем описании используют взаимозаменяемо. Делеционный вариант может встречаться в природе (например, варианты сплайсинга) или может быть сконструирован искусственно, предпочтительно с помощью средств генных технологий. Как правило, белок или домен белка, от которого имеет происхождение делеционный вариант, представляет собой белок дикого типа. Однако делеционные варианты, которые можно использовать в настоящем изобретении, могут также иметь происхождение от гомологов, ортологов или паралогов родительского полипептида или от искусственно сконструированных вариантов при условии, что делеционные варианты проявляют по меньшей мере одну биологическую активность родительского полипептида. Предпочтительно делеционный вариант (или фрагмент) имеет делецию вплоть до 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 или 100 аминокислот на его N-конце и/или не его C-конце и/или внутри последовательности по сравнению с родительским полипептидом.

Используемый в настоящем описании «вариант» может быть альтернативно или дополнительно охарактеризован определенной степенью идентичности последовательности с родительским полипептидом, от которого он имеет происхождение. Вариант WT/эталонного SHC/HAC или производного SHC/HAC по настоящему изобретению может обладать по меньшей мере 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%, по меньшей мере 81%, по меньшей мере 82%, по меньшей мере 83%, по меньшей мере 84%, по меньшей мере 85%, по меньшей мере 86%, по меньшей мере 87%, по меньшей мере 88%, по меньшей мере 89%, по меньшей мере 90%, по меньшей мере 91%, по меньшей мере 92%, по меньшей мере 93%, по меньшей мере 94%, по меньшей мере 95%, по меньшей мере 96%, по меньшей мере 97%, по меньшей мере 98% или по меньшей мере 99% идентичностью последовательности соответствующего эталонного полипептида или соответствующего эталонного полинуклеотида.

Выражение «по меньшей мере 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%, по меньшей мере 81%, по меньшей мере 82%, по меньшей мере 83%, по меньшей мере 84%, по меньшей мере 85%, по меньшей мере 86%, по меньшей мере 87%, по меньшей мере 88%, по меньшей мере 89%, по меньшей мере 90%, по меньшей мере 91%, по меньшей мере 92%, по меньшей мере 93%, по меньшей мере 94%, по меньшей мере 95%, по меньшей мере 96%, по меньшей мере 97%, по меньшей мере 98% или по меньшей мере 99% идентичности последовательности» на протяжении всего описания используют по отношению к сравнениям полипептидных и полинуклеотидных последовательностей.

Полинуклеотид, принадлежащий к семейству любого из раскрытых в настоящем описании ферментов, или белок можно идентифицировать на основании его подобия релевантному гену или белку соответственно. Например, идентификация может быть основана на идентичности последовательности. В некоторых предпочтительных воплощениях изобретение характеризуется изолированными молекулами нуклеиновой кислоты, по меньшей мере на 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%, по меньшей мере на 81%, по меньшей мере на 82%, по меньшей мере на 83%, по меньшей мере на 84%, по меньшей мере на 85%, по меньшей мере на 86%, по меньшей мере на 87%, по меньшей мере на 88%, по меньшей мере на 89%, по меньшей мере на 90%, по меньшей

мере на 91%, по меньшей мере на 92%, по меньшей мере на 93%, по меньшей мере на 94%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98% или по меньшей мере на 99% идентичными (a) молекуле нуклеиновой кислоты, кодирующей полипептид SEQ ID NO: 5-163 (см. таблицы 14-17 и таблицу 4a, приведенные в настоящем описании); (b) нуклеотидной последовательности SEQ ID NO: 6-168, 169, 170, 172, 174 и 176 (см. таблицы 14-17 и таблицу 4a, приведенные в настоящем описании) и (c) молекуле нуклеиновой кислоты, включающей в себя сегмент из по меньшей мере 30 (например, по меньшей мере 30, 40, 50, 60, 80, 100, 125, 150, 175, 200, 250, 300, 400, 500, 600, 700, 800, 850, 900, 950, 1000 или 1010) нуклеотидов SEQ ID NO: 6-168, 169, 170, 172, 174 и 176 (см. таблицы 14-17 и таблицу 4a, приведенные в настоящем описании).

Предпочтительно рассматриваемый полипептид и эталонный полипептид проявляют указанную идентичность последовательности на протяжении непрерывного отрезка из 20, 30, 40, 45, 50, 60, 70, 80, 90, 100 или более аминокислот. Предпочтительно рассматриваемый полинуклеотид и эталонный полинуклеотид проявляют указанную идентичность последовательности на протяжении непрерывного отрезка из 60, 90, 120, 135, 150, 180, 210, 240, 270, 300 или более нуклеотидов. В случае сравнения двух последовательностей и отсутствия уточненной эталонной последовательности, по сравнению с которой следует определять идентичность последовательности, идентичность последовательности следует рассчитывать по отношению к более длинной из двух подлежащих сравнению последовательностей, если конкретно не указано иное. Если эталонная последовательность указана, идентичность последовательности определяют на основании всей длины эталонной последовательности, указанной SEQ ID NO: 1, 2, 3 и/или 4, если конкретно не указано иное.

Например, пептидная последовательность, состоящая из 130 аминокислот, сравниваемая с аминокислотами полноразмерного эталонного SHC, содержащего 631 аминокислотный остаток, может проявлять максимальный процент идентичности последовательности 20,6% (130/631×100), а последовательность длиной 300 аминокислот может проявлять максимальный процент идентичности последовательности 47,5% (300/631×100). Подобие нуклеотидных и аминокислотных последовательностей, т.е. процент идентичности последовательности, можно определить посредством выравнивания последовательностей. Такие выравнивания можно осуществлять с помощью нескольких известных в данной области техники алгоритмов, предпочтительно с помощью математического алгоритма Karlin и Altschul (Karlin & Altschul (1993) Proc. Natl. Acad. Sci. USA 90: 5873-5877), с помощью алгоритма hmmalign (пакет программ HMMER, http://hmmer.wustl.edu/) или с помощью алгоритма CLUSTAL (Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994) Nucleic Acids Res. 22, 4673-80), доступного, например, на сайте http://www.ebi.ac.uk/Tools/clustalw/ или на сайте http://www.ebi.ac.uk/Tools/clustalw2/index.html, или на сайте http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_clustalw.html или с помощью программы GAP (математического алгоритма университета Айовы), которую использовали при выравнивании последовательностей WTSHC, как представлено в таблице 18 настоящего документа, или с помощью математического алгоритма Myers и Miller (1989 - Cabios 4: 11-17), раскрытого и используемого при выравнивании последовательностей WTSHC в таблице 19, представленной в настоящем описании.

Предпочтительно используют параметры по умолчанию, представленные на сайте http://www.ebi.ac.uk/Tools/clustalw/ или http://www.ebi.ac.uk/Tools/clustalw2/index.html.

Степень идентичности последовательностей (совпадения последовательностей)

можно рассчитать, используя, например, программы BLAST, BLAT или BlastZ (или BlastX). Подобный алгоритм включен в программы BLASTN и BLASTP Altschul et al (1990) J. Mol. Biol. 215, 403-410. Поиск полинуклеотидов BLAST проводят с помощью программы BLASTN, балл = 100, длина слова = 12, с получением полинуклеотидных последовательностей, гомологичных нуклеиновым кислотам, кодирующим релевантный белок.

Поиск белков BLAST выполняют с помощью программы BLASTP, балл = 50, длина слова = 3, с получением аминокислотных последовательностей, гомологичных полипептиду SrKO. Для получения выравнивания с использованием гэпов в сравнительных целях используют Gapped BLAST, как описано в публикации Altschul et al (1997) Nucleic Acids Res. 25, 3389-3402. При использовании программ BLAST и Gapped BLAST используют параметры соответствующих программ по умолчанию. Анализ совпадения последовательностей может быть дополнен установленными методами картирования по гомологии, такими как Shuffle-LAGAN (Brudno M., Bioinformatics 2003b, 19 Suppl 1:154-162) или метод марковских случайных полей. При ссылке на процент идентичности последовательностей в настоящей заявке эти проценты вычисляют по отношению ко всей длине более длинной последовательности, если конкретно не указано иное.

Сенсорный анализ продукта

В результате биоконверсии гомофарнезола в (-)-амброксид в соответствии с настоящим изобретением (-)-амброксид получают в качестве преобладающего соединения, но также могут быть получены соединения, отличающиеся от (-)-амброксида, которые могут придавать или не придавать приятные ольфакторные ноты смеси биоконверсии и, таким образом, могут положительным или отрицательным образом влиять на сенсорный характер конечного продукта (-)-амброксида. Соответственно, проводят сенсорный анализ, используя общепризнанные сенсорные тесты, которые проводят подготовленные эксперты (например, парфюмеры), и, таким образом, это тестирование может способствовать определению ольфакторной релевантности конечного продукта относительно эталонного продукта для химически релевантного целевого продукта. Как показывает сенсорный анализ в примере 22, удаление из (-)-амброксида одного или более соединений, представляющих собой побочные продукты, может улучшить запах остаточного соединения ((-)-амброксида) даже в том случае, когда удаляемые соединения как таковые не имеют запаха. Таким образом, усиление запаха (-)-амброксида наблюдали в отсутствие соединений II, III и IV.

Аспекты изобретения

1. Способ получения (-)-амброксида или смеси, содержащей (-)-амброксид, где (3E,7E)-гомофарнезол (EEH) или смесь стереоизомеров, содержащую EEH, конвертируют ферментативным путем в (-)-амброксид или в смесь, содержащую (-)-амброксид, где ферментативную конверсию осуществляют с использованием фермента SHC/HAC в реакционных условиях, подходящих для получения (-)-амброксида, и где смесь стереоизомеров, содержащая EEH, по существу состоит из изомеров гомофарнезола, выбранных из группы, состоящей из [(3E,7E) и [(3Z,7E)] и/или [(3E,7E) и (3E,7Z)] и/или [(3Z,7E), (3E,7E) и (3E,7Z)], также обозначенных как [EE:EZ], [EE:ZE] и [EE:EZ:ZE], соответственно.

2. Способ получения (-)-амброксида или смеси, содержащей (-)-амброксид, где (3E,7E)-гомофарнезол (EEH) или смесь стереоизомеров, содержащую EEH, ферментативно конвертируют с получением (-)-амброксида или смеси, содержащей (-)-амброксид, где

ферментативную конверсию осуществляют с использованием фермента SHC/HAC в реакционных условиях, подходящих для получения (-)-амброксида, и где в случае проведения реакции в присутствии солюбилизирующего агента, Тритон X-100 или тауродезоксихолат не используют в комбинации с ферментом SHC/HAC дикого типа.

3. Способ получения (-)-амброксида или смеси, содержащей (-)-амброксид, где (3E,7E) -гомофарнезол (EEH) или смесь стереоизомеров, содержащую EEH, конвертируют ферментативным путем в (-)-амброксид или в смесь, содержащую (-)-амброксид, где ферментативную конверсию осуществляют с использованием фермента SHC/HAC в реакционных условиях, подходящих для получения (-)-амброксида, и где смесь стереоизомеров, содержащая EEH, по существу состоит из изомеров гомофарнезола, выбранных из группы, состоящей из [(3E,7E) и [(3Z,7E)] и/или [(3E,7E) и (3E,7Z)] и/или [(3Z,7E), (3E,7E) и (3E,7Z)], также обозначенных как [EE:EZ], [EE:ZE] и [EE:EZ:ZE], соответственно, и где реакцию проводят в трехфазной системе, содержащей водную фазу, твердую фазу и масляную фазу.

4. Способ в соответствии с параграфом 1, или с параграфом 2, или с параграфом 3, который осуществляют с использованием полипептидной последовательности фермента SHC/HAC, выбранной из группы, состоящей из SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3; SEQ ID NO: 4, или производного SHC/HAC, выбранного из таблицы 1, таблицы 5, таблицы 2, таблицы 6, таблицы 3, таблицы 7, таблицы 4, таблицы 8 или таблицы 13, таблицы 14, или выбранного из 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 и/или 178 или последовательности с по меньшей мере 30% идентичностью, по меньшей мере 40% идентичностью, по меньшей мере 50% идентичностью, или по меньшей мере 60% идентичностью, или по меньшей мере 70% идентичностью, или по меньшей мере 80% идентичностью, или по меньшей мере 90% идентичностью, или по меньшей мере 95% идентичностью, или по меньшей мере 96% идентичностью, или по меньшей мере 97% идентичностью, или по меньшей мере 98% идентичностью, или по меньшей мере 99% идентичностью SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3 или SEQ ID NO: 4.

5. Способ по любому из параграфов 1-4, где используют рекомбинантные клетки-хозяева, продуцирующие фермент SHC/HAC.

6. Способ в соответствии с параграфом 4 или с параграфом 5, где нуклеотидная последовательность, кодирующая фермент SHC/HAC, выбрана из группы, состоящей из SEQ ID NO: 165, 166, 167, 168, 169 или SEQ ID NO: 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40 и/или 170, 172, 174 и/или 176.

7. Способ в соответствии с любым из параграфов 16, где конверсию гомофарнезола в (-)-амброксид проводят при температуре в диапазоне от 30°C до 60°C, при pH в диапазоне приблизительно 4-8.

8. Способ в соответствии с любым из параграфов 17, где конверсию гомофарнезола в (-)-амброксид проводят с использованием одного или более чем одного из реакционных условий для фермента, представляющего собой SHC/HAC дикого типа или производное фермента SHC/HAC, которые представлены в таблице 24 или в таблице 24a, предпочтительно при pH в диапазоне от 5,0 до 6,2, предпочтительно при температуре 35°C.

9. Способ в соответствии с любым из параграфов 3-8, где соотношение SDS/клетки находится в диапазоне от 10:1 до 20:1, предпочтительно составляет 16:1 при соотношении биокатализатора и EEH, которое составляет приблизительно 2:1.

10. Способ в соответствии с любым из параграфов 3-9, где массовое соотношение биокатализатора к гомофарнезолу находится в диапазоне от приблизительно 0,5-2:1,

предпочтительно составляет приблизительно 1:1 или 0,5:1.

11. Способ в соответствии с любым из параграфов 3-10, где стадии выращивания клеток и реакции биоконверсии осуществляют в одном реакционном сосуде.

12. Способ в соответствии с параграфом 2, где субстрат гомофарнезола содержит один или более чем один стереоизомер гомофарнезола.

13. Способ в соответствии с параграфом 12, где субстрат гомофарнезола содержит или по существу состоит из двух стереоизомеров гомофарнезола.

14. Способ в соответствии с параграфом 13, где субстрат гомофарнезола содержит или по существу состоит из стереоизомеров EE:EZ.

15. Способ в соответствии с параграфом 14, где гомофарнезол содержит или по существу состоит из смеси стереоизомеров EE:EZ в массовых соотношениях, выбранных из группы, состоящей из: 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; и 60:40.

16. Способ в соответствии с параграфом 15, где гомофарнезол содержит или по существу состоит из смеси стереоизомеров EE:EZ в массовом соотношении, выбранном из группы, состоящей из: EE:EZ 92:8; EE:EZ 90:10; EE:EZ 80:20; EE:EZ 86:14; EE:EZ 70:30; EE:EZ 69:31; и EE:EZ 66:34.

17. Способ в соответствии с параграфом 15 или с параграфом 16, где гомофарнезол содержит или по существу состоит из смеси стереоизомеров EE:EZ в массовом соотношении 80:20.

18. Способ в соответствии с любым из параграфов 1-17, где (-)-амброксид получают с примесью по меньшей мере одного или более чем одного из побочных продуктов (II), (IV) и/или (III).

19. Способ в соответствии с любым из параграфов 1-18, где (-)-амброксид выделяют из биоконверсионной реакционной смеси с использованием стадии экстракции/дистилляции органическим растворителем или паром или фильтрования.

20. Способ в соответствии с любым из параграфов 1-19, где (-)-амброксид выделяют из твердой фазы биоконверсионной реакционной смеси с использованием стадии экстракции/дистилляции органическим растворителем или паром.

21. Способ в соответствии с параграфом 19 или с параграфом 20, где (-)-амброксид выделяют из реакционной смеси с использованием органического растворителя.

22. Способ в соответствии с параграфом 21, где (-)-амброксид выделяют из реакционной смеси с использованием этанола или толуола.

23. Способ в соответствии с любым из параграфов 19-22, где (-)-амброксид селективно кристаллизуют с использованием органического растворителя.

24. Способ в соответствии с параграфом 23, где (-)-амброксид по существу не содержит побочных продуктов (II), (IV) и/или (III).

25. Способ в соответствии с любым из параграфов 1-24, где (-)-амброксид получают в диапазоне концентраций приблизительно 125-200 г/л.

26. (-)-Амброксид, получаемый способом по любому из параграфов 1-25, где (-)-амброксид имеет порог восприятия запаха от приблизительно 0,1 до приблизительно 0,5 нг/л.

27. (-)-Амброксид по параграфу 26 в твердой форме, предпочтительно в аморфной или кристаллической форме.

28. Способ получения продукта, содержащего (-)-амброксид, включающий включение в продукт (-)-амброксида по любому из параграфа 26 или параграфа 27.

29. Способ по параграфу 28, где продукт представляет собой ароматизирующий продукт, косметический продукт, чистящий продукт, моющий продукт или мыльный продукт.

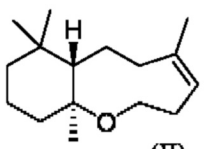
30. Ароматизирующий продукт, или косметический продукт, или потребительский продукт для ухода, содержащий (-)-амброксид по любому из параграфа 26 или параграфа 27.

31. Ароматизирующая композиция, или косметическая композиция, или потребительская композиция для ухода, содержащая (-)-амброксид по параграфу 26 или параграфу 27 и один или более чем один дополнительный компонент.

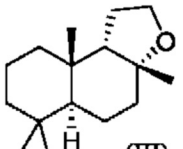
32. Применение (-)-амброксида по параграфу 26 или параграфу 27 в качестве составной части ароматизирующего продукта, или косметического продукта, или потребительского продукта, такого как продукт для ухода за тканями, гигиено-косметический продукт, продукт для заботы о красоте и/или чистящий продукт.

33. Способ увеличения, усиления или придания аромата ароматизирующей композиции, включающий стадию смешивания с указанной ароматизирующей композицией продукта, увеличивающего или усиливающего аромат, полученного в соответствии со способом, включающим стадии:

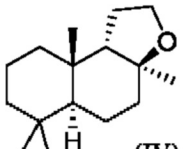
(а) приготовления реакционной смеси, содержащей (-)-амброксид в смеси с одним или более чем одним из соединений побочных продуктов (II), (III) или (IV).



(II)



(III)



(IV)

(б) экстрагирования (-)-амброксида в смеси с одним или более чем одним из соединений побочных продуктов (II), (III) или (IV); и

(в) селективной кристаллизации (-)-амброксида из экстракционной смеси;

где (-)-амброксид получают путем ферментативной конверсии (3E,7E)-гомофарнезола (EEH) или смеси стереоизомеров, содержащей EEH, с использованием фермента SHC/HAC в реакционных условиях, приемлемых для получения (-)-амброксида, и где смесь стереоизомеров, содержащая EEH, по существу состоит из изомеров гомофарнезола, выбранных из группы, состоящей из [(3E,7E) и [(3Z,7E)] и/или [(3E,7E) и (3E,7Z)] и/или [(3Z,7E), (3E,7E) и (3E,7Z)], также обозначенную как [EE:EZ], [EE:ZE] и [EE:EZ:ZE], соответственно.

34. Способ по параграфу 33, где реакцию проводят в трехфазной системе, содержащей водную фазу, твердую фазу и масляную фазу.

35. Способ в соответствии с параграфом 33, или с параграфом 34, где процесс осуществляют с использованием фермента SHC/HAC с полипептидной последовательностью, выбранной из группы, состоящей из SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3; SEQ ID NO: 4, или производного SHC/HAC, выбранного из таблицы 1, таблицы 5, таблицы 2, таблицы 6, таблицы 3, таблицы 7, таблицы 4, таблицы 8 или таблицы 14, или выбранного из SEQ ID NO: 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 171, 173, 175 и/или 178 или последовательности с по меньшей мере 30% идентичностью, по меньшей мере 40% идентичностью, по меньшей мере 50% идентичностью, или по меньшей мере 60% идентичностью, или по меньшей мере 70% идентичностью, или по меньшей мере 80% идентичностью, или по меньшей мере 90% идентичностью, или по меньшей мере 95% идентичностью, или по меньшей мере 96% идентичностью, или по меньшей мере 97% идентичностью, или по меньшей мере 98%

идентичностью, или по меньшей мере 99% идентичностью SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3 или SEQ ID NO: 4.

36. Способ по любому из параграфов 33-35, где используют рекомбинантные клетки-хозяева, продуцирующие фермент SHC/HAC.

37. Способ в соответствии с параграфом 35 или с параграфом 36, где нуклеотидная последовательность, кодирующая фермент SHC/HAC, выбрана из группы, состоящей из SEQ ID NO: 165, 166, 167, 168, 169 или SEQ ID NO: 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40 и/или 170, 172, 174 и/или 176.

38. Способ в соответствии с любым из параграфов 33-37, где конверсию гомофарнезола в (-)-амброксид проводят при температуре в диапазоне от 30°C до 60°C, при pH в диапазоне приблизительно 4-8.

39. Способ в соответствии с любым из параграфов 33-38, где конверсию гомофарнезола в (-)-амброксид проводят с использованием одного или более чем одного из реакционных условий для фермента SHC/HAC дикого типа или производного фермента SHC/HAC, представленных в таблице 24 или в таблице 24a, предпочтительно при pH в диапазоне от 5,0 до 6,2, предпочтительно при температуре 35°C.

40. Способ в соответствии с любым из параграфов 34-39, где соотношение SDS/клетки находится в диапазоне от 10:1 до 20:1, предпочтительно составляет 16:1 при соотношении биокатализатора и EEH, которое составляет приблизительно 2:1.

41. Способ в соответствии с любым из параграфов 34-40, где массовое соотношение биокатализатора к гомофарнезолу находится в диапазоне от приблизительно 0,5-2:1, предпочтительно составляет приблизительно 1:1 или 0,5:1.

42. Способ в соответствии с любым из параграфов 34-41, где стадии выращивания клеток и реакции биоконверсии проводят в одном реакционном сосуде.

43. Способ в соответствии с любым из параграфов 33-42, где гомофарнезол содержит или по существу состоит из смеси стереоизомеров EE:EZ в массовых соотношениях, выбранных из группы, состоящей из: 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; и 60:40.

44. Способ в соответствии с параграфом 43, где гомофарнезол содержит или по существу состоит из смеси стереоизомеров EE:EZ в массовых соотношениях, выбранных из группы, состоящей из: EE:EZ 92:08; EE:EZ 90:10; EE:EZ 80:20; EE:EZ 86:14; EE:EZ 70:30; EE:EZ 69:31; и EE:EZ 66:34.

45. Способ в соответствии с параграфом 43 или с параграфом 44, где гомофарнезол содержит или по существу состоит из смеси стереоизомеров EE:EZ в массовом соотношении 80:20.

46. Способ в соответствии с любым из параграфов 33-45, где (-)-амброксид получают с примесью по меньшей мере одного или более чем одного из побочных продуктов (II), (IV) и/или (III).

47. Способ в соответствии с любым из параграфов 33-46, где (-)-амброксид выделяют из биоконверсионной реакционной смеси с использованием стадии экстракции/дистилляции органическим растворителем или паром или фильтрования.

48. Способ в соответствии с любым из параграфов 33-47, где (-)-амброксид выделяют из твердой фазы биоконверсионной реакционной смеси с использованием стадии экстракции/дистилляции органическим растворителем или паром.

49. Способ в соответствии с параграфом 47 или с параграфом 48, где (-)-амброксид выделяют из реакционной смеси с использованием органического растворителя.

50. Способ в соответствии с параграфом 49, где (-)-амброксид выделяют из реакционной смеси с использованием этанола или толуола.

51. Способ в соответствии с любым из параграфов 47-49, где (-)-амброксид селективно кристаллизуют с использованием органического растворителя.

52. Способ в соответствии с параграфом 51, где (-)-амброксид по существу не содержит побочных продуктов (II), (IV) и/или (III).

53. Способ в соответствии с любым из параграфов 33-52, где (-)-амброксид получают в диапазоне концентраций приблизительно 125-200 г/л.

54. Способ по любому из параграфов 33-53, где (-)-амброксид имеет порог восприятия запаха от приблизительно 0,1 до приблизительно 0,5 нг/л.

Дополнительные аспекты изобретения

1. Производное сквален-гопен-циклазы (SHC)/гомофарнезол-амброксид-циклазы (HAC), содержащее аминокислотную последовательность, имеющую 1-50 мутаций, независимо выбранных из замен, делеций или инсерций, относительно SEQ ID NO: 1.

2. Производное SHC/HAC в соответствии с параграфом 1, где производное SHC содержит аминокислотную последовательность, имеющую от 1 до 40 мутаций, 1-30 мутаций, 1-20 мутаций, 1-10 мутаций или 1-6 мутаций относительно SEQ ID NO: 1.

3. Производное SHC/HAC в соответствии с параграфом 1 или с параграфом 2, где производное SHC содержит аминокислотную последовательность, обладающую по меньшей мере 40% идентичностью, по меньшей мере 50% идентичностью, или по меньшей мере 60% идентичностью, или по меньшей мере 70% идентичностью, или по меньшей мере 80% идентичностью, или по меньшей мере 90% идентичностью, или по меньшей мере 95% идентичностью, или по меньшей мере 96% идентичностью, или по меньшей мере 97% идентичностью, или по меньшей мере 98% идентичностью, или по меньшей мере 99% идентичностью относительно SEQ ID NO: 1.

4. Производное SHC/HAC в соответствии с параграфом 3, где вариант SHC содержит аминокислотную последовательность, обладающую по меньшей мере 95% идентичностью SEQ ID NO: 1.

5. Производное SHC/HAC, содержащее 1-10 мутаций, независимо выбранных из замен, делеций или инсерций относительно SEQ ID NO: 1, где одна или более мутаций, отличающихся от мутации активного центра SHC, локализована/локализованы в домене 2 фермента SHC (Фиг. 19 и/или 20).

6. Производное SHC/HAC в соответствии с любым из параграфов 1-5, где одна или более мутаций относительно SEQ ID NO: 1 выбраны из таблицы 1, где, если выбрана только одна мутация, она представляет собой не F601Y.

7. Производное SHC/HAC по параграфу 6, где по меньшей мере 2, 3, 4, 5, 6, 7, 8, 9 или 10 мутаций выбраны из таблицы 1 или таблицы 5.

8. Производное SHC/HAC по параграфу 2, содержащее аминокислотную последовательность, которая имеет вплоть до 6 мутаций относительно SEQ ID NO: 1 и содержит по меньшей мере замены F601Y или M132R в комбинации с по меньшей мере одной или более из F129L и/или I432T.

9. Производное SHC/HAC в соответствии с параграфом 7, содержащее аминокислотную последовательность, имеющую вплоть до 8 изменений аминокислот относительно SEQ ID NO: 1 и, которое содержит одно или более чем одно изменение аминокислот в положении, выбранном из группы, состоящей из положений 77, 129, 132, 192, 224, 432, 579, 601 и 605 относительно SEQ ID NO: 1, где производное SHC/HAC обладает повышенной ферментативной активностью HAC относительно SEQ ID NO: 1.

10. Производное SHC/HAC в соответствии с параграфом 9, содержащее одну или более замен, выбранных из группы, состоящей из: T77A, F129L, M132R, I92V, A224V, I432T, Q579H, F601Y и/или F605W относительно SEQ ID NO: 1.

11. Производное SHC/HAC в соответствии с параграфом 10, содержащее F601Y.

12. Производное SHC/HAC в соответствии с параграфом 10, содержащее F129L.

13. Производное SHC/HAC в соответствии с параграфом 10, содержащее F601Y и F129L.

14. Производное SHC/HAC в соответствии с параграфом 10, содержащее M132R и I432T.

15. Производное SHC/HAC в соответствии с параграфом 14, дополнительно содержащее аминокислотную замену A224V.

16. Производное SHC/HAC в соответствии с параграфом 14, дополнительно содержащее F601Y.

17. Производное SHC/HAC в соответствии с параграфом 14, дополнительно содержащее F129L.

18. Производное SHC/HAC в соответствии с параграфом 17, дополнительно содержащее F601Y.

19. Производное SHC/HAC в соответствии с параграфом 11, дополнительно содержащее Q579H.

20. Производное SHC/HAC в соответствии с параграфом 10, содержащее T77A, и I92V, и F129L.

21. Производное SHC/HAC в соответствии с любым из параграфов 1-20, имеющее аминокислотную последовательность, выбранную из группы, состоящей из SEQ ID NO: 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39 и/или 171.

22. Изолированная нуклеотидная последовательность, кодирующая производное SHC в соответствии с любым из параграфов 1-21.

23. Изолированная нуклеотидная последовательность в соответствии с любым из параграфов 22, где нуклеотидная последовательность выбрана из группы, состоящей из SEQ ID NO: 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40 и/или 170.

24. Конструкция, содержащая нуклеотидную последовательность по параграфу 22 или параграфу 23.

25. Конструкция в соответствии с параграфом 24, содержащая промотор, функционально связанный с нуклеотидной последовательностью по параграфу 22 или 23.

26. Конструкция по параграфу 25, где промотор представляет собой индуцибельный или конститутивный промотор.

27. Вектор, содержащий конструкцию по любому из параграфов 24-26.

28. Вектор по параграфу 27, представляющий собой плазмиду.

29. Вектор в соответствии с параграфом 28, способный направлять экспрессию в клетках-хозяевах, выбранных из прокариотических, дрожжевых, растительных клеток-хозяев и клеток-хозяев насекомых.

30. Конструкция по любому из параграфов 24-26 или вектор в соответствии с любым из параграфов 27-29, где конструкция или вектор способны к интеграции в геном клетки-хозяина, выбранной из прокариотических, дрожжевых, растительных клеток-хозяев и клеток-хозяев насекомых.

31. Рекомбинантная клетка-хозяин, содержащая нуклеотидную последовательность в соответствии с параграфом 22 или 23, или конструкцию в соответствии с любым из параграфов 24-26 или 30, или вектор в соответствии с любым из параграфов 27-30.

32. Рекомбинантная клетка-хозяин в соответствии с параграфом 31, выбранная из группы прокариотических клеток-хозяев, состоящей из бактерий рода Escherichia, Streptomyces, Bacillus, Pseudomonas, Lactobacillus и Lactococcus.

33. Рекомбинантная клетка-хозяин по параграфу 32, представляющая собой клетку-хозяина E. coli.

34. Рекомбинантная клетка-хозяин по параграфу 33, в которой происходит гиперэкспрессия гена, кодирующего производное SHC/HAC.

35. Способ получения производного SHC/HAC в соответствии с любым из параграфов 1-21, включающий стадию культивирования одной или более чем одной рекомбинантной клетки-хозяина в соответствии с любым из параграфов 31-34 в условиях, обеспечивающих возможность продуцирования производного фермента SHC/HAC.

36. Способ по параграфу 35, где культивирование клеток проводят в условиях, приемлемых для продуцирования биокатализатора.

37. Способ получения (-)-амброксида, включающий конверсию гомофарнезола в (-)-амброксид с использованием рекомбинантной клетки-хозяина в соответствии с любым из параграфов 31-34 или с использованием рекомбинантной клетки-хозяина, содержащей SEQ ID NO: 169 или SEQ ID NO: 165, кодирующей WT SHC/HAC, где при использовании WT SHC/HAC биоконверсию гомофарнезола в (-)-амброксид проводят с солюбилизирующим агентом, отличающимся от Тритона X-100 или тауродезоксихолата.

38. Способ в соответствии с параграфом 37, где конверсию гомофарнезола в (-)-амброксид проводят в подходящих реакционных условиях биоконверсии для WT SHC/HAC или производного фермента SHC/HAC.

39. Способ в соответствии с параграфом 37 или 38, где конверсию гомофарнезола в (-)-амброксид проводят при подходящих значениях pH, температуры, концентраций солюбилизирующего агента для WT SHC/HAC или производного фермента SHC/HAC.

40. Способ в соответствии с параграфом 39, где конверсию гомофарнезола в (-)-амброксид проводят при температуре в диапазоне от 30°C до 60°C, при pH в диапазоне приблизительно 4-8 и в присутствии солюбилизирующего агента, отличающегося от Тритона X-100 или тауродезоксихолата, для фермента WT SHC/HAC.

41. Способ в соответствии с любым из параграфов 37-40, где конверсию гомофарнезола в (-)-амброксид проводят с использованием одного или более одного из реакционных условий для фермента WT SHC/HAC или производного фермента SHC/HAC, представленных в таблице 24 или в таблице 24a.

42. Способ по любому из параграфов 37-41, где массовое соотношение биокатализатора к гомофарнезолу находится в диапазоне от приблизительно 0,5:1-2:1, предпочтительно составляет приблизительно 1:1 или 0,5:1.

43. Способ в соответствии с любым из параграфов 37-42, где стадии выращивания клеток и реакции биоконверсии проводят в одном реакционном сосуде.

44. Способ в соответствии с любым из параграфов 37-43, где субстрат гомофарнезол содержит один или более стереоизомеров гомофарнезола.

45. Способ по параграфу 44, где субстрат гомофарнезол содержит два стереоизомера гомофарнезола.

46. Способ по параграфу 45, где субстрат гомофарнезол содержит стереоизомеры EE:EZ.

47. Способ в соответствии с любым из параграфов 44-46, где гомофарнезол содержит смесь стереоизомеров EE:EZ в массовых соотношениях, выбранных из группы, состоящей из: 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 и 70:30.

48. Способ по параграфу 47, где гомофарнезол содержит смесь стереоизомеров EE: EZ в массовом соотношении, выбранном из группы, состоящей из: EE:EZ 90:10; EE:EZ 80:20; EE:EZ 86:14; EE:EZ 70:30; EE:EZ 69:31; и EE:EZ 66:34.

49. Способ по параграфу 35 или 36, где гомофарнезол содержит смесь стереоизомеров EE:EZ в массовом соотношении 80:20.

50. Способ по любому из параграфов 37-49, где (-)-амброксид получают с примесью по меньшей мере одного или более чем одного из побочных продуктов (II), (IV) и/или (III).

51. Способ в соответствии с любым из параграфов 37-50, где (-)-амброксид выделяют из биоконверсионной реакционной смеси с использованием стадии экстракции/дистилляции органическим растворителем или паром, или выделяют непосредственно кристаллы (-)-амброксида из биоконверсионной реакционной смеси с помощью фильтрования.

52. Способ в соответствии с параграфом 51, где (-)-амброксид выделяют из реакционной смеси с использованием органического растворителя.

53. Способ по параграфу 52, где (-)-амброксид селективно кристаллизуют с использованием органического растворителя.

54. Способ по параграфу 52 или 53, где (-)-амброксид по существу не содержит побочных продуктов (II), (IV) и/или (III).

55. (-)-Амброксид, получаемый способом по любому из параграфов 51-54.

56. (-)-Амброксид по параграфу 55 в твердой форме, предпочтительно в аморфной или кристаллической форме.

57. Способ получения продукта, содержащего (-)-амброксид, включающий включение в продукт (-)-амброксида по параграфу 55 или 56, предпочтительно в ароматизирующий продукт, косметический продукт, чистящий продукт, моющий продукт или мыльный продукт.

58. Ароматизирующий продукт, или косметический продукт, или потребительский продукт для ухода, содержащие (-)-амброксид по параграфу 55 или 56.

59. Ароматизирующая композиция, или косметическая композиция, или потребительская композиция для ухода, содержащая (-)-амброксид по параграфу 55 или 56 и один или более чем один дополнительный компонент.

60. Применение (-)-амброксида по параграфу 55 или 56 в качестве составной части ароматизирующего продукта, или косметического продукта, или потребительского продукта для ухода, такого как продукт для ухода за тканями, гигиено-косметический продукт, продукт для заботы о красоте и/или чистящий продукт.

61. Применение производного фермента SHC/HAC по любому из параграфов 1-21, нуклеотидной последовательности в соответствии с параграфами 22 или 23, конструкции в соответствии с любым из параграфов 22-26 или 30, вектора в соответствии с любым из параграфов 27-30 или рекомбинантной клетки-хозяина в соответствии с любым из параграфов 31-34 или рекомбинантной клетки-хозяина, экспрессирующей WT SHC/HAC, для биоконверсии гомофарнезола в (-)-амброксид, где фермент SHC/HAC применяют для реакции биоконверсии с солюбилизирующим агентом, отличающимся от Тритона X-100.

62. Способ получения (-)-амброксида или смеси стереоизомеров (-)-амброксида, где (3E,7E)-гомофарнезол (EEH) или смесь стереоизомеров (3E,7E)-гомофарнезола ферментативно конвертируют с получением (-)-амброксида или смеси стереоизомеров (-)-амброксида, где ферментативную конверсию осуществляют в реакционных условиях,

приемлемых для получения (-)-амброксида, и при проведении реакции в присутствии солюбилизирующего агента Тритон X-100 не используют в комбинации с ферментом WT SHC/HAC.

63. Способ в соответствии с параграфом 62, который осуществляют с использованием фермента SHC/HAC, выбранного из группы, состоящей из AacSHC (SEQ ID NO: 1), ZmoSHC1 (SEQ ID NO: 2), ZmoSHC2, (SEQ ID NO: 3); BjpSHC (SEQ ID NO: 4), производного фермента SHC/HAC, выбранного из таблицы 1, таблицы 5, таблицы 2, таблицы 6, таблицы 3, таблицы 7, таблицы 4 и/или таблицы 8 или последовательности с по меньшей мере 30% идентичностью, по меньшей мере 40% идентичностью, по меньшей мере 50% идентичностью, или по меньшей мере 60% идентичностью, или по меньшей мере 70% идентичностью, или по меньшей мере 80% идентичностью, или по меньшей мере 90% идентичностью, или по меньшей мере 95% идентичностью, или по меньшей мере 96% идентичностью, или по меньшей мере 97% идентичностью, или по меньшей мере 98% идентичностью, или по меньшей мере 99% идентичностью с SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3 и/или SEQ ID NO: 4.

64. Способ в соответствии с параграфом 63, где конверсию гомофарнезола в (-)-амброксид проводят при температуре в диапазоне от 30°C до 60°C, при pH в диапазоне 4-8 и в присутствии солюбилизирующего агента, отличающегося от Тритона X-100, для фермента WT SHC.

65. Способ в соответствии с параграфом 64, где реакционные условия для WT SHC/HAC или каждого производного SHC/HAC представлены в таблице 24 или в таблице 24a.

66. Способ в соответствии с любым из параграфов 62-65, включающий (a) культивирование одной или более чем одной рекомбинантной клетки-хозяина, экспрессирующей фермент WT SHC или производное фермента SHC в условиях, обеспечивающих возможность экспрессии полипептида WT SHC или производного SHC/HAC, перед конверсией E, E-гомофарнезола в (-)-амброксид.

67. Способ в соответствии с параграфом 66, где стадию культивирования и последующую стадию конверсии проводят в одном и том же реакционном сосуде в различных реакционных условиях.

68. Способ в соответствии с параграфом 67, где стадию культивирования осуществляют в диапазоне pH от приблизительно 6 до приблизительно 7, а стадию конверсии гомофарнезола в (-)-амброксид осуществляют в диапазоне pH приблизительно 4,8-5,5.

69. Способ в соответствии с любым из параграфов 62-68, где субстрат гомофарнезол содержит стереоизомеры EE:EZ.

70. Способ по параграфу 69, где гомофарнезол содержит смесь стереоизомеров EE:EZ в массовом соотношении, выбранном из группы, состоящей из: EE:EZ 90:10; EE:EZ 80:20; EE:EZ 86:14; EE:EZ 70:30; EE:EZ 69:31; и EE:EZ 66:34.

71. Способ по параграфу 70, где гомофарнезол содержит EE:EZ в массовом соотношении 80:20.

72. Способ по любому из параграфов 62-71, где (-)-амброксид получают с примесью по меньшей мере одного или более чем одного из побочных продуктов (II), (IV) и/или (III).

73. Способ по любому из параграфов 62-72, где (-)-амброксид выделяют из реакционной смеси с использованием стадии экстракции/дистилляции органическим растворителем или паром или фильтрования.

74. Способ в соответствии с параграфом 73, где (-)-амброксид выделяют из

реакционной смеси с использованием органического растворителя.

75. Способ по параграфу 74, где (-)-амброксид селективно кристаллизуют с использованием органического растворителя.

76. Способ по параграфу 74 или 75, где (-)-амброксид по существу не содержит побочных продуктов (II), (IV) и/или (III).

77. (-)-Амброксид, получаемый способом по любому из параграфов 72-76.

78. (-)-Амброксид по параграфу 26 в твердой форме, предпочтительно в аморфной или кристаллической форме.

79. Способ получения продукта, включающий включение в продукт (-)-амброксида по любому из параграфа 77 или параграфа 78.

80. Способ по параграфу 79, где продукт представляет собой ароматизирующий продукт, косметический продукт, чистящий продукт, моющий продукт или мыльный продукт.

81. Ароматизирующий продукт, или косметический продукт, или потребительский продукт для ухода, содержащий (-)-амброксид по параграфу 77 или 78.

82. Ароматизирующая композиция, или косметическая композиция, или потребительская композиция для ухода, содержащая (-)-амброксид по параграфу 77 или 78 и один или более чем один дополнительный компоненто.

83. Применение (-)-амброксида по параграфу 77 или 78 в качестве составной части ароматизирующего продукта или потребительского косметического продукта для ухода.

Дополнительные аспекты изобретения (ZmoSHC1)

1. Производное сквален-гопен-циклазы (SHC)/гомофарнезол-амброксид-циклазы (HAC), содержащее аминокислотную последовательность, имеющую 1-50 мутаций, независимо выбранных из замен, делеций или инсерций, относительно SEQ ID NO: 2.

2. Производное SHC/HAC в соответствии с параграфом 1, где производное SHC содержит аминокислотную последовательность, имеющую от 1 до 40 мутаций, 1-30 мутаций, 1-20 мутаций, 1-10 мутаций или 1-6 мутаций относительно SEQ ID NO: 2.

3. Производное SHC/HAC в соответствии с параграфом 1 или с параграфом 2, где производное SHC содержит аминокислотную последовательность, обладающую по меньшей мере 40% идентичностью, по меньшей мере 50% идентичностью, или по меньшей мере 60% идентичностью, или по меньшей мере 70% идентичностью, или по меньшей мере 80% идентичностью, или по меньшей мере 90% идентичностью, или по меньшей мере 95% идентичностью, или по меньшей мере 96% идентичностью, или по меньшей мере 97% идентичностью, или по меньшей мере 98% идентичностью, или по меньшей мере 99% идентичностью относительно SEQ ID NO: 2.

4. Производное SHC/HAC в соответствии с параграфом 3, где вариант SHC содержит аминокислотную последовательность, обладающую по меньшей мере 95% идентичностью относительно SEQ ID NO: 2.

5. Производное SHC/HAC, содержащее 1-10 мутаций, независимо выбранных из замен, делеции или инсерций относительно SEQ ID NO: 2, где одна или более чем одна мутация, отличающаяся от мутации активного центра SHC, локализована/локализованы в домене 2 фермента SHC (Фиг. 19 и/или 20).

6. Производное SHC/HAC в соответствии с любым из параграфов 15, где одна или более мутаций относительно SEQ ID NO: 2 выбраны из таблицы 2, где, если выбрана только одна мутация, она представляет собой не F668Y.

7. Производное SHC/HAC по параграфу 6, где по меньшей мере 2, 3, 4, 5, 6, 7, 8, 9 или 10 мутаций выбраны из таблицы 2 и/или таблицы 6.

8. Производное SHC/HAC по параграфу 2, содержащее аминокислотную последовательность, которая имеет вплоть до 6 мутаций относительно SEQ ID NO: 2 и содержит по меньшей мере замены F668Y или Y185R в комбинации с по меньшей мере одной или более из F182L и/или I498T.

9. Производное SHC/HAC по параграфу 7, содержащее аминокислотную последовательность, имеющую вплоть до 8 изменений аминокислот относительно SEQ ID NO: 2, и которое содержит одно или более чем одно изменение аминокислот в положении, выбранном из группы, состоящей из положений 129, 145, 182, 185, 282, 498, 647 и 668 относительно SEQ ID NO: 2, где производное SHC/HAC обладает повышенной ферментативной активностью HAC относительно SEQ ID NO: 2.

10. Производное SHC/HAC в соответствии с параграфом 9, содержащее одну или более замен, выбранных из группы, состоящей из: S129A, V145V, F182L, Y185R, G282V, I498T, H646H и F668Y относительно SEQ ID NO: 2.

11. Производное SHC/HAC в соответствии с параграфом 10, содержащее F668Y.

12. Производное SHC/HAC в соответствии с параграфом 10, содержащее F182L.

13. Производное SHC/HAC в соответствии с параграфом 10, содержащее F668Y и F182L.

14. Производное SHC/HAC в соответствии с параграфом 10, содержащее Y185R и I498T.

15. Производное SHC/HAC в соответствии с параграфом 14, дополнительно содержащее G282V.

16. Производное SHC/HAC в соответствии с параграфом 14, дополнительно содержащее F668Y.

17. Производное SHC/HAC в соответствии с параграфом 14, дополнительно содержащее F182L.

18. Производное SHC/HAC в соответствии с параграфом 17, дополнительно содержащее F668Y.

19. Производное SHC/HAC в соответствии с параграфом 11, дополнительно содержащее H646H.

20. Производное SHC в соответствии с параграфом 10, содержащее S129A, и V145V, и F182L.

21. Производное SHC/HAC в соответствии с любым из параграфов 1-20, имеющее аминокислотную последовательность, выбранную из группы, состоящей из SEQ ID NO: 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73 и/или 75.

22. Изолированная нуклеотидная последовательность, кодирующая производное SHC в соответствии с любым из параграфов 1-21.

23. Изолированная нуклеотидная последовательность в соответствии с параграфом 22, где нуклеотидная последовательность выбрана из группы, состоящей из SEQ ID NO: 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74 и/или 76.

24. Конструкция, содержащая нуклеотидную последовательность по параграфу 22 или параграфу 23.

25. Конструкция в соответствии с параграфом 24, содержащая промотор, функционально связанный с нуклеотидной последовательностью по параграфу 22 или 23.

26. Конструкция по параграфу 25, где промотор представляет собой индуцибельный или конститутивный промотор.

27. Вектор, содержащий конструкцию в соответствии с любым из параграфов 24-26.

28. Вектор по параграфу 27, представляющий собой плазмиду.

29. Вектор в соответствии с параграфом 28, способный направлять экспрессию в клетках-хозяевах, выбранных из прокариотических, дрожжевых, растительных клеток-хозяев и клеток-хозяев насекомых.

30. Конструкция по любому из параграфов 24-26 или вектор в соответствии с любым из параграфов 27-29, где конструкция или вектор способны к интеграции в геном клетки-хозяина, выбранной из прокариотических, дрожжевых, растительных клеток-хозяев и клеток-хозяев насекомых.

31. Рекомбинантная клетка-хозяин, содержащая нуклеотидную последовательность в соответствии с параграфом 22 или 23, или конструкцию в соответствии с любым из параграфов 24-26 или 30, или вектор в соответствии с любым из параграфов 27-30.

32. Рекомбинантная клетка-хозяин в соответствии с параграфом 31, выбранная из группы прокариотических клеток-хозяев, состоящей из бактерий рода Escherichia, Streptomyces, Bacillus, Pseudomonas, Lactobacillus и Lactococcus.

33. Рекомбинантная клетка-хозяин по параграфу 32, представляющая собой клетку-хозяина E. coli.

34. Рекомбинантная клетка-хозяин по параграфу 33, в которой происходит гиперэкспрессия гена, кодирующего производное SHC/HAC.

35. Способ получения производного SHC/HAC в соответствии с любым из параграфов 1-21, включающий стадию культивирования одной или более чем одной рекомбинантной клетки-хозяина в соответствии с любым из параграфов 31-34 в условиях, обеспечивающих возможность продуцирования производного фермента SHC/HAC.

36. Способ по параграфу 35, где культивирование клеток проводят в условиях, приемлемых для продуцирования био катализатор а.

37. Способ получения (-)-амброксида, включающий конверсию гомофарнезола в (-)-амброксид с использованием рекомбинантной клетки-хозяина в соответствии с любым из параграфов 31-34 или с использованием рекомбинантной клетки-хозяина, содержащей SEQ ID NO: 166, кодирующей WT SHC/HAC, где при использовании WT SHC/HAC биоконверсию гомофарнезола в (-)-амброксид проводят с солюбилизирующим агентом, отличающимся от Тритона X-100.

38. Способ в соответствии с параграфом 37, где конверсию гомофарнезола в (-)-амброксид проводят в подходящих реакционных условиях биоконверсии для WT SHC/HAC или производного фермента SHC/HAC.

39. Способ в соответствии с параграфом 37 или 38, где конверсию гомофарнезола в (-)-амброксид проводят при подходящих значениях pH, температуры, концентраций солюбилизирующего агента для WT SHC/HAC или производного фермента SHC/HAC.

40. Способ в соответствии с параграфом 39, где конверсию гомофарнезола в (-)-амброксид проводят при температуре в диапазоне от 30°C до 60°C, при pH в диапазоне приблизительно 4-8 и в присутствии солюбилизирующего агента, отличающегося от Тритона X-100, для фермента WT SHC/HAC.

41. Способ в соответствии с любым из параграфов 37 40, где конверсию гомофарнезола в (-)-амброксид проводят с использованием одного или более чем одного из реакционных условий для фермента WT SHC/HAC или производного фермента SHC/HAC, представленных в таблице 24 или в таблице 24a.

42. Способ по любому из параграфов 37-41, где массовое соотношение биокатализатора к гомофарнезолу находится в диапазоне от приблизительно 0,5:1-2:1, предпочтительно составляет приблизительно 1:1 или 0,5:1.

43. Способ в соответствии с любым из параграфов 37-42, где стадии выращивания клеток и реакции биоконверсии проводят в одном реакционном сосуде.

44. Способ в соответствии с любым из параграфов 37-43, где субстрат гомофарнезол содержит один или более чем один стереоизомер гомофарнезола.

45. Способ по параграфу 44, где субстрат гомофарнезол содержит два стереоизомера гомофарнезола.

46. Способ по параграфу 45, где субстрат гомофарнезол содержит стереоизомеры EE:EZ.

47. Способ в соответствии с любым из параграфов 44 46, где гомофарнезол содержит смесь стереоизомеров EE:EZ в массовых соотношениях, выбранных из группы, состоящей из: 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 и 70:30.

48. Способ по параграфу 47, где гомофарнезол содержит смесь стереоизомеров EE:EZ в массовом соотношении, выбранном из группы, состоящей из: EE:EZ 90:10; EE:EZ 80:20; EE:EZ 86:14; EE:EZ 70:30; EE:EZ 69:31; и EE:EZ 66:34.

49. Способ по параграфу 35 или 36, где гомофарнезол содержит смесь стереоизомеров EE:EZ в массовом соотношении 80:20.

50. Способ по любому из параграфов 37-49, где (-)-амброксид получают с примесью по меньшей мере одного или более чем одного из побочных продуктов (II), (IV) и/или (III).

51. Способ по любому из параграфов 37-50, где (-)-амброксид выделяют из биоконверсионной реакционной смеси с использованием этапа экстракции/дистилляции органическим растворителем или паром или фильтрования.

52. Способ в соответствии с параграфом 51, где (-)-амброксид выделяют из реакционной смеси с использованием органического растворителя.

53. Способ по параграфу 52, где (-)-амброксид селективно кристаллизуют с использованием органического растворителя.

54. Способ по параграфу 52 или 53, где (-)-амброксид по существу не содержит побочных продуктов (II), (IV) и/или (III).

55. (-)-Амброксид, получаемый способом по любому из параграфов 51-54.

56. (-)-Амброксид по параграфу 55 в твердой форме, предпочтительно в аморфной или кристаллической форме.

57. Способ получения продукта, содержащего (-)-амброксид, включающий включение в продукт (-)-амброксида по параграфу 55 или 56, предпочтительно в ароматизирующий продукт, косметический продукт, чистящий продукт, моющий продукт или мыльный продукт.

58. Ароматизирующий продукт, или косметический продукт, или потребительский продукт для ухода, содержащее (-)-амброксид по параграфу 55 или 56.

59. Ароматизирующая композиция, или косметическая композиция, или потребительская композиция для ухода, содержащая (-)-амброксид по параграфу 55 или 56 и один или более чем один дополнительный компонент.

60. Применение (-)-амброксида по параграфу 55 или 56 в качестве составной части ароматизирующего продукта, или косметического продукта, или потребительского продукта для ухода, такого как продукт для ухода за тканями, гигиено-косметический продукт, продукт для заботы о красоте и/или чистящий продукт.

61. Применение производного фермента SHC/HAC по любому из параграфов 1-21, нуклеотидной последовательности в соответствии с параграфами 22 или 23, конструкции в соответствии с любым из параграфов 22-26 или 30, вектора в соответствии с любым из параграфов 27-30 или рекомбинантной клетки-хозяина в соответствии с любым из

параграфов 31 34 или рекомбинантной клетки-хозяина, экспрессирующей WT SHC/HAC, для биоконверсии гомофарнезола в (-)-амброксид, где фермент SHC/HAC используют для реакции биоконверсии с солюбилизирующим агентом, отличающимся от Тритона X-100.

Дополнительные аспекты изобретения (ZmoSHC2)

1. Производное сквален-гопен-циклазы (8HC)/гомофарнезол-амброксид-циклазы (HAC), содержащее аминокислотную последовательность, имеющую 1-50 мутаций, независимо выбранных из замен, делеций или инсерций, относительно SEQ ID NO: 3.

2. Производное SHC/HAC в соответствии с параграфом 1, где производное SHC содержит аминокислотную последовательность, имеющую от 1 до 40 мутаций, 1 30 мутаций, 1 20 мутаций, 1-10 мутаций или 1-6 мутаций относительно SEQ ID NO: 3.

3. Производное SHC/HAC в соответствии с параграфом 1 или с параграфом 2, где производное SHC содержит аминокислотную последовательность, обладающую по меньшей мере 40% идентичностью, по меньшей мере 50% идентичностью, или по меньшей мере 60% идентичностью, или по меньшей мере 70% идентичностью, или по меньшей мере 80% идентичностью, или по меньшей мере 90% идентичностью, или по меньшей мере 95% идентичностью, или по меньшей мере 96% идентичностью, или по меньшей мере 97% идентичностью, или по меньшей мере 98% идентичностью, или по меньшей мере 99% идентичностью относительно SEQ ID NO: 3.

4. Производное SHC/HAC в соответствии с параграфом 3, где вариант SHC содержит аминокислотную последовательность, обладающую по меньшей мере 95% идентичностью относительно SEQ ID NO: 3.

5. Производное SHC/HAC, содержащее 1-10 мутаций, независимо выбранных из замен, делеций или инсерций относительно SEQ ID NO: 3, где одна или более чем одна мутация, отличающаяся от мутации активного центра SHC, локализована в домене 2 фермента SHC (Фиг. 19 и/или 20).

6. Производное SHC/HAC в соответствии с любым из параграфов 1-5, где одна или более чем одна мутация относительно SEQ ID NO: 3 выбрана из таблицы 3, где, если выбрана только одна мутация, она не является F620Y.

7. Производное SHC/HAC по параграфу 6, где по меньшей мере 2, 3, 4, 5, 6, 7, 8, 9 или 10 мутаций выбраны из таблицы 3 и/или таблицы 7.

8. Производное SHC/HAC по параграфу 2, содержащее аминокислотную последовательность, которая имеет вплоть до 6 мутаций относительно SEQ ID NO: 3 и содержит по меньшей мере замены F620Y или I140R в комбинации с по меньшей мере одной или более из F137L и/или I450T.

9. Производное SHC/HAC по параграфу 7, содержащее аминокислотную последовательность, имеющую вплоть до 8 изменений аминокислот относительно SEQ ID NO: 3 и содержащую одно или более чем одно изменение аминокислот в положении, выбранном из группы, состоящей из положений 85, 100, 137, 140, 233, 450, 598 и 620 относительно SEQ ID NO: 3, где производное SHC/HAC обладает повышенной ферментативной активностью HAC относительно SEQ ID NO: 3.

10. Производное SHC/HAC в соответствии с параграфом 9, содержащее одну или более чем одну замену, выбранную из группы, состоящей из: G85A, V100V, F137L, I140R, V233V, I450T, N598H и F620Y относительно SEQ ID NO: 3.

11. Производное SHC/HAC в соответствии с параграфом 10, содержащее F620Y.

12. Производное SHC/HAC в соответствии с параграфом 10, содержащее F137L.

13. Производное SHC/HAC в соответствии с параграфом 10, содержащее F620Y HF137L.

14. Производное SHC/HAC в соответствии с параграфом 10, содержащее I140R и I450T.

15. Производное SHC/HAC в соответствии с параграфом 14, дополнительно содержащее V233V.

16. Производное SHC/HAC в соответствии с параграфом 14, дополнительно содержащее F620Y.

17. Производное SHC/HAC в соответствии с параграфом 14, дополнительно содержащее F137L.

18. Производное SHC/HAC в соответствии с параграфом 17, дополнительно содержащее F620Y.

19. Производное SHC/HAC в соответствии с параграфом 11, дополнительно содержащее N598H.

20. Производное SHC в соответствии с параграфом 10, содержащее G85A, и V100V, HF137L.

21. Производное SHC/HAC в соответствии с любым из параграфов 1 20, имеющее аминокислотную последовательность, выбранную из группы, состоящей из SEQ ID NO: 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109 и/или 111.

22. Изолированная нуклеотидная последовательность, кодирующая производное SHC в соответствии с любым из параграфов 121.

23. Изолированная нуклеотидная последовательность в соответствии с параграфом 22, где нуклеотидная последовательность выбрана из группы, состоящей из SEQ ID NO: 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110 и/или 112.

24. Конструкция, содержащая нуклеотидную последовательность по параграфу 22 или параграфу 23.

25. Конструкция в соответствии с параграфом 24, содержащая промотор, функционально связанный с нуклеотидной последовательностью по параграфу 22 или 23.

26. Конструкция по параграфу 25, где промотор представляет собой индуцибельный или конститутивный промотор.

27. Вектор, содержащий конструкцию в соответствии с любым из параграфов 24-26.

28. Вектор по параграфу 27, представляющий собой плазмиду.

29. Вектор в соответствии с параграфом 28, способный направлять экспрессию в клетках-хозяевах, выбранных из прокариотических, дрожжевых, растительных клеток-хозяев и клеток-хозяев насекомых.

30. Конструкция по любому из параграфов 24-26 или вектор в соответствии с любым из параграфов 27-29, где конструкция или вектор способны к интеграции в геном клетки-хозяина, выбранной из прокариотических, дрожжевых, растительных клеток-хозяев и клеток-хозяев насекомых.

31. Рекомбинантная клетка-хозяин, содержащая нуклеотидную последовательность в соответствии с параграфом 22 или 23, или конструкцию в соответствии с любым из параграфов 24-26 или 30, или вектор в соответствии с любым из параграфов 27-30.

32. Рекомбинантная клетка-хозяин в соответствии с параграфом 31, выбранная из группы прокариотических клеток-хозяев, состоящей из бактерий рода Escherichia, Streptomyces, Bacillus, Pseudomonas, Lactobacillus и Lactococcus.

33. Рекомбинантная клетка-хозяин по параграфу 32, представляющая собой клетку-хозяина E. coli.

34. Рекомбинантная клетка-хозяин по параграфу 33, в которой происходит гиперэкспрессия гена, кодирующего производное SHC/HAC.

35. Способ получения производного SHC/HAC в соответствии с любым из параграфов 121, включающий стадии: (a) культивирование одной или более чем одной рекомбинантной клетки-хозяина в соответствии с любым из параграфов 31 34 в условиях, обеспечивающих возможность продуцирования производного фермента SHC/HAC.

36. Способ по параграфу 35, где культивирование клеток проводят в условиях, приемлемых для продуцирования биокатализатора.

37. Способ получения (-)-амброксида, включающий конверсию гомофарнезола в (-)-амброксид с использованием рекомбинантной клетки-хозяина в соответствии с любым из параграфов 31 34 или с использованием рекомбинантной клетки-хозяина, содержащей SEQ ID NO: 167, кодирующей WT SHC/HAC, где при использовании WT SHC/HAC биоконверсию гомофарнезола в (-)-амброксид проводят с солюбилизирующим агентом, отличающимся от Тритона X-100.

38. Способ в соответствии с параграфом 37, где конверсию гомофарнезола в (-)-амброксид проводят в приемлемых реакционных условиях биоконверсии для WT SHC/HAC или производного фермента SHC/HAC.

39. Способ в соответствии с параграфом 37 или 38, где конверсию гомофарнезола в (-)-амброксид проводят при приемлемых значениях pH, температуры, концентраций солюбилизирующего агента для WT SHC/HAC или производного фермента SHC/HAC.

40. Способ в соответствии с параграфом 39, где конверсию гомофарнезола в (-)-амброксид проводят при температуре в диапазоне от 30°C до 60°C, при pH в диапазоне приблизительно 4-8 и в присутствии солюбилизирующего агента, отличающегося от Тритона X-100 для фермента WT SHC/HAC.

41. Способ в соответствии с любым из параграфов 37-40, где конверсию гомофарнезола в (-)-амброксид проводят с использованием одного или чем одного более из реакционных условий для фермента WT SHC/HAC или производного фермента SHC/HAC, представленных в таблице 24 или в таблице 24a.

42. Способ по любому из параграфов 37-41, где массовое соотношение биокатализатора к гомофарнезолу находится в диапазоне от приблизительно 0,5:1-2:1, предпочтительно приблизительно 1:1 или 0,5:1.

43. Способ в соответствии с любым из параграфов 37-42, где этапы выращивания клеток и реакции биоконверсии проводят в одном реакционном сосуде.

44. Способ в соответствии с любым из параграфов 37-43, где субстрат гомофарнезол содержит один или более чем один стереоизомер гомофарнезола.

45. Способ по параграфу 44, где субстрат гомофарнезол содержит два стереоизомера гомофарнезола.

46. Способ по параграфу 45, где субстрат гомофарнезол содержит стереоизомеры EE:EZ.

47. Способ в соответствии с любым из параграфов 44-46, где гомофарнезол содержит смесь стереоизомеров EE:EZ в массовых соотношениях, выбранных из группы, состоящей из: 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 и 70:30.

48. Способ по параграфу 47, где гомофарнезол содержит смесь стереоизомеров EE:EZ в массовом соотношении, выбранном из группы, состоящей из: EE:EZ 90:10; EE:EZ 80:20; EE:EZ 86:14; EE:EZ 70:30; EE:EZ 69:31; и EE:EZ 66:34.

49. Способ по параграфу 35 или 36, где гомофарнезол содержит смесь стереоизомеров EE:EZ в массовом соотношении 80:20.

50. Способ по любому из параграфов 37-49, где (-)-амброксид получают с примесью

по меньшей мере одного или более чем одного из побочных продуктов (II), (IV) и/или (III).

51. Способ по любому из параграфов 37-50, где (-)-амброксид выделяют из биоконверсионной реакционной смеси с использованием этапа экстракции/дистилляции органическим растворителем или паром или фильтрования.

52. Способ в соответствии с параграфом 51, где (-)-амброксид выделяют из реакционной смеси с использованием органического растворителя.

53. Способ по параграфу 52, где (-)-амброксид селективно кристаллизуют с использованием органического растворителя.

54. Способ по параграфу 52 или 53, где (-)-амброксид по существу не содержит побочных продуктов (II), (IV) и/или (III).

55. (-)-Амброксид, получаемый способом по любому из параграфов 51-54.

56. (-)-Амброксид по параграфу 55 в твердой форме, предпочтительно в аморфной или кристаллической форме.

57. Способ получения продукта, содержащего (-)-амброксид, включающий включение в продукт (-)-амброксида по параграфу 55 или 56, предпочтительно в ароматизирующий продукт, косметический продукт, чистящий продукт, моющий продукт или мыльный продукт.

58. Ароматизирующий продукт, или косметический продукт, или потребительский продукт для ухода, содержащее (-)-амброксид по параграфу 55 или 56.

59. Ароматизирующая композиция, или косметическая композиция, или потребительская композиция для ухода, содержащая (-)-амброксид по параграфу 55 или 56 и один или более дополнительных компонентов.

60. Применение (-)-амброксида по параграфу 55 или 56 в качестве составной части ароматизирующего продукта, или косметического продукта, или потребительского продукта для ухода, такого как продукт для ухода за тканями, гигиено-косметический продукт, продукт для заботы о красоте и/или чистящий продукт.

61. Применение производного фермента SHC/HAC по любому из параграфов 1-21, нуклеотидной последовательности в соответствии с параграфами 22 или 23, конструкции в соответствии с любым из параграфов 22-26 или 30, вектора в соответствии с любым из параграфов 27-30 или рекомбинантной клетки-хозяина в соответствии с любым из параграфов 31-34 или рекомбинантной клетки-хозяина, экспрессирующей WT SHC/HAC, для биоконверсии гомофарнезола в (-)-амброксид, где фермент SHC/HAC применяют для реакции биоконверсии с солюбилизирующим агентом, отличающимся от Тритона X-100.

Дополнительные аспекты изобретения (BJpSHC)

1. Производное сквален-гопен-циклазы (SHC)/гомофарнезол-амброксид-циклазы (HAC), содержащее аминокислотную последовательность, имеющую 1-50 мутаций, независимо выбранных из замен, делеций или инсерций, относительно SEQ ID NO: 4.

2. Производное SHC/HAC в соответствии с параграфом 1, где производное SHC содержит аминокислотную последовательность, имеющую от 1 до 40 мутаций, 1-30 мутаций, 1-20 мутаций, 1-10 мутаций или 1-6 мутаций относительно SEQ ID NO: 4.

3. Производное SHC/HAC в соответствии с параграфом 1 или с параграфом 2, где производное SHC содержит аминокислотную последовательность, обладающую по меньшей мере 40% идентичностью, по меньшей мере 50% идентичностью, или по меньшей мере 60% идентичностью, или по меньшей мере 70% идентичностью, или по меньшей мере 80% идентичностью, или по меньшей мере 90% идентичностью, или по меньшей мере 95% идентичностью, или по меньшей мере 96% идентичностью, или по

меньшей мере 97% идентичностью, или по меньшей мере 98% идентичностью, или по меньшей мере 99% идентичностью относительно SEQ ID NO: 4.

4. Производное SHC/HAC в соответствии с параграфом 3, где вариант SHC содержит аминокислотную последовательность, обладающую по меньшей мере 95% идентичностью относительно SEQ ID NO: 4.

5. Производное SHC/HAC, содержащее 1-10 мутаций, независимо выбранных из замен, делеций или инсерций относительно SEQ ID NO: 4, где одна или более чем одна мутация, отличающаяся от мутации активного центра SHC, локализована в домене 2 фермента SHC (Фиг. 19 и/или 20).

6. Производное SHC/HAC в соответствии с любым из параграфов 15, где одна или более мутаций относительно SEQ ID NO: 4 выбраны из таблицы 4, где, если выбрана только одна мутация, она не является F628Y.

7. Производное SHC/HAC по параграфу 6, где по меньшей мере 2, 3, 4, 5, 6, 7, 8, 9 или 10 мутаций выбраны из таблицы 4 и/или таблицы 8.

8. Производное SHC/HAC по параграфу 2, содержащее аминокислотную последовательность, которая имеет вплоть до 6 мутаций относительно SEQ ID NO: 4 и содержит по меньшей мере замены F628Y или I140R в комбинации с по меньшей мере одной или более из F137L и/или I450T.

9. Производное SHC/HAC по параграфу 7, содержащее аминокислотную последовательность, имеющую вплоть до 8 изменений аминокислот относительно SEQ ID NO: 4 и содержащую одно или более чем одно изменение аминокислот в положении, выбранном из группы, состоящей из положений 88, 104, 141, 144, 241, 459, 607 и 628 относительно SEQ ID NO: 4, где производное SHC/HAC обладает повышенной ферментативной активностью HAC относительно SEQ ID NO: 4.

10. Производное SHC/HAC в соответствии с параграфом 9, содержащее одну или более замен, выбранных из группы, состоящей из: A88A, V104V, F141L, Y144R, V241V, I459T, M607H и F628Y относительно SEQ ID NO: 4.

11. Производное SHC/HAC в соответствии с параграфом 10, содержащее F628Y.

12. Производное SHC/HAC в соответствии с параграфом 10, содержащее F141L.

13. Производное SHC/HAC в соответствии с параграфом 10, содержащее F628Y HF141L.

14. Производное SHC/HAC в соответствии с параграфом 10, содержащее Y144R и I459T.

15. Производное SHC/HAC в соответствии с параграфом 14, дополнительно содержащее V241V.

16. Производное SHC/HAC в соответствии с параграфом 14, дополнительно содержащее F628Y.

17. Производное SHC/HAC в соответствии с параграфом 14, дополнительно содержащее F141L.

18. Производное SHC/HAC в соответствии с параграфом 17, дополнительно содержащее F628Y.

19. Производное SHC/HAC в соответствии с параграфом 11, дополнительно содержащее M607H.

20. Производное SHC в соответствии с параграфом 10, содержащее S129A, и V145V, HF182L.

21. Производное SHC/HAC в соответствии с любым из параграфов 1 20, имеющее аминокислотную последовательность, выбранную из группы, состоящей из SEQ ID NO: 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145 и/или

147.

22. Изолированная нуклеотидная последовательность, кодирующая производное SHC в соответствии с любым из параграфов 1-21.

23. Изолированная нуклеотидная последовательность в соответствии с параграфом 22, где нуклеотидная последовательность выбрана из группы, состоящей из SEQ ID NO: 114, 116, 118, 120, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146 и/или 148.

24. Конструкция, содержащая нуклеотидную последовательность по параграфу 22 или параграфу 23.

25. Конструкция в соответствии с параграфом 24, содержащая промотор, функционально связанный с нуклеотидной последовательностью по параграфу 22 или 23.

26. Конструкция по параграфу 25, где промотор представляет собой индуцибельный или конститутивный промотор.

27. Вектор, содержащий конструкцию в соответствии с любым из параграфов 24-26.

28. Вектор по параграфу 27, представляющий собой плазмиду.

29. Вектор в соответствии с параграфом 28, способный направлять экспрессию в клетках-хозяевах, выбранных из прокариотических, дрожжевых, растительных клеток-хозяев и клеток-хозяев насекомых.

30. Конструкция по любому из параграфов 24-26 или вектор в соответствии с любым из параграфов 27-29, где конструкция или вектор способны к интеграции в геном клетки-хозяина, выбранной из прокариотических, дрожжевых, растительных клеток-хозяев и клеток-хозяев насекомых.

31. Рекомбинантная клетка-хозяин, содержащая нуклеотидную последовательность в соответствии с параграфом 22 или 23, или конструкцию в соответствии с любым из параграфов 24-26 или 30, или вектор в соответствии с любым из параграфов 27-30.

32. Рекомбинантная клетка-хозяин в соответствии с параграфом 31, выбранная из группы прокариотических клеток-хозяев, состоящей из бактерий рода Escherichia, Streptomyces, Bacillus, Pseudomonas, Lactobacillus и Lactococcus.

33. Рекомбинантная клетка-хозяин по параграфу 32, представляющая собой клетку-хозяина E. coli.

34. Рекомбинантная клетка-хозяин по параграфу 33, в которой происходит гиперэкспрессия гена, кодирующего производное SHC/HAC.

35. Способ получения производного SHC/HAC в соответствии с любым из параграфов 1-21, включающий стадии: (a) культивирование одной или более рекомбинантных клеток-хозяев в соответствии с любым из параграфов 31 34 в условиях, обеспечивающих возможность продуцирования производного фермента SHC/HAC.

36. Способ по параграфу 35, где культивирование клеток проводят в условиях, приемлемых для продуцирования био катализатор а.

37. Способ получения (-)-амброксида, включающий конверсию гомофарнезола в (-)-амброксид с использованием рекомбинантной клетки-хозяина в соответствии с любым из параграфов 31-34 или с использованием рекомбинантной клетки-хозяина, содержащей SEQ ID NO: 168, кодирующей WT SHC/HAC, где при использовании WT SHC/HAC биоконверсию гомофарнезола в (-)-амброксид проводят с солюбилизирующим агентом, отличающимся от Тритона X-100.

38. Способ в соответствии с параграфом 37, где конверсию гомофарнезола в (-)-амброксид проводят в приемлемых реакционных условиях биоконверсии для WT SHC/HAC или производного фермента SHC/HAC.

39. Способ в соответствии с параграфом 37 или 38, где конверсию гомофарнезола в (-)-амброксид проводят при приемлемых значениях pH, температуры, концентраций солюбилизирующего агента для WT SHC/HAC или производного фермента SHC/HAC.

40. Способ в соответствии с параграфом 39, где конверсию гомофарнезола в (-)-амброксид проводят при температуре в диапазоне от 30°C до 60°C, при pH в диапазоне приблизительно 4-8 и в присутствии солюбилизирующего агента, отличающегося от Тритона X-100, для фермента WT SHC/HAC.

41. Способ в соответствии с любым из параграфов 37 40, где конверсию гомофарнезола в (-)-амброксид проводят с использованием одного или более из реакционных условий для фермента WT SHC/HAC или производного фермента SHC/HAC, представленных в таблице 24 или в таблице 24a.

42. Способ по любому из параграфов 37-41, где массовое соотношение биокатализатора к гомофарнезолу находится в диапазоне от приблизительно 0,5:1-2:1, предпочтительно приблизительно 1:1 или 0,5:1.

43. Способ в соответствии с любым из параграфов 37-42, где этапы выращивания клеток и реакции биоконверсии проводят в одном реакционном сосуде.

44. Способ в соответствии с любым из параграфов 27-31, где субстрат гомофарнезол содержит один или более стереоизомеров гомофарнезола.

45. Способ по параграфу 44, где субстрат гомофарнезол содержит два стереоизомера гомофарнезола.

46. Способ по параграфу 45, где субстрат гомофарнезол содержит стереоизомеры EE:EZ.

47. Способ в соответствии с любым из параграфов 44-46, где гомофарнезол содержит смесь стереоизомеров EE:EZ в массовых соотношениях, выбранных из группы, состоящей из: 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 и 70:30.

48. Способ по параграфу 47, где гомофарнезол содержит смесь стереоизомеров EE:EZ в массовом соотношении, выбранном из группы, состоящей из: EE:EZ 90:10; EE:EZ 80:20; EE:EZ 86:14; EE:EZ 70:30; EE:EZ 69:31; и EE:EZ 66:34.

49. Способ по параграфу 35 или 36, где гомофарнезол содержит смесь стереоизомеров EE:EZ в массовом соотношении 80:20.

50. Способ по любому из параграфов 37-49, где (-)-амброксид получают с примесью по меньшей мере одного или более чем одного из побочных продуктов (II), (IV) и/или (III).

51. Способ по любому из параграфов 37-50, где (-)-амброксид выделяют из биоконверсионной реакционной смеси с использованием этапа экстракции/дистилляции органическим растворителем или паром или фильтрования.

52. Способ в соответствии с параграфом 51, где (-)-амброксид выделяют из реакционной смеси с использованием органического растворителя.

53. Способ по параграфу 52, где (-)-амброксид селективно кристаллизуют с использованием органического растворителя.

54. Способ по параграфу 52 или 53, где (-)-амброксид по существу не содержит побочных продуктов (II), (IV) и/или (III).

55. (-)-Амброксид, получаемый способом по любому из параграфов 51-54.

56. (-)-Амброксид по параграфу 55 в твердой форме, предпочтительно в аморфной или кристаллической форме.

57. Способ получения продукта, содержащего (-)-амброксид, включающий включение

в продукт (-)-амброксида по параграфу 55 или 56, предпочтительно в ароматизирующий продукт, косметический продукт, чистящий продукт, моющий продукт или мыльный продукт.

58. Ароматизирующий продукт, или косметический продукт, или потребительский продукт для ухода, содержащее (-)-амброксид по параграфу 55 или 56.

59. Ароматизирующая композиция, или косметическая композиция, или потребительская композиция для ухода, содержащая (-)-амброксид по параграфу 55 или 56 и один или более чем один дополнительный компонент.

60. Применение (-)-амброксида по параграфу 55 или 56 в качестве составной части ароматизирующего продукта, или косметического продукта, или потребительского продукта для ухода, такого как продукт для ухода за тканями, гигиено-косметический продукт, продукт для заботы о красоте и/или чистящий продукт.

61. Применение производного фермента SHC/HAC по любому из параграфов 1-21, нуклеотидной последовательности в соответствии с параграфами 22 или 23, конструкции в соответствии с любым из параграфов 22-26 или 30, вектора в соответствии с любым из параграфов 27-30 или рекомбинантной клетки-хозяина в соответствии с любым из параграфов 31-34 или рекомбинантной клетки-хозяина, экспрессирующей WT SHC/HAC, для биоконверсии гомофарнезола в (-)-амброксид, где фермент SHC/HAC применяют для реакции биоконверсии с солюбилизирующим агентом, отличающимся от Тритона X-100.

В другом аспекте предложена модель кристаллической структуры (CMS) SHC, основанная на структурных координатах SHC, с аминокислотной последовательностью описанного в настоящем описании SHC или его производного. CMS SHC содержит домен связывающего кармана сквалена/гомофарнезола (squalene/homofarnesol binding pocket domain, SHBD), содержащий связывающий карман сквалена/гомофарнезола (SHBP) и субстрат сквален/гомофарнезол, связанный с SBD (например, см. Фиг. 19 и 20). Эта модель кристаллической структуры SHC (CMS) облегчает тестирование потенциальных производных-кандидатов фермента SHC/HAC in-silico.

Таким образом, в других воплощениях в настоящем изобретении предложен способ скрининга на фермент (например, производное SHC/HAC), способный к связыванию SHBD, включающий использование CMS SHC/HAC. В другом аспекте в настоящем изобретении предложен способ скрининга на фермент (например, эталонного фермента SHC или производного SHC/HAC), способный к связыванию SHBP, где способ включает приведение SHBP в контакт с тестируемым соединением (например, производным SHC) и определение наличия или отсутствия связывания тестируемого соединения с указанным SHBP. В некоторых воплощениях изобретения способ предназначен для скрининга тестируемого соединения (например, модулятора), полезного при модулировании активности производного фермента SHC.

В другом аспекте в настоящем изобретении предложен способ прогнозирования, имитации или моделирования молекулярных характеристик и/или молекулярных взаимодействий эталонного SHC и/или производного SHC/HAC со связывающим доменом сквалена/гомофарнезола (SHBD), включающий использование компьютерной модели, где в указанной компьютерной модели включены, использованы или изображены структурные координаты связывающего домена сквалена/гомофарнезола, как определено выше, для получения изображения указанного лиганд-связывающего домена и, возможно, вывода указанного изображения на дисплей.

Во всем тексте настоящего описания и следующей ниже формулы изобретения, если контекстом не требуется иное, под словом «содержать» и его вариантами, такими как

«содержит» и «содержащий», следует понимать включение указанного целого числа, или стадии, либо группы целых чисел или стадий, но не исключение любого другого целого числа, или стадии, либо группы целых чисел или стадий. Термин «содержащий» также означает «включающий» и также «состоящий из», например, композиция, «содержащая» X, может состоять исключительно из X или может включать что-либо еще дополнительное, например, X+Y. Необходимо отметить, что используемые в настоящем описании и прилагаемой формуле изобретения формы единственного числа включают в себя соответствующие формы множественного числа, если контекстом явным образом не требуется иное. В качестве примера, ссылка на «ген» или «фермент» является ссылкой на «один или более чем один ген» или «один или более чем один фермент».

Должно быть понятно, что изобретение не ограничено конкретной методологией, протоколами и реагентами, описанными в настоящем описании, поскольку они могут варьировать. Должно быть также понятно, что используемая в настоящем описании терминология предназначена только для цели описания конкретных воплощений и не предназначена для ограничения объема настоящего изобретения, которое будет ограничено только прилагаемой формулой изобретения. Если не определено иное, все используемые в настоящем описании технические и научные термины имеют те же значения, которые обычно понятны специалистам в данной области техники. В соответствии с настоящим изобретением могут быть применены традиционные методы молекулярной биологии, микробиологии и рекомбинантных ДНК, которые входят в компетенцию специалистов в данной области техники.

Данное изобретение не ограничено в его применении к деталям конструкции и порядком расположения компонентов, приведенных в последующем описании или проиллюстрированных в графических материалах. Возможны другие воплощения изобретения и его применение на практике или осуществление различными путями. Используемые здесь формулировки и терминология также предназначены для описания настоящего изобретения, и их не следует рассматривать как ограничивающие.

Предпочтительно используемые в настоящем описании термины определены, как описано в публикации «A multilingual glossary of biotechnological terms: (IUPAC Recommendations))), Leuenberger, H.G.W, Nagel, B. and Kolbl, H. eds. (1995), Helvetica Chimica Acta, CH-4010 Basel, Switzerland).

Во всем тексте данного описания цитируется несколько документов. Каждый из цитируемых в настоящем описании документов (включая все патенты, заявки на патенты, научные публикации, спецификации производителя, номера последовательностей по каталогу GenBank и т.д.), приведенных выше или ниже, полностью включен в настоящий документ посредством ссылки.

Описанные в настоящем описании примеры являются иллюстративными для настоящего изобретения и не предназначены для его ограничения. В настоящем описании раскрыты различные воплощения настоящего изобретения. Многочисленные модификации и вариации описанных и проиллюстрированных в настоящем описании методов могут быть выполнены без отклонения от сущности и объема изобретения. Соответственно, следует понимать, что примеры являются только иллюстративными и не имеют ограничительного характера для объема изобретения.

**Таблица 10: приведен номер доступа для AacSHC**

| SHC штамм-источник (название SHC) | Ссылка | Номер доступа |
|---|---|---|
| *Alicyclobacillus acidocaldarius* (WT AacSHC) | JP2009-060799 (Kao) Neumann *et al* Biol Chem (1986) 367; 723-729 | NBRC15652 |

**Таблица 11: приведен номер доступа для ZmoSHC**

| штамм-источник SHC (название SHC) | Ссылка | Штамм и Номер доступа | SEQ ID NO: в соответствии с WO 2010139719 US2012/0135477 |
|---|---|---|---|
| *Zymomonas mobilis* (WT ZmoSHC) | WO2010139719 US2012/0135477 | ATCC31821 PF62207_2 Номер доступа Genpept No AAV90172 | SEQ ID NO: 1 SEQ ID NO: 2 |
| | Reipen *et al* (1995) Microbiology 141: 155-161. | EMBL/Genbank Accession No. для X80766 | |

**Таблица 12: приведены источники других ферментов SHC в WO2010139719**

| штамм-источник SHC (название SHC) | Ссылка | Номер доступа | SEQ ID NO: в соответствии с WO 2010139719 US2012/0135477 |
|---|---|---|---|
| *Bradryhizobium japonicum* (WT BjpSHC) | WO2010139719 US2012/0135477 | PF62207_5 | SEQ ID NO: 5 |
| *Burkholderia ambifaria* | WO2010139719 US2012/0135477 | | SEQ ID NO: 6 |
| *Burkholderia ambifaria* | WO2010139719 US2012/0135477 | | SEQ ID NO: 7 |
| *Bacillus anthracis* | WO2010139719 US2012/0135477 | | SEQ ID NO: 8 |
| *Frankia alni* | WO2010139719 US2012/0135477 | | SEQ ID NO: 9 |
| *Rhodopseudomonas palent* | WO2010139719 US2012/0135477 | | SEQ ID NO: 10 |

**Таблица 13: WT AacSHC, аминокислотная и нуклеотидная SEQ ID NO:**

| Аминокислотная SEQ ID NO: | Штамм | Нуклеотидная SEQ ID NO: |
|---|---|---|
| SEQ ID NO: 1 | WT AacSHC *Alicyclobacillus acidocaldarius* AB007002.1 | 169 |
| SEQ ID NO: 1 | (GI: 218288697) *Alicyclobacillus acidocaldarius* ZP_03492960.1 | 165 |

**Таблица 14: Производное AacSHC, аминокислотная и нуклеотидная SEQ ID NO:**

| Аминокислотная SEQ ID NO: | Мутация (-и) | Производное SHC название | Нуклеотидная 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 |
| 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 |

**Таблица 15: WT ZmoSHC1 и производное ZmoSHC1, аминокислотная и нуклеотидная SEQ ID NO:**

| Аминокислотная SEQ ID NO: | Штамм/мутация (-и) | Производное SHC название | Нуклеотидная SEQ ID NO: |
|---|---|---|---|
| SEQ ID NO: 2 | WT ZmoSHC1 ZmoSHC1 (GI: 56552444) ATCC 31821] | Эталонная последовательность ZmoSHC1 SEQ ID NO: 2 в WO 2010/139719 | 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 |

| | | | |
|---|---|---|---|
| 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 |

**Таблица 16: WT ZmoSHC2 и производное ZmoSHC2, аминокислотная и нуклеотидная SEQ ID NO:**

| Аминокислотная SEQ ID NO: | Штамм/мутация (-и) | Производное SHC название | Нуклеотидная SEQ ID NO: |
|---|---|---|---|
| SEQ ID NO: 3 | WT ZmoSHC2 ZmoSHC2 (GI: 677871) | Эталонная последовательность ZmoSHC1 | 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 |
| 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 |

**Таблица 17: WT BjpSHC1 и производное BjpSHC1, аминокислотная и нуклеотидная SEQ ID NO:**

| Аминокислотная SEQ ID NO: | Штамм/мутация (-и) | Производное SHC название | Нуклеотидная SEQ ID NO: |
|---|---|---|---|
| SEQ ID NO: 4 | WT *B. japonicum* 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 |
| 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 |

SEQ ID NO: 1 (Alicyclobacillus acidocaldarius), AacSHC

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDR
VDRDRMEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQ
KALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGS
WARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRAL
HGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPA
FIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAG
EWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPD
ERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHIPFCDFGEVTDPPSEDV
TAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSA
LKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAW
ALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLGYTMYR
HVFPTLALGRYKQAIERR

SEQ ID NO: 2 (Zymomonas mobilis), ZmoSHC1

MGIDRMNSLSRLLMKKIFGAEKTSYKPASDTIIGTDTLKRPNRRPEPTAKVDKTIFKT
MGNSLNNTLVSACDWLIGQQKPDGHWVGAVESNASMEAEWCLALWFLGLEDHPLR
PRLGNALLEMQREDGSWGVYFGAGNGDINATVEAYAALRSLGYSADNPVLKKAAA
WIAEKGGLKNIRVFTRYWLALIGEWPWEKTPNLPPEIIWFPDNFVFSIYNFAQWARAT
MVPIAILSARRPSRPLRPQDRLDELFPEGRARFDYELPKKEGIDLWSQFFRTTDRGLH
WVQSNLLKRNSLREAAIRHVLEWIIRHQDADGGWGGIQPPWVYGLMALHGEGYQL
YHPVMAKALSALDDPGWRHDRGESSWIQATNSPVWDTMLALMALKDAKAEDRFTP
EMDKAADWLLARQVKVKGDWSIKLPDVEPGGWAFEYANDRYPDTDDTAVALIALS
SYRDKEEWQKKGVEDAITRGVNWLIAMQSECGGWGAFDKDNNRSILSKIPFCDFGES
IDPPSVDVTAHVLEAFGTLGLSRDMPVIQKAIDYVRSEQEAEGAWFGRWGVNYIYGT
GAVLPALAAIGEDMTQPYITKACDWLVAHQQEDGGWGESCSSYMEIDSIGKGPTTPS
QTAWALMGLIAANRPEDYEAIAKGCHYLIDRQEQDGSWKEEEFTGTGFPGYGVGQTI
KLDDPALSKRLLQGAELSRAFMLRYDFYRQFFPIMALSRAERLIDLNN

SEQ ID NO: 3 (Zymomonas mobilis), ZmoSHC2

MTVSTSSAFHHSPLSDDVEPIIQKATRALLEKQQQDGHWVFELEADATIPAEYILLKH
YLGEPEDLEIEAKIGRYLRRIQGEHGGWSLFYGGDLDLSATVKAYFALKMIGDSPDAP
HMLRARNEILARGGAMRANVFTRIQLALFGAMSWEHVPQMPVELMLMPEWFPVHI
NKMAYWARTVLVPLLVLQALKPVARNRRGILVDELFVPDVLPTLQESGDPIWRRFFS
ALDKVLHKVEPYWPKNMRAKAIHSCVHFVTERLNGEDGLGAIYPAIANSVMMYDA
LGYPENHPERAIARRAVEKLMVLDGTEDQGDKEVYCQPCLSPIWDTALVAHAMLEV
GGDEAEKSAISALSWLKPQQILDVKGDWAWRRPDLRPGGWAFQYRNDYYPDVDDT
AVVTMAMDRAAKLSDLHDDFEESKARAMEWTIGMQSDNGGWGAFDANNSYTYLN
NIPFADHGALLDPPTVDVSARCVSMMAQAGISITDPKMKAAVDYLLKEQEEDGSWF
GRWGVNYIYGTWSALCALNVAALPHDHLAVQKAVAWLKTIQNEDGGWGENCDSY
AIDYSGYEPMDSTASQTAWALLGLMAVGEANSEAVTKGINWLAQNQDEEGLWKE
DYYSGGGFPRVFYLRYHGYSKYFPLWALARYRNLKKANQPIVHYGM

SEQ ID NO: 4 (Bradyrhizobium japonicum), BjpSHC

MTVTSSASARATRDPGNYQTALQSTVRAAADWLIANQKPDGHWVGRAESNACMEA
QWCLALWFMGLEDHPLRKRLGQSLLDSQRPDGAWQVYFGAPNGDINATVEAYAAL
RSLGFRDDEPAVRRAREWIEAKGGLRNIRVFTRYWLALIGEWPWEKTPNIPPEVIWFP
LWFPFSIYNFAQWARATLMPIAVLSARRPSRPLPPENRLDALFPHGRKAFDYELPVKA
GAGGWDRFFRGADKVLHKLQNLGNRLNLGLFRPAATSRVLEWMIRHQDFDGAWG
GIQPPWIYGLMALYAEGYPLNHPVLAKGLDALNDPGWRVDVGDATYIQATNSPVW
DTILTLLAFDDAGVLGDYPEAVDKAVDWVLQRQVRVPGDWSMKLPHVKPGGWAFE
YANNYYPDTDDTAVALIALAPLRIIDPKWKAKGIDEAIQLGVDWLIGMQSQGGGWG
AFDKDNNQKILTKIPFCDYGEALDPPSVDVTAHIIEAFGKLGISRNHPSMVQALDYIRR
EQEPSGPWFGRWGVNYVYGTGAVLPALAAIGEDMTQPYIGRACDWLVAHQQADGG
WGESCASYMDVSAVGRGTTTASQTAWALMALLAANRPQDKDAIERGCMWLVERQ
SAGTWDEPEFTGTGFPGYGVGQTIKLNDPALSQRLMQGPELSRAFMLRYGMYRHYF
PLMALGRALRPQSHS

SEQ ID NO: 149 (Burkholderia ambifaria)

MNDLTEMATLSAGTVPAGLDAAVASATDALLAAQNADGHWVYELEADSTIPAEYV
LLVHYLGETPNLELEQKIGRYLRRVQQADGGWPLFTDGAPNISASVKAYFALKVIGD
DENAEHMQRARRAIQAMGGAEMSNVFTRIQLALYGAIPWRAVPMMPVEIMLLPQW
FPFHLSKVSYWARTVIVPLLVLNAKRPIAKNPRGVRIDELFVDPPVNAGLLPRQGHQS

PGWFAFFRVVDHALRAADGLFPNYTRERAIRQAVSFVDERLNGEDGLGAIYPAMAN
AVMMYDVLGYAEDHPNRAIARKSIEKLLVVQEDEAYCQPCLSPVWDTSLAAHALLE
TGDARAEEAVIRGLEWLRPLQILDVRGDWISRRPHVRPGGWAFQYANPHYPDVDDT
AVVAVAMDRVQKLKHNDAFRDSIARAREWVVGMQSSDGGWGAFEPENTQYYLNN
IPFSDHGALLDPPTADVSGRCLSMLAQLGETPLNSEPARRALDYMLKEQEPDGSWYG
RWGMNYVYGTWTALCALNAAGLTPDDPRVKRGAQWLLSIQNKDGGWGEDGDSYK
LNYRGFEQAPSTASQTAWALLGLMAAGEVNNPAVARGVEYLIAEQKEHGLWDETR
FTATGFPRVFYLRYHGYRKFFPLWALARYRNLKRNNATRVTFGL

SEQ ID NO: 151 (Burkholderia ambifaria)

MIRRMNKSGPSPWSALDAAIARGRDALMRLQQPDGSWCFELESDATITAEYILMMHF
MDKIDDARQEKMARYLRAIQRLDTHGGWDLYVDGDPDVSCSVKAYFALKAAGDSE
HAPHMVRARDAILELGGAARSNVFTRILLATFGQVPWRATPFMPIEFVLFPKWVPISM
YKVAYWARTTMVPLLVLCSLKARARNPRNIAIPELFVTPPDQERQYFPPARGMRRAF
LALDRVVRHVEPLLPKRLRQRAIRHAQAWCAERMNGEDGLGGIFPPIVYSYQMMDV
LGYPDDHPLRRDCENALEKLLVTRPDGSMYCQPCLSPVWDTAWSTMALEQARGVA
VPEAGAPASALDELDARIARAYDWLAERQVNDLRGDWIENAPADTQPGGWAFQYA
NPYYPDIDDSAVVTAMLDRRGRTHRNADGSHPYAARVARALDWMRGLQSRNGGFA
AFDADCDRLYLNAIPFADHGALLDPPTEDVSGRVLLCFGVTKRADDRASLARAIDYV
KRTQQPDGSWWGRWGTNYLYGTWSVLAGLALAGEDPSQPYIARALAWLRARQHA
DGGWGETNDSYIDPALAGTNAGESTSNCTAWALLAQMAFGDGESESVRRGIAYLQS
VQQDDGFWWHRSHNAPGFPRIFYLKYHGYTAYFPLWALARYRRLAGGVSAAGAHA
VPASTGADAAIA

SEQ ID NO: 153 (Bacillus anthracis)

MLLYEKAHEEIVRRATALQTMQWQDGTWRFCFEGAPLTDCHMIFLLKLLGRDKEIEP
FVERVASLQTNEGTWKLHEDEVGGNLSATIQSYAALLASKKYTKEDANMKRAENFI
QERGGVARAHFMTKFLLAIHGEYEYPSLFHLPTPIMFLQNDSPFSIFELSSSARIHLIPM
MLCLNKRFRVGKKLLPNLNHIAGGGGEWFREDRSPVFQTLLSDVKQIISYPLSLHHK
GYEEIERFMKERIDENGTLYSYATASFYMIYALLALGHSLQSSMIQKAIAGITSYIWK
MERGNHLQNSPSTVWDTALLSYALQEAQVSKDNKMIQNATAYLLKKQHTKKADWS
VHAPALTPGGWGFSDVNTTIPDIDDTTAVLRALARSRGNKNIDNAWKKGGNWIKGL
QNNDGGWGAFEKGVTSKLLAKLPIENASDMITDPSTPDITGRVLEFFGTYAQNELPEK
QIQRAINWLMNVQEENGSWYGKWGICYLYGTWAVMTGLRSLGIPSSNPSLTRAASW
LEHIQHEDGGWGESCHSSVEKRFVTLPFSTPSQTAWALDALISYYDTETPAIRKGVSY
LLSNPYVNERYPTGTGLPGAFYIRYHSYAHIYPLLTLAHYIKKYRK

SEQ ID NO: 155 (Frankia alni)

MPAGVGVLVWLDQRLRAMGRPDLVTTTGGAEIPFVLVAATASTVGVALALRRPRHP
VGWLFLALGGVLLLSGGTQGYAAYGAVARPGRLPAADLVAIYADAGFIPWLVLVAL
ILHLTPTGRPLSARWGRIALATAVAGGLWLLVGLVTTETMQPPFQSVTNPLLIGGPLG
PLLVARRVLGLATGAGVVLAAVSLIVRFRRSVDVERRQLLWVAVAAVPLPVLMAAS
FAASYAGNNTAAGLAAATLIGLLAIGAGLAIGQYHLYDVEEILSRAVTYLLVSGLLA
ASYATVVIVVGQSLAGRTGRSQISAVLATLAAVAVTAPAYRKIQEGVDRRFSRRRFE
TLQVIRRYLRDPDPDVAVEEVLRRALGDPTLAVAYLVDDRRQWVSADGQPANPGNS
FMAAVEVYRRGRPIARVTFDRGRAQPGLVRAAAATAATAELDNAGLRAAVALQLVE
VRQSRTRIAAAQIFAERRTIERNLIIDGAQQRLLALALQLRAVQLGGDEASLRQAISTGI
DQLQAAVVELRELANGLIIPAVLADGGLAAALDDVAARTPVPIKISAPDRRYPPDLEA
AAWFIACEAMANAVKHAHPTTIAVDVSAPDGQLIVEVRDDGIGGAQPSGPGLRGIAD
RAEAFGGSLTVHTDPGTGTTIRAILHRRSPLSSGRRSVMIEGCVDVVAVRRFRCRSSR
GSGSRRRRSSWRCGGICGSRCRTGMSRSCSRNAASKLIT

SEQ ID NO: 157 (Rhodopseudomonas patent)

MDSILAPRADAPRNIDGALRESVQQAADWLVANQKPDGHWVGRAETNATMEAQW
CLALWFLGLEDHPLRVRLGRALLDTQRPDGAWHVFYGAPNGDINATVEAYAALRSL
GHRDDEEPLRKARDWILSKGGLANIRVFTRYWLALIGEWPWEKTPNILPEVIWLPTW
FPFSIYNFAQWARATLMPIAVLSAHRPSRPLAPQDRLDALFPQGRDSFNYDLPARLGA
GVWDVIFRKIDTILHRLQDWGARRGPHGIMRRGAIDHVLQWIIRHQDYDGSWGGIQP
PWIYGLMALHTEGYAMTHPVMAKALDALNEPGWRIDIGDATFIQATNSPVWDTMLS
LLAFDDAGLGERYPEQVERAVRWVLKRQVLVPGDWSVKLPDVKPGGWAFEYANNF
YPDTDDTSVALMALAPFRHDPKWQAEGIEDAIQRGIDWLVAMQCKEGGWGAFDKD
NDKKILAKIPFCDFGEALDPPSADVTAHIIEAFAKVGLDRNHPSIVRALDYLKREQEPE
GPWFGRWGVNYVYGTGAVLPALAAIGEDMRQPYIARACDWLIARQQANGGWGESC
VSYMDAKQAGEGTATASQTAWALMALIAADRPQDRDAIERGCLYLTETQRDGTWQ
EVHYTGTGFPGYGVGQTIKLNDPLLSKRLMQGPELSRSFMLRYDLYRHYFPMMAIGR
VLRQRGDRSGH

SEQ ID NO: 159 (Streptomyces coelicolor)

MTATTDGSTGASLRPLAASASDTDITIPAAAAGVPEAAARATRRATDFLLAKQDAEG
WWKGDLETNVTMDAEDILLRQFLGIQDEETTRAAALFIRGEQREDGTWATFYGGPG
FLSTTIEAYVALRLAGDSPEAPHMARAAEWIRSRGGIASARVFTRIWLALFGWWKW

DDLPELPPELIYFPTWVPLNIYDFGCWARQTIVPLTIVSAKRPVRPAPFPLDELIITDPA
RPNPPRPLAPVASWDGAFQRIDKALIIAYRKVAPRRLRRAAMNSAARWIIERQENDG
CWGGIQPPAVYSVIALYLLGYDLEHPVMRAGLESLDRFAVWREDGARMIEACQSPV
WDTCLATIALADAGVPEDHPQLVKASDWMLGEQIVRPGDWSVKRPGPPGGWAFEF
IINDNYPDIDDTAEVVLALRRVRIIIIDPERVEKAIGRGVRWNLGMQSKNGAWGAFDV
DNTSAFPNRLPFCDFGEVIDPPSADVTAHVVEMLAVEGLAHDPRTRRGIQWLLDAQE
TDGSWFGRWGVNYVYGTGSVIPALTAAGLPTSHPAIRRAVRWLESVQNEDGGWGE
DLRSYRYVREWSGRGASTASQTGWALMALLAAGERDSKAVERGVAWLAATQRED
GSWDEPYFTGTGFPWDFSINYNLYRQVFPLTALGRYVHGEPFAKKPRAADAPAEAAP
AEVKGS

SEQ ID NO: 169

ATGGCTGAGCAGTTGGTGGAAGCGCCGGCCTACGCGCGGACGCTGGATCGCGCG
GTGGAGTATCTCCTCTCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTC
TGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTTGTGCCACATTCTCGATCG
CGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCG
CGAGGACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGAC
CATCGAGGCGTACGTCGCGCTCAAGTATATCGGCATGTCGCGCGACGAGGAGCC
GATGCAGAAGGCGCTCCGGTTCATTCAGAGCCAGGGCGGGATCGAGTCGTCGCG
CGTGTTCACGCGGATGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGT
GCCCATGGTCCCGCCGGAGATCATGTTCCTCGGCAAGCGCATGCCGCTCAACATC
TACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTCTCGATTGTGATGA
GCCGCCAGCCGGTGTTCCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGA
GACCGACGTGCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTT
CGACGCGCTCGACCGGGCGCTGCACGGGTATCAGAAGCTGTCGGTGCACCCGTTC
CGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCA
AGATTCTCGACATGACGCAGCATCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGA
GCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGGCTTCCATCTCG
CCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGG
CCGATCACGACCGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCA
CGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAACCTCAAGCCGGGCGGGTTCG
CGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGT
GTGGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCGGGACGCCAT

GACGAAGGGATTCCGCTGGATTGTCGGCATGCAGAGCTCGAACGGCGGTTGGGG
CGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACATCCCGTTCTGCGAC
TTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGT
GTTTCGGCAGCTTCGGGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGA
ATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGGTTCGGTCGTTGGGGCGT
CAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATC
GACACGCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAG
AACCCGGACGGCGGCTGGGGCGAGGACTGCCGCTCGTACGAGGATCCGGCGTAC
GCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGGCCTGGGCGCTGATGGCGCTC
ATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTC
GTGGAGACGCAGCGCCCGGACGGCGGCTGGGATGAGCCGTACTACACCGGCACG
GGCTTCCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCCACGTGTTTCCGAC
GCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

Вариант 101A10 (SEQ ID NO: 30)

ATGGCTGAGCAGTTGGTGGAAGCACCGGCCTACGCGCGGACGCTGGATCGCGCG
GTGGAGTATCTCCTCTCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTC
TGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTTGTGCCACATTCTCGATCG
CGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCG
CGAGGACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGAC
CATCGAGGCGTACGTCGCGCTCAAGTATATCGGCATGTCGCGCGACGAGGAGCC
GATGCAGAAGGCGCTCCGGTTCATTCAGAGCCAGGGCGGGATCGAGTCGTCGCG
CGTGTTCACGCGGATGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGT
GCCCATGGTCCCGCCGGAGATCATGTTCCTCGGCAAGCGCATGCCGCTCAACATC
TACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTCTCGATTGTGATGA
GCCGCCAGCCGGTGTTCCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGA
GACCGACGTGCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTT
CGACGCGCTCGACCGGGCGCTGCACGGGTATCAGAAGCTGTCGGTGCACCCGTTC
CGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCA
AGATTCTCGACATGACGCAGCATCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGA
GCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGGCTTCCATCTCG
CCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGG
CCGATCACGACCGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCA

CGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAACCTCAAGCCGGGCGGGTTCG
CGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGT
GTGGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCGGGACGCCAT
GACGAAGGGATTCCGCTGGATTGTCGGCATGCAGAGCTCGAACGGCGGTTGGGG
CGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACATCCCGTTCTGCGAC
TTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGT
GTTTCGGCAGCTTCGGGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGA
ATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGGTTCGGTCGTTGGGGCGT
CAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATC
GACACGCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAG
AACCCGGACGGCGGCTGGGGCGAGGACTGCCGCTCGTACGAGGATCCGGCGTAC
GCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGGCCTGGGCGCTGATGGCGCTC
ATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCATTACCTC
GTGGAGACGCAGCGCCCGGACGGCGGCTGGGATGAGCCGTACTACACCGGCACG
GGCTACCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCCACGTGTTTCCGA
CGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

Вариант 101A10 (SEQ ID NO: 29)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDR
VDRDRMEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQ
KALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGS
WARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRAL
HGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPA
FIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAG
EWILDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPD
ERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHIPFCDFGEVTDPPSEDV
TAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSA
LKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAW
ALMALIAGGRAESEAARRGVHYLVETQRPDGGWDEPYYTGTGYPGDFYLGYTMYR
HVFPTLALGRYKQAIERR

Вариант 111C8 (SEQ ID NO: 28)

ATGGCTGAGCAGTTGGTGGAAGCGCCGGCCTACGCGCGGACGCTGGATCGCGCG
GTGGAGTATCTCCTCTCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTC
TGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTTGTGCCACATTCTCGATCG

CGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCG
CGAGGACGGCGCGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGAC
CGTCGAGGCGTACGTCGCGCTCAAGTATATCGGCATGTCGCGCGACGAGGAGCC
GATGCAGAAGGCGCTCCGGTTCATTCAGAGCCAGGGCGGGATCGAGTCGTCGCG
CGTGCTCACGCGGATGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGT
GCCCATGGTCCCGCCGGAGATCATGTTCCTCGGCAAGCGCATGCCGCTCAACATC
TACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTCTCGATTGTGATGA
GCCGCCAGCCGGTGTTCCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGA
GACCGACGTGCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTT
CGACGCGCTCGACCGGGCGCTGCACGGGTATCAGAAGCTGTCGGTGCACCCGTTC
CGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCA
AGATTCTCGACATGACGCAGCATCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGA
GCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGGCTTCCATCTCG
CCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGG
CCGATCACGACCGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCA
CGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAACCTCAAGCCGGGCGGGTTCG
CGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGT
GTGGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCGGGACGCCAT
GACGAAGGGATTCCGCTGGATTGTCGGCATGCAGAGCTCGAACGGCGGTTGGGG
CGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACATCCCGTTCTGCGAC
TTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGT
GTTTCGGCAGCTTCGGGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGA
ATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGGTTCGGTCGTTGGGGCGT
CAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATC
GACACGCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAG
AACCCGGACGGCGGCTGGGGCGAGGACTGCCGCTCGTACGAGGATCCGGCGTAC
GCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGGCCTGGGCGCTGATGGCGCTC
ATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTC
GTGGAGACGCAGCGCCCGGACGGCGGCTGGGATGAGCCGTACTACACCGGCACG
GGCTTCCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCCACGTGTTTCCGAC
GCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

Вариант 111C8 (SEQ ID NO: 27)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDR
VDRDRMEKIRRYLLHEQREDGAWALYPGGPPDLDTTVEAYVALKYIGMSRDEEPMQ
KALRFIQSQGGIESSRVLTRMWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGS
WARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRAL
HGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPA
FIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAG
EWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPD
ERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHIPFCDFGEVTDPPSEDV
TAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSA
LKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAW
ALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLGYTMYR
HVFPTLALGRYKQAIERR

Вариант SHC215G2 (SEQ ID NO: 22)

ATGGCTGAGCAGTTGGTGGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCG
GTGGAGTATCTCCTCTCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTC
TGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTTGTGCCACATTCTCGATCG
CGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCG
CGAGGACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGAC
CATCGAGGCGTACGTCGCGCTCAAGTATATCGGCATGTCGCGCGACGAGGAGCC
GATGCAGAAGGCGCTCCGGTTCATTCAGAGCCAGGGCGGGATCGAGTCGTCGCG
CGTGTTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGT
GCCCATGGTCCCGCCGGAGATCATGTTCCTCGGCAAGCGCATGCCGCTCAACATC
TACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTCTCGATTGTGATGA
GCCGCCAGCCGGTGTTCCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGA
GACCGACGTGCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTT
CGACGCGCTCGACCGGGTGCTGCACGGGTATCAGAAGCTGTCGGTGCACCCGTTC
CGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCA
AGATTCTCGACATGACGCAGCATCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGA
GCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGGCTTCCATCTCG
CCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGGCTGCGGGGCTTCCGG
CCGATCACGACCGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCA
CGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAACCTCAAGCCGGGCGGGTTCG

CGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGT
GTGGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCGGGACGCCAT
GACGAAGGGATTCCGCTGGATTGTCGGCATGCAGAGCTCGAACGGCGGTTGGGG
CGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACACCCCGTTCTGCGAC
TTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGT
GTTTCGGCAGCTTCGGGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGA
ATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGGTTCGGTCGTTGGGGCGT
CAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATC
GACACGCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAG
AACCCGGACGGCGGCTGGGGCGAGGACTGCCGCTCGTACGAGGATCCGGCGTAC
GCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGGCCTGGGCGCTGATGGCGCTC
ATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTC
GTGGAGACGCAGCGCCCGGACGGCGGCTGGGATGAGCCGTACTACACCGGCACG
GGCTTCCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCCACGTGTTTCCGAC
GCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

Вариант SHC215G2 (SEQ ID NO: 21)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDR
VDRDRMEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQ
KALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGS
WARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRVL
HGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPA
FIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAG
EWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPD
ERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHTPFCDFGEVTDPPSEDV
TAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSA
LKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAW
ALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLGYTMYR
HVFPTLALGRYKQAIERR

Вариант SHC3 (SEQ ID NO: 26)

ATGGCTGAGCAGTTGGTGGAAGCGCCGGCCTACGCGCGGACGCTGGATCGCGCG
GTGGAGTATCTCCTCTCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTC
TGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTTGTGCCACATTCTCGATCG
CGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCG

CGAGGACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGAC
CATCGAGGCGTACGTCGCGCTCAAGTATATCGGCATGTCGCGCGACGAGGAGCC
GATGCAGAAGGCGCTCCGGTTCATTCAGAGCCAGGGCGGGATCGAGTCGTCGCG
CGTGTTCACGCGGATGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGT
GCCCATGGTCCCGCCGGAGATCATGTTCCTCGGCAAGCGCATGCCGCTCAACATC
TACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTCTCGATTGTGATGA
GCCGCCAGCCGGTGTTCCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGA
GACCGACGTGCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTT
CGACGCGCTCGACCGGGCGCTGCACGGGTATCAGAAGCTGTCGGTGCACCCGTTC
CGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCA
AGATTCTCGACATGACGCAGCATCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGA
GCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGGCTTCCATCTCG
CCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGG
CCGATCACGACCGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCA
CGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAACCTCAAGCCGGGCGGGTTCG
CGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGT
GTGGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCGGGACGCCAT
GACGAAGGGATTCCGCTGGATTGTCGGCATGCAGAGCTCGAACGGCGGTTGGGG
CGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACATCCCGTTCTGCGAC
TTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGT
GTTTCGGCAGCTTCGGGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGA
ATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGGTTCGGTCGTTGGGGCGT
CAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATC
GACACGCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAG
AACCCGGACGGCGGCTGGGGCGAGGACTGCCGCTCGTACGAGGATCCGGCGTAC
GCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGGCCTGGGCGCTGATGGCGCTC
ATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTC
GTGGAGACGCAGCGCCCGGACGGCGGCTGGGATGAGCCGTACTACACCGGCACG
GGCTACCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCCACGTGTTTCCGA
CGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

Вариант SHC3 (SEQ ID NO: 25)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDR

VDRDRMEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQ
KALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGS
WARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDAIDRAL
HGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPA
FIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAG
EWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPD
ERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHIPFCDFGEVTDPPSEDV
TAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSA
LKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAW
ALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGYPGDFYLGYTMYR
HVFPTLALGRYKQAIERR

Вариант SHC10 (SEQ ID NO: 32)

ATGGCTGAGCAGTTGGTGGAAGCGCCGGCCTACGCGCGGACGCTGGATCGCGCG
GTGGAGTATCTCCTCTCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTC
TGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTTGTGCCACATTCTCGATCG
CGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCG
CGAGGACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGAC
CATCGAGGCGTACGTCGCGCTCAAGTATATCGGCATGTCGCGCGACGAGGAGCC
GATGCAGAAGGCGCTCCGGTTCATTCAGAGCCAGGGCGGGATCGAGTCGTCGCG
CGTGCTCACGCGGATGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGT
GCCCATGGTCCCGCCGGAGATCATGTTCCTCGGCAAGCGCATGCCGCTCAACATC
TACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTCTCGATTGTGATGA
GCCGCCAGCCGGTGTTCCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGA
GACCGACGTGCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTT
CGACGCGCTCGACCGGGCGCTGCACGGGTATCAGAAGCTGTCGGTGCACCCGTTC
CGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCA
AGATTCTCGACATGACGCAGCATCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGA
GCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGGCTTCCATCTCG
CCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGG
CCGATCACGACCGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCA
CGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAACCTCAAGCCGGGCGGGTTCG
CGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGT

GTGGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCGGGACGCCAT
GACGAAGGGATTCCGCTGGATTGTCGGCATGCAGAGCTCGAACGGCGGTTGGGG
CGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACATCCCGTTCTGCGAC
TTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGT
GTTTCGGCAGCTTCGGGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGA
ATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGGTTCGGTCGTTGGGGCGT
CAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATC
GACACGCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAG
AACCCGGACGGCGGCTGGGGCGAGGACTGCCGCTCGTACGAGGATCCGGCGTAC
GCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGGCCTGGGCGCTGATGGCGCTC
ATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTC
GTGGAGACGCAGCGCCCGGACGGCGGCTGGGATGAGCCGTACTACACCGGCACG
GGCTTCCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCCACGTGTTTCCGAC
GCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

Вариант SHC10 (SEQ ID NO: 31)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDR
VDRDRMEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQ
KALRFIQSQGGIESSRVLTRMWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGS
WARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRAL
HGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPA
FIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAG
EWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPD
ERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHIPFCDFGEVTDPPSEDV
TAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSA
LKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAW
ALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLGYTMYR
HVFPTLALGRYKQAIERR

Вариант SHC26 (SEQ ID NO: 24)

ATGGCTGAGCAGTTGGTGGAAGCTCCGGCCTACGCGCGGACGCTGGATCGCGCG
GTGGAGTATCTCCTCTCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTC
TGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTTGTGCCACATTCTCGATCG
CGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCG
CGAGGACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGAC

CATCGAGGCGTACGTCGCGCTCAAGTATATCGGCATGTCGCGCGACGAGGAGCC
GATGCAGAAGGCGCTCCGGTTCATTCAGAGCCAGGGCGGGATCGAGTCGTCGCG
CGTGTTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGT
GCCCATGGTCCCGCCGGAGATCATGTTCCTCGGCAAGCGCATGCCGCTCAACATC
TACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTCTCGATTGTGATGA
GCCGCCAGCCGGTGTTCCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGA
GACCGACGTGCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTT
CGACGCGCTCGACCGGGCGCTGCACGGGTATCAGAAGCTGTCGGTGCACCCGTTC
CGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCA
AGATTCTCGACATGACGCAGCATCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGA
GCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGGCTTCCATCTCG
CCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGG
CCGATCACGACCGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCA
CGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAACCTCAAGCCGGGCGGGTTCG
CGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGT
GTGGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCGGGACGCCAT
GACGAAGGGATTCCGCTGGATTGTCGGCATGCAGAGCTCGAACGGCGGTTGGGG
CGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACACCCCGTTCTGCGAC
TTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGT
GTTTCGGCAGCTTCGGGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGA
ATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGGTTCGGTCGTTGGGGCGT
CAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATC
GACACGCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAG
AACCCGGACGGCGGCTGGGGCGAGGACTGCCGCTCGTACGAGGATCCGGCGTAC
GCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGGCCTGGGCGCTGATGGCGCTC
ATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTC
GTGGAGACGCAGCGCCCGGACGGCGGCTGGGATGAGCCGTACTACACCGGCACG
GGCTTCCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCCACGTGTTTCCGAC
GCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

Вариант SHC26 (SEQ ID NO: 23)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDR
VDRDRMEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQ

KALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGS
WARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRAL
HGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPA
FIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAG
EWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPD
ERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHTPFCDFGEVTDPPSEDV
TAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSA
LKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAW
ALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLGYTMYR
HVFPTLALGRYKQAIERR

Вариант SHC30 (SEQ ID NO: 34)

ATGGCTGAGCAGTTGGTGGAAGCGCCGGCCTACGCGCGGACGCTGGATCGCGCG
GTGGAGTATCTCCTCTCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTC
TGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTTGTGCCACATTCTCGATCG
CGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCG
CGAGGACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGAC
CATCGAGGCGTACGTCGCGCTCAAGTATATCGGCATGTCGCGCGACGAGGAGCC
GATGCAGAAGGCGCTCCGGTTCATTCAGAGCCAGGGCGGGATCGAGTCGTCGCG
CGTGCTCACGCGGATGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGT
GCCCATGGTCCCGCCGGAGATCATGTTCCTCGGCAAGCGCATGCCGCTCAACATC
TACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTCTCGATTGTGATGA
GCCGCCAGCCGGTGTTCCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGA
GACCGACGTGCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTT
CGACGCGCTCGACCGGGCGCTGCACGGGTATCAGAAGCTGTCGGTGCACCCGTTC
CGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCA
AGATTCTCGACATGACGCAGCATCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGA
GCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGGCTTCCATCTCG
CCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGG
CCGATCACGACCGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCA
CGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAACCTCAAGCCGGGCGGGTTCG
CGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGT
GTGGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCGGGACGCCAT

GACGAAGGGATTCCGCTGGATTGTCGGCATGCAGAGCTCGAACGGCGGTTGGGG
CGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACATCCCGTTCTGCGAC
TTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGT
GTTTCGGCAGCTTCGGGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGA
ATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGGTTCGGTCGTTGGGGCGT
CAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATC
GACACGCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAG
AACCCGGACGGCGGCTGGGGCGAGGACTGCCGCTCGTACGAGGATCCGGCGTAC
GCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGGCCTGGGCGCTGATGGCGCTC
ATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTC
GTGGAGACGCAGCGCCCGGACGGCGGCTGGGATGAGCCGTACTACACCGGCACG
GGCTACCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCCACGTGTTTCCGA
CGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

Вариант SHC30 (SEQ ID NO: 33)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDR
VDRDRMEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQ
KALRFIQSQGGIESSRVLTRMWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGS
WARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRAL
HGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPA
FIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAG
EWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPD
ERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHIPFCDFGEVTDPPSEDV
TAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSA
LKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAW
ALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGYPGDFYLGYTMYR
HVFPTLALGRYKQAIERR

Вариант SHC31 (SEQ ID NO: 36)

ATGGCTGAGCAGTTGGTGGAAGCGCCGGCCTACGCGCGGACGCTGGATCGCGCG
GTGGAGTATCTCCTCTCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTC
TGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTTGTGCCACATTCTCGATCG
CGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCG
CGAGGACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGAC
CATCGAGGCGTACGTCGCGCTCAAGTATATCGGCATGTCGCGCGACGAGGAGCC

GATGCAGAAGGCGCTCCGGTTCATTCAGAGCCAGGGCGGGATCGAGTCGTCGCG
CGTGCTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGT
GCCCATGGTCCCGCCGGAGATCATGTTCCTCGGCAAGCGCATGCCGCTCAACATC
TACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTCTCGATTGTGATGA
GCCGCCAGCCGGTGTTCCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGA
GACCGACGTGCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTT
CGACGCGCTCGACCGGGCGCTGCACGGGTATCAGAAGCTGTCGGTGCACCCGTTC
CGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCA
AGATTCTCGACATGACGCAGCATCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGA
GCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGGCTTCCATCTCG
CCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGG
CCGATCACGACCGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCA
CGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAACCTCAAGCCGGGCGGGTTCG
CGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGT
GTGGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCGGGACGCCAT
GACGAAGGGATTCCGCTGGATTGTCGGCATGCAGAGCTCGAACGGCGGTTGGGG
CGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACACCCCGTTCTGCGAC
TTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGT
GTTTCGGCAGCTTCGGGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGA
ATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGGTTCGGTCGTTGGGGCGT
CAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATC
GACACGCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAG
AACCCGGACGGCGGCTGGGGCGAGGACTGCCGCTCGTACGAGGATCCGGCGTAC
GCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGGCCTGGGCGCTGATGGCGCTC
ATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTC
GTGGAGACGCAGCGCCCGGACGGCGGCTGGGATGAGCCGTACTACACCGGCACG
GGCTTCCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCCACGTGTTTCCGAC
GCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

Вариант SHC31 (SEQ ID NO: 35)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDR
VDRDRMEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQ
KALRFIQSQGGIESSRVLTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGS

WARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRAL
IIGYQKLSVIIPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPA
FIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAG
EWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPD
ERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHIPFCDFGEVTDPPSEDV
TAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSA
LKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAW
ALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDFYLGYTMYR
HVFPTLALGRYKQAIERR

Вариант SHC32 (SEQ ID NO: 38)

ATGGCTGAGCAGTTGGTGGAAGCGCCGGCCTACGCGCGGACGCTGGATCGCGCG
GTGGAGTATCTCCCTCTCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTC
TGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTTGTGCCACATTCTCGATCG
CGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCG
CGAGGACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGAC
CATCGAGGCGTACGTCGCGCTCAAGTATATCGGCATGTCGCGCGACGAGGAGCC
GATGCAGAAGGCGCTCCGGTTCATTCAGAGCCAGGGCGGGATCGAGTCGTCGCG
CGTGTTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGT
GCCCATGGTCCCGCCGGAGATCATGTTCCTCGGCAAGCGCATGCCGCTCAACATC
TACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTCTCGATTGTGATGA
GCCGCCAGCCGGTGTTCCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGA
GACCGACGTGCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTT
CGACGCGCTCGACCGGGCGCTGCACGGGTATCAGAAGCTGTCGGTGCACCCGTTC
CGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCA
AGATTCTCGACATGACGCAGCATCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGA
GCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGGCTTCCATCTCG
CCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGG
CCGATCACGACCGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCA
CGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAACCTCAAGCCGGGCGGGTTCG
CGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGT
GTGGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCGGGACGCCAT
GACGAAGGGATTCCGCTGGATTGTCGGCATGCAGAGCTCGAACGGCGGTTGGGG

CGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACACCCCGTTCTGCGAC
TTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGT
GTTTCGGCAGCTTCGGGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGA
ATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGGTTCGGTCGTTGGGGCGT
CAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATC
GACACGCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAG
AACCCGGACGGCGGCTGGGGCGAGGACTGCCGCTCGTACGAGGATCCGGCGTAC
GCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGGCCTGGGCGCTGATGGCGCTC
ATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTC
GTGGAGACGCAGCGCCCGGACGGCGGCTGGGATGAGCCGTACTACACCGGCACG
GGCTACCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCCACGTGTTTCCGA
CGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

Вариант SHC32 (SEQ ID NO: 37)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDR
VDRDRMEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQ
KALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGS
WARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRAL
HGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPA
FIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHIDRLVKAG
EWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPD
ERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHTPFCDFGEVTDPPSEDV
TAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSA
LKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAW
ALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGYPGDFYLGYTMYR
HVFPTLALGRYKQAIERR

Вариант SHC33 (SEQ ID NO: 40)

ATGGCTGAGCAGTTGGTGGAAGCGCCGGCCTACGCGCGGACGCTGGATCGCGCG
GTGGAGTATCTCCTCTCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTC
TGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTTGTGCCACATTCTCGATCG
CGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCG
CGAGGACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGAC
CATCGAGGCGTACGTCGCGCTCAAGTATATCGGCATGTCGCGCGACGAGGAGCC
GATGCAGAAGGCGCTCCGGTTCATTCAGAGCCAGGGCGGGATCGAGTCGTCGCG

CGTGCTCACGCGGAGGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGT
GCCCATGGTCCCGCCGGAGATCATGTTCCTCGGCAAGCGCATGCCGCTCAACATC
TACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTCTCGATTGTGATGA
GCCGCCAGCCGGTGTTCCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGA
GACCGACGTGCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTT
CGACGCGCTCGACCGGGCGCTGCACGGGTATCAGAAGCTGTCGGTGCACCCGTTC
CGCCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCA
AGATTCTCGACATGACGCAGCATCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGA
GCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGGCTTCCATCTCG
CCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGG
CCGATCACGACCGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCA
CGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAACCTCAAGCCGGGCGGGTTCG
CGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGT
GTGGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCGGGACGCCAT
GACGAAGGGATTCCGCTGGATTGTCGGCATGCAGAGCTCGAACGGCGGTTGGGG
CGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACACCCCGTTCTGCGAC
TTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGT
GTTTCGGCAGCTTCGGGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGA
ATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGGTTCGGTCGTTGGGGCGT
CAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATC
GACACGCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAG
AACCCGGACGGCGGCTGGGGCGAGGACTGCCGCTCGTACGAGGATCCGGCGTAC
GCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGGCCTGGGCGCTGATGGCGCTC
ATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTC
GTGGAGACGCAGCGCCCGGACGGCGGCTGGGATGAGCCGTACTACACCGGCACG
GGCTACCCAGGGGATTTCTACCTCGGCTACACCATGTACCGCCACGTGTTTCCGA
CGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

Вариант SHC33 (SEQ ID NO: 39)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDR
VDRDRMEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQ
KALRFIQSQGGIESSRVLTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGS
WARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRAL

HGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPA
FIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAG
EWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPD
ERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHTPFCDFGEVTDPPSEDV
TAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSA
LKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAW
ALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGYPGDFYLGYTMYR
HVFPTLALGRYKQAIERR

Вариант F605W(SEQ ID NO: 170)

ATGGCTGAGCAGTTGGTGGAAGCGCCGGCCTACGCGCGGACGCTGGATCGCGCG
GTGGAGTATCTCCTCTCCTGCCAAAAGGACGAAGGCTACTGGTGGGGGCCGCTTC
TGAGCAACGTCACGATGGAAGCGGAGTACGTCCTCTTGTGCCACATTCTCGATCG
CGTCGATCGGGATCGCATGGAGAAGATCCGGCGGTACCTGTTGCACGAGCAGCG
CGAGGACGGCACGTGGGCCCTGTACCCGGGTGGGCCGCCGGACCTCGACACGAC
CATCGAGGCGTACGTCGCGCTCAAGTATATCGGCATGTCGCGCGACGAGGAGCC
GATGCAGAAGGCGCTCCGGTTCATTCAGAGCCAGGGCGGGATCGAGTCGTCGCG
CGTGTTCACGCGGATGTGGCTGGCGCTGGTGGGAGAATATCCGTGGGAGAAGGT
GCCCATGGTCCCGCCGGAGATCATGTTCCTCGGCAAGCGCATGCCGCTCAACATC
TACGAGTTTGGCTCGTGGGCTCGGGCGACCGTCGTGGCGCTCTCGATTGTGATGA
GCCGCCAGCCGGTGTTCCCGCTGCCCGAGCGGGCGCGCGTGCCCGAGCTGTACGA
GACCGACGTGCCTCCGCGCCGGCGCGGTGCCAAGGGAGGGGGTGGGTGGATCTT
CGACGCGCTCGACCGGGCGCTGCACGGGTATCAGAAGCTGTCGGTGCACCCGTTC
CGCCGCGCGGCCGAGATCCGCGCCTTGGACTGGTTGCTCGAGCGCCAGGCCGGA
GACGGCAGCTGGGGCGGGATTCAGCCGCCTTGGTTTTACGCGCTCATCGCGCTCA
AGATTCTCGACATGACGCAGCATCCGGCGTTCATCAAGGGCTGGGAAGGTCTAGA
GCTGTACGGCGTGGAGCTGGATTACGGAGGATGGATGTTTCAGGCTTCCATCTCG
CCGGTGTGGGACACGGGCCTCGCCGTGCTCGCGCTGCGCGCTGCGGGGCTTCCGG
CCGATCACGACCGCTTGGTCAAGGCGGGCGAGTGGCTGTTGGACCGGCAGATCA
CGGTTCCGGGCGACTGGGCGGTGAAGCGCCCGAACCTCAAGCCGGGCGGGTTCG
CGTTCCAGTTCGACAACGTGTACTACCCGGACGTGGACGACACGGCCGTCGTGGT
GTGGGCGCTCAACACCCTGCGCTTGCCGGACGAGCGCCGCAGGCGGGACGCCAT
GACGAAGGGATTCCGCTGGATTGTCGGCATGCAGAGCTCGAACGGCGGTTGGGG
CGCCTACGACGTCGACAACACGAGCGATCTCCCGAACCACATCCCGTTCTGCGAC

TTCGGCGAAGTGACCGATCCGCCGTCAGAGGACGTCACCGCCCACGTGCTCGAGT
GTTTCGGCAGCTTCGGGTACGATGACGCCTGGAAGGTCATCCGGCGCGCGGTGGA
ATATCTCAAGCGGGAGCAGAAGCCGGACGGCAGCTGGTTCGGTCGTTGGGGCGT
CAATTACCTCTACGGCACGGGCGCGGTGGTGTCGGCGCTGAAGGCGGTCGGGATC
GACACGCGCGAGCCGTACATTCAAAAGGCGCTCGACTGGGTCGAGCAGCATCAG
AACCCGGACGGCGGCTGGGGCGAGGACTGCCGCTCGTACGAGGATCCGGCGTAC
GCGGGTAAGGGCGCGAGCACCCCGTCGCAGACGGCCTGGGCGCTGATGGCGCTC
ATCGCGGGCGGCAGGGCGGAGTCCGAGGCCGCGCGCCGCGGCGTGCAATACCTC
GTGGAGACGCAGCGCCCGGACGGCGGCTGGGATGAGCCGTACTACACCGGCACG
GGCTTCCCAGGGGATTGGTACCTCGGCTACACCATGTACCGCCACGTGTTTCCGA
CGCTCGCGCTCGGCCGCTACAAGCAAGCCATCGAGCGCAGGTGA

Вариант F605W (SEQ ID NO: 171)

MAEQLVEAPAYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDR
VDRDRMEKIRRYLLHEQREDGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEEPMQ
KALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGS
WARATVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRAL
HGYQKLSVHPFRRAAEIRALDWLLERQAGDGSWGGIQPPWFYALIALKILDMTQHPA
FIKGWEGLELYGVELDYGGWMFQASISPVWDTGLAVLALRAAGLPADHDRLVKAG
EWLLDRQITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPD
ERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPNHIPFCDFGEVTDPPSEDV
TAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSA
LKAVGIDTREPYIQKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAW
ALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFPGDWYLGYTMYR
HVFPTLAIAGRYKQAIERR

SEQ ID NO: 166 (ZmoSHC1)

ATGGGTATTGACAGAATGAATAGCTTAAGTCGCTTGTTAATGAAGAAGATTTTCG
GGGCTGAAAAAACCTCGTATAAACCGGCTTCCGATACCATAATCGGAACGGATA
CCCTGAAAAGACCGAACCGGCGGCCTGAACCGACGGCAAAAGTCGACAAAACGA
TATTCAAGACTATGGGGAATAGTCTGAATAATACCCTTGTTTCAGCCTGTGACTG
GTTGATCGGACAACAAAAGCCCGATGGTCATTGGGTCGGTGCCGTGGAATCCAAT
GCTTCGATGGAAGCAGAATGGTGTCTGGCCTTGTGGTTTTTGGGTCTGGAAGATC
ATCCGCTTCGTCCAAGATTGGGCAATGCTCTTTTGGAAATGCAGCGGGAAGATGG
CTCTTGGGGAGTCTATTTCGGCGCTGGAAATGGCGATATCAATGCCACGGTTGAA

GCCTATGCGGCCTTGCGGTCTTTGGGGTATTCTGCCGATAATCCTGTTTTGAAAAA
AGCGGCAGCATGGATTGCTGAAAAAGGCGGATTAAAAAATATCCGTGTCTTTACC
CGTTATTGGCTGGCGTTGATCGGGGAATGGCCTTGGGAAAAGACCCCTAACCTTC
CCCCTGAAATTATCTGGTTCCCTGATAATTTTGTCTTTTCGATTTATAATTTTGCCC
AATGGGCGCGGGCAACCATGGTGCCGATTGCTATTCTGTCCGCGAGACGACCAAG
CCGCCCGCTGCGCCCTCAAGACCGATTGGATGAACTGTTTCCAGAAGGCCGCGCT
CGCTTTGATTATGAATTGCCGAAAAAAGAAGGCATCGATCTTTGGTCGCAATTTT
TCCGAACCACTGACCGTGGATTACATTGGGTTCAGTCCAATCTGTTAAAGCGCAA
TAGCTTGCGTGAAGCCGCTATCCGTCATGTTTTGGAATGGATTATCCGGCATCAG
GATGCCGATGGCGGTTGGGGTGGAATTCAGCCACCTTGGGTCTATGGTTTGATGG
CGTTACATGGTGAAGGCTATCAGCTTTATCATCCGGTGATGGCCAAGGCTTTGTC
GGCTTTGGATGATCCCGGTTGGCGACATGACAGAGGCGAGTCTTCTTGGATACAG
GCCACCAATAGTCCGGTATGGGATACAATGTTGGCCTTGATGGCGTTAAAAGACG
CCAAGGCCGAGGATCGTTTTACGCCGGAAATGGATAAGGCCGCCGATTGGCTTTT
GGCTCGACAGGTCAAAGTCAAAGGCGATTGGTCAATCAAACTGCCCGATGTTGA
ACCCGGTGGATGGGCATTTGAATATGCCAATGATCGCTATCCCGATACCGATGAT
ACCGCCGTCGCTTTGATCGCCCTTTCCTCTTATCGTGATAAGGAGGAGTGGCAAA
AGAAAGGCGTTGAGGACGCCATTACCCGTGGGGTTAATTGGTTGATCGCCATGCA
AAGCGAATGTGGCGGTTGGGGAGCCTTTGATAAGGATAATAACAGAAGTATCCTT
TCCAAAATTCCTTTTTGTGATTTCGGAGAATCTATTGATCCGCCTTCAGTCGATGT
AACGGCGCATGTTTTAGAGGCCTTTGGCACCTTGGGACTGTCCCGCGATATGCCG
GTCATCCAAAAAGCGATCGACTATGTCCGTTCCGAACAGGAAGCCGAAGGCGCG
TGGTTTGGTCGTTGGGGCGTTAATTATATCTATGGCACCGGTGCGGTTCTGCCTGC
TTTGGCGGCGATCGGTGAAGATATGACCCAGCCTTACATCACCAAGGCTTGCGAT
TGGCTGGTCGCACATCAGCAGGAAGACGGCGGTTGGGGCGAAAGCTGCTCTTCCT
ATATGGAGATTGATTCCATTGGGAAGGGCCCAACCACGCCGTCCCAGACTGCTTG
GGCTTTGATGGGGTTGATCGCGGCCAATCGTCCCGAAGATTATGAAGCCATTGCC
AAGGGATGCCATTATCTGATTGATCGCCAAGAGCAGGATGGTAGCTGGAAAGAA
GAAGAATTCACCGGCACCGGATTCCCCGGTTATGGCGTGGGTCAGACGATCAAGT
TGGATGATCCGGCTTTATCGAAACGATTGCTTCAAGGCGCTGAACTGTCACGGGC
GTTTATGCTGCGTTATGATTTTTATCGGCAATTCTTCCCGATTATGGCGTTAAGTC
GGGCAGAGAGACTGATTGATTTGAATAATTGA

**Таблица 18: расчеты процента идентичности последовательности с**

**помощью алгоритмов программ Blast и GAP для фермента WT AacSHC относительно ферментов WT SHC, раскрытых в WO2010/0139719 (BASF)**

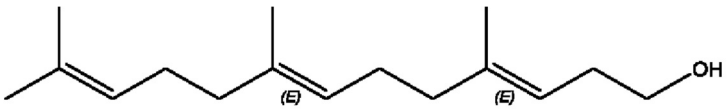
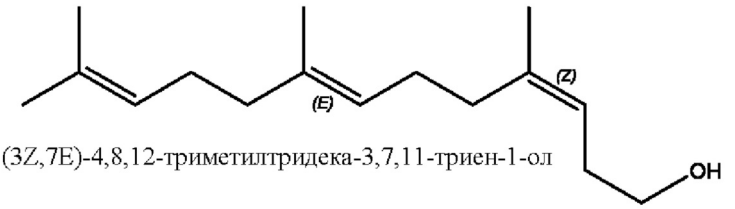
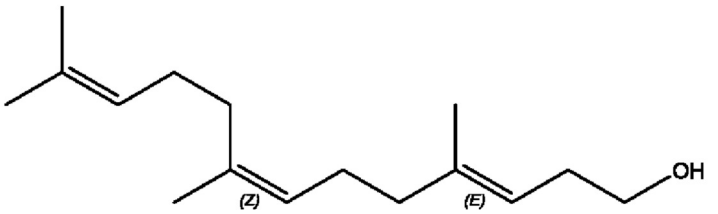
| Выравнивание последовательностей WT AacSHC (631AA, SEQ ID NO: 1) по сравнению с | % идентичности карт Blast | % идентичности GAP | Источник WT SHC (НАС) | Длина (АК) |
|---|---|---|---|---|
| WO 2010139719 SEQ ID NO: 1 | 44% | 40% | *Zymomonas mobilis* | 725 |
| WO 2010139719 SEQ ID NO: 2 | 44% | 40% | *Zymomonas mobilis* | 725 |
| WO 2010139719 SEQ ID NO: 5 | 44% | 42% | *Bradyrhizobium japonicum* | 684 |
| WO 2010139719 SEQ ID NO: 6 | 41% | 38% | *Burkholderia ambifaria 6* | 657 |
| WO 2010139719 SEQ ID NO: 7 | 40% | 37% | *Burkholderia ambifaria* | 682 |
| WO 2010139719 SEQ ID NO: 8 | 34% | 32% | *Bacillus anthracis* | 617 |
| WO 2010139719 SEQ ID NO: 9 | 37% | н. о. | *Frankia alni* | 720 |
| WO 2010139719 SEQ ID NO: 10 | 45% | 42% | *Rhodopseudomonas palent* | 685 |
| WO 2010139719 SEQ ID NO: 11 | 52% | 46% | *Streptomyces coelicolor* | 679 |

**Таблица 19: расчеты процента идентичности последовательности с помощью алгоритмов программы Blast и Huang и Miller для фермента WT AacSHC относительно ферментов WT ZmoSHC1 и WT ZmoSHC2, раскрытых в Seitz (2012)**

| Сравниваемые последовательности SHC | % идентичности карт Blast | Процент идентичности как раскрыто в Seitz (2012, диссертация) |
|---|---|---|
| AacSHC/HAC по сравнению с ZmoSHC1 | 44% | 41% |
| AacSHC/HAC по сравнению с ZmoSHC2 | 39% | 37% |
| AacSHC/HAC по сравнению с AacSHC | 98% | НО |
| ZmoSHC1 по сравнению с ZmoSHC2 | 36% | 34% |
| AacSHC по сравнению с ZmoSHC2 | 39% | 37% |

Диссертация Miriam Seitz доступна по ссылке http://elib.unistuttgart.de/handle/11682/1400

**Таблица 20: Номенклатура для гомофарнезола**

| Соединение | Сокращение | Наименование и структура |
|---|---|---|
| *E,E*-гомофарнезол | EEH |  (3E,7E)-4,8,12-триметилтридека-3,7,11-триен-1-ол |
| *E,Z*-гомофарнезол | EZH |  (3Z,7E)-4,8,12-триметилтридека-3,7,11-триен-1-ол |
| *Z,E*-гомофарнезол | ZEH |  (3E,7Z)-4,8,12-триметилтридека-3,7,11-триен-1-ол |

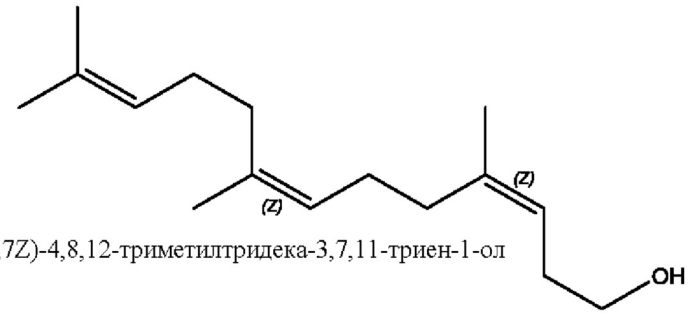
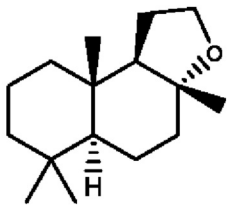
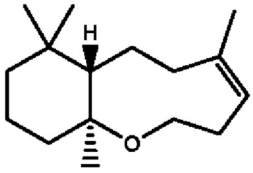
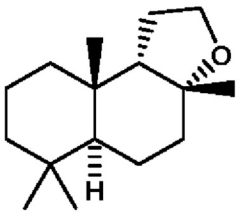
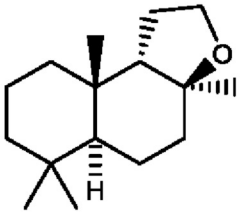
| | | |
|---|---|---|
| Z,Z-гомофарнезол | ZZH |  (3Z,7Z)-4,8,12-триметилтридека-3,7,11-триен-1-ол |

**Таблица 21: Номенклатура для продуктов реакции**

| Соединение | Описание | Наименование и структура |
|---|---|---|
| (I) | (-)-Амброксид |  (3aR,5aS,9aS,9bR)-3a,6,6,9a-тетраметилдодекагидронафто[2,1-b]фуран |
| (II) | Макроцикл |  (7aS,11aS,Z)-5,8,8,11a-тетраметил-2,3,6,7,7a,8,9,10,11,11a-декагидробензо[b]оксонин |
| (III) | 9b-эпи-амброксид |  (3aR,5aS,9aS,9bS)-3a,6,6,9a-тетраметилдодекагидронафто[2,1-b]фуран |
| (IV) | Escher *et al* (1990) |  (3aS,5aS,9aS,9bS)-3a,6,6,9a-тетраметилдодекагидронафто[2,1-b]фуран |

Escher S, Giersch W., Niclass Y, Bemardinello G and Ohloff G (1990). Configuration-odor

relationships in 5β-Ambrox. Helv. Chim. Acta 73, 1935-1947.

КРАТКОЕ ОПИСАНИЕ ГРАФИЧЕСКИХ МАТЕРИАЛОВ

Для лучшего понимания настоящего изобретения приведена ссылка на сопроводительные графические материалы, в которых:

На Фиг. 1-4 показано выравнивание последовательностей отдельных производных AacSHC относительно AacSHC SEQ ID NO: 1. На Фиг. 1 сверху вниз изображены SEQ ID NO:: 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 и SEQ ID NO: 39.

На Фиг. 5 показана карта плазмиды.

На Фиг. 6 показаны относительные значения активности HAC AacSHC дикого типа и производных AacSHC, представленных в таблице 24, в стандартных условиях (pH 6,0, 55°C, 0,050% SDS, клетки до значения $OD_{650\ нм}$ 10).

На Фиг. 7a показаны профили активности HAC производных AacSHC относительно WT AacSHC с использованием гомофарнезола качества EEH:EZH 87:13 и гомофарнезола чистоты 96% (определенной с использованием NMR).

На Фиг. 7b показано относительное повышение активности производных AacSHC относительно WT AacSHC (4 ч (начальная скорость) и выход через 22 ч) с использованием гомофарнезола качества EEH:EZH 87:13 и гомофарнезола чистоты 96% (определенной с использованием NMR).

На Фиг. 8a показаны профили активности HAC производных AacSHC относительно WT AacSHC с использованием гомофарнезола качества EEH:EZH 92:08 и гомофарнезола чистоты 100% (определенной методом NMR).

На Фиг. 8b показано относительное повышение активности производных AacSHC относительно WT AacSHC (4 ч (начальная скорость) и выход через 22 ч) с использованием гомофарнезола качества EEH:EZH 92:08 и гомофарнезола чистоты 100% (определенной методом NMR).

На Фиг. 9a показаны профили активности HAC производных AacSHC, приведенных в таблице 24, относительно WT AacSHC с использованием гомофарнезола качества EEH:EZH 66:33 и гомофарнезола чистоты 76% (определенной методом NMR).

На Фиг. 9b показано относительное повышение активности производных AacSHC относительно WT SHC (4 ч (начальная скорость) и выход через 22 ч) с использованием гомофарнезола качества EEH:EZH 66:33 и гомофарнезола чистоты 76% (определенной методом NMR).

На Фиг. 10 показаны результаты определения активности HAC для трех производных SHC, проявляющих приблизительно 10-кратное (215G2), 7-кратное (SHC26) и 6-кратное (SHC32) повышение активности по сравнению с ферментом AacSHC/HAC дикого типа.

На Фиг. 11 показана наблюдаемая конверсия E,E-гомофарнезола в амброксид производным SHC/HAC (215G2 SHC) и WT AacSHC. После 7 часов реакции (оценка начальной скорости реакции) конверсия вариантом 215G2 SHC была в 13 раз выше по сравнению с конверсией, достигаемой SHC дикого типа. После 48 часов реакции конверсия вариантом была приблизительно в 8 раз выше, чем для фермента дикого типа.

На Фиг. 12 показаны полученные продукты реакции (амброксид и продукт (IV)) при использовании EEH в качестве исходного материала (для биоконверсии WT SHC и/или производным SHC/HAC); и полученные продукты реакции ((-)-амброксид (I) и продукты (II), (IV) и (III) (см. таблицу 21) при использовании EE:EZ в качестве исходного материала); для простоты ссылки соединения I-IV можно идентифицировать следующим

образом:

I: (3aR,5aS,9aS,9bR)-3a,6,6,9a-тетраметилдодекагидронафто[2,1-b]фуран

(-)-Амброксид

II: (7aS, 11aS,Z)-5,8,8,11a-тетраметил-2,3,6,7,7a,8,9,10,11,11a-декагидробензо[b]оксонин

IV: (3aR,5aS,9aS,9bS)-3a,6,6,9a-тетраметилдодекагидронафто[2,1-b]фуран

III: (3aS,5aS,9aS,9bS)-3a,6,6,9a-тетраметилдодекагидронафто[2,1-b]фуран

9-эпи-Амброксид

На Фиг. 13 показан анализ GC продуктов реакции для амброксида и продуктов (II), (IV) и (III) в таблице 25.

На Фиг. 14 показан анализ GC продуктов реакции для амброксида и продуктов (II), (IV) и (III) в таблице 25.

На Фиг. 15 приведены сравнительные данные для активности варианта 215G2SHC в количественном анализе цельноклеточной биоконверсии в присутствии Тритона X-100 или SDS.

На Фиг. 16 показана процентная доля конвертированного EEH для различных соотношений SDS/клетки.

На Фиг. 17 показан % конверсии EEH в стандартной реакции биоконверсии (как описано в примере 7) для трех различных концентраций SDS.

На Фиг. 18 показан % конверсии EEH в стандартной реакции биоконверсии (как описано в примере 7) для трех различных значений pH.

На Фиг. 19 показана локализация мутаций, идентифицированных в вариантах SHC/HAC 101A10, 111C8 и 215G2, на кристаллической структуре SHC (в цвете): красный для варианта 215G2; пурпурный (винно-красный) для варианта 101A10 и зеленый для варианта 111C8. Для аминокислот, идентифицированных как ответственные за повышенную активность, боковые цепи выделены желтым цветом в сокристаллизованном аналоге субстрата. Другие мутации для идентифицированных вариантов без повышения активности отмечены синим цветом. Отмечено, что синие мутации распределены приблизительно пополам (т.е. 50:50) по 2 доменам фермента, при этом идентифицированные благоприятные мутации AacSHC локализованы, в основном (удаленно друг от друга), в домене 2. Единственное исключение составляет мутация F601Y, расположенная вблизи активного центра.

На Фиг. 20 показаны следующие мутации (черным и белым цветом): мутации, не оказывающие благоприятного действия на активность SHC/HAC, показаны черным цветом, они распределены по 2 доменам фермента SHC. Серым цветом показаны мутации, идентифицированные в вариантах SHC (101A10, 111C8 и 215G2), проявляющие повышенную активность SHC/HAC, и они расположены за единственным исключением в домене 2 фермента SHC. Выделена боковая цепь мутаций, вносящих вклад в повышенную активность вариантов.

На Фиг. 21 показана область клонирования и экспрессии плазмиды pET-28a(+); номера SEQ ID для последовательностей на Фиг. 21 приведены ниже:

pET 28a (нуклеотидная последовательность): SEQ ID NO: 179;

pET 28a (аминокислотная последовательность): SEQ ID NO: 180;

pET 28b (нуклеотидная последовательность): SEQ ID NO: 181;

pET 28b (аминокислотная последовательность): SEQ ID NO: 182;

pET 28c (нуклеотидная последовательность): SEQ ID NO: 183; и

pET 28c (аминокислотная последовательность): SEQ ID NO: 184.

На Фиг. 22 показана объемная производительность 1,5x концентрированной реакционной смеси биоконверсии EEH, содержащей 375 г/л клеток, 188 г/л EEH, 2,33%

SDS, по сравнению с объемной производительностью регулярной биоконверсии, которая была проведена параллельно при 125 г/л EEH, 250 г/л клеток, 1,55% SDS (пример 7).

На Фиг. 23 показана регулярная биоконверсия (125 г/л EEH, 250 г/л клеток, 1,55% SDS), которая была проведена, как описано в примере 7, но с заменой цитратного буферного раствора pH 5,4 на 0,5% или 0,9% NaCl, при этом все остальные параметры реакции не менялись. Параллельно в качестве контроля проводили биоконверсию в цитратном буферном растворе.

На Фиг. 24 показано изменение твердофазной экстракции (-)-амброксида по мере проведения отмывок толуолом в виде % от количества (-)-амброксида, исходно присутствовавшего в 200 мл всего реакционного бульона (за счет соотношения объемов бульон/толуол % в первом экстракте превышает 100%).

На Фиг. 25 показано изменение твердофазной экстракции (-)-амброксида по мере проведения отмывок этанолом в виде процентной доли от исходно присутствующего количества (-)-амброксида. После 4 отмывок (всего 640 мл EtOH, т.е. 3,2x начального объема всей реакционной смеси или 8x объема твердой фазы) было извлечено приблизительно 99% (-)-амброксида, первоначально присутствующего в реакционном бульоне.

ПРИМЕРЫ

Во избежание сомнений все ссылки на WT SHC и варианты SHC относятся к WT AacSHC (SEQ ID NO: 1) и его вариантам (например, перечисленным в таблице 23 и/или в таблице 24).

ПРИМЕР 1

Получение биокатализатора

Методы 1

Получение плазмиды SHC

Ген, кодирующий сквален-гопен-циклазу Alicyclobacillus acidocaldarius (AacSHC) встраивали в плазмиду pET-28a(+), где он находился под контролем IPTG-индуцибельного промотора T7 для продуцирования белка в Escherichia coli (см. Фиг. 5 и 21). Этой плазмидой трансформировали штамм E. coli BL21(DE3), используя стандартный протокол трансформации с тепловым шоком.

Культуры в колбах Эрленмейера

Для продуцирования белка использовали либо комплексную (LB), либо минимальную питательную среду. Одним из примеров успешно используемой минимальной питательной среды является среда M9.

Приготовление среды

Выбранную по умолчанию минимальную питательную среду готовили, как описано ниже, для 350 мл культуры: к 35 мл исходного раствора лимонной кислоты/фосфата (133 г/л $KH_2PO_4$, 40 г/л $(NH_4)_2HPO_4$, 17 г/л лимонной кислоты. $H_2O$ с pH, доведенным до 6,3) добавляли 307 мл $H_2O$, при необходимости pH доводили до 6,8 32% раствором NaOH. После стерилизации в автоклаве добавляли 0,850 мл 50% раствора $MgSO_4$, 0,035 мл раствора микроэлементов (состав приведен в следующем разделе), 0,035 мл раствора тиамина и 7 мл 20% раствора глюкозы.

Продуцирование биокатализатора SHC (продуцирование биокатализатора)

Для продуцирования биокатализатора в небольшом масштабе (SHC дикого типа или вариантов SHC) 350 мл культуры (питательной среды с добавлением 50 мкг/мл канамицина) инокулировали из предварительной культуры штамма E. coli BL21(DE3), содержащего плазмиду для продуцирования SHC. Клетки выращивали до оптической

плотности приблизительно 0,5 ($OD_{650\ нм}$) при 37°C при постоянном перемешивании (250 об/мин).

Затем продуцирование белка индуцировали добавлением IPTG до концентрации 300 мкМ с последующей инкубацией в течение 5-6 часов при постоянном встряхивании. Наконец, полученную в результате биомассу собирали центрифугированием, промывали 50 мМ буфером Трис-HCl pH 7,5. Клетки хранили в виде осадков при 4°C или -20°C до дальнейшего использования. Как правило, из 1 литра культуры получали от 2,5 до 4 грамм клеток независимо от используемой питательной среды.

Ферментации готовили и проводили в реакторах Infors HT емкостью 750 мл. В ферментационный сосуд добавляли 168 мл деионизированной воды. Реакционный сосуд был оборудован всеми необходимыми датчиками ($pO_2$, pH, отбора проб, гашения пены), устройством подачи C+N и бутылками с гидроксидом натрия и простерилизован в автоклаве. После стерилизации в автоклаве в реактор добавляют:

20 мл 10x буферного раствора фосфата/лимонной кислоты

14 мл 50% глюкозы

0,53 мл раствора $MgSO_4$

2 мл раствора $(NH_4)_2SO_4$

0,020 мл раствора микроэлементов

0,400 мл раствора тиамина

0,200 мл исходного раствора канамицина

Устанавливали следующие рабочие параметры: pH=6,95, $pO_2$=40%, T=30°C, перемешивание при 300 об/мин. Каскад: установка числа об/мин 300, минимум 300, максимум 1000, установка скорости потока л/мин 0,1, минимум 0, максимум 0,6. Контроль гашения пенообразования: 1:9.

В ферментер вносили посевную культуру до значения $OD_{650\ нм}$ 0,4-0,5. Эту посевную культуру выращивали в питательной среде LB (с канамицином) при 37°C, 220 об/мин в течение 8 ч. Сначала ферментацию проводили в периодическом режиме в течение 11,5 ч, после чего начинали подачу C+N с раствором для подпитки (стерилизованный раствор глюкозы (143 мл $H_2O$+ 35 г глюкозы), в который после стерилизации было добавлено: 17,5 мл раствора $(NH_4)_2SO_4$, 1,8 мл раствора $MgSO_4$, 0,018 мл раствора микроэлементов, 0,360 мл раствора тиамина, 0,180 мл исходного раствора канамицина. Подпитку проводили при постоянной скорости потока приблизительно 4,2 мл/ч. Измерения глюкозы и $NH_4^+$ проводили извне для оценки доступности источников C и N в культуре. Обычно концентрация глюкозы оставалась очень низкой.

Культуры выращивали в течение суммарно 25 часов, где они, как правило, достигали значения $OD_{650\ нм}$ 40-45. Затем продуцирование SHC начинали путем добавления в ферментер IPTG до концентрации приблизительно 1 мМ (в виде однократного добавления IPTG или в течение периода 3 4 часа с использованием инфузионного шприца), установки температуры на 40°C и $pO_2$ на 20%. Индукцию продуцирования SHC продолжали в течение 16 ч при 40°C. По окончании индукции клетки собирали центрифугированием, промывали 0,1 М буферным раствором лимонной кислоты/цитрата натрия с pH 5,4 и хранили в виде осадков при температуре 4°C или -20°C до дальнейшего использования.

Результаты 1a

Как правило, при всех прочих неизменных условиях удельная активность

продуцируемого биокатализатора была выше при использовании минимальной среды по сравнению с комплексной средой. Индукцию успешно осуществляли при температуре 30 или 37°C. Было отмечено, что при проведении индукции при 40-43°C получали биокатализатор с более высокой удельной активностью.

Результаты 1b

В приведенной ниже таблице 22 для 2 примеров показан объем культуры, оптическая плотность и количество клеток, как в начале, так и в конце индукции, а также количество собранной биомассы (сырая масса).

**Таблица 22**

| | Объем в начале индукции (мл) | $OD_{650\ нм}$ Расчетное количество | клеток в начале индукции (г) | Объем в конце индукции (мл) | $OD_{650\ нм}$ в конце индукции | Собранные клетки (г) |
|---|---|---|---|---|---|---|
| Пример 1 | 273 | 40 | 10,9 | 342 | 55 | 28 |
| Пример 2 | 272 | 44 | 12,0 | 341 | 57 | 23 |

$OD_{650\ нм}$ на момент инокуляции: 0,45 (пример 1) и 0,40 (пример 2). Начальные объемы: 205 мл.

ПРИМЕР 2

Получение вариантов SHC и скрининг активности

Методы 2

Во избежание сомнений EE соответствует (3E,7E); смесь EZ соответствует (3Z,7E); ZE соответствует (7Z,3E); ZZ соответствует (7Z,3Z); и EEH соответствует (3E,7E).

Программу изменения фермента проводили, используя ген SHC дикого типа (WT) Alicyclobacillus acidocaldarius (AacSHC) в качестве матрицы (GenBank M73834, Swissprot P33247). Была создана библиотека из приблизительно 10500 вариантов SHC, которую подвергали скринингу на поиск вариантов, проявляющих повышенную способность к замыканию цикла EEH. Скрининг проводили в реакциях в цитратном буфере с pH 6,0 (0,150 мл), содержащем 4 г/л EEH и 0,050% SDS, при 55°C и при постоянном перемешивании.

При отборе лучших вариантов для валидации стандартный тест проводили в цитратном буферном растворе pH 6,0, содержащем 4 г/л EEH, 0,050% SDS, клетки, экспрессирующие варианты SHC, до значения $OD_{600\ нм}$ 10,0. Конечный объем составлял 1 мл, реакционные смеси инкубировали при 55°C и энергично перемешивали с помощью магнитной мешалки. Отбор проб реакционной смеси в различные моменты времени позволял исследовать профили активности (конверсию EEH в (-)-амброксид), определяемые с помощью анализа методом газовой хроматографии (см. приведенные ниже методы анализа).

Из данного цикла валидации были получены 3 варианта с повышенной активностью замыкания цикла EEH (101A10, 111C8 и 215G2), и впоследствии в этих 3 вариантах было идентифицировано в общей сложности 8 мутаций. Впоследствии было проведено исследование мутаций, чтобы определить, какие из них полезны в отношении замыкания цикла EEH с образованием амброксида. Кроме этого производного AacSHC, был сконструирован другой вариант AacSHC, который содержал все идентифицированные полезные мутации (SHC33, как показано в таблице 23 ниже). Условия скрининга были

следующими: 4 г/л ЕЕН, клетки до $OD_{650\,нм}$ 10,0, SDS до 0,05% и 0,1% (2 концентрации), и реакции проводили при 55°C при постоянном перемешивании.

Результаты 2a

**Таблица 23: Мутации в оцениваемых производных фермента AacSHC**

| 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 |

Результаты 2b

Из трех отобранных мутаций (101A10, 111C8 и 215G2) лучшую активность показала мутация 215G2.

ПРИМЕР 3

Оптимизированные реакционные условия с вариантами SHC

Исследуемые параметры реакции: температура, концентрация SDS и pH

Методы 3

Реакционные условия для вариантов производных SHC, идентифицированных в таблице 23, были индивидуально оптимизированы в отношении температуры, pH и концентрации SDS. С этой целью клетки E. coli трансформировали плазмидой для продуцирования индивидуальных вариантов, культивировали их в колбах Эрленмейера и индуцировали продуцирование SHC, как описано выше. Таким образом, гарантировали, что все культуры содержат одинаковые или очень сходные количества SHC. Клетки собирали центрифугированием, промывали 0,1 М нитратным буферным раствором (pH 6,0) и хранили при -20°C до дальнейшего использования.

Результаты 3

Итоговые результаты этого исследования оптимизации приведены ниже в таблице. Цикл оптимизации также проводили с SHC дикого типа.

В следующей таблице 24 показаны оптимальные реакционные условия для дикого типа и вариантов, рассматриваемые для характеризации каждого производного фермента SHC/HAC.

**Таблица 24 Оптимальные реакционные условия для производных фермента SHC**

| SHC | Температура (°C) | pH | [SDS] (% масс/масс) |
|---|---|---|---|
| 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) |

Обсуждение 3

В примере 3 показаны различия, отмеченные в реакционных условиях для производных фермента SHC по сравнению с WT SHC. Для вариантов SHC наблюдали значимые отклонения оптимальных значений температуры, pH и концентрации SDS от SHC дикого типа. Лишь небольшое число мутаций оказали значимый эффект на оптимальные реакционные условия биоконверсии. Для определения оптимальных реакционных условий для отдельных вариантов SHC реакции проводили при нагрузке субстрата 4 г/л EEH и клеток, продуцирующих фермент SHC дикого типа или его производные, при оптимальной плотности $OD_{650\ нм}$ 10,0.

Температура

Данные в таблице 24 показывают неожиданный результат, что, тогда как фермент WT SHC обладает оптимальной активностью при 55°C (в диапазоне 45-60°C), ряд производных SHC обладает оптимальной активностью при 35°C (34-50°C). Применение производных SHC по настоящему изобретению в способах получения (-)-амброксида из E,E-гомофарнезола при пониженных температурах реакции обладает значительными экономическими преимуществами для получения (-)-амброксида в промышленном масштабе.

Солюбилизирующий агент

SDS был отобран и идентифицирован из длинного перечня возможных солюбилизирующих агентов, которые не были пригодными в реакции биоконверсии (дополнительную информацию см. в примере 14).

SDS лучше чем, например, Тритон X-100, в отношении скорости и выхода реакции (оба агента тестировали при использовании 4 г/л EEH и 125 г/л EEH, как приведено в примере 7).

ПРИМЕР 4

Тестирование активности вариантов SHC по сравнению с ферментом WT SHC в стандартных условиях

Методы 4

Продуцирование вариантов для сравнения относительной активности биокатализаторов (как описано в таблице 24) описано ниже. Клетки E. coli трансформировали плазмидой для продуцирования одного из вариантов SHC, и впоследствии клетки E. coli культивировали в питательной среде LB при 37°C и 280 об/мин, выращивали до значения $OD_{650нм}$ 0,50 и индуцировали продуцирование фермента добавлением IPTG. Индукцию продолжали в течение 5,5 часов при 37°C, 280 об/мин. Клетки собирали центрифугированием, промывали 0,1 M цитратным буферным раствором pH 6,0 и хранили при -20°C до дальнейшего использования. При сравнении активности вариантов SHC (см. Фиг. 6) образец реакционной смеси наносили на полиакриламидный гель с SDS (SDS-PAGE) для анализа содержания SHC в реакционных смесях. Этот анализ подтвердил, что все реакционные смеси содержат идентичные количества фермента SHC.

Результаты 4a

На Фиг. 6 показаны относительные значения активности SHC дикого типа и его вариантов в стандартных условиях (pH 6,0, 55°C, 0,050% SDS, клетки до значения $OD_{650нм}$ 10). Было также отмечено, что SHC дикого типа и по меньшей мере протестированные варианты SHC в соответствии с примерами настоящего изобретения толерантны к растворителю. Это означает, что в биоконверсионную реакционную смесь можно добавлять выбранные не смешиваемые с водой растворители (вплоть до почти 100%).

Результаты 4b

При использовании варианта 215G2 SHC заметного эффекта на активность этого варианта при добавлении в реакционную смесь NaCl (тестируемые концентрации от 5 до 100 мМ (только)) не наблюдали. Кроме того, при добавлении NaCl вплоть до 100 мМ или до 154 мМ (0,9% NaCl) отрицательного влияния на активность SHC для варианта 215G2 не показано. Этот результат позволяет предположить, что при проведении реакции биоконверсии в физиологическом растворе NaCl (0,9%) или подобном ему растворе и поддержании pH при соответствующем значении (например, приблизительно 5,4 (5,2-5,6)) реакцию биоконверсии можно проводить в отсутствие буфера, но в присутствии физиологического раствора NaCl или подобного ему раствора.

Обсуждение 4

На Фиг. 6 проиллюстрировано ранжирование активности отдельных вариантов и ферментов SHC дикого типа в отношении конверсии EEH в (-)-амброксид.

ПРИМЕР 5

Профили активности WT SHC и производного SHC

Методы 5

Тест на активность проводили в 0,1 M цитратном буфере в объеме 5 мл при постоянном встряхивании при 900 об/мин на аппарате Heidolph Synthesis 1. Значение pH используемого буферного раствора, температура, при которой проводили реакцию, и концентрация SDS (додецилсульфата натрия) в реакционной смеси зависели от используемого варианта SHC (дикого типа или варианта). Сводные данные оптимальных условий для каждого из тестируемых вариантов приведены выше в таблице 24.

В качестве исходного материала использовали гомофарнезол чистоты 96% и субстрат гомофарнезол с соотношением EEH:EZH 87:13.

Во избежание сомнений, смесь EE:EZ представляет собой смесь изомеров ((3E,7E) и

(3Z,7E).

Результаты 5

Используемый гомофарнезол: EEH:EZH 87:13, чистота (NMR): 96%.

Результаты проведения стандартного теста в оптимизированных условиях, показанные на Фиг. 7B (профили активности производных SHC относительно WT SHC) и на Фиг. 7A, показывают относительное повышение активности производных SHC относительно WT SHC (4 ч (начальная скорость) и выход через 22 ч).

Используемый гомофарнезол: EEH:EZH 92:08, чистота (NMR): 100%

Результаты проведения стандартного теста в оптимизированных условиях, показанные на Фиг. 8B (профили активности производных AacSHC относительно WT AacSHC) и на Фиг. 8A, показывают относительное повышение активности производных AacSHC относительно WT SHC (4 ч (начальная скорость) и выход через 22 ч).

Используемый гомофарнезол: EEH:EZH 66:33, чистота (NMR): 76%

Результаты проведения стандартного теста в оптимизированных условиях, показанные на Фиг. 9B (профили активности производных AacSHC, как показано в таблице 24, относительно WT AacSHC) и на Фиг. 9A, показывают относительное повышение активности производных AacSHC относительно WT SHC (4 ч (начальная скорость) и выход через 22 ч).

Обсуждение 5

Основной вывод состоит в том, что независимо от качества используемого субстрата гомофарнезола четыре лучших производных фермента SHC были ранжированы в следующем порядке: 215GSHC, SHC26, SHC32 и SHC3.

ПРИМЕР 6

Определение баланса массы из реакционных смесей, полностью экстрагированных растворителем

Методы 6

Все условия оставались неизменными, и для каждого варианта проводили 2 реакции. В качестве субстрата использовали гомофарнезол. После 4 часов и 22 часов инкубации продукт реакции и непрореагировавший субстрат полностью экстрагировали для каждого из вариантов отмывками в общем количестве 6 равным объемом трет-бутил-метилового эфира (MTBE/tBME). Содержание гомофарнезола и амброксида в каждой из отмывок определяли с помощью анализа GC. Общее количество образовавшегося амброксида и остаточного гомофарнезола рассчитывали по калибровочным кривым, которые строили с использованием растворов, приготовленных из аутентичного амброксида и гомофарнезола.

Результаты 6

Результаты на Фиг. 10 показали, что с используемым субстратом было подтверждено, что 3 лучших варианта проявляют приблизительно 10-кратное (215G2), 7-кратное (SHC26) и 6-кратное (SHC32) повышение активности по сравнению с ферментом SHC дикого типа.

ПРИМЕР 7

Результативность биотрансформации при 125 г/л E,E-гомофарнезола (EEH)

Методы 7

Цель состояла в определении повышения волюметрической производительности с использованием варианта 215G2SHC. Для оптимизации условий тестовой реакции, включающих параметры pH, концентрацию клеток и концентрацию SDS, был создан дизайн серии экспериментальных исследований (ДЭИ). Реакционные условия были следующими: 125 г/л EEH (из гомофарнезола EE:EZ 86:14), 250 г/л клеток, 1,55% SDS,

при этом реакцию проводили при 35°C в 0,1 М цитратном буфере pH 5,4.

Типичную реакцию (общий объем 150 г) проводят, как описано ниже, в ферментерах Infors емкостью 0,75 л. В реакционный сосуд загружают необходимое количество гомофарнезола, соответствующее 18,75 г EEH. Добавляют 2,33 г SDS из 15,5% (масс/масс.) раствора, приготовленного в 0,1 М цитратном буферном растворе pH 5,4. Суспензию клеток готовят из клеток E. coli, которые продуцировали вариант 215G2 SHC, путем суспендирования клеток в 0,1 М цитратном буферном растворе pH 5,4. После определения сырой массы клеток этой суспензии центрифугированием в течение 10 мин при 10°C и 17210g в реакционный сосуд добавляют необходимый объем клеток с целью введения в реакционную смесь 37,5 г клеток. Объем реакционной смеси дополняют до 150 г необходимым количеством реакционного буферного раствора. Реакцию проводят при 37°C при постоянном перемешивании при 900 об/мин. Регулирование pH проводят, используя 40% водный раствор лимонной кислоты. Образцы реакционной смеси (1 мл) отбирают в различные моменты времени, экстрагируют 5 объемами MTBE/tBME (5 мл). Содержание гомофарнезола и амброксида в реакционной смеси определяли с помощью анализа методом GC после осветления фазы растворителя центрифугированием (настольная центрифуга, 13000 об/мин, 2 мин), 10-кратное разведение MTBE/tBME.

Такую же реакцию проводили с клетками E. coli, которые продуцировали фермент SHC дикого типа. В этом случае реакцию проводили при 55°C в 0,1 М цитратном буферном растворе pH 6,0. Сводные данные реакционных условий для этого примера приведены в строке 2 таблицы 24a ниже. Реакционные условия, представленные в строке 1 таблицы 24a ниже, взяты из предыдущих примеров (например, примеров 3-5).

**Таблица 24a: в строке 2 представлены реакционные условия для примера 7**

| SHC | Температура (°C) | pH | [SDS] (%, масс./масс.) | [EEH] (г/л) | Биокатализатор (г/л) (сырая масса клеток) |
|---|---|---|---|---|---|
| 215G2 | 35°C | 5,4 (5,0–6,2) | 0,06 | 4 | 1,45 клеток |
| 215G2 | 35°C | 5,4 (5,0-6,2) | 1,55 | 125 | 250 клеток |

Результаты 7

На Фиг. 11 представлена наблюдаемая конверсия EEH в амброксид 2 ферментами. После 7 часов реакции (оценка начальной скорости реакции) конверсия вариантом 215G2 SHC была в 13 раз выше по сравнению с конверсией, достигаемой SHC дикого типа. После 48 часов реакции конверсия вариантом была приблизительно в 8 раз выше, чем для фермента дикого типа.

Общие примечания 7

Концентрация клеток

Все концентрации клеток (г/л) в реакционных смесях, описанных в этом примере, указаны в виде сырой массы клеток. Концентрацию в виде сырой массы клеток (г/л) суспензии клеток определяют после центрифугирования образца этой суспензии клеток в течение 10 мин при 17210g и 4°C.

Корреляция между г/л клеток и $OD_{650\ нм}$

При использовании биоконверсии 125 г/л EEH с 215G2 SHC или WT SHC 250 г/л клеток в этой реакции соответствует значению $OD_{650\ нм}$ приблизительно равному 172

в данной реакции. При тестировании различных препаратов биокатализатора наблюдали изменения в отношении значения $OD_{650\ нм}$ к количеству биокатализатора. При использовании биокатализатора при 4 г/л EEH, но применяя клетки до значения $OD_{650\ нм}$ 10,0, определили, что $OD_{650\ нм}$ 10,0 эквивалентна 1,45 г/л клеток.

Обсуждение 7

Данные показывают, что оптимизированный и эффективный способ биоконверсии HAC был разработан с использованием относительно высоких концентраций субстрата EEH (125 г/л) по сравнению с документами уровня техники, где раскрыта концентрация субстрата гомофарнезола, составляющая лишь от приблизительно 0,2 г/л (см. JP 2009060799) до приблизительно 2,36 г/л (10 мМ) (см. WO 2010/139719 A2, US 2012/0135477 A1) и Seitz et al (2012), цитируется выше).

ПРИМЕР 8

Анализ GC

Методы 8

Для количественного определения содержания EEH и амброксида образцы экстрагировали соответствующим объемом трет-бутилметилового эфира (MBTE/tBME). Фракцию растворителя отделяли от водной фазы центрифугированием, после чего проводили анализ методом газовой хроматографии. 1 мкл фазы растворителя инъецировали (коэффициент деления потока 3) на колонку 30 м × 0,32 мм × 0,25 мкм Zebron ZB-5. Хроматографию на колонке проводили при постоянном потоке (4 мл/мин $H_2$) с градиентом температуры: 100°C, 15°C/мин до 200°C, 120°C/мин до 240°C, 4 мин при 240°C, что приводило в результате к разделению амброксида, EEH и EZH. Температура на входе составляла 200°C, температура детектора: 300°C.

Конверсию EEH рассчитывали, исходя из площадей пиков, соответствующих амброксиду и EEH, по следующей формуле:

конверсия (%) = 100 × (Площадь$_{амброксид_пик}$/(Площадь$_{амброксид_пик}$ + Площадь$_{E,E\text{-}гомофарнезол_пик}$))

Подлинность продукта реакции амброксида подтверждали посредством GC-MC (зарегистрированные значения и интенсивности: 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%)).

Обсуждение 8

Извлечение продукта проводили либо путем экстракции растворителем, либо экстракции паром. Используемые растворители представляли собой, например, MTBE или гексан: изопропанол (3:2). Реакционную смесь многократно экстрагировали равными объемами растворителя, и фракции растворителя подвергали анализу GC до тех пор, пока ни субстрат, ни продукт больше не обнаруживались. Как правило, было достаточно от 5 до 6 отмывок. Альтернативно, экстракцию продуктов реакции осуществляли паром.

ПРИМЕР 9

Реакция, протекающая в одном реакционном сосуде

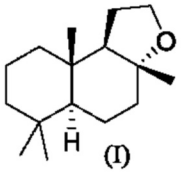
Методы 9

Проводили ферментацию 200 мл с E. coli BL21 (DE3), трансформированной плазмидой pET28a(+) 215G2 SHC для продуцирования 215G2 SHC с N-концевой His-меткой, используя стандартный протокол выращивания и индукции, описанный выше. В конце фазы индукции аэрацию отключали, температуру устанавливали на 35°C, pH доводили до 5,5 лимонной кислотой, а скорость мешалки - до 500 об/мин. Объем культуры рассчитывали с учетом всех добавлений, проведенных в процессе выращивания культуры

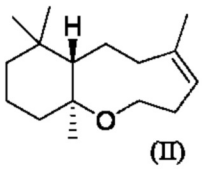
(подпитки и потребления основания). В соответствии с этим объемом и OD культуры в ферментер добавляли необходимое количество SDS. Добавляли EEH до 4 г/л. Из реакционной смеси отбирали образцы в различные моменты времени, образцы (150-300 мкл) экстрагировали 700 мкл MTBE для анализа GC. Конверсия EEH в амброксид происходила непосредственно в бульоне для культивирования. Реакцию проводили в общей сложности в течение 22,5 дней, в течение которых многократно добавляли EEH.

Результаты 9

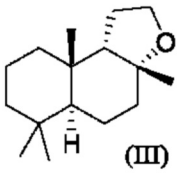
По достижении завершения реакции 10,6 г EEH циклизовалось с получением амброксида. Продукты реакции (структуры приведены ниже) экстрагировали паром и количественно извлекали из реакционной смеси.



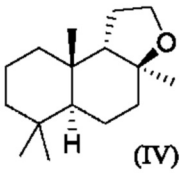

(I)




(II)




(III)




(IV)

Примечание относительно продуктов реакции

При конверсии гомофарнезола EE:EZ 87:13 посредством SHC образуются продукты реакции: амброксид, продукты (II), (IV) и (III), как представлено на Фиг. 12, и их соотношение отражает соотношение EE:EZ исходного материала.

При использовании в качестве исходного материала EEH образуются только (-)-амброксид (I) и продукт (IV).

При использовании в качестве исходного материала EZH (3Z,7E) образуются только продукты (II) и (III).

Однако при использовании смеси EEH и EZH образуются амброксид (I) и продукты (II), (IV) и (III).

При 100% конверсии EE:EZ 66:34 будет получено 66%:34% ((амброксид + (IV)):((II) + (III)).

При проведении экстракции паром экстрагируются все 4 продукта амброксид и продукты (II), (IV) и (III), а в результате этапа кристаллизации образуется амброксид чистоты 99% (GC) при выходе по меньшей мере 70%.

Обсуждение 9

Эти данные демонстрируют, что получение (-)-амброксида возможно в биоконверсионной реакционной среде или в реакционной системе «реакции, протекающей в одном реакционном сосуде», и что селективное обогащение амброксидом достигается в результате экстракции паром и кристаллизации.

Если исходное вещество гомофарнезола представляет собой смесь изомеров EE и EZ (например, 86:14), из каждого из этих изомеров образуется 2 продукта (всего 4), при этом (-)-амброксид, тем не менее, является основным компонентом неочищенного продукта и преобладающим компонентом в кристаллизованном материале (чистота 99,1%). (+)-Амброксид не был обнаружен.

ПРИМЕР 10

Конверсия смесей гомофарнезола EE:EZ

Во избежание сомнений, EE соответствует (3E,7E); смесь EZ соответствует (3Z,7E);

ZE соответствует (7Z,3E); ZZ соответствует (7Z,3Z); EEH соответствует (3E,7E); и EZH соответствует (3Z,7E).

Методы 10

Смеси EE:EZ подвергали биоконверсии в следующих реакционных условиях: суммарное содержание гомофарнезола составляло 146 г/л при концентрации клеток 250 г/л и SDS 1,55%, и использовали следующие субстраты гомофарнезола (смеси гомофарнезола EE:EZ):

EE:EZ 86:14 (самое высокое содержание EEH для данного примера),

EE:EZ 69:31 (самое низкое содержание EEH для данного примера),

EE:EZ 80:20

EEH:EZH 70:30

Биоконверсия смеси гомофарнезола 7E, 3E/7E, 3Z

Биоконверсию проводили с использованием следующих реакционных условий:

Реакцию (суммарный объем 150,1 г) проводили в 0,1 М буферном растворе лимонной кислоты/цитрата натрия pH 5,4 в ферментере Infors HT емкостью 750 мл, содержащем общее количество гомофарнезола 146 г/л, используя субстрат гомофарнезол, который представлял собой смесь 7E,3E:7E,3Z 86:14, 250 г/л клеток (полученных в соответствии со способом примера 1) и 1,55% SDS. Реакцию проводили при 35°C при постоянном перемешивании (800 об/мин), регулирование pH проводили, используя 10-40% водный раствор лимонной кислоты. Из реакционной смеси отбирали образцы в различные моменты времени, образцы экстрагировали растворителем для анализа GC. Было отмечено, что конверсия гомофарнезола происходит одинаково быстро при 2 качественных вариантах гомофарнезола (EE:EZ 86:14 и EE:EZ 69:31).

Результаты 10

Конверсию обоих изомеров E,E- и E,Z-гомофарнезола наблюдали при проведении биоконверсии 125 г/л E,E-гомофарнезола из материала EEH:EZH 86:14 с использованием WT SHC и одного конкретного производного SHC (215G2 SHC). Таким образом, фермент SHC дикого типа из Alicyclobacillus acidocaldarius образует такие же продукты реакции (т.е. амброксид, продукты (II), (IV) и (III)) из материала EEH:EZH 86:14, как и варианты SHC из таблицы 23 из смесей EEH:EZH. На Фиг. 13 и 14 приведен анализ GC продуктов реакции для амброксида и продуктов (II), (IV) и (III).

Обсуждение 10

В результате биоконверсии гомофарнезола в амброксид в соответствии с настоящим изобретением образуется (-)-амброксид как преобладающее соединение, но могут также образоваться идентифицированные выше соединения, отличающиеся от (-)-амброксида (например, соединения (II), (IV) и (III)) которые могут придавать или не придавать благоприятные ольфакторные ноты продукту (-)-амброксиду. Как показано выше, в условиях селективной кристаллизации амброксид можно отделить от других побочных продуктов ((II), (IV) и (III)). Соответственно, в случае отрицательного влияния на сенсорный характер конечного продукта амброксида селективное отделение продуктов (II), (IV) и (III) от конечного продукта (-)-амброксида повышает его ценность в качестве аорматизирующего пордукта, или корригента, или косметического продукта, или потребительского продукта для ухода. Сенсорный анализ проводят с использованием хорошо разработанных сенсорных тестов, используемых подготовленными парфюмерами. Чистота конечного продукта (-)-амброксида может быть показателем ольфакторного качества продукта в том случае, если продукт сам по себе, в основном, определяет желаемый сенсорный профиль.

ПРИМЕР 11

Конверсия EEH из смеси EE:EZ:ZE:ZZ-гомофарнезола

Методы 11

EE:EZ:ZE:ZZ-Гомофарнезол 40:26:20:14 использовали в качестве субстрата для конверсии EEH с использованием 215G2 SHC. В сравнительных целях также использовали другой гомофарнезол EE:EZ 2:1 или 93:07.

Конверсию смеси EE:EZ:ZE:ZZ-гомофарнезола исследовали с вариантом 215G2 SHC, но не в оптимизированных условиях. Реакционные условия были следующими: pH 5,8 в 100 мМ цитратном буфере, 0,10% SDS, 40°C. При проведении всех реакций при постоянной концентрации 2 г/л EEH (с соответственно измененными общими концентрациями гомофарнезола) наблюдали описанные ниже коэффициенты конверсии EEH.

Результаты 11

Наблюдали следующие коэффициенты конверсии смеси изомеров гомофарнезола:

| | |
|---|---|
| EE:EZ2:150-55% | EE:EZ2:150-55% |
| EE:EZ93:778% | EE:EZ93:778% |
| EE:EZ:ZE:ZZ 40:26:20:146% | EE:EZ:ZE:ZZ 40:26:20:146% |

Обсуждение 11

Помимо наблюдаемых значений выхода, эти данные показывают, что вариант 215G2 SHC способен осуществлять конверсию EEH в амброксид из комплексной смеси EE:EZ:ZE:ZZ гомофарнезола. Как и ожидали, наблюдали, что более низкий коэффициент конверсии приводил в результате к снижению выхода амброксида. Этот результат согласуется с точкой зрения, что изомеры гомофарнезола, отличающиеся от EEH, могут конкурировать с EEH за доступ к производным фермента SHC/HAC и, следовательно, могут действовать как конкурентные ингибиторы и/или альтернативные субстраты конверсии EEH в (-)-амброксид.

ПРИМЕР 12

Сравнительные данные для цельноклеточной биоконверсии с использованием Тритона X-100 и SDS

Методы 12

Клетки-хозяева E. coli выращивали в соответствии с протоколом Методов 4 из Примера 4. Реакцию биоконверсии с использованием варианта 215G2SHC проводили в соответствии со стандартным тестом примера 4. В качестве наиболее подходящих реакционных условий для варианта 215G2 SHC были выбраны: концентрация гомофарнезола 4 г/л при количестве клеток до значения $OD_{650\ нм}$ 10,0 в буферном растворе лимонной кислоты/фосфата натрия 0,1 М pH 5,4, 35°C и SDS в концентрации 0,07%.

Результаты 12

На Фиг. 15 приведено сравнение активности варианта 215G2SHC в количественном анализе цельноклеточной биоконверсии при использовании Тритона X-100 в диапазоне концентрации от 0,005% до 0,48% и SDS в концентрации 0,07%.

Обсуждение 12

Данные показывают, что максимальная активность с использованием Тритоном X-100 составляла лишь около 20% активности, полученной с использованием SDS.

ПРИМЕР 13

Отношение SDS/клетки

Методы 13

Реакцию биоконверсии осуществляли в соответствии с Методами 4 из Примера 4,

используя концентрацию субстрата EEH 4 г/л, клетки при значении $OD_{650\ нм}$ 5,0, в результате которой получили производное фермента 215G2 SHC.

Результаты 13

Результаты, представленные на Фиг. 16, показывают процентную долю конвертированного EEH для различных соотношений SDS/клетки.

На Фиг. 16 показано, что процентная доля конверсии EEH в (-)-амброксид с использованием различных значений соотношения SDS/клетки зависит от соотношения SDS/клетки. Это соотношение должно быть тщательно установлено для достижения максимальной конверсии.

Например, в случае слишком низкой концентрации SDS может наблюдаться субоптимальная конверсия гомофарнезола. С другом стороны, например, при слишком высокой концентрации SDS может быть риск повреждения биокатализатора за счет разрушения интактной клетки микроорганизма и/или денатурации/инактивации фермента SHC/HAC. При проведении реакции биоконверсии в соответствии с Методами 7 из Примера 7 с использованием 125 г/л EEH и 250 г/л биокатализатора наилучший протокол биоконверсии показал соотношение [SDS]/[клетки] 16:1.

Обсуждение 13

Результаты показывают, что существует некоторая степень независимости между концентрацией солюбилизирующего агента (SDS), количеством биомассы и концентрацией субстрата (EEH). В качестве примера, при повышении концентрации субстрата гомофарнезола для эффективного протекания реакции биоконверсии требуются достаточные количества биокатализатора и солюбилизирующего агента (SDS).

ПРИМЕР 14

Тестирование возможных солюбилизирующих агентов для применения в реакции биоконверсии

Методы 14

В качестве возможного заменителя SDS тестировали различные солюбилизирующие агенты (представленные в таблице 26 ниже) в реакциях замыкания цикла EEH ферментом 215G2 SHC, используя такие же условия, как в стандартном тесте (4 г/л EEH, клетки до значения $OD_{650\ нм}$ 10,0). С использованием стандартного теста также тестировали возможность усиления активности (кумулятивный эффект) за счет комбинирования SDS при его оптимальной концентрации (0,060-0,070%) с другими используемыми солюбилизирующими агентами, оптимальные концентрации которых определяли индивидуально на основании скрининга, проведенного с этими соединениями (см. таблицу 26 ниже). Кроме того, также тестировали «глубокие эвтектические растворители» и ионные жидкости, известные как способствующие солюбилизации нерастворимых в воде соединений.

Результаты 14

В следующей таблице 26 приведены сводные данные о солюбилизирующих агентах (например: поверхностно-активных веществах, детергентах, усилителях растворимости и т.п.), протестированных к настоящему времени в реакциях замыкания цикла EEH ферментом 215G2 SHC. Ни в одном случае не наблюдали повышения активности по сравнению с контрольной реакцией, проведенной с использованием SDS в диапазоне концентрации 0,060-0,070%. Активность, наблюдаемая при использовании этих соединений по отдельности при концентрации, определенной как оптимальная, составляли лишь приблизительно 20% от полученных в контрольных реакциях с SDS. Было отмечено, что при полном отсутствии добавления солюбилизирующего агента

достигалась 20% конверсия EEH. При использовании SDS и добавлении дополнительного солюбилизирующего агента (при концентрации, определенной в тесте как оптимальная) синергического эффекта не наблюдали. Наблюдали, скорее, снижение процентной доли конверсии EEH. На основании этого исследования был сделан вывод, что в условиях тестирования соединения вовсе не улучшают конверсию EEH; скорее, было получено их нежелательное воздействие на замыкание цикла, и из исследованных солюбилизирующих агентов SDS наиболее полезен. Кроме того, при тестировании «глубоких эвтектических растворителей» и ионных жидкостей, которые, как известно, способствуют солюбилизации нерастворимых в воде соединений, также не было получено положительных результатов.

**Таблица 26: приведен перечень солюбилизирующих агентов, которые были протестированы в реакции биоконверсии**

| **Солюбилизирующий агент** | **Тестируемый диапазон концентрации** |
|---|---|
| Каприлилсульфобетаин | 0,19–3,0% |
| CHAPS | 0,020–0,18% |
| Cremophor EL | 0,0063–0,5% |
| Диметилсульфоксид | 0,00032–0,2% |
| Гексадецилпиридиния хлорида моногидрат | 0,013–0,5% |
| Миристилсульфобетаин | 0,0009–0,03% |
| Нонидет P40 | 0,005–0,16% |
| Октил-β-D-глюкопиранозид | 0,0008–0,6% |
| Пальмитилсульфобетаин | 0,00000003–0,03% |
| Pluronic P-105 | 0,000074–0,018% |
| Соли четвертичного аммония (например, тетраметиламмония бромид) | 20–160 мМ |
| Тауродезоксихолата натрия гидрат | 0,05 0,4% |
| Stepan® | 0,01 0,6% |
| Thesit® | 0,05 0,8% |
| Тритон X-100 | 0,005 0,32% |
| Твин 20 | 0,05 0,8% |
| Твин 80 | 0,01 0,4% |

Обсуждение 14

Автором изобретения был выбран и идентифицирован SDS в качестве полезного солюбилизирующего агента из длинного перечня других солюбилизирующих агентов, которые не оказались полезными в реакции биоконверсии гомофарнезола в (-)-амброксид по настоящему изобретению.

ПРИМЕР 15

Чувствительность реакции биоконверсии к концентрации SDS

Методы 15

Применяемыми условиями являются условия стандартной биоконверсии (как описано в примере 7) при 125 г/л и 250 г/л биокатализатора и 1,55% SDS. Также были протестированы две другие концентрации SDS (1,40% и 1,70% SDS). Все концентрации SDS приведены в % масс./масс.

Условия стандартной реакции биоконверсии (как описано в примере 7) при 125 г/л и 250 г/л биокатализатора и 1,55% SDS также использовали для тестирования различных значений pH.

Контрольную реакцию проводили при pH 5,4 в 0,1 М цитратном буферном растворе. Для проведения реакций при низком pH использовали 0,1 М ацетатный буферный раствор.

Результаты 15

Данные, представленные на Фиг. 17, показывают, что реакция биоконверсии оказалась менее чувствительной к изменениям концентрации SDS, чем при тестировании активности HAC в стандартном тесте при 4 г/л EEH и применяемом количестве клеток при $OD_{650\ нм}$ 10,0.

Данные, представленные на Фиг. 18, показывают, что при применении реакции биоконверсии оказалась менее чувствительной к изменениям pH, чем при тестировании активности HAC в стандартном тесте при 4 г/л EEH и применяемом количестве клеток при $OD_{650\ нм}$ 10,0.

Обсуждение 15

Эти данные показывают робастность реакции биоконверсии при 125 г/л EEH и 250 г/л клеток по отношению к тестируемым показателям - диапазону концентрации SDS и диапазону pH.

ПРИМЕР 16

Локализация идентифицированных мутаций SHC/HAC на кристаллической структуре

Положения мутаций, идентифицированных в вариантах AacSHC/HAC, отмечены на Фиг. 19 следующим образом: красным для варианта 215G2; пурпурным (винно-красным) для варианта 101A10 и зеленым для варианта 111C8. Для аминокислот, идентифицированных как ответственные за повышенную активность, боковые цепи выделены желтым цветом в сокристаллизованном аналоге субстрата. Другие мутации для идентифицированных вариантов без повышения активности отмечены синим цветом. Отмечено, что синие мутации распределены приблизительно пополам (т.е. 50:50) по 2 доменам фермента, при этом идентифицированные благоприятные мутации AacSHC локализованы, в основном (удаленно друг от друга), в домене 2. Единственное исключение составляет мутация F601Y, расположенная вблизи активного сайта. Если рассматривать только оба производных фермента SHC/HAC, представляющих собой 215G2 и 111C8, все мутации локализуются в домене 2. На Фиг. 20 та же информация представлена черным и белым цветом.

Результаты 16

Все полезные мутанты (красный/зеленый/пурпурный), соответствующие 215G2, 111C8 и 101A10, локализованы, в основном (кроме одного мутанта F601Y), в домене 2 (Wendt et at (1997) Science 277: 1811) кристаллической структуры SHC (как показано на Фиг. 19). Комбинации полезных мутаций SHC пронумерованы в соответствии с AacSHC дикого типа (SEQ ID NO: 1).

Обсуждение 16

Кристаллическая структура полезна для идентификации производных SHC/HAC с требуемым соотношением структуры/активности, особенно в отношении конверсии гомофарнезола в (-)-амброксид. Полезный этап предварительного отбора может ограничиться отбором аминокислотных остатков, локализованных в домене 2 кристаллической структуры SHC/HAC (см. Фиг. 19 и 20).

ПРИМЕР 17

Получение гомофарнезола

Методы 17

Общие условия анализов

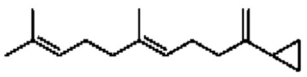
Неполярная ЖХ/МС: 50°C/2 мин, 20°C/мин 200°C, 35°C/мин 270°C. Система ЖХ/МС Agilent 5975C MSD с системой ЖХ HP серия 7890A. Неполярная колонка: BPX5 производства компании SGE, 5% фенил 95% диметилполисилоксан 0,22 мм × 0,25 мм × 12 м. Газ-носитель: Гелий. Температура инжектора: 230°C. Разделение потоков 1:50. Скорость потока: 1,0 мл/мин. Линия передачи: 250°C. МС-квадруполь: 106°C. МС-источник: 230°C.

A) Получение N-метил-N-нитрозомочевины (MNU) в THF

Раствор мочевины (175 г, 2,9 моль) и метиламина гидрохлорида (198 г, 2,9 моль) в воде (400 мл) нагревают с обратным холодильником (105°C) в течение 3,5 ч при перемешивании. При 40°C добавляют раствор $NaNO_2$ (101 г, 1,45 моль) в воде (200 мл). Через 15 мин добавляют THF (1000 мл), в результате чего получают 2-фазную прозрачную смесь. Концентрированную $H_2SO_4$ (110 г, 1,1 моль) добавляют при 0-5°C и перемешивают в течение 1,5 ч. Еще через 0,5 ч при 0-5°C разделяют две прозрачные фазы при 25°C. Органическую фазу (A) (1065 мл, теоретически 1,35 M) выдерживают в течение нескольких дней при 0-5°C или сразу переносят в реактор для циклопропанирования.

После разделения фаз водную фазу дважды экстрагируют THF (2x 1:1). В результате получают 1100 мл фазы B и 1075 фазы C. Если фаза A обеспечивает 51% конверсию концевого алкена в циклопропан при последующей реакции циклопропанирования, фаза B образует менее 0,5% циклопропана, а фаза C не дает обнаружимой конверсии. Авторы изобретения заключили, что после первого разделения фаз экстрагируется более 99% MNU. Поэтому после первого разделения фаз водную фазу обычно отбрасывают (из органической фазы A) после обработки концентрированным водным раствором KOH и уксусной кислотой.

B) Получение E-Δ-фарнезена с использованием MNU в THF



N-метил-N-нитрозомочевину 1,35 M в THF (136 мл, 184 ммоль) добавляют по каплям при 0°C к быстро перемешиваемой смеси E-бета-фарнезена (CAS 18794-84-8) (25 г, 122 ммоль) и водного раствора KOH (50 мл, 40%) при 0-5°C. После добавления 4 мл раствора MNU добавляют $Pd(acac)_2$ (7,4 мг, 0,024 ммоль, 0,02%), предварительно растворенный в 0,5 мл дихлорметана.

Остальной раствор MNU добавляют в течение 4 ч при 0-5°C. На этой стадии GC показала 28% не преобразованного E-β-фарнезена, 65% желаемого моноциклопропана (показано выше) и 3% бициклопропанированного соединения 5. После 16 ч при 25°C добавляют уксусную кислоту (100 мл) при 0-5°C с последующим добавлением трет-бутил-метилового эфира (250 мл). После разделения фаз органическую фазу промывают 2 M HCl (250 мл), и водную фазу экстрагируют трет-бутил-метиловым эфиром (250 мл).

Объединенные органические слои промывают водой (2× 100 мл), 10% водным раствором NaOH (2× 100 мл) и водой (2× 100 мл), высушивают над $MgSO_4$, фильтруют и концентрируют с получением 26,9 г светло-желтой жидкости, которая содержит 9% E-β-фарнезена, 82% желаемого соединения моноциклопропана и 6% бициклопропанированного побочного продукта.

Далее желаемое соединение можно выделить путем очистки дистилляцией. В результате добавления 1 г $K_2CO_3$ (1 г) и дистилляции на колонке из рулонной стали 30 см при 40-60 мбар получают 147 г соединения моноциклопропана (68%, скорректированный выход) при 135-145°C. Фракции объединяют с получением 92 г соединения моноциклопропана 100% чистоты.

Аналитические данные E-Δ фарнезена:

1H-NMR ($CDCl_3$, 400 МГц): 5.1 (2 т, 2 H), 4.6 (2 H), 2.2 (2 H), 2.1 (4 H), 2.0 (2 H), 1.7 (s, 3 H), 1.6 (2 s, 6 H), 1.3 (1 H), 0.6 (2 H), 0.45 (2 H) млн$^{-1}$. 13C-NMR ($CDCl_3$, 400 МГц): 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) млн$^{-1}$. ЖХ/МС: 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%). ИК (пленка): 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). Аналитически вычислено для $C_{16}H_{26}$: C, 88,00; H, 12,00. Обнаружено: C, 87,80; H, 12,01.

C) Получение (7E)-4,8,12-триметилтридека-3,7,11-триен-1-ола ((7E)-гомофарнез ола)

Смесь (E)-(6,10-диметилундека-1,5,9-триен-2-ил)циклопропана (E-Δ фарнезена) (1 г, 4,6 ммоль), додекана (0,2 г, 1,15 ммоль, внутренний стандарт) и L-(+)-винной кислоты (1 г, 6,9 ммоль) в трубке под давлением нагревают при перемешивании при 150°C. Через 18 ч и после полной конверсии (в соответствии с GC) смесь наливают на смесь воды (50 мл) и толуола (50 мл).

Фазы разделяют, и водную фазу экстрагируют толуолом (50 мл). Объединенные органические слои промывают концентрированным водным раствором $Na_2CO_3$ (50 мл) и концентрированным раствором NaCl (2× 50 мл), высушивают над $MgSO_4$, фильтруют и выпаривают при пониженном давлении с получением коричневатой смолы (1,35 г), которую смешивают с 30% водным раствором KOH (4,3 мл) и перемешивают при 25°C в течение 2 ч. С помощью анализа методом GC выявлено образование 96% (7E)-4,8,12-триметилтридека-3,7,11-триен-1-ола в соответствии с внутренним стандартом. Отношение E/Z 68:22. Данные анализа E-изомера согласуются с данными литературы, см., например, P. Kocienski, S. Wadman J. Org. Chem. 54, 1215 (1989).

Результаты 17

Данные показывают получение гомофарнезола, приемлемого для биоконверсии в (-)-амброксид.

Обсуждение 17

Способ получения гомофарнезола также подробно описан в двух одновременно поданных заявках на патент - PCT/EP2014/072882 (WO2015/059290) и PCT/EP2014/072891 (WO2015/059293) - полное содержание которых включено в настоящий документ посредством ссылки.

ПРИМЕР 18

Реакция, протекающая в одном реакционном сосуде

В данном эксперименте: Проводили (i) ферментацию штаммом E. coli,

продуцирующим вариант 215G2 SHC (например, как описано в примере 1) с последующей (ii) конверсией EEH непосредственно в ферментационной смеси. Поскольку 3 параметра [клетки], [EEH] и [SDS] (г/л) связаны между собой, было необходимо скорректировать параметры [EEH] и [SDS] в доступном объеме ферментационной смеси в зависимости от концентрации клеток (г/л), полученной в конце ферментации. Цель состояла в конверсии 125 г/л EEH с использованием 250 г/л клеток при концентрации SDS 1,55%. Для обеспечения надлежащей биоконверсии клетки должны находиться в состоянии покоя в статусе истощения источника глюкозы. Аэрацию отключали.

Методы 18

Ферментация:

Для обеспечения достаточно точного определения объема ферментационного бульона в реакторе в конце ферментации регистрировали все объемы отбираемых образцов, а также объемы всех добавлений в ферментер (подпитки, основания, кислоты и т.д.).

Определение концентрации клеток в ферментационной смеси:

Для определения сырой массы клеток (г/л) отбирали образец ферментационной смеси (5-10 мл) при постоянном перемешивании и помещали в центрифужную пробирку. Регистрировали массу образца. Образец центрифугировали в течение 10 мин при 17210g и 4°C (например, 12000 об/мин, ротор SS-34, центрифуга Sorvall RC3B). Супернатант осторожно отбирали пипеткой и регистрировали массу осадка. Концентрацию клеток на основании сырой массы определяют как $г_{клеток}/л_{смеси}$ или $г_{клеток}/г_{смеси}$.

Объем ферментационной смеси в ферментере определяли с учетом всех добавлений и отборов. В случае нахождения ферментера на весах массу ферментационной смеси определяли путем взвешивания, в ином случае 1 мл принимали за 1 г.

Определение необходимых количеств гомофарнезола и SDS:

В соответствии с определенной концентрацией клеток и объемом ферментационной смеси определяли количество E,E-гомофарнезола и SDS для добавления в реактор с целью сохранения одинакового соотношения между 3, как описано в примере биоконверсии примера 9: 125 г/л EEH, 250 г/л клеток, 1,55% SDS.

Осуществление биоконверсии:

1. Температуру устанавливали на 35°C. Аэрацию отключали.

2. В ферментационную смесь добавляли рассчитанное количество гомофарнезола.

3. Необходимое количество SDS осторожно добавляли из водного исходного 15,5% раствора SDS.

4. Реакционную смесь тщательно перемешивали в течение приблизительно 15 мин при 800 $млн^{-1}$.

5. Регистрировали pH реакционной смеси (внутренний электрод для измерения PH).

6. Образец (приблизительно 1 мл) отбирали в 15 мл пробирку Falcon. Добавляли приблизительно 5 мл деионизированной воды, и после тщательного перемешивания регистрировали pH на внешнем калиброванном электроде.

7. pH в реакторе пошагово устанавливали на 5,4 (значение, измеренное на внешнем калиброванном электроде) с использованием 85% $H_3PO_4$, при этом контролируя pH на внешнем электроде, как описано выше (6.).

8. В ходе биоконверсии регулировали pH, используя, например, 10 25% $H_3PO_4$ и 32% NaOH.

10. Отбор образцов реакционной смеси: приблизительно 1 мл реакционной смеси помещали в пробирку Falcon емкостью 15 мл. Добавляли приблизительно 5 мл MTBE. Образец экстрагировали при энергичном перемешивании. Аликвоту центрифугировали

в настольной центрифуге в течение 1 мин при максимальной скорости (пробирка Eppendorf). 100 л фазы растворителя добавляли в виалу для GC, содержащую 900 л MTBE. В течение первого дня биоконверсии образцы отбирали один раз в 1-1,5 часа. В последующие дни отбирали только 3 образца.

11. 1 л фазы растворителя анализировали на содержание амброксида и EEH, как описано в примере 8.

12. Конверсию EEH (%) рассчитывали как $100 \times (\text{амброксид}_{\text{площадь}}/(\text{амброксид}_{\text{площадь}} + \text{EEH}_{\text{площадь}}))$.

Результаты 18

Результаты демонстрируют, что ферментацию и конверсию EEH проводили в одном реакционном сосуде в реакторе KLF2000 (Bioengineering) в масштабе 1,9 литров. 251 г/л клеток обеспечивали конверсию 238 г EEH (251 г/л клеток) до не менее 93% за 47 часов. При измерении через 93 ч после начала конверсия составляла 99%.

Подобный эксперимент в одном реакционном сосуде проводили в реакторе Infors HT емкостью 0,75 л. После ферментации в соответствии со стандартным протоколом (пример 1) добавляли клетки из реактора, собранные из других экспериментов по ферментации, проведенных параллельно по такому же протоколу. Полученный в результате объем смеси составлял 479 г. Концентрация клеток была определена как 313,7 г/л, что составляло 1,25x концентрацию клеток при стандартной биоконверсии (250 г/л клеток). В реактор добавляли соответствующие количества EEH и SDS. Конверсия 75,1 г EEH (эквивалентно 157 г/л EEH в этом примере) составляла 98% менее чем за 90 ч. Этот результат демонстрирует возможность проведения ферментации и конверсии EEH в одном реакционном сосуде при концентрации не менее 125 г/л EEH, если при проведении ферментации используют клетки, продуцирующие вариант 215G2 SHC, при достаточно высокой плотности клеток.

Обсуждение 18

Преимущественно была получена 99% конверсия субстрата, что весьма полезно с коммерческой точки зрения при использовании дорогостоящего исходного материала (например, EEH).

ПРИМЕР 19

Повышение волюметрической продуктивности

Методы 19

Чтобы дополнительно повысить волюметрическую продуктивность, использовали 1,5x концентрированную биоконверсионную реакционную смесь, содержащую 375 г/л клеток, 188 г/л EEH, 2,33% SDS. Параллельно проводили стандартную биоконверсию с использованием 125 г/л EEH, 250 г/л клеток, 1,55% SDS (пример 7). 2 реакции проводили в реакторах Infors HT емкостью 0,750 л, все остальные параметры оставались неизменными.

Результаты 19

Результаты на Фиг. 22 демонстрируют, что процентная доля конверсии через 75 ч после начала составляла 88% при 1,5x биоконверсии по сравнению с 95% при стандартной биоконверсии. Процентная доля конверсии через 96 ч после начала составляла 93% конверсии EEH при 1,5x биоконверсии по сравнению с 97% при стандартной биоконверсии. Процентная доля конверсии при 1,5x биоконверсии составляла 96% от полученной при стандартной биоконверсии, было отмечено, что со временем перемешивание при 1,5x биоконверсии становилось более затруднительным по мере исчезновения маслянистого гомофарнезола, замещаемого твердыми продуктами реакции. Этим можно объяснить несколько сниженный уровень конверсии при 1,5x

биоконверсии. Использование реактора, оборудованного улучшенным перемешивающим устройством, может улучшить конверсию ЕЕН при 1,5х биоконверсии. Этот результат показывает, что можно проводить реакции биоконверсии при 188 г/л ЕЕН или более при условии достижения эффективного перемешивания, поскольку эффективность перемешивания, по-видимому, является единственным ограничением системы.

Производительность (-)-амброксида

«Производительность (-)-амброксида» относится к количеству извлекаемого (-)-амброксида в граммах на литр объема биотрансформации в час времени биоконверсии (т.е. времени после добавления субстрата). В связи с этим и со ссылкой на Фиг. 22 производительность (-)-амброксида рассчитывают, как описано ниже:

Биоконверсия 125 г/л ЕЕН (250 г/л клеток)

производительность через 1,25 ч: 10,3 грамма на литр в час

производительность через 8,25 ч: 6,3 грамма на литр в час

производительность через 21,25 ч: 4,1 грамма на литр в час

Биоконверсия 187,5 г/л ЕЕН (375 г/л клеток)

производительность через 1,25 ч: 12,2 грамма на литр в час

производительность через 8,25 ч: 8,2 грамма на литр в час

производительность через 21,25 ч: 5,5 грамма на литр в час

Можно считать, что производительность, рассчитанная через приблизительно 6-8 часов после начала, является репрезентативной для начальной скорости реакции, которая лучше всего описывает максимальный коэффициент конверсии системы.

При типичных биоконверсиях с использованием 125 г/л ЕЕН с 250 г/л клеток показана производительность амброксида от 6,3 до 8,5 граммов на литр в час через приблизительно 6 8 часов (репрезентативная для начальной скорости реакции).

ПРИМЕР 20

Замена реакционного буферного раствора раствором NaCl

Методы 20

Стандартную биоконверсию (125 г/л ЕЕН, 250 г/л клеток, 1,55% SDS) проводили, как описано в примере 7, но с заменой цитратного буферного раствора pH 5,4 на 0,5% или 0,9% NaCl, при этом все остальные параметры реакции не менялись. Параллельно в качестве контроля проводили биоконверсию в цитратном буферном растворе.

Результаты 20

Результаты на Фиг. 23 показывают, что коэффициент конверсии ЕЕН был одинаковым при проведении реакций в буферном растворе и в 0,9% растворе NaCl. Коэффициент конверсии был ниже при проведении реакции только в 0,5% NaCl. Этот результат показывает возможность проведения биоконверсии в отсутствие обеспечиваемой буфером точной регуляции pH и обеспечение достаточной ионной силы.

ПРИМЕР 21

Экстракция твердой фазы из реакционного бульона

С учетом того, что (-)-амброксид нерастворим в воде и не является жидким при температурах ниже приблизительно 75°C, эти свойства рассматривали как возможное преимущество для экстракции продукта из твердой фазы биотрансформации с использованием либо смешиваемых с водой (например, этанола), либо не смешиваемых с водой (например, толуола) растворителей.

Методы 21

200 мл реакционной смеси центрифугировали, чтобы отделить твердую фазу от

жидкой (водной) (Sorvall GS3, 5000 об/мин, 10 мин, 10°C). Таким образом отделяли приблизительно 80 мл твердого осадка от приблизительно 120 мл жидкой фазы. Анализ (газовая хроматография, пример 8) водной фазы после экстракции MTBE показал, что она содержит не более чем приблизительно 0,3% (-)-амброксида, исходно присутствующего в 200 мл реакционной смеси. Для экстракции амброксида из твердой фазы использовали толуол и 99% этанол.

Результаты 21

Экстракция толуолом:

80 мл твердой фазы экстрагировали 6 раз 45 мл толуола (приблизительно 1А объема твердой фазы, при энергичном перемешивании в течение 30 с и центрифугировании (Sorvall GS3, 5000 об/мин, 10 мин, 10°C). Проводили анализ методом GC для определения содержания (-)-амброксида в фазе растворителя. Более 99,5% (-)-амброксида, исходно присутствующего в реакционной смеси, экстрагировали в результате 6 экстракций, при которых общий объем толуола составляет 1,3 5х исходного объема реакционной смеси (200 мл) или 3,4х объема твердой фазы. На графике на Фиг. 24 показано изменение экстракции по мере проведения отмывок толуолом в виде % от количества (-)-амброксида, первоначально присутствовавшего в 200 мл всей реакционной смеси (за счет соотношения объема реакционная смесь/толуол % в первом экстракте превышает 100%).

Экстракция этанолом:

80 мл твердой фазы экстрагировали (Infors Multifors HT, 35°C, 1000 об/мин, 30 мин) приблизительно 160 мл (2 объема) 99% этанола с последующим центрифугированием. В ходе процедуры экстракции кристаллизация амброксида не происходила. На графике на Фиг. 25 показано, что после 4 отмывок (всего 640 мл EtOH, т.е. 3,2х начального объема всей реакционной смеси или 8х от объема твердой фазы) было извлечено приблизительно 99% (-)-амброксида, первоначально присутствующего в реакционной смеси. На первой стадии экстракции необходимо достаточное количество этанола, чтобы предотвратить кристаллизацию амброксида (растворимость в этаноле). При использовании только 1 или ½ объема твердой фазы на первой стадии экстракции получили липкую пасту, с которой было трудно обращаться, и (-)-амброксид кристаллизовался в виде игольчатых кристаллов на осадке после центрифугирования. Температура, по-видимому, не является фактором, ответственным за эту кристаллизацию (экстракцию и центрифугирование тестировали при комнатной температуре и приблизительно при 35°C-40°C).

За образование кристаллов, по-видимому, ответственна концентрация (-)-амброксида в фазе EtOH, а также соотношение EtOH/вода в жидкой фазе (остаточное содержание влаги в твердой фазе). Тем не менее было отмечено, что объем этанола можно уменьшить до 1 объема твердой фазы.

Примечание 21

Поскольку (-)-амброксид не находится в жидкой фазе при комнатной температуре, он выделяется вместе с биомассой и может быть экстрагирован органическим растворителем (например, смешиваемым с водой растворителем (например, этанолом) или не смешиваемым с водой растворителем (например, толуолом). Стадия центрифугирования, которая отделяет (-)-амброксид в твердой фазе реакционной смеси, обладает преимуществом, поскольку позволяет уменьшить количество растворителя, необходимое для экстракции (-)-амброксида.

ПРИМЕР 22

Сенсорный анализ

Цель: проведение сенсорного анализа (-)-амброксида и побочных продуктов (соединений II, III и IV), образующихся в «неочищенном» экстракте и в «кристаллизованном» экстракте.

Результат 22(a)

Результатом преобразования EEH является (-)-амброксид (соединение I) и изомер (-)-амброксида (соединение IV).

Результат 22(b)

Результатом биотрансформации EZH является макроциклический простой эфир (соединение II) и 9b-эпи-амброксид (соединение III).

Результат 22(c)

Неочищенная композиция (-)-амброксида содержит соединения I, II, III и IV, где % каждого присутствующего соединения составляет 87,1, 2,8, 2,5 и 7,6%, соответственно.

Результат 22(d)

Композиция материала после селективной кристаллизации (в лабораторном масштабе) содержит те же соединения, присутствующие в количестве 99,1, 0,1, 0,1 и 0,7%, соответственно.

Были получены следующие результаты сенсорного анализа:

(-)-Амброксид: OTH 0,2 нг/л (OTH представляет собой порог восприятия запаха).

Соединение IV из EEH: слабый, IsoE, древесный, GC-TH 5-10 нг.

Соединение II из EZH: «без запаха» (GC-TH более 500 нг) (GC-TH представляет собой порог обнаружения).

Соединение III из EZH: GC-TH приблизительно в 10 раз выше, чем для амброксида (около 2 нг).

Вывод

Общее процентное содержание каждого из 3 побочных продуктов (соединений II, III и IV) в «неочищенном» экстракте составляет приблизительно 3%.

Общее процентное содержание каждого из 3 побочных продуктов (соединений II, III и IV) в «кристаллизованном» экстракте составляет приблизительно 1% (лабораторный масштаб).

Сенсорный анализ 3 побочных продуктов (соединений II, III и IV) показал более слабый запах по сравнению с (-)-амброксидом.

В действительности запах 9b-эпи-амброксида (соединения III) приблизительно в 10 раз слабее, чем (-)-амброксида, что позволяет предположить, что он по существу не имеет запаха.

Как показывает сенсорный анализ, удаление из (-)-амброксида одного или более соединений, представляющих собой побочные продукты, может улучшить запах остаточного соединения (т.е. (-)-амброксида) даже в том случае, когда удаляемые соединения как таковые не имеют запаха.

Таким образом, усиление запаха амброксида наблюдали в отсутствие соединений II, III и IV.

ПРИМЕР 23

Извлечение амброксида экстракцией паром

Методы 23

Полученная в результате чистота неочищенного (экстрагированного паром) и кристаллизованного (-)-амброксида

Реакционную смесь реакции биотрансформации EE:EZ 86:14 экстрагировали паром, и продукт реакции кристаллизовали, как описано ниже. Погон от дистилляции паром собирали в виде двухфазной смеси. Органическую фазу сохраняли, а водную фазу

отбрасывали. Состав органической фазы анализировали методом GC, и результаты приведены в таблице 25 ниже (см. «неочищенный»). Впоследствии органическую фазу концентрировали до сухости. Затем к неочищенному высушенному продукту добавляли этанол, и смесь подогревали до растворения продукта. При комнатной температуре медленно добавляли воду, и (-)-амброксид кристаллизуется при периодическом перемешивании и охлаждении в ледяной бане.

Результаты 23

В таблице 25 ниже также показаны результаты анализа методом GC продуктов, полученных после этапа экстракции/дистилляции паром («неочищенный»), и кристаллизованного продукта ((-)-амброксида). В таблице 25 ссылки на EZH и EEH относятся к (3Z,7E) гомофарнезолу и 7E,3E гомофарнезолу соответственно.

В таблице 25 показано, что из конкретного исходного материала (EEH:EZH 86:14) при использовании WT SHC или производного SHC образуется желаемый конечный продукт (-)-амброксид и очень специфичная смесь побочных продуктов (II, IV и III). Данные исследования селективной кристаллизации показывают значительное обогащение (-)-амброксидом (I) практически при отсутствии побочных продуктов (II), (IV) или (III), обнаруживаемых в кристаллизованном образце. Соответственно, из данной смеси EE:EZ получают ольфакторно чистый продукт (-)-амброксид, который селективно кристаллизуется относительно несложным и экономичным путем.

**Таблица 25: показаны результаты анализа методом GC для кристаллизованного продукта.**

| | Площадь пика (GC) | | | | | | Амброксид |
|---|---|---|---|---|---|---|---|
| | (II) | (IV) | (III) | Амброксид | EZH | EEH | (%) |
| **Неочищенный** | 215073 | 190376 | 588769 | 6751605 | 13429 | 14184 | 86,9 |
| **Кристаллизованный** | 10088 | 8951 | 64625 | 9032941 | 0 | 0 | 99,1 |

Обсуждение 23

Экстракция паром/фильтрация являются экологически благоприятными способами выделения амброксида, поскольку обеспечивают удобное выделение амброксида без использования растворителей с сопутствующей инактивацией биокатализатора.

Резюме 23

(-)-Амброксид, полученный при использовании реакции биоконверсии, можно экстрагировать с использованием растворителя из всей реакционной смеси (например, используя не смешиваемый с водой растворитель, или путем экстракции/дистилляции паром, или путем фильтрования) или из твердой фазы (например, используя смешиваемый с водой растворитель) способами, известными специалистам в данной области техники.

SEQUENCE LISTING

<110> Givaudan SA
<120> Enzymes and Applications thereof
<130> 30578
<150> GB 1507207.7
<151> 2015-04-24
<160> 188
<170> PatentIn version 3.3
<210> 1

<211> 631
<212> PRT
<213> Alicyclobacillus acidocaldarius
<400> 1

Met Ala Glu Gln Leu Val Glu Ala Pro Ala Tyr Ala Arg Thr Leu Asp
1 5 10 15
Arg Ala Val Glu Tyr Leu Leu Ser Cys Gln Lys Asp Glu Gly Tyr Trp
20 25 30
Trp Gly Pro Leu Leu Ser Asn Val Thr Met Glu Ala Glu Tyr Val Leu
35 40 45
Leu Cys His Ile Leu Asp Arg Val Asp Arg Asp Arg Met Glu Lys Ile
50 55 60
Arg Arg Tyr Leu Leu His Glu Gln Arg Glu Asp Gly Thr Trp Ala Leu
65 70 75 80
Tyr Pro Gly Gly Pro Pro Asp Leu Asp Thr Thr Ile Glu Ala Tyr Val
85 90 95
Ala Leu Lys Tyr Ile Gly Met Ser Arg Asp Glu Glu Pro Met Gln Lys
100 105 110
Ala Leu Arg Phe Ile Gln Ser Gln Gly Gly Ile Glu Ser Ser Arg Val
115 120 125
Phe Thr Arg Met Trp Leu Ala Leu Val Gly Glu Tyr Pro Trp Glu Lys
130 135 140
Val Pro Met Val Pro Pro Glu Ile Met Phe Leu Gly Lys Arg Met Pro
145 150 155 160
Leu Asn Ile Tyr Glu Phe Gly Ser Trp Ala Arg Ala Thr Val Val Ala
165 170 175
Leu Ser Ile Val Met Ser Arg Gln Pro Val Phe Pro Leu Pro Glu Arg
180 185 190
Ala Arg Val Pro Glu Leu Tyr Glu Thr Asp Val Pro Pro Arg Arg Arg
195 200 205
Gly Ala Lys Gly Gly Gly Gly Trp Ile Phe Asp Ala Leu Asp Arg Ala
210 215 220
Leu His Gly Tyr Gln Lys Leu Ser Val His Pro Phe Arg Arg Ala Ala
225 230 235 240
Glu Ile Arg Ala Leu Asp Trp Leu Leu Glu Arg Gln Ala Gly Asp Gly
245 250 255
Ser Trp Gly Gly Ile Gln Pro Pro Trp Phe Tyr Ala Leu Ile Ala Leu
260 265 270
Lys Ile Leu Asp Met Thr Gln His Pro Ala Phe Ile Lys Gly Trp Glu
275 280 285
Gly Leu Glu Leu Tyr Gly Val Glu Leu Asp Tyr Gly Gly Trp Met Phe
290 295 300
Gln Ala Ser Ile Ser Pro Val Trp Asp Thr Gly Leu Ala Val Leu Ala
305 310 315 320
Leu Arg Ala Ala Gly Leu Pro Ala Asp His Asp Arg Leu Val Lys Ala
325 330 335
Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala
340 345 350

Val Lys Arg Pro Asn Leu Lys Pro Gly Gly Phe Ala Phe Gln Phe Asp
355 360 365
Asn Val Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Val Trp Ala
370 375 380
Leu Asn Thr Leu Arg Leu Pro Asp Glu Arg Arg Arg Arg Asp Ala Met
385 390 395 400
Thr Lys Gly Phe Arg Trp Ile Val Gly Met Gln Ser Ser Asn Gly Gly
405 410 415
Trp Gly Ala Tyr Asp Val Asp Asn Thr Ser Asp Leu Pro Asn His Ile
420 425 430
Pro Phe Cys Asp Phe Gly Glu Val Thr Asp Pro Pro Ser Glu Asp Val
435 440 445
Thr Ala His Val Leu Glu Cys Phe Gly Ser Phe Gly Tyr Asp Asp Ala
450 455 460
Trp Lys Val Ile Arg Arg Ala Val Glu Tyr Leu Lys Arg Glu Gln Lys
465 470 475 480
Pro Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Leu Tyr Gly
485 490 495
Thr Gly Ala Val Val Ser Ala Leu Lys Ala Val Gly Ile Asp Thr Arg
500 505 510
Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn
515 520 525
Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala
530 535 540
Tyr Ala Gly Lys Gly Ala Ser Thr Pro Ser Gln Thr Ala Trp Ala Leu
545 550 555 560
Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg
565 570 575
Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp
580 585 590
Glu Pro Tyr Tyr Thr Gly Thr Gly Phe Pro Gly Asp Phe Tyr Leu Gly
595 600 605
Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr
610 615 620
Lys Gln Ala Ile Glu Arg Arg
625 630

<210> 2
<211> 725
<212> PRT
<213> 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
465 470 475 480

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
705 710 715 720

Ile Asp Leu Asn Asn
725

<210> 3
<211> 658
<212> PRT
<213> 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
625 630 635 640
Ala Arg Tyr Arg Asn Leu Lys Lys Ala Asn Gln Pro Ile Val His Tyr
645 650 655
Gly Met

<210> 4
<211> 684
<212> PRT
<213> Bradyrhizobium japonicum
<400> 4

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

```
Leu Gly Phe Arg Asp Asp Glu Pro Ala Val Arg Arg Ala Arg Glu Trp
        115                 120                 125
Ile Glu Ala Lys Gly Gly Leu Arg Asn Ile Arg Val Phe Thr Arg Tyr
    130                 135                 140
Trp Leu Ala Leu Ile Gly Glu Trp Pro Trp Glu Lys Thr Pro Asn Ile
145                 150                 155                 160
Pro Pro Glu Val Ile Trp Phe Pro Leu Trp Phe Pro Phe Ser Ile Tyr
                165                 170                 175
Asn Phe Ala Gln Trp Ala Arg Ala Thr Leu Met Pro Ile Ala Val Leu
            180                 185                 190
Ser Ala Arg Arg Pro Ser Arg Pro Leu Pro Pro Glu Asn Arg Leu Asp
        195                 200                 205
Ala Leu Phe Pro His Gly Arg Lys Ala Phe Asp Tyr Glu Leu Pro Val
    210                 215                 220
Lys Ala Gly Ala Gly Gly Trp Asp Arg Phe Phe Arg Gly Ala Asp Lys
225                 230                 235                 240
Val Leu His Lys Leu Gln Asn Leu Gly Asn Arg Leu Asn Leu Gly Leu
                245                 250                 255
Phe Arg Pro Ala Ala Thr Ser Arg Val Leu Glu Trp Met Ile Arg His
            260                 265                 270
Gln Asp Phe Asp Gly Ala Trp Gly Gly Ile Gln Pro Pro Trp Ile Tyr
        275                 280                 285
Gly Leu Met Ala Leu Tyr Ala Glu Gly Tyr Pro Leu Asn His Pro Val
    290                 295                 300
Leu Ala Lys Gly Leu Asp Ala Leu Asn Asp Pro Gly Trp Arg Val Asp
305                 310                 315                 320
Val Gly Asp Ala Thr Tyr Ile Gln Ala Thr Asn Ser Pro Val Trp Asp
                325                 330                 335
Thr Ile Leu Thr Leu Leu Ala Phe Asp Asp Ala Gly Val Leu Gly Asp
            340                 345                 350
Tyr Pro Glu Ala Val Asp Lys Ala Val Asp Trp Val Leu Gln Arg Gln
        355                 360                 365
Val Arg Val Pro Gly Asp Trp Ser Met Lys Leu Pro His Val Lys Pro
    370                 375                 380
Gly Gly Trp Ala Phe Glu Tyr Ala Asn Asn Tyr Tyr Pro Asp Thr Asp
385                 390                 395                 400
Asp Thr Ala Val Ala Leu Ile Ala Leu Ala Pro Leu Arg His Asp Pro
                405                 410                 415
Lys Trp Lys Ala Lys Gly Ile Asp Glu Ala Ile Gln Leu Gly Val Asp
            420                 425                 430
Trp Leu Ile Gly Met Gln Ser Gln Gly Gly Gly Trp Gly Ala Phe Asp
        435                 440                 445
Lys Asp Asn Asn Gln Lys Ile Leu Thr Lys Ile Pro Phe Cys Asp Tyr
    450                 455                 460
Gly Glu Ala Leu Asp Pro Pro Ser Val Asp Val Thr Ala His Ile Ile
465                 470                 475                 480
Glu Ala Phe Gly Lys Leu Gly Ile Ser Arg Asn His Pro Ser Met Val
                485                 490                 495
```

Gln Ala Leu Asp Tyr Ile Arg Arg Glu Gln Glu Pro Ser Gly Pro Trp
500 505 510
Phe Gly Arg Trp Gly Val Asn Tyr Val Tyr Gly Thr Gly Ala Val Leu
515 520 525
Pro Ala Leu Ala Ala Ile Gly Glu Asp Met Thr Gln Pro Tyr Ile Gly
530 535 540
Arg Ala Cys Asp Trp Leu Val Ala His Gln Gln Ala Asp Gly Gly Trp
545 550 555 560
Gly Glu Ser Cys Ala Ser Tyr Met Asp Val Ser Ala Val Gly Arg Gly
565 570 575
Thr Thr Thr Ala Ser Gln Thr Ala Trp Ala Leu Met Ala Leu Leu Ala
580 585 590
Ala Asn Arg Pro Gln Asp Lys Asp Ala Ile Glu Arg Gly Cys Met Trp
595 600 605
Leu Val Glu Arg Gln Ser Ala Gly Thr Trp Asp Glu Pro Glu Phe Thr
610 615 620
Gly Thr Gly Phe Pro Gly Tyr Gly Val Gly Gln Thr Ile Lys Leu Asn
625 630 635 640
Asp Pro Ala Leu Ser Gln Arg Leu Met Gln Gly Pro Glu Leu Ser Arg
645 650 655
Ala Phe Met Leu Arg Tyr Gly Met Tyr Arg His Tyr Phe Pro Leu Met
660 665 670
Ala Leu Gly Arg Ala Leu Arg Pro Gln Ser His Ser
675 680

<210> 5
<211> 631
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<400> 5

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

130 135 140
Val Pro Met Val Pro Pro Glu Ile Met Phe Leu Gly Lys Arg Met Pro
145 150 155 160
Leu Asn Ile Tyr Glu Phe Gly Ser Trp Ala Arg Ala Thr Val Val Ala
165 170 175
Leu Ser Ile Val Met Ser Arg Gln Pro Val Phe Pro Leu Pro Glu Arg
180 185 190
Ala Arg Val Pro Glu Leu Tyr Glu Thr Asp Val Pro Pro Arg Arg Arg
195 200 205
Gly Ala Lys Gly Gly Gly Gly Trp Ile Phe Asp Ala Leu Asp Arg Ala
210 215 220
Leu His Gly Tyr Gln Lys Leu Ser Val His Pro Phe Arg Arg Ala Ala
225 230 235 240
Glu Ile Arg Ala Leu Asp Trp Leu Leu Glu Arg Gln Ala Gly Asp Gly
245 250 255
Ser Trp Gly Gly Ile Gln Pro Pro Trp Phe Tyr Ala Leu Ile Ala Leu
260 265 270
Lys Ile Leu Asp Met Thr Gln His Pro Ala Phe Ile Lys Gly Trp Glu
275 280 285
Gly Leu Glu Leu Tyr Gly Val Glu Leu Asp Tyr Gly Gly Trp Met Phe
290 295 300
Gln Ala Ser Ile Ser Pro Val Trp Asp Thr Gly Leu Ala Val Leu Ala
305 310 315 320
Leu Arg Ala Ala Gly Leu Pro Ala Asp His Asp Arg Leu Val Lys Ala
325 330 335
Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala
340 345 350
Val Lys Arg Pro Asn Leu Lys Pro Gly Gly Phe Ala Phe Gln Phe Asp
355 360 365
Asn Val Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Val Trp Ala
370 375 380
Leu Asn Thr Leu Arg Leu Pro Asp Glu Arg Arg Arg Arg Asp Ala Met
385 390 395 400
Thr Lys Gly Phe Arg Trp Ile Val Gly Met Gln Ser Ser Asn Gly Gly
405 410 415
Trp Gly Ala Tyr Asp Val Asp Asn Thr Ser Asp Leu Pro Asn His Ile
420 425 430
Pro Phe Cys Asp Phe Gly Glu Val Thr Asp Pro Pro Ser Glu Asp Val
435 440 445
Thr Ala His Val Leu Glu Cys Phe Gly Ser Phe Gly Tyr Asp Asp Ala
450 455 460
Trp Lys Val Ile Arg Arg Ala Val Glu Tyr Leu Lys Arg Glu Gln Lys
465 470 475 480
Pro Asp Gly Ser Trp Phe Gly Arg Trp Gly Val Asn Tyr Leu Tyr Gly
485 490 495
Thr Gly Ala Val Val Ser Ala Leu Lys Ala Val Gly Ile Asp Thr Arg
500 505 510
Glu Pro Tyr Ile Gln Lys Ala Leu Asp Trp Val Glu Gln His Gln Asn

515 520 525

Pro Asp Gly Gly Trp Gly Glu Asp Cys Arg Ser Tyr Glu Asp Pro Ala
530 535 540

Tyr Ala Gly Lys Gly Ala Ser Thr Pro Ser Gln Thr Ala Trp Ala Leu
545 550 555 560

Met Ala Leu Ile Ala Gly Gly Arg Ala Glu Ser Glu Ala Ala Arg Arg
565 570 575

Gly Val Gln Tyr Leu Val Glu Thr Gln Arg Pro Asp Gly Gly Trp Asp
580 585 590

Glu Pro Tyr Tyr Thr Gly Thr Gly Phe Pro Gly Asp Phe Tyr Leu Gly
595 600 605

Tyr Thr Met Tyr Arg His Val Phe Pro Thr Leu Ala Leu Gly Arg Tyr
610 615 620

Lys Gln Ala Ile Glu Arg Arg
625 630

<210> 6
<211> 1896
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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 cctgttgcac gagcagcgcg aggacggcgc gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggatgtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacatcccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560

gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 7
<211> 631
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<400> 7

Met Ala Glu Gln Leu Val Glu Ala Pro Ala Tyr Ala Arg Thr Leu Asp
1 5 10 15
Arg Ala Val Glu Tyr Leu Leu Ser Cys Gln Lys Asp Glu Gly Tyr Trp
20 25 30
Trp Gly Pro Leu Leu Ser Asn Val Thr Met Glu Ala Glu Tyr Val Leu
35 40 45
Leu Cys His Ile Leu Asp Arg Val Asp Arg Asp Arg Met Glu Lys Ile
50 55 60
Arg Arg Tyr Leu Leu His Glu Gln Arg Glu Asp Gly Thr Trp Ala Leu
65 70 75 80
Tyr Pro Gly Gly Pro Pro Asp Leu Asp 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
145 150 155 160
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> 8

<211> 1896

<212> DNA

<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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 cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accgtcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggatgtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacatcccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 9
<211> 631
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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> Artificial sequence

<220>

<223> Синтетическое производное AacSHC

<400> 10

atggctgagc agttggtgga agcgccggcc tacgcgcgga cgctggatcg cgcggtggag 60

tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120

acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180

atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240

tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300

atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360

ggcgggatcg agtcgtcgcg cgtgctcacg cggatgtggc tggcgctggt gggagaatat 420

ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480

ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540

atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600

accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660

ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720

gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780

attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840

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

<210> 11
<211> 631
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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

225 230 235 240

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> 12
<211> 1896
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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 cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggaggtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacatcccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 13
<211> 631
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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
145 150 155 160
Leu Asn Ile Tyr Glu Phe Gly Ser Trp Ala Arg Ala Thr Val Val Ala
165 170 175
Leu Ser Ile Val Met Ser Arg Gln Pro Val Phe Pro Leu Pro Glu Arg
180 185 190
Ala Arg Val Pro Glu Leu Tyr Glu Thr Asp Val Pro Pro Arg Arg Arg
195 200 205
Gly Ala Lys Gly Gly Gly Gly Trp Ile Phe Asp Ala Leu Asp Arg Val
210 215 220
Leu His Gly Tyr Gln Lys Leu Ser Val His Pro Phe Arg Arg Ala Ala
225 230 235 240
Glu Ile Arg Ala Leu Asp Trp Leu Leu Glu Arg Gln Ala Gly Asp Gly
245 250 255
Ser Trp Gly Gly Ile Gln Pro Pro Trp Phe Tyr Ala Leu Ile Ala Leu
260 265 270
Lys Ile Leu Asp Met Thr Gln His Pro Ala Phe Ile Lys Gly Trp Glu
275 280 285
Gly Leu Glu Leu Tyr Gly Val Glu Leu Asp Tyr Gly Gly Trp Met Phe
290 295 300
Gln Ala Ser Ile Ser Pro Val Trp Asp Thr Gly Leu Ala Val Leu Ala
305 310 315 320
Leu Arg Ala Ala Gly Leu Pro Ala Asp His Asp Arg Leu Val Lys Ala

325 330 335
Gly Glu Trp Leu Leu Asp Arg Gln Ile Thr Val Pro Gly Asp Trp Ala
340 345 350
Val Lys Arg Pro Asn Leu Lys Pro Gly Gly Phe Ala Phe Gln Phe Asp
355 360 365
Asn Val Tyr Tyr Pro Asp Val Asp Asp Thr Ala Val Val Val Trp Ala
370 375 380
Leu Asn Thr Leu Arg Leu Pro Asp Glu Arg Arg Arg Arg Asp Ala Met
385 390 395 400
Thr Lys Gly Phe Arg Trp Ile Val Gly Met Gln Ser Ser Asn Gly Gly
405 410 415
Trp Gly Ala Tyr Asp Val Asp Asn Thr Ser Asp Leu Pro Asn His 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> 14
<211> 1896
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<400> 14
atggctgagc agttggtgga agcgccggcc tacgcgcgga cgctggatcg cgcggtggag 60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120

acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggatgtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg tgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacatcccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 15
<211> 631
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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 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 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> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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 cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggatgtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200

acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacaccccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 17
<211> 631
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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
305 310 315 320
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 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 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> 18
<211> 1896
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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 cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggatgtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacatcccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcattac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 19
<211> 631
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC

<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> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<400> 20

atggctgagc agttggtgga agcgccggcc tacgcgcgga cgctggatcg cgcggtggag 60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggatgtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480

ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacatcccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
tacccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 21
<211> 631
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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 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> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<400> 22

atggctgagc agttggtgga agctccggcc tacgcgcgga cgctggatcg cgcggtggag 60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggaggtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg tgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacaccccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560

gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 23
<211> 631
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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 cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggaggtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacaccccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 25
<211> 631
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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
385 390 395 400
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> 26

<211> 1896

<212> DNA

<213> Artificial sequence

<220>

<223> Синтетическое производное AacSHC

<400> 26

atggctgagc agttggtgga agcgccggcc tacgcgcgga cgctggatcg cgcggtggag 60

tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120

acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180

atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240

tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300

atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360

ggcgggatcg agtcgtcgcg cgtgttcacg cggatgtggc tggcgctggt gggagaatat 420

ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480

ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540

atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600

accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660

ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720

gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780

attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840

ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacatcccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
tacccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 27
<211> 631
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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 cctgttgcac gagcagcgcg aggacggcgc gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accgtcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgctcacg cggatgtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacatcccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 29
<211> 631
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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
225 230 235 240
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 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> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<400> 30
atggctgagc agttggtgga agcaccggcc tacgcgcgga cgctggatcg cgcggtggag 60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120

acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggatgtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacatcccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcattac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
tacccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 31
<211> 631
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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
145 150 155 160
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> 32
<211> 1896
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<400> 32

atggctgagc agttggtgga agcgccggcc tacgcgcgga cgctggatcg cgcggtggag 60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgctcacg cggatgtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200

acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacatcccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 33
<211> 631
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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 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 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> Artificial sequence

<220>

<223> Синтетическое производное AacSHC

<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 cctgttgcac 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 aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacatcccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
tacccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 35

<211> 631

<212> PRT

<213> Artificial sequence

<220>

<223> Синтетическое производное AacSHC

<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
305 310 315 320
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> 36
<211> 1896
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<400> 36

atggctgagc agttggtgga agcgccggcc tacgcgcgga cgctggatcg cgcggtggag 60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgctcacg cggaggtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480

ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacaccccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 37
<211> 631
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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 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> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<400> 38

atggctgagc agttggtgga agcgccggcc tacgcgcgga cgctggatcg cgcggtggag 60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggaggtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacaccccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560

gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
tacccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 39
<211> 631
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<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
145 150 155 160
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 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> Artificial sequence
<220>
<223> Синтетическое производное AacSHC
<400> 40

atggctgagc agttggtgga agcgccggcc tacgcgcgga cgctggatcg cgcggtggag 60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgctcacg cggaggtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacaccccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
tacccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 41
<211> 725
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<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                  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 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 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> 42

<211> 2178

<212> DNA

<213> Artificial sequence

<220>

<223> Синтетическое производное ZmoSHC1

<220>

<221> misc_feature

<222> (387)..(387)

<223> n представляет собой a, c, g, или t
<400> 42
atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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
gccttgtggt ttttgggtct ggaagatcat ccgcttcgtc caagattggg caatgctctt 360
ttggaaatgc agcgggaaga tggcgcntgg ggagtctatt tcggcgctgg aaatggcgat 420
atcaatgcca cggttgaagc ctatgcggcc ttgcggtctt tggggtattc tgccgataat 480
cctgttttga aaaaagcggc agcatggatt gctgaaaaag gcggattaaa aaatatccgt 540
gtctttaccc gttattggct ggcgttgatc ggggaatggc cttgggaaaa gacccctaac 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 tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgttaa aagacgccaa 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 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 taattatatc 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gacccagcct 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 ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178
<210> 43
<211> 725
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<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
465 470 475 480
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
705 710 715 720
Ile Asp Leu Asn Asn
725

<210> 44
<211> 2178
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1

<400> 44

atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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
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 gacccctaac 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 tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgttaa aagacgccaa 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 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 taattatatc 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gacccagcct 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 ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

<210> 45

<211> 725

<212> PRT

<213> Artificial sequence

<220>

<223> Синтетическое производное ZmoSHC1

<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
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> 46
<211> 2178
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<220>

<221> misc_feature
<222> (544)..(546)
<223> ntn представляет собой ttr или ctn
<400> 46

atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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
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
gtcntnaccc gttattggct ggcgttgatc ggggaatggc cttgggaaaa gacccctaac 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 tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgttaa aagacgccaa 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 ccgtggggtt aattggttga tcgccatgca aagcgaatgt 1440
ggcggttggg gagcctttga taaggataat aacagaagta tcctttccaa aattcctttt 1500
tgtgatttcg gagaatctat tgatccgcct tcagtcgatg taacggcgca tgttttagag 1560
gcctttggca ccttgggact gtcccgcgat atgccggtca tccaaaaaagc gatcgactat 1620
gtccgttccg aacaggaagc cgaaggcgcg tggtttggtc gttggggcgt taattatatc 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gacccagcct 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 ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

<210> 47
<211> 725
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<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
625 630 635 640
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> 48
<211> 2178
<212> DNA
<213> Artificial sequence

<220>
<223> Синтетическое производное ZmoSHC1
<220>
<221> misc_feature
<222> (553)..(555)
<223> ngn представляет собой agr или cgn
<400> 48
atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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
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 gtngntggct ggcgttgatc ggggaatggc cttgggaaaa gacccctaac 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 tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgttaa aagacgccaa ggccgaggat cgttttacgc cggaaatgga taaggccgcc 1200
gattggcttt tggctcgaca ggtcaaagtc aaaggcgatt ggtcaatcaa actgcccgat 1260
gttgaacccg gtggatgggc atttgaatat gccaatgatc gctatcccga taccgatgat 1320
accgccgtcg ctttgatcgc ccttcctct 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 taattatatc 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gacccagcct 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 ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178
<210> 49
<211> 725
<212> PRT
<213> Artificial sequence

<220>
<223> Синтетическое производное ZmoSHC1
<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

385 390 395 400

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

705 710 715 720

Ile Asp Leu Asn Asn

725

<210> 50

<211> 2178
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<220>
<221> misc_feature
<222> (846)..(846)
<223> n представляет собой a, c, g, или t
<400> 50

atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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
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 gacccctaac 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
cgtgtnttac attgggttca gtccaatctg ttaaagcgca atagcttgcg tgaagccgct 900
atccgtcatg ttttggaatg gattatccgg catcaggatg ccgatggcgg ttggggtgga 960
attcagccac cttgggtcta tggtttgatg gcgttacatg gtgaaggcta tcagctttat 1020
catccggtga tggccaaggc tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgttaa aagacgccaa 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 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 taattatatc 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gacccagcct 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 ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

<210> 51

<211> 725
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<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                 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 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> 52
<211> 2178
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<220>
<221> misc_feature
<222> (1494)..(1494)
<223> n представляет собой a, c, g, или t
<400> 52

atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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
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 gacccctaac 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 tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgttaa aagacgccaa 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 ccgtggggtt 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 gacccagcct 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 ggcgctgaac tgtcacgggc gtttatgctg 2100

cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

<210> 53
<211> 725
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<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
545 550 555 560
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> 54
<211> 2178
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<400> 54

atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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
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 gacccctaac 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 tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgttaa aagacgccaa ggccgaggat cgttttacgc cggaaatgga taaggccgcc 1200
gattggcttt tggctcgaca ggtcaaagtc aaaggcgatt ggtcaatcaa actgcccgat 1260
gttgaacccg gtggatgggc atttgaatat gccaatgatc gctatcccga taccgatgat 1320
accgccgtcg ctttgatcgc ccttcctct 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 taattatatc 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gacccagcct 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 ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160

attgatttga ataattga 2178

<210> 55

<211> 725

<212> PRT

<213> Artificial sequence

<220>

<223> Синтетическое производное ZmoSHC1

<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> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<400> 56

atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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
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 gacccctaac 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 tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgttaa aagacgccaa 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 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 taattatatc 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gacccagcct 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 ataycccggt tatggcgtgg gtcagacgat caagttggat 2040
gatccggctt tatcgaaacg attgcttcaa ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

<210> 57
<211> 725
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<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
145 150 155 160
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 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
385 390 395 400
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> 58

<211> 2178

<212> DNA

<213> Artificial sequence

<220>

<223> Синтетическое производное ZmoSHC1

<220>

<221> misc_feature

<222> (553)..(555)

<223> ngn представляет собой agr или cgn

<220>

<221> misc_feature

<222> (846)..(846)

<223> n представляет собой a, c, g, или t

<220>

<221> misc_feature

<222> (1494)..(1494)

<223> n представляет собой a, c, g, или t

<400> 58

atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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

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 gtngntggct ggcgttgatc ggggaatggc cttgggaaaa gacccctaac 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

cgtgtnttac attgggttca gtccaatctg ttaaagcgca atagcttgcg tgaagccgct 900

atccgtcatg ttttggaatg gattatccgg catcaggatg ccgatggcgg ttggggtgga 960

attcagccac cttgggtcta tggtttgatg gcgttacatg gtgaaggcta tcagctttat 1020

catccggtga tggccaaggc tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080

ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140

atggcgttaa aagacgccaa 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 ccgtggggtt 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 gacccagcct 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 ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

<210> 59
<211> 725
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<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
465 470 475 480
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> 60
<211> 2178
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<220>
<221> misc_feature
<222> (553)..(555)
<223> ngn представляет собой agr или cgn
<220>
<221> misc_feature
<222> (1494)..(1494)
<223> n представляет собой a, c, g, или t
<400> 60

atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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
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 gtngntggct ggcgttgatc ggggaatggc cttgggaaaa gacccctaac 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 tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140

atggcgttaa aagacgccaa 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 ccgtggggtt 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 gacccagcct 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 ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

<210> 61
<211> 725
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<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> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<400> 62

atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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
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 gacccctaac 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 tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgttaa aagacgccaa ggccgaggat cgttttacgc cggaaatgga taaggccgcc 1200

gattggcttt tggctcgaca ggtcaaagtc aaaggcgatt ggtcaatcaa actgcccgat 1260
gttgaacccg gtggatgggc atttgaatat gccaatgatc gctatcccga taccgatgat 1320
accgccgtcg ctttgatcgc ccttttcctct 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 taattatatc 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gacccagcct 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 ataycccggt tatggcgtgg gtcagacgat caagttggat 2040
gatccggctt tatcgaaacg attgcttcaa ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

<210> 63
<211> 725
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<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                 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 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> 64
<211> 2178
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<220>
<221> misc_feature
<222> (387)..(387)
<223> n представляет собой a, c, g, или t
<220>
<221> misc_feature
<222> (544)..(546)
<223> ntn представляет собой ttr или ctn
<400> 64

atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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
gccttgtggt ttttgggtct ggaagatcat ccgcttcgtc caagattggg caatgctctt 360
ttggaaatgc agcgggaaga tggcgcntgg ggagtctatt tcggcgctgg aaatggcgat 420
atcaatgcca cggttgaagc ctatgcggcc ttgcggtctt tggggtattc tgccgataat 480
cctgttttga aaaaagcggc agcatggatt gctgaaaaag gcggattaaa aaatatccgt 540
gtcntnaccc gttattggct ggcgttgatc ggggaatggc cttgggaaaa gacccctaac 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 tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgttaa aagacgccaa 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 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 taattatatc 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gacccagcct 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 ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

<210> 65
<211> 725
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<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

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> 66

<211> 2178

<212> DNA

<213> Artificial sequence

<220>

<223> Синтетическое производное ZmoSHC1

<400> 66

atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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

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 gacccctaac 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 tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgttaa aagacgccaa 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 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 taattatatc 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gacccagcct 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 ataycccggt tatggcgtgg gtcagacgat caagttggat 2040
gatccggctt tatcgaaacg attgcttcaa ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

<210> 67
<211> 725
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<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                 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
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> 68
<211> 2178
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<220>
<221> misc_feature
<222> (544)..(546)
<223> ntn представляет собой ttr или ctn
<400> 68

atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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
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
gtcntnaccc gttattggct ggcgttgatc ggggaatggc cttgggaaaa gacccctaac 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 tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgttaa aagacgccaa 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 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 taattatatc 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gacccagcct 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 ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

<210> 69
<211> 725
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<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
225 230 235 240
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>  70
<211>  2178
<212>  DNA
<213>  Artificial sequence
<220>
<223>  Синтетическое производное ZmoSHC1
<220>
<221>  misc_feature
<222>  (544)..(546)
<223>  ntn представляет собой ttr или ctn
<400>  70
atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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
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
gtcntnaccc gttattggct ggcgttgatc ggggaatggc cttgggaaaa gacccctaac 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 tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgttaa aagacgccaa 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 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 taattatatc 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gacccagcct 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 ataycccggt tatggcgtgg gtcagacgat caagttggat 2040
gatccggctt tatcgaaacg attgcttcaa ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

<210> 71
<211> 725
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<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> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<220>
<221> misc_feature
<222> (544)..(546)
<223> ntn представляет собой ttr или ctn
<220>
<221> misc_feature
<222> (553)..(555)
<223> ngn представляет собой agr или cgn
<220>
<221> misc_feature

<222> (1494)..(1494)
<223> n представляет собой a, c, g, или t
<400> 72
atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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
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
gtcntnaccc gtngntggct ggcgttgatc ggggaatggc cttgggaaaa gacccctaac 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 tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgttaa aagacgccaa 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 ccgtggggtt 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 gacccagcct 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 ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178
<210> 73
<211> 725
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<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> Artificial sequence
<220>

<223> Синтетическое производное ZmoSHC1
<220>
<221> misc_feature
<222> (553)..(555)
<223> ngn представляет собой agr или cgn
<220>
<221> misc_feature
<222> (1494)..(1494)
<223> n представляет собой a, c, g, или t
<400> 74

atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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
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 gtngntggct ggcgttgatc ggggaatggc cttgggaaaa gacccctaac 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 tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgttaa aagacgccaa 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 ccgtggggtt aattggttga tcgccatgca aagcgaatgt 1440
ggcggttggg gagcctttga taaggataat aacagaagta tcctttccaa aacnccttt 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 gacccagcct 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 ataycccggt tatggcgtgg gtcagacgat caagttggat 2040
gatccggctt tatcgaaacg attgcttcaa ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

<210> 75

<211> 725
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<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                 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> 76
<211> 2178
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<220>
<221> misc_feature
<222> (544)..(546)
<223> ntn представляет собой ttr или ctn
<220>
<221> misc_feature
<222> (553)..(555)
<223> ngn представляет собой agr или cgn
<220>
<221> misc_feature
<222> (1494)..(1494)
<223> n представляет собой a, c, g, или t
<400> 76

atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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
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
gtcntnaccc gtngntggct ggcgttgatc ggggaatggc cttgggaaaa gacccctaac 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 tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgttaa aagacgccaa 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 ccgtggggtt 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 gacccagcct 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 ataycccggt tatggcgtgg gtcagacgat caagttggat 2040
gatccggctt tatcgaaacg attgcttcaa ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

<210> 77
<211> 658
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<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

625 630 635 640
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> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<220>
<221> misc_feature
<222> (255)..(255)
<223> n представляет собой a, c, g, или t
<400> 78

atgactgtat cgacttcctc ggcttttcat catagcccgt 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 gcgcntggtc tttgttttat ggtggtgatc ttgatttgag cgccacggtc 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 tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt 600
atcttggttg atgaattatt tgtgccggat gttttaccga cccttcagga aagcggtgac 660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat 720
tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt 780
ttgaatggtg aagacgggtt gggtgctatt tatccggcga ttgccaatag cgtcatgatg 840
tatgatgcct tgggatatcc cgaaaaccat 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 cggcctgatc tcagacccgg 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
ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat 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 cggtggtttt 1860
ccccgtgttt tttatcttcg gtatcacggt tattccaaat attttcctct ttgggcttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 79
<211> 658
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<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> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2

<400> 80

atgactgtat cgacttcctc ggcttttcat catagcccgt 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 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 tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt 600
atcttggttg atgaattatt tgtgccggat gttttaccga cccttcagga aagcggtgac 660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat 720
tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt 780
ttgaatggtg aagacgggtt gggtgctatt tatccggcga ttgccaatag cgtcatgatg 840
tatgatgcct tgggatatcc cgaaaaccat 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 cggcctgatc tcagacccgg 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
ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat 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 cggtggtttt 1860
ccccgtgttt tttatcttcg gtatcacggt tattccaaat attttcctct ttgggcttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 81
<211> 658
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2

<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
385 390 395 400
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> 82

<211> 1977

<212> DNA

<213> Artificial sequence

<220>

<223> Синтетическое производное ZmoSHC2

<220>

<221> misc_feature

<222> (409)..(411)

<223> ntn представляет собой ttr или ctn

<400> 82

atgactgtat cgacttcctc ggcttttcat catagcccgt 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 gattctcctg atgcgcctca tatgcttcga 360

gccagaaatg aaattttggc acggggtggg gcgatgcgtg ccaatgtcnt nacacgtatt 420

caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaatgcc cgtagagttg 480

atgttgatgc cggaatggtt 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 gaaggctatt catagctgtg tccattttgt gaccgagcgt 780
ttgaatggtg aagacgggtt gggtgctatt tatccggcga ttgccaatag cgtcatgatg 840
tatgatgcct tgggatatcc cgaaaaccat 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 cggcctgatc tcagacccgg 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
ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat 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 cggtggtttt 1860
ccccgtgttt tttatcttcg gtatcacggt tattccaaat attttcctct ttgggcttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 83
<211> 658
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<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
545 550 555 560

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> 84

<211> 1977

<212> DNA

<213> Artificial sequence

<220>

<223> Синтетическое производное ZmoSHC2

<220>

<221> misc_feature

<222> (418)..(420)

<223> ngn представляет собой agr или cgn

<400> 84

atgactgtat cgacttcctc ggcttttcat catagcccgt 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 gattctcctg atgcgcctca tatgcttcga 360
gccagaaatg aaattttggc acggggtggg gcgatgcgtg ccaatgtctt tacacgtngn 420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaatgcc cgtagagttg 480
atgttgatgc cggaatggtt 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 gaaggctatt catagctgtg tccattttgt gaccgagcgt 780
ttgaatggtg aagacgggtt gggtgctatt tatccggcga ttgccaatag cgtcatgatg 840
tatgatgcct tgggatatcc cgaaaaccat 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 cggcctgatc tcagacccgg 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
ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat 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 cggtggtttt 1860
ccccgtgttt tttatcttcg gtatcacggt tattccaaat attttcctct ttgggcttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 85
<211> 658
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<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
145 150 155 160
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> 86

<211> 1977

<212> DNA

<213> Artificial sequence

<220>

<223> Синтетическое производное ZmoSHC2

<400> 86

atgactgtat cgacttcctc ggcttttcat catagcccgt 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 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 tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt 600

atcttggttg atgaattatt tgtgccggat gttttaccga cccttcagga aagcggtgac 660

cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat 720

tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt 780

ttgaatggtg aagacgggtt gggtgctatt tatccggcga ttgccaatag cgtcatgatg 840

tatgatgcct tgggatatcc cgaaaaccat 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 cggcctgatc tcagacccgg 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

ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560

gcattgaatg tggccgcttt accccatgat 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 cggtggtttt 1860

ccccgtgttt tttatcttcg gtatcacggt tattccaaat attttcctct ttgggcttta 1920

gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 87
<211> 658
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<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
305 310 315 320
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> 88
<211> 1977
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2

<220>
<221> misc_feature
<222> (1350)..(1350)
<223> n представляет собой a, c, g, или t
<400> 88
atgactgtat cgacttcctc ggcttttcat catagcccgt 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 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 tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt 600
atcttggttg atgaattatt tgtgccggat gttttaccga cccttcagga aagcggtgac 660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat 720
tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt 780
ttgaatggtg aagacgggtt gggtgctatt tatccggcga ttgccaatag cgtcatgatg 840
tatgatgcct tgggatatcc cgaaaaccat 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 cggcctgatc tcagacccgg 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
ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggcttgg 1620
ctgaaaacta ttcaaaatga agatggtggt tggggtgaaa attgcgatag ctatgcccct 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> 89
<211> 658
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<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> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<400> 90

atgactgtat cgacttcctc ggcttttcat catagcccgt 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 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 tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt 600
atcttggttg atgaattatt tgtgccggat gttttaccga cccttcagga aagcggtgac 660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat 720
tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt 780
ttgaatggtg aagacgggtt gggtgctatt tatccggcga ttgccaatag cgtcatgatg 840
tatgatgcct tgggatatcc cgaaaaccat 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 cggcctgatc tcagacccgg 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
ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggcttgg 1620
ctgaaaacta ttcaaaatga agatggtggt 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
ccccgtgttt tttatcttcg gtatcacggt tattccaaat attttcctct ttgggcttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 91
<211> 658
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<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 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> 92

<211> 1977

<212> DNA

<213> Artificial sequence

<220>

<223> Синтетическое производное ZmoSHC2

<400> 92

atgactgtat cgacttcctc ggcttttcat catagcccgt 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 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 tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt 600

atcttggttg atgaattatt tgtgccggat gttttaccga cccttcagga aagcggtgac 660

cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat 720

tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt 780

ttgaatggtg aagacgggtt gggtgctatt tatccggcga ttgccaatag cgtcatgatg 840

tatgatgcct tgggatatcc cgaaaaccat 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 cggcctgatc tcagacccgg 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
ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat 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 cggtggttay 1860
ccccgtgttt tttatcttcg gtatcacggt tattccaaat attttcctct ttgggcttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 93
<211> 658
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<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
305 310 315 320
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> 94
<211> 1977
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<220>
<221> misc_feature
<222> (418)..(420)
<223> ngn представляет собой agr или cgn
<220>
<221> misc_feature
<222> (1350)..(1350)
<223> n представляет собой a, c, g, или t
<400> 94

atgactgtat cgacttcctc ggcttttcat catagcccgt 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 gattctcctg atgcgcctca tatgcttcga 360
gccagaaatg aaattttggc acggggtggg gcgatgcgtg ccaatgtctt tacacgtngn 420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaatgcc cgtagagttg 480
atgttgatgc cggaatggtt 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 gaaggctatt catagctgtg tccattttgt gaccgagcgt 780
ttgaatggtg aagacgggtt gggtgctatt tatccggcga ttgccaatag cgtcatgatg 840
tatgatgcct tgggatatcc cgaaaaccat 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 cggcctgatc tcagacccgg 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
ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat 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 cggtggtttt 1860
ccccgtgttt tttatcttcg gtatcacggt tattccaaat attttcctct ttgggcttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 95
<211> 658
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<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> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<220>
<221> misc_feature
<222> (418)..(420)
<223> ngn представляет собой agr или cgn
<220>
<221> misc_feature
<222> (1350)..(1350)
<223> n представляет собой a, c, g, или t
<400> 96

atgactgtat cgacttcctc ggcttttcat catagcccgt 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 gattctcctg atgcgcctca tatgcttcga 360
gccagaaatg aaattttggc acggggtggg gcgatgcgtg ccaatgtctt tacacgtngn 420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaatgcc cgtagagttg 480
atgttgatgc cggaatggtt 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 gaaggctatt catagctgtg tccattttgt gaccgagcgt 780
ttgaatggtg aagacgggtt gggtgctatt tatccggcga ttgccaatag cgtcatgatg 840
tatgatgcct tgggatatcc cgaaaaccat 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 cggcctgatc tcagacccgg 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
ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggcttgg 1620
ctgaaaacta ttcaaaatga agatggtggt tggggtgaaa attgcgatag ctatgcccct 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> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<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> Artificial sequence

<220>

<223> Синтетическое производное ZmoSHC2

<400> 98

atgactgtat cgacttcctc ggcttttcat catagcccgt 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 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 tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt 600
atcttggttg atgaattatt tgtgccggat gttttaccga cccttcagga aagcggtgac 660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat 720
tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt 780
ttgaatggtg aagacgggtt gggtgctatt tatccggcga ttgccaatag cgtcatgatg 840
tatgatgcct tgggatatcc cgaaaaccat 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 cggcctgatc tcagacccgg 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
ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat 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 cggtggttay 1860
ccccgtgttt tttatcttcg gtatcacggt tattccaaat attttcctct ttgggcttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 99
<211> 658
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<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
145 150 155 160
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> 100
<211> 1977
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<220>
<221> misc_feature
<222> (255)..(255)
<223> n представляет собой a, c, g, или t
<220>
<221> misc_feature
<222> (409)..(411)
<223> ntn представляет собой ttr или ctn
<400> 100

atgactgtat cgacttcctc ggcttttcat catagcccgt 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 gcgcntggtc tttgttttat ggtggtgatc ttgatttgag cgccacggtc 300
aaagcctatt ttgccttgaa aatgatcgga gattctcctg atgcgcctca tatgcttcga 360
gccagaaatg aaattttggc acggggtggg gcgatgcgtg ccaatgtcnt nacacgtatt 420

caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaatgcc cgtagagttg 480
atgttgatgc cggaatggtt 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 gaaggctatt catagctgtg tccattttgt gaccgagcgt 780
ttgaatggtg aagacgggtt gggtgctatt tatccggcga ttgccaatag cgtcatgatg 840
tatgatgcct tgggatatcc cgaaaaccat 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 cggcctgatc tcagacccgg 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
ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggcttgg 1620
ctgaaaacta ttcaaaatga agatggtggt tggggtgaaa attgcgatag ctatgcccтt 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> 101
<211> 658
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<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
305 310 315 320
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 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> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<400> 102

atgactgtat cgacttcctc ggcttttcat catagcccgt 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 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 tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt 600
atcttggttg atgaattatt tgtgccggat gttttaccga cccttcagga aagcggtgac 660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat 720
tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt 780
ttgaatggtg aagacgggtt gggtgctatt tatccggcga ttgccaatag cgtcatgatg 840
tatgatgcct tgggatatcc cgaaaaccat 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 cggcctgatc tcagacccgg 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
ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat catttagctg ttcagaaagc tgtggcttgg 1620
ctgaaaacta ttcaaaatga agatggtggt tggggtgaaa attgcgatag ctatgccctt 1680
gattatagcg gatacgagcc gatggattcg acggcttccc aaacagcatg ggctttattg 1740
ggcttgatgg ctgttgggga agctaattcc gaggccgtga caaagggtat acaytggttg 1800
gcacaaaatc aggatgaaga aggattgtgg aaagaagatt attatagtgg cggtggttay 1860
ccccgtgttt tttatcttcg gtatcacggt tattccaaat attttcctct ttgggcttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 103
<211> 658
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<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
545 550 555 560
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> 104
<211> 1977
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<220>
<221> misc_feature
<222> (409)..(411)
<223> ntn представляет собой ttr или ctn
<400> 104

atgactgtat cgacttcctc ggcttttcat catagcccgt 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 gattctcctg atgcgcctca tatgcttcga 360
gccagaaatg aaattttggc acggggtggg gcgatgcgtg ccaatgtcnt nacacgtatt 420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaatgcc cgtagagttg 480
atgttgatgc cggaatggtt 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 gaaggctatt catagctgtg tccattttgt gaccgagcgt 780
ttgaatggtg aagacgggtt gggtgctatt tatccggcga ttgccaatag cgtcatgatg 840
tatgatgcct tgggatatcc cgaaaaccat 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 cggcctgatc tcagacccgg 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
ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat 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 cggtggtttt 1860

ccccgtgttt tttatcttcg gtatcacggt tattccaaat attttcctct ttgggcttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 105
<211> 658
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<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 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
545 550 555 560
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> 106
<211> 1977
<212> DNA
<213> Artificial sequence

<220>
<223> Синтетическое производное ZmoSHC2
<220>
<221> misc_feature
<222> (409)..(411)
<223> ntn представляет собой ttr или ctn
<400> 106
atgactgtat cgacttcctc ggcttttcat catagcccgt 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 gattctcctg atgcgcctca tatgcttcga 360
gccagaaatg aaattttggc acggggtggg gcgatgcgtg ccaatgtcnt nacacgtatt 420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaatgcc cgtagagttg 480
atgttgatgc cggaatggtt 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 gaaggctatt catagctgtg tccattttgt gaccgagcgt 780
ttgaatggtg aagacgggtt gggtgctatt tatccggcga ttgccaatag cgtcatgatg 840
tatgatgcct tgggatatcc cgaaaaccat 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 cggcctgatc tcagacccgg 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
ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat 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 cggtggttay 1860
ccccgtgttt tttatcttcg gtatcacggt tattccaaat attttcctct ttgggcttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977
<210> 107
<211> 658
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<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
225 230 235 240
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> 108
<211> 1977
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<220>
<221> misc_feature
<222> (409)..(411)
<223> ntn представляет собой ttr или ctn
<220>
<221> misc_feature
<222> (418)..(420)
<223> ngn представляет собой agr или cgn

<220>
<221> misc_feature
<222> (1350)..(1350)
<223> n представляет собой a, c, g, или t
<400> 108
atgactgtat cgacttcctc ggcttttcat catagcccgt 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 gattctcctg atgcgcctca tatgcttcga 360
gccagaaatg aaattttggc acggggtggg gcgatgcgtg ccaatgtcnt nacacgtngn 420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaatgcc cgtagagttg 480
atgttgatgc cggaatggtt 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 gaaggctatt catagctgtg tccattttgt gaccgagcgt 780
ttgaatggtg aagacgggtt gggtgctatt tatccggcga ttgccaatag cgtcatgatg 840
tatgatgcct tgggatatcc cgaaaaccat 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 cggcctgatc tcagacccgg 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
ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat 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 cggtggtttt 1860
ccccgtgttt tttatcttcg gtatcacggt tattccaaat attttcctct ttgggcttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977
<210> 109
<211> 658
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<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                  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 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> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<220>
<221> misc_feature
<222> (418)..(420)
<223> ngn представляет собой agr или cgn
<220>
<221> misc_feature
<222> (1350)..(1350)
<223> n представляет собой a, c, g, или t
<400> 110
atgactgtat cgacttcctc ggcttttcat catagcccgt 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 gattctcctg atgcgcctca tatgcttcga 360
gccagaaatg aaattttggc acggggtggg gcgatgcgtg ccaatgtctt tacacgtngn 420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaatgcc cgtagagttg 480
atgttgatgc cggaatggtt 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 gaaggctatt catagctgtg tccattttgt gaccgagcgt 780
ttgaatggtg aagacgggtt gggtgctatt tatccggcga ttgccaatag cgtcatgatg 840
tatgatgcct tgggatatcc cgaaaaccat 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 cggcctgatc tcagacccgg 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
ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat 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 cggtggttay 1860
ccccgtgttt tttatcttcg gtatcacggt tattccaaat attttcctct ttgggcttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 111
<211> 658
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<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> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<220>
<221> misc_feature
<222> (409)..(411)
<223> ntn представляет собой ttr или ctn
<220>
<221> misc_feature
<222> (418)..(420)
<223> ngn представляет собой agr или cgn
<220>
<221> misc_feature
<222> (1350)..(1350)
<223> n представляет собой a, c, g, или t
<400> 112

atgactgtat cgacttcctc ggcttttcat catagcccgt 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 gattctcctg atgcgcctca tatgcttcga 360
gccagaaatg aaattttggc acggggtggg gcgatgcgtg ccaatgtcnt nacacgtngn 420
caattagctc tgttcggggc aatgtcatgg gagcatgtcc ctcaaatgcc cgtagagttg 480
atgttgatgc cggaatggtt 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 gaaggctatt catagctgtg tccattttgt gaccgagcgt 780
ttgaatggtg aagacgggtt gggtgctatt tatccggcga ttgccaatag cgtcatgatg 840
tatgatgcct tgggatatcc cgaaaaccat 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 cggcctgatc tcagacccgg 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
ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat 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 cggtggttay 1860
ccccgtgttt tttatcttcg gtatcacggt tattccaaat attttcctct ttgggcttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 113
<211> 684
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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
305 310 315 320
Val Gly Asp Ala Thr Tyr Ile Gln Ala Thr Asn Ser Pro Val Trp Asp
325 330 335
Thr Ile Leu Thr Leu Leu Ala Phe Asp Asp Ala Gly Val Leu Gly Asp
340 345 350
Tyr Pro Glu Ala Val Asp Lys Ala Val Asp Trp Val Leu Gln Arg Gln
355 360 365
Val Arg Val Pro Gly Asp Trp Ser Met Lys Leu Pro His Val Lys Pro
370 375 380
Gly Gly Trp Ala Phe Glu Tyr Ala Asn Asn Tyr Tyr Pro Asp Thr Asp
385 390 395 400
Asp Thr Ala Val Ala Leu Ile Ala Leu Ala Pro Leu Arg His Asp Pro
405 410 415
Lys Trp Lys Ala Lys Gly Ile Asp Glu Ala Ile Gln Leu Gly Val Asp
420 425 430
Trp Leu Ile Gly Met Gln Ser Gln Gly Gly Gly Trp Gly Ala Phe Asp
435 440 445
Lys Asp Asn Asn Gln Lys Ile Leu Thr Lys Ile Pro Phe Cys Asp Tyr
450 455 460
Gly Glu Ala Leu Asp Pro Pro Ser Val Asp Val Thr Ala His Ile Ile

465 470 475 480
Glu Ala Phe Gly Lys Leu Gly Ile Ser Arg Asn His Pro Ser Met Val
485 490 495
Gln Ala Leu Asp Tyr Ile Arg Arg Glu Gln Glu Pro Ser Gly Pro Trp
500 505 510
Phe Gly Arg Trp Gly Val Asn Tyr Val Tyr Gly Thr Gly Ala Val Leu
515 520 525
Pro Ala Leu Ala Ala Ile Gly Glu Asp Met Thr Gln Pro Tyr Ile Gly
530 535 540
Arg Ala Cys Asp Trp Leu Val Ala His Gln Gln Ala Asp Gly Gly Trp
545 550 555 560
Gly Glu Ser Cys Ala Ser Tyr Met Asp Val Ser Ala Val Gly Arg Gly
565 570 575
Thr Thr Thr Ala Ser Gln Thr Ala Trp Ala Leu Met Ala Leu Leu Ala
580 585 590
Ala Asn Arg Pro Gln Asp Lys Asp Ala Ile Glu Arg Gly Cys Met Trp
595 600 605
Leu Val Glu Arg Gln Ser Ala Gly Thr Trp Asp Glu Pro Glu Phe Thr
610 615 620
Gly Thr Gly Phe Pro Gly Tyr Gly Val Gly Gln Thr Ile Lys Leu Asn
625 630 635 640
Asp Pro Ala Leu Ser Gln Arg Leu Met Gln Gly Pro Glu Leu Ser Arg
645 650 655
Ala Phe Met Leu Arg Tyr Gly Met Tyr Arg His Tyr Phe Pro Leu Met
660 665 670
Ala Leu Gly Arg Ala Leu Arg Pro Gln Ser His Ser
675 680

<210> 114
<211> 2055
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<400> 114

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 gccaagggcg 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 cgcctggggc 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 gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac 1200
gacaccgcgg tcgcgctgat cgcgctggcg ccactgcgcc acgatccgaa atggaaggcc 1260
aaagggatcg acgaggctat ccagctcggt gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagatcccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggt 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 ctcgccgcca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttcccgggc tacggcgtcg gccagaccat caagctgaac 1920
gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc cttcatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

<210> 115
<211> 684
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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
145 150 155 160

Pro Pro Glu Val Ile Trp Phe Pro Leu Trp Phe Pro Phe Ser Ile Tyr
165 170 175
Asn Phe Ala Gln Trp Ala Arg Ala Thr Leu Met Pro Ile Ala Val Leu
180 185 190
Ser Ala Arg Arg Pro Ser Arg Pro Leu Pro Pro Glu Asn Arg Leu Asp
195 200 205
Ala Leu Phe Pro His Gly Arg Lys Ala Phe Asp Tyr Glu Leu Pro Val
210 215 220
Lys Ala Gly Ala Gly Gly Trp Asp Arg Phe Phe Arg Gly Ala Asp Lys
225 230 235 240
Val Leu His Lys Leu Gln Asn Leu Gly Asn Arg Leu Asn Leu Gly Leu
245 250 255
Phe Arg Pro Ala Ala Thr Ser Arg Val Leu Glu Trp Met Ile Arg His
260 265 270
Gln Asp Phe Asp Gly Ala Trp Gly Gly Ile Gln Pro Pro Trp Ile Tyr
275 280 285
Gly Leu Met Ala Leu Tyr Ala Glu Gly Tyr Pro Leu Asn His Pro Val
290 295 300
Leu Ala Lys Gly Leu Asp Ala Leu Asn Asp Pro Gly Trp Arg Val Asp
305 310 315 320
Val Gly Asp Ala Thr Tyr Ile Gln Ala Thr Asn Ser Pro Val Trp Asp
325 330 335
Thr Ile Leu Thr Leu Leu Ala Phe Asp Asp Ala Gly Val Leu Gly Asp
340 345 350
Tyr Pro Glu Ala Val Asp Lys Ala Val Asp Trp Val Leu Gln Arg Gln
355 360 365
Val Arg Val Pro Gly Asp Trp Ser Met Lys Leu Pro His Val Lys Pro
370 375 380
Gly Gly Trp Ala Phe Glu Tyr Ala Asn Asn Tyr Tyr Pro Asp Thr Asp
385 390 395 400
Asp Thr Ala Val Ala Leu Ile Ala Leu Ala Pro Leu Arg His Asp Pro
405 410 415
Lys Trp Lys Ala Lys Gly Ile Asp Glu Ala Ile Gln Leu Gly Val Asp
420 425 430
Trp Leu Ile Gly Met Gln Ser Gln Gly Gly Gly Trp Gly Ala Phe Asp
435 440 445
Lys Asp Asn Asn Gln Lys Ile Leu Thr Lys Ile Pro Phe Cys Asp Tyr
450 455 460
Gly Glu Ala Leu Asp Pro Pro Ser Val Asp Val Thr Ala His Ile Ile
465 470 475 480
Glu Ala Phe Gly Lys Leu Gly Ile Ser Arg Asn His Pro Ser Met Val
485 490 495
Gln Ala Leu Asp Tyr Ile Arg Arg Glu Gln Glu Pro Ser Gly Pro Trp
500 505 510
Phe Gly Arg Trp Gly Val Asn Tyr Val Tyr Gly Thr Gly Ala Val Leu
515 520 525
Pro Ala Leu Ala Ala Ile Gly Glu Asp Met Thr Gln Pro Tyr Ile Gly
530 535 540

Arg Ala Cys Asp Trp Leu Val Ala His Gln Gln Ala Asp Gly Gly Trp
545 550 555 560
Gly Glu Ser Cys Ala Ser Tyr Met Asp Val Ser Ala Val Gly Arg Gly
565 570 575
Thr Thr Thr Ala Ser Gln Thr Ala Trp Ala Leu Met Ala Leu Leu Ala
580 585 590
Ala Asn Arg Pro Gln Asp Lys Asp Ala Ile Glu Arg Gly Cys Met Trp
595 600 605
Leu Val Glu Arg Gln Ser Ala Gly Thr Trp Asp Glu Pro Glu Phe Thr
610 615 620
Gly Thr Gly Phe Pro Gly Tyr Gly Val Gly Gln Thr Ile Lys Leu Asn
625 630 635 640
Asp Pro Ala Leu Ser Gln Arg Leu Met Gln Gly Pro Glu Leu Ser Arg
645 650 655
Ala Phe Met Leu Arg Tyr Gly Met Tyr Arg His Tyr Phe Pro Leu Met
660 665 670
Ala Leu Gly Arg Ala Leu Arg Pro Gln Ser His Ser
675 680

<210> 116
<211> 2055
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<400> 116

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 gccaagggcg 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 cgcctggggc 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 gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac 1200
gacaccgcgg tcgcgctgat cgcgctggcg ccactgcgcc acgatccgaa atggaaggcc 1260
aaagggatcg acgaggctat ccagctcggt gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagatcccg 1380

ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggt 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 ctcgccgcca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttcccgggc tacggcgtcg gccagaccat caagctgaac 1920
gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc cttcatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

<210> 117
<211> 684
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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
305 310 315 320
Val Gly Asp Ala Thr Tyr Ile Gln Ala Thr Asn Ser Pro Val Trp Asp
325 330 335
Thr Ile Leu Thr Leu Leu Ala Phe Asp Asp Ala Gly Val Leu Gly Asp
340 345 350
Tyr Pro Glu Ala Val Asp Lys Ala Val Asp Trp Val Leu Gln Arg Gln
355 360 365
Val Arg Val Pro Gly Asp Trp Ser Met Lys Leu Pro His Val Lys Pro
370 375 380
Gly Gly Trp Ala Phe Glu Tyr Ala Asn Asn Tyr Tyr Pro Asp Thr Asp
385 390 395 400
Asp Thr Ala Val Ala Leu Ile Ala Leu Ala Pro Leu Arg His Asp Pro
405 410 415
Lys Trp Lys Ala Lys Gly Ile Asp Glu Ala Ile Gln Leu Gly Val Asp
420 425 430
Trp Leu Ile Gly Met Gln Ser Gln Gly Gly Gly Trp Gly Ala Phe Asp
435 440 445
Lys Asp Asn Asn Gln Lys Ile Leu Thr Lys Ile Pro Phe Cys Asp Tyr
450 455 460
Gly Glu Ala Leu Asp Pro Pro Ser Val Asp Val Thr Ala His Ile Ile
465 470 475 480
Glu Ala Phe Gly Lys Leu Gly Ile Ser Arg Asn His Pro Ser Met Val
485 490 495
Gln Ala Leu Asp Tyr Ile Arg Arg Glu Gln Glu Pro Ser Gly Pro Trp
500 505 510
Phe Gly Arg Trp Gly Val Asn Tyr Val Tyr Gly Thr Gly Ala Val Leu
515 520 525
Pro Ala Leu Ala Ala Ile Gly Glu Asp Met Thr Gln Pro Tyr Ile Gly
530 535 540
Arg Ala Cys Asp Trp Leu Val Ala His Gln Gln Ala Asp Gly Gly Trp
545 550 555 560
Gly Glu Ser Cys Ala Ser Tyr Met Asp Val Ser Ala Val Gly Arg Gly
565 570 575
Thr Thr Thr Ala Ser Gln Thr Ala Trp Ala Leu Met Ala Leu Leu Ala
580 585 590
Ala Asn Arg Pro Gln Asp Lys Asp Ala Ile Glu Arg Gly Cys Met Trp
595 600 605
Leu Val Glu Arg Gln Ser Ala Gly Thr Trp Asp Glu Pro Glu Phe Thr

610 615 620
Gly Thr Gly Phe Pro Gly Tyr Gly Val Gly Gln Thr Ile Lys Leu Asn
625 630 635 640
Asp Pro Ala Leu Ser Gln Arg Leu Met Gln Gly Pro Glu Leu Ser Arg
645 650 655
Ala Phe Met Leu Arg Tyr Gly Met Tyr Arg His Tyr Phe Pro Leu Met
660 665 670
Ala Leu Gly Arg Ala Leu Arg Pro Gln Ser His Ser
675 680

<210> 118
<211> 2055
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<220>
<221> misc_feature
<222> (421)..(423)
<223> ntn представляет собой ttr или 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 cgctcgctgg gcttccgcga cgacgagccg 360
gcggtgcgcc gggcgcggga atggatcgag gccaagggcg gcctgcgcaa catccgcgtc 420
ntnacccgct 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 cgcctggggc 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 gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac 1200
gacaccgcgg tcgcgctgat cgcgctggcg ccactgcgcc acgatccgaa atggaaggcc 1260
aaagggatcg acgaggctat ccagctcggt gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagatcccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggt 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 ctcgccgcca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttcccgggc tacggcgtcg gccagaccat caagctgaac 1920
gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc cttcatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

<210> 119
<211> 684
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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
145 150 155 160
Pro Pro Glu Val Ile Trp Phe Pro Leu Trp Phe Pro Phe Ser Ile Tyr
165 170 175
Asn Phe Ala Gln Trp Ala Arg Ala Thr Leu Met Pro Ile Ala Val Leu
180 185 190
Ser Ala Arg Arg Pro Ser Arg Pro Leu Pro Pro Glu Asn Arg Leu Asp
195 200 205
Ala Leu Phe Pro His Gly Arg Lys Ala Phe Asp Tyr Glu Leu Pro Val
210 215 220
Lys Ala Gly Ala Gly Gly Trp Asp Arg Phe Phe Arg Gly Ala Asp Lys
225 230 235 240
Val Leu His Lys Leu Gln Asn Leu Gly Asn Arg Leu Asn Leu Gly Leu
245 250 255
Phe Arg Pro Ala Ala Thr Ser Arg Val Leu Glu Trp Met Ile Arg His
260 265 270

Gln Asp Phe Asp Gly Ala Trp Gly Gly Ile Gln Pro Pro Trp Ile Tyr
275 280 285
Gly Leu Met Ala Leu Tyr Ala Glu Gly Tyr Pro Leu Asn His Pro Val
290 295 300
Leu Ala Lys Gly Leu Asp Ala Leu Asn Asp Pro Gly Trp Arg Val Asp
305 310 315 320
Val Gly Asp Ala Thr Tyr Ile Gln Ala Thr Asn Ser Pro Val Trp Asp
325 330 335
Thr Ile Leu Thr Leu Leu Ala Phe Asp Asp Ala Gly Val Leu Gly Asp
340 345 350
Tyr Pro Glu Ala Val Asp Lys Ala Val Asp Trp Val Leu Gln Arg Gln
355 360 365
Val Arg Val Pro Gly Asp Trp Ser Met Lys Leu Pro His Val Lys Pro
370 375 380
Gly Gly Trp Ala Phe Glu Tyr Ala Asn Asn Tyr Tyr Pro Asp Thr Asp
385 390 395 400
Asp Thr Ala Val Ala Leu Ile Ala Leu Ala Pro Leu Arg His Asp Pro
405 410 415
Lys Trp Lys Ala Lys Gly Ile Asp Glu Ala Ile Gln Leu Gly Val Asp
420 425 430
Trp Leu Ile Gly Met Gln Ser Gln Gly Gly Gly Trp Gly Ala Phe Asp
435 440 445
Lys Asp Asn Asn Gln Lys Ile Leu Thr Lys Ile Pro Phe Cys Asp Tyr
450 455 460
Gly Glu Ala Leu Asp Pro Pro Ser Val Asp Val Thr Ala His Ile Ile
465 470 475 480
Glu Ala Phe Gly Lys Leu Gly Ile Ser Arg Asn His Pro Ser Met Val
485 490 495
Gln Ala Leu Asp Tyr Ile Arg Arg Glu Gln Glu Pro Ser Gly Pro Trp
500 505 510
Phe Gly Arg Trp Gly Val Asn Tyr Val Tyr Gly Thr Gly Ala Val Leu
515 520 525
Pro Ala Leu Ala Ala Ile Gly Glu Asp Met Thr Gln Pro Tyr Ile Gly
530 535 540
Arg Ala Cys Asp Trp Leu Val Ala His Gln Gln Ala Asp Gly Gly Trp
545 550 555 560
Gly Glu Ser Cys Ala Ser Tyr Met Asp Val Ser Ala Val Gly Arg Gly
565 570 575
Thr Thr Thr Ala Ser Gln Thr Ala Trp Ala Leu Met Ala Leu Leu Ala
580 585 590
Ala Asn Arg Pro Gln Asp Lys Asp Ala Ile Glu Arg Gly Cys Met Trp
595 600 605
Leu Val Glu Arg Gln Ser Ala Gly Thr Trp Asp Glu Pro Glu Phe Thr
610 615 620
Gly Thr Gly Phe Pro Gly Tyr Gly Val Gly Gln Thr Ile Lys Leu Asn
625 630 635 640
Asp Pro Ala Leu Ser Gln Arg Leu Met Gln Gly Pro Glu Leu Ser Arg
645 650 655

Ala Phe Met Leu Arg Tyr Gly Met Tyr Arg His Tyr Phe Pro Leu Met
660 665 670

Ala Leu Gly Arg Ala Leu Arg Pro Gln Ser His Ser
675 680

<210> 120
<211> 2055
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<220>
<221> misc_feature
<222> (430)..(432)
<223> ngn представляет собой agr или 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 gtcgctgctc 240
gacagccagc gcccggacgg cgcctggcag gtctatttcg gcgcccccaa tggcgacatc 300
aacgcgactg tcgaggccta tgccgcgctc cgctcgctgg gcttccgcga cgacgagccg 360
gcggtgcgcc gggcgcggga atggatcgag gccaagggcg gcctgcgcaa catccgcgtc 420
ttcacccgcn 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 atggatgatc 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
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 ccagctcggt gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagatcccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggt 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 ctcgccgcca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttcccgggc tacggcgtcg gccagaccat caagctgaac 1920
gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc cttcatgctc 1980

cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

<210> 121
<211> 684
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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

305 310 315 320
Val Gly Asp Ala Thr Tyr Ile Gln Ala Thr Asn Ser Pro Val Trp Asp
325 330 335
Thr Ile Leu Thr Leu Leu Ala Phe Asp Asp Ala Gly Val Leu Gly Asp
340 345 350
Tyr Pro Glu Ala Val Asp Lys Ala Val Asp Trp Val Leu Gln Arg Gln
355 360 365
Val Arg Val Pro Gly Asp Trp Ser Met Lys Leu Pro His Val Lys Pro
370 375 380
Gly Gly Trp Ala Phe Glu Tyr Ala Asn Asn Tyr Tyr Pro Asp Thr Asp
385 390 395 400
Asp Thr Ala Val Ala Leu Ile Ala Leu Ala Pro Leu Arg His Asp Pro
405 410 415
Lys Trp Lys Ala Lys Gly Ile Asp Glu Ala Ile Gln Leu Gly Val Asp
420 425 430
Trp Leu Ile Gly Met Gln Ser Gln Gly Gly Gly Trp Gly Ala Phe Asp
435 440 445
Lys Asp Asn Asn Gln Lys Ile Leu Thr Lys Ile Pro Phe Cys Asp Tyr
450 455 460
Gly Glu Ala Leu Asp Pro Pro Ser Val Asp Val Thr Ala His Ile Ile
465 470 475 480
Glu Ala Phe Gly Lys Leu Gly Ile Ser Arg Asn His Pro Ser Met Val
485 490 495
Gln Ala Leu Asp Tyr Ile Arg Arg Glu Gln Glu Pro Ser Gly Pro Trp
500 505 510
Phe Gly Arg Trp Gly Val Asn Tyr Val Tyr Gly Thr Gly Ala Val Leu
515 520 525
Pro Ala Leu Ala Ala Ile Gly Glu Asp Met Thr Gln Pro Tyr Ile Gly
530 535 540
Arg Ala Cys Asp Trp Leu Val Ala His Gln Gln Ala Asp Gly Gly Trp
545 550 555 560
Gly Glu Ser Cys Ala Ser Tyr Met Asp Val Ser Ala Val Gly Arg Gly
565 570 575
Thr Thr Thr Ala Ser Gln Thr Ala Trp Ala Leu Met Ala Leu Leu Ala
580 585 590
Ala Asn Arg Pro Gln Asp Lys Asp Ala Ile Glu Arg Gly Cys Met Trp
595 600 605
Leu Val Glu Arg Gln Ser Ala Gly Thr Trp Asp Glu Pro Glu Phe Thr
610 615 620
Gly Thr Gly Phe Pro Gly Tyr Gly Val Gly Gln Thr Ile Lys Leu Asn
625 630 635 640
Asp Pro Ala Leu Ser Gln Arg Leu Met Gln Gly Pro Glu Leu Ser Arg
645 650 655
Ala Phe Met Leu Arg Tyr Gly Met Tyr Arg His Tyr Phe Pro Leu Met
660 665 670
Ala Leu Gly Arg Ala Leu Arg Pro Gln Ser His Ser
675 680
<210> 122

<211> 2055
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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 gtcgctgctc 240
gacagccagc gcccggacgg cgcctggcag gtctatttcg gcgcccccaa tggcgacatc 300
aacgcgactg tcgaggccta tgccgcgctc cgctcgctgg gcttccgcga cgacgagccg 360
gcggtgcgcc gggcgcggga atggatcgag gccaagggcg 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 cgcctggggc 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 gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac 1200
gacaccgcgg tcgcgctgat cgcgctggcg ccactgcgcc acgatccgaa atggaaggcc 1260
aaagggatcg acgaggctat ccagctcggt gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagatcccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggt 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 ctcgccgcca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttcccgggc tacggcgtcg gccagaccat caagctgaac 1920
gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc cttcatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055
<210> 123
<211> 684
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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                 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
625 630 635 640
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> 124
<211> 2055
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<220>
<221> misc_feature
<222> (1377)..(1377)
<223> n представляет собой a, c, g, или t

<400> 124

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 gccaagggcg 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 cgcctggggc 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 gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac 1200
gacaccgcgg tcgcgctgat cgcgctggcg ccactgcgcc acgatccgaa atggaaggcc 1260
aaagggatcg acgaggctat ccagctcggt gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagacnccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggt 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 ctcgccgcca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttcccgggc tacggcgtcg gccagaccat caagctgaac 1920
gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc cttcatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

<210> 125
<211> 684
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1

<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
625 630 635 640
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> 126
<211> 2055
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<400> 126

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 gccaagggcg 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 cgcctggggc 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 gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac 1200
gacaccgcgg tcgcgctgat cgcgctggcg ccactgcgcc acgatccgaa atggaaggcc 1260
aaagggatcg acgaggctat ccagctcggt gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagatcccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggt 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 ctcgccgcca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcca ytggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttcccgggc tacggcgtcg gccagaccat caagctgaac 1920
gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc cttcatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

<210> 127
<211> 684
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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                 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> 128
<211> 2055
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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 gtcgctgctc 240
gacagccagc gcccggacgg cgcctggcag gtctatttcg gcgcccccaa tggcgacatc 300
aacgcgactg tcgaggccta tgccgcgctc cgctcgctgg gcttccgcga cgacgagccg 360
gcggtgcgcc gggcgcggga atggatcgag gccaagggcg 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 cgcctggggc 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 gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac 1200
gacaccgcgg tcgcgctgat cgcgctggcg ccactgcgcc acgatccgaa atggaaggcc 1260
aaagggatcg acgaggctat ccagctcggt gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagatcccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggt 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 ctcgccgcca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg ttayccgggc tacggcgtcg gccagaccat caagctgaac 1920
gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc cttcatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

<210> 129
<211> 684
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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
305 310 315 320
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
625 630 635 640
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> 130
<211> 2055
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<220>
<221> misc_feature
<222> (430)..(432)
<223> ngn представляет собой agr или cgn
<220>
<221> misc_feature
<222> (1377)..(1377)
<223> n представляет собой a, c, g, или 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 gtcgctgctc 240
gacagccagc gcccggacgg cgcctggcag gtctatttcg gcgcccccaa tggcgacatc 300
aacgcgactg tcgaggccta tgccgcgctc cgctcgctgg gcttccgcga cgacgagccg 360
gcggtgcgcc gggcgcggga atggatcgag gccaagggcg gcctgcgcaa catccgcgtc 420
ttcacccgcn 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 atggatgatc 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
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 ccagctcggt gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagacnccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggt 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 ctcgccgcca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttcccgggc tacggcgtcg gccagaccat caagctgaac 1920
gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc cttcatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

<210> 131
<211> 684
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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 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                 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
625 630 635 640
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> 132
<211> 2055
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<220>
<221> misc_feature
<222> (430)..(432)
<223> ngn представляет собой agr или cgn
<220>
<221> misc_feature
<222> (1377)..(1377)
<223> n представляет собой a, c, g, или t
<400> 132

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 gccaagggcg gcctgcgcaa catccgcgtc 420
ttcacccgcn 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 atggatgatc 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
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 ccagctcggt gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagacnccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggt 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 ctcgccgcca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttcccgggc tacggcgtcg gccagaccat caagctgaac 1920
gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc cttcatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

<210> 133
<211> 684
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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 cgcctggcag gtctatttcg gcgcccccaa tggcgacatc 300
aacgcgactg tcgaggccta tgccgcgctc cgctcgctgg gcttccgcga cgacgagccg 360
gcggtgcgcc gggcgcggga atggatcgag gccaagggcg 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 cgcctggggc 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 gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac 1200
gacaccgcgg tcgcgctgat cgcgctggcg ccactgcgcc acgatccgaa atggaaggcc 1260
aaagggatcg acgaggctat ccagctcggt gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagatcccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggt 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 ctcgccgcca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg ttayccgggc tacggcgtcg gccagaccat caagctgaac 1920
gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc cttcatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

<210> 135
<211> 684
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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 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 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> 136
<211> 2055
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<220>
<221> misc_feature
<222> (421)..(423)
<223> ntn представляет собой ttr или 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 gtcgctgctc 240
gacagccagc gcccggacgg cgcctggcag gtctatttcg gcgcccccaa tggcgacatc 300
aacgcgactg tcgaggccta tgccgcgctc cgctcgctgg gcttccgcga cgacgagccg 360
gcggtgcgcc gggcgcggga atggatcgag gccaagggcg gcctgcgcaa catccgcgtc 420
ntnacccgct 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 cgcctggggc 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 gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac 1200
gacaccgcgg tcgcgctgat cgcgctggcg ccactgcgcc acgatccgaa atggaaggcc 1260
aaagggatcg acgaggctat ccagctcggt gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagatcccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggt 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 ctcgccgcca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttcccgggc tacggcgtcg gccagaccat caagctgaac 1920
gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc cttcatgctc 1980

cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

<210> 137
<211> 684
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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 gcttccgcga cgacgagccg 360
gcggtgcgcc gggcgcggga atggatcgag gccaagggcg 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 cgcctggggc 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 gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac 1200
gacaccgcgg tcgcgctgat cgcgctggcg ccactgcgcc acgatccgaa atggaaggcc 1260
aaagggatcg acgaggctat ccagctcggt gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagatcccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggt 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 ctcgccgcca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcca ytggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg ttayccgggc tacggcgtcg gccagaccat caagctgaac 1920
gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc cttcatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055
<210> 139
<211> 684
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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                  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 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> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<220>
<221> misc_feature
<222> (421)..(423)
<223> ntn представляет собой ttr или 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 cgcctggcag gtctatttcg gcgcccccaa tggcgacatc 300
aacgcgactg tcgaggccta tgccgcgctc cgctcgctgg gcttccgcga cgacgagccg 360
gcggtgcgcc gggcgcggga atggatcgag gccaagggcg gcctgcgcaa catccgcgtc 420
ntnacccgct 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 cgcctggggc 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 gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac 1200
gacaccgcgg tcgcgctgat cgcgctggcg ccactgcgcc acgatccgaa atggaaggcc 1260
aaagggatcg acgaggctat ccagctcggt gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagatcccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggt 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 ctcgccgcca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttcccgggc tacggcgtcg gccagaccat caagctgaac 1920
gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc cttcatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

<210> 141
<211> 684
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1

<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> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<220>
<221> misc_feature
<222> (421)..(423)
<223> ntn представляет собой ttr или 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 gccaagggcg gcctgcgcaa catccgcgtc 420
ntnacccgct 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 cgcctggggc 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 gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac 1200
gacaccgcgg tcgcgctgat cgcgctggcg ccactgcgcc acgatccgaa atggaaggcc 1260
aaagggatcg acgaggctat ccagctcggt gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagatcccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggt 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 ctcgccgcca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg ttayccgggc tacggcgtcg gccagaccat caagctgaac 1920
gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc cttcatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

<210> 143
<211> 684
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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
625 630 635 640
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> 144
<211> 2055
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<220>
<221> misc_feature
<222> (421)..(423)
<223> ntn представляет собой ttr или ctn
<220>
<221> misc_feature
<222> (430)..(432)
<223> ngn представляет собой agr или cgn
<220>
<221> misc_feature
<222> (1377)..(1377)
<223> n представляет собой a, c, g, или t
<400> 144
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 gcgccccaa tggcgacatc 300
aacgcgactg tcgaggccta tgccgcgctc cgctcgctgg gcttccgcga cgacgagccg 360
gcggtgcgcc gggcgcggga atggatcgag gccaagggcg gcctgcgcaa catccgcgtc 420
ntnacccgcn 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 atggatgatc 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
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 ccagctcggt gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagacnccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggt 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 ctcgccgcca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttcccgggc tacggcgtcg gccagaccat caagctgaac 1920
gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc cttcatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

<210> 145
<211> 684
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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 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
385 390 395 400
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> 146
<211> 2055
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<220>
<221> misc_feature
<222> (430)..(432)
<223> ngn представляет собой agr или cgn
<220>
<221> misc_feature
<222> (1377)..(1377)
<223> n представляет собой a, c, g, или t
<400> 146

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 gccaagggcg gcctgcgcaa catccgcgtc 420
ttcacccgcn 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 atggatgatc 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
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 ccagctcggt gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagacnccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggt 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 ctcgccgcca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg ttayccgggc tacggcgtcg gccagaccat caagctgaac 1920
gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc cttcatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

<210> 147
<211> 684
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<220>
<221> misc_feature
<222> (421)..(423)
<223> ntn представляет собой ttr или ctn
<220>
<221> misc_feature
<222> (430)..(432)
<223> ngn представляет собой agr или cgn
<220>
<221> misc_feature
<222> (1377)..(1377)
<223> n представляет собой a, c, g, или 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 gtcgctgctc 240
gacagccagc gcccggacgg cgcctggcag gtctatttcg gcgcccccaa tggcgacatc 300
aacgcgactg tcgaggccta tgccgcgctc cgctcgctgg gcttccgcga cgacgagccg 360
gcggtgcgcc gggcgcggga atggatcgag gccaagggcg gcctgcgcaa catccgcgtc 420
ntnacccgcn 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 atggatgatc 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
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 ccagctcggt gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagacnccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggt 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 ctcgccgcca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg ttayccgggc tacggcgtcg gccagaccat caagctgaac 1920
gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc cttcatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

<210> 149
<211> 657
<212> PRT
<213> 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                 320
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                 400
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
465 470 475 480

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> 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
145 150 155 160
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
225 230 235 240
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
305 310 315 320
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
385 390 395 400
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> 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
385 390 395 400
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
465 470 475 480
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
545 550 555 560
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> 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 tggtgtggcg 360
cgtgctcatt ttatgacgaa gtttttatta gcaattcatg gagaatatga atatccttca 420
ctctttcatt taccaacacc aatcatgttt ttacagaatg attccccctt tagtatattt 480
gaattaagta gctcagcacg tattcattta attccgatga tgctatgttt aaataaaaga 540
tttcgagtag ggaaaaagtt attaccaaat ttaaatcaca ttgcgggcgg aggcggagaa 600
tggtttcggg 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 aaaggggggca 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
gggatttgtt atctatatgg tacgtgggct gttatgactg gtttacggtc actcggaatt 1500
ccgtctagca atccttcatt gacacgagca gcttcatggc ttgaacatat acagcatgaa 1560
gatggtggtt ggggagaatc atgccacagt agtgtggaga aaaggttcgt tactttacca 1620

tttagtacac catcccaaac tgcatgggcg 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> 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
625 630 635 640
Gly Thr Thr Ile Arg Ala Leu Leu His Arg Arg Ser Pro Leu Ser Ser
645 650 655
Gly Arg Arg Ser Val Met Ile Glu Gly Cys Val Asp Val Val Ala Val
660 665 670
Arg Arg Phe Arg Cys Arg Ser Ser Arg Gly Ser Gly Ser Arg Arg Arg
675 680 685
Arg Ser Ser Trp Arg Cys Gly Gly Ile Cys Gly Ser Arg Cys Arg Thr

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 15

Gly Ala Leu Arg Glu Ser Val Gln Gln Ala Ala Asp Trp Leu Val Ala
20 25 30

Asn Gln Lys Pro Asp Gly His Trp Val Gly Arg Ala Glu Thr Asn Ala
35 40 45

Thr Met Glu Ala Gln Trp Cys Leu Ala Leu Trp Phe Leu Gly Leu Glu
50 55 60

Asp His Pro Leu Arg Val Arg Leu Gly Arg Ala Leu Leu Asp Thr Gln
65 70 75 80

Arg Pro Asp Gly Ala Trp His Val Phe Tyr Gly Ala Pro Asn Gly Asp
85 90 95

Ile Asn Ala Thr Val Glu Ala Tyr Ala Ala Leu Arg Ser Leu Gly His
100 105 110

Arg Asp Asp Glu Glu Pro Leu Arg Lys Ala Arg Asp Trp Ile Leu Ser
115 120 125

Lys Gly Gly Leu Ala Asn Ile Arg Val Phe Thr Arg Tyr Trp Leu Ala
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
210 215 220

Ala Gly Val Trp Asp Val Ile Phe Arg Lys Ile Asp Thr Ile Leu His
225 230 235 240

Arg Leu Gln Asp Trp Gly Ala Arg Arg Gly Pro His Gly Ile Met Arg
245 250 255

Arg Gly Ala Ile Asp His Val Leu Gln Trp Ile Ile Arg His Gln Asp
260 265 270

Tyr Asp Gly Ser Trp Gly Gly Ile Gln Pro Pro Trp Ile Tyr Gly Leu
275 280 285

Met Ala Leu His Thr Glu Gly Tyr Ala Met Thr His Pro Val Met Ala

290 295 300

Lys Ala Leu Asp Ala Leu Asn Glu Pro Gly Trp Arg Ile Asp Ile Gly
305 310 315 320
Asp Ala Thr Phe Ile Gln Ala Thr Asn Ser Pro Val Trp Asp Thr Met
325 330 335
Leu Ser Leu Leu Ala Phe Asp Asp Ala Gly Leu Gly Glu Arg Tyr Pro
340 345 350
Glu Gln Val Glu Arg Ala Val Arg Trp Val Leu Lys Arg Gln Val Leu
355 360 365
Val Pro Gly Asp Trp Ser Val Lys Leu Pro Asp Val Lys Pro Gly Gly
370 375 380
Trp Ala Phe Glu Tyr Ala Asn Asn Phe Tyr Pro Asp Thr Asp Asp Thr
385 390 395 400
Ser Val Ala Leu Met Ala Leu Ala Pro Phe Arg His Asp Pro Lys Trp
405 410 415
Gln Ala Glu Gly Ile Glu Asp Ala Ile Gln Arg Gly Ile Asp Trp Leu
420 425 430
Val Ala Met Gln Cys Lys Glu Gly Gly Trp Gly Ala Phe Asp Lys Asp
435 440 445
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
465 470 475 480
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
545 550 555 560
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
625 630 635 640
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> 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> 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                 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 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> Zymomonas mobilis
<400> 162

atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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
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 gacccctaac 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 tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgttaa aagacgccaa 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 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 taattatatc 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gacccagcct 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 ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

<210> 163
<211> 725
<212> PRT

<213> Zymomonas mobilis
<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                  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 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> Zymomonas mobilis
<400> 164
atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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
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 gacccctaac 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 tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgttaa aagacgccaa 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 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 taattatatc 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gacccagcct 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 ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178
<210> 165
<211> 1896
<212> DNA
<213> Alicyclobacillus acidocaldarius
<400> 165
atggctgagc agttggtgga agcgccggcc tacgcgcgga cgctggatcg cgcggtggag 60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180

atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggatgtggc tggcgctggt gggagaatat 420
ccatgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgtatgccg 480
ctcaacattt acgagtttgg ctcgtgggcc cgggcgaccg tcgtggcgat ctcaattgtc 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgat 600
accgacgtgc ctccgcgccg gcgcggcgcc aagggaggcg gcgggcgaat cttcgacgcg 660
ctggatcgcg ccctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccggg ccttggactg gctgctcgag cgccaggccg gagacggcag ttggggcggg 780
attcagccgc cctggtttta tacgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggttg ggagggcctc gagctgtacg gagtggacct cgactacggc 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggtcttgc cgtgctcgcg 960
ctgcgcgccg cggggcttcc ggccgatcac gaccggttgg tcaaggcggg cgagtggctt 1020
ttggaccggc agatcaccgt gccgggagac tgggcggtga agcgcccgaa cctcaaaccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tattacccgg acgtcgacga cacggccgtc 1140
gtggtctggg cgctgaacag ccttcgcttg ccggacgagc gccgcaggcg ggacgtgatg 1200
acgaaggggt tccgctggat cgtcggtatg cagagttcca acggcggctg gggcgcgtac 1260
gacgtcgaca acacgagcga tctgccaaac cacatcccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcggagga 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 caagggcgcg agcaccccgt cgcagacggc ttgggcgctg 1680
atggcactca tcgcgggcgg cagggcggag tcggattccg tgcgccgcgg cgtgcaatat 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccgggcg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 166
<211> 2178
<212> DNA
<213> Zymomonas mobilis
<400> 166

atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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
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 gacccctaac 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 tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgttaa aagacgccaa 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 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 taattatatc 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gacccagcct 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 ggcgctgaac tgtcacgggc gtttatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

<210> 167
<211> 1977
<212> DNA
<213> Zymomonas mobilis
<400> 167

atgactgtat cgacttcctc ggcttttcat catagcccgt 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 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 tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt 600
atcttggttg atgaattatt tgtgccggat gttttaccga cccttcagga aagcggtgac 660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat 720
tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt 780
ttgaatggtg aagacgggtt gggtgctatt tatccggcga ttgccaatag cgtcatgatg 840
tatgatgcct tgggatatcc cgaaaaccat 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 cggcctgatc tcagacccgg 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
ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat 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 cggtggtttt 1860
ccccgtgttt tttatcttcg gtatcacggt tattccaaat attttcctct ttgggcttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 168
<211> 2055
<212> DNA
<213> Bradyrhizobium japonicum
<400> 168

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 gccaagggcg 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 cgcctggggc 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 gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac 1200
gacaccgcgg tcgcgctgat cgcgctggcg ccactgcgcc acgatccgaa atggaaggcc 1260
aaagggatcg acgaggctat ccagctcggt gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagatcccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggt 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 ctcgccgcca atcgccccca ggacaaggac 1800

gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttcccgggc tacggcgtcg gccagaccat caagctgaac 1920
gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc cttcatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

<210> 169
<211> 1896
<212> DNA
<213> Alicyclobacillus acidocaldarius

<400> 169
atggctgagc agttggtgga agcgccggcc tacgcgcgga cgctggatcg cgcggtggag 60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggatgtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacatcccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg atttctacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 170
<211> 1896
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное AacSHC

<400> 170

atggctgagc agttggtgga agcgccggcc tacgcgcgga cgctggatcg cgcggtggag 60
tatctcctct cctgccaaaa ggacgaaggc tactggtggg ggccgcttct gagcaacgtc 120
acgatggaag cggagtacgt cctcttgtgc cacattctcg atcgcgtcga tcgggatcgc 180
atggagaaga tccggcggta cctgttgcac gagcagcgcg aggacggcac gtgggccctg 240
tacccgggtg ggccgccgga cctcgacacg accatcgagg cgtacgtcgc gctcaagtat 300
atcggcatgt cgcgcgacga ggagccgatg cagaaggcgc tccggttcat tcagagccag 360
ggcgggatcg agtcgtcgcg cgtgttcacg cggatgtggc tggcgctggt gggagaatat 420
ccgtgggaga aggtgcccat ggtcccgccg gagatcatgt tcctcggcaa gcgcatgccg 480
ctcaacatct acgagtttgg ctcgtgggct cgggcgaccg tcgtggcgct ctcgattgtg 540
atgagccgcc agccggtgtt cccgctgccc gagcgggcgc gcgtgcccga gctgtacgag 600
accgacgtgc ctccgcgccg gcgcggtgcc aagggagggg gtgggtggat cttcgacgcg 660
ctcgaccggg cgctgcacgg gtatcagaag ctgtcggtgc acccgttccg ccgcgcggcc 720
gagatccgcg ccttggactg gttgctcgag cgccaggccg gagacggcag ctggggcggg 780
attcagccgc cttggtttta cgcgctcatc gcgctcaaga ttctcgacat gacgcagcat 840
ccggcgttca tcaagggctg ggaaggtcta gagctgtacg gcgtggagct ggattacgga 900
ggatggatgt ttcaggcttc catctcgccg gtgtgggaca cgggcctcgc cgtgctcgcg 960
ctgcgcgctg cggggcttcc ggccgatcac gaccgcttgg tcaaggcggg cgagtggctg 1020
ttggaccggc agatcacggt tccgggcgac tgggcggtga agcgcccgaa cctcaagccg 1080
ggcgggttcg cgttccagtt cgacaacgtg tactacccgg acgtggacga cacggccgtc 1140
gtggtgtggg cgctcaacac cctgcgcttg ccggacgagc gccgcaggcg ggacgccatg 1200
acgaagggat tccgctggat tgtcggcatg cagagctcga acggcggttg gggcgcctac 1260
gacgtcgaca acacgagcga tctcccgaac cacatcccgt tctgcgactt cggcgaagtg 1320
accgatccgc cgtcagagga cgtcaccgcc cacgtgctcg agtgtttcgg cagcttcggg 1380
tacgatgacg cctggaaggt catccggcgc gcggtggaat atctcaagcg ggagcagaag 1440
ccggacggca gctggttcgg tcgttggggc gtcaattacc tctacggcac gggcgcggtg 1500
gtgtcggcgc tgaaggcggt cgggatcgac acgcgcgagc cgtacattca aaaggcgctc 1560
gactgggtcg agcagcatca gaacccggac ggcggctggg gcgaggactg ccgctcgtac 1620
gaggatccgg cgtacgcggg taagggcgcg agcaccccgt cgcagacggc ctgggcgctg 1680
atggcgctca tcgcgggcgg cagggcggag tccgaggccg cgcgccgcgg cgtgcaatac 1740
ctcgtggaga cgcagcgccc ggacggcggc tgggatgagc cgtactacac cggcacgggc 1800
ttcccagggg attggtacct cggctacacc atgtaccgcc acgtgtttcc gacgctcgcg 1860
ctcggccgct acaagcaagc catcgagcgc aggtga 1896

<210> 171

<211> 631

<212> PRT

<213> Artificial sequence

<220>

<223> Синтетическое производное AacSHC

<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 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> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<400> 172

atgggtattg acagaatgaa tagcttaagt cgcttgttaa 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
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 gacccctaac 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 tttgtcggct ttggatgatc ccggttggcg acatgacaga 1080
ggcgagtctt cttggataca ggccaccaat agtccggtat gggatacaat gttggccttg 1140
atggcgttaa aagacgccaa 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 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 taattatatc 1680
tatggcaccg gtgcggttct gcctgctttg gcggcgatcg gtgaagatat gacccagcct 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 ggcgctgaac tgtcacgggc gtggatgctg 2100
cgttatgatt tttatcggca attcttcccg attatggcgt taagtcgggc agagagactg 2160
attgatttga ataattga 2178

<210> 173
<211> 725
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC1
<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 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 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
705 710 715 720
Ile Asp Leu Asn Asn
725

<210> 174
<211> 1977
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<400> 174

atgactgtat cgacttcctc ggcttttcat catagcccgt 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 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 tttacaggcg ttaaagcctg tcgcccgtaa tcggcgcggt 600
atcttggttg atgaattatt tgtgccggat gttttaccga cccttcagga aagcggtgac 660
cctatatggc gtcgtttttt ttcggcactt gataaggtat tgcataaagt agaaccttat 720
tggccgaaaa atatgcgcgc gaaggctatt catagctgtg tccattttgt gaccgagcgt 780
ttgaatggtg aagacgggtt gggtgctatt tatccggcga ttgccaatag cgtcatgatg 840
tatgatgcct tgggatatcc cgaaaaccat 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 cggcctgatc tcagacccgg 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
ggtagctggt tcgggcgttg gggtgtcaat tacatatatg gcacatggtc ggccttatgt 1560
gcattgaatg tggccgcttt accccatgat 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 cggtggtttt 1860
ccccgtgttt ggtatcttcg gtatcacggt tattccaaat attttcctct ttgggcttta 1920
gcgcgctatc gcaatttgaa aaaagccaat cagccgattg ttcattatgg gatgtaa 1977

<210> 175
<211> 658
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное ZmoSHC2
<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                 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 Trp
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> 176
<211> 2055
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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 cgcctggcag gtctatttcg gcgcccccaa tggcgacatc 300
aacgcgactg tcgaggccta tgccgcgctc cgctcgctgg gcttccgcga cgacgagccg 360
gcggtgcgcc gggcgcggga atggatcgag gccaagggcg 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 cgcctggggc 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 gcaggtgcgc gtgcccggcg actggtcgat gaagctgccg 1140
catgtcaagc ccggcggctg ggcgttcgaa tacgccaaca actactatcc cgacacggac 1200
gacaccgcgg tcgcgctgat cgcgctggcg ccactgcgcc acgatccgaa atggaaggcc 1260
aaagggatcg acgaggctat ccagctcggt gtcgactggc tgatcggcat gcagagccag 1320
ggcggcggct ggggcgcgtt cgacaaggac aacaaccaga agatcctgac caagatcccg 1380
ttctgcgatt atggcgaggc gctcgatccg ccctcggtcg acgtcaccgc ccacatcatc 1440
gaggcgttcg gcaagctcgg catctcgcgc aaccatccgt cgatggtgca ggcgctggac 1500
tatattcgcc gtgagcagga gccgagcggt 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 ctcgccgcca atcgccccca ggacaaggac 1800
gcgatcgagc gtggctgcat gtggctggtc gagcgccagt cggccggcac ctgggacgag 1860
ccggaattca ccggcaccgg tttcccgggc tacggcgtcg gccagaccat caagctgaac 1920

gatcccgcgc tgtcgcagcg gctgatgcag ggcccggaat tgtcccgcgc ctggatgctc 1980
cgctacggca tgtaccgcca ctacttcccg ctgatggcgc tcggccgcgc cctacgcccg 2040
cagagtcata gctag 2055

<210> 177
<211> 684
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное BjapSHC1
<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 80
Asp Ser Gln Arg Pro Asp Gly Ala Trp Gln Val Tyr Phe Gly Ala Pro
85 90 95
Asn Gly Asp Ile Asn Ala Thr Val Glu Ala Tyr Ala Ala Leu Arg Ser
100 105 110
Leu Gly Phe Arg Asp Asp Glu Pro Ala Val Arg Arg Ala Arg Glu Trp
115 120 125
Ile Glu Ala Lys Gly Gly Leu Arg Asn Ile Arg Val Phe Thr Arg Tyr
130 135 140
Trp Leu Ala Leu Ile Gly Glu Trp Pro Trp Glu Lys Thr Pro Asn Ile
145 150 155 160
Pro Pro Glu Val Ile Trp Phe Pro Leu Trp Phe Pro Phe Ser Ile Tyr
165 170 175
Asn Phe Ala Gln Trp Ala Arg Ala Thr Leu Met Pro Ile Ala Val Leu
180 185 190
Ser Ala Arg Arg Pro Ser Arg Pro Leu Pro Pro Glu Asn Arg Leu Asp
195 200 205
Ala Leu Phe Pro His Gly Arg Lys Ala Phe Asp Tyr Glu Leu Pro Val
210 215 220
Lys Ala Gly Ala Gly Gly Trp Asp Arg Phe Phe Arg Gly Ala Asp Lys
225 230 235 240
Val Leu His Lys Leu Gln Asn Leu Gly Asn Arg Leu Asn Leu Gly Leu
245 250 255
Phe Arg Pro Ala Ala Thr Ser Arg Val Leu Glu Trp Met Ile Arg His
260 265 270
Gln Asp Phe Asp Gly Ala Trp Gly Gly Ile Gln Pro Pro Trp Ile Tyr
275 280 285
Gly Leu Met Ala Leu Tyr Ala Glu Gly Tyr Pro Leu Asn His Pro Val
290 295 300

Leu Ala Lys Gly Leu Asp Ala Leu Asn Asp Pro Gly Trp Arg Val Asp
305 310 315 320
Val Gly Asp Ala Thr Tyr Ile Gln Ala Thr Asn Ser Pro Val Trp Asp
325 330 335
Thr Ile Leu Thr Leu Leu Ala Phe Asp Asp Ala Gly Val Leu Gly Asp
340 345 350
Tyr Pro Glu Ala Val Asp Lys Ala Val Asp Trp Val Leu Gln Arg Gln
355 360 365
Val Arg Val Pro Gly Asp Trp Ser Met Lys Leu Pro His Val Lys Pro
370 375 380
Gly Gly Trp Ala Phe Glu Tyr Ala Asn Asn Tyr Tyr Pro Asp Thr Asp
385 390 395 400
Asp Thr Ala Val Ala Leu Ile Ala Leu Ala Pro Leu Arg His Asp Pro
405 410 415
Lys Trp Lys Ala Lys Gly Ile Asp Glu Ala Ile Gln Leu Gly Val Asp
420 425 430
Trp Leu Ile Gly Met Gln Ser Gln Gly Gly Gly Trp Gly Ala Phe Asp
435 440 445
Lys Asp Asn Asn Gln Lys Ile Leu Thr Lys Ile Pro Phe Cys Asp Tyr
450 455 460
Gly Glu Ala Leu Asp Pro Pro Ser Val Asp Val Thr Ala His Ile Ile
465 470 475 480
Glu Ala Phe Gly Lys Leu Gly Ile Ser Arg Asn His Pro Ser Met Val
485 490 495
Gln Ala Leu Asp Tyr Ile Arg Arg Glu Gln Glu Pro Ser Gly Pro Trp
500 505 510
Phe Gly Arg Trp Gly Val Asn Tyr Val Tyr Gly Thr Gly Ala Val Leu
515 520 525
Pro Ala Leu Ala Ala Ile Gly Glu Asp Met Thr Gln Pro Tyr Ile Gly
530 535 540
Arg Ala Cys Asp Trp Leu Val Ala His Gln Gln Ala Asp Gly Gly Trp
545 550 555 560
Gly Glu Ser Cys Ala Ser Tyr Met Asp Val Ser Ala Val Gly Arg Gly
565 570 575
Thr Thr Thr Ala Ser Gln Thr Ala Trp Ala Leu Met Ala Leu Leu Ala
580 585 590
Ala Asn Arg Pro Gln Asp Lys Asp Ala Ile Glu Arg Gly Cys Met Trp
595 600 605
Leu Val Glu Arg Gln Ser Ala Gly Thr Trp Asp Glu Pro Glu Phe Thr
610 615 620
Gly Thr Gly Phe Pro Gly Tyr Gly Val Gly Gln Thr Ile Lys Leu Asn
625 630 635 640
Asp Pro Ala Leu Ser Gln Arg Leu Met Gln Gly Pro Glu Leu Ser Arg
645 650 655
Ala 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> Alicyclobacillus acidocaldarius
<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
305 310 315 320
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 Ser Leu Arg Leu Pro Asp Glu Arg Arg Arg Arg Asp Val 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 Arg
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 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
545 550 555 560
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
580 585 590
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> 179
<211> 382
<212> DNA
<213> Artificial sequence
<220>
<223> Синтетическая область клонирования/экспрессии плазмиды pET-28a(+)
<220>
<221> CDS
<222> (109)..(285)
<400> 179
agatctcgat cccgcgaaat taatacgact cactataggg gaattgtgag cggataacaa 60

ttcccctcta 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 15
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 30 35
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 50
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
ttttttg 382

<210> 180
<211> 58
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическая конструкция
<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> Artificial sequence
<220>
<223> Синтетическая область клонирования/экспрессии плазмиды pET-28b(+)
<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> Artificial sequence
<220>
<223> Синтетическая конструкция
<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> Artificial sequence
<220>
<223> Синтетическая область клонирования/экспрессии плазмиды pET-28c(+)
<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> Artificial sequence
<220>
<223> Синтетическая конструкция
<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> Artificial sequence
<220>
<223> Синтетическое производное SHC
<220>
<221> VARIANT
<222> (77)..(77)
<223> T77X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P,

Q, R, S, T, V, W или Y
<220>
<221> VARIANT
<222> (92)..(92)
<223> I92X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y
<220>
<221> VARIANT
<222> (129)..(129)
<223> F129X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.
<220>
<221> VARIANT
<222> (132)..(132)
<223> M132X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.
<220>
<221> VARIANT
<222> (224)..(224)
<223> A224X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.
<220>
<221> VARIANT
<222> (432)..(432)
<223> I432X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.
<220>
<221> VARIANT
<222> (579)..(579)
<223> Q579X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.
<220>
<221> VARIANT
<222> (601)..(601)
<223> F601X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или 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 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 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

465 470 475 480
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 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> Artificial sequence
<220>
<223> Синтетическое производное SHC
<220>
<221> VARIANT
<222> (129)..(129)
<223> S129X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.
<220>
<221> VARIANT
<222> (145)..(145)
<223> V145X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.
<220>
<221> VARIANT
<222> (182)..(182)
<223> F182X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.
<220>
<221> VARIANT
<222> (185)..(185)
<223> Y185X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.
<220>

<221> VARIANT
<222> (282)..(282)
<223> G282X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.
<220>
<221> VARIANT
<222> (498)..(498)
<223> I498X имеет X, выбранный из: A, B, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.
<220>
<221> VARIANT
<222> (646)..(646)
<223> H646X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.
<220>
<221> VARIANT
<222> (668)..(668)
<223> F668X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или 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
625 630 635 640
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
705 710 715 720
Ile Asp Leu Asn Asn
725

<210> 187
<211> 658
<212> PRT
<213> Artificial sequence
<220>
<223> Синтетическое производное SHC
<220>
<221> VARIANT
<222> (85)..(85)
<223> G85X имеет X, выбранный из: A, C, D, E, F, G, H, I, K , L, M, N, P, Q, R, S, T, V, W или Y.
<220>
<221> VARIANT
<222> (100)..(100)
<223> V100X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.
<220>
<221> VARIANT
<222> (137)..(137)
<223> F137X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.
<220>
<221> VARIANT
<222> (140)..(140)
<223> I140X имеет X, выбранный из: A,, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.
<220>
<221> VARIANT
<222> (233)..(233)
<223> V233X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.
<220>

<221> VARIANT
<222> (450)..(450)
<223> I450X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.
<220>
<221> VARIANT
<222> (598)..(598)
<223> N598X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y
<220>
<221> VARIANT
<222> (620)..(620)
<223> F620X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или 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
465 470 475 480
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 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
625 630 635 640
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> Artificial sequence

<220>

<223> Синтетическое производное SHC

<220>

<221> VARIANT

<222> (88)..(88)

<223> A88X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

<220>

<221> VARIANT

<222> (104)..(104)

<223> V104X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

<220>

<221> VARIANT

<222> (141)..(141)

<223> F141X, имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

<220>

<221> VARIANT

<222> (144)..(144)

<223> Y144X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

<220>

<221> VARIANT

<222> (241)..(241)

<223> V241X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

<220>

<221> VARIANT

<222> (459)..(459)

<223> I459X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

<220>

<221> VARIANT

<222> (607)..(607)

<223> M607X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или Y.

<220>

<221> VARIANT

<222> (628)..(628)

<223> F628X имеет X, выбранный из: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W или 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
305 310 315 320

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 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                 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 Xaa 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 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
```

## (57) Формула изобретения

1. Способ получения (3aR,5aS,9aS,9bR)-3a,6,6,9a-тетраметилдодекагидронафто[2,1-b]фурана ((-)-амброксида) или смеси изомеров, содержащей (-)-амброксид, где смесь изомеров, содержащую (3E,7E)-гомофарнезол (EEH), ферментативно конвертируют в (-)-амброксид или смесь изомеров, содержащую (-)-амброксид, где ферментативную конверсию осуществляют с использованием фермента SHC/HAC (сквален-гопен-циклаза/гомофарнезол-амброксид-циклаза), обладающего активностью гомофарнезол-амброксид-циклазы и содержащего аминокислотную последовательность, имеющую

по меньшей мере 30% идентичность с SEQ ID NO: 1, в реакционных условиях, подходящих для получения (-)-амброксида, где смесь изомеров, содержащая EEH, выбрана из по меньшей мере одной из [(3E,7E) и (3Z,7E)], [(3E,7E) и (3E,7Z)] или [(3Z,7E), (3E,7E) и (3E,7Z)], также обозначаемых как [EE:EZ], [EE:ZE] и [EE:EZ:ZE] соответственно, и где (-)-амброксид присутствует в твердой фазе в виде кристаллов или в аморфной форме.

2. Способ по п. 1, где используют рекомбинантные целые клетки, продуцирующие фермент SHC/HAC.

3. Способ по п. 2, где используют додецилсульфат натрия (SDS) с рекомбинантными клетками-хозяевами, продуцирующими фермент SHC/HAC.

4. Способ по п. 3, где соотношение SDS/клетки находится в диапазоне от 10:1 до 20:1.

5. Способ по п. 1, где конверсию гомофарнезола в (-)-амброксид проводят при температуре в диапазоне от 30 до 60°C и при pH в диапазоне от приблизительно 4 до приблизительно 8.

6. Способ по п. 5, где конверсию гомофарнезола в (-)-амброксид проводят при температуре от приблизительно 34°C до приблизительно 50°C и при pH в диапазоне от приблизительно 5 до приблизительно 6,2.

7. Способ по п. 1, где субстрат гомофарнезол содержит изомеры EE:EZ.

8. Способ по п. 1, где гомофарнезол содержит смесь изомеров EE:EZ в массовых соотношениях, выбранных из группы, состоящей из: 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 и/или 60:40.

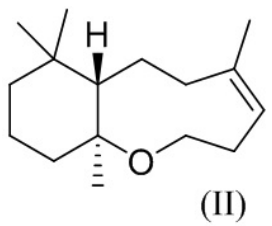
9. Способ по п. 1, где гомофарнезол содержит смесь изомеров EE:EZ в массовом соотношении, выбранном из группы, состоящей из: EE:EZ 92:08; EE:EZ 90:10; EE:EZ 80:20; EE:EZ 86:14; EE:EZ 70:30; EE:EZ 69:31; и/или EE:EZ 66:34.

10. Способ по п. 9, где гомофарнезол содержит EE:EZ в массовом соотношении 80:20.

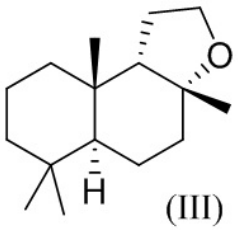
11. Способ по п. 1, где (-)-амброксид выделяют из смеси с использованием стадии экстракции/дистилляции органическим растворителем или паром либо фильтрования.

12. Способ по п. 11, где (-)-амброксид селективно кристаллизуют с использованием органического растворителя.

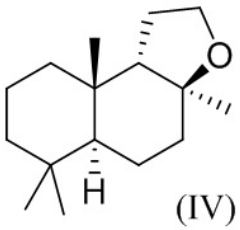
13. Способ по п. 12, где (-)-амброксид по существу не содержит побочных продуктов (II), (IV) и/или (III)



(II)



(III)



(IV)

14. Реакционная смесь биоконверсии для получения (-)-амброксида, содержащая водную фазу, твердую фазу и масляную фазу, получаемая способом по любому из пп. 1-10, где твердая фаза содержит продукт реакции (-)-амброксид в твердой форме.

15. Выделенный продукт реакции для ароматизации, содержащий (-)-амброксид в твердой форме, получаемый способом по любому из пп. 1-13, где указанный продукт реакции содержит 80-97% (-)-амброксида.

16. Продукт реакции по п. 14 или 15, где (-)-амброксид находится в аморфной форме.

17. Продукт реакции по п. 14 или 15, где (-)-амброксид находится в кристаллической форме.

18. Реакционная смесь или продукт реакции по любому из пп. 14-17, где продукт реакции содержит 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, или 89,5%, или 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97% (-)-амброксида.

19. Реакционная смесь или продукт реакции по любому из пп. 14-18, где определение (-)-амброксида производится по площади пика, измеренного с помощью газовой хроматографии.

20. Реакционная смесь или продукт реакции по любому из пп. 14-19 с практически полным отсутствием обнаружения побочных продуктов (II), (IV) или (III) в твердой (кристаллической) форме.

21. Реакционная смесь или продукт реакции по любому из пп. 14-20, где количество продукта (-)-амброксида составляет от приблизительно 20 г/л до приблизительно 200 г/л.

22. Способ получения продукта, содержащего (-)-амброксид, включающий включение продукта реакции по любому из пп. 15-21 в продукт.

23. Способ по п. 22, где продукт представляет собой ароматизирующий продукт, косметический продукт, моющий продукт или мыльный продукт.

24. Применение продукта реакции по любому из пп. 15-21 в качестве составной части ароматизирующего продукта, или косметического продукта, или потребительского продукта.

1

```
SHCwt   MAEQLVEAPAYARTLDRAVEYILSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP 86
SHC1    MAEQLVEAPAYARTLDRAVEYILSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP 86
SHC2    MAEQLVEAPAYARTLDRAVEYILSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGSWALYPGGPP 86
SHC3    MAEQLVEAPAYARTLDRAVEYILSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP 86
SHC4    MAEQLVEAPAYARTLDRAVEYILSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP 86
SHC5    MAEQLVEAPAYARTLDRAVEYILSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP 86
SHC6    MAEQLVEAPAYARTLDRAVEYILSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP 86
SHC7    MAEQLVEAPAYARTLDRAVEYILSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP 86
SHC8    MAEQLVEAPAYARTLDRAVEYILSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP 86
SHC9    MAEQLVEAPAYARTLDRAVEYILSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP 86
SHC10   MAEQLVEAPAYARTLDRAVEYILSCQKDEGYWWGPLLSNVTMEAEYVLLCHILDRVDRDRMEKIRRYLLHEQREDGTWALYPGGPP 86

SHCwt   DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARA 172
SHC1    DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARA 172
SHC2    DLDTTMEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVLTRMWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARA 172
SHC3    DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARA 172
SHC4    DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARA 172
SHC5    DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVLTRMWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARA 172
SHC6    DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARA 172
SHC7    DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVLTRMWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARA 172
SHC8    DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVLTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARA 172
SHC9    DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVFTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARA 172
SHC10   DLDTTIEAYVALKYIGMSRDEEPMQKALRFIQSQGGIESSRVLTRRWLALVGEYPWEKVPMVPPEIMFLGKRMPLNIYEFGSWARA 172
```

ФИГ. 1

2

```
SHCwt  TVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW 258
SHC1   TVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW 258
SHC2   TVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW 258
SHC3   TVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRVLHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW 258
SHC4   TVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW 258
SHC5   TVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW 258
SHC6   TVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW 258
SHC7   TVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW 258
SHC8   TVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW 258
SHC9   TVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW 258
SHC10  TVVALSIVMSRQPVFPLPERARVPELYETDVPPRRRGAKGGGGWIFDALDRALHGYQKLSVHPFRRAAEIRALDWLLERQAGDGSW 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
```

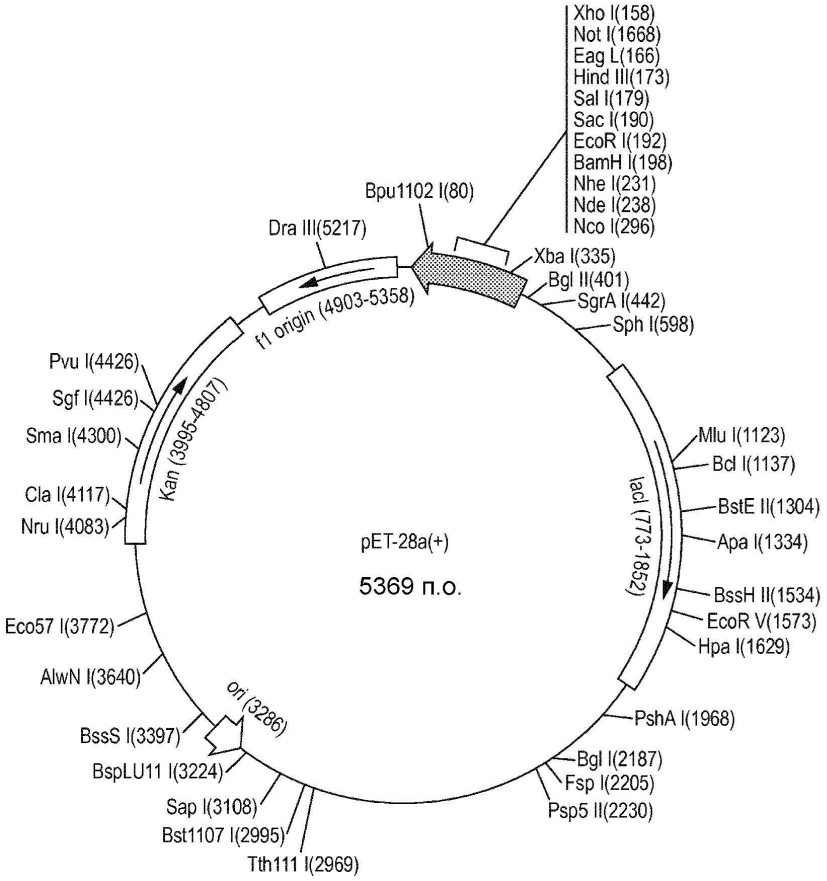
ФИГ. 2

```
SHCwt ITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPN 430
SHC1  ITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPN 430
SHC2  ITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPN 430
SHC3  ITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPN 430
SHC4  ITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPN 430
SHC5  ITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPN 430
SHC6  ITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPN 430
SHC7  ITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPN 430
SHC8  ITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPN 430
SHC9  ITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPN 430
SHC10 ITVPGDWAVKRPNLKPGGFAFQFDNVYYPDVDDTAVVVWALNTLRLPDERRRRDAMTKGFRWIVGMQSSNGGWGAYDVDNTSDLPN 430

SHCwt HIPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYI 516
SHC1  HIPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYI 516
SHC2  HIPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYI 516
SHC3  HIPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYI 516
SHC4  HIPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYI 516
SHC5  HIPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYI 516
SHC6  HIPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYI 516
SHC7  HIPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYI 516
SHC8  HIPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYI 516
SHC9  HIPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYI 516
SHC10 HIPFCDFGEVTDPPSEDVTAHVLECFGSFGYDDAWKVIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREPYI 516
```

ФИГ. 3

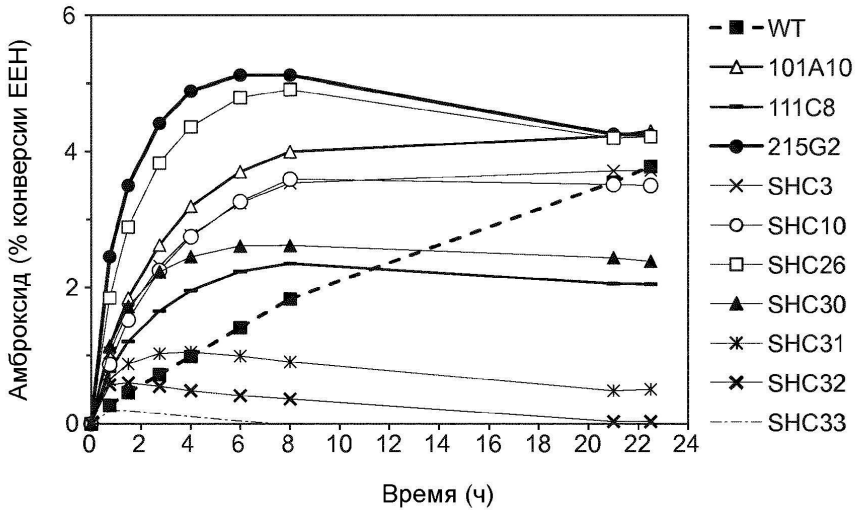
```
SHCwt   QKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFP 602
SHC1    QKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVYYLVETQRPDGGWDEPYYTGTGYP 602
SHC2    QKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFP 602
SHC3    QKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFP 602
SHC4    QKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGYP 602
SHC5    QKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFP 602
SHC6    QKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGFP 602
SHC7    QKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGYP 602
SHC8    QKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQILVETQRPDGGWDEPYYTGTGFP 602
SHC9    QKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGYP 602
SHC10   QKALDWVEQHQNPDGGWGEDCRSYEDPAYAGKGASTPSQTAWALMALIAGGRAESEAARRGVQYLVETQRPDGGWDEPYYTGTGYP 602

SHCwt   GDFYLGYTMYRHVFPTLALGRYKQAIERR 631
SHC1    GDFYLGYTMYRHVFPTLALGRYKQAIERR 631
SHC2    GDFYLGYTMYRHVFPTLALGRYKQAIERR 631
SHC3    GDFYLGYTMYRHVFPTLALGRYKQAIERR 631

SHC4    GDFYLGYTMYRHVFPTLALGRYKQAIERR 631
SHC5    GDFYLGYTMYRHVFPTLALGRYKQAIERR 631
SHC6    GDFYLGYTMYRHVFPTLALGRYKQAIERR 631
SHC7    GDFYLGYTMYRHVFPTLALGRYKQAIERR 631
SHC8    GDFYLGYTMYRHVFPTLALGRYKQAIERR 631
SHC9    GDFYLGYTMYRHVFPTLALGRYKQAIERR 631
```

ФИГ. 4


Xho I(158)
Not I(1668)
Eag L(166)
Hind III(173)
Sal I(179)
Sac I(190)
EcoR I(192)
BamH I(198)
Nhe I(231)
Nde I(238)
Nco I(296)
Bpu1102 I(80)
Dra III(5217)
Xba I(335)
Bgl II(401)
SgrA I(442)
Sph I(598)
f1 origin (4903-5358)
Pvu I(4426)
Sgf I(4426)
Sma I(4300)
Kan (3995-4807)
Cla I(4117)
Nru I(4083)
pET-28a(+)
5369 п.о.
MIu I(1123)
Bcl I(1137)
lacI (773-1852)
BstE II(1304)
Apa I(1334)
BssH II(1534)
EcoR V(1573)
Hpa I(1629)
Eco57 I(3772)
AlwN I(3640)
ori (3286)
PshA I(1968)
BssS I(3397)
Bgl I(2187)
BspLU11 I(3224)
Fsp I(2205)
Sap I(3108)
Psp5 II(2230)
Bst1107 I(2995)
Tth111 I(2969)

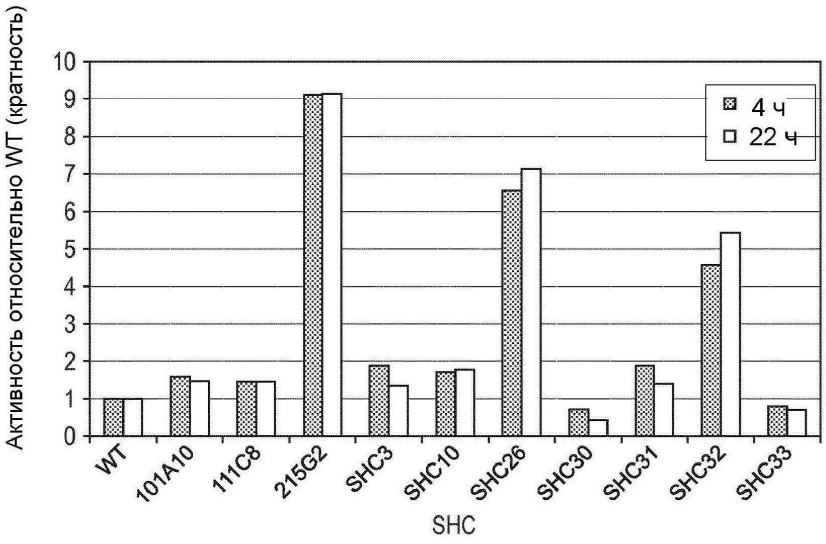
ФИГ. 5

6

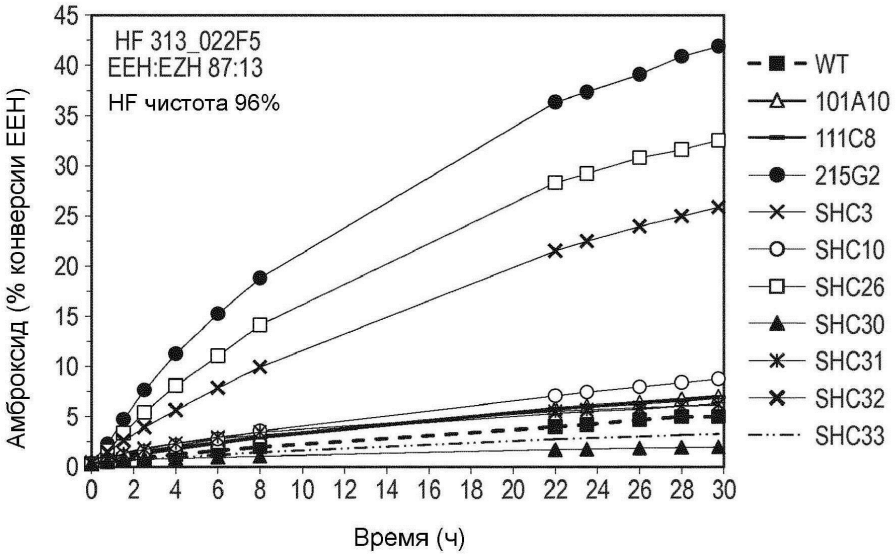

Амброксид (% конверсии EEH)
Время (ч)
WT
101A10
111C8
215G2
SHC3
SHC10
SHC26
SHC30
SHC31
SHC32
SHC33

ФИГ. 6

7

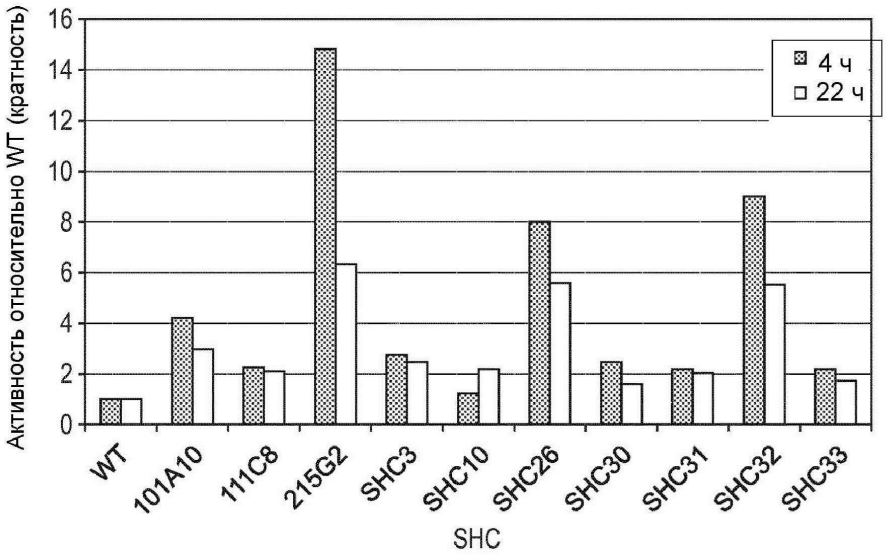



ФИГ. 7a

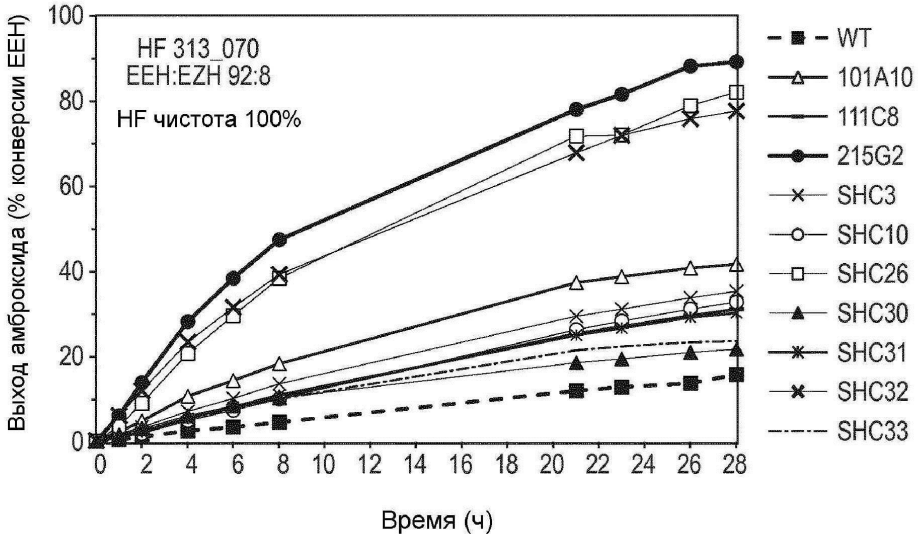



ФИГ. 7b

8

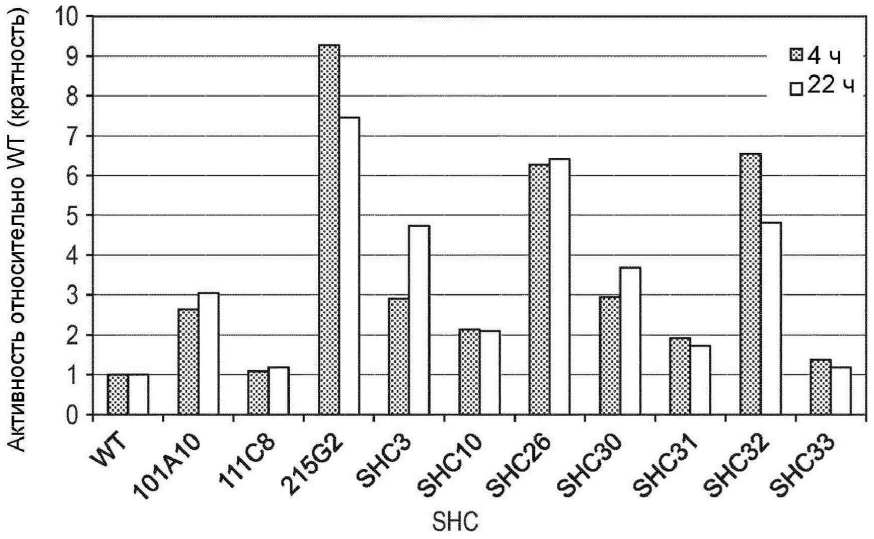



ФИГ. 8a

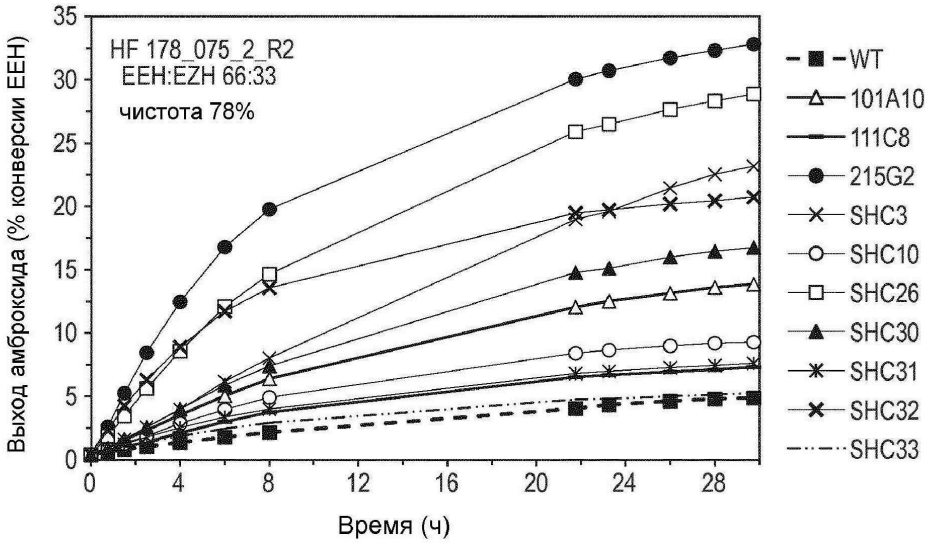



ФИГ. 8b

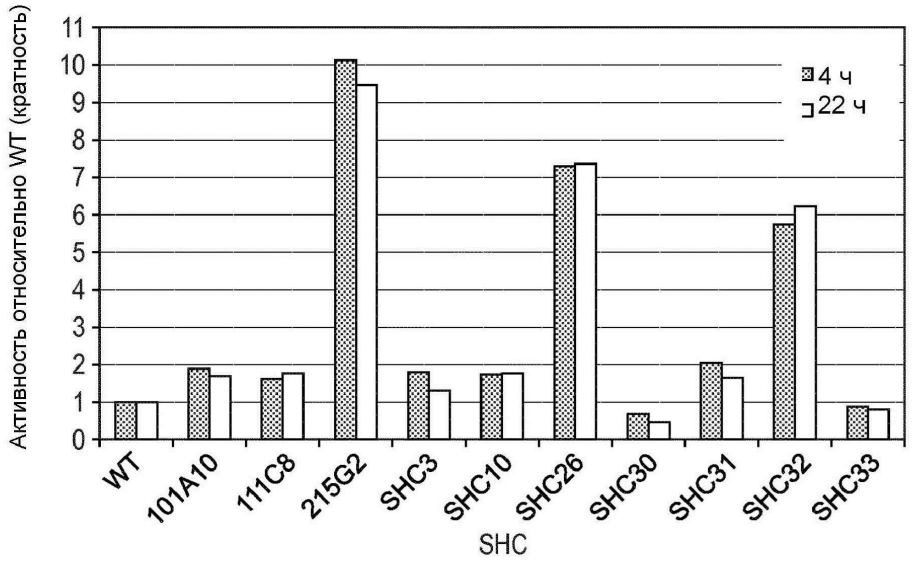
9


Активность относительно WT (кратность)
10
9
8
7
6
5
4
3
2
1
0
4 ч
22 ч
WT
101A10
111C8
215G2
SHC3
SHC10
SHC26
SHC30
SHC31
SHC32
SHC33
SHC

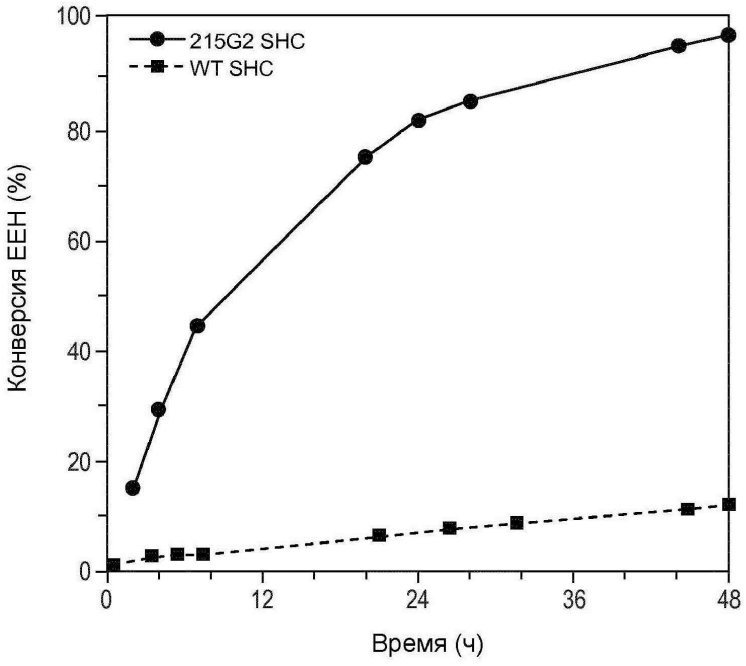
ФИГ. 9a


HF 178_075_2_R2
EEH:EZH 66:33
чистота 78%
Выход амброксида (% конверсии EEH)
35
30
25
20
15
10
5
0
0
4
8
12
16
20
24
28
Время (ч)
WT
101A10
111C8
215G2
SHC3
SHC10
SHC26
SHC30
SHC31
SHC32
SHC33

ФИГ. 9b

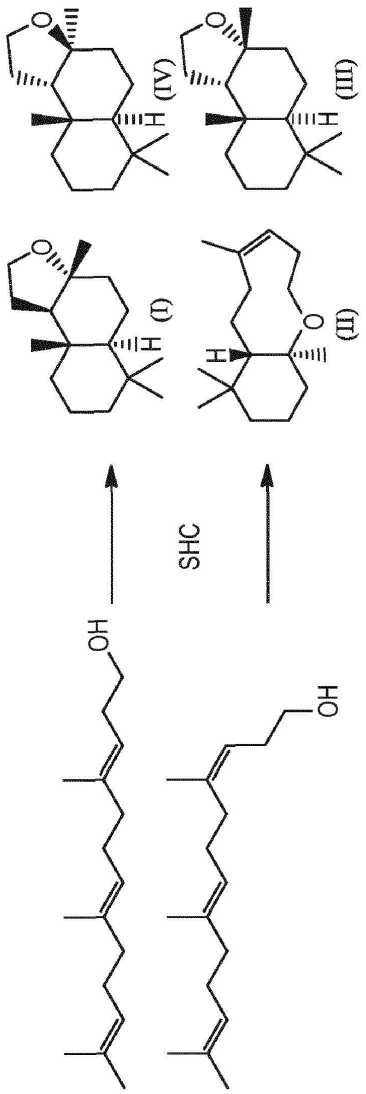



ФИГ. 10



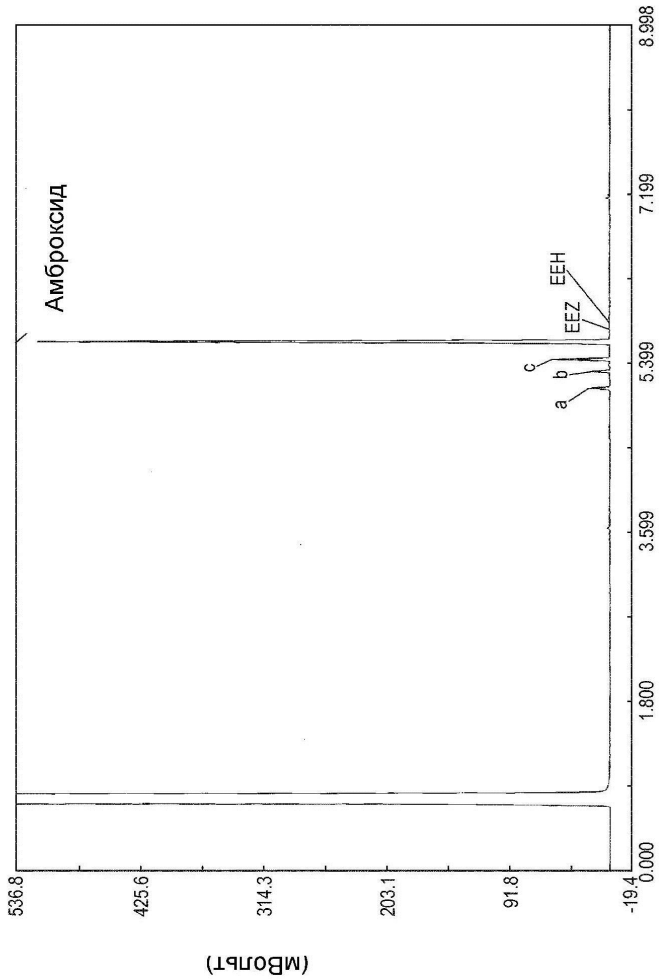

ФИГ. 11

11

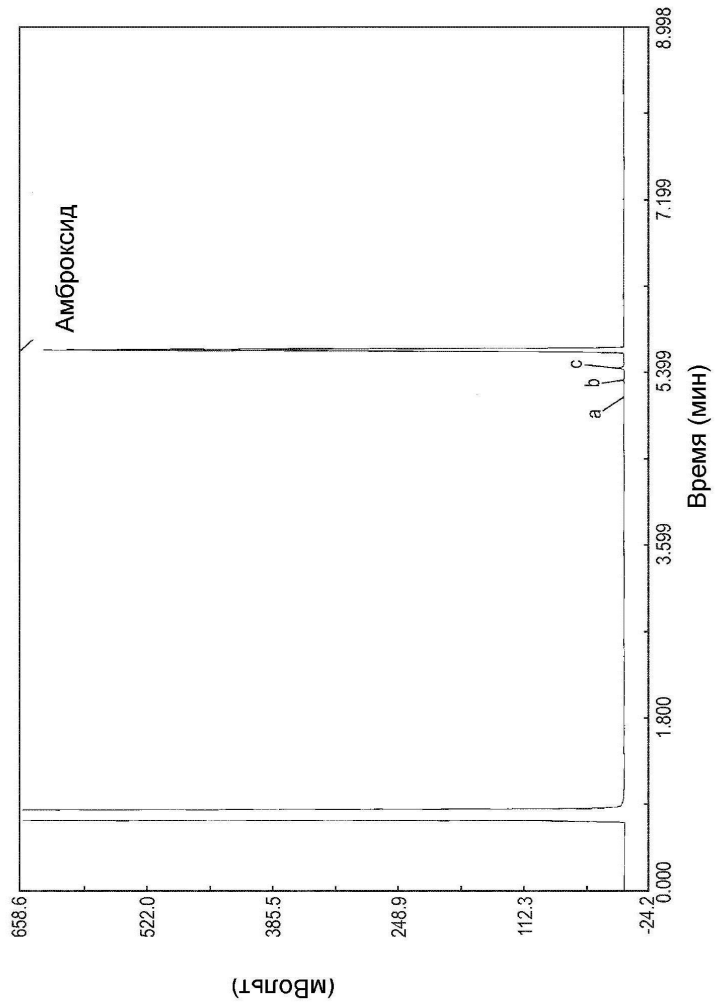



ФИГ. 12

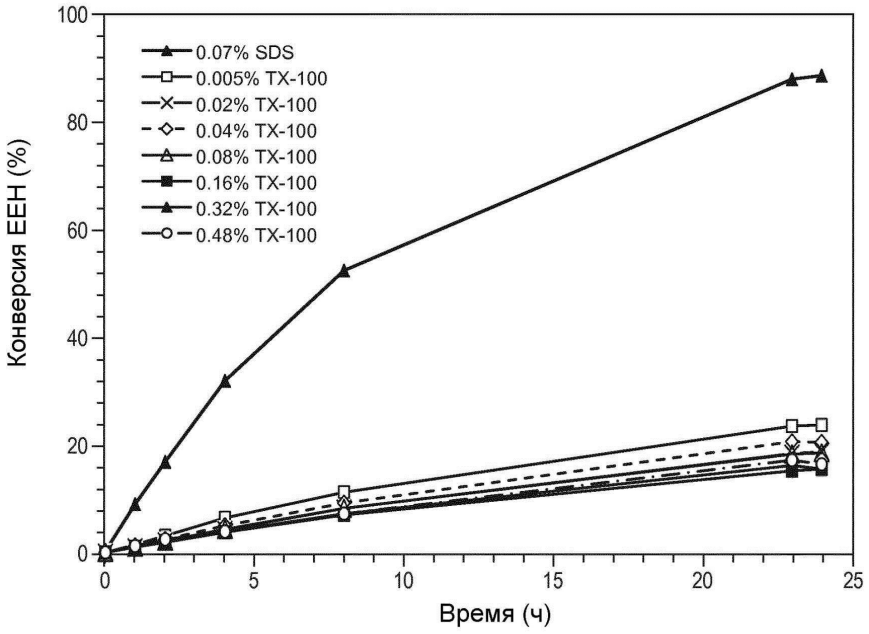
12


Амброксид
EEH
EEZ
c
b
a
536.8
425.6
314.3
203.1
91.8
-19.4
0.000
1.800
3.599
5.399
7.199
8.998
Время (мин)
(мВольт)

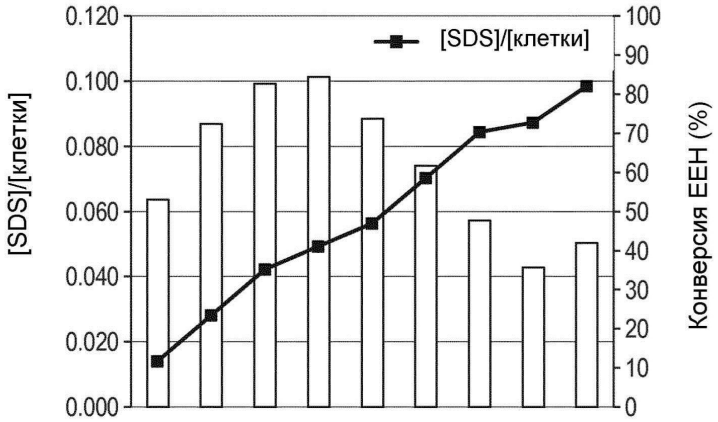
ФИГ. 13


Амброксид
a
b
c
658.6
522.0
385.5
248.9
112.3
-24.2
0.000
1.800
3.599
5.399
7.199
8.998
Время (мин)
(мВольт)

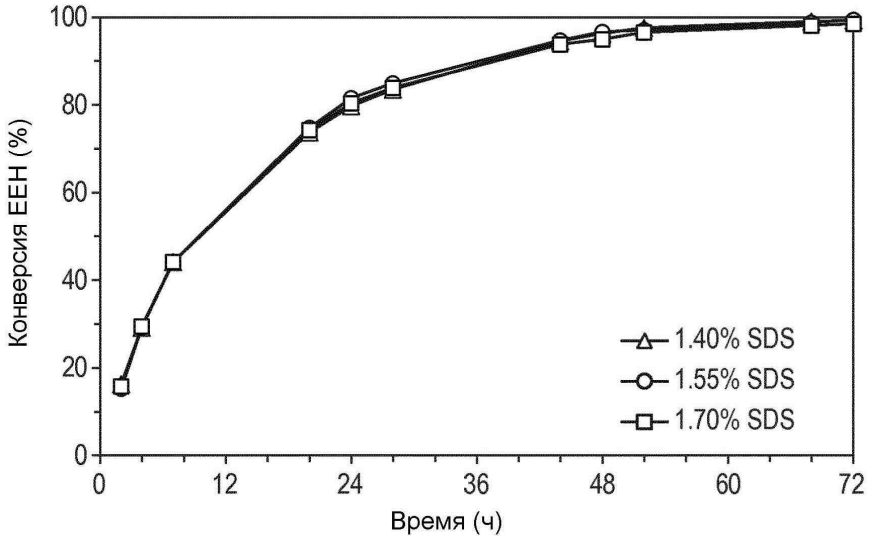
ФИГ. 14


0.07% SDS
0.005% TX-100
0.02% TX-100
0.04% TX-100
0.08% TX-100
0.16% TX-100
0.32% TX-100
0.48% TX-100
Конверсия ЕЕН (%)
Время (ч)
0
20
40
60
80
100
0
5
10
15
20
25

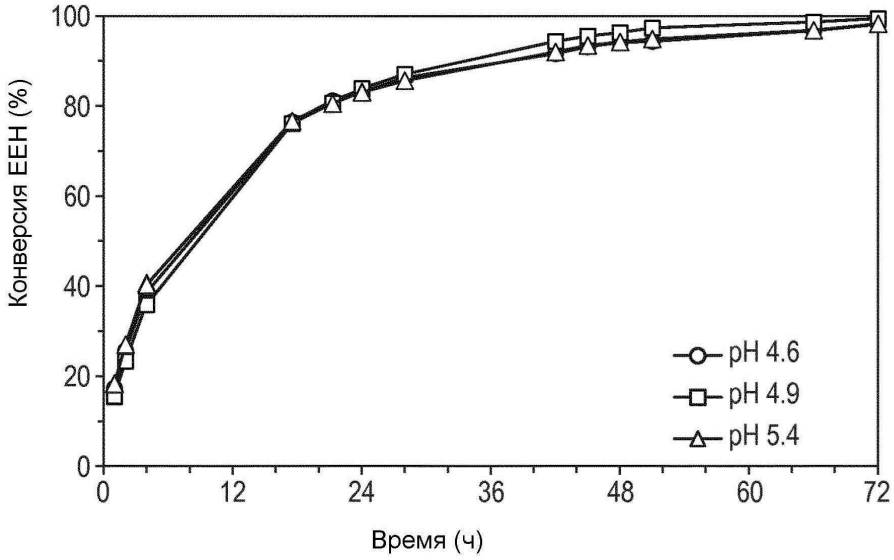
ФИГ. 15


[SDS]/[клетки]
[SDS]/[клетки]
Конверсия ЕЕН (%)
0.000
0.020
0.040
0.060
0.080
0.100
0.120
0
10
20
30
40
50
60
70
80
90
100

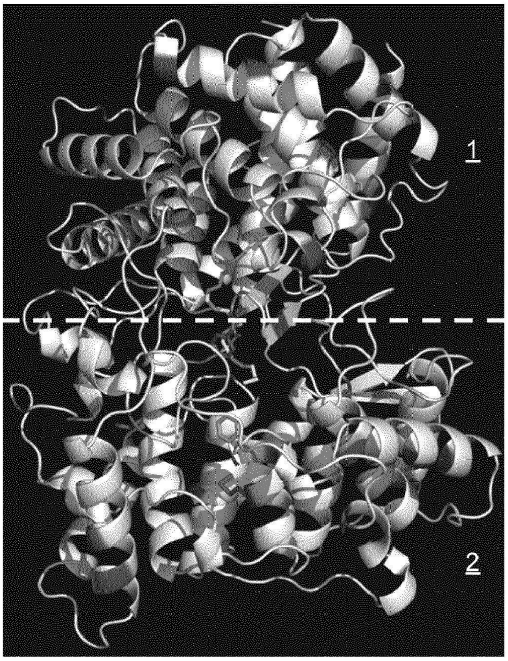
ФИГ. 16


Конверсия ЕЕН (%)
100
80
60
40
20
0
0
12
24
36
48
60
72
Время (ч)
1.40% SDS
1.55% SDS
1.70% SDS

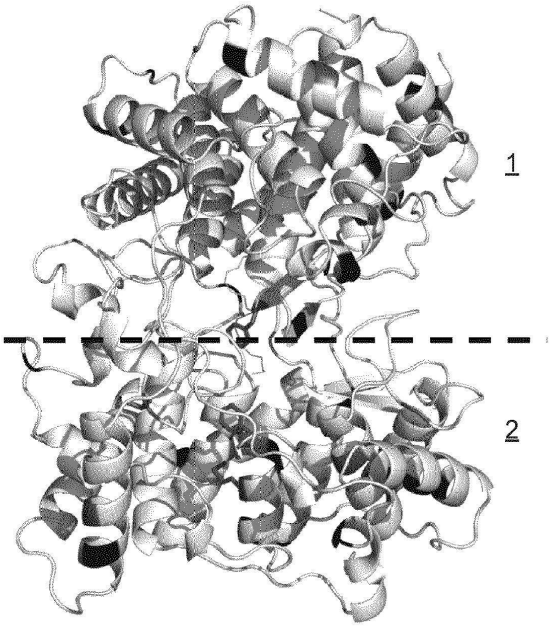
ФИГ. 17


Конверсия ЕЕН (%)
100
80
60
40
20
0
0
12
24
36
48
60
72
Время (ч)
pH 4.6
pH 4.9
pH 5.4

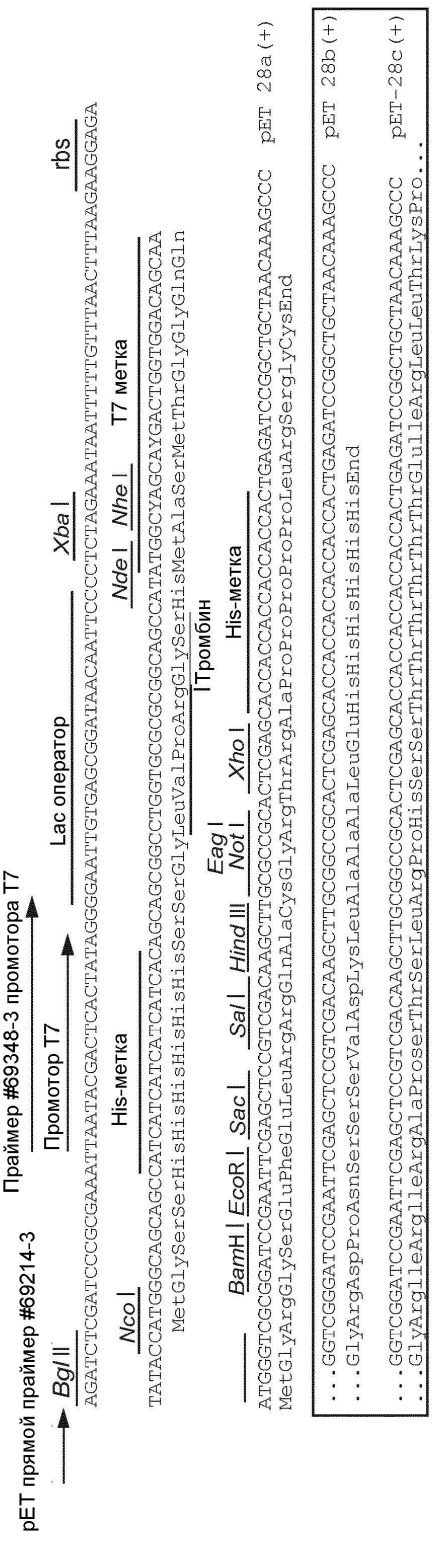
ФИГ. 18




ФИГ. 19

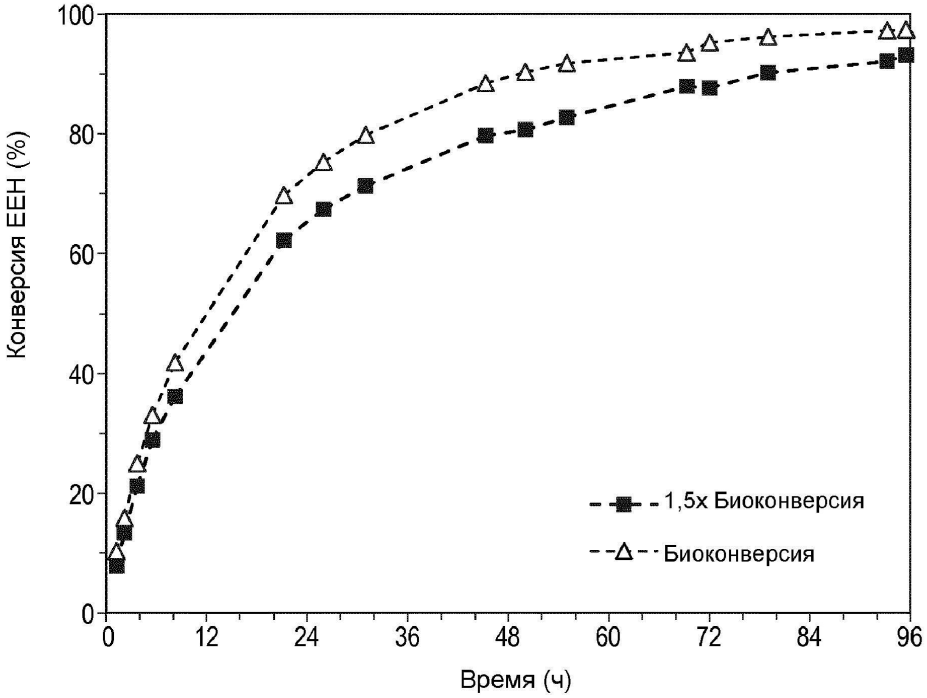



ФИГ. 20

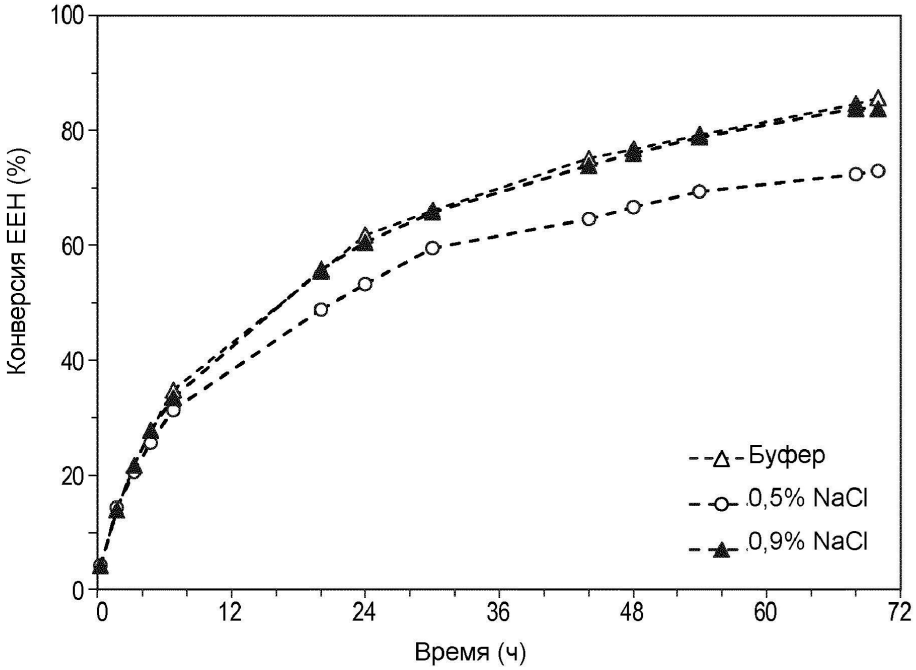



Область клонирования/экспрессии pET-28a-c(+)

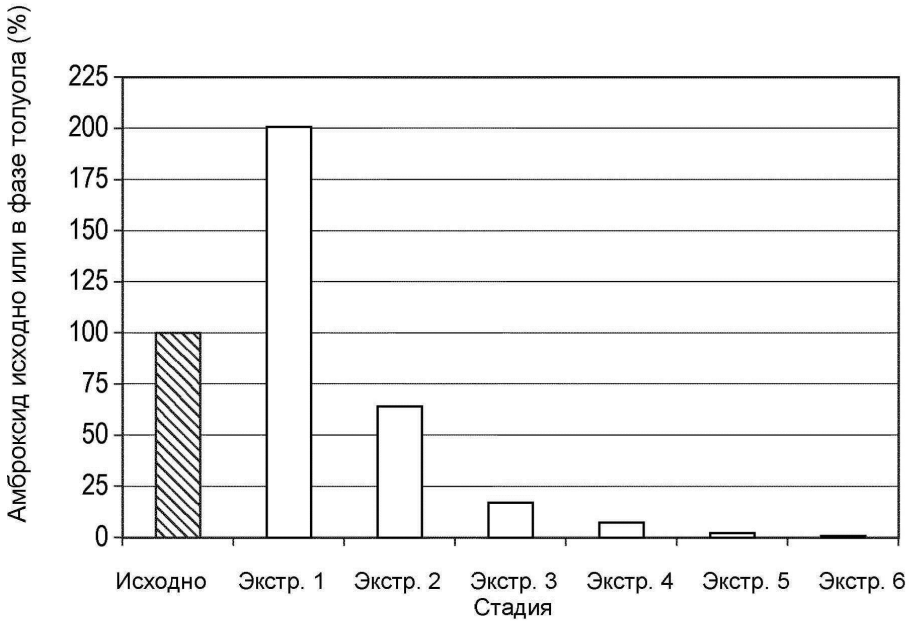
ФИГ. 21


Конверсия ЕЕН (%)
100
80
60
40
20
0
0
12
24
36
48
60
72
84
96
Время (ч)
1,5x Биоконверсия
Биоконверсия

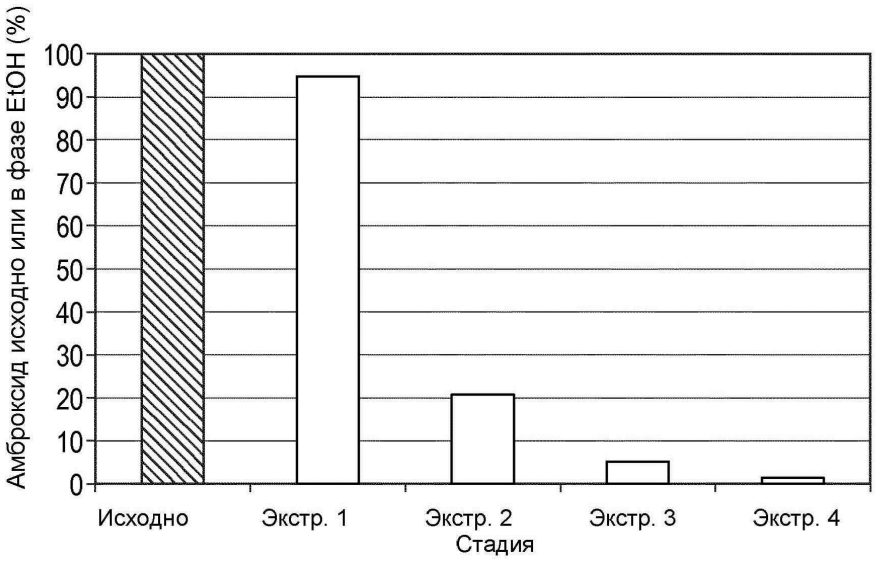
ФИГ. 22


100
80
60
40
20
0
Конверсия ЕЕН (%)
0
12
24
36
48
60
72
Время (ч)
Буфер
0,5% NaCl
0,9% NaCl

ФИГ. 23


Амброксид исходно или в фазе толуола (%)
225
200
175
150
125
100
75
50
25
0
Исходно
Экстр. 1
Экстр. 2
Экстр. 3
Экстр. 4
Экстр. 5
Экстр. 6
Стадия

ФИГ. 24


Амброксид исходно или в фазе EtOH (%)
100
90
80
70
60
50
40
30
20
10
0
Исходно
Экстр. 1
Экстр. 2
Экстр. 3
Экстр. 4
Стадия

ФИГ. 25