



(51) МПК

C12N 1/21 (2006.01)*C12P 13/12* (2006.01)*C12R 1/19* (2006.01)

ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ,
ПАТЕНТАМ И ТОВАРНЫМ ЗНАКАМ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(21), (22) Заявка: 2003131993/13, 03.11.2003

(24) Дата начала отсчета срока действия патента:
03.11.2003

(43) Дата публикации заявки: 10.05.2005

(45) Опубликовано: 27.04.2006 Бюл. № 12

(56) Список документов, цитированных в отчете о
поиске: US 6040160, 21.03.2000. US 6218168,
17.04.2001. RU 2175351 C2, 27.10.2001.

Адрес для переписки:

117279, Москва, ул. Миклухо-Маклая, 55А, ЗАО
"Фирма "Центр патентных услуг", пат.пов.
Е.А.Харченко

(72) Автор(ы):

Зиятдинов Михаил Харисович (RU),
Редькина Екатерина Игоревна (RU),
Гусятинер Михаил Маркович (RU)

(73) Патентообладатель(и):

Закрытое акционерное общество "Научно-
исследовательский институт Аджиномото-
Генетика" (ЗАО АГРИ) (RU)

(54) БАКТЕРИЯ, ПРИНАДЛЕЖАЩАЯ К РОДУ *ESCHERICHIA*, - ПРОДУЦЕНТ L-ЦИСТЕИНА И СПОСОБ ПОЛУЧЕНИЯ L-ЦИСТЕИНА

(57) Реферат:

Изобретение относится к биотехнологии и представляет собой способ получения L-цистеина с использованием бактерии, принадлежащей к роду *Escherichia*, в котором продукция L-аминокислоты указанной бактерией увеличена за счет усиления экспрессии генов кластера *cusPTWAM*. Согласно

данному способу бактерию выращивают в питательной среде, содержащей тиосульфат, и затем полученный и накопленный L-цистеин выделяют из культуральной среды. Заявленное изобретение позволяет получать L-цистеин с высокой степенью эффективности. 2 н. и 7 з.п. ф-лы, 2 табл.



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY,
PATENTS AND TRADEMARKS

(51) Int. Cl.

C12N 1/21 (2006.01)**C12P 13/12** (2006.01)**C12R 1/19** (2006.01)**(12) ABSTRACT OF INVENTION**(21), (22) Application: **2003131993/13, 03.11.2003**(24) Effective date for property rights: **03.11.2003**(43) Application published: **10.05.2005**(45) Date of publication: **27.04.2006 Bull. 12**

Mail address:

**117279, Moskva, ul. Miklukho-Maklaja, 55A,
ZAO "Firma "Tsentr patentnykh uslug",
pat.pov. E.A.Kharchenko**

(72) Inventor(s):

**Zijatdinov Mikhail Kharisovich (RU),
Red'kina Ekaterina Igorevna (RU),
Gusjatiner Mikhail Markovich (RU)**

(73) Proprietor(s):

**Zakrytoe aktsionernoe obshchestvo "Nauchno-
issledovatel'skij institut Adzhinomoto-
Genetika" (ZAO AGRI) (RU)**

(54) BACTERIUM BELONGING TO GENUS ESCHERICHIA AS PRODUCER OF L-CYSTEINE AND METHOD FOR PREPARING L-CYSTEINE

(57) Abstract:

FIELD: biotechnology, microbiology, amino acids.

SUBSTANCE: invention represents a method for preparing L-cysteine by using bacterium belonging to genus *Escherichia* wherein production of L-amino acid by indicated bacterium is increased due to enhancing expression of genes of

cluster *cysPTWAM*. Microorganisms are cultured in nutrient medium containing thiosulfate followed by isolation of prepared and accumulated L-cysteine from cultural medium. Invention provides preparing L-cysteine with high degree of effectiveness.

EFFECT: improved preparing method.

9 cl, 2 tbl, 5 ex

Область техники.

Настоящее изобретение относится к биотехнологии, в частности к способу получения L-цистеина методом ферментации, и, более конкретно, к генам, полученным из бактерии *Escherichia coli*. Эти гены являются полезными для улучшения продукции L-цистеина.

5 Предшествующий уровень техники

Традиционно L-аминокислоты в промышленном масштабе могут быть получены методом ферментации с использованием штаммов микроорганизмов, полученных из природных источников, или их мутантов, специально модифицированных для того, чтобы

10 увеличить продукцию L-аминокислот.
 Описано множество методов увеличения продукции L-аминокислот, например, путем трансформации микроорганизма рекомбинантной ДНК (см., например, патент США 4278765). Указанные методы основаны на повышении активности ферментов, вовлеченных в биосинтез аминокислот, и/или уменьшении чувствительности целевого фермента к ингибированию продуцируемой L-аминокислотой по принципу обратной связи (см.,
 15 например, выложенную патентную заявку Японии №56-18596 (1981), WO 95/16042 или патенты США 5661012 и 6040160).

Синтез L-цистеина из неорганической серы является основным механизмом, благодаря которому восстановленная сера включается в органические соединения микроорганизмов, таких как *Salmonella* и *Escherichia coli*. В этом процессе неорганический сульфат,
 20 самый распространенный источник утилизируемой серы аэробной биосферы, поступает внутрь клетки и восстанавливается до сульфида, который, в свою очередь, включается в L-цистеин с использованием механизма, аналогичного фиксации аммония в глутамин или глутамат. Функцию транспорта сульфата в клетку выполняет сульфатпермеаза, кодируемая генами *cysTWA* и *sbp* (sulphate binding protein - сульфат-связывающий белок). Два
 25 дополнительных механизма фиксации серы были описаны для *S.typhimurium* и *E.coli*. Первый способ фиксации осуществляется благодаря реакции тиосульфата с O-ацетил-L-серином, катализируемой O-ацетилсерин(тиол)-лиазой-B, кодируемой геном *cysM*, с образованием тиосульфат S-сульфоцистеина, который в дальнейшем восстанавливается до L-цистеина. При этом механизме транспорт тиосульфата в клетку
 30 осуществляется тиосульфатпермеазой, кодируемой генами *cysPTWA*. Кроме того, сульфид может включаться в O-ацетил-L-серин с помощью реакции, катализируемой O-ацетилсерин(тиол)-лиазой-A или B, кодируемых генами *cysK* и *cysM*, соответственно, с образованием L-цистеина. Второй механизм - это реакция взаимодействия O-сукцинил-L-гомосерина с сульфидом с образованием гомоцистеина, катализируемая цистатионин-γ-синтазой.
 35 (*Escherichia coli* and *Salmonella*, Second Edition, Editor in Chief: F.C.Neidhardt, ASM Press, Washington D.C., 1996).

Был опубликован процесс получения L-треонина путем ферментации микроорганизмов, принадлежащих к семейству *Enterobacteriaceae*, в котором экспрессия как минимум одного или более генов пути биосинтеза цистеина, выбранного (ых) из группы, включающей в
 40 себя гены *cysG*, *cysB*, *cysZ*, *cysK*, *cysM*, *cysA*, *cysW*, *cysU*, *cysP*, *cysD*, *cysN*, *cysC*, *cysJ*, *cysI*, *cysH*, *cysE* и *sbp* усилена (WO 03006666 A2).

Но в настоящее время нет данных, описывающих тот факт, что повышение экспрессии кластера генов *cysPTWAM* в клетках бактерии-производителя L-цистеина, растущей на среде, содержащей тиосульфат, может приводить к улучшению продукции L-цистеина.

45 Описание изобретения

Целью настоящего изобретения является увеличение продуктивности штаммов, продуцирующих L-цистеин, и предоставление способа получения L-цистеина с использованием указанных штаммов.

50 Данная цель была достигнута путем усиления экспрессии кластера генов *cysPTWA*, совместно с геном *cysM*, кодирующих систему транспорта в клетку сульфата/тиосульфата и O-ацетилсерин(тиол)-лиазу-B соответственно, что приводит к увеличению продукции L-цистеина у соответствующего штамма - производителя L-цистеина в случае, когда указанный штамм культивируется в среде, содержащей тиосульфат. Таким образом, было совершено

настоящее изобретение.

Настоящее изобретение включает в себя следующее:

1. Бактерия, принадлежащая к роду *Escherichia* - продуцент L-цистеина, где указанная бактерия модифицирована таким образом, что экспрессия генов кластера *cysPTWAM* усилена.

2. Бактерия, в соответствии с п.1, в которой экспрессия генов кластера *cysPTWAM* усилена за счет увеличения количества копий генов кластера *cysPTWAM* или за счет модификации последовательности контроля экспрессии таким образом, что в результате такой модификации экспрессия этих генов усиливается.

3. Бактерия, в соответствии с п.2, в которой количество копий указанных генов увеличено в результате трансформации указанной бактерии многокопийным вектором, содержащим гены кластера *cysPTWAM*.

4. Бактерия, в соответствии с п.3, в которой природный промотор указанного кластера заменен на более сильный промотор.

5. Бактерия, в соответствии с п.4, в которой количество копий указанных генов увеличено путем интеграции дополнительных копий генов кластера *cysPTWAM* в хромосому указанной бактерии.

6. Бактерия, в соответствии с п.5, в которой гены кластера *cysPTWAM* получены из бактерии, принадлежащей роду *Escherichia*.

7. Бактерия, в соответствии с п.6, где бактерия дополнительно модифицирована таким образом, что экспрессия открытой рамки считывания *udeD* усилена.

8. Способ получения L-цистеина, включающий стадии выращивания бактерии в соответствии с пунктами с 1 по 7, в питательной среде, содержащей тиосульфат, и выделения L-цистеина из культуральной жидкости.

9. Способ в соответствии с п.8, в котором у указанной бактерии усилена экспрессия генов биосинтеза L-цистеина.

Настоящее изобретение более детально будет описано ниже.

Бактерией, согласно настоящему изобретению, является бактерия, принадлежащая к роду *Escherichia* - продуцент L-цистеина, в которой усилена экспрессия генов кластера *cysPTWAM*, приводящая к повышению продукции L-цистеина. Конкретно, бактерией согласно настоящему изобретению является бактерия, принадлежащая к роду *Escherichia* - продуцент L-цистеина, модифицированная таким образом, что экспрессия генов кластера *cysPTWAM* указанной бактерией усилена.

Термин «бактерия, обладающая способностью к продукции L-цистеина» означает бактерию, способную накапливать L-цистеин в среде, когда такая бактерия согласно настоящему изобретению культивируется, в питательной среде. Способность к продукции L-цистеина может быть придана или усилена селекцией. Используемый здесь термин «бактерия, обладающая способностью к продукции L-цистеина» также означает бактерию, способную производить и накапливать в культуральной среде L-цистеин в количествах, больших, чем природный или родительский штаммы, и, прежде всего, означает, что микроорганизм способен производить и накапливать в среде L-цистеин в концентрациях не меньше, чем 0,5 г/л, более предпочтительно не меньше, чем 1,0 г/л.

Термин «бактерия, принадлежащая к роду *Escherichia*» означает, что бактерия относится к роду *Escherichia* в соответствии с классификацией, известной специалисту в области микробиологии. В качестве примера микроорганизма, принадлежащего к роду *Escherichia*, использованного в настоящем изобретении, может быть упомянута бактерия *Escherichia coli* (*E.coli*).

Термин "модифицирована таким образом, что экспрессия гена(ов) усилена" означает, что уровень экспрессии гена(ов) повысился, по сравнению с не модифицированным, например, природным штаммом. В качестве примера можно привести случай, когда увеличено количество экспрессируемого(ых) гена(ов) в пересчете на клетку или случай, когда повышен уровень экспрессии гена(ов). Количество копий экспрессируемого гена можно измерить, например, с помощью рестрикции хромосомной ДНК, и последующей

гибридизацией по Саузерну, с использованием сконструированного на основе нуклеотидной последовательности гена, зонда, флюоресцентной гибридации in situ (FISH) и так далее. Уровень экспрессии генов измеряется с помощью различных методов, включая гибридацию по Нозерну, количественную обратную транскрипцию - ПЦР (RT-PCR) и подобные им. Далее, в качестве примера природного штамма, который является эталоном для сравнения, может быть упомянут штамм *Escherichia coli* K-12. В результате увеличения экспрессии гена(ов) появляется эффект увеличения накопления L-цистеина в питательной среде.

Увеличение экспрессии генов кластера *cysPTWAM* в бактериальной клетке может быть достигнуто путем увеличения количества копий генов кластера *cysPTWAM* или модификацией нуклеотидной последовательности контроля экспрессии указанных генов таким образом, что в результате модификации экспрессия этих генов усиливается.

В качестве генов кластера *cysPTWAM*, могут быть использованы гены, полученные из бактерии, принадлежащей к роду *Escherichia* и гены, полученные из другой бактерии, такой как *Salmonella*. Среди вышеуказанных генов предпочтительны гены, полученные из бактерии, принадлежащей к роду *Escherichia*.

Нуклеотидная последовательность генов кластера *cysPTWAM* из *Escherichia coli* уже определена: для гена *cysP* - нуклеотиды с 2540532 по 2541548 в последовательности с инвентарным номером NC_000913.1, gi: 16130350 в базе данных GenBank, (SEQ ID NO: 1), для гена *cysT(cysU)* - нуклеотиды с 2539699 по 2540532 в последовательности с инвентарным номером NC_000913.1, gi:16130349 в базе данных GenBank, (SEQ ID NO: 3), для гена *cysW* - нуклеотиды с 2538824 по 2539699 в последовательности с инвентарным номером NC_000913.1, gi: 16132224 в базе данных GenBank, (SEQ ID NO: 5), для гена *cysA* - нуклеотиды с 2537737 по 2538834 в последовательности с инвентарным номером NC_000913.1, gi:16130348: в базе данных GenBank, (SEQ ID NO: 7), для гена *cysM* - нуклеотиды с 2536692 по 2537603 в последовательности с инвентарным номером NC_000913.1, gi:16130347 в базе данных GenBank, (SEQ ID NO: 9). Кластер генов *cysPTWAM* располагается на хромосоме штамма *E.coli* K-12 между локусом d2420 и локусом b2426. Таким образом, эти гены могут быть получены с помощью ПЦР (полимеразная цепная реакция; по White, T.J. et al., Trends Genet., 5,185 (1989)) с использованием затравок, синтезированных на основе нуклеотидной последовательности указанного гена. Гены из других микроорганизмов, кодирующие транспортную систему сульфата/тиосульфата, могут быть получены таким же образом.

Примером генов кластера *cysPTWAM*, полученных из *Escherichia coli*, являются следующие ДНК:

- ген *cysP*, кодирующий следующий белок (A) или (B):

(A) белок, который представлен аминокислотной последовательностью, приведенной в списке последовательностей под номером SEQ ID NO:2;

(B) белок, который представлен аминокислотной последовательностью, включающей делеции, замены, вставки или добавление одной или нескольких аминокислот в аминокислотную последовательность, приведенную в списке последовательностей под номером SEQ ID NO:2, и который обладает активностью периплазматического тиосульфат-связывающего белка;

- ген *cysT*, кодирующий следующий белок (C) или (D):

(C) белок, который представлен аминокислотной последовательностью, приведенной в списке последовательностей под номером SEQ ID NO:4;

(D) белок, который представлен аминокислотной последовательностью, включающей делеции, замены, вставки или добавление одной или нескольких аминокислот в аминокислотную последовательность, приведенную в списке последовательностей под номером SEQ ID NO:4, и который обладает активностью пермеазы сульфат/тиосульфат транспортной системы, вместе с белками (E) или (F), и (G) или (H);

- ген *cysW*, кодирующий следующий белок (E) или (F):

(E) белок, который представлен аминокислотной последовательностью, приведенной в

списке последовательностей под номером SEQ ID NO:6;

(F) белок, который представлен аминокислотной последовательностью, включающей делеции, замены, вставки или добавление одной или нескольких аминокислот в аминокислотную последовательность, приведенную в списке последовательностей под номером SEQ ID NO:6, и который обладает активностью пермеазы сульфат/тиосульфат транспортной системы, вместе с белками (C) или (D), и (G) или (H);

- ген *cysA*, кодирующий следующий белок (G) или (H):

(G) белок, который представлен аминокислотной последовательностью, приведенной в списке последовательностей под номером SEQ ID NO:8;

(H) белок, который представлен аминокислотной последовательностью, включающей делеции, замены, вставки или добавление одной или нескольких аминокислот в аминокислотную последовательность, приведенную в списке последовательностей под номером SEQ ID NO:8, и который обладает активностью пермеазы сульфат/тиосульфат транспортной системы, вместе с белками (C) или (D), и (E) или (F);

- ген *cysM*, кодирующий следующий белок (I) или (J):

(I) белок, который представлен аминокислотной последовательностью, приведенной в списке последовательностей под номером SEQ ID NO:10;

(J) белок, который представлен аминокислотной последовательностью, включающей делеции, замены, вставки или добавление одной или нескольких аминокислот в аминокислотную последовательность, приведенную в списке последовательностей под номером SEQ ID NO:8, и который обладает активностью O-ацетилсерин(тиол)-лиазы-B.

Термин "активность пермеазы сульфат/тиосульфат транспортной системы" означает активность белка, транспортирующего тиосульфат в клетку из внешней среды. Термин "активность O-ацетилсерин(тиол)-лиазы-B" означает каталитическую активность реакции между O-ацетил-L-серинном и тиосульфатом с образованием S-сульфоцистеина. Наличие указанных активностей может быть определено, к примеру, методом комплементации, с использованием бактерий, имеющих мутации в соответствующих генах.

ДНК, кодирующая вышеуказанные белки согласно настоящему изобретению включает в себя ДНК, кодирующую белок, который может содержать делеции, замены, вставки или добавление одной или нескольких аминокислот в аминокислотную последовательность в одной или более позициях белков (A), (C), (E), (G) и (I), при условии, что они не вызывают потерю активности указанного белка. Несмотря на то, что количество «нескольких» аминокислот различается в зависимости от положения и типа аминокислотного остатка в трехмерной структуре белка, оно может быть от 2 до 30, предпочтительно от 2 до 20, и более предпочтительно от 2 до 10 для белка, состоящего примерно из 300 аминокислотных остатков. ДНК, кодирующая практически такие же белки, как и белки, описанные в пунктах (A), (C), (E), (G) и (I) могут быть получены, например, модификацией нуклеотидной последовательности, кодирующей белок, описанный в пункте (A), (C), (E), (G) или (I) с использованием сайт-направленного мутагенеза таким образом, что один или несколько аминокислотных остатков будут удалены, заменены, введены или добавлены. Модифицированная подобным образом ДНК может быть получена традиционными методами, использующими обработку химическими реагентами и содержание в условиях, вызывающих мутации. К подобного рода обработкам относятся обработка ДНК, кодирующей белки согласно настоящему изобретению, с помощью гидроксилamina или обработка бактерии, содержащей ДНК, с помощью УФ-излучения или химического реагента, такого как N-метил-N'-нитро-N-нитрозогуанидин или азотистая кислота.

ДНК, содержащая кластер генов *cysPTWAM*, включает в себя варианты, которые могут быть найдены в различных штаммах или вариантах бактерий, принадлежащих к роду *Escherichia*, в виду природного разнообразия.

К методам увеличения экспрессии генов относятся методы увеличения числа копий гена. Введение гена в вектор, способный к функционированию в бактерии, принадлежащей к роду *Escherichia*, увеличивает число копий указанного гена. Для подобных целей могут

быть предпочтительно использованы многокопийные векторы. Примерами многокопийных векторов являются pBR322, pUC19, pBluescript KS⁺, pACYC177, pACYC184, pAYC32, pMW119, pET22b и подобные им. Кроме того, усиление экспрессии гена может быть достигнуто путем введения некоторого числа копий гена в бактериальную хромосому, например, методом гомологичной рекомбинации или подобным ему.

Трансформация бактерии с помощью ДНК, кодирующей белок, означает введение указанной ДНК в клетку бактерии, например, с помощью традиционных методов, для того, чтобы усилить экспрессию генов в бактериальной клетке.

Кроме того, усиление экспрессии генов может быть достигнуто помещением ДНК согласно настоящему изобретению под контроль сильного промотора взамен природного. Термин "природный промотор" означает существующую в природном организме область ДНК, расположенную перед открытой рамкой считывания (ORF - opened reading frame) гена и способствующую экспрессии этого гена. Общая трансляция генов кластера *cysPTWA* свидетельствует о том, что эти гены экспрессируются как одна транскрипционная единица с промотора, расположенного прямо перед геном *cysP*. В хромосоме *E.coli* ген *cysM* отделяют от гена *cysA* всего 174 п.о., и этот ген также может являться частью оперона *cysPTWA*, который на хромосоме транскрибируется против часовой стрелки (*Escherichia coli* and *Salmonella*, Second Edition, Editor in Chief: F.C.Neidhardt, ASM Press, Washington D.C., 1996). Сила промотора определяется частотой акта инициации синтеза РНК. Методы оценки силы промотора и примеры сильных промоторов описаны у Deuschle, U., Kammerer, W., Gentz, R., Bujard, H. (Promoters in *Escherichia coli*: a hierarchy of in vivo strength indicates alternate structures. *EMBO J.* 1986, 5, 2987-2994).

Усиление трансляции может быть достигнуто путем введения в ДНК согласно настоящему изобретению более эффективной последовательности Shine-Dalgarno (SD последовательности) вместо природной SD последовательности, где под SD последовательностью подразумевается область находящейся на цепи перед стартом кодоном мРНК, взаимодействующая с 16SРНК рибосомы (Shine J. and Dalgarno L., *Proc. Natl. Acad. Sci. USA*, 1974, 71,4, 1342-6). Термин "природная SD последовательность" означает SD последовательность, существующую в природном организме. SD последовательность гена $\phi 10$ из фага T7 может служить примером эффективной SD последовательности (Olins P.O. et al, *Gene*, 1988, 73, 227-235).

Использование сильного промотора может быть совмещено с увеличением количества копий генов в клетке. Также возможно увеличить количество копий генов кластера *cysPTWAM* путем комбинирования интеграции одного или нескольких генов указанного кластера с введением одного или нескольких генов кластера на многокопийном векторе.

Методы получения хромосомной ДНК, гибридизации, ПЦР, получения ДНК плазмид, расщепления и лигирования ДНК, трансформации, отбора олигонуклеотидов в качестве праймеров и другие подобные методы являются обычными методами, хорошо известными для специалиста в указанной области техники. Перечисленные методы описаны в Sambrook, J., and Russell D., "Molecular Cloning A Laboratory Manual, Third Edition", Cold Spring Harbor Laboratory Press (2001) или подобных изданиях.

Бактерия согласно настоящему изобретению может быть получена путем введения вышеуказанных ДНК в бактерию, уже обладающую способностью к продукции L-цистеина. С другой стороны, бактерия согласно настоящему изобретению может быть получена путем придания бактерии, уже содержащей указанные ДНК, способности к продукции L-цистеина.

В качестве родительских штаммов, в которых активности белков согласно настоящему изобретению будут повышены, могут быть использованы бактерии, обладающие способностью к продукции L-цистеина, принадлежащие к роду *Escherichia*, такие как штаммы *E.coli* JM15 и *E.coli* MG1655, трансформированные различными аллелями *cysE*, кодирующими устойчивую к ингибированию по принципу обратной связи сериновую ацетилтрансферазу (патент США 6,218,168, патентная заявка России 2003121601 соответственно), штамм *E.coli* W3110, обладающий генами, кодирующими белок,

способный к экскреции токсичных для клетки соединений (патент США 5972663), штаммы E.coli, содержащие цистеиновую десульфогидразу со сниженной активностью (патентная заявка Японии JP 11155571A2), штамм E.coli W3110, с повышенной активностью позитивного регулятора транскрипции цистеинового регулона, кодируемого геном *cysB* (заявка PCT WO 0127307 A1), и подобные им штаммы.

Ранее было показано, что ген *udeD*, кодирующий мембранный протеин, не участвующий в пути биосинтеза ни одной из L-аминокислот, может повысить продукцию L-цистеина в случае, когда дополнительные копии указанного гена введены в клетки соответствующего штамма-продуцента (патент США 5972663). Таким образом, желательно, чтобы бактерия-продуцент L-цистеина была бы в дальнейшем модифицирована таким образом, чтобы экспрессия открытой рамки считывания *udeD* была усилена.

Бактерия согласно настоящему изобретению может иметь повышенную экспрессию генов биосинтеза L-цистеина. Такие гены, позитивно влияющие на биосинтез L-цистеина, включают в себя ген *cysE*, кодирующий устойчивую к ингибированию по принципу обратной связи сериновую ацетилтрансферазу, и подобные гены.

Способ согласно настоящему изобретению включает в себя способ продукции L-цистеина, включающий стадии культивирования бактерий, согласно настоящему изобретению, в питательной среде с целью продукции и накопления L-цистеина в питательной среде, и выделения L-цистеина из культуральной жидкости.

В способе согласно настоящему изобретению выращивание бактерии, накопление и выделение L-цистеина из культуральной жидкости может быть осуществлено способом, подобным способу, традиционно используемому для продукции аминокислот методом ферментации с использованием бактерий. Питательная среда для выращивания может быть как синтетической, так и натуральной, при условии, что она содержит источники углерода, азота и серы, минеральные соединения и, если необходимо, питательные добавки в количестве, необходимом для роста бактерии. Источники углерода включают в себя различные углеводы, такие как глюкоза и сахароза, и различные органические кислоты. В зависимости от способности к их усвоению используемыми бактериями, могут быть использованы спирты, такие как этанол и глицерин. В качестве источников азота используются аммиак, различные соли аммония, такие как сульфат аммония, другие соединения азота, такие как амины, природные источники азота, такие как пептон, гидролизат соевых бобов и микробный ферментализат. В качестве источников серы, согласно настоящему изобретению, используются тиосульфаты. В качестве минеральных соединений используются однозамещенный фосфат калия, хлорид натрия, хлорид кальция, соли магния, соли железа, соли марганца, и подобные им.

При необходимости в среду могут быть добавлены и некоторые другие соединения. Например, если для роста микроорганизмам требуется метионин (метиониновая ауксотрофность) достаточное количество метионина может быть добавлено в среду культивирования.

Выращивание проводят предпочтительно в аэробных условиях, таких как взбалтывание и аэрация с перемешиванием, при температуре от 20°C до 42°C, предпочтительно от 34°C до 40°C. Значение pH обычно находится в пределах от 5 до 9, предпочтительно в пределах от 6.5 до 7.2. pH среды может быть скорректировано аммиаком, карбонатом кальция, различными кислотами, основаниями и буферами. Обычно выращивание в течение от 1 до 5 дней приводит к накоплению целевой L-аминокислоты в культуральной жидкости.

После выращивания нерастворимые компоненты, такие как клетки, могут быть удалены из питательной среды методом центрифугирования или фильтрации на мембране, после чего целевая L-аминокислота может быть выделена и очищена методами ионного обмена, концентрации или кристаллизации.

Наилучший способ осуществления изобретения.

Настоящее изобретение будет детальнее разъяснено со ссылкой на следующие примеры. В примерах все аминокислоты имеют L-конфигурацию, если не указано другое. В качестве матрицы для всех описанных ниже ПЦР была использована хромосомная ДНК

штамма *E.coli* MG1655 (ВКПМ В-6195).

Пример 1. Конструирование плазмиды содержащей ген *udeD*, мутантный ген *cysE* и мутантный ген *serA5*.

5 Ранее было показано, что ген *udeD* кодирует трансмембранный белок, положительно влияющий на продукцию цистеина (патент США 5972663). Поэтому ген *udeD* был амплифицирован с помощью ПЦР и клонирован. Затем мутантный ген *cysEX*, кодирующий устойчивую к ингибированию цистеином по принципу обратной связи сериновую ацетилтрансферазу, описанную в патенте США 6218168, был амплифицирован с помощью ПЦР и клонирован.

10 Также мутантный ген *serA5* кодирующий фосфоглицератдегидрогеназу, устойчивую к ингибированию серином по принципу обратной связи, описанную в патенте США 6180373, был амплифицирован с помощью ПЦР и клонирован. Амплификация мутантного гена *serA5* необходима для повышения внутриклеточной концентрации серина, являющегося предшественником L-цистеина.

15 В завершение, с помощью ПЦР был амплифицирован и клонирован промотор P_{ompA} , при этом использовались праймеры, приведенные в списке последовательностей под номерами SEQ ID No.11 (праймер PrOMPFAF) и No. 12 (праймер PrOMPAP). Праймер PrOMPFAF содержит на 5'-конце сайт узнавания рестриктазы Sall. Этот сайт Sall был также введен в прямые праймеры, использованные для амплификации генов *udeD*, *cysEX* и *serAS*. Таким образом, сайт Sall был использован для объединения промотора P_{otrA} и каждого из указанных генов. Праймер PrOMPAP содержит на 5'-конце сайт узнавания рестриктазы PaeI. Эти рестрикционные сайты были введены для последующего объединения всех генов на одной плазмиде, используемой для трансформации.

25 Затем все три гена, каждый из которых находился под промотором P_{ompA} , были клонированы на вектор pACYC184. Таким образом была получена плаزمида pACYC-DES.

Пример 2. Клонирование генов *cysPTWA* из *E.coli*.

Гены *cysPTWA*, кодирующие белки системы транспорта сульфата/тиосульфата в клетку, были амплифицированы с использованием праймеров, приведенных в списке последовательностей под номерами SEQ ID No.13 (праймер Mz025) и No.14 (праймер Mz026), для последующего клонирования. Праймер Mz025 представляет собой нуклеотидную последовательность, комплементарную последовательности, расположенной, начиная с 315 п.о. перед старт-кодоном гена *cysP* с введенным на 5'-конец праймера сайтом узнавания рестриктазы PaeI. Праймер Mz026 представляет собой нуклеотидную последовательность, комплементарную последовательности, начиная с 13 п.о. после стоп-кодона гена *cysA*, с введенным на 5'-конец праймера сайтом узнавания рестриктазы Sall. Полученный ПЦР фрагмент, содержащий кластер генов *cysPTWA* под собственным промотором, был обработан смесью рестриктаз PaeI и Sall, и вставлен в вектор pMW119, также предварительно обработанный теми же рестриктазами. Таким образом была получена плазмида pMW-PTWA.

40 Пример 3. Клонирование гена *cysM* из *E.coli*.

Ген *cysM*, кодирующий O-ацетилсерин(тиол)-лиазу-В был клонирован с помощью ПЦР, с использованием праймеров, приведенных в списке последовательностей под номерами SEQ ID No.15 (праймер *cysMF*) и No.16 (праймер *cysMR*). Праймер *cysMF* содержит сайт рестрикции Sall, введенный в 5'-конец. Праймер *cysMR* содержит сайт рестрикции XbaI, введенный в 5'-конец.

Промотор P_{cysK} был получен с помощью ПЦР с использованием праймеров, приведенных в списке последовательностей под номерами SEQ ID No.17 (праймер P_{cysKF}) и No.18 (праймер P_{cysKR}). Нуклеотидная последовательность праймера P_{cysKF} комплементарна последовательности, начиная с 6 п.о. перед старт кодоном гена *cysK* с введенным на 5'-конце сайтом рестрикции Sall. Нуклеотидная последовательность праймера P_{cysKR} комплементарна нуклеотидной последовательности, начиная с 301 п.о. перед старт-кодоном гена *cysK* с введенным на 5'-конце сайтом рестрикции PaeI. Затем два полученных продукта ПЦР объединили путем обработки полученной смеси

рестриктазой Sall, с последующим лигированием и клонированием в интегративный вектор pMW119int. Вектор pMW119int был сконструирован на основе коммерчески доступного вектора pMW119 путем введения сайтов фазмиды Mu attR и attL, необходимых для последующей Mu-интеграции. Таким образом была получена плаزمида pMW-P_{cysK}-cysM.

5 Пример 4. Конструирование штамма E.coli с повышенной активностью системы деградации L-цистеина.

Ранее гены tnaA и metC были описаны как гены, кодирующие белки основного пути деградации цистеина (патентная заявка Японии JP 2002271463). Таким образом, был сконструирован штамм E.coli MG1655 с неактивной системой деградации L-цистеина

10 (tnaA⁻, metC⁻).

Сначала была получена мутация в гене metC путем обработки нитрозогуанидином (NTG) с последующей многократной процедурой пенициллинового обогащения. Был отобран мутант, способный расти на цистатионине и не растущий на гомосерине. Затем в полученный штамм metC⁻ стандартной процедурой P1 трансдукции (Sambrook et al, "Molecular Cloning A Laboratory Manual, Second Edition", Cold Spring Harbor Laboratory Press (1989)) был перенесен разрушенный ген tnaA из штамма VKPM B-7427 (tnaA300::Tn10(Tc^R)).

Таким образом был получен штамм E.coli MT.

20 Пример 5. Эффект влияния усиленной экспрессии генов кластера cysPTWAMua на продукцию L-цистеина.

Штамм E.coli MT использовался в качестве родительского штамма для оценки эффекта влияния усиленной экспрессии генов кластера cysPTWAM на продукцию L-цистеина.

Сначала ген cysM интегрировали в хромосому штамма MT с использованием плазмиды pMW-P_{cysK}-cysM по стандартной процедуре Mu-интеграции.

25 Затем плазмиды pACYC-DES и pMW-PTWA были последовательно введены в полученные трансдуктанты MTintCYSM с образованием штаммов MTintCYSM/pACYC-DES и MTintCYSM/pACYC-DES/pMW-PTWA.

30 Оба штамма MTintCYSM/pACYC-DES и MTintCYSM/pACYC-DES/pMW-PTWA выращивались ночь на качалке при 34°C в 2 мл питательного бульона с добавками 100 мг/л ампициллина и 25 мкг/мл стрептомицина. 0.2 мл полученных культур были инокулированы в 2 мл ферментационной среды, содержащей хлорамфеникол (30 мг/л) и ампициллин (100 мг/л) в пробирках 20×200 мм, и выращивались при 34°C в течение 42 часов на качалке при 250 об/мин. Использовался следующий состав ферментационной

35 среды: 15.0 г/л (NH₄)₂SO₄, 1.5 г/л KH₂PO₄, 1.0 г/л MgSO₄, 20.0 г/л CaCO₃, 0.1 мг/л тиамин, 1% L-бульона, 4% глюкозы, 300 мг/л L-метионина и различные концентрации Na₂S₂O₃.

После выращивания количество накопленного в среде L-цистеина определяли с помощью метода, описанного Gaitonde, M.K. (Biochem. J., 104:2, 627-33 (1967)).

Полученные данные представлены в Таблице 1.

40

Таблица 1		
Штаммы	Концентрация тиосульфата, г/л	Концентрация L-цистеина, г/л
MTintCYSM/pACYC-DES	2.5	3.1
	5.0	3.4
	7.0	3.8
	10.0	2.8
MTintCYSM/pACYC-DES/pMW-PTWA	2.5	4.1
	5.0	5.6
	7.0	6.2
	10.0	5.5

45

50 Для ферментации в мини-ферментерах по одной петле каждого из штаммов MTintCYSM/pACYC-DES и MTintCYSM/pACYC-DES/pMW-PTWA, выращенных на L-агаре, перенесли в L-бульон и выращивали при 34°C со взбалтыванием (140 об/мин) до достижения оптической плотности OD₅₄₀≈2.0. Затем 25 мл посевной культуры внесли 250 мл ферментационной среды и выращивали при 34°C со взбалтыванием (1500 об/мин) в

течение 48 часов.

Состав среды для ферментации в мини-ферментерах (г/л):

5	Триптон	2.0
	Дрожжевой экстракт	1.0
	(NH ₄) ₂ SO ₄	5.0
	KH ₂ PCO ₄	1.5
	NaCl	0.5
	MgSO ₄ ·7H ₂ O	0.3
10	CaCl ₂ ·2H ₂ O	0.015
	FeSCO ₄ ·7H ₂ O	0.075
	Цитрат натрия·2H ₂ O	1.0
	Глюкоза	100.0
	Метионин	0.45

15 В среду для ферментации, начиная с 6 часа от начала ферментации, подавался тиосульфат со скоростью 0.4 г/л·ч.

После выращивания количество накопленного в ферментационной среде L-цистеина определялось, как описано выше. Полученные данные представлены в Таблице 2.

Таблица 2.	
Штаммы	Концентрация L-цистеина, г/л
MTintCYSM/pACYC-DES1	3.9
MTintCYSM/pACYC-DES/pMW-PTWA	6.6

25 Как видно из Таблиц 1 и 2, усиленная экспрессия генов кластера *cysPTWAM* приводит к повышению продукции цистеина у штамма MT.

30

35

40

45

50

СПИСОК ПОСЛЕДОВАТЕЛЬНОСТЕЙ

<110> Ajinomoto-Genetika Research Institute

5 <120> METHOD FOR PRODUCING L-CYSTEINE USING BACTERIA
 BELONGING TO THE GENUS ESCHERICHIA

<140>
 <141>

10 <160> 18

<170> PatentIn Ver. 2.1

<210> 1
 <211> 1017
 <212> DNA

15 <213> Escherichia coli

<220>
 <221> CDS
 <222> (1)..(1017)

20 <400> 1
 atg gcc gtt aac tta ctg aaa aag aac tca ctc gcg ctg gtc gct tct 48
 Met Ala Val Asn Leu Leu Lys Lys Asn Ser Leu Ala Leu Val Ala Ser
 1 5 10 15

ctg ctg ctg gcg ggc cat gta cag gca acg gaa ctg ctg aac agt tct 96
 Leu Leu Leu Ala Gly His Val Gln Ala Thr Glu Leu Leu Asn Ser Ser
 20 25 30

25 tat gac gtc tcc cgc gag ctg ttt gcc gcc ctg aat ccg ccg ttt gag 144
 Tyr Asp Val Ser Arg Glu Leu Phe Ala Ala Leu Asn Pro Pro Phe Glu
 35 40 45

30 caa caa tgg gca aaa gat aac ggc ggc gac aaa ctg acg ata aaa caa 192
 Gln Gln Trp Ala Lys Asp Asn Gly Gly Asp Lys Leu Thr Ile Lys Gln
 50 55 60

tct cat gcc ggg tca tca aaa cag gcg ctg gcg att tta cag ggc tta 240
 Ser His Ala Gly Ser Ser Lys Gln Ala Leu Ala Ile Leu Gln Gly Leu
 65 70 75 80

35 aaa gcc gac gtt gtc act tat aac cag gtg acc gac gta caa atc ctg 288
 Lys Ala Asp Val Val Thr Tyr Asn Gln Val Thr Asp Val Gln Ile Leu
 85 90 95

40 cac gat aaa ggc aag ctg atc ccg gcc gac tgg cag tcg cgc ctg ccg 336
 His Asp Lys Gly Lys Leu Ile Pro Ala Asp Trp Gln Ser Arg Leu Pro
 100 105 110

aat aat agc tcg ccg ttc tac tcc acc atg ggc ttc ctg gtg cgt aag 384
 Asn Asn Ser Ser Pro Phe Tyr Ser Thr Met Gly Phe Leu Val Arg Lys
 115 120 125

45 ggt aac ccg aag aat atc cac gat tgg aac gac ctg gtg cgc tcc gac 432
 Gly Asn Pro Lys Asn Ile His Asp Trp Asn Asp Leu Val Arg Ser Asp
 130 135 140

gtg aag ctg att ttc ccg aac ccg aaa acg tcg ggt aac gcg cgt tat 480
 Val Lys Leu Ile Phe Pro Asn Pro Lys Thr Ser Gly Asn Ala Arg Tyr

50

RU 2 275 425 C2

	145		150		155		160										
	acc	tat	ctg	gcg	gca	tgg	ggc	gca	gcg	gat	aaa	gct	gac	ggt	ggt	gac	528
	Thr	Tyr	Leu	Ala	Ala	Trp	Gly	Ala	Ala	Asp	Lys	Ala	Asp	Gly	Gly	Asp	
				165						170						175	
5	aaa	ggc	aaa	acc	gaa	cag	ttt	atg	acc	cag	ttc	ctg	aaa	aac	ggt	gaa	576
	Lys	Gly	Lys	Thr	Glu	Gln	Phe	Met	Thr	Gln	Phe	Leu	Lys	Asn	Val	Glu	
				180						185					190		
	gtg	ttc	gat	act	ggc	ggt	cgt	ggc	gcg	acc	acc	act	ttt	gcc	gag	cgc	624
	Val	Phe	Asp	Thr	Gly	Gly	Arg	Gly	Ala	Thr	Thr	Thr	Phe	Ala	Glu	Arg	
10				195						200					205		
	ggc	ctg	ggc	gat	gtg	ctg	att	agc	ttc	gaa	tcg	gaa	gtg	aac	aac	atc	672
	Gly	Leu	Gly	Asp	Val	Leu	Ile	Ser	Phe	Glu	Ser	Glu	Val	Asn	Asn	Ile	
				210						215					220		
15	cgt	aaa	cag	tat	gaa	gcg	cag	ggc	ttt	gaa	gtg	gtg	att	ccg	aaa	acc	720
	Arg	Lys	Gln	Tyr	Glu	Ala	Gln	Gly	Phe	Glu	Val	Val	Ile	Pro	Lys	Thr	
						230						235				240	
	aac	att	ctg	gcg	gaa	ttc	ccg	gtg	gcg	tgg	ggt	gat	aaa	aac	gtg	cag	768
	Asn	Ile	Leu	Ala	Glu	Phe	Pro	Val	Ala	Trp	Val	Asp	Lys	Asn	Val	Gln	
					245						250					255	
20	gcc	aac	ggt	acg	gaa	aaa	gcc	gcc	aaa	gcc	tat	ctg	aac	tgg	ctc	tat	816
	Ala	Asn	Gly	Thr	Glu	Lys	Ala	Ala	Lys	Ala	Tyr	Leu	Asn	Trp	Leu	Tyr	
				260						265					270		
	agc	ccg	cag	gcg	caa	acc	atc	atc	acc	gac	tat	tac	tac	cgc	gtg	aat	864
	Ser	Pro	Gln	Ala	Gln	Thr	Ile	Ile	Thr	Asp	Tyr	Tyr	Tyr	Arg	Val	Asn	
25				275					280					285			
	aac	ccg	gag	gtg	atg	gac	aaa	ctg	aaa	gac	aaa	ttc	ccg	cag	acc	gag	912
	Asn	Pro	Glu	Val	Met	Asp	Lys	Leu	Lys	Asp	Lys	Phe	Pro	Gln	Thr	Glu	
				290				295					300				
30	ctg	ttc	cgc	gtg	gaa	gac	aaa	ttt	ggc	tcc	tgg	ccg	gaa	gtg	atg	aaa	960
	Leu	Phe	Arg	Val	Glu	Asp	Lys	Phe	Gly	Ser	Trp	Pro	Glu	Val	Met	Lys	
						310						315				320	
	acc	cac	ttc	acc	agc	ggc	ggc	gag	tta	gac	aag	ctg	tta	gcg	gcg	ggg	1008
	Thr	His	Phe	Thr	Ser	Gly	Gly	Glu	Leu	Asp	Lys	Leu	Leu	Ala	Ala	Gly	
35						325						330				335	
	cgt	aac	tga														1017
	Arg	Asn															
40	<210>	2															
	<211>	338															
	<212>	PRT															
	<213>	Escherichia coli															
	<400>	2															
45	Met	Ala	Val	Asn	Leu	Leu	Lys	Lys	Asn	Ser	Leu	Ala	Leu	Val	Ala	Ser	
	1				5					10					15		
	Leu	Leu	Leu	Ala	Gly	His	Val	Gln	Ala	Thr	Glu	Leu	Leu	Asn	Ser	Ser	
				20					25						30		
	Tyr	Asp	Val	Ser	Arg	Glu	Leu	Phe	Ala	Ala	Leu	Asn	Pro	Pro	Phe	Glu	
				35				40					45				
	Gln	Gln	Trp	Ala	Lys	Asp	Asn	Gly	Gly	Asp	Lys	Leu	Thr	Ile	Lys	Gln	
50				50				55					60				

RU 2 275 425 C2

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

30 <210> 3
<211> 834
<212> DNA
<213> Escherichia coli

<220>
35 <221> CDS
<222> (1)..(834)

<400> 3
atg ttt gct gtc tcc tcc aga cgc gtg ctg ccg ggc ttt acc tta agc 48
Met Phe Ala Val Ser Ser Arg Arg Val Leu Pro Gly Phe Thr Leu Ser
1 5 10 15
40 ctc ggc acc agt ctg ctg ttt gtg tgc ctg att ttg ctg ctg ccg ctc 96
Leu Gly Thr Ser Leu Leu Phe Val Cys Leu Ile Leu Leu Leu Pro Leu
20 25 30
toc gcg ctg gtg atg caa ctg gcc cag atg agc tgg gcg cag tac tgg 144
Ser Ala Leu Val Met Gln Leu Ala Gln Met Ser Trp Ala Gln Tyr Trp
35 40 45
45 gag gtg atc acc aac ccg cag gtg gtc gcg gcc tac aaa gta acg ctg 192
Glu Val Ile Thr Asn Pro Gln Val Val Ala Ala Tyr Lys Val Thr Leu
50 55 60

50

RU 2 275 425 C2

ctg tcg gcg ttt gtg gca tcg att ttt aac ggc gtt ttc ggt ctg ctg 240
 Leu Ser Ala Phe Val Ala Ser Ile Phe Asn Gly Val Phe Gly Leu Leu
 65 70 75 80

5 atg gcg tgg atc cta acc cgc tat cgc ttc cca ggc cgc acg ctg ctt 288
 Met Ala Trp Ile Leu Thr Arg Tyr Arg Phe Pro Gly Arg Thr Leu Leu
 85 90 95

gat gcg ctg atg gat tta ccc ttt gcg ctg cca acg gct gtc gcc ggt 336
 Asp Ala Leu Met Asp Leu Pro Phe Ala Leu Pro Thr Ala Val Ala Gly
 100 105 110

10 tta acg ctg gcc tcg ctc ttt tcc gta aac ggt ttt tac ggt gaa tgg 384
 Leu Thr Leu Ala Ser Leu Phe Ser Val Asn Gly Phe Tyr Gly Glu Trp
 115 120 125

ctg gcg aag ttt gat atc aaa gtc acc tat aca tgg ctg ggg att gcg 432
 Leu Ala Lys Phe Asp Ile Lys Val Thr Tyr Thr Trp Leu Gly Ile Ala
 15 130 135 140

gtg gct atg gcc ttt acc agc att ccg ttt gtg gtg cgt acc gtg cag 480
 Val Ala Met Ala Phe Thr Ser Ile Pro Phe Val Val Arg Thr Val Gln
 145 150 155 160

20 ccg gtg ctg gaa gag tta ggc ccg gaa tat gaa gaa gcg gcg gaa acg 528
 Pro Val Leu Glu Glu Leu Gly Pro Glu Tyr Glu Glu Ala Ala Glu Thr
 165 170 175

ctt ggt gca acg cgc tgg cag agt ttc tgc aaa gtg gtg ctg ccg gag 576
 Leu Gly Ala Thr Arg Trp Gln Ser Phe Cys Lys Val Val Leu Pro Glu
 180 185 190

25 ctt tct ccg gcg ctg gtg gcg ggc gtg gcg ctg tcg ttt acc cgt agt 624
 Leu Ser Pro Ala Leu Val Ala Gly Val Ala Leu Ser Phe Thr Arg Ser
 195 200 205

ctt ggt gaa ttt ggc gcg gtg att ttt atc gcc gga aat atc gcg tgg 672
 Leu Gly Glu Phe Gly Ala Val Ile Phe Ile Ala Gly Asn Ile Ala Trp
 30 210 215 220

aag acg gaa gtg acg tcg ctg atg att ttt gtg cgc tta cag gag ttt 720
 Lys Thr Glu Val Thr Ser Leu Met Ile Phe Val Arg Leu Gln Glu Phe
 225 230 235 240

35 gat tac ccg gca gcg agc gcg att gct tcg gtg atc ctc gcg gca tct 768
 Asp Tyr Pro Ala Ala Ser Ala Ile Ala Ser Val Ile Leu Ala Ala Ser
 245 250 255

ctg ctg ctg ctg ttc tca att aac act ctg caa agt cgc ttt ggt cgg 816
 Leu Leu Leu Leu Phe Ser Ile Asn Thr Leu Gln Ser Arg Phe Gly Arg
 260 265 270

40 cgt gtg gta ggt cat taa 834
 Arg Val Val Gly His
 275

45 <210> 4
 <211> 277
 <212> PRT
 <213> Escherichia coli

<400> 4
 Met Phe Ala Val Ser Ser Arg Arg Val Leu Pro Gly Phe Thr Leu Ser

50

RU 2 275 425 C2

```

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

30 <210> 5
    <211> 876
    <212> DNA
    <213> Escherichia coli

    <220>
35 <221> CDS
    <222> (1)..(876)

    <400> 5
atg gcg gaa gtt acc caa ttg aag cgt tat gac gcg cgc ccg att aac 48
Met Ala Glu Val Thr Gln Leu Lys Arg Tyr Asp Ala Arg Pro Ile Asn
1          5          10          15
40 tgg ggc aaa tgg ttt ctg att ggc atc ggg atg ctg gtt tcg gcg ttc 96
Trp Gly Lys Trp Phe Leu Ile Gly Ile Gly Met Leu Val Ser Ala Phe
20          25          30

atc ctg ctg gtg ccg atg att tac atc ttc gtg cag gca ttc agc aag 144
Ile Leu Leu Val Pro Met Ile Tyr Ile Phe Val Gln Ala Phe Ser Lys
35          40          45

ggg ctg atg ccg gtt tta cag aat ctg gcc gat ccg gac atg ctg cac 192
Gly Leu Met Pro Val Leu Gln Asn Leu Ala Asp Pro Asp Met Leu His
50          55          60

50

```


RU 2 275 425 C2

	gcc atc tgg ctg acg gtg atg atc gcg ctg att gcc gta ccg gta aac	240
	Ala Ile Trp Leu Thr Val Met Ile Ala Leu Ile Ala Val Pro Val Asn	
	65 70 75 80	
5	ctg gtg ttc ggc att ctg ctg gcc tgg ctg gtg acg cgc ttt aac ttc	288
	Leu Val Phe Gly Ile Leu Leu Ala Trp Leu Val Thr Arg Phe Asn Phe	
	85 90 95	
	cct gga cgc cag tta ctg ctg acg cta ctg gac att ccg ttt gcc gta	336
	Pro Gly Arg Gln Leu Leu Leu Thr Leu Leu Asp Ile Pro Phe Ala Val	
	100 105 110	
10	tcg ccg gtg gtt gcc ggt ctg gtg tat ttg ctg ttc tac ggc tct aac	384
	Ser Pro Val Val Ala Gly Leu Val Tyr Leu Leu Phe Tyr Gly Ser Asn	
	115 120 125	
	ggc ccg ctc ggc ggt tgg ctc gac gag cat aac ctg caa att atg ttc	432
	Gly Pro Leu Gly Gly Trp Leu Asp Glu His Asn Leu Gln Ile Met Phe	
15	130 135 140	
	tcc tgg ccg gga atg gtg ctg gtc acc atc ttc gtg acg tgt ccg ttt	480
	Ser Trp Pro Gly Met Val Leu Val Thr Ile Phe Val Thr Cys Pro Phe	
	145 150 155 160	
20	gtg gtg cgc gaa ctg gtg ccg gtg atg tta agc cag ggc agc cag gaa	528
	Val Val Arg Glu Leu Val Pro Val Met Leu Ser Gln Gly Ser Gln Glu	
	165 170 175	
	gac gaa gcg gcg att ttg ctt ggc gcg tcc ggc tgg cag atg ttc cgt	576
	Asp Glu Ala Ala Ile Leu Leu Gly Ala Ser Gly Trp Gln Met Phe Arg	
	180 185 190	
25	cgc gtc aca tta ccg aac atc cgc tgg gcg ctg ctt tat ggc gtg gtg	624
	Arg Val Thr Leu Pro Asn Ile Arg Trp Ala Leu Leu Tyr Gly Val Val	
	195 200 205	
	ttg acc aac gcc ccg gca att ggc gag ttt ggc gcg gtg tcg gtg gtt	672
	Leu Thr Asn Ala Arg Ala Ile Gly Glu Phe Gly Ala Val Ser Val Val	
30	210 215 220	
	tcc ggc tcg att ccg ggc gaa acc ctg tcg ctg ccg tta cag att gaa	720
	Ser Gly Ser Ile Arg Gly Glu Thr Leu Ser Leu Pro Leu Gln Ile Glu	
	225 230 235 240	
35	ttg ctg gag cag gac tac aac acc gtc ggc tcc ttt acc gct gcg gcg	768
	Leu Leu Glu Gln Asp Tyr Asn Thr Val Gly Ser Phe Thr Ala Ala Ala	
	245 250 255	
	ctg tta acg ctg atg gcg att atc acc ctg ttt tta aaa agt atg ttg	816
	Leu Leu Thr Leu Met Ala Ile Ile Thr Leu Phe Leu Lys Ser Met Leu	
	260 265 270	
40	cag tgg cgc ctg gag aat cag gaa aaa cgc gca cag cag gag gaa cat	864
	Gln Trp Arg Leu Glu Asn Gln Glu Lys Arg Ala Gln Gln Glu Glu His	
	275 280 285	
45	cat gag cat tga	876
	His Glu His	
	290	
	<210> 6	
	<211> 291	
	<212> PRT	
50		

<213> Escherichia coli

<400> 6

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

<210> 7

35 <211> 1098
 <212> DNA
 <213> Escherichia coli

<220>

<221> CDS

40 <222> (1)..(1098)

<400> 7

atg agc att gag att gcc aat att aag aag tcg ttt ggt cgc acc cag 48
 Met Ser Ile Glu Ile Ala Asn Ile Lys Lys Ser Phe Gly Arg Thr Gln
 1 5 10 15
 45 gtg ctg aac gat atc tca ctg gat att cct tca ggt cag atg gtc gcg 96
 Val Leu Asn Asp Ile Ser Leu Asp Ile Pro Ser Gly Gln Met Val Ala
 20 25 30
 ttg ctg ggg cgg tcc ggt tcc ggg aaa acc acg ctg ctg cgc att atc 144
 Leu Leu Gly Pro Ser Gly Ser Gly Lys Thr Thr Leu Leu Arg Ile Ile
 50

RU 2 275 425 C2

	35	40	45	
	gcc ggg ctg gag cat caa acc agc ggg cat att cgc ttc cac ggc acc			192
	Ala Gly Leu Glu His Gln Thr Ser Gly His Ile Arg Phe His Gly Thr			
	50	55	60	
5				
	gac gtg agc cgc ctg cac gca cgt gat cgt aaa gtc ggt ttc gtg ttc			240
	Asp Val Ser Arg Leu His Ala Arg Asp Arg Lys Val Gly Phe Val Phe			
	65	70	75 80	
10				
	cag cat tac gcg ctg ttc cgc cat atg acg gtg ttc gac aat atc gct			288
	Gln His Tyr Ala Leu Phe Arg His Met Thr Val Phe Asp Asn Ile Ala			
	85	90	95	
	ttt ggc ctg acg gtg ctg ccg cgt cgc gag cgc ccg aat gcc gca gcc			336
	Phe Gly Leu Thr Val Leu Pro Arg Arg Glu Arg Pro Asn Ala Ala Ala			
	100	105	110	
15				
	atc aaa gcg aaa gtg aca aaa ttg ctg gaa atg gtc cag ctt gcc cat			384
	Ile Lys Ala Lys Val Thr Lys Leu Leu Glu Met Val Gln Leu Ala His			
	115	120	125	
20				
	ctg gcg gat cgt tat ccg gcg cag ctt tcc ggc ggc cag aaa cag cgc			432
	Leu Ala Asp Arg Tyr Pro Ala Gln Leu Ser Gly Gly Gln Lys Gln Arg			
	130	135	140	
	gtg gcg ctg gcg cgc gcg ctg gct gtg gaa ccg caa att ctg ctg ctt			480
	Val Ala Leu Ala Arg Ala Leu Ala Val Glu Pro Gln Ile Leu Leu Leu			
	145	150	155 160	
25				
	gat gaa ccg ttt ggc gcg ctg gat gcg cag gtg cgt aaa gag ctg cgt			528
	Asp Glu Pro Phe Gly Ala Leu Asp Ala Gln Val Arg Lys Glu Leu Arg			
	165	170	175	
	cgc tgg ctg cgt caa ctc cat gaa gaa cta aaa ttc acc agc gtt ttt			576
	Arg Trp Leu Arg Gln Leu His Glu Glu Leu Lys Phe Thr Ser Val Phe			
	180	185	190	
30				
	gtg acc cac gat cag gaa gaa gcg acc gaa gta gct gat cgt gta gtt			624
	Val Thr His Asp Gln Glu Glu Ala Thr Glu Val Ala Asp Arg Val Val			
	195	200	205	
	gtg atg agc cag ggc aat att gaa cag gct gac gcg ccg gat cag gta			672
	Val Met Ser Gln Gly Asn Ile Glu Gln Ala Asp Ala Pro Asp Gln Val			
	210	215	220	
35				
	tgg cgc gaa ccg gcg acc cgt ttt gtg ctc gaa ttt atg ggc gaa gtg			720
	Trp Arg Glu Pro Ala Thr Arg Phe Val Leu Glu Phe Met Gly Glu Val			
	225	230	235 240	
	aac cgc ctg cag gga acc att cgc ggc ggg cag ttc cat gtt gcc gcg			768
	Asn Arg Leu Gln Gly Thr Ile Arg Gly Gly Gln Phe His Val Gly Ala			
	245	250	255	
40				
	cat cgc tgg ccg ctg ggc tac aca cct gcg tat cag ggg ccg gtg gat			816
	His Arg Trp Pro Leu Gly Tyr Thr Pro Ala Tyr Gln Gly Pro Val Asp			
	260	265	270	
45				
	ctc ttc ctg cgc cct tgg gaa gtg gat atc agc cgc cgt acc agc ctc			864
	Leu Phe Leu Arg Pro Trp Glu Val Asp Ile Ser Arg Arg Thr Ser Leu			
	275	280	285	
	gat tcg ccg ctg ccg gta cag gta ctg gaa gcc agc ccg aaa ggt cac			912
	Asp Ser Pro Leu Pro Val Gln Val Leu Glu Ala Ser Pro Lys Gly His			
50				

RU 2 275 425 C2

	290		295		300															
	tac	acc	caa	tta	gtg	gtg	cag	ccg	ctg	ggg	tgg	tac	aac	gaa	ccg	ctg	960			
	Tyr	Thr	Gln	Leu	Val	Val	Gln	Pro	Leu	Gly	Trp	Tyr	Asn	Glu	Pro	Leu				
	305					310					315					320				
5	acg	gtc	gtg	atg	cat	ggc	gac	gat	gcc	ccg	cag	cgt	ggc	gag	cgt	tta	1008			
	Thr	Val	Val	Met	His	Gly	Asp	Asp	Ala	Pro	Gln	Arg	Gly	Glu	Arg	Leu				
					325					330						335				
	ttc	ggt	ggt	ctg	caa	cat	gcg	cgg	ctg	tat	aac	ggc	gac	gag	cgt	atc	1056			
	Phe	Val	Gly	Leu	Gln	His	Ala	Arg	Leu	Tyr	Asn	Gly	Asp	Glu	Arg	Ile				
10				340					345						350					
	gaa	acc	cgc	gat	gag	gaa	ctt	gct	ctc	gca	caa	agc	gcc	tga	1098					
	Glu	Thr	Arg	Asp	Glu	Glu	Leu	Ala	Leu	Ala	Gln	Ser	Ala							
			355					360					365							
15	<210>	8	<211>	365	<212>	PRT	<213>	Escherichia coli												
	<400>	8																		
20	Met	Ser	Ile	Glu	Ile	Ala	Asn	Ile	Lys	Lys	Ser	Phe	Gly	Arg	Thr	Gln				
	1				5					10					15					
	Val	Leu	Asn	Asp	Ile	Ser	Leu	Asp	Ile	Pro	Ser	Gly	Gln	Met	Val	Ala				
			20						25					30						
	Leu	Leu	Gly	Pro	Ser	Gly	Ser	Gly	Lys	Thr	Thr	Leu	Leu	Arg	Ile	Ile				
			35					40					45							
25	Ala	Gly	Leu	Glu	His	Gln	Thr	Ser	Gly	His	Ile	Arg	Phe	His	Gly	Thr				
	50					55						60								
	Asp	Val	Ser	Arg	Leu	His	Ala	Arg	Asp	Arg	Lys	Val	Gly	Phe	Val	Phe				
	65				70						75					80				
	Gln	His	Tyr	Ala	Leu	Phe	Arg	His	Met	Thr	Val	Phe	Asp	Asn	Ile	Ala				
					85						90					95				
	Phe	Gly	Leu	Thr	Val	Leu	Pro	Arg	Arg	Glu	Arg	Pro	Asn	Ala	Ala	Ala				
				100						105					110					
30	Ile	Lys	Ala	Lys	Val	Thr	Lys	Leu	Leu	Glu	Met	Val	Gln	Leu	Ala	His				
				115						120					125					
	Leu	Ala	Asp	Arg	Tyr	Pro	Ala	Gln	Leu	Ser	Gly	Gly	Gln	Lys	Gln	Arg				
				130						135					140					
	Val	Ala	Leu	Ala	Arg	Ala	Leu	Ala	Val	Glu	Pro	Gln	Ile	Leu	Leu	Leu				
	145				150						155					160				
35	Asp	Glu	Pro	Phe	Gly	Ala	Leu	Asp	Ala	Gln	Val	Arg	Lys	Glu	Leu	Arg				
					165						170					175				
	Arg	Trp	Leu	Arg	Gln	Leu	His	Glu	Glu	Leu	Lys	Phe	Thr	Ser	Val	Phe				
					180						185					190				
	Val	Thr	His	Asp	Gln	Glu	Glu	Ala	Thr	Glu	Val	Ala	Asp	Arg	Val	Val				
				195							200					205				
	Val	Met	Ser	Gln	Gly	Asn	Ile	Glu	Gln	Ala	Asp	Ala	Pro	Asp	Gln	Val				
				210							215				220					
	Trp	Arg	Glu	Pro	Ala	Thr	Arg	Phe	Val	Leu	Glu	Phe	Met	Gly	Glu	Val				
						230						235				240				
	Asn	Arg	Leu	Gln	Gly	Thr	Ile	Arg	Gly	Gly	Gln	Phe	His	Val	Gly	Ala				
					245						250					255				
	His	Arg	Trp	Pro	Leu	Gly	Tyr	Thr	Pro	Ala	Tyr	Gln	Gly	Pro	Val	Asp				
45					260						265					270				
	Leu	Phe	Leu	Arg	Pro	Trp	Glu	Val	Asp	Ile	Ser	Arg	Arg	Thr	Ser	Leu				
											280					285				
	Asp	Ser	Pro	Leu	Pro	Val	Gln	Val	Leu	Glu	Ala	Ser	Pro	Lys	Gly	His				
												300								
	290							295												
	Tyr	Thr	Gln	Leu	Val	Val	Gln	Pro	Leu	Gly	Trp	Tyr	Asn	Glu	Pro	Leu				
50																				

RU 2 275 425 C2

```

305          310          315          320
Thr Val Val Met His Gly Asp Asp Ala Pro Gln Arg Gly Glu Arg Leu
          325          330          335
Phe Val Gly Leu Gln His Ala Arg Leu Tyr Asn Gly Asp Glu Arg Ile
          340          345          350
5  Glu Thr Arg Asp Glu Glu Leu Ala Leu Ala Gln Ser Ala
          355          360          365

<210> 9
<211> 912
10 <212> DNA
    <213> Escherichia coli

<220>
<221> CDS
<222> (1)..(912)

15 <400> 9
    gtg agt aca tta gaa caa aca ata ggc aat acg cct ctg gtg aag ttg 48
    Val Ser Thr Leu Glu Gln Thr Ile Gly Asn Thr Pro Leu Val Lys Leu
        1          5          10          15

    cag cga atg ggg ccg gat aac ggc agt gaa gtg tgg tta aaa ctg gaa 96
    Gln Arg Met Gly Pro Asp Asn Gly Ser Glu Val Trp Leu Lys Leu Glu
        20          25          30

    ggc aat aac ccg gca ggt tcg gtg aaa gat cgt gcg gca ctt tcg atg 144
    Gly Asn Asn Pro Ala Gly Ser Val Lys Asp Arg Ala Ala Leu Ser Met
        35          40          45

25  atc gtc gag gcg gaa aag cgc ggg gaa att aaa ccg ggt gat gtc tta 192
    Ile Val Glu Ala Glu Lys Arg Gly Glu Ile Lys Pro Gly Asp Val Leu
        50          55          60

    atc gaa gcc acc agt ggt aac acc ggc att gcg ctg gca atg att gcc 240
    Ile Glu Ala Thr Ser Gly Asn Thr Gly Ile Ala Leu Ala Met Ile Ala
30  65          70          75          80

    gcg ctg aaa ggc tat cgc atg aaa ttg ctg atg ccc gac aac atg agc 288
    Ala Leu Lys Gly Tyr Arg Met Lys Leu Leu Met Pro Asp Asn Met Ser
        85          90          95

35  cag gaa cgc cgt gcg gcg atg cgt gct tat ggt gcg gaa ctg att ctt 336
    Gln Glu Arg Arg Ala Ala Met Arg Ala Tyr Gly Ala Glu Leu Ile Leu
        100          105          110

    gtc acc aaa gag cag ggc atg gaa ggt gcg cgc gat ctg gcg ctg gag 384
    Val Thr Lys Glu Gln Gly Met Glu Gly Ala Arg Asp Leu Ala Leu Glu
        115          120          125

40  atg gcg aat cgt ggc gaa gga aag ctg ctc gat cag ttc aat aat ccc 432
    Met Ala Asn Arg Gly Glu Gly Lys Leu Leu Asp Gln Phe Asn Asn Pro
        130          135          140

    gat aac cct tat gcg cat tac acc acc act ggg ccg gaa atc tgg cag 480
    Asp Asn Pro Tyr Ala His Tyr Thr Thr Thr Gly Pro Glu Ile Trp Gln
45  145          150          155          160

    caa acc ggc ggg cgc atc act cat ttt gtc tcc agc atg ggg acg acc 528
    Gln Thr Gly Gly Arg Ile Thr His Phe Val Ser Ser Met Gly Thr Thr
        165          170          175

```

50

RU 2 275 425 C2

ggc act atc acc ggc gtc tca cgc ttt atg cgc gaa caa tcc aaa ccg 576
 Gly Thr Ile Thr Gly Val Ser Arg Phe Met Arg Glu Gln Ser Lys Pro
 180 185 190

5 gtg acc att gtc ggc ctg caa ccg gaa gag ggc agc agc att ccc ggc 624
 Val Thr Ile Val Gly Leu Gln Pro Glu Glu Gly Ser Ser Ile Pro Gly
 195 200 205

att cgc cgc tgg cct acg gaa tat ctg ccg ggg att ttc aac gct tct 672
 Ile Arg Arg Trp Pro Thr Glu Tyr Leu Pro Gly Ile Phe Asn Ala Ser
 210 215 220

10 ctg gtg gat gag gtg ctg gat att cat cag cgc gat gcg gaa aac acc 720
 Leu Val Asp Glu Val Leu Asp Ile His Gln Arg Asp Ala Glu Asn Thr
 225 230 235 240

atg cgc gaa ctg gcg gtg cgg gaa gga ata ttc tgt ggc gtc agc tcc 768
 Met Arg Glu Leu Ala Val Arg Glu Gly Ile Phe Cys Gly Val Ser Ser
 245 250 255

15 ggc ggc gcg gtt gcc gga gca ctg cgg gtg gca aaa gct aac cct gac 816
 Gly Gly Ala Val Ala Gly Ala Leu Arg Val Ala Lys Ala Asn Pro Asp
 260 265 270

20 gcg gtg gtg gtg gcg atc atc tgc gat cgt ggc gat cgc tac ctt tct 864
 Ala Val Val Val Ala Ile Ile Cys Asp Arg Gly Asp Arg Tyr Leu Ser
 275 280 285

acc ggg gtg ttt ggg gaa gag cat ttt agc cag ggg gcg ggg att taa 912
 Thr Gly Val Phe Gly Glu Glu His Phe Ser Gln Gly Ala Gly Ile
 290 295 300

25
 <210> 10
 <211> 303
 <212> PRT
 <213> Escherichia coli

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

35 Ile Val Glu Ala Glu Lys Arg Gly Glu Ile Lys Pro Gly Asp Val Leu
 50 55 60
 Ile Glu Ala Thr Ser Gly Asn Thr Gly Ile Ala Leu Ala Met Ile Ala
 65 70 75 80
 Ala Leu Lys Gly Tyr Arg Met Lys Leu Leu Met Pro Asp Asn Met Ser
 85 90 95

40 Gln Glu Arg Arg Ala Ala Met Arg Ala Tyr Gly Ala Glu Leu Ile Leu
 100 105 110
 Val Thr Lys Glu Gln Gly Met Glu Gly Ala Arg Asp Leu Ala Leu Glu
 115 120 125
 Met Ala Asn Arg Gly Glu Gly Lys Leu Leu Asp Gln Phe Asn Asn Pro
 130 135 140

45 Asp Asn Pro Tyr Ala His Tyr Thr Thr Thr Gly Pro Glu Ile Trp Gln
 145 150 155 160
 Gln Thr Gly Gly Arg Ile Thr His Phe Val Ser Ser Met Gly Thr Thr
 165 170 175
 Gly Thr Ile Thr Gly Val Ser Arg Phe Met Arg Glu Gln Ser Lys Pro
 180 185 190
 Val Thr Ile Val Gly Leu Gln Pro Glu Glu Gly Ser Ser Ile Pro Gly

50

RU 2 275 425 C2

```

      195                200                205
Ile Arg Arg Trp Pro Thr Glu Tyr Leu Pro Gly Ile Phe Asn Ala Ser
  210                215                220
Leu Val Asp Glu Val Leu Asp Ile His Gln Arg Asp Ala Glu Asn Thr
  225                230                235                240
5 Met Arg Glu Leu Ala Val Arg Glu Gly Ile Phe Cys Gly Val Ser Ser
      245                250                255
Gly Gly Ala Val Ala Gly Ala Leu Arg Val Ala Lys Ala Asn Pro Asp
      260                265                270
Ala Val Val Val Ala Ile Ile Cys Asp Arg Gly Asp Arg Tyr Leu Ser
      275                280                285
10 Thr Gly Val Phe Gly Glu Glu His Phe Ser Gln Gly Ala Gly Ile
      290                295                300

```

```

<210> 11
<211> 34
15 <212> DNA
    <213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 11
20 agctgagtcg accgcctcgt tatcatccaa aatc                                     34

<210> 12
<211> 33
<212> DNA
25 <213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 12
30 agctgagcat gcactaattt tccttgcgga ggc                                     33

<210> 13
<211> 32
<212> DNA
35 <213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 13
    agctgagcat gctgatggcg gcagcacact gc                                     32

40 <210> 14
    <211> 33
    <212> DNA
    <213> Artificial Sequence

45 <220>
    <223> Description of Artificial Sequence: primer

    <400> 14
        agctgatcta gattcactca acctatcagg cgc                                     33

50

```

<210> 15
 <211> 35
 <212> DNA
 <213> Artificial Sequence

5 <220>
 <223> Description of Artificial Sequence: primer

<400> 15
 agctgagtcg acgtgagtac attagaacaa acaat 35

10 <210> 16
 <211> 33
 <212> DNA
 <213> Artificial Sequence

15 <220>
 <223> Description of Artificial Sequence: primer

<400> 16
 agctgatcta gaagtctccg atgctattaa tcc 33

20 <210> 17
 <211> 34
 <212> DNA
 <213> Artificial Sequence

25 <220>
 <223> Description of Artificial Sequence: primer

<400> 17
 agctgagtcg actccttaac tgtatgaaat tggg 34

30 <210> 18
 <211> 33
 <212> DNA
 <213> Artificial Sequence

35 <220>
 <223> Description of Artificial Sequence: primer

<400> 18
 agctgagcat gccagcctg tttacgatga tcc 33

40

Формула изобретения

1. Бактерия, принадлежащая к роду *Escherichia*, - продуцент L-цистеина, модифицированная таким образом, что экспрессия генов кластера *cysPTWAM* усилена.

45 2. Бактерия по п.1, отличающаяся тем, что экспрессия генов кластера *cysPTWAM* усилена за счет увеличения количества копий генов кластера *cysPTWAM* или за счет модификации последовательности контроля экспрессии таким образом, что в результате такой модификации экспрессия этих генов усиливается.

3. Бактерия по п.2, отличающаяся тем, что количество копий указанных генов увеличено в результате трансформации указанной бактерии многокопийным вектором, содержащим гены кластера *cysPTWAM*.

50 4. Бактерия по п.3, отличающаяся тем, что природный промотор указанного кластера заменен на более сильный промотор.

5. Бактерия по п.4, отличающаяся тем, что количество копий указанных генов

увеличено путем интеграции дополнительных копий генов кластера *cysPTWAM* в хромосому указанной бактерии.

6. Бактерия по п.5, отличающаяся тем, что гены кластера *cysPTWAM* получены из бактерии, принадлежащей роду *Escherichia*.

5 7. Бактерия по п.6, отличающаяся тем, что бактерия дополнительно модифицирована таким образом, что экспрессия открытой рамки считывания *udeD* усилена.

8. Способ получения L-цистеина, включающий стадии выращивания бактерии по любому из пп.1-7, в питательной среде, содержащей тиосульфат, и выделения L-цистеина из культуральной жидкости.

10 9. Способ по п.8, отличающийся тем, что у указанной бактерии усилена экспрессия генов биосинтеза L-цистеина.

15

20

25

30

35

40

45

50