

(19)



**Евразийское
патентное
ведомство**

(11) **035117**

(13) **B1**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ

(45) Дата публикации и выдачи патента
2020.04.29

(21) Номер заявки
201791118

(22) Дата подачи заявки
2015.11.17

(51) Int. Cl. **C07K 14/34** (2006.01)
C12N 15/10 (2006.01)
C12N 15/70 (2006.01)

(54) КОДОН-ОПТИМИЗИРОВАННЫЙ ПОЛИНУКЛЕОТИД ДЛЯ ЭКСПРЕССИИ CRM₁₉₇ НА ВЫСОКОМ УРОВНЕ

(31) 4045/CHE/2014

(32) 2014.11.20

(33) IN

(43) 2017.10.31

(86) PCT/IN2015/000427

(87) WO 2016/079755 2016.05.26

(71)(73) Заявитель и патентовладелец:
БАЙОЛОДЖИКАЛ И ЛИМИТЕД (IN)

(72) Изобретатель:
**Акшай Гоэл, Рави Пратар Нараян
Мишра, Нарендер Дев Мантена,
Махима Датла (IN)**

(74) Представитель:
Медведев В.Н. (RU)

(56) CN-A-103266125

ALESSANDRA STEFAN ET AL.:
"Overexpression and purification of the recombinant diphtheria toxin variant CRM197 in", JOURNAL OF BIOTECHNOLOGY, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, NL, vol. 156, no. 4, 15 August 2011 (2011-08-15), pages 245-252, XP028119257, ISSN: 0168-1656, DOI: 10.1016/J.J BIOTEC.2011.08.024 [retrieved on 2011-08-25] abstract page 246 - page 248

(57) Изобретение относится к экспрессии на высоком уровне бактериального токсоида или белка-токсина, представляющего фармакологический интерес, посредством оптимизированной новой полинуклеотидной последовательности, и к хозяину, трансформированному указанным полинуклеотидом. В частности, изобретение относится к способу высокой продукции полипептида CRM₁₉₇, где полинуклеотид по изобретению используют для трансформации подходящего хозяина, что приводит к сверхэкспрессии соответствующих белков, и к способу выделения экспрессируемого полипептида. Более конкретно, настоящее изобретение относится к экспрессии CRM₁₉₇ на высоком уровне Escherichia coli и к способу его выделения и очистки.

B1

035117

035117

B1

Область изобретения

Изобретение относится к экспрессии на высоком уровне бактериального токсина с помощью новой оптимизированной полинуклеотидной последовательности и к хозяину, трансформированному указанным полинуклеотидом.

Изобретение также относится к способу продукции на высоком уровне полипептида CRM₁₉₇, где полинуклеотид по настоящему изобретению используют для трансформации подходящего хозяина, что приводит к сверхэкспрессии соответствующих белков, и к способу выделения экспрессируемого полипептида.

Уровень техники, к которому относится изобретение

Дифтерийный токсин (DT) представляет собой белковый экзотоксин, который синтезируется и секретируется *Corynebacterium diphtheriae*. Токсигенные штаммы *Corynebacterium diphtheriae* содержат ген токсина, включающий бактериофагальный лизоген. Зрелая форма DT синтезируется в качестве содержащего 535 аминокислот единого полипептида, который происходит из исходного пропептида 536, который претерпевает протеолиз в положениях 190, 192 и 193 с образованием зрелого токсина. Этот сплайсинг или протеолиз приводит к двум субъединицам, А и В, которые связаны вместе дисульфидным мостиком (Moskang et al., Biol. Chem. 264: 15709-15713, 1989). Субъединица А представляет собой каталитически активную часть NAD-зависимой ADP-рибозилтрансферазы. Она ответственна за обеспечение инактивации фактора элонгации 2 (EF-2), и, таким образом, подавляет синтез белка в клетке-мишени.

Дифтерийный токсин является в высокой степени цитотоксичным; одна молекула может быть летальной для клетки, и доза 10 нг/кг может привести к гибели животных и человека. В процесс предоставления искусственного иммунитета посредством инъекции этого токсина должен быть включен процесс детоксификации, чтобы сделать его безопасным для употребления реципиентом. Обычно его детоксифицируют посредством химической модификации природных форм DT таким образом, чтобы он все еще сохранял антигенность, требуемую для вакцинного препарата.

Затем была внедрена генетически детоксифицированная форма дифтерийного токсина, известная как Cross Reacting Material 197 или CRM₁₉₇; которая по существу сохраняет иммуногенные свойства перекрестного реагирования DT.

CRM₁₉₇ используют для получения конъюгатных вакцин, включая вакцины против *Corynebacterium diphtheria*, гепатита В, *Bordetella pertussis*, *Clostridium tetani*, *Neisseria meningitidis*, *Streptococcus pneumoniae*, *Haemophilus influenzae*. Он был получен посредством мутагенеза с использованием нитрозогуанидина токсигенного коринефага β, который затем использовали для инфицирования *Corynebacterium diphtheria*. (Uchida et al., Nature New Biology (1971) 233; 8-11, Nucleic Acids Res. 1984 May 25; 12 (10) : 4063-9).

CRM₁₉₇ исследовали в отношении его потенциального применения в качестве усилителя DT или вакцинного антигена. Белок CRM₁₉₇ имеет ту же молекулярную массу, что и DT; но отличается заменой одного основания в субъединице А т.е. заменой основания в полинуклеотидной последовательности DT дикого типа, где замена гуанина на аденин приводит аминокислотной замене в положении 52, что приводит к глутаминовой кислоте в CRM₁₉₇ вместо глицина (Giannini G. et al., 1984). Эта точечная мутация приводит к значительному снижению токсичности и делает CRM₁₉₇ безопасным для применения у человека.

Продукция значительных количеств дифтерийного токсина, такого как CRM₁₉₇, для применения в вакцинах, была затруднена вследствие низкого уровня экспрессии в бактериях дикого типа. Эта проблема ранее была решена Bishai et al., (J. Bacteriol. 189:5140-5151) посредством экспрессии CRM₁₉₇ в *Escherichia coli*, которые описали экспрессию рекомбинантного слитого белка, содержащего дифтерийный токсин (включая сигнальную последовательность tox); однако это привело к продукции деградировавшего белка. Низкий выход в активной форме также ассоциирован с деградацией, ненадлежащей укладкой или обоими из них, в зависимости от конкретных характеристик, например, размера и вторичной структуры, токсина. Таким образом, как и в случае большинства биофармацевтических средств, существует дополнительная потеря экспрессируемого белка, которая происходит в ходе стадий очистки CRM₁₉₇, причем поддержание биологически активной формы CRM₁₉₇ является трудным. Таким образом, существует потребность в достижении высокого уровня экспрессии бактериального токсина CRM₁₉₇ в активной форме.

В WO 2011/042516 описан способ получения бактериального токсина посредством периплазматической экспрессии, включающий стадии а) выращивания культуры бактериальных клеток-хозяев, содержащих экспрессирующий вектор, в котором конкретная сигнальная последовательность связана с последовательностью бактериального токсина, и б) индукции экспрессии полипептида, содержащего конкретную сигнальную последовательность, связанную с бактериальным токсином, чтобы бактериальный токсин экспрессировался периплазматически.

В WO 2013/178974 A1 описан способ внутриклеточной экспрессии CRM₁₉₇ у хозяина *Escherichia coli*, содержащего экспрессирующий вектор, содержащий ген, кодирующий CRM₁₉₇, функционально связанный с промотором, и по меньшей мере одну полностью палиндромную последовательность оператора.

В WO 2015/134402 A1 описан способ получения рекомбинантного CRM₁₉₇ в хозяине *Escherichia coli*

с уменьшенным геномом, включающий инкубацию *Escherichia coli* с уменьшенным геномом, содержащих экспрессирующий вектор, содержащий нуклеотидную последовательность, кодирующую белок CRM₁₉₇, слитый с сигнальной последовательностью, которая контролирует перенос белка CRM₁₉₇ в периплазму, функционально связанную с последовательностью контроля экспрессии, в условиях, подходящих для экспрессии рекомбинантного белка CRM₁₉₇, где достигается выход по меньшей мере 1 г растворимого CRM на литр и где нативный родительский штамм *Escherichia coli* представляет собой штамм 12, предпочтительно K12 MG1655.

В US 2012/0128727 A1 описана выделенная молекула нуклеиновой кислоты, которая кодирует полипептид CRM₁₉₇, экспрессирующий вектор, содержащий выделенную молекулу нуклеиновой кислоты, и способ рекомбинантной продукции слитого белка CRM₁₉₇-метка, включающий культивирование рекомбинантной клетки в условиях, способствующих продукции указанного слитого белка CRM₁₉₇-метка, и выделение указанного слитого белка.

В US 2012/0289688 A1 описан способ периплазматической экспрессии рекомбинантного полипептида путем (A) выращивания культуры грамотрицательных клеток-хозяев; и (B) индукции экспрессии полипептида, чтобы белок экспрессировался периплазматически; где в ходе экспрессии выполняется одно или несколько из следующих условий: (i) pH на стадии a) ниже, чем pH на стадии b); (ii) температура на стадии a) выше, чем температура на стадии b); или (iii) скорость подачи субстрата на стадии a) превышает скорость подачи субстрата на стадии b).

В US 2014/0050758 A1 описан способ периплазматической экспрессии бактериального токсоида, включающий стадии:

a) выращивания культуры грамотрицательных клеток-хозяев в среде для ферментации, где клетку-хозяина трансформируют полинуклеотидом и где полинуклеотид кодирует бактериальный токсид и периплазматическую сигнальную последовательность; индукции экспрессии бактериального токсоида;

b) обеспечения созревания клетки-хозяина, где стадия обеспечения созревания включает: I) воздействие на клетку-хозяина посредством pH-шока; II) инкубацию клетки-хозяина без добавления питательных веществ; или III) воздействие на клетку-хозяина температуры ниже -20°C; и

c) экстракции бактериального токсоида из клетки-хозяина, где процесс экстракции включает осмотический шок, где грамотрицательная клетка-хозяин выбрана из группы, состоящей из *Escherichia coli*, *Pseudomonas* и *Moraxella*, где клетка-хозяин является живой на стадии b) и где способ осуществляют в ферментере, который содержит 10-5000 л культуры.

В US 8530171 описан способ получения рекомбинантного белка-токсина в клетке-хозяине *Pseudomonas*, причем указанный способ включает лигирование в экспрессирующий вектор нуклеотидной последовательности, кодирующей белок-токсин; трансформацию клетки-хозяина *Pseudomonas* экспрессирующим вектором; и культивирование трансформированной клетки-хозяина *Pseudomonas* в культуральной среде, пригодной для экспрессии рекомбинантного белка-токсина; где рекомбинантный белок-переносчик представляет собой CRM₁₉₇, и где рекомбинантный белок продуцируется с выходом растворимого или активного белка CRM₁₉₇ от приблизительно 0,2 г/л до приблизительно 12 г/л.

Конъюгированные полисахаридные вакцины, в которых используется CRM₁₉₇ в качестве белка-носителя, одобрены для применения у человека. Они включают: Menveo® (Novartis Vaccines and Diagnostics), представляющую собой вакцину, показанную для предупреждения инвазивного менингококкового заболевания, вызываемого *Neisseria meningitidis* подгрупп А, С, Y и W-135; Menjugate® (Novartis Vaccines), представляющую собой конъюгатную вакцину против менингококка группы С; и Prevnar® (Wyeth Pharmaceuticals, Inc.), представляющую собой вакцину против детской пневмонии, которая нацелена на тридцать серотипов *Streptococcus pneumoniae*, и HibTITER® (Wyeth), представляющую собой вакцину против *Haemophilus influenzae* типа b. Кроме того, CRM₁₉₇ имеет потенциальное применение в качестве усиливающего антигена для вакцинации против *C. diphtheria*, и его исследуют в качестве белка-носителя для применения в других вакцинах.

Недавно возрос интерес к CRM₁₉₇ вследствие его потенциального противоопухолевого действия, связанного с его способностью связывать растворимую форму HB-EGF (Mekada et al., публикация патента США № 2006/0270600A1). Эта противоопухолевая функция свойственна не только CRM₁₉₇, но также и другим нетоксичным производным токсина DT (например, двойной мутант DT52E148K или слитый белок GST-DT). Эти мутанты были сконструированы с использованием ПЦР, начиная с гена, кодирующего CRM₁₉₇. Однако в указанных исследованиях целый CRM₁₉₇ был продуцирован с использованием культур *C. diphtheria*, выращиваемых при 35°C в течение 16-17 ч. CRM₁₉₇ очищали из супернатанта посредством первоначальной преципитации с сульфатом аммония, за которой следовали три последовательных стадии ионообменной и гидрофобной хроматографии (Mekada et al.).

Таким образом, очевидна потребность в альтернативном способе получения CRM₁₉₇ с высоким выходом и экономичным образом. Таким образом, способ экономичной продукции CRM₁₉₇ может значительно облегчить исследование, разработку и производство вакцин.

Задача изобретения

Главной задачей настоящего изобретения является предоставление оптимизированного полинук-

леотида для экспрессии CRM₁₉₇ на высоком уровне.

Другой задачей является предоставление регулируемого способа контроля экспрессии полипептида для получения CRM₁₉₇.

Другой задачей является предоставление способа экспрессии на высоком уровне для коммерческой продукции CRM₁₉₇ в чистой форме с высоким выходом.

Сущность изобретения

Настоящее изобретение относится оптимизированной полинуклеотидной последовательности, содержащей SEQ ID NO: 2 и ее варианты, которые по меньшей мере на 70% гомологичны указанной оптимизированной полинуклеотидной последовательности SEQ ID NO: 2.

В другом варианте осуществления настоящее изобретение относится оптимизированной полинуклеотидной последовательности (SEQ ID NO: 2) и ее структурным вариантам, выбранным из, но не ограничиваясь ими, SEQ ID NO. 3, 4, 5, 6, 7, 8, 9 и 10, пригодным для экспрессии полипептида CRM₁₉₇ на высоком уровне.

В другом варианте осуществления настоящее изобретение относится к оптимизированной полинуклеотидной последовательности (SEQ ID NO: 2) и ее вариантам, таким как SEQ ID NO: 3, 4, 5, 6, 7, 8, 9 и 10, которые по меньшей мере на 70-88% гомологичны указанной оптимизированной полинуклеотидной последовательности SEQ ID NO: 2.

Кроме того, настоящее изобретение относится к способу получения полипептида, включающему стадии:

- a) выбора оптимизированной полинуклеотидной последовательности, по существу состоящей из SEQ ID NO: 2, или ее вариантов, которые по меньшей мере на 70% гомологичны SEQ ID NO: 2,
- b) необязательно лигирования полинуклеотидных последовательностей стадии (a) в подходящий вектор,
- c) встраивания или трансформации полинуклеотидной последовательности в клетку-хозяина *Escherichia coli*,
- d) культивирования трансформированной клетки в культуральной среде для экспрессии полипептида на высоком уровне,
- e) поддержания индуцирующей температуры между 10 и 40°C для продукции полипептида,
- f) экстракции бактериального полипептида из клетки-хозяина с последующей очисткой для получения чистого полипептида с высоким выходом.

Полипептид, полученный, как описано выше, можно использовать в качестве белка-носителя для получения конъюгированных иммуногенных препаратов.

Краткое описание чертежей

Фиг. 1 иллюстрирует электрофоретическое разделение геля SDS-PAGE, соответствующее экспрессируемому CRM₁₉₇, который кодируется полинуклеотидом SEQ ID NO: 2; где дорожка 1: стандартный маркер молекулярной массы; дорожки 2 и 3: CRM₁₉₇, который отделен и очищен от общих клеточных белков в экстрактах культуры *E. coli* и разделен на невосстанавливаемом и восстанавливаемом SDS-PAGE, соответственно.

Фиг. 2 - анализ с использованием вестерн-блоттинга очищенного CRM₁₉₇ с использованием поликлональных антител кролика; дорожка 1: маркер молекулярной массы. Дорожки 2 и 3 включают образцы CRM₁₉₇, подвергнутые электрофорезу в восстанавливающих и невосстанавливающих условиях, соответственно.

Фиг. 3 - SDS PAGE демонстрирует очищенную растворимую фракцию, дорожка 1: эталонный белок; дорожка 2: маркер молекулярной массы белка; дорожка 3-10: объединенный полипептид CRM₁₉₇.

Фиг. 4 - электрофоретическое разделение (SDS PAGE 12%), демонстрирующее исследование, проведенное при солиubilизации CRM₁₉₇, дорожка 1: солиubilизованная мочевиной фракция CRM₁₉₇; дорожка 2: супернатант; дорожка 3: образец первого промывания осадка; дорожка 4: образец, объединенный после 2-го промывания осадка.

Фиг. 5 - эксклюзионная хроматография (SEC-ВЭЖХ), где главный элюируемый пик демонстрирует присутствие CRM₁₉₇ в образце.

Фиг. 6 - отпечаток пептидной массы (масс-спектрометрия) для полипептида CRM₁₉₇ для определения идентичности первичной аминокислотной последовательности. Рекombинантный CRM (BioE rCRM) по настоящему изобретению обладал 100% сходством последовательности с эталонной последовательностью CRM₁₉₇.

Фиг. 7 - подтверждение N-концевой последовательности rCRM₁₉₇ посредством деградации способом Эдмана. Последовательность из 10 аминокислот GADDWDSSK (N-концевой ацетил) демонстрирует начальную часть очищенного полипептида. Первая аминокислота идентифицирована как G.

Фиг. 8 - спектры CD рекомбинантного CRM (BioE rCRM) по настоящему изобретению перекрывались с эталонным CRM₁₉₇. Также анализировали параметр вторичной структуры, и он продемонстрировал сходство с эталоном. Этот результат демонстрирует, что рекомбинантный CRM (BioE rCRM) по настоящему изобретению структурно сходен с эталонным CRM₁₉₇.

Фиг. 9 - подтверждение структурной эквивалентности рекомбинантного CRM (BioE rCRM) по на-

стоящему изобретению с эталоном с использованием флуоресцентного анализа. Наложение профиля DSF рекомбинантного CRM (BioE rCRM) по настоящему изобретению и эталонного CRM₁₉₇ (C7 CRM). Данные подтверждают сходство рекомбинантного CRM (BioE rCRM) по настоящему изобретению с эталонным C7-CRM₁₉₇.

Фиг. 10 - подтверждение дисульфидных связей. Рекомбинантный CRM (BioE rCRM) по настоящему изобретению анализировали в отношении присутствия правильных дисульфидных связей в белке.

Подтверждено, что присутствуют две дисульфидные связи: первая связь связывает аминокислоты 186-201 и вторая связь связывает аминокислоты 461-471. Способ масс-спектрометрии использовали для анализа дисульфидных связей в CRM₁₉₇.

Фиг. 11 - подтверждение антигенного сходства рекомбинантного CRM (BioE rCRM) по настоящему изобретению с эталонным CRM₁₉₇ с использованием ELISA, специфичного к CRM₁₉₇. Все CRM, происходящие из других источников, продемонстрировали сходный профиль распознавания моноклональными антителами.

Подробное описание изобретения

Экспрессия полипептидов для применения в качестве фармацевтических продуктов или вакцин требует достижения высокой биомассы и/или продуктивности линии клеток-хозяев. Эффективность продукции полипептида может быть значительно снижена в отсутствие множества факторов, которые включают использование оптимальной полинуклеотидной последовательности, кодирующей этот полипептид. Известно, что генетический код демонстрирует вырожденность, которая обеспечивает варьирование полинуклеотидных последовательностей, кодирующих одну и ту же аминокислотную последовательность. Скорость синтеза аминокислотной цепи является определяющим фактором конечных уровней экспрессии с индивидуального гена, который влияет на формат экспрессирующей конструкции. Таким образом конструкция оптимальной полинуклеотидной последовательности является важной для определения конечных уровней экспрессии полипептида и должна хорошо регулироваться. Это включает, но не ограничивается ими, частоту, с которой кодоны встречаются в организме или, в случае искусственных носителей или векторов, ближайшую желаемую частоту. Это в свою очередь отражает распространенность тРНК или частоту собственных клеточных тРНК, ввиду чего должны быть тщательно выбраны паттерны выбора синонимических кодонов. Дополнительные факторы также включают потенциал к образованию вторичных структур, уровни мРНК и стабильность РНК, последующие необходимые для проведения манипуляции, пути синтеза и т.д. Встречаемость этих структур должна тщательно регулироваться, поскольку выбор этих паттернов различается при оптимизации индивидуальных представляющих интерес белков и хозяев экспрессии.

Таким, основным вариантом осуществления настоящего изобретения относится к оптимизированной полинуклеотидной последовательности (SEQ ID NO: 2) и ее структурным вариантам.

В другом варианте осуществления изобретение относится к оптимизированной полинуклеотидной последовательности (SEQ ID NO: 2) и ее структурным вариантам, выбранным из, но не ограничиваясь ими, SEQ ID NO: 3, 4, 5, 6, 7, 8, 9 и 10, обладающим 70% или более сходством, пригодным для экспрессии на высоком уровне полипептида CRM₁₉₇.

Периплазматическая экспрессия относится к секреции экспрессируемого продукта (такого как бактериальный токсин или дифтерийный токсин) с заданного представляющего интерес гена в периплазматическое пространство в клетке-хозяине.

Цитоплазматическая экспрессия относится к экспрессии белка в цитоплазматическом компартменте клетки, заключенной в клеточную мембрану.

Индукция экспрессии относится к стадии, проводимой для индукции экспрессии с полинуклеотида, так что продукт образуется с увеличенной скоростью, это может вовлекать добавление подходящего индуктирующего агента, такого как IPTG, арабиноза, мальтоза и т.п.

Оптимизированная последовательность по настоящему изобретению является пригодной для других вариантов SEQ ID NO: 2, выбранных из, но не ограничиваясь ими, SEQ ID NO: 3, 4, 5, 6, 7, 8, 9, 10, и также включает последовательности для продукции производных SEQ ID NO: 1, дифтерийного токсина дикого типа, который сохраняет те же воспалительные и иммуностимулирующие свойства, и способен связываться с клеточным рецептором HB-EGF, но отличается от CRM₁₉₇ одной аминокислотной заменой и лишен клеточной токсичности в отношении хозяина-мишени.

Полинуклеотидная последовательность CRM₁₉₇ может происходить из последовательности дифтерийного токсина (Greenfield, L. et al., 1983, Proc. Natl. Acad. Sci. USA 80:6853-6857), или может быть получена с использованием аминокислотной последовательности CRM₁₉₇, приведенной Giannini G. et al., (1984) в качестве ссылки. Полинуклеотидную последовательность дикого типа, полученную таким образом, оптимизировали для высокой экспрессии в клетке-хозяине, более предпочтительно, в граммотрицательной клетке, более предпочтительно, в *Escherichia coli* в качестве клетки-хозяина. Такую полинуклеотидную последовательность по настоящему изобретению можно получать посредством химического синтеза или с помощью процедуры сборки.

В другом варианте осуществления предусматривается способ внутриклеточной экспрессии CRM₁₉₇ в клетке-хозяине, где экспрессирующая конструкция с регуляторной последовательностью обеспечивает

экспрессию полинуклеотида по изобретению. Эта полинуклеотидная последовательность может быть связана с сигнальной последовательностью для направленного транспорта кодируемого полипептида. Она может быть функционально связана с периплазматической сигнальной последовательностью, которая обеспечивает нацеливание экспрессии на секрецию в периплазматическое пространство хозяина.

Также настоящее изобретение относится к продукции на высоком уровне CRM₁₉₇, где периплазматической экспрессии достигают путем предоставления подходящей индуцирующей температуры для экспрессии полинуклеотида без какой-либо гетерологичной последовательности для направленного транспорта в периплазматическое пространство.

Необязательно полинуклеотид по изобретению также может быть связан с полинуклеотидами полипептидов-меток. Также известно, что присутствие метки повышает стабильность и растворимость белка в цитоплазме и служит для его последующей очистки.

Эти метки могут быть связаны на 5'-конце или 3'-конце, отдельно или в комбинации связанной с множественным мечением, с олигонуклеотидной последовательностью, которая кодирует полипептид-метку для способствования их цитоплазматической стабильности и/или последующей очистке, с использованием матриц и смол с высокой аффинностью в отношении различных пептидов-меток. Различные метки, которые можно использовать согласно изобретению, включают метки HA (гемагглютинин), MYC-метку, Strep II, FLAG, HPC (тяжелая цепь белка C), глутатион-S-трансферазу (GST), мальтоза-связывающий белок (MBP), целлюлоза-связывающий белок (CBD) и хитин-связывающий белок (CBP).

Полинуклеотид по изобретению также может быть встроен в векторную конструкцию, содержащую регуляторную последовательность, с использованием молекулярных способов, хорошо известных в данной области (См. Sambrook et al., *Molecular Cloning*, 2nd ed., (1989)). Она включает, но не ограничивается ими, подходящий промотор, ориджин репликации, участок связывания рибосомы, последовательность терминации транскрипции, селективные маркеры и участок множественного клонирования. В частности, предпочтительной является плаزمид с эффективной и специфической конструкцией; такая как плазмид, включающая промотор T7, специфичный к ферменту РНК-полимеразе фага T7. Такие способы могут представлять собой, но не ограничиваться ими, способы, описанные в патентной заявке США № 2012/0128727; патенте США № 5055294; патенте США № 5128130; патенте США № 5281532; патенте США № 4695455; патенте США № 4861595; патенте США № 4755465 и патенте США № 5169760. Плазмидная система для продукции белка CRM₁₉₇ в *Corynebacterium diphtheriae* также описана в патенте США № 5614382.

В одном варианте осуществления клетка-хозяин представляет собой грамтрицательную клетку-хозяина. Экспрессирующие системы клеток-хозяев, таких как *Escherichia coli*, *Bacillus* sp., *Pseudomonas* sp., широко обсуждаются для продукции белков. В одном варианте осуществления полинуклеотид по изобретению предпочтительно используют для внутриклеточной экспрессии CRM₁₉₇ в *Escherichia coli*, где штамм-хозяин выбран из BL21 (DE3), BL21 A1, HMS174 (DE3), DH5 α , W3110, B834, origami, Rosetta, NovaBlue (DE3), Lemo21 (DE3), T7, ER2566 и C43 (DE3).

В предпочтительном варианте осуществления настоящего изобретения относится к полинуклеотидной последовательности (SEQ ID NO: 2), кодирующей бактериальный токсин, который необязательно можно лигировать в вектор, а затем вводить его в клетку-хозяина. Введение в клетку-хозяина можно проводить любым из способов, известных в данной области. Такое введение или трансформацию можно проводить посредством физического или химического способа трансформации. Затем проводят селекцию преобразованных колоний на чашках Петри с добавлением антибиотика.

Подходящие векторы, используемые в рамках настоящего изобретения, включают, но не ограничиваются ими, pET9a, pET3a, pET3b, pET3c, pET5a, pET5b, pET5c, pET9b, pET9c, pET12a, pTWIN1, pTWIN2, pET12b, pET12c, pET17b и, как правило, все векторы, которые имеют сильный промотор фага T7 (например pRSETA, B и C [Invitrogen]) и pTYB1, pTYB2, pTYB3 и pTYB4.

В другом варианте осуществления клетки *Escherichia coli* используют для экспрессии полинуклеотида, кодирующего CRM₁₉₇. Встроенный полинуклеотид проверяют в отношении правильной ориентации и положения посредством секвенирования. Полученную конструкцию используют для трансформации клеток-хозяев любым из известных химических или физических способов. Например, для трансформации хозяина используют электропорацию клеток-хозяев с помощью электрического поля в диапазоне от 6,5 кВ·см⁻¹ до 25 кВ·см⁻¹, являющуюся предпочтительным способом в рамках настоящего изобретения. Этим клеткам позволяют расти в течение от 30 мин до 120 мин при 25-40°C в подходящей среде, такой как среда LB или SOC, а затем переносят в чашки Петри с селективной средой на 10-36 ч, при 25-40°C, где отбирают положительные колонии, содержащие полинуклеотиды по изобретению.

Селекцию положительных колоний можно проводить с маркерами или без них. Подходящие маркеры, которые можно использовать, выбирают из, но не ограничиваясь ими, антибиотиков, таких как ампициллин, канамицин и т.п.

Полинуклеотид, кодирующий полноразмерный белок CRM₁₉₇, клонируют рядом с промотором *lad* T7, который контролирует экспрессию белка в штаммах *Escherichia coli*, положительных по полимеразе T7. Экспрессия полинуклеотида строго контролируется промотором T7, который индуцируется в при-

сутствии IPTG или самоиндукционных средах.

Параметры культивирования хозяина оптимизируют для экспрессии белка CRM₁₉₇ на высоком уровне. В одном варианте осуществления оптимизируют компоненты культуральных сред, условия культивирования, включающие температуру роста, концентрацию индуцирующих агентов и время индукции. Используемую культуральную среду можно выбирать из, но не ограничиваясь ими, сред с определенным химическим составом, LB (Луриа-Бертани), TB (Terrific Broth), SOB (Super Optimal Broth), SOC (Super Optimal Broth с катаболическим репрессором), бульона YT (дрожжевой экстракт и триптон), Super broth, обогащенных сред, минимальных сред, минеральных сред и т.п. Ингредиенты сред включают, но не ограничиваются ими, источник углерода, например, такой как глюкоза, сахароза или глицерин, источник органического азота, такой как пептон, триптон, аминокислоты или дрожжевой экстракт, используют источник неорганического азота, и он может быть выбран, например, из солей аммония, водного раствора аммиака и газообразного аммиака, добавок, например, дополненных низкими уровнями аминокислот, витаминов, пептонов или других ингредиентов. Культуральную среду можно получать с использованием способов, известных в данной области.

Трансформированные клетки-хозяева можно исследовать в отношении экспрессии в малом объеме, таком как 5-50 мл, в LB, terrific broth или среде с определенным химическим составом. Экспрессию можно подвергать воздействию различных концентраций индукторов на уровне приблизительно 0,01 мМ, приблизительно 0,05 мМ, приблизительно 0,2 мМ, приблизительно 0,3 мМ, приблизительно 0,4 мМ, приблизительно 0,5 мМ, приблизительно 0,6 мМ, приблизительно 0,7 мМ, приблизительно 0,8 мМ, приблизительно 0,9 мМ и приблизительно 1 мМ. Экспрессию полипептида определяют с использованием электрофореза, предпочтительно электрофореза SDS PAGE, и исследуют в качестве полос сверхэкспрессии, окрашенных кумасси бриллиантовым синим (-).

Затем клетки-хозяева инокулируют в культуры в 500-мл флаконах; и им позволяют расти в оптимальных условиях, при которых экспрессия CRM₁₉₇ продолжается от 16 ч до 32 ч. После культивирования при постоянном встряхивании и предпочтительно в аэробных условиях, клетки собирают и лизируют. Для лизиса клеток можно использовать любые из известных способов, причем предпочтительный способ включает лизирующий буфер, содержащий детергент в соответствующей концентрации. После лизиса белковый компонент объединяют на одной или нескольких стадиях центрифугирования. Лизис проводят в буфере, содержащем Tris-HCl 20-50 мМ, pH 7,5-8,5, NaCl 100-150 мМ, детергент 0,5-1,5% и ингибитор протеаз 0,5-1,5%, при встряхивании.

В одном варианте осуществления температура индукции для экспрессии составляет от 10 до 40°C. В одном варианте осуществления CRM₁₉₇ получают при температуре индукции регулируемым образом, где, когда температуру индукции поддерживают на уровне от 10 до 20°C, более 80% экспрессируемого CRM₁₉₇ присутствует в растворимой фракции. В другом варианте осуществления, когда температуру индукции поддерживают на уровне от 25 до 40°C, более 80% полученного экспрессируемого CRM₁₉₇ представляет собой нерастворимую фракцию в качестве цитоплазматических телец включения, из которых его очищают после стадии солиubilизации из объединенной цитоплазматической фракции.

В одном варианте осуществления осуществления цитоплазматические тельца включения солиubilизируют в различных концентрациях мочевины, по существу в 1 М мочеvine, или 2 М мочеvine, или 3 М мочеvine, или 4 М мочеvine, или 5 М мочеvine, или 6 М мочеvine, или 7 М мочеvine, или 8 М мочеvine, или 9 М мочеvine.

В конкретном варианте осуществления выход растворимого CRM₁₉₇ составляет приблизительно 0,1 г/л, 0,25 г/л, 0,5 г/л, приблизительно 1 г/л, приблизительно 1,5 г/л, приблизительно 2 г/л, приблизительно 2,5 г/л, приблизительно 3 г/л, приблизительно 3,5 г/л, приблизительно 4 г/л, приблизительно 4,5 г/л, приблизительно 4,5 г/л, приблизительно 5 г/л.

В конкретном варианте осуществления выход нерастворимого CRM₁₉₇ составляет приблизительно 0,25 г/л, 0,5 г/л, приблизительно 1 г/л, приблизительно 1,5 г/л, приблизительно 2 г/л, приблизительно 2,5 г/л, приблизительно 3 г/л, приблизительно 3,5 г/л, приблизительно 4 г/л, приблизительно 4,5 г/л, приблизительно 4,5 г/л, приблизительно 5 г/л.

Экспрессируемый белок очищают с использованием ионообменной хроматографической колонки, а затем аффинной хроматографии.

Таким образом, изобретение вовлекает более одной последовательной стадии очистки, и также используется величина pI CRM₁₉₇ на стадии ионообменной хроматографии, где его отделяют от других контаминирующих белков. Наконец, количество CRM₁₉₇ определяют с использованием анализа ВСА/Брэдфорд/Лоури и визуализируют в 10-12% акриламидном геле (SDS-PAGE).

Идентификацию полипептида проводят с использованием вестерн-блоттинга и сходных иммуноанализов. Чистоту и целостность очищенного полипептида измеряют способами SDS-PAGE и ВЭЖХ. Выход белка, экспрессированного таким образом, составляет 500-1000 мг/л культуральной среды, и его затем можно варьировать путем модулирования добавок в культуру и условий, а также стадий очистки. Способ по изобретению также обеспечивает промышленно применимый способ регуляции времени индукции, а затем модулирование pH и температуры стадий хроматографии обеспечивает простой, недорогой и нетрудоемкий способ. Это исключает потребность в существенных шагах, вовлекающих

приготовление буферов или набора, или их рабочего раствора. В процессе удаления метки отсутствует необходимость в предоставлении дополнительных буферов, или солей, или ферментов, или оборудования. В конкретных вариантах осуществления из очищенного полипептида CRM₁₉₇ без труда удаляют аминокислоту метионин, присутствие которой является нежелательным в конечном белке CRM₁₉₇ и удаление которой требует дополнительных стадий очистки. Полипептид, полученный таким образом, находится в активной и нативной форме; из него без труда удален нежелательный метионин в качестве первой аминокислоты, без необходимости в дополнительных стадиях. Аминокислотную последовательность CRM₁₉₇ анализировали с использованием инструментов *in silico*/биоинформатических инструментов; они продемонстрировали гидрофобность белка приблизительно 38,4%. Обнаружено, что изоэлектрическая точка CRM₁₉₇ составляет приблизительно 5,81. Белок CRM₁₉₇ содержал 4 аминокислотных остатка цистеина и 21 остаток пролина. Рефолдинг полипептида подтверждают посредством функциональных анализов путем измерения активности эндонуклеазы в отношении ДНК. Биофизическое подтверждение/подтверждение вторичной структуры проводят посредством анализа кругового дихроизма (CD) (фиг. 8) и дифференциальной сканирующей флуориметрии (DSF) (фиг. 9) полипептида и его сравнивают с коммерчески доступными полипептидами (Sigma Aldrich).

В другом варианте осуществления подтверждали присутствие правильной дисульфидной связи и сравнивали с коммерчески доступными полипептидами CRM₁₉₇ (Sigma Aldrich) (фиг. 10). Также подтверждали правильность аминокислотной последовательности продуцированного полипептида посредством расщепления полипептида CRM₁₉₇ множеством протеаз и картирования аминокислотной последовательности (фиг. 6). N-концевую аминокислотную последовательность продуцированного полипептида подтверждают посредством секвенирования способом деградации Эдмана (фиг. 7).

В предпочтительном варианте осуществления настоящее изобретение относится к оптимизированной полинуклеотидной последовательности (SEQ ID NO: 2) и ее структурным вариантам, имеющим равное или не более чем 70% сходство, предпочтительно сходство от 85 до 99%, пригодным для экспрессии на высоком уровне полипептида CRM₁₉₇.

В другом предпочтительном варианте осуществления настоящее изобретение относится к оптимизированной полинуклеотидной последовательности (SEQ ID NO: 2), пригодной для экспрессии на высоком уровне полипептида CRM₁₉₇ в клетке *Escherichia coli*.

В другом предпочтительном варианте осуществления настоящее изобретение относится к экспрессии на высоком уровне дифтерийного токсина, или CRM₁₉₇, или их вариантов, с использованием нуклеиновой кислоты SEQ ID NO: 2 или ее варианта в клетке грамотрицательной бактерии, предпочтительно в *Escherichia coli*, включающей стадии:

- a) выбора гена SEQ ID NO: 2 или его варианта, который кодирует полипептид CRM₁₉₇,
- b) субклонирования гена SEQ ID NO: 2 в экспрессирующий вектор,
- c) трансформации клетки-хозяина *Escherichia coli* экспрессирующим вектором стадии b;
- d) культивирования трансформированной клетки-хозяина в культуральной среде, пригодной для экспрессии белка-токсина;
- e) индукции экспрессии слитого белка посредством добавления IPTG в качестве индуцирующего агента при температуре в диапазоне от 30 до 40°C,
- f) экстракции бактериального токсоида в нерастворимой форме из клетки-хозяина, и
- g) очистки CRM₁₉₇ в чистой форме с выходом более 0,5 мг/л.

Очистку проводят с использованием хроматографии. Способ хроматографии может представлять собой аффинную хроматографию, гель-фильтрацию, высокоэффективную жидкостную хроматографию (ВЭЖХ) или ионообменную хроматографию, или комбинацию двух или более. Предпочтительно, когда CRM₁₉₇ представляет собой связанный с меткой слитый белок, для отделения CRM₁₉₇ от других белков можно использовать аффинную хроматографию.

В другом предпочтительном варианте осуществления простая стадия, вовлекающая сдвиг температуры и pH в условиях колонки, также облегчает элюирование CRM₁₉₇ из связанной с ним метки. Более конкретно, для отделения CRM₁₉₇ от метки можно использовать pH в диапазоне 6,5-8,5 и температуру в диапазоне 4-30°C.

В других вариантах осуществления CRM₁₉₇, полученный в соответствии с настоящим изобретением, используют в конъюгате с полисахаридными молекулами, выделенными из *Salmonella typhi*, *Salmonella paratyphi*, *Pneumococcus*, *Haemophilus influenzae*, *Meningococcus*, *Streptococcus pneumoniae* и других патогенных бактерий.

В другом варианте осуществления CRM₁₉₇, полученный в соответствии с настоящим изобретением, используют в качестве конъюгированного носителя для вакцин, таких как вакцины против *Salmonella typhi*, *Salmonella paratyphi*, *Pneumococcus*, *Haemophilus influenzae*, *Meningococcus*, *Streptococcus pneumoniae* и других патогенных бактерий.

Более конкретно, настоящее изобретение проиллюстрировано с помощью следующих примеров. Однако следует понимать, что настоящее изобретение не ограничивается этими примерами никоим образом, но включает его варианты в пределах описанных в настоящем описании параметров, как может быть известно специалистам в данной области.

Пример 1.

Стадия (i): синтез нового гена CRM₁₉₇.

Полноразмерный ген CRM₁₉₇ оптимизировали в соответствии с использованием кодонов *Escherichia coli*. Для оптимизации гена CRM₁₉₇ использовали следующие параметры: предпочтительное использование кодонов, содержание GC, вторичная структура мРНК, выбранные желаемые паттерны, выбранные нежелательные паттерны, повторяющиеся последовательности (прямой повтор, инвертированный повтор и двойной повтор), участки распознавания ферментами рестрикции (делеция или инсерция).

Оптимизированный ген CRM₁₉₇ (SEQ ID NO: 2) клонировали в участок множественного клонирования плазмидного вектора pUC57 с использованием участков рестрикции BamHI и SapI с получением pUC57 CRM₁₉₇. Векторы, содержащие ген CRM₁₉₇, трансформировали в хозяина *Escherichia coli* DH5α и проводили селекцию клонов на чашке с LB+канамицин^f. Присутствие и правильность гена CRM₁₉₇ в pUC57 подтверждали посредством рестрикционного расщепления плазмиды pUC57_CRM₁₉₇ посредством AgeI (расположен в гене CRM₁₉₇) и NdeI (расположен в плазмиде pUC57). Кроме того, последовательность CRM₁₉₇ подтверждали посредством ПЦР и секвенирования ДНК.

Стадия (ii): встраивание CRM₁₉₇ в экспрессирующий вектор pTWIN1.

Escherichia coli DH5α, содержащие pUC57_CRM₁₉₇, выращивали в течение ночи в LB+канамицин в объеме 50 мл. Бактерии центрифугировали и осадок использовали для выделения плазмиды. Выделение плазмиды проводили с помощью набора Qiagen plasmid mini-prep kit с использованием инструкций производителя. Выделенную плазмиду количественно определяли посредством nono-drop.

CRM₁₉₇ (SEQ ID NO: 2) вырезали из pUC57, 5 мкг плазмиды расщепляли эндонуклеазами рестрикции BamHI и SapI. Расщепленную плазмиду разделяли на 1% агарозном геле и полосу, соответствующую гену CRM₁₉₇ (SEQ ID NO: 2, ~1,6 т.п.н.) очищали с помощью набора Qiagen Gel extraction kit с использованием инструкций изготовителя. Затем 5 мкг экспрессирующей плазмиды pTWIN1 также расщепляли BamHI и SapI с получением рестрикционных участков, совместимых с геном CRM₁₉₇. Расщепленную pTWIN1 также очищали из геля с помощью набора Qiagen Gel extraction kit с использованием инструкций изготовителя.

Расщепленный ген CRM₁₉₇ лигировали в pTWIN1 с использованием набора для лигирования ДНК на основе лигазы T4 (Promega) с использованием инструкций изготовителя. Вектор (pTWIN1) и вставку (CRM₁₉₇) смешивали в соотношении 1:3, 1:4, 1:5 в присутствии ДНК-лигазы T4 и буферов в объеме реакции 20 мкл. Смесь для лигирования инкубировали в течение ночи при 16°C. На следующее утро 5 мкл смеси для лигирования добавляли/трансформировали в экспрессирующего хозяина BL21-DE3 *Escherichia coli*. BL21 трансформировали с использованием протокола химической трансформации. Смесь для лигирования+клетки BL21 инкубировали на льду в течение 30 мин. После инкубации проводили тепловой шок в течение 45 с при 42°C. Образец охлаждали при комнатной температуре и к нему добавляли 500 мкл среды SOC. Пробирку с трансформантами инкубировали в течение 2 ч при 37°C при 200 об/мин. Из этой смеси 100 мкл высевали на чашку с LB+ампициллин для скрининга трансформантов.

Селекцию экспрессирующих CRM₁₉₇ трансформантов BL21-DE3 проводили на следующее утро в чашках с бульоном Луриа+ампициллин. Из них отбирали 5 клонов, выросших на бульоне Луриа+ампициллин, и выращивали в 10 мл среды Луриа+ампициллин в течение ночи при 37°C, 200 об./мин. Культуру центрифугировали и плазмиду экстрагировали из клеточного осадка с использованием набора для экстракции плазмиды Qiagen.

Для подтверждения правильности клона 2 мкг плазмиды расщепляли эндонуклеазами рестрикции AgeI и AраI, соответственно. Участок AgeI находится в CRM₁₉₇ и AраI находится в pTWIN1. Таким образом, двойное расщепление обоими ферментами использовали для подтверждения правильности клона. Клон был обозначен как pTWIN1 CRM₁₉₇ (BL21-DE3). Более того, клоны подтверждали посредством ПЦР с использованием праймеров, специфичных к гену CRM₁₉₇, и секвенирования ДНК. Исходные культуры BL21, экспрессирующих CRM₁₉₇, в глицерине, получали путем выращивания бактерий в 10 мл среды Луриа+ампициллин в течение ночи. На следующее утро в культуру добавляли 40% стерилизованный глицерин и аликвоты объемом 1 мл распределяли во флаконы для замораживания. Флаконы хранили при -80 градусах для дальнейшего применения в анализе экспрессии.

Стадия (iii): подтверждение экспрессии CRM₁₉₇.

Клон BL21, который хранили при -80 градусах, наносили штрихами на чашку со средой Луриа+ампициллин. Чашку инкубировали в течение ночи при 37°C. Единичную колонию отбирали и инокулировали в 50 мл среды Луриа+ампициллин во флаконе объемом 150 мл. Флакон инкубировали при 37°C, 200 об/мин, до OD₆₀₀=1. После достижения желаемой OD отбирали 5 мл культуры, которую использовали в качестве неиндуцированной культуры. Неиндуцированную культуру хранили на льду до применения. Для индукции экспрессии гена CRM₁₉₇ 0,5 мМ IPTG добавляли к оставшимся 45 мл культуры и флакон далее инкубировали в течение дополнительных 4 ч при 30 градусах и вращении 200 об/мин. Индуцированную культуру собирали через 4 ч и экспрессию CRM₁₉₇ исследовали посредством SDS-PAGE (фиг. 1) и вестерн-блоттинга (фиг. 2).

Для анализа SDS-PAGE 1 мл индуцированной и неиндуцированной культуры (обе нормализованные до OD600=1) отбирали в 1,5-мл пробирку Eppendorf. Пробирку центрифугировали и осадок ресуспендировали в 50 мкл PBS. В эту суспензию добавляли 50 мкл буфера для нанесения SDS-PAGE с восстановителем (2×). Смесь кипятили при 100°C в течение 5 мин. Образец охлаждали при комнатной температуре и 20 мкл неиндуцированной и индуцированной культуры наносили на гель с 4-12% Tris-глицином. Гель разделяли в течение 1,5 ч при 150 В. Гели извлекали и инкубировали в красителе кумаси бриллиантовом синем в течение 1 ч. После окрашивания гель обесцвечивали в растворе для обесцвечивания, содержащем 40% метанол=10% уксусную кислоту в течение 3 ч. Экспрессию CRM₁₉₇ визуализировали в качестве белка размером ~58 кДа, который был виден только в индуцированной культуре. Для вестерн-блоттинга отдельный набор гелей разделяли аналогичным образом с SDS-PAGE, и гель подвергали блоттингу на PVDF (поливинилидендифторидная мембрана). Мембрану подвергали иммуноблоттингу с использованием антитела против CRM₁₉₇. На вестерн-блоте CRM₁₉₇ выглядел как отдельная иммунореактивная полоса на уровне ~58 кДа. В неиндуцированной культуре не наблюдали полосы, специфической для CRM₁₉₇. Этот эксперимент подтверждает, что клон, полученный в настоящем исследовании, может экспрессировать белок rCRM₁₉₇. Этот клон далее использовали для крупномасштабной продукции и очистки CRM₁₉₇.

Стадия (iv): ферментация и очистка CRM₁₉₇ из *Escherichia coli* BL21.

Флакон с клетками *Escherichia coli* BL21 объемом один мл инокулировали в 50 мл среды LB+Amp и выращивали в течение ночи при 37°C, 200 об/мин. Ферментацию проводили в масштабе 20 л. Клетки *Escherichia coli* инокулировали в ферментер и культивировали при 30 градусах Цельсия. Культуру индуцировали 0,5 mM IPTG при OD600=20. Через 12 ч после индукции культуру для ферментации собирали и получали клеточный осадок посредством центрифугирования. Клеточный осадок лизировали механически в гомогенизаторе. Тельца включения (которые содержат желаемый белок CRM₁₉₇) выделяли центрифугированием клеточных лизатов. Супернатант удаляли и осадок, который содержал тельца включения (IB), оставляли. IB гомогенизировали посредством ресуспендирования осадка в 8 М мочеvine и белок очищали ионообменной хроматографией. Проводили количественное определение CRM₁₉₇ на уровне ферментации. Лизаты цельных клеток разделяли на SDS-PAGE вместе с известным количеством BSA в качестве стандарта. Количественное определение CRM₁₉₇, который выглядел как полоса размером ~58 кДа на SDS-PAGE (фиг. 3) проводили посредством денситометрического анализа. В качестве эталона для количественного определения использовали BSA. Было обнаружено, что общее количество белка составляло 1,4 г CRM₁₉₇/литр клеток *Escherichia coli* BL21.

Тест солюбилизации CRM₁₉₇, полученного, как описано выше, проводили посредством электрофоретического разделения геля (SDS PAGE 12%), демонстрирующего тест, проведенный в отношении солюбилизации CRM₁₉₇, дорожка 1: солюбилизированная в мочеvine фракция CRM₁₉₇; дорожка 2: супернатант; дорожка 3: образец после первого промывания осадка; дорожка 4: образец, объединенный после 2-го промывания осадка (фиг. 4).

Присутствие CRM₁₉₇ в образце определяли посредством эксклюзионной хроматографии (SEC-ВЭЖХ), где основной элюированный пик демонстрирует присутствие CRM₁₉₇ в образце (фиг. 5).

Первичную аминокислотную последовательность CRM₁₉₇, полученного, как описано выше, определяли с помощью отпечатка пептидной массы (масс-спектрометрия), и она имела 100% сходство последовательности с эталонной последовательностью CRM₁₉₇ (фиг. 6).

А. Пептидный гидролизат BE rCRM, проанализированный посредством LC-MS.

В. Охват расщепленной трипсином, Glu-C и Asp-N последовательности CRM₁₉₇ (идентичный SEQ ID NO: 1, т.е. демонстрирует 100% гомологию).

N-концевую последовательность rCRM₁₉₇, полученную, как описано выше, подтверждали способом деградации Эдмана.

Аминокислотная последовательность из 10 аминокислот GADDWSSK (N-концевой ацетил) демонстрирует начальную часть очищенного полипептида. Первая аминокислота идентифицирована как G (фиг. 7).

Спектры CD CRM₁₉₇, полученного, как описано выше, перекрывались с эталонным CRM₁₉₇. Результат демонстрирует, что рекомбинантный CRM₁₉₇, полученный, как описано выше, структурно сходен с эталонным CRM₁₉₇ (фиг. 9).

Проводили подтверждение дисульфидных связей CRM₁₉₇, полученного способом, описанным выше (фиг. 10), и с помощью масс-спектрометрии было подтверждено, что в белке присутствует две дисульфидные связи: первая связывает аминокислоты 186 и 201, и вторая связь связывает аминокислоты 461 и 471.

Таблица I

Нумерация цистина	Пептидная последовательность
186	GQDAMYEYMAQACAGNR (SEQ ID NO: 11)
201	SVGSSLSCINLDWDVIR (SEQ ID NO: 12)
461	CR (SEQ ID NO: 13)
471	AIDGDVTFCRPK (SEQ ID NO: 14)

Таблица II

Комбинация триптических пептидных цепей	Теоретическое значение моно-м/z	Наблюдаемое значение моно-м/z
Cys 186- 201	935, 6706	935, 6694
Cys 461- 471	913, 9558	913, 9539

Антигенное сходство CRM₁₉₇, полученного описанным выше способом, с эталонным CRM₁₉₇ подтверждали посредством ELISA, специфического для CRM₁₉₇. Все CRM, происходящие из различных источников, продемонстрировали сходный профиль распознавания моноклональными антителами (фиг. 11).

Перевод чертежей

Фиг. 1:

1. kDa - кДа;
2. CRM₁₉₇ band (~58kDa) - полоса CRM₁₉₇ (~58 кДа).

Фиг. 2:

1. kDa - кДа;
2. CRM₁₉₇ band (~58kDa) - полоса CRM₁₉₇ (~58 кДа).

Фиг. 3:

1. kDa - кДа.

Фиг. 4:

1. Urea Solubilized pellet - солюбилизированный в мочеvine осадок;
2. Supernatant - супернатант;
3. Pellet Wash-1 - промывание осадка 1;
4. Pellet Wash-2 - промывание осадка 2.

Фиг. 5:

1. CRM₁₉₇ Protein - белок CRM₁₉₇;
2. Minutes - минуты.

Фиг. 6:

1. Trypsin - трипсин.

Фиг. 7:

1. Initial Yield - первоначальный выход;
2. Repetitive Yield - повторный выход;
3. Percent Yield - процентный выход;
4. Observed - наблюдаемый;
5. Reference - эталон;
6. Cycle - цикл.

Фиг. 8:

1. Reference rCRM₁₉₇ - эталонный rCRM₁₉₇;
2. nm - нм;
3. Wavelength - длина волны.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Оптимизированная полинуклеотидная последовательность, кодирующая Cross Reacting Material 197 (CRM₁₉₇), содержащая SEQ ID NO: 2, и ее варианты, которые по меньшей мере на 88% гомологичны указанной оптимизированной полинуклеотидной последовательности SEQ ID NO: 2.

2. Варианты SEQ ID NO: 2 по п.1, выбранные из SEQ ID NO: 3, 4, 5, 6, 8 и 9.

3. Способ получения CRM₁₉₇, включающий стадии:

- a) выбора оптимизированной полинуклеотидной последовательности, кодирующей CRM₁₉₇, по существу состоящей из SEQ ID NO: 2, или ее вариантов, которые по меньшей мере на 88% гомологичны SEQ ID NO: 2,
- b) лигирования полинуклеотидных последовательностей стадии (a) в подходящий вектор,
- c) встраивания или трансформации полинуклеотидной последовательности в клетку-хозяина *Escherichia coli*,
- d) культивирования трансформированной клетки в культуральной среде для экспрессии CRM₁₉₇ на высоком уровне,
- e) поддержания индуцирующей температуры между 10 и 40°C для продукции CRM₁₉₇,
- f) экстракции CRM₁₉₇ из клетки-хозяина с последующей очисткой для получения чистого CRM₁₉₇ с высоким выходом.

4. Способ по п.3, в котором подходящий вектор представляет собой плазмидный вектор, выбранный из pET9a, pET3a, pET3b, pET3c, pET5a, pET5b, pET5c, pET9b, pET9c, pET12a, pTWIN1, pTWIN2, pET12b, pET12c, pET17b.

5. Способ по п.3, в котором штамм *Escherichia coli* выбран из BL21 (DE3), BL21 A1, HMS174 (DE3), DH5 α , W3110, B834, origami, Rosetta, NovaBlue (DE3), Lemo21 (DE3), T7, ER2566 и C43 (DE3).

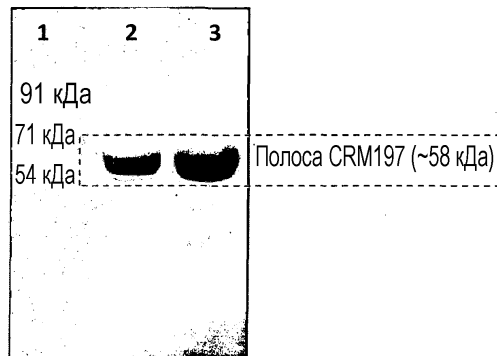
6. Способ по п.3, в котором выход CRM₁₉₇ составляет приблизительно 0,1 г/л, 0,25 г/л, 0,5 г/л, приблизительно 1 г/л, приблизительно 1,5 г/л, приблизительно 2 г/л, приблизительно 2,5 г/л, приблизительно 3 г/л, приблизительно 3,5 г/л, приблизительно 4 г/л, приблизительно 4,5 г/л, приблизительно 5 г/л.

7. Способ по п.3, в котором периплазматической экспрессии достигают путем обеспечения подходящей температуры индукции экспрессии полинуклеотида без какой-либо гетерологичной последовательности для направленного транспорта в периплазматическое пространство.

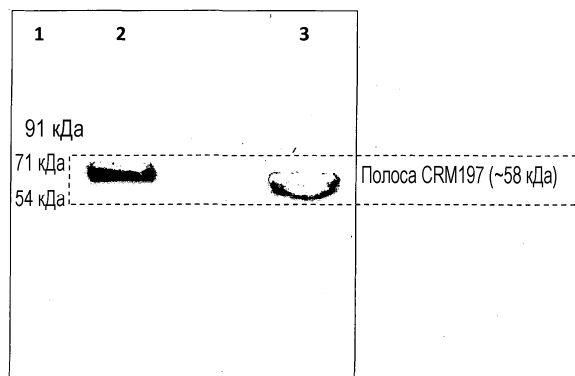
8. Способ по п.7, в котором температура индукции находится в диапазоне от 10 до 40°C.

9. Способ по п.6, в котором указанный CRM₁₉₇ конъюгирован с полисахаридными молекулами, выделенными из *Salmonella typhi*, *Salmonella paratyphi*, *Pneumococcus*, *Haemophilus influenzae*, *Meningococcus*, *Streptococcus pneumoniae* и других патогенных бактерий.

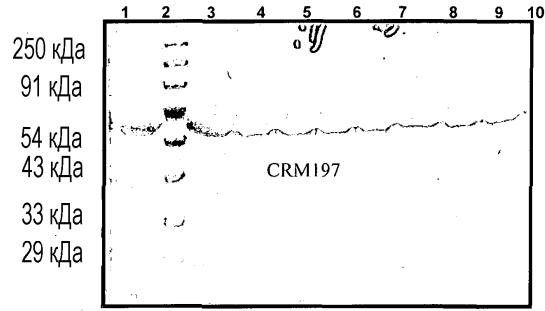
10. Способ по п.6, в котором указанный CRM₁₉₇ используют в качестве конъюгированного носителя для вакцин, таких как вакцины против *Salmonella typhi*, *Salmonella paratyphi*, *Pneumococcus*, *Haemophilus influenzae*, *Meningococcus*, *Streptococcus pneumoniae* и других патогенных бактерий.



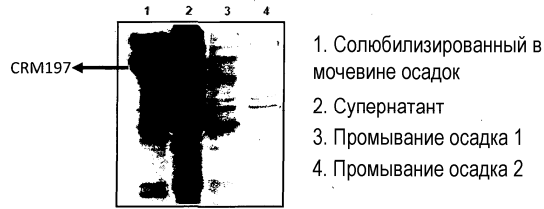
Фиг. 1



Фиг. 2

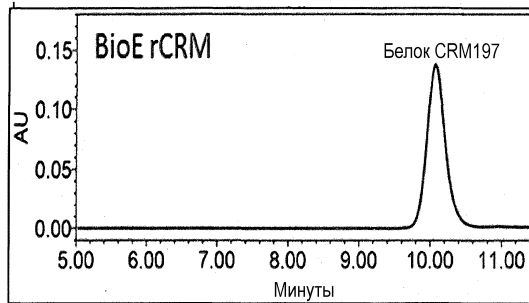


Фиг. 3

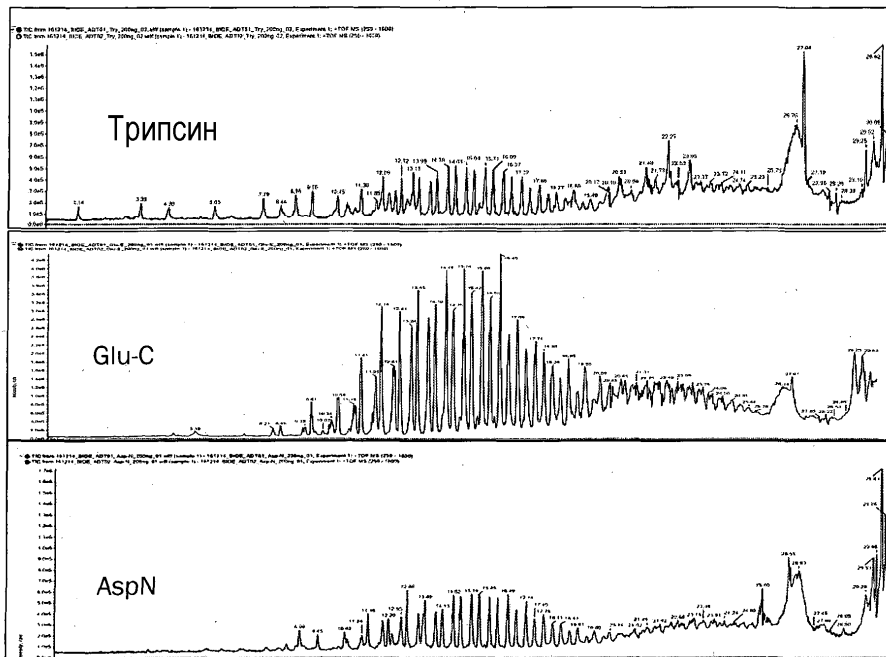


1. Солюбилизованный в мочеvine осадок
2. Супернатант
3. Промывание осадка 1
4. Промывание осадка 2

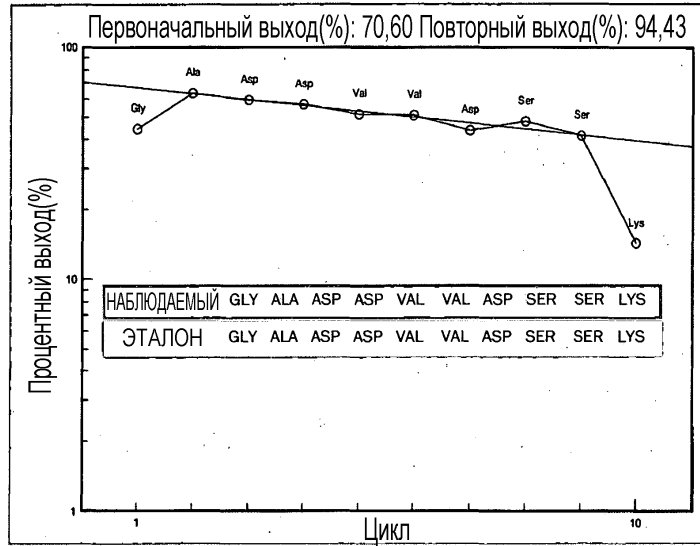
Фиг. 4



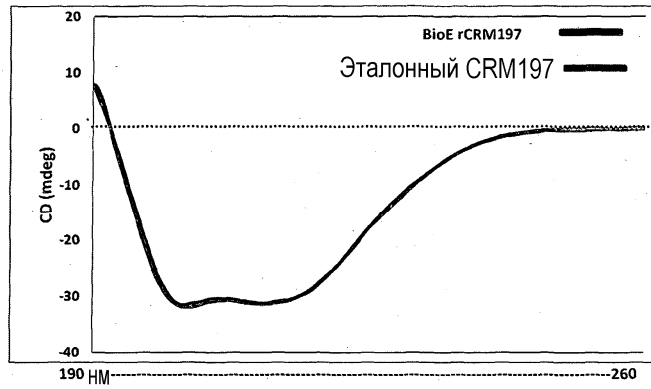
Фиг. 5



Фиг. 6

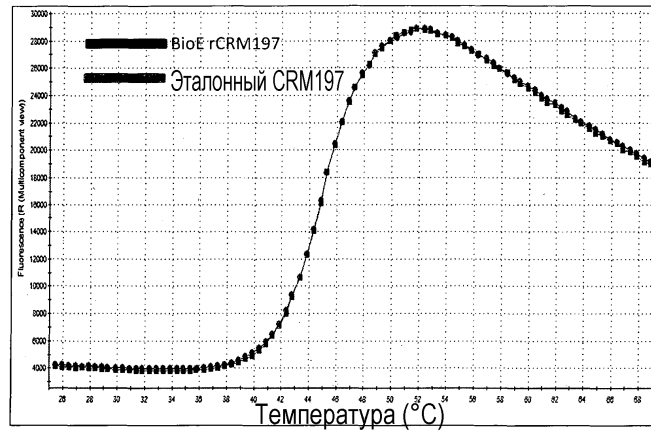


Фиг. 7

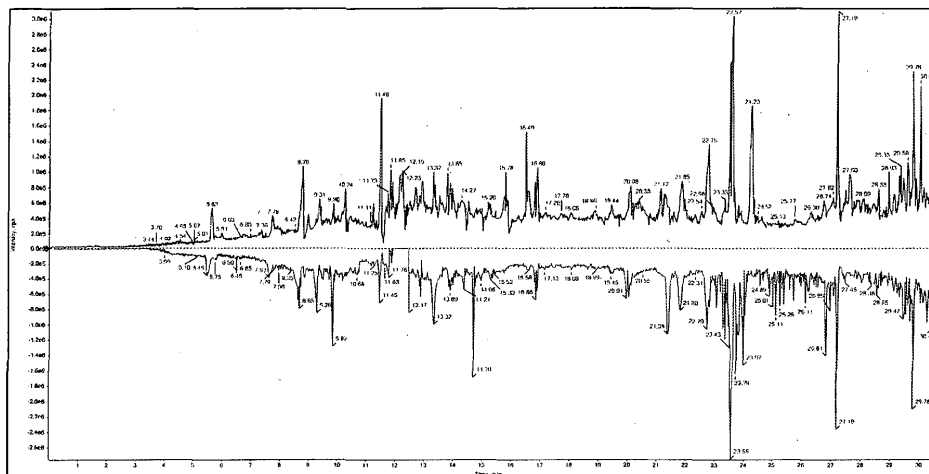


Длина волны

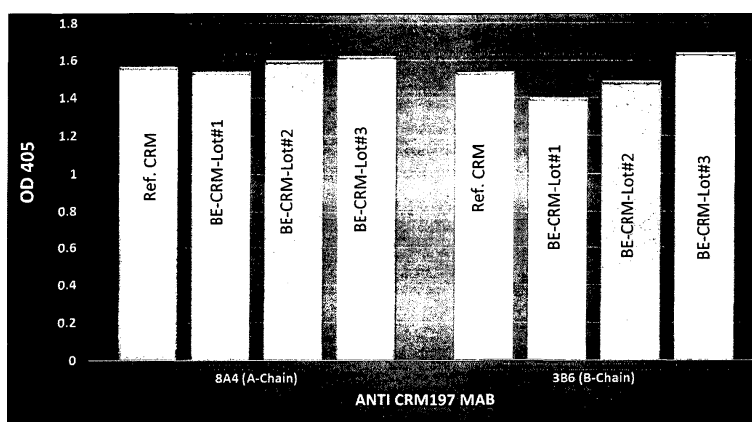
Фиг. 8



Фиг. 9



Фиг. 10



Фиг. 11