

Санкт-Петербургский Государственный университет

Биологический факультет

Кафедра биохимии

Сидоренкова Ольга Ирэковна

Исследование CasX нуклеазы из Deltaproteobacteria

Выпускная квалификационная работа
по направлению подготовки Биология
основная образовательная программа бакалавриата «Биология»

Работа выполнена в НИК «НаноБио» СПбПУ
в лаборатории молекулярной микробиологии Константина Северинова
директор НИК «НаноБио» к.ф.-м.н. Ходорковский Михаил Алексеевич

Научный руководитель:
к.б.н., доцент кафедры биохимии Цветкова Елена Викторовна

Научный консультант:
младший научный сотрудник Сколтех Федорова Яна Витальевна

Санкт-Петербург

2019

Содержание

Список условных обозначений и сокращений	4
Введение	5
1. Обзор литературы	7
1.1. Общий принцип организации CRISPR-Cas систем	7
1.2. Использование CRISPR-Cas систем в генной инженерии	14
1.3. Разнообразие Cas белков 2 класса – Cas12a, Cas12b, Cas12c, Cas13a	16
1.4. CasX – отдельное семейство редакторов генома	18
2. Материалы и методы	22
2.1. Праймеры	22
2.2. Среды и растворы, использованные в работе	23
2.3. Гель-электрофорез в агарозном геле и ПААГ	24
2.4. Молекулярное клонирование, методы работы с бактериями и плазмидами	25
2.5. Синтез рекомбинантного белка в клетках-продуцентах <i>Escherichia coli</i> штамма Rosetta (DE3) и анализ его растворимости	27
2.6. Выделение и хроматографическая очистка рекомбинантного белка	28
2.7. Синтез направляющей РНК	30
2.8. Синтез ДНК-мишеней	30
2.9. Проведение эндонуклеазной реакции <i>in vitro</i>	31
3. Результаты	31
3.1. Конструирование плазмидного вектора для экспрессии системы рЕТ21a CasX	31
3.2. Клонирование рЕТ21a CasX	33
3.3. Синтез белка DpbCasX в клетках-продуцентах <i>E. coli</i> штамма Rosetta (DE3) и анализ его растворимости	34
3.4. Конструирование плазмидного вектора с последовательностью МВР и повторное клонирование	35
3.5. Выделение рекомбинантного белка DpbCasX_МВР и его хроматографическая очистка	37
3.6. Проведение эндонуклеазной реакции <i>in vitro</i>	41
4. Обсуждение	44
Выводы	47

Благодарности	48
Список использованной литературы	49

Список условных обозначений и сокращений

CRISPR – короткие палиндромные повторы, регулярно расположенные группами (clustered regularly interspaced short palindromic repeats)

Cas – CRISPR ассоциированные гены (CRISPR associated)

ZFN – нуклеазы с доменом «цинковые пальцы» (Zn-fingers nucleases)

TALEN – нуклеазоподобные активаторы транскрипции (Transcription Activator-Like Effector Nucleases)

PAM – мотив смежный с протоспейсером (Protospacer adjacent motif)

sgРНК – гибридная молекула РНК из crРНК и tracrРНК

DISARM – гены, ассоциированные с системой рестрикции-модификации (Defence Island System Associated with Restriction–Modification)

HNH – нуклеазный домен белка Cas9, который расщепляет цепь ДНК-мишени, комплементарную crРНК

RuvC – нуклеазный домен белка Cas9, который расщепляет цепь ДНК-мишени, некомплементарную crРНК

NLS – сигнал ядерной локализации (Nuclear Localization Signal)

HEPN – нуклеотид-связывающие домены эукариот и прокариот (Higher Eukaryotes and Prokaryotes Nucleotide-binding domain)

GFP – зеленый флуоресцентный белок (Green Fluorescent Protein)

OBD – олигонуклеотид-связывающий домен (Oligonucleotide-Binding Domain)

NTSB – домен, связывающий цепь ДНК-мишени, некомплементарную crРНК (Non-Target-Strand-Binding)

TSL – домен, связывающий цепь ДНК-мишени, комплементарную crРНК (Target-Strand-Loading)

IPTG – изопропил- β -D-1-тиогалактопиранозид (Isopropyl β -D-1-thiogalactopyranoside)

β Мэ – β -меркаптоэтанол

Введение

Изменение ДНК живых организмов представляет большой интерес для области биотехнологии и медицины. Целевое редактирование последовательности ДНК позволяет изучать функции генов, вносить изменения в геном путем удаления дефектных или же введения новых последовательностей. На сегодняшний день большая часть подходов для изменения генома основана на создании направленных «разрезов» в геномной ДНК, которые затем исправляются внутриклеточными системами репарации, что позволяет вставлять новые последовательности путем рекомбинации или нокаутировать гены, за счет негомологичного воссоединения образовавшихся после разрыва молекулы концов.

Для внесения изменений в конкретном сайте генома используют мегануклеазы, нуклеазы с доменом «цинковые пальцы» Zn-fingers nucleases (ZFN) и нуклеазоподобные активаторы транскрипции Transcription Activator-Like Effector Nucleases (TALEN), последовательности автосплайсируемых интронов [Guha, Wai, Hausner, 2017].

С 2014 года в качестве инструмента генного редактирования широкое применение получили CRISPR-Cas системы адаптивного иммунитета бактерий [Doudna, Charpentier, 2014]. Локус CRISPR и ассоциированные с ним гены Cas помогают бактериям запоминать, узнавать и уничтожать чужеродные генетические элементы. CRISPR – своеобразная копилка участков чужеродного генетического материала. Cas белок работает в комплексе с направляющей РНК, которая комплементарно связывается с ДНК-мишенью. Благодаря этому эндонуклеаза узнает целевой сайт и может вносить по нему двунитевые разрывы в молекулу. Гибридные РНК можно синтезировать искусственно и вместе с эффекторным белком Cas направлять на разнообразные мишени.

Система проста и удобна в использовании, с ее помощью можно редактировать ДНК эукариотических организмов с высокой точностью до одного нуклеотида. Наиболее известная и хорошо изученная система – CRISPR/Cas9 бактерии *Streptococcus pyogenes*, и она уже активно используется в генной инженерии [Jinek и др., 2013; Tang и др., 2017; Yao и др., 2018]. Но SpCas9 обладает рядом недостатков: нуклеаза неспецифична и может вызывать нежелательные мутации, большой размер гена SpCas9 не позволяет осуществлять его эффективную доставку аденоассоциированными вирусами, выбор мишеней ограничен NGG-фланкированными сайтами. Поэтому характеристика неизученных Cas-белков из разных бактерий открывает возможности для нахождения новых, наиболее эффективных инструментов для геномного редактирования.

Целью данной работы было продемонстрировать активность CasX нуклеазы из Deltaproteobacteria.

Для достижения этой цели был поставлен ряд задач:

1. Создать генетическую конструкцию для экспрессии системы pET21a CasX.
2. Получить рекомбинантный белок DpbCasX.
3. Показать в условиях *in vitro* нуклеазную активность белка DpbCasX.

1. Обзор литературы

1.1. Общий принцип организации системы CRISPR/Cas у бактерий

CRISPR Cas системы – защитные системы бактерий и архей

У архей и бактерий существуют системы «врожденного» и «приобретенного» иммунитета, которые защищают их при заражении бактериофагами или другими чужеродными генетическими элементами. «Врожденный» иммунитет представлен системами рестрикции-модификации и abortивной бактериофагальной инфекции [Du и др., 2014]. Ранее считалось, что прокариоты не обладают адаптивным иммунитетом. Сейчас к нему относят CRISPR-Cas и мало изученный механизм DISARM [Ofir и др., 2018]. Задача CRISPR-Cas систем – обеспечить защиту бактериальной клетки от повторного заражения чужеродными генетическими элементами. В состав систем входят два основных элемента: CRISPR-кассета и прилегающий к ней кластер генов cas (Рис. 1).

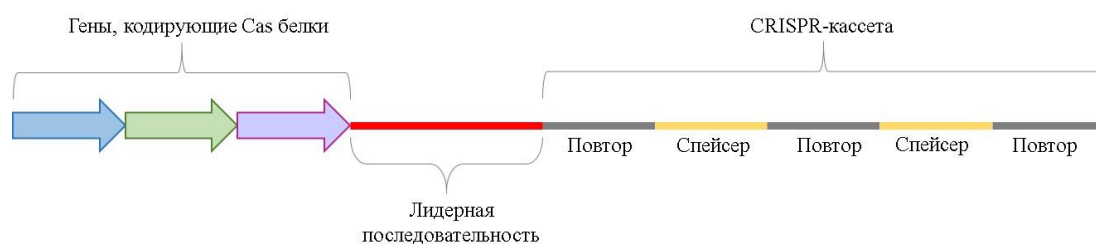


Рис. 1. Структура локуса CRISPR.

CRISPR (сокращение от clustered regularly interspaced palindromic repeats) — это последовательность бактериального генома, представляющая собой кассету, состоящую из повторяющихся последовательностей, прямых повторов (DR – direct repeats). Прямые повторы разделены уникальными последовательностями – спейсерами, совпадающими с различными генетическими элементами (бактериофаги, плазмиды, транспозоны), с которыми в течение жизни могла сталкиваться бактерия или ее предки. Последовательности спейсеров являются частью генома клетки и поэтому передаются её потомкам при делении. Прямые повторы и спейсеры имеют, как правило, длину от 25 до 40 нуклеотидов [Westra, Buckling, Fineran, 2014]. Прямые повторы консервативны среди родственных систем CRISPR-Cas, что дополняет и расширяет уже имеющуюся классификацию на основании cas-генов [Kunin, Sorek, Hugenholtz, 2007]. При попадании в клетку любого экзогенного материала происходит добавление новых спейсеров в CRISPR-кассету. Таким способом бактерии приспосабливаются к окружающей среде. Они способны быстро эволюционировать вслед за постоянно изменяющимися вирусами и совершенствовать собственные молекулярные механизмы защиты [Takeuchi и др., 2012].

К CRISPR массиву примыкает АТ-богатая регуляторная лидерная последовательность, выступающая в качестве промотора и обеспечивающая постоянную экспрессию последовательности кассеты в длинную пре-crisprРНК, которая затем процессируется в зрелые crisperРНК [Deltcheva и др., 2011]. Второй компонент системы – CRISPR-ассоциированные cas-гены, кодирующие cas-белки. Cas белки совместно с РНК, транскрибирующимися с локуса CRISPR, проявляют нуклеазную активность в отношении чужеродной ДНК. Узнавание белком Cas нужной последовательности для расщепления ДНК осуществляется посредством комплементарного связывания с ней РНК (Рис.2).

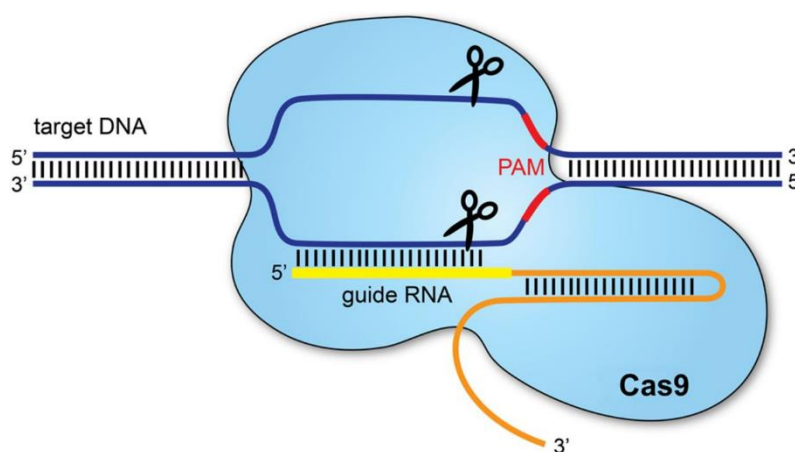


Рис. 2. Схема работы эффекторного комплекса на примере эндонуклеазы Cas9 [Redman и др.,2012]

В зависимости от состава эффекторного модуля, расщепляющего ДНК, CRISPR-Cas системы разделяют на два класса [Макарова и др., 2015].

Табл. 1. Классификация CRISPR-Cas систем.

1 класс	2 класс
Мультибелковый эффекторный комплекс. Сюда относят системы I (Cascade), III (Csm/Cmr), IV(Cfs) типа	Включают лишь один эффекторный белок. II (эндонуклеаза Cas9), V (Cpf1) тип систем

Работу CRISPR-Cas систем можно разделить на три этапа:

1. Адаптация – встраивание новых спейсеров в кассету
2. Экспрессия – экспрессия Cas генов и синтез направляющих РНК
3. Интерференция – узнавание ДНК-мишеней и их деградация

На этапе адаптации комплекс специализированных Cas белков связывается с чужеродной ДНК, перемещается вдоль нее до обнаружения 2-4 буквенного мотива, называемого PAM (Protospacer Adjacent Motif). После узнавания участка комплекс «вырезает» часть ДНК – протоспейсер. Затем белковый комплекс встраивает этот новый фрагмент в начало CRISPR-массива, добавляя новую спейсерную последовательность. Так формируется иммунологическая память (Рис.3).

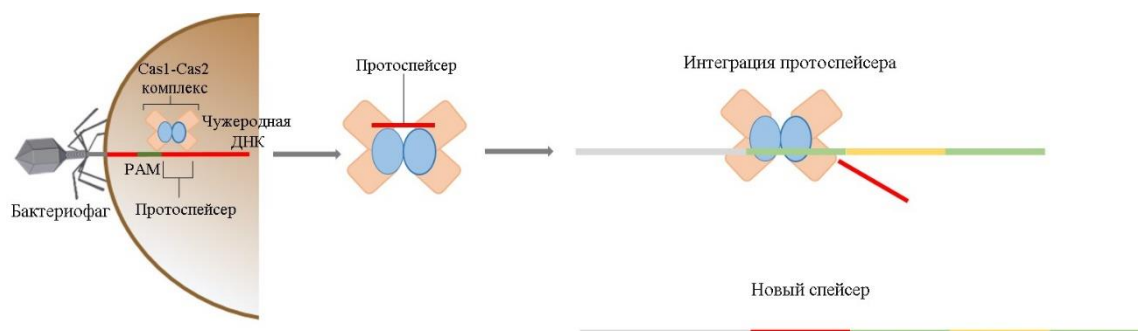


Рис. 3. Этап адаптации.

Адаптация – сложный, многоступенчатый процесс. Несмотря на разнообразие CRISPR-Cas систем, основные компоненты адаптационного аппарата консервативны и представлены белками Cas1 и Cas2, действующими в виде комплекса. По сравнению с этапами экспрессии и интерференции молекулярный механизм адаптации долгое время оставался неизвестным. Для изучения принципов этого процесса применяют экспериментальные системы трех типов. Наличие или отсутствие вставки проверяют проведением ПЦР с использованием праймеров, кодирующих лидерную последовательность и первый спейсер в кассете. Анализ продуктов реакции методом гелелектрофореза или секвенированием показывает, осуществилась ли интеграция спейсера в геном [Yosef, Goren, Qimron, 2012]. Следующий подход основан на положительном отборе клонов, которые становятся устойчивыми к антибиотику при интеграции нового спейсера в CRISPR-кассету. Это происходит за счет того, что вставка нового спейсера сдвигает рамку считывания, содержащую стоп-кодон, запуская тем самым транскрипцию гена, кодирующего антибиотик [Díez-villaseñor и др., 2013]. Третий подход включает сборку *in vitro* адаптационного комплекса из белков Cas1 и Cas2 и меченных ДНК-мишеней, и его инкубацию с плазмидой-акцептором нового спейсера. Интеграция спейсера в геномную последовательность приводит к релаксации суперскрученной плазмидной ДНК или линейзации молекулы. Обе формы ДНК легко детектируются с помощью гелелектрофореза [James K. Nuñez, Amy S.Y. Lee, 2015]. Этими методами удалось установить, что обычно комплекс Cas1-Cas2 вставляет новую спейсерную последовательность между лидерной последовательностью и первым повтором CRISPR-

кассеты. Лидерная последовательность помимо промотора и регуляторной области содержит на проксимальном конце короткую последовательность, необходимую для распознавания спейсера. Считается, что область соединения лидерной последовательности и первого повтора является ключевой для распознавания адаптационным комплексом [Wei и др., 2015].

- Молекулярный механизм адаптации

- 1) *Приобретение спейсера*

Разрывы ДНК, которые образуются в остановленных вилках репликации, способствуют приобретению спейсера. В бактериальном геноме есть Chi-сайты (горячие точки рекомбинации), вблизи которых происходит гомологичная рекомбинация при восстановлении двуцепочечных разрывов ДНК. Эти короткие последовательности служат сигналом для активации фермента RecBCD (геликаза-нуклеаза), запускающего репарационную рекомбинацию. Продукты деградации ДНК RecBCD являются источником спейсеров, а Chi-сайты ограничивают размер будущих вставок (Рис.4). Собственная ДНК бактерии защищена высокой плотностью Chi-сайтов на хромосоме, которые будут быстро снижать активность фермента RecBCD [Levy и др., 2015]. Некоторые фаги приспособляются к такому механизму защиты, экспрессируя ингибиторы RecBCD и формируя другой механизм автономной рекомбинации или насыщая свои геномы Chi-сайтами [Vobay, Touchon, Rocha, 2013].

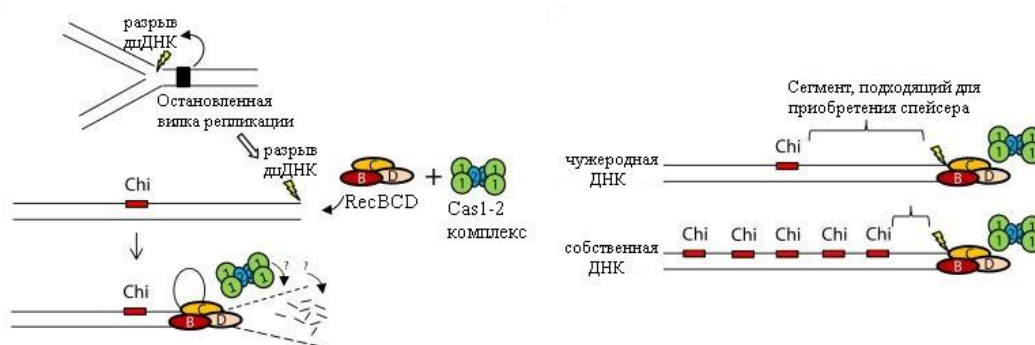


Рис. 4. Модель, объясняющая предпочтение чужеродной ДНК при приобретении спейсера [Levy и др., 2015].

Адаптационный комплекс представляет собой гетерогексамер Cas1-Cas2, с которым протоспейсер связывается не полностью, а с образованием выступов на 5'- и 3'-концах. В димере Cas1 есть специальный «карман» для распознавания PAM-комплементарной последовательности на 3'-конце. После ее узнавания, она ориентируется в «кармане»

таким образом, что происходит расщепление одноцепочечных выступов. На 3'-конце образуется гидроксильная группа, которая выполняет каталитическую функцию [Wang и др., 2015].

2) Интеграция спейсера в геном

Адаптационный комплекс Cas1-Cas2 выполняет функцию интегразы. Вставка нового спейсера в геном осуществляется в два этапа (Рис.5):

- На первом 3'-ОН группа протоспейсера атакует минус-цепь ДНК, в области соединения первого повтора и спейсера.
- На втором вторая 3'-ОН группа на противоположном конце протоспейсера осуществляет нуклеофильную атаку плюс-цепи по месту соединения лидерной последовательности и первого повтора.

Так спейсер полностью интегрируется в геном. На завершающем этапе ДНК-полимераза достраивает цепи и их сшивает ДНК-лигаза. Этот процесс напоминает интеграцию ретровируса в геном или транспозицию ДНК [James K. Nuñez, Amy S.Y. Lee, 2015]. Для систем II класса характерно во время адаптации участие эффекторного белка, что было показано для эндонуклеазы Cas9. Белок на этом этапе не проявляет нуклеазной активности, а участвует в узнавании PAM [Wei, Terns, Terns, 2015].

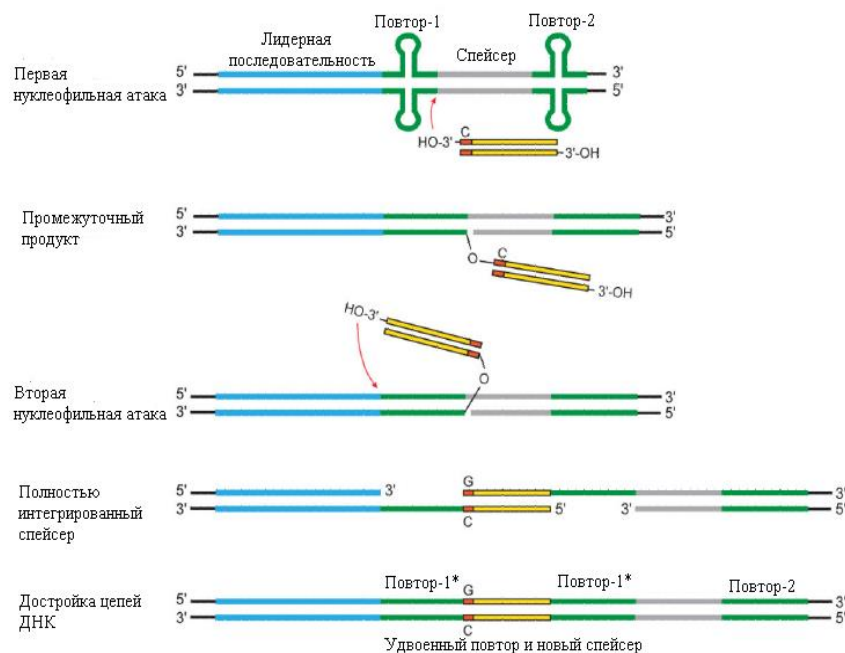


Рис. 5. Модель интеграции нового спейсера в CRISPR-кассету [James K. Nuñez, Amy S.Y. Lee, 2015].

- *Адаптационный комплекс: белки Cas1 и Cas2*

Стабильный гетерогексамерный комплекс состоит из двух димеров белка Cas1, объединенных димером Cas2. Cas1 является эндонуклеазой и выполняет каталитическую функцию. Cas2 может расщеплять ДНК и РНК, но не участвует в приобретении нового спейсера. Считается, что Cas2 направляет Cas1 на лидерную последовательность [Davies и др., 2014].

В случае CRISPR-Cas систем, действующих против чужеродных РНК-молекул, спейсеры приобретаются не из ДНК, а из РНК при помощи обратной транскриптазы, также закодированной в CRISPR-последовательности [Silas и др., 2016; Silas и др., 2017].

Во время этапа экспрессии кассета CRISPR транскрибируется с образованием длинной последовательности пре-crРНК, которая дальше процессируется в зрелые crРНК. В системах I, III и IV типа пре-crРНК обрабатывается рибонуклеазами из мультибелкового комплекса, у V типа, а именно Crf1, эффекторный белок сам принимает участие в процессинге пре-crРНК [Brouns и др., 1996; Ryan N. Jackson, 2015; Zetsche и др., 2015]. Для систем II типа необходимым является присутствие трансактивирующей РНК (tracrРНК). Она комплементарно связывается с участками повторов на пре-crРНК. Эти дуплексы распознаются РНКазой III и фрагментируются на зрелые субъединицы [Deltcheva и др., 2011]. Такая зрелая crРНК молекула состоит из последовательности спейсера и смежного с ним прямого повтора. Идентичные повторы формируют особую структуру, распознаваемую эффектором, crРНК служит гидом для нацеливания нуклеазы на ДНК-мишень.

Завершающая стадия – интерференция или расщепление чужеродного участка вредоносного вируса или плазмиды, который распознается благодаря комплементарному слиянию направляющей crРНК с протоспейсером. Обязательное наличие PAM рядом с целевой последовательностью предотвращает разрушение собственной ДНК. В зависимости от класса системы эффектором может быть, как один белок (системы II, V типов), так и целый белковый комплекс (I, II и IV тип). Системы I типа кодируют эффекторный комплекс Cascade, II и V типа – один белок-нуклеазу (например, Cas9 и Crf1), III типа – комплексы Csm или Cmr, а IV тип – комплекс из белков Csf (Рис.6).

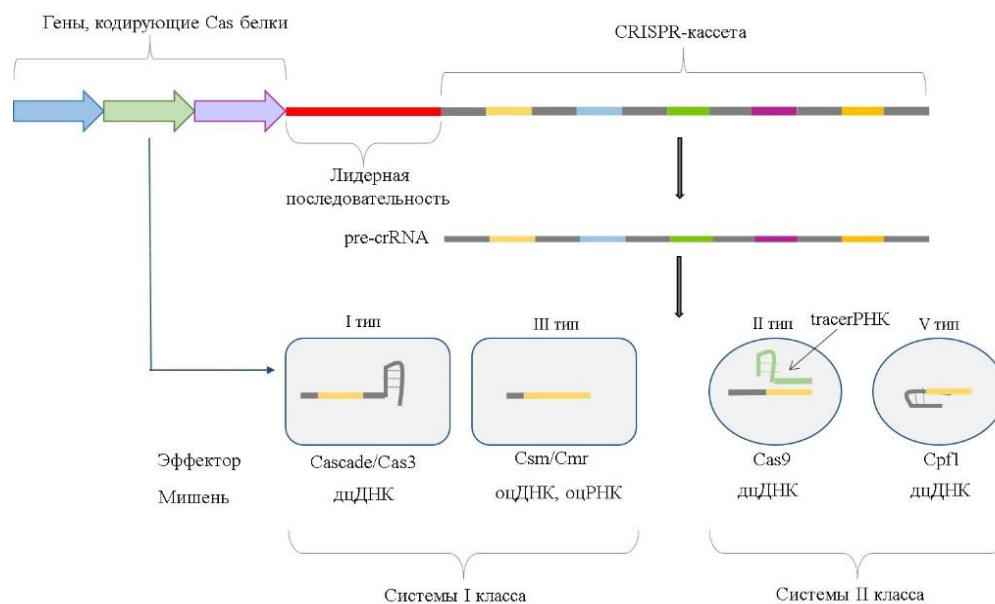


Рис. 6. Типы crRNA разных эффекторных нуклеаз.

- *I mun – Cascade/Cas3*

Рибонуклеопротеиновый комплекс Cascade состоит из пяти белков (Cse1, Cse2, Cas5e, Cas6e, Cas7) и, направляемый crRNA, распознает ДНК-мишень, облегчая рекрутинг геликазы-нуклеазы Cas3. При связывании с протоспейсером в Cascade происходят конформационные изменения, облегчающие изгиб ДНК для деградации ее Cas3 [Westra и др., 2012].

- *III mun – Csm/Cmr*

В III типе CRISPR-Cas систем выделяют 2 подтипа – III-A (Csm) и III-B (Cmr). Csm состоит из пяти субъединиц, Cmr из шести. Оба эффекторных комплекса обладают двойной нуклеазной активностью – расщепляют одноцепочечные РНК и ДНК [Natoum-aslan и др., 2015]. Подробный механизм расщепления ДНК-мишеней системами III типа пока что остается не до конца ясным.

Мультисубъединичные комплексы систем I и III типа очень схожи по структуре, что позволяет предположить их дивергентную эволюцию от одного общего предка [Ryan N. Jackson, 2015].

- *IV mun (Csf)*

Функционирование CRISPR-Cas систем этого типа пока еще мало изучено – экспериментально не показано взаимодействие белковых компонентов с crRNA. Но

имеющиеся данные говорят о том, что эффектор включает в себя четыре белка Csf, подобных субъединицам из Cascade [Özcan и др., 2019].

- *II и V тип на примере белков Cas9 и Cpf1 (Cas12a)*

В системах II и V типа за распознавание и деградацию целевой последовательности отвечает только один белок-эффектор. Механизм интерференции лучше всего исследован у эндонуклеазы Cas9. В состав нуклеазы Cas9 входят два активных сайта – HNH, который расщепляет цепь ДНК, комплементарную crРНК, и RuvC, «разрезающий» другую цепь. Для правильной работы белка необходима помимо crРНК – tracrРНК, которая является важным участником процессинга pre-crРНК и стимулирует нуклеазную активность Cas9 [Jinek и др., 2012]. В отличие от Cas9, Cpf1 требует наличия только crРНК и включает один нуклеазный домен RuvC [Zetsche и др., 2015]. Кристаллографический белок предполагает наличие второго каталитического сайта между мотивами RuvC аналогичного домену HNH – Nuc [Yamano и др., 2016].

1.2.Использование CRISPR-Cas систем в генной инженерии

Технология генного редактирования требует разработки высокоточных инструментов, чтобы внесение программируемых изменений касалось только целевых областей генома. Эта возможность появилась с открытием CRISPR-Cas систем: направляемые короткими РНК нуклеазы индуцируют расщепление в определенном месте геномной последовательности. Мутации, приводящие к изменению аминокислотного состава в активных сайтах эндонуклеаз позволяют создавать белки-никазы, которые вносят «разрез» только в одну цепь ДНК, что снижает возможную мутагенную активность при гомологичной рекомбинации и негомологичном соединении концов [Cong и др., 2013]. Кроме того, разные направляющие последовательности РНК могут быть закодированы в единый CRISPR-массив для того, чтобы одновременно отредактировать несколько сайтов в геноме.

Первым был описан и лучше всего изучен белок Cas9 из бактерии *Streptococcus pyogenes* (SpCas9). Поэтому сегодня он получил широкое применение в качестве инструмента генного редактирования, существенно упростив технологию изменения ДНК. Экспериментально было продемонстрировано, что для расщепления ДНК достаточно экспрессии следующих компонентов: tracr- и crРНК, РНКазы III и эндонуклеазы Cas9. Клонирование CRISPR/Cas системы в клетках *E. coli* предотвращало последующую трансформацию фагом или плазмидой, несущими совпадающие со спейсерами последовательности и PAM [Saprunauskas и др., 2011]. В условиях *in vitro* система также

эффективна. Комплекс РНК-белок эффективно расщепляет плазмидную ДНК, имеющую PAM, но для деградации одноцепочечных молекул ДНК присутствие PAM необязательно [Gasiunas и др., 2012]. Трехкомпонентная система работает и в эукариотических клетках линии НЕК 293FT и требует для нацеливания в ядро двух молекул NLS. Параллельно с этим была создана химерная конструкция с частичным слиянием tracr- и crРНК, которая также успешно направляла белок Cas9 на ДНК-мишени [Cong и др., 2013]. PAM для эндонуклеазы Cas9 – это последовательность NGG, где N – это любой нуклеотид. Расщепление белком Cas9 плазмидной и линейной ДНК сайт-специфично: «разрез» вносится через 3 нуклеотида после PAM. Некомплементарная crРНК цепь ДНК расщепляется по одному или по нескольким сайтам через 3-8 нуклеотидов после PAM [Jinek и др., 2012]. Расщепление осуществляется двумя каталитическими доменами нуклеазы Cas9: HNH и RuvC. HNH домен отвечает за расщепление нити, комплементарной crРНК и представляет собой двуцепочечный антипараллельный β -лист, окруженный четырьмя α -спиралями, в активном центре которого располагаются остатки гистидина 840, мутации по которому приводят к полной дезактивации фермента, аспарагина 863 и 854 и аспартата 839 [Nishimasu и др., 2016]. Предполагается, что расщепление нити ДНК происходит следующим образом: не поделенная электронная пара азота имидазольного кольца гистидина активирует молекулу воды для атаки на фосфатную группу нуклеотида. Далее происходит разрыв 3'-5'-фосфодиэфирной связи. Остальные аминокислотные остатки активного центра координационно связаны с ионом магния, что увеличивает электрофильность атома фосфора. Расщепление нити ДНК, некомплементарной направляющей РНК, происходит за счет RuvC домена. Этот сайт имеет более сложную структуру и состоит из шестицепочечного β -листа, окруженного четырьмя α -спиралями, и дополнительной β -складкой. Активный центр включает аминокислотные остатки гистидина 983, аспартата 10 и 986, глутамата 762. Замена любого из этих аминокислотных остатков приводила к потере функций фермента [Nishimasu и др., 2016]. Механизм расщепления ДНК такой же, как и у домена HNH. Анализ аминокислотного состава активных центров других Cas белков по этим же положениям позволяет оценивать их нуклеазную активность.

Белок Cas9 активен в комплексе с двумя типами РНК: crРНК, содержащей последовательность спейсера, и tracrРНК (транс-активирующей), которая кодирует последовательность прямого повтора. Связываясь с РНК и целевой ДНК, белок претерпевает значительные конформационные изменения. Между доменами белка образуется «канал», в котором происходит комплементарное спаривание оснований

направляющей РНК и ДНК-мишени [Jinek и др., 2014]. Но Cas9 можно запрограммировать, используя гибрид двух направляющих РНК (sgРНК), соединенных линкерной петлей (Рис.7) [Jinek и др., 2012], и упростить систему всего до двух компонентов.

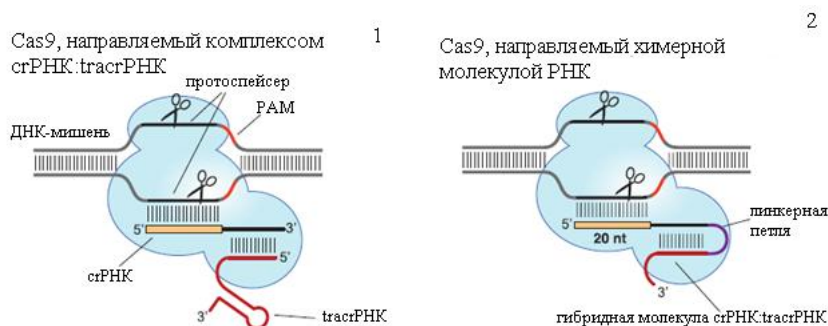


Рис. 7. Схема расщепления ДНК-мишени белком Cas9. (1) В комплексе с двумя направляющими РНК. (2) с гибридной молекулой РНК [Jinek и др., 2012].

Благодаря возможности легко направлять Cas белки на конкретные сайты генома системы CRISPR-Cas9 быстро набрали популярность. Их стали использовать для модификации и редактирования геномов большого количества клеток и организмов. При помощи этой технологии были созданы модельные объекты для изучения некоторых заболеваний. Например, стало возможным искусственно вызывать рак печени у мышей, нацеливая CRISPR-Cas9 на гены-супрессоры опухолей [Хуе и др., 2015]. Также данная методика позволяет анализировать функции генов методом скрининга [Zhou и др., 2014]. Технология имеет большой потенциал в генной терапии, разработке лекарств против ранее неизлечимых заболеваний, модификации растений для развития сельского хозяйства.

Изучение технологии CRISPR-Cas показало, какое огромное влияние она сможет оказать на развитие генной инженерии. Быстрое совершенствование этого инструмента, и его эффективное внедрение поднимают серьезные этические вопросы, которые пока еще остаются неразрешенными.

1.3.Разнообразие Cas белков 2 класса – Cas12a, Cas12b, Cas12c, Cas13a

Поскольку система CRISPR-Cas9 имеет ряд недостатков, большой интерес представляет поиск новых Cas белков, относящихся ко второму классу. Системы первого класса встречаются гораздо чаще, но их не используют в генной инженерии, потому что в клетке гораздо легче собрать компактный «инструмент» из двух или трех компонентов, чем мультисубъединичный белковый комплекс.

Одним из первых был подробно охарактеризован белок FnCas12a (Cpf1) – представитель V типа. Это крупный белок-эндонуклеаза (длиной около 1300 аминокислот). В отличие от SpCas9, FnCas12a направляется только одной молекулой РНК и имеет Т-богатый РАМ (TTN). Также он вносит двуниевые разрывы с образованием выступов, что может облегчить процесс негомологичного соединения концов (Рис.8). FnCas12a выделяют в отдельный тип, так как у него обнаружен только один нуклеазный домен – RuvC, а адаптационный модуль кодирует белки cas1, cas2 и cas4, что больше похоже на системы 1 класса [Makarova и др., 2015]. FnCas12a обладает РНКазной и ДНКазной активностью одновременно. Он самостоятельно процессирует pre-crРНК в зрелые crРНК, содержащие последовательность спейсера и прямого повтора [Fonfara и др., 2016]. Такие crРНК гораздо меньше по размеру, чем дуплекс crРНК-tracrРНК. Проверка активности белков-ортологов Cas12a из разных семейств в условиях *in vitro* показала их активность. Но только лишь 3 из 8 нуклеаз проявляли нуклеазную активность в эукариотических клетках, AsCas12a, LbCas12a, и как позже оказалось FnCas12a [Zetsche и др., 2015]. Эти отличительные особенности Cas12a позволяют избежать трудностей, возникающих при использовании SpCas9.

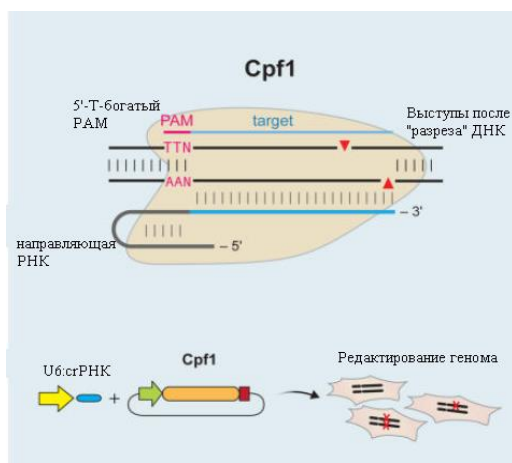


Рис. 8. Схема расщепления белком Cas12a(Cpf1) ДНК-мишени [Zetsche и др., 2015].

При помощи методов биоинформатики были идентифицированы и затем отнесены к эффекторам второго класса группы белков C2c1 (Cas12b), C2c2 (Cas13a), C2c3 (Cas12c) [Shmakov и др., 2015]. Для группы C2c1 предполагается наличие крупного эффекторного белка, а в адаптационном модуле генов cas1, cas2, cas4. Похожую структуру имеют локусы C2c3 – ген cas1 и крупный эффектор, при этом для этих белков описаны необычно короткие спейсеры по 17-18 нуклеотидов. Их доменная организация похожа на Cpf1 наличием RuvC-подобного домена и С-концевой областью. Сходство проявляется и в наличии АТ-богатого РАМ [Shmakov и др., 2015]. Для C2c1 была продемонстрирована его

нуклеазная активность в условиях *in vitro* при инкубации с sgРНК при высоких температурах. Локусы C2c2 состоят только из гена основного белка и массива CRISPR. Возможно, гены белков-участников адаптации находятся в другом месте генома. Анализ аминокислотных последовательностей не показал сходства эффектора C2c2 с другими белками, но были обнаружены нуклеотид-связывающие домены HEPN, обладающие свойствами РНКазы. Авторы выносят их в отдельный VI тип [Shmakov и др., 2015].

Поиск и описание новых белков второго класса важны для понимания их эволюции. Считается, что они происходят от мобильных элементов, а именно транспозаз типа TnpB путем случайной вставки в геном, и являются далекими потомками разных типов белков первого класса.

CRISPR-Cas системы разных бактерий уникальны, и поэтому изучение и характеристика функционирования Cas белков обеспечит универсальность и специфичность инструментов генного редактирования.

1.4. CasX – отдельное семейство редакторов генома

Большинство ферментов CRISPR-Cas систем получают из изолированных бактерий, что ограничивает исследователей. Методами метагеномики можно получить доступ к генетическому материалу огромного количества некультивируемых бактерий. Так были обнаружены две новые системы – CRISPR-CasX и CRISPR-CasY из микробных сообществ подземных вод и отложений [Burstein и др., 2017].

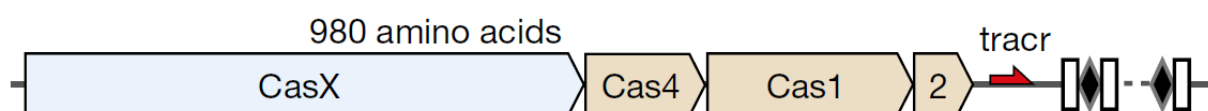


Рис.9. Структура локуса CRISPR-CasX [Burstein и др., 2017].

CRISPR-CasX системы были найдены в бактериях Planctomycetes и Deltaproteobacteria в 2017 году. Они включают в себя белки cas1, cas2, cas4 и новые эффекторные белки CasX, имеющие относительно малый размер (около 980 аминокислотных остатков) (Рис.9). Система содержит также CRISPR-кассету со спейсерными последовательностями длиной 20 нуклеотидов и прямыми повторами длиной 23 нуклеотида. Сравнение с гомологичными последовательностями показало, что С-конец белка проявляет гомологию с RuvC доменом систем V типа (Cpf1), а 630 аминокислотных остатков N-конца не обнаруживают сходства с каким-либо известным белком, что позволяет считать его новым эффектором второго класса (Рис.10).

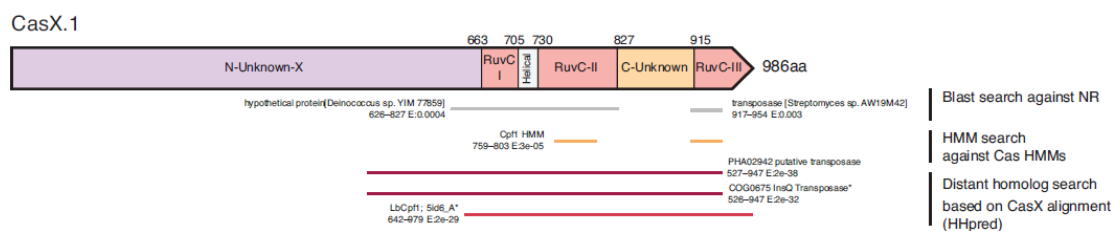


Рис.10. Сравнительный анализ CasX с гомологичными белками Cas [Burstein и др.,

Но CasX нельзя отнести к системам V типа, поскольку для них не характерно сочетание *tracrRNA* и отдельных белков *cas1*, *cas2*, *cas4*. Кроме того, последовательность *cas1* филогенетически далека от других подобных белков среди CRISPR-Cas систем V типа [Burstein и др., 2017]. Также размер белка CasX (980 аминокислот) нетипично мал по сравнению с белками данного типа (*Cpf1*, *C2c1* и *C2c3*) длиной 1200 аминокислот. Активность системы была показана в клетках *E. coli*: плазмиду, включающую последовательность гена CasX и короткий массив CRISPR, трансформировали в бактериальные клетки. Экспрессия CRISPR-CasX системы блокировала последующую трансформацию плазмидой, содержащей ДНК-мишень [Burstein и др., 2017]. РАМ-скрининг, проведенный Burstein и коллегами, показал строгое предпочтение белком последовательности TTCN на 5'-конце протоспейсера.

Несмотря на показанную активность CRISPR-CasX системы в бактериальных клетках, в 2017 году было неизвестно какие минимальные компоненты системы необходимы для ее работы в условиях *in vitro*, а следовательно и в эукариотических клетках. Поэтому в данной работе было решено синтезировать *in vitro* все компоненты CRISPR-CasX системы и проверить нуклеазную активность комплекса *DpbCasX* – *tracrRNA* – *crRNA in vitro*. В процессе нашей работы (февраль 2019 года) вышла статья группы Дженнифер Дудны, посвященная кристаллографической структуре белка CasX [Liu и др., 2019], где показали активность белка CasX *in vitro*, опередив наши исследования, но в то же время дав новые данные для продолжения работы. Проведенный в ней филогенетический анализ предполагает, что CasX возник из транспозазы типа *TrnB* путем его независимой инсерции в локус CRISPR. Необычная структура может сулить особый молекулярный механизм работы белка по сравнению с остальными эндонуклеазами второго класса. Чтобы проверить это предположение система CRISPR-CasX из бактерий *Planctomycetes* и *Deltaproteobacteria* (далее *Plm* и *Dpb*, соответственно) была создана в условиях *in vitro*. Дополнительно, при помощи криоэлектронной микроскопии были представлены модели четвертичной структуры белка [Liu и др., 2019]. *DpbCasX* расщепляет двуцепочечную ДНК, комплементарную гибридной направляющей

РНК. Эффективность расщепления снижается при инкубировании белка с одноцепочечной ДНК, что несвойственно для нуклеаз V типа, тоже имеющих всего один активный сайт –RuvC. Например, активность белка Cpf1 увеличивается в несколько раз именно при связывании одноцепочечной ДНК. Сравнение последовательностей RuvC доменов среди Cas белков помогло обнаружить потенциальный состав активного центра CasX: Asp672, Glu769 и Asp935. Мутирование по этим сайтам приводит к дезактивации белка. Помимо этого, была продемонстрирована активность CRISPR-CasX системы из бактерий Deltaproteobacteria и Planctomycetes в эукариотических клетках. DpbCasX с ограниченной эффективностью разрушает целевой ген флуоресцентного белка GFP в клетках эмбриональной почки человека HEK293T. PlmCasX – проявляет более высокую активность и расщепляет ДНК примерно в 30% случаев (например, для широко используемого SpCas9 эффективность составляет в среднем 55%).

Тройной комплекс белок-РНК-мишень анализировали при помощи криоэлектронной микроскопии, чтобы понять, как CasX связывает ДНК. При этом был идентифицирован доменный состав белка CasX (Рис.11). Эффекторный домен RuvC и олигонуклеотид связывающий домен OBD – элементы, характерные для Cas белков. С разделенными цепями ДНК в комплексе были обнаружены два новых сайта: NTSB и TSL, которые связывают не комплементарную направляющей РНК и комплементарную цепи ДНК соответственно. Структура TSL указывает на то, что этот домен возможно участвует в разделении цепей ДНК. Но анализ белковой конструкции в отсутствие домена NTSB показал, что именно этот сайт отвечает за «раскручивание» ДНК. Helical I и II – соединяют домены между собой [Liu и др., 2019].

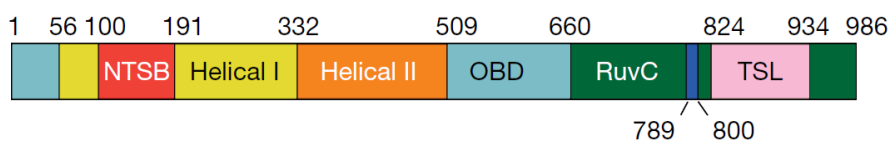


Рис.11. Доменный состав CasX [Liu и др., 2019].

CasX может вызывать деградацию ДНК в двух конформационных состояниях. В первом цепь ДНК, не комплементарная гидовой РНК, находится в активном сайте RuvC, а вторая удерживается Helical I и II. RuvC расщепляет связанную цепь, затем смещается и может расщепить вторую (Рис.12).

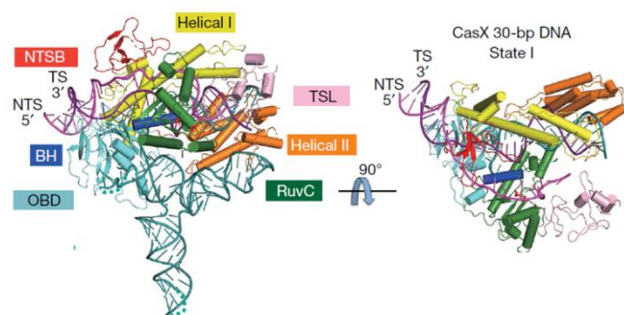


Рис.12. Конформация I тройного комплекса CasX-sgРНК-мишень ДНК, где NTS – цепь, не комплементарная sgРНК, TS – комплементарная [Liu и др., 2019].

Во втором дуплекс ДНК-sgРНК принимает изогнутую форму, благодаря чему RuvC напрямую связывается с цепью ДНК, комплементарной sgРНК (Рис.13).

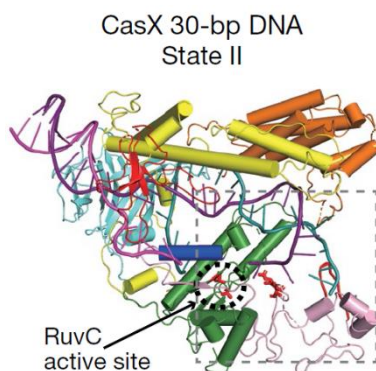


Рис.13. Конформация II тройного комплекса CasX-sgРНК-мишень ДНК [Liu и др., 2019].

Экспериментально продемонстрировано, что конформация I оказывается предпочтительнее в 70% случаев.

Таким образом, белок CasX из-за своей уникальной структурной организации, происхождения и образования специфического тройного комплекса может быть вынесен в отдельное семейство белков CasX второго класса, несмотря на сильную схожесть с Cas-нуклеазами V типа. Компактный размер и высокая эффективность его работы как в *in vitro*, так и *in vivo* условиях в бактериальных и эукариотических клетках делает CasX перспективным объектом для изучения в качестве нового инструмента геномного редактирования.

2.2. Среды и растворы, использованные в работе

LB (Luria-Bertani): 10г триптон (VWR Life Science, США), 5г дрожжевой экстракт (Helicon, Россия), 10г NaCl (Helicon, Россия), смешивали в 1л дистиллированной H₂O и автоклавировали при 121°C 60 мин.

TB (Terrific Broth): 12г триптона (VWR Life Science, США), 24г дрожжевого экстракта (Helicon, Россия), 4 мл глицерола (ITW Reagents, США) растворили в 900 мл дистиллированной H₂O. Отдельно приготовили 100 мл фосфатного буфера (0.17МКН₂РO₄, 0.72М К₂НРO₄ (Helicon, Россия)), рН 7.5. Оба раствора автоклавировали при 121°C 60 мин. После, охладив емкости, добавили стерильный фосфатный буфер в раствор.

Буферные растворы для проведения электрофореза:

ТАЕ: 2М Трис (VWR Life Science, США), 1М уксусная кислота (Helicon, Россия), 50ММ ЭДТА (Helicon, Россия), рН 8.4.

Трис-глициновый буфер: 0.025М Трис (VWR Life Science, США), 0.192М глицин (Helicon, Россия), рН 8.3.

Буферы для проведения хроматографической очистки белка:

Буфер А: 500ММ NaCl (Helicon, Россия), 50ММ HEPES (Helicon, Россия), 10% глицерол (ITW Reagents, США), 1ММ β-меркаптоэтанол (βМэ, VWR Life Science, США), рН 7.5.

Буфер Б: 500ММ NaCl (Helicon, Россия), 300ММ имидазол (Acros Organic, США), 50ММ HEPES, 1ММ βМэ (VWR Life Science, США), 10% глицерол (ITW Reagents, США), рН 7.5.

Буфер В: 300ММ NaCl (Helicon, Россия), 50ММ HEPES (Helicon, Россия), 10% глицерол (ITW Reagents, США), рН 7.5.

Буферный раствор для проведения in vitro реакции:

Реакционный буфер: 20ММ HEPES (Helicon, Россия), 10ММ MgCl₂ (Helicon, Россия), 150ММ KCl (Helicon, Россия), 1% глицерол (ITW Reagents, США), рН 7.5.

2.3. Гель-электрофорез в агарозном геле и ПААГ

Электрофорез в 1,5% агарозном геле

Для приготовления геля 1,5 г сухой агарозы (Helicon, Россия) добавляли в 100 мл ТАЕ 1х буфера. Взвесь агарозы нагревали в микроволновой печи до полного ее растворения. После охлаждения раствора добавляли 3 мкл бромистого этидия (Helicon, Россия) и тщательно перемешивали. Полученную смесь заливали на подложку для геля и оставляли застывать. Затем гель переносили в камеру для горизонтального электрофореза SE-1, добавляли достаточное количество буфера ТАЕ 1х и наносили маркер и исследуемые образцы в лунки. Подавали напряжение 50В. Анализ изображения проводили при помощи системы гель-документирования Chemidoc MP System (Bio Rad, США) с программным обеспечением Imagelab.

Электрофорез в ПААГ в денатурирующих условиях по Лэммли [Lämmli, 1970]

Первым этапом было приготовление 10% разделяющего геля. Для этого смешивали компоненты (см. Табл.3) в фальконе, заливали между специальными стеклами, оставляли застывать. Затем смешивали в фальконе компоненты для приготовления 6% концентрирующего геля (см. Табл.3) и заливали между стеклами, поместив гребенку шириной 1см, оставляли застывать гель. После стекла помещали в камеру для вертикального электрофореза Mini-PROTEIN Tetra и добавляли достаточное количества Трис-глицинового буфера 1х с добавлением 1% SDS. Далее наносили пробы, которые предварительно инкубировали 5 мин при температуре 95°C в термоциклере T100 Thermal Cycler (Bio-Rad, США) с добавлением 10% SDS и βМэ. Подавали напряжение 40В в концентрирующем геле и 100В в разделяющем. Окрашивание геля проводили красителем Coomassie Brilliant Blue G-250 (Thermo Fisher Scientific, США). Краситель отмывали 10% раствором уксусной кислоты (Helicon, Россия) и затем дистиллированной водой. Анализ изображения проводили при помощи системы гель-документирования Chemidoc MP System (Bio Rad, США) с программным обеспечением Imagelab.

Табл. 3. Состав 10% разделяющего геля и 6% концентрирующего.

Компонент	Разделяющий 10% гель	Концентрирующий 6% гель
Дистиллированная вода	2 мл	1,4 мл
АА:бисАА 29:1 (Helicon, Россия)	1,67 мл	330 мкл
Трис (VWR Life Science, США) 1,5М, рН=8.8 и рН=6.8 соответственно	1,25 мл	253 мкл
SDS 10% (Helicon, Россия)	50 мкл	20 мкл
APS 10% (Helicon, Россия)	50 мкл	20 мкл
TEMED (Helicon, Россия)	3 мкл	2 мкл

2.4. Молекулярное клонирование, методы работы с бактериями и плазмидами

Последовательность гена белка CasX была заказана на ресурсе Addgene.

Конструирование плазмидных векторов pET21a CasX и pET21a CasX_MBP было проведено при помощи программы Benchling.

Клонирование гена интереса в плазмидном векторе включает следующие этапы:

А) ПЦР геномной последовательности белка CasX

Аmplification последовательности генов, кодирующих белок CasX и белок CasX_MBP, осуществляли при помощи Takara Taq Premix в количестве 5 единиц активности на микролитр с 50 нанограммами геномной ДНК, общий объем реакционной смеси – 25 мкл. Для ПЦР взяли пары праймеров BP463-BP463 и BP493-BP494. Далее пробы инкубировали в термоциклере T100 Thermal Cycler (Bio-Rad, США) согласно следующей программе: (98°C – 10с, 55°C – 10с, 72°C – 30с) – 30 раз. Визуализацию продуктов реакции проводили методом гель-электрофореза ДНК в 1,5% агарозном геле. Затем ДНК выделяли при помощи набора GeneJET Gel Extraction Kit (Thermo Fisher Scientific, США). Концентрацию ДНК измеряли на спектрофотометре NanoDrop2000C (Thermo Scientific, США) при длине волны 280нм по формуле Бугера-Ламберта-Бера.

Б) Рестрикция плазмидного вектора pET21a

Рестрикцию плазмиды производили ферментами NdeI (Thermo Fisher Scientific, США) и XhoI (Thermo Fisher Scientific, США) в количестве 10 единиц активности с 1 мкг плазмидной ДНК в течение 40 минут при температуре 37°C в термоциклере T100 Thermal Cycler (Bio-Rad, США), общий объем реакционной смеси 20 мкл. Для контроля полноты прохождения рестрикции, методом гель-электрофореза в 1,5% агарозном геле анализировали пробы до и после рестрикции. Далее рестрицированную плазмиду выделяли при помощи набора GeneJET Gel Extraction Kit (Thermo Fisher Scientific, США). Концентрацию плазмидной ДНК измеряли на спектрофотометре NanoDrop2000C при длине волны 280нм формуле Бугера-Ламберта-Бера.

В) Сборка полученных фрагментов ДНК методом Гибсона

Праймеры для амплификации последовательности белка CasX и CasX_MBP подбирали таким образом, чтобы полученный ПЦР-фрагмент и плаزمида pET21a имели перекрывающиеся концевые участки длиной около 20 нуклеотидов, включающие в себя сайты рестрикции, по которым будет производиться вставка. Линеаризованную плазмиду и ПЦР-фрагмент смешивали в соотношении 1:3 и добавляли такой же объем фермента NeBuilder Hi-Fi DNA Assembly Master Mix (NEB, США). Реакционную смесь инкубировали 60 минут при 50°C. После чего трансформировали ее в бактериальные клетки и высевали на среду с антибиотиком ампициллином.

Г) Трансформация бактерий *Escherichia coli*

Аликвоту компетентных клеток *Escherichia coli* штамма DH5α размораживали на льду. Затем с помощью микропипетки вносили 5 мкл реакционной смеси, аккуратно перемешивали носиком пипетки и оставляли на льду. Спустя 15 мин подвергали клетки тепловому шоку при 42°C. По истечении минуты пробирки снова помещали в лед еще на 2 минуты. Потом добавляли 300 мкл среды LB и клетки выращивали в течение 45 минут при температуре 37°C в термошейкере Multitron Pro (INFORS HT, Швейцария), выставив его скорость в 500 оборотов в минуту. Высевали 150 мкл полученной смеси на чашки Петри с агаризованной LB средой, содержащей антибиотик ампициллин 1 мг/мл, и инкубировали в течение ночи при 37°C. Чашки с бактериями хранили при температуре 4°C.

Д) Рестрикционный анализ

Результаты клонирования оценивали при помощи рестрикционного анализа. Для этого с помощью микропипетки отбирали бактериальные колонии-трансформанты и переносили в фалькон, наполненный 5 мл среды LB, содержащей антибиотик ампициллин в концентрации 1 мг/мл, и инкубировали в термошейкере Multitron Pro (INFORS HT, Швейцария) в течение ночи при 37°C. На следующий день клеточную культуру центрифугировали на центрифуге Centrifuge 5804 R (Eppendorf, Германия) в течение 15 мин при 3024g. Из клеточного осадка выделяли плазмиду pET21a CasX и pET21a CasX_MBP при помощи GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific, США). Концентрацию плазмидной ДНК измеряли на спектрофотометре NanoDrop2000C (Thermo Scientific, США при длине волны 280нм по формуле Бугера-Ламберта-Бера.

Далее полученные плазмиды рестрицировали при помощи ферментов Eco72I (Thermo Fisher Scientific, США) и Eco32I (Thermo Fisher Scientific, США) для pET21a CasX и Eco72I (Thermo Fisher Scientific, США) для pET21a CasX_MBP. Ферменты в количестве 10 единиц активности инкубировали с 200 нг плазмидной ДНК в течение 40 минут при температуре 37°C в термоциклере T100 Thermal Cycler (Bio-Rad, США), общий объем реакционной смеси 20 мкл. Визуализацию продуктов рестрикции проводили методом гель-электрофореза в 1,5% агарозном геле, анализировали пробы до и после рестрикции.

2.5. Синтез белка в клетках-продуцентах *Escherichia coli* штамма Rosetta (DE3) и анализ растворимости белка

Выделенные плазмиды pET21a CasX и pET21a CasX_MBP трансформировали аналогичным образом в клетки-продуценты *E.coli* штамма Rosetta (D3) для последующего синтеза белка. Переносили колонии в фалькон, наполненный 5 мл среды LB, содержащей антибиотик ампициллин 1 мг/мл, и инкубировали в термошейкере Multitron Pro (INFORS HT, Швейцария) в течение ночи при 37°C. После клеточную культуру помещали в стеклянную колбу с 500 мл среды TB и инкубировали в термошейкере Multitron Pro (INFORS HT, Швейцария) при 37°C до тех пор, пока оптическая плотность бактериальной культуры не достигла 0,6 при длине волны 600нм. Измерение проводили на спектрофотометре Cary 50 UV-Vis (Varian, Австралия). Отбирали пробу 1 мл из колбы для электрофоретического анализа. Затем в колбы добавляли 1mM индуктор лактозного оперона IPTG (Thermo Fisher Scientific, США). Клеточные культуры с добавленным индуктором инкубировали в течение ночи при 16°C. На следующий день отбирали 2

пробы объемом 1 мл для оценки растворимости белка и электрофоретического анализа. Остальное содержимое колбы откручивали на центрифуге HERMLE Z 513 K (Hermle Labortechnik, Германия) 30 мин при 3500g при температуре 4°C. По завершению центрифугирования в предварительно охлажденные фальконы объемом 50 мл перелили 40 мл супернатанта, которыми далее ресуспензировали на льду клеточный осадок. Полученную суспензию откручивали на центрифуге HERMLE Z 513 K (Hermle Labortechnik, Германия) в течение 15 мин при 3500g при температуре 4°C. Супернатант удаляли и осадок хранили при -20°C.

Оценка растворимости белка

Взятую для анализа пробу откручивали на центрифуге MiniSpin (Eppendorf, Германия) 6 мин при 12000g для осаждения клеток. Затем добавляли 1 мл буфера для лизиса клеток (буфер А) и центрифугировали при тех же параметрах. Из супернатанта и разбавленного буфером А осадка отбирали аликвоты по 12 мкл и анализировали методом гель-электрофореза в ПААГ по Лэммли.

2.6. Выделение и хроматографическая очистка рекомбинантного белка

- *Подготовительный этап*

Все компоненты для приготовления буферных растворов смешивали в 250 мл дистиллированной воды и охлаждали до 4°C и при этой же температуре устанавливали pH 7,5 с использованием растворов 1mM HCl и 10mM NaOH. Затем буферный раствор переносили при помощи фильтра BIOFIL FPE-214-500 с порами 0,22 мкм в стерильную емкость и на 20 мин помещали в УВ-ванну для дегазации.

Осадок клеток размораживали на льду, добавляли 15 мл буфер для лизиса клеток (буфер А), ресуспензировали, добавляли лизоцим (Helicon, Россия) в концентрации 1мг/мл для разрушения клеточных стенок бактерий и оставляли на 20 мин на льду. Затем суспензию соницировали на ультразвуковом гомогенизаторе Labsonic P (Sartorius, Германия) с амплитудой 30% и импульсным циклом 0,2 в течение получаса с интервалом 30 сек, откручивали на центрифуге Centrifuge 5810 R (Eppendorf, Германия) при 16000g 40 мин при температуре 4°C. Полученный лизат на льду фильтровали через фильтрующую шприцевую насадку PES Biofil с порами 0,45 мкм и сразу наносили на Ni HisTrap HP колонку.

- *Металл-аффинная хроматография на колонке Ni HisTrap HP*

Хроматографическую очистку рекомбинантного белка проводили на хроматографе низкого давления ÄKTApurifier (GE Healthcare, США). Перед установлением колонки

насосы и помпы промывали буфером А. Далее устанавливали в хроматограф колонку Ni HisTrap HP (7x25 мм, объем 1 мл, средний размер частиц 34 мкм), уравнивали буфером А и наносили лизат, содержащий рекомбинантный белок, со скоростью 1 мл/мин. Затем колонку дополнительно промывали еще 15 мл буфера А. Элюировали белки, связанные с Ni-подложкой, используя градиент концентрации имидазола. Подавали по 20 мл сначала 7% концентрацию буфера Б, а после 100% концентрацию буфера Б со скоростью 1 мл/мин и собирали фракции по 1мл. Полученные фракции откручивали на центрифуге Centrifuge 5415 R (Eppendorf, Германия) в течение 30 мин при 16000g, предварительно отобрав аликвоты по 12 мкл для электрофоретического анализа в 10% ПААГ.

- *Гель-фильтрационная хроматография на колонке Superdex 200 10/30 GL*

Устанавливали в хроматограф колонку Superdex 200 10/30 GL (30см x 10мм, объем 24 мл, средний размер частиц 13 мкм), уравнивали буфером А, наносили супернатант из открученных фракций. Подавали на колонку 20 мл буфера В со скоростью 0,5 мл/мин и собирали фракции объемом 600 мкл. Из отобранных фракций после гель-фильтрационной хроматографии отбирали аликвоты по 12 мкл для электрофоретического анализа, оставшееся содержимое помещали в концентратор типа Центрикон и откручивали на центрифуге Centrifuge 5810 R (Eppendorf, Германия) 15 мин при 5000g при 4°C. После чего разделяли весь объем белкового раствора на аликвоты объемом 10 мкл и замораживали с использованием жидкого азота. Замороженные аликвоты хранили при -80°C. Аликвоты полученных фракций анализировали методом электрофореза в ПААГ по Лэммли.

- *Измерение концентрации белка*

Подготавливали два типа проб:

- Для построения калибровочной кривой, использовали раствор BSA с известным содержанием белка: 0.5, 1, 2, 3 и 4 мкг.
- Для определения концентрации целевого белка использовали аликвоты объемом 0.5, 1, 2 и 3 мкл.

Все 9 проб смешивали анализировали методом гель-электрофореза в ПААГ по Лэммли. Концентрацию белка рассчитывали методом денситометрии при помощи системы геледокументирования Chemidoc MP System (Bio Rad, США) с программным обеспечением Imagelab.

2.7. Синтез направляющей РНК

sgРНК	GGCGCGUUUAUCCAUUACUUUGGAGCAGUCCAGCGACUAUGUCGUAUGGA CGAAGCGCUUAUUUAUCGGAGAGAAACCGAUAAGUAAAACGCAUCAAAAGUCC UGCAGCAGAAAAUCAAA
-------	--

На основе известной последовательности sgРНК были сконструированы праймеры для синтеза направляющей РНК – ВР642 и ВР744 – с помощью Hi-Scribe Kit T7 (NEB, США).

Этап I. Синтез матрицы ДНК методом ПЦР с использованием праймеров ВР642 и ВР744.

Этап II. Транскрипция на основе T7 RNA polymerase (NEB, США) и смеси нуклеотидов Synthesis Kit. Смешав компоненты в пробирке, инкубировали в термоциклере T100 Thermal Cycler (Bio-Rad, США) 4 часа при 37°C. По завершению процесса добавили 1мкл DNase I (NEB, США) для разрушения матричной цепи и инкубировали в течение 30 мин при 37°C.

Этап III. Синтезированную РНК выделяли при помощи набора RNA Clean & Concentrator (Zymo Reseach, США).

Этап IV. Измеряли концентрацию на спектрофотометре NanoDrop2000C (Thermo Scientific, США при длине волны 280нм по формуле Бугера-Ламберта-Бера.

2.8. Синтез ДНК-мишеней

Для синтеза первой ДНК-мишени использовали праймеры ВР641 и Amp-R. ПЦР производили при помощи полимеразы Encyclo (Евроген, Россия) в термоциклере T100 Thermal Cycler (Bio-Rad, США) согласно следующей программе: 95 °C – 3 мин, (95°C – 30 с, 63 °C – 30 с, 72 °C – 1 мин) – 32 раза, 72 °C – 5 мин. Затем проводили рестрикцию синтезированного фрагмента, инкубируя его с рестриктазой NdeI (Thermo Fisher Scientific, США) в количестве 10 единиц активности с 2мкг ДНК в течение 40 минут при температуре 37°C, общий объем реакционной смеси 20мкл. Для синтеза второй ДНК-мишени брали праймеры ВР641 и ВР662. ПЦР производили также при помощи полимеразы Encyclo (Евроген, Россия) в термоциклере T100 Thermal Cycler (Bio-Rad, США) согласно следующей программе: 95 °C – 3 мин, (95°C – 30 с, 63 °C – 30 с, 72 °C – 1 мин) – 32 раза, 72 °C – 5 мин.

Анализ продуктов реакций проводили методом гель-электрофореза ДНК в 1,5% агарозном геле. Затем ДНК выделяли при помощи набора GeneJET Gel Extraction Kit (Thermo Fisher Scientific, США). Концентрацию ДНК измеряли на спектрофотометре NanoDrop2000C (Thermo Scientific, США) при длине волны 280нм по формуле Бугера-Ламберта-Бера.

2.9. Проведение эндонуклеазной реакции *in vitro*

Полученные белок и sgРНК инкубировали в соотношении 1:1, 1:5, 1:20 с ДНК-мишенью в термоциклере T100 Thermal Cycler (Bio-Rad, США) 90 минут при 37 °С. Через 1,5 ч добавляли в каждую пробирку по 1мкл рибонуклеазы А (NEB, США) и инкубировали 15 мин при 37 °С. Далее добавляли по 1 мкл протеиназы К (NEB, США), инкубировали 30 мин при 37 °С. Анализ продуктов реакций проводили методом гель-электрофореза ДНК в 1,5% агарозном геле.

3. Результаты

В ходе данной работы были собраны компоненты для проведения *in vitro* реакции: белок CasX, sgRNA и ДНК-мишень.

3.1 Конструирование плазмидного вектора для экспрессии системы рЕТ21а CasX

Для получения рекомбинантного белка DpbCasX в бактериальном векторе рЕТ21а был клонирован ген белка CasX из Deltaproteobacteria.

Плазида рЕТ21а CasX должна включать (Рис. 14):

- Ген устойчивости к антибиотику ампициллину – селективный маркер для отбора клонов (ингибирует действие бета-лактамазы)
- Промотор гена устойчивости к ампициллину (AmpR)
- Последовательность гена DpbCasX
- Последовательность, кодирующую несколько аминокислотных остатков гистидина (His-Tag) на конце целевого гена, необходимую для последующей металл-аффинной хроматографии. Аминокислотные остатки обратимо взаимодействуют с Ni, содержащемуся в наполнителе колонки, что позволит отделить белок от остальных примесей, выделяющихся из бактерий-продуцентов.
- Вектор рЕТ21 а содержит ориджин репликации рBR322, обеспечивающий среднюю копияность плазмиды внутри бактериальной клетки

- Целевой ген DpbCasX находится под контролем промотора T7 (последовательность, заимствованная из генома бактериофага T7). T7 промотор обеспечивает эффективный синтез целевого белка в клетках-продуцентах.
- Последовательность LacO лактозного оперона, регулируемая репрессором lacI, расположенную между целевым геном и T7 промотором. Связываясь с LacO, последовательность lacI препятствует продвижению РНК-полимеразы и транскрипции гена интереса. Добавление IPTG (аналога лактозы) позволяет освободить LacO последовательность от ингибитора, что приводит к синтезу белка. Такая конструкция индуцирует синтез целевого продукта только при достижении оптимального количества клеток-продуцентов в культуре.
- Перед последовательностью, кодирующей DpbCasX, находится RBS (ribosomal binding site) – сайт связывания рибосом (при последующей трансляции мРНК)
- После гена DpbCasX находится терминатор транскрипции бактериофага T7

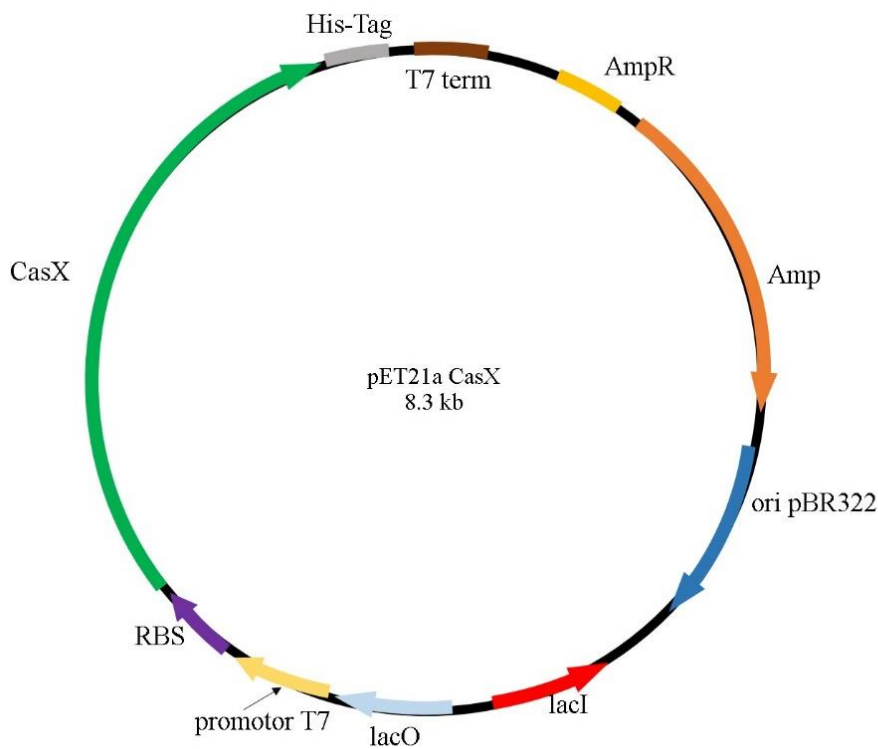


Рис.14. Карта плазмиды pET21a CasX

3.2 Клонирование рЕТ21а_CasX

Для получения плазмиды рЕТ21а_CasX была проведена вставка последовательности, кодирующей DpbCasX, в плазмидный вектор рЕТ21а. Для этого был получен ПЦР-фрагмент последовательности гена DpbCasX длиной 3000 пар нуклеотидов. Для линейаризации плазмиды рЕТ21а была проведена ее рестрикция с помощью ферментов NdeI и XhoI. Далее эти двуцепочечные фрагменты ДНК сшивали методом Гибсона при помощи фермента NEBuilder HiFi DNA Assembly Master Mix. Продукты реакции Гибсона трансформировали в бактериальные клетки *Escherichia coli* штамма DH5α и высевали на чашки Петри с агаризованной питательной средой и ампициллином. В качестве контроля использовали клетки, в которые были трансформированы плазмиды, также рестрицированные и обработанные ферментом реакции Гибсона, но не содержащие фрагмент гена CasX. Из выросших колоний выделяли плазмидную ДНК. Для проверки наличия необходимой вставки был проведен рестрикционный анализ с использованием ферментов Eco72I и Eco32I. На геле ожидали увидеть фрагменты размером 1900, 1700 и 4700 пар нуклеотидов.

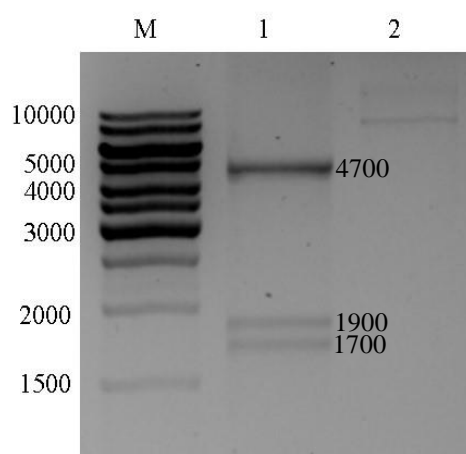


Рис.15. Электрофореграмма продуктов рестрикции в 1,5% агарозном геле плазмиды рЕТ21а CasX ферментами Eco72I и Eco32I, где

М-маркер 1kb Ladder; 1-рЕТ21а CasX, обработанная Eco72I, Eco32I; 2-рЕТ21а CasX до рестрикции;

Как видно из рисунка 15 плазида рЕТ21а_CasX, обработанная ферментами Eco72I и Eco32I, расщепляется на фрагменты ожидаемого размера.

Для дополнительной проверки, плазида рЕТ21а_CasX была отправлена в компанию Евроген на секвенирование методом Сэнгера. Секвенирование, также, как и

рестрикторный анализ, подтвердило получение плазмидного вектора pET21a_CasX с правильно ориентированной вставкой.

3.3 Синтез белка DpbCasX в клетках-продуцентах *E. coli* штамма Rosetta(DE3) и анализ его растворимости

Полученный вектор pET21a_CasX был трансформирован в клетки *E. coli* штамма Rosetta (DE3), которые служат продуцентами рекомбинантного белка. Далее культуру клеток выращивали в среде LB с ампициллином до достижения оптической плотности OD = 0.6. Затем к клеткам добавляли индуктор лактозного оперона IPTG, который так же, как и аллолактоза связывается с lac-репрессором. Благодаря этому гены lac-оперона могут транскрибироваться за счет освобождения T7 промотора от репрессора для посадки РНК-полимеразы. Отличие IPTG от аллолактозы состоит в том, что из-за наличия атома серы в его молекуле соединение не гидролизуется, то есть концентрация индуктора в клетке остается постоянной. После роста клеток и синтеза ими белка отбирали аликвоты по 1 мл, осаждали клетки центрифугированием при 3000g, добавляли к ним буфер А, откручивали на центрифуге при 16000g для осаждения крупных агрегатов и отбирали растворимую часть (супернатант). Осадок растворяли в буфере А. Пробы анализировали в 10% ПААГ. Ожидаемый размер белка – 113 кДа. Из рисунка 16 видно, что проба нерастворимой части клеточного лизата характеризуется значительной гетерогенностью, но в ней присутствует доминирующая белковая полоса, соответствующая молекулярной массе около 113 кДа. В растворимой части лизата, напротив, полоса, соответствующая целевому белку отсутствует, что говорит о нерастворимости DpbCasX в данных условиях его синтеза и лизиса клеток-продуцентов. Для повышения растворимости белка были опробованы различные составы буферов для лизиса: варьировались рН, ионная сила растворов, использовался глицерол и Triton X100. Однако при изменении условий выделения DpbCasX все равно детектировался в осадке.

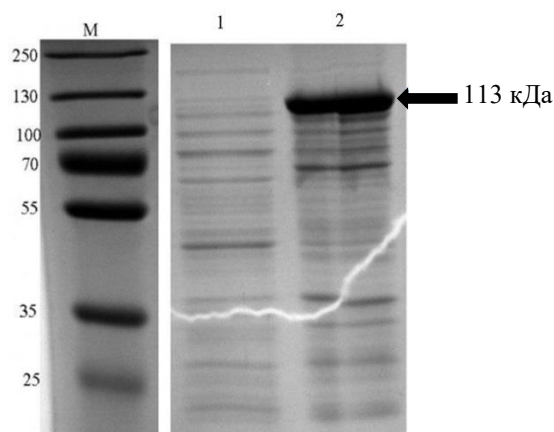


Рис. 16. Электрофореграмма в ПААГ супернатанта и осадка клеточного лизата после индукции синтеза белка DpbCasX, где

М-маркер PageRuler, 1-рЕТ21а CasX супернатант; 2-рЕТ21а CasX осадок;

Поэтому было решено прибегнуть к созданию генетической конструкции, несущей ген МВР слитый с последовательностью гена DpbCasX. Экспрессия этой системы может увеличить растворимость целевого белка.

3.4 Конструирование плазмидного вектора с последовательностью гена МВР и повторное клонирование

Для получения вектора рЕТ21а CasX_МВР было проведено клонирование ПЦР-фрагмента последовательности кодирующей белок МВР в ранее созданный в этой работе плазмидный вектор рЕТ21а_CasX (Рис.17). Сначала был получен ПЦР-продукт последовательности белка МВР с матрицы рЕТ LIC МВР (имеющийся в лаборатории коммерческий вектор), затем плазида рЕТ21а_CasX была линейаризована по сайту NdeI и при помощи фермента NEBuilder HiFi DNA Assembly Master Mix были сшиты данные ДНК-фрагменты. Реакционная смесь была трансформирована в бактериальные клетки *Escherichia coli* штамма DH5 α . С чашки был взят клон для проверки рестрикцией. Для этого использовали рестриктазу Eco72I. Размер ожидаемых продуктов – ок. 1700 п.н. и 7700 п.н. (Рис.18).

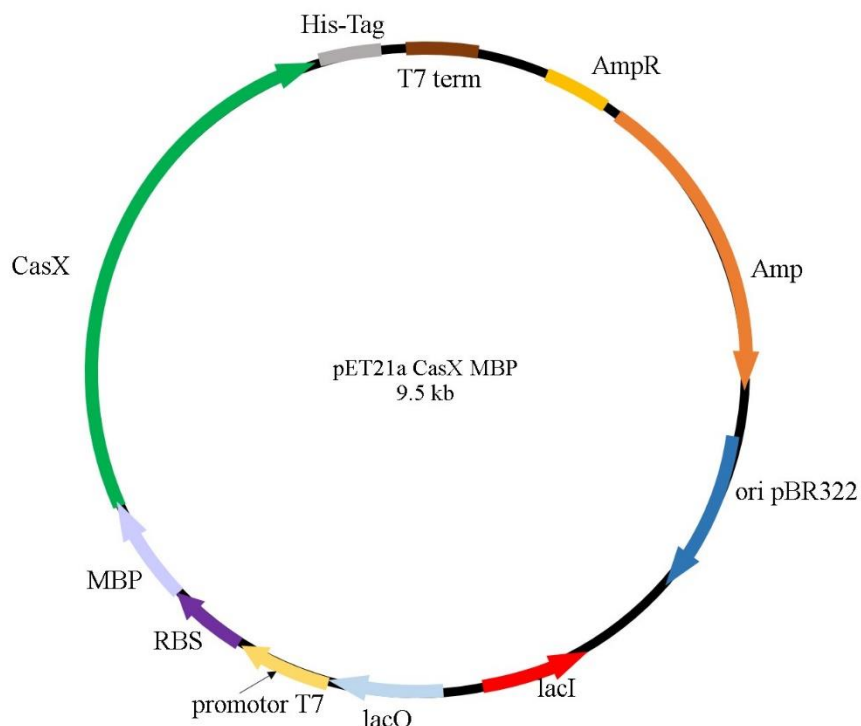


Рис.17. Карта плазмиды pET21a CasX MBP

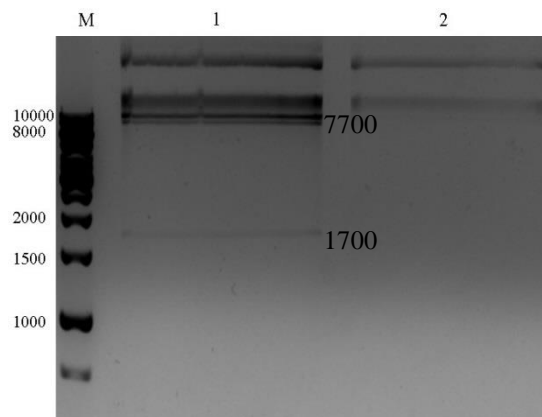


Рис.18. Электрофореграмма продуктов рестрикции плазмиды pET21a CasX_MB P рестриктазой Eco72I в 1,5% агарозном геле, где

M-маркер 1kb Ladder, 1- плазида pET21a CasX_MB P, обработанная ферментом Eco72I; 2-контроль, плазида pET21a CasX_MB P

3.5 Выделение рекомбинантного белка DpbCasX_MBP и его хроматографическая очистка

Добавление MBP увеличило растворимость рекомбинантного белка CasX, что позволило перейти к следующему шагу работы – его выделению из клеток-продуцентов и очистке. Для синтеза белка DpbCasX_MBP вектор pET21a CasX MBP был трансформирован в клетки *E. coli* штамма Rosetta. Проверку эффективности синтеза белка проводили на двух разных клонах, взятых с чашки после трансформации. Культуры клеток выращивали в среде ТВ до достижения ими оптической плотности OD=0.6. Далее к клеткам добавляли с индуктор лактозного оперона, предварительно отобрав пробы до начала индукции. После увеличения биомассы клеток были отобраны аликвоты для электрофоретического анализа в 10% ПААГ. Ожидаемый размер белка DpbCasX_MBP – 155.5 кДа, поскольку молекулярная масса белка MBP составляет 42.5 кДа.

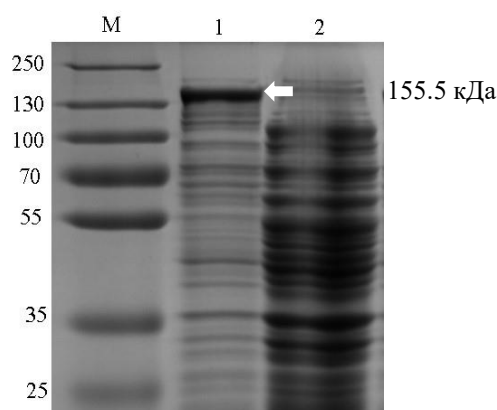


Рис.19. Электрофореграмма в 10% ПААГ клеточной культуры после и до индукции синтеза белка DpbCasX_MBP, где

M- маркер PageRuler; 1- pET CasX_MBP после; 2- pET CasX_MBP - до;

Из рисунка 19 видно значительное увеличение количества белка в пробах, отобранных после индукции. Что позволяет судить о том, что белка в клетках синтезировано достаточно для выделения.

Осадок клеток с синтезированным белком DpbCasX_MBP был лизирован в буфере А, соницирован, откручен на центрифуге при 16000g и профильтрован. Полученный лизат подвергали двум стадиям хроматографической очистки для получения чистого препарата белка, пригодного для проведения биохимических реакций.

Хроматографическая очистка белка включала в себя последовательно металл-аффинную и гель-фильтрационную хроматографии.

Первый этап очистки рекомбинантного белка DpbCasX_MBP был проведен на колонке Ni HisTrap HP. DpbCasX_MBP за счет наличия на С-конце последовательности из шести гистидиновых аминокислотных остатков может связываться с наполнителем данной колонки. Адсорбция гистидин-меченного белка основана на обратимом взаимодействии аминокислотных остатков гистидина с никелем, удерживаемом агарозной смолой. Связавшийся белок можно элюировать более сильным хелатирующим агентом (в данном случае, имидазолом), который будет конкурировать с Ni за связывание с His-tag меткой. Хроматографию проводили на установке АКТАPurifier, элюцию белка с колонки детектировали по наличию поглощения при длине волны 280 нм.

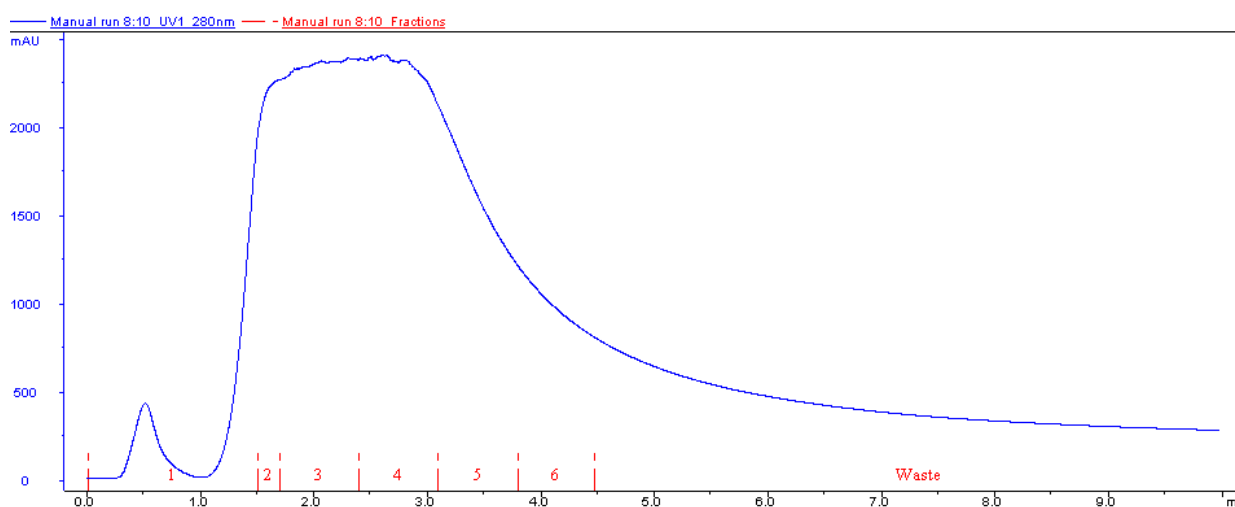


Рис.20. Профиль элюции белковых фракций при проведении аффинной хроматографии на Ni HisTrap HP колонке. По оси ординат – поглощение при длине волны 280нм, mAU (милли-единицы абсорбции), по оси абсцисс – объем элюции, мл

На рисунке 20 представлены результаты первой стадии хроматографической очистки на колонке Ni HisTrap HP. По протяженному пику поглощения при длине волны 280 нм можно судить о выходе большого количества рекомбинантного белка. Фракции 3 и 4 были отобраны для следующего этапа очистки белка на колонке Superdex 200 10/30 GL.

На колонке Superdex 200 10/30 GL проводили гель-фильтрационную хроматографию. Этот тип хроматографии позволяет разделять вещества по размеру их молекул. Неподвижная фаза содержит поры, в которых задерживаются при прохождении мелкие молекулы, а крупные выходят первыми. Анализ результатов гель-фильтрации проводили по калибровочной хроматограмме с известными молекулярными массами белков, проходящих через колонку на определенный миллилитр буфера (Рис.21). Зная молекулярную массу исследуемого белка, можно оценить, на каком объеме элюции он будет выходить с колонки. Рекомбинантный белок DpbCasX_MBP 155.5 кДа – приблизительно на 12 миллилитре.

Column Superdex 200 10/300 GL

- Sample:
1. Thyroglobulin (M, 669 000) 5 mg/ml
 2. Ferritin (M, 440 000) 0.4 mg/ml
 3. BSA (M, 67 000) 8 mg/ml
 4. β -lactoglobulin (M, 35 000) 2.5 mg/ml
 5. Ribonuclease A (M, 13 700) 5 mg/ml
 6. Cytochrome C (M, 13 600) 1.5 mg/ml
 7. Aprotinin (M, 6 512) 2 mg/ml
 8. Vitamin B12 (M, 1 355) 0.1 mg/ml

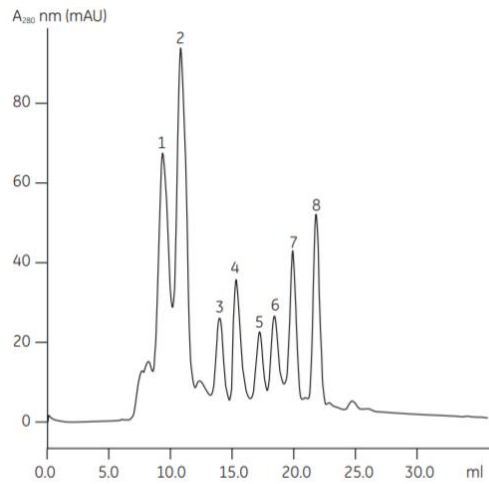


Рис.21. Типичная хроматограмма из функционального теста Superdex 200 10/300 GL

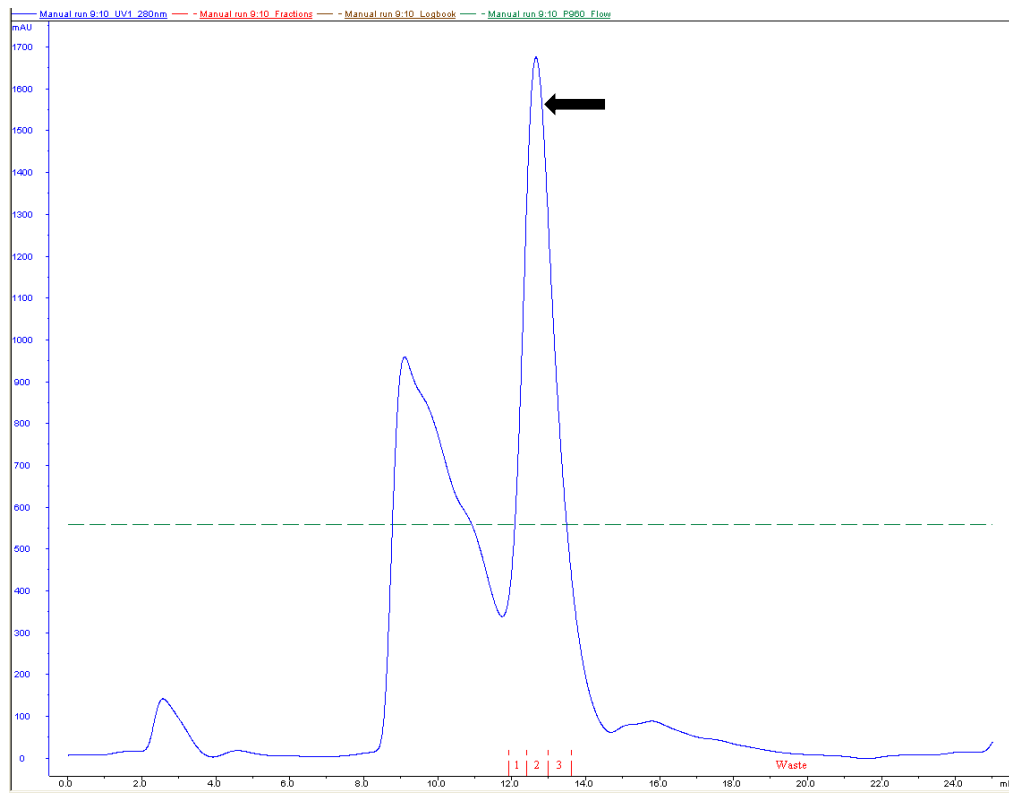


Рис.22. Профиль элюции белковых фракций при проведении гель-фильтрации на колонке Superdex 200 10/30 GL. По оси ординат – поглощение при длине волны 280нм, mAU (милли-единицы абсорбции), по оси абсцисс – объем элюции, мл

На хроматограмме на рисунке 22 стрелкой указан пик, содержащий фракцию с мономерным целевым белком DpbCasX_MBP. Фракция 2 была отобрана для дальнейшей работы. Для электрофоретического анализа в денатурирующих условиях в 10% ПААГ была взята аликвота 12 мкл, а остальное содержимое фракции 2 было сконцентрировано через фильтры типа Центрикон с номинальной отсекаемой молекулярной массой 30 кДа и разделено на аликвоты по 10 мкл для дальнейшего использования. Пробы замораживали и хранили при -80°C .

Подтвердить наличие в выделенных фракциях исследуемого белка и оценить степень его гомогенности можно при помощи электрофоретического анализа в 10% ПААГ. Из рисунка 23 видно, что в полученных фракциях действительно содержится белок DpbCasX_MBP молекулярной массой 155.5 кДа, и практически не имеется примесей.

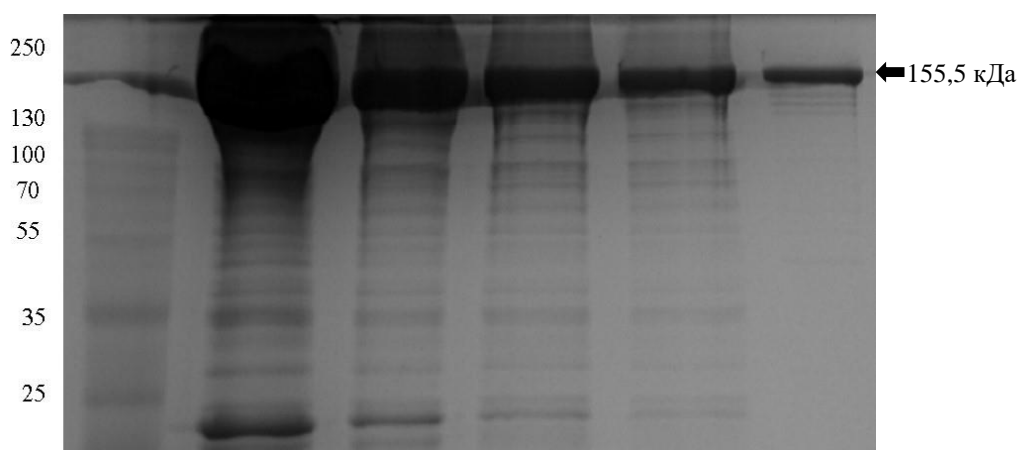


Рис.23. Электрофореграмма белковых фракций в 10% ПААГ после двух этапов хроматографической очистки, где

М-маркер PageRuler;

1,2,3,4-фракции 3,4,5,6 после Ni HisTrap колонки;

5- фракция 2 после колонки Superdex200.

Для того чтобы рассчитать количество белка требующееся для реакции *in vitro* необходимо знать его концентрацию. Измерение концентрации было проведено при помощи электрофореза в 10% ПААГ по калибровочной кривой BSA с известным содержанием белка (Рис.24). Раствор BSA брали в следующих количествах: 0.5, 1, 2, 3 и 4 мкг, белок DpbCasX_MBP объемом 0.5, 1, 2 и 3 мкл.

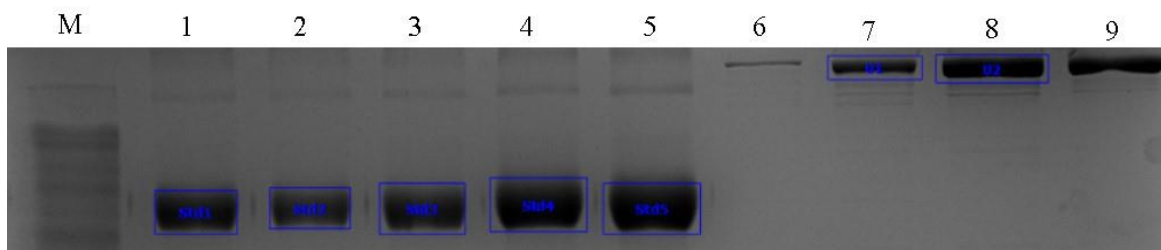


Рис.24. Электрофореграмма определения концентрации белка CasX MBP, где

М-маркер PageRuler, 1,2,3,4,5- раствор БСА с известным содержанием белка: 0.5, 1, 2, 3 и 4 мкг; 6,7,8,9- белок CasX MBP объемом объемом 0.5, 1, 2 и 3 мкл.

Для сравнения интенсивности «бэндов» калибровочной кривой с искомым белком была использована программа ImageLab.

Концентрация выделенного белка равна 0,75 мг/мл.

Таким образом, был получен мономер рекомбинантного белка DpbCasX_MBP – первый компонент для проведения реакции *in vitro*. После этого была искусственно синтезирована молекула гибридной РНК, последовательность которой была позаимствована статьи [Liu и др., 2019], посвященной CasX и вышедшей во время работы над проектом.

3.6 Проведение эндонуклеазной реакции *in vitro*

Нуклеазную активность белка можно оценить в реакции *in vitro*. По наличию или отсутствию расщепления соответствующей ДНК-мишени можно судить об активности исследуемого белка.

Нами были подготовлены все необходимые компоненты системы: рекомбинантный белок DpbCasX_MBP, направляющая sgРНК и ДНК-мишень, включающая сайт узнавания и известный РАМ. Расщепление белком целевой последовательности происходит через несколько нуклеотидов после РАМ. Если комплекс белок-РНК активен, то после внесения эндонуклеазой разрыва образуются два фрагмента разной длины, детектируемые методом гель-электрофореза (Рис.25).

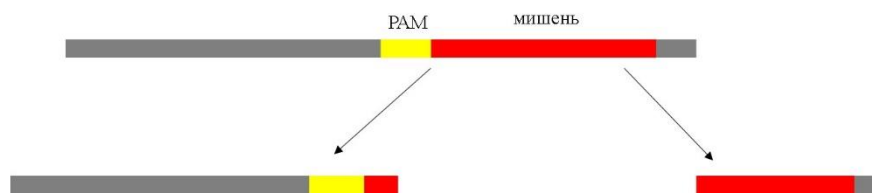


Рис.25. Схема расщепления ДНК-мишени

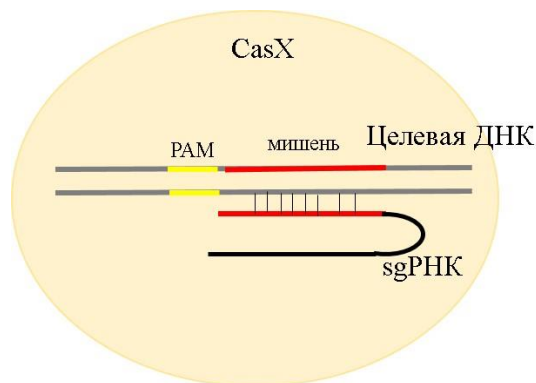


Рис.26. Схема работы эффекторного комплекса *in vitro*.

Совместная инкубация полученных компонентов позволит понять, обладает ли DpbCasX_МРВ нуклеазной активностью. При первом проведении реакции *in vitro* были взяты разные соотношения белок:РНК (1:1, 1:5, 1:20) и разные концентрации белка (100nM и 400nM).

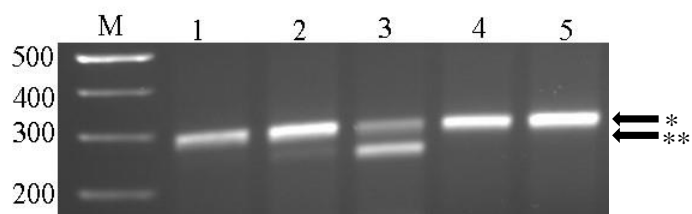


Рис.27. Электрофореграмма продуктов реакции *in vitro* расщепления ДНК-мишени белком DpbCasX в 1,5% агарозном геле, где

М-маркер 100bp DNA Ladder;

1-CasX 100nM:sgРНК 120nM, 1:1; 2- CasX 100nM:sgРНК 2mkM, 1:20;

3- CasX 400nM:sgРНК 2mkM, 1:5;

4- контроль, 100nM CasX без добавления sgРНК

5- контроль, 400nM CasX без добавления sgРНК

*- исходная ДНК; **- после расщепления комплексом.

Рисунок 27 показывает, что белок обладает нуклеазной активностью. Наиболее эффективно CasX расщепляет ДНК-мишень в соотношении белок:РНК 1:5 в концентрации белка 400nM.

После того, как было установлено оптимальное соотношение компонентов для протекания реакции, было решено определить температуру, при которой будет наиболее эффективна активность нуклеазы CasX. Выбранный диапазон температур: от 20°C до 50°C с интервалом в 5°C.

В этом эксперименте было взято установленное в предыдущем опыте соотношение компонентов реакции. В контрольную пробу не добавляется sgРНК.

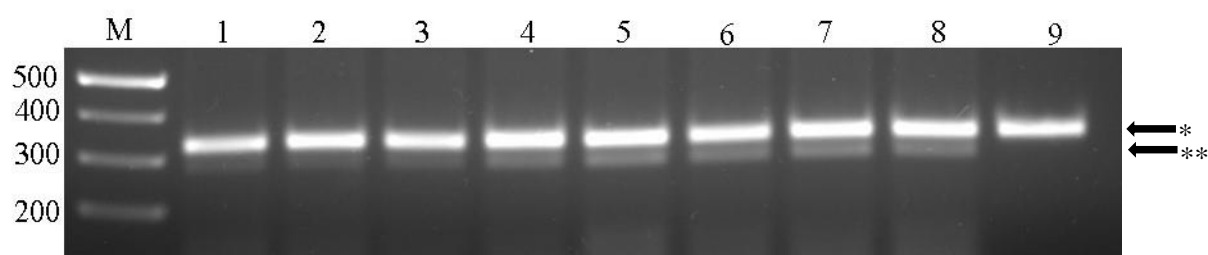


Рис. 28. Электрофореграмма продуктов *in vitro* реакции в 1,5% агарозном геле в диапазоне от 20°C до 50°C с интервалом в 5°C, где

М-100bp DNA Ladder; 1-20°C; 2-25°C; 3-30°C; 4-35°C; 5-37°C; 6-40°C; 7-45°C; 8-50°C; 9- контроль

*- исходная ДНК; **- после расщепления комплексом.

Из рисунка 28 можно судить о том, что эндонуклеаза CasX расщепляет ДНК-мишень при любой из выбранных нами температур, но эффективнее всего – в диапазоне от 35 до 50°C.

4. Обсуждение

Характеристика нуклеазной активности Cas белка важна для использования его в качестве инструмента генного редактирования в эукариотических клетках. Поэтому основным методом исследования CRISPR-Cas систем является проведение эндонуклеазной реакции *in vitro* [Anders, Jinek, 2014; Karvelis, Gasiunas, Siksny, 2013; Liu и др., 2015]. Для этого нужно собрать систему из следующих компонентов: эндонуклеаза Cas, направляющая молекула РНК (crРНК и tracrРНК или химерная молекула sgРНК) и ДНК-мишень. Инкубация подготовленных компонентов и последующая визуализация методом геле-электрофореза в 1,5% агарозном геле позволит оценить наличие нуклеазной активности у исследуемого белка. Анализ работы рибонуклеопротеинового комплекса в условиях *in vitro* необходим для того, чтобы определить минимальный состав эффектора, установить оптимальное соотношение РНК:белок, подобрать температуру для максимально эффективного расщепления ДНК-мишени комплексом.

Активность системы CRISPR-CasX до недавнего времени была показана только *in vivo* в бактериальных клетках [Burstein и др., 2017]. Целью данной работы было показать нуклеазную активность белка DpbCasX в условиях *in vitro*. Создание активной системы *in vitro* позволило бы определить минимальный набор компонентов необходимый для специфического расщепления ДНК-мишени.

При выполнении данной работы мы столкнулись с рядом трудностей.

Так как белки обладают высокой лабильностью, то их выделение требует подбора оптимальных условий (рН, концентрация солей, температура и пр.). Продукция рекомбинатного белка в клетках-продуцентах может вести к образованию крупных агрегатов внутри клетки, неправильному сворачиванию и образованию белком телец включения. Такой агрегированный белок уже не будет принимать нативную конформацию и окажется непригодным для выделения и очистки методом хроматографии. Выяснилось, что во время экспрессии системы рЕТ21a CasX в клетках-продуцентах весь рекомбинантный белок, синтезируемый клетками *E. coli* штамма Rosetta (DE3), выпадал в осадок. Изменение рН и ионной силы лизисного буфера, а также добавление глицерола или тритона X100 не увеличивало растворимость белка.

Из литературы известно [Kapust, Waugh, 1999], что для увеличения растворимости рекомбинантного белка используют конструкции с мальтоза-связывающим белком (МВР), которые являются популярными и универсальными для повышения растворимости

белков. Наиболее вероятной причиной увеличения растворимости при использовании этого подхода считается то, что белок MBP быстро и правильно укладывается сразу же после освобождения с рибосом, увеличивая шансы целевого белка не выпасть в осадок.

Большинство N- и C-концов эффекторных Cas-нуклеаз доступны для добавления к ним различных аминокислотных последовательностей, что позволяет создавать таким образом белки слияния с ненарушенной нуклеазной активностью. В связи с отсутствием на момент начала нашей работы данных о кристаллической структуре DpbCasX было решено добавить ген, кодирующий белок MBP, к последовательности DpbCasX, чтобы белок MBP находился на N-конце белка DpbCasX, и проверить растворимость химерного белка в этом случае. Действительно, добавление MBP к белку DpbCasX помогло преодолеть проблему его нерастворимости.

В связи с огромной практической значимостью CRISPR-Cas систем изучение новых Cas эффекторов является быстро развивающейся и высококонкурентной областью. В процессе нашей работы в феврале 2019 года была выпущена статья группы Дженнифер Дудны, посвященная кристаллографической структуре CasX [Liu и др., 2019], где уже была показана активность белка *in vitro*. Эта работа опередила наши исследования, но в то же время дала новые полезные сведения для продолжения наших экспериментов. Из данных, представленных в статье, мы позаимствовали составы буферов для выделения, которые помогли увеличить выход белка DpbCasX_MBП, а также последовательность гибридной sgРНК (изначально, планировалось использовать двухкомпонентную систему tracr- и cr-РНК).

В ходе реакций *in vitro* нам удалось воспроизвести результаты группы Дженнифер Дудны и показать, что DpbCasX обладает нуклеазной активностью. Наличие MBP-последовательности на N-конце не мешает работе эндонуклеазы, но, возможно, если отщепить ферментами последовательность MBP, то эффективность DpbCasX повысится.

Определение температурного диапазона, в котором эндонуклеаза DpbCasX активна, является важным пунктом его характеристики для дальнейшего использования в биотехнологии. Применяемая в геномной инженерии эндонуклеаза SpCas9 является ферментом широкого спектра действия, так как может расщеплять ДНК в температурном диапазоне от 20 до 70°C [Mougiakos и др., 2017]. Белок Cas12a, наиболее схожий с CasX, более чувствителен к температуре при исследованиях по редактированию генома [Malzahn и др., 2019]. Проведенный нами эксперимент демонстрирует, что DpbCasX может быть эффективен при различных температурах. Поскольку нуклеаза DpbCasX_MBП

лучше расщепляет ДНК-мишень при более высоких температурах, то предпочтительнее его использовать в клетках животных, а не растений. Кроме того, небольшой размер белка позволит использовать для его доставки вирусные векторы. Таким образом, белок DpbCasX может быть перспективным объектом изучения в качестве потенциального инструмента генного редактирования.

Выводы

1. Создан вектор на основе плазмиды pET21a для экспрессии гена CasX_MBP в клетках *E. coli* штамма Rosetta (D3).
2. Получен рекомбинантный белок DpbCasX_MBP.
3. Анализ нуклеазной активности белка DpbCasX_MBP в условиях *in vitro* показал, что при инкубации с направляющей РНК эндонуклеаза расщепляет ДНК-мишень, содержащую РАМ.
4. DpbCasX_MBP эффективно расщепляет ДНК-мишень в диапазоне температур от 35 до 50°C.

Благодарности

Хочу выразить благодарность моему научному руководителю Цветковой Елене Викторовне за помощь в оформлении работы, ценные указания и замечания. Огромную благодарность приношу научному консультанту Федоровой Яне Витальевне за помощь в ведении исследования и освоении методик, за советы и замечания в ходе выполнения и написания данной работы. Благодарю Селькову Полину Анатольевну за методические рекомендации во время работы в лаборатории, Побегалова Георгия Евгеньевича и Арсениева Анатолия Николаевича за помощь при проведении хроматографии.

Благодарю весь коллектив НИК «Нанобио» во главе с Ходорковским Михаилом Алексеевичем за оказанную мне помощь и предоставление материальной базы и оборудования для проведения этого исследования.

Список литературы

1. Anders C., Jinek M. In vitro Enzymology of Cas9 *Carolin* // 2014. С. 1–20.
2. Bobay L.M., Touchon M., Rocha E.P.C. Manipulating or Superseding Host Recombination Functions: A Dilemma That Shapes Phage Evolvability // *PLoS Genet.* 2013. Т. 9. № 9.
3. Brouns S.J.J. и др. Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes // 1996. Т. 272. № June. С. 1641–1643.
4. Burstein D. и др. New CRISPR–Cas systems from uncultivated microbes *microbes* // *Nat. Publ. Gr.* 2017. Т. 542. № 7640. С. 237–241.
5. Cong L. и др. Multiplex Genome Engineering Using CRISPR/Cas Systems // *Science* (80-.). 2013.
6. Davies C. и др. Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity // *Nat. Struct. Mol. Biol.* 2014. Т. 21. № 6. С. 528–534.
7. Deltcheva E. и др. CRISPR RNA maturation by trans-encoded small RNA and host factor // 2011. Т. 471. № 7340. С. 602–607.
8. Díez-villaseñor C. и др. CRISPR-spacer integration reporter plasmids reveal distinct genuine acquisition specificities among CRISPR-Cas I-E variants of *Escherichia coli* // 2013. Т. 10. № 5. С. 792–802.
9. Doudna J.A., Charpentier and E. The new frontier of genome engineering with CRISPR-Cas9 // 2014. Т. 346. № 6213.
10. Dy R.L. и др. A widespread bacteriophage abortive infection system functions through a Type IV toxin-antitoxin mechanism // *Nucleic Acids Res.* 2014. Т. 42. № 7. С. 4590–4605.
11. Fonfara I. и др. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA // *Nature.* 2016. Т. 532. № 7600. С. 517–521.
12. Gasiunas G. и др. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria // *Proc. Natl. Acad. Sci.* 2012. Т. 109. № 39. С. E2579–E2586.
13. Guha T.K., Wai A., Hausner G. Programmable Genome Editing Tools and their Regulation for Efficient Genome Engineering // *Comput. Struct. Biotechnol. J.* 2017. Т. 15. С. 146–160.
14. Hatoum-aslan A. и др. Co-transcriptional DNA and RNA Cleavage during Co-

- transcriptional DNA and RNA Cleavage during Type III CRISPR-Cas Immunity // *Cell*. 2015. T. 161. № 5. С. 1164–1174.
15. James K. Nuñez, Amy S.Y. Lee A.E. and J.A.D. Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity // *Nature*. 2015. T. 519. № Xxxxx. С. 193–198.
16. Jinek M. и др. A Programmable Dual-RNA – Guided // 2012. T. 337. № August. С. 816–822.
17. Jinek M. и др. RNA-programmed genome editing in human cells // *Elife*. 2013. T. 2. С. 1–9.
18. Jinek M. и др. Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation // *Science* (80-.). 2014. T. 343. № 6176. С. 1–28.
19. Kapust R.B., Waugh D.S. Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused // 1999. С. 1668–1674.
20. Karvelis T., Gasiunas G., Siksnys V. Programmable DNA cleavage in vitro by Cas9 // *Biochem. Soc. Trans.* 2013. T. 41. № 6. С. 1401–1406.
21. Kunin V., Sorek R., Hugenholtz P. Evolutionary conservation of sequence and secondary structures in CRISPR repeats // *Genome Biol.* 2007. T. 8. № 4.
22. Lämml U.K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4 // *Nature*. 1970. T. 227. № 5259. С. 680–685.
23. Levy A. и др. CRISPR adaptation biases explain preference for acquisition of foreign DNA // 2015. T. 520. № 7548. С. 505–510.
24. Liu J.-J. и др. CasX enzymes comprise a distinct family of RNA-guided genome editors // *Nature*. 2019.
25. Liu Y. и др. In Vitro CRISPR / Cas9 System for Efficient Targeted DNA Editing // *MBio*. 2015. T. 6. № Q. С. Q.
26. Makarova K.S. и др. An updated evolutionary classification of CRISPR–Cas systems // 2015. T. 13. № 11. С. 722–736.
27. Malzahn A.A. и др. Application of CRISPR-Cas12a temperature sensitivity for improved genome editing in rice, maize, and Arabidopsis // *BMC Biol.* 2019. T. 17. № 1. С. 1–14.
28. Mougiakos I. и др. Characterizing a thermostable Cas9 for bacterial genome editing and

- silencing // *Nat. Commun.* 2017. Т. 8. № 1.
29. Nishimasu H. и др. Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA // 2016. Т. 118. № 24. С. 6072–6078.
30. Ofir G. и др. DISARM is a widespread bacterial defence system with broad anti-phage activities // *Nat. Microbiol.* 2018. Т. 3. № 1. С. 90–98.
31. Özcan A. и др. Type IV CRISPR RNA processing and effector complex formation in *Aromatoleum aromaticum* // *Nat. Microbiol.* 2019. Т. 4. № 1. С. 89–96.
32. Ryan N. Jackson B.W. A conserved structural chassis for mounting versatile CRISPR RNA-guided immune responses // 2015. Т. 40. № 4. С. 1291–1296.
33. Saprunauskas R. и др. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli* // *Nucleic Acids Res.* 2011. Т. 39. № 21. С. 9275–9282.
34. Shmakov S. и др. Discovery and Functional Characterization of Article Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems // *Mol. Cell.* 2015. С. 1–13.
35. Silas S. и др. Direct CRISPR spacer acquisition from RNA by a natural reverse transcriptase-Cas1 fusion protein // *Science* (80-.). 2016. Т. 351. № 6276.
36. Silas S. и др. On the Origin of Reverse Transcriptase-Using CRISPR-Cas Systems and Their Hyperdiverse, Enigmatic Spacer Repertoires // *MBio.* 2017. Т. 8. № 4.
37. Takeuchi N. и др. Nature and intensity of selection pressure on crispr-associated genes // *J. Bacteriol.* 2012. Т. 194. № 5. С. 1216–1225.
38. Tang L. и др. CRISPR/Cas9-mediated gene editing in human zygotes using Cas9 protein // *Mol. Genet. Genomics.* 2017. Т. 292. № 3. С. 525–533.
39. Wang J. и др. Structural and Mechanistic Basis of PAM-Dependent Spacer Acquisition in CRISPR-Cas Systems // *Cell.* 2015. Т. 163. № 4. С. 840–853.
40. Wei Y. и др. Sequences spanning the leader-repeat junction mediate CRISPR adaptation to phage in *Streptococcus thermophilus* // *Nucleic Acids Res.* 2015. Т. 43. № 3. С. 1749–1758.
41. Wei Y., Terns R.M., Terns M.P. Cas9 function and host genome sampling in type II-A CRISPR–cas adaptation // *Genes Dev.* 2015. Т. 29. № 4. С. 356–361.

42. Westra E.R. и др. CRISPR Immunity Relies on the Consecutive Binding and Degradation of Negatively Supercoiled Invader DNA by Cascade and Cas3 // *Mol. Cell*. 2012. Т. 46. № 5. С. 595–605.
43. Westra E.R., Buckling A., Fineran P.C. CRISPR-Cas systems: Beyond adaptive immunity // *Nat. Rev. Microbiol.* 2014. Т. 12. № 5. С. 317–326.
44. Xue W. и др. CRISPR-mediated direct mutation of cancer genes in the mouse liver *Wen* // 2015. Т. 344. № 6188. С. 1173–1178.
45. Yamano T. и др. Crystal structure of Cpf1 in complex with guide RNA and target DNA // *Acta Crystallogr. Sect. D Struct. Biol.* 2016. Т. 165. № 4. С. 949–962.
46. Yao R. и др. CRISPR-Cas9/Cas12a biotechnology and application in bacteria // *Synth. Syst. Biotechnol.* 2018. Т. 3. № 3. С. 135–149.
47. Yosef I., Goren M.G., Qimron U. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli* // *Nucleic Acids Res.* 2012. Т. 40. № 12. С. 5569–5576.
48. Zetsche B. и др. Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System // *Cell*. 2015. Т. 163. № 3. С. 759–771.
49. Zhou Y. и др. High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells // *Nature*. 2014. Т. 509. № 7501. С. 487–491.