Автономная некоммерческая образовательная организация высшего образования
«Научно-технологический университет «Сириус»

На правах рукописи



Бровин Андрей Николаевич

Оптимизация белкового транс-сплайсинга с помощью интеинов для применения в терапии наследственных заболеваний

1.5.3. Молекулярная биология

Диссертация

на соискание ученой степени

кандидата биологических наук

Научный руководитель:

Кандидат биологических наук

Карабельский Александр Владимирович

Федеральная территория «Сириус»

2026

ОГЛАВЛЕНИЕ

ОГЛАВЛЕНИЕ 2
ВВЕДЕНИЕ 5
ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ 13
1.1. Интеины и их роль в современной биотехнологии 13
1.2. Общие характеристики интеинов для белкового сплайсинга 16
1.2.1. Определение интеинов и их роль в живых системах 16
1.2.2. Механизм белкового сплайсинга в общем виде 18
1.3. Структурные особенности интеинов на уровне аминокислотной последовательности 21
1.3.1. Структурные особенности интеинов группы DnaE 25
1.3.2. Структурные особенности интеинов группы DnaE на уровне 3D-модели 27
1.3.3. Сравнение кинетических характеристик интеинов группы DnaE 29
1.3.4. Влияние положения и состава аминокислот экстеинов при сплайсинге с участием интеинов группы DnaE 31
1.3.5. Кросс-реактивные свойства интеинов группы DnaE 34
1.4. Примеры применения белкового сплайсинга в современной биотехнологии 36
1.4.1. Биосенсоры и методы мечения белков на основе интеинов 36
1.4.2. Биотехнологические приложения интеинов для получения рекомбинантных белков 38
1.4.3. Применение интеинов в генной терапии 41
1.4.3.1. Интеины для заместительной генной терапии 41
1.4.3.2. Интеины для редактирующей генной терапии 45
1.5. Разработка препаратов генной терапии на основе двухвекторных систем 49
1.5.1. Обзор актуальных мишеней, используемых для терапии с помощью AAV 49
1.5.2. Офтальмологические препараты как модель для терапии наследственных заболеваний 50
1.5.4. Эпидемиология болезни Штаргардта 53
1.5.5. Современное состояние терапии синдрома Штаргардта 55
1.6. Основные выводы по итогам первой главы 57
ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ 58
2.1. Генная инженерия для получения генетических конструкций 58
2.1.1. Аминокислотные последовательности интеинов: поиск в базах данных, выравнивание и анализ 58

2.1.2. Алгоритм выбора сайтов разделения целевой молекулы для интеинов группы DnaE....58
2.1.3. Получение генетических конструкций, кодирующих интеины....59
2.2. Компьютерное моделирование для визуализации молекул....61
2.3. Клеточные тесты для оптимизации условий трансфекции и валидации данных компьютерного моделирования и анализа....62
2.3.1. Поддержание клеточных линий....62
2.3.2. Трансфекция и трансдукция адгезионных клеток....63
2.3.3. Динамический анализ с помощью Incucyte....63
2.3.4. Цитометрический анализ клеточных популяций....64
2.3.5. Вестерн-блоттинг....64
2.3.6. Получение AAV в клетках HEK293....65
2.3.7. Получение липидных наночастиц....68
2.4. Работа с лабораторными животными....69
2.4.1. Содержание лабораторных животных....69
2.4.2.Проведение процедуры интравитреальной инъекции....70
2.4.3. Анализ тканей гистологическими методами....70
ГЛАВА 3. РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ....72
3.1. Дизайн экспериментальной части работы....72
3.2. Выбор интеинов и анализ аминокислотной последовательности экстеина....73
3.2.1. Выбор интеинов для экспериментальной работы....73
3.2.2. Анализ аминокислотной последовательности GFP для разделения по сайтам сплайсинга для интеинов группы DnaE....75
3.4.3. Моделирование получаемых предсплайсинговых молекул....76
с помощью AlphaFold 3....76
3.4.4. Функциональная проверка сайтов для разделения GFP....78
с помощью интеина Ssp DnaE....78
3.5. Оптимизация выхода продукта реакции сплайсинга GFP....81
3.6. Оптимизация сплайсинга с помощью интеинов....84
и их кросс-реактивного взаимодействия....84
3.7. Применение оптимальной интеиновой пары для раздельной доставки GFP с помощью AAV в клетки разного происхождения....88
3.7.1. Вестерн-блот-анализ клеток после трансфекции....88
3.7.2. Трансдукция клеточных линий с помощью AAV5 GFP Npu DnaE....91
3.8. Получение и тестирование препарата для терапии синдрома Штаргардта....98
3.8.1. Дизайн генетической конструкции с ABCA4....98

3.8.2. Вестерн-блот-анализ ABCA4 после трансфекции клеток HEK293 101
3.8.3. Тестирование тропизма AAV к тканям глаза in vivo.....103
3.8.4. Тестирование платформ доставки in vivo.....105
3.8.5. Анализ экспрессии ABCA4 in vivo.....107
ЗАКЛЮЧЕНИЕ.....110
СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ 113
СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ 114
СПИСОК ЛИТЕРАТУРЫ.....115
ПРИЛОЖЕНИЕ А 121
Сопроводительная информация для получения генетических конструкций.....121
Таблица А.1 — Праймеры для синтеза нуклеотидной последовательности интеинов ..121

ВВЕДЕНИЕ

Актуальность темы исследования

Наследственные или орфанные заболевания являются самыми сложными для терапии и диагностики в современной медицине. Ключевая проблема лечения орфанных заболеваний может быть продемонстрирована на контрасте высокого разнообразия мутаций при крайне низкой частоте встречаемости каждой из них. Так, например, самые распространённые генетические синдромы имеют частоту встречаемости 1:5000 в то время, как онкологические заболевания диагностируют 100 раз чаще. Сопоставимые цифры в диагностике (в год проводится 400 тысяч консультаций по орфанным заболеваниям и 700 тысяч по онкологическим) можно обнаружить по причине все еще не выявленных взаимосвязей генотип-фенотип в большинстве случаев. Перечисленные данные демонстрируют несовершенство природных механизмов репарации и репликации, которые приводят к возникновению *de novo* мутаций, являющихся причиной двух наиболее серьезных и сложных причин смертности и снижения качества жизни человека в настоящее время.

Второй ключевой проблемой лечения наследственных заболеваний является необходимость применения крайне высокотехнологичных и дорогостоящих методов, которые не позволяют применять их массово. К основным подходам терапии относятся различные манипуляции с генетическим материалом человека, предполагающие компенсацию функции мутантной гена с помощью доставки нормальной версии гена или редактирование мутантного участка гена. Каждый из подходов имеет ряд ограничений, связанных с биоэтическими или технологическими ограничениями.

В настоящее время терапия наследственных заболеваний стала возможна благодаря развитию вирусных векторных платформ. Для доставки генетических конструкций используют векторы на основе лентивирусов или аденоассоциированных вирусов (AAV) и др., а в последние несколько лет все чаще используют системы на основе липидных наночастиц. Безопасность и эффективность вирусных векторов на основе AAV была подтверждена в многочисленных клинических исследованиях, одобрены 8 препаратов для генозаместительной терапии.

Существующее ограничение емкости генома AAV (~ 4,5 тысячи пар нуклеотидов) инициировало многочисленные поисковые исследования, направленные на решение проблемы генотерапии с использованием крупных генов. Наиболее универсальным способом на сегодняшний день является разделение терапевтической конструкции на

составные части для упаковки в отдельные вирусные векторы. После трансдукции восстановление кодирующей последовательности или пептида происходит на уровне ДНК, РНК или пост-трансляционно на уровне белка.

Одной из наиболее эффективных методик восстановления функциональной молекулы белка из частей является технология белкового транс-сплайсинга, которая реализуется за счет особого класса белков, называемых интеинами. Высокая скорость и аффинность взаимодействия интеинов, автокаталитическая активность и способность ковалентно соединять белковые последовательности сделали данную технологию популярной в различных областях биотехнологии.

С момента открытия в 1987 году область приложения интеинового транс-сплайсинга постоянно расширялась, а за последние пять лет появилось множество примеров ее использования в генной терапии. Эта технология имеет несколько особенностей, связанных с механизмом взаимодействия как интеинов, так и экстеинов между собой, что обеспечивает и потенциальную возможность управлять эффективностью реакции сплайсинга. Основной задачей данной работы было детальное изучение механизмов белкового транс-сплайсинга с целью создания в дальнейшем эффективных препаратов, доставляющих протяженные кодирующие последовательности генов в соматические клетки.

В качестве основного инструмента для исследований использовали молекулу зеленого флуоресцентного белка (GFP), которую необходимо было разделить на две части и далее восстановить *in vitro* в функциональный белок с помощью интеинов из ферментативной группы эпсилон субъединицы третьей полимеразы (DnaE). Были проведены работы по изучению и оптимизации механизма белкового сплайсинга, подбору условий для увеличения выхода реакции при трансфекции и трансдукции на клеточных моделях, а также проведена первичная апробация технологии *in vivo* в рамках разработки препарата для заместительной генной терапии одной из форм наследственной ретинопатии.

Степень разработанности темы исследования

В доступных базах данных не обнаружено публикаций на русском языке, полностью соответствующих тематике настоящего исследования. Три работы под авторством П.Л. Старокадомского носят описательный характер основных механизмов белкового сплайсинга и одного примера прикладного применения интеинов для очистки гормона роста человека. В иностранных базах данных тематика представлена в количестве более

1500 работ по различным вопросам, включающих фундаментальные и прикладные исследования.

Несмотря на большое количество работ в теме белкового сплайсинга, остается много неизвестных, поскольку сборка рекомбинантных белков требует индивидуального подхода к каждому отдельно взятому варианту интеина и экстеина. Доступной методики или алгоритма для разработки генно-терапевтических препаратов на основе интеинового транс-сплайсинга не существует. Для проверки работоспособности интеинов *in vitro* в большинстве статей используется культура клеток человека HEK293, что существенно ограничивает понимание эффективности механизма белкового сплайсинга.

Цель работы

Целью данной работы является изучение, применение и оптимизация механизма белкового транс-сплайсинга с помощью интеинов для создания новых эффективных препаратов генозаместительной терапии с использованием аденоассоциированных вирусных векторов.

Задачи работы

1. Выбор оптимальной интеиновой пары для эффективной сборки крупных белков с помощью применения технологии транс-сплайсинга.
2. Конструирование плазмид для визуализации и оценки эффективности белкового транс-сплайсинга в клетках человека по восстановлению флуоресценции репортерного белка.
3. Оптимизация условий трансфекции и трансдукции *in vitro* и *in vivo* для сборки целевой белковой молекулы в клетках человека.
4. Применение полученных результатов для создания прототипа препарата для генозаместительной терапии.
5. Проведение первичных экспериментов на мышах для проверки работоспособности конструкций.

Научная новизна

В диссертационной работе изложены основные экспериментальные результаты по изучению механизмов белкового транс-сплайсинга с помощью интеинов. Основные результаты касаются применения методов генной и белковой инженерии для получения генетических конструкций, позволяющих проверить ряд гипотез о работе системы белкового сплайсинга.

Основным результатом работы является разработка универсального алгоритма, позволяющего рассчитать теоретические сайты разделения любого экстеинового белка. В ходе экспериментальной валидации было продемонстрировано наличие новых сайтов для белка GFP и доказана работоспособность единственного раннее использованного сайта 69/70. Для белка ABCA4 было найдено множество сайтов разделения, однако экспериментальная проверка показала, что не все сайты одинаково эффективно реализуют реакцию сплайсинга. Таким образом, было показано, что наличие цистеина на C+1 положении является важным, но не ключевым фактором в восстановлении экстеиновой молекулы.

Дополнительно были изучены свойства кросс-реактивного взаимодействия между интеинами одной ферментативной группы DnaE. В ходе исследования были обнаружены ранее не описанные комбинации интеинов, влияющие на изменение эффективности транс-сплайсинга как в положительную (NpuN+SspC), так и в отрицательную (SspN+AvaC) сторону. Результат экспериментальных данных может далее быть использован для изучения структурных особенностей интеинов группы DnaE в совокупности с результатами выравнивания аминокислотных последовательностей и определения роли аминокислот в положениях C6, C9, C21 и C32 в области каталитических доменов.

Отдельно была продемонстрирована возможность применения раздельной интеиновой системы доставки *GFP* в AAV векторах пятого серотипа не только на двух модельных перевиваемых клеточных линиях HEK293 и ARPE19, но и на первичных культурах клеток FbP13, что подтверждает единство работы механизма белковый транс-сплайсинга с помощью интеинов в большинстве типов клеток человека.

В заключении работы был произведен трансфер технологии раздельной сборки и доставки ABCA4, в результате чего был получен прототип препарата для терапии синдрома Штаргардта и отработаны оптимальные условия для введения препарата в глаз мыши.

Теоретическая значимость диссертационного исследования

Значимой теоретической составляющей работы является подробное описание общего алгоритма применения интеинов для получения разделенных двухвекторных систем на основе аденоассоциированных вирусов. Каждый из описанных этапов имеет ряд теоретических и практических особенностей, однако наиболее важным результатом является разработка и применение алгоритма по поиску сайтов для разделения целевой молекулы на экстеины.

Практическая значимость диссертационного исследования

Главной целью исследовательской работы является разработка и применение технологии транс-сплайсинга для создания генотерапевтических препаратов нового поколения с наибольшим выходом по количеству собранного белка после раздельной доставки кодирующих участков гена. Основной результат заключается в получении эффективного прототипа препарата для терапии синдрома Штаргардта. Прототип был протестирован *in vitro и in vivo* и продемонстрировал свою эффективность.

Методология и методы исследования.

В рамках данной работы был использованы методы молекулярной и клеточной биологии, гистологии, генной и белковой инженерии, биотехнологии, данные были обработаны в соответствии с методами статистически.

Для получения генетических конструкций использовали методы химического синтеза для получения праймеров, полимеразной цепной реакции для сборки и мутагенеза генетических конструкций, рестриктазно-лигазные методы клонирования, секвенирование по Сэнгеру для подтверждения корректности получаемых последовательностей.

Для изучения свойств кросс-реактивности использовали транзиентную химическую трансфекцию эукариотических клеток и методы биоаналитики (проточная цитометрия, флуоресцентная микроскопия).

Для получения AAV использовали суспензионную клеточную культуру HEK293, трансфекцию и культивирование клеток в бессывороточной среде, осветление культуральной жидкости, концентрирование и хроматографическую очистку вирусных частиц с последующим анализом титра методом ПЦР. Для оценки сборки молекулы белка после транс-сплайсинга использовали вестерн-блоттинг в случае *in vitro* экспериментов на клеточных культурах и иммуногистохимическое окрашивание в экспериментах *in vivo.*

Основные положения, выносимые на защиту:

1. Разработанный программный код позволяет по аминокислотной последовательности анализировать белки для разделения на две части интеином DnaE. На основе работы данного алгоритма в работе впервые были обнаружены сайты разделения 47/48 для белка GFP и множество сайтов для белка ABCA4. В ходе экспериментальной проверки была доказана роль оптимального аминокислотного состава как на N-, так и на C-концевой части экстеина для прохождения реакции сплайсинга.

2. Работоспособность системы раздельной доставки GFP с помощью интеина Npu DnaE в AAV векторе пятого серотипа была продемонстрирована в клетках HEK293, ARPE19 и FbP13, что подтверждает единство работы механизма интеинового сплайсинга в клетках человека.

3. Свойство кросс-реактивного взаимодействия присутствует во всех сочетаниях между интеинами Ssp, Ava, Npu DnaE в случае использования общего сайта разделения GFP по сайту 69/70. Продемонстрировано, что пара SspN+AvaC полностью блокирует реакцию сплайсинга, а пара NpuN+SspC незначительно увеличивает выход реакции. Результат экспериментальных данных может далее быть использован для изучения структурных особенностей интеинов группы DnaE в совокупности с результатами выравнивания аминокислотных последовательностей и определения роли аминокислот в положениях C6, C9, C21 и C32 в области каталитических доменов.

4. Сравнение форматов доставки генетического материала и методов введения препаратов в ткани сетчатки глаза демонстрирует преимущество вирусных векторных систем по сравнению с доставкой конструкций при помощи липидных наночастиц. Липидные наночастицы позволяют обеспечить экспрессию GFP с мРНК через 24 часа после введения, однако они локализуются в клетках пигментного эпителия. Вирусные частицы AAV восьмого серотипа эффективно трансдуцируют как клетки пигментного эпителия, так и фоторецепторы.

Степень достоверности и апробация результатов

Достоверность полученных результатов была подтверждена методами математической статистики с учетом соблюдения условия трех биологических и трех технических повторов, достаточных для статистической проверки гипотез при анализе биологических данных.

Для апробации научных результатов были подготовлены и представлены устные сообщения на конференциях: VIII молодежная школа-конференция по молекулярной биологии и генетическим технологиям Института Цитологии РАН (11–14 октября 2022 г., Институт Цитологии РАН, Санкт-Петербург), Конференция "Биохимия человека 2024" (17–19 октября 2024 года, Кафедра биохимии им. Т.Т. Берёзова Лаборатория экспериментальной онкологии НИИ молекулярной и клеточной медицины Медицинского института Российского университета дружбы народов, Москва), Саммит разработчиков лекарственных препаратов «Сириус.Биотех 2026» (20–22 мая 2026 года, Научно-технологический университет «Сириус», ФТ Сириус), научная конференция в рамках XIII Российского Форума Биотехнологий OpenBio (21–25 сентября 2026 года, Наукоград Кольцово).

Личный вклад

Автором самостоятельно были получены все генетические конструкции для сборки GFP методом транс-сплайсинга, описываемые в работе. Дизайн, выполнение и анализ клеточных экспериментов были проведены автором самостоятельно. *In vivo* эксперименты по сравнению тропности AAV серотипов были проведены автором самостоятельно, начиная от внутриглазных инъекций до гистологических исследований и анализа изображений. Анализ, обработка данных, получение графиков и изображений проводились автором в программе GraphPad Prism 8.2.1 полностью самостоятельно. Тексты и черновики приведенных публикаций на русском языке был составлен автором лично, в переводе на английский язык принимали участие все соавторы работ.

Дизайн генетических конструкций для препарата AAV8-ABCA4-Npu DnaE был составлен автором работы, а плазмиды были получены совместно с основным соавтором публикаций Сабанцевым М.О. в рамках написания и выполнения магистерской ВКР. *In vivo* эксперименты по тестированию препарата AAV8-ABCA4-Npu DnaE были выполнены основным соавтором публикаций Сабанцевым М.О. в рамках написания и выполнения магистерской ВКР.

Публикации

Материалы диссертационной работы отражены в четырех научных экспериментальных публикациях, журналах, индексируемых в международных базах данных Web of Science/Scopus, из которых две публикаций Q1 и две в К1.

Объем и структура работы

Диссертационная работа состоит из введения, обзора литературы, разделов с описанием результатов работы, заключения, выводов, списка публикаций по теме диссертации, списка литературы (105 источников). Работа изложена на 123 страницах, содержит 33 рисунка и 8 таблиц.

Благодарности

Автор выражает глубокую признательность научному руководителю к.б.н. Карабельскому А.В., коллегам: к.б.н. Малоголовкину А.С., к.б.н. Егорову А.Д., соавторам: PhD Минской Е.С., к.х.н. Чувпило С.А., к.х.н. Гурееву М.А., Сабанцеву М.О. за совместную работу, поддержку и ценные комментарии, оказанные на всех этапах выполнения работы.

Также автор выражает благодарность сотрудникам ресурсных центров Генетической инженерии, Биотехнологических продуктов, Клеточных технологий, Доклинических исследований Университета «Сириус» за создание условий, необходимых для реализации научных проектов. Отдельно автор выражает благодарность заведующему лабораторией генно-клеточной терапии Института регенеративной медицины Московского государственного университета имени М.В. Ломоносова д.м.н. Макаревичу П.И. за помощь в проведении *in vivo* экспериментов и Потаповой А.М. за методическую помощь в проведении *in vivo* экспериментов. Также автор выражает благодарность заведующему лабораторией редактирования генома ФГБНУ «МГНЦ» к.б.н. Смирнихиной С.А. за помощь в организации и проведении in vitro экспериментов на первичных клетках человека.

ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ

1.1. Интеины и их роль в современной биотехнологии

Интеины служат универсальным инструментом молекулярной биологии и биотехнологии. В одной из недавних публикаций (Anastassov, 2024) они названы новым «швейцарским ножом» синтетической биологии [1]. Такое сравнение отражает широкий спектр их применения. Интеины используют для нехроматографической очистки белков, разработки биосенсоров и биологических логических элементов, сборки антител, получения белков, токсичных для клеток продуцентов, а также в заместительной и редактирующей генной терапии. Область их применения продолжает расширяться (Nanda, 2020) [2].

Для демонстрации растущего интереса к применению интеинов в биотехнологии и биологии был проведён обзорный анализ статей в базе данных PubMed на портале NCBI с интервалом в 10 лет за период с 2004 по 2024 год (База данных PubMed) [3]. Суммарное количество публикаций с ключевым словом «интеины» за указанный 20-летний период составляет 1417 работ, а среднее количество публикаций — 74 работы в год, что свидетельствует о востребованности и актуальности рассматриваемой тематики.

Не менее важным является распределение статей по тематическим направлениям, которые можно разделить на четыре основных категории исследований: изучение теоретических основ белкового сплайсинга и структурных особенностей интеинов (выделено синим), биотехнологические приложения интеинов для сборки белковых комплексов (выделено светло-коричневым), разработка биосенсоров для визуализации клеточных и белковых процессов (выделено светлосерым) и применение в области генной терапии наследственных заболеваний (выделено светло-синим). Визуализация данного распределения статей по тематике с шагом в 10 лет представлена на Рисунке 1, где явно прослеживается тенденция к переходу от теоретических исследований к практическим и наблюдается возрастающий интерес к применению технологии в генной терапии.

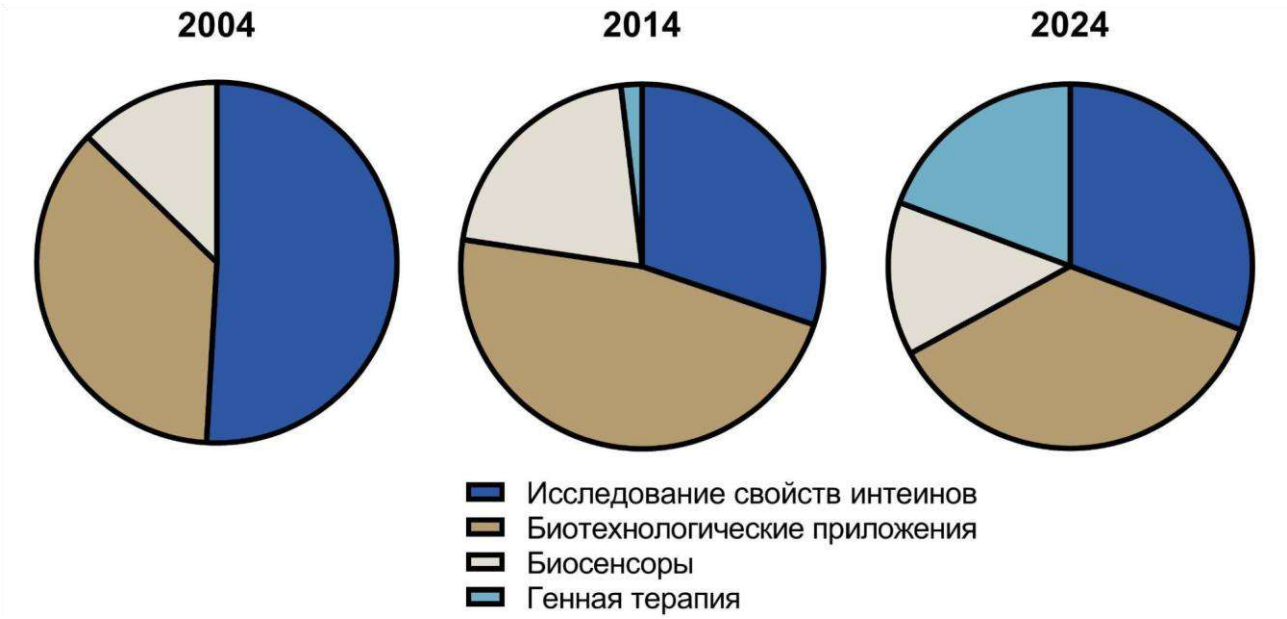


Рисунок 1 — Распределение статей по тематике исследований и годам публикации по данным PubMed NCBI. Иллюстрация была подготовлена автором работы с использованием GraphPad Prism 8.1 на основе данных PubMed NCBI [3]

Тематика данной работы связана с одной из ключевых проблем генной терапии — преодолением ограничения упаковочной ёмкости AAV. В Таблице 1 представлены характеристики часто используемых интеинов, применяемых в генной терапии для раздельной доставки и сборки крупных белков. В таблице обобщены результаты обзора наиболее актуальных статей за последние 5 лет. Температура в Таблице 1 соответствует условиям определения приведённых кинетических параметров и не должна рассматриваться как универсальный температурный оптимум. Значения скорости зависят от последовательностей экстеинов, места искусственного расщепления, состава буфера, концентрации фрагментов и конструкции белка. Кинетические данные приведены по результатам исследований [4], [5]; наличие экспериментальных структур проверено по базе RCSB PDB.

Для сравнительного анализа были отобраны природные и искусственно расщепленные интеины, наиболее широко используемые в исследованиях белкового транс-сплайсинга или применяемые в качестве эталонных систем. Интеины SspDnaE, NpuDnaE и AvaDnaE относятся к группе DnaE цианобактерий, SspDnaB и RmaDnaB локализованы в репликативных хеликазах DnaB, тогда как gp41-1 и NrdJ1 были обнаружены при анализе метагеномных последовательностей.

Таблица 1 — Кинетические и структурные характеристики часто используемых расщепленных интеинов

Интеин	Род и вид организма	Кинетика реакции	◦T, C	Экспериментальная 3D-структура
Rma DnaB	Rhodothermus marinus DSM 4252	$k = (6{,}9 \pm 2{,}2) \times 10^{-5}$ с$^{-1}$; $t_{1/2} \approx 167$ мин	25	Экспериментальная структура в PDB не обнаружена
Ssp DnaE	Synechocystis sp. PCC 6803	$k \approx 1{,}5 \times 10^{-4}$ с$^{-1}$; $t_{1/2} \approx 76$ мин	37	PDB 1ZD7, 1ZDE, 3NZM, 4GIG
Ssp DnaB	Synechocystis sp. PCC 6803	$k = 7{,}1 \times 10^{-4}$ с$^{-1}$; $t_{1/2} \approx 16{,}3$ мин	25	PDB 1MI8; варианты 6FRE, 6FRG, 6FRH
Ava DnaE	Anabaena variabilis ATCC 29413	$k = 3{,}1 \times 10^{-2}$ с$^{-1}$; $t_{1/2} = 23$ с	37	Экспериментальная структура в PDB не обнаружена
Npu DnaE	Nostoc punctiforme PCC 73102	$k = 3{,}7 \times 10^{-2}$ с$^{-1}$; $t_{1/2} = 19$ с	37	PDB 2KEQ, 4KL5, 4LX3, 4QFQ
NrdJ-1	Морской метагеном; род и вид организма не установлены	$k = 9{,}8 \times 10^{-2}$ с$^{-1}$; $t_{1/2} \approx 7$ с	37	Экспериментальная структура в PDB не обнаружена
gp41-1	Метагеном; род и вид организма не установлены	$k = 1{,}8 \times 10^{-1}$ с$^{-1}$; $t_{1/2} \approx 4$ с	45	Экспериментальная структура в PDB не обнаружена

Данные в таблице упорядочены по возрастанию скорости реакции белкового сплайсинга. Наиболее быстрыми среди рассмотренных вариантов являются интеины gp41-1 и NrdJ-1, сплайсинг которых занимает считанные секунды. Оба интеина были выявлены при анализе метагеномных последовательностей, поэтому организмы их происхождения не установлены, что затрудняет интерпретацию естественного биологического контекста и эволюционных связей. Кроме того, кинетика gp41-1 была измерена при 45 ◦C, поэтому ее непосредственное сопоставление с данными, полученными при 37 ◦C, и перенос результатов на условия культивирования клеток человека требуют осторожности. Аналогичное ограничение относится к SspDnaB, кинетические параметры которого определяли при 25 ◦C.

Наиболее наглядными объектами для сравнительного анализа являются интеины группы DnaE, кинетика которых была исследована при 37℃ - температуре, наиболее актуальной для применения в эукариотических системах. Эти интеины локализованы в гомологичных белках-хозяинах DnaE, но существенно различаются по скорости реакции, что позволяет сопоставлять их с учетом близкого эволюционного родства.

Высокая скорость сплайсинга, гомология последовательностей, наличие трёхмерных структур и температурные условия реакции делают интеины группы DnaE удобной моделью для структурных и экспериментальных исследований. Все последующие анализы и эксперименты будут проводиться именно на интеинах данной ферментативной группы. В следующих разделах рассмотрим общие характеристики интеинов и их структурные особенности на примере группы DnaE.

1.2. Общие характеристики интеинов для белкового сплайсинга

1.2.1. Определение интеинов и их роль в живых системах

Интеины—группа вспомогательных белков, инициирующих и реализующих реакцию белкового сплайсинга. Своё название интеины получили от сокращения «**inte**rnal prot**ein**» по аналогии с генетическим сплайсингом и понятием интронов «**intr**agenic regi**on**», где первая часть указывает на направление действия функционального элемента (Perler, 1994) [6]. Номенклатура интеинов аналогична номенклатуре рестриктаз: за трёхбуквенным обозначением родовой и видовой принадлежности организма следует обозначение гена-хозяина, с которым ассоциирован интеин. Из обзорного номенклатурного анализа известных природных интеинов можно сделать два основных вывода [7].

Во-первых, интеины распространены в трех основных доменах жизни (бактерии, археи и эукариоты), однако наибольшее видовое разнообразие демонстрируют бактерии (124 вида), что показано на Рисунке 2A. Во-вторых, интеины и их хост-гены реализуют базовые механизмы репликационного цикла (Рисунок 2B), что обеспечило их сохранение во всех трёх доменах жизни, где представлен 21 общий белок-хозяин. Намного реже белки-хозяева участвуют в процессах синтеза РНК, деградации белков и АТФ-зависимого транспорта, однако первые интеины были обнаружены в ходе изучения именно вакуолярной АТФазы в аскомицетах Neurospora crassa [6]. Было отмечено, что основной белок АТФазы с молекулярной массой более 500 кДа содержит высоко консервативную основную субъединицу массой 67 кДа, которая гомологична аналогичным ферментам у растений и дрожжей. Этот результат послужил основанием для дальнейших исследований консервативных регионов АТФаз на уровне последовательностей, а затем — функций кодируемых белков, что позволило открыть явление белкового сплайсинга и обосновать причины присутствия интеинов в разных эволюционных группах организмов.

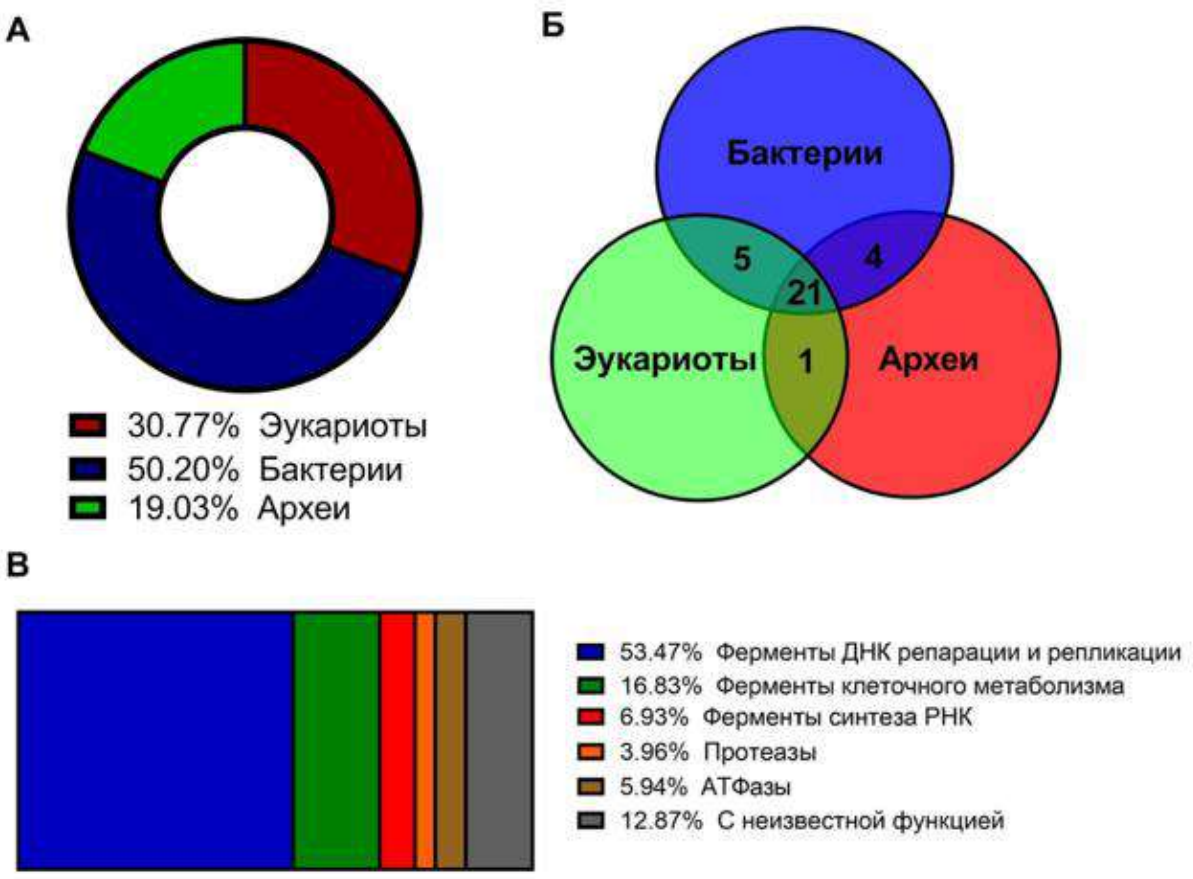


Рисунок 2 — Распределение интеинов: А — по доменам жизни с указанием числа видов; Б — по количеству генов-хозяев, уникальных и общих для трёх доменов жизни; В — по функциям белков-хозяев

Вопросу эволюции и распространения интеинов посвящено множество обзоров, ключевой темой которых является их нишевая роль у бактерий и редкая встречаемость у эукариот [2,9-11]. В нескольких работах встречается предположение о вирусном происхождении интеинов, однако практически все исследователи характеризуют их как «эгоистичные» или «паразитарные» мобильные генетические элементы [8]. Сохранение и распространение интеинов поддерживаются исключительно тесной связью с такими базовыми клеточными процессами, как репликация, репарация и синтез нуклеиновых кислот.

1.2.2. Механизм белкового сплайсинга в общем виде

Несмотря на явно выраженную паразитическую функцию интеинов, можно предположить, что их основная роль заключается в снижении затрат клеточных ресурсов на одновременный синтез больших белковых субъединиц.

Интеины позволяют разделить один белок на меньшие части и далее восстановить целую молекулу менее ресурсозатратным методом белкового сплайсинга, основанного на химических реакциях. Отдельным преимуществом является возможность организации доменов как в составе одной белковой молекулы для последующего цис-сплайсинга, так и в составе двух независимых белковых субъединиц для последующего транс-сплайсинга. Для наглядности варианты белкового сплайсинга схематично изображены на Рисунке 3.

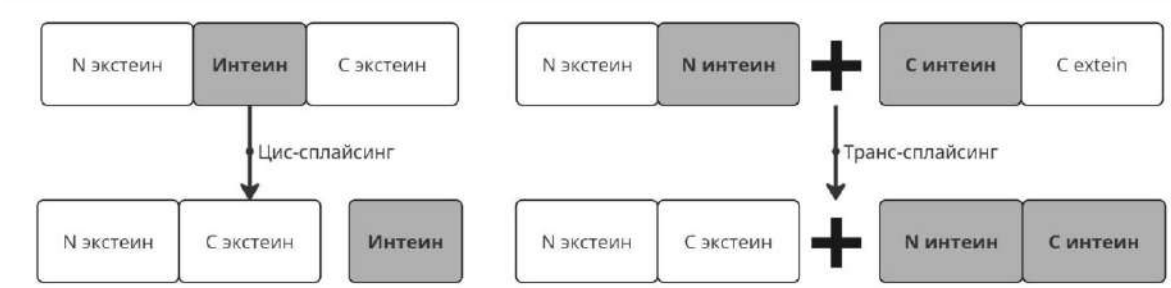


Рисунок 3 — Схема обмена белковыми доменами в рамках одной (слева - цис-сплайсинг) или двух молекул (справа - транс-сплайсинг) с помощью интеинов

Описанная перестройка белковых доменов обеспечивается способностью интеинов осуществлять белковый сплайсинг. Данная реакция является аналогом РНК-сплайсинга, но за счет биохимии белковых взаимодействий проходит более точно и направленно. Ключевой особенностью интеинового сплайсинга является полностью автономный и автокаталитический характер самой реакции, которая протекает по механизму, представленному на Рисунке 4.

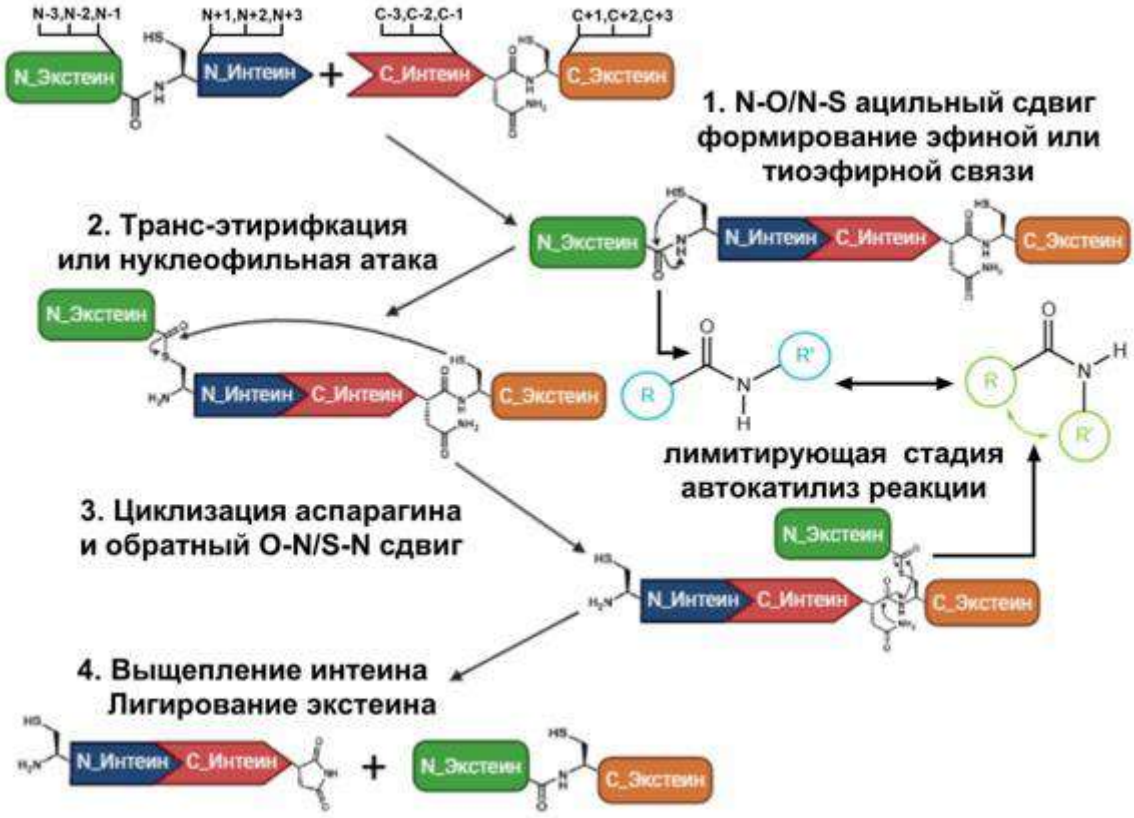


Рисунок 4 — Основные этапы белкового сплайсинга с участием интеинов

Последовательность этапов реакции, представленных на Рисунке 4, описана ниже:

1. N-S/N-O-ацильный сдвиг — первый и ключевой этап реакции, в ходе которого происходит разрыв исходной пептидной цепи при образовании эфирной или тиоэфирной связи между N-экстеином и первым остатком интеина (атомом кислорода серина или атомом серы цистеина) [9]. В результате данной реакции формируется высокоактивная молекула, однако данный процесс является энергетически невыгодным, поскольку осуществляется переход из исходного стабильного состояния в новое реакционноспособное состояние. Для преодоления данного барьера реакции требуется источник энергии, который формируется благодаря изменению энергетически невыгодной цис-конформации последних аминокислот экстеина и первых остатков интеина [12].

Таким образом, данная стадия является ключевой, поскольку за счёт компенсации разницы энергетических состояний, обеспечивающей необходимый автокатализ реакции, запускается весь каскад последующих реакций. Перечисленные условия накладывают ряд существенных ограничений на состав N-экстеиновой последовательности, что

подтверждается анализом крайне ограниченного аминокислотного состава в N-1/C-1 и N+1/C+1 положениях в последовательностях природных интеинов.

2.Трансэтерификация — второй этап, в ходе которого происходит пространственный перенос С-части к N-части экстеина с помощью нуклеофильной атаки. Сформированный в ходе образования эфира сайт на N-части является определённым участком взаимодействия с ОН-группой на С-части, в результате чего реакция проходит по базовым законам органической химии. Благодаря простоте и высокой скорости данного механизма нуклеофильная атака также встречается в ходе синтеза нуклеотидных последовательностей [13] и при работе других белковых комплексов, включая протеазы и полимеразы [14]. В данном случае результатом реакции является разветвленное промежуточное соединение с двумя N-концами, смещенным зарядом и сложной структурой, которая образуется исключительно благодаря формированию свободного эфира на первом этапе.

3.Последующая циклизация аспарагина или глутамина упрощает разветвленную структуру с помощью образования сукцинимидного кольца, которое замыкает разомкнутый цикл на С-конце интеина. Здесь также важно обратить внимание на роли аспарагина в положении С-1 и гистидина в положении С-2 интеина. Положение аспарагина строго зафиксировано с точки зрения органической химии для замыкания цикла, а положение гистидина существенно ускоряет процесс циклизации [15]. В результате интеин выщепляется в виде отдельной молекулы, а на экстеиновой молекуле спонтанно происходит O-N/S-N сдвиг, который превращает эфирную связь, образовавшуюся на первом этапе, в обычную пептидную, но уже в другом окружении аминокислот.

4. Итогом реакции сплайсинга является пространственная «перетасовка» белковых доменов с образованием большей экстеиновой и меньшей интеиновой субъединицы. При этом важно отметить, что после восстановления пептидной цепи необходимы благоприятные условия для формирования корректной вторичной и третичной структуры белка. Если фолдинг интеина определяется его природной аминокислотной последовательностью, то условия фолдинга экстеина варьируются в зависимости от свойств конкретного разделяемого белка.

Отдельно стоит упомянуть и о роли прямого и обратного ацильных сдвигов, которые не только запускают реакцию сплайсинга, но и объясняют раздельное выщепление N- и С-экстеинов на этапе сплайсинга, поскольку сдвиг происходит спонтанно благодаря запасу энергии, обусловленному благоприятной конформацией аминокислотных остатков на стыке интеина и экстеина. В результате реакции могут образовываться три варианта продуктов, которые представлены на Рисунке 5.



Рисунок 5 — Результат реакции ацильного сдвига в случаях автосплайсинга каждой из частей и сплайсинга при взаимодействии двух субъединиц

Важно отметить, что реакция сплайсинга определяется как структурными особенностями и физико-химическими взаимодействиями отдельных частей интеинов, так и влиянием отдельных аминокислот, запускающих саму реакцию сплайсинга. В следующем разделе уточним информацию о структурных особенностях интеинов.

1.3. Структурные особенности интеинов на уровне аминокислотной последовательности

Материал следующих разделов изложен в порядке усложнения уровней организации белка: от первичной аминокислотной последовательности до четвертичной структуры белковых комплексов. Структурные особенности рассмотрены на примере интеинов группы DnaE, используемых в дальнейшей экспериментальной работе.

Сначала рассмотрим некоторые базовые структурные особенности интеинов, которые обеспечат необходимое понимание функционального разделения аминокислотной последовательности на домены.

Во многом определение структуры и функций интеиновых доменов стало возможным после сравнительного эволюционного анализа на основе множественных выравниваний аминокислотных последовательностей и определения консенсусных

областей. Отметим три основных работы, заложивших основы для дальнейших исследований.

Обобщённая структура интеина, представленная на Рисунке 6, была описана в работе [16], в которой интеин подразделяли на три функциональные части: N-сплайсинговый домен, DOD-экзонуклеазный или линкерный домен, C-сплайсинговый домен. N- и C-части образуют Hint-домен (horseshoe-like structure), реализующий реакцию сплайсинга, а эндонуклеазный домен DOD (или LAGLIDADG) обеспечивает распространение, или «хоминг», интеинов в популяции.

Эндонуклеазная активность позволяет выщепленному интеину переходить от исходной «+»-аллели на «-»-аллель после сплайсинга путем разрезания непалиндромной дцДНК на уровне белка и дальнейшего встраивания последовательности интеинов после репарации двухцепочечного разрыва с использованием в качестве матрицы аллельной ДНК, содержащей последовательность интеина. За счет простоты и универсальности механизма хоминга осуществляется эффективный горизонтальный и вертикальный перенос интеина между организмами. Однако стоит заметить, что мини- и сплит интеины, которые не обладают DOD-доменом, лишены возможности эффективного распространения в популяции по стандартному механизму и используют альтернативные стратегии горизонтального переноса [2].

Более подробное деление на домены с восемью буквенными обозначениями было представлено в работе [17]. По результатам выравниваний 36 наиболее гомологичных интеинов из трех доменов жизни были определены наиболее консервативные аминокислоты. Наибольшую концентрацию вариабельных позиций наблюдали в B-домене N-части и H-домене DOD-части, в то время как A-домен N-части, а также F- и G-домены C-части были представлены в наиболее консервативном виде. Последний факт обусловлен образованием наиболее важного функционального активного центра на месте стыковки, обеспечивающего прохождение реакции сплайсинга по механизму, описанному ранее.

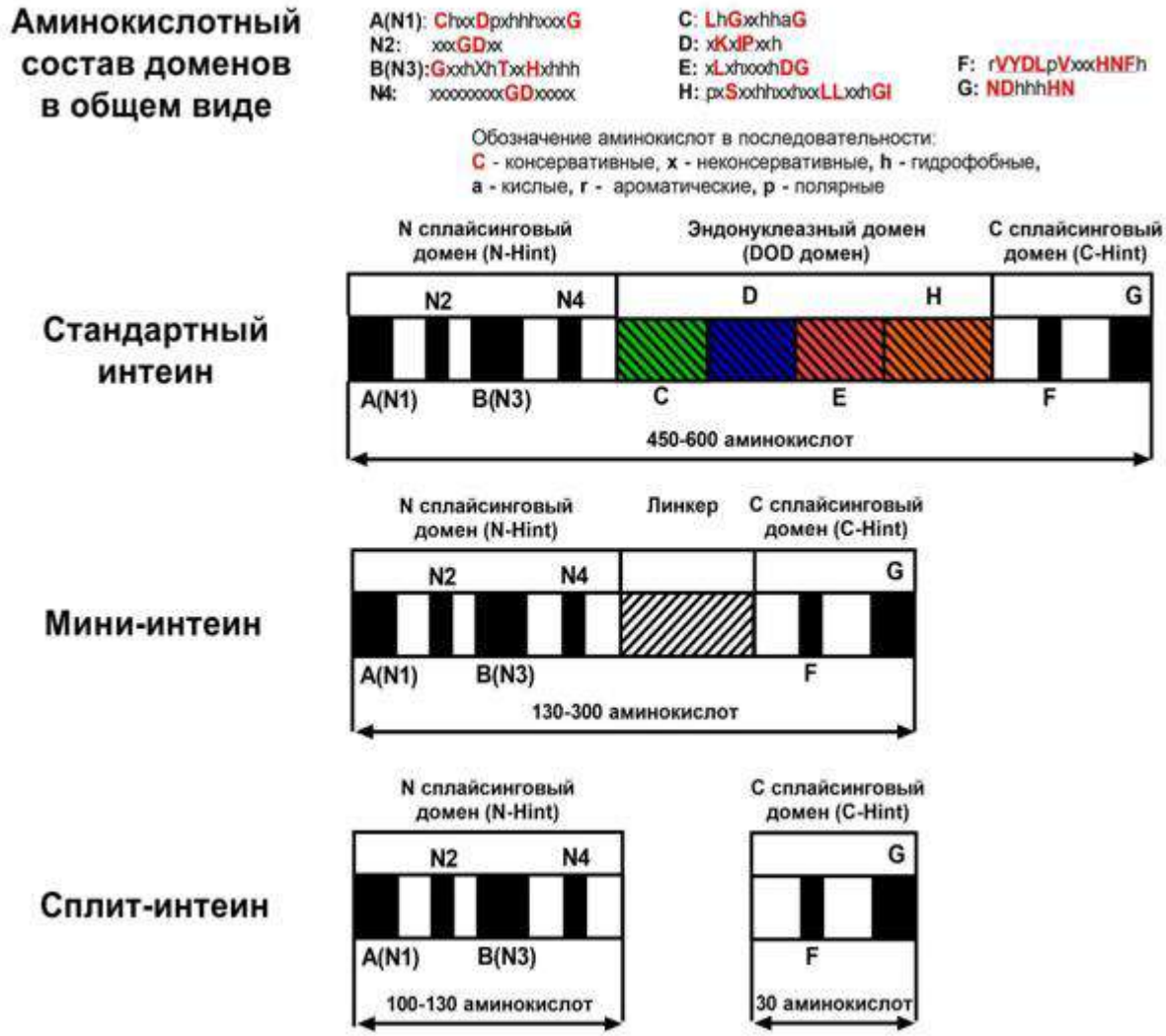


Рисунок 6 — Доменная организация аминокислотной последовательности стандартного и мини интеина [16,17]

Эндонуклеазный домен (в частности, регион C, или LAGLIDADG) также содержит ряд консервативных остатков, обеспечивающих хоминг. В остальных регионах присутствуют отдельные консервативные аминокислотные остатки, функции которых будут рассмотрены далее.

Другой вариант деления последовательности интеина представлен в работе [18], где выделены 15 доменов. В большинстве случаев их границы совпадают с границами доменов в классификации Perler, однако в N сплайсинговой части дополнительно выделены домены N2 и N4. При этом регион N2 представлен во всех интеинах, а N4 — не во всех. C-сплайсинговый домен дополнительно разделён на пять автономных частей (CAD); также проведено сравнение доменов CAD5 и N2.

Вывод о высокой гомологии этих областей не позволил однозначно определить их функции, поскольку на момент исследования было доступно недостаточно экспериментально определённых структур. Тем не менее наиболее успешным оказалось деление по Perler, поскольку большая часть выводов Pietrokovski не нашла дальнейшего подтверждения.

Ещё одна закономерность обсуждалась сначала Perler [17], а затем Ellila [19] при сравнении интеинов из разных организмов и генов-хозяев. При выравнивании природных последовательностей было выявлено преобладание цистеина или серина в положении N+1, а также мотива XHN на C-конце домена G. В положении C+1 встречаются цистеин, серин и треонин, что также обусловлено химией реакции сплайсинга. Остатки в положениях N-1, C+2 и C+3 имеют случайное распределение, что необходимо учитывать при использовании интеинов для получения рекомбинантных белков.

Следующий этап анализа аминокислотных последовательностей был проведён [20] уже при наличии экспериментально определённых структур и на выборке наиболее гомологичных интеинов одной группы хостовых белков третьей субъединицы эпсилон-полимеразы. По имеющимся дополнительным данным о кинетических параметрах реакции сплайсинга и результатах выравнивания был сделан вывод о функционально значимых регионах в сплит-интеинах группы DnaE. Рассмотрим наиболее показательный результат данного исследования на Рисунке 7.

Используя предыдущую классификацию, отметим наличие всего четырех доменов у сплит-интеинов (A, B, F, G), которые формируют основные функциональные активные центры. Отметим также, что домены B и F содержат повторяющийся мотив KDH(N/K)F, который образует бета-лист в N- и C-частях, что соответствует гомологии доменов N2 и CAD5. Другие функционально значимые консервативные бета-листы формируются аминокислотами в положениях N+1. . . N+4 и C-3. . . C-1, а остальные альфа-спирали и бета-листы содержат больше различий. Отметим, что наиболее вариабельные аминокислоты в основном сосредоточены в неструктурированных областях и большем бета-листе C-части интеина. Для наглядного представления связей между доменами и вторичными структурами рассмотрим визуализацию целой молекулы NpuDnaE в следующем разделе.

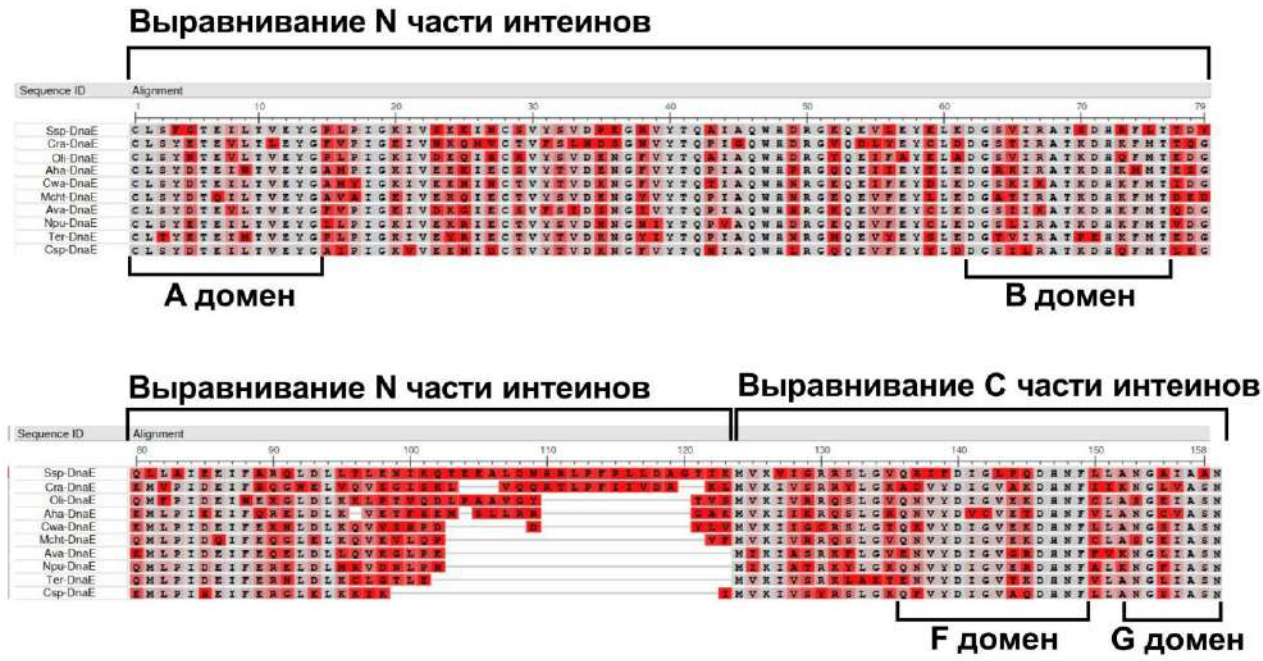


Рисунок 7 — Выравнивание аминокислотных последовательностей 10 сплит-интеинов группы DnaE с разделением на консервативные и вариабельные домены. Далее приведена расшифровка обозначений: Ssp-DnaE — Synechocystis sp. PCC 6803; Cra-DnaE — Cylindrospermopsis raciborskii CS-505; Oli-DnaE — Oscillatoria limnetica strain Solar Lake; Aha-DnaE — Aphanothece halophytica; Cwa-DnaE — Crocosphaera watsonii WH8501; Mcht-DnaE — Microcoleus chthonoplastes PCC 7420; Ava-DnaE — Anabaena variabilis ATCC 29413; Npu-DnaE — Nostoc punctiforme PCC 73102; Ter-DnaE — Trichodesmium erythraeum IMS101; Csp-DnaE — Cyanothece sp. PCC 8801.

1.3.1. Структурные особенности интеинов группы DnaE

Взаимосвязь аминокислотной последовательности и пространственной структуры наиболее наглядно можно рассмотреть на примере интеинов семейства DnaE.

Наиболее изученные природные разделённые интеины группы DnaE происходят из цианобактерий. К ним относятся Ssp DnaE из Synechocystis sp. PCC 6803, Npu DnaE из Nostoc punctiforme и Ava DnaE из Anabaena variabilis, которые далее будут рассматриваться подробнее. Эти интеины имеют родственные аминокислотные последовательности и общий контекст расположения в белке-хозяине, однако различаются по скорости сплайсинга, чувствительности к условиям реакции и экстеиновому окружению [4, 20, 21].

Название DnaE связано с белком-хозяином — каталитической αсубъединицей бактериальной ДНК-полимеразы III. Этот фермент участвует в репликации генома, обеспечивая копирование ДНК при подготовке клетки к делению. В составе полимеразного

комплекса DnaE непосредственно катализирует удлинение новой цепи ДНК посредством последовательного присоединения дезоксирибонуклеотидов к её 3′-концу в соответствии с последовательностью матрицы. Таким образом, DnaE является компонентом репликационного аппарата, тогда как встроенный в неё интеин выполняет иную функцию — обеспечивает созревание белка-хозяина посредством белкового сплайсинга [21].

Природная разделённая организация подробно описана для Ssp DnaE. Nи C-концевые части полимеразной субъединицы кодируются отдельными генами DnaE-N и DnaE-C, расположенными в разных участках генома. Продукт первого гена содержит N-концевой фрагмент полимеразы, соединённый с N фрагментом интеина, а продукт второго C-фрагмент интеина, соединённый с C-концевой частью полимеразы. После взаимодействия интеиновых фрагментов происходит транс-сплайсинг: интеиновые последовательности удаляются, а фрагменты полимеразы, выступающие в роли экстеинов, соединяются пептидной связью. В результате восстанавливается непрерывная последовательность DnaE. Интеин не сохраняется в составе зрелой полимеразной субъединицы и сам не осуществляет синтез ДНК [21].

Применение интеинов DnaE в эукариотических системах основано не на переносе функции бактериальной полимеразы, а на возможности использовать их сплайсинговую активность для соединения фрагментов других белков. В рекомбинантных конструкциях природные экстеины заменяют N- и C-концевыми частями целевой молекулы. Природная разделённая организация делает эти системы пригодными для работы с двумя отдельно экспрессируемыми цепями. Дополнительным преимуществом отдельных представителей, включая Npu DnaE и Ava DnaE, является высокая активность в модельных реакциях при 37 °C — температуре, используемой при культивировании многих клеток млекопитающих. При этом Ssp DnaE в сопоставимых тестовых системах характеризуется меньшей скоростью, поэтому высокую активность нельзя автоматически приписывать всем представителям семейства [4, 20].

Для генной терапии такое свойство позволяет распределить последовательность крупного целевого белка между двумя векторными кассетами и восстановить белок после совместной доставки и экспрессии обеих конструкций в одной клетке. В случае AAV этот подход обходит ограничение вместимости одного вектора, не увеличивая физическую упаковочную ёмкость капсида. Возможность интеин-опосредованной сборки крупных белков в тканях млекопитающих показана, в частности, при исследовании доставки ABCA4 в сетчатку [22]. Однако активность интеина в модельной реакции не гарантирует эффективной сборки любого белка: результат зависит от экстеинового окружения, свойств фрагментов и условий их экспрессии [4, 22]. Поэтому выбор DnaE для настоящей работы

определяется сочетанием изученности, природной разделённой организации и активности конкретных представителей, а их пригодность для целевых конструкций требует экспериментальной проверки.

1.3.2. Структурные особенности интеинов группы DnaE на уровне 3D-модели

Для сопоставления результатов выравнивания с пространственной организацией интеина на Рисунке 8 показано расположение выделенных участков в структуре Npu DnaE из Nostoc punctiforme (PDB ID: 4KL5). Панели A и B представляют выравнивание последовательностей десяти интеинов группы DnaE, а панель C позволяет сопоставить отмеченные в нём области с элементами трёхмерной структуры 4KL5 [23, 24].

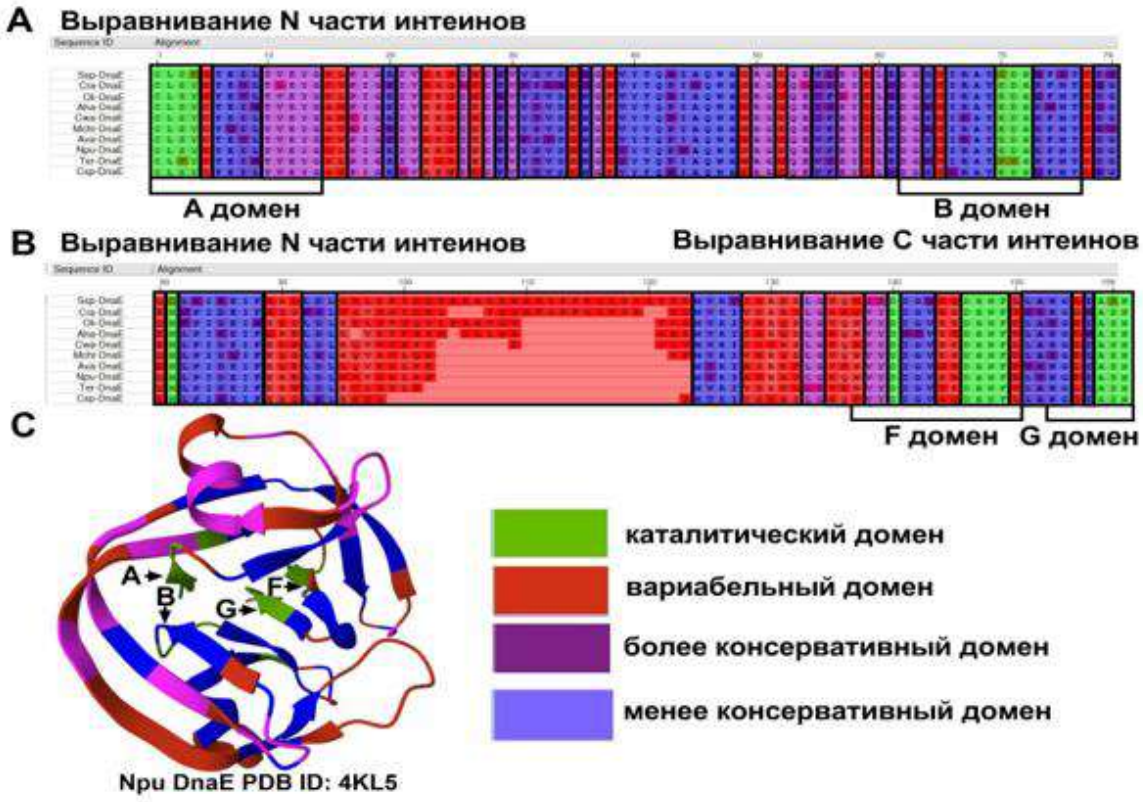


Рисунок 8 — Сопоставление выравнивания аминокислотных последовательностей десяти интеинов группы DnaE с пространственной структурой Npu DnaE.
A — начальная область N концевых фрагментов; B — продолжение N-концевых фрагментов и C-концевые фрагменты; C — расположение выделенных участков в структуре Npu DnaE (PDB ID: 4KL5). Рамками обозначены блоки A, B, F и G. Согласно цветовой схеме рисунка, зелёным выделены каталитически значимые участки, красным — вариабельные, фиолетовым — более консервативные, сине-фиолетовым — менее консервативные участки.

Представленное выравнивание демонстрирует неравномерное распределение консервативных и вариабельных областей. N-фрагмент содержит протяжённые консервативные последовательности между блоками A и B; наибольшее разнообразие аминокислот наблюдается во второй половине N-фрагмента. C-фрагмент отличается небольшой длиной и чередованием вариабельных и консервативных областей. Несмотря на присутствие замен в C-концевых фрагментах, блоки F и G содержат хорошо сохраняющиеся позиции, выделенные зелёным цветом.

Сопоставление аннотации последовательности со структурой показывает, что блоки A, B, F и G, удалённые друг от друга в линейной последовательности, сближаются в сложенной структуре и образуют единый каталитический центр. Более консервативные области N-фрагмента, выделенные синим цветом, необходимы для формирования структурного каркаса. Он обеспечивает подковообразную укладку N-фрагмента, сформированную преимущественно β-структурами с двумя короткими α-спиралями (Shah, 2012) [4]. При этом C-часть имеет относительно свободную укладку из-за небольшого размера образующейся после трансляции молекулы и малого числа элементов вторичной структуры. Наибольшая вариабельность аминокислот наблюдается в неструктурированной области между спиралями и листами. Эти различия наиболее наглядны при сравнении C-концевых участков N-фрагментов, общая длина которых варьирует от 102 до 123 аминокислотных остатков.

Обратим внимание также на общую закономерность, характерную для мини-интеинов группы DnaE, — асимметрию размеров фрагментов: у Ssp DnaE N-фрагмент включает 123 аминокислотных остатка, а C-фрагмент — 36; у Npu DnaE соответствующие длины составляют 102 и 36 остатков [25, 26]. Вместе обе части образуют каталитическую α-субъединицу ДНК-полимеразы III, участвующую в репликации бактериального генома. Для сборки данного белкового комплекса важны пост-трансляционное формирование отдельных субъединиц и их первоначальное электростатическое взаимодействие, позволяющие завершить упорядочение N-структуры. Этот механизм получил название «захват и коллапс» (capture and collapse) (Shah, 2013) [27]. Следовательно, при сравнительном анализе DnaE значение имеют не только остатки активного центра, но и аминокислоты, определяющие межфрагментные контакты и формирование общей пространственной структуры. Далее сравним последовательности внутри семейства DnaE с учётом структурных особенностей и кинетики реакции интеиновых пар.

1.3.3. Сравнение кинетических характеристик интеинов группы DnaE

Внутрисемейная вариабельность проявляется в скорости сплайсинга, количестве образуемого продукта, длине N-фрагмента и аминокислотном составе обоих фрагментов интеина. На Рисунке 9 сопоставлены результаты выравнивания последовательностей DnaE, значения времени полуреакции сплайсинга и дендрограмма сходства последовательностей.

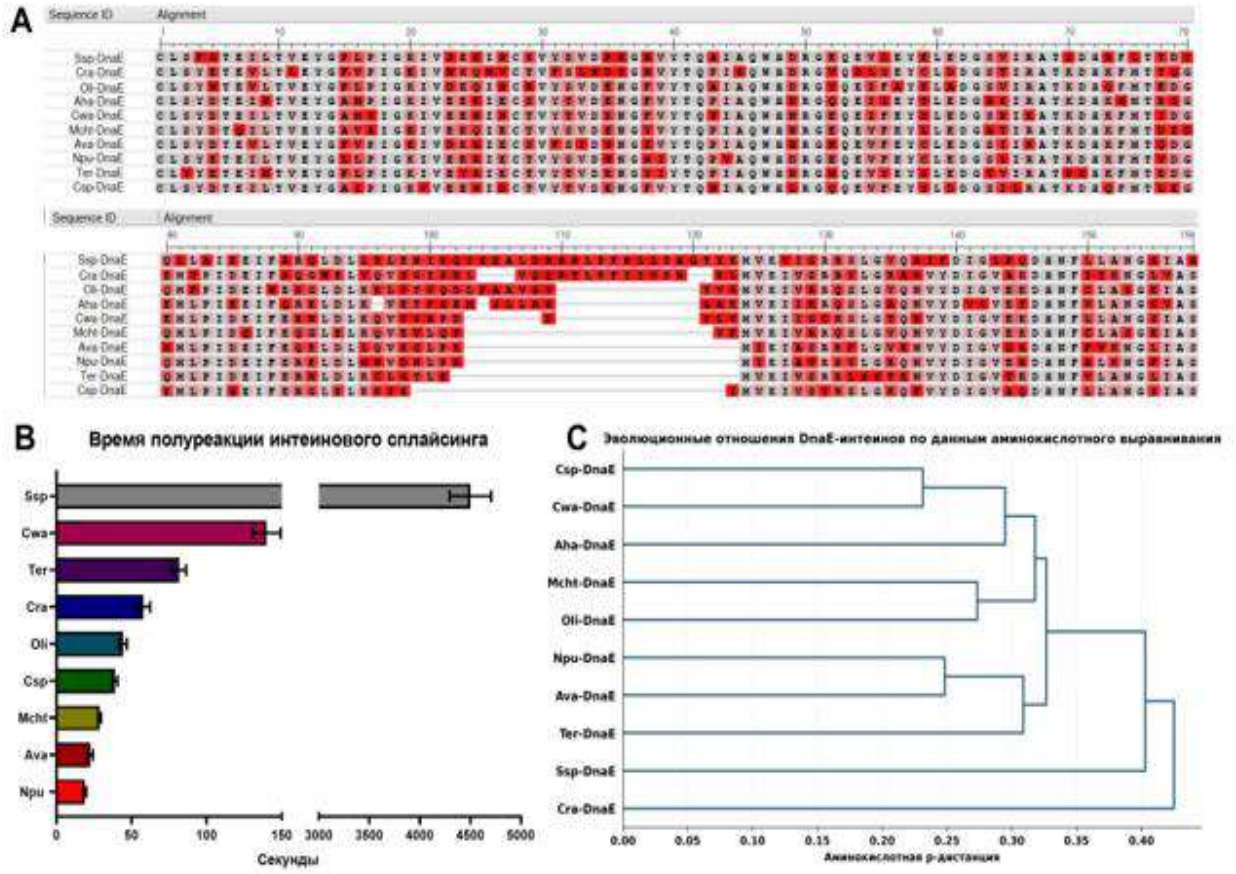


Рисунок 9 — Сопоставление последовательностной вариабельности и кинетических характеристик интеинов группы DnaE.

A — выравнивание аминокислотных последовательностей десяти интеинов с цветовым выделением различий; разрывы отражают различия в длине сопоставляемых участков. B — время полуреакции сплайсинга для девяти представителей, выраженное в секундах; горизонтальная ось имеет разрыв для отображения значения Ssp DnaE. C — дендрограмма, построенная по результатам аминокислотного выравнивания; горизонтальная ось обозначена как аминокислотная p-дистанция.

На панели A видны как сохраняющиеся позиции, так и многочисленные замены, распределённые вдоль последовательностей. Особенно заметны различия в длине участка перед началом C-фрагмента: наиболее протяжённая последовательность в этой области представлена у Ssp DnaE, тогда как у ряда других интеинов, включая Ava DnaE и Npu DnaE, ей соответствуют существенно более короткие участки. Таким образом, внутрисемейные различия включают не только замены аминокислот, но и изменение длины отдельных областей.

Панель B показывает широкий разброс времени полуреакции. У Npu DnaE и Ava DnaE оно составляет приблизительно 20 с, у Cwa DnaE — около 140 с, а у Ssp DnaE — около 4500 с. Меньшему времени полуреакции соответствует более быстрое протекание сплайсинга и большее количество синтезированного продукта в условиях одного эксперимента. Разрыв оси времени подчёркивает существенное отличие Ssp DnaE от остальных интеинов в данном семействе.

Дендрограмма на панели C отражает сходство исследованных последовательностей. На ней выделяются несколько парных кластеров: Npu DnaE–Ava DnaE, Csp DnaE–Cwa DnaE и Mcht DnaE–Oli DnaE. При этом сходство последовательностей не всегда сопровождается сходством кинетических характеристик: Npu DnaE и Ava DnaE близки по времени полуреакции, тогда как Csp DnaE и Cwa DnaE заметно различаются по этому показателю. Ssp DnaE и Cra DnaE присоединяются к остальной группе при более высоких значениях дистанции, но существенно отличаются друг от друга по скорости сплайсинга.

В данном сравнении наиболее контрастными являются интеины Npu DnaE и Ssp DnaE: для них доступны трёхмерные структуры, отношение кинетических показателей достигает 200, а длина N-частей различается на 21 аминокислотный остаток. Такое сочетание характеристик удобно для экспериментальной проверки гипотез о получении более эффективного интеина.

Отметим также, что при выравнивании структур всех 10 интеинов в каталитически активной области была продемонстрирована полная идентичность последовательностей, что связано с критической ролью аминокислот на границе с экстеином. Однако если консервативность аминокислот интеина определяется его принадлежностью к ферментативному семейству, то состав экстеинов в биотехнологических конструкциях может меняться в зависимости от того, какой белок требуется разделить и затем собрать посредством сплайсинга. В следующем разделе рассмотрим влияние аминокислотного состава экстеинов на сплайсинг с участием интеинов группы DnaE.

1.3.4. Влияние положения и состава аминокислот экстеинов при сплайсинге с участием интеинов группы DnaE

Как описано в пункте 1.2.2, ацильные сдвиги со стороны экстеинов являются начальной и завершающей стадиями реакции сплайсинга. Промежуточная реакция, связанная с нуклеофильной атакой, определяется интеином, аминокислотный состав которого остаётся консервативным в пределах выбранного ферментативного семейства. Экстеины же могут варьировать в зависимости от задачи сборки крупного белка. Далее рассмотрим выводы, сделанные на ранних этапах изучения механизма транс-сплайсинга.

Одной из первых работ в этой области стало исследование [15], в котором были получены следующие результаты:

1. Замена серина на аланин в положении N+1 и изолейцина на лизин в положении N+2 блокирует реакцию на этапе образования эфирной связи и сплайсинга. Замены на цистеин и треонин в том же положении снижают эффективность, но не полностью блокируют эти этапы.
2. При замене аспарагина на аспарагиновую кислоту в положении C-1 реакция сплайсинга начинается, однако неполное замыкание цикла препятствует завершению этапа циклизации. В результате со временем накапливается промежуточное разветвлённое соединение с двумя экстеиновыми N-концами и интеином без образования отдельных молекул экстеина и интеина. Это указывает на блокировку реакции на предпоследнем этапе.
3. Замена гистидина на лейцин в положении C-2 позволяет реакции завершиться, однако количество продукта увеличивается только при повышении температуры до 65 градусов. Это, в свою очередь, указывает на возникновение энергетического барьера, преодоление которого требует большей подвижности частиц.
4. Замены серина на аланин или глицин в положении C+1 блокируют сплайсинг при сохранении остальных реакций. Замены на цистеин и треонин в том же положении снижают эффективность, но не полностью блокируют эти этапы.
5. Комбинированная замена серина на аланин в N+1 и аспарагина на аспарагиновую кислоту в C-1 полностью блокировала реакцию на всех этапах.
6. Замена лейцина в положении C-5 и тирозина в положении C-4 на метионин приводила к изменению эффективности сплайсинга в разных условиях.

Первый и наиболее очевидный вывод заключается в ключевой роли серина (цистеина и треонина) в положениях N-1 и C+1 экстеинов, определяющих протекание каждой полуреакции и сплайсинга в целом. Второй вывод касается жесткого закрепления аспарагина в положении C-1 и меньшего эффекта в случае замены гистидина в положении C-2. Замены остальных аминокислот, удалённых от места стыковки интеинов и экстеинов, влияют на прохождение реакции, но уже менее значительно и предсказуемо с точки зрения базовой химии белкового сплайсинга. Все описанные мутации вносили в конструкцию из мальтозы в качестве экстеина и интеина Psp-GBD Pol.

Во второй фундаментальной работе Cheriyan [28] также подчёркнута критическая роль положения и состава аминокислот на стыке интеинов и экстеинов. При этом анализ расширен до трёх аминокислотных остатков на каждом из реакционноспособных концов. По результатам экспериментального перебора всех аминокислот в обозначенных положениях авторы составили таблицу, включавшую частотное распределение 483 вариантов сборки белка устойчивости к антибиотику канамицину с помощью интеина NpuDnaE.

Для дальнейшего упрощения авторы перешли к балльной системе оценки влияния каждой аминокислоты и её положения, что продемонстрировано в Таблице 2.

Основные выводы, которые можно сделать из данного скринингового анализа:

1. Характер реакции сплайсинга определяют как N-, так и C-экстеины, однакоцистеин в положении C+1 играет решающую роль, поскольку его свободная OH/SH-группа осуществляет нуклеофильную атаку.

2. Второй по значимости аминокислотой является триптофан в положении C+2,однако в большинстве природных систем в этом положении встречается фенилаланин. Тем не менее указанная модификация привела к увеличению скорости реакции сплайсинга в 1,5 раза, что говорит об однозначном положительном влиянии триптофана на реакцию.

3. Влияние аминокислот N-экстеина оценено в меньшей степени, однако с точкизрения химии реакции образование эфирной связи является ключевым компонентом на стадии прямого и обратного ацильного сдвига, необходимого для разрыва исходной и итоговой пептидной цепи.

Данное исследование предоставило подробные сведения об аминокислотном составе концевых последовательностей экстеинов и интеинов группы DnaE. Важно отметить, что данные таблицы специфичны для интеинов каждой отдельно взятой ферментативной группы. В заключительном теоретическом разделе рассмотрим специфичность взаимодействия фрагментов разных интеиновых пар при одинаковом экстеиновом окружении.

Таблица 2 — Вклад от состава и положения каждой из аминокислот экстеинов на эффективность реакции транс-сплайсинга при использовании интеина NpuDnaE

	N-экстеин			C-экстеин		
3-х буквенный код/положение аминокислоты	-3	-2	-1	+1	+2	+3
C	0,46	0,4	0,8	50	0,07	0,07
M	0,13	0,73	1,26	20	0,07	2,25
W	0,07	0,4	0,99	0	29,42	0,07
K	1,59	0,93	4,44	0	0,07	0
N	0,93	1,26	2,19	0	0	8,61
H	1,06	1,26	2,19	0	0	8,61
Y	0,2	0,6	2,05	0	0,13	5,23
A	0,63	0,83	1,92	0	0	0
G	4,67	2,68	0,17	0	0	0,03
Q	1,39	0,86	0,33	0	0	0,27
D	1,39	1,52	0	0	0	1,26
E	1,26	3,51	0,07	0	0	5,5
V	0,56	0,5	0,23	0	0,03	0,1
F	0,13	0,73	1,26	0	0,07	2,25
R	1,81	0,2	0,86	0	0	0
S	0,99	1,24	1,37	0	0	0,07
P	0,8	1,52	0	0	0	0
T	0,63	0,89	1,59	0	0	0,96
L	0,11	0,57	0,75	0	0	0,82
I	0,07	0,53	0	0	0	0,4

1.3.5. Кросс-реактивные свойства интеинов группы DnaE

Описанные ранее особенности интеинов группы DnaE — ограниченная вариабельность экстеинов и полная идентичность интеиновых последовательностей в области каталитических остатков — ставят вопрос о специфичности взаимодействия фрагментов разных интеиновых пар при одинаковых экстеинах. Это свойство отдельно исследовалось в нескольких работах [29-35].

Основной предпосылкой этих исследований была необходимость разделить экстеиновую молекулу на три части и восстановить её правильную структуру с помощью интеинов с близкими кинетическими характеристиками и эффективностью сплайсинга [29]. В результате исследования была обнаружена возможность образования как промежуточных продуктов реакции, так и продуктов, не содержащих центральную часть.

В работе также подчёркивается, что использование двух разных интеинов не всегда решает эту задачу из-за их высокой гомологии. Следующее упоминание встречается в работе Iwai [30], в которой исследовали NpuDnaE для выявления элементов последовательности, влияющих на эффективность реакции сплайсинга. Для решения данной задачи использовали интеины SspDnaE и NpuDnaE, поскольку ранее было известно, что они различаются по эффективности сплайсинга в 20 раз. Выравнивание последовательностей выявило различие в длине N-частей интеинов на 20 аминокислот, а также ряд замен, распределённых по всей последовательности. Дополнительно в работе оценивается влияние аминокислоты в положении +2 в линкере CXNGT, необходимом для транс-сплайсинга. Подчёркивается, что интеин SspDnaE менее стабилен и активен только в присутствии фенилаланина, триптофана и тирозина. При этом также подчёркивается, что N- и C-части интеинов полностью взаимозаменяемы и комбинация NpuN+SspC имеет ту же эффективность, что и комбинация NpuN+NpuC. Единственное объяснение данного факта в контексте приведенного исследования заключалось в том, что вариация 20 аминокислот в N-части не являлась специфичной для реакции транс-сплайсинга.

Дальнейшее изучение данной комбинации было проведено в работе [4], где помимо выравнивания аминокислотных последовательностей проводили статическое моделирование «химер» SspDnaE_N+NpuDnaE_C и NpuDnaE_N+ SspDnaE_C. Также экспериментально подтвердили активность химерных вариантов с исходными последовательностями и с рядом групповых мутаций. Наиболее наглядными вариантами для сравнения оказались сочетания NpuDnaE_N+NpuDnaE_C; SspDnaE_N+NpuDnaE_C и вариант SspDnaE_N с заменами L56F, S70K, A83P и E85D. Оказалось, что мутации в

SspDnaE_N в сочетании с NpuDnaE_C в 1,5 раза увеличивают скорость реакции сплайсинга, однако подробного объяснения наблюдаемого эффекта в работе не приводится.

Еще одно недавнее упоминание кросс-реактивного свойства интеинов из группы DnaE встречается в работе [31]. Изучали комбинацию интеинов группы DnaE: AvaDnaE_N и NpuDnaE_C. NpuDnaE_C использовали как интеин с быстрой кинетикой и высоким выходом продукта сплайсинга, а AvaDnaE_N — как самый близкий эволюционный гомолог. В работе также со ссылкой на источник [32] указывается, что AvaDnaE_N является предпочтительным сплайсингпартнёром для NpuDnaE_C.

Помимо исследований взаимодействий между интеинами одной группы, в работе [34] продемонстрировали возможность взаимодействия между представителями одной ферментативной группы и блокировку реакции в случае перекрестного взаимодействия между разными группами (NspDnaE, TerDnaE3, SspDnaB). Общая схема изучения свойства кросс-реактивности представлена на Рисунке 10.

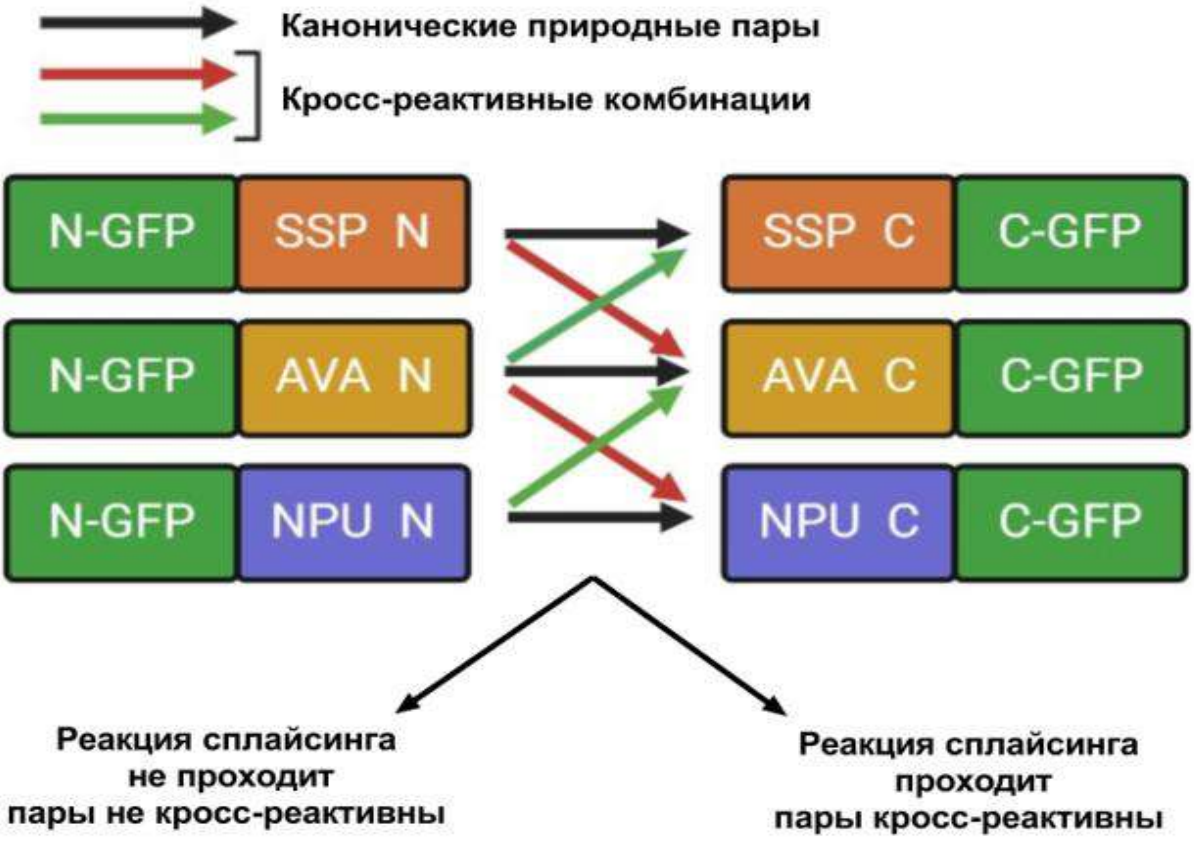


Рисунок 10 — Общая схема изучения свойства кросс-реактивности для интеинов, реализующих реакцию транс-сплайсинга

Из приведенных интеинов NspDnaE и TerDnaE-3 относятся к одной группе и одному классу белков, что делает их взаимодействия между собой более предсказуемыми, чем взаимодействия с интеином SspDnaB, который исходно реализует реакцию цис-сплайсинга, но был адаптирован методами генной инженерии для транс-сплайсинга. В результате исследования кросс-реактивных взаимодействий положительные данные были получены лишь для пары NspDnaE и TerDnaE-3, что объясняется их высоким функциональным сходством. Для пар NspDnaE/SspDnaB и TerDnaE-3/SspDnaB положительных результатов реакции транс-сплайсинга зафиксировано не было.

Выявленные различия в активности указывают на различия в структуре и функциях, а выравнивание аминокислотных последовательностей выявляет протяжённые несовпадающие участки в С-частях интеинов групп DnaE и DnaB. Это свойство позволяет совместно использовать данные интеины для сборки крупных белков из трёх и более частей, однако необходимо учитывать различия в скорости отдельных реакций сплайсинга.

Число рассмотренных исследований кросс-реактивности невелико, но совокупность имеющихся данных создаёт основу для дальнейшего изучения этого свойства. Такие исследования могут помочь охарактеризовать функциональную роль доменов и отдельных аминокислот в механизме сплайсинга. Далее подробнее рассмотрим применение реакции транс-сплайсинга для решения биологических и биотехнологических задач.

1.4. Примеры применения белкового сплайсинга в современной биотехнологии

Рассмотренные общие свойства интеинов и механизмы белкового транссплайсинга отражают основные направления фундаментальных исследований по данной тематике. Далее кратко рассмотрим прикладные направления, обозначенные на Рисунке 1: биосенсоры, биотехнологические приложения и генная терапия. Основное внимание уделим технологии транс-сплайсинга, позволяющей собирать крупные белковые молекулы.

1.4.1. Биосенсоры и методы мечения белков на основе интеинов

Селективное мечение белков и создание биосенсоров, позволяющих исследовать внутриклеточный транспорт и белок-белковые взаимодействия, востребованы как в фундаментальных, так и в прикладных исследованиях. Химическая гибкость интеинов и автономность их работы позволяют относительно просто создавать биомаркеры.

Схема модификации белков для визуализации представлена на Рисунке 11. Подход заключается в разделении белка на две части и добавлении к одной из них метки (флуорофора, биотина и т. д.). Ключевое преимущество использования интеинов заключается в природной асимметрии размеров их Nи C-частей, позволяющей получать меньшую часть (длиной от 6 до 36 аминокислот) методами химического синтеза с необходимыми модификациями [35]. Последующий транс-сплайсинг позволяет получить полноразмерную молекулу с необходимой меткой на N- или C-конце.

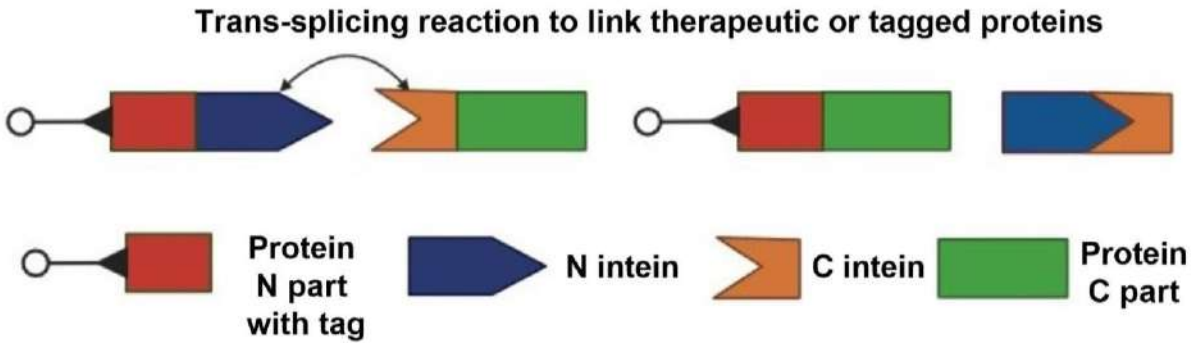


Рисунок 11 — Схема мечения белка с помощью интеин-опосредованного транс-сплайсинга. N- и C-фрагменты целевого белка (красный и зелёный) соединяются с сохранением метки на N-конце, а фрагменты интеина (синий и оранжевый) удаляются из состава продукта.

В биосенсорах используют белковый транс-сплайсинг для соединения фрагментов репортера: их сближение и последующая сборка позволяют преобразовать исследуемое событие в регистрируемый сигнал. Например, восстановление активности разделённой люциферазы с помощью интеина SspDnaE использовали для изучения регулируемого переноса белков в ядро [36,37].

Другой пример — система SIMPL2, в которой взаимодействие исследуемых белков способствует сближению фрагментов интеина и ковалентному соединению связанных с ними репортёрных фрагментов; последующая сборка люциферазного репортера обеспечивает биолюминесцентный сигнал [38].

Таким образом, интеины позволяют как вводить метки в белковые молекулы, так и связывать регистрацию сигнала с локализацией или взаимодействием белков, что востребовано при изучении клеточных процессов и тестировании биологически активных соединений.

1.4.2. Биотехнологические приложения интеинов для получения рекомбинантных белков

Развитие методов визуализации белков с помощью меток способствовало разработке методов их очистки. Наиболее полный и подробный обзор методов очистки белков за 30-летний период представлен в работе [39]. Из всего многообразия реализованных систем следует отметить две наиболее важные: CIVPS (Controllable Intervening Protein Sequences) [40] и IMPACT (Intein Mediated Purification with an Affinity Chitin-binding Tag) [41].

В системе IMPACT хитинсвязывающий домен обеспечивает удержание белковой конструкции на сорбенте, а контролируемое расщепление на границе интеина позволяет высвободить целевой белок. В зависимости от конструкции расщепление регулируют изменением pH, температуры или добавлением тиольных реагентов; отдельные варианты системы также позволяют получать циклические формы белков. Альтернативный подход основан на взаимодействии фрагмента интеина, закреплённого на сорбенте, с комплементарным фрагментом, связанным с целевым белком [42]. Схема очистки белка с использованием расщеплённого интеина представлена на Рисунке 12.

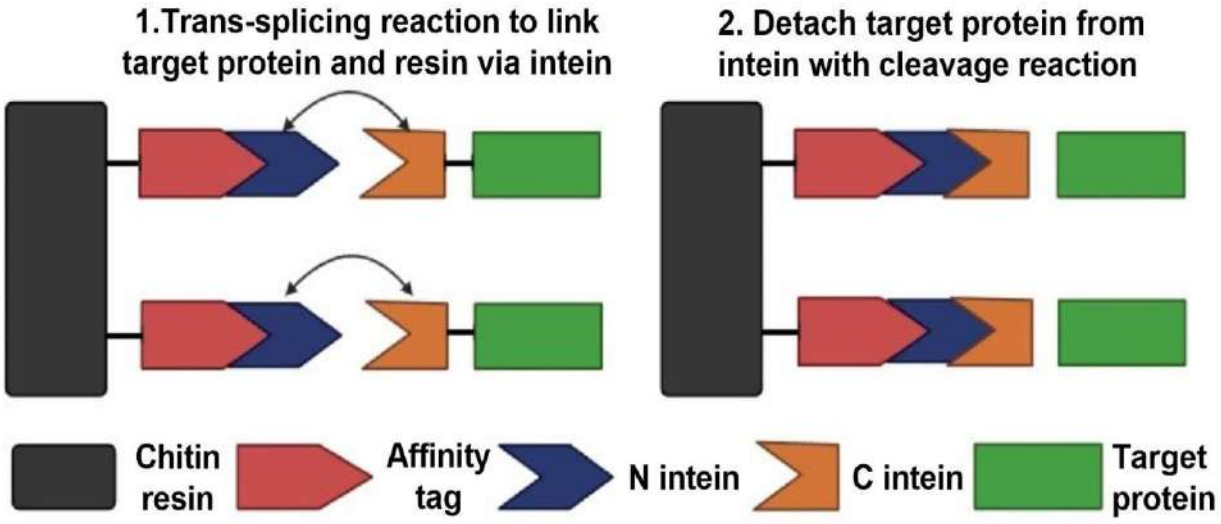


Рисунок 12 — Схема очистки белка с использованием расщеплённого интеина: 1—связывание целевого белка с хитиновым сорбентом через взаимодействие N- и C-фрагментов интеина; 2 — высвобождение целевого белка при расщеплении связи на границе с интеином. Аффинная метка показана красным, N-фрагмент интеина — синим, C-фрагмент — оранжевым, целевой белок — зелёным.

Второе важное применение белкового сплайсинга — сборка крупных белковых комплексов. Раздельное получение фрагментов позволяет снизить нагрузку на клетки, увеличить выход белка и защитить клетки-продуценты от воздействия его функционально активной полноразмерной формы, что особенно важно при получении протеаз в дрожжах и бактериях [43]. Для решения этой задачи протеазу разделили на два фрагмента, кодирующие последовательности которых объединили с последовательностями соответствующих частей интеина. Фрагменты нарабатывали раздельно и затем очищали с использованием хитин связывающего домена. Протеазу собирали в буферной системе на колонке посредством транс-сплайсинга, после чего концентрировали через полупроницаемую мембрану с заданным порогом отсечения.

В этой же работе обсуждалось применение гибридной пары интеинов NpuN+SspC, которая обеспечивала в 1,5 раза больший выход белка по сравнению с исходной парой NpuN+NpuC.

Достаточно сложной задачей на сегодняшний день остается синтез больших биспецифических антител без ошибочного перекрестного взаимодействия тяжелых и легких цепей. В результате дизайна, предложенного в работе [44], соединение фрагментов в области шарнирного участка позволило комбинировать Fab-домены различной специфичности и уменьшать число нежелательных вариантов сборки тяжёлых и лёгких цепей (Рисунок 14). Таким образом, интеины обеспечивают модульное получение белковых продуктов, сборку которых сложно контролировать при совместной экспрессии всех компонентов.

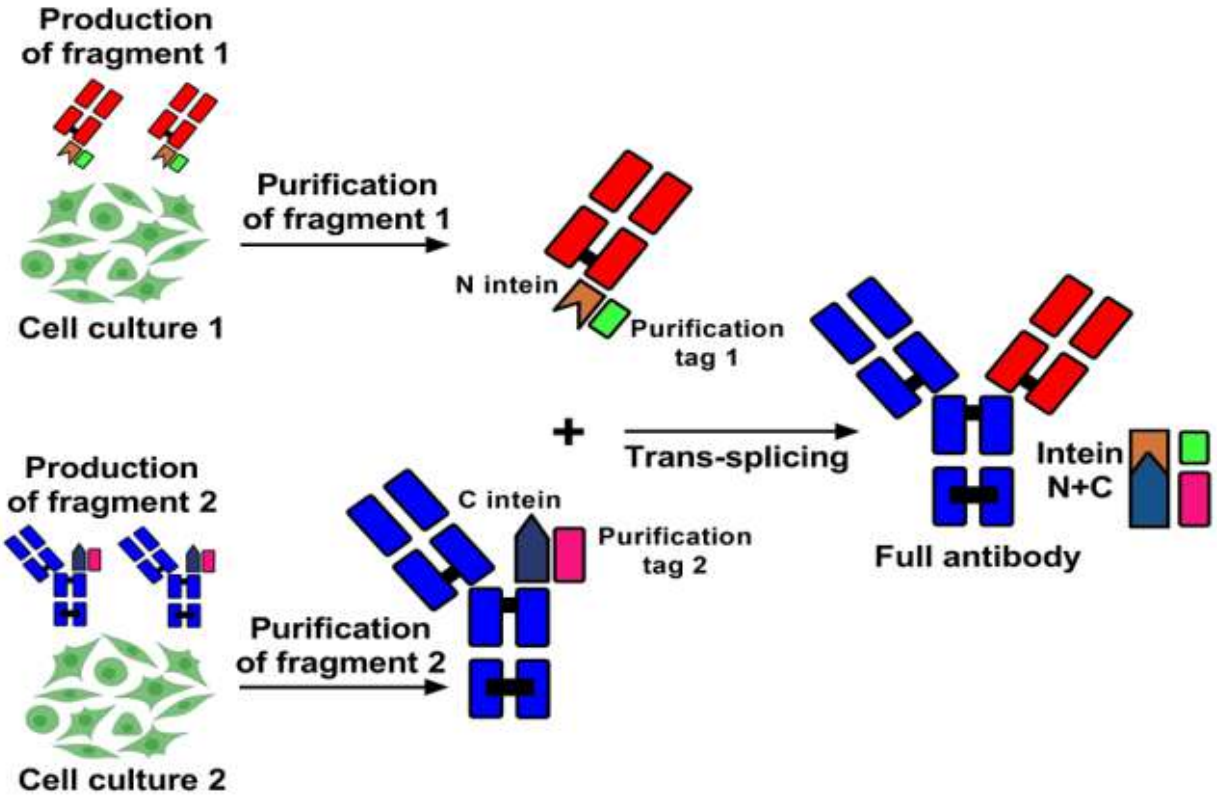


Рисунок 14 — Схема получения биспецифического антитела с помощью интеинопосредованного транс-сплайсинга. Два фрагмента антитела, соединённые с комплементарными фрагментами интеина и аффинными метками, синтезируют в отдельных клеточных культурах и очищают независимо. Последующий транс-сплайсинг объединяет фрагменты антитела в единую молекулу с удалением интеина и связанных с ним меток.

Ключевое преимущество этого подхода к получению биспецифических антител заключается в раздельном синтезе фрагментов, образующих при сборке только одну комбинацию. Белковый транс-сплайсинг уменьшает количество возможных вариантов сборки, снижает иммуногенность, позволяет проводить аффинную очистку антител, комбинировать Fab-домены и получать разнообразные биспецифические антитела.

Таким образом, рассмотренные примеры иллюстрируют основные области биотехнологического применения интеинов и технологии транс-сплайсинга. Далее сосредоточим внимание на применении интеинов в генной терапии для решения аналогичной задачи — сборки крупных белков в клетках после раздельной доставки генетического материала.

1.4.3. Применение интеинов в генной терапии

1.4.3.1. Интеины для заместительной генной терапии

Согласно статистическим данным, представленным на Рисунке 1, в течение последних 10 лет наблюдается растущий интерес к применению белкового трансплайсинга в генной терапии наследственных заболеваний. Более подробный анализ позволяет разграничить исследования до и после 2020 года: появление модифицированной технологии геномного редактирования на основе CRISPR/Cas9 существенно изменило стратегии редактирования генома.

Генная терапия наследственных заболеваний в период с 2010 по 2020 год предполагала использование стратегии заместительной доставки здоровой копии трансгена с помощью аденоассоциированных вирусов, поскольку данный тип доставки характеризовался низкой иммуногенностью, стабильной эписомальной экспрессией и высокой биологической безопасностью [45]. Ключевым препятствием для использования AAV является ограничение упаковочной емкости. Это ограничение сужает возможности их использования для терапии тысяч наследственных заболеваний, поэтому параллельно развивались и другие платформы [46]. Однако проблема возможной неспецифической интеграции ВИЧ в геном человека является ключевым ограничением в применении этой платформы, поэтому использование методов, расширяющих упаковочную ёмкость аденоассоциированного вируса, стало популярным направлением исследований в период с 2005 по 2020 год.

Подходы к преодолению ограничения упаковочной ёмкости аденоассоциированного вируса можно разделить на следующие направления: генетическая перегрузка экспрессионной кассеты до 6 тысяч нуклеотидов [47], минимизация кодирующей последовательности доставляемого трансгена до 4 тысяч нуклеотидов [48], оптимизация регуляторных элементов кассеты [49], использование многовекторных систем для раздельной доставки трансгена [50]. Для каждого из подходов за время исследований была сформирована собственная технологическая ниша, однако наибольшее количество работ посвящено раздельной доставке генетического материала с помощью нескольких вирусов. После раздельной доставки трансгена необходимо восстановить его полную последовательность, чтобы получить функциональную молекулу на уровне белка. Варианты решения данной задачи представлены на Рисунке 15.

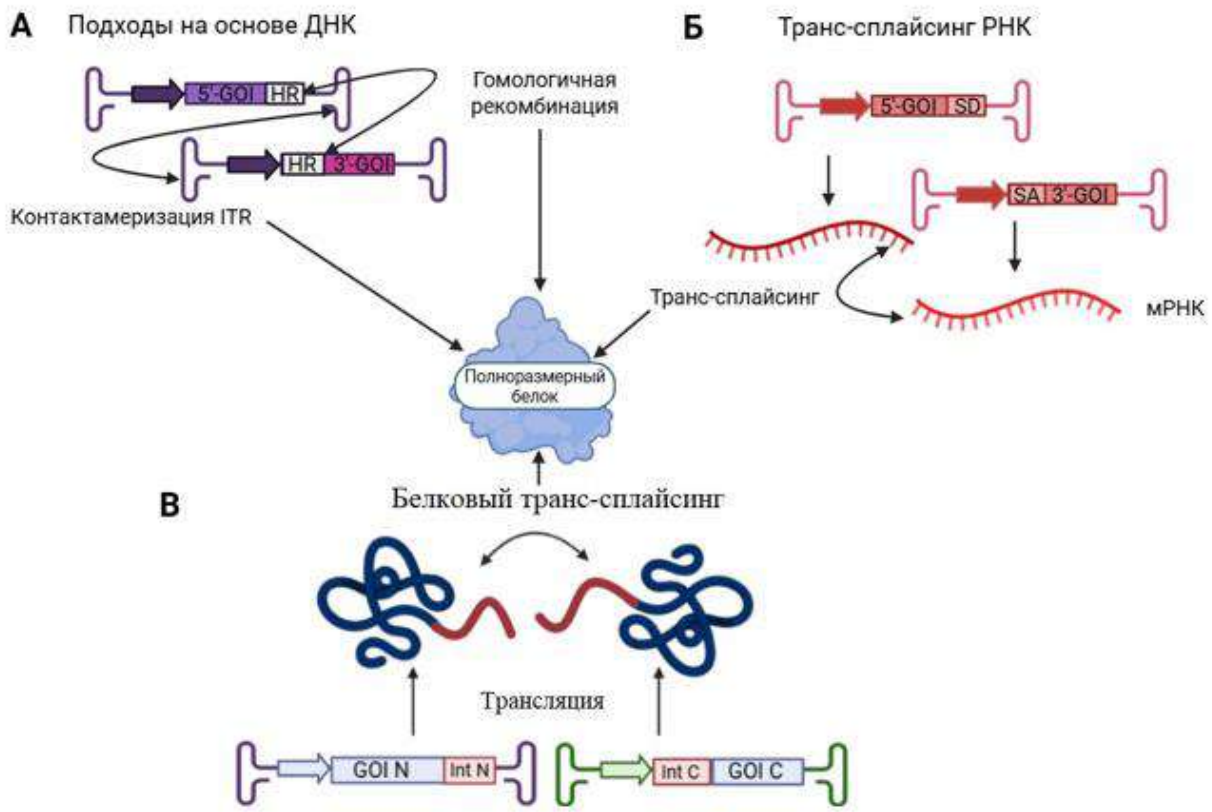


Рисунок 15 — Подходы к восстановлению полноразмерного белка после раздельной доставки фрагментов трансгена: А — подходы на уровне ДНК, основанные на конкатемеризации и гомологичной рекомбинации; Б — транс-сплайсинг РНК; В — интеин-опосредованный белковый транс-сплайсинг

Теоретически решить данную задачу можно на уровне конкатенации ДНК, мРНК-сплайсинга и белкового сплайсинга с помощью интеинов [51]. По результатам экспериментов наибольшей результативностью обладает белковый сплайсинг, поскольку синтез белков происходит быстро в ядре и цитоплазме, синтезируемые молекулы обладают высокой аффинностью и быстро взаимодействуют благодаря свойствам интеинов, а реакция полностью автономна и не требует присутствия кофакторов и катализаторов.

Таким образом, перечисленные факторы в совокупности с положительными результатами исследований *in vitro и in vivo* сделали интеины популярным инструментом заместительной генной терапии наследственных заболеваний. В Таблице 3 приведён список основных прототипов препаратов, для которых использовалась технология белкового транс-сплайсинга на основе интеинов.

Таблица 3 — Препараты для заместительной терапии наследственных заболеваний на основе интеинов

Ген мишень	Размер кДНК	Целевой орган	Интеин	Серотип AAV	*In-vitro* тесты	*In-vivo* тесты	Источник	Год
F8	7054	Клетки печени	SspDnaB	-	HEK293 COS7	-	[51]	2010
F8	7054	Клетки печени	SspDnaB	-	-	C57BL/6	[52]	2013
F8	7054	Клетки печени	NpuDnaE	AAV8	HEK293	HemA 129SF8tm1Kaz	[53]	2022
ATP7B	4398	Клетки печени	NpuDnaE	AAV8	HepG2	C57BL/6 KO	[54]	2022
DMD	6337	Мышечная ткань,фибробласты	SspDnaB Rma DnaB	AAV1	HEK293	mdx	[55]	2008
DMD	8525	Мышечная ткань,фибробласты	DnaE, gp-41, IMPDH, Nrdj	AAV9 MYO	HEK293	C57BL/6,mdx4cv	[56]	2024
CEP290	7437	Клетки фоторецепторы	NpuDnaE	AAV8	HEK293	C57BL/6;	[57]	2019
ABCA4	6819	Клетки фоторецепторы	NpuDnaE	AAV8	HEK293	C57BL/6;	[57]	2019
ABCA4	6819	Клетки фоторецепторы	SspDnaB NpuDnaE	AAV8	HEK293	C57BL/6 129S-Abca4tm1Ght/J	[58]	2023
OTOF	7125	клетки внутреннего уха	NpuDnaE Rma DnaB	AAV9 PHP.eB	HEK293	C57BL/6 KO Otof	[59]	2022

Основной вывод, который можно сделать, анализируя Таблицу 3, заключается в высокой востребованности технологии белкового сплайсинга для сборки таких крупных белков, как восьмой фактор свертывания крови, дистрофин в разных изоформах и ряд белков, отвечающих за сенсорные функции организма.

Одним из наиболее востребованных направлений является терапия гемофилии А из-за высокой частоты её встречаемости — 1:5000 случаев, отсутствия альтернативных способов доставки полного транскрипта и относительно небольшой вводимой дозы [51-53]. Дистрофия Дюшена до сих пор является самым сложным с точки зрения терапии синдромом, поскольку в своей полной форме транскрипт составляет 14 тысяч нуклеотидов, что увеличивает количество вариантов патогенных мутаций до нескольких десятков и требует введения высоких доз [55-56].

Наследственные заболевания органов слуха и зрения (синдром Штаргардта, амавроз Лебера, аутосомно-рецессивная глухота девятого типа) являются удобными клиническими моделями для концептуальной проверки препаратов передовой терапии, поскольку эти органы изолированы от общей иммунной системы, их клетки высоко дифференцированы и медленно обновляются, а размеры органов относительно невелики, что позволяет использовать небольшие дозы препаратов для локального введения [57-59].

Отметим также, что информация о половине из перечисленных препаратов появилась в течение последних пяти лет. Повышение эффективности и подбор оптимальных условий для корректного разделения белка и восстановления его функциональной копии с привлечением отдельного комплекса методов вычислительной биологии составляют значительную часть ранних этапов разработки генотерапевтических препаратов с использованием интеинов [56].

Для каждого разделяемого белка необходимо *de novo* определить сайты разделения для сплайсинга, которые являются уникальными для каждого интеина в отдельно взятой ферментативной группе.

Далее мы рассмотрим одно усовершенствование, которое сократило время разработки препаратов, многократно увеличило число потенциальных мишеней и принципиально изменило подход к терапии наследственных заболеваний.

1.4.3.2. Интеины для редактирующей генной терапии

В 2012 и 2020 годах произошли два важных события в истории современной молекулярной биологии, принципиально изменивших подход к терапии наследственных заболеваний. В 2012 году группой Charpentier была открыта система редактирования генома на основе бактериального белка CRISPR/Cas9 [60]. На протяжении долгого времени технологию редактирования использовали в основном для вырезания генов при получении нокаутных клеточных и животных моделей. В 2020 году за данное открытие была присуждена Нобелевская премия по химии, однако за год до этого в лаборатории Liu была предложена модификация, которую назвали прайм-редактированием [61].

Принципиальным отличием прайм-редакторов стало добавление обратной транскриптазы к белковому комплексу модифицированного Cas9 и расширение гидовой РНК за счёт участка связывания праймера и последовательности для редактирования гена. Так, стандартный белковый комплекс CRISPR/Cas9 увеличился в размерах с минимальных 4,2 до 5,5 тысяч нуклеотидов, однако при этом многократно увеличилась точность наведения редактора, появилась возможность осуществлять замены или инсерции генов и вносить одноцепочечный разрыв ДНК вместо двухцепочечного, что существенно уменьшает риск нецелевых хромосомных перестроек. Такое существенное улучшение точности и безопасности системы редактирования привело к значительному снижению эффективности редактирования, которая в среднем составляет 2-10%. Над увеличением эффективности работают исследователи по всему миру, и на текущий момент известно уже о 7-м поколении прайм-редакторов, способных проводить редактирование *in vitro* с эффективностью 80-95% [62].

Тем не менее, несмотря на постепенное решение проблемы эффективности редактирования, задача доставки комплекса в живые системы с каждой модификацией усложняется, поскольку размер увеличивается пропорционально количеству вносимых в белок изменений. Первоначально задачу доставки решали с помощью физических методов переноса плазмид (микроинъекции, электропорация, соникация, лазерное облучение), поскольку данные подходы не ограничивают размер доставляемой конструкции, однако требуют соответствующего технического оснащения и высокой квалификации специалиста, способного выполнить данную процедуру [63].

Альтернативные методы используют биологические платформы доставки в виде вирусных векторов (аденовирус, аденоассоциированный вирус, лентивирус), а также химические системы на основе полимеров и липидных наночастиц.

Развитие химических технологий позволило существенно упростить процесс получения наночастиц для доставки мРНК без существенных ограничений на размеры плазмиды, однако стабильность и эффективность работы данных решений пока требуют дополнительных исследований. Наиболее проверенными и стабильными системами доставки генетического материала по-прежнему остаются вирусные векторы, однако и они имеют ряд существенных ограничений по вместимости (в случае аденоассоциированного вируса) и биологической безопасности (в случае аденовируса и лентивируса).

Поскольку платформа аденоассоциированного вируса зарекомендовала себя в качестве эффективного и безопасного инструмента доставки больших генов, то разработка системы раздельной доставки комплекса прайм-редакторов была лишь вопросом времени. Технологическая задача сборки прайм-редакторов с помощью интеина сводится к нахождению сайта сплайсинга, оптимального для интеина заданной ферментативной группы. Общая схема разделения прайм-редактора, его сборки и функциональных компонентов представлена на Рисунке 16.

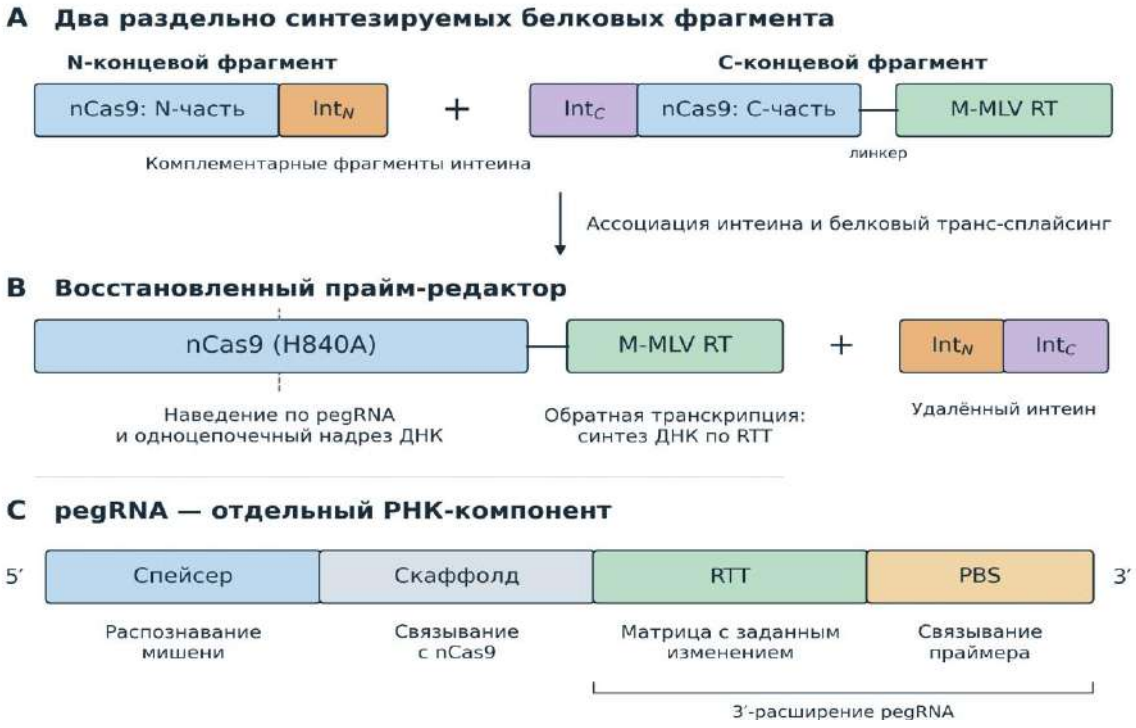


Рисунок 16 — Схема раздельной экспрессии иинтеин-опосредованной сборки прайм-редактора:

A — два белковых фрагмента с комплементарными частями интеина;

B — восстановленный прайм-редактор, включающий никазу nCas9 (H840A) и модифицированную обратную транскриптазу M-MLV RT;

C — функциональные участки pegRNA. RTT — матрица обратной транскрипции, PBS — участок связывания праймера.

Отметим, что при фиксированном выборе интеина сборка комплекса прайм-редактора происходит по одной схеме и может быть использована с разными гидовыми РНК. Это преимущество позволило в крайне сжатые сроки создать несколько препаратов для терапии частых и редких заболеваний, частичный список которых представлен в Таблице 4.

Можно сделать вывод о существенном расширении спектра заболеваний, для терапии которых могут использоваться генетические редакторы. Основное внимание исследователей сместилось на нейродегенеративные заболевания мозга, при этом ограничение упаковочной ёмкости перестало быть ключевым для трансгена, но стало лимитирующим для редактора. Стоит также отметить сложность терапевтической технологии, поскольку генетические конструкции, кодирующие редактор, необходимо доставить в клетки, локализовать белки в ядре, провести реакцию транс-сплайсинга для сборки комплекса, направить редактор с помощью гидовых РНК и далее произвести анализ результатов редактирования с помощью полногеномного секвенирования. По стоимости и сложности эта технология значительно превосходит заместительную терапию, однако позволяет однократно произвести редактирование, которое зафиксируется в клетках на протяжении всей жизни редактируемого организма. Открытыми остаются вопросы эффективности редактирования и нецелевых эффектов, их решение требует дальнейших исследований. В следующем разделе рассмотрим разработку генотерапевтических препаратов на основе двух векторных систем.

Таблица 4. Препараты для редактирующей терапии наследственных заболеваний на основе интеинов

Ген мишень	Размер кДНК	Целевой орган	Редактор	Интеин	Серотип AAV	*In-vitro* тесты	*In-vivo* тесты	Источник	Год публикаци
Pcsk9	3640	Клетки печени	ABE	NpuDnaE Rma DnaB	AAV8	HEK293T, HeLa	C57BL/6	[64]	2020
Npc1	3909	Клетки мозга	CBE	NpuDnaE CfaDnaE	AAV9	HEK293T	C57BL/6 KO Npc1tm	[65]	2020
Dnmt1	4896	Клетки фоторецепторы и глиальные	PE3	Rma DnaB	AAV1 AAV8	HEK293T, HeLa, ARPE-19	CD1	[66]	2021
Pcsk9	3640	Клетки печени	PE2, PE3	Npu DnaE	AAV8	HEK293T	C57BL/6	[67]	2022
Rpe65	1600	Эпителиальные клетки сетчатки глаза	PE2, PE3	NpuDnaE Rma DnaB	AAV8	HEK293-rd12	C57BL/6 KO rd12	[68]	2022
Dnmt1	4896	Клетки мозга	PE2, PE4	Npu DnaE	AAV9	HEK293T, Neuro-2A	C57BL/6	[69]	2023
ABCA3 RP1 PR11	5112 483 705	Клетки эмбриональной почки	CBE	Npu DnaE	AAV2	HEK293T	-	[70]	2025

1.5. Разработка препаратов генной терапии на основе двухвекторных систем

1.5.1. Обзор актуальных мишеней, используемых для терапии с помощью AAV

В таблицах 3 и 4 перечислены уже разработанные прототипы препаратов на основе интеинов для заместительной и редактирующей терапии с помощью прайм-редакторов. Ключевым достоинством применения прайм-редакторов является возможность быстрой настройки системы для однократного применения редактора оснований или более протяженных фрагментов. Подбор гидовых РНК не ограничивает исследователя в размере доставляемой конструкции, поскольку происходит целевое редактирование коротких фрагментов генов.

Основное ограничение по размеру касается необходимости внесения в сам комплекс CRISPR/Cas модификаций, которые способны увеличивать эффективность редактирования в несколько раз [63]. Тем не менее вопрос биологической безопасности применения технологии редактирования, обусловленный возможными нецелевыми эффектами, остаётся нерешённым. Число и разнообразие разрабатываемых препаратов на основе данного подхода постоянно увеличиваются с 2020 года, а одобрение FDA препарата Casgevy позволит увеличить вероятность регистрации препаратов в будущем [70].

По количеству зарегистрированных препаратов лидером в области заместительной генной терапии по-прежнему остаётся платформа на основе аденоассоциированных векторов.

На текущий момент FDA одобрено шесть препаратов на основе аденofассоциированных вирусов для терапии спинально-мышечной атрофии (Solgensma), мышечной дистрофии (Elevidys), дистрофии сетчатки (Luxturna), гемофилии А (Roctavian) и В (Hemgenix) [71]. Из перечисленных препаратов Luxturna выделяется тем, что использует доставку небольших доз вирусных частиц в глаз, а остальные препараты требуют расчёта дозы пропорционально массе пациента. Помимо малых доз, офтальмологические препараты имеют ряд особенностей, которые подробнее рассмотрим в следующем разделе.

1.5.2. Офтальмологические препараты как модель для терапии наследственных заболеваний

Основное преимущество использования глаза как мишени для генной терапии заключается в его относительно небольшом размере, что делает достаточным для трансдукции всего $1*10^8$ – $1*10^9$ VG на каждый глаз [72]. Также глаз является иммунопривилегированным органом, что позволяет многократно уменьшить ответ со стороны иммунных клеток крови, отделенных от тканей глаза гематоофтальмологическим барьером. Тем не менее у глаза присутствует система локального иммунного ответа, которая сосредоточена в основном на внешних тканях глаза, таких как роговица, конъюнктива и слёзная железа [73].

Внутренние отделы глаза содержат множество слоев высокодифференцированных клеток, которые выполняют основную сенсорную функцию, связанную с обработкой светового сигнала и практически не содержат клеток, выполняющих защитную или барьерную функцию. Вторым благоприятным для генной терапии свойством клеток сетчатки является их окончательная дифференцировка, которая делает клетки неспособными к делению и обновлению [74].

Таким образом, трансдукция данных клеток обеспечивает максимально длительный эффект, поскольку влияние факторов иммунной системы минимизировано, а обновление клеток либо невозможно, либо происходит крайне редко. Поэтому для проверки эффективности генной терапии *in vitro* была выбрана именно одна из форм наследственно обусловленной слепоты. Среди перечисленных вариантов офтальмологических мишеней белок ABCA4 является наиболее удобной мишенью для разработки двухвекторных систем.

1.5.3. Обзор мишени ABCA4 для заместительной генной терапии

ABCA4 — ген, кодирующий одноименный белок, который является членом семейства транспортеров АТФ-связывающих кассет (ABC). Белок ABCA4 функционирует как флиппаза, играющая важнейшую роль в поддержании нормальной работы системы фототрансдукции. Его основная функция — перенос различных молекулярных метаболитов, образующихся в процессе зрительного цикла, в ходе транспорта из просвета наружных сегментов фоторецепторов в цитоплазму [75].

Этот процесс жизненно важен для правильной утилизации ретинальдегида – побочного продукта изомеризации ретиналя в фоторецепторах под воздействием света.

Перемещая эти молекулы, ABCA4 помогает предотвратить накопление потенциально вредных веществ, таких как производные ретиноидов, токсичных для фоторецепторов и при высоких концентрациях способных привести к дегенеративным заболеваниям сетчатки. Активность ABCA4 необходима для поддержания уровня токсичных метаболитов в норме. Схематично роль белка представлена на Рисунке 17.

ABCA4 состоит из нескольких структурных компонентов, которые обеспечивают его функционирование в качестве белка-транспортера. Это крупный белок, состоящий примерно из 2,3 тысяч аминокислот. ABCA4 содержит два больших трансмембранных домена, каждый из которых состоит из 6 трансмембранных спиралей, встроенных в мембраны дисков наружного сегмента фоторецепторных клеток [76]. Эти трансмембранные области образуют канал, обеспечивающий транспорт молекул через мембрану. ABCA4 имеет два нуклеотидсвязывающих домена (NBD), расположенных на его цитоплазматической стороне, которые отвечают за связывание и гидролиз АТФ. Линкерные области соединяют трансмембранный и АТФ-связывающий (TMD, NBD) домены, способствуя конформационным изменениям, необходимым для осуществления транспорта. Также между спиралями имеются внеклеточные и цитоплазматические петли, которые участвуют в связывании субстратов и взаимодействии с другими клеточными компонентами. Общая структура ABCA4 позволяет ему функционировать в качестве флиппазы, активно транспортирующей такие молекулы, как производные ретинальдегида, из просвета наружного сегмента фоторецепторных клеток в цитоплазму, что критически важно для правильной работы сетчатки и поддержания зрительного цикла.

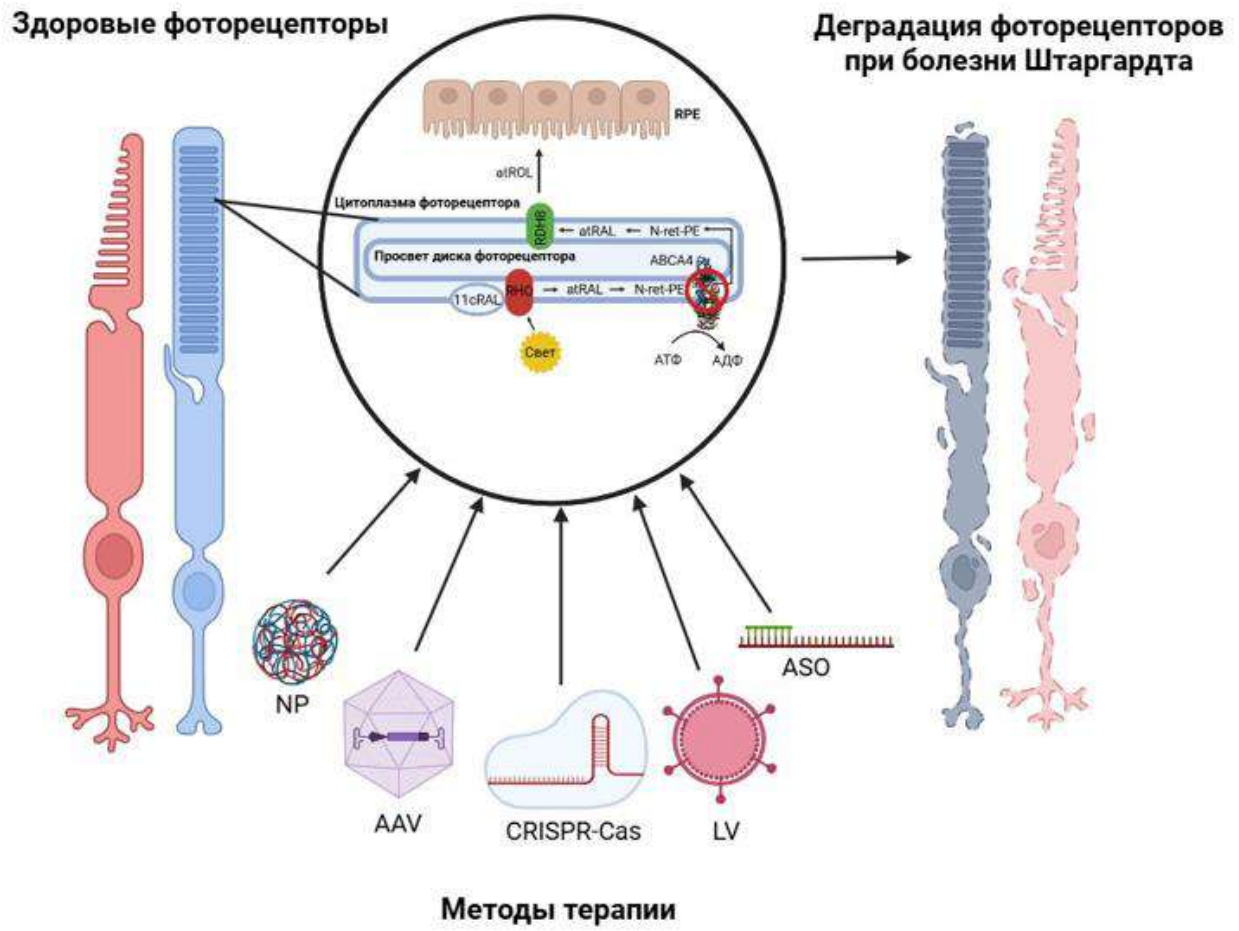


Рисунок 17 — Схематическое представление роли ABCA4 в метаболизме ретиноидов и подходов к терапии болезни Штаргардта. Слева показаны здоровые фоторецепторы, справа — их дегенерация при заболевании; в центре — схема транспорта и превращений ретиноидов в наружном сегменте фоторецептора. Внизу обозначены терапевтические подходы с использованием NP, AAV, CRISPR-Cas, LV и ASO.

При болезни Штаргардта мутации в гене ABCA4 нарушают нормальное функционирование клеток пигментного эпителия сетчатки. Эти изменения вызывают сбой в транспорте транс-ретиналя, что приводит к его накоплению в фоторецепторных клетках. Это накопление приводит к образованию токсичных бисретиноидов, таких как A2E, которые накапливаются в клетках пигментного эпителия сетчатки в виде липофусцина. Чрезмерное накопление липофусцина нарушает нормальную работу как фоторецепторов, так и клеток пигментного эпителия.

Кроме того, присутствие A2E и других бисретиноидов делает клетки пигментного эпителия более уязвимыми к окислительному стрессу, поскольку гранулы липофусцина

генерируют реактивные формы кислорода под воздействием света, вызывая окислительное повреждение и провоцируя апоптоз [78] или ферроптоз [79].

Также ABCA4 участвует в рециркуляции ретинальдегида в пигментном эпителии сетчатки. Ретинальдегид высвобождается при протеолизе родопсина в эндолизосомах эпителия после фагоцитоза внешних сегментов фоторецепторов. Таким образом, несмотря на то что уровень экспрессии ABCA4 в пигментном эпителии составляет всего 1% от уровня экспрессии в фоторецепторах, дефицит этого белка может играть роль в патогенезе болезни Штаргардта [80].

1.5.4. Эпидемиология болезни Штаргардта

STGD1 — самая распространенная причина наследственной потери зрения с вероятностью проявления 1-5 на 10000 новорожденных [81]. Заболевание чаще всего манифестирует в возрасте 6-12 лет и по мере прогрессирования нередко приводит к полной потере зрения. Однако существуют формы с поздним проявлением симптомов заболевания [82]. В нескольких исследованиях, проведённых среди пациентов с клинически подтверждённым диагнозом болезни Штаргардта, были выявлены мутации гена ABCA4, имеющие значительное распространение в российской популяции. Водномизних, проведённом на выборке из 51 пациента из 10 этнических групп, было установлено, что 70,5% случаев ретинопатий были связаны с мутациями в этом гене. Это подтверждает важность ABCA4 как ключевого гена, ассоциированного с болезнью Штаргардта и в российской популяции. Более того, 8 из 10 новых мутаций, которые были обнаружены в рамках этого исследования, характерны для неславянских популяций, что подчёркивает необходимость изучения специфичных для региона мутаций [83,84].

Точное число людей с болезнью Штаргардта 1 типа (STGD1) в России неизвестно из-за редкости заболевания и отсутствия единого реестра пациентов, но ориентировочную оценку можно сделать на основе мировых эпидемиологических данных. STGD1 — это аутосомно-рецессивное заболевание, встречающееся с частотой примерно 1 случай на 8 000–10 000 человек. С учётом численности населения России около 146 миллионов человек в 2025 году общее количество носителей мутации может достигать 14 600 человек, а расчетное число больных (включая гомозигот и компаунд-гетерозигот) — от 3 500 до 5 000. Точная генетическая идентификация носителей мутации осложняется постановкой некорректного диагноза в ходе клинической диагностики по фенотипическим признакам.

Фактическое количество больных STGD1 в России, несомненно, превышает официальную статистику, а для строгого эпидемиологического описания целесообразно формирование единого национального реестра наследственных заболеваний сетчатки.

Исследования показывают, что в российской популяции существует широкий спектр мутаций, ассоциированных с болезнью Штаргардта, а заболевание характеризуется фенотипической вариабельностью. Например, среди 38 пациентов с клинически подтверждённым диагнозом STGD1 были выявлены 12 мутаций, которые ранее не описывались в международной литературе. Эти мутации могут быть специфичны для российской популяции, что подтверждает необходимость проведения более широких исследований, направленных на идентификацию новых мутаций, характерных для региона. Кроме того, фенотипическая вариабельность заболевания также велика, что делает диагностику и лечение болезни более сложными. У пациентов наблюдаются как классические формы заболевания с ранним началом, так и более поздние формы с медленным прогрессированием.

Одним из ключевых результатов исследования стало выявление двух мутаций, которые широко распространены в российской популяции. Это комплексный аллель c.[1622T>C;3113C>T] и вариант c.5882G>A. Эти мутации были обнаружены у значительного числа пациентов и ассоциируются с различными клиническими проявлениями болезни. Например, комплексный аллель c.[1622T>C;3113C>T] был выявлен у 38,5% пациентов с заболеваниями сетчатки, что подчеркивает его важность как основной мутации для этой патологии в российской популяции. В то же время мутация c.5882G>A была выявлена у 5,7% пациентов, что подтверждает её значимость в патогенезе болезни. У пациентов с такими мутациями наблюдается более позднее начало заболевания и медленное его прогрессирование, что может оказывать влияние на клиническую практику. Сочетание этих мутаций может приводить к более позднему и менее агрессивному проявлению болезни [85].

Таким образом, распространённость болезни Штаргардта в Российской Федерации имеет свои особенности, связанные с определёнными мутациями гена ABCA4, которые чаще всего встречаются в российской популяции. Наиболее распространёнными являются мутации c.[1622T>C;3113C>T] и c.5882G>A, которые играют ключевую роль в патогенезе заболевания.

1.5.5. Современное состояние терапии синдрома Штаргардта

На текущий момент не существует зарегистрированного и одобренного препарата для терапии синдрома Штаргардта, что обусловлено особенностями и сложностью доставки терапевтических молекул в ткани глаза, а также размером кодирующей последовательности гена *ABCA4*.

Основные усилия по разработке препаратов сосредоточены на фармакологических методах поддерживающей терапии, основанных на модуляции ретиноидного цикла, снижении токсичности бисретиноидов и усилении нейропротекторных свойств клеток.

Самой очевидной стратегией, применяемой в терапии на основе малых молекул, является снижение концентрации токсичных продуктов зрительного цикла, таких как A2E и липофусцин [86]. Достичь поставленной цели можно с помощью ингибиторной терапии на основе препаратов, блокирующих белок RPE65, блокирование которого замедляет регенерацию зрительного пигмента и снижает образование вредных метаболитов. Препарат Emixustat, который принадлежит к классу ингибиторов, показал свою эффективность на третьей стадии клинических испытаний, в рамках которой пациенты получали 10 мг эмиксустата ежедневно, после чего наблюдалось полное восстановление амплитуды b-волн [87].

Другие малые молекулы — антагонисты RBP4, предназначенные для снижения уровня ретинола, такие как фенретинид, A1120, BPN-14136 и LBS-008, — не показали положительного эффекта, и их исследования были завершены на стадии доклинических испытаний или второй фазы клинических испытаний [88]. Основной недостаток терапии на основе малых молекул заключается в необходимости постоянного поддержания уровня действующих веществ, которые влияют на утилизацию метаболитов, но не устраняют исходную причину их образования. Более перспективными являются современные методы терапии, способные оказывать более продолжительный эффект и воздействовать непосредственно на причину заболевания на уровне генов. Это может существенно улучшить состояние пациента в долгосрочной перспективе.

Для современной терапии повсеместно используются препараты на основе нуклеиновых кислот и различные системы их доставки в клетки эукариот. Препараты данного класса подразделяют по типу используемой системы доставки генетической информации на невирусные и вирусные.

К невирусным системам относятся pH-чувствительные наночастицы на основе аминолипидов, способные доставлять препараты РНК без ограничения длины транскрипта благодаря возможности регулировать размер получаемых частиц [89-93].

Ключевым ограничением при применении наночастиц является их невысокая стабильность и непродолжительная экспрессия трансгена из-за низкой стабильности молекул РНК.

Препарат EIAV-ABCA4, созданный на основе вируса анемии лошадей, несет полноразмерную копию трансгена *ABCA4* и использовался для субретинального введения на первой и второй стадиях клинических испытаний в дозах от $1{,}0\times10^6$ до $1{,}8\times10^6$ трансдукционных единиц на глаз. По результатам клинических испытаний с участием 22 пациентов эффективность препарата не была подтверждена с достаточной статистической достоверностью, однако выраженных побочных эффектов также не выявили [94].

На платформе аденоассоциированных вирусов удалось получить более выраженные положительные результаты с препаратом VG801. Данный препарат использует технологию разделения трансгена *ABCA4* на два фрагмента мРНК для дальнейшей сборки на уровне РНК-сплайсинга с раздельной доставкой на основе модифицированного аденоассоциированного вируса vgAAV, который обладает большей тропностью к клеткам сетчатки. Полученный препарат в 2024 году был одобрен FDA для проведения клинических испытаний. Также препарат был отобран для участия в инициативе FDA RDEA (Rare Disease Endpoint Advancement) [95].

Отдельно отметим препараты на основе антисмысловых олигонуклеотидов, которые могут доставляться как с помощью вирусных, так и невирусных систем. Ключевым отличием данного метода является использование коротких одноцепочечных последовательностей олигонуклеотидов, комплементарных смысловой цепи. В результате гибридизации смысловой и антисмысловой РНК образуется дуплекс, который блокирует синтез мутантного белка за счет корректировки сплайсинга мРНК. Наиболее значимым представителем данного класса является препарат QR-1011, который корректировал одну из наиболее распространенных мутаций ABCA4 c.5461-10T>C. При применении препарата наблюдались существенное восстановление темновой адаптации и снижение уровня липофусцина. Тем не менее применение антисмысловых олигонуклеотидов, как и малых молекул, требует регулярного введения, что является ключевым ограничением терапии [96].

За 20 лет развития биотехнологии наблюдается рост интереса к применению препаратов на основе доставки нуклеиновых кислот, которые способны эффективно воздействовать на первопричину заболевания, а не только облегчать отдельные симптомы. Технологии на основе доставки полноразмерной версии гена *ABCA4* с помощью аденоассоциированных вирусов и липидных наночастиц являются наиболее результативными и перспективными.

1.6. Основные выводы по итогам первой главы

Интеины, будучи от природы паразитическими вспомогательные белками архей, бактерий и одноклеточных эукариот, способны реализовывать реакцию белкового сплайсинга без дополнительных активаторов или катализаторов реакции. Основным механизмом, запускающим реакцию, является ацильный сдвиг, который возможен при энергетически невыгодной цис-конформации шести аминокислот на стыке экстеинов и интеинов.

Исходную информацию об оптимальном составе аминокислот, необходимых для инициации реакции, исследователи получили на примере природных белков-хозяев, содержащих интеины. Далее путём экспериментального перебора аминокислот было показано влияние каждой отдельно взятой аминокислоты экстеина на реакцию транс-сплайсинга. Роли всех аминокислот внутри структуры интеина до сих пор остаются неизвестными из-за отсутствия достаточного количества экспериментально определённых структур. По известным структурам интеинов была показана ключевая роль доменов A, B, F и G и их ближайшего окружения, выполняющих каталитические функции в процессе белкового сплайсинга.

Большая часть попыток увеличить эффективность реакции сплайсинга привела к модификации предпоследних аминокислот на С-конце С-фрагмента интеина для увеличения локальной стабильности G-домена, которая ускоряет циклизацию аспарагина, завершающую реакцию сплайсинга. Другие методы модификации, включают в себя вариацию целых частей интеинов из пары, которые способны вступать в реакцию кросс-реактивного взаимодействия при высокой гомологии основных каталитических доменов.

Комбинирование оптимизационных методов позволило создать несколько модифицированных или синтетических интеинов, которые нашли своё применение в одноступенчатой аффинной хроматографии и сборке токсичных для клетки белков и биспецифических антител. В последнее время быстро растёт интерес к применению интеинов в области заместительной и редактирующей генной терапии наследственных заболеваний. Эффективность заместительной терапии была продемонстрирована при сборке терапевтических белков *in vitro* и *in vivo* на примере разработки прототипов препаратов для лечения гемофилии А, дистрофии Дюшена, синдрома Штаргардта.

ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ

2.1. Генная инженерия для получения генетических конструкций

В разделе описаны методы получения генетических конструкций *de novo* с помощью химического синтеза олигонуклеотидов в соответствии с заданными схемами белковых комплексов для транс-сплайсинга. Рассмотрены выравнивание и анализ аминокислотных последовательностей интеинов, преобразование аминокислотной последовательности в нуклеотидную с оптимизацией кодонов, сборка генов из смеси олигонуклеотидов, получение рекомбинантных фрагментов генов с помощью перекрытий и удлинения цепи в ходе ПЦР, алгоритм выбора сайтов разделения белка, разработка схем клонирования собранных последовательностей, трансформация бактерий, скрининг положительных клонов и сайт-направленный мутагенез.

2.1.1. Аминокислотные последовательности интеинов: поиск в базах данных, выравнивание и анализ

Аминокислотные последовательности интеинов были взяты из базы данных InBase [97]. С учётом рассмотренных ранее статей основное внимание уделяли природным интеинам группы DnaE, поскольку данное семейство является самым изученным и многочисленным, а его представители обладают высокой транс-сплайсинговой активностью и эффективностью. Также высокая гомология последовательностей и эволюционное родство позволили оценить роль отдельных доменов и аминокислот с помощью выравнивания аминокислотных последовательностей в программе COBALT NCBI NIH [98]. Для визуализации различий в аминокислотных последовательностях использовали настройки отображения Frequency-Based Difference.

2.1.2. Алгоритм выбора сайтов разделения целевой молекулы для интеинов группы DnaE

Для определения оптимальных сайтов разделения молекулы целевого белка использовали алгоритм на основе данных об экстеиновых остатках из таблицы в работе [28]. Алгоритм был адаптирован для расчётов в среде Jupyter Notebook на языке Python 3.0 с помощью подпрограммы, сгенерированной средствами искусственного интеллекта DeepSeek 3.1. После разделения последовательности целевого белка к N-концевой части

добавляли N-концевой фрагмент интеина, а к С-концевой части — С-концевой фрагмент интеина в одной рамке считывания. При выборе сайта в первую очередь учитывали максимальный балл оценки экстеиновых остатков, а затем — результаты визуального анализа предсказанной структуры.

2.1.3. Получение генетических конструкций, кодирующих интеины

Для создания плазмид, содержащих последовательности, кодирующие интеины группы DnaE, использовали инструмент виртуального дизайна Benchling [99]. С помощью сервиса была проведена кодонная оптимизация аминокислотной последовательности для уменьшения GC-состава нуклеотидной последовательности до 50%. Далее к готовой ДНК-последовательности добавляли сайты рестрикции BamHI на 5’-конец и HindIII на 3’-конец. Итоговую последовательность разделяли на олигонуклеотиды по 80 пар оснований с помощью сервиса DNA Works [100] при заданной температуре отжига попарно перекрывающихся участков 60 градусов. Полный перечень праймеров, использованных в работе, приведён в приложении (Приложение А Таблица А.1). Синтез олигонуклеотидов проводили в ресурсном центре генетической инженерии на синтезаторе ДНК/РНК ASM-2000 (ООО «БИОССЕТ», Россия).

Сборку конструкций проводили по протоколу двухэтапной ПЦР с перекрытием [101] с использованием набора Tersus Plus PCR kit (PK221, Evrogen, Россия). Для прведения ПЦР в пробирку Ultraflux на 200 мкл (3220-00, SSIbio, США) с помощью наконечников Vertex на 10 (4137NS0, SSIbio, США) и 200 мкл (4237NS0, SSIbio, США) последовательно добавляли 4 мкл 5X буфера Tersus Red, 0,5 мкл смеси дезоксинуклеозидтрифосфатов 10 мМ dNTP, 0,5 мкл 50X смеси полимераз Tersus, 2 мкл 10-пикомолярной смеси олигонуклеотидов, 12 мкл деионизованной воды и 1 мкл ДНК-матрицы.

Амплификацию проводили с помощью T100 Thermal Cycler (1861096, BioRad, США) с программой амплификации: первичная денатурация — 95°C, 5:00; в цикле: денатурация — 95°C, 00:10; отжиг — 60°C, 00:20; элонгация — 72°C, 00:30 или 1:00 в зависимости от длины ПЦР-продукта. На первом этапе собирали фрагмент из олигонуклеотидов при 20 циклах амплификации, на втором этапе — амплифицировали целевой ПЦР-продукт для его выделения из смеси при 40 циклах амплификации. После получения последовательностей интеинов по аналогичной методике проводили сращивание по перекрытию с фрагментами последовательности *GFP*.

Для подтверждения сборки итоговый ПЦР-продукт наносили на нативный 1%-ный агарозный гель (агароза типа 1 с низким EEO, CAS 9012-36-6, PanReac Applichem, США) на основе водного ТАЕ-буфера (0,04 М Трис-(гидроксиметил) аминометан, 1 мМ ЭДТА,

0,02 М уксусная кислота, вода высокой степени очистки, pH 8,3) в объёме 2-5 мкл и анализировали с помощью электрофореза при напряжении 90 В в течение 20 минут с применением маркера длин фрагментов Step50 (S-8055, Biolabmix, Россия). Результат визуализировали с помощью ChemiDoc MP Imaging System (12003154, BioRad, США) и Image Lab Software 4.1 (1709690, BioRad, США). После подтверждения соответствия длины

ПЦР-продукта ожидаемой его очищали на колонке с помощью набора CleanUp Mini (BC023L, Evrogen, Россия) по протоколу производителя и далее использовали для клонирования рестриктазно-лигазным методом.

Для клонирования и дальнейшей сборки AAV на основе итоговой конструкции использовали набор для сборки AAV-5 CMV Expression System (Cell Biolabs, Inc., США) включающий три плазмиды pAAV-MCS Expression Vector (Cell Biolabs, Inc., Part No. VPK-410, США), pAAV-RC5 Vector (Cell Biolabs, Inc.,Part No. VPK-425, США), pHelper Vector (Cell Biolabs, Inc., Part No. 340202, США) [102]. Схема клонирования была разработана с помощью программного обеспечения SnapGene с использованием высокоактивных эндонуклеаз рестрикции, совместимых в одной буферной системе.

Для реакции рестрикции в пробирку на 1,5 мл (022364111, Eppendorf, США) добавляли 5 мкл 10X буфера rCutSmart (B6004S, New England Biolabs, США), 1-2 мкг очищенного ПЦР-продукта или плазмидного вектора, 0,5 мкл эндонуклеазы рестрикции BamHI-HF (R3136L, New England Biolabs, США) и рестриктазу HindIII-HF (R3104L, New England Biolabs, США); деионизованную воду добавляли до общего объёма 50 мкл. Смесь инкубировали 4 часа при 37 градусах в термостате после чего к обработанному ферментами вектору добавляли 1 мкл щелочной фосфатазы rSAP (M0371L, New England Biolabs, США) на 1 мкг пДНК и инкубировали 40 минут при 37 градусах, после чего инактивировали ферменты при 65 градусах в течение 10 минут. Итоговую смесь очищали с помощью набора CleanUp Mini по протоколу производителя.

Для лигирования в пробирку объёмом 200 мкл добавляли очищенный линеаризованный дефосфорилированный вектор pAAV-MCS в количестве 100 нг, очищенный фрагмент вставки в количестве 70 нг, что соответствует соотношению вставки и вектора 5:1 по массе, 2 мкл лигазы T4 (M0202L, New England Biolabs, США), разведённой в 10 раз в 10X буфере T4 DNA Ligase Reaction Buffer (B0202S, New England Biolabs, США), и деионизованную воду до 10 мкл на реакцию. Смесь инкубировали при температуре 16 градусов в течение ночи или в течение 2 часов при комнатной температуре в термостате. Далее 5 мкл лигазной смеси использовали для химической трансформации бактериального штамма DH5A по стандартному протоколу [103].

После получения клонов на чашке проводили скрининг на наличие вставки с помощью праймеров AAV_Seq_F (CCTCTTATCTTCCTCCCACAGCTCCTG), AAV_Seq_R (TTCCAGGGCCAGGAGAGGCA) и набора 5X ScreenMix (PK041L, Evrogen, Россия). Для этого в пробирку на 200 мкл добавляли 2 мкл 5X ScreenMix, 0,5 мкл AAV_Seq_F с концентрацией 10 пмоль/мкл, 0,5 мкл AAV_Seq_R с концентрацией 10 пмоль/мкл, 7 мкл деионизованной воды и материал бактериальной колонии. Материал той же колонии дополнительно пересевали на новую чашку для последующего выращивания культуры в 3мл жидкой LB-среды с антибиотиком и выделения плазмидной ДНК методом минипрепа.

ПЦР проводили в условиях, описанных ранее. Результат визуализировали с помощью гель-электрофореза. Положительные клоны проверяли секвенированием по Сенгеру на приборе ABI 3500 (Thermo Fisher Scientific Inc., США) и после подтверждения выращивали в 200 мл жидкой LB-среды с антибиотиком для последующего выделения плазмидной ДНК методом мидипрепа. Очистку проводили с помощью наборов Plasmid Miniprep 2.0 (BC221L, Evrogen, Россия) и Plasmid Midiprep 3.0 (BC224, Evrogen, Россия) по протоколу производителя. Очищенную плазмиду использовали для экспериментов по трансфекции адгезионных клеток и сборке AAV в суспензионных клетках.

2.2. Компьютерное моделирование для визуализации молекул

Для моделирования предсплайсинговых комплексов в AlphaFold 3 использовали аминокислотные последовательности белка GFP и интеина SspDnaE. Подготовку входных последовательностей проводили в соответствии с дизайном белковых конструкций для трёх сайтов разделения GFP: 47/48, 69/70 и 127/128. Для каждого варианта последовательность GFP разделяли на N- и C-концевые фрагменты по соответствующей границе между аминокислотными остатками.

В состав первой полипептидной цепи входили N-концевой фрагмент GFP и N-концевая часть SspDnaE, а в состав второй — C-концевая часть SspDnaE и C-концевой фрагмент GFP. Таким образом, порядок компонентов от N- к Сконцу соответствовал схемам GFPN–SspDnaEN и SspDnaEC–GFPC. Последовательности задавали в однобуквенном аминокислотном коде, сохраняя порядок остатков в соответствующих конструкциях. Такое представление описывало две отдельные полипептидные цепи до реакции сплайсинга, а не ковалентно соединённый полноразмерный GFP.

Каждую пару последовательностей использовали для отдельного задания на моделирование комплекса в AlphaFold 3. Полученные модели сопоставляли по взаимному расположению фрагментов GFP и частей интеина, а также по сближению и ориентации соединяемых концов экстеинов. Анализ носил качественный характер и был направлен на выявление геометрических различий между вариантами разделения.

2.3. Клеточные тесты для оптимизации условий трансфекции и валидации данных компьютерного моделирования и анализа

2.3.1. Поддержание клеточных линий

Для работы с адгезионными клеточными линиями человека HEK293 (85120602, ECACC), ARPE19 (ИБР РАН, Россия), FbP13 (МГМЦ им. Бочкова, Россия) использовали среду DMEM High Glucose (D5648, Sigma-Aldrich, Германия) с 10% FBS (S181H-500, Biowest, Франция), 0,25%-ный раствор трипсина с солями Хенкса (П036п, ПанЭко, Москва, Россия) и DPBS. После размораживания клетки культивировали на 25-сантиметровых матрасах с адгезивным покрытием (707003, Nest, Китай) до 2-го пассажа и далее использовали для экспериментов. Для пересева отбирали среду серологической пипеткой на 5 мл, добавляли 500 мкл DPBS и затем удаляли раствор для удаления остатков среды.

К клеткам на пластике добавляли 500 мкл трипсина и инкубировали в термостате 4 минуты. Далее добавляли 500 мкл DPBS и центрифугировали в пробирке типа Falcon на 15 мл (10-9151, Biologix, Китай). Супернатант с остатками трипсина удаляли, а к клеточному осадку добавляли 1 мл среды с сывороткой. Для подсчёта клеток смешивали 20 мкл суспензии и 20 мкл трипанового синего (15250061, Thermo Scientific, США) и наносили 20 мкл смеси на слайд (A25750, Thermo Scientific, США) для счётчика клеток Countess™ II FL (AMQAF1000, Thermo Scientific, США). После подсчёта часть клеток использовали для дальнейшего культивирования на адгезионном матрасе T25 (600 тысяч клеток для линии HEK293, 1 миллион для линий ARPE19 и FbP13), а оставшиеся клетки — для экспериментов. Клетки культивировали при 37 °C и 5% CO_2; для суспензионных культур использовали перемешивание при 170 rpm. Клетки пересевали с учётом индивидуальной скорости роста, не превышая максимальную плотность 2 млн/мл.

2.3.2. Трансфекция и трансдукция адгезионных клеток

При трансфекции и трансдукции использовали суспензию клеток, число которых предварительно определяли после культивирования. Такой вариант трансфекции называли обратной. Для трансфекции использовали PEI (764892,SigmaAldrich-Merck Life Science LLC, США), разведённый в DPBS до рабочей концентрации 1 мкг/мл, 22,75 мл среды DMEM High Glucose (D5648, Sigma-Aldrich, Германии), FBS (S181H-500, Biowest, Франция), 0,25%-ный раствор трипсина с солями Хенкса (П036п, ПанЭко,Россия) и планшеты с адгезионным покрытием (703001, 712001, 702001, 701001, Nest, Китай). В пробирку на 2 мл добавляли разведённый до концентрации 1 мкг/мл PEI, среду DMEM без сыворотки до 100 мкл и плазмидную ДНК в соответствии с заданной нагрузкой. После инкубации трансфекционной смеси добавляли 900 мкл клеточной суспензии с заданным количеством клеток, ресуспендировали в 1 мл и добавляли в лунку 6-луночного планшета. Альтернативно трансфекцию проводили по аналогичному протоколу, уменьшив количество клеток и реактивов пропорционально рабочей площади. Далее инкубировали клетки в течение 16 часов, после чего заменяли среду на среду с 5%-ным содержанием сыворотки. Трансдукцию проводили аналогично, но вместо трансфекционной смеси использовали очищенный и стерильно профильтрованный препарат вирусных частиц после хроматографии. За клетками наблюдали в течение 48 или 72 часов.

2.3.3. Динамический анализ с помощью Incucyte

Качественный и полуколичественный анализ клеток проводили с помощью системы прижизненного мониторинга Incucyte S3 (Sartorius AG, Германия), которая автоматически сканировала планшеты в трёх основных оптических каналах (Bright field, Green object, Red object) каждые 2 часа в течение 48 часов. В качестве основной метрики для анализа использовали число зелёных объектов на микрофотографии. Настройки одного сканирования включали съёмку 12 полей с регистрацией флуоресценции в зелёном канале при экспозиции 300 мс. Дополнительно выполняли вычитание фоновой флуоресценции. С помощью данного метода оценивали влияние исследуемых условий на динамику накопления сигнала, что позволяло выбрать оптимальные условия.

2.3.4. Цитометрический анализ клеточных популяций

Проточная цитометрия как независимый метод количественного анализа популяции клеток, флуоресцирующих в зелёном канале, позволяет измерить общее количество флуоресцентных клеток, определить процентное соотношение мёртвых и живых клеток, а также эффективность доставки генетического материала с помощью трансфекции и трансдукции и другие параметры. В первую очередь в данном исследовании определяли эффективность трансфекции и трансдукции по количеству флуоресцирующих FITC+ клеток для оптимизации условий. Для проведения проточной цитометрии клетки после трипсинизации ресуспендировали в FAX-буфере (DPBS с 5% FBS и 5 мМ ЭДТА). Полученные образцы вносили в лунки 96-луночного планшета с плоским дном (3300, Costar Corning, США) и далее переходили к работе с цитометром CytoFlex(A00-1-1102,Beckman Coulter, США). Перед началом работы систему промывали раствором 0,001% азида натрия (S2002, Sigma-Aldrich-Merck Life Science LLC, США), моющим средством и деионизованной водой. Далее образцы анализировали при заданных напряжениях детекторов FSC, SSC, FHS, FITC (FSC -17 В, SSC -17 В, FITC - 4 В, PE - 1 В).

2.3.5. Вестерн-блоттинг

Полуколичественный анализ исследуемых белковых прекурсоров и продуктов транс-сплайсинга проводили методом вестерн-блоттинга с использованием специфичных антител. Для подготовки клеток использовали 5 мл PBS с добавлением 10 мкл PMSF и 100 мкл коктейля ингибиторов протеаз cOmplete (11697498001, Roche, Швейцария) в объёме 500 мкл на образец. Суспензию гомогенизировали с помощью соникатора в течение 40 секунд при 40% мощности и центрифугировали 10 минут при 13 000 rcf. Далее определяли концентрацию белка с помощью NanodropOneC в режиме A280 (1 ABS/ 1 mg). С учётом измеренной концентрации в каждую пробу вносили 100 мкг белкового материала и доводили объём до 50 мкл с помощью PBS.

После определения концентраций к образцам добавляли по 15 мкл четырёхкратного буфера Лэммли (4% SDS, 10% 2-меркаптоэтанола, 20% глицерина, 0,004% бромфенолового синего, 0,125 М Tris-HCl pH 6,8) и далее пробы инкубировали в течение 10 минут при 100°C. Итоговый препарат наносили в количестве 25 мкл в одну лунку разделяющего 12,5%-ного геля (нижний разделяющий гель, верхний концентрирующий гель). Для анализа белковых молекул использовали маркер Precision Plus Protein™ Kaleidoscope™ Prestained Protein Standards (1610375, BioRad, США).

Электрофорез проводили при 150 В в течение 1 часа в однократном Running-буфере. Белки из полученного геля переносили на PVDF-мембрану (Immobilon®-P, IPVH00010, Merck Life Science LLC, США), которую предварительно активировали с помощью 80%-ного этанола инкубированием в течение 10 минут для малых и средних белков. Для больших белков (ABCA4 и другие) использовали мокрый перенос в течение 12 часов.

Далее мембрану отмывали водой и использовали для полусухого переноса с помощью Trans-Blot Turbo Transfer System (1704150, BioRad, США) и буфера (25 мМ TrisOH, 190 мМ glycine, 0,1% SDS, pH 8,3) с параметрами переноса 2 А в течение 20 минут. Далее мембрану блокировали в 5%-ном растворе сухого молока (1706404, BioRad, США) в TBS-T в течение 2 часов при перемешивании 250 rpm.

Далее мембрану инкубировали с первичными поликлональными кроличьими антителами anti-GFP (AB01101231265, Evrogen, Россия), разведёнными в 5%-ном растворе сухого молока в соотношении 1:2000, или с моноклональными мышиными антителами anti-6His-HRP (MCA5995P, BioRad, США) в течение ночи. После инкубации мембрану отмывали в TBS-T 3 раза в течение 5 минут, затем добавляли вторичные антитела Goat Anti-Rabbit IgG H&L HRP (HAF008, R&D Systems, США) к кроличьим anti-GFP и инкубировали 2 часа.

Отмытую мембрану проявляли с помощью ECL (1705062, BioRad, США) и визуализировали с помощью системы Chemidoc по каналу Colorimetric. После проявки процедуру повторяли с антителами anti-GAPDH (8G4R, Hytest, Россия) для нормирования сигнала по референсному белку.

Для визуализации ABCA4 использовали мышиные моноклональные антитела ABCA4 Antibody 3F4 (Santa Cruz, sc-65672, США) в разведении 1:1000 в 5%-ном растворе молока; инкубация длилась 16 часов. Для визуализации использовали вторичные HRP-меченые антитела Goat Anti-Mouse IgG H&L HRP (Abcam, ab97023, США) в разведении 1:5000 в 5%-ном растворе молока; инкубация длилась 2 часа.

2.3.6. Получение AAV в клетках HEK293

Для раздельной доставки генетического материала с помощью AAV и восстановления функций белка использовали набор плазмид AAV-5 Packaging System, в который входили pAAV-MCS Expression Vector (Cell Biolabs, Inc., Part No. VPK-410, США), pAAV-RC5 Vector (Cell Biolabs, Inc., Part No. VPK-425, США), pHelper Vector (Cell Biolabs, Inc., Part No. 340202, США).

Протокол трансфекции суспензионных клеток ранее был описан в работе [110] и применён для наработки в 600 мл суспензии. Для этого в литровые колбы Эрленмейера

(784211, Nest, Китай) добавляли 300 мл суспензии клеток HEK293, культивируемых в среде BalanCD HEK293 (91165, FUJIFILM, Китай), с концентрацией 1 млн/мл. Для транзиентной трёхплазмидной трансфекции в пробирки на 50 мл (0030122178, Eppendorf, США) добавляли 2,25 мл PEI (764892, Sigma-Aldrich-Merck LifeScience LLC,США), разведённого в DPBS до концентрации 1мкг/мл, и 22,75 мл среды DMEM/F12 без глутамина и сыворотки (С470п, ПанЭко, Москва, Россия) и инкубировали 5 минут для полного растворения после перемешивания. Во вторую пробирку добавляли плазмиды в суммарном количестве 1,5 мкг на 1 млн клеток в молярном соотношении pAAV_MCS-Helper-RepCap 2/5 - 2:2:5 и доводили средой до 25 мл. Суспензию с плазмидами добавляли по капле к разведённому в среде PEI через шприцевой PES-фильтр с размером пор 0,22 мкм (431224, Corning, США) и инкубировали полученную смесь 12 минут. Полученную смесь добавляли по капле с помощью серологических пипеток (91025, SPL LIFESCIENCES, Китай) в колбы с суспензией клеток по 25 мл. Через 24 часа к суспензии добавляли BalanCD HEK293 Feed (91166, FUJIFILM, Китай) и культивировали ещё 96 часов в шейкере-инкубаторе (Multitron Standard, Infors, Швейцария) при 180 rpm, 37°C и 80% влажности.

Для выделения вирусных частиц к суспензии в колбах добавляли 10% Tween 20 (P1379, Sigma-Aldrich-Merck Life Science LLC, США) до конечной концентрации 0,05% и инкубировали 1 час в условиях культивирования. Далее к суспензии добавляли 1 М хлорид магния (M9272, Sigma-Aldrich-Merck Life Science LLC, США) до концентрации 1 мМ и бензоназу с концентрацией 5 млн ед/мл (101675, Millipore-Merck Life Science LLC, США) до конечной концентрации 30 ед/мл и инкубировали 1 час в условиях культивирования. Полученную суспензию разливали в стаканы для центрифугирования объёмом 400 мл (5910707-002, Eppendorf, США) и центрифугировали 10 минут при 3500 rpm для осаждения клеточного дебриса. Супернатант смешивали с диатомитом Celite® 503 (22151, Sigma-Aldrich-Merck Life Science LLC, США) из расчёта 1 г диатомита на 100 мл супернатанта и далее фильтровали через систему стерильной фильтрации с размером пор 0,22 мкм (344001, Nest, Китай). Фильтрат концентрировали с помощью системы тангенциальной фильтрации Vivaflow 200 TFF Cassette PES с фильтром на 100 кДа (VF20P4, Sartorius AG, Германия) и перистатического насоса Masterflex L/S с головкой Easy-Load (77921-65, Masterflex, Германия) со скоростью 100 мл/мин. Перед началом работ систему промывали водой MilliQ и PBS (02-0869-00, Millipore-Merck Life Science LLC, США) в объёме 200 мл и далее производили концентрирование образцов до объёма 50 мл. По завершении работ систему промывали 0,1 М раствором NaOH в объёме 200 мл и далее отмывали водой MilliQ в объёме 1 литра.

Для аффинной хроматографии полученный препарат вирусных частиц фильтровали через шприцевой PES-фильтр с размером пор 0,22 мкм и далее использовали колонку HiTrap AVB Sepharose HP (28411211, Cytiva, Швеция) с сорбентом AVB Sepharose High Performance (28411210, Cytiva, США) и систему хроматографии среднего давления NGC Discover 10 (7880009, BioRad, США). Хроматографическую очистку AAV5 проводили по протоколу производителя [111] с использованием следующих буферов: уравновешивающего (20 мМ Tris, 200 мМ NaCl, pH 7,8), промывочного (20 мМ Tris, 200 мМ NaCl, pH 7,8), элюирующего (20 мМ Tris, 200 мМ NaCl, pH 7,8), регенерирующего (0,1 М citrate, pH 2,1), очищающего (10 мМ NaOH, pH 2,1). После хроматографии корректировали pH элюата с помощью 1 М TrisOH.

Для оценки эффективности хроматографии и определения конечной концентрации вирусных частиц использовали метод ПЦР в режиме реального времени по протоколу и зстатьи [112]. Перед началом анализа проводили пробоподготовку, включавшую последовательную обработку препарата вирусных частиц ДНКазой I и протеиназой К. Для этого в пробирку на 200 мкл добавляли 2 мкл десятикратного буфера для ДНКазы (100 мМ Tris-HCl (pH 7), 30 мМ MgCl2, 30 мМ CaCl2), 1 мкл ДНКазы I (EM-1250, Биолабмикс, Россия), 2 мкл препарата AAV и 15 мкл воды до общего объёма 20 мкл. Смесь инкубировали 30 минут при 37 градусах и инактивировали 10 минут при 55 градусах в амплификаторе T100. Далее к смеси добавляли 20 мкл протеиназы К (EP-1200, Биолабмикс, Россия), разведённой до концентрации 0,1 ед/мкл. Смесь инкубировали 55 минут при 55 градусах и инактивировали 5 минут при 95 градусах в амплификаторе T100.

Выделенный препарат геномной векторной ДНК амплифицировали с помощью праймеров AAV2ITR_F (5′-GGAACCCCTAGTGATGGAGTT-3′), AAV2ITR_R(5′-CGGCCTCAGTGAGCGA-3′) и зонда (5′-FAM-CACTCCCTCTCTGCGCGCTCG-BHQ1-3′) из статьи [112]. Для построения стандартной кривой использовали линеаризованную форму плазмидного вектора pAAV-MCS Expression Vector с концентрацией 10 нг/мкл, что соответствовало наибольшей концентрации в 10^9 копий гена/мл, и серию последовательных 10-кратных разведений до 1 пг/мкл, что соответствовало наименьшей концентрации 1E4 копий гена/мл. Для проведения ПЦР в пробирку на 1,5 мл добавляли компоненты смеси из расчёта на 110 реакций: 800 мкл двукратной готовой смеси БиоМастер HS-qPCR Hi-ROX, 110 мкл деионизованной воды, по 5 мкл прямого праймера, обратного праймера и зонда с концентрацией 100 пмоль/мкл. Далее готовую смесь перемешивали и распределяли по 10 мкл на дно каждой ячейки 96-луночного планшета (3401-40, SSIbio, США) после чего добавляли по 5 мкл препарата вирусного генома на стенку индивидуальной ячейки. Далее планшет заклеивали оптически прозрачной плёнкой UltraFlux® Standard PCR Film (3622-

00, SSIbio, США) и центрифугировали 1 минуту при 200 rcf. Анализ проводили с помощью системы ПЦР в режиме реального времени StepOnePlus (437659, Thermo Scientific, США) с применением режима измерений Quantification-Standard Curve в приложении StepOneSoftWare V2.3 software. Критериями приемлемости для анализа являлись линейность стандартной кривой более 0,9, эффективность реакции в диапазоне от 85 до 100%, стандартное отклонение порогового цикла для одного образца не более 0,2, соответствие значений концентрации образца линейному диапазону определяемых концентраций. Полученные данные использовали для трансдукции клеток в условиях *in vitro u in vivo* в выбранной дозе.

2.3.7. Получение липидных наночастиц

Липидные наночастицы (LNP), содержащие мРНК *GFP*, получали с использованием композиции LNP-5 на основе двух ионизируемых липидов, описанной Kunyk и соавт. [104]. Опубликованная композиция была разработана для доставки мРНК репортёрных белков и самоамплифицирующейся РНК; в настоящем исследовании в качестве доставляемого компонента использовали мРНК, кодирующую GFP. Липидная фаза содержала эквимолярную смесь ионизируемых липидов ALC-0315 и SM-102, а также 1,2-дистеароил-snглицеро-3-фосфохолин (DSPC), холестерин и 1,2-димиристоил-rac-глицеро3-метоксиполиэтиленгликоль-2000 (DMG-PEG-2000). Суммарная доля двух ионизируемых липидов составляла 50 мол.%. Молярное соотношение суммы ионизируемых липидов, DSPC, DMG-PEG-2000 и холестерина составляло 50:10:1,5:38,5 соответственно. Таким образом, доля каждого из липидов ALC-0315 и SM-102 составляла 25% от общего количества липидов.

Липидные компоненты растворяли в этаноле, а мРНК *GFP* разводили в ацетатном буфере с pH 4,0. Количество липидов рассчитывали таким образом, чтобы обеспечить отношение азота к фосфату (N/P), равное 9,4. Этанольную липидную фазу и водную фазу, содержащую мРНК, объединяли путём быстрого микрофлюидного смешивания, обеспечивающего спонтанную самосборку наночастиц с включённой РНК. После формирования частиц дисперсию разбавляли и проводили замену буфера на фосфатно-солевой буфер (PBS) с pH 7,4 для удаления этанола и неинкапсулированных низкомолекулярных компонентов. Полученный препарат пропускали через стерильный фильтр из полиэфирсульфона (PES) с размером пор 0,22 мкм и хранили при 4 °C до использования.

Гидродинамический диаметр частиц и индекс полидисперсности определяли методом динамического светорассеяния. Концентрацию РНК и эффективность

инкапсуляции определяли с помощью флуоресцентного анализа с РНКсвязывающим реагентом, сравнивая результаты анализа образцов в отсутствие и в присутствии детергента. Препараты считали пригодными для введения, если средний диаметр частиц находился в диапазоне 80–120 нм, индекс полидисперсности был ниже 0,2, а эффективность инкапсуляции РНК превышала 80%, что соответствовало характеристикам оптимизированной композиции LNP-5, приведённым в публикации [104]. Перед введением в глаз суспензию LNP разводили стерильным PBS таким образом, чтобы вводимый объём 1 мкл содержал 100 нг мРНК *GFP*.

2.4. Работа с лабораторными животными

В данном разделе описаны методы работы со здоровыми лабораторными мышами дикого типа линии C57BL/6, использованными для проверки вирусных векторных систем доставки модельных и целевых белков в ткани сетчатки глаза. Рассмотрены условия содержания лабораторных животных, процедура интравитреальной инъекции, извлечение и подготовка органов к иммуногистохимическому окрашиванию, а также анализ полученных изображений.

2.4.1. Содержание лабораторных животных

Для экспериментов *in vivo* использовали 50 самок лабораторных мышей линии C57BL/6 (ИЦИГ СО РАН, Новосибирск, Россия) без нарушений в развитии в возрасте 8-12 недель массой 20-25 г. Животных содержали в стандартных условиях вивария при температуре 22±2°C и относительной влажности 50±10% с 12-часовым световым циклом (свет с 8:00 до 20:00). Мышей размещали группами по 5 особей в пластиковых клетках размером 30×20×15 см со стерильной древесной подстилкой, которую заменяли каждые 3 дня.

Экспериментальные группы формировали из минимального количества животных на одну группу в количестве трех биологических повторов. Количетство групп было пропорционально количеству используемых препаратов и одной группы отрицательного контроля с введением PBS.

Животные имели неограниченный доступ к стандартному гранулированному корму и фильтрованной воде. Вентиляция помещения обеспечивала 12-15 воздухообменов в час.

Все манипуляции с животными проводили в дневное время с 9:00 до 15:00 для минимизации стрессового воздействия.

2.4.2. Проведение процедуры внутриглазной инъекции

Перед введением препарата AAV животное подготавливали следующим образом. После фиксации мыши в руке на роговицу глаза наносили смесь растворов атропина (0,5%) и фенилэфрина (2,5%) пипеткой Пастера. После нанесения капли раствор оставляли на роговице в течение одной минуты. Процедуру повторяли до явного увеличения внутриглазного давления и раскрытия зрачка. После подготовки глаз животное анестезировали внутримышечной инъекцией смеси ксилазина (6 мг/мл, 100 мкл) и золетила (10 мг/мл, 100 мкл), дополнительно применяя увлажняющие глазные капли для предотвращения высыхания роговицы и образования катаракты.

Во время наркоза мышь фиксировали под бинокуляром и далее с помощью иглы размером G27 прокалывали роговицу для введения иглы шприца Гамильтона G33. В отверстие размером G27 вводили иглу от шприца Гамильтона G33 на глубину 0,5 мм для интравитреального введения и на 1,5 мм для субретинального метода введения. Контроль субретинальной инъекции производился с помощь микроскопа при анализе глазного дна. После укола медленно вводили необходимый объём (не более 1 мкл) препарата и далее медленно извлекали иглу. Глаз после инъекции обрабатывали глазным антибиотиком тетрациклином и мазью для увлажнения глаз. После инъекции шприц Гамильтона троекратно промывали раствором PBS с помощью 5-8 нажатий на плунжер шприца. Для консервации шприц Гамильтона троекратно промывали 70%-ным раствором спирта, затем деионизованной водой и PBS.

2.4.3. Анализ тканей гистологическими методами

По окончании эксперимента мышей умерщвляли методом цервикальной дислокации и затем извлекали глазные яблоки из глазниц с помощью хирургических инструментов. После извлечения удаляли все прилегающие соединительные ткани и промывали глаз в растворе PBS. Для быстрой фиксации глазные яблоки инкубировали 4 часа в 4%-ном растворе параформальдегида и затем разделяли орган на отдельные ткани с помощью хирургических инструментов.

Для гистологического анализа глаз инкубировали 24 часа в 4%-ном растворе параформальдегида. Для увеличения плотности глаза препарат погружали в 30%-ный раствор сахарозы на 24-48 часов до полного оседания органа на дно пробирки.

Для подготовки к получению криосрезов образцы помещали в матричную среду Tissue-Tek® O.C.T. Compound (4583, Sakura, США) в специальные формы, которые затем погружали в чашку Петри с изопропанолом для обеспечения равномерного охлаждения. После этого формы быстро замораживали в жидком азоте. Замороженные блоки хранили при -80°C до получения срезов.

Срезы толщиной 20 мкм получали на криостате Leica CM1520 (Leica Microsystems, Германия) при температуре камеры -20°C. Полученные срезы переносили на предметные стёкла (BioVitrum, 0313-7171-16) и последовательно промывали в PBS для удаления OCT-среды. Затем проводили блокирование неспецифического связывания 50%-ной фетальной бычьей сывороткой (FBS) в PBS в течение 1 часа при комнатной температуре.

После блокирования срезы инкубировали с первичными антителами (те же описаны в разделе 2.3.5.) к целевому белку (anti-GFP (AB01101231265, Евроген, Россия) в случае GFP или anti-ABCA4 (Santa Cruz, sc-65672, США) в случае ABCA4), конъюгированными с активным эфиром AF647, в разведении 1:200 в 5%-ном растворе BSA в течение ночи при 4°C. На следующий день срезы трижды промывали PBS.

Далее срезы заключали в соответствующую монтажную среду LumiMount®DAPI (53144, Lumiprobe, США), после чего образцы анализировали на конфокальном микроскопе.

Для визуализации использовали конфокальный микроскоп ZEISS LSM 980. Сигнал целевых антител, меченных Alexa Fluor 647, регистрировали в дальней красной области спектра; справочные максимумы возбуждения и эмиссии конъюгатов этого красителя составляют приблизительно 650 и 668 нм соответственно. Для регистрации автофлуоресценции в зелёном канале использовали канал FITC. Спектральными ориентирами для этого канала служат максимумы возбуждения и эмиссии FITC — приблизительно 495 и 519 нм соответственно. Для детекции ядер использовали канал DAPI: при связывании с ДНК максимумы возбуждения и эмиссии DAPI составляют приблизительно 358 и 461 нм соответственно, что соответствует синей флуоресценции. Для съёмки изображений использовали объектив с 20-кратным увеличением. Приведённые спектральные максимумы не следует отождествлять с длинами волн фактически использованных лазерных линий и границами окон регистрации. Эти параметры определяются настройками съёмки и должны быть указаны по данным исходного эксперимента.

Обработку и анализ изображений срезов сетчатки проводили с использованием программного обеспечения ZEN 3.4 Blue edition.

ГЛАВА 3. РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

3.1. Дизайн экспериментальной части работы

Прикладная цель работы заключается в разработке препарата для терапии наследственных ретинопатий (на примере синдрома Штаргардта) на основе технологии белкового транс-сплайсинга, оптимизированной для повышения эффективности реакции белкового сплайсинга. В качестве основного критерия оценки работоспособности системы транс-сплайсинга была использована количественная метрика — эффективность транс-сплайсинга, оцениваемая по доле GFP-положительных клеток в общей популяции методом проточной цитометрии. Данный анализ позволяет комплексно оценить эффективность трансфекции и условия, влияющие на эффективность сборки разделённого белка GFP. Отдельно отметим, что оптимизация условий трансфекции проводили на модельном белке GFP, поскольку он удобен для флуоресцентной визуализации в клетках и генно-инженерных манипуляций.

Для достижения поставленной цели был разработан план оптимизации условий сборки сначала модельного белка GFP, а затем крупного белка ABCA4 для терапии синдрома Штаргардта с использованием оптимизированной пары интеинов.

На первом этапе проводили поиск, обзор и отбор наиболее изученных и широко используемых природных интеинов для дальнейших исследований. По результатам анализа литературы были выбраны интеины группы DnaE, которые отличаются высокой скоростью сплайсинга и наличием известных 3D-структур, позволяющих проводить дальнейший мутагенез. По данным литературы также был произведен подбор сайтов разделения молекул GFP и ABCA4, специфичных для интеинов группы DnaE. В ходе тестирования проверялись как уже известные сайты, так и вновь обнаруженные при проведённом анализе.

На следующем этапе производили оптимизацию условий трансфекции с использованием первых полученных конструкций в клетках линии HEK293. В ходе оптимизации варьировали концентрации каждого из компонентов трансфекционной смеси для увеличения эффективности реакции сплайсинга, определяемой по проценту GFP-положительных клеток.

После подбора оптимальных условий трансфекции проверяли гипотеу о повышении эффективности сборки целевого белка путём сравнения интеинов группы DnaE (SspDnaE, AvaDnaE и NpuDnaE) с разной кинетикой реакции. В ходе сравнения проверяли как канонические пары, так и варианты кросс-реактивных взаимодействий.

После выбора наиболее эффективной пары интеинов плазмидные конструкции использовали для получения генотерапевтических препаратов, которые далее тестировали на клеточных культурах и *in vivo*. В ходе работы сравнивались методы введения препаратов, а также платформы доставки на основе вирусных (аденоассоциированные вирусы) и невирусных (липидные наночастицы) частиц.

На основе результатов, полученных на модельном белке, были проведены аналогичные исследования на примере препарата для терапии синдрома Штаргардта.

3.2. Выбор интеинов и анализ аминокислотной последовательности экстеина

3.2.1. Выбор интеинов для экспериментальной работы

На основе данных, приведенных в литературном обзоре (см. Таблицу 1 и раздел 1.3) наиболее наглядными объектами для сравнительного анализа являются интеины группы DnaE, кинетика которых была исследована при 37 С, что является наиболее актуальным для применения в эукариотических системах. Эти интеины локализованы в гомологичных белках-хозяинах DnaE, но существенно различаются по скорости реакции, что позволяет сопоставлять их с учетом близкого эволюционного родства и экспериментально проверить свойство кросс-реактивного взаимодействия.

Для экспериментальной работы были выбраны интеины Ssp DnaE, Ava DnaE, Npu DnaE, которые обладают одинаковыми каталитическими доменами, но разной кинетикой реакции сплайсинга, что может быть использовано для проверки экспериментальных гипотез при условии одинакового дизайна разделения целевого белка. Ранее в таблице 2 уже рассматривались аминокислоты, влияющие на эффективность сплайсинга для интеинов данной ферментативной группы. Эту информацию применили для анализа белков GFP и ABCA4. На основе данных таблицы 2 был разработан программный код для оценки последовательности любого белка и поиска потенциальных сайтов его разделения. Фрагмент кода представлен на Рисунке 18.

```
SCORE_MATRIX = {
    "D": [1.39, 1.52, 0.00, 0.00, 0.00, 1.26],
    "E": [1.26, 3.51, 0.07, 0.00, 0.00, 5.50],
    "N": [0.93, 1.26, 2.19, 0.00, 0.00, 8.61],
    "Q": [1.39, 0.86, 0.33, 0.00, 0.00, 0.27],
    "H": [1.06, 1.26, 2.19, 0.00, 0.00, 8.61],
    "K": [1.59, 0.93, 4.44, 0.00, 0.07, 0.00],
    "R": [1.81, 0.20, 0.86, 0.00, 0.00, 0.00],
    "S": [0.99, 1.24, 1.37, 0.00, 0.00, 0.07],
    "C": [0.46, 0.40, 0.80, 50.00, 0.07, 0.07],
    "T": [0.63, 0.89, 1.59, 0.00, 0.00, 0.96],
    "P": [0.80, 1.52, 0.00, 0.00, 0.00, 0.00],
    "G": [4.67, 2.68, 0.17, 0.00, 0.00, 0.03],
    "A": [0.63, 0.83, 1.92, 0.00, 0.00, 0.00],
    "V": [0.56, 0.50, 0.23, 0.00, 0.03, 0.10],
    "I": [0.07, 0.53, 0.00, 0.00, 0.00, 0.40],
    "L": [0.11, 0.57, 0.75, 0.00, 0.00, 0.82],
    "M": [0.13, 0.73, 1.26, 20.00, 0.07, 2.25],
    "F": [0.13, 0.73, 1.26, 0.00, 0.07, 2.25],
    "Y": [0.20, 0.60, 2.05, 0.00, 0.13, 5.23],
    "W": [0.07, 0.40, 0.99, 0.00, 29.42, 0.07],
}
POSITIONS = ["-3", "-2", "-1", "+1", "+2", "+3"]
def analyse_6mers(sequence):
    # Удаляем пробелы, переносы строк и стоп-кодон
    sequence = sequence.upper()
    sequence = sequence.replace(" ", "")
    sequence = sequence.replace("\n", "")
    sequence = sequence.replace("*", "")
    # Проверяем аминокислоты
    unknown = set(sequence) - set(SCORE_MATRIX.keys())
    if unknown:
        raise ValueError(
            f"В последовательности обнаружены неизвестные символы: {unknown}"
        )
    results = []
    # Все возможные 6-меры
    for i in range(len(sequence) - 5):
        sixmer = sequence[i:i+6]
        # Баллы для каждой позиции
        scores = [
            SCORE_MATRIX[aa][position]
            for aa, position in zip(sixmer, range(6))
        ]
        total_score = sum(scores)
        # Координаты в исходной последовательности
        start = i + 1
        end = i + 6
        # Позиции Cys внутри 6-мера
        cys_positions = [
            start + j
            for j, aa in enumerate(sixmer)
            if aa == "C"
        ]
        results.append({
            "Start": start,
            "End": end,
            "6-mer": sixmer,
            "-3": scores[0],
            "-2": scores[1],
            "-1": scores[2],
            "+1": scores[3],
            "+2": scores[4],
            "+3": scores[5],
            "Total score": total_score,
            "Cys present": "Yes" if "C" in sixmer else "No",
            "Cys position":
                ", ".join(map(str, cys_positions))
                if cys_positions else ""
        })
    df = pd.DataFrame(results)
    # Сортировка от большего балла к меньшему
    df = df.sort_values(
        by="Total score",
        ascending=False
    ).reset_index(drop=True)
    # Номер строки после сортировки
    df.insert(0, "Rank", range(1, len(df) + 1))
    return df
```

Рисунок 18 — Фрагмент кода программы для оценки аминокислотной последовательности белка и поиска сайтов его разделения с помощью интеинов группы DnaE

Таким образом, был получен универсальный инструмент, позволяющий выбирать сайты разделения экстеиновых белков с учетом их аминокислотной последовательности. Метод основан на биохимических свойствах аминокислот и их влиянии на катализ реакции сплайсинга. Однако он не позволяет предсказать функциональность и корректную конформацию получаемых молекул для белков с неустановленной структурой. Поэтому дальнейшие эксперименты по выбору сайта сплайсинга проводили на молекуле GFP, для которой известны аминокислотная последовательность и структура.

3.2.2. Анализ аминокислотной последовательности GFP для разделения по сайтам сплайсинга для интеинов группы DnaE

По результатам работы алгоритма была составлена Таблица 5, в которой приведены шестиаминокислотные последовательности с указанием мест разделения и числовой оценкой каждого сайта.

Таблица 5 — Результаты ранжирования потенциальных сайтов разделения GFP для интеинов группы DnaE

Позиция в рейтинге	Последовательность сайта разделения	Оценка сайта по баллам	Позиция N-1/C+1 аминокислоты	Тип вторичных структур в сайте разделения	Источник
1	GVQCFS	55,64	69/70	альфа-спираль	[63]
2	KFICTT	52,32	47/48	бета-лист	-
3	LPVPWP	34,32	55/56	альфа-спираль	-
4	KSAMPE	30,25	86/87	альфа-спираль	[113]
5	TLGMDE	26,87	228/229	неструктурированная область	-
6	RDHMVL	26,37	217/218	бета-лист	-
7	PDHMKQ	24,85	77/78	альфа-спираль	-
8	VYIMAD	22,42	152/153	бета-лист	-
9	GEGEGD	22,29	33/34	бета-лист	-
10	GDVNGH	19,7	21/22	неструктурированная область	-
11	ADKQKN	19,64	156/157	неструктурированная область	[62,114]
76 позиции пропущены					
87	KGIDFK	5,93	127/128	неструктурированная область	[115]
98 позиции пропущены					
185	TTLTYG	2,8	64/65	неструктурированная область	[113]
49 позиции пропущены					
234	LVELDG	0,82	17/18	неструктурированная область	-

В результате работы алгоритма было сгенерировано 234 варианта разделения молекулы GFP с указанием числовой оценки для каждого сайта сплайсинга. В первую очередь были рассмотрены позиции с наивысшим рейтингом, выделенные зеленым цветом. Сайт 69/70 получил наивысшую оценку благодаря наличию цистеина в положении C+1 (50 баллов) и комбинации GVQ в N-концевой части, обеспечивающей 5,5 балла. Следует отдельно отметить, что данный сайт неоднократно использовался в нескольких работах и может быть выбран в качестве положительного контроля. Сайт 47/48 также содержит цистеин в аналогичном положении, однако комбинация KFI в N-концевой части обеспечивает лишь 2,32 балла, поскольку изолейцин в положении N-1 дает нулевой вклад в реакцию сплайсинга. Ранее данный сайт не был описан в работах других авторов.

Остальные позиции с рейтингом от 20 до 30 баллов содержат метионин или триптофан в положении C+2, однако не встречаются ни в одной из рассмотренных работ. В литературе упоминается сайт 127/128 с последовательностью KGIDFK, который использовали для разделения GFP с помощью интеина VDE. При анализе данного сайта с помощью алгоритма для интеинов DnaE была получена оценка в 5,93 балла, что указывает на специфичность выбора сайта разделения молекулы для интеинов определенной группы. Таким образом, для дальнейшей экспериментальной проверки были выбраны следующие сайты:

- 69/70 — положительный контроль сплайсинга (ранее опубликованный сайт);
- 47/48 — экспериментальный вариант (ранее не опубликован);
- 127/128 — отрицательный контроль сплайсинга
  (сайт, ранее опубликованный для интеина VDE и белка GFP).

В следующем разделе представлено моделирование получаемых структур с помощью AlphaFold 3, проведенное для предварительной оценки возможной конформации молекул после их разделения интеинами.

3.4.3. Моделирование получаемых предсплайсинговых молекул с помощью AlphaFold 3

Для моделирования использовали аминокислотные последовательности белка GFP и интеина SspDnaE, экспериментальные структуры которых известны. Ключевым достоинством AlphaFold 3 является возможность построения моделей на основе аминокислотных или нуклеотидных последовательностей. Каждый из трех вариантов

разделения белка GFP был загружен в систему в соответствии с дизайном, учитывающим выбранные сайты разделения. Результаты моделирования представлены на Рисунке 19.

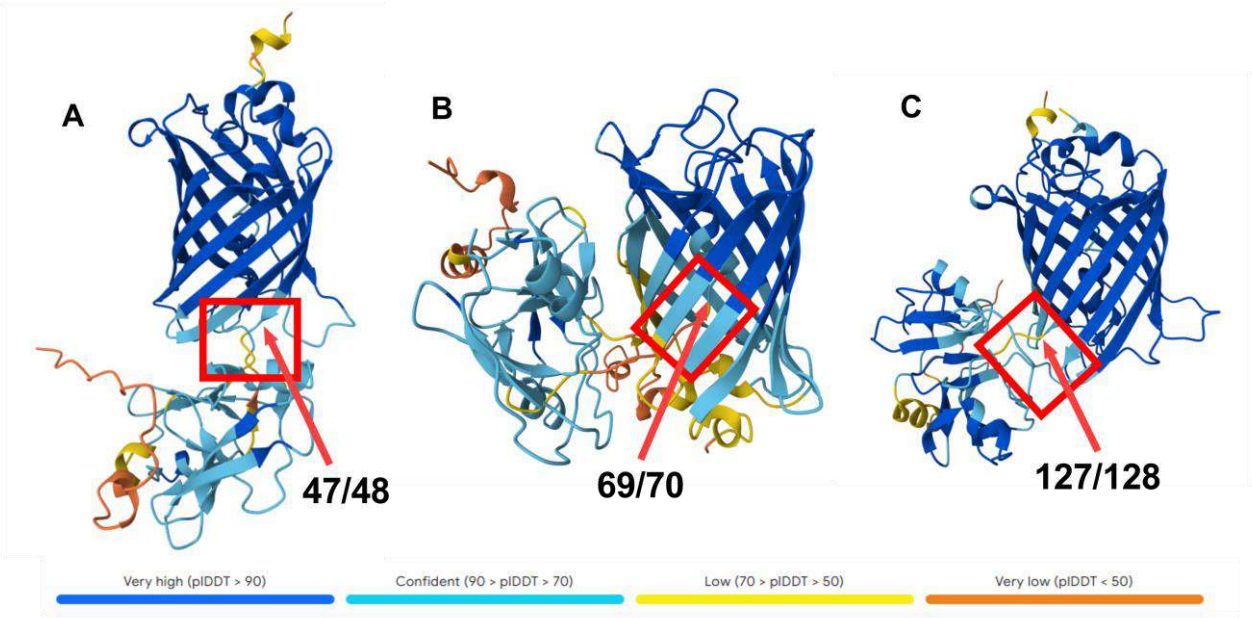


Рисунок 19 — Результаты моделирования предсплайсинговых комплексов GFP/SspDnaE с помощью AlphaFold 3 для сайтов разделения:
A — модель комплекса с сайтом разделения GFP между аминокислотными остатками 47 и 48; B — модель комплекса с сайтом разделения между остатками 69 и 70;
C — модель комплекса с сайтом разделения между остатками 127 и 128.
Красными рамками и стрелками обозначены области сайтов разделения GFP.
Цветовая шкала отражает локальную уверенность предсказания структуры по показателю pLDDT: синий — очень высокая (более 90), голубой — высокая (70–90), жёлтый — низкая (50–70), оранжевый — очень низкая (менее 50).

На Рисунке 19 представлены модели комплексов, соответствующих трем сайтам разделения GFP. Во всех вариантах N- и C-концевые части SspDnaE взаимодействуют друг с другом, а фрагменты GFP формируют общий комплекс. Основные различия между моделями наблюдаются во взаимном расположении интеина и частей GFP, а также в области их стыковки.

В модели с сайтом разделения 47/48 (Рисунок 19A) фрагменты GFP и SspDnaE образовывали общий комплекс, однако концы частей GFP находились на значительном расстоянии друг от друга. Поэтому для их соединения необходима конформационная перестройка, которая может затруднять протекание реакции сплайсинга. Для сайта 69/70

(Рисунок 19B) концы фрагментов GFP располагались значительно ближе, однако область их стыковки характеризовалась большей подвижностью и менее определенной взаимной ориентацией. В модели с сайтом 127/128 (Рисунок 19C) также наблюдали близкое расположение концов GFP, при этом общая структура комплекса была более упорядоченной и устойчивой.

Таким образом, наиболее благоприятная геометрия для последующего сплайсинга была предсказана для сайтов 69/70 и 127/128. Однако укладка и пространственная ориентация не дают полной информации о функциональности представленных молекул, поэтому полученные результаты требовали экспериментальной проверки *in vitro.*

Таким образом, инструмент может быть использован для наглядной визуализации прекурсорных молекул, однако по полученным моделям сложно сделать однозначный вывод об их реакционной способности.

3.4.4. Функциональная проверка сайтов для разделения GFP с помощью интеина Ssp DnaE

Для проверки работоспособности алгоритма разделения молекулы GFP и результатов предсказания предсплайсинговых структур использовали интеин SspDnaE, нуклеотидная последовательность которого была синтезирована из праймеров, представленных в приложении. После синтеза N- и C-конструкции SspDnaE соединяли с соответствующими фрагментами последовательности белка GFP в области выбранных сайтов разделения. Полученные конструкции клонировали в вектор pAAV-CMV-MCS по сайтам BamHI и HindIII.

В результате был сформирован набор из шести плазмид, которые использовали для прямой трансфекции с помощью полиэтиленимина по стандартному протоколу. Суммарное количество ДНК составляло 2 мкг (по 1 мкг каждой плазмиды из пары) на 1 млн клеток, а соотношение ДНК:PEI — 1:1. Через 48 ч после трансфекции клетки анализировали с помощью флуоресцентной микроскопии и проточной цитометрии. Результаты представлены на Рисунке 20.

В нетрансфицированных клетках HEK293 флуоресценция практически отсутствовала, тогда как трансфекция плазмидой pEGFP приводила к выраженному накоплению GFP и использовалась в качестве положительного контроля (Рисунок 20 A, B). Среди конструкций с интеином SspDnaE наиболее интенсивную флуоресценцию наблюдали для сайта 69/70 (Рисунок 20E).

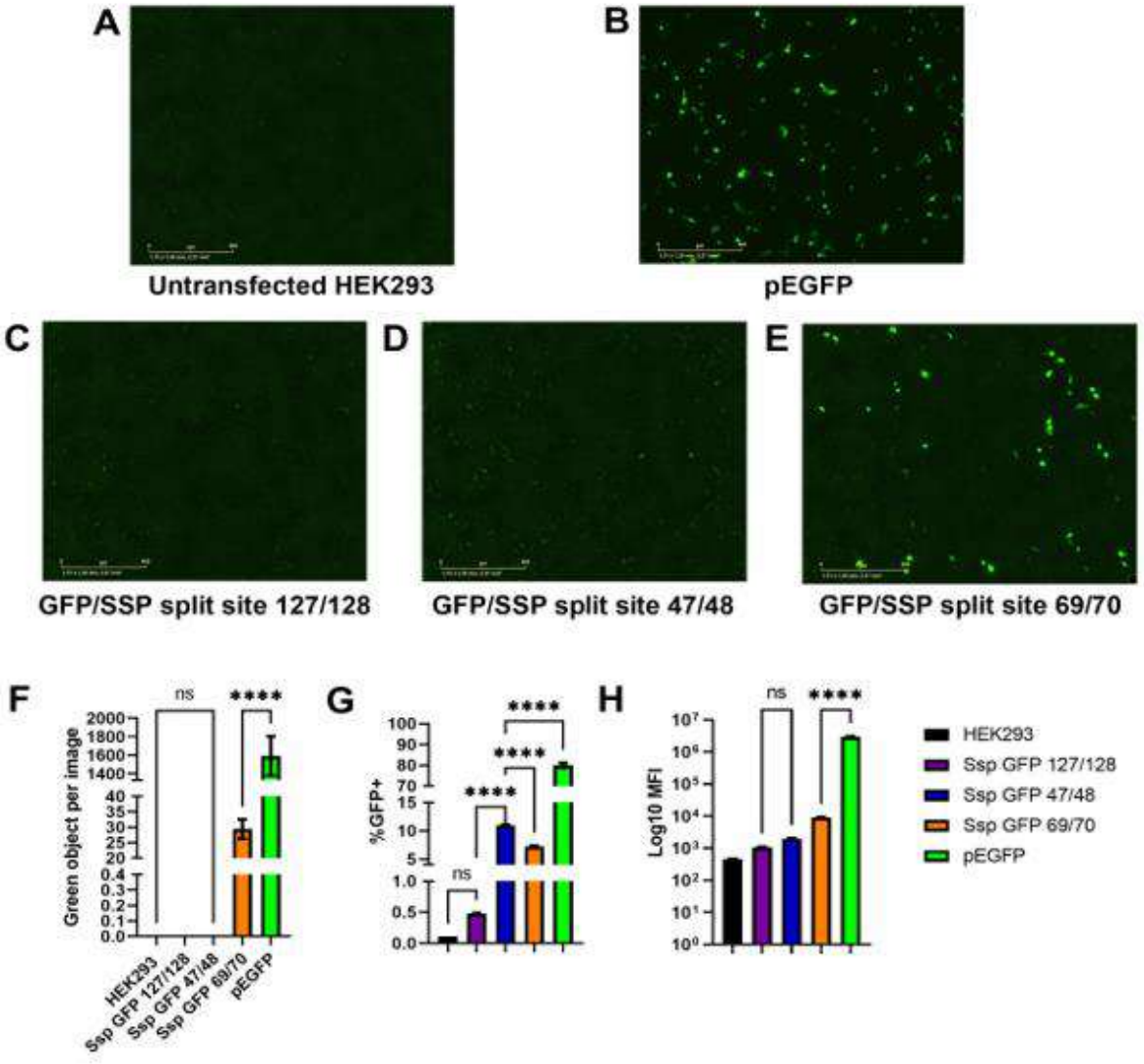


Рисунок 20 — Результаты функциональной проверки сайтов разделения GFP 47/48, 69/70 и 127/128 с использованием интеина SspDnaE: флуоресцентная микроскопия и проточная цитометрия через 48 ч после трансфекции.

A–E — флуоресцентная микроскопия:

A — нетрансфицированные клетки (отрицательный контроль);

B — клетки после трансфекции плазмидой pEGFP (положительный контроль);

C — сайт разделения GFP 127/128; D — сайт 47/48; E — сайт 69/70.

F — число зелёных флуоресцентных объектов на изображение Incucyte;

G — доля GFP-положительных клеток, определённая методом проточной цитометрии;

H — средняя интенсивность флуоресценции (MFI) по данным проточной цитометрии, представленная в логарифмической шкале.

Для сайта 47/48 регистрировался более слабый сигнал (Рисунок 20D), а конструкция 127/128 практически не отличалась от нетрансфицированных клеток (Рисунок 20C).

Визуальный анализ микроскопических изображений (Рисунок 20 A–E) выявил выраженную флуоресценцию для сайта сплайсинга 69/70 и положительного контроля. Результаты количественной оценки числа флуоресцентных объектов представлены на Рисунке 20 F.

Данные проточной цитометрии (Рисунок 20 G, H) подтвердили результаты микроскопии. Доля GFP-положительных клеток составляла около 7% для сайта 69/70 и около 10% для сайта 47/48, тогда как для сайта 127/128 она оставалась ниже 1% и статистически не отличалась от отрицательного контроля (Рисунок 20 G). Несмотря на более высокую долю GFP-положительных клеток для сайта 47/48, средняя интенсивность флуоресценции была максимальной для конструкции 69/70, ниже для 47/48 и находилась на фоновом уровне для 127/128 (Рисунке 20 H).

Полученные результаты можно объяснить следующим образом. Сайт 69/70 подтвердил возможность его использования в качестве положительного контроля, что может быть связано с ненулевым вкладом каждой аминокислоты в последовательности экстеинового окружения. Дополнительным фактором могла быть подвижная и доступная для реакции структура предсплайсингового комплекса, благоприятствующая протеканию сплайсинга.

Результаты для экспериментального сайта 47/48 показали, что одного наличия цистеина в положении C+1 недостаточно для эффективной реакции. Последовательность N-экстеина имела низкую суммарную оценку 2,32 балла и содержала изолейцин в положении N-1, вклад которого в эффективность сплайсинга был принят равным нулю. Кроме того, физическая разобщенность концов GFP в предсказанной структуре могла дополнительно снижать эффективность реакции. При микроскопическом и цитометрическом анализе наблюдались отдельные GFP-положительные клетки, однако по большинству показателей эффективность сплайсинга оставалась низкой. Для сайта 127/128, использованного в качестве отрицательного контроля сплайсинга, как и в нетрансфицированных клетках, значимой флуоресценции не наблюдали. Следовательно, доступность концов GFP сама по себе не обеспечивала протекания реакции при отсутствии функционально значимого аминокислотного окружения сайта сплайсинга.

Таким образом, для функционального восстановления молекулы в первую очередь важен аминокислотный состав N- и C-концевого экстеинового окружения интеина, включая наличие подходящих остатков в положениях N-1 и C+1.

Дополнительным благоприятным фактором является конформация комплекса, обеспечивающая доступность реакционного центра и близкое расположение соединяемых концов. Не менее важно учитывать специфичность экстеиновых сайтов для каждого типа интеинов, поскольку положение и свойства аминокислот непосредственно влияют на катализ реакции. Для последующих экспериментов с интеинами группы DnaE был выбран сайт 69/70. Предположение о возможности его использования с другими интеинами данного семейства будет проверено в дальнейших экспериментах. Кроме того, необходимо оптимизировать условия сборки целевого белка, поскольку при использованных условиях доля GFP-положительных клеток для наиболее эффективного сайта составляла около 7%, что недостаточно для потенциального терапевтического применения.

3.5. Оптимизация выхода продукта реакции сплайсинга GFP

В предыдущем разделе была продемонстрирована система плазмид для сборки GFP с помощью интеина SspDnaE. При использовании аминокислотного сайта 69/70 эффективность реакции сплайсинга составляла всего 7%, что недостаточно для терапевтического применения.

Для решения данной проблемы в первую очередь подбирали оптимальные условия доставки плазмидной ДНК с помощью трансфицирующего агента, в качестве которого использовали полиэтиленимин. Выбор трансфицирующего агента был обусловлен технологией получения частиц AAV, важным требованием которой является возможность быстрого и простого масштабирования объемов наработки вирусных частиц. Протоколы масштабирования предполагают использование одних и тех же компонентов с пропорциональным увеличением количества реагентов в соответствии с числом клеток. В этом отношении полиэтиленимин имеет преимущества перед известными аналогами на основе липосом.

Для трансфекции с помощью полиэтиленимина использовали три компонента: смесь плазмидной ДНК, полиэтиленимин и клетки в адгезионном или суспензионном состоянии. В ходе оптимизации варьировали концентрациии и соотношения компонентов, а также применяли методы прямой и обратной трансфекции. Дополнительно оценивали выживаемость клеток, чтобы определить оптимальные концентрации компонентов, не приводящие к её существенному снижению. Таким образом, условия трансфекции оптимизировали по критерию максимальной эффективности сплайсинга GFP, оцениваемой по доле GFP положительных клеток, с учётом их выживаемости. Результаты проведенных этапов оптимизации представлены на Рисунке 21.

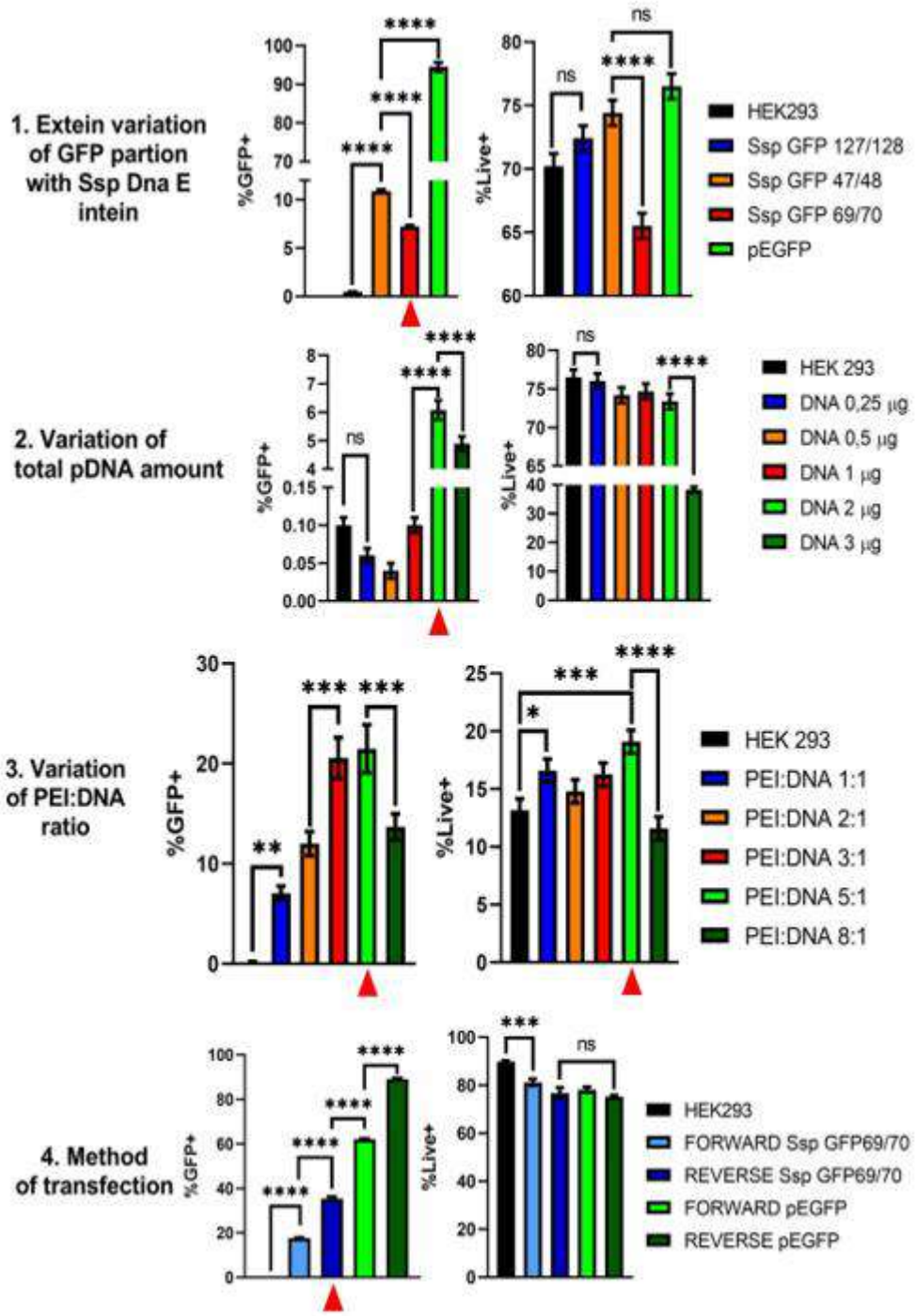


Рисунок 21 — Оптимизация условий трансфекции для повышения эффективности восстановления GFP с помощью интеина SspDnaE:
1 — сравнение сайтов разделения GFP;
2 — подбор суммарного количества плазмидной ДНК;
3 — подбор соотношения PEI:ДНК;
4 — сравнение прямого и обратного методов трансфекции,
красной стрелкой указаны оптимальные условия

Исходная эффективность сплайсинга для сайта 69/70 составляла 7% (Рисунок 21.1) при следующих условиях прямой трансфекции: суммарное количество ДНК — 2 мкг (по 1 мкг каждой плазмиды из пары) на 1 млн клеток, соотношение PEI:ДНК — 1:1. Следует отметить, что доля GFP-положительных клеток для сайта 47/48 составляла 10%, однако яркие флуоресцентные объекты при анализе с помощью микроскопии практически отсутствовали (см. результаты в предыдущем разделе). Отметим также, что в варианте 69/70 наблюдали падение выживаемости клеток на 10% относительно остальных вариантов, которое вносит вклад в количество GFP-положительных клеток, однако средняя интенсивность при этом не падает.

Детекция флуоресцеции в случае 47/48 может быть также связана с высокой экспрессией С-концевого фрагмента GFP, содержащего большую часть β-бочонка и хромофорную область, что потенциально способно создавать автофлуоресцентный или ложноположительный сигнал. Вместе с тем отсутствие интесивной флуоресценции по MFI и при микроскопии (Рисунок 20) не подтверждает функциональное восстановление GFP. Поэтому вопрос о восстановлении функции белка после сплайсинга требует проведения дополнительных клеточных или биохимических тестов на уровне функциональной активности.

Для последующих экспериментов по оптимизации состава трансфицирующей смеси использовали сайт 69/70. Оптимизация включала подбор плазмидной нагрузки, соотношения PEI:ДНК, а также сравнение процедур прямой и обратной трансфекции.

На первом этапе подбирали плазмидную нагрузку в диапазоне от 0,25 до 3 мкг суммарной ДНК. Плазмиды из пары добавляли в равных массовых долях при фиксированном соотношении PEI:ДНК — 1:1 и прямом методе трансфекции. Результаты эксперимента представлены на Рисунке 21.2, где отчетливо видна разница между использованием 1 и 2 мкг суммарной плазмидной ДНК. При нагрузке 2 мкг эффективность достигала 7% без снижения жизнеспособности клеток, которое отчетливо наблюдается в случае 3 мкг. Поэтому в последующих экспериментах использовали суммарную плазмидную нагрузку 2 мкг ДНК.

На следующем этапе оптимизировали соотношение PEI:ДНК путем изменения массы PEI в диапазоне от 2 до 16 мкг при фиксированной массе суммарной ДНК 2 мкг. Результаты эксперимента представлены на Рисунке 21.3. Переход от соотношения 2:1 к 3:1 сопровождался двукратным увеличением эффективности — с 11 до 22%.

Дальнейшее увеличение количества трансфицирующего агента (3:1 и 5:1) не приводило к существенному повышению эффективности сборки белка GFP, но увеличивало долю погибших клеток, что указывало на токсическое действие избытка

полиэтиленимина в случае соотношения 8:1. Для дальнейших экспериментов использовали 2 мкг суммарной ДНК и 10 мкг PEI, что соответствовало соотношению PEI:ДНК 5:1 и обеспечивало оптимальный баланс между эффективностью трансфекции и жизнеспособностью клеток.

На заключительном этапе оптимизации сравнивали прямую и обратную трансфекцию. При обратной трансфекции трансфицирующую смесь добавляли к клеткам, находившимся в суспензии. Это обеспечивало более равномерное взаимодействие комплексов PEI:ДНК с клетками во всем объеме пробирки. При прямой трансфекции смесь добавляли к уже прикрепленным клеткам, что затрудняло точный контроль числа клеток и равномерности распределения реагентов по клеточному монослою. Результаты сравнения двух протоколов представлены на Рисунок 21.4.

При трансфекции одной плазмидой, кодирующей полноразмерный GFP, эффективность достигала 80% при обратной трансфекции и 60% при прямой. При использовании пары плазмид эффективность сборки GFP составляла 17% при прямой трансфекции и 35% при обратной. Таким образом, применение обратной трансфекции приводило к двукратному увеличению эффективности сборки GFP посредством транс-сплайсинга, тогда как при доставке полноразмерного GFP прирост составил 30%.

Таким образом, оптимальный протокол трансфекции клеток HEK293 набором плазмид SspDnaEGFP69/70 с использованием полиэтиленимина включал 2 мкг суммарной плазмидной ДНК при соотношении PEI:ДНК 5:1 и добавление трансфицирующей смеси к клеточной суспензии, то есть проведение обратной трансфекции. Максимальная эффективность сборки белка GFP в этих условиях составляла 35%, что все еще недостаточно для терапевтических приложений, требующих эффективности, близкой к 70-75%. Для дальнейшего увеличения эффективности сборки модельного белка использовали быстрые интеины Ava и Npu, которые относятся к группе DnaE и совместимы с выбранным сайтом разделения экстеинового белка.

3.6. Оптимизация сплайсинга с помощью интеинов и их кросс-реактивного взаимодействия

В предыдущих разделах было показано, что интеин SspDnaE способен эффективно собирать белок GFP по сайту 69/70 с максимальной эффективностью 35%. В таблице 6 приведены данные о кинетике реакции сплайсинга для других интеинов группы DnaE,

обладающих сходными ферментативными свойствами и консервативными каталитическими доменами.

Для проверки возможности использования выбранного сайта и сравнения каталитической активности быстрых и медленных интеинов группы DnaE были заказаны праймеры для сборки генетических конструкций. Дизайн предсплайсинговых белков был полностью аналогичен дизайну парной системы SspDnaE/GFP 69/70, а корректность сборки конструкций проверяли с помощью секвенирования по Сенгеру.

Таблица 6 — Сопоставление скорости реакции сплайсинга и эффективности восстановления GFP по сайту 69/70

Интеин	Время сплайсинга, с	Эффективность, %
SspDnaE	4560	58
AvaDnaE	23	66
NpuDnaE	19	81

Для трансфекции использовали оптимизированный протокол, описанный в предыдущем разделе. Результаты сравнительного эксперимента представлены на Рисунке 22. По результатам визуального анализа микрофотографий, полученных с помощью флуоресцентной микроскопии (Рисунок 22 B–E), было сложно однозначно оценить различия между образцами. Однако количественный анализ числа зеленых объектов (Рисунок 22 F) выявил статистически значимые различия между вариантом NpuDnaE/GFP 69/70 и вариантами AvaDnaE/GFP 69/70 и SspDnaE/GFP 69/70. Аналогичные различия были подтверждены методом проточной цитометрии (Рисунок 22 G, H). Эффективность сплайсинга возрастала при увеличении скорости реакции, то есть при уменьшении времени сплайсинга, однако данная зависимость была нелинейной.

Таким образом, было экспериментально показано, что выбранный сайт разделения GFP 69/70 совместим с используемыми интеинами группы DnaE. Эффективность сборки экстеина возрастала при увеличении скорости реакции сплайсинга, однако зависимость не была прямо пропорциональной. Интеин NpuDnaE обеспечивал восстановление GFP с эффективностью около 80%, что приближалось к значению положительного контроля с прямым синтезом полноразмерного флуоресцентного белка.

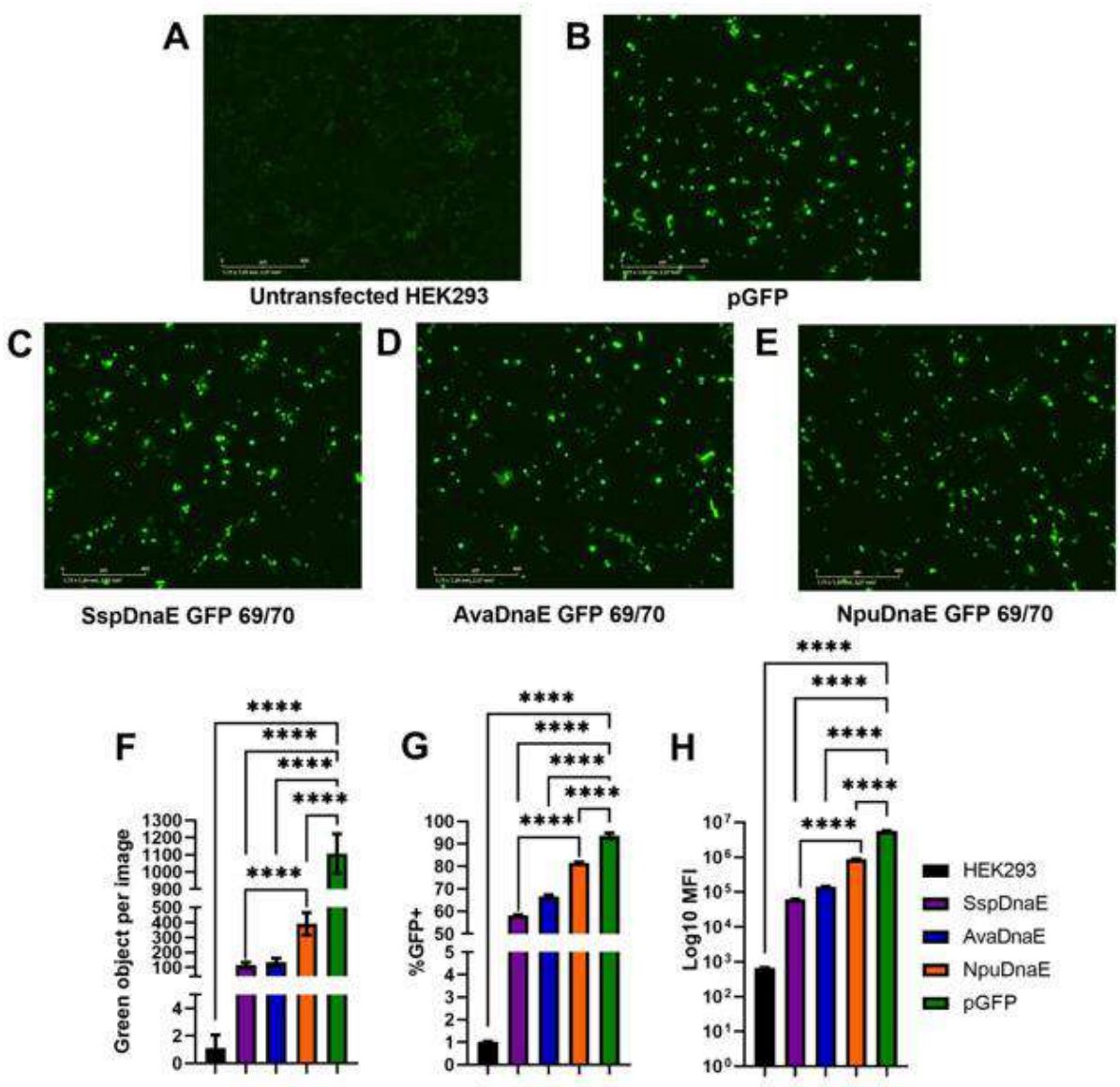


Рисунок 22 — Сравнение эффективности восстановления GFP по сайту 69/70 с помощью интеинов SspDnaE, AvaDnaE и NpuDnaE:
A — нетрансфицированные клетки HEK293; B — положительный контроль pGFP;
C–E — клетки после трансфекции конструкциями SspDnaE, AvaDnaE и NpuDnaE соответственно; F — количество флуоресцентных объектов;
G — доля GFP-положительных клеток; H — средняя интенсивность флуоресценции

С учетом совместимости выбранного сайта разделения и сходства каталитических доменов актуальным оставался вопрос о специфичности взаимодействия интеинов одной группы. Это свойство может иметь важное значение при разделении крупного белка более чем на две части. Для исследования кроссреактивности N- и C-концевые части трех интеинов группы DnaE комбинировали во всех возможных сочетаниях. Результаты сравнительного эксперимента представлены на Рисунке 23.

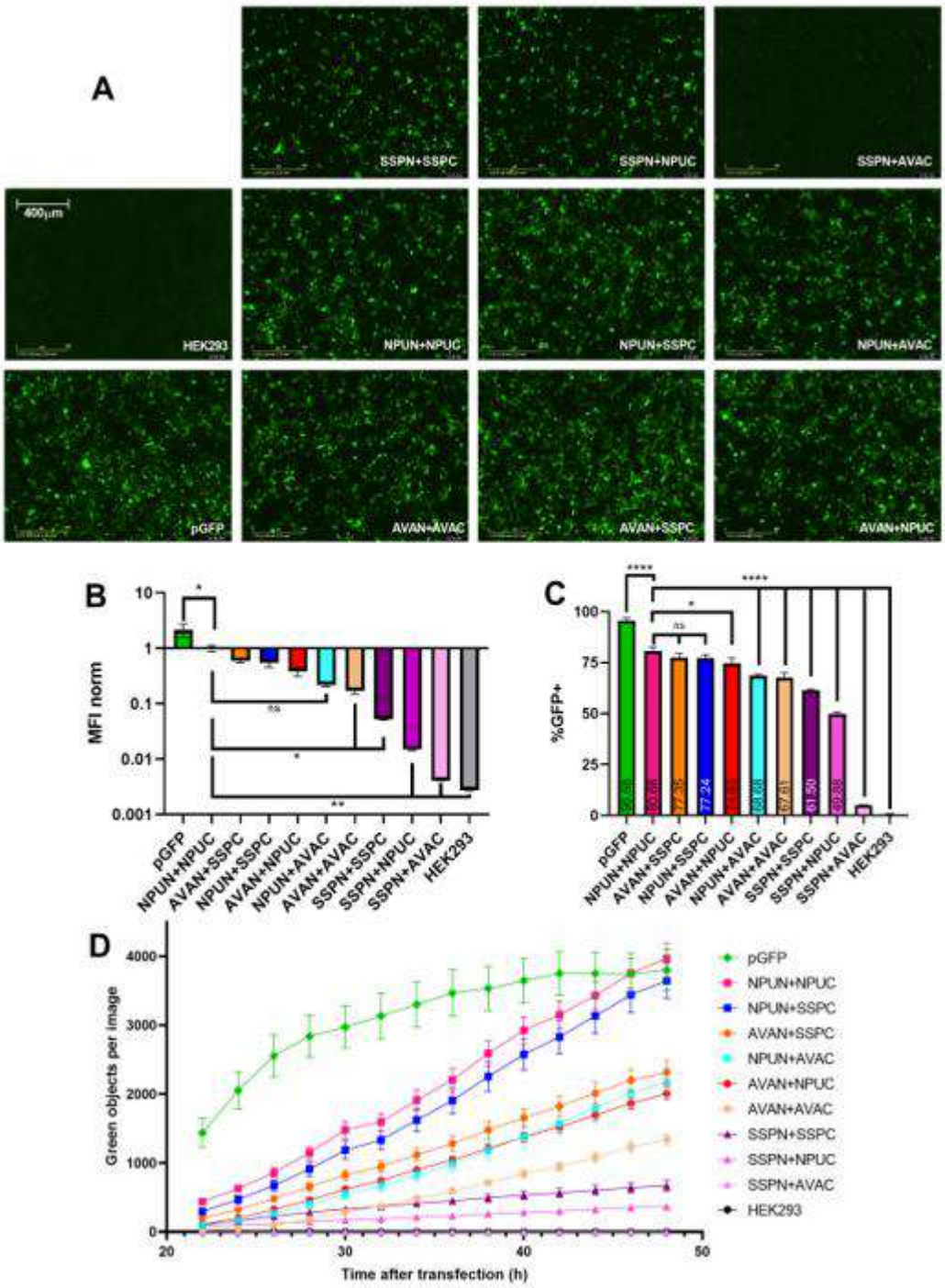


Рисунок 23 — Исследование кросс-реактивности N- и C-концевых частей интеинов SspDnaE, NpuDnaE и AvaDnaE при восстановлении GFP по сайту 69/70:
A — флуоресцентные микрофотографии клеток HEK293 после трансфекции девятью комбинациями фрагментов интеинов;
B — нормированная средняя интенсивность флуоресценции (MFI), представленная в логарифмической шкале; C — доля GFP-положительных клеток.
D — динамика накопления числа зелёных флуоресцентных объектов на изображение после трансфекции (до 48 ч).

На микрофотографиях (Рисунок 23 A) видно, что реакция сплайсинга протекала как при гомологичном, так и при большинстве гетерологичных сочетаний фрагментов интеинов. Исключением являлся вариант SspN+AvaC, для которого восстановление GFP практически не наблюдали. Данный результат указывает на структурную несовместимость этих фрагментов и создает предпосылки для более детального сравнительного анализа C-концевых частей интеинов.

Анализ динамики накопления сигнала (Рисунок 23 D) показал, что пары NpuN+NpuC и NpuN+SspC через 72 ч после трансфекции обеспечивали накопление GFP, сопоставимое с положительным контролем, представленным трансфекцией плазмидой, кодирующей полноразмерный GFP. Для остальных комбинаций наблюдали более низкую эффективность восстановления белка. Анализ методом проточной цитометрии (Рисунок 23 B, C) позволил подробнее охарактеризовать различия между комбинациями.

В большинстве случаев эффективность сплайсинга была ниже, чем для наиболее эффективного варианта NpuN+NpuC. В группах NpuN+NpuC, NpuN+SspC и AvaN+SspC с эффективность была близка к 80%. Все активные комбинации потенциально могут быть использованы в дальнейшей работе, однако пара NpuN+NpuC продемонстрировала наилучшие показатели эффективности и кинетики реакции, поэтому для последующих этапов была выбрана именно она.

3.7. Применение оптимальной интеиновой пары для раздельной доставки GFP с помощью AAV в клетки разного происхождения

3.7.1. Вестерн-блот-анализ клеток после трансфекции

После выбора оптимальной интеиновой пары NpuN+NpuC для белка GFP плазмиды были модифицированы путём добавления последовательности, кодирующей шестигистидиновую метку на C-конце каждой из конструкций pAAVCMV-GFP69_N-NpuN и pAAV-CMV-NpuC-GFP70_C для визуализации процесса транс-сплайсинга на уровне белка. Результаты электрофореза белков, их переноса на мембрану и окрашивания специфичными антителами anti-GFP и anti-6His представлены на Рисунке 24.

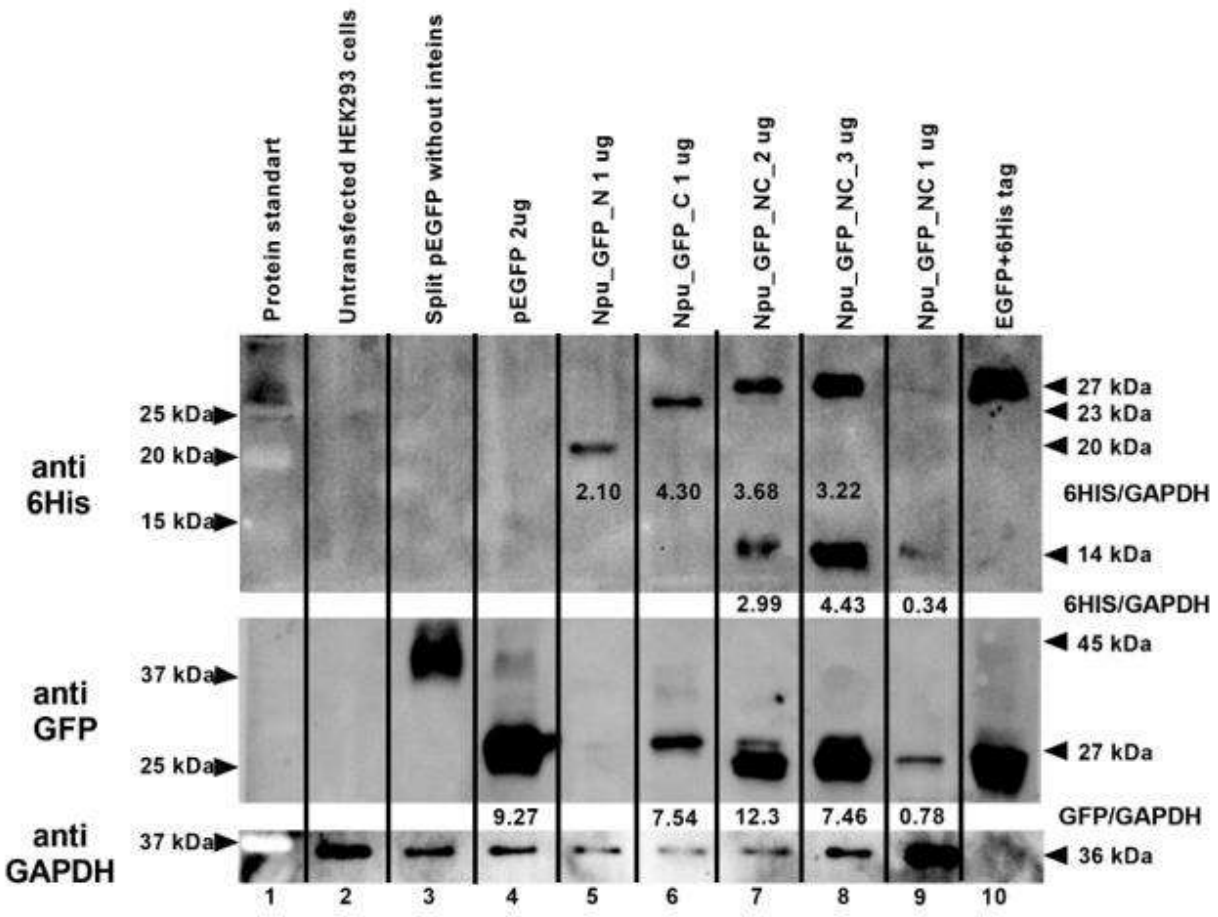


Рисунок 24 — Результаты вестерн-блоттинга со специфическими антителами: верхняя панель — мембрана с антителами к 6×His-tag, средняя панель — мембрана с антителами к GFP, нижняя панель — мембрана с антителами к GAPDH

Для комплексной проверки антител и работоспособности системы белкового сплайсинга использовали следующие контроли. Для проверки специфичности антител anti-GFP использовался контрольный очищенный белок EGFP_6His_tag, наработанный в системе E. coli (дорожка 10). Для контроля эффективности трансфекции HEK293 использовали плазмиду pAAV-CMV-GFP, не содержащую метку с гистидинами (дорожка 4). Для контроля автосплайсинга GFP использовалась пара плазмид с разделенным белком по сайту 69/70 без интеинов (дорожка 3). Нетрансфицированные клетки HEK293 использовали для проверки антител anti-GAPDH (дорожка 2).

Качественный анализ мембраны, последовательно окрашенной тремя типами антител (anti-6His tag, anti-GFP, anti-GAPDH) со стриппингом между этапами окрашивания, позволили сделать следующие выводы. Антитела к шестигистидиновой метке позволяют

детектировать как предсплайсинговые (дорожки 5 и 6), так и постсплайсинговые молекулы (дорожки 7–9).

Нормализация сигнала по GAPDH показывает, что С-часть нарабатывается примерно в два раза эффективнее, чем N-часть, что может быть связано с меньшей длиной С-фрагмента интеина и большей доступностью нуклеотидной последовательности для трансляции белка. Результаты нормализации также указывают на близкие пропорции количества белка до сборки (N_GFP+Npu_N — 2,1, Npu_C+GFP_C — 4,3) и после сборки GFP (GFP — 3,68, Npu_N+ Npu_C — 2,99) при использовании одинакового количества плазмидной ДНК по отдельности и вместе. При увеличении количества плазмид в 1,5 раза (дорожка 8) количество GFP сокращается, а количество интеина увеличивается, что говорит об ограниченной работоспособности реакции сплайсинга целевого белка.

При анализе мембраны, окрашенной антителами anti-GFP, важно отметить две основные особенности.

Во-первых, комбинация плазмид с разделенным GFP без интеинов приводит к образованию белкового конструкта некорректной формы и размера, который не обладает флуоресцентными свойствами, но содержит сайт связывания антител (дорожка 3).

Во-вторых, из двух предсплайсинговых молекул распознается только С часть, поскольку она содержит более крупный фрагмент белка GFP (с 69-й по 238-ю аминокислоту) и сайт связывания антител (дорожки 5 и 6). Отметим также, что С-фрагмент (дорожка 6) образует полосу, которая также наблюдается выше основной полосы GFP на дорожках 7 и 8, что говорит о возможной неполной реакции с N-частью, которая синтезируется вдвое менее интенсивно.

Однозначный вывод в этом случае сделать невозможно, поскольку N-часть не распознается антителами к GFP. Отметим также, что при добавлении вдвое меньшего количества плазмиды полоса над GFP не детектируется (дорожка 9), а при добавлении в полтора раза большего количества полосы GFP и NpuC_GFP_C сливаются в один фрагмент, что говорит о возможном некотором ограничении образования целевой молекулы в реакции сплайсинга.

При анализе мембраны, окрашенной антителами anti-GAPDH, можно отметить неравномерное распределение белка в образцах при нанесении одинакового количества тотального белка — 20 мкг на лунку. Лучше всего данный эффект проявляется на дорожках 2, 7 и 9, где белки распределяются по массам и количеству, детектируемому каждым типом антиген-специфичных антител. Количественная нормализация по GAPDH позволяет компенсировать данные различия и сделать выводы об эффективности синтеза белка и ограничениях его сборки.

Основной вывод, который можно сделать из вестерн-блот анализа, касается диспропорции синтеза предсплайсинговых молекул и последующего возможного ограничения реакции сборки целевой молекулы при накоплении непрореагировавшей С-части. Данную диспропорцию можно компенсировать путём изменения соотношения плазмид в паре до N:C – 2:1 при сохранении суммарного количества добавляемой ДНК на уровне 2 мкг. Второй вывод касается получения неправильно собранного белка GFP в ходе реакции автосплайсинга, которая приводит к образованию продукта с некорректными формой, размерами и функциональными свойствами. Вестерн-блоттинг является удобным методом визуализации белок-белковых взаимодействий интеинов при использовании антител, способных распознавать предсплайсинговые молекулы.

После подтверждения работоспособности интеинового сплайсинга на уровне сборки белка с помощью антител и восстановления флуоресцентных свойств GFP с помощью микроскопии и цитометрии плазмидные конструкции могут быть использованы при получении AAV для раздельной доставки GFP.

3.7.2. Трансдукция клеточных линий с помощью AAV5 GFP Npu DnaE

Для демонстрации эффективной доставки разделенного GFP в клетки разного происхождения использовали AAV, поскольку не все клеточные линии могут быть эффективно трансфицированы с помощью полиэтиленимина. Для трансдукции были выбраны следующие клетки человека из биобанка лаборатории:

HEK293 – линия здоровых клеток эмбриональной почки человека, которая обладает высокой восприимчивостью к трансфекции и трансдукции большинством серотипов AAV. Использована в качестве положительного контроля.

ARPE19 – линия здоровых клеток пигментного эпителия сетчатки глаза человека, которая используется в качестве модельного объекта для тестирования офтальмологических препаратов. Обладает большей селективностью по отношению к серотипам AAV и хуже трансфицируется по сравнению с HEK293.

FbP13 — доступнеы для работы первичные фибробласты человека с мутацией в гене MT-ND4. Их использовали в качестве модели, максимально приближённой к условиям *in vivo*. Клетки обладают явно выраженной селективностью к трансдукции AAV и не трансфицируются с помощью полиэтиленимина. Использованы в качестве экспериментальной модели.

Для раздельной доставки GFP был использован AAV 5-го серотипа, который обладает широким тканевым тропизмом. Известно, что пятый серотип без модификаций способен трансдуцировать клетки глаза (пигментный эпителий, глиальные клетки), гепатоциты, миобласты и фибробласты. Инженерная модификация капсида с помощью пептидов и мутаций может повысить его специфичность по отношению к определённой ткани.

Методом классической транзиентной трёхплазмидной трансфекции был получен AAV пятого серотипа для доставки целого GFP. Для выбора оптимальной дозы трансдукции AAV5 GFP использовали наиболее восприимчивую к трансдукции клеточную линию HEK293 и вирусную нагрузку в диапазоне от $5*10^3$ до $2*10^5$ VG на клетку с шагом увеличения в два раза. Результат эксперимента по трансдукции HEK293 представлен на Рисунке 25.

Визуальный анализ микрофотографий (Рисунок 25 A–G) показал достоверное увеличение сигнала уже при дозе $5*10^3$ VG на клетку по сравнению с отрицательным контролем. Количественные метрики на основе анализа изображений (Рисунок 25 H) и интенсивности флуоресценции (Рисунок 25 I) позволяют визуализировать степенное увеличение сигнала, поскольку в них отсутствует ограничение изза абсолютных единиц измерений. Измерение эффективности в относительных единицах (Рисунок 25 J), с одной стороны, ограничивает диапазон измерений значениями от 0 до 100%, а с другой — позволяет представить дозозависимость в логарифмическом виде. Так намного заметнее проявляется эффект насыщения и выход на плато, который начинается с дозы в $1*10^5$ VG на клетку.

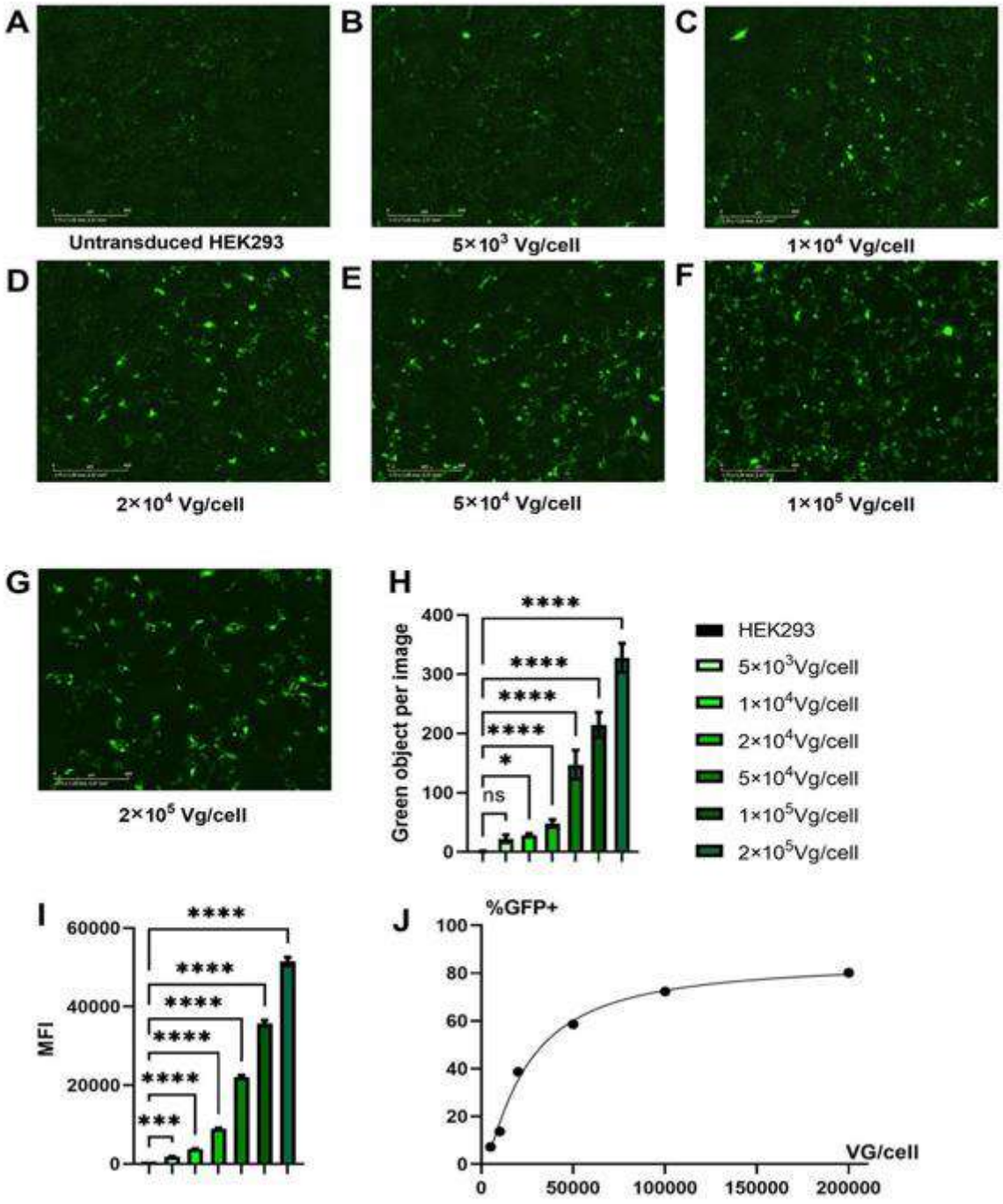


Рисунок 25 — Оценка эффективности трансдукции AAV5 GFP.

A–F — флуоресцентная микроскопия клеток HEK293 через 72 часа после трансдукции с возрастающими дозами AAV5 GFP.

H — анализ интенсивности зеленой флуоресценции с помощью IncuCyte.

I — анализ методом проточной цитометрии: значения MFI для каждого образца,

J - кривая зависимости «доза — эффект» для клеток GFP+

ns — статистически незначимо, p-значение > 0,05; (*) p-значение < 0,05;

(**) p-значение < 0,01; (***) p-значение < 0,001; (****) p-значение < 0,0001

Дальнейшее увеличение дозы в два раза приводит к увеличению эффективности всего на 6%, что сигнализирует об избыточной нагрузке на клетки и невозможности доставить большее количество ДНК-материала с помощью вирусных частиц. Поэтому целесообразность дальнейшего использования дозы $1*10^5$ VG на клетку определяется эффективностью трансдукции с помощью AAV пятого серотипа. Для дальнейших экспериментов с раздельной доставкой GFP использовали именно эту дозу для каждого вектора, поскольку препараты для реализации белкового транс-сплайсинга предполагают парную доставку частей в двух векторах, что существенно увеличивает нагрузку на клетки.

Для следующего эксперимента также использовали клеточную линию HEK293 при доставке разделённого и полноразмерного GFP с помощью AAV5. Сравнительный эксперимент на контрольной линии проводили для оценки эффективности сборки целого белка с помощью двух векторной доставкию. Результат сравнительного эксперимента, при котором использовалась доза $1*10^5$ VG на клетку каждого вируса, представлен на Рисунке 26.

По результатам эксперимента по флуоресценции можно визуально оценить эффективность доставки генетического материала *GFP* и сборки функционального белка (Рисунок 26 A–C). Количественный анализ выявил разницу в числе зелёных объектов в 3,5 раза между вариантами доставки целого и разделённого белка (Рисунке 26 D). По данным проточной цитометрии эффективность доставки сборки белка GFP составила 74%, а раздельная доставка и сборка GFP интеинами обеспечила достижение флуоресцентного сигнала до уровня 51% флуоресцентного белка (Рисунке 26 E). Разница в 1,5 раза по эффективности и в 3,5 раза по количеству зеленых объектов говорит о естественном ограничении раздельных систем доставки, которое надо учитывать при проектировании двухвекторных систем.

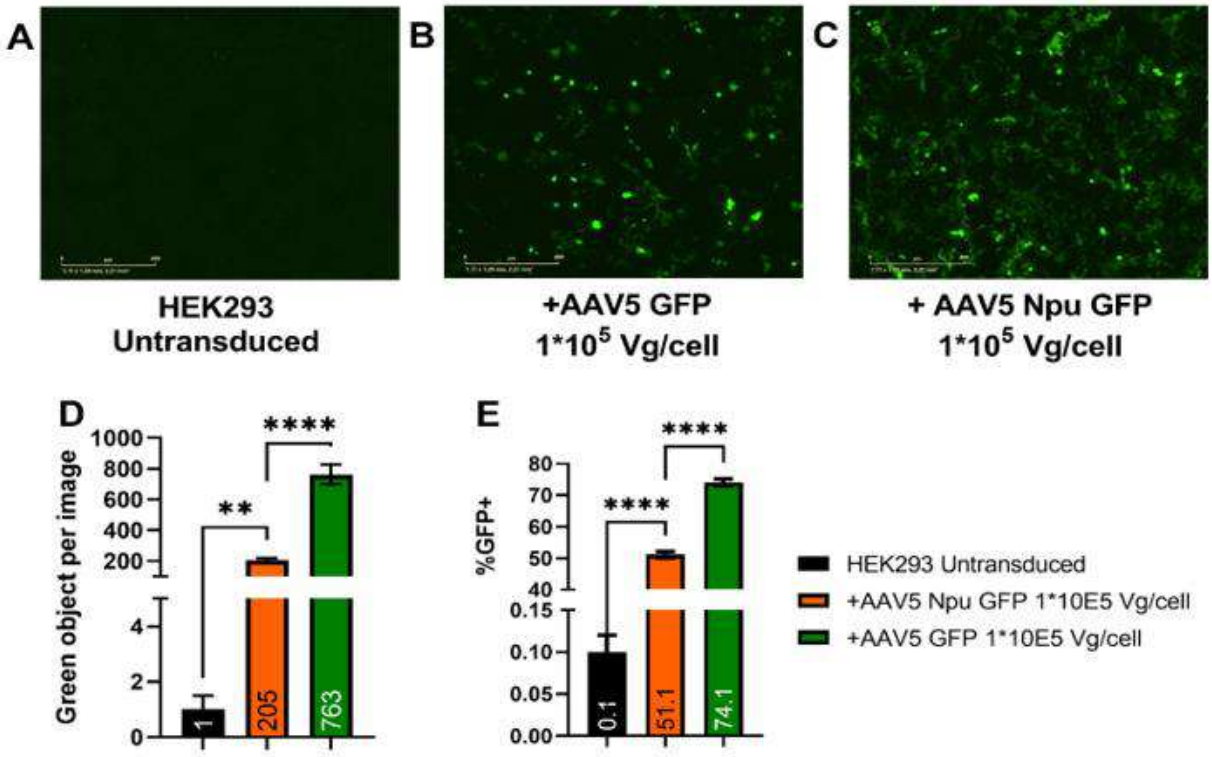


Рисунок 26 — Оценка сборки GFP после котрансдукции с AAV5, доставляющим фрагменты белка: A–C - клетки HEK293 через 72 часа после трансдукции. D - анализ интенсивности флуоресценции с помощью программы IncuCyte в клетках HEK293, E -Анализ методом проточной цитометрии клеток GFP+ и средней интенсивности флуоресценции в клетках HEK293.
ns — статистически незначимо, p-значение >0,05; (**) p-значение <0,01; (****) p-значение <0,0001

Рассмотрим, какими показателями обладает аналогичная система доставки разделенного и целого GFP в тканеспецифичных системах.

Для следующего эксперимента использовали линию ARPE19 как клеточную модель пигментного эпителия сетчатки. Эксперимент был проведён одновременно с экспериментом на HEK293 в идентичных условиях. Результат эксперимента представлен на Рисунке 27.

По качественным характеристикам результаты микроскопии (Рисунок 27 A–C) идентичны данным на Рисунке 27, однако визуально можно отметить меньшее количество ярко-зелёных объектов на фотографии, что проявляется в результатах количественного анализа на графике (Рисунок 27 D). При этом анализ единичных клеток методом проточной цитометрии показал, что разница в эффективности сборки GFP близка к результатам, полученным для клеток HEK293 (Рисунок 27 E).

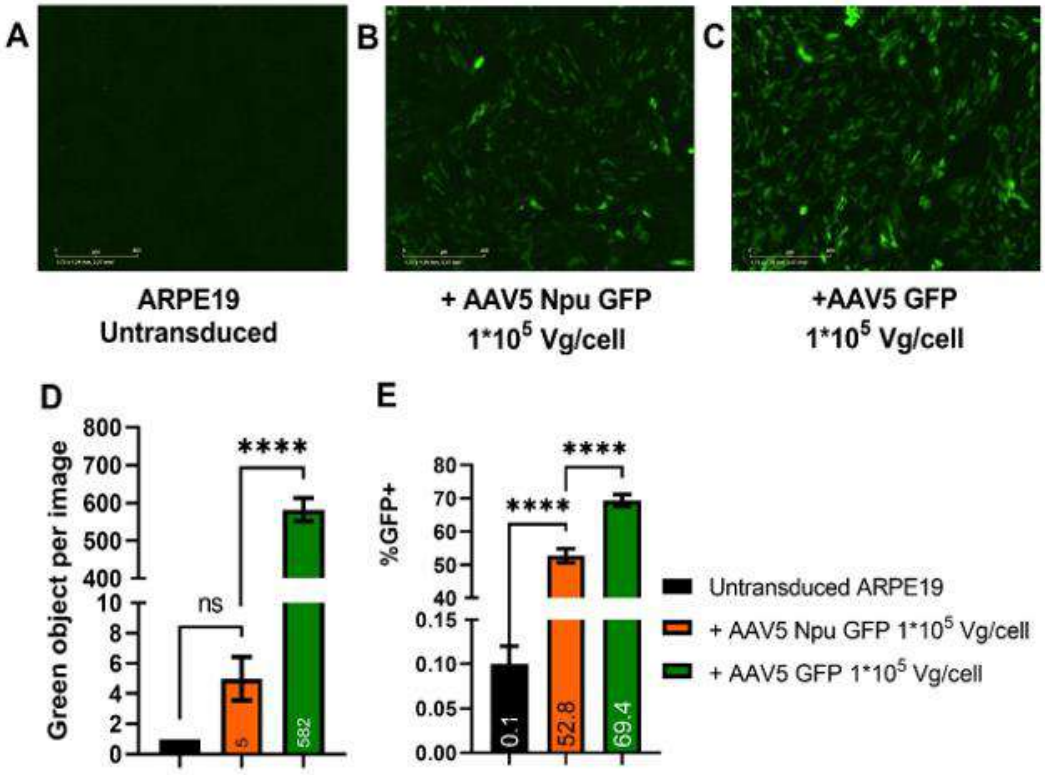


Рисунок 27 — Оценка сборки GFP после котрансдукции с AAV5, доставляющим фрагменты белка. A–C - клетки ARPE19 через 72 часа после трансдукции. D - анализ интенсивности флуоресценции с помощью IncuCyte в клетках ARPE19, E - анализ методом проточной цитометрии клеток GFP+ и средней интенсивности флуоресценции в клетках ARPE19. ns — статистически незначимо, p-значение >0,05; (**) p-значение <0,01; (****) p-значение <0,0001

Данный результат можно объяснить близостью эпителиальных свойств линий HEK293 и ARPE19, а также их принадлежностью к клеточным линиям и одинаковой восприимчивостью к трансдукции AAV5.

Для последнего эксперимента использовались первичные фибробласты, которые отличаются большей селективностью по отношению к AAV, ограниченным числом циклов деления и другими свойствами первичных культур. Эксперимент был проведён одновременно с экспериментами на HEK293 и ARPE19 в идентичных условиях. Результат эксперимента представлен на рисунке Рисунке 28.

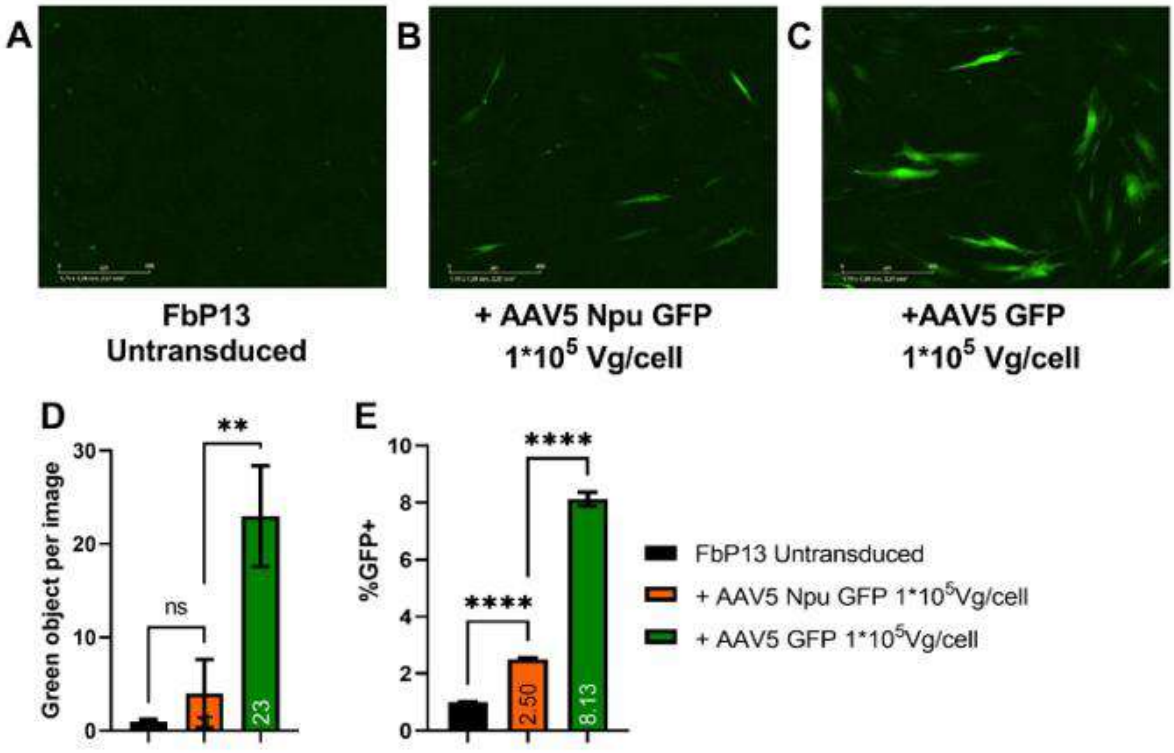


Рисунок 28 — Оценка сборки GFP после котрансдукции с AAV5, доставляющим фрагменты белка. A–C клетки FbP13 через 72 часа после трансдукции. D - анализ интенсивности флуоресценции с помощью IncuCyte в клетках FbP13. E - анализ методом проточной цитометрии клеток GFP+ и средней интенсивности флуоресценции в клетках FbP13. ns — статистически незначимо, p-значение >0,05; (**) p-значение <0,01; (****) p-значение <0,0001

Микроскопия позволяет качественно оценить трансдукцию единичных клеток. Фибробласты отличаются крупными размерами, что наглядно видно на фотографиях (Рисунок 28 B и C). Количественная обработка изображений показала небольшую, но достоверную разницу в количестве зеленых объектов в 1,5 раза между доставкой целого и разделенного GFP (Рисунок 28 D). При анализе данных проточной цитометрии наблюдается различие в 3 раза по эффективности доставки и сборки GFP (Рисунок 28 E). При этом значения эффективности были крайне низки: 2,5% для сборки и 8,3% для доставки целого GFP. Для увеличения эффективности трансдукции клеток фибробластов может быть использован серотип с другим тропизмом, а также дополнительные реактивы в виде полибрена и других добавок.

Основной вывод, который касается испытаний трансдукции трёх клеточных линий, заключается в том, что доставка и сборка раздельной конструкции с интеинами происходит в каждом случае. Этот факт позволяет подтвердить широкий тропизм пятого серотипа AAV и использовать его для дальнейших экспериментов *in vivo* на сетчатке глаза. Вторым фактом является универсальность работы интеиновой системы под контролем CMV-промотора, которая позволяет синтезировать белки в широком спектре клеток. Общей также является разница в эффективности сборки и доставки целого белка в 1,5 раза. В случае клеточных линий показатель эффективности сборки составляет 50%, что является важным ограничением при использовании двухвекторных систем. В случае первичных клеток ограничением становится избирательная восприимчивость к конкретному серотипу AAV, которую необходимо проверять отдельно для каждого типа органов. Полученные выводы будут использованы в экспериментах *in vivo* на тканях глаза.

3.8. Получение и тестирование препарата для терапии синдрома Штаргардта

3.8.1. Дизайн генетической конструкции с ABCA4

Дальнейшая серия экспериментов посвящена разработке препарата для терапии одной из форм наследственной слепоты на примере синдрома Штаргардта. Раздел объединяет этапы получения конструкций и вирусных частиц, тестирования *in vitro* и тестирования *in vivo* на здоровых мышах линии C57BL/6.

На первом этапе производили работу с генетическими конструкциями на основе вектора H46562, который содержит кодирующую последовательность гена ABCA4 (транскрипт NM_000350) длиной 6819 нуклеотидов. Оптимизация кодонов данной последовательности и синтез гена были произведены компанией Vazyme (Vazyme International LLC, Китай).

В первую очередь для создания препарата с разделённым геном ABCA4 был произведён анализ аминокислотной последовательности по алгоритму, представленному в разделе 3.4.2. Результат работы алгоритма представлен в Таблице 7. Отметим, что сайты разделения белка ABCA4 по данному алгоритму специфичны для интеиновой системы DnaE.

Таблица 7. Оценка сайтов для разделения молекулы белка ABCA4 по 6-символьным последовательностям.

Позиция в рейтинге	Последовательность сайта разделения	Оценка сайта по баллам	Позиция N-1/C+1 аминокислоты	Использован ранее в литературе
1	ALSCAN	1797	61,18	-
2	EYPCGN	1455	60,47	-
3 позиции пропущены				
6	LPECPE	1502	57,20	-
7	GNRCLK	1444	56,79	-
8	GIFCNV	75	56,56	-
9	EGTCSC	1177	55,60	[58]
2 позиции пропущены				
12	RLYCSG	1140	54,46	[58]
13	LGMCPQ	1008	54,32	-
14	SPSCRC	1488	53,95	-
4 позиции пропущены				
19	STTCPA	1188	53,47	-
2 позиции пропущены				
22	LKNCFG	1150	53,33	[58]
2 позиции пропущены				
31	TCSCSS	1179	52,47	-
Остальные 12 позиции с цистеином на N+1 пропущены				

По результатам анализа последовательности было обнаружено множество сайтов, содержащих цистеин на N+1 положении, что сразу делает рейтинг лучших довольно протяженным (43 позиции). Для сокращения позиций мы в списке пропустили те, в которых наблюдали явную диспропорцию в получаемых фрагментах. Также используем ограничение для каждой из частей в 1180 аминокислот (соответствует 3,5 тысячам пар оснований), которые являются ограничением для кодирующей ДНК в одном AAV векторе.

Таким образом список сократился до 6 позиций, которые выделены цветом в таблице 7. Зеленый цвет соответствует вариантам, которые были проверены в статье Li [58]. Среди них вариант 1150 демонстрировал наилучший результат, поскольку позволяет получить максимальное количество сплайсированного продукта ABCA4, в то время как 1140 и 1177 отличаются смещением пропорции в сторону наработки несплайсированного N фрагмента с меньшим количеством целевого белка. С помощью алгоритма был дополнительно найден следующий цистеин на сайте 1179, однако с наибольшей вероятностью эффективность разделения будет аналогична 1177. Остальные варианты с желтым цветом 1008 и 1188 смещают пропорцию и могут перегрузить один из векторов, но могут потенциально быть использованы для проверки.

Таким образом в результате расчетов и анализа уже проверенных вариантов было принято решение использовать сайт 1150, который гарантированно обеспечит эффективную сборку белка ABCA4. Именно этот сайт был далее использован для получения генетических конструкций с разделнным *ABCA4*. Схематично дизайн конструкций представлен на Рисунке 29.

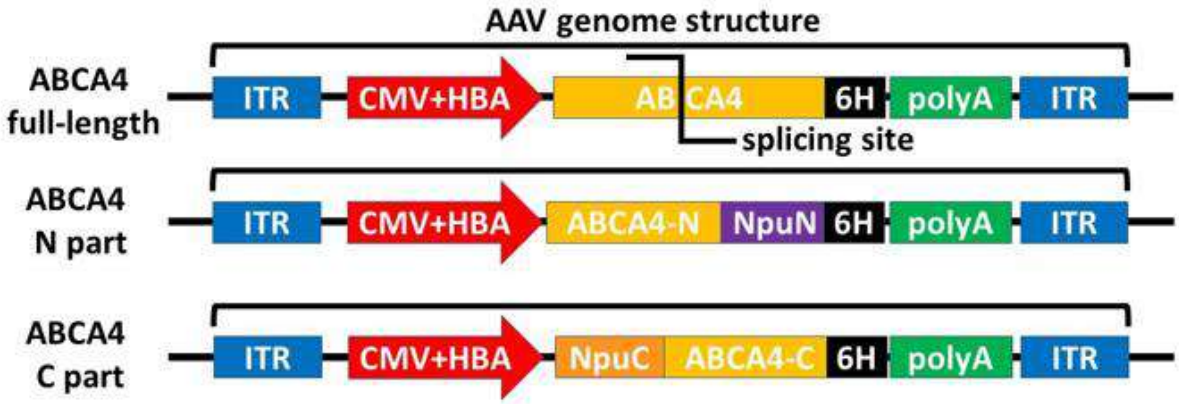


Рисунок 29 — Схема генетических конструкций для экспрессии полноразмерного ABCA4 и его фрагментов, предназначенных для восстановления белка посредством транс-сплайсинга с участием интеина NpuDnaE.
Сверху — контрольная конструкция, кодирующая полноразмерный ABCA4; отмечен сайт разделения белка, соответствующий позиции 1150.
В центре — конструкция, кодирующая N-концевой фрагмент ABCA4, соединённый с N-концевой частью интеина (NpuN).
Снизу — конструкция, кодирующая C-концевую часть интеина (NpuC), соединённую с C-концевым фрагментом ABCA4.
Обозначения: ITR — инвертированные концевые повторы AAV;
CMV+HBA — регуляторный блок экспрессии; 6H — шестигистидиновая метка;
polyA — сигнал полиаденилирования

3.8.2. Вестерн-блот-анализ ABCA4 после трансфекции клеток HEK293

Итоговый набор плазмид ABCA4_1150_Npu N+C использовали для трансфекции клеток HEK293, чтобы визуализировать результат сборки с помощью интеинового транс-сплайсинга. Данные вестерн-блот-анализа представлены на Рисунке 30.

При окрашивании с помощью антител к GAPDH наблюдали одинаковый уровень эндогенного белка «домашнего хозяйства», что говорит о выравнивании образцов по суммарной белковой массе. При окрашивании с помощью антител anti-6-HIS полос не наблюдали, что, возможно, говорит о недоступности сайта метки для антител из-за конформации экстеиновых частей. При окрашивании с помощью антител к ABCA4 наблюдали одну полосу на уровне 260 кДа в варианте трансфекции с помощью двух плазмид для сборки, что совпадает с теоретической массой белка ABCA4. При этом полученные результаты не совпали с данными статьи Li [58] при использовании антител anti-ABCA4 3F4 (Santa Cruz, sc-65672).

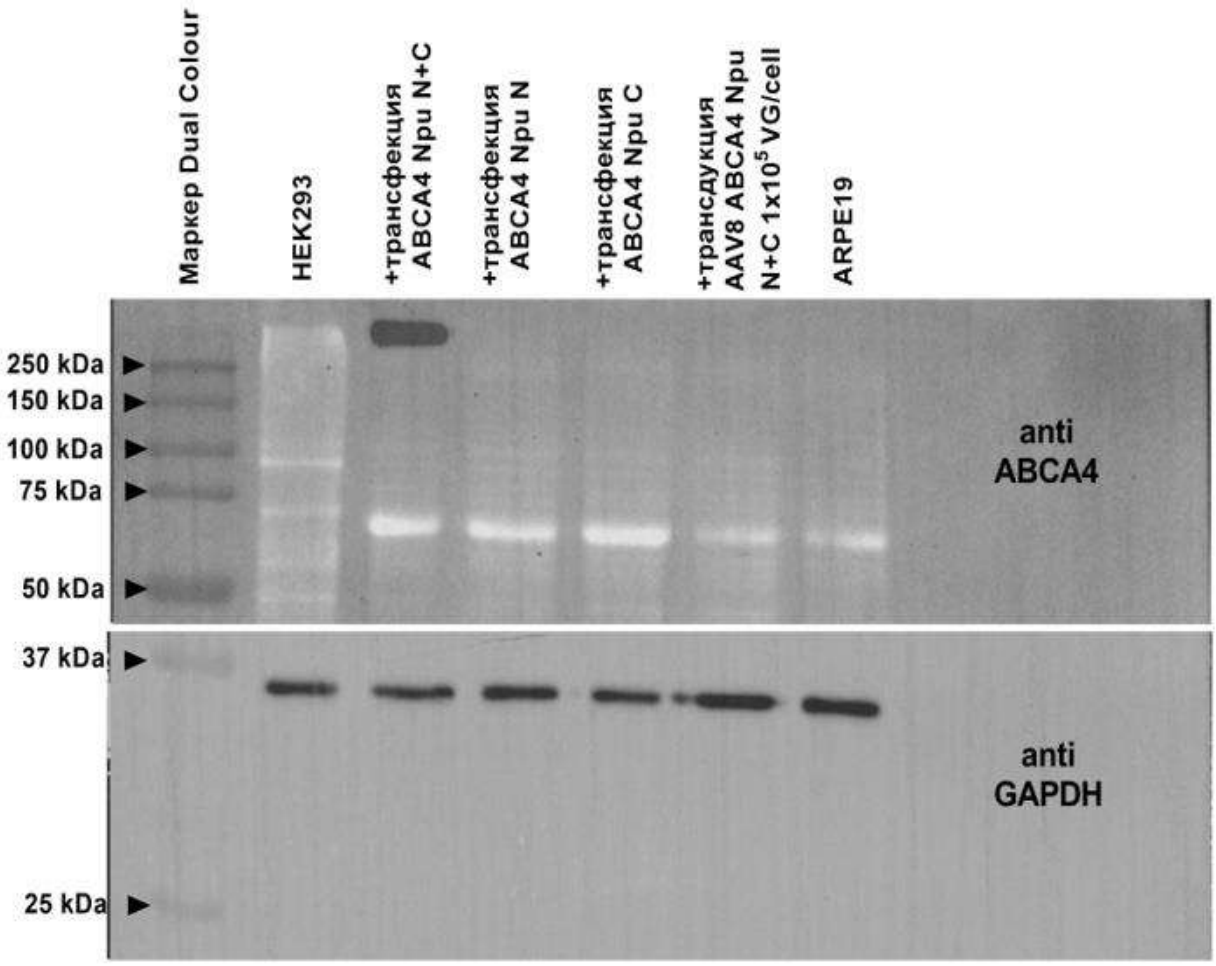


Рисунок 30 — Вестерн-блот-анализ ABCA4 в клеточных лизатах.
Верхняя панель — выявление ABCA4 антителами anti-ABCA4;
нижняя панель — выявление GAPDH антителами anti-GAPDH (контроль загрузки).
Дорожки слева направо: 1 — маркер молекулярной массы Dual Colour;
2 — нетрансфицированные клетки HEK293;
3 — клетки HEK293 после котрансфекции конструкциями ABCA4 Npu N и ABCA4 Npu C;
4 и 5 — клетки HEK293 после трансфекции только N-концевой или только C-концевой конструкцией соответственно;
6 — клетки HEK293 после трансдукции AAV8 ABCA4 Npu N+C при указанной на рисунке дозе $1*10^5$~VG/клетку;
7 — клетки ARPE19. N и C обозначают конструкции, кодирующие соответствующие фрагменты ABCA4, соединённые с частями интеина NpuDnaE.
На дорожке 3 верхней панели выявляется полоса в области ожидаемой молекулярной массы ABCA4 (около 260~кДа). Положения маркерных белков и их молекулярные массы указаны слева.

Антитела данного производителя отличаются распознаванием полной формы белка ABCA4 размером 220 кДа, в то время как антитела anti-ABCA4 5B4 (Merck, MABN2440) распознают эпитоп на N-конце белка длиной в 250 аминокислот. Поэтому антитела Merck MABN2440 способны распознавать как целую молекулу, так и её N-концевую часть, а антитела Santa Cruz sc-65672 распознают только полную молекулу.

Отдельно отметим отсутствие белка ABCA4 в клетках как HEK293, так и ARPE19, что говорит об отсутствии возможности использовать данные клеточные линии в качестве положительного контроля. Единственной коммерчески доступной линией, которая может быть использована для клеточных тестов офтальмологических мишеней, является линия иммортализованных фоторецепторов мыши 661W. Из-за логистических и санкционных ограничений данная линия не могла быть доставлена в Российскую Федерацию, и дальнейший эксперимент целесообразнее всего было проводить на животных.

3.8.3. Тестирование тропизма AAV к тканям глаза *in vivo*

В качестве модели для доставки генетического материала были использованы здоровые мыши линии C57BL/6. На первом этапе работы с животными отрабатывали условия эффективной доставки материала с помощью векторов, экспрессирующих ген, кодирующий *GFP*.

В первую очередь сравнивали селективность трансдукции AAV разных серотипов по отношению к тканям глаза, в частности к сетчатке, которая состоит из разных функциональных слоёв клеток разного происхождения. Ранее на примере линии ARPE19 отмечали способность AAV5 эффективно трансдуцировать клетки эпителия сетчатки, однако эффективность доставки на уровне целого глаза необходимо было проверить. В некоторых работах коллег ранее отмечали способность AAV шестого серотипа трансдуцировать первичные стволовые клетки, фибробласты, адипоциты и другие типы клеток в условиях, максимально приближённых к условиям *in vivo*, однако на сетчатке данный серотип не проверяли.

По многочисленным литературным данным, для создания офтальмологических препаратов используется AAV восьмого серотипа, что было учтено в следующем эксперименте при сравнении трансдуцирующей способности серотипов. При каждой инъекции использовали интравитреальный метод введения, а доза составляла 1×10^8 VG на глаз в 1 мкл раствора.

После инъекций животных содержали 14 дней, а затем проводили хирургическое разделение глаза на отделы с последующим анализом зелёных объектов в ткани на флуоресцентном микроскопе с минимальной проводкой и фиксацией. Результат качественного анализа представлен на Рисунке 31.

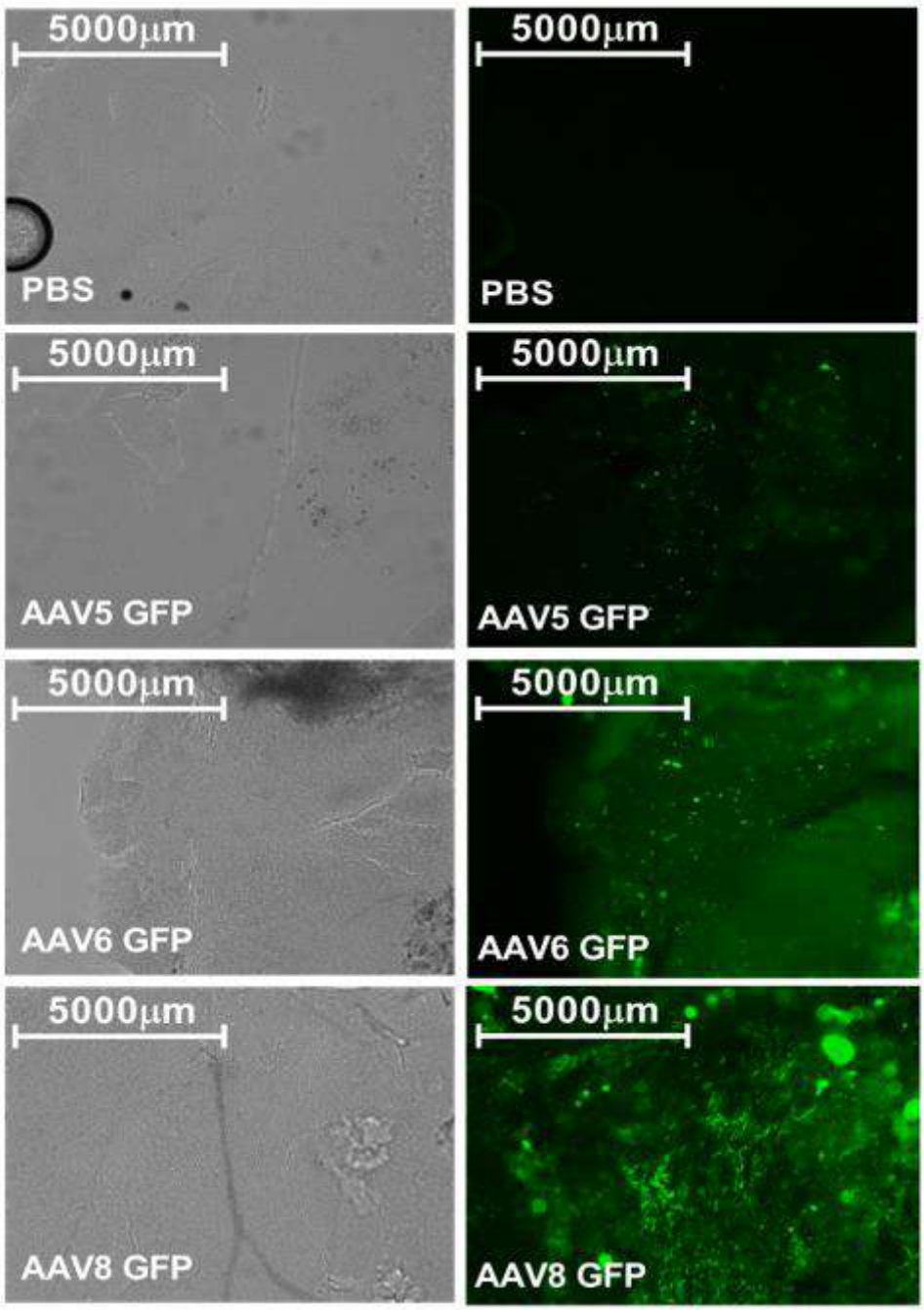


Рисунок 31 — Репрезентативные изображения в светлом поле (слева) и изображения флуоресценции GFP (справа) целых препаратов сетчатки после интравитреального введения PBS или векторов AAV5 GFP, AAV6 GFP и AAV8 GFP

Контрольный образец PBS показывал только фоновую флуоресценцию, тогда как сетчатки, обработанные вектором, демонстрируют серотипзависимые особенности распределения GFP-положительных областей и клеток.

Визуализация целых препаратов сетчатки с минимальной фиксацией удобно использовать для быстрой и качественной оценки трансдуцирующей способности серотипов AAV, что уменьшает количество вариантов на следующих этапах анализа.

В сетчатках, обработанных PBS, наблюдали фоновую флуоресценцию при отсутствии ярких флуоресцентных точек. AAV5 GFP приводил к появлению относительно редкого сигнала GFP, ограниченного несколькими областями сетчатки. AAV6 GFP при этом обеспечивал более широкий и интенсивный сигнал, хотя распределение оставалось гетерогенным, с кластерами GFP-положительных структур, разделёнными областями с более слабой экспрессии. AAV8 GFP приводил к появлению наиболее сильного и обширного сигнала флуоресценции в анализируемых полях зрения, включая области с высокой плотностью GFP-положительных структур. Неравномерный сигнал, наблюдаемый во всех группах векторов, указывает на то, что интравитреальное введение привело к региональной, а не равномерной трансдукции клеток сетчатки.

Поскольку флуоресценция целых препаратов не позволяет установить, какие слои сетчатки или типы клеток были трансдуцированы, эти наблюдения представли собой качественный скрининг, за которым последовал послойный гистологический анализ. Короткий срок выполнения процедуры скрининга (2 часа) позволил провести сравнение серотипов AAV в тот же день и быстро выбрать AAV восьмого серотипа в качестве наиболее эффективного трандуцирующего вектора.

3.8.4. Тестирование платформ доставки *in vivo*

Для сопоставления платформ доставки *GFP* были исследованы срезы сетчатки после интравитреального и субретинального введения липидных наночастиц (LNP) и векторов AAV8 GFP, поскольку к невирусным платформам наблюдается особенный интерес из-за возможности доставлять мРНК без ограничений на размеры кодирующей ДНК. Поскольку для терапии *ABCA4* использовали pH-чувствительные наночастицы, то использование липидных наночастиц актуально для сравненения с вирусной платформой в ходе разработи препарата. Контролями служили образцы после введения PBS соответствующим способом. Распределение флуоресценции оценивали по отдельным каналам и совмещённым изображениям (Рисунке 32).

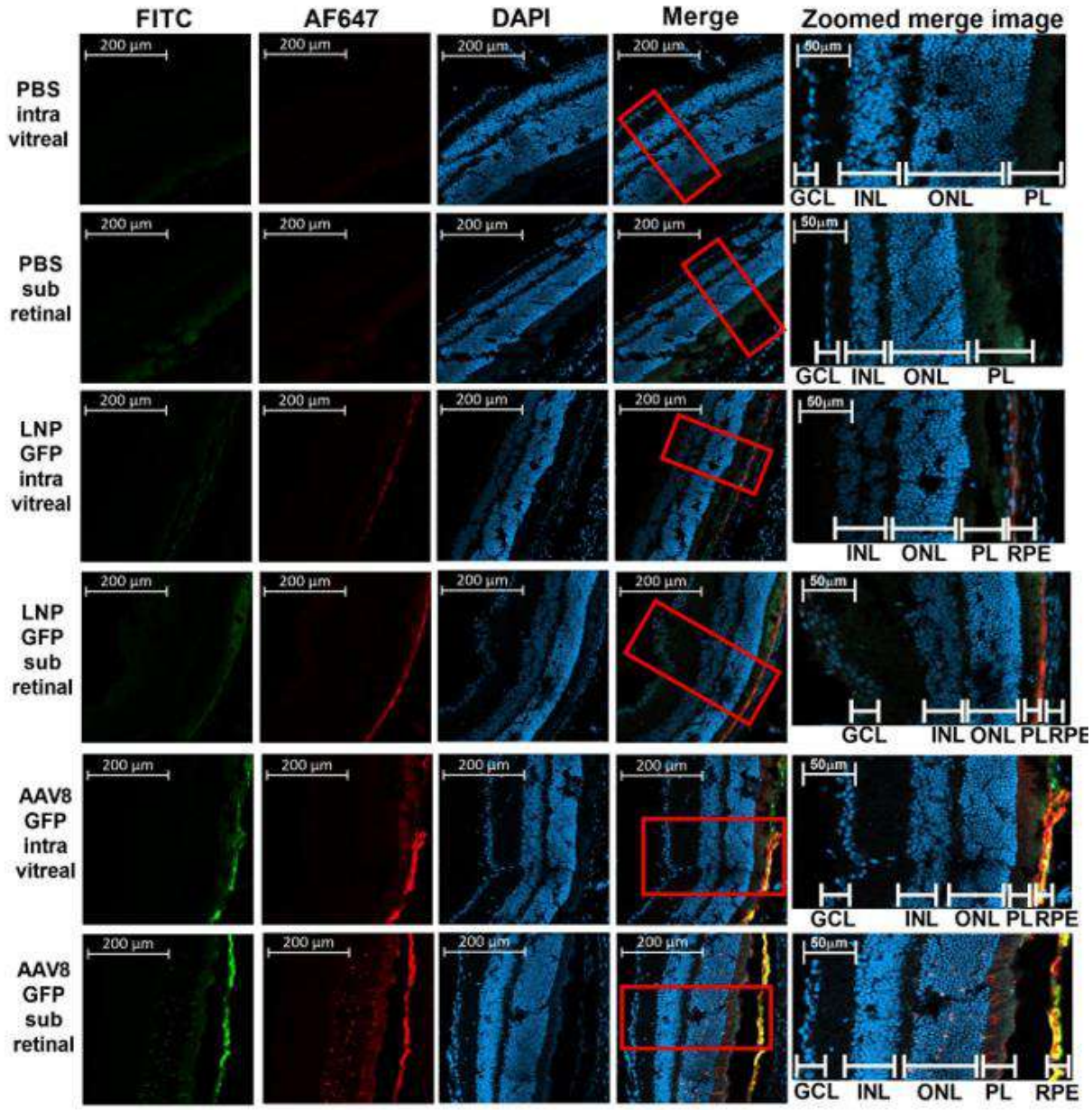


Рисунок 32 — Сравнение распределения флуоресцентного сигнала в срезах сетчатки при использовании LNP GFP и AAV8 GFP с разными способами введения. Строки сверху вниз: PBS, интравитреальное введение; PBS, субретинальное введение; LNP GFP, интравитреальное введение; LNP GFP, субретинальное введение; AAV8 GFP, интравитреальное введение; AAV8 GFP, субретинальное введение. Столбцы слева направо: FITC (зелёный канал), AF647 (красный канал), DAPI (ядра клеток,синий канал), совмещение каналов и увеличенные фрагменты областей, выделенных красными рамками. Сокращения: GCL — слой ганглиозных клеток; INL — внутренний ядерный слой; ONL — наружный ядерный слой; PL — наружные сегменты, RPE — пигментный эпителий сетчатки. Стрелками отмечены отдельные клетки RPE, наружные сегменты фоторецепторов и фоторецепторы согласно обозначениям на изображении. Масштабные отрезки: 200 мкм на обзорных изображениях и 50 мкм на увеличенных фрагментах.

В контрольных образцах PBS наблюдали слабый зелёный сигнал, преимущественно в наружных отделах сетчатки, при низкой интенсивности красного канала. После введения LNP GFP при обоих способах доставки красная флуоресценция была заметна главным образом в области пигментного эпителия. На увеличенных фрагментах выделены отдельные клетки RPE; выраженного распространения сигнала на фоторецепторы, сопоставимого с показанным для AAV8, в этих полях зрения не наблюдали.

В образцах AAV8 GFP зелёный и красный сигналы были визуально более выраженными, чем в представленных образцах LNP GFP и PBS. При интравитреальном введении AAV8 яркие участки пространственного совмещения каналов располагались в области RPE и рядом с наружными сегментами фоторецепторов. При субретинальном введении, помимо сигнала в области RPE, на увеличенном фрагменте видны красные флуоресцентные структуры, обозначенные на изображении как фоторецепторы. Таким образом, наблюдаемая картина различалась не только по выраженности сигнала, но и по его распределению относительно структур сетчатки.

Представленные изображения подтверждают целесообразность дальнейшего использования AAV8 как вектора для доставки экспрессионных кассет в наружные отделы сетчатки. При этом сравнение носило скорее качественный характер: по отдельным полям зрения нельзя было определить долю клеток с экспрессией GFP или установить количественное преимущество одного способа введения. Для такого вывода в дальнейшем необходим дополнительный анализ сопоставимых областей срезов при одинаковых настройках регистрации, с учётом фоновой флуоресценции и вариабельности между животными.

3.8.5. Анализ экспрессии *ABCA4 in vivo*

Для оценки распределения сигнала ABCA4 в сетчатке после введения парных конструкций AAV8 ABCA4 N+C использовали: контроль PBS без антител anti-ABCA4, контроль PBS с антителами anti-ABCA4 и опытный образец AAV8 ABCA4 N+C с антителами anti-ABCA4. Анализ отдельных каналов и их совмещённых изображений позволял сопоставить фоновую флуоресценцию с окрашиванием контрольного и опытного образцов (Рисунок 33).

В контрольном образце PBS без антител anti-ABCA4 наблюдалась зелёная флуоресценция, наиболее заметная в наружных отделах сетчатки, тогда как красный сигнал был слабым. Наличие зелёного сигнала в этом контроле показывает, что его нельзя непосредственно интерпретировать как результат выявления ABCA4.

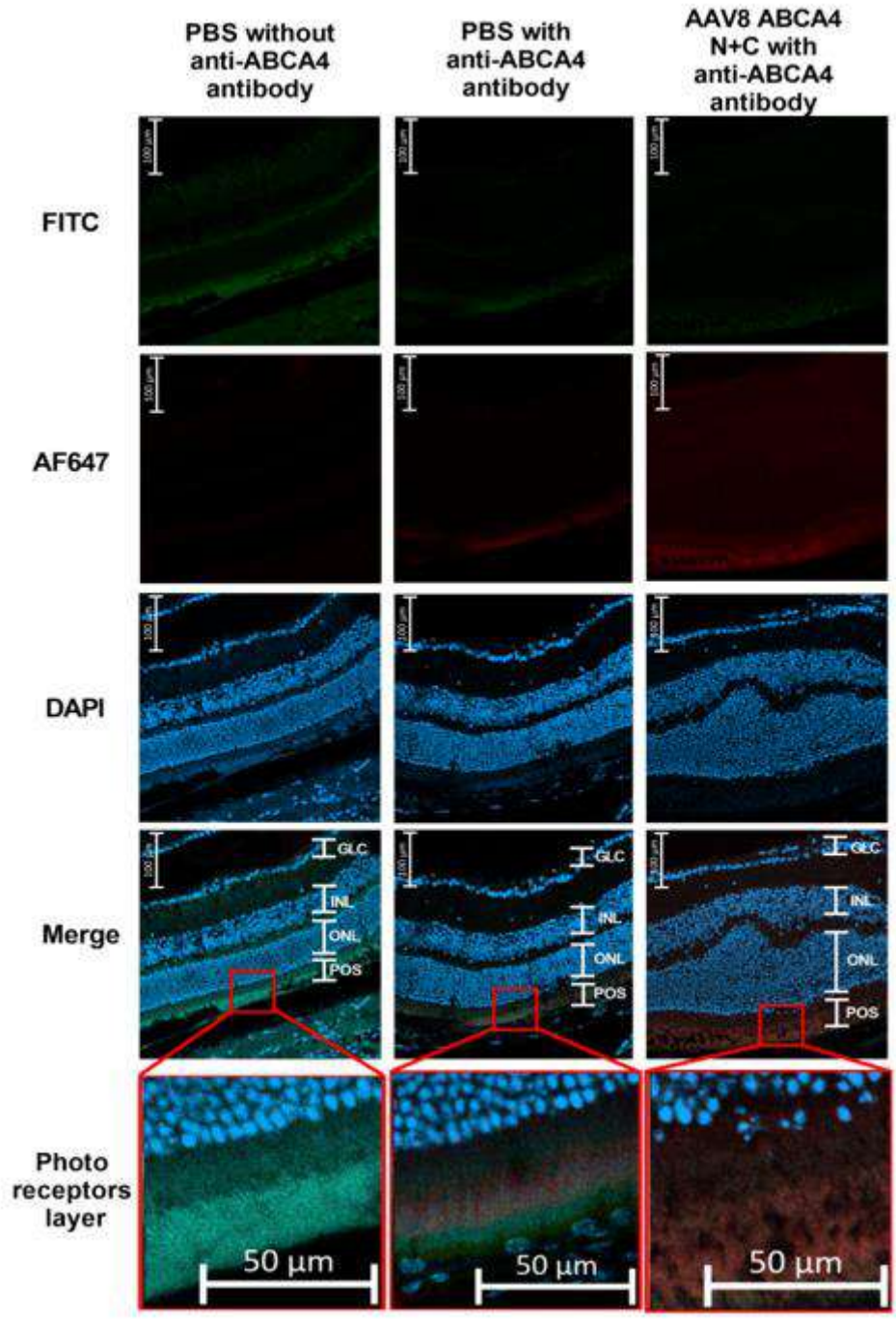


Рисунок 33 — Иммунофлуоресцентный анализ срезов сетчатки после введения AAV8 ABCA4 N+C. Столбцы слева направо:
PBS без антител anti-ABCA4; PBS с антителами anti-ABCA4;
AAV8 ABCA4 N+C с антителами anti-ABCA4.
Строки сверху вниз: FITC (зелёный канал), AF647 (красный канал),
DAPI (ядра клеток, синий канал), совмещение каналов и увеличенные фрагменты области фоторецепторов, выделенной красными рамками.
Обозначения слоёв на изображении: GLC — слой ганглиозных клеток; INL — внутренний ядерный слой; ONL — наружный ядерный слой; POS — наружные сегменты фоторецепторов. Масштабные отрезки: 100 мкм на обзорных изображениях и 50 мкм на увеличенных фрагментах.

При окрашивании образца PBS антителами anti-ABCA4 в красном канале становилась заметна полоса в области наружных сегментов фоторецепторов. Следовательно, при оценке опытного образца необходимо учитывать не только фон без антител, но и исходное окрашивание сетчатки в контроле с антителами, поскольку первичные антимышиные антитела ABCA4 детектируют эндогенный белок в тканях мыши.

В образце AAV8 ABCA4 N+C красный сигнал в представленном поле зрения был визуально более выраженным и распространялся на более широкую область, чем в контроле PBS с антителами anti-ABCA4. На совмещённом изображении и увеличенном фрагменте сигнал обнаруживали в области, обозначенной как наружные сегменты фоторецепторов. Окрашивание DAPI позволяло соотнести распределение сигнала с расположением ядерных слоёв сетчатки.

Таким образом, представленные изображения согласуются с усилением иммунореактивного сигнала ABCA4 после введения парных конструкций AAV8 ABCA4 N+C. Этот результат носит качественный характер и требует количественного сопоставления образцов при одинаковых условиях окрашивания и регистрации. Поскольку использовались здоровые мыши, необходимо учитывать возможный вклад эндогенного ABCA4 наряду с фоновым окрашиванием. По данным изображениям нельзя разграничить эндогенный белок и продукт, полученный после доставки конструкций, установить полноразмерность продукта транс-сплайсинга или подтвердить его функциональную активность. Поэтому наблюдаемое усиление сигнала следует рассматривать как основание для дальнейшей проверки восстановления ABCA4, а не как самостоятельное доказательство терапевтической эффективности. Для определения терапевтического эффекта необходимо использовать функциональный тест, например изменение уровня A2E метаболитов, которое может свидетельствовать доказательством восстановления функции ABCA4 в качестве флиппазы, обеспечивающей корректный транспорт данных веществ.

ЗАКЛЮЧЕНИЕ

Интеины являются универсальной и автономной системой для сборки белковых комплексов. Спектр задач, решаемых с помощью интеинов, охватывает ключевые области работы с белками: визуализацию и нехроматографическую очистку, раздельную сборку крупных белковых комплексов для обхода технологических ограничений [1,2]. Отдельно отметим растущую роль интеинов в области генной терапии, где данный инструмент применяется для сборки целых терапевтических белков после раздельной доставки и сборки генетических редакторов [31,32]. Основными мишенями для терапии крупных белков, где уже применяются интеины, являются гены *F8, DMD, ABCA4, CEP290*, однако с ростом роли редакторов данный список будет расширяться с каждым годом всё больше [51-59]. Тем не менее при разработке систем для транс-сплайсинга следует отметить несколько значимых аспектов.

Во-первых, сплайсинг интеинов зависит от состава аминокислот в области каталитических доменов, к которым относятся три аминокислоты на концах экстеинов и три первых аминокислоты в A и G-домене интеина. Полная консервативность аминокислот в каталитических доменах интеинов внутри семейства DnaE была продемонстрирована с помощью выравниваний аминокислотных последовательностей, в то время как состав экстеинов контекстно зависит от каждого отдельно взятого белка, используемого для сборки. Для интеинов группы DnaE экспериментально была получена таблица, которая определяет вклад каждой аминокислоты в реакцию сплайсинга в зависимости от положения относительно места разделения. На основе данной таблицы был составлен алгоритм, с помощью которого были определены новые и известные сайты разделения GFP и ABCA4.

По результату сборки и разделения GFP в данной работе было показано, что не все варианты разделения одинаково эффективно восстанавливают структуру и функциональность целевой молекулы. По результату сборки и разделения ABCA4 в работе Li [58] установлено, что в зависимости от используемого сайта эффективность сборки целевой молекулы также может различаться внутри одного эксперимента. Предсказать работоспособность сборки возможно только экспериментально, поскольку многоэтапный процесс химических взаимодействий между предсплайсированными молекулами содержит слишком много переменных для полноценного компьютерного моделирования реакции.

Во-вторых, подробное исследование группы DnaE показало, что за счёт единства каталитических доменов возможно кросс-реактивное взаимодействие N- и C-частей.

С одной стороны, данное свойство не позволяет разделить большую молекулу более чем на три части с помощью одной интеиновой системы, с другой стороны, открывает

возможность исследования индивидуальных структурных особенностей интеинов, отличающих их между собой по скорости и эффективности взаимодействия каноничных и неканоничных пар. Впервые изучен результат транс-сплайсинга при использовании парных сочетаний парных сочетаний интеинов NpuN+SspC, NpuN+AvaC, NpuN+NpuC, которые оказались способны поддерживать эффективность сплайсинга на уровне 75%. Однако при этом было показано, что при использовании пары SspN+AvaC полностью блокирует реакцию сплайсинга. При этом комбинация SspN+NpuC сохраняет эффективность сборки на уровне 50%, а NpuC и AvaC отличаются по восьми аминокислотным позициям: C6, C9, C11, C12, C21, C27, C28, C32. С учётом несложной и известной структуры NpuC в дальнейшем можно произвести углублённый структурный анализ данной молекулы и изучить роль этих аминокислот, окружающих каталитические домены, на эффективность сплайсинга.

Для практического использования рекомендуется использовать интеин в канонической паре NpuN + NpuC, который позволяет получить эффективную сборку как после трансфекции (75–80%), так и после трансдукции (45–50%).

В-третьих, в ходе работы мы последовательно оптимизировали процесс транс-сплайсинга за счет выбора наиболее функционального сайта разделения молекулы GFP, подбора концентраций каждого из компонентов трансфекции, выбора наиболее быстрого и эффективного интеина в рамках одной ферментативной группы. Итогом оптимизации трансфекции стало увеличение эффективности трансфекции и сборки целевой молекулы от 7% на старте (сайт GFP 69/70, условия трансфекции: количество суммарной ДНК — 2 мкг, соотношение PEI:ДНК — 1:1, тип трансфекции — прямая, интеин SspDnaE) до 80% (сайт GFP 69/70, условия трансфекции: количество суммарной ДНК — 2 мкг, соотношение PEI:ДНК — 5:1, тип трансфекции — обратная, интеин NpuDnaE). Интересно, что использование этого протокола с последующей поставкой вестерн-блота (после нормализации сигнала) приводило к диспропорции синтеза N- и C-частей предсплайсированных частей GFP_Npu при одинаковых условиях трансфекции. В результате диспропорции синтеза в ходе реакции C-фрагмент остаётся непрореагировавшим, что может быть скомпенсировано с помощью применения пропорции соотношения плазмид в общей смеси N:C как 2:1. Аналогичную гипотезу о диспропорции синтеза предсплайсированных молекул в случае ABCA4 подтвердить не удалось, поскольку антитела anti-ABCA4 распознают только полную форму белка, а anti-6-His не распознают данный эпитоп из-за сложной конформации молекулы.

В-четвёртых, на этапе трансдукции было показано, что эффективность сборки в 1,5–2 раза ниже, чем в случае доставки и синтеза целой молекулы напрямую. Тем не менее на

клеточных линиях HEK293 и ARPE19 удалось получить 50% GFP-положительных клеток после котрансдукции с AAV5 и разделённым GFP в дозе 10^5 VG/клетку. Трансдукция более тканеспецифичных клеток — фибробластов FbP13 и клеток сетчатки целого глаза мыши — показала качественный результат с низкой эффективностью, для увеличения которой целесообразно было заменить вектор доставки AAV другого серотипа. В случае сетчатки глаза наибольшей тропностью из природных серотипов обладает AAV8 [57,58], который и был использован далее для сравнения платформ доставки генетического материала для терапии наследственных ретинопатий.

При разработке препарата для терапии наследственных ретинопатий было проверено два платформенных решения по доставке генетического материала GFP: вирусная платформа доставки оцДНК на основе AAV8 и невирусная платформа доставки мРНК на основе липидных наночастиц (LNP), изготовленных по протоколу, аналогичному используемому компанией Moderna. В ходе эксперимента показано, что липидные наночастицы без тропных модификаций доставляют мРНК в слой пигментного эпителия менее эффективно, чем AAV8. Также вирусная система на основе AAV8 способна доставлять материал в фоторецепторы, что намного важнее для терапии большинства типов ретинопатий. Поэтому на текущий момент AAV8 является эффективным природным серотипом, который может быть использован для разработки препаратов от ретинопатий, однако стоит отметить, что также существуют инженерные варианты капсидов, которые также могут быть использованы для этой задачи [105].

В результате отработки технологии белкового транс-сплайсинга на примере GFP была получена технологическая схема для разработки препаратов с раздельной доставкой любого белка на две части. В результате данной работы создан прототип терапевтического препарата для раздельной доставки ABCA4. Раздление было осуществлено по сайту 1150 с помощью интеина NpuDnaE на основе AAV8. Сборку целевого белка подтвердили Вестерн-блотом и криогистологическим анализом тканей после введения препарата *in vivo*. Для проверки функциональной активности собранного белка в дальнейшем планируется использовать нокаутную мышиную линию C57BL/6JCya-Abca4em3/Cya, например, функциональный тест по уменьшению концентрации A2E-метаболитов и поведенческий тест на животной модели через полгода-год после введения препарата [58].

СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ

AAV – аденоассоциированные вирусы

ABCA – АТФ-зависимый мембранный транспортер

ABE – редактор оснований аденина

CBE – редактор оснований цитозина

DMD – мышечная дистрофия Дюшенна

F8/VIII – фактор свертывания крови VIII

FACS – сортировка клеток, активируемых флуоресценцией

FDA - управление по санитарному надзору за качеством пищевых продуктов и медикаментов США

GFP – зеленый флуоресцентный белок

АТФ – аденозинтрифосфат

ДНК – дезоксирибонуклеиновая кислота

кДНК – кодирующая ДНК

ПЦР – полимеразная цепная реакция

РНК – рибонуклеиновая кислота

ЯМР – ядерный магнитный резонанс

СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

Публикации в изданиях, входящих в наукометрические базы Web of Science и Scopus:

1. **Brovin, A.,** Minskaia, E., Sabantsev, M., Chuvpilo, S., & Karabelsky, A. Protein trans-splicing: optimization of intein-mediated GFP assembly as a model for the development of gene therapy //Frontiers in Bioengineering and Biotechnology. – 2024. – Т. 12. – С. 1488912.

2. Sabantsev, M. O., **Brovin, A. N.***, Gureev, M. A., Porozov, Y. B., Chuvpilo, S. A., & Karabelsky, A. V., Application of Rational Design and Molecular Metadynamics for the Estimation of Changes in Trans-Splicing Efficiency during the Mutagenesis of Ssp DnaE Intein //ACS Bio Med Chem Au 2025, Т. 5, № 4, С. 738–752

3. **Brovin, A. N.,** Sabantsev, M. O., Gureev, M. A., Chuvpilo, S. A., & Karabelsky, A. V., «Intein cross-reactivity enable modulation of trans-splicing efficiency» //Biological Communications. – 2025. – Т. 70. – №. 4. – С. С. 255-267

4. **Бровин А.Н.,** Сабанцев М.О., Игнатьева Е.В., Байбарин К.А., Карабельский А.В. Современные генетические технологии для терапии синдрома Штаргардта. Часть 1. Невирусные методы доставки терапевтических конструкций. Офтальмология. 2026;23(3):512-520. https://doi.org/10.18008/1816-5095-2026-3-512-520

Публикации в других изданиях, в сборниках трудов конференций:

1. Бровин А.Н., Карабельский А.В., Разработка тест системы для апробации методов доставки крупных генов с помощью аденоассоциированных вирусных векторов //Цитология. материалы viii молодежной школы-конференции по молекулярной биологии и генетическим технологиям института цитологии ран (11–14 октября 2022 г., институт Цитологии РАН, Санкт-Петербург) – 2022 – Т.64, №7, С. 611-780.

2. Бровин А.Н., Сабанцев М.О., Чувпило С.А., Минская Е.С., Егоров А.Д., Карабельский А.В Тест-система визуализации белкового транс-сплайсинга молекулы gfp с интеинами группы dnae для задач генной терапии наследственных заболеваний, ассоциированных с крупными генами // Биохимия человека: материалы всероссийской конференция с международным участием, 17–19 октября 2024 г – 2024 – С.102

СПИСОК ЛИТЕРАТУРЫ

1. Anastassov S., Filo M., Khammash M. Inteins: A Swiss army knife for syntheticbiology // Biotechnology Advances. – 2024. – С. 108349.
2. Nanda A. et al. Inteins in science: evolution to application // Microorganisms. –2020. – Т. 8. – № 12. – С. 2004.
3. База данных PubMed [Электронный ресурс]. URL: https://pubmed.ncbi.nlm.nih.gov/?term=inteins&sort=pubdate&sort_order=asc&size=100 (Дата обращения: 1.08.2025).
4. Shah N. H. et al. Ultrafast protein splicing is common among cyanobacterial split inteins: implications for protein engineering //Journal of the American Chemical Society. – 2012. – Т. 134. – №. 28. – С. 11338-11341.
5. Carvajal-Vallejos P. et al. Unprecedented rates and efficiencies revealed for newnatural split inteins from metagenomic sources // Journal of Biological Chemistry. – 2012. – Т. 287. – № 34. – С. 28686-28696.
6. Perler F. B. et al. Protein splicing elements: inteins and exteins–a definition ofterms and recommended nomenclature // Nucleic acids research. – 1994. – Т. 22. – № 7. – С. 1125.
7. Perler F. B. InBase: the intein database // Nucleic acids research. – 2002. – Т. 30. – № 1. – С. 383-384.
8. Bowman E.J., Tenney K., Bowman B.J. Isolation of genesen coding the Neurospora vacuolar ATPase. Analysis of vma-1 encoding the 67-kDa subunit reveals homology to other ATPases // Journal of Biological Chemistry. – 1988. – Т. 263. – № 28. – С. 13994-14001.
9. Gogarten J. P. et al. Inteins: structure, function, and evolution // Annual Reviewsin Microbiology. – 2002. – Т. 56. – № 1. – С. 263-287.
10. Belfort M. Back to basics: structure, function, evolution and application of homingendonucleases and inteins // Homing endonucleases and inteins. – Berlin, Heidelberg : Springer Berlin Heidelberg, 2005. – С. 1-10.
11. Pietrokovski S. Intein spread and extinction in evolution // TRENDS in Genetics.– 2001. – Т. 17. – № 8. – С. 465-472.
12. Starokadomskyy P. L. Protein splicing and its evolution in eukaryotes //Biopolymers and Cell. – 2010. – Т. 26. – № 1. – С. 5-12.
13. Cuesta S. A. et al. Comparative study of the nucleophilic attack step in theproteases catalytic activity: A theoretical study // Molecular Physics. – 2020. – Т. 118. – № 14. – С. e1705412.
14. Gregory M. T. et al. Multiple deprotonation paths of the nucleophile 3-OH in theDNA synthesis reaction // Proceedings of the National Academy of Sciences. – 2021. – Т. 118. – № 23. – С. e2103990118.
15. Xu M. Q., Perler F. B. The mechanism of protein splicing and its modulation bymutation // The EMBO journal. – 1996. – Т. 15. – № 19. – С. 5146-5153.
16. Starokadomskyy P. L. Protein splicing //Molecular Biology. – 2007. – Т. 41. – С. 278-293.
17. Perler F. B., Olsen G. J., Adam E. Compilation and analysis of intein sequences //Nucleic acids research. – 1997. – Т. 25. – №. 6. – С. 1087-1093
18. Pietrokovski S. Modular organization of inteins and C-terminal autocatalytic domains //Protein science. – 1998. – Т. 7. – №. 1. – С. 64-71.

19. Ellilä S., Jurvansuu J. M., Iwaï H. Evaluation and comparison of protein splicing by exogenous inteins with foreign exteins in Escherichia coli //Febs Letters. – 2011. – Т. 585. – №. 21. – С. 3471-3477.
20. Iwai, H., Züger, S., Jin, J., and Tam, P. H. 2006. Highly efficient protein trans-splicing by a naturally split DnaE intein from Nostoc punctiforme. FEBS Letters 580(7):1853–1858.
21. Wu H., Hu Z., Liu X.-Q. Protein trans-splicing by a split intein encoded in a split DnaE gene of Synechocystis sp. PCC6803 // Proceedings of the National Academy of Sciences. – 1998. – Vol. 95, № 16. – P. 9226–9231.
22. Tornabene P. et al. Intein-mediated protein trans-splicing expands adeno-associated virus transfer capacity in the retina // Science translational medicine. – 2019. – Т. 11. – № 492. – С. eaav4523.
23. Aranko A. S. et al. Intermolecular domain swapping induces intein-mediated protein alternative splicing // Nature Chemical Biology. – 2013. – Т. 9. – № 10. – С. 616-622
24. База данных RCSB PDB [Электронный ресурс]. URL: \url{https://www.rcsb.org} (Дата обращения: 1.08.2025).\par
25. Wu H., Hu Z., Liu X.-Q. Protein trans-splicing by a split intein encoded in a split DnaE gene of Synechocystis sp. PCC6803 // Proceedings of the National Academy of Sciences. – 1998. – Vol. 95, № 16. – P. 9226–9231. DOI: 10.1073/pnas.95.16.9226.
26. Oeemig J. S. et al. Solution structure of DnaE intein from Nostoc punctiforme: structural basis for the design of a new split intein suitable for site-specific chemical modification // FEBS Letters. – 2009. – Vol. 583, № 9. – P. 1451–1456. DOI: 10.1016/j.febslet.2009.03.058.
27. Shah N. H., Eryilmaz E., Cowburn D., Muir T. W. Naturally Split Inteins Assemble through a ``Capture and Collapse" Mechanism // Journal of the American Chemical Society. – 2013. – Vol. 135, № 49. – P. 18673–18681. DOI: 10.1021/ja4104364.
28. Cheriyan M. et al. Faster protein splicing with the Nostoc punctiforme DnaEintein using non-native extein residues // Journal of Biological Chemistry. – 2013. – Т. 288. – № 9. – С. 6202-6211.
29. Otomo, T., Ito, N., Kyogoku, Y., and Yamazaki, T. 1999. NMR observation of selected segments in a larger protein: Central- segment isotope labeling through intein-mediated ligation. Biochemistry 38(49):16040–16044. https://doi.org/10.1021/bi991902j
30. Iwai, H., Züger, S., Jin, J., and Tam, P. H. 2006. Highly efficient protein trans-splicing by a naturally split DnaE intein from Nostoc punctiforme. FEBS Letters 580(7):1853–1858. https://doi.org/10.1016/j.febslet.2006.02.045
31. Ray, D. M., Flood, J. R., and David, Y. 2023. Harnessing Split-Inteins as a Tool for the Selective Modification of Surface Receptors in Live Cells. ChemBioChem 24(3). https://doi.org/10.1002/cbic.202200487
32. Vila-Perelló M., Muir T. W. Biological applications of protein splicing //Cell. – 2010. – Т. 143. – №. 2. – С. 191-200.
33. Song, H., Meng, Q., and Liu, X. Q. 2012. Protein Trans-Splicing of an Atypical Split Intein Showing Structural Flexibility and Cross-Reactivity. PLoS ONE 7(9):7–12. https://doi.org/10.1371/journal.pone.0045355
34. Pritchett L. C. for the degree of Master of Science. – 2005.
35. Wu H., Hu Z., Liu X.-Q. Protein trans-splicing by a split intein encoded in a split DnaE gene of Synechocystis sp. PCC6803 // Proceedings of the National Academy of Sciences. – 1998. – Vol. 95, № 16. – P. 9226–9231. DOI: 10.1073/pnas.95.16.9226.

36. Kim S. B. et al. High-throughput sensing and noninvasive imaging of proteinnuclear transport by using reconstitution of split Renilla luciferase // Proceedings of the National Academy of Sciences. – 2004. – Т. 101. – № 32. – С. 11542-11547.
37. Ozawa T. et al. A genetic approach to identifying mitochondrial proteins // Naturebiotechnology. – 2003. – Т. 21. – № 3. – С. 287-293.
38. Yao Z. et al. A split intein and split luciferase-coupled system for detectingprotein-protein interactions // Molecular systems biology. – 2025. – Т. 21. – № 2. – С. 107-125.
39. Prabhala S. V., Gierach I., Wood D. W. The evolution of intein-based affinitymethods as reflected in 30 years of patent history // Frontiers in Molecular Biosciences. – 2022. – Т. 9. – С. 857566.
40. Comb, Donald G., et al. "Modification of protein by use of a controllableinterveining protein sequence."U.S. Patent No. 5,496,714. 5 Mar. 1996.
41. New England Biolabs Inc (2008). IMPACTTM-TWIN - Purification, Ligation andCyclization of Recombinant Proteins Using Self-Cleavable Affinity Tags. Ipswich, USA: New England Biolabs, 1–17.
42. Muir T. W. et al. Method of using split inteins, conjugates to generate proteinconjugates : пат. 10526401 США. – 2020.
43. Shi C. et al. A general purification platform for toxic proteins based on inteintrans-splicing // Applied microbiology and biotechnology. – 2014. – Т. 98. – № 22. – С. 9425-9435.
44. Han L. et al. Efficient generation of bispecific IgG antibodies by split inteinmediated protein trans-splicing system // Scientific Reports. – 2017. – Т. 7. – № 1. – С. 8360.
45. Minskaia E. et al. Viral vectors in gene replacement therapy // Biochemistry(Moscow). – 2023. – Т. 88. – № 12. – С. 2157-2178.
46. Wolff J. H., Mikkelsen J. G. Delivering genes with human immunodeficiencyvirus-derived vehicles: still state-of-the-art after 25 years // Journal of biomedical science. – 2022. – Т. 29. – № 1. – С. 79.
47. Rankovic V. et al. Overloaded adeno-associated virus as a novel gene therapeutictool for otoferlin-related deafness // Frontiers in Molecular Neuroscience. – 2021. – Т. 13. – С. 600051.
48. Kolesnik V. V. et al. Optimization strategies and advances in the research anddevelopment of AAV-based gene therapy to deliver large transgenes // Clinical and Translational Medicine. – 2024. – Т. 14. – № 3. – С. e1607.
49. Choi J. H. et al. Optimization of AAV expression cassettes to improve packagingcapacity and transgene expression in neurons // Molecular brain. – 2014. – Т. 7. – № 1. – С. 17.
50. Tornabene P., Trapani I. Can adeno-associated viral vectors deliver effectivelylarge genes? // Human gene therapy. – 2020. – Т. 31. – № 1-2. – С. 47-56.
51. Zhu F.X.et al. Protein trans-splicingbaseddual-vectordeliveryofthecoagulation factor VIII gene // Science China Life Sciences. – 2010. – Т. 53. – № 6. – С. 683-689.
52. Zhu F.X.et al. Inter-chaindisulfidebondimprovedproteintrans-splicingincreases plasma coagulation activity in C57BL/6 mice following portal vein FVIII gene delivery by dual vectors // Science China Life Sciences. – 2013. – Т. 56. – № 3. – С. 262-267.
53. Esposito F. et al. Liver gene therapy with intein-mediated F8 trans-splicingcorrects mouse haemophilia A // EMBO molecular medicine. – 2022. – Т. 14. – № 6. – С. e15199.

54. Padula A. et al. Full-length ATP7B reconstituted through protein trans-splicingcorrects Wilson disease in mice // Molecular Therapy Methods & Clinical Development. – 2022. – T. 26. – C. 495-504.
55. Li J. et al. Protein trans-splicing as a means for viral vector-mediated in vivo genetherapy // Human gene therapy. – 2008. – T. 19. – № 9. – C. 958-964.
56. Tasfaout H. et al. Split intein-mediated protein trans-splicing to express largedystrophins // Nature. – 2024. – T. 632. – № 8023. – C. 192-200.
57. Tornabene P. et al. Intein-mediated protein trans-splicing expands adenoassociated virus transfer capacity in the retina // Science translational medicine. – 2019. – T. 11. – № 492. – C. eaav4523.
58. Li R. et al. Split AAV8 mediated ABCA4 expression for gene therapy of mouse stargardt disease (STGD1) // Human Gene Therapy. – 2023. – T. 34. – № 13-14. – C. 616-628.
59. Tang H. et al. Hearing of Otof-deficient mice restored by trans-splicing of N- andC-terminal otoferlin // Human genetics. – 2023. – T. 142. – № 2. – C. 289-304.
60. Jinek M.et al. A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity // science. – 2012. – T. 337. – № 6096. – C. 816-821.
61. Anzalone A. V. et al. Search-and-replace genome editing without double-strandbreaks or donor DNA // Nature. – 2019. – T. 576. – № 7785. – C. 149-157.
62. Lee J., Kweon J., Kim Y. Emerging trends in prime editing for precision genomeediting // Experimental & Molecular Medicine. – 2025. – C. 1-11.
63. Yang W. et al. Progress of delivery methods for CRISPR-Cas9 // Expert Opinionon Drug Delivery. – 2022. – T. 19. – № 8. – C. 913-926.
64. Chen Y. et al. Development of highly efficient dual-AAV split adenosine baseeditor for in vivo gene therapy // Small Methods. – 2020. – T. 4. – № 9. – C. 2000309.
65. Levy J. M. et al. Cytosine and adenine base editing of the brain, liver, retina, heart and skeletal muscle of mice via adeno-associated viruses // Nature biomedical engineering. – 2020. – T. 4. – № 1. – C. 97-110
66. Zhi S. et al. Dual-AAV delivering split prime editor system for in vivo genomeediting // Molecular Therapy. – 2022. – T. 30. – № 1. – C. 283-294.
67. Zheng C. et al. A flexible split prime editor using truncated reverse transcriptaseimproves dual-AAV delivery in mouse liver // Molecular Therapy. – 2022. – T. 30. – № 3. – C. 1343-1351.
68. She K. et al. Dual-AAV split prime editor corrects the mutation and phenotype inmice with inherited retinal degeneration // Signal transduction and targeted therapy. – 2023. – T. 8. – № 1. – C. 57.
69. Davis J. R. et al. Efficient prime editing in mouse brain, liver and heart with dualAAVs // Nature biotechnology. – 2024. – T. 42. – № 2. – C. 253-264.
70. Yu W. et al. Optimized dual-AAV base editor delivery system with enhancedediting efficiency and virion production titer // Synthetic and Systems Biotechnology. – 2025. – T. 10. – № 3. – C. 697-706.
71. FDA News Release [Электронный ресурс]. URL: https://www.fda.gov/ne ws-events/press-announcements/fda-approves-first-gene-therapies-t reat-patients-sickle-cell-disease (Дата обращения: 1.08.2025).

72. Approved Cellular and Gene Therapy Products [Электронный ресурс]. URL: https://www.fda.gov/vaccines-blood-biologics/cellular-gene-the rapy-products/approved-cellular-and-gene-therapy-products (Дата обращения: 1.08.2025).
73. AlEissa M. M. et al. Advances in Precision Therapeutics and Gene Therapy Applications for Retinal Diseases: Impact and Future Directions // Genes. – 2025. – Т. 16. – № 7. – С. 847
74. Niederkorn J. Y. The eye sees eye to eye with the immune system: the 2019 proctor lecture // Investigative ophthalmology & visual science. – 2019. – Т. 60. – № 13. – С. 4489-4495.
75. Krajewska J.B., Waszczykowska A. Gene therapy strategies in ophthalmology—an overviewofcurrentdevelopmentsandfutureprospects//JournalofAppliedGenetics. – 2025. – С. 1-13.
76. Molday R.S. et al. Structure and function of ABCA4 and its role in the visual cycle and Stargardt macular degeneration // Prog. Retin. Eye Res. 2022. Vol. 89. P. 101036.
77. Xie T. et al. Structural basis of substrate recognition and translocation by human ABCA4 // Nat. Commun. 2021. Vol. 12, № 1. P. 3853.
78. Ng E.S.Y. et al. Membrane Attack Complex Mediates Retinal Pigment EpitheliumCell Death in Stargardt Macular Degeneration // Cells. 2022. Vol. 11, № 21. P. 3462. 85.
79. Yang B. et al. Exposure of A2E to blue light promotes ferroptosis in the retinal pigment epithelium // Cell. Mol. Biol. Lett. 2025. Vol. 30, № 1. P. 22.
80. Lenis T.L. et al. Expression of ABCA4 in the retinal pigment epithelium and itsimplications for Stargardt macular degeneration // Proc. Natl. Acad. Sci. Proceedings of the National Academy of Sciences, 2018. Vol. 115, № 47. P. E11120–E11127.
81. Tanna P. et al. Stargardt disease: clinical features, molecular genetics, animalmodels and therapeutic options // Br. J. Ophthalmol. BMJ Publishing Group Ltd, 2017. Vol. 101, № 1. P. 25–30.
82. Li C.H.Z. et al. Study of Late-Onset Stargardt Type 1 Disease: Characteristics,Genetics, and Progression // Ophthalmology. 2024. Vol. 131, № 1. P. 87–97.
83. Zolnikova I.V. et al. Stargardt disease-associated mutation spectrum of a RussianFederation cohort // Eur. J. Med. Genet. 2017. Vol. 60, № 2. P. 140–147.
84. Zolnikova I.V. et al. Clinical and Genetic Correlations of Inherital Retinal Diseasewith Mutations in the ABCA4 Gene by Patients of the Russian Population: 4 // Ophthalmol. Russ. 2021. Vol. 18, № 4. P. 897–907.
85. Kadyshev V.V. et al. Major Contribution of c.[1622T>C;3113C>T] Complex Allele and c.5882G>A Variant in ABCA4-Related Retinal Dystrophy in an Eastern European Population: 22 // Int. J. Mol. Sci. Multidisciplinary Digital Publishing Institute, 2023. Vol. 24, № 22. P. 16231.
86. Dugel P.U. et al. Phase ii, randomized, placebo-controlled, 90-day study ofemixustat hydrochloride in geographic atrophy associated with dry age-related macular degeneration // Retina Phila. Pa. 2015. Vol. 35, № 6. P. 1173–1183.
87. Kubota R.et al. Randomised study evaluating the pharmaco dynamic femixustat hydrochloride in subjects with macular atrophy secondary to Stargardt disease // Br. J. Ophthalmol. 2022. Vol. 106, № 3. P. 403–408.
88. Fujinami K. et al. Stargardt macular dystrophy and therapeutic approaches // Br.J. Ophthalmol. 2024. Vol. 108, № 4. P. 495–505.
89. Sun D. et al. Non-viral Gene Therapy for Stargardt Disease with ECO/pRHO-

ABCA4 Self-Assembled Nanoparticles // Mol. Ther. 2020. Vol. 28, № 1. P. 293–303.
90. Sun D. et al. Stable Retinoid Analogue Targeted Dual pH-Sensitive SmartLipid ECO/pDNA Nanoparticles for Specific Gene Delivery in the Retinal Pigment Epithelium // ACS Appl. Bio Mater. 2020. Vol. 3, № 5. P. 3078–3086.
91. Sun D. et al. Formulation and efficacy of ECO/pRHO-ABCA4-SV40 nanoparticles for nonviral gene therapy of Stargardt disease in a mouse model // J. Controlled Release. 2021. Vol. 330. P. 329–340.
92. Sun D. et al. Effective gene therapy of Stargardt disease with PEG-ECO/pGRK1-ABCA4-S/MARnanoparticles//Mol.Ther.-NucleicAcids.2022.Vol.29.P.823–835.
93. SunD.et al. Intravitreal Delivery of PEGylated-ECOPlasmid DNA Nanoparticles for Gene Therapy of Stargardt Disease // Pharm. Res. 2024. Vol. 41, № 4. P. 807–817.
94. Parker M.A. et al. Three-Year Safety Results of SAR422459 (EIAV-ABCA4) Gene Therapy in Patients With ABCA4-Associated Stargardt Disease: An OpenLabel Dose-Escalation Phase I/IIa Clinical Trial, Cohorts 1-5: Results of Gene Therapy in Patients With Stargardt Disease // Am. J. Ophthalmol. 2022. Vol. 240. P. 285–301.
95. VeonGen - ViGeneron Announces FDA Clearance of IND for Novel mRNA Trans-Splicing Gene Therapy VG801 to Treat Stargardt Disease and Other ABCA4Linked Retinal Dystrophies [Электронный ресурс]. URL: https://veongen.co m/press/vigeneron-announces-fda-clearance-of-ind-for-novel-mrna-t rans-splicing-gene-therapy-vg801-to-treat-stargardt-disease-and-o ther-abca4-linked-retinal-dystrophies/ (Дата обращения: 1.08.2025).
96. Kaltak M. et al. QR-1011 restores defective ABCA4 splicing caused by multiple severe ABCA4 variants underlying Stargardt disease // Sci. Rep. 2024. Vol. 14, № 1. P. 684.
97. Perler F. B. InBase: the intein database // Nucleic acids research. – 2002. – T.30. – № 1. – C. 383-384.
98. COBALT NCBI NIH [Электронный ресурс] URL: https://www.ncbi.nlm.nih.gov/tools/cobalt/cobalt.cgi (Дата обращения: 1.08.2025).
99. Benchlink [Электронный ресурс].URL:https://benchling.com/and2bro/ f/lib_50rYfweB-cftr/seq_akpdupg7-cftr-codon-optimized-1-1767/edit (Дата обращения: 1.08.2025).
100. DNA Works NIH [Электронный ресурс]. URL: https://hpcwebapps.cit .nih.gov/dnaworks/ (Дата обращения: 1.08.2025).
101. Young, L. Two-step total gene synthesis method. / L. Young, Q. Dong // Nucleicacids research. – 2004. – T. 32. – № 7. – C. 1-6.
102. Cell Biolabs [Электронный ресурс]. URL: https://www.cellbiolabs.co m/aav-5-cmv-expression-system (Дата обращения: 1.08.2025).
103. Addgene Bacterial Transformation Protocol [Электронный ресурс]. URL: ht tps://www.addgene.org/protocols/bacterial-transformation/ (Дата обращения: 1.08.2025).
104. Kunyk, D.; Vasileva, O.; Bespalov, M.; Dzharullaeva, A.; Mazunina, E.; Chubarova, A.; Seliverstov, M.; Smirnykh, D.; Shevyrev, D.; Shtraus, S.; et al. Dual ionizable lipids (ALC0315 and SM102) in a novel lipid nanoparticle formulation enhance in vivo delivery and timedependent expression of messenger RNA and self-amplifying RNA. ChemMedChem 2026, 21, e70335.
105. Gurtsieva, D., Minskaia, E., Zhuravleva, S., Subcheva, E., Sakhibgaraeva, E., Brovin, A., ... & Karabelsky, A. (2024). Engineered AAV2. 7m8 serotype shows significantly higher transduction efficiency of ARPE-19 and HEK293 cell lines compared to AAV5, AAV8 and AAV9 serotypes. Pharmaceutics, 16(1), 138.

ПРИЛОЖЕНИЕ А

Сопроводительная информация для получения генетических конструкций

Таблица А.1 — Праймеры для синтеза нуклеотидной последовательности интеинов

Набор праймеров для синтеза нуклеотидной последовательности	
SspDnaE	
Название праймера	**Последовательность праймера (5' → 3')**
Ssp_N_1_F	TGCCTTTCTTTCGGAACTGAGATCCTTACCGTTGAGTACGGACCACTTC
Ssp_N_2_R	AATTTCCTCAGAAACGATCTTACCAATAGGAAGTGGTCCGTACTCAACGGT
Ssp_N_3_F	CTATTGGTAAGATCGTTTCTGAGGAAATTAACTGCTCAGTGTACTCTGTTGATCC
Ssp_N_4_R	AGCCTGAGTGTAAACTCTTCCTTCTGGATCAACAGAGTACACTGAGCAGTT
Ssp_N_5_F	AGAAGGAAGAGTTTACACTCAGGCTATCGCACAATGGCACGATAGGG
Ssp_N_6_R	TCGTACTCGAGAACCTCTTGTTCACCCCTATCGTGCCATTGTGCGAT
Ssp_N_7_F	GTGAACAAGAGGTTCTCGAGTACGAGCTTGAAGATGGATCCGTTATTCGT
Ssp_N_8_R	AGTCAAGAATCTATGGTCAGAGGTAGCACGAATAACGGATCCATCTTCAAGC
Ssp_N_9_F	GCTACCTCTGACCATAGATTCTTGACTACAGATTATCAGCTTCTCGCTATCGA
Ssp_N_10_R	AGATCAAGTTGCCTAGCAAAGATTTCCTCGATAGCGAGAAGCTGATAATCTGT
Ssp_N_11_R	TGTGGTTGTCAAGAGCCTCTTCTGTCTGCTTGATGTTCTCCAAAGTAAGG
Ssp_N_12_F	CAGAAGAGGCTCTTGACAACCACAGACTTCCATTCCCTTTGCTCGATG
Ssp_N_13_R	AATTATAAGCTTTTACTTGATGGTTCCAGCATCGAGCAAAGGGAATGGAAGTC
Ssp_N_14_F	GGAAATCTTTGCTAGGCAACTTGATCTCCTTACTTTGGAGAACATCAAGCAGA
BamHI_Ssp_C_1_F	AATTTA***GGATCC***ATGGTTAAGGTGATTGGAAGACGTTCT
Ssp_C_2_R	CCGATATCGAAGATCCTTTGAACACCAAGAGAACGTCTTCCAATCACCTT
Ssp_C_3_F	GTTCAAAGGATCTTCGATATCGGATTGCCACAAGACCACAACTTTCTTCT
Ssp_C_4_R	GCAATTGGCAGCGATGGCACCATTAGCGAGAAGAAAGTTGTGGTCTTGTG
Ssp Dna E GFP сайт разделения 47/48	
BamHI_GFP_1_F	AATTTA***GGATCC***ATGGTGAGCAAGGGCGAGGA
SSPN_GFP_47_R	CCGAAAGAAAGGCAGATGAACTTCAGGGTCAGCT
SSPN_GFP_48_F	CCCTGAAGTTCATCTGCCTTTCTTTCGGAACTGAGATCCTTACC
Ssp_N_R_HindIII	AATTAT***AAGCTT***TTACTTGATGGTTCCAGCATCGAGCA
BamHI_Ssp_C_1_F	AATTTA***GGATCC***ATGGTTAAGGTGATTGGAAGACGTTCT
SSPC_GFP_47_F	CCGAAAGAAAGGCAGATGAACTTCAGGGTCAGCT
SSPC_GFP_48_R	CGGTGGTGCAATTGGCAGCGATGGCACC
GFP_238_R_HindIII	AATTAT***AAGCTT***TTACTTGTACAGCTCGTCCATGCCG
Ssp Dna E GFP сайт разделения 69/70	
BamHI_GFP_1_F	AATTTA***GGATCC***ATGGTGAGCAAGGGCGAGGA
SspN_GFP_62_69_R	AGGATCTCAGTTCCGAAAGAAAGGCACTGCACGCCGTAGGTCAGGGTGG
SspN_GFP_52_62_F	GCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCG
Ssp_N_R_HindIII	AATTAT***AAGCTT***TTACTTGATGGTTCCAGCATCGAGCA
BamHI_Ssp_C_1_F	AATTTA***GGATCC***ATGGTTAAGGTGATTGGAAGACGTTCT
SspC_GFP_70_F	CTAATGGTGCCATCGCTGCCAATTGCTTCAGCCGCTACCCCGACCACATG
SspC_GFP_70_R	GCATGGCGGACTTGAAGAAGTCGTGCTGCTTCATGTGGTCGGGGTAGCGG
GFP_238_R_HindIII	AATTAT***AAGCTT***TTACTTGTACAGCTCGTCCATGCCG

Продолжение таблицы 1

Ssp Dna E GFP сайт разделения 127/128	
Название праймера	**Последовательность праймера (5' → 3')**
BamHI_GFP_1_F	AATTTA***GGATCC***ATGGTGAGCAAGGGCGAGGA
GFP_Ssp_N_R	CTTGATGGTTCCAGCATCGAGCA
GFP_Ssp_N_F	GCTGAAGGGCATCTGCCTTTCTTTC
Ssp_N_R_HindIII	AATTAT***AAGCTT***TTACTTGATGGTTCCAGCATCGAGCA
BamHI_Ssp_C_1_F	AATTTA***GGATCC***ATGGTTAAGGTGATTGGAAGACGTTCT
Ssp_C_GFP_R	GCTGCCAATTGCGACTTCAAGGAG
Ssp_C_GFP_F	CTCCTTGAAGTCGCAATTGGCAGC
GFP_238_R_HindIII	AATTAT***AAGCTT***TTACTTGTACAGCTCGTCCATGCCG
Набор праймеров для сборки AvaDnaE перекрытием	
N_AVA_1_F	GCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACC
N_AVA_2_R	GACAGTCAAAACTTCGGTGTCGTATGACAGACACTGCA CGCCGTAGGTCAGGGTGGTCACGAGGGTGG
N_AVA_3_F	GACACCGAAGTTTTGACTGTCGAGTATGGATTCGTGCC AATAGGCGAGATAGTCGACAAGGGGATCGA
N_AVA_4_R	GGCTGGGTGTATACAATACCGTTGCTATCGATGGAAAA CACACTGCACTCGATCCCCTTGTCGACTAT
N_AVA_5_F	CGGTATTGTATACACCCAGCCGATCGCTCAATGGCACC ACAGGGGCAAGCAGGAGGTGTTTGAGTATT
N_AVA_6_R	GAATTTATGATCCTTTGTGGCTTTAATGATAGATCCGTC TTCGAGGCAATACTCAAACACCTCCTGCT
N_AVA_7_F	TAAAGCCACAAAGGATCATAAATTCATGACGCAAGACG GCAAGATGTTACCTATTGATGAAATCTTCG
N_AVA_8_R	TTTCGGGAAGCCCTTTCACCTGCAGCAGATCGAGCTCCT GTTCGAAGATTTCATCAATAGGTAACATC
N_AVA_9_R_HindIII	AATTTA***AAGCTT***TTCGGGAAGCCCTTTCACCTG
BamHI_C_AVA_1_F	AATTTA***GGATCC***ATGATCAAAATTGCTAGCAGAAAGTTCCTCGGTGTGGA
C_AVA_2_R	TTCACGAAAAAATTGTGATCCCTTCCGACGCCTATGTCGT AAACGTTCTCCACACCGAGGAACTTTCT
C_AVA_3_F	AGGGATCACAATTTTTTCGTGAAGAATGGGCTGATCGCCT CCAACTGCTTCAGCCGCTACCCCGACCA
C_AVA_4_R	TGAAGAAGTCGTGCTGCTTCATGTGGTCGGGGTAGCGG
Ava Dna E GFP сайт 69/70	
BamHI_GFP_1_F	AATTTA***GGATCC***ATGGTGAGCAAGGGCGAGGA
N_AVA_2_R	GACAGTCAAAACTTCGGTGTCGTATGACAGACACTGCA CGCCGTAGGTCAGGGTGGTCACGAGGGTGG
N_AVA_1_F	GCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACC
N_AVA_9_R_HindIII	AATTTA***AAGCTT***TTCGGGAAGCCCTTTCACCTG
BamHI_C_AVA_1_F	AATTTA***GGATCC***ATGATCAAAATTGCTAGCAGAAAGTTCCTCGGTGTGGA
C_AVA_4_R	TGAAGAAGTCGTGCTGCTTCATGTGGTCGGGGTAGCGG
C_AVA_3_F	AGGGATCACAATTTTTTCGTGAAGAATGGGCTGATCGCCT CCAACTGCTTCAGCCGCTACCCCGACCA
GFP_238_R_HindIII	AATTAT***AAGCTT***TTACTTGTACAGCTCGTCCATGCCG

Продолжение таблицы 1

Набор праймеров для сборки NpuDnaE перекрытием	
Npu DnaE	
Название праймера	**Последовательность праймера (5' → 3')**
BamHI_GFP_NPU_N_1_F	AATTTA***GGATCC***ATGGTGAGCAAGGGCGAGG
NPU_N_2_F	GGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGTCT AAGCTATGAAACGGAAATATTGACAGTAGAATA
NPU_N_3_R	AGTACATTCGATGCGCTTTTCTACAATTTTACCAATCGGTAATA ATCCATATTCTACTGTCAATATTTCCGTTTCA
NPU_N_4_F	GAAAAGCGCATCGAATGTACTGTTTATAGCGTTGATAATAATG GAAATATTTATACACAACCTGTAGCACAATGGC
NPU_N_5_R	AATCAATGAACCATCTTCCAAACAATACTCAAACACCTCTTGTT CTCCGCGATCGTGCCATTGTGCTACAGGTTGT
NPU_N_6_F	TGTTTGGAAGATGGTTCATTGATTCGGGCAACAAAAGACCATA AGTTTATGACTGTTGATGGTCAAATGTTGCCAA
NPU_N_7_R	TTCGGCAAATTATCAACCCGCATCAAATCCAATTCACGTTCAAA TATTTCATCAATTGGCAACATTTGACCATCAA
NPU_N_8_F	CGGGTTGATAATTTGCCGAATGAGCTCCTCGAGGGCAGCAGCC ATCATCATCATCATCACAGCAGCGGCTAAAAGC
NPU_N_9_R_6H_HindIII	AATTAT***AAGCTT***TTAGCCGCTGCTGTGATG
BamHI_NPU_C_1_F	AATTTA***GGATCC***ATGATCAAAATAGCCACACG
NPU_C_2_R	TGGTCGCGCTCAACTCCAATGTCATAGACATTTTGTTTGCCTAA ATATTTACGTGTGGCTATTTTGATCATGGATCCTA
NPU_C_3_F	GGAGTTGAGCGCGACCATAATTTTGCACTCAAAAATGGCTTCA TAGCTTCTAATTGCTTCAGCCGCTACCCCGACCACA
NPU_C_4_R	AGATGGTGCGCTCCTGGACGTAGCCTTCGGGCATGGCGGACT TGAAGAAGTCGTGCTGCTTCATGTGGTCGGGGTAGCG
Npu Dna E GFP сайт 69/70	
BamHI_GFP_1_F	AATTTA***GGATCC***ATGGTGAGCAAGGGCGAGGA
NPUN_GFP_69_R	CTGCACGCCGTAGGTCAGG
NPU_N_2_F	GGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCA GTGTCTAAGCTATGAAACGGAAATATTGACAGTAGAATA
NPU_N_9_R_6H_HindIII	AATTAT***AAGCTT***TTAGCCGCTGCTGTGATG
BamHI_NPU_C_1_F	AATTTA***GGATCC***ATGATCAAAATAGCCACACG
NPU_C_5_R	AGATGGTGCGCTCCTGGACG
NPUC_GFP_70_F	TGCTTCAGCCGCTACCCCG
GFP_238_R_HindIII	AATTAT***AAGCTT***TTACTTGTACAGCTCGTCCATGCCG