Автономная некоммерческая образовательная организация
высшего образования
«**НАУЧНО–ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ «СИРИУС**»

*На правах рукописи*



**Дахневич Анастасия Ярославовна**

# Направленная эволюция ортогональных пар аминоацил–тРНК–синтетаза/тРНК для встройки неприродных аминокислот в пептиды при помощи биосинтеза

1.5.3. Молекулярная биология

ДИССЕРТАЦИЯ
на соискание ученой степени
кандидата биологических наук

Научный руководитель:
PhD
Иванов Роман Алексеевич

Федеральная территория «Сириус» – 2026

**ОГЛАВЛЕНИЕ**

**ВВЕДЕНИЕ.....5**
**1. ОБЗОР ЛИТЕРАТУРЫ.....13**
1.1 Новые белковые лекарственные препараты..... 13
1.1.1 Возможности применения неприродных аминокислот при создании белковых лекарственных препаратов..... 13
1.1.2 Классификация и способы модификации белков путем введения неприродных аминокислот.....17
1.1.3 Биотехнологические подходы получения белковых препаратов, содержащих неприродные аминокислоты в последовательности.....22
1.2 Расширение генетического кода..... 25
1.2.1 Природное расширение генетического кода на примере селеноцистеина и пирролизина.....27
1.2.2 Структурные особенности пирролизиновой аминоацил–тРНК–синтетазы.....28
1.2.3 Классификация PylRS.....33
1.3 Методы инженерии PylRS для расширения спектра включаемых НПА..... 36
1.3.1 Структурная модификация PylRS для изменения субстратной специфичности..... 37
1.3.2 Рациональная инженерия С–концевого каталитического домена PylRS.....38
1.3.3 Рациональная инженерия N–концевого домена и междоменного линкера..... 41
1.3.4 Направленная эволюция аминоацил–тРНК–синтетаз..... 43
1.3.5 PACE как метод направленной эволюции PylRS.....44
1.4 Способы декорации липидных наночастиц..... 48
**2. МАТЕРИАЛЫ И МЕТОДЫ.....53**
2.1 Общие материалы и методы.....53
2.1.1 Питательные среды и растворы..... 53
2.1.2 Бактериальные штаммы.....53
2.1.3 Плазмиды и гены.....54
2.1.4 Конструирование олигонуклеотидов и проведение ПЦР.....54
2.1.5 Приготовление электрокомпетентных клеток и электропорация.....55
2.1.6 Секвенирование методом Сэнгера.....56
2.2 Сборка плазмидных конструкций.....56
2.2.1 Получение фаговых библиотек, несущих варианты гена PylRS..... 56
2.2.2 Сборка плазмид негативной селекции..... 58
2.2.3 Сборка плазмид флуоресцентной тест системы.....59
2.3. Фаг ассоциированная эволюция..... 60
2.3.1 Цикл фаг–ассоциированной эволюции (PANCE).....60
2.3.2 Определение селективности АРС полученных в результате PANCE..... 61
2.3.3 Определение активности АРС полученных в результате PANCE методом теста флуоресценции sfGFP..... 61
2.3.4 Подготовка библиотеки для секвенирования на Illumina MiSeq..... 62
2.3.5 Обработка результатов MiSeq..... 63

2.3.6 Моделирование белковых структур....63
2.4 Наработка и очистка белков....64
2.4.1 Наработка энтеропептидазы....65
2.4.2 Наработка Эксендин–4 и DARPin mCD8....65
2.4.4 Обращенно–фазовая хроматографическая очистка образцов....66
2.4.5 Электрофоретическое разделение белков в ПААГ....67
2.4.6 Масс–спектрометрический анализ....67
2.4.7 Клик–химическая конъюгация....68
2.5 Создание липидных наночастиц....69
2.5.1 Транскрипция и очистка РНК in vitro....69
2.5.2 Формуляция липидных наночастиц....69
2.5.3 Физико–химические характеристики ЛНЧ....70
2.6 Клеточные тесты....71
2.6.1 Выделение и активация мышиных спленоцитов....71
2.6.2 Цитометрическое окрашивание мышиных спленоцитов....72
2.6.3 Клеточные тесты активности агонистов GLP–1R....72
2.6.4 Цитометрическое окрашивание клеток, экспрессирующих GLP–1R....74
2.7 Статистическая обработка данных....75
**3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЯ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ....76**
3.1 Разработка и валидация платформы направленной эволюции PylRS....76
3.1.1 Получение генетических последовательностей вариантов MmPylRS и ChPylRS....76
3.1.2 Оптимизация системы направленной эволюции PylRS методом PANCE....78
3.1.3 Разработка отрицательной селекции....81
3.1.4 Анализ динамики эволюции PylRS и накопления замен....85
3.1.5 Оценка активности отобранных вариантов PylRS....88
3.1.6 Структурный анализ выявленных аминокислотных замен PylRS....93
3.2 Разработка платформы сайт–специфической функционализации липидных наночастиц на основе белков с НПА....100
3.2.1 Функционализация липидных наночастиц при помощи DARPin....100
3.2.1.1 Получение DARPin mCD8....100
3.2.1.2 Получение и характеристика липидных наночастиц, декорированных DARPin....105
3.2.1.3 Функциональные тесты с ЛНЧ–DARPin....110
3.2.2 Функционализация липидных наночастиц при помощи Эксендин–4....113
3.2.2.1 Создание генетических конструкций для наработки Эксендина–4....113
3.2.2.2 Оптимизация условий наработки Эксендина–4....115
3.2.2.3 Получение рекомбинантной энтеропептидазы для удаления фьюжн–тага....121
3.2.2.4 Получение Exe4 с pAzF и его характеризация....124
3.2.2.5 Оценка эффективности связывания Exe4 с рецептором....128
3.2.2.6 Получение и характеристика липидных наночастиц, декорированных

Эксендин–4..... 131
**ЗАКЛЮЧЕНИЕ..... 139**
**ВЫВОДЫ..... 143**
**СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ..... 145**
**СПИСОК ЛИТЕРАТУРЫ..... 148**
Приложение А
Среды и растворы..... 159
Приложение Б
Используемые плазмиды..... 162
Приложение В
Праймеры для сборки ChPylRS..... 166
Приложение Г
Последовательности нуклеотидов, аминокислот и праймеры для сборки Эксендина–4..... 168
Приложение Д
Последовательности нуклеотидов, аминокислот и праймеры для сборки hEKL..... 169
Приложение Е
Последовательности нуклеотидов, аминокислот и праймеры для сборки DARPin mCD8... 173

# ВВЕДЕНИЕ

**Актуальность и степень разработанности темы исследования.**

В природе белки состоят из 22 аминокислот, из которых 20 являются основными, а две – дополнительными: селеноцистеин и пирролизин. Помимо протеиногенных, в природе существует более 500 неприродных аминокислот (НПА), которые не встраиваются в белки посредством рибосомального синтеза. НПА, введенные в аминокислотную последовательность, обеспечивают наличие нехарактерных химических групп, что может обеспечить рекомбинантным белкам уникальные свойства, создает возможности их селективной химической модификации и расширяет перспективы применения. Согласно метаанализам и систематическим обзорам, в мировой медицинской практике одобрено и находится на разных стадиях клинических испытаний более ста десяти белковых препаратов, содержащих различные неприродные аминокислотные фрагменты.

Данный подход обладает рядом преимуществ: высокой эффективностью, возможностью автоматизации процесса, сокращением времени синтеза и упрощением выделения целевого продукта. К основным недостаткам относятся высокая стоимость синтеза, значительный расход дорогостоящих реагентов и органических растворителей, а также образование большого количества химических отходов, что ограничивает экономическую и экологическую эффективность метода. Развивающимся направлением современной биофармацевтики стало расширение генетического кода с помощью инженерных ортогональных пар тРНК и аминоацил–тРНК–синтетаз (АРС), что дает возможность осуществлять сайт–специфическое встраивание неприродных аминокислот в рекомбинантные белки и пептиды в процессе их трансляции клетками–продуцентами. Таким образом, НПА может быть прецизионно направлена в любую заранее выбранную позицию в ходе наработки в традиционных экспрессионных системах. Такой подход не только снижает стоимость синтезируемых субстанций, но и расширяет возможности получения белков и пептидов большего размера, имеющих сложную пространственную организацию.

Для сайт–направленного включения НПА в белки требуется наличие в клетке соответствующих механизмов. Одним из важнейших элементов расширения генетического кода является наличие свободного кодона, в настоящее время используется амбер стоп–кодон из–за его малой распространенности. Помимо свободного кодона, важным является наличие ортогональных пар АРС/тРНК. Ортогональными являются те АРС, которые не взаимодействуют с протеиногенными аминокислотами и тРНК клетки–хозяина, при этом комплементарны

синтетическому аппарату клетки и избирательно распознают одну из НПА в качестве субстрата. Ортогональные пары АРС/тРНК были найдены и адаптированы к модельным объектам в начале века. Наибольшее распространение получила пирролизил–тРНК–синтетаза (PylRS) из архей рода *Methanosarcina*. Эта АРС является полностью ортогональной, так как не имеет перекрестных взаимодействий с эндогенными элементами системы трансляции ни у бактерий, ни у эукариот, обладает большим карманом связывания аминокислот в активном центре и лишена домена редактирования, что важно для распознавания большого количества НПА. В то же время $тРНК_{Pyl}$ является природным супрессором амбер стоп–кодона, при этом антикодон не является элементом идентичности для PylRS, вследствие чего данную систему можно использовать для перекодирования UAG–кодона. Для дальнейшего развития технологии актуальной задачей становится оптимизация синтетаз для повышения эффективности их работы как с точки зрения селективности, так и с точки зрения биохимических характеристик.

Одним из путей решения этой задачи является использование методов направленной эволюции. В период с 2017 по 2020 год был разработан метод PACE (фаг–ассоциированная непрерывная эволюция), успешно адаптированный для работы с АРС. Суть метода заключается в организации зависимости скорости размножения фага от активности гена интереса, встроенного в его геном. Несмотря на небольшой срок использования, метод продемонстрировал высокую эффективность. Однако до сих пор он применяется лишь в нескольких зарубежных научных центрах. Оптимизация данного метода и разработка новых ортогональных пар АРС/тРНК позволят расширить возможности генетически направленного включения НПА в белки.

Особый интерес представляют НПА с азидогруппой, обеспечивающие селективную конъюгацию посредством реакции азид–алкинового циклоприсоединения (SPAAC) в мягких условиях без медных катализаторов. Такой подход позволяет контролировать положение модификации и стехиометрию продукта, снижая образование позиционных изомеров и повышая однородность белковых конъюгатов. Применение данного подхода для функционализации поверхности липидных наночастиц (ЛНЧ) белками позволяет целенаправленно изменять их взаимодействие с клетками, повышать специфичность связывания с клетками–мишенями и улучшать адресность доставки мРНК. Это открывает перспективы для разработки систем доставки терапевтических молекул в Т–лимфоциты для CAR–Т–терапии *in vivo*, а также в бета–клетки островков Лангерганса для создания новых подходов к профилактике и лечению сахарного диабета 1–го типа. Перспективность метода подтверждается клиническими исследованиями белковых препаратов с азидосодержащими

НПА в составе (NCT04829604 и NCT04009681), а его универсальность может быть продемонстрирована на примере получения функциональных пептидов и биотерапевтических препаратов с заранее заданными свойствами.

**Целью** работы является создание биотехнологической платформы для разработки и производства инновационных лекарственных препаратов на основе белков с неприродными аминокислотами. Для достижения цели были поставлены следующие **задачи**:

1. Оптимизация протокола направленной эволюции при помощи бактериофагов для повышения специфичности пар АРС/тРНК для встраивания целевых НПА и улучшения каталитических характеристик;
2. Изучение роли аминокислотных замен эволюционировавших вариантов PylRS с использованием функционального тестирования и структурного моделирования;
3. Исследование эффективности включения НПА в белки при помощи вариантов PylRS с использованием флуоресцентной тест–системы;
4. Разработка лабораторной технологии получения Exe4, пептидного агонистов GLP–1 рецептора, и рекомбинантного DARPin, специфичного к мышиному CD8, с 4–Азидо–L–фенилаланином;
5. Разработка метода функционализации липидных наночастиц белками с использованием безмедного азид–алкинового циклоприсоединения;
6. Исследование фармакологической активности полученных липидных наночастиц на клеточных моделях GLP–1R– и CD8–положительных клеток;

**Научная новизна исследования** заключается в разработке и валидации метода расширения генетического кода бактерий с использованием платформы направленной эволюции.

Реализована стратегия направленной эволюции как нативной пирролизил–тРНК–синтетазы *Methanosarcina mazei* (MmPylRS), так и химерного белка на основе N–домена *Methanosarcina barkeri* и С домена *Methanosarcina mazei* (ChPylRS) с двойной аминокислотной заменой Asn346Ala/Cys348Ala с заданной специфичностью к четырем производным фенилаланина и тирозина: О–аллил–L–тирозина, О–метил–L–тирозина, 4–азидо–L–фенилаланина и 2–хлоро–L–фенилаланина. В работе оптимизирован протокол фаг–ассоциированной эволюции PANCE посредством интеграции циклов отрицательной селекции, проводимой в отсутствие целевых НПА, что позволило эффективно подавлять неспецифическую активность ортогональных АРС.

В результате направленной эволюции тРНК–синтетазы идентифицирован ряд новых, ранее не описанных в литературе аминокислотных замен. В частности, обнаружена конвергентно возникшая во всех эволюционных линиях замена Met300Thr в каталитическом домене MmPylRS. Показано, что появление гидроксильной группы треонина в положении 300 перестраивает сеть водородных связей активного центра фермента, стабилизируя объемные ароматические радикалы аналогов тирозина и фенилаланина, использованных в данном исследовании.

Впервые разработана и структурно охарактеризована панель из четырех уникальных ортогональных синтетаз с субстратной специфичностью: AllYPylRS, MethYPylRS, AzFPylRS и 2ClFPylRS (в структурном каркасе химерного белка ChPylRS). Новые замены в N–домене (Glu88Lys, Arg55Cys, Asp2Asn), согласно предложенному механизму, способствуют перераспределению связей с тРНК с N– на C–домен, приводя к образованию более стабильного комплекса АРС–тРНК. Также установлена роль ряда новых аминокислотных замен C–домена (например, Glu210Lys, Glu357Lys, Asn280Asp), в расширении субстратного кармана, образовании новых водородных связей с карбоксильной группой производных тирозина и стабилизации третичной структуры фермента.

Кроме того, в работе впервые осуществлен биосинтез функциональных белков DARPin, специфичных к мышиному антигену CD8 (mCD8) и пептидных агонистов GLP–1 рецептора, несущих направлено встроенный 4–азидо–L–фенилаланин (pAzF) на C–конце молекулы, что открывает новый биотехнологический путь для их последующей конъюгации с липидными наночастицами для таргетной доставки мРНК.

**Теоретическая и практическая значимость работы** заключается в расширении фундаментальных представлений о пластичности и границах адаптивности трансляционного аппарата живых систем. Результаты исследования раскрывают молекулярные механизмы эволюционной перестройки аминоацил–тРНК–синтетаз II класса под новые аминокислоты, содержащие пара–, мета– и орто–замещенные бензольных колец.

Практическая значимость диссертационного исследования определяется созданием готовой технологической платформы для высокоэффективной биоортогональной модификации белковых молекул. Сконструированная серия рекомбинантных штаммов–продуцентов *E. coli*, экспрессирующих оптимизированные пары PylRS/тРНК$_{Pyl}$, позволяет осуществлять препаративную наработку полноразмерных целевых белков с уровнем встраивания НПА в диапазоне 89–98%.

Разработанные модифицированные белки DARPin anti–mCD8 и Exe4 с остатком pAzF на C–конце молекулы успешно применены для сайт–направленной безмедной клик–химической конъюгации (SPAAC) с липидными наночастицами (ЛНЧ), что обеспечивает адресную доставку терапевтических мРНК в CD8 и GLP–1R положительные клетки.

**Методология и методы исследования** диссертационной работы базируются на фундаментальных принципах молекулярной биологии, белковой инженерии и направленной эволюции ферментов, реализуемых через концепцию расширения генетического кода бактерий путем создания искусственных ортогональных пар «аминоацил–тРНК–синтетаза/тРНК». Экспериментальная стратегия исследования основана на модернизации платформы фаг–ассоциированной эволюции PANCE за счет интеграции прерывистых циклов негативной селекции, что обеспечивает элиминацию фоновой активности ферментов к каноническим аминокислотам. В данной работе используется широкий спектр молекулярных методов клонирования (например, изотермической сборки с помощью Gibson Assembly, SLIC или Golden Gate Assembly) для создания генетических векторов; микробиологические методы скрининга клонов на селективных средах и оптимизации условий ферментации штаммов–продуцентов *E. coli*; биоинформатические подходы для высокопроизводительного секвенирования (NGS, Illumina MiSeq) эволюционных пулов и конформационного 3D–моделирования активных центров синтетаз (AlphaFold3, PyMOL). Оценка эффективности супрессии амбер–кодона осуществлялась флуоресцентным анализом репортерного белка sfGFP, а препаративное выделение и очистка целевых молекул DARPin anti–mCD8 и пептидов GLP–1 выполнялись методами металлоаффинной (Ni–IMAC), ионообменной хроматографии и высокоэффективной жидкостной (ВЭЖХ) хроматографии. Достоверность и точность сайт–специфического встраивания неприродных аминокислот подтверждались методами электрофореза (SDS–PAGE) и времяпролетной масс–спектрометрии. Для создания конъюгатов белков, содержащих pAzF, использовались методы биоортогональной химии для проведения реакций азид–алкинового циклоприсоединения (SPAAC) с целью направленной клик–конъюгации полученных модифицированных белков и пептидов с ЛНЧ.

**Положения, выносимые на защиту.**

1. Аминокислотные замены в MmPylRS обеспечивают изменение субстратного профиля синтетазы в сторону O–аллил–L–тирозина (Asp2Asn, Glu88Lys, Ala107Thr, Asn280Asp, Met300Thr, Asn346Ala, Cys348Ala, Asn420Ser), O–метил–L–тирозина (Met17Leu, Glu39Lys, Pro109Ser, Met300Thr, Glu337Lys, Asn346Ala, Cys348Ala, Asp408Gly) и 4–азидо–L–фенилаланина (Asp2Asn, Arg55Cys, Met300Thr, Tyr306His, Asn346Ala,

Cys348Ala, Glu357Lys), обеспечивая десятикратное повышение эффективности включения соответствующих неприродных аминокислот по сравнению с исходной синтетазой, не проходившей эволюцию, при сохранении кросс–специфичности к структурно сходным субстратам.

2. Химерная двухсубъединичная PylRS, объединяющая N–концевой фрагмент PylRS *Methanosarcina barkeri* (остатки 1–102) с заменами His62Tyr, Ala100Ser и Pro101Arg и вторую субъединицу, содержащую фрагмент PylRS *M. barkeri* (остатки 107–149) и фрагмент PylRS *M. mazei* (остатки 185–454) с заменами Asn241Ser, Glu249Lys, Met300Thr, Asn346Ala и Cys348Ala, обеспечивает селективное включение 2–хлоро–L–фенилаланина в белки при помощи биосинтеза.
3. Сайт–специфическое включение 4–азидо–L–фенилаланина в белки Exe4 и DARPin mCD8 позволяет проводить их биоортогональную конъюгацию и последующую функционализацию липидных наночастиц, обеспечивая рецептор–зависимое таргетирование GLP–1R–положительных и CD8–положительных клеток и повышение эффективности трансфекции.

**Степень достоверности и апробация результатов работы.** Достоверность результатов диссертационного исследования обеспечивается применением современных молекулярно–биологических, биоинженерных, биохимических и клеточно–биологических методов, включая генетическое конструирование, секвенирование, гетерологичную экспрессию рекомбинантных белков, хроматографическую очистку и функциональное тестирование полученных пептидов на клеточных моделях. Достоверность полученных данных подтверждается использованием контрольных образцов, проведением независимых экспериментальных повторов и применением адекватных методов статистической обработки.

Результаты диссертационной работы были представлены на итоговой аттестации научного центра трансляционной медицины АНОО ВО «Университет «Сириус», а также в формате устного доклада на Саммите разработчиков лекарственных препаратов «Сириус.Биотех», 2026, Федеральная территория «Сириус», Россия.

По результатам, представленным в диссертационном исследовании, было опубликовано 3 печатные работы, в рецензируемых изданиях, индексируемых в базах данных Scopus и Web of Science:

1. **Dakhnevich A,** Kazakova A, Iliushin D, Ivanov RA. Pyrrolysine Aminoacyl–tRNA Synthetase as a Tool for Expanding the Genetic Code. Int J Mol Sci. 2025 Jan 10;26(2):539. doi: 10.3390/ijms26020539. PMID: 39859254; PMCID: PMC11764691.

2. **Dakhnevich A**, Miasoutova S, Iliushin D, Rozanov A, Ivanov R. Expanding the genetic code: phage–driven evolution of pyrrolysyl–synthetase for site–specific incorporation of synthetic phenylalanine and tyrosine derivatives. Front Mol Biosci. 2026 Mar 17;13:1737987. doi: 10.3389/fmolb.2026.1737987. PMID: 41924351; PMCID: PMC13035772.
3. **Dakhnevich, A.**; Pateev, I.; Skvortsov, I.A.; Yarmachkova, S.; Shevyrev, D.; Chubarova, A.; Gasina, M.; Cherepanova, N.; Siaglova, E.D.; Reshetnikov, V.; et al. Genetic Code Expansion Enables Oriented, Site–Specific Conjugation of DARPins to Lipid Nanoparticles for Selective mRNA Delivery to CD8+ T Cells. Int. J. Mol. Sci. 2026, 27, 7755. https://doi.org/10.3390/ijms27177755

**Личный вклад автора**

Основной объем экспериментальных исследований выполнен автором лично либо при его непосредственном участии. Автором проведен анализ литературы, разработан и оптимизирован протокол PANCE с интеграцией циклов негативной селекции, сконструированы плазмидные векторы, проведены молекулярно–генетические исследования, анализ белков и валидация протокола SPAAC. Совместно с соавторами соответствующих публикаций выполнены селекция фаговых популяций, подготовка и биоинформатический анализ данных высокопроизводительного секвенирования, структурное моделирование вариантов PylRS, оптимизация ферментации и выделение модифицированных белков и пептидов, а также эксперименты с наночастицами *in vitro*. Автор принимал непосредственное участие в интерпретации масс–спектрометрических данных. Подготовка научных публикаций, патентных заявок, диссертации и автореферата осуществлена автором лично; научный руководитель Иванов Р.А. и соавторы принимали непосредственное участие в обсуждении результатов, их интерпретации и подготовке совместных публикаций.

**Объем и структура диссертации**

Диссертация изложена на 154 страницах и построена по следующему плану: введение, обзор литературы, описание материалов и методов, результаты и обсуждения, выводы, приложение. Библиографический список включает 211 источников. Работа проиллюстрирована 43 рисунками и содержит 5 таблиц.

# 1. ОБЗОР ЛИТЕРАТУРЫ

## 1.1 Новые белковые лекарственные препараты

Белковые и пептидные препараты занимают центральное место в современной фармацевтической индустрии благодаря высокой специфичности взаимодействия с биологическими мишенями и применяются для лечения онкологических, сердечно–сосудистых и метаболических заболеваний, а также в иммуномодулирующей и заместительной терапии [1,2]. Однако потенциал классических рекомбинантных белков ограничен однообразием стандартного генетического кода. В природе белки состоят из 22 аминокислот, из которых 20 являются основными, а две – дополнительными: селеноцистеин и пирролизин. Помимо протеиногенных аминокислот, в природе существует более 500 неприродных или некодируемых аминокислот, которые не включаются в белки посредством рибосомального синтеза [3]. Еще больше вариантов может быть синтезировано химическим путем. В последние годы в биофармацевтике значительное внимание уделяется разработке препаратов с включением НПА. Принципиальная важность этого подхода заключается в возможности введения НПА в строго заданную позицию полипептидной цепи с последующим использованием ее уникальной химической группы для изменения фармакологических свойств молекулы, биоконъюгации или создания новой биологической функции [4–6].

### 1.1.1 Возможности применения неприродных аминокислот при создании белковых лекарственных препаратов

Введение неприродных аминокислот в структуру пептидных препаратов уже реализовано в ряде лекарственных средств, применяемых в клинической практике. Так например, semaglutide (Ozempic™/Wegovy™, Novo Nordisk) содержит 2–аминоизомасляную кислоту (Aib) и C18–жирнокислотный фрагмент: Aib способствует стабилизации α–спиральной конформации и повышению устойчивости к протеолизу, тогда как липидный фрагмент обеспечивает связывание с альбумином и увеличивает период полувыведения препарата до примерно 7 дней, что позволяет применять его один раз в неделю [7]. Другими примерами

являются октреотид™ (Octreotide, Деко) и ланреотид™ (Somatuline, Ipsen) – синтетические аналоги соматостатина, содержащие D–фенилаланин и D–триптофан. Введение аминокислот в D–конфигурации повышает устойчивость пептидов к протеолитической деградации и способствует поддержанию конформации, необходимой для взаимодействия с рецептором [8,9].

Клинически реализованные примеры демонстрируют, что включение НПА позволяет целенаправленно изменять свойства пептидных молекул, включая конформационную стабильность, устойчивость к протеолизу и продолжительность действия. Дальнейшее развитие этого подхода связано с использованием НПА для направленной оптимизации взаимодействия терапевтических белков и пептидов с молекулярными мишенями. Показательным примером является человеческий PD–1, в котором включение флюоросульфат–L–тирозина (FSY) в положение 129 вместо аланина обеспечило ковалентное взаимодействие с остатком His69 белка PD–L1. Полученный вариант PD–1(FSY) продемонстрировал усиленное противоопухолевое действие *in vivo* за счет образования более стабильного ковалентного комплекса PD–1/PD–L1 [10]. Другим примером служит модификация тритриптицина – катионного антимикробного пептида, обладающего противоопухолевым потенциалом. Замена Arg на Lys и последующее введение неприродных производных Lys позволили повысить селективность производных тритриптицина в отношении раковых клеток и усилить их способность к быстрой пермеабилизации цитоплазматической мембраны, что сопровождалось выраженным цитотоксическим эффектом в отношении опухолевых клеток [11].

Помимо модификации уже известных терапевтических молекул, НПА используются при конструировании принципиально новых лекарственных средств. Отдельное направление связано с разработкой макроциклических пептидных препаратов, в которых введение неприродных структурных элементов расширяет химическое пространство потенциальных терапевтических соединений. Так, Zampaloni и соавт. разработали зосурабалпин – макроциклический пептидный антибиотик, воздействующий на липополисахаридный транспортер $LptB_2FGC$ у *Acinetobacter baumannii*. Полученное соединение проявляет выраженную активность против штаммов *A. baumannii*, устойчивых к карбапенемам, и продемонстрировало эффективность в доклинических моделях инфекции [12].

С фармацевтической точки зрения одним из наиболее перспективных и технологически развитых направлений применения НПА является сайт–специфическая модификация белков, обеспечивающая контролируемое введение биортогональных функциональных групп в строго определённые положения полипептидной цепи. Наличие таких функциональных групп позволяет осуществлять последующую селективную химическую модификацию белков

посредством биортогональных реакций, не затрагивающих другие реакционноспособные участки молекулы. К таким реакциям относятся азид–алкиновое циклоприсоединение (клик–химия) [13], реакция Штаудингера (для синтеза фосфазенов) и инверсная реакция Дильса – Альдера (для создания пиридиновых колец и других гетероциклов) [14]. Этот подход применим для PEG–илирования белков, присоединения липидных фрагментов, флуоресцентных меток, радионуклидов, олигонуклеотидов и лекарственных веществ [15–18].

В отличие от традиционных подходов, основанных преимущественно на химической модификации реакционноспособных остатков лизина или цистеина, локализованных на поверхности белка [19,20]. При применении традиционных методов модификации наличие нескольких доступных реакционных центров обусловливает стохастический характер протекания реакции и, как следствие, образование гетерогенной популяции белковых молекул, различающихся по степени и сайтам модификации. Это существенно осложняет контроль молекулярной структуры продукта, его последующую очистку и аналитическую характеристику. Генетическое включение НПА, содержащей уникальную биортогональную функциональную группу, позволяет локализовать реакционноспособный центр в заранее определенном участке полипептидной цепи и обеспечить последующее селективное введение целевых функциональных компонентов.

Особенно значительный потенциал сайт–специфическая конъюгация приобрела в области разработки антитело–лекарственных конъюгатов (antibody–drug conjugates, ADC). Для данной группы препаратов принципиальное значение имеют число молекул цитотоксического агента, приходящихся на одну молекулу антитела, положение конъюгации, стабильность линкера, а также сохранение структурной целостности и способности антитела к специфическому связыванию с мишенью. В случае традиционных методов конъюгации вариабельность указанных параметров обусловливает структурную неоднородность получаемого препарата и потенциально может влиять на его фармакокинетические, фармакодинамические и терапевтические характеристики. Генетическое включение р–ацетил–фенилаланина (pAF) или других биортогональных НПА в заранее выбранные положения антитела позволяет создавать строго определенные реакционные центры, которые в дальнейшем используются для селективного химического присоединения лекарственного агента [21,22]. Такой подход обеспечивает более высокий уровень контроля и позволяет минимизировать гетерогенность продукта, связанную с вариабельностью сайтов конъюгации.

Одним из наиболее известных примеров такого подхода является ARX788 – сайт–специфический анти–HER2 ADC, в котором pAF генетически вводится в антитело и

используется для последующего присоединения цитотоксического компонента. В доклинических исследованиях ARX788 продемонстрировал высокую противоопухолевую активность, в том числе в моделях HER2–положительных и HER2–low опухолей, а также в моделях резистентности к T–DM1 [22]. Клинические исследования подтвердили возможность применения этой платформы у пациентов с HER2–положительными солидными опухолями. Важным преимуществом препарата является его конструктивная однородность, достигаемая благодаря сайт–специфическому включению НПА и контролируемой конъюгации [23].

Применение НПА находят и в области флуоресцентной визуализации белков и исследования динамики белок–белковых взаимодействий в живых клетках. Использование крупных генетически кодируемых флуоресцентных белков, таких как GFP, не всегда оптимально, поскольку значительный размер может влиять на пространственную организацию, внутриклеточную локализацию и функциональную активность исследуемого белка, особенно если предметом исследования является небольшой или мембраносвязанный белок [24]. Альтернативным подходом является сайт–специфическое включение НПА, обладающих флуоресцентными свойствами либо содержащих биортогональные функциональные группы для последующего присоединения органических флуорофоров, что позволяет проводить высокоселективную маркировку белка без введения крупного флуоресцентного домена. Такой подход обеспечивает минимальное структурно–функциональное воздействие метки и открывает возможности для прижизненной визуализации локализации, транспорта и динамики белков [25].

Отдельное направление применения НПА связано с сайт–специфической липидизацией белков и пептидов. Липидные фрагменты способны взаимодействовать с сывороточным альбумином, благодаря чему увеличивается время циркуляции модифицированного препарата. Генетически кодируемые неприродные аминокислоты могут использоваться для сайт–специфического PEG–илирования белков. В отличие от традиционного случайного PEG–илирования, такая стратегия позволяет заранее выбирать положение полимерного фрагмента и тем самым оптимизировать влияние модификации на биологическую активность и фармакокинетику. Подобный подход представляет интерес для гормонов, цитокинов, ферментов и других терапевтических белков, для которых короткий период полувыведения является существенным ограничением клинического применения [26].

Отдельного внимания заслуживает использование сайт–специфически введенных НПА в тераностике, поскольку позволяет сочетать адресную доставку биологически активного соединения с радионуклидной визуализацией или терапией. Природные аминокислоты плохо

удерживают тяжелые металлы, поэтому для этих целей используют НПА способные присоединять хелатирующие группы. Экспериментально продемонстрирована возможность сайт–специфического введения DOTA в рекомбинантные антитела с последующим радиомечением $^{68}Ga$ и $^{111}In$ при сохранении способности полученных конъюгатов к специфическому связыванию с молекулярной мишенью [27,28].

Формирующийся рынок белковых препаратов с НПА постепенно переходит от преимущественно исследовательской стадии к практическому применению в создании ADC, разработке белковых препаратов с оптимизированными фармакокинетическими и фармакодинамическими характеристиками, повышенной стабильностью и специфичностью, а также возможностью многофункциональной и сайт–специфической конъюгации. Таким образом, технологии расширения химического многообразия белков постепенно переходят из области фундаментальной биологии в платформу молекулярного конструирования биофармацевтических препаратов. В связи с этим особый интерес представляют методы введения НПА, позволяющие селективно модифицировать уже сформированные белковые молекулы и получать функционально заданные биоконъюгаты.

### 1.1.2 Классификация и способы модификации белков путем введения неприродных аминокислот

Расширение химического многообразия белков за пределы двадцати двух канонических аминокислот может осуществляться посредством химического синтеза пептидных фрагментов, содержащих заданные НПА и их последующем объединении с другими фрагментами, либо химической трансформации уже сформированной белковой молекулы. Такой подход позволяет получать белки с точно определенным составом и локализацией модифицированных аминокислот, включая D–аминокислоты, аминокислоты с измененными боковыми цепями, флуоресцентные и фотоактивируемые остатки и другие [29,30].

В зависимости от способа формирования полипептидной цепи химические методы введения НПА условно можно разделить на три основные группы: полный химический синтез, полусинтетические методы на основе химического лигирования и постсинтетическую сайт–селективную химическую модификацию белков. При полном химическом синтезе вся полипептидная цепь формируется посредством последовательного присоединения

аминокислотных остатков, преимущественно методом твердофазного пептидного синтеза (Solid–Phase Peptide Synthesis, SPPS). SPPS, разработанный Брюсом Меррифилдом, является одним из основных методов химического получения пептидов. Метод был впервые описан Меррифилдом в 1963г. и основан на последовательном присоединении защищенных аминокислот к растущей полипептидной цепи, ковалентно связанной с нерастворимым полимерным носителем [31].

Типичный цикл SPPS включает последовательное выполнение нескольких стадий: присоединение первой аминокислоты к функционализированному полимерному носителю; удаление временной защитной группы с аминогруппы растущей цепи, например Fmoc–группы с использованием пиперидина; активацию карбоксильной группы следующей аминокислоты с помощью конденсирующего реагента, например HATU или PyBOP, и её последующее присоединение к свободной аминогруппе пептида; промывку смолы для удаления избытка реагентов и побочных продуктов [32]. После завершения каждого цикла защитная группа удаляется, и последовательность операций повторяется до получения пептида заданной длины (рисунок 1.1).

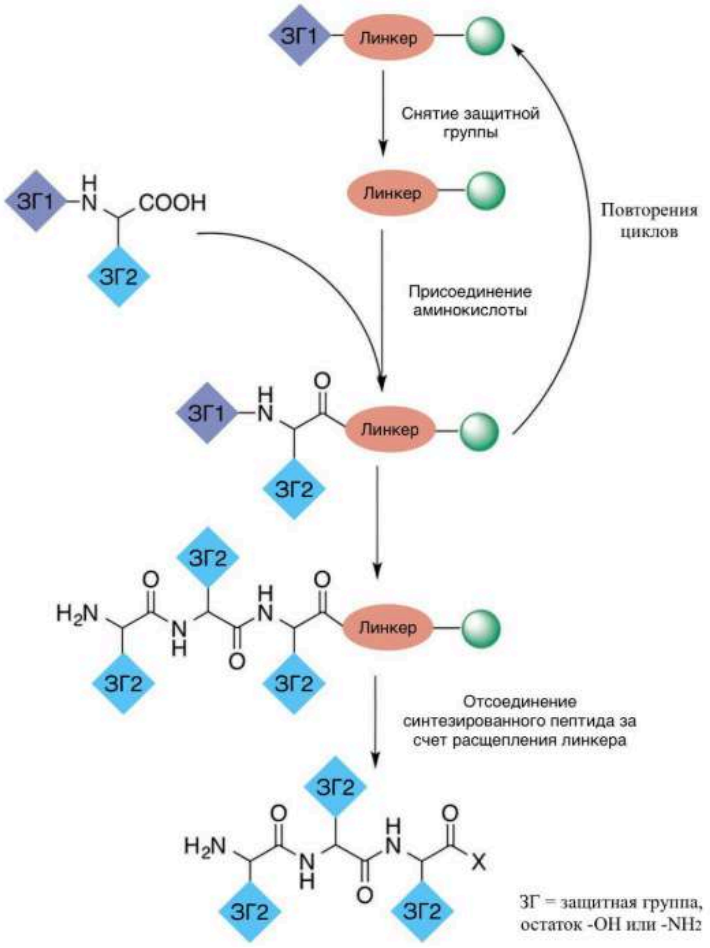



Рисунок 1.1 – Метод твердофазного химического синтеза пептидов. Адаптировано из [32]

Особое значение SPPS имеет для химического введения неприродных аминокислот, поскольку НПА могут непосредственно использоваться в качестве строительных блоков на соответствующей стадии сборки полипептидной цепи. Это позволяет получать пептиды с заданным расположением НПА, а также сочетать их с различными модификациями боковых цепей и N– или C–концевых участков молекулы.

Основным ограничением метода является накопление побочных продуктов при многократном повторении циклов синтеза. Даже неполное протекание отдельных стадий приводит к образованию укороченных последовательностей, которые могут сохраняться в конечном продукте. Кроме того, по мере увеличения длины цепи возрастает вероятность неполного сочетания аминокислот, агрегации растущего пептида и снижения общей эффективности синтеза. Поэтому классический SPPS наиболее эффективно применяется для синтеза пептидов относительно небольшой длины, тогда как получение более крупных

белковых молекул обычно требует комбинации химического синтеза отдельных фрагментов с последующим химическим лигированием [33].

Нативное химическое лигирование (native chemical ligation, NCL) – хемоселективная реакция, позволяющая соединять незащищенные пептидные фрагменты в водном растворе. Классический вариант реакции основан на взаимодействии С–концевого тиоэфира одного фрагмента с N–концевым цистеиновым остатком другого с последующей перегруппировкой и образованием нативной пептидной связи. Развитие NCL значительно расширило возможности химического синтеза белков, поскольку позволило объединять несколько относительно коротких синтетических фрагментов в полипептиды большей длины [29,34].

Особое значение NCL имеет для введения НПА в строго определенные участки белковой молекулы при лигировании положение неприродного остатка задаётся его расположением в исходном синтетическом пептидном фрагменте. Это позволяет получать однородные белковые продукты, содержащие D–аминокислоты, нестандартные функциональные группы, флуоресцентные метки, стабильные изотопы и другие химически заданные структурные элементы. Расширение классической NCL включает использование селенольных и тиольных производных НПА, а также последующие реакции десульфуризации и деселенизации, позволяющие формировать нативные аминокислотные остатки после проведения лигирования [35].

Развитием данного подхода является экспрессируемое белковое лигирование (expressed protein ligation, EPL), объединяющее рекомбинантные и синтетические компоненты. В рамках EPL один фрагмент белка получают биотехнологически и переводят в реакционноспособную форму, тогда как второй, содержащий НПА, получают химическим синтезом [36]. Последующее хемоселективное соединение этих компонентов позволяет включать неприродные аминокислотные остатки в определенные участки достаточно крупных белков, сохраняя преимущества рекомбинантного получения основной части молекулы.

Альтернативным вариантом получения НПА при помощи химических методов является сайт–селективные модификации уже сформированных белков. НПА создаётся непосредственно в нативной полипептидной цепи, химически преобразуется природный аминокислотный остаток либо используется его уникальная реакционная группа для последующего введения заданного функционального фрагмента. Селективность таких превращений определяется различиями в реакционной способности боковых цепей, их доступности и микроокружении, а также использованием хемоселективных и биортогональных реакций, протекающих в мягких водных условиях [37]. Наиболее широко в подобных реакциях используется цистеин, тиольная

группа которого обладает высокой нуклеофильностью, что создает предпосылки для высокой хемоселективности. Сульфгидрильная группа цистеина может вступать в реакции с малеимидами, α–галогенкарбонильными соединениями и другими электрофильными реагентами, обеспечивая образование устойчивых ковалентных связей с функциональными молекулами [38]. При необходимости уникальный реакционный центр может быть создан посредством замены исходного остатка на цистеин в выбранной позиции белка, после чего проводится его селективная химическая модификация. Особый интерес представляет химическое превращение цистеина в дегидроаланин (Dha) – неприродный аминокислотный остаток, обладающий выраженными электрофильными свойствами [39]. Такой подход существенно расширяет спектр доступных химических модификаций: посредством Dha в белковую структуру могут быть введены аналоги липидирования, гликозилирования, фосфорилирования, ацетилирования и метилирования лизина [17].

Химические методы позволяют получать белки с НПА за счет непосредственного химического синтеза, фрагментного лигирования или селективной химической трансформации уже сформированной белковой структуры. Наиболее высоким уровнем контроля над первичной структурой характеризуются тотальный синтез и полусинтетические стратегии на основе NCL/EPL, тогда как постсинтетическая модификация предоставляет возможность работать с готовыми белковыми молекулами.

Несмотря на разнообразие химических методов синтеза, данное направление имеет множество недостатков. Необходимость многоступенчатого синтеза неизбежно приводит к накоплению ошибок на каждой стадии. Часто это вызывает низкий выход целевого продукта и проблемы со стереоселективностью – образованием смесей энантиомеров, разделение которых требует сложного хроматографического оборудования и дополнительных затрат [40]. Низкие выходы, в свою очередь, ведут к накоплению значительного количества непрореагировавших исходных и побочных продуктов, что приводит к образованию неутилизированных отходов и загрязнению окружающей среды. Фторирование и окисление характеризуются низкой воспроизводимостью и требуют строго контролируемых условий, что усложняет стандартизацию процесса. Совокупность этих факторов делает химический синтез трудно масштабируемым. Кроме того, постоянная потребность в дорогостоящих реагентах и защитных группах в сочетании с необходимостью работы с агрессивными химическими веществами делает этот подход экономически затратным и не всегда оправданным для фармацевтического производства.

В связи с этим следующим этапом развития технологии стало использование биотехнологических подходов, при которых НПА включаются непосредственно в полипептидную цепь в ходе трансляции в клеточных или бесклеточных системах. Этот подход позволяет встраивать НПА непосредственно в генетически заданную последовательность белка и тем самым существенно расширить возможности его молекулярного конструирования.

### 1.1.3 Биотехнологические подходы получения белковых препаратов, содержащих неприродные аминокислоты в последовательности

Биотехнологические подходы к встраиванию НПА в аминокислотную последовательность белка представляют собой активно развивающееся направление современной фармацевтики [1]. Биотехнологические методы опираются на применение микроорганизмов, которые способны синтезировать белковые соединения из доступного и недорогого сырья, что исключает использование дорогостоящих катализаторов и растворителей, включая пиперидин и диметилформамид. Биологические подходы работают со строгой стереоспецифичностью, исключая образование нежелательных энантиомеров и необходимость их последующего разделения. Наконец, ферментативные и рибосомные системы встраивают НПА непосредственно в растущую пептидную цепь с сохранением необходимой конформации, тогда как при химическом синтезе это достигается только через многоступенчатые этапы [41]. Это позволяет получать целевые соединения из относительно доступных субстратов и проводить реакции в мягких водных условиях, что потенциально облегчает масштабирование и снижает потребность в органических растворителях и сложных стадиях защиты и деблокирования функциональных групп [42–44].

Для получения пептидов и белков, содержащих НПА, применяются несколько различных биотехнологических платформ. К ним относятся бесклеточные рибосомные системы трансляции, нерибосомный пептидный синтез (NRPS), хемобиологические и хемоэнзиматические стратегии, сочетающие химическую подготовку фрагментов с их ферментативным объединением [45–47].

Одной из наиболее гибких платформ является FIT–система (Flexible In vitro Translation), разработанная группой Хироаки Суги [45]. Это бесклеточная система трансляции, в которой используется реконструированный трансляционный аппарат и Flexizyme – короткая

искусственная молекула РНК (рибозим). Это ферменты способны присоединять НПА в виде эфира к выбранным тРНК, после чего полученные аминоацил–тРНК используются рибосомой при синтезе пептида при помощи очищенного набора факторов трансляции, ферментов, АТФ, извлеченных из клетки. В отличие от обычной аминоацил–тРНК–синтетазы, обладающей высокой специфичностью к своему природному субстрату, Flexizyme допускает значительно более широкий набор аминокислот и тем самым позволяет перепрограммировать трансляцию непосредственно *in vitro*.

Сочетание FIT с технологией мРНК дисплея привело к созданию платформы RaPID (Random non–standard Peptides Integrated Discovery), которая позволяет одновременно синтезировать и скринировать огромные библиотеки макроциклических пептидов [46]. Благодаря включению НПА и возможности последующей внутри– или межмолекулярной циклизации были получены библиотеки, содержащие структуры, значительно выходящие за пределы химического пространства обычных рибосомных пептидов. Теоретически размер таких библиотек может достигать порядка $10^{12}$–$10^{13}$ различных последовательностей, что делает RaPID особенно привлекательной платформой для поиска лигандов к белковым мишеням, которые трудно блокировать низкомолекулярными соединениями.

Нерибосомный пептидный синтез представляет еще один подход к внеклеточному биотехнологическому синтезу белков с НПА [48]. Он осуществляется крупными мультиферментными комплексами – нерибосомными пептид–синтетазами, организованными в функциональные модули. Классический модуль включает аденилирующий (A), пептидилпереносящий (PCP) и конденсационный (C) домены. A–домен распознаёт и активирует определённый мономер с образованием аминоацил–аденилата, после чего активированный остаток переносится на PCP–домен посредством 4–фосфопантетеиновой группы; C–домен катализирует образование пептидной связи между соседними мономерами [49]. Принципиальное отличие NRPS от рибосомного синтеза заключается в широком химическом диапазоне используемых блоков. Нерибосомные пептид–синтетазы способны включать D–аминокислоты, N–метилированные и β–аминокислоты, гидроксикислоты и другие непротеиногенные мономеры [45]. Дополнительные ферментативные домены обеспечивают эпимеризацию, гетероциклизацию, метилирование и другие посттрансляционные модификации, благодаря чему формируются сложные структуры природных пептидных антибиотиков и других биологически активных соединений [50].

Отдельную группу составляют хемоэнзиматические методы синтеза пептидов (CEPS, chemoenzymatic peptide synthesis), сочетающие химическое получение пептидных фрагментов с

их последующим ферментативным лигированием. CEPS подходит для получения длинных белковых препаратов, поскольку позволяет частично компенсировать ограничения последовательного SPPS, связанные с накоплением укороченных продуктов, агрегацией и снижением выхода при увеличении длины цепи. Такой подход позволяет разделить задачу синтеза на несколько стадий: короткие или средние фрагменты получают посредством SPPS, после чего соединяют с использованием пептидных лигаз. Одним из наиболее изученных инструментов является Omniligase–1 – инженерный вариант пептид–лигазы, способный катализировать образование пептидной связи между относительно крупными фрагментами в водной среде [47,51]. Экспериментально показана возможность применения Omniligase–1 для сборки пептидов терапевтического назначения. В частности, были разработаны варианты CEPS для синтеза GLP–1 агонистов, включая лираглутид и семаглутид, в которых химически полученные фрагменты объединяются ферментативно с последующей установкой липидного фрагмента. Для лираглутида описана ферментативная сборка 31–аминокислотной последовательности с последующим палмитоилированием боковой цепи Lys [52]. Экспериментальные данные показывают, что при оптимизации лигазы и положения соединения фрагментов возможно достижение высокой степени конверсии при проведении реакции в водной среде. Однако важно учитывать, что CEPS не устраняет полностью ограничения химического синтеза: необходимость предварительного получения фрагментов, их очистки и последующего лигирования сохраняется. Поэтому данный метод представляет скорее гибридную платформу, сочетающую химическую точность SPPS с селективностью ферментативного катализа, чем полноценную замену биосинтезу.

Несмотря на технологическое разнообразие FIT, NRPS и CEPS, их объединяет одно существенное ограничение: включение конкретной НПА в значительной степени зависит от субстратной специфичности используемой ферментной или трансляционной системы. Для Flexizyme спектр доступных аминокислот значительно шире, однако конкретные комбинации НПА, тРНК и условий трансляции требуют экспериментальной оптимизации. При применении NRPS состав продукта определяется специфичностью А–доменов и архитектурой ферментного комплекса. Используя метод CEPS выбор фрагментов и положение лигирования зависят от субстратной специфичности используемой лигазы.

Таким образом, перечисленные биотехнологические платформы значительно расширяют возможности получения НПА содержащих белков, однако для создания универсальной биотехнологической платформы для введения НПА в белки требуется высокая точность и масштабируемость процесса. Эту задачу решает расширения генетического кода, позволяющее

перепрограммировать трансляционный аппарат клетки так, чтобы выбранная НПА включалась рибосомой непосредственно в растущую полипептидную цепь.

## 1.2 Расширение генетического кода

Существует мнение, что современный набор аминокислот, вероятно, сформировался в результате длительного процесса расширения и последующей стабилизации системы трансляции. Аминоацил–тРНК–синтетазы и тРНК относятся к числу наиболее древних компонентов трансляционного аппарата, а их филогенетические и структурные особенности позволяют реконструировать отдельные этапы установления генетического кода [53,54]. Согласно эволюционным моделям, ранние системы трансляции могли использовать более ограниченный и химически отличный набор аминокислот, который постепенно расширялся. Последующая специализация АРС и тРНК, а также повышение точности распознавания аминокислот и кодонов способствовали фиксации современного генетического кода [55]. При этом нельзя исключать, что на ранних этапах эволюции в трансляционных системах могли использоваться аминокислотные мономеры, которые в современных организмах не относятся к каноническим 20 аминокислотам [56].

В современных организмах включение аминокислот в строго определённые положения полипептидной цепи обеспечивается взаимодействием кодонов мРНК, тРНК и соответствующих аминоацил–тРНК–синтетаз. Большинство организмов использует 61 смысловой кодон для кодирования 20 канонических аминокислот, тогда как UAA, UAG и UGA обычно выполняют функцию стоп–сигналов и обеспечивают завершение трансляции. Однако данное соответствие не является абсолютно универсальным. Существуют природные системы, в которых стоп–кодоны выполняют другую функцию: включение дополнительных аминокислот. Наиболее известными примерами являются селеноцистеин и пирролизин [57,58]. Именно этот природный феномен впоследствии был использован как концептуальная основа для разработки искусственного расширения генетического кода и создания систем, позволяющих направленно включать НПА в заданные положения терапевтических белков.

Расширение генетического кода представляет собой комплекс методов молекулярной биологии, позволяющих включать НПА в последовательность белков при помощи рибосомального синтеза [59]. Существует два основных варианта реализации расширения

генетического кода. Первая стратегия – включение НПА, являющихся структурными аналогами природных аминокислот, при помощи модифицированных эндогенных АРС клетки–хозяина [60,61]. Подобные системы были разработаны как на основе тирозил–тРНК–синтетазы *E. coli* (EcoTyrRS), так и на основе лейцил–тРНК–синтетазы (EcoLeuRS) путем изменения антикодона родственной тРНК и последовательных раундов отбора для смещения специфичности в сторону выбранных НПА [60,61]. Такая стратегия получила название переназначения смыслового кодона.

Вторая стратегия – сайт–специфическое включение НПА в ответ на стоп– или четырехбуквенный кодон экзогенными парами АРС и тРНК. Для сайт–направленного встраивания НПА в белки должно быть выполнено несколько условий. Одним из ключевых элементов является наличие свободного кодона. Обычно сайт–специфическое расширения генетического кода использует нонсенс–супрессорные тРНК, которые встраивают НПА в ответ на стоп–кодон [59]. В настоящее время в качестве такого кодона используется амбер стоп–кодон UAG [62]. Данный кодон мало распространен в большинстве организмов и редко используется в значимых белках, что позволяет задействовать его без ущерба для жизнеспособности клетки [63,64]. Другой подход – использование квадруплетных кодонов, который не требует перекодирования и предлагает потенциальное расширение с 64 до 256 кодонов [65,66].

Помимо свободного кодона важно наличие ортогональных пар АРС/тРНК [67]. Ортогональными АРС являются такие ферменты, которые не взаимодействуют с протеиногенными аминокислотами и тРНК клетки–хозяина, при этом комплементарны синтетическому аппарату клетки и избирательно распознают одну из НПА в качестве субстрата (Рисунок 1.2).

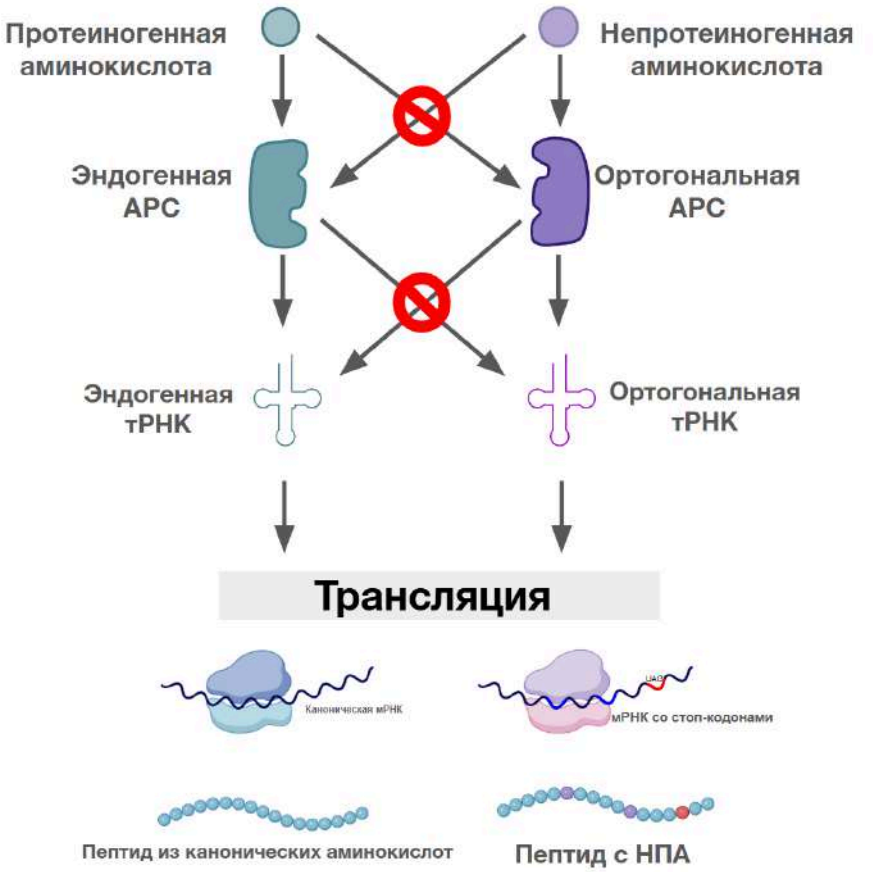



Рисунок 1.2 – Естественная трансляция, заканчивающаяся образованием пептида, содержащего природные аминокислоты. Ортогональная трансляция, приводящая к образованию пептида, содержащего непротеиногенную аминокислоту, которая включается при считывании квадруплетных кодонов (синий) или амбер стоп–кодона (красный). Перечеркнутые стрелки указывают на отсутствие или минимальную активность, что соответствует принципу ортогональности

### 1.2.1 Природное расширение генетического кода на примере селеноцистеина и пирролизина

К особой группе аминокислот, расширяющих стандартный набор из 20 протеиногенных аминокислот, относятся условно неприродные аминокислоты – селеноцистеин (Sec) и пирролизин (Pyl). Несмотря на их природное происхождение и генетически запрограммированное включение в состав белков, они обнаруживаются лишь у ограниченного числа организмов и используют нестандартные механизмы кодирования. Так включение Sec в

полипептидную цепь происходит в ответ на стоп–кодон UGA и требует наличия специфического цис–регуляторного элемента SECIS (selenocysteine insertion sequence), а также специализированных компонентов трансляционного аппарата [68]. В системе кодирования селеноцистеина UGA приобретает дополнительную функцию, выступая не только сигналом терминации, но и кодоном для включения Sec.

В отличие от Sec, Pyl представляет собой другой природный вариант расширения генетического кода, основанный на использовании специализированной пары PylRS/тРНК$_{Pyl}$. Включение Pyl в состав белков первоначально было выявлено у представителей семейства *Methanosarcinaceae*, археи которого используют метиламины в качестве субстратов для роста и экспрессируют Pyl–содержащие белки [58]. В условиях дефицита кислорода в ходе эволюции сформировались микробные консорциумы, включающие ацетогенные и метаногенные микроорганизмы, между которыми устанавливаются тесные метаболические взаимодействия. Метаногены, потребляя водород, поддерживают его низкую концентрацию и тем самым создают термодинамически благоприятные условия для жизнедеятельности ацетогенов, тогда как образующиеся ацетаты могут далее вовлекаться метаногенами в процессы метаногенеза [69]. Ключевую роль в метаболизме метиламинов при этом играют метилтрансферазы, в частности монометиламинметилтрансфераза (MMA–MT). Именно при анализе гена, кодирующего этот фермент, было обнаружено наличие стоп–кодона UAG непосредственно в рамке считывания. Первоначально такая организация гена казалась необычной, однако последующие исследования показали, что UAG в данном случае не приводит к терминации трансляции, а служит сигналом для включения Pyl [70].

### 1.2.2 Структурные особенности пирролизиновой аминоацил–тРНК–синтетазы

В 1998 году у *Methanosarcina barkeri* были выявлены гены, необходимые для включения Pyl в положение, соответствующее амбер стоп–кодону [69]. В дальнейшем гомологичные метилтрансферазы и компоненты системы биосинтеза и кодирования Pyl были обнаружены у других представителей архей и бактерий [58,71]. Было установлено, что включение Pyl осуществляется за счет совместного действия пирролизил–тРНК–синтетазы и пирролизиновой тРНК (тРНК$_{Pyl}$). PylRS селективно распознаёт свободный пирролизин и катализирует его присоединение к 3'–гидроксильной группе тРНК$_{Pyl}$, после чего аминоацилированная тРНК

доставляет Pyl к рибосоме и обеспечивает его включение в полипептидную цепь в ответ на кодон UAG в процессе стандартной трансляции [72]. Благодаря наличию отдельной ортогональной пары АРС/тРНК система пирролизина рассматривается как природный пример сайт–специфического расширения генетического кода, при котором дополнительная аминокислота генетически запрограммированно включается в строго определённое положение полипептидной цепи.

Следует отметить, что гены PylRS (*pylS*) и тРНК Pyl (*PylT*) расположены в общем кластере генов pyl, примыкающем к кластеру генов метиламин–метилтрансферазы. [58] (Рисунок 1.3). Также в кластере pyl расположены три других гена: *pylB*, *pylC* и *pylD*, вовлеченные в синтез Pyl из лизина (Lys). Ген pylB кодирует фермент radical S–adenosyl–L–methionine [73,74], который катализирует синтез (3R)–3–метил–D–орнитина из Lys; его карбоксилат затем связывается с амином боковой цепи Lys под действием продукта гена pylC, члена семейства карбамоилфосфатсинтетаз, при помощи АТФ. Получившийся в результате этой реакции (3R)–3–метил–D–орнитил–Nε–L–Lys затем подвергается окислению под действием продукта гена pylD, NAD–зависимой дегидрогеназы, после чего происходит одновременная внутримолекулярная конденсация с образованием Pyl [75].

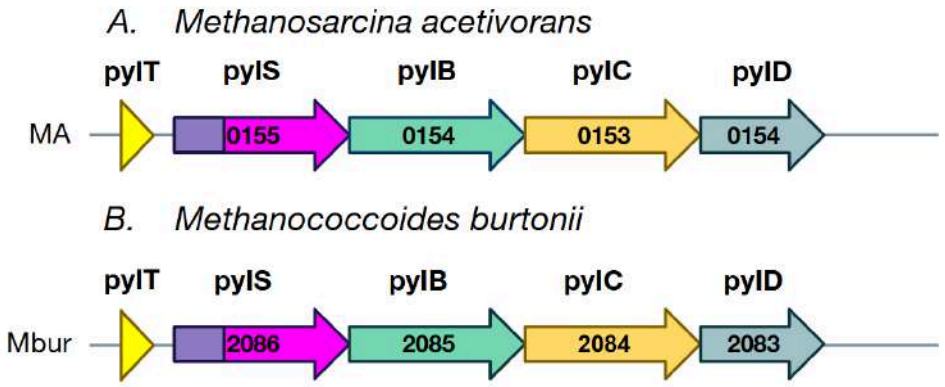



Рисунок 1.3 – Структура кластера Pyl в геномах двух представителей родов (А – *Methanosarcina*, В – *Methanococcoides*) семейства метаногенных архей *Methanosarcinaceae*. Число, которым помечен каждый ген, является номером локуса в аннотированном геноме. PylRS (PylS) и $тРНК_{Pyl}$ (PylT) участвуют в генетическом кодировании пирролизина. pylB, pylC и pylD участвуют в биосинтезе пирролизина

Исследования кристаллических структур PylRS из разных видов показали, что структура фермента напоминает другие аминоацил–тРНК–синтетазы класса II и в большинстве случаев состоит из двух доменов: С–концевого (PylSc) и N–концевого (PylSn). Как и другие АРС класса II, PylRS существует в виде димера и имеет три петли активного центра. PylSc является консервативным каталитическим доменом, его трехмерная организация напоминает организацию других синтетаз класса II: он также имеет ядро в виде семинитевого антипараллельного β–листа, окруженного несколькими α–спиралями, которые образуют складку Россмана для связывания АТФ [76]. Особое сходство наблюдается между PylRS и бактериальной фенилаланил–тРНК–синтетазой (PheRS), несмотря на то, что идентичность последовательностей довольно низкая (< 30%) [77]. Филогенетический анализ указывает на то, что PylRS произошел от PheRS вследствие дупликации гена PheRS, причем возникновение PylRS произошло еще до последнего универсального общего предка (LUCA) [78,79]. При этом у обоих ферментов гидрофобные карманы, в которых связываются циклоалкановая группа Pyl и фенильная группа фенилаланина (Phe), имеют одинаковую организацию. И у PylRS, и у PheRS внутренняя поверхность связывающих карманов образована ароматическими остатками, а сам карман закрыт петлей, также содержащей ароматический остаток (Tyr–384 в случае PylRS и Phe–260 в случае PheRS). Эта петля функционирует также как субстратспецифичный элемент в других синтетазах класса II для их родственных аминоацильных субстратов [77]. Однако существенное различие между PylRS и PheRS заключается в аминокислотном составе: PylRS содержит ароматические остатки, которые образуют водородные связи с пиррольным кольцом Pyl, в то время как PheRS содержит ароматические остатки, которые не способны к образованию водородных связей, что обусловлено фенильной группой у субстрата PheRS [80].

Наиболее важными аминокислотами связывающего кармана PylRS, определяющими специфичность, являются Tyr–384 и Asn–346. Взаимодействие с Tyr–384 позволяет Pyl принять правильную ориентацию; это достигается благодаря образованию водородных связей между γ–OH тирозина и азотом пиррольного кольца, а также α–аминогруппой [77]. При связывании правильного субстрата боковая цепь Tyr закрывает гидрофобный карман, полностью окружая Pyl. Asn–346 также образует водородные связи с Pyl через амидный азот боковой цепи N346 и амидный кислород боковой цепи Pyl [70]. Хотя PylRS прочно связывает Pyl, этот фермент способен ацилировать $тРНК_{Pyl}$ аналогами Pyl с различными модификациями боковой цепи субстрата, особенно если модификация приходится на область пирролина [81]. Это уникальная ситуация, поскольку другие аминоацил–тРНК–синтетазы относятся к аминоацилированию производными родственной аминокислоты гораздо более строго, а некоторые даже имеют

редактирующий домен, гидролизующий комплекс неправильной аминокислоты и тРНК. Однако поскольку PylRS распознает боковую цепь Pyl главным образом за счёт относительно неспецифических гидрофобных взаимодействий и не имеет редактирующего домена, он способен аминоацилировать крайне широкий спектр аминокислот [82].

Помимо С–концевого каталитического домена ферменты PylRS некоторых организмов обладают уникальным N–концевым тРНК–связывающим доменом, который, по–видимому, не имеет структурной или последовательной гомологии ни с одним известным РНК–связывающим белком, при этом демонстрирует высокое структурное сходство с Т– и вариабельными петлями $тРНК_{Pyl}$ [83]. Взаимодействие между PylRS и $тРНК_{Pyl}$ происходит таким же образом, как и с другими синтетазами класса II: PylRS приближается к $тРНК_{Pyl}$ со стороны основной бороздки акцепторного стержня, где формируются α–спираль тРНК–связывающего домена 1 и С–концевая область вместе с выпуклым доменом соседней субъединицы в димере PylSc, образуя сайт связывания, в котором расположена акцепторная спираль $тРНК_{Pyl}$ (рисунок 1.4).

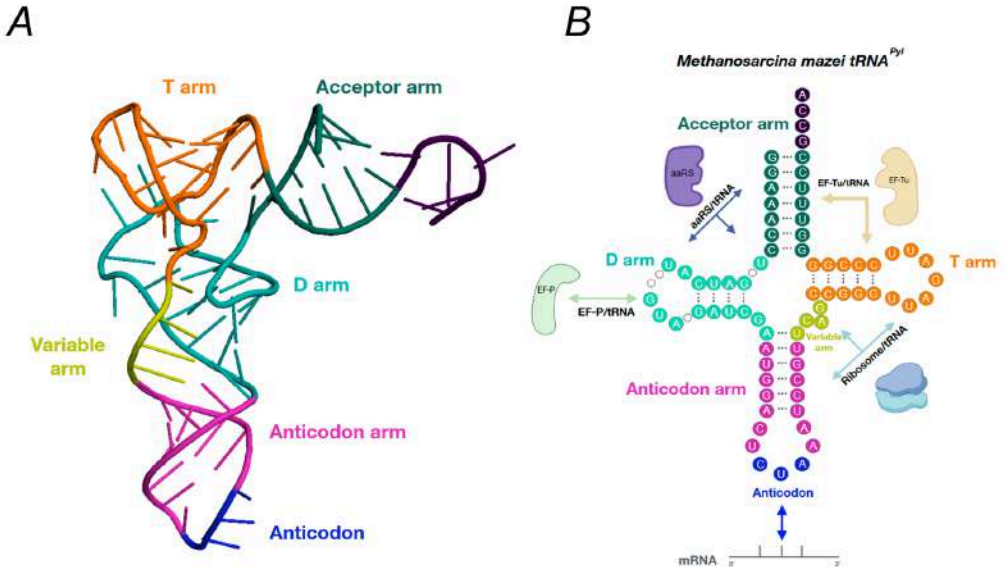



Рисунок 1.4 – Структура тРНК с доменами, выделенными разными цветами. (А) Трехмерная структура тРНК *Methanosarcina mazei* (PDB: 1N78) состоит из акцепторного плеча (зеленое), D–плеча (светло–зеленое), антикодонного плеча (фиолетовое), вариабельного плеча (желтое) и Т–плеча (оранжевое). (В) Вторичная структура структуры тРНК, использующая ту же цветовую гамму, с частями, которые взаимодействуют с различными факторами

Это взаимодействие направляет акцепторную спираль к каталитическому участку, где фермент, подобно другим синтетазам класса II, ацилирует 3'–OH акцепторного стержня [84]. В общей сложности 31 остаток бактериального Dh–PylSc из *Desulfitobacterium hafniense* участвует в водородных связях или стэкинг–взаимодействиях с 17 основаниями в Dh–тРНК$_{Pyl}$, при этом четыре из этих остатков принадлежат соседней субъединице димера PylSc [85].

Структурные особенности PylRS во многом определяют пригодность пирролизиновой системы для искусственного расширения генетического кода. В отличие от большинства природных АРС, PylRS сочетает широкую субстратную пластичность с высокой ортогональностью взаимодействия с трансляционным аппаратом клетки. Это позволяет использовать пару PylRS/тРНК$_{Pyl}$ в качестве универсальной платформы для сайт–специфического включения НПА в белки различных организмов.

Одним из ключевых свойств пары PylRS/тРНК$_{Pyl}$ является её высокая ортогональность. Компоненты данной системы практически не взаимодействуют с эндогенными аминоацил–тРНК–синтетазами и тРНК клеток–хозяев ни у бактерий, ни у эукариот, что позволяет использовать её в широком спектре организмов [83]. Другой важной особенностью PylRS является строение её активного центра. Фермент обладает относительно объёмным карманом связывания субстрата и не содержит отдельного редактирующего домена, характерного для ряда других АРС. Это обеспечивает высокую структурную и химическую пластичность активного центра и позволяет распознавать широкий спектр неприродных аминокислот, включая производные пирролизина с разветвленными функциональными группами [77,86]. Именно это свойство стало основой для направленной инженерии PylRS и создания её вариантов с заданной специфичностью к различным НПА.

Дополнительное преимущество связано с особенностями распознавания тРНК$_{Pyl}$. В природной системе эта тРНК является супрессором янтарного стоп–кодона UAG. При этом, в отличие от многих других АРС, PylRS не использует антикодон тРНК$_{Pyl}$ в качестве основного элемента идентичности [87,88]. Такая особенность имеет важное практическое значение: система может непосредственно применяться для перекодирования UAG без дополнительной модификации антикодона. Более того, антикодон тРНК$_{Pyl}$ может быть заменён без существенного нарушения аминоацилирования, что позволяет переназначать другие кодоны для включения НПА и расширять возможности генетического кодирования [78].

Таким образом, сочетание ортогональности, широкой субстратной специфичности, отсутствия антикодон–зависимого распознавания тРНК и природной способности к супрессии UAG делает пару PylRS/тРНК$_{Pyl}$ одной из наиболее универсальных платформ для расширения

генетического кода. Эти свойства, в свою очередь, создали основу для последующей инженерии PylRS и разработки систем, позволяющих контролируемо вводить широкий спектр НПА в заданные позиции белков.

### 1.2.3 Классификация PylRS

Существующее множество ферментов PylRS можно разделить на три класса: «PylSn+PylSc», слитые «PylSn–PylSc» и «ΔPylSn» (Рисунок 1.5). Отличия между классами связаны с особенностями экспрессии двух доменов PylRS. Класс слитых PylSc–PylSn (впервые обнаружен у *Methanosarcina mazei*, MmPylRS) характеризуется C– и N–доменами, соединенными линкером, тогда как у класса PylSn (впервые описан у *Desulfitobacterium hafniense*, DhPylRS) C– и N–домены являются двумя отдельными белками, которые после трансляции объединяются в единый PylRS. Последний класс, ΔPylSn (впервые обнаружен у *Candidatus Methanomethylophilus alvus*, MaPylRS), лишен тРНК–связывающего N–домена, но при этом остается высокоактивным в *Escherichia coli* [58,80,89].




Рисунок 1.5 – Доменная организация трех классов ферментов: «PylSn + PylSc», «слитые PylSn–PylSc» и «ΔPylSn»

В линкерной части ферментов класса слитых «PylSn–PylSc» среди представителей семейства *Methanosarcinaceae* наблюдается значительная вариабельность: например, у

*Methanococcoides burtonii* длина линкера составляет всего 14 аминокислот, в то время как у *Myceliophthora thermophila* линкер состоит из 72 аминокислот [90]. Расшифрованные на данный момент кристаллические структуры изолированного N–концевого домена PylRS из *M.mazei* показывают, что этот белок сворачивается в компактную глобулу, стабилизированную ионом цинка, и контактирует с T– и вариабельными петлями $тРНК_{Pyl}$, тогда как для кристаллических структур CTD показано взаимодействие с другой стороной $тРНК_{Pyl}$ (Suzuki et al., 2017). Подобная тесная связь между PylSn, PylSc и вариабельной петлей $тРНК_{Pyl}$ стерически объясняет ортогональность PylRS данного класса и высокую специфичность к $тРНК_{Pyl}$: более крупное вариабельное плечо канонических тРНК будет препятствовать их связыванию с PylRS [76,87]. Кроме того, в отличие от АРС для большинства канонических аминокислот, между антикодоном $тРНК_{Pyl}$ и PylSc или PylSn PylRS нет взаимодействий, что облегчает применение PylRS для кодирования неприродных аминокислот в различных кодонах [85].

В классе PylSn+PylSc домены экспрессируются из двух отдельных генов как два разных белка, после чего собираются в PylRS. Биологическая роль такого разделения неясна, как и роль PylSn в данном классе PylRS. Известно, что DhPylRS проявляет значительную активность *in vivo* даже при отсутствии экспрессии РНК–связывающего домена [91], что позволяет предположить, что его экспрессия не обязательна. Возможно, биологическая функция NTD заключается в привлечении $тРНК_{Pyl}$ к каталитическому домену – это подтверждается высокой аффинностью связывания РНК–связывающего домена с $тРНК_{Pyl}$ [83]. Также было показано, что помимо рекрутирования $тРНК_{Pyl}$ PylSn может стабилизировать её в растворе, способствуя её связыванию с CTD и последующему аминоацилированию.

Третий класс «ΔPylSn» был открыт недавно, и особенностью его представителей является полное отсутствие PylSn при сохранении активного PylSc [92]. Ферменты ΔPylSn часто используют, когда в белок необходимо включить сразу две различные НПА, поскольку они взаимно ортогональны ферментам из класса слитых «PylSn–PylSc» [93]. ΔPylSn обнаруживают у недавно описанных, преимущественно некультивируемых метилотрофных метаногенов 7–го порядка, получивших название *Methanomassiliicoccales* [80,94–96]. Из–за отсутствия РНК–связывающего домена и нетипичной структуры $тРНК_{Pyl}$ изначально было неясно, активен ли PylRS в этих организмах. Однако, как и в случае с другими известными Pyl–включающими организмами, ген, кодирующий MtmB, содержал кодон UAG в рамке считывания, что указывало на генетическое кодирование Pyl [95]. Наиболее перспективным среди данного класса PylRS был фермент MaPylRS, выделенный из *Candidatus Methanomethylophilus alvus*, который был значительно активнее *in vivo*, чем широко используемый MmPylRS [89]. Впоследствии

активность MaPylRS была подтверждена несколькими исследованиями [92,97,98], а кристаллическая структура фермента определена двумя независимыми группами (PDB: 6EZD, 6JP2). Интересно, что, несмотря на высокую активность этого фермента при отсутствии РНК–связывающего домена, в его кристаллической структуре не было обнаружено никаких примечательных особенностей: MaPylRS демонстрировал высокий уровень гомологичности и структурного выравнивания с MmPylRS и DhPylRS. Следовательно, высокая активность автономного каталитического домена может быть обусловлена особенностями $тРНК_{Pyl}$ этих организмов, которая из–за отсутствия PylSn перед реакцией аминоацилирования должна самостоятельно принять правильную ориентацию.

В настоящее время известно, что гены PylRS присутствуют у 156 видов различных анаэробных бактерий и архей [78]. Гомологи PylSc выявлены у 75 видов архей, относящихся к восьми типам: *Euryarchaeota*, *Thermoplasmatota*, *Asgardarchaeota*, *Hydrothermarchaeota*, *Thaumarchaeota*, *Bathyarchaeota*, *Verstraetearchaeota* и *Korarchaeota*. При этом из этих 75 идентифицированных генов большая часть (47) принадлежит к классу слитых ферментов «PylSn–PylSc». Кроме того, все организмы, имеющие слитые «PylSn–PylSc», относятся к семейству *Methanosarcinaceae*. Их родственники из того же отряда *Methanosarcinales*, *Methermicoccus shengliensis* и *Methanomicrobia archaeon*, несут ферменты классов «ΔPylSn» и «PylSn+PylSc» соответственно.

Для 21 PylRS, обнаруженных у архей, невозможно точно идентифицировать класс фермента, так как геномы этих организмов не прочитаны полностью. Однако, учитывая, что ни в одном из геномов не удалось обнаружить PylSn, можно предположить, что данные PylRS относятся к классу «ΔPylSn». Это предположение дополнительно подтверждается тем, что имеющиеся гены PylSc обладают высоким сходством с PylRS из класса «ΔPylSn». Большинство продуцентов этих синтетаз относится к отряду метаногенных архей *Methanomassiliicoccales*, обитающих в пищеварительной системе животных [99]. Помимо *Methanomassiliicoccales*, ферменты класса «ΔPylSn» также присутствуют у архей типов Euryarchaeota, Bathyarchaeota и *Thaumarchaeota*. Примечательно, что ферменты класса «PylSn+PylSc» редко встречаются у архей: из 75 исследованных генов они выявлены лишь в 7 случаях [78].

81 ген PylRS обнаружен в геномах бактерий. Чаще всего PylRS выявлялись у представителей типа *Firmicutes*, особенно у видов отряда *Clostridiales* (41 ген). Помимо *Firmicutes*, пирролизиновые синтетазы найдены у 14 дельтапротеобактерий, двух актинобактерий и одной спирохеты. Во всех случаях бактериальные PylRS принадлежали к классу «PylSn+PylSc».

С учетом таксономического распределения организмов, кодирующих Pyl, и результатов недавних филогенетических анализов предполагается, что бактерии приобрели способность инкорпорировать Pyl в результате горизонтального переноса генов от *Methanomassiliicoccales*, которые, в свою очередь, унаследовали эту способность от предка *Methanosarcinaceae* [78,80,100,101].

## 1.3 Методы инженерии PylRS для расширения спектра включаемых НПА

PylRS обладает рядом структурных и функциональных особенностей, благодаря которым эта система стала одной из наиболее широко используемых платформ для сайт–специфического включения НПА. Совокупность этих свойств позволяет модифицировать субстратную специфичность PylRS без существенного нарушения взаимодействия с тРНК и использовать одну пару PylRS/тРНК$_{Pyl}$ для включения большого числа структурно различных НПА [77,102]. На сегодняшний день с использованием PylRS–систем в белки генетически включено уже более 300 различных НПА и других, включая α–гидроксикислоты и не–α–аминокислоты [79]. Современные стратегии инженерии PylRS включают насыщенный мутагенез нескольких положений активного центра, рациональное конструирование вариантов и создание химерных ферментов, объединяющих структурные элементы различных классов PylRS [103].

### 1.3.1 Структурная модификация PylRS для изменения субстратной специфичности

Первые исследования преимущественно были сосредоточены на использовании нативной PylRS и её вариантов для включения производных пирролизина. Структурный анализ MmPylRS показал, что активный центр фермента формирует относительно объемный карман связывания с аминокислотой, способный вместить боковые цепи различного размера и химической природы [77]. Последующие исследования подтвердили, что такая архитектура обеспечивает высокую полиспецифичность PylRS и позволяет использовать её для включения НПА, существенно отличающихся от природного субстрата [77,102]. В частности, были получены варианты PylRS, способные распознавать НПА с азидными, алкиновыми,

алкеновыми, кетоновыми и другими функциональными группами, что расширило возможности последующей биортогональной модификации белков [83,104,105].

Важным этапом развития ортогональных систем на основе PylRS стало установление структурного сходства с бактериальной PheRS [106]. Это наблюдение позволило предположить возможность изменения субстратной специфичности PylRS в сторону ароматических аминокислот. Существенное значение имели остатки N346 и C348 (для MmPylRS): изменение этих позиций приводило к перестройке субстрат–связывающего кармана и изменению его способности размещать объемные ароматические заместители. Полученные варианты позволили включать в белки производные фенилаланина с алкиновыми, азидными, кетоновыми и галогенированными ароматическими группами [105,107,108].

Позже, наряду с ароматическими аминокислотами, последовательное изменение аминокислотных остатков, формирующих активный центр PylRS, позволило получить варианты фермента с повышенной эффективностью аминоацилирования различных производных лизина [102,109]. Особенно показательным примером является генетически программируемое включение Nε–ацетиллизина (AcK) – природной посттрансляционной модификации лизина. Создание соответствующих данной НПА вариантов PylRS позволило вводить AcK непосредственно в заданное положение рекомбинантного белка и тем самым исследовать функциональные последствия ацетилирования без необходимости сложной химической модификации [109].

В дальнейшем было показано, что субстратная пластичность PylRS выходит далеко за пределы производных тирозина и лизина. Так, были идентифицированы варианты PylRS, способные включать НПА с существенно меньшими боковыми цепями, что потребовало целенаправленного уменьшения объема субстрат–связывающего кармана [110].

### 1.3.2 Рациональная инженерия С–концевого каталитического домена PylRS

Структурные исследования показали, что С-концевой домен фермента содержит каталитический центр и основной участок связывания аминокислоты, поэтому изменение отдельных остатков в этой области позволяет регулировать размер, гидрофобность и геометрию субстратного кармана. Благодаря расшифровке кристаллических структур PylRS были идентифицированы аминокислотные остатки в активных центрах, точечные изменения в

которых позволили расширить репертуар встраиваемых НПА [76]. Так, важными для связывания субстрата PylRS, а также её гомологов, считаются следующие положения в субстратсвязывающем кармане: Tyr306, Leu309, Asn346, Cys348 и Tyr384 – именно они образуют сеть водородных связей с Pyl [111–114].

Одними из первых были получены данные о том, что всего двух аминокислотных замен Asn346Ala и Cys348Leu (или Cys348Lys) достаточно, чтобы Mm PylRS имел высокую активность по отношению к Phe [83,106,115,116]. Было показано, что при введении замены Asn346Ala значительно уменьшается связывание PylRS с Pyl и устраняются стерические препятствия для связывания фенилаланина [117]. При введении данной замены остается большое пустое пространство в активном сайте вокруг пара–положения Phe, что позволяет использовать мутанта Asn346Ala PylRS для встраивания больших НПА, содержащих разветвленные боковые цепи. Замена Cys348 на аминокислоту с более крупной боковой цепью, которая занимает пространство пирролидина Pyl в активном центре, приводит к увеличению взаимодействий со связанным Phe. Важным наблюдением оказалось то, что замена Cys348 на аминокислотный остаток меньшего размера позволяет распознавать производные Phe с большими пара–заместителями [104]. Двойная замена Cys348Ala и Asn346Ala (PylRS–AA) приводит к увеличению кармана активного центра, который обеспечивает пара–замещенному фенилаланину большую структурную гибкость для связывания с активным сайтом [105]. Мутант PylRS–AA имеет широкий спектр субстратов, включающий производные Phe с разнообразным диапазоном орто– , мета– и пара – заместителей, а также производные гистидина, однако хуже ацилирует Phe [118].

Также было показано, что аминокислотные замены Tyr384Phe или Tyr384Trp увеличивают активность PylRS в отношении как Pyl, так и НПА [101]. Существует предположение, что водородная связь Tyr384 закрывает активный сайт, который служит для защиты нестабильного пирролизиладенилата, когда клеточная концентрация тРНК Pyl низка [99] и вступает в специфические водородные связи как с азотом пиррола, так и с α–аминогруппой субстрата [77]. Эти данные позволяют предположить, что водородная связь между α–аминогруппой субстрата и Tyr384 формируется гибко и не важна для каталитической активности PylRS. Следовательно, PylRS с заменой в данной позиции может распознавать большее количество субстратов, чем нативный фермент.

Еще одним важным положением в PylRS для увеличения субстратной специфичности к НПА является Tyr306Ala. Сообщалось, что сконструированный мутант PylRS Tyr306Ala имеет расширенный карман связывания, а аминокислотная замена остатка тирозина в фенилаланин

(Tyr384Phe) увеличивает эффективность встраивания НПА *in vivo*. Таким образом подобный двойной мутант может аминоацилировать аминокислоты с разветвленными боковыми цепями. Соответственно, белки могут быть модифицированы с помощью напряженных циклооктенов и октинов [119,120]. Было также показано, что двойной мутант может встраивать фурансодержащие НПА, которые значительно больше, чем Pyl [121].

Также известно, что замена лейцина в 309 положении на меньший по размеру остаток аланина (Leu309Ala) создает пространство для взаимодействия гидрофобных и объемных остатков в положениях 348 и 306. В одном из исследований совместно аминокислотные замены Cys348Phe и Leu309Ala приводят к образованию активного сайта AcKRS, который меньше и более гидрофобен, чем у нативного PylRS, что объясняет, почему AcKRS не проявляет активность в отношении более крупного субстрата Pyl [122].

Все перечисленные выше замены улучшают эффективность встраивания неприродных аминокислот не только в классе слитых «PylSn–PylSc», но и в других классах фермента. Так например были разработаны пары ISO4–G1 PylRS [72]. На основании выравниваний последовательностей Ala121, Leu124, Tyr125, Met128, Asn165, Val16, Leu227 и Trp237 в аминокислотном связывающем кармане ISO4–G1 PylRS соответствуют ранее описанным положениям Ala302, Leu305, Tyr306, Leu309, Asn346, Cys348, Leu407 и Trp417 соответственно в Mm PylRS. Для включения объемных производных Lys, таких как ZLys, mAzZLys, N ε –(p –азидобензилоксикарбонил)–L–лизин (pAzZLys), N ε –(p –этинилбензилоксикарбонил)–L–лизин ( pEtZLys) и TCO*Lys, аминокислотные замены в положениях Tyr306 и Leu309 в Mm PylRS были транспонированы в ISO4–G1 PylRS. Как упоминалось ранее аминокислотная замена Tyr125Ala (Tyr306Ala Mm PylRS) увеличивает карман активного центра в том числе в ISO4–G1 PylRS, который затем становится пригодным для размещения громоздких НПА. В ISO4–G1 PylRS боковая цепь Met128 (Leu309 в MmPylRS) выступает в карман связывания аминокислот, что уменьшает размер кармана по сравнению с Mm PylRSc. Замены Leu и Ala в положении 128 увеличивают внутреннее пространство кармана активного центра.

Таким образом, модификации в активном центре PylRS позволили включать в белки НПА, содержащие алкеновые, алкиновые, азидные, бензольные, фурановые и серосодержащие группировки [52,70,76,77].

Классический рациональный подход к проектированию, основанный на знании трехмерной структуры, часто нацелен на ограниченный набор ключевых аминокислот в активном центре [124]. Наиболее очевидными мишенями являются остатки, формирующие субстрат–связывающий карман, поскольку изменение их свойств непосредственно влияет на

распознавание НПА. Однако исследования PylRS показали, что функционально значимые аминокислотные замены могут располагаться и за пределами каталитического центра. В частности, изменения N–концевого домена способны существенно повышать эффективность включения НПА, несмотря на то что данный домен непосредственно не участвует в формировании активного центра [110].

### 1.3.3 Рациональная инженерия N–концевого домена и междоменного линкера

Хотя большинство исследований по разработке модифицированных PylRS нацелены на карман связывания субстрата, замены в других частях фермента также могут улучшить включение НПА. Поскольку замены в N–концевом домене и линкерной части между доменами не влияют на распознавание субстрата, их можно трансплантировать в другие ранее идентифицированные варианты PylRS, чтобы обеспечить аналогичные улучшения активности.

Как говорилось ранее, Т петля и вариабельная петля тРНК$_{Pyl}$ имеют гидрофильные взаимодействия с PylRSn, а его усечение почти полностью исключает специфичность связывания PylRS с тРНК$_{Pyl}$ [125]. Так крайне важным аспектом в классе слитых ферментов является наличие гибкой петли между С– и N–концевыми доменами. Было показано, что добавление линкера обеспечивает дополнительную гибкость для потенциального регулирования взаимодействия между доменами PylRS и изменяет интерфейс между PylRS и тРНК$_{Pyl}$, в свою очередь, повышая эффективность включения НПА [126]. Как говорилось ранее, PylRSn менее растворима, чем PylRSc, поэтому расщепление на два домена снижает эффективность фермента. Однако стало известно, что в бактериальных клетках MmPylRS расщепляется после транскрипции в линкерной части в положениях Lys110 и Ala189. Аминокислотная замена Pro188Gly поддерживает структуру PylRSn в правильном положении, что нарушает распознавание протеазой в Ala189 и не дает PylRS расщепиться на два домена [127]. Следовательно для класса слитых ферментов модификация линкера позволяет дополнительно регулировать взаимодействия между доменами и изменяет связывание с тРНК$_{Pyl}$, в свою очередь, повышая эффективность аминоацилирования [128].

Также было обнаружено, что эволюционировавший PylRS, несущий три аминокислотные замены в PylRSn Arg19His/His29Arg/Thr122Ser, повышает эффективность подавления стоп–кодона при добавлении BocK [129]. Еще одной заменой, которая увеличивает

каталитические свойства PylRS и сродство к тРНК$_{Pyl}$, является  His62Tyr. Предполагается, что His62Tyr нарушает две водородные связи между PylRS и фрагментами Т–петли тРНК$_{Pyl}$, хотя также устанавливает новые более слабые контакты с фосфатным фрагментом Gly21 и основанием A20 [87].

Также для увеличения эффективности аминоацилирования новой аминокислотой необходимы аминокислотные замены Arg61Lys, His63Y и Ser193Arg. Arg61Lys и His63Tyr находятся в PylRSn на границе между N–концевой частью и областью связывания Т–петли тРНК. В частности, боковая цепь гуанидина Arg61 имеет полярное взаимодействие с фосфодиэфирным остовом около 58A тРНК$_{Pyl}$, а также с остатком Arg52 в пределах PylSn. Кроме того, боковая цепь имидазола His62 образует водородные связи с аминогруппой боковой цепи Lys85. Аминокислотные замены Arg61Lys и His62Y могут нарушить эти взаимодействия. Замена Ser193Arg находится в С–каталитическом домене и расположена на границе с областью связывания D–петли тРНК. Хотя Ser193 находится относительно далеко от PylRSn, он не взаимодействует с остатками рядом с каталитическим сайтом. Было выдвинуто предположение, что за счет этой замены каталитическое ядро PylRS расширяется локально, чтобы НПА с более короткими объемными боковыми цепями могли связаться с тРНК$_{Pyl}$ [126].

Важным следствием рационального дизайна PylRS для применения в расширении генетического кода является то, что подобные манипуляции не всегда приводят к формированию строго моноспецифичной АРС. Показано, что отдельные эволюционировавшие варианты PylRS способны распознавать широкий спектр структурно различных НПА. Для одного из таких вариантов была продемонстрирована активность в отношении библиотеки из 313 НПА, а структурный анализ показал возможность связывания различных субстратов в нескольких альтернативных ориентациях [130]. Полученный результат имеет принципиальное значение для инженерии генетического кода, поскольку показывает, что создание отдельной АРС для каждой НПА не всегда необходимо. Вместо этого можно получить полиспецифичный вариант PylRS, способный распознавать целое семейство структурно родственных соединений. Вместе с тем высокая полиспецифичность часто сопровождается снижением эффективности аминоацилирования отдельных субстратов. Поэтому дальнейшая инженерия PylRS направлена не только на расширение субстратного спектра, но и на восстановление каталитической активности в отношении конкретных НПА. Для повышения эффективности расширения генетического кода необходимо сосредоточиться на повышении растворимости, активности аминоацилирования и совместимости с тРНК, рибосомой и другими компонентами трансляции

в клеточных системах. Именно сочетание этих требований делает инженеринг АРС одной из центральных задач расширения генетического кода [79,131].

### 1.3.4 Направленная эволюция аминоацил–тРНК–синтетаз

В случаях, когда рациональное прогнозирование оптимальных аминокислотных и нуклеотидных замен затруднено, используют направленную эволюцию. Её принцип заключается в создании библиотеки вариантов PylRS с последующим отбором или скринингом ферментов, демонстрирующих требуемую активность в отношении выбранной НПА. Преимущество такого подхода состоит в возможности выявлять комбинации замен, эффект которых трудно предсказать по структуре белка.

В общем случае направленная эволюция АРС включает создание клонов, селекцию или скрининг функциональных вариантов с последующим повторением этих этапов. Для формирования библиотек используют случайный мутагенез, error–prone PCR, насыщенный мутагенез отдельных аминокислотных позиций, рекомбинацию генетических фрагментов, а также комбинации рационального дизайна с последующей экспериментальной оптимизацией [131]. Процесс направленной эволюции АРС обычно проводят *in vivo*, так как основным требованием к модифицированным парам АРС/тРНК является их ортогональность, предотвращающая перекрестные реакции аминоацилирования, что быстрее и проще проверять именно в *in vivo* системах. Вероятность обнаружения нового варианта АРС, способного эффективно встраивать НПА в белки, повышается по мере увеличения разнообразия библиотек АРС и при использовании фенотипического признака, позволяющего точно и быстро определять предпочтительные варианты АРС.

Отбор часто основан на использовании репортерных или селекционных систем, в которых функциональность АРС связывается с подавлением стоп–кодона. Например, в последовательность гена селекционного маркера вводят амбер–кодон UAG, а эффективность исследуемого варианта АРС оценивают по восстановлению экспрессии соответствующего белка. В разработанных для PylRS системах использовались, в частности, хлорамфеникол–ацетилтрансфераза и барназа, что позволяло проводить последовательные положительные и отрицательные раунды отбора. Дополнительное применение флуоресцентных

репортеров и FACS позволяет не только отбирать функциональные варианты, но и количественно сравнивать эффективность включения НПА в больших библиотеках [102,131].

Следующим этапом развития методов инженерии АРС стало применение непрерывной направленной эволюции, прежде всего PACE. Его ключевая особенность заключается в том, что он позволяет провести более 30 циклов эволюции белка в течение 24 часов без вмешательства исследователей. Возможность проводить лабораторную эволюцию в автоматическом или самоподдерживающемся режиме, с использованием природоподобных подходов, основанных на использовании принципов организации процессов эволюционного давления, может значительно повысить эффективность получения оптимальных вариантов белка. В системе PACE функциональная активность исследуемой АРС непосредственно связана с способностью бактериофага к размножению. Использование PACE позволило преодолеть одно из ограничений традиционной направленной эволюции – относительно небольшое число циклов отбора. При стандартном подходе обычно проводят несколько последовательных раундов положительного и отрицательного отбора, каждый из которых требует создания библиотеки, трансформации клеток, выделения клонов и их последующего анализа. В результате число поколений эволюции остается ограниченным человеческим фактором, а полученные АРС нередко характеризуются сниженной каталитической эффективностью по сравнению с исходными ферментами [132]. PACE, напротив, позволяет проводить сотни поколений эволюции в течение относительно короткого времени, что существенно увеличивает доступное эволюционное пространство.

Таким образом, направленная эволюция постепенно перешла от преимущественного рационального поиска аминокислотных замен в активном центре к комплексной оптимизации всей молекулярной архитектуры АРС.

### 1.3.5 PACE как метод направленной эволюции PylRS

Для получения белков с измененными свойствами используют методы направленного мутагенеза, позволяющие имитировать естественную эволюцию в лабораторных условиях и объединять процессы мутагенеза и отбора наиболее удачных белков. Соответственно, основным преимуществом направленной эволюции по сравнению с классическим мутагенезом является минимизация участия человека [133].

Вместо того, чтобы связывать активность белков–мишеней (в данном случае аминоацил–тРНК–синтетазы) с ростом клеток–хозяев, что может легко привести к фоновым заменам в результате адаптивной эволюции, PACE связывает активность синтетазы с производством способного к заражению фага M13. В методе фаг–ассоциированной эволюции активность белка интереса связывается с жизненным циклом фага так, чтобы при повышении уровня активности или специфичности целевого белка, бактериофаги, производящие такой вариант, получали эволюционное преимущество над остальными [134]. Этот подход использует особенности жизненного цикла нитевидных бактериофагов M13, в которых белок pIII играет ключевую роль [135]. Этот белок кодируется геном gIII и участвует как в начальных стадиях инфекции, так и в заключительных стадиях сборки, высвобождения и отсоединения вириона от клеточной мембраны. Для реализации метода PANCE ген gIII в бактериофаге (SP) заменяется геном, кодирующим целевой белок PylRS. Вместо фага gIII содержится на вспомогательной плазмиде (AP) в клетках–хозяевах, и его экспрессия запускается желаемой активностью синтетазы. Варианты SP, которые индуцируют более высокий уровень продукции pIII, будут заражать большее количество клеток и сохранятся в работе, в то время как нефункциональные или низкоактивные варианты SP выходят из отбора [134]. Повышение эффективности PylRS с новым субстратом увеличивает скорость продукции белка pIII и, следовательно, скорость репликации фага.

Эволюционный процесс в PACE можно условно разделить на два взаимосвязанных этапа: генерацию генетического разнообразия за счет накопления замен и последующий отбор вариантов с требуемым фенотипом. В условиях PACE изменения возникают непосредственно в закодированном в геноме бактериофага гене интереса, поскольку именно он непрерывно проходит циклы репликации и отбора. Однако естественная частота возникновения нуклеотидных замен у *E. coli* относительно невысока, вследствие чего без дополнительной стимуляции формирование библиотеки вариантов происходило бы недостаточно быстро для эффективной направленной эволюции. Для повышения скорости возникновения нуклеотидных замен в системе PACE используют специальную плазмиду мутагенеза (MP, mutagenesis plasmid), содержащую набор генов, изменяющих точность репликации и репарации ДНК [136].

В состав MP могут входить *dnaQ926*, *dam* и *seqA*, а также дополнительные гены мутагенеза. Ген *dnaQ926* кодирует мутантную ε–субъединицу ДНК–полимеразы III со сниженной 3′→5′–экзонуклеазной корректирующей активностью, что приводит к повышению частоты ошибок при репликации ДНК и, соответственно, накоплению нуклеотидных замен [137]. Мутации в dam, кодирующем ДНК–аденинметилазу, нарушают нормальное

метилирование адениновых остатков в последовательностях GATC и изменяют функционирование систем репарации и рекомбинации ДНК, способствуя повышению генетической вариабельности [138]. Дополнительный вклад вносит модификация seqA, продукт которого участвует в регуляции повторной инициации репликации; нарушение его функции изменяет динамику репликации и усиливает мутагенез [139]. В совокупности эти компоненты создают в клетке условия для ускоренного накопления генетических вариантов, которые затем подвергаются отбору в ходе PACE.

Последовательности генов мутагенеза размещают на отдельной плазмиде под контролем индуцируемого промотора арабинозного оперона, что позволяет регулировать интенсивность мутагенеза [136]. Такая регуляция имеет значение, поскольку постоянная гипериндукция мутагенеза может снижать жизнеспособность клеток и одновременно увеличивать вероятность накопления нежелательных изменений в геноме *E. coli*. В то время как ременное и контролируемое включение мутагенеза позволяет поддерживать баланс между скоростью замен и сохранением жизнеспособности клеточной популяции, что является одним из ключевых условий эффективной работы PACE.

Следует отметить, что для реализации PACE требуется сборка системы с камерой, в которую постоянно подается культура свежих клеток–хозяев и обеспечивается их непрерывный поток. Создание такой системы и её автоматизация могут вызвать затруднения. Однако существует упрощенный с технической точки зрения вариант – PANCE (phage–assisted non–continuous evolution) [86]. Этот подход представляет собой прерывистую версию PACE, при которой циклы мутагенеза и отбора проводятся вручную путем перезаражения бактериофагов из одной пробы в следующую. При этом нет необходимости в сложной аппаратуре, что делает PANCE доступным для большего числа лабораторий, но более трудоемким и длительным с точки зрения реализации [77, 86].

Процесс PANCE и PACE состоит из поочередно сменяющих друг друга этапов положительной и негативной селекции. На этапе положительной селекции главная задача состоит в получении фагов, белок интереса которых проявляет наибольшую активность. На этапе отрицательной селекции происходит коррекция и отбор вариантов фагов с наименьшей нецелевой активностью. В случае аминоацил тРНК синтетаз это соединение тРНК с нецелевыми аминокислотами [132].

В основе дизайна PACE и PANCE для направленной эволюции АРС используются две различные стратегии, благодаря которым будет происходить супрессия амбер–кодона и экспрессироваться pIII, необходимый для фаговой инфекции. Особенность первой стратегии

заключается в том, что амбер–кодон находится в гене Т7–РНК–полимеразы, которая осуществляет транскрипцию gIII с промотора Т7. При этом ген Т7–РНК–полимеразы лежит на дополнительной плазмиде (complementary plasmid, CP). В случае, когда в ответ на стоп–кодон происходит встраивание аминокислоты в последовательность Т7–РНК–полимеразы, осуществляется экспрессия gIII. Этот подход приводит к усиленной продукции pIII, поскольку каждое событие подавления амбер–кодона дает начало множеству транскриптов gIII. В рамках первой стратегии PACE были подобраны специальные участки в Т7–РНК–полимеразе, замены которых на НПА из широкого спектра вариантов не приведёт к ингибированию ферментативной активности. Также было выяснено число амбер–кодонов в гене Т7–РНК–полимеразы, необходимое для осуществления полноразмерной трансляции полимеразы, полностью зависимого от ортогональной пары АРС/тРНК.

Вторая, более строгая стратегия, отличается тем, что преждевременный стоп–кодон находится уже в gIII, соответственно, CP в данной системе отсутствует. Супрессия амбер–кодона приводит к прямой трансляции полноразмерного pIII и получающийся фаг обладает способностью заражать бактериальные клетки. Для реализации второй стратегии отбора преждевременные UAG стоп–кодоны находились в gIII, аминокислотные остатки были выбраны так, чтобы они не участвовали в связывании pIII ни с белком TolA клетки хозяина, ни с F–пилями клетки–хозяина, соответственно, их изменение не приведет к нарушениям фаговой способности к заражению. При этом для давления отбора было достаточно наличие одного преждевременного стоп кодона в кодирующей последовательности gIII, чтобы размножение бактериофага зависело от активности АРС.

Если на первых этапах отбора, называемых в совокупности положительной селекцией, достигается получение наиболее активных АРС, то на следующем этапе, именуемым отрицательной селекцией, необходимо элиминировать синтетазы, обладающие неспецифической активностью. Отрицательный отбор осуществляется при помощи доминантно–отрицательной версии pIII, pIII–neg (N–C83, содержит интактные домены N1 и N2, но имеет мутантный домен С с внутренней делецией 70 аминокислот) [132]. Остаточный домен С достаточен для обеспечения прикрепления к фаговой частице, но не может катализировать отсоединение зарождающегося фага от мембраны клетки–хозяина. Отрицательный отбор проходит в отсутствии НПА в среде. При таком давлении отбора нежелательная активность АРС приводит к синтезу белка gIII–neg, при помощи которого фаги, несущие субоптимальные АРС, выходят из отбора и более не могут заражать клетки.

Метод PACE показал себя удобным для изменения свойств АРС, однако его использование не ограничено лишь эволюцией синтетаз. Использование направленной фаг–ассоциированной эволюции доступно для многих белков, а сам метод может быть модернизирован в зависимости от преследуемой цели.

Развитие методов инженерии PylRS демонстрирует последовательный переход от модификации отдельных аминокислотных остатков к многоступенчатой оптимизации всей ортогональной системы. Инженерия PylRS представляет собой совокупность взаимодополняющих стратегий, включающих структурно–рациональный дизайн, направленную и непрерывную эволюцию, модификацию N–концевого домена и линкерных областей, а также вычислительное прогнозирование. На начальном этапе рациональный дизайн используется преимущественно для изменения геометрии субстрат–связывающего кармана и расширения спектра распознаваемых НПА. Последующая модификация N–концевого домена и линкерных участков позволяет дополнительно оптимизировать взаимодействие синтетазы с тРНК и повысить эффективность включения НПА в процессе трансляции. Методы направленной эволюции, включая PACE, в свою очередь позволяют выявлять комбинации замен, которые трудно предсказать исключительно на основании структурных моделей. Дополнительные возможности предоставляет применение вычислительных методов, включая машинное обучение, для предварительного прогнозирования наиболее перспективных вариантов PylRS.

Комплексное применение этих подходов позволяет одновременно расширять спектр доступных НПА и поддерживать высокие уровни селективности и эффективности их включения.

### 1.4 Способы декорации липидных наночастиц

Наночастицы обладают многочисленными преимуществами как носители лекарственных средств, включая наноразмерность, высокое отношение площади поверхности к объему, возможность избирательного нацеливания и контролируемое высвобождение. Системы липидных наночастиц (ЛНЧ) предлагают биосовместимый метод доставки и быстро внедряются в РНК-терапию [140]. Успешное использование мРНК-ЛНЧ в вакцинах против инфекционных заболеваний открыло широкие перспективы для их применения в иммунотерапии рака и редактировании генома [141]. Основные преимущества ЛНЧ как невирусной системы доставки включают более низкую иммуногенность, отсутствие риска интеграции вирусного генома и повышенную биосовместимость [142–144]. Однако ключевым ограничением стандартных ЛНЧ остается отсутствие тканеспецифичности: при системном

введении большая часть наночастиц неспецифически накапливается в печени и селезенке [145,146]. Современные попытки изменить биораспределение ЛНЧ направлены на модификацию липидного состава [147,148]. Такие изменения могут влиять на взаимодействие ЛНЧ с клеточными мембранами и перенаправлять их в другие клеточные популяции, но эту стратегию трудно адаптировать для воздействия на отдельные типы клеток. Чтобы преодолеть это ограничение, для целенаправленной доставки необходимо модифицировать или декорировать поверхность ЛНЧ лигандами, распознающими специфические рецепторы на клетках-мишенях [149]. Наночастицы могут быть функционализированы белками путем адсорбции, ковалентного связывания или использования молекул-адаптеров. При иммобилизации антител конъюгация должна обеспечивать требуемое количество биомолекул на одну наночастицу и их правильную ориентацию [150].

Связывание лиганда с поверхностью ЛНЧ чаще всего осуществляется за счет ковалентной конъюгации с функционализированными молекулами полиэтиленгликоля (ПЭГ), входящими в состав липидной оболочки. Традиционные подходы к ковалентной сшивке белков с липидными компонентами ЛНЧ опираются на классические реакции карбодиимидного сочетания, амидного связывания с помощью активированных эфиров или тиол–малеимидной химии [151] (рисунок 1.6).

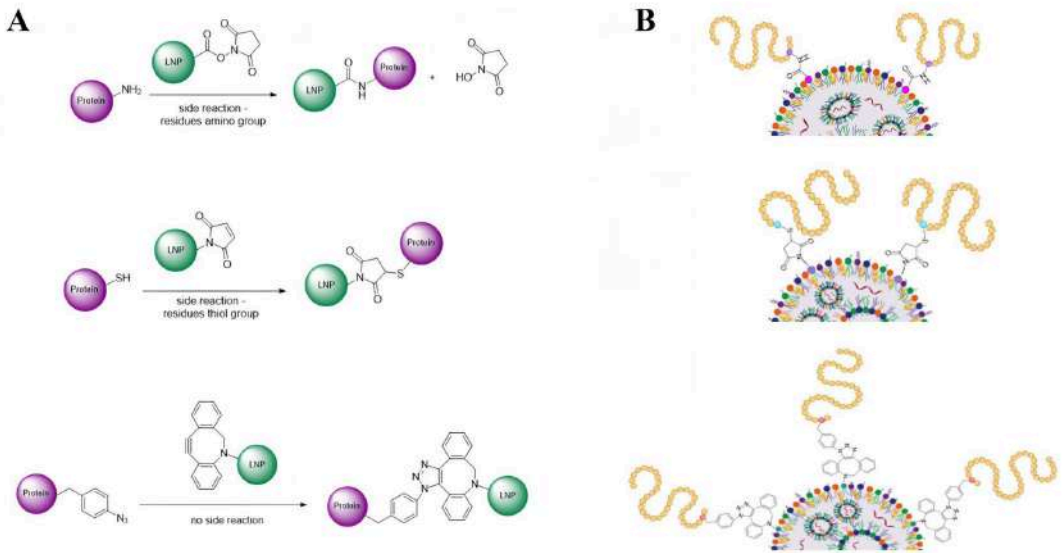



Рисунок 1.6 – (A) Различные методы конъюгации для присоединения белков к ЛНЧ. (B) Схематическое изображение белка на поверхности ЛНЧ при разных типах конъюгации

Модификация по остаткам лизина с использованием липидов, активированных N–гидроксисукцинимидными эфирами, либо прямая сшивка карбоксилированных липидов с белком с помощью EDC/NHS, является наиболее доступным методом [152]. Однако, согласно литературным данным, этот подход обладает рядом недостатков. Из–за большого количества доступных аминогрупп на поверхности реакция протекает стохастически, приводя к образованию гетерогенной смеси белков с различной степенью модификации. Непредсказуемое связывание белка также вызывает его неправильную пространственную ориентацию на поверхности ЛНЧ [153]. Кроме того, наблюдается неоднородность индекса полидисперсности липидных наночастиц, поскольку неконтролируемое сшивание белков между несколькими наночастицами или мультимеризация белков на одной частице нарушают структурную стабильность липидного слоя. Нарушение однородной упаковки липидов приводит к локальным дефектам, провоцирует агрегацию, увеличение размеров наночастиц и невоспроизводимость соотношения белок/наночастица от партии к партии [154]. Кроме того использование EDC/NHS–активации для функционализации наночастиц белками сопряжено с кинетическими и структурными ограничениями, поскольку эффективность конъюгации критически зависит от баланса pH, ионной силы буфера и контроля побочных реакций [155,156].

Для достижения большей селективности часто применяется реакция Михаэля между малеимид–функционализированными липидами (DSPE–PEG–Mal) и свободными сульфгидрильными группами цистеинов белка [20]. Несмотря на успешное применение одобренного FDA иммунодепрессанта на основе ПЭГилированного тиол-малеимидного аддукта Cimzia® (цертолизумаб пегол), родственные конъюгаты в контексте разработки и декорирования ЛНЧ могут иметь ограничения. Хотя данная реакция протекает в относительно мягких условиях и характеризуется высокой скоростью, ее адаптация для модификации ЛНЧ не всегда возможна. Токсикологические исследования продемонстрировали, что остаточные нереагировавшие малеимидные группы на поверхности липидных наночастиц вызывают выраженную иммуногенность и активацию системы комплемента [157]. Было показано, что свободный малеимид на поверхности наноструктур активно связывается с сывороточным альбумином крови *in vivo*. Кластеризация альбумина на поверхности ЛНЧ запускает альтернативный путь каскада комплемента, что приводит к ускоренному выведению частиц из кровотока, снижению уровня тромбоцитов до 50% и выраженной системной токсичности. Другой проблемой является химическая нестабильность конъюгата: сукцинимид–тиоэфирная связь подвергается обратимому распаду *in vivo* [158]. Может происходить ретро–Михаэлевское

раскрытие (retro–Michael reaction). В присутствии высоких концентраций эндогенных тиолов, например, глутатиона или свободных цистеинов плазмы, происходит постепенное отщепление таргетирующего белка от поверхности ЛНЧ, что может приводить к преждевременному высвобождению инкапсулированного содержимого [159]. Кроме того, малеимидное кольцо подвергается спонтанному гидролизу в водных растворах. С одной стороны, раскрытое кольцо стабилизирует конъюгат и останавливает ретро–реакцию. С другой стороны, этот процесс протекает стохастически, создавая химическую гетерогенность (смесь закрытых и раскрытых форм), что усложняет стандартизацию [160].

Ориентированная вставка подразумевает, что белок крепится к липидной наночастице строго определенным участком, оставляя активный центр открытым. В случае тиол–малеимидного взаимодействия добиться этого сложно. Во-первых, если в белке есть несколько доступных свободных цистеинов, малеимид будет связываться со всеми случайным образом. Конструкция закрепится в хаотичных пространственных ориентациях, частично экранируя активный центр белка [161]. Однако этого можно избежать, выбрав в качестве таргетирующего агента белки или пептиды без цистеинов в составе. Во-вторых, чтобы сделать вставку ориентированной, на C- или N-конец белка часто искусственно вводят цистеиновый таг. Однако наличие свободных тиольных групп часто приводит к нежелательной димеризации или агрегации молекул при наработке или очистке [162]. Чтобы провести конъюгацию, необходимо освободить неправильно связанные цистеины с помощью восстановителей, например DTT или TCEP [163]. В сложных белках и полноразмерных антителах образование S–S-связи критично для формирования третичной структуры и междоменных связей. Попытки их частичного восстановления для направленной конъюгации могут приводить к полной денатурации, нарушению структуры и потере функциональности биомолекул.

Из–за ограничений классических методов химической конъюгации возникает необходимость перехода к биоортогональным стратегиям, позволяющим повысить однородность и воспроизводимость получаемых конъюгатов. Конъюгаты белков с направленной модификацией неоднократно демонстрировали фармакологические профили, в целом превосходящие профили их гетерогенных аналогов. В качестве альтернативы конъюгации между малеимидом на поверхности частицы и цистеином в составе белка всё больший интерес вызывает биоортогональная клик–химия, в частности безмедное азид–алкиновое циклоприсоединение, индуцируемое напряжением цикла (SPAAC) [164]. Этот метод требует предварительного введения в белок неприродных функциональных групп, обеспечивающих селективное взаимодействие без необходимости использования катализаторов на основе меди.

Введение НПА с азидогруппой в последовательность белка обеспечивает высокую селективность конъюгации и позволяет контролировать как положение модификации, так и стехиометрию конечного продукта. В отличие от традиционных методов, основанных на взаимодействии с природными функциональными группами белка, генетически направленное включение НПА снижает вероятность образования смеси позиционных изомеров и способствует получению более однородных конъюгатов. Применение этого метода недавно получило широкую огласку благодаря клиническим испытаниям (NCT04829604 и NCT04009681) терапевтических препаратов на основе белков, содержащих НПА. Применение SPAAC позволяет проводить конъюгацию в мягких условиях без использования медных катализаторов, что особенно важно для сохранения структуры и биологической активности белков. Сочетание генетически направленного введения функциональных групп и биоортогональных реакций может стать важным инструментом создания воспроизводимых белковых конъюгатов с заранее заданными свойствами и расширить возможности разработки современных биотерапевтических препаратов.

# 2. МАТЕРИАЛЫ И МЕТОДЫ

## 2.1 Общие материалы и методы

### 2.1.1 Питательные среды и растворы

В данной работе были использованы производные фенилаланина и тирозина: О–метил–L–тирозин (oMeY), 4–азидо–L–фенилаланин (pAzF) (Tokyo Chemical Industry Co., Ltd., Япония), О–аллил–L–тирозин (oAllylY) (синтезированный из О–бензил–L–тирозина в Центре медицинской химии Университета Сириус) и 2–хлор–L–фенилаланин (2–ClF) (Shanghai Acmec Biochemical Co., Китай). Максимальная концентрация О–аллил–L–тирозина составляла 5 мМ, других НПА – 10 мМ. НПА растворяли в среде 2×YT (составы сред для прокариотических культур, использованных в данном исследовании, перечислены в приложении А, таблица А1) в ультразвуковой бане при температуре не выше 30°C, затем проводили фильтрацию через шприцевые фильтры с размером пор 0,22мкм (Jet Bio Filtration, Китай)

Для приготовления селективных сред использовались антибиотики (Gold Biotechnology, США) в следующих рабочих концентрациях: карбенициллин – 50 мкг/мл; спектиномицин – 50 мкг/мл; хлорамфеникол – 30 мкг/мл; канамицин – 30 мкг/мл; тетрациклин – 15 мкг/мл; стрептомицин – 50 мкг/мл.

### 2.1.2 Бактериальные штаммы

В работе был использован штамм *E.coli* NEB® Stable (*F′ proA+B+ lacIq* Δ*(lacZ)M15 zzf::Tn10 (TetR)/* Δ*(ara–leu) 7697 araD139 fhuA* Δ*lacX74 galK16 galE15 e14– Φ80dlacZ*Δ*M15 recA1 relA1 endA1 nupG rpsL (StrR) rph spoT1* Δ*(mrr–hsdRMS–mcrB*).

Для наработки нуклеиновых кислот использовали штамм *E.cloni® 10G* (*F – mcrA Δ(mrr–hsdRMS–mcrBC) endA1 recA1 Φ80dlacZΔM15 ΔlacX74 araD139 Δ(ara, leu)7697 galU galK rpsL nupG λ– tonA (StrR)*).

Для наработки бактериофага и проведения PACE использовали штамм *E.coli S2060* (Addgene #105064). *(F' proA+B+ Δ(lacIZY) zzf::Tn10 lacIQ1 PN25–tetR luxCDE Ppsp(AR2) lacZ luxR Plux groESL / endA1 recA1 galE15 galK16 nupG rpsL ΔlacIZYA araD139 Δ(ara,leu)7697 mcrA Δ(mrr–hsdRMS–mcrBC) proBA::pir116 araE201 ΔrpoZ Δflu ΔcsgABCDEFG ΔpgaC λ–)*

Для наработки пептида легкой цепи человеческой энтеропептидазы использовался штамм *SHuffle® T7 express E. coli* (*fhuA2 lacZ::T7 gene1 [lon]. ompT ahpC gal λatt::pNEB3–r1–cDsbC (SpecR, lacIq) ΔtrxB sulA11 R(mcr–73::miniTn10—TetS)2 [dcm]. R(zgb–210::Tn10 —TetS) endA1 Δgor Δ(mcrC–mrr)114::IS10*).

Для наработки белков был выбран штамм *T7 express E. coli (NEB)* (*fhuA2 lacZ::T7 gene1 [lon]. ompT gal sulA11 R(mcr–73::miniTn10—TetS)2 [dcm]. R(zgb–210::Tn10—TetS) endA1 Δ(mcrC–mrr)114::IS10*).

### 2.1.3 Плазмиды и гены

Плазмиды pDB023f (Addgene #99217), pDB038 (Addgene #99210), pDB038a (Addgene #99210), pDB038b (Addgene #99210), pDB021CH(+) (Addgene #99209), SplitC и SplitD и MP6 были получены из addgene. Плазмиды Rosetta, pET и pET30a+ были переданы из Института цитологии и генетики СО РАН. Все используемые плазмиды перечислены в приложении А, таблица А2.

### 2.1.4 Конструирование олигонуклеотидов и проведение ПЦР

Олигонуклеотиды для проведения ПЦР конструировались с помощью программы SnapGene и онлайн–ресурса OligoAnalyzer. В SnapGene выделялась область комплементарная праймерам, а в OligoAnalyzer проверялось образование вторичных структур и температуру отжига. Все олигонуклеотиды были синтезированы в ресурсном центре генетической инженерии АНОО ВО НТУ «Сириус» с использованием синтезатора ДНК/РНК ASM–800ET (BioSset, Россия).

Все ПЦР проводились в амплификаторах Bio–Rad T100, США, ПЦР в реальном времени (ПЦР–РВ) осуществлялась на приборе CFX96 Touch System (Bio–Rad, США). Для сборки генов

*de novo* использовалась Taq Hot Start полимераза (Евроген, Россия), для амплификации генов с плазмид — Q5®High–Fidelity DNA Polymerase (New England Biolabs, США).

После ПЦР пробы хранились на –20˚С и проверялись электрофорезом в 1,5% агарозном геле в камерах для горизонтального электрофореза (Bio–Rad, США). Визуализация фрагментов в геле проводилась при помощи бромистого этидия и гель–документирующей системы ImageLab (Bio–Rad, США).

Ампликоны, фактическая длина которых совпадала с ожидаемой, были использованы в дальнейшей работе и очищены от остатков ПЦР–смеси при помощи набора магнитных частицах KAPA HyperPure Beads (KAPABIOSYSTEMS, США) по протоколу производителя.

### 2.1.5 Приготовление электрокомпетентных клеток и электропорация

Клеточная культура, выращенная до оптической плотностью ($OD_{600}$) равной 0,6 в среде LB, помещалась на лед для инкубации 1ч. После чего суспензию осветляли центрифугированием 15 минут при 6000 rcf при помощи Eppendorf Centrifuge 5910R (Eppendorf Германия). Клеточный осадок последовательно промывали охлажденными до 4℃ 5%–ным и 10%–ным растворами глицерина. После последней отмывки осадок ресуспендировали в 200 мкл 10% глицерина, аликвотировали по 30 мкл, замораживали в жидком азоте и хранили при −80℃.

Трансформацию клеток проводили методом электропорации. К аликвоте электрокомпетентных клеток добавляли плазмидную ДНК, после чего электропорацию проводили в кюветах Bio–Rad Gene Pulser Cuvette с использованием электропоратора Bio–Rad MicroPulser (Bio–Rad, США). После электропорации клеточную суспензию переносили в среду SOC и инкубировали в течение 1ч при 37℃. Затем клетки высевали на чашки Петри с соответствующими антибиотиками и инкубировали в течение 18 ч при 37℃ в термостате Biosan Smart (BioTherm, РФ).

### 2.1.6 Секвенирование методом Сэнгера

Образцы ДНК для секвенирования представляли собой либо плазмиды (100 нг/мкл), либо ПЦР–продукты (20–50 нг/мкл), а также праймеры с концентрацией 3 пмоль/мкл, исходя из расчета что на одну реакцию нужно 5 мкл образца ДНК и 5 мкл праймера. Секвенирование проводилось в ресурсном центре генетической инженерии АНОО ВО НТУ «Сириус» с использованием набора BigDye Terminator v3.1 Cycle Sequencing Kit (Thermo Scientific™, США) на приборе Applied Biosystems 3730xl DNA Analyzer (48 Capillaries, 50 sm) (Thermo Scientific™, США) с использованием POP–7 Polymer.

## 2.2 Сборка плазмидных конструкций

Ампликоны, полученные с матрицы плазмиды и собранных *de novo*, содержали перекрывающиеся концевые участки для сборки методом Гибсона либо сайты рестрикции соответствующих рестриктаз. Трансформацию проводили методом электропорации согласно протоколу п.1.2.5. Наработку и очистку плазмидной ДНК осуществляли одинаково для всех конструкций. Трансформантов высевали на LB–агар с селективным антибиотиком и инкубировали при 37°C в течение 16–18ч, если не указано иное. По 10–15 колоний переносили в 7 мкл деионизированной воды и использовали как матрицу для скрининговой ПЦР. Лучшие клоны выращивали в 5–7 мл LB с селективным антибиотиком в лабораторном шейкере инкубаторе 16 часов при 37℃ 180 об/мин. После чего плазмидную ДНК выделяли с использованием Plasmid Miniprep Kit (Biolabmix, Россия). Секвенирование методом Сэнгера проводили согласно протоколу п.2.1.6.

### 2.2.1 Получение фаговых библиотек, несущих варианты гена PylRS

Для реализации PACE использовали нитчатый бактериофаг M13, в котором ген pIII заменяли последовательностью АРС археи *Methanosarcina mazei.* Также создали химерную синтетазу ChPylRS, содержащую N–домен MbPylRS из *M. barkeri* и C–домен MmPylRS из *M.*

*mazei*, соединенные нуклеотидным линкером. N–домен MbPylRS был собран из синтетических олигонуклеотидов указанных в приложении Б. В С–доменах обеих синтетаз вносили аминокислотные замены Asn346Ala/Cys348Ala.

Фаг собирали методом Golden Gate по сайту SapI из фрагментов pBT29–splitC (Addgene #122598), pBT29–splitD (Addgene #122599) и APC, кодирующей MmPylRS_mut или ChPylRS_mut [165]. Очищенные фрагменты смешивали в соотношении 1:1:4 (C:D:APC) с SapI и T4–лигазным буфером. После рестрикции при 37 °C в течение 30 мин добавляли T4–лигазу и проводили сборку: 16°C – 20 мин; 37°C – 5 мин; 16°C – 5 мин (30 циклов); 16°C – 20 мин; 65°C – 20 мин; 80°C – 10 мин. Сборку подтверждали ПЦР с фланкирующими ген интереса праймерами, после чего очищенную лигазную смесь трансформировали методом электропорации в клетки S2060 с плазмидой pJC175e (Addgene #79219). Клетки инкубировали 16ч в 2×YT без антибиотиков, после чего наличие фага с целевой вставкой подтверждали ПЦР. Клетки центрифугировали (8000 rcf, 3 мин), супернатант фильтровали через фильтр 0,22 мкм и хранили при 4 °C.

Для получения отдельных бляшек клетки S2060/pJC175e выращивали до $OD_{600}$ = 0,6 в 2×YT с ампициллином или карбенициллином. Фаговый супернатант серийно разводили 1:100 (4 разведения), смешивали 600 мкл клеток с 40 мкл соответствующего разведения фага и инкубировали 5 мин при комнатной температуре. Смесь добавляли к 4 мл верхнего агара 2×YT (~55°C) с X–Gal (0,04%) и выливали на нижний агар 2×YT с карбенициллином. Чашки инкубировали 16ч при 37°C. В качестве положительного контроля использовали VCSM13 ($10^6$).

Полученные бляшки переносили в 20 мкл 2×YT, после чего 10 мкл суспензии использовали для хранения при 4°C, а оставшуюся часть прогревали при 80°C в течение 30 мин и анализировали ПЦР с фланкирующими праймерами. Ампликоны очищали с помощью KAPA HyperPure Beads (KAPABIOSYSTEMS, США) и подтверждали корректность вставки секвенированием по Сэнгеру согласно п. 1.2.6. Корректные клоны выращивали в 2 мл 2×YT в течение ночи при 37°C без антибиотиков. Супернатант после центрифугирования и фильтрации использовали для дальнейшей работы.

Титр фаговых библиотек определяли методом ПЦР–РВ или анализа бляшек на S2060/pJC175e. Для ПЦР–РВ фаг известного титра серийно разводили 1:10 (8 разведений) для построения калибровочной кривой. 1 мкл образца неизвестного титра добавляли к 24 мкл реакционной смеси с праймерами( SplitC_gV_for_qPCR_F 5’–CACCGTTCATCTGTCCTCTTTCAA–3’; SplitC_gV_for_qPCR_R 5’–CGACCTGCTCCATGTTACTTAGC–3’; SplitC_gV_for_qPCR_probe

5’–FAM–CATAAGGGAACCGAACTGACCAAC–BHQ1–3’) и HS–qPCR (2×) (Биолабмикс, Россия). ПЦР проводили по протоколу производился: отжиг при 60 °C, элонгация – 15 с, всего 40 циклов.

### 2.2.2 Сборка плазмид негативной селекции

Для проведения экспериментов по негативной селекции были собраны две плазмиды: pDB007ns2a–neg и pDB016–tRNApyl.

Плазмида pDB007ns2a–neg была собрана из двух доноров: из pDB007ns2a (Addgene #99214)) получен фрагмент, содержащий ген *gIII* без амбер стоп–кодонов, из pDB023f (Addgene #99217) – фрагмент с геном T7–РНК–полимеразы, содержащим два амбер стоп–кодона. Обе плазмиды расщепляли рестриктазами NcoI и XbaI (SibEnzyme, Россия). Полученные фрагменты очищали при помощи магнитных частиц KAPA Pure Beads (KAPA Biosystems, США) в соответствии с протоколом производителя. Сборку проводили с использованием T4 ДНК–лигазы (SybEnzyme, Россия). После лигирования смесь также очищали с помощью магнитных частиц перед последующей трансформацией с высевом на чашку с селективным антибиотиком ампицилином. Дополнительно, для верификации сборки pDB007ns2a–neg проводили рестрикционный анализ с ферментом BglII (SibEnzyme, Россия). Правильно собранная конструкция давала фрагменты длиной 5787 и 2570 п.н.

Для получения плазмиды pDB016–tRNApyl использовали pDB016 (Addgene #99220) и pDB038 (Addgene #99210) в качестве доноров необходимых фрагментов. Из плазмиды pDB038 амплифицировали ген $мРНК_{Pyl}$. Праймеры были сконструированы таким образом, чтобы ампликоны несли сайты рестрикции NcoI с обоих концов. Плазмиду pDB016 и ампликоны с геном $мРНК_{Pyl}$ расщепляли рестриктазой NcoI (SibEnzyme, Россия) и проводили выделение ДНК на магнитных частицах. Далее готовили смесь для сборки с T4 ДНК–лигазой. Аналогично эксперименту по получению плазмиды pDB007ns2a–neg, проводили очистку пробы на магнитных частицах и трансформацию в электрокомпетентные клетки *E.cloni* с последующим высевом на чашку, содержащую LB агар с добавлением спектиномицина.

### 2.2.3 Сборка плазмид флуоресцентной тест системы

Для включения НПА в белки были созданы плазмидные конструкции pHPyl–PylRS–GlnS, содержащие АРС, под контролем химерного глутаминового промотора, а также тРНК$_{Pyl}$. Сборка проводилась из двух фрагментов: первый из которых содержал ген АРС с промотором и терминатором, достроенными при помощи длинных олигонуклеотидов к концам ампликона, и второго, содержащего ориджин репликации и регуляторную область, а также ген устойчивости к хлорамфениколу и тРНК$_{Pyl}$.

Другая плазмида pET–sfGFP–TAG содержит последовательность sfGFP (суперфолдер GFP) с амбер–кодоном в 27 положении. Плазмида была собрана из трех фрагментов, первый содержал последовательность sfGFP, второй ориджин репликации и регуляторные последовательности, третий ген устойчивости к канамицину. Первый фрагмент был амлифицирован с матрицы sfGFP–C1 (Addgene #54579), второй и третий фрагмент амплифицированы с матрицы pET. Праймеры были сконструированы таким образом, чтобы ампликоны несли сайты рестрикции SapI с обоих концов. Сборку проводили методом Golden Gate, используя фермент SapI (New England Biolabs, США) по протоколу описанному в п.2.2.1.

Для создания генетических конструкций, содержащих Эксендин–4, последовательности гена (нуклеотидная и аминокислотная последовательности указаны в Приложении В) была собрана из олигонуклеотидов путем нескольких последовательных ПЦР с перекрывающимися концевыми участками (олигонуклеотиды для синтеза указаны в приложении В таблица В1). В качестве векторной последовательности использовалась плазмида pET30(a+). Сборку проводили методом Gibson Assembly. Смесь после сборки дополнительно очищали с помощью магнитных частиц перед последующей трансформацией. Таким образом была получена конструкция pET30(a+)_His8_TRX–A_Enterokinase_Exe4. Для наработки белка с pAzF в 40 положении Эксендин–4 был введен амбер–кодон UAG. Конструкция была получена методом сайт направленного мутагенеза при помощи перекрывающихся олигонуклеотидов, матрицей для ПЦР выступала ранее полученная конструкция pET30(a+)_His8_TRX–A_Enterokinase_Exe4.

Последовательность, кодирующая каталитически активную легкую цепь человеческой энтеропептидазы (hEKL, human enterokinase light chain), была взята из общедоступных баз данных, собрана методом ПЦР из перекрывающихся олигонуклеотидов с использованием фланкирующих праймеров. Последовательности праймеров, нуклеотидная и аминокислотная последовательнгости hEKL указаны в Приложении Г. Ген, кодирующий hEKL был клонирован в

плазмиду pET30(a+)_His8_TRX–A_Enterokinase. В качестве векторной последовательности использовалась плазмида pET30(a+) (Addgene #175021).

Аналогичным образом были собраны плазмиды, содержащие генетическую последовательность DARPin анти mCD8 (MSE10) [166]. Последовательности праймеров, нуклеотидная и аминокислотная последовательности DARPin указаны в Приложении Д.

## 2.3. Фаг ассоциированная эволюция

### 2.3.1 Цикл фаг–ассоциированной эволюции (PANCE)

В настоящей работе PANCE включал четыре раунда положительной селекции. Перед каждым раундом *E. coli* S2060 (Addgene #105064) трансформировали необходимыми плазмидами. В первом раунде использовали pDB023f (Addgene #99217), содержащую два амбер стоп–кодона в гене РНК–полимеразы T7, и pDB021CH(+) (Addgene #99209) с геном gIII под промотором T7. Во втором, третьем и четвертом раундах применяли pDB038, pDB038a и pDB038b (Addgene #99210), содержащие соответственно один, два и три стоп–кодона в gIII. Плазмиду мутагенеза pMP6C использовали во всех раундах.

Положительная селекция проводилась в среде 2xYT, дополненной 2–10 мМ соответствующей АРС НПА и необходимыми антибиотиками. Для выращивания клеток на чашках Петри использовали 2×YT с 1,5% агаром и соответствующими антибиотиками. Единичные колонии переносили в жидкую 2×YT с антибиотиками и 100 мМ глюкозы и выращивали 16 ч при 37 °C и 180 об/мин. Ночную культуру пересевали в свежую 2×YT с антибиотиками и 2–10мМ НПА и выращивали до $OD_{600}$=0,6–0,7. Мутагенез индуцировали добавлением 5 мМ арабинозы, после этого клетки инфицировали селективным фагом предыдущего раунда ($10^6$ БОЕ/мл) и инкубировали ночь при 37 °C 180 об/мин. Перед следующим раундом клетки центрифугировали (8000 об/мин, 5 мин), супернатант фильтровали через фильтр 0,22 мкм.

Титр фага определяли методом количественной ПЦР–РВ. К 1 мкл фагосодержащего супернатанта добавляли реакционную смесь с праймерами и зондом согласно п.2.2.1 и HS–qPCR (Биолабмикс, Россия). Амплификацию проводили на BIO–RAD CFX96 (Bio–Rad,

США), калибровочную кривую строили по серийным разведениям фага известного титра. Расчет титра выполняли в Microsoft Excel.

Отрицательную селекцию проводили с использованием плазмид pDB016+tRNAPyl и pDB007nsa2–neg. Инфекцию проводили в 2×YT с необходимыми антибиотиками без НПА при $OD_{600}$=0,5–0,6, используя супернатант предыдущего раунда с концентрацией селективного фага $10^6$ БОЕ/мл. Инкубацию инфекционной смеси проводили не более 4 ч.

### *2.3.2 Определение селективности АРС полученных в результате PANCE*

Клетки содержащие необходимые плазмиды, в зависимости от этапа PANCE добавляли: в среду с соответствующей АРС НПА, концентрацией 1–10 ммоль/мл в зависимости от этапа; либо в среду 2xYT в отсутствии НПА. Заражение и культивирование проводили по протоколам, описанным ранее для положительного отбора п.2.3.1. По результатам ПЦР–РВ оценивали разницу в титрах бактериофага. Тест проводили после каждого этапа PANCE. При разнице титров на 1,5–2 порядка и более в случае добавления в среду НПА, относительно условий без НПА, эволюционную линию переводили на следующий этап PANCE. В остальных случаях проводили несколько этапов негативной селекции, после чего повторяли тест.

### *2.3.3 Определение активности АРС полученных в результате PANCE методом теста флуоресценции sfGFP*

Ген АРС, кодирующий вариант синтетазы, получали методом ПЦР с Fusion ДНК–полимеразой (Thermo Fisher Scientific, США) в качестве матрицы использовали пул бактериофагов после последнего этапа PANCE. В качестве вектора использовали фрагмент плазмиды pHPyl–PylRS–GlnS; сборку проводили методом Gibson Assembly. Полученную плазмиду pHPyl–PylRS–GlnS, содержащую ген АРС совместно с pET–sfGFP–27TAG трансформировали в электрокомпетентные клетки BL21 Star. Клетки выращивали на LB–агаре с канамицином и хлорамфениколом при 37°C в течение ночи. Отдельные колонии культивировали в 800 мкл среды 2×YT с антибиотиками при 37°C и 750 об/мин в течение ночи.

Для оценки активности синтетаз в 96–луночный планшет с глубокими лунками вносили 760 мкл среды 2×YT с 2 мМ соответствующей НПА или без неё и 40 мкл ночной культуры (в трех повторах). Экспрессию индуцировали добавлением 1 мМ IPTG, после чего культуры выращивали 18ч при 37°C 750 об/мин. По 200 мкл культур переносили в плоскодонные планшеты Costar и измеряли флуоресценцию sfGFP (λex = 485 нм, λem = 528 нм) и $OD_{600}$ на приборе CLARIOstar Plus (BMG Labtech, Германия).

Нормализацию проводили только по оптической плотности культуры, а вычитание фонового сигнала выполняли для соответствующих культур, с добавлением и без добавления НПА соответственно. В качестве положительного контроля использовали клетки, экспрессирующие sfGFP без стоп–кодонов, отрицательными контролями служили клетки с sfGFP–27TAG и PylRS дикого типа, а также чистая культуральная среда.

### 2.3.4 Подготовка библиотеки для секвенирования на Illumina MiSeq

Основной протокол, использованный в ходе получения библиотеки генов – AmpliSeq [167]. Для подготовки генов вариантов *PylRS* к секвенированию на Illumina MiSeq, проводили амплификацию фрагментов длиной 400 – 450 пар нуклеотидов. Для ПЦР реакции использовали набор Q5 Hot Start High–Fidelity 2X Master Mix (NEB, США). Пары праймеров были разработаны таким образом, чтобы обеспечить покрытие полной длины гена дважды. Полученные ампликоны использовались в качестве матрицы для второй ПЦР, в ходе которой к ним присоединяли необходимые технические последовательности и баркоды. Очистку продуктов амплификации от остаточных праймеров и праймер–димеров после каждого раунда PCR выполняли с использованием магнитных частиц KAPA HyperPure Beads (KAPA Biosystems, США) по стандартному протоколу производителя.

Общую библиотеку сформировали путем смешивания всех очищенных индексированных ампликонов в эквимолярных количествах. Полученную смесь дополнительно два раза очищали на магнитных частицах, обеспечив тем самым высокое качество библиотеки для последующего секвенирования. Количественный анализ проводили с использованием флуориметра Qubit 4 Fluorometer (Thermo Fisher Scientific, США). Итоговое качество библиотеки оценивали на капиллярном электрофорезе с использованием прибора Agilent TapeStation 4150 (Agilent Technologies, США). Концентрация библиотеки составляла более 20 нг,

что соответствовало требованиям для загрузки на платформу Illumina MiSeq. Секвенирование на Illumina MiSeq проводили в ресурсном центре геномных исследований АНОО ВО НТУ «Сириус».

### 2.3.5 Обработка результатов MiSeq

Обработку данных секвенирования начинали с удаления ридов короче 50 нуклеотидов и адаптерных последовательностей с помощью Trimmomatic v0.39. Качество данных до и после фильтрации оценивали в FastQC v0.12.1. Отфильтрованные риды выравнивали на референсную последовательность гена с помощью Bowtie2 v2.4, удаляя невыровненные риды. Размер целевых фрагментов составлял 400–480 п. н. SAM–файлы конвертировали в BAM с помощью samtools view v1.6. Глубину покрытия оценивали командой samtools depth, распределение размеров вставок – с помощью Picard CollectInsertSizeMetrics. Варианты выявляли с использованием FreeBayes v1.3.9 при качество выравнивания не ниже 20, минимальное покрытие десятикратное, частота альтернативного аллеля от 1%. Консенсусные последовательности формировали в bcftools v1.9 с помощью bcftools consensus. Визуальный контроль выравниваний и выявленных вариантов проводили в Integrative Genomics Viewer v2.12.

### 2.3.6 Моделирование белковых структур

Выравнивание последовательностей и поиск нуклеотидных замен в вариантах проводились с использованием программы MEGA (Molecular, Evolutionary, and Genetic Analysis). Структурные аналоги идентифицировали с помощью базы данных PDB (Protein Data Bank). В качестве шаблонов для исследования были выбраны наиболее похожие кристаллические структуры: для N–домена – структура 5V6X, имеющая 87% идентичности последовательности с эталонной (Seq Identity = 87%); для C–домена – структура 2ZCE (Seq Identity = 99%).

Трехмерные модели эволюционировавших ферментов были построены с использованием онлайн–сервиса SWISS–MODEL. Качество построенных моделей оценивалось с помощью

параметра глобальной оценки качества модели (GMQE): GMQE = 0,79 для N–домена и 0,82 для C–домена. Анализ пространственной структуры, визуализация модели и введение точечных нуклеотидных замен были выполнены с использованием PyMOL версии 3.1.5.1.

## 2.4 Наработка и очистка белков

Этап первичной металл–хелатной хроматографической очистки образцов проводили аналогично для всех белков. По истечению времени инкубации жидкость с клетками центрифугировали в Eppendorf Centrifuge 5910R (Eppendorf, Германия) при 6000 об/мин в течение 15–20 мин при 4°C. Супернатант сливали, а клеточный осадок ресуспендировали в 30 мл буфера для ресуспендрования (Приложение А, Таблица А5). Пробирку типа Фалькон с клеточной суспензией помещали в контейнер со льдом и подвергали ультразвуковому разрушению клеточной стенки в гомогенизаторе UC–650 Ultrasonic cell disruptor, (MIULAB, Китай) при цикличном параметре 15мин, 7 сек включения, 5 сек выключения. В дальнейшем пробирку с разрушенными клетками центрифугировали при 12000 об/мин в течении 1ч при 4°C. Супернатант отфильтровывали через шприцевой мембранный фильтр с размером пор 0,22мкм (Пущинские лаборатории, РФ).

Очистку белков проводили методом IMAC при помощи ÄKTA Avant 150 (Cytiva, США) с использованием колонки Diba Omnifit® EZ (10×100 мм) с никелевым сорбентом Nuvia IMAC (Bio–Rad, США). Перед нанесением образцы фильтровали через шприцевой фильтр 0,22–0,45 мкм. Образцы наносили со скоростью 1 мл/мин (время контакта 2 мин), после чего несвязавшиеся компоненты удаляли 10 колоночными объемами (CV) буфера А (50мМ Трис–HCl, 5мМ имидазол, 300мМ NaCl, pH=7,5) при скорости потока 1 мл/мин. Элюцию проводили ступенчатым градиентом 10, 20 и 60% буфера Б (50мМ Трис–HCl, 500мМ имидазол, pH=7,5) при скорости 1 мл/мин. Все стадии проводили при комнатной температуре. Фракции проскока и элюата собирали автоматически с помощью фракционного коллектора ÄKTA Avant.

### 2.4.1 Наработка энтеропептидазы

Электрокомпетентные клетки *SHuffle® T7 express E. coli* ко–трансформировали pET30a(+)_His8_TRX–A_hEKL_Enterokinase совместно с pGro7 (Addgene #228495). После получения колоний на чашке Петри, проводили посев ночной культуры в среду LB (Приложение А, Таблица А.1) с добавлением канамицина и хлорамфеникола, растили при 37°C, 180 об/мин в ночь. В 1л жидкой среды LB вносились: соответствующие антибиотики, арабиноза до конечной концентрации 10мкМ, ночная культура до достижения оптической плотности $OD_{600}$=0,1. Клетки культивировали при 30°C до достижения оптической плотности $OD_{600}$=1,2. После этого добавляли ИПТГ до конечной концентрации 0,7мМ. После индукции экспрессии белка клетки растили в шейкере инкубаторе при 28°C, 24–30ч. Очищенную металл–хелатной хроматографией энтеропептидазу подвергали диализу в целлюлозном мешке с размером пор 10кДа (Thermo Fisher Scientific, США), помещенную в большой сосуд с буфером для хранения (pH 7,5, 25mM Tris, 2mM CaCl2, 200mM NaCl) и магнитной мешалкой 16–18 часов. Полученный белок разводили до концентрации 0,1мг/мл и использовали в дальнейших экспериментах.

### 2.4.2 Наработка Эксендин–4 и DARPin mCD8

Для наработки DARPin и Эксендина–4 (Exe4) с НПА электрокомпетентные клетки *T7 express* ко–трансформировали двумя плазмидами: pET30(a+)_His8_DDDDK_CD8_DARPin_ncAA или pET30(a+)_His8_TRX–A_DDDDK_Exe4 с амбер–стоп кодоном и плазмидой pHPyl–AzFPylRS–GlnS, содержащей ортогональную АРС, которая включает pAzF в процессе трансляции. Для наработки белков без НПА клетки трансформировали лишь плазмидами pET30a__His10_DDDDK_DARPin_mCD8_GG–C–GS или pET30(a+)_His8_TRX–A_DDDDK_Exe4 без добавления НПА в среду.

Трансформанты высевали на чашки, содержащие канамицин и хлорамфеникол, если необходимо, инкубировали при 37°C, 16–18 часов. После образования колоний отдельные клоны скринировали для подтверждения наличия правильной плазмиды. Лучшие клоны высевали в 7мл жидкой среды LB, инкубировали при 37°C 180 об/мин, 16–18 часов. Далее ночную культуру добавляли до $OD_{600}$=0,1 в колбу с 500 мл индукционной среды на основе

бульона 2xYT, дополненной соответствующими антибиотиками и 2мМ pAzF. Клетки растили до $OD_{600}$=0,6 при 37°C, 180 об/мин. Добавляли 0,5мМ ИПТГ и инкубировали 24ч при 18°C, 180 об/мин.

Очищенные белки подвергали диализу в целлюлозном мешке с размером пор 10кДа (Thermo Fisher Scientific, США) против буфера для хранения (pH 7,5, 25mM Tris, 2mM CaCl2). На каждый 1мг полученного Exe4 добавляли 10мкг энтеропептидазы. Смесь белков инкубировали при 30°C, 12–16 часов.

Целевой Exe4 очищали при помощи обращенно–фазной ВЭЖХ (раздел 2.4.4). Полученные фракции анализировали методом SDS–PAGE.

### 2.4.4 Обращенно–фазовая хроматографическая очистка образцов

Очистку Exe4 проводили методом полупрепаративной обращенно–фазовой ВЭЖХ на системе Agilent 1260 Infinity II (Agilent, США) с использованием колонки Gemini® NX–C18 (150×2,0 мм, 5 мкм, 110 Å; Phenomenex, США) с предколонкой. В качестве подвижных фаз использовали 30 мМ ацетат аммония pH 5,5 в Milli–Q (буфер А) и 30 мМ ацетат аммония pH 5,5 в 70% ацетонитриле HPLC–grade (буфер Б). Образец растворяли в стартовом буфере (5% Б) до концентрации 1 мг/мл и фильтровали через нейлоновый фильтр 0,22 мкм. Объем инъекции составлял 500 мкл, скорость потока – 0,5 мл/мин, температура колонки – 30 °C.

Разделение осуществляли по следующей программе: 0 – 15 мин – 5% Б; 15–16 мин – 5–30% Б; 16–36 мин – 30–40% Б; 36–46 мин – 40–95% Б; 46–56 мин – 95% Б; 56–56,5 мин – возврат к 5% Б; 56,5–80 мин – уравновешивание при 5% Б. Перед основной очисткой проводили 1–2 пробных запуска с объемом 20–50 мкл для определения времени удерживания. Целевой пик элюирован при 45,5–46,5 мин (~40% Б) и автоматически собирался по времени удерживания.

После очистки колонку промывали 95% ацетонитрилом и хранили в 50% ацетонитриле. Хроматограммы обрабатывали в OpenLab CDS v2.7 (Agilent). Объединенные целевые фракции хранили при 4 °C или лиофилизировали в течение 16–18 ч.

### 2.4.5 Электрофоретическое разделение белков в ПААГ

Для визуализации белка, полученного после хроматографической очистки, использовали метод SDS–PAGE (электрофорез белков в полиакриламидном геле) в системе для вертикального электрофореза от Bio–Rad (Bio–Rad, США). Были использованы два вида геля: 20% триciновые и 16% глициновые, каждый из которых использовался в зависимости от белка. Составы гелей и буферов представлены в приложении А, таблицы А2 и А3.

Пробу образца 2–10 мкг белка нагревали до 95 °С в течение 10 мин в 4×Laemmli Sample Buffer (Bio–Rad, США) с добавлением 5% β–меркаптоэтанола, вносили в гель. В пустую лунку вносили маркер RAV–11 (Biolabmix, Россия) в объеме 3 мкл. По завершении электрофореза гель извлекали из стекол, промывали деионизированной водой и инкубировали в 5% глутаровом альдегиде 20 минут (если использовался трициновый гель). Белковые полосы окрашивали красителем Coomassie Brilliant Blue Ultrafast Staining solution (Servicebio, Китай) в соответствии с инструкциями производителя. После электрофореза были получены флуоресцентные изображения при 650 нм, а также колориметрическое изображение геля, окрашенного Кумасси, с помощью Системы гель–документирования Bio–Rad ChemiDoc Imaging System (Bio–Rad, США).

### 2.4.6 Масс–спектрометрический анализ

Характеризация белков методом масс–спектрометрии проводилась в ресурсном центре масс–спектрометрии АНОО ВО НТУ «Сириус». Анализ образцов методом ВЭЖХ–МС выполнялся с использованием высокоразрешающего квадрупольного времяпролетного масс–спектрометра Bruker maXis II 4G ETD (Bruker, Биллерика, Массачусетс, США) с источником электро распылительной ионизации (ESI) и диссоциации, индуцированной столкновениями (CID), соединенного с системой сверх высокоэффективной жидкостной хроматографии Thermo Scientific UltiMate 3000 (Thermo Scientific™, США). Анализ проводился с использованием хроматографической колонки Waters BioResolve RP, 2,7 мкм, 2,1 мм × 50 мм, в режиме обращенно–фазовой элюции. Подвижная фаза А представляла собой воду с 0,1% муравьиной кислотой и 0,02% ТФК; подвижная фаза В представляла собой ацетонитрил с 0,1%

муравьиной кислотой. Температура колонки поддерживалась на уровне 60°C. Каждый образец, содержащий 10 нг, разбавляли в пять раз подвижной фазой А.

Последовательность инъекций включала раствор эталонного материала IgGk NIST 8671, все образцы и контрольный раствор (подвижная фаза А). Данные собирались масс–спектрометром в режиме сканирования в диапазоне 800–7000 м/z. Полученные данные обрабатывались с помощью программного обеспечения Byos® Protein Metrics (Byos, Бостон, Массачусетс, США). Была проведена оценка пригодности системы по критериям приемлемости, выполнена деконволюция, сравнены теоретические и измеренные массы, а также оценена относительная распространенность всех обнаруженных изоформ белка.

### 2.4.7 Клик–химическая конъюгация

DARPin или TRX–A–Exe4 разводили до концентрации 1мг/мл в буфере (pH 7.5, 25mM Tris). DBCO–Cy5 (номер Кат. 30F0, приобретен у Lumiprobe, Россия) разводили до 1мг/мл в ДМСО. DSPE–PEG(2000)–DBCO (1,2–distearoyl–sn–glycero–3–phosphoethanolamine–N–[dibenzocyclooctyl(polyethylene glycol)–2000]) (Avanti, USA) разводили до концентрации 10мг/мл в ДМСО. Раствор белка смешивали с избытком флуоресцентной метки или липида в молярном соотношении 1:5 для полного связывания. Раствор инкубировали в течение 16–18 часов при комнатной температуре. По завершении непрореагировавшее соединение удаляли из образцов при помощью центрикона с отсечкой массой 5кДа (Millipore, Германия). Концентрацию определяли с помощью набора для анализа белка QuDye® Protein (Lumiprobe, Россия).

## 2.5 Создание липидных наночастиц

### 2.5.1 Транскрипция и очистка РНК *in vitro*

Использованная в работе конструкция ДНК, кодирующая мРНК eGFP и мРНК Kate2S, была приготовлена на основе коммерческого вектора pSmart (Lucigen, Middleton, WI, США),

линеаризованного эндонуклеазой AhlI (СибЭнзим, Москва, Россия) для транскрипции *in vitro*. Конструкция мРНК включала кэп–структуру CapAG, вирусную 5'UTR TPL длиной 245 нуклеотидов, кодирующую последовательность eGFP, 3'UTR от вакцинной мРНК мРНК–1273 (Moderna) и поли(А)–хвост из 114 нуклеотидов. Синтез мРНК проводили в буфере для IVT из набора *in vitro* мРНК–20 (Биолабмикс, Россия). Реакционная смесь также содержала 10 мМ DTT (Биолабмикс, Россия), 1 е.а./мкл ингибитора РНКаз RiboCare (Евроген, Россия), 18 е.а./мкл РНК–полимеразы T7 (Биолабмикс, Россия), 500 нг линеаризованной матричной ДНК, 2,4 мМ аналога кэпа AG (Биолабмикс, Россия) и по 3 мМ каждого из четырех рибонуклеозидтрифосфатов (Биолабмикс, Россия). Вместо UTP использовали m1ΨTP. Итоговый объем реакционной смеси составлял 50 мкл. Реакции инкубировали в течение 2 ч при 37°C. Матричную ДНК расщепляли термолабильной ДНКазой (2 ед., 15 мин, 37°C). мРНК очищали с помощью VAHTS RNA Clean Beads (Vazyme, Китай), а качество оценивали методом капиллярного гель–электрофореза на приборе Qsep (Bioptic, Чанчжоу, Китай).

### 2.5.2 Формуляция липидных наночастиц

Для инкапсуляции мРНК в ЛНЧ использовали автоматизированную систему Dolomite Microfluidics (Великобритания). В качестве водной фазы применяли 10 мМ цитратный буфер, pH 3 (RusChem, Россия), в качестве органической фазы – абсолютный этанол (Hayman, Индия). Концентрация РНК составляла 200 нг/мкл, общая концентрация липидов – 10 мкг/мкл. Липидная смесь включала холестерин (Sigma–Aldrich, Merck, США), SM–102 и ALC–0315 (Sinopeg, Китай), DSPC (Avanti, США) и DMG–PEG2000 (Avanti, США) в молярном соотношении 42,70:23,15:23,15:9,40:1,60 соответственно (LNP–5). Соотношение водной и органической фаз составляло 3:1, общая скорость потока – 1,5 мл/мин.

Полученные ЛНЧ диализовали в мембранах Float–A–Lyzer G2 (MWCO 10 кДа, эфир целлюлозы; Repligen, США) против PBS (pH 7,4) в течение 12 ч при 4°C при интенсивном перемешивании с однократной заменой буфера. Для конъюгации ЛНЧ с белками к частицам, содержащим 1,33 или 1,5 мол.% DMG–PEG2000, добавляли DSPE–PEG2000–DBCO–DARPin или DSPE–PEG2000–DBCO–Exe4 в количестве 0,27 или 0,1 мол.% соответственно. После 1ч инкубации при низкой температуре и мягком перемешивании ЛНЧ диализовали против буфера

25 мМ Tris–HCl (pH 7,5) с использованием центрифужных концентраторов MWCO 100 кДа (Jet Biofil, Китай).

### 2.5.3 Физико–химические характеристики ЛНЧ

*Динамическое рассеяние света*. Размер ЛНЧ и индекс полидисперсности (PDI) определяли на приборе Zetasizer Ultra (Malvern Panalytical). Образцы разбавляли в 100 раз раствором PBS (pH 7,4), уравновешивали при комнатной температуре и анализировали в пластиковой кювете. Для каждого образца проводили три измерения по 100 серий до получения стабильного значения.

*Дзета–потенциал*. ζ–потенциал измеряли на Zetasizer Ultra с U–образной кюветой. ЛНЧ разбавляли в 100 раз в деионизированной воде, уравновешивали при комнатной температуре и проводили три измерения по трем зондами, до 100 измерений до стабилизации показаний. Относительное стандартное отклонение не превышало 5%.

*Эффективность инкапсуляции РНК*. Эффективность инкапсуляции (EE, %) определяли по ранее описанной методике [168]. ЛНЧ лизировали в TE–буфере (pH 7,5) с 1% Triton X–100 и окрашивали SYBR Green (Еврогене, Россия) до и после лизиса. Концентрацию мРНК определяли по калибровочным кривым стандарта мРНК в TE–буфере с Triton X–100 и без него, рассчитывая содержание инкапсулированной РНК по разнице сигналов лизированных и нелизированных частиц.

*Сканирующая просвечивающая электронная микроскопия (СТЭМ)*. СТЭМ–изображения получали на Crossbeam 550 (Carl Zeiss, Германия) с детектором ПЭМ при ускоряющем напряжении 30 кВ и токе 200 пА. Свежеприготовленные ЛНЧ (10 мг/мл липидов) разбавляли PBS в 10 раз, 10 мкл суспензии наносили на парафиновую пленку и переносили на каплю медную СТЭМ–сетку на 2 мин. Затем добавляли 10 мкл 1% водного ацетата уранила, инкубировали 1 мин и высушивали при комнатной температуре. Изображения регистрировали сразу после подготовки образцов.

## 2.6 Клеточные тесты

### 2.6.1 Выделение и активация мышиных спленоцитов

Эксперименты на животных проводились в соответствии с международными рекомендациями по содержанию и обращению с лабораторными животными. Проведение экспериментов было одобрено комитетом по этике Университета науки и технологий «Сириус» (№ 6.1) 15 января 2024 года.

Селезенку 8–недельной самки мыши линии Balb/c механически измельчали, суспензию пропускали через клеточный фильтр 70 мкм с промывкой холодным PBS. Спленоциты собирали в 20 мл PBS и дважды центрифугировали при 300 × g, 5 мин, 4°C с промежуточной отмывкой 10 мл холодного PBS. Клетки ресуспендировали в RPMI–1640 (ПанЭко, Россия) с 10% FBS (Cytiva, США), 2 мМ L–глутамина (Gibco, США) и 1% пенициллина/стрептомицина (Gibco, США) и замораживали в жидком азоте по 10 млн клеток/мл в криосреде (50% FBS, 40% RPMI–1640, 10% DMSO; 1 мл/криовиалу).

Для активации Т–клеток спленоциты размораживали при 37°C, промывали и ресуспендировали в OptiVitro® T Cell Serum–free Medium (ExCell Bio, Китай). Добавляли mouse CD3/CD28 T Cell Activation Beads (Elabscience, Китай) в соотношении спленоциты:частицы 3:1 и recombinant hIL–2 (25 IU/мл; R&D Systems, Великобритания). Клетки культивировали в 48–луночных TC–treated планшетах (Jet BIOFIL, Китай) по 500 мкл/лунку при исходной концентрации 2 млн клеток/мл при 37 °C в течение 3 суток. После культивирования магнитные частицы удаляли на магнитном штативе. Очищенные клетки использовали для трансфекции согласно п. 2.7.5 либо для окрашивания антителами и DARPin согласно п. 2.7.6.

### 2.6.2 Цитометрическое окрашивание мышиных спленоцитов

Специфичность связывания DARPin с рецептором CD8 оценивали на спленоцитах мыши, активированных в течение 3 суток, методом проточной цитометрии на BD LSRFortessa. По $1\times10^6$ клеток инкубировали в 300 мкл FACS–буфера (PBS, 1% FCS, 2 мМ EDTA) с FcR–блокирующими антителами (Elabscience®, #E–AB–F0997A®) 15 мин при 4°C, после чего

добавляли CD3–PE (Elabscience®, #E–AB–F1103UD) и CD4–AF488 (Elabscience®, #E–AB–F1097L). DARPin–Cy5 вносили в концентрации до 0,05 мкг/мл. После инкубации 40 мин в темноте при 4 °C клетки дважды отмывали и окрашивали DAPI (0,1 мкг/мл) для исключения мертвых клеток. Специфичность оценивали по профилю и доле положительных клеток относительно контроля CD8–APC.

Способность ЛНЧ, декорированных anti–CD8–DBCO–DARPin, специфично таргетировать CD8+ лимфоциты оценивали *in vitro* по экспрессии GFP на мышиных спленоцитах. Клетки после активации засевали в 96–луночный U–bottom TC–treated планшет (Corning Inc., USA) в количестве 80 000 на лунку в 100 мкл среды (состав описан в п.2.6.4). К клеткам добавляли 750 нг/лунку мРНК–ЛНЧ в объеме 20 мкл буфера. К нетрансфецированному контролю добавляли такой же объем буфера, но без ЛНЧ и тщательно ресуспендировали пипетированием.

На следующий день после внесения ЛНЧ клетки отмывали PBS без белка и окрашивали FVS780 (BD Horizon™, #565388) 15 мин при 4°C. После отмывки FACS–буфером клетки окрашивали CD3–AF700 (BioLegend, #100216), CD4–ElabFluor® Violet 450 (Elabscience®, #E–AB–F1353UQ) и CD8–APC (Elabscience®, #E–AB–F1104UE) в течение 30 мин при 4°C. После отмывки анализ проводили на BD LSRFortessa.

### 2.6.3 Клеточные тесты активности агонистов GLP–1R

Для оценки биологической активности агонистов GLP–1R *in vitro* использовали репортерную клеточную линию HEK293T–CREluc–GLP1R–GFP [169,170]. В качестве положительного контроля применяли семаглутид (Novo Nordisk, Дания). Клетки выращивали до 90% конфлюэнтности, снимали трипсинизацией, промывали ростовой средой и суспендировали в DMEM (PanEco, Россия) с 4,5 г/л глюкозы, 2 мМ L–глутамина и 10% FBS (Cytiva HyClone, США) до концентрации $5\times10^5$ клеток/мл. Клетки высевали в черные непрозрачные 96–луночные планшеты (SPL Lifesciences, Южная Корея) по $5\times10^4$ клеток/лунку и инкубировали 24 ч при 37°C и 5% $CO_2$.

Активность Exe4 без НПА, полученного и очищенного в идентичных условиях, предварительно верифицировали относительно семаглутида. После подтверждения воспроизводимости остальные образцы сравнивали с Exe4.

Исследуемые агонисты растворяли в PBS до 0,1 мг/мл, затем последовательными разведениями (шаг ≤10) получали исходную концентрацию 223,4 нМ и далее 10 точек титрации с шагом разведения 5. В лунки с клетками добавляли по 50 мкл раствора пептида в трех независимых повторах. Отрицательный контроль содержал только среду. Клетки инкубировали 20 ч при 37°C и 5% $CO_2$.

Активность люциферазы определяли с использованием AbiLux Firefly Luciferase Assay Kit (Абисенс, Россия). Из лунок удаляли по 70 мкл супернатанта, к оставшимся 80 мкл добавляли 40 мкл субстрата люциферин Д с лизирующим буфером и инкубировали 3–5 мин при комнатной температуре на шейкере BIOSAN PST–60HL–4 (500 об/мин). Люминесценцию регистрировали на CLARIOstar (BMG LABTECH, Германия).

Сигнал нормализовали относительно максимального ответа агониста (100%) и отрицательного контроля (0%). Кривые «доза–эффект» и EC50 рассчитывали в GraphPad Prism v8.0 и выше с использованием четырёх параметрической логистической модели нелинейной регрессии. Данные принимали при $R^2 \geq 0,95$, CV технических повторов ≤20%, наклоне кривой 0,9–1,4 и различии нижнего и верхнего плато ≤20%; нормальность распределения оценивали критерием Шапиро–Уилка. Для оценки относительной активности рассчитывали относительную потенцию (RP) каждого образца относительно Exe–4 и геометрическое среднее RP биологических повторов. Статистический анализ включал тест Граббса на выбросы, F–тест Фишера для сравнения дисперсий и t–тест Стьюдента для сравнения средних. Для F– и t–тестов использовали $\alpha = 0,05$; статистически значимыми считали различия при $p \leq 0,05$.

### 2.6.4 Цитометрическое окрашивание клеток, экспрессирующих GLP–1R

Окрашивание конъюгатом Exe4–Cy5 клеток человеческой линии HEK293T проводили в 100 мкл буфера PBS (1% FBS) при температуре 4°C в течение 30 минут, далее клетки дважды отмывали в буфере PBS (300g, 5 мин) и анализировали флуоресценцию на проточном цитометре CytoFLEX S (Beckman Coulter, США). В качестве маркера жизнеспособности использовали краситель DAPI (0,05 мкг/мл), который добавляли в образец непосредственно перед анализом. В ходе анализа регистрировали не менее 10 000 событий в гейте живых клеток ($DAPI^-$). Флуоресценцию GFP+ (GLP–1R+) клеток детектировали в канале 525/40 нм (лазер 488 нм), а флуоресцентный сигнал от метки Cy5 регистрировали в канале 660/20 нм (лазер 638 нм).

Способность ЛНЧ, декорированных Exe4, специфично таргетировать GLP–1R оценивали *in vitro* по экспрессии Kate2S. Клетки HEK293T wt и HEK293T GLP1R+/GFP+ культивировали в среде DMEM (Cytiva, США) с добавлением 10% FBS (Cytiva, США), PenStrep 1× (ПанЭко, Россия) и 2 мМ L–Glu. За 48 ч до трансфекции клетки высевали в 48–луночный планшет по 400 мкл питательной среды на лунку. В монокультурах высевали по $3\times10^4$ клеток HEK293T wt или HEK293T GLP1R+ GFP+ на лунку. В ко–культурах высевали по $1{,}5\times10^4$ клеток каждого типа, суммарно $3\times10^4$ клеток на лунку. Клетки инкубировали при 37 °C и 5% $CO_2$.

За 1 ч до трансфекции питательную среду в лунках заменяли на 400 мкл свежей среды того же состава. Использовали липидные наночастицы состава LNP–5, описанного ранее (раздел 2.5.2). ЛНЧ содержали репортерную РНК Kate2S и краситель sPcPHBA. В контрольные лунки без трансфекции ЛНЧ не добавляли. В экспериментальные лунки вносили по 250 нг РНК в составе ЛНЧ: группы с немодифицированными ЛНЧ и группы с ЛНЧ, декорированными Exe4 (ЛНЧ–Exe4). После добавления частиц клетки инкубировали при 37 °C и 5% $CO_2$.

В образцах с предварительной блокировкой GLP–1R после замены среды в лунки добавляли конъюгат Exe4–PEG–DBCO в 10–кратном избытке относительно количества Exe4, содержащегося в порции 250 нг РНК–ЛНЧ. Клетки выдерживали 30 мин при 37 °C и 5% $CO_2$, после чего вносили 250 нг РНК в составе ЛНЧ–Exe4.

Через 24 ч после трансфекции клетки снимали с планшета раствором Версена (ПанЭко, Россия), инкубируя 15 мин при комнатной температуре. Затем клетки отмывали FACS–буфером (PBS, 1% FBS, 0.05% азида натрия) двумя последовательными центрифугированиями при 300g в течение 5 мин. Для оценки жизнеспособности перед анализом добавляли DAPI (0.1 мкг/мл). Анализ проводили на проточном цитометре BD LSRFortessa (BD, США). Детекцию осуществляли в следующих каналах: DAPI – V450/50; GLP–1R/GFP – B530/30; sPcPHBA – YG585/15; Kate2S – Y710/50.

## 2.7 Статистическая обработка данных

Все данные представлены как средние значения ± SD, как указано в подписях к рисункам. Нормальность распределения количественных показателей проверяли с помощью критериев Колмогорова–Смирнова и Шапиро–Уилка. Непараметрический критерий Манна–Уитни использовался для проверки значимости различий между независимыми

количественными выборками, не имеющими нормального распределения. Для сравнения различий между тремя или более независимыми выборками применяли однофакторный дисперсионный анализ с последующим тестом множественных сравнений Шидака. Все результаты получены как минимум в трех биологических повторах, если иное не указано в подписях к рисункам. Статистический анализ и графики были созданы в GraphPad Prism 9.0.0 (GraphPad Software Inc, Сан–Диего, Калифорния, США), а также при помощи R 4.4.1 (The R Foundation for Statistical Computing) в среде RStudio (Posit Software, PBC formerly RStudio, PBC). Различия считались достоверными при уровне значимости $p < 0,05$.

# 3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЯ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ

## 3.1 Разработка и валидация платформы направленной эволюции PylRS

### 3.1.1 Получение генетических последовательностей вариантов MmPylRS и ChPylRS

Наиболее часто используемыми ферментами для включения НПА являются пирролизил–тРНК–синтетазы, выделенные из *M. mazei* и *M. barkeri*. Поскольку MmPylRS обладает большей стабильностью и активностью, она была выбрана в качестве одного из белков–мишеней для направленной эволюции [105,171]. Для более эффективной экспрессии в клетках *E. coli* кодонный состав гена был оптимизирован. Для получения варианта PylRS, который с большей эффективностью способен включать НПА с бензольным кольцом, в последовательность MmPylRS были введены аминокислотные замены Asn346Ala и Cys348Ala. Они были выбраны потому, что, находясь в активном центре, в кармане связывания аминокислоты, существенно влияют на способность PylRS узнавать различные производные фенилаланина, имеющие в м– и о–положениях алкиновые, азидные, кетоновые и галогенбензольные группы [105,107]. Таким образом был собран ген MmPylRS_mut. Схематическое изображение показано на рисунке 3.1 (А). Полученная генетическая последовательность была клонирована в плазмиду pHpyl–MmPylRS–GlnS.

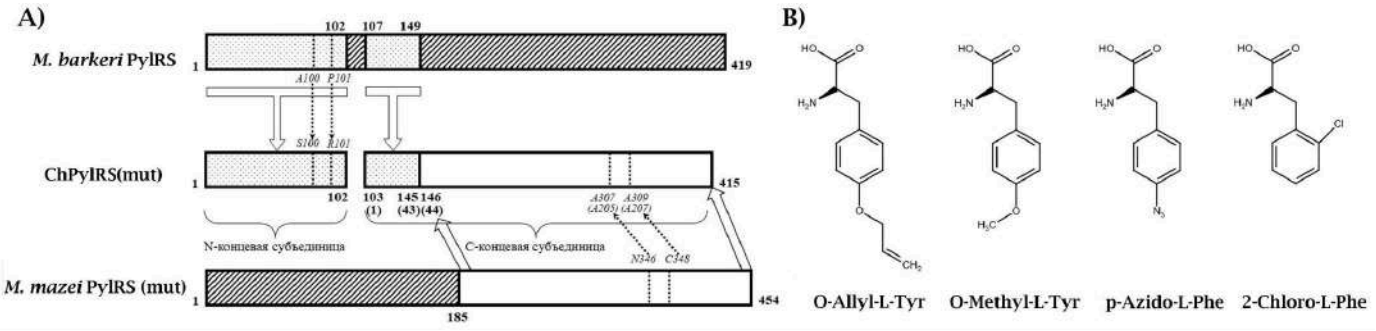



Рисунок 3.1 – **(А)** Структура PylRS с внесенными аминокислотными заменами, использованная в данном исследовании. **(В)** Структуры неприродных аминокислот, использованных в данном исследовании

Поиск в литературе позволил обнаружить вариант химерного белка ChPylRS, в котором N–домен взят из *M. barkeri*, а C–домен – из *M. mazei*, причем между доменами присутствует разрыв [132]. У такого варианта PylRS активность в отношении пирролизина, а также растворимость выше, что выгодно отличает его от ферментов дикого типа. Интересно, что подобная структура с двумя эффекторными доменами и разрывом между ними изначально относится к классу «PylSn+PylSc». В PylRS этого класса домены экспрессируются из двух отдельных генов как два разных белка, после чего собираются в один функциональный PylRS [91]. Благодаря улучшенным свойствам данного варианта было решено использовать его в работе по направленной эволюции. Схематическое изображение гена ChPylRS_mut представлено на рисунке 3.1 (А). Модифицированная химерная конструкция состоит из двух независимо транслирующихся субъединиц: одна содержит 1–102 аминокислоты пирролизил–тРНК–синтетазы *M. barkeri* с аминокислотными заменами A100S и P101R, другая содержит от N– к C–концу 107–149 аминокислоты пирролизил–тРНК–синтетазы *M. barkeri* и 185–454 аминокислоты пирролизил–тРНК–синтетазы *M. mazei* с внесенными аминокислотными заменами Asn346Ala и Cys348Ala (нумерация аминокислот относительно полноразмерной последовательности MmPylRS).

Для сборки гена химерной синтетазы большая часть C–домена была взята из ранее полученного MmPylRS_mut, а последовательность N–домена с фрагментом C–домена из MbPylRS была собрана из синтетических олигонуклеотидов. В качестве матрицы для амплифкации C–концевого фрагмента был выбран ген MmPylRS_mut из плазмиды pHpyl–MmPylRS–GlnS при помощи праймеров с увеличенными концевыми участками, необходимыми для сборки методом Гибсона. Фрагменты гена синтетазы Mb из 12 олигонуклеотидов были собраны по 4 праймера (1–4, 5–8 и 9–12). 12 олигонуклеотидов примерно 30–50 пн длиной содержали перекрывающиеся участки гена и собирались по 4 в небольшие полинуклеотиды длиной примерно 120–150 пн. Получившиеся фрагменты сшивали при помощи концевых олигонуклеотидов с увеличенными концевыми участками, необходимыми для сборки методом Гибсона. После чего полученный ампликон длиной 489 п.н. использовали для сшивки с ампликоном MmPylRS. Очищенные ПЦР–продукты вдалее использовались для сборки в плазмиду pHpyl–ChPylRS–GlnS методом Гибсона.

В результате были получены два гена синтетаз, которые в дальнейшем использовались в работах по направленной эволюции в отношении четырех производных фенилаланина и тирозина, представленных на рисунке 3.1В. MmPylRS_mut была использована в эксперименте с O–аллил–L–тирозина (oAllyY, синтетаза AllYPylRS), 4–азидо–L–фенилаланина (pAzF,

синтетаза AzFPylRS), O–метил–L–тирозина (oMethY, синтетаза MethYPylRS), тогда как ChPylRS_mut проходила эволюцию в отношении 2–хлорфенилаланина (2–ClF, синтетаза 2ClFPylRS).

### 3.1.2 Оптимизация системы направленной эволюции PylRS методом PANCE

В методе фаг–ассоциированной эволюции активность исследуемого белка связывают с жизненным циклом фага так, чтобы при повышении уровня активности или специфичности целевого белка бактериофаги, кодирующие в геноме такой вариант, получали эволюционное преимущество над остальными. Для работы был адаптирован вариант ранее описанной системы PANCE [87].

Для реализации метода PANCE ген gIII в SP заменяют геном, кодирующим целевой белок PylRS. Клетки–хозяева содержат вспомогательную плазмиду (AP) с геном gIII, уровень экспрессии которого зависит от активности PylRS. Селективное давление определяют расположение и количество амбер стоп–кодонов, концентрация НПА и время культивирования бактериофагов. Для генов интереса с низкой начальной активностью, к которым относится PylRS, необходимы несколько раундов положительной селекции с постепенно возрастающим селективным давлением (рисунок 3.2).

Направленная эволюция методом PANCE в данной работе представляет собой четырехэтапный процесс. На первом этапе были использованы клетки *E. coli* S2060, трансформированные двумя AP: в первой амбер стоп–кодон представлен единожды в гене T7–РНКП, тогда как ген gIII, необходимый для формирования вирулентных бактериофагов, расположен под контролем T7–промотора во второй вспомогательной плазмиде (рисунок 3.2, стратегия 1). Такой подход повышает продукцию pIII, поскольку каждое событие подавления амбер стоп–кодона может приводить к образованию нескольких транскриптов gIII.

На втором этапе клетки трансформировали одной AP, содержащей gIII с одним амбер стоп–кодоном, а также тРНК$_{Pyl}$ (рисунок 3.2, стратегия 2). Репликация фага зависит от успешного аминоацилирования тРНК$_{Pyl}$ выбранной PylRS и последующего включения соответствующей НПА в gIII. На этом этапе селективное давление возрастает, поскольку теперь одна копия gIII образуется за счет введения одной НПА в ответ на стоп–кодон.

Концентрацию целевой аминокислоты в среде на этом этапе поддерживали на уровне 10мМ (для всех НПА, кроме oAllylY, порог растворимости которого составлял 5 мМ). На каждом последующем этапе условия отбора был усложнен: плазмида для третьего этапа PANCE несла ген gIII с двумя амбер стоп–кодонами, а для четвертого этапа – с тремя (рисунок 3.2). Кроме того, на четвертом этапе PANCE концентрация НПА постепенно снижалась с 10 мМ в начале до 2 мМ в конце, что создает условия для отбора наиболее эффективных PylRS. Предпринимались попытки перенести бактериофаги в условия с 1 мМ исследуемых НПА. Однако результаты нескольких тестов адаптации при меньшем количестве НПА показали, что наиболее стабильные различия в титрах между присутствием и отсутствием соответствующих НПА в среде наблюдаются при конечной концентрации 2 мМ.

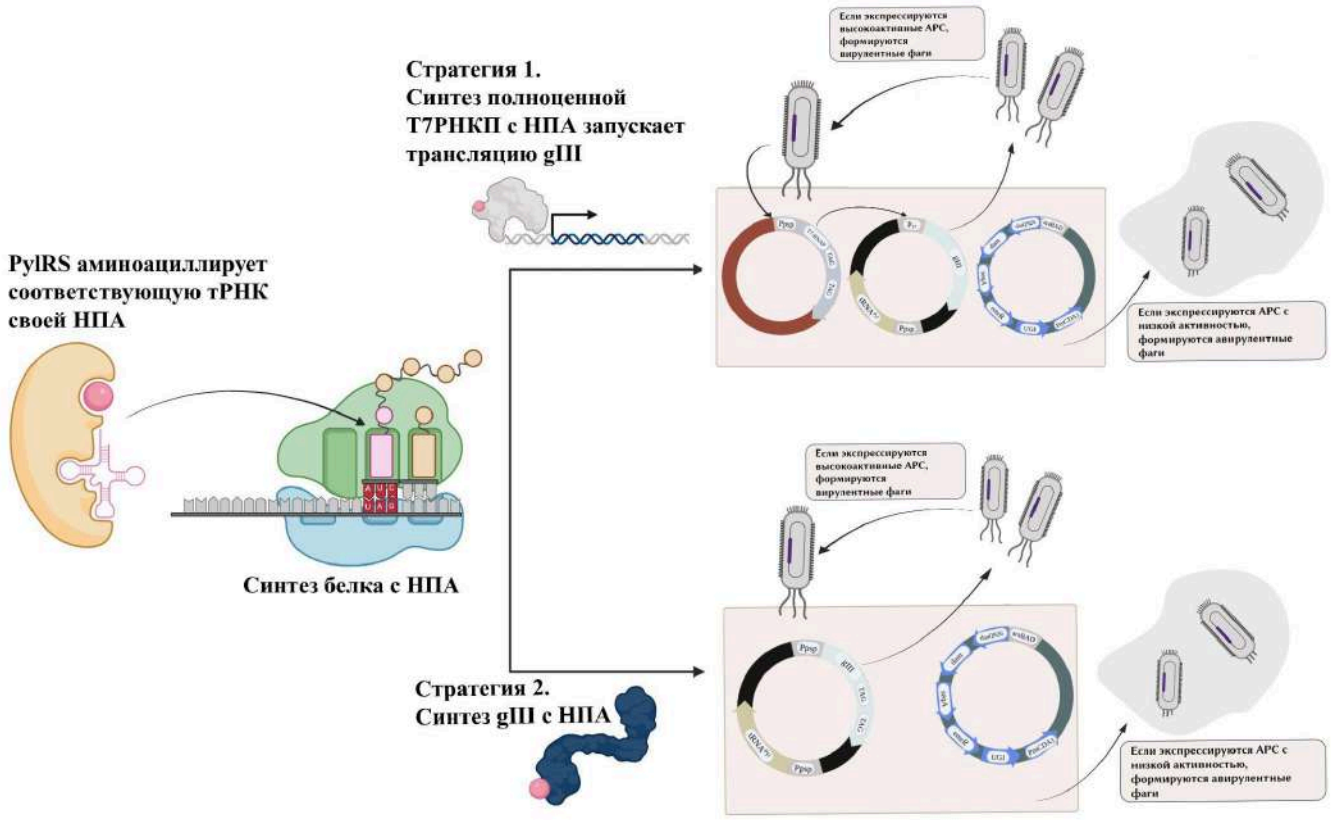



Рисунок 3.2 – Стратегии связывания активности PylRS с жизнеспособностью бактериофага. **Стратегия 1:** аминоацилирование тРНКPyl НПА обеспечивает трансляцию полноразмерной РНК–полимеразы T7 из транскрипта, содержащего амбер стоп–кодоны. Затем РНК–полимераза T7 запускает экспрессию gIII при помощи промотора T7 (PT7). **Стратегия 2:** аминоацилирование тРНКPyl НПА позволяет транслировать полноразмерный pIII из мРНК gIII, содержащей амбер стоп–кодон

Степень адаптации PylRS в SP оценивали по росту титра бактериофага (рисунок 3.3).

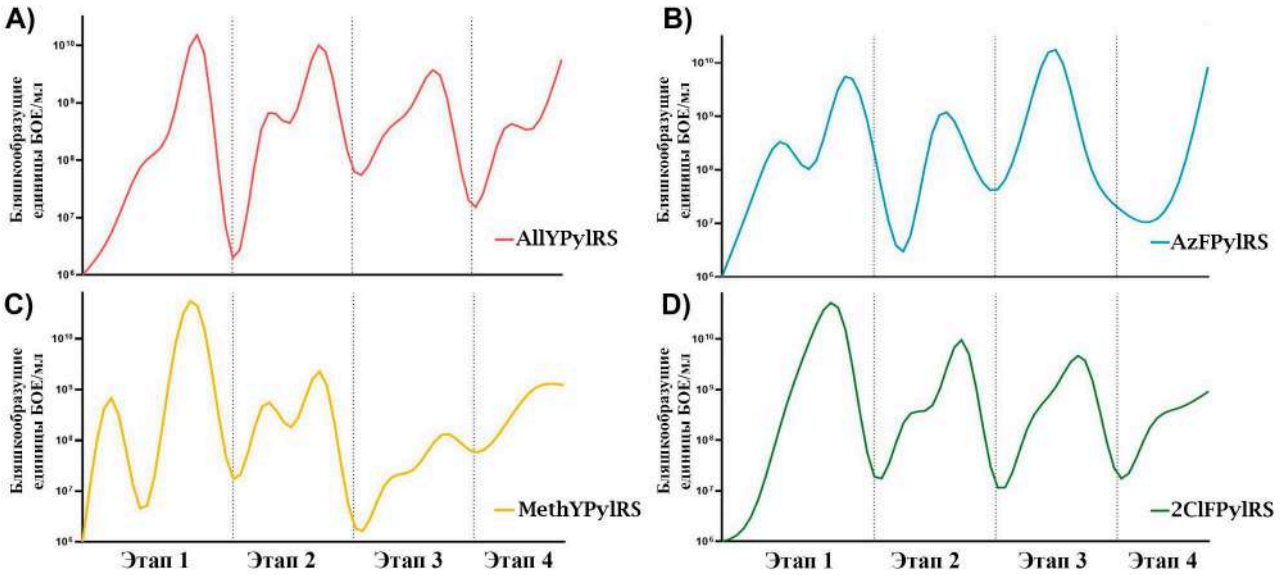



Рисунок 3.3 – Сводные данные по количеству бляшкообразующих единицы на мл фагового супернатанта на 4 этапах положительного отбора PANCE **(A)** O–аллил–L–тирозина (AllYPylRS), **(B)** 4–азидо–L–фенилаланина (AzFPylRS), **(C)** O–метил–L–тирозина (MethYPylRS) и **(D)** 2–хлорфенилаланина (2ClFPylRS). Титры представляют собой индивидуальные измерения, основанные на среднем значении трех технических повторов анализа ПЦР–РВ

На графиках прослеживаются определенные закономерности зависимости титров бактериофагов от условий PANCE. При переводе бактериофага на новый этап селекции скорость его размножения снижалась, что приводило к уменьшению числа частиц на мл клеточной суспензии. Далее, по мере прохождения этапа селекции, титр бактериофагов увеличивался. Например, на третьем и четвертом этапах эволюции наблюдался устойчивый рост, при котором к концу раунда положительного отбора титр достигал максимальных значений порядка $10^{8,5}$–$10^{9,5}$ БОЕ/мл. При переходе к более строгим условиям титр резко снижался ниже $10^{8}$ БОЕ/мл. Такие изменения связаны с появлением и закреплением нуклеотидных и аминокислотных замен, положительно влияющих на активность АРС.

Важно подчеркнуть, что рост титра не отражает функциональную специфичность вариантов PylRS, а служит лишь общим маркером сохранения активности фермента в эволюционной линии. Если при переходе на следующий этап PANCE титр SP необратимо падает, это свидетельствует о необходимости вернуться к предыдущему этапу положительного отбора.

### 3.1.3 Разработка отрицательной селекции

Перечисленные ранее этапы, в совокупности называемые положительной селекцией, приводят к формированию и закреплению нуклеотидных и аминокислотных замен в вариантах PylRS в ходе эволюционного отбора, тогда как последующий этап, называемый отрицательной селекцией, включает элиминацию синтетаз с неспецифической активностью. Этот этап требует плазмиды, содержащей доминантно–негативную форму gIII (gIII–neg). В отличие от положительной селекции, где используется интактный gIII, gIII–neg содержит C$^-$домен с делецией 70 аминокислот, отсутствие которых приводит к образованию фаговой частицы, не способной катализировать свое отсоединение от клетки–хозяина [172]. Однако описанные ранее в литературе методы отрицательной селекции выполняются одновременно с положительным отбором и включают индукцию экспрессии доминантно–негативной формы gIII параллельно с положительной селекцией. Это может приводить к преждевременному удалению ферментов с потенциально полезной активностью, особенно при добавлении в среду НПА, необходимых для положительного отбора во время адаптации АРС к новому субстрату. В связи с этими ограничениями в данном исследовании отрицательная селекция проводилась вручную как отдельный этап после каждого этапа положительного отбора, без добавления НПА.

Для этого были сконструированы две плазмиды: pDB007ns2a–neg и pDB016+tRNApyl. Плазмида pDB016+tRNApyl несет gIII–neg под контролем T7$^-$промотора и тРНК$_{Pyl}$. Плазмида pDB007ns2a–neg содержит интактный ген gIII и ген T7–РНК–полимеразы под контролем промотора фагового шока. Схематическое изображение плазмид и принципа отрицательной селекции показано на рисунке 3.4.

В разработанной в рамках данного диссертационного исследования системе в условиях активации РНК$^-$полимеразы T7, когда PylRS не распознает неприродные аминокислоты и вместо этого включает природные аминокислоты в позицию амбер стоп–кодона в транскрипте

РНК–полимеразы T7, уровень продукции gIII–neg достаточен для доминантного ингибирования сборки инфекционных фагов. Активность плазмиды pDB007ns2a–neg имеет решающее значение, поскольку она обеспечивает сборку активных вирусных частиц, которые будут передаваться дальше по эволюционной линии. Без экспрессии интактного gIII все SP, в том числе с целевой активностью, были бы утеряны. Отрицательный отбор проводили в условиях без добавления НПА к клеткам.

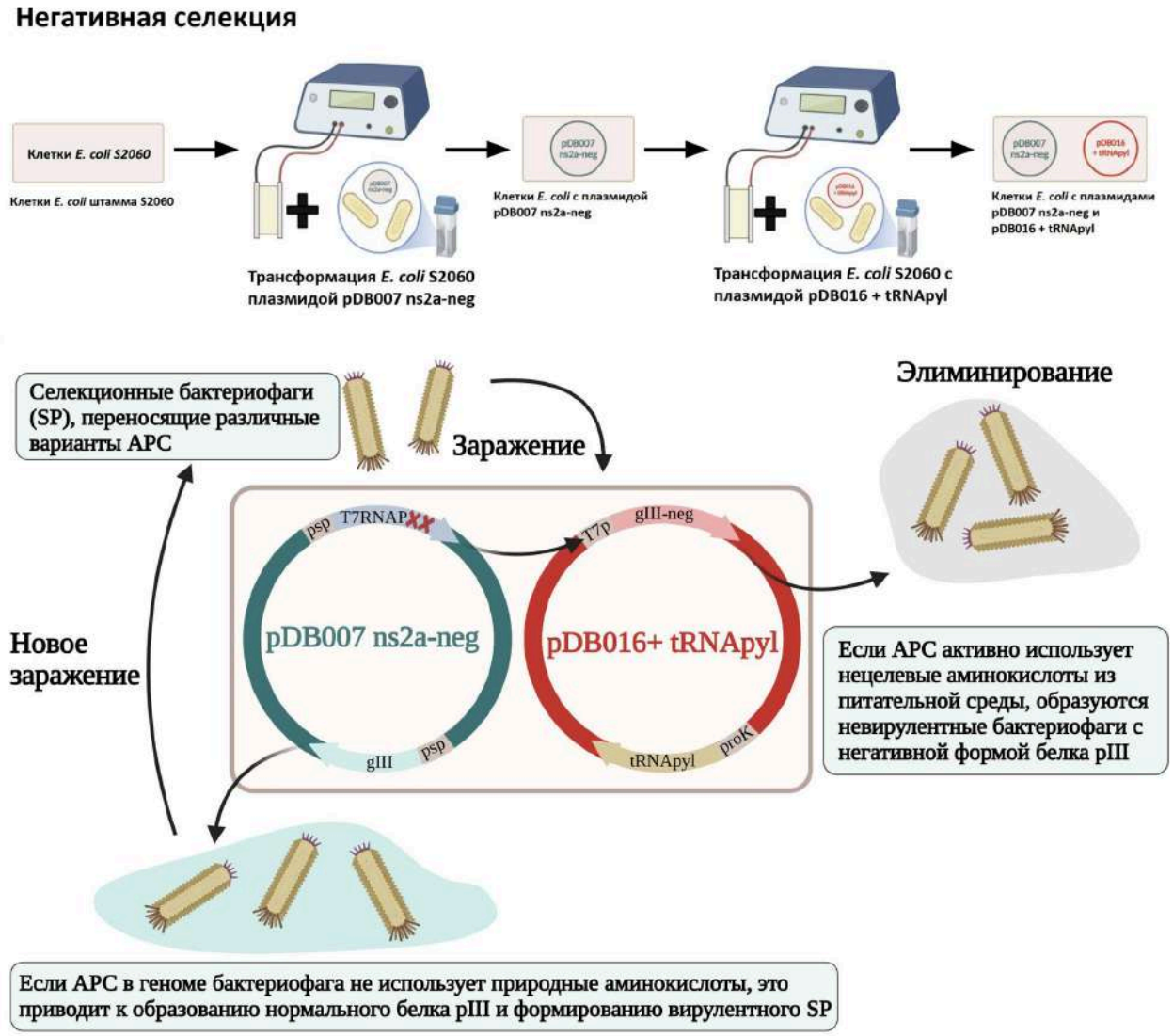



Рисунок 3.4 – Отрицательная селекция: используется плазмида с доминантно–отрицательной формой gIII (gIII–neg) без амбер стоп–кодонов под контролем промотора T7; T7RNAP под контролем psp. Совместно экспрессируется с интактным gIII под контролем psp. Эти плазмиды действуют совместно

Когда экспрессия gIII–neg субоптимальна или неполна, фаговые частицы, несущие варианты PylRS с нежелательной, но не максимальной базальной активностью по отношению к каноническим аминокислотам, могут избегать отрицательной селекции. Чтобы минимизировать этот риск, в данном исследовании была разработана стратегия с несколькими уровнями контроля. После каждого этапа положительного отбора проводился не один, а несколько последовательных раундов отрицательной селекции, что значительно увеличивало вероятность элиминации частиц, избежавших селекции из–за временного или количественного дефицита gIII–neg в отдельных клетках.

Ключевым элементом протокола была обязательная проверка эффективности последовательных раундов отрицательного отбора путем сравнения титров фагов на клеточных культурах с набором плазмид, соответствующим строгости предшествующего этапа положительной селекции, без плазмиды мутагенеза, в присутствии и отсутствии НПА.

Были подготовлены две клеточные культуры: одна – с добавлением соответствующей аминокислоты в концентрации 2–10 мМ в зависимости от этапа, предшествовавшего отрицательной селекции, другая – выращенная в среде 2xYT без НПА. Обе культуры заражали стандартным титром $10^6$ БОЕ/мл из последнего раунда отрицательной селекции. Различия в титрах бактериофагов оценивалось по результатам ПЦР–РВ. Если при добавлении НПА в среду титры отличались от условий без НПА на 1,5–2 порядка и более, считалось, что PylRS достигла достаточной специфичности для дальнейшей направленной эволюции и перехода к следующему этапу. Недостаточное различие рассматривалось как индикатор возможного избегания отрицательной селекции и основание для возвращения к предыдущему этапу эволюции.

После 12 раундов четырех этапов положительного отбора и дополнительного раунда отрицательной селекции для каждой эволюционной линии бактериофаги были протестированы в самых строгих условиях отбора – в присутствии соответствующих НПА и без них. Титры показали, что эволюционировавшие линии фагов реплицировались значительно эффективнее в присутствии НПА, что указывает на появление более активных вариантов PylRS (рисунок 3.5).

Использованная итеративная стратегия обратной связи была специально разработана для обеспечения высокой строгости отбора и минимизации вероятности обогащения библиотек эволюционировавших ферментов вариантами, включающими природные аминокислоты. В рамках этого подхода многоступенчатая валидация успешно компенсировала неопределенность в уровнях экспрессии pIII/gIII–neg.

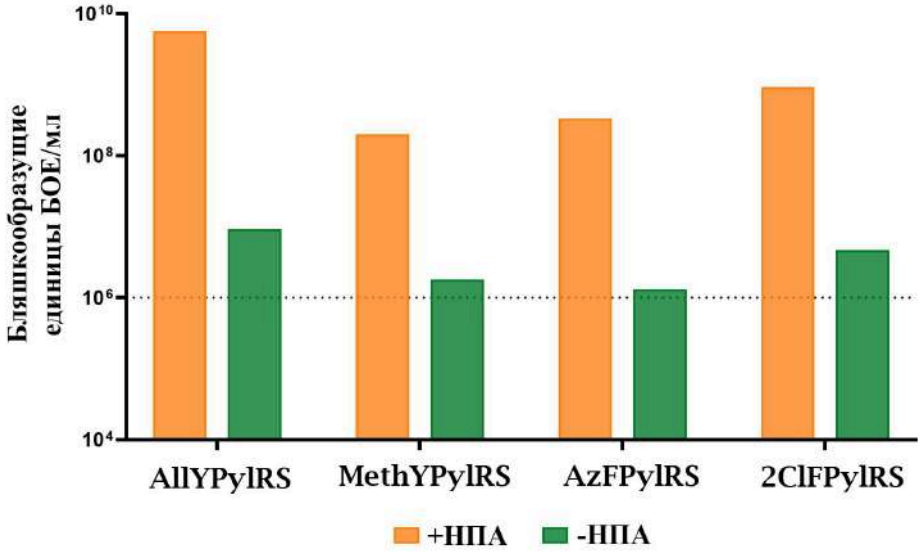



Рисунок 3.5 – Измерения титров фагов: сравнение эволюционировавших фагов в присутствии и отсутствии неприродной аминокислоты. Приведенные значения титров представляют собой результаты отдельных измерений, рассчитанные как среднее арифметическое трех технических повторов анализа методом количественной ПЦР

### 3.1.4 Анализ динамики эволюции PylRS и накопления замен

Всего для каждой АРС с соответствующей НПА было проведено 50 раундов направленной эволюции. Для идентификации и сравнения приобретенных нуклеотидных замен выполнено ампликон–ориентированное секвенирование на платформе Illumina MiSeq (таблица 3.1). Это позволило проследить формирование эволюционных траекторий. По результатам секвенирования видно, что после PANCE формируется не одна доминирующая траектория, а полиморфная популяция вариантов PylRS, причем состав нуклеотидных и аминокислотных замен зависит от НПА. Едиными для всех популяций остаются предварительно введенные аминокислотные замены Asn346Ala и Cys348Ala (Asn205Ala и Cys207Ala в С–домене ChPylRS), которые сохранялись во всех линиях.

Таблица 3.1 – Нуклеотидные и аминокислотные замены, выявленные методом MiSeq, и их частота в вариантах PylRS на различных этапах PANCE

| AllYPylRS | | | | | | |
|---|---|---|---|---|---|---|
| Номер нуклеотида | Нуклеотидная замена | Аминокислотная замена | Частота Этап 1 | Частота Этап 2 | Частота Этап 3 | Частота Этап 4 |
| 4 | G => A | Asp2Asn | 0,7010 | 0,9996 | 0,9995 | 0,9990 |
| 262 | G => A | Glu88Lys | | | | 0,9200 |
| 549 | G => A | Val183Val | | 0,4363 | 0,6933 | 0,7663 |
| 899 | T => C | Met300Thr | 0,6347 | 0,7089 | 0,8255 | 0,6886 |
| | | | | | | |
| MethYPylRS | | | | | | |
| Номер нуклеотида | Нуклеотидная замена | Аминокислотная замена | Частота Этап 1 | Частота Этап 2 | Частота Этап 3 | Частота Этап 4 |
| 192 | G => A | Met17Leu | | | 0,9986 | 0,9986 |
| 899 | T => C | Met300Thr | 0,7921 | 0,7182 | 0,8315 | 0,7150 |
| 1009 | G => A | Glu337Lys | | | 0,9941 | 0,9985 |
| 1223 | A => G | Asp408Gly | 0,5076 | 0,9991 | 0,9988 | 0,9989 |
| | | | | | | |
| AzFPylRS | | | | | | |
| Номер нуклеотида | Нуклеотидная замена | Аминокислотная замена | Частота Этап 1 | Частота Этап 2 | Частота Этап 3 | Частота Этап 4 |
| 4 | G => A | Asp2Asn | | | 0,9990 | 0,9979 |
| 750 | C => T | Ile250Ile | | | | 0,8659 |
| 899 | T => C | Met300Thr | 0,8383 | 0,7344 | 0,8249 | 0,4588 |
| 916 | T => C | Tyr306His | | 0,1631 | 0,9979 | 0,9982 |
| 1069 | G => A | Glu357Lys | | | | 0,4403 |
| | | | | | | |
| 2ClFPylRS | | | | | | |
| Номер нуклеотида | Нуклеотидная замена | Аминокислотная замена | Частота Этап 1 | Частота Этап 2 | Частота Этап 3 | Частота Этап 4 |
| 184 | C => T | His62Tyr | 0,1651 | 0,3816 | 0,9616 | 0,9975 |
| 623 | A => G | Asn202Ser | | 0,1274 | 0,0763 | 0,9660 |
| 800 | T => C | Met261Thr | 0,8122 | 0,9989 | 0,9986 | 0,9989 |

Наиболее интересным результатом стало независимое появление замены Met300Thr (Met261Thr в ChPylRS) во всех четырех эволюционных линиях: AllYPylRS, MethYPylRS, AzFPylRS и 2ClFPylRS. Уже на первом этапе положительного отбора Met300Thr присутствовала с высокой частотой в AllYPylRS (63,47%), MethYPylRS (79,21%), AzFPylRS (83,83%) и 2ClFPylRS (81,22%) и сохранялась на высоком уровне в течение нескольких этапов. Такое независимое появление одной и той же аминокислотной замены при различных селекционных траекториях указывает на ее вероятную общую функциональную роль – не столько в распознавании конкретной боковой группы НПА, сколько в перестройке каталитического центра и оптимизации взаимодействия с ароматическими производными фенилаланина и тирозина, использованными в данном исследовании. Однако на финальной стадии PANCE частота встречаемости данной замены снижалась до 68,86% в AllYPylRS, 71,50% в MethYPylRS и особенно резко – до 45,88% в AzFPylRS. Это указывает на раннее селективное преимущество Met300Thr, однако недостаточное для окончательной адаптации фермента. По–видимому, на последующих этапах другие замены могут становиться более выгодными и частично вытеснять варианты, содержащие Met300Thr.

Конвергентная динамика отмечена также для Asp2Asn. В AllYPylRS она присутствовала уже на первом этапе с частотой 70,10% и практически полностью фиксировалась на втором (99,96%), сохраняясь на уровне около 99,9% до последней стадии. В AzFPylRS эта же замена отсутствовала на первых двух этапах, но, появившись на третьем, с высокой частотой сохранялась до конца (99,79%).

Для AzFPylRS наиболее выраженное обогащение характерно для Tyr306His: после появления на втором этапе с частотой 16,31% ее доля резко возросла до 99,79%. При этом особенно показательно, что на третьем этапе эволюции частота замены Tyr306His практически фиксируется одновременно с сохранением Met300Thr на уровне 82,49%. Следовательно, наиболее вероятно, что Tyr306His не просто сменила раннюю Met300Thr, а стала частью новой, более эффективной комбинации аминокислотных замен.

Особенно интересна замена Asn100Ser в линии 2ClFPylRS: ее частота сначала снизилась с 12,74% на втором этапе до 7,63% на третьем, но затем резко увеличилась до 96,60% на четвертом. Такая динамика может указывать на то, что замена, которая на ранних этапах не давала существенного преимущества, могла стать значительно более выгодной после появления других замен. Это может свидетельствовать о возможном наличии эпистатических взаимодействий между аминокислотными заменами.

Некоторые замены быстро увеличивали свою частоту на ранних этапах и затем сохранялись на высоком уровне до конца протокола PANCE. Например, Asp408Gly в MethYPylRS, частота которой увеличилась с 50,8% до 99,9% между первым и вторым этапами. Такая быстрая стабилизация может указывать на сильное положительное влияние этих замен на ранних этапах эволюции. В то же время ряд замен появлялся только на поздних этапах: Met17Leu и Glu337Lys в MethYPylRS возникли на третьем этапе сразу с частотами 99,86% и 99,41%, а Glu88Lys в AllYPylRS – только на четвертом с частотой 92,00%.

В совокупности эти данные показывают, что эволюция PylRS при использовании метода PANCE протекает нелинейно. Четыре эволюционные линии развиваются на основе общего эволюционного каркаса с различающимися поздними ответвлениями. Общим элементом является замена Met300Thr, тогда как отдельные эволюционные линии затем используют разные комбинации аминокислотных замен. Одни, вероятно, обеспечивают раннее повышение общей активности PylRS и быстро распространяются в популяции, тогда как другие становятся полезными только на более поздних этапах и могут участвовать в дальнейшей оптимизации активности или специфичности фермента. Это наглядно демонстрирует, что конечный вариант PylRS является не результатом простого накопления независимо полезных замен, а динамического отбора комбинаций аминокислотных замен, в которых эффект одной замены может зависеть от присутствия других.

### 3.1.5 Оценка активности отобранных вариантов PylRS

Для оценки активности полученных вариантов АРС был проведен флуоресцентный тест с использованием белка sfGFP в системе экспрессии на основе *E. coli* BL21 Star. В рамках данной работы тест-система состояла из двух плазмид: pHpyl и pET. По окончании PANCE АРС из четырех эволюционных библиотек были амплифицированы для встраивания в плазмиду pHpyl. АРС находилась в плазмиде pHPyl–PylRS–GlnS под контролем конститутивного промотора GlnS – химерного промотора кишечной палочки. Кроме того, по литературным данным была подобрана и встроена между сайтом +1 промотора GlnS и стартовым кодоном гена 5'–нетранслируемая область, которая, согласно биоинформатическим предсказаниям, максимизирует эффективность инициации трансляции [173].

sfGFP в плазмиде pET–sfGFP–27TAG, в отличие от стандартного GFP, быстрее созревает и обладает толерантностью к НПА в своем составе. Флуоресценция sfGFP использовалась для измерения активности PylRS. Добавление амбер стоп–кодона в рамку считывания sfGFP приводит к преждевременной терминации трансляции при отсутствии АРС, обеспечивающей встраивание НПА. Если PylRS по каким–либо причинам не может аминоацилировать тРНК соответствующей НПА, синтез белка–репортера также обрывается, и флуоресценция не наблюдается. Амбер стоп–кодон был расположен в 27–м положении, которое, согласно литературным данным, является наиболее распространенным для замены на производные тирозина и фенилаланина [173].

Для анализа флуоресценции была проведена нормализация по оптической плотности культуры, а также вычитание фонового сигнала для среды с добавлением и без добавления НПА для соответствующих культур. Отношение флуоресценции к $OD_{600}$ клеток, экспрессирующих варианты PylRS, сравнивалось с клетками, экспрессирующими PylRS дикого типа, как в присутствии, так и в отсутствии соответствующих аминокислот. Для каждой линии было протестировано более 22 клонов, а результаты для лучших из них представлены на рисунке 3.6.

Флуоресцентный анализ показал, что варианты PylRS, созданные с использованием PANCE, эффективно включают четыре целевые неприродные аминокислоты. Каждая из четырех PylRS демонстрирует высокую активность по отношению к соответствующей НПА и на порядок превосходит синтетазу, не проходившую эволюцию. Без соответствующей НПА в среде эволюционировавшие клоны демонстрировали низкий уровень экспрессии sfGFP, тогда как добавление НПА приводило к сверхэкспрессии sfGFP. У некоторых клонов с наиболее активными вариантами синтетаз интенсивность флуоресценции превышала 50% интенсивности положительного контроля. Этот результат высок и сопоставим с данными других исследований, использующих анализы на основе sfGFP [126,174]. Для дальнейшего анализа включения НПА были отобраны наиболее активные клоны каждой PylRS в эксперименте с соответствующей аминокислотой (обозначены зелеными стрелками на рисунке 3.6).

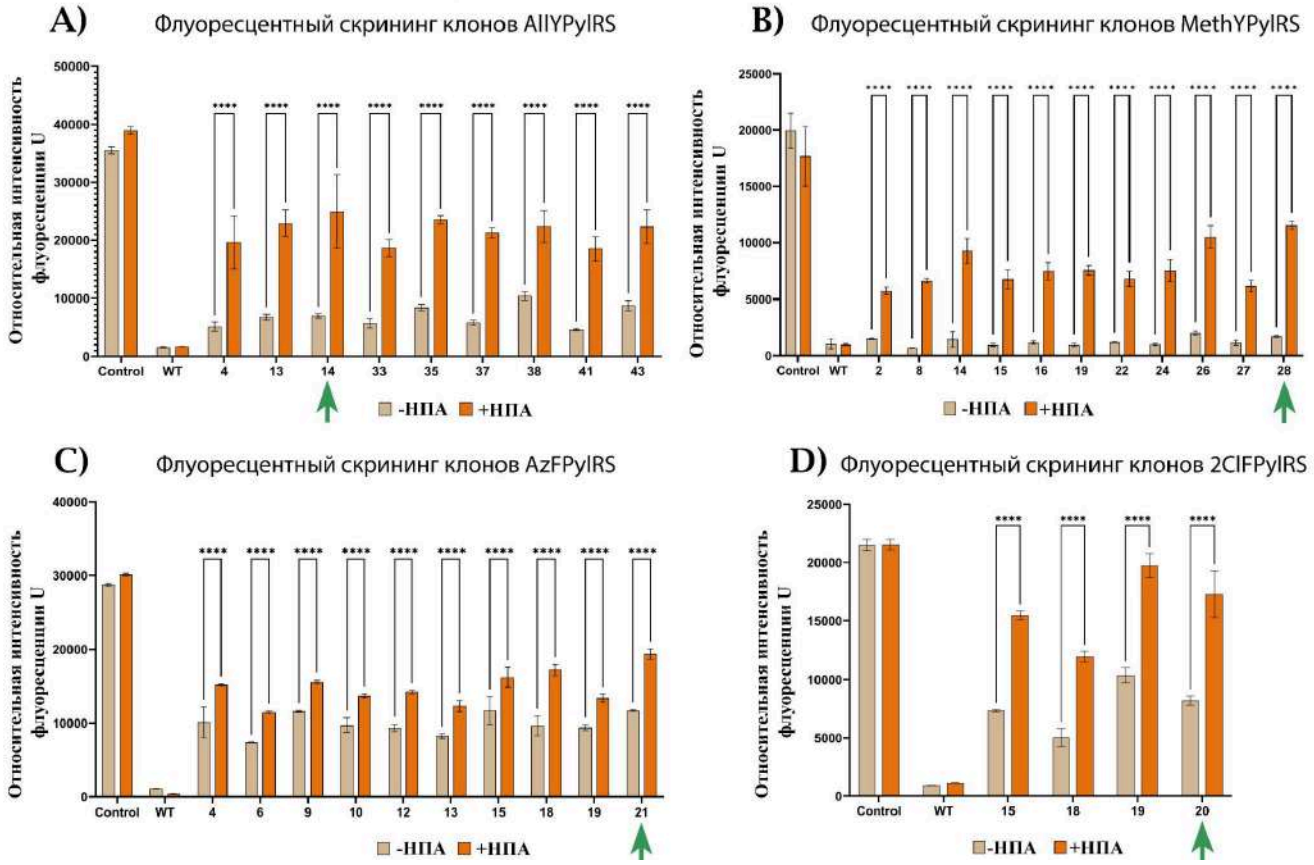



Рисунок 3.6 – Измерения флуоресценции наиболее активных вариантов с добавлением и без добавления **(A)** O–аллил–L–тирозина (AllYPylRS), **(B)** O–метил–L–тирозина (MethYPylRS), **(C)** 4–азидо–L–фенилаланина (AzFPylRS) и **(D)** 2–хлорфенилаланина (2ClFPylRS). Положительный контроль: клетки, экспрессирующие sfGFP без стоп–кодона. WT: клетки, содержащие исходную PylRS, которая не подвергалась PANCE (MmPylRS_mut для O–аллил–L–тирозина, O–метил–L–тирозина и 4–азидо–L–фенилаланина; ChPylRS_mut для 2–хлорфенилаланина). Данные представлены как среднее значение ± SD (n=6). ****p<0,0001. Для обработки использованы односторонний дисперсионный анализ, тест множественного сравнения по Шидаку. Зеленые стрелки указывают на клоны, выбранные для дальнейшего анализа

Ввиду структурного сходства всех четырех НПА, использованных в этом исследовании, была изучена перекрестная активность эволюционировавшей PylRS в отношении нецелевых субстратов (рисунок 3.7). Синтетазы MmPylRS_mut и ChPylRS_mut в целом сохраняли низкую активность, однако активность первой в отношении oAllylY и 2–ClF изначально была относительно высокой. Как и ожидалось, каждая из полученных методом эволюции синтетаз

продемонстрировала наибольшую активность включения целевой аминокислоты в sfGFP. Однако была выявлена характерная кросс–специфичность: AllYPylRS включала 2–ClF (36%), тогда как 2ClFPylRS использовала oAllylY (18,2%), при этом взаимодействие этих ферментов с остальными неприродными аминокислотами было значительно слабее. Также следует отметить сходные уровни флуоресценции, наблюдаемые для субстратов oMeY (34,7%) и pAzF (30,7%) в случае MethYPylRS, что указывает на ее способность к перекрестному взаимодействию.

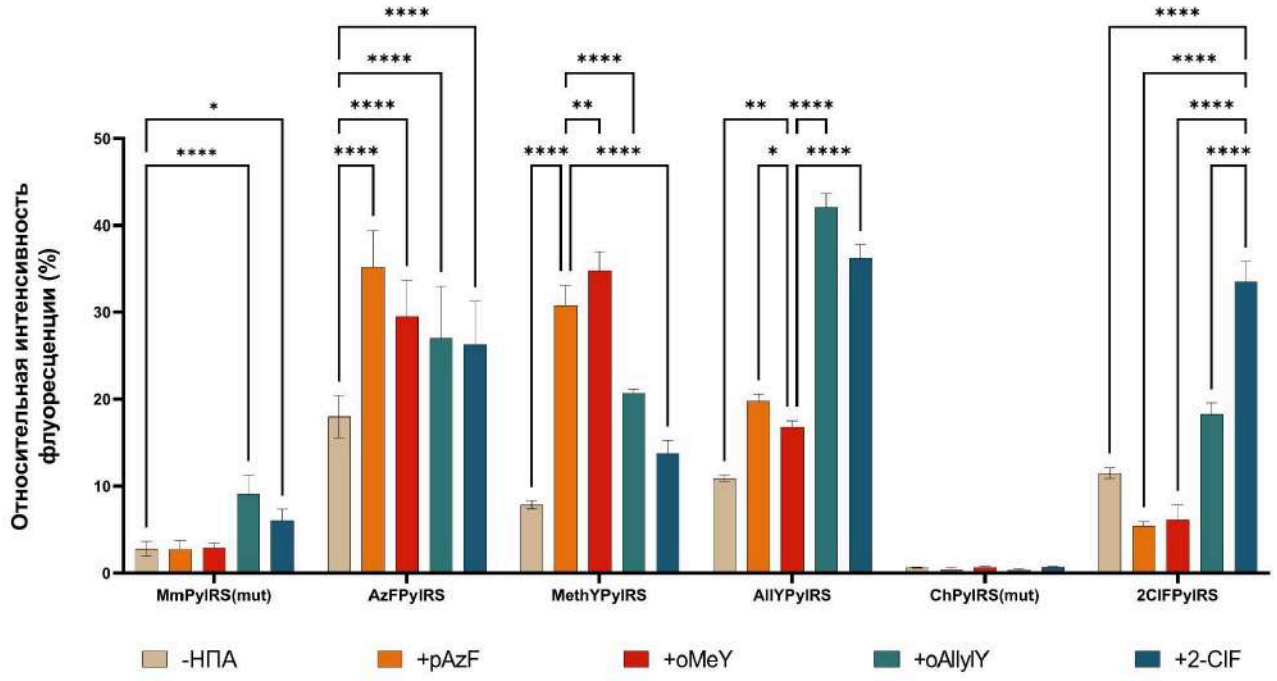



Рисунок 3.7 – Относительная интенсивность флуоресценции с и без добавления О–аллил–L–тирозина, О–метил–L–тирозина, 4–азидо–L–фенилаланина и 2–хлорфенилаланина для наиболее активных вариантов каждой эволюционной линии и изначальных синтетаз (MmPylRS_mut и ChPylRS_mut). Данные представлены как среднее значение ± SD (n=9). *p<0,05, **p<0,01, ***p<0,001, ****p<0,0001. Для обработки использованы односторонний дисперсионный анализ, тест множественного сравнения по Шидаку

В то же время AzFPylRS проявила высокую активность как в отношении своего основного субстрата pAzF (35,2%), так и других НПА: oMeY (29,5%), oAllylY (27%) и 2–ClF (26,2%). Несмотря на различия во флуоресценции, их выраженность была относительно

небольшой, что указывает на отсутствие резкой субстратной избирательности и свидетельствует о достаточно широкой субстратной специфичности AzFPylRS.

Эти результаты показывают, что повышение активности PylRS не всегда приводит к увеличению специфичности. Фермент может эффективнее включать целевую НПА, но при этом сохранять способность включать структурно сходные аминокислоты. Вероятно, изменения в кармане связывания, которые позволяют ферменту лучше взаимодействовать с целевой НПА, одновременно облегчают связывание других похожих субстратов.

Для всесторонней оценки активности включения целевых НПА очищенные образцы sfGFP подвергали масс–спектрометрическому анализу. Клетки *E. coli* BL21 Star одновременно трансформировали плазмидой pHPyl–PylRS–GlnS, несущей соответствующий эволюционированный ген PylRS, и плазмидой pET–sfGFP–27TAG, кодирующей репортерный белок со стоп–кодоном TAG в позиции 27. Экспрессию индуцировали в присутствии 2 мМ целевой НПА. Анализ интактной массы подтвердил ожидаемую молекулярную массу белков, продемонстрировав специфическое включение целевого аналога в sfGFP в заданной позиции. Кроме того, количественно определили долю белков, содержащих НПА, относительно общего пула идентифицированных полноразмерных белков. Во всех экспериментальных условиях доля целевого продукта превышала 89% (рисунок 3.8).

| | | AllYPylRS | MethYPylRS | AzFPylRS | 2ClFPylRS |
|---|---|---|---|---|---|
| **Название** | **Масса** | | | | |
| sfGFP | 27452.6 | 4.1 | 1.3 | 9.6 | 10.7 |
| sfGFP_F27oAllylY | 27508.7 | **95.9** | | | |
| sfGFP_F27oMetY | 27482.6 | | **98.7** | | |
| sfGFP_F27pAzF | 27493.6 | | | **90.4** | |
| sfGFP_F272-ClF | 27486.6 | | | | **89.3** |

Рисунок 3.8 – Относительные значения доли sfGFP, содержащего целевую НПА (показаны синим цветом), полученные методом масс–спектрометрического анализа

Наиболее интенсивные минорные пики в масс–спектрах соответствовали массе белка с включенным фенилаланином – характерному артефакту, обусловленному неполным подавлением трансляции стоп–кодона и конкурентным включением природной аминокислоты. Современные методы направленной эволюции, включая PANCE, нацелены на максимальную

специфичность; однако достижение абсолютной (100%) зависимости от НПА в живых клетках остается сложной задачей. Остаточные фоновые уровни обусловлены сложной клеточной физиологией, включая конкурентные процессы трансляции.

### 3.1.6 Структурный анализ выявленных аминокислотных замен PylRS

Высокопроизводительное секвенирование на платформе MiSeq позволило идентифицировать основные линии, сформированные в ходе направленной эволюции. Последующий скрининг с использованием флуоресцентного репортерного анализа выявил лучшие клоны, содержащие как консервативные, так и новые, ранее не описанные минорные замены, определяющие улучшенные функциональные свойства PylRS (таблица 3.2).

Таблица 3.2 – Аминокислотные замены PylRS, возникшие в результате PANCE, и их локализация в белке, а также соответствующие аминокислотные остатки в *M. mazei*, определенные методом множественного выравнивания последовательностей. Замены, обнаруженные с помощью секвенирования на платформе Illumina MiSeq, выделены зеленым цветом. Замены, обнаруженные в лучших по итогам флуоресцентного скрининга клонах, методом секвенирования по Сэнгеру, выделены красным цветом. Замены, возникшие независимо во всех клонах, выделены синим цветом. Замены, не приводящие к замещению аминокислот, выделены серым цветом

| AllYPylRS | | | | | | |
|---|---|---|---|---|---|---|
| Замена | Asp2Asn | Glu88Lys | Ala107Thr | Asn280Asp | Met300Thr | Asn420Ser |
| Локализация | Карман связывания с тРНК в NTD | Карман связывания с тРНК в NTD | Вариабельный линкер между NTD и CTD | Каталитический домен | Каталитический домен Активный центр | Каталитический домен Активный центр |
| Назначение | Образует карман для связывания тРНК | Образует карман для связывания тРНК | Изменение гибкости междоменного линкера | Стабилизация конструкции | Образует карман для связывания аминокислоты | Образует карман для связывания аминокислоты |
| Соответствие *M.mazei* | Asp2 | Glu88 | Ala107 | Asn280 | Met300 | Asn420 |

Продолжение таблицы 3.2

| | | | | | | |
|---|---|---|---|---|---|---|
| MethYPylRS | | | | | | |
| Замена | Met17Leu | Glu39Lys | Pro109Ser | Met300Thr | Glu337Lys | Asp408Gly |
| Локализация | Карман связывания с тРНК в NTD | Карман связывания с тРНК в NTD | Вариабельный линкер между NTD и CTD | Каталитический домен Активный центр | Каталитический домен Мотив 2 | Каталитический домен Активный центр |
| Назначение | Образует карман для связывания тРНК | Образует карман для связывания тРНК | Изменение гибкости междоменного линкера | Образует карман для связывания аминокислоты | Образует карман для связывания аминокислоты | Образует карман для связывания аминокислоты |
| Соответствие *M.mazei* | Met17 | Glu39 | Pro109 | Met300 | Glu337 | Asp408 |
| | | | | | | |
| AzFPylRS | | | | | | |
| Замена | Asp2Asn | Arg55Cys | Ile250Ile | Met300Thr | Tyr306His | Glu357Lys |
| Локализация | Карман связывания с тРНК в NTD | Карман связывания с тРНК в NTD | Каталитический домен Мотив 1 | Каталитический домен Активный центр | Каталитический домен Активный центр | Каталитический домен |
| Назначение | Образует карман для связывания тРНК | Образует карман для связывания тРНК | | Образует карман для связывания аминокислоты | Образует карман для связывания аминокислоты | Стабилизация конструкции |
| Соответствие *M.mazei* | Asp2 | Arg55 | Ile250Ile | Met300 | Tyr306 | Glu357 |
| | | | | | | |
| 2ClFPylRS | | | | | | |
| Замена | His62Tyr | Asn202Ser | Glu210Lys | Met261Thr | | |
| Локализация | Карман связывания с тРНК в NTD | Каталитический домен | Каталитический домен Мотив 1 | Каталитический домен Активный центр | | |
| Назначение | Образует карман для связывания тРНК | Стабилизирует конструкцию | Участвует в связывании с АТФ | Образует карман для связывания аминокислоты | | |
| Соответствие *M.mazei* | His62 | Asn241 | Glu249 | Met300 | | |

Эволюция PylRS с различными НПА привела к накоплению множественных аминокислотных замен в N–концевом домене. Наблюдаются конвергентные замены: AllPylRS и

AzFPylRS имеют одну и ту же замену Asp2Asn, которая, согласно литературным данным, существенно повышает каталитическую активность фермента. Показано, что Asp2Asn в три раза увеличивает активность PylRS по отношению к 3BrF [124,132]. Предполагается, что механизм связан с нарушением водородных связей между синтетазой и Т–петлей тРНК$_{Pyl}$, что позволяет формироваться альтернативным контактам, таким как Glu88Lys и Arg55Cys, обнаруженным в AllPylRS и AzFPylRS соответственно. Аминокислотная замена Glu88Lys ранее в литературе не описывалась, тогда как замена в 55-м положении, как известно, участвует в образовании водородных связей с нуклеотидом A46 тРНК [124], а замена Arg55Cys может изменить локальное электростатическое окружение и систему водородных связей в области этого контакта. Структурная модель предполагает, что остаток Cys55 может образовывать альтернативные контакты с остовом тРНК (рисунок 3.9A).

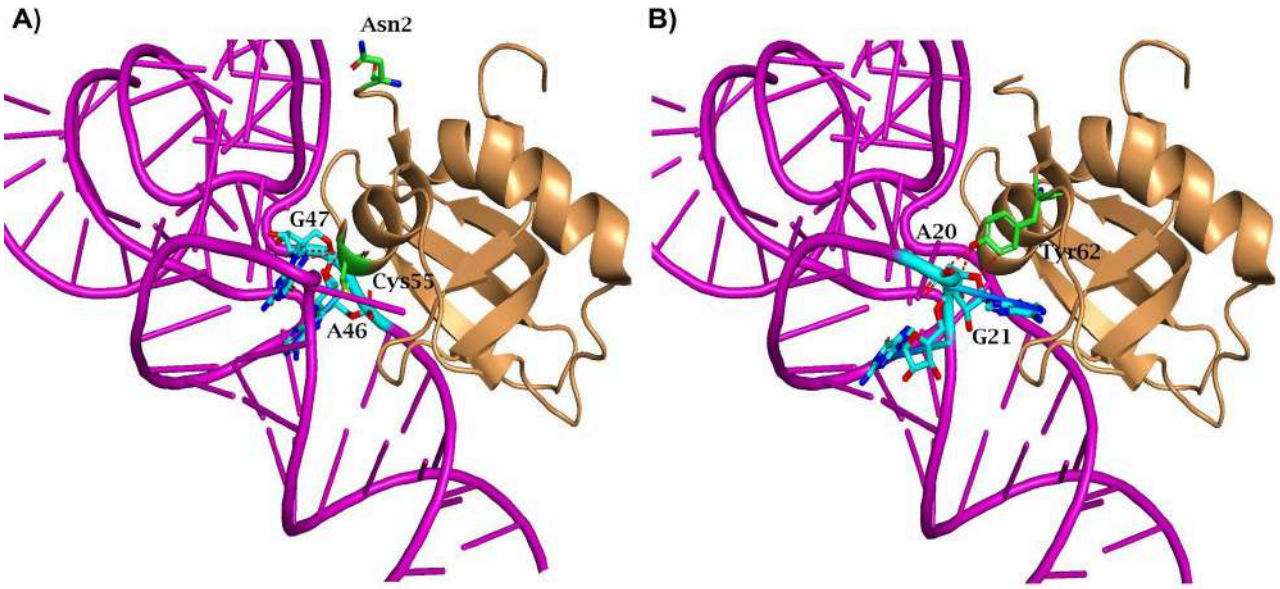



Рисунок 3.9 – Структура комплекса NTD MmPylRS с тРНК$_{Pyl}$ из **(A)** AzFPylRS и **(B)** 2ClFPylRS. Аминокислотные замены, произошедшие во время PANCE, обозначены зеленым цветом. Красные пунктирные линии представляют потенциальные водородные связи. Структура основана на записи PDB 5V6X

В случае 2ClFPylRS была обнаружена аминокислотная замена His62Tyr, ранее описанная для ChPylRS, которая повышает эффективность подавления амбер стоп–кодона [132,175].

Согласно результатам 3D–моделирования и литературным данным, замена в положении 62 нарушает две водородные связи между PylRS и фрагментами Т–петли тРНК$_{Pyl}$, хотя это частично компенсируется слабыми взаимодействиями с фосфатом G21 и основанием A20 (рисунок 3.9 В) [87]. Дополнительные взаимодействия тРНК с С–концевым доменом ChPylRS, по–видимому, усиливаются заменами Asn202Ser и Glu210Lys (соответствующими Asn241 и Glu249 в *M.mazei*). Хотя эти аминокислотные замены ранее не описывались в литературе, их роль может быть связана со стабилизацией фермента. В частности, замена Glu210Lys затрагивает высококонсервативный MOTIF 1 – канонический структурный элемент, характерный для АРС класса II [176]. Согласно литературным данным, MOTIF 1 играет ключевую роль в димеризации PylSn и PylSc, что имеет решающее значение для формирования функционального фермента, способного координировать молекулы АТФ, тРНК и некодирующих аминокислот [77]. Замена Glu210Lys, вероятно, приводит к образованию дополнительных водородных связей, которые, в свою очередь, могут усиливать контакты с тРНК$_{Pyl}$ за счет изменения конформационной динамики доменов.

Аминокислотные замены Met17Leu и Glu39Lys, которые закрепились в MethYPylRS в процессе эволюции, ранее не рассматривались как мишени для рационального дизайна PylSn PylRS. Можно предположить, что их механизм действия аналогичен описанным ранее: замена Met17Leu, вероятно, повышает стабильность структуры белка, тогда как Glu39Lys формирует сеть контактов с перераспределением электронной плотности, способствуя смещению взаимодействий с тРНК$_{Pyl}$. Замена Glu337Lys в MethYPylRS также подтверждает гипотезу о перераспределении взаимодействий. Она локализована в PylSc, в высококонсервативном MOTIF 2, который отвечает за связывание АТФ и тРНК [77]. Аналогичный механизм связывания тРНК PylSc ранее был описан для DhPylRS [177]. Эта АРС, принадлежащая к классу «PylSn+PylSc», проявляет значительную активность *in vivo* даже в отсутствие полноценного РНК-связывающего домена [91]. Показано, что её С-концевой домен образует обширную сеть контактов с тРНК, включая ключевые взаимодействия в положениях Lys166:C72 и His168:C74 (соответствующих Lys366 и His368 в PylRS *M. mazei*), которые связывают тРНК вблизи A76, что имеет решающее значение для аминоацилирования [177]. Таким образом, замена Glu337Lys в MethYPylRS, по-видимому, направлена на установление столь же прочных контактов с тРНК$_{Pyl}$ со стороны PylSc.

В совокупности эти данные позволяют предположить, что выбранные изменения в PylSn снижают сродство этого домена к тРНК. Это способствует перераспределению взаимодействий с тРНК в сторону PylSc, что в конечном итоге приводит к формированию более прочного и

специфичного комплекса PylRS/тРНК и повышению эффективности аминоацилирования с новыми субстратами.

Помимо замен, обнаруженных в N–домене, большое количество замен, как и ожидалось, было сосредоточено в активном центре PylRS. Особый интерес представляет аминокислотная замена Met300Thr, которая возникла независимо во всех изученных ферментах и ранее не описана в литературе (рисунок 3.10).

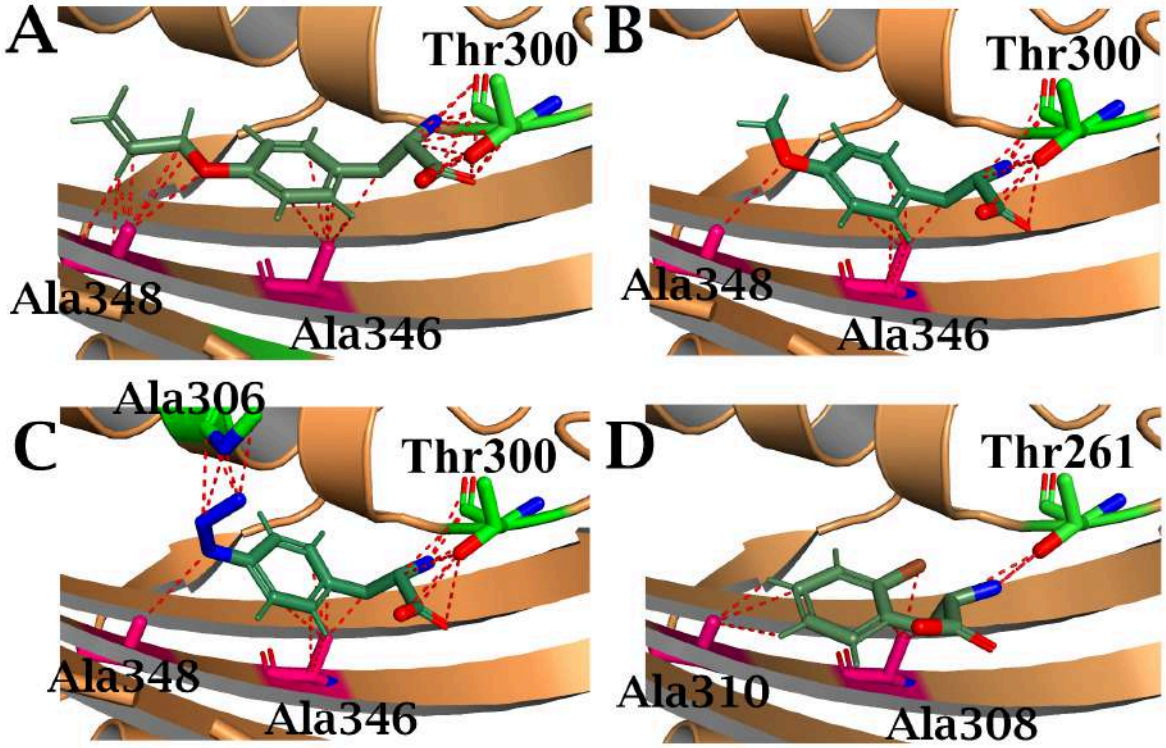



Рисунок 3.10 – Структура аминокислотных связывающих карманов **(A)** AllYPylRS, **(B)** MethYPylRS, **(C)** AzFPylRS и **(D)** 2ClFPylRS в комплексе с соответствующей неприродной аминокислотой. Аминокислотные замены обозначены зеленым цветом. Красные пунктирные линии представляют потенциальные водородные связи. Структуры основаны на записи PDB 2ZCE

Известно, что остатки Met300 и Ala302 расположены вблизи α–аминогруппы Pyl в каталитическом ядре PylRS, но их боковые цепи слишком удалены от α–аминогруппы Pyl, чтобы образовывать стабильные ван–дер–ваальсовы взаимодействия. Поэтому PylRS

распознает α–аминогруппу субстрата косвенно – через сеть водородных связей [70,108]. Также известно, что аминокислотная замена Ala302Thr приводит к появлению гидроксильной группы, образующей выступ вблизи β–слоев активного центра и определяющей основание сайта связывания аминокислоты. В частности, остаток Thr302 может непосредственно образовывать водородную связь с α–карбоксильной группой субстрата без участия молекулы воды [117]. Для оценки функционального вклада аминокислотной замены, конвергентно возникшей во всех исследованных эволюционных линиях, была проведена замена Met300Thr в исходном, не подвергавшемся эволюции, варианте MmPylRS_mut (рисунок 3.11).

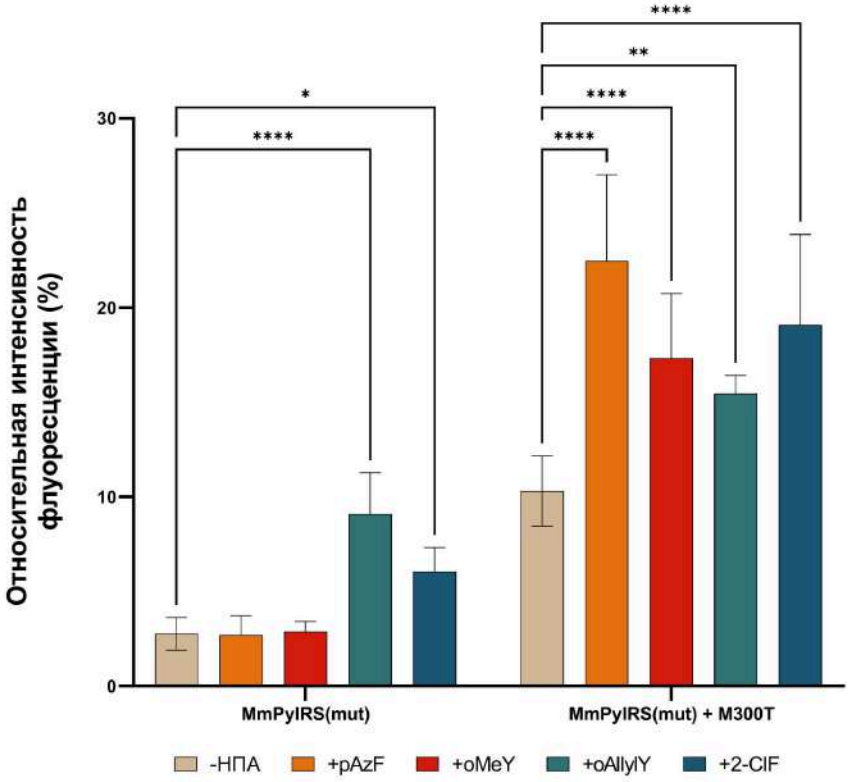



Рисунок 3.11 – Относительная интенсивность флуоресценции с и без добавления O–аллил–L–тирозина, O–метил–L–тирозина, 4–азидо–L–фенилаланина и 2–хлорфенилаланина для MmPylRS_mut и MmPylRS_mut с внесенной аминокислотной заменой Met300Thr. Данные представлены как среднее значение ± SD (n=9). *p<0,05, **p<0,01, ****p<0,0001. Для обработки использованы односторонний дисперсионный анализ, тест множественного сравнения по Шидаку

Относительная интенсивность флуоресценции, отражающая активность синтетазы, не проходившей эволюцию, в отношении исследованных НПА, составляла приблизительно 2–3% для pAzF и oMetY и возрастала до 9 и 6% для oAllylY и 2–ClF соответственно. Введение исследуемой аминокислотной замены приводило к заметному увеличению активности по отношению ко всем исследованным НПА. При этом максимальная активность варианта наблюдалась для pAzF (~22%), однако для остальных субстратов также сохранялись достаточно высокие значения (15–19%).

Данные функционального тестирования напрямую подтверждают выводы, основанные на анализе динамики аминокислотной замены и 3D–моделирования. Met300Thr не повышает сродство к одной конкретной НПА, а проявляется в виде общего существенного повышения активности синтетазы в отношении производных фенилаланина и тирозина. Следовательно, введение данной аминокислотной замены приводит к увеличению сродства PylRS к широкому спектру НПА, исследованных в данной работе, тогда как остальные замены, вероятно, в большей степени определяют субстратную специфичность фермента. Полученные результаты представляют особый интерес в контексте конвергентного возникновения данной замены в независимых эволюционных линиях и позволяют рассматривать ее как потенциально значимый функциональный элемент адаптации синтетазы, а также использовать как отправную точку для дальнейшей эволюции к НПА аналогам тирозина и фенилаланина.

Кроме того, выявленные в PylSc аминокислотные замены также представляют значительный интерес: Ala107Thr, Asn280Asp и Asn420Ser в AllYPylRS, Asp408Gly в MethYPylRS и Tyr306His и Glu357Lys в AzFPylRS. По-видимому, эти замены стабилизируют третичную структуру фермента и увеличивают карман связывания аминокислоты. Такое структурное изменение оптимизирует фермент для аминоацилирования тРНК крупными нециклическими аминокислотами, в частности фенилаланином и производными тирозина с громоздкими заместителями. Примечательно, что среди всех только один гомологичный остаток, Tyr306, ранее был описан в литературе. Было показано, что замена Tyr306Ala значительно увеличивает размер кармана связывания PylRS, тем самым повышая эффективность включения нециклических аминокислот [119,178]. На основании литературных данных и 3D–моделирования можно предположить, что Tyr306His также расширяет активный центр фермента и обеспечивает слабые водородные взаимодействия с азидогруппой pAzF.

Возникшие множественные замены в отдельных вариантах PylRS могут способствовать наблюдаемому повышению активности и улучшению распознавания НПА. Хотя имеющиеся данные не позволяют количественно разделить индивидуальные и совокупные эффекты,

некоторые из описанных в рамках диссертации аминокислотных замен неоднократно обнаруживались в литературе. Такие комбинации могут отражать компенсаторные взаимодействия, сохраняющие стабильность или активность фермента, но при этом изменяющие его способность распознавать субстрат или тРНК.

С другой стороны, некоторые аминокислотные замены по отдельности могут не оказывать существенного влияния, но становятся полезными на генетическом фоне, сформировавшемся в ходе эксперимента PANCE. Это особенно актуально для замен, возникающих одновременно в N–концевом домене, С–концевом домене и кармане связывания аминокислоты, где изменения в одном участке могут влиять на функциональные последствия в другом. Таким образом, эволюционные траектории, наблюдаемые в ходе данного исследования, могут отражать не только накопление независимых полезных изменений, но и формирование коадаптированных сетей.

## 3.2 Разработка платформы сайт–специфической функционализации липидных наночастиц на основе белков с НПА

### 3.2.1 Функционализация липидных наночастиц при помощи DARPin

#### 3.2.1.1 Получение DARPin mCD8

Цитотоксические CD8+ Т–лимфоциты играют ключевую роль в адаптивном противоопухолевом и противовирусном иммунном ответе, непосредственно лизируя трансформированные клетки за счет секреции перфорина и гранзимов и экспрессии лиганда Fas. Однако в микроокружении опухоли эти клетки претерпевают прогрессирующую дисфункцию и метаболическое истощение, что ограничивает эффективность стандартных методов лечения [179]. Целенаправленная доставка терапевтических нуклеиновых кислот с использованием липидных наночастиц может преодолеть этот барьер, защищая молекулы от деградации нуклеазами и обеспечивая локальное перепрограммирование цитотоксических Т–клеток *in vivo* при минимальной системной токсичности [180,181]. Модификация ЛНЧ для точного нацеливания на CD8+ Т–клетки или антигенпрезентирующие клетки индуцирует

антигенспецифическую пролиферацию и восстанавливает цитотоксический потенциал лимфоцитов, что критически важно для преодоления резистентности опухоли [180]. Кроме того, целенаправленная доставка терапевтической мРНК путем конъюгации ЛНЧ с анти–CD8–лигандами устраняет дорогостоящую стадию CAR–T–терапии *ex vivo* и минимизирует системную гепатотоксичность [182].

Традиционно для адресной доставки используются моноклональные антитела, но их высокая молекулярная масса, сложная пространственная структура и высокая стоимость производства ограничивают их широкое применение [183]. Каркасные белки, включая искусственные белки, миметики антител и сконструированные белки-повторы анкирина (Designed Ankyrin Repeat Proteins, DARPins), рассматриваются как эффективная альтернатива антителам [161]. DARPin – класс искусственно сконструированных небольших (13–16 кДа) однодоменных белков, имитирующих антитела и основанных на природных анкириновых повторах [184]. Благодаря отсутствию дисульфидных связей, высокой термодинамической стабильности и возможности экспрессии с высоким выходом в бактериальных системах эти белки являются перспективной альтернативой в качестве лигандов для адресной доставки ЛНЧ, избирательного воздействия на опухолевые антигены и создания диагностических платформ [161,185].

Обладая высокой аффинностью, низкой молекулярной массой и высокой стабильностью, белки DARPin могут специфически связываться с поверхностными антигенами иммунных клеток, включая мышиный маркер CD8 (mCD8) Т-лимфоцитов мыши. В данной диссертационной работе был выбран DARPin на основе клона MSE10 [166]. Для прямого сравнения функционализации поверхности частиц с использованием классического тиол-малеимидного и безмедного азид-алкинового циклоприсоединения в белковую последовательность был добавлен гибкий линкер, содержащий либо цистеин (DARPin-Cys), либо амбер стоп-кодон для последующего введения pAzF (DARPin-pAzF). Данная конструкция также содержала гис-таг на N-конце, для удобства очистки (рисунок 3.12).

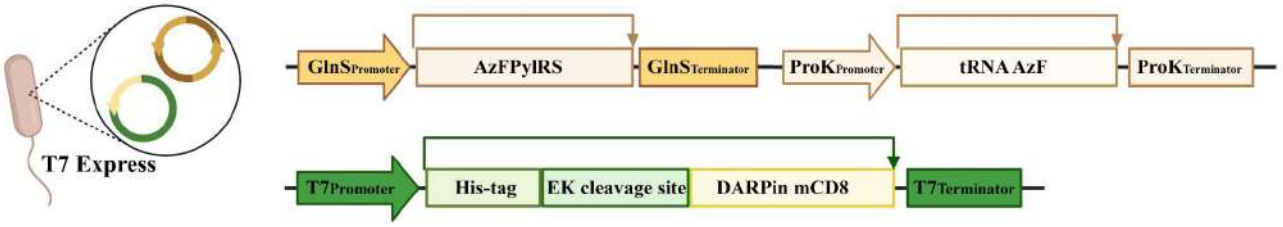



Рисунок 3.12 – Схематическое изображение штамма–продуцента DARPin mCB8, содержащего две плазмиды, и схема их открытых рамок считывания

Наработка DARPin с pAzF проводилась в штамме *E. coli* T7 express, ко–трансформированном плазмидами, содержащими ген DARPin, и плазмидой, содержащей AzFPylRS и соответствующую ей тРНК (рисунок 3.12). В среду для культивирования добавлялись 2 мМ НПА. Аналогичным способом, но без добавления дополнительных плазмидных конструкций и НПА в среду, проводилась наработка белка DARPin–Cys.

Молекулярная масса DARPin–pAzF и DARPin–Cys, рассчитанная на основе полученной аминокислотной последовательности, составила 19,1 кДа. Определнный методом SDS–PAGE размер слитого белка хорошо соответствовал рассчитанной молекулярной массе (рисунок 3.13).

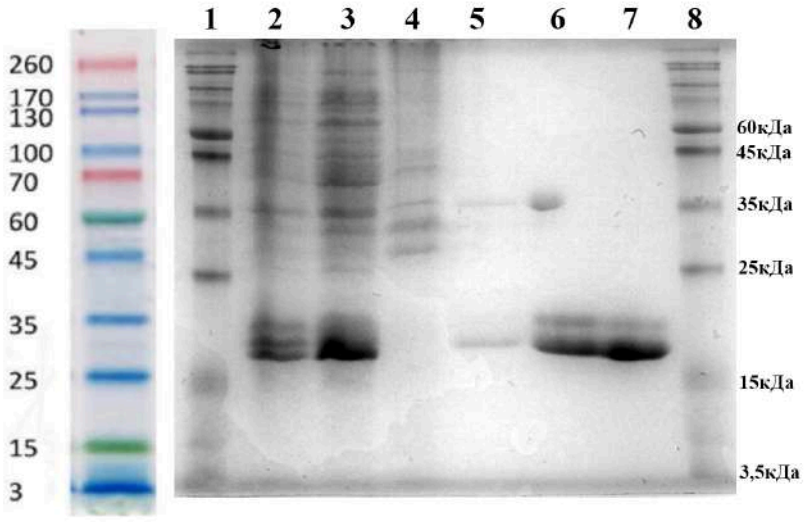



Рисунок 3.13 – Электрофореграмма 16% глициновый ПААГ, ожидаемая масса DARPin 19,1кДа: 1 и 8 – Маркер RAV11 (Biolabmix, Россия); 2 – Клетки после индукции; 3 – Растворимая фракция белка; 4 – Нерастворимая фракция белка; 5 и 6 – образец DARPin–Cys после очистки и диализа, видно образование димера двух DARPin при помощи Cys–Cys мостика (38,2кДа); 7 – Очищенный образец DARPin–pAzF

Белки были очищены из лизатов трансформированных *E. coli* методом аффинной хроматографии с иммобилизованными металлами. Высокая чистота белка была подтверждена анализом SDS–PAGE (рисунок 3.13). Выход белков составлял около 15–20 мг/л.

Для оптимизации протоколов биоортогональной конъюгации DARPin–pAzF был выбран краситель Cy5–DBCO. Хорошо известно, что SPAAC с производными дибензоциклооктина идеально подходит для биоконъюгации, поскольку не требует дополнительных реагентов и характеризуется высокой константой скорости [186,187]. Использование флуоресцентного красителя в качестве модельного субстрата для реакции SPAAC позволило оптимизировать условия сайт–специфической конъюгации и подтвердить успешное включение неприродной аминокислоты методом гель–электрофореза перед переходом к более сложным функциональным конъюгатам (рисунок 3.14).

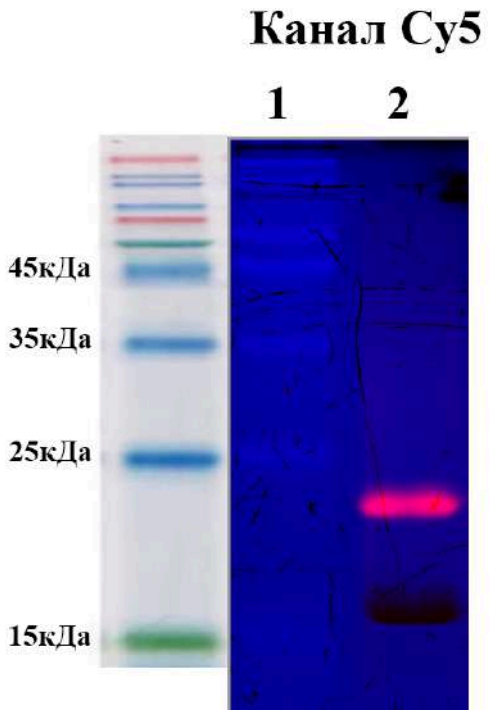



Рисунок 3.14 – Электрофореграмма 16% глициновый ПААГ образец DARPin–pAzF (19,1кДа) совместно с примесным белком без pAzF в составе конъюгация с DBCO–Cy5 флуоресценция в геле. Возбуждение DBCO–Cy5 – 630 нм, излучение – 650 нм

Реакцию конъюгации проводили путем добавления Cy5–DBCO в 25 мМ Tris–буфер при комнатной температуре на протяжении ночи. При возбуждении на длине волны 630 нм DARPin–Cy5 проявлял характерную флуоресценцию при длине волны 650 нм. Изображения, полученные в результате анализа, показали, что Cy5–DBCO вступает в реакцию конъюгации с DARPin.

Конъюгат DARPin–pAzF с флуоресцентной меткой Cy5 был успешно применен для окрашивания мышиных Т–клеток. В образце, где клетки одновременно окрашивали коммерческими антителами к CD3 и CD4 и анти–CD8 DARPin, популяция CD8+ Т–клеток была интенсивно окрашена в канале Cy5 и четко отделена от популяции CD4+ клеток (рисунок 3.15).

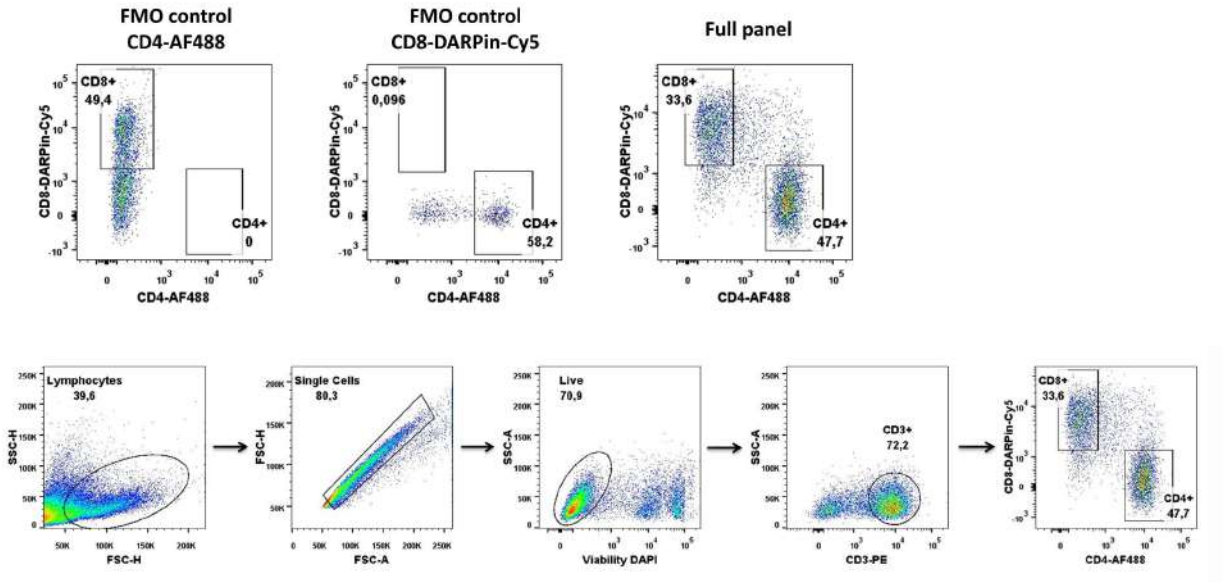



Рисунок 3.15 – Стратегия гейтирования при проточной цитометрии мышиных Т–клеток. После гейтирования по параметрам светорассеяния и исключения дуплетов (FSC–A/FSC–H) жизнеспособные ($DAPI^-$) CD3+–Т–клетки анализировали на предмет экспрессии CD4 (с использованием коммерческих антител) и CD8. Для детекции CD8 применяли DARPin, конъюгированный с флуорофором Cy5; сайт–специфическое присоединение осуществляли методом SPAAC с использованием стратегии конъюгации pAzF–DBCO

Этот анализ подтверждает не только конъюгацию DBCO–Cy5 с белком, но и то, что DBCO, присоединенный к С–концу белка посредством pAzF, не препятствует эффективному

связыванию DARPin со своим эпитопом рецептора CD8, что открывает возможности для функционализации поверхности ЛНЧ при сохранении аффинных свойств белка DARPin.

### 3.2.1.2 Получение и характеристика липидных наночастиц, декорированных DARPin

После отработки методики конъюгации с флуоресцентной меткой была выполнена функционализация белка липидом DSPE-PEG(2000)-DBCO с последующим включением полученного конъюгата (IPAD-1) в состав ЛНЧ. Реакции SPAAC протекают в мягких физиологических условиях при нейтральном pH и комнатной температуре, что позволяет сохранить нативную структуру белка.

Для инкапсуляции мРНК eGFP и доставки её к CD8+ Т–лимфоцитам использовалось микрожидкостное смешивание липидной фазы с водной фазой, содержащей мРНК. ЛНЧ включали два ионизируемых липида (SM–102 и ALC–0315 в эквимолярном соотношении), холестерин, DSPC и DMG–PEG–2000 в молярном соотношении 23,15:23,15:42,7:9,4:1,6 соответственно (рисунок 3.16 А). Такая композиция ранее успешно применялась для доставки мРНК и была названа LNP–5 [188].

Все препараты LNP–5, включая препараты, предназначенные для CD8+ клеток, должны соответствовать общепринятым критериям контроля качества материалов, содержащих ЛНЧ–мРНК: среднее значение Z (80–120 нм), PDI (0,05–0,20) и эффективность инкапсулирования (EE, >85%). Получение ЛНЧ с использованием DSPE–PEG–2000–DBCO в качестве сокомпонента при исходном смешивании липидов в различных молярных соотношениях (0,27–1,60 мол.%) позволило получить частицы с высоким значением PDI (>0,2) и средним гидродинамическим диаметром (Z) >200 нм после диализа против PBS (pH 7,5). Ситуация значительно улучшилась после замены буфера для удаления непрореагировавших компонентов PBS на Tris–HCl с аналогичным значением pH. Однако последующее прямое декорирование путем добавления к частицам с DBCO DARPin–pAzF приводило к полной потере инкапсулированного содержимого на этапе очистки методом тангенциальной фильтрации.

Вместо прямой декорации была выбрана функционализация ЛНЧ методом постсинтетической модификации (поствставки) с использованием конъюгата IPAD–1 (рисунок

3.31 C). Согласно последним литературным данным, поствставка липида в уже сформированные ЛНЧ считается более простым методом функционализации, поскольку не требует добавления дополнительных липидов на этапе формирования частиц [164,189].

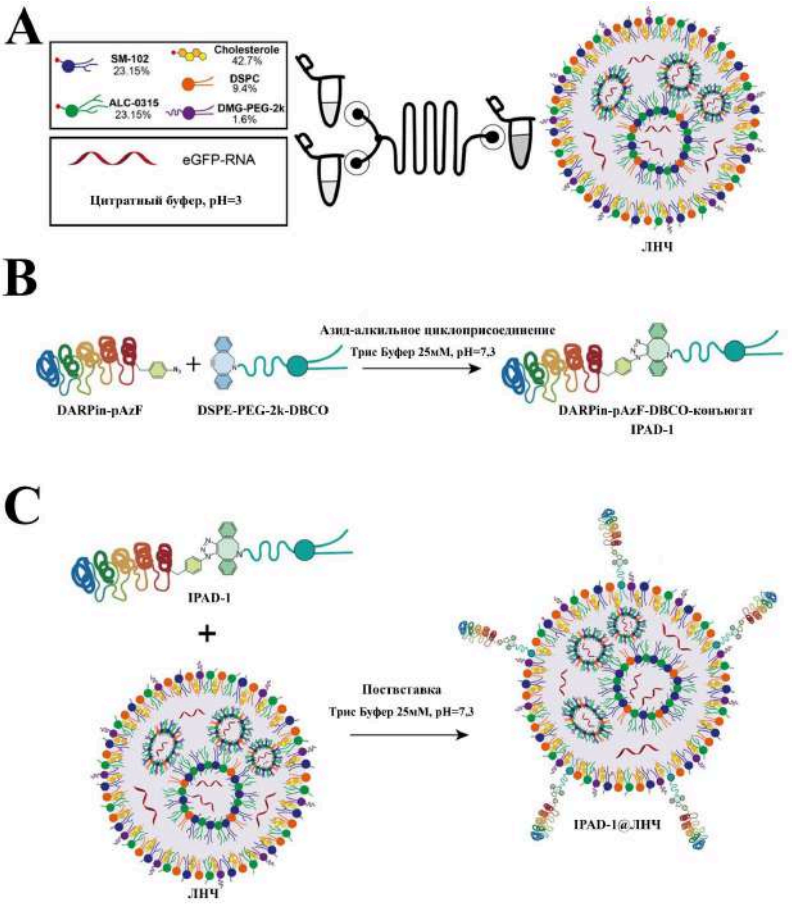



Рисунок 3.16 – Разработка состава и модификация поверхности LNP–5 методом поствставки с использованием конъюгата DARPin–pAzF и липида DSPE–PEG–2000–DBCO (IPAD–1)

Молярное соотношение IPAD–1 относительно общего количества липидов в LNP–5 было выбрано в диапазоне 0,10–0,27 мол.% по аналогии с данными O. Heidenreich и соавторов [190], полученными для поствставки DSPE–PEG–2000–DBCO, конъюгированного с пептидом azido–LDV или красителем azido–Cy3. Для сравнения физико–химических свойств были приготовлены образцы с содержанием IPAD–1 0,10 и 0,27 мол.% соответственно, выделенные методом тангенциальной фильтрации в двух буферах (рисунок 3.17).

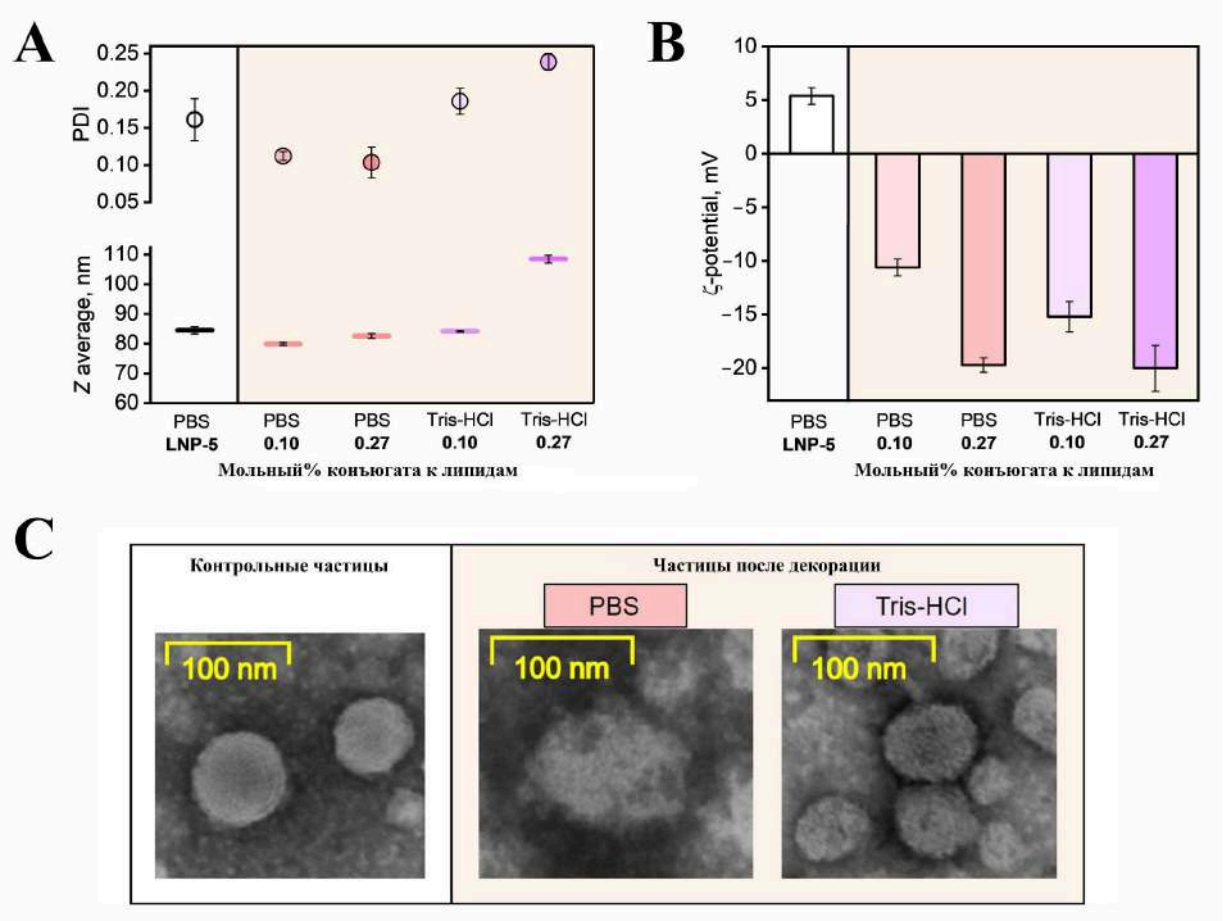



Рисунок 3.17 – Сравнение ζ–потенциала **(A)**, индекса полидисперсности **(B)** и среднего гидродинамического диаметра (Z-среднее значение, B) в зависимости от природы буфера для LNP–5 до и после модификации IPAD-1 (0,10 и 0,27 мол.%). **(C)** – микрофотографии исходных LNP–5 и после модификации, диализированных против PBS и Tris-HCl

Почти во всех случаях поствставка не приводила к увеличению среднего гидродинамического диаметра, за исключением частиц, диализированных против буфера PBS и содержащих 0,27 мол.% IPAD–1. В этом случае увеличение концентрации белково–липидного конъюгата в ЛНЧ привело к росту среднего гидродинамического диаметра до 108 нм и PDI до 0,24 (рисунок 3.17B), что указывает на нарушение сферической структуры, приводящее к агрегации наночастиц и высвобождению инкапсулированной мРНК.

СТЭМ с негативным контрастированием уранилацетатом позволяла визуализировать структуру и морфологию наноносителей. Контрольные наночастицы ЛНЧ имели правильную сферическую форму с четко очерченными границами, как ранее наблюдалось для аналогичных систем на основе LNP–5 [188]. Однако при модификации наночастиц и их диализе с

использованием PBS их морфология изменялась: преимущественно наблюдались бесформенные фрагменты, лишенные исходной сферической формы (рисунок 3.17C). Такое структурное нарушение в PBS может указывать на то, что электростатические взаимодействия с IPAD–1 в этой среде дестабилизируют липидный бислой или индуцируют фазовый переход, приводящий к распаду ЛНЧ. Наночастицы, полученные в буфере Tris–HCl, напротив, сохраняли более выраженную округлую структуру, их поверхность казалась шероховатой, рыхлой и покрытой плотным молекулярным слоем (рисунок 3.17C). Об успешном внедрении конъюгата IPAD–1 свидетельствовали изменение ζ–потенциала и СТЭМ.

В таблице 3.3 приведены все полученные физико–химические характеристики наноносителей.

Таблица 3.3 – Физико–химическая характеристика исследованных LNP–5 и LNP–5@IPAD–1

| ЛНЧ | Z средний диаметр, nm | PDI | Частицы/мл | ζ–потенциал, mV | EE, % |
|---|---|---|---|---|---|
| LNP–5 (без DARPin, PBS) | 84.5 ± 1.24 | 0.161 | $4.49 \cdot 10^{12}$ | +5.4 ± 0.77 | 100 |
| LNP–5 + IPAD–1 (0.10 мол%, PBS) | 79.9 ± 0.45 | 0.112 | $3.02 \cdot 10^{12}$ | –10.6 ± 0.77 | 91 |
| LNP–5 + IPAD–1 (0.10 мол%, Tris) | 82.6 ± 0.76 | 0.103 | $3.46 \cdot 10^{12}$ | –19.7 ± 0.66 | 90 |
| LNP–5 + IPAD–1 (0.27 мол%, PBS) | 84.2 ± 0.32 | 0.186 | $2.25 \cdot 10^{12}$ | –15.2 ± 1.43 | 97 |
| LNP–5 + IPAD–1 (0.27 мол%, Tris) | 108.5 ± 1.32 | 0.239 | $3.42 \cdot 10^{12}$ | –20.0 ± 2.14 | 91 |

Все ЛНЧ обладали высокой степенью инкапсуляции до и после введения (EE 90–100%). Примечательно, что, несмотря на высокий EE (97%), наблюдаемый при содержании 0,27 мол.% IPAD–1 и использовании PBS в качестве буфера для модификации и диализа декорированных частиц, общая концентрация частиц на единицу объема, измеренная методом MADLS, снизилась вдвое по сравнению с исходным показателем LNP–5. Напротив, при использовании Tris–HCl количество частиц всех типов оставалось на уровне примерно $3,5 \times 10^{12}$ частиц/мл. Таким образом, было показано, что выбор буфера критически влияет на структурную целостность модифицированных ЛНЧ. Tris–HCl помогает сохранить целостность наночастиц, способствуя формированию развитого пористого поверхностного слоя.

Благодаря поствставке IPAD–1 и всестороннему физико–химическому изучению ЛНЧ были определены оптимальный буфер и соотношение IPAD–1 к PEG для получения липидных

наночастиц, подходящих для последующих цитофлуориметрических экспериментов. Хотя можно было использовать как 0,10, так и 0,27 мол.% IPAD–1, состав с 0,10 мол.% в буфере Tris–HCl продемонстрировал лучшие коллоидные свойства (средний Z ≈ 83 нм, PDI ≈ 0,10) и был выбран для функциональных исследований. Этот оптимизированный состав применяли во всех последующих исследованиях по трансфекции.

Дополнительно были подготовлены частицы с прямой декорацией через тиол–малиемидную конъюгацию с помощью DARPin–Cys. Физико–химические свойства контрольных ЛНЧ, конъюгированных через малеимид, были сопоставимы со свойствами частиц, декорированных IPAD–1, и обобщены в таблице 3.4.

Таблица 3.4 – Физико–химическая характеристика исследованных LNP–5, LNP–5–Mal и LNP–5–Mal с декором DARPin–Cys

| ЛНЧ | Z средний диаметр, nm | PDI | ζ–потенциал, mV |
|---|---|---|---|
| ЛНЧ–5 (0) PBS | 84.5 ± 1.24 | 0.16 | + 5.4 ± 0.77 |
| ЛНЧ–5–Mal (0.27) PBS | 85.9 ± 0.90 | 0.10 | – 4.6 ± 0.20 |
| ЛНЧ–5–Mal (0.27) + DARPin–Cys PBS | 90.8 ± 1.60 | 0.10 | – 8.7 ± 1.30 |

### 3.2.1.3 Функциональные тесты с ЛНЧ–DARPin

Специфичность таргетирования оценивалась *in vitro* путем селективной трансфекции мышиных CD8+ Т-лимфоцитов мРНК-ЛНЧ eGFP, содержащих анти-CD8 DARPin-pAzF (группа IPAD-1). Все функциональные эксперименты проводились с использованием оптимального состава 0,10 мол.% IPAD-1 в буфере Tris-HCl. Для сравнения использовались немодифицированные ЛНЧ, ЛНЧ с пассивно адсорбированным на поверхности DARPin-pAzF (группа DARPin) и ЛНЧ, декорированные контрольным анти-CD8 DARPin, конъюгированным через тиол–малиемидную связь с DMG-PEG-2000-Mal (группа DARPin-Cys-Mal).

Нетрансфицированные клетки служили отрицательным контролем (рисунок 3.18). Все эксперименты были проведены в трех повторах.

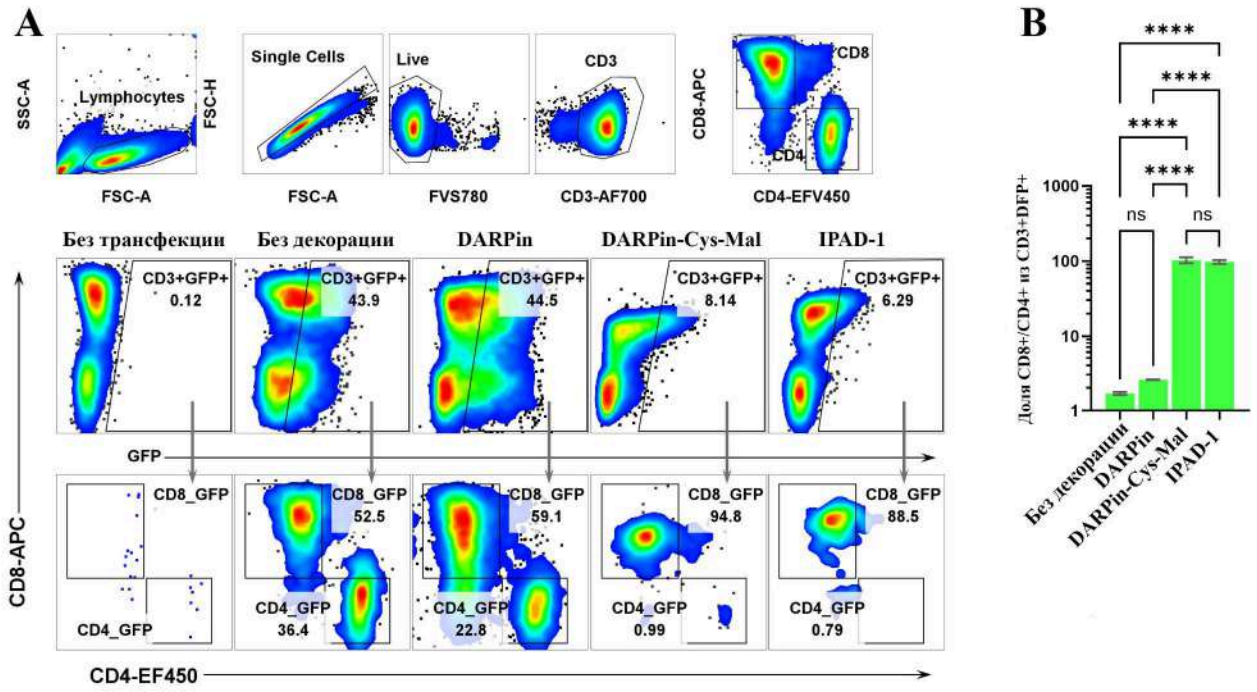



Рисунок 3.18 – Специфичность ЛНЧ, нацеленного на CD8. **(A)** Стратегия гейтирования и репрезентативные графики, демонстрирующие экспрессию GFP в CD3+ Т–клетках и распределение между CD8+ и CD4+ подгруппами после обработки. **(B)** Соотношение CD8/CD4 среди CD3+/GFP+ лимфоцитов (*n*=3). Данные представлены как среднее значение ± SD. *p < 0,02, **p<0,005, ns – незначимо. Для обработки использованы односторонний дисперсионный анализ, тест множественного сравнения по Шидаку

Анализ экспрессии GFP в CD4+ и CD8+ клетках показал, что немодифицированные ЛНЧ, а также ЛНЧ с пассивной абсорбцией неспецифически трансфицируют обе популяции лимфоцитов. Напротив, направленная конъюгация обоих вариантов анти-CD8 DARPin приводила к резкому увеличению доли GFP-позитивных событий в целевой популяции CD8+ лимфоцитов и одновременно снижала неспецифическую трансфекцию CD4+ лимфоцитов до фонового уровня. Блокирование рецептора CD8 на Т-клетках коммерческим антителом (клон 53-6.7) перед добавлением частиц, украшенных IPAD-1, уменьшало трансфекцию примерно в 20 раз, практически до нуля (рисунок 3.19).

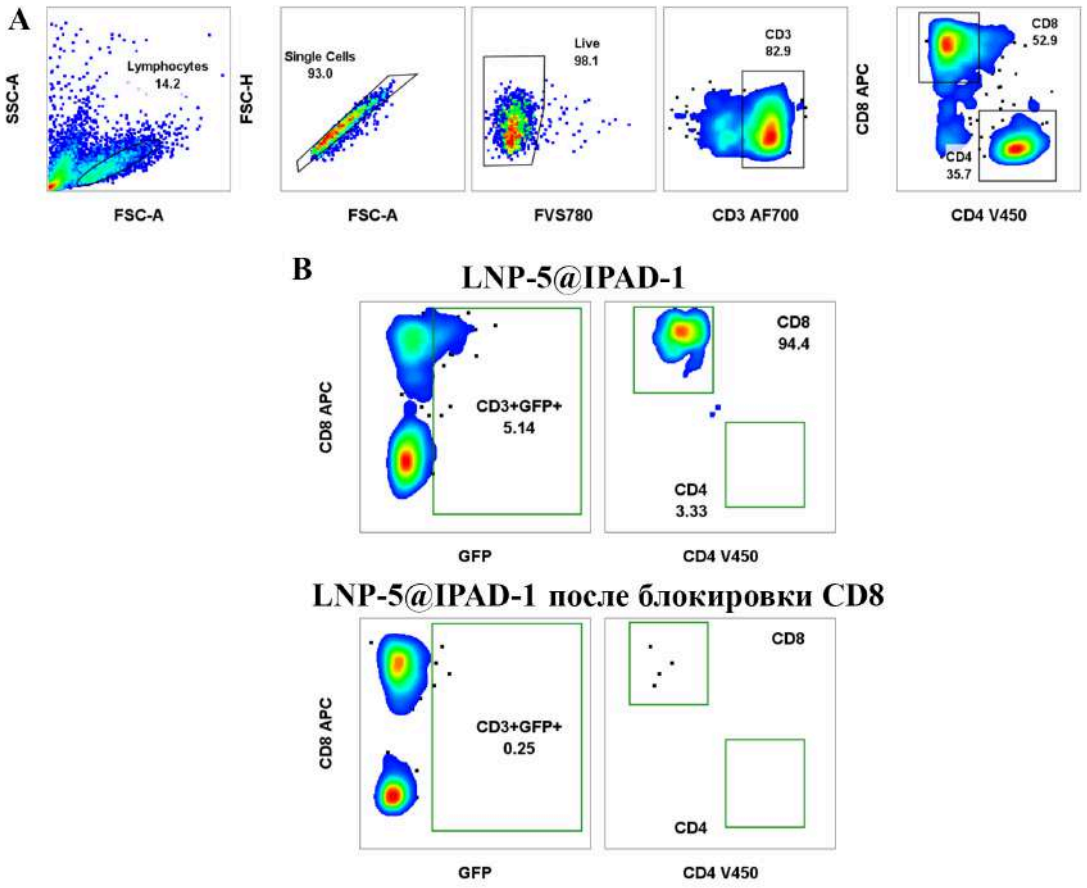



Рисунок 3.19 – Специфичность ЛНЧ, нацеленных на CD8. (А) Стратегия гейтирования и (Б) репрезентативные графики, показывающие экспрессию GFP в CD3–Т–клетках и распределение между CD8+ и CD4+ подгруппами после трансфекции (*n*=2)

Чтобы оценить специфичность воздействия, были рассчитаны соотношения CD8+ клеток к CD4+ клеткам среди GFP⁻позитивных Т⁻лимфоцитов. Как и ожидалось, для немодифицированных ЛНЧ это значение было минимальным (1,7 ± 0,07), что отражает отсутствие специфичности. Пассивная абсорбция частицами DARPin, содержащего pAzF, не приводила к специфической трансфекции CD8+ лимфоцитов, о чем свидетельствует распределение CD4+GFP+ и CD8+GFP+ клеток, которое оставалось сопоставимым с распределением в группе, трансфицированной немодифицированными частицами PEG. Это наблюдение подчеркивает, что для достижения специфичности функционализированных ЛНЧ требуется использование конъюгата белок–липид, а не только одного белка, поскольку пассивная адсорбция белка недостаточна для обеспечения целевой селективности. В то же время оба варианта направленной конъюгации ЛНЧ с рецептором CD8 приводили к

значительному увеличению показателя селективности до ~90 (рисунок 3.18В). Таким образом, эффективность и специфичность нацеливания ЛНЧ с использованием биоортогональной клик-химии сопоставимы с показателями, достигаемыми классическим методом тиол–малиемидной конъюгации.

Недавно Buchholz и соавторы разработали платформу для декорирования ЛНЧ слитым белком ApoE2–DARPin, используя тот же клон MSE10 DARPin, который применялся в данном исследовании [166]. После трансфекции активированных мышиных спленоцитов они получили аналогичный результат: было трансфицировано приблизительно 7,5% клеток CD3+, и все они были CD8 положительными. Эти результаты хорошо согласуются с данными, полученными в диссертации, как с точки зрения общей эффективности трансфекции, так и селективности для целевой популяции CD8+.

Однако при сравнении неокрашенных ЛНЧ обнаруживается расхождение. В исследовании Buchholz использование таких частиц приводило к незначительной трансфекции мышиных Т-клеток (<1%), тогда как в данном исследовании при тех же условиях было получено 43,9% eGFP-позитивных клеток среди общей популяции CD3+. Это заметное различие, вероятно, обусловлено различиями в липидном составе двух препаратов. Если ЛНЧ Buchholz содержали ионизируемый липид Dlin-MC3-DMA, то в состав LNP–5 входила эквимолярная комбинация SM-102 и ALC-0315. В результате очевидное преимущество частиц, декорированных при помощи DARPin, в данном исследовании сравнивается с относительно высокой базовой трансфекцией неокрашенного препарата. Эта гипотеза подтверждается работой Escalona-Rayo и соавторов, которые продемонстрировали на мышиных дендритных клетках, полученных из костного мозга (BMDC), что как поглощение клетками, так и экспрессия eGFP зависят от типа ионизируемого липида, с ранжированием эффективности SM-102 > ALC-0315 > Dlin-MC3-DMA [191].

Эти наблюдения подчеркивают, что результат трансфекции целевых систем ЛНЧ определяется сложным взаимодействием множества параметров. Выбор ионизируемого липида является критическим фактором, определяющим эффективность доставки, но его недостаточно для обеспечения специфичности к типу клеток. Не менее важен и метод декорирования: пассивная адсорбция целевых белков не обеспечивает селективности, тогда как ковалентная конъюгация с помощью четко определенных химических соединений обеспечивает стехиометрическое, ориентированное отображение лиганда на поверхности ЛНЧ. В данном исследовании в результате реакции SAAC между DBCO и остатком pAzF в DARPin успешно достигается стабильная, стехиометрическая, ориентированная конфигурация.

### 3.2.2 Функционализация липидных наночастиц при помощи Эксендин–4

#### 3.2.2.1 Создание генетических конструкций для наработки Эксендина–4

Эксендин–4, один из применяемых в терапии пептидов миметиков глюкагоноподобного пептида 1 (GLP–1), был обнаружен в секрете аризонского ядозуба (*Heloderma suspectum*) [192]. Он состоит из 39 аминокислот и, хотя имеет только 53% гомологии с человеческим GLP–1, является полным агонистом GLP–1R [193]. Важной особенностью Exe4 является наличие аланина вместо глицина во втором положении последовательности. Такая замена препятствует расщеплению пептида дипептидилпептидазой–4 (DPP–4), вследствие чего Exe4 характеризуется высокой метаболической стабильностью и более продолжительным временем циркуляции по сравнению с эндогенным GLP–1 [194]. Указанные свойства делают Exe4 перспективной основой для дальнейшей молекулярной модификации с целью получения аналогов с заданными фармакокинетическими и биологическими характеристиками.

Для включения pAzF в белковую последовательность Exe4 было выбрано 40–е положение, находящееся вне последовательности Эксендина–4, в природе состоящей из 39 аминокислот. Это положение было выбрано из–за минимального влияния на конформацию пептида. Анализ литературы показал, что замены в N-концевом участке пептида влияют на его способность связываться с рецептором, тогда как замены в С-концевой части не приводят к снижению аффинности [195]. В структуру нативного пептида в 40–е и 41–е положения были добавлены тирозин и лизин в качестве группы сравнения белка с pAzF и без него.

Выбор лизина в 41–е положение был продиктован особенностями трансляционного механизма белков с НПА. Размещение амбер-кодона, чья роль стоп-кодона была заменена на кодирование pAzF, непосредственно после последнего триплета Exe4 создает риск его конкурентного распознавания фактором терминации трансляции RF1, что приводит к преждевременному завершению синтеза. Добавление следующего за pAzF лизина в рамке считывания снижает вероятность такой конкуренции. В этом случае рибосома распознает сигналы к продолжению синтеза белка, что должно обеспечивать эффективное включение НПА.

В связи с небольшим размером Exe4 (~4,2 кДа) для его эффективной продукции в прокариотической системе была разработана конструкция, обеспечивающая экспрессию

пептида в составе фьюжн белка. Такой подход позволяет повысить стабильность целевого продукта и предотвратить его деградацию в клетке. На N–конец пептида была добавлена нуклеотидная последовательность, кодирующая тиоредоксин А (TRX–A), обладающий высокой растворимостью и способствующий сохранению целевого пептида в растворимой фракции при высокой интенсивности экспрессии [196]. Для последующей аффинной очистки в конструкцию также включили шестигистидиновую метку (His–tag), обеспечивающую последующее выделение методом металлоаффинной хроматографии. После выделения целевого белка из клеточного лизата и хроматографической очистки аффинный таг и TRX–A остаются ковалентно связанными с пептидом, что требует их последующего удаления. С этой целью между последовательностями Exe4 и TRX–A был клонирован сайт протеолиза энтеропептидазы, высокоспецифичной протеазы, расщепляющей субстрат без остатков аминокислот на N–концевой части целевого белка. Для отщепления фьюжн–тагов использовали энтеропептидазу собственного производства.

Для получения фьюжн–белка Эксендина–4 (TRX–A–Exe4) был создан экспрессионный вектор на основе pET30(a+). Полученная кассета находилась под контролем T7–промотора и T7–терминатора (рисунок 3.20).

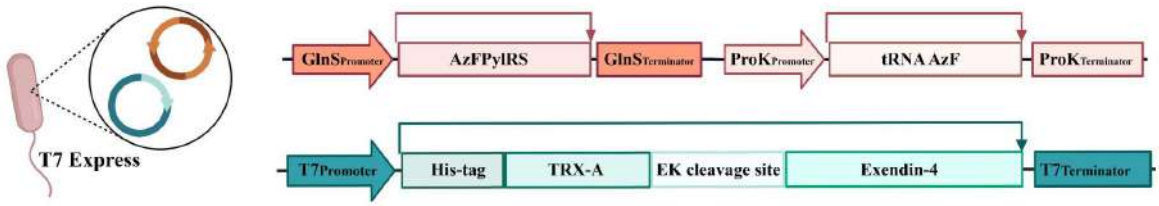



Рисунок 3.20 – Схематическое изображение штамма–продуцента Эксендина–4 с pAzF (Exe4–pAzF), содержащего две плазмиды, и схематическое изображение их открытых рамок считывания

Для последующего сайт–специфического включения pAzF в заданные положения Exe4 в его генетическую последовательность вводили амбер стоп–кодон (TAG). Для этого применяли метод сайт–направленного мутагенеза, позволяющий осуществлять направленную замену нуклеотидов в заранее определенном участке гена.

### 3.2.2.2 Оптимизация условий наработки Эксендина–4

Отсутствие в структуре белка TRX–A–Exe4 дисульфидных мостиков и необходимости посттрансляционной модификации существенно упрощает выбор бактериальной системы экспрессии и условий культивирования. Оптимизацию условий наработки проводили при помощи конструкции, содержащей тирозин в 40 положении. В каждый из перечисленных ниже штаммов трансформировали плазмиду pET30(a+)_His8_TRX–A_Enterokinase_Exe4.

На первом этапе была проведена валидация разработанной экспрессионной конструкции с целью определить, обеспечивает ли она эффективную трансляцию Exe–4 и его накопление в растворимой фракции. С учетом этих особенностей в качестве первого продуцирующего штамма использовали *E. coli BL21(DE3)*. Данный штамм содержит в геноме последовательность Т7 РНК–полимеразы, экспрессия которой регулируется промотором lacUV5, что позволяет индуцировать транскрипцию рекомбинантного гена добавлением IPTG. Дополнительным преимуществом *BL21(DE3)* является отсутствие протеаз Lon и OmpT, что снижает риск внутриклеточного протеолиза рекомбинантного продукта [197]. Экспрессию проводили при двух температурах после индукции экспрессии белка – 37°C и 18°C. Через 20ч после индукции анализировали образцы лизатов бактерий до индукции, а также растворимую и нерастворимую фракции клеточного лизата методом SDS–PAGE (рисунок 3.21).

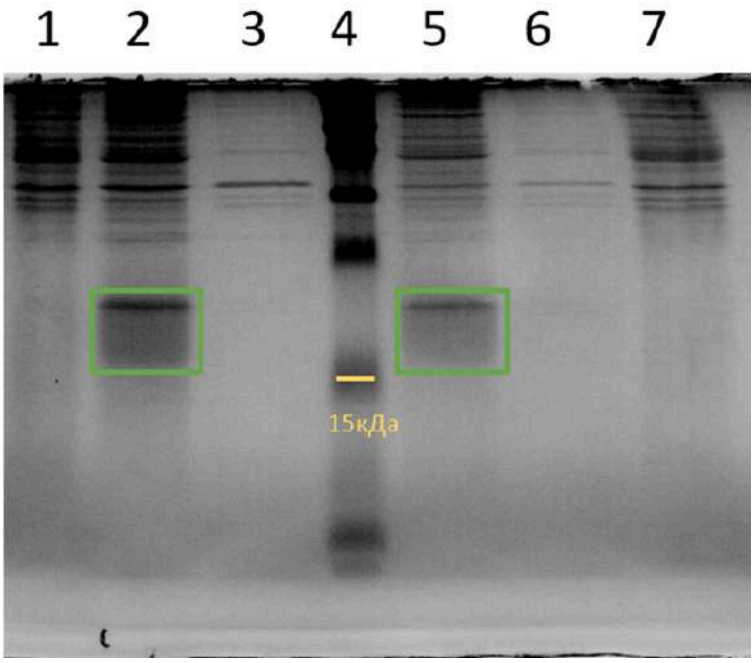



Рисунок 3.21 – Электрофореграмма 16% SDS–PAGE, ожидаемая масса TRX–A–Exe4 (выделена в зеленом квадрате) – 18.629 кДа: 1– Штамм *E. coli* BL21(DE3) до индукции ИПТГ; 2– Растворимая фракция при 37$^0$C; 3– Нерастворимая фракция при 37$^0$C; 4– Маркер RAV11 (Biolabmix, Россия); 5– Растворимая фракция при 18$^0$C; 6– Нерастворимая фракция при 18$^0$C; 7– Клетки до индукции

Электрофоретический анализ показал наличие целевого продукта в обоих температурных условиях, при этом основная его часть обнаруживалась в растворимой фракции. Выраженного накопления продукта в нерастворимой фракции не наблюдалось, что свидетельствует о корректной работе конструкции и способности полученного фьюжн белка сохранять растворимость при экспрессии в *E. coli*. Таким образом, уже на этапе первичной проверки конструкция продемонстрировала пригодность для дальнейшей наработки белка.

После подтверждения работоспособности разработанной экспрессионной конструкции и получения целевого продукта преимущественно в растворимой фракции следующим этапом стало определение наиболее подходящего штамма–продуцента. Для сравнительной оценки были выбраны штаммы *E. coli*, различающиеся по интенсивности и характеру T7–зависимой экспрессии.

*E. coli T7 Express* представляет собой производный от *BL21(DE3)* штамм, в котором ген T7–РНК–полимеразы встроен в хромосомный lac–оперон (в отличие от классического *BL21(DE3)*, где он расположен внутри профага λ). Это обеспечивает более высокую

стабильность клеток и снижает фоновую экспрессию целевого белка до индукции IPTG. Такая система обеспечивает эффективную экспрессию генов под контролем Т7–промотора и может использоваться для получения как растворимых, так и мембранных белков [198].

В отличие от Т7 Express, штаммы *E. coli* С41(DE3) и С43(DE3) были разработаны для более контролируемой экспрессии рекомбинантных белков. С41(DE3) получен на основе BL21(DE3) путем отбора вариантов, способных переносить экспрессию потенциально токсичных белков. Снижение интенсивности Т7–зависимой экспрессии уменьшает метаболическую нагрузку на клетки и может способствовать повышению выхода функционального продукта. Этот штамм часто используется для экспрессии мембранных, секреторных и токсичных белков, а также крупных мультидоменных ферментов. Штамм С43(DE3), являющийся производным С41(DE3), характеризуется еще более мягким режимом экспрессии и повышенной устойчивостью клеток к стрессу, связанному с продукцией рекомбинантного белка. Вследствие этого его применение особенно актуально для белков, склонных к неправильному сворачиванию и образованию нерастворимых агрегатов [199].

Эффективность экспрессии белка в растворимой и нерастворимой фракциях была оценена методом SDS–PAGE в 16% глициновом ПААГ (рисунок 3.22). Все штаммы культивировали в идентичных условиях при $37^0$С.

По результатам SDS–PAGE наиболее выраженное накопление целевого фьюжн–белка в растворимой форме наблюдалось при использовании штамма *E. coli* Т7 Express и культивировании при 37 °С. Выбранные условия применяли для дальнейшего получения TRX–A–Exe4 в препаративном количестве.

После наработки белка в увеличенном объеме проводили металлаффинную хроматографию на Ni–содержащем сорбенте. Элюцию осуществляли ступенчатым повышением концентрации имидазола от 10 до 60%. Основной пик целевого продукта наблюдался при концентрации буфера Б 60% (рисунок 3.23).

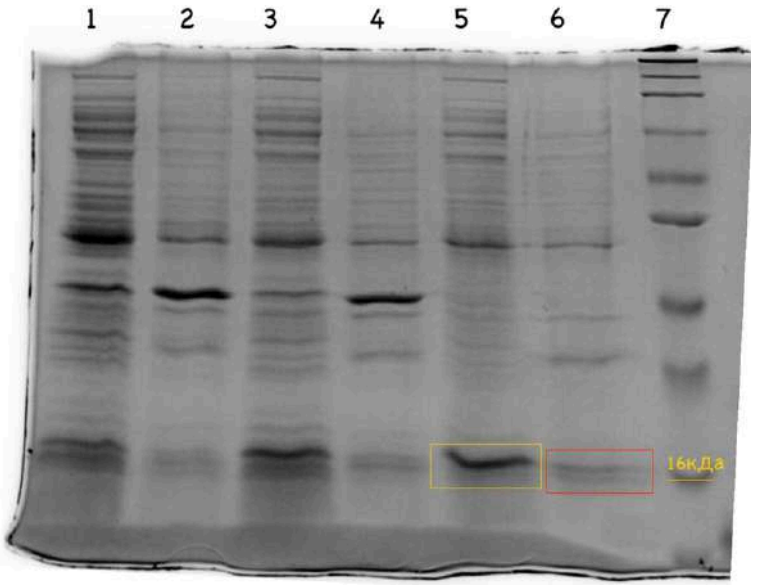



Рисунок 3.22 – Электрофореграмма электрофореза в 16% SDS–PAGE после индукции ИПТГ: 1 – Штамм C41(DE3); 2 – Штамм C41(DE3) растворимая фракция; 3 – Штамм C43(DE3) нерастворимая фракция; 4 – Штамм C43(DE3) растворимая фракция; 5 – Штамм *T7 express* нерастворимая фракция; 6 – Штамм *T7 express* растворимая фракция; 7 – Маркер RAV11

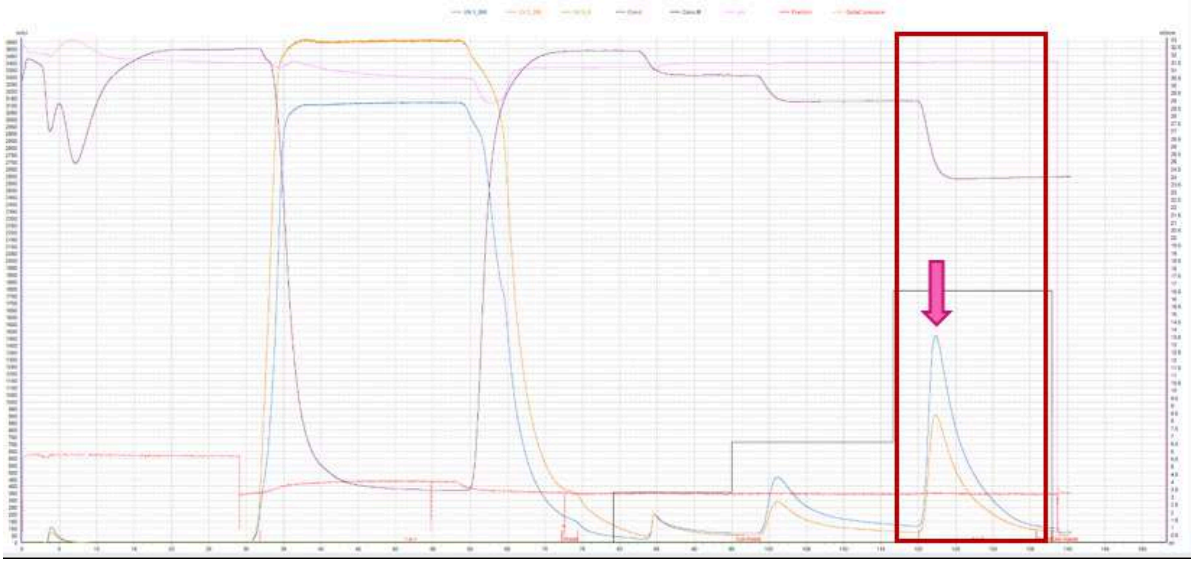


Рисунок 3.23 – Хроматограмма полученная после выделения белка TRX–A–Exe4 методом металл–аффинной ступенчатой хроматографии с использованием сорбента с иммобилизованным никелем (Ni). Стрелкой указан пик элюции целевого белка при концентрации буфера Б 60%

Полученный элюат дополнительно анализировали методом SDS–PAGE. Электрофорез в полиакриламидном геле образцов после хроматографической очистки представлен на рисунке 3.24. Целевой продукт с расчетной молекулярной массой 18,629 кДа преимущественно обнаруживался в растворимой фракции и во фракции элюата при 60% буфера Б, тогда как в нерастворимой фракции и проскоке его содержание было существенно ниже. Совокупность результатов электрофоретического и хроматографического анализа подтвердила работоспособность разработанной системы экспрессии и позволила определить условия, обеспечивающие получение и выделение растворимого фьюжн–белка TRX–A–Exe4.

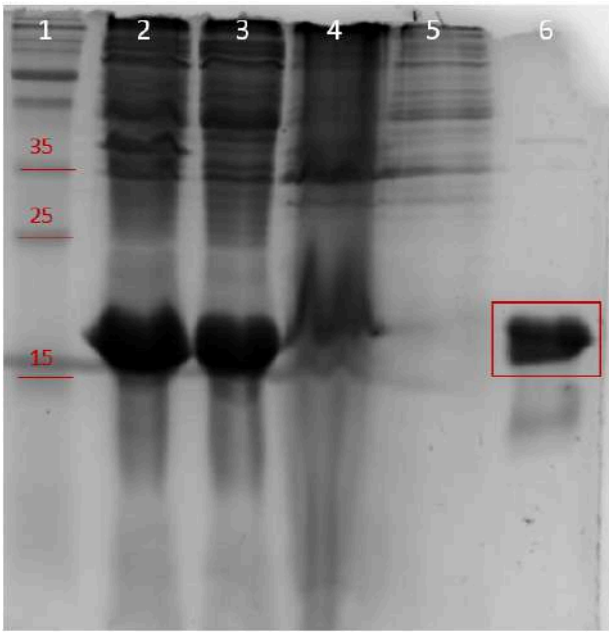



Рисунок 3.24 – Электрофореограмма 16% глицин ПААГ, ожидаемая масса белка 18.629 кДа: 1 – Маркер RAV11; 2 – Клетки после индукции ИПТГ; 3 – Растворимая фракция; 4 – Нерастворимая фракция; 5 – Слив (flow–through), или проскок; 6 – Элюат из фракции с концентрацией буфера Б 60%

Таким образом, удалось оптимизировать условия для наработки фьюжн–белка агониста GLP–1 рецептора Exe4, связанного с TRX–A. На следующем этапе выбранные условия адаптировали для получения вариантов TRX–A–Exe4–pAzF. Для этого использовался штамм *E.coli* T7 Express, одновременно несущий две плазмиды: pHPyl–AzPylRS–GlnS, кодирующую ортогональную систему для включения pAzF, и

pET30(a+)_His8_TRX–A_Enterokinase_Exe4_ncAA, содержащую ген фьюжн–белка TRX–A–Exe4–pAzF с амбер стоп–кодоном в 40 положении. Наработку варианта Exe4 проводили в присутствии соответствующей НПА, которую добавляли в культуральную среду до конечной концентрации 2 мМ.

### 3.2.2.3 Получение рекомбинантной энтеропептидазы для удаления фьюжн–тага

Энтеропептидаза представляет собой высокоспецифичную сериновую протеазу II типа, участвующую в пищеварении у млекопитающих. Зрелый фермент состоит из тяжелой и легкой цепей, соединенных дисульфидной связью; каталитическая активность локализована в легкой цепи массой около 35–62 кДа [200]. По механизму действия энтеропептидаза относится к протеазам химотрипсинового семейства и обладает выраженной субстратной специфичностью. Фермент распознает последовательность $(Asp)_4$–Lys↓X и гидролизует пептидную связь непосредственно после остатка Lys [201]. Важным преимуществом такого сайта является возможность удаления N–концевых фьюжн–тагов без добавления дополнительных аминокислот к целевой последовательности. Благодаря этому энтеропептидаза широко используется для удаления аффинных и солюбилизирующих тегов при получении рекомбинантных белков [202,203]. В настоящей работе данный фермент применяли для отделения TRX–A от Exe4 после его экспрессии и первичной очистки. В данном исследовании использовалась рекомбинантная легкая цепь человеческой энтеропептидазы (hEKL), поскольку она характеризуется примерно десятикратно большей каталитической эффективностью (kcat/KM) по сравнению с бычьим ферментом [204].

Вместе с тем получение активной hEKL в бактериальной системе представляет собой отдельную технологическую задачу. Аминокислотная последовательность энтеропептидазы содержит девять цистеинов, из которых формируются четыре дисульфидные связи, образование которых требует соответствующих окислительно–восстановительных условий и правильного пространственного сворачивания белка. Поэтому стандартная цитоплазматическая экспрессия в *E. coli* может приводить к образованию неправильных конформационных форм и снижению выхода функционального фермента. Для повышения вероятности получения активной энтеропептидазы была разработана комплексная стратегия оптимизации конструкции и системы экспрессии. Замена свободного цистеина на серин (C112S) в аминокислотной

последовательности hEKL помогла снизить количество потенциальных пар для образования дисульфидных связей, ограничив число альтернативных путей сворачивания и сместив равновесие в сторону активной конформации фермента. Дополнительным решением стало использование TRX–A в качестве N–концевого фьюжн–тага. Благодаря своим физико–химическим свойствам TRX–A повышает растворимость сложных эукариотических белков в прокариотических системах экспресии [196]. Также было показано, что он выступает донором тиольных и дисульфидных групп, что может способствовать образованию дисульфидных связей целевого белка [205]. Кроме того, TRX–A защищает целевой белок от протеолитической деградации и повышает его устойчивость к локальному перегреву, возникающему при ультразвуковой гомогенизации. Это связано с высокой термостабильностью самого тиоредоксина, который выдерживает нагревание до 80–90°C без необратимой денатурации и частично передает это свойство целевому белку [206].

Для получения экспрессионной конструкции ген hEKL собирали из синтетических олигонуклеотидов методом ПЦР. Полученный фрагмент клонировали методом Gibson Assembly в ранее разработанную плазмидную систему на основе pET30a(+). В отличие от конструкции для Exe–4 в данном случае His–tag расположен на C–концевом участке фермента. Экспрессия полученной кассеты находилась под контролем T7–промотора и T7–терминатора (рисунок 3.25).

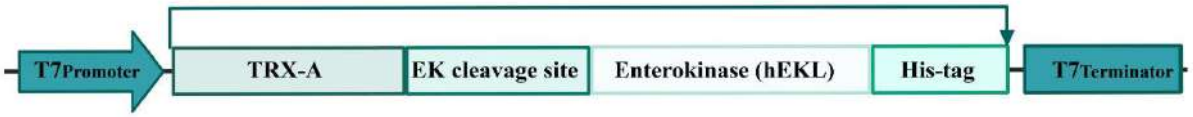



Рисунок 3.25 – Схематическое изображение открытой рамки считывания плазмиды, содержащей ген hEKL

Энтеропептидаза имеет четыре дисульфидные связи, образование которых требует кислую pH среды, поэтому в прокариотических системах наработка даже при оптимизированной генетической последовательности затруднена. В связи с этим в качестве штамма–продуцента был выбран SHuffle® T7 express *E. coli*, который за счет окислительной среды цитоплазмы обеспечивает формирование S–S связей [207]. Дополнительно была проведена трансформация вспомогательной плазмидой p.Gro7, адаптированной для

поддержания корректного сворачивания белков с дисульфидными связями и содержащей шапероновый комплекс GroEL/GroES [208].

При подборе оптимальных условий культивирования штамма–продуцента в качестве основы была использована работа Ebrahimifard et al [33]. По литературным данным, генетические последовательности белков–шаперонов p.Gro7 имеют прокариотическое происхождение и принадлежат грамотрицательным бактериям вида *E. coli*. В связи с этим их белки эволюционно адаптированы к функционированию в диапазоне умеренных температур (25–37°C) [209]. Исходя из этого, культивирование до индукции наработки целевого белка проводилось при 30 °C. Для определения оптимального температурного режима наработки целевого белка hEKL в растворимой фракции экспрессию белка индуцировали ИПТГ по ранее подобранному протоколу и культивировали при трех температурах: 25°C, 28°C и 30°C. Согласно литературным данным, корректно свернутый белок энтеропептидазы с правильно сформированными дисульфидными связями не визуализируется при электрофорезе в ПААГ [204]. В связи с этим количественную оценку выхода белка проводили по результатам аффинной хроматографической очистки (рисунок 3.26), функциональную – по эффективности протеолитического отщепления метки TRX–A от Exe4.

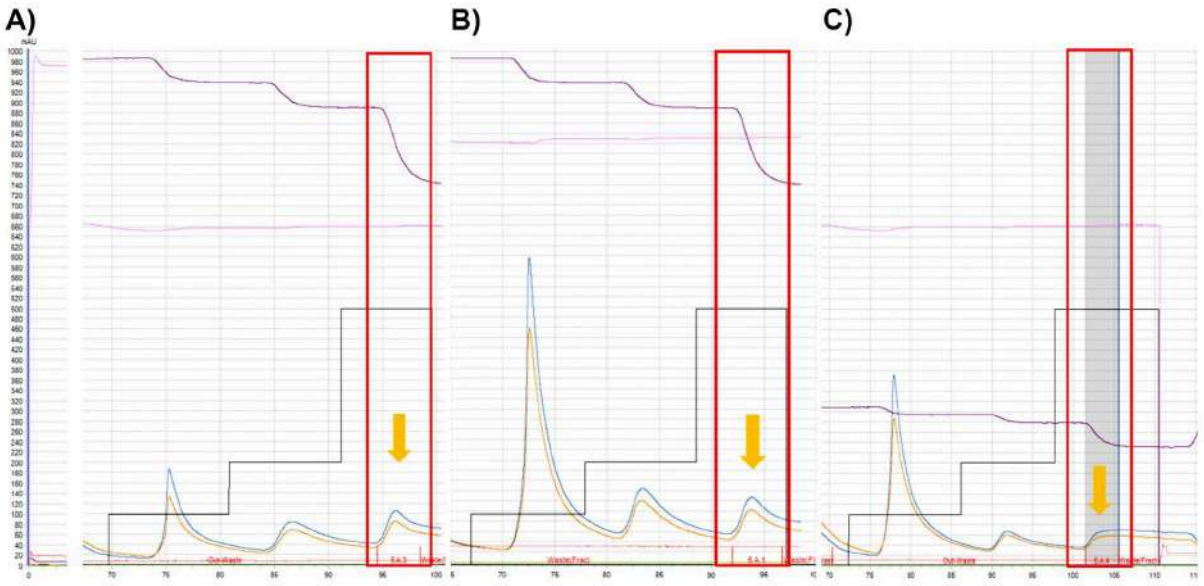



Рисунок 3.26 – Хроматограмма полученных образцов: (**A)** – hEKL, полученная при условиях наработки 25°C, концентрация буфера Б 50%; (**B)** – hEKL, полученная при условиях наработки 28°C, концентрация буфера Б 50%; (**C)** – hEKL, полученная при условиях наработки 30°C, концентрация буфера Б 50%

Для всех исследованных образцов hEKL были получены воспроизводимые хроматограммы с близким профилем элюции. Основной пик целевого белка наблюдался при концентрации элюирующего буфера Б 50% (400 мМ имидазола).

В результате был оптимизирован протокол, при котором оптимальными условиями получения фермента являются: рост ночной культуры с $OD_{600}$ = 0,1 при 30°C, индукция 0,7 мМ ИПТГ при достижении $OD_{600}$ = 1,2, инкубация 24–30 ч при 28°C. Очищенный фермент энтеропептидазы в дальнейшем использовали для протеолитического отщепления тиоредоксина А от слитого с ним белка Exe4.

### 3.2.2.4 Получение Exe4 с pAzF и его характеризация

При выборе условий для определения активности рекомбинантной hEKL учитывались литературные данные о зависимости активности фермента от состава реакционной среды. Показано, что энтеропептидаза сохраняет каталитическую активность в широком диапазоне pH (4.5–9.5), однако оптимальные условия для проведения стандартного ферментативного анализа соответствуют слабощелочной среде, в частности pH около 7.5–8.0. Существенное значение имеет наличие ионов $Ca^{2+}$, необходимых для проявления каталитической активности фермента. В связи с этим в качестве основы реакционного буфера целесообразно использовать Tris–HCl с добавлением $CaCl_2$. При этом применение фосфатных буферов и высоких концентраций NaCl нежелательно, поскольку они могут приводить к снижению активности энтеропептидазы. После аффинной хроматографической очистки TRX–A–Exe4 и hEKL проводили диализ против буфера 25мМ Tris, 2мМ CaCl2, pH 7,5.

Для проверки активности энтеропептидазы подбирались различные соотношения целевого белка и фермента: 1:1, 1:2, 1:4, 1:10, 1:100 (рисунок 3.27). Протеолиз проводился при комнатной температуре в течение ночи. Качественный анализ активности протеолиза энтеропептидазы осуществляли методом электрофореза в 16% трициновом ПААГ. По результатам анализа для полного протеолиза в дальнейшем было выбрано соотношение 1 мг/мл целевого белка к 1/10 объема (10мкг) раствора энтеропептидазы.

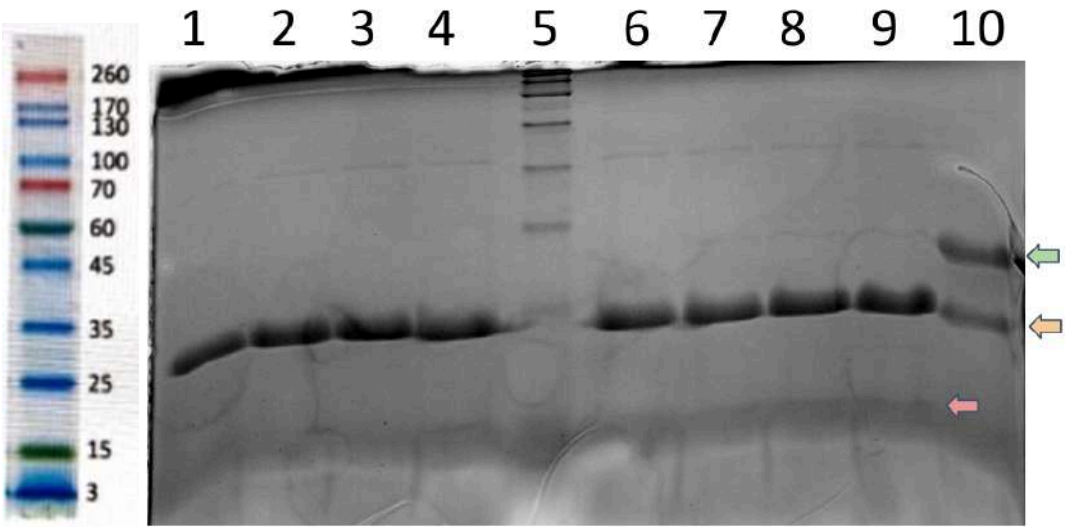



Рисунок 3.27 – Электрофорез в 16% трициновом ПААГ после растительной эксендина–4 к энтеропептидазе. Ожидаемая масса TRX–A–Exe4=18.629 кДа (зеленая стрелка); TRX–A=14,2 кДа (оранжевая стрелка); Exe4=4,4 кДа (розовая стрелка): 1 – hEKL полученная при 28$^0$C к TRX–A–Exe4 1:1; 2 – hEKL полученная при 28$^0$C к TRX–A–Exe4 1:2; 3 – hEKL полученная при 28$^0$C к TRX–A–Exe4 1:4; 4 – hEKL полученная при 28$^0$C к TRX–A–Exe4 1/100:1; 5 – Маркер RAV11; 6 – hEKL полученная при 25$^0$C к TRX–A–Exe4 1:1; 7 – hEKL полученная при 25$^0$C к TRX–A–Exe4 1:2; 8 – hEKL полученная при 25$^0$C к TRX–A–Exe4 1:5; 9 – hEKL полученная при 25$^0$C к TRX–A–Exe4 100:1

Очистку Exe4, содержащего и не содержащего pAzF, после протеолиза, проводилась обращенно–фазовая высокоэффективная жидкостная хроматография (ВЭЖХ). Полученные образцы после очистки были лиофилизированы и перед проведением последующих экспериментов растворены в 25мМ Tris–HCl.

Моделирование реакции SPAAC проводилось с использованием флуоресцентного красителя аналогично DARPin–pAzF. Анализ проводился с помощью SDS–PAGE и флуоресцентной визуализации геля (рисунок 3.28). При возбуждении на длине волны 630 нм Exe4–Cy5 проявлял характерное флуоресцентное излучение при 650 нм. Полученные изображения показали, что Cy5–DBCO эффективно конъюгировал с Exe4.

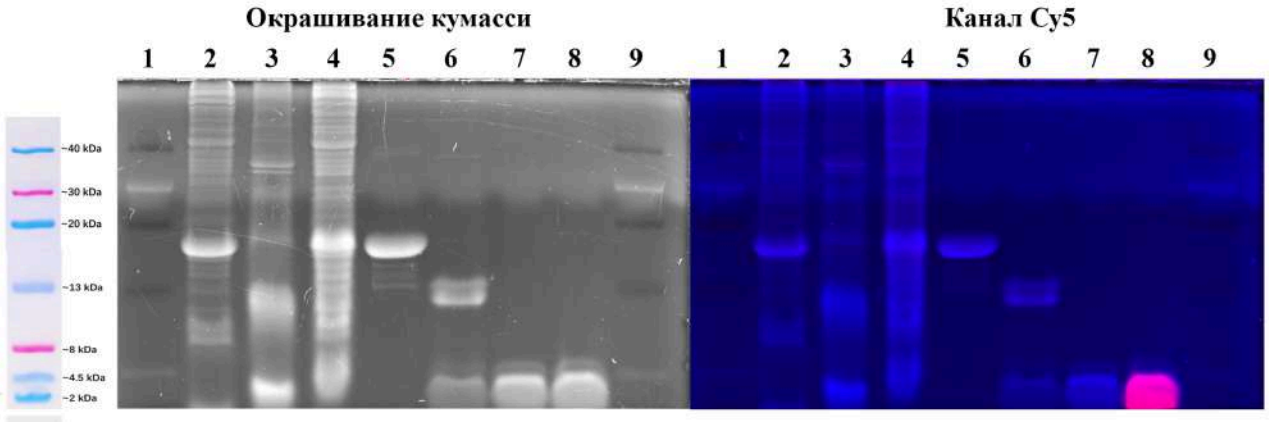



Рисунок 3.28 – Электрофореграмма 16% трицин SDS–PAGE Exe4 окрашенного кумасси и в канале Cy5: 1 и 9 – маркер белков IX; 2 – Клетки после индукции ИПТГ; 3 – Нерастворимая фракция; 4 – Растворимая фракция; 5 – TRX–A–Exe4 после очистки Ni–NTA; 6 – TRX–A–Exe4 после протеолиза; 7 – Exe4 очищенный, при помощи обращенно–фазной хроматографии; 8 – Exe4 очищенный, при помощи обращенно–фазной хроматографии, конъюгированный с DBCO–Cy5

Для подтверждения способности Exe4–pAzF специфически связываться со своим рецептором на поверхности клеток было проведено лиганд–рецепторное флуоресцентное окрашивание клеточной линии HEK293T, представленной в двух вариантах: (1) дикотипный (без экспрессии GLP–1R, далее HEK wt) и (2) с экзогенной экспрессией GLP1–R (коэкспрессия с eGFP, далее HEK GLP–1R). Окрашивание клеток проводили Exe4–Cy5.

Для подбора оптимальной концентрации конъюгата Exe4–Cy5 в цитофлуориметрическом анализе клетки HEK wt (рисунок 3.29A) и HEK GLP–1R (рисунок 3.29B) были окрашены серийными разведениями конъюгата 1:20, 1:200 и 1:2000 с исходной концентрацией 0.25 мг/мл. Разведение 1:20 приводило к окрашиванию 99.8% клеток HEK GLP–1R, однако наблюдалось неспецифическое взаимодействие около 74% клеток HEK wt. Это можно объяснить избытком конъюгата в разведении 1:20, что вызывало неспецифический сигнал от клеток без экзогенной экспрессии GLP–1R на поверхности. При разведении 1:200 99% клеток HEK GLP–1R окрашивались Exe4–Cy5, тогда как сигнал от клеток HEK wt не наблюдался, что свидетельствует о высокой специфичности конъюгата при окрашивании клеток в разведении 1:200. При более высоком разведении 1:2000 эффективность флуоресценции клеток HEK GLP–1R снижалась, и в гейт окрашенных событий попадало лишь 83% клеток. Таким образом, было установлено не только оптимальное разведение 1:200 Exe4–Cy5 для

цитофлуориметрического окрашивания клеток HEK, с итоговой концентрацией конъюгата в клеточной суспензии 1,25 мкг/мл, но и высокая специфичность Exe4–Cy5 в отношении рецептора GLP–1R.

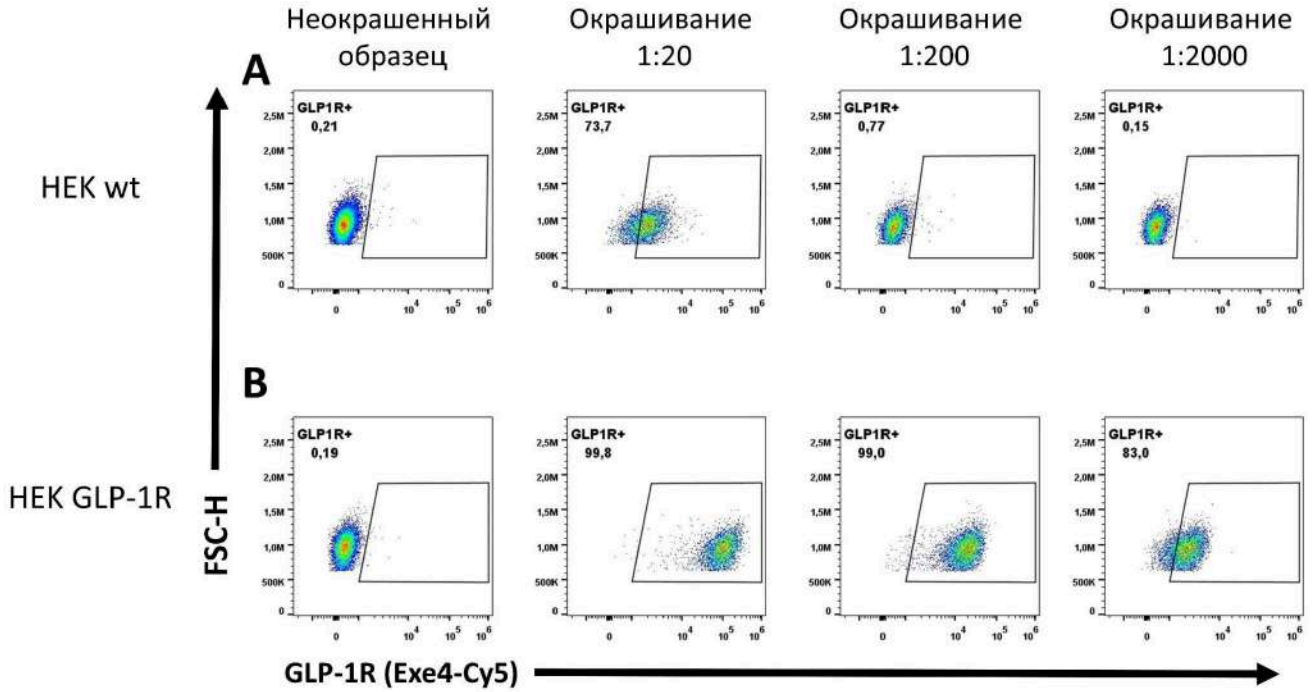



Рисунок 3.29 – Цитофлуориметрический анализ связывания Exe–4–Cy5 с клетками: (**A**) контрольными HEK wt и (**B**) HEK GLP–1R при разведениях конъюгата 1:20, 1:200 и 1:2000

После установления оптимального разведения Exe4–Cy5 для цитофлуориметрического анализа были подготовлены образцы, в которых клетки HEK wt и HEK GLP–1R смешивались в примерном соотношении 1:1, после чего окрашивали Exe4–Cy5 в разведении 1:200 (рисунок 3.30B). Подтверждена высокая специфичность конъюгата по отношению к линии, экспрессирующей GLP–1R, несмотря на присутствие в образце линии HEK wt без экспрессии рецептора, потенциально способной конкурировать за связывание с Exe4–Cy5.

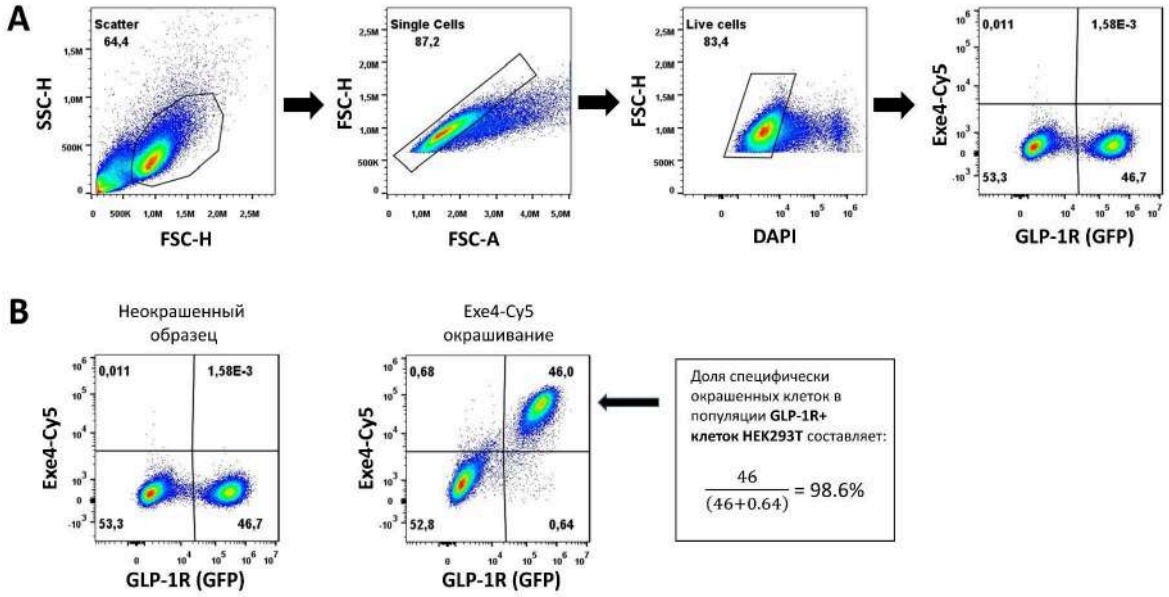


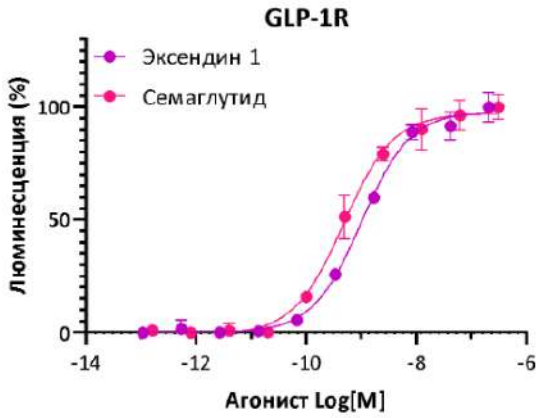

Рисунок 3.30 – Цитофлуориметрический анализ связывания Exe–4–Cy5 со смешанной культурой HEK293 wt и HEK293 GLP–1R: (**A**) репрезентативная стратегия гейтирования окрашенных клеточных образцов (**B**) распределение клеток, окрашенных конъюгатом при разведении 1:200, по сигналу от GFP (GLP1–R) и сигналу от флуоресцентной метки Cy5

### 3.2.2.5 Оценка эффективности связывания Exe4 с рецептором

Оценка активности агониста GLP–1–рецептора Exe4 с НПА была проведена с использованием эукариотической стабильной репортерной клеточной линии HEK293T–CREluc–GLP1R–GFP [169,170], созданной на основе эмбриональных клеток почки человека HEK293T. Эта клеточная линия предназначена для высокопроизводительного скрининга агонистов GLP–1R и оценки их биологической активности. Клетки экспрессируют человеческий GLP–1R и репортерный ген Firefly Luciferase. Принцип работы системы основан на природном цАМФ–зависимом сигнальном пути: связывание агониста с GLP–1R приводит к активации аденилатциклазы и повышению внутриклеточного уровня цАМФ, что сопровождается активацией протеинкиназы А и фосфорилированием транскрипционного фактора CREB. В результате взаимодействия CREB с регуляторными элементами CRE, расположенными перед геном люциферазы, происходит индукция ее экспрессии, а интенсивность люминесцентного сигнала пропорциональна степени активации рецептора.

Дополнительно конструкция содержит последовательность T2A и ген eGFP, что позволяет контролировать экспрессию GLP–1R по уровню флуоресценции.

В качестве положительного контроля был использован лекарственный пептидный препарат семаглутид (NovoNordisk, Дания). Анализ проводился в четырех независимых экспериментах с тремя техническими повторами для каждого агониста. Согласно полученным кривым «доза–ответ» (рисунок 3.31), немодифицированный Exe4 продемонстрировал практически полную активацию GLP–1R, достигая максимального ответа 98,1%. Значение $EC_{50}$ составило 1,010 нМ, а $pEC_{50}$ – 8,995. Полученные результаты свидетельствуют о высокой агонистической активности Exe4 в данной клеточной системе и подтверждают ее пригодность для последующего сравнения биологической активности модифицированных вариантов пептида, содержащих НПА.




| | Exe4 | Семаглутид |
|---|---|---|
| pEC50 | 8,995 | 9,321 |
| EC50 | 1,01 нМ | 0,478 нМ |

Рисунок 3.31 – Кривые доза–зависимость для эксендина–4 и семаглутида при активации GLP–1R в клеточной линии HEK293T–CREluc–GLP 1R–GFP

Аналогичные параметры для семаглутида составили 97,3% и 0,478 нМ (pEC50 = 9,321) соответственно. Относительная активность Exe4 составила 57,5% от активности семагутида, что примерно в 2 раза меньше и согласуется с литературными данными, характеризующими Exe4 как менее активный агонист GLP–1 рецептора по сравнению с семаглутидом примерно в 2 раза [210].

Для подтверждения рецепторной селективности Exe4 была оценена его способность активировать родственные рецепторы – GIPR и GCGR (рисунок 3.32).

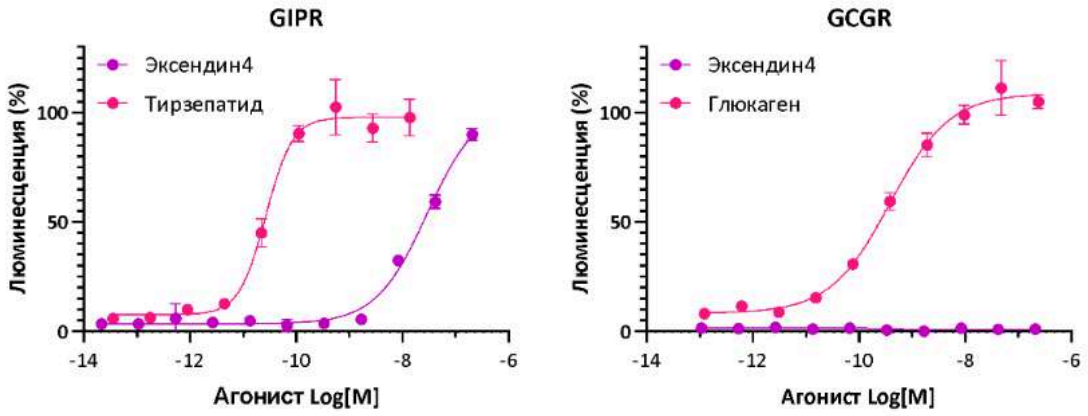
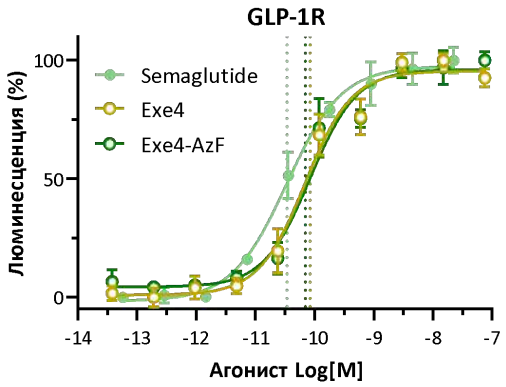



Рисунок 3.32 – оценка селективности эксендина–4 по отношению рецепторов GIPR и GCGR в сравнении с тирзепатидом и глюкагеном соответственно

В отношении рецептора GIPR Exe4 проявлял лишь минимальную агонистическую активность, тогда как тирзепатид (двойной агонист GLP–1/GIP) демонстрировал выраженную дозозависимую активацию. Для рецептора глюкагона (GCGR) Exe4 вообще не вызывал активации, в отличие от глюкагона, использованного в качестве положительного контроля. Полученный Exe4 с двумя дополнительными концевыми аминокислотами обладает высокой селективностью к GLP–1R и может быть использован как агонист сравнения при тестировании новых аналогов.

На следующем этапе была оценена активность модифицированного аналога Exe4–pAzF, содержащего pAzF в 40–м положении, относительно нативного Exe4 (рисунок 3.33).




| Пептид | EC50 |
|---|---|
| Exe4 | 1,016 нМ |
| Exe4-pAzF | 1,008 нМ |
| Семаглутид | 0,478 нМ |

Рисунок 3.33 – Сравнение активности Exe4, Exe4–pAzF и семаглутида в отношении GLP–1R в клеточной линии HEK293T–CREluc–GLP 1R–GFP

Значения $EC_{50}$ для Ex4 и Exe4–pAzF составили 1,016 нМ и 1,008 нМ соответственно, что соответствует относительной активности Exe4–pAzF на уровне 82,2% от нативного эксендин–4. Статистически значимых различий активности между Exe4–pAzF и Exe4 по отношению к GLP–1R выявлено не было, что свидетельствует о том, что введение pAzF в 40–е положение не нарушает взаимодействие пептида с GLP–1–рецептором.

### 3.2.2.6 Получение и характеристика липидных наночастиц, декорированных Эксендин–4

После отработки методики конъюгирования Exe4 с флуоресцентной меткой была выполнена функционализация белка липидом DSPE-PEG(2000)-DBCO с последующим включением полученного конъюгата (Exe4–DBCO–PEG) в состав ЛНЧ. Реакция конъюгирования с липидом проводилась до протеолитического отщепления TRX–A. Конъюгат после отщепления фьюжн–тага был очищен при помощи обращенно–фазовой ВЭЖХ. Качество полученных фракций оценивали методом аналитической ВЭЖХ (рисунок 3.34). По результатам анализа содержание минорных примесей в очищенном препарате не превышало 5%.

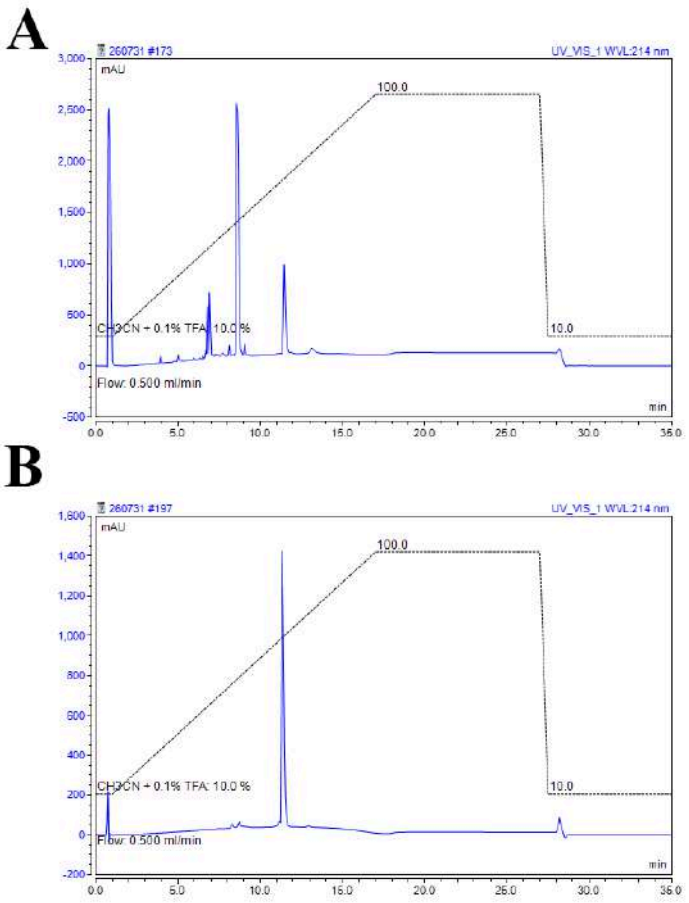



Рисунок 3.34 – Образец Exe4–DBCO–PEG после протеолиза (A) до и (B) после очистки при помощи ВЭЖХ

Для инкапсуляции мРНК–Kate2S использовалась ранее разработанная платформа LNP–5 с добавлением субфталоцианина бора(III) с аксиальной 4–карбоксифенокси–группой (sPcPHBA) [211] в молярном соотношении 1:100 ко всем липидам, входящим в состав частиц. Указанное соединение является флуоресцентным красителем, необходимым для оценки связывания частиц с клетками. Поствставку Exe4–DBCO–PEG осуществлялась по ранее разработанному для DARPin протоколу в количестве 0,10 мол.% с последующим диализом против буфера Tris–HCl. Отличительной характеристикой ЛНЧ, декорированных Exe4, было изменение ζ–потенциала (рисунок 3.35). Исходный ЛНЧ, содержащий краситель sPcPHBA, демонстрировал отрицательный ζ–потенциал (–12 мВ). Отрицательный заряд был обусловлен наличием карбоксильной группы в аксиальной группе красителя; данная полярная группа, по–видимому, локализуется во внешнем липидном слое ЛНЧ, экранируя заряд частиц. Декорированные

эксендин–4 частицы, напротив, характеризовались сменой знака поверхностного заряда (до +5мВ), что свидетельствует о включении липидированного белка во внешний липидный слой.

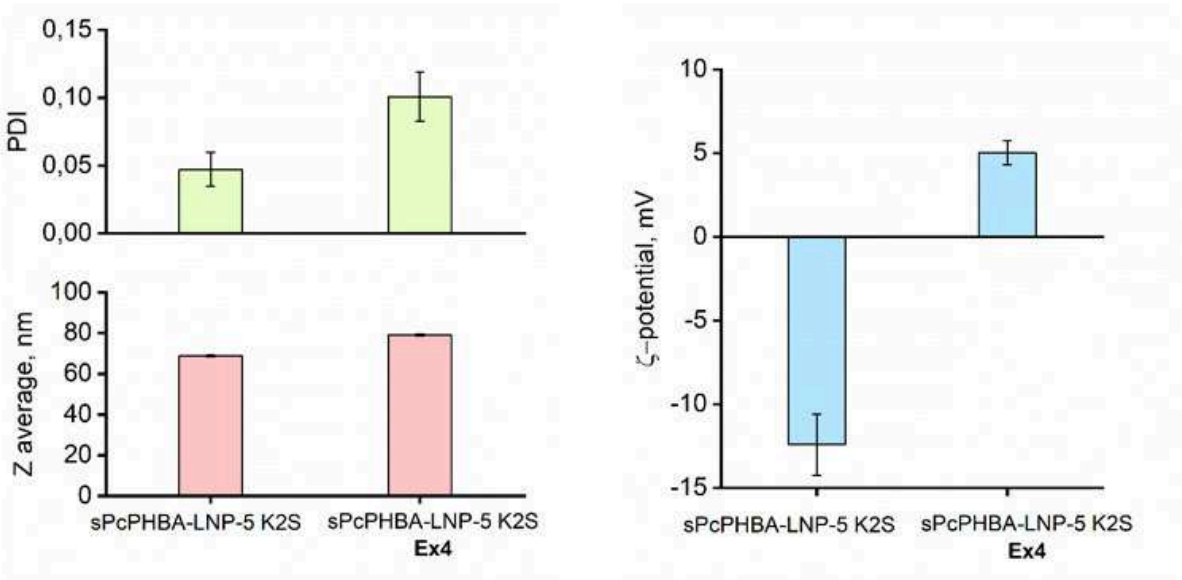



Рисунок 3.35 – Сравнение индекса полидисперсности и среднего гидродинамического диаметра (Z-average) наночастиц до и после декорирования Exe4–DBCO–PEG (0,10мол.%)

В таблице 3.5 приведены все полученные физико–химические характеристики наноносителей. Все ЛНЧ обладают высокой степенью инкапусляции до и после введения Exe4–DBCO–PEG (EE 90+%).

Таблица 3.5 – Физико–химическая характеристика исследованных ЛНЧ без декорации и ЛНЧ@Exe4

| ЛНЧ | Z средний диаметр, nm | PDI | Частицы/мл | ζ–потенциа, mV | EE, % |
|---|---|---|---|---|---|
| sPcPHBA–LNP–5 (без Exe4, Tris) | 68.81 ± 0.31 | 0.05 ± 0.01 | $3.10 \cdot 10^{12}$ | –12.41 ± 1.81 | 99 |
| sPcPHBA–LNP–5@Exe4 (Tris) | 79.05 ± 0.45 | 0.10 ± 0.02 | $1.33 \cdot 10^{12}$ | +5.02 ± 0.71 | 92 |

Специфичность таргетирования была оценена *in vitro* путем трансфекции образцов мРНК-ЛНЧ Kate2S, содержащих или не содержащих Exe4–DBCO–PEG. Анализ проводился в

монокультурах HEK293T wt и HEK293T GLP1R+ GFP+, а также в ко–культуре, где отдельно оценивались субпопуляции GFP− (wt) и GFP+ (GLP1R+). Нетрансфицированные клетки служили отрицательным контролем. Все эксперименты были проведены в трех повторах. Через 24 ч после трансфекции методом проточной цитометрии оценивали долю Kate2S+ клеток (%), sPcPHBA+ клеток (%) и среднюю интенсивность флуоресценции Kate2S (MFI) среди живых клеток (DAPI−) в соответствии со стратегией гейтирования, приведенной на рисунке 3.36A.

При анализе ко–культуры клеток по окрашиванию sPcPHBA, входящего в состав ЛНЧ, статистически достоверных различий в MFI sPcPHBA+ клеток между группами обнаружено не было (рисунок 3.36C). Полученный результат, вероятно, отражает высокую эффективность ассоциации частиц с клеточной поверхностью: количество ЛНЧ, добавленных к клеткам, составляет порядка $10^{10}$ частиц на лунку, что представляет собой значительный избыток и обеспечивает практически полное окрашивание клеток sPcPHBA независимо от наличия декорации Exe4. Вместе с тем связывание ЛНЧ с клеточной мембраной, регистрируемое по сигналу sPcPHBA, само по себе не свидетельствует о проникновении частиц в клетку, высвобождении РНК из эндосом и трансляции белка с матрицы РНК. В связи с этим более репрезентативными показателями успешной трансфекции являются доля Kate2S+ клеток и медианная интенсивность флуоресценции (MFI) по белку Kate2S, которые отражают завершенность внутриклеточного пути доставки и эффективность экспрессии целевой РНК.

В ко–культуре добавление ЛНЧ без декорации обеспечивало 84,5% Kate2S+ клеток с MFI 3475,3 среди клеток дикого типа HEK293T wt (рисунок 3.36B). Декорирование ЛНЧ–Exe4 привело к небольшому увеличению доли Kate2S+ клеток до 88,8%, при этом MFI (Kate2S) статистически значимо снижалась до 2955. Блокировка GLP–1R снижала MFI до 1961,3 при доле Kate2S+ клеток 85,0% (рисунок 3.36D).

При анализе GLP–1R+ клеток в составе ко–культуры трансфекция немодифицированными ЛНЧ обеспечивала 70,1% Kate2S+ клеток с MFI 1987,3 (рисунок 3.36B). Использование ЛНЧ–Exe4 приводило к статистически достоверному повышению доли Kate2S+ клеток до 91,7% и выраженному почти пятикратному увеличению MFI до 8432,0. Предварительная блокировка GLP–1R снижала долю Kate2S+ клеток до 48,9% и MFI до 1138,7 (рисунок 3.36D). Таким образом, в ко–культуре таргетный эффект ЛНЧ–Exe4 был специфичен для GLP1R–экспрессирующих клеток, что подтверждает рецептор–опосредованное усиление трансфекции с использованием декорированных частиц.

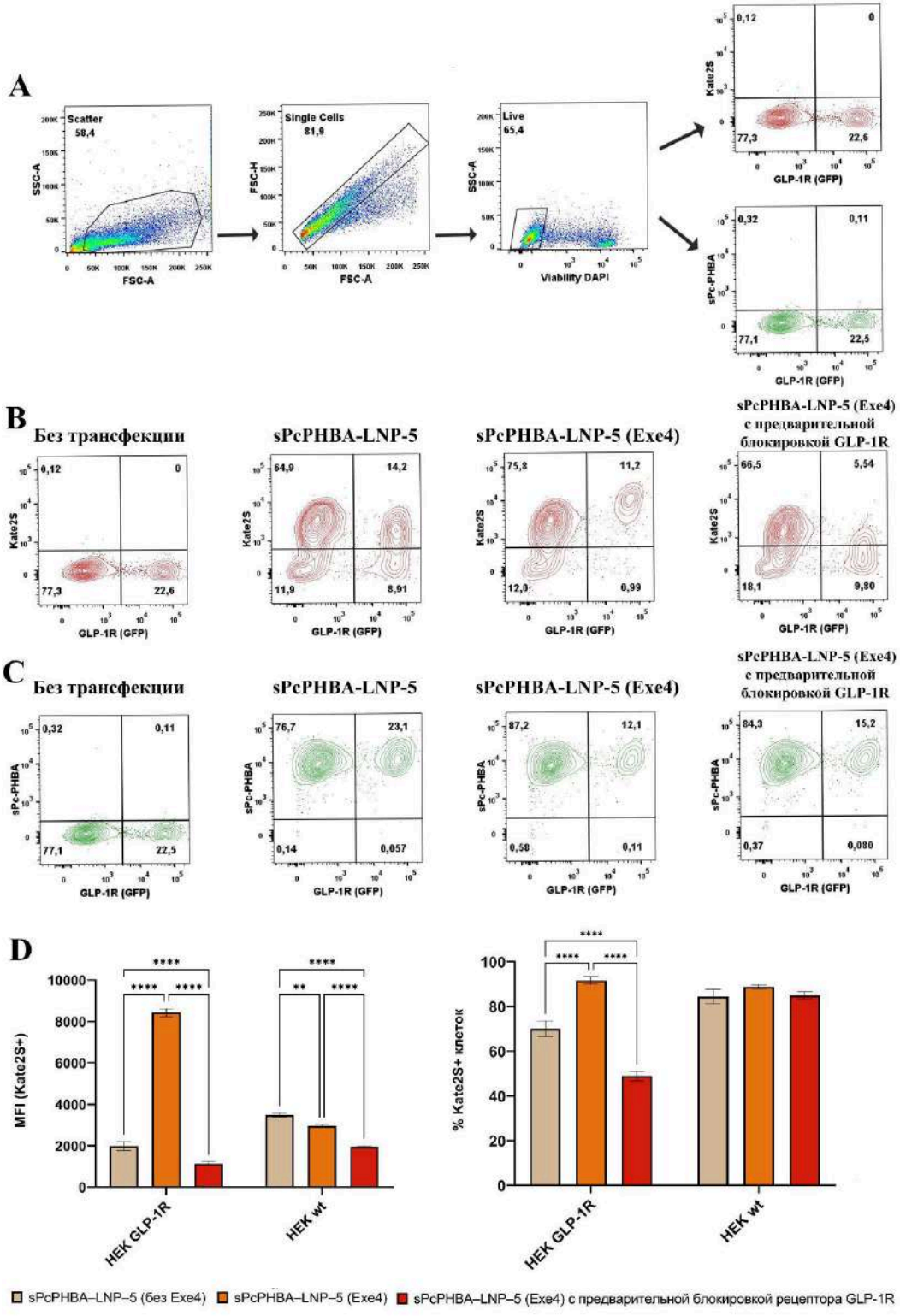



Рисунок 3.36 – Специфичность ЛНЧ, нацеленных на GLP–1R. **(A)** Стратегия гейтирования и репрезентативные графики, демонстрирующие экспрессию Kate2S **(B)** и окрашивание sPcPHBA **(C)** в ко–культуре клеток HEK wt и HEK GLP–1R(GFP)+. (D) MFI и количество Kate2S+ клеток в ко–культуре (n=3). Данные представлены как среднее значение ± SD. **p<0,01, ****p<0,0001. Для обработки использованы односторонний дисперсионный анализ, тест множественного сравнения по Шидаку

Профиль трансфекции wt–клеток в монокультурах соответствовал таковому в ко–культуре (рисунок 3.37 A и B). В монокультуре HEK293T wt добавление ЛНЧ без модификации приводило к трансфекции 89,9% клеток со средним уровнем MFI 3953,7 (рисунок 3.37B). ЛНЧ с Exe4 (ЛНЧ–Exe4) не вызывали статистически достоверных изменений по доле Kate2S+ клеток (89,3%), при этом наблюдалось небольшое снижение MFI до 3432,7. Предварительная блокировка рецептора GLP–1R конъюгатом Exe4–PEG–DBCO перед добавлением ЛНЧ–Exe4 снижала MFI до 2146,7 и долю Kate2S+ клеток до 85,0% (рисунок 3.37C).

Выраженное рецептор–опосредованное усиление трансфекции клеток HEK293T GLP1R+ GFP+, наблюдавшееся в формате ко–культуры, тотмечалось и при трансфекции монокультуры (рисунок 3.37A). Трансфекция немодифицированными ЛНЧ обеспечивала 60,3% Kate2S+ клеток с MFI 1716,3, тогда как использование ЛНЧ–Exe4 приводило к статистически достоверному увеличению доли Kate2S+ клеток до 90,1%, так и 4,1–кратному повышению MFI до 7107,3 (рисунок 3.37C). Предварительная блокировка GLP–1R вызывала резкое снижение обоих показателей: доля Kate2S+ клеток составила 29,5%, MFI – 1052,7. Таким образом, в клетках, экспрессирующих GLP–1R, декорирование ЛНЧ лигандом Exe4 обеспечивает выраженное усиление трансфекции, которое отменяется блокадой рецептора.

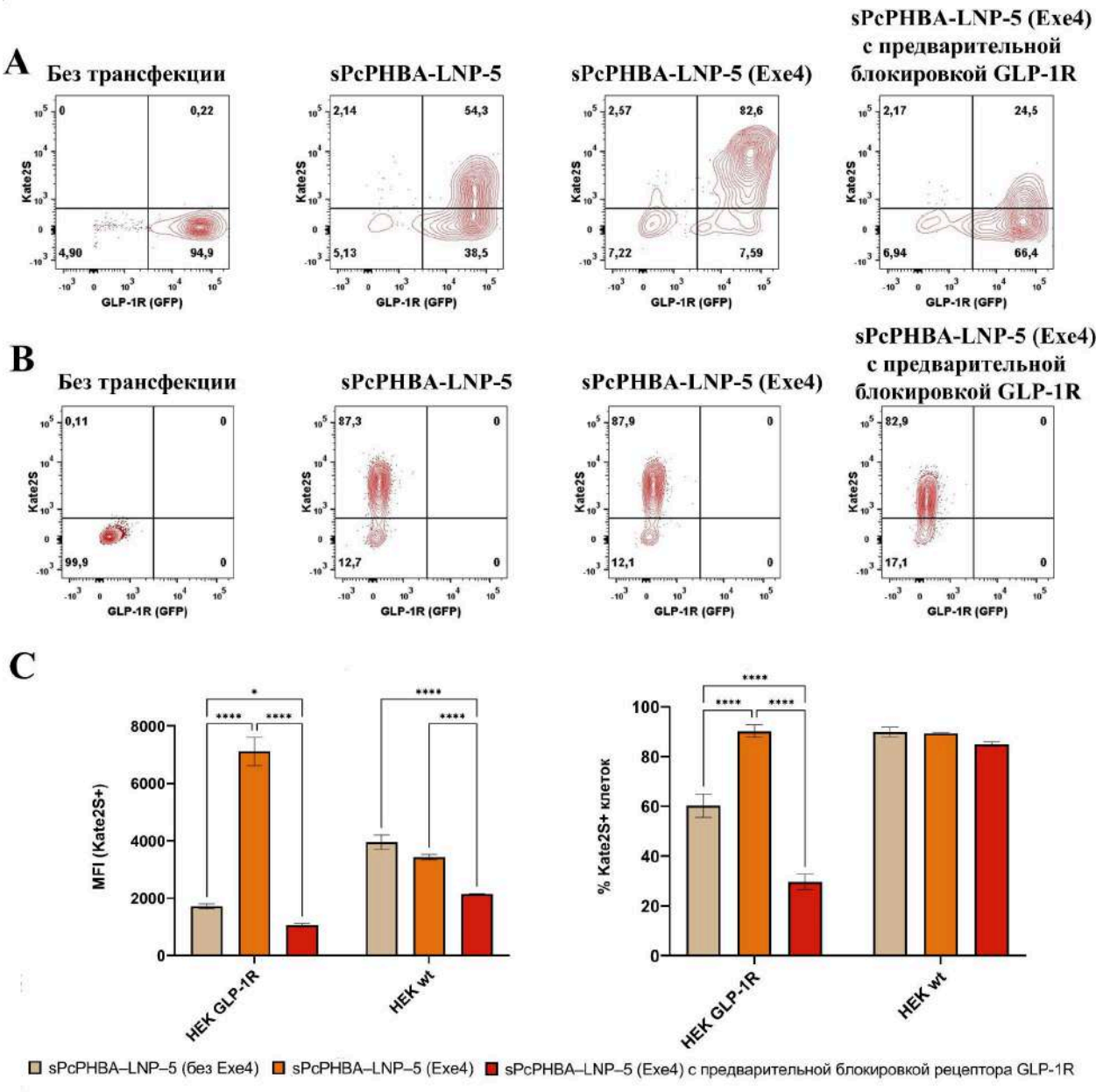



Рисунок 3.37 – Специфичность ЛНЧ, нацеленных на GLP–1R. Репрезентативные графики, демонстрирующие экспрессию Kate2S в HEK GLP–1R(GFP)+ (A) и HEK wt (B). (C) MFI и количество Kate2S+ клеток в монокультурах (n=3). Данные представлены как среднее значение ± SD. **p<0,01, ****p<0,0001. Для обработки использованы односторонний дисперсионный анализ, тест множественного сравнения по Шидаку

На основе полученных данных по таргетной доставке мРНК с помощью ЛНЧ, функционализированных DARPin и Exe4, можно заключить, что подход, основанный на сайт-специфическом включении pAzF в аминокислотную последовательность целевого белка с

последующей биоортогональной конъюгацией на поверхности ЛНЧ, представляет собой универсальную платформенную технологию. Предлагаемая стратегия не ограничивается белками, использованными в данном диссертационном исследовании, и применима к широкому спектру рекомбинантных пептидов. Благодаря высокой селективности и мягким условиям реакции азидно-алкинового циклоприсоединения этот метод обеспечивает контролируемую, пространственно направленную функционализацию наноносителей, открывая возможности для гибкого декорирования поверхности наночастиц.

# ЗАКЛЮЧЕНИЕ

В рамках данной диссертационной работы была успешно реализована и оптимизирована многоступенчатая схема направленной эволюции пирролизил–аминоацил–тРНК–синтетаз с использованием метода PANCE. Ключевым нововведением стала разработка протокола негативного отбора, в котором селекция в отсутствие НПА проводилась после каждого этапа позитивного отбора, что обеспечивало эффективное подавление неспецифических или «условно активных» вариантов фермента, способных функционировать без целевого субстрата. Этот итеративный подход, основанный на чередовании этапов, направленных на усиление целевой функции и устранение вариантов с субоптимальными свойствами, существенно повышал селективное давление в пользу синтетаз, активность которых действительно зависит от присутствия НПА. Весь процесс контролировался с помощью стандартизированных параметров: оптической плотности культуры, времени культивирования и титра инфекционных вирусных частиц, что обеспечивало воспроизводимый и управляемый отбор вариантов синтетаз с улучшенными каталитическими свойствами и строгой субстратной специфичностью. В результате были получены варианты PylRS, адаптированные к четырем ароматическим НПА: O–метил–L–тирозину (oMeY), 4–азидо–L–фенилаланину (pAzF), O–аллил–L–тирозину (oAllylY) и 2–хлор–L–фенилаланину (2–ClF).

Полученные варианты фермента – AllYPylRS, MethYPylRS, AzFPylRS и 2ClFPylRS – характеризуются повышенной активностью включения соответствующих НПА по сравнению с исходными вариантами, не проходившими эволюцию. Масс–спектрометрический анализ подтвердил образование полноразмерного sfGFP с соответствующей НПА в заданной позиции. Для исследованных вариантов синтетаз доля белка, содержащего целевую НПА, превышала 89% среди идентифицированных продуктов. Эти результаты подтверждают высокую активность разработанных систем расширения генетического кода и демонстрируют возможность использования модифицированных PylRS для получения белков с НПА.

Одновременно было установлено, что повышение активности PylRS не всегда сопровождается ростом его субстратной специфичности. Анализ перекрестной реактивности показал, что некоторые эволюционировавшие варианты сохраняют значительную активность в отношении структурно родственных НПА. В частности, AzFPylRS сохраняла высокую активность как в отношении своего основного субстрата pAzF (35,2%), так и других НПА: oMeY (29,5%), oAllylY (27%) и 2–ClF (26,2%), а MethYPylRS – в отношении oMeY и pAzF на

уровне 34,7 и 30,7% соответственно. Полученные результаты показывают, что активность и субстратная специфичность PylRS могут изменяться независимо друг от друга, что необходимо учитывать при создании синтетаз с заданными свойствами.

Новизна данного исследования заключается в оптимизации методологии, позволяющей получать новые оригинальные варианты синтетаз с улучшенными свойствами. Анализ последовательностей выявил аминокислотные замены в различных функциональных областях PylRS, включая карман связывания аминокислоты, N– и C–концевые домены фермента. Полученные данные указывают на то, что повышение активности PylRS к выбранным НПА может быть связано не только с изменениями непосредственно в области кармана связывания аминокислоты, но и с перестройкой взаимодействий между отдельными участками фермента и тРНК. Такие изменения могут свидетельствовать о зависимости селективного преимущества отдельных замен от генетического фона и указывать на возможное наличие эпистатических взаимодействий между ними. Сравнение независимых эволюционных траекторий выявило сочетание конвергентных и субстрат–специфичных изменений. Повторное появление отдельных замен может свидетельствовать о существовании общих адаптивных решений, повышающих функциональную активность PylRS в отношении ароматических НПА, тогда как различия в составе и динамике других замен между вариантами, специфичными к oAllylY, oMeY, pAzF и 2–ClF, указывают на последующую адаптацию фермента к структурным особенностям конкретного субстрата. Результаты эволюции и скрининга АРС показывают, что взаимодействие между НПА и ферментом может усиливаться множественными аминокислотными заменами, которые формируют и тонко настраивают селективность и активность. Это открывает новые возможности для углубленного изучения PylRS и создает основу для исследования других субстратов.

Особый интерес представляет аминокислотная замена Met300Thr, которая независимо возникала в нескольких эволюционных линиях. Ее раннее появление и высокая частота на первых этапах эволюции позволяют предположить, что данная замена может обеспечивать общее преимущество при адаптации PylRS к ароматическим НПА. Вместе с тем отсутствие полной фиксации Met300Thr на поздних этапах эволюции указывает на то, что она, вероятно, не является единственным определяющим фактором повышения активности и может представлять собой одну из ранних адаптивных замен, после которой происходит дальнейшая оптимизация фермента за счет дополнительных изменений. Данные функционального тестирования напрямую подтверждают выводы, полученные на основе анализа динамики накопившихся замен и 3D–моделирования. Аминокислотная замена Met300Thr не повышает сродство к одной

конкретной НПА, а приводит к общему существенному повышению активности синтетазы в отношении производных фенилаланина и тирозина, использованных в данном исследовании. Введение данной аминокислотной замены увеличивает сродство PylRS к выбранным НПА, тогда как остальные замены, вероятно, в большей степени определяют субстратную специфичность фермента.

Практическая значимость полученных результатов особенно очевидна для pAzF–содержащих белков, поскольку азидогруппа предоставляет возможность последующей биоортогональной модификации белка. Возможность такого применения была продемонстрирована на примере двух белков: Exe4 и DARPin mCD8. Полученные белки с pAzF в составе сохраняли способность взаимодействовать с соответствующими рецепторами и могли быть сайт–специфически модифицированы посредством безмедной клик–химической реакции SPAAC с DBCO–содержащими соединениями.

Модифицированные DARPin и Exe4 использовались для функционализации липидных наночастиц. Белки с азидогруппой предварительно конъюгировали с DSPE–PEG–DBCO посредством SPAAC, после чего полученный белок–липидный конъюгат вводили в предварительно сформированные ЛНЧ методом поствставки. Показано, что добавление конъюгата белка с липидом в концентрации 0,10 мол.% от общего содержания липидов с последующим диализом против буфера 25 мМ Tris pH 7,5, позволяет эффективно модифицировать поверхность наночастиц без существенного нарушения их физико–химических характеристик.

Функциональные исследования показали, что полученные DARPin–ЛНЧ обеспечивали эффективное и селективное таргетирование CD8 положительных клеток. По сравнению с немодифицированными частицами и системами с пассивной адсорбцией белка ковалентно модифицированные DARPin–ЛНЧ демонстрировали повышенный уровень трансфекции CD8+ клеток при минимальном воздействии на клетки, не экспрессирующие CD8. Блокирование CD8–рецептора существенно снижало эффективность трансфекции, что подтверждает рецептор–зависимый механизм таргетирования.

Декорирование липидных наночастиц Exe4 также обеспечивает рецептор–опосредованное повышение эффективности трансфекции клеток, экспрессирующих GLP–1R. Усиление экспрессии репортерного белка Kate2S, проявляющееся увеличением доли Kate2S–положительных клеток и интенсивности флуоресценции, существенно снижается при предварительном блокировании GLP–1R. Полученные результаты подтверждают потенциал эксендин–4–модифицированных липидных наночастиц для таргетной доставки РНК в

GLP–1R–положительные клетки. Таким образом, было показано, что подход, основанный на сайт–специфическом включении pAzF в аминокислотную последовательность целевого белка с его последующей биоортогональной конъюгацией на поверхности ЛНЧ, представляет собой универсальную платформенную технологию.

Совокупность полученных в настоящей работе результатов подтверждает перспективность разработанных систем PylRS не только как инструментов для сайт–специфического введения НПА, но и как основы для последующего создания функционально модифицированных белков и нанобиотехнологических платформ. Выявленные аминокислотные замены и особенности их эволюционной динамики могут быть использованы для рационального проектирования PylRS с заданными характеристиками активности и субстратной специфичности, а разработанный подход PANCE – для получения синтетаз, способных эффективно распознавать и включать другие НПА. Перспективными направлениями дальнейшей работы являются повышение субстратной специфичности полученных PylRS, количественная оценка механизмов действия выявленных замен, а также применение разработанных систем для создания адресной доставки биомолекул.

# ВЫВОДЫ

1. Оптимизирован протокол направленной эволюции пирролизиновых аминоацил–тРНК–синтетаз (PylRS) на основе фаг–ассоциированной эволюции PANCE, включающий этап отрицательной селекции после каждого раунда положительного отбора и контроль её эффективности в присутствии и отсутствии неприродных аминокислот (НПА), что позволяет получать высокоактивные варианты PylRS при одновременном снижении фонового неспецифического аминоацилирования.
2. Получены варианты PylRS, адаптированные к включению O–метил–L–тирозина (oMeY), O–аллил–L–тирозина (oAllylY), 4–азидо–L–фенилаланина (pAzF) и 2–хлор–L–фенилаланина (2–ClF) в белки, причем доля полноразмерного sfGFP с целевой НПА превышает 89% среди идентифицированных полноразмерных продуктов по данным масс–спектрометрического анализа.
3. Обнаруженные аминокислотные замены в последовательности синтетаз, ранее не рассматривавшиеся в качестве мишеней для рационального проектирования, затрагивают не только карман связывания субстрата, но и перестраивают взаимодействия между отдельными участками фермента и тРНК, что указывает на сложный эпистатический характер формирования активности и специфичности к новому субстрату.
4. Показано, что конвергентно возникшая аминокислотная замена Met300Thr в каталитическом домене MmPylRS повышает активность синтетазы в отношении производных тирозина и фенилаланина. В совокупности с анализом эволюционных траекторий и результатами структурного моделирования это указывает на роль Met300Thr в общей адаптации MmPylRS к ароматическим НПА, тогда как дополнительные замены могут вносить вклад в настройку субстратной специфичности.
5. Сайт–специфическое включение 4–азидо–L–фенилаланина в Exe4 и DARPin mCD8 обеспечивает возможность их биоортогональной конъюгации с DBCO–содержащими соединениями посредством реакции SPAAC без изменения их специфичности по отношению к соответствующему рецептору.
6. Разработанный метод поствставки белок–липидных конъюгатов Exe4 и DARPin mCD8 в предварительно сформированные липидные наночастицы (ЛНЧ) в концентрации 0,10 мол.% от общего содержания липидов с последующим диализом против буфера 25 мМ

Tris, pH 7,5, обеспечивает ковалентную функционализацию поверхности ЛНЧ при сохранении их физико–химических характеристик.

7. Функционализация ЛНЧ конъюгатами Exe4 и DARPin mCD8 повышает эффективность трансфекции GLP–1R– и CD8–положительных клеток. Снижение эффективности трансфекции при блокировании соответствующих рецепторов подтверждает рецептор–опосредованное повышение эффективности трансфекции целевой популяции клеток.

## СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ

**2–ClF** – 2–хлор–L–фенилаланин

**2ClFPylRS** – пирролизил–тРНК–синтетаза, специфичная к 2–хлор–L–фенилаланину

**AcK** – Nε–ацетиллизин

**AllYPylRS** – пирролизил–тРНК–синтетаза, специфичная к O–аллил–L–тирозину

**AP** – вспомогательная плазмида (auxiliary plasmid)

**AzFPylRS** – пирролизил–тРНК–синтетаза, специфичная к 4–азидо–L–фенилаланину

**ChPylRS** – химерная пирролизил–тРНК–синтетаза на основе N–домена *Methanosarcina barkeri* и C–домена *Methanosarcina mazei*

**ChPylRS_mut** – мутантная форма химерной пирролизил–тРНК–синтетазы

**CP** – дополнительная плазмида (complementary plasmid)

**Cy5–DBCO** – конъюгат флуоресцентного красителя Cy5 с дибензоциклооктином

**DARPin–Cys** – сконструированный белок–повтор анкирина, содержащий остаток цистеина

**DARPin–Cys–Mal** – DARPin против CD8, конъюгированный с DMG–PEG–2000–Mal посредством малеимид–тиольной связи

**DARPin–pAzF** – сконструированный белок–повтор анкирина, содержащий 4–азидо–L–фенилаланин

**DARPins** – сконструированные белки–повторы анкирина (Designed Ankyrin Repeat Proteins)

**DBCO** – дибензоциклооктин (dibenzocyclooctyne)

**Dh–PylSc** – каталитический C–домен пирролизил–тРНК–синтетазы из *Desulfitobacterium hafniense*

**DSPE–PEG–Mal** – конъюгат дис–стеароилфосфатидилэтаноламина с полиэтиленгликолем и малеимидной группой

**EcoLeuRS** – лейцил–тРНК–синтетаза *Escherichia coli*

**EcoTyrRS** – тирозил–тРНК–синтетаза *Escherichia coli*

**EPL** – лигирование экспрессированного белка (expressed protein ligation)

**Exe4** – эксендин–4

**FIT** – система гибкой бесклеточной трансляции (Flexible In vitro Translation)

**hEKL** – лёгкая цепь человеческой энтеропептидазы (human enteropeptidase light chain)

**IPAD–1** – конъюгат белка с липидом DSPE–PEG(2000)–DBCO

**LNP–5** – липидные наночастицы, включающие два ионизируемых липида (SM–102 и ALC–0315), холестерин, DSPC и DMG–PEG–2000 в молярном соотношении 23.15:23.15:42.7:9.4:1.6 соответственно

**LUCA** – последний универсальный общий предок (Last Universal Common Ancestor)

**MaPylRS** – пирролизил–тРНК–синтетаза из *Candidatus Methanomethylophilus alvus*

**MbPylRS** – пирролизил–тРНК–синтетаза из *Methanosarcina barkeri*

**MP** – плазмида мутагенеза (mutagenesis plasmid)

**mCD8** – мышиный антиген CD8

**MethYPylRS** – пирролизил–тРНК–синтетаза, специфичная к O–метил–L–тирозину

**MmPylRS** – пирролизил–тРНК–синтетаза из *Methanosarcina mazei*

**MmPylRS_mut** – мутантная форма пирролизил–тРНК–синтетазы из *Methanosarcina mazei*

**NCL** – нативное химическое лигирование (native chemical ligation)

**NRPS** – нерибосомный синтез пептидов (non–ribosomal peptide synthesis)

**Nε–(p–азидобензилоксикарбонил)–L–лизин (pAzZLys)** – неприродная аминокислота, содержащая азидную группу

**Nε–(p–этинилбензилоксикарбонил)–L–лизин (pEtZLys)** – неприродная аминокислота, содержащая этинильную группу

**oAllylY** – O–аллил–L–тирозин

**oMeY** – O–метил–L–тирозин

**PACE** – непрерывная эволюция фагов (phage–assisted continuous evolution)

**PANCE** – непрерывная эволюция с использованием фагов и без непрерывного культивирования (phage–assisted non–continuous evolution)

**pAzF** – 4–азидо–L–фенилаланин

**pIII–neg** – вариант белка pIII с интактными доменами N1 и N2 и мутантным C–доменом с внутренней делецией 70 аминокислот

**PEG** – полиэтиленгликоль (polyethylene glycol)

**PylB** – ген кластера пирролизина, кодирующий фермент синтеза S–аденозил–L–метионина

**PylC** – ген кластера пирролизина, кодирующий белок семейства карбамоилфосфатсинтетаз

**PylD** – ген кластера пирролизина, кодирующий НАД–зависимую дегидрогеназу

**PylRS** – пирролизил–тРНК–синтетаза

**PylRS–AA** – вариант пирролизил–тРНК–синтетазы с заменой аминокислот в позициях 346 и 348 на аланин

**PylSc** – каталитический домен пирролизил–тРНК–синтетазы

**PylSn** – N– и T–концевые домены пирролизил–тРНК–синтетазы

**PylSn+PylSc** – пирролизил–тРНК–синтетазы, домены которых экспрессируются в виде отдельных белков, кодируемых разными генами

**PylSn–PylSc** – пирролизил–тРНК–синтетазы, домены которых соединены линкером

**ΔPylSn** – пирролизил–тРНК–синтетазы, содержащие только С–концевой домен

**PylS** – ген кластера пирролизина, кодирующий пирролизил–тРНК–синтетазу

**PylT** – ген кластера пирролизина, кодирующий тРНК пирролизина

**RaPID** – метод интегрированного скрининга случайных нестандартных пептидов (Random non–standard Peptides Integrated Discovery)

**SPAAC** – азид–алкиновое циклоприсоединение без участия меди (strain–promoted azide–alkyne cycloaddition)

**SPPS** – твердофазный синтез пептидов (Solid–Phase Peptide Synthesis)

**TRX–A** – тиоредоксин А (Thioredoxin A)

**АРС** – аминоацил–тРНК–синтетаза

**НПА** – неприродная аминокислота

**СТЭМ** – сканирующая просвечивающая электронная микроскопия

**тРНК** – транспортная рибонуклеиновая кислота

**тРНК$_{Pyl}$** – транспортная РНК пирролизина

**ВЭЖХ** – высокоэффективная жидкостная хроматография

## СПИСОК ЛИТЕРАТУРЫ


1. Kumar V. и др. Therapeutic proteins: developments, progress, challenges, and future perspectives // 3 Biotech. 2024. Т. 14, № 4. С. 112.
2. Ebrahimi S. B., Samanta D. Engineering protein–based therapeutics through structural and chemical design // Nat Commun. 2023. Т. 14, № 1. С. 2411.
3. Walsh C. T., O'Brien R. V., Khosla C. Nonproteinogenic Amino Acid Building Blocks for Nonribosomal Peptide and Hybrid Polyketide Scaffolds // Angew Chem Int Ed Engl. 2013. Т. 52, № 28. С. 7098–7124.
4. Huang Y., Liu T. Therapeutic applications of genetic code expansion // Synthetic and Systems Biotechnology. 2018. Т. 3, № 3. С. 150.
5. Huang Y. и др. Genetic Code Expansion: Recent Developments and Emerging Applications // Chem Rev. 2025. Т. 125, № 2. С. 523–598.
6. Natter Perdiguero A., Deliz Liang A. Practical Approaches to Genetic Code Expansion with Aminoacyl–tRNA Synthetase/tRNA Pairs // Chimia (Aarau). 2024. Т. 78, № 1–2. С. 22–31.
7. Lau J. и др. Discovery of the Once–Weekly Glucagon–Like Peptide–1 (GLP–1) Analogue Semaglutide // J Med Chem. 2015. Т. 58, № 18. С. 7370–7380.
8. Bo Q. и др. Structural insights into the activation of somatostatin receptor 2 by cyclic SST analogues // Cell Discov. 2022. Т. 8, № 1. С. 47.
9. Li Y.–G. и др. Structural insights into somatostatin receptor 5 bound with cyclic peptides // Acta Pharmacol Sin. 2024. Т. 45, № 11. С. 2432–2440.
10. Li Q. и др. Developing Covalent Protein Drugs via Proximity–Enabled Reactive Therapeutics // Cell. 2020. Т. 182, № 1. С. 85–97.e16.
11. Arias M. и др. Selective anticancer activity of synthetic peptides derived from the host defence peptide tritrpticin // Biochim Biophys Acta Biomembr. 2020. Т. 1862, № 8. С. 183228.
12. Zampaloni C. и др. A novel antibiotic class targeting the lipopolysaccharide transporter // Nature. 2024. Т. 625, № 7995. С. 566.
13. Scinto S. L. и др. Bioorthogonal chemistry // Nat Rev Methods Primers. 2021. Т. 1. С. 30.
14. Reddington S. и др. Genetically encoding phenyl azide chemistry: new uses and ideas for classical biochemistry // Biochem Soc Trans. 2013. Т. 41, № 5. С. 1177–1182.
15. Liu C. C., Schultz P. G. Adding new chemistries to the genetic code // Annu Rev Biochem. 2010. Т. 79. С. 413–444.
16. Samieipour F., Dianat–Moghadam H., Khanahmad H. Recent developments in bioconjugation: From strategies to design and clinical applications // Biomedicine & Pharmacotherapy. 2025. Т. 192. С. 118593.
17. Yu C., Bao G., Sun W. Recent Advances on Dehydroalanine–Specific Modification and Diversification of Peptides and Proteins // Chembiochem. 2025. Т. 26, № 20. С. e202500349.
18. Gorzeń O. и др. Engineering unnatural amino acids in peptide linkers enables cathepsin–selective antibody–drug conjugates for HER2–positive breast cancer // J Control Release. 2025. Т. 387. С. 114269.
19. Stephanopoulos N., Francis M. B. Choosing an effective protein bioconjugation strategy // Nat Chem Biol. 2011. Т. 7, № 12. С. 876–884.
20. Chen Y.–X., Triola G., Waldmann H. Bioorthogonal chemistry for site–specific labeling and surface immobilization of proteins // Acc Chem Res. 2011. Т. 44, № 9. С. 762–773.
21. Skidmore L. и др. ARX788, a Site–specific Anti–HER2 Antibody–Drug Conjugate, Demonstrates Potent and Selective Activity in HER2–low and T–DM1–resistant Breast and Gastric Cancers // Mol Cancer Ther. 2020. Т. 19, № 9. С. 1833–1843.

22. Nagaraja Shastri P. и др. Nonclinical Development of Next–generation Site–specific HER2–targeting Antibody–drug Conjugate (ARX788) for Breast Cancer Treatment // Mol Cancer Ther. 2020. Т. 19, № 9. С. 1822–1832.
23. Hu X. и др. ACE–Breast–02: a randomized phase III trial of ARX788 versus lapatinib plus capecitabine for HER2–positive advanced breast cancer // Signal Transduct Target Ther. 2025. Т. 10, № 1. С. 56.
24. Elia N. Using unnatural amino acids to selectively label proteins for cellular imaging: a cell biologist viewpoint // FEBS J. 2021. Т. 288, № 4. С. 1107–1117.
25. Peng T., Hang H. C. Site–Specific Bioorthogonal Labeling for Fluorescence Imaging of Intracellular Proteins in Living Cells // J Am Chem Soc. 2016. Т. 138, № 43. С. 14423–14433.
26. Ding W. и др. Computational design and genetic incorporation of lipidation mimics in living cells // Nat Chem Biol. 2024. Т. 20, № 1. С. 42–51.
27. Honarvar H. и др. Position for site–specific attachment of a DOTA chelator to synthetic affibody molecules has a different influence on the targeting properties of 68Ga– compared to 111in–labeled conjugates // Mol Imaging. 2014. Т. 13.
28. Chen L.–A. и др. Site–specific Conjugation of 6 DOTA Chelators to a CA19–9–targeting scFv–Fc Antibody for Imaging and Therapy // J Med Chem. 2023. Т. 66, № 15. С. 10604–10616.
29. Nilsson B. L., Soellner M. B., Raines R. T. Chemical synthesis of proteins // Annu Rev Biophys Biomol Struct. 2005. Т. 34. С. 91–118.
30. Hodgson D. R. W., Sanderson J. M. The synthesis of peptides and proteins containing non–natural amino acids // Chem Soc Rev. 2004. Т. 33, № 7. С. 422–430.
31. Amblard M. и др. Methods and protocols of modern solid phase Peptide synthesis // Mol Biotechnol. 2006. Т. 33, № 3. С. 239–254.
32. Amblard M. и др. Fundamentals of modern peptide synthesis // Methods Mol Biol. Clifton, N.J., 2005. Т. 298. С. 3–24.
33. Paravizzini S. J. и др. Backbone Protecting Groups for Enhanced Peptide and Protein Synthesis // Angew Chem Int Ed Engl. International ed. in English, 2025. Т. 64, № 33. С. e202509939.
34. Kent S. B. H. Total chemical synthesis of proteins // Chem Soc Rev. 2009. Т. 38, № 2. С. 338–351.
35. Conibear A. C. и др. Native chemical ligation in protein synthesis and semi–synthesis // Chem Soc Rev. 2018. Т. 47, № 24. С. 9046–9068.
36. Muir T. W. Semisynthesis of proteins by expressed protein ligation // Annu Rev Biochem. 2003. Т. 72. С. 249–289.
37. Spicer C. D., Davis B. G. Selective chemical protein modification // Nat Commun. 2014. Т. 5. С. 4740.
38. Chalker J. M. и др. Chemical modification of proteins at cysteine: opportunities in chemistry and biology // Chem Asian J. 2009. Т. 4, № 5. С. 630–640.
39. Bernardes G. J. L. и др. Facile conversion of cysteine and alkyl cysteines to dehydroalanine on protein surfaces: versatile and switchable access to functionalized proteins // J Am Chem Soc. 2008. Т. 130, № 15. С. 5052–5053.
40. Mueller L. K. и др. Challenges and Perspectives in Chemical Synthesis of Highly Hydrophobic Peptides // Front Bioeng Biotechnol. 2020. Т. 8. С. 162.
41. Kang M. и др. Harnessing the power of an expanded genetic code toward next–generation biopharmaceuticals // Curr Opin Chem Biol. 2018. Т. 46. С. 123–129.
42. Wendisch V. F. Microbial production of amino acids and derived chemicals: synthetic biology approaches to strain development // Curr Opin Biotechnol. 2014. Т. 30. С. 51–58.
43. Koketsu K., Mitsuhashi S., Tabata K. Identification of homophenylalanine biosynthetic genes from the cyanobacterium Nostoc punctiforme PCC73102 and application to its microbial production by Escherichia coli // Appl Environ Microbiol. 2013. Т. 79, № 7. С. 2201–2208.

44. Kim D., Sung D., Lee J. W. Expanding the genetic code: Strategies for noncanonical amino acid incorporation in biopolymer // Bioresour Technol. 2025. T. 432. C. 132691.
45. Ohuchi M., Murakami H., Suga H. The flexizyme system: a highly flexible tRNA aminoacylation tool for the translation apparatus // Curr Opin Chem Biol. 2007. T. 11, № 5. C. 537–542.
46. Bag S. S., Saraogi I., Guo J. Editorial: Expansion of the Genetic Code: Unnatural Amino Acids and their Applications // Front Chem. 2022. T. 10. C. 958433.
47. Yazawa K., Numata K. Recent advances in chemoenzymatic peptide syntheses // Molecules. Basel, Switzerland, 2014. T. 19, № 9. C. 13755–13774.
48. Iacovelli R., Bovenberg R. A. L., Driessen A. J. M. Nonribosomal peptide synthetases and their biotechnological potential in Penicillium rubens // J Ind Microbiol Biotechnol. 2021. T. 48, № 9–10. C. kuab045.
49. Bozhüyük K. A. J. и др. Evolution–inspired engineering of nonribosomal peptide synthetases // Science. New York, N.Y., 2024. T. 383, № 6689. C. eadg4320.
50. Kudo F., Miyanaga A., Eguchi T. Structural basis of the nonribosomal codes for nonproteinogenic amino acid selective adenylation enzymes in the biosynthesis of natural products // J Ind Microbiol Biotechnol. 2019. T. 46, № 3–4. C. 515–536.
51. Schmidt M., Nuijens T. Chemoenzymatic Synthesis of Linear– and Head–to–Tail Cyclic Peptides Using Omniligase–1 // Methods Mol Biol. Clifton, N.J., 2019. T. 2012. C. 43–61.
52. Rizzo S. и др. A Sustainable Chemo–Enzymatic Approach to the Synthesis of Liraglutide // ACS Sustainable Chem. Eng. 2024. T. 12, № 45. C. 16791–16799.
53. van der Gulik P. T. S., Speijer D. How Amino Acids and Peptides Shaped the RNA World // Life (Basel). 2015. T. 5, № 1. C. 230–246.
54. Wetzel R. Evolution of the aminoacyl–tRNA synthetases and the origin of the genetic code // J Mol Evol. 1995. T. 40, № 5. C. 545–550.
55. Ilardo M. и др. Adaptive Properties of the Genetically Encoded Amino Acid Alphabet Are Inherited from Its Subsets // Sci Rep. 2019. T. 9. C. 12468.
56. Lei L., Burton Z. F. Evolution of Life on Earth: tRNA, Aminoacyl–tRNA Synthetases and the Genetic Code // Life (Basel). 2020. T. 10, № 3. C. 21.
57. Xu X.–M. и др. Targeted insertion of cysteine by decoding UGA codons with mammalian selenocysteine machinery // Proceedings of the National Academy of Sciences. Proceedings of the National Academy of Sciences, 2010. T. 107, № 50. C. 21430–21434.
58. Srinivasan G., James C. M., Krzycki J. A. Pyrrolysine encoded by UAG in Archaea: charging of a UAG–decoding specialized tRNA // Science. 2002. T. 296, № 5572. C. 1459–1462.
59. de la Torre D., Chin J. W. Reprogramming the genetic code // Nat Rev Genet. 2021. T. 22, № 3. C. 169–184.
60. Italia J. S. и др. Resurrecting the Bacterial Tyrosyl–tRNA Synthetase/tRNA Pair for Expanding the Genetic Code of Both E. coli and Eukaryotes // Cell Chem Biol. 2018. T. 25, № 10. C. 1304–1312.e5.
61. Liu W. и др. Genetic incorporation of unnatural amino acids into proteins in mammalian cells // Nat Methods. 2007. T. 4, № 3. C. 239–244.
62. Johnson D. B. F. и др. RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites // Nat Chem Biol. 2011. T. 7, № 11. C. 779–786.
63. Povolotskaya I. S. и др. Stop codons in bacteria are not selectively equivalent // Biol Direct. 2012. T. 7. C. 30.
64. Lajoie M. J. и др. Genomically recoded organisms expand biological functions // Science. 2013. T. 342, № 6156. C. 357–360.
65. Phelps S. S. и др. Translocation of a tRNA with an extended anticodon through the ribosome // J Mol Biol. 2006. T. 360, № 3. C. 610–622.

66. Xi Z. и др. Using a quadruplet codon to expand the genetic code of an animal // Nucleic Acids Research. 2021. Т. 50, № 9. С. 4801.
67. Mukai T. и др. Rewriting the Genetic Code // Annu Rev Microbiol. 2017. Т. 71. С. 557–577.
68. Mariotti M. и др. SECISearch3 and Seblastian: new tools for prediction of SECIS elements and selenoproteins // Nucleic Acids Res. 2013. Т. 41, № 15. С. e149.
69. Maeder D. L. и др. The Methanosarcina barkeri genome: comparative analysis with Methanosarcina acetivorans and Methanosarcina mazei reveals extensive rearrangement within methanosarcinal genomes // J Bacteriol. 2006. Т. 188, № 22. С. 7922–7931.
70. Wan W., Tharp J. M., Liu W. R. Pyrrolysyl–tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool // Biochim Biophys Acta. 2014. Т. 1844, № 6. С. 1059–1070.
71. Zhang Y., Gladyshev V. N. High content of proteins containing 21st and 22nd amino acids, selenocysteine and pyrrolysine, in a symbiotic deltaproteobacterium of gutless worm Olavius algarvensis // Nucleic Acids Res. 2007. Т. 35, № 15. С. 4952–4963.
72. Yanagisawa T. и др. Crystal Structure of Pyrrolysyl–tRNA Synthetase from a Methanogenic Archaeon ISO4–G1 and Its Structure–Based Engineering for Highly–Productive Cell–Free Genetic Code Expansion with Non–Canonical Amino Acids // International Journal of Molecular Sciences. 2023. Т. 24, № 7. С. 6256.
73. Quitterer F. и др. Crystal structure of methylornithine synthase (PylB): insights into the pyrrolysine biosynthesis // Angew Chem Int Ed Engl. 2012. Т. 51, № 6. С. 1339–1342.
74. Gaston M. A. и др. The complete biosynthesis of the genetically encoded amino acid pyrrolysine from lysine // Nature. 2011. Т. 471, № 7340. С. 647–650.
75. Krzycki J. A. The path of lysine to pyrrolysine // Curr Opin Chem Biol. 2013. Т. 17, № 4. С. 619–625.
76. Nozawa K. и др. Pyrrolysyl–tRNA synthetase–tRNA(Pyl) structure reveals the molecular basis of orthogonality // Nature. 2009. Т. 457, № 7233. С. 1163–1167.
77. Kavran J. M. и др. Structure of pyrrolysyl–tRNA synthetase, an archaeal enzyme for genetic code innovation // Proc Natl Acad Sci U S A. 2007. Т. 104, № 27. С. 11268–11273.
78. Guo L.–T. и др. Ancestral archaea expanded the genetic code with pyrrolysine // J Biol Chem. 2022. Т. 298, № 11. С. 102521.
79. Koch N. G., Budisa N. Evolution of Pyrrolysyl–tRNA Synthetase: From Methanogenesis to Genetic Code Expansion // Chem Rev. 2024. Т. 124, № 16. С. 9580–9608.
80. Borrel G. и др. Unique characteristics of the pyrrolysine system in the 7th order of methanogens: implications for the evolution of a genetic code expansion cassette // Archaea. 2014. Т. 2014. С. 374146.
81. Polycarpo C. R. и др. Pyrrolysine analogues as substrates for pyrrolysyl–tRNA synthetase // FEBS Lett. 2006. Т. 580, № 28–29. С. 6695–6700.
82. Yanagisawa T. и др. Expanded genetic code technologies for incorporating modified lysine at multiple sites // Chembiochem. 2014. Т. 15, № 15. С. 2181–2187.
83. Krahn N. и др. Engineering aminoacyl–tRNA synthetases for use in synthetic biology // Enzymes. 2020. Т. 48. С. 351–395.
84. Englert M. и др. Aminoacylation of tRNA 2'– or 3'–hydroxyl by phosphoseryl– and pyrrolysyl–tRNA synthetases // FEBS Lett. 2013. Т. 587, № 20. С. 3360–3364.
85. Tharp J. M., Ehnbom A., Liu W. R. tRNAPyl: Structure, function, and applications // RNA Biol. 2018. Т. 15, № 4–5. С. 441–452.
86. Yanagisawa T. и др. Crystallographic studies on multiple conformational states of active–site loops in pyrrolysyl–tRNA synthetase // J Mol Biol. 2008. Т. 378, № 3. С. 634–652.
87. Suzuki T. и др. Crystal structures reveal an elusive functional domain of pyrrolysyl–tRNA synthetase // Nature chemical biology. 2017. Т. 13, № 12. С. 1261.

88. Ambrogelly A. и др. Pyrrolysine is not hardwired for cotranslational insertion at UAG codons // Proc Natl Acad Sci U S A. 2007. Т. 104, № 9. С. 3141–3146.
89. Willis J. C. W., Chin J. W. Mutually orthogonal pyrrolysyl–tRNA synthetase/tRNA pairs // Nat Chem. 2018. Т. 10, № 8. С. 831–837.
90. Meng K. и др. Unconventional genetic code systems in archaea // Frontiers in Microbiology. 2022. Т. 13. С. 1007832.
91. Fladischer P. и др. A Semi–Rationally Engineered Bacterial Pyrrolysyl–tRNA Synthetase Genetically Encodes Phenyl Azide Chemistry // Biotechnol J. 2019. Т. 14, № 3. С. e1800125.
92. Yamaguchi A. и др. Pyrrolysyl–tRNA Synthetase with a Unique Architecture Enhances the Availability of Lysine Derivatives in Synthetic Genetic Codes // Molecules. 2018. Т. 23, № 10. С. 2460.
93. Tharp J. M. и др. Genetic Encoding of Three Distinct Noncanonical Amino Acids Using Reprogrammed Initiator and Nonsense Codons // ACS Chem Biol. 2021. Т. 16, № 4. С. 766–774.
94. Borrel G. и др. Phylogenomic Data Support a Seventh Order of Methylotrophic Methanogens and Provide Insights into the Evolution of Methanogenesis // Genome Biology and Evolution. 2013. Т. 5, № 10. С. 1769.
95. Borrel G. и др. Comparative genomics highlights the unique biology of Methanomassiliicoccales, a Thermoplasmatales–related seventh order of methanogenic archaea that encodes pyrrolysine // BMC Genomics. 2014. Т. 15. С. 679.
96. Mihajlovski A., Alric M., Brugère J.–F. A putative new order of methanogenic Archaea inhabiting the human gut, as revealed by molecular analyses of the mcrA gene // Res Microbiol. 2008. Т. 159, № 7–8. С. 516–521.
97. Meineke B. и др. Methanomethylophilus alvus Mx1201 Provides Basis for Mutual Orthogonal Pyrrolysyl tRNA/Aminoacyl–tRNA Synthetase Pairs in Mammalian Cells // ACS Chemical Biology. 2018. Т. 13, № 11. С. 3087.
98. Beránek V., Willis J. C. W., Chin J. W. An Evolved Methanomethylophilus alvus Pyrrolysyl–tRNA Synthetase/tRNA Pair Is Highly Active and Orthogonal in Mammalian Cells // Biochemistry. 2019. Т. 58, № 5. С. 387–390.
99. Cozannet M. и др. New Insights into the Ecology and Physiology of Methanomassiliicoccales from Terrestrial and Aquatic Environments // Microorganisms. 2020. Т. 9, № 1. С. 30.
100. Matassi G. Horizontal gene transfer drives the evolution of Rh50 permeases in prokaryotes // BMC Evol Biol. 2017. Т. 17, № 1. С. 2.
101. Fournier G. P., Huang J., Gogarten J. P. Horizontal gene transfer from extinct and extant lineages: biological innovation and the coral of life // Philos Trans R Soc Lond B Biol Sci. 2009. Т. 364, № 1527. С. 2229–2239.
102. Dunkelmann D. L., Chin J. W. Engineering Pyrrolysine Systems for Genetic Code Expansion and Reprogramming // Chemical Reviews. 2024. Т. 124, № 19. С. 11008.
103. Koch N. G., Budisa N. Focused Engineering of Pyrrolysyl–tRNA Synthetase–Based Orthogonal Translation Systems for the Incorporation of Various Noncanonical Amino Acids // Methods Mol Biol. Clifton, N.J., 2023. Т. 2676. С. 3–19.
104. Wang Y.–S. и др. A rationally designed pyrrolysyl–tRNA synthetase mutant with a broad substrate spectrum // J Am Chem Soc. 2012. Т. 134, № 6. С. 2950–2953.
105. Wang Y.–S. и др. Genetic incorporation of twelve meta–substituted phenylalanine derivatives using a single pyrrolysyl–tRNA synthetase mutant // ACS Chem Biol. 2013. Т. 8, № 2. С. 405–415.
106. Wang Y.–S. и др. The de novo engineering of pyrrolysyl–tRNA synthetase for genetic incorporation of L–phenylalanine and its derivatives // Mol Biosyst. 2011. Т. 7, № 3. С. 714–717.
107. Tharp J. M. и др. Genetic incorporation of seven ortho–substituted phenylalanine derivatives // ACS Chem Biol. 2014. Т. 9, № 4. С. 884–890.

108. Kobayashi T. и др. Recognition of non–alpha–amino substrates by pyrrolysyl–tRNA synthetase // J Mol Biol. 2009. Т. 385, № 5. С. 1352–1360.
109. Neumann H., Peak–Chew S. Y., Chin J. W. Genetically encoding N(epsilon)–acetyllysine in recombinant proteins // Nat Chem Biol. 2008. Т. 4, № 4. С. 232–234.
110. Koch N. G. и др. Engineering Pyrrolysyl–tRNA Synthetase for the Incorporation of Non–Canonical Amino Acids with Smaller Side Chains // Int J Mol Sci. 2021. Т. 22, № 20. С. 11194.
111. Fu C. и др. Genetically Encoding a Lipidated Amino Acid for Extension of Protein Half–Life in vivo // Angew Chem Int Ed Engl. 2019. Т. 58, № 5. С. 1392–1396.
112. Wang N. и др. Genetically Encoding Fluorosulfate–l–tyrosine To React with Lysine, Histidine, and Tyrosine via SuFEx in Proteins in Vivo // J Am Chem Soc. 2018. Т. 140, № 15. С. 4995–4999.
113. Tian Y. и др. Genetically Encoded 2–Aryl–5–carboxytetrazoles for Site–Selective Protein Photo–Cross–Linking // J Am Chem Soc. 2017. Т. 139, № 17. С. 6078–6081.
114. Li S. и др. Genetically encoded chemical crosslinking of carbohydrate // Nat Chem. 2023. Т. 15, № 1. С. 33–42.
115. Ko J. и др. Pyrrolysyl–tRNA synthetase variants reveal ancestral aminoacylation function // FEBS letters. 2013. Т. 587, № 19. С. 3243.
116. Crnković A. и др. Pyrrolysyl–tRNA synthetase, an aminoacyl–tRNA synthetase for genetic code expansion // Croatica chemica acta. Arhiv za kemiju. 2016. Т. 89, № 2. С. 163.
117. Takimoto J. K. и др. Stereochemical basis for engineered pyrrolysyl–tRNA synthetase and the efficient in vivo incorporation of structurally divergent non–native amino acids // ACS Chem Biol. 2011. Т. 6, № 7. С. 733–743.
118. Gong X. и др. Update of the Pyrrolysyl–tRNA Synthetase/tRNAPyl Pair and Derivatives for Genetic Code Expansion // J Bacteriol. 2023. Т. 205, № 2. С. e0038522.
119. Yanagisawa T. и др. Multistep engineering of pyrrolysyl–tRNA synthetase to genetically encode N(epsilon)–(o–azidobenzyloxycarbonyl) lysine for site–specific protein modification // Chem Biol. 2008. Т. 15, № 11. С. 1187–1197.
120. Yanagisawa T. и др. Structural Basis for Genetic–Code Expansion with Bulky Lysine Derivatives by an Engineered Pyrrolysyl–tRNA Synthetase // Cell Chem Biol. 2019. Т. 26, № 7. С. 936–949.e13.
121. Schmidt M. J. и др. Structural basis of furan–amino acid recognition by a polyspecific aminoacyl–tRNA–synthetase and its genetic encoding in human cells // Chembiochem. 2014. Т. 15, № 12. С. 1755–1760.
122. Umehara T. и др. N–acetyl lysyl–tRNA synthetases evolved by a CcdB–based selection possess N–acetyl lysine specificity in vitro and in vivo // FEBS Lett. 2012. Т. 586, № 6. С. 729–733.
123. Passioura T., Suga H. Reprogramming the genetic code in vitro // Trends Biochem Sci. 2014. Т. 39, № 9. С. 400–408.
124. Zhang Q. и др. Machine learning–guided evolution of pyrrolysyl–tRNA synthetase for improved incorporation efficiency of diverse noncanonical amino acids // Nat Commun. 2025. Т. 16, № 1. С. 6648.
125. Jiang R., Krzycki J. A. PylSn and the homologous N–terminal domain of pyrrolysyl–tRNA synthetase bind the tRNA that is essential for the genetic encoding of pyrrolysine // J Biol Chem. 2012. Т. 287, № 39. С. 32738–32746.
126. Jiang H.–K. и др. Linker and N–Terminal Domain Engineering of Pyrrolysyl–tRNA Synthetase for Substrate Range Shifting and Activity Enhancement // Front Bioeng Biotechnol. 2020. Т. 8. С. 235.

127. Cho C.–C. D. и др. Pyrrolysyl–tRNA Synthetase Activity can be Improved by a P188 Mutation that Stabilizes the Full–Length Enzyme // Journal of molecular biology. 2022. Т. 434, № 8. С. 167453.
128. Mukai T. и др. Reassignment of a rare sense codon to a non–canonical amino acid in Escherichia coli // Nucleic Acids Res. 2015. Т. 43, № 16. С. 8111–8122.
129. Sharma V. и др. Evolving the N–Terminal Domain of Pyrrolysyl–tRNA Synthetase for Improved Incorporation of Noncanonical Amino Acids // Chembiochem. 2018. Т. 19, № 1. С. 26–30.
130. Guo L.–T. и др. Polyspecific pyrrolysyl–tRNA synthetases from directed evolution // Proceedings of the National Academy of Sciences. Proceedings of the National Academy of Sciences, 2014. Т. 111, № 47. С. 16724–16729.
131. Schmidt M. J., Summerer D. Directed Evolution of Orthogonal Pyrrolysyl–tRNA Synthetases in Escherichia coli for the Genetic Encoding of Noncanonical Amino Acids // Methods Mol Biol. Clifton, N.J., 2018. Т. 1728. С. 97–111.
132. Bryson D. I. и др. Continuous directed evolution of aminoacyl–tRNA synthetases // Nat Chem Biol. 2017. Т. 13, № 12. С. 1253–1260.
133. Dalby P. A. Strategy and success for the directed evolution of enzymes // Curr Opin Struct Biol. 2011. Т. 21, № 4. С. 473–480.
134. Miller S. M., Wang T., Liu D. R. Phage–assisted continuous and non–continuous evolution // Nat Protoc. Nature Publishing Group, 2020. Т. 15, № 12. С. 4101–4127.
135. Hay I. D., Lithgow T. Filamentous phages: masters of a microbial sharing economy // EMBO Rep. 2019. Т. 20, № 6. С. e47427.
136. Esvelt K. M., Carlson J. C., Liu D. R. A system for the continuous directed evolution of biomolecules // Nature. Nature Publishing Group, 2011. Т. 472, № 7344. С. 499–503.
137. Fijalkowska I. J., Schaaper R. M. Mutants in the Exo I motif of Escherichia coli dnaQ: defective proofreading and inviability due to error catastrophe // Proc Natl Acad Sci U S A. 1996. Т. 93, № 7. С. 2856–2861.
138. Horton J. R. и др. Structure and substrate recognition of the Escherichia coli DNA adenine methyltransferase // J Mol Biol. 2006. Т. 358, № 2. С. 559–570.
139. Lu M. и др. SeqA: a negative modulator of replication initiation in E. coli // Cell. 1994. Т. 77, № 3. С. 413–426.
140. Kremsner P. G. и др. Safety and immunogenicity of an mRNA–lipid nanoparticle vaccine candidate against SARS–CoV–2 : A phase 1 randomized clinical trial // Wien Klin Wochenschr. 2021. Т. 133, № 17–18. С. 931–941.
141. He W. и др. Diverse nanoparticles deliver mRNA to enhance tumor immunotherapy // BMB Rep. 2025. Т. 58, № 3. С. 124–132.
142. Sivamaruthi B. S. и др. A Review of Nanoparticle–Mediated Delivery Systems for RNA–Based Therapeutics // Recent Pat Nanotechnol. 2025.
143. Sasaki K. и др. mRNA–Loaded Lipid Nanoparticles Targeting Dendritic Cells for Cancer Immunotherapy // Pharmaceutics. 2022. Т. 14, № 8. С. 1572.
144. Li X. и др. Nanoparticle technology for mRNA: Delivery strategy, clinical application and developmental landscape // Theranostics. 2024. Т. 14, № 2. С. 738–760.
145. Cullis P. R., Hope M. J. Lipid Nanoparticle Systems for Enabling Gene Therapies // Mol Ther. 2017. Т. 25, № 7. С. 1467–1475.
146. Pateev I. и др. Biodistribution of RNA Vaccines and of Their Products: Evidence from Human and Animal Studies // Biomedicines. 2023. Т. 12, № 1. С. 59.
147. Cheng Q. и др. Selective organ targeting (SORT) nanoparticles for tissue–specific mRNA delivery and CRISPR–Cas gene editing // Nat Nanotechnol. 2020. Т. 15, № 4. С. 313–320.

148. Vasileva O. и др. Composition of lipid nanoparticles for targeted delivery: application to mRNA therapeutics // Front Pharmacol. 2024. Т. 15. С. 1466337.
149. Mitchell M. J. и др. Engineering precision nanoparticles for drug delivery // Nat Rev Drug Discov. 2021. Т. 20, № 2. С. 101–124.
150. Marques A. C. и др. Functionalizing nanoparticles with cancer–targeting antibodies: A comparison of strategies // Journal of Controlled Release. 2020. Т. 320. С. 180–200.
151. Lin Y., Cheng Q., Wei T. Surface engineering of lipid nanoparticles: targeted nucleic acid delivery and beyond // Biophys Rep. 2023. Т. 9, № 5. С. 255–278.
152. Walsh S. J. и др. Site–selective modification strategies in antibody–drug conjugates // Chem Soc Rev. 2021. Т. 50, № 2. С. 1305–1353.
153. Suleiman E. и др. Conjugation of Native–Like HIV–1 Envelope Trimers onto Liposomes Using EDC/Sulfo–NHS Chemistry: Requirements and Limitations // Pharmaceutics. 2020. Т. 12, № 10. С. 979.
154. Xia Y., Sun J., Liang D. Aggregation, fusion, and leakage of liposomes induced by peptides // Langmuir. 2014. Т. 30, № 25. С. 7334–7342.
155. Jans H. и др. Impact of pre–concentration to covalently biofunctionalize suspended nanoparticles // Nanotechnology. 2010. Т. 21, № 34. С. 345102.
156. Totaro K. A. и др. Systematic Investigation of EDC/sNHS–Mediated Bioconjugation Reactions for Carboxylated Peptide Substrates // Bioconjug Chem. 2016. Т. 27, № 4. С. 994–1004.
157. Zaleski M. H. и др. Conjugation Chemistry Markedly Impacts Toxicity and Biodistribution of Targeted Nanoparticles, Mediated by Complement Activation // Adv Mater. Deerfield Beach, Fla., 2025. Т. 37, № 5. С. e2409945.
158. Baldwin A. D., Kiick K. L. Tunable degradation of maleimide–thiol adducts in reducing environments // Bioconjug Chem. 2011. Т. 22, № 10. С. 1946–1953.
159. O'Brien Laramy M. и др. Chemistry, manufacturing and controls strategies for using novel excipients in lipid nanoparticles // Nat Nanotechnol. 2025. Т. 20, № 3. С. 331–344.
160. Shen B.–Q. и др. Conjugation site modulates the in vivo stability and therapeutic activity of antibody–drug conjugates // Nat Biotechnol. 2012. Т. 30, № 2. С. 184–189.
161. Plückthun A. Designed ankyrin repeat proteins (DARPins): binding proteins for research, diagnostics, and therapy // Annu Rev Pharmacol Toxicol. 2015. Т. 55. С. 489–511.
162. Maruani A. и др. A plug–and–play approach to antibody–based therapeutics via a chemoselective dual click strategy // Nat Commun. 2015. Т. 6. С. 6645.
163. Chudasama V., Maruani A., Caddick S. Recent advances in the construction of antibody–drug conjugates // Nat Chem. 2016. Т. 8, № 2. С. 114–119.
164. Chen M. Z. и др. A versatile antibody capture system drives specific in vivo delivery of mRNA–loaded lipid nanoparticles // Nat Nanotechnol. 2025. Т. 20, № 9. С. 1273–1284.
165. Thuronyi B. W. и др. Continuous evolution of base editors with expanded target compatibility and improved activity // Nat Biotechnol. 2019. Т. 37, № 9. С. 1070–1079.
166. Theuerkauf S. A. и др. ApoE2–DARPin fusion proteins enable selective RNA transfer to CD8 T cells by lipid nanoparticles // J Control Release. 2025. Т. 388, № Pt 2. С. 114377.
167. Cruaud P. и др. High–throughput sequencing of multiple amplicons for barcoding and integrative taxonomy // Sci Rep. 2017. Т. 7. С. 41948.
168. Kozlova A. и др. A Cap–Optimized mRNA Encoding Multiepitope Antigen ESAT6 Induces Robust Cellular and Humoral Immune Responses Against Mycobacterium tuberculosis // Vaccines. Multidisciplinary Digital Publishing Institute, 2024. Т. 12, № 11. С. 1267.
169. Tolkacheva E. V. и др. Novel Small Molecule GLP–1R Agonists Based on 1H–Benzo[d]imidazole–5–Carboxylic Acid Scaffold // Molecules. Basel, Switzerland, 2026. Т. 31, № 7. С. 1129.

170. Толкачева Е. В. и др. Производные 3–(5–((4–(4–(бензилокси)пиримидин–2–ил)пиперазин–1–ил)метил)–4h–1,2,4–триазол–3–ил) бензойной кислоты в качестве низкомолекулярных агонистов рецептора глюкагоноподобного пептида–1: пат. RU2844125C1 USA. 2025.
171. Dakhnevich A. и др. Pyrrolysine Aminoacyl–tRNA Synthetase as a Tool for Expanding the Genetic Code // Int J Mol Sci. 2025. Т. 26, № 2. С. 539.
172. Carlson J. C. и др. Negative selection and stringency modulation in phage–assisted continuous evolution // Nat Chem Biol. 2014. Т. 10, № 3. С. 216–222.
173. Beattie A. T., Dunkelmann D. L., Chin J. W. Quintuply orthogonal pyrrolysyl–tRNA synthetase/tRNAPyl pairs // Nat Chem. 2023. Т. 15, № 7. С. 948–959.
174. Odoi K. A. и др. Nonsense and Sense Suppression Abilities of Original and Derivative Methanosarcina mazei Pyrrolysyl–tRNA Synthetase–tRNAPyl Pairs in the Escherichia coli BL21(DE3) Cell Strain // PLoS One. 2013. Т. 8, № 3. С. e57035.
175. Tseng H.–W. и др. Expanding the Scope of Orthogonal Translation with Pyrrolysyl–tRNA Synthetases Dedicated to Aromatic Amino Acids // Molecules. Multidisciplinary Digital Publishing Institute, 2020. Т. 25, № 19. С. 4418.
176. Eriani G. и др. Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs // Nature. 1990. Т. 347, № 6289. С. 203–206.
177. Nozawa K. и др. Pyrrolysyl–tRNA synthetase–tRNAPyl structure reveals the molecular basis of orthogonality: 7233 // Nature. Nature Publishing Group, 2009. Т. 457, № 7233. С. 1163–1167.
178. Schloßhauer J. L. и др. Enhancing the performance of a mutant pyrrolysyl–tRNA synthetase to create a highly versatile eukaryotic cell–free protein synthesis tool // Sci Rep. Nature Publishing Group, 2023. Т. 13, № 1. С. 15236.
179. Zhao Y. и др. Nanoparticles that modulate immune cells – an important strategy for the future treatment of tumors // Front. Immunol. Frontiers, 2026. Т. 17.
180. Aunins E. A. и др. An Il12 mRNA–LNP adjuvant enhances mRNA vaccine–induced CD8 T cell responses // Sci Immunol. 2025. Т. 10, № 108. С. eads1328.
181. Yum S. и др. DNA delivered by lipid nanoparticles induces CD8+ T cell–dependent antitumor responses and enhances anti–PD–L1 therapy // JCI Insight. 2026. Т. 11, № 6. С. e197404.
182. Hunter T. L. и др. In vivo CAR T cell generation to treat cancer and autoimmune disease // Science. New York, N.Y., 2025. Т. 388, № 6753. С. 1311–1317.
183. Tombácz I. и др. Highly efficient CD4+ T cell targeting and genetic recombination using engineered CD4+ cell–homing mRNA–LNPs // Mol Ther. 2021. Т. 29, № 11. С. 3293–3304.
184. Hartmann J. и др. A Library–Based Screening Strategy for the Identification of DARPins as Ligands for Receptor–Targeted AAV and Lentiviral Vectors // Mol Ther Methods Clin Dev. 2018. Т. 10. С. 128–143.
185. Siegler E. и др. Designed Ankyrin Repeat Proteins as Her2 Targeting Domains in Chimeric Antigen Receptor–Engineered T Cells // Hum Gene Ther. 2017. Т. 28, № 9. С. 726–736.
186. Jang S. и др. Development of a simple method for protein conjugation by copper–free click reaction and its application to antibody–free Western blot analysis // Bioconjug Chem. 2012. Т. 23, № 11. С. 2256–2261.
187. Wang Y. и др. Noncanonical amino acids as doubly bio–orthogonal handles for one–pot preparation of protein multiconjugates // Nat Commun. 2023. Т. 14, № 1. С. 974.
188. Kunyk D. и др. Dual Ionizable Lipids (ALC0315 and SM102) in a Novel Lipid Nanoparticle Formulation Enhance In Vivo Delivery and Time–Dependent Expression of Messenger RNA and Self–Amplifying RNA // ChemMedChem. 2026. Т. 21, № 11. С. e70335.
189. Li X. и др. A versatile method for conjugating lipid nanoparticles on T cells through combination of click chemistry and metabolic glycoengineering // Biotechnol J. 2024. Т. 19, № 1. С. e2300339.

190. Swart L. E. и др. A robust post–insertion method for the preparation of targeted siRNA LNPs // Int J Pharm. 2022. Т. 620. С. 121741.
191. Escalona–Rayo O. и др. In vitro and in vivo evaluation of clinically–approved ionizable cationic lipids shows divergent results between mRNA transfection and vaccine efficacy // Biomed Pharmacother. 2023. Т. 165. С. 115065.
192. Joosten L. и др. Enhanced Specific Activity by Multichelation of Exendin–3 Leads To Improved Image Quality and In Vivo Beta Cell Imaging // Mol Pharm. 2018. Т. 15, № 2. С. 486–494.
193. Zheng Z. и др. Glucagon–like peptide–1 receptor: mechanisms and advances in therapy // Signal Transduct Target Ther. 2024. Т. 9. С. 234.
194. Gupta V. Glucagon–like peptide–1 analogues: An overview // Indian J Endocrinol Metab. 2013. Т. 17, № 3. С. 413–421.
195. Fang Z. и др. Ligand–Specific Factors Influencing GLP–1 Receptor Post–Endocytic Trafficking and Degradation in Pancreatic Beta Cells // Int J Mol Sci. 2020. Т. 21, № 21. С. 8404.
196. LaVallie E. R. и др. A thioredoxin gene fusion expression system that circumvents inclusion body formation in the E. coli cytoplasm // Biotechnology (N Y). Nature Publishing Company, 1993. Т. 11, № 2. С. 187–193.
197. Jeong H., Kim H. J., Lee S. J. Complete Genome Sequence of Escherichia coli Strain BL21 // Genome Announc. 2015. Т. 3, № 2. С. e00134–15.
198. Studier F. W. T7 Expression Systems for Inducible Production of Proteins from Cloned Genes in E. coli // Curr Protoc Mol Biol. 2018. Т. 124, № 1. С. e63.
199. Miroux B., Walker J. E. Over–production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels // J Mol Biol. 1996. Т. 260, № 3. С. 289–298.
200. Yang X. и др. The Global Status and Trends of Enteropeptidase: A Bibliometric Study // Front Med (Lausanne). 2022. Т. 9. С. 779722.
201. Light A., Janska H. Enterokinase (enteropeptidase): comparative aspects // Trends in Biochemical Sciences. 1989. Т. 14, № 3. С. 110–112.
202. Yin X. и др. Expression and purification of exendin–4, a GLP–1 receptor agonist, in Escherichia coli // Protein Expr Purif. 2005. Т. 41, № 2. С. 259–265.
203. Huang L. и др. Functional expression and purification of bovine enterokinase light chain in recombinant Escherichia coli // Prep Biochem Biotechnol. 2007. Т. 37, № 3. С. 205–217.
204. Kim Y. S. и др. Expression and purification of soluble and active human enterokinase light chain in Escherichia coli // Biotechnol Rep (Amst). Amsterdam, Netherlands, 2021. Т. 30. С. e00626.
205. Ganjave S. D. и др. Effect of Thioredoxin Tag, Oxidizing Environment, and Temperature on the Global Metabolome of E. Coli Strains // Appl Biochem Biotechnol. 2026. Т. 198, № 1. С. 111–128.
206. Schenkel M. и др. Heat treatment of thioredoxin fusions increases the purity of α‐helical transmembrane protein constructs // Protein Sci. 2021. Т. 30, № 9. С. 1974–1982.
207. Fathi–Roudsari M., Akhavian–Tehrani A., Maghsoudi N. Comparison of Three Escherichia coli Strains in Recombinant Production of Reteplase // Avicenna J Med Biotechnol. 2016. Т. 8, № 1. С. 16–22.
208. Ishii N. GroEL and the GroEL–GroES Complex // Subcell Biochem. 2017. Т. 83. С. 483–504.
209. Nishihara K. и др. Chaperone coexpression plasmids: differential and synergistic roles of DnaK–DnaJ–GrpE and GroEL–GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in Escherichia coli // Appl Environ Microbiol. 1998. Т. 64, № 5. С. 1694–1699.

210. Thompson A. и др. Molecular Characterisation of Small Molecule Agonists Effect on the Human Glucagon Like Peptide–1 Receptor Internalisation // PLoS One. 2016. Т. 11, № 4. С. e0154229.
211. Skvortsov I. A. и др. Monomeric and Dimeric Boron (III) Subphthalocyanines Functionalized with 4–Hydroxy–Benzoic Acid as Potential Photosensitizers and Photocatalysts in Sulfoxidation // ChemPlusChem. John Wiley & Sons, Ltd, 2024. Т. 89, № 10. С. e202400319.

# Приложение А

## Среды и растворы

Таблица А.1 – Среды используемые для выращивания прокариотических культур

| **Название** | **Состав** |
|---|---|
| LB | триптон — 10 г/л, дрожжевой экстракт — 10 г/л, хлорид натрия — 5 г/л, агаризованная среда содержала агар — 9 г/л, pH = 8.0 |
| 2xYT | триптон — 16 г/л, дрожжевой экстракт — 10 г/л, хлорид натрия — 5 г/л. Агаризованная среда 2xYT содержала агар — в концентрации 1,5% в случае нижнего агара, и 0,6% в слечае верхнего. pH = 7.0 |
| SOC | триптон — 20 г/л, дрожжевой экстракт — 5,5 г/л, хлорид натрия — 0,5 г/л, хлорид калия — 10 г/л, хлорид магния — 0,95 г/л, сульфат магния — 1,2 г/л, глюкоза — 3,6 г/л, pH 7.0 |

Таблица А.2 – Растворы, используемые для приготовления полиакриамидного геля для анализа белков в глициновой или трициновой буферной системе.

| **Состав** | **Трис–глициновый** | **Трис–трициновый** |
|---|---|---|
| | **Концентрирующий гель (4%), мл** | |
| Вода | 1.63 | 2.74 |
| 30% Acril/bis 29:1 | 0.35 | 0.53 |
| 10% SDS | 0.03 | – |
| Буфер (2.5M Tris–HCL pH 8.8) | – | 0.61 |
| Буфер (0.5M Tris–HCL pH 6.8) | 0.67 | – |

Продолжение таблицы А.2

| | | |
|---|---|---|
| | **Концентрирующий гель (4%), мкл** | |
| 10% APS | 13.33 | 36 |
| TEMED | 2.67 | 4 |
| | | |
| | **Разделяющий гель, мл** | |
| | 16% | 20% |
| Вода | 1.12 | – |
| 30% Acril/bis 29:1 | 2.83 | 5.33 |
| 10% SDS | 0.05 | – |
| Буфер (2.5M Tris–HCL pH 8.8) | – | 3.20 |
| Буфер (1.5M Tris–HCL pH 8.8) | 1.33 | – |
| | **Разделяющий гель, мкл** | |
| 10% APS | 26.67 | 18 |
| TEMED | 2.67 | 4.4 |

Таблица А.3 — Состав буферных растворов для проведения SDS–PAGE

| **Состав** | **Трис–трициновый** |
|---|---|
| Буфер | Трис–боратный ЭДТА (TBE), x1 |
| Трис–(гидроксиметил) аминометан | 3.03г (0.025М) |
| Трицин | 4.5г |
| SDS | 0.5г (0.05%) |
| Деионизированная вода | до 1л |
| **Состав** | **Трис–глициновый** |
| Буфер | Трис–глицин с додецилсульфатом натрия (SDS PAGE), x10 |
| Глицин | 1.92М |
| SDS | 1% |
| Трис | 0.25М |
| pH | 8.8 |

# Приложение Б

## Используемые плазмиды

Таблица Б.1 – Используемые плазмиды. Зелёным цветом выделены конструкции, собранные в рамках данной работы

| **Название** | **Устойчивость** | **Рамка считывания 1** | | **Рамка считывания 2** | |
|---|---|---|---|---|---|
| | | **Промотор** | **Ген** | **Промотор** | **Ген** |
| pMP6C | chlorR | $P_{psp}$ | *dnaQ926, dam, seqA, emrR, ugi, cda1, dam* | $P_c$ | *AraC* |
| pDB021CH(+) | ampR | $P_{T7}$ | *gIII, luxAB* | $P_{proK}$ | $tRNA^{Pyl}$ |
| pDB038 | spcR | $P_{psp}$ | *gIII(P29*)* | $P_{proK}$ | $tRNA^{Pyl}$ |
| pDB038a | spcR | $P_{psp}$ | *gIII(P29*, Y184*), luxAB* | $P_{proK}$ | $tRNA^{Pyl}$ |
| pDB038b | spcR | $P_{psp}$ | *gIII(P29*, P83*, Y184*), luxAB* | $P_{proK}$ | $tRNA^{Pyl}$ |
| pDB016+tRNAPyl | spcR | $P_{T7}$ | *gIII–neg* | $P_{proK}$ | $tRNA^{Pyl}$ |

Продолжение таблицы Б.1

| | | | | | |
|---|---|---|---|---|---|
| pDB007ns2a_neg | specR | $P_{psp}$ | *gIII* | $P_{psp}$ | *T7–RNAP* |
| pHpyl–GlnS | chlorR | GlnS | – | $P_{proK}$ | $tRNA^{Pyl}$ |
| pHpyl–MmPylRS–GlnS | chlorR | GlnS | *MmPylRS_mut* | $P_{proK}$ | $tRNA^{Pyl}$ |
| pHpyl–AllYPylRS–GlnS | chlorR | GlnS | *AllYPylRS* | $P_{proK}$ | $tRNA^{Pyl}$ |
| pHpyl–AzFPylRS–GlnS | chlorR | GlnS | *AzFPylRS* | $P_{proK}$ | $tRNA^{Pyl}$ |
| pHpyl–MethYPylRS–GlnS | chlorR | GlnS | *MethYPylRS* | $P_{proK}$ | $tRNA^{Pyl}$ |
| pHpyl–ChPylRS–GlnS | chlorR | GlnS | *ChPylRS_mut* | $P_{proK}$ | $tRNA^{Pyl}$ |
| pHpyl–2ClFPylRS–GlnS | chlorR | GlnS | *2ClFPylRS* | $P_{proK}$ | $tRNA^{Pyl}$ |
| pET–sfGFP–27TAG | kanR | $P_{T7}$ | *Superfolder GFP (27*)* | – | – |

Продолжение таблицы Б.1

| | | | | | |
|---|---|---|---|---|---|
| SP–MmPylRS | none | PgIII | *Mm PylRS* | | |
| SP–chPylRS mut | none | GlnS | – | $P_{proK}$ | *tRNA*$^{Pyl}$ |
| pET30a__His10_DDDDK_DARPin_mCD8_GG–ncAA–GS | kanR | $P_{T7}$ | anti_mCD8_DARPin_GG–*–GS | – | – |
| pET30a__His10_DDDDK_DARPin_mCD8_GG–C–GS | kanR | $P_{T7}$ | anti_mCD8_DARPin_GG–C–GS | – | – |
| pS_AG_TPL_eGFP_Mod | ampR | $P_{T7}$ | eGFP | – | – |
| pET30(a+)_His8_TRX–A_DDDDK_Exe4_40ncAA | kanR | $P_{T7}$ | HisTag_TRX–A_Exe4_* | – | – |
| pET30(a+)_His8_TRX–A_DDDDK_Exe4 | kanR | $P_{T7}$ | HisTag_TRX–A_Exe4_YK | | |
| pET30(a+)_TRX–A_Enterokinase_DDDDK_His8 | kanR | $P_{T7}$ | TRX–A_Enterokinase_8xHis | – | – |

# Приложение В

## Праймеры для сборки ChPylRS

Таблица В.1 – праймеры для сборки ChPylRS

| Название | Описание | Последовательность |
|---|---|---|
| chPylRS–N–1–to pUC18_F | Сборка фрагмента Mb | 5’–TCCCCTGATTCTGTggatccATGGACAAAAAGCCCTTAGATGTACTAATATCTGCTACCGGC–3’ |
| chPylRS–N–2 | | 5’–ATCTTATGCAGAGTGCCCGTGCGGCTCATCCACAGGCCGGTAGCAGATATTAGTACATCT–3’ |
| chPylRS–N–3 | | 5’–ACGGGCACTCTGCATAAGATTAAACACTACGAAGTTTCCCGCTCCAAAATCTACATTGAG–3’ |
| chPylRS–N–4 | | 5’–CTGCTGTTGTTGACTACCAGGTGGTCACCGCACGCCATCTCAATGTAGATTTTGGAGCGGG–3’ |
| chPylRS–N–5 | | 5’–CACCTGGTAGTCAACAACAGCAGATCCTGTCGTACCGCACGTGCGTTTCGTCATCACAAG–3’ |
| chPylRS–N–6 | | 5’–CTCATCGCTAACACGGCAACGTTTGCACGTCTTACGATACTTGTGATGACGAAACGCACG–3’ |

Продолжение таблицы В.1

| | | |
|---|---|---|
| chPylRS–N–7 | Сборка фрагмента Mb | 5’–GTTGCCGTGTTAGCGATGAGGACATCAATAACTTCTTGACCCGTAGCACCGAAGGTAAAA–3’ |
| chPylRS–N–8 | | 5’–TTACTTGCGGCTCGACACCACTTTCACCTTAACGCTGGTTTTACCTTCGGTGCTACGGGT–3’ |
| chPylRS–N–9 | | 5’–TGGTGTCGAGCCGCAAGTAAtaggagattttcaacATGccgaaatctgttctcgtgcgcc–3’ |
| chPylRS–N–10 | | 5’–ggtagacgctttcgcagaaaccggattttccagcggtttcggcgcacgagaaacagattt–3’ |
| chPylRS–N–11 | | 5’–gtttctgcgaaagcgtctaccgacacctctcgttctgttccgtctccggcgaaatctacc–3’ |
| chPylRS–N–12–to pUC18 | | 5’–TTTCGTCTCTAGAtgcagaggtcggaaccggagagttcggggtagatttcgccggagacg–3’ |
| pylRS(ch)C–to–N_F | Амплификация фрагмента Mm | 5’–gacctctgcaTCTGCTCCGGCGCTGACT–3’ |
| chPylRS–C–1–to pUC18_R | | 5’–TTCGTCTCTAGAGGTtAGAGGTTGGTGCTGATACCGT–3’ |

## Приложение Г

### Последовательности нуклеотидов, аминокислот и праймеры для сборки Эксендина–4

Нуклеотидная последовательность Эксендин–4

*CATGGTGAAGGTACTTTCACGTCCGATCTTTCCAAACAAATGGAAGAGGAAGCAGTCCGTCTTTTTATTGAATGGTTGAAGAATGGTGGTCCGTCGAGCGGCGCTCCGCCACCGTCTTATAAATAA*

Аминокислотная последовательность Эксендин–4

*HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPSYK**

Таблица Г.1 – Праймеры для сборки фрагментов гена *Эксендин–4*

| **Название** | **Прямой** | **Обратный** |
|---|---|---|
| 1130–Exendin_4_2_F(223) | GATGATGATGATAAACATGGTGAAGGTACTTTCACGTCCG | AAAAGACGGACTGCTTCCTCTTCCATTTGTTTGGAAAGATCGGACGTGAAAGTAC |
| Exendin_4_2_R(224)–1112 | AAGAGGAAGCAGTCCGTCTTTTTATTGAATGGTTGAAGAATGGTGGTCCGTCGAGCG | GTTAGCAGCCGGTTatttataAGACGGTGGCGGAGCGCCGCTCG |

## Приложение Д

### Последовательности нуклеотидов, аминокислот и праймеры для сборки hEKL

Нуклеотидная последовательность hEKL

*ATTGTGGGCGGTTCTaacgcgaaagaaggcgcatggccgtgggtggtcggcctgtattacggcggtcgcctgctttgcgg cgcatctctggtcagtagcgattggctggtgtccgctgcacattgtgtgtacggccgtaacctggaaccatccaaatggaccgcgattctgggtct gcatatgaagagtaatctgacttctccgcaaaccgtgccgcgtctgattgatgaaattgtgattaacccgcattataaccgccgccgtaaagata atgatattgccatgatgcacctggaatttaaggttaattacaccgattatatccagccgatcagcttgcctgaagagaaccaggtcttcccacctg gtcgcaactgctccattgcaggctggggcaccgtggtttatcagggtaccaccgccaacattctgcaggaggcagacgtgccgctgttgtctaac gagcgctgtcagcaacagatgccggagtataatattaccgaaaatatgatttgtgctggttacgaagaaggcggtattgatagttgccaaggcg actccggcggtccgttaatgtgccaagaaaacaatcgttggtttttagccggcgttaccagttttggttacaaatgcgccttacctaaccgtccagg tgtgtacgcccgcgttagccgctttaCCGAATGGATTCAGTCCTTTCTGCAC*

Аминокислотная последовательность hEKL

*IVGGSNAKEGAWPWVVGLYYGGRLLCGASLVSSDWLVSAAHCVYGRNLEPSKWTAILGLHM KSNLTSPQTVPRLIDEIVINPHYNRRRKDNDIAMMHLEFKVNYTDYIQPISLPEENQVFPPGRNCSIA GWGTVVYQGTTANILQEADVPLLSNERCQQQMPEYNITENMICAGYEEGGIDSCQGDSGGPLMCQ ENNRWFLAGVTSFGYKCALPNRPGVYARVSRFTEWIQSFLH**

Таблица Д.1 Праймеры для наработки фрагментов гена *hEKL*

| **Название** | **Прямой** | **Обратный** |
|---|---|---|
| 1136_hEKL_TrxA – 1.2 1_hEKL | GATGATGATAAAATTGTG GGCGGTTCTaacgcgaaagaag gcgca | cgccaaaatacaggccgaccacccac ggccatgcgccttctttcgcgtt |
| 2_hEKL – 3_hEKL | tcggcctgtattttggcggtcgcctgcttt gcggcgcatctctggtcag | cgtacacacaatgtgcagcggacacca gccaatcgctactgaccagagatgcgc cg |
| 4_hEKL – 5_hEKL | cgctgcacattgtgtgtacggccgtaac ctggaaccatccaaatggaccgcg | gcggagaagtcagattactcttcatatgc agacccagaatcgcggtccatttggatg g |

Продолжение таблицы Г.1

| | | |
|---|---|---|
| 6_hEKL – 7_hEKL | gaagagtaatctgacttctccgcaaaccgtgccgcgtctgattgatgaaattgtg | tatctttacggcggcggttataatgcgggttaatcacaatttcatcaatcagacgcg |
| 8_hEKL – 9_hEKL | taaccgccgccgtaaagataatgatattgccatgatgcacctggaatttaaggttaat | caggcaagctgatcggctggatataatcggtgtaattaaccttaaattccaggtgcatc |
| 10_hEKL – 11_hEKL | agccgatcagcttgcctgaagagaaccaggtcttcccacctggtcgcaactg | gataaaccacggtgccccagcctgcaatggagcagttgcgaccaggtggg |
| 12_hEKL – 13_hEKL | ggggcaccgtggtttatcagggtaccaccgccaacattctgcaggagg | gcgctcgttagacaacagcggcacgtctgcctcctgcagaatgttggc |
| 14_hEKL – 15_hEKL | gctgttgtctaacgagcgctgtcagcaacagatgccggagtataatattaccg | caataccgccttcttcgtaaccagcacaaatcatattttcggtaatattatactccggcatc |
| 16_hEKL – 17_hEKL | ggttacgaagaaggcggtattgatagttgccaaggcgactccggcggtccgtt | cggctaaaaaccaacgattgttttcttggcacattaacggaccgccggagtc |
| 18_hEKL – 19_hEKL | caatcgttggtttttagccggcgttaccagttttggttacaaatgcgccttacc | GGGCGTACACACCTGGACGGTTAGGTAAGGCGCATTtGTAACCAAAAc |
| 20_hEKL – 1138_hEKL_TrxA | GTCCAGGTGTGTACGCCCGCGTTagcCGCTTTACCGAATGGATTCAGTCCT | gTGGTGGTGGTGGTGGTGGTGCAGAAAGGACTGAATCCATTCGG |
| 1139_TrxA_hEKL – 1138_hEKL_TrxA | GCACCACCACCACCACCACCACTGAGA | gTGGTGGTGGTGGTGGTGGTGCAGAAAGGACTGAATCCATTCGG |
| 1137_hEKL_TrxA – 1138_hEKL_TrxA | CCTGGCCGGTTCTGGTTCTGGCGATGATGATGATAAA | gTGGTGGTGGTGGTGGTGGTGCAGAAAGGACTGAATCCATTCGG |

**Приложение Е**

**Последовательности нуклеотидов, аминокислот и праймеры для сборки DARPin mCD8**

Нуклеотидная последовательность DARPin mCD8

*GATCTGGGTAAGAAACTGCTGGATGCTGCGAATGCCGGTCAGGATGATGAAGTGCGCAT CCTGATGGCGAATGGCGTGGATGTGAACGCGAGTGATAATTGGGGCTACACGCCGCTTCATTTA GCCGCCGATACCGGACATTTGGAAATCGTTGATGTTCTGTTGGCGCACGGGGCCGATGTAAACG CATCCAACTGGTACGGCACCACGCCCCTGCACAGTGCGGCTCAGAGTGGGCATCTGGAGATTG TGGACGTTTTGTTAGCGAACGGTGCAGACGTGAACGCATCGAACTCCAACGGGGAAACGCCAC TGCACTCTGCTGCAAATCGCGGCCATCTGGAGATTGTGGATGTGTTACTGGCGCATGGCGCAGA TGTAAACGCAAACACAACCAGCGGCAAAACGCCGTTCGATCTTGCAATTGATAATGGCAATGAAG ATATTGCGGAAGTCCTGCAGAAAGCGGCCGGAGGATGCGGCAGT*

Аминокислотная последовательность DARPin mCD8

*DLGKKLLDAANAGQDDEVRILMANGVDVNASDNWGYTPLHLAADTGHLEIVDVLLAHGAD VNASNWYGTTPLHSAAQSGHLEIVDVLLANGADVNASNSNGETPLHSAANRGHLEIVDVLLAHGAD VNANTTSGKTPFDLAIDNGNEDIAEVLQKAAGGCGS**

Таблица Е.1 Праймеры для наработки фрагментов гена *DARPin mCD8*

| **Название** | **Нуклеотидная последовательность** |
|---|---|
| CD8_DARPin_1 | CTGGTTCTGGCGATGATGATGATAAAGATC TGGGTAAGAAACTGCTGGATGCTGCGAAT GC |
| CD8_DARPin_2 | CCATCAGGATGCGCACTTCATCATCCTGAC CGGCATTCGCAGCATCCAGC |
| CD8_DARPin_3 | GAAGTGCGCATCCTGATGGCGAATGGCGT GGATGTGAACGCGAGTGATAATTGGG |
| CD8_DARPin_4 | GGTATCGGCGGCTAAATGAAGCGGCGTGTA GCCCCAATTATCACTCGCGTTCAC |
| CD8_DARPin_5 | GCTTCATTTAGCCGCCGATACCGGACATTT GGAAATCGTTGATGTTCTGTTGGCGCACG |
| CD8_DARPin_6 | GCCGTACCAGTTGGATGCGTTTACATCGGC CCCGTGCGCCAACAGAACATCAA |
| CD8_DARPin_7 | GCATCCAACTGGTACGGCACCACGCCCCTG CACAGTGCGG |

Продолжение таблицы Е.1

| | |
|---|---|
| CD8_DARPin_8 | GCTAACAAAACGTCCACAATCTCCAGATGC CCACTCTGAGCCGCACTGTGCAGGGG |
| CD8_DARPin_9 | GGAGATTGTGGACGTTTTGTTAGCGAACGG TGCAGACGTGAACGCATCGAACTCCAAC |
| CD8_DARPin_10 | CGATTTGCAGCAGAGTGCAGTGGCGTTTCC CCGTTGGAGTTCGATGCGTTCACG |
| CD8_DARPin_11 | CTGCACTCTGCTGCAAATCGCGGCCATCTG GAGATTGTGGATGTGTTACTGGCG |
| CD8_DARPin_12 | CGCTGGTTGTGTTTGCGTTTACATCTGCGC CATGCGCCAGTAACACATCCACAATC |
| CD8_DARPin_13 | CGCAAACACAACCAGCGGCAAAACGCCGT TCGATCTTGCAATTGATAATGGCAATGAAG |
| CD8_DARPin_14 | GGCCGCTTTCTGCAGGACTTCCGCAATATC TTCATTGCCATTATCAATTGCAAGATCG |