

РОССИЙСКАЯ АКАДЕМИЯ НАУК
СИБИРСКОЕ ОТДЕЛЕНИЕ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ НАУКИ
ИНСТИТУТ ХИМИЧЕСКОЙ БИОЛОГИИ И ФУНДАМЕНТАЛЬНОЙ МЕДИЦИНЫ

на правах рукописи

Жданова Полина Викторовна

**Структурные основы узнавания в белково-нуклеиновых комплексах:
масс-спектрометрия с использованием водородно-дейтериевого обмена**

1.5.4 – биохимия

ДИССЕРТАЦИЯ

на соискание ученой степени

кандидата химических наук

Научный руководитель:

к.х.н. Коваль Владимир Васильевич

Новосибирск – 2022

Список сокращений

APCI – химическая ионизация при атмосферном давлении

API – ионизация при атмосферном давлении

APPI – фотоионизация при атмосферном давлении

ECD - диссоциация с захватом электронов

ESI - ионизация электрораспылением

ETD - диссоциация с переносом электронов

HDX - водородно-дейтериевый обмен

HDX-MS – водородно-дейтериевый обмен с последующей масс-спектрометрией

IgG - иммуноглобулины G

LC-MS – жидкостная хроматография с масс-спектрометрией

MS – масс-спектрометрия

NEIL - endonuclease VIII-like

PDB – Protein Data Bank

PIС - коктейль-ингибитор протеазы

PLIMSTEX - protein–ligand interactions by mass spectrometry, titration, and H/D exchange
(взаимодействия белок-лиганд с помощью масс-спектрометрии, титрования и H/D-обмена)

RMSD - среднеквадратическое отклонение

SDS - додецилсульфат натрия

sgРНК - одиночная направляющая РНК

smFRET – резонансный перенос энергии флуоресценции одиночной молекулы

SUPREX – stability of unpurified proteins from rates of H/D exchange (стабильность неочищенных белков по скорости обмена H/D)

ТСЕР - трис(2-карбоксиэтил)фосфин

ВЭЖХ – высокоэффективная жидкостная хроматография

ДТТ - 1,4-дителиотреитол

ИПТГ - изопропил-β-Дтиогалактопиранозид

КД – круговой дихроизм

Крио-ЭМ – криоэлектронная микроскопия

МАЛДИ - матрично-активированная лазерная десорбция/ионизация

РСА – рентгеноструктурный анализ

ТЕМЕД - *N,N,N',N'*-тетраметилэтилендиамин

Трис - трис(гидроксиметил)аминометан

Фурье-ИК – инфракрасная спектроскопия с преобразованием Фурье

ЭДТА – этилендиаминтетраацетат

ЯМР – ядерный магнитный резонанс

Оглавление

1. Введение	5
2. Обзор литературы <i>Метод водородно-дейтериевого обмена с последующим масс-спектрометрическим анализом</i>	7
2.1. <i>История изучения структуры белка</i>	8
2.2. <i>Точность и разнообразие биофизических методов изучения структуры</i>	9
2.3. <i>Развитие метода обмена водорода на дейтерий: от ЯМР к MS</i>	11
2.4. <i>Метод водородно-дейтериевого обмена с последующим масс-спектрометрическим анализом</i>	14
2.5. <i>Преимущества метода HDX-MS</i>	18
2.6. <i>Ограничения метода HDX-MS</i>	19
2.7. <i>Механизмы реакции обмена</i>	19
2.8. <i>Кинетика процессов обмена</i>	20
2.9. <i>Области применения метода HDX-MS</i>	23
2.10. <i>Заключение</i>	29
3. Материалы и методы	30
3.1. Материалы	30
3.1.1. Общие материалы	30
3.1.2. Стандартные буферы и смеси	31
3.2. Методы	31
3.2.1. Выделение плазмидной ДНК	31
3.2.2. Экспрессия, выделение и очистка рекомбинантного белка hNEIL2	32
3.2.3. Электрофорез в системе Лэммли	33

3.2.4.	Эксперименты по водородно-дейтериевому обмену с последующей масс-спектрометрией hNEIL2	34
3.2.5.	Эксперименты по водородно-дейтериевому обмену с последующей масс-спектрометрией SpCas9 и комплекса SpCas9 с sgPHK.....	36
3.2.6.	Компьютерное моделирование	38
4.	Обсуждение результатов;	40
4.1.	<i>hNEIL2</i>	40
4.1.1.	Подготовка к HDX-MS экспериментам.....	41
4.1.2.	HDX-MS анализ динамики hNEIL2	42
4.1.3.	МД-анализ конформации hNEIL2.....	46
4.2.	<i>SpCas9</i> и <i>SpCas9-PHK</i>	52
4.2.1.	Компьютерное моделирование апоCas9 комплекса Cas9- PHK.	53
4.2.2.	HDX-MS анализ динамики SpCas9 и SpCas9-PHK	58
5.	Заключение.....	67
6.	Выводы.....	72
	Список цитированной литературы.....	74

1. Введение

Большинство биологических процессов в живых организмах не проходят без участия белков. Белки - это многогранный класс макромолекул, вовлеченный в каждый клеточный процесс. Белки участвуют в репликации и транскрипции ДНК, в производстве и модификации других белков, а также контролируют процесс деления клеток, метаболизм и передачу информации внутрь клетки и за ее пределы. Представление о процессах, происходящих в живых организмах, требует понимания как функционируют белковые молекулы, как в свободном состоянии, так и во всевозможных комплексах. Следовательно, необходимо знать особенности строения и укладки белков, а также механизмов их действия.

Разрешить структуру, а также исследовать динамические особенности функционирования макромолекул способны методы биофизики. Большинство известных структур белков и белковых комплексов получены с помощью метода рентгеноструктурного анализа (РСА). Популярными методами также являются спектроскопия ядерного магнитного резонанса (ЯМР) и криоэлектронная микроскопия (крио-ЭМ). С развитием масс-спектрометрических методов анализа широкое распространение получил метод водородно-дейтериевого обмена, ранее доступный в основном только для ЯМР-спектроскопии, и сразу зарекомендовал себя отличной альтернативой уже существующим методам.

Цель данной работы заключается в детальном изучении структурных особенностей узнавания в белково-нуклеиновых комплексах методом водородно-дейтериевого обмена метода водородно-дейтериевого обмена с последующим масс-спектрометрическим анализом (HDX-MS). Для достижения цели были поставлены следующие задачи:

1. Исследовать структуру NEIL2 человека методом HDX-MS.
2. Получить и проанализировать теоретические модели белка hNEIL2.

3. Провести сравнение экспериментальных и теоретических данных, тем самым разрешив структуру hNEIL2 в растворе.
4. Исследовать структуру Cas9 и комплекса Cas9 с направляющей РНК методом HDX-MS.
5. Получить и проанализировать теоретические модели Cas9 и его комплекса с РНК, а также тройного комплекса Cas9-РНК-ДНК.
6. Сравнить данные HDX эксперимента с результатами компьютерного моделирования, выделить ключевые различия в конформации Cas9 при формировании комплекса с РНК в растворе.

2. Обзор литературы *Метод водородно-дейтериевого обмена с последующим масс-спектрометрическим анализом*

Одними из важных компонентов всех живых организмов являются белковые молекулы, выполняющие широкий спектр функций. Белки являются комбинацией 20 аминокислот, а их многочисленные биологические функции обусловлены большим разнообразием, вариаций последовательностей и посттрансляционных модификаций. Благодаря эволюционным процессам живых организмов белки приобрели специальные способности, по которым они могут быть разделены по классам. На сегодняшний день выделяют следующие функциональные классы белков: ферменты катализируют различные внутриклеточные и внеклеточные реакции, структурные белки поддерживают структурную жесткость клетки, транспортные белки облегчают трансмембранный поток определенных материалов, регуляторные белки работают как датчики и переключатели для регулирования экспрессии генов и активности белков, моторные белки облегчают макроскопические движения, а сигнальные белки передают сообщения для облегчения коммуникации различных клеточных компонентов. Последовательность аминокислот определяет пространственную структуру белка, а биологическая функция белка и механизмы его действия определяются этой трехмерной структурой.

Структура белков несет на себе печать окружения, в котором они находятся и функционируют. В зависимости от общего строения белка и среды его обитания можно выделить три класса: фибриллярные, мембранные и глобулярные белки. Так, фибриллярные белки образуют массивные агрегаты. Структуры таких молекул обычно высоко регулярны за счет взаимодействия между цепями. Мембранные белки частично находятся в мембране, а частично выступают в воду. Часть белка, находящаяся в мембране, как и фибриллярные белки, имеет регулярную структуру, но ограничивается толщиной мембраны. Глобулярные белки существуют в воде и обладают структурой,

которая обусловлена взаимодействием структурных частей белка между собой, сближением в пространстве далеких по цепи групп и взаимодействием с кофакторами.

Поскольку функциональные механизмы действия белков закодированы в трехмерной структуре, то понимание структуры белковых молекул является одним из ключевых вопросов биологии. Изучение структуры белков и макромолекулярных комплексов позволяет приоткрыть особенности механизмов взаимодействия белков с субстратами, белок-белковое взаимодействие и т.д.

2.1. История изучения структуры белка

Первая рентгеновская фотография белковой структуры, кристаллического пепсина, была получена в 1934 году британским кристаллографом Джоном Берналом и кристаллографом Дороти Кроуфут (Ходжкин) [1].

Значительный прогресс в понимании структуры биомолекул пришел в период 50-60-х годов XX века. В 1958 году Дж. Кендрию была расшифрована с разрешением 6 Å структура миоглобина кашалота [2], а в 1960 году М. Перутцем была представлена структура гемоглобина с разрешением 5,5 Å [3]. В 1962 году авторы получили Нобелевскую премию по химии за исследование глобулярных белков. Данные исследования стали важным поворотом и открыли новые перспективы в исследовании высшей структуры белков.

Со времен пионерских работ рентгеноструктурный анализ не теряет лидирующих позиций, а прогресс в определении структур биомолекул только ускоряется, а подавляющее большинство структур в базах данных получены с помощью метода РСА.

Важную роль в структурном анализе белков сыграл метод спектроскопии ядерного магнитного резонанса. Первопроходцами в данной области являются Рихард Эрнст и Курт Вютрих, а также Эд Бакс, Мариус Клор, Анджела Гроненборн. Нобелевская премия по

химии 1991 г. была присуждена профессору Ричарду Р. Эрнсту за важные методологические разработки в области спектроскопии ЯМР. Благодаря разработкам Эрнста [4–6] произошел значительный рост чувствительности и разрешения метода. В свою очередь Курт Вютрих показал применимость метода ЯМР к биологическим молекулам [7]. В 2002 году Вютрих стал Нобелевским лауреатом по химии, разделив премию с Джоном Фенном и Коичи Танака за успешное применение методов к биологическим молекулам. В 1990 году Эдриан Бакс с коллегами описали новый подход для изучения белков [8]. Авторы изучали белок кальмодулин посредством последовательного отнесения основных резонансов ^1H , ^{13}C и ^{15}N . Мариус Клор сыграл ключевую роль в развитии трех- и четырехмерной ЯМР-спектроскопии [9], в определение структуры больших белковых комплексов методом ЯМР в растворе [10], а также проанализировал и охарактеризовал динамику белка с помощью ЯМР [11].

В последние десятилетия благодаря разработке различных стратегий мечения, развития методов сбора и анализа данных, метод ЯМР-спектроскопии позволяет анализировать структуру и динамику белков в условиях, близких к нативным, как для водорастворимых, так и для мембранных белков. Кроме того, метод ЯМР широко востребован для изучения взаимодействий между различными молекулами, для получения информации о скорости химических реакций и для решения многих других задач.

2.2. Точность и разнообразие биофизических методов изучения структуры

На ряду с РСА и ЯМР существует большое разнообразие методов изучения структурных особенностей белковых молекул и их комплексов. Например, криоэлектронная микроскопия, резонансный перенос энергии флуоресценции одиночной молекулы (smFRET), круговой дихроизм (КД) и др.

Методы различаются точностью с которой разрешается структура молекулы (рис.1). Точность, близкую к атомарной ($<4 \text{ \AA}$) может обеспечить рентгеноструктурный анализ и

крио-ЭМ. Общую характеристику вторичной структуры белка могут предоставить такие методы как КД-спектроскопия, инфракрасная спектроскопия с преобразованием Фурье (FT-IR, Фурье-ИК), резонансная спектроскопия комбинационного рассеяния.

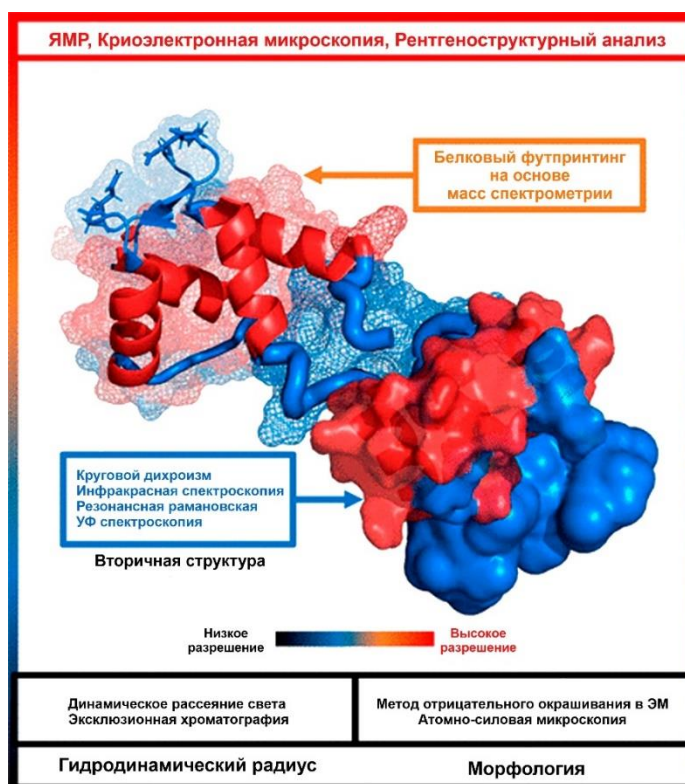


Рисунок 1. Пространственное разрешение различных биофизических методов анализа, показанных на примере структуры белка кальмодулина (PDB ID: 1CDF). Рисунок адаптирован из [12].

Любой метод исследования имеет свои преимущества и ограничения. Так, при исследовании методом крио-ЭМ анализируемый объект быстро замораживается при криогенных температурах (ниже 120 К). При таких температурах молекулы воды не образуют кристаллов и не повреждают структуру белка, а также минимизируются повреждения от пучка электронов. Данный метод используется для характеристики

белков, белковых комплексов и мембранных белков. Предпочтительными объектами исследования являются белки и комплексы с массами больше 100 кДа [13].

Несмотря на то, что существенное количество структур получено с помощью рентгеноструктурного анализа, метод имеет ряд серьезных ограничений. В первую очередь, РСА требует кристаллизации белка, что осуществимо не для всех белков. Рентгеновским лучам «трудно» увидеть атомы водорода, которые очень важны для биологических функций. Биологические молекулы функционируют в жидкой среде, в то время как при кристаллизации молекула принимает определенную конформацию.

ЯМР-спектроскопия предоставляет информацию о динамике белка в растворе, а для нерастворимых белков разработан твердотельный ЯМР. Методы внутриклеточного ЯМР обеспечивают самое высокое разрешение для определения структуры белка, но у этого метода низкая чувствительности и существенным недостатком является относительно большое количество используемого образца (несколько миллиграмм). Наряду с этим сбор данных и их анализ трудоемкий и время затратный процесс. Метод ЯМР отлично проявил себя для малых белков, однако метод может быть адаптирован для объектов с массой до 1 МДа [14].

Метод масс-спектрометрического анализа, аналогично ЯМР, позволяет исследовать структуру белка в растворе. Метод MS имеет более низкое разрешение, речь идет не об атомах, а более высших порядках организации. Однако данный метод анализа имеет значительно меньшее количество ограничений, балансируя между пространственным разрешением и производительностью.

2.3. Развитие метода обмена водорода на дейтерий: от ЯМР к MS

Феномен обмена водорода на дейтерий (HDX) впервые был описан в 1950-х годах Каем Ульриком Линнерстрём-Лангом и его сотрудниками из Калсбергской лаборатории (Дания, Копенгаген). Линнерстрём-Ланг провел серию новаторских исследований

водородного обмена на различных белках и полипептидах [15]. Очищенный и высушенный образец белка растворяли в D₂O и инкубировали при 37°C и pH 3, в определенные моменты времени реакция обмена останавливалась путем замораживания образца до -80°C. Избыток воды удаляли путем лиофилизации и сушкой в вакууме над P₂O₅ в течение нескольких дней. Содержание дейтерия определяли путем обратной реакции с водой. Полученное содержание дейтерия в воде определяли по плотности в градиентной трубке.

С появлением и распространением метода ЯМР-спектроскопии исследователи стали его активно использовать для анализа обмена водорода на дейтерий в белках. HDX стал инструментом для анализа сворачивания белка с варьированием таких параметров, как температура, pH и использование денатурирующих агентов. Благодаря HDX ЯМР были идентифицированы глобальные, субглобальные и локальные события разворачивания и сыграли решающую роль в развитии теории фолдинга белков [16].

Во времена бурного развития методов ЯМР в масс-спектрометрическом анализе использовался в основном метод ионизации электронным ударом. Этим способом была возможность исследовать только малые органические молекулы. В 1966 году был впервые представлен метод химической ионизации [17], а впервые в 1972 году Дональд Хант комбинировал этот подход с водородно-дейтериевым обменом [18]. Фрейзер с коллегами сообщили об анализе последовательных реакциях обмена дейтерия в газовой фазе методом масс-спектрометрии ионно-циклотронного резонанса. Позже, в 1992 году Смит с соавторами сообщили о методе качественного исследования структуры белковых молекулярных ионов в газовой фазе с помощью газофазных реакций для бычьего проинсулина и α -лактальбумина [19].

Изотопный обмен для масс-спектрометрии приобрел большую популярность после разработки методов мягкой ионизации. Так, в 1988 году Танакой был описан метод

матрично-активированной лазерной десорбции/ионизации [20], а Фенном – электроспреей [21]. Именно эти методы открыли новые возможности применения водородно-дейтериевого обмена для масс-спектрометрического анализа макромолекул в растворах. Активно развивающееся направление привлекло внимание многих ученых и возрос интерес к исследованиям посвященных использованию водородно-дейтериевого обмена для изучения конформаций белков, межбелковых взаимодействий и переходов между промежуточными продуктами сворачивания белков.

Разработка источников ионизации при атмосферном давлении (API) показало возможность проведения реакции обмена H/D непосредственно во время ионизации при атмосферном давлении. Первые такие эксперименты были проведены с использованием электрораспыления [22] и термораспыления [23]. Затем реакция обмена водорода на дейтерий в процессе ионизации была продемонстрирована многими другими исследователями для электроспрея, химической ионизации при атмосферном давлении (APCI)[24], фотоионизации при атмосферном давлении (APPI) [25], и других типов источников ионов [26].

Наряду с усовершенствованием методов ионизации, важным шагом стала классическая работа Чжана и Смита [27]. Ученые разработали подход протеолитического расщепления с этапом жидкостной хроматографии и последующей масс-спектрометрией (LC-MS), что стало дополнением к существовавшему методу ЯМР и радиоактивному анализу. Большинство исследований, направленные на пространственное разрешение структуры являются вариациями классического исследования и дают информацию о специфическом поведении регионов внутри белка.

2.4. Метод водородно-дейтериевого обмена с последующим масс-спектрометрическим анализом

Метод водородно-дейтериевого обмена с последующей масс-спектрометрией (HDX-MS) получил распространение только в конце прошлого века и все ещё набирает популярность в области структурной биологии. Как понятно из названия, суть метода заключается в реакции обмена водорода на дейтерий и последующей детекцией на масс-спектрометре. В основе большинства исследований методом HDX-MS лежит стратегия продолжительного мечения объекта исследования в D₂O. Включение дейтерия отслеживается в течение определенных промежутков времени, от секунд до нескольких часов или дней. В процессе обмена масса белка постепенно увеличивается согласно кинетическим параметрам. Такая стратегия применяется для наблюдения структурной динамики белковых комплексов.

В процессе обмена участвуют атомы водорода как боковых цепей, так и остова, в частности, амидные протоны. Однако реакция обмена (как прямая, так и обратная) в боковых остатках аминокислот происходят очень быстро и не влияет на изменение конечной массы белка или пептидов при дальнейшем анализе.

Общая схема эксперимента показана на рис. 2. Сначала белок подвергается мечению в D₂O, реакция обмена проводится определённое время. Затем, аликвота белка отбирается, и реакция останавливается при пониженной температуре и pH. Для исключения реакции обратного обмена образцы немедленно замораживаются в жидком азоте и перемещают на сухой лед или в морозильную камеру на -80°C. Непосредственно перед анализом образцы размораживаются и загружаются в систему. Аликвота белка проходит через протеолитическую колонку и подвергается ферментативному расщеплению, а получившийся пул пептидов задерживается на колонке-ловушке. Расщепление происходит при комнатной температуре и занимает в основном от 2 до 5 минут. Накопление пептидов и дальнейший анализ проводится при 0°C для снижения реакции

обратного обмена. Протеолитически полученные пептиды смываются с колонки-ловушки градиентом растворителей и анализируются на масс-спектрометре. Полученные спектры обрабатываются в специальном программном обеспечении. На основании полученных данных происходит сопоставление со структурой из базы данных или с результатами компьютерного моделирования, что позволяет установить структурную динамику белковой молекулы. Подробные механизмы обмена будут описаны далее.

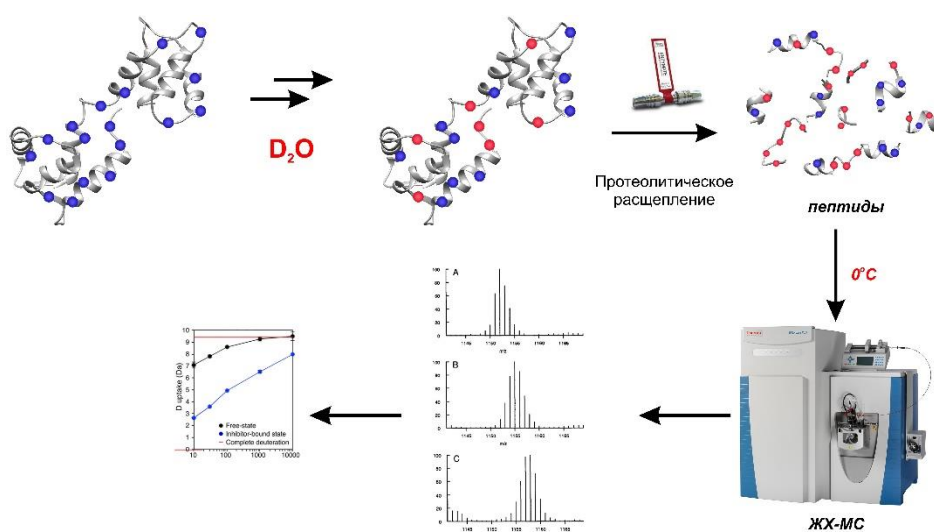


Рисунок 2. Общая схема проведения реакции обмена водорода на дейтерий.

Есть две стратегии проведения экспериментов обмена. Первый и наиболее распространенный – инкубация исследуемого белка или комплекса в D_2O . Другой подход заключается в выдерживании образца в D_2O , а затем добавляют H_2O , чтобы вызвать обратный обмен. Второй метод используется редко и далее он обсуждаться не будет.

Помимо этого, существует несколько сценариев обмена водорода на дейтерий «восходящий» (bottom-up), «нисходящий» (top-down), «среднесходящий» (middle-down) (рис. 3).

Восходящий сценарий обмена на сегодняшний день является наиболее представленным подходом, вкратце он описан выше. В данном подходе белок участвует в реакции обмена при комнатной температуре и рD ~7,6, которая останавливается в определенные временные точки путем снижения температуры до 0°C и рН до 2,5. Затем следует этап протеолитического расщепления протеазами, проявляющими активность в кислых условиях. Пептиды, полученные в процессе расщепления, разделяются ВЭЖХ при пониженной температуре и анализируются масс-спектрометрическими методами. Анализ доменов белка предоставляет структурную информацию на уровне разрешения пептидов. Для данного подхода отсутствует ограничение по размеру исследуемых белков.

Стратегия нисходящего H/D обмена получила меньшее распространение, хотя является мощной альтернативой восходящего метода анализа. Данный сценарий может увеличить пространственное разрешение, что достигается прямой фрагментацией в газовой фазе белка посредством электронного механизма диссоциации. В нисходящем подходе отсутствует процедура расщепления аналита. Благодаря этой особенности шаг ВЭЖХ может быть проведен при температурах ниже 0 °C, что максимально снижает возможность протекания реакции обратного обмена. Наиболее популярным объектом для исследования данным методом являются антитела. Нисходящая стратегия проведения реакции имеет наибольшее значение для изучения белков с большими массами. Однако применимость разрешенного метода HDX/ECD-MS для малых пептидов продемонстрировал Рэнд с коллегами [28], а для белка 17 кДа применимость метода была показана Пэном (Pan) с соавторами [29].

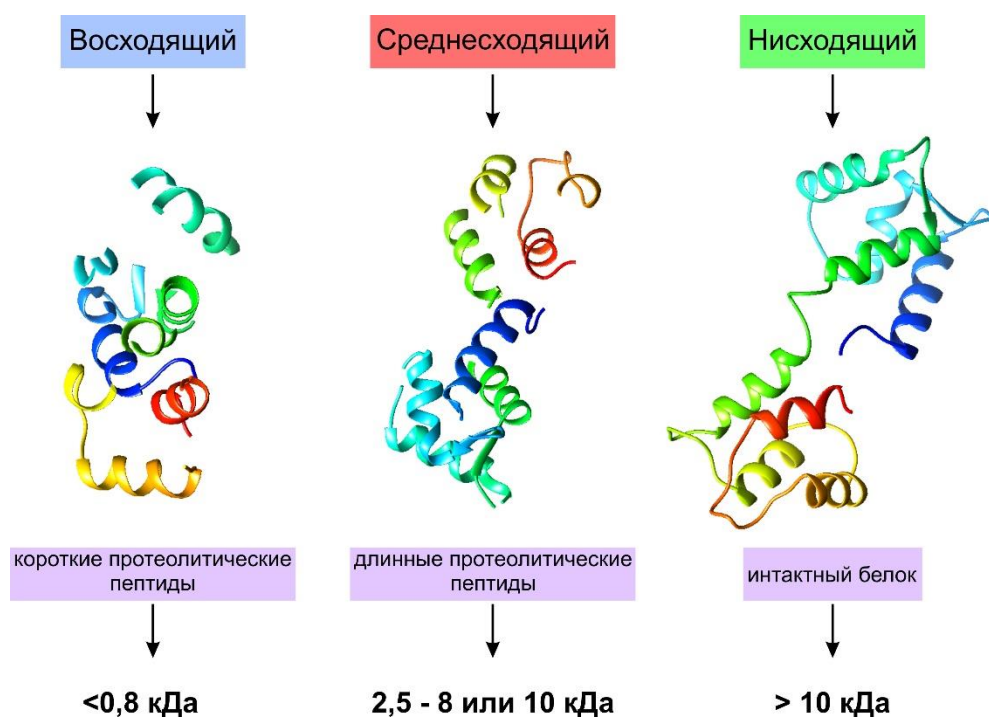
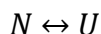


Рисунок 3. Стратегии проведения экспериментов по Н/D обмену.

Сочетанием положительных аспектов двух предыдущих стратегий обладает среднесходящий HDX. Данный метод отлично подходит для изучения больших белков или тех белков, для которых нисходящий метод не эффективен. Стратегия заключается в расщеплении белков в растворе на ограниченное количество специфических фрагментов при низкой температуре и в кислых условиях. Разделение фрагментов происходит с помощью ВЭЖХ с последующей фрагментацией в масс-спектрометре. Среднесходящий сценарий может дать информацию со 100 % покрытием (для 150 кДа mAb) и разрешением порядка одного аминокислотного остатка [30].

Метод Н/D обмена может дать информацию о механизмах взаимодействия белков с субстратами. Это является возможным, поскольку связывание лиганда повышает термодинамическую стабильность структуры, приводя к защите от обмена протонов, в сравнении с нативным белком. Такие процессы для простоты восприятия принято

описывать двумя состояниями, равновесным переходом между нативной конформацией белка (N) и развернутым состоянием U:



Для регистрации и характеристики короткоживущих переходных состояний структурирования используется импульсный HDX. Белок подвергается краткому импульсу HDX (порядок миллисекунд) в определенные времена протекания реакции. Серии измерений связаны с варьированием временных интервалов между сворачиванием и импульсом. Таким образом появляется возможность получить детальное понимание процессов и событий, приводящих белок из разупорядоченного состояния к нативной конформации. Импульсный HDX также требует основных условий, рD 8-10.

$$pD = pH + 0,4$$

2.5. *Преимущества метода HDX-MS*

Положительными аспектами применения HDX в белковых структурах является то, что атомы водорода, связанные с углеродом, практически не участвуют в обмене, водород в боковой цепи белка обменивается с высокой скоростью, неуловимой для LC-MS. Амидные атомы водорода белкового остова вступают в реакцию обмена со скоростью, достаточной для детекции, что позволяет установить вторичные элементы структуры, поскольку обмену подвергается каждая аминокислота, за исключением пролина.

Обмен является функцией таких характеристик, как водородные связи и доступность поверхности растворителю. Медленное протекание процесса обмена говорит о том, что поверхность экранирована от растворителя и/или существует водородное связывание. Быстрый обмен, в свое время, является индикатором доступных растворителю частей белковой структуры и/или отсутствия водородных связей. Изменение представленности

доменов растворителю могут быть отслежены путем измерения количества и скорости включения дейтерия в NH связи.

2.6. Ограничения метода HDX-MS

Ограничений у метода HDX-MS немного, по сравнению со многими биофизическими методами анализа. Одной из проблем является получение большого покрытия последовательности, в частности для белков с большим количеством дисульфидных мостиков. Проблема преодолевается присутствием подходящего восстанавливающего агента, однако в таких условиях существует опасность деактивации пепсина, содержащего три -S-S- связи. В процессе протеолитического расщепления обычно находится 5-10 аминокислотных остатков, а в около идеальных условиях разрешение увеличивается практически до одной аминокислоты. Для преодоления длинных пептидов процесс расщепления проводят при высоких давлениях.

Вторая сложность – анализ данных HDX, однако разработанное и совершенствующееся программное обеспечение значительно облегчает задачу.

2.7. Механизмы реакции обмена

Реакция обмена водорода на дейтерий катализируется как основаниями, так и кислотами. Схемы механизмов H/D обмена в различных условиях приведены на рисунке 4.

В основном катализе (рис. 4а) амидный протон отщепляется вследствие нуклеофильной атаки ионом OH^- (или OD^-) с образованием амидат-иона, который последовательно репротонируется с отщеплением протона от молекулы воды[31]. Кислотный катализ может протекать по двум механизмам: через N-протонирование (рис. 4б) и O-протонирование (рис. 4в). Механизм N-протонирования протекает по схожему с катализом основаниями механизму. Первичная реакция - протонирование атома азота H_3O^+ (или D_3O^+) с последующим депротонированием интермедиата. Альтернативный

механизм катализа кислотой предложил Перрен [32]. В процессе реакции протонируется атом кислорода с помощью H_3O^+ (или D_3O^+), что приводит к окислению протона NH. Протон отщепляется водой, образуя имидиевую кислоту с последующим протонированием D_3O^+ с образованием амида.

При физиологических значениях pH основным механизмом изотопного обмена в белках является катализируемый основаниями.

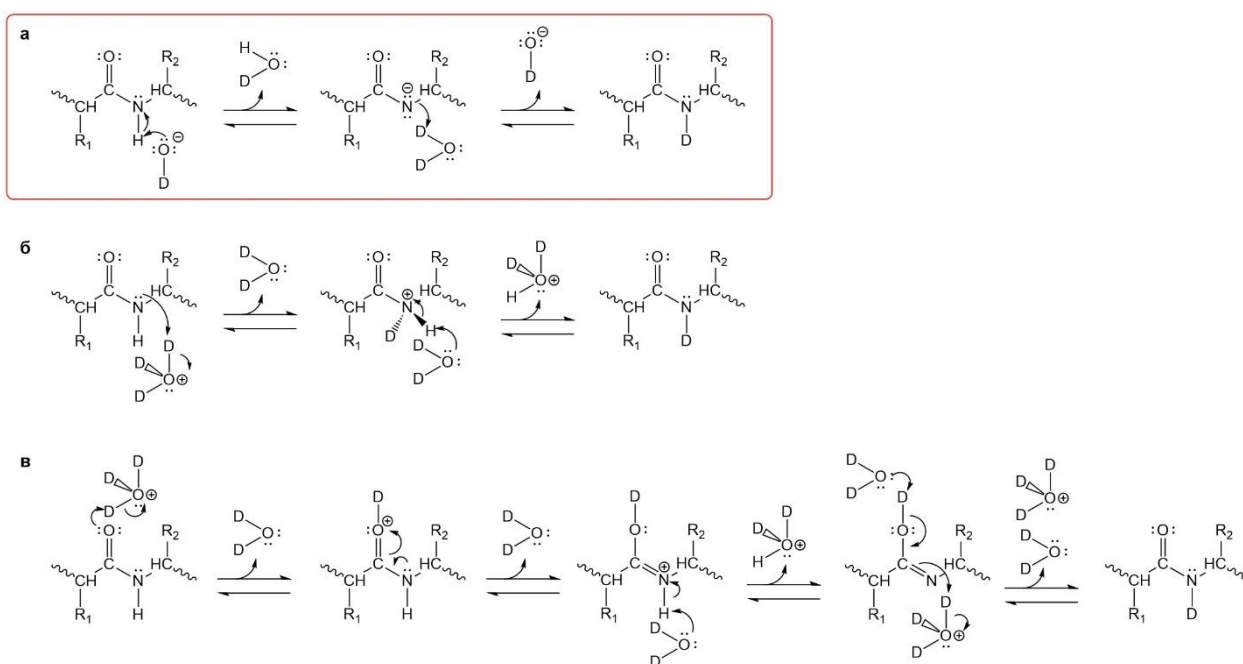


Рисунок 4. Механизмы обмена водорода на дейтерий, катализируемый основаниями (а) и кислотами (б и в). Для кислотного катализа показаны два возможных механизма протекания реакции: N-протонирование (б) и O-протонирование (в).

2.8. Кинетика процессов обмена

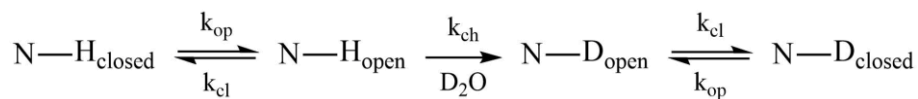
Ключевым процессом водородно-дейтериевого обмена является скорость процесса обмена $\text{N-H} \rightarrow \text{N-D}$. Относительные уровни обмена для каждой части белка описываются как функция от времени мечения:

$$\text{уровень дейтерирования } (t) = \frac{m(t) - m_0}{m_{100} - m_0}$$

где, $m(t)$ – масса центроида (средняя масса) целевого пептида, m_{100} и m_0 представленные данные для полного мечения и не меченного образца соответственно.

В модели Линдерстрема-Ланга обмен защищенными амидами происходит на разных уровнях структурного развертывания. В зависимости от точного местоположения амида могут возникать одна или несколько кратковременных высокоэнергетических конформационных флуктуаций, которые на мгновение разрушают внутреннюю водородную связь амида и делают возможной образование водородной связи с растворителем. Эта коллективная подвижность нативно свернутых белков была описана Линдерстрём-Лангом как дыхательные движения белков [48].

Обмен водорода на дейтерий может быть охарактеризован «химической» константой скорости k_{ch} . Также используются понятия открытого и закрытого состояний белка с соответствующими константами, k_{op} и k_{cl} , соответственно. Общая схема обмена описывается следующим образом:



где каждая N-H группа может быть представлена уникальной комбинацией k_{op} , k_{cl} и k_{ch} .

На процесс обмена влияет структурное разворачивание и сворачивание белка, а также внутренняя скорость обмена (k_{ch}), таким образом константа скорости обмена k_{HDX} :

$$k_{HDX} = \frac{k_{op}k_{ch}}{k_{op} + k_{cl} + k_{ch}}$$

В условиях нативного состояния после разворачивания белки быстро возвращаются в свернутое состояние и $k_{cl} \gg k_{op}$, соответственно уравнение приобретает вид:

$$k_{HDX} = \frac{k_{op}k_{ch}}{k_{cl} + k_{ch}}$$

Представленная схема реакции обычно описывается двумя лимитирующими состояниями, EX1 и EX2. В первом состоянии наблюдаемая скорость обмена равна константе скорости реакции разворачивания, $k_{HDX} = k_{op}$. Этот процесс называется кинетикой EX1, или коррелированным обменом. Большинство белков в физиологических условиях имеют достаточно стабильное нативное состояние, а рефолдинг после кратковременных событий разворачивания происходит гораздо быстрее, чем скорость химического обмена, $k_{cl} \gg k_{ch}$, и $k_{HDX} = \frac{k_{op}k_{ch}}{k_{cl}}$. В условиях протекания данного процесса происходит единичное событие разворачивания, а множество циклов разворачивания/сворачивания может происходить до того, как амидный водород подвергнется изотопному обмену. Такая схема называется кинетикой EX2 или некоррелированным обменом.

В масс-спектрометрических экспериментах процессы EX1 и EX2 различимы (рис. 5а, б). Например, для экспериментального наблюдения кинетики EX1 можно использовать денатуранты или экстремальные значения pH для разворачивания белков. Однако в спектрах, полученных экспериментальным путем иногда проявляется смешанная кинетика EX1/EX2 или EXX (рис. 5в). Такое явление наблюдается в том случае, когда кинетика EX2 и EX1 возникают одновременно или может наблюдаться, когда разные кинетики проявляются в разных областях одного и того же белка или пептида. Альтернативным объяснением может являться то, что одни и те же атомы амидных водородов могут иметь разные переходы разворачивания как в коррелированных, так и в некоррелированных процессах [33].

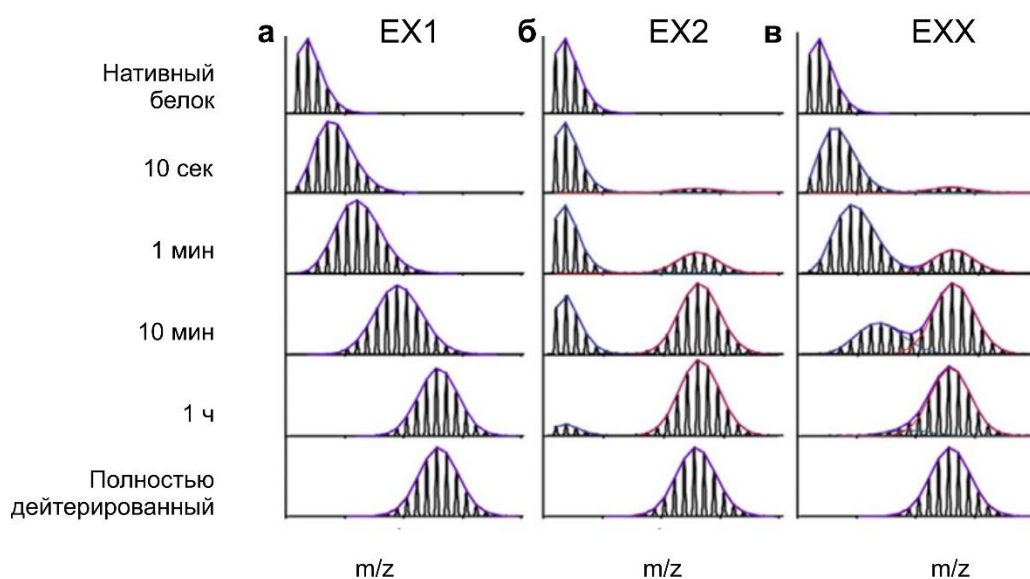


Рисунок 5. Экспериментально полученные масс спектры, отражающие кинетику процессов обмена: **а.** EX1, коррелированный обмен, **б.** EX2, некоррелированный обмен, и **в.** EXX, смешанная кинетика. Изображение адаптировано из [34].

2.9. Области применения метода HDX-MS

В 2015 году Пиррон с коллегами проанализировал публикации по HDX-MS за 30 месяцев (2012-2014 годы) [35]. Полученные результаты показали, что 68,9 % статей приходилось на долю исследовательских институтов, 4,7 % - на коммерческие компании и 1,7 % - на государственные проекты (рис. 6а). Также 24,3 % занимали публикации из смешанных областей: государственные + академические проекты занимают 11,1 %, производство + академия – 13,2 %, а проекты, поддержанные всеми тремя институтами, внесли вклад всего 0,4 %.

а

б

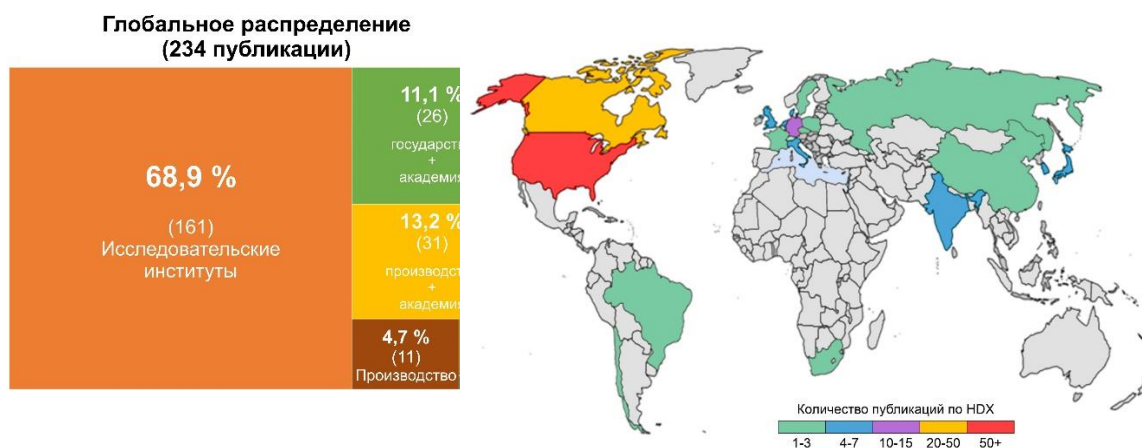


Рисунок 6. Статистика публикационной активности в течение 30 месяцев (2012-2014 гг.) в области HDX-MS по количеству статей и организациям (а) и по странам (б). Рисунки адаптированы из [35].

Как видно на рисунке 6б, в изученный период наибольший вклад в публикационную активность был сделан США и Канадой. Сейчас сохраняется данная тенденция, однако все больше исследований публикуются учеными из стран Европы и России.

Объектами изучения являются белки и их комплексы с различными субстратами. Существует несколько ярко выраженных направлений исследований методом HDX-MS. Просторную нишу занимают исследования структуры белка, куда относят немембранные белки. Исследование структурных особенностей белков делает важный вклад в понимание их функций. Большинство таких экспериментов показывают сравнение по меньшей мере двух конформаций и/или форм белка. Полученные данные интерпретируются с помощью кристаллической структуры одной из изучаемых форм, если таковая имеется.

Комбинация метода HDX-MS с другой структурной информацией, полученной методом РСА, ЯМР, крио-ЭМ или методами компьютерного моделирования, позволяет сделать более точные выводы о структуре белка.

Огромный пласт работ посвящен изучению мембранных белков. Мембранные белки сложны в изучении классическими методами биофизики, поскольку имеют большой размер, не кристаллизуются, а данные, полученные методами ЯМР сложны в интерпретации. Метод HDX-MS имеет ряд преимуществ при изучении данных белков, поскольку требуется небольшое количество материала, есть возможность отделить белки от липидов. На конформацию мембранных белков сильное влияние оказывает окружение, то есть мембрана посредством электростатических и гидрофобных взаимодействий. Трансмембранные белки более сложный объект изучения, чем периферийные белки, поскольку высокогидрофобные области охватывают мембрану и осложняют отделение от липидной составляющей и расщепление. Реакции обмена таких белков проводят в присутствии детергента или в мембранных миметиках. Основным требованием является возможность снижения количества или полное удаление дополнительных компонентов перед ионизацией.

Исследователи также заинтересованы в различных механизмах сворачивания/разворачивания белковых молекул. Как говорилось ранее, на обмен во время фолдинга оказывает влияние формирование водородных связей и доступность структурных элементов растворителю. Для изучения механизмов этих процессов используются денатурирующие агенты, а для мониторинга обратных процессов денатурант разбавляют. Поскольку реакции сворачивания/разворачивания протекают быстро, необходимо использовать быстрые способы наблюдения за изменениями. Таким методом является импульсное мечение дейтерием в течение 10 с. Короткие импульсы позволяют зафиксировать не денатурированные или незащищенные участки белка и тем самым проследить процессы сворачивания/разворачивания. Наряду с данными фолдинга могут быть исследованы факторы, влияющие на его протекание, например, давление, модификация различных групп, мутации, связывание и др.

Белки шапероны принимают участие в сворачивании некоторых белков. Эксперименты HDX-MS позволяют понять не только вариативность процессов сворачивания, но и как функционируют шапероны.

Особый интерес представляет изучение процессов взаимодействия белков с другими белками, нуклеиновыми кислотами или малыми молекулами. Регуляция активности и функций белка зачастую происходит путем взаимодействия с другими молекулами. Понимание происходящих событий может дать критически важную информацию о состояниях, связанных с заболеваниями.

Взаимодействие белков с точки зрения дизайна терапевтических препаратов находит отражение в картировании эпитопов (часть макромолекулы антигена, распознаваемая иммунной системой). Так, например, изучение структуры белка, связывающего фактор H (fHbp *Neisseria meningitidis*) методом HDX-MS, дало возможность быстрого получения практически полной информации о структуре эпитопа [36]. Авторы проводили сравнение таких методов как HDX-MS, PCA, фаговый дисплей и пептидный массив.

Исследование белков, взаимодействующих с нуклеиновыми кислотами, необходимо для понимания механизмов, протекающих при различных процессах, например, репарации. Эта фундаментальная информация позволяет в дальнейшем разрабатывать стратегии подавления или увеличения эффективности работы ферментов для задач медицины, биотехнологии и т.д.

Также белки изучаются в составе лекарственных препаратов. Как минимум половина объема исследований отведена моноклональным антителам и в частности IgG, широко используемым в фармакологической индустрии. HDX-MS играет важную роль в структурной характеристике антител, поскольку они обладают большой массой.

Белковые препараты могут быть склонны к агрегации, индуцированной окружающей средой или процессами производства. Важно понимать, какие белковые агрегаты могут обладать токсическим эффектом для пациентов. Помимо антител актуальным вопросом характеристики структуры стоит для других белков. Ярким примером является рекомбинантный фактор коагуляции IX (rFIX) [37]. Важной составляющей биофармацевтической ниши являются работы по изучению инсулина методом HDX-MS, включая анализ различных форм и вариантов инсулина с различными терапевтическими свойствами.

Отдельно стоит выделить разработку и усовершенствование методов HDX-MS. Первым направлением является работа связанная с увеличением покрытия белка, модификацией методов протеолитического расщепления, восстановления дисульфидных мостиков с использованием электрохимической ячейки, аффинный захват [38]. Разработка методов, снижающие обратный обмен для МАЛДИ [39] и электроспрея [40], а также условия быстрого смешивания для H/D мечения [41]. Наряду с этим усовершенствование методов LC-MS включают проведение экспериментов с охлаждением ниже нуля для минимизации обратного обмена [33], использование ECD, ETD фрагментации в нисходящих стратегиях HDX-MS [42].

Гросс с коллегами разработал метод LC/ESI-MS для наблюдения за взаимодействием белок-лиганд посредством MS, титрованием водородно-дейтериевый обмен ("PLIMSTEX") [43]. Получаемы кривые показывают количество обмененных амидных протонов как функцию концентрации лиганда от заданного времени обмена.

В своей работе Фитцджеральд с соавторами описывают альтернативный метод, основанный на использовании МАЛДИ масс-спектрометрии и упоминается как SUPREX (stability of unpurified proteins from rates of H/D exchange) [27]. Описанный авторами метод обладает уникальной применимостью к смесям белков. Специфично, что количество

обмениваемых водородов является функцией от концентрации денатуранта от фиксированного времени обмена. Повышенная концентрация денатурантов требуется для вовлечения белка в разворачивание в присутствии связанного лиганда. Анализ методом SUPREX подходит для скрининга больших библиотек соединений с целью идентификации высоко афинных лигандов для выбранных белковых мишеней.

2.10. Заключение

История становления методов структурной биологии начинается в первой половине XX века и получает широкое развитие начиная с его середины. Бурно развивающиеся кристаллографические методы постепенно дополняются подходами ядерного магнитного резонанса, а в последствии и методами масс-спектрометрии. Имея небольшое количество ограничений, метод водородно-дейтериевого обмена получает все большее распространение благодаря балансу между разрешением и производительностью. Изучение процессов Помимо фундаментальных исследований метод HDX-MS находит применение в биофармацевтике, биотехнологических разработках и др.

Усовершенствование подходов к проведению экспериментов, анализа образцов, инструментальных методов и развитие программного обеспечения позволяет получать большее количество информации и упрощает интерпретацию данных, что делает HDX-MS привлекательным методом для широкого спектра объектов. На ряду с этим, увеличивается применимость метода с точки зрения процессов, происходящих в белке, влияющих на его активность или механизм действия.

3. Материалы и методы

3.1. Материалы

3.1.1. Общие материалы

В работе были использованы *N,N'*-метиленбисакриламид, *N,N,N',N'*-тетраметилэтилендиамин (ТЕМЕД), борная кислота, глицерин, агар, бромистый этидий, NaCl, KCl, MgCl₂, агароза, триптон, дрожжевой экстракт, этилендиаминтетраацетат (ЭДТА), акриламид, трис(гидроксиметил)аминометан (Трис), канамицин 1,4-дителиотреитол (ДТТ), изопропил-β-Дтиогаляктопиранозид (ИПТГ), (все — «Helicon», Россия), коктейль-ингибитор протеазы (PIC), реагент Бредфорда, муравьиная кислота, оксид дейтерия (D₂O), имидазол (все — «Sigma», США), персульфат аммония, ацетонитрил (все — «AppliChem», Германия), бромфеноловый синий, набор белковых маркеров 10-250 кДа, SDS (все — «Диа-м», Россия), кумасси голубой («VWR Life Science», США), набор для очистки плазмид NucleoBond® Xtra Midi («Macherey-Nagel», Германия). Все реактивы и растворители квалификации х.ч. и ос.ч. Во всех экспериментах, использовали бидистиллированную или деионизированную воду.

При выделении белков для хроматографии использовались колонки: гепариновая колонка HiTrapTM Heparin 1 мл («Cytiva», США) и хелатирующая колонка: HiTrapTM Chelating HP 1 мл («Cytiva», США)

В HDX-MS экспериментах использовались колонки: аналитическая - ACQUITY UPLCTM Peptide BEH C18 1,7 мкм, 1x10 мм; «Waters Corporation», США) и колонка-ловушка: VanGuardTM BEH C18 1,7 мкм, 2,1x5 мм («Waters Corporation», США)

Плазида pET24a со вставкой гена hNEIL2 была синтезирована коммерчески.

Белок SpCas9 был любезно предоставлен Степановым Г.А. (Лаборатория геномного редактирования, ИХБФМ СО РАН?).

3.1.2. Стандартные буферы и смеси

Буфер для лизиса клеток: 20 мМ HEPES KOH, pH = 7,6; 150 мМ NaCl; 20 мМ Имидазол; 1 % тритон; 5 % глицерин; 2 мМ ЭДТА; 5 мМ MgCl₂.

Буфер А для хелатирующей колонки: 20 мМ HEPES KOH, pH = 7,6; 500 мМ NaCl; 30 мМ Имидазол; 0,025 % NP-40.

Буфер Б для хелатирующей колонки: 20 мМ HEPES KOH, pH = 7,6; 500 мМ NaCl; 500 мМ Имидазол; 0,025 % NP-40.

Буфер А для гепариновой колонки: 20 мМ HEPES KOH, pH = 7,6; 50 мМ NaCl.

Буфер Б для гепариновой колонки: 20 мМ HEPES KOH, pH = 7,6; 1 М NaCl.

SOC: 2 % триптон (w/v), 0,5 % дрожжевой экстракт (w/v), 10 мМ NaCl, 2,5 мМ KCl, 10 мМ MgCl₂, 20 мМ глюкоза.

Среда LB: 10 г/л триптона, 5 г/л дрожжевого экстракта, 10 г/л NaCl.

Трис-глициновый буфер: 25 мМ Трис-глицин (pH 8,3), 0,1 % SDS.

Буфер для проведения H/D обмена: 20 мМ HEPES KOD, pH = 7,6; 150 мМ NaCl; 1 мМ TCEP в D₂O.

3.2. Методы

3.2.1. Выделение плазмидной ДНК

Для трансформации плазмидной ДНК использовали электрокомпетентные клетки *E. coli* NovaXG (Novagen, Германия). 1 мкл плазмиды hNEIL2 с концентрацией 0,1 мкг/мл смешали с 40 мкл клеток в предварительно охлажденной кювете для электропорации, которую затем поместили в электропоратор (Амаха Biosystems, Германия), 2,1 кВ. В кювету незамедлительно добавили 300 мкл SOC, аккуратно пипетировали 2-3 раза и перенесли в пробирку. Клетки в питательной среде инкубировали при 37 °С в течение 1 ч

и перемешивании (250 об/мин). На чашки с агаризованной средой LB, содержащей антибиотик канамицин (50 мкг/мл) высевали клетки и инкубировали в течение ночи при 37 °С. В колбу объемом 250mL добавили 20mL LB. Одиночную колонию переносили в 20 мл среды LB, содержащую 50 мкг/мл канамицина и оставляли инкубироваться в течение ночи при 37 °С и 220 об/мин. В колбу, содержащую 180 мл среды LB добавили 200 мкл канамицина (конечная концентрация 50 мкг/мл) и затем 20 мл Nova XG NEIL2 и инкубировали при 37 °С, 220 об/мин в течение ночи. Из ночной культуры выделяли плазмидную ДНК с помощью набора NucleoBond® Xtra Midi по стандартному протоколу производителя.

3.2.2. Экспрессия, выделение и очистка рекомбинантного белка hNEIL2

Экспрессия. Плазмидную ДНК трансформировали в электрокомпетентные клетки *E. coli* Rosetta2 (DE3) (Novagen, Германия) по протоколу, описанному выше. 190 мл LB среды, содержащей 50 мкг/мл канамицина инокулировали 10 мл ночной культуры *E. coli* Rosetta2 (DE3), несущей плазмиду pET24a-hNEIL2. Культуру выращивали при 37 °С на качалке со скоростью 220 об./мин до оптической плотности $OD_{595} = 0,6-0,7$. Затем 100 мл клеточной культуры добавляли к 900 мл LB с 50 мкг/мл канамицина и аналогично предыдущему шагу инкубировали до достижения $OD_{595} \sim 0,6$. Колбу оставляли при 4 °С перед началом индукции. Синтез рекомбинантного белка индуцировали добавлением ИПТГ до концентрации 50 мкМ (475 мкл 1 М), рост культуры продолжали в течение ночи при 15 °С.

Выделение. Клетки осаждали центрифугированием при 5000×g в течение 7 мин. при 4 °С и распускали в 10 мл буфера для лизиса. К полученной суспензии добавляли ДТТ (12 мкл 1 М) и коктейль ингибитора протеаз (120 мкл 1 М) и дважды пропускали через French Press (Thermo Fisher Scientific, США). Клеточный дебрис отделяли центрифугированием при 20000×g в течение 1 ч. при 4 °С. В супернатант добавляли NaCl

(4 М) и имидазол (3 М) в необходимом количестве, чтобы конечная концентрация была аналогичной буферу А для хелатирующей колонки.

Очистка. Первой ступенью очистки белка является хроматография через хелатирующую колонку. Колонку (HiTrapTM Chelating HP, 1 мл) подготавливали последовательным промыванием 5 мл H₂O, 5 мл ЭДТА 0.05 М, 5 мл H₂O, 4 мл NiCl₂ 0.1 М и уравнивали 5 мл буфера А. На подготовленную колонку наносили разбавленный супернатант с целевым белком. Колонку промывали 5 мл буфера А и вели элюцию градиентом 30–500 мМ имидазола, объем градиента составлял 10 мл, фракции собирали объемом 0,8 мл. Фракции в районе основного пика поглощения на 280 нм проанализировали методом электрофореза (4-12 % SDS-ПААГ) и объединяли те, которые содержали целевой белок. Фракции разбавляли буфером до концентрации соли буфера А для гепариновой колонки (HiTrapTM Heparin, 1 мл), содержащего 20 мМ HEPES буфер (pH 7,6), 50 мМ NaCl и наносили на уравновешенную буфером А колонку. Колонку промывали 5 мл нижнего буфера (А) и вели элюцию градиентом 50–1000 мМ NaCl, объем градиента составлял 10 мл. Фракции собирали по 0,8 мл и в районе основного пика поглощения на 280 нм анализировали электрофорезом в 4-12 % SDS-ПААГ с окрашиванием кумасси голубым. Фракции, содержащие только полосу, соответствующую подвижности hNEIL2, объединяли и добавляли ДТТ до конечной концентрации 1 мМ и 10% (v/v) глицерин. Аликвоты белка хранили при температуре –80 °С для дальнейшей работы. Общую концентрацию выделенных белков определяли методом Брэдфорда с использованием БСА для построения калибровочных кривых.

3.2.3. Электрофорез в системе Лэммли

Анализ белковых фракций проводили методом электрофореза по Лэммли. Концентрирующий гель состоял из 5 % акриламида, 125 мМ Трис-НСl (pH 6,8), 0,05 %-ного персульфата аммония (ПСА), 0,1%-ный ТЕМЕДа и 0,1 %-ного SDS, а разделяющий гель – из 12 % акриламида, 375 мМ Трис-НСl (pH 8,8), 0,05 %-ного ПСА, 0,1 %-ного

ТЕМЕДа и 0,1 %-ного SDS. Образцы перед нанесением инкубировали с краской () при 95 °С в течение 5 минут. Электрофорез проводили в трис-глициновом буфере с градиентом при напряжении 14–20 В/см. Для визуализации результатов гель окрашивали раствором Кумасси, а затем отмывали раствором, содержащим этанол и уксусную кислоту.

3.2.4. Эксперименты по водородно-дейтериевому обмену с последующей масс-спектрометрией hNEIL2

Реакция обмена. Перед началом обмена концентрацию очищенного белка hNEIL2 доводили до 5,4 мкМ в 20 мМ HEPES (pH 7,6), 150 мМ NaCl с добавлением 1 мМ трис(2-карбоксиэтил)фосфина (ТСЕР). После 1 ч инкубации при комнатной температуре реакцию обмена инициировали добавлением 160 мкл дейтерированного буфера (20 мМ HEPES (pH 7,6), 150 мМ NaCl, 1 мМ ТСЕР в D₂O) к 20 мкл уравновешенного раствора белка hNEIL2. Конечное содержание дейтерия составляло ~90 %, что способствовало однонаправленному обмену. H/D обмен останавливали в определенные временные точки (10 с, 30 с, 1 мин, 2 мин, 5 мин, 10 мин, 30 мин, 1 и 2 ч), смешивая аликвоту 20 мкл реакционной смеси с 40 мкл охлажденного буфера с pH = 2,5, содержащего 2,5 % муравьиную кислоту и 4 М мочевины. Раствор быстро перемешивали и немедленно замораживали в жидком азоте, затем переносили на сухой лёд, хранили при -80 °С не более месяца. Эксперименты проводили в трёх повторах для каждой временной точки. Недейтерированные контроли обрабатывали по аналогичному протоколу.

Анализ образцов. 10 пмоль меченого hNEIL2 расщепляли в течение 2 минут при скорости потока 100 мкл/мин растворителя А (0,15% муравьиной кислоты, pH 2,5) с использованием колонки с иммобилизованным пепсином при 20 °С. Полученные после расщепления пептиды задерживались, концентрировались и обессоливали на колонке-ловушке при 0 °С и потоке 100 мкл/мин буфера А. Затем пептиды разделяли на аналитической колонке в течение 8 минут линейным градиентом 5-30 % растворителя Б

(0,15 % муравьиной кислоты в ацетонитриле, рН 2,5) и 2 минуты линейным градиентом 30-40 % при 0 °С и скоростью потока 40 мкл/мин. Поток ЖХ направляли на масс-спектрометр Synapt G2-Si HDMS, оснащенный стандартным источником ионизации электрораспылением (ESI). Масс-спектры получали в режиме положительных ионов и разрешения в диапазоне m/z 50–1950. В случае образцов, не содержащих дейтерий, использовали те же хроматографические условия. Для оптимизации эффективности фрагментации использовались четыре различных линейных изменения энергии столкновения ловушек MS^E (10–30 В, 15–35 В, 20–45 В и 10–45 В).

После каждого анализа белка пепсиновую колонку промывали дважды раствором 1 % муравьиной кислоты, 5 % ацетонитрила и 1,5 М гуанидиний хлорида (рН 1,7). Между каждым анализом образца делали инъекции буфера А для исключения переноса пептидов.

Анализ данных. Спектры недеитерированных образцов обрабатывались с использованием ProteinLynx Global Server v3.0.1. Поиск пептидов проводили по базе данных, состоящей из последовательности hNEIL2, пепсина и смешанных последовательностей двух белков. Параметры рабочего процесса и обработки PLGS приведены в таблице 1. Дальнейший отбор пептидов проводили в программе DynamX v3.0 (параметры представлены в таблице 1). В результате обработки был отобран 81 пептид и 96,7 % покрытие последовательности белка hNEIL2.

Таблица 1. Параметры обработки результатов HDX-MS.

PLGS	
Протеолитический фермент	Неспецифичный
Скорость ложного обнаружения	100

Минимальное совпадение фрагментных ионов пептидов и белка	3 и 7
Минимум совпадение пептидов/белка	2
Низкий и повышенный энергетический порог	200, 40 счетов
Окно коррекции контрольной массы	0.25 Да при 556.2771 Да/e

DynamX

Минимальная интенсивность	5000
Минимальный счет PLGS	> 6,5
Максимальная ошибка МН ⁺	10 ppm
Минимальные продукты/аминокислота	0,1
Пороговое значение файлов	2

3.2.5. Эксперименты по водородно-дейтериевому обмену с последующей масс-спектрометрией SpCas9 и комплекса SpCas9 с sgРНК

Реакция обмена. Перед началом обмена у белка Cas9 заменили буфер, поскольку он содержал 50 % глицерина, что затрудняет проведение реакции обмена. Смену буфера проводили с помощью фильтров Amicon® Ultra (Merck Millipore, США). Конечный буфер содержал 20 мМ HEPES (pH 7,6), 150 мМ NaCl, 10 мМ MgCl₂ и 1 мМ ТСЕР. Исходный раствор белка инкубировали 1 ч при комнатной температуре, затем добавляли дейтерированный буфер (20 мМ HEPES (pD 7,6), 150 мМ NaCl, 10 мМ MgCl₂, 1 мМ ТСЕР в D₂O) к уравновешенному раствору белка Cas9, так что его конечная концентрация составила 0,6 мкМ. В случае комплекса Cas9 с sgРНК белок также инкубировался в течение часа при комнатной температуре, а РНК добавляли непосредственно перед

началом реакции обмена в соотношении белок:РНК 1:10. Обмен вели определенное время (10 с, 30 с, 1 мин, 2 мин, 5 мин, 10 мин, 30 мин, 1, 2, 4, 6 и 8 ч) и останавливали смешивая аликвоту 20 мкл реакционной смеси с 40 мкл охлажденного буфера, содержащего 2,5 % муравьиную кислоту и 4 М мочевины, рН = 2,5 и немедленно замораживали в жидком азоте. Образцы хранили при -80 °С не более месяца. Эксперименты проводили в трёх повторах для каждой временной точки. Недейтерированные контроли обрабатывали по аналогичному протоколу.

Анализ образцов. 10 пмоль меченого Cas9 расщепляли в течение 2 минут при скорости потока 100 мкл/мин растворителя А (0,15% муравьиной кислоты, рН 2,5) с использованием колонки с иммобилизованным пепсином при комнатной температуре. Белок подвергался протеолитическому расщеплению на пептиды, которые задерживались, концентрировались и обессоливали на колонке-ловушке при 0 °С и потоке 100 мкл/мин буфера А. Полученные пептиды разделяли на аналитической колонке в течение 8 минут линейным градиентом 5-30 % растворителя Б (0,15 % муравьиной кислоты в ацетонитриле, рН 2,5) и 2 минуты линейным градиентом 30-40 % при 0 °С и скоростью потока 40 мкл/мин. Поток ЖХ направляли на масс-спектрометр HF QExactive, оснащенный стандартным источником ионизации электрораспылением (ESI). Масс-спектры регистрировали в режиме положительных ионов и разрешения в диапазоне m/z 200–3000, с использованием метода Full MS – ddMS2.

После каждого анализа белка пепсиновую колонку промывали как описано ранее.

Анализ данных. Идентификацию пептидов проводили путем обработки спектров меченных образцов в Proteome Discoverer v 2.2 с использованием локальной базы данных, содержащей последовательность Cas9. Дальнейшая валидация идентифицированных пептидов проводилась с помощью программного обеспечения HDExaminer. Покрытие последовательности белка Cas9 составила 64%.

3.2.6. Компьютерное моделирование

hNEIL2. Компьютерное моделирование белка человека *hNEIL2* начинали с моделирования по гомологии. Гомологичные модели получали с помощью сервера SWISS-MODEL [44] используя в качестве шаблонов кристаллографические структуры *oNEIL2* (PDB ID: 6VJI [45]) и *mvNEIL2/3* (4MB7[46]). Для проведения молекулярной динамики (МД) из моделей были удалены неструктурированные петли, сформированные аминокислотными остатками 58-125, а разрыв был сшивали и структуру оптимизировали с помощью Modeller [47]. Ион Zn^{2+} помещали в цинковый палец в окружении Cys-291, His-295, Cys-315, Cys-318 и положение уточняли с помощью сервера CheckMyMetal [48].

Молекулярную динамику для стартовой структуры проводили с использованием программного обеспечения Amber20 с ускорением GPU в неявной модели растворителя. Для молекулы белка использовали силовое ff14SB, а для иона цинка – поле TIP3P. Для удаления нежелательных контактов энергию системы минимизировали, а затем постепенно нагревали систему от 20 К до 300 К в течение 100 пс. Далее уравнивали систему в течение 1 нс и проводили 100 нс молекулярную динамику при 300 К, сохраняя траектории каждую 1 пс. Полученные траектории анализировали с помощью UCSF Chimera, водородные связи определяли с помощью MDTRA [49].

SpCas9, *SpCas9-sgPНК* и *SpCas9-sgPНК-ДНК*. Моделирование Cas9 по гомологии выполняли с помощью сервера Phyre2. Структура с идентификатором 4CMQ из базы данных PDB была выбрана в качестве шаблона с идентичностью 92 %. Недостающие участки белка достраивали с помощью Modeller. Структуру белка дополнили тремя ионами Mg^{2+} . Один ион магния окружен Asp-10, Glu-762 и Glu-766 остатками аминокислот, второй ион находится рядом с первым, в окружении Asp-10, Glu-766 и His-983, а третий – His-1297 и Asp-1328 аминокислотными остатками. Полученную конфигурацию использовали в качестве стартовой модели для молекулярной динамики апофермента Cas9. Координаты молекул sgPНК и ДНК для формирования комплекса

получили из кристаллической структуры комплекса *Staphylococcus aureus* Cas9-sgРНК-ДНК (5CZZ в базе данных PDB).

Молекулярную динамику полученных систем проводили в модели неявного растворителя с использованием Amber 20 с GPU-ускорением. Силовое поле ff14SB использовали для белка, силовое поле OL3 для РНК и TIP3P для ионов магния. После шага минимизации систему постепенно нагревали, начиная от 20 К до 300 К в течение 100 пс. Этап уравнивания с последующей МД проводили в течение 50 нс при 300 К в ансамбле с постоянным давлением. Координаты атомов сохраняли каждую 1 пс. Программное обеспечение UCSF Chimera использовали для анализа траекторий и создания молекулярной визуализации.

4. Обсуждение результатов;

4.1. *hNEIL2*

Повреждение ядерной ДНК, такое как алкилирование, дезаминирование, апуринизация и окисление – частое явление, в результате которого ежедневно в каждой клетке человека возникает около 30 000 повреждений [50,51]. Репарация большинства окисленных оснований ДНК происходит посредством высоко консервативного пути репарации оснований, который инициируется иссечением поврежденного основания специфичными ДНК-гликозилазами [50][52]. Большинство гликозилаз может быть разделено на три структурных суперсемейства: α/β -фолд, спираль-шпилька-спираль (HhH), и спираль-два поворота- спираль (H2TH). Суперсемейство H2TH включает бактериальные ферменты Fpg/MutM и эндонуклеазу VIII (Nei) и их эукариотические аналоги NEIL (endonuclease VIII-like): NEIL1, NEIL2, и NEIL3.

Ферменты NEIL впервые были обнаружены в 2000-х и сразу привлекли внимание к своей функциональной избыточности, предполагающую некоторые специализированные роли в репарации ДНК. Например, белки NEIL способны использовать в качестве субстратов отличные от В формы ДНК, включая одноцепочечные ДНК, G-квадруплексы, «глазки», вилки и трех- или четырех-нитевые сшивки. В частности, NEIL2 предпочитает такие субстраты, содержащий глазки, R- и D-петли [53,54]. Однако биологическая роль NEIL2 до конца не ясна, поскольку его специфичность к поврежденным основаниям пересекается с аналогичной у NEIL1, NEIL3 и NTHL1.

NEIL2 наиболее близок к бактериальным белкам Fpg и Nei и в отличие от NEIL1 и NEIL3 содержит один цинковый палец. Однако, отличительной чертой является то, что это Cys₃His цинковый палец, тогда как у Fpg, Nei и NEIL3 - Cys₄.

4.1.1. Подготовка к HDX-MS экспериментам

Белок hNEIL2 экспрессировали в штамме E.coli Rosetta (DE3) при 15 С. Очистку рекомбинантного белка проводили в две стадии. Первый этап очистки hNEIL2 с концевой гистиридиновой меткой проводили с использованием хелатирующей колонки HiTrap Chelating, собранные фракции анализировали электрофорезом в полиакриламидном геле (рис. 7А). Гепариновую колонку использовали на втором этапе очистки, фракции, соответствующие пику белка, анализировали с помощью SDS-ПААГ (рис. 7Б). Фракции, содержащие hNEIL2 объединяли и хранили при -80 °С в буфере, содержащем 10% глицерин и 1 мМ ДТТ.

Перед проведением экспериментов водородно-дейтериевого обмена необходимо подтвердить массу рекомбинантного hNEIL2 и его гомогенность с помощью масс-спектрометрического анализа. Молекулярная масса белка, определенная с помощью LC-MS анализа равна $37631,98 \pm 0,11$ Да (рис. 8). Охарактеризованный рекомбинантный hNEIL2 были использован в последующих экспериментах HDX-MS.

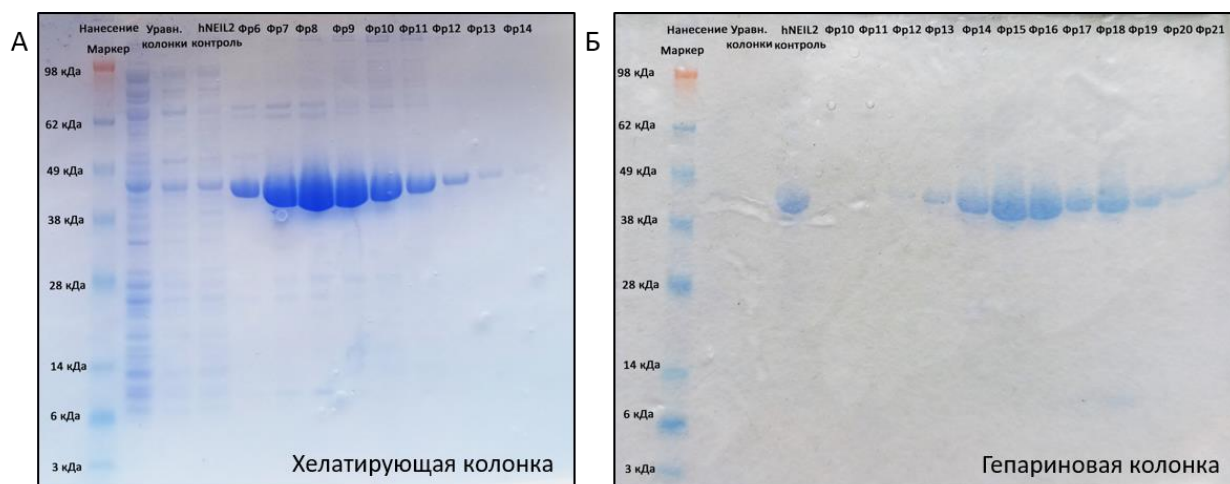


Рисунок 7. Электрофорез в полиакриламидном геле фракций белка после очистки на хелатирующей колонке (А) и гепариновой колонке (Б).

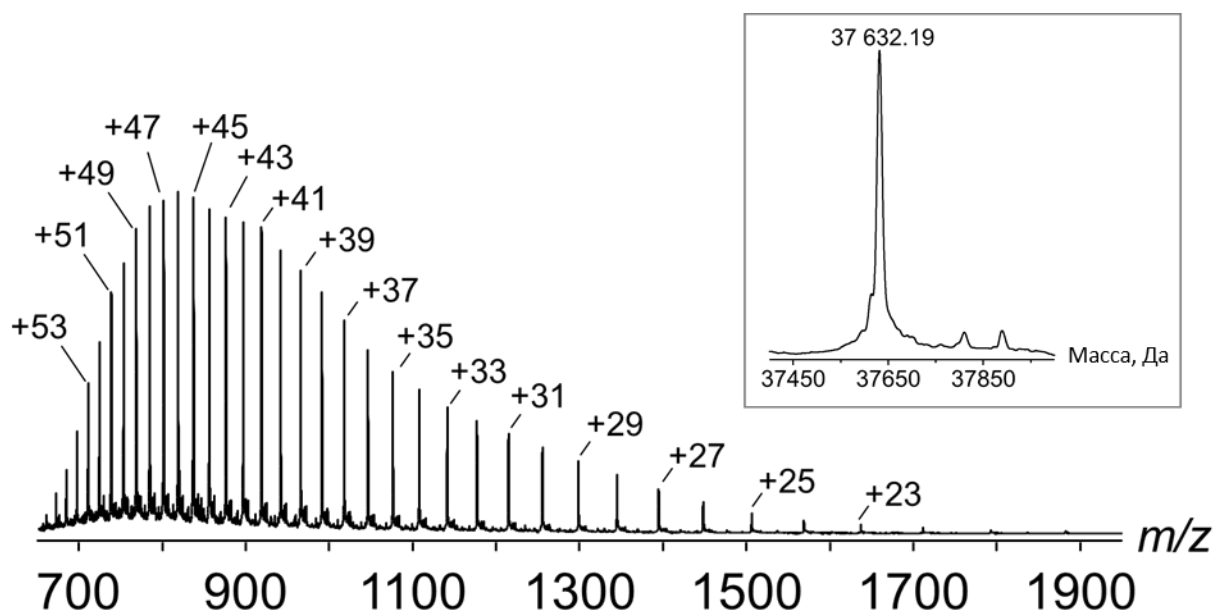


Рисунок 8. Масса рекомбинантного белка hNEIL2, определенная методом LC-MS.

4.1.2. HDX-MS анализ динамики hNEIL2

Исследование структуры hNEIL2 начинали с подготовки образца белка (20 мкМ) и его инкубирования при комнатной температуре. Белок подвергали обмену в дейтериевом буфере (pD 7,6) при комнатной температуре. H/D обмен останавливали через временные интервалы 10 с, 30 с, 1 мин, 2 мин, 5 мин, 10 мин, 30 мин, 1 и 2 ч. После протеолитического расщепления пепсином на колонке полученные пептиды подвергали масс-спектрометрическому анализу. Покрывие последовательности hNEIL2 составило 96,7%. Единственным неразрешенным участком стал фрагмент со 176 по 181 аминокислоту. Уже с первых секунд реакции наблюдали значительную разницу в количестве обменянного дейтерия для различных участков белка. В тоже время остатки 56-129 из пролин-богатой области показали наибольший уровень H/D обмена, 65%. Быстрый обмен характерен для неструктурированных участков структуры белка, например, петель, а также для некоторых структурных элементов, находящихся на поверхности. Так, наиболее интенсивный и быстрый обмен наблюдали в петле β3/β4,

уникальной для гликозилаз hNEIL2. Быстрый обмен происходил в аминокислотных остатках 181-196, образующих линкер между N-концевым и C-концевым доменами, который хорошо упорядочен в структурах oNEIL2 и mvNEIL2/3. Результат реакции обмена в каждой временной точке показан на графике (рис. 9А). Благодаря тому, что неструктурированная петля 56-129 содержит большое количество пролина (10 из 74 остатков), то не может быть эффективно расщеплена пепсином. По этой причине пептиды, покрывающие петлю $\beta 3/\beta 4$, были нетипично длинными, что затрудняло автоматический анализ программой DynamX и пептиды этой области добавляли вручную. Аминокислотные остатки 37-55 и 238-265 в течение первых 10 с практически не демонстрировали обмена. Низкую скорость обмена на протяжении всего времени мечения зафиксировали и в пептидах 1-55, 195-230, 237-266 и 270-322, что характерно для элементов со стабильной вторичной структурой. Для некоторых пептидных групп наблюдали бимодальное распределение пиков, отражающее кинетику EX1 (рис. 9А). Пептиды EX1 принадлежат ядру β -сэндвича в N-концевом домене и мотиву H2TH и его окружению в C-концевом домене, что позволяет предположить, что эти части белка являются наиболее стабильными.

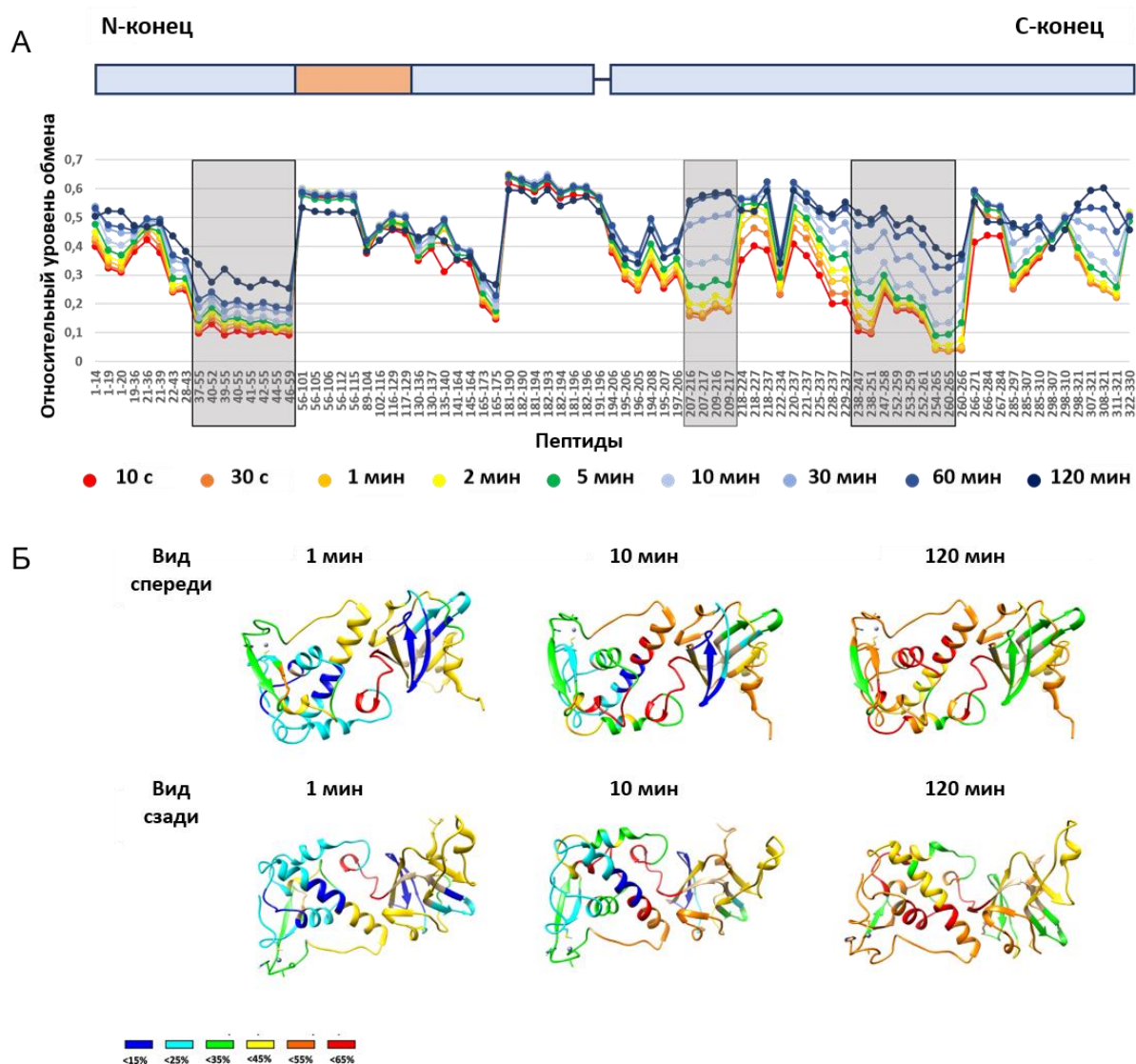


Рисунок 9. Временная эволюция Н/Д обмена hNEIL2. **А**, Относительные уровни обмена пептидов, зарегистрированные в определенные временные точки (от N-конца к С-концу). Серым цветом показаны участки с кинетикой EX1. **Б**, Структуры hNEIL2, полученные методами молекулярной динамики с наложением экспериментального распределения уровней дейтерия в 10 с, 10 и 120 мин.

Наличие большого неупорядоченного фрагмента на N-конце hNEIL2 может влиять на стабильность домена или всего белка. Анализ данных, полученных методом кругового дихроизма позволяет оценить термостабильность белка. Эксперименты КД hNEIL2 проводили в интервале температур 20- 80 °С с последующим разложением спектров для модели двух состояний с использованием алгоритма выпуклых ограничений [55] (рис. 10).

Анализ данных выявил единственный переход при $40,8 \pm 0,5$ °С. Разделить спектры на большее количество состояний не удалось, что позволяет предположить эквивалентность термостабильности обоих белковых доменов. Полученный при 40,8 °С конформационный переход свидетельствует о том, что hNEIL2 принадлежит группе наименее термостабильных белков протеома человека [56].

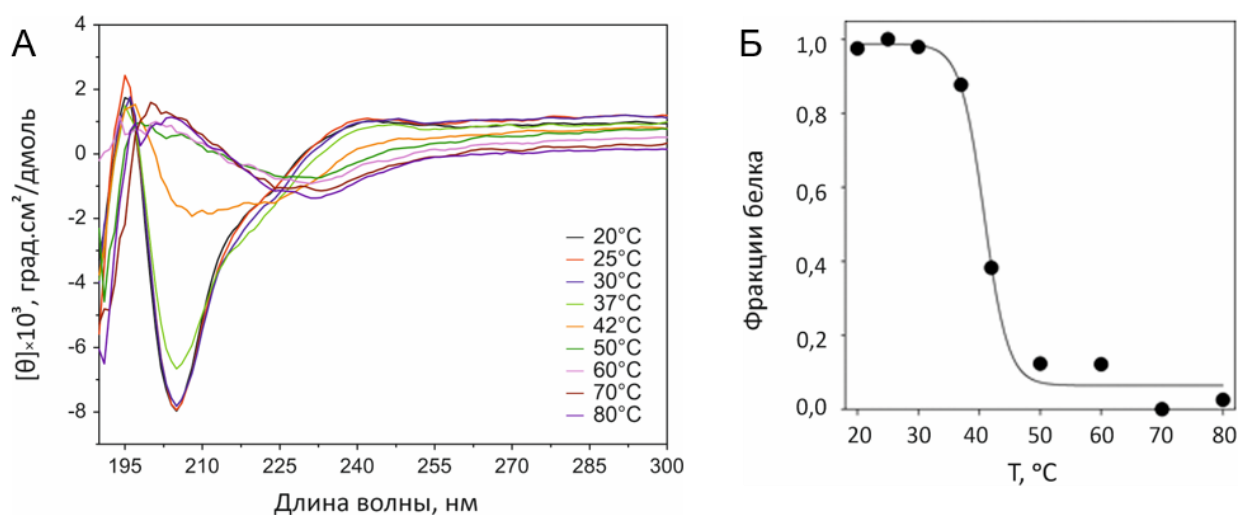


Рисунок 10. Анализ плавления белка с помощью кругового дихроизма. **А**, КД спектры hNEIL2 в диапазоне температур 20-80 °С. **Б**, доля свернутого белка, рассчитанная из полных спектров для модели с двумя состояниями с использованием алгоритма выпуклых ограничений.

Сведения о конформационной стабильности hNEIL2 были дополнительно получены методом дифференциальной сканирующей флуориметрии (рис. 11). Анализ спектров флуоресценции триптофана (Trp) моделью двух состояний показали $T_m = 47,6$ °С. N-концевой домен содержит пять остатков Trp, а С-концевой – три. Полученная T_m может свидетельствовать о большей стабильности N-концевого домена, несмотря на большую неструктурированную часть, по сравнению с С-концевым.

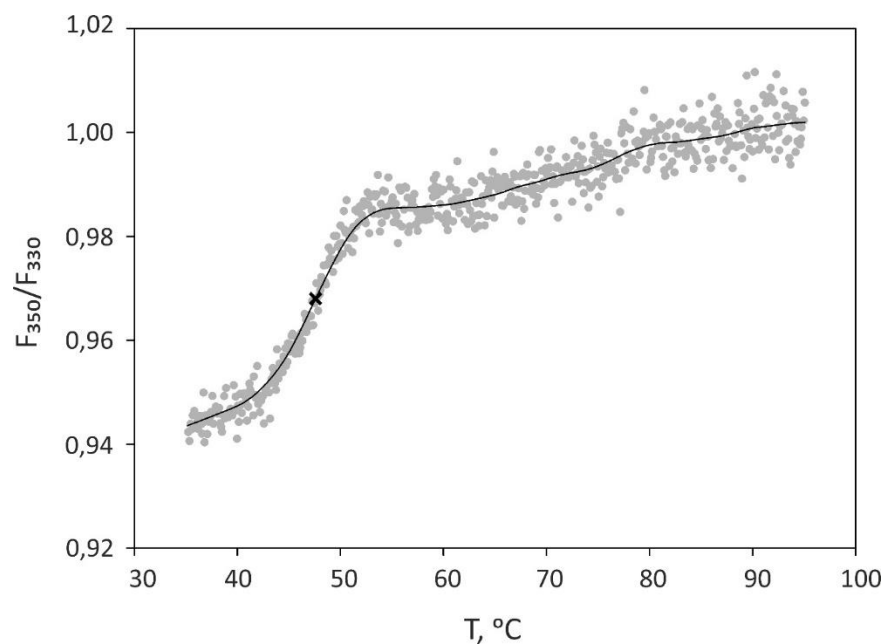


Рисунок 11. Профиль плавления hNEIL2 анализировали с помощью флуоресценции Tnp. F350/F330, отношение эмиссии при 350 нм и 330 нм. Tm отмечен X

4.1.3. МД-анализ конформации hNEIL2

На сегодняшний день в базе данных PDB трехмерная структура hNEIL2 отсутствует. Для интерпретации полученных экспериментальных данных необходима полноразмерная модель белка. Поэтому моделированием по гомологии были получены теоретические структуры, одна из которых основана на «открытой» модели oNEIL2, другая — на «закрытой» mvNEIL2/3 (рис. 12).

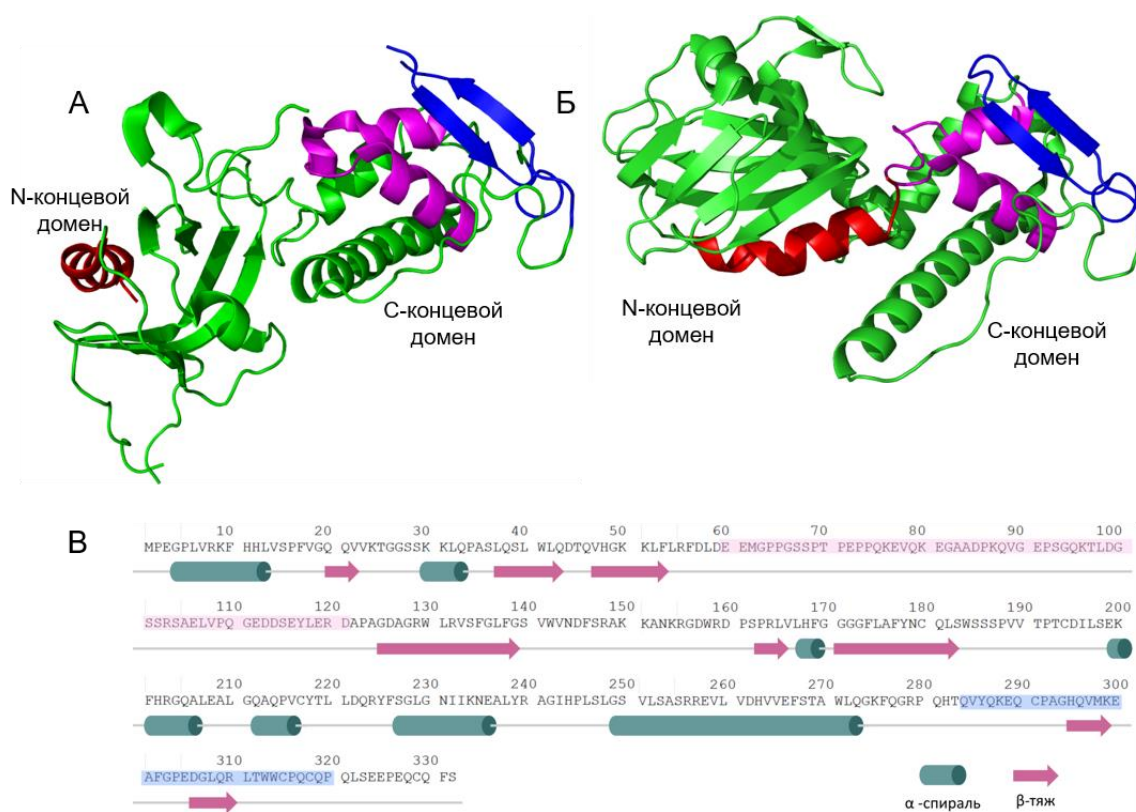


Рисунок 12. **А**, структура oNEIL2 (PDB ID 6VJI), «открытая» модель. **Б**, структура mvNEIL2/3 (PDB ID 4MB7), «закрытая» модель. **В**, Схема структурных элементов белка hNEIL2. Цилиндром показаны α -спирали, стрелкой – β -тяги. Синим цветом выделены аминокислоты, образующие цинковый палец, розовым показана неупорядоченная петля.

В моделях NEIL2 пограничные остатки неструктурированной петли удаляли, поскольку в структуре oNEIL2 они направлены от N-концевого домена, что свидетельствует об отсутствии взаимодействия петли с остальной частью белка. Для каждой модели МД выполняли дважды со случайными начальными значениями для оценки воспроизводимости результатов. Системы уравнивались в течение 80 нс, а последние 20 нс использовали для анализа. Несмотря на то, что уравновешенные траектории обладают смещенными значениями междоменных расстояний от исходных моделей, они были четко разделены по открытой-закрытой конформации (рисунок 14).

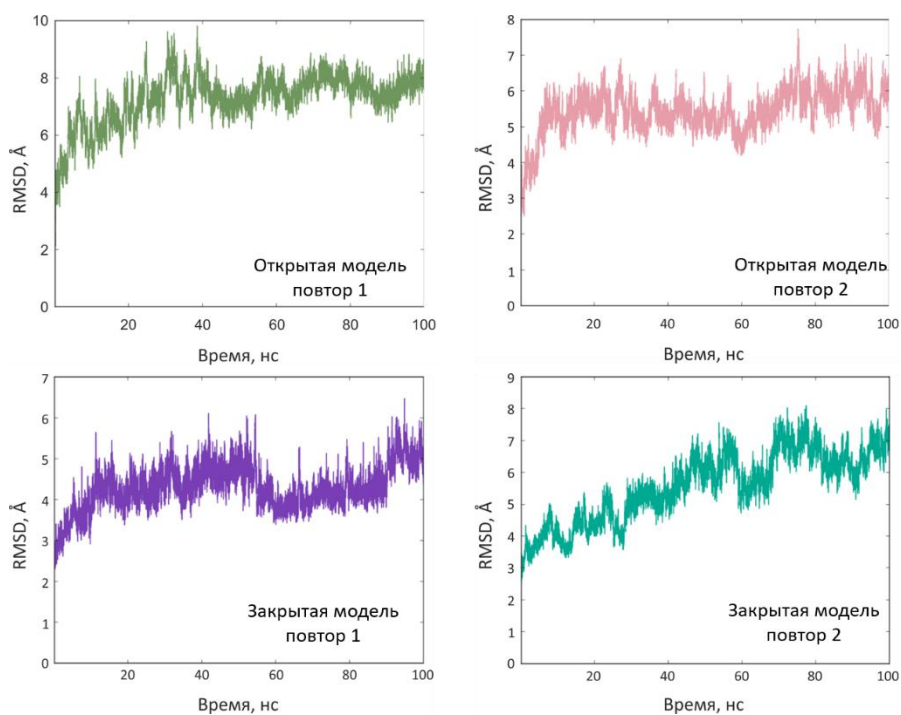


Рисунок 13. Флуктуации RMSD белка hNEIL2 на протяжении всей траектории молекулярной динамики.

Наложение относительного уровня дейтерия для каждого пептида на теоретические структуры дало общее качественное соответствие между эффективностью поглощения и стабильностью области белка при МД. Первая область берет начало с N-конца и продолжается до пептида 46–59. Этот участок белка содержит N-концевую α -спираль с каталитическими остатками Pro-1 и Glu-2 и часть β -сэндвич-домена вплоть до неструктурированной петли $\beta 3/\beta 4$ (рис. 12B). Неструктурированная петля образована пептидами, включающими остатки 56–129, и демонстрирует быстрое поглощение дейтерия. Быстрый H/D обмен наблюдается примерно до 196 аминокислотного остатка, включая край N-концевого домена, междоменный линкер и начало C-концевого домена. Остальная часть C-концевого домена, по-видимому, хорошо свернута и в основном показывает медленную кинетику обмена.

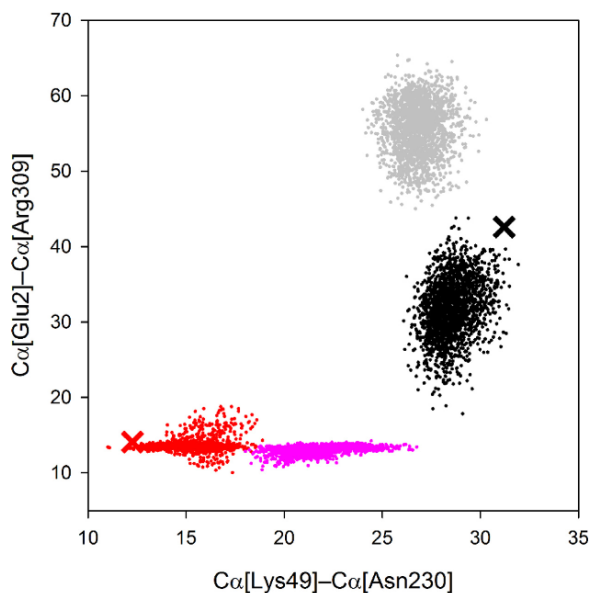


Рисунок 14. Расстояние между $C\alpha$ атомами Lys49 и Asn230, а также между Glu2 и Arg309, наблюдаемы в течение последних 20 нм МД.

Благодаря своей способности генерировать конформационные ансамбли, подходящие для анализа статистическими методами, МД хорошо подходит для оценки отдельных водородных связей, образованных амидными группами белкового остова. Для сопоставления данных МД с результатами HDX использовали заселенность водородных связей в качестве основного инструмента для детального анализа полученных траекторий, в частности для разделения открытой и закрытой конформации hNEIL2 (рис. 14).

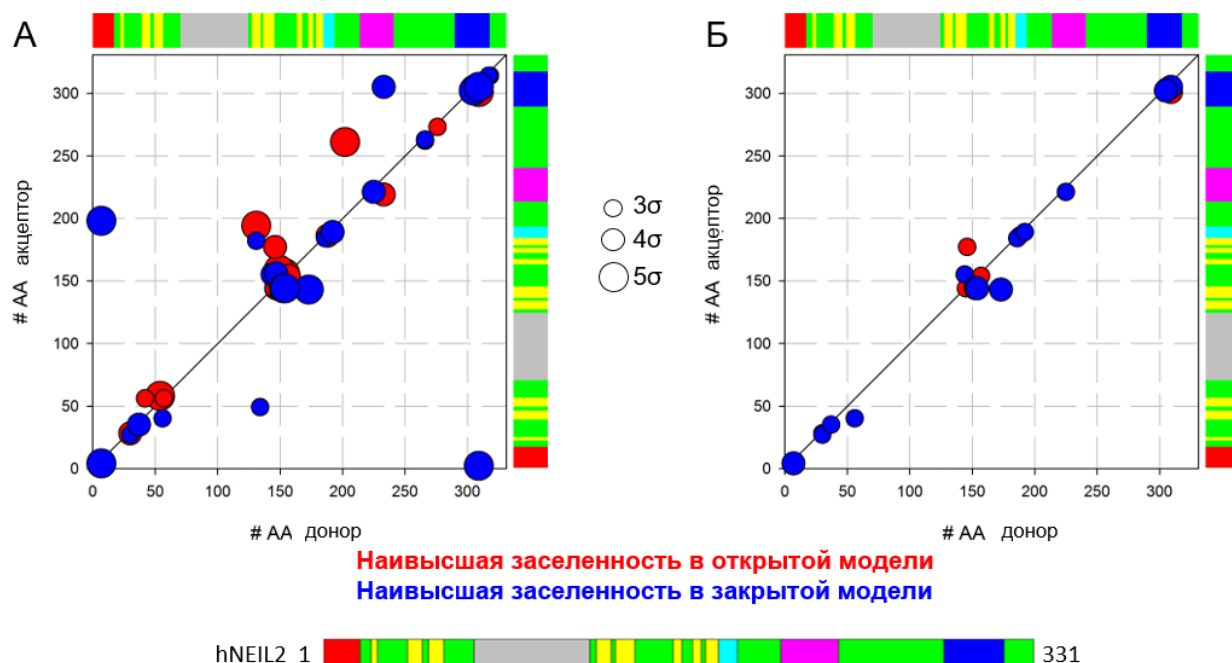


Рисунок 15. Локально перегруппированные водородные связи hNEIL2 в MD. Наибольшие различия в заполнении всех внутримолекулярных водородных связей (А) и амидных водородных связей (Б) между открытой и закрытой конформациями. Цветные полосы схематически изображают ключевые структурные элементы hNEIL2. Красным показана спираль на N-конце, розовым - H2TH мотив, синим - цинковый палец β-сэндвич – желтым, цианом – безцинковый палец. Серым цветом обозначены участки, соответствующие неупорядоченным участкам hNEIL2.

Независимые эксперименты показали следующие межсерийные корреляции: для открытой модели $r = 0,833$ для всех водородных связей, $r = 0,846$ для амидных водородных связей основной цепи и для закрытой модели $r = 0,803$ и $r = 0,809$ соответственно. Чтобы понять в каких областях происходят локальные перестройки в открытом и закрытом состояниях, усреднили заполнение каждой отдельной водородной связи в обоих повторах для каждой модели и выбрали те, которые демонстрируют наибольшую разницу между двумя конформациями. Наиболее заметные изменения в

характере водородных связей наблюдали в петлях $\beta 1/\beta 2$ и $\beta 5/\beta 6$ в N-концевом домене, междоменном линкере и цинковом пальце (рис. 12B). Закрытая модель продемонстрировала четкие междоменные взаимодействия с участием N-концевой спирали, цинкового пальца и части C-концевого домена около линкера.

Наибольшей информативностью для различения преобладающей конформации hNEIL2 в растворе обладают остатки с наибольшей разницей в доступности амидного водорода. Амиды остова, менее вовлеченные во внутримолекулярные водородные связи в предпочтительном состоянии, должны обмениваться быстрее. Для данного анализа усреднили доступность амидов основной цепи в обоих повторах МД для каждой модели, отобрали 20% остатков с наибольшей доступностью амида в открытом состоянии и 20% остатков с наибольшей доступностью амида в закрытом состоянии. Экспериментально определенные доли поглощения для каждого амида основной цепи в каждый момент времени усредняли по всем пептидам, содержащим этот остаток, и рассчитывали коэффициенты корреляции между поглощением и доступностью амида для каждого момента времени. Наилучшую корреляцию с экспериментальными данными продемонстрировали амиды основной цепи открытой модели. Таким образом, данные HDX лучше согласуются с открытым состоянием hNEIL2, преобладающим в растворе.

Наиболее необычным структурным элементом NEIL2 является большая, преимущественно неупорядоченная вставка между $\beta 3$ и $\beta 4$ тяжами в N-концевом домене (остатки 59–129 в hNEIL2, рис. 12B). Пептиды из этой области продемонстрировали очень быстрый H/D обмен, свидетельствующий о свободном доступе воды на амиды основной цепи, что согласуется с рентгеновской структурой oNEIL2, где электронная плотность в соответствующей области не наблюдается. Примечательно, что присутствие этой петли, по-видимому, не дестабилизирует N-концевой домен по отношению к остальной части белка. Несмотря на неупорядоченность и в целом низкую консервативность, наличие петли $\beta 3/\beta 4$ у всех амниот указывает на некоторую функцию этого структурного

элемента. Одним из вариантов является его вовлечение в белок-белковые взаимодействия. hNEIL2 может быть связан с другими ферментами BER или с РНК-полимеразой II и факторами транскрипции YB-1 и NF-κB [57,58]. Например, показано, что взаимодействие NEIL2 с POLβ, LIG3α и YB-1 происходит через N-концевой домен. Также можно предположить, что неупорядоченная область NEIL2 может быть вовлечена в разделение фаз жидкость-жидкость, явление, которое все чаще считается очень важным для динамической организации внутриядерных процессов, включая передачу сигналов повреждения ДНК и восстановление.

4.2. *SpCas9* и *SpCas9*-РНК

Эндонуклеаза Cas9 - важный компонент инструментов геномного редактирования на основе CRISPR-Cas. Достижение высокой специфичности и эффективности Cas9 при направленном расщеплении ДНК является основной проблемой, ограничивающей клиническое применение системы CRISPR-Cas9. Глубокое понимание механизма Cas9 и его структурно-функциональных взаимосвязей необходимо для разработки стратегий точного редактирования генов.

Функцией эндонуклеазы Cas9 из *Streptococcus pyogenes* является защита бактериальных клеток от чужеродной ДНК. В сочетании со специфической направляющей РНК (sgРНК) Cas9 создает двухцепочечный разрыв в ДНК-мишени рядом с мотивом PAM, прилегающим к протоспейсеру (50-NGG в *S. pyogenes*). Полученный разрыв может быть дополнительно отредактирован одним из двух путей репарации ДНК: гомологичной рекомбинацией (HDR) или негомологичным соединением концов (NHEJ). Клинические применения биологии CRISPR ограничены необходимостью снижения нецелевой активности эндонуклеаз Cas. Стратегии повышения специфичности системы включают разработку «высокоточных» вариантов Cas9. Успех в решении этих задач напрямую связан с четким пониманием отношений структура-функция в комплексе Cas9-sgРНК-ДНК.

4.2.1. Компьютерное моделирование apoCas9 комплекса Cas9- РНК.

На сегодняшний день в базе данных PDB для *Streptococcus pyogenes* содержится более 30 кристаллических структур белка Cas9 и его комплексов. Однако особенностями данных структур является то, что, либо они разрешены не полностью, либо структуры представлены для усеченной последовательности белка. Для сопоставления поведения белка в виде апофермента или комплекса, описанного методом HDX-MS с данными компьютерного моделирования необходимо иметь структуру полноразмерного белка. По этой причине мы провели компьютерное моделирование по гомологии с использованием сервера Phyre2 [59] в результате которого наиболее подходящей структурой была предложена структура Cas9 из *Streptococcus pyogenes* с PDB ID: 4CMQ [60]. Неразрешенные участки структуры дополняли с помощью Modeller [47], а ионы Mg^{2+} помещали в структуру белка вручную. Поскольку исходная кристаллическая структура из базы данных не содержала нуклеиновой кислоты, для получения комплекса Cas9-sgРНК и Cas9-sgРНК-ДНК начальные координаты для нуклеиновых кислот брали из кристаллической структуры комплекса Cas9-sgРНК-ДНК с идентификатором PDB 5CZZ [61]. Последовательности РНК и ДНК модифицировали с помощью UCSF Chimera 1.15 [62] и LEaP (Amber20) [63]. Полученная модель Cas9-РНК содержала 1360 аминокислот, 85 нуклеотидов и три иона Mg^{2+} , а модель комплекса с ДНК дополнительно 20 нуклеотидов. Модель комплекса отличалась от используемой в экспериментах длиной молекулы РНК на 18 нуклеотидов. В теоретической модели РНК свободные концевые участки были сокращены, так как они вносят дополнительные возмущения в систему, однако все ключевые фрагменты сохранялись. Структура комплекса фермент-РНК-ДНК приведена на рисунке 16.

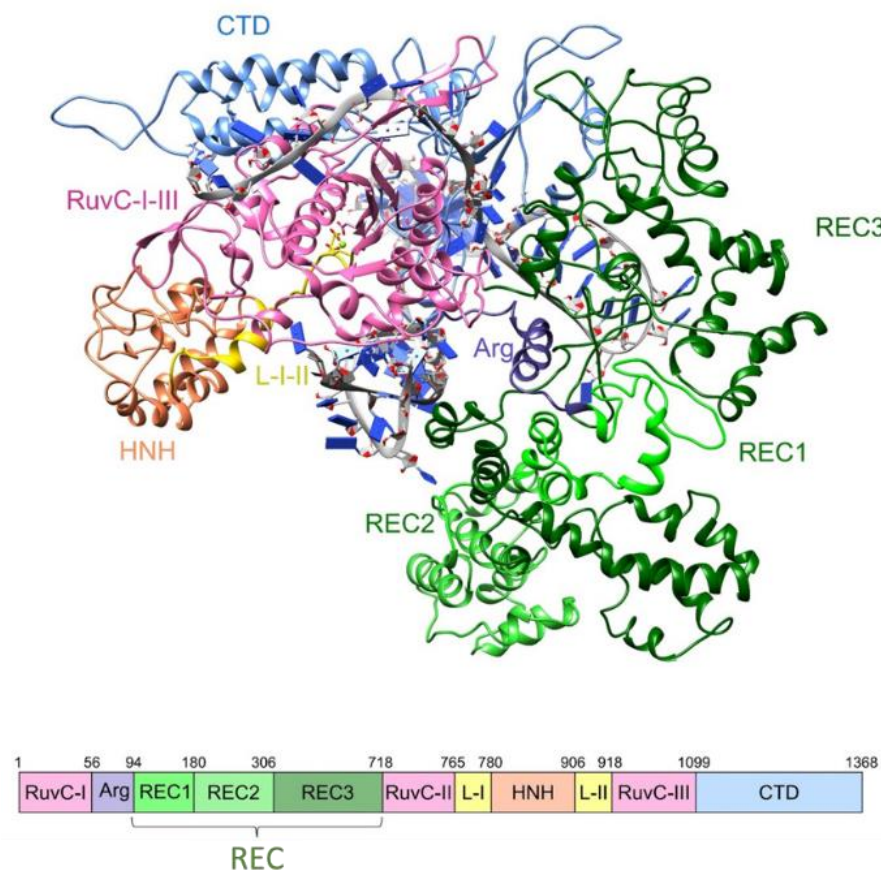


Рисунок 16. Структура Cas9 из *S. pyogenes* в комплексе с sgРНК-ДНК (PDB ID: 4CMQ). Каждый домен белка окрашен согласно диаграмме доменной организации внизу.

Анализ молекулярно-динамических траекторий показало стабильность белка в форме апофермента и в составе комплекса. Конформационную динамику белка количественно определяли с помощью значений RMSD (среднеквадратичное отклонение от референсной структуры по Ca атомам белкового остова). На рисунке 17А показаны значения RMSD для основных доменов Cas9. Так, апофермент на протяжении всей 50 нс МД показал стабильность, все домены белка достигали равновесия быстро, в течение ~5 нс, и колебались в пределах 1 Å на протяжении остального времени динамики.

В комплексе Cas9 с РНК эволюция значений RMSD во времени для всех доменов, за исключением REC lobe, показала хорошую стабильность (рис. 17Б). Для домена REC характер колебаний после 18 нс скачкообразно возрос до 8 Å. Максимальные флуктуации

компонентов доли NUC находятся в пределах 6,2 Å. В частности, среднее значение RMSD доменов RuvC и HNH составило ~ 4,0 Å и ~ 6,2 Å соответственно. В течение 8 нс флуктуации С-концевого домена было стабильным и находилось в районе ~4,8 Å, а затем увеличилось до ~6,2 Å. В литературных данных описан сдвиг REC3 на ~65 Å в сторону домена HNH при связывании Cas9 с sgРНК, что говорит о лабильности данного домена.

Добавление в систему ДНК-субстрата повлияло на характер профиля RMSD некоторых доменов (рис. 17В). Так, домен HNH и α -спиральный Arg-мост продемонстрировали более высокий уровень стабильности во время МД симуляции по сравнению с комплексом Cas9-sgРНК. Максимальные флуктуации компонентов доли NUC находятся в пределах 7 Å. В частности, среднее значение RMSD для RuvC и HNH составило ~5 и ~3,4 Å соответственно. Домены REC и CTD продемонстрировали более низкий уровень стабильности. Молекулярная динамика REC описывалась постепенным увеличением RMSD до ~11,5 Å. При изучении значений RMSD отдельных компонентов доли REC (REC1, REC2 и REC3), было обнаружено, что эти домены колеблются по-разному. Средние значения RMSD доменов REC1 и REC2 составили ~7,2 и ~5,1 Å соответственно. Для сравнения, среднеквадратичное отклонение REC3 достигло ~10–11 Å, что существенно способствовало наблюдаемой нестабильности доли REC.

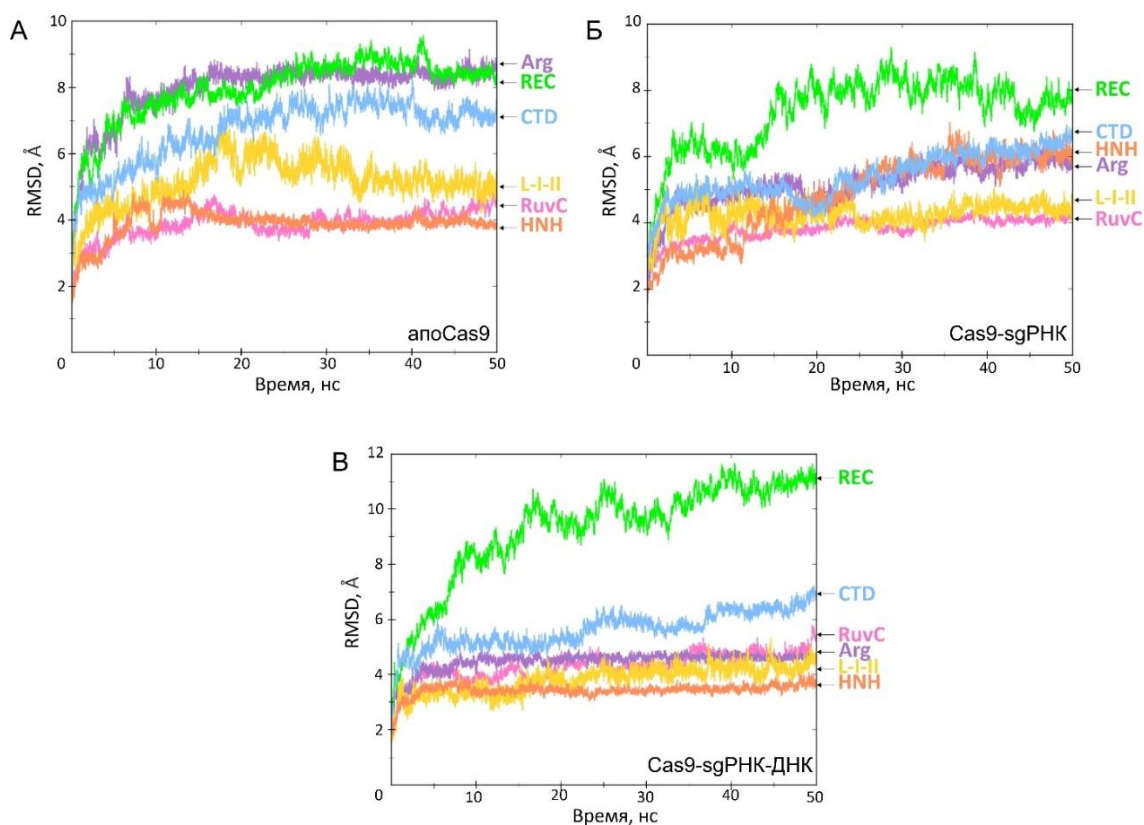


Рисунок 17. Значения RMSD (в ангстремах), рассчитанные для 50 нс МД траекторий для апоCas9 (А), его комплекса с sgPHK (Б) и тройного комплекса Cas9-sgPHK-ДНК (В). Зеленым показан домен REC, голубым - CTD, RuvC – розовым, L-I-II - желтым, Arg – фиолетовым и оранжевым – HNH домены Cas9.

Ионы Mg^{2+} в структурах находятся в двух доменах: RuvC и CTD. Положение ионов металла оставалось постоянным на протяжении всей траектории МД (рис. 18). В домене RuvC два иона магния координировались Asp-10, Glu-762, Glu-766, Asp-986 и His-983 на равных расстояниях ($\sim 2 \text{ \AA}$). Во время молекулярной динамики расстояние между аминокислотным остатком Glu-762 и вторым ионом Mg^{2+} сохранялось в области 2 \AA , однако координация происходила разными атомами кислорода. В то время как один из атомов кислорода координировал ион металла, другой атом отдалялся от иона на $\sim 4 \text{ \AA}$, то есть в процессе МД происходило «переключение» между двумя кислородными атомами Glu-762. Третий ион Mg^{2+} координировался карбоксилатными группами Asp-1299, Glu-1304, Glu-1307 и Asp-1328, находящимися в α -спирали. Координация данного иона

боковой цепью Glu-802 домена HNH, как и в случае второго иона Mg^{2+} , происходила со сменой координирующих атомов кислорода. Стоит отметить, что в течение последних 10 нс расстояние между карбоксильной группой Asp-1328 и Mg^{2+} увеличилось до ~ 4 Å. Стабильность положения ионов металлов при моделировании демонстрирует актуальность всей представленной модели комплекса.

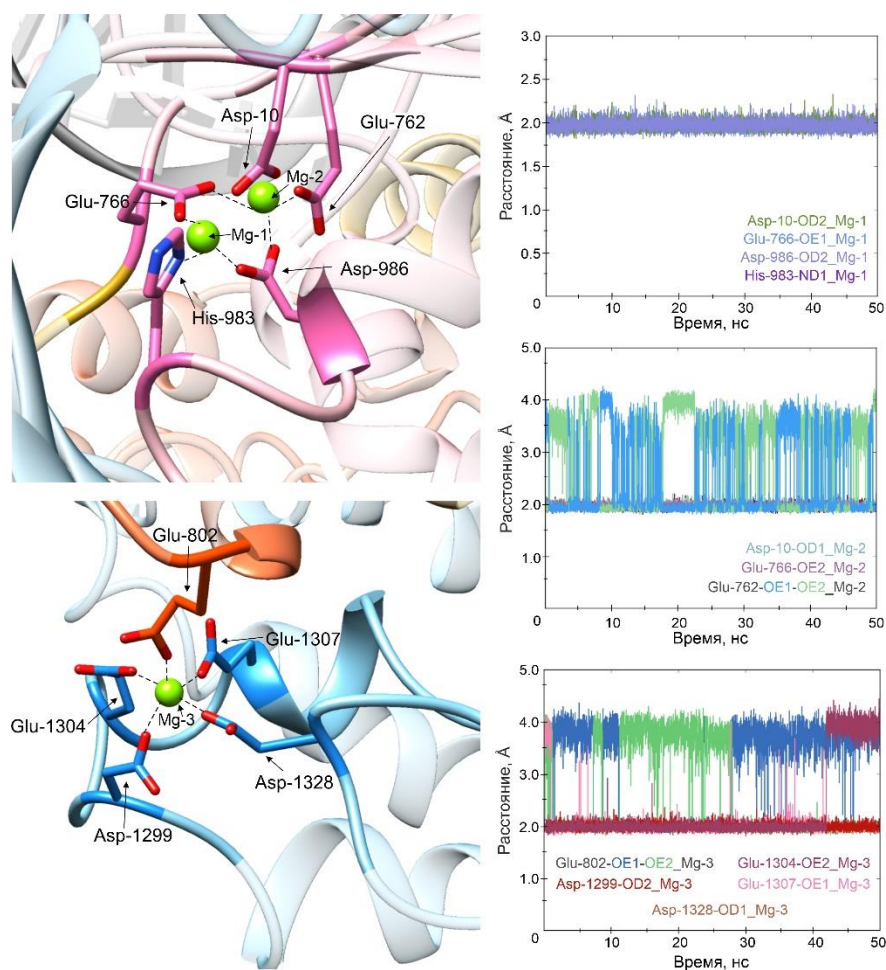


Рисунок 18. МД-моделирование координационных центров металлов в комплексе Cas9-sgРНК-ДНК. **А**, координация ионов магния ($Mg-1$, $Mg-2$, зеленые сферы) аминокислотными остатками домена RuvC (розовый). **Б**, локализация Mg^{2+} ($Mg-3$, зеленая сфера) в домене CTD (голубой). Координирующий остаток Glu-802 принадлежит домену HNH (оранжевый). Графики справа демонстрируют колебания расстояния между металлом и функциональными группами координирующих его аминокислот в ходе МД.

4.2.2. HDX-MS анализ динамики SpCas9 и SpCas9-PHK

Пептидная карта белка – это пул пептидов, покрывающих последовательность исследуемого объекта, которые получены в ходе протеолитического расщепления. Для установления структуры белка Cas9 методом HDX-MS мы в первую очередь составили пептидную карту в условиях аналогичных реакции обмена. Для этого раствор белка уравнили при комнатной температуре в течение 1 ч, аликвоту добавили в буфер для остановки процесса обмена. Расщепление белка проводили на колонке с иммобилизованным пепсином, а полученные пептиды разделяли на аналитической колонке и анализировали масс-спектрометрическим методом. Покрытие последовательности Cas9 после анализа пептидов составило 64 %. Полученную карту использовали для количественного анализа данных H/D обмена.

Реакцию обмена водорода на дейтерий проводили для апоCas9 при комнатной температуре, останавливая в определенные временные точки начиная от 10 с, заканчивая 8 ч путем добавления аликвоты реакционной смеси в охлажденный «гасящий» буфер. Фракции белка анализировали масс-спектрометрически после протеолитического расщепления пепсином. Полученные спектры на первом этапе обрабатывали с помощью программы HDExaminer, затем данные были обработаны вручную для повышения точности результатов. На рисунке 19 приведена тепловая карта наблюдаемого H/D обмена. На карте уровень обмена показан цветами от темно-синего до красного. Низкий уровень обмена, оттенки синего, соответствует 10-30 % включенного в остов дейтерия, 31-70 % соответствует среднему уровню обмена и показан градиентом от зеленого до темно-желтого, оттенками оранжевого и красным показаны высокие значения обменного дейтерия в амидах, 71-100 %. На характер (скорость) обмена водорода на дейтерий в белковом остове влияет природа элементов вторичной структуры и их относительное расположение в пространстве. Исходя из этого мы разделили полученные

пептиды на четыре группы. Пептиды с низким уровнем обмена (<30 %) отнесли к первой группе. Данные участки структуры расположены внутри белковой глобулы и амидные протоны экранированы от растворителя. В случае апофермента наибольшее число таких пептидов обнаружено в доменах RuvC III, REC2,3 и CTD, а также по одному пептиду - в доменах RuvC I,II и REC1. Вторую группу образуют пептиды доменов RuvCII-III, REC1-3, CTD и HNH, продемонстрировавшие средний уровень обмена, который со временем практически не изменился. Третью группу составили пептиды с быстрым, высоким и стабильным в течение 8 ч реакции уровнем включения дейтерия. Обычно такое поведение характерно для неструктурированных участков структуры белка. В нашем случае в данную группу вошли петли и α -спирали из домена REC2 и линкера L-II. В четвертую группу мы включили пептиды обмен в которых происходил постепенно, что присуще участкам белка с выраженной вторичной структурой. Пептиды этой группы находятся в доменах REC2, REC3, RuvC III и CTD, в линкеры L-II и Arg и расположены на поверхности РНК-связывающей области.

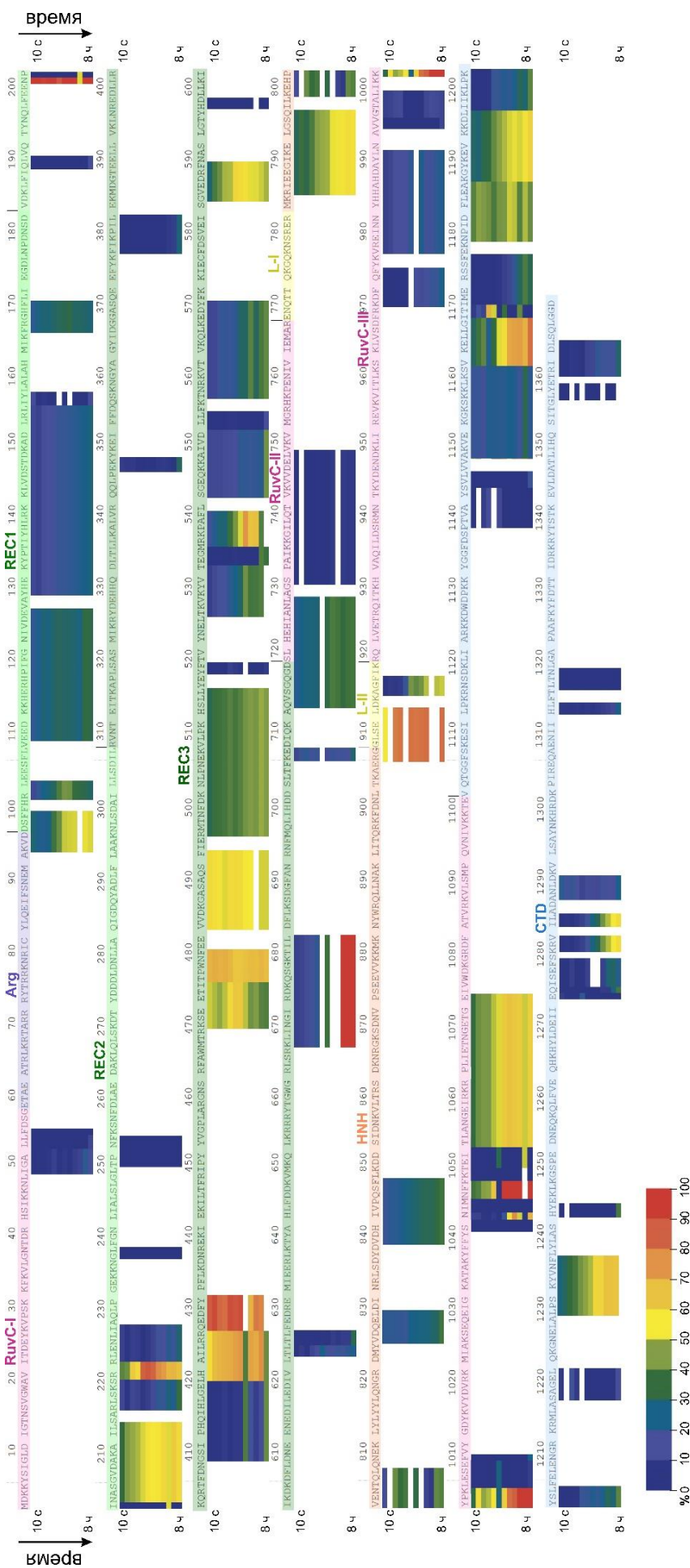


Рисунок 19. Карта дейтерирования apoCas9. Столбцы окрашены в соответствии с относительным процентом включения дейтерия, как указано в условных обозначениях, от темно-синего до красного цвета. Поглощение дейтерия от 10% до 30% считалось «низким», поглощение от 31% до 70% — «средним», поглощение от 71% до 90% — «высоким». Зависимый от времени рост уровня поглощения рассматривался как «постепенный». Для каждого пептида уровень дейтерирования рассчитывается для каждого из экспериментальных периодов инкубации HDX и показан стрелками слева и справа от каждого белкового сегмента: 10 с, 30 с, 1, 2, 5, 10, 30 мин, 1, 2, 4, 6 и 8 ч.



Рисунок 20. Карта дейтерирования комплекса Cas9-sgРНК. Столбцы окрашены в соответствии с относительным процентом включения дейтерия, как указано в условных обозначениях, от темно-синего до красного цвета. Поглощение дейтерия от 10% до 30% считалось «низким», поглощение от 31% до 70% — «средним», поглощение от 71% до 90% — «высоким». Зависимый от времени рост уровня поглощения рассматривался как «постепенный». Для каждого пептида уровень дейтерирования рассчитывается для каждого из экспериментальных периодов инкубации HDX и показан стрелками слева и справа от каждого белкового сегмента: 10 с, 30 с, 1, 2, 5, 10, 30 мин, 1, 2, 4, 6 и 8 ч.

Изучение структурно-динамических особенностей взаимодействия Cas9 с направляющей РНК проводили с помощью HDX-MS, как описано ранее. Для формирования комплекса белок-РНК компоненты смешивали непосредственно началом реакции обмена в мольном соотношении 1:10. Количество обнаруженных пептидов было ниже, чем в случае апофермента (рис. 20). Для классификации пептидов также использовали четыре группы, а затем сравнивали поведение H/D обмена в апоCas9 и комплексе Cas9-РНК. Различия в поглощении дейтерия одними и теми же пептидами в разных состояниях белка свидетельствовали о значительном изменении конформации некоторых доменов в процессе связывания РНК ферментом.

Поскольку используемый подход HDX-MS на сегодняшний день не позволяет получить структуру белка с атомарным разрешением, то наилучшим способом интерпретации данных является наложение результатов эксперимента на доступные в базах данных кристаллические структуры или на теоретически полученные модели. Для анализа конформационных изменений в структуре белка измеренное относительное поглощение дейтерия за 1 и 480 минут мы спроецировали на смоделированные структуры апоCas9 и его комплекса с РНК. На ряду с этим, результаты HDX мы сопоставили с кристаллическими структурами Cas9 (4CMQ) и Cas9-sgРНК (4ZT0). Незавершенные структуры, взятые из PDB, дополнили с помощью Modeller. На рисунке 21 показаны структурные элементы SpCas9. Результаты сравнения показали хорошую корреляцию между результатами обмена и трёхмерной моделью белка. Так, α -спиральный междоменный линкер Arg находится в тесном контакте с направляющей РНК в сформированном комплексе. По данным HDX пептид 92-97, находящийся в данном линкере, в процессе формирования комплекса практически не подвержен обменным процессам, когда в апоферменте обмен протекает постепенно. Такое поведение говорит о том, что sgРНК экранирует α -спиральный участок от растворителя. Это же подтверждается уменьшением значений RMSD во время молекулярной динамики в

структуре апоCas9 и Cas9-РНК. Линкерный участок L-II представлен петлей и α -спиральным фрагментом. Пептид 905-911 в апоCas9 расположен на поверхности белка, в то время как в комплексе он окружен неупорядоченными частями доменов RuvC и HNH. Соответственно скорость обмена в связанном состоянии белка у пептида из L-II снижается, что коррелирует с литературными данными. Так, линкер L-II по данным крио-ЭМ тройного комплекса Cas9 с РНК-ДНК дуплексом претерпевает конформационные изменения для облегчения смещения домена HNH [64].

При связывании фермента с направляющей РНК конформация доменов RuvC II и III становится менее упорядоченной, о чем говорит обнаруженное повышение уровня поглощения пептидов. Данные МД показывают наличие взаимодействия α -спирали 719-727 с нуклеотидами G11-G14 РНК, которые трансформируют данную спираль в петлю, что способствует увеличению дейтерирования. В домене HNH апоCas9 содержится четыре пептида, относящихся к группе среднего уровня обмена. Взаимодействие белка с sgРНК для данных пептидов увеличивает их доступность для растворителя.

Пептиды доли REC имеют хорошее покрытие последовательности, а наибольшее количество обменивающихся пептидов зарегистрировано в REC3 домене. Относительная конформационная стабильность проявляется у домена REC1, где низкий и средний уровни H/D обмена в пептидах практически не изменяется во время формирования комплекса. Пептиды домена REC2 в процесс обмена вступают неохотно и среди них преобладает низкий уровень дейтерирования. Однако пептид 202-212 в составе апофермента подвергался обмену постепенно (40-60 %), располагаясь на поверхности, а при взаимодействии белка с РНК характер обмена изменился на более низкий (30-50 %). Согласно результатам молекулярной динамики, наибольший вклад в значение RMSD Cas9 вносит домен REC3. В свою очередь HDX-MS эксперименты показали широкий спектр пептидов с различным уровнем обмена водорода на дейтерий, многие из которых меняли доступность для растворителя после процесса комплексообразования. Так, α -спираль 532-

539 и петлевые структуры 426–430, 494–514, 532–539, 582–587 в апоферменте показали средние и высокие уровни H/D обмена, а в связанном состоянии уровни обмена снижались. Данные структурные элементы контактируют с РНК и были погружены внутрь белковой глобулы. Пептиды 465–477, 622–623, 665–680 и 712–718 наоборот, показали повышение уровня обмена дейтерия в комплексе. Первый пептид вступает в межмолекулярные контакты с 56–58 нуклеотидами РНК, приближаясь к поверхности белка. Пептид 712-718, соединяющий доли REC и NUC, в несвязанном состоянии обладает средним уровнем обмена, но при связывании трансформируется в пептид с высоким и быстрым обменом. В теоретической структуре данный участок белка в процессе связывания РНК смещается на поверхность глобулы. Таким образом, как данные HDX, так и данные молекулярной динамики подтверждают высокий уровень подвижности домена REC3 при взаимодействии Cas9 с sgРНК.

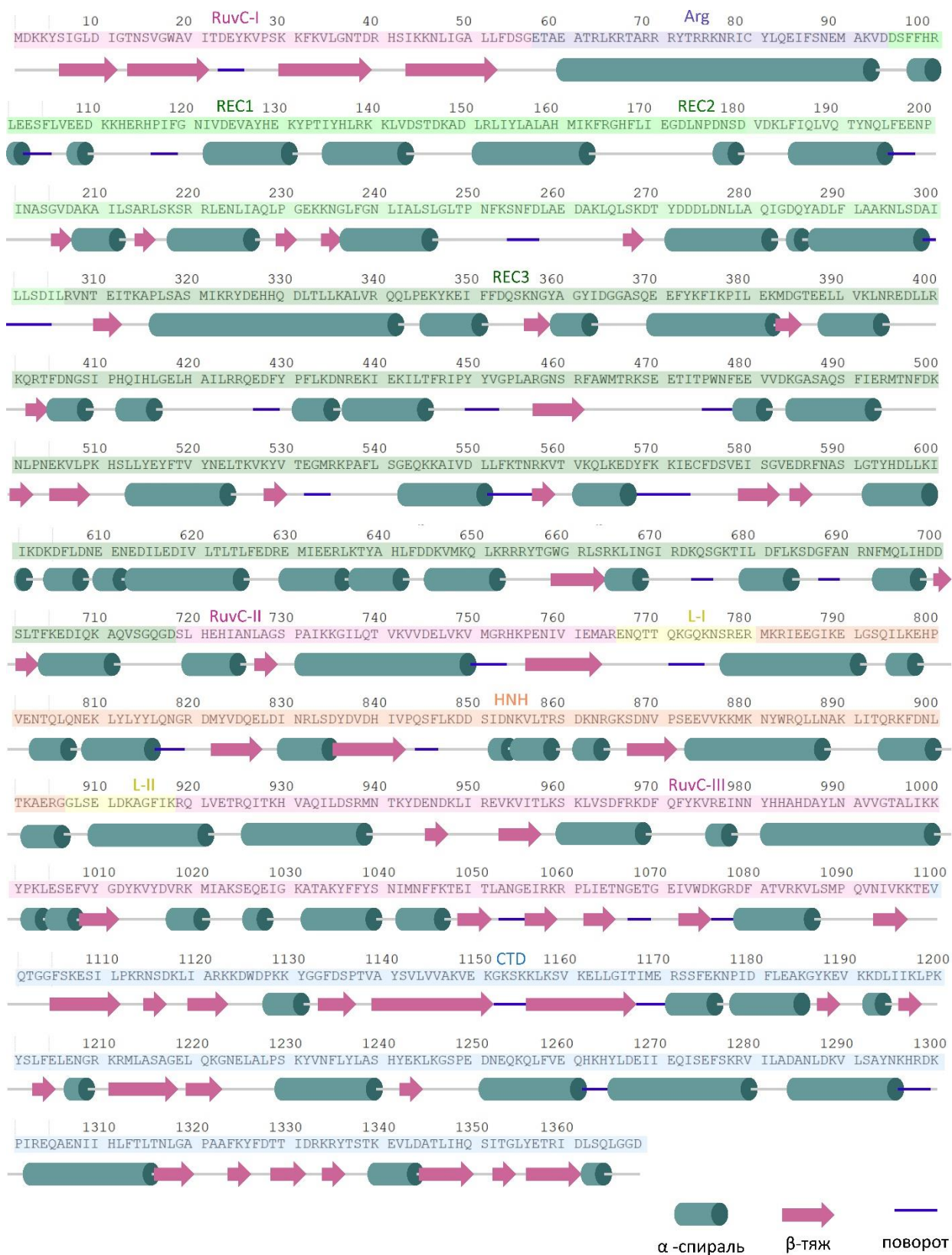


Рисунок 21. Схема структурных элементов SpCas9. Цилиндром показаны α-спирали, стрелкой – β-тяжи, синим цветом обозначены повороты. Цветом показаны домены:

зеленым - REC, голубым - CTD, RuvC – розовым, L-I-II - желтым, Arg – фиолетовым и оранжевым – HNH.

C-концевой домен в основном расположен на поверхности белковой глобулы, а в его структуре преобладают петли, соединенные короткими мотивами α -спиралей и β -листов. Такая топология и структура предопределяет эффективный обмен водорода на дейтерий растворителя. Однако экспериментальные данные HDX для апоCas9 показали, что для большинства пептидов этого домена характерен низкий уровень обмена. Все стадии HDX-MS эксперимента проводили так, что вероятность обратного обмена была минимальна, а различие в структурах может быть обусловлено кристаллической и растворенной формой белка. Причина пониженного обмена может заключаться в белок-белковом взаимодействии или олигомеризацией белков в растворе, однако для подтверждения этого нужны дополнительные эксперименты. Снижение скорости обмена в экспонированных областях, может быть обусловлено присутствием структурированных молекул воды вблизи аминокислотного остова. При образовании комплекса Cas9 с РНК конформация CTD домена заметно перестраивалась. Так, существенно возрастал уровень обменного дейтерия в пептидах 1147–1159, 1271–1276, и со средней эффективностью – в пептидах 1177–1185, 1315. –1317, 1358–1362. Наряду с этим пептид 1160-1165 (β -лист) показал понижение в эффективности обмена, что согласуется с моделью комплекса, где он участвует в координации sgРНК.

5. Заключение

Метод водородно-дейтериевого обмена с последующей масс-спектрометрией является мощной альтернативой широко распространенным методам биофизики в вопросе определения трехмерной структуры белков и их комплексов. Благодаря проведению реакции H/D обмена в растворе, можно наблюдать за эволюцией процессов комплексообразования в течение времени, что не доступно для большинства существующих методов. Наряду с этим для экспериментов требуется относительно небольшое количество образца, несколько десятков пикомоль, в то время как для ЯМР-спектроскопии необходимо несколько миллиграмм белка.

С помощью комбинации метода HDX-MS с методами компьютерного моделирования нами были получены данные о структурно-динамических особенностях некоторых апоферментов и их комплексов.

Первым объектом стал белок человека hNEIL2. Данный белок является участником в пути эксцизионной репарации оснований и относится к классу ферментов ДНК-гликозилаз/АП-лиаз. hNEIL2 проявляет субстратную специфичность к окисленным производным цитозина, предпочитая двухцепочечные молекулы ДНК, содержащие некомплементарные участки, так называемые глазки. Структуру белка NEIL2 опоссума удалось разрешить лишь в 2021 году методом рентгеновской кристаллографии [64]. Особенностью структуры является протяженный неупорядоченный участок между N- и C-концевым доменом белка, который в существенной мере осложняет кристаллизацию фермента. По этой причине белок oNEIL2 авторы кристаллизовали частями. Аналогичным способом разрешить структуру белка hNEIL2 оказалось невозможным.

Мы провели эксперименты методом обмена водорода на дейтерий с последующим масс-спектрометрическим анализом для hNEIL2. Реакцию обмена проводили во временном диапазоне от 10 с до 2 ч в трех повторах. В результате анализа полученных

данных покрытие последовательности белка составило 96,7 %, неразрешенной осталась только область со 176 по 181 аминокислоту. Пролин-богатая неупорядоченная петля, сформированная аминокислотными остатками 56-129, в процессе протеолитического расщепления оставалась практически не тронутой протеазой, но несмотря на это была разрешена. С первых секунд реакции данный фрагмент показал высокий уровень H/D обмена, что характерно для неструктурированных участков структуры белка, таких как петли. Быстрый обмен также наблюдали в петле $\beta 3/\beta 4$ и в линкере 181-196 между двумя концевыми доменами. На графике (рис. 9А) четко видны области, характерные для элементов вторичной структуры. Так, для пептидов С-концевой области постепенное увеличение количества поглощенного дейтерия показало наличие в составе α -спиралей. Спектры некоторых групп пептидов описывались бимодальным распределением, указывающее на кинетику EX1. Пептиды EX1 входят в ядро β -сэндвича в N-концевом домене и в мотив H2TH и его окружению в С-концевом домене.

Неупорядоченный участок в hNEIL2 может повлиять на стабильность белка или его доменов. С помощью экспериментов методами КД-спектроскопии и дифференциальной сканирующей флуориметрии нами было установлено, что белок hNEIL2 можно отнести к группе наименее стабильных белков и его N-концевой домен, несмотря на большую неупорядоченность проявляет большую термостабильность относительно С-конца.

Для уточнения полученной на основании метода HDX-MS структуры фермента было проведено компьютерное моделирование белка hNEIL2. Поскольку для ферментов класса H2TH известны как открытая, так и закрытая конформация, то моделирование проводили для обоих случаев с целью выявления характерного для апофермента состояния в растворе. Подготовленные к молекулярной динамике структуры были получены на основании кристаллических структур из базы данных PDB и не содержали неупорядоченного участка (аминокислоты 58-125). Симуляцию молекулярной динамики проводили в двух повторах для открытого и закрытого состояний фермента. Анализ МД

траекторий и последующее сравнение результатов с данными H/D обмена показали, что наиболее предпочтительным состоянием апофермента hNEIL2 является открытая конформация.

Таким образом, комбинацией методов удалось установить структуру белка hNEIL2 в растворе и показать, что она соответствует открытому состоянию фермента, а также, что среди белков человеческого протеома данный фермент принадлежит к группе наименее термостабильных.

Вторым объектом изучения стал фермент системы геномного редактирования CRISPR/Cas – Cas9. Фермент, полученный из *Streptococcus pyogenes*, связываясь с направляющей РНК (sgРНК), вносит двухцепочечный разрыв в целевую ДНК. С целью детального изучения механизма узнавания и расщепления субстрата в растворе мы применили метод HDX-MS для белка SpCas9 и его комплекса с sgРНК. Поскольку белок SpCas9 (~158кДа) значительно больше фермента hNEIL2 (~37кДа), диапазон времени реакции был выбран от 10 с до 8 ч. Эксперименты проводили в трех повторах как для апо-Cas9, так и для его комплекса с РНК. При изучении фермент-субстратного комплекса методом HDX-MS важно, чтобы весь белок находился в связанном состоянии, поэтому комплекс формировали в мольном отношении белок : нуклеиновая кислота 1:10. В результате протеолитического расщепления недеитерированного белка покрытие последовательности было 64 %. Обработка полученных после обмена спектров показала видимые отличия в структуре несвязанного белка и в процессе формирования комплекса. Пептиды были поделены на четыре группы согласно наблюдаемому характеру обмена. В первую группу были внесены пептиды с низким уровнем обмена (10-30 %), во вторую – со средним (31-70 %), в третью группу относили пептиды с высоким уровнем обмена (71-100 %) и четвертую группу составили пептиды с постепенным увеличением уровня включенного дейтерия. Для более детального понимания структурно-динамических особенностей образования комплексов было проведено компьютерное моделирование

белка и его комплексов с нуклеиновыми кислотами, а проанализированные молекулярно-динамические траектории использовались для сравнения с экспериментальными данными.

Для апофермента наибольшее количество пептидов, труднодоступных для обмена наблюдалось в доменах RuvC III, REC2 CTD. Во вторую группу вошли пептиды, принадлежащие доменам RuvCII-III, REC1-3, CTD и HNH. Характерное для неупорядоченных или наиболее доступных пептидов поведение продемонстрировали линкер L-II и петли и α -спирали из домена REC2. Пептиды последней группы продемонстрировали обмен водорода на дейтерий, присущий элементам вторичной структуры. В данную группу вошли пептиды доменов REC2, REC3, RuvC III и CTD, в линкерах L-II и Arg, относящиеся к РНК-связывающей области. При формировании комплекса Cas9 с РНК характер Н/D обмена изменился. Так, например, пептид линкера L-II продемонстрировал снижение уровня обмена при комплексообразовании, смещаясь с поверхности белка внутрь глобулы. Описанное в литературных источниках смещение домена REC3 в сторону домена HNH, было подтверждено данными HDX-MS эксперимента и результатами МД. Пептид 92-97 Arg-моста в апоферменте постепенно вступал в реакцию обмена водорода на дейтерий, в то время как в комплексе с РНК обменных процессов практически не наблюдалось. Такое существенное изменение в уровне поглощения дейтерия говорит об экранировании данного участка α -спирали молекулой sgРНК. Уменьшение значений RMSD для этого домена в составе комплекса подтверждает полученные данные.

Согласно результатам молекулярной динамики, наибольший вклад в изменение стабильности Cas9 в составе комплекса вносит домен REC и в частности REC3. HDX-MS эксперименты показали, что пептиды данной области относятся к различным группам и многие из них в процессе формирования комплекса меняли характер Н/D обмена. Некоторые пептиды со средним уровнем обмена в апоферменте при связывании белка снижали уровень поглощенного дейтерия. Данные аминокислотные остатки были

погружены внутрь глобулы и некоторые из них контактировали с sgРНК. Другие пептиды наоборот при формировании комплекса приближались к поверхности белка, взаимодействуя с нуклеотидами РНК. По данным HDX-MS экспериментов и симуляции МД домен REC3 показал высокую подвижность при взаимодействии Cas9 с sgРНК.

Домен CTD также показал заметную перестройку при связывании Cas9 с РНК, поскольку для большинства пептидов уровень дейтерирования возрастал. Однако для β -лист, образованного аминокислотами 1160-1165, уровень обменного дейтерия понижался, что показывает его участие в координации sgРНК.

В результате проведенного исследования структурно-динамических особенностей апо-Cas9 и его комплекса с направляющей РНК показаны конформационные изменения, происходящие при формировании подготовленного к расщеплению ДНК-субстрата комплекса. При взаимодействии фермент-РНК происходят значительные изменения в положении REC3 домена, Arg-моста и С-концевого домена.

Таким образом, на примере двух различных ферментов показаны возможности применения метода HDX-MS при изучении структурных и динамических аспектов существования белковой молекулы в растворе в несвязанном состоянии и в процессе связывания с лигандом. Представленные эксперименты демонстрируют возможность детального изучения механизмов узнавания и действия ферментами субстратов, а также отслеживания изменения конформации в различных состояниях молекулы белка вне зависимости от ее размера и особенностей строения.

6. Выводы

1. HDX-MS эксперименты для белка hNEIL2 показали наличие как структурированных участков последовательности, так и неструктурированных. Так, быстрый H/D обмен наблюдали для неупорядоченных петель, образованных 56-129 аминокислотными остатками, а также в случае линкера (181-196) между N- и C-концевыми доменами. В некоторых участках структуры наблюдали кинетику EX1.
2. Моделированием по гомологии получены теоретические модели hNEIL2, отражающие открытую и закрытую конформацию белка в растворе. Для каждой системы проведены расчеты методами МД-моделирования в двух повторах. Показана стабильность системы, а также возможность различить состояния молекулы белка открыто-закрыто. Оценка заселенности водородных связей МД траекторий показала хорошие межсерийные корреляции. В открытой модели наиболее заметные изменения в характере водородных связей показаны для петель $\beta 1/\beta 2$ и $\beta 5/\beta 6$ в N-концевом домене, междоменном линкере и цинковом пальце. Закрытая модель продемонстрировала четкие междоменные взаимодействия с участием N-концевой спирали, цинкового пальца и части C-концевого домена около линкера.
3. Данные, полученные комбинацией методов HDX-MS и МД указывают на то, что в hNEIL2 может преобладать открытое состояние, поскольку экспериментально определенная эффективность обмена лучше коррелирует с доступностью амидов в модели с открытой конформацией.
4. На основании анализа HDX-MS показано значительное изменение конформации некоторых доменов в процессе связывания Cas9 с РНК.
5. Получены теоретические модели апоCas9, Cas9-sgРНК и Cas9-sgРНК-ДНК. Показано, что в процессе МД апоCas9 остается стабильным, в то время как при формировании комплекса с РНК домен REC показал высокую лабильность. Добавление в систему молекулы ДНК повлияло на увеличение стабильности таких доменов, как HNH и

Arg, в то время как домен REC также показал высокую подвижность. Показано, что положение ионов Mg^{2+} на протяжении всех МД траекторий оставалось стабильным. В тройном комплексе Cas9-sgРНК-ДНК показано, что в процессе координации двух ионов магния происходит «переключение» между координирующими атомами кислорода в аминокислотных остатках Glu-762 и Glu-802.

6. С использованием сравнения результатов метода HDX-MS и МД показано, что доступность растворителя в модельной структуре коррелирует с наблюдаемым относительным поглощением дейтерия. В частности, пептиды из доменов RuvC II-III и REC2-3 лежат внутри структуры и окружены другими белковыми мотивами. Пептиды со средним уровнем обмена в доменах HNH, RuvC III и REC1-3, в основном расположены на поверхности белка. Связывание направляющей РНК вызывает существенные конформационные изменения в доменах RuvC III, REC3, HNH и CTD, линкерах Arg и L-II и менее выраженные изменения в доменах REC1-2 и RuvC II. Полученные результаты демонстрируют высокий уровень пластичности домена REC3, который считается «сенсором» формирования полного гибрида РНК:ДНК.

Список цитированной литературы

1. Bernal J.D., Crowfoot D. X-Ray Photographs of Crystalline Pepsin // Nature. 1934. Vol. 133. P. 794–795.
2. Kendrew J.C., Bodo G., Dintzis H.M., Parrish R.G., Wyckoff H. A three-dimensional model of the myoglobin molecule obtained by X-ray analysis // Nature. 1958. Vol. 181. P. 662–666.
3. Perutz M., Rossmann M., Cullis A.F., Muirhead H., Will G., North A.C.T. Structure of haemoglobin // Nature. 1960. Vol. 185. P. 416–422.
4. Ernst R.R., Anderson W.A. Application of fourier transform spectroscopy to magnetic resonance // Rev. Sci. Instrum. 1966. Vol. 37, № 1. P. 93–102.
5. Aue W.P., Bartholdi E., Ernst R.R. Two-dimensional spectroscopy. Application to nuclear magnetic resonance // J. Chem. Phys. 1976. Vol. 64, № 5. P. 2229–2246.
6. Jeener J., Meier B.H., Bachmann P., Ernst R.R. Investigation of exchange processes by two-dimensional NMR spectroscopy // J. Chem. Phys. 1979. Vol. 71, № 11. P. 4546–4553.
7. Wüthrich K. The way to NMR structures of proteins // Nat. Struct. Biol. 2001. Vol. 8, № 11. P. 923–925.
8. Ikura M., Kay L.E., Bax A. A Novel Approach for Sequential Assignment of ^1H , ^{13}C , and ^{15}N Spectra of Larger Proteins: Heteronuclear Triple-Resonance Three-Dimensional NMR Spectroscopy. Application to Calmodulin // Biochemistry. 1990. Vol. 29, № 19. P. 4659–4667.
9. Clore G.M., Gronenborn A.M. Structures of larger proteins in solution: Three- and four-dimensional heteronuclear NMR spectroscopy // Science. 1991. Vol. 252, № 5011. P. 1390–1399.
10. Clore G.M., Venditti V. Structure, dynamics and biophysics of the cytoplasmic protein-protein complexes of the bacterial phosphoenolpyruvate: Sugar phosphotransferase system // Trends Biochem. Sci. 2013. Vol. 38, № 10. P. 515–530.
11. Clore G.M., Driscoll P.C., Gronenborn A.M., Wingfield P.T. Analysis of the Backbone Dynamics of Interleukin- 1β Using Two-Dimensional Inverse Detected Heteronuclear ^{15}N - ^1H NMR Spectroscopy // Biochemistry. 1990. Vol. 29, № 32. P. 7387–7401.
12. Liu X.R., Zhang M.M., Gross M.L. Mass Spectrometry-Based Protein Footprinting for

High Order Structure Analysis: Fundamentals and Applications // Chem. Rev. 2021. Vol. 120, № 10. P. 4355–4454.

13. Thompson R.F., Walker M., Siebert C.A., Muench S.P., Ranson N.A. An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology // Methods. 2016. Vol. 100, № 2016. P. 3–15.

14. Mainz A., Religa T.L., Sprangers R., Linser R., Kay L.E., Reif B. NMR spectroscopy of soluble protein complexes at one mega-dalton and beyond // Angew. Chemie - Int. Ed. 2013. Vol. 52, № 33. P. 8746–8751.

15. Hvidt A., Linderstrøm-Lang K. Exchange of hydrogen atoms in insulin with deuterium atoms in aqueous solutions // BBA - Biochim. Biophys. Acta. 1954. Vol. 14, № C. P. 574–575.

16. Bai Y., Englander S.W. Future Directions in Folding The Multi- State Nature of Protein Structure // Proteins Struct. Funct. Genet. 1996. Vol. 24. P. 145–151.

17. Munson M.S.B., Field F.H. Chemical Ionization Mass Spectrometry // J. Am. Chem. Soc. 1966. Vol. 88, № 19. P. 4337–4345.

18. Hunt D.F., McEwen C.N., Upham R.A. Determination of Active Hydrogen in Organic Compounds by Chemical Ionization Mass Spectrometry // Anal. Chem. 1972. Vol. 44, № 7. P. 1292–1294.

19. Zehl M., Rand K.D., Jensen O.N., Jorgensen T.J.D. Electron transfer dissociation facilitates the measurement of deuterium incorporation into selectively labeled peptides with single residue resolution // J. Am. Chem. Soc. 2008. Vol. 130, № 51. P. 17453–17459.

20. Tanaka K., Waki H., Ido Y., Akita S., Yoshida Y., Yoshida T., Matsuo T. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry // Rapid Commun. Mass Spectrom. 1988. Vol. 2, № 8. P. 151–153.

21. Fenn J.B., Mann M., Meng C.K., Wong S.F., Whitehouse C.M. Electrospray Ionization for Mass Spectrometry of Large Biomolecules // Science. 1989. Vol. 246. P. 64–71.

22. Hemling M.E., Conboy J.J., Bean M.F., Mentzer M., Carr S.A. Gas phase hydrogen / deuterium exchange in electrospray ionization mass spectrometry as a practical tool for structure elucidation // J. Am. Soc. Mass Spectrom. 1994. Vol. 5, № 5. P. 434–442.

23. Siegel M.M. Hydrogen-Deuterium Exchange Studies Utilizing a Thermospray Mass Spectrometer Interface // Anal. Chem. 1988. Vol. 60, № 19. P. 2090–2095.

24. Davies N.W., Smith J.A., Molesworth P.P., Ross J.J. Hydrogen/deuterium exchange on aromatic rings during atmospheric pressure chemical ionization mass spectrometry // *Rapid Commun. Mass Spectrom.* 2010. Vol. 24, № 7. P. 1105–1110.
25. Ahmed A., Kim S. Atmospheric pressure photo ionization hydrogen/deuterium exchange mass spectrometry - A method to differentiate isomers by mass spectrometry // *J. Am. Soc. Mass Spectrom.* 2013. Vol. 24, № 12. P. 1900–1905.
26. Attygalle A.B., Gangam R., Pavlov J. Real-time monitoring of in situ gas-phase H/D exchange reactions of cations by atmospheric pressure helium plasma ionization mass spectrometry (HePI-MS) // *Anal. Chem.* 2014. Vol. 86, № 1. P. 928–935.
27. Hopper E.D., Roulhac P.L., Campa M.J., Patz Jr. E.F., Fitzgerald M.C. Throughput and Efficiency of a Mass Spectrometry-Based Screening Assay for Protein-Ligand Binding Detection // *J Am Soc Mass Spectrom.* 2008. Vol. 19, № 9. P. 1303–1311.
28. Rand K.D., Adams C.M., Zubarev R.A., Jørgensen T.J.D., Uppsala S.-. Electron Capture Dissociation Proceeds with a Low Degree of Intramolecular Migration of Peptide Amide Hydrogens // *J. Am. Chem. Soc.* 2008. Vol. 130, № 4. P. 1341–1349.
29. Pan J., Han J., Borchers C.H., Konermann L. Hydrogen / Deuterium Exchange Mass Spectrometry with Top-Down Electron Capture Dissociation for Characterizing Structural Transitions of a 17 kDa Protein // *J. Am. Chem. Soc.* 2009. Vol. 131. P. 12801–12808.
30. Fornelli L., Ayoub D., Aizikov K., Beck A., Tsybin Y.O. Middle-Down Analysis of Monoclonal Antibodies with Electron Transfer Dissociation Orbitrap Fourier Transform Mass Spectrometry // *Anal. Chem.* 2014. Vol. 86. P. 3005–3012.
31. Berger B.A., Loewenstein A., Meiboom S. Nuclear Magnetic Resonance Study of the Protolysis and Ionization of N-Methylacetamide // *J. Am. Chem. Soc.* 1959. Vol. 81, № 1. P. 62–67.
32. Perrin C.L. Proton Exchange in Amides: Surprises from Simple Systems // *Acc. Chem. Res.* 1989. Vol. 22, № 8. P. 268–275.
33. Amon S., Trelle M.B., Jensen O.N., Jørgensen T.J.D. Spatially resolved protein hydrogen exchange measured by subzero-cooled chip-based nanoelectrospray ionization tandem mass spectrometry // *Anal. Chem.* 2012. Vol. 84, № 10. P. 4467–4473.
34. Guttman M., Lee K.K. Isotope Labeling of Biomolecules: Structural Analysis of Viruses

by HDX-MS // *Methods Enzymol.* 2016. Vol. 566. P. 405–426.

35. Pirrone G.F., Iacob R.E., Engen J.R. Applications of hydrogen/deuterium exchange MS from 2012 to 2014 // *Anal. Chem.* 2015. Vol. 87, № 1. P. 99–118.

36. Malito E., Faleri A., Surdo P. Lo, Veggi D., Maruggi G., Grassi E., Cartocci E., Bertoldi I., Genovese A., Santini L., et al. Defining a protective epitope on factor H binding protein, a key meningococcal virulence factor and vaccine antigen // *Proc. Natl. Acad. Sci. U. S. A.* 2013. Vol. 110, № 9. P. 3304–3309.

37. Houde D., Berkowitz S.A. Conformational Comparability of Factor IX–Fc Fusion Protein, Factor IX, and Purified Fc Fragment in the Absence and Presence of Calcium // *J. Pharm. Sci.* 2012. Vol. 101, № 7. P. 1688–1700.

38. Bou-Assaf G.M., Chamoun J.E., Emmett M.R., Fajer P.G., Marshall A.G. Advantages of isotopic depletion of proteins for hydrogen/deuterium exchange experiments monitored by mass spectrometry // *Anal. Chem.* 2010. Vol. 82, № 8. P. 3293–3299.

39. Lemaire P., Debois D., Smargiasso N., Quinton L., Gabelica V., De Pauw E.A. Use of 1,5-diaminonaphthalene to combine matrix-assisted laser desorption/ionization in-source decay fragmentation with hydrogen/deuterium exchange // *Rapid Commun. Mass Spectrom.* 2013. Vol. 27, № 16. P. 1837–1846.

40. Walters B.T., Ricciuti A., Mayne L., Englander S.W. Minimizing back exchange in the hydrogen exchange - mass spectrometry experiment // *J. Am. Soc. Mass Spectrom.* 2013. Vol. 23. P. 2132–2139.

41. Keppel T.R., Weis D.D. Analysis of disordered proteins using a simple apparatus for millisecond quench-flow H/D exchange // *Anal. Chem.* 2013. Vol. 85, № 10. P. 5161–5168.

42. Pan J., Borchers C.H. Top-down structural analysis of posttranslationally modified proteins by Fourier transform ion cyclotron resonance-MS with hydrogen/deuterium exchange and electron capture dissociation // *Proteomics.* 2013. Vol. 13, № 6. P. 974–981.

43. Zhu M.M., Rempel D.L., Du Z., Gross M.L. Quantification of Protein - Ligand Interactions by Mass Spectrometry , Titration , and H / D Exchange : PLIMSTEX // *J. Am. Chem. Soc.* 2003. Vol. 125. P. 5252–5253.

44. Waterhouse A., Bertoni M., Bienert S., Studer G., Tauriello G., Gumienny R., Heer F.T., De Beer T.A.P., Rempfer C., Bordoli L., et al. SWISS-MODEL: Homology modelling of protein

structures and complexes // *Nucleic Acids Res.* Oxford University Press, 2018. Vol. 46, № W1. P. W296–W303.

45. Bryant Jeri L, Boughter John D, Gong Suzhen, LeDoux Mark S H.D.H. Unique Structural Features of Mammalian NEIL2 DNA Glycosylase Prime Its Activity for Diverse DNA Substrates and Environments // *Physiol. Behav.* 2021. Vol. 29, № 1. P. 29–42.

46. Prakash A., Eckenroth B.E., Averill A.M., Imamura K., Wallace S.S., Doublie S. Structural Investigation of a Viral Ortholog of Human NEIL2/3 DNA Glycosylases // *DNA Repair.* 2013.

47. Fiser A., Šali A. MODELLER: Generation and Refinement of Homology-Based Protein Structure Models // *Methods Enzymol.* 2003. Vol. 374. P. 461–491.

48. Zheng H., Cooper D.R., Porebski J., Shabalin I.G., Handing K.B., Minor W., Porebski P.J., Shabalin I.G., Handing K.B., Minor W. CheckMyMetal: a macromolecular metal-binding validation tool // *Acta Crystallogr. Sect. D Struct. Biol.* International Union of Crystallography, 2017. Vol. 73, № 3. P. 223–233.

49. Popov A. V., Vorobjev Y.N., Zharkov D.O. MDTRA: A molecular dynamics trajectory analyzer with a graphical user interface // *J. Comput. Chem.* 2013. Vol. 34, № 4. P. 319–325.

50. Friedberg E.C., Walker G.C., Siede W., Wood R.D., Schultz R.A., Ellenberger T. DNA repair and mutagenesis // *Trends in Genetics.* 2006. Vol. 12, № 2.

51. Wallace S.S., Murphy D.L., Sweasy J.B. Base excision repair and cancer // *Cancer Letters.* NIH Public Access, 2012. Vol. 327, № 1–2. P. 73–89.

52. Zharkov D.O. Base excision DNA repair // *Cell. Mol. Life Sci.* 2008. Vol. 65, № 10. P. 1544–1565.

53. Dou H., Mitra S., Hazra T.K. Repair of Oxidized Bases in DNA Bubble Structures by Human DNA Glycosylases NEIL1 and NEIL2 // *J. Biol. Chem.* 2003. Vol. 278, № 50. P. 49679–49684.

54. Makasheva K.A., Endutkin A. V., Zharkov D.O. Requirements for dna bubble structure for efficient cleavage by helix-two-turn-helix dna glycosylases // *Mutagenesis.* 2020. Vol. 35, № 1. P. 119–128.

55. Greenfield N.J. Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions // *Nat. Protoc.* 2007.

Vol. 1, № 6. P. 2527–2535.

56. Leuenberger P., Gansch S., Kahraman A., Cappelletti V., Boersema P.J., Von Mering C., Claassen M., Picotti P. Cell-wide analysis of protein thermal unfolding reveals determinants of thermostability // *Science* (80-.). 2017. Vol. 355, № 6327.
57. Das A., Wiederhold L., Leppard J.B., Kedar P., Prasad R., Wang H., Boldogh I., Karimi-Busheri F., Weinfeld M., Tomkinson A.E., et al. NEIL2-initiated, APE-independent repair of oxidized bases in DNA: Evidence for a repair complex in human cells // *DNA Repair* (Amst). 2006. Vol. 5, № 12. P. 1439–1448.
58. Das S., Chattopadhyay R., Bhakat K.K., Boldogh I., Kohno K., Prasad R., Wilson S.H., Hazra T.K. Stimulation of NEIL2-mediated oxidized base excision repair via YB-1 interaction during oxidative stress // *J. Biol. Chem.* 2007. Vol. 282, № 39. P. 28474–28484.
59. Mezulis S., Yates C.M., Wass M.N., E Sternberg M.J., Kelley L.A., Mezulis S., Yates C.M., Wass M.N., Sternberg M.J.E. The Phyre2 web portal for protein modeling, prediction and analysis // *Nat. Protoc.* 2015. Vol. 10, № 6. P. 845–858.
60. Jinek M., Jiang F., Taylor D.W., Sternberg S.H., Kaya E., Ma E., Anders C., Hauer M., Zhou K., Lin S., et al. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation // *Science* (80-.). American Association for the Advancement of Science, 2014. Vol. 343, № 6176.
61. Nishimasu H., Ran F.A., Hsu P.D., Konermann S., Shehata S.I., Dohmae N., Ishitani R., Zhang F., Nureki O. Crystal structure of Cas9 in complex with guide RNA and target DNA // *Cell*. Elsevier Inc., 2014. Vol. 156, № 5. P. 935–949.
62. Pettersen E.F., Goddard T.D., Huang C.C., Couch G.S., Greenblatt D.M., Meng E.C., Ferrin T.E. UCSF Chimera - A visualization system for exploratory research and analysis // *J. Comput. Chem. J Comput Chem*, 2004. Vol. 25, № 13. P. 1605–1612.
63. Case D.A., Belfon K., Ben-Shalom I.Y., Brozell S.R., Cerutti D.S., Cheatham T.E., Cruzeiro I.V.W.D., Darden T.A., Duke R.E., Giambasu G., et al. AMBER 2020. University of California, San Francisco, 2020.
64. Eckenroth B.E., Cao V.B., Averill A.M., Dragon J.A., Doublé S. Unique Structural Features of Mammalian NEIL2 DNA Glycosylase Prime Its Activity for Diverse DNA Substrates and Environments // *Structure*. 2021. Vol. 29, № 1. P. 29-42.e4.

Приложение

Последовательности белков и нуклеиновых кислот.

hNEIL2

MPEGPLVRKFHHLVSPFVGQVVKTTGGSSKKLQPASLQSLWLQDTQVHGKKLFLRFDLDE
EMGPPGSSPTPEPPQKEVQKEGAADPKQVGEPSGQKTLDGSSRSAELVPQGEDDSEYLER
DAPAGDAGRWLRVSFGLFGSVVWNDFSRACKANKRGDWRDPSRLVLHFGGGGFLAFYNC
QLSWSSSPVVTPTCDILSEKFHRGQALEALGQAQPVCYTLLDQRYFSGLGNIIKNEALYR
AGIHPLSLGSLVLSASRREVLVDHVVEFSTAWLQGKFQGRPQHTQVYQKEQCPAGHQVMKE
AFGPEDGLQRLTWWCPQCQPQLSEEPEQCQFS

SpCas9

MDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE
ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHERHPFIG
NIVDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSD
VDKLFQILVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGN
LIALSLGLTPNFKSNFDLAEDAQLSKDQYDDDLNLLAQIGDQYADFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA
GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELH
AILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEE
VVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFL
SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWG
RLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSL
HEHIANLAGSPAIKKILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRER
MKRIEEDIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH
IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL

TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRK
MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
ATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA
YSLVLVAKVEKGKSKKLKSVKELLGITIMERSSEKPNPIDFLEAKGYKEVKKDLIILPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLLKGS PEDNEQKQLFVE
QHKHYLDEIEQISEFSKR VILADANLDK VLSAYNKHRDKPIREQAENIIHLFTLTNLGA
PAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD

sgPHK

5' GGUUGGACAUGCUCGACAUUCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA
GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUU 3'

DHK

5' GATCTCCCGAATGTCGAGCATGTCCA ACTCAC 3'