

СОДЕРЖАНИЕ

Том 66, номер 5, 2021

Колонка главного редактора	675
Михаил Валентинович Ковальчук (к 75-летию со дня рождения)	677

ОБЗОРЫ

Комплексный подход к изучению структуры и функций сердечных потенциалзависимых ионных каналов <i>Ю. Г. Качер, М. Г. Карлова, Г. С. Глухов, Х. Чжан, Е. В. Заклязьминская, Ж. Лоуссарн, О. С. Соколова</i>	681
Структурные исследования РНК-белковых комплексов в пострибосомную эру <i>С. В. Тищенко, А. О. Михайлина, Н. В. Леконцева, Е. А. Столбоушкина, Е. Ю. Никонова, О. С. Никонов, А. Д. Никулин</i>	697
Структурные исследования мембранного гликопротеина-предшественника бета-амилоида и белков, участвующих в процессе его протеолиза <i>А. С. Урбан, Я. В. Бершацкий, К. В. Павлов, Э. В. Бочаров</i>	708

ДИФРАКЦИЯ И РАССЕЯНИЕ ИОНИЗИРУЮЩИХ ИЗЛУЧЕНИЙ

Влияние хлоридов одно- и двухвалентных металлов на олигомерный состав кристаллизационных растворов лизоцима и дальнейший рост кристаллов <i>М. А. Марченкова, П. В. Конарев, А. С. Бойкова, К. Б. Ильина, Ю. В. Писаревский, М. В. Ковальчук</i>	723
Поиск эффективного алгоритма для восстановления решений при анализе данных малоуглового рассеяния от раствора кремнезоля <i>А. Е. Крюкова, П. В. Конарев, В. В. Волков</i>	730
Электронная дифракция микрокристаллов на примере лизоцима <i>Р. А. Камышинский, В. А. Кралин, М. Ю. Чесноков, В. Р. Самыгина, А. С. Орехов</i>	738

СТРУКТУРА МАКРОМОЛЕКУЛЯРНЫХ СОЕДИНЕНИЙ

Пространственная структура рекомбинантной термофильной рибокиназы <i>Thermus species 2.9</i> в комплексе с аденозиндифосфатом <i>В. И. Тимофеев, Ю. А. Абрамчик, Т. И. Муравьева, Н. Е. Жухлистова, Р. С. Есипов, И. П. Куранова</i>	742
Кристаллизация в условиях микрогравитации и структура атомного разрешения уридинфосфорилазы из <i>Vibrio cholerae</i> <i>П. А. Эйстрих-Геллер, С. В. Рубинский, В. Р. Самыгина, А. Г. Габдулхаков, М. В. Ковальчук, А. С. Миронов, А. А. Лашков</i>	750
Роль конформационных изменений гексамерной молекулы бактериальных уридинфосфорилаз в связывании субстратов <i>К. М. Поляков, Н. Н. Мордкович, Т. Н. Сафонова, А. Н. Антипов, Н. А. Окорокова, П. В. Дороватовский, В. П. Вейко</i>	759
Структура мутантной формы белка-регулятора трансляции Hfq с удлиненной петлей L4 <i>А. А. Алипов, Н. В. Леконцева, А. О. Михайлина, М. С. Фандо, С. В. Тищенко, А. Д. Никулин</i>	765

Влияние миссенс-мутации Pe222Thr в SsoIF2 γ на сродство γ - и β -субъединиц aIF2	
<i>О. С. Никонов, О. В. Кравченко, Н. А. Невская, Е. А. Столбоушкина, М. Б. Гарбер, С. В. Никонов</i>	771
Пространственная структура термостабильной трансаминазы D-аминокислот из археи <i>Methanocaldococcus jannaschii</i> DSM 2661	
<i>К. М. Бойко, А. Ю. Николаева, А. К. Бакунова, Т. Н. Стеханова, Т. В. Ракитина, В. О. Попов, Е. Ю. Безсуднова</i>	776
Кристаллизация и предварительное рентгеновское исследование C-концевого фрагмента миомезина-2	
<i>Е. П. Ламбер, Н. Пинотсис</i>	782
Пространственная структура точечного мутанта эстеразы PMGL2	
<i>Д. А. Горяйнова, А. Ю. Николаева, М. В. Крюкова, Л. Е. Петровская, Д. А. Корженевский, Е. А. Крюкова, Д. А. Долгих, К. М. Бойко</i>	785
Компьютерное моделирование структур обратимо переключаемых флуоресцентных белков с LOV-доменами	
<i>Ю. И. Метелешко, М. Г. Хренова, А. В. Немухин</i>	789
О тактике <i>ab initio</i> -поиска формы белковых частиц по данным малоуглового рентгеновского рассеяния	
<i>В. В. Волков</i>	793
Олигомеризация церулоплазмينا под действием ионов меди	
<i>М. В. Петухов, А. В. Соколов, В. А. Костевич, В. Р. Самыгина</i>	802
Формирование в растворе структур высокого порядка CBS-пирофосфатазой из <i>D. hafniense</i>	
<i>Л. А. Дадинова, Е. Ю. Сошинская, Ю. М. Чесноков, Р. А. Камышинский, А. Л. Васильев, Э. В. Штыкова</i>	807
Супрамолекулярные комплексы тетрапептидов, способные индуцировать конформационные переходы бета-домена альфа-лактальбумина человека	
<i>Я. А. Забродская, А. В. Швецов, Ю. П. Гармай, Д. В. Лебедев, Р. Даттани, В. В. Егоров</i>	815
Изучение конформационной подвижности GroEL методами криоэлектронной микроскопии и молекулярной динамики	
<i>И. С. Панина, А. А. Мамчур, И. А. Ярошевич, Д. В. Зленко, Е. Б. Пичкур, С. С. Кудрявцева, В. И. Муронец, О. С. Соколова, Т. Б. Станишинева-Коновалова</i>	821
Определение околоатомной структуры олигомерных белков <i>E. coli</i> без предварительной очистки	
<i>Е. Б. Пичкур, В. И. Микиртумов, О. В. Тихонова, Н. И. Деркачева, Л. П. Курочкина, О. С. Соколова</i>	829

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ

RuFeprstr: программное дополнение к системе молекулярной графики RuMol для расчетов свободной энергии связывания лиганда с рецептором

А. А. Лашков, И. В. Толмачев, П. А. Эйстрих-Геллер, С. В. Рубинский 836

КОЛОНКА ГЛАВНОГО РЕДАКТОРА

DOI: 10.31857/S002347612105009X

Вниманию читателей представлен тематический выпуск, посвященный работам российских исследователей в области структурной биологии.

Структурная биология – междисциплинарная наука на стыке физики, математики, молекулярной биологии, биохимии, биофизики, изучающая трехмерные структуры отдельных биомолекул (в первую очередь белков) и сложных молекулярных машин. Знание структуры и ее динамических изменений позволяет детально изучить аспекты функционирования биомолекул и использовать эти знания для разработки лекарств, создания и совершенствования биотехнологических цепочек. Современная структурная биология располагает целым арсеналом методов, которые используются в зависимости от объектов. Рентгеноструктурный анализ белков (или белковая кристаллография), зародившийся в 50-х годах прошлого века, до сих пор является основным методом. Пионером белковой кристаллографии в России был Б.К. Вайнштейн, создавший в 1958 г. лабораторию белковой кристаллографии в Институте кристаллографии РАН (сейчас ФНИЦ “Кристаллография и фотоника” РАН). Работы этой лаборатории, выполненные на мировом уровне, внесли весомый вклад в развитие направления в целом. Доля биологических структур, полученных рентгеноструктурным анализом, на настоящий момент составляет более 80%. Ключевым компонентом успеха метода является получение высококачественных кристаллов исследуемых объектов. В сборнике большая часть обзоров и статей посвящена рентгеноструктурному анализу белков, в том числе исследованию процессов кристаллизации и применению специальных техник кристаллизации. Статья Эйстрих-Геллер и др. демонстрирует, как использование микрогравитации помогает получить кристаллы, рассеивающие до атомного разрешения (1.06 Å) для белка уридинфосфорилазы. Работа выполнена в рамках программы ЦНИИМашРоскосмоса по кристаллизации белков на Российском сегменте МКС в рамках Федеральной космической программы России. ФНИЦ “Кристаллография и фотоника” РАН (ранее Институт кристаллографии РАН) является постановщиком этих экспериментов с 2005 г. Структура такого высокого разрешения позволила увидеть детали строения, которые не были известны для этого белка ранее. Такой ре-

зультат оказался возможен, в том числе, благодаря использованию синхротронного излучения. В вошедших в сборник статьях по белковой кристаллографии результаты получены с использованием синхротронов. Такие работы традиционно ведутся в ФНИЦ “Кристаллография и фотоника” РАН, Курчатовском институте, а также институтах Российской академии наук ФИЦ “Фундаментальные основы биотехнологии”, Институте белка и Институте молекулярной биологии. Принятая в Российской Федерации “Федеральная научно-техническая программа развития синхротронных и нейтронных исследований и исследовательской инфраструктуры на 2019–2027 годы” подчеркивает важность развития этих работ.

В статьях Эйстрих-Геллер и др., Метелешко и др. продемонстрировано применение математического моделирования для анализа кристаллических структур. Отдельная статья посвящена разработке утилиты к программе молекулярной графики *rumol* для анализа взаимодействия лиганда с белком в кристаллических структурах.

Ядерно-магнитный резонанс (ЯМР) позволяет определить структуру белков с молекулярной массой до ~100 кДа в растворе с высокой точностью. Использование ЯМР в ряде работ по получению структур белков, играющих роль в возникновении болезни Альцгеймера, продемонстрировано в обзоре Урбана и др. В обзоре показаны преимущества ЯМР для исследования строения неструктурированных белков и выигрешность использования нескольких структурных методов для исследования сложных биологических объектов и процессов. Другой бурно развивающийся метод – криоэлектронная микроскопия (криоЭМ) – позволяет определять структуры макромолекул и макромолекулярных комплексов большого размера также без получения кристаллов. Достижения этого метода представлены в обзоре Качер и др. и проиллюстрированы в статьях Паниной и др., Пичкур и др. (работы выполнены совместно МГУ и Курчатовским институтом). Ограничения по размеру исследуемых объектов постоянно смещаются в сторону менее крупных образцов. Метод криоЭМ занял свое место в ряде методов структурной биологии высокого разрешения с 2013 г. с появлением новых сверхбыстрых детекторов и имеет блестящие перспективы.

Еще одним методом, активно используемым в исследовании структур биологических объектов, является малоугловое рентгеновское рассеяние (МУРР). Этот метод, позволяющий определять структуры низкого разрешения в растворе при физиологических условиях, отличается простотой эксперимента и имеет широкие границы применимости. Статьи сотрудников ФНИЦ “Кристаллография и фотоника” РАН и Курчатовского института демонстрируют его возможности в сочетании с использованием синхротронного излучения. Оно позволяет исследовать образцы более низких концентраций и получать данные более высокого качества. В работе М.А. Марченковой и др., направленной на изучение процесса образования белковых кристаллов, метод МУРР был применен для изучения олигомерного состава кристаллизационных растворов лизоцима в зависимости от типа осадителя. Пред-

ставлены также статьи по развитию методов анализа малоугловых экспериментальных данных.

Выпуск дает общее представление о широте спектра работ в области структурной биологии, которые ведутся в настоящий момент в России, о возможностях и перспективах этой междисциплинарной науки и могут быть интересны как структурным биологам, так и ученым, работающим в смежных областях.

*Главный редактор журнала “Кристаллография”
член-корреспондент РАН
Профессор М.В. Ковальчук*

*Приглашенный редактор академик РАН,
доктор химических наук В.О. Попов*

*Приглашенный редактор
кандидат физико-математических наук
В.Р. Самыгина*

МИХАИЛ ВАЛЕНТИНОВИЧ КОВАЛЬЧУК (К 75-ЛЕТИЮ СО ДНЯ РОЖДЕНИЯ)

DOI: 10.31857/S0023476121050246



21 сентября 2021 г. исполняется 75 лет Михаилу Валентиновичу Ковальчуку – одному из ведущих ученых-физиков России, доктору физико-математических наук, профессору, члену-корреспонденту РАН. Он широко известен как ученый-экспериментатор в области рентгеновской физики и кристаллографии, основоположник использования рентгеновского и синхротронного излучений для исследования структуры поверхности, нанобиоматериалов и систем, много сделавший для сохранения и развития науки в России.

В 1970 г. Михаил Валентинович окончил физический факультет Ленинградского университета, где специализировался в области рентгеновской физики. Эту тематику он продолжал развивать в Институте кристаллографии им. А.В. Шубникова АН СССР, где его трудовой путь начался в 1970 г. с должности стажера-исследова-

теля. С самого начала направление его исследований было связано с развитием рентгенодифракционных методов изучения структуры различных материалов, тонких кристаллических слоев на поверхности кристаллов.

В 1978 г. М.В. Ковальчук защитил кандидатскую диссертацию на тему “Метод трехкристального рентгеновского спектрометра и исследование структурного совершенства тонких кристаллических слоев”. В последующие годы он занялся разработкой физических и методических основ принципиально нового метода изучения конденсированных сред – структурно-чувствительной спектроскопии с использованием стоячих рентгеновских волн. Этот метод впервые позволил совместить уникальные возможности рентгеновской спектроскопии и дифракции, в связи с чем был создан ряд экспериментальных лабораторий.

Одна из них – лаборатория рентгеновской оптики и синхротронного излучения – создана по инициативе М.В. Ковальчука, возглавившего ее.

В 1988 г. М.В. Ковальчук защитил докторскую диссертацию на тему “Метод стоячих рентгеновских волн в исследовании структуры приповерхностных слоев полупроводников”, внося определяющий вклад в создание и практическое использование метода стоячих волн. Им был выполнен цикл пионерских работ по изучению выхода фотоэлектронов в условиях двух- и многоволновой дифракции, развивающих новое направление, связанное с созданием фазочувствительных рентгеновских методов – основы рентгено-голографических исследований. Уникальная структурная чувствительность метода стоячих рентгеновских волн стала фундаментом для превращения рентгеновских методов в инструмент для изучения структуры поверхности, тонких кристаллических слоев и двумерных систем. Эти исследования М.В. Ковальчука были позже отмечены Российской академией наук премией имени Е.С. Федорова.

Пятнадцать лет (с 1998 по 2013 г.) Михаил Валентинович возглавлял Институт кристаллографии. В этот период Институт был переориентирован на решение принципиально новых научных задач, связанных с созданием и развитием наноматериалов, нанодиагностики, нано-

биотехнологий и космического материаловедения. М.В. Ковальчуком были сформулированы новые научные приоритеты, создана современная научно-исследовательская инфраструктура, обеспечивающая уникальные исследовательские возможности, в том числе для кристаллизации белков, формирования систем молекулярных пленок. В 2016 г. по инициативе М.В. Ковальчука на базе Института кристаллографии (путем объединения с четырьмя академическими институтами) была создана новая междисциплинарная структура, получившая название “Федеральный научно-исследовательский центр “Кристаллография и фотоника”. Объединенный ученый совет нового Центра возглавил М.В. Ковальчук.

С начала 1990-х годов параллельно с исследованиями в области полупроводниковых кристаллов и пленок, рентгеновской оптики М.В. Ковальчук значительное внимание уделял синхротронному излучению. Используя свой опыт работы на ряде европейских источников синхротронного излучения, он приступил к разработке новых экспериментальных станций сначала для Зеленоградского, а затем и Курчатовского синхротронов. В 1999 г. по инициативе академика Е.П. Велихова Михаил Валентинович был назначен директором-организатором Курчатовского центра синхротронного излучения.

За несколько лет М.В. Ковальчук и его коллеги реализовали масштабный научный проект по разработке, созданию и вводу в эксплуатацию комплекса уникального научно-исследовательского оборудования — экспериментальных станций на пучках первого в России специализированного источника синхротронного излучения КИСИ-Курчатов.

Реализация синхротронного проекта — серьезный вклад М.В. Ковальчука в развитие экспериментальных и технологических возможностей России для проведения исследований в области фундаментальных наук, материаловедения, нано- и биотехнологий, молекулярной биологии и медицины, а также в развитие методов диагностики с атомарным разрешением. Эта работа отмечена Премией Правительства РФ.

В 2005 г. М.В. Ковальчук назначен директором Российского научного центра “Курчатовский институт”.

В этот период М.В. Ковальчуком была выдвинута и под его руководством реализована созданным им коллективом ученых и специалистов концепция формирования национальной сети уникальных экспериментальных установок класса мегасайенс как базы для восстановления и развития всего отечественного научно-технологического комплекса.

В рамках реализации этой концепции М.В. Ковальчуком и возглавляемым им коллективом был разработан и осуществлен ряд проектов мегасайенс национального и глобального масштабов. В их числе — глубокая модернизация и оснащение уникальным экспериментальным оборудованием единственного в России Курчатовского специализированного источника синхротронного излучения, создание и запуск крупнейшего в мире источника нейтронов — реакторного комплекса “ПИК” в г. Гатчине Ленинградской области.

Реализация этих проектов позволила не только восстановить и укрепить высокотехнологичный производственно-технологический потенциал страны, но и сформировать уникальную национальную междисциплинарную инфраструктуру исследований и разработок, консолидировавшую научно-технологическое сообщество и ставшую базой для развития в России важнейших научно-технологических направлений.

В том числе, речь идет о нанотехнологическом проекте, ведущим идеологом и организатором которого стал М.В. Ковальчук, в рамках реализации которого созданы принципиально новые технологии и материалы, обеспечившие научно-технологический прорыв в целом ряде ключевых отраслей экономики, таких как энергетика, транспорт, медицина, инфокоммуникационные технологии и системы.

По инициативе и при непосредственном участии М.В. Ковальчука в Курчатовском институте сформирована научная программа, ориентированная прежде всего на проведение междисциплинарных научных исследований на таких крупных исследовательских установках, как ускорительные и лазерные комплексы, термоядерные реакторы, нейтронный и синхротронный источники.

Для осуществления этой программы Указом Президента РФ в 2010 г. был запущен пилотный проект по созданию в России первого национального исследовательского центра.

Распоряжением Правительства от 31 декабря 2010 г. М.В. Ковальчук назначен первым директором Национального исследовательского центра “Курчатовский институт”, под эгидой которого сегодня объединена значительная часть ядерно-физического, материаловедческого и генетического потенциала страны.

Междисциплинарные исследования, начатые М.В. Ковальчуком в Институте кристаллографии и продолженные в Курчатовском институте, вышли на новый уровень с развитием им принципиально нового научного направления — конвергенции нано-, био-, инфо-, когнитивных и социогуманитарных (НБИКС) наук и технологий. М.В. Ковальчук впервые выдвинул концепцию формирования природоподобной техносферы

как закономерного этапа развития человеческой цивилизации, сформулировал идеологию и основные принципы использования конвергенции НБИКС наук и технологий как инструмента создания природоподобной техносферы. Под руководством и при непосредственном участии М.В. Ковальчука в НИЦ “Курчатовский институт” создан не имеющий мировых аналогов Курчатовский комплекс НБИКС-природоподобных технологий, где осуществляются масштабные исследования и разработки в области создания технологий, основанных на принципах функционирования живой природы.

Таким образом, была создана база для научно-технологического прорыва в исследовании процессов жизнедеятельности живой природы, сформировано ядро метрологического комплекса принципиально новой природоподобной технологической базы России. Сегодня эти направления реализуются в рамках Федеральной научно-технической программы развития синхротронных и нейтронных исследований и исследовательской инфраструктуры, а также Федеральной научно-технической программы развития генетических технологий на 2019–2027 гг.

Благодаря активной позиции М.В. Ковальчука в последние годы существенно укрепилась позиция России в таких глобальных международных научных проектах, как Большой адронный коллайдер (LHC, CERN), Международный экспериментальный термоядерный реактор (ITER), Европейский ускорительный центр по исследованию тяжелых ионов и антипротонов (FAIR). М.В. Ковальчук – инициатор и идеолог участия Российской Федерации в качестве одного из основных партнеров в проекте создания Европейского рентгеновского лазера на свободных электронах XFEL в г. Гамбурге.

Значимый вклад вносят М.В. Ковальчук и руководимый им коллектив в развитие исследований и разработок в области термоядерного синтеза и плазменных технологий.

Под его руководством в НИЦ “Курчатовский институт” осуществлены создание и пуск крупнейшей в стране инновационной термоядерной установки мирового уровня Токамак Т-15МД. Она станет базовым элементом инфраструктуры термоядерных исследований в России и обеспечит выполнение работ в поддержку проекта Международного термоядерного экспериментального реактора ITER.

М.В. Ковальчук является руководителем приоритетного технологического направления по технологиям прямого преобразования энергии. Под его руководством разработаны новые энерготехнологии и системы энергообеспечения для морской техники, основанные на принципах прямого преобразования энергии.

Под руководством М.В. Ковальчука и при непосредственном его участии совместно с ГК “Росатом” разработаны концепция и прототипы базовых элементов инфраструктуры распределенной малой ядерной энергетики на основе технологий прямого преобразования энергии, в частности для энергообеспечения Арктического и других труднодоступных регионов.

В последнее десятилетие в НИЦ “Курчатовский институт” благодаря инициативам М.В. Ковальчука возрожден целый ряд важных направлений прикладных исследований и разработок в целях национальной безопасности. Указом Президента Российской Федерации на М.В. Ковальчука возложено общее руководство научно-исследовательскими проектами Военного инновационного технополиса “Эра” Минобороны России.

М.В. Ковальчуком предложена концепция, сформулированы идеология и основные принципы ускоренного безопасного развития ядерной медицины на базе ядерно-физических научных центров. В настоящее время эта концепция реализуется в рамках Федеральной научно-технической программы развития синхротронных и нейтронных исследований и исследовательской инфраструктуры на 2019–2027 гг.

М.В. Ковальчук стал инициатором развертывания в НИЦ “Курчатовский институт” работ, основанных на использовании естественно-научных методов для изучения объектов культурного наследия. Эти исследования проводятся совместно с Институтом археологии РАН, Государственным историческим музеем, Государственным музеем изобразительных искусств им. А.С. Пушкина и являются новым важным этапом в развитии исторического материаловедения.

М.В. Ковальчук внес определяющий вклад в формирование кадровой базы междисциплинарных исследований и разработок. Под его руководством и при непосредственном участии создана инновационная система образования в области конвергентных наук и технологий, в том числе им организованы первый в мире факультет нано-, био-, информационных и когнитивных технологий в Московском физико-техническом институте, базовый факультет в Национальном исследовательском ядерном университете МИФИ.

Также НИЦ “Курчатовский институт” реализует масштабный проект междисциплинарного довузовского образования (Курчатовский проект), в котором участвуют несколько десятков школ Москвы, Санкт-Петербурга, Пскова, республики Крым.

Большое значение имеют реализованные по инициативе и под руководством М.В. Ковальчука проекты возрождения культурно-просветительских площадок Дома ученых им. А.П. Алексан-

дрова в Москве и Мемориальной дачи И.В. Курчатова в Крыму.

М.В. Ковальчук с 2001 г. является членом Президиума Совета при Президенте Российской Федерации по науке и образованию. Также он является председателем Национального комитета кристаллографов России, с 2006 г. – главным редактором журнала “Кристаллография” РАН, с 2019 г. – главным редактором журнала “Российские нанотехнологии”. В 2020 г. М.В. Ковальчук стал главным редактором журнала “Вестник Военного инновационного технополиса “Эра”, созданного по его предложению.

М.В. Ковальчук уделяет большое внимание популяризации научных знаний, с 2007 г. он яв-

ляется автором и ведущим научно-популярных телевизионных циклов “Истории из будущего” и “Картина мира с Михаилом Ковальчуком”. В рамках этих циклов вышло более 300 передач, рассказывающих о наиболее актуальных темах и направлениях науки и технологий.

М.В. Ковальчук – полный кавалер орденов “За заслуги перед Отечеством”, лауреат премий Правительства РФ в области науки и техники и в области образования.

Редколлегия и Редакция журнала “Кристаллография” от души поздравляют Михаила Валентиновича с юбилеем! Коллеги и друзья желают ему творческой энергии, вдохновения, новых достижений и открытий на благо российской науки.

УДК 577.25

КОМПЛЕКСНЫЙ ПОДХОД К ИЗУЧЕНИЮ СТРУКТУРЫ И ФУНКЦИЙ СЕРДЕЧНЫХ ПОТЕНЦИАЛЗАВИСИМЫХ ИОННЫХ КАНАЛОВ

© 2021 г. Ю. Г. Качер¹, М. Г. Карлова¹, Г. С. Глухов¹, Х. Чжан², Е. В. Заклязьминская³,
Ж. Лоуссарн⁴, О. С. Соколова^{1,2,*}

¹ Московский государственный университет им. М.В. Ломоносова, Москва, Россия

² Университет МГУ-ППИ в Шэньчжэне, Шэньчжэнь, Китай

³ Российский научный центр хирургии им. ак. Б.В. Петровского, Москва, Россия

⁴ Научно-исследовательское подразделение института грудной клетки,
Национальный институт здравоохранения и медицинских исследований, Нант, Франция

*E-mail: sokolova@mail.bio.msu.ru

Поступила в редакцию 03.09.2020 г.

После доработки 15.09.2020 г.

Принята к публикации 15.09.2020 г.

Мембранные белки, в частности ионные каналы, стали объектом пристального внимания структурной протеомики в середине XX века. Методы изучения ионных каналов разнообразны и включают в себя структурные (рентгеноструктурный анализ (РСА), криоэлектронная просвечивающая микроскопия (крио-ПЭМ), анализ с помощью рентгеновских лазеров на свободных электронах) и функциональные (пэтч-кламп). В обзоре рассмотрена эволюция подходов к изучению структуры сердечных ионных каналов, дано представление о новых приемах структурной биологии в приложении к каналам, в том числе об использовании липо- и нанодисков, обсуждается вклад электрофизиологических исследований, молекулярной динамики для получения полной картины о структуре и функционировании сердечных ионных каналов. За годы исследований многие методы были разработаны и адаптированы для изучения ионных каналов. Электрофизиологические исследования стали мощным инструментом для расшифровки механизмов ионной проводимости и селективности, стробирования и регулирования, а также тестирования ингибиторных молекул. Реконструкция атомной структуры ионных каналов стала возможной благодаря активному развитию сначала РСА и крио-ПЭМ, а в настоящий момент — за счет использования рентгеновских лазеров на свободных электронах.

DOI: 10.31857/S0023476121050076

ОГЛАВЛЕНИЕ

Введение

1. Сердечные ионные каналы и каналопатии: зачем нужно знать структуру ионных каналов?
2. Путь к структуре полноразмерного ионного канала
3. Структурно-функциональные исследования ионных каналов
4. Электронная микроскопия для исследования ионных каналов
5. Революция в структурной биологии — крио-ПЭМ структуры ионных каналов с высоким разрешением
6. Новые стратегии для выделения ионных каналов
 - 6.1. Почему так сложно изучать мембранные белки?
 - 6.2. Нанодиски

6.3. Липодиски

7. Изучение динамики ионных каналов с помощью рентгеновских лазеров на свободных электронах

Заключение

ВВЕДЕНИЕ

XX век ознаменовал расцвет структурной биологии — получено колоссальное количество данных, накоплен реальный потенциал для реконструкции всех процессов в живых организмах. Завершились два десятилетия XXI века, расшифровано более 160 000 третичных и четвертичных структур белков; в их числе крупные молекулярные машины, рибосома [1], мембранные белки [2, 3], белки цитоскелета [4], ферменты и др. Оптимисты прогнозируют расшифровку архитектур всех известных белков и их функций через 30 лет. Это означает не только понимание структуры ин-

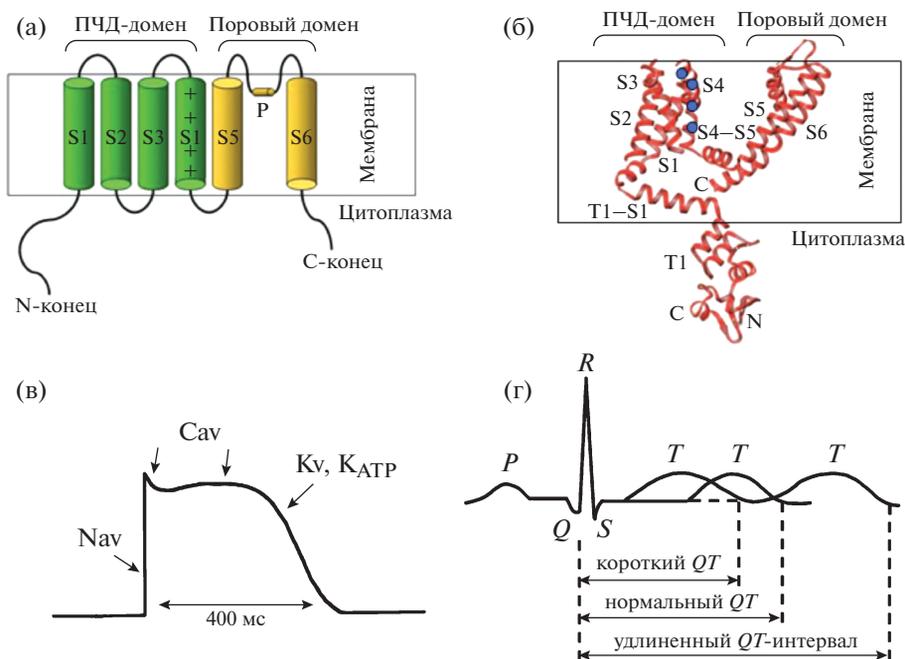


Рис. 1. Строение и функционирование ПЗ-каналов. Схема одной α -субъединицы Kv-канала (а). Обозначены трансмембранные сегменты S1–S6 и петля P, формирующая пору канала. Заряженные аргинины на сегменте S4 обозначены плюсами. ПЧД – потенциал-чувствительный домен. Кристаллическая структура одной α -субъединицы канала Kv1.2 (PDB ID 2A79) (б). Схема сердечного ПД (в). Стрелками отмечены ПЗ-каналы, отвечающие за различные фазы ПД. ЭКГ в норме и при синдромах LQT и SQT (r).

дивидуальных макромолекул и их взаимодействий с лигандами, но и межмолекулярных взаимодействий в сложных белковых комплексах. Изменение конформационного состояния белка отражается на его функциональной активности [5]. Исчерпывающее знание структур всех доменов белковой молекулы дает возможность правильно интерпретировать ее конформационные изменения при активации и ингибировании и, более того, управлять этими процессами [6].

Вслед за геномикой, протеомикой, транскриптомикой и метаболомикой появилась отрасль науки «структурная протеомика». Понимание биологических механизмов на уровне атомного разрешения позволяет применять эту информацию в медицине и биотехнологии. Особенно это актуально для изучения мембранных белков, в частности ионных каналов и транспортеров, принимающих участие в регуляции жизнедеятельности каждой клетки организма, в особенности кардиомиоцитов.

Ионные каналы – большой функциональный класс интегральных трансмембранных белков. Большинство ионных каналов имеет ротационную симметрию, образующую несколько идентичными субъединицами или гомологичными доменами, располагающимися вокруг поры канала. Наиболее распространенная классификация ионных каналов – по механизму управле-

ния – включает следующие группы: каналы, управляемые напряжением, лигандом/лигандами, липидами, температурой или светом, а также механически управляемые.

Управляемые напряжением или потенциал-зависимые (ПЗ) ионные каналы – самая обширная группа. Их роль в возбудимых клетках заключается в деполяризации мембран в ответ на изменение потенциала. Их функционирование определяется слаженной работой трех структурных единиц: сенсора напряжения, поры и ворот канала (рис. 1а). Большинство ПЗ-каналов, проводящих ионы K^+ , Ca^{2+} и Na^+ , состоит из четырех трансмембранных доменов с шестью сегментами (рис. 1б), каждый из которых пересекает липидный бислой.

Потенциалы действия (ПД) в сердечных клетках существенно отличаются от обнаруженных в других типах клеток. Это связано с наличием особых «сердечных» ионных каналов (активнее всего экспрессирующихся в кардиомиоцитах) и кинетикой их работы. За деполяризацию мембраны миоцитов миокарда млекопитающих отвечают быстрые натриевые каналы, за фазу плато – кальциевые ионные каналы, важнейшую роль в определении потенциала покоя в фазе реполяризации ПД играют калиевые каналы (рис. 1в).

1. СЕРДЕЧНЫЕ ИОННЫЕ КАНАЛЫ И КАНАЛОПАТИИ: ЗАЧЕМ НУЖНО ЗНАТЬ СТРУКТУРУ ИОННЫХ КАНАЛОВ?

Ритмичные сокращения сердца на протяжении всей жизни человека обеспечиваются синхронной генерацией и распространением электрических импульсов, к которым приводит слаженная работа нескольких десятков ионных каналов. Катионные ПЗ-каналы, которые кодируются генами *KCNQ1*, *KCNH2*, *KCNE1*, *KCNE2*, *KCNJ2*, *TRPM4*, играют важную роль в реполяризации сердечных клеток и поддержании нормальной продолжительности сердечного ПД.

Семейство Kv7.x или *KCNQ* включает в себя пять типов каналов Kv7.1–Kv7.5. Канал Kv7.1, кодируемый геном *KCNQ1*, преимущественно экспрессируется в тканях сердца, а также в эпителиальных тканях и гладкой мускулатуре. Каналы Kv7.2–Kv7.5 (*KCNQ2–5*) в основном встречаются в тканях нервной системы [7]. В сердечной мышечной ткани Kv7.1 экспрессируется совместно с вспомогательными бета-субъединицами *KCNE1*, *KCNE2* и *KCNE3* для формирования функционально активного канала, обеспечивающего медленный ток реполяризации [8] и, следовательно, отвечающего за длительность сердечного ПД. К настоящему моменту описано 109 мутаций (в основном точечных замен), затрагивающих активность канала Kv7.1, изменяющих кинетику активации, взаимодействие с регуляторными субъединицами, либо нарушающих транспорт в мембрану клетки, приводящих к развитию синдрома удлиненного интервала QT (*longQT*, *LQTS*) на электрокардиограмме (*ЭКГ*) (рис. 1г) и синдрому внезапной смерти. Кроме *LQTS* некоторые из этих мутаций ассоциированы с синдромом Джервелла–Ланге–Нильсена [9], состоянием, при котором нарушения сердечного ритма сопровождаются нарушениями слуха. Для данного канала описаны также активирующие мутации, вызывающие увеличение тока и ведущие к укорочению ПД сердца (*синдром укороченного интервала QT*, *shortQT*, *SQTS*) и фибрилляции предсердий [10].

Правильная работа ионных каналов может быть нарушена в результате мутаций, влияющих на количество экспрессируемых белковых субъединиц, первичную и вторичную структуру, а также посттрансляционные модификации белка. В результате генетически детерминированных изменений суммарная проницаемость через канал для ионов может увеличиваться (нарушается механизм инактивации или закрывания канала, увеличивается экспрессия) либо уменьшаться (нарушается активация, пропуская способность, снижается экспрессия белка). Примечательно, что различные по своему функциональному эффекту мутации могут приводить к клинически противоположным *ЭКГ*-феноменам, и

соответствующие аритмогенные синдромы формируют аллельные серии заболеваний. Одним из хорошо изученных примеров аллельной серии заболеваний служат синдромы, развивающиеся в результате различных мутаций в гене *KCNH2*. Мутации, реализующиеся по типу “*loss of function*” (потеря функции) (например, замены N410D, A561T, G610S, A614V, N996I и др.), приводят к снижению суммарного исходящего калиевого тока из клетки, что задерживает время реполяризации и выражается удлинением интервала QT [11]. Эта генетическая форма представляет собой второй по частоте подтип наследственного синдрома *LQT*, заболевания с высоким риском развития желудочковой тахикардии типа “*пируэт*” и внезапной сердечной смерти [12]. На мутации в этом гене приходится ~35% всех случаев генетически подтвержденных случаев заболевания [13].

Мутации в гене *KCNH2*, реализующиеся по типу “*gain of function*” (усиление функции), ведут к тому, что выведение из клетки положительно заряженных ионов калия увеличивается, что приводит к ускорению реполяризации и проявляется укорочением интервала QT на *ЭКГ* (рис. 1г). Электрокардиографические и клинические проявления таких мутаций более многообразны, включают синдром *SQT* (замена T618I), синдром Бругада (например, T152I, R164C, W927G, R1135H и т.д.), семейную фибрилляцию предсердий (замена N588K) [14–16].

Функционирование канала Kv11.1 (кодируемого геном *KCNH2*) может быть нарушено не только мутациями, но и взаимодействием с широким спектром лекарственных препаратов. Блокирование канала Kv11.1 с развитием вторичного удлинения сердечного ПД (и, как следствие, увеличение продолжительности интервала QT) представляет собой один из очень опасных побочных эффектов широкого круга лекарств. Первым препаратом, выведенным в 1998 г. с фармацевтического рынка в связи с блокадой канала Kv11.1 и риском развития желудочковых аритмий, был блокатор H1-гистаминовых рецепторов терфенадин, а сейчас таких препаратов насчитывается не менее десятка [17]. В настоящее время все новые фармацевтические субстанции тестируются на их способность блокировать сердечные ионные каналы, главным образом Kv11.1. Понимание особенностей строения и функционирования ионных каналов, в частности Kv11.1, очень важно для рационального дизайна новых лекарственных средств любой направленности. Было показано, что некоторые аминокислоты в S6-доме спирали (Thr623, Phe656) и в основании поровой спирали (Thr623, Ser624, Val625) особенно чувствительны к лекарственным взаимодействиям [18]. Независимо от того, в какую сторону смещается продолжительность реполяризации кардиомиоцита, основной угрозой для носителей мута-

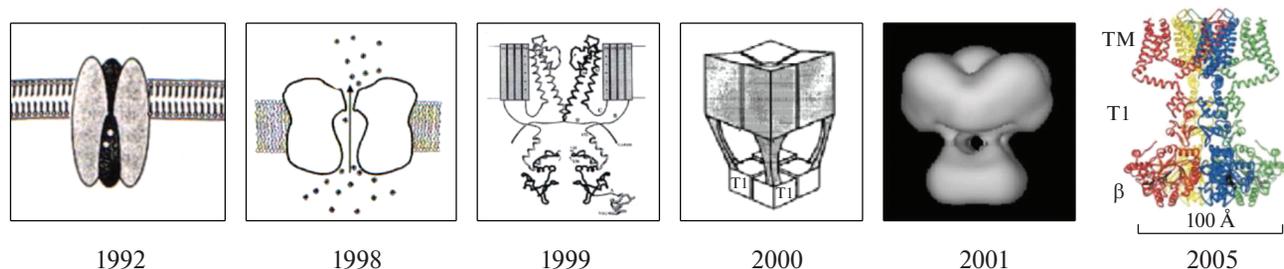


Рис. 2. Путь к структуре Kv-канала дрозофилы *Shaker*. Слева направо: 1992 г. — показано, что канал является тетрамерным [20]; 1998 г. — появление данных о водной поре канала; к 1999 г. опубликованы структуры порового домена (гомологичный канал KcsA [21]) и тетрамеризационного домена [22]; 2000 г. — гипотеза “висячей гондолы” [23]; 2001 г. — первая ПЭМ-структура канала [24] подтвердила упомянутую выше гипотезу; 2005 г. — кристаллическая структура химерного полноразмерного Kv1.1/1.2-канала с β -субъединицей [25].

ций в гене *KCNH2* является повышенный риск развития желудочковых нарушений ритма и внезапной смерти.

Эффективность лечения разных наследственных сердечных каналопатий различна. В настоящее время доступны как хирургические (имплантация антиаритмических устройств и левосторонняя стеллэктомия), так и медикаментозные подходы. При синдроме LQT высокую (но не абсолютную) эффективность показали β -блокаторы [19]. Однако успехи лекарственной терапии других каналопатий (семейные нарушения проводимости, синдром Бругада, SQT) не столь велики. Поэтому детальное изучение структурных и функциональных изменений нормальных и мутантных ионных каналов необходимо для разработки новых молекул, способных восстановить или компенсировать генетический дефект.

Повышение доступности высокоэффективного секвенирования привело к экспоненциальному росту выполняемых генетических тестов. Естественным результатом роста выполняемых диагностик стало выявление большого числа новых, не охарактеризованных редких вариантов, клиническое значение которых трудно интерпретировать. Детальное понимание структуры ионных каналов, понимание функциональной значимости каждого аминокислотного остатка помогут разработать предиктивные модели, способные достоверно прогнозировать значение новых уникальных вариантов.

2. ПУТЬ К СТРУКТУРЕ ПОЛНОРАЗМЕРНОГО ИОННОГО КАНАЛА

К концу XX века были успешно клонированы многие гены, кодирующие ДНК ионных каналов эукариотического происхождения, а также их бактериальных аналогов. Однако четвертичная структура большинства ионных каналов долгое время оставалась неизвестной, так как они с большим трудом поддаются трехмерной кристал-

лизации. Выводы о структуре и функционировании каналов делались на основе непрямых данных мутационного анализа, электрофизиологии и молекулярной динамики (МД) (рис. 2).

В отсутствие прямых структурных данных были разработаны комплексные подходы, позволяющие решить сложнейшую задачу определения процессов, лежащих в основе активации Kv-каналов. Эта задача частично была решена к началу 90-х годов XX века, когда были предложены первые модели активации ПЗ Kv-каналов (обзор [26]). По данным цистеинового сканирования и электрофизиологии установили, что Kv-канал является тетрамером (рис. 2); при активации канала сенсор потенциала S4 под действием изменения трансмембранного потенциала движется поперек мембраны [27]; спираль S4 тянет линкер S4–S5, что приводит к нарушению взаимодействия S4–S5/S6 и открытию поры канала [28] (рис. 3а). Однако некоторые модели активации, как потом выяснилось, были ошибочны [29]. Для их валидации необходимы были реальные структурные данные.

Первая кристаллическая структура бактериального (*Streptomyces lividans*) ионного канала KcsA с разрешением 3.2 Å была получена в лаборатории будущего Нобелевского лауреата Р. МакКиннона и опубликована в 1998 г. [21]. Эта работа за прошедшие 22 года была процитирована более 7000 раз. Так как структура поры гомологична у большинства катионных каналов [30], структура KcsA открыла неопределимые возможности для структурных биологов и биофизиков (рис. 4).

Атомная структура KcsA позволила использовать МД для моделирования конформаций не только этого канала [31], но и гомологичных каналов эукариот [32, 33] (рис. 3в). Для получения хорошо дифрагирующих кристаллов авторы удалили внутриклеточные С-концы субъединиц канала KcsA. Поэтому необходимо было доказать функциональность полученной конструкции — начался ренессанс методов электрофизиологии.

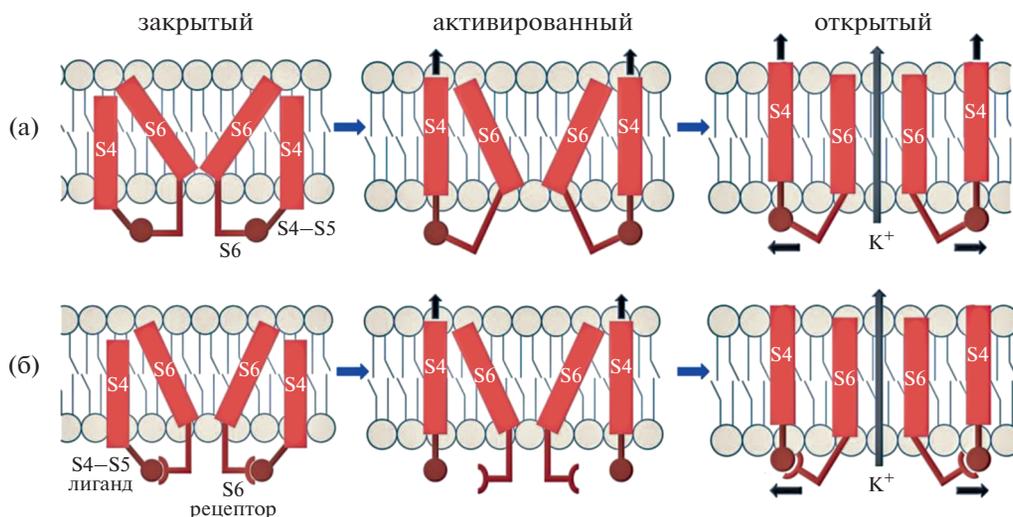


Рис. 3. Две модели соединения между сенсором напряжения и порвыми субъединицами ионных ПЗ-каналов: а – механистическая модель, характерная для каналов семейства *Shaker*; б – лиганд-рецепторная модель, характерная для каналов EAG, ERG, KCNQ. Рисунок адаптирован из [28].

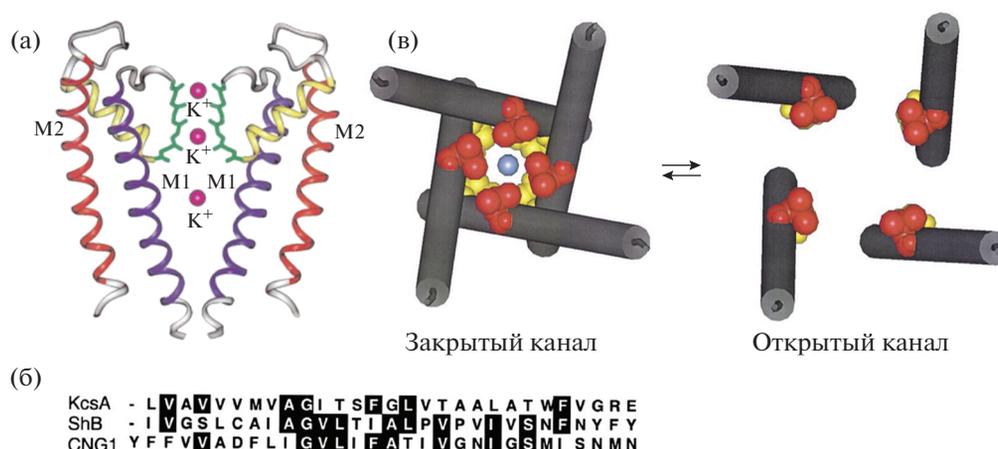


Рис. 4. Кристаллическая структура бактериального рН-зависимого калиевого канала KcsA позволила провести молекулярное моделирование гомологичного канала. Три иона калия расположены на внешней и внутренней позициях селективного фильтра (S1 и S3) и в центре водной полости (а). Выравнивание последовательностей канала KcsA, *Shaker* и гомологичного CNG1-канала (б). Модель конформационных изменений в CNG1-канале при открытии и закрытии ворот канала, построенная на основе МД (в) [33].

3. СТРУКТУРНО-ФУНКЦИОНАЛЬНЫЕ ИССЛЕДОВАНИЯ ИОННЫХ КАНАЛОВ

Фундаментальное понимание того, как структура катионных каналов меняется при мутациях, возможно только при сочетании структурных исследований с функциональными (электрофизиологическими) подходами. Характеристика активности ионных каналов с помощью электрофизиологических методов является основным источником знаний как в физиологическом, так и патофизиологическом контексте, что, в свою очередь, помогает в разработке соответствующих фармакологических препаратов.

Электрофизиология в XX веке внесла огромный вклад в расшифровку молекулярных механизмов ионной проводимости, ионной селективности, стробирования каналов, а также регулирования ионных каналов дополнительными белками [34–39].

В контексте патофизиологии ионных каналов электрофизиологические инструменты, в частности метод пэтч-кламп [40], использовались в тысячах исследований для подтверждения того, что варианты генов, кодирующие часть комплекса ионных каналов, являются причиной различных патологий. Хорошим примером в данном случае

является синдром LQT. Довольно простая схема моногенного наследования заболевания, вероятно, обусловила то, что оно долгое время находилось в центре внимания научной общественности после идентификации нескольких мутаций, в основном в генах *KCNQ1*, *KCNH2* и *SCN5A* [41, 42]. Во многих биофизических исследованиях идентифицированная замена была введена в плазмиду, кодирующую канал, и проведены эксперименты *in vitro* с помощью метода пэтч-кламп для проверки того, приводит ли мутация к модификации ионного тока.

Значительные достижения в области высокопроизводительного секвенирования и геномного анализа позволили исследовать более сложные формы наследственных заболеваний, такие как синдром Бругада [43]. Этот синдром характеризуется фибрилляцией желудочков и повышенным риском внезапной сердечной смерти, как и при синдроме LQT, но патофизиология несколько отличается, и ЭКГ-мониторинг позволяет легко отличить синдром LQT от синдрома Бругада (подъем сегмента ST на кардиограмме). При синдроме LQT большинство случаев связано со специфической мутацией, в основном в генах *KCNQ1*, *KCNH2* и *SCN5A*. При синдроме Бругада мутация в генах была идентифицирована только у 20% пациентов, несмотря на очевидное семейное происхождение патологии. Ситуация осложнена тем, что во многих случаях при идентифицированных в семье мутациях (в основном ген *SCN5A*, кодирующий ПЗ-натриевый канал Nav1.5) некоторые члены семьи являются лишь ее носителями (неполная пенетрантность), а другие члены семьи имеют патологию (фенокопию) [44]. В этом случае методы электрофизиологии представляют еще больший интерес, чем в случае синдрома LQT, так как они позволяют выяснить, приводит ли обнаруженная мутация к дисфункции канала и может ли она частично участвовать в патогенезе.

Что касается патофизиологии ионных каналов, инструменты электрофизиологии также полезны для определения того, как нарушение молекулярного механизма может привести к конкретной патологии. Например, исследования пэтч-кламп показали, что для активности ПЗ-каналов Kv7.1 и Kv11.1 требуется фосфолипид фосфатидилинозитол-4,5-бисфосфат (PI(4,5)P2) и что некоторые мутации при синдроме LQT изменяют взаимодействие этого полярного липида с каналом, что приводит к потере каналом своей функции [45–47].

Наконец, изучение внутримолекулярных взаимодействий в ПЗ-каналах с помощью электрофизиологических методов привело к разработке пептидов, имитирующих эти взаимодействия и либо ингибирующих, либо активирующих целевую группу ионных каналов [48–51]. Такие исследова-

ния могут в долгосрочной перспективе привести к разработке новых терапевтических средств.

Во всех аспектах, обсуждаемых выше, могут быть также полезны высокопроизводительные аналоги пэтч-кламп. Интерес к функциональной валидации каналов в сочетании с многократным увеличением количества тестируемых вариантов гена (более 1000 в канале Kv11.1 (hERG)) позволил разработку методов фенотипирования генов [52, 53].

Важно упомянуть об ограничениях техники пэтч-кламп, особенно потому, что это может приводить к неправильной интерпретации функции ионных каналов. Типичным случаем может служить субъединица KCNE1, первоначально названная MinK; она ошибочно считалась ионным каналом на основании того, что при экспрессии в ооцитах *Xenopus* наблюдали медленно активирующийся ПЗ-калиевый ток. Несмотря на очень осторожную интерпретацию результатов в первой статье, дальнейшие публикации подтвердили эту гипотезу [54, 55]. Однако ток можно было наблюдать только тогда, когда РНК вводили в ооциты *Xenopus*, но не в других моделях, что все же вызвало некоторые сомнения [56]. Только клонирование гена канала *KCNQ1* (первоначально названного KVLQT1) и демонстрация того, что этот канал изначально присутствует в ооцитах *Xenopus* и генерирует медленно активирующийся ток, управляемый напряжением только в присутствии KCNE1, определенно продемонстрировали, что KCNE1 не кодирует порообразующую субъединицу [38, 39], а является дополнительной субъединицей канала *KCNQ1*.

4. ЭЛЕКТРОННАЯ МИКРОСКОПИЯ ДЛЯ ИССЛЕДОВАНИЯ ИОННЫХ КАНАЛОВ

В 90-х годах XX века для исследования структуры ионных каналов начинают применять электронную микроскопию. Каналы выделяют из мембран клеток с использованием детергентов (DDM, CHAPS, Digitonin и других), наносят на углеродные сетки для последующей окраски солями тяжелых металлов (рис. 2д) [24, 57–60] либо используют появившийся в 80-х годах XX века метод криоэлектронной просвечивающей микроскопии (крио-ПЭМ) [61, 62].

Преимуществом крио-ПЭМ является то, что белковые молекулы находятся в водном окружении, поэтому их форма и конформация сохраняются. Частицы во льду не имеют преимущественной ориентации по сравнению с частицами на поверхности углерода. В то же время в крио-ПЭМ соотношение сигнал/шум, как правило, низкое, поскольку низок контраст исследуемых объектов. Для того чтобы повысить контраст, необходимо объединить данные, полученные от очень боль-

шого числа проекций. Реконструкция макромолекул по данным крио-ПЭМ продолжается уже более четырех десятилетий, однако на начальных этапах она позволяла получать трехмерные структуры с разрешением 15–18 Å только очень больших белков ионных каналов, таких как рианодинчувствительный рецептор (RyR) [63, 64].

Радиационное повреждение образцов является существенным препятствием для получения изображений с высоким разрешением [65]. Для снижения радиационного повреждения во время выбора зоны съемки, выравнивания и фокусировки пучка электронов применяются специальные системы “низкой дозы”, чтобы блокировать луч до последнего шага – получения изображения. Малые выдержки приводят к потере разрешения, шумным изображениям и, как следствие, недостатку данных с высоким разрешением. Это, в свою очередь, усложняет последующий процесс обработки изображений для определения трехмерной структуры объекта. Определение углов ориентации глобулярных молекул особенно чувствительно к шумам на необработанных изображениях, а также к ошибкам в процедурах выравнивания и классификации. Недостаточные размеры Kv-каналов (10–15 нм в диаметре) приводили к потере контраста изображения в микроскопе и падению разрешения реконструкции.

Для очень коротких выдержек становится критически важной способностью детектора воспринимать каждый попавший на него электрон. Такие устройства были сконструированы в XXI веке, они позволяют напрямую детектировать электроны (прямой детектор электронов, ПДЭ) и получить за секундную выдержку от 16 до 400 кадров. Благодаря специальной процедуре обработки изображения одной и той же белковой частицы со всех полученных кадров можно выровнять относительно друг друга и сложить для увеличения соотношения сигнал/шум. Таким образом, компенсируется движение частиц во льду, происходящее за время даже секундной выдержки [66].

5. РЕВОЛЮЦИЯ В СТРУКТУРНОЙ БИОЛОГИИ – КРИО-ПЭМ-СТРУКТУРЫ ИОННЫХ КАНАЛОВ С ВЫСОКИМ РАЗРЕШЕНИЕМ

За 15 лет после опубликования первой кристаллической структуры ионного канала исследователям с помощью рентгеноструктурного анализа (РСА) удалось расшифровать атомную структуру некоторого количества прокариотических [21, 67, 68], архейных [69] и эукариотических (рис. 2е) каналов [70–72].

Революция в структурной биологии мембранных белков началась с опубликования в 2013 г. первой атомной структуры ионного канала TRP

[73], *de novo* полученной с помощью крио-ПЭМ, что, в частности, связано с началом серийного выпуска камер ПДЭ. Разработка нового программного обеспечения для 3D-реконструкции [74] ускорила во много раз сбор и анализ сотен тысяч и миллионов изображений отдельных частиц каналов. В 2015 г. метод крио-ПЭМ был объявлен методом года по версии журнала Nature, а еще через 2 года, в 2017 г., трое ученых-основателей метода – Р. Хендерсон, Ж. Дюбуше и Й. Франк – удостоились Нобелевской премии. И это было только начало. В течение последующих 4 лет технический прорыв позволил получить многие до того неизвестные структуры ионных каналов, включая сердечные каналы KCNQ1 и KCNE2 [75, 76].

Первые крио-ПЭМ-исследования с высоким разрешением каналов семейства EAG показали, что модель активации, основанная на данных о структуре каналов семейства *Shaker* [26, 77], не подходит для описания работы семейств каналов, кодируемых генами *KCNH* [3, 49, 75, 78].

Структура мембранного домена каналов семейства *Shaker* имеет характерную особенность, открытую после расшифровки кристаллической структуры химерного канала Kv1.1/Kv1.2 [71], – ПЧД взаимодействует не со своим поровым доменом, а с доменом, принадлежащим соседней субъединице (рис. 5а, 5в).

Такая архитектура получила название “*domain swapped*” – доменный обмен. Подобная организация трансмембранных спиралей была возможна из-за длинного S4–S5 линкера у каналов семейства *Shaker*. Формирование доменного обмена оказалось невозможно для каналов EAG, так как у них оказался очень короткий S4–S5 линкер, и спираль S5 взаимодействует со своим же ПЧД (рис. 5б, 5г) [75].

Сравнение структуры каналов: химерного Kv1.2/2.1 [71] и Kv11.1 в открытой конформации [3] с каналом Kv10.1, пора которого была закрыта за счет взаимодействия с кальмодулином (CaM) [75], привело исследователей к новой гипотезе функционирования каналов семейства EAG (рис. 6). Отличие ее от предыдущей модели состоит в том, что линкеру отводится меньшая роль, а основное изменение конформации S6 происходит за счет латерального смещения S4- и S5-сегментов. Роль шарнира, который обеспечивает изменение конформации S6, выполняет остаток глицина (G648 у Kv11.1 и G460 у Kv10.1) [3].

В результате было выдвинуто предположение, что активация каналов семейства EAG происходит по лиганд-рецепторному механизму (рис. 3б). В данном случае линкер S4–S5 действует как внешний лиганд, который связывает нижнюю часть спирали S6 и блокирует канал в закрытом состоянии, когда спирали S4 ПЧД находятся в со-

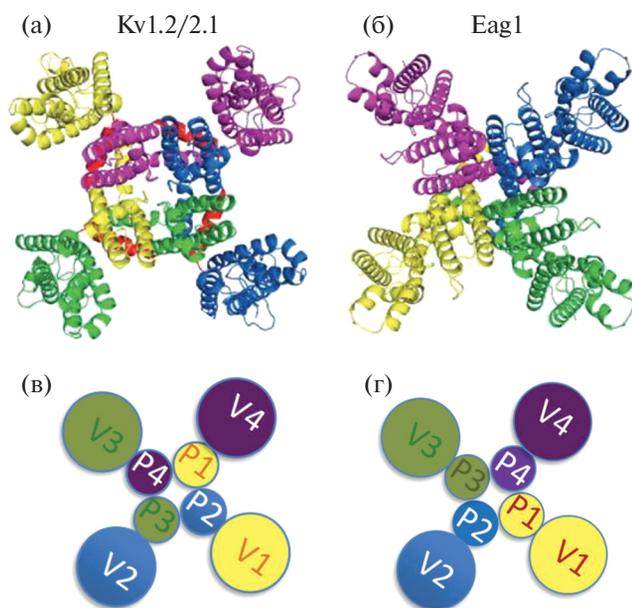


Рис. 5. Различия в расположении доменов у различных Kv-каналов по данным РСА и крио-ПЭМ: а – канал Kv1.1/1.2 [25]; б – канал Kv10.1 [75]; в, г – схематическое изображение взаимодействия поровых доменов и ПЧД в соответствующих каналах.

стоянии “внизу” при отрицательном потенциале. Когда спирали S4 находятся в состоянии “вверху”, линкер S4–S5 выходит из кармана связывания с S6, и канал открывается. Этот молекулярный механизм хорошо согласуется с результатами электрофизиологических экспериментов [79], сделанных, в частности, на каналах hERG [49].

6. НОВЫЕ СТРАТЕГИИ ДЛЯ ВЫДЕЛЕНИЯ ИОННЫХ КАНАЛОВ

6.1. Почему так сложно изучать мембранные белки?

Структурные исследования требуют получения целевого белка в большом количестве и с высокой степенью чистоты. К настоящему времени разработано множество различных подходов к наработке и очистке белков. Тем не менее для мембранных белков именно подготовка образца является узким местом при определении структуры, в особенности экстракция целевого белка из клеточных мембран. Дeterгент, который неподходящим образом имитирует гидрофобное окружение исследуемого белка, может также значительно изменить его структуру.

В таких условиях субнанометрового разрешения реконструкция 30 лет назад позволяла достичь лишь электронная кристаллография: структуры бактериородопсина в пурпурной мембране [80] и ацетилхолинового рецептора в электрическом органе ската *Torpedo marmorata* [81] бы-

ли получены этим методом. Позднее научились моделировать природные двумерные кристаллы, вместо нативных мембран использовали очищенные молекулы каналов и путем диализа удаляли детергент из смеси, заменяя его на липиды, однако этот подход оказался достаточно сложен для воспроизведения и до сих пор применяется лишь в нескольких лабораториях в мире.

Для получения реконструкций белков каналов в мембранном окружении разрабатывались и другие специальные подходы, такие, например, как сферическая реконструкция [82]. После вставки в маленькую сферическую везикулу из липидов мембранный белок приобретал строго определенную ориентацию относительно мембраны, и его положение на проецируемом изображении везикулы напрямую определяло два из трех определяемых углов Эйлера. Анализ изображений везикул во льду показал, что их плотность хорошо описывается простой моделью рассеяния электронов на мембране. Компьютерное моделирование показало, что этот метод теоретически может улучшить реконструкцию мембранных белков. В [83] его применили для изучения кластеризации грамицидина в липосомах, однако высокого разрешения этот метод достичь не позволил.

Значение липидов ближайшего окружения показано для многих белков, например, взаимодействие PI(4,5)P2 с Kv7 [45] и взаимодействие KcsA с анионными липидами [84]. Удаление этих липидов при сольubilизации детергентами может приводить не только к функциональным нарушениям, но и к структурным перестройкам. В связи с этим постоянно ведется как поиск новых детергентов, так и разработка принципиально новых подходов к сольubilизации мембранных белков. Относительно недавно были разработаны два похожих, однако основанных на различных соединениях подходы – использование нанодисков и липодисков в структурных исследованиях мембранных белков.

6.2. Нанодиски

Нанодиски – это альтернативный и более эффективный подход к выделению мембранных белков. Это имитаторы мембранного окружения (мембраномиметики), которые можно использовать для поддержания мембранных белков в растворимой форме для дальнейших структурных исследований. Как следует из названия, нанодиски представляют собой дискообразные наноразмерные бислои фосфолипидов, окруженные молекулами амфипатического α -спирального белка MSP, который выступает в роли пояса, ограничивающего размер диска (рис. 7а, 7в).

Изначально метод был основан на использовании аполипопротеина ApoA1 человека. Липопротеин

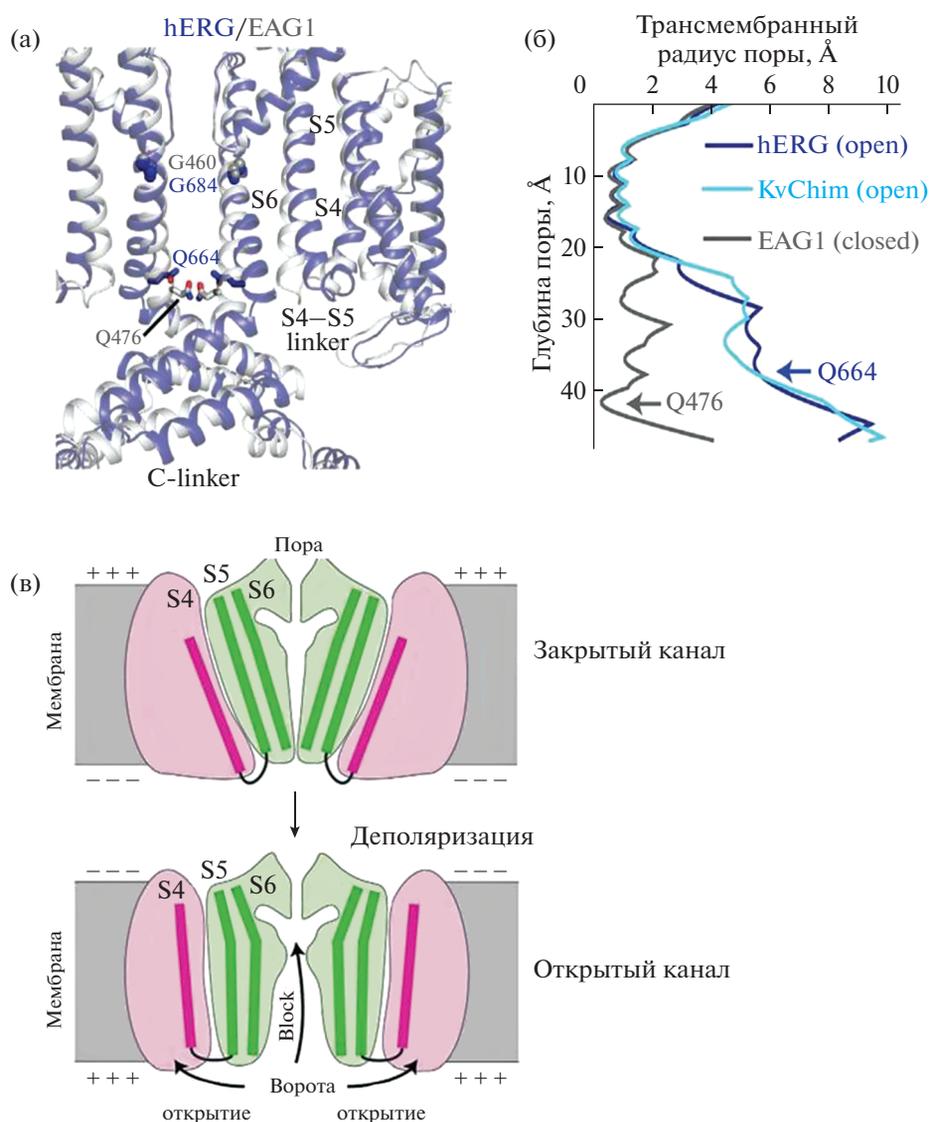


Рис. 6. Сравнение структур закрытого и открытого Кv-каналов. Различия в структурах трансмембранной части и прилегающих к поро цитоплазматических доменов открытого канала Kv11.1 (темно-серый) и закрытого канала Kv10.1 (светло-серый) (а). Сравнения профилей поры каналов, структуры которых получены в открытом (hERG и Kv1.2/2.1) и закрытом (EAG1) состоянии (б). Схема активации каналов семейства EAG (в).

теины высокой плотности состоят из липидов и холестерина, основной аполипопротеин ApoA1 может принимать различные структурные формы. Он синтезируется в печени, затем из формы без липидов постепенно обретает вид шарика из липидов, холестерина и сложных эфиров холестерина, проходя через временную дискоидную стадию. Именно диски и стали прототипом водорастворимых наночастиц [87].

Включение мембранного белка в нанодиск происходит самопроизвольно, так как он взаимодействует с гидрофобными хвостами липидов. Нанодиски могут использоваться в большом диапазоне концентраций при комнатной и физиологических температурах и храниться в течение не-

скольких месяцев при 4°C с минимальной агрегацией. При соблюдении правильных условий можно добиться монодисперсных образцов с контролируемым размером по сравнению с липосомами, бицеллами, мицеллами. На данный момент существует множество модифицированных вариантов белка MSP [88], в том числе коммерчески доступных.

Количественный и качественный состав фосфолипидов нанодисков можно варьировать для имитации биологических особенностей тех или иных мембран. Самые популярные синтетические липиды, используемые для сборки нанодисков – фосфатидилхолин (POPC), димиристоилфосфатидилхолин (DMPC), дипальмитоилфос-

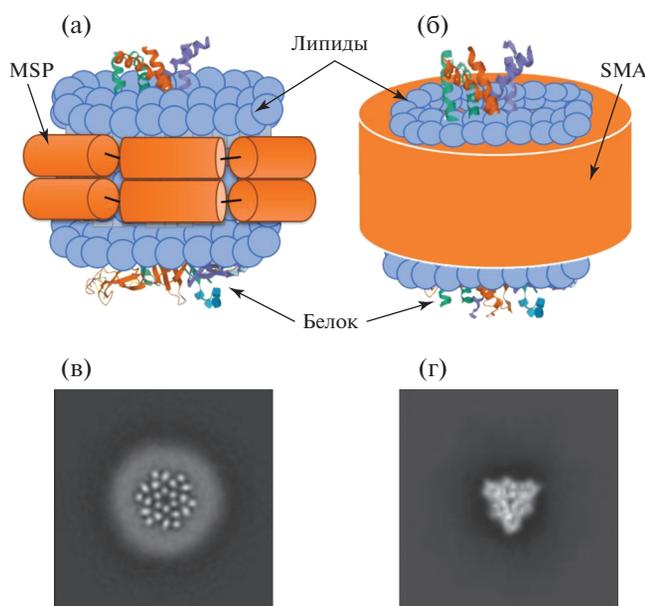


Рис. 7. Строение комплексов нанодиска и липодиска с мембранным белком. Схематическое представление нанодиска (а) и липодиска (б) с инкорпорированным мембранным белком. Внизу – суммарные крио-ПЭМ-проекции каналов в нано- и липодиске: в – канал dORAI в MSP [85]. Стрелка указывает на дополнительную плотность, соответствующую нанодиску; г – канал ASIC в SMA [86]. Отметим, что липодиск не образует дополнительной плотности вокруг мембранной части канала в отличие от нанодиска.

фатидилхолин (**ДРРРС**), а также смеси с заряженными фосфолипидами, такими как PI(4,5)P2, фосфатидилсерин (**РС**), фосфатидилглицерин (**РР**), фосфатидилэтаноламин (**РЕ**).

Публикации многих структур ионных каналов, выделенных в нанодисках за последние 5 лет, однозначно говорят об успешности метода (табл. 1). Возникли новые варианты нанодисков, которые адаптированы, в том числе, для больших белков и комплексов [89, 90], поскольку размер нанодиска может иметь решающее значение для поддержания нативного олигомерного состояния белка, инкорпорированного в нанодиски. Более того, для многих типов нанодисков уже опубликованы оптимизированные протоколы [91] и подобраны оптимальные молярные отношения смеси липидов.

6.3. Липодиски

Одним из новейших методов выделения белков в естественном липидном окружении является использование сополимера стирола и малеиновой кислоты (**СМА**). Особенностью данного полимера является способность переходить между линейной и спиральной конформациями в зависимости от pH и ионной силы среды. Чередующиеся гидрофобные (стирольные) и гидрофиль-

ные (остатки малеиновой кислоты) фрагменты **СМА** делают его амфипатическим и, следовательно, способным встраиваться в биологические мембраны. Встраиваясь в мембрану, полимер разрушает ее и формирует фрагменты мембраны округлой формы размером 10–40 нм, окруженные поясом полимера [106] (рис. 7б). Такие частицы назвали **SMALP** – липидные частицы, образованные сополимером малеиновой кислоты (**SMA lipid particles**). Термин Lipodisk® (липодиск) был введен компанией Malvern Cosmeceutics Ltd. для обозначения патентованных смесей полимер/липид, разрабатывавшихся как носители для доставки гидрофобных фармацевтических субстанций. В настоящее время термин “липодиск” или “нативный нанодиск” в основном применяется в отношении дисковидных мембранных структур, образованных полимерами, в противовес термину “нанодиск”, который используется для обозначения структур, сформированных производными аполипопротеинов.

Влияние состава полимера (соотношение остатков стирола и малеиновой кислоты), pH среды и ионной силы раствора на эффективность растворения модельной мембраны подробно исследовано в [107]. Полимер остается растворимым в диапазоне pH 7–9, при понижении pH начинает агрегировать. Однако полимер с соотношением остатков стирола и малеиновой кислоты 1.4 : 1 выдерживает понижение pH среды до 4. Высокая ионная сила раствора снижает растворимость полимера, однако повышает эффективность сольubilизации мембран. Интересно, что диаметр формируемых липодисков зависит в большей степени от липидного состава мембраны, чем от характеристик полимера [108]. Длина молекул полимера при заданном соотношении остатков стирола и малеиновой кислоты влияет не столько на размеры частиц, сколько на их стабильность [109]. На модельной системе (липосомы, образованные смесью POPC/POPG в соотношении 9 : 1) было показано, что на размер формируемых липодисков влияет соотношение массы липида и полимера: после достижения критической концентрации полимера, необходимой для полного разрушения везикул, дальнейшее увеличение концентрации полимера приводит к уменьшению размеров липодисков [110].

Влияние липидного состава бислоя на кинетику формирования **SMALP** исследовано в [111, 112]. Показано, что длина ацильной части липидов практически не влияет на размер, однако определяет кинетику формирования **SMALP**. При растворении мембран решающее значение имеют особенности упаковки – фаза, толщина мембраны, латеральное давление, плотность заряда, температура, ионная сила раствора. Более тонкие мембраны с низким латеральным давлением, низкой плотностью заряда на поверхности

Таблица 1. Ионные каналы, структуры которых получены в нанодисках

Ионный канал	Год	Частицы MSP	Липиды	Метод	Разрешение, Å	Литература
KCC4	2020	MSP1D1	DOPE, POPC, POPS	крио-ПЭМ	3.65	[92]
VSD4-NavAb	2020	MSP1E3	DMPC	ПЭМ		[93]
KCNQ1 + KCNE3	2020	MSP2N2	полярные липиды сои	крио-ПЭМ	3.1	[94]
V/A-type ATPase	2019	MSP1E3D1	POPC и полярные липиды <i>T. thermophilus</i>	крио-ПЭМ	3.5	[95]
TMEM16F	2019	MSP2N2	полярные липиды сои, POPC, POPE, POPS	крио-ПЭМ	~4	[96]
AdeB	2019	MSP1E3D1	липидный экстракт <i>E. coli</i>	крио-ПЭМ	2.98	[97]
LRRC8A	2019	MSP1E3D1	POPC	крио-ПЭМ	4.18	[98]
H ⁺ -АТФаза	2019	MSP1E3D1	вакуолярные липиды дрожжей	ПЭМ		[99]
Orai	2019	MSP1E3D1	POPC, POPG, POPE	крио-ПЭМ	5.7	[85]
TRPM4	2018	MSP2N2	полярные липиды сои	крио-ПЭМ	~3	[100]
OSCA1.2	2018	MSP2N2	полярные липиды сои	крио-ПЭМ	3.1	[88]
TPC1	2018	MSP E3D1 и сапозин А	полярные липиды сои	крио-ПЭМ	3.7	[101]
Kv1.2–2.1	2018	MSP1E3D1	POPC, POPG, POPE	крио-ПЭМ	3.3	[102]
VDAC-1	2018	DCND	POPC, POPG с холестерином	ПЭМ		[90]
Kv7.1	2017	MSP2N2	без экзогенных липидов	ПЭМ	25	[103]
TcdA1	2016	MSP1D1	POPC	крио-ПЭМ	3.46	[104]
TRPV1	2016	MSP2N2 и MSP1E3	липиды сои	крио-ПЭМ	без лиганда – 3.2, с агонистом – 2.9, с антагонистом – 3.4	[105]
CorA	2016	MSP1D1	POPC, POPG	крио-ПЭМ	3.8	[102]

Примечание. DMPG – димиритоил-глицеро-фосфорилглицерин, POPG – пальмитоил-олеоил-глицеро-фосфоглицерин, POPE – пальмитоил-олеоил-глицерофосфэтаноламин, POPS – 1-пальмитоил-олеоил-глицерофосфосерин, DOPE – диолеоил-фосфатидил-этаноламин.

и при высокой концентрации NaCl растворяются быстрее и с большим выходом SMALP. Показано, что полимер не имеет сродства к каким-либо определенным липидам и в SMALP соотношение липидов остается таким же, как было в исходной мембране. Термодинамика процесса формирования SMALP исследована в [113] на модельных мембранах. Влияние липидного состава на формирование SMALPs отражено в кинетике солубилизации эукариотических клеток: по данным флуоресцентной микроскопии внутриклеточные мембраны в присутствии SMA разрушаются быстрее и эффективнее, чем плазматическая мембрана [114]. Более медленное разрушение плазматической мембраны может быть отчасти связано с присутствием большого количества структурных белков.

Если мембрана содержит молекулы белка, то при добавлении полимера белки оказываются заключены в формирующиеся липодиски, т.е. мембранный белок оказывается вырезан из мембраны с сохранением некоторого слоя окружающих

белок липидов. Первоначально это было продемонстрировано для липосом, образованных DMPC и содержащих бактериородопсин [115]. Позже при помощи полимера SMA были экстрагированы ABC-транспортёры из мембранных фракций различных клеток [116], комплекс IV из митохондрий клеток дрожжей [117], калиевый ПЗ-канал человека Kv7.1 из нефракционированных эукариотических клеток [118] и другие белки. Мембранные белки, заключенные в липодиски, оказываются достаточно стабильными и могут быть очищены и проанализированы различными биохимическими методами.

Значительным преимуществом использования SMA является полное отсутствие детергента в протоколе очистки. Как следствие, выделенные таким образом белки могут быть извлечены вместе с их естественной липидной средой и лигандами. Такой подход позволяет повысить стабильность и сохранить функциональную активность белков и белковых комплексов, что было показано при очистке аденозинового рецептора челове-

Таблица 2. Ионные каналы, структуры которых получены в SMA с помощью крио-ПЭМ

Ионный канал	Год	Разрешение, Å	Литература
KimA (<i>Bacillus subtilis</i>)	2020	3.7	[120]
ASIC1	2020	2.8, 3.7	[86]
AcrB (<i>Salmonella Typhimurium</i>)	2020	4.6	[127]
AcrB (<i>E. coli</i>)	2018	8.8	[128]
AcrB (<i>E. coli</i>)	2018	3.2	[129]
Alternative complex III (ACIII) (<i>Flavobacterium johnsoniae</i>)	2018	3.4	[130]

ка A2A [119], калиевого ПЗ-канала KcsA [120], транспортера AcrB *E. coli* [121], фотосинтетических реакционных центров *Rhodobacter (Rba.) sphaeroides* [122]. Часто SMA повышает общий выход солюбилизованного белка, а также позволяет провести очистку в одну стадию [116, 118].

Сохранение ближайшего липидного окружения при экстрагировании белков полимером SMA открывает широкие возможности для анализа их естественного липидного окружения и поиска липид-белковых взаимодействий [123–125]. В [126] совместное экстрагирование липидов с мембранными белками было использовано для определения липидного состава микродоменов мембран *Saccharomyces cerevisiae*.

Преимущества полимера SMA при солюбилизации мембранных белков делают его перспективным в структурных исследованиях. В табл. 2 приведены данные о структурах белков, полученных в составе липодисков, опубликованные за период 2018–2020 гг. Несмотря на сравнительно небольшое количество статей, исследования каналов с помощью SMA становятся популярны и позволяют получить субнанометровое разрешение реконструкций.

7. ИЗУЧЕНИЕ ДИНАМИКИ ИОННЫХ КАНАЛОВ С ПОМОЩЬЮ РЕНТГЕНОВСКИХ ЛАЗЕРОВ НА СВОБОДНЫХ ЭЛЕКТРОНАХ

С появлением рентгеновских лазеров на свободных электронах (РЛСЭ) и микрофокусных станций на синхротронных источниках излучения серийная кристаллография становится важнейшим методом для исследования мембранных белков [131]. Новый подход к получению дифракционных данных позволяет собрать структурную информацию с нано- и микрокристаллов при комнатной температуре с минимальным влиянием радиационного разрушения при помощи ультрабыстрого сбора данных за время, меньшее, чем характерное время глобального радиационного повреждения кристалла. Сочетание огромной яркости вспышки свыше 10 Вт/см² (сфокусированной на область ~100 нм²) и короткой дли-

тельности порядка 10 фс приводит к уникальным принципиальным возможностям этого инструмента в структурной биологии.

Мембранные белки при попытках кристаллизовать их часто имеют тенденцию к образованию очень маленьких кристаллов, размером около микро- или нанометра. Качество дифракции таких кристаллов обычно значительно хуже по сравнению с кристаллами растворимых белков аналогичного размера. В таких случаях добиться лучшего разрешения кристаллической структуры можно увеличением мощности рентгеновского источника, причем протяженность импульса должна быть меньше характерного времени разрушения кристалла, т.е. несколько фемтосекунд. Не так давно для кристаллизации мембранных белков стали использовать липидную кубическую фазу (LQP) [132]. Этот подход и разработка специальных инжекторов для вязких растворов позволили разработать метод серийной фемтосекундной кристаллографии (рис. 8а). Он заключается в сборе данных дифракции со множества микрокристаллов в LQP, причем с каждого кристалла успевают получить только одну картину дифракции в одной ориентации. Большой массив данных совмещается в один набор. Этот метод позволил получить структуры таких каналов, как GPCR-рецептор [133] и бактериородопсин [134]. Светоиндуцированная изомеризация ретиналя бактериородопсина является одной из самых быстрых реакций, известных в биологии, но теперь с помощью РЛСЭ оказалось возможно наблюдать сверхбыстрые фотохимические реакции, включая индуцированные светом молекулярные движения внутри ретиналя в пределах фемтосекунд с атомным разрешением (рис. 8б) [134].

ЗАКЛЮЧЕНИЕ

Мы являемся свидетелями бурного развития структурных исследований ионных каналов. За последние 5 лет с помощью крио-ПЭМ получено более 50 новых структур различных ионных каналов, что в несколько раз больше, чем полученных с помощью РСА за 20 лет. Крио-ПЭМ дает возможность получать реконструкции полноразмерных белков,

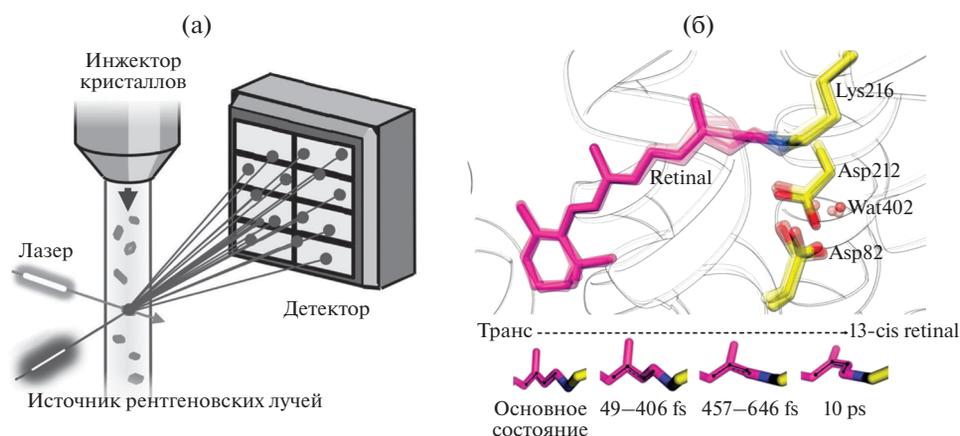


Рис. 8. Времыразрешенная серийная фемтосекундная кристаллография: а – схема установки, б – светоиндуцированная изомеризация ретиналя бактериородопсина [134].

которые ранее было невозможно закристаллизовать целиком из-за их размера и сложной организации, например рецептора глутамата [135].

С развитием структурной биологии электрофизиология получила новое развитие и используется для валидации новых структур. Функциональные исследования легко проводить в реальном времени: открытие и закрытие канала занимает секунды. Можно исследовать динамику сердечных каналов, что очень важно для понимания причин тяжелых заболеваний, которые вызывают мутации в генах ионных каналов. Кардиомиоциты, полученные из индуцированных плюрипотентных стволовых клеток, экспрессируют многие белки, которые присутствуют и в зрелых кардиомиоцитах, что позволяет получить данные одновременно для десятков и сотен сердечных ионных каналов.

До сих пор результатом структурных исследований являлось дискретное количество “застывших” конформаций, которые могли быть приведены в движение только благодаря МД с ограничением по времени порядка микросекунд. Однако и эти ограничения, оказалось, можно преодолеть. С развитием лазеров на свободных электронах метод серийной кристаллографии с временным разрешением позволил получить динамические картины активации ионных каналов [134].

Работа выполнена при поддержке Российского фонда фундаментальных исследований в части ионных каналов в SMA (грант № 18-504-12045), в части структуры и физиологии каналов KCNH2 и KCNQ1 (грант № 20-54-15004). Ж. Лоуссарн получил поддержку в рамках гранта CNRS PRC RUSSIE 2019 (PRC № 2773).

СПИСОК ЛИТЕРАТУРЫ

1. Natchiar S.K., Myasnikov A.G., Kratzat H. et al. // Nature. 2017. V. 551. № 7681. P. 472. <https://doi.org/10.1038/nature24482>
2. Suo Y., Wang Z., Zubcevic L. et al. // Neuron. 2020. V. 105. № 5. P. 882. <https://doi.org/10.1016/j.neuron.2019.11.023>
3. Wang W., MacKinnon R. // Cell. 2017. V. 169. № 3. P. 422. <https://doi.org/10.1016/j.cell.2017.03.048>
4. Chakraborty S., Jasnin M., Baumeister W. // Protein Sci. 2020. V. 29. № 6. P. 1302. <https://doi.org/10.1002/pro.3858>
5. Owji A.P., Qingqing Z., Changyi J. et al. // Nat. Struct. Mol. Biol. 2020. V. 27. № 4. P. 382. <https://doi.org/10.1038/s41594-020-0402-z>
6. Schmid S., Hugel T. // Elife. 2020. V. 9. <https://doi.org/10.7554/eLife.57180>
7. Jespersen T., Grunnet M., Olesen S.P. // Physiology (Bethesda). 2005. V. 20. P. 408. <https://doi.org/10.1152/physiol.00031.2005>
8. Peroz D., Rodriguez N., Choveau F. et al. // J. Physiol. 2008. V. 586. № 7. P. 1785. <https://doi.org/10.1113/jphysiol.2007.148254>
9. Duggal P., Vesely M.R., Wattanasirichaigoon D. et al. // Circulation. 1998. V. 97. № 2. P. 142. <https://doi.org/10.1161/01.cir.97.2.142>
10. Chen Y.H., Xu S.-J., Bendahhou S. et al. // Science. 2003. V. 299. № 5604. P. 251. <https://doi.org/10.1126/science.1077771>
11. Smith J.L., Anderson C.L., Burgess D.E. et al. // J. Arrhythm. 2016. V. 32. № 5. P. 373. <https://doi.org/10.1016/j.joa.2015.11.009>
12. Koponen M., Havulinna A.S., Marjamaa A. et al. // BMC Med. Genet. 2018. V. 19. № 1. P. 56. <https://doi.org/10.1186/s12881-018-0574-0>
13. Tester D.J., Ackerman M.J. // Methodist Debakey Cardiovasc. J. 2014. V. 10. № 1. P. 29. <https://doi.org/10.14797/mdcj-10-1-29>
14. Hu D., Yang L., Jiancheng Z. et al. // JACC Clin. Electrophysiol. 2017. V. 3. № 7. P. 727. <https://doi.org/10.1016/j.jacep.2016.11.013>
15. Wang Q.I., Seiko O., Ding W.-G. et al. // J. Cardiovasc. Electrophysiol. 2014. V. 25. № 5. P. 522. <https://doi.org/10.1111/jce.12361>

16. *Hong K., Bjerregaard P., Gussak I. et al.* // J. Cardiovasc. Electrophysiol. 2005. V. 16. № 4. P. 394.
<https://doi.org/10.1046/j.1540-8167.2005.40621.x>
17. *Hishigaki H., Kuhara S.* // Database (Oxford). 2011. P. bar017.
<https://doi.org/10.1093/database/bar017>
18. *Sanguinetti M.C., Chen J., Fernandez D. et al.* // Novartis Found Symp. 2005. V. 266. P. 159.
19. *Etheridge S.P., Asaki S.Y., Niu M.C.* // Curr. Opin. Cardiol. 2019. V. 34. № 1. P. 46.
<https://doi.org/10.1097/HCO.0000000000000587>
20. *Li M., Jan Y.N., Jan L.Y.* // Science. 1992. V. 257. № 5074. P. 1225.
<https://doi.org/10.1126/science.1519059>
21. *Doyle D.A., Cabral J.M., Pfuetzner R.A. et al.* // Science. 1998. V. 280. № 5360. P. 69.
<https://doi.org/10.1126/science.280.5360.69>
22. *Kreusch A., Pfaffinger P.J., Stevens C.F. et al.* // Nature. 1998. V. 392. № 6679. P. 945.
<https://doi.org/10.1038/31978>
23. *Kobertz W.R., Williams C., Miller C.* // Biochemistry. 2000. V. 39. № 34. P. 10347.
<https://doi.org/10.1021/bi001292j>
24. *Sokolova O., Kolmakova-Partensky L., Grigorieff N.* // Structure. 2001. V. 9. № 3. P. 215.
[https://doi.org/10.1016/s0969-2126\(01\)00578-0](https://doi.org/10.1016/s0969-2126(01)00578-0)
25. *Long S.B., Campbell E.B., Mackinnon R.* // Science. 2005. V. 309. № 5736. P. 897.
<https://doi.org/10.1126/science.1116269>
26. *Grizel A.V., Glukhov G.S., Sokolova O.S.* // Acta Naturae. 2014. V. 6. № 4. P. 10.
27. *Glauner K.S., Mannuzzu L.M., Gandhi C.S., Isacoff E.Y.* // Nature. 1999. V. 402. № 6763. P. 813.
<https://doi.org/10.1038/45561>
28. *Ferrer T., Rupp J., Piper D.R., Tristani-Firouzi M.* // J. Biol. Chem. 2006. V. 281. № 18. P. 12858.
<https://doi.org/10.1074/jbc.M513518200>
29. *Lee S.Y., Lee A., Chen J. et al.* // Proc. Natl. Acad. Sci. U. S. A. 2005. V. 102. № 43. P. 15441.
<https://doi.org/10.1073/pnas.0507651102>
30. *Yellen G.* // Nature. 2002. V. 419. № 6902. P. 35.
<https://doi.org/10.1038/nature00978>
31. *Holyoake J., Domene C., Bright J. et al.* // Eur. Biophys. J. 2004. V. 33. № 3. P. 238.
<https://doi.org/10.1007/s00249-003-0355-2>
32. *Jensen M., Jogini V., Borhani D. et al.* // Science. 2012. V. 336. № 6078. P. 229.
<https://doi.org/10.1126/science.1216533>
33. *Flynn G.E., Zagotta W.N.* // Neuron. 2001. V. 30. № 3. P. 689.
[https://doi.org/10.1016/s0896-6273\(01\)00324-5](https://doi.org/10.1016/s0896-6273(01)00324-5)
34. *MacKinnon R., Yellen G.* // Science. 1990. V. 250. № 4978. P. 276.
<https://doi.org/10.1126/science.2218530>
35. *Yool A.J., Schwarz T.L.* // Nature. 1991. V. 349. № 6311. P. 700.
<https://doi.org/10.1038/349700a0>
36. *Liu Y., Holmgren M., Jurman M. et al.* // Neuron. 1997. V. 19. № 1. P. 175.
[https://doi.org/10.1016/s0896-6273\(00\)80357-8](https://doi.org/10.1016/s0896-6273(00)80357-8)
37. *Schroede B.C., Waldegger S., Fehr S. et al.* // Nature. 2000. V. 403. № 6766. P. 196.
<https://doi.org/10.1038/35003200>
38. *Barhanin J., Lesage F., Guillemare E. et al.* // Nature. 1996. V. 384. № 6604. P. 78.
<https://doi.org/10.1038/384078a0>
39. *Sanguinetti M.C., Curran M., Zou A. et al.* // Nature. 1996. V. 384. № 6604. P. 80.
<https://doi.org/10.1038/384080a0>
40. *Loussouarn G., Baró I., Escand D.* // Methods Mol. Biol. 2006. V. 337. P. 167.
<https://doi.org/10.1385/1-59745-095-2:167>
41. *Keating M.* // Circulation. 1992. V. 85. № 6. P. 1973.
<https://doi.org/10.1161/01.cir.85.6.1973>
42. *Tester D.J., Will M., Haglund C. et al.* // Heart Rhythm. 2005. V. 2. № 5. P. 507.
<https://doi.org/10.1016/j.hrthm.2005.01.020>
43. *Gourraud J.B., Barc J., Thollet A. et al.* // Front. Cardiovasc. Med. 2016. V. 3. P. 9.
<https://doi.org/10.3389/fcvm.2016.00009>
44. *Probst V., Wilde A., Barc J. et al.* // Circ. Cardiovasc. Genet. 2009. V. 2. № 6. P. 552.
<https://doi.org/10.1161/CIRCGENET-ICS.109.853374>
45. *Loussouarn G., Park K.-H., Bellocq C. et al.* // EMBO J. 2003. V. 22. № 20. P. 5412.
<https://doi.org/10.1093/emboj/cdg526>
46. *Rodriguez N., Amarouch M., Montnach J. et al.* // Biophys. J. 2010. V. 99. № 4. P. 1110.
<https://doi.org/10.1016/j.bpj.2010.06.013>
47. *Park K.H., Piron J., Dahimene S. et al.* // Circ. Res. 2005. V. 96. № 7. P. 730.
<https://doi.org/10.1161/01.RES.0000161451.04649.a8>
48. *Choveau F.S., Rodriguez N., Abderemane F. et al.* // J. Biol. Chem. 2011. V. 286. № 1. P. 707.
<https://doi.org/10.1074/jbc.M110.146324>
49. *Malak O.A., Es-Salah-Lamoureux Z., Loussouarn G.* // Sci. Rep. 2017. V. 7. № 1. P. 113.
<https://doi.org/10.1038/s41598-017-00155-2>
50. *Malak O.A., Gluhov G., Grizel A. et al.* // J. Biol. Chem. 2019. V. 294. № 16. P. 6506.
<https://doi.org/10.1074/jbc.RA119.007626>
51. *Malak O.A., Abderemane-Ali F., Wei Y. et al.* // Sci. Rep. 2020. V. 10. № 1. P. 5852.
<https://doi.org/10.1038/s41598-020-62615-6>
52. *Ng C.A., Perry M.D., Liang W. et al.* // Heart Rhythm. 2020. V. 17. № 3. P. 492.
<https://doi.org/10.1016/j.hrthm.2019.09.020>
53. *Vanoye C.G., Desai R., Fabreet K. et al.* // Circ. Genom. Precis. Med. 2018. V. 11. № 11. P. e002345.
<https://doi.org/10.1161/CIRCGEN.118.002345>
54. *Takumi T., Ohkubo H., Nakanishi S.* // Science. 1988. V. 242. № 4881. P. 1042.
<https://doi.org/10.1126/science.3194754>
55. *Ben-Efraim I., Bach D., Shai Y.* // Biochemistry. 1993. V. 32. № 9. P. 2371.
<https://doi.org/10.1021/bi00060a031>
56. *Lesage F., Attali B., Lakey J. et al.* // Receptors Channels. 1993. V. 1. № 2. P. 143.

57. *van Huizen R., Czajkowsky D., Shi D. et al.* // FEBS Lett. 1999. V. 457. № 1. P. 107.
[https://doi.org/10.1016/s0014-5793\(99\)01021-2](https://doi.org/10.1016/s0014-5793(99)01021-2)
58. *Li M., Li M., Unwin N. et al.* // Curr. Biol. 1994. V. 4. № 2. P. 110.
[https://doi.org/10.1016/s0960-9822\(94\)00026-6](https://doi.org/10.1016/s0960-9822(94)00026-6)
59. *Orlova E.V., Papakosta M., Booy F. et al.* // J. Mol. Biol. 2003. V. 326. № 4. P. 1005.
[https://doi.org/10.1016/s0022-2836\(02\)00708-8](https://doi.org/10.1016/s0022-2836(02)00708-8)
60. *Sokolova O., Accardi A., Gutierrez D. et al.* // Proc. Natl. Acad. Sci. U. S. A. 2003. V. 100. № 22. P. 12607.
<https://doi.org/10.1073/pnas.2235650100>
61. *Wolf M., Eberhart A., Glossmann H. et al.* // J. Mol. Biol. 2003. V. 332. № 1. P. 171.
[https://doi.org/10.1016/s0022-2836\(03\)00899-4](https://doi.org/10.1016/s0022-2836(03)00899-4)
62. *Jiang Q.X., Wang D.N., MacKinnon R.* // Nature. 2004. V. 430. № 7001. P. 806.
<https://doi.org/10.1038/nature02735>
63. *Samsó M., Wagenknecht T., Allen P.D.* // Nat. Struct. Mol. Biol. 2005. V. 12. № 6. P. 539.
<https://doi.org/10.1038/nsmb938>
64. *Serysheva I.I., Ludtke S., Baker M. et al.* // Proc. Natl. Acad. Sci. U. S. A. 2008. V. 105. № 28. P. 9610.
<https://doi.org/10.1073/pnas.0803189105>
65. *Mishyna M., Volokh O., Danilova Y. et al.* // Micron. 2017. V. 96. P. 57.
<https://doi.org/10.1016/j.micron.2017.02.004>
66. *Brilot A.F., Chen J., Cheng A. et al.* // J. Struct. Biol. 2012. V. 177. № 3. P. 630.
<https://doi.org/10.1016/j.jsb.2012.02.003>
67. *Corringer P.J., Baaden M., Bocquet N. et al.* // J. Physiol. 2010. V. 588. Pt. 4. P. 565.
<https://doi.org/10.1113/jphysiol.2009.183160>
68. *Kuo A., Gulbis J., Antcliff J. et al.* // Science. 2003. V. 300. № 5627. P. 1922.
<https://doi.org/10.1126/science.1085028>
69. *Jiang Y., Lee A., Chenet J. et al.* // Nature. 2003. V. 423. № 6935. P. 33.
<https://doi.org/10.1038/nature01580>
70. *Kawate T., Michel J., Birdsong W., Gouaux E.* // Nature. 2009. V. 460. № 7255. P. 592.
<https://doi.org/10.1038/nature08198>
71. *Long S.B., Tao X., Campbell E., MacKinnon R.* // Nature. 2007. V. 450. № 7168. P. 376.
<https://doi.org/10.1038/nature06265>
72. *Brohawn S.G., del Mármol J., MacKinnon R.* // Science. 2012. V. 335. № 6067. P. 436.
<https://doi.org/10.1126/science.1213808>
73. *Liao M., Cao E., Julius D. et al.* // Nature. 2013. V. 504. № 7478. P. 107.
<https://doi.org/10.1038/nature12822>
74. *Scheres S.H.* // Methods Enzymol. 2016. V. 579. P. 125.
<https://doi.org/10.1016/bs.mie.2016.04.012>
75. *Whicher J.R., MacKinnon R.* // Science. 2016. V. 353. № 6300. P. 664.
<https://doi.org/10.1126/science.aaf8070>
76. *Sun J., MacKinnon R.* // Cell. 2017. V. 169. № 6. P. 1042.
<https://doi.org/10.1016/j.cell.2017.05.019>
77. *Jiang Y., Ruta V., Chen J. et al.* // Nature. 2003. V. 423. № 6935. P. 42.
<https://doi.org/10.1038/nature01581>
78. *Lörinczi É., Camilo Gómez-Posada J., de la Peña P. et al.* // Nat. Commun. 2015. V. 6. P. 6672.
<https://doi.org/10.1038/ncomms7672>
79. *Ferrer T., Rupp J., Piper D. et al.* // J. Biol. Chem. 2006. V. 281. № 18. P. 12858.
<https://doi.org/10.1074/jbc.M513518200>
80. *Schertler G.F., Villa C., Henderson R.* // Nature. 1993. V. 362. № 6422. P. 770.
<https://doi.org/10.1038/362770a0>
81. *Unwin N.* // Nature. 1995. V. 373. № 6509. P. 37.
<https://doi.org/10.1038/373037a0>
82. *Jiang Q.X., Chester D.W., Sigworth F.J.* // J. Struct. Biol. 2001. V. 133. № 2–3. P. 119.
<https://doi.org/10.1006/jsbi.2001.4376>
83. *Antonenko Y.N., Gluhov G., Firsov A. et al.* // Phys. Chem. Chem. Phys. 2015. V. 17. № 26. P. 17461.
<https://doi.org/10.1039/c5cp02047f>
84. *Marius P.S., East Alvis J., Lee A.* // Biophys. J. 2005. V. 89. № 6. P. 4081.
<https://doi.org/10.1529/biophysj.105.070755>
85. *Liu X., Wu G., Yu Y. et al.* // PLoS Biol. 2019. V. 17. № 4. P. e3000096.
<https://doi.org/10.1371/journal.pbio.3000096>
86. *Yoder N., Gouaux E.* // Elife. 2020. V. 9.
<https://doi.org/10.7554/eLife.56527>
87. *Peters-Libeu C.A., Newhouse Y., Hall S. et al.* // J. Lipid. Res. 2007. V. 48. № 5. P. 1035.
<https://doi.org/10.1194/jlr.M600545-JLR200>
88. *Johansen N.T., Tidemand F., Nguyen Tam T. et al.* // FEBS J. 2019. V. 286. № 9. P. 1734.
<https://doi.org/10.1111/febs.14766>
89. *Zhao Z., Zhang M., Hogle J. et al.* // J. Am. Chem. Soc. 2018. V. 140. № 34. P. 10639.
<https://doi.org/10.1021/jacs.8b04638>
90. *Nasr M.L., Baptista D., Strauss M. et al.* // Nat. Methods. 2017. V. 14. № 1. P. 49.
<https://doi.org/10.1038/nmeth.4079>
91. *Hagn F., Nasr M.L., Wagner G.* // Nat. Protoc. 2018. V. 13. № 1. P. 79.
<https://doi.org/10.1038/nprot.2017.094>
92. *Reid M.S., Kern D.M., Brohawn S.G.* // Elife. 2020. V. 9.
<https://doi.org/10.7554/eLife.52505>
93. *Gardill B., Huang J., Tu L. et al.* // Sci. Rep. 2020. V. 10. № 1. P. 1130.
<https://doi.org/10.1038/s41598-020-58002-w>
94. *Sun J., MacKinnon R.* // Cell. 2020. V. 180. № 2. P. 340.
<https://doi.org/10.1016/j.cell.2019.12.003>
95. *Zhou L., Sazanov L.A.* // Science. 2019. V. 365. № 6455.
<https://doi.org/10.1126/science.aaw9144>
96. *Feng S., Dang S., Han T. et al.* // Cell. Rep. 2019. V. 28. № 2. P. 567e4.
<https://doi.org/10.1016/j.celrep.2019.06.023>
97. *Su C.C., Morgan C., Kambakam S. et al.* // mBio. 2019. V. 10. № 4.
<https://doi.org/10.1128/mBio.01295-19>

98. Kern D.M., Oh S., Hite R. et al. // *Elife*. 2019. V. 8. <https://doi.org/10.7554/eLife.42636>
99. Sharma S., Oot R., Khan M., Wilkens S. // *J. Biol. Chem.* 2019. V. 294. № 16. P. 6439. <https://doi.org/10.1074/jbc.RA119.007577>
100. Autzen H.E., Myasnikov A., Campbell M. et al. // *Science*. 2018. V. 359. № 6372. P. 228. <https://doi.org/10.1126/science.aar4510>
101. Kintzer A.F., Green E., Dominik P. et al. // *Proc. Natl. Acad. Sci. U. S. A.* 2018. V. 115. № 39. P. E9095. <https://doi.org/10.1073/pnas.1805651115>
102. Matthies D., Dalmas O., Borgnia M. et al. // *Cell*. 2016. V. 164. № 4. P. 747. <https://doi.org/10.1016/j.cell.2015.12.055>
103. Shenkarev Z.O., Karlova M., Kulbatskii D. et al. // *Biochemistry (Mosc)*, 2018. V. 83. № 5. P. 562. <https://doi.org/10.1134/S0006297918050097>
104. Gatsogiannis C., Merino F., Prumbaum D. et al. // *Nat. Struct. Mol. Biol.* 2016. V. 23. № 10. P. 884. <https://doi.org/10.1038/nsmb.3281>
105. Gao Y., Cao E., Julius D., Cheng Y. // *Nature*. 2016. V. 534. № 7607. P. 347. <https://doi.org/10.1038/nature17964>
106. Tonge S.R., Tighe B.J. // *Adv. Drug. Deliv. Rev.* 2001. V. 53. № 1. P. 109. [https://doi.org/10.1016/s0169-409x\(01\)00223-x](https://doi.org/10.1016/s0169-409x(01)00223-x)
107. Scheidelaar S., Koorengel M., van Walree C. et al. // *Biophys. J.* 2016. V. 111. № 9. P. 1974. <https://doi.org/10.1016/j.bpj.2016.09.025>
108. Colbasevici A., Voskoboynikova N., Orekhov P. et al. // *Biochim. Biophys. Acta Biomembr.* 2020. V. 1862. № 5. P. 183207. <https://doi.org/10.1016/j.bbamem.2020.183207>
109. Domínguez Pardo J.J., Koorengel M., Uwugiaren N. et al. // *Biophys. J.* 2018. V. 115 (1). P. 129. <https://doi.org/10.1016/j.bpj.2018.05.032>
110. Zhang R., Sahu I., Liu L. et al. // *Biochim. Biophys. Acta*. 2015. V. 1848. № 1. Pt. B. P. 329. <https://doi.org/10.1016/j.bbamem.2014.05.008>
111. Scheidelaar S., Koorengel M., Dominguez Pardo J. et al. // *Biophys. J.* 2015. V. 108. № 2. P. 279. <https://doi.org/10.1016/j.bpj.2014.11.3464>
112. Dominguez Pardo J.J., Dörr J., Iyer A. et al. // *Eur. Biophys. J.* 2017. V. 46. № 1. P. 91. <https://doi.org/10.1007/s00249-016-1181-7>
113. Cuevas Arenas R., Klingler J., Vargas C., Keller S. // *Nanoscale*. 2016. V. 8. № 32. P. 15016. <https://doi.org/10.1039/c6nr02089e>
114. Dörr J.M., van Coevorden-Hameete M., Hoogenraad C., Killian J. // *Biochim. Biophys. Acta. Biomembr.* 2017. V. 1859. № 11. P. 2155. <https://doi.org/10.1016/j.bbamem.2017.08.010>
115. Knowles T.J., Finka R., Smith C. et al. // *J. Am. Chem. Soc.* 2009. V. 131. № 22. P. 7484. <https://doi.org/10.1021/ja810046q>
116. Gulati S., Jamshad M., Knowles T. et al. // *Biochem. J.* 2014. V. 461. № 2. P. 269. <https://doi.org/10.1042/BJ20131477>
117. Long A.R., O'Brien C., Malhotra K. et al. // *BMC Biotechnol.* 2013. V. 13. P. 41. <https://doi.org/10.1186/1472-6750-13-41>
118. Karlova M.G., Voskoboynikova N., Gluhov G. et al. // *Chem. Phys. Lipids*. 2019. V. 219. P. 50. <https://doi.org/10.1016/j.chemphyslip.2019.01.013>
119. Jamshad M., Charlton J., Lin Y. et al. // *Biosci. Rep.* 2015. V. 35. № 2. <https://doi.org/10.1042/BSR20140171>
120. Dörr J.M., Koorengel M., Schäfer M. et al. // *Proc. Natl. Acad. Sci. U. S. A.* 2014. V. 111. № 52. P. 18607. <https://doi.org/10.1073/pnas.1416205112>
121. Postis V., Rawson S., Mitchell J. et al. // *Biochim. Biophys. Acta*. 2015. V. 1848. № 2. P. 496. <https://doi.org/10.1016/j.bbamem.2014.10.018>
122. Swainsbury D.J., Scheidelaar S., van Grondelle R. et al. // *Angew. Chem. Int. Ed. Engl.* 2014. V. 53. № 44. P. 11803. <https://doi.org/10.1002/anie.201406412>
123. Schmidt V., Sidore M., Bechara C. et al. // *Biochim. Biophys. Acta. Biomembr.* 2019. V. 1861. № 2. P. 431. <https://doi.org/10.1016/j.bbamem.2018.10.017>
124. Swainsbury D.J.K., Proctor M., Hitchcock A. et al. // *Biochim. Biophys. Acta Bioenerg.* 2018. V. 1859. № 3. P. 215. <https://doi.org/10.1016/j.bbabi.2017.12.005>
125. Prabudiansyah I., Kusters I., Caforio A., Driessen A. // *Biochim. Biophys. Acta*. 2015. V. 1848. № 10. Pt. A. P. 2050. <https://doi.org/10.1016/j.bbamem.2015.06.024>
126. van 't Klooster J.S., Cheng T., Sikkema H. et al. // *Elife*. 2020. V. 9. <https://doi.org/10.7554/eLife.57003>
127. Parmar M., Rawson S., Scarff C. et al. // *Biochim. Biophys. Acta Biomembr.* 2018. V. 1860. № 2. P. 378. <https://doi.org/10.1016/j.bbamem.2017.10.005>
128. Hernando M., Orriss G., Perodeau J. et al. // *Biochim. Biophys. Acta Biomembr.* 2020. V. 1862. № 5. P. 183191. <https://doi.org/10.1016/j.bbamem.2020.183191>
129. Johnson R.M., Fais C., Parmar M. et al. // *Microorganisms*. 2020. V. 8. № 6. P. 943. <https://doi.org/10.3390/microorganisms8060943>
130. Sun C., Benlekbir S., Venkatakrishnan P. et al. // *Nature*. 2018. V. 557. № 7703. P. 123. <https://doi.org/10.1038/s41586-018-0061-y>
131. Standfuss J. // *Curr. Opin. Struct. Biol.* 2019. V. 57. P. 63. <https://doi.org/10.1016/j.sbi.2019.02.001>
132. Caffrey M., Cherezov V. // *Nat. Protoc.* 2009. V. 4. № 5. P. 706. <https://doi.org/10.1038/nprot.2009.31>
133. Cherezov V., Rosenbaum D., Hanson M. et al. // *Science*. 2007. V. 318. № 5854. P. 1258. <https://doi.org/10.1126/science.1150577>
134. Nogly P., Weinert T., James D. et al. // *Science*. 2018. V. 361. № 6398. P. eaat0094. <https://doi.org/10.1126/science.aat0094>
135. Twomey E.C., Sobolevsky A.I. // *Biochemistry*. 2018. V. 57. № 3. P. 267. <https://doi.org/10.1021/acs.biochem.7b00891>

УДК 548.73

СТРУКТУРНЫЕ ИССЛЕДОВАНИЯ РНК-БЕЛКОВЫХ КОМПЛЕКСОВ В ПОСТРИБОСОМНУЮ ЭРУ

© 2021 г. С. В. Тищенко¹, А. О. Михайлина¹, Н. В. Леконцева¹, Е. А. Столбоушкина¹,
Е. Ю. Никонова¹, О. С. Никонов¹, А. Д. Никулин^{1,*}

¹ Институт белка РАН, Пущино, Россия

*E-mail: nikulin@vega.protres.ru

Поступила в редакцию 15.09.2020 г.

После доработки 02.10.2020 г.

Принята к публикации 02.10.2020 г.

Структурные исследования РНК-белковых комплексов являются важным этапом в понимании многих молекулярных механизмов в клетке, например, таких как регуляция белкового синтеза и РНК-шаперонная активность белков. В обзоре рассмотрены различные объекты, исследующиеся в Институте белка РАН. На основе анализа структур комплексов рибосомного белка L1 со специфическими участками на мРНК и рРНК представлены принципы, по которым осуществляется регуляция белком трансляции мРНК своего оперона. Описаны исследования гетеротримерного фактора инициации трансляции IF2 архей и эукариот, приведены данные о взаимодействии глицил-тРНК-синтетазы с IRES вирусов. Приведены результаты исследований взаимодействия молекул РНК с одним из функционально важных участков белка Hfq и выявлены различия в РНК-связывающих свойствах Hfq и архейных белков семейства Lsm.

DOI: 10.31857/S0023476121050210

ОГЛАВЛЕНИЕ

Введение

1. Регуляторные комплексы рибосомного белка L1 с мРНК
2. Определение структуры и исследования РНК-связывающих свойств глобального регулятора трансляции – белка Hfq
3. Исследования РНК-связывающих свойств архейных Lsm белков
4. Исследование функции регуляторов транскрипции UxuR и ExuR
5. Структурные и функциональные исследования глицил-тРНК-синтетазы человека
6. Структурные исследования гетеротримерного фактора инициации трансляции 2

Заключение

ВВЕДЕНИЕ

Определение структур бактериальных рибосом с помощью методов рентгеноструктурного анализа и криоэлектронной микроскопии на рубеже XX и XXI веков стало прорывным событием в исследованиях процесса биосинтеза белка в клетках. Важной стартовой частью этого проекта стало определение структур отдельных компонентов рибосомы, а именно, рибосомных белков

и их комплексов с фрагментами рибосомной РНК. Они позволили ускорить интерпретацию получаемых огромных карт электронной плотности рибосом, а также описать подвижные участки рибосомных субчастиц, которые не видны на структурах целых рибосом. В Институте белка РАН (ИБ РАН) найдены условия кристаллизации 70S рибосом и 30S рибосомных субчастиц *Thermus thermophilus*, определены структуры более пятнадцати рибосомных белков и ряда рРНК-белковых комплексов, что внесло существенный вклад в реализацию масштабного “рибосомного проекта”. Разработан комплексный подход к определению структур РНК-белковых комплексов, включающий в себя рациональный дизайн фрагментов РНК, образующих наиболее стабильные комплексы с белками, кристаллизацию этих комплексов, получение дифракционных данных на лабораторной установке Proteum (Bruker-AXS) и на источниках синхротронного излучения BESSY, DESY, ESRF, MAX IV с последующим определением структур методами многоволновой аномальной дифракции и мультимолекулярного замещения. Накопленный опыт по структурно-функциональным исследованиям рРНК-белковых комплексов позволил перейти к решению новых задач, связанных с изучением целого ряда регуляторных комплексов. Исследования регуляции трансляции и транскрипции в

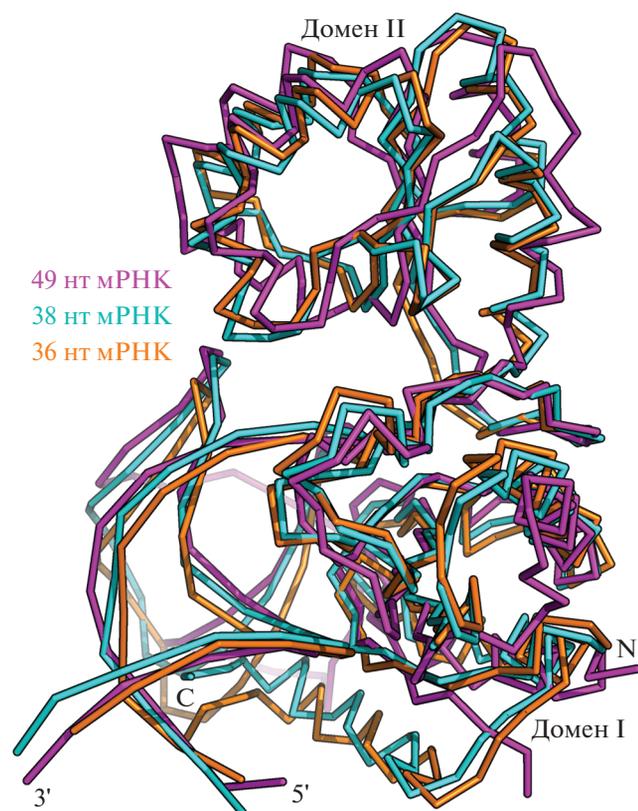


Рис. 1. Наложение структур комплексов рибосомного белка L1 с фрагментами мРНК длиной 49 нуклеотидов (серый, PDB ID 1U63), 38 нуклеотидов (белый, PDB ID 1ZHO) и 36 нуклеотидов (черный, PDB ID 2HW8). Указаны 5'- и 3'-концы РНК, домены I и II белка L1.

клетках на атомарном уровне позволяют находить ответы на новые вопросы, стоящие перед научным сообществом. В предлагаемом обзоре представлены краткие результаты наиболее интересных работ сотрудников ИБ РАН в этом направлении.

1. РЕГУЛЯТОРНЫЕ КОМПЛЕКСЫ РИБОСОМНОГО БЕЛКА L1 С МРНК

Рибосомный двухдоменный белок L1 является составной частью L1-выступа большой субчастицы рибосомы; у прокариот этот белок также регулирует трансляцию белков своего оперона по принципу обратной связи. В *E. coli* белок L1 связывается с участком мРНК перед первым геном L1 оперона [1], в экстремально-термофильных бактериях *Thermus thermophilus* и *Thermotoga maritima* имеются два участка связывания на мРНК [2], в археях *Methanococcus vannielii* и *Methanococcus jannaschii* регуляторный участок на мРНК располагается в начале кодирующей последовательности первого гена L1 оперона [3, 4]. Исследование регуляторных свойств рибосомного белка L1

явилось естественным продолжением работ по изучению структурной основы биосинтеза белка.

Белок L1 – единственный рибосомный белок, регуляторные свойства которого детально исследованы на основе сравнительного анализа структур комплексов L1-рРНК и L1-мРНК. Для получения кристаллизуемых препаратов комплекса L1-мРНК применена стратегия, разработанная в ИБ РАН при исследовании структуры комплекса рибосомного белка S8 со специфическим фрагментом 16S рРНК [5]. Были получены 10 фрагментов мРНК длиной от 36 до 49 нуклеотидов, специфически связывающихся с белком L1. Менялась длина и нуклеотидный состав двух спиралей, фланкирующих ассиметричную петлю, содержащую консервативные нуклеотиды. Верхняя полноразмерная спирально-петельная часть фрагментов мРНК, не содержащая консервативных нуклеотидов, в укороченных вариантах была заменена четырехнуклеотидной петлей UUCG [6]. Семь фрагментов представляли собой различные вариации L1-связывающего фрагмента мРНК *M. vannielii*, три фрагмента являлись вариациями специфического участка мРНК *M. jannaschii*. Высокую мозаичность кристаллов удалось понизить добавлением солей ртути в кристаллизационную каплю. Были определены структуры трех комплексов: комплекса L1-мРНК (49 нт) из *M. jannaschii* (разрешение 3.4 Å [7]) и двух гетерологичных комплексов, образованных белком L1 *T. thermophilus* и короткими фрагментами мРНК *M. vannielii*, длиной 38 (разрешение 2.6 Å [8]) и 36 нт (разрешение 2.1 Å [9]).

Белок L1 в перечисленных комплексах контактировал преимущественно с атомами сахарофосфатного остова мРНК. Сравнение структур рибосомного [10, 11] и полученных регуляторных комплексов белка L1 показало, что белок образует сеть недоступных растворителю водородных связей с мотивом, содержащим строго консервативные нуклеотиды и расположенным в месте сочленения двух спиралей, структурно инвариантным как на мРНК, так и на рРНК. Со стороны белка L1 в этом взаимодействии участвуют расположенные рядом петли домена I (рис. 1).

Для исследования особенностей взаимодействия белка L1 со специфическими фрагментами мРНК и рРНК применяли комбинированный структурно-функциональный подход, добавив к структурному анализу метод поверхностного плазмонного резонанса, который позволяет наблюдать кинетику взаимодействия молекул в реальном времени [12]. Специфический фрагмент мРНК отличается от L1-связывающего фрагмента рРНК сильно укороченными участками, соответствующими спирали 78 и петле Б рибосомной РНК (рис. 2). Это приводит к понижению сред-

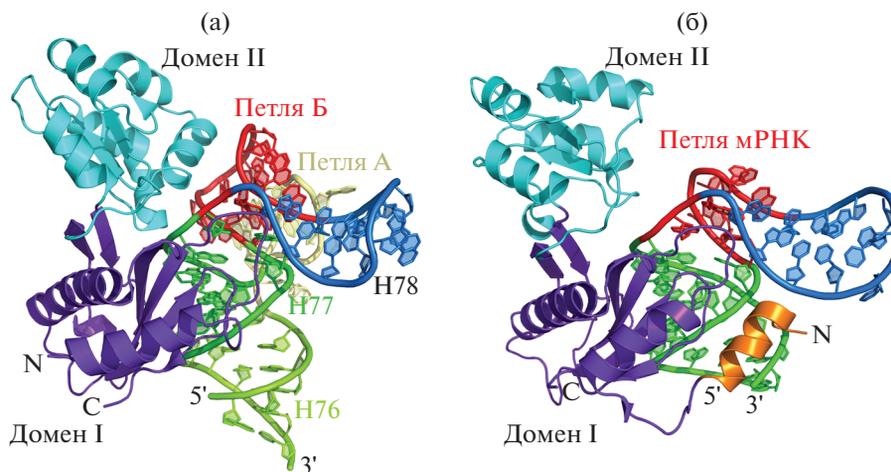


Рис. 2. Сравнение структур комплексов рибосомного белка L1 с фрагментом 23S рРНК (PDB ID 1MZP) (а) и белка L1 с мРНК (PDB ID 2HW8) (б). Указаны элементы структур РНК и белка.

ства белка L1 к мРНК на несколько порядков по сравнению с комплексом L1-рРНК.

Основная консервативная область контакта белка L1 расположена на его домене I и может быть разделена на два участка: один из них включает в себя консервативную триаду Thr-Met-Gly, образующую закрытые от растворителя водородные связи с нуклеотидами спирали H77 рРНК [13], другой – N-концевую часть спирали α_1 , взаимодействующую с нуклеотидами спирали H78 [14]. В комплексах белка L1 с мРНК и рРНК оба участка взаимодействия на домене I имеют одинаковые консервативные водородные связи, однако количество неконсервативных контактов в комплексе с мРНК уменьшается, что значительно снижает сродство белка к матричной РНК. Был получен домен I белка L1 в изолированном виде и показано, что он достаточен для взаимодействия как с рРНК, так и с мРНК [15, 16], а также обладает регуляторными свойствами целого белка [17]. Аминокислотные остатки домена II белка в рибосомном комплексе образуют дополнительные контакты с рРНК, что еще больше увеличивает стабильность рибосомного комплекса. Такое комплексное исследование комплексов белка L1 с мРНК и рРНК выявило причину разницы в сродстве белка к двум разным мишеням и позволило раскрыть основу регуляции трансляции рибосомным белком L1 по механизму обратной связи.

2. ОПРЕДЕЛЕНИЕ СТРУКТУРЫ И ИССЛЕДОВАНИЯ РНК-СВЯЗЫВАЮЩИХ СВОЙСТВ ГЛОБАЛЬНОГО РЕГУЛЯТОРА ТРАНСЛЯЦИИ – БЕЛКА Hfq

РНК-связывающий бактериальный белок Hfq принадлежит к обширному семейству Sm/Lsm

белков и в бактериях выполняет роль многофункционального белка-регулятора трансляции [18]. Это небольшой белок размером ~ 10 кДа, формирующий в растворе четвертичную структуру в виде гексамера [18]. В 1979 г. белок Hfq был идентифицирован как хозяйский фактор, необходимый для репликации (+)-цепи РНК бактериофага Q β [19]. В дальнейшем было выявлено, что делеция гена *hfq* в клетках *Escherichia coli* приводит к изменению продукции более чем 30 различных белков, включая белки репарации ДНК [20], вовлеченные в метаболизм железа [21] и синтез σ^s -субъединицы бактериальной РНК-полимеразы [22]. В большинстве случаев регуляция трансляции белком Hfq реализуется посредством *транс*-кодируемых малых регуляторных РНК (**мрРНК**; в английской литературе используют аббревиатуру sRNA). Белок Hfq входит в когорту так называемых “РНК-шаперонов”, включающих в себя такие бактериальные регуляторные белки, как CsrA, ProQ, KhpA, KhpB, SpoVG, способствующие взаимодействию мрРНК со своими мишенями на мРНК [23]. Этот способ регуляции позволяет контролировать адаптацию бактерий к изменению внешней среды за короткое время. Hfq – наиболее универсальный белок среди РНК-шаперонов и обнаружен в большинстве бактерий. Он стимулирует взаимодействие мрРНК с комплементарными участками на мРНК с последующей репрессией или стимуляцией трансляции, а также регулирует время жизни мРНК за счет взаимодействия с их поли(А)3'-концевыми участками.

В начале 2000 г. были получены первые структуры белка из *Staphylococcus aureus* [24], укороченной формы белка из *E. coli* [25] и полноразмерного белка из *Pseudomonas aeruginosa* [26], что позволило предложить ряд гипотез о том, как белок

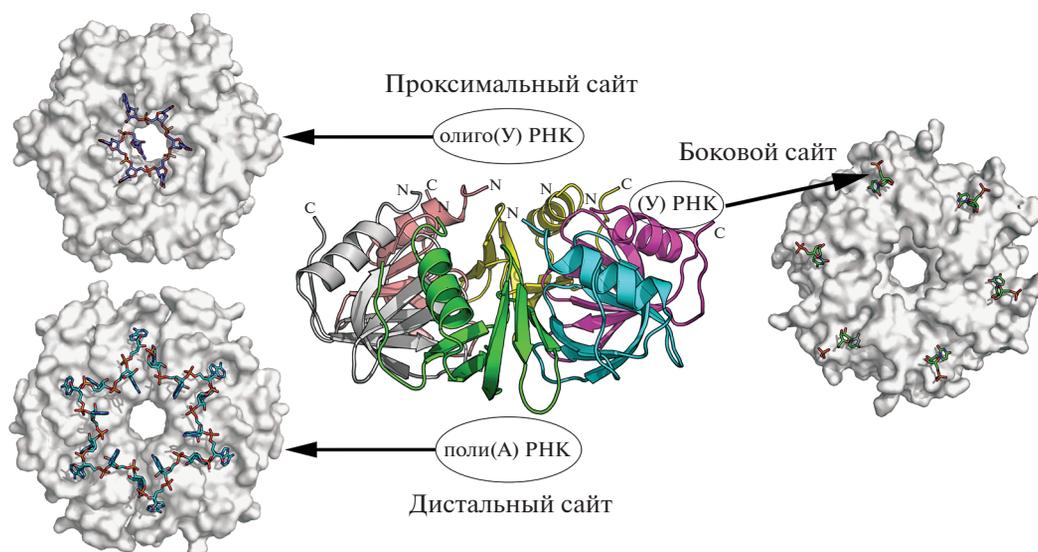


Рис. 3. Расположение участков взаимодействия белка Hfq с РНК. В центре — вид на гексамер белка с боковой стороны. Указаны N- и C-концы мономеров белка. Выделены области связывания РНК: уридин-специфический проксимальный сайт, аденин-специфический дистальный сайт и один из шести уридин-специфических боковых сайтов. Слева вверху показана структура комплекса белка Hfq с 5'-GUUUUUA-3' РНК [25] (вид с проксимальной стороны гексамера; PDB ID 1KQ2), слева внизу — структура комплекса Hfq с олиго(A) РНК [68] (вид со стороны дистального сайта; PDB ID 3GIB), справа посередине — структура комплекса Hfq с УМФ [28] (вид со стороны проксимального сайта; PDB ID 4J6X). Молекулы РНК и нуклеотидов показаны шаростержневой моделью на поверхности гексамеров белка Hfq.

способствует взаимодействию двух разных РНК. Были идентифицированы два основных участка связывания РНК на поверхности гексамера белка: уридин-специфический участок на “проксимальной” поверхности белка (со стороны N-концов мономеров белка) и аденин-специфический на противоположной, “дистальной”, поверхности (рис. 3).

Структуры комплексов Hfq с короткими фрагментами РНК показали детали узнавания оснований РНК аминокислотными остатками белка и объяснили его высокую специфичность к определенным последовательностям РНК. Интересно, что для разрушения вторичной структуры РНК белком Hfq не требуется дополнительной энергии в виде расщепления АТФ или ГТФ, однако ряд биохимических данных позволял предположить наличие АТФазной активности белка [27]. Мы получили структуру комплекса белка Hfq с АТФ и, исходя из особенностей контактов АТФ с аминокислотными остатками белка, сделали вывод, что вероятность специфической АТФазной активности Hfq крайне мала. Параллельно проверили возможность других одиночных нуклеотидов связываться с поверхностью белка и оказалось, что Hfq обладает высоким сродством к УТФ и ЦТФ, но не связывает ГТФ. Положение связанных нуклеотидов и ориентация их оснований полностью соответствовали таковым в структурах комплексов Hfq с олигонуклеотидами [28]. На ос-

нове этих наблюдений была предложена методика картирования участков связывания одноцепочечных РНК на поверхности РНК-связывающих белков [29, 30], которая в дальнейшем использовалась при исследовании гомологичных архейных Lsm белков. Более того, удалось идентифицировать третий, “латеральный” участок взаимодействия белка Hfq с РНК (рис. 3). Он расположен на боковой поверхности гексамера белка в области N-концевой спирали мономера и, как оказалось, является местом встречи и взаимодействия двух РНК — мРНК и мРНК. Данное открытие было сделано одновременно с идентификацией этого участка биохимическими методами, что подтвердило достоверность наших результатов [31–34]. Таким образом, мы впервые показали на структурном уровне, как происходит фиксация цепи РНК с последующим ее взаимодействием с белковым партнером. Эти данные внесли существенный вклад в дальнейшее развитие модели взаимодействия мРНК с мРНК с участием белка Hfq и процессов регуляции трансляции, опосредованное белками РНК-шаперонами.

3. ИССЛЕДОВАНИЯ РНК-СВЯЗЫВАЮЩИХ СВОЙСТВ АРХЕЙНЫХ LSM БЕЛКОВ

Архейные Lsm белки (или SmAP — Sm Archaeal Proteins), как и бактериальные белки Hfq, принадлежат к семейству Sm/Lsm и имеют аналогичную третичную структуру из пяти-тяжелого β -ли-

ста с N-концевой α -спиралью [35]. Геномы различных видов архей могут кодировать от одного до трех паралога SmAP белков, классифицируемых как SmAP1, SmAP2 и SmAP3. Интересно, что SmAP1 и SmAP2 из клеток одного вида имеют меньшую степень гомологии, чем SmAP1 (или SmAP2) из клеток разных организмов. Однако все они имеют строгий консенсус по аминокислотной последовательности RGXX (где X – положительно заряженный аминокислотный остаток) в отличие от консенсуса YKNAI бактериальных белков Hfq (рис. 4а). Кроме этого, в отличие от большинства бактериальных белков Hfq архейные SmAP белки имеют удлиненную петлю L4, соединяющую тяжи β 3 и β 4, и лишены протяженной неупорядоченной C-концевой части. Наиболее заметное отличие SmAP от бактериальных гомологов – формирование ими четвертичной структуры в виде гомогептамеров (за редким исключением), а не гомогексамеров (рис. 4б). Этим свойством они ближе к своим эукариотическим гомологам – белкам Sm/Lsm, которые участвуют в процессинге РНК в клетках и, как правило, входят в состав различных РНП, таких как сплайсома. В отличие от бактериальных и эукариотических гомологов функции SmAP в архейных клетках не установлены. Имеются данные, что они связываются с рядом “незрелых” РНК, а также с белками, участвующими в процессинге РНК в клетках [36–39], что указывает на их возможную вовлеченность в созревание РНК, однако прямых доказательств функциональной роли SmAP белков пока не имеется.

Задачей проводимых исследований стало определение РНК-связывающих и РНК-шаперонных свойств белков SmAP с целью определения возможной функции этих белков в археях. Для этого применили описанную выше технику идентификации РНК-связывающих участков на поверхности SmAP белков, определив ряд структур комплексов белков с рибонуклеотидами. Полученные данные показали, что SmAP белки имеют только один, уридин-специфический РНК-связывающий сайт на проксимальной стороне гомогептамера [40, 41]. Таким образом, сложно предположить, что SmAP белки могут связывать две разные молекулы РНК и участвовать в регуляции трансляции посредством мРНК. Однако, как было показано ранее, SmAP белки могут разрушать структуру целевой РНК без участия второй, комплементарной, РНК, что потенциально позволяет этим белкам взаимодействовать с мРНК, изменять ее пространственную структуру и таким образом влиять на процесс трансляции. Следовательно, SmAP белки нельзя исключать из числа возможных регуляторов трансляции в археях, но для полного выяснения их функции требуются дополнительные исследования.

4. ИССЛЕДОВАНИЕ ФУНКЦИИ РЕГУЛЯТОРОВ ТРАНСКРИПЦИИ UXR И EXUR

Экспрессия генов регулируется не только на уровне трансляции, но и на уровне транскрипции. Одним из примеров этого является регуляция метаболического процесса усвоения D-гексуроновых кислот, таких как D-глюкуронат и D-галактуронат (метаболический путь Эшвелла) [42, 43]. В бактериях этот путь является альтернативой наиболее известному гликолитическому пути метаболизма и играет заметную роль для выживания в стрессовых условиях. Процессы утилизации гексурононов находятся под контролем двух тесно взаимосвязанных регуляторов – UxuR и ExuR [44, 45]. Эти белки являются паралогами, так как ExuR, по-видимому, появился в результате дубликации гена *uxuR* [46], и относятся к семейству регуляторов GntR. К настоящему времени получены пространственные структуры нескольких белков этого семейства [46, 47], однако степень их гомологии с UxuR и ExuR составляет не более 25% [48]. Целью данной работы стало определение структур этих двух белков и их комплексов с регуляторными областями ДНК, в том числе в присутствии лигандов – глюкороната и галакторуната. Это позволило бы на молекулярном уровне описать регуляцию экспрессии генов, вовлеченных в метаболизм гексурононов.

Получение белков в препаративных количествах было осложнено высокой склонностью белков UxuR и ExuR к агрегации и их токсичностью для клеток-продуцентов, поэтому пришлось тщательно подбирать условия их получения [49]. Эксперименты по определению сродства белков к целевым участкам ДНК подтвердили наличие двух различающихся участков связывания UxuR и ExuR, причем присутствие сахаров меняло сродство белков к ДНК. Нестабильность белков в условиях кристаллизации не позволила получить кристаллы белков в свободном состоянии и в комплексах с лигандами и/или фрагментами ДНК, поэтому структуры таких комплексов промоделированы с помощью методов гибкого докинга на основе структур гомологичных белков [50]. Это позволило локализовать предположительный участок связывания сахаров в области C-концевых доменов UxuR, совпадающий с местом, предсказанным ранее по данным мутагенеза.

Применение современных методов молекулярного моделирования и молекулярной динамики позволяет получить интересные данные по взаимодействию молекул, однако только определение структур комплексов может подтвердить достоверность расчетных моделей.

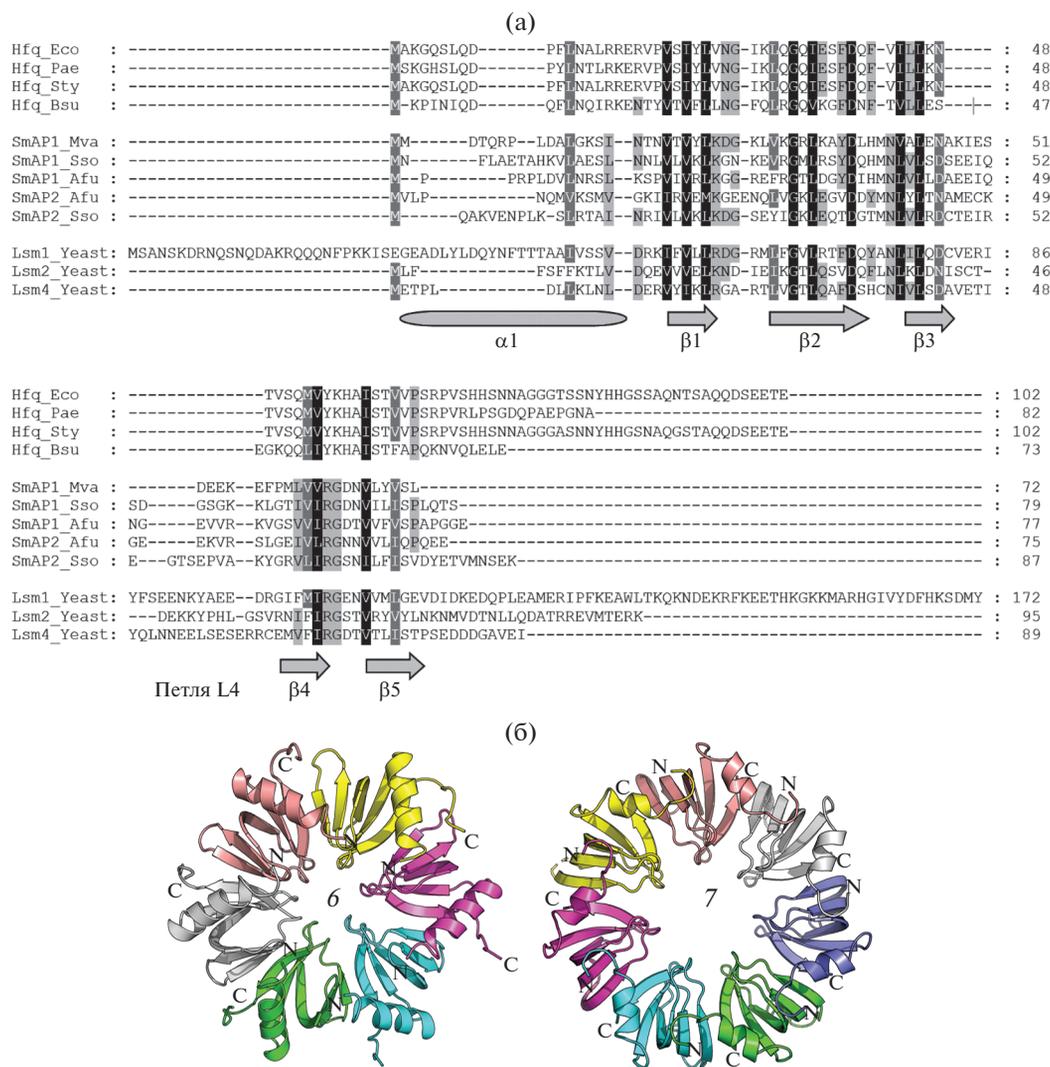


Рис. 4. Выравнивание аминокислотных последовательностей представителей белков Hfq (Hfq_Eco – *E. coli*, Hfq_Pae – *P. aeruginosa*, Hfq_Sty – *Salmonella typhimurium*, Hfq_Bsu – *Bacillus subtilis*), SmAP (SmAP1_Mva – *Methanococcus vannielii*, SmAP1_Sso – *Sulfolobus solfataricus*, SmAP1-Afu – *Archeoglobus fulgidus*) и эукариотических Lsm из *Saccharomyces cerevisiae*. Внизу показаны элементы вторичной структуры, соответствующие аминокислотным последовательностям белков. Выравнивание сделано в программе T-Coffee с учетом информации по 3D-структурам белков (Espresso) [69] (a). Схематическое представление структур гексамерного бактериального белка Hfq из *P. aeruginosa* (слева, PDB ID 1U1S) и гептамерного архейного белка SmAP из *M. vannielii* (справа, PDB ID 5MK1). Вид со стороны α -спиралей мономеров белков (проксимальной стороны) (б).

5. СТРУКТУРНЫЕ И ФУНКЦИОНАЛЬНЫЕ ИССЛЕДОВАНИЯ ГЛИЦИЛ-ТРНК-СИНТЕТАЗЫ ЧЕЛОВЕКА

Энтеровирусы, а также некоторые родственные им вирусы инициируют синтез своих белков в зараженной эукариотической клетке посредством кэп-независимого механизма. В этом случае специфическая вторичная структура определенного участка вирусной мРНК (IRES) делает возможной сборку трансляционного аппарата непосредственно в месте, близком к старт-кодону. В отличие от кэп-зависимой инициации трансляции механизмы IRES-зависимой инициации весьма разнообразны. IRES из различных

источников могут быть абсолютно не похожи друг на друга: не найдено элемента структуры, общего для всех IRES, или какой-либо заметной гомологии, общей для всех их последовательностей. Такое разнообразие вызвало определенные трудности классификации IRES с момента их открытия. Однако интенсивные исследования вирусов, их мРНК и механизмов ее трансляции позволили к настоящему моменту выделить ряд общих черт, позволяющих вполне определенно отличать одни типы IRES от других.

Большинство вирусных IRES обладают стабильной вторичной и третичной структурами, способствующими их эффективному связыва-

нию с 40S субчастицей [51]. Инициация трансляции на всех исследованных на сегодняшний день IRES типа I зависит от наличия ITAF (англ. *IRES trans-acting factors*, представляют собой специальные факторы трансляции), список которых до сих пор полностью не определен. Одним из ITAFs, необходимых для эффективной инициации трансляции полиовирусной мРНК, является глицил-тРНК-синтетаза человека (**GlyRS**). Этот белок взаимодействует с апикальной частью домена V IRES полиовируса, которая мимикрирует под антикодонную шпильку глициновой тРНК и содержит глициновый антикодон. GlyRS имеет трехдоменную структуру, при этом за узнавание глицинового антикодона отвечает отдельный домен, называемый **ABD** (Anticodon Binding Domain). К настоящему моменту определены структуры GlyRS как в свободной форме, так и в комплексе с глициновой тРНК, что позволило детально изучить механизм узнавания белком антикодонной петли тРНК [52, 53], однако структура комплекса GlyRS с вирусной мРНК пока не определена.

Ранее были получены фрагменты домена V IRES первого типа (IRES I) и определен минимальный фрагмент, взаимодействующий с GlyRS [54]. Также было показано, что глицил-тРНК-синтетаза человека специфически связывает фрагменты IRES I вирусов, таксономически удаленных друг от друга, с близкими константами. Это указывает на универсальность механизма регуляции инициации трансляции при помощи глицил-тРНК-синтетазы для таких вирусов и на схожесть пространственных структур соответствующих участков их мРНК [55].

Сродство глицил-тРНК-синтетазы человека к фрагментам IRES I примерно в 3 раза превышает сродство к родственной тРНК [55], что дает основания предполагать схожий механизм взаимодействия белка и РНК в обоих случаях. В отличие от тРНК «антикодонная петля» в составе апикальной части пятого домена IRES содержит не пять, а шесть нуклеотидов. Пространственная структура V домена IRES I до сих пор неизвестна. Однако удалось построить теоретическую модель этого фрагмента и его комплекса с ABD доменом глицил-тРНК-синтетазы человека (рис. 5), основываясь на имеющейся структурной информации [56]. Модель имеет повышенное (по сравнению с гомологичной частью tRNA^{Gly}) число водородных связей с белком, что позволяет объяснить полученные нами биохимические данные.

6. СТРУКТУРНЫЕ ИССЛЕДОВАНИЯ ГЕТЕРОТРИМЕРНОГО ФАКТОРА ИНИЦИАЦИИ ТРАНСЛЯЦИИ 2

У бактерий фактор инициации трансляции 2 (**IF2**) является шестидоменным мономерным

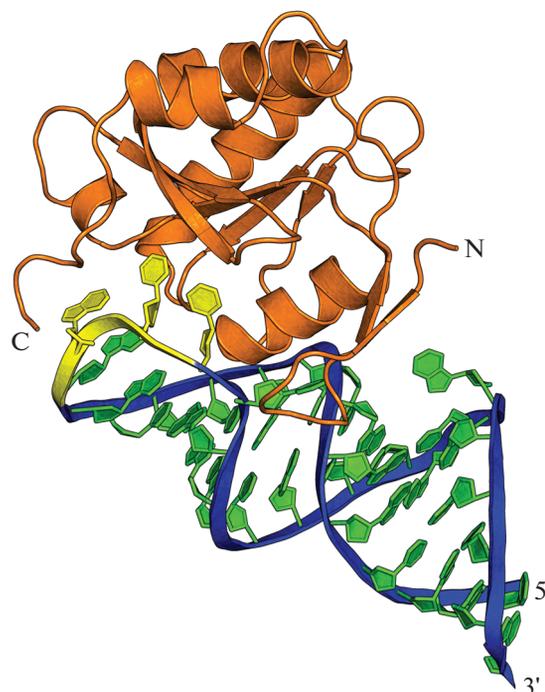


Рис. 5. Расчетная модель комплекса антикодон-связывающего домена глицил-тРНК-синтетазы человека с фрагментом IRES полиовируса.

белком, который способствует связыванию инициаторной аминоксил-тРНК с малой рибосомной субчастицей. У архей и эукариот фактор инициации трансляции 2 (**e/aIF2**) состоит из трех различных субъединиц: α , β , γ . Функционально он аналогичен бактериальному IF2, однако ни одна из субъединиц e/aIF2 не имеет гомологии с бактериальным IF2. Субъединицы эукариотического eIF2 по размеру больше, чем гомологичные им субъединицы архейного aIF2. В присутствии ГТФ гетеротримерный фактор e/aIF2 специфически узнает инициаторную Met-тРНК_i и образует с ней тройной комплекс e/aIF2–GTP–Met-тРНК_i. В такой форме e/aIF2 доставляет Met-тРНК_i в Р-участок малой рибосомной субчастицы, где за счет кодон-антикодонного взаимодействия происходит точное узнавание иницирующего кодона на мРНК с последующим гидролизом ГТФ. В комплексе с ГДФ фактор инициации трансляции IF2 покидает рибосому, после чего происходит обмен связанного ГДФ на ГТФ, и фактор снова способен образовывать прочный комплекс с Met-тРНК_i. Таким образом, фактор инициации трансляции IF2 выполняет очень важную функцию в клетке: запускает синтез белка на рибосоме.

Впервые структура гетеротримерного фактора инициации трансляции IF2 была определена в лаборатории ИБ РАН [57] и до настоящего времени она остается единственной кристаллической

структурой полноразмерного фактора инициации трансляции eIF2 эукариотического типа. Структура aIF2 была определена из термофильной археи *Sulfolobus solfataricus* и в дальнейшем использовалась для построения модели β -субъединицы эукариотического фактора eIF2 в составе пространственной структуры дрожжевого 48S преинициаторного комплекса (48S PIC), которая исследована методом криоэлектронной микроскопии [58]. Общими очертаниями гетеротримерный фактор инициации трансляции 2 напоминает английскую букву L и состоит из конформационно-стабильной центральной части и двух подвижных “крыльев” [57]. Минимальной и достаточной для связывания инициаторной Met-тРНК является конформационно-стабильная центральная часть aIF2 – γ -субъединица с третьим доменом α -субъединицы aIF2 α_{D3} γ .

Практически одновременно были расшифрованы две кристаллические структуры тройного комплекса aIF2–ГТФ–Met-тРНКi с использованием архейного фактора инициации трансляции IF2. В 2012 г. французская группа ученых получила кристаллы тройного комплекса, содержащего полноразмерный гетеротримерный фактор aIF2 $\alpha\beta\gamma$, однако построить и уточнить модель для подвижной β -субъединицы не удалось в силу недостаточно ясной электронной плотности [59]. В 2013 г. в лаборатории ИБ РАН были получены кристаллы тройного комплекса, содержащего коровую конформационно-стабильную часть фактора aIF2 α_{D3} γ [60]. Обе структуры тройного комплекса дали понимание того, как фактор aIF2 отличает инициаторную тРНК от элонгаторной тРНК. Инициаторная тРНК обращена к белку aIF2 своей D-стороной, тогда как фактор элонгации EF-Tu, структурно схожий с aIF2 γ , взаимодействует с элонгаторной тРНК по T-шпильке [61]. Таким образом, различия в нуклеотидной последовательности T-шпильки инициаторной и элонгаторных тРНК могут предотвращать связывание инициаторной тРНК с EF-Tu, а различия в D-шпильке могут предотвращать связывание элонгаторных тРНК с aIF2.

При сравнении двух структур тройного комплекса aIF2–ГТФ–Met-тРНКi было обнаружено, что несмотря на то, что в обеих структурах тРНК обращена к белку своей D-стороной, взаимодействует она с aIF2 несколько по-разному (рис. 6). В структуре тройного комплекса, полученной нашей группой, акцепторный стебель тРНК образует контакты со всеми тремя доменами γ -субъединицы и aIF2 α_{D3} : ССА-конец тРНК локализован в расщелине между G(I)- и III-доменами γ -субъединицы, II-домен γ -субъединицы aIF2 взаимодействует с 5'-концом Met-тРНКi, а белок aIF2 α_{D3} контактирует с локтем тРНК [60]. В модели aIF2–ГТФ–Met-тРНКi, полученной группой француз-

ских коллег [59], акцепторный стержень тРНК не взаимодействует с III-доменом γ -субъединицы, а локоть тРНК контактирует с aIF2 α_{D1} и aIF2 α_{D2} , которые согласно биохимическим данным проявляют неспецифическую РНК-связывающую активность [62, 63].

В сотрудничестве с лабораторией И. Шатского (НИИ ФХБ МГУ) в экспериментах *in vitro* показано, что архейный фактор aIF2 способен поддерживать рибосомное сканирование [64]. С помощью метода тоупринтинга выявили, что если добавить архейный фактор aIF2 в систему инициации трансляции млекопитающих, в которой отсутствует его эукариотический гомолог, то 48S инициаторный комплекс образуется, и идет сканирование мРНК в поисках иницирующего AUG-кодона. Эта совместная работа раскрыла особенность второй стадии инициации трансляции – присоединение большой субчастицы к 48S инициаторному комплексу. На этом этапе в работу включаются дополнительные факторы: eIF5 и eIF5B (промотируют присоединение большой субчастицы). У архей aIF5 отсутствует, но мы добавляли eIF5 в систему трансляции с архейным вторым фактором, и оказалось, что если заменить eIF2 на aIF2, то 80S рибосомный комплекс не образуется. Этот результат демонстрирует, что в археях способ стимуляции ГТФазной активности aIF2 отличается от того, который реализуется у эукариот.

Дальнейшие наши исследования в области архейного фактора aIF2 привели к тому, что был определен ряд пространственных структур γ -субъединицы aIF2, представляющих разные функциональные состояния этой рибосомной ГТФазы: от активного (ON) ГТФ-связанного до неактивного (OFF) ГДФ-связанного состояния, проходя через промежуточные формы, когда белок может быть связан одновременно с ГДФ и неорганическим фосфатом P_i или не содержать нуклеотида вовсе [65]. Сравнение структур впервые продемонстрировало смещение доменов II и III относительно домена G(I), хотя и не столь значительное как в его структурном гомологе, факторе элонгации трансляции EF-Tu. Скоординированные перестройки происходят и в петлях-переключателях Switch 1 и 2. Петля Switch 1 играет роль “застежки” для нуклеотид-связывающего кармана. Обнаружено, что конформация этого переключателя не позволяет полностью закрыть карман в aIF2 γ в отличие от EF-Tu. Он остается открытым на протяжении всего цикла работы aIF2 γ , и γ -фосфату ничто не препятствует покинуть белок после гидролиза ГТФ. Это в свою очередь указывает на возможность спонтанного гидролиза ГТФ у фактора aIF2 на рибосоме и объясняет отсутствие у архей фактора aIF5, стимулирующего ГТФазную активность aIF2.

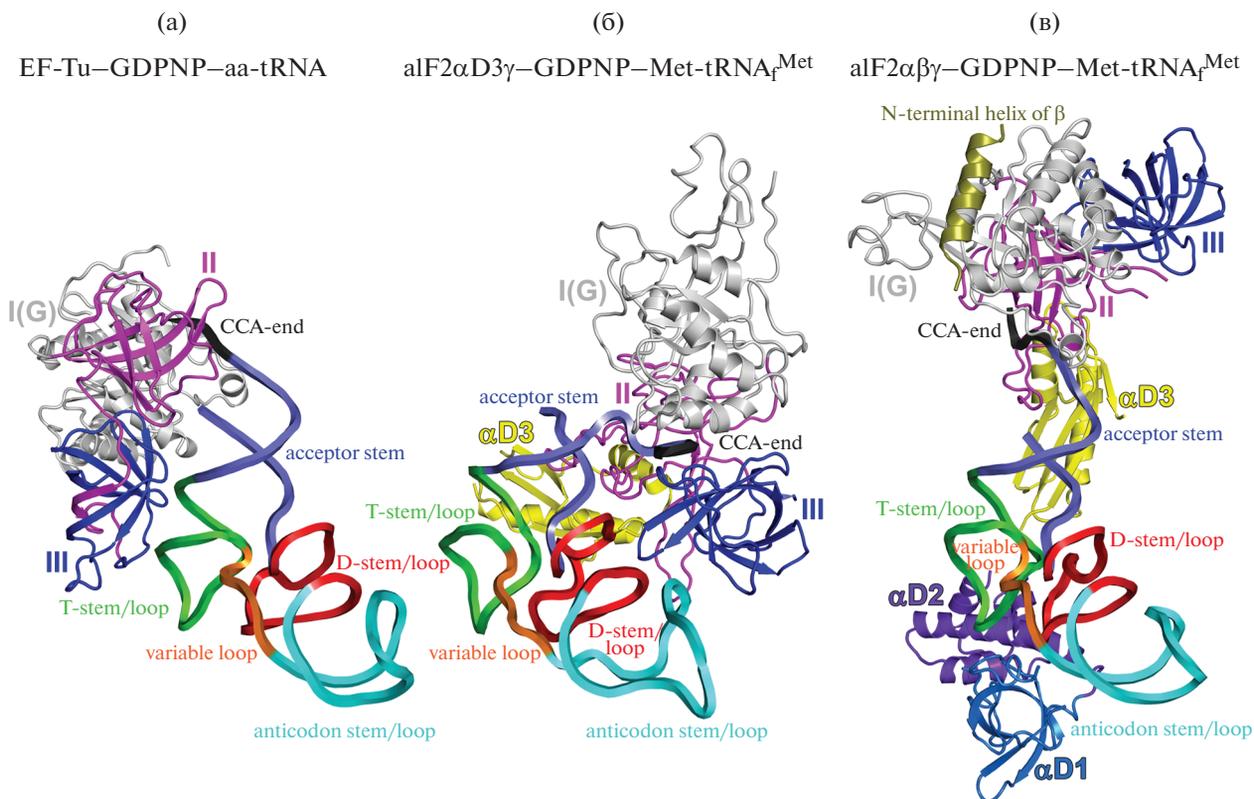


Рис. 6. Сравнение структур комплексов тРНК с EF-Tu (PDB ID 1TTT, слева), aIF2 α D3 γ (PDB ID 3QSY, в центре) и aIF2 α β γ (PDB ID 3V11, справа).

При анализе структуры aIF2 γ , нуклеотид-связывающий карман которой находится в состоянии интермедиата при переходе из ГТФ- в ГДФ-связанную форму, были обнаружены три кластера молекул воды, образующие водные мостики между γ -фосфатом ГТФ и двумя консервативными остатками Asp19 и Asp93 [66]. Эти водные мостики могут быть использованы для переноса протона, образующегося при каталитической реакции – гидролизе ГТФ.

Кроме этого, выявлена дополнительная петля (третья) у γ -субъединицы aIF2, которая может работать как переключатель во время гидролиза ГТФ [67]. В результате изменения конформации этой петли растворитель (молекулы воды) получает доступ к γ -фосфату ГТФ. Главным привратником на пути к γ -фосфату ГТФ является Met45. Интересно, что у эукариотической γ -субъединицы eIF2 вместо Met45 расположена более крупная аминокислота – изолейцин, и для открытия “ворот” требуется приложить усилия, чтобы не встретить стерические препятствия, что приводит к необходимости использовать белковый фактор eIF5, индуцирующий гидролиз ГТФ. А у фактора элонгации трансляции EF-Tu и вовсе отсутствует эта петля. В совокупности эти и изложенные выше данные лишней раз указывают на вероятность

отсутствия ГТФаза-активирующего фактора у архей и спонтанного гидролиза ГТФ у фактора aIF2.

ЗАКЛЮЧЕНИЕ

Накопленный в Институте белка РАН опыт структурно-функциональных исследований РНК-белковых комплексов в рамках “рибосомного проекта” успешно использован при выполнении ряда исследований принципов регуляции трансляции и транскрипции функционально важными объектами. Среди них исследование авторегуляции трансляции универсально-консервативного рибосомного белка L1 в бактериях и археях, регулятора трансляции Hfq и его архейных гомологов, регуляторов транскрипции UxuR и ExuR, исследование дополнительной функции глицил-тРНК-синтетазы человека, изучение гетеротримерного фактора инициации трансляции IF2.

В обзоре кратко представлены результаты проведенных исследований. На молекулярном уровне показана природа различного родства рибосомного белка L1 к специфически узнаваемым участкам мРНК и рРНК, что позволяет регулировать трансляцию своего оперона по принципу обратной связи. Для белка Hfq выявлены структурные основы узнавания молекул РНК в функцио-

нально важном боковом РНК-связывающем участке белка. Исследованы РНК-связывающие свойства ряда архейных белков семейства Lsm и показаны существенные отличия их РНК-связывающих свойств по сравнению с бактериальным гомологом — белком Hfq. Получены структурные данные о взаимодействии глицил-тРНК-синтазы с IRES вирусов. Проведены обширные комплексные исследования гетеротримерного фактора инициации трансляции IF2 архей и эукариот.

Особую благодарность авторы выражают родоначальникам структурно-функциональных исследований в Институте белка РАН — Марии Борисовне Гарбер, Станиславу Владимировичу Нионову и Наталии Александровне Невской.

Работа выполнена при поддержке следующих фондов: проект “Регуляторные комплексы рибосомного белка L1 с мРНК” — Российский фонд фундаментальных исследований (РФФИ) (гранты № 03-04-48327, 04-04-49634, 05-04-48338, INTAS YSF № 04-83-3842); проект “Исследования РНК-связывающих свойств белка Hfq” — РФФИ (гранты № 04-04-48556, 07-04-00296, 13-04-00783); проект “Исследования РНК-связывающих свойств архейных Lsm белков” — Российский научный фонд (РНФ) (грант № 14-14-00496), РФФИ (грант № 18-04-00222); проект “Исследование функции регуляторов транскрипции UxuR и ExuR” — РНФ (грант № 18-14-00322); проект “Структурные и функциональные исследования глицил-тРНК-синтазы человека” — РНФ (грант № 15-14-00028), РФФИ (грант № 19-34-90135); проект “Структурные исследования гетеротримерного фактора инициации трансляции IF2” — РФФИ (грант № 18-04-01331).

СПИСОК ЛИТЕРАТУРЫ

- Gourse R.L., de Boer H.A., Nomura M. // Cell. 1986. V. 44. № 1. P. 197.
[https://doi.org/10.1016/0092-8674\(86\)90498-8](https://doi.org/10.1016/0092-8674(86)90498-8)
- Михайлина А.О., Костарева О.С., Нионова Е.Ю. и др. // Молекулярная биология. 2018. Т. 52. № 1. С. 98.
<https://doi.org/10.7868/S0026898418010135>
- Mayer C., Köhrer C., Gröbner P., Piendl W. // Mol. Microbiol. 1998. V. 27. № 2. P. 455.
<https://doi.org/10.1046/j.1365-2958.1998.00693.x>
- Hanner M., Mayer C., Köhrer C. et al. // J. Bacteriol. 1994. V. 176. № 2. P. 409.
<https://doi.org/10.1128/jb.176.2.409-418.1994>
- Tishchenko S.V., Vassilieva J.M., Platonova O.B. et al. // Biochem. 2001. V. 66. № 9. P. 948.
<https://doi.org/10.1023/a:1012353122174>
- Ennifar E., Nikulin A., Tishchenko S. et al. // J. Mol. Biol. 2000. V. 304. № 1. P. 35.
<https://doi.org/10.1006/jmbi.2000.4204>
- Nevskaya N., Tishchenko S., Gabdulkhakov A. et al. // Nucl. Acids Res. 2005. V. 33. № 2. P. 478.
<https://doi.org/10.1093/nar/gki194>
- Nevskaya N., Tishchenko S., Volchkov S. et al. // J. Mol. Biol. 2006. V. 355. № 4. P. 747.
<https://doi.org/10.1016/j.jmb.2005.10.084>
- Tishchenko S., Nikonova E., Nikulin A. et al. // Acta Cryst. D. 2006. V. 62. (Pt. 12). P. 1545.
<https://doi.org/10.1107/S0907444906041655>
- Nikulin A., Eliseikina I., Tishchenko S. et al. // Nat. Struct. Biol. 2003. V. 10. № 2. P. 104.
<https://doi.org/10.1038/nsb886>
- Tishchenko S., Gabdulkhakov A., Nevskaya N. et al. // Acta Cryst. D. 2012. V. 68. (Pt. 8). P. 1051.
<https://doi.org/10.1107/S0907444912020136>
- Kostareva O., Tishchenko S., Nikonova E. et al. // J. Mol. Recognit. 2011. V. 24. № 4. P. 524.
<https://doi.org/10.1002/jmr.1063>
- Tishchenko S., Kostareva O., Gabdulkhakov A. et al. // Acta Cryst. D. 2015. V. 71. (Pt. 2). P. 376.
<https://doi.org/10.1107/S1399004714026248>
- Костарева О.С., Невская Н.А., Тищенко С.В. и др. // Молекулярная биология. 2018. Т. 52. № 1. С. 106.
<https://doi.org/10.7868/S0026898418010147>
- Tishchenko S., Nikonova E., Kljashorny V. et al. // Nucl. Acids Res. 2007. V. 35. № 21. P. 7389.
<https://doi.org/10.1093/nar/gkm898>
- Tishchenko S., Kljashorny V., Kostareva O. et al. // J. Mol. Biol. 2008. V. 383. № 2. P. 301.
<https://doi.org/10.1016/j.jmb.2008.08.058>
- Korepanov A.P., Kostareva O.S., Bazhenova M.V. et al. // Protein J. 2015. V. 34. № 2. P. 103.
<https://doi.org/10.1007/s10930-015-9602-5>
- Updegrove T.B., Zhang A., Storz G. // Curr. Opin. Microbiol. 2016. V. 30. P. 133.
<https://doi.org/10.1016/j.mib.2016.02.003>
- de Fernandes M.T.F., Hayward W.S., August T.J. // J. Biol. Chem. 1972. V. 247. № 3. P. 824.
- Schuppli D., Miranda G., Tsui H.C. et al. // Proc. Natl. Acad. Sci. U. S. A. 1997. V. 94. № 19. P. 10239.
<https://doi.org/10.1073/pnas.94.19.10239>
- Vecerek B., Moll I., Afonyushkin T. et al. // Mol. Microbiol. 2003. V. 50. № 3. P. 897.
<https://doi.org/10.1046/j.1365-2958.2003.03727.x>
- Muffler A., Traulsen D.D., Fischer D. et al. // J. Bacteriol. 1997. V. 179. № 1. P. 297.
<https://doi.org/10.1128/JB.179.1.297-300.1997>
- Jørgensen M.G., Pettersen J.S., Kallipolitis B.H. // Biochim. Biophys. Acta — Gene Regul. Mech. 2020. V. 1863. № 5. P. 194504.
<https://doi.org/10.1016/j.bbgrm.2020.194504>
- Schumacher M.A., Pearson R.F., Möller T. et al. // EMBO J. 2002. V. 21. № 13. P. 3546.
<https://doi.org/10.1093/emboj/cdf322>
- Sauter C., Basquin J., Suck D. // Nucl. Acids Res. 2003. V. 31. № 14. P. 4091.
<https://doi.org/10.1093/nar/gkg480>
- Nikulin A., Stolboushkina E., Perederina A. et al. // Acta Cryst. D. 2005. V. 61. (Pt. 2). P. 141.
<https://doi.org/10.1107/S0907444904030008>
- Arлуison V., Mutyam S.K., Mura C. et al. // Protein Sci. 2007. V. 16. № 9. P. 1830.
<https://doi.org/10.1110/ps.072883707>
- Murina V., Lekontseva N., Nikulin A. // Acta Cryst. D. 2013. V. 69. (Pt. 8). P. 1504.
<https://doi.org/10.1107/S090744491301010X>

29. *Nemchinova M., Balobanov V., Nikonova E. et al.* // *Protein J.* 2017. V. 36. № 3. P. 157.
<https://doi.org/10.1007/s10930-017-9709-y>
30. *Balobanov V., Lekontseva N., Mikhaylina A., Nikulin A.* // *Methods Mol. Biol.* 2020. V. 2113. P. 251.
https://doi.org/10.1007/978-1-0716-0278-2_17
31. *Chen J., Morita T., Gottesman S.* // *Front. Cell. Infect. Microbiol.* 2019. V. 9. № 6. P. 201.
<https://doi.org/10.3389/fcimb.2019.00201>
32. *Sauer E., Schmidt S., Weichenrieder O.* // *Proc. Natl. Acad. Sci.* 2012. V. 109. № 24. P. 9396.
<https://doi.org/10.1073/pnas.1202521109>
33. *Robinson K.E., Orans J., Kovach A.R. et al.* // *Nucl. Acids Res.* 2013. V. 42. № 4. P. 2376.
<https://doi.org/10.1093/nar/gkt1171>
34. *Panja S., Schu D.J., Woodson S.* // *Nucl. Acids Res.* 2013. V. 41. № 15. P. 7536.
<https://doi.org/10.1093/nar/gkt521>
35. *Мурина В.Н., Никулин А.Д.* // *Успехи биологической химии.* 2011. Т. 51. С. 133.
36. *Fischer S., Benz J., Spath B. et al.* // *J. Biol. Chem.* 2010. V. 285. № 45. P. 34429.
<https://doi.org/10.1074/jbc.M110.118950>
37. *Maier L.K., Benz J., Fischer S. et al.* // *Biochimie.* 2015. V. 117. P. 129.
<https://doi.org/10.1016/j.biochi.2015.02.023>
38. *Märtenz B., Bezerra G., Kreuter M. et al.* // *Life.* 2015. V. 5. № 2. P. 1264.
<https://doi.org/10.3390/life5021264>
39. *Märtenz B., Hou L., Amman F. et al.* // *Nucl. Acids Res.* 2017. V. 45. № 13. P. 7938.
<https://doi.org/10.1093/nar/gkx437>
40. *Nikulin A., Mikhailina A., Lekontseva N. et al.* // *J. Biomol. Struct. Dyn.* 2017. V. 35. № 8. P. 1615.
<https://doi.org/10.1080/07391102.2016.1189849>
41. *Lekontseva N., Mikhailina A., Fando M. et al.* // *Biochimie.* 2020. V. 175. P. 1.
<https://doi.org/10.1016/j.biochi.2020.05.001>
42. *Ashwell G.* // *Methods Enzymol.* 1962. V. 5. P. 190.
[https://doi.org/10.1016/S0076-6879\(62\)05205-2](https://doi.org/10.1016/S0076-6879(62)05205-2)
43. *Peekhaus N., Conway T.* // *J. Bacteriol.* 1998. V. 180. № 14. P. 3495.
<https://doi.org/10.1128/JB.180.14.3495-3502.1998>
44. *Ritzenthaler P., Blanco C., Mata-Gilsinger M.* // *Mol. Gen. Genet.* 1985. V. 199. № 3. P. 507.
<https://doi.org/10.1007/BF00330766>
45. *Utz C.B., Nguyen A.B., Smalley D.J. et al.* // *J. Bacteriol.* 2004. V. 186. № 22. P. 7690.
<https://doi.org/10.1128/JB.186.22.7690-7696.2004>
46. *Suvorova I.A., Tutukina M.N., Ravcheev D.A. et al.* // *J. Bacteriol.* 2011. V. 193. № 15. P. 3956.
<https://doi.org/10.1128/JB.00277-11>
47. *Fillenberg S.B., Friess M.D., Korner S. et al.* // *PLoS One* 2016. V. 11. № 6. P. e0157691.
<https://doi.org/10.1371/journal.pone.0157691>
48. *Tutukina M.N., Potapova A.V., Vlasov P.K. et al.* // *J. Biomol. Struct. Dyn.* 2016. V. 34. № 10. P. 2296.
<https://doi.org/10.1080/07391102.2015.1115779>
49. *Bessonova T.A., Lekontseva N.V., Shvyreva U.S. et al.* // *Protein Expr. Purif.* 2019. V. 161. P. 70.
<https://doi.org/10.1016/j.pep.2019.05.001>
50. *Пуртов Ю.А., Тутукина М.Н., Никулин А.Д., Озо-льин О.Н.* // *Биофизика.* 2019. Т. 64. № 1. С. 061.
<https://doi.org/10.1134/S0006350919010160>
51. *Nikonov O.S., Chernykh E.S., Garber M.B., Nikonova E.Y.* // *Biochemistry (Moscow).* 2017. V. 82. № 13. P. 1615.
<https://doi.org/10.1134/S0006297917130041>
52. *Qin X., Hao Z., Tian Q. et al.* // *J. Biol. Chem.* 2014. V. 289. № 29. P. 20359.
<https://doi.org/10.1074/jbc.M114.557249>
53. *Qin X., Deng X., Chen L., Xie W.* // *J. Mol. Biol.* 2016. V. 428. № 18. P. 3603.
<https://doi.org/10.1016/j.jmb.2016.05.018>
54. *Никонова Е.Ю., Михайлина А.О., Леконцева Н.В. и др.* // *Биофизика.* 2016. Т. 61. № 2. С. 277.
<https://doi.org/10.1134/S0006350916020135>
55. *Никонова Е.Ю., Михайлина А.О., Немчинова М.С. и др.* // *Молекулярная биология.* 2018. Т. 52. № 1. С. 10.
<https://doi.org/10.7868/S0026898418010020>
56. *Никонов О.С., Немчинова М.С., Кляшторный В.Г. и др.* // *Молекулярная биология.* 2018. Т. 52. № 1. С. 112.
<https://doi.org/10.7868/S0026898418010159>
57. *Stolboushkina E., Nikonov S., Nikulin A. et al.* // *J. Mol. Biol.* 2008. V. 382. № 3. P. 680.
<https://doi.org/10.1016/j.jmb.2008.07.039>
58. *Llácer J.L., Hussain T., Marler L. et al.* // *Mol. Cell.* 2015. V. 59. № 3. P. 399.
<https://doi.org/10.1016/j.molcel.2015.06.033>
59. *Schmitt E., Panvert M., Lazennec-Schurdevin C. et al.* // *Nat. Struct. Mol. Biol.* 2012. V. 19. № 4. P. 450.
<https://doi.org/10.1038/nsmb.2259>
60. *Stolboushkina E., Nikonov S., Zelinskaya N. et al.* // *J. Mol. Biol.* 2013. V. 425. № 6. P. 989.
<https://doi.org/10.1016/j.jmb.2012.12.023>
61. *Nissen P., Kjeldgaard M., Thirup S.r. et al.* // *Science.* 1995. V. 270. № 5241. P. 1464.
<https://doi.org/10.1126/science.270.5241.1464>
62. *Yatime L., Schmitt E., Blanquet S., Mechulam Y.* // *J. Biol. Chem.* 2004. V. 279. № 16. P. 15984.
<https://doi.org/10.1074/jbc.M311561200>
63. *Pedulla N.* // *Nucl. Acids Res.* 2005. V. 33. № 6. P. 1804.
<https://doi.org/10.1093/nar/gki321>
64. *Dmitriev S.E., Stolboushkina E.A., Terenin I.M. et al.* // *J. Mol. Biol.* 2011. V. 413. № 1. P. 106.
<https://doi.org/10.1016/j.jmb.2011.08.026>
65. *Nikonov O., Stolboushkina E., Arkhipova V. et al.* // *Acta Cryst. D.* 2014. V. 70. (Pt. 3). P. 658.
<https://doi.org/10.1107/S1399004713032240>
66. *Nikonov O., Kravchenko O., Arkhipova V. et al.* // *Biochimie.* 2016. V. 121. P. 197.
<https://doi.org/10.1016/j.biochi.2015.11.029>
67. *Nikonov O., Kravchenko O., Nevskaya N. et al.* // *Acta Cryst. D.* 2019. V. 75. (Pt. 4). P. 392.
<https://doi.org/10.1107/S20597983190002304>
68. *Link T.M., Valentin-Hansen P., Brennan R.G.* // *Proc. Natl. Acad. Sci. USA.* 2009. V. 106. № 46. P. 19292.
<https://doi.org/10.1073/pnas.0908744106>
69. *Notredame C., Higgins D.G., Heringa J.* // *J. Mol. Biol.* 2000. V. 302. № 1. P. 205.
<https://doi.org/10.1006/jmbi.2000.4042>

УДК 577.352.332

СТРУКТУРНЫЕ ИССЛЕДОВАНИЯ МЕМБРАННОГО ГЛИКОПРОТЕИНА-ПРЕДШЕСТВЕННИКА БЕТА-АМИЛОИДА И БЕЛКОВ, УЧАСТВУЮЩИХ В ПРОЦЕССЕ ЕГО ПРОТЕОЛИЗА

© 2021 г. А. С. Урбан^{1,2,3,*}, Я. В. Бершацкий^{1,2}, К. В. Павлов³, Э. В. Бочаров^{1,2}

¹ Институт биоорганической химии им. М.М. Шемякина и Ю.А. Овчинникова РАН, Москва, Россия

² Московский физико-технический институт, Долгопрудный, Россия

³ Федеральный научно-клинический центр физико-химической медицины, Москва, Россия

*E-mail: anatoly.urban@gmail.com

Поступила в редакцию 07.10.2020 г.

После доработки 14.10.2020 г.

Принята к публикации 23.11.2020 г.

Болезнь Альцгеймера – самый распространенный вид нейродегенеративного заболевания в мире. Ее клинические проявления объясняются избирательной дегенерацией нейронов в участках коры головного мозга, ответственных за когнитивное восприятие и память. Вне нервных клеток в местах контакта нейронов накапливается амилоидный пептид, который собирается в упорядоченные тяжи, фибриллы, образующие так называемые амилоидные бляшки. Установлено, что амилоидный пептид (бета-амилоид или $A\beta_{1-38}$ – $A\beta_{1-43}$), агрегирующий и формирующий амилоидные бляшки в мозге, является продуктом последовательного расщепления мембранного гликопротеина-предшественника β -амилоида β - и γ -секретазами в плазматической мембране нейронов. В обзоре освещены результаты структурных исследований данного процесса и основных белков, участвующих в нем.

DOI: 10.31857/S0023476121050222

ОГЛАВЛЕНИЕ

Введение
1. Многодоменная структура белка-предшественника бета-амилоида
1. Многодоменная структура белка-предшественника бета-амилоида
1.1. GFLD/HBD
1.2. CuBD
1.3. E1
1.4. KPI
1.5. E2
1.6. Полноразмерный эктодомен
1.7. AICD
1.8. TM домен
1.9. A β
1.10. Олигомеры
2. А-секретаза
3. В-секретаза
4. Г-секретаза
Заключение

ВВЕДЕНИЕ

Стремительное увеличение доли пожилых людей, нуждающихся в уходе и медицинской поддержке, является серьезной социальной и экономической проблемой в развитых странах. Процесс старения включает в себя различные нейродегенеративные заболевания, нарушающие сложную сеть метаболических и сигнальных путей. Болезнь Альцгеймера – самый распространенный в мире вид нейродегенеративного заболевания, которым ежегодно заболевает несколько миллионов человек [1]. С увеличением продолжительности жизни вероятность развития этого вида деменции возрастает многократно. В патогенезе болезни Альцгеймера задействованы различные пептиды и белки, в том числе амилоидогенные, и прежде всего бета-амилоидный пептид (A β) и его белок-предшественник, а также мембранные секретазы, расщепляющие его до различных A β -изоформ. Настоящий обзор посвящен структурным исследованиям данного процесса современными методами структурной биологии, такими как рентгеноструктурный анализ (РСА), малоугловое рассеяние рентгеновских лучей (МУРР), ядерный магнитный резонанс (ЯМР) и криоэлектронная микроскопия.

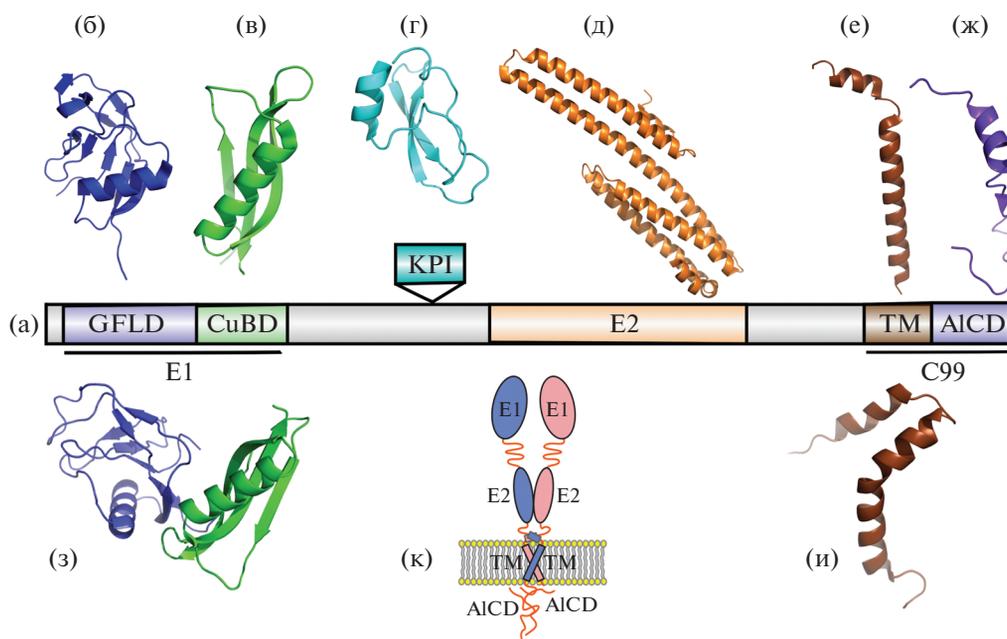


Рис. 1. Схема строения APP человека (а) и структуры его отдельных доменов (б–з): б – GFLD – домен, подобный фактору роста (PDB:4PQD); в – CuBD – медьсвязывающий домен (PDB:2FJZ); г – KPI – домен, гомологичный ингибиторам протеаз типа Кунитц (PDB:1AAP); д – домен E2 (PDB:3NYL); е – трансмембранный домен (PDB:2LLM); ж – AICD – внутриклеточный домен (PDB:2ROZ); з – полноразмерный домен E1, различными цветами показаны входящие в него поддомены GFLD и CuBD (PDB:4PWQ); и – фрагмент C99, состоящий из трансмембранного и внутриклеточного доменов, показана структура трансмембранного домена, так как внутриклеточный домен находится в неупорядоченном состоянии (PDB:2LPI); к – схематическое изображение общей доменной организации APP.

1. МНОГОДОМЕННАЯ СТРУКТУРА БЕЛКА-ПРЕДШЕСТВЕННИКА БЕТА-АМИЛОИДА

Гликопротеин-предшественник β -амилоида (APP – amyloid precursor protein) экспрессируется почти во всех животных клетках, исследованных на сегодня, причем с выраженным эволюционным консерватизмом. Хотя изучению белка APP посвящены тысячи работ, его клеточные функции до сих пор полностью неясны. Был сделан ряд предположений, что белок APP является рецептором холестерина в рафтах мембран нейронов [2], регулирует метаболизм железа в нейронах [3], а также является частью системы врожденного иммунитета в нервной системе человека [4]. Это типичный многодоменный мембранный белок первого типа, пронизывающий мембрану один раз, т.е. он имеет внеклеточную, внутриклеточную и трансмембранную части (рис. 1к). Как и многие мембранные рецепторы типа I, белок APP димеризуется латерально в плазматической мембране. Полноразмерный белок APP включает в себя следующие структурные домены: цистин-богатый домен, имеющий сходство с фактором роста (GFLD – growth factor-like domain), гепарин-связывающий домен (HBD), медьсвязывающий домен (CuBD), вместе образующие домен E1, кислый домен Aс, домен, гомологичный ингибиторам сериновых протеаз типа Кунитц (KPI),

домен E2, трансмембранный домен, содержащий последовательность $A\beta$, и внутриклеточный домен AICD (APP intracellular domain) (рис. 1а). Вследствие большого размера (~750 аминокислотных остатков), димеризации в мембране и подвижности доменов друг относительно друга структура высокого разрешения полноразмерного белка APP не получена до сих пор, поэтому пространственные структуры доменов были определены отдельно различными структурными методами.

1.1. GFLD/HBD

Для цистин-богатого домена, имеющего сходство с фактором роста (28–123), содержащего сайт связывания гепарина (HBD) [5], методом PCA определена структура с разрешением 1.8 Å (PDB:1MWP, рис. 1б) [6]. Структура домена состоит из девяти β -тяжей и одной α -спирали, которые складываются в компактный глобулярный домен. Большинство тяжей очень короткие, шесть из них содержат только три остатка. В основе домена лежит пятицепочечный антипараллельный β -лист. Центральным лист покрыт с одной стороны двумя двухцепочечными листами и с другой – спиралью и длинной петлей. Домен содержит три дисульфидных мостика, которые связывают его части вместе [6]. Поверхность домена

в целом имеет положительный заряд, показано ее взаимодействие с фосфоинозитол бифосфатами и трифосфатами [7]. Структура этого домена в свободном состоянии также была определена с разрешением 1.33 Å (PDB:4PQD, не опубликовано). Домен способен образовывать комплекс с двухвалентным ионом меди, структура и свойства которого описаны в [8] (PDB:4JFN, 1.75 Å).

1.2. CuBD

Структура изолированного медьсвязывающего домена CuBD была определена методом ЯМР в растворе ([9], PDB:1OWT) и позднее методом РСА с разрешением 1.61 Å ([10], PDB:2FJZ, рис. 1в). Структура состоит из α -спирали (остатки 147–159) и β -листа из трех тяжей (остатки 133–139, 162–167 и 181–188). Дисульфидная связь между Cys-133 и Cys-187 соединяет тяжи 1 и 3, а между Cys-158 и Cys-186 – α -спираль с тяжем 3. Дисульфидная связь Cys-144–Cys-174 соединяет две петли на другом конце молекулы. Ввиду малого количества гидрофобных остатков в окрестностях этой дисульфидной связи можно предположить, что она очень важна для стабилизации структуры в области, которая не имеет вторичной структуры. В дополнение к трем дисульфидным связям в домене есть небольшое гидрофобное ядро (Leu136, Trp150, Val153, Ala154, Leu165, Met170, Val182 и Val185), стабилизирующее структуру. Поверхность CuBD сильно заряжена с несколькими областями отрицательного (Glu156, Glu160 и Glu183, Asp167 и Asp131) и положительного (Lys132, Lys134, Lys161, His147, His151 и Lys155) электростатического потенциала.

В [9] на основе данных электронного парамагнитного резонанса предложен механизм восстановления иона меди Cu^{2+} до Cu^+ , опосредованный взаимодействием ионов меди с CuBD. Так как сайт связывания меди имеет достаточно жесткую тетраэдрическую геометрию, предпочтительную для связывания иона Cu^+ , в то время как для Cu^{2+} энергетически выгодной является гексаэдрическая геометрия, связывание меди сопровождается восстановлением Cu^{2+} до Cu^+ . В отсутствие других восстановителей это приводит к окислению Met170, входящего в сайт связывания. Однако вряд ли имеет место такая реакция *in vivo*, так как приводит к ослаблению взаимодействия сайта с ионом, а восстановителем могут выступать аскорбат или тиолы.

1.3. E1

Ряд работ посвящен исследованию полноразмерного домена E1. Была получена кристаллическая структура гепарин-индуцированного димера домена E1 [11] (PDB:3KTM). Низкие *B*-факторы остатков на границе доменов, сравнимые с

B-факторами внутренних остатков, заметно сниженные по сравнению с независимо определенными структурами отдельных поддоменов, говорят о том, что домен E1 образует единую структурную единицу и не состоит из двух независимых доменов, соединенных гибким линкером. Сам линкер, являясь высоко консервативным фрагментом молекулы, не принимает никакой стандартной вторичной структуры, но соединяет оба субдомена, как zipper, и является частью гидрофобной междоменной границы раздела. Граница раздела стабилизирована одним солевым мостиком и несколькими сетями водородных связей. Также интересно, что Cu^{2+} -связывающий сайт находится вблизи границы раздела между GFLD и CuBD. В [12] была получена структура E1-домена в комплексе с ионом меди (PDB:4JFN), где он скоординирован с атомами Nε2 двух консервативных остатков гистидина (His108 и His110). Она подтвердила, что сайт связывания иона находится на границе раздела доменов. Интересно, что связывание меди вызывает восстановление дисульфидного мостика (Cys98–Cys105), хотя очистку и кристаллизацию не проводили в восстанавливающих условиях, а также частичную дестабилизацию домена E1, как показано с помощью анализа *B*-факторов и спектроскопии кругового дихроизма. Также было подтверждено наличие двух пар остатков гистидина, которые координируют медь в GFLD (His108/His110) и CuBD (His147/His151), и что связывание ими ионов меди вызывает димеризацию E1-домена как *in vitro*, так и *in vivo*.

Структура домена E1 субатомного разрешения определена в [13] (PDB:4PWQ, 1.4 Å, рис. 1з). На основе ее анализа, а также данных динамического рассеяния света и точечного мутагенеза было установлено, что упаковка субдоменов GFLD и CuBD внутри E1 носит pH-зависимый характер. При нейтральных и слабощелочных значениях pH образуется более открытая конформация комплекса, а при кислых – более компактная. Специфические водородные связи между двумя карбоксилатами Asp177 и Glu87, а также между Asn89 и His147 являются основными детерминантами этого pH-управляемого конформационного переключения в APP–E1.

1.4. KPI

Первым кристаллизованным фрагментом APP был домен, гомологичный ингибиторам протеаз типа Кунитц (KPI), присутствующий только в не нейронных изоформах (sAPP α 751 и sAPP α 770) [14]. Поскольку этот домен кодируется одним экзоном и вставляется в середину последовательности APP, а также с учетом того, что он довольно гомологичен другим ингибиторам протеазы типа Кунитц, вероятно, что он представляет собой не-

зависимую единицу фолгинга и что структура изолированного домена хорошо представляет структуру домена в контексте полноразмерного APP. Домен ингибирует сериновые протеазы, как и гомологичные ему ингибиторы, структура его комплекса с бычьим трипсином была решена в [15] (PDB:1AAP, рис. 1г).

1.5. E2

Еще одна область внеклеточной части APP с определенной пространственной структурой и высокой консервативностью – E2-домен – присутствует, в том числе, и в нейрональной изоформе APP695. Первая структура этого сегмента APP была определена методом ЯМР в растворе (PDB:1TKN). Она включает в себя остатки 460–569, что соответствует поддомену E2, устойчивому к протеолизу трипсином [16]. Структура состоит из четырех α -спиралей, соединенных короткими петлями. Позднее были опубликованы структуры полноразмерного домена E2, определенные методом PCA (PDB:3NYL [17], PDB:3NYJ [17], рис. 1д). Полноразмерный домен E2 состоит из шести α -спиралей. N-концевая часть включает спирали 1 и 2, половину спирали 3. Ее наиболее заметной особенностью является двуспиральный элемент (colied-coil), образованный спиралями 2 и 3. Отсутствие какого-либо видимого поворота спиралей друг относительно друга может являться следствием нарушения в гептадном повторе, характерном для канонических двуспиральных белков. C-концевая часть содержит вторую половину спирали 3 и спирали 4, 5 и 6. Эта структура состоит в основном из трехпечечных антипараллельных двуспиральных элементов: спирали 3 и 5 параллельны, тогда как 4 находится в противоположной ориентации. C-концевой пучок спиралей обладает левосторонней закрученностью в отличие от N-концевого. Эти два структурных элемента, объединенные непрерывной спиралью 3, находятся на ее противоположных сторонах и напрямую не взаимодействуют друг с другом. Центральная спираль 3 включает 56 аминокислот и имеет длину 90 Å. Она слегка изогнута в средней части, хотя не содержит остатков пролина. Изгиб может быть следствием взаимодействия со спиралями 1/2 и 4/5 либо образования димера E2. Обе описанные кристаллические структуры E2 содержат домен в виде антипараллельного димера и почти полностью повторяют друг друга. Кроме того, в [17] показано, что домен E2 образует димер в растворе, и димеризация усиливается в присутствии гепарина. В дополнение к этому высокое разрешение структуры позволило идентифицировать в ней два сайта связывания фосфат-анионов, которые в физиологических условиях участвуют в связывании гепарина.

Как и E1, домен E2 обладает способностью связывать ионы Zn^{2+} и Cu^{2+} . Структуры комплексов решены в [18] (PDB:3UMI, 3UMK). Связывание ионов происходит в двух сайтах с разной спецификой. Аффинность к ионам меди оказывается гораздо больше, чем к ионам цинка. Данные поверхностного плазмонного резонанса, спектроскопии кругового дихроизма и ограниченного протеолиза показали, что связывание металлов в растворе приводит к значительным конформационным изменениям. Анализ структуры комплексов подтверждает это. Связывание ионов требует определенной координационной геометрии и, следовательно, определенной конформации домена. Так, при связывании меди происходит поворот N-концевого субдомена на 12° . Координация меди требует ориентации четырех координирующих гистидинов (His313, His382, His432 и His436) в отличие от цинка, в случае которого молекула воды заменяет в качестве лиганда His313. Таким образом, домен E2 может находиться по крайней мере в трех разных конформационных состояниях, различающихся гибкостью: наиболее гибкая свободная форма; более жесткая форма, связанная с цинком; и еще более ограниченный во внутримолекулярных движениях комплекс с медью [18].

1.6. Полноразмерный эктодомен

Полноразмерный эктодомен APP является самостоятельной биологически активной молекулой, биологические функции которой определяются фрагментами sAPP α и sAPP β , продуцируемыми α - и β -секретазами соответственно. Структура высокого разрешения этого фрагмента до сих пор не определена. Однако в [19] методом МУРР и молекулярного моделирования была исследована структура фрагментов sAPP α_{695} sAPP α_{770} . Из кривых рассеяния были определены функции попарного распределения расстояний $P(r)$ для этих молекул, анализ которых позволил установить максимальные размеры белков D_{max} и их радиусы вращения R_g . Максимумы длины, полученные из функций $P(r)$, оказались равными 135 и 150, а радиусы вращения 42.6 ± 0.3 и 46.5 ± 2 Å для sAPP α_{695} и sAPP α_{770} соответственно. Практически такие же результаты дал анализ графиков Гинье, также подтвердивший монодисперсность исследованных образцов. Из соотношения между D_{max} и R_g , а также общей формы распределения $P(r)$ можно сделать вывод о вытянутой форме исследованных частиц. На основе тех же данных МУРР было проведено моделирование *ab initio* полноразмерного sAPP α с использованием в качестве модели цепочки псевдоостатков и с применением метода симулированного отжига для минимизации расхождения между расчетной и экспериментально измеренной кривой $P(r)$ [20].

Для проверки моделей по ним были рассчитаны отношения гидродинамических объемов и молекулярных масс $sAPP\alpha$ и $sAPP\beta$, которые показали хорошую сходимость к экспериментально измеренным значениям.

Известные кристаллические структуры доменов GFLD, CuBD, KPI и E2 были вписаны в структуру низкого разрешения, полученную по данным МУРР методом моделирования твердого тела [21], что позволило впервые рассмотреть модель структуры полноразмерного эктодомена APP. Также в [15] методом МУРР показано, а с помощью аналитической гель-фильтрации и ультрацентрифугирования подтверждено, что при взаимодействии с высокомолекулярным гепарином $sAPP\alpha_{695}$ образует комплекс с соотношением белок:гепарин 2:1, т.е. происходит гепарин-индуцированная димеризация эктодомена APP в растворе.

1.7. AICD

Внутриклеточный домен APP, содержащий около 50 аминокислотных остатков в свободном состоянии, не имеет стабильной упорядоченной структуры, что было показано методом ЯМР в растворе [22] и что типично для внутриклеточных доменов многих трансмембранных белков. Однако нестабильные элементы вторичной структуры, о наличии которых свидетельствуют измеренные химические сдвиги H_N и $H\alpha$, могут являться зачатками структуры, которую принимает домен при взаимодействии с внутриклеточными лигандами. Структурные изменения в AICD при взаимодействии с белком Fe65 были зарегистрированы методом ЯМР, а структура комплекса из 31-аминокислотного фрагмента AICD и домена PTB2 Fe65 решена методом РСА [23] (PDB:3DXD). При связывании с Fe65 C-концевая часть AICD образует структуру, состоящую из двух α -спиралей, соединенных коротким β -листом. На основе анализа структуры мутантов T668E и T668A предполагаем механизм разрушения комплекса при фосфорилировании T668, согласно которому присоединение фосфата приводит к нарушению кэп-бокса N-концевой спирали AICD и ее дестабилизации. Аналогичная структура была получена с 32-аминокислотным C-концевым пептидом AICD и белком мыши Fe65L1 методом ЯМР [24] (PDB:2ROZ, рис. 1ж).

1.8. Трансмембранный домен

Полноразмерный трансмембранный домен APP включает в себя остатки Gly700-Leu723, часто к нему добавляют примембранный регион Gln686-Lys699, так как они образуют общую структурную единицу, взаимодействующую с мембраной. Пептид слишком мал, гидрофобен и

подвижен, поэтому его сложно использовать как для получения кристаллов, так и исследовать методом криоэлектронной микроскопии. Основным методом исследования этого домена стала спектроскопия ЯМР. Пептид нерастворим в воде, поэтому необходим подбор мембраномоделирующей среды для поддержания его в нативной конформации. Так как размеры липосом и нанодисков не позволяют эффективно использовать ЯМР-спектроскопию для их изучения, структуры трансмембранного домена APP определены с использованием детергентных мицелл как мембраномоделирующей среды.

Первая структура полноразмерного трансмембранного домена опубликована в [25] (PDB:2LLM, рис. 1е) (использованы мицеллы додецилфосфохолина в качестве мембраномоделирующей среды), она включала в себя остатки Gln686-Lys726, т.е. весь трансмембранный и примембранный участки. Структура состоит из двух α -спиралей: короткой примембранной (Lys687-Asp694), взаимодействующей с поверхностью мембраны, и длинной трансмембранной спирали (Gly700-Leu723), связанных через небольшую подвижную петлю. Вблизи остатков Gly708-Gly709 трансмембранной спирали – небольшой 20° -ный изгиб. Топология структуры была подтверждена с помощью анализа динамики и метода спиновых меток. Также косвенно с помощью анализа изменений химических сдвигов при добавлении в среду холестерил гемисукцината – растворимого аналога холестерина – было подтверждено наличие в петле, соединяющей две спирали, сайта связывания холестерина [2]. В [26] опубликована структура димера Gln686-Lys726 (PDB:2LOH), в которой трансмембранные α -спирали образуют параллельный левозакрученный димер – расширенный гептадный мотив $I^{702}X_3M^{706}X_2G^{709}X_3A^{713}X_2I^{716}X_3I^{720}X_2I^{723}$. Примембранная спираль, хотя и не участвует в димеризации непосредственно, также испытывает влияние перехода домена в димерную форму, что подтверждается анализом изменений химического сдвига. На другой стороне спирали находится GG4, который часто является мотивом димеризации трансмембранных доменов односпиральных трансмембранных белков [27], однако в данном димере он не задействован, и, возможно, участвует во взаимодействии трансмембранного домена с холестерином. Димер имеет площадь контакта $\sim 630 \text{ \AA}^2$, равномерно распределенную вдоль трансмембранной спирали, и энергию димеризации ~ 1.4 ккал/моль.

Еще одна структура [28], решенная методом ЯМР, включает трансмембранный домен вместе с внутриклеточным доменом AICD (PDB:2LP1, рис. 1и). В качестве мембраномоделирующей среды были использованы мицеллы LMPG. Структура состоит из неупорядоченного N-кон-

цевого сегмента, переходящего в примембранную спираль, взаимодействующую с поверхностью мембраны. Она, в свою очередь, соединена с трансмембранной спиралью с характерным изгибом в центре, как и в [25]. Далее идут С-петля и С-концевая спираль из девяти аминокислотных остатков, погруженная в мембрану. N-концевая часть структуры в целом повторяет (PDB:2LLM), а С-концевая представляет собой почти неупорядоченное состояние AICD. Также, но уже при использовании в качестве среды бицелл DHPC–DMPC с содержанием холестерина 0–20 мол. % в [28] показано связывание холестерина в районе N-концевой части трансмембранной спирали и N-концевой межспиральной петли с константой релевантной физиологической концентрации холестерина в мембране клеток. Связывание вовлекает незадействованный в димеризации GG4-мотив, что проливает свет на его физиологическую роль. Однако в [29] при исследовании изолированного примембранных спирали и петли, была получена альтернативная правозакрученная структура димера с ассоциацией через мотив GX₃A (PDB:2LZ3) с углом между спиралями 22°. Также была показана роль наследственной патогенной мутации V715M, дестабилизирующей спиральную структуру. Для другой наследственной мутации L723P установлен эффект дестабилизации С-концевой части трансмембранной спирали с увеличением доступа воды к пептидному остову, что может приводить к заглужению N-концевой стороны спирали и ускорять ее протеолитическое расщепление γ -секретазой [30].

1.9. A β

β -амилоидные пептиды – наиболее часто исследуемые фрагменты APP. Это связано с тем, что именно β -амилоидный пептид – основной компонент амилоидных бляшек, обнаруживаемых в тканях мозга, пораженных болезнью Альцгеймера [31]. Более точно, основными компонентами амилоидных отложений являются пептиды A β _{1–40} и A β _{1–42}, или, в другой нумерации, APP672–711 и APP672–713. Первая опубликованная структура A β _{1–40} была определена методом ЯМР в растворе (PDB:1AML, рис. 2а) [32]. В ней выделяются две α -спиральные области: от Gln15 до Asp23 и от Pe31 до Met35. Однако в качестве среды была использована смесь 40% трифторэтанола и 60% воды, которая предотвращает агрегацию пептида, но далека от его нативного мембранного окружения. В частности, она может способствовать образованию искусственно индуцированной α -спиральной структуры [33], следствием может быть достаточно большое среднее квадратичное отклонение 4.8 Å по 20 лучшим рассчитанным моделям. Структура того же пептида была опре-

делена с использованием додецилсульфата натрия (SDS) как мембраномоделирующей среды (PDB:1BA4, рис. 2б) [34]. В этой структуре наблюдается только одна протяженная α -спираль Gln15–Val36 с изгибом в области Gly25–Asn27. Хотя две части спирали проявляют заметную подвижность друг относительно друга, среднеквадратичное отклонение по остаткам 15–36 составляло ~ 1.5 Å, а N-концевая часть структуры находится в неупорядоченном состоянии. В [35] удалось провести структурные исследования пептида A β _{1–40} в водном окружении, используя специальный протокол дезагрегации и работая с очень малыми – 77 мкМ – концентрациями пептида. Такие условия наиболее приближены к условиям *in vivo*. В полученной методом ЯМР структуре (PDB:2LFM, рис. 2г) центральная часть пептида H13–D23 образует спираль 3_{10} , что подтверждается наличием *i,i+2* контактов, выявленных методом спектроскопии ядерного эффекта Оверхаузера (ЯЭО). Эта спираль имеет гидрофобные участки с обеих сторон, с которыми взаимодействуют гидрофобные остатки в почти неструктурированных N- и С-концах пептида. Хотя на N- и С-концах отсутствует стабильная вторичная структура, они не являются полностью неструктурированными и содержат много коротких элементов α - и β -структур, образовавшихся в результате локальных взаимодействий между боковыми цепями концов и центральной спиральной структурой. Такая структура может быть переходным звеном между α -спиральной структурой, характерной для трансмембранного домена APP, и β -структурой – основой β -амилоидных олигомеров и фибрилл [36]. Для пептида A β _{17–34} была показана α -спиральная конформация в областях Phe20–Ser26 и Lys28–Gly33 с перегибом около остатков Ser26–Asn27–Lys28 (PDB:2MJ1). Для стабилизации пептида в растворимом состоянии на N- и С-концы были добавлены по два остатка глутаминовой кислоты [37], что позволило провести измерения с концентрацией пептида в образце ~ 3 мг/мл в физиологически релевантных условиях без использования детергентов или органических растворителей.

Перегиб в середине пептида, который разбивает его на две спирали, предположительно, позволяют пептиду претерпевать структурные изменения от α -спирали к β -листу, когда две половинки сближаются. 42-членный пептид A β _{1–42} является более сложным объектом для изучения. Имея два дополнительных гидрофобных остатка на С-конце, он еще более склонен к быстрой агрегации в растворе, чем A β _{1–40}. Тем не менее, растворив пептид в гексафторэтаноле (HFIP) и добавляя необходимое количество воды, можно получить A β _{1–42} в растворенном состоянии в почти водной среде и исследовать его свойства. При ва-

рыировании доли воды в смеси от 20 до 99% с помощью спектроскопии кругового дихроизма и ЯМР удалось наблюдать постепенный переход от α -спиральной конформации к β -структурной [38]. Большая часть молекул переходит к β -структурной конформации при концентрации воды около 85%, но поскольку за этим следует достаточно быстрая агрегация пептида, для получения структуры методом ЯМР была выбрана концентрация воды 70%, что соответствует молярной доле воды 92.9%. Структура (PDB:1Z0Q, рис. 2в) состоит из двух α -спиралей, соединенных петлей: N-концевая спираль достаточно устойчива, со среднеквадратичным отклонением 1.67 Å; C-концевая спираль теряет стабильность, укорачивается, распространяясь лишь до остатка Gly33, и для некоторых рассчитанных конформеров начинает проявлять свойства, характерные для β -укладки.

Все структуры, содержащие α -спиральные элементы, имеют сходство с интактным трансмембранным доменом APP, также обладающим α -спиральной конформацией, и, по всей видимости, представляют разные стадии переходного процесса от α -спирального трансмембранного домена к β -структурным димерам, олигомерам и фибриллам, который реализуется после расщепления полноразмерного APP секретазами. Первая структура фибрилл была получена методом ЯМР в растворе в [39]. Для ее получения образцы фибрилл растворяли в DMSO/TFA, что позволило получить агрегаты, достаточно мелкие для изучения методом ЯМР. Так как комплексы, полученные диссоциацией фибрилл по такому протоколу, все еще слишком большие для получения ограничений методом ЯЭО, для установления водородных связей и солевых мостиков, участвующих в стабилизации внутри- и межмолекулярных взаимодействий, был применен обмен H/D, сайт-направленный мутагенез и данные твердотельного ЯМР. Остатки 1–17 в данной структуре (PDB:2BEG, рис. 2д) неупорядочены, а остатки 18–42 образуют мотив β -тяж–поворот– β -тяж. β -тяжи соседних мономеров образуют два параллельных β -слоя. Межмолекулярная связь стабилизирована солевым мостиком между Asp23 и Lys28 соседних мономеров. Межмолекулярные контакты боковых цепей образуются между нечетными остатками тяжа 1 *n*-й молекулы и четными остатками тяжа 2*n*–1-й молекулы. Это приводит к неспаренной валентности для взаимодействия на концах фибрилл и объясняет их рост. Фибриллы другой структуры, состоящие из пептидов $A\beta_{1-40}$, были исследованы методом твердотельного ЯМР [40]. Фибриллы содержат параллельные β -тяжи из остатков 11–22 и 30–39, соединенные неупорядоченной вставкой 23–29. На основе данных ЯМР и просвечивающей электронной микроскопии была построена модель фибриллы, имеющая трехосевую симметрию от-

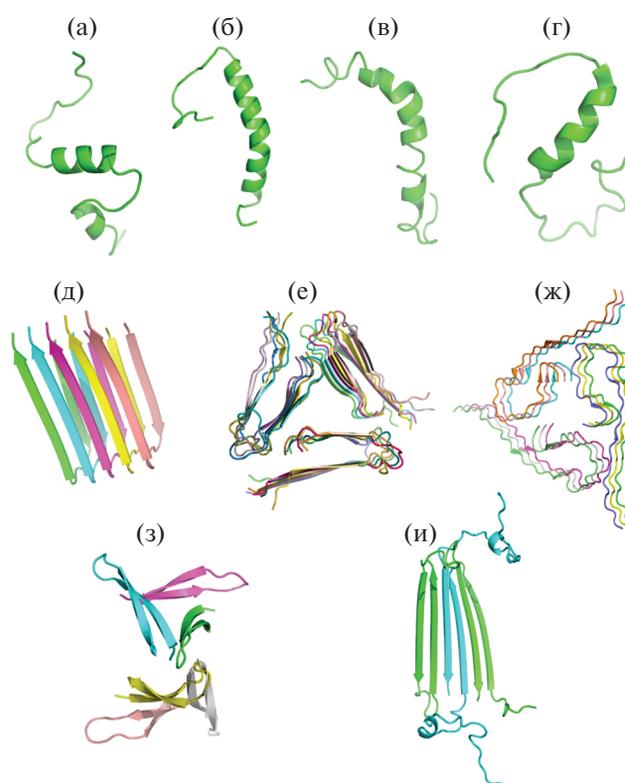


Рис. 2. Различные структуры β -амилоидных пептидов (а–г), образованных ими фибрилл (д–ж) и олигомеров (з, и): а – $A\beta_{1-40}$ в смеси TFE:H₂O 40:60 (PDB:1AML); б – $A\beta_{1-40}$ в мицеллах SDS (PDB:1BA4); в – $A\beta_{1-42}$ в смеси HFIP:H₂O 30:70 (PDB:1Z0Q); г – $A\beta_{1-40}$ в воде при концентрации пептида 77 μ M (PDB:2LFM); д – структура фибриллы $A\beta_{1-42}$, определенная методом ЯМР в растворе (PDB:2BEG); е – структура фибриллы $A\beta_{1-40}$ с трехосевой симметрией, построенная на основе данных ЯМР и просвечивающей электронной микроскопии (PDB:2LMQ); ж – структура фибриллы, построенная на основе данных ЯМР в твердом теле с использованием в качестве затравки биоматериала больных болезнью Альцгеймера (PDB:2M4J); з – структура додекамера $A\beta_{17-36}$, полученная методом рентгеновской кристаллографии (PDB:5HOY); и – структура тетрамера $A\beta_{1-42}$ в мицеллах DPC, полученная методом ЯМР в растворе, цветом выделены мономеры с различной структурой.

носительно оси ее роста (PDB:2LMQ, рис. 2е). Похожая структура (PDB:2M4J, рис. 2ж) была получена также с помощью данных ЯМР в твердом теле, но в качестве стартовых точек роста фибрилл использовали биологический материал, выделенный из мозга пациентов с клинически подтвержденной болезнью Альцгеймера [41]. Это является важным свидетельством в пользу биологической релевантности результатов, так как молекулярная структура и свойства фибрилл, полученных *in vitro*, повторяют структуру и свойства

тех образований, которые являлись начальными точками роста [42].

1.10. Олигомеры

На данный момент преобладает мнение, что наиболее токсичными для клеток формами β -амилоидных пептидов являются олигомеры [43–45]. Изучение структуры и свойств олигомерных форм осложняется трудностью получения их монодисперсных образцов, достаточно стабильных для проведения структурных исследований. В ряде работ по данным молекулярного моделирования, атомно-силовой микроскопии, электронной микроскопии, электрофизиологии и других методов были предложены модели образования олигомеров и мембранных пор и каналов на их основе. Так, была предложена модель формирования гексамера $A\beta_{1-42}$, принимающего форму β -бочонка. Собираясь в 36-членный агрегат, способный пронизывать мембрану, он увеличивает ее проводимость для ионов, в частности Ca^{2+} [46]. Электрофизиологические измерения показывают, что пептиды $A\beta_{1-42}$ способны образовывать как минимум три типа каналов с различной проводимостью, в то время как $A\beta_{1-40}$ каналов не образует [47].

Олигомерное состояние пептидов $A\beta$ является, по всей видимости, наиболее токсичным и, по крайней мере, в условиях *in vitro* транзитным с характерным временем жизни несколько часов или дней в зависимости от условий эксперимента [48, 49]. Для стабилизации олигомерных агрегатов и исследования их структуры применяют различные экспериментальные методики. В случае пептида $A\beta_{17-36}$, дополнительно стабилизовав дисульфидным мостиком изгиб, необходимый для взаимодействия его N- и C-концевых частей, удалось определить структуру додекамера методом PCA (PDB:5HOY, рис. 23). Структура додекамера $A\beta_{17-36}$ состоит из четырех тримерных блоков. В центральной полости, образованной четырьмя тримерными единицами, две молекулы Jeffamine M-600, использованного при кристаллизации, стабилизируют додекамерную структуру за счет гидрофобных взаимодействий с остатками в $A\beta_{17-36}$ [50]. Для уточнения структуры было проведено моделирование методом молекулярной динамики с использованием крупнозернистого (coarse grained) силового поля и модели диссипативной динамики частиц (DPD – dissipative particle dynamics) пептида $A\beta_{9-42}$ с экспериментально определенной структурой в качестве начального состояния. В этом моделировании C-концевые области спонтанно вписываются в центральную полость без каких-либо дополнительных ограничений на C-концевых участках. Таким образом, они участвуют в формировании гидро-

фобного ядра, стабилизирующего додекамерную структуру [51]. Додекамеры, в свою очередь, способны собираться в пору. Ее внешний диаметр составляет около 11 нм, а диаметр отверстия в центре поры ~2 нм. Толщина поры составляет ~5 нм, что сопоставимо с толщиной двухслойной липидной мембраны. Пять додекамеров собираются, образуя кольцевую структуру поры [50]. Олигомеры, использованные для кристаллизации, связываются с антителом A11, которое взаимодействует с олигомерными формами $A\beta$, в том числе с додекамерами массой 56 кДа, обнаруженными *in vivo* и обладающими токсичностью [52], но не взаимодействует с фибриллярными формами [53]. Это позволяет говорить о годности примененной модели и релевантности результатов.

Совсем другая структура олигомерной субъединицы, сформированной $A\beta_{1-42}$, была получена методом ЯМР в растворе с использованием детергента DPC в качестве мембраномоделирующей среды [54]. В таких условиях пептиды формировали тетрамеры, состоящие из структурно различных субъединиц. Один тип субъединиц $A\beta_{1-42}$ формирует β -шпильку из двух β -цепей, $\beta 1$ и $\beta 2$ – от G9 до Ala21 и от Gly29 до Val40. В субъединице второго типа остатки от Leu17 до Phe20 образуют короткую α -спираль, тогда как остатки Gly29–Phe41 образуют β -цепь $\beta 3$. Таким образом, ядро тетрамера состоит из шести β -тяжей, соединенных всего двумя β -поворотами, а также двух коротких и двух длинных, содержащих α -спиральный сегмент, гибких N-концов. Среднеквадратичное отклонение структуры (PDB:6RHY, рис. 2и), рассчитанное по атомам основной цепи, составляет 0.77 Å. Чтобы отнести сигналы и уточнить характер взаимодействия между субъединицами в различных конформациях, были проведены эксперименты с использованием смеси $A\beta_{1-42}$ и $A\beta_{17-42}$, лишённого сегмента, образующего α -спираль в субъединице второго типа и не способного сложиться в β -шпильку первого типа. Тетрамеры, структура которых была определена, способны самопроизвольно образовывать октамеры со структурой β -сэндвича, как показано методами эксклюзионной хроматографии/спектрометрии ионной подвижности и с помощью моделирования методом молекулярной динамики [54]. Также с помощью моделирования и электрофизиологических измерений была показана способность таких олигомеров формировать поры в липидном бислое.

2. А-СЕКРЕТАЗА

А-секретазы – протеазы, расщепляющие APP в α -сайте, что приводит к образованию неамилоидогенного пептида р3 и внеклеточной части sAPP α . Такую активность проявляют белки семейства ADAM (A Disintegrating And Metallopro-

teinase). В геноме человека закодировано 22 белка, относящихся к семейству ADAM. Все они расщепляют трансмембранные белки вблизи поверхности клетки в ходе процесса, называемого шеддингом. Субстраты ADAM включают факторы роста, цитокины, хемокины, рецепторы и молекулы адгезии. Таким образом, белки ADAM играют важную роль в межклеточной адгезии и передаче сигналов [55]. Хотя активность в отношении APP способны проявлять ADAM9 и ADAM17, основным путем, приводящим к р3 и sAPP α , является протеолиз ADAM10 [56, 57]. Помимо APP биологически важными субстратами ADAM10 являются Notch [58], эфринный рецептор [59], проэпидермальный фактор роста [60], HER2 [61], TNF- α [62] и другие белки. Металлопротеиназы ADAM представляют собой трансмембранные белки I типа с модульной доменной организацией. ADAM10, как и его ближайший гомолог – ADAM17 (также известен как TACE – TNF- α converting enzyme), имеет эктодомен, состоящий из N-концевой сигнальной последовательности. Далее следует так называемый продомен, за ним металлопротеиназный, дезинтегриновый и цистеин-богатый домены. После них идет трансмембранный домен, пронизывающий мембрану клетки, и небольшой цитоплазматический домен.

Как и многие другие члены семейства ADAM, ADAM10 экспрессируется как профермент, включающий инактивирующий продомен 10–213, а удаление при протеолизе конвертазой PC7 и фурином продомена в ходе созревания фермента приводит к его активации [63]. Продомен является селективным ингибитором ADAM10 с IC50 около 50 нм и не взаимодействует с другими белками того же семейства [64]. Изолированный продомен образует димер в растворе и по данным спектроскопии кругового дихроизма имеет стабильный фолд, состоящий из α - и β -структурных элементов [64]. По всей видимости, продомен также играет роль шаперона, необходимого для созревания ADAM10, так как экспрессия укороченного варианта белка, лишённого продомена, приводит к протеолитически неактивному продукту, а коэкспрессия изолированного продомена вместе с ним восстанавливает протеолитическую функцию [63]. Был описан ряд мутаций, затрагивающих продомен ADAM10, ассоциированных с поздним развитием болезни Альцгеймера [65]. На данный момент нет структурных данных о механизмах ингибирования протеолитической активности и специфичности этого взаимодействия.

Впервые структура дезинтегринового и цистеин-богатого доменов (PDB:2A07) была определена для ADAM10 из *Bos taurus* методом РСА с разрешением 2.9 Å [59]. Пары доменов образуют непрерывную удлинённую, слегка изогнутую

структуру длиной около 75 Å. Дезинтегриновый домен содержит два β -тяжа, а цистеин-богатый домен – семь коротких β -тяжей и три короткие α -спирали. Помимо небольших гидрофобных ядер в цистеин-богатом домене и на границе между двумя доменами общая структура в основном стабилизируется набором дисульфидных связей. Дезинтегриновый домен ADAM10 (остатки 455–550) обладает структурной гомологией с дезинтегринами триместатином [66] и ингибитором свертывания крови [67]. Дисульфидные связи в дезинтегриновом домене ADAM10 топологически идентичны связям в триместатине. Богатый цистеином домен (остатки 551–646) вместе с дезинтегриновым доменом образует небольшое гидрофобное ядро, формируя непрерывную структуру, которая обладает небольшой междоменной гибкостью, а дисульфидный мостик (Cys530–Cys562) между двумя доменами дополнительно стабилизирует их взаимодействие. Цистеин-богатый домен имеет характерный для ADAM-белков α/β -фолд, стабилизированный пятью дисульфидными связями. Один из трех его β -листов состоит из трех центральных β -тяжей, которые контактируют с тремя α -спиралями, а два других β -листа состоят из двух N-концевых и двух C-концевых β -тяжей соответственно. N-концевой β -лист контактирует с α -спиралью цистеин-богатого домена с одной стороны и с дезинтегриновым доменом с другой. На поверхности цистеин-богатого домена находится достаточно большой отрицательно заряженный карман, через который происходит взаимодействие с субстратом [59].

Для изучения взаимодействия механизма специфичности субстрата и роли каждого из доменов в [68] методом РСА была определена структура полноразмерного эктодомена ADAM10 человека (PDB:6BE6, рис. 3а). Внеклеточный домен ADAM10 принимает компактную конформацию, напоминающую острие стрелы, где каталитический – металлопротеиназный – домен окружен дезинтегриновым и цистеин-богатыми доменами. Структура, вероятно, представляет аутоингибированную конформацию, поскольку расположение богатого цистеином домена в устье пептидсвязывающей бороздки частично перекрывает активный сайт. Как и в структуре из [59] (PDB:2A07), дезинтегриновый и цистеин-богатый домены содержат мало вторичной структуры и принимают чашевидную форму, а их структура стабилизируется сетью дисульфидных связей. Каталитический домен ADAM10 похож на активные центры других ферментов из семейства ADAM, матриксных металлопротеаз и металлопротеиназ из яда змей. Одна сторона кармана, в котором расположен активный сайт, образована пятицепочечным β -листом, вторая – центральной спиралью. Эта спираль включает мотив HEX₂N, часть каталитического центра HEX₂NX₂GX₂N

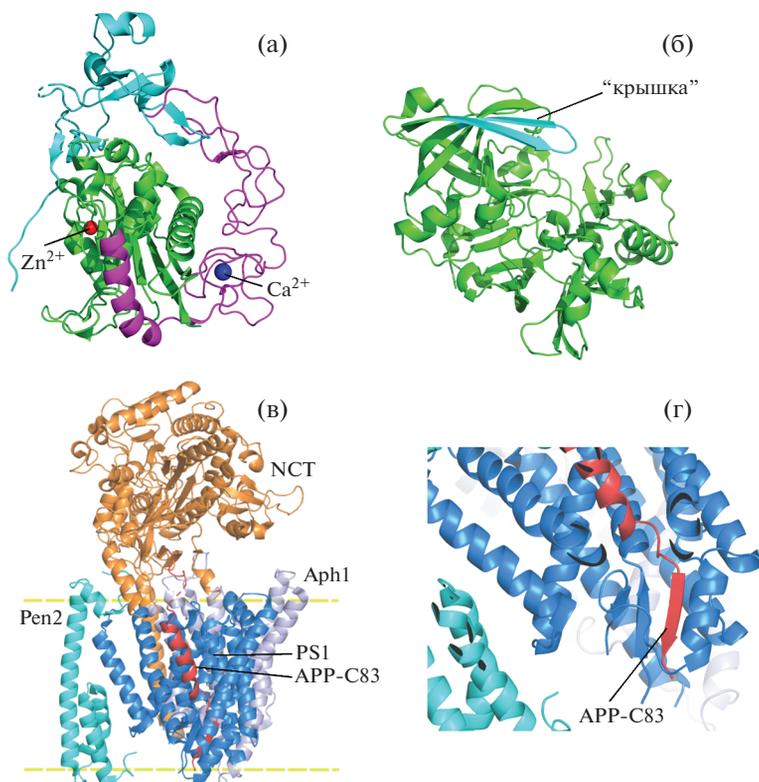


Рис. 3. Структура эктодомена ADAM10 (PDB:6BE6), определенная методом PCA (а), сферами обозначены координированные ионы цинка и кальция. Структура эктодомена свободного BACE1 (PDB:1SGZ) (б): отмечена “крышка”. Структура γ -секретазного комплекса (в), содержащего никастрин (NCT), Aph1, Pen2, пресенилин (PS1), APP-C83, линиями схематически изображена поверхность мембраны. Увеличенный фрагмент структуры γ -секретазного комплекса, на котором показано образование β -слоя сегментами APP-C83 и PS1 (г).

[69], содержащий два гистидина, которые координируют связанный ион цинка (His383 и His387) и остаток глутамата (Glu384), участвующий в кислотном катализе. Третий гистидин цинксвязывающего мотива (His393) расположен на петле, следующей за этой спиралью.

Интересной особенностью данной структуры является то, что каждый активный центр фермента в кристалле занят С-концом соседней белковой субъединицы, имитируя процесс узнавания субстрата. Один из центров координации цинка занят С-концевым атомом кислорода из остатка Gly655. То есть С-концевой гистидиновый таг, используемый для очистки белка, был отщеплен во время выделения или кристаллизации, что привело к случайному формированию комплекса фермент–продукт. Из анализа структуры следует, что цистеин-богатый домен препятствует доступу белковых субстратов к карманам S1' и S2', что приводит к автоингибированию. Наложение структуры эктодомена ADAM10 на структуру комплекса дезинтегринового и цистеин-богатого домена с модулирующим активностью ADAM10 антителом 8C7 [70] (PDB:5L0Q) показывает, что каталитический домен ADAM10 и тя-

желая цепь 8C7 будут перекрываться, что приведет к значительному стерическому конфликту. То есть связывание антитела 8C7 приведет к увеличению доступности активного сайта и снятию автоингибирования. По биохимическим данным связывание 8C7 действительно увеличивает активность ADAM10 [68].

3. В-СЕКРЕТАЗА

В-секретаза – первый фермент, осуществляющий процессинг APP по амилоидогенному пути. Именно с расщепления APP в β -сайте начинается серия протеолитических реакций, приводящих к образованию пептидов A β 40 и A β 42. Ферментативная активность β -секретазы определяется двумя близкими гомологами BACE1 и BACE2. Активность BACE1 превалирует в нервной ткани и повышается с возрастом, и в случае развития болезни Альцгеймера [71] в последнее время упрочняется мнение, что BACE2 также может вносить вклад в развитие патологии [72]. BACE1 экспрессируется преимущественно в нейронах, и ее активность важна для нормальной миелинизации аксонов [73], что опосредовано, по всей видимо-

сти, активностью BACE1 в отношении субстрата Neuregulin 1 (Nrg1) [74]. Среди других субстратов BACE1 важные биологические функции выполняют лиганд рецептора Notch Jag1 [75], натриевый потенциал-зависимый канал $\beta 2$ Nav $\beta 2$ [76], калиевые потенциал-зависимые каналы KCNE1 и KCNE2 [77], ассоциированный с судорогами 6-подобный белок Sez6L [78]. Методом протеомики было идентифицировано 68 потенциальных субстратов BACE1 [79], позднее с помощью более детального анализа секретома к ним было добавлено еще 23 [78].

BACE1 представляет собой интегральный трансмембранный гликопротеин I типа с отщепляемой сигнальной последовательностью из 21 остатка, продомен из 23 остатков с большим эктодоменом из 421 остатка, одним трансмембранным доменом из 22 аминокислот и коротким цитоплазматическим сегментом из 24 остатков [80]. Продомен проявляет лишь слабые свойства ингибитора активности BACE1 в отличие от похожих доменов других аспартатных протеаз, однако помогает фолдингу фермента [81]. Известно, что взаимодействие BACE1 с липидным бислоем критически важно для протеолиза APP и BACE1 с делетированным трансмембранным доменом и не активно в отношении APP *in vivo* [82].

Структура эктодомена BACE1 в комплексе с ингибитором пептидной природы изучена в [83]. Структура была получена методом молекулярного замещения с использованием в качестве начальной модели человеческого пепсина, который на 22% гомологичен BACE1. Протеазный домен BACE1 (PDB:1FKN) имеет двулопастную структуру, типичную для аспартатных протеаз, и демонстрирует высокую степень консервативности с пепсином и BACE2 [84] (PDB:2EWY). Субстрат, связывающий карман и содержащий активный сайт, расположен между N- и C-концевыми долями и частично прикрыт “крышкой”, представляющей собой антипараллельную шпильку, которая контролирует доступ субстрата к активному сайту и протеолитическую специфичность. В целом структура почти полностью повторяет структуру пепсина, за исключением восьми коротких вставок, образующих петли или спирали на поверхности глобулы и несколько увеличивающих общую поверхность молекулы. Из 21 остатка продомена только последние шесть видны на картах электронной плотности. Остальные, вероятно, подвижны, что согласуется с тем, что неструктурированный просегмент вытесняется из щели активного сайта ингибитором. В структуре свободного эктодомена (PDB:1SGZ, рис. 2б) “крышка” оказывается зафиксированной в еще более открытом состоянии водородной связью между Tyr41 и Gly74, в то время как в структуре со связанным ингибитором Tyr41 взаимодействует с его боковыми цепями P1 и P2'. При этом остается

пространство 6.5 Å, достаточное для ограниченного прохода субстрата к активному центру [85].

Структура проливает свет на патологический механизм “шведской” наследственной мутации APP KM670/671NL. Подсайты BACE1 S1 и S3, состоящие в основном из гидрофобных остатков, имеют конформации, сильно отличающиеся от пепсина из-за отсутствия пепсиновой спирали 111–114. Боковые цепи ингибитора P3 – валин – и P1 – лейцин – плотно упакованы друг относительно друга и имеют сильные гидрофобные контакты с ферментом, особенно P1, который взаимодействует с Tyr71 и Phe108. В APP дикого типа остатки P2 и P1, смежные с сайтом расщепления секретазой, – это лизин и метионин соответственно. “Шведский” мутант APP содержит аспарагин и лизин в этих положениях, что приводит к более сильному межмолекулярному взаимодействию и 60-кратному увеличению K_{cat}/K_m APP [86].

Трансмембранный домен BACE1 является крайне интересным объектом, так как содержит уникальный мотив $M^{462}X_3C^{466}X_3M^{470}X_3C^{474}X_3C^{478}$, способный к тримеризации и взаимодействию с ионами меди [87]. Поскольку активность BACE1 является ключевым звеном в патогенезе болезни Альцгеймера, определение его структуры дало начало большому количеству исследований, цель которых – найти клинически эффективные и безопасные ингибиторы этого фермента. На данный момент методом PCA получено более 400 пространственных структур комплексов BACE1 с различными ингибиторами. Ряд из них находится на поздних стадиях клинических исследований.

4. Г-СЕКРЕТАЗА

Г-секретаза завершает протеолитический процессинг APP, предварительно расщепленного α - или β -секретазой. Ее активность приводит к образованию фрагмента р3 и β -амилоидных пептидов соответственно, а также внутриклеточного фрагмента AICD. Г-секретаза обладает несколькими характерными и крайне интересными особенностями. Во-первых, в отличие от α - и β -секретаз, активные домены которых растворимы, протеолиз γ -секретазой происходит в толще мембраны [88]. Во-вторых, γ -секретаза, по всей видимости, не имеет ярко выраженной субстратной специфичности и выполняет роль “трансмембранной протеасомы”, расщепляя трансмембранные домены различных белков [89, 90]. Всего насчитывается более 90 трансмембранных белков I типа, трансмембранные домены которых расщепляет γ -секретаза, после того как их эктодомены были отщеплены шеддазами, такими как α - и β -секретазы [90]. В-третьих, γ -секретаза способ-

на взаимодействовать со своим субстратом несколько раз, на каждом последующем этапе отщепляя от него по три–четыре аминокислотных остатка и таким образом создавая набор продуктов различной длины [91, 92]. Именно так образуются пептиды $A\beta_{1-40}$ и $A\beta_{1-42}$, изменения в абсолютном и относительном количестве которых сопровождают патогенез болезни Альцгеймера. Кроме того, именно в последовательности PS1 – компонент γ -секретазного комплекса находится большинство мутаций, ассоциированных с ранним развитием болезни Альцгеймера [93].

Таким образом, γ -секретазы является ключевым ферментом, участвующим в образовании β -амилоидных пептидов, и изучение ее структуры необходимо для понимания молекулярных механизмов развития этого заболевания.

Γ -секретазы, а точнее γ -секретазный комплекс, является мультисубъединичным ферментом и состоит из четырех трансмембранных белков: PS (presenilin, представлен белками PS1 или PS2), NCT (nicastrin), Aph1 (anterior pharynx defective 1), Pen2 (presenilin enhancer 2) [94]. Комплекс включает в себя 20 трансмембранных спиралей и имеет молекулярную массу ~170 кДа, в результате гликозилирования никастрина к нему добавляются еще 30–70 кДа [88]. Пресенилин осуществляет каталитическую функцию, никастрин – функцию распознающего субстрат центра, остальные субъединицы действуют как необходимые кофакторы [95]. В ходе созревания комплекса происходит эндопротеолиз пресенилина с образованием N- и C-концевых фрагментов, являющихся частями активной γ -секретазы [96]. Так как компоненты комплекса претерпевают сложный процесс созревания и по отдельности не проявляют биологической активности (в частности, не удается получить каталитически активный изолированный пресенилин), важным шагом на пути к пониманию структурных основ организации γ -секретазы стала разработка системы экспрессии функциональной γ -секретазы в клетках CHO (Chinese hamster ovary – клетки яичника китайского хомячка) [97], позволяющей получать достаточные для структурных исследований количества γ -секретазы.

Следующим важным шагом на пути решения этой задачи стало развитие метода криоэлектронной микроскопии, который позволил получить структуру функционального γ -секретазного комплекса с субатомным разрешением. Структура была опубликована почти одновременно двумя независимыми группами: (PDB:4UIS) [98] с разрешением 4.4 Å и (PDB:5A63) [99]. В обеих работах для построения структуры пресенилина использовали структуру белка PSN – гомолога пресенилина из архей, совпадающего по последовательности с пресенилином на 19%, – получен-

ную методом РСА. Структура γ -секретазы имеет подковообразную форму. Внеклеточная область γ -секретазы в основном состоит из внеклеточного домена никастрина. Единственная трансмембранная (ТМ) спираль никастрина взаимодействует с ТМ1, ТМ5, ТМ7 Aph-1. Pen-2 содержит три трансмембранные спирали, две из которых не проходят через мембрану насквозь, а соединяются в ее толще. ТМ1 и ТМ3 Pen-2 взаимодействуют с ТМ4 из пресенилина. Основная часть подковы состоит из 16 трансмембранных спиралей пресенилина и Aph-1. Девять трансмембранных спиралей пресенилина значительно различаются по длине: так, ТМ9 содержит 30 остатков, а ТМ7 – только 18. Расстояние между атомами С α двух каталитических остатков пресенилина Asp257 и Asp385 – около 10.6 Å, что значительно больше, чем в других аспаргатных протеазах, но они находятся на разном расстоянии от поверхности мембраны. Каталитический центр расположен рядом с мотивом PAL на ТМ9, который, как считается, играет роль в распознавании субстрата [100]. Вероятно, связывание вызывает выравнивание этих двух остатков аспартата и последующий катализ. Расположение каталитических остатков свидетельствует о том, что белки-субстраты могут получать доступ к активному центру латерально с выпуклой стороны подковы. Мутации в гене пресенилина, ассоциированные с болезнью Альцгеймера, затрагивают 135 аминокислот, из которых на 53 приходится две или более мутаций (<http://www.alzforum.org/mutations>). Среди этих 53 остатков 35 находятся в областях с разрешенной электронной плотностью и структурой. Таких остатков восемь в ТМ5, семь в ТМ6, но ни одного в ТМ1. На эти 35 аминокислот приходится в общей сложности 101 патогенная мутация. Эти остатки группируются в два кластера. Первый находится на внутренней поверхности трансмембранных спиралей ТМ2–ТМ5. Среди 20 затронутых остатков в этих трансмембранных спиральных боковые цепи 18 обращены к внутренней стороне ТМ2–5. То есть затрагивается только одна сторона каждой спирали. Второй кластер расположен на внутренней стороне ТМ6–ТМ9 вблизи каталитических остатков Asp257 и Asp385.

На данный момент кульминацией структурных исследований γ -секретазы стала работа [101], в которой методом криоэлектронной микроскопии с разрешением 2.6 Å определена структура комплекса γ -секретазы с ее субстратом – трансмембранным доменом APP (PDB:6IYC, рис. 3в). Для стабилизации комплекса фермент–субстрат в обе последовательности были введены остатки цистеина: мутация Q112C для PS1 и V695C для APP-C83, а чтобы избежать расщепления субстрата, каталитический остаток Asp385 в PS1 был мутирован на Ala. Так как каталитическая активность необходима для автопротеолиза пресени-

лина, он был экспрессирован в виде N- и C-концевых фрагментов. По сравнению с γ -секретазой, не содержащей субстрата, гибкая трансмембранная спираль 2 (TM2) PS1 упорядочивается при связывании с APP—C83 и участвует в его распознавании, а C-концевой участок TM6 PS1 распадается на петлю и короткую спираль. Трансмембранная спираль APP тесно взаимодействует с пятью окружающими трансмембранными спиралями (TM2, TM3, TM5, TM6 и TM7) пресенилина, в том числе с теми, на которые приходится большая часть патогенных мутаций. C-концевой сегмент APP образует β -тяж, который вместе с двумя индуцированными β -тяжами пресенилина формирует β -лист (рис. 3г). В результате пептидная связь, непосредственно предшествующая β -листу, оказывается подставлена каталитическому сайту γ -секретазы для расщепления.

ЗАКЛЮЧЕНИЕ

Развитие методов изучения пространственной структуры биомолекул позволило полностью описать на молекулярном уровне процесс образования β -амилоидных пептидов, в том числе описать каскад событий, происходящих при последовательном протеолизе белка APP. Полученные структурные данные позволяют выявить ряд факторов, приводящих к развитию болезни Альцгеймера, и составляют практическую основу для разработки эффективных целевых фармакологических соединений для ее эффективной терапии и ранней диагностики.

Работа выполнена при поддержке Российского научного фонда: раздел 1 (Многодоменная структура белка-предшественника бета-амилоида) — грант № 20-64-46027, разделы 2–4 (α -, β -, γ -секретазы) — грант № 19-15-00427.

СПИСОК ЛИТЕРАТУРЫ

1. O'Brien R.J., Wong P.C. // *Annu. Rev. Neurosci.* 2011. V. 34. P. 185. <https://doi.org/10.1146/annurev-neuro-061010-113613>
2. Beel A.J., Mobley C.K., Kim H.J. et al. // *Biochemistry.* 2008. V. 47 (36). P. 9428. <https://doi.org/10.1021/bi800993c>
3. Duce J.A., Tsatsanis A., Cater M.A. et al. // *Cell.* 2010. V. 142 (6). P. 857. <https://doi.org/10.1016/j.cell.2010.08.014>
4. Soscia S.J., Kirby J.E., Washicosky K.J. et al. // *PLoS ONE / Ed. Bush A.I.* 2010. V. 5 (3). P. e9505. <https://doi.org/10.1371/journal.pone.0009505>
5. Small D.H., Nurcombe V., Reed G. et al. // *J. Neurosci. J. Soc. Neurosci.* 1994. V. 14 (4). P. 2117.
6. Rossjohn J., Cappai R., Feil S.C. et al. // *Nat. Struct. Biol.* 1999. V. 6 (4). P. 327. <https://doi.org/10.1038/7562>
7. Dawkins E., Gasperini R., Hu Y. et al. // *J. Neurosci. Res.* 2014. V. 92 (11). P. 1478. <https://doi.org/10.1002/jnr.23422>
8. Baumkötter F., Schmidt N., Vargas C. et al. // *J. Neurosci.* 2014. V. 34 (33). P. 11159. <https://doi.org/10.1523/JNEUROSCI.0180-14.2014>
9. Barnham K.J., McKinsty W.J., Multhaup G. et al. // *J. Biol. Chem.* 2003. V. 278 (19). P. 17401. <https://doi.org/10.1074/jbc.M300629200>
10. Kong G.K.-W., Adams J.J., Harris H.H. et al. // *J. Mol. Biol.* 2007. V. 367 (1). P. 148. <https://doi.org/10.1016/j.jmb.2006.12.041>
11. Dahms S.O., Hoefgen S., Roeser D. et al. // *Proc. Natl. Acad. Sci.* 2010. V. 107 (12). P. 5381. <https://doi.org/10.1073/pnas.0911326107>
12. Baumkötter F., Schmidt N., Vargas C. et al. // *J. Neurosci.* 2014. V. 34 (33). P. 11159. <https://doi.org/10.1523/JNEUROSCI.0180-14.2014>
13. Hoefgen S., Dahms S.O., Oertwig K., Than M.E. // *J. Mol. Biol.* 2015. V. 427 (2). P. 433. <https://doi.org/10.1016/j.jmb.2014.12.005>
14. Hynes T.R., Randal M., Kennedy L.A. et al. // *Biochemistry.* 1990. V. 29 (43). P. 10018. <https://doi.org/10.1021/bi00495a002>
15. Scheidig A.J., Hynes T.R., Pelletier L.A. et al. // *Protein Sci.* 1997. V. 6 (9). P. 1806. <https://doi.org/10.1002/pro.5560060902>
16. Dulubova I., Ho A., Huryeva I. et al. // *Biochemistry.* 2004. V. 43 (30). P. 9583. <https://doi.org/10.1021/bi049041o>
17. Lee S., Xue Y., Hu J. et al. // *Biochemistry.* 2011. V. 50 (24). P. 5453. <https://doi.org/10.1021/bi101846x>
18. Dahms S.O., Könnig I., Roeser D. et al. // *J. Mol. Biol.* 2012. V. 416 (3). P. 438. <https://doi.org/10.1016/j.jmb.2011.12.057>
19. Gralle M., Oliveira C.L.P., Guerreiro L.H. et al. // *J. Mol. Biol.* 2006. V. 357 (2). P. 493. <https://doi.org/10.1016/j.jmb.2005.12.053>
20. Svergun D.I., Petoukhov M.V., Koch M.H. // *Biophys. J.* 2001. V. 80 (6). P. 2946.
21. Konarev P.V., Petoukhov M.V., Svergun D.I. // *J. Appl. Cryst.* 2001. V. 34 (4). P. 527. <https://doi.org/10.1107/S0021889801006100>
22. Ramelot T.A., Gentile L.N., Nicholson L.K. // *Biochemistry.* 2000. V. 39 (10). P. 2714. <https://doi.org/10.1021/bi992580m>
23. Radzimanowski J., Simon B., Sattler M. et al. // *EMBO Rep.* 2008. V. 9 (11). P. 1134. <https://doi.org/10.1038/embor.2008.188>
24. Li H., Koshiba S., Hayashi F. et al. // *J. Biol. Chem.* 2008. V. 283 (40). P. 27165. <https://doi.org/10.1074/jbc.M803892200>
25. Nadezhdin K.D., Bocharova O.V., Bocharov E.V., Arseniev A.S. // *Acta Naturae.* 2011. V. 3 (1). P. 69.
26. Nadezhdin K.D., Bocharova O.V., Bocharov E.V., Arseniev A.S. // *FEBS Lett.* 2012. V. 586 (12). P. 1687. <https://doi.org/10.1016/j.febslet.2012.04.062>
27. Mackenzie K.R. // *Chem. Rev.* 2006. V. 106 (5). P. 1931. <https://doi.org/10.1021/cr0404388>

28. Barrett P.J., Song Y., Van Horn W.D. et al. // Science. 2012. V. 336 (6085). P. 1168.
<https://doi.org/10.1126/science.1219988>
29. Chen W., Gamache E., Rosenman D.J. et al. // Nat. Commun. 2014. V. 5 (1). P. 3037.
<https://doi.org/10.1038/ncomms4037>
30. Bocharov E.V., Nadezhdin K.D., Urban A.S. et al. // ACS Chem. Biol. 2019. V. 14 (1). P. 1573.
<https://doi.org/10.1021/acscchembio.9b00309>
31. Hardy J., Allsop D. // Trends Pharmacol. Sci. 1991. V. 12 (10). P. 383.
32. Sticht H., Bayer P., Willbold D. et al. // Eur. J. Biochem. 1995. V. 233 (1). P. 293.
https://doi.org/10.1111/j.1432-1033.1995.293_1.x
33. Fraga A.S., Esteves A.C., Micaelo N. et al. // Int. J. Biol. Macromol. 2012. V. 50 (2). P. 323.
<https://doi.org/10.1016/j.ijbiomac.2011.12.018>
34. Coles M., Bicknell W., Watson A.A. et al. // Biochemistry. 1998. V. 37 (31). P. 11064.
<https://doi.org/10.1021/bi972979f>
35. Vivekanandan S., Brender J.R., Lee S.Y., Ramamoorthy A. // Biochem. Biophys. Res. Commun. 2011. V. 411 (2). P. 312.
<https://doi.org/10.1016/j.bbrc.2011.06.133>
36. Fezoui Y., Teplow D.B. // J. Biol. Chem. 2002. V. 277 (40). P. 36948.
<https://doi.org/10.1074/jbc.M204168200>
37. Fonar G., Samson A.O. // Biosci. Rep. 2014. V. 34 (6). P. e00155.
<https://doi.org/10.1042/BSR20140094>
38. Tomaselli S., Esposito V., Vangone P. et al. // Chem. Bio. Chem. 2006. V. 7 (2). P. 257.
<https://doi.org/10.1002/cbic.200500223>
39. Lühns T., Ritter C., Adrian M. et al. // Proc. Natl. Acad. Sci. 2005. V. 102 (48). P. 17342.
<https://doi.org/10.1073/pnas.0506723102>
40. Paravastu A.K., Leapman R.D., Yau W.-M., Tycko R. // Proc. Natl. Acad. Sci. 2008. V. 105 (47). P. 18349.
<https://doi.org/10.1073/pnas.0806270105>
41. Lu J.-X., Qiang W., Yau W.-M. et al. // Cell. 2013. V. 154 (6). P. 1257.
<https://doi.org/10.1016/j.cell.2013.08.035>
42. Kodali R., Williams A.D., Chemuru S., Wetzel R. // J. Mol. Biol. 2010. V. 401 (3). P. 503.
<https://doi.org/10.1016/j.jmb.2010.06.023>
43. Zhao L.N., Long H., Mu Y., Chew L.Y. // Int. J. Mol. Sci. 2012. V. 13 (6). P. 7303.
<https://doi.org/10.3390/ijms13067303>
44. Sengupta U., Nilson A.N., Kaye R. // E Bio. Medicine. 2016. V. 6. P. 42.
<https://doi.org/10.1016/j.ebiom.2016.03.035>
45. Rauk A. // Dalton Trans. Camb. Engl. 2008. № 10. P. 1273.
<https://doi.org/10.1039/b718601k>
46. Shafirir Y., Durell S., Arispe N., Guy H.R. // Proteins. 2010. V. 78 (16). P. 3473.
<https://doi.org/10.1002/prot.22853>
47. Bode D.C., Baker M.D., Viles J.H. // J. Biol. Chem. 2017. V. 292 (4). P. 1404.
<https://doi.org/10.1074/jbc.M116.762526>
48. Yu Y., Yin T., Peng Q. et al. // ACS Sens. 2019. V. 4 (2). P. 471.
<https://doi.org/10.1021/acssensors.8b01493>
49. Shea D., Hsu C.-C., Bi T.M. et al. // Proc. Natl. Acad. Sci. 2019. V. 116 (18). P. 8895.
<https://doi.org/10.1073/pnas.1820585116>
50. Kreutzer A.G., Hamza I.L., Spencer R.K., Nowick J.S. // J. Am. Chem. Soc. 2016. V. 138 (13). P. 4634.
<https://doi.org/10.1021/jacs.6b01332>
51. Kawai R., Chiba S., Okuwaki K. et al. // ACS Chem. Neurosci. 2020. V. 11 (3). P. 385.
<https://doi.org/10.1021/acscchemneuro.9b00602>
52. Lesne S.E. // Swiss Med. Wkly. 2014. V. 144. № 4546. P. w14021.
<https://doi.org/10.4414/smw.2014.14021>
53. Ono K., Yamada M. // J. Neurochem. 2011. V. 117 (1). P. 19.
<https://doi.org/10.1111/j.1471-4159.2011.07187.x>
54. Ciudad S., Puig E., Botzanowski T. et al. // Nat. Commun. 2020. V. 11 (1). P. 3014.
<https://doi.org/10.1038/s41467-020-16566-1>
55. Weber S., Saftig P. // Development. 2012. V. 139 (20). P. 3693.
<https://doi.org/10.1242/dev.076398>
56. Asai M., Hattori C., Szabó B. et al. // Biochem. Biophys. Res. Commun. 2003. V. 301 (1). P. 231.
[https://doi.org/10.1016/s0006-291x\(02\)02999-6](https://doi.org/10.1016/s0006-291x(02)02999-6)
57. Kuhn P.-H., Wang H., Dislich B. et al. // EMBO J. 2010. V. 29 (17). P. 3020.
<https://doi.org/10.1038/emboj.2010.167>
58. Kovall R.A., Gebelein B., Sprinzak D., Kopan R. // Dev. Cell. 2017. V. 41 (3). P. 228.
<https://doi.org/10.1016/j.devcel.2017.04.001>
59. Janes P.W., Saha N., Barton W.A. et al. // Cell. 2005. V. 123 (2). P. 291.
<https://doi.org/10.1016/j.cell.2005.08.014>
60. Arduise C., Abache T., Li L. et al. // J. Immunol. 2008. V. 181 (10). P. 7002.
<https://doi.org/10.4049/jimmunol.181.10.7002>
61. Feldinger K., Generali D., Kramer-Marek G. et al. // Oncotarget. 2014. V. 5 (16). P. 6633.
<https://doi.org/10.18632/oncotarget.1955>
62. Mezyk-Kopeć R., Bzowska M., Stalińska K. et al. // Cytokine. 2009. V. 46 (3). P. 309.
<https://doi.org/10.1016/j.cyto.2009.03.002>
63. Anders A., Gilbert S., Garten W. et al. // FASEB J. 2001. V. 15 (10). P. 1837.
<https://doi.org/10.1096/fj.01-0007fje>
64. Moss M.L., Bomar M., Liu Q. et al. // J. Biol. Chem. 2007. V. 282 (49). P. 35712.
<https://doi.org/10.1074/jbc.M703231200>
65. Vassar R. // Neuron. 2013. V. 80 (2). P. 250.
<https://doi.org/10.1016/j.neuron.2013.09.031>
66. Fujii Y., Okuda D., Fujimoto Z. et al. // J. Mol. Biol. 2003. V. 332 (5). P. 1115.
[https://doi.org/10.1016/S0022-2836\(03\)00991-4](https://doi.org/10.1016/S0022-2836(03)00991-4)
67. Senn H., Klaus W. // J. Mol. Biol. 1993. V. 232 (3). P. 907.
<https://doi.org/10.1006/jmbi.1993.1439>

68. *Seegar T.C.M., Killingsworth L.B., Saha N. et al.* // Cell. 2017. V. 171 (7). P. 1638.
<https://doi.org/10.1016/j.cell.2017.11.014>
69. *Fahrenholz F., Gilbert S., Kojro E. et al.* // Ann. N. Y. Acad. Sci. 2000. V. 920. P. 215.
70. *Atapattu L., Saha N., Chheang C. et al.* // J. Exp. Med. 2016. V. 213 (9). P. 1741.
<https://doi.org/10.1084/jem.20151095>
71. *Ahmed R.R., Holler C.J., Webb R.L. et al.* // J. Neurochem. 2010. V. 112 (4). P. 1045.
<https://doi.org/10.1111/j.1471-4159.2009.06528.x>
72. *Wang Z., Xu Q., Cai F. et al.* // JCI Insight. 2019. V. 4. № 1. P. e123431.
<https://doi.org/10.1172/jci.insight.123431>
73. *Willem M., Garratt A.N., Novak B. et al.* // Science. 2006. V. 314 (5799). P. 664.
<https://doi.org/10.1126/science.1132341>
74. *Hu X., Hicks C.W., He W. et al.* // Nat. Neurosci. 2006. V. 9 (12). P. 1520.
<https://doi.org/10.1038/nn1797>
75. *Hu X., He W., Luo X. et al.* // Cell Rep. 2013. V. 4 (1). P. 40.
<https://doi.org/10.1016/j.celrep.2013.06.005>
76. *Kim D.Y., Carey B.W., Wang H. et al.* // Nat. Cell. Biol. 2007. V. 9 (7). P. 755.
<https://doi.org/10.1038/ncb1602>
77. *Sachse C.C., Kim Y.H., Agsten M. et al.* // FASEB J. 2013. V. 27 (6). P. 2458.
<https://doi.org/10.1096/fj.12-214056>
78. *Kuhn P.-H., Koroniak K., Hogg S. et al.* // EMBO J. 2012. V. 31 (14). P. 3157.
<https://doi.org/10.1038/emboj.2012.173>
79. *Hemming M.L., Elias J.E., Gygi S.P., Selkoe D.J.* // PloS One. 2009. V. 4 (12). P. e8477.
<https://doi.org/10.1371/journal.pone.0008477>
80. *Venugopal C., Demos C.M., Rao K.S.J. et al.* // CNS Neurol. Disord. Drug. Targets. 2008. V. 7 (3). P. 278.
81. *Shi X.-P., Chen E., Yin K.-C. et al.* // J. Biol. Chem. 2001. V. 276 (13). P. 10366.
<https://doi.org/10.1074/jbc.M009200200>
82. *Yan R., Han P., Miao H. et al.* // J. Biol. Chem. 2001. V. 276 (39). P. 36788.
<https://doi.org/10.1074/jbc.M104350200>
83. *Hong L.* // Science. 2000. V. 290 (5489). P. 150.
<https://doi.org/10.1126/science.290.5489.150>
84. *Ostermann N., Eder J., Eidhoff U. et al.* // J. Mol. Biol. 2006. V. 355 (2). P. 249.
<https://doi.org/10.1016/j.jmb.2005.10.027>
85. *Hong L., Tang J.* // Biochemistry. 2004. V. 43 (16). P. 4689.
<https://doi.org/10.1021/bi0498252>
86. *Lin X., Koelsch G., Wu S. et al.* // Proc. Natl. Acad. Sci. 2000. V. 97 (4). P. 1456.
<https://doi.org/10.1073/pnas.97.4.1456>
87. *Bittner H.J., Guixà-González R., Hildebrand P.W.* // Biochim. Biophys. Acta BBA – Biomembr. 2018. V. 1860 (5). P. 1105.
<https://doi.org/10.1016/j.bbamem.2018.01.019>
88. *Zhang X., Li Y., Xu H., Zhang Y.* // Front. Cell. Neurosci. 2014. V. 8. P. 427.
<https://doi.org/10.3389/fncel.2014.00427>
89. *Kopan R., Ilagan M.X.G.* // Nat. Rev. Mol. Cell. Biol. 2004. V. 5 (6). P. 499.
<https://doi.org/10.1038/nrm1406>
90. *Haapasalo A., Kovacs D.M.* // J. Alzheimers Dis. JAD. 2011. V. 25 (1). P. 3.
<https://doi.org/10.3233/JAD-2011-101065>
91. *Xu X.* // J. Alzheimers Dis. JAD. 2009. V. 16 (2). P. 211.
<https://doi.org/10.3233/JAD-2009-0957>
92. *Shah S., Lee S.-F., Tabuchi K. et al.* // Cell. 2005. V. 122 (3). P. 435.
<https://doi.org/10.1016/j.cell.2005.05.022>
93. *Shea Y.-F., Chu L.-W., Chan A.O.-K. et al.* // J. Formos. Med. Assoc. 2016. V. 115 (2). P. 67.
<https://doi.org/10.1016/j.jfma.2015.08.004>
94. *De Strooper B.* // Neuron. 2003. V. 38 (1). P. 9.
[https://doi.org/10.1016/s0896-6273\(03\)00205-8](https://doi.org/10.1016/s0896-6273(03)00205-8)
95. *Shah S., Lee S.-F., Tabuchi K. et al.* // Cell. 2005. V. 122 (3). P. 435.
<https://doi.org/10.1016/j.cell.2005.05.022>
96. *Dries D.R., Yu G.* // Curr. Alzheimer Res. 2008. V. 5 (2). P. 132.
<https://doi.org/10.2174/156720508783954695>
97. *Alattia J.-R., Matasci M., Dimitrov M. et al.* // Biotechnol. Bioeng. 2013. V. 110 (7). P. 1995.
<https://doi.org/10.1002/bit.24851>
98. *Sun L., Zhao L., Yang G. et al.* // Proc. Natl. Acad. Sci. 2015. V. 112 (19). P. 6003.
<https://doi.org/10.1073/pnas.1506242112>
99. *Bai X., Yan C., Yang G. et al.* // Nature. 2015. V. 525 (7568). P. 212.
<https://doi.org/10.1038/nature14892>
100. *Sato C., Takagi S., Tomita T., Iwatsubo T.* // J. Neurosci. 2008. V. 28. № 24. P. 6264.
<https://doi.org/10.1523/JNEUROSCI.1163-08.2008>
101. *Zhou R., Yang G., Guo X. et al.* // Science. 2019. V. 363 (6428). P. eaaw0930.
<https://doi.org/10.1126/science.aaw0930>

ДИФРАКЦИЯ И РАССЕЯНИЕ ИОНИЗИРУЮЩИХ ИЗЛУЧЕНИЙ

УДК 548.5, 548.73, 544.77.023.5

ВЛИЯНИЕ ХЛОРИДОВ ОДНО- И ДВУХВАЛЕНТНЫХ МЕТАЛЛОВ НА ОЛИГОМЕРНЫЙ СОСТАВ КРИСТАЛЛИЗАЦИОННЫХ РАСТВОРОВ ЛИЗОЦИМА И ДАЛЬНЕЙШИЙ РОСТ КРИСТАЛЛОВ

© 2021 г. М. А. Марченкова^{1,2}, П. В. Конарев^{1,2}, А. С. Бойкова^{1,2,*}, К. Б. Ильина^{1,2},
Ю. В. Писаревский^{1,2}, М. В. Ковальчук^{1,2}

¹ Институт кристаллографии им. А.В. Шубникова ФНИЦ “Кристаллография и фотоника” РАН, Москва, Россия

² Национальный исследовательский центр “Курчатовский институт”, Москва, Россия

*E-mail: boykova.irk@yandex.ru

Поступила в редакцию 03.09.2020 г.

После доработки 23.10.2020 г.

Принята к публикации 26.11.2020 г.

Методом малоуглового рассеяния рентгеновских лучей исследовано влияние типа осадителя (LiCl, NaCl, KCl, NiCl₂ и CuCl₂) на образование олигомеров (димеров и октамеров) в кристаллизационных растворах лизоцима при двух концентрациях белка. Из этих же растворов выращены кристаллы для установления влияния олигомерного состава на рост кристаллов. На основе данных, полученных в настоящей и предыдущих работах, описывающих влияние концентрации осадителя, продемонстрирована обратно пропорциональная зависимость суммарного содержания октамеров и димеров от порядкового номера катиона, что согласуется с увеличением активности ионов в лиотропном ряду для Li⁺, Na⁺, K⁺ и увеличением ионных радиусов для Li⁺, Na⁺, K⁺ и для Ni²⁺, Cu²⁺. Показано, что уменьшение концентрации белка в кристаллизационном растворе ведет к уменьшению объемной доли октамеров при неизменяющейся объемной доли димеров и понижению вероятности появления кристаллов.

DOI: 10.31857/S0023476121050131

ВВЕДЕНИЕ

Изучение пространственной структуры белков, функционирующих в живых организмах, позволяет разобраться, каким образом биологические молекулы способны выполнять свои функции в природе.

На сегодняшний день около 90% пространственных структур макромолекул, депонированных в белковый банк данных (Protein Data Bank, PDB), определено методом рентгеноструктурного анализа (РСА). Однако ограничением данного метода является требование монокристалличности образца. Поэтому исследование основных механизмов кристаллизации биологических макромолекул представляет фундаментальный интерес как для развития кристаллографии, так и для оптимизации подбора условий кристаллизации и сокращения времени роста кристаллов.

В последние годы обозначился переход от классической схемы кристаллообразования к двухступенчатой, включающей в себя образование промежуточных кластеров-прекурсоров. Одними из первых предположения такого рода высказали сотрудники Института кристаллографии РАН профессор Н.Н. Шефтал в 1957 г. при анализе кристаллизации из газовой фазы и Р.О. Гриз-

дейл в 1968 г. при анализе кристаллизации из растворов [1–3]. Активное развитие этого подхода началось в XXI веке. К настоящему времени имеются обширные данные по этой тематике [4–6]. Особое место занимает цикл работ, выполненный в ФНИЦ “Кристаллография и фотоника” РАН под руководством М.В. Ковальчука, в котором методами малоуглового рассеяния рентгеновских лучей (МУРР) и нейтронов (МУРН) впервые экспериментально обнаружены предкристаллизационные кластеры-прекурсоры в кристаллизационных растворах нескольких белков (лизоцим [7, 8], протеиназа [9], термолизин [10] и аминотрансфераза [11]) и кристаллизационных растворах неорганического соединения дигидрофосфата калия [12]. Для модельного белка лизоцима показано, что в условиях роста кристаллов лизоцима тетрагональной сингонии в растворе образуются олигомерные частицы белка — димеры и октамеры [7, 8], последние являются кластерами-прекурсорами кристалла. При добавлении к раствору лизоцима таких осадителей, как LiCl, NaCl, KCl, CoCl₂, NiCl₂ и CuCl₂, приводящих к росту кристаллов тетрагональной сингонии, объемная доля октамеров увеличива-

ется в следующем порядке: для одновалентных ионов: $K^+ - Na^+ - Li^+$, для двухвалентных: $Cu^{2+} - Ni^{2+} - Co^{2+}$ [13]. Наличие таких лиотропных рядов (или рядов Гофмейстера, которые располагают ионы по силе их воздействия на различные свойства, в случае белка – его растворимость и стабильность) обусловлено их влиянием на какие-либо параметры исследуемой системы [14].

В настоящей работе продолжено изучение структуры растворов лизоцима в условиях роста кристаллов тетрагональной сингонии при добавлении осадителей $NaCl$, KCl , $LiCl$, $NiCl_2$ и $CuCl_2$ в зависимости от концентрации белка, осадителя и температуры на станциях P12 EMBL BioSAXS (DESY, Гамбург, Германия) и BM29 BioSAXS (ESRF, Гренобль, Франция).

МАТЕРИАЛЫ И МЕТОДЫ

Материалы и подготовка образцов. Для приготовления образцов использовали белок лизоцим из куриного яйца производства Sigma-Aldrich (CAS № 12650-88-3). Для приготовления маточных растворов использовали следующие неорганические соли: $NaCl$ (CAS № 7647-14-5, Helicon), KCl (CAS № 7447-40-7, abcr GmbH), $LiCl$ (ТУ 6-09-3751-83, Лаверна Стройинжиниринг), $CoCl_2$ (CAS № 7791-13-1, Alta Aesar), $NiCl_2$ (CAS № 7791-20-0, Alta Aesar) и $CuCl_2$ (CAS № 7447-39-4, Acros Organics). Все растворы были приготовлены с использованием ультрачистой воды Millipore (сопротивление воды 18 МОм·см). Белок и соли растворяли в 0.2 М натрий-ацетатном буфере, pH 4.5. Растворы солей фильтровали с помощью мембранных шприцевых фильтров Millex с размером пор 0.22 мкм, раствор белка центрифугировали в течение 10 мин с частотой 10000 об./мин. Начальная концентрация в маточном растворе белка – 80 мг/мл, начальные концентрации всех солей в маточных растворах – 0.8 и 0.4 М.

Методика МУРР-измерений. Перед проведением измерений методом МУРР маточные растворы лизоцима и солей смешивали друг с другом в равных объемах. Эксперименты были проведены на станциях P12 EMBL BioSAXS источника синхротронного излучения PETRA III (DESY, Гамбург, Германия) и BM29 BioSAXS Европейского источника синхротронного излучения (Гренобль, Франция).

Описание эксперимента на станции P12 EMBL BioSAXS (DESY, Гамбург, Германия). Энергия составляла 10 кэВ ($\lambda = 0.124$ нм), в качестве детектора сигнала использовали двухкоординатный детектор PILATUS 6M, позволяющий проводить регистрацию относительно слабых сигналов рассеяния. Расстояние образец–детектор составляло 3.0 м, данные МУРР записывали в диапазоне величин вектора обратного рассеяния $s = 0.02$ –

7.0 нм⁻¹, что соответствует разрешению 300–0.9 нм в реальном пространстве. Измерения проводили с использованием специализированной ячейки для образцов МУРР, состоящей из горизонтального термостатируемого в диапазоне от 278 до 323 К кварцевого капилляра со стенками толщиной 50 мкм и диаметром 1.7 мм, размещенного в специализированном корпусе из нержавеющей стали. Время экспозиции составляло 50 мс, было сделано 20 снимков для каждого измерения образца. Более детальное описание станции приведено в [15]. Объем образца в каждом измерении составлял 40 мкл. Измерения проводили при температуре 20°C.

Описание эксперимента на станции BM 29 BioSAXS (ESRF, Гренобль, Франция). Энергия составляла 12.4 кэВ, в качестве детектора сигнала использовали двухкоординатный детектор Pilatus 1M. Расстояние образец–детектор составляло 2.9 м. Исследуемые образцы помещали в специальную термостатируемую роботизированную систему в кюветы из полистирола объемом 200 мкл, нагрев которых осуществлялся одновременно. Первоначально образцы нагревали до 20°C, затем температуру понижали до 10°C. Далее раствор из кюветы с помощью робота помещали в проточный кварцевый капилляр диаметром 1.8 мм, который использовался при измерениях [16]. Исследуемый раствор равномерно продвигался по капилляру, при этом пучок попадал в одну и ту же точку на капилляре, но все время в новую часть образца. За время движения образца по капилляру было сделано 10 снимков. Время экспозиции каждого измерения составляло 1 с, сечение пучка на образце – 700 мкм². Объем образца в каждом измерении составлял 50 мкл.

Методика обработки экспериментальных данных. Усреднение сигнала от буферного раствора, вычитание усредненного сигнала от буфера из экспериментальных данных рассеяния раствором белков и нормировку на концентрацию белка выполняли с помощью программы PRIMUS, входящей в программный пакет ATSAS [17, 18]. В результате получены экспериментальные кривые интенсивности I от модуля вектора рассеяния s (где $s = 4\pi \sin \theta / \lambda$, 2θ – угол рассеяния, λ – длина волны) для растворов белка в различных условиях. Угловой диапазон составлял $0.03 < s < 5.0$ нм⁻¹. При сравнении последовательных кадров радиационного повреждения на исследуемых образцах не обнаружено. После первичной обработки экспериментальные кривые малоуглового рассеяния обрабатывали с помощью программы OLIGOMER [18] для определения объемных долей мономеров и олигомеров разного порядка. Расчет теоретических кривых олигомерных компонентов проводили с помощью программы CRY SOL [19]. В качестве мономерного компо-

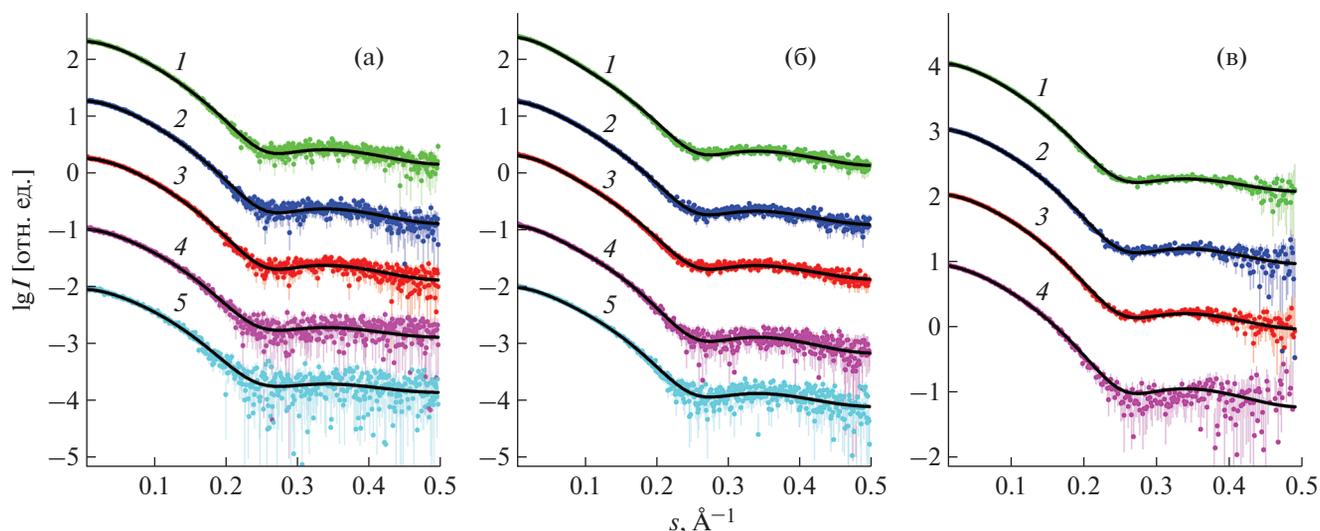


Рис. 1. Экспериментальные кривые МУПП от раствора лизоцима с добавлением осадителей LiCl (1), KCl (2), NaCl (3), NiCl₂ (4) и CuCl₂ (5) и теоретические приближения смесью олигомеров, рассчитанные программой OLIGOMER (черные линии) для концентрации осадителей и концентраций лизоцима соответственно: а – 0.4 М, 20 мг/мл, P12 (EMBL, Гамбург, Германия); б – 0.4 М, 40 мг/мл, P12 (EMBL, Гамбург, Германия); в – 0.43 М, 20 мг/мл, BM 29 BioSAXS (ESRF, Гренобль, Франция). Кривые смещены по вертикали для лучшей визуализации.

нента была взята кристаллографическая структура мономера лизоцима (PDB ID: 4WLD), а модели димера, тетрамера, гексамера и октамера получены по методике, описанной в [7]. Качество приближения оценивали с помощью минимизации невязки χ^2 между экспериментальными данными и теоретическими модельными приближениями по формуле, приведенной в [13].

Кристаллизация. Маточные растворы, приготовленные для измерений методом МУПП на станции P12 EMBL BioSAXS (DESY, Гамбург, Германия), использовали также для кристаллизации (*материалы и методы*). Кристаллизацию осуществляли методом диффузии в парах в варианте сидячей капли [20] с помощью кристаллизационного робота Mosquito-LCP (EMBL, Гамбург, Германия), объем капли составлял 200 нл (100 нл маточного раствора белка + 100 нл маточного раствора осадителя). Рост кристаллов проводился в автоматизированной системе визуализации ROCK IMAGER при температуре 19°C. Система позволяет наблюдать рост кристаллов белков и фотографировать капли в течение длительного времени несколько раз (в “нулевой” день (сразу после загрузки кристаллизационного планшета) и далее в 1, 3, 7, 14, 28, 56 и 84 день). Осадители использовали те же (NaCl, KCl, LiCl, NiCl₂ и CuCl₂). Конечные концентрации в капле составляли для лизоцима 40 и 20 мг/мл, для осадителя – 0.4 М.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Результаты моделирования данных МУПП и тенденции изменения объемной доли олигомеров.

В серии экспериментов, выполненных на станциях малоуглового рассеяния P12 (EMBL, Гамбург) и BM 29 BioSAXS (ESRF, Гренобль), были проведены измерения растворов при концентрации лизоцима 20 и 40 мг/мл и концентрации осадителей 0.4 и 0.2 М при температуре 20°C. В качестве осадителей использовали неорганические соли – хлориды щелочных (NaCl, KCl и LiCl) и переходных металлов (NiCl₂ и CuCl₂). Для сравнения проведены измерения растворов лизоцима без осадителей. Экспериментальные и теоретические кривые, рассчитанные с помощью программы OLIGOMER, показаны на рис. 1.

Каждая комбинация типа осадителя и концентрации белка была измерена несколько раз на станциях P12 EMBL BioSAXS (DESY) и BM 29 BioSAXS (ESRF); в табл. 1 приведены усредненные результаты обработки экспериментальных данных.

Рассчитанные кривые МУПП от олигомерных смесей для растворов белка с осадителем хорошо совпадают с экспериментальными данными во всем угловом диапазоне, значения невязки χ^2 не превышают 1.6, что свидетельствует о правильности предложенной модели обработки. Отметим, что данные, полученные в одном эксперименте при повторных измерениях, различаются незначительно (в некоторых случаях на величину погрешности обработки).

Исследования объемной доли олигомеров в кристаллизационных растворах лизоцима при одинаковых условиях (концентрация белка 20 мг/мл и концентрация осадителей 0.4 М), проведенные

Таблица 1. Средние значения объемных долей димеров (Д) и октамеров (О) (%) лизоцима в растворах с осадителями LiCl, NaCl, KCl, CuCl₂ и NiCl₂, полученные в экспериментах на станциях P12 EMBL BioSAXS (DESY, Hamburg) и BM 29 BioSAXS (ESRF, Grenoble)

	C _{ос} , М	C _б , МГ/МЛ	T, °С	LiCl		NaCl		KCl		NiCl ₂		CuCl ₂	
				О	Д	О	Д	О	Д	О	Д	О	Д
DESY*	0.4	40	20	6.70	17.75	5.15	16.25	4.65	14.65	5.55	8.85	4.05	6.4
DESY*	0.4	20	20	2.45	17.00	2.15	14.1	2.25	13.65	2.6	7.4	1.95	7.55
ESRF**	0.2	20	20	1.13	8.6	0.875	5.7	0.43	3.8	1.43	3.1	0.7	0.5
ESRF [13]	0.2	20	20	1.70	7.3	1.8	5.4	1.9	5.0	2.4	1.9	1.1	0
ESRF [13]	0.2	20	10	1.9	13.1	2.2	10.6	2.4	10.5	2.8	7.9	1.7	3.5
ESRF [13]	0.1	20	20	0.1	6.2	0.1	4.9	0.3	4.5	1.0	0	0	0
ESRF [13]	0.1	20	10	0.2	10.5	0.1	10.1	0.3	9.3	1.4	3.7	0	0.1
ESRF ***	0.43	20	20	2	12.6	1.86	10.3	2.27	9.7	2.46	9.83		

Примечание. Одним цветом отмечены эксперименты при одинаковых условиях; C_{ос}, C_б – концентрации осадителя и белка соответственно.

* Среднее по двум экспериментам (округления нет).

** Среднее по трем экспериментам.

*** Среднее по времени (три эксперимента).

на разных источниках синхротронного излучения, показали, что усредненные по нескольким измерениям объемные доли олигомеров отличаются друг от друга в относительных пределах 1–15% (наибольшее различие в 20% с осадителем LiCl) для октамеров и 25–30% для димеров. Кроме того, при выполнении повторных экспериментов на одной и той же станции объемные доли димеров при определенных условиях могли различаться в абсолютных пределах на 1.2–1.5%, а для октамеров – в пределах 0.3–0.5%, с учетом данной неоднозначности, можно заключить, что результаты, представленные в табл. 3, измеренные при различных условиях, находятся в согласии друг с другом. Однако объемная доля тетрамеров и гексамеров во всех экспериментах равняется 0%. Разброс значений между экспериментами на разных станциях и между повторными экспериментами в одной серии могут быть вызваны следующими факторами: неизбежными различиями в протоколе приготовления (временной период между добавлением осадителя и фактическим измерением), разной степенью ослабления пучка во избежание радиационного повреждения образца, изначально разной интенсивностью падающего пучка, разной степенью “чистоты” буферного раствора и наличием “примесей”, стабильностью пучка на самих экспериментальных станциях и эффективностью заполнения образцов роботом.

Октамеры в среднем демонстрируют более стабильное поведение (их объемные доли при сравнении с данными, полученными в различных сериях экспериментов на различных источниках, слабо меняются по сравнению с содержанием димеров), в то время как димеры, судя по всему, являются более нестабильными образованиями, и

их содержание может значительно меняться в ходе проведения различных экспериментов.

Во всех случаях объемная доля октамеров максимальна для NiCl₂ и минимальна для CuCl₂. Исключение составляет LiCl (концентрация осадителя 0.4 М, белка 40 мг/мл), где объемная доля сильно повысилась в сравнении с условиями для NiCl₂.

Зависимость объемной доли октамеров от порядкового номера Li, Na, K при комнатной температуре различается для разных измерений (табл. 2): для данных ESRF (строки 4–7) она прямо пропорциональна порядковому номеру элемента, для данных DESY (строки 1, 2) и ESRF (строка 3) – обратно пропорциональна.

Для димеров во всех измерениях (кроме строк 2 и 8) их объемная доля увеличивается при использовании разных осадителей в следующем порядке (от наименьшего к наибольшему): Cu²⁺, Ni²⁺, K⁺, Na⁺, Li⁺.

Объемная доля димеров и октамеров увеличивается для такой же, как в случае с димерами, последовательности ионов Cu²⁺, Ni²⁺, K⁺, Na⁺, Li⁺ или уменьшается в ряду Li⁺, Na⁺, K⁺, Ni²⁺, Cu²⁺, что согласуется с увеличением активности ионов в лиотропном ряду Гофмейстера для Li⁺, Na⁺, K⁺ и увеличением ионных радиусов для Li⁺, Na⁺, K⁺ и для Ni²⁺, Cu²⁺.

Временные измерения МУРР. Для условий кристаллизации, где концентрация белка составляла 20 мг/мл, а концентрация осадителей LiCl, NaCl, KCl, NiCl₂ – 0.43 М, на станции BM29 (ESRF, Гренобль, Франция) было проведено исследование в промежутках времени 0* мин (*0 мин в дан-

Таблица 2. Тенденция изменения объемной доли олигомеров в кристаллизационном растворе лизоцима с осадителями – хлоридами металлов

	$C_{ос}, M$	$C_6,$ мг/мл	$T, ^\circ C$	Изменение объемной доли октамеров в ряду от меньшего содержания к большему	Изменение объемной доли димеров в ряду от меньшего содержания к большему	Изменение объемной доли октамеров и димеров в ряду от меньшего содержания к большему
DESY*	0.4	40	20	$Cu^{2+} < K^+ < Na^+ < Ni^{2+} < Li^+$	$Cu^{2+} < Ni^{2+} < K^+ < Na^+ < Li^+$	$Cu^{2+} < Ni^{2+} < K^+ < Na^+ < Li^+$
DESY*	0.4	20	20	$Cu^{2+} < Na^+ < K^+ < Li^+ < Ni^{2+}$	$Ni^{2+} < Cu^{2+} < K^+ < Na^+ < Li^+$	$Cu^{2+} < Ni^{2+} < K^+ < Na^+ < Li^+$
ESRF**	0.2	20	20	$K^+ < Cu^{2+} < Na^+ < Li^+ < Ni^{2+}$	$Cu^{2+} < Ni^{2+} < K^+ < Na^+ < Li^+$	$Cu^{2+} < Ni^{2+} \approx K^+ < Na^+ < Li^+$
ESRF [13]	0.2	20	20	$Cu^{2+} < Li^+ < Na^+ < K^+ < Ni^{2+}$	$Cu^{2+} < Ni^{2+} < K^+ < Na^+ < Li^+$	$Cu^{2+} < Ni^{2+} < K^+ < Na^+ < Li^+$
ESRF [13]	0.2	20	10	$Cu^{2+} < Li^+ < Na^+ < K^+ < Ni^{2+}$	$Cu^{2+} < Ni^{2+} < K^+ < Na^+ < Li^+$	$Cu^{2+} < Ni^{2+} < K^+ \approx Na^+ < Li^+$
ESRF [13]	0.1	20	20	$Cu^{2+} < Li^+ = Na^+ < K^+ < Ni^{2+}$	$Cu^{2+} = Ni^{2+} < K^+ < Na^+ < Li^+$	$Cu^{2+} < Ni^{2+} < K^+ < Na^+ < Li^+$
ESRF [13]	0.1	20	10	$Cu^{2+} < Li^+ \approx Na^+ \approx K^+ < Ni^{2+}$	$Cu^{2+} < Ni^{2+} < K^+ < Na^+ < Li^+$	$Cu^{2+} < Ni^{2+} < K^+ < Na^+ < Li^+$
ESRF ***	0.43	20	20	$Na^+ < Li^+ < K^+ < Ni^{2+}$	$K^+ < Na^+ < Ni^{2+} < Li^+$	$K^+ < Na^+ < Ni^{2+} < Li^+$

Примечание. Одним цветом отмечены эксперименты при одинаковых условиях.

* Среднее по двум экспериментам (округления нет).

** Среднее по трем экспериментам.

*** Среднее по времени (три эксперимента).

Таблица 3. Объемная доля олигомерных фракций (димеров и октамеров), радиус инерции R_g и значение невязки χ^2 для растворов лизоцима при концентрации белка 20 мг/мл и концентрации осадителей 0.43 M

Осади- тель	0 мин				100 мин				170 мин			
	$R_g, \text{Å}$	Д, %	О, %	χ^2	$R_g, \text{Å}$	Д, %	О, %	χ^2	$R_g, \text{Å}$	Д, %	О, %	χ^2
NaCl	18.6	10.1	1.9	0.98	18.5	10.3	1.8	0.91	18.6	10.6	1.9	0.87
KCl [21]	19.1	9.4	2.3	0.92	18.9	10.0	2.2	0.92	19.1	9.7	2.3	0.90
LiCl	18.9	12.0	2.1	1.06	18.8	12.8	2.0	1.18	18.8	13.0	2.0	1.13
NiCl ₂	19.1	9.9	2.3	1.01	19.4	8.9	2.6	1.21	19.4	10.7	2.5	1.30

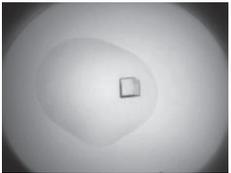
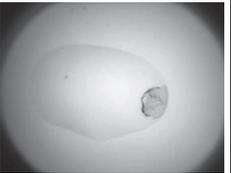
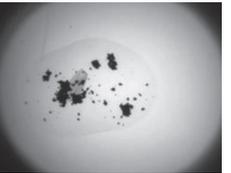
ном случае означает фактически время начала измерения, а не смешивания; от смешивания раствора белка и раствора осадителя до начала измерений проходило ~10–15 мин), 100 и 170 мин при температуре 20°C. Результаты обработки экспериментальных данных приведены в табл. 3.

С течением времени (от 0 до 170 мин) после смешения содержание октамеров и димеров в растворе изменяется незначительно. Это позволяет убрать вклад времени между добавлением осадителя к белку и фактическим измерением раствора из существенных факторов, влияющих на разброс значений содержания олигомеров между экспериментами.

Рост кристаллов в исследованных условиях. Рост кристаллов проводили из тех же маточных растворов, приготовленных для исследования методом МУРР на станции P12 (EMBL, Гамбург, Германия). Для каждого типа осадителя была поставлена кристаллизация в трех каплях. В подписях к рисункам указано, в скольких опытах образовались кристаллы лизоцима, также приведены средние объемные доли октамеров в растворах по данным МУРР.

В табл. 4 приведены фотографии кристаллов, выращенных в условиях с концентрацией белка 40 и 20 мг/мл. Фотографии сделаны по истечении 56 дней (исключение составляют фотографии кристалла с осадителями KCl и CuCl₂, сделанные

Таблица 4. Фотографии кристаллов лизоцима, выращенных из исследуемых растворов при температуре 20°C, концентрациях белка 40 и 20 мг/мл и осадителей 0.4 М

C_6 , мг/мл	LiCl	NaCl	KCl	NiCl ₂	CuCl ₂
40	 3/3, 6.70%	 3/3, 5.15%	 3/3, 4.65%	 3/3, 5.55%	 3/3, 4.05%
20	 2/3, 2.45%	 3/3, 2.15%	 1/3*, 2.25%	 1/3, 2.6%	 0/3*, 1.95%

Примечание. Показана одна из трех капель на 56 день кристаллизации, указано, в скольких каплях из трех выросли кристаллы, приведены средние объемные доли октамеров в растворах по данным МУРР (табл. 1)

* Фотография сделана на 84 день.

на 84 день). Все приведенные кристаллы выросли через день, кроме кристаллов в растворах с осадителями KCl и CuCl₂ (концентрация белка 20 мг/мл), которые выросли спустя 84 дня.

При концентрации лизоцима в капле 40 мг/мл кристаллы выросли во всех трех каплях при объемной доле октамеров в растворе от 4.05 до 6.70% (наибольшее значение наблюдалось при осадителе LiCl, наименьшее – при CuCl₂). Уменьшение концентрации белка с 40 до 20 мг/мл, ведущее и к уменьшению объемной доли октамеров при неизменяющейся объемной доли димеров в растворе, уменьшает вероятность появления кристалла вплоть до полного отсутствия кристаллов в случае использования осадителя CuCl₂ (1.95% октамеров). При использовании других осадителей наличие октамеров в растворе (при значении их средней объемной доли от 2.15% и больше) коррелирует с тем фактом, что в данных растворах наблюдался рост кристаллов лизоцима.

ЗАКЛЮЧЕНИЕ

Методом малоуглового рассеяния рентгеновского излучения на станциях P12 EMBL BioSAXS (DESY, Гамбург, Германия) и BM 29 BioSAXS (ESRF, Гренобль, Франция) определен олигомерный состав кристаллизационных растворов лизоцима с использованием разных осадителей NaCl, KCl, LiCl, NiCl₂ и CuCl₂. Для разных концентраций белка, осадителя и для разных температур показано, что в растворах присутствуют только димеры и октамеры, а также подтверждены выявленные ранее тенденции увеличения объемной доли октамеров при большем пересыщении [13].

Исследования объемной доли олигомеров в кристаллизационных растворах лизоцима при одинаковых условиях, проведенные на разных источниках синхротронного излучения, показали, что усредненные по нескольким измерениям объемные доли олигомеров отличаются друг от друга в относительных пределах 1–15% для октамеров и 25–30% для димеров. Однако зависимость суммарной объемной доли октамеров и димеров от порядкового номера катиона обратно пропорциональна порядковому номеру элемента: объемная доля димеров и октамеров уменьшается в ряду Li⁺, Na⁺, K⁺, Ni²⁺, Cu²⁺, что согласуется с увеличением активности ионов в лиотропном ряду Гофмейстера для Li⁺, Na⁺, K⁺ и увеличением ионных радиусов для Li⁺, Na⁺, K⁺ и для Ni²⁺, Cu²⁺.

Двукратное уменьшение концентрации белка в кристаллизационном растворе, ведущее и к уменьшению объемной доли октамеров при неизменяющейся объемной доли димеров, понижает вероятность появления кристалла вплоть до полного отсутствия кристаллов в случае использования осадителя CuCl₂ (1.95% октамеров). При использовании других осадителей наличие октамеров в растворе (при значении их средней объемной доли от 2.15% и больше) коррелирует с тем фактом, что в данных растворах наблюдался рост кристаллов лизоцима. Таким образом, удалось показать, что образование предкристаллизационной олигомерной фазы на начальной стадии кристаллизации в растворе может являться одним из необходимых условий того, что в растворе вырастет белковый кристалл.

Экспериментальные данные МУРР были собраны на станции P12, управляемой EMBL в Гам-

бурге, на накопителе PETRA III (DESY, Гамбург, Германия) и на станции VM 29 (ESRF, Гренобль, Франция). Авторы выражают благодарность А.Ю. Грузинову и Су Jeffries (P12) и Mark Tully (VM29) за помощь в проведении экспериментов.

Работа выполнена при частичной финансовой поддержке Министерства науки и высшего образования в рамках выполнения работ по Государственному заданию ФНИЦ “Кристаллография и фотоника” РАН, Российского фонда фундаментальных исследований (гранты № 18-32-20070 мол_а_вед, 19-29-12042 мк), НИЦ “Курчатовский институт” (приказ № 1360) и iNEXT [6938].

СПИСОК ЛИТЕРАТУРЫ

1. *Асхабов А.М.* // Записки Рос. минерал. о-ва. 2019. Т. 148. № 6. С. 1.
2. *Шефталъ Н.Н.* // Успехи физ. наук. 1957. Т. 62. С. 191.
3. *Гриздейл Р.О.* // Теория и практика выращивания кристаллов. М: Металлургия, 1968. С. 176.
4. *Асхабов А.М.* // Зап. Рос. минерал. о-ва. 2016. Т. 145. № 5. С. 17.
5. *Vekilov P.G.* // Nanoscale. 2010. V. 2. P. 2346. <https://doi.org/10.1039/C0NR00628A>
6. *Alexander E.S., Driessche V., Kellermeier M. et al.* New Perspectives on mineral nucleation and growth. From solution precursors to solid materials. Springer, 2017. 380 p.
7. *Kovalchuk M.V., Blagov A.E., Dyakova Y.A. et al.* // Cryst. Growth Des. 2016. V. 16. P. 1792. <https://doi.org/10.1021/acs.cgd.5b01662>
8. *Boikova A.S., Dyakova Y.A., Ilina K.B. et al.* // Acta Cryst. D. 2017. V. 73. № 7. P. 591. <https://doi.org/10.1107/S2059798317007422>
9. *Бойкова А.С., Дьякова Ю.А., Ильина К.Б. и др.* // Кристаллография. 2018. Т. 63. № 6. С. 857. <https://doi.org/10.1134/S0023476118060061>
10. *Kovalchuk M.V., Boikova A.S., Dyakova Y.A. et al.* // J. Biomol. Struct. Dyn. 2019. V. 37. № 12. P. 3058. <https://doi.org/10.1080/07391102.2018.1507839>
11. *Marchenkova M.A., Konarev P.V., Rakitina T.V. et al.* // J. Biomol. Struct. Dyn. 2020. V. 38. P. 2939. <https://doi.org/10.1080/07391102.2019.1649195>
12. *Ковальчук М.В., Алексеева О.А., Благов А.Е. и др.* // Кристаллография. 2019. Т. 64. № 1. С. 10. <https://doi.org/10.1134/S0023476119010156>
13. *Дьякова Ю.А., Бойкова А.С., Ильина К.Б. и др.* // Кристаллография. 2019. Т. 64. № 1. С. 15. <https://doi.org/10.1134/S0023476119010065>
14. *Sedlak E., Stagg L., Wittung-Stafshede P.* // Arch. Biochem. Biophys. 2008. № 479. P. 69. <https://doi.org/10.1016/j.abb.2008.08.013>
15. *Blanchet C.E., Spilotros A., Schwemmer F. et al.* // J. Appl. Cryst. 2015. V. 48. № 2. P. 431. <https://doi.org/10.1107/S160057671500254X>
16. *Pernot P., Round A., Barrett R. et al.* // J. Synchrotron Rad. 2013. V. 20. P. 660. <https://doi.org/10.1107/S0909049513010431>
17. *Franke D., Petoukhov M.V., Konarev P.V. et al.* // J. Appl. Cryst. 2017. V. 50. P. 1212. <https://doi.org/10.1107/S1600576717007786>
18. *Konarev P.V., Volkov V.V., Sokolova A.V. et al.* // J. Appl. Cryst. 2003. V. 36. P. 1277. <https://doi.org/10.1107/S0021889803012779>
19. *Svergun D.I., Barberato C., Koch M.H.J.* // J. Appl. Cryst. 1995. V. 28. P. 768. <https://doi.org/10.1107/S0021889895007047>
20. *Ковальчук М.В., Просеков П.А., Марченкова М.А. и др.* // Кристаллография. 2014. Т. 59. № 5. С. 749. <https://doi.org/10.7868/S0023476114050105>
21. *Kovalchuk M.V., Boikova A.S., Dyakova Y.A. et al.* // Thin Solid Films. 2019. V. 677. P. 13. <https://doi.org/10.1016/j.tsf.2019.02.051>

ДИФРАКЦИЯ И РАССЕЯНИЕ ИОНИЗИРУЮЩИХ ИЗЛУЧЕНИЙ

УДК 538.911

ПОИСК ЭФФЕКТИВНОГО АЛГОРИТМА ДЛЯ ВОССТАНОВЛЕНИЯ РЕШЕНИЙ ПРИ АНАЛИЗЕ ДАННЫХ МАЛОУГЛОВОГО РАССЕЯНИЯ ОТ РАСТВОРА КРЕМНЕЗОЛЯ

© 2021 г. А. Е. Крюкова^{1,2,*}, П. В. Конарев^{1,2}, В. В. Волков¹

¹ Институт кристаллографии им. А.В. Шубникова, ФНИЦ “Кристаллография и фотоника” РАН, Москва, Россия

² Национальный исследовательский центр “Курчатовский институт”, Москва, Россия

*E-mail: KrukovaAE@yandex.ru

Поступила в редакцию 08.06.2020 г.

После доработки 17.06.2020 г.

Принята к публикации 19.06.2020 г.

Представлено сравнение устойчивости восстановления решения задачи определения объемных распределений частиц по размерам на примере данных малоуглового рассеяния от раствора кремнезоля, состоящего из монодисперсных растворов TM-50 и SM. Исследован ряд итеративных алгоритмов: квазиньютоновский градиентный метод минимизации в варианте Бройдена–Флетчера–Голдфарба–Шанно с простыми ограничениями на параметры, метод имитации отжига, их комбинация, а также симплекс-метод Нелдера–Мида. В результате сравнения границ устойчивости для трех пар параметров раствора кремнезоля (радиусов и полидисперсностей частиц) методы выстроены в порядке возрастания их эффективности. Наиболее подходящим из предложенных подходов оказалась комбинация градиентного метода минимизации и имитации отжига для двухкомпонентной полидисперсной системы раствора кремнезоля с малоперекрывающимся объемным распределением частиц по размерам. Подход применим к исследованию систем различной природы, в том числе растворов белковых комплексов.

DOI: 10.31857/S0023476121050106

ВВЕДЕНИЕ

Подавляющее большинство систем в природе, представляющих интерес для научных исследований в различных отраслях, в частности медицине, биологии и промышленности, являются многокомпонентными и полидисперсными. Описание и расшифровка структуры такого рода систем, не исследованных ранее, в настоящее время являются актуальной задачей. Одним из эффективных методов по восстановлению структуры исследуемого объекта является метод малоуглового рассеяния (МУР) [1–5]. Для анализа полидисперсных объектов в растворе по данным МУР разработан ряд программ, таких как MIXTURE [1], GNOM [6], SASFIT [7], SASVIEW [8], McSAS [9] и другие.

Ранее для оценки качества работы программы MIXTURE (с реализованным методом переменной метрики в варианте Бройдена–Флетчера–Голдфарба–Шанно с простыми ограничениями на параметры [10], от англ. Broymden–Fletcher–Goldfarb–Shanno algorithm, **BFGS**) было проведено исследование на устойчивость решения задачи определения структурных параметров наночастиц для теоретических данных МУР, рассчитанных от различных модельных систем [11, 12]. Использо-

вались системы, состоящие из сферических частиц двух типов, имеющих перекрывающиеся и неперекрывающиеся объемные распределения по размерам. Кроме того, к наборам данных добавлялся шум двух типов (гауссовский и пуассоновский) с различным относительным уровнем и исследовалась степень отклонения найденного решения от заданной теоретической модели. Было показано, что диапазон области стартовых значений параметров, позволяющий восстановить истинное решение, зависит от многих факторов, в частности от относительного вклада параметров компонентов смеси в общую кривую рассеяния и от уровня и вида шума, содержащегося в данных.

На следующем этапе полученный результат проверяли на примере реальной экспериментальной системы – раствора кремнезольей, представляющего собой смесь частиц TM-50 и SM [13]. Предварительно с помощью программы GNOM [6] было выявлено наличие двух типов полидисперсных компонентов. Далее в рамках двухкомпонентной модели сферических частиц исследовалась возможность восстановления решений программой MIXTURE. Установлено, что характеры границ устойчивости решений схожи для

теоретических и экспериментальных данных, что свидетельствует о стабильности работы алгоритма, реализованного в программе MIXTURE. В то же время ограничения областей стартовых значений параметров, при которых возможно восстановление, указывает на попадание в локальные минимумы в процессе поиска решения.

Вследствие этого возникла идея проверить эффективность работы модифицированной версии программы MIXTURE, в алгоритме которой реализован метод имитации отжига (англ. Simulated annealing, SA) [14], в ряде случаев позволяющий находить глобальный минимум целевой функции. Результаты такого исследования приведены в [15]. В качестве объекта исследования был выбран тот же раствор кремнезоля, что и в [13]. Анализ сходимости решения проводили при одновременном варьировании только одной пары параметров – средних радиусов частиц кремнезоля R_1 и R_2 . В результате было установлено, что обе модификации программы (с использованием BFGS- и SA-процедур) дали схожие результаты, но при этом границы для BFGS-процедуры оказались шире.

Далее была рассмотрена возможность комбинированного использования упомянутых выше алгоритмов минимизации: квазиньютоновского градиентного спуска и метода имитации отжига [16]. Предполагалось, что это улучшит возможности восстановления распределений частиц по размерам. Действительно, такой подход позволил расширить границы параметров, при старте с которых могут быть найдены истинные значения системы.

Цель данной работы – систематическое сравнение границ устойчивого приближения для программы с различными минимизационными схемами. Для этого провели исследования на устойчивость упомянутых выше модификаций программы MIXTURE (BFGS, SA, BFGS + SA) для трех пар параметров: (R_1, R_2) , (dR_1, R_2) и (R_1, dR_2) , где dR_1 и dR_2 – полидисперсности двух типов частиц кремнезоля. Помимо этого исследовали возможности алгоритма с применением симплекс-метода Нелдера–Мида (англ. Nelder–Mead, NM) [17]. Исследования проводили на примере экспериментальных данных МУР от полидисперсного раствора кремнезоля.

КРЕМНЕЗОЛИ

Раствор, выбранный в качестве объекта исследования, получен путем смешения монодисперсных растворов кремнезольей Ludox TM-50TM и SMTM. Образцы TM-50 (50 мас. % SiO₂ и 0.3 мас. % Na) и SM (30 мас. % SiO₂ и 0.2 мас. % Na) изготовлены компанией Grace Davidson [18]. Несмотря на то что размеры (диаметры) частиц, заявленные про-

изводителем (для сухих образцов), составляют 70 Å для SM и 220 Å для TM-50, ранее было показано [19], что в растворе они увеличиваются до 100 Å для SM и 270 Å для TM-50. Раствор кремнезоля разводили водой в соотношении 1:100, чтобы избежать эффекта межчастичного взаимодействия, при этом центрифугирование не использовали, так как растворы изначально стабилизировались с использованием гидроксида натрия NaOH для предотвращения агрегации.

ЭКСПЕРИМЕНТАЛЬНАЯ УСТАНОВКА

Эксперимент по исследованию раствора кремнезоля выполнен на установке “АМУР-К” [20] (ФНИЦ “Кристаллография и фотоника” РАН, Москва, Россия) и проведен следующим образом: образец кремнезоля помещали в стеклянный капилляр, а затем в вакуумную камеру. Расстояние образец–детектор составляло 700 мм. Время измерений одного образца составляло 10 мин. Экспериментальные данные нормировали на интенсивность падающего пучка, после чего в них вводили поправку на коллимационные искажения.

Установка “АМУР-К” оснащена однокоординатным позиционно-чувствительным детектором ОДЗМ при фиксированной длине волны излучения λ , равной 0.1542 нм (CuK $_{\alpha}$ -линия острофокусной трубки, монохроматор из пиролитического графита), и коллимационной системой Кратки. Сечение рентгеновского пучка составляло 0.2×8 мм², область углов рассеяния соответствовала диапазону значений модуля вектора рассеяния

$$0.1 < s < 10.0 \text{ нм}^{-1} \quad (1)$$

$$\left(|s| = \frac{4\pi \sin \theta}{\lambda}, 2\theta - \text{угол рассеяния} \right).$$

В результате эксперимента получена кривая МУР от раствора кремнезоля (рис. 1а, точки). Наилучшее приближение к экспериментальным данным, рассчитанное программой MIXTURE (рис. 1а, сплошная линия), и объемное распределение по радиусам частиц представлены на рис. 1б. Также приведено приближение, при котором решение не попало в глобальный минимум функции (рис. 1а, пунктирная линия). Согласно эксперименту радиусы равны 49 и 140 Å, что согласуется с результатами в [19]. Параметры модели, восстановленные с использованием четырех модификаций программы MIXTURE, представлены в табл. 1. Видно, что результаты восстановления схожи.

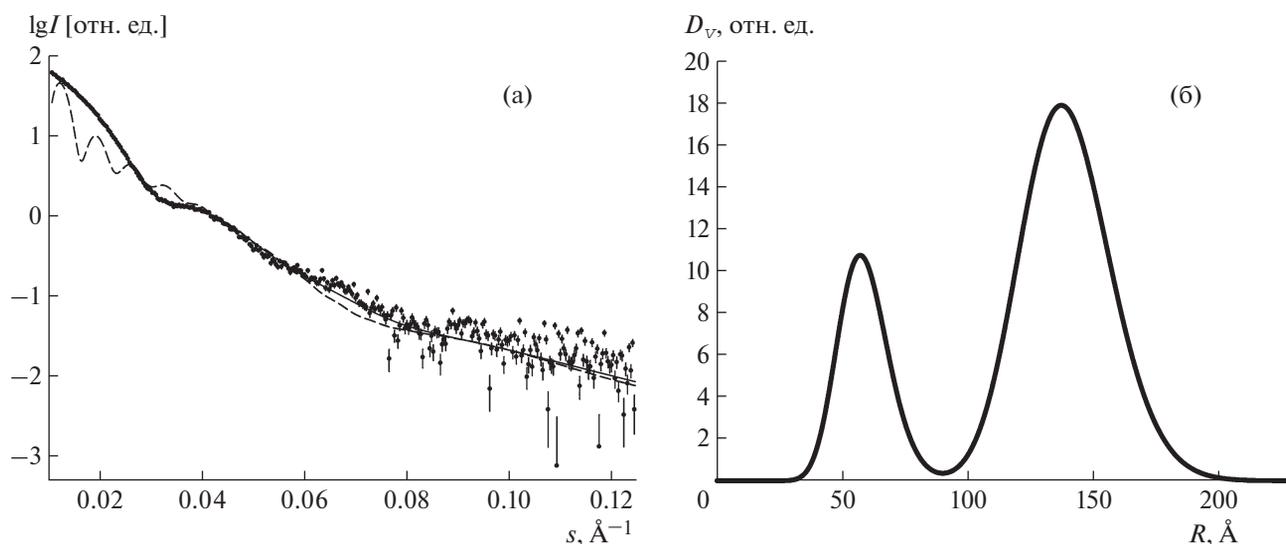


Рис. 1. Экспериментальная кривая МУР от раствора кремнезоля (точки). Наилучшее приближение к экспериментальным данным, рассчитанное программой MIXTURE (сплошная линия). Приближение, при котором решение не попало в глобальный минимум функции (пунктирная линия) (а). Объемное распределение по радиусам сферических частиц, найденное программой MIXTURE, соответствующее наилучшему приближению (б).

МЕТОД РАСЧЕТА ИНТЕНСИВНОСТИ МАЛОУГЛОВОГО РАССЕЯНИЯ ОТ МНОГОКОМПОНЕНТНОЙ ПОЛИДИСПЕРСНОЙ СИСТЕМЫ

Интенсивность рассеяния многокомпонентной системы в случае разбавленной системы можно представить в виде

$$I(s) = \sum_k v_k I_k(s), \quad (2)$$

где суммирование проводится по разным компонентам, v_k – относительная объемная доля k -го компонента, $I_k(s)$ – интенсивность k -го компонента.

Для полидисперсной системы частиц интенсивность каждого компонента можно представить в виде

$$I_k(s) = \int D_k(R) V_k(R) [\Delta\rho_k(R)]^2 i_{0k}(s, R) dR, \quad (3)$$

где R – размер частицы, $D_k(R)$ – нормированное объемное распределение частиц по размеру для k -го компонента, $V_k(R)$ – эффективный объем k -го компонента, $\Delta\rho_k(R)$ – рассеивающий контраст k -го компонента, $i_{0k}(s, R)$ – нормированный формфактор k -го компонента.

Для расчета функции объемного распределения k -го компонента использовали распределение Шульца:

$$D(R) = \frac{1}{R_0} (z+1)^{z+1} \frac{1}{\Gamma(z+1)} \left(\frac{R}{R_0}\right)^z \exp\left[-\frac{(z+1)R}{R_0}\right], \quad (4)$$

где $z = \left(\frac{R_0}{\Delta R}\right)^2 - 1$, R_0 – среднее значение размера частицы, ΔR – дисперсия распределения, $\Gamma(z)$ – гамма-функция.

Критерием качества решения системы служит значение квадратичной невязки χ^2 , которое в результате поиска распределения частиц по размерам минимизируется:

$$\chi^2 = \frac{1}{N-1} \sum_j \left[\frac{I(s_j) - c I_{\text{calc}}(s_j)}{\sigma(s_j)} \right]^2, \quad (5)$$

где N – число экспериментальных точек, c – шкалирующий коэффициент, $I(s_j)$ и $I_{\text{calc}}(s_j)$ – экспериментальная и рассчитанная интенсивности соответственно, $\sigma(s_j)$ – экспериментальная ошибка, соответствующая модулю вектора рассеяния s_j .

Алгоритм, использующий формулы (2)–(4), реализован в программе MIXTURE. Такой под-

Таблица 1. Найденные структурные параметры системы, состоящей из двух типов сферических наночастиц кремнезоля – ТМ-50 и SM

	ТМ-50		SM	
	$R_1, \text{\AA}$	$dR_1, \text{\AA}$	$R_2, \text{\AA}$	$dR_2, \text{\AA}$
Метод Нелдера–Мида (NM)	59.8	10.0	140.0	17.9
Метод имитации отжига (SA)	58.7	13.1	139.4	17.9
Квазиньютоновский градиентный метод (BFGS)	58.8	10.0	139.4	17.9
BFGS + SA	58.8	10.0	139.4	17.9

ход применим к исследованию систем различной природы, в частности он успешно использовался для количественного описания температурного структурного перехода в трехкомпонентных микроэмульсиях [21], анализа ансамбля состояний самособирающихся икосаэдрических капсул фермента люмазинсинтазы в нативном состоянии и при наличии мутаций [22], а также изучения олигомерного состава ассоциатов наночастиц оксида железа, внедренных в капсулы вируса мозаики Бром [23].

Основным фактором, затрудняющим поиск глобального минимума функции, является нелинейная зависимость расчетных данных от параметров модели. При стандартном анализе для каждого параметра задают диапазон возможных значений, в пределах которого ищут решения. В зависимости от сложности используемой модели, стартовых значений параметров и шума экспериментальных данных решения могут существенно различаться. Поэтому анализ устойчивости и возможности восстановления теоретических данных приобретает исключительно важный характер.

ИСПОЛЬЗУЕМЫЕ МЕТОДЫ МИНИМИЗАЦИИ

В ходе работы проводили исследование на устойчивость алгоритмов четырех модификаций программы MIXTURE. В каждой из которых использовали различные минимизационные схемы: метод BFGS [10], SA [14], комбинация этих методов (BFGS + SA), а также NM [17].

Алгоритм BFGS является итерационным методом, относится к классу квазиньютоновских методов, в которых матрица вторых производных (гессиан) функции вычисляется не напрямую, а в некотором приближении, основываясь на сделанных до этого шагах. Метод достаточно эффективен, не требует больших временных затрат для поиска решения, но при этом решение может остановиться в локальном минимуме целевой функции, находящемся далеко от глобального экстремума целевой функции.

Метод SA является стохастическим методом поиска глобального экстремума. Он является наиболее подходящим для минимизации функций с большим числом переменных, что идеально подходит для многих задач МУР. Алгоритм основан на методе случайных испытаний, запускается с произвольной модели, параметры которой затем изменяются случайным образом в соответствии со схемой Метрополиса [24] до тех пор, пока решение не станет соответствовать экспериментальным данным, например, по критерию χ^2 . В отличие от метода квазиньютоновского градиентного спуска, где на каждом следующем шаге принимается решение с меньшим χ^2 (5), в методе

SA с определенной долей вероятности (определяемой так называемой “температурой отжига”) может приниматься худшее решение, это дает возможность программе выходить из локальных минимумов и продолжать спуск. По мере продолжения поиска значение температуры монотонно снижается, с использованием шкалирующего коэффициента 0.93–0.95 после каждых 1000–10000 вариаций модели. Однако поиск решения с помощью такого алгоритма занимает длительное время (так как требуется достаточно большое количество температурных итераций).

Комбинация BFGS и SA представляет собой поочередное использование этих алгоритмов минимизации. Схема поиска минимума показана на рис. 2. В целях поиска первого (возможно, локального) минимума нелинейной целевой функции (1) проводят старт программы MIXTURE, использующей градиентный метод минимизации, с некоторых значений параметров двухкомпонентной полидисперсной модели N_1 и N_2 (где N – один из структурных параметров R_1 , R_2 , dR_1 либо dR_2). Далее с оптимизированным методом BFGS значений проводится запуск модифицированной версии MIXTURE с использованием метода имитации отжига для того, чтобы вывести систему «из равновесия», причем со всего лишь одной итерацией по температуре, поэтому вычисления не занимают длительного времени. После этого с новых полученных значений программа запускается вновь с градиентным методом и так повторяется до тех пор, пока параметры модели не перестанут изменяться. За пороговое значение их изменчивости брали 1–5% от величины, полученной на предыдущей итерации.

Метод NM, известный как метод деформируемого многогранника (часто используемое название “симплекс-метод” неверно, так как этот термин относится к другому, конфигурационному методу поиска), простой и в то же время эффективный, позволяющий оптимизировать функции многих переменных без вычисления градиентов. Суть метода заключается в последовательном перемещении вершин многогранника в пространстве параметров модели в сторону точки экстремума. Однако если метод находит локальный экстремум типа “плоского узкого оврага”, то многогранник может сжаться настолько, что поиск экстремума прекратится из-за ограниченной точности вычислений. В этом случае достаточно эффективным оказывается повторный запуск с новой конфигурации вершин. Метод отличается высокой надежностью и позволяет после нескольких запусков найти решение, более точное, чем ранее рассмотренные алгоритмы, хотя и требует длительных расчетов.

Несмотря на то что все эти методы являются достаточно эффективными, в ряде случаев только

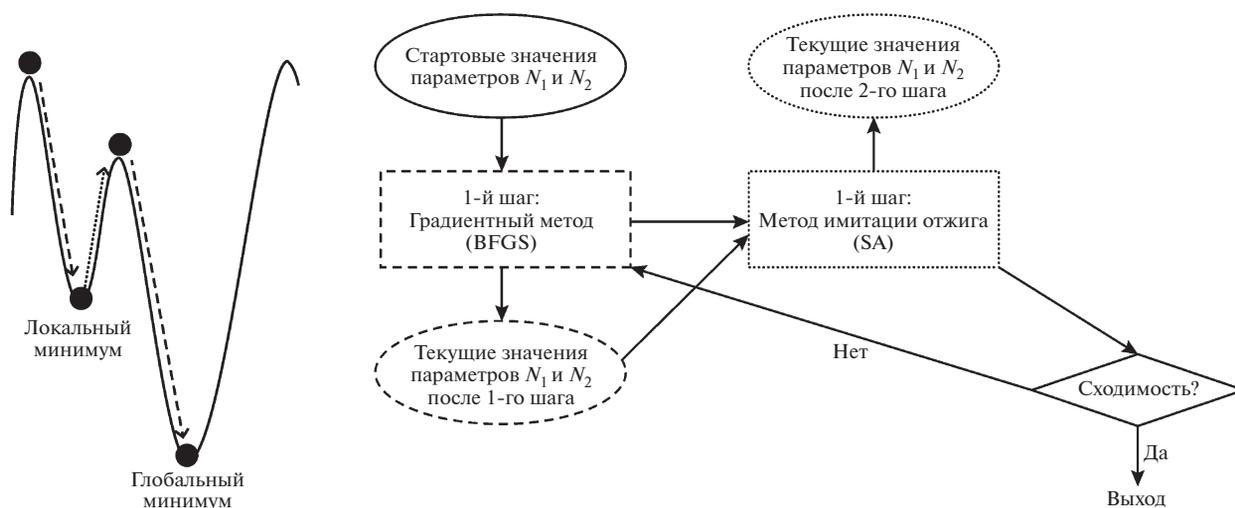


Рис. 2. Схема поиска глобального минимума целевой функции при комбинированном использовании методов BFGS и SA.

экспериментальным путем можно определить, будет ли используемый метод подходить для поставленной задачи. Поэтому с целью нахождения наиболее эффективного алгоритма для решения задачи восстановления объемного распределения частиц по размерам для двухкомпонентной полидисперсной системы были протестированы все упомянутые выше модификации программы MIXTURE.

ИССЛЕДОВАНИЕ УСТОЙЧИВОСТИ ВОССТАНОВЛЕНИЯ РЕШЕНИЙ

Приближение проводили в рамках модели, состоящей из двух типов сферических частиц с радиусами R_1 , R_2 , полидисперсностями dR_1 , dR_2 и объемными долями V_1 , V_2 . Исследовали сходимость расчетов к правильному решению при одновременном варьировании стартовых значений следующих пар параметров: (R_1, R_2) , (dR_1, R_2) , (R_1, dR_2) , тогда как остальные параметры фиксировались на истинных значениях, найденных программой MIXTURE (табл. 1) с относительными объемными долями $V_1 = 0.25$ для меньших частиц кремнезоля (SM) и $V_2 = 0.75$ для больших частиц кремнезоля (TM-50). В табл. 1 эти значения не приведены, так как алгоритмы всех исследуемых модификаций программы MIXTURE устойчивы к изменению объемных долей во всем диапазоне от 0 до 1.

Диапазоны, в пределах которых выбирали стартовые значения параметров, были следующими: для радиусов R_1 10.0–96.0 Å и R_2 100.0–500.0 Å, для полидисперсностей dR_1 и dR_2 0.5–50.0 Å. Диапазоны были выбраны таким образом, чтобы радиусы меньших и больших частиц не

коррелировали друг с другом. При этом шаг расчетной сетки по R_1 и R_2 составлял 2 Å, шаг сетки по dR_1 и dR_2 – 0.5 Å. Таким образом, для пары параметров (R_1, R_2) были посчитаны 8844 точки, для (dR_1, R_2) – 20 100 точек, для (R_1, dR_2) – 4400 точек. На основе этих данных построены двумерные карты, представленные на рис. 3. Верхняя панель соответствует результатам, полученным модифицированной версией программы MIXTURE с реализованным методом NM, вторая – программе с методом SA, третья – с BFGS, четвертая (нижняя) панель – комбинации методов BFGS и SA. Первый столбец (слева) – для пары параметров (R_1, R_2) , второй – для (dR_1, R_2) , третий – для (R_1, dR_2) . Область светлого цвета соответствует нахождению истинного решения (в случае, когда значения параметров, полученных в результате моделирования программой MIXTURE, находились в пределах 5% от найденных ранее оптимальных значений). Область темного цвета соответствует всем остальным решениям, существенно отличающимся от истинного решения. Пунктирной линией отмечены истинные (оптимальные) значения параметров, восстановленные программой MIXTURE, указанные в табл. 1.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Анализ двумерных карт позволил установить, что самые широкие границы устойчивости соответствуют комбинации методов BFGS и SA для всех трех пар параметров. При этом можно выстроить ряд алгоритмов в порядке возрастания их эффективности: NM, SA, BFGS, BFGS + SA, поэтому на рис. 3 они расположены в соответствующем порядке, сверху вниз.

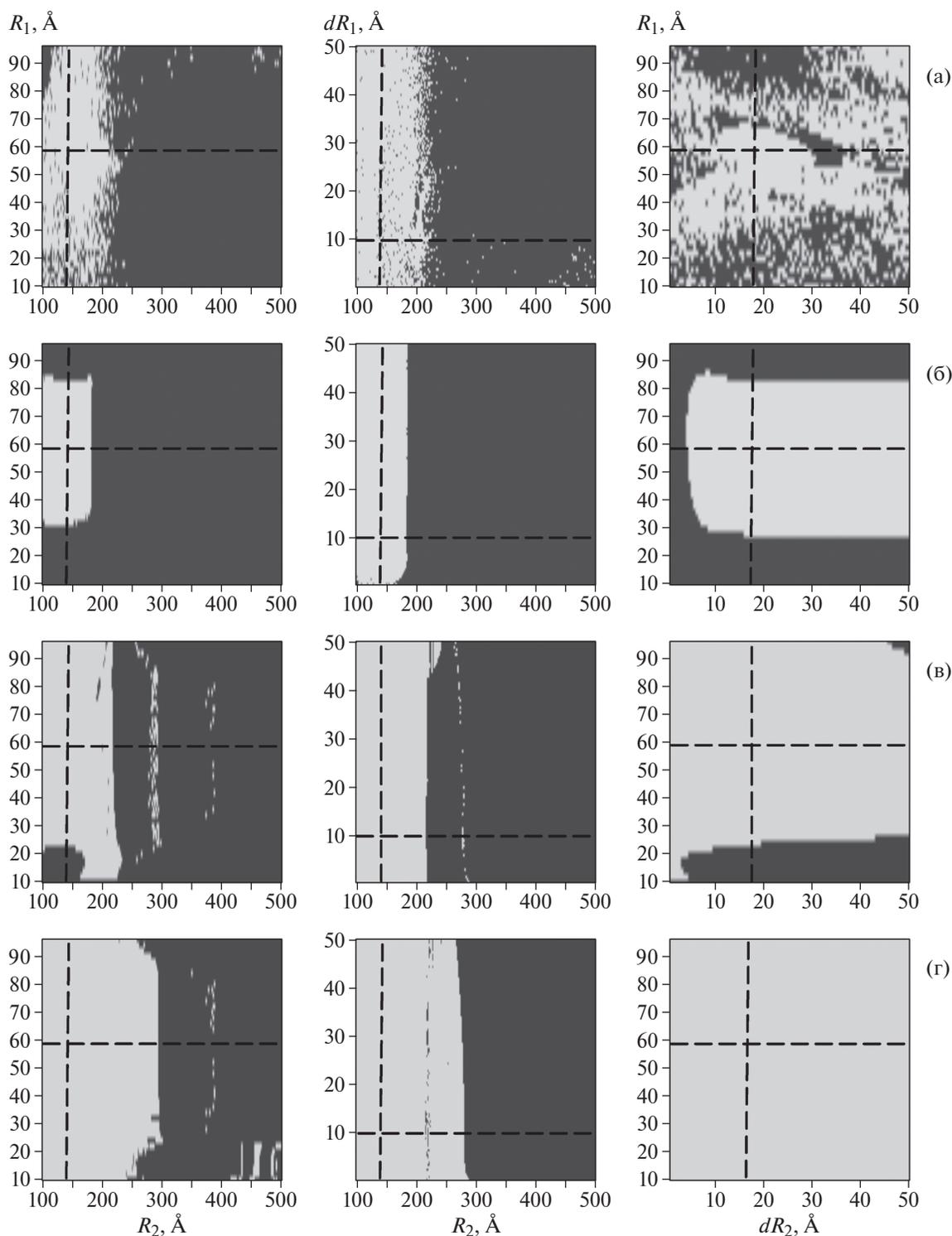


Рис. 3. Контурные графики устойчивости восстановления решения программой MIXTURE в зависимости от стартовых значений радиусов сферических частиц кремнезоля и их полидисперсностей (R_1, R_2) , (dR_1, R_2) , (R_1, dR_2) . Пересечение пунктирных линий соответствует точным (истинным) значениям параметров. Области темного цвета – “неудача” (отличие от истинных параметров более чем на 5%), области светлого цвета – “успешное” нахождение решения (найденные параметры находятся в пределах 5% от истинных). Используемые алгоритмы минимизации: NM (а), SA (б), BFGS (в), BFGS + SA (г).

Помимо этого, в некоторых случаях области нахождения правильных решений имеют достаточно сложные границы фрактального вида, с отдельными «очагами» неустойчивости внутри данных диапазонов, что свидетельствует о разной чувствительности методов минимизации к определенным параметрам модели.

Анализируя двумерные карты, отметим, что метод NM оказался наименее эффективным по сравнению с другими использованными методами. Несмотря на то что его границы устойчивости шире, чем для метода SA, сравнительно легко попасть в область неустойчивости при незначительном отклонении от истинного значения. Для других трех методов можно говорить об имеющихся тенденциях. Для параметров R_1 и R_2 прослеживаются следующие закономерности: стартовые значения допустимо варьировать в пределах 50, 65 и 80% для R_1 и 15, 65, 100% для R_2 от истинного значения для методов SA, BFGS, BFGS + SA соответственно. Во всех случаях при значении R_2 более 300 Å не удалось добиться восстановления данных. То есть стартовое значение данного параметра может не более чем в 2 раза превышать его истинное значение. При этом решения являются устойчивыми к варьированию параметров dR_1 и dR_2 во всем исследуемом диапазоне.

Также отметим, что метод SA теоретически мог бы позволить найти более широкие границы устойчивости, если бы существенно увеличили число вариаций модели при каждой итерации по температуре, но это на порядок увеличило бы время расчетов, что, как полагаем, неприемлемо с практической точки зрения. Кроме того, метод SA требует длительной работы по поиску оптимального режима снижения температуры, общего рецепта для этого до сих пор не выработано. В случае комбинированной схемы BFGS + SA метод SA используется исключительно с целью выхода из локального минимума и последующего продолжения поиска градиентным методом, для этой цели достаточно использования одной итерации по температуре.

ЗАКЛЮЧЕНИЕ

Исследована устойчивость решений программы MIXTURE для двухкомпонентного раствора кремнезоля с использованием следующих методов минимизации: градиентного метода (BFGS), метода имитации отжига (SA), сочетания этих двух методов (BFGS + SA), а также симплекс-метода Нелдера–Мида (NM). Анализ полученных данных позволяет распределить методы в порядке возрастания их эффективности: NM, SA, BFGS, BFGS + SA. Таким образом, алгоритм поиска с последовательным комбинированием методов минимизации BFGS и SA для систем сфериче-

ских частиц с малоперекрывающимися объемными распределениями частиц по размерам позволяет максимально эффективно восстанавливать функцию $D_v(r)$ в более широких диапазонах стартовых значений параметров, чем с использованием других предложенных алгоритмов. Разработанные подходы применимы к исследованию систем различной природы, в том числе для растворов белковых комплексов.

Работа выполнена при поддержке Министерства науки и высшего образования РФ в рамках выполнения работ по Государственному заданию ФНИЦ «Кристаллография и фотоника» РАН, а также при частичной финансовой поддержке Российского фонда фундаментальных исследований (грант № 19-32-90190 Аспиранты) и Программы деятельности НИЦ КИ на 2018–2022 г. Малоугловые измерения выполнены с использованием оборудования ЦКП ФНИЦ «Кристаллография и фотоника» РАН при поддержке Министерства науки и высшего образования России (проект RFMEFI62119X0035).

СПИСОК ЛИТЕРАТУРЫ

1. Svergun D.I., Konarev P.V., Volkov V.V. et al. // J. Chem. Phys. 2000. V. 113. P. 1651. <https://doi.org/10.1063/1.481954>
2. Konarev P.V., Petoukhov M.V., Volkov V.V., Svergun D.I. // J. Appl. Cryst. 2006. V. 39. P. 277. <https://doi.org/10.1107/S0021889806004699>
3. Petoukhov M.V., Franke D., Shkumatov A.V. et al. // J. Appl. Cryst. 2012. V. 45. P. 342. <https://doi.org/10.1107/S0021889812007662>
4. Konarev P.V., Volkov V.V., Sokolova A.V. et al. // J. Appl. Cryst. 2003. V. 36. P. 1277. <https://doi.org/10.1107/S0021889803012779>
5. Svergun D.I., Feigin L.A. X-Ray and Small-Angle Neutron Scattering. Moscow: Nauka, 1986.
6. Svergun D.I. // J. Appl. Cryst. 1992. V. 25. P. 495. <https://doi.org/10.1107/S0021889892001663>
7. Bressler I., Kohlbrecher J., Thünemann A.F. // J. Appl. Cryst. 2015. V. 48. P. 1587. <https://doi.org/10.1107/S1600576715016544>
8. Doucet M., Cho J.H., Gervaise A. et al. // SasView version 4.2.2. Zenodo. 2019. <https://doi.org/10.5281/zenodo.2652478>
9. Bressler I., Pauw B.R., Thünemann A.F. // J. Appl. Cryst. 2015. V. 48. P. 962. <https://doi.org/10.1107/S1600576715007347>
10. Gill P.E., Murray W., Wright M.H. Practical Optimization. London: Academic Press, 1981.
11. Kryukova A.E., Konarev P.V., Volkov V.V. // J. Physics: Conf. Ser. 2017. V. 941. P. 012069. <https://doi.org/10.1088/1742-6596/941/1/012069>
12. Крюкова А.Е., Конарев П.В., Волков В.В. // Кристаллография. 2018. Т. 63. № 1. С. 32. <https://doi.org/10.7868/S0023476118010113>

13. *Крюкова А.Е., Козлова А.С., Конарев П.В. и др.* // Кристаллография. 2018. Т. 63. № 4. С. 524.
<https://doi.org/10.1134/S002347611804015X>
14. *Kirkpatrick S., Gelatt C.D., Vecchi M.P.* // Science. 1983. V. 220. P. 671.
<https://doi.org/10.1126/science.220.4598.671>
15. *Kryukova A.E., Konarev P.V., Volkov V.V., Asadchikov V.E.* // J. Mol. Liquids. 2019. V. 283. P. 221.
<https://doi.org/10.1016/j.molliq.2019.03.070>
16. *Kryukova A.E., Konarev P.V., Volkov V.V., Asadchikov V.E.* // Phys. Atomic Nuclei. 2019. V. 82. P. 1576.
<https://doi.org/10.1134/S1063778819110097>
17. *Nelder J.A., Mead R.* // Computer J. 1965. V. 7. P. 308.
<https://doi.org/10.1093/comjnl/7.4.308>
18. *Dumont F., Warlus J., Watillon A. et al.* // Interface Sci. 1990. V. 138. P. 543.
[https://doi.org/10.1016/0021-9797\(90\)90236-H](https://doi.org/10.1016/0021-9797(90)90236-H)
19. *Asadchikov V.E., Volkov V.V., Volkov Yu.O. et al.* // JETP Lett. 2011. V. 94. P. 625.
<https://doi.org/10.1134/S0021364011190040>
20. *Mogilevskii L.Yu., Dembo A.T., Svergun D.I., Feigin L.A.* // Crystallography Reports. 1984. V. 29. P. 587.
21. *Svergun D.I., Konarev P.V., Volkov V.V. et al.* // J. Chem. Phys. 2000. V. 113. P. 1651.
<https://doi.org/10.1063/1.481954>
22. *Zhang X., Konarev P.V., Petoukhov M.V. et al.* // J. Mol. Biol. 2006. V. 362. P. 753.
<https://doi.org/10.1016/j.jmb.2006.07.037>
23. *Malyutin A.G., Cheng H., Sanchez-Felix O.R. et al.* // ACS Appl. Mater. Interfaces. 2015. V. 7. P. 12089.
<https://doi.org/10.1021/acsami.5b02278>
24. *Metropolis N., Rosenbluth A.W., Rosenbluth M.N. et al.* // J. Chem. Phys. 1953. V. 21. P. 1087.
<https://doi.org/10.1063/1.1699114>

ДИФРАКЦИЯ И РАССЕЯНИЕ ИОНИЗИРУЮЩИХ ИЗЛУЧЕНИЙ

УДК 537.533.73

ЭЛЕКТРОННАЯ ДИФРАКЦИЯ МИКРОКРИСТАЛЛОВ НА ПРИМЕРЕ ЛИЗОЦИМА

© 2021 г. Р. А. Камышинский^{1,2,3}, В. А. Кралин^{1,3}, М. Ю. Чесноков^{1,2},
В. Р. Самыгина^{1,2}, А. С. Орехов^{1,2,3,*}

¹ *Национальный исследовательский центр “Курчатовский институт”, Москва, Россия*

² *Институт кристаллографии им. А.В. Шубникова ФНИЦ “Кристаллография и фотоника” РАН, Москва, Россия*

³ *Московский физико-технический институт (национальный исследовательский университет),
Долгопрудный, Россия*

*E-mail: Orekhov.anton@gmail.com

Поступила в редакцию 15.12.2020 г.

После доработки 13.01.2021 г.

Принята к публикации 18.01.2021 г.

Электронная дифракция микрокристаллов – новый мощный подход криоэлектронной микроскопии, позволяющий определять структуру трехмерных кристаллов по поворотной серии дифракционных картин. В отличие от рентгеновской кристаллографии настоящий метод позволяет изучать микро- и даже наноразмерные кристаллы, при этом для получения структуры с высоким пространственным разрешением достаточно одного или нескольких кристаллов, что существенно снижает необходимые объемы образца. Рассмотрены методические аспекты применения электронной дифракции микрокристаллов на примере классического модельного белка – лизоцима.

DOI: 10.31857/S0023476121050088

ВВЕДЕНИЕ

Классическим способом получения трехмерных структур белков является рентгеновская кристаллография, основанная на взаимодействии рентгеновских лучей с кристаллической решеткой. Однако ввиду относительно слабого взаимодействия рентгеновского излучения с образцом [1] необходимы кристаллы размером более 20 мкм [2, 3], получение которых является крайне трудоемким и затруднительным для ряда объектов [4], что накладывает ограничения на применимость метода. Для получения наборов дифракционных данных с кристаллов размером более ~5 мкм используются источники синхротронного излучения, а в последние годы большие надежды по решению структур с использованием микрокристаллов меньших размеров возлагаются на серийную микрокристаллографию, в том числе фемтосекундную кристаллографию с использованием лазеров на свободных электронах [5, 6] (например, European XFEL). Данный подход подразумевает получение нескольких тысяч дифракционных картин отдельных микрокристаллов. Несмотря на увеличение количества успешно решенных структур [7], к настоящему моменту подобные установки являются уникальными, а для успешного эксперимента необходимы большие объемы изучаемых образцов [6].

Второе десятилетие двадцатого века отмечено бурным развитием методов просвечивающей криоэлектронной микроскопии (**крио-ЭМ**) и взрывным ростом их использования в клеточной и структурной биологии [8]. Наиболее известные подходы крио-ЭМ – метод анализа одиночных частиц (англ. SPA – Single Particle Analysis) и криоэлектронная томография – подразумевают получение прямых изображений изучаемого объекта с последующей обработкой и восстановлением 3D-структуры [9]. Как правило, для успешной 3D-реконструкции с получением структуры высокого (ниже 4 Å) пространственного разрешения толщина изучаемого объекта не должна превышать 100–200 нм, а его молекулярная масса в зависимости от конфигурации микроскопа и используемого детектора должна быть не менее ~40–150 кДа [10, 11].

Вторая группа подходов крио-ЭМ базируется на использовании данных электронной дифракции. Так, электронная 2D-кристаллография позволяет восстанавливать структуру тонких (менее 1 мкм) белковых кристаллов по двумерным дифракционным картинам [12]. Однако в силу необходимости использования относительно больших доз электронов для получения данных высокого разрешения и, как следствие, радиационных повреждений облучаемых областей [13] необходимо получение больших серий двумерных ди-

фракционных картин от десятков и сотен различных областей двумерных кристаллов с последующей ресурсоемкой обработкой данных.

Напротив, для получения дифракционной картины с высоким разрешением от трехмерных белковых кристаллов достаточно крайне низких электронных доз ($\sim 1 \text{ э}/\text{Å}^2$). Основываясь на сказанном выше, в 2013 г. группа под руководством Т. Гонена предложила новый метод — электронную дифракцию микрокристаллов (ЭДМ, англ. MicroED — Microcrystal Electron Diffraction) [14–16], которая заключается в получении непрерывной поворотной дифракционной серии от витрифицированных белковых кристаллов в широком угловом диапазоне [17] (как правило, в пределах $\pm 70^\circ$). Такой подход позволяет использовать белковые кристаллы существенно меньшего размера, чем в рентгеновской кристаллографии. При этом ввиду сходства экспериментальных процедур [18] обработка полученных данных осуществляется в стандартных кристаллографических программных пакетах [19], а достигаемое пространственное разрешение подчас превосходит результаты, полученные с помощью XFEL [20].

Безусловно, для метода ЭДМ сохраняется основное ограничение крио-ЭМ — толщина образца не должна превышать нескольких сотен нанометров. Однако метод применим для кристаллов и большего размера — например, для кристаллов толщиной до 100 мкм возможно получение тонких срезов толщиной до 100–200 нм травлением с помощью фокусированного ионного пучка в криогенном режиме (**крио-ФИП**). Кроме того, метод ЭДМ позволяет успешно восстанавливать 3D-структуры белков и белковых комплексов существенно меньшей молекулярной массы, чем требуется для других методов крио-ЭМ [21]. Для реализации метода используется то же оборудование, что и для других методов крио-ЭМ: источник с полевой эмиссией, быстрый и высокочувствительный детектор [22] (Falcon 3/Falcon 4, Gatan K2/K3) и энергетический фильтр [20, 23]. Можно предположить, что с ростом популярности крио-ЭМ по всему миру и увеличением доступности соответствующего оборудования метод ЭДМ может стать одним из основных методов структурной биологии уже в ближайшие годы. Так, к настоящему времени с помощью ЭДМ решены более 100 структур, причем некоторые из них — впервые [21]. В представленной работе впервые в России методом ЭДМ получены дифракционные картины тонких срезов кристаллов модельного белка лизоцима, приготовленных с использованием крио-ФИП.

ПОЛУЧЕНИЕ МИКРОКРИСТАЛЛОВ ЛИЗОЦИМА

Для кристаллизации использовали белок лизоцим из куриного яйца производства Sigma-Aldrich (CAS № 12650-88-3). Белок растворяли в 0.1 М натрий-ацетатном буфере, pH 4.5, в концентрации 200 мг/мл. Кристаллы размером 20–50 мкм выращены методом кристаллизации в объеме. Перед кристаллизацией раствор белка центрифугировали и смешивали в соотношении 1 : 1 с кристаллизационным раствором (300 мМ хлорида натрия, 0.1 М натрий-ацетатного буфера, pH 4.5). Кристаллы размером 15–35 мкм вырастали при температуре 4°C и имели классическую морфологию кристаллов тетрагональной кристаллической формы лизоцима.

ВИТРИФИКАЦИЯ МИКРОКРИСТАЛЛОВ

На первом этапе в течение 30 с осуществляли обработку электронно-микроскопической медной сетки с углеродной подложкой тлеющим разрядом с помощью прибора Pelco EasiGlow при силе тока 25 мА и давлении в камере 0.26 мбар. Далее на подготовленную сетку, помещенную в камеру прибора Vitrobot Mark IV (температура 4°C , влажность $\sim 95\%$), наносили 3 мкл раствора кристаллов лизоцима, после чего излишки раствора удаляли двусторонним сжатием фильтровальной бумагой и осуществляли процедуру витрификации — быстрое погружение сетки с образцом в жидкий этан при температуре жидкого азота с образованием тонкого слоя аморфного льда (рис. 1а). В процессе нанесения образца на сетку температура в камере системы Vitrobot составляла 4°C , влажность — не менее 95%.

ПОЛУЧЕНИЕ ТОНКОГО СРЕЗА МИКРОКРИСТАЛЛОВ

Полученную сетку с витрифицированным раствором микрокристаллов лизоцима переносили в предварительно откачанную до высокой степени вакуума ($\sim 10^{-5}$ Па) камеру растрового электронно-ионного микроскопа Versa 3D, оснащенного криогенной приставкой QuorumPRT3010. Далее с помощью газовой-инжекционной системы осуществляли четыре последовательных цикла напыления защитного проводящего слоя платины при температуре 26°C на поверхность сетки в течение 3–5 с. Пауза между циклами составляла не менее 10 с, что позволило избежать артефактов напыления, снизить накопление заряда и защитить поверхность образца от радиационных повреждений в процессе эксперимента (рис. 1б).

На рис. 2а представлено изображение, полученное в криогенном растворе электронном микроскопе (крио-РЭМ), микрокристаллов лизоцима на поверхности сетки. Большинство кри-

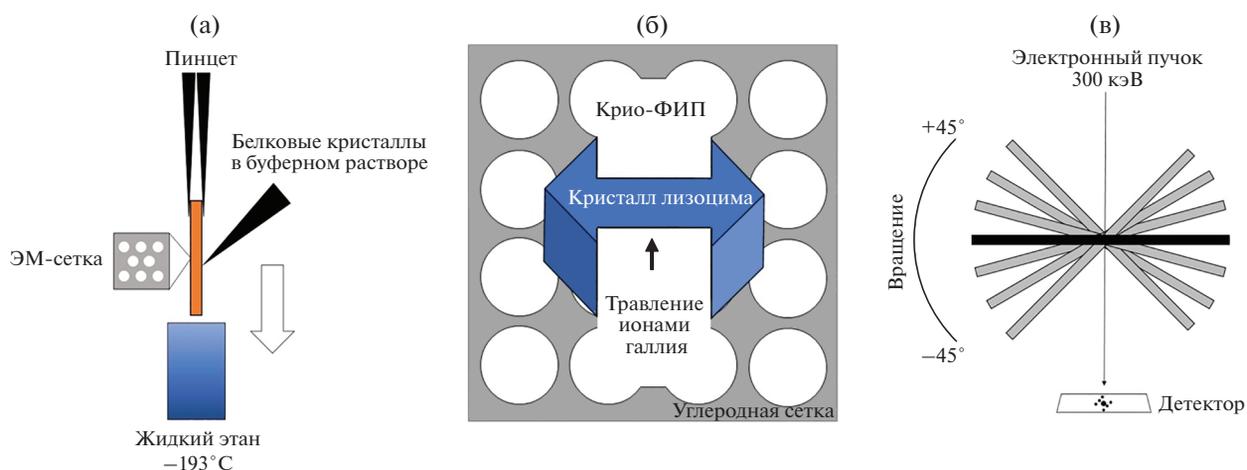


Рис. 1. Схемы экспериментов: а – заморозка кристаллов, б – приготовление тонкого среза замороженного образца с помощью крио-ФИП, в – получение поворотной серии изображений.

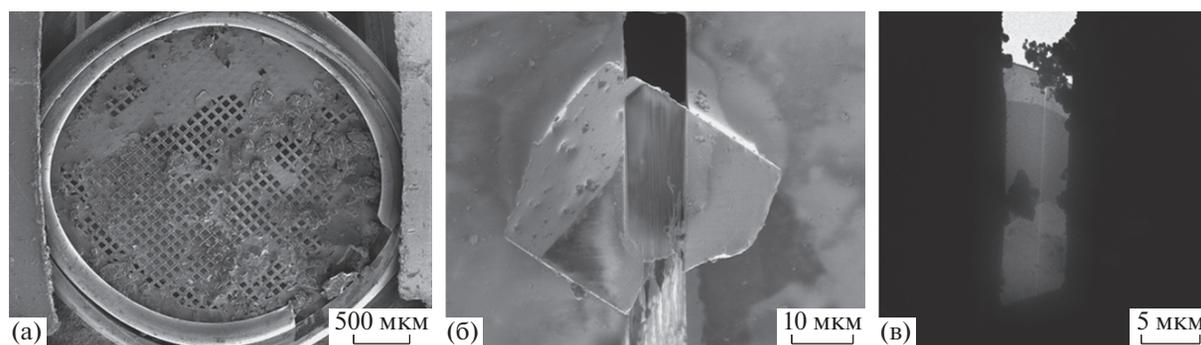


Рис. 2. Подготовка тонкого среза микрокристалла с помощью крио-ФИП: а – общий вид сетки, б – промежуточный этап приготовления тонкого среза кристалла, в – крио-ЭМ-изображение полученного среза.

сталлов расположено непосредственно на медной поддерживающей сетке, что существенно ограничивает выбор возможных областей для проведения эксперимента. Микрокристаллы размером до ~ 50 мкм, расположенные на некотором удалении от медных перемычек сетки, подвергались травлению с помощью крио-ФИП в условиях “скользящего пучка”. Для этого столик с образцом поворачивали на угол 12° – 20° и осуществляли утонение (рис. 1б, 2б) при ускоряющем напряжении 30 кВ с последовательным понижением силы тока от 1 нА до 30 пА для минимизации радиационных повреждений и снижения области аморфизации приповерхностного слоя. Температура столика с образцом в процессе эксперимента составляла -175°C .

ЭЛЕКТРОННАЯ ДИФРАКЦИЯ МИКРОКРИСТАЛЛОВ

Полученные сетки с тонкими срезами микрокристаллов лизоцима (рис. 1в) переносили в крио-ЭМ Titan Krios 60-300, оснащенный корректором сферических aberrаций (Image corrector). Исследование проводили при ускоряющем

напряжении 300 кВ и длине камеры 1.35 м. В результате с помощью камеры Ceta было получено видео дифракционной серии при непрерывном повороте сетки в диапазоне от -45° до 45° и со скоростью вращения держателя $1.5^\circ/\text{с}$. Поток электронов составил ~ 0.2 э/Å²/с, а полная доза ~ 12 э. Набор данных осуществлялся в течение 60 с, частота обновления изображений составила ~ 4 кадра/с.

На рис. 3 приведены примеры дифракционных картин микрокристаллов лизоцима, полученные при наклоне держателя на -30° (рис. 3а), 0° (рис. 3б), 30° (рис. 3в). Наблюдается падение пространственного разрешения дифракционных картин от 2.2 Å при -30° до 2.9 Å при 30° во время съемки поворотной серии, вызванное накоплением радиационных повреждений на облучаемой области.

ЗАКЛЮЧЕНИЕ

Получен набор дифракционных данных методом электронной дифракции микрокристаллов от монокристалла размером ~ 35 мкм. Продемонстрированы основные технические аспекты мето-

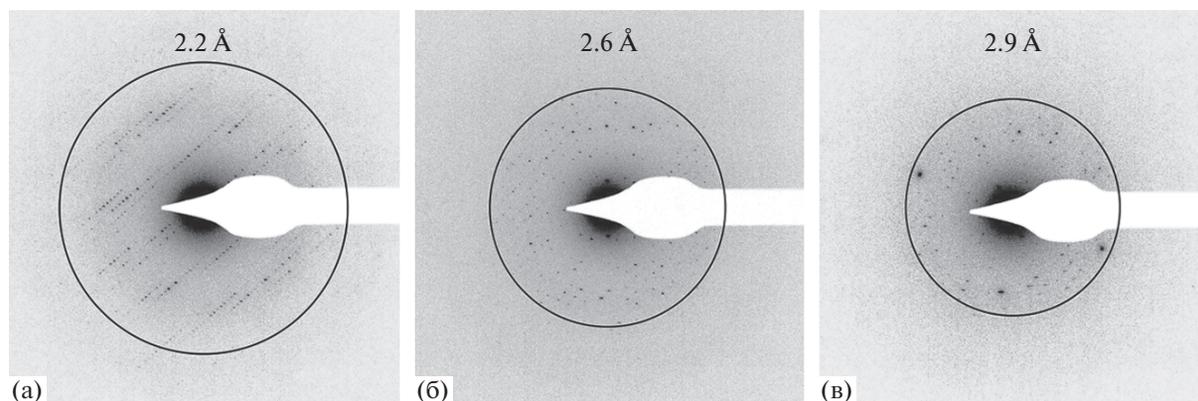


Рис. 3. Электронная дифракция микрокристаллов. Изображения поворотной дифракционной серии, полученные под углами -30° (а), 0° (б), 30° (в).

да – витрификация и препарирование микрокристаллов методом фокусирования ионного травления, а также получение поворотной дифракционной серии в режиме облучения образца малыми дозами электронов с помощью крио-ЭМ. Качество полученных экспериментальных данных указывает на возможность расшифровки структуры лизоцима. Помимо определения структуры трехмерных кристаллов при помощи крио-ЭМ ЭДМ актуальна для предварительной характеристики микрокристаллических суспензий для экспериментов по серийной кристаллографии с использованием синхротронного излучения или лазеров на свободных электронах.

Работа выполнена при финансовой поддержке НИЦ “Курчатовский Институт” (приказ от 02.07.2020 № 1056) и Российского фонда фундаментальных исследований (грант № 19-29-12054).

СПИСОК ЛИТЕРАТУРЫ

- Henderson R. // Q. Rev. Biophys. 1995. V. 28. № 2. P. 171. <https://doi.org/10.1017/S003358350000305X>
- Diederichs K., Wang M. // Protein Crystallography. Methods and Protocols / Eds. Wlodawer A. et al. New York: Humana Press, 2017. P. 239. https://doi.org/10.1007/978-1-4939-7000-1_10
- Sliz P., Harrison S.C., Rosenbaum G. // Structure. 2003. V. 11. № 1. P. 13. [https://doi.org/10.1016/S0969-2126\(02\)00910-3](https://doi.org/10.1016/S0969-2126(02)00910-3)
- McPherson A., Gavira J.A. // Acta Cryst. F. 2014. V. 70. № 1. P. 2. <https://doi.org/10.1107/S2053230X13033141>
- Chapman H.N., Fromme P., Barty A. et al. // Nature. 2011. V. 470. № 7332. P. 73. <https://doi.org/10.1038/nature09750>
- Zatsepin N.A., Li C., Colasurdo P., Nannenga B.L. // Curr. Opin. Struct. Biol. 2019. V. 58. P. 286. <https://doi.org/10.1016/j.sbi.2019.06.004>
- Johansson L.C., Stauch B., Ishchenko A., Cherezov V. // Trends Biochem. Sci. 2017. V. 42. № 9. P. 749. <https://doi.org/10.1016/j.tibs.2017.06.007>
- Kuhlbrandt W. // Science. 2014. V. 343. № 6178. P. 1443. <https://doi.org/10.1126/science.1251652>
- Камышинский Р.А., Чесноков Ю.М., Орехов А.С. // Кристаллография. 2020. Т. 65. № 5. С. 774. <https://doi.org/10.31857/S0023476120050094>
- Fan X., Wang J., Zhang X. et al. // Nat. Commun. 2019. V. 10. № 1. P. 2386. <https://doi.org/10.1038/s41467-019-10368-w>
- Wu M., Lander G.C. // Curr. Opin. Struct. Biol. 2020. V. 64. P. 9. <https://doi.org/10.1016/j.sbi.2020.05.007>
- Gonen T., Cheng Y., Sliz P. et al. // Nature. 2005. V. 438. № 7068. P. 633. <https://doi.org/10.1038/nature04321>
- Glaeser R.M. // J. Ultrastruct. Res. 1971. V. 36. № 3–4. P. 466. [https://doi.org/10.1016/S0022-5320\(71\)80118-1](https://doi.org/10.1016/S0022-5320(71)80118-1)
- Shi D., Nannenga B.L., Iadanza M.G., Gonen T. // eLife. 2013. V. 2. <https://doi.org/10.7554/eLife.01345>
- Nannenga B.L., Shi D., Leslie A.G.W., Gonen T. // Nat. Methods. 2014. V. 11. № 9. P. 927. <https://doi.org/10.1038/nmeth.3043>
- Nannenga B.L., Gonen T. // Curr. Opin. Struct. Biol. 2016. V. 40. P. 128. <https://doi.org/10.1016/j.sbi.2016.09.007>
- Shi D., Nannenga B.L., de la Cruz M.J. et al. // Nat. Protoc. 2016. V. 11. № 5. P. 895. <https://doi.org/10.1038/nprot.2016.046>
- Gemmi M., Mugnaioli E., Gorelik T.E. et al. // ACS Cent. Sci. 2019. V. 5. № 8. P. 1315. <https://doi.org/10.1021/acscentsci.9b00394>
- Hattne J., Reyes F.E., Nannenga B.L. et al. // Acta Cryst. A. 2015. V. 71. № 4. P. 353. <https://doi.org/10.1107/S2053273315010669>
- Nannenga B.L., Gonen T. // Nat. Methods. 2019. V. 16. № 5. P. 369. <https://doi.org/10.1038/s41592-019-0395-x>
- Nguyen C., Gonen T. // Curr. Opin. Struct. Biol. 2020. V. 64. P. 51. <https://doi.org/10.1016/j.sbi.2020.05.018>
- Hattne J., Martynowycz M.W., Penczek P.A., Gonen T. // IUCrJ. 2019. V. 6. № 5. P. 921. <https://doi.org/10.1107/S2052252519010583>
- Maki-Yonekura S., Hamaguchi T., Naitow H. et al. // Microscopy. 2021. V. 70. № 2. P. 232. <https://doi.org/10.1093/jmicro/dfaa052>

СТРУКТУРА МАКРОМОЛЕКУЛЯРНЫХ
СОЕДИНЕНИЙ

УДК 538.911

ПРОСТРАНСТВЕННАЯ СТРУКТУРА РЕКОМБИНАНТНОЙ
ТЕРМОФИЛЬНОЙ РИБОКИНАЗЫ *Thermus species* 2.9
В КОМПЛЕКСЕ С АДЕНОЗИНДИФОСФАТОМ

© 2021 г. В. И. Тимофеев^{1,2}, Ю. А. Абрамчик³, Т. И. Муравьева³, Н. Е. Жухлистова¹,
Р. С. Есипов³, И. П. Куранова^{1,2,*}

¹ Институт кристаллографии им. А.В. Шубникова ФНИЦ “Кристаллография и фотоника” РАН, Москва, Россия

² Национальный исследовательский центр “Курчатовский институт”, Москва, Россия

³ Институт биорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова РАН, Москва, Россия

*E-mail: inna@crys.ras.ru

Поступила в редакцию 08.09.2020 г.

После доработки 22.09.2020 г.

Принята к публикации 24.09.2020 г.

Пространственная структура рекомбинантной термофильной рибокиназы *Thermus species* 2.9 установлена при разрешении 2.4 Å с использованием кристаллов, выращенных в невесомости методом встречной диффузии, и дифракционного набора, полученного на источнике синхротронного излучения SPring-8 (Япония). В активном центре фермента локализована молекула аденозиндифосфата и описано ее окружение. Показано, что пространственная структура рибокиназы *Thermus species* 2.9 весьма сходна со структурами других представителей ферментов этого семейства.

DOI: 10.31857/S0023476121050209

ВВЕДЕНИЕ

Рибокиназа из термофильного штамма бактерий *Thermus species* 2.9 (*Th. sp.* 2.9) относится к семейству углеводных рибокиназ (ЕС 2.7.1.15). Белки этого семейства катализируют фосфорилирование рибозы посредством переноса γ -фосфата аденозинтрифосфата (АТФ) на 5'-гидроксиметильную группу сахара [1, 2]. Реакция протекает в присутствии АТФ и ионов магния [3, 4]. D-рибоза является широко распространенным углеводом, вовлечена в производство энергии, является компонентом РНК и ДНК, входит в состав кофакторов многих ферментов. Через внешнюю клеточную мембрану рибоза диффундирует в периплазму, где связывается с рибозасвязывающими белками, транспортирующими ее через внутреннюю мембрану [5]. Катализируемое рибокиназой фосфорилирование 5'-гидроксиметильной группы является первой ступенью, вовлекающей рибозу в дальнейшие метаболические превращения [6–8]. Продукт реакции D-рибоза–5-фосфат используется для синтеза нуклеотидов, триптофана и гистидина или вовлекается в пентозафосфатный путь. Рибоза-5-фосфат инкорпорируется в АТФ или другие высокоэнергетические соединения. Фосфорибозилпирофосфатсинтетаза катализирует превращение рибоза-5-фосфата в 5-фосфорибозил-1-пирофосфат (5-PRPP), кото-

рый вовлечен в биосинтетические реакции, приводящие к биосинтезу пуриновых и пиримидиновых оснований, кофакторов ферментов и аминокислот гистидина и триптофана [9, 10]. Рибокиназа участвует также в биохимических процессах, обеспечивающих повторное использование освобождающейся при расщеплении нуклеотидов рибозы [11].

Интерес к структуре и свойствам рибокиназ возрос в связи с применением этих белков в каскадной стратегии энзиматического синтеза биологически активных производных нуклеозидов, который протекает при последовательном действии рекомбинантной рибокиназы, фосфопентамутаза, нуклеозидфосфорилазы [12–16]. Многие производные нуклеозидов, модифицированные по гетероциклическому основанию и углеводной группе, являются важными терапевтическими препаратами против вирусных инфекций и злокачественных опухолей [17, 18]. Широкое использование каскадной стратегии ограничено высокой субстратной специфичностью используемых ферментов и их недостаточной стабильностью. Часть этих ограничений можно обойти, используя ферменты из термофильных организмов. Ферменты термофильных микроорганизмов менее чувствительны к структуре субстрата, активны при температуре 70–80°C и существенно повышают скорость реакции, увеличивая раствори-

Таблица 1. Статистические характеристики экспериментального набора рентгеновских данных и результаты уточнения структуры рибокиназы *Th. sp. 2.9*

Обработка набора	
Пространственная группа	$P2_1$
$a, b, c, \text{Å}$	44.2, 155.6, 83.0
β , град	98.2
Разрешение, Å	30.0–2.40 (2.53–2.40)*
Количество независимых рефлексов	41928
Полнота набора, %	96.9 (97.7)
$I/\sigma(I)$	5.29 (4.47)
$R_{\text{mgd-F}}$, %	7.7 (14.0)
Уточнение	
PDB_ID	6ZNX
Разрешение, Å	30.00–2.40 (2.46–2.40)
Количество рефлексов	39758
R_{crist} , %	23.0
R_{free} , %	27.1
Количество атомов	8553
B -фактор, Å ²	
Средний	36.21
RMS	
По длинам связей, Å	0.007
По углам, град	1.635
Карта Рамачандрана	
Наиболее благоприятные области	97.1%
Допустимые области	2.9%
Запрещенные области	0.0%

* В скобках приведены значения для последнего слоя.

мость гетероциклических субстратов [19–21]. В [22] было показано, что рекомбинантная рибокиназа из термофильного штамма *Th. sp. 2.9* как катализатор первой ступени реакции каскадного синтеза имеет ряд преимуществ. Фермент имеет максимум активности при 85°C, углеводным субстратом данной рибокиназы кроме D-рибозы может служить 2-дезоксид-рибоза. Знание пространственной структуры дает возможность рационального конструирования новых мутантных форм фермента с более широкой селективностью, наиболее удобных для биотехнологического применения. В настоящей работе при разрешении 2.4 Å установлена пространственная структура рекомбинантной термофильной рибокиназы *Th. sp. 2.9* в комплексе с аденозиндифосфатом (АДФ). Структура решена с использованием дифракционного набора, полученного на синхротроне SPring-8 (Япония) для кристаллов, выра-

щенных в невесомости методом встречной диффузии. Показано, что по типу укладки полипептидной цепи фермент сходен с другими белками семейства рибокиназ.

МАТЕРИАЛЫ И МЕТОДЫ

Кристаллизация и сбор дифракционных данных. Для кристаллизации использовали препарат рекомбинантной рибокиназы, полученной и очищенной по методике [21]. Кристаллы были выращены в невесомости методом встречной диффузии из растворов белка с концентрацией 11.7 мг/мл в 0.25 М Трис-буфере, pH 8.0, содержащем 50 мМ KCl, 1 мМ АТФ, 5 мМ MgCl₂, 5% глицерина, 0.04% NaN₃ [23]. Концентрация сульфата аммония в осадителе (Na-цитратный буфер, pH 5.5, содержащий 45 мМ KCl, 0.1 М Na-цитрата, pH 5.5, 5 мМ MgCl₂, 5 мМ АТФ, 0.04% NaN₃) была увеличена до 15.5% в отличие от концентрации, используемой в [23].

Дифракционный набор для кристалла, предварительно замороженного в токе жидкого азота, собран при 100 К на станции BL41XU синхротрона SPring-8 (Япония), оснащенной детектором EIGER. Для сбора данных использовали метод вращения, длина волны составляла 0.8 Å, угол вращения 360°, угол качания 0.1°, расстояние кристалл–детектор 400 мм. Набор был обработан с использованием программы iMosflm [24]. Кристаллы фермента описываются пр. гр. $P2_1$, содержат четыре молекулы в независимой части ячейки, $a = 44.210$, $b = 155.550$, $c = 83.040$ Å, $\beta = 98.20^\circ$. Статистика набора представлена в табл. 1.

Решение и уточнение структуры. Структура рибокиназы *Th. sp. 2.9* в комплексе с АДФ определена методом молекулярного замещения с помощью программы BALBES [25]. В качестве стартовой модели использованы координаты рибокиназы *Vibrio cholerae* O395 в комплексе с рибозой, АДФ и ионом натрия (PDB_ID: 4XDA) [26].

Уточнение структуры проведено по программе REFMAC [27]. Ручную правку модели осуществляли с помощью программы Coot [28], используя карты электронной плотности, рассчитанные с коэффициентами $2|F_o| - |F_c|$ и $|F_o| - |F_c|$. Также с помощью программы Coot с использованием карт электронной плотности были локализованы молекулы воды и АДФ. Молекула АДФ, вписанная в карту электронной плотности, приведена на рис. 1.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Кристаллы для определения пространственной структуры рекомбинантной рибокиназы из термофильного штамма бактерий *Th. sp. 2.9* были выращены в невесомости методом встречной диффузии по технологии [29] из растворов белка,

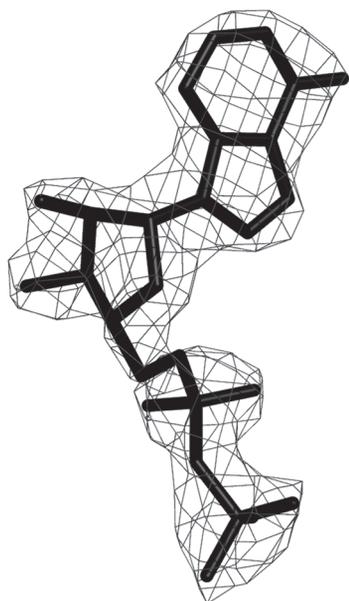


Рис. 1. Электронная плотность молекулы АДФ, рассчитанная с коэффициентами $|F_o| - |F_c|$ при 2σ .

содержащих ионы магния и АТФ после оптимизации концентрации осадителя в условиях [23]. В результате разрешение дифракционного набора выращенных кристаллов было повышено с 2.87 до 2.4 Å.

Пространственная структура рибозиназы *Th. sp. 2.9* решена методом молекулярного замещения с использованием в качестве стартовой модели координат рибозиназы *Vibrio cholerae* (PDB_ID: 4XDA) и уточнена при разрешении 2.4 Å. Статистические параметры набора и результаты уточнения приведены в табл. 1. Кристаллы фермента (пр. гр. $P2_1$) содержат четыре молекулы белка (A, B, C, D) в независимой части элементарной ячейки (рис. 2). В каждой молекуле фермента локализована связанная в активном центре молекула АДФ с заселенностью, близкой к единице.

При совмещении по C α -атомам четырех молекул фермента, расположенных в независимой части ячейки, величина среднеквадратичного отклонения между парами молекул A–B, A–C, A–D равна соответственно 0.587, 0.760 и 0.570 Å. Значительные отклонения заметны на участках, соединяющих большой и малый домены, в сегментах, соединяющих стрэнды малого домена, и в петле, соединяющей спирали α_5 и α_6 .

Укладка полипептидной цепи в молекуле рибозиназы *Th. sp. 2.9*. Пространственные структуры известны для ряда белков семейства рибозиназ, в том числе рибозиназы человека и ряда микробных рибозиназ: из *E. coli*, *Vibrio cholera* O395, *Leishmania donovani*, *Thermotoga maritima* [26, 30,

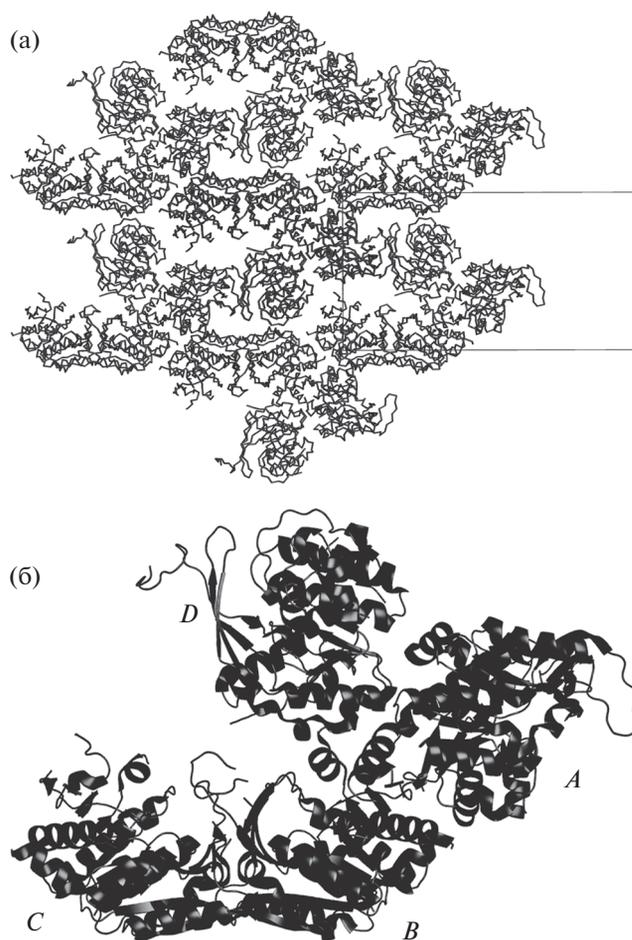


Рис. 2. Укладка мономеров в решетке кристаллов комплекса рибозиназы *Th. sp. 2.9* с АДФ (а); независимая часть элементарной ячейки кристаллов комплекса рибозиназы *Th. sp. 2.9* с АДФ – четыре мономера белка A, B, C, D (б).

31]. Показано, что, хотя идентичность аминокислотных последовательностей между белками этого семейства менее 30%, рибозиназы сходны типом укладки полипептидной цепи, имеют два общих мотива в аминокислотной последовательности, а субъединица фермента состоит из двух доменов. Еще одной чертой, общей для большинства рибозиназ с известной структурой, является димерное строение биологически активной формы фермента.

Типичная для рибозиназ укладка полипептидной цепи обнаружена и в молекуле термофильной рибозиназы *Th. sp. 2.9*. Полипептидная цепь молекулы рибозиназы *Th. sp. 2.9*, содержащая 301 аминокислотный остаток (а.о.), включает в себя 13 β -стрэндов, 10 α -спиралей и 3 η -спирали (η -спирали 3_{10}). Как и во всех белках семейства рибозиназ, молекула состоит из двух доменов: большого каталитического глобулярного α/β -домена и выступающего из него малого β -домена,

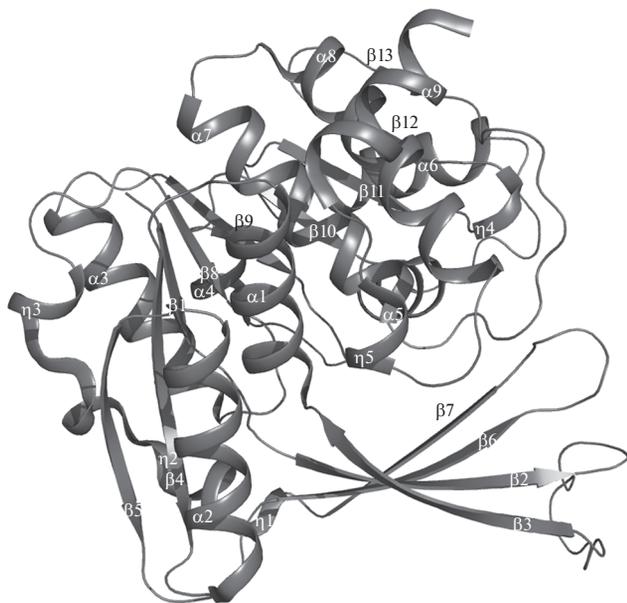


Рис. 3. Укладка полипептидной цепи в мономере комплекса рибокиназы *Th. sp. 2.9* с АДФ. Элементы вторичной структуры обозначены стрелками (β -стрэнды) и спиралями: α и 3_{10} (η) – спирали.

который представляет собой изогнутый четырех-стрэндовый β -слой (рис. 3). Малый β -домен образован двумя вставками в центральную α/β -укладку. Каждая из вставок состоит из пары антипараллельных β -стрэндов $\beta 2$ – $\beta 3$ (а.о. 8–36) и стрэндов $\beta 6$ – $\beta 7$ (а.о. 90–108). Большой домен состоит из центрального β -слоя, образованного девятью β -стрэндами, которые с обеих сторон окружены α -спиралями. Первые шесть параллельных β -стрэндов домена вместе с примыкающими α -спиралями имеют топологию классического NAD-связывающего домена Россмана [32], характерную для белков, связывающих нуклеотиды. Однако в рибокиназе к этому слою примыкают еще три стрэнда, соединенные друг с другом короткими поворотами. Три С-концевые β -стрэнда, продолжающие центральный слой ($\beta 11$ – $\beta 13$), антипараллельны и соединены короткими поворотами типа шпильки. Обе стороны С-концевого участка большого домена закрыты α -спиралями 6 и 7 с одной стороны и 8 и 9 с другой. Большой и малый домены соединены короткими неупорядоченными участками, которые дают возможность доменам двигаться друг относительно друга. На коротких переходных участках, соединяющих стрэнд $\beta 1$ большого домена со стрэндом $\beta 2$ малого домена или стрэнд $\beta 3$ малого домена со спиралью $\alpha 1$ большого домена, повышенное содержание остатков глицина и пролина, они входят в один из двух общих для всего семейства мотивов аминокислотных последовательностей [11]. Остатки глицина востребованы по стерическим и

конформационным причинам. Они увеличивают пространство для возможного изменения взаимного расположения доменов, а остатки пролина способствуют изменению направления полипептидной цепи. Повышенное содержание пролина и глицина наблюдается и в переходных между доменами участках $\beta 7$ – $\eta 2$ и $\beta 5$ – $\beta 6$. Температурные факторы а.о. фрагмента 85–89, соединяющего стрэнды $\beta 5$ и $\beta 6$, и фрагмента Pro108–Pro116 между стрэндом $\beta 7$ и спиралью $\eta 2$ выше среднего по структуре.

Для сравнения на рис. 4 представлены последовательности и распределение элементов вторичной структуры в мономерах рибокиназы *Th. sp. 2.9* и *E. coli*. Степень идентичности аминокислотных последовательностей между этими ферментами превышает 40%. При совмещении по $\text{C}\alpha$ -атомам субъединицы А рибокиназы *Th. sp. 2.9* с субъединицей *E. coli* среднеквадратичное отклонение равно 1.445 Å. Распределение элементов вторичной структуры, а также количество а.о. в α -спиралях и β -стрэндах практически одинаковы в обоих белках.

Активный центр фермента, включающий в себя центры связывания рибозы и АДФ, расположен на участке поверхности каталитического α/β -домена, находящегося между большим и малым доменами (рис. 5). Известно, что катализируемая ферментом реакция протекает посредством нуклеофильной атаки атомов кислорода $\text{O}5'$ гидроксиметильных групп рибозы γ -фосфата АТФ [11]. Нуклеофильные свойства кислорода гидроксиметильной группы рибозы увеличиваются благодаря водородной связи между гидроксильной группой и остатком аспарагиновой кислоты, действующей как каталитическое основание. В молекуле рибокиназы *E. coli* – это Asp255, в рибокиназе *Th. sp. 2.9* – Asp248. Атомы азота основной цепи остатков глицина, окружающие остаток аспарагиновой кислоты, участвуют в образовании анионной дыры, стабилизирующей переходное состояние реакции [11]. Сближение субстратов, необходимое для протекания реакции, – результат перехода фермента из открытого в закрытое состояние сразу после связывания рибозы [30]. Конформационные изменения, обеспечивающие этот переход, сопровождаются движением доменов. В закрытой конформации малый домен изолирует от растворителя область активного центра, в котором связана рибоза. Поэтому малому домену приписывают роль крышки, контролирующей доступ растворителя к активному центру.

О строении участка молекулы, в котором происходит связывание рибозы, можно судить, сравнивая структуру молекулы рибокиназы *Th. sp. 2.9* и структуру комплекса рибокиназы *E. coli* с рибозой и АДФ [11]. Связывание рибозы имеет место

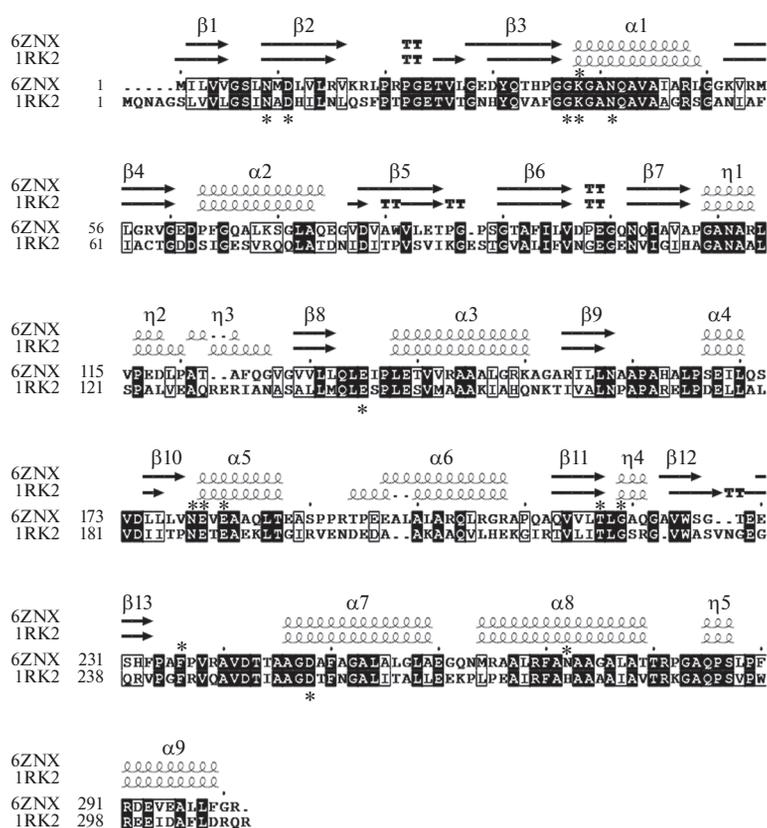


Рис. 4. Сравнение аминокислотных последовательностей и вторичных структур рибокиназ *Th. sp. 2.9* и *E. coli*. Элементы вторичной структуры α и 3_{10} (η) спирали и β -стрэнды обозначены спиралями и стрелками соответственно. Черным цветом выделены идентичные в обоих белках аминокислотные остатки, звездочками — аминокислотные остатки, участвующие в связывании молекул рибозы и АДФ.

на поверхности большого домена, расположенной наиболее близко к сегменту, связывающему большой и малый домены. В прямых контактах с рибозой в рибокиназе *E. coli* участвуют а.о. стрэнда $\beta 1$ Asn14, Asp16 (Asn9, Asp11 в *Th. sp. 2.9*), Gly42 (Gly37 в *Th. sp. 2.9*), расположенные между $\beta 3$ и $\alpha 1$, остатки спирали $\alpha 1$ Lys43 и Asn46 (Lys38 и Asn41 в *Th. sp. 2.9*), расположенные между $\alpha 8$ и $\alpha 3$, и Asp255 (Asp248 в *Th. sp. 2.9*) спирали $\alpha 7$. Почти все перечисленные а.о. относятся к α/β -домену и взаимодействуют с полярной стороной рибозного кольца. Гидрофобная поверхность сахарного кольца взаимодействует в основном с алифатическими боковыми цепями внутренней стороны малого домена. Отметим, что в рибокиназе рибоза связана с областью большого домена, который имеет топологию Россмана, и в большинстве белков связываются нуклеотиды. Все взаимодействующие с рибозой а.о. инвариантны не только в структурах рибокиназ из *E. coli* и *Th. sp. 2.9*, но и входят в один из аминокислотных мотивов, характеризующих семейство рибокиназ [11, 30]. Хотя некоторые рибокиназы, в том числе рибокиназы *Th. sp. 2.9* и *E. coli*, могут использовать дезоксирибозу как

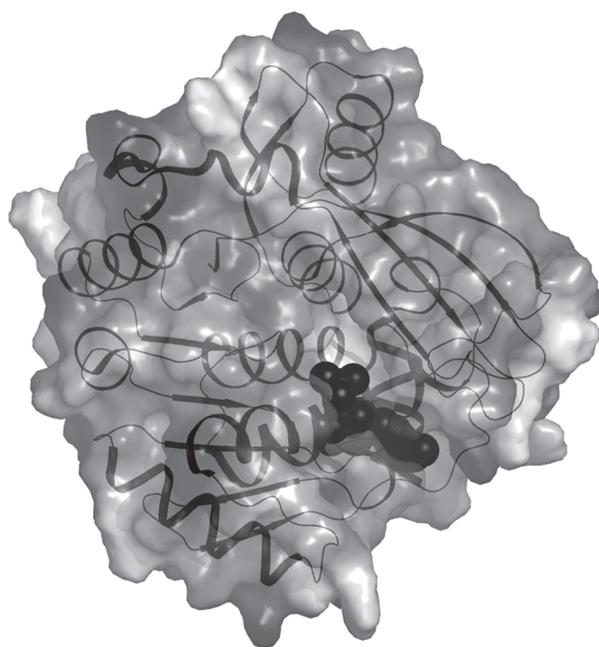


Рис. 5. Молекула АДФ в активном центре субъединицы рибокиназы *Th. sp. 2.9*.

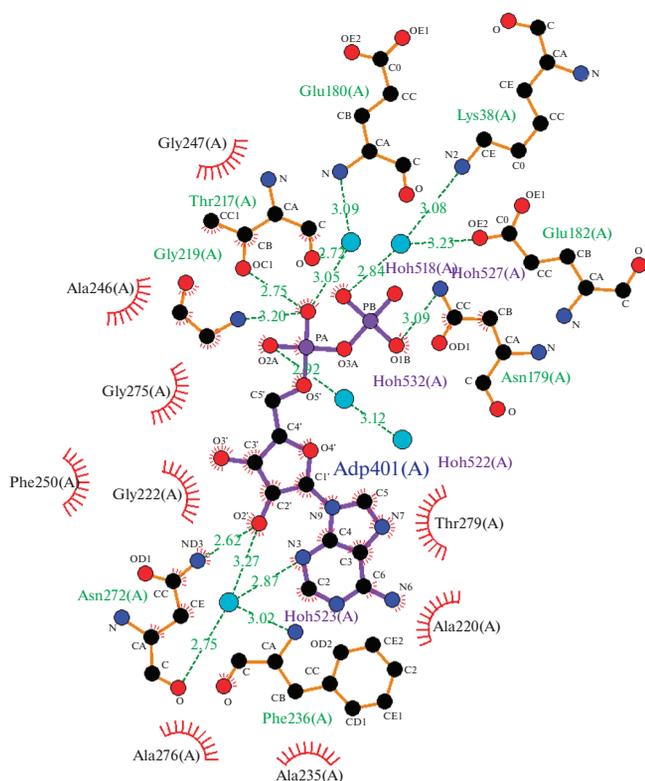


Рис. 6. Ближайшее окружение молекулы АДФ в активном центре рибокиназы *Th. sp. 2.9* с АДФ.

субстрат [33], в рибокиназе *E. coli* это приводит к потере трех водородных связей с а.о. Asn14, Asp16, Gly42 (Asn9, Asp11, Gly37 в *Th. sp. 2.9* соответственно), что объясняет предпочтительность рибозы как субстрата. В рибокиназе из *E. coli* связанная рибоза почти полностью закрыта малым доменом, который действует как крышка над центром связывания, а также второй субъединицей димера.

Молекула АДФ, локализованная в молекуле рибокиназы *Th. sp. 2.9*, образовалась в результате гидролиза АТФ, присутствующего в кристаллизационном растворе, и занимает в активном центре положение, схожее с положением молекулы АДФ в субъединице рибокиназы *E. coli*. Молекула АДФ локализована в полости между двумя доменами, но непосредственно связана только с а.о., находящимися на поверхности α/β -домена (рис. 5). По сравнению с рибозасвязывающим центром АДФ-связывающий участок находится на большем расстоянии от переходного участка между двумя доменами сегмента. Поэтому при переходе к закрытой конформации фермента молекула АДФ не накрывается полностью малым доменом и остается доступной для растворителя. В связывании молекулы АДФ участвуют а.о. нескольких фрагментов вторичной структуры: пяти спиралей ($\alpha 1, \alpha 5, \alpha 7, \alpha 8, \alpha 9$) и двух стрэндов ($\beta 12$ и $\beta 13$). Осно-

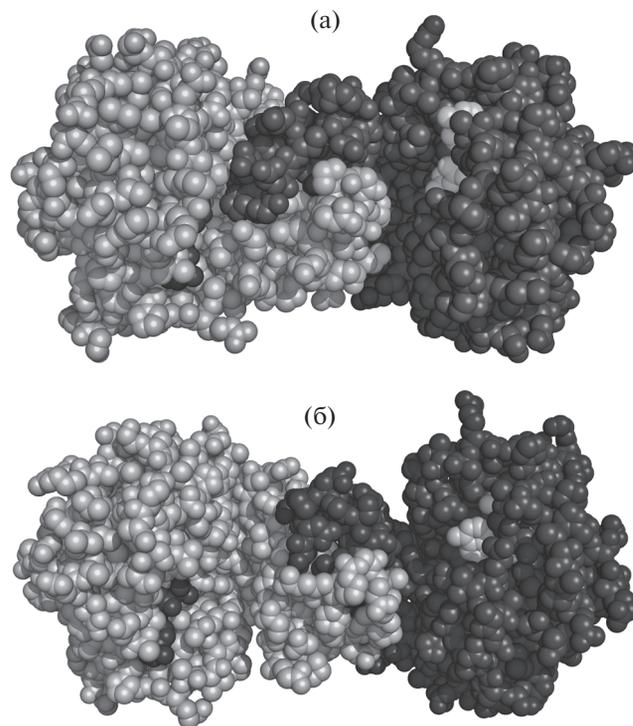


Рис. 7. Молекулы рибокиназы, построенные из мономеров *D, A* (а) и *B, C* (б). В активных центрах мономеров показаны молекулы АДФ.

вание аденина локализовано в петле, соединяющей стрэнд $\beta 13$ со спиралью $\alpha 7$ (а.о. 235–245) (рис. 5). Оксигруппы рибозного фрагмента расположены по одну сторону рибозного кольца. Оксигруппа при втором атоме углерода образует прямую водородную связь с атомом OD2_Asp272 спирали $\alpha 8$, а через молекулу воды — с азотом основной цепи Phe236 стрэнда $\beta 13$. Вторая гидроксильная группа рибозы связана с атомом кислорода основной цепи Gly222 стрэнда $\beta 12$. Фосфатные группы молекулы АДФ направлены в сторону центра связывания рибозы. Атомы кислорода α -фосфатной группы связаны непосредственно с атомом азота основной цепи Gly219 и OG1_Thr217 (стрэнд $\beta 11$), а через молекулу H_2O — с атомом азота основной цепи Glu180 спирали $\alpha 5$. Один из атомов кислорода β -фосфата связан с атомом азота амидогруппы Asn79, который находится между стрэндом $\beta 10$ и спиралью $\alpha 5$, а через молекулу воды — с OE2_Glu182 ($\alpha 5$) и Ne_Lys38 ($\alpha 1$) (рис. 6). Почти все а.о., непосредственно участвующие в связывании АДФ в рибокиназах *Th. sp. 2.9* и *E. coli*, инвариантны в обоих белках. Имеется только одна замена. Со второй оксигруппой рибозного кольца в рибокиназе *Th. sp. 2.9* связан остаток Asp272, а в рибокиназе *E. coli* соответствующую позицию занимает His279.

Особенности пространственной структуры рибокиназы Th. sp. 2.9. Несмотря на значительное

сходство пространственных структур рибоскиназы *Th. sp. 2.9* и *E. coli*, оба фермента различаются четвертичной структурой. Фермент из *E. coli*, как и большинство рибоскиназ с известной пространственной структурой, функционирует как димер, в котором субъединицы связаны посредством взаимодействия малых доменов — их β -стрэнды образуют связь, называемую β -claspstructure [34]. В кристаллах рибоскиназы *E. coli* (пр. гр. $P6_122$) в независимой части элементарной ячейки содержится один мономер, а димер, являющийся функционально активной формой фермента, образуется при действии оси второго порядка. Граница раздела между молекулами димера формируется вследствие взаимодействий между β -слоями малых доменов. Упакованные почти перпендикулярно друг другу, они образуют сплюснутый баррель. Такая структура была названа β -claspstructure [34]. Непосредственно в катализе обе субъединицы димера функционируют независимо, однако при переходе фермента в закрытое состояние вторая субъединица димера закрывает часть активного центра, где связывается рибоза, что обеспечивает надежную изоляцию активного центра от растворителя во время реакции. Анализ расположения молекул в независимой части элементарной ячейки рибоскиназы *Th. sp. 2.9* показывает, что связь между протомерами через малые домены отсутствует. Однако анализ кристаллической упаковки, проведенный с использованием сервиса PISA [35], показывает, что димер может быть образован путем объединения субъединиц *A* и *D* после применения оператора $-x, y - 1/2, -z + 1$ или путем объединения субъединиц *B* и *C* после применения оператора $x - 1, y, z$. Данные димеры структурно аналогичны (рис. 7), площади контактирующих поверхностей, как и энергия формирования комплекса, согласно сервису PISA близки (3810 \AA^2 и -28.9 ккал/моль и 3410 \AA^2 и -29.9 ккал/моль соответственно), однако количество водородных связей различается (табл. 2). Четыре молекулы рибоскиназы *Th. sp. 2.9*, расположенные в независимой части элементарной ячейки, несколько различаются взаимным расположением большого и малого доменов. В молекулах *A, B, C, D* расстояние между α -атомами а.о. Gly100 малого домена и Val241 большого домена равно 7.42, 12.33, 18.1 и 5.71 \AA соответственно. Оба а.о. расположены на противоположных участках полипептидной цепи, ограничивающей активный центр, и могут характеризовать степень его открытости. Большая подвижность а.о. малого домена в рибоскиназе *Th. sp. 2.9* может вносить вклад в различия кинетических параметров между рибоскиназами *Th. sp. 2.9* и *E. coli*, как это показано в [22].

Работа выполнена при поддержке ОКР “МКС (Наука)” в рамках Федеральной космической программы 2016–2025 в части решения и уточне-

Таблица 2. Контакты между мономерами *D* и *A*, связанными оператором симметрии $-x, y - 1/2, -z + 1$, и мономерами *B* и *C*, связанными оператором симметрии $x - 1, y, z$

Мономер <i>D</i> (x, y, z)	Мономер <i>A</i> ($-x, y - 1/2, -z + 1$)	Расстояние, \AA
N_GLU24	OE1_GLN103	3.19
OH_TYR31	O_VAL106	2.72
N_ILE104	O_GLU24	3.24
N_VAL106	O_VAL26	3.04
O_ARG18	N_VAL106	3.44
O_PRO20	N_ILE104	3.28
OH_TYR31	N_ALA107	3.71
OH_TYR31	N_VAL106	3.70
O_ASN102	N_GLY23	2.57
OE1_GLN103	NE2_GLN103	2.71
O_ILE104	N_VAL26	3.11
O_VAL106	N_GLY28	2.68
Мономер <i>B</i> (x, y, z)	Мономер <i>C</i> ($x - 1, y, z$)	Расстояние, \AA
N_VAL26	O_ILE104	3.36
N_GLY28	O_VAL106	3.05
OE1_GLN103	NE2_GLN103	2.83
O_GLU24	N_ILE104	3.68

ния структуры белка, а также Министерства науки и высшего образования РФ в рамках Государственного задания ФНИЦ “Кристаллография и фотоника” РАН в части анализа пространственной структуры.

СПИСОК ЛИТЕРАТУРЫ

1. Mathews I., Erion M.D., Ealick S.E. // Biochemistry. 1998. V. 37. P. 15607. <https://doi.org/10.1021/bi9815445>
2. Schumacher M.A., Scott D.M., Mathews I.I. et al. // J. Mol. Biol. 2000. V. 296. P. 549. <https://doi.org/10.1006/jmbi.1999.3474>
3. Schimmel S.D., Hoffee P., Horecker B.L. // Arch. Biochem. Biophys. 1974. V. 164. P. 560. [https://doi.org/10.1016/0003-9861\(74\)90067-8](https://doi.org/10.1016/0003-9861(74)90067-8)
4. Andersson C., Mowbray S. // J. Mol. Biol. 2002. V. 315. P. 409. <https://doi.org/10.1006/jmbi.2001.5248>
5. Furlong C.E. // Methods Enzymol. 1982. V. 90. P. 467. [https://doi.org/10.1016/S0076-6879\(82\)90173-2](https://doi.org/10.1016/S0076-6879(82)90173-2)
6. Park J., Gupta R.S. // Cell. Mol. Life Sci. 2008. № 65. P. 2875. <https://doi.org/10.1007/s00018-008-8123-1>
7. Anderson A., Cooper R.A. // Biochim. Biophys. Acta. 1969. V. 177. P. 163. [https://doi.org/10.1016/0304-4165\(69\)90081-6](https://doi.org/10.1016/0304-4165(69)90081-6)

8. *Lopilato J.E., Garwin J.L., Emr S.D. et al.* // J. Bacteriol. 1984. V. 158 P. 665.
<https://doi.org/10.1128/JB.158.2.665-673.1984>
9. *Switzer R.L.* // J. Biol. Chem. 1969. V. 244. P. 2854.
10. *Becker M.A., Raivio K.O., Bakay B. et al.* // J. Clin. Invest. 1980. V. 65. P. 109.
<https://doi.org/10.1172/JCI109640>
11. *Sigrell J.A., Cameron A.D., Jones T.A., Mowbray S.L.* // Structure. 1998. V. 6. P. 183.
[https://doi.org/10.1016/s0969-2126\(98\)00020-3](https://doi.org/10.1016/s0969-2126(98)00020-3)
12. *Chuvikovskiy D.V., Esipov R.S., Skoblov Y.S. et al.* // Bioorg. Med. Chem. 2006. V. 14. P. 6327.
<https://doi.org/10.1016/j.bmc.2006.05.057>
13. *Miroshnikov A.I., Esipov R.S., Muravyova T.I. et al.* // Open Conf. Proc. J. 2010. V. 1. P. 98.
14. *Fateev I.V., Antonov K.V., Konstantinova I.D. et al.* // Beilstein J. Org. Chem. 2014. V. 10. P. 1657.
<https://doi.org/10.3762/bjoc.10.173>
15. *Stepchenko V.A., Seela F., Esipov R.S. et al.* // Synth. Lett. 2012. V. 23(10). P. 1541.
<https://doi.org/10.1055/s-0031-129067916>
16. *Konstantinova I.D., Antonov K.V., Fateev I.V. et al.* // Synthesis. 2011. V. 10. P. 1555.
<https://doi.org/10.1055/s-0030-1260010>
17. *Mikhailopulo I.A., Miroshnikov A.I.* // Biotechnol. Acta. 2013. V. 6. P. 3.
18. *Mikhailopulo I.A., Miroshnikov A.I.* // Acta Naturae. 2010. V. 2. P. 6.
<https://doi.org/10.32607/20758251-2010-2-2-36-58>
19. *Zhou X., Mikhailopulo I.A., Cruz-Bournazou N., Neubauer P.* // J. Mol. Catalysis B: Enzymatic. 2015. V. 115. P. 119.
<https://doi.org/10.1016/j.molcatb.2015.02.009>
20. *Zhou X., Szeker K., Jiao L.-Y. et al.* // Adv. Synthesis Catalysis. 2015. V. 357. № 6. P. 1237.
<https://doi.org/10.1002/adsc.201400966>
21. *Esipov R.S., Abramchik Yu.A., Fateev I.V. et al.* // Acta Naturae. 2016. V. 8. № 3 (30). P. 160.
22. *Esipov R.S., Abramchik Yu.A., Fateev I.V. et al.* // Biofarm. Z. 2016. V. 8(2). P. 3.
23. *Абрамчик Ю.А., Тимофеев В.И., Муравьева Т.И. и др.* // Кристаллография. 2016. Т. 61. № 6. С. 940.
<https://doi.org/10.7868/S0023476116060023>
24. *Battye T.G., Kontogiannis L., Johnson O. et al.* // Acta Cryst. D. 2011. V. 67. № 4. P. 271. <https://doi.org/https://doi.org/10.1107/S0907444910048675>
25. *Long F., Vagin A.A., Young P., Murshudov G.N.* // Acta Cryst. D. 2008. V. 64. № 1. P. 125.
<https://doi.org/10.1107/S0907444907050172>
26. *Paul R., Patra M.D., Sen U.* // Adv. Exp. Med. Biol. 2015. V. 842. P. 293.
<https://doi.org/10.1107/S2053230X18000109>
27. *Murshudov G.N., Vagin A.A., Dodson E.J.* // Acta Cryst. D. 1997. V. 53. P. 240.
<https://doi.org/10.1107/S0907444996012255>
28. *Emsley P., Cowtan K.* // Acta Cryst. D. 2004. V. 60. P. 2126.
<https://doi.org/10.1107/S0907444904019158>
29. *Куранова И.П., Смирнова Е.А., Абрамчик и др.* // Кристаллография. 2011. Т. 56. № 5. С. 944.
<https://doi.org/10.1134/S1063774511050154>
30. *Sigrell J.A., Cameron A.D., Mowbray S.L.* // J. Mol. Biol. 1999. V. 290. P. 1009.
<https://doi.org/10.1006/jmbi.1999.2938>
31. *Gatreddi S., Vasudevan D., Qureshi I.A.* // Int. J. Biol. Macromol. 2019. V. 136. P. 253.
<https://doi.org/10.1016/j.ijbiomac.2019.06.001>
32. *Rossmann M.G., Moras D., Olsen K.W.* // Nature. 1974. V. 250. P. 194199.
33. *Jargiello P.* // Biochim. Biophys. Acta. 1982. V. 698. P. 78.
[https://doi.org/10.1016/0167-4781\(82\)90187-7](https://doi.org/10.1016/0167-4781(82)90187-7)
34. *Chothia C., Janin J.* // Biochemistry. 1982. V. 21. P. 3955.
<https://doi.org/10.1021/bi00260a009>
35. *Krissinel E., Henrick K.* // J. Mol. Biol. 2007. V. 372. P. 774.
<https://doi.org/10.1016/j.jmb.2007.05.022>

СТРУКТУРА МАКРОМОЛЕКУЛЯРНЫХ СОЕДИНЕНИЙ

УДК 544.1+548.737+577.1+577.3

КРИСТАЛЛИЗАЦИЯ В УСЛОВИЯХ МИКРОГРАВИТАЦИИ И СТРУКТУРА АТОМНОГО РАЗРЕШЕНИЯ УРИДИНФОСФОРИЛАЗЫ ИЗ *Vibrio cholerae*

© 2021 г. П. А. Эйстрих-Геллер¹, С. В. Рубинский¹, В. Р. Самыгина^{1,2}, А. Г. Габдулхаков¹,
М. В. Ковальчук^{1,2}, А. С. Миронов³, А. А. Лашков^{1,*}

¹ Институт кристаллографии им. А.В. Шубникова ФНИЦ “Кристаллография и фотоника” РАН, Москва, Россия

² Национальный исследовательский центр “Курчатовский институт”, Москва, Россия

³ Институт молекулярной биологии им. В.А. Энгельгардта РАН, Москва, Россия

*E-mail: alashkov83@gmail.com

Поступила в редакцию 08.09.2020 г.

После доработки 22.09.2020 г.

Принята к публикации 28.09.2020 г.

Уридинфосфорилазы известны в качестве одной из ключевых мишеней при разработке новых противоопухолевых и антипаразитарных препаратов. В условиях микрогравитации на Международной космической станции методом встречной диффузии в капиллярах выращены кристаллы уридинфосфорилазы из патогенной бактерии *Vibrio cholerae*. Определена ее пространственная структура атомного (1.04 Å) разрешения (RCSB ID PDB: 6Z9Z). Впервые в пространственной структуре уридинфосфорилазы в отсутствие связанных специфических лигандов обнаружены альтернативные конформации протяженных областей (β-стрендов и смежных петель) белковой молекулы. Эти альтернативные конформации, возможно, связаны с функционированием фермента. Методами конформационного анализа с использованием марковских моделей показано, что переход между конформациями возможен и без связывания фермента с лигандом.

DOI: 10.31857/S0023476121050052

ВВЕДЕНИЕ

При развитии опухолевых заболеваний в клетках человека повышается потребность в азотистых пиримидиновых основаниях, к которым относятся уридин и тимидин. Азотистые основания можно получить с помощью обратимой каталитической реакции фосфоролитического расщепления пиримидиновых нуклеозидов. Катализатором этой реакции выступает уридинфосфорилаза (UPh). Экспрессия этого белка возрастает в опухолевых клетках [1–3], поэтому действие некоторых противоопухолевых лекарственных препаратов направлено на регуляцию метаболизма азотистых оснований и их производных. К одному из типов таких препаратов относят лекарства, основанные на ингибировании фермента UPh. Эти лекарства регулируют метаболизм азотистых оснований, а именно уридина и тимидина, в ресинтезе которых принимает участие UPh.

В настоящее время активно изучается энзиматическая специфичность UPh, разрабатываются обратимые конкурентные ингибиторы, снижение концентрации которых приводит к восстановлению активности UPh. Изучены и описаны структуры белка и его комплексов из разных организ-

мов с разнообразными лигандами с разрешением до 1.2 Å [4–13].

Для разработки новых лекарственных препаратов – ингибиторов UPh – необходимо понимание механизмов функционирования белка, чему способствует более точное исследование структуры белка и его комплексов методом рентгеноструктурного анализа (РСА). Как показывает практика авторов работы, анализ структур атомного и субатомного разрешения позволяет выявить не только структурные особенности белка и его комплексов, но и с помощью анализа альтернативных конформаций, анизотропных температурных факторов и системы протонирования активного центра приблизиться к экспериментальному определению механизма функционирования фермента как биологической машины [6, 10, 11].

Одним из перспективных методов кристаллизации макромолекул является кристаллизация в условиях микрогравитации. Согласно [14–16] кристаллы, выращенные в невесомости, имеют более упорядоченную внутреннюю структуру, более низкую степень двойникования и лучшую рассеивающую способность по сравнению с кристаллами, выращенными в наземных условиях.

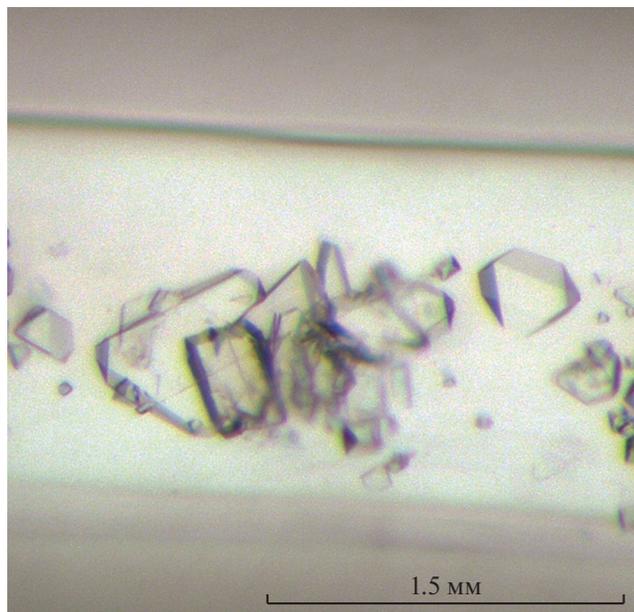


Рис. 1. Кристаллы уридинфосфорилазы из *V. cholerae*, выращенные в условиях микрогравитации.

В работе представлены результаты РСА кристаллов уридинфосфорилазы из *Vibrio cholerae* (*VchUPh*), выращенных в условиях микрогравитации на Международной космической станции (МКС). Приведен анализ обнаруженных альтернативных конформаций протяженных элементов (β -стрендов и петель) третичной структуры белка при отсутствии связанных специфических лигандов. Подобные альтернативные конформации в структурах UPh атомного разрешения наблюдались лишь в ее комплексах с функционально значимыми лигандами, имеющими неполную заселенность [6, 10, 11].

МАТЕРИАЛЫ И МЕТОДЫ

Выделение и очистка *VchUPh*. Очистку препарата *VchUPh* проводили методом двухэтапной ионообменной хроматографии с использованием наполнителя Бутил-сефароза на первом этапе и Q-сефароза (Amersham Pharmacia Biotech) — на втором. По результатам SDS-гель-электрофореза гомогенность препарата *VchUPh* составляла 96%. Методика наработки биомассы продуцента, выделения и очистки *VchUPh* описана в [8].

Кристаллизация. В [8, 10] описана кристаллизация белка методом диффузии в парах при концентрации белка 15 мг/мл (в 0.1 М Трис-НCl), состав противораствора: 0.2 М $MgCl_2 \cdot 6H_2O$, 15% (w/v) полиэтиленгликоля (ПЭГ) 4000, 0.1 М Трис-НCl, pH 8.5.

Для эксперимента в условиях микрогравитации условия кристаллизации методом диффузии

в парах первоначально были адаптированы к методу встречной диффузии в капиллярах. В качестве начальных взяты известные условия кристаллизации, варьировались лишь концентрация хлорида магния от 0.1 до 0.4 М и концентрация ПЭГ 4000 от 15 до 25% с шагом 2%. При кристаллизации использовали метод встречной диффузии в капиллярах [17]. Кристаллизацию проводили на МКС. Параллельно в лаборатории проводил наземный контрольный эксперимент с идентичными кристаллизационными условиями. В контрольном эксперименте пригодные для РСА кристаллы выросли в течение двух недель. Для сбора рентгенодифракционных данных использовали кристаллы, выращенные в условиях микрогравитации в 0.3 М $MgCl_2 \cdot 6H_2O$, 19% (w/v) ПЭГ 4000, 0.1 М Трис-НCl, pH 8.5 при температуре 293 К (рис. 1).

Сбор и обработка данных. Дифракционные данные с кристалла *VchUPh* получены при температуре 100 К на станции BL41XU синхротрона SPring-8 (Хёго, Япония). В качестве криопротектанта использовали раствор, содержащий 20% глицерина и все компоненты кристаллизационного раствора. С помощью программ комплекса XDS [18] проведена обработка экспериментального набора интенсивностей отражений. Шкалирование интенсивностей и расчет модулей структурных факторов проведены в программе SCALA [19]. Параметры съемки и итоговая статистика дифракционных данных приведены в табл. 1.

Решение и уточнение структуры. Структура решена методом молекулярного замещения по программе Molrep [20]. В качестве стартовой модели использована структура *VchUPh* (ID PDB: 6EYP). Структура уточнена при помощи программы *phenix.refine* [21] в сочетании с ручной коррекцией в программе Coot [22]. В структуре локализованы связанные с белком неспецифические лиганды: глицерин, 1,2-этанediол, диэтиленгликоль, ионы магния, натрия. Температурные факторы атомов рассчитывали в анизотропном приближении. На последних этапах уточнения к модели добавлены атомы водорода. Корректность результатов уточнения структуры оценивали в программах ProCheck [23], MolProbity [24] и PDB Validation Server (<https://validate-rcsb.wwpdb.org>). Основные значения параметров уточнения структуры приведены в табл. 1. Структура комплекса депонирована в RCSB PDB (ID PDB: 6Z9Z).

Молекулярное моделирование методом классической молекулярной динамики (МД). Для определения наиболее вероятных конформаций участков основной цепи, формирующих активный центр молекулы, и оценки стабильности связывания фермента с молекулой глицерина проводили МД-моделирование в пакете программ GROMACS (версия 2018.8) [25] с использованием набора пол-

ноатомных силовых полей CHARMM [26] версии C36m [27]. Обработку моделей лигандов осуществляли с помощью WEB-сервиса CGenFF [28]. Молекулы воды описывали трехцентровой моделью TIP3P. К системе были добавлены ионы магния и хлора в концентрации, соответствующей 0.3 М $MgCl_2$. Расчет дальних электростатических взаимодействий проводили методом частица–сетка, используя функцию кубической интерполяции. Для описания ван-дер-ваальсовых взаимодействий использовали функцию сглаживания “force-switch” в интервале 10–12 Å. Давление в системе контролировали с помощью баростата Паринелло–Рахмана [29] на уровне 1 бар. Температура МД-системы поддерживалась постоянной с помощью термостата шкалирования скоростей частиц (V-rescale [30]) на уровне 293 К, что соответствовало температуре роста кристаллов. Длину траектории МД задавали равной 100 нс. Для интегрирования с шагом 2 фс уравнений движения Ньютона использовали алгоритм с перескоками. Всего было получено четыре траектории димера CD: две траектории А-конформера с молекулой глицерина и без нее и две траектории В-конформера также с молекулой глицерина и без нее.

Определение относительной аффинности связывания глицерина с активным центром VchUPh методом расчета линейной энергии взаимодействия (LIE, [31]) осуществляли в программе gmx_lie программного пакета GROMACS. Траектории МД-симуляции белок–лиганд вновь обрабатывали, используя метод поля реакции при отсечке 15 Å для расчета потенциальной энергии, кулоновского взаимодействия лиганда с окружением. Новую обработку траекторий осуществляли в пределах времени, в течение которого глицерин не покидал область активного центра UPh (4.2 нс для альтернативной А-конформации и 1 нс для В-конформации).

Для дальнейшего исследования результатов молекулярного моделирования использовали метод марковских состояний [32] с помощью программного комплекса PyEmma [33]. Для первоначального понижения размерности данных до набора переменных, определяющего конформационные свойства полипептидной цепи, использовали значения синуса и косинуса двугранных углов основной цепи D VchUPh для участков аминокислотных остатков (а.о.) 90–107, 144–153, 215–226. Для дальнейшего понижения размерности конформационного пространства использовали метод анализа независимых компонентов с запаздыванием по времени [34]. На этом этапе корреляционные функции рассчитывали таким образом, чтобы коэффициенты автокорреляции были максимальными, а корреляции между различными переменными были равны нулю, что позволило дополнительно понизить размерность путем преоб-

Таблица 1. Характеристики получения экспериментального набора интенсивностей, обработки данных, решения и уточнения структуры

Характеристика набора данных	
Длина волны, Å	0.7
Детектор	EIGER 16M
Расстояние от кристалла до детектора, мм	190
Угол осцилляции, град	0.1
Пр. гр.	$P2_1$
$a, b, c, \text{Å}$	93.186, 97.061, 93.227
$\alpha = \gamma, \beta, \text{град}$	90.00; 119.965
Разрешение, Å	46.59–1.04 (1.08–1.04)*
Полное число отражений	2680710 (268036)*
Число независимых отражений	682103 (67467)*
Полнота набора, %	99.15 (98.57)*
Повторяемость	3.9 (4.0)*
Среднее значение $\langle I/\sigma(I) \rangle$	12.64 (1.56)*
Общий температурный фактор Вильсона, Å ²	10.16
$R_{\text{merge}}, \%$	5.56 (82.41)*
$R_{\text{meas}}, \%$	6.41 (94.98)*
$R_{\text{pim}}, \%$	3.16 (46.72)*
$CC_{1/2}, \%$	99.8 (63.6)*
Уточнение структуры	
Срезка набора, $\min(\sigma(F)/ F)$	2.0
Число рефлексов в рабочем наборе	679672 (67460)*
Число рефлексов в тестовом наборе	33812 (3492)*
$R_{\text{work}}, \%$	12.56 (23.27)*
$R_{\text{free}}, \%$	14.20 (24.36)*
Cruickshank DPI, Å	0.029
Число уточняемых неводородных атомов	
Белка	12501
Лигандов	105
Воды	1688
R.m.s.d. от “идеальной” геометрии	
по длинам валентных связей, Å	0.007
по валентным углам, град	0.96
Среднее значение B-фактора для всех атомов, Å ²	16.27
Макромолекулы	14.39
Лигандов	32.47
Воды	29.23
Статистика Рамачандрана	
Число а.о. в наиболее благоприятных областях, %	98.79
Число а.о. в разрешенных областях, %	0.81
Число а.о. в неблагоприятных областях, %	0.4
Число редко встречающихся ротамеров, %	0.73
ID PDB	6Z9Z

Примечание. $CC_{1/2}$ – коэффициент корреляции Пирсона между двумя случайно выбранными группами измеренных интенсивностей отражений, составляющих каждая половину от всего набора.

* В скобках приведены значения для зоны высокого разрешения.

разования к так называемым “независимым координатам”. Выбор времени запаздывания основывался на знании аналогичных времен, типичных для схожих конформационных изменений в других белках [35], поэтому дальнейшее “подстраивание” проводили после анализа предполагаемых временных интервалов переходов между основными состояниями полученной марковской модели [36]. Для рассматриваемых участков молекулы (а.о. 90–107, 144–153, 215–226) устанавливали времена запаздываний 3, 4 и 3 нс соответственно.

Перед построением марковской модели проводили дискретизацию траекторий при помощи алгоритма кластеризации k -средних. Число кластеров, равное тридцати для всех трех рассматриваемых участков полипептидной цепи, подбиралось с помощью эвристического метода, описанного в документации PyEmm на основе функции оценки VAMP-2 Score [37]. На основе проведенной кластеризации построена марковская модель, а именно: осуществлен подсчет количества переходов из одного состояния в другое, построена матрица переходов, вычислены ее собственные значения и собственные векторы. Проекция центроидов кластеров, полученных на этапе дискретизации, на собственные векторы с наибольшими собственными значениями определяли положение основных состояний.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Кристаллы VchUPh и данные их рентгеноструктурного анализа. В условиях микрогравитации на МКС выращены высокосовершенные кристаллы уридинфосфорилазы из патогенной бактерии *Vibrio cholerae*. Определена пространственная структура VchUPh с рекордно высоким разрешением – 1.04 Å.

Пространственная организация структуры фермента VchUPh аналогична рассмотренным ранее структурам VchUPh [6, 8–10]. Каждая субъединица молекулы VchUPh, представляющей собой тороидальный гомогексамер диаметром ~106 Å, включает в себя 253 а.о. (молекулярная масса = 27.5 кДа). По данным программы DSSP [38] вторичная структура мономера представлена восемью β -стрендами (28% а.о.) и восемью α -спиралями (32% а.о.). В межмономерной области каждого гомодимера гексамерной молекулы локализован ион Na^+ , координирующий три пары одноименных а.о. смежных субъединиц гомодимера. В каждом димере находится по два идентичных активных центра фермента, образованных а.о. соседних субъединиц гомодимера.

Активный центр VchUPh. Активные центры молекулы свободны от специфических лигандов, но большинство рибозосвязывающих участков

нуклеозидсвязывающих сайтов, образованных а.о. Glu197, Thr93 и His7 соседней субъединицы в гомодимере [10], заняты неспецифически связанными молекулами этандиола и глицерина. Это можно объяснить схожестью химической природы многоатомных спиртов и моносахаридных остатков, являющихся альдегидо- или кето-спиртами. В большинстве активных центров молекулы глицерина связаны водородными связями лишь с одним аминокислотным остатком рибозосвязывающего сайта (Glu197 или His7).

В ранее решенной с разрешением 1.22 Å структуре нелигандированной VchUPh – 6EYP – в большинстве активных центров глицерин находится в двойном положении: один из конформеров связан водородными связями лишь с карбоксильной группой Glu197, другой конформер связан с обоими а.о. (Glu197, His7).

В рассматриваемой структуре связанные с рибозосвязывающими участками активного центра молекулы спиртов находятся в одной конформации, но иногда с неполной заселенностью. В активном центре субъединицы D молекула глицерина связана как с имидазольной группой His7 соседней субъединицы (NE2_His7/C – 3.0 Å – O1_GOL), так и с карбоксильной группой Glu197 (OE2_Glu197/D – 2.7 Å – O3_GOL, OE1_Glu197/D – 3.3 Å – O3_GOL) (рис. 2).

В случае связывания молекулы глицерина одновременно с His7 и Glu197 ее конформация и расположение в сайте связывания UPh похожи на конформацию части C3'–O3'–C4'–O4'–C5'–O5' рибозной компоненты субстратов, псевдосубстратов или ингибиторов UPh (уридин, тимидин, цитидин, 2,2-ангидроуридин). Но имеется и существенное отличие: Thr93, формирующий подвижную стенку активного центра [6, 10, 11], не образует водородных связей с молекулой спирта ни в одном активном центре как рассматриваемой структуры, так и 6EYP, а его боковая цепь находится на расстоянии > 4–5 Å от атомов лигандов. В отличие от глицерина субстраты и ингибиторы связываются с боковой цепью Thr93 посредством O4'-гидроксигруппы, положению которой в молекуле глицерина соответствует O2-гидроксигруппа.

Дополнительное исследование связывания молекулы глицерина в активном центре проведено методом МД. В качестве стартовых моделей для МД брали гомодимер CD в разных альтернативных положениях а.о. 91–106, образующих “стенку” активного центра D-субъединицы. Именно в этом гомодимере разница между альтернативными конформациями наибольшая. В обоих случаях лиганд лишь небольшую часть времени траектории находился в связанном состоянии (~4.2 нс и 1 нс для A и B-конформаций соответственно) (рис. 3).

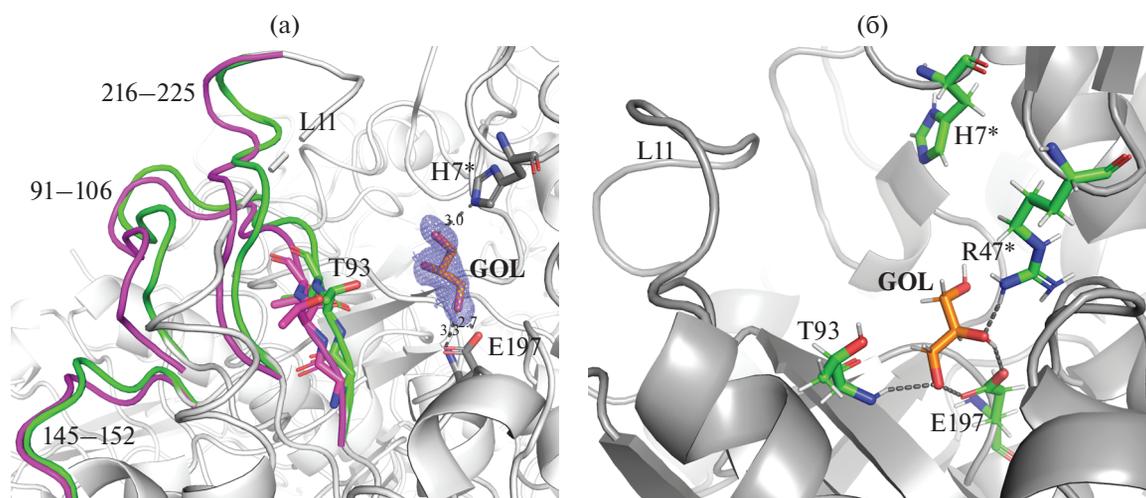


Рис. 2. Пространственная организация ферментативного центра D(C) *VchUPh* (ID PDB: 6Z9Z). Показан фрагмент карты электронной плотности $2F_o - F_c$. Пунктиром показаны полярные контакты доноров и акцепторов водородных связей. Аминокислотные остатки углеродного остова и ход главной цепи в разных альтернативных конформациях структуры показаны разными цветами; * – а.о. С-субъединицы; L11 – петля-“шлагбаум” (а). Структура активного центра фермента (D-субъединица, A-конформация) с минимальным значением ΔG_{LIE} относительной аффинности связывания глицерина с *VchUPh* (б).

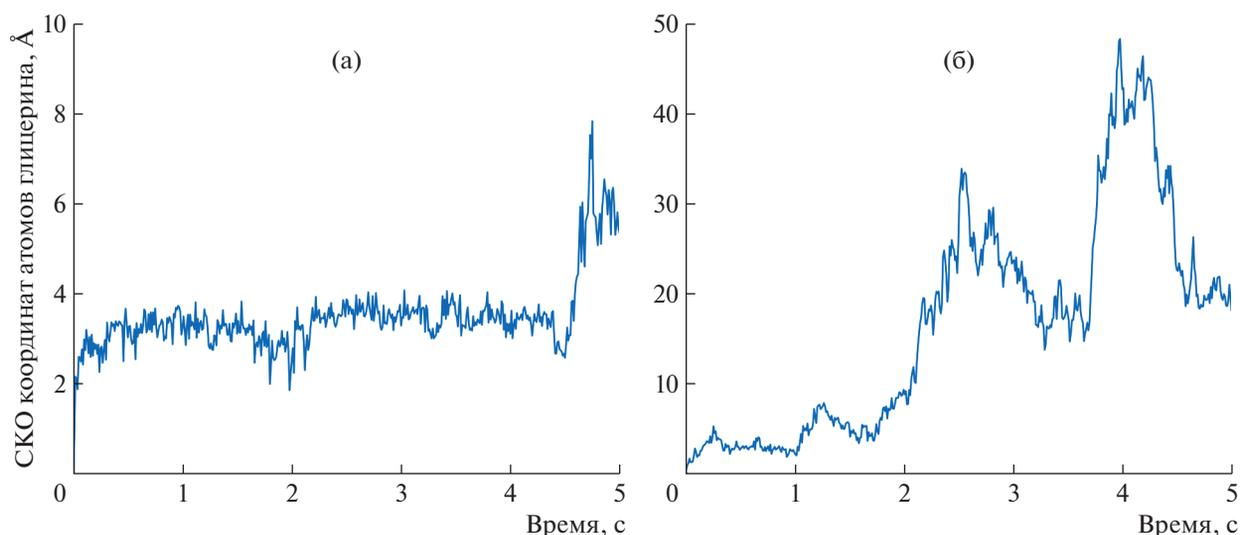


Рис. 3. График зависимости среднеквадратичного отклонения координат неводородных атомов глицерина в сравнении с их начальным значением от времени молекулярной динамики (первые 6 нс) в случае A-конформации фрагментов структуры *VchUPh* (а); B-конформации (б).

В случае A-конформации, когда фрагмент 91–106 находился ближе к лиганду (рис. 2а), связывание с активным центром было относительно стабильным в течение ~ 4.2 нс (рис. 3а), и средняя оценка свободной энергии связывания составляла $\langle \Delta G_{LIE} \rangle = -14.4$ кДж/моль. Однако в течение нескольких первых пикосекунд конформация лиганда принципиально менялась по отношению к экспериментально определенной, и минимальное значение оценки свободной энергии ($\Delta G_{LIE}^{\min} =$

-38.8 кДж/моль) соответствовало этой новой конформации глицерина в сайте связывания (рис. 2б). В этой конформации лиганд не связан водородными связями с атомами боковой цепи His7/C, зато связывается с аминок группой основной цепи Thr93/D и атомами боковой цепи Arg47/C. Arg47 входит в фосфатсвязывающий сайт активного центра, а Thr93 в $\beta 5$ -стренд, образующий стенку активного центра.

В случае B-конформации фрагмента 91–106, когда тот находился дальше от лиганда (рис. 2а),

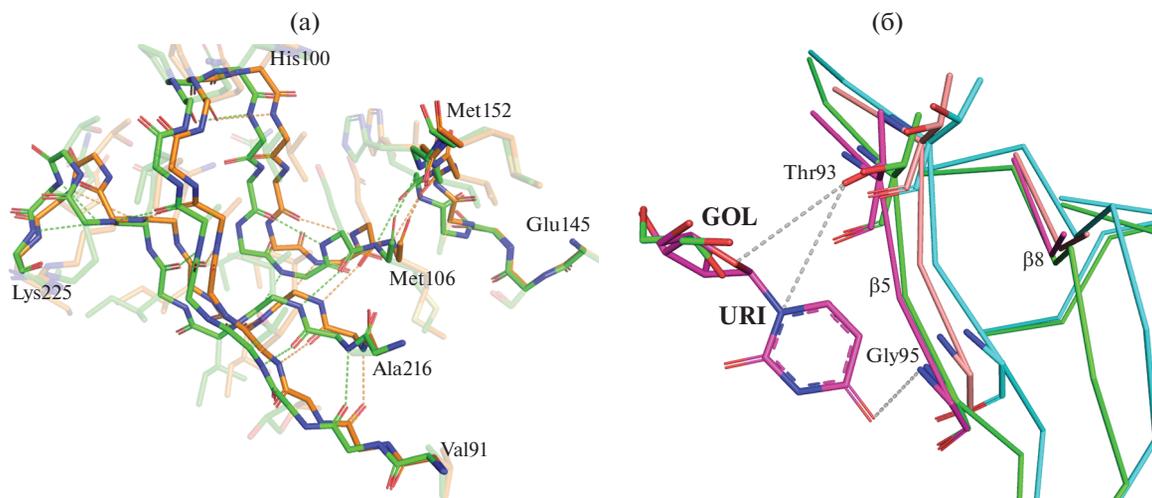


Рис. 4. Сеть водородных связей между атомами главной цепи фрагментов пространственной структуры субъединицы D *VchUPh* в области альтернативных конформаций. Альтернативные конформации показаны разными цветами углеродного остова молекулы (а). Наложение пространственных структур фрагментов β 5- и β 8-стрендов нелигандированной *VchUPh* (ID PDB: 6Z9Z) и ее комплекса с уридином ((ID PDB: 5M2T) (б).

связывание было стабильным в течение значительно меньшего времени (~ 1 нс) (рис. 3б), и средняя оценка свободной энергии выше ($\langle \Delta G_{LIE} \rangle = -7.9$ кДж/моль) в сравнении с А-конформацией. В течение всего времени пребывания в активном центре фермента молекула глицерина не образовывала устойчивых водородных связей с Thr93, однако, так же как и для А-конформации, связывалась с Arg47/С фосфатсвязывающего сайта. Учитывая все изложенное выше, можно сделать вывод о зависимости стабильности и аффинности связывания глицерина от положения и конформации β 5-стренда, образующего стенку активного центра, однако в любом случае стабильность связывания глицерина с ферментом ниже, чем у его субстратов и конкурентных ингибиторов.

Петля-“шлагбаум” L11 [10] во всех активных центрах находится в открытом положении, как и в структуре 6EYP, что логично, так как этандиол и глицерин не являются субстратами UPh, а фосфатсвязывающий сайт свободен от ортофосфат-аниона.

Анализ альтернативных конформаций функционально значимых областей VchUPh в нелигандированном состоянии. В субъединицах А, С, F боковая цепь Arg47 фосфатсвязывающего сайта находится в альтернативных конформациях. Альтернативные конформации этого а.о. наблюдались ранее [6, 10, 11] во всех структурах, где фосфатсвязывающий сайт не лигандирован ортофосфат- либо сульфат-анионом. Подвижность боковых цепей аргинина фосфатсвязывающего сайта, как было показано ранее, играет важную роль в связывании и позиционировании ортофосфат-аниона относительно молекулы нуклеотида [10].

Большой интерес вызывают альтернативные конформации пространственно близко расположенных участков 91–106, 145–152 и 216–225 а.о. D-субъединицы. В альтернативных конформациях находятся не только боковые цепи этих а.о., но и главная цепь этих фрагментов (рис. 2).

В рассмотренных прежде структурах высокого разрешения комплексов *VchUPh* с субстратами [10], псевдосубстратом – 6-метилурацилом [6], ингибитором – 2,2'-ангидроуридином [11] также наблюдаются двойные положения фрагментов β 5- и β 8-стрендов при связывании фермента с лигандом, имеющим неполную заселенность. При этом одно из двойных положений фрагментов элементов третичной структуры соответствует лигандированному состоянию, а второе – нелигандированному. Однако, во-первых, в двойной конформации находились значительно меньшие участки структуры (92–95 и 217–219 а.о. в структуре комплекса *VchUPh* с 2,2'-ангидроуридином [11]). Во-вторых, двойные положения наблюдались при частичной заселенности лиганда. В-третьих, а.о. одного из двойных положений фрагмента β 5-стренда связывались с молекулой лиганда либо непосредственно (уридин, тимидин, 2,2'-ангидроуридин), либо через одну молекулу воды (6-метилурацил). В структуре, описываемой в данной работе, заселенность молекулы глицерина в активном центре D-субъединицы полная, и ее атомы не образуют значимых связей с атомами а.о. 92–94.

Фрагментам субъединицы D, находящимся в альтернативных конформациях, соответствуют следующие элементы вторичной структуры:

участок 91–96 β 5-стренда, неупорядоченные фрагменты с поворотными элементами 97–106 и 145–149, начальный фрагмент 150–152 β 7-стренда,

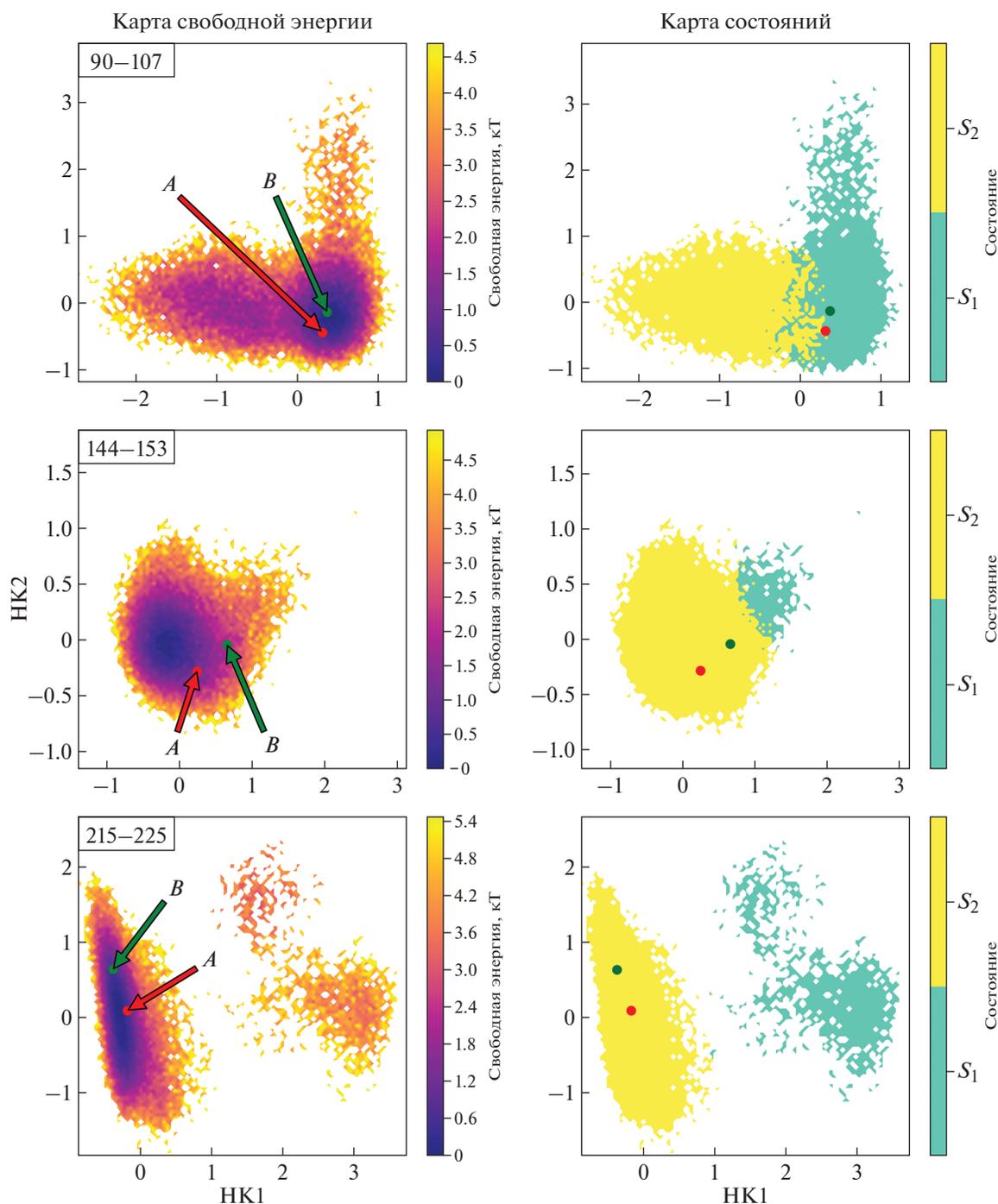


Рис. 5. Карта свободной энергии конформаций трех фрагментов пространственной структуры *VchUPh*, построенная по данным марковских моделей как функция двух самых медленных независимых компонент (левый столбец). Карта самых долгоживущих метастабильных состояний (правый столбец). Конформации (A и B) фрагментов, полученные из данных рентгеноструктурного анализа, выделены стрелками.

концевой фрагмент 216–221 β 8-стренда, начальный фрагмент петли-“шлагбаума” L11 (221–225).

Атомы главной цепи этих фрагментов взаимодействуют друг с другом посредством сети водородных связей (рис. 4а), и их движения в молекуле, по-видимому, согласованы. В [10] предложена

схема конформационных изменений, индуцированных связанным специфическим лигандом, которые приводят к закрытию активного центра. Изменения происходят в результате согласованного сдвига β 5-стренда, β 8-стренда и связанной с последним петлей-“шлагбаумом” L11. Однако в

рассматриваемой структуре, судя по всему, имеет место не зависящее от связывания рибозосвязывающим сайтом специфического лиганда движение $\beta 5$ - и $\beta 8$ -стрендов. При этом не происходит полного перехода сайта связывания в закрытую конформацию, так как, во-первых, нет взаимодействия фосфатсвязывающего сайта с фосфат-анионом и, во-вторых, возможно, для фиксации закрытой конформации необходимо значимое взаимодействие молекулы нуклеозида или его аналога с атомами а.о. 93–95 $\beta 5$ -стренда (рибозной части с Thr93, а пиримидиновой – с основной цепью Gly95).

Менее вероятная гипотеза состоит в том, что в рассматриваемой структуре молекула глицерина, связываясь в рибозосвязывающем сайте, индуцирует конформационные изменения элементов структуры. Однако никаких значимых взаимодействий ни напрямую, ни через молекулы воды между атомами молекулы глицерина и атомами фрагмента 93–95 $\beta 5$ -стренда обнаружено не было. Кроме того, даже в случае, когда молекула глицерина связывается лишь с одним а.о. нуклеозидсвязывающего сайта (субъединица А), все равно обнаруживаются аналогичные наблюдаемым в субъединице D альтернативные конформации фрагмента $\beta 5$ -стренда.

Сравнение альтернативных конформаций в структурах атомного разрешения нелигандированной *VchUPh* и ее комплекса с уридином (ID PDB: 5M2T, 1.03 Å) выявило, что альтернативные конформации $\beta 5$ - и $\beta 8$ -стрендов в целом соответствуют наблюдаемым в рассматриваемой структуре (рис. 4б). Но атомы $\beta 5$ -стренда в конформации, соответствующей в структуре 5M2T закрытому состоянию активного центра, в особенности участка 93–94, находятся ближе к лиганду и противоположной стенке активного центра (рис. 4б). Это подтверждает гипотезу о том, что согласованные изменения конформации $\beta 5$ - и $\beta 8$ -стрендов возможны и без специфического связывания активного центра с субстратами или конкурентными ингибиторами, но переход в полностью закрытое состояние и его фиксация возможны лишь при связывании с атомами а.о. $\beta 5$ -стренда со специфическими лигандами фермента. Атомы уридина образуют в свою очередь водородные связи лишь с атомами а.о. 93–95 $\beta 5$ -стренда одной (более “закрытой”) конформации.

С целью доказательства или опровержения гипотезы о спонтанном, не зависящем от связывания с лигандом, переходе между конформациями рассматриваемых фрагментов структуры проведен конформационный анализ с использованием моделей марковских состояний. На рис. 5 (левый столбец) изображена карта свободной энергии кластеров конформаций фрагментов структуры в зависимости от значения первых двух (самых

“медленных”) независимых компонент. Обе экспериментально определенные конформации соответствуют субминимальному значению свободной энергии, обе находятся в пределах одной потенциальной ямы, потенциальных барьеров между ними не наблюдается. При анализе основных стационарных самых «долгоживущих» состояний марковской модели выявлено, что экспериментально определенные конформации относятся к одному стационарному состоянию (рис. 5, правый столбец), однако для участка 90–107 точки находятся близко к границе другого марковского состояния. Из этих данных можно сделать вывод об относительно свободном переходе между экспериментально определенными конформациями фрагментов структуры (а.о. 91–106, 145–152 и 216–225), который не индуцируется связыванием с лигандом, а происходит самопроизвольно.

ЗАКЛЮЧЕНИЕ

Методом встречной диффузии в капиллярах в условиях микрогравитации на МКС выращены высокосоввершенные кристаллы уридинфосфорилазы из патогенной бактерии *Vibrio cholerae*. Определена ее пространственная структура атомного (1.04 Å) разрешения. Это рекордно высокое разрешение для данного фермента. Впервые в пространственной структуре уридинфосфорилазы в отсутствие связанных специфических лигандов обнаружены альтернативные конформации протяженных областей ($\beta 5$ -, $\beta 7$ -, $\beta 8$ -стрендов и смежных неупорядоченных участков) белковой молекулы. На основании анализа рентгеноструктурных данных и конформационного анализа методами молекулярной динамики и марковских моделей установлено, что согласованные изменения конформации $\beta 5$ -, $\beta 7$ - и $\beta 8$ -стрендов возможны и без специфического связывания аминокислотных остатков активного центра с субстратами, псевдосубстратами или конкурентными ингибиторами. Переход активного центра в полностью закрытое состояние и его фиксация возможны лишь при связывании лигандов с атомами а.о. $\beta 5$ -стренда со специфическими лигандами фермента.

Выявлено, что молекулы глицерина могут связываться в рибозосвязывающем участке активного центра уридинфосфорилазы в конформации, близкой к конформации фрагментов природных субстратов и ингибиторов фермента. Но аффинность и стабильность подобного связывания велики в сравнении с субстратами и конкурентными ингибиторами и зависят, помимо всего прочего, от конформации и положения $\beta 5$ -стренда фермента.

Работа выполнена в рамках Федеральной космической программы 2016–2025 (МКС “Наука”) (в части кристаллизации методом встречной диффузии и сбора дифракционных данных), при фи-

нансовой поддержке Российского фонда фундаментальных исследований (проект № 19-29-12054) в части поиска условий кристаллизации, при поддержке Министерства науки и высшего образования в рамках выполнения работ по Государственному заданию ФНИЦ “Кристаллография и фотоника” РАН (в части решения, уточнения и анализа пространственной структуры, в том числе с использованием молекулярной динамики и марковских моделей). Работа выполнена с использованием инфраструктуры ЦКП “Информатика” ФИЦ ИУ РАН [39].

СПИСОК ЛИТЕРАТУРЫ

1. *Katsumata K., Tomioka H., Sumi T. et al. // Cancer Chemother. Pharmacol.* 2003. V. 51 (2). P. 155. <https://doi.org/10.1007/s00280-003-0583-2>
2. *Kançakı A., Takebayashi Y., Bando H. et al. // Int. J. Cancer.* 2002. V. 97 (5). P. 631. <https://doi.org/10.1002/ijc.10105>
3. *Luccioni C., Beaumatin J., Bardot V. et al. // Int. J. Cancer.* 1994. V. 58 (4). P. 517. <https://doi.org/10.1002/ijc.2910580411>
4. *Лашков А.А., Жухлистова Н.Е., Сотниченко С.Е. и др. // Кристаллография.* 2010. Т. 55. С. 44.
5. *Caradoc-Davies T.T., Cutfield S.M., Lamont I.L. et al. // J. Mol. Biol.* 2004. V. 337. P. 337. <https://doi.org/10.1016/j.jmb.2004.01.039>
6. *Прокофьев И.И., Лашков А.А., Габдулхаков А.Г. и др. // Кристаллография.* 2018. Т. 63. С. 415. <https://doi.org/10.7868/S0023476118030116>
7. *Lashkov A.A., Zhukhlistova N.E., Gabdulkhakov A.G. et al. // Acta Cryst. D.* 2010. V. 66. P. 51. <https://doi.org/10.1107/S0907444909044175>
8. *Lashkov A.A., Gabdulkhakov A.G., Prokofev I.I. et al. // Acta Cryst. F.* 2012. V. 68. P. 1394. <https://doi.org/10.1107/S1744309112041401>
9. *Prokofev I.I., Lashkov A.A., Gabdulkhakov A.G. et al. // Acta Cryst. F.* 2014. V. 70. P. 60. <https://doi.org/10.1107/S2053230X13031877>
10. *Прокофьев И.И., Габдулхаков А.Г., Балаев В.В. и др. // Кристаллография.* 2016. Т. 61. С. 919. <https://doi.org/10.7868/S0023476116060138>
11. *Эйстрих-Геллер П.А., Рубинский С.В., Прокофьев И.И. и др. // Кристаллография.* 2020. Т. 65. С. 271. <https://doi.org/10.31857/S002347612002006X>
12. *Daa Silva Neto A.M., Torini de Souza J.R., Romanello L. et al. // Biochimie.* 2016. V. 125. P. 12. <https://doi.org/10.1016/j.biochi.2016.02.007>
13. *Roosild T.P., Castronovo S., Villosio A. et al. // J. Struct. Biol.* 2011. V. 176. P. 229. <https://doi.org/10.1016/j.jsb.2011.08.002>
14. *Куранова И.П., Ковальчук М.В. // Природа.* 2014. № 3. С. 12.
15. *Snell E.H., Helliwell J.R. // Rep. Prog. Phys.* 2005. V. 68 (4). P. 799. <https://doi.org/10.1088/0034-4885/68/4/r02>
16. *Бойко К.М., Тимофеев В.И., Самыгина В.Р. и др. // Кристаллография.* 2016. Т. 61. С. 691.
17. *Абрамчик Ю.А., Тимофеев В.И., Муравьева Т.И. и др. // Кристаллография.* 2016. Т. 61. № 6. С. 940.
18. *Kabsch W. // Acta Cryst. D.* 2010. V. 66. P. 125. <https://doi.org/10.1107/S0907444909047337>
19. *Evans P. // Acta Cryst. D.* 2006. V. 62. Pt. 1. P. 72. <https://doi.org/10.1107/S0907444905036693>
20. *Vagin A., Teplyakov A. // J. Appl. Cryst.* 1997. V. 30 (6). P. 1022. <https://doi.org/10.1107/S0021889897006766>
21. *Afonine P.V., Grosse-Kunstleve R.W., Echols N. et al. // Acta Cryst. D.* 2012. V. 68. P. 352. <https://doi.org/10.1107/S0907444912001308>
22. *Emsley P., Lohkamp B., Scott W.G. et al. // Acta Cryst. D.* 2010. V. 66. P. 486. <https://doi.org/10.1107/S0907444910007493>
23. *Laskowski R.A., MacArthur M.W., Moss D.S. et al. // J. Appl. Cryst.* 1993. V. 26 (2). P. 283. <https://doi.org/10.1107/S0021889892009944>
24. *Davis I.W., Leaver-Fay A., Chen V.B. et al. // Nucl. Acids Res.* 2007. V. 35. P. 375. <https://doi.org/10.1093/nar/gkm216>
25. *Van Der Spoel D., Lindahl E., Hess B. et al. // J. Comput. Chem.* 2005. V. 26. P. 1701. <https://doi.org/10.1002/jcc.20291>
26. *MacKerell A.D. // Abstracts Papers Am. Chem. Soc.* 1998. V. 216. P. 696.
27. *Huang J., Rauscher S., Nawrocki G. et al. // Nature Methods.* 2016. V. 14. P. 71. <https://doi.org/10.1038/nmeth.4067>
28. *Vanommeslaeghe K., Hatcher E., Acharya C. et al. // J. Comput. Chem.* 2010. V. 31. P. 671. <https://doi.org/10.1002/jcc.21367>
29. *Parrinello M., Rahman A. // J. Appl. Phys.* 1981. V. 52. P. 7182. <https://doi.org/10.1063/1.328693>
30. *Bussi G., Parrinello M. // Comput. Phys. Commun.* 2008. V. 179. P. 26. <https://doi.org/10.1063/1.3073889>
31. *Aqvist J., Marelus J. // Comb. Chem. High Throughput Screen.* 2001. V. 4 (8). P. 613. <https://doi.org/10.2174/1386207013330661>
32. *Schutte C., Fischer A., Huisinga W., Deuflhard P. // J. Comput. Phys.* 1999. V. 151. P. 146. [https://doi.org/10.1016/S1570-8659\(03\)10013-0](https://doi.org/10.1016/S1570-8659(03)10013-0)
33. *Wehmeyer C., Scherer M.K., Hempel T. et al. // Living J. Comp. Mol. Sci.* 2019. V. 1. P. 5965. <https://doi.org/10.33011/livecoms.1.1.5965>
34. *Perez-Hernandez G., Paul F., Giorgino T. et al. // J. Chem. Phys.* 2013. V. 139. P. 015102. <https://doi.org/10.1063/1.4811489>
35. *Husic B.E., Pande V.S. // J. Chem. Phys.* 2017. V. 147. P. 176101. <https://doi.org/10.1063/1.5002086>
36. *Trendelkamp-Schroer B., Wu H., Paul F., Noé F. // J. Chem. Phys.* 2015. V. 143. P. 174101. <https://doi.org/10.1063/1.4934536>
37. *McGibbon R.T., Pande V.S. // J. Chem. Phys.* 2015. V. 142. P. 124105. <https://doi.org/10.1063/1.4916292>
38. *Kabsch W., Sander C. // Biopolymers.* 1983. V. 22. P. 2577. <https://doi.org/10.1002/bip.360221211>
39. *Положение о ЦКП “Информатика”.* <http://www.frccsc.ru/ckp>

СТРУКТУРА МАКРОМОЛЕКУЛЯРНЫХ СОЕДИНЕНИЙ

УДК 548.73 + 577.1

РОЛЬ КОНФОРМАЦИОННЫХ ИЗМЕНЕНИЙ ГЕКСАМЕРНОЙ МОЛЕКУЛЫ БАКТЕРИАЛЬНЫХ УРИДИНФОСФОРИЛАЗ В СВЯЗЫВАНИИ СУБСТРАТОВ

© 2021 г. К. М. Поляков^{1,*}, Н. Н. Мордкович², Т. Н. Сафонова², А. Н. Антипов²,
Н. А. Окорокова², П. В. Дороватовский³, В. П. Вейко^{2,**}

¹ Институт молекулярной биологии им. В.А. Энгельгардта РАН, Москва, Россия

² Институт биохимии им. А.Н. Баха, Федеральный исследовательский центр “Фундаментальные основы биотехнологии” РАН, Москва, Россия

³ Национальный исследовательский центр “Курчатовский институт”, Москва, Россия

*E-mail: kmpolyakov@gmail.com

**E-mail: vladveiko@yahoo.com

Поступила в редакцию 17.09.2020 г.

После доработки 19.11.2020 г.

Принята к публикации 19.11.2020 г.

Получены кристаллы и определены трехмерные структуры мутантных форм уридинфосфоорилазы из *Shewanella oneidensis* MR-1 по остатку треонина активного центра. Показано, что петля 161–179, ответственная за узнавание нуклеозида, участвует в стабилизации гексамерной структуры белка, а ее разупорядочение существенно облегчает проникновение молекулы нуклеозида в активный центр фермента. Обсуждается роль конформационных изменений в функционировании фермента.

DOI: 10.31857/S0023476121050192

ВВЕДЕНИЕ

Уридинфосфоорилаза (УФ, КФ 2.4.2.3) – ключевой фермент пиримидинового обмена [1–3]. УФ катализирует расщепление гликозидной связи С–N в урдине с образованием урацила и рибозо-1'-фосфата. Этот фермент участвует в катаболизме пиримидиновых нуклеозидов, играет важную роль в гомеостатической регуляции концентрации уридина в плазме, он обнаружен в большинстве тканей и опухолей практически всех живых организмов. Интерес к исследованию молекулярных основ действия УФ базируется на данных о том, что в клетках злокачественных новообразований у человека наблюдается повышенный уровень активности УФ [4]. Показано, что экспрессия УФ в некоторых опухолях в 2–4 раза выше по сравнению с соседними нормальными тканями [5]. Изучение структурно-функциональной взаимосвязи в УФ может существенно упростить дизайн высокоэффективных ингибиторов данного фермента, позволяющих подавить развитие и метастазирование опухоли.

К настоящему времени решены кристаллические структуры семи представителей УФ бактерий. Все эти ферменты формируют гексамерные молекулы с точечной симметрией 32. На основании структурных исследований для УФ было сде-

лано предположение, что связывание нуклеозида происходит в открытой конформации фермента [6]. Доступ в активный центр мономера в открытой конформации обеспечивается за счет смещения петли, предшествующей С-концевой α -спирали, и частичного разупорядочения этих участков цепи. Ранее аналогичные конформационные изменения структуры мономеров были установлены для пуриновых нуклеозидфосфоорилаз [7].

Несмотря на активное исследование УФ из разных источников, молекулярный механизм реакции, катализируемой этим белком, остается не до конца выясненным. Особый интерес представляет остаток треонина активного центра, инвариантный в структурах УФ [8], который по данным рентгеноструктурного анализа (РСА) осуществляет непосредственный контакт с обоими субстратами – уридином и ионом фосфата [6]. Для более детального изучения роли этого остатка в функционировании УФ ранее были получены две мутантные формы УФ из *Shewanella oneidensis* MR-1 (SoУФ) по остатку треонина 91 активного центра, замененного на аланин и серин (T91A и T91S соответственно), и исследованы их ферментативные характеристики [9]. В настоящей работе получены кристаллы этих мутантов и определены их пространственные структуры методом РСА. В результате исследования выявлены

конформационные изменения для мутантной формы T91A, проявляющиеся как разупорядочение функциональной петли 161–179, которая включает в себя остатки, участвующие в связывании основания нуклеозида. Такие конформационные изменения приводят к повышению доступности активного центра молекулы белка для нуклеозида даже по сравнению с открытой формой фермента. В работе обсуждается роль этих изменений в связывании субстратов.

МАТЕРИАЛЫ И МЕТОДЫ

Получение мутантных форм уридинфосфориллазы. Мутантные формы УФ из *Sh. oneidensis* MR-1 по остатку треонина активного центра (T91S и T91A) были получены ранее, их ферментативные характеристики приведены в [9]. Рекombинантные белки выделены в гомогенном состоянии по методике, описанной в [8], их первичные структуры подтверждены методом MALDI-TOF/TOF-масс-спектрометрического анализа их триптических гидролизатов.

Концентрацию белка определяли по методу Бредфорда [10] с окраской реагентом “Bio-Rad Protein Assay” (“Bio-Rad”, США). В качестве стандарта использовали раствор бычьего сывороточного альбумина (“Sigma”, США). Ферментативную активность рекомбинантных УФ и их мутантных форм определяли в калий-фосфатном буфере по методике [11].

Сохранение мутантными формами SoУФ четвертичной структуры подтверждали методом аналитической гель-фильтрации на колонке Tricorn 10/300 с сорбентом Superdex 200 с использованием прибора АКТА FPLC (“GE Healthcare”, Великобритания) в 10 mM натрий-фосфатном буфере, pH 7.4, содержащем 150 mM NaCl. В качестве белков-маркеров использовали набор “Gel Filtration Calibration Kits” (GE Healthcare Life Sciences, Великобритания), а также рекомбинантную SoУФ из *Sh. oneidensis* MR-1.

Кристаллизацию мутантных форм SoУФ проводили методом “висячей капли” посредством диффузии в парах при комнатной температуре в термостатированной комнате с использованием стандартных наборов для кристаллизации (скрининга) глобулярных белков компании Hampton Research. Для кристаллизации использовали лиофильно высушенный препарат, растворенный в деионизированной воде (MilliQ) с концентрацией 20 мг/мл. Раствор белка центрифугировали непосредственно перед кристаллизацией на микроцентрифуге в течение 10 мин со скоростью 18000 об./мин для удаления инородных частиц и агрегатов. Для приготовления каждой капли использовалось 2.0 мкл раствора белка, который смешивался с противораствором в соотношении

1:1. Объем противораствора в резервуаре составлял 400 мкл. Кристаллизацию проводили в пластиковых планшетах для висячих капель “Linbro” фирмы Hampton Research с использованием силиконированных стекол диаметром 22 мм фирмы Hampton Research. Кристаллы мутантной формы T91S были получены с использованием буфера 0.1 M Bis-Tris (pH 6.5), содержащего 0.2 M сульфата аммония и полиэтиленгликоль (ПЭГ) 3350 (25 мас. %), в качестве противораствора. Кристаллы мутантной формы T91A получены с использованием раствора 1.0 M моногидрата сульфата лития с добавлением ПЭГ 8000 (2 мас. %). Кристаллы, пригодные для PCA, выросли в течение одной–двух недель.

Сбор и обработка дифракционных данных. Наборы дифракционных данных собраны до разрешения 1.65 и 1.90 Å для T91S и T91A соответственно на станции белковой кристаллографии “Белок” синхротрона Национального исследовательского центра “Курчатовский институт” при температуре 100 K. Перед сбором данных кристаллы помещали в крио-растворы, приготовленные из кристаллизационных растворов с добавлением 25% глицерина. Обработку дифракционных данных проводили по программам XDS и XSCALE [12].

Решение и уточнение структур мутантов. Структуры решены методом молекулярного замещения по программе MOLREP [13] с использованием в качестве стартовой модели нативной УФ из *Shewanella oneidensis* MR-1 [6], из которой предварительно были удалены молекулы воды. Кристаллографическое уточнение структуры проводили по программе REFMAC5 [14]. Исправление моделей структур с визуальным контролем параметров стереохимии проводили по программе COOT [15]. Структура T91S решена с разрешением 1.65 Å и уточнена до значения $R_{\text{work}} = 19.4\%$. Уточнение структуры мутантной формы T91S проводили с учетом двойникования. Структура T91A решена с разрешением 1.90 Å и уточнена до значения $R_{\text{work}} = 15.6\%$. Рисунки приготовлены по программе CCP4MG [16]. Кристаллографические расчеты были проведены с использованием комплекса программ CCP4 [17]. Межмолекулярные контакты в олигомерах анализировали по программе PISA [18]. В табл. 1 приведены статистические характеристики наборов рентгеноструктурных данных и уточненных структур.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Кристаллы мутантных форм T91A и T91S выращены, как указано выше, в различных условиях. Для всех кристаллографически независимых мономеров обеих структур электронная плотность подтверждает замены аминокислотного

Таблица 1. Кристаллографические данные и статистические характеристики наборов рентгеноструктурных данных и уточнения структур

Структура	Мутант T91S	Мутант T91A
Станция сбора данных	Станция Белок, Курчатовский институт	Станция Белок, Курчатовский институт
Пространственная группа	$P2_1$	$P3$
$a, b, c, \text{Å}; \alpha, \beta, \gamma, \text{град}$	91.40, 95.49, 91.41; 90, 120.1, 90	91.66, 91.66, 47.56; 90, 90, 120
Разрешение, Å	1.65 (1.75–1.65)*	1.90 (2.02–1.90)
Полнота набора, %	97.4 (88.9)	99.4 (98.5)
Повторяемость	4.9	6.4
$I/\sigma(I)$	8.8 (3.08)	15.6 (3.24)
$R_{\text{meas}}, \%$	13.9 (54.9)	12.3 (68.8)
CC(1/2), %	99.3 (88.3)	99.8 (86.9)
Среднее значение В-фактора для всех атомов, Å^2	14.3	19.6
Число двойниковых доменов кристалла	6**	1
$R_{\text{work}}, \%$	19.4	15.6
$R_{\text{free}}, \%$	23.3	22.1
Число мономеров в независимой части ячейки	6	2
Число уточняемых атомов белка	10946	3590
Число уточняемых молекул воды	839	295
R.m.s.d. от идеальной геометрии, по длинам валентных связей Å	0.020	0.016
R.m.s.d. от идеальной геометрии по валентным углам, град	2.1	2.3

* В скобках приведены значения для последнего слоя.

** Заселенности доменов 0.085, 0.411, 0.137, 0.225, 0.051, 0.091.

остатка в положении 91. В активном центре структуры T91S локализованы молекула урацила в одном мономере, сульфат-ионы в четырех мономерях и молекула глицерина в двух мономерях. Кроме того, в этой структуре локализованы два сульфат-иона, расположенные на оси третьего порядка гексамерной молекулы фермента. В структуре T91A надежно идентифицированы четыре молекулы глицерина, шесть ионов сульфата и ион лития, расположенный на некристаллографической оси симметрии, связывающей независимые мономеры. Однако в этой структуре не удается найти ионы сульфата, расположенные на оси третьего порядка. Несмотря на то что в независимой части элементарной ячейки структуры мутанта T91S содержится шесть кристаллографически независимых мономеров, а в структуре мутанта T91A — два мономера, функциональными олигомерами в обеих структурах являются сходные гексамерные структуры.

Гексамер T91S сформирован шестью кристаллографически независимыми мономерами и имеет некристаллографическую симметрию 32 (рис. 1). Гексамер T91S состоит из трех димеров, связанных некристаллографической осью третьего

порядка. Мономеры в димере связаны некристаллографическими осями симметрии второго порядка. Все мономеры находятся в закрытой конформации. Петля, предшествующая концевой α -спирали (остатки 220–235), частично разупорядочена во всех мономерях. На оси некристаллографической симметрии третьего порядка с двух сторон гексамера локализованы ионы сульфата, координированные остатками аргининов 175 трех мономеров.

Гексамер T91A состоит из несимметричных димеров, связанных кристаллографической осью третьего порядка (рис. 2). Оба независимых мономера (A и D) находятся в открытой конформации. Отметим, что остатки петли 161–179 в мономере A не локализованы. Эта петля содержит глутамин 163 и аргинин 165, участвующие в связывании основания нуклеозида, а также аргинин 175, связывающий ион сульфата/фосфата на оси третьего порядка. Во втором мономере (мономере D) эта петля локализована на картах электронной плотности, но атомы соответствующих остатков имеют более высокие температурные факторы (около 40Å^2) по сравнению со средним температурным фактором по молекуле (19.6Å^2).

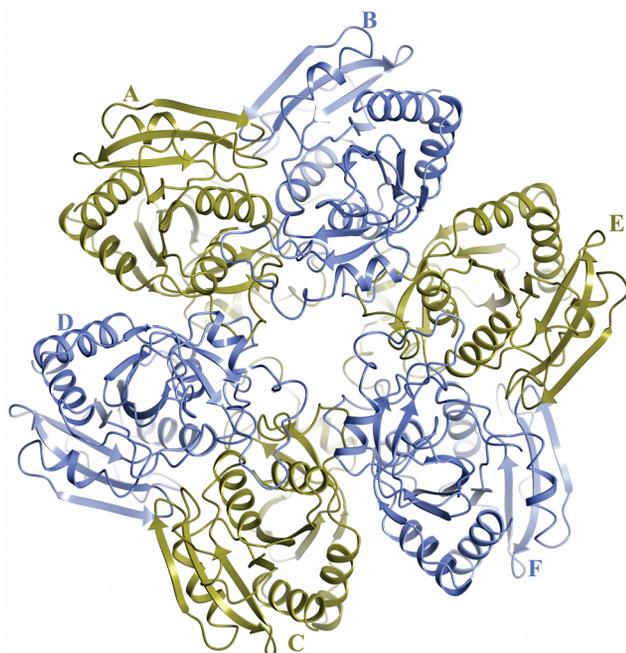


Рис. 1. Ход полипептидной цепи в гексамерной молекуле T91S. Гексамер образован шестью кристаллографически независимыми мономерами.

Такое структурное различие мономеров и приводит к несимметричности гексамерной молекулы. Подобно мономерам в структуре T91S три мономера в T91A (мономер D и его симметричные копии) образуют контакты в области оси третьего порядка (рис. 2).

По оценкам программы PISA [18] гексамеры мутантных форм УФ стабильны. Энергии диссоциации гексамеров равны 70.6 и 36.0 ккал/моль для T91S и T91A соответственно. Такая разница в энергии диссоциации объясняется вкладом связывания двух сульфат-ионов, присутствующих на оси третьего порядка в структуре T91S (около 26.0 ккал/моль), и упорядочением в этой структуре петли 161–179.

В структурах обоих мутантов остаток аргинина 175 в мономерах с упорядоченной петлей 161–179 принимает участие в стабилизации гексамеров за счет водородных связей с карбонильными атомами кислорода соседних мономеров (рис. 3). Кроме того, аргинин 175 образует солевые мостики с атомами ионов сульфата, расположенными на оси третьего порядка. Отметим, что в структуре T91A ионы сульфата на кристаллографической оси не выявлены. Сравнение структур T91A и T91S в области связывания этого сульфат-иона показано на рис. 3. Связывание одного иона сульфата на оси третьего порядка приводит к увеличению свободной энергии диссоциации гексамера на 13 ккал/моль.

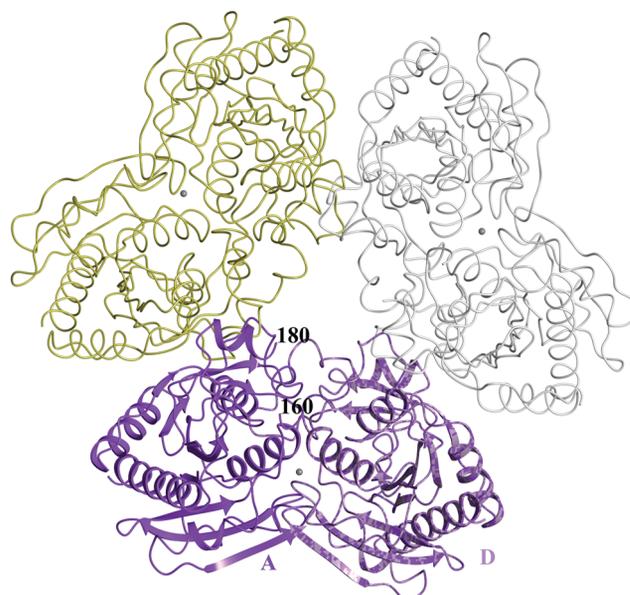


Рис. 2. Ход полипептидной цепи в гексамерной молекуле T91A. Кристаллографически независимые мономеры A и D показаны как модели с β -слоями и α -спиралями. Димеры, связанные с исходным димером кристаллографической осью третьего порядка, показаны в виде ленточной модели с разными оттенками. В мономере A обозначены номера остатков, между которыми не локализована петля (160–180).

Сравнение мономера D структуры T91S в закрытой конформации и мономера A структуры T91A в открытой конформации с разупорядоченной петлей 161–179 приведено на рис. 4. Смещение петли в районе остатка 224 от активного центра, наблюдаемое в открытой конформации, может обеспечивать доступ молекулы нуклеозида в активный центр фермента. Дополнительное разупорядочение

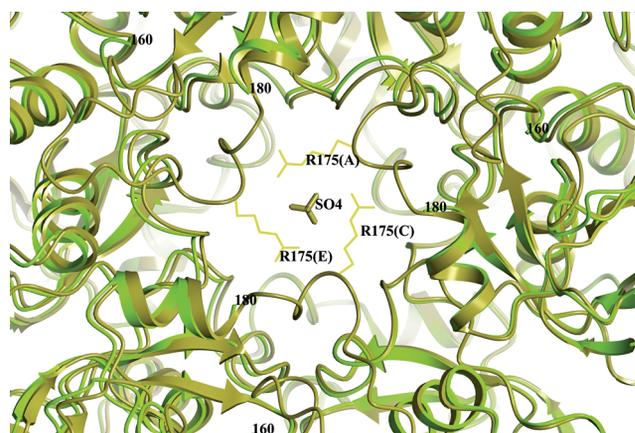


Рис. 3. Сравнение структур T91A (мономеры A, E и C; темный цвет) и T91S (мономер A; светлый цвет) для области связывания сульфат-иона на оси третьего порядка гексамера.

порядочение петли 161–179 в структуре мономера в открытой конформации значительно облегчает встраивание молекулы нуклеозида в активный центр фермента.

Анализ структур УФ из PDB, решенных с разрешением выше 2 Å, показал, что все УФ из бактерий образуют гексамеры. В большинстве структур гексамеры имеют некристаллографическую ось симметрии третьего порядка. В таких структурах гексамеры симметричны и аналогичны структуре гексамера T91A. Однако для четырех бактериальных УФ из семи исследованных были получены кристаллические модификации, содержащие в структуре гексамеры с кристаллографической осью третьего порядка. При этом кристаллографически независимые мономеры в димерах могут быть как симметричными (*E. coli*, PDB_ID 1rxu; *Yersinia pseudotuberculosis*, PDB_ID 4ny1), так и несимметричными (структуры мутантной формы T91A из *Shewanella oneidensis* и УФ из *Salmonella typhimurium* (PDB_ID oxf, 3hsw, 1y1t)). В случае формирования несимметричного димера одна из молекул димера частично разупорядочена (отсутствуют остатки в районе петли 161–179).

Считается, что субстраты способны проникать в активный центр фермента, находящегося в открытой конформации, и связываться там. Однако сравнение структур мономеров в закрытой конформации из симметричных димеров и структур мономеров несимметричных димеров с разупорядоченной петлей 161–179 (на примере структур мономеров мутантных форм T91S и T91A) показывает, что проникновение молекулы нуклеозида в активный центр существенно облегчается в случае разупорядоченной петли 161–179 в несимметричных гексамерах (рис. 4).

Можно предположить, что встраивание субстрата в активный центр УФ происходит именно в мономеры с разупорядоченной петлей в районе остатков 161–179. При этом объемная молекула нуклеозида встраивается в активный центр первой и стабилизирует петлю в районе остатков 161–179 (два остатка из этой петли, глутамин и аргинин, образуют связи с атомами основания). Связывание нуклеозида приводит к стабилизации мономера УФ в открытой конформации. Далее происходит связывание фосфат-иона в активном центре фермента, приводящее к изменению конформации мономера на закрытую. Одновременно с этим гексамерная молекула стабилизируется за счет связывания сульфат/фосфат-ионов на оси третьего порядка. Ферментативная реакция происходит в мономерах в закрытой конформации. Благодаря меньшему размеру (по сравнению с нуклеозидом) продуктов реакции выход основания нуклеотида и рибозофосфата из активного центра не требует разупорядочения петли в районе остатков 161–179.

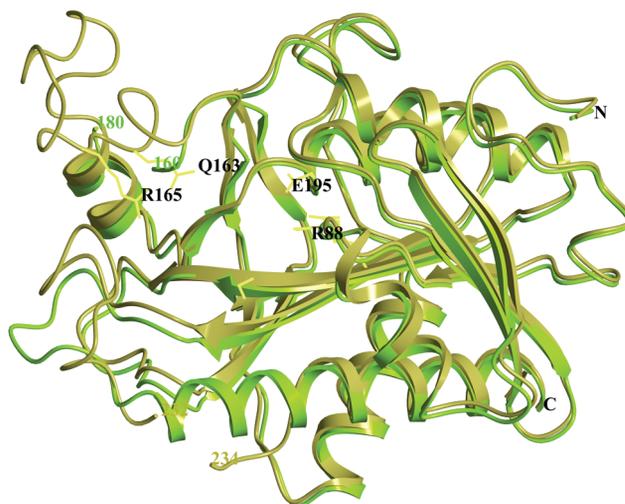


Рис. 4. Сравнение мономера А структуры T91A (темный цвет) и мономера D структуры T91S (светлый цвет). Мономер А структуры T91A находится в открытой конформации (остаток 224 смещен в сторону от остатка 165, N-концевой фрагмент последней α -спирали упорядочен). В этой структуре не локализована петля 160–180. Мономер D структуры T91S находится в закрытой конформации (остаток 224 смещен в сторону остатка 165, N-концевой фрагмент последней α -спирали упорядочен). В этой структуре не локализована петля 232–234. Для структуры T91A показаны остатки активного центра.

Присутствие ионов фосфата/сульфата на оси третьего порядка в структуре гексамера (рис. 4) может стабилизировать структуру за счет взаимодействия этих ионов с остатками аргинина 175 трех мономеров. По-видимому, ионы фосфата/сульфата принимают участие в разупорядочении петли в районе 186 остатка.

Таким образом, на основании анализа трехмерных структур мутантов T91A и T91S показано, что конформационные изменения, связанные с разупорядочением гибкой петли в районе остатков 161–179 бактериальной УФ, играют важную роль в связывании субстрата и осуществлении ферментативной реакции.

Работы в части кристаллизации мутантных форм УФ, сбора дифракционных наборов данных проведены в рамках выполнения работ по Государственному заданию Института биохимии им. А.Н. Баха, ФНИЦ “Фундаментальные основы биотехнологии”, в части получения мутантных форм УФ, определения каталитической активности белка, решения, уточнения и анализа трехмерных структур мутантов – при поддержке Российского фонда фундаментальных исследований (грант № 18-04-00784а). При проведении исследований использовали оборудование Центра коллективного пользования “Промышленные биотехнологии” Федерального исследовательского центра “Фундаментальные основы биотехнологии” РАН.

СПИСОК ЛИТЕРАТУРЫ

1. *Cao D., Pizzorno G.* // *Drugs Today (Barc)*. 2004. V. 40. P. 431.
<https://doi.org/10.1358/dot.2004.40.5.850491>
2. *Cappiello M., Mascia L., Scolozzi C. et al.* // *Biochim. Biophys. Acta*. 1998. V. 1425. P. 273.
[https://doi.org/10.1016/s0304-4165\(98\)00071-3](https://doi.org/10.1016/s0304-4165(98)00071-3)
3. *Tozzi M.G., Camici M., Mascia L. et al.* // *FEBS J*. 2006. V. 273. P. 1089.
<https://doi.org/10.1111/j.1742-4658.2006.05155.x>
4. *Luccioni C., Beaumatin J., Bardot V. et al.* // *Int. J. Cancer*. 1994. V. 58. P. 517.
<https://doi.org/10.1002/ijc.2910580411>
5. *Watanabe S., Uchida T.* // *Biochem. Biophys. Res. Commun*. 1995. V. 216. P. 265.
<https://doi.org/10.1006/bbrc.1995.2619>
6. *Safonova T.N., Mikhailov S.N., Veiko V.P. et al.* // *Acta Cryst. D*. 2014. V. 70. P. 3310.
<https://doi.org/10.1107/S1399004714024079>
7. *Koellner G., Bzowska A., Wielgus-Kutrowska B. et al.* // *J. Mol. Biol.* 2002. V. 315. P. 351.
<https://doi.org/10.1006/jmbi.2001.5211>
8. *Мордкович Н.Н., Манувера В.А., Вейко В.П. и др.* // *Биотехнология*. 2012. № 1. С. 21.
9. *Мордкович Н.Н., Сафонова Т.Н., Антипов А.Н. и др.* // *Прикладная биохимия и микробиология*. 2018. Т. 54. № 1. С. 16.
<https://doi.org/10.7868/S0555109918010026>
10. *Bradford M.M.* // *Anal. Biochem*. 1976. V. 2. P. 248.
<https://doi.org/10.1006/abio.1976.9999>
11. *Leer J.C., Hammer-Jespersen K., Schwartz M.* // *Eur. J. Biochem*. 1977. V. 75. P. 217.
<https://doi.org/10.1111/j.1432-1033.1977.tb11520.x>
12. *Kabsch W.* // *Acta. Cryst. D*. 2010. V. 66. P. 125.
<https://doi.org/10.1107/S0907444909047337>
13. *Vagin A.A., Teplyakov A.V.* // *J. Appl. Cryst*. 1997. V. 30. P. 1022.
<https://doi.org/10.1107/S0021889897006766>
14. *Murshudov G.N., Skubák P., Lebedev A.A. et al.* // *Acta. Cryst. D*. 2011. V. 67. P. 355.
<https://doi.org/10.1107/S0907444911001314>
15. *Emsley P., Lohkamp B., Scott W. et al.* // *Acta. Cryst. D*. 2010. V. 66. P. 486.
<https://doi.org/10.1107/S0907444910007493>
16. *McNicholas S., Potterton E., Wilson K.S. et al.* // *Acta. Cryst. D*. 2011. V. 67. P. 386.
<https://doi.org/10.1107/S0907444911007281>
17. *Winn M.D., Ballard C.C., Cowtan K.D. et al.* // *Acta. Cryst. D*. 2011. V. 67. P. 235.
<https://doi.org/10.1107/S0907444910045749>
18. *Krissinel E., Henrick K.* // *J. Mol. Biol.* 2007. V. 372. P. 774.
<https://doi.org/10.1016/j.jmb.2007.05.022>

СТРУКТУРА МАКРОМОЛЕКУЛЯРНЫХ СОЕДИНЕНИЙ

УДК 548.73

СТРУКТУРА МУТАНТНОЙ ФОРМЫ БЕЛКА-РЕГУЛЯТОРА ТРАНСЛЯЦИИ Hfq С УДЛИНЕННОЙ ПЕТЛЕЙ L4

© 2021 г. А. А. Алипов¹, Н. В. Леконцева¹, А. О. Михайлина¹, М. С. Фандо¹,
С. В. Тищенко¹, А. Д. Никулин^{1,*}

¹ Институт белка РАН, Пущино, Россия

*E-mail: nikulin@vega.protres.ru

Поступила в редакцию 17.09.2020 г.

После доработки 03.10.2020 г.

Принята к публикации 06.10.2020 г.

Бактериальный белок Hfq принадлежит к обширному семейству Lsm белков, имеющих консервативную третичную структуру и формирующих стабильную четвертичную структуру в виде колец из нескольких мономеров. Бактериальные белки формируют гексамеры, а их архейные и эукариотические гомологи – гептамеры. Причина формирования высоко гомологичными Lsm белками с одинаковой третичной структурой мультимеров с различающимся числом мономеров не установлена и является предметом активной дискуссии. Экспериментально проверено влияние вариабельной по длине петли L4 на четвертичную структуру белка Hfq. Полученный белок сохраняет гексамерную четвертичную структуру, а удлиненная петля отходит от поверхности центральной части белка в разные стороны. Таким образом, не удалось изменить четвертичную структуру белка, что подтверждает ее высокую стабильность. При этом РНК-связывающие свойства мутантной формы белка практически не изменились.

DOI: 10.31857/S0023476121050027

ВВЕДЕНИЕ

РНК-связывающий белок Hfq является регулятором трансляции большого числа бактериальных генов [1–3]. Он способствует специфическому взаимодействию малых регуляторных РНК с мРНК, что приводит к активации или репрессии трансляции. Белок Hfq принадлежит к обширному семейству Lsm белков, представители которого имеются в бактериях, археях и эукариотах. Их объединяет консервативная пятиклетчатая β -структура с N-концевой α -спиралью общей протяженностью ~70 аминокислотных остатков (а.о.). На C-конце у Lsm белков наблюдается неупорядоченная и не консервативная последовательность от 3 до 30 а.о. [4]. Характерная особенность Lsm белков – формирование ими кольцевой четвертичной структуры из шести (бактериальные Hfq) или семи (эукариотические Sm/Lsm и архейные SmAP) мономеров [4]. Имеются единичные примеры формирования белками этого семейства пентамеров [5] и октамеров [6], но это исключение из общего правила. Начиная со времени объединения Lsm белков в один класс, перед исследователями стоял вопрос о причине формирования столь близкими по пространственной структуре белками четвертичной структуры с разным числом мономеров. Выдвигались гипотезы о влия-

нии на четвертичную структуру C-концевой части белка [7–10] и петли L4 между тяжами $\beta 3$ и $\beta 4$ [11, 12]. Первая гипотеза не получила подтверждения, поскольку все известные бактериальные белки Hfq формируют только гексамеры, несмотря на различную длину C-концов – от 3 до 25 а.о., а архейные и эукариотические белки формируют гептамеры, несмотря на значительное различие C-концевых последовательностей. Более того, экспериментально показано, что делеция C-концевой части белка Hfq не сказывается на его четвертичной структуре [13, 14]. Вторая гипотеза имеет больше оснований: все гексамерные белки имеют короткую последовательность петли L4 (включая Hfq-подобный архейный белок из *Methanococcus jannaschii*), а все гептамерные – удлиненную последовательность этой петли. Исключением является белок AF-SM2 (SmAP2 белок из археи *Archeoglobus fulgidus*), который может менять четвертичную структуру в зависимости от pH раствора или присутствия РНК [12]. Эта гипотеза неоднократно обсуждалась в литературе, но экспериментальной проверки влияния длины петли на четвертичную структуру Lsm белков не проводилось. В представленной работе получена генно-инженерная конструкция, несущая ген белка Hfq из *P. aeruginosa* со вставкой последовательности, кодирующей протяженную петлю L4

```

                *           20           *           40           *
AF-SM1 :      MPPRPLDVLNRSLSKSPVIVRLKGGREFRGTLDGYDIHMNLVLLDA
PaeHfqL4: MSKGHSLQDPYLNTLRKERVPVSIYLVNGIKLQGQIESFDQFVILLKNTA
PaeHfq : MSKGHSLQDPYLNTLRKERVPVSIYLVNGIKLQGQIESFDQFVILLKNT-

                60           *           80           *
AF-SM1 :      EEIQNGEVVRKVGSVVIRGDTVVFVSPAPGGE
PaeHfqL4: EEIQNGEVVRKVVSQMVYKHAISTVVPSRPVRLPSGDQPAEPGNA
PaeHfq : -----VSQMVYKHAISTVVPSRPVRLPSGDQPAEPGNA

```

Рис. 1. Сравнение последовательностей белков AF-SM1 из *Archaeoglobus fulgidus*, Hfq из *Pseudomonas aeruginosa* дикого типа (PaeHfq) и с удлиненной петлей L4 (PaeHfqL4). Выделена последовательность удлиненной петли L4.

архейного SmAP белка. Белок был выделен, охарактеризованы его РНК-связывающие свойства, определена пространственная структура с помощью рентгеноструктурного анализа.

МАТЕРИАЛЫ И МЕТОДЫ

Клонирование и получение штамма-продуцента. Генетическую конструкцию, кодирующую ген белка PaeHfqL4 со вставкой, получили методом сайт-направленного мутагенеза, следуя протоколу QuikChange® Site-Directed Mutagenesis Kit (Stratagene, США). Нуклеотидная последовательность вставки соответствовала последовательности гена SmAP белка AF-SM1 из *Archaeoglobus fulgidus* (UNIPROT O29386), кодирующего длинную петлю L4 (рис. 1). Используемые праймеры (Синтол, Россия) содержали необходимую вставку:

F-PaeHfqL4:

5'-ТТCAGAACGGCGAAGTGGTGCCTAAAG
TGCTCAGCCAGATGGTTTACA-3'

R-PaeHfqL4:

5'-CACTTCGCCGTTCTGAATTTCTTCCGC-
GGTGTCTTTCAGCAGGATGA-3'.

Для амплификации ДНК использовали *KOD Hot Start* ДНК-полимеразу. ДНК-матрицей в ПЦР служила плаزمида рЕТ-22b(+)-PaeHfq [15], несущая ген белка Hfq из *P. aeruginosa* дикого типа.

Выделение и очистку белков дикого типа PaeHfq и его мутантной формы PaeHfqL4 проводили по схеме, разработанной в [15]. Дополнительной стадией очистки стала гель-фильтрация. Сконцентрированный до 1 мл препарат белка наносили на колонку со смолой Superdex G75 (объем колонки 120 мл), уравновешенную буфером 200 мМ NaCl, 50 мМ Трис-НСl, рН 8.0.

Получение фрагментов РНК. Для анализа РНК-белковых взаимодействий использовали фрагменты двух разных РНК. Фрагмент малой некодирующей РНК DsrA длиной 42 нуклеотида (**DsrA42**; 5'-GGGUGUAAACGAAUUUUUAUUC-GUUUGAAAUUCGUUACACCC-3') и фрагмент 5'-НТО *rpoS* мРНК длиной 40 нуклеотидов (**rpoS40**; 5'-GGCAUUUUGAAAUUCGUUACAAG-

GGGAAAUCCGUAAACCC-3') нарабатывали транскрипцией *in vitro* с рестрицированных по сайту EcoRI плазмид рUC18, содержащих гены фрагментов РНК под контролем промотора РНК-полимеразы фага Т7. Получение и очистка DsrA42 РНК описаны в [16]. Фрагмент гена мРНК rpoS40 клонировали в вектор рUC18 с использованием перекрывающихся праймеров (Синтол, Россия), причем прямой олигонуклеотид содержал последовательность Т7 промотора:

F-rpoS40:

5'-СТАCTGCAAAGCTТААТACGACTCACTA
TAGGGCAUUUUGAAAUUCGUUACAAGG-3'

R-rpoS40:

5'-АТСТАGCCC GG GTT TACGGATTTCCCC
TTGTAACGAATTTС-3'.

Очистку наработанного транскрипцией *in vitro* фрагмента мРНК rpoS40 проводили по методике для очистки DsrA42 [16].

Кинетический анализ взаимодействия белков PaeHfq и PaeHfqL4 с фрагментами РНК проводили методом поверхностного плазмонного резонанса [17] на системе ProteOn XPR36 (Bio-Rad, США). Присоединение к фрагментам РНК биотинилированного олигонуклеотида проводили как описано в [16].

Биотинилированный фрагмент РНК наносили на сенсорный NLC-чип (Bio-Rad, США) с иммобилизованным авидином до достижения единиц ответа 500–1000. Поверхность чипа отмывали от несвязавшейся РНК 0.1%-ным раствором додецилсульфата натрия. Затем пропускали препарат белка с различной концентрацией (0.15–1.15 нМ) со скоростью 30 мкл/мин. Серии разведений из пяти концентраций анализатора (PaeHfq или PaeHfqL4) готовили с использованием буфера, содержащего 50 мМ Трис-НСl, рН_{25°C} 7.5, 175 мМ NaCl, 2 мМ MgCl₂, 0.05% Tween-20. Время ассоциации составляло 360 с, диссоциации – 2500 с. Все эксперименты по связыванию проводили при температуре 25°C. Серию из трех–пяти сенсogramм обсчитывали в программе BIAEvaluation с использованием модели Ленгмюра (1:1).

Кристаллизацию белка *RaeHfqL4* проводили методом диффузии паров в висящей капле при 23°C. Препарат белка (1–2 мкл) с концентрацией 12 мг/мл в 200 мМ NaCl, 50 мМ Трис-HCl, pH 8.0, смешивали с равным объемом противораствора. Кристаллы белка получены в условиях № 3 набора JBScreen Nuc-Pro 1 (Jena Bioscience, Германия): 5% ПЭГ400, 50 мМ PIPES 7.0, 30 мМ MgCl₂. Кристаллы появлялись в течение 3–5 дней и достигали размера 0.4 × 0.4 × 0.02 мм. Перед замораживанием кристаллы вымачивали в криорастворе с увеличенным содержанием ПЭГ 400.

Получение наборов кристаллографических данных. Дифракционный набор данных получен на детекторе DECTRIS PILATUS3S 6M линии ID23-1 синхротрона ESRF (Гренобль, Франция) с кристалла, охлажденного до температуры 100 К, на длине волны 0.97242 Å. Для сбора данных использовали метод качания–вращения с углом осцилляции 0.2°, суммарный угол вращения составил 200°. Первичные дифрактограммы были обработаны в программе XDS [18] и интегрированы в программе AIMLESS комплекса CCP4 [19]. Детали результата обработки и статистические характеристики набора данных представлены в табл. 1.

Определение и уточнение структуры белка *RaeHfqL4*. Структуру белка определяли методом молекулярного замещения с использованием программы MorDA [20]. Исходная модель для решения была выбрана программой автоматически, на основе заданной аминокислотной последовательности белка *RaeHfqL4*. Найденное решение оценивали на соответствие экспериментальным данным по картам электронной плотности (ЭП) 2mFo-DFc и mFo-DFc. Уточнение структуры белка осуществляли в программе PHENIX [21]. На первом этапе проводили уточнение ориентации мономеров белка в гексамере, рассматривая каждый мономер в качестве независимого “твёрдого тела” с приписанным общим B-фактором для всех атомов, входящих в одну из шести субъединиц (rigid body refinement). На втором этапе уточнения использовали протокол “симуляции отжига” – молекулярной динамики с ограничениями по имеющимся экспериментальным дифракционным данным (simulated annealing). На этом этапе уточняли индивидуальные значения B-фактора для каждого атома.

После этих этапов в результирующей карте ЭП идентифицировали участок, который соответствовал добавленной к структуре белка Hfq петли L4 длиной 13 а.о., и в соответствующий участок ЭП вписывали недостающие а.о. с помощью программы SOOT [22]. После завершения итеративного дополнения аминокислотных последовательностей мономеров белка до полных проводили несколько раундов уточнения структуры. На первых этапах использовали протокол “симуля-

Таблица 1. Статистические характеристики экспериментальных данных и уточнения структуры белка *RaeHfqL4* (PDB ID 6XYJ)

Длина волны, на которой получили дифракционные данные, Å	0.9724
Область разрешения, Å	50–2.79 (2.89–2.79)
Пр. гр. ячейки кристалла	C ₂ ₁
<i>a</i> , <i>b</i> , <i>c</i> , Å; α, β, γ, град	76.8, 112.6, 78.0; 90.0, 119.4, 90.0
Общее число отражений (рефлексов)	52850 (4964)
Из них уникальных	14417 (1396)
Избыточность	3.7 (3.6)
Полнота, %	98.7 (93.0)
Среднее значение <i>I</i> /σ(<i>I</i>)	12.2 (0.8)
V-фактор Вильсона	76.2
<i>R</i> -merge	0.079 (1.60)
<i>R</i> -meas	0.093 (1.88)
<i>R</i> -pim	0.047 (0.965)
CC _{1/2}	1.0 (0.56)
CC*	1.0 (0.85)
Число отражений, использованных при уточнении	14320 (1320)
Число отражений, использованных для расчета <i>R</i> -free	710 (68)
<i>R</i> _{work}	0.218 (0.435)
<i>R</i> _{free}	0.279 (0.410)
CC _{work}	0.954 (0.604)
CC _{free}	0.929 (0.651)
Число неводородных атомов в модели	3862
СКО для длин связей, Å	0.013
СКО для валентных углов, град	1.79
Средний B-фактор для атомов	105.8
Число аминокислотных остатков на карте Рамачандрана	
В предпочтительных областях, %	93
В допустимых областях, %	7
В запрещенных областях, %	0

Примечание. Значения в скобках приведены для области высокого разрешения 2.89–2.79 Å.

ции отжига” с уточнением индивидуальных значений B-фактора для каждого атома. На последних этапах уточнения, когда проводили тонкую подгонку значений длин связей, валентных и торсионных углов для подготовки финальной модели, использовали протокол без “симуляции отжига”, но с оптимизацией весовых параметров соотношения геометрии/свободной энергии модели.

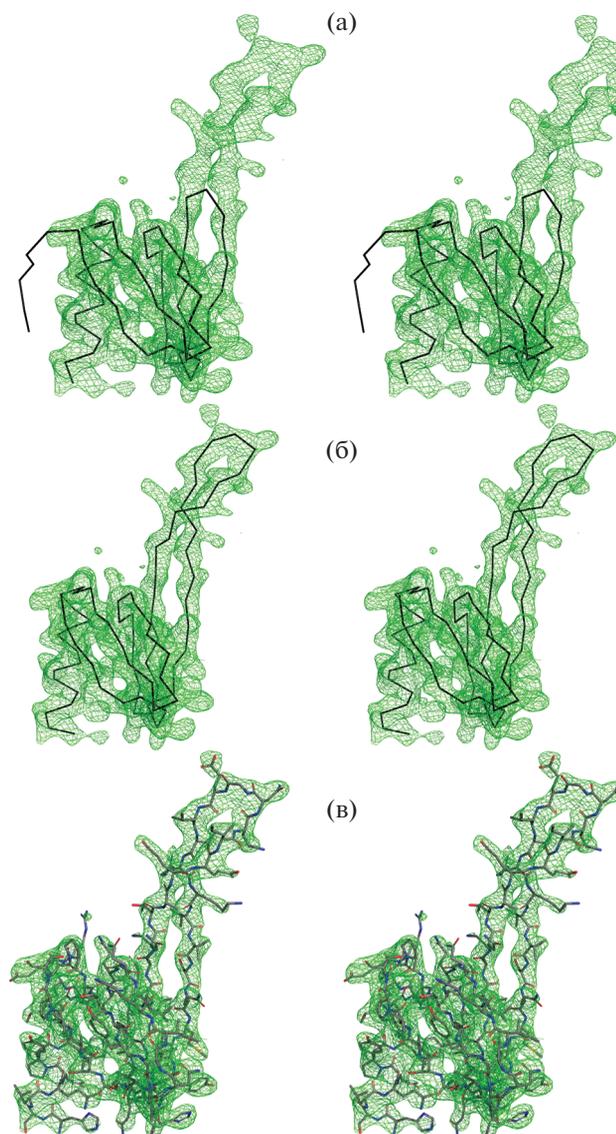


Рис. 2. Этапы дополнения исходной структуры белка PaеHfq до полной структуры мутантной формы белка с удлиненной петлей L4 на примере одного из мономеров белка (стереопары). Карта электронной плотности 2Fo-DF показана в виде сетки со срезкой 1 σ , разрешение 2.8 Å; а – наложение α -цепи исходной структуры PaеHfq на карту ЭП; б – наложение α -цепи модифицированной модели с добавленными а.о. удлиненной петли L4; в – наложение достроенной модели PaеHfqL4 на карту ЭП.

Перед депонированием структуры белка в банк данных проводили тщательную проверку геометрии белка с помощью сервиса MolProbity, встроенного в программный комплекс PHENIX. Найденные ошибки исправляли с помощью программы молекулярной графики COOT и подвергали повторному уточнению. Результирующая структура была уточнена в программе REFMAC комплекса CCP4 и депонирована в банк данных с присвоением PDBID 6XYJ. Статистические ха-

рактеристики уточнения структуры приведены в табл. 1.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Структура белка Hfq из P. aeruginosa с удлиненной петлей L4. Для определения структуры белка Hfq из *P. aeruginosa* с удлиненной петлей L4 (PaеHfqL4) использовали метод молекулярного замещения. Поскольку ожидалось, что число мономеров в исследуемом белке будет отличаться от их количества в исходной структуре, для решения задачи молекулярного замещения использовали программу MorDA. Она позволяет при поиске решения задачи молекулярного замещения в автоматическом режиме тестировать ряд возможных вариантов исходных моделей, автоматически отбираемых из банка данных на основе гомологии первичных структур. Такой режим работы программы в данном случае позволяет значительно ускорить нахождение решения при неопределенности числа мономеров в асимметричной части ячейки кристалла. В результате работы программы наилучшей исходной моделью был выбран мономер белка Hfq из *P. aeruginosa*, причем асимметричная часть кристаллической ячейки содержала гексамер белка, а не гептамер, как ожидалось. Карта ЭП, рассчитанная по ориентированной модели гексамера белка, показала наличие для каждого из мономеров белка неопisanного участка плотности в области между тяжами β 3 и β 4 (рис. 2). Этот участок был дополнен а.о., соответствующими последовательности вставки в петлю L4. Модель уточнена до хороших значений стереохимических параметров (с учетом разрешения кристаллографических данных 2.8 Å).

Полученная структура мутантной формы белка PaеHfq с удлиненной петлей L4 показала, что мономер белка сохраняет трехмерную структуру из N-концевой α -спирали и пяти β -тяжей. Вставка в петлю L4 из 13 а.о. в белке PaеHfqL4 приобрела конформацию β -шпильки (рис. 3). Однако в отличие от архейных гомологов, где петля L4 формирует плотно прилегающую к поверхности гептамера β -шпильку, в рассматриваемом случае удлиненная петля L4 отходит от основной части белка в разные стороны, а белок остается гексамером (рис. 4). В результате ожидаемого изменения четвертичной структуры белка не произошло. Возможными причинами этого могут быть компактность и высокая стабильность гексамера Hfq, не позволяющая сближать петли к его центру, а различия в последовательностях бактериального и архейных белков не способствуют формированию гидрофобных контактов и/или водородных связей между а.о. шпильки L4 и коровой частью белка, что способствовало бы “притягиванию” шпильки к ее поверхности. Таким образом, пока нельзя полностью отвергать гипотезу о су-

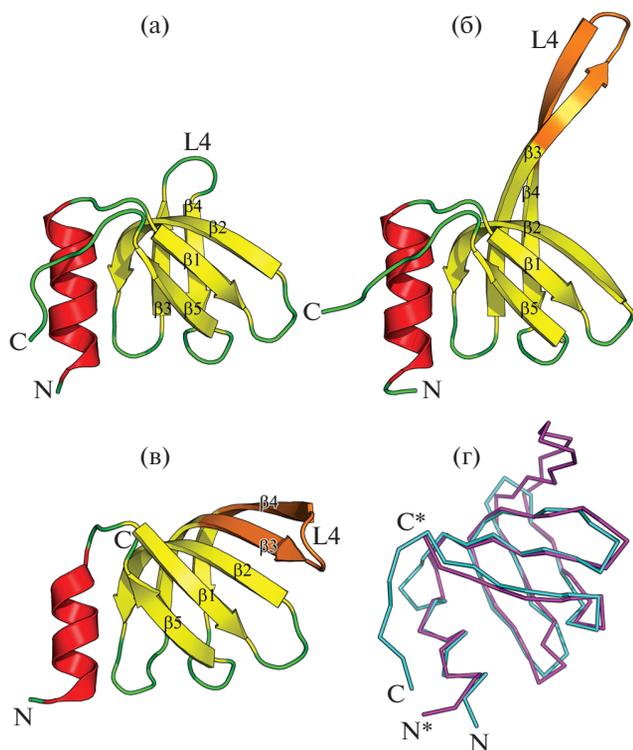


Рис. 3. Структура мономеров исходного белка PaеHfq (а), мутантной формы белка PaеHfqL4 (б), мономера SmAP белка из *Archaeoglobus fulgidus* (AF-SM1) (в). Обозначены N- и C-концы мономеров, элементы β -слоя, положение петли L4. Цветом выделены удлиненные относительно PaеHfq петли L4 в мономерах PaеHfqL4 и AF-SM1, имеющие конформацию β -шпиклек. Ориентация мономеров одинаковая относительно центральной части белка. Наложение главных цепей мономеров белков Hfq (светло-серый) и AF-SM1 (черный) (г). Обозначены N- и C-концы мономеров.

ществленном влиянии длины петли L4 на четвертичную структуру Lsm белков. Окончательный ответ можно получить только при создании мутантной формы Hfq с прилегающей к поверхности мономера длинной петли L4.

Влияние удлинения петли L4 на РНК-связывающие свойства белка PaеHfq. Известно, что белок Hfq действует как РНК-шаперон, облегчая взаимодействие между малыми регуляторными РНК (мРНК) и их мРНК-мишенями [2]. Так, после взаимодействия Hfq с лидерной последовательностью *rpoS* мРНК происходит связывание мРНК с DsrA мРНК [23, 24]. Формирование тройного комплекса мРНК–Hfq–мРНК приводит к образованию дуплекса *rpoS*–DsrA и последующей инициации трансляции мРНК. Участки связывания мРНК и мРНК на белке различаются: мРНК имеют, как правило, уридин-богатые участки и взаимодействуют с так называемым уридин-связывающим участком на белке, расположенным вблизи петли L4, а мРНК связывается с противоположной стороны гексамера [2]. Третий РНК-связывающий участок белка Hfq расположен на боковой поверхности гексамера белка и является “местом встречи” мРНК и мРНК [25–28]. Для анализа влияния увеличенной петли L4 на взаимодействие белков со специфическими фрагментами мРНК DrsA и мРНК *rpoS* были определены константы диссоциации РНК-белковых комплексов (табл. 2). Показано, что удлинение петли L4 практически не изменило сродство белка к фрагменту лидерной последовательности мРНК *rpoS*. Поскольку участок связывания мРНК удален от петли L4, это вполне ожидаемый результат. Сродство мутантной формы белка к

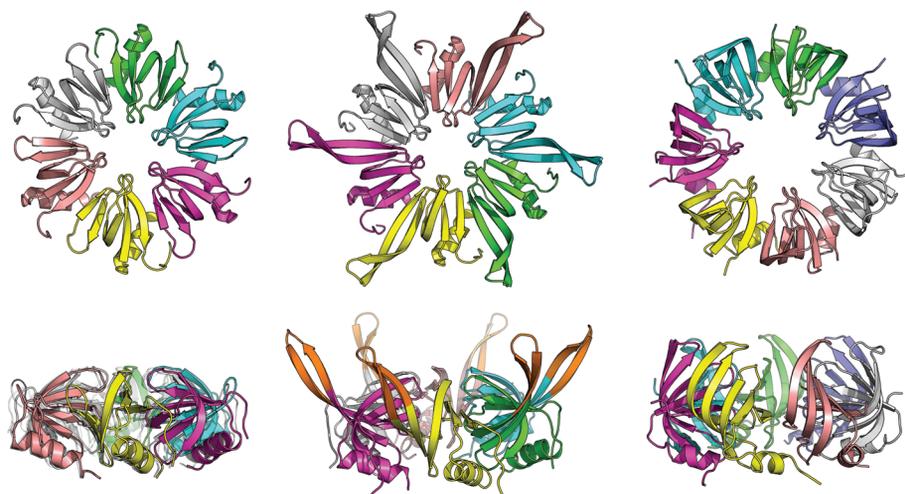


Рис. 4. Сравнение четвертичной структуры белка PaеHfq дикого типа (слева), его мутантной формы PaеHfqL4 (в центре) и архейного гомолога AF-SM1 (справа). Вверху – вид со стороны дистальной (противоположной стороне с α -спиралями) поверхности белков; внизу – вид сбоку. Участки петель L4 выделены цветом.

Таблица 2. Константы диссоциации (K_D) РНК-белковых комплексов

	K_D , нМ	
	РacHfq	РacHfqL4
DsrA42	0.1 ± 0.01	0.04 ± 0.004
RpoS40	0.03 ± 0.002	0.02 ± 0.007

DsrA мРНК несколько увеличилось, что можно объяснить наличием дополнительных положительно заряженных а.о.в удлинённой петле. Тем не менее удлинённая петля L4 в целом не оказала значительного влияния на РНК-связывающие свойства белка Hfq, поскольку находится в отдалении от его РНК-связывающих участков.

Работа выполнена при поддержке Российского фонда фундаментальных исследований, проект № 18-04-00222.

СПИСОК ЛИТЕРАТУРЫ

1. Васильева Ю.М., Гарбер М.Б. // Молекулярная биология. 2002. Т. 36. № 6. С. 1.
2. Updegrave T.B., Zhang A., Storz G. // Curr. Opin. Microbiol. 2016. V. 30. P. 133. <https://doi.org/10.1016/j.mib.2016.02.003>
3. Sauer E. // RNA Biol. 2013. V. 10. № 4. P. 610. <https://doi.org/10.4161/rna.24617>
4. Мурина В.Н., Никулин А.Д. // Успехи биологической химии. 2011. Т. 51. С. 133.
5. Das D., Kozbial P., Axelrod H.L. et al. // Proteins. 2010. V. 75. № 2. P. 296. <https://doi.org/10.1002/prot.22360>
6. Naidoo N., Harrop S.J., Sobti M. et al. // J. Mol. Biol. 2008. V. 377. № 5. P. 1357. <https://doi.org/10.1016/j.jmb.2008.01.007>
7. Thore S., Mayer C., Sauter C. et al. // J. Biol. Chem. 2003. V. 278. № 2. P. 1239. <https://doi.org/10.1074/jbc.M207685200>
8. Fortas E., Piccirilli F., Malabirade A. et al. // Biosci. Rep. 2015. V. 35. № 2. P. 1. <https://doi.org/10.1042/BSR20140128>
9. Arluison V., Folichon M., Marco S. et al. // Eur. J. Biochem. 2004. V. 271. № 7. P. 1258. <https://doi.org/10.1111/j.1432-1033.2004.04026.x>
10. Beich-Frandsen M., Vecerek B., Konarev P.V. et al. // Nucl. Acids Res. 2011. V. 39. № 11. P. 4900. <https://doi.org/10.1093/nar/gkq1346>
11. Sauter C., Basquin J., Suck D. // Nucl. Acids Res. 2003. V. 31. № 14. P. 4091. <https://doi.org/10.1093/nar/gkg480>
12. Törö I., Basquin J., Teo-Dreher H., Suck D. // J. Mol. Biol. 2002. V. 320. № 1. P. 129. [https://doi.org/10.1016/S0022-2836\(02\)00406-0](https://doi.org/10.1016/S0022-2836(02)00406-0)
13. Beich-Frandsen M., Vecerek B., Sjoblom B. et al. // Acta Cryst. F. 2011. V. 67. № 5. P. 536. <https://doi.org/10.1107/S174430911100786X>
14. Olsen A.S., Møller-Jensen J., Brennan R.G., Valentin-Hansen P. // J. Mol. Biol. 2010. V. 404. № 2. P. 173. <https://doi.org/10.1016/j.jmb.2010.09.038>
15. Nikulin A., Stolboushkina E., Perederina A. et al. // Acta Cryst. D. 2005. V. 61. № 2. P. 141. <https://doi.org/10.1107/S0907444904030008>
16. Nikulin A., Mikhailina A., Lekontseva N. et al. // J. Biomol. Struct. Dyn. 2017. V. 35. № 8. P. 1615. <https://doi.org/10.1080/07391102.2016.1189849>
17. Katsamba P.S., Park S., Laird-Offringa I.A. // Methods. 2002. V. 26. № 2. P. 95. [https://doi.org/10.1016/S1046-2023\(02\)00012-9](https://doi.org/10.1016/S1046-2023(02)00012-9)
18. Kabsch W. // Acta Cryst. D. 2010. V. 66. № 2. P. 125. <https://doi.org/10.1107/S0907444909047337>
19. Winn M.D., Ballard C.C., Cowtan K.D. et al. // Acta Cryst. D. 2011. V. 67. № 4. P. 235. <https://doi.org/10.1107/S0907444910045749>
20. Vagin A., Lebedev A. // Acta Cryst. A. 2015. V. 71. № a1. P. s19. <https://doi.org/10.1107/S2053273315099672>
21. Afonine P.V., Grosse-Kunstleve R.W., Echols N. et al. // Acta Cryst. D. 2012. V. 68. № 4. P. 352. <https://doi.org/10.1107/S0907444912001308>
22. Emsley P., Cowtan K. // Acta Cryst. D. 2004. V. 60. № 12. P. 2126. <https://doi.org/10.1107/S0907444904019158>
23. Sledjeski D.D., Whitman C., Zhang A. // J. Bacteriol. 2005. V. 183. № 6. P. 1997. <https://doi.org/10.1128/JB.183.6.1997-2005.2001>
24. Vecerek B., Beich-Frandsen M., Resch A., Bläsi U. // Nucl. Acids Res. 2010. V. 38. № 4. P. 1284. <https://doi.org/10.1093/nar/gkp1125>
25. Chen J., Morita T., Gottesman S. // Front. Cell. Infect. Microbiol. 2019. V. 9. № 6. P. 201. <https://doi.org/10.3389/fcimb.2019.00201>
26. Sauer E., Schmidt S., Weichenrieder O. // Proc. Natl. Acad. Sci. 2012. V. 109. № 24. P. 9396. <https://doi.org/10.1073/pnas.1202521109>
27. Robinson K.E., Orans J., Kovach A.R. et al. // Nucl. Acids Res. 2013. V. 42. № 4. P. 2736. <https://doi.org/10.1093/nar/gkt1171>
28. Panja S., Schu D.J., Woodson S. A. // Nucl. Acids Res. 2013. V. 41. № 15. P. 7536. <https://doi.org/10.1093/nar/gkt521>

СТРУКТУРА МАКРОМОЛЕКУЛЯРНЫХ
СОЕДИНЕНИЙ

УДК 577.322.63

ВЛИЯНИЕ МИССЕНС-МУТАЦИИ Pе222Thr В SsoIF2γ
НА СРОДСТВО γ- И β-СУБЪЕДИНИЦ aIF2© 2021 г. О. С. Никонов^{1,*}, О. В. Кравченко¹, Н. А. Невская¹, Е. А. Столбоушкина¹,
М. Б. Гарбер¹, С. В. Никонов¹¹ Институт белка РАН, Пущино, Россия

*E-mail: alik@vega.protres.ru

Поступила в редакцию 08.09.2020 г.

После доработки 16.12.2020 г.

Принята к публикации 16.12.2020 г.

Фактор инициации трансляции 2 эукариот (eIF2) и архей (aIF2) доставляет заряженную инициаторную тРНК (Met-tRNA^{iMet}) на малую рибосомную субчастицу. Фактор e/aIF2 состоит из трех субъединиц (α, β, γ) и функционирует как гетеротримерный комплекс. Ранее было показано, что мозиготная миссенс-мутация Pе222Thr в γ-субъединице человеческого IF2, нарушающая взаимодействие β- и γ-субъединиц, ведет к умственной отсталости и микроцефалии. В представленной работе показано, что аналогичная мутация в γ-субъединице фактора инициации трансляции 2 архей, который гомологичен эукариотическому белку, не изменяет ни конформацию этого белка, ни его сродство к aIF2β.

DOI: 10.31857/S0023476121050155

ВВЕДЕНИЕ

В эукариотах и археях фактор инициации трансляции 2 (e/aIF2) играет ключевую роль в инициации биосинтеза белка. В ГТФ-связанной форме он доставляет инициаторную метионил-тРНК на малую субчастицу рибосомы. Структурные перестройки, возникающие в 43S преинициаторном комплексе, состоящем из малой рибосомной субчастицы 40S, связанной с факторами инициации трансляции eIF1, eIF1A, eIF3, и тройным комплексом eIF2-Met-tRNA^{iMet}-GTP, способствуют быстрому гидролизу ГТФ даже в отсутствие мРНК [1]. После узнавания старт-кодона и удаления неорганического фосфата (Pi) eIF2 переходит в ГДФ-связанную форму и диссоциирует из инициаторного комплекса [2, 3], оставляя инициаторную тРНК в Р-сайте малой рибосомной субчастицы и открывая возможности для дальнейших этапов биосинтеза белка.

Фактор инициации трансляции 2 состоит из трех субъединиц (α, β, γ) и функционирует как гетеротримерный комплекс. Центральную роль в формировании e/aIF2 играет γ-субъединица. Она взаимодействует как с α-, так и с β-субъединицами, тогда как α- и β-субъединицы не взаимодействуют друг с другом. Присутствие α-субъединицы необходимо для связывания инициаторной метионил-тРНК [4, 5]; β-субъединица в эукариотическом факторе при образовании инициаторного комплекса 43S взаимодействует с мРНК [6].

Дефекты в связывании γ- и β-субъединиц, обусловленные миссенс-мутацией Pе222Thr в человеческом IF2γ (HsaIF2γ), приводят к X-хромосомному неврологическому заболеванию, характеризующемуся умственной отсталостью и микроцефалией [7]. В дрожжах аналогичная мутация ухудшает правильный выбор старт-кодона и функционирование SceIF2 *in vivo*, причем негативная роль указанной мутации может быть минимизирована суперэкспрессией гена белка eIF2β [7]. Структурное обоснование негативной роли миссенс-мутации в этих факторах в настоящее время невозможно, так как до сих пор не определена структура ни одного эукариотического фактора инициации трансляции 2 с атомным разрешением.

Ранее было показано, что *in vitro* aIF2 может функционально заменять eIF2 в связывании Met-tRNA^{iMet} с эукариотической рибосомой и в сканировании матрицы [8]. В настоящее время известны кристаллические структуры γ-субъединиц из *Pyrococcus abyssi* (PabIF2γ) [9], *Methanococcus jannaschii* (MjaIF2γ) [10], *Sulfolobus solfataricus* (SsoIF2γ) [11, 13] и *Pyrococcus furiosus* (PfuIF2γ) [12]. Каждая γ-субъединица состоит из трех доменов, N-концевой домен (G-домен) представляет собой ГТФазу и отвечает за основные функции белка, в том числе за узнавание β-субъединицы [12–14]. G-домены структур aIF2 из разных организмов (за исключением лабильных переключателей) могут быть наложены друг на друга с

Таблица 1. Статистические характеристики дифракционного набора и кристаллографического уточнения *Ile181Thr SsoIF2 γ*

Статистика набора	
Пр. гр.	<i>I</i> 23
$a = b = c$, Å; $\alpha \leq \beta \leq \gamma$, град	186.59; 90.0
Длина волны, Å	1.54
Пределы разрешения, Å	26.39–2.1 (2.2–2.1)
Общее число отражений	62755 (8124)
Число уникальных отражений	18403 (3357)
Полнота, %	99.9 (100.0)
R_{merge} , %	10.8 (34.28)
Избыточность	3.41 (2.42)
Среднее $I/\sigma(I)$	8.02 (1.99)
Статистика уточнения	
Диапазон разрешения, Å	26.38–2.10 (2.13–2.10)
Число молекул в асимметричной части	1
Число отражений	62736 (2594)
Размер тестовой выборки, %	5
R_{work} , %	18.2 (23.21)
R_{free} , %	20.6 (25.06)
Средний температурный фактор, Å ²	24.3
Среднеквадратичные отклонения	
Длины связей, Å	0.005
Валентные углы, град	1.102
Число остатков на карте Рамачандрана	
Наиболее предпочтительные районы, %	97.8
Дополнительно разрешенные районы, %	2.2
PDB ID	6R8T

Примечание. Данные в скобках соответствуют интервалу наиболее высокого разрешения.

$\text{g.m.s.d.} < 0.65 \text{ \AA}$ [14]. Это предполагает наличие только незначительных локальных изменений в сайте узнавания β -субъединицы. Со стороны последней во всех известных структурах в узнавании γ -субъединицы участвует α -спираль [12–14]. В *HsaIF2 β* участок 171–189 аминокислотной последовательности, соответствующий спирали $\alpha 1$ в структуре факторов из архей [15], скорее всего также имеет спиральную конформацию.

Тогда можем предположить, что миссенс-мутация будет иметь одинаковое влияние на целостность гетеротримера во всех *e/aIF2*. В настоящее время проверить это предположение возможно только для архей, так как только для них известны

пространственные структуры *IF2* или его субъединицы. Для исследования выбрали *SsoIF2 γ* с мутацией *Ile181Thr*, которая является аналогом мутации *Ile222Thr* в *HsaIF2 γ* . Полученная в работе структура мутантной формы *Ile181Thr SsoIF2 γ* оказалась идентична структуре белка *SsoIF2 γ* дикого типа (4rjl) [16]. Встраивание полученной структуры в структуру $\alpha\beta\gamma$ -гетеротримера [13] не выявило стерических или других препятствий для связывания β - и γ -субъединиц.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Получение *Ile181Thr SsoIF2 γ* . Суперпродукцию и очистку мутантной формы *Ile181Thr SsoIF2 γ* проводили так же, как и белка *SsoIF2 γ* дикого типа [17]. Непосредственно перед кристаллизацией белок *Ile181Thr SsoIF2 γ* смешивали с нуклеотидом GDPCP и проводили эксперименты по кристаллизации в условиях, описанных в [18] для мутантной формы *Sso aIF2 γ* ($\Delta 37$ –47) с GDPCP.

Определение структуры мутантного белка. Дифракционные данные собраны в Институте белка РАН (Пушино, Россия) на генераторе с вращающимся анодом (Bruker AXS MICROSTAR) с CCD-детектором (Bruker PLATINUM 135). Данные обрабатывали с помощью программного комплекса PROTEUM^{plus} (Bruker AXS). Стартовые фазы получены методом молекулярного замещения в программе Phaser [19], принадлежащей комплексу кристаллографических программ CCP4 [20]. В качестве стартовой модели использовали структуру белка *SsoIF2 γ* дикого типа высокого разрешения в комплексе с GDPCP (PDB ID 4rjl) [16]. Полученную структуру уточняли при разрешении 2.1 Å с использованием программного комплекса PHENIX [21]. Ручную правку и модификацию модели осуществляли с помощью программного комплекса Coot [22]. Статистика сбора данных и кристаллографического уточнения представлена в табл. 1. Координаты и структурные факторы помещены в банк белковых данных.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

В настоящее время структура $\beta\gamma$ -комплекса *IF2* известна для двух архей: *S. solfataricus* [13, 14] и *P. furiosus* [12], что позволяет надежно выделить область межмолекулярного взаимодействия $\beta\gamma$. Фрагменты аминокислотных последовательностей *a/eIF2 γ* и *a/eIF2 β* , ответственные за формирование межмолекулярного взаимодействия $\beta\gamma$, представлены на рис. 1. Список доступных структур субъединиц *aIF2* с литературными ссылками и PDB-кодами для каждой структуры приведены в табл. 2. Для исследования влияния миссенс-мутации в *aIF2 γ* на сродство β - и γ -субъединиц в архейном *IF2* выбрали структуру *SsoIF2 γ* прежде всего потому, что гетеротример *SsoIF2* определен



Рис. 1. Сравнение аминокислотных последовательностей фрагментов α IF2 и ϵ IF2 (табл. 2), ответственных за формирование области межмолекулярного взаимодействия $\beta\gamma$. Остатки, взаимодействующие в межмолекулярной области β - и γ -субъединиц SsoIF2, помечены звездочками; а – сравнение фрагмента последовательностей IF2 γ архейных и эукариотических белков; б – сравнение аминокислотных последовательностей участка β -субъединицы архейных и эукариотических белков, ответственного за связывание γ -субъединицы; на сером фоне показаны идентичные остатки, на черном – остатки, кардинально различающиеся в архейных и эукариотических последовательностях.

с более высоким разрешением (2.15 Å), чем гетеродимер PfuIF2 (2.8 Å). Кроме того, в кристалле гетеродимера PfuIF2 N-конец β -субъединицы имеет плотный контакт с доменом II γ -субъединицы, что, по-видимому, искажает его спиральную конформацию, которая сохраняется в SsoIF2 γ .

Кристаллическая структура мутантной формы Pe181Thr SsoIF2 γ в комплексе с GDPCP (аналогом ГТФ) определена при разрешении 2.1 Å. Наложение этой структуры на структуру аналогичного комплекса белка дикого типа [16] показывает их полную идентичность (средняя квадратичная ошибка равна 0.184 Å для всех C α -атомов). Фрагмент структуры, ответственный за контакт с β -субъединицей, показан на рис. 2. В месте точечной мутации боковая цепь треонина повторяет

валиновую часть Pe181, положения атомов OG1 и CG1 совпадают в пределах ошибки измерения. В мутантной форме атом OG1 связан с двумя молекулами воды (S107 и S177), в структуре дикого типа вода вблизи CG1 отсутствует. Таким образом, единственное следствие миссенс-мутации в SsoIF2 γ – появление полярного остатка на поверхности при полном сохранении конформации белка.

Остаток в положении 181 включен в область межмолекулярного взаимодействия $\beta\gamma$ и может оказывать влияние на сродство комплекса. Замена γ -субъединицы дикого типа в структурах SsoIF2 $\alpha\beta\gamma$ и PfuIF2 $\beta\gamma$ на мутантную форму Pe181Thr SsoIF2 γ показывает, что место мутации в области межмолекулярного взаимодействия $\beta\gamma$ как SsoIF2, так и PfuIF2 остается доступным растворителю. Обе молекулы воды, входящие в бли-

Таблица 2. Опубликованные структуры субъединиц α IF2

Организм	Обозначение	Объект	PDB-код	Разрешение, Å	Литература
<i>S. solfataricus</i>	Sso	SsoIF2 γ	4rjl	1.64	[16]
		SsoIF2 $\alpha\beta\gamma$	3cw2	2.80	[14]
		SsoIF2 $\alpha\beta\gamma_{incomp}$	2qn6	2.15	[13]
<i>P. furiosus</i>	Pfu	PfuIF2($\beta\gamma$)	2d74	2.80	[12]
<i>P. abyssi</i>	Pab	PabIF2 γ	1kk3	1.90	[9]
<i>M. jannaschii</i>	Mja	MjaIF2 γ	1s0u	2.40	[10]

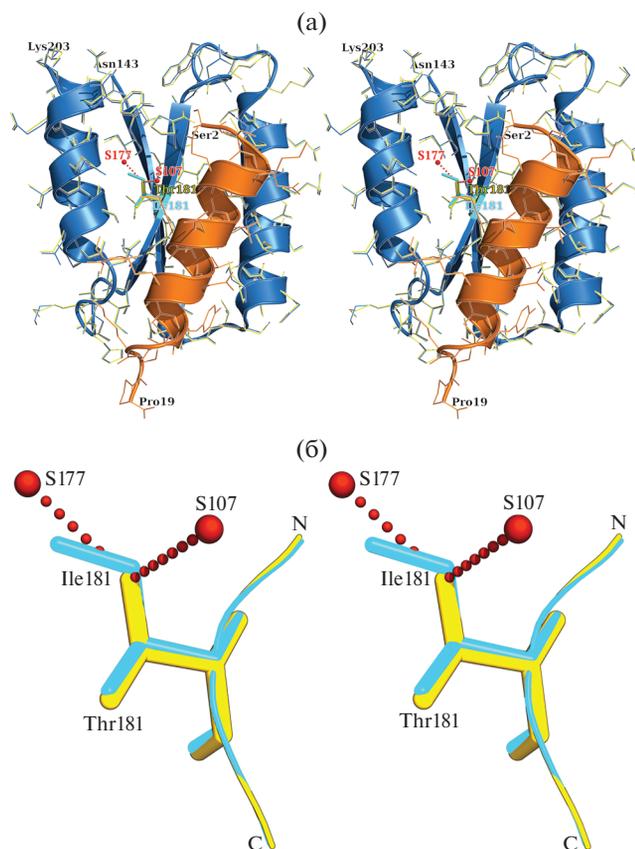


Рис. 2. Наложение участков (Asn143-Lys203) структур SsoIF2 γ дикого типа и мутантной формы Ile181Thr SsoIF2 γ , определяющих связывание β -субъединицы: а – вторичная структура и аминокислотные остатки SsoIF2 γ дикого типа показаны черным, аминокислотные остатки мутантной формы – белым. N-концевая α -спираль (Ser2-Pro19) SsoIF2 β , участвующая в формировании $\beta\gamma$ -комплекса, показана серым; б – положения атомов CG1 в изолейцине белка дикого типа и OG1 в треонине мутантной формы совпадают в пределах ошибки измерения (0.26 Å). В мутантной форме остаток Thr181 связан с двумя молекулами воды.

жайшее окружение Thr181 в свободной мутантной форме SsoIF2 γ , сохраняют свое положение при образовании $\beta\gamma$ -гетеродимера, что позволяет этому остатку образовать все возможные для него водородные связи. Таким образом, анализ структурных данных показывает, что миссенс-мутация в SsoIF2 не создает стерических или каких-либо других ограничений, мешающих связыванию субъединиц. Это подтверждают проведенные биохимические эксперименты по связыванию мутантной формы Ile181Thr SsoIF2 γ с SsoIF2 β дикого типа (рис. 3). Обе субъединицы по-прежнему связываются друг с другом со средством в соотношении 1 : 1.

Сравнение аминокислотных последовательностей архейных и эукариотических факторов инициации трансляции 2 (рис. 1) показывает, что

Оптическая плотность, отн. ед.

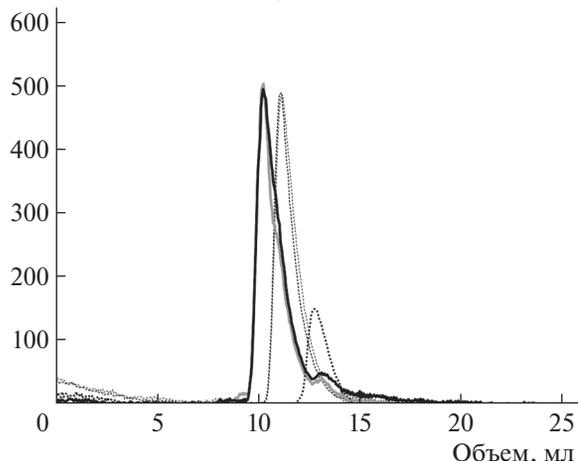


Рис. 3. Профиль элюции комплекса SsoIF2 $\beta\gamma$ и его компонентов, полученный с использованием гель-хроматографии и колонки Superdex 75 10/30. Тонкая черная пунктирная линия – SsoIF2 γ дикого типа, тонкая серая пунктирная линия – Ile181Thr SsoIF2 γ , черная пунктирная линия – SsoIF2 β , черная сплошная линия – SsoIF2 $\beta\gamma$, серая сплошная линия – Ile181Thr SsoIF2 $\beta\gamma$.

некоторые остатки области межмолекулярного взаимодействия $\beta\gamma$, сохраняющие идентичность в археях, кардинально меняются в случае эукариот. Это остатки в позициях 167, 177, 187, 189, 194 и 197 (номенклатура SsoIF2 γ). Возможно, именно замены в каких-то из этих положений при наличии миссенс-мутации влияют на средство β - и γ -субъединиц в эукариотических IF2. Исследование влияния замен в этих положениях на средство Ile181Thr SsoIF2 γ и SsoIF2 β SsoIF2 является целью дальнейшей работы.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант № 18-04-01331-а).

СПИСОК ЛИТЕРАТУРЫ

1. *Algire M.A., Maag D., Lorsch J.R.* // *Mol. Cell.* 2005. V. 20. P. 251. <https://doi.org/10.1016/j.molcel.2005.09.008>
2. *Kap L.D., Lorsch J.R.* // *Annu. Rev. Biochem.* 2004. V. 73. P. 657. <https://doi.org/10.1146/annurev.biochem.73.030403.080419>
3. *Jackson R.J., Hellen C.U., Pestova T.V.* // *Nat. Rev. Mol. Cell. Biol.* 2010. V. 11. P. 113. <https://doi.org/10.1038/nrm2838>
4. *Pedulla N., Palermo R., Hasenohrl D. et al.* // *Nucl. Acids Res.* 2005. V. 33. P. 1804. <https://doi.org/10.1093/nar/gki321>
5. *Yatime L., Schmitt E., Blanquet S., Mechulam Y.* // *J. Biol. Chem.* 2004. V. 279. P. 15984. <https://doi.org/10.1074/jbc.M311561200>

6. *Laurino J.P., Thompson G.M., Pacheco E., Castilho B.A.* // Mol. Cell. Biol. 1999. V. 19. P. 173.
<https://doi.org/10.1128/MCB.19.1.173>
7. *Borck G., Shin B-S., Stiller B. et al.* // Mol. Cell. 2012. V. 48. P. 641.
<https://doi.org/10.1016/j.molcel.2012.09.005>
8. *Dmitriev S.E., Stolboushkina E.A., Terenin I.M. et al.* // J. Mol. Biol. 2011. V. 413. P. 106.
<https://doi.org/10.1016/j.jmb.2011.08.026>
9. *Schmitt E., Blanquet S., Mechulam Y.* // EMBO J. 2002. V. 21. P. 1821.
<https://doi.org/10.1093/emboj/21.7.1821>
10. *Roll-Mecak A., Alone P., Cao C. et al.* // J. Biol. Chem. 2004. V. 279. P. 10634.
<https://doi.org/10.1074/jbc.M310418200>
11. *Yatime L., Mechulam Y., Blanquet S., Schmitt E.* // Structure. 2006. V. 14. P. 119.
<https://doi.org/10.1016/j.str.2005.09.020>
12. *Sokabe M., Yao M., Sakai N., Toya S., Tanaka I.* // Proc. Natl. Acad. Sci. USA. 2006. V. 103. P. 13016.
<https://doi.org/10.1073/pnas.0604165103>
13. *Yatime L., Mechulam Y., Blanquet S., Schmit E.* // Proc. Natl. Acad. Sci. USA. 2007. V. 104. P. 18445.
<https://doi.org/10.1073/pnas.0706784104>
14. *Stolboushkina E., Nikonov S., Nikulin A. et al.* // J. Mol. Biol. 2008. V. 382. P. 680.
<https://doi.org/10.1016/j.jmb.2008.07.039>
15. *Thompson G.M., Pacheco E., Melo E.O., Castilho B.A.* // Biochem. J. 2000. V. 347. P. 703.
<https://doi.org/10.1042/bj3470703>
16. *Nikonov O., Kravchenko O., Arkhipova V. et al.* // Biochimie. 2016. V. 121. P. 197.
<https://doi.org/10.1016/j.biochi.2015.11.029>
17. *Nikonov O., Stolboushkina E., Nikulin A. et al.* // J. Mol. Biol. 2007. V. 373. P. 328.
<https://doi.org/10.1016/j.jmb.2007.07.048>
18. *Nikonov O., Stolboushkina E., Arkhipova V. et al.* // Acta Cryst. D. 2014. V. 70. P. 658.
<https://doi.org/10.1107/S1399004713032240>
19. *Storoni L.C., McCoy A.J., Read R.J.* // Acta Cryst. D. 2004. V. 60. P. 432.
<https://doi.org/10.1107/S0907444903028956>
20. Collaborative Computational Project. Number 4 // Acta Cryst. D. 1994. V. 50. P. 760.
<https://doi.org/10.1107/S0907444994003112>
21. *Adams P.D., Grosse-Kunstleve R.W., Hung L.W. et al.* // Acta Cryst. D. 2002. V. 58. P. 1948.
<https://doi.org/10.1107/S0907444902016657>
22. *Emsley P., Lohkamp B., Scott W., Cowtan K.* // Acta Cryst. D. 2010. V. 66. P. 486.
<https://doi.org/10.1107/S0907444910007493>

СТРУКТУРА МАКРОМОЛЕКУЛЯРНЫХ
СОЕДИНЕНИЙ

УДК 548.73

ПРОСТРАНСТВЕННАЯ СТРУКТУРА ТЕРМОСТАБИЛЬНОЙ
ТРАНСАМИНАЗЫ D-АМИНОКИСЛОТ ИЗ АРХЕИ *Methanocaldococcus*
jannaschii DSM 2661

© 2021 г. К. М. Бойко^{1,*}, А. Ю. Николаева², А. К. Бакунова¹, Т. Н. Стеханова¹, Т. В. Ракитина²,
В. О. Попов^{1,2}, Е. Ю. Безсуднова¹

¹ Институт биохимии им. А.Н. Баха РАН, Федеральный исследовательский центр “Фундаментальные основы биотехнологии” РАН, Москва, Россия

² Национальный исследовательский центр “Курчатовский институт”, Москва, Россия

*E-mail: boiko_konstantin@inbi.ras.ru

Поступила в редакцию 14.07.2020 г.

После доработки 14.08.2020 г.

Принята к публикации 14.08.2020 г.

Пиридоксаль-5'-фосфат-зависимые трансминазы (PLP) катализируют стереоспецифический перенос аминокислоты с аминокислоты или амина на кетон или кетокислоту. Трансминазы участвуют в метаболизме аминокислот у всех организмов; на основе этого суперсемейства ферментов активно разрабатываются биокатализаторы стереоселективного аминирования органических соединений для тонкого органического синтеза. Представлены краткая биохимическая характеристика и кристаллическая структура термостабильной трансминазы I типа укладки PLP-связывающего домена из термофильной археи *Methanocaldococcus jannaschii* DSM 2661, полученная с разрешением 1.8 Å. Описаны строение функционального димера фермента и организация его активного центра, проведено сравнение с ближайшими гомологами.

DOI: 10.31857/S0023476121050039

ВВЕДЕНИЕ

Пиридоксаль-5'-фосфат-зависимые трансминазы (PLP) (ТА, аминотрансферазы, ЕС 2.6.1.X) катализируют обратимый стереоселективный перенос аминокислоты с аминотрансферата (амин/аминокислота) на кетосубстрат (кетон/кетокислота/альдегид) с образованием хирального амина/аминокислоты и нового кетосоединения [1]. В клетках всех организмов ТА являются ключевыми ферментами метаболизма аминокислот. В биотехнологии ТА применяются для стереоселективного аминирования органических соединений, на основе природных ТА разрабатываются биокатализаторы для синтеза оптически активных аминов и неприродных аминокислот [2–6].

Принятая на сегодня классификация ТА составлена по множественному выравниванию последовательностей в базе данных семейств белковых доменов PFAM и объединяет, ориентируясь на субстратную специфичность, ТА в классы и подклассы (семейства) по паре реагирующих субстратов: донору аминокислоты/акцептору аминокислоты (механизм трансминирования приведен ниже). Имеются шесть основных классов ТА [1]. Кроме того, все ТА разделяются по типу укладки

PLP-связывающего домена на ферменты I или IV типа укладки [7, 8].

Любой фермент характеризуется оптимальным природным субстратом (или набором субстратов). Для всех ТА такими субстратами являются α-аминокислоты и их кетоаналоги, однако некоторые ферменты в определенных условиях активны с отличными от природных субстратов соединениями, т.е. проявляют разнородную (промишквитетную) активность (*promiscuous activity*). Такая активность представляет интерес для целей биотехнологии, так как позволяет использовать ферменты в синтезе неприродных соединений [9].

Трансминазы III класса относятся к I типу PLP-укладки и являются примером ферментов с разнородной активностью, т.е. проявляют активность не только с α-, β-, γ- и ε-аминокислотами и их кетоаналогами, но и с (S)-ароматическими и алифатическими аминами, а также с альдегидами и кетонами [10, 11]. Трансминазы III класса, катализирующие аминирование (S)-8-амино-7-оксононаноата (КАРА) с образованием 7,8-диаминопеларгоновой кислоты (ДАРА), составляют подсемейство ДАРА ТА и в клетках участвуют в синтезе биотина [12]. Большинство ДАРА ТА узкоспецифичны, используют S-аденозил-L-мети-

онин (SAM) в качестве аминодонора. Среди ТА III класса – это одни из самых “медленных” ферментов [13, 14]. В [15] была охарактеризована DAPA TA из *Psychrobacter cryohalolentis*, проявляющая активность не только к природным SAM, DAPA и KAPA, но и к (S)-(-)-1-фенилэтиламину (S-PEA) и ряду альдегидов и кетонов. Для проявления такой разнородной субстратной специфичности в активном центре фермента присутствует ряд аминокислотных замен, а также повышенная мобильность отдельных остатков по сравнению с аминокислотными остатками в активном центре канонической DAPA TA из *Escherichia coli*.

Исследованная в настоящей работе термостабильная DAPA TA из археи *Methanocaldococcus jannaschii* DSM 2661 (BioA, EC 2.6.1.62) гомологична описанной ранее DAPA TA из *P. cryohalolentis*. Охарактеризована неприродная активность BioA в реакции с S-PEA и его энантиомером R-PEA с кетосубстратом изобутаналем, получена пространственная структура фермента с разрешением 1.8 Å и проведены ее детальный анализ и сравнение со структурами гомологичных ТА. Полученные данные дополняют многообразие структур ТА I типа укладки, а также могут быть полезны при разработке новых термостабильных биокатализаторов, проявляющих максимальную активность и стабильность при температурах 50°C и выше, а также стабильных в присутствии органических растворителей, улучшающих растворимость органических субстратов в биотехнологических процессах.

МАТЕРИАЛЫ И МЕТОДЫ

Клонирования, экспрессия и очистка BioA. Ген BioA был получен полимеразной цепной реакцией (ПЦР) с использованием геномной ДНК *M. jannaschii* DSM 2661 [16] в качестве матрицы. Продукт ПЦР лигировали в плазмиду pET-21b (Novagen, Germany) и модифицировали согласно методике, описанной в [17], для получения целевого белка, на N-конце слитого с His-тагом и сайтом узнавания TEV протеазы. Трансформированные клетки *E. coli* BL21Star (Stratagene, USA) выращивали в среде LB/ампицилин при 37°C до $A_{600} = 0.8$, экспрессию индуцировали 1 mM изопропил-β-тиогалактопиранозидом. После инкубирования при 25°C клетки собирали центрифугированием, далее ресуспендировали в 50 mM Трис-НСl-буфере, pH 7.5, содержащем 200 mM NaCl, 20 mM имидазола, 10% глицерина, 2 mM β-меркаптоэтанола, 0.1% Тритона X-100 и 1 mM фенолметилсульфонилфторида, и разрушали ультразвуком. Грубый экстракт центрифугировали при 10000 g в течение 25 мин. Супернатант наносили на колонку HisTrap HP (GE Healthcare, USA), уравновешенную 50 mM Трис-НСl-буфером, pH 7.5, содержащим 500 mM NaCl, 20 mM имидазола и 0.1% Тритона X-100. BioA с 6-His-тагом

элюировали линейным градиентом имидазола (20–500 mM) в том же буфере без Тритона X-100. Фракцию с целевым ферментом инкубировали с десятикратным молярным избытком PLP при 4°C в течение 12 ч, переносили в буфер 50 mM Трис-НСl, pH 7.9, содержащий 100 mM NaCl, 1 mM ЭДТА, 1 mM β-меркаптоэтанола и TEV протеазу (1 мг на 10 мг белка). Раствор инкубировали при комнатной температуре в течение 2 ч, диализировали против 50 mM Трис-НСl-буфера, pH 7.5, содержащего 500 mM NaCl, 20 mM имидазола, и наносили на колонку HisTrap HP. Проскок собирали, концентрировали и наносили на колонку Superdex 200 10/300 GL (GE Healthcare), уравновешенную 50 mM Трис-НСl-буфером, pH 7.5, содержащим 50 mM NaCl и 15 мкМ PLP. Очищенный BioA в холоформе хранили при –20°C с добавлением 50% глицерина. Чистоту белка подтверждали при помощи ДДС-ПААГ электрофореза. Концентрацию белка определяли методом Бредфорда.

Определение активности BioA и стационарных параметров реакции. Удельную активность BioA в реакции с R-PEA и S-PEA с изобутаналем в качестве аминокцептора рассчитывали из начального линейного участка зависимости накопления продукта от времени. За единицу ферментативной активности U принимали количество BioA, катализирующее образование 1 мкМ ацетофенона в минуту. За образованием ацетофенона следили по увеличению оптической плотности при 245 нм ($\Delta\epsilon = 12 \text{ mM}^{-1} \text{ cm}^{-1}$) на спектрофотометре Evolution 300 UV-Vis (Thermo Scientific, USA). Стандартную реакцию проводили при 65°C в 100 mM бикарбонатном буфере, pH 10, с добавлением 90 мкМ PLP при концентрации каждого субстрата 10 mM.

Для получения кинетических параметров варьировали концентрацию одного субстрата при постоянной концентрации второго (концентрация для изобутанала составляла 12 mM, для R-PEA – 10 mM). Измерения проводили при 35°C в 75 mM натрий-фосфатном буфере, pH 8, с добавлением 90 мкМ PLP. Кинетические параметры получали, обрабатывая экспериментальные данные согласно уравнению Михаэлиса–Ментен, используя Origin 7.5 (Origin Lab, USA).

Кристаллизацию BioA проводили методом диффузии в парах, как описано в [18]. Финальные условия кристаллизации включали 0.2 M MgCl₂; 0.1 M Tris, pH 8.5; 19% PEG 3350, кристаллизация велась при температуре +20°C. Кристаллы палочкообразной формы размером ~100 × 100 × 50 мкм росли в течение семи дней.

Сбор и обработка дифракционных данных. Решение и уточнение структуры. Перед сбором дифракционных данных кристаллы вылавливали петлей и переносили в криораствор, содержащий

кроме компонентов, входящих в противораствор, 25% глицерина. После чего кристалл в петле замораживали в парах азота. Дифракционные наборы собирали при температуре 100 К на станции ID29 синхротронного источника ESRF (Франция). Дифракционная картина фиксировалась детектором Pilatus6M. Для расчета стратегии сбора данных использовали программу Best [19], предложившую следующие параметры: длина волны 0.969 Å, угол вращения 287°, угол качания 0.1°, расстояние кристалл–детектор 302 мм. Набор дифракционных данных обработан с использованием программы iMosflm [20]. Статистика набора данных приведена в табл. 1.

Решение структуры проведено методом молекулярного замещения при помощи программы MOLREP [21]. В качестве стартовой модели использовали структуру холо-формы синтазы диаминопералгоновой кислоты из *Bacillus subtilis* (код PDB 3DOD). Кристаллографическое уточнение структуры проведено с использованием программ Refmac5 [22] и Coot [23] до достижения R -факторами следующих значений: $R_{\text{work}} = 17.7\%$, $R_{\text{free}} = 22.2\%$ (табл. 2).

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Активность BioA и кинетические параметры полной реакции. Удельная активность выделенных препаратов BioA с R-PEA и S-PEA соответственно составила 42 ± 2 и 21 ± 1 мУ/мг при 65°C в 100 мМ бикарбонатном буфере, pH 10. Кинетические параметры катализируемой BioA реакции между R-PEA и изобутаналем приведены в табл. 3. Можно сделать вывод, что BioA стабилен и активен в щелочных условиях при 65°C, однако проявляет низкую стереоспецифичность к неканоническим субстратам.

Анализ структуры BioA. Кристаллы BioA принадлежат пр. гр. $P1$. В независимой части элементарной ячейки кристалла находятся две субъединицы белка, образующие функциональный димер [1, 8]. В обеих субъединицах в электронной плотности отсутствуют восемь аминокислот с N-конца, а также участок 190–209, что, по-видимому, связано с их относительной подвижностью. Парное сравнение субъединиц между собой дает RMSD по C α -атомам, не превышающее 0.14 Å², свидетельствующее об отсутствии между ними значимых различий.

Общая архитектура субъединицы BioA типична для TA I типа укладки и состоит из двух доменов – малого N-концевого (остатки 1–58 и 359–466) и большого C-концевого (остатки 59–358) (рис. 1а) [7]. N-концевой домен BioA содержит пять α -спиралей и два β -слоя по три тяжа каждый, а C-концевой – 9 α -спиралей, окружающих β -слой, состоящий из семи тяжей.

Таблица 1. Кристаллографические данные и параметры съемки кристалла трансаминазы BioA

Пр. гр.	$P1$
$a, b, c, \text{Å}$	61.23, 61.38, 63.07;
$\alpha, \beta, \gamma, \text{град}$	106.11, 93.48, 96.47
$T, \text{К}$	100
$\lambda, \text{Å}$	0.969
Разрешение, Å	60.56–1.80 (1.84–1.80)*
Число независимых рефлексов	71599 (4097)
Повторяемость	2.7 (2.6)
Полнота набора, %	88.5 (86.1)
$I/\sigma(I)$	6.4 (2.4)
$R_{\text{meas}}, \%$	12.6 (53.3)
$CC_{1/2}, \%$	98.9 (75.5)

* В скобках приведены значения для последнего слоя.

Таблица 2. Данные уточнения структуры трансаминазы BioA

$R_{\text{fact}}, \%$	17.7
$R_{\text{free}}, \%$	22.2
Общий средний B -фактор	22.0
Средний B -фактор по белку	21.8
Средний B -фактор по лигандам	23.3
Средний B -фактор по растворителю	27.2
Число неводородных атомов	
Белок	6958
Лиганды	3
Растворитель	446
Всего	7407
Среднеквадратичные отклонения	
Длины связей, Å	0.013
Валентные углы, град	1.842
График Рамачандрана	
Наиболее благоприятные, %	94.5
Допустимые, %	4.4
Код PDB	6ZHK

Таблица 3. Кинетические параметры катализируемой BioA реакции между R-PEA и изобутаналем при 65°C

Субстрат	$k_{\text{cat}}, \text{с}^{-1}$	$K_m, \text{мМ}$	$k_{\text{cat}}/K_m, \text{с}^{-1} \text{мМ}^{-1}$
R-PEA	0.033 ± 0.001	0.9 ± 0.1	0.037 ± 0.005
Изобутаналь	0.0162 ± 0.0005	0.19 ± 0.03	0.08 ± 0.02

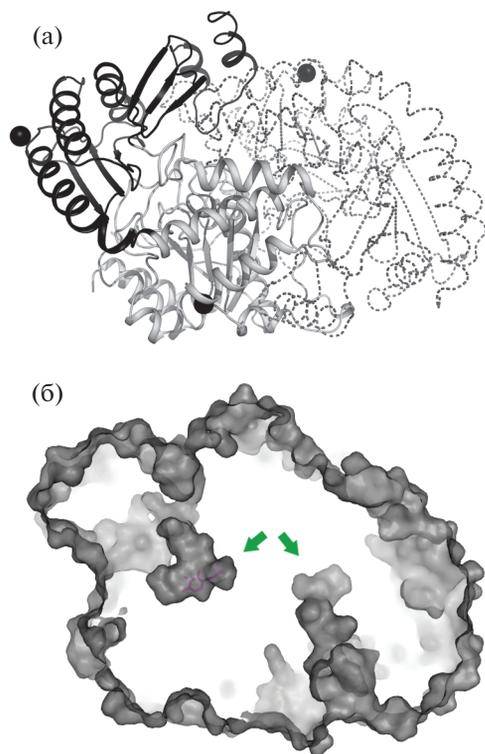


Рис. 1. Структура BioA. Димер BioA (а). N-концевой домен одной из субъединиц димера показан в виде черной ленточной модели, С-концевой – в виде серой ленточной модели. Вторая субъединица димера показана в виде прерывистой ленточной модели. Ионы магния показаны сферами. Субстратный канал BioA (б). Серым показана поверхность димера BioA (центральный срез) и каналы, ведущие от поверхности молекулы в активный центр фермента. Молекула РМР взята из структуры 6I01 и показана палочковой моделью для одной субъединицы. Активный центр каждой субъединицы обозначен стрелкой.

Анализ контактов между субъединицами из независимой части показал, что BioA аналогично многим другим ТА [1, 8, 15] является димером в кристалле. Скрытая при формировании димера поверхность (buried surface area) составляет 5678 Å² (28.5% поверхности субъединицы). Димер стабилизирован девятью солевыми мостиками и 50 водородными связями. Оценка гидрофобных взаимодействий, сделанная по выигрышу

свободной энергии при формировании димерного межмолекулярного контакта, дает величину –96.3 ккал/моль (табл. 4).

Димер BioA является минимальной каталитической единицей фермента, что связано с устройством активного центра ТА, включающего в себя остатки обеих субъединиц, а именно – двух доменов одной субъединицы и С-концевого домена соседней. Активный центр BioA находится глубоко в центре каждой из субъединиц димера и соединен с поверхностью протяженным каналом длиной ~15 Å (рис. 1б). Полость активного центра сформирована остатками W63, D121, G122, A123, E124, V126, Y155, H156, G157, E234, D268, V270, A271, G336, H337 и T338, а также каталитическим лизином K301. Для данного класса ферментов известно, что помимо каталитических остатков в активном центре находится молекула кофактора – РЛР, ковалентно связанная в структуре холоформы с каталитическим остатком лизина и отсоединяющаяся от него в процессе катализа с образованием свободной РМР-формы, способной покинуть активный центр. Анализ активного центра BioA показал, что в обеих субъединицах фермента отсутствует разностная электронная плотность для кофактора, поэтому каталитический остаток лизина K301 находится в свободной форме. Возможно, в случае BioA отсутствие кофактора вызвано условиями кристаллизации.

Сравнение субъединицы BioA с гомологичными структурами показало, что наибольшее сходство фермент имеет со структурами синтазы диаминопералгоновой кислоты из *Bacillus subtilis* (RMSD – 1.08 Å², код PDB – 3DRD), термостабильной ω-трансминазы из *Thermomicrobium roseum* (RMSD – 1.50 Å², код PDB – 6I01) и ТА из *Pseudomonas fluorescens* (RMSD – 1.57 Å², код PDB – 6S54) (рис. 2а). Несмотря на сходство структур субъединиц, указанные ферменты имеют различное олигомерное состояние в кристалле – димер (3DRD и 6I01) и тетрамер (6S54), что, впрочем, типично для ТА [1]. Тем не менее все указанные олигомерные состояния включают в себя аналогичный BioA димер (рис. 2б), что позволяет провести их взаимное сравнение. Анализ контактов упомянутых димеров (табл. 4) показывает, что ди-

Таблица 4. Сравнение димерных межмолекулярных контактов BioA и гомологичных ферментов

	BioA	3DRD [24]	6I01 [25]	6S54* [26]
Поверхность, скрытая при формировании димера, Å ² (процент от поверхности субъединицы)	5678 (28.5%)	4289 (22.1%)	5698 (28.3%)	5440 (27.6%)
Водородные связи	50	35	38	64
Солевые мостики	9	4	17	8
Гидрофобные взаимодействия (ΔG ^{int} , ккал/моль)	–96.3	–63.6	–81.1	–88.6

* Димерный контакт, аналогичный BioA.

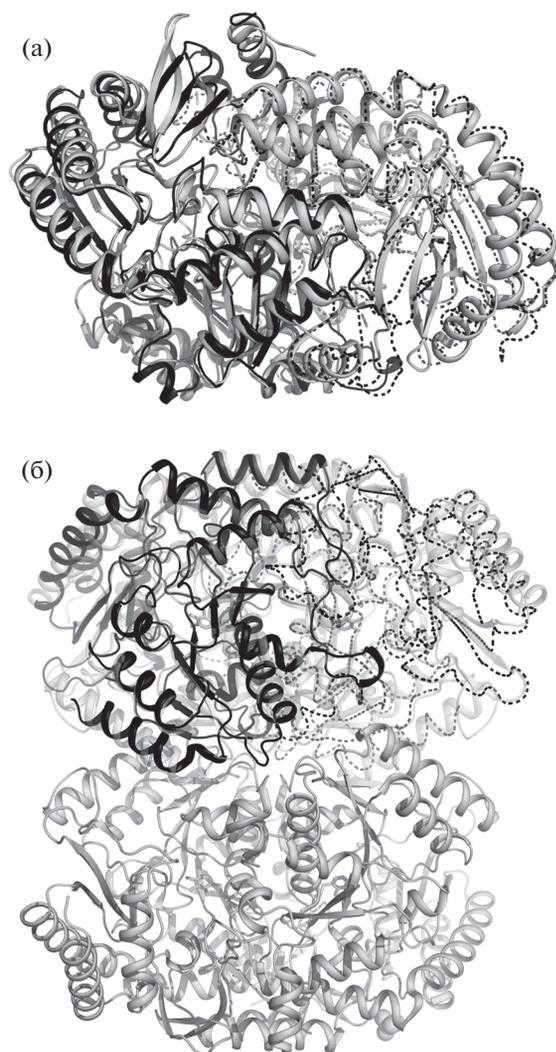


Рис. 2. Сравнение структур BioA и ближайших гомологов. Наложение димеров BioA (субъединицы показаны в виде сплошной и прерывистой ленточных моделей соответственно) и 3DRD (светло-серый) (а). Наложение димера BioA (окраска аналогична панели А) и тетрамера 6S54 (светло-серый) (б). Димер 6S54, аналогичный димеру BioA, показан полупрозрачно для ясности.

мерное межмолекулярное взаимодействие данного класса ферментов весьма прочно даже в случае белков из мезофильных организмов (например, 6S54). При этом в случае тетрамера 6S54 межмолекулярный контакт, аналогичный димеру BioA, значительно превосходит по прочности остальные, что подчеркивает ключевую роль димеризации для функционирования фермента.

Сравнение активных центров BioA и гомологичных ферментов позволило установить, что связанный кофактор присутствует не во всех сравниваемых структурах. Так, помимо BioA он отсутствует в структурах 3DRD и в одной из субъединиц 6S54, при этом во второй субъединице об-

наружена РМР-форма кофактора. В структуре 6S54 ковалентно связанный PLP находится во всех четырех субъединицах из независимой части. Анализ окружения области связывания кофактора показал, что в случае BioA остатки, формирующие окружение PLP, в основном консервативны, при этом в структуре 3DRD отсутствует петля 143–173 (нумерация по 3DRD), закрывающая активный центр от растворителя. В отсутствие кофактора в BioA вакантное положение его фосфатной группы занято боковой группой остатка Н337, имеющей двойное положение, чего не наблюдается в структуре гомологов, где крупная боковая группа гомологичного ароматического остатка отвернута от кофактора. Неконсервативный S240 в BioA (аланин у гомологов) создает полярную область на входе в полость активного центра. Еще одним отличием BioA является наличие вблизи активного центра пары неконсервативных полярных остатков – D158 и Y25, боковые группы которых образуют водородную связь между собой, а также с каталитическим Y155. В положении, аналогичном Y25, у гомологов находится гидрофобный остаток. Таким образом, в случае BioA пара D158–Y25 может дополнительно стабилизировать структуру активного центра, являясь следствием адаптации к повышенной температуре.

Таким образом, исследована структура термостабильной ТА из археи *Methanocaldococcus jannaschii* DSM 2661, проведено ее сравнение со структурами гомологичных ферментов и выявлены различия в активном центре, которые могут отражать субстратную специфичность BioA.

Работы по экспрессии и выделению белка выполнены при финансовой поддержке Российского научного фонда (грант № 19-14-00164), кристаллизация и рентгеноструктурный эксперимент выполнены при поддержке Российского Федерального космического агентства (проект КЭ (ЦР) “Кристаллизатор”), работы по кристаллографическому уточнению и анализу полученной структуры выполнены при поддержке Министерства науки и высшего образования РФ.

СПИСОК ЛИТЕРАТУРЫ

1. *Bezsudnova E.Y., Boyko K.M., Popov V.O.* // Biochemistry (Mosc.) 2017. V. 82. P. 1572. <https://doi.org/10.1134/S0006297917130028>
2. *Fuchs M., Farnberger J.E., Kroutil W.* // Eur. J. Org. Chem. 2015. V. 2015. P. 6965. <https://doi.org/10.1002/ejoc.201500852>
3. *Guo F., Berglund P.* // Green Chem. 2017. V. 19. P. 333. <https://doi.org/10.1039/C6GC02328B>
4. *Slabu I., Galman J.L., Turner N.J. et al.* // Acs. Catalysis. 2017. V. 7. P. 8263. <https://doi.org/10.1021/acscatal.7b02686>

5. *Steffen-Munsberg F., Vickers C., Kohls H. et al.* // *Bio-technol. Adv.* 2015. V. 33. P. 566.
<https://doi.org/10.1016/j.biotechadv.2014.12.012>
6. *Wu S., Snajdrova R., Moore J.C. et al.* // *Angew. Chem. Int. Ed. Engl.* 2020. V. 60. № 1. P. 88.
<https://doi.org/10.1002/anie.202006648>
7. *Grishin N.V., Phillips M.A., Goldsmith E.J.* // *Protein Sci.* 1995. V. 4. P. 1291.
<https://doi.org/10.1002/pro.5560040705>
8. *Bezsudnova E.Y., Popov V.O., Boyko K.M.* // *Appl. Microbiol. Biotechnol.* 2020. V. 104. P. 2343.
<https://doi.org/10.1007/s00253-020-10369-6>
9. *Hult K., Berglund P.* // *Trends Biotechnol.* 2007. V. 25. P. 231.
<https://doi.org/10.1016/j.tibtech.2007.03.002>
10. *Kaulmann U., Smithies K., Smith M.E.B. et al.* // *Enzyme Microbial Technology.* 2007. V. 41. P. 628.
<https://doi.org/10.1016/j.enzmictec.2007.05.011>
11. *Steffen-Munsberg F., Matzel P., Sowa M.A. et al.* // *Appl. Microbiol. Biotechnol.* 2016. V. 100. P. 4511.
<https://doi.org/10.1007/s00253-015-7275-9>
12. *Stoner G.L., Eisenberg M.A.* // *J. Biol. Chem.* 1975. V. 250. P. 4037.
[https://doi.org/10.1016/S0021-9258\(19\)41381-1](https://doi.org/10.1016/S0021-9258(19)41381-1)
13. *Cobessi D., Dumas R., Pautre V. et al.* // *Plant Cell.* 2012. V. 24. P. 1608.
<https://doi.org/10.1105/tpc.112.097675>
14. *Mann S., Ploux O.* // *Febs J.* 2006. V. 273. P. 4778.
<https://doi.org/10.1111/j.1742-4658.2006.05479.x>
15. *Bezsudnova E.Y., Stekhanova T.N., Popinako A.V. et al.* // *Appl. Microbiol. Biotechnol.* 2018. V. 102. P. 9621.
<https://doi.org/10.1007/s00253-018-9310-0>
16. *Bult C.J., White O., Olsen G.J. et al.* // *Science.* 1996. V. 273. P. 1058.
<https://doi.org/10.1126/science.273.5278.1058>
17. *Boyko K., Gorbacheva M., Rakitina T. et al.* // *Acta Cryst. F.* 2015. V. 71. P. 24.
<https://doi.org/10.1107/S2053230X14025333>
18. *Boyko K.M., Lipkin A.V., Popov V.O. et al.* // *Crystallography Reports.* 2013. V. 58. P. 442.
<https://doi.org/10.1134/S106377451105004X>
19. *Bourenkov G.P., Popov A.N.* // *Acta Cryst. D.* 2006. V. 62. P. 58.
<https://doi.org/10.1107/S0907444905033998>
20. *Battye T.G., Kontogiannis L., Johnson O. et al.* // *Acta Cryst. D.* 2011. V. 67. P. 271.
<https://doi.org/10.1107/S0907444910048675>
21. *Vagin A.A., Isupov M.N.* // *Acta Cryst. D.* 2001. V. 57. P. 1451.
<https://doi.org/10.1107/s0907444901012409>
22. *Vagin A.A., Steiner R.A., Lebedev A.A. et al.* // *Acta Cryst. D.* 2004. V. 60. P. 2184.
<https://doi.org/10.1107/S0907444904023510>
23. *Emsley P., Lohkamp B., Scott W.G. et al.* // *Acta Cryst. D.* 2010. V. 66. P. 486.
<https://doi.org/10.1107/S0907444910007493>
24. *Dey S., Lane J.M., Lee R.E. et al.* // *Biochemistry.* 2010. V. 49. P. 6746.
<https://doi.org/10.1021/bi902097j>
25. *Kwon S., Lee J.H., Kim C.M. et al.* // *Sci Rep.* 2019. V. 9. P. 6958.
<https://doi.org/10.1038/s41598-019-43490-2>
26. *Roura Padrosa D., Alaux R., Smith P. et al.* // *Front Bio-eng Biotechnol.* 2019. V. 7. P. 282.
<https://doi.org/10.3389/fbioe.2019.00282>

СТРУКТУРА МАКРОМОЛЕКУЛЯРНЫХ СОЕДИНЕНИЙ

УДК 548.73

КРИСТАЛЛИЗАЦИЯ И ПРЕДВАРИТЕЛЬНОЕ РЕНТГЕНОВСКОЕ ИССЛЕДОВАНИЕ С-КОНЦЕВОГО ФРАГМЕНТА МИОМЕЗИНА-2

© 2021 г. Е. П. Ламбер¹, Н. Пинотсис^{2,*}

¹ Институт трансляционной биомедицины Санкт-Петербургского государственного университета, Санкт-Петербург, Россия

² Институт структурной и молекулярной биологии, Биркбек Колледж, Лондон, Великобритания

*E-mail: n.pinotsis@mail.cryst.bbk.ac.uk, n.pinotsis@bbk.ac.uk

Поступила в редакцию 21.08.2020 г.

После доработки 14.09.2020 г.

Принята к публикации 22.09.2020 г.

Рекомбинантный С-концевой фрагмент миомезина-2 очищен до гомогенного состояния. Методом диффузии паров растворителя найдены условия кристаллизации С-концевого фрагмента миомезина-2, и выращены кристаллы данного белка для рентгеновского исследования. Кристаллы принадлежат пр. гр. *P222*, и имеют параметры элементарной ячейки: $a = 114.2$, $b = 142.0$, $c = 111.7$ Å. От одного из выращенных кристаллов на синхротроне ESRF собран дифракционный набор до разрешения 3.4 Å, пригодный для определения пространственной структуры С-концевого фрагмента миомезина-2.

DOI: 10.31857/S0023476121050118

ВВЕДЕНИЕ

В течение долгого времени группа кардиомиопатий (КМП) считалась заболеваниями с неизвестной этиологией. Однако с развитием молекулярно-генетических методов и молекулярной кардиологии была обнаружена связь этой группы заболеваний с генетическими дефектами синтеза различных белков. Считается, что на долю генетически обусловленных КМП приходится ~30% случаев [1, 2]. В настоящее время известно, что возникновение гипертрофической кардиомиопатии (ГКМП) связано с мутациями генов саркомерных белков [3] и, как следствие, нарушением структуры и функции сократительной мышечной единицы – саркомера.

Предположительно, белки семейства миомезина (миомезин-1, миомезин-2 и миомезин-3) [4–7] играют роль структурных звеньев, связывающих миозиновые волокна в саркомере. Миомезин-2 (изначально носивший название М-белка) наименее изучен [6]. Миомезин-2 интересен тем, что он экспрессируется в основном в быстро сокращающихся мышцах и саркомерах миокарда [8]. Недавние генетические исследования обнаружили у пациентов с диагнозом ГКМП мутации в гене миомезина-2 [9]. Кроме того, мутации миомезина-2 были обнаружены у пациентов с пороком сердца, тетрадой Фалло [10, 11]. Эти данные позволяют предположить, что миомезин-2 играет важную роль в работе сердца. Соответственно, исследования структуры и функции белка миомезина-2 важны как для изучения его роли в работе сердечной мышцы, так и с целью выявления ме-

ханизмов, приводящих к возникновению сердечной патологии, включая ГКМП и тетраду Фалло.

Миомезин-2 состоит из 13 доменов и имеет молекулярную массу 165 кДа [6]. N-концевой фрагмент миомезина-2 предположительно отвечает за связывание с миозином [12], за ним следуют два Ig-домена, пять FnIII-доменов и еще пять Ig-доменов, расположенных на С-конце. Два последних Ig-домена, mp12-mp13, представляют особый интерес, так как одна из мутаций, приводящих к развитию порока сердца, тетраде Фалло, была обнаружена именно на этом участке миомезина-2. По этой причине в настоящей работе исследован С-концевой фрагмент миомезина-2 методом белковой кристаллографии.

Перспективы использования результатов данного исследования в медицине в будущем вызывают интерес к изучению пространственной структуры миомезина-2. В настоящей работе представлена методика получения и очистки рекомбинантного С-концевого фрагмента mp12-mp13 миомезина-2. Найдены условия кристаллизации и получены кристаллы mp12-mp13. От выращенных кристаллов собран дифракционный набор до разрешения 3.4 Å, пригодный для решения пространственной структуры С-концевого фрагмента миомезина-2.

МЕТОДЫ

Выделение рекомбинантного С-концевого фрагмента mp12-mp13 миомезина-2. Рекомбинантный белок, человеческий С-концевой фрагмент миомезина-2 mp12-mp13, длиной 211 аминокислотных остатков и имеющий молекулярную массу

Таблица 1. Статистические характеристики дифракционных данных С-концевого фрагмента mp12-mp13 миомезина-2

Пр. гр.	$P222_1$
$a, b, c, \text{Å}$	114.23, 141.95, 111.68
α, β, γ , град	90.0, 90.0, 90.0
Длина волны, Å	0.976582
Разрешение, Å	47.87–3.41 (3.50–3.41)
Количество независимых рефлексов	25371 (1808)
Полнота набора	99.5 (97.0)
$I/\sigma(I)$	6.66 (0.38)
Повторяемость	4.42 (4.50)
$CC(1/2)$	99.9 (28.2)
R_{meas}	15.5 (403.8)

Примечание. В скобках приведены значения для слоя высшего разрешения.

23.7 кДа, был экспрессирован в бактериальных штаммах *E. coli* BL21 Star с использованием ранее описанного протокола аутоиндукции [13]. Клетки центрифугировали ($6000 \times g$, 15 мин) и ресуспендировали в лизирующем буфере, содержащем 25 мМ Tris-HCl, pH 7.5, 0.3 М NaCl, 5 мМ β -меркаптоэтанол (β МЕ), 10 мМ имидазола, 5% глицерина, 1 таблетку ингибиторов протеаз (Complete, EDTA-free, Roche) и 0.25 мг/мл лизоцима. Клетки лизировали в гомогенизаторе EmulsiFlex-C3 (Avestin), неочищенный экстракт центрифугировали при $50000 \times g$ в течение 45 мин. Супернатант наносили на колонку HisTrap объемом 5 мл (GE Healthcare), калиброванную лизисным буфером и подключенную к хроматографической системе АКТА (GE Healthcare). Стадии промывки выполняли множественными объемами лизирующего буфера, а также буфером с высоким содержанием соли (25 мМ Tris-HCl, pH 7.5, 1 М NaCl, 5 мМ β МЕ, 10 мМ имидазола, 5% глицерина). Белок элюировали, используя градиент имидазола (буфер для элюирования: 25 мМ Tris-HCl, pH 7.5, 150 мМ NaCl, 5 мМ β МЕ, 0.6 М имидазола, 5% глицерина). Для отрезания гистидинового тага (6His-tag) использовали протокол расщепления на колонке, описанный в [13]. Затем элюированный белок диализовали в течение ночи против буфера А (25 мМ Tris-HCl, pH 7.5, 5% глицерина) и загружали на колонку Resource Q (GE Healthcare). Белок элюировали градиентом NaCl с концентрацией ~ 0.05 – 0.1 М NaCl. Элюированный белок дополнительно концентрировали и загружали в колонку Superdex 200 16/60 (GE Healthcare), уравновешенную буфером SEC (25 мМ Tris-HCl, pH 7.5, 0.15 М NaCl, 5 мМ β МЕ, 5% глицерина). Качество белка оценивали с помощью SDS-PAGE. Чистота рекомбинантного белка составляла более 95%.

Кристаллизация. Поиск условий кристаллизации проводили методом диффузии паров растворителя, используя раствор белка с концентрацией 12 мг/мл в 25 мМ Tris-HCl-буфере, pH 7.5, содержащем 150 мМ NaCl, 5% глицерина, 5 мМ β МЕ и коммерческие наборы фирм Hampton-Research и Molecular Dimensions. Образование кристаллов наблюдали при использовании в качестве осаждающего раствора 20%-ного полиэтиленгликоля 3350 и 0.2 М трехосновного моногидрата цитрата калия при температуре 16°C.

Сбор и обработка дифракционных данных. Кристаллы были предварительно заморожены в криорастворе, содержащем 15% MPD. Дифракционные наборы собраны на синхротроне ESRF (European Synchrotron Radiation Facility, Франция) при 100 К на станции синхротронного излучения id29. Наилучший дифракционный набор собран до разрешения 3.4 Å , статистика этого набора приведена в табл. 1. Данные обрабатывали с использованием пакета программ XDS [14].

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Рекомбинантный белок, С-концевой фрагмент mp12-mp13 человеческого миомезина-2, был очищен до гомогенного состояния. Первоначальный поиск условий кристаллизации проведен методом диффузии паров растворителя в сидящей капле с использованием коммерческих кристаллизационных наборов фирм Hampton Research и Molecular Dimensions. Кристаллы были обнаружены в условиях B12 коммерческого кристаллизационного набора JCSG-plus фирмы Molecular Dimensions (рис. 1). Дифракционные наборы от замороженных кристаллов собраны при 100 К на источнике синхротронного излучения ESRF. Пример дифракционной картины приведен на рис. 2. Приведенные дифракционные характеристики (табл. 1) показывают, что удалось получить набор данных для кристаллов С-концевого фрагмента mp12-mp13 миомезина-2. Наиболее вероятными являются значения коэффициента Мэтьюса в диапазоне 2.39–3.18, соответствующие шести–восемью молекулам в независимой части и 48.5–61.37% растворителя в ячейке. В данный момент структура находится в процессе решения, одновременно ведется работа по оптимизации условий кристаллизации с целью улучшения дифракционного качества кристаллов.

Степень идентичности аминокислотных последовательностей миомезина-2 и более изученного миомезина-1, который обладает высокой эластичностью (его молекулы могут растягиваться в несколько раз и возвращаться в прежнее состояние без внутренних повреждений) [15], составляет 47%. Несмотря на довольно высокую степень гомологии, функция и локализация миомезина-1 и миомезина-2 различаются: миомезин-1 расположен параллельно миозиновым фи-

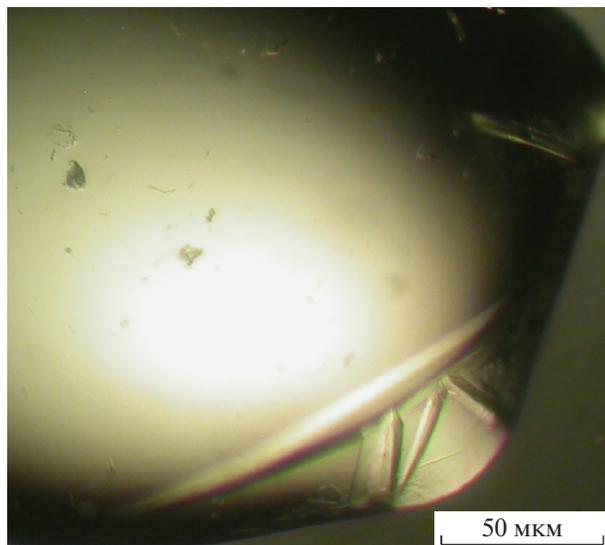


Рис. 1. Кристаллы С-концевого фрагмента $\mu\text{p12-}\mu\text{p13}$ миомезина-2, выращенные методом диффузии паров растворителя.

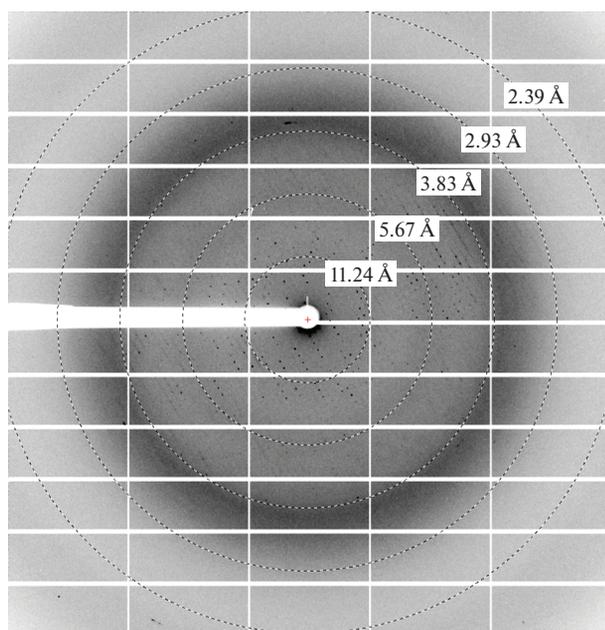


Рис. 2. Дифракционная картина, полученная с кристалла С-концевого фрагмента $\mu\text{p12-}\mu\text{p13}$ миомезина-2 на синхротроне ESRF.

ламентарам, миомезин-2 – перпендикулярно [4]. Известно, что С-концевой фрагмент миомезина-1 $\mu\text{p12-}\mu\text{p13}$ отвечает за гомодимеризацию, играющую роль в сборке М-линии саркомера [16]. Получение пространственной структуры С-концевого фрагмента миомезина-2 позволит установить различия между миомезином-1 и миомезином-2, что в дальнейшем поможет установить функцию миомезина-2. Дифракционный набор, представленный в работе, является первым набором, полученным для миомезина-2, и является важным шагом на пути определения простран-

ственной структуры миомезина-2, играющего важную роль в работе сердца.

Авторы выражают благодарность И. Агарковой (Швейцария) за дискуссии при подготовке статьи.

Наборы данных собраны на станции синхротронного излучения id29 на синхротроне ESRF (Гренобль, Франция). Авторы выражают благодарность сотрудникам id29 за помощь при получении дифракционных наборов данных.

Работа выполнена в рамках госзадания Санкт-Петербургского государственного университета (проект № 51143531).

СПИСОК ЛИТЕРАТУРЫ

1. *Towbin J.A., Hejmancik J.F., Brink P. et al.* // *Circulation*. 1993. V. 87. № 6. P. 1854. <https://doi.org/10.1161/01.cir.87.6.1854>
2. *Tsubata S., Bowles K.R., Vatta M. et al.* // *J. Clin. Invest.* 2000. V. 106. № 5. P. 655. <https://doi.org/10.1172/JCI9224>
3. *Marian A.J., Roberts R.* // *J. Mol. Cell. Cardiol.* 2001. V. 33. № 4. P. 655. <https://doi.org/10.1006/jmcc.2001.1340>
4. *Obermann W.M., Gautel M., Steiner F. et al.* // *J. Cell. Biol.* 1996. V. 134. № 6. P. 1441. <https://doi.org/10.1083/jcb.134.6.1441>
5. *Grove B.K., Kurer V., Lehner C. et al.* // *J. Cell. Biol.* 1984. V. 98. № 2. P. 518. <https://doi.org/10.1083/jcb.98.2.518>
6. *Masaki T., Takaiti O.* // *J. Biochem.* 1974. V. 75. № 2. P. 367. <https://doi.org/10.1093/oxfordjournals.jbchem.a130403>
7. *Schoenauer R., Lange S., Hirschy A. et al.* // *J. Mol. Biol.* 2008. V. 376. № 2. P. 3381. <https://doi.org/10.1016/j.jmb.2007.11.048>
8. *Agarkova I., Perriard J.-C.* // *Trends Cell. Biol.* 2005. V. 15. № 9. P. 477. <https://doi.org/10.1016/j.tcb.2005.07.001>
9. *Dorn C.* Identification of Disease-Related Genes in Congenital Heart Defects using Next-Generation Sequencing. Dissertation. Dr. RER. NAT. Free University of Berlin, 2014. 173 p.
10. *Apitz C., Webb G.D., Redington A.N.* // *Lancet*. 2009. V. 374. № 9699. P. 1462. [https://doi.org/10.1016/S0140-6736\(09\)60657-7](https://doi.org/10.1016/S0140-6736(09)60657-7)
11. *Troelsen K.D.* Myomesin 2 characterisation of a candidate gene for congenital heart defects. Dissertation. DR. RER. NAT. Copenhagen University, 2017. 145 p.
12. *Obermann W.M., van der Ven P.F., Steiner F. et al.* // *Mol. Biol. Cell.* 1998. V. 9. № 4. P. 829. <https://doi.org/10.1091/mbc.9.4.829>
13. *Mlynek G., Lehner A., Neuhold J. et al.* // *Amino Acids*. 2014. V. 46. № 6. P. 1565. <https://doi.org/10.1007/s00726-014-1699-x>
14. *Kabsch W.* // *Acta Cryst. D*. 2010. V. 66. № 2. P. 125. <https://doi.org/10.1107/S0907444909047337>
15. *Pinotsis N., Chatziefthimiou S.D., Berkemeier F. et al.* // *PLoS Biol.* 2012. V. 10. № 2. P. e1001261. <https://doi.org/10.1371/journal.pbio.1001261>
16. *Lange S., Himmel M., Auerbach D. et al.* // *J. Mol. Biol.* 2005. V. 345. № 2. P. 289. <https://doi.org/10.1016/j.jmb.2004.10.040>

СТРУКТУРА МАКРОМОЛЕКУЛЯРНЫХ СОЕДИНЕНИЙ

УДК 548.73

ПРОСТРАНСТВЕННАЯ СТРУКТУРА ТОЧЕЧНОГО МУТАНТА ЭСТЕРАЗЫ PMGL2

© 2021 г. Д. А. Горяйнова¹, А. Ю. Николаева², М. В. Крюкова², Л. Е. Петровская³,
Д. А. Корженевский², Е. А. Крюкова³, Д. А. Долгих^{3,4}, К. М. Бойко^{1,*}

¹ Институт биохимии им. А.Н. Баха РАН, Федеральный исследовательский центр “Фундаментальные основы биотехнологии” РАН, Москва, Россия

² Национальный исследовательский центр “Курчатовский институт”, Москва, Россия

³ Институт биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова РАН, Москва, Россия

⁴ Московский государственный университет им. М.В. Ломоносова, Москва, Россия

*E-mail: boiko_konstantin@inbi.ras.ru

Поступила в редакцию 17.07.2020 г.

После доработки 17.07.2020 г.

Принята к публикации 03.08.2020 г.

Ферменты из экстремофильных организмов представляют повышенный интерес для задач биотехнологии, поскольку обладают природной адаптацией к экстремальным условиям, имеющим место в технологических процессах. Липазы являются обширным классом гидролитических ферментов, которые катализируют расщепление сложноэфирных связей в триацилглицеридах, а также находят многочисленные применения в биотехнологии. Исследована кристаллическая структура точечного мутанта эстеразы PMGL2, ген которой обнаружен в результате скрининга метагеномной библиотеки ДНК из вечномерзлого грунта Сибири. Структура получена с разрешением 1.5 Å. Проведено сравнение структур мутантной формы и дикого типа PMGL2 на уровне строения субъединицы, функционального димера и активного центра фермента.

DOI: 10.31857/S0023476121050064

ВВЕДЕНИЕ

В настоящее время одним из важных направлений научных исследований в биотехнологии является изучение возможностей практического использования микроорганизмов, обитающих в экстремальных условиях (экстремофилов). Такие микроорганизмы представляют особый интерес для биотехнологического применения, поскольку являются природными продуцентами ферментов, способных функционировать в нестандартных и часто экстремальных условиях: при повышенной или пониженной температуре, в присутствии высокой солености или органических растворителей и т.д. [1, 2].

Примером практически значимых ферментов являются липазы, являющиеся обширным классом гидролитических ферментов, которые катализируют расщепление сложноэфирных связей в триацилглицеридах, а также находят многочисленные применения в различных биотехнологических процессах [3, 4]. Липолитические ферменты имеют схожую пространственную структуру — двудоменный α/β -гидролазный фолд, включающий в себя каноническую каталитическую триаду, состоящую из остатков Ser-His-Asp/Glu.

Согласно существующей классификации, основанной на наличии консервативных мотивов, липазы делятся на несколько семейств — I–VIII [5–7], пространственные структуры представителей каждого из которых к настоящему моменту установлены. Бактериальные ферменты, относящиеся к семейству IV, по своей структуре и свойствам схожи с гормон-чувствительными липазами млекопитающих и потому были названы бактериальными гормон-чувствительными липазами (bHSL). Ферменты этого семейства в свою очередь подразделяются на два класса на основании последовательности мотива активного центра — GD(E)SAG и GTSAG [8–10]. В последнем случае допускается вариативность остатка, предшествующего каталитическому серину. В отличие от хорошо изученных ферментов класса GDSAG ферменты второго класса изучены значительно меньше — установлены структуры всего трех ферментов. Одним из них является эстераза PMGL2, ген которой обнаружен в результате скрининга метагеномной библиотеки ДНК из вечномерзлого грунта. Эстераза относится к семейству HSL [11] и характеризуется наличием неканонического аминокислотного остатка (a.o.) в положении, предшествующем ка-

Таблица 1. Кристаллографические данные и параметры съемки кристалла эстеразы PMGL2

Пр. гр.	$P2_1$
$a, b, c, \text{Å}$	47.15, 92.53, 74.41
$\alpha, \beta, \gamma, \text{град}$	90.00, 106.61, 90.00
$T, \text{К}$	100
$\lambda, \text{Å}$	1.0
Разрешение, Å	56.48–1.50 (1.53–1.50)*
Число независимых рефлексов	95216 (4656)
Повторяемость	6.3 (6.3)
Полнота набора, %	97.6 (95.9)
$I/\sigma(I)$	16.1 (3.1)
$R_{\text{meas}}, \%$	7.3 (61.2)
$CC_{1/2}, \%$	99.9 (86.5)

* В скобках приведены значения для последнего слоя.

Таблица 2. Данные уточнения эстеразы PMGL2

$R_{\text{fact}}, \%$	16.8
$R_{\text{free}}, \%$	19.5
Общий средний B -фактор	18.5
Средний B -фактор по белку	17.9
Средний B -фактор по лигандам	27.2
Средний B -фактор по растворителю	25.8
Число неводородных атомов	
Белок	4833
Лиганды	24
Растворитель	414
Всего	5286
Среднеквадратичные отклонения	
Длины связей, Å	0.012
Валентные углы, град	1.891
График Рамачандрана	
Наиболее благоприятные, %	96.7
Допустимые, %	2.8
Код PDB	6ZL7

талитическому S174, в консервативном мотиве активного центра — $^{172}\text{GXSA}^{176}$. В случае PMGL2 в этом положении находится остаток цистеина — C173. С целью выяснения функциональной роли данного остатка цистеина был сконструирован ряд мутантных вариантов PMGL2, содержащих в этом положении остатки треонина, аспарагиновой кислоты и серина, встречающиеся у гомологичных ферментов [12].

В настоящей работе исследована кристаллическая структура точечного мутанта C173S эстеразы PMGL2 (m3PMGL2) с разрешением 1.5 Å . Прове-

дено сравнение структур мутантной формы и дикого типа PMGL2 на уровне строения субъединицы, функционального димера и активного центра фермента.

МАТЕРИАЛЫ И МЕТОДЫ

Очистка и выделение белка, кристаллизация. Очистку и выделение мутантной формы m3PMGL2 проводили согласно [10]. Очищенный белок концентрировали до 10 мг/мл при помощи концентраторов Millipore для последующей кристаллизации.

Кристаллизацию m3PMGL2 проводили при температуре +20°C методом диффузии в парах (висячая капля) [10]. Финальные условия кристаллизации следующие: 250 mM MgCl_2 , 100 mM HEPES 12% PEG 3350, pH 7.5.

Сбор и обработка дифракционных данных. Решение и уточнение структуры. Непосредственно перед сбором данных кристаллы m3PMGL2 кратковременно вымачивали в кристаллизационном растворе, содержащем дополнительно 25% глицерина в качестве криопротектора, после чего кристаллы вылавливали при помощи петили и помещали под струю жидкого азота. Дифракционные данные собраны на станции синхротронного источника SPring-8 (Harima Science Garden, Япония) при температуре 100 К. Дифракционная картина фиксировалась детектором Pilatus6M. Экспериментальные данные проиндексированы и обработаны с использованием программы iMOSFLM [13]. С помощью программы Pointless [14] установлена пр. гр. $P2_1$. Статистика набора данных приведена в табл. 1.

Структура m3PMGL2 получена с разрешением 1.50 Å при помощи метода молекулярного замещения с применением программы MOLREP [15]. В качестве стартовой модели использовали структуру PMGL2 дикого типа (код PDB6QIN). Уточнение структуры проводили с помощью программы REFMAC5 из набора CCP4 [16]. Визуальный осмотр карт электронной плотности и уточнение модели проводили с использованием программы COOT [17]. Уточнение структуры выполняли до достижения R -факторами следующих значений: $R_{\text{work}} = 16.8\%$, $R_{\text{free}} = 19.5\%$ (табл. 2).

В случае m3PMGL2 в независимой части элементарной ячейки кристалла находятся две белковые цепи (в цепи A — 317 остатков, в цепи B — 320 остатков), один ион магния, две молекулы PEG 3350 и 453 молекулы растворителя. Электронная плотность отсутствует для 16 и 13 N-концевых аминокислот для цепей A и B соответственно, что может быть связано с их относительной подвижностью. Также в обеих цепях молекулы для а.о. 221–226 электронная плотность имела



Рис. 1. Наложение структуры m3PMGL2 (темный) на структуру PMGL2 дикого типа (светлый). Расположение петли 221–226 обозначено стрелкой.

низкий уровень, позволивший, однако, достроить эти участки.

Анализ структуры проводили с использованием программы COOT и системы молекулярной графики PyMOL, Version 1.9.0.0 (Schrodinger, USA). Сравнение и наложение структур проводили с помощью сервиса PDBeFold [18], а контакты анализировали с использованием сервиса PDBePISA [19].

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Кристаллы m3PMGL2 принадлежат пр. гр. $P2_1$. В независимой части элементарной ячейки кристалла находятся две субъединицы белка, образующие функциональный димер.

Структура m3PMGL2 имеет типичный α/β -гидролазный фолд, состоящий из двух канонических доменов – CAP-домена и каталитического домена. Домен CAP содержит две α -спирали ($\alpha 1$ и $\alpha 2$). Каталитический домен состоит из восьми β -тяжей ($\beta 1$ – $\beta 8$), окруженных девятью α -спиралями ($\alpha 3$ – $\alpha 11$) [10], и содержит в активном центре каталитический остаток серина – S174.

Анализ контактов между субъединицами из независимой части показал, что m3PMGL2 аналогично многим другим эстеразам является димером в кристалле. Скрытая при формировании димера поверхность (buried surface area) составляет 1508 \AA^2 (11.3% поверхности субъединицы). Димер стабилизирован 10 солевыми мостиками и 20 водородными связями. Оценка гидрофобных межмолекулярных взаимодействий в димере, сделанная по выигрышу свободной энергии при формировании межмолекулярного контакта, да-

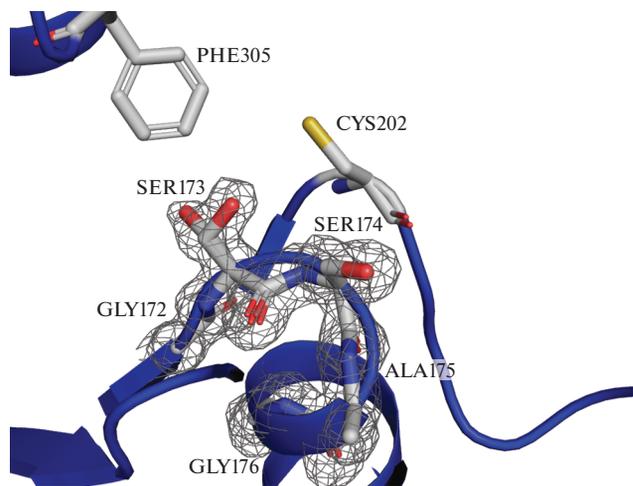


Рис. 2. Электронная плотность 2FOFC (серая сетчатая поверхность) вокруг остатка S173. Уровень срезы – 1σ . Палочковой моделью показано двойное положение остатка серина и окружающие его остатки активного центра (подписаны). Темным цветом показан ход полипептидной цепи белка.

ет величину -7.5 ккал/моль. Полученные значения схожи с таковыми для дикого типа PMGL2 [10].

Проведенный анализ показал, что структуры субъединиц m3PMGL2 и дикого типа фермента также весьма схожи ($\text{RMSD} = 0.181 \text{ \AA}$). Наложение субъединиц продемонстрировало, что элементы вторичной структуры расположены идентично, что говорит об отсутствии влияния точечной замены на общую структуру фермента (рис. 1). Основные различия между структурами обнаружены в петлях 221–226. В молекуле PMGL2 дикого типа а.о. 221–226 не имели электронной плотности [10] в отличие от молекулы m3PMGL2, где эта область имеет электронную плотность, хотя и низкого уровня.

Полость активного центра m3PMGL2 сформирована остатками G172, S173, A175, G176. В активном центре m3PMGL2 на месте цистеина, предшествующего каталитическому S174, находится серин, что подтверждает проведенную мутацию (рис. 2). Электронная плотность для S173 демонстрирует, что он имеет двойное положение: в одном положении его боковая группа ориентирована по направлению к C202, а в другом – к боковой группе F305. В последнем положении S173 способен образовывать водородную связь с боковой группой E120 через молекулу растворителя.

ЗАКЛЮЧЕНИЕ

Исследована структура мутантной формы эстеразы PMGL2, проведено ее сравнение со структурой белка дикого типа и показано, что на уровне структуры димера, субъединицы и активного

центра различий в строении двух форм PMGL2 не обнаружено. Полученные результаты позволяют утверждать, что замена цистеина на серин в положении 173 не влияет на структуру фермента, при этом биохимические характеристики мутантной формы близки к таковым для дикого типа [12].

Работы по экспрессии и выделению белка выполнены при финансовой поддержке Российского фонда фундаментальных исследований (грант № 18-04-00491), работы по кристаллографическому уточнению и анализу полученной структуры выполнены при поддержке Министерства науки и высшего образования РФ. Статья подготовлена по материалам кристаллизационных исследований, выполненных в ЦР (КЭ) “Кристаллизатор” (Шифр ЦР (КЭ) “МКС (Наука)” ФНИЦ-1) на российском сегменте Международной космической станции.

СПИСОК ЛИТЕРАТУРЫ

1. Kumar L.A.G., Singh B. // *Biotechnology*. 2011. V. 10. P. 121.
<https://doi.org/10.3923/biotech.2011.121.135>
2. Karan R., Capes M.D., Dassarma S. // *Aquat. Biosyst.* 2012. V. 8. P. 4.
<https://doi.org/10.1186/2046-9063-8-4>
3. Gaur R., Hemamalini R., Khare S.K. // *Current Developments in Biotechnology and Bioengineering: Production, Isolation and Purification of Industrial Products*. 2017. P. 175.
<https://doi.org/10.1016/B978-0-444-63662-1.00008-7>
4. Casas-Godoy L., Duquesne S., Bordes F. et al. // *Methods Mol. Biol.* 2012. V. 861. P. 3.
https://doi.org/10.1007/978-1-61779-600-5_1
5. Ollis D.L., Cheah E., Cygler M. et al. // *Protein Eng.* 1992. V. 5. P. 197.
<https://doi.org/10.1093/protein/5.3.197>
6. Arpigny J.L., Jaeger K.E. // *Biochem. J.* 1999. V. 343. P. 177.
7. Nardini M., Dijkstra B.W. // *Curr. Opin. Struct. Biol.* 1999. V. 9. P. 732.
[https://doi.org/10.1016/S0959-440X\(99\)00037-8](https://doi.org/10.1016/S0959-440X(99)00037-8)
8. Jeon J.H., Lee H.S., Kim J.T. et al. // *Appl. Microbiol. Biotechnol.* 2012. V. 93. P. 623.
<https://doi.org/10.1007/s00253-011-3433-x>
9. Li P.Y., Ji P., Li C.Y. et al. // *J. Biol. Chem.* 2014. V. 289. P. 19031.
<https://doi.org/10.1074/jbc.M114.574913>
10. Boyko K.M., Kryukova M.V., Petrovskaya L.E. et al. // *PLOS ONE*. 2020. V. 15 (1). P. e0226838.
<https://doi.org/10.1371/journal.pone.0226838>
11. Petrovskaya L.E., Novototskaya-Vlasova K.A., Spirina E.V. et al. // *FEMS Microbiol. Ecol.* 2016. V. 92. P. fiw046.
<https://doi.org/10.1093/femsec/fiw046>
12. Kryukova M.V., Petrovskaya L.E., Novototskaya-Vlasova K.A. et al. // *Biochemistry*. 2020. V. 85. P. 831.
<https://doi.org/10.1134/S0006297920060085>
13. Battye T.G., Kontogiannis L., Johnson O. et al. // *Acta Cryst. D*. 2011. V. 67. P. 271.
<https://doi.org/10.1107/S0907444910048675>
14. Evans P. // *Acta Cryst. D*. 2006. V. 62. P. 72.
<https://doi.org/10.1107/S0907444905036693>
15. Vagin A., Teplyakov A. // *J. Appl. Cryst.* 1997. V. 30. P. 1022.
<https://doi.org/10.1107/S0021889897006766>
16. Collaborative Computational Project N. // *Acta Cryst. D*. 1994. V. 50. P. 760.
<https://doi.org/10.1107/S0907444994003112>
17. Emsley P., Cowtan K. // *Acta Cryst. D*. 2004. V. 60. P. 2126.
<https://doi.org/10.1107/S0907444904019158>
18. Krissinel E., Henrick K. // *Acta Cryst. D*. 2004. V. 60. P. 2256.
<https://doi.org/10.1107/S0907444904026460>
19. Krissinel E., Henrick K. // *J. Mol. Biol.* 2007. V. 372. P. 774.
<https://doi.org/10.1016/j.jmb.2007.05.022>

СТРУКТУРА МАКРОМОЛЕКУЛЯРНЫХ СОЕДИНЕНИЙ

УДК 577.322.9

КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ СТРУКТУР ОБРАТИМО ПЕРЕКЛЮЧАЕМЫХ ФЛУОРЕСЦЕНТНЫХ БЕЛКОВ С LOV-ДОМЕНАМИ

© 2021 г. Ю. И. Метелешко¹, М. Г. Хренова^{1,2,*}, А. В. Немухин^{1,3}

¹ Московский государственный университет им. М.В. Ломоносова, Москва, Россия

² Институт биохимии им. А.Н. Баха, Федеральный исследовательский центр “Фундаментальные основы биотехнологии” РАН, Москва, Россия

³ Институт биохимической физики им. Н.М. Эмануэля РАН, Москва, Россия

*E-mail: wasabiko@lcc.chem.msu.ru

Поступила в редакцию 06.08.2020 г.

После доработки 14.09.2020 г.

Принята к публикации 16.09.2020 г.

Методами молекулярного моделирования впервые построены полноатомные трехмерные структуры новых обратимо переключаемых флавинодержущих белков с LOV-доменами (Light, Oxygen, Voltage), используемых в микроскопии сверхвысокого разрешения. Для вариантов белков rsLOV1 и rsLOV2 известны первичные последовательности, но не пространственные модели. При построении моделей rsLOV1 и rsLOV2 за основу взята структура PDB ID: 2MWG фоторецепторного белка YtvA из *Bacillus subtilis* и выполнены необходимые точечные замены аминокислотных остатков. Построенные модели заключены в водные оболочки, для которых выполнены расчеты молекулярно-динамических траекторий и проведен динамический сетевой анализ. Показано, что динамические свойства нативной и мутантных структур различаются прежде всего в отношении α -спирали, ответственной за передачу сигнала при фотовозбуждении.

DOI: 10.31857/S0023476121050143

ВВЕДЕНИЕ

Интерес к флавинодержущим белкам с LOV-доменами (Light, Oxygen, Voltage) объясняется их применением в качестве флуоресцентных маркеров в живых системах, а в последнее время – в составе оптогенетических конструкций [1, 2]. По сравнению с известными маркерами семейства зеленого флуоресцентного белка флуоресцентные белки на основе LOV-доменов обладают рядом преимуществ: небольшой размер, применимость в анаэробных условиях, возможность применения в широком диапазоне pH (от 4 до 11) [3]. В качестве хромофора большинство LOV-доменов связывает флавиномононуклеотид, широко распространенный в клетках.

Фотопереключение белков с LOV-доменом основано на реакциях образования и разрыва ковалентной связи между атомом серы консервативного аминокислотного остатка цистеина и C4a-атомом флавина. В активном, флуоресцентном состоянии “ON” связи между цистеином и флавином нет, но при освещении синим цветом (488 нм) формируется связь (Cys)S–C4a, и белок переходит в нефлуоресцентное состояние “OFF”. В природных объектах образование ковалентно-связанного аддукта запускает ряд структурных

изменений белка, при которых сигнал передается белку-эффектору. При освещении УФ-светом (408 нм) аддукт распадается, и восстанавливается состояние “ON”.

Первые попытки создать обратимо фотопереключаемый маркер были предприняты для белка YtvA-LOV из *Bacillus subtilis* [4]. Однако слабая флуоресценция и низкая эффективность переключения во флуоресцентное состояние под действием УФ-света не способствовали практическому применению природного белка. В [5] были получены новые варианты обратимо фотопереключаемых белков, названных rsLOV1 и rsLOV2, демонстрирующие повышенную яркость и улучшенную кинетику переключения. Данные белки были синтезированы олигонуклеотид-направленным мутагенезом на основе YtvA-LOV.

В отличие от исходного природного белка YtvA-LOV, для которого известна пространственная структура (рис. 1), для rsLOV1 и rsLOV2 ранее была известна только первичная последовательность. Знание пространственной структуры необходимо для анализа путей передачи сигнала при фотопереключенном возбуждении.

Для решения задачи построения трехмерных полноатомных структур rsLOV1 и rsLOV2 в данной работе использованы современные методы

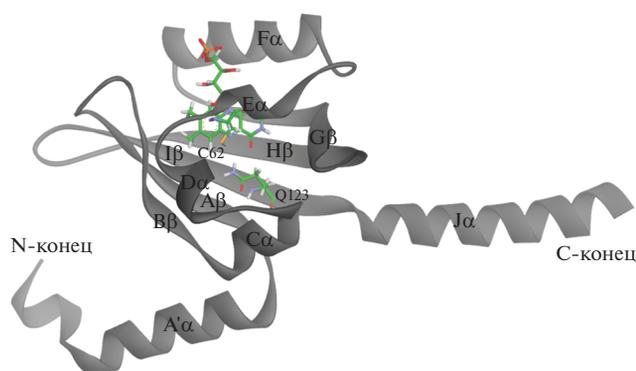


Рис. 1. Структура домена YtvA-LOV, построенная на основе PDB ID: 2MWG. Стержневой моделью показаны хромофор и аминокислотные остатки Cys62 и Gln123.

молекулярного моделирования. Анализ литературы показывает, что подобные подходы могут оказаться успешными при изучении структурных изменений, происходящих в LOV-доменах [6].

МАТЕРИАЛЫ И МЕТОДЫ

Полноатомные молекулярные модели rsLOV1 и rsLOV2, мутантов YtvA-LOV, получены на основе известной структуры YtvA *Bacillus subtilis* (PDB ID: 2MWG). Модели построены с тем же количеством аминокислотных остатков, что и в [5]: 1-146 для YtvA-LOV и 1-142 для rsLOV1 и rsLOV2. Для получения структур rsLOV1 и rsLOV2 проведены точечные замены в соответствии с известными первичными последовательностями (рис. 2).

Боковые цепи аминокислот лизина и аргинина предполагались положительно заряженными, глутаминовой и аспарагиновой кислот – отрицательно заряженными. Модельные системы были сольватированы в водных оболочках размером $94 \times 74 \times 76 \text{ \AA}^3$ для YtvA-LOV и $91 \times 74 \times 76 \text{ \AA}^3$ для rsLOV1 и rsLOV2. Для нейтрализации зарядов систем были добавлены ионы натрия и хлора таким

образом, чтобы ионная сила раствора составляла 0.15 М.

Молекулярно-динамические (МД) расчеты проводили с использованием программного пакета NAMD [7]. Для каждой системы рассчитано по три МД-траектории длительностью 100 нс (шаг интегрирования 1 фс). Динамику Ланжевена использовали для сохранения температуры 300 К и давления 1 атм в NPT-ансамбле. Для электростатических взаимодействий использовали алгоритм PME [8]. Энергии и силы вычисляли с силовым полем CHARMM36 [9] для белковой макромолекулы, TIP3P для молекул воды [10]. Для хромофора использовали параметры и типы атомов CGenFF [11], оптимизированные в [12]. Сетевой динамический анализ выполняли с использованием программ VMD [13] и Carma [14, 15].

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Анализ полученных МД-траекторий показывает, что самое активное движение наблюдается для A'α- и Jα-спиралей (рис. 1), на рис. 3 представлены все варианты их положения. Как видно из рисунка, A'α-спираль либо контактирует с петлей Hβ-Iβ (при этом N-концевая часть находится или под спиралью (рис. 3а), или за ней (рис. 3в)), либо находится ниже N-концевой части белка (рис. 3б, 3г, 3е). Последние два завитка Jα-спирали обладают повышенной подвижностью, причем в некоторых траекториях они стремятся потерять свою структуру (рис. 3в), а в одной траектории, для YtvA-LOV, половина Jα-спирали полностью развернулась (рис. 3е). Для rsLOV2 также наблюдалось необычное движение A'α- и Jα-спиралей навстречу друг другу (на рис. 3г показано их максимальное сближение), после чего они вернулись в обычные позиции. Отметим, что для YtvA-LOV во всех траекториях A'α-спираль располагается ниже N-концевой части белка (рис. 3д, 3е).

YtvA-LOV	1	M A S F Q S F G I P G Q L E V I K K A L D H I R V G V V I T D P A L E D N P I V Y V N Q G F V Q M T
rsLOV1	1	M T R F Q S F G I S G Q L E V I K K A L D H I R V G V V I T D P A R E D N P I V Y V N Q G F V Q M T
rsLOV2	1	M T R F Q S F G I S G Q L E V I K K A L D H I R V G V V I T D P A R E D N P V Y V N N G F V Q M T
YtvA-LOV	51	G Y E T E E I L G K N C R F L Q G K H T D P A E V D I I R T A L Q N K E P V T V Q I I N Y K K D G T
rsLOV1	51	G Y E A E E I L G K N C R I L Q G E H T D P A E V D I I R T A L Q N K E P V T V Q I I L N Y R K D G T
rsLOV2	51	G Y E A E E I L G K S C R I L Q G E H T D P A E V D I I R T A L Q N K E P V T V Q I I L N Y R K D G T
YtvA-LOV	101	M F W N E L N I D P M E I E D K T Y F V G I Q N D I T K Q K E Y E K L L E D S L T E I T A L
rsLOV1	101	M F W N L L H I V P I V I E G K T Y F V G N Q N D I T K Q K E Y E K L L E R P H R G ----
rsLOV2	101	M F W N L L H I V P I V I E G K T Y F V G N Q N D I T K Q K E Y E K L L E R P H R G ----

Рис. 2. Выравненные аминокислотные последовательности YtvA-LOV, rsLOV1 и rsLOV2. Цветом выделены различия.

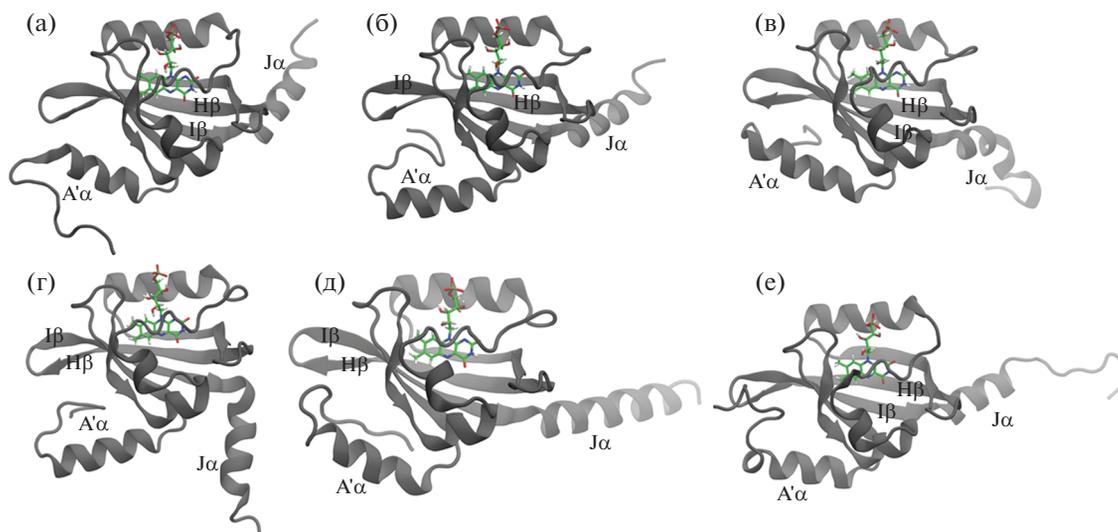


Рис. 3. Характерные конфигурации из молекулярно-динамических траекторий: а, б, г – rsLOV2, в – rsLOV1, д, е – YtvA-LOV.

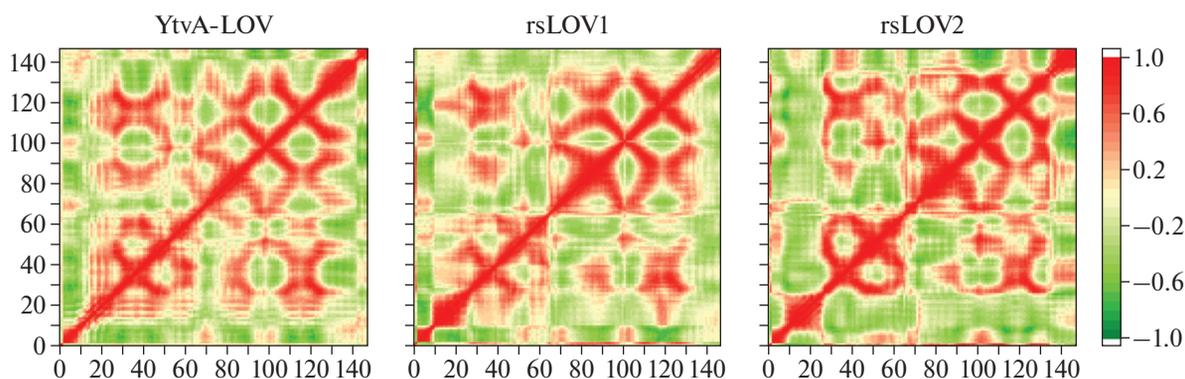


Рис. 4. Дисперсионно-ковариационные матрицы. По осям отложены номера узлов сети. Узлы 0 и 1 отвечают флавиону (изоалоксазиновому кольцу и рибитиловому хвосту соответственно), узлы от 2 до 148 соответствуют аминокислотам 1–146.

Дополнительную информацию о сходстве и различии структур исследуемых систем предоставляет сетевой динамический анализ. На рис. 4 представлены дисперсионно-ковариационные матрицы. Области, отвечающие скоррелированному движению аминокислотных остатков, отмечены более темными цветами; у rsLOV1 и rsLOV2 они шире, чем у YtvA-LOV. Существенное отличие наблюдается для rsLOV2 в области между 30-м и 62-м остатками, что, вероятно, связано с тем, что там находятся дополнительные мутации, отсутствующие у rsLOV1 (I39V, Q44N, N61S).

Для исследуемых систем также рассмотрены изменения, происходящие в хромофорсодержащем кармане. Самой подвижной боковой цепью в этой области обладает остаток Gln123 (рис. 1). Известно, что аналогичный остаток глутамин сохраняется во всех LOV-доменах и играет важную роль в передаче сигнала [6]. Чтобы проиллюстрировать поведение боковой цепи Gln123, были

рассмотрены такие параметры, как двугранный угол NE2–CD–CG–CB и расстояние между атомом NE2 глутамин и N5 флавина (рис. 5). Из графиков видно, что для природного белка YtvA-LOV характерны расстояния от 4 до 5.5 Å и величины угла от 10° до 70° (рис. 6а). Для белков rsLOV1 и rsLOV2 наблюдается большая заселенность состояний в области около 3–3.5 Å и от 150° до –120° (рис. 6б). Также для rsLOV2 наблюдается слабо заселенное состояние с расстояниями больше 6 Å (рис. 6в), где боковая цепь Gln123 направлена в сторону Jα-спирали.

Для численного сравнения подвижности боковой цепи глутамин для каждой системы была подсчитана доля кадров в МД-траекториях, в которых расстояние NE–N5 не превышало 3.5 Å: 2.6% для YtvA-LOV, 15.9% для rsLOV1 и 13.7% для rsLOV2.

Таким образом, впервые исследованы структуры новых обратимо переключаемых белков

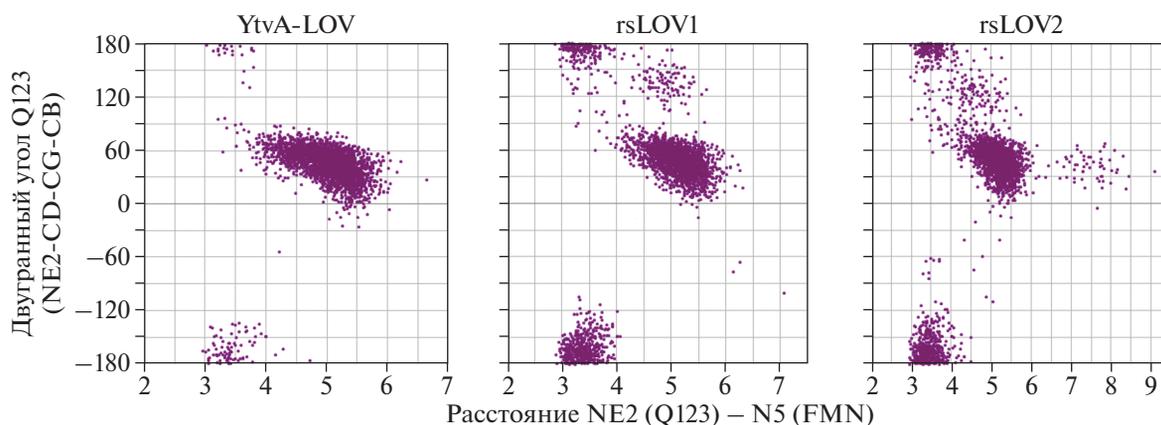


Рис. 5. Графики зависимости расстояния NE2–N5 (Å) от угла NE2–CD–CG–CB для Gln123.

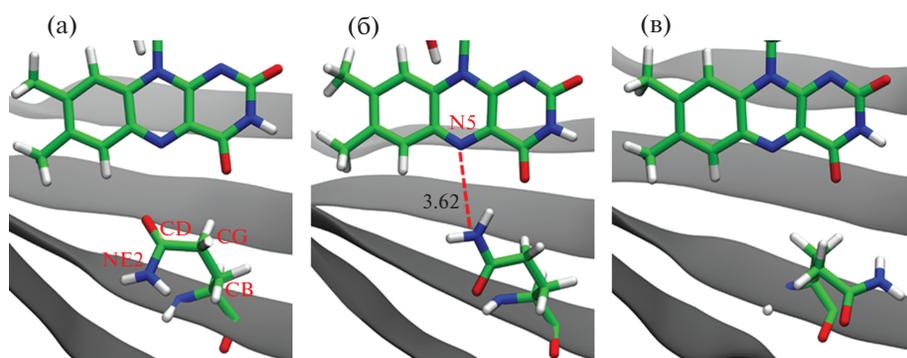


Рис. 6. Конформеры Gln123 с разной величиной угла NE2–CD–CG–CB: а – 50°, б – 160°, в – 50°.

rsLOV1 и rsLOV2 и проведено их сравнение со структурой YtvA-LOV. Выявлены отличия в поведении Gln123, находящегося в хромофорсвязывающем кармане. В белках rsLOV1 и rsLOV2 боковая цепь Gln123 обладает большей подвижностью и гибкостью, чем в YtvA-LOV.

Работа выполнена при поддержке Российского научного фонда (проект № 17-13-01051) с использованием оборудования Центра коллективного пользования сверхвысокопроизводительными вычислительными ресурсами МГУ им. М.В. Ломоносова.

СПИСОК ЛИТЕРАТУРЫ

1. Shcherbakova D.M., Shemetov A.A., Kaberniuk A.A. et al. // *Annu. Rev. Biochem.* 2015. V. 84. P. 519. <https://doi.org/10.1146/annurev-biochem-060614-034411>
2. Losi A., Gardner K.H., Möglich A. // *Chem. Rev.* 2018. V. 118. P. 10659. <https://doi.org/10.1021/acs.chemrev.8b00163>
3. Mukherjee A., Walker J., Weyant K.B. et al. // *PLoS ONE.* 2013. V. 8. P. e64753. <https://doi.org/10.1371/journal.pone.0064753>
4. Losi A., Gärtner W., Raffelberg S. et al. // *Photochem. Photobiol. Sci.* 2013. V. 12. P. 231. <https://doi.org/10.1039/C2PP25254F>
5. Gregor C., Sindenstein S.C., Andresen M. et al. // *Sci. Rep.* 2018. V. 8. P. 1. <https://doi.org/10.1038/s41598-018-19947-1>
6. Freddolino P.L., Gardner K.H., Schulten K. // *Photochem. Photobiol. Sci.* 2013. V. 12. P. 1158. <https://doi.org/10.1039/c3pp25400c>
7. Phillips J.C., Braun R., Wang W. et al. // *J. Comput. Chem.* 2005. V. 26. P. 1781. <https://doi.org/10.1002/jcc.20289>
8. Darden T., York D., Pedersen L. // *J. Chem. Phys.* 1993. V. 98. P. 10089. <https://doi.org/10.1063/1.464397>
9. Best R.B., Zhu X., Shim J. et al. // *J. Chem. Theory Comput.* 2012. V. 8. P. 3257. <https://doi.org/10.1021/ct300400x>
10. Jorgensen W.L., Madura J.D. // *J. Am. Chem. Soc.* 1983. V. 105. P. 1407. <https://doi.org/10.1021/ja00344a001>
11. Vanommeslaeghe K., Hatcher E., Acharya C. et al. // *J. Comput. Chem.* 2010. V. 31. P. 671. <https://doi.org/10.1002/jcc.21367>
12. Khrenova M.G., Meteleshko Y.I., Nemukhin A.V. // *J. Phys. Chem. B.* 2017. V. 121. P. 10018. <https://doi.org/10.1021/acs.jpcc.7b07533>
13. Humphrey W., Dalke A., Schulten K. // *J. Mol. Graph.* 1996. V. 14. P. 33. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5)
14. Glykos N.M. // *J. Comput. Chem.* 2006. V. 27. P. 1765. <https://doi.org/10.1002/jcc.20482>
15. Koukos P.I., Glykos N.M. // *J. Comput. Chem.* 2013. V. 34. P. 2310. <https://doi.org/10.1002/jcc.23381>

СТРУКТУРА МАКРОМОЛЕКУЛЯРНЫХ
СОЕДИНЕНИЙ

УДК 539.219.1

О ТАКТИКЕ *AB INITIO*-ПОИСКА ФОРМЫ БЕЛКОВЫХ ЧАСТИЦ
ПО ДАННЫМ МАЛОУГЛОВОГО РЕНТГЕНОВСКОГО РАССЕЯНИЯ© 2021 г. В. В. Волков^{1,*}¹ Институт кристаллографии им. А.В. Шубникова ФНИЦ “Кристаллография и фотоника” РАН, Москва, Россия

*E-mail: volkicras@mail.ru

Поступила в редакцию 23.09.2020 г.

После доработки 02.01.2021 г.

Принята к публикации 12.03.2021 г.

Определение 3D-формы частиц по одномерным данным малоуглового рассеяния от растворов макромолекул неоднозначно. В силу плохой обусловленности обратной задачи, решение неустойчиво и зависит от параметров алгоритма поиска. Следовательно, на практике необходимо не только оценивать степень стабильности решения, но и подбирать параметры как метода решения, так и самой модели. Рассмотрена тактика поиска, заключающаяся в последовательном определении набора моделей формы частицы, представленной в виде структуры, состоящей из плотноупакованных шариков малого размера. Набор решений получают при варьировании относительного вклада членов целевой функции: критерия отклонения модельной кривой рассеяния от экспериментальной, штрафов за рыхлость и разрыв структуры тела, за отклонение среднего числа контактов шариков от заданного значения. Приведены примеры решения модельных задач и определения формы молекул по измерениям, выложенным в банке малоугловых данных и моделей SASBDB.

DOI: 10.31857/S0023476121050234

ВВЕДЕНИЕ

К настоящему времени разработано несколько методов *ab initio*-восстановления формы с использованием модели малых объемных элементов [1–3]. Поиск 3D-формы частицы осуществляется минимизацией суммарной квадратичной невязки между экспериментальной кривой рассеяния и теоретической, рассчитываемой от модели структуры, представленной набором плотноупакованных элементов малого объема. В качестве элементов берут шарики или кубики такого размера, чтобы дифракция на их упаковке влияла на кривую рассеяния за пределами экспериментального диапазона. Кубические элементы позволяют более однородно заполнить пространство, однако и в случае шариков поправку на однородность легко учесть с помощью поправочного коэффициента. Поиск решения (пространственного размещения шариков в модели) ведут с помощью методов глобальной минимизации, например моделированием отжига (*simulated annealing* [4]) [1, 2] или с помощью генетических алгоритмов [3]. Алгоритм, основанный на поиске решения методом моделирования отжига, оказался исключительно эффективным в силу ряда алгоритмических решений, в частности из-за отказа от вариации на каждом шаге поиска всех параметров модели (число параметров равно числу узлов в прямом пространстве, в которых могут быть раз-

мещены шарики, заполняющие пространство и моделирующие электронную плотность модели). Вместо этого проводят варьирование только одного параметра, выбираемого случайным образом. При этом экспериментальные данные рассеяния приближают с наложением ограничений (прежде всего, плотность и неразрывность структуры) на решение.

ТЕОРЕТИЧЕСКАЯ ЧАСТЬ

Коротко основная идея моделирования, реализованная в программе DAMMIN [1], заключается в следующем. Область поиска (например, сфера с диаметром, равным максимальному размеру частицы D_{\max}) разбивается на N ($\sim 10^3$ – 10^5) узлов с плотнейшей упаковкой. D_{\max} определяют из функции парных расстояний $p(r)$, вычисленной из экспериментальных данных $I(s)$, например, с помощью программы GNOM [5] из пакета ATSAS [6] по формуле

$$p(r) = \frac{1}{2\pi^2} \int_{s=0}^{\infty} I(s) \frac{\sin(sr)}{sr} ds, \quad (1)$$

где r – длина отрезка в структуре, $s = 4\pi \sin \theta / \lambda$ – модуль вектора рассеяния, 2θ – угол рассеяния, λ – длина волны излучения. Каждому j -му узлу приписывается индекс X_j , обозначающий фазу,

которой принадлежит данный узел ($X_j = 0$ означает растворитель, $X_j = 1$ – частица). Таким образом, структура модели описывается конфигурационным вектором длиной N . В методе DAMMIN интенсивность рассеяния от такой модели быстро рассчитывается с помощью мультипольного разложения функции амплитуды рассеяния от объемного элемента. Тактика моделирования отжига заключается в старте с произвольной модели и ее последующих случайных модификациях до тех пор, пока теоретическое рассеяние от модели не станет приближать данные. В обычном методе Монте-Карло на каждом шаге используют две модели – лучшую и пробную, полученную случайной вариацией лучшей модели. Пробная модель отвергается, если она хуже, и принимается в качестве лучшей в противном случае, после чего цикл повторяют. Однако такой метод, несмотря на его стохастическую природу, на практике не способен найти решение за приемлемое время, застревая в широких локальных минимумах целевой функции. В методе отжига в качестве текущей структуры принимают модель, которая не обязательно лучше предыдущей. Случайную вариацию проводят относительно текущей модели и получают пробную. Вероятность принять худшую пробную модель в качестве новой текущей зависит от некоторого параметра, который называют в литературе “температурой” T . Большая температура означает высокую вероятность, вычисляемую как $\exp(-\Delta/T)$, принять в качестве новой текущей модели худшую, если Δ (пробное значение целевой функции минус текущее) больше 0 (решение хуже). Если пробная модель оказывается лучше, т.е. $\Delta < 0$, то она всегда принимается в качестве новой текущей. В начале поиска температуру выбирают достаточно высокой, чтобы частота принятия худших решений в качестве текущих превышала частоту обновления лучших моделей. Это заставляет программу “блуждать” по области поиска и в случае попадания модели в окрестность другого локального минимума, в конце концов, выбираться из текущего. По мере роста числа испытаний температуру периодически снижают. Чаще всего используют мультипликативный закон $T_{next} = T_{curr}F$, где F – фактор отжига, равный в программе DAMMIN 0.9–0.98. Значение целевой функции уменьшается по аналогии с уменьшением внутренней энергии системы по мере снижения температуры, из-за чего метод и получил название “моделирование отжига” [4].

Поскольку для определения формы используются данные рассеяния, диапазон которых соответствует низкому пространственному разрешению, искомая модель также должна иметь более низкое разрешение по отношению к размеру объемного элемента. Поэтому на модель накладываются условия связности и плотности. Полная целевая функция, использованная в данной работе,

имеет вид штрафной функции, зависящей от элементов конфигурационного вектора:

$$\Phi(\mathbf{X}) = w_R R + w_D P_D + w_L P_L + w_C P_C + \dots \quad (2)$$

Здесь R – критерий сходства модельной и экспериментальных кривых рассеяния

$$R = \left\{ \frac{\sum_{i=1}^N [(I_{\text{exp}}(s_i) - \xi I_{\text{mod}}(s_i)) W(s_i)]^2}{\sum_{i=1}^N [I_{\text{exp}}^2(s_i) W^2(s_i)]} \right\}, \quad (3)$$

где $\xi = \frac{(\mathbf{I}_{\text{exp}} \cdot \mathbf{I}_{\text{mod}})}{\|\mathbf{I}_{\text{exp}}\|^2}$ – МНК-множитель, совмещающий кривые рассеяния, умноженные на $W(s)$ – весовую функцию, которую назначают в виде

$$W(s) = \begin{cases} s^n, & n = 0, 1, 2, 3, 4 \\ \text{если } s > s_{\max[I_{\text{exp}}(s)s^n]}, \\ 0.5 \left\{ s^n + \frac{\max[I_{\text{exp}}(s)s^n]}{I_{\text{exp}}(s)} \right\} \\ \text{если } s \leq s_{\max[I_{\text{exp}}(s)s^n]} \end{cases} \quad (4)$$

Умножение на $W(s)$ позволяет выравнивать вклад отклонений вдоль кривой рассеяния: при $n = 0$ расчет невязки (3) происходит на исходной шкале данных (что имеет смысл при спаде интенсивности не более одного порядка), с увеличением n кривая интенсивности трансформируется в контур, проходящий через максимальное значение на некотором $s = s_{\max[I_{\text{exp}}(s)s^n]}$ и спадающий до некоторого значения, тем большего, чем больше n . Тем самым достигается возможность ослабить вклад в суммарную невязку начального участка, который наиболее подвержен искажающему влиянию рассеяния от агрегатов частиц, и одновременно увеличить вклад малоинтенсивного участка при больших s . Нижняя строка в (4) служит для того, чтобы ослабление начального участка было не слишком большим. На практике n рекомендуется выбирать таким образом, чтобы величина $\frac{\max[I_{\text{exp}}(s)s^n]}{\min[I_{\text{exp}}(s)s^n]}$ была в диапазоне 10–50. При таком взвешивании обычное в МНК-методе деление невязок на стандартное отклонение экспериментальных шумов не применяется.

Для адекватного описания формы частицы и исключения влияния на кривую рассеяния упаковок шариков их диаметр должен быть достаточно мал. Для обеспечения этого число узлов в области поиска выбирают не менее 2000–5000 для компактных частиц и до 10000–30000 для анизометричных тел. При этом размер шариков оказывается в 5–10 раз меньше, чем пространственное

разрешение структурной модели, и для описания однородных областей структуры они должны находиться в плотноупакованном состоянии. Для обеспечения этого в целевую функцию добавлены штрафные члены, отражающие требования неразрывности структуры $w_D P_D$ и ее диффузности $w_L P_L$. Величину P_D вычисляют как отношение общего числа шариков структуры к числу шариков, составляющих максимальный связный домен. P_L (штраф за “рыхлость”) вычисляют по формуле $P_L = 1 - \langle 1 - \exp(-N_e) + \exp(-12) \rangle_N$, где 12 – максимальное число контактов шарика с соседями, N_e – фактическое число контактов, $\langle \rangle_N$ означает усреднение числа контактов, приведенное к одному шарiku по структуре. Для весовых коэффициентов w_D и w_L рекомендованы значения 0.01–0.001 из соображения примерного равенства вкладов от R и штрафов в точке минимума функционала (2). P_C – штраф за относительное смещение центра тяжести частицы из центра области поиска. Этот член предотвращает “прилипание” модели к границе области и тем самым влияние границы на форму.

Намного более эффективный алгоритм поиска шариковых моделей реализован в программе DAMMIF [2]. Если DAMMIN осуществляет поиск модели в ограниченной области пространства и вхолостую просматривает много узлов, находящихся вдали от сформировавшегося структурного кластера, то DAMMIF всегда работает со связной структурой в неограниченном пространстве. Можно сказать, что DAMMIN “собирает” частицу, а DAMMIF ее “выращивает”, работая при этом в 5–10 раз быстрее и делая значительно больше случайных модификаций. Тем не менее постоянная связность графа структуры может ограничивать ее изменчивость в ходе поиска, поэтому для определения разнообразия возможных моделей программу DAMMIF (как и DAMMIN) следует запускать несколько раз с разными параметрами: числом модификаций в цикле постоянной температуры, числом принимаемых модификаций до перехода к следующей температуре, начальной температурой и фактором ее снижения, балансом весов штрафов. Такая работа слишком затратная по времени и требует от пользователя большого опыта решения модельных задач.

В данной работе рассмотрен модифицированный алгоритм DAMMINV, автоматизирующий установку значений параметров поиска. Близкие результаты в некоторых случаях могут быть получены и с использованием DAMMIN и DAMMIF при соответствующей подготовке серии текстовых файлов заданий или повторном запуске программ в диалоговом режиме. Модифицированная версия DAMMINV ориентирована на последовательный поиск 10–15 решений в режиме варьирующейся величины весовых коэффициентов при

штрафных членах (“переключающийся режим”, “alternating mode”). В DAMMINV введен дополнительный штраф за наличие шариков, не находящихся в контакте с основной структурой (наподобие алгоритма DAMMIF, но радиус контакта может быть увеличен для повышения степени варьируемости модели; подробнее ниже). Для моделирования диффузных структур введен также дополнительный штраф за отклонение среднего числа контактов шариков от заданного значения, который заменяет штраф за рыхлость (например, штраф за отклонение числа контактов от 2 фиксирует поиск структуры в виде рыхлого клубка).

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Предлагаемая последовательность шагов поиска формы частицы

1. Выбрать правило взвешивания интенсивности рассеяния W (формула (4)). Для этого можно воспользоваться программой SASPLOT из [6], в которой можно оценить площади левой и правой половин кривой рассеяния. В большинстве случаев степень n в формуле (4) должна обеспечивать 3–10-кратное доминирование площади левой (начальной) половины взвешенной кривой, тем большее, чем шире угловой диапазон данных рассеяния. В DAMMINV выбор делается автоматически.

Важное замечание: при расчете функции парных расстояний при наличии небольшой агрегации частиц не следует искусственно уменьшать D_{\max} во избежание артефактов в решении. Артефакты выражаются в виде расплывания структуры вблизи поверхности области поиска. Вместо этого лучше увеличивать начальный угол в данных, отбрасывая область, искаженную рассеянием от больших образований. В таком случае артефакты в моделях формы часто выглядят как легко различимые отдельные домены или “хвосты”.

2. Для быстрого предварительного поиска модели можно установить больший радиус шарика, чем предлагается по умолчанию, из расчета 3–5 тысяч узлов в области поиска (или, например, удвоенный радиус в DAMMIF). Так как разрешение данных малоуглового рассеяния (МУР) невысокое, детали модели, меньшие чем 1/5–1/10 от максимального диаметра частицы, отражают только наличие малоразмерных внутренних неоднородностей, но не их фактическую структуру. Заметим также, что в программе DAMMIF можно увеличить число успешных модификаций модели (критерий перехода к следующей температуре), по умолчанию их доля составляет 0.1 от максимального числа модификаций при постоянной температуре и ее можно увеличить до 0.3–0.5. Далее будет показано, что в некоторых случаях такое

увеличение приводит к нахождению более адекватной структуры.

3. Стартовую температуру T_{init} и множитель ее снижения T_F при использовании программ DAMMIN и DAMMIF можно не варьировать. Однако слишком большая T_{init} в DAMMIN приводит к большому числу начальных температурных циклов, в которых структура остается в виде случайного диффузного облака без существенного уменьшения невязки. В DAMMINV начальная температура устанавливается равной $T_{\text{init}} = 2.5\sigma_{\Phi(X)}$, где $\sigma_{\Phi(X)}$ — стандартное отклонение изменений целевой функции (2), вычисляемое с помощью таблицы разностей по значениям $\Phi(X)$, полученным путем 100–300 вариаций стартовой модели (обычно это случайным образом заполненная область поиска). Такой выбор T_{init} обеспечивает начало снижения $\Phi(X)$ уже на втором–третьем температурном цикле в ходе поиска. В ходе поиска программа может дополнительно уменьшать температуру в случае, если критерий сходства формы экспериментальной и модельной кривых рассеяния (угол между векторами данных) становится больше 3° . Это предотвращает превращение сформированной к данному моменту модели в полностью случайное расположение шариков.

4. Осуществить поиск решения при $w_R = 1.0$ (уравнение (2)). В DAMMINV для повышения эффективности применен режим, названный SBS (Skip Bulk Solvent), при котором пропускаются модификации, размещающие новый шарик в узел, не контактирующий с шариками текущей модели, если относительное число оторванных шариков в модели становится меньше 3%. Это позволяет программе осуществить большее число вариаций связной структуры без увеличения их максимально допустимого числа и штрафа за наличие отдельно расположенных шариков. Этот режим почти аналогичен реализованному в алгоритме DAMMIF, но радиус контакта может превышать радиус упаковки шариков в сетке узлов (в этом случае штраф за отдельные шарики остается ненулевым). Для обеспечения большей варьированности модели при поиске при включении SBS-режима температуру повышают в 1.5 раза.

5. Перед поиском следующего решения вес w_R уменьшают в 10 раз, назначают T_{init} в 1.5–2 раза меньше, чем на шаге 3. Вес штрафа за смещение центра тяжести частицы от центра области поиска w_C увеличивают в 4 раза. Старт процедуры отжига проводят с полученной модели и получают решение с несколько худшим значением R , но лучшими значениями штрафов. Результат, полученный на шаге 4, обычно не рассматривают в качестве решения, если невязка R значительно ухудшается. Цель этого шага — получение модели, в большей степени удовлетворяющей требованиям штрафов.

6. Веса всех штрафных членов возвращают в первоначальные значения, кроме штрафа за смещение центра тяжести частицы от центра области поиска, который уменьшают до значения 5% от первоначального (это безопасно, так как малая величина смещения уже получена на шаге 4). Устанавливают $T_{\text{init}} = 1.5\sigma_{\Phi(X)}$ или в 1.5–2 раза меньше первоначального значения, чтобы сохранить основные черты модели. Отметим, что минимизация штрафа за рыхлость тела нередко приводит к ориентированию частицы в области поиска так, что некоторые участки ее поверхности выстраиваются вдоль плоскостей упаковки шариков, и форма оказывается граненой. Величина штрафа $w_L P_L$ становится настолько малой, что полученная форма может сохраняться в следующих решениях. Чтобы этого избежать, на этом шаге в DAMMINV предусмотрен поворот тела относительно (фиксированной) области поиска на эйлеровы углы вокруг осей X и Z с соответствующим пересчетом координат шариков, чтобы направления граней не совпадали с плоскостями упаковки пространственных узлов. Так как новые координаты шариков теперь не совпадают с пространственной сеткой, структура частицы пересчитывается на ближайшие узлы. Переходят к шагу 3, на котором программа продолжает поиск решения, которое должно улучшить невязку, это основная задача данной фазы поиска. Итерации продолжают до достижения общего числа циклов поиска с фиксированной температурой 500–600.

Рассмотренная тактика направлена на получение решений, отличающихся друг от друга, и вместе с тем на повышение вероятности найти более глубокий минимум целевой функции. По разбросу полученных структур (обычно 6–10 моделей) с помощью процедур DAMAVER или DAMCLUST можно оценить стабильность решения, как это сделано, например, в [7].

Еще одним отличием программы DAMMINV является учет рассеяния от шариков модели, которое дает существенный вклад в области больших углов. Форм-фактор шариков заменен на рассеяние от эллипсоидов [8] с пропорциями осей 0.71 : 1.0 : 1.4, соответствующими типовым тепловым колебаниям атомов. Это уменьшает влияние на форму кривой рассеяния дифракции от пространственной упаковки элементов, что позволяет более адекватно определять структуры, модели которых состоят из небольшого числа шариков (менее 50–100), без необходимости добавлять или вычитать константу из данных рассеяния. Если такая необходимость существует (например, при неадекватном вычитании рассеяния от кюветы с растворителем), в программе предусмотрена соответствующая опция.

Однозначность определения структуры может быть выражена в терминах дисперсии решений

для конкретной задачи. Информацию о степени разброса моделей дают программы DAMAVER и DAMCLUST. Из-за отсутствия линейной связи между параметрами структуры (координатами шариков) и входными данными рассеяния выражение для дисперсии решений аналитически получить невозможно. Тем не менее можно дать ответ на вопрос о степени однозначности (или устойчивости), проводя анализ величин парных корреляций моделей в серии решений, которые получены методом случайного поиска и, вследствие этого, независимы. Конечно, такой ответ всегда является неполным, и для получения дополнительной информации об однозначности в пакете ATSAS предусмотрены дополнительные процедуры (программа AMBIMETER [6, 9]).

В общем случае сложно различить неоднозначность решения задачи определения структурной модели и его неустойчивость. И то, и другое приводит к тому, что находимые модели частиц отличаются друг от друга, если варьировать угловой диапазон данных рассеяния, проводить поиск с разных стартовых приближений, изменять параметры самого метода минимизации. Для получения окончательного ответа необходимо проводить серию численных экспериментов с варьированием условий поиска, анализировать отдельные решения с точки зрения “физического смысла”, отбирать группы похожих решений (например, по R -фактору в прямом пространстве) и т.п. В итоге анализ данных МУР оказывается довольно сложной экспертной задачей, хотя нередко и удается получить окончательный ответ без проведения большого количества численных экспериментов. В данной публикации не рассматриваются все влияющие на решение факторы, так как их последовательный анализ может быть предметом не одной статьи.

Примеры экспериментальных данных брали из базы моделей белковых частиц, определенных по данным МУР [<https://www.sasbdb.org/data>]. Образцы отбирали по линейности графика Гинье и не слишком большому уровню шума.

Пример неоднозначности модели частицы. Факт неоднозначности отмечали многие исследователи, однако в литературе до сих пор не было примеров двух разных по структуре тел, МУР от которых было бы идентично. Идентичность кривых рассеяния означает, что телам должны соответствовать одинаковые функции парных расстояний $p(r)$ (1), из чего следует идентичность максимальных размеров D_{\max} , объемов V и радиусов инерции R_g [8]. Это ограничивает число возможных форм частиц. В литературе существует пример расположения пяти точек на плоскости двумя разными способами при идентичных наборах парных расстояний между вершинами графов и дополнительно идентичности наборов площадей

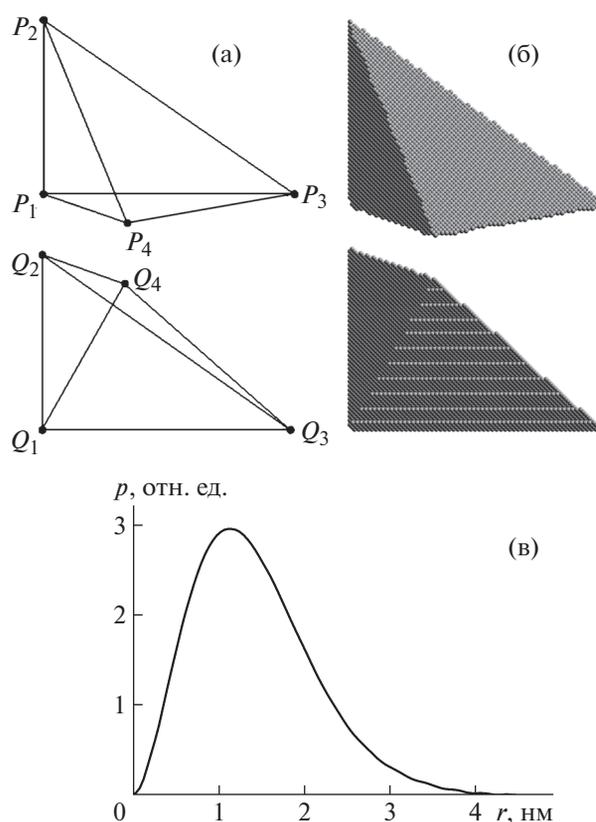


Рис. 1. Пример двух пирамидальных тел одинаковой высоты относительно плоскости рисунка (точки P_4 и Q_4 вынесены из плоскости), обладающих идентичными функциями парных расстояний $p(r)$: а – опубликованная плоская модель [10]; б – объемные шариковые модели, составленные примерно из 20 000 шариков; в – вид функций парных расстояний, рассчитанных по шариковым моделям (совпадают с точностью до толщины линии на графике).

треугольников [10]. Эксперимент с опубликованным примером показал, что, если для обеих структур вывести центральную точку из плоскости на одинаковые расстояния (что обеспечивает равенство объемов полученных пирамид) и, полагая тела однородными по плотности, численно рассчитать функции расстояний $p(r)$ с уменьшающимся радиусом шариков, можно увидеть, что кривые распределений стремятся к совпадению. Формы этих моделей показаны на рис. 1.

Влияние параметра взвешивания интенсивности рассеяния. На рис. 2 представлены кривые рассеяния от частицы рис. 3 (1), взвешенные согласно формуле (4). Модель на рис. 3 (2) получена с помощью программы DAMMIN с установками по умолчанию с весом $W = s^2$. Тот же вес для DAMMINV, несмотря на перемежающийся режим, не привел к удовлетворительному восстановлению формы (рис. 3 (3)), тогда как $W = s^1$ (рис. 2 (2)) позволяет найти близкую к точной мо-

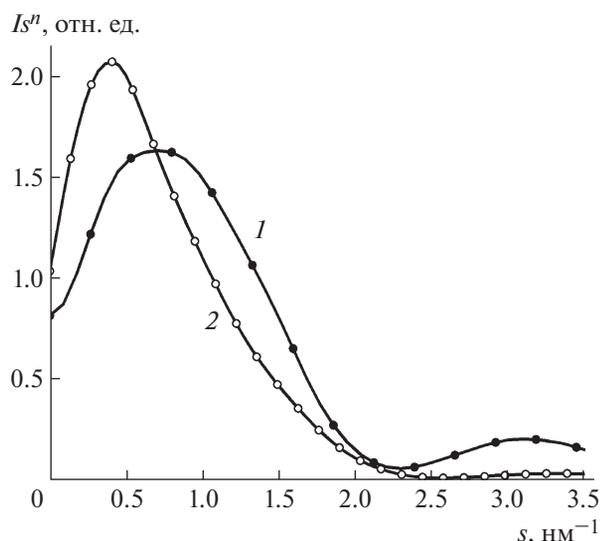


Рис. 2. Взвешенные интенсивности рассеяния, используемые для расчета невязки R . Точки – теоретическое рассеяние от частицы, представленной на рис. 3 (1), сплошная линия – типичная кривая рассеяния от найденных моделей, показанных на рис. 3 (2–4). На графике: 1 – весовая функция $W = s^2$, 2 – $W = s^1$.

дели форму (рис. 3 (4)). Значения критерия подгонки R рассчитаны для невзвешенных данных.

Поиск формы частиц в перемежающемся режиме. Другой пример, демонстрирующий работу поиска с периодическим изменением относительных вкладов членов целевой функции, представлен на рис. 4, 5. В случае программы DAMMINV применяли вес $W = s^1$. Остальные параметры поиска оставляли по умолчанию. Во всех случаях значение R составляло $(1 \pm 0.2) \times 10^{-4}$. Видно, что одиночные поиски DAMMIN и DAMMIF с установками по умолчанию и DAMMINV (формы 4–6 на рис. 5) не приводят к сколько-нибудь приемлемому результату, несмотря на хорошее соответствие кривых рассеяния (рис. 4). Последовательный поиск в перемежающемся режиме на пятой итерации дает приемлемое восстановление формы, сохраняющееся на следующих итерациях. Как отмечено ранее, увеличение числа принятых при фиксированной температуре модификаций относительно полного числа испытаний до 0.5 (100 000 вместо 20 000) в данном случае привело к нахождению адекватной модели во всех трех случаях.

Формы молекулы бычьего альбумина как пример анализа данных с частичной агрегацией белка. Приведем результаты анализа типичного сложного случая изучения белка в растворе. На рис. 6а показаны кривые МУР от раствора альбумина. Расчет функции парных расстояний показал, что распределение имеет протяженное плечо в обла-

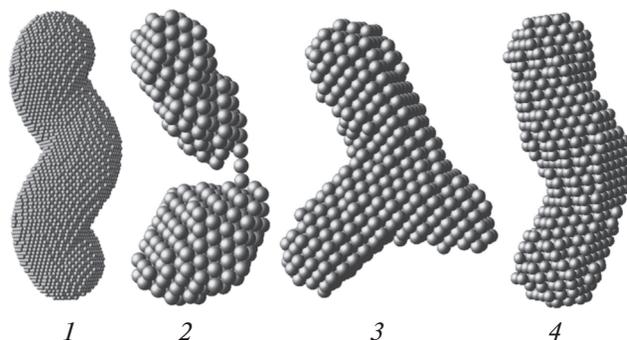


Рис. 3. Типичные результаты восстановления формы частицы (1) по теоретическим данным рассеяния, показанным на рис. 2: 2 – программа DAMMIN ($R = 1.3 \times 10^{-3}$, $W = s^2$), 3 – DAMMINV ($R = 0.73 \times 10^{-4}$, $W = s^2$, перемежающийся режим), 4 – DAMMINV ($R = 0.71 \times 10^{-4}$, $W = s^1$, перемежающийся режим).

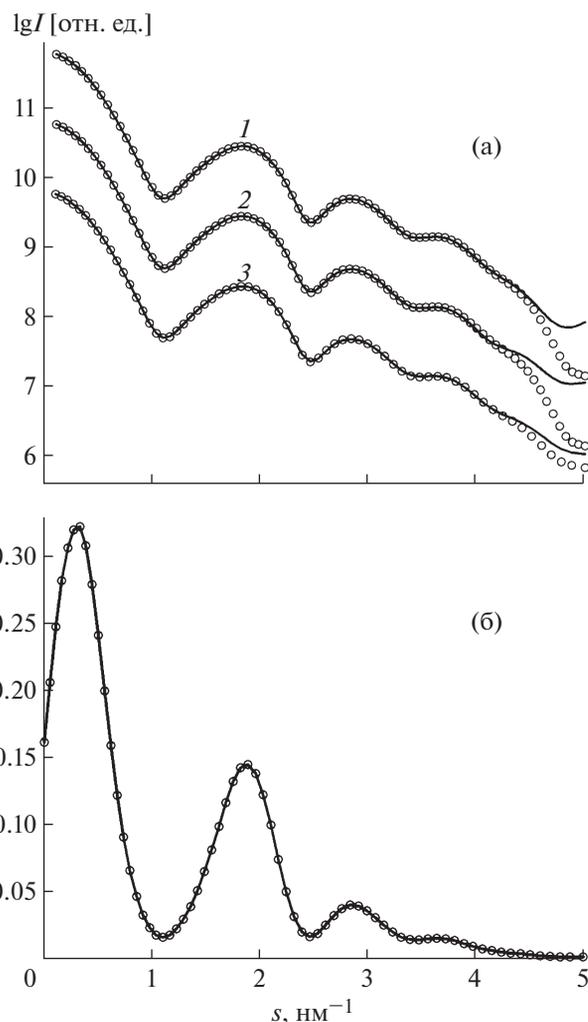


Рис. 4. Результаты подгонки модельной интенсивности рассеяния (точки), рассчитанной от структуры 1 на рис. 5. Сплошными линиями показано рассеяние от найденных структур: 1 – DAMMIN, 2 – DAMMIF, 3 – DAMMINV (структура 6 на рис. 5) (а). Взвешенная нормированная интенсивность рассеяния для расчета невязки R (б).

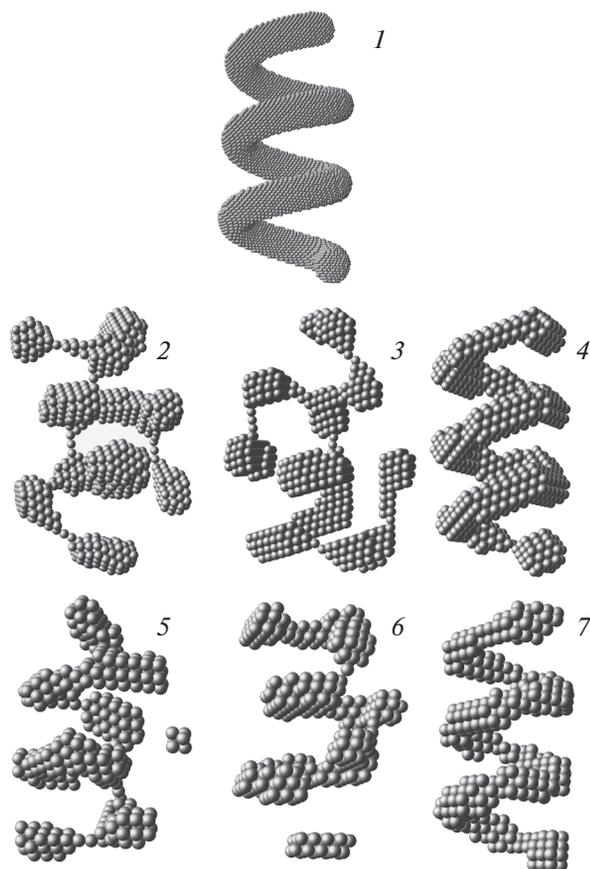


Рис. 5. Результаты моделирования исходной структуры (1): программа DAMMIN (2); DAMMIF с установками по умолчанию (3), DAMMIF с увеличенным до 100 000 числом принятых при фиксированной температуре модификаций (4); DAMMINV – 5 – после первой итерации, 6 – после третьей, 7 – после пятой. Нечетные номера итераций соответствуют шагу 4 перемежающейся моды, который соответствует исходному (большому) относительному вкладу члена R в целевой функции.

сти больших, более 9 нм, длин отрезков (рис. 7), что свидетельствует о частичной агрегации белка в растворе, хотя область Гинье прямолинейна (рис. 6б) и об этом не свидетельствует. Соответствующие радиусы инерции, найденные по области Гинье и по функции парных расстояний, равны соответственно 3.25 ± 0.02 и 3.26 ± 0.04 нм. Их близость также не свидетельствует об агрегации, хотя радиус инерции, рассчитанный в программе CRY SOL [11] по кристаллической структуре белка, даже с учетом гидратной оболочки меньше: 2.84–2.89 нм. Действительно, типичные модели формы молекулы (рис. 8), найденные в перемежающемся режиме поиска программой DAMMINV, демонстрируют наличие артефактов, искусственно увеличивающих длину максимальной хорды в структуре. Формы молекул определяли с использованием взвешивания $W = s^1$, так как взвешива-

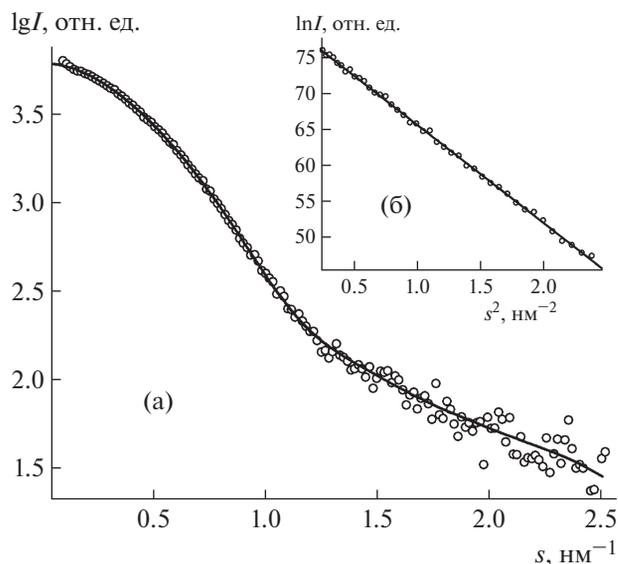


Рис. 6. Экспериментальная интенсивность рассеяния от раствора бычьего альбумина (точки) и типичная модельная кривая (линия) от структур, приведенных на рис. 8 (а); график Гинье для оценки радиуса инерции R_g (б).

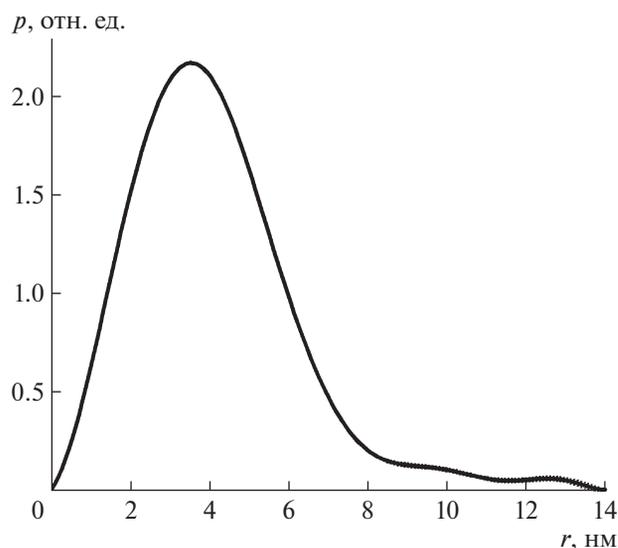


Рис. 7. Функция парных расстояний, рассчитанная по данным рис. 6а.

ние $W = s^2$ слишком сильно увеличивает интенсивность рассеяния на больших углах и результат оказывается несколько хуже. Так как артефакты формы возникают в решениях в разных местах структуры, усредненная с помощью DAMAVER структура не показывает хорошего совпадения с экспериментом ($\chi^2 = 10.4$) и здесь не приводится. Основной результат данного исследования – варианты формы макромолекулы после отбрасыва-

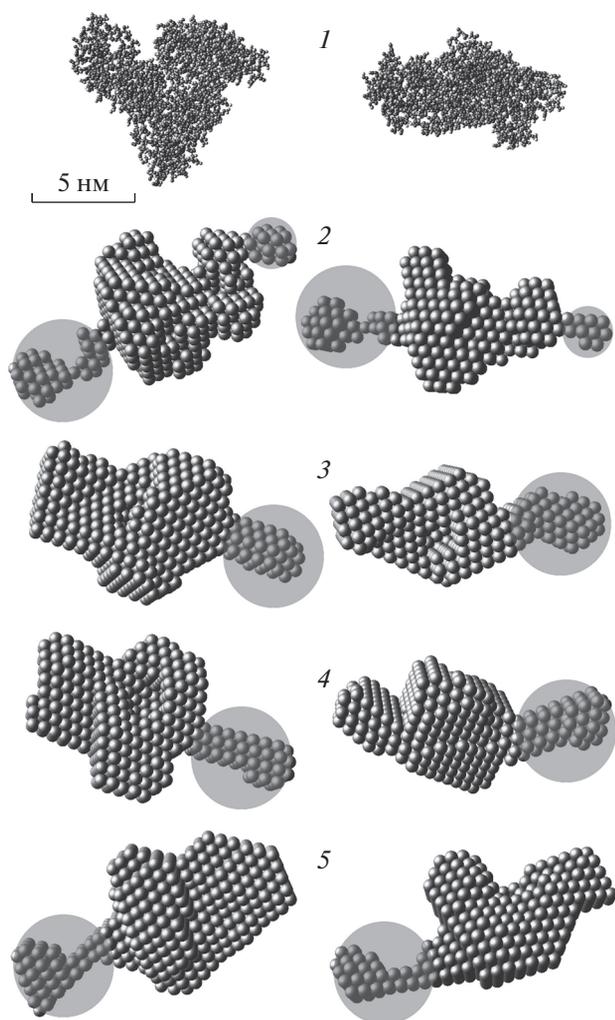


Рис. 8. Кристаллическая структура молекулы бычьего альбумина из белковой базы данных 3V03 (1) и последовательные решения DAMMINV после шага 3 (2–5). Кругами помечены артефакты из-за агрегации (предположительно – частичной агрегации молекул в растворе).

ния артефактов агрегации, помеченных на рис. 8 серыми кругами. Эти варианты в силу предлагаемой тактики поиска должны отличаться друг от друга, охватывая круг возможных форм.

Определение формы молекул по опубликованным данным. Данные для анализа взяты из базы данных МУР от белковых растворов SASBDB [12]. Данные рассеяния и ссылки на публикации не приведены, так как их можно найти в базе по коду образца. По данным рассеяния рассчитаны распределения по расстояниям $p(r)$ без принудительного уменьшения максимального расстояния D_{\max} . Теоретические кривые рассеяния от полученных моделей формы описывают данные с точностью не хуже $\chi^2 = 1.12$ и не представляют интереса для

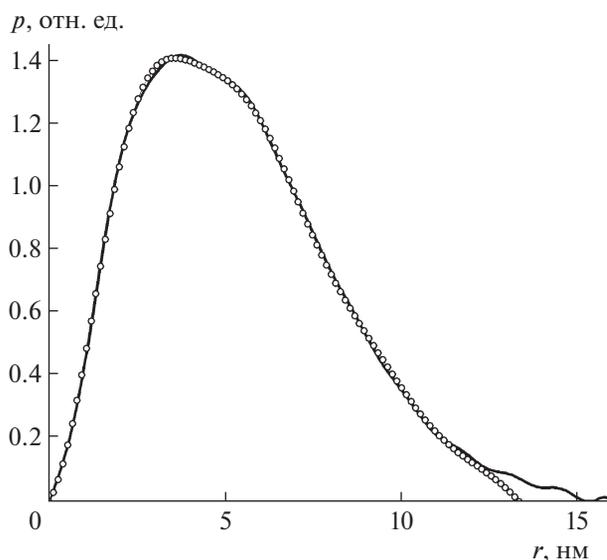


Рис. 9. Функции парных расстояний для образца FRET биосенсора, код данных в базе SASDAF5. Точки – опубликованные данные, линия – рассчитанное распределение, которое было использовано в качестве входных данных для программы DAMMINV.

обсуждения. Поэтому приводим только наборы полученных структур.

На рис. 9 показаны опубликованная и рассчитанная кривые распределения по расстояниям для структуры SASDAF5 (FRET биосенсора), рис. 10 представляет опубликованные и рассчитанные структуры. Видно, что расчет с использованием расширенной $p(r)$ предоставляет структуры, более близкие по своему характеру к молекулярной модели, что может служить ее подтверждением. Пустоты малого размера в найденных моделях, как было сказано ранее, не являются деталями структуры, а отражают только попытку описать разреженность структуры в прилегающей области. Анализ структур с помощью DAMCLUST показал небольшой разброс моделей, и на рис. 10 представлены наиболее различающиеся.

ЗАКЛЮЧЕНИЕ

Представлены некоторые результаты численных экспериментов по определению формы частиц из данных МУР от разбавленных изотропных монодисперсных систем и предложена схема последовательных расчетов с помощью программы моделирования шариковыми структурами. Весь спектр возможных случаев охватить в одной работе невозможно, но главным результатом проведенного исследования служит утверждение: предварительное моделирование с использованием искусственных структур и последовательный поиск решений с варьированием параметров алгоритма поиска могут помочь ответить на во-

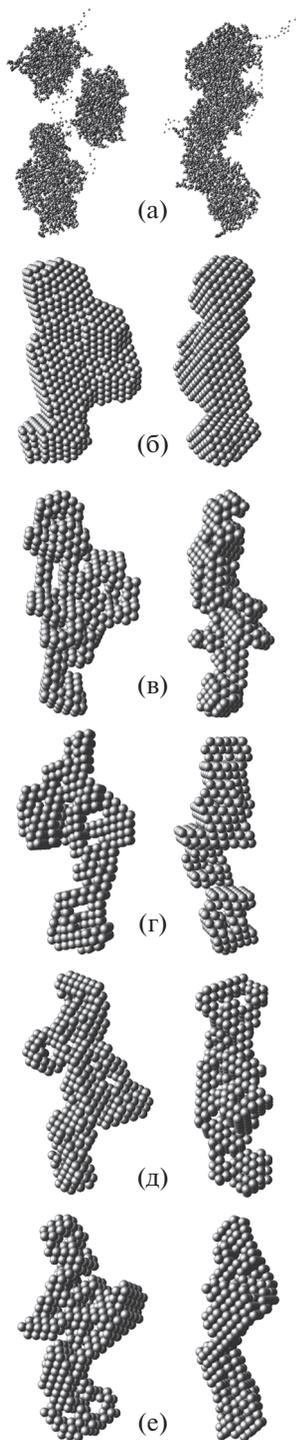


Рис. 10. Модели частицы FRET биосенсора (код SASDAF5): а, б – атомарная и шариковая опубликованные модели; в–е – найденные с помощью DAMMINV по данным рис. 9. Правый столбец – структуры повернуты на 90° относительно вертикальной оси.

прос о надежности полученных решений в исследовании белкового объекта.

Развитие методологии анализа данных МУР выполняется в тесном сотрудничестве с группой

Д.И. Свергуна BIOSAXS [<https://www.embl-hamburg.de/biosaxs/>] в Европейской лаборатории молекулярной биологии (EMBL с/о DESY, Hamburg), которому автор выражает глубокую благодарность.

Работа выполнена при финансовой поддержке Российского научного фонда (грант № 19-14-00244) и Министерства науки и высшего образования в рамках выполнения работ по Государственному заданию ФНИЦ “Кристаллография и фотоника” РАН в части развития библиотеки подпрограмм общего назначения.

СПИСОК ЛИТЕРАТУРЫ

1. *Svergun D.I.* // *Biophys. J.* 1999. V. 76. P. 2879. [https://doi.org/10.1016/S0006-3495\(99\)77443-6](https://doi.org/10.1016/S0006-3495(99)77443-6)
2. *Franke D., Svergun D.I.* // *J. Appl. Cryst.* 2009. V. 42. P. 342. <https://doi.org/10.1107/S0021889809000338>
3. *Chacon P., Moran F., Díaz J.F. et al.* // *Biophys. J.* 1998. V. 74. № 6. P. 2760. [https://doi.org/10.1016/S0006-3495\(98\)77984-6](https://doi.org/10.1016/S0006-3495(98)77984-6)
4. *Kirkpatrick S., Gelatt C.D., Vecchi M.P.* // *Science.* 1983. V. 220. P. 671. <https://doi.org/10.1126/science.220.4598.671>
5. *Semenyuk A.V., Svergun D.I.* // *J. Appl. Cryst.* 1991. V. 24. P. 537. <https://doi.org/10.1107/S002188989100081X>
6. *Manalastas-Cantos K., Konarev P.V., Hajizadeh N.R. et al.* // *J. Appl. Cryst.* 2021. V. 54. P. 1. <https://doi.org/10.1107/S1600576720015368>
7. *Volkov V.V., Svergun D.I.* // *J. Appl. Cryst.* 2003. V. 36. P. 860. <https://doi.org/10.1107/S0021889803000268>
8. *Свергун Д.И., Фейгин Л.А.* Рентгеновское и нейтронное малоугловое рассеяние. М.: Наука, 1986. 280 с.
9. *Petoukhov M.V., Svergun D.I.* // *Acta Cryst. D.* 2015. V. 71. P. 1051. <https://doi.org/10.1107/S1399004715002576>
10. *Boutin M., Kemper G.* // *arXiv:math/0304192 [math.AC].* 2003. P. 1. <https://arxiv.org/abs/math/0304192>
11. *Svergun D.I., Barberato C., Koch M.H.J.* // *J. Appl. Cryst.* 1995. V. 28. P. 768. <https://doi.org/10.1107/S0021889895007047>
12. Small Angle Scattering Biological Data Bank SASBDB // <https://www.sasbdb.org/>

СТРУКТУРА МАКРОМОЛЕКУЛЯРНЫХ
СОЕДИНЕНИЙ

УДК 548.73

ОЛИГОМЕРИЗАЦИЯ ЦЕРУЛОПЛАЗМИНА
ПОД ДЕЙСТВИЕМ ИОНОВ МЕДИ

© 2021 г. М. В. Петухов^{1,2}, А. В. Соколов³, В. А. Костевич³, В. Р. Самыгина^{1,4,*}

¹ Институт кристаллографии им. А.В. Шубникова ФНИЦ “Кристаллография и фотоника” РАН, Москва, Россия

² Институт физической химии и электрохимии им. А.Н. Фрумкина РАН, Москва, Россия

³ Институт экспериментальной медицины, Санкт-Петербург, Россия

⁴ Национальный исследовательский центр “Курчатовский институт”, Москва, Россия

*E-mail: lera@crys.ras.ru

Поступила в редакцию 22.04.2020 г.

После доработки 02.05.2020 г.

Принята к публикации 27.05.2020 г.

Методом малоуглового рентгеновского рассеяния исследована олигомеризация медьсодержащей ферроксидазы церулоплазмина при наличии не связанной с белком меди в растворе. Церулоплазмин является мономером с молекулярной массой ~132 кДа. Добавление хлорида меди в концентрации 10 мкМ приводит к димеризации 60% белка. Анализ межсубъединичных контактов показал, что димеризация может препятствовать образованию белок-белковых комплексов церулоплазмина с белками лейкоцитов миелопероксидазой и пероксидазой эозинофилов. Образование этих комплексов является частью защитной реакции организма на воспалительные процессы. Также, вероятно, димерный церулоплазмин теряет способность связывать и окислять один из субстратов – *p*-фенилендиамин.

DOI: 10.31857/S0023476121050179

ВВЕДЕНИЕ

Олигомеризация белков может являться причиной некоторых нейродегенеративных заболеваний, таких как болезнь Альцгеймера, Паркинсона и других [1, 2]. Триггерами олигомеризации могут служить ионы металлов. Индуцированная металлами агрегация может встречаться как у частично разупорядоченных белков, так и у глобулярных. Например, ионы хрома, кобальта и никеля способствуют агрегации альбумина [3]. Часто встречаются данные, свидетельствующие о том, что агрегацию белков способны вызывать ионы цинка и меди [4–7]. Медь присутствует в организме человека в концентрации 16–20 мкМ в крови, 15 мкМ в синаптической щели и 0.5–2.5 мкМ в спинномозговой жидкости [8]. Ее концентрация в спинномозговой жидкости повышена в 2.2 раза у пациентов с болезнью Альцгеймера [9]. При болезни Альцгеймера концентрация так называемой свободной (диализуемой) меди в плазме крови достигает 5 мкМ [10].

Церулоплазмин (ЦП) – медьсодержащая многофункциональная гликозилированная ферроксидаза с молекулярной массой (ММ) ~132 кДа, циркулирующая в растворимой форме в плазме крови. Благодаря способности превращать токсичное двухвалентное железо в трехвалентное он

является природным антиоксидантом. ЦП содержит от 40 до 70% меди плазмы крови [11]. Большинство исследований такими методами, как гель-фильтрация, электрофорез в полиакриламидном геле в неденатурирующих условиях и малоугловое рентгеновское рассеяние (МУРР), показывают, что ЦП млекопитающих является мономером [12–15]. Ранее в литературе была описана так называемая димерная изоформа ЦП с молекулярной массой ~200 кДа [16]. Димер ЦП был обнаружен в плазме крови здоровых доноров и секретировался культурой гепатоцитов (HepG2). Опыты, в которых при кросс-сшивке мономерного ЦП с помощью дитиобиссукцинимидилпропионата образовался продукт с ММ ~200 кДа, а при ограниченном протеолизе мономерного и димерного ЦП образовались идентичные фрагменты, позволили заключить, что обнаруженная ими изоформа ЦП является димером, а не продуктом посттрансляционной модификации или альтернативного сплайсинга [16]. Полипептидная цепь белка свернута в глобулу, состоящую из шести доменов, расположенных в форме псевдогексамера, соединенных длинными петлями. Благодаря ионам меди в структуре ЦП помимо ферроксидазной активности обладает свойствами оксидазы, NO-оксидазы, NO₂-синтазы, су-

Таблица 1. Структурные параметры ЦП, полученные по данным МУРР

	R_g , нм	D_{\max} , нм	V_p , нм ³	ММ, кДа	χ^2	Объемные доли мономер : димер
Эксперимент	4.2 ± 0.1	15 ± 1	305 ± 10	190 ± 20	4.03	$(0.4 \pm 0.01) : (0.6 \pm 0.01)$
Мономер ЦП	2.9	9.0	195	129.1*		
Димер ЦП	4.2	15	390	258.2**		

* [28].

** Значение получено исходя из молекулярной массы (ММ) мономера в [28].

пероксиддисмутазы. Структурные исследования показывают, что ЦП человека содержит шесть–восемь центров связывания меди, из которых шесть основных сайтов всегда заполнены, а заполнение двух лабильных сайтов наблюдается не всегда [13]. ЦП крысы содержит на один лабильный сайт больше [17]. Обнаружено, что при избытке меди ЦП подвержен олигомеризации. Было исследовано олигомерное состояние ЦП в растворе методом МУРР в присутствии хлорида меди и обнаружена частичная димеризация белка. В результате была получена трехмерная модель димера низкого разрешения, определены зоны контактов мономер–мономер.

МАТЕРИАЛЫ И МЕТОДЫ

Выделение и очистка церулоплазмينا. Для получения препарата мономерного ЦП с $A_{610}/A_{280} > 0.049$ плазму крови с добавлением 1 мМ этилендиаминтетрауксусной кислоты и 0.1 мМ фенолметилсульфонил фторида (PMSF) подвергали хроматографии на UNOsphere Q и неомидин-агарозе [18]. Далее ЦП концентрировали до 80 мг/мл с помощью центрифужной ячейки Vivaspin 20 ($M > 100$ кДа), трижды заменяя 0.1 М Hepes-NaOH pH 7.5.

Малоугловое рентгеновское рассеяние. Предварительные измерения малоуглового рассеяния ЦП были проведены на станции БиоМУР Курчатовского центра синхротронного излучения [19]. Финальные измерения МУРР были проведены на станции EMBL P12 BioSAXS (кольцо PETRA-III синхротрона DESY, Гамбург, Германия) [20] в диапазоне векторов рассеяния $0.023 < s < 5.0$ нм⁻¹ ($s = 4\pi \sin \theta / \lambda$, где 2θ – угол рассеяния и $\lambda = 0.124$ нм – длина волны). Линия P12 оснащена оборудованием для автоматической смены образцов и двумерным детектором Pilatus 2M (фирма DECTRIS, Швейцария). Время экспозиции одного образца – 20 кадров по 0.05 с каждый. При сравнении временных кадров радиационного повреждения зафиксировано не было. Данные были собраны при концентрации ЦП 1.5–7.0 мг/мл (11–53 мкМ) как в присутствии 10 мкМ CuCl₂, так и без добавления соли меди. Первичную обработку экспериментальных данных проводили в соот-

ветствии со стандартными процедурами [21]. В частности, усреднение сигнала от буфера и его вычитание осуществляли с использованием программы PRIMUS [22]. Рассеяние в нулевой угол $I(0)$ и радиус инерции R_g вычислены с помощью приближения Гинье [23]. Программу обратного фурье-преобразования GNOM [24] использовали для оценки максимального размера частиц D_{\max} . Эффективные ММ частиц в растворе получены путем сравнения с рассеянием на эталонном образце бычьего сывороточного альбумина (66 кДа). В качестве дополнительного контроля ММ использовали исключенный (породовский) объем V_p [25, 26].

Олигомерное равновесие. Альтернативные модели димера ЦП, полученные с помощью докинга, применяли для приближения экспериментальных данных образца, содержащего мономеры и димеры. Анализ смесей мономер–димер осуществляли по программе OLIGOMER [22], которая приближает данные рассеяния линейной комбинацией кривых рассеяния компонентов смеси. Варьируемыми параметрами являются объемные доли компонентов, а приближение осуществляется путем минимизации невязки χ^2 между экспериментальными данными $I_{\text{exp}}(s)$ и рассчитанной кривой $I_{\text{calc}}(s)$.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Экспериментальная кривая МУРР для раствора ЦП с добавлением 10 мкМ хлорида меди показана на рис. 1. Соответствующие общие параметры растворенных частиц приведены в табл. 1. В табл. 1 также приведены теоретические значения R_g , D_{\max} , V_p для кристаллографической модели и ММ мономера по данным масс-спектрометрии [27], поскольку этот метод намного точнее определяет ММ белка, чем теоретические расчеты по кристаллографической модели (118 кДа) или аминокислотной последовательности ЦП (122.2 кДа) из-за гликозилирования, типичного для белков млекопитающих [27]. Полученные экспериментально для образца церулоплазмينا ММ и V_p свидетельствуют о том, что в растворе находятся как мономеры, так и димеры ЦП в со-

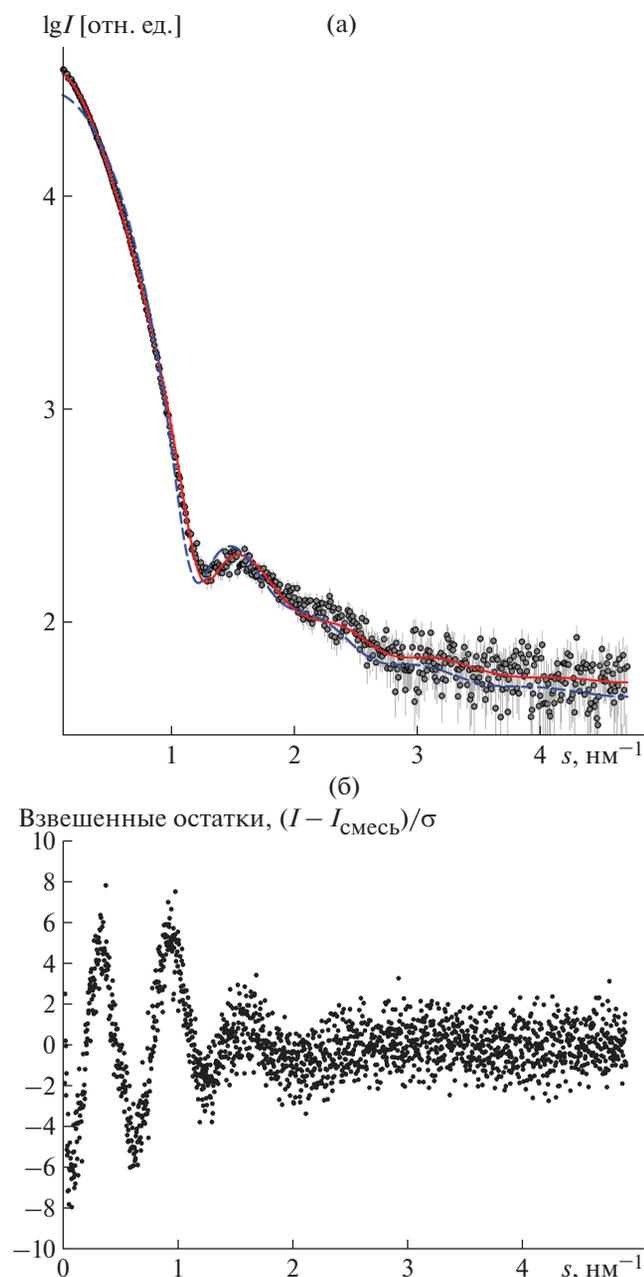


Рис. 1. Кривые МУРР раствора церулоплазмина (а): кружки — экспериментальные данные, штриховая линия — приближение кристаллографической моделью мономера, сплошная линия — приближение смесью мономера и димера в соотношении 40 : 60%; распределение взвешенных аминокислотных остатков в случае приближения смесью $(I(s_i) - I_{\text{смесь}}(s_i)) / \sigma(s_i)$ (б).

отношении 40 : 60%. Действительно, кривая МУРР, рассчитанная для мономера ЦП, систематически отличается от экспериментальных данных в области малых и средних углов. Олигомерное равновесие в смеси было охарактеризовано с помощью линейной комбинации кривых рассеяния, рассчитанных для кристаллографической модели

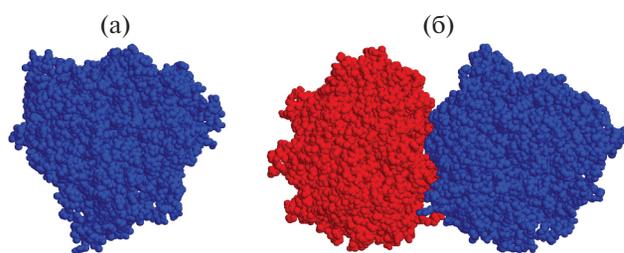


Рис. 2. Кристаллографическая модель мономера церулоплазмина (а) и оптимальная модель димера, полученная докингом (б).

ЦП (код pdb 4enz) и моделей димера, построенных с использованием сервера ClusPro [28]. Всего было рассмотрено двадцать возможных вариантов димеризации ЦП. При помощи программы OLIGOMER и докинга найден оптимальный вариант димера, а также определены объемные доли мономера и димера, которые совпали со значениями, полученными из анализа ММ и V_p . На рис. 2, 3 представлена оптимальная модель димера, а на рис. 1 — приближение экспериментальных данных МУРР смесью 40% кристаллографического мономера с 60% оптимального варианта димера. Несколько высокое значение χ^2 , видимо, связано с недооцененными ошибками эксперимента. Контакт мономеров расположен далеко от сайтов связывания меди, как основных, так и лабильных.

Анализ областей димеризации позволил приблизительно определить основные зоны контактов. Петля 698–710 субъединицы *A* взаимодействует с доменом III субъединицы *B*. С этим доменом взаимодействует, вероятно, и Trp669 субъединицы *A* (рис. 4). Trp669 и Ser662 относятся к сайту связывания *p*-фенилендиамина. Этот аро-

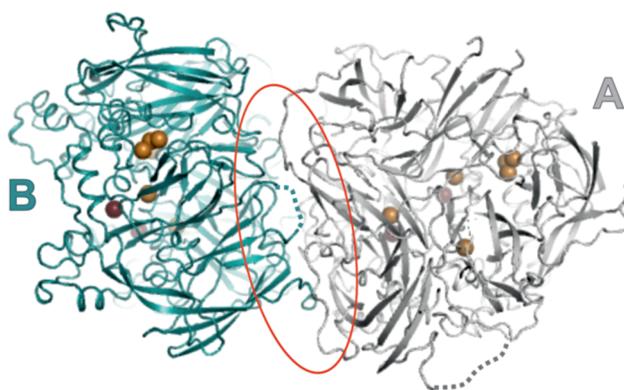


Рис. 3. Модель димера церулоплазмина: ионы меди изображены сферами, темным цветом показаны ионы в лабильных сайтах связывания меди, зона контакта мономер–мономер выделена овалом, *A* и *B* — субъединицы.

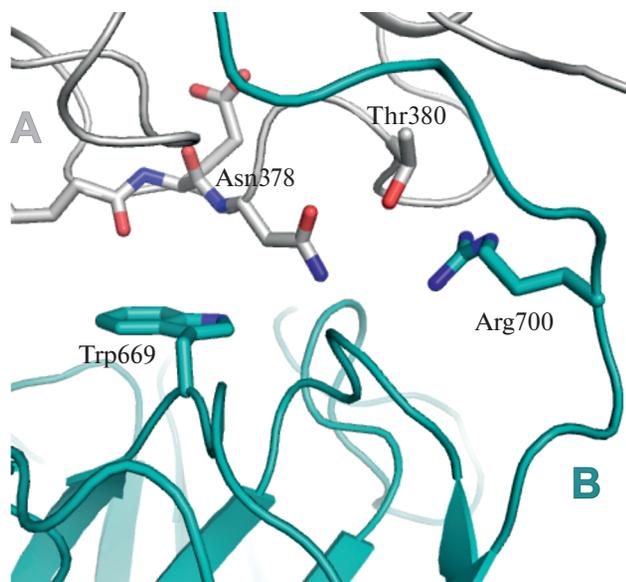


Рис. 4. Внутримерный контакт в области *p*-фенилендиаминсвязывающего сайта.

матический амин является классическим субстратом для измерения оксидазной активности ЦП [29, 30]. Следовательно, образование димера может блокировать связывание этого субстрата и, соответственно, его окисление.

Другой контакт расположен в области петли 885–890, соединяющей домены V и VI субъединицы A с доменом IV субъединицы B. Способность ионов металлов вызывать агрегацию конкретного белка зависит от многих параметров. В их число входят тип металла и белка, внутриклеточная концентрация ионов металла, структура белка, содержание и взаимное расположение

потенциальных лигандирующих групп в боковых цепях аминокислот. К таким группам относятся тиоловая, имидазольная и карбоксилатная [31]. Предпочтительным кандидатом на роль лиганда ионов меди является гистидин [32]. Используемый метод не позволяет точно определить сайты связывания металла, тем более что в зонах контактов расположены преимущественно лабильные междоменные петли, которые могут иметь различные конформации. Конформацию одного из участков ЦП, а.а 885–890, предположительно вовлеченного в межмолекулярное взаимодействие, ни разу не удалось установить при рентгеноструктурном анализе [13].

Аминокислотные последовательности областей зон межмолекулярных контактов не содержат гистидины. Однако потенциальный сайт связывания иона меди с типичными лигандами может быть сформирован боковыми цепями Glu704 (субъединицы B) и Met580 (субъединицы A), расположенными на α -спирали (рис. 5a). Другой потенциальный сайт включает в себя Glu 419 (субъединицы A) и Glu712 (субъединицы B) (рис. 5б). Аргинин также является предпочтительным лигандом для Cu^{2+} [32], и нельзя исключить, что Arg701 (субъединицы B) (рис. 5a) может быть вовлечен в связывание ионов металла, как и соседний Arg700 (рис. 4). Потенциальная роль димеризации ЦП под действием ионов меди вытекает из анализа областей олигомеризации. Петля 885–890 участвует в обнаруженных *in vivo* белок-белковых взаимодействиях ЦП [13], ингибируя в них миелопероксидазу и пероксидазу эозинофилов [33]. Таким образом, ЦП предотвращает синтез хлорноватистой кислоты (НОС) и окисление ряда ароматических субстратов. НОС является мощным антибактериальным агентом и частью

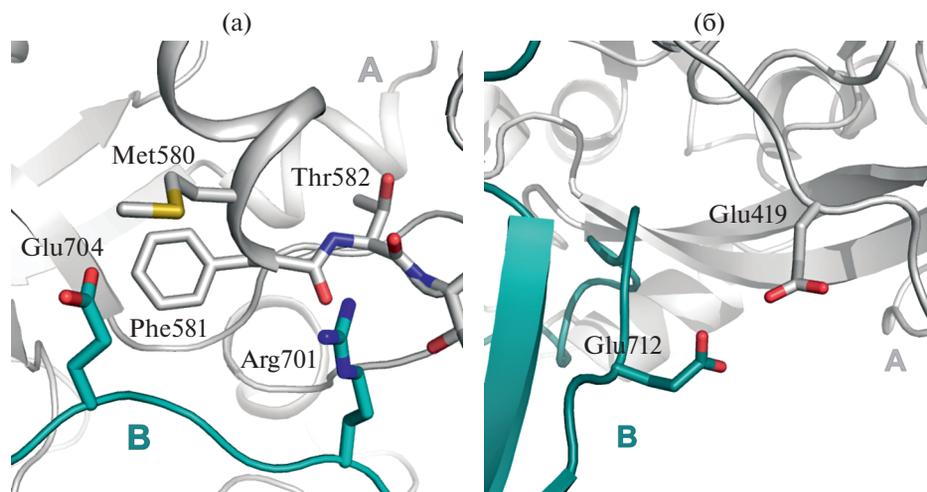


Рис. 5. Потенциальные сайты связывания ионов свободной меди: а — контакт в области Met580, б — контакт в области Glu419.

врожденного иммунного ответа. Однако НОСИ повреждает все виды биополимеров, и ее избыток может приводить к галогенирующему стрессу, сопровождающему хронические воспалительные заболевания, в том числе атеросклероз [13, 34]. Расположение в области димеризации петли 885–890 экранирует ее и, предположительно, будет мешать образованию упомянутых выше белок-белковых комплексов, лишая ЦП ингибирующего действия на миелопероксидазу и пероксидазу эозинофилов. Таким образом, в результате действия избытка ионов меди эффективность реализации части функций ЦП может существенно снижаться.

Работа выполнена при финансовой поддержке Совета по грантам Президента РФ (грант № МД-1901.2020.4) в части выделения и очистки ЦП, Министерства науки и высшего образования РФ в рамках Государственного задания ФНИЦ “Кристаллография и фотоника” РАН в части обработки малоугловых данных и при поддержке НИЦ “Курчатовский институт” (Приказ № 1360) в части измерения экспериментальных кривых МУРР и моделирования димера ЦП.

СПИСОК ЛИТЕРАТУРЫ

1. *Chiti F., Dobson C.M.* // *Annu. Rev. Biochem.* 2006. V. 75. P. 333. <https://doi.org/10.1146/annurev.biochem.75.101304.123901>
2. *Soto C.* // *Nat. Rev. Neurosci.* 2003. V. 4. P. 49. <https://doi.org/10.1038/nrn1007>
3. *Hedberga X.Y.S., Dobryden I., Chaudhary H. et al.* // *Colloids Surf. B: Biointerfaces.* 2019. V. 173. P. 751. <https://doi.org/10.1016/j.colsurfb.2018.10.061>
4. *Brown D.R., Qin K., Herms J.W. et al.* // *Nature.* 1997. V. 390. P. 684. <https://doi.org/10.1038/37783>
5. *Atwood C.S., Moir R.D., Huang X. et al.* // *J. Biol. Chem.* 1998. V. 273. P. 12817. <https://doi.org/10.1074/jbc.273.21.12817>
6. *Paik S.R., Shin H.J., Lee J.H. et al.* // *Biochem. J.* 1999. V. 340. P. 821.
7. *Morgan C.J., Gelfand M., Atreya C. et al.* // *J. Mol. Biol.* 2001. V. 309. P. 339. <https://doi.org/10.1006/jmbi.2001.4661>
8. *Brown D.R., Qin K., Herms et al.* // *Nature (London).* 1997. V. 390. P. 684. <https://doi.org/10.1038/37783>
9. *Multhaup G., Schlicksupp A., Hesse et al.* // *Science.* 1996. V. 271. P. 1406. <https://doi.org/10.1126/science.271.5254.1406>
10. *Squitti R., Quattrocchi C.C., Salustri C. et al.* // *Prion.* 2008. V. 2. P. 23. <https://doi.org/10.4161/pri.2.1.6297>
11. *Lindely M.C.* // *Metallomics.* 2016. V. 8. P. 887. <https://doi.org/10.1039/C6MT00103C>
12. *Sabatucci A., Vachette P., Vasilyev V.B. et al.* // *J. Mol. Biol.* 2007. V. 371. P. 1038. <https://doi.org/10.1016/j.jmb.2007.05.089>
13. *Samygina V.R., Sokolov A.V., Bourenkov G. et al.* // *PlosOne.* 2013. V. 8. P. e67145. <https://doi.org/10.1371/journal.pone.0067145>
14. *Sokolov A.V., Dadinova L.A., Petoukhov M.V. et al.* // *Biochem. (Moscow).* 2018. V. 83. P. 701. <https://doi.org/10.1134/S000629791806007X>
15. *Samygina V.R., Sokolov A.V., Pulina M.O. et al.* // *Crystallography Reports.* 2008. V. 53. P. 655. <https://doi.org/10.1134/S1063774508040172>
16. *Sato M., Schilsky M.L., Stockert R.J. et al.* // *J. Biol. Chem.* 1990. V. 265. P. 2533.
17. *Samygina V.R., Sokolov A.V., Bourenkov G. et al.* // *Metallomics.* 2017. V. 9. P. 1828. <https://doi.org/10.1039/C7MT00157F>
18. *Соколов А.В., Костевич В.А., Романико Д.Н. и др.* // *Биохимия.* 2012. Т. 77. С. 775.
19. *Peters G.S., Zakharchenko O.A., Konarev P.V. et al.* // *Nucl. Instrum. Methods Phys. Res. A.* 2019. V. 945. P. 162616. <https://doi.org/10.1016/j.nima.2019.162616>
20. *Blanchet C.E., Spilotos A., Schwemmer F. et al.* // *J. Appl. Cryst.* 2015. V. 48. P. 431. <https://doi.org/10.1107/S160057671500254X>
21. *Feigin L.A., Svergun D.I.* *Structure Analysis by Small-Angle X-Ray and Neutron Scattering.* New York: Plenum Press. XIII, 1987. 335 p. <https://doi.org/10.1007/978-1-4757-6624-0>
22. *Konarev P.V., Volkov V.V., Sokolova A.V. et al.* // *J. Appl. Cryst.* 2003. V. 36. P. 1277. <https://doi.org/10.1107/S0021889803012779>
23. *Guinier A.* // *Ann. Phys. (Paris).* 1939. V. 12. P. 161. <https://doi.org/10.1051/anphys/19391120161>
24. *Svergun D.I.* // *J. Appl. Cryst.* 1992. V. 25. P. 495. <https://doi.org/10.1107/S0021889892001663>
25. *Porod G.* // *General Theory, in Small-Angle X-ray Scattering / Eds. Glatter O., Kratky O.* London: Academic Press, 1982. P. 17.
26. *Petoukhov M.V., Franke D., Shkumatov A.V. et al.* // *J. Appl. Cryst.* 2012. V. 45. P. 342. <https://doi.org/10.1107/S0021889812007662>
27. *Boivin S., Aouffen M., Fournier A. et al.* // *Biochem. Biophys. Res. Commun.* 2001. V. 288. P. 1006. <https://doi.org/10.1006/bbrc.2001.5881>
28. *Kozakov D., Hall D.R., Xia B et al.* // *Nature Protocols.* 2017. V. 12. P. 255. <https://doi.org/10.1038/nprot.2016.169>
29. *Ravin H.A.* // *J. Lab. Clin. Med.* 1961. V. 58. P. 161.
30. *Varfolomeeva E.Y., Semenova E.V., Sokolov A.V. et al.* // *Free Radical Res.* 2016. V. 50. P. 909. <https://doi.org/10.1080/10715762.2016.1197395>
31. *Tamás M.J., Fauvet B., Christen Ph. et al.* // *Curr. Genet.* 2018. V. 64. P. 177. <https://doi.org/10.1007/s00294-017-0748-x>
32. *Wu Zh., Fernandez-Lima A., Russel D.H.* // *J. Am. Soc. Mass. Spectr.* 2010. V. 21. P. 522. <https://doi.org/10.1016/j.jasms.2009.12.020>
33. *Sokolov A.V., Kostevich V.A., Zakharova E.T. et al.* // *Free Radical Res.* 2015. V. 46. P. 800. <https://doi.org/10.3109/10715762.2015.1005615>
34. *Panasenko O.M., Torkhovskaya T.I., Gorudko I.V. et al.* // *Biochem. (Moscow).* 2020. V. 85. (S. 1.) P. 34. <https://doi.org/10.1134/S0006297920140035>

СТРУКТУРА МАКРОМОЛЕКУЛЯРНЫХ СОЕДИНЕНИЙ

УДК 539.26

ФОРМИРОВАНИЕ В РАСТВОРЕ СТРУКТУР ВЫСОКОГО ПОРЯДКА CBS-ПИРОФОСФАТАЗОЙ ИЗ *D. hafniense*

© 2021 г. Л. А. Дадина^{1,*}, Е. Ю. Сошинская¹, Ю. М. Чесноков^{1,2}, Р. А. Камышинский^{1,2},
А. Л. Васильев^{1,2}, Э. В. Штыкова¹

¹ Институт кристаллографии им. А.В. Шубникова ФНИЦ “Кристаллография и фотоника” РАН, Москва, Россия

² Национальный исследовательский центр “Курчатовский институт”, Москва, Россия

*E-mail: lubovmsu@mail.ru

Поступила в редакцию 03.09.2020 г.

После доработки 03.09.2020 г.

Принята к публикации 19.10.2020 г.

Для решения вопроса об олигомерном состоянии CBS-пирофосфатазы (CBS-PPase) дикого типа из *D. hafniense* проведено исследование этого фермента двумя независимыми структурными методами: с помощью малоуглового рентгеновского рассеяния и криогенной просвечивающей электронной микроскопии. Впервые обнаружено формирование в концентрированном растворе белка стабильных структур высокого порядка – крупных спиралевидных ассоциатов. Также впервые показано, что образование таких структур является обратимым процессом и при присоединении лиганда белок переходит в форму тетрамера, в которой обычно существует в разбавленных растворах. Результаты работы важны для понимания функциональных особенностей CBS-PPase, в том числе для понимания патогенеза некоторых заболеваний.

DOI: 10.31857/S0023476121050040

ВВЕДЕНИЕ

Олигомеризация является хорошо известным свойством белков и наблюдается во всех биологических системах [1]. Подсчитано, что по крайней мере 35% всех клеточных белков существует в виде олигомеров [2, 3]. С биологической точки зрения олигомеризация важна, поскольку она сохраняет жизненно важные молекулярные ресурсы, создавая из повторяющихся функционально значимых единиц более крупные полипептидные формирования, обладающие новыми функциональными свойствами, например кооперативностью.

Белковые олигомеры могут быть гомо- или гетеро-олигомерами и варьироваться от димеров до структур высокого порядка: капсидов – внешней белковой оболочки вирусов, и фибрилл – нитевидных белковых структур в клетках и тканях млекопитающих. Некоторые белки образуют только одно конкретное активное олигомерное состояние. Хорошо известным примером такого белка является гемоглобин, существующий в красных кровяных тельцах как гетеро-тетрамер [4]. Другие белки в процессе выполнения своих функций могут находиться в динамическом равновесии между несколькими состояниями с разной активностью, переключаясь между этими состояниями при определенных условиях. Отме-

тим, что во многих случаях такая динамическая сборка и разборка олигомеров играют центральную роль в регуляции активности белка, например в случае актина, который индуцирует подвижность клеток с помощью этого механизма [5].

Образование таких белковых структур высокого порядка, как фибриллы, часто лежит в основе патогенеза нейродегенеративных заболеваний, встречающихся как у человека, так и у животных. Амилоидные фибриллы и олигомеры связаны с большим количеством заболеваний человека, включая болезнь Альцгеймера и прионные болезни. Формирование белковых агрегатов, отражающее конформационную и термодинамическую нестабильность белка, можно рассматривать как общий компонент патогенеза [6]. Недавно стало известно, что даже такие хорошо изученные и полезные белки, как, например, лизоцим, являющийся антибактериальным агентом и разрушающий клеточные стенки бактерий, при определенных условиях могут образовывать патогенные формирования и превращаться в амилоидные структуры [7]. Поэтому важной областью исследования структуры и функции белковых макромолекул является изучение процессов их олигомеризации и формирования структур высокого порядка.

В недавних исследованиях структурных особенностей неорганических CBS-пирофосфатаз (CBS-PPase) из разных бактериальных клеток (*D. hafniense*, *E. lenta* и *E. harbinense*) было показано, что в растворе этот фермент образует устойчивый тетрамер [8, 9], хотя имеется устоявшееся мнение, что CBS-PPase является гомодимером [10]. Отметим, что структурные исследования CBS-PPase и определение их олигомерного состояния очень важны для установления механизмов передачи сигнала между активным и регуляторным центрами, понимания молекулярных основ регуляции ферментов этого семейства и определения механизма подавления их каталитической активности.

CBS-PPase относятся к семейству II растворимых неорганических PPase, которые являются наиболее каталитически активными среди всех типов PPase. Их активные сайты располагаются между DHH- и DHNA2-доменами [11], а автоингибирующая регуляторная вставка в DHH-доме-не [12]. Эта вставка образована двумя доменами цистатион- β -синтазы (CBS) и одним доменом DRTGG (названного в честь консервативных мотивов Asp-Arg и Thr-Gly-Gly).

Известно, что CBS-домены участвуют в образовании структур более высокого порядка, формируя из двух модулей Бейтмана гомодимерные формы, состоящие уже из четырех CBS-доменов [13]. При этом CBS-домены способны связывать лиганды, в качестве которых в CBS-PPase выступают различные фосфатные производные аденозина [13–15].

Несмотря на достигнутые успехи в изучении структуры и базовых молекулярных механизмов функционирования этих ферментов [16, 17], имеющих данных недостаточно для установления механизма регуляции белка и определения пути передачи сигнала от регуляторного к активному центру. Вопрос об олигомерном состоянии белка по-прежнему остается открытым из-за противоречивости данных [8–10] и отсутствия структуры высокого разрешения полноразмерного фермента. На данный момент известны лишь несколько кристаллических структур канонических PPase семейства II [18–20] и регуляторной вставки из *Clostridium perfringens*, в которой пара CBS- и DRTGG-домен образует гомодимер. Но выяснить механизм передачи сигнала между активным и регуляторным центрами невозможно без определения строения полноразмерного фермента, в связи с чем структурный анализ полноразмерной CBS-PPase в условиях, наиболее близких к физиологическим, остается актуальной задачей. Поэтому цель настоящей работы – структурный анализ четвертичной структуры и олигомерного состояния белка CBS-PPase из *D. hafniense* (dh-PPase) с помощью взаимодополняющих ме-

тодов структурной диагностики – малоуглового рентгеновского рассеяния (МУРР) и криогенной просвечивающей электронной микроскопии (крио-ПЭМ).

Результаты работы могут стать шагом к пониманию принципиально важных функциональных особенностей CBS-PPase и способствовать устранению пробелов в понимании патогенеза некоторых заболеваний.

МАТЕРИАЛЫ И МЕТОДЫ

Выделение и очистка CBS-PPase. Образцы бактериальной неорганической пирофосфатазы dh-PPase предоставлены научной группой из Института физико-химической биологии им. А.Н. Белозерского МГУ им. М.В. Ломоносова. Для приготовления препаратов dh-PPase использовали методики, описанные в [15]. Чистота препаратов была подтверждена SDS-PAGE, как описано в [10]. Концентрацию белка определяли спектрофотометрически, используя значение $A_{280}^{0.1\%}$, равное 0.478, рассчитанное из аминокислотного состава dh-PPase в программе ProtParam [21]. Концентрация белка варьировалась от 0.5 до 9 мг/мл. Все эксперименты проводили в 0.1 М Mops-KOH-буфере, pH 7.2, 2 мМ MgCl₂, 0.1 мМ CoCl₂, 150 мМ KCl.

Эксперимент и анализ данных МУРР. Измерения кривых МУРР проводили в Гамбурге на синхротроне DESY, накопительном кольце PetraIII линии P12. Линия оснащена роботизированной системой для автоматической смены образцов, а также 2D-детектором Pilatus2M (фирма DECTRIS, Швейцария). Интенсивность рассеяния с длиной волны рентгеновского излучения $\lambda = 0.124$ нм была измерена в области значений волновых векторов $0.08 < s < 7$ нм⁻¹, где $s = (4\pi \sin \theta)/\lambda$, 2θ – угол рассеяния. Для предотвращения искажения результатов из-за возможных радиационных повреждений было снято по 50 кривых рассеяния для буфера (перед серией экспериментов для образца и после) и образца. В ходе обработки полученных данных радиационных повреждений выявлено не было. Первичная обработка полученных данных была проведена с использованием стандартных процедур [22].

Для исключения влияния возможного межчастичного взаимодействия и образования агрегатов также использовали процедуру онлайн эксклюзионной хроматографии в сочетании с малоугловым рентгеновским рассеянием (SEC-SAXS). Данный метод позволяет получить рассеяние от монодисперсного раствора. Эксперимент проводили с использованием колонки Superdex 200 10/30 (GE Healthcare) в режиме непрерывного потока из колонок SEC, элюируемых с использованием

периодов воздействия 1 св общей сложности 3600 с (объем одной колонки).

Для получения максимальной эффективности восстановления исходных данных экспериментальные кривые были обрезаны так, чтобы на выходе получилось $N = 10-20$ шенноновских каналов:

$$N = \frac{s}{\Delta s}, \quad (1)$$

где $\Delta s = \frac{\pi}{D_{\max}}$.

Исходя из данного условия, для анализа и модельных расчетов использовали наиболее информативный интервал волновых векторов $0.08 < s < 2 \text{ нм}^{-1}$.

Радиус инерции R_g был получен с использованием приближения Гинье в области малых значений вектора рассеяния s . Из наклона прямолинейного участка $\ln I(s)$ от s^2 графика Гинье экспериментальных данных определялась интенсивность рассеяния в нулевой угол $I(0)$, пропорциональная молекулярной массе:

$$I_{\text{exp}}(s) = I(0) \exp(-s^2 R_g^2/3). \quad (2)$$

Данное соотношение справедливо в области $(sR_g) < 1.3$.

Молекулярную массу (ММ) рассчитывали двумя способами: с помощью Байесовского подхода (ММ_В) [23] и на основе расчета объема Порода V_p (ММ_р). ММ_р находили из установленно-го эмпирического соотношения между V_p и ММ_р, которое для белков равно 1.65 [24].

Функцию распределения по расстояниям $p(r)$ рассчитывали с помощью программы GNOM [25] по уравнению

$$p(r) = \frac{1}{2\pi^2} \int_0^\infty sr I(s) \sin(sr) ds. \quad (3)$$

По функции $p(r)$ определяется максимальный размер белка D_{\max} из условия $p(r) = 0$ при $r > D_{\max}$.

Параметр D_{\max} необходим для *ab initio*-моделирования, которое реализовано в программе DAMMIN [26]. Данный алгоритм поиска формы низкого разрешения основан на моделировании отжига в ограниченном объеме (сфере диаметром D_{\max}), а конечный результат достигается за счет минимизации невязки χ^2 с экспериментальными данными:

$$\chi^2 = \frac{1}{N-1} \sum_j \left[\frac{I_{\text{exp}}(s_j) - c I_{\text{calc}}(s_j)}{\sigma(s_j)} \right]^2, \quad (4)$$

где N – число экспериментальных точек, $I_{\text{exp}}(s_j)$ и $\sigma(s_j)$ – экспериментальные интенсивности и их ошибки, $I_{\text{calc}}(s_j)$ – интенсивность, вычисленная от модели, c – шкалирующий множитель.

В качестве дополнительного метода восстановления структуры белка был использован гибридный подход (программа CORAL [27]), сочетающий в себе моделирование жесткими телами на основе структур высокого разрешения (PDB ID: 2haw, 3l31) и *ab initio*-алгоритм для достраивания недостающих фрагментов белка. При этом недостающие аминокислоты моделируются с помощью виртуальных остатков, свернутых в белково-подобную цепь так, чтобы рассеяние от цельной структуры, рассчитанное в программе CRY SOL [28], соответствовало экспериментальному рассеянию, минимизируя невязку χ^2 .

Полученные двумя способами модели белка проанализированы с помощью программы SUPCOMB [29] для определения степени различия в их пространственной организации. Значение нормализованного пространственного отклонения (NSD), близкое к 1.0, показывает, что модели схожи.

Вспомогательную программу GEN_DAM_body [В.В. Волков, не опубликовано, используется с разрешения автора] применяли для построения спирали с заданными параметрами. Расчет интенсивности рассеяния от полученной модели проводили с помощью программы DAM2IS из программного комплекса ATSAS [27].

Анализ неоднозначности полученных *ab initio*-моделей выполняли с помощью программы AMBIMETER [30]. Прогнозирование форм белков и их структурную классификацию проводили с помощью DATCLASS [24].

Для анализа количественного состава олигомерной смеси dh-PPase в растворе использовали программу OLIGOMER [22]. С учетом интенсивности рассеяния каждой компоненты в смеси, $I_i(s)$, программа приближает экспериментальную кривую рассеяния к их линейной комбинации, чтобы определить объемную долю w_i каждой компоненты. Уравнение

$$I(s) = \sum (w_i I_i(s)) \quad (5)$$

решается относительно w_i методом неотрицательных наименьших квадратов, чтобы минимизировать расхождение между экспериментальной и расчетной кривыми рассеяния.

Эксперимент по криогенной просвечивающей электронной микроскопии. Исследование проводили с помощью крио-ПЭМ TitanKrios (Thermo Fisher Scientific, США), оснащенного высокочувствительной системой прямого детектирования электронов Falcon II (Thermo Fisher Scientific, США), а также корректором сферических aberrаций (Image corrector, CEOS, Германия) при ускоряющем напряжении 300 кВ. Перед крио-ПЭМ-исследованием проводили процедуру витрификации образцов. Электронно-микроскопические медные сетки, покрытые тонким слоем аморфного углерода, были гидрофилизированы в тлеющем

разряде (30 с, 25 мА) с помощью установки Pelco easi Glow (Ted Pella, США). Далее на сетки наносили 3 мкл препарата с образцом и с помощью установки Vitrobot Mark IV (Thermo Fisher Scientific, США) проводили процедуру витрификации образцов в жидком этане, охлажденном до температуры жидкого азота. В результате частицы образца были зафиксированы в тонком слое аморфного льда, что позволило исследовать их в нативном состоянии и минимизировать радиационные повреждения объекта в ходе дальнейших экспериментов. В течение всего эксперимента влажность в камере системы Vitrobot Mark IV составляла 95–100%, температура 4°C.

Набор данных крио-ПЭМ проводили в автоматизированном режиме с помощью программного обеспечения EPU (Thermo Fisher Scientific, США). Изображения получены при увеличении 59000 (размер пикселя 1.107 Å), суммарная доза электронов, прошедших через единицу площади образца составила ~ 60 э/Å². Предобработка стеков изображений – коррекция дрейфа, оценка параметров функции передачи контраста (CTF) и выбор одиночных проекций объекта интереса – выполнена с помощью программного пакета Warp [31]. ~ 318000 потенциальных проекций одиночных частиц были выделены из 606 стеков исходных изображений после предобработки и экспортированы в программный пакет CryoSPARC [32] для проведения двумерной классификации. На следующем этапе были отобраны классы, содержащие изображения ~ 65000 частиц dh-PPase, и проведена финальная двумерная классификация.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Поскольку белок CBS-PPase из бактерии *D. hafniense* был получен при концентрации 1 мг/мл, для определения его олигомерного состояния в растворе образец концентрировали до 9 мг/мл. Кривая малоуглового рассеяния концентрированного раствора dh-PPase (рис. 1, кривая 1) в области малых углов демонстрирует резкое увеличение интенсивности рассеяния и выраженные максимумы в диапазоне значений вектора рассеяния $0.027 < s < 0.25$ нм⁻¹. Такой профиль рассеяния отражает присутствие в растворе крупных рассеивающих объектов (кластеров), имеющих строго определенную форму. Наблюдаемые максимумы не исчезают при последовательном разбавлении раствора до концентрации 0.5 мг/мл (рис. 1, кривые 2–4), что говорит о высокой стабильности образованных структур. Для определения их формы по выделенному диапазону данных ($0.027 < s < 0.25$ нм⁻¹), соответствующему 16 шенноновским каналам, было проведено *ab initio*-восстановление с помощью программы DAMMIN [26]. Соответствующая функция распределения по расстояниям $p(r)$ представлена на рис. 1в.

Максимальный размер D_{\max} , определенный из функции $p(r)$, равен 210 нм, а поперечное сечение рассеивающего объекта ~ 90 нм. Наличие периодически повторяющихся слабо выраженных максимумов на профиле функции $p(r)$ свидетельствовало о присутствии упорядоченных повторяющихся элементов в структуре кластеров dh-PPase. При восстановлении программой DAMMIN формы белка этими периодически повторяющимися элементами оказались витки слегка искаженной спирали, размеры которой соответствовали размерам, определенным из анализа функции $p(r)$, т.е. спиральному телу длиной 210 нм и диаметром 90 нм (рис. 1г). Толщина витка спирали оказалась равной 35 нм.

Наличие спиральной структуры подтверждено независимым моделированием с помощью геометрических тел. Для этого была построена шариковая модель спирального тела, внешние размеры которого определялись из анализа функции распределений по расстояниям $p(r)$. От полученной спирали была рассчитана интенсивность малоуглового рассеяния и проведено сравнение с экспериментальными данными МУРР (рис. 1б). Профиль рассеяния от шариковой модели идеальной спирали соответствует профилю экспериментальной кривой (рис. 1б), а имеющееся расхождение обусловлено неидеальностью структуры dh-PPase в растворе, поскольку моделирование геометрическими телами представляет собой метод грубого приближения и не может учитывать все особенности реальных структур, а также подвижность и гибкость их отдельных фрагментов. Однако проведенное моделирование дает основание утверждать, что при концентрировании dh-PPase действительно формирует в растворе большие спиралевидные ассоциаты (рис. 1а, кривые 1–4).

Поскольку в растворе dh-PPase присутствуют крупные белковые ассоциаты-кластеры, для получения данных рассеяния только от минимальных по размеру стабильных форм белка использовали метод SEC-SAXS. После проведения процедуры SEC-SAXS dh-PPase была получена в монодисперсном состоянии, о чем свидетельствуют характерная форма кривой рассеяния (рис. 2а, кривая 1) и график в координатах Гинье с достаточно длинным прямолинейным участком (рис. 2б). Радиус инерции R_g , рассчитанный с использованием приближения Гинье в области малых значений вектора рассеяния s , оказался равен 4.95 нм.

График Кратки (рис. 2в) для этой фракции белка имеет характерный колоколообразный вид, что указывает на то, что образец в целом компактный и имеет ограниченное число разупорядоченных гибких областей [33]. Компактность полученной структуры также подтверждена с помо-

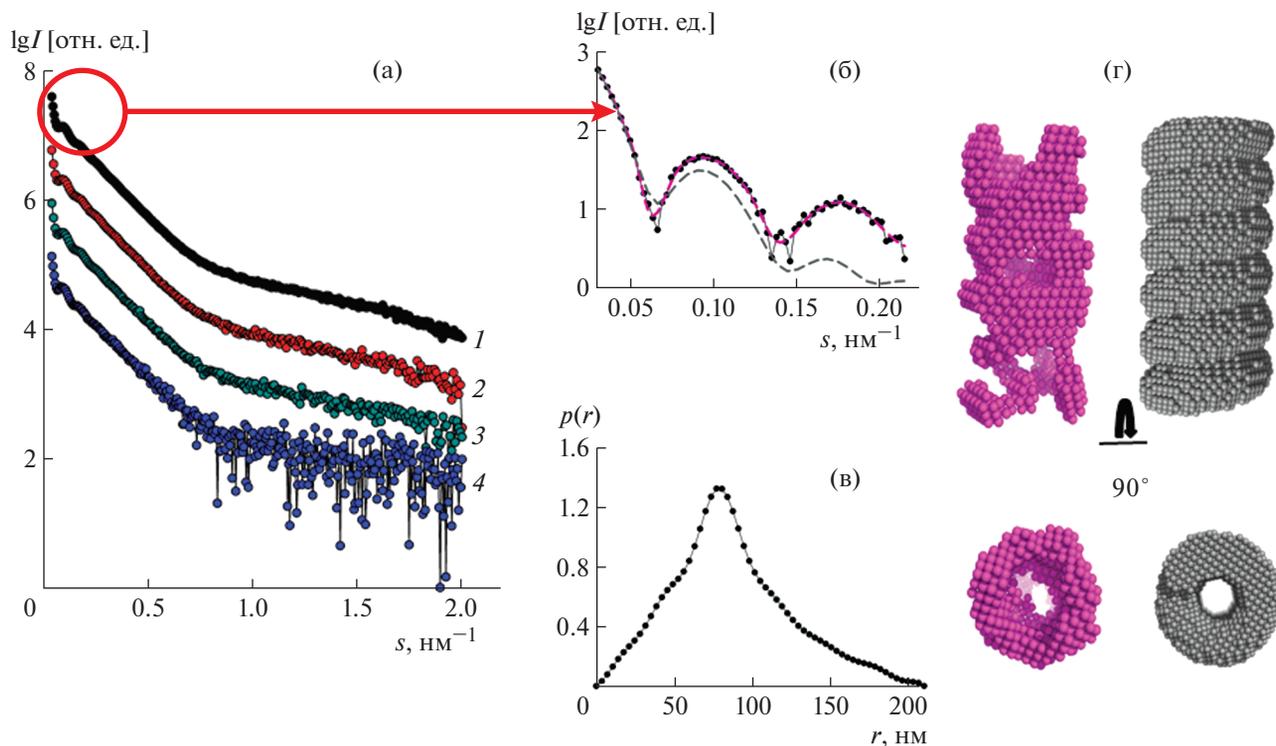


Рис. 1. Структурный анализ ассоциатов полноразмерного белка dh-PPase (а); кривые рассеяния от растворов белка dh-PPase с концентрацией 9 (1), 4.5 (2), 2.5 (3), 0.5 мг/мл (4). Анализ структуры ассоциатов: начальная часть ($0.027 < s < 0.25 \text{ нм}^{-1}$) экспериментальной кривой 1 (точки), модельная кривая от *ab initio*-модели (сплошная линия) и рассеяние от идеальной спирали (пунктирная линия) (б). Функция распределения по расстояниям $p(r)$ для ассоциатов (в). Восстановленное по функции $p(r)$ с помощью программы DAMMIN спиральное тело, идеальная спираль с теми же параметрами (г).

стью программы прогнозирования форм белков по кривой рассеяния DATCLASS [24].

Значения ММ, определенных по экспериментальной кривой, исходя из эмпирического соотношения Породовского объема к ММ белка [24] и обобщенного Байесовского подхода [23], составляют 250 и 243 кДа соответственно. В целом полученные значения ММ и радиуса инерции R_g соответствуют тетрамеру белка (табл. 1).

Используя функцию распределения по расстояниям $p(r)$, методом *ab initio*, реализованном в программе DAMMIN [26], получена форма белка низкого разрешения (рис. 2д, серые шарики), рассеяние от которой хорошо соответствует экспериментальным данным $\chi^2 = 1.1$ (рис. 2а, кривая 2, пунктирная кривая). Так как полученные

структурные характеристики (табл. 1) указывают на то, что белок образует тетрамер в растворе, т.е. димер димеров, при восстановлении формы использовали симметрию $P2$.

Поскольку решение обратной задачи восстановления трехмерной формы по одномерной кривой рассеяния является неоднозначным, с помощью программы AMBIMETER [30] проведена оценка неоднозначности полученной модели. Как правило, индекс неоднозначности ~ 1.5 предполагает, что найдено уникальное, т.е. практически единственное решение, в то время как оценка выше 2.5 говорит о неоднозначном восстановлении формы. В данном случае индекс неоднозначности оказался меньше 1.5, что указывает на уникальность полученной модели.

Таблица 1. Макромолекулярные характеристики dh-PPase после процедуры онлайн эксклюзионной хроматографии

Образец	R_g , нм	V_p , нм ³	D_{max} , нм	ММ _р , кДа	ММ _в , кДа
dh-PPase	4.95 ± 0.10	412 ± 20	18.3 ± 1	250 ± 10	243 (доверительный интервал: 195–264)

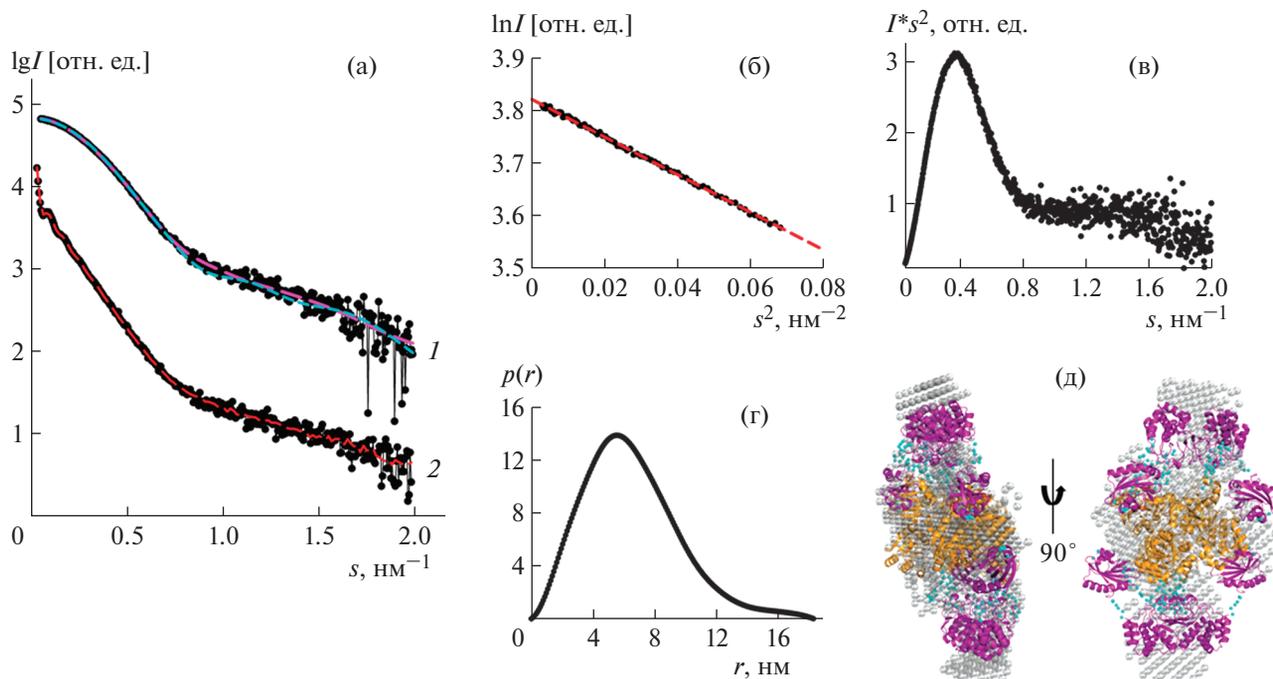


Рис. 2. Сравнение кривой рассеяния dh-PPase после онлайн эксклюзионной хроматографии (точки) и модельных кривых, полученных с помощью программ DAMMIN (удлиненная пунктирная линия) и CORAL (короткая пунктирная линия) (1); сравнение кривой рассеяния от концентрированного раствора dh-PPase (точки) с модельной кривой, полученной программой OLIGOMER (пунктирная линия) (2) (а). График Гинье (б); график в координатах Кратки (в); функция распределения по расстояниям для dh-PPase (r). Модели dh-PPase, полученные с помощью программ DAMMIN (шарики) и CORAL (ленты – каталитические домены DHH и DHHA2, регуляторные домены CBS и DRTGG, шарики – восстановленные программой CORAL линкеры между доменами) (д).

В качестве дополнительного метода определения структуры фермента использовали гибридное моделирование и программу CORAL [27]. Данный метод позволяет получить более детальную структуру dh-PPase и определить относительное положение каталитической и регуляторной частей белка в растворе. Для моделирования использовали известные структуры высокого разрешения отдельных частей CBS-PPase: регуляторная часть, состоящая из CBS- и DRTGG-домена (PDB ID: 3I31), и каталитическая часть, т.е. DHH- и DHHA2-домены, структура которой была получена в [8]. Поскольку каждая из частей является димерной, они были разделены на мономерные составляющие с условием контакта между соответствующими мономерами, а каталитическая и регуляторная части были соединены между собой достроенными программой CORAL гибкими линкерами. Для получения тетрамера была применена симметрия $P2$. Полученная модель представлена на рис. 2д и хорошо соответствует экспериментальным данным с $\chi^2 = 1.62$.

Как можно видеть, модели, полученные двумя независимыми методами, хорошо согласуются друг с другом (рис. 2д), что подтверждается также количественной оценкой с использованием нормализованного пространственного расхождения

NSD [30]. При сравнении моделей, полученных с помощью программ DAMMIN и CORAL, значение NSD равно 1.5. Это свидетельствует о хорошем совпадении структур, полученных разными методами.

Учитывая неоднозначность решения структурных задач в малоугловом рассеянии и противоречивость данных [8–10], необходимо независимое подтверждение существования dh-PPase в растворе в виде тетрамера в качестве минимальной стабильной формы этого белка. С этой целью был использован метод крио-ПЭМ.

На рис. 3 представлены результаты 2D-классификации частиц.

Анализ данных крио-ПЭМ также приводит к выводу, что dh-PPase представляет собой тетрамер. Однако из-за большого количества доменов, соединенных гибкими фрагментами белковой цепи, наблюдается высокая подвижность доменов, что затрудняет получение трехмерной структуры высокого разрешения. Тем не менее форма проекций частиц, представленных на рис. 3, и их размер соответствуют моделям, полученным по данным МУРР, следовательно, тетрамер является наименьшей устойчивой структурой dh-PPase.

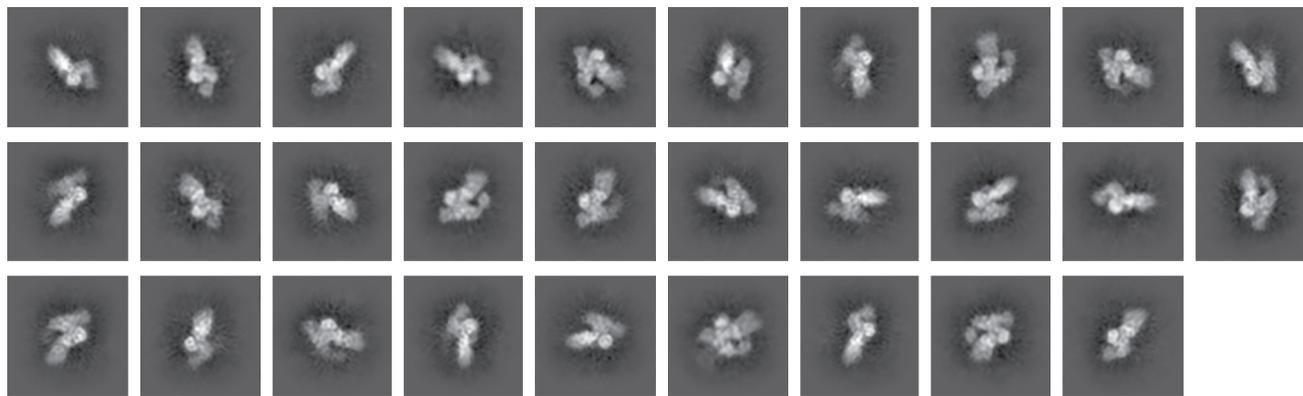


Рис. 3. Результаты двумерной классификации проекций одиночных частиц dh-PPase. Ширина одиночного изображения (квадрата) – 28 нм.

Однако, как показали метод онлайн эксклюзионной хроматографии и последующий анализ, в концентрированном (9 мг/мл) растворе белок существует в виде крупных спиралевидных кластеров. Но можно предположить и одновременное существование какого-то количества тетрамеров dh-PPase. Применение программы OLIGOMER [22] показало, что раствор dh-PPase действительно представляет собой смесь, состоящую из объемной доли $w_i = 0.93$ спиралевидных кластеров и объемной доли $w_i = 0.07$ тетрамеров белка (рис. 2а, кривая 2). Профиль кривой МУРР такой смеси в основном соответствует рассеянию от крупных удлиненных тел. То есть при достаточно высокой концентрации в растворе происходит почти полная ассоциация белка в большие упорядоченные кластеры. Эти ассоциаты являются термодинамически и структурно устойчивыми и не распадаются на отдельные макромолекулы даже при разведении раствора dh-PPase почти в 20 раз (рис. 1а).

Для ответа на вопрос об обратимости такой ассоциации при выполнении белком его функций, например, при присоединении лиганда, к концентрированному раствору dh-PPase был добавлен диаденозинтетрафосфат (Ar_4A) в нужной концентрации (1000 мкМ). Оказалось, что в присутствии лиганда спиралевидные кластеры распадаются, а раствор белка представляет собой смесь небольшого количества агрегатов с тетрамерами dh-PPase, т.е. при присоединении субстрата минимальной рабочей единицей фермента действительно является тетрамер [9].

Таким образом, наблюдаемая ассоциация dh-PPase не случайна, обратима и отражает определенные функциональные свойства фермента.

ЗАКЛЮЧЕНИЕ

Полученные результаты структурных исследований двумя независимыми методами, МУРР и

крио-ПЭМ, свидетельствуют о тетрамерной структуре dh-PPase в растворе. Этот вывод подтверждается и в процессе взаимодействия фермента с лигандом. Добавление последнего к концентрированному раствору dh-PPase, находящемуся в основном в виде крупных спиралевидных ассоциатов, приводит к исчезновению ассоциатов и получению монодисперсного раствора тетрамера белка. Образование упорядоченных структур высокого порядка этого фермента, как и обратимость этого явления, наблюдались в настоящей работе впервые. Однако природу и значимость этого явления еще только предстоит определить соответствующими методами биохимии и биологии.

Авторы выражают благодарность научной группе из Института физико-химической биологии им. А.Н. Белозерского МГУ им. М.В. Ломоносова и лично В.А. Анашкину за предоставление образцов.

Работа выполнена при поддержке Министерства науки и высшего образования РФ в рамках выполнения работ по Государственному заданию ФНИЦ “Кристаллография и фотоника” РАН.

СПИСОК ЛИТЕРАТУРЫ

1. *Mayssam H.A., Barbara I.* // Bioorg. Med. Chem. 2005. V. 13. P. 5013. <https://doi.org/10.1016/j.bmc.2005.05.037>
2. *Jones S., Thornton J.M.* // Proc. Natl. Acad. Sci. 1996. V. 93. P. 13. <https://doi.org/10.1073/pnas.93.1.13>
3. *Goodsell D.S., Olson A.J.* // Annu. Rev. Biophys. Biomol. Struct. 2000. V. 29. P. 105. <https://doi.org/10.1146/annurev.biophys.29.1.105>
4. *Fermi G., Perutz M.F., Shaanan B. et al.* // J. Mol. Biol. 1984. V. 175. P. 159. [https://doi.org/10.1016/0022-2836\(84\)90472-8](https://doi.org/10.1016/0022-2836(84)90472-8)
5. *Ono S.* // Int. Rev. Cytol. 2007. V. 258. P. 1. [https://doi.org/10.1016/S0074-7696\(07\)58001-0](https://doi.org/10.1016/S0074-7696(07)58001-0)

6. Сахаров В.Н., Лутвицкий П.Ф. // Вестн. РАМН. 2016. Т. 71. С. 46.
<https://doi.org/10.15690/vramn635>
7. Granel B., Valleix S., Serratrice J. et al. // *Medicine*. 2006. V. 85. P. 66.
<https://doi.org/10.1097/01.md.0000200467.51816.6d>
8. Dadinova L.A., Soshinskaia E.Yu., Jeffries C.M. et al. // *Biomolecules*. 2020. V. 10. P. 564.
<https://doi.org/10.3390/biom10040564>
9. Dadinova L.A., Anashkin V.A., Shtykova E.V. // *Crystallography Reports*. 2020. V. 65. P. 749.
<https://doi.org/10.1134/S1063774520050053>
10. Parfenyev A., Salminen A., Halonen P. et al. // *J. Biol. Chem.* 2001. V. 276. P. 24511.
<https://doi.org/10.1074/jbc.M101829200>
11. Aravind L., Koonin E. // *Trends Cell. Biol.* 1998. V. 23. P. 17.
12. Baykov A.A., Anashkin V.A., Salminen A. et al. // *FEBS Lett.* 2017. V. 591. P. 3225.
<https://doi.org/10.1002/1873-3468.12877>
13. Anashkin V.A., Salminen A., Tuominen H.K. et al. // *J. Biol. Chem.* 2015. V. 290. P. 27594.
<https://doi.org/10.1074/jbc.M115.680272>
14. Jamsen J., Tuominen H., Salminen A. et al. // *Biochem. J.* 2007. V. 408. P. 327.
<https://doi.org/10.1042/BJ20071017>
15. Salminen A., Anashkin V.A., Lahti M. et al. // *J. Biol. Chem.* 2014. V. 289. P. 22865.
<https://doi.org/10.1074/jbc.M114.589473>
16. Anashkin V.A., Salminen A., Osipova E. et al. // *ACS Omega*. 2019. V. 4. P. 15549.
<https://doi.org/10.1021/acsomega.9b01879>
17. Anashkin V.A., Salminen A., Vorobjeva N.N. et al. // *Biochem. J.* 2016. V. 473. P. 2097.
<https://doi.org/10.1042/BCJ20160293>
18. Merckel M., Fabrichny I., Salminen A. et al. // *Structure*. 2001. V. 94. P. 289.
[https://doi.org/10.1016/s0969-2126\(01\)00587-1](https://doi.org/10.1016/s0969-2126(01)00587-1)
19. Ahn S., Milner A., Fütterer K. et al. // *J. Mol. Biol.* 2001. V. 313. P. 797.
<https://doi.org/10.1006/jmbi.2001.5070>
20. Rantanen M., Lehtiö L., Rajagopal L. et al. // *Acta Cryst. D*. 2007. V. 63. P. 738.
<https://doi.org/10.1107/S0907444907019695>
21. Gasteiger E., Hoogland C., Gattiker A. et al. *The Proteomics Protocols Handbook*. N.J.: Humana Press, 2005. 607 p.
22. Konarev P.V., Volkov V.V., Sokolova A.V. et al. // *J. Appl. Cryst.* 2003. V. 36. P. 1277.
<https://doi.org/10.1107/S0021889803012779>
23. Hajizadeh N.R., Franke D., Jeffries C.M. et al. // *Sci. Rep.* 2018. V. 8. P. 7204.
<https://doi.org/10.1038/s41598-018-25355-2>
24. Franke D., Petoukhov M.V., Konarev P.V. et al. // *J. Appl. Cryst.* 2017. V. 50. P. 1212.
<https://doi.org/10.1107/S1600576717007786>
25. Svergun D.I. // *J. Appl. Cryst.* 1992. V. 25. P. 495.
<https://doi.org/10.1107/S0021889892001663>
26. Svergun D.I. // *Biophys. J.* 1999. V. 76. P. 2879.
[https://doi.org/10.1016/S0006-3495\(99\)77443-6](https://doi.org/10.1016/S0006-3495(99)77443-6)
27. Petoukhov M.V., Franke D., Shkumatov A.V. et al. // *J. Appl. Cryst.* 2012. V. 45. P. 342.
<https://doi.org/10.1107/S0021889812007662>
28. Svergun D.I., Barberato C., Koch M.H.J. // *J. Appl. Cryst.* 1995. V. 28. P. 768.
<https://doi.org/10.1107/S0021889895007047>
29. Kozin M., Svergun D. // *J. Appl. Cryst.* 2001. V. 34. P. 33.
<https://doi.org/10.1107/S0021889800014126>
30. Petoukhov M.V., Svergun D.I. // *Acta Cryst. D*. 2015. V. 71. P. 1051.
<https://doi.org/10.1107/S1399004715002576>
31. Tegunov D., Cramer P. // *Nat. Methods*. 2019. V. 16. P. 1146.
<https://doi.org/10.1038/s41592-019-0580-y>
32. Punjani A., Rubinstein J., Fleet D. et al. // *Nat. Methods*. 2017. V. 14. P. 290.
<https://doi.org/10.1038/nmeth.4169>
33. Jacques D.A., Guss J.M., Svergun D.I. et al. // *Acta Cryst. D*. 2012. V. 68. P. 620.
<https://doi.org/10.1107/S0907444912012073>

СТРУКТУРА МАКРОМОЛЕКУЛЯРНЫХ СОЕДИНЕНИЙ

УДК 577

СУПРАМОЛЕКУЛЯРНЫЕ КОМПЛЕКСЫ ТЕТРАПЕПТИДОВ, СПОСОБНЫЕ ИНДУЦИРОВАТЬ КОНФОРМАЦИОННЫЕ ПЕРЕХОДЫ БЕТА-ДОМЕНА АЛЬФА-ЛАКТАЛЬБУМИНА ЧЕЛОВЕКА

© 2021 г. Я. А. Забродская^{1,2,3,4}, А. В. Швецов^{1,2,4}, Ю. П. Гармай¹, Д. В. Лебедев^{1,4},
Р. Датгани⁵, В. В. Егоров^{1,4,6,*}

¹ Национальный исследовательский центр “Курчатовский институт” – ПИЯФ, Гатчина, Россия

² Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия

³ Научно-исследовательский институт гриппа им. А.А. Смородинцева Минздрава России, Санкт-Петербург, Россия

⁴ Национальный исследовательский центр “Курчатовский институт”, Москва, Россия

⁵ Европейский центр синхротронного излучения, Гренобль, Франция

⁶ Институт экспериментальной медицины, Санкт-Петербург, Россия

*E-mail: egorov_vv@pnpi.nrcki.ru

Поступила в редакцию 12.09.2020 г.

После доработки 23.12.2020 г.

Принята к публикации 23.12.2020 г.

При помощи метода малоуглового рентгеновского рассеяния показано, что тетрапептид, являющийся индуктором фибриллогенеза для фрагмента бета-домена альфа-лактальбумина человека, активен в форме супрамолекулярных комплексов. Образование таких комплексов не детектировалось при использовании микроскопии и аналитической гель-фильтрации. Молекулярно-динамическое моделирование поведения ансамбля изучаемых тетрапептидов методом свободной диффузии подтвердило их склонность к образованию олигомеров, наблюдаемую в эксперименте. Полученные данные и методические подходы могут быть применены при разработке новых препаратов пептидной природы, способных к модулированию активности белков путем воздействия на их пространственную структуру.

DOI: 10.31857/S002347612105026X

ВВЕДЕНИЕ

Способность биомолекул к самоорганизации играет ключевую роль в функционировании живых систем. Для изучения надмолекулярных комплексов применяется широкий спектр биофизических и биохимических методов, таких как электронная и атомно-силовая микроскопия, хроматография, электрофорез в геле, динамическое светорассеяние. Однако некоторые олигомерные комплексы, существующие только в состоянии динамического равновесия олигомер–мономер в растворе, разрушаются при исследовании, так как взаимодействия между мономерами менее энергетически выгодны, чем взаимодействия между мономером и, например, поверхностью хроматографического сорбента или подложкой для проведения исследования с помощью микроскопии. В то же время динамическое светорассеяние не обладает достаточным разрешением для определения морфологических характеристик изучаемых комплексов.

В последнее время появляется все больше данных о роли коротких пептидов (образовавшихся, например, в результате ограниченного протеолиза белков [1] или при реализации альтернативных механизмов трансляции [2]) в регуляции активности белков, опосредованной воздействием на их пространственную структуру [3]. Индукция конформационных переходов в целевых вирусных белках или маркерных белках опухолей с применением пептидов используется для создания новых препаратов, направленных на борьбу с инфекционными и онкологическими заболеваниями [4–6].

В [7, 8] для изучения процесса индуцированных пептидами конформационных переходов в полипептидах использовался модельный пептид WT (участок бета-домена альфа-лактальбумина человека протяженностью 17 аминокислотных остатков (GYDTQAIIVENNESTEYGG)). Выбор этого пептида в качестве модели обусловлен, в том числе, перспективой его использования в качестве основы для противоопухолевого препарата, специфически воздействующего на клетки опу-

Таблица 1. Первичные структуры пептидов WT, L и R

Пептид	Аминокислотная последовательность
WT	GYDTQAIVENNESTEY G
L	GYDT
R	TDY G

Примечание. Жирным шрифтом выделен аминокислотный остаток, замененный на близкий по свойствам в пептиде R.

холей молочной железы [9]. Было показано, что добавление в раствор тетрапептидов, гомологичных по первичной структуре C- или N-концевым участкам пептида WT (пептиды L и R соответственно), приводит к быстрому переходу WT в бета-структурированную форму и индуцирует образование амилоидоподобных фибрилл. Первичные структуры пептидов WT, L и R приведены в табл. 1. Аминокислотная замена остатка глутаминовой кислоты TDY**G** на аспарагиновую TEY**G** обусловлена тем, что пептид TEY**G** оказался крайне гигроскопичным и неустойчивым в растворе.

При индукции фибрилlogenеза не наблюдалось значимого уменьшения числа молекул индуктора в растворе и разницы в морфологии фибрилл, образовавшихся в присутствии и в отсутствие индуктора [10]. Была предложена модель, объясняющая способность тетрапептидов индуцировать конформационные переходы [11]. Вкратце при помощи молекулярно-динамического (МД) моделирования было показано, что тетрапептиды способны стабилизировать бета-структурированную, способную к образованию надмолекулярных комплексов, форму пептида WT. При этом в рамках модели после образования олигомеров WT тетрапептиды высвобождаются из комплексов и способны индуцировать фибрилlogenез дальше по механизму, сходному с ферментативной реакцией. В то же время предложенная модель не могла объяснить тот факт, что индукторы при добавлении в концентрациях ниже критической (менее 0.5 мг/мл) не вызывали увеличения скорости фибрилlogenеза. Данное исследование посвящено уточнению модели индукции фибрилlogenеза в описываемой модельной системе.

МАТЕРИАЛЫ И МЕТОДЫ

Пептиды WT, L и R. Пептиды GYDTQAIVENNESTEY**G** (WT) и GYDT (L), TDY**G** (R) были синтезированы в НПФ “Верта”, чистота выше 90%.

Атомно-силовую микроскопию (АСМ) пептидных агрегатов WT проводили на микроскопе Bio-Solver Pro с измерительной головкой Smena-B с

использованием полуконтактного режима, зонд NSG03 (NT-MDT, Россия). Образцы растворяли в концентрации 0.7 мг/мл в фосфатно-солевом буфере (ФСБ), pH 7.4, наносили на свежерасщепленную слюду (SPI Supplies), инкубировали в течение 1 мин, а затем промывали большим количеством деионизированной воды MilliQ.

Лазерная корреляционная спектроскопия. Измерения квазиупругого рассеяния света проводили с использованием ЛКС-спектрометра ЛКС-03 (Intox, Россия). Пептиды L и R растворяли в ФСБ в концентрациях 0.1, 0.2, 0.5, 0.7 и 1 мг/мл. Спектральный анализ проводили с использованием программного обеспечения прибора.

Спектрофотометрия. Концентрацию пептидов определяли с использованием спектрофотометра Hitachi U-3310 в 1 см кварцевых кюветках. Спектры обрабатывали с применением программного обеспечения прибора, коэффициент молярной экстинкции ($1280 \text{ M}^{-1} \text{ cm}^{-1}$) был вычислен в пептидном калькуляторе PepCalc [12].

Малоугловое рассеяние рентгеновских лучей (МУРР). Эксперименты по МУРР проводили на источнике синхротронного излучения ESRF (Гренобль, Франция) на спектрометре ID02 в соответствии с правилами использования [13]. Регистрацию спектров пептидов, растворенных в ФСБ в концентрациях 0.7 мг/мл, проводили с использованием детекторов, расположенных на расстоянии 1.2 и 5 м, в результате диапазон переданных импульсов q составил $0.014\text{--}6 \text{ nm}^{-1}$ ($q = 4\pi \sin(\theta)/\lambda$, где 2θ – угол рассеяния, λ – длина волны). Спектры регистрировали при комнатной температуре. Для контроля радиационных повреждений образца в каждой серии измерений было зарегистрировано 50 кривых МУРР с экспозицией 15 мс и интервалом 1 с. При отсутствии систематических изменений во всех спектрах серии полученные результаты усредняли. Обработка данных включала в себя вычитание сигнала от чистого растворителя (ФСБ), зарегистрированного в аналогичных условиях, а затем устранение некогерентного рассеяния с использованием линейного приближения спектров в координатах Iq^4 vs q^4 , как описано в [14]. Дальнейшую обработку и аппроксимацию экспериментальных данных проводили с помощью программного обеспечения SasView [15] и Origin2015.

Молекулярное моделирование. Модели пептидов были построены с помощью программы Rmold [16]. В дальнейшем с помощью GROMACS [17] была построена система в виде кубического бокса размером 100^3 \AA^3 , содержащая по 32 мономера пептидов TDY**G** и GYDT, а также атомы Na^+ и Cl^- в числе, соответствующем 50 мМ концентрации NaCl. Для моделирования взаимодействия тетрапептидов между собой применяли метод МД в ре-

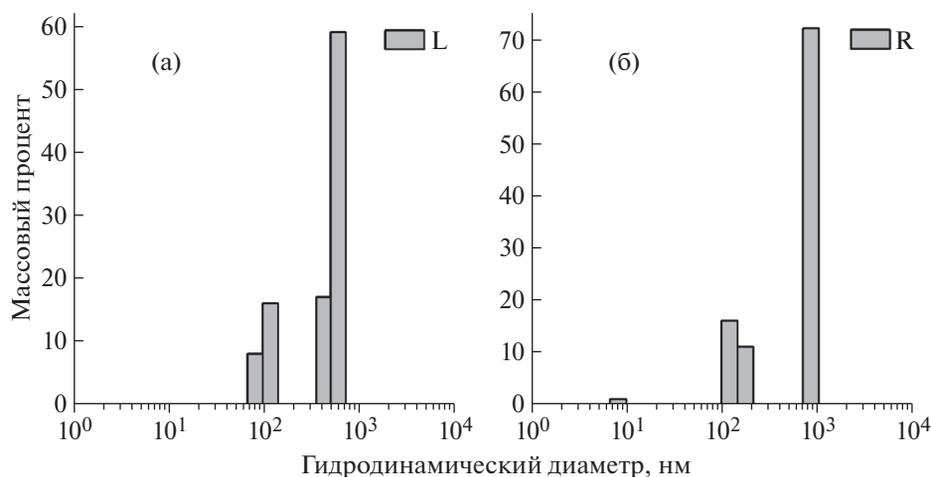


Рис. 1. Результаты обработки спектров динамического светорассеяния пептидов L (0.5 мг/мл) (а) и R (0.5 мг/мл) (б).

жиге свободной диффузии. Для этого использовали следующий протокол.

Подготовленную стартовую конфигурацию помещали в бокс с молекулами воды. Размер бокса составлял 100^3 \AA^3 . Для пептидов использовали поле amber14sb [18], для воды — tip3r [19]. Далее следовали этапы минимизации энергии системы и первая стадия уравнивания. Начальные скорости рассчитывали, исходя из распределения Максвелла для температуры 310 К. Чтобы избежать разрушения системы, на все тяжелые атомы пептидов накладывали дополнительный потенциал в виде чаши, который ограничивал их движение. При расчете использовали термостат и баростат Берендсена [20]. Уравнивание растворителя проводили в течение 5 нс симуляции с почти неподвижным белком. Последнюю полученную конфигурацию использовали в качестве стартовой на второй стадии уравнивания, при этом снимали действие потенциалов, ограничивающих движение атомов. Использовали термостат Нозе-Гувера [21–24] и баростат Паринелло-Рамана [25, 26]. Уравнивание проводили в течение 10 нс. Далее проводили МД-симуляцию в течение 500 нс. Результаты данного этапа моделирования использовали в дальнейшем для анализа.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Наличие критической концентрации индуктора конформационных переходов, ниже которой он не оказывал эффекта на четвертичную структуру полипептида, наблюдаемое при изучении воздействия пептидов L и R на пептид WT, было обнаружено и ранее при изучении низкомолекулярных индукторов конформационных переходов [27]. Тогда было показано, что активность низкомолекулярного индуктора триазавирина

проявляется только в виде супрамолекулярных комплексов, образующихся при превышении определенной критической концентрации. Этот эффект был обусловлен необходимостью кооперативности взаимодействия индуктор–полипептид для воздействия на конформацию; кооперативность обуславливалась существованием индуктора в виде нековалентно связанных линейных гомоолигомеров [28]. Вероятно, пептиды L и R в концентрациях, превышающих критическую, существуют в растворе в виде супрамолекулярных комплексов, способных модулировать четвертичную структуру пептида WT. В пользу существования олигомерных форм пептидов R и L свидетельствовало наличие светорассеяния при превышении критических концентраций для пептидов (0.5 мг/мл для L, 0.2 мг/мл для R). Результаты обработки спектров приведены на рис. 1.

Различие в критических концентрациях может быть обусловлено наличием внутримолекулярных взаимодействий в молекулах пептида L (что следует из результатов МД-моделирования). Повидимому, при образовании структур более высокого порядка такие взаимодействия могут препятствовать олигомеризации при более низких концентрациях, характерных для пептида R, образующего преимущественно межмолекулярные связи.

При исследовании пептида R методом гель-фильтрации молекулярная масса олигомеров соответствовала тетрамеру пептида, и более высокомолекулярные формы не детектировались. Исследование надмолекулярных структур тетрапептидов-индукторов методом электронной микроскопии [10] не показало наличия олигомеров пептида R. Результаты АСМ образцов пептидов L и R свидетельствовали об отсутствии упорядоченных структур на подложке и были сравнимы с результатами микроскопии чистого буфера

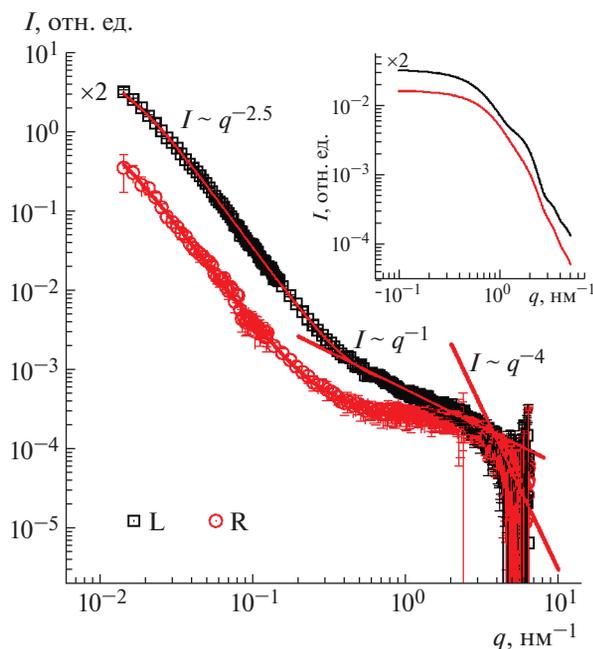


Рис. 2. Спектры МУРР на растворах пептидов L (квадраты) и R (круги). На врезке расчетные спектры олигомерных агрегатов, построенных методом молекулярного моделирования (см. подпись к рис. 4). Интенсивность рассеяния (I) спектра пептида L была умножена на 2 для наглядного представления данных.

(данные не показаны). В то же время все эти результаты не позволяли объяснить факт наличия критической концентрации индуктора фибриллогенеза.

Результаты МУРР (рис. 2) подтверждают присутствие в растворах обоих пептидов в концентрации выше критической крупных структур. В диапазоне малых амплитуд переданных импульсов ($q < 0.3 \text{ nm}^{-1}$) зависимость интенсивности рассеяния I от вектора рассеяния q близка к степенной с показателем степени 2.5. Данные в этом диапазоне могут быть аппроксимированы моделью рассеяния масс-фракталом ($D = 2.50 \pm 0.02$ и $D = 2.42 \pm 0.10$ для пептидов L и R соответственно) ограниченного размера [29], при этом характерные корреляционные длины ξ для агрегатов, образуемых пептидами L и R, оказываются близки между собой (70 ± 5 и около 110 нм соответственно) и находятся в хорошем соответствии с результатами динамического светорассеяния. Для пептида L характерный размер отдельных рассеивающих объектов составлял $\sim 10\text{--}15$ нм.

В диапазоне более высоких амплитуд переданного импульса ($q = 0.5\text{--}1.2 \text{ nm}^{-1}$) интенсивность рассеяния для пептида L в целом следует степенному закону $I \sim q^{-1}$, что соответствует линейным структурам, которые могут образовываться за счет взаимодействия между мономерами пептидов, формируя олигомерную цепь. Изолированные линейные олигомеры пептидов, полученные

в результате молекулярного моделирования, спектры которых приведены на рис. 2 (врезка), в эксперименте не наблюдаются и, как можно предположить, не являются устойчивыми, ассоциируя в более крупные глобулярные агрегаты.

Отметим, что хроматографический анализ растворов пептидов L и R в концентрациях, превышающих критические, показывал наличие в растворе только тетрамеров указанных пептидов, причем спектрофотометрическое определение концентрации пептидов в растворе после разделения свидетельствовало о сохранении количества пептидов в процессе хроматографии. Таким образом, все детектируемые при помощи динамического светорассеяния и МУРР ассоциаты пептидов диссоциировали в процессе хроматографии на тетрамеры.

Для моделирования олигомеризации пептидов использовали метод МД в режиме свободной диффузии. Данный метод позволяет получить состояния, которые невозможно моделировать с использованием молекулярного докинга, в том числе неустойчивые олигомеры, существующие только в растворе. В каждый бокс для моделирования поместили по 32 молекулы L или R. Несмотря на то что оба пептида в процессе МД-моделирования показали способность к образованию устойчивой трехмерной динамической сетки в растворе за счет взаимодействия друг с другом, анализ вторичной структуры показал различающиеся результаты (рис. 3). Если для L в большей степени характерны структуры мономеров, характеризующиеся как повороты, то в случае R наблюдалось образование структур, соответствующих межмолекулярным β -листам. По всей видимости, способность образовывать различные виды β -структур является основой для образования различных трехмерных сеток пептидов L и R в растворе.

На рис. 4 представлены модели структур 32-мерных пептидов L и R, полученные в результате МД-моделирования в режиме свободной диффузии.

Таким образом, показано, что пептиды L и R способны к образованию супрамолекулярных комплексов в растворе в модели *in silico*, при этом структуры олигомерных комплексов для этих пептидов существенно различаются.

В [30–32] показано, что некоторые короткие пептиды способны к образованию надмолекулярных структур, в том числе амилоидоподобных фибрилл, в растворе. Тот факт, что короткие пептиды способны образовывать конформационно активные супрамолекулярные комплексы [33] вкупе со способностью коротких пептидов к проникновению через многие клеточные барьеры, открывает вероятность существования новых механизмов регуляции активности белков в норме и

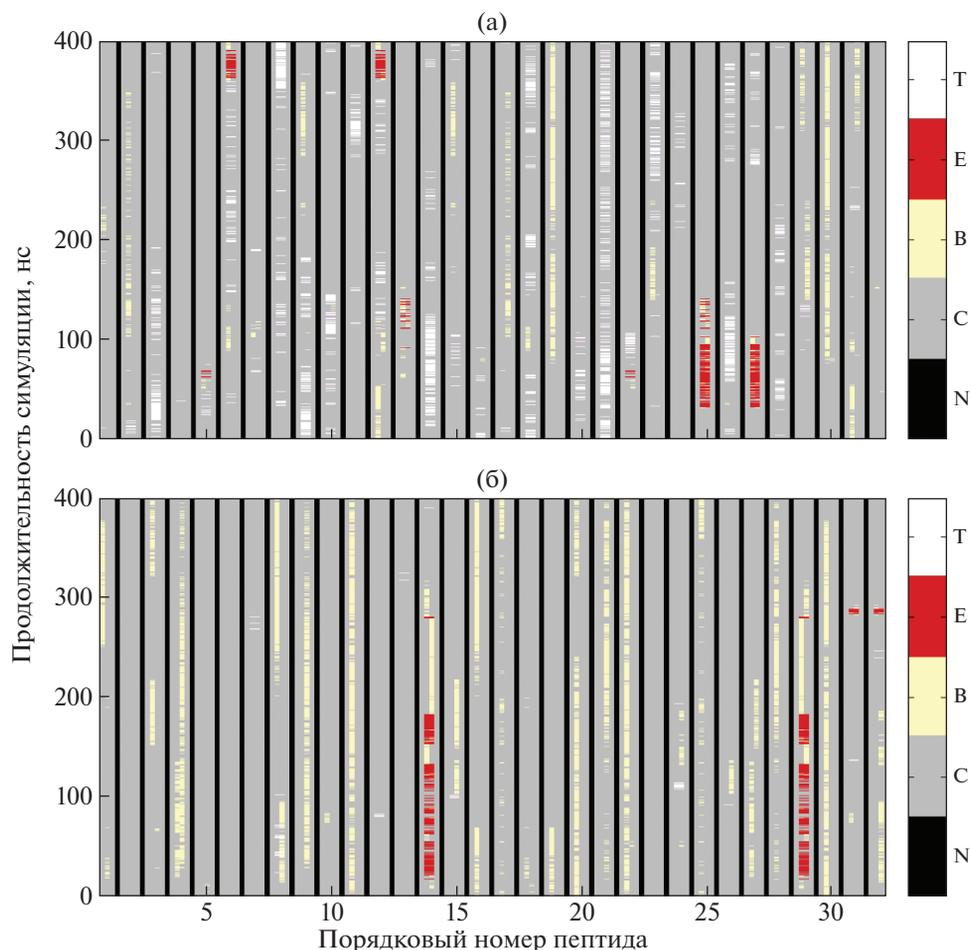


Рис. 3. Карта вторичной структуры для пептидов L (а) и R (б) (Т – бета-изгиб, Е – бета-структура в составе стопки, В – аминокислотный остаток в отдельном бета-мостике, N – неупорядоченная структура, С – клубок). Каждая прямоугольная секция представляет собой результаты молекулярной динамики одного из 32 пептидов, порядковый номер пептидов указан на оси абсцисс с шагом 5.

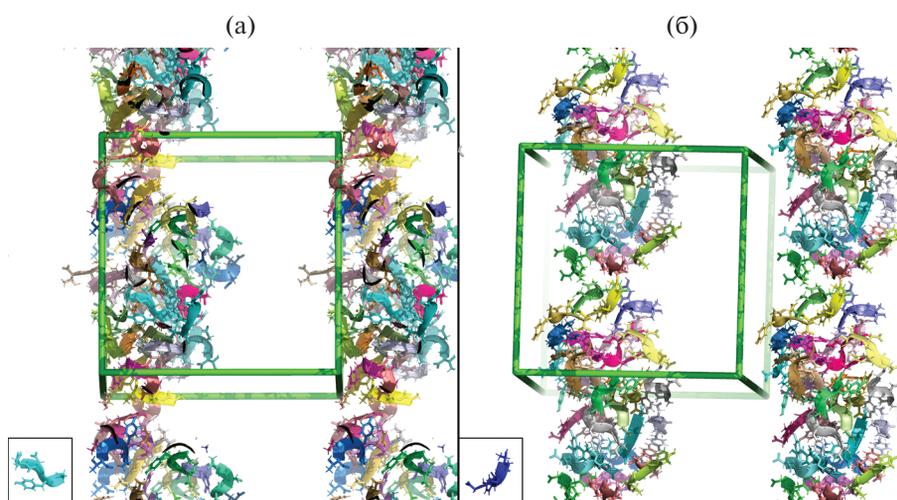


Рис. 4. Модели структур 32-мерных пептидов L (а) и R (б). Элементарная периодическая ячейка размером $100 \times 100 \times 100 \text{ \AA}$ показана в виде куба. В каждой ячейке содержится 32 молекулы пептида L (а) и R (б), представленные комбинацией лент (gibbons) и палочек (sticks). Мономеры окрашены случайным образом в разные цвета, примеры каждого мономера в том же масштабе приведены во врезках на соответствующей части рисунка.

при патологии и возможность разработки новых биологически активных веществ пептидной природы. Метод МУРР позволяет определять способность пептидов к образованию подобных неустойчивых комплексов, а в сочетании с методом МД-моделирования создавать их модели. Полученные с помощью комбинации указанных методов модели супрамолекулярных пептидных комплексов могут быть использованы как при изучении механизма действия препаратов пептидной природы, так и при изучении механизмов действия природных супрамолекулярных пептидных комплексов.

ЗАКЛЮЧЕНИЕ

В активной концентрации пептидные индукторы (тетрапептиды GYDT и TDYG) фибриллогенеза пептида, соответствующего по первичной структуре фрагменту бета-домена альфа-лактальбумина человека (GYDTQAIVENNESTEYG), образуют в растворе олигомеры, не выявляемые при помощи аналитической хроматографии и атомно-силовой микроскопии, но детектируемые при помощи методов динамического рассеяния и МУРР. Результаты МД-моделирования тетрапептидов в растворе согласуются с данными МУРР.

Кривые МУРР получены в ESRF, Гренобль, Франция, эксперимент № LS2508 @ID02. Работа выполнена при поддержке НИЦ “Курчатовский институт” (приказ № 1363 от 25.06.2019 г.).

СПИСОК ЛИТЕРАТУРЫ

1. *Mangione P.P., Porcari R., Gillmore J.D. et al.* // Proc. Natl. Acad. Sci. 2014. V. 111. № 4. P. 1539. <https://doi.org/10.1073/pnas.1317488111>
2. *Andrews S.J., Rothnagel J.A.* // Nat. Rev. Genet. 2014. V. 15. № 3. P. 193. <https://doi.org/10.1038/nrg3520>
3. *Wu H.* // Cell. 2013. V. 153. № 2. P. 287. <https://doi.org/10.1016/j.cell.2013.03.013>
4. *Zabrodskaya Y.A., Lebedev D.V., Egorova M.A. et al.* // Biophys. Chem. 2018. V. 234. P. 16. <https://doi.org/10.1016/j.bpc.2018.01.001>
5. *Michiels E., Roose K., Gallardo R. et al.* // Nat. Commun. 2020. V. 11. № 1. P. 1. <https://doi.org/10.1038/s41467-020-16721-8>
6. *Gallardo R., Ramakers M., De Smet F. et al.* // Science. 2016. V. 354. № 6313. P. aah4949. <https://doi.org/10.1126/science.aah4949>
7. *Egorov V.V., Solovyov K.V., Grudinina N.A. et al.* // Protein Pept. Lett. 2007. V. 14. № 5. P. 471. <https://doi.org/10.2174/092986607780782858>
8. *Egorov V.V., Гармай Ю.П., Соловьев К.В. и др.* // Докл. РАН. 2007. Т. 414. № 6. С. 828.
9. *Tuohy V., Jaini R., Johnson J. et al.* // Cancers. 2016. V. 8. № 6. P. 56. <https://doi.org/10.3390/cancers8060056>
10. *Egorov V.V., Lebedev D.V., Shaldzhyan A.A. et al.* // Prion. 2014. V. 8. № 5. P. 369. <https://doi.org/10.4161/19336896.2014.983745>
11. *Кадочников В.В., Егоров В.В., Швецов А.В. и др.* // Кристаллография. 2016. V. 61. № 1. P. 107. <https://doi.org/10.7868/s0023476116010082>
12. PepCalc.com – Peptide calculator. <https://pepcalc.com/>
13. *Narayanan T., Sztucki M., Van Vaerenbergh P. et al.* // J. Appl. Cryst. 2018. V. 51. № 6. P. 1511. <https://doi.org/10.1107/S1600576718012748>
14. *Svergun D.I., Koch M.H.J.* // Reports Prog. Phys. 2003. V. 66. № 10. P. 1735. <https://doi.org/10.1088/0034-4885/66/10/R05>
15. SasView – Small Angle Scattering Analysis. <http://www.sasview.org/>
16. LLC Schrödinger. The PyMOL Molecular Graphics System, Version 2.4. 2020. <https://pymol.org/2/>
17. *Abraham M.J., Murtola T., Schulz R. et al.* // SoftwareX. 2015. V. 1–2. P. 19. <https://doi.org/10.1016/J.SOFTX.2015.06.001>
18. *Maier J.A., Martinez C., Kasavajhala K. et al.* // J. Chem. Theory Comput. 2015. V. 11. № 8. P. 3696. <https://doi.org/10.1021/acs.jctc.5b00255>
19. *Jorgensen W.L., Chandrasekhar J., Madura J.D. et al.* // J. Chem. Phys. 1983. V. 79. № 2. P. 926. <https://doi.org/10.1063/1.445869>
20. *Berendsen H.J.C., Postma J.P.M., van Gunsteren W.F. et al.* // J. Chem. Phys. 1984. V. 81. № 8. P. 3684. <https://doi.org/10.1063/1.448118>
21. *Nosé S.* // J. Chem. Phys. 1984. V. 81. № 1. P. 511. <https://doi.org/10.1063/1.447334>
22. *Hoover W.G.* // Phys. Rev. A. 1985. V. 31. № 3. P. 1695. <https://doi.org/10.1103/PhysRevA.31.1695>
23. *Martyna G.J., Klein M.L., Tuckerman M.* // J. Chem. Phys. 1992. V. 97. № 4. P. 2635. <https://doi.org/10.1063/1.463940>
24. *Martyna G.J., Tuckerman M.E., Tobias D.J., Klein M.L.* // Mol. Phys. 1996. V. 87. № 5. P. 1117. <https://doi.org/10.1080/00268979600100761>
25. *Parrinello M., Rahman A.* // J. Appl. Phys. 1981. V. 52. № 12. P. 7182. <https://doi.org/10.1063/1.328693>
26. *Nosé S., Klein M.L.* // Mol. Phys. 1983. V. 50. № 5. P. 1055. <https://doi.org/10.1080/00268978300102851>
27. *Shvetsov A.V., Zabrodskaya Y.A., Nekrasov P.A., Egorov V.V.* // J. Biomol. Struct. Dyn. 2018. V. 36. № 10. P. 2694. <https://doi.org/10.1080/07391102.2017.1367329>
28. *Zabrodskaya Y.A., Shvetsov A.V., Tsvetkov V.B., Egorov V.V.* // J. Biomol. Struct. Dyn. 2019. V. 37. № 12. P. 3041. <https://doi.org/10.1080/07391102.2018.1507837>
29. *Sorensen C.M., Cai J., Lu N.* // Langmuir. 1992. V. 8. № 8. P. 2064. <https://doi.org/10.1021/la00044a029>
30. *Gazit E.* // Prion. 2007. V. 1. № 1. P. 32. <https://doi.org/10.4161/pri.1.1.4095>
31. *Rechtes M., Porat Y., Gazit E.* // J. Biol. Chem. 2002. V. 277. № 38. P. 35475. <https://doi.org/10.1074/jbc.M206039200>
32. *Haspel N., Zanuy D., Ma B. et al.* // J. Mol. Biol. 2005. V. 345. № 5. P. 1213. <https://doi.org/10.1016/j.jmb.2004.11.002>
33. *Zhou J., Du X., Xu B.* // Prion. 2015. V. 9. № 2. P. 110. <https://doi.org/10.1080/19336896.2015.1022021>

СТРУКТУРА МАКРОМОЛЕКУЛЯРНЫХ
СОЕДИНЕНИЙ

УДК 577.322.54

ИЗУЧЕНИЕ КОНФОРМАЦИОННОЙ ПОДВИЖНОСТИ GroEL
МЕТОДАМИ КРИОЭЛЕКТРОННОЙ МИКРОСКОПИИ
И МОЛЕКУЛЯРНОЙ ДИНАМИКИ

© 2021 г. И. С. Панина¹, А. А. Мамчур², И. А. Ярошевич², Д. В. Зленко², Е. Б. Пичкур³,
С. С. Кудрявцева⁴, В. И. Муронец⁴, О. С. Соколова², Т. Б. Станишнева-Коновалова^{2,*}

¹ Институт биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова, Москва, Россия

² Московский государственный университет им. М.В. Ломоносова, Биологический факультет, Москва, Россия

³ Национальный исследовательский центр “Курчатовский институт”, Москва, Россия

⁴ Московский государственный университет им. М.В. Ломоносова, Факультет биоинженерии и биоинформатики,
Москва, Россия

*E-mail: stanishneva-konovalova@mail.bio.msu.ru

Поступила в редакцию 06.08.2020 г.

После доработки 06.08.2020 г.

Принята к публикации 25.03.2021 г.

Бактериальный шаперонин GroEL представляет собой сложный белковый олигомер кольцевой структуры, способствующий сворачиванию других белков путем их инкапсуляции в полости. Существует крайне мало структурной информации о неупорядоченном С-концевом фрагменте субъединиц GroEL, который участвует в процессе сворачивания субстратного белка. Представлена 3D-реконструкция апо-формы GroEL, полученная с помощью криоэлектронной микроскопии (крио-ЭМ) с разрешением 3.02 Å и дополненная расчетами методом молекулярной динамики (МД). Результаты крио-ЭМ и МД хорошо согласуются и демонстрируют различную подвижность доменов субъединиц белка. Данные МД предсказывают динамику и сеть внутримолекулярных контактов С-концевых участков белка. Эти результаты имеют большое значение для последующего изучения механизма сворачивания белков в полости GroEL.

DOI: 10.31857/S0023476121050167

ВВЕДЕНИЕ

Шапероны – белки, отвечающие за сворачивание других макромолекул и обеспечивающие образование их нативной трехмерной структуры. Шапероны также называют белками теплового шока (англ. Heat Shock Proteins, HSP), поскольку их активность сильно возрастает при повышении температуры [1]. В группу шаперонов входят белки, имеющие схожие функции, но различающиеся по структуре и специфичности к субстрату [2]. Одним из семейств шаперонов являются шаперонины (HSP60). Обычно они представляют собой структуру из двух колец, каждое из которых содержит несколько субъединиц, различающихся по строению и расположению у разных организмов [3].

Выделяют две группы шаперонинов. К первой относятся шаперонины бактерий, а также шаперонины хлоропластов и митохондрий. Комплексы этой группы образуют симметричные кольца, каждое из которых состоит из семи одинаковых субъединиц. Шаперонины первой группы в своем функциональном цикле взаимодействуют с

кошаперонином, образующим “крышку” над полостью кольца и инкапсулирующим белок-субстрат. Вторая группа шаперонинов объединяет белки, представленные в цитозоле эукариот и архей. Шаперонины второй группы тоже состоят из двух колец, однако каждое из них включает в себя восемь или девять субъединиц. При этом субъединицы внутри одного кольца могут различаться по структуре [4]. Кроме того, некоторые бактериофаги имеют в геноме последовательности, кодирующие ортологи GroEL [5, 6]. Примечательно, что шаперонины бактериофагов могут значительно отличаться от бактериальных: шаперонин бактериофага ОВР *Pseudomonas fluorescens* представляет собой асимметричную однокольцевую структуру из семи субъединиц [7].

Наиболее изученным шаперонином является прокариотический комплекс GroEL/GroES, который в нормальных условиях широко представлен в клетках *E. coli*. Он взаимодействует с ненативными конформациями различных белков, предотвращая их неправильное сворачивание и агрегацию, и обладает слабой АТФ-азной актив-

ностью [8]. По своей структуре бактериальный шаперонин GroEL представляет собой сложный олигомерный белковый комплекс, состоящий из 14 идентичных субъединиц с молекулярной массой (ММ) 57 кДа каждая, объединенных в два кольца по семь субъединиц. Для осуществления своей функции ему необходимо взаимодействовать с GroES [9], состоящим из семи идентичных субъединиц с ММ 10 кДа каждая, объединенных в куполообразную кольцевую структуру. Каждая субъединица гептамерного кольца GroEL состоит из трех доменов: апикального, промежуточного и экваториального. Апикальный домен взаимодействует с белками-субстратами и GroES, а экваториальный связывает АТФ [10].

Между аминокислотными остатками GroEL формируются контакты, часть из которых меняется в ходе функционального цикла. В апо-форме (т.е. форме без нуклеотидов и белка-субстрата) связь колец друг с другом обеспечивают солевые мостики между остатками экваториальных доменов: Lys105-Ala109 и Glu461-Arg452 [11–15]. Также солевыми мостиками связаны соседние субъединицы внутри одного кольца (Arg197-Glu386). Считается, что этот контакт, стабилизирующий структуру олигомера в апо-форме, разрушается при связывании АТФ, что запускает конформационные перестройки белкового комплекса [16, 17]. Кроме того, существует несколько внутрисубъединичных солевых мостиков: Asp83-Lys327, Glu209-Arg58 и Glu409-Arg501. Первые два мостика наиболее характерны для шаперонинов в связанном с нуклеотидом состоянии, но в отсутствие кошаперонина. Такие структуры более открыты, чем апо-форма (апикальный домен сильнее удален от экваториального), но конформационные перестройки еще не завершены. Контакт между аминокислотами Glu409 и Arg501 является очень стабильным и не разрушается даже при связывании АТФ и переходе в наиболее открытую конформацию субъединицы [16, 17].

Выяснить структуру и механизм работы комплекса GroEL/GroES во многом помогли методы рентгеноструктурного анализа и криоэлектронной микроскопии (крио-ЭМ), с помощью которых было получено более 150 структур комплекса (по данным базы Protein Data Bank), в том числе на разных стадиях функционального цикла. Однако у структурных методов есть ряд ограничений, в частности они не могут расшифровать подвижные участки молекул. У GroEL 23 аминокислоты С-конца каждой субъединицы образуют подвижный хвост, который выходит из экваториального домена в полость кольца [18]. Этот участок отсутствует на реконструкциях комплекса, но важен для его функционирования. Последние 13 аминокислот С-конца представляют собой четыре повтора последовательности Gly-Gly-Met и одну аминокислоту метионин (гидрофобная по-

следовательность [GGM]₄M), а предыдущие аминокислоты содержат отрицательные заряды. Гидрофобная последовательность является консервативным участком для всех гомологов GroEL [19]. Удаление этой последовательности не вызывает гибели бактерий, однако снижает их устойчивость к тепловому шоку [20]. Замена последовательности на [AAA]₄A привела к тому, что скорость сворачивания субстрата была ниже, чем для дикого типа, но больше, чем для мутанта с полностью удаленными С-концами [21]. Исследование гидрофильной части С-концевого фрагмента (аминокислоты 526–531), в котором гидрофильный участок заменялся на нейтральный или гидрофобный, продемонстрировало, что мутанты значительно медленнее восстанавливали нативную структуру субстратного белка роданазы [22]. В других работах было показано, что при удалении С-концов значительно снижалась скорость сворачивания субстратов Рубиско [23] и GFP [24]. Хотя структурные методы не дали результатов об атомарной структуре С-концевых участков, они позволили обнаружить некоторые закономерности в их динамике. В [25] методом крио-ЭМ показано, что С-концевые фрагменты *цис*- и *транс*-колец отклоняются друг от друга и взаимодействуют с субстратными красителями, которые использовались в данном эксперименте. В том же исследовании при рассмотрении комплекса GroEL–Рубиско отмечено, что субстратный белок контактирует с апикальными доменами и С-концами.

Вычислительные методы, включая метод молекулярной динамики (МД), могут выступать в качестве дополнения к структурным, моделируя подвижные участки. С помощью МД ранее проводилось изучение С-концевых фрагментов GroEL и было показано, как меняется положение С-концов в ходе функционального цикла и как на него влияет присутствие нуклеотидов [26]. В этом исследовании использовались АТФ- и АДФ-связанные кристаллические структуры GroEL, а положения всех остатков, кроме С-концевых, были зафиксированы. В настоящей работе использована полученная авторами структура апо-формы GroEL дикого типа по данным крио-ЭМ, построена ее атомарная модель с С-концевыми участками и проведен расчет МД-системы без дополнительных ограничений. Результаты МД указывают на различия в подвижности доменов субъединиц, что согласуется с вариациями локального разрешения структуры из крио-ЭМ, а также описывают динамику С-концевых участков в апо-форме.

МЕТОДЫ

Наработка и очистка GroEL. Продукцию шаперонина GroEL осуществляли, используя замороженную бактериальную культуру W3110 *E. coli* с

плазмидой рOF 39. Для возобновления культуры 50 мл бактериальной среды LB (Luria Broth, Sigma), содержащей 50 мкг/мл ампициллина, заражали замороженной культурой и растили в течение ночи при 37°C и перемешивании 200 об./мин. Далее брали шесть колб с 200 мл среды LB, содержащей 50 мкг/мл ампициллина, и добавляли ночную культуру до оптической плотности 0.1 при 600 нм. Продукцию GroEL проводили в течение ночи при 37°C и перемешивании 200 об./мин. Полученную культуру центрифугировали 20 мин при 5000 g и 4°C. Далее отмывали бактериальные клетки (осадок) 50 мМ Трис-НСl-буфером, рН 8.0, от среды, а затем суспендировали в лизирующем буфере (100 мМ Tris-НСl, рН 8.1), содержащем 0.1 мМ ЭДТА, 10 мМ ДТТ и 0.2 мг/мл ингибиторов протеаз (Sigma). Бактериальные клетки разрушали с помощью ультразвуковой обработки (Fisher Bioblock, Illkirch, Франция) пятью импульсами по 40 с при амплитуде 50%. Суспензию центрифугировали 30 мин при 11 000 об./мин, после чего последовательно высаливали супернатант сухим (NH₄)₂SO₄ до 30 и 80%. Далее центрифугировали препарат при тех же условиях и растворяли осадок в буфере В (50 мМ Tris-НСl, рН 7.2, 0.1 мМ ЭДТА, 2 мМ ДТТ). Полученный раствор диализовали против буфера В в течение ночи. Диализат наносили на колонку с DEAE-Sephrose fast flow (Sigma), уравновешенную буфером В. Элюцию белка проводили градиентом 0–500 мМ NaCl. Хроматографию проводили со скоростью 1–3 мл/мин на хроматографической системе Akta Prime с программным обеспечением Unicorn Control (Amersham Biosciences, Piscataway, США). Собирали фракции в объеме 8 мл и анализировали их методом электрофореза в полиакриламидном геле в присутствии SDS. Полученный препарат GroEL, содержащий ~350 мМ NaCl (стабилизирует белок), нагревали в круглодонной колбе на водяной бане до 58°C, после чего остужали до комнатной температуры и добавляли Mg-АТР до конечной концентрации 2 мМ. Затем прогревали смесь до той же температуры в течение 2–3 мин. Данная стадия позволяет GroEL пройти свой естественный цикл и выпустить в среду денатурированные белки *E. coli*, которые могли сохраниться в полости шаперонина во время выделения. Денатурировавшие белки осаждали центрифугированием в течение 10 мин при 11 000 g. Препарат GroEL диализовали против буфера В в течение ночи, а затем повторно очищали на DEAE-Sephrose fast flow по указанной выше методике. Собирали фракции в объеме 6 мл и также анализировали их методом электрофореза в полиакриламидном геле в присутствии SDS. Фракции, содержащие очищенный GroEL, объединяли. Очищенный препарат GroEL высаливали сухим (NH₄)₂SO₄ до 80% и хранили при 4°C.

Получение структуры GroEL по данным крио-ЭМ. Данные крио-ЭМ получены в Ресурсном центре зондовой и электронной микроскопии НИЦ “Курчатовский институт”. Вначале поддерживающие сетки для электронной микроскопии с периодическими отверстиями в аморфной пленке углерода (Quantifoil R1.2/1.3, Quantifoil) были обработаны в тлеющем разряде с помощью установки PELCO easiGlow при стандартных условиях: время обработки образца – 30 с, сила тока – 0.25 мА, остаточное давление в камере – 0.26 мБар. Далее на сетку наносили 3 мкл образца GroEL и проводили его витрификацию в камере установки Vitrobot Mark IV при следующих параметрах: сила сжатия при промакивании (Blot force) – 0 единиц, время промакивания (Blot time) – от 2.5 до 3.5 с, температура в камере – 4.5°C, влажность в камере – 95–100%. Затем замороженные сетки переносили в криоэлектронный микроскоп Titan Krios 60–300, оборудованный высокоэффективным детектором электронов Falcon II, где проводили съемку изображений с использованием программного обеспечения EPU (FEI). Обработку изображений и построение реконструкций проводили с помощью программных пакетов Warp [27] и CryoSPARC [28]. Всего было снято 675 стеков изображений, из которых 590 стеков были отобраны для дальнейшего анализа. При помощи программного пакета Warp провели коррекцию дрейфа, оценку параметров функции передачи контраста (CTF), а также выбор на изображениях одиночных проекций объекта исследования. Из исходных изображений после пре-процессинга были выделены 47 149 проекций одиночных частиц и экспортированы в программный пакет CryoSPARC, где была проведена их двухмерная классификация. Далее выбрали 15 классов, входящие в них частицы были подвергнуты процедуре трехмерной классификации. После чего отобрали 29 700 частиц, из которых строили 3D-реконструкцию GroEL с учетом симметрии D7. Разрешение структуры было оценено по критерию FSC = 0.143 и составило 3.02 Å.

Расчет траекторий молекулярной динамики. Атомарная модель апо-формы GroEL для расчетов МД построена в программе COOT [29] по структуре из крио-ЭМ. 23 остатка С-концевого фрагмента (526–548), не выявленные в плотности, были добавлены вручную с использованием стандартной утилиты Builder программы PyMOL версии 2.2.3 (www.pymol.org). Моделирование проводили с использованием программного пакета Gromacs [30] версии 2020.1 в силовом поле a99SB-disp [31] при постоянной температуре (300 К) и постоянном давлении (1 атм) с применением периодических граничных условий. Температура и давление поддерживались постоянными с помощью алгоритмов V-rescale [32] и Parrinello-Rahman [33] соответственно. Белок помещали в ячей-

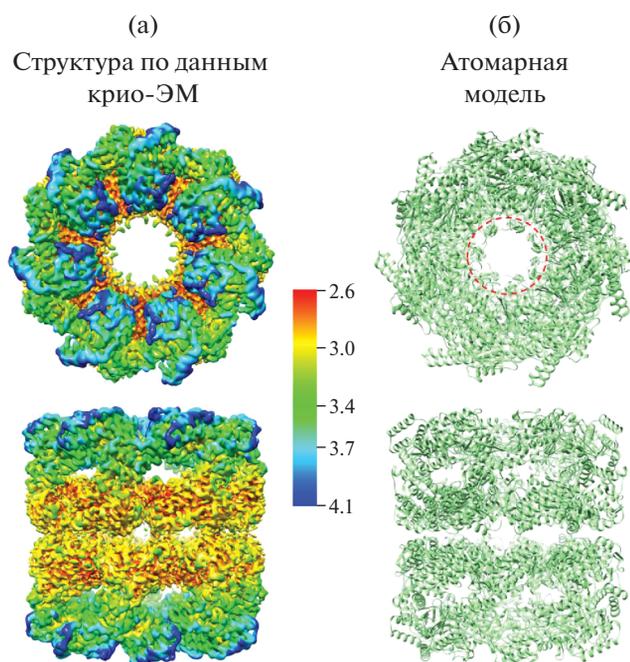


Рис. 1. Структура шаперонина GroEL по данным крио-ЭМ, окрашенная в соответствии с локальным разрешением (а). Атомарная модель, использованная в расчетах МД (б). Пунктирная окружность указывает на положение достроенных С-концевых участков.

ку с водой (модель TIP4P-D) $17 \times 17 \times 18$ нм, содержащей 150 мМ NaCl, включая противоионы для нейтрализации суммарного заряда белка (519 Na^+ и 253 Cl^-). Электростатические взаимодействия рассчитывали с помощью алгоритма PME с радиусом отсечки 1.2 нм. Радиус отсечки для ван-дер-ваальсовых сил составлял 1.2 нм. Временной шаг интегрирования составил 2 фс. Перед расчетом МД была проведена релаксация системы, включающая этап минимизации энергии с последующим нагреванием системы от 5 до 300 К в течение 5 нс. Длительность траектории МД – 250 нс.

Значения RMSD и RMSF для $\text{C}\alpha$ -атомов, а также попарное расстояние между остатками или атомами рассчитаны с помощью стандартных утилит пакета Gromacs. В качестве критерия существования контакта (по типу “солевой мостик”) между аминокислотами выбран порог 0.3 нм, который соответствует сумме ван-дер-ваальсовых радиусов $sp^3 \text{ N}$ и $sp^3 \text{ O}$ [34]. Программное обеспечение MDLovoFit использовали для дифференциального расчета RMSD [35]. С помощью MDLovoFit проводили анализ флуктуаций структуры посредством выравнивания всех подмножеств $\text{C}\alpha$ -атомов белка, соответствующих различным долям ϕ – от общего числа $\text{C}\alpha$, и поиска подгруппы с наименьшим значением RMSD.

Для сравнения результатов крио-ЭМ и молекулярной динамики на основании траектории МД построена 3D-карта рассеивающей плотности шаперонина GroEL. Для этого был создан алгоритм для Python 3.6.9 [36], который переводит множество кадров МД в единую 3D-матрицу. Измерения матрицы соответствуют абсциссе, ординате и аппликату пространства, которое содержит координаты шаперонина GroEL, а каждый элемент матрицы пропорционален вероятности нахождения в соответствующем объеме (вокселе) атомов молекулярной модели. В данной работе размер вокселя матрицы соответствовал $1 \times 1 \times 1 \text{ \AA}^3$, а в формировании значений элементов матрицы участвовали только тяжелые атомы молекулярной модели. Для усреднения использовали 25000 кадров молекулярной траектории, которые были предварительно выровнены прогрессивным алгоритмом программного пакета Gromacs. 23 аминокислоты, расположенные на С-концах субъединиц, при выравнивании не учитывали.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Среднее разрешение структуры тетрадекамера GroEL, полученной по данным крио-ЭМ, составило 3.02 \AA . На рис. 1а представлена поверхность GroEL, окрашенная в соответствии с локальным разрешением, которое варьирует от 2.5 до 4.1 \AA . Значения от 2.5 до 3 \AA характерны для области экваториальных доменов, ответственных за межкольцевые взаимодействия и связывание АТФ. В зоне апикальных доменов, связывающих белки-субстраты и GroES, разрешение падает до 4.1 \AA . Эти особенности распределения локального разрешения структуры согласуются с результатами [37] и указывают на большую подвижность апикального домена относительно экваториального. С помощью полученной карты рассеивающей плотности построили атомарную модель тетрадекамера GroEL. 23 аминокислоты С-конца были достроены и добавлены каждой субъединице (рис. 1б).

С этой моделью провели расчет траектории МД длиной 250 нс. На основании траектории была построена 3D-карта рассеивающей плотности шаперонина GroEL (рис. 2), которая аналогично плотности из крио-ЭМ отражает вероятность обнаружить атом (рассеивающий центр) в соответствующей позиции структуры. Карта рассеивающей плотности, полученная из траектории МД, согласуется с картой плотности из крио-ЭМ, демонстрируя большие значения элементов матрицы в области экваториального домена по сравнению с апикальным. Согласие между результатами крио-ЭМ, полученными усреднением структуры по ансамблю частиц, и результатами МД, полученными усреднением по времени, указывает на

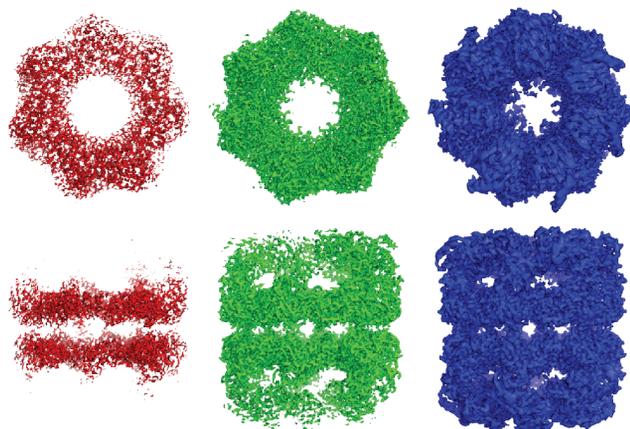


Рис. 2. Карта рассеивающей плотности, полученная на основании траектории МД. Представлены три изометрических поверхности, соответствующие высокому (0.39, слева), среднему (0.21, центр) и низкому (0.11, справа) уровням элементов 3D-матрицы рассеивающей плотности шаперонина GroEL.

достаточность собранного в крио-ЭМ числа частиц и продолжительности расчета МД.

Расчет среднеквадратичных флуктуаций (RMSF) α -атомов каждой субъединицы GroEL в траекториях МД показал, что домены этого белка различаются по своей мобильности, а основной вклад в конформационную подвижность вносят аминокислотные остатки С-конца и апикального домена (рис. 3). Среднее значение RMSF для экваториального домена составляет 0.15 нм, для интермедиального – 0.19 нм, для апикального – 0.27 нм, а для С-концевых остатков превышает 1 нм. Согласно полученным данным наиболее высокая величина флуктуаций упорядоченных частей белка (за исключением С-концов) наблюдается для остатков β -листов 6 и 7 и α -спиралей К

и L, расположенных в начале и конце апикального домена соответственно, и может достигать значения 0.9 нм.

Значения среднеквадратичного отклонения (RMSD) координат α -атомов от их положения в начальной структуре для всех 14 субъединиц по результатам МД находятся в пределах от 0.27 до 0.33 нм. Отклонения от начальной позиции атомов апикального домена существенно превышают таковые для экваториального и интермедиального доменов: средние значения RMSD составляют 0.27, 0.19 и 0.17 нм соответственно. Чтобы выяснить, является ли вычисленное значение RMSD апикального домена следствием большего отклонения его начального (стартового) состояния от равновесного (достигнутого в ходе МД) или его большей подвижности, дополнительно был проведен анализ RMSD атомов относительно их положения в предшествующей структуре отдельно для каждого домена, а также выравнивание всевозможных подмножеств атомов одной субъединицы с целью выявления наиболее изменчивых областей с использованием программного пакета MDLoFit [35]. Проведенный анализ позволил выявить полный набор α -атомов одной субъединицы с наименьшим RMSD. Было установлено, что по меньшей мере 60% атомов каждой субъединицы формируют “неподвижное ядро” и могут быть наложены на их исходные позиции в пределах 0.2 нм. Как и в [18], в данном моделировании не все субъединицы претерпевали изменения синхронно, что привело к переменному составу подвижной и неподвижной частей. Тем не менее для 11 из 14 субъединиц набор остатков неподвижного ядра совпадает на 79%. Рисунок 4а демонстрирует усредненное изменение во времени субъединицы целиком (средняя линия) и выявленных стабильных (нижняя линия) и высоко подвижных участков (верхняя линия). Ато-

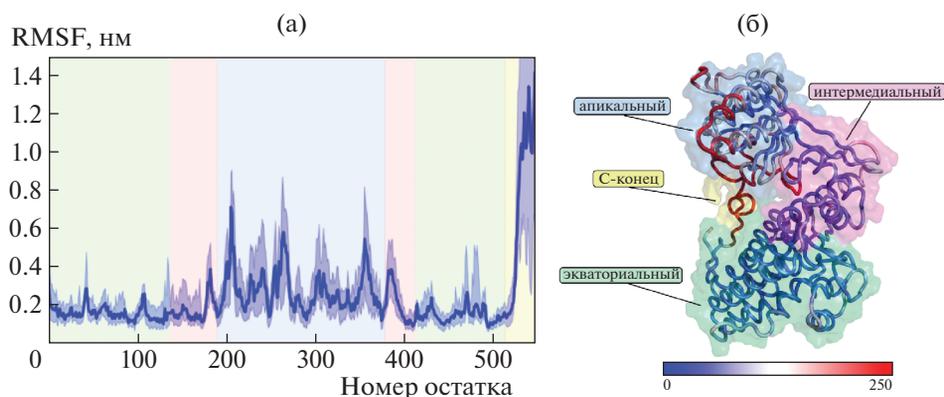


Рис. 3. Различия в подвижности доменов GroEL в ходе траекторий МД. График RMSF α -атомов субъединицы белка (а). Жирной линией показаны усредненные по 14 субъединицам значения. Бледной широкой полосой показан диапазон значений RMSF каждого остатка. Ход основной цепи (показаны только α -атомы) субъединицы GroEL (б), раскрашенный в соответствии со шкалой значений В-фактора (внизу).

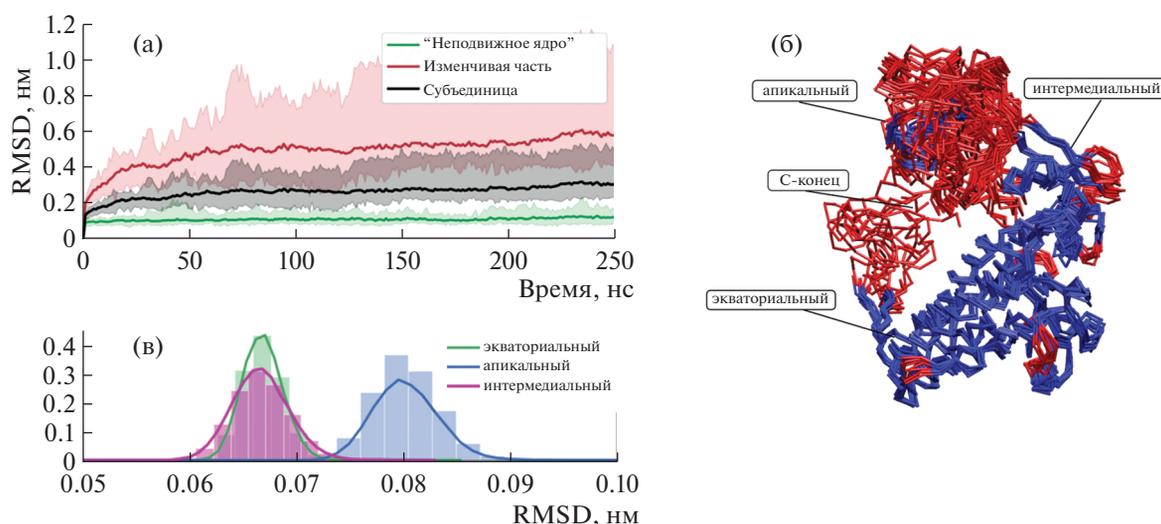


Рис. 4. Оценка RMSD α -атомов одной субъединицы. RMSD как функции времени α -атомов: субъединицы целиком; 60% атомов, составляющих неподвижное ядро субъединицы; оставшиеся 40% атомов, претерпевающих наибольшие изменения. Соответствующим бледным цветом закрашены диапазоны значений RMSD для 14 субъединиц (а). Суперпозиция фреймов МД. Неподвижное ядро показано темным, 40% наиболее подвижных атомов – светлым (б). Структуры ориентированы подобным рис. 3б образом. Видно, что в подвижную часть входят апикальный домен, С-концы и участки петель других доменов. Профиль распределения значений RMSD атомов относительно их положения в предшествующей структуре (Δt 50 пс) экваториального, интермедиального и апикального доменов (в).

мы, входящие в подвижную группу, отклоняются от исходной структуры более чем на 0.6 нм и преимущественно входят в состав апикального домена, а также неупорядоченных С-концов (рис. 4б, светлым). Неподвижное ядро быстро приходит к равновесному положению и не изменяет своей структуры на протяжении проведенной МД (рис. 4б, темным).

Анализ RMSD относительно предшествующего по времени положения атомов показывает различную скорость изменения положения атомов в разных доменах (рис. 4в). За выбранный интервал времени (500 пс) экваториальный и интермедиальный домены изменяют свои структуры в пределах 0.06–0.07 нм, тогда как изменения RMSD

апикального домена лежат в диапазоне 0.07–0.09 нм. Проведенный анализ подвижности шаперонина GroEL по результатам МД дает численные оценки разнородной подвижности доменов этого белка и дополняет качественную картину структурной динамики этой макромолекулы, полученную в рамках крио-ЭМ.

В рамках работы особое внимание уделено динамике пяти физиологически значимых аминокислотных контактов (рис. 5). Несмотря на отсутствие в стартовой модели непосредственного контакта между боковыми цепями внутрисубъединичных пар остатков Arg58–Glu209 и Asp83–Lys327, эти солевые мостики образуются в ходе динамики. Расстояние между парой тяжелых атомов солевого мостика Glu409–Arg501 остается постоянным (0.28 нм) на протяжении всего моделирования. Время жизни солевых мостиков, формируемых между субъединицами одного кольца, составляет ~95% от времени МД.

Стабилизация межмолекулярного взаимодействия колец достигается за счет двух (в каждой паре субъединиц) солевых мостиков Arg452–Glu461. Анализ корреляции между образующимися контактами выявил, что наличие одного из контактов в паре связано с уменьшением вероятности образования другого (табл. 1). Описанное в литературе взаимодействие между кольцами посредством контактов Lys105–Ala109 не было реализовано в ходе МД, однако остатки Lys105–Ala109 одной субъединицы образуют перманентную водородную связь между атомами основной цепи.

Таблица 1. Коэффициенты корреляции (r) для пары контактов Arg452–Glu461 и Glu461–Arg452 соответствующих субъединиц, находящихся в межмолекулярном контакте двух колец

Пара взаимодействующих субъединиц	r
A–K	–0.824
B–J	–0.693
C–I	–0.500
D–H	–0.541
E–N	–0.564
F–M	–0.923
G–L	–0.897

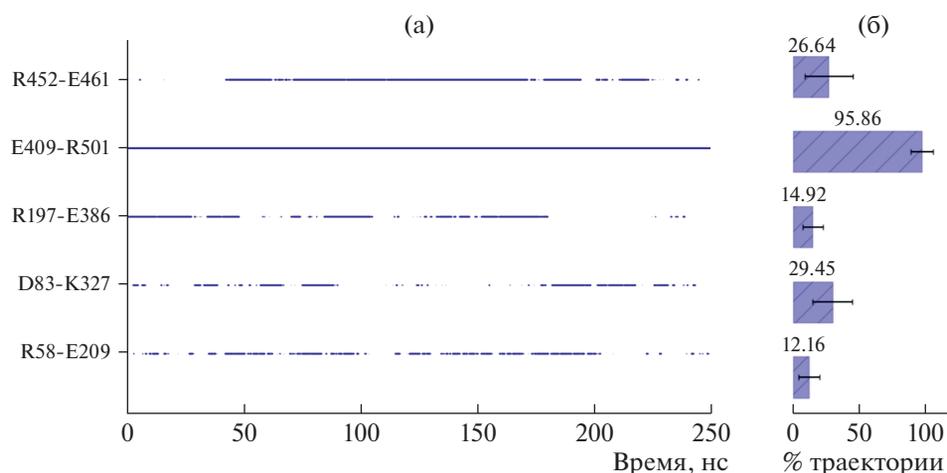


Рис. 5. Наблюдаемые в ходе МД солевые мостики, образованные между доменами и субъединицами GroEL: а – карта солевых мостиков. По оси ординат указана пара остатков, образующих солевой мостик; б – диаграмма среднего для 14 субъединиц времени жизни изучаемых контактов, выраженное в процентах от времени моделирования.

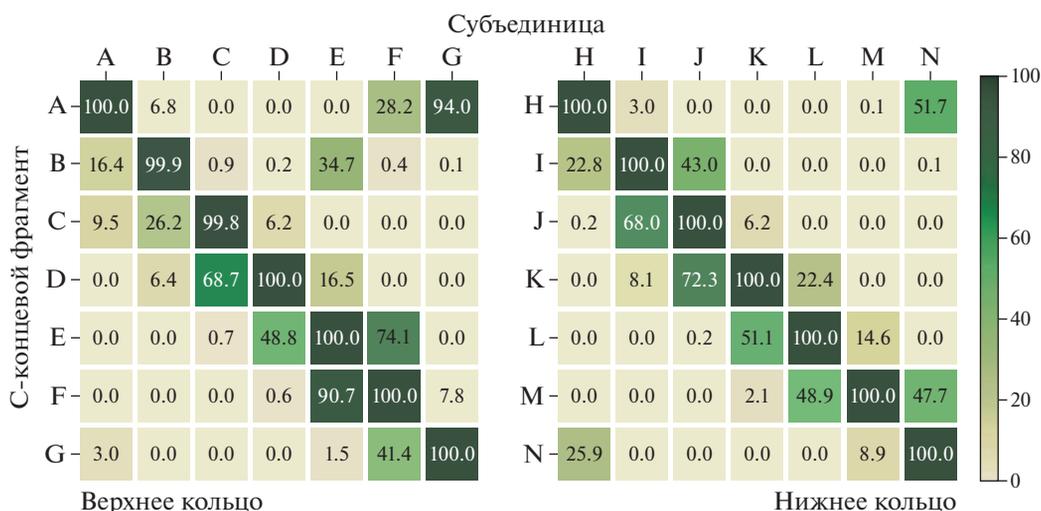


Рис. 6. Тепловая карта матрицы контактов С-концевых фрагментов с субъединицами. Цветовая шкала соответствует заселенности попарных контактов, выраженной в процентах от времени траектории.

В выполненных расчетах высокоподвижные С-концевые фрагменты белка формируют неупорядоченную сеть контактов между собой, а также взаимодействуют с другими доменами субъединиц своего кольца. Карта контактов (рис. 6) показывает, что помимо взаимодействия со своей субъединицей С-концы часто связываются с двумя соседними субъединицами в кольце.

Замечено, что существует различие в частоте образования контактов С-конца субъединицы с двумя соседними субъединицами кольца. С большей вероятностью С-конец связывается с внутренней поверхностью субъединицы, расположенной следующей по часовой стрелке (при взгляде со стороны апикальных доменов). Это наблюдение можно объяснить структурой белка, в которой С-конец структурированной части эква-

ториального домена одной субъединицы расположен ближе к следующей по часовой стрелке субъединице, что делает более вероятным образование контакта. Отметим, что С-концы редко образуют контакты с субъединицами, расположенными более чем через одного соседа, хотя в ходе проведенной МД были обнаружены контакты между удаленными друг от друга субъединицами В и Е.

Проведенное исследование иллюстрирует эффективное взаимодействие между экспериментальным (крио-ЭМ) и теоретическим (МД) подходами для изучения структурной динамики белка на примере шаперонина GroEL. Рассеивающая плотность, полученная в результате исследования крио-ЭМ, является не только отправной точкой для создания молекулярной модели для проведе-

ния МД, но и позволяет качественно оценить относительную подвижность отдельных доменов белкового комплекса. Как показано в работе, такая качественная оценка может быть сравнена с результатом расчета МД и являться способом его верификации. В свою очередь метод МД позволяет оценить количественные характеристики молекулярной подвижности, динамический характер образования и разрыва конкретных связей, а также рассмотреть крайне подвижные части белка, структуру которых невозможно реконструировать при помощи крио-ЭМ.

Исследование выполнено за счет гранта Российского научного фонда (проект № 19-74-20055).

СПИСОК ЛИТЕРАТУРЫ

1. *Maio A.D.* // Shock. 1999. V. 11. № 1. P. 1.
<https://doi.org/10.1097/00024382-199901000-00001>
2. *Bose D., Chakrabarti A.* // IUBMB Life. 2017. V. 69. № 9. P. 647.
<https://doi.org/10.1002/iub.1656>
3. *Skjaerven L., Cuellar J., Martinez A., Valpuesta J.M.* // FEBS Lett. 2015. V. 589. P. 2522.
<https://doi.org/10.1016/j.febslet.2015.06.019>
4. *Yebenes H., Mesa P., Munoz I.G. et al.* // Trends Biochem. Sci. 2011. V. 36. №8. P. 424.
<https://doi.org/10.1016/j.tibs.2011.05.003>
5. *Hertveldt K., Lavigne R., Pleteneva E. et al.* // J. Mol. Biol. 2005. V. 354. № 3. P. 536.
<https://doi.org/10.1016/j.jmb.2005.08.075>
6. *Kurochkina L.P., Semenyuk P.I., Orlov V.N. et al.* // J. Virology. 2012. V. 86. № 18. P. 10103.
<https://doi.org/10.1128/jvi.00940-12>
7. *Stanishneva-Konovalova T.B., Semenyuk P.I., Kurochkina L.P. et al.* // J. Struct. Biol. 2020. V. 209. № 2. P. 107439.
<https://doi.org/10.1016/j.jsb.2019.107439>
8. *Sparrer H., Lilie H., Buchner J.* // J. Mol. Biol. 1996. V. 258. №1. P. 74.
<https://doi.org/10.1006/jmbi.1996.0235>
9. *Grallert H., Buchner J.* // J. Struct. Biol. 2001. V. 135. № 2. P. 95.
<https://doi.org/10.1006/jsbi.2001.4387>
10. *Horwich A.L., Farr G.W., Fenton W.A.* // Chem. Rev. 2006. V. 106. № 5. P. 1917.
<https://doi.org/10.1021/cr040435v>
11. *Lorimer G.H., Fei X., Ye X.* // Philos. Trans. R. Soc. London. B. 2018. V. 373 (1749). P. 20170179.
<https://doi.org/10.1098/rstb.2017.0179>
12. *Ludtke S.J., Baker M.L., Chen D.-H. et al.* // Structure. 2008. V. 16 (3). P. 441.
<https://doi.org/10.1016/j.str.2008.02.007>
13. *Ranson N.A., Clare D.K., Farr G.W. et al.* // Nat. Struct. Mol. Biol. 2006. V. 13 (2). P. 147.
<https://doi.org/10.1038/nsmb1046>
14. *Saibil H.R., Fenton W.A., Clare D.K., Horwich A.L.* // J. Mol. Biol. 2013. V. 425. № 9. P. 1476.
<https://doi.org/10.1016/j.jmb.2012.11.028>
15. *Sot B., Galán A., Valpuesta J.M. et al.* // J. Biol. Chem. 2002. V. 277. № 37. P. 34024.
<https://doi.org/10.1074/jbc.m205733200>
16. *Ma J., Sigler P.B., Xu Z., Karplus M.* // J. Mol. Biol. 2000. V. 302. № 2. P. 303.
<https://doi.org/10.1006/jmbi.2000.4014>
17. *Ranson N.A., Farr G.W., Roseman A.M. et al.* // Cell. 2001. V. 107. № 7. P. 869.
[https://doi.org/10.1016/s0092-8674\(01\)00617-1](https://doi.org/10.1016/s0092-8674(01)00617-1)
18. *Piana S., Shaw D.E.* // J. Phys. Chem. B. 2018. V. 122. № 49. P. 11440.
<https://doi.org/10.1021/acs.jpcc.8b07366>
19. *Brocchieri L., Karlin S.* // Protein Science. 2008. V. 9. № 3. P. 476.
<https://doi.org/10.1110/ps.9.3.476>
20. *Mclennan N.F., Girshovich A.S., Lissin N.M. et al.* // Mol. Microbiol. 1993. V. 7. № 1. P. 49.
<https://doi.org/10.1111/j.1365-2958.1993.tb01096.x>
21. *Tang Y.-C., Chang H.-C., Roeben A. et al.* // Cell. 2006. V. 125. № 5. P. 903.
<https://doi.org/10.1016/j.cell.2006.04.027>
22. *Machida K., Kono-Okada A., Hongo K. et al.* // J. Biol. Chem. 2008. V. 283. № 11. P. 6886.
<https://doi.org/10.1074/jbc.m708002200>
23. *Weaver J., Rye H.S.* // J. Biol. Chem. 2014. V. 289. № 33. P. 23219.
<https://doi.org/10.1074/jbc.m114.577205>
24. *Ishino S., Kawata Y., Taguchi H. et al.* // J. Biol. Chem. 2015. V. 290. № 24. P. 15042.
<https://doi.org/10.1074/jbc.m114.633636>
25. *Chen D.-H., Madan D., Weaver J. et al.* // Cell. 2013. V. 153. № 6. P. 1354.
<https://doi.org/10.1016/j.cell.2013.04.052>
26. *Dalton K.M., Frydman J., Pande V.S.* // PLoS ONE. 2015. V. 10. P. e0117724. № 3.
<https://doi.org/10.1371/journal.pone.0117724>
27. *Tegunov D., Cramer P.* // Nat. Methods. 2019. V. 16. № 11. P. 1146.
<https://doi.org/10.1038/s41592-019-0580-y>
28. *Punjani A., Rubinstein J.L., Fleet D.J., Brubaker M.A.* // Nat. Methods. 2017. V. 14. № 3. P. 290.
<https://doi.org/10.1038/nmeth.4169>
29. *Emsley P., Lohkamp B., Scott W.G., Cowtan K.* // Acta Cryst. D. 2010. V. 66. № 4. P. 486.
<https://doi.org/10.1107/s0907444910007493>
30. *Abraham M.J., Murtola T., Schulz R. et al.* // SoftwareX. 2015. V. 1. P. 19.
<https://doi.org/10.1016/j.softx.2015.06.001>
31. *Robustelli P., Piana S., Shaw D.E.* // Proc. Natl Acad. Sci. 2018. V. 115 (21).
<https://doi.org/10.1073/pnas.1800690115>
32. *Bussi G., Donadio D., Parrinello M.* // J. Chem. Phys. 2007. V. 126. № 1. P. 014101.
<https://doi.org/10.1063/1.2408420>
33. *Parrinello M., Rahman A.* // J. Appl. Phys. 1981. V. 52. №12. P. 7182.
<https://doi.org/10.1063/1.328693>
34. *Tsai J., Taylor R., Chothia C., Gerstein M.* // J. Mol. Biol. 1999. V. 290. № 1. P. 253.
<https://doi.org/10.1006/jmbi.1999.2829>
35. *Martínez L.* // PLoS ONE. 2015. V. 10. № 3. P. e0119264.
<https://doi.org/10.1371/journal.pone.0119264>
36. Python Release Python 3.6.9. *Python.org*. Retrieved from <https://www.python.org/downloads/release/python-369/>
37. *Roh S.-H., Hryc C.F., Jeong H.-H. et al.* // Proc. Natl Acad. Sci. 2017. V. 114. № 31. P. 8259.
<https://doi.org/10.1073/pnas.1704725114>

СТРУКТУРА МАКРОМОЛЕКУЛЯРНЫХ СОЕДИНЕНИЙ

УДК 577.22

ОПРЕДЕЛЕНИЕ ОКОЛОАТОМНОЙ СТРУКТУРЫ ОЛИГОМЕРНЫХ БЕЛКОВ *E. coli* БЕЗ ПРЕДВАРИТЕЛЬНОЙ ОЧИСТКИ

© 2021 г. Е. Б. Пичкур^{1,2}, В. И. Микиртумов³, О. В. Тихонова⁴, Н. И. Деркачева⁵,
Л. П. Курочкина⁶, О. С. Соколова^{3,*}

¹ Национальный исследовательский центр “Курчатовский институт”, Москва, Россия

² Петербургский институт ядерной физики им. Б.П. Константинова НИЦ “Курчатовский институт”,
Гатчина, Россия

³ Московский государственный университет им. М.В. Ломоносова, Москва, Россия

⁴ Научно-исследовательский институт биомедицинской химии им. В.Н. Ореховича, Москва, Россия

⁵ Московский государственный медико-стоматологический университет им. А.И. Евдокимова, Москва, Россия

⁶ Московский государственный университет им. М.В. Ломоносова,
НИИ физико-химической биологии им. А.Н. Белозерского, Москва, Россия

*E-mail: sokolova@mail.bio.msu.ru

Поступила в редакцию 04.05.2020 г.

После доработки 04.05.2020 г.

Принята к публикации 29.05.2020 г.

С помощью криоэлектронной просвечивающей микроскопии изучен грубо очищенный экстракт белков *E. coli*, на основе 2D-классификации проекций выделены классовые суммы, содержащие проекции двух белков: β -галактозидазы и каталитического дигидролипоилтрансукцинилазного домена 2-оксоглутаратдегидрогеназного комплекса, идентифицированных в экстракте методом тандемной масс-спектрометрии. Решены их структуры с разрешением, близким к атомному. В результате моделирования структуры каталитического домена 2-оксоглутаратдегидрогеназного комплекса *de novo* получена атомная модель, выявляющая различия в положении некоторых аминокислотных остатков активного центра по сравнению с ранее полученными кристаллическими структурами.

DOI: 10.31857/S0023476121050180

ВВЕДЕНИЕ

В последние десятилетия структурные методы биологии претерпели бурное развитие, что выразилось в определении новых атомных структур различных белков. Криоэлектронная просвечивающая микроскопия (крио-ПЭМ) не так давно была признана в качестве одного из основных инструментов при изучении белковых структур с высоким разрешением наряду с рентгеноструктурным анализом и ядерно-магнитной спектроскопией (ЯМР) [1]. Использование этого метода резко возросло в связи с определенными технологическими достижениями. К ним относятся не только новые приборы, прямые детекторы электронов, фазовые пластинки, но и такие программные средства, как Relion [2], cryoSPARC [3], Chimera [4, 5] и Coot [6]. В результате удалось добиться впечатляющих результатов при изучении белков и их комплексов. Ярким примером преимущества крио-ПЭМ может служить полученная за месяц структура S-белка коронавируса 2019-nCoV [7]. Эти данные позволили выявить

принципиальные различия в структуре близкородственных вирусов, вызывающих инфекции SARS и COVID-19.

В настоящее время узким местом при получении данных с высоким разрешением остается биохимическая очистка белков. Этот этап трудно автоматизировать, поскольку он в значительной степени опирается на человеческий опыт. Сохраняется ряд проблем, связанных с трудностями выбора системы экспрессии белка, метода его очистки и концентрации функционального белка. Цель состоит в том, чтобы свести к минимуму количество шагов очистки, минимизировать их влияние на белок и получить достаточное количество гомогенного препарата для дальнейшего структурного анализа.

В представленной работе к решению этой проблемы подошли с технологической стороны. При осаждении рекомбинантного белка из экстракта продуцировавших его клеток *E. coli* сульфатом аммония получен препарат, содержащий наряду с целевыми некоторыми клеточными белками. Препарат

был изучен в крио-ПЭМ и решены структуры двух идентифицированных масс-спектрометрией белков *E. coli*: β -галактозидазы и каталитического дигидролипоилтрансукцинилазного домена 2-оксоглутаратдегидрогеназного комплекса (КД-ОДК) с разрешением, близким к атомному. В результате моделирования структуры КД-ОДК *de novo* построена атомная модель, выявляющая различия в положении некоторых аминокислотных остатков (а.о.) активного центра по сравнению с известными кристаллическими структурами [8, 9].

МЕТОДИКА ИССЛЕДОВАНИЙ

Подготовка белков *E. coli*. Выделение и осаждение белков *E. coli* проводили в соответствии с процедурой, описанной в [10]. Клетки BL21(DE3) *E. coli*, продуцирующие рекомбинантный шаперонин бактериофага ОВР, наращивали в большом объеме среды 2xTY при температуре 37°C до оптической плотности $A_{600} = 0.7$. Экспрессию белка индуцировали добавлением изопропил- β -D-тиогалактопиранозида до конечной концентрации 1 мМ и растили клеточную культуру в течение 3.5 ч при 25°C. Клетки осаждали центрифугированием, ресуспендировали в 50 мМ трис-HCl-буфере (pH 7.5), подвергали ультразвуковой обработке с помощью дезинтегратора Virsonic 100 (Vertis, США) и центрифугировали (13000 g) для удаления клеточного дебриса. Нуклеиновые кислоты осаждали путем добавления в супернатант 1/10 объема 30% (w/v) раствора сульфата стрептомицина с последующим центрифугированием. Белки осаждали из супернатанта путем добавления насыщенного раствора сульфата аммония до конечной концентрации 30% (w/v). После центрифугирования осадок белка растворяли в 50 мМ Трис-HCl-буфере (pH 7.5), содержащем 100 мМ NaCl, и промывали 30 объемами буфера (50 мМ Трис-HCl, pH 7.5 с добавлением 10 мМ $MgCl_2$ и 100 мМ KCl) на фильтрах Amicon (Millipore, США). Состав белков анализировали с помощью тандемной масс-спектрометрии.

Масс-спектрометрический анализ. Гидролиз белка в растворе проводили с использованием трипсина (Sequencing Grade Modified, Promega, Madison, WI, USA), как описано в [12]. Анализ полученных при гидролизе пептидов осуществляли с использованием системы высокоэффективной жидкостной хроматографии Ultimate 3000 RSLCnano (Thermo Scientific, США), соединенной с масс-спектрометром Q ExactiveTM HF (Thermo Scientific, США). Пептиды перед аналитическим разделением наносили на обогащающую колонку Acclaim μ -Precolumn (0.5 \times 3 мм, 5 мкм размер частиц; Thermo Scientific) при скорости потока 10 мкл/мин в течение 5 мин в изократическом режиме подвижной фазы Б (2% ацетонитрила, 0.1% муравьиной кислоты). Затем пептиды

разделяли на колонке Acclaim Pepmap C18 (75 мкм \times 150 мм, 2 мкм размер частиц; Thermo Scientific) в градиентном режиме элюирования. Градиент формировали подвижными фазами А (0.1% муравьиной кислоты) и Б (80% ацетонитрила, 0.1% муравьиной кислоты) при скорости потока 0.3 мкл/мин. Колонку промывали 2%-ной подвижной фазой Б в течение 5 мин, после чего концентрацию подвижной фазы Б линейно увеличивали до 35% в течение 40 мин, затем за 5 мин линейно увеличивали концентрацию фазы Б до 99%. После промывки колонки в течение 5 мин при 99% фазы Б концентрацию Б снижали до начальных условий – 2% фазы Б за 5 мин. Общая длительность анализа составила 60 мин.

Масс-спектрометрический анализ проводили в трех технических повторах на гибридном орбитальном масс-спектрометре Q ExactiveTM HF (Thermo Scientific, США) в режиме положительной ионизации с использованием источника NESI (Thermo Scientific, США). Для масс-спектрометрического анализа были установлены следующие параметры настроек: напряжение на эмиттере 2.1 кВ, температура капилляра 240°C. Панорамное сканирование проводили в диапазоне масс от 300 до 1500 m/z, тандемное сканирование фрагментных ионов от нижней границы 100 m/z до верхней границы, определяемой зарядным состоянием прекурсорного иона, но не более 2000 m/z. Для тандемного сканирования учитывали только ионы от $z = 2+$ до $z = 6+$ по зарядному состоянию. Максимальное число разрешенных для синхронной изоляции ионов в режиме MS2 было установлено не более 20-ти. Максимальное время накопления для прекурсорных ионов составило не более 50 мс, для фрагментных ионов – не более 110 мс.

Идентификацию белков по масс-спектрам проводили в поисковой системе MASCOT [13] с использованием базы данных аминокислотных последовательностей SwissProt с ограничением видовой принадлежности организма исследуемого образца *E. coli*. Для оценки FDR была сгенерирована база данных ложных последовательностей белков в виде обратных последовательностей аминокислот протеома *E. coli*. Параметры поиска установлены следующим образом: расщепляющий фермент – трипсин с возможностью пропуска одного сайта расщепления последовательности, точность определения масс-моноизотопных пиков пептидов ± 5 ppm, точность определения масс в спектрах MS/MS ± 0.01 Да. В качестве обязательной и возможной модификации было учтено карбамидометилирование цистеина и окисление метионина соответственно. Для валидации сопоставлений спектров и пептидов PSM, идентификации пептидов и идентификации белков использовали величину FDR не более 1.0%. Белки рассматривали в качестве достоверно иденти-

фицированных, если для них было обнаружено по крайней мере два уникальных пептида, прошедших валидацию. Безметковая количественная оценка содержания белков проходила на основе эмпирического показателя $emPAI$ [14].

Криоэлектронная просвечивающая микроскопия. Препарат белков *E. coli* (3.0 мкл) наносили на сетку с углеродной подложкой, имеющей регулярные отверстия диаметром 1.2 мкм и периодом 1.3 мкм (R1.2/1.3, Quantifoil). Сетки предварительно в течение 30 с обрабатывали тлеющим разрядом в установке для гидрофилизации PELCO easyGlow (Ted Pella, США) при давлении в камере 0.26 мБар и силе тока 25 мА. Сетки с нанесенным белком моментально замораживали в жидком этане в приборе Vitrobot Mark IV (Thermo Fisher Scientific) при 100%-ной влажности и температуре +4.5°C. Стеки микрографий по 38 кадров (4017 стеков) снимали в крио-микроскопе Titan Krios (Thermo Fisher Scientific) с прямым детектором электронов Falcon II в следующих условиях: ускоряющее напряжение – 300 кВ, увеличение – 75000×, размер пикселя – 0.86 Å, поток электронов ~100 э/Å². Стеки были предварительно обработаны в программе Warp [11] для оценки параметров функции передачи контраста (CTF) и коррекции дрейфа на отдельных кадрах стека. Изображения белковых молекул были выбраны с помощью нейронной сети VoxNet и вырезаны из микрографий с помощью рамки 256 × 256 пикселей с получением набора из 470 117 частиц. Из них более 250 000 частиц представляли ОВР шаперонин [10]. 2D-классификацию проводили в программе CryoSPARC [3]. 2D-классы, соответствующие различным белкам, выбирали вручную с последующим генерированием модели *ab-initio* для каждого класса в отдельности. Затем выполнили 3D-уточнение структуры каждого класса с использованием предварительной модели без наложения симметрии. На следующем этапе полученные структуры были использованы в качестве ссылок для дальнейшей 3D-классификации исходного набора частиц. Финальные 3D-структуры белка реконструировали с наложенной симметрией: диэдральная D₂ для β-галактозидазы, октаэдрическая (O) для КД-ОДК. Окончательные реконструкции включали 99 246 частиц для β-галактозидазы и 13 770 частиц для КД-ОДК. Разрешение реконструкций оценивали в соответствии с критерием FSC = 0.143, оно составило 2.6 Å для β-галактозидазы и 3.2 Å для КД-ОДК.

De novo моделирование структуры белка. Моделирование проводили с помощью программ Coot [6] и Phenix [15]. Разрешение карты КД-ОДК, полученное в крио-ПЭМ, было достаточным для идентификации боковых цепей и однозначного определения последовательности данного белка. Уникальная часть карты КД-ОДК извлечена с помощью алгоритма *phenix.mapbox* и NCS-операто-

ров, полученных с помощью алгоритма *phenix.mapsymmetry* после задания симметрии группы точек “O”, атомная модель этой части построена автоматически с помощью алгоритма *phenix.maptomodel*. Полученную модель проверяли вручную в Coot, большие диапазоны остатков несоответствия и пустые зоны заполняли аланиновыми остатками; при необходимости формировали новые пептидные связи таким образом, чтобы модель состояла из одной цепи. После этого боковые цепи аланинов в модели были заменены на боковые цепи аминокислот, соответствующих последовательности C-концевой части белка КД-ОДК (174–405 а.о.). С помощью программы Coot провели ручную правку всех остатков для максимально точного соответствия карте электростатического потенциала при сохранении правильной геометрии остатков; для всех боковых цепей были выбраны ротамеры с их оптимальной ориентацией. Для отредактированной модели уникальной части структуры были применены операторы NCS, полученные с помощью алгоритма *phenix.mapsymmetry*.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Цель работы заключалась в том, чтобы показать, что крио-ПЭМ и обработка изображений позволяют получить 3D-структуры белковых молекул без тонкой очистки целевых белков. Для этого получили коктейль белков из клеток *E. coli*, гетерологично экспрессирующих вирусный шаперонин (ОВР) [10]. Рекомбинантный шаперонин осаждали из клеточного экстракта 30%-ным сульфатом аммония [16], при этом некоторое количество клеточных белков также оказалось в осадке. Идентификацию смеси осажденных белков *E. coli* проводили с использованием высокоэффективной жидкостной хроматографии с тандемной масс-спектрометрией. Результаты представлены в табл. 1. Относительные количества совместно осажденных белков оценивали с помощью показателя $emPAI$, рассчитанного в программе MASCOT. Наибольшее соответствие последовательности (39%) получено для β-галактозидазы (P00722), остальные семь белков представлены несколькими пептидами каждый.

Смесь белков после осаждения и отмычки моментально заморозили в Vitrobot MARK IV (Thermo Fisher Scientific) и исследовали в крио-ПЭМ с использованием стандартного протокола (описан выше). Для идентификации различных белков провели два последовательных раунда 2D-классификации в CryoSPARC [3]. Первую классификацию использовали для исключения из набора “мусорных” классов, после второй (рис. 1) выделили 40 классов, разделенных визуально на три набора данных, каждый из которых представлял отдельный белок.

Таблица 1. Список достоверно идентифицированных белков *E. coli*

Uniprot ID	Белок	Score	MW, kDa	Sequence Coverage, %	emPAI
P00722	Beta-galactosidase (BGAL_ECOLI)	4240	117.35	39	3.06
P0A9L8	Pyrroline-5-carboxylate reductase (P5CR_ECOLI)	395	28.36	21	1.18
P0A9P0	Dihydrolipoyl dehydrogenase (DLDH_ECOLI)	484	50.94	18	0.99
P05055	Polyribonucleotide nucleotidyltransferase (PNP_ECOLI)	394	77.11	14	0.65
P0A6F5	60 kDa chaperonin (CH60_ECOLI)	396	57.46	15	0.65
P0A9H3	Inducible lysine decarboxylase (LDCI_ECOLI)	275	81.61	7	0.54
P0AFG7	Succinyltransferase component of 2-oxoglutarate dehydrogenase complex (ODO2_ECOLI)	311	43.98	6	0.34
P0AES4	DNA gyrase subunit A (GYRA_ECOLI)	41	97.13	3	0.07

Три олигомерных белка: транзистентно экспрессируемый гептамерный шаперонин фага ОВР *P. fluorescens* [10] и два белка *E. coli* – тетрамерная β -галактозидаза и 24-мерный КД-ОДК, имеют частицы сопоставимого размера, 12–15 нм в поперечнике (рис. 1). Кроме этого, наблюдались классы, соответствующие более мелкому асимметричному белку неопределенной природы. Другие белки, пептиды которых были обнаружены при помощи масс-спектрометрии (табл. 1), не смогли идентифицировать на крио-ПЭМ-изображениях. Возможно, эти белки не образуют в используемых условиях стабильные олигомерные комплексы или их концентрации слишком малы для исследования методом крио-ПЭМ.

Наиболее распространенным белком *E. coli* в исследуемой пробе по данным тандемной масс-спектрометрии (табл. 1) и 2D-классификации (рис. 1) была β -галактозидаза (99246 частиц). Она кодируется геном *lacZ* и образует стабильный гомотетрамер с молекулярной массой ~464 кДа [17]. На крио-ПЭМ-изображениях он имеет хорошо узнаваемую форму и может быть легко отделен при помощи 2D-классификации (рис. 1). Стандартный подход к анализу изображений [18] позволил получить 3D-реконструкцию β -галактозидазы с симметрией D2 и разрешением 2.6 Å (рис. 2а–2в).

В состав 2-оксоглутаратдегидрогеназного комплекса *E. coli* входят три различных фермента: Е1 (2-оксоглутаратдекарбоксилаза), Е2 (дигидроли-

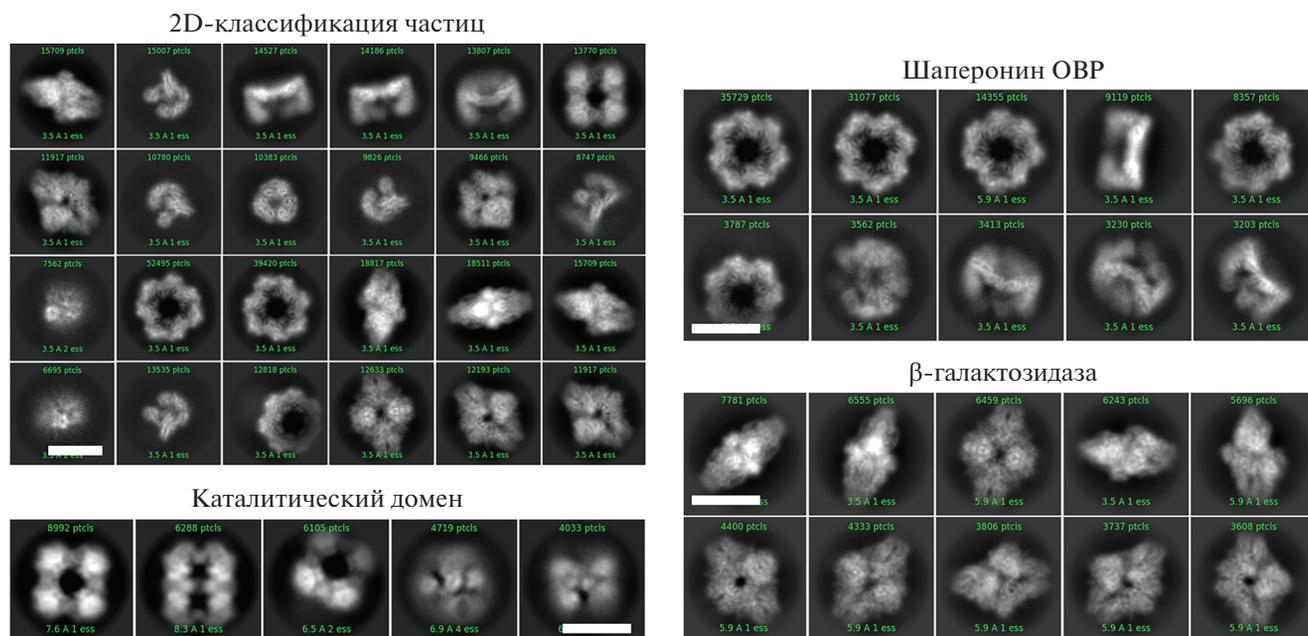


Рис. 1. Крио-ПЭМ-анализ экстракта белков *E. coli*: 2D-классификация и выделенные классы отдельных белков. Масштабные отрезки – 10 нм.

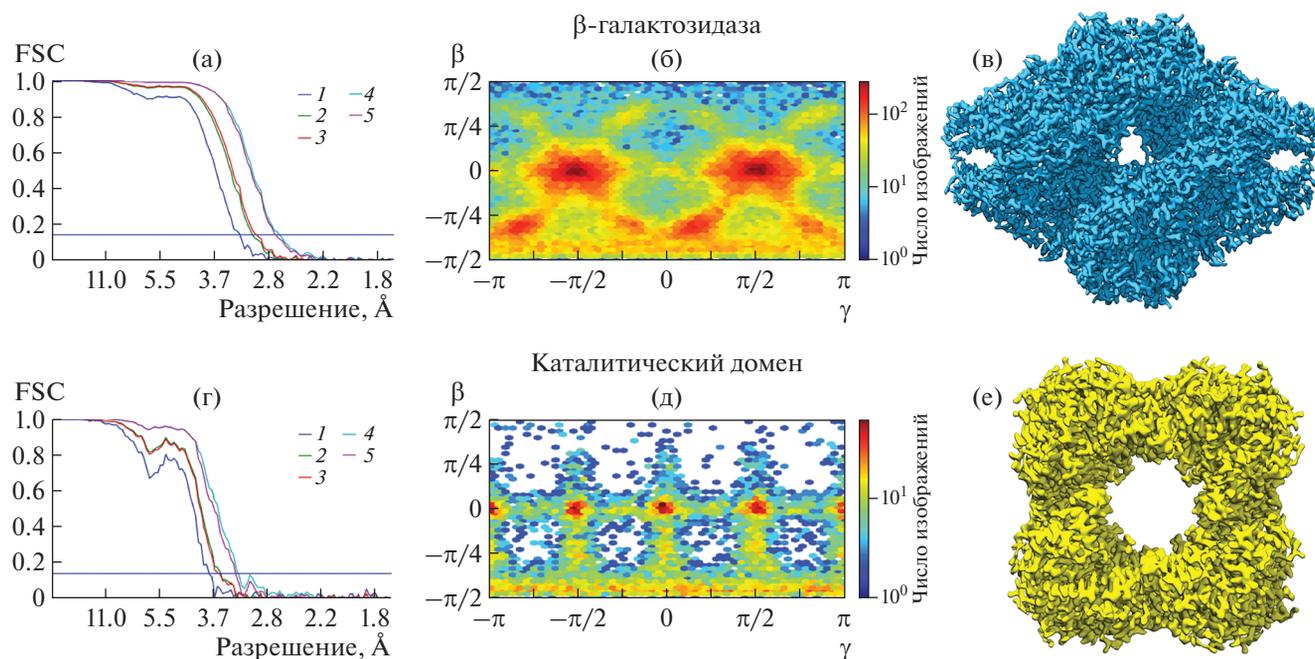


Рис. 2. Построение трехмерных структур белков *E. coli*: β -галактозидазы (а–в) и КД-ОДК (г–е). Кривые функций корреляции (FSC) для соответствующих структур (а, г): 1 – без маски, 2 – сферическая маска, 3 – свободная маска, 4 – узкая маска, 5 – после коррекции; распределение проекций частиц относительно вертикальной оси z сферы Эйлера (б, д). β – угол между проекцией и осью z ; γ – угол поворота вокруг оси z ; изображения выполнены в сguoSPARC [3]. 3D-реконструкция β -галактозидазы (в); 3D-реконструкция КД-ОДК (е).

поилтрансукцинилаза) и E3 (дигидролипоилдегидрогеназа). Фермент E2 катализирует перенос сукциниловой группы из S-сукцинил-дигидролипоила на кофермент А (КоА) [19]. Его полно-размерный мономер длиной 405 а.о. имеет структуру, в которой можно выделить три домена: N-концевой липоилсвязывающий домен, E3-связывающий домен и C-концевой каталитический домен. Молекулярная масса всего белка ~ 44 кДа.

В полученном белковом коктейле *E. coli* с помощью тандемной масс-спектрометрии идентифицировано только два значимых пептида для фермента E2 (один из N-концевого липоилсвязывающего домена и один из C-концевого каталитического домена). Тем не менее на полученных крио-ПЭМ-изображениях C-концевой домен E2 (КД-ОДК) оказался вторым по численности белком после 2D-классификации (13770 частиц). 2D-классификация показала характерные частицы с октаэдрической симметрией (рис. 1). Линейные размеры 2D-классов схожи с размерами ранее определенных кристаллических структур КД-ОДК [8, 9]. При сопоставлении полученных данных с результатами [8, 20] оказалось, что в полученной криоэлектронной структуре отсутствует плотность для N-концевых E3-связывающих и липоилсвязывающих доменов, а также соединенных с ними линкеров. Ранее предполагалось, что эти домены могут быть удалены эндогенными

протеазами [9]. Структуры изолированных N-концевых доменов были решены с помощью ЯМР [21, 22].

Отметим, что кристаллы этого домена были ранее случайно получены при попытке кристаллизации рекомбинантной амидазы *Arabidopsis thaliana*, экспрессированной в гетерологичной системе *E. coli* [9]. Предполагалось, что остатки гистидина на поверхности каталитического домена могут способствовать связыванию его с никельсодержащей смолой [23], однако в данном исследовании мы не использовали Ni-аффинную хроматографию.

Реконструкция в сguoSPARC с октаэдрической симметрией КД-ОДК имеет разрешение 3.2 Å (рис. 2г–2е). Высокое разрешение позволило определить структуру каталитического домена *de novo*. Модель структуры создавали с помощью программы Coot [6] и трехмерных координат для каждого атома (рис. 3а).

Сравнение крио-ПЭМ-структуры с рентгеновской структурой (PDB ID 6rbg) показало, что существуют некоторые различия в ориентациях боковых цепей ключевых аминокислот в активном центре: D382 и D185 (рис. 3а, вставки). Основной, пока не решенный, вопрос относительно структуры КД-ОДК – это причины изменений ориентаций боковых цепей, происходящие в активном участке при связывании субстрата. Ранее

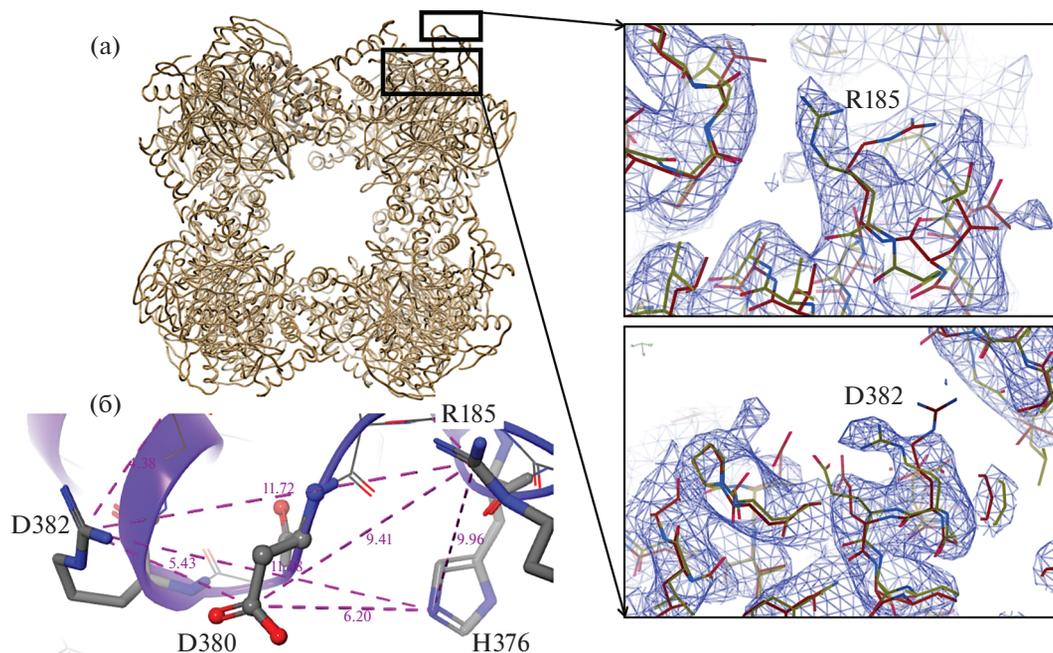


Рис. 3. Структура каталитического домена 2-оксоглутарат-дегидрогеназного комплекса. Атомная модель КД-ОДК, выполненная *de novo*; рамкой отмечено положение активного центра, вставки – различия в позициях аргининов R185 и R382 активного центра. Показаны кристаллическая структура из [9] и структура из данного исследования. Карта электростатического потенциала показана на уровне 2σ (а). Расстояния между ключевыми аминокислотами активного центра КД-ОДК (б).

было высказано предположение, что остаток H375 инициирует первую ступень катализа путем депротонирования тиоловой группы КоА, которая затем атакует карбонильный атом сукцинильного остатка, связанного с дигидролипамидом. Разрушение промежуточного соединения приводит к образованию сукцинил-КоА и протонированной дигидролипоиловой группы [8]. Также предполагалось, что при связывании КоА в активном центре аспарагин D380 образует солевые мостики с гистидином H376 и с одним из двух остатков аргинина: R185 или R382 [20]. Для этого D380 должен менять свою ориентацию в зависимости от субстрата. Полученная с помощью крио-ПЭМ структура позволила точно локализовать боковые цепи R185 и R382; эти боковые цепи были не упорядочены в раннем рентгеновском исследовании (PDB ID 1e2o) и выявлены позже (PDB ID 6prb). Сравнение модели PDB:6prb с полученной в данной работе (рис. 3а, вставки) показало, что в модели *de novo* остатки Arg повернуты. Измеренные расстояния между ключевыми а.о. в активном центре составили: D380-H376 = 6.2, D380-R382 = 5.4, D380-R185 = 9.4 Å (рис. 3б). Таким образом, аспарагин D380 располагается примерно посередине между гистидином H376 и аргинином R380 и его взаимодействия с этими остатками, по-видимому, носят ионный характер. С аргинином R185 аспарагин D380 согласно рассмотренной модели не взаимодействует. Для

более точного определения структуры активного центра фермента необходимо получить реконструкцию КД-ОДК с субстратом.

ЗАКЛЮЧЕНИЕ

Крио-ПЭМ становится незаменимым инструментом в структурных исследованиях биологических макромолекул. До сих пор для получения структур с высоким разрешением требовались высоко очищенные препараты белков. Однако 3D-классификация гетерогенных частиц, реализованная в программах Relion [2] и cryoSPARC [3] в сочетании с масс-спектрометрией, как показывает практика, позволяет обойтись без специальной очистки [24]. В настоящей работе использовали крио-ПЭМ и tandemную масс-спектрометрию для идентификации белковых молекул в грубо очищенной белковой пробе. С помощью анализа изображений удалось выделить две выборки частиц, соответствующие двум белкам *E. coli* и получить две 3D-структуры: β -галактозидазы и КД-ОДК. Оба белка находились в нативном состоянии и не имели рекомбинантных тегов, используемых для очистки. Разрешение, близкое к атомному, позволило выполнить *de novo*-моделирование атомной модели КД-ОДК. Сравнение полученной структуры с опубликованной кристаллической структурой этого же домена [9] показало, что активный сайт в растворе

находится в несколько иной конформации, чем в кристалле.

Авторы выражают благодарность В.Н. Новоселскому за помощь в подготовке рис. 3б.

Эксперименты проведены на оборудовании Уникальной научной установки “3D-ЕМС” МГУ (при поддержке Министерства науки и высшего образования РФ, уникальный идентификатор RFMEFI61919X0014). Получение крио-ПЭМ-данных с высоким разрешением проведено с использованием оборудования Ресурсного центра зондовой и электронной микроскопии “Нанозонд” НИЦ “Курчатовский институт”. Расчеты проведены с использованием вычислительных ресурсов Федерального центра коллективного пользования “Комплекс моделирования и обработки данных исследовательских установок мега-класса” НИЦ “Курчатовский институт” (<http://ckp.nrcki.ru/>). Масс-спектрометрические измерения проведены на оборудовании центра коллективного пользования “Протеом человека” ИБМХ.

Исследование выполнено при поддержке Междисциплинарной научно-образовательной школы Московского государственного университета “Молекулярные технологии живых систем и синтетическая биология” в рамках научного проекта государственного задания МГУ № 121032500056-3.

СПИСОК ЛИТЕРАТУРЫ

1. Соколова О.С., Курпичников М.П., Шайтан К.В. и др. Современные методы изучения структуры и функций ионных каналов. М.: Товарищество научных изданий КМК, 2020. 316 с.
2. Scheres S.H. // *Methods Enzymol.* 2016. V. 579. P. 125. <https://doi.org/10.1016/bs.mie.2016.04.012>
3. Punjani A., Rubinstein J.L., Fleet D.J. et al. // *Nat. Methods.* 2017. V. 14. P. 290. <https://doi.org/10.1038/nmeth.4169>
4. Yang Z., Lasker K., Schneidman-Duhovny D. et al. // *J. Struct. Biol.* 2012. V. 179. P. 269. <https://doi.org/10.1016/j.jsb.2011.09.006>
5. Pettersen E.F., Goddard T.D., Huang C.C. et al. // *J. Comput. Chem.* 2004. V. 25. P. 1605. <https://doi.org/10.1002/jcc.20084>
6. Emsley P., Lohkamp B., Scott W.G. et al. // *Acta Cryst. D.* 2010. V. 66. P. 486. <https://doi.org/10.1107/S0907444904019158>
7. Wrapp D., Wang N., Corbett K.S. et al. // *Science.* 2020. V. 367. P. 1260. <https://doi.org/10.1126/science.abb2507>
8. Knapp J.E., Carroll D., Lawson J.E. et al. // *Protein Sci.* 2000. V. 9. P. 37. <https://doi.org/10.1110/ps.9.1.37>
9. Andi B., Soares A. S., Shi W. et al. // *Acta Cryst. F.* 2019. V. 75. P. 616. <https://doi.org/10.1107/S2053230X19011488>
10. Stanishneva-Konovalova T.B., Semenyuk P.I., Kurochkina L.P. et al. // *J. Struct. Biol.* 2020. V. 209. P. 107439. <https://doi.org/10.1016/j.jsb.2019.107439>
11. Tegunov D., Cramer P. // *Nat. Methods.* 2019. V. 16. P. 1146. <https://doi.org/10.1038/s41592-019-0580-y>
12. Anselm V., Novikova S., Zgoda V. // *Int. J. Mol. Sci.* 2017. V. 18. <https://doi.org/10.3390/ijms18081763>
13. Perkins D.N., Pappin D.J., Creasy D.M. et al. // *Electrophoresis.* 1999. V. 20. P. 3551. [https://doi.org/10.1002/\(SICI\)1522-2683\(19991201\)20:18<3551::AID-ELPS3551>3.0.CO;2-2](https://doi.org/10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2)
14. Ishihama Y., Oda, Y., Tabata T. et al. // *Mol. Cell. Proteomics.* 2005. V. 4. P. 1265. <https://doi.org/10.1074/mcp.M500061-MCP200>
15. Afonine P.V., Poon B.K., Read R.J. et al. // *Acta Cryst. D.* 2018. V. 74. P. 531. <https://doi.org/10.1107/S2059798318006551>
16. Wingfield P. // *Curr. Protoc. Protein Sci.* 2001. Appendix 3. <https://doi.org/10.1002/0471140864.psa03fs13>
17. Jacobson R.H., Zhang X.J., DuBose R.F. et al. // *Nature.* 1994. V. 369. P. 761. <https://doi.org/10.1038/369761a0>
18. Zivanov J., Nakane T., Forsberg B.O. et al. // *Elife.* 2018. P. e42166. <https://doi.org/10.7554/eLife.42166>
19. Perham R.N. // *Biochemistry.* 1991. V. 30. P. 8501. <https://doi.org/10.1021/bi00099a001>
20. Knapp J.E., Mitchell D.T., Yazdi M.A. et al. // *J. Mol. Biol.* 1998. V. 280. P. 655. <https://doi.org/10.1110/ps.9.1.37>
21. Riccaud P.M., Howard M.J., Roberts E.L. et al. // *J. Mol. Biol.* 1996. V. 264. P. 179. <https://doi.org/10.1006/jmbi.1996.0632>
22. Robien M.A., Clore G.M., Omichinski J.G. et al. // *Biochemistry.* 1992. V. 31. P. 3463. <https://doi.org/10.1021/bi00128a021>
23. Bolanos-Garcia V.M., Davies O.R. // *Biochim. Biophys. Acta.* 2006. V. 1760. P. 1304. <https://doi.org/10.1016/j.bbagen.2006.03.027>
24. Verbeke E.J., Mallam, A.L., Drew K. et al. // *Cell. Rep.* 2018. V. 24. P. 259. <https://doi.org/10.1016/j.celrep.2018.06.022>

RYFERPRESTR: ПРОГРАММНОЕ ДОПОЛНЕНИЕ К СИСТЕМЕ МОЛЕКУЛЯРНОЙ ГРАФИКИ RYMOI ДЛЯ РАСЧЕТОВ СВОБОДНОЙ ЭНЕРГИИ СВЯЗЫВАНИЯ ЛИГАНДА С РЕЦЕПТОРОМ

© 2021 г. А. А. Лашков^{1,*}, И. В. Толмачев², П. А. Эйстрих-Геллер¹, С. В. Рубинский¹

¹ Институт кристаллографии им. А.В. Шубникова ФНИЦ “Кристаллография и фотоника” РАН, Москва, Россия

² Школа № 179, Москва, Россия

*E-mail: alashkov83@gmail.com

Поступила в редакцию 06.05.2020 г.

После доработки 17.05.2020 г.

Принята к публикации 19.05.2020 г.

Расчеты свободной энергии связывания лигандов с рецептором, чаще всего белковой природы, находят широкое применение в рациональном дизайне новых биологически активных соединений. Равновесные методы, основанные на применении молекулярной динамики и “алхимических” преобразований, такие как термодинамическое интегрирование и метод Беннета, являются “золотым стандартом” подобных вычислений. Технически эти методы требуют введения дополнительных ограничений на взаимное расположение атомов в системе лиганд–рецептор. Разработано дополнение к системе молекулярной графики RYMOI, облегчающее введение требуемых ограничений и аналитический расчет поправки при вычислении свободной энергии связывания лиганда с рецептором (белком).

DOI: 10.31857/S002347612105012X

ВВЕДЕНИЕ

Создание новых лекарственных препаратов является одним из самых ресурсозатратных процессов в современном мире. Разработка биологически активных соединений – многостадийный процесс. На первом этапе нужно отобрать несколько веществ – кандидатов – для дальнейших доклинических и клинических испытаний. Если известна молекулярная мишень действия разрабатываемого биологически активного соединения, то проводят либо химический, либо виртуальный скрининг библиотеки соединений на степень аффинности к молекулярной мишени. Во втором случае крайне важными являются количественные оценки степени связывания молекулы биологически активного соединения (лиганда) с молекулярной мишенью – рецептором (чаще всего белковой молекулой).

Виртуальный скрининг, основанный на молекулярном докинге, т.е. подборе структуры молекулы лиганда, аффинной к структуре рецептора, представляет собой многостадийный процесс [1]. В качестве структуры рецептора принято использовать экспериментально определенную либо методом ядерного магнитного резонанса, либо методом рентгеноструктурного анализа с атомным разрешением пространственную структуру макромолекулы. Количество лигандов на каждой

стадии виртуального скрининга существенно уменьшается по сравнению с предыдущей. Для оценки степени аффинности на первом этапе приемлемо применять высокопроизводительные оценочные функции [2]. Однако у каждого класса оценочных функций есть свои недостатки. Например, оценочные функции, основанные на силовых полях, требуют значительных вычислительных ресурсов, а эмпирические функции не отличаются высокой точностью [3]. Оценочные функции применимы в основном лишь для качественной оценки аффинности связывания и ранжирования результатов докинга [2].

Для количественной оценки аффинности связывания лиганда с рецептором необходимо применение более точных, но и более ресурсоемких методов, основанных на молекулярной динамике (МД): либо моделирование доставки лиганда к месту связывания, либо изменение характера связей лиганда с рецептором на основе так называемых “алхимических” преобразований [4]. На сегодняшний день среди методов, основанных на равновесных “алхимических” преобразованиях, широкое применение получили: термодинамическое интегрирование [4, 5], метод отношения вероятности принятия шага Беннета [6], метод множественного отношения вероятности принятия шага Беннета [7]. Технически эти методы требуют

введения дополнительных ориентационных ограничений на взаимное расположение атомов в системе лиганд–рецептор для предотвращения при отключении взаимодействий лиганда с рецептором и изменении параметра связи Кирквуда [8] выталкивания лиганда из его сайта связывания. Лиганд оказывается в окружении, сильно отличающемся от окружения лиганда в сайте связывания. Это противоречит сути равновесных методов расчета свободной энергии, основанных на “алхимических” преобразованиях.

Для удобного введения ориентационных ограничений на положение лиганда и расчета поправки к свободной энергии связывания при введенных ограничениях авторами разработано программное дополнение к системе молекулярной графики PyMol [9]. Альтернативой является прямая правка файлов настроек МД-симуляции с использованием текстового редактора с поиском индексов атомов в файле структуры, что представляет собой довольно трудоемкий процесс, сопровождающийся повышением вероятности ошибок.

ТЕРМОДИНАМИЧЕСКИЙ ЦИКЛ

Для расчета свободной энергии связывания часто используется термодинамический цикл, описанный в [1]. Для расчета необходимо провести моделирование с последовательным изменением параметра связи [8] атомов лиганда для систем лиганд–рецептор и лиганд–растворитель. Параметр связи меняется последовательно сначала для кулоновских взаимодействий, потом для взаимодействий, описываемых потенциалом Леннарда-Джонса. Ограничения на положение атомов лиганда в сайте связывания рецептора также необходимо вводить постепенно до начала изменения параметров связи. Это нужно для корректного расчета изменения энергии связывания, обусловленного введением ограничений в системе лиганд–рецептор в процессе моделирования.

В случае цикла, в котором в отдельных ветвях реализуются переходы молекулы лиганда из связанного состояния в несвязанное в системах лиганд–растворитель (“solv”) и лиганд–рецептор (“prot”), свободная энергия связывания рецептора с лигандом равна:

$$\Delta G_{\text{bind}}^0 = \Delta G_{\text{coul+vdw}}^{\text{solv}} + \Delta G_{\text{restr_on}}^{\text{solv}} - \Delta G_{\text{restr+coul+vdw}}^{\text{prot}} \quad (1)$$

где $\Delta G_{\text{restr+coul+vdw}}^{\text{prot}}$ — изменение свободной энергии в системе лиганд–рецептор в процессе перехода лиганда из связанного состояния в несвязанное, но с введенными дополнительными ограничениями, $\Delta G_{\text{coul+vdw}}^{\text{solv}}$ — изменение свободной энергии в системе лиганд–растворитель в процессе перехода лиганда из взаимодействующего с растворите-

лем состояния в невзаимодействующее, $\Delta G_{\text{restr_on}}^{\text{solv}}$ — аналитически рассчитываемая поправка на то, что в системе лиганд–растворитель не вводятся соответствующие ограничения на положение атомов лиганда.

В случае заряженного лиганда при использовании методов, основанных на суммировании по Эвальду [10], для расчета кулоновских взаимодействий и периодических граничных условий необходимо в формулу (1) вводить дополнительные слагаемые, описанные в [11]. В случае применения схемы отсечки либо сдвигового потенциала для расчета дисперсионного члена потенциала Леннарда-Джонса в формулу (1) желателен внесение поправки, описанную в [12]. При использовании метода суммирования по Эвальду для расчета дисперсионного члена потенциала Леннарда-Джонса [13] эта поправка не требуется.

ПРИМЕНЕНИЕ ОРИЕНТАЦИОННЫХ ОГРАНИЧЕНИЙ В ПАКЕТЕ GROMACS

Начиная с версии 2016 в пакете программ МД GROMACS [14] имеется возможность введения ограничений на взаимное положение атомов, даже относящихся к разным молекулам. Ограничения могут быть введены на расстояние между двумя атомами, угол между двумя лучами, выходящими из одного атома и проходящими через два других, двугранный угол между двумя плоскостями, которые проходят через четыре атома. Ограничения могут быть описаны разными функциями потенциальной энергии. Чаще всего используют гармоническое соотношение между отклонением параметра от равновесного значения (r_0 , θ_0 , ϕ_0) и потенциальной энергией (U) при введенном ограничении:

$$U(r) = \frac{K_r}{2} (r - r_0)^2, \quad (2)$$

$$U(\theta) = \frac{K_\theta}{2} (\theta - \theta_0)^2, \quad (3)$$

$$U(\phi) = \frac{K_\phi}{2} (\phi - \phi_0)^2. \quad (4)$$

В [15] описана общая схема введения ориентационных ограничений на положение лиганда в системе лиганд–рецептор, включающая одно ограничение на расстояние между атомом рецептора (a) и лиганда (A), два ограничения на углы $b-a-A$ и $a-A-B$ и три ограничения на двугранные углы $c-b-a-A$, $b-a-A-B$, $a-A-B-C$ (рис. 1). Там же даны рекомендации по выбору значений коэффициентов жесткости K . В большинстве случаев они должны лежать в диапазоне 5–50 ккал/(моль Å²) (ккал/(моль рад²) в случае ограничений по углам и двугранным углам). В [16] при вычислении свободной энергии связывания ингибитора и суб-

стра с уридинфосфорилазой из холерного вибриона использовали коэффициенты жесткости в диапазоне 10–20 ккал/(моль Å²) (ккал/(моль рад²)).

Для описания рассматриваемых в настоящей работе ограничений в программном пакете GROMACS в основном файле топологии, начиная с версии 2016, добавлена секция [intermolecular interaction]. Она должна быть последней секцией в этом файле. В этой секции необходимо описать три подсекции [bonds], [angles] и [dihedrals]. Каждое ограничение (из шести) описывается одной строкой в соответствующей подсекции. В строке задаются глобальные индексы атомов, на взаимное положение которых будет накладываться ограничение, тип функции потенциальной энергии (в соответствии с документацией GROMACS), равновесные значения параметра ограничения (в нанометрах для расстояния между атомами и в градусах для углов) и константы жесткости для двух состояний: A ($\lambda = 0$) и B ($\lambda = 1$). В задаче вычисления свободной энергии связывания лиганда с белком K ($\lambda = 0$) полагают равным нулю, а равновесные значения параметров ограничений для состояний A и B равными между собой. При моделировании реализуется переход из состояния A в состояние B через промежуточные состояния – число состояний и значение λ для них контролируют с помощью настроек моделирования.

Как показывает практика авторов настоящей и других работ [1], при использовании в методе МД алгоритма жестких ограничений LINCS [17] для описания валентных связей атомов водорода с другими атомами дополнительные ограничения на атомы водорода накладывать не следует.

ВВЕДЕНИЕ ОРИЕНТАЦИОННЫХ ОГРАНИЧЕНИЙ В ПАКЕТЕ NAMD

В пакете МД NAMD [18] также доступен расчет свободной энергии связывания лиганда с рецептором методами возмущения свободной энергии или термодинамического интегрирования. Возможность введения ориентационных ограничений присутствует за счет механизма коллективных переменных [19]. Однако в одном процессе в отличие от GROMACS невозможно совместить последовательное плавное включение ограничений и последующее отключение взаимодействий лиганда с рецептором. Тем не менее можно раздельно проводить эти преобразования, для каждого из которых нужен отдельный файл конфигурации, написанный на языке TCL (Tool Command Language) и содержащий описание коллективных переменных через индексы взаимодействующих атомов и параметры гармонического потенциала для них. Для составления шаблона файлов настроек ограничений в программном пакете

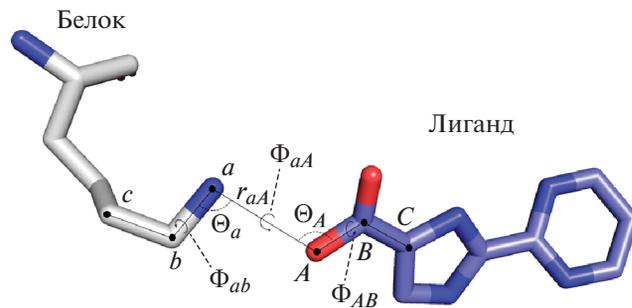


Рис. 1. Схема, описанная в [15], для введения дополнительных ограничений на положение лиганда в сайте связывания рецептора (белка).

NAMD также можно использовать описываемое в настоящей работе программное обеспечение.

ВЫЧИСЛЕНИЕ АНАЛИТИЧЕСКОЙ ПОПРАВКИ НА ВВЕДЕННЫЕ ОГРАНИЧЕНИЯ

Необходимую для вычисления ΔG_{bind} поправку на энергию введенных ограничений для системы лиганд–растворитель с поправкой на стандартное состояние 1 М можно вычислить аналитически по формуле [15]:

$$\Delta G_{\text{rest_off}}^{\text{solv}} = -RT \ln \left[\frac{8\pi^2 V_0}{r_{aA}^2 \sin \theta_a \sin \theta_A} \times \frac{\sqrt{K_{r_{aA}} K_{\theta_a} K_{\theta_A} K_{\phi_{ba}} K_{\phi_{aA}} K_{\phi_{AB}}}}{(2\pi RT)^3} \right], \quad (5)$$

где r_{aA} – длина псевдосвязи [нм], θ_a, θ_A – угловые ограничения, $K_{r_{aA}}, K_{\theta_a}, K_{\theta_A}, K_{\phi_{ba}}, K_{\phi_{aA}}, K_{\phi_{AB}}$ – константы жесткости для ограничений на расстояние, углы и двугранные углы, V_0 – объем стандартного состояния (1.66 нм³), R – универсальная газовая постоянная, T – температура [К],

$$\Delta G_{\text{rest_on}}^{\text{solv}} = -\Delta G_{\text{rest_off}}^{\text{solv}}.$$

Аналитическая поправка, вычисленная по формуле (5), не учитывает возможное влияние ограничений на симметричные группы молекулы в том случае, когда параметры ограничивают переход из одной симметричной формы в другую [20] или имеется высокий потенциальный барьер между симметричными формами, не связанный с применением ограничений. Эта дополнительная поправка должна быть вычислена отдельно и добавлена в формулу (1). Так, если химическая группа преобразуется осью симметрии N -го порядка, то $\Delta G_{\text{sym}} = -RT \ln(N)$. Например, для фенильной группы ($N = 2$) $\Delta G_{\text{sym}} = -0.413$ ккал/моль.

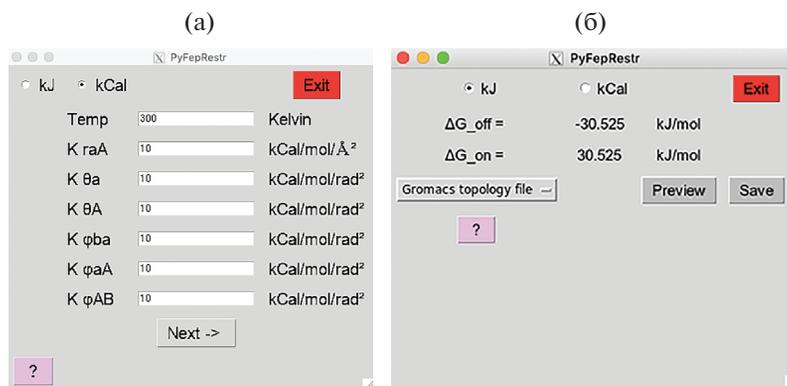


Рис. 2. Форма графического интерфейса пользователя программного дополнения, предназначенная для ввода входных параметров (а), для просмотра результирующего значения $\Delta G_{\text{restr}}^{\text{sol}}v$ и записи в файл топологии ограничений, визуально описываемых на предыдущем этапе работы с программным дополнением (б).

ОПИСАНИЕ ПРОГРАММНОГО ДОПОЛНЕНИЯ К СИСТЕМЕ МОЛЕКУЛЯРНОЙ ГРАФИКИ PYMOL

Программное дополнение к системе молекулярной графики PyMol – PyFepRestr – предназначено для визуализации задания атомов, участвующих в ориентационных ограничениях, задания констант жесткости, автоматической конвертации используемых единиц измерения, аналитического расчета поправки на введенные ориентационные взаимодействия и записи в файл топологии программного пакета МД GROMACS настроек ограничений. Программное дополнение написано на языке программирования Python 2.7, но совместимо и с Python 3+. Для реали-

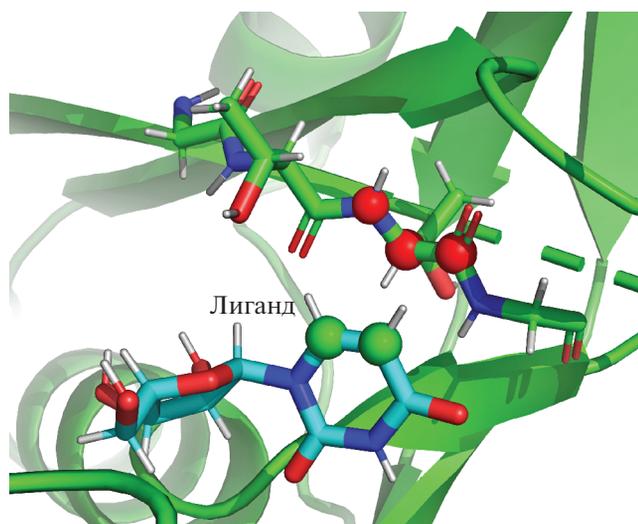


Рис. 3. Окно просмотра структуры системы молекулярной графики PyMol на этапе визуального описания вводимых ограничений. Шарами разного цвета обозначены атомы лиганда и рецептора, участвующие в задании ориентационных ограничений.

зации графического интерфейса использован модуль Tkinter стандартной библиотеки подпрограмм. Для взаимодействия с PyMol в программе применяют интерфейс прикладного программирования, реализованный с помощью модулей `pymol.cmd` и `pymol.wizard`. Программа совместима с версиями PyMol, начиная с 1.5.0.6.

При запуске дополнения отображается первая форма графического интерфейса пользователя (рис. 1а), в которой необходимо указать ограничения констант жесткости и абсолютную температуру МД-моделирования. Значения констант можно задавать как в кДж/моль, так и в ккал/моль. После задания параметров управление передается главному окну PyMol (рис. 3), где в визуальном режиме на структуре комплекса лиганд–рецептор отмечают атомы, по которым строятся ограничения, удерживающие лиганд в сайте связывания в процессе МД-моделирования. После указания шести атомов отображается следующая форма графического интерфейса пользователя (рис. 2б) с рассчитанными значениями поправки $\Delta G_{\text{restr}}^{\text{sol}}v$ на введенные ограничения. Также настройки ограничений для программного пакета GROMACS можно отобразить на экране и записать в основной файл топологии. Возможны отображение и запись шаблона файла настройки ограничений через механизм коллективных переменных для программного пакета NAMD. Перевод единиц измерения констант жесткости в принятые в программном пакете МД осуществляется автоматически.

ЗАКЛЮЧЕНИЕ

Авторами разработано дополнение к системе молекулярной графики PyMol, облегчающее введение ориентационных ограничений на положение лиганда в сайте связывания и аналитический расчет поправки при вычислении свободной энергии

связывания лиганда с рецептором – PyFepRestr. Описаны программа и используемые в ней алгоритмы, дано теоретическое обоснование. Программа опубликована на сервисе GitHub под лицензией MIT (https://github.com/tolmv/plugin_for_Pymol).

Работа выполнена с использованием оборудования ЦКП ФНИЦ “Кристаллография и фотоника” РАН при поддержке Министерства науки и высшего образования РФ (проект RFMEFI62119X0035) в рамках Государственного задания ФНИЦ “Кристаллография и фотоника” РАН.

СПИСОК ЛИТЕРАТУРЫ

1. *Aldeghi M., Heifetz A., Bodkin M.J. et al.* // Chem. Sci. 2016. V. 7. P. 207. <https://doi.org/10.1039/c5sc02678d>
2. *Kitchen D.B., Decornez H., Furr J.R. et al.* // Nat. Rev. Drug Discov. 2004. V. 3. P. 935. <https://doi.org/10.1038/nrd1549>
3. *Li J., Fu A., Zhang L. et al.* // Interdiscipl. Sci. Comput. Life. Sci. 2019. V. 11. P. 320. <https://doi.org/10.1007/s12539-019-00327-w>
4. *Chipot C.* // Wiley Interdiscipl. Rev.: Comput. Mol. Sci. 2014. V. 4 (1). P. 71. <https://doi.org/10.1002/wcms.1157>
5. *Klimovich P.V., Shirts M.R., Mobley D.L.* // J. Comput. Aided Mol. Des. 2015. V. 29 (5). P. 397. <https://doi.org/10.1007/s10822-015-9840-9>
6. *Bennett C.H.* // J. Comput. Phys. 1976. V. 22 (2). P. 245. [https://doi.org/10.1016/0021-9991\(76\)90078-4](https://doi.org/10.1016/0021-9991(76)90078-4)
7. *Shirts M.R., Chodera J.D.* // J. Chem. Phys. 2008. V. 129. P. 124105. <https://doi.org/10.1063/1.2978177>
8. *Kirkwood J.* // J. Chem. Phys. 1935. V. 3. P. 300. <https://doi.org/10.1063/1.1749657>
9. *DeLano W.L., Lam J.W.* // Abstr. Papers Am. Chem. Soc. 2005. V. 230. P. 1371.
10. *Hünenberger P.H., McCammon J.A.* // Biophys. Chem. 1999. V. 78. P. 69. [https://doi.org/10.1016/s0301-4622\(99\)00007-1](https://doi.org/10.1016/s0301-4622(99)00007-1)
11. *Rocklin G.J., Mobley D.L., Dill K.A. et al.* // J. Chem. Phys. 2013. V. 139. P. 184103. <https://doi.org/10.1063/1.4826261>
12. *Shirts M.R., Mobley D.L., Chodera J.D. et al.* // J. Phys. Chem. 2007. V. 111 (45). P. 13052. <https://doi.org/10.1021/jp0735987>
13. *Wennberg C.L., Murtola T., Hess B. et al.* // Chem. Theory Comput. J. 2013. V. 9. P. 3527. <https://doi.org/10.1021/ct400140n>
14. *Van Der Spoel D., Lindahl E., Hess B. et al.* // J. Comput. Chem. 2005. V. 26. P. 1701. <https://doi.org/10.1002/jcc.20291>
15. *Boresch S., Tettinger F., Leitgeb M. et al.* // J. Phys. Chem. B. 2003. V. 107. P. 9535. <https://doi.org/10.1021/jp0217839>
16. *Эйтсрих-Геллер П.А., Рубинский С.В., Прокофьев И.И. и др.* // Кристаллография. 2020. Т. 65. № 2. С. 271. <https://doi.org/10.31857/S002347612002006x>
17. *Hess B., Bekker H., Berendsen H.J.C. et al.* // J. Comput. Chem. 1997. V. 18. P. 1463. [https://doi.org/10.1002/\(SICI\)1096-987X\(199709\)18:123.0.CO;2-H](https://doi.org/10.1002/(SICI)1096-987X(199709)18:123.0.CO;2-H)
18. *Phillips J.C.* // J. Comput. Chem. 2005. V. 26. P. 1781. <https://doi.org/10.1002/jcc.20289>
19. *Fiorin G., Klein M.L.* // Mol. Phys. 2013. V. 111. P. 3345. <https://doi.org/10.1080/00268976.2013.813594>
20. *Mobley D.L., Chodera J.D., Dill K.A.* // J. Chem. Phys. 2006. V. 125 (8). P. 084902. <https://doi.org/10.1063/1.2221683>