

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР  
«КУРЧАТОВСКИЙ ИНСТИТУТ»

На правах рукописи



КОРДОНСКАЯ ЮЛИЯ ВЛАДИМИРОВНА

**Молекулярная динамика олигомеров белков в кристаллизационных  
растворах**

Специальность 1.3.8. Физика конденсированного состояния

Диссертация на соискание ученой степени  
кандидата физико-математических наук

Научный руководитель:  
член-корреспондент РАН,  
доктор физико-математических наук  
Дьякова Юлия Алексеевна

Москва 2025

## Оглавление

<b>ВВЕДЕНИЕ</b> .....	4
<b>ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ</b> .....	14
1.1. Вычислительные подходы к предсказанию белковых структур .....	15
1.2. Белковая кристаллизация.....	18
1.3. Молекулярная динамика как инструмент исследования биомолекулярных систем.....	24
1.3.1. Физико-математический аппарат метода МД.....	29
1.3.2. Силовое поле.....	31
1.3.3. Периодические граничные условия .....	34
1.3.4. Способы ускорения МД-расчетов .....	36
1.4. Применение молекулярной динамики для исследования механизмов кристаллизации белков.....	38
1.4.1. МД-исследования кластеров-прекурсоров кристалла лизоцима: идентификация и роль ионов .....	42
1.4.2. Влияние ионов осадителя на стабильность олигомеров лизоцима при различных температурах .....	45
1.5. Заключение.....	50
<b>ГЛАВА 2. АДАПТАЦИЯ МЕТОДОВ МОЛЕКУЛЯРНОЙ ДИНАМИКИ К ИССЛЕДОВАНИЯМ КРИСТАЛЛИЗАЦИОННЫХ РАСТВОРОВ БЕЛКОВ</b> .....	53
2.1. Проведение вычислительного эксперимента .....	54
2.1.1. Построение моделей .....	54
2.1.2. МД-моделирование в атомистическом силовом поле.....	56
2.1.3. Анализ результатов .....	58
2.2. Определение оптимальных размеров ячейки моделирования .....	59
2.2.1. Влияние размеров ячейки моделирования и концентрации ионов осадителя на стабильность димера лизоцима .....	59
2.2.2. Минимально допустимый размер ячейки моделирования на основе сравнения стабильности олигомеров лизоцима в кристаллизационном растворе.....	63
2.3. Применимость крупнозернистого силового поля MARTINI для моделирования белковых олигомеров в кристаллизационном растворе .....	69

2.4. Заключение.....	77
<b>ГЛАВА 3. СТРУКТУРА И ТЕРМОСТАБИЛЬНОСТЬ ДИМЕРОВ, ОБРАЗУЮЩИХСЯ В КРИСТАЛЛИЗАЦИОННОМ РАСТВОРЕ ПРОТЕИНАЗЫ К</b> .....	<b>79</b>
3.1. Уточнение структуры димера протеиназы К, образующегося в ее кристаллизационном растворе.....	80
3.2. Влияние температуры на стабильность кластера-прекурсора кристалла протеиназы К.....	85
3.2.1. <i>Влияние температуры на сохранение структурной целостности димера E</i> .....	86
3.2.2. <i>Влияние температуры на флуктуации (RMSF) атомов димера E</i> .....	87
3.2.3. <i>Кристаллизация протеиназы К при 60°C</i> .....	88
3.3. Заключение.....	90
<b>ГЛАВА 4. СТРУКТУРА И ТЕРМОСТАБИЛЬНОСТЬ ГЕКСАМЕРОВ, ОБРАЗУЮЩИХСЯ В КРИСТАЛЛИЗАЦИОННОМ РАСТВОРЕ ТЕРМОЛИЗИНА</b> .....	<b>92</b>
4.1. Уточнение структуры гексамера термолизина, образующегося в его кристаллизационном растворе.....	93
4.2. Влияние температуры на стабильность кластера-прекурсора кристалла термолизина .....	97
4.3. Заключение.....	103
<b>Выводы и основные результаты работы</b> .....	<b>105</b>
<b>Список литературы</b> .....	<b>107</b>

## ВВЕДЕНИЕ

Белки играют ключевую роль в живых организмах, выполняя широкий спектр функций: катализ биохимических реакций (ферменты), формирование клеточного каркаса (цитоскелет, коллаген), транспорт веществ (гемоглобин), передачу сигналов (гормоны, рецепторы), иммунную защиту (антитела), движение (актин, миозин), хранение и регуляция генетической информации (гистоны, транскрипционные факторы), а также резервное питание (альбумины, казеин). Эти процессы возможны благодаря уникальной пространственной организации белковых молекул, при которой незначительные изменения в конформации могут радикально влиять на функциональность. Каталитические и регуляторные свойства белков определяются взаимным расположением аминокислотных остатков, формирующих активные центры ферментов и участки связывания лигандов [1]. Установление **пространственной структуры белков имеет как фундаментальное значение для понимания механизмов их функционирования, так и прикладное**, в частности, для направленного дизайна лекарственных препаратов (драг-дизайна). Кроме того, структурные данные используются в биотехнологиях для конструирования ферментов с заданными свойствами и создания биосенсоров, а в случае термофильных белков они необходимы для создания новых промышленных биокатализаторов, устойчивых к экстремальным условиям.

Установление атомарной 3D-структуры белков остается ключевой задачей структурной биологии, однако традиционные методы, такие как рентгеноструктурный анализ (РСА) и ядерный магнитный резонанс (ЯМР) и криоэлектронная микроскопия (крио-ЭМ), сталкиваются с рядом ограничений по разрешению и применимости. Например, РСА, с помощью которого установлено более 80 % структур [2], требует получения высококачественных монокристаллов, что часто становится непреодолимым барьером для мембранных белков или белков с неупорядоченными доменами [3]. Современные вычислительные подходы, включая алгоритмы машинного обучения, значительно

продвинулись в предсказании структур, но не учитывают влияние лигандов, посттрансляционных модификаций и условий среды, требуя экспериментальной верификации [4].

Процесс кристаллизации белков является сложной экспериментальной задачей, чувствительной к множеству параметров: концентрации белка и осадителя, pH, температуре, ионной силе и другим факторам [5]. Несмотря на автоматизацию скрининга условий, поиск оптимальных параметров часто проводится эмпирически, методом проб и ошибок, что делает его трудоемким и дорогостоящим [6]. Получение кристаллов термофильных белков представляет особую методологическую сложность, поскольку их стабильность часто требует поддержания высоких температур. Это обуславливает необходимость как систематического изучения механизмов их кристаллизации, так и разработки специализированных подходов к высокотемпературной кристаллизации [7].

Тем не менее, экспериментальные методы, такие как малоугловое рассеяние рентгеновских лучей (МУРР) и нейтронов (МУРН), выявили наличие промежуточной фазы кристаллизации белков, характеризующейся образованием олигомеров — стабильных ассоциатов белков, предположительно являющихся трехмерными фрагментами будущей кристаллической структуры [8–10]. Более того, установлено, что наличие таких олигомеров или, так называемых «кластеров-прекурсоров», является необходимым, хотя и недостаточным условием появления кристалла [10].

**Обнаружение критической роли кластеров-прекурсоров в формировании кристаллической решетки белка** указывает на особую важность их изучения. Отсутствие детального понимания их структурной организации и динамического поведения существенно ограничивает дальнейшую разработку обоснованных стратегий получения белковых кристаллов. Для преодоления этого барьера необходимо комплексно исследовать:

- молекулярные механизмы образования и диссоциации олигомеров;
- роль специфических межмолекулярных взаимодействий (водородных связей,

электростатических и гидрофобных контактов) в агрегации белков;

- влияние параметров раствора (температура, pH, ионный состав и т. д.) на стабильность промежуточных структур;
- механизмы встраивания кластеров-прекурсоров в формирующуюся кристаллическую решетку.

Компьютерное моделирование, в частности метод молекулярной динамики (МД), представляет собой перспективный инструмент для решения таких задач, поскольку оно позволяет:

- изучать и визуализировать процессы с атомарным разрешением;
- анализировать временную динамику на фемтосекундном масштабе;
- контролируемо варьировать практически любые параметры раствора (температуру, pH, ионный состав и т. д.);
- анализировать межмолекулярные взаимодействия.

Такой подход обеспечивает уникальное сочетание атомарного и временного разрешения, недоступного для большинства экспериментальных методов, что открывает новые возможности для изучения предкристаллизационной фазы растворов. Примером эффективности данной методики стали исследования двух типов октамеров лизоцима с помощью МД [11], в ходе которых был выявлен стабильный тип октамера, образующийся в кристаллизационном растворе лизоцима. Эти результаты продемонстрировали значительный потенциал МД-моделирования для детального изучения ранних этапов кристаллизации.

Тем не менее, традиционная МД, использующая атомистические модели, эффективна для изучения локальных процессов (стабильность конкретного олигомера) на временах до нескольких микросекунд, но не способна охватить процессы самосборки олигомеров из мономеров или их дальнейшей агрегации. Кроме того, применение классической МД для изучения белковой кристаллизации требует тщательной проработки методических аспектов

моделирования, таких как определение оптимальных размеров ячейки моделирования во избежание артефактов.

Таким образом, в настоящее время особую актуальность приобретает развитие и применение методов молекулярного моделирования, позволяющих получить детальную картину поведения белков на ранних стадиях кристаллизации, в частности, исследовать механизмы формирования, стабильность и динамику олигомерных прекурсоров кристаллов. Понимание этих процессов необходимо для перехода от традиционного эмпирического поиска условий кристаллизации к принципиально новому рациональному дизайну экспериментов.

**Цель работы:** развитие и применение подхода к установлению особенностей самоорганизации белковых молекул в кристаллизационных растворах, основанного на оценке стабильности олигомеров белков с помощью метода молекулярной динамики.

Для достижения поставленной цели необходимо было решить следующие **задачи:**

1. Оценить возможность использования силового поля MARTINI – крупнозернистого описания модели, позволяющего ускорить расчеты по сравнению с атомистическим подходом, – для моделирования кристаллизационных растворов белков.
2. Определить минимальные размеры ячейки моделирования, исключающие артефакты периодических граничных условий (для ускорения МД-расчетов кристаллизационных растворов белков).
3. Уточнить структуру кластеров-прекурсоров (олигомеров) кристаллов протеиназы К и термолизина.
4. Определить влияние температуры на стабильность выявленных кластеров-прекурсоров кристаллов протеиназы К и термолизина (в диапазонах температур 20–80°C и 10–90°C соответственно).

### **Научная новизна.**

В работе впервые:

1. Доказана возможность использования крупнозернистого силового поля MARTINI для МД-моделирования олигомеров глобулярных белков в кристаллизационных растворах;
2. Определены минимально допустимые размеры ячейки моделирования для исследования олигомеров глобулярных белков в кристаллизационных растворах с помощью МД;
3. Установлены наиболее стабильные олигомеры (кластеры-прекурсоры), выделенные из кристаллических структур протеиназы К и термолизина;
4. Определена зависимость стабильности выявленных кластеров-прекурсоров кристаллов протеиназы К и термолизина от температуры. Обнаружено, что они стабильны в широком диапазоне температур: 20–60 и 10–90°C соответственно;
5. Получены кристаллы протеиназы К после инкубирования ее раствора при 60°C – температуре, предсказанной с помощью МД;
6. Показана предсказательная способность развитого подхода к установлению новых перспективных температур белковой кристаллизации, основанного на оценке стабильности кластеров-прекурсоров кристаллов белков с помощью МД.

### **Практическая значимость.**

Развитый в работе подход к исследованию кристаллизационных растворов белков, основанный на сочетании крупнозернистой и атомистической МД, позволяет раскрыть детали механизмов кристаллизации, недоступные или труднодоступные для экспериментальных измерений.

Данный подход предоставляет возможность моделирования различных этапов белковой кристаллизации за счет оптимального выбора силового поля (уровня детализации модели), позволяя:

- моделировать длительные процессы (например, агрегацию олигомеров) ускоренно с помощью крупнозернистой МД;

- моделировать локальные процессы (например, диссоциацию олигомеров) с помощью атомистической МД при оптимальных параметрах (минимально допустимых размерах ячейки моделирования).

Получение кристаллов протеиназы К после инкубации при 60°C – температуре, предсказанной на основе анализа термостабильности кластера-прекурсора ее кристалла с помощью МД, – не только подтверждает перспективность развитого в работе МД-подхода для теоретически обоснованного выбора температуры кристаллизации белков, но и позволит установить пространственную структуру протеиназы К в нативном состоянии.

#### **Положения, выносимые на защиту:**

1. Подтверждение применимости крупнозернистого силового поля MARTINI для ускоренного молекулярно-динамического моделирования поведения олигомеров глобулярных белков в кристаллизационных растворах.
2. Для корректной оценки стабильности олигомеров глобулярных белков в кристаллизационных растворах на основе атомистического молекулярно-динамического моделирования, минимальное расстояние между атомами белка и гранью периодической ячейки должно составлять не менее одного радиуса отсечки невалентных взаимодействий.
3. Среди шести типов димеров, выделенных из кристаллической структуры протеиназы К, и четырех типов гексамеров, выделенных из кристаллической структуры термолизина, наибольшей стабильностью (по данным молекулярно-динамического моделирования) обладают димер типа «Е» и гексамер типа «А».
4. Димер типа «Е» протеиназы К и гексамер типа «А» термолизина сохраняют стабильность в температурных пределах 20–60 и 10–90°C соответственно.

#### **Личный вклад автора:**

Все результаты, представленные в работе, получены лично автором или при ее непосредственном участии.

Автор непосредственно принимала участие в построении всех указанных в диссертационной работе моделей олигомеров белков (лизоцима, протеиназы К и

термолизина); проведении их моделирования методом молекулярной динамики; обработке, анализе и интерпретации результатов вычислений.

Обсуждение результатов и их интерпретация проводились совместно с научным руководителем и соавторами публикаций.

**Апробация работы:** основные результаты были представлены на международных и российских конференциях:

1. **Ю.В. Кордонская**, В.И. Тимофеев, Ю.А. Дьякова, М.А. Марченкова, Ю.В. Писаревский, Д.Д. Подшивалов, А.С. Бойкова, М.В. Ковальчук «Исследование начальной стадии кристаллизации лизоцима методом молекулярной динамики» / 53-я Школа ПИЯФ по Физике Конденсированного Состояния (Москва, 2019);

2. **Кордонская Ю.В.**, Тимофеев В.И., Ильина К.Б., Марченкова М.А., Подшивалов Д.Д., Бойкова А.С., Дьякова Ю.А., Писаревский Ю.В., Ковальчук М.В. «Исследование начальной стадии процесса кристаллизации белка лизоцима методом молекулярной динамики» / VIII-я международная молодежная научная школа-конференция "Современные проблемы физики и технологий", НИЯУ МИФИ (Москва, 2019);

3. **Ю.В. Кордонская**, В.И. Тимофеев, Ю.А. Дьякова, М.А. Марченкова, Ю.В. Писаревский, М.В. Ковальчук «Моделирование молекулярной динамики олигомеров лизоцима в растворе при различных температурах» / XVI Курчатовская междисциплинарная молодежная научная школа, НИЦ «Курчатовский институт» (Москва, 2019);

4. **Ю.В. Кордонская**, В.И. Тимофеев, Ю.А. Дьякова, М.А. Марченкова, Ю.В. Писаревский, М.В. Ковальчук «Исследование стабильности олигомеров лизоцима в предкристаллизационном растворе методами молекулярной динамики» / 63-я Всероссийская научная конференция МФТИ (2020);

5. **Yu.V. Kordonskaia**, V.I. Timofeev, Yu.A. Dyakova, M.A. Marchenkova, Yu.V. Pisarevsky, M.V. Kovalchuk «Precipitant ions influence on lysozyme mono- and

oligomers stability investigated by molecular dynamics simulation» / Moscow Conference on Computational Molecular Biology (Moscow, 2021);

6. **Yu.V. Kordonskaia**, V.I. Timofeev, Yu.A. Dyakova, M.A. Marchenkova, Yu.V. Pisarevsky, M.V. Kovalchuk «Study of the behaviour of lysozyme oligomers in solutions by the molecular dynamics method at different temperatures» / 25th General Assembly and Congress of the International Union of Crystallography 2021 (Prague, 2021);

7. **Ю.В. Кордонская**, В.И. Тимофеев, Ю.А. Дьякова, М.А. Марченкова, Ю. В. Писаревский, М. В. Ковальчук «Влияние размера ячейки моделирования и концентрации ионов осадителя на поведение димера тетрагонального лизоцима» / Кластер конференций 2021 (Иваново, 2021);

8. **Ю.В. Кордонская**, В.И. Тимофеев, Ю.А. Дьякова, М.А. Марченкова, Ю. В. Писаревский, М. В. Ковальчук «Изменение свободной энергии при формировании кристаллического контакта между мономерами лизоцима при различных физико-химических условиях» / 64-я Всероссийская научная конференция МФТИ (Москва, 2021)

9. **Ю.В. Кордонская**, В.И. Тимофеев, Ю.А. Дьякова, М.А. Марченкова, Ю.В. Писаревский, М.В. Ковальчук «Молекулярная динамика олигомеров лизоцима в кристаллизационных растворах с различными осадителями» / XVII Международная научная конференция «Актуальные вопросы биологической физики и химии. БФФХ-2022» (Севастополь, 2022);

10. **Ю.В. Кордонская**, В.И. Тимофеев, Ю.А. Дьякова, М.А. Марченкова, Ю.В. Писаревский, М.В. Ковальчук "Определение кластера-прекурсора кристаллов белка протеиназы К" / OPENBIO-2022 (Наукоград Кольцово, 2022);

11. **Ю.В. Кордонская**, В.И. Тимофеев, Ю.А. Дьякова, М.А. Марченкова, Ю.В. Писаревский, М.В. Ковальчук «Установление кластера-прекурсора кристалла олигопептидазы» / Синхротронные и нейтронные методы исследования конденсированных фаз (Москва, 2024);

12. **Ю.В. Кордонская**, В.И. Тимофеев, Ю.А. Дьякова, М.А. Марченкова, Ю.В. Писаревский, М.В. Ковальчук «Кластеры-прекурсоры кристаллов протеиназы К и термолизина: определение и температурная стабильность» / *Материалы будущих технологий* (Москва, 2025).

**Публикации:** основные результаты, вошедшие в диссертационную работу, изложены в 7-ми публикациях, размещенных в рецензируемых научных изданиях из списка ВАК (и индексируемых в базах данных Scopus и Web of Science).

A1. **Ю.В. Кордонская**, В.И. Тимофеев, Ю.А. Дьякова, М.А. Марченкова, Ю.В. Писаревский, М.В. Ковальчук «Влияние размера ячейки моделирования и концентрации ионов осадителя на поведение димера тетрагонального лизоцима» // *Кристаллография*. 2021. Т. 66. №3. С. 478.

A2. **Yu.V. Kordonskaya**, V.I. Timofeev, M.A. Marchenkova, P. V. Konarev «Identification of the Precursor Cluster in the Crystallization Solution of Proteinase K Protein by Molecular Dynamics Methods» // *Crystals*. 2022. V. 12. I. 4. P. 484.

A3. **Yu.V. Kordonskaya**, V.I. Timofeev, Yu.A. Dyakova, M.A. Marchenkova, Yu.V. Pisarevsky, Silvestrova, S.Y., M.V. Kovalchuk. «Unusual Temperature Behavior of Stability of Proteinase K Dimer Formed in Crystallization Solution Defined by Molecular Dynamics» // *Crystals*. 2022. V. 12. I. 11. P. 1645.

A4. **Yu.V. Kordonskaia**, V.I. Timofeev, Yu.A. Dyakova, M.A. Marchenkova, Yu.V. Pisarevsky, S.Y. Silvestrova, M.V. Kovalchuk. «Identification of the precursor cluster in thermolysin crystallization solution by molecular dynamics methods» // *Mendeleev Commun*. 2023. V. 33. I. 2. P. 225.

A5. **Ю.В. Кордонская**, В.И. Тимофеев, Ю.А. Дьякова, М.А. Марченкова, Ю.В. Писаревский, М.В. Ковальчук «Влияние температуры на стабильность кластера-прекурсора кристалла термолизина» // *Кристаллография*. 2024. Т. 69. № 4. С. 694.

A6. **Ю.В. Кордонская**, В.И. Тимофеев, Ю.А. Дьякова, М.А. Марченкова, Ю.В. Писаревский, М.В. Ковальчук «Проверка применимости крупнозернистого

силового поля MARTINI для моделирования белковых олигомеров в кристаллизационном растворе» // *Кристаллография*. 2024. Т. 69. № 5. С. 885.

А7. **Ю.В. Кордонская**, И.Ф. Гарипов, В.И. Тимофеев, М.А. Марченкова, Ю.А. Дьякова, Ю.В. Писаревский, М.В. Ковальчук «Влияние размеров ячейки моделирования на стабильность димеров и гексамеров лизоцима в кристаллизационном растворе» // *Российские нанотехнологии*. 2024. Т. 19. № 3. С. 291.

## ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ

Белки выполняют широкий спектр биологических функций, включая каталитическую, структурную, транспортную, регуляторную и защитную. Эти функции напрямую зависят от трехмерной структуры молекулы, которая, в свою очередь, формируется из аминокислотной последовательности. Активные центры ферментов, участки связывания лигандов и олигомерные интерфейсы формируются в результате специфического взаимного расположения аминокислотных остатков, и малейшие изменения в укладке белка могут радикально повлиять на его функцию. Ключевая роль пространственной структуры белка в реализации его биологических функций делает **определение структуры белков фундаментальной задачей современной структурной биологии** и важной основой для практических применений, таких как рациональный дизайн лекарственных препаратов [12].

Особую роль структура играет у белков термофильных организмов, оптимальная температура функционирования которых составляет 60–80°C (у гипертермофилов она выше 80°C). Термофильные белки часто обладают повышенной жесткостью полипептидной цепи и стабильностью в экстремальных условиях, что делает их ценными биокатализаторами для пищевой, целлюлозно-бумажной, фармацевтической промышленности, а также производства биотоплива и компостирования. При этом по сравнению с мезофильными белками — термофильные остаются существенно менее изученными. Согласно данным Protein Data Bank (PDB, [2]), они составляют лишь ~1 % от общего числа депонированных структур, что объясняется комплексом методологических трудностей:

- сложность культивирования экстремофильных организмов-продуцентов, а также выделения из них термофильных белков и очистки таких белков при сохранении нативной структуры;

- необходимость поддержания стабильных высокотемпературных условий на всех этапах работы и сохранения стабильности белков при переходе от нативных условий к кристаллизационным.

### 1.1. Вычислительные подходы к предсказанию белковых структур

Рентгеноструктурный анализ (РСА) является основным экспериментальным методом определения пространственной организации белков с высоким разрешением. Более 80 % (из представленных в базе PDB [2]) экспериментально установленных структур, получены именно этим способом. Тем не менее, РСА требует получения хорошо упорядоченных кристаллов, что зачастую становится основным ограничивающим фактором, особенно для мембранных и гибких белков. В качестве альтернатив выступают ядерный магнитный резонанс и криоэлектронная микроскопия. ЯМР подходит для исследования небольших растворимых белков (до ~50 кДа) в условиях, приближенных к физиологическим, обеспечивая уникальными данными о динамике молекулы в растворе. Однако он уступает РСА и крио-ЭМ в пространственном разрешении и неприменим для крупных комплексов. Революция в крио-ЭМ за последние годы повысила разрешение метода (до ~1.2 Å), однако оно по-прежнему ниже, чем у РСА (~1.0–1.5 Å для белков). Кроме того, крио-ЭМ хуже подходит для выявления альтернативных конформаций и изучения малых белков (<100 кДа).

Сложности, связанные с экспериментальными барьерами, обусловили необходимость разработки вычислительных подходов к предсказанию трехмерной структуры белков на основе их аминокислотной последовательности [13]. Ранние методы основывались на физико-химических и структурных принципах. Среди них можно выделить:

- Гомологичное моделирование, которое основывается на предположении, что белки с близкими последовательностями имеют сходную пространственную структуру. Метод включает построение модели целевого белка на основе одной или нескольких известных структур-гомологов. Он демонстрирует

высокую точность при наличии родственной структуры с идентичностью последовательности выше 30–40 %.

- Методы *de novo* и *ab initio*, использующие фундаментальные физические принципы. Они предполагают поиск глобального минимума свободной энергии среди возможных конформаций белковой цепи. Наиболее известный подход в этом направлении — фрагментарное моделирование, реализованное, например, в пакете Rosetta [14]. Однако эти методы крайне ресурсоёмки, и их применение ограничено малым числом аминокислотных остатков.
- Молекулярная динамика, которая позволяет исследовать поведение различных стартовых конформаций белковой цепи во времени. При наличии соответствующих силовых полей МД может использоваться как средство поиска устойчивых структур, особенно если комбинируется с другими методами, например метадиNAMикой [15], для ускорения переходов между состояниями.

Несмотря на свою фундаментальную основу, эти подходы имеют ограничения: они зависят от качества исходных моделей, энергетических функций и алгоритмов поиска по конформационному пространству. Кроме того, в отсутствие структурных гомологов или при высокой гибкости целевого белка точность предсказания снижается [16].

В последние годы произошел прорыв в области предсказания структур, связанный с развитием алгоритмов машинного обучения, в частности — нейросетевых моделей глубокого обучения. Эти подходы основаны на выявлении сложных эмпирических закономерностей в больших массивах данных о структуре белков. Яркий пример — алгоритм AlphaFold2 [17], который выявляет статистические корреляции между аминокислотными последовательностями родственных белков и достигает точности, близкой к экспериментальной. Данная программа генерирует модель с трехмерными координатами атомов белка и предоставляет оценку локальной достоверности предсказания для каждого

остатка, что позволяет критически оценить надежность различных участков модели [17]. Возможности AlphaFold2 включают:

- высокую точность предсказания для стабильных глобулярных доменов, часто достигающую суб-ангстремного разрешения;
- способность генерировать модели даже при отсутствии близких структурных гомологов, если доступны достаточно информативные многомерные выравнивания (MSA);
- значительно более высокую скорость получения результата (от минут до часов) по сравнению с экспериментальными методами;
- возможность предсказания структуры мультибелковых комплексов с помощью специализированной версии AlphaFold-Multimer [18].

Однако AlphaFold2 не является универсальным решением и имеет ряд ограничений, подчеркивающих необходимость экспериментальных данных:

- переменная точность: точность существенно ниже для неупорядоченных или гибких участков полипептидной цепи, некоторых трансмембранных белков и белков с недостаточным количеством гомологичных последовательностей в базах данных;
- статичность модели: алгоритм предсказывает одну или несколько наиболее вероятных статичных конформаций, не отражая конформационную динамику, внутренние движения или структурные переходы, которые могут быть критичными для функциональной активности белка;
- изолированность молекулы: предсказание выполняется для изолированной белковой цепи и не учитывает явным образом влияние специфического окружения (pH, ионный состав, лиганды, взаимодействие с мембранами).

Несмотря на ограничения, методы машинного обучения, такие как AlphaFold2, являются ценным дополнением к существующим экспериментальным подходам. Они предоставляют высококачественные начальные структуры для

детального исследования методами МД (например, для изучения динамики, взаимодействия с лигандами или поведения в специфических условиях) [19]. Предсказанные модели существенно помогают в интерпретации экспериментальных данных (крио-ЭМ, МУРР) и ускоряют генерацию и проверку гипотез в структурной биологии [20].

Однако, несмотря на впечатляющий прогресс таких методов, они лучше подходят для компактных глобулярных доменов и хуже справляются с гибкими или разупорядоченными участками полипептидной цепи [21,22]. Кроме того, успех применения AlphaFold2 во многом зависит от наличия информативных MSA близкородственных последовательностей [23]. Таким образом, **методы машинного обучения — это мощное дополнение, но не альтернатива экспериментальным техникам** определения пространственной структуры белков.

## 1.2. Белковая кристаллизация

**Рентгеноструктурный анализ** по-прежнему является **наиболее надежным и точным методом определения пространственной организации** белковых молекул. Высокое разрешение позволяет локализовать водородные связи и ионы, но требует выращивания крупных и высокоупорядоченных монокристаллов. Хотя РСА стал доступным и широко применяемым инструментом, именно этап получения пригодных для анализа кристаллов остаётся наиболее трудоемким, непредсказуемым и ограничивающим фактором всего процесса. Кристаллизация белков в значительной степени до сих пор осуществляется во многом эмпирически — методом подбора условий, опирающегося на предыдущий опыт и результаты. Даже при автоматизированном скрининге условий лишь малая доля проб приводит к пригодным для дифракции образцам, а воспроизводимость экспериментов низка из-за зависимости от множества тонких параметров [24,25].

С накоплением знаний о механизмах зарождения и роста белковых кристаллов были разработаны более рациональные подходы к их получению и

повышению дифракционного качества. Тем не менее, универсальной методологии, обеспечивающей успех, пока не существует, что обусловлено комплексом факторов, включая разнообразие специфических физико-химических свойств белков и их чувствительность к условиям среды.

Процесс кристаллизации, как правило, начинается с этапа скрининга условий, включающего систематическое варьирование компонентов и условий раствора: типа и концентрации осадителя, буфера, pH, температуры, ионной силы и др. Общепринятой практикой выращивания кристалла белка является постепенное снижение его растворимости до состояния пересыщения [26]. Это осуществляется путем добавления в белковый раствор осадителей: концентрированных растворов солей, амфифильных органических соединений, а также водорастворимых полимеров, главным образом полиэтиленгликолей различной молекулярной массы [27].

Механизм действия осадителей заключается в изменении характера межмолекулярного взаимодействия, обусловленным как перераспределением зарядов вблизи поверхности белка (за счет экранирования поверхностных зарядов молекул [28]), так и модификацией свойств растворителя. При высоких ионных силах реализуется эффект «высаливания», при котором ионы соли конкурируют с белком за молекулы воды, что приводит к дестабилизации гидратной оболочки белка и увеличению вероятности образования межмолекулярных контактов [29]. Органические растворители, дополнительно снижая диэлектрическую проницаемость среды, усиливают электростатическое притяжение между заряженными группами на поверхности белка. Полимеры высокого молекулярного веса (такие как полиэтиленгликоль), в свою очередь, индуцируют оба этих эффекта одновременно [30].

Следствием такого воздействия осадителей является переход от преобладающего электростатического отталкивания к эффективным притягивающим взаимодействиям, что способствует формированию олигомерных агрегатов, выступающих прекурсорами кристаллического ядра [28]. Данный

принцип лежит в основе рационализованного подхода к подбору кристаллизационных условий, заключающемся в предварительном анализе потенциальных кристаллизационных растворов до появления в них кристаллов с помощью МУРР [8]. Условия, при которых обнаружены белковые кластеры, рассматриваются как наиболее перспективные для последующей кристаллизации. Растворы, содержащие исключительно мономеры (или функциональные олигомеры), напротив, исключаются из дальнейшего эксперимента. Этот подход позволяет ускорить и направить поиск оптимальных параметров кристаллизации. В заключительном разделе литературного обзора будут рассмотрены детали исследований предкристаллизационной фазы раствора лизоцима с помощью сочетания методов МУРР и МД [11,31].

Помимо состава среды, на растворимость и кристаллизацию белка существенно влияют такие факторы, как рН и температура. Допустимый диапазон этих параметров ограничен пределами стабильности нативной конформации молекулы [27].

С термодинамической точки зрения, переход белка из раствора в кристаллическое состояние сопровождается понижением энтропии, поскольку упорядоченное расположение молекул в решетке ограничивает их степени свободы. Однако этот процесс компенсируется снижением свободной энергии за счет образования многочисленных специфических межмолекулярных связей, что и служит движущей силой кристаллизации.

Белковые кристаллы обладают рядом особенностей, отличающих их от кристаллов других молекул. Из-за хиральности белков (состоящих исключительно из L-аминокислот) и отсутствия центра инверсии, они могут кристаллизоваться лишь в 65 из 230 возможных пространственных групп. При этом 90 % белков кристаллизуются всего в 10 группах. Сложный рельеф поверхности белковой молекулы и преобладание полярных и заряженных участков ведут к ее интенсивному связыванию с молекулами воды и ионами, что приводит к высокому содержанию растворителя в кристалле (30–90 %). Кроме

того, плотность межмолекулярных контактов в белковом кристалле значительно ниже, чем в кристаллах низкомолекулярных соединений, что делает их особенно хрупкими и чувствительными к механическим воздействиям.

Процесс кристаллизации условно подразделяется на две стадии: нуклеация (образование зародышей) и рост кристаллов [32]. Эти стадии отражены на фазовой диаграмме системы «белок–осадитель», которая строится в координатах концентрации белка и концентрации осадителя (рис. 1, [6]).

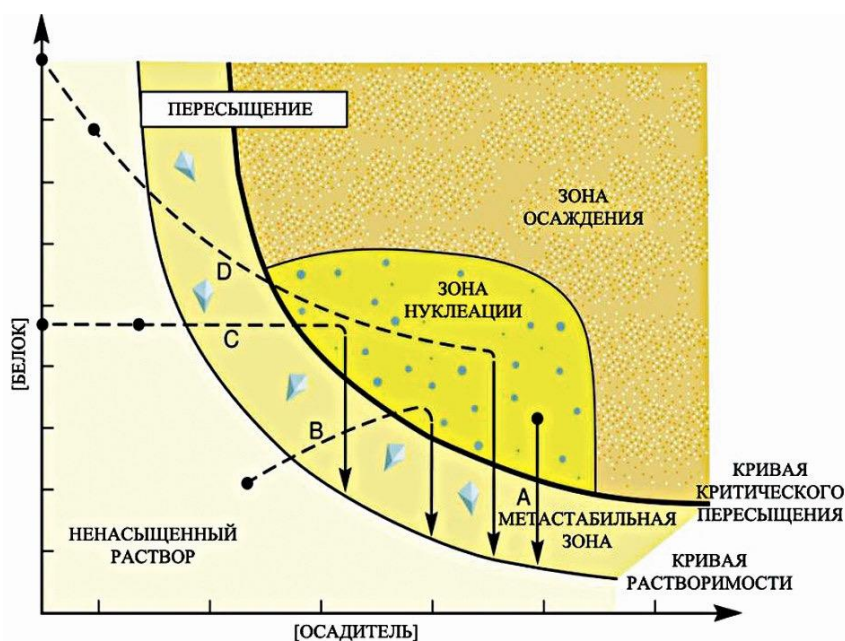


Рисунок 1 – Фазовая диаграмма, отражающая состояние системы «белок–осадитель» в зависимости от соотношения их концентраций. Стрелками схематично обозначено протекание кристаллизации в соответствии с наиболее распространенными методами: объемная кристаллизация (А), диффузия паров (В), диализ (С) и свободная диффузия (D) [6].

Зародыши кристалла, как правило, образуются в лабильной зоне (нуклеации): при очень высоких степенях пересыщения, достигающих 1000 %. Из-за чрезмерного пересыщения в зоне осаждения возможно образование аморфного осадка или поликристаллов. При гомогенной нуклеации объединяются сами белковые молекулы: по мере увеличения уровня пересыщения повышается вероятность столкновений молекул и формирования кластеров, которые при достижении критического размера становятся устойчивыми ядрами роста. При

гетерогенной нуклеации зародыш формируется на твердых поверхностях, которые снижают энергетический барьер образования кристаллического ядра.

Метастабильная зона фазовой диаграммы характеризуется умеренной степенью пересыщения, достаточной для роста кристаллов, но при которой нуклеация маловероятна. Именно в этой зоне создаются условия для роста упорядоченных кристаллов, которые растут при минимальной скорости агрегации, поскольку «неправильно» встроившиеся в кристаллическую решетку молекулы могут диссоциировать и повторно присоединиться в ориентации, соответствующей упаковке кристалла.

Для улучшения качества кристаллов возможно разделение стадий нуклеации и роста, например, путем введения заранее полученных зародышей в условия метастабильной зоны. Прекращение роста кристалла происходит при достижении равновесного состояния между кристаллом и раствором.

Ключевым фактором, влияющим на морфологию и совершенство кристалла, является механизм массопереноса — транспорта единиц роста к поверхности кристалла. В естественных (земных) гравитационных условиях массоперенос включает как диффузию, так и конвекцию [33,34]. Последняя возникает из-за градиентов плотности и температуры, возникающих при кристаллизации. Такие потоки могут приводить к неравномерному пересыщению на различных гранях кристалла, ускоренному росту, образованию дефектов и сростков. Для подавления конвекции и перехода к более однородному диффузионному росту применяются методы микрогравитационной кристаллизации, включая эксперименты на борту МКС [33].

Особый интерес представляет задача кристаллизации термофильных белков, характеризующихся высокой термической стабильностью, поскольку выращивание их кристаллов возможно при относительно высоких температурах, например, при 65°C [35], что значительно выше традиционных температур белковой кристаллизации: ~4–23°C. Такая высокотемпературная кристаллизация обладает рядом преимуществ [35]:

- увеличение размера и качества кристаллов — кристаллы, выращенные при высоких температурах, оказываются крупнее по сравнению с традиционными методами, что способствует улучшению разрешающей способности и визуализации структурных деталей (до 1.7 Å);
- сокращение времени кристаллизации — благодаря высокой термической прочности, процесс формирования кристаллов происходит быстрее, иногда за 2 часа, что значительно экономит время эксперимента по сравнению с обычными методами (обычно 5 дней);
- улучшение разрешающей способности структуры — использование высоких температур способствует получению более четкого изображения электронной плотности;
- поддержание функциональной конформации белка — высокотемпературная кристаллизация позволяет сохранять нативную структуру термофильных белков, что важно для изучения механизмов их функционирования;
- возможность кристаллизации сложных и крупных комплексов термофильных белков — высокая температура помогает кристаллизовать более крупные молекулы, в том числе ферменты со сложными субстратами или олигосахаридами, что обычно затруднено при низких температурах.

Однако при получении кристаллов термофильных белков можно столкнуться и со специфическими сложностями:

- проблемы с контролем растворимости — при повышенных температурах растворимость белков меняется, и нередко возникает необходимость в специальных условиях или добавках, чтобы предотвратить сворачивание или агрегацию, к тому же не все растворители и буферные растворы подходят для высокотемпературной кристаллизации;
- контроль скорости испарения и осаждения — при высоких температурах интенсивно происходит испарение воды, что усложняет поддержание правильного уровня насыщения и баланс между скоростью кристаллизации и ростом больших, качественных кристаллов;

- технические ограничения — требуются специальные установки для нагрева при высоких температурах и условия, при которых кристаллы не разрушатся или не испарятся полностью, а также технические сложности по управлению микроокружением внутри камеры.

Таким образом, **получение пригодного для РСА кристалла остается самым непредсказуемым этапом в структурной биологии**, во многом определяющим успех всего исследования. Сложность белковой кристаллизации обусловлена как особенностями белковых молекул, так и спецификой их кристаллов. До настоящего времени этот процесс невозможно полностью подчинить формализованным подходам, что подчеркивает важность дальнейших исследований в данной области [36].

### 1.3. Молекулярная динамика как инструмент исследования биомолекулярных систем

Метод молекулярной динамики (МД) зародился как вычислительный подход для изучения движения атомов и молекул, его первое применение к моделированию белка (бычьего панкреатического ингибитора трипсина) было опубликовано в 1977 г. [37]. За прошедшие десятилетия МД претерпела колоссальное развитие, превратившись из специализированного инструмента для теоретиков в неотъемлемую часть современной структурной биологии, биофизики и дизайна лекарств. В 2013 году М. Карплюсу, М. Левитту и А. Уоршелу была присуждена Нобелевская премия по химии *«за развитие многомасштабных моделей сложных химических систем»* — это признание во многом основывалось на их пионерских работах в области МД-моделирования биомолекул [38,39].

Сегодня МД является одним из ведущих вычислительных подходов в молекулярной биологии, успешно дополняя экспериментальные методы, такие как РСА, крио-ЭМ, ЯМР и электронный парамагнитный резонанс. Практически все современные структуры, полученные с помощью РСА или ЯМР, проходят

стадию уточнения и валидации с помощью МД, поскольку только в рамках симуляции можно детально воспроизвести и проанализировать поведение молекулы в условиях, приближенных к реальным экспериментальным (температура, давление, ионная сила, состав растворителя и т. д.).

В отличие от экспериментальных методов, часто предоставляющих усредненную или статичную картину, МД моделирует «молекулярное видео» исследуемой системы с фемтосекундным временным и атомистическим пространственным разрешением. Это позволяет не только визуализировать динамику каждого атома, но и количественно описывать широкий спектр фундаментальных биомолекулярных процессов, включая конформационные переходы белков, механизмы их фолдинга и агрегации, процессы связывания с лигандами, а также понимать, как различные факторы (мутации, посттрансляционные модификации, изменения во внешней среде) влияют на структуру, динамику и функциональную активность биомакромолекул [19].

К основным областям применения МД в биофизике и структурной биологии относятся:

- изучение динамики биомолекулярных систем: отслеживание траекторий атомов дает представление о естественных движениях биомолекул в растворе – от быстрых локальных флуктуаций до медленных крупномасштабных конформационных изменений, происходящих на различных временных масштабах;
- расчет термодинамических характеристик: согласно эргодической гипотезе, усреднение термодинамических параметров системы по времени (аналогичное усреднение по достаточно длинной МД-траектории) эквивалентно их усреднению по статистическому ансамблю. Это позволяет получать расчетные величины, близкие к экспериментально измеряемым параметрам и делает МД незаменимой для оценки свободной энергии различных процессов (например,

связывания лиганда, конформационных переходов, образования межмолекулярных комплексов) и других характеристик;

- исследование конформационного пространства: МД позволяет исследовать доступные энергетически выгодные конформационные состояния молекул и их комплексов. Это находит применение в задачах предсказания структуры гибких участков белков, моделирования связывания лигандов (докинг) и анализе конформационных переходов.

Следует отметить, что МД практически не имеет альтернативных методов исследования динамики, однако для задач, связанных с усредненными свойствами и изучением конформационного пространства, возможно использование методов Монте-Карло [40].

Стремительный рост популярности и значимости МД за последние десятилетия обусловлен несколькими взаимосвязанными факторами.

Во-первых, благодаря экспоненциальному росту числа экспериментально определенных структур биомакромолекул (прежде всего, благодаря успехам РСА и крио-ЭМ), которые служат надежными начальными моделями для симуляции и одновременно являются причиной возникновения новых вопросов (касающихся динамики и механизмов функционирования), на которые может помочь ответить моделирование.

Во-вторых, из-за бурного развития аппаратного обеспечения: от специализированных суперкомпьютеров до общедоступных высокопроизводительных графических процессоров, – и совершенствования программных пакетов (таких как GROMACS [41], NAMD [42], AMBER [43], CHARMM [44], Desmond [45] и OpenMM [46]), превративших МД из крайне ресурсоемкого метода, доступного лишь небольшому числу специалистов, в мощный и относительно доступный инструмент для широкого круга исследователей [47,48]. Современные программные комплексы значительно

проще в использовании и обладают развитыми средствами для подготовки моделей и анализа результатов, а также поддерживают различные силовые поля.

В-третьих, сами физические модели (силовые поля) и алгоритмы, лежащие в основе МД, претерпели значительное усовершенствование, что привело к существенному повышению точности и надежности получаемых результатов.

Основная идея МД-моделирования заключается в численном решении классических уравнений движения Ньютона для системы взаимодействующих частиц (атомов). Исходя из начальных положений и скоростей всех атомов в биомолекулярной системе (например, белок в водном окружении или липидном бислое), рассчитываются силы (согласно параметрам силового поля), действующие на каждый атом со стороны всех остальных атомов. Затем путем двойного интегрирования уравнений движения вычисляются новые координаты, которые становятся начальными для следующего итерационного шага (обычно порядка нескольких фемтосекунд). Последовательность таких конфигураций формирует траекторию системы – пошаговую эволюцию ее поведения во времени.

МД-моделирование является уникальным и эффективным исследовательским инструментом по нескольким причинам. Во-первых, оно позволяет отслеживать координаты и скорости каждого атома в каждый момент времени, предоставляя уровень детализации, недостижимый для большинства экспериментальных техник. Во-вторых, условия моделирования (начальная структура, состав окружения, наличие лигандов, мутаций, посттрансляционных модификаций, протонирование остатков, температура, давление и т. д.) точно известны и могут тщательно контролироваться исследователем, что позволяет изучать влияние отдельных факторов.

Общая схема типичного МД-моделирования состоит из следующих основных этапов:

1. Построение начальной модели системы, включающее получение или предсказание трехмерной структуры макромолекулы, добавление недостающих атомов (например, атомов водорода, часто отсутствующих в рентгеноструктурных данных), создание соответствующего окружения (молекулы растворителя, ионы для нейтрализации системы и создания нужной ионной силы, липидный бислой для мембранных белков и т.д.).
2. Минимизация потенциальной энергии системы, которая необходима для снятия стерических напряжений (наложений атомов) и нефизичных контактов в исходной модели.
3. Уравновешивание системы: постепенное доведение системы до заданных условий (температуры, давления) в рамках выбранного термодинамического ансамбля (например,  $NVT$ , а затем  $NPT$ ) для достижения равновесного состояния перед началом моделирования продуктивных данных.
4. «Продуктивное» МД-моделирование: симуляция, в ходе которой генерируется и записывается траектория движения системы (координаты, скорости, силы) в течение заданного времени.
5. Анализ результатов: обработка полученных траекторий, включающая расчет разнообразных структурных, динамических и термодинамических параметров, статистический анализ, визуализацию траектории.

При анализе результатов МД важно учитывать хаотическую природу моделируемых молекулярных систем. Незначительные изменения в начальных условиях (например, начальных скоростях атомов) могут приводить к существенно различным уникальным траекториям при сохранении общих статистических свойств. Поэтому для получения надежных выводов, особенно при сравнении различных систем или условий, обычно необходимо проведение нескольких независимых симуляций и последующий статистический анализ усредненных данных [19].

### 1.3.1. Физико-математический аппарат метода МД

Наиболее точное описание движения и взаимодействия атомов, учитывающее поведение электронов, требует решения уравнений квантовой механики. На сегодняшний день это возможно только для систем, содержащих небольшое количество атомов — до нескольких сотен. Однако при моделировании макроскопических биомолекулярных ансамблей, содержащих от  $10^3$  до  $10^9$  атомов, прямое квантово-механическое вычисление становится вычислительно неэффективным. Для таких масштабов метод МД заменяет «точные» электронные взаимодействия классическими моделями, в которых каждая частица представлена материальной точкой (в классической МД одной точке соответствует один атом) с массой  $m_i$ , координатой  $\mathbf{r}_i$  и скоростью  $\mathbf{v}_i$ , где  $i$  — номер атома. Динамика множества этих точек определяется вторым законом Ньютона ( $\mathbf{F}_i = m_i \mathbf{a}_i$ ) и вычисляется с помощью решения системы дифференциальных уравнений:

$$\begin{cases} m_i \frac{d\mathbf{v}_i}{dt} = \mathbf{F}_i \\ \frac{d\mathbf{r}_i}{dt} = \mathbf{v}_i \end{cases}, \quad i = 1 \dots N,$$

где  $N$  — число атомов,  $\mathbf{F}_i$  — суммарная сила, действующая на  $i$ -ый атом со стороны остальных атомов и внешних полей (если они заданы). Начальные координаты, как правило, определяются из экспериментальных структур (например, из данных РСА или крио-ЭМ). Для численного интегрирования уравнений движения обычно используется временной шаг  $\sim 2$  фс, чтобы надежно описать быстрые колебания С–Н связей (период  $\sim 11$  фс).

Если начальные скорости не заданы пользователем, то они могут быть сгенерированы, например, в соответствии со статистическим распределением Максвелла-Больцмана при определенной температуре:

$$p(v_i) = \sqrt{\frac{m_i}{2\pi kT}} \exp\left(-\frac{m_i v_i^2}{2k_b T}\right),$$

где  $p(v_i)$  – функция плотности вероятности,  $k_b$  – постоянная Больцмана.

Силы  $F_i$  вычисляются с помощью дифференцирования потенциальной энергии всей системы  $U(\mathbf{r}_1 \dots \mathbf{r}_N)$ :

$$\mathbf{F}_i = -\frac{\partial U(\mathbf{r}_1 \dots \mathbf{r}_N)}{\partial \mathbf{r}_i}.$$

Суммарная потенциальная энергия включает в себя вклады валентных  $U_{val}$  и невалентных  $U_{nonbond}$  взаимодействий:

$$U = U_{val} + U_{nonbond},$$

$$U_{val} = U_{bond} + U_{angle} + U_{tors}, \quad U_{nonbond} = U_{LJ} + U_{coulomb}.$$

Каждый потенциал описывается функцией, соответствующей типу взаимодействия. Ковалентная связь между двумя атомами представляется гармоническим потенциалом, где  $b_{ij}$  – равновесная длина связи,  $k_{ij}^b$  – постоянный коэффициент жесткости:

$$U_{bond}(r_{ij}) = \frac{1}{2} k_{ij}^b (r_{ij} - b_{ij})^2.$$

Уравнение, описывающее потенциал валентного угла  $\theta_{ijk}$  между тремя атомами также представлено гармоническим потенциалом:

$$U_{angle}(\theta_{ijk}) = \frac{1}{2} k_{ijk}^\theta (\theta_{ijk} - \theta_{ijk}^0)^2.$$

Потенциал торсионных углов  $\phi_{ijkl}$  задается периодической Фурье-функцией:

$$U_{tors}(\phi_{ijkl}) = k_\phi (1 + \cos(n\phi - \phi_s)).$$

Невалентные взаимодействия описываются потенциалами Леннарда–Джонса и Кулона. Потенциал Леннарда–Джонса иногда заменяется потенциалом Букингема, который более реалистичен в области сил отталкивания, но его вычисление требует больших затрат:

$$U_{LJ}(r_{ij}) = \frac{C_{ij}^{(12)}}{r_{ij}^{12}} - \frac{C_{ij}^{(6)}}{r_{ij}^6},$$

$$U_{Bh}(r_{ij}) = A_{ij} \exp(-B_{ij}r_{ij}) - \frac{C_{ij}}{r_{ij}^6}.$$

Кулоновское взаимодействие между двумя заряженными частицами определяется выражением:

$$U_{coulomb}(r_{ij}) = k \frac{q_i q_j}{\varepsilon_r r_{ij}},$$

где  $q_i, q_j$  – парциальные заряды атомов,  $\varepsilon_r$  – диэлектрическая проницаемость среды,  $k = 1/4\pi\varepsilon_0$  – постоянная Кулона.

В настоящем разделе были рассмотрены основные потенциальные функции, используемые в МД, вид и количество которых зависят от выбора силового поля, поскольку оно содержит как аналитические выражения для расчета потенциальной энергии, так и набор параметров и констант (таких как  $m_i, k_{ij}^b, b_{ij}, k_{ijk}^\theta$  и др.), которые индивидуальны для каждого типа атомов и определяются экспериментально или с помощью квантово-механических расчетов.

### 1.3.2. Силовое поле

Силы в МД рассчитываются как (отрицательный) градиент потенциальной энергии, аналитический вид функции и параметры которой задаются моделью, известной как молекулярно-механическое силовое поле. Такие приближенные модели позволяют воспроизводить ключевые характеристики квантово-

механических и экспериментальных данных. Типичное силовое поле включает в себя описание электростатических взаимодействий, предпочтительных длин ковалентных связей и других межатомных взаимодействий, отвечающих за структуру и динамику молекул. Несмотря на значительные улучшения за последние десятилетия [49], силовые поля остаются приближенными моделями, что следует учитывать при интерпретации результатов моделирования. Кроме того, в классической МД ковалентные связи остаются фиксированными в течение всего процесса моделирования, что требует предварительного задания состояний протонирования и дисульфидных мостиков [50]. Для изучения химических реакций, сопровождающихся разрывом или образованием ковалентных связей, а также фотохимических процессов, применяются гибридные методы квантовой механики/молекулярной механики (QM/MM). В таких подходах небольшая реакционноспособная часть молекулы описывается методами квантовой химии, тогда как остальная часть моделируется в рамках классической МД [51].

Наиболее широко используемыми в атомистическом моделировании силовыми полями являются различные версии AMBER [52], CHARMM [53] и OPLS [54]. Несмотря на схожесть вида базовых функций, силовые поля различаются аналитическими выражениями для потенциалов и значениями параметров, входящих в эти выражения. Такие различия обусловлены специализацией каждого силового поля для моделирования определенных классов систем (белков в растворе, твердых тел и др.), что определяет их специфические преимущества и ограничения. Например, силовое поле CHARMM36m, являющееся частным случаем поля CHARMM (CGenFF [53]), характеризуется тщательно оптимизированными и верифицированными параметрами для белков, липидов и лигандов. В свою очередь, поле A99SB-disp демонстрирует высокую точность при моделировании неупорядоченных белковых структур [55]. Поле OPLS3, вероятно, обладает наиболее оптимизированными параметрами для описания лигандов, хотя его

проприетарный статус существенно ограничивает возможности независимой оценки [54].

**Крупнозернистые** (coarse-grained) **силовые поля**, такие как MARTINI [56], предлагают альтернативный подход, при котором группы атомов объединяются в эффективные взаимодействующие частицы (крупные «зерна»). Такое описание модели обеспечивает существенное снижение вычислительной нагрузки при сохранении ключевых физико-химических характеристик системы. В частности, MARTINI нашло широкое применение при моделировании липидных бислоев (мембран), белково-лигандных комплексов и процессов самосборки макромолекул. Основное преимущество поля заключается в оптимальном соотношении между вычислительной эффективностью и физической достоверностью получаемых результатов.

В основе параметризации MARTINI, аналогично атомистическим силовым полям, лежат как экспериментальные данные, так и результаты квантово-химических расчетов. Ключевым отличием данного поля является ориентация на воспроизведение, прежде всего термодинамических характеристик системы. Параметризация основана на принципе эмпирического подбора параметров межчастичных взаимодействий с целью корректного воспроизведения свободной энергии переноса химических групп между фазами раствора. Для валидации модели используется сравнение результатов крупнозернистой МД как с экспериментальными данными, так и с результатами моделирования с помощью более точных атомистических силовых полей, таких как OPLS или GROMOS. Все частицы в MARTINI классифицируются по полярности, заряду и уровню водорастворимости. Для описания межмолекулярных взаимодействий применяются усеченные потенциалы Леннарда–Джонса и кулоновский потенциал для заряженных частиц.

Последняя версия силового поля (MARTINI 3 [57]) включает существенные усовершенствования, направленные на более точное воспроизведение плотности,

коэффициентов диффузии, геометрии агрегации и устойчивости вторичных структур белков, за счет:

- введения новых типов эффективных частиц;
- оптимизации межчастичных взаимодействий;
- учета специфических эффектов (водородные связи, электронная поляризуемость).

Модульная архитектура MARTINI обеспечивает высокую степень гибкости моделирования и позволяет создавать крупнозернистые модели широкого спектра биомолекул — белков, липидов, нуклеиновых кислот и их комплексов — с разумным балансом между точностью и вычислительной эффективностью. Благодаря этому MARTINI стал мощным инструментом для моделирования процессов, происходящих на микро- и миллисекундных временных масштабах, включая агрегацию белков, фазовые переходы в мембранах и поведение макромолекул в растворе.

Таким образом, выбор между атомистическими и крупнозернистыми подходами зависит от конкретных задач моделирования, требуемой точности и доступности вычислительных ресурсов.

### *1.3.3. Периодические граничные условия*

При проведении МД-моделирования жидкости или газа с заданными параметрами давления или плотности система помещается в непроницаемую ячейку моделирования. Для малых систем, где существенная доля атомов находится вблизи границ ячейки, особенно значимой становится проблема граничных эффектов. Классический способ минимизации краевых эффектов — применение периодических граничных условий: моделируемая система помещается в элементарную ячейку, которая окружена своими точными копиями во всех направлениях. Таким образом, границы ячейки становятся проницаемыми:

атом, пересекающий границу ячейки, появляется с ее противоположной стороны (рис. 2).

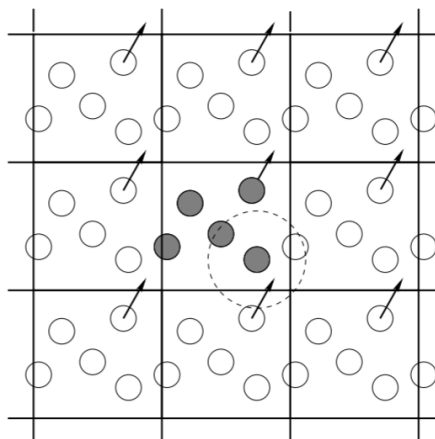


Рисунок 2 – Схематическое изображение периодических граничных условий. При пересечении частицей границы ячейки моделирования она появляется с противоположной стороны через соответствующий периодический образ.

При реализации периодических граничных условий используется принцип ближайшего образа, при котором для расчета короткодействующих сил между атомами учитывается только ближайший периодический образ каждой частицы. Тогда радиус отсечки  $R_c$  (радиус действия) этих взаимодействий не должен превышать половину длины наиболее короткого ребра ( $\mathbf{a}$ ,  $\mathbf{b}$  или  $\mathbf{c}$ ) ячейки:

$$R_c < \frac{1}{2} \min(\|\mathbf{a}\|, \|\mathbf{b}\|, \|\mathbf{c}\|).$$

Хотя такой подход устраняет проблемы, связанные с граничными эффектами, он вводит новые особенности. Для кристаллических систем периодичность естественна, но при моделировании неперiodических систем, таких как жидкости или растворы, возникают артефакты, связанные с искусственной периодичностью структуры.

Однако ошибки, вызванные периодическими граничными условиями, как правило, менее критичны, чем артефакты от границы с вакуумом, что позволяет

получить в итоге более корректные результаты по сравнению с изолированной системой.

#### *1.3.4. Способы ускорения МД-расчетов*

Одним из фундаментальных ограничений классической МД является высокая вычислительная трудоемкость метода. Для обеспечения численной стабильности и точности моделирования временной шаг интегрирования должен быть достаточно малым — как правило, порядка 1–2 фс. При этом многие представляющие интерес биомолекулярные процессы, такие как глобальные конформационные переходы белков, ассоциация молекул или связывание лигандов, характеризуются существенно более длительными временными масштабами протекания — от наносекунд до миллисекунд. Соответственно, моделирование даже относительно простого биомолекулярного процесса может потребовать от  $10^6$  до  $10^9$  итерационных шагов, на каждом из которых пересчитываются силы взаимодействия между всеми парами частиц в радиусе отсечки, что эквивалентно миллионам межатомных операций. Такие расчеты требуют значительных вычислительных ресурсов и длительного времени выполнения.

Использование графических процессоров или специализированных архитектур (например, суперкомпьютера), позволяющих параллельно обрабатывать большие массивы межатомных взаимодействий, кардинально ускоряет моделирование. Тем не менее, их применение ограничено высокой стоимостью, сложностью масштабирования и вычислительными барьерами для достижения миллисекундных временных масштабов. В качестве эффективной альтернативы существуют стратегии, основанные на моделях Маркова, которые позволяют реконструировать долговременную кинетику биомолекулярных процессов через статистический анализ множества коротких независимых траекторий [58].

Дополнительные возможности повышения эффективности расчетов предоставляют специальные методы усиленной выборки, которые позволяют системе обходить энергетические барьеры и исследовать редкие события за ограниченное время. К наиболее распространенным подходам относятся:

- «управляемая МД» (steered MD) — метод направленного воздействия на молекулу в сторону заданной конформации [59];
- метадинамика — техника, использующая добавление потенциалов смещения для предотвращения возврата системы в уже исследованные области конформационного пространства [60];
- «метод обмена репликами» (replica exchange MD) — параллельное моделирование при различных температурах с последующим обменом конфигурациями между репликами [61];
- «ускоренная МД» (accelerated MD) и «метод зонтичных выборок» (umbrella sampling) — подходы, направленные на декомпозицию энергетического ландшафта системы (уменьшение высоты энергетических барьеров) [62].

Одной из наиболее эффективных стратегий увеличения доступного временного масштаба моделирования является использование моделей, включая силовое поле MARTINI [56], при которых несколько атомов объединяются в одну крупную частицу — зерно. Это существенно снижает число степеней свободы системы, позволяя использовать увеличенные временные шаги и моделировать более длительные процессы, выходящие за рамки возможностей классического атомистического подхода [63]. При этом повышение эффективности моделирования достигается за счет снижения разрешения и потери точных деталей взаимодействий на атомарном уровне, что требует осторожной интерпретации результатов и, при необходимости, последующей валидации на основе атомистических расчетов.

#### 1.4. Применение молекулярной динамики для исследования механизмов кристаллизации белков

В последние десятилетия методы вычислительной биологии, в частности МД, стали активно применяться для исследования структуры белковых кристаллов, различных аспектов их образования и роста, а также для анализа поведения белков в предкристаллизационных растворах. МД-симуляции в области кристаллизации белков преследуют несколько ключевых целей: от фундаментального понимания механизмов нуклеации (образования первичного зародыша кристалла) и выявления метастабильных олигомерных состояний белка в растворе, до решения прикладных задач, таких как предсказание оптимальных условий кристаллизации и уточнение динамической структуры уже сформированного кристалла.

МД-моделирование также вносит значительный вклад в интерпретацию и дополнение экспериментальных данных о структуре и динамике белков непосредственно в кристаллическом состоянии. Моделирование белка в периодической ячейке, параметры которой соответствуют элементарной ячейке кристалла, позволяет воспроизвести фрагмент кристаллической структуры и детально изучать его динамические свойства [64]. Кроме того, сравнение амплитуд флуктуаций атомов, полученных из МД-траекторий, с экспериментальными температурными В-факторами, а также сопоставление смоделированной плотности воды с локализацией сайтов связывания молекул растворителя в экспериментальной структуре, позволяют валидировать используемую модель силового поля. Так, в работах [65,66] для кристалла стрептавидин-биотинового комплекса было показано, что включение в симуляцию реальных компонентов кристаллизационного раствора (солей и криопротекторов, таких как глицерол) существенно улучшает корреляцию (коэффициент Пирсона 0.88) между расчетными атомными флуктуациями и экспериментальными В-факторами по сравнению с моделированием в чистой воде. Подобные исследования подчеркивают, что белковый кристалл не является абсолютно

статичной структурой, а представляет собой динамическую систему со значительными конформационными колебаниями и высокой подвижностью молекул растворителя, причем до 70–80% экспериментально установленных позиций молекул воды могут быть успешно воспроизведены в МД-симуляциях [65–67].

Дополнительные исследования элементарной ячейки кристалла виллина продемонстрировали, что добавление высокой концентрации ионов осадителя в кристаллический растворитель предотвращает крупные трансформации асимметричных единиц и потерю симметрии ячейки без заметного влияния на гибкость белка [68]. Аналогичный эффект может быть достигнут путем нейтрализации определенных заряженных остатков, что указывает на их потенциальную роль в электростатической дестабилизации кристаллической упаковки. Эти результаты способствуют не только лучшему пониманию внутрикристаллических взаимодействий, но подчеркивают необходимость максимально точного воспроизведения условий кристаллизации при моделировании, особенно для адекватной оценки стабильности межмолекулярных взаимодействий.

Современные направления в области МД-исследований белковых кристаллов включают увеличение размеров моделируемой системы до суперячеек (нескольких элементарных ячеек), а также расширение временного диапазона симуляций до микросекунд, что позволяет достичь еще лучшего согласования с экспериментом и выявить корреляции между динамическими изменениями в кристалле и функциональными движениями белка [69].

Помимо анализа уже сформированных кристаллов, МД применяется для предсказания оптимальных условий кристаллизации. Одним из таких инструментов является МД-расчет второго осмотического вириального коэффициента ( $B_{22}$ ), отражающего усредненный характер взаимодействий между парами молекул белка в растворе. Значение  $B_{22}$ , вычисляемое с помощью МД-

симуляций на основе анализа потенциала средней силы или радиальных функций распределения, коррелирует с растворимостью белка и, согласно эмпирическому правилу Джорджа–Уилсона, может служить предиктором диапазона условий (например, концентрации осадителя), в которых кристаллизация наиболее вероятна [70].

Тем не менее, прямое моделирование нуклеации белковых кристаллов из пересыщенного раствора остается исключительно трудоемкой задачей. Это обусловлено как стохастической природой процесса, так и тем, что он часто выходит за пределы достижимого временного диапазона классической МД (обычно до нескольких микросекунд). Поэтому при изучении механизмов нуклеации применяются методы ускоренной выборки, включая метадинамику, «сидинговое» моделирование (введение затравки), моделирование при постоянном химическом потенциале и другие техники, позволяющие системе преодолевать энергетические барьеры и исследовать редкие события [71].

Например, методы «зонтичной выборки» (*umbrella sampling*, [72]) и «направленного потока» (*forward flux sampling*, [73]) позволили исследовать механизм формирования первичных межбелковых контактов при кристаллизации и изучить диссоциацию уже сформированных кристаллов. Как продемонстрировано в работе [74], процесс кристаллизации инициируется слабыми, но дальнедействующими потенциалами: электростатическими взаимодействиями (на расстоянии до 10 нм) и неспецифическими поверхностными контактами между полярными группами. Эти обратимые взаимодействия обеспечивают первичную ориентацию молекул, тогда как последующая термодинамическая стабилизация кристаллической решетки достигается за счет образования специфических гидрофобных контактов, возникающих на финальных стадиях упаковки. Установлено также, что для кластеров с высокой площадью межмолекулярного взаимодействия наблюдаются многоступенчатые механизмы диссоциации.

Для моделирования процессов самосборки на еще больших пространственных и временных масштабах, недоступных для прямых атомистических симуляций, успешно применяются крупнозернистые силовые поля. Они доказали свою применимость для описания ассоциации белков, обеспечивая результаты, согласующиеся с данными атомистического моделирования. Например, симуляции кристаллических структур инсулина выявили различия в жесткости разных полиморфов и влияние межмолекулярной упаковки на локальную подвижность цепей [75]. Кроме того, моделирование помогло обнаружить причины рН-зависимых конформационных переходов, а также изучить динамику и стабильность кристаллических форм инсулина, что важно для разработки новых форм лекарств.

В дополнение к МД, для описания макроскопических аспектов кристаллизации и фазового поведения белковых растворов применяются и другие теоретические подходы, например, фазово-полевые модели. Так, в работе [76] использовались фазово-полевые уравнения для описания двухфазного поведения раствора белков, что продемонстрировало трехстадийный механизм кристаллизации: гомогенное формирование капель насыщенной белковой фазы, гетерогенная нуклеация кристаллов на их поверхности и возможное последующее растворение части частиц. Примечательно, что выводы о двухстадийной нуклеации, при которой сначала образуются локальные высокоплотные аморфные скопления белка, которые затем служат центрами для формирования упорядоченных кристаллитов, находят подтверждение и в МД-исследовании [76]. Такая картина хорошо согласуется с современными экспериментальными и теоретическими концепциями неклассической нуклеации белков.

Анализ литературы последних лет показал, что методы МД предоставляют ценную и зачастую уникальную информацию как о механизмах кристаллизации белков, так и об их внутрикристаллической динамике на атомном уровне.

#### *1.4.1. МД-исследования кластеров-прекурсоров кристалла лизоцима: идентификация и роль ионов*

Ключевым шагом к пониманию процессов, происходящих на ранних стадиях кристаллизации белков, стали результаты цикла работ, посвященных изучению структуры и динамики кластеров-прекурсоров кристаллов белков с помощью комплексного применения методов малоуглового рентгеновского рассеяния (МУРР) и МД. Модельной системой в этих исследованиях является лизоцим куриного яйца куриного яйца (HEWL) — один из наиболее изученных объектов в этом направлении, благодаря обширной экспериментальной базе данных, включая МУРР и рентгеноструктурный анализ.

Согласно данным МУРР [4,61] перед началом кристаллизации лизоцима (в тетрагональной сингонии) в его растворе образуются димеры и октамеры белка, тогда как промежуточные олигомеры (например, тетрамеры и гексамеры) и кластеры большего размера отсутствуют. Анализ кристаллической упаковки тетрагонального лизоцима (пр. гр.  $R4_32_12$ ) выявил два структурно-различных типа октамеров, обозначенных как А и В, с молекулярными объемами  $117 \text{ нм}^3$  и  $140 \text{ нм}^3$  соответственно (рис. 3). Однако точность измерений МУРР не позволила однозначно определить пространственную структуру октамера, формирующегося в кристаллизационном растворе лизоцима.

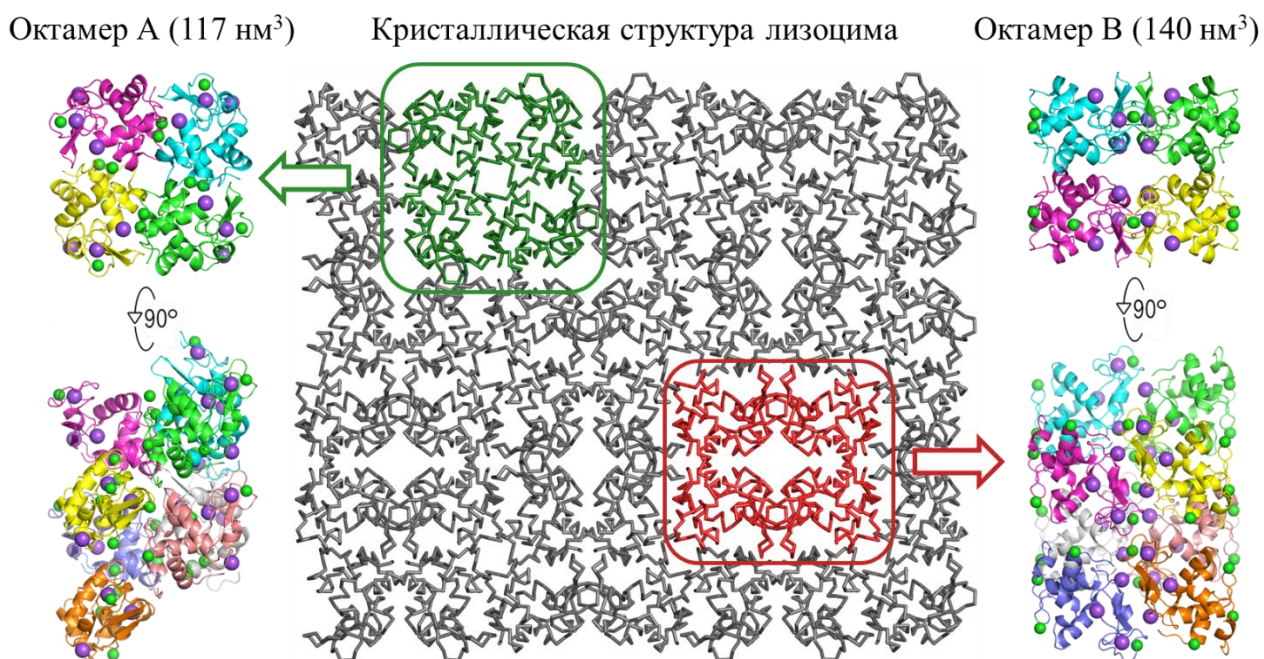


Рисунок 3 – Проекция кристаллической структуры лизоцима тетрагональной сингонии и два типа октамеров (А и В) – возможных единиц роста кристалла.

МД-моделирование, проведенное в работе [11], позволило устранить неопределенность в структуре кластера-прекурсора путем сравнения стабильности двух типов октамеров, определяемой сохранением межмолекулярных контактов, характерных для кристаллической упаковки. В качестве количественного критерия использовалась площадь межмолекулярных интерфейсов, рассчитанная с помощью программы PISA [78] на основе траекторий МД (рис. 4). Полученные результаты показали, что в условиях, соответствующих кристаллизационным (в присутствии осадителя NaCl), октамер типа А демонстрирует устойчивость и сохраняет кристаллические контакты, тогда как октамер типа В подвергается быстрому распаду на олигомеры меньшего порядка (рис. 4). Это позволило сделать вывод о том, что именно октамер типа А с наибольшей вероятностью представляет собой стабильный прекурсор, образующийся в кристаллизационном растворе лизоцима. Дополнительно было обнаружено, что в отсутствие ионов NaCl оба типа октамеров теряют

структурную целостность и диссоциируют, что подчеркивает ключевую роль ионов осадителя в стабилизации предкристаллизационных кластеров.

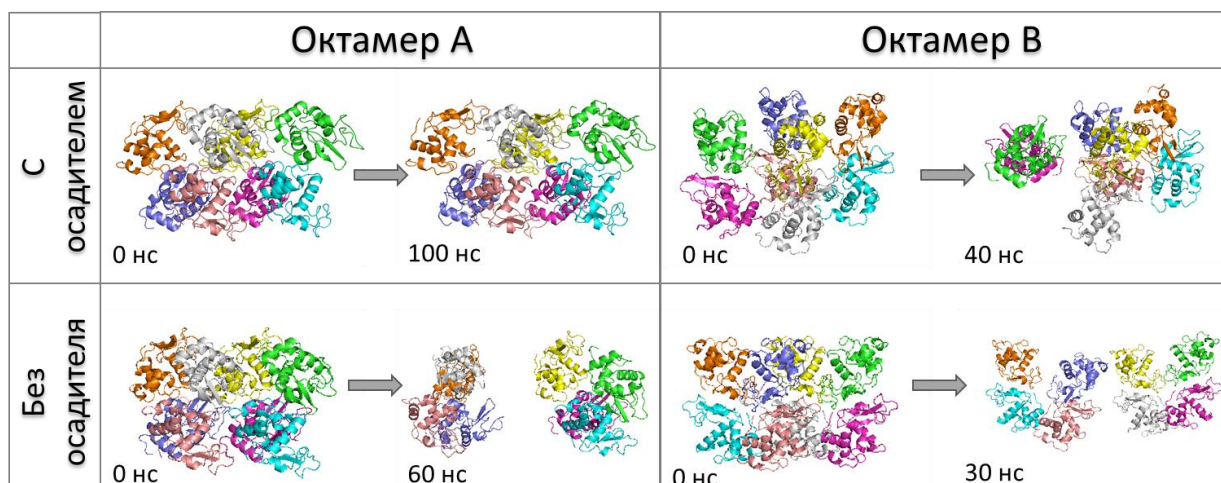


Рисунок 4 – Изменение структуры октамеров А и В в ходе их молекулярной МД в присутствии осадителя NaCl (сверху) и в его отсутствие (снизу) в растворе.

Аналогичным образом была смоделирована МД димеров, тетрамеров и гексамеров, соответствующих фрагментам октамеров типов А и В, в водном растворе как с добавлением NaCl, так и без него. В результате было обнаружено, что тетрамеры и гексамеры обоих типов (А и В) являются нестабильными независимо от наличия осадителя в растворе: они быстро диссоциируют на мономеры, что полностью соответствует экспериментальным данным МУРР, свидетельствующим об отсутствии этих олигомерных форм в предкристаллизационных растворах лизоцима. В частности, нестабильность тетрамеров может быть обусловлена слабостью специфических межмолекулярных контактов, например, наличием единственной водородной связи (между аргинином и треонином), легко разрывающейся в ходе динамики.

Поведение димеров оказалось более сложным: димер типа А сохранял структурную стабильность только в присутствии ионов NaCl, тогда как в воде без осадителя быстро диссоциировал. В отличие от него, димер типа В устойчив как в условиях с добавлением NaCl, так и в чистой воде. Тем не менее, несмотря на стабильность димера В *in silico*, вероятность его образования в реальном кристаллизационном растворе ограничена вследствие необходимости

преодоления высокого энергетического барьера для формирования специфических межмолекулярных контактов, характерных для его структуры.

#### *1.4.2. Влияние ионов осадителя на стабильность олигомеров лизоцима при различных температурах*

Разработанный в [11] методологический подход к изучению поведения белковых олигомеров в кристаллизационных растворах был далее расширен в исследовании [31], где особое внимание было уделено температурной зависимости стабильности мономеров и октамеров лизоцима. В рамках данного исследования анализировались исключительно кластеры типа А, поскольку именно они, как было показано в предыдущем разделе, с наибольшей вероятностью формируются в растворе до начала кристаллизации лизоцима [11].

В качестве исходных моделей использовались фрагменты двух экспериментально определенных кристаллических структур лизоцима (тетрагональной сингонии с PDB ID: 4WLD и 6QWY), различающихся по составу и распределению ионов осадителя NaCl, встроенных в кристаллическую решетку. Структура молекулы лизоцима 4WLD содержит лишь один ион натрия, в то время как в структуре 6QWY присутствуют три иона Na<sup>+</sup> и четыре иона Cl<sup>-</sup>, координированные к белку. Различия в составе ионного окружения позволили исследовать влияние ионной среды на стабильность межмолекулярных взаимодействий как в олигомерах, так и в отдельных мономерах, при различных температурах моделирования.

Для систематической оценки вклада ионной среды в стабилизацию межмолекулярных взаимодействий в олигомерных ассоциатах в работе [31] методом МД были смоделированы мономеры и октамеры лизоцима с различным количеством встроенных в структуру ионов осадителя. Было рассмотрено три варианта: модель полностью без ионов (на основе структуры 4WLD), модель с единственным ионом натрия, координированным к подвижной петле 60–73 (также на основе 4WLD), и модель, содержащая полный набор

кристаллографически определенных ионов (три иона  $\text{Na}^+$  и четыре иона  $\text{Cl}^-$  на мономер) из структуры 6QWY.

Моделирование происходило в диапазоне температур от 5 до 40°C, с целью воспроизвести условия, в которых проводился эксперимент МУРР. При этом особое внимание уделялось области низких температур, поскольку, согласно результатам МУРР, наибольшая концентрация димеров и октамеров лизоцима в кристаллизационных растворах наблюдается при температуре около 10°C [79].

Для количественной оценки стабильности мономеров и октамеров лизоцима по результатам МД-симуляций были построены профили среднеквадратичных флуктуаций (RMSF)  $\text{C}_\alpha$ -атомов. Согласно данным на рис. 5а, в растворе, содержащем осадитель, мономер лизоцима без кристаллографически связанных ионов демонстрирует наибольшую структурную стабильность, однако различия между тремя моделями в целом незначительны. Влияние координированных ионов на структурную устойчивость молекулы лизоцима в водном растворе без ионов  $\text{Na}^+$  и  $\text{Cl}^-$  (рис. 5б) еще меньше.

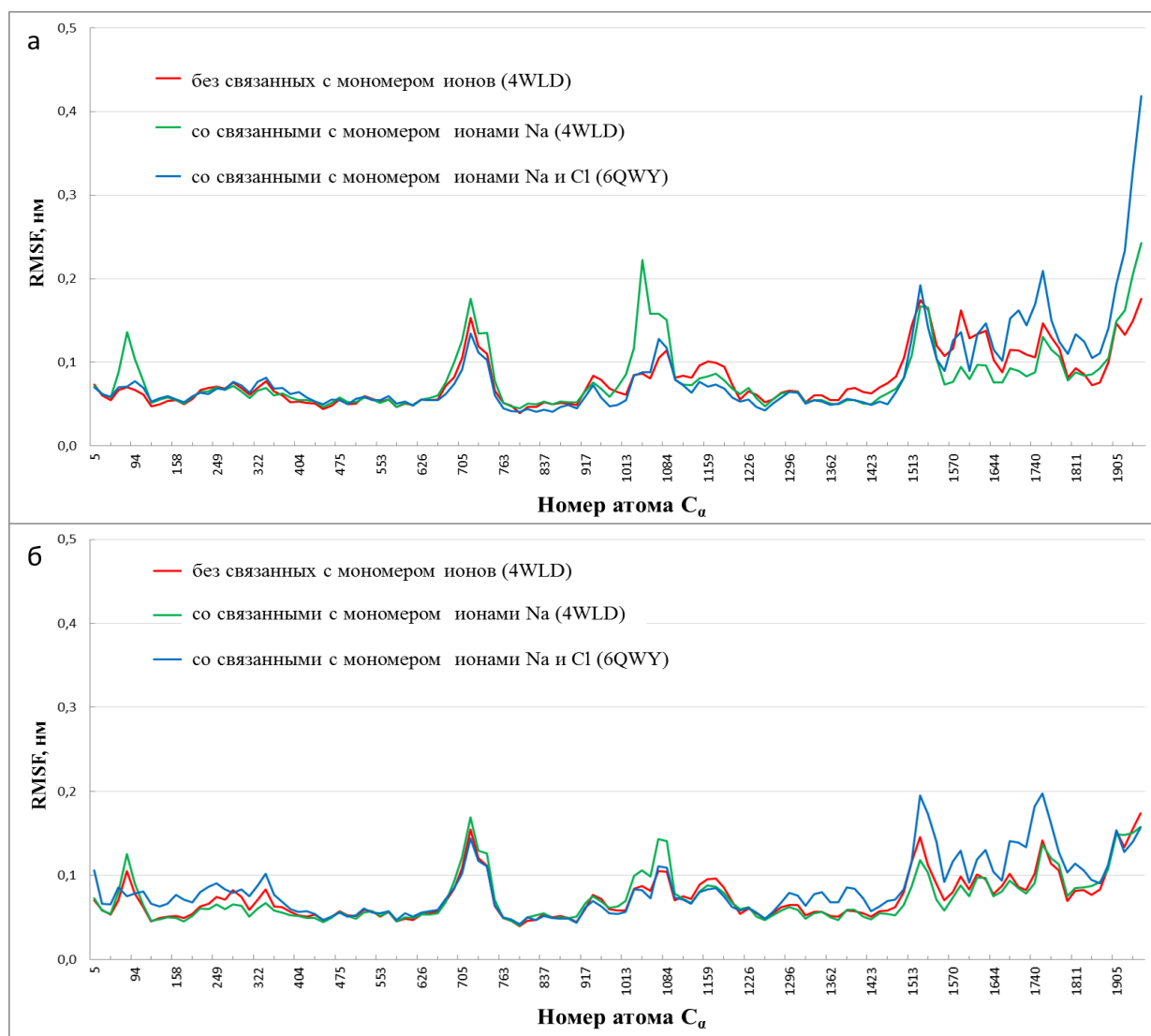


Рисунок 5 – RMSF  $C_{\alpha}$ -атомов мономера лизоцима, смоделированного с осадителем (*a*) и без (*б*) в растворе. Молекулам со встроенными ионами  $Na^{+}$  и  $Cl^{-}$  соответствуют кривые синего цвета; молекулам только с ионами  $Na^{+}$  – зеленого; молекулам без связанных ионов – красного.

Тем не менее, сравнение стабильности моделей в средах, различающихся ионным составом (раствор с ионами осадителя или водный раствор), указывает на минимальные отличия в локальной подвижности отдельных участков молекулы лизоцима, в частности в области  $C_{\alpha}$ -атомов с номерами 1850–1950. Эти данные указывают, что главная роль ионов, вероятно, заключается не в прямой стабилизации структуры мономера, а в модификации поверхностного распределения зарядов, создающего условия, благоприятные для кристаллической

упаковки. Дополнительным подтверждением этой гипотезы служат экспериментальные данные по формированию положительно заряженной поверхности пленки лизоцима в Ленгмюровской после внесения кристаллизационного раствора с осадителем [80]. На границе пленки локализуются анионы  $\text{Cl}^-$ , что свидетельствует о значимости ионного окружения в организации структуры кластеров-прекурсоров.

На рис. 6 представлены результаты МД-моделирования в виде графиков среднеквадратичных флуктуаций (RMSF)  $\text{C}_\alpha$ -атомов октамеров лизоцима в кристаллизационном растворе при различных температурах и с разными наборами ионов осадителя, встроенных в структуру белка. Сравнительный анализ графиков показывает, что увеличение температуры приводит к монотонному росту RMSF, т. е. снижению стабильности, только в случае, когда в моделировании учитываются ионы как  $\text{Na}^+$ , так и  $\text{Cl}^-$  (рис. 6в). Это поведение соответствует экспериментальным данным МУРР [81], согласно которым с повышением температуры концентрация октамеров в растворе уменьшается. Таким образом, полученные результаты указывают на то, что согласование МД-моделирования с экспериментом достигается только при учете полного набора кристаллографически определенных ионов осадителя в модели октамера.

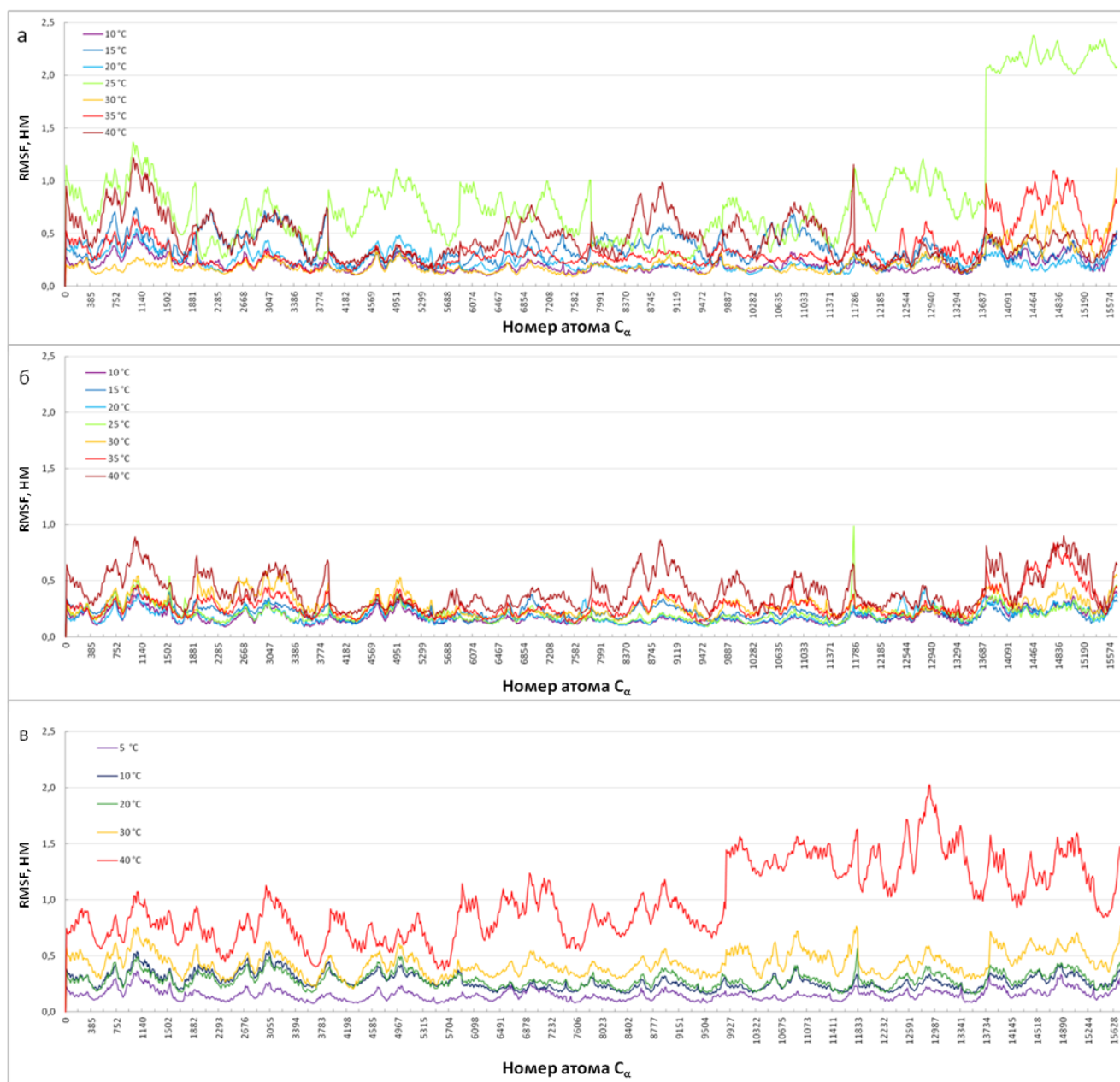


Рисунок 6 – RMSF  $C_{\alpha}$ -атомов октамера лизоцима, смоделированного в кристаллизационном растворе при различных температурах. Модели различаются составом связанных с белком ионов осадителя: (а) без ионов; (б) только с ионами  $Na^{+}$ ; (в) с ионами  $Na^{+}$  и  $Cl^{-}$ .

Моделирование показало, что в случае мономера наличие связанных ионов оказывает лишь незначительный дестабилизирующий эффект, тогда как для октамеров температурная устойчивость структуры воспроизводится корректно только при учете всех ионов, координированных с белком в кристалле. Полученные данные согласуются с результатами МУРР и подчеркивают, что

кристаллизацию белка определяет не только состав раствора, но и участие ионов, встроенных в структуру белка. Эти ионы, вероятно, способствуют перераспределению поверхностных зарядов и стабилизации межмолекулярных взаимодействий, критичных для образования прекурсоров кристаллической решетки.

### 1.5. Заключение

Пространственная организация белка играет ключевую роль в реализации его биологических функций, определяя специфичность его взаимодействий с лигандами, кофакторами и другими молекулами [12]. Установление трехмерной структуры белков остается одной из фундаментальных задач структурной биологии, необходимой как для базового понимания молекулярных механизмов, так и для прикладных целей, включая рациональный дизайн лекарств. Хотя экспериментальные методы, такие как РСА, крио-ЭМ и ЯМР, предоставляют высокоточную структурную информацию, их применение сопряжено со значительными техническими и временными затратами, а также ограничено в отношении отдельных классов белков [3].

Это стимулировало активное развитие вычислительных подходов к предсказанию трехмерной структуры белка на основе его аминокислотной последовательности, таких как гомологичное моделирование и методы, основанные на физических принципах, включая МД [13]. Однако точность этих методов снижается при отсутствии близких структурных аналогов. Современные алгоритмы машинного обучения показывают высокую точность предсказаний стабильных глобулярных доменов. Тем не менее, они имеют ограничения, связанные с невозможностью учета условий среды, присутствия лигандов и конформационной подвижности молекул. Поэтому, несмотря на свою эффективность, они представляют собой важное дополнение к экспериментальным методам, но не могут полностью их заменить [4].

На сегодняшний день РСА является основным и наиболее точным методом определения пространственной структуры белка, однако его применение требует получения хорошо упорядоченных белковых кристаллов, выращивание которых остается наиболее трудоемким и непредсказуемым этапом всего исследования. Как правило, кристаллизация осуществляется путем эмпирического перебора параметров пересыщенного раствора: состава осадителя, концентраций компонентов, рН, температуры и других факторов. Рационализация этого процесса ограничена отсутствием фундаментального понимания механизмов нуклеации и роста белкового кристалла.

Современные концепции белковой кристаллизации все чаще рассматривают ее как двухступенчатый процесс, включающий образование промежуточных олигомерных кластеров-прекурсоров, служащих ядрами кристаллической фазы [8,10,81]. Дальнейшее развитие теории требует углубленного изучения природы и поведения обнаруженных предкристаллизационных агрегатов. В этом контексте метод МД зарекомендовал себя как мощный и гибкий инструмент, эффективно дополняющий экспериментальные подходы на ранних этапах кристаллизации.

МД-моделирование позволяет с атомарной точностью исследовать динамику межмолекулярных взаимодействий, идентифицировать устойчивые формы комплексов и количественно оценивать влияние различных факторов на структурную стабильность белков. Особое преимущество метода заключается в возможности изучения условий, которые затруднены или недоступны для непосредственного экспериментального наблюдения.

В частности, для кристаллизационных растворов лизоцима МД позволила идентифицировать наиболее вероятную структуру прекурсорного октамера, подтвердить нестабильность тетрамеров и гексамеров, а также продемонстрировать важную роль учета координированных ионов осадителя при моделировании предкристаллизационных олигомеров.

Таким образом, МД-моделирование служит не только методом интерпретации экспериментальных наблюдений, но и надежным инструментом теоретического анализа, способным направлять и ускорять исследования в

области белковой кристаллизации. Использование МД способствует переходу от эмпирического подбора условий к теоретически обоснованному подходу получения белковых кристаллов.

## ГЛАВА 2. АДАПТАЦИЯ МЕТОДОВ МОЛЕКУЛЯРНОЙ ДИНАМИКИ К ИССЛЕДОВАНИЯМ КРИСТАЛЛИЗАЦИОННЫХ РАСТВОРОВ БЕЛКОВ

МД-моделирование процесса белковой кристаллизации — от свободных мономеров в растворе до формирования трехмерной кристаллической решетки — представляет собой крайне ресурсоемкую задачу, требующую экстремально длительных симуляций и масштабных вычислительных мощностей. Это обусловлено как высокой степенью свободы биомолекул, так и длительностью процессов кристаллизации, зачастую выходящих за пределы временных масштабов, достижимых классической МД. В этой связи особенно актуальными становятся подходы, направленные на моделирование отдельных ключевых стадий кристаллизации, имеющих решающее значение для последующего формирования кристалла.

В цикле экспериментальных и теоретических работ [8,10,81] было показано, что до начала роста кристалла в белковых растворах формируются стабильные олигомерные структуры — так называемые кластеры-прекурсоры. Данные агрегаты являются важным элементом двухступенчатой модели нуклеации и, как предполагается, выступают в роли зародышей кристаллической решетки. Это открывает возможность сосредоточиться не на моделировании всего процесса кристаллизации, а на изучении стабильности и поведения предкристаллизационных олигомеров, уже известных из кристаллических структур.

Методологический подход, применяемый в настоящей работе, основывается на использовании данных о кристаллической упаковке: из кристаллической структуры целевого белка выделяется фрагмент (олигомер), предположительно соответствующий прекурзору, после чего он помещается в модельный раствор, воспроизводящий условия кристаллизационного эксперимента (включая состав осадителя, ионную силу, рН и температуру). Далее проводится МД-моделирование полученной системы, анализируется поведение олигомера и

оценивается его стабильность. Такой подход позволяет выявлять молекулярные механизмы, лежащие в основе ассоциации белков в растворе, и оценивать роль различных межмолекулярных взаимодействий во время формирования кристалла.

Таким образом, методика направлена на создание вычислительной модели, воспроизводящей предкристаллизационную фазу белкового раствора и может служить основой для рационального поиска условий кристаллизации, а также углубленного анализа механизмов формирования белковых кристаллов.

Кроме того, в рамках диссертационной работы рассматриваются способы повышения вычислительной эффективности (ускорения) МД-симуляций за счет оптимизации размеров ячейки моделирования и применения крупнозернистых моделей, в частности силового поля MARTINI. Использование минимально допустимых объемов симуляционной ячейки позволяет существенно сократить число частиц в системе при сохранении корректного воспроизведения ключевых межмолекулярных взаимодействий. Применение крупнозернистого подхода, в свою очередь, позволяет моделировать процессы на более продолжительных временных интервалах и в системах с существенно большим числом молекул, чем это возможно при использовании атомистического описания. Таким образом, расширяются возможности численного изучения процессов ассоциации и кристаллизации белков, в том числе в условиях, труднодоступных для экспериментального наблюдения.

## 2.1. Проведение вычислительного эксперимента

### 2.1.1. Построение моделей

В настоящей работе модели олигомеров белков (лизозима куриного яичного белка – HEWL, протеиназы К и термолизина) выделялись из соответствующих кристаллических структур, установленных методом РСА и депонированных в Protein Data [2].

Для каждого исследуемого белка с помощью программы РuMOL (версия 1.5 [79]) был воссоздан фрагмент его кристаллической структуры путем применения операторов симметрии к координатам белковой молекулы для ее распространения в соответствии с пространственной группой (**пр. гр.**) кристаллов. Параметры белковых кристаллов и растворов, из которых они были выращены, представлены в табл. 1. Асимметричная единица элементарной ячейки данных кристаллов состоит из одной молекулы белка. Кристаллы лизоцима и протеиназы К относятся к тетрагональной сингонии, а термолизина – к гексагональной.

Таблица 1. Характеристики кристаллов исследуемых белков (идентификатор PDB, пространственная группа и количество ионов, встроенных в молекулу белка в кристалле) и растворов, из которых были выращены соответствующие кристаллы (рН, тип и концентрация осадителя).

<b>Белок</b>	Лизоцим	Протеиназа К	Термолизин
<b>PDB ID</b>	6QWY	7A68	3DNZ
<b>Пр. гр.</b>	P4 <sub>3</sub> 2 <sub>1</sub> 2	P4 <sub>3</sub> 2 <sub>1</sub> 2	P6 <sub>1</sub> 22
<b>рН</b>	4.5	8.0	6.0
<b>Осадитель</b>	NaCl	NaNO <sub>3</sub>	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>
<b>Концентрация осадителя, М</b>	0.4	0.5	0.75
<b>Количество ионов, связанных с белком в кристалле*</b>	3 Na <sup>+</sup> + 4 Cl <sup>-</sup>	2 Na <sup>+</sup>	4 Ca <sup>2+</sup> + 1 Zn <sup>2+</sup>

\* Число ионов осадителя (в случае Na<sup>+</sup> и Cl<sup>-</sup>) или лигандов (в случае Ca<sup>2+</sup> и Zn<sup>2+</sup>), приходящихся на одну молекулу белка.

Из фрагментов кристаллической структуры протеиназы К и термолизина выделялись все возможные типы димеров и гексамеров соответственно и записывались в отдельный pdb-файл. Для лизоцима использовалась построенная таким же способом модель октамера – наиболее вероятного кластера, образующегося в его кристаллизационном растворе, и выделенного из него димера [11]. Модели гексамера лизоцима получены путем удаления димера из

структуры октамера. Кристаллическая вода удалялась из pdb-файлов олигомеров, а ионы осадителя или лиганды, связанные с белком в кристалле были оставлены (табл. 1).

Все остатки, которые были обнаружены в кристаллах с множественной заселенностью, имеют одинаковую вероятность нахождения в одной из двух конформаций (А или В), так как их коэффициент заселенности составляет 0.5. Поскольку вероятность нахождения в одной из двух конформаций (А или В) одинаковая, то была сохранена только конформация А.

### 2.1.2. МД-моделирование в атомистическом силовом поле

Все расчеты и подготовка структур к ним были выполнены в программном пакете GROMACS (версий 5.0.4, 2021 и 2023 [41]) и проводились одинаково для всех моделей олигомеров, если не указано иное. Состояния протонирования аминокислотных остатков при рН кристаллизационного буфера (табл. 1) были заданы с помощью сервера PROPKA (версий 2.0.0 и 3.4.1 [82]). В качестве силового поля использовалось Amber ff99SB-ILDN [83], так как набор усовершенствованных параметров торсионных углов боковых цепей некоторых аминокислот в этой версии приводит к увеличению точности результатов моделирования по сравнению с более ранними версиями поля Amber [43]. При уравнивании систем и вычислении продуктивной динамики применялись периодические граничные условия вдоль трех измерений, а дальнедействующие электростатические взаимодействия учитывались методом «частица-сетка» Эвальда (PME, [84]) с кубической интерполяцией и шагом сетки в обратном пространстве 0.16 нм. Нековалентные взаимодействия учитывались только для атомов, находящихся на расстоянии не более 1 нм друг от друга. Длины связей в олигомерах поддерживались постоянными с помощью алгоритма LINCS [85].

Подготовленные структуры олигомеров помещались в центр кубической ячейки моделирования, при этом минимальное расстояние между ее краем и молекулой протеиназы К и термолизина составляло 1 нм. При проверке влияния

размеров ячейки на стабильность олигомеров это расстояние составляло 1, 1.5, 2, 2.5 и 3 нм.

Остальное пространство ячейки заполнялось четырехчастичной (состоящей из одного атома кислорода, двух атомов водорода и одного мнимого атома, имитирующего неподеленную электронную пару атома кислорода) моделью воды TIP4P-Ew (Transferable Intermolecular Potential 4 Point - Ewald [86]), разработанной специально для выполнения МД расчетов методом суммирования зарядов по Эвальду (PME [84]). Для воспроизведения кристаллизационного раствора, содержащего осадитель, молекулы воды заменялись его ионами ( $\text{Na}^+$ ,  $\text{Cl}^+$ ,  $\text{NO}_3^-$ ,  $\text{NH}_4^+$ ,  $\text{SO}_4^{2-}$ ) таким образом, чтобы концентрация осадителя в ячейке соответствовала условиям кристаллизационного эксперимента. Поскольку силовое поле Amber ff99SB-ILDN не содержит параметров для ионов  $\text{NO}_3^-$ ,  $\text{NH}_4^+$  и  $\text{SO}_4^{2-}$ , трехмерные структуры нитрита и аммония были получены из PDBChem [87] (код: NO3 и NH4 соответственно), а структура сульфата – с помощью модуля PLMD (Peptide Ligand Molecular Dynamics) из пакета MDAnalysis [88,89]. Топологии данных ионов были сгенерированы с применением пакета Antechamber [90]. Использование алгоритма суммирования по Эвальду (PME) для расчета дальних электростатических взаимодействий формально требует нейтральности суммарного заряда молекулярной системы, поэтому в каждую ячейку дополнительно добавлялось незначительное количество ионов  $\text{Na}^+$  и  $\text{Cl}^-$ .

Перед вычислением продуктивной МД моделируемые системы подвергались минимизации энергии методом наискорейшего спуска (50 000 шагов) до тех пор, пока максимальная сила, действующая на каждый атом, становилась не более  $1000 \text{ кДж}\cdot\text{М}^{-1}\cdot\text{нм}^{-1}$ . Данный шаг необходим для избегания вычислительных ошибок, к которым могут приводить слишком большие силы, возникающие из-за имеющихся в исходной кристаллической структуре и появившихся в процессе добавления воды атомов «наложенных» друг на друга или расположенных слишком близко. Для задания нужной температуры и плотности молекул системы последовательно подвергались уравниванию в *NVT*- и *NPT*-ансамблях (по

100 пс) с использованием модифицированного термостата Берендсена с масштабируемыми скоростями (V-rescale) [91] и баростата Паринелло-Рамана [92] соответственно. Расчеты продуктивной МД производились в *NPT*-ансамбле также с помощью модифицированного термостата Берендсена и баростата Паринелло-Рамана при давлении, составляющим 1 атм. Список взаимодействующих атомов, находящихся на расстоянии не более 4 ковалентных связей друг от друга, обновлялся каждые 10 шагов. Интегрирование происходило по стандартному алгоритму leap-frog [93], шаг интегрирования составлял 2 фс.

### 2.1.3. Анализ результатов

Стабильность олигомеров оценивались с помощью графиков RMSF (Root Mean Square Fluctuations) атомов  $C_\alpha$ . Значения RMSF служат мерой гибкости полипептидной цепи, поскольку они показывают степень отклонения атомов  $C_\alpha$  от их среднего положения. Чем выше значение RMSF атома, тем он подвижнее.

Изменение структуры олигомеров в ходе МД анализировалось на основе графиков RMSD (Root Mean Square Deviation), то есть среднеквадратичного отклонения всех атомов  $C_\alpha$  от исходной (в данном случае кристаллической) структуры. Разница между RMSF и RMSD состоит в том, что RMSF отражает колебания каждого атома относительно его среднего положения в течение всего времени моделирования, а RMSD – отклонение структуры (всех атомов  $C_\alpha$ ) от исходной в зависимости от времени.

Размеры белка оценивались с помощью радиуса инерции  $R_g$ , характеризующем его компактность. Этот параметр рассчитывается как среднеквадратичное расстояние всех атомов в белке от центра масс молекулы.

Перед анализом траекторий артефакты, возникающие из-за квазибесконечных периодических граничных условий, устранялись с помощью команды *gmx trjconv* с флагом *-pbc nojump*, восстанавливающей целостность олигомеров. После чего при выполнении команд *gmx rmsf*, *gmx rms* и *gmx gyrate* для всех атомов  $C_\alpha$

вычислялись значения RMSF, RMSD и  $R_g$ , соответственно. Структурное выравнивание траекторий атомов лизоцима к исходному положению при вычислении RMSD производилось применением команды *gmx trjconv* с флагом *-fit rot+trans*, т. к. движение белка как твердого тела (параллельные переносы и вращения) не влияет на его конформационные изменения.

## 2.2. Определение оптимальных размеров ячейки моделирования

### 2.2.1. Влияние размеров ячейки моделирования и концентрации ионов осадителя на стабильность димера лизоцима

В программном пакете GROMACS [41], используемом для моделирования в данной работе, применяются периодические граничные условия, чтобы избежать реальных фазовых границ. Таким образом, при исследовании жидкой фазы в систему вносится неестественная для неё периодичность. Влияние этой искусственной периодичности тем сильнее, чем меньше моделируемая система, но для большого числа атомов ( $>10^4$ – $10^5$ ) артефакты, как правило, незначительны.

Однако в работе [94] авторы установили, что стабильность гемоглобина в  $T_0$ -состоянии (без лиганда) критически зависит от размеров ячейки моделирования. Известно, что тетрамеры гемоглобина стабильны в растворе как с лигандом ( $R_4$ ), так и без него ( $T_0$ ), но во время МД-симуляций гемоглобина его  $T_0$ -состояние устойчиво только тогда, когда ячейка моделирования содержит в десять раз больше молекул воды, чем при её стандартных размерах для таких симуляций (рис. 7). В остальных же случаях  $T_0$ -структура гемоглобина быстро переходила к  $R$ -подобному строению. Полученные результаты позволяют предположить, что такая большая ячейка необходима для проявления гидрофобного эффекта, стабилизирующего  $T_0$ -тетрамер.

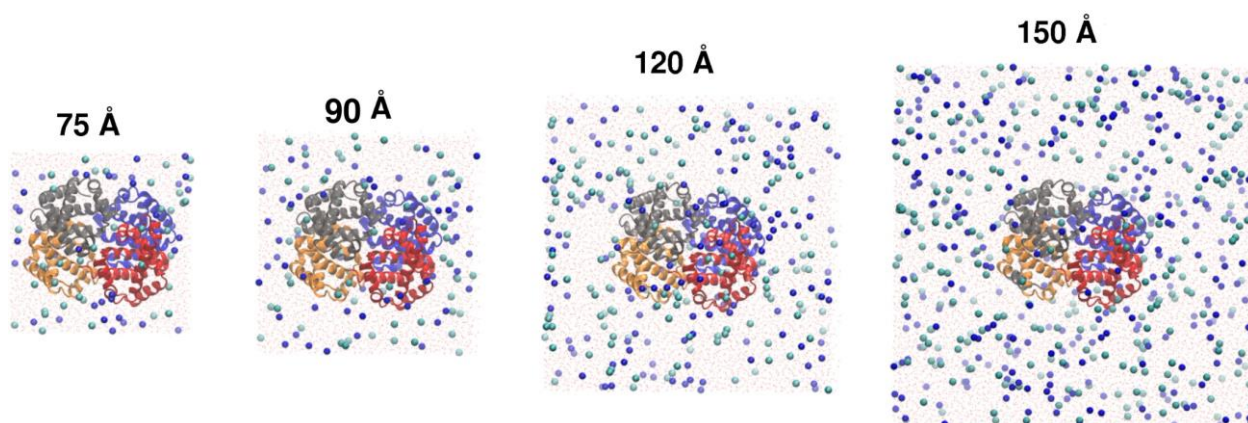


Рисунок 7 – Тетрамеры гемоглобина в ячейках моделирования разного размера. Сверху указаны длины ребер соответствующих ячеек. Сферами представлены ионы  $\text{Na}^+$  и  $\text{Cl}^-$ , добавленные для нейтрализации системы и достижения биологически приемлемой концентрации соли 0.15 М. Тетрамер стабилен только в самой большой ячейке с длиной ребра 150 Å.

Вследствие данного наблюдения в имеющей практическое значение работе [95] изучалось влияние размеров ячейки моделирования на стабильность димера лизоцима, являющегося фрагментом структуры тетрагонального кристалла, для получения в дальнейшем достоверных результатов при МД-исследованиях взаимодействия белковых олигомеров с растворами. Симуляции проводились в двух ячейках моделирования разного объема и при трех концентрациях осадителя в растворе.

Размеры кубических ячеек моделирования задавали таким образом, чтобы расстояние между любым атомом белка и краем ячейки в одном эксперименте было не менее 1 нм, в другом – не менее 2.5 нм, длина ребра ячейки составила 8.6 и 11.6 нм соответственно. Для воспроизведения кристаллизационных условий димеры лизоцима моделировались в водном растворе с осадителем. Часть молекул воды заменялась ионами натрия и хлора таким образом, чтобы концентрация соли составляла 0.4 М (25 мг/мл) для ячейки с ребром 11.6 нм и 0.2, 0.4, 0.6 М для ячейки с ребром 8.6 нм (табл. 2).

Таблица 2. Отличия параметров двух наборов проведенных моделирований.

	Длина грани ячейки, нм	Концентрация осадителя NaCl, М
1-й набор	8.6	0.4
	11.6	
2-й набор	8.6	0.2
		0.4
		0.6

По результатам моделирования МД были построены графики RMSF атомов  $C_{\alpha}$ . На рис. 8 приведены RMSF  $C_{\alpha}$ -атомов димера лизоцима в растворе при концентрации осадителя 0.4 М, смоделированных в ячейках разного размера (с ребром, составляющим 1.0 и 2.5 нм). Из рис. 8 следует, что на относительно короткой траектории в 10 нс размер ячейки моделирования не оказывает существенного влияния на подвижность атомов белка. Тем не менее, следует отметить некоторую разницу в подвижностях атомов на определённых участках белка. Наибольшая разница (превышающая 0.028 нм) в значениях RMSF между димерами в разных ячейках наблюдается у аминокислотных остатков: SER81, ALA82, GLY102, TRP111, ARG112, ASN113, GLY117, GLY126 в полипептидной цепи А; и остатков HIP15 (HIS в протонированном состоянии), PRO70, GLY71, THR89, ALA90, ARG114, LYS116, GLY117, THR118, ASP119, ALA122 в цепи В.

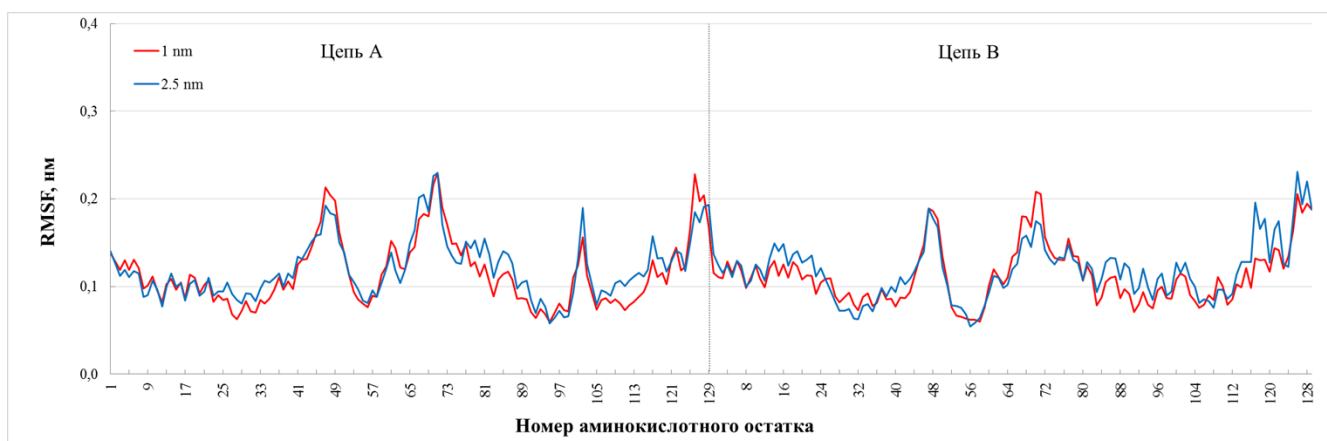


Рисунок 8 – RMSF  $C_{\alpha}$ -атомов димера лизоцима в растворе с концентрацией осадителя 0.4 М при расстоянии между краем ячейки моделирования и ближайшим к нему атомом белка 1 нм (красная кривая) и 2.5 нм (синяя).

На рис. 9 представлены значения RMSF для  $C_{\alpha}$ -атомов димера лизоцима в растворе при концентрациях осадителя 0.2, 0.4 и 0.6 М. Видно, что влияние концентрации осадителя на подвижность атомов белка на траектории такой же длительности наиболее заметно. Так, для ASN37, ALA42, ASN44, ARG45, THR47, ARG68, THR69, PRO70, ALA107, ALA122, TRP123, TRP124, GLY126, LEU129 в цепи А и GLY67, ARG68, THR69 в цепи В разница между значениями RMSF для одних и тех же остатков, но при разных концентрациях осадителя (попарно) составляет более 0.05 нм. Для PRO70 в цепи А разница между значениями RMSF составляет более 0.1 нм для систем с концентрацией NaCl 0.2 и 0.6 М.

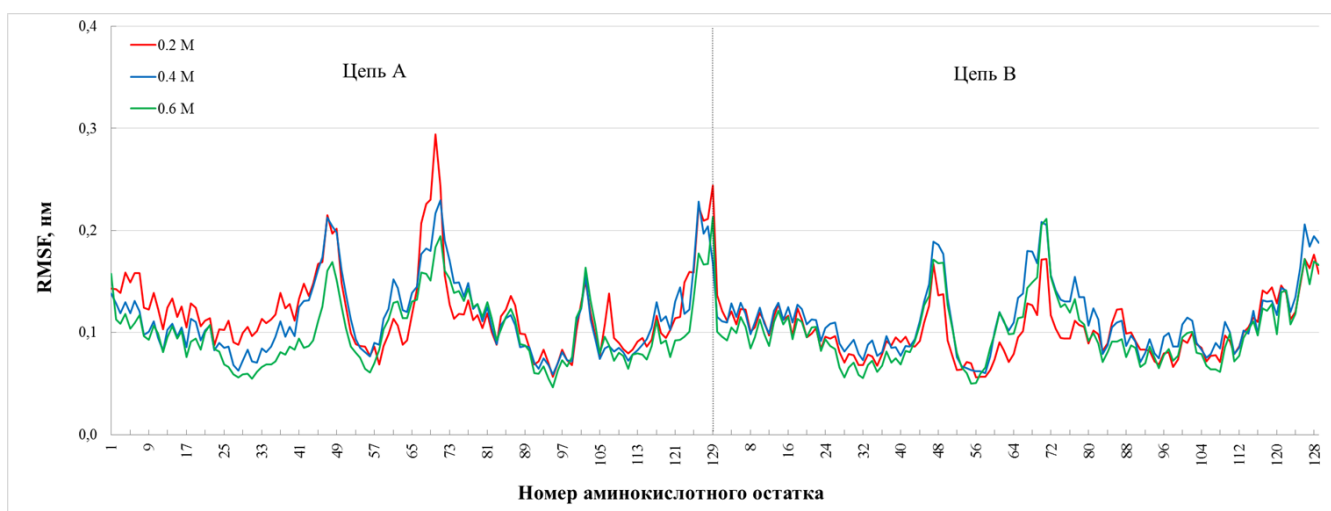


Рисунок 9 – RMSF C $\alpha$ -атомов димера лизоцима в растворе при концентрациях осадителя 0.2 (красная кривая), 0.4 (синяя) и 0.6 М (зеленая).

Сравнение полученных результатов показало, что при оценке стабильности димеров лизоцима с помощью МД концентрация осадителя в растворе играет более существенную роль, чем размеры ячейки моделирования. Даже на относительно короткой траектории (10 нс) разница в значениях RMSF, усреднённых по всем атомам C $\alpha$ , составила 0.13 Å для димеров в разных ячейках. В то время как разница между RMSF димеров, смоделированных при концентрациях осадителя 0.4 и 0.6 М, составила 0.14 Å; при концентрациях 0.2 и 0.4 М – 0.16 Å; и при концентрациях 0.2 и 0.6 М – 0.19 Å. Следует отметить, что с увеличением количества ионов хлорида натрия в диапазоне от 0.2 до 0.6 М мономер 1 в составе димера стабилизируется постепенно, при этом зависимость устойчивости мономера 2 от концентрации осадителя не является монотонной.

### *2.2.2. Минимально допустимый размер ячейки моделирования на основе сравнения стабильности олигомеров лизоцима в кристаллизационном растворе*

Аналогичные исследования были проведены для димера и гексамера типа А (выделенных из октамера А) на более продолжительном промежутке времени (700 нс – 1 мкс) и для большего числа ячеек (пяти). Следует отметить, что по результатам МУРР и МУРН [8,10] в кристаллизационном растворе лизоцима (после добавления осадителя NaCl) были обнаружены лишь димеры и октамеры, а тетрамеров и гексамеров в нем не наблюдалось. Поэтому размеры ячейки считались допустимыми, если димер в ней являлся стабильнее гексамера. Поскольку олигомеры имеют разный объем, то минимальные размеры ячеек для разных олигомеров тоже различаются. Поэтому был введен универсальный параметр размера ячейки – минимальное расстояние между ее гранями и атомами белка (далее – отступ).

Кубические ячейки моделирования были заданы таким образом, чтобы отступ составлял 1, 1.5, 2, 2.5 и 3 нм, соответственно, длины ребер ячеек равнялись 8.6, 9.7, 10.6, 11.6 и 12.6 нм для димера и 10.9, 11.9, 12.9, 13.9 и 14.9 нм в случае гексамера (табл. 3). Для каждой ячейки продолжительность динамики димеров составила 1 мкс, а гексамеров – 700 нс, т. к. за это время все гексамеры успели полностью потерять связи между некоторыми мономерами и дальнейшие расчёты не имели смысла для исследуемой задачи, но потребовали бы большое количество времени и вычислительных ресурсов.

Важно отметить, что задание объема ячеек так, чтобы отступ составлял менее 1 нм при учете нековалентных взаимодействий в радиусе 1 нм (как в данной и большинстве работ) может приводить к существенным артефактам. Вследствие периодических граничных условий копии белка в соседних виртуальных ячейках будут оказывать нереалистично большое влияние друг на друга.

Таблица 3. Длина ребра ячеек и параметры оценки стабильности олигомеров в соответствии с заданным минимальным расстоянием между гранью ячейки и атомами олигомера (отступом). Значения RMSF, RMSD и  $R_g$  усреднены по всем атомам  $C_\alpha$ . Для удобства сравнения жирным шрифтом выделены величины наименее стабильных димеров и наиболее стабильных гексамеров.

Отступ, нм	Длина ребра ячейки, нм		RMSF <sup>max</sup> , нм		RMSF <sup>max</sup> - RMSF <sup>min</sup> , нм		RMSD <sup>max</sup> , нм		$R_g^{\max}$ - $R_g^{\min}$ , нм	
	Д*	Г**	Д	Г	Д	Г	Д	Г	Д	Г
1	8.6	10.9	0.4	2	0.3	1.4	0.4	2.5	<b>0.2</b>	1.6
1.5	9.7	11.9	0.6	3	<b>0.6</b>	2.2	0.8	5.6	<b>0.2</b>	3.5
2	10.6	12.9	0.3	<b>1.4</b>	0.3	<b>1.2</b>	0.5	1.6	<b>0.2</b>	<b>0.8</b>
2.5	11.6	13.9	<b>0.7</b>	3.9	<b>0.6</b>	3.5	<b>1</b>	<b>1.2</b>	<b>0.2</b>	3.7
3	12.6	14.9	0.3	1.6	0.3	1.3	0.4	2.8	<b>0.2</b>	1.4

\* Д – димер

\*\* Г – гексамер

Как видно из рис. 10а и табл. 3, максимальное значение RMSF для гексамеров в каждой из ячеек составляет не менее 1.4 нм (для отступа 2 нм). В то же время, для димеров, независимо от размера ячейки, RMSF не превышает 0.7 нм (для отступа 2.5 нм). Это означает, что даже самое низкое из максимальных значений RMSF гексамеров (1.4 нм) в два раза превышает самое высокое значение RMSF среди всех димеров (0.7 нм). Наименьший разброс значений RMSF гексамеров (для отступа 2 нм) составляет 1.2 нм, что в 2 раза больше, чем у димеров с наибольшим разбросом RMSF – 0.6 нм (для отступа 1.5 и 2.5 нм).

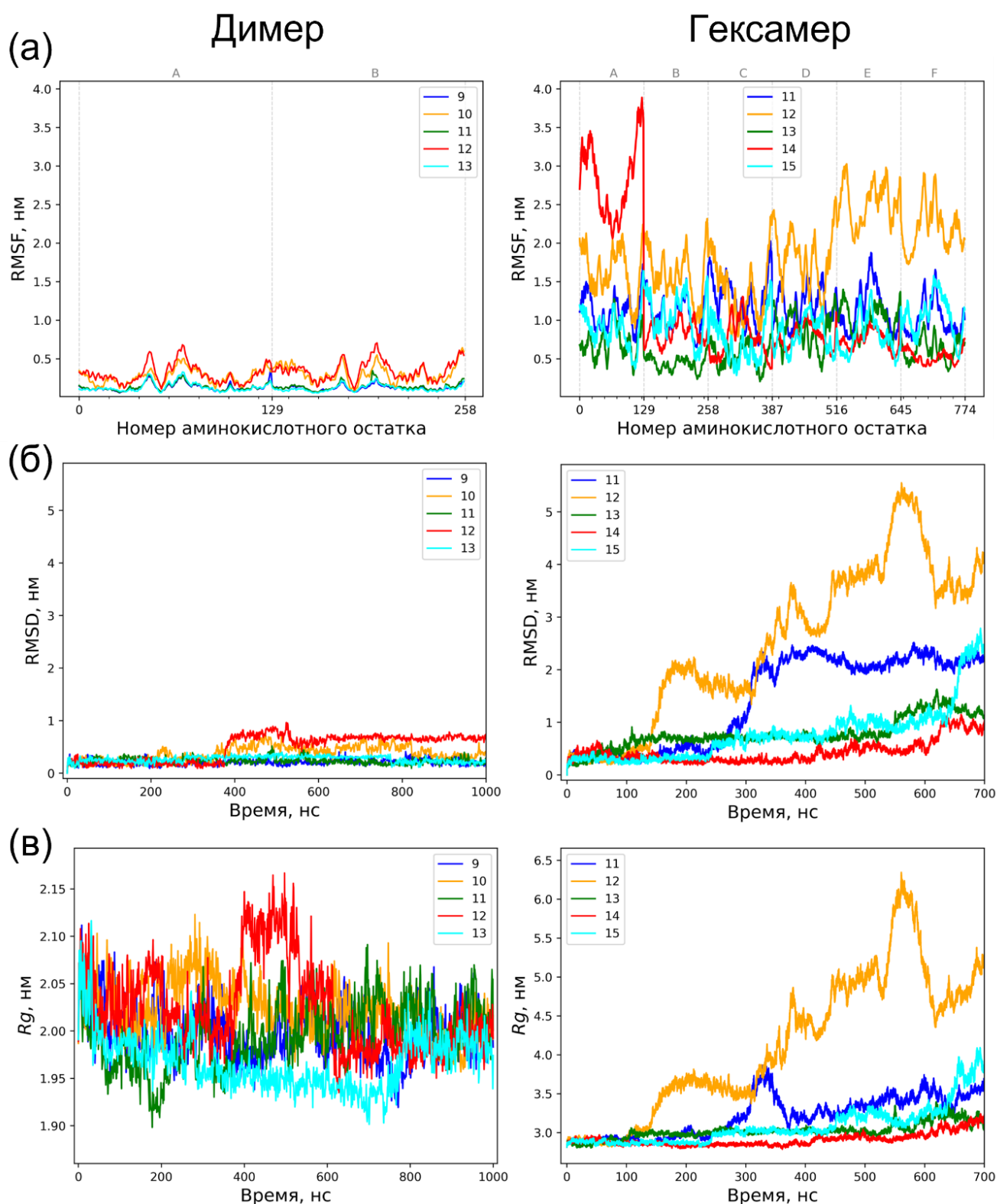


Рисунок 10 – RMSF (а), RMSD (б) и радиус инерции ( $R_g$ , в) димеров (слева) и гексамеров (справа)  $C_\alpha$ -атомов лизоцима, смоделированных в ячейках с различными длинами граней ячеек: 9, 10, 11, 12 и 13 для димера и 11, 12, 13, 14 и 15 для гексамера.

Из рис. 10б и табл. 3 следует, что среди гексамеров наименьшие изменения в белковой структуре наблюдаются в ячейке с отступом 2.5 нм, поскольку соответствующая кривая (красная) на рис. 10б подавляющую часть времени находится ниже кривых для остальных гексамеров. Кроме того, максимальное и конечное (на 700-й наносекунде) значения RMSD составляют 1.2 и 0.7 нм соответственно, что меньше, чем у остальных гексамеров. Максимальное значение RMSD димера с самыми большими деформациями (для отступа 2.5 нм) составляет почти столько же – 1 нм. Однако структура данного димера быстро перешла в стабильную конформацию (приблизительно на 570 наносекунде) с RMSD, составляющим 0.7 нм от начальной структуры, и оставалась в ней до конца моделирования, т. к. затем RMSD отклонялось от 0.7 нм лишь незначительно. В то же время структура гексамера не успела стабилизироваться за 700 нс и ее RMSD, вероятно, продолжило бы расти.

Радиус инерции димера, как следует из рис. 10в и табл. 3 в любой ячейке изменялся в пределах 0.2 нм, в то время как у гексамера с наиболее стабильным объемом (в ячейке с отступом 2 нм) это значение составило 0.8 нм, что в 4 раза больше, чем у любого димера.

Визуальная оценка траекторий показала, что все гексамеры полностью потеряли связи между несколькими парами мономеров, а димеры, наоборот, оставались целыми, хотя один изменил конформацию (при отступе 2.5 нм). Для наглядности на рис. 11 приведены самые нестабильные димеры и наиболее устойчивые гексамеры.

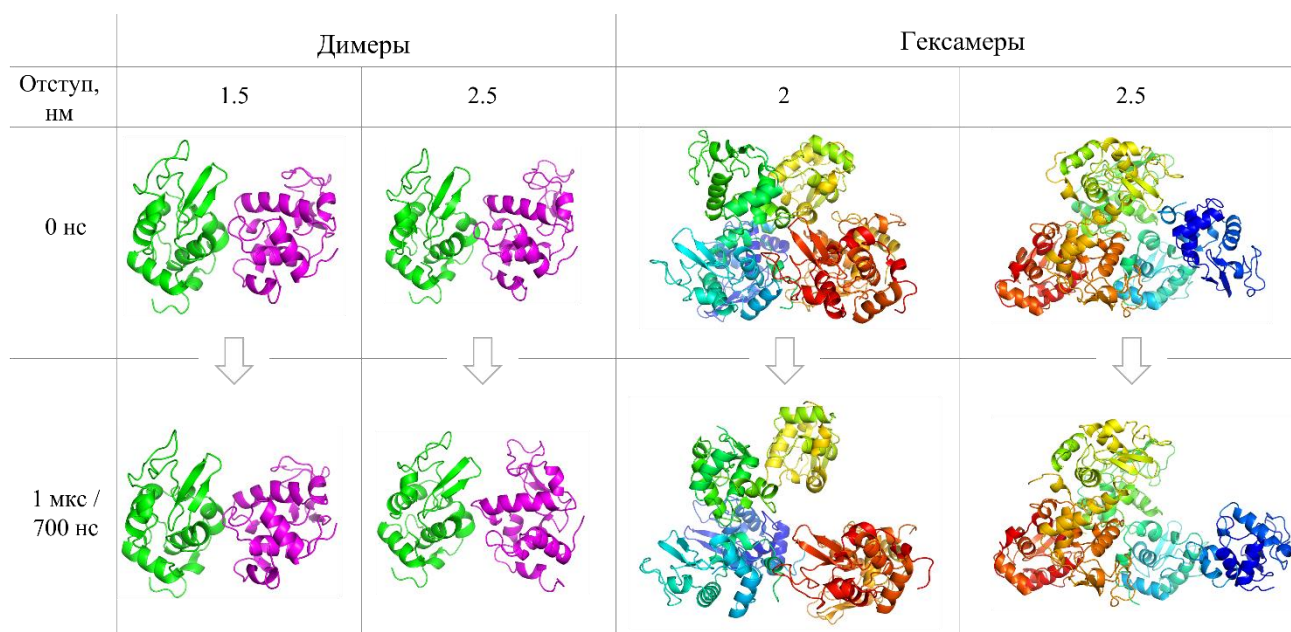


Рисунок 11 – Структура наименее стабильных димеров и наиболее устойчивых гексамеров лизоцима в начале (0 нс) и в конце (1 мкс для димеров и 700 нс для гексамеров) моделирования. Наглядно показано, что самые нестабильные димеры устойчивее самых стабильных гексамеров, поскольку некоторые пары молекул гексамера потеряли все связи между собой: при отступе 2 нм пропали все связи между желтым мономером и синим, оранжевым и красным мономерами; при отступе 2.5 нм – между желтым и синим, бирюзовым мономерами.

Таким образом, по всем вышеперечисленным критериям гексамеры лизоцима менее стабильны, чем димеры независимо от размеров ячейки моделирования. Более того самый стабильный гексамер оказался менее устойчив, чем самый нестабильный димер. Поскольку с помощью МУРР [8] определено, что в кристаллизационном растворе лизоцима образуются димеры (и октамеры), но отсутствуют гексамеры (и тетрамеры) белка, то молекулярно-динамические результаты согласуются с экспериментальными при любых из рассматриваемых размеров ячейки моделирования.

Время МД расчетов напрямую зависит от количества атомов в ячейке, то есть от ее объема (по крайней мере, при исследовании растворов). Следовательно, при изучении кристаллизационных растворов белков целесообразно использовать

ячейку минимально возможных размеров. Среди исследуемых в данной работе ячеек, при которых результаты моделирования корректны (то есть среди всех рассматриваемых), наименьшими размерами обладают ячейки с длиной ребра 8.6 и 10.9 нм для димера и гексамера соответственно, то есть при минимальном расстоянии между белком и гранью, составляющим 1 нм. Дальнейшее уменьшение ячейки уже может привести к артефактам, так как нековалентные взаимодействия учитываются в радиусе 1 нм.

### 2.3. Применимость крупнозернистого силового поля MARTINI для моделирования белковых олигомеров в кристаллизационном растворе

«Полноатомное» (атомистическое) моделирование, при котором каждому атому ставится в соответствие одна частица, имеет ряд ограничений, особенно в отношении пространственных и временных масштабов молекулярных систем и процессов, которые можно эффективно исследовать. Одним из способов их преодоления является использование крупнозернистых силовых полей, в которых несколько атомов специфическим образом объединяют в одну крупную (псевдо)частицу (или «зерно», рис. 12). Это уменьшает количество взаимодействующих частиц в ячейке моделирования, что, в свою очередь, сокращает число арифметических операций, выполняемых как при расчёте сил, действующих на каждую частицу (атом или зерно), так и численном решении уравнений движения каждой частицы. Более того, шаг интегрирования в крупнозернистом подходе увеличивается по сравнению с полноатомным моделированием, поскольку, во-первых, основное из множества допущений такой параметризации – игнорирование некоторых атомистических степеней свободы, например, высокочастотных колебаний связей. В результате, взаимодействия между частицами становятся эффективными, а энергетический ландшафт значительно упрощается, что позволяет существенно увеличить скорость выборки за счет потери деталей. Во-вторых, структурные и термодинамические свойства малочувствительны к размеру к шагу интегрирования [56]. Ускорение расчетов на

2–3 порядка открывает новые возможности для моделирования более крупных систем и более продолжительных процессов, значительно расширяя горизонты исследований в области молекулярного моделирования.

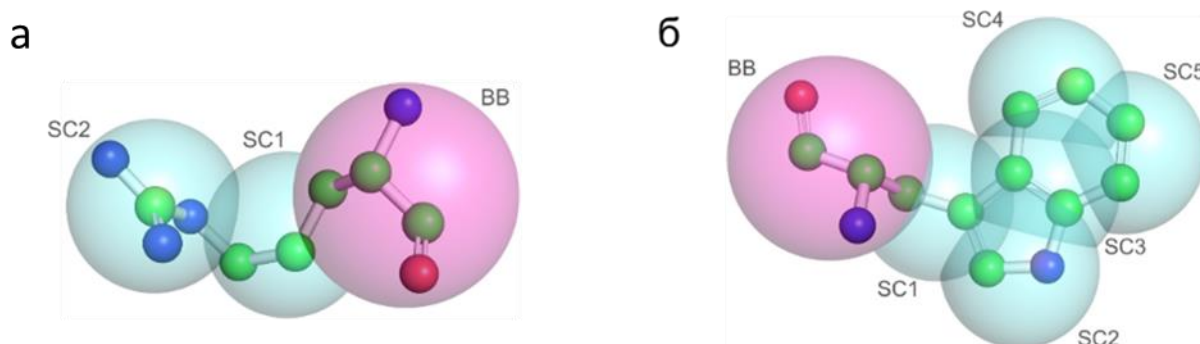


Рисунок 12 – Соответствие атомистической и крупнозернистой (в поле MARTINI) структуры аргинина (*а*) и триптофана (*б*). Атомы углерода, азота и кислорода представлены зелеными, синими и красными маленькими сферами, соответственно. Для наглядности атомы водорода скрыты. Полупрозрачными сферами показаны частицы в MARTINI: розовые (BB – Backbone) относятся к основной цепи белка, а голубые (SC – Side Chain) – к боковой цепи.

Среди крупнозернистых силовых полей одним из наиболее популярных и универсальных является поле MARTINI [97], изначально созданное для моделирования липидов, но затем его применение было распространено на белки [98]. Как правило, в этом поле одним зерном (центром взаимодействия) задаются четыре неводородных атома, за исключением кольцевых фрагментов молекул, которые представляются бóльшим числом частиц (рис. 12). Также в MARTINI в одно зерно объединяется четыре молекулы воды. Зерна различаются по полярности и относительной силе взаимодействия друг с другом.

Несмотря на эффективность метода изучения белковой кристаллизации с помощью моделирования олигомеров, выделенных из кристаллической структуры [11], применение атомистической МД ограничено вычислительными ресурсами. Даже для относительно небольших систем (порядка  $10^5$  атомов, например, олигомер лизоцима в водном растворе) доступные времена моделирования, как

правило, не превышают 1 мкс. При этом для воспроизведения, например, образования олигомера (путем связывания мономеров), являющегося стохастическим и относительно редким событием, необходимо наличие большого количества мономеров в ячейке моделирования. Это неизбежно приводит к многократному увеличению размера системы, делая вычисление длительных траекторий практически невозможным в рамках полноатомного подхода и обосновывая необходимость применения крупнозернистых силовых полей.

Таким образом, проверка применимости крупнозернистого МД-моделирования для изучения кристаллизационных растворов белков, безусловно, представляет собой важную и актуальную задачу. Для ее решения с помощью крупнозернистого силового поля MARTINI [97] в кристаллизационном растворе была смоделирована динамика двух типов (А и В) октамеров лизоцима, являющихся фрагментами его кристаллической (тетрагональной) структуры. Результаты симуляции были сопоставлены с данными, полученными ранее путем вычисления атомистической МД [11]. В [11] было установлено, что лишь один из типов октамеров (А) представляет собой кластер-прекурсор кристалла лизоцима, в то время как другой (В) – распадается на составные элементы.

Подготовка моделей и проведение МД-расчетов происходили по методике, описанной в Разделе 2.1, за исключением приведенных далее особенностей.

Топология двух типов октамеров лизоцима (А и В) в формате силового поля MARTINI 3 [97,98] была сгенерирована с помощью скрипта martinize2.py [99] на основе атомистических структурных моделей [11]. В результате чего одна молекула лизоцима, состоящая из 129 аминокислотных остатков (а.о.), вместо 1022 атомов стала содержать 304 зерна (в 3.4 раза меньше, рис. 13).

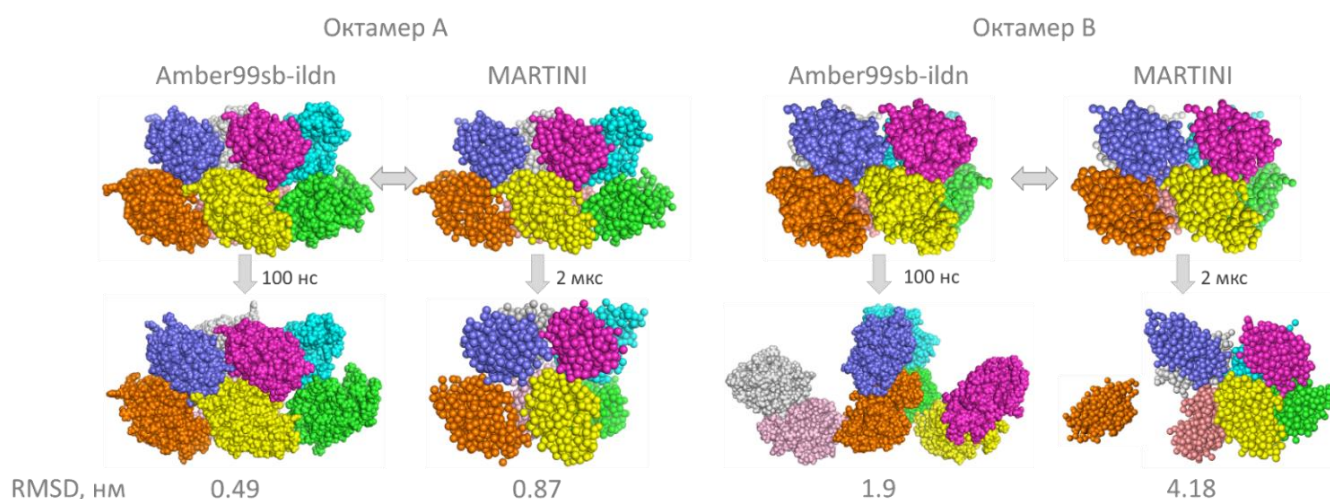


Рисунок 13 – Структура октамеров А (слева) и В (справа) в начале (сверху) и в конце (снизу) моделирования в полноатомном (Amber99sb-ildn) и крупнозернистом (MARTINI) силовых полях. Молекулы (полипептидные цепи) лизоцима в составе октамеров выделены разными цветами. В Amber99sb-ildn одна сфера соответствует одному атому, в MARTINI – группе атомов (зерну). Одна молекула содержит 1022 атома в Amber99sb-ildn и 304 зерна в MARTINI. Снизу приведены попарные значения RMSD между начальной и конечной структурой. Наглядно показано, что в обоих полях октамер В распадается, в то время как октамер А остается более стабильным.

Добавление молекул воды и ионов осадителя NaCl в концентрации 0.4 М происходило с использованием скрипта *insane.py* [100], при этом минимальное расстояние между белком и гранью ячейки составило 7 нм.

Перед запуском «продуктивной» МД моделируемые системы подвергались минимизации энергии методом наискорейшего спуска до тех пор, пока максимальная сила, действующая на любое зерно, не становилась ниже 100 кДж/(моль·нм). После чего ячейки уравнивались в течение 50 нс с помощью баростата «C-rescale» [101] в *NPT*-ансамбле. Расчеты продуктивной МД проводились в *NPT*-ансамбле с использованием термостата «V-rescale» [91] при 20°C и баростата Парринелло-Рамана при 1 атм [66]. Интегрирование уравнений движения выполнялось с шагом 20 фс. Электростатические взаимодействия рассчитывались методом реакционного поля [102] и вычислялись явно только в

сфере с радиусом 1.1 нм, за пределами которой диэлектрическая проницаемость оставалась однородной ( $\epsilon_r = 15$ ). Продолжительность каждой рассчитанной траектории составила 2 мкс.

По результатам моделирования был проведен сравнительный анализ структурной стабильности исследуемых комплексов. На рисунках 14а, 14б представлены графики RMSF  $C_\alpha$ -атомов; интервал между делениями горизонтальной оси составляет 129, что соответствует количеству а.о. в одной полипептидной цепи лизоцима.

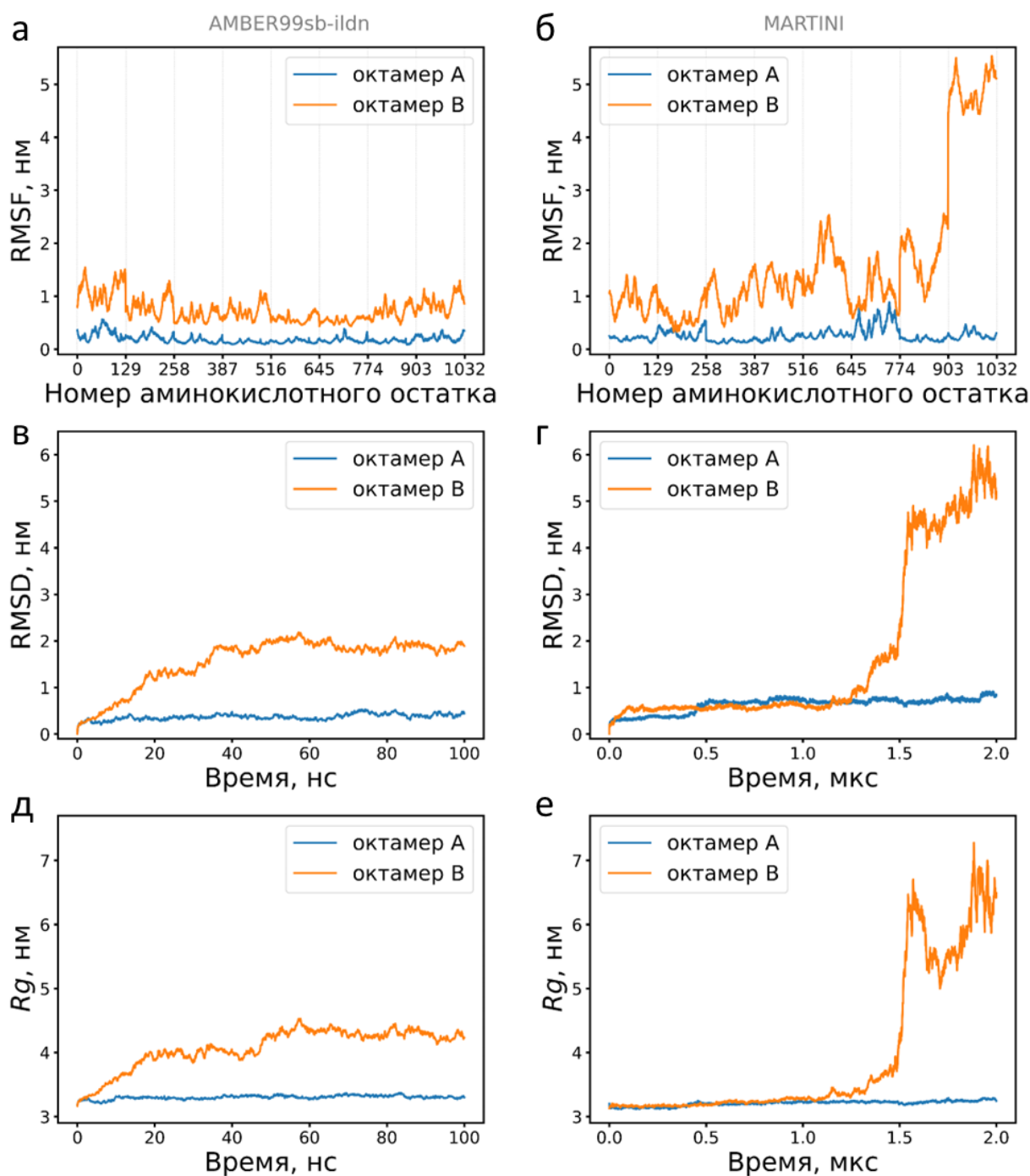


Рисунок 14 – Сравнение стабильности октамеров А и В в кристаллизационном растворе в ходе МД, смоделированной в разных силовых полях: полноатомном Amber99sb-ildn (слева) и крупнозернистом MARTINI (справа). Устойчивость оценивалась с помощью критериев RMSF (а, б), RMSD (в, г) и  $R_g$  (д, е). Комплексный структурно-динамический анализ подтверждает более высокую стабильность октамера А по сравнению с октамером В, что указывает на согласованность результатов, полученных в силовых полях Amber99SB-ILDN и MARTINI.

Как следует из рисунка 14б, при моделировании в поле MARTINI молекула в составе октамера В с номерами остатков 904–1032 демонстрирует крайне высокие значения RMSF (более 5 нм), что свидетельствует о ее существенной нестабильности. Визуализация финальной конформации (рис. 13) подтверждает, что данная субъединица (выделена оранжевым) полностью утратила контакты с остальной частью комплекса.

В случае использования поля Amber99sb-ildn картина дестабилизации октамера В незначительно отличалась. Все молекулы октамера В демонстрируют схожие структурно-динамические флуктуации (рис. 14а), что согласуется с существенным изменением формы комплекса на рис. 13. К концу моделирования (по истечении 100 нс) каждая субъединица октамера В сохранила связь как минимум с одним (в случае «концевых» молекул) или двумя соседями. Это указывает на частичную, но не полную, диссоциацию комплекса. Тем не менее, наблюдаемый разрыв контактов между отдельными субъединицами однозначно указывает на термодинамическую нестабильность октамера В.

Различия в графиках RMSF для октамера А в рассматриваемых силовых полях менее существенны (рис. 14а, 14б). В поле MARTINI молекулы, состоящие из а.о. 130–258 и 646–774, демонстрируют несколько повышенную подвижность, тогда как значения RMSF для остальных цепей флуктуируют около величины 0.25 нм (рис. 14б). Визуальный анализ (рис. 13) показывает, что целостность октамера А сохраняется в обоих силовых полях, несмотря на более заметные изменения формы в крупнозернистой модели. Данное различие объясняется тем, что длительность симуляции в MARTINI (2 мкс) в 20 раз превышала время полноатомной траектории (100 нс), что позволило комплексу накопить бóльшие структурные отклонения от начальной конфигурации.

В поле MARTINI структура октамера А претерпела наиболее значительные изменения через 0.5 мкс после начала симуляции (рост RMSD практически в 2 раза, рис. 14г), а затем перешла в относительно стабильное состояние, сохраняющееся на протяжении оставшихся 1.5 мкс. Устойчивость конформации в

течение бóльшей части времени привело к отмеченным ранее невысоким флуктуациям от усредненной структуры (RMSF, рис. 14б). При попарном выравнивании начальной и конечной конфигураций октамера А значение RMSD в крупнозернистом поле (0.87 нм) почти в 2 раза превышает аналогичный показатель для атомистической симуляции (0.49 нм, рис. 13).

В поле Amber99sb-ildn октамер В начинает дестабилизироваться гораздо раньше, чем в MARTINI (рис. 14в, 14г). В полноатомном поле значимые структурные изменения (RMSD ~1 нм) наблюдаются уже к 18-й нс (рис. 14в), тогда как в крупнозернистой модели аналогичный рост RMSD наблюдается лишь по истечении 1.3 мкс (рис. 14г). Тем не менее, амплитуда структурных трансформаций в поле MARTINI оказалась существенно выше: максимальное значение RMSD достигает 6 нм (рис. 14г), что в 3 раза больше, чем при расчетах в Amber99sb-ildn (рис. 14в). Значение RMSD между начальной и конечной структурами октамера В в крупнозернистом поле оказалось примерно в 2 раза больше, чем в атомистическом (4.18 против 1.9 нм, рис. 13).

Аналогичные тенденции демонстрируют графики радиуса гирации октамера В (рис. 14д, 14е). Анализ зависимостей RMSD и  $R_g$  указывает на различие в механизмах дестабилизации: в поле MARTINI структурный переход носит резкий характер (в районе 1.5 мкс), тогда как в Amber99sb-ildn изменения протекают более плавно.

Совокупный анализ параметров устойчивости (RMSF, RMSD,  $R_g$ , рис. 14) свидетельствует о том, что октамер А обладает существенно более высокой стабильностью по сравнению с октамером В, независимо от используемого силового поля (Amber99sb-ildn или MARTINI). Этот вывод также подтверждается визуализацией структурных изменений и значениями RMSD, рассчитанными между начальными и конечными конформациями (рис. 13).

Таким образом, результаты МД-моделирования октамеров А и В, выполненного с использованием полноатомного и крупнозернистого подходов,

демонстрируют хорошее согласие. Выявленные различия в деталях динамического поведения не влияют на достоверность качественных выводов, полученных с помощью MARTINI. Это подтверждает применимость данного крупнозернистого силового поля для моделирования кристаллизационных растворов белков и оценки стабильности олигомерных кластеров.

#### 2.4. Заключение

Ввиду чрезвычайно высоких вычислительных затрат, характерных для моделирования полного процесса кристаллизации белка, в настоящей работе используется более рациональный подход: метод атомистической МД применяется для изучения поведения и устойчивости олигомерных ассоциатов, формирующихся на предкристаллизационной стадии. Такой подход обеспечивает существенную экономию вычислительных ресурсов и позволяет сосредоточиться на ключевых межмолекулярных взаимодействиях, лежащих в основе зарождения кристаллической фазы.

Модели олигомеров выделяются из известной кристаллической структуры белка, размещенной в базе PDB [2]. Каждый из олигомеров помещается в центр ячейки моделирования, остальное пространство которой заполняется либо водой, либо водой с добавлением ионов осадителя, имитируя, таким образом, кристаллизационный раствор. Затем энергия всей системы минимизируется, и с помощью уравнивания в *NVT*- и *NPT*-ансамблях задаются необходимые температура и давление в ячейке. Далее производятся вычисления «продуктивной» МД, результатами которых являются траектории движения всех атомов (в виде файлов с координатами всех атомов в каждый момент времени). Все симуляции производились с использованием программы GROMACS [41] в силовых полях Amber ff99SB-ILDN [83] и MARTINI [97]. Анализ поведения олигомеров и изменения их структур производится на основе графиков RMSF, RMSD и  $R_g$ , а также визуализации результатов с помощью таких программ как VMD [77] и PyMol [79].

В ходе исследования влияния размеров ячейки моделирования на стабильность димеров и гексамеров лизоцима было выявлено, что, во-первых, поведение димеров больше зависит от концентрации осадителя, чем от объема ячейки. Во-вторых, **определены минимально допустимые размеры ячейки**, при которых результаты исследования стабильности олигомеров белков в кристаллизационных растворах остаются корректными: при минимальном расстоянии между атомами белка и гранью ячейки, составляющим 1 нм. В целях экономии вычислительных ресурсов все дальнейшие атомистические расчеты проводились при данном отступе – 1 нм.

Сравнительный анализ результатов МД-моделирования двух типов октамеров лизоцима (выделенных из кристаллической структуры) в силовых полях Amber99sb-ildn и MARTINI выявил, что октамер А демонстрирует более высокую стабильность по сравнению с октамером В, вне зависимости от выбора поля. Согласованность результатов, полученных в рамках обоих подходов, обосновывает **возможность использования поля MARTINI для эффективного исследования кристаллизационных растворов**, в частности, идентификации кластеров-прекурсоров белковых кристаллов.

Результаты, представленные в Главе 2 опубликованы в работах [A1, A6, A7].

### ГЛАВА 3. СТРУКТУРА И ТЕРМОСТАБИЛЬНОСТЬ ДИМЕРОВ, ОБРАЗУЮЩИХСЯ В КРИСТАЛЛИЗАЦИОННОМ РАСТВОРЕ ПРОТЕИНАЗЫ К

Одним из ключевых этапов кристаллизации является формирование в пересыщенном растворе устойчивых олигомерных ассоциатов – кластеров-прекурсоров, которые служат «строительными блоками» для растущей кристаллической решетки [8,10]. Идентификация структуры и изучение свойств таких прекурсоров способствует пониманию механизмов, происходящих на ранних стадиях самосборки белков, и позволяет выявить факторы, определяющие успешность кристаллизации.

В настоящей главе объектом исследований является протеиназа К (**ProtK**) из гриба *Parengyodontium album* — сериновая протеаза, получившая широкое распространение в биотехнологии благодаря своей исключительной стабильности и высокой каталитической активности. Способность фермента формировать кристаллы в трех различных сингониях (тетрагональной, моноклинной и орторомбической) дополнительно обуславливает интерес к изучению механизмов его самоорганизации.

Отличительными чертами ProtK являются высокая протеолитическая активность в отношении широкого спектра субстратов, включая ДНКазы и РНКазы, и уникальная устойчивость к денатурирующим условиям. ProtK сохраняет функциональные свойства в широком диапазоне температур (до 70–80°C с оптимумом при 56°C [103]), а также в присутствии детергентов (например, додецилсульфата натрия), хелатирующих агентов (этилендиаминтетрауксусной кислоты) и высоких концентраций мочевины [104]. Сочетание этих свойств сделало ProtK незаменимым инструментом в молекулярной биологии для выделения и очистки нуклеиновых кислот (ДНК, РНК); в клинической диагностике для подготовки образцов к анализу (например, ПЦР); в криминалистике для экстракции ДНК из сложных биоматериалов, а также в белковой химии для структурно-функциональных исследований [105].

Несмотря на обширное применение, молекулярные механизмы, лежащие в основе структурной стабильности и способности ProtK к самоорганизации, остаются недостаточно изученными. Понимание влияния температуры на активность и структурные свойства фермента, в том числе на его способность к формированию олигомеров, является ключевым для оптимизации существующих технологических регламентов и разработки новых подходов к его применению.

Экспериментальные данные, полученные ранее методом МУРР, показали, что предкристаллизационные растворы ProtK содержат димеры [106]. Примечательно, что образование (в кристаллизационном растворе ProtK) именно димеров наблюдалось независимо от природы используемого осадителя (например, нитрата или хлорида натрия). Несмотря на то, что метод МУРР позволил определить олигомерный тип прекурсоров (димеры), их пространственная организация (взаимная ориентация мономеров) оставалась неизвестна. При этом выявление конкретной структуры димера критически важно для понимания специфики межмолекулярных контактов, стабилизирующих прекурсор, и механизма дальнейшего роста кристалла.

Следует отметить, что при исследованиях роста белковых кристаллов одинаковой сингонии (тетрагональной, пр. гр.  $P4_32_12$ ) в растворе лизоцима были обнаружены и димеры, и октамеры, в то время как в растворе ProtK формируются только димеры. Таким образом, тип кластера-прекурсора не всегда однозначно определяется сингонией кристаллической решетки, хотя в случае тетрагональной сингонии наблюдаются и определенные сходства (наличие димеров) в структуре растворов лизоцима и ProtK на предкристаллизационной стадии.

### 3.1. Уточнение структуры димера протеиназы K, образующегося в ее кристаллизационном растворе

Целью данного раздела являлась идентификация наиболее вероятной структуры димерного кластера-прекурсора кристалла ProtK с использованием методов молекулярного моделирования.

Анализ межмолекулярных контактов в кристалле ProtK (PDB ID: 7A68), проведенный с помощью сервера PDBePISA [78], показал, что тетрагональная кристаллическая упаковка ProtK допускает существование шести уникальных типов димеров, которые потенциально могли бы выполнять роль прекурсоров. Данные типы димеров, обозначенные как А, В, С, D, Е и F (рис. 15), различаются характером межмолекулярного интерфейса и объемом.

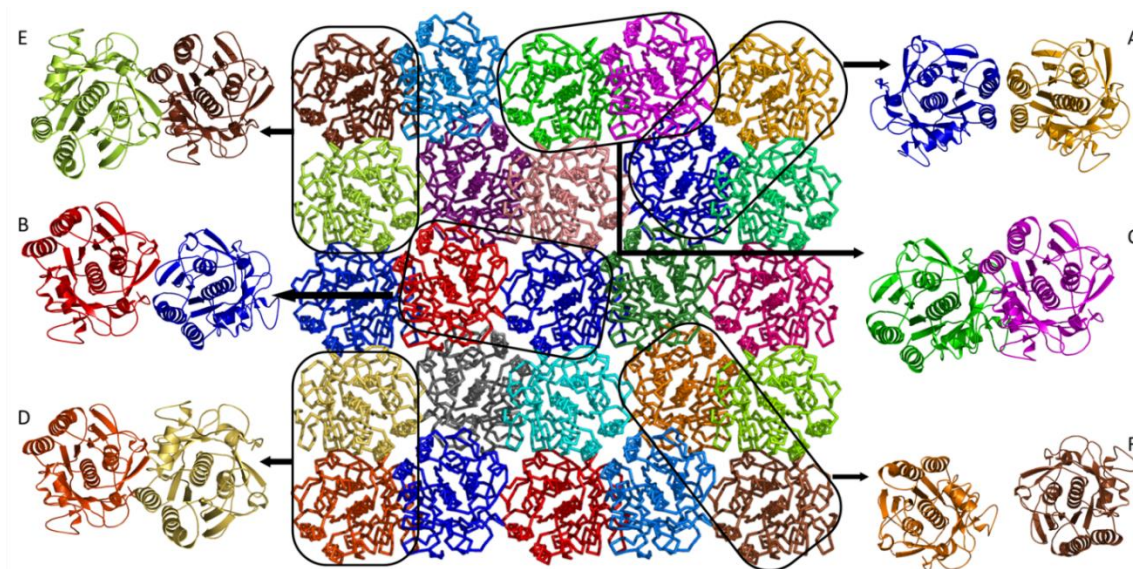


Рисунок 15 – Фрагмент тетрагональной кристаллической структуры ProtK (в центре) и выделенные из нее возможные типы кластеров-прекурсоров (димеры А–F).

Для установления пространственной организации димера, являющегося кластером-прекурсором кристалла ProtK, был применен метод МД, успешно зарекомендовавший себя при решении аналогичной задачи для лизоцима [11]. Данный *in silico* подход предполагает, что в реальном предкристаллизационном растворе преимущественно присутствуют димеры наиболее стабильного типа.

В настоящем разделе проведено МД-моделирование поведения каждого из шести типов димеров А–F в условиях, приближенных к экспериментальным условиям кристаллизации: в водном растворе, содержащем 0.5 М нитрата натрия в качестве осадителя, при рН 8.0 и температуре 20°C. Расчеты выполнялись с использованием программного пакета GROMACS 2021 [41] и атомистического

силового поля Amber ff99SB-ILDN [83]. Размеры ячейки моделирования задавались так, чтобы минимальное расстояние между ее гранью и белком составляло 1 радиус отсечки невалентных взаимодействий (в данной работе 1 нм), что, обеспечивает корректность расчетов и высокую вычислительную скорость (как было показано в Разделе 2.2). Детальное описание протокола приведено в Разделе 2.1.

Количественная оценка и сравнение стабильности димеров проводились на основе анализа трех стандартных структурно-динамических параметров, рассчитанных на основе полученных траекторий для  $C_{\alpha}$ -атомов: RMSF (среднеквадратичные флуктуации) характеризует подвижность полипептидной цепи; RMSD (среднеквадратичное отклонение) отражает структурные изменения относительно исходной кристаллической конформации; радиус инерции ( $R_g$ ) служит мерой компактности белковой структуры.

Сравнение значений RMSF (рис. 16 и табл. 4) выявило значительные различия в гибкости полипептидных цепей димеров. Наименьшие флуктуации атомов и, соответственно, наибольшая стабильность наблюдались для димера типа E, среднее значение RMSF для которого составило всего 0.15 нм. Это существенно (в 1.6 раза) ниже, чем у следующего по стабильности димера C (0.24 нм), и более чем в 3 раза ниже, чем у наименее стабильного димера F (0.48 нм).

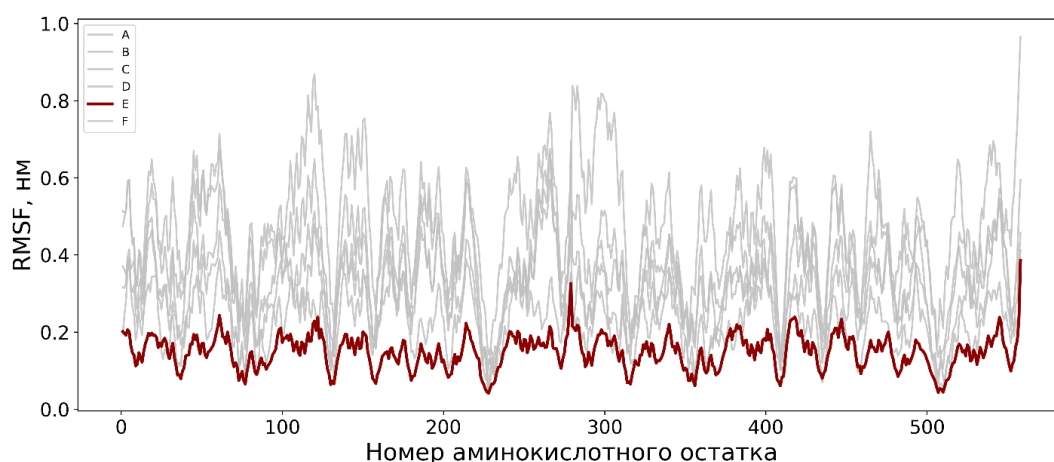


Рисунок 16 – RMSF атомов  $C_{\alpha}$  шести структурно-различных димеров ProtK (A-F) в кристаллизационном растворе.

Таблица 4. Значения RMSF, усредненные по всем атомам  $C_{\alpha}$  димеров ProtK.

Димер	A	B	C	D	E	F
Среднее значение RMSF, нм	0.28	0.37	0.24	0.36	0.15	0.48

Минимальные структурные изменения в ходе 100-нс моделирования подтверждают данный вывод (рис. 17). Структура димера E лучше остальных сохранила свою исходную кристаллическую конформацию к концу моделирования; RMSD димера E не превышало  $\sim 0.25$  нм в течение 95 % времени симуляции. Димер F напротив демонстрировал быстрые и значительные структурные изменения (RMSD достигало  $\sim 1.5$  нм). Интересно, что хотя димер C стабильнее димера A согласно среднему значению RMSF (0.24 против 0.28 нм), его структура начинала отклоняться от исходной раньше (уже после 8 нс), чем у димера A (после 45 нс).

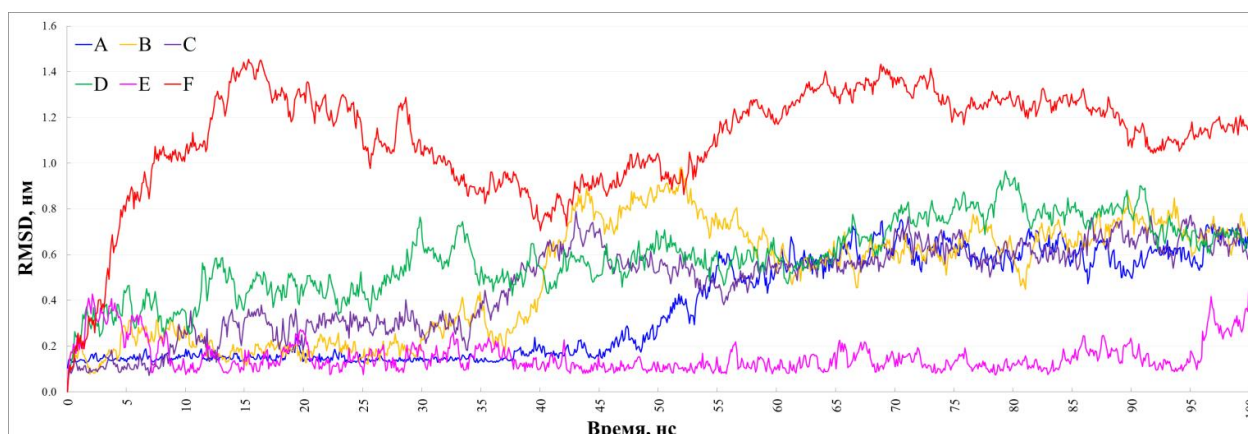


Рисунок 17 – RMSD атомов  $C_{\alpha}$  шести структурно-различных димеров ProtK (A-F) в кристаллизационном растворе.

Анализ динамики радиуса инерции ( $R_g$ , рис. 18) также указывает на уникальную стабильность димера E: несмотря на незначительное уменьшение объема на 95 нс симуляции, размеры димера E оставались практически неизменными на протяжении 95 % времени моделирования, в отличие от других димеров, показавших тенденции либо к уменьшению (F, D), либо к увеличению размеров (A, C).

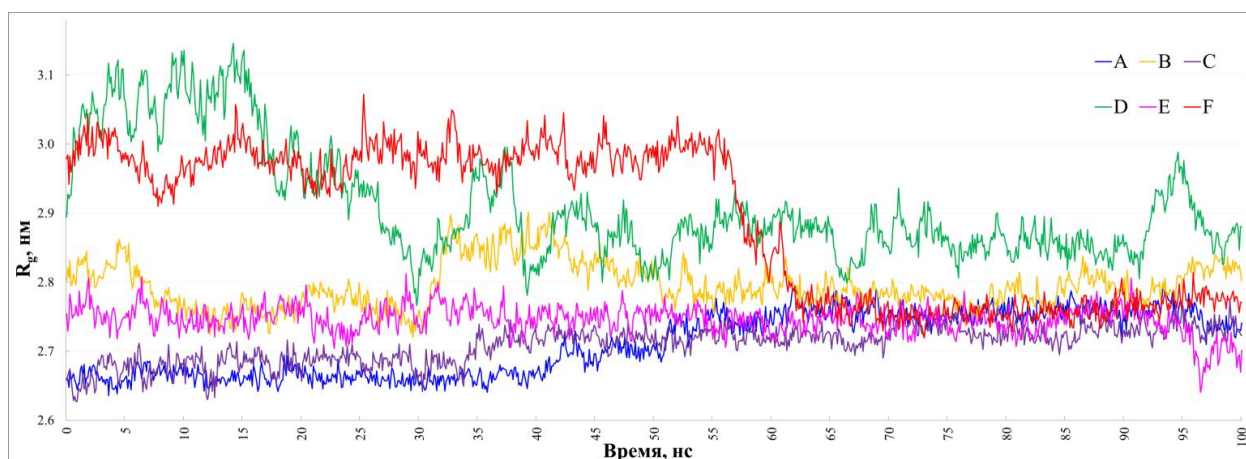


Рисунок 18 – Радиус инерции ( $R_g$ ) шести структурно-различных димеров ProtK (A-F) в кристаллизационном растворе.

Дополнительным и независимым подтверждением корректности идентификации димера E как кластера-прекурсора являются результаты моделирования экспериментальных данных МУРР для раствора ProtK с использованием программы OLIGOMER [107]. Сравнение теоретических кривых рассеяния для смесей мономеров с каждым из трех наиболее стабильных димеров (A, C, E) показало, что наилучшее соответствие экспериментальным данным МУРР (минимальное значение  $\chi^2$ ) достигается именно для смеси мономеров и димеров типа E (табл. 5)

Таблица 5. Анализ экспериментальных данных МУРР с использованием смеси мономеров и димеров ProtK в присутствии 1 М  $\text{NaNO}_3$  [106]. Теоретические кривые для мономерной и димерной составляющих рассчитаны в программе CRY SOL [108]. Приведены значения наилучшего соответствия ( $\chi^2$ ) и объемные доли димеров ProtK ( $v$ ).

мономер+димер А		мономер+димер С		мономер+димер Е	
$\chi^2$	$v, \%$	$\chi^2$	$v, \%$	$\chi^2$	$v, \%$
0.75	7.4	0.76	6.9	<b>0.71</b>	7.3

Таким образом, совокупный анализ результатов МД-моделирования (RMSF, RMSD,  $R_g$ ) и их сопоставление с экспериментальными данными МУРР позволили

однозначно идентифицировать димер типа E как наиболее стабильную конформацию среди шести структурно возможных кандидатов. Этот результат позволяет заключить, что димер E является наиболее вероятным кластером-прекурсором кристалла ProtK, формирующимся на ранних стадиях ее кристаллизации. Последующие исследования в данной главе посвящены детальному изучению температурной стабильности именно этого установленного типа димера.

### 3.2. Влияние температуры на стабильность кластера-прекурсора кристалла протеиназы K

Изучение влияния температуры на стабильность кластеров-прекурсоров кристалла термофильного фермента ProtK имеет как фундаментальное значение для понимания механизмов роста кристаллов, так и практическое — для оптимизации протоколов кристаллизации. Поэтому в настоящем разделе с помощью МД была детально исследована структурно-динамическая стабильность образующегося в кристаллизационном растворе ProtK димера E (см. предыдущий раздел) в широком температурном диапазоне от 20 до 80°C, включающем зону максимальной ферментативной активности ProtK (50–60°C).

Для обеспечения статистической достоверности результатов для каждой из 19 исследуемых температур выполнено по три независимых моделирования длительностью 100 нс (всего 57 симуляций). Протокол симуляций соответствовал описанному в Разделе 2.1.

В дополнение к анализу RMSF C $\alpha$ -атомов проведена оценка динамики межмолекулярного интерфейса. Площадь контакта между мономерами рассчитывалась как разность между суммарной площадью поверхности (доступной для растворителя) отдельных мономеров и площадью поверхности димера. Этот параметр напрямую отражает целостность структуры, поскольку отсутствие площади контакта свидетельствует о диссоциации димера.

### 3.2.1. Влияние температуры на сохранение структурной целостности димера E

Анализ динамики площади межмолекулярных контактов (рис. 19) показал, что полная диссоциация димера E на мономеры (определяемая длительным разрывом контакта, при котором площадь составляет 0 нм<sup>2</sup>) является исключительно редким событием. Она была зафиксирована лишь в 4 из 57 проведенных симуляций (~7 %) при температурах 38, 63, 67 и 70°C. Следует отметить, что даже в этих редких случаях потеря контакта была обратимой в рамках 100 нс в трех из четырех симуляций (при 38, 63 и 70°C). Несмотря на устойчивость прекурсора к необратимым структурным изменениям, обнаружена область температур 63–70°C, соответствующая относительно высокой частоте его диссоциации.

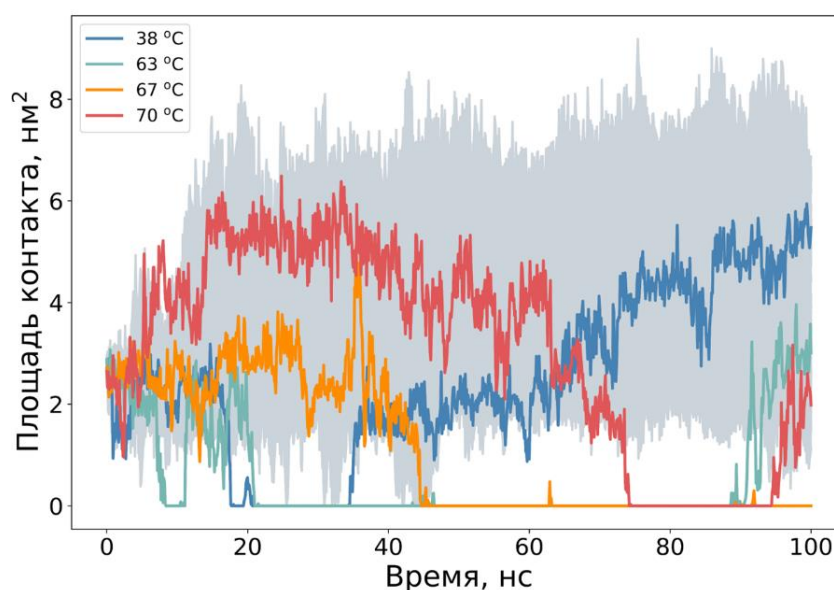


Рисунок 19 – Динамика площади межмолекулярного контакта в димере E ProtK в диапазоне температур 20–80°C. Цветные кривые соответствуют траекториям с длительной диссоциацией (площадь составляет 0 нм<sup>2</sup>) при 38, 63, 67 и 70°C. В остальных моделированиях (~93 %, показаны серым) площадь контакта либо сохранялась, либо достигала 0 нм<sup>2</sup> лишь кратковременно, что указывает на сохранение структурной целостности димера E в течение 100 нс.

### 3.2.2. Влияние температуры на флуктуации (RMSF) атомов димера E

Несмотря на тенденцию к увеличению RMSF с повышением температуры (рис. 20, общая структурная стабильность димера E сохраняется в широком диапазоне (20–60°C), в частности при температурах, близких к области максимальной ферментативной активности ProtK (50–60°C). В интервале 20–60°C наблюдаются умеренные значения RMSF (от ~0.15 до ~0.43 нм, рис. 20), при этом стандартные отклонения между независимыми траекториями низкие ( $\leq 0.3$  нм), что указывает на стабильное и воспроизводимое поведение прекурсора. Наибольшие значения RMSF (до 1.04 нм при 67°C) и стандартных отклонений (до 1.23 нм при 67°C) смещены в область 63–70°C, что указывает на резкое снижение стабильности сразу после зоны максимальной ферментативной активности ProtK (50–60°C). При 80°C прекурсор демонстрирует небольшую стабилизацию (RMSF =  $0.34 \pm 0.09$  нм) по сравнению с интервалом 63–70°C. При этом в самом диапазоне пиковой каталитической активности (50–60°C) димер E сохраняет структурную стабильность (RMSF = 0.32–0.33 нм).

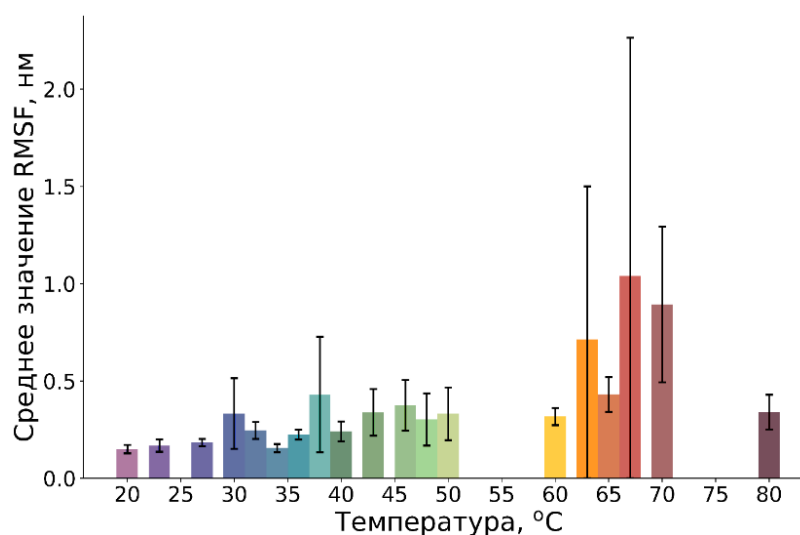


Рисунок 20 – Значения RMSF, усредненные по всем атомам  $C_\alpha$  кластера-прекурсора кристалла ProtK, смоделированного в кристаллизационном растворе при различных температурах (от 20 до 80°C). Указанные погрешности (на основании несмещенной оценки дисперсии) составляют одно стандартное отклонение при усреднении RMSF по трем независимым моделированиям для каждой температуры.

Таким образом, температурный диапазон наибольших флуктуаций (RMSF) атомов димера E совпадает с областью относительно частой диссоциации 63–70°C, что подтверждает наибольшую нестабильность прекурсора именно в данном интервале температур.

### 3.2.3. Кристаллизация протеиназы K после инкубации при 60°C

Известно, что молекула ProtK сохраняет общую стабильность до 80°C, а ее ферментативная активность достигает максимума при 50–60°C [103]. В результате моделирования МД кластера-прекурсора ее кристалла (димера E) обнаружено, что наблюдаемое начало его дестабилизации (закрывающееся в повышении значений RMSF и учащении случаев диссоциации) при 63°C следует сразу за областью максимальной каталитической активности ProtK (50–60°C), и не совпадает с температурой начала денатурации – 80°C. Это позволяет предположить, что верхний температурный предел стабильности предкристаллизационных олигомеров определяется не общей термостабильностью белковой молекулы, а температурой его ферментативной активности.

Обнаруженная в работе устойчивость кластера-прекурсора кристалла ProtK (димера E) в диапазоне температур 20–60°C указывает на возможность роста ее кристаллов при нетипично высоких для большинства белков температурах, а именно при 50–60°C.

Важно отметить, что данная гипотеза, основанная на результатах МД-моделирования, была подтверждена экспериментально. Кристаллизация ProtK была успешно осуществлена методом диффузии паров в режиме висячей капли после ее инкубации при температуре 60°C (рис. 21). Белковый раствор с концентрацией 20 мг/мл в 0.4 М Tris-буфере (pH 8.0) смешивался в равных пропорциях (по 1.5 мкл) с резервуарным раствором, содержащим один из осадителей:  $\text{Li}_2\text{SO}_4$ ,  $(\text{NH}_4)_2\text{SO}_4$  или  $\text{NaNO}_3$  (концентрация 1 М в том же буфере).

Вишняя капля объемом 3 мкл уравнивалась над 300 мкл резервуарного раствора.

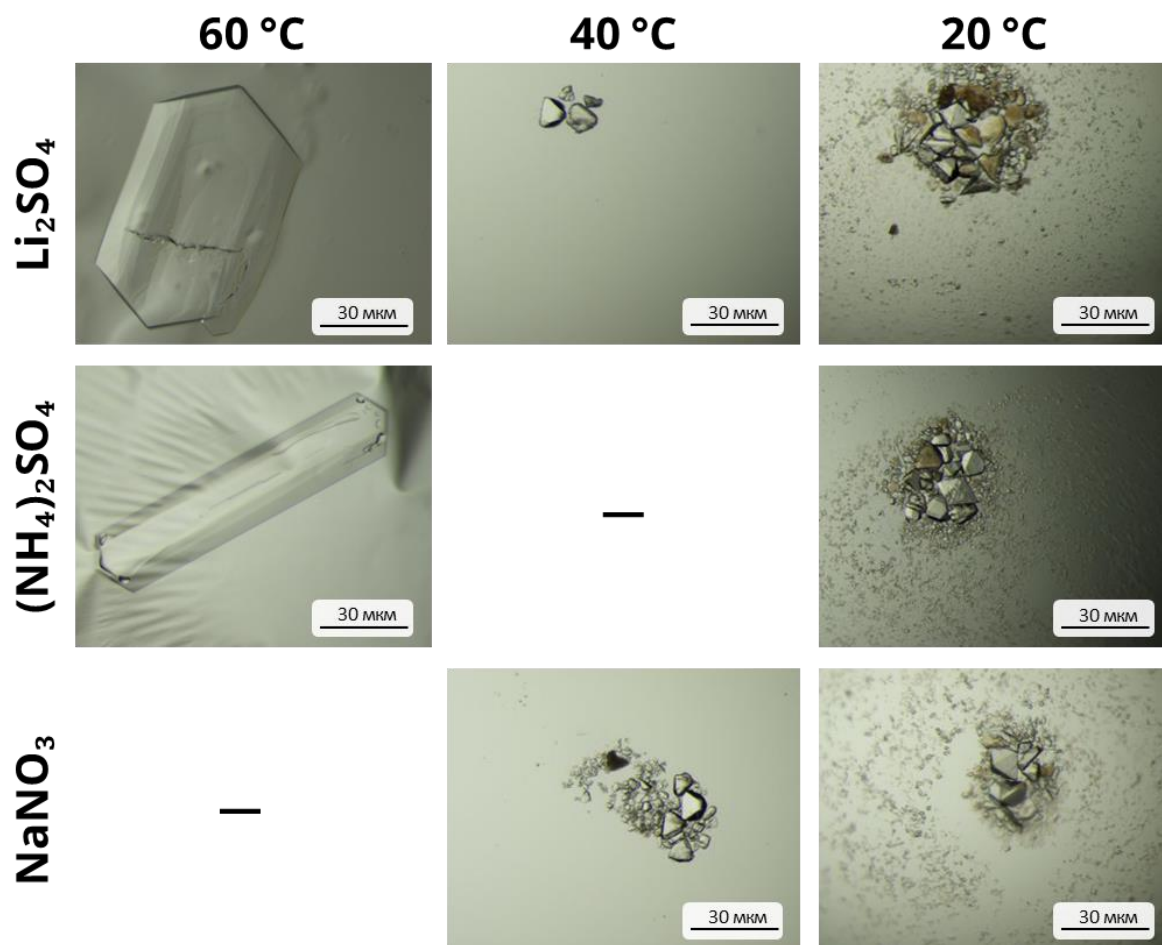


Рисунок 21 – Сравнение наиболее крупных кристаллов ProtK, выращенных после инкубации при 60°C, а также при 40 и 20°C в растворах с тремя осадителями: Li<sub>2</sub>SO<sub>4</sub>, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> и NaNO<sub>3</sub>. Заметно, что при 60°C кристаллы крупнее, а при 40 и 20°C кристаллизация более массовая.

Экспериментальные результаты подтверждают не только адекватность использованного МД-подхода для оценки стабильности прекурсоров, но и его предсказательную силу для определения перспективных температурных диапазонов при выращивании белковых кристаллов. В частности, предсказанная с помощью МД кристаллизация ProtK при 60°C открывает новые значимые возможности для более точной интерпретации механизма ее действия на основе структурных данных, поскольку кристаллизация фермента при температуре его максимальной каталитической активности позволяет зафиксировать нативную,

функционально значимую конформацию белка в кристаллической решетке. Кроме того, проведение кристаллизации при повышенных температурах может способствовать изменению кинетики и термодинамики процесса самосборки, потенциально приводя к образованию альтернативных кристаллических форм (полиморфных модификаций). Получение таких полиморфов позволяет выявить новые конформационные состояния белка, важные для понимания его конформационных переходов и адаптации к различным условиям.

### 3.3. Заключение

В Главе 3 представлены результаты комплексного исследования ранних стадий кристаллизации ProtK с применением метода МД.

Во-первых, на основе сравнительного анализа стабильности шести структурно-различных димеров, являющихся фрагментами кристаллической структуры ProtK, была однозначно **идентифицирована структура наиболее вероятного кластера-прекурсора – димера типа E**. Установление структурной организации «строительного блока» кристалла является принципиально важным шагом для понимания фундаментальных механизмов белковой кристаллизации и служит необходимой основой для дальнейшего моделирования процессов нуклеации и роста кристаллов ProtK.

Во-вторых, была детально исследована температурная стабильность идентифицированного кластера-прекурсора в широком диапазоне 20–80°C. В результате МД-моделирования была **установлена структурная устойчивость димера E в диапазоне температур 20–60°C**, подтвержденная как анализом межмолекулярных контактов, показавшим крайне низкую вероятность диссоциации в этом температурном интервале (лишь в ~7 % симуляций при 38, 63, 67 и 70°C), так и умеренными значениями флуктуаций атомов (RMSF до ~0.43 нм).

Обнаруженная в работе значительная стабильность кластера-прекурсора кристалла ProtK в диапазоне температур 20–60°C позволила выдвинуть **обоснованную гипотезу о перспективности проведения ее кристаллизации при нетипично высоких для большинства белков температурах (50–60°C)**. Данное теоретическое **предположение подтвердилось экспериментально**: кристаллы ProtK были успешно получены после инкубации при температуре 60°C. Этот результат не только демонстрирует предсказательную способность используемого в работе МД-подхода для определения оптимальных условий кристаллизации, но и открывает новые возможности для установления каталитически значимой конформации ProtK.

Результаты, представленные в Главе 3 опубликованы в работах [A2, A3].

#### ГЛАВА 4. СТРУКТУРА И ТЕРМОСТАБИЛЬНОСТЬ ГЕКСАМЕРОВ, ОБРАЗУЮЩИХСЯ В КРИСТАЛЛИЗАЦИОННОМ РАСТВОРЕ ТЕРМОЛИЗИНА

Термолизин (**ThL**), термостабильная металлопротеиназа из *Bacillus thermoproteolyticus*, наряду с ProtK относится к классу термофильных ферментов, проявляющих максимальную ферментативную активность при температуре  $\sim 70^{\circ}\text{C}$  [109]. Благодаря высокой стабильности и каталитической активности в экстремальных условиях, ThL широко применяется в биотехнологии и промышленности. В частности, он используется для специфического расщепления пептидов и белков, что востребовано при секвенировании белков и подготовке образцов для масс-спектрометрии. Кроме того, ThL применяется для синтеза пептидов в условиях обратной реакции, включая получение фармакологически значимых соединений и аналогов биологически активных пептидов. В пищевой промышленности ThL используется для модификации белков с целью улучшения их вкусовых и функциональных свойств, а также для получения низкокалорийных подсластителей на основе дипептидов, таких как аспартам.

Таким образом, ThL представляет собой универсальный фермент, обладающий как фундаментальной ценностью для изучения механизмов действия термофильных протеаз, так и широким спектром практических приложений в медицине, фармацевтике и пищевой индустрии. Сравнительное исследование предкристаллизационных стадий ProtK и ThL важно не только для выявления общих закономерностей формирования кластеров-прекурсоров у термофильных протеаз, но и для определения специфических особенностей, влияющих на кристаллизационный потенциал ThL.

Аналогично подходу, успешно примененному для ProtK в Главе 3, в данной главе для ThL решались две похожие задачи. Для этого была применена методология МД-моделирования, оптимизированная в Главе 2 (использование ячейки минимально допустимых размеров).

Первая задача — уточнение структуры кластера-прекурсора кристалла ThL. Несмотря на данные МУРР, указывающие на гексамерный характер прекурсоров [110], анализ кристаллической структуры (PDB ID: 3DNZ, пр. гр.  $R6_122$ ) выявляет четыре конформационно различных типа гексамера. Для определения пространственной организации гексамерных комплексов, образующихся в кристаллизационном растворе ThL, проведена сравнительная оценка стабильности четырех типов потенциальных прекурсоров.

Вторая задача — изучение структурно-динамической устойчивости идентифицированного кластера в широком температурном диапазоне 10–90°C, включающем область его оптимальной ферментативной активности: 50–70°C [109]. Данный подход позволяет оценить влияние температуры на целостность прекурсора, сопоставить ее с ферментативной активностью ThL и предложить новые потенциальные условия кристаллизации.

#### 4.1. Уточнение структуры гексамера термолизина, образующегося в его кристаллизационном растворе

Для определения наиболее вероятной структуры кластера-прекурсора кристалла ThL, олигомерный тип (гексамер) которого был установлен методом МУРР [110], был проведен анализ его гексагональной кристаллической структуры (PDB ID: 3DNZ). Данный анализ, выполненный с помощью программы PyMOL и операторов симметрии (пр. гр.  $R6_122$ ) позволил выделить четыре уникальных гексамерных комплекса А, В, С и D (рис. 22), различающихся взаимной ориентацией мономеров и площадью межмолекулярных контактов. Данные четыре структуры рассматривались в качестве кандидатов на роль прекурсора, формирующего кристаллическую решетку ThL.

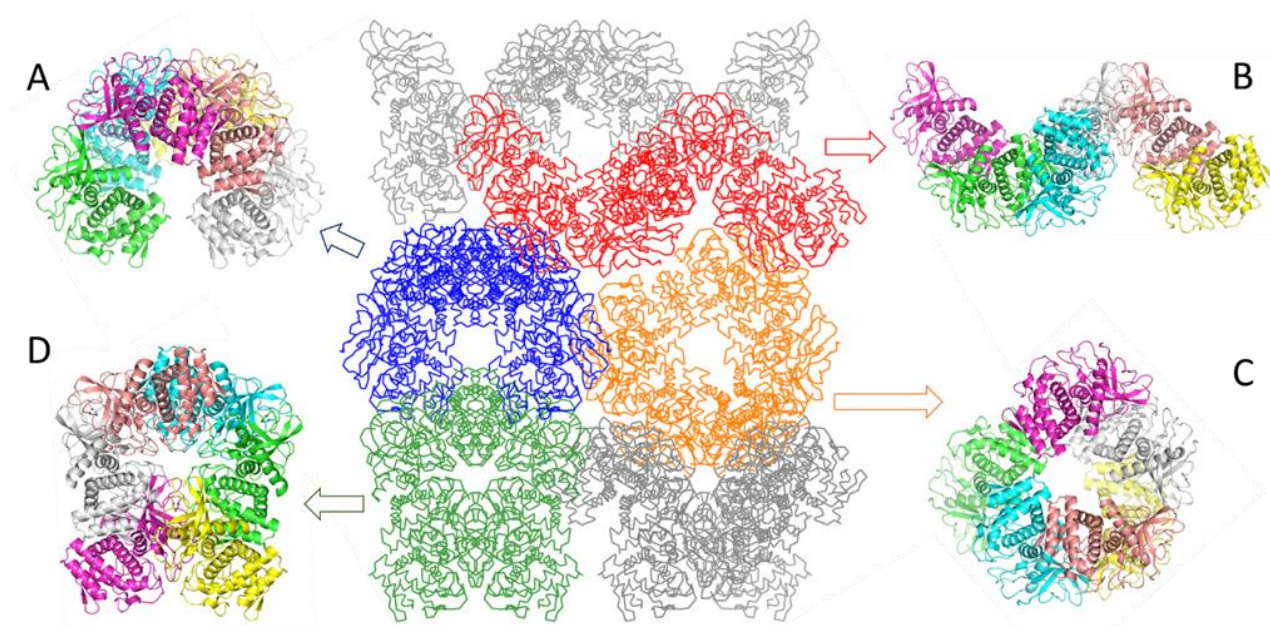


Рисунок 22 – Фрагмент гексагонального кристалла термолизина (в центре) и выделенные из него потенциальные кластеры-прекурсоры (гексамеры А–D).

Оценка относительной стабильности каждого из четырех кандидатов проводилась на основе результатов атомистического МД-моделирования в условиях, соответствующих кристаллизационному эксперименту: в водном растворе, содержащем 0.75 М сульфата аммония в качестве осадителя, при 20°C. Расчеты были выполнены с использованием пакета GROMACS 2021 [41] и силового поля Amber ff99SB-ILDN [59]. Оптимизированный протокол моделирования (описанный в Главе 2) включал использование ячейки с минимальным расстоянием между ее гранью и белком, составляющим 1 нм, что обеспечивало корректность моделирования невалентных взаимодействий. Длительность МД-траектории составила 100 нс для каждого гексамера. Детальный протокол моделирования соответствует представленному в Разделе 2.1. Стабильность оценивалась с помощью стандартных структурно-динамических характеристик RMSF, RMSD и радиуса инерции  $R_g$ , рассчитанных для  $C_\alpha$ -атомов.

Согласно графикам RMSF (рис. 23) наименьшие значения флуктуаций  $C_\alpha$ -атомов, свидетельствующие о повышенной структурной жесткости полипептидных цепей, наблюдаются для гексамеров типа А и D. Однако

однозначно определить наиболее стабильный из данных типов структур на основе характеристики RMSF не удалось, поскольку усредненные по всем  $C_{\alpha}$ -атомам значения RMSF оказались слишком близки: 0.29 нм для гексамера А и 0.30 нм – для D.

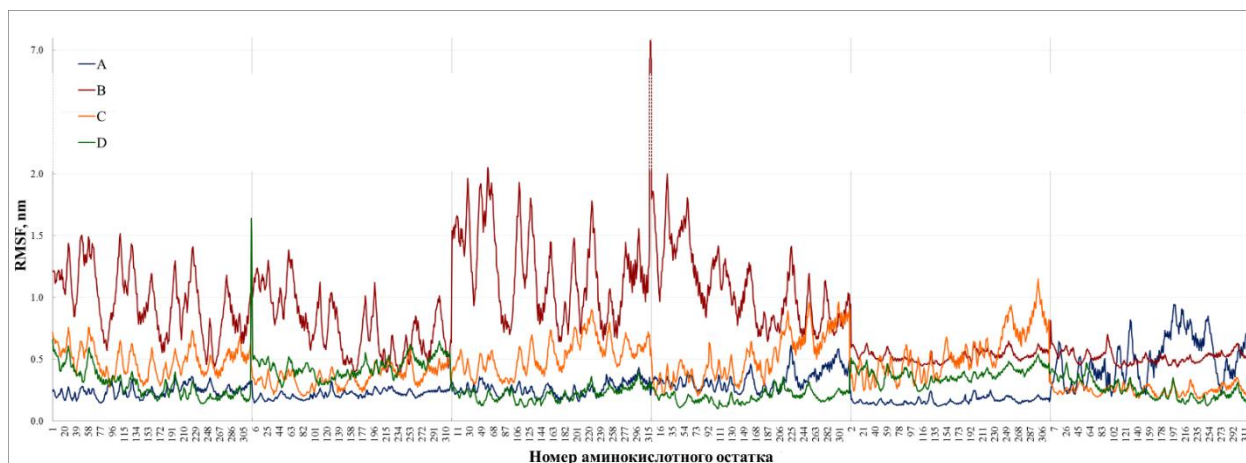


Рисунок 23 – RMSF атомов  $C_{\alpha}$  четырех структурно-различных гексамеров А–D ThL в кристаллизационном растворе.

Однозначно выявить наиболее стабильный тип гексамера позволил анализ RMSD (рис. 24), согласно которому структура гексамера А испытывала наименьшие отклонения от исходной кристаллической конформации на протяжении всей 100-нс симуляции (RMSD  $\sim$ 0.5 нм в интервале 10–100 нс). Начиная примерно с 10 нс, значение RMSD для гексамера D было стабильно выше, чем для гексамера А, достигая  $\sim$ 0.8 нм к концу симуляции, что указывает на его меньшую структурную устойчивость по сравнению с типом А. Гексамеры В и С демонстрировали еще большие отклонения (RMSD  $\sim$ 2.3 нм для В и  $\sim$ 1.4 нм для С к 100 нс), подтверждая их более низкую стабильность, что согласуется с данными RMSF.

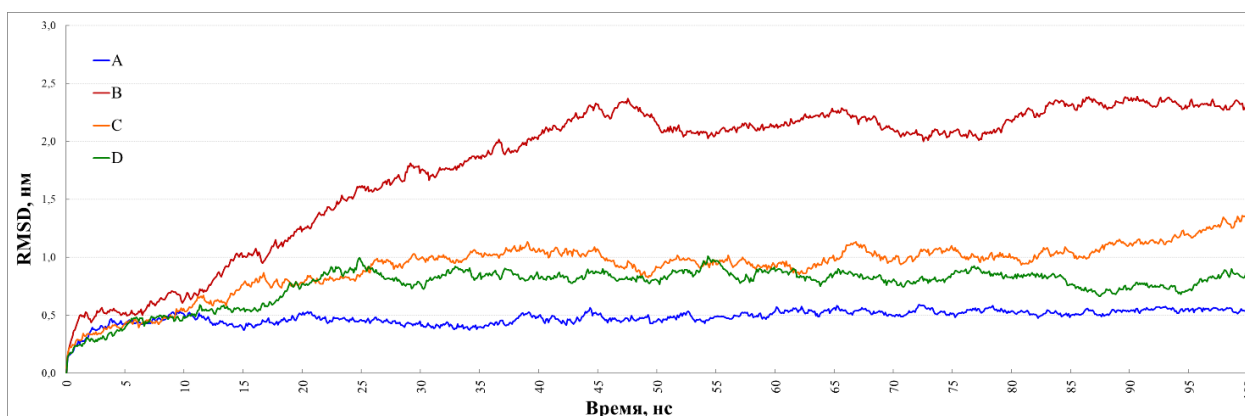


Рисунок 24 – RMSD атомов  $C_{\alpha}$  четырех структурно-различных гексамеров А–D ThL в кристаллизационном растворе.

Данные о динамике радиуса инерции (рис. 25) дополнительно подтверждают выводы: гексамер А показал наиболее стабильное значение  $R_g$  ( $\sim 4.1$ – $4.2$  нм), близкое к исходному, особенно после 70 нс симуляции. Для гексамеров С и D наблюдалось умеренное (менее выраженное, чем у типа В) увеличение объема.

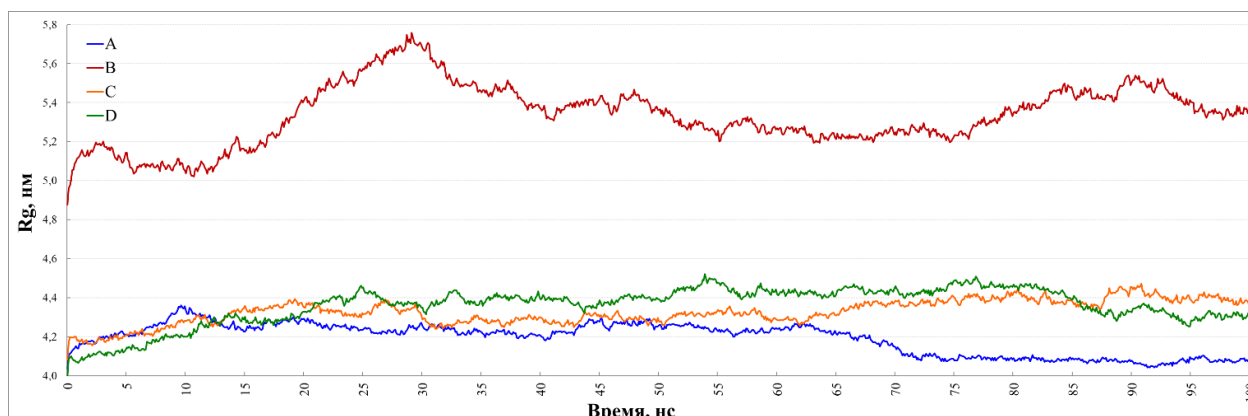


Рисунок 25 – Радиус инерции  $R_g$  четырех структурно-различных гексамеров А–D ThL в кристаллизационном растворе.

Следует отметить, что рассчитанные по результатам МД значения  $R_g$  для гексамера А ( $\sim 4.1$ – $4.2$  нм) заметно превышают экспериментальные данные, полученные методом МУРР ( $R_g \sim 2.3$  нм). Обнаруженное расхождение является ожидаемым, поскольку метод МУРР определяет усредненное значение радиуса инерции для всего ансамбля частиц в растворе, в котором доминируют мономеры (90 %), в то время как доля димеров (4–8 %) и гексамеров (1–3 %) существенно ниже. Экспериментальное значение радиуса инерции ( $\sim 2.3$  нм) лишь

незначительно превышает соответствующее значение для мономера ( $\sim 2.1$  нм), что обусловлено вкладом присутствующих в растворе олигомеров. Поэтому результаты МД, показывающие стабильность гексамера А с  $R_g \sim 4.2$  нм, не противоречат данным МУРР, а дополняют их, подтверждая принципиальную возможность существования таких устойчивых гексамеров в растворе.

Комплексный анализ характеристик RMSF, RMSD и  $R_g$  однозначно указывает на то, что **гексамер типа А обладает наибольшей стабильностью** среди четырех рассмотренных кандидатов, что позволяет идентифицировать его как **наиболее вероятный кластер-прекурсор** гексагонального кристалла ThL, образующийся в его кристаллизационном растворе. Последующее исследование температурной стабильности (Раздел 4.2) посвящено именно этому типу структуры.

#### 4.2. Влияние температуры на стабильность кластера-прекурсора кристалла термолизина

Термолизин является термостабильным ферментом, сохраняющим структурную и функциональную стабильность при высоких температурах (до 80–90°C) с оптимумом каталитической активности при 70°C [109]. Понимание поведения кластера-прекурсора его кристалла в широком температурном диапазоне представляет значительный интерес как для изучения механизмов термостабильности белковых олигомеров, так и для определения потенциальных условий кристаллизации. В связи с этим была детально исследована стабильность идентифицированного кластера-прекурсора (гексамера А) методом МД в диапазоне температур от 10 до 90°C.

Множественные независимые МД-моделирования (по 3 или 5 траекторий длительностью 100 нс для каждой температуры, всего 38 симуляций) были проведены при 10 температурах: в интервале 10–90°C с шагом 10°C и при 15°C. Протокол моделирования полностью соответствовал описанному в Разделе 2.1.

Стабильность оценивалась с помощью критерия RMSF C $\alpha$ -атомов (с расчетом среднего значения и стандартного отклонения по независимым траекториям для каждой температуры) и площади межмолекулярных контактов (схема вычислений приведена в Разделе 3.2) между мономерами в гексамере для оценки его структурной целостности.

Основным результатом исследования стала демонстрация высокой термостабильности и структурной целостности гексамера А во всем диапазоне температур 10–90°C. Наиболее показательным является результат анализа межмолекулярных контактов: полная диссоциация гексамера на олигомеры меньшего порядка или мономеры не наблюдалась ни в одной из 38 проведенных симуляций (рис. 26). Исчезновение площади межмолекулярного контакта, свидетельствующее о разрыве связей, было зафиксировано лишь кратковременно и только в одном из 38 моделирований (при 50°C), для двух межмолекулярных интерфейсов (между первым и вторым мономерами, а также между пятым и последним). Этот результат подчеркивает значительную прочность и стабильность межмолекулярных взаимодействий, формирующих гексамер А.

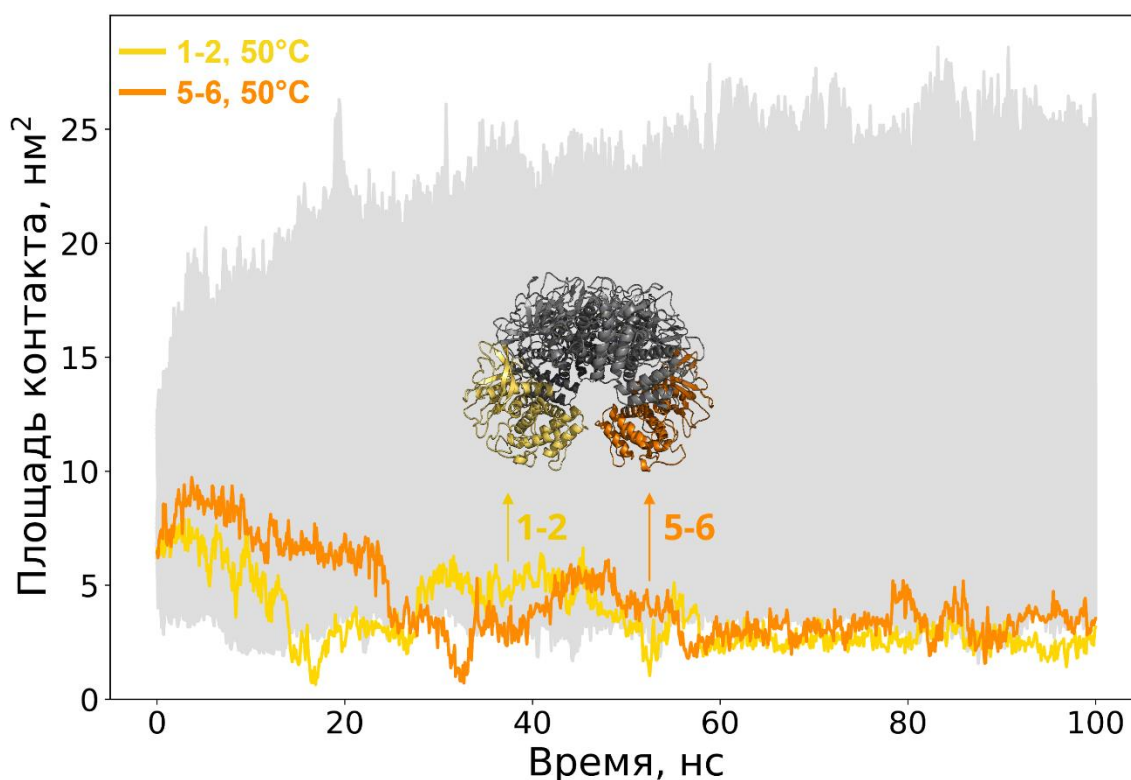


Рисунок 26 – Динамика площади всех (5) межмолекулярных контактов в гексамере А термолизина в диапазоне температур 10–90°C. Цветным кривым соответствует единственная траектория с кратковременным разрывом контактов: площадь = 0 нм<sup>2</sup> сразу между двумя парами молекул (между 1 и 2, между 5 и 6 мономерами) при температуре 50°C. В остальных моделированиях (показаны серым) площадь контакта не достигала 0 нм<sup>2</sup>, указывая на сохранение структурной целостности гексамера А в течение 100 нс.

Несмотря на общую тенденцию к увеличению структурной гибкости (росту RMSF) с повышением температуры (рис. 27), примечательно, что даже при 90°C усредненные по траекториям (трем для 10–40°C и 90°C; или пяти для 50–80°C) значения RMSF остаются умеренными (~0.33 нм), указывая на отсутствие признаков глобальной дестабилизации. Важно отметить, что аминокислотные остатки, формирующие активный центр ThL (выделены черным маркером на рис. 27), сохраняют низкие значения RMSF (относительно атомов в их окрестности) во всем диапазоне температур, указывая на стабильность функционально важных областей олигомера.

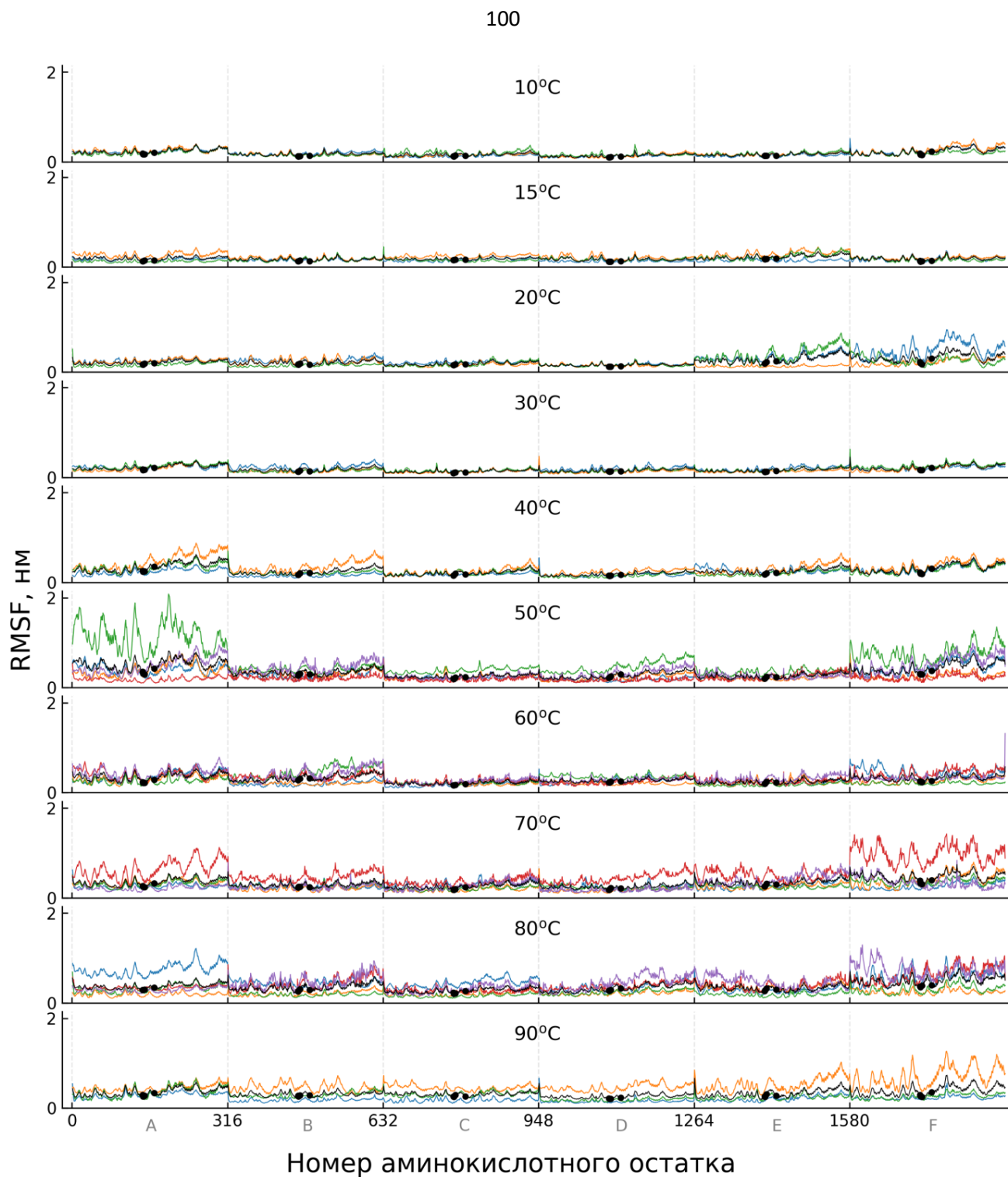


Рисунок 27 – Значения RMSF кластера-прекурсора (гексамера А) кристалла термолизина, смоделированного в кристаллизационном растворе при температурах в диапазоне 10–90°C. Каждая цветная кривая соответствует одному моделированию, черная – значениям RMSF, усредненным по всем независимым моделированиям (трем или пяти) при определенной температуре.

Анализ флуктуаций отдельных мономеров внутри гексамера показал (рис. 27), что наиболее подвижными являются мономеры 1 и 6, имеющие наименьшую площадь «соприкосновения» с соседними молекулами (рис. 28), тогда как мономеры 3 и 4, расположенные в центре, остаются наиболее стабильными.

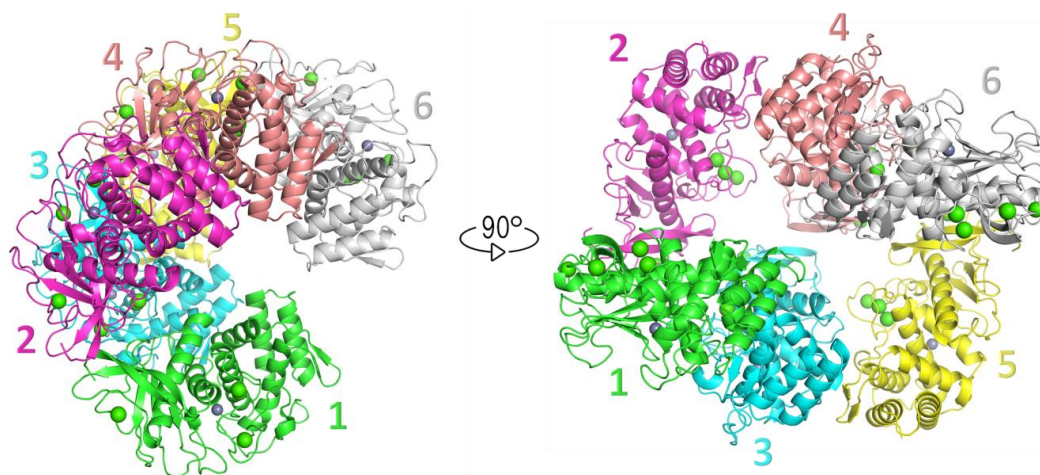


Рисунок 28 – Кластер-прекурсор (гексамер А) кристалла термолизина в двух проекциях. Цифрами 1–6 обозначены мономеры, составляющие гексамер. Зелеными сферами показаны ионы кальция, серыми – цинка.

Особый интерес представляет сравнительный анализ температурной зависимости стабильности кластеров-прекурсоров кристаллов двух исследованных термофильных протеаз – термолизина (гексамер А) и протеиназы К (димер Е). Стабильность оценивалась с помощью значений RMSF, усредненных по независимым траекториям для каждой температуры (рис. 20 и 29). Оба прекурсора демонстрируют высокую общую стабильность в широких диапазонах температур, подтверждая гипотезу о возможности формирования устойчивых предкристаллизационных ассоциатов термофильных белков при повышенных температурах. Однако характер зависимости RMSF от температуры обоих белков различается. Гексамер А ThL показывает относительно умеренные флуктуации во всем диапазоне 10–90°C, оставаясь структурно целостным и со средними значениями RMSF, не превышающими ~0.4 нм даже при 90°C. Димер Е ProtK, напротив, устойчив до 60°C, но демонстрирует резкий пик нестабильности (RMSF ~1.0 нм) в области 63–70°C, лишь частично стабилизируясь при 80°C. Таким

образом, прекурсор ThL оказывается значительно стабильнее прекурсора ProtK при температурах выше  $\sim 60^\circ\text{C}$ .

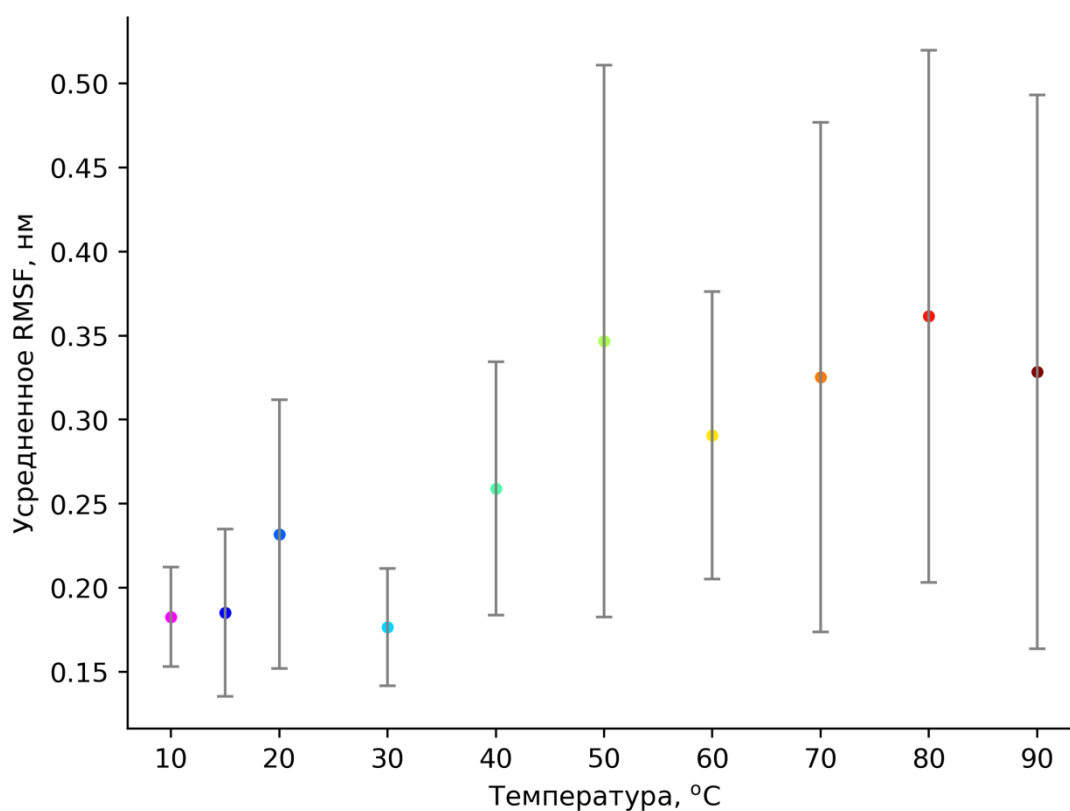


Рисунок 29 – Значения RMSF, усредненные сначала по  $C_\alpha$ -атомам кластера-прекурсора кристалла термолизина, а затем по независимым траекториям для каждой температуры (от 10 до  $90^\circ\text{C}$ ). Указанные погрешности составляют одно стандартное отклонение при усреднении RMSF по независимым моделированиям.

Полученные для ThL результаты позволяют уточнить гипотезу о связи стабильности прекурсора с функциональной активностью фермента, сформулированную в Главе 3 на основе данных для протеиназы K. Если для ProtK начало дестабилизации прекурсора ( $\sim 63^\circ\text{C}$ ) коррелировало с областью снижения ее активности (оптимум  $50\text{--}60^\circ\text{C}$ ), то для ThL прекурсор остается стабильным значительно выше температуры его оптимальной активности ( $\sim 70^\circ\text{C}$ ), вплоть до  $90^\circ\text{C}$ . Это указывает на то, что хотя динамика, связанная с пиком активности, может влиять на стабильность прекурсора (как в случае ProtK), верхний температурный предел существования прекурсора не всегда жестко связан с

температурой максимальной активности, а может определяться общей высокой термостабильностью белка и спецификой межмолекулярных контактов, как у ThL.

Оценка стандартных отклонений RMSF между независимыми траекториями (характеризующая воспроизводимость динамического поведения гексамера А при разных температурах, рис. 29) показала, что наименьший разброс значений RMSF вокруг среднего (т.е. наилучшая предсказуемость динамики прекурсора) в диапазоне температур 40–90°C наблюдается при 60°C. Это может указывать на снижение конформационной гетерогенности прекурсоров именно при данной температуре. Учитывая высокую общую стабильность гексамера А во всем диапазоне 10–90°C, можно предположить, что кристаллизация ThL потенциально возможна в этом широком интервале температур. Однако температура около 60°C представляется особо перспективной, так как сочетает значительную структурную стабильность прекурсора высокой степенью воспроизводимости его поведения, что может быть благоприятным фактором для формирования высокоупорядоченных кристаллов из-за снижения вероятности дефектов, вызванных хаотичными конформационными переходами.

#### 4.3. Заключение

В данной главе решены задачи по идентификации структуры и изучению температурной стабильности кластера-прекурсора кристалла термолизина. С помощью сравнительного МД-моделирования **установлена пространственная структура прекурсора (гексамер типа А)** среди четырех возможных кандидатов. Последующее МД-моделирование этого гексамера в широком диапазоне температур (10–90°C) продемонстрировало его **высокую структурную целостность и термостабильность**: диссоциация не наблюдалась ни в одной из 38 симуляций, а потеря межмолекулярного контакта была крайне редким и кратковременным событием.

Практическим следствием **установленной высокой термостабильности гексамера А в диапазоне 10–90°C** является вывод о **принципиальной возможности проведения кристаллизации термолизина в этом чрезвычайно широком температурном интервале**. Это существенно расширяет поле для экспериментального скрининга и оптимизации условий кристаллизации, включая нетрадиционно высокие температуры. При этом температура около **60°C**, характеризующаяся наибольшей предсказуемостью поведения прекурсора и близостью к пику каталитической активности, может представлять особый интерес для получения наиболее качественных и упорядоченных кристаллов.

Результаты, представленные в Главе 4 опубликованы в работах [A4, A5].

## ВЫВОДЫ И ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

1. **Предложен и апробирован подход к исследованию кристаллизационных растворов белков**, основанный на сочетании «упрощенной» крупнозернистой и традиционной атомистической МД, позволяющий эффективно изучать различные этапы белковой кристаллизации за счет оптимального выбора силового поля (уровня детализации модели).
2. МД-моделирование двух типов октамеров (А и В) лизоцима с использованием атомистического (Amber99sb-ildn) и крупнозернистого (MARTINI) силовых полей показало, что октамер А стабильнее октамера В независимо от выбора поля. Согласованность полученных результатов **подтверждает корректность применения поля MARTINI** для исследования кристаллизационных растворов глобулярных белков.
3. Установлено, что увеличение объема ячейки при моделировании МД олигомеров белков в кристаллизационных растворах не приводит к существенным изменениям в их поведении. **Определены размеры ячейки моделирования, достаточные** для исследования кристаллизационных растворов белков с помощью МД, – при расстоянии между атомами белка и гранью ячейки, составляющем 1 радиус отсечки невалентных взаимодействий (1 нм).
4. Сравнительный анализ динамики шести возможных типов димеров протеиназы К в кристаллизационном растворе выявил, что максимальную стабильность проявляет димер Е. Это позволяет **идентифицировать димер Е** как наиболее вероятную **единицу роста (кластер-прекурсор) кристалла протеиназы К**.
5. Обнаружена высокая **стабильность димера Е** (кластера-прекурсора кристалла) протеиназы К в температурном диапазоне **20–60°C**. Выдвинута обоснованная **гипотеза о возможности кристаллизации** данного белка в указанном интервале температур.
6. Предсказанная на основе МД-результатов перспективность проведения кристаллизации протеиназы К при нетипично высоких для большинства

белков температурах (50–60 °С) **подтверждена экспериментально**: кристаллы протеиназы К были успешно получены после инкубации при температуре 60°С. Результаты указывают на **адекватность использованного МД-подхода** для оценки стабильности олигомеров и его **предсказательную способность** в определении перспективных температур белковой кристаллизации.

7. Сравнительный анализ динамики четырех возможных типов гексамеров термолизина в кристаллизационном растворе выявил, что максимальную стабильность проявляет гексамер А. Это позволяет **идентифицировать гексамер А** как наиболее вероятную **единицу роста (кластер-прекурсор) кристалла термолизина**.
8. Обнаружена высокая **стабильность гексамера А** (кластера-прекурсора кристалла) термолизина в температурном диапазоне **10–90°С**. Выдвинута обоснованная **гипотеза о возможности кристаллизации** данного белка при температуре 60°С.

## СПИСОК ЛИТЕРАТУРЫ

1. Denesyuk A. I. Structural Catalytic Core in Subtilisin-like Proteins and Its Comparison to Trypsin-like Serine Proteases and  $\alpha/\beta$ -Hydrolases / A. I. Denesyuk, K. Denessiouk, M. S. Johnson, V. N. Uversky // *International Journal of Molecular Sciences*. — 2024. — Vol. 25. — № 22. — P. 11858. — DOI: 10.3390/ijms252211858.
2. Berman H. M. The Protein Data Bank / H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne // *Nucleic Acids Research*. — 2000. — Vol. 28. — № 1. — P. 235–242. — DOI: 10.1093/nar/28.1.235.
3. Laganowsky A. Membrane Proteins Bind Lipids Selectively to Modulate Their Structure and Function / A. Laganowsky, E. Reading, T. M. Allison, M. B. Ulmschneider, M. T. Degiacomi, A. J. Baldwin, C. V. Robinson // *Nature*. — 2014. — Vol. 510. — № 7503. — P. 172–175. — DOI: 10.1038/nature13419.
4. Terwilliger T. C. AlphaFold Predictions Are Valuable Hypotheses and Accelerate but Do Not Replace Experimental Structure Determination / T. C. Terwilliger, D. Liebschner, T. I. Croll, C. J. Williams, A. J. McCoy, B. K. Poon, P. V. Afonine [et al.] // *Nature Methods*. — 2024. — Vol. 21. — № 1. — P. 110–116. — DOI: 10.1038/s41592-023-02087-4.
5. Chen R. Q. An Ignored Variable: Solution Preparation Temperature in Protein Crystallization / R. Q. Chen, L. Q. Lu, Q. D. Cheng, L. B. Ao, C. Y. Zhang, H. Hou, Y. M. Liu, D. W. Li, D. C. Yin // *Scientific Reports*. — 2015. — Vol. 5. — P. 7797. — DOI: 10.1038/srep07797.
6. Chayen N. E. Turning Protein Crystallisation from an Art into a Science / N. E. Chayen // *Current Opinion in Structural Biology*. — 2004. — Vol. 14. — № 5. — P. 577–583. — DOI: 10.1016/j.sbi.2004.08.002.
7. Kruglikov A. Proteins from Thermophilic *Thermus Thermophilus* Often Do Not Fold Correctly in a Mesophilic Expression System Such as *Escherichia Coli* / A. Kruglikov, Y. Wei, X. Xia // *ACS Omega*. — 2022. — Vol. 7. — № 42. — P. 37797–37806. — DOI: 10.1021/acsomega.2c04786.
8. Kovalchuk M. V. Investigation of the Initial Crystallization Stage in Lysozyme Solutions by Small-Angle X-Ray Scattering / M. V. Kovalchuk, A. E. Blagov, Y. A. Dyakova, A. Y. Gruzinov, M. A. Marchenkova, G. S. Peters, Yu. V. Pisarevsky, V. I. Timofeev, V. V. Volkov // *Crystal Growth & Design*. — 2016. — Vol. 16. — № 4. — P. 1792–1797. — DOI: 10.1021/acs.cgd.5b01662.

9. Марченкова М. А. In situ-исследования состояния молекул лизоцима на стадии начала процесса кристаллизации методом малоуглового рентгеновского рассеяния / М. А. Марченкова, В. В. Волков, А. Е. Благов, Ю. А. Дьякова, К. Б. Ильина, Е. Ю. Терещенко, В. И. Тимофеев, Ю. В. Писаревский, М. В. Ковальчук // Кристаллография. — 2016. — Т. 61. — № 1. — С. 10–15. — DOI: 10.7868/S0023476116010148.
10. Boikova A. S. Octamer Formation in Lysozyme Solutions at the Initial Crystallization Stage Detected by Small-Angle Neutron Scattering / A. S. Boikova, Y. A. Dyakova, K. B. Ilina, P. V. Konarev, A. E. Kryukova, A. I. Kuklin, M. A. Marchenkova, B. V. Nabatov, A. E. Blagov, Y. V. Pisarevsky [et al.] // Acta Crystallographica Section D: Structural Biology. — 2017. — Vol. 73. — № 7. — P. 591–599. — DOI: 10.1107/S2059798317007422.
11. Кордонская Ю. В. Исследование поведения олигомеров белка лизоцима в растворах методом молекулярной динамики / Ю. В. Кордонская, В. И. Тимофеев, Ю. А. Дьякова, М. А. Марченкова, Ю. В. Писаревский, Д. Д. Подшивалов, М. В. Ковальчук // Кристаллография. — 2018. — Т. 63. — № 6. — С. 902–905. — DOI: 10.1134/S002347611806019X.
12. Dill K. A. Principles of Protein Folding — A Perspective from Simple Exact Models / K. A. Dill, S. Bromberg, K. Yue, H. S. Chan, K. M. Ftebig, D. P. Yee, P. D. Thomas // Protein Science. — 1995. — Vol. 4. — № 4. — P. 561–602. — DOI: 10.1002/pro.5560040401.
13. Johnson M. S. A Structural Basis for Sequence Comparisons: An Evaluation of Scoring Methodologies / M. S. Johnson, J. P. Overington // Journal of Molecular Biology. — 1993. — Vol. 233. — № 4. — P. 716–738. — DOI: 10.1006/jmbi.1993.1548.
14. Rohl C. A. Protein Structure Prediction Using Rosetta / C. A. Rohl, C. E. M. Strauss, K. M. S. Misura, D. Baker // Methods in Enzymology. — 2004. — Vol. 383. — P. 66–93. — DOI: 10.1016/S0076-6879(04)83004-0.
15. Barducci A. Metadynamics / A. Barducci, M. Bonomi, M. Parrinello // Wiley Interdisciplinary Reviews: Computational Molecular Science. — 2011. — Vol. 1. — № 5. — P. 826–843. — DOI: 10.1002/wcms.31.
16. Zhang Y. The Protein Structure Prediction Problem Could Be Solved Using the Current PDB Library / Y. Zhang, J. Skolnick // Proceedings of the National Academy of Sciences of the United States of America. — 2005. — Vol. 102. — № 4. — P. 1029–1034. — DOI: 10.1073/pnas.0407152101.

17. Jumper J. Highly Accurate Protein Structure Prediction with AlphaFold / J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko [et al.] // *Nature*. — 2021. — Vol. 596. — № 7873. — P. 583–589. — DOI: 10.1038/s41586-021-03819-2.
18. Evans R. Protein Complex Prediction with AlphaFold-Multimer / R. Evans, M. O’Neill, A. Pritzel, N. Antropova, A. Senior, T. Green, A. Žídek, R. Bates, S. Blackwell, J. Yim [et al.] // *bioRxiv*. — 2021. — DOI: 10.1101/2021.10.04.463034.
19. Hollingsworth S. A. Molecular Dynamics Simulation for All / S. A. Hollingsworth, R. O. Dror // *Neuron*. — 2018. — Vol. 99. — № 6. — P. 1129–1143. — DOI: 10.1016/j.neuron.2018.08.011.
20. Terwilliger T. C. Improved AlphaFold Modeling with Implicit Experimental Information / T. C. Terwilliger, B. K. Poon, P. V. Afonine, C. J. Schlicksup, T. I. Croll, C. Millán, J. S. Richardson, R. J. Read, P. D. Adams // *Nature Methods*. — 2022. — Vol. 19. — № 11. — P. 1376–1382. — DOI: 10.1038/s41592-022-01645-6.
21. Shaw D. E. Atomic-Level Characterization of the Structural Dynamics of Proteins / D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. Shan [et al.] // *Science*. — 2010. — Vol. 330. — № 6002. — P. 341–346. — DOI: 10.1126/science.1187409.
22. Wayment-Steele H. K. Predicting Multiple Conformations via Sequence Clustering and AlphaFold2 / H. K. Wayment-Steele, A. Ojoawo, R. Otten, J. M. Apitz, W. Pitsawong, M. Hömberger, S. Ovchinnikov, L. Colwell, D. Kern // *Nature*. — 2024. — Vol. 625. — № 7996. — P. 832–839. — DOI: 10.1038/s41586-023-06832-9.
23. Mirdita M. ColabFold: Making Protein Folding Accessible to All / M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, M. Steinegger // *Nature Methods*. — 2022. — Vol. 19. — № 6. — P. 679–682. — DOI: 10.1038/s41592-022-01488-1.
24. Cooper D. R. X-Ray Crystallography: Assessment and Validation of Protein-Small Molecule Complexes for Drug Discovery / D. R. Cooper, P. J. Porebski, M. Chruszcz, W. Minor // *Expert Opinion on Drug Discovery*. — 2011. — Vol. 6. — № 8. — P. 771–782. — DOI: 10.1517/17460441.2011.585154.
25. Куранова И. П. Кристаллы для изучения белковых структур / И. П. Куранова, М. В. Ковальчук // *Природа*. — 2014. — № 3. — С. 12–21.
26. McPherson A. Introduction to Protein Crystallization / A. McPherson, J. A. Gavira // *Acta Crystallographica Section F: Structural Biology Communications*. — 2014. — Vol. 70. — № 1. — P. 2–20. — DOI: 10.1107/S2053230X13033141.

27. Бландел Т. Кристаллография белка / Т. Бландел, Л. Джонсон. — М. : Мир, 1979. — 584 с.
28. Ducruix A. Protein Interactions as Seen by Solution X-Ray Scattering Prior to Crystallogenesis / A. Ducruix, J. P. Guilloteau, M. Riès-Kautt, A. Tardieu // *Journal of Crystal Growth*. — 1996. — Vol. 168. — № 1–4. — P. 28–39. — DOI: 10.1016/0022-0248(96)00359-4.
29. Yamanaka M. Optimization of Salt Concentration in PEG-Based Crystallization Solutions / M. Yamanaka, K. Inaka, N. Furubayashi, M. Matsushima, S. Takahashi, H. Tanaka, S. Sano, M. Sato, T. Kobayashi, T. Tanaka // *Journal of Synchrotron Radiation*. — 2011. — Vol. 18. — № 1. — P. 84–87. — DOI: 10.1107/S0909049510035995.
30. Majeed S. Enhancing Protein Crystallization through Precipitant Synergy / S. Majeed, G. Ofek, A. Belachew, C. C. Huang, T. Zhou, P. D. Kwong // *Structure*. — 2003. — Vol. 11. — № 9. — P. 1061–1070. — DOI: 10.1016/S0969-2126(03)00185-0.
31. Kordonskaya Y. V. Precipitant Ions Influence on Lysozyme Oligomers Stability Investigated by Molecular Dynamics Simulation at Different Temperatures / Y. V. Kordonskaya, M. A. Marchenkova, V. I. Timofeev, Y. A. Dyakova, Y. V. Pisarevsky, M. V. Kovalchuk // *Journal of Biomolecular Structure and Dynamics*. — 2021. — Vol. 39. — № 18. — P. 7223–7230. — DOI: 10.1080/07391102.2020.1803138.
32. Feher G. Nucleation and Growth of Protein Crystals: General Principles and Assays / G. Feher, Z. Kam // *Methods in Enzymology*. — 1985. — Vol. 114. — P. 77–112. — DOI: 10.1016/0076-6879(85)14006-1.
33. Куранова И. П. Кристаллизация белков на Земле и в невесомости / И. П. Куранова // *Поверхность. Рентгеновские, синхротронные и нейтронные исследования*. — 2004. — № 6. — С. 4–12.
34. Rosenberger F. Inorganic and Protein Crystal Growth-Similarities and Differences / F. Rosenberger // *Journal of Crystal Growth*. — 1986. — Vol. 76. — № 3. — P. 618–636. — DOI: 10.1016/0022-0248(86)90179-X.
35. Zhu Z. High-Temperature Crystallization Method for the GH57 Family Hyperthermophilic Amylopullulanase from *Aquifex Aeolicus* / Z. Zhu, W. Wang, L. Huang, Q. Xu, H. Zhou, M. Li, F. Yu, Q. Wang // *Crystal Growth & Design*. — 2024. — Vol. 24. — № 24. — P. 10350-10357. — DOI: 10.1021/acs.cgd.4c01306.

36. Тимофеев В. И. Кристаллизация белков как центральное звено получения трёхмерной структуры белка / В. И. Тимофеев, И. П. Куранова // *Биофизика и медицинская техника*. — 2015. — Т. 1. — С. 334–353.
37. McCammon J. A. Dynamics of Folded Proteins / J. A. McCammon, B. R. Gelin, M. Karplus // *Nature*. — 1977. — Vol. 267. — № 5612. — P. 585–590. — DOI: 10.1038/267585a0.
38. Levitt M. Refinement of Protein Conformations Using a Macromolecular Energy Minimization Procedure / M. Levitt, S. Lifson // *Journal of Molecular Biology*. — 1969. — Vol. 46. — № 2. — P. 269–279. — DOI: 10.1016/0022-2836(69)90421-5.
39. Lifson S. Consistent Force Field for Calculations of Conformations, Vibrational Spectra, and Enthalpies of Cycloalkane and N-Alkane Molecules / S. Lifson, A. Warshel // *The Journal of Chemical Physics*. — 1968. — Vol. 49. — № 11. — P. 5116–5129. — DOI: 10.1063/1.1670007.
40. Metropolis N. The Monte Carlo Method / N. Metropolis, S. Ulam // *Journal of the American Statistical Association*. — 1949. — Vol. 44. — № 247. — P. 335–341. — DOI: 10.1080/01621459.1949.10483310.
41. Abraham M. J. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers / M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl // *SoftwareX*. — 2015. — Vol. 1–2. — P. 19–25. — DOI: 10.1016/j.softx.2015.06.001.
42. Phillips J. C. Scalable Molecular Dynamics with NAMD / J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, K. Schulten // *Journal of Computational Chemistry*. — 2005. — Vol. 26. — № 16. — P. 1781–1802. — DOI: 10.1002/jcc.20289.
43. Case D. A. Amber 2022 / D. A. Case, H. M. Aktulga, K. Belfon, I. Y. Ben-Shalom, J. T. Berryman, S. R. Brozell, D. S. Cerutti, T. E. Cheatham III, G. A. Cisneros, V. W. D. Cruzeiro [et al.]. — University of California, San Francisco, 2022.
44. Brooks B. R. CHARMM: The Biomolecular Simulation Program / B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch [et al.] // *Journal of Computational Chemistry*. — 2009. — Vol. 30. — № 10. — P. 1545–1614. — DOI: 10.1002/jcc.21287.
45. Bowers K. J. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters / K. J. Bowers, D. E. Chow, H. Xu, R. O. Dror, M. P. Eastwood, B. A. Gregersen, J. L. Klepeis, I. Kolossvary, M. A. Moraes, F. D. Sacerdoti [et al.] //

- Proceedings of the 2006 ACM/IEEE Conference on Supercomputing. — 2006. — P. 43. — DOI: 10.1109/SC.2006.54.
46. Eastman P. OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics / P. Eastman, J. Swails, J. D. Chodera, R. T. McGibbon, Y. Zhao, K. A. Beauchamp, L. P. Wang, A. C. Simmonett, M. P. Harrigan, C. D. Stern [et al.] // PLOS Computational Biology. — 2017. — Vol. 13. — № 7. — P. e1005659. — DOI: 10.1371/journal.pcbi.1005659.
47. Salomon-Ferrer R. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald / R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand, R. C. Walker // Journal of Chemical Theory and Computation. — 2013. — Vol. 9. — № 9. — P. 3878–3888. — DOI: 10.1021/ct400314y.
48. Stone J. E. Evaluation of Emerging Energy-Efficient Heterogeneous Computing Platforms for Biomolecular and Cellular Simulation Workloads / J. E. Stone, M. J. Hallock, J. C. Phillips, J. R. Peterson, Z. Luthey-Schulten, K. Schulten // 2016 IEEE 30th International Parallel and Distributed Processing Symposium (IPDPS). — 2016. — P. 89–100. — DOI: 10.1109/IPDPSW.2016.130.
49. Lindorff-Larsen K. Systematic Validation of Protein Force Fields against Experimental Data / K. Lindorff-Larsen, P. Maragakis, S. Piana, M. P. Eastwood, R. O. Dror, D. E. Shaw // PLOS One. — 2012. — Vol. 7. — № 2. — P. e32131. — DOI: 10.1371/journal.pone.0032131.
50. Goh G. B. Constant pH Molecular Dynamics of Proteins in Explicit Solvent with Proton Tautomerism / G. B. Goh, B. S. Hulbert, H. Zhou, C. L. Brooks // Proteins: Structure, Function, and Bioinformatics. — 2014. — Vol. 82. — № 7. — P. 1319–1331. — DOI: 10.1002/prot.24499.
51. Senn H. M. QM/MM Methods for Biomolecular Systems / H. M. Senn, W. Thiel // Angewandte Chemie International Edition. — 2009. — Vol. 48. — № 7. — P. 1198–1229. — DOI: 10.1002/anie.200802019.
52. Wang J. Development and Testing of a General Amber Force Field / J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case // Journal of Computational Chemistry. — 2004. — Vol. 25. — № 9. — P. 1157–1174. — DOI: 10.1002/jcc.20035.
53. Huang J. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins / J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de

Groot, H. Grubmüller, A. D. MacKerell // *Nature Methods*. — 2016. — Vol. 14. — № 1. — P. 71–73. — DOI: 10.1038/nmeth.4067.

54. Harder E. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins / E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyan, M. K. Dahlgren, J. L. Knight [et al.] // *Journal of Chemical Theory and Computation*. — 2016. — Vol. 12. — № 1. — P. 281–296. — DOI: 10.1021/acs.jctc.5b00864.

55. Robustelli P. Developing a Molecular Dynamics Force Field for Both Folded and Disordered Protein States / P. Robustelli, S. Piana, D. E. Shaw // *Proceedings of the National Academy of Sciences of the United States of America*. — 2018. — Vol. 115. — № 21. — P. 4758–4766. — DOI: 10.1073/pnas.1800690115.

56. Marrink S. J. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations / S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, A. H. De Vries // *The Journal of Physical Chemistry B*. — 2007. — Vol. 111. — № 27. — P. 7812–7824. — DOI: 10.1021/jp071097f.

57. Souza P. C. T. Martini 3: A General Purpose Force Field for Coarse-Grained Molecular Dynamics / P. C. T. Souza, R. Alessandri, J. Barnoud, S. Thallmair, I. Faustino, F. Grünewald, I. Patmanidis, H. Abdizadeh, B. M. H. Bruininks, T. A. Wassenaar [et al.] // *Nature Methods*. — 2021. — Vol. 18. — № 4. — P. 382–388. — DOI: 10.1038/s41592-021-01098-3.

58. Schwantes C. R. Perspective: Markov Models for Long-Timescale Biomolecular Dynamics / C. R. Schwantes, R. T. McGibbon, V. S. Pande // *The Journal of Chemical Physics*. — 2014. — Vol. 141. — № 9. — P. 090901. — DOI: 10.1063/1.4895044.

59. Schlitter J. Targeted Molecular Dynamics: A New Approach for Searching Pathways of Conformational Transitions / J. Schlitter, M. Engels, P. Krüger // *Journal of Molecular Graphics*. — 1994. — Vol. 12. — № 2. — P. 84–89. — DOI: 10.1016/0263-7855(94)80072-3.

60. Laio A. Metadynamics: A Method to Simulate Rare Events and Reconstruct the Free Energy in Biophysics, Chemistry and Material Science / A. Laio, F. L. Gervasio // *Reports on Progress in Physics*. — 2008. — Vol. 71. — № 12. — P. 126601. — DOI: 10.1088/0034-4885/71/12/126601.

61. Maragliano L. A Temperature Accelerated Method for Sampling Free Energy and Determining Reaction Pathways in Rare Events Simulations / L. Maragliano, E.

- Vanden-Eijnden // *Chemical Physics Letters*. — 2006. — Vol. 426. — № 1–3. — P. 168–175. — DOI: 10.1016/j.cplett.2006.05.062.
62. Hamelberg D. Accelerated Molecular Dynamics: A Promising and Efficient Simulation Method for Biomolecules / D. Hamelberg, J. Mongan, J. A. McCammon // *The Journal of Chemical Physics*. — 2004. — Vol. 120. — № 24. — P. 11919–11929. — DOI: 10.1063/1.1755656.
63. Marrink S. J. Perspective on the Martini Model / S. J. Marrink, D. P. Tieleman // *Chemical Society Reviews*. — 2013. — Vol. 42. — № 16. — P. 6801–6822. — DOI: 10.1039/C3CS60093A.
64. Cerutti D. S. Molecular Dynamics Simulations of Macromolecular Crystals / D. S. Cerutti, D. A. Case // *Wiley Interdisciplinary Reviews: Computational Molecular Science*. — 2019. — Vol. 9. — № 4. — P. e1402. — DOI: 10.1002/wcms.1402.
65. Cerutti D. S. Simulations of a Protein Crystal: Explicit Treatment of Crystallization Conditions Links Theory and Experiment in the Streptavidin-Biotin Complex / D. S. Cerutti, I. Le Trong, R. E. Stenkamp, T. P. Lybrand // *Biochemistry*. — 2008. — Vol. 47. — № 46. — P. 12065–12077. — DOI: 10.1021/bi800894u.
66. Cerutti D. S. Dynamics of the Streptavidin - Biotin Complex in Solution and in Its Crystal Lattice: Distinct Behavior Revealed by Molecular Simulations / D. S. Cerutti, I. Le Trong, R. E. Stenkamp, T. P. Lybrand // *The Journal of Physical Chemistry B*. — 2009. — Vol. 113. — № 19. — P. 6971–6985. — DOI: 10.1021/jp9010372.
67. Cerutti D. S. Simulations of a Protein Crystal with a High Resolution X-Ray Structure: Evaluation of Force Fields and Water Models / D. S. Cerutti, P. L. Freddolino, R. E. Duke, D. A. Case // *The Journal of Physical Chemistry B*. — 2010. — Vol. 114. — № 40. — P. 12811–12824. — DOI: 10.1021/jp105813j.
68. Kuzmanic A. Dependence of Protein Crystal Stability on Residue Charge States and Ion Content of Crystal Solvent / A. Kuzmanic, B. Zagrovic // *Biophysical Journal*. — 2014. — Vol. 106. — № 3. — P. 677–686. — DOI: 10.1016/j.bpj.2013.12.022.
69. Klyshko E. Functional Protein Dynamics in a Crystal / E. Klyshko, J. S. H. Kim, L. McGough, V. Valeeva, E. Lee, R. Ranganathan, S. Rauscher // *Nature Communications*. — 2024. — Vol. 15. — № 1. — P. 3442. — DOI: 10.1038/s41467-024-47473-4.
70. Lopez A. J. Quantifying Protein–Protein Interactions in Molecular Simulations / A. J. Lopez, P. K. Quoika, M. Linke, G. Hummer, J. R. Kofinger // *The Journal of*

Physical Chemistry B. — 2020. — Vol. 124. — № 23. — P. 4673–4685. — DOI: 10.1021/acs.jpcc.9b11802.

71. Finney A. R. Molecular Simulation Approaches to Study Crystal Nucleation from Solutions: Theoretical Considerations and Computational Challenges / A. R. Finney, M. Salvalaglio // Wiley Interdisciplinary Reviews: Computational Molecular Science. — 2024. — Vol. 14. — № 1. — P. e1697. — DOI: 10.1002/wcms.1697.

72. Torrie G. M. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling / G. M. Torrie, J. P. Valleau // Journal of Computational Physics. — 1977. — Vol. 23. — № 2. — P. 187–199. — DOI: 10.1016/0021-9991(77)90121-8.

73. Allen R. J. Forward Flux Sampling for Rare Event Simulations / R. J. Allen, C. Valeriani, Rein Ten Wolde P. // Journal of Physics: Condensed Matter. — 2009. — Vol. 21. — № 46. — P. 463102. — DOI: 10.1088/0953-8984/21/46/463102.

74. Taudt A. Simulation of Protein Association: Kinetic Pathways towards Crystal Contacts / A. Taudt, A. Arnold, J. Pleiss // Physical Review E. — 2015. — Vol. 91. — № 3. — P. 033311. — DOI: 10.1103/PhysRevE.91.033311.

75. Triandafillidis D. P. The T2 Structure of Polycrystalline Cubic Human Insulin / D. P. Triandafillidis, F. Karavassili, M. Spiliopoulou, A. Valmas, M. Athanasiadou, G. Nikolaras, S. Fili, P. Kontou, M. W. Bowler, C. T. Chasapis [et al.] // Acta Crystallographica Section D: Structural Biology. — 2023. — Vol. 79. — № 5. — P. 374–386. — DOI: 10.1107/S2059798323001328.

76. L’Vov P. E. Two-Step Mechanism of Macromolecular Nucleation and Crystallization: Field Theory and Simulations / P. E. L’Vov, A. R. Umantsev // Crystal Growth & Design. — 2021. — Vol. 21. — № 1. — P. 366–382. — DOI: 10.1021/acs.cgd.0c01224.

77. Krissinel E. Inference of Macromolecular Assemblies from Crystalline State / E. Krissinel, K. Henrick // Journal of Molecular Biology. — 2007. — Vol. 372. — № 3. — P. 774–797. — DOI: 10.1016/j.jmb.2007.05.022.

78. Schrödinger L. The PyMOL Molecular Graphics System, Version 1.8 [Электронный ресурс] / Schrödinger, LLC. — Thomas Holder. — 2015. — <https://pymol.org> (дата обращения: 16.09.2023).

79. Kovalchuk M. V. Structural Characteristics of Lysozyme Langmuir Layers Grown on a Liquid Surface from an Oligomeric Mixture Formed during the Early Stages of Lysozyme Crystallization / M. V. Kovalchuk, A. S. Boikova, Y. A. Dyakova, K. B.

Irina, P. V. Konarev, M. A. Marchenkova, Y. V. Pisarevskiy, P. A. Prosekov, A. V. Rogachev, A. Y. Seregin // *Thin Solid Films*. — 2019. — Vol. 677. — P. 13–21. — DOI: 10.1016/j.tsf.2019.02.051.

80. Marchenkova M. A. In Situ Study of the State of Lysozyme Molecules at the Very Early Stage of the Crystallization Process by Small-Angle X-Ray Scattering / M. A. Marchenkova, V. V. Volkov, A. E. Blagov, Y. A. Dyakova, K. B. Ilina, E. Y. Tereschenko, V. I. Timofeev, Y. V. Pisarevsky, M. V. Kovalchuk // *Crystallography Reports*. — 2016. — Vol. 61. — № 1. — P. 5–10. — DOI: 10.1134/S1063774516010144.

81. Dolinsky T. J. PDB2PQR: An Automated Pipeline for the Setup of Poisson-Boltzmann Electrostatics Calculations / T. J. Dolinsky, J. E. Nielsen, J. A. McCammon, N. A. Baker // *Nucleic Acids Research*. — 2004. — Vol. 32. — Web Server issue. — P. W665–W667. — DOI: 10.1093/nar/gkh381.

82. Lindorff-Larsen K. Improved Side-Chain Torsion Potentials for the Amber Ff99SB Protein Force Field / K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, D. E. Shaw // *Proteins: Structure, Function, and Bioinformatics*. — 2010. — Vol. 78. — № 8. — P. 1950–1958. — DOI: 10.1002/prot.22711.

83. Essmann U. A Smooth Particle Mesh Ewald Method / U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen // *The Journal of Chemical Physics*. — 1995. — Vol. 103. — № 19. — P. 8577–8592. — DOI: 10.1063/1.470117.

84. Hess B. LINCS: A Linear Constraint Solver for Molecular Simulations / B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije // *Journal of Computational Chemistry*. — 1997. — Vol. 18. — № 12. — P. 1463–1472. — DOI: 10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.

85. Horn H. W. Development of an Improved Four-Site Water Model for Biomolecular Simulations: TIP4P-Ew / H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura, T. Head-Gordon // *The Journal of Chemical Physics*. — 2004. — Vol. 120. — № 20. — P. 9665–9678. — DOI: 10.1063/1.1683075.

86. Dimitropoulos D. Using MSDchem to Search the PDB Ligand Dictionary / D. Dimitropoulos, J. Ionides, K. Henrick // *Current Protocols in Bioinformatics*. — 2006. — Vol. 15. — № 1. — P. 14.3.1-14.3.21. — DOI: 10.1002/0471250953.bi1403s15.

87. Michaud-Agrawal N. MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations / N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, O.

Beckstein // *Journal of Computational Chemistry*. — 2011. — Vol. 32. — № 10. — P. 2319–2327. — DOI: 10.1002/jcc.21787.

88. Gowers R. J. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations / R. J. Gowers, M. Linke, J. Barnoud, T. J. E. Reddy, M. N. Melo, S. L. Seyler, J. Domá nski, D. L. Dotson, S. Buchoux, I. M. Kenney [et al.] // *Proceedings of the 15th Python in Science Conference*. — 2016. — P. 98–105. — DOI: 10.25080/Majora-629e541a-00e.

89. Sousa Da Silva A. W. ACPYPE - AnteChamber PYthon Parser InterfacE / A. W. Sousa Da Silva, W. F. Vranken // *BMC Research Notes*. — 2012. — Vol. 5. — P. 367. — DOI: 10.1186/1756-0500-5-367.

90. Berendsen H. J. C. Molecular Dynamics with Coupling to an External Bath / H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. Dinola, J. R. Haak // *The Journal of Chemical Physics*. — 1984. — Vol. 81. — № 8. — P. 3684–3690. — DOI: 10.1063/1.448118.

91. Parrinello M. Strain Fluctuations and Elastic Constants / M. Parrinello, A. Rahman // *The Journal of Chemical Physics*. — 1982. — Vol. 76. — № 5. — P. 2662–2666. — DOI: 10.1063/1.443248.

92. Van Gunsteren W. F. A Leap-Frog Algorithm for Stochastic Dynamics / W. F. Van Gunsteren, H. J. C. Berendsen // *Molecular Simulation*. — 1988. — Vol. 1. — № 3. — P. 173–185. — DOI: 10.1080/08927028808080941.

93. El Hage K. Valid Molecular Dynamics Simulations of Human Hemoglobin Require a Surprisingly Large Box Size / K. El Hage, F. Hédin, P. K. Gupta, M. Meuwly, M. Karplus // *eLife*. — 2018. — Vol. 7. — P. e35560. — DOI: 10.7554/eLife.35560.

94. Кордонская Ю. В. Влияние размера ячейки моделирования и концентрации ионов осадителя на поведение димера тетрагонального лизоцима / Ю. В. Кордонская, В. И. Тимофеев, Ю. А. Дьякова, М. А. Марченкова, Ю. В. Писаревский, М. В. Ковальчук // *Кристаллография*. — 2021. — Т. 66. — № 3. — С. 478–481. — DOI: 10.31857/S0023476121030103.

95. Marrink S. J. Comment on "On Using a Too Large Integration Time Step in Molecular Dynamics Simulations of Coarse-Grained Molecular Models" by M. Winger, D. Trzesniak, R. Baron and W. F. van Gunsteren, *Phys. Chem. Chem. Phys.*, 2009, 11, 1934 / S. J. Marrink, X. Periole, D. P. Tieleman, A. H. De Vries // *Physical*

Chemistry Chemical Physics. — 2010. — Vol. 12. — № 9. — P. 2254–2256. — DOI: 10.1039/B915293H.

96. Monticelli L. The MARTINI Coarse-Grained Force Field: Extension to Proteins / L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, S. J. Marrink // *Journal of Chemical Theory and Computation*. — 2008. — Vol. 4. — № 5. — P. 819–834. — DOI: 10.1021/ct700324x.

97. Kroon P. C. Martinize2 and Vermouth: Unified Framework for Topology Generation / P. C. Kroon, F. Grünewald, J. Barnoud, M. van Tilburg, P. C. T. Souza, T. A. Wassenaar, S.-J. Marrink // *arXiv*. — 2022. — DOI: 10.48550/arXiv.2212.01191.

98. Wassenaar T. A. Computational Lipidomics with Insane: A Versatile Tool for Generating Custom Membranes for Molecular Simulations / T. A. Wassenaar, H. I. Ingólfsson, R. A. Böckmann, D. P. Tieleman, S. J. Marrink // *Journal of Chemical Theory and Computation*. — 2015. — Vol. 11. — № 5. — P. 2144–2155. — DOI: 10.1021/acs.jctc.5b00209.

99. Bernetti M. Pressure Control Using Stochastic Cell Rescaling / M. Bernetti, G. Bussi // *The Journal of Chemical Physics*. — 2020. — Vol. 153. — № 11. — P. 114107. — DOI: 10.1063/5.0020514.

100. Hünenberger P. H. Alternative Schemes for the Inclusion of a Reaction-Field Correction into Molecular Dynamics Simulations: Influence on the Simulated Energetic, Structural, and Dielectric Properties of Liquid Water / P. H. Hünenberger, W. F. Van Gunsteren // *The Journal of Chemical Physics*. — 1998. — Vol. 108. — № 15. — P. 6117–6134. — DOI: 10.1063/1.476022.

101. Xu C. Investigation of the Thermal Stability of Proteinase K for the Melt Processing of Poly(L-Lactide) / C. Xu, A. Battig, B. Schartel, R. Siegel, J. Senker, I. Von Der Forst, C. Unverzagt, S. Agarwal, A. Möglich, A. Greiner // *Biomacromolecules*. — 2022. — Vol. 23. — № 11. — P. 4841–4850. — DOI: 10.1021/acs.biomac.2c01008.

102. Ebeling W. Proteinase K from *Tritirachium Album* Limber / W. Ebeling, N. Hennrich, M. Klockow, H. Metz, H. D. Orth, H. Lang // *European Journal of Biochemistry*. — 1974. — Vol. 47. — № 1. — P. 91–97. — DOI: 10.1111/j.1432-1033.1974.tb03671.x.

103. Boom R. Rapid and Simple Method for Purification of Nucleic Acids / R. Boom, C. J. A. Sol, M. M. M. Salimans, C. L. Jansen, P. M. E. Wertheim-Van Dillen,

J. Van Der Noordaa // *Journal of Clinical Microbiology*. — 1990. — Vol. 28. — № 3. — P. 495–503. — DOI: 10.1128/jcm.28.3.495-503.1990.

104. Бойкова А. С. Исследование предкристаллизационной стадии раствора (влияния температуры и типа осадителя) протеиназы К методом малоуглового рассеяния рентгеновского излучения / А. С. Бойкова, Ю. А. Дьякова, К. Б. Ильина, П. В. Конарев, А. Е. Крюкова, М. А. Марченкова, Ю. В. Писаревский, М. В. Ковальчук // *Кристаллография*. — 2018. — Т. 63. — № 6. — С. 857–862. — DOI: 10.1134/S0023476118060061.

105. Konarev P. V. PRIMUS: A Windows PC-Based System for Small-Angle Scattering Data Analysis / P. V. Konarev, V. V. Volkov, A. V. Sokolova, M. H. J. Koch, D. I. Svergun // *Journal of Applied Crystallography*. — 2003. — Vol. 36. — № 5. — P. 1277–1282. — DOI: 10.1107/S0021889803012779.

106. Svergun D. CRY SOL—a Program to Evaluate X-Ray Solution Scattering of Biological Macromolecules from Atomic Coordinates / D. Svergun, C. Barberato, M. H. J. Koch // *Journal of Applied Crystallography*. — 1995. — Vol. 28. — № 6. — P. 768–773. — DOI: 10.1107/S0021889895007047.

107. Adekoya O. A. The Thermolysin Family (M4) of Enzymes: Therapeutic and Biotechnological Potential / O. A. Adekoya, I. Sylte // *Chemical Biology & Drug Design*. — 2009. — Vol. 73. — № 1. — P. 7–16. — DOI: 10.1111/j.1747-0285.2008.00757.x.

108. Kovalchuk M. V. Pre-Crystallization Phase Formation of Thermolysin Hexamers in Solution Close to Crystallization Conditions / M. V. Kovalchuk, A. S. Boikova, Y. A. Dyakova, K. B. Ilina, P. V. Konarev, A. E. Kryukova, M. A. Marchenkova, Y. V. Pisarevsky, V. I. Timofeev // *Journal of Biomolecular Structure and Dynamics*. — 2019. — Vol. 37. — № 12. — P. 3058–3064. — DOI: 10.1080/07391102.2018.1507839.