

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
имени М.В.ЛОМОНОСОВА  
БИОЛОГИЧЕСКИЙ ФАКУЛЬТЕТ



На правах рукописи

Шайтан Алексей Константинович

**Интегративное моделирование структуры и динамики  
биомакромолекулярных комплексов**

Специальность 03.01.09 —  
«Математическая биология, биоинформатика»

Диссертация на соискание учёной степени  
доктора физико-математических наук

Научный консультант:  
д-р биол. наук, академик РАН  
Кирпичников Михаил Петрович

Москва — 2021

## Оглавление

	Стр.
<b>Введение</b> . . . . .	<b>7</b>
Публикации автора по теме диссертации . . . . .	16
<b>Глава 1. Методы молекулярного и интегративного     моделирования в структурной биологии</b> . . . . .	<b>25</b>
1.1 Введение . . . . .	25
1.2 Методы атомистического суперкомпьютерного моделирования . . . . .	38
1.2.1 Силовые поля . . . . .	38
1.2.2 Методы молекулярной динамики . . . . .	43
1.2.3 Периодические граничные условия . . . . .	44
1.2.4 Статистические ансамбли в МД моделировании . . . . .	46
1.2.5 Ланжевеновская и диссипативная динамика частиц . . . . .	49
1.2.6 Расчет невалентных взаимодействий . . . . .	52
1.3 Методы расчета поверхностей в молекулярном моделировании . . . . .	55
1.4 Методы моделирования ДНК в динуклеотидном приближении . . . . .	56
1.4.1 Описание геометрии ДНК . . . . .	56
1.4.2 Эмпирические константы жесткости изгиба ДНК . . . . .	60
1.4.3 Программные продукты . . . . .	65
1.4.4 Алгоритмы интегративного моделирования . . . . .	65
1.5 Методы моделирования супрануклеосомной структуры хроматина . . . . .	67
1.6 Основы некоторых экспериментальных методов . . . . .	73
1.6.1 Методы футпринтинга ДНК . . . . .	73
1.6.2 FRET спектроскопия и микроскопия . . . . .	75
1.6.3 Методы геномного анализа в исследовании структуры хроматина . . . . .	82
1.7 Выводы главы 1 . . . . .	86
<b>Глава 2. Применение методов молекулярной динамики для     изучения нуклеосом</b> . . . . .	<b>89</b>
2.1 Введение в структуру и динамику нуклеосом . . . . .	89
2.1.1 Обзор динамических мод нуклеосом . . . . .	93
2.1.2 Факторы, влияющие на динамику нуклеосом: недавние примеры . . . . .	95

	Стр.
2.1.3	Тонкие детали динамики нуклеосом . . . . . 99
2.2	Моделирование нуклеосом с линкерными участками ДНК . . . . 102
2.2.1	Введение . . . . . 103
2.2.2	Результаты . . . . . 105
2.2.3	Обсуждение . . . . . 126
2.2.4	Материалы и методы . . . . . 129
2.3	Моделирование коровых частиц нуклеосом на масштабах свыше 10 микросекунд . . . . . 134
2.3.1	Введение . . . . . 134
2.3.2	Материалы и методы . . . . . 137
2.3.3	Результаты и обсуждение . . . . . 139
2.3.4	Благодарности . . . . . 156
2.4	Выводы главы 2 . . . . . 156

<b>Глава 3.</b>	<b>Интегративное моделирование с применением данных экспериментов по футпринтингу ДНК . . . . . 159</b>
3.1	Методы количественной обработки и структурной интерпретации экспериментов по расщеплению ДНК гидроксил радикалами с помощью программы HYDROID . . . . . 159
3.1.1	Введение . . . . . 161
3.1.2	Разработка протокола . . . . . 163
3.1.3	Применение метода . . . . . 164
3.1.4	Сравнение с другими методами . . . . . 165
3.1.5	Преимущества и ограничения . . . . . 166
3.1.6	Обзор протокола HYDROID . . . . . 167
3.1.7	Конвейер HYDROIDexp . . . . . 168
3.1.8	Конвейер HYDROIDpred . . . . . 172
3.1.9	Визуализация и сравнение данных . . . . . 173
3.1.10	Планирование экспериментов . . . . . 176
3.1.11	Примеры данных . . . . . 181
3.1.12	Материалы . . . . . 183
3.1.13	Описание протокола . . . . . 184
3.1.14	Ожидаемые результаты . . . . . 191
3.1.15	Благодарности . . . . . 196

3.2	Создание моделей нуклеосом по данным гидроксильного футпринтинга . . . . .	196
3.2.1	Введение . . . . .	197
3.2.2	Материалы и методы . . . . .	202
3.2.3	Результаты . . . . .	209
3.2.4	Обсуждение . . . . .	223
3.2.5	Благодарности . . . . .	228
3.3	Выводы главы 3 . . . . .	228
<b>Глава 4. Моделирование комплексов нуклеосом с белками</b>		
	<b>хроматина . . . . .</b>	<b>230</b>
4.1	Введение . . . . .	230
4.1.1	О прогрессе методов структурной биологии в изучении супрануклеосомной организации . . . . .	238
4.2	Комплексы нуклеосомы и РНК-полимеразы . . . . .	239
4.2.1	Введение . . . . .	239
4.2.2	Создание моделей комплекса в положении активного центра +49 (нулевая петля) . . . . .	240
4.2.3	Выявление отрицательно заряженных регионов на предполагаемых гистон-взаимодействующих поверхностях Pol II . . . . .	241
4.2.4	Более высокий отрицательный заряд области ZLS коррелирует с более сильным барьером +45 и более эффективной выживаемостью нуклеосом при транскрипции . . . . .	246
4.2.5	Создание моделей комплекса полимеразы и нуклеосомы в положении активного центра +42 . . . . .	250
4.3	Конформационные перестройки нуклеосом при взаимодействии с комплексом FАСТ . . . . .	257
4.4	Модели взаимодействия нуклеосом с белком CENP-C . . . . .	260
4.4.1	Введение . . . . .	261
4.4.2	Построение модель взаимодействия на основе данных футпринтинга ДНК . . . . .	262
4.4.3	Обсуждение . . . . .	265

4.5	Анализ комплексов нуклеосомы и гистона H1 . . . . .	267
4.5.1	Совместное влияние аминокислотной последовательности гистона H1 и нуклеотидной последовательности ДНК на структуру хроматосомы: анализ методами молекулярного моделирования . . . . .	267
4.5.2	Моделирование структуры ДНК в хроматосоме по данным FRET . . . . .	276
4.6	Выводы главы 4 . . . . .	281

## **Глава 5. Интегративное моделирование амилоидоподобных фибрилл . . . . .**

5.1	Амилоиды и материалы на основе амилоидов . . . . .	283
5.1.1	Строение природных амилоидных фибрилл и их полиморфизм . . . . .	284
5.1.2	Экспериментальные методы, используемые для изучения амилоидоподобных волокон . . . . .	290
5.1.3	Самосборка и образование фибрилл . . . . .	290
5.1.4	Функциональные амилоиды и синтетические амилоидные агрегаты . . . . .	294
5.2	Мультимасштабное интегративное моделирование фибрилл из тиофен-пептидных олигомеров . . . . .	305
5.2.1	Формулировка задачи . . . . .	305
5.2.2	Молекулярно-механическая модель . . . . .	306
5.2.3	Построение и анализ периодических укладок . . . . .	308
5.2.4	Моделирование фибриллярных агрегатов в растворителе . . . . .	314
5.2.5	Крупномасштабное моделирование морфологии агрегатов в вакууме . . . . .	333
5.2.6	Адсорбция агрегатов на поверхности . . . . .	344
5.2.7	Обсуждение . . . . .	352
5.2.8	Выводы раздела . . . . .	373
5.3	Моделирование пептидных фибрилл EF-C, взаимодействующих с ретровирусами . . . . .	373
5.3.1	Описание экспериментальных результатов . . . . .	374
5.3.2	Детали моделирования . . . . .	381

	Стр.
5.3.3 Моделирование молекулярной динамики . . . . .	381
5.4 Выводы главы 5 . . . . .	385
<b>Заключение . . . . .</b>	<b>386</b>
Обсуждение результатов . . . . .	386
Выводы диссертационной работы . . . . .	390
Перспективы . . . . .	392
Благодарности . . . . .	394
<b>Список сокращений и условных обозначений . . . . .</b>	<b>395</b>
<b>Список литературы . . . . .</b>	<b>397</b>
<b>Список рисунков . . . . .</b>	<b>482</b>
<b>Список таблиц . . . . .</b>	<b>488</b>

## Введение

Понимание механизмов функционирования живых систем на молекулярном и супрамолекулярном уровнях является одной из ключевых задач биологии XXI века. Такое понимание важно не только с фундаментальной точки зрения, но и имеет прямые практические приложения в области медицины, здравоохранения, биоинженерии и биотехнологии. Особенностью биологических молекул является формирование ими сложных пространственных структур, которые могут перестраиваться в ходе различных функциональных процессов. Исследование *структурно-динамических* моделей биомакромолекул является одним из ключевых подходов к изучению механизмов функционирования живой материи. Методы компьютерного молекулярного моделирования являются неотъемлемой частью этого подхода. Среди методов построения моделей важное место занимают методы, использующие экспериментальные данные рентгеноструктурного анализа (РСА), ядерно-магнитного резонанса (ЯМР), криоэлектронной микроскопии (криоЭМ). В настоящее время они сочетаются с методами молекулярной динамики (МД), Монте-Карло, позволяющих проводить *компьютерные эксперименты*, имитационное моделирование молекулярных систем. В этой области существуют и ограничения, которые затрудняют исследования биологических систем значительных размеров таких, как большинство биомакромолекулярных комплексов. Сложность интерпретации и самого получения данных РСА, ЯМР, криоЭМ возрастает по мере увеличения размера и конформационной подвижности исследуемой системы. По состоянию на 2020 год 92% структур депонированные в базе PDB имеют молекулярный вес менее 75 кДа. Возможности исследования больших биомакромолекулярных комплексов (от 100 кДа и более) методами молекулярной динамики ограничиваются вычислительными возможностями и качеством параметризации атом-атомных взаимодействий (так называемых силовых полей). Несмотря на то, что благодаря прогрессу компьютерных технологий, методом МД стал возможен расчет систем, состоящих из сотен тысяч и миллионов атомов на временах десятков микросекунд (а с использованием специализированных суперкомпьютеров и на миллисекундных временах), временные масштабы многих физиологически важных процессов, включая важные моды динамики многих биомакромолекулярных комплексов, находятся пока за

пределами возможностей моделирования в субсекундном и секундном диапазоне времен. С ростом размера биомолекулярной системы неточности задания атом-атомных взаимодействий суммируются и сильнее отражаются на качестве воспроизведения ее динамики, в то время как количество динамических состояний увеличивается, а разница свободной энергии между ними уменьшается. *Преодоление вышеописанных ограничений с целью построения и анализа структурно-динамических моделей биомакромолекулярных комплексов является важной задачей.* На современном этапе одним из основных путей решения данной задачи является разработка и совершенствование подходов, которые используют для построения моделей как расчет физических взаимодействий между атомами, так и различные экспериментальные данные. К таким экспериментальным данным могут относиться как данные классических методов структурной биологии (данные РСА, ЯМР, криоЭМ), так и данные спектральных методов, различных видов микроскопии, биохимического анализа.

Методы молекулярного компьютерного моделирования зародились в 1940-ых годах с появлением первых вычислительных систем, однако активное их применение к биологическим молекулам стало возможным в 1970-ых – 1980-ых годах. Активно стали развиваться как подходы к моделированию биомолекул методом МД, так и оптимизационные методы моделирования структуры на основе экспериментальных данных (главным образом РСА и ЯМР). Методы МД основаны на представлении молекулярной системы в виде набора атомов, взаимодействующих согласно законам классической механики Ньютона, и численном решении уравнений движения. На сегодняшний день разработана обширная методическая база для моделирования белков, нуклеиновых кислот, липидов методами молекулярной динамики, разработаны наборы функциональных форм и параметров, описывающих взаимодействия между атомами (например, семейства силовых полей CHARMM, AMBER, OPLS и др.), алгоритмы численного интегрирования уравнений движения, поддержания температуры, давления, использования параллельных суперкомпьютерных технологий, созданы программные комплексы для расчетов (например, AMBER, CHARMM, GROMACS, NAMD, PUMA и др.). В то же время уровень развития методов МД, как с точки зрения доступных времен моделирования, так и с точки зрения качества силовых полей, не позволяет использовать их полностью независимым образом в отрыве от экспериментальных данных. Стандартной практикой для расчетов методом МД является



использование в качестве стартовой конформации биомакромолекулярной системы структурной модели, полученной на основе данных РСА, ЯМР, криоЭМ.

Важный вклад в построение структурных моделей биомакромолекулярных системы вносят комбинированные вычислительные подходы, которые, кроме данных РСА, ЯМР, криоЭМ, могут использовать различные экспериментальные данные, получаемые в ходе биофизических, биохимических, функциональных, спектроскопических и других экспериментов. Такого рода эксперименты могут предоставлять информацию о расстояниях между введенными метками в белке или ДНК (например, методы FRET, ЭПР), характере укладки белковой цепи (например, ИК-, КД-спектроскопия, дифракция рентгеновских лучей на фибриллах), реакционной доступности химических групп (методы футпринтинга, химического сшивания) и др. Такой подход, основанный на интеграции различных экспериментальных данных при построении структурно-динамических моделей, в литературе называется *интегративным моделированием* [1]<sup>1</sup>. Программы, используемые для реконструкции структур по данным экспериментов ЯМР, вероятно, первыми начали позволять использовать некоторые типы данных из других экспериментов, например, данные малоуглового рентгеновского и нейтронного рассеяния [2]. Другим примером является использование в методах МД дополнительных членов потенциальной энергии, которые зависят от экспериментально полученных карт электронной плотности низкого разрешения методами электронной микроскопии или РСА. Такой подход позволят деформировать начальные структуры биомолекул для придания им экспериментально наблюдаемой формы [3]. Разрабатываются программные подходы, предоставляющие возможность одновременно использовать данные различных биофизических и биохимических экспериментов (например, карты электронной плотности и расстояния между группами атомов) [4]. В целом подходы интегративного моделирования могут различаться по (i) уровню представления модели (например, атомистическое или огрубленное представление модели), (ii) способу оценки соответствия модели экспериментальным данным (использование различных силовых полей или оценочных функций), (iii) методам поиска структуры, соответствующей экспериментальным данным, включая использование различных алгоритмов оптимизации, в том числе минимизации энергии, расчетов методом Монте-Карло, (iv) заданию

<sup>1</sup>Ссылки на список литературы приведены цифрами, ссылки на список статей автора – цифрами с буквой А.

типа варьируемых параметров (координаты атомов, групп атомов, обобщенные координаты реакции). В настоящее время универсальных подходов интегративного моделирования не существует, такие подходы требуют специальных методов и программных реализаций под конкретный тип моделируемых систем (например, комплексы белок-ДНК, амилоидоподобные фибриллы, мембранные системы и т.д.) и конкретный набор экспериментальных данных.

**Целью** данной работы является разработка интегративных подходов к построению и методов анализа структурно-динамических моделей больших ДНК-белковых комплексов и амилоидоподобных фибрилл.

В работе поставлены и решены следующие основные **задачи**:

1. Разработать подходы к интегративному моделированию ДНК-белковых комплексов методом молекулярной динамики и проанализировать динамику нуклеосом с атомистическим уровнем детализации в микросекундном временном диапазоне.
2. На основе различных экспериментальных данных разработать интегративные подходы к моделированию и верификации моделей комплексов ДНК и белков.
3. Для интегративного моделирования биомакромолекулярных комплексов разработать методы анализа экспериментальных данных по расщеплению ДНК гидроксильными радикалами. Применить данные методы для изучения организации ДНК в нуклеосомах.
4. Разработать подходы и методы построения моделей нуклеосом с белками хроматина. Применить данные подходы для изучения комплексов нуклеосом с белками H1, CENP-C, комплексом белков FACT.
5. Разработать и применить интегративные подходы по построению атомистических моделей амилоидоподобных фибрилл на основе экспериментальных данных спектроскопии и микроскопии.

**Объект и предмет исследований.** Предмет исследования состоит в разработке новых системных подходов и алгоритмов построения структурно-динамических моделей биомакромолекулярных комплексов на основе методов молекулярного моделирования с использованием различных экспериментальных данных. Объектами исследования являлись ДНК-белковые комплексы, нуклеосомы, комплексы нуклеосом с белками хроматина, амилоидоподобные фибриллы, состоящие из пептидов и конъюгатов пептидов с синтетическими полимерами.

## **Научная новизна**

Предложен оригинальный интегративный подход к созданию моделей ДНК-белковых комплексов с использованием атомистического моделирования биомакромолекул методами молекулярной динамики, огрубленного молекулярного моделирования ДНК, с учетом свойств симметрии ДНК-белковых комплексов и экспериментальных данных по футпринтингу ДНК, распределению электронной плотности в комплексах, спектроскопических данных о расстояниях между мечеными нуклеотидами.

Впервые на атомистическом уровне в микросекундном временном диапазоне рассчитаны траектории молекулярной динамики нуклеосом и определены микроконформационные состояния ДНК в процессах “дыхания” и отворачивания ДНК от октамера гистонов.

Установлена конформация нуклеосом и их комплексов с рядом белков хроматина (белком СENP-C, РНК полимеразы, белками комплекса FACT), установлены структуры ряда амилоидоподобных фибрилл.

Методами мультимасштабного молекулярного моделирования с использованием данных ИК-, КД-спектроскопии, рентгеновской дифракции, атомно-силовой микроскопии установлены структуры ряда амилоидоподобных фибрилл.

## **Теоретическая и практическая значимость**

Разработанные в данной работе методы и подходы позволяют решать ряд актуальных научных и практических задач, связанных с изучением структурно-динамических характеристик биомакромолекулярных комплексов. Результаты работы применимы для изучения упаковки ДНК в хроматине на нуклеосомном/супрануклеосомном уровнях и упаковки пептидов в амилоидоподобных фибриллах. Структурно-динамические модели нуклеосом и их комплексов могут быть использованы в разработке лекарственных препаратов, направленных на модуляцию эпигенетических механизмов в клетке, например, для поиска ингибиторов связывания белков хроматина с нуклеосомами или ингибиторов посттрансляционных модификаций гистонов. Полученные модели амилоидоподобных фибрилл служат основой для дизайна функциональных самособирающихся филаментов (электропроводящих фибрилл, фибрилл, усиливающих вирусную трансдукцию).

## **Методология и методы исследования.**

Разработанные в работе комплексный подход к изучению структуры и динамики биомакромолекулярных комплексов основан на комбинации различных методов атомистической молекулярной динамики, огрубленного моделирования, оптимизации геометрии молекулярной системы при заданных ограничениях на экспериментально определенные параметры системы, отбора конформаций молекулярной системы, соответствующих определенным экспериментальным параметрам, оцифровки экспериментальных данных, с целью определения численных значений параметров, используемых при моделировании.

Методы атомистической молекулярной динамики позволяют проследить за характером временной эволюции, оценить оптимальность стартовой структуры молекулярной системы, изучить ее конформационную релаксацию, охарактеризовать ансамбль динамических состояний стартовой структуры. В результате МД расчетов становится возможным оценить различные макропараметры системы, связанные с набором динамических состояний системы, и сравнить их значения с экспериментальными данными. В случае нуклеосом таким параметром, например, является доступность нуклеотидов для атаки и расщепления гидроксильными радикалами, которую можно сравнить с результатами экспериментов по гидроксильному футпринтингу. В случае амилоидоподобных фибрилл – шаг спирали фибриллы, который можно сравнить с данными атомно-силовой микроскопии, а структурные факторы – с данными рентгеновской дифракции.

Используемые нами методы огрубленного моделирования основаны на представлении ДНК в виде последовательности Уотсон-Криковских пар оснований, где каждая пара представляется в виде плоского абстрактного элемента. Взаимное расположение соседних вдоль по цепи элементов определяется шестью параметрами Rise, Shift, Slide, Twist, Tilt, Roll (три отвечают за поступательные, а три за вращательные степени свободы). Функция потенциальной энергии задается в гармоническом динуклеотидном приближении, то есть зависит от квадратов смещений этих параметров вдоль цепи ДНК. Такое представление ДНК позволяет достаточно быстро проводить оптимизацию ее геометрии путем поиска минимума энергии. Задание внешних экспериментальных ограничений, например, ограничений на расстояния между определенными нуклеотидами, измеренные методом Ферстеровского резонансного переноса

энергии (FRET), позволяет рассчитывать модели, удовлетворяющие экспериментальным данным.

Важным методологическим подходом в нашем исследовании являлся также метод (автоматизированного или полу-автоматизированного) перебора возможной структурной организации биомакромолекулярного комплекса среди большого набора детерминированных вариантов и оценка его соответствия экспериментальным данным, а также оценка его относительной энергетической выгоды. Таким подходом решались, в частности, задачи по поиску типов укладки бета-листов в амилоидоподобных фибриллах. Например, в данных ИК- и КД-спектров проявление пиков на определенной длине волны можно связать с наличием в структуре бета-листов с параллельной или антипараллельной укладкой бета-нитей, данные рентгеновской дифракции с определенным расстоянием между бета-листами. По этим параметрам можно проводить отбор конформаций генерируемых молекулярных систем, соответствующих определенной стартовой структуре и ее эволюции. Аналогично, при построении моделей нуклеосом с определенным положением ДНК, можно использовать данные гидроксильного футпринтинга ДНК для отбора моделей с правильным положением ДНК на нуклеосоме.

При моделировании методами молекулярной динамики использовались программные пакеты NAMD, LAMMPS и Gromacs различных версий, силовые поля AMBER, CHARMM, PCFF с различными моделями воды и параметрами ионов. Большинство процедур обработки данных, автоматизации и огрубленного моделирования производились с применением языка Python, библиотек NumPy, SciPy, MDAnalysis и других. В ходе работы был разработан ряд программных библиотек для обработки экспериментальных данных (HYDORID, HYDROID\_GUI), а также алгоритмы и конвейеры для моделирования нуклеосом, комплексов нуклеосом с белками хроматина, амилоидоподобных фибрилл.

### **Основные положения, выносимые на защиту:**

1. Для построения структурно-динамических моделей сложных биомакромолекулярных комплексов концептуально обосновано применение новых интегративных подходов на основе сочетания методов атомистического и огрубленного молекулярного моделирования, методов учета разнородных экспериментальных данных рентгеноструктурного анализа, атомно-силовой и электронной микроскопии, футпринтинга ДНК,

- ИК-, КД-спектроскопии, измерений расстояний между флуоресцентными метками на основе эффекта Ферстеровского резонансного переноса энергии.
2. С использованием разработанного интегративного подхода возможно создание атомистических моделей нуклеосом и комплексов нуклеосом с белками хроматина, при этом учет симметрии белковых комплексов позволяет значительно повысить точность построения молекулярных моделей на основе данных футпринтинга ДНК.
  3. Интегративное моделирование позволяет воспроизвести на атомистическом уровне функциональную динамику нуклеосом, важную с точки зрения эпигенетической регуляции функционирования генома, включая крупномасштабные конформационные перестройки структуры ДНК-белковых комплексов (углы входа-выхода ДНК в нуклеосоме, диффузия гистоновых хвостов вдоль ДНК), а также позволяет обнаружить новые моды динамической подвижности, связанные с изменением конформации ДНК, перестройкой взаимодействий гистоновых хвостов, деформацией глобулярных доменов гистонов.
  4. Разработанный подход интегративного моделирования применим для построения молекулярных моделей амилоидоподобных фибрилл, реконструкции укладки пептидов в фибриллах и установления связи между морфологией фибриллы и межмолекулярной укладкой пептидов.

**Достоверность** полученных результатов обеспечивается их публикацией в рецензируемых журналах международного уровня с высокими импакт-факторами. Результаты находятся в соответствии с результатами, полученными другими авторами. Материалы диссертационной работы докладывались и обсуждались на научных семинарах в МГУ, ряде международных университетов и исследовательских центров (включая, Национальные Институты Здоровья, США, Университет г. Ульма, Германия, Университет Джонса Хопкинса, США, Классический Университет Центрального Китая, Китай и др.). Основные результаты работы докладывались на следующих научных конференциях: Biophysical Society 64th Annual Meeting Сан-Диего, США, 15-19 февраля 2020; The 44th FEBS Congress, Краков, Польша, 6-11 июля 2019; Multiscale Modeling of Chromatin: Bridging Experiment with Theory, Les Houches, Франция, 31 марта - 5 апреля 2019; Keystone Symposia on Molecular and Cellular Biology “Genomic instability and DNA repair”, Санта Фе, США, 2-7 апреля 2017; Biophysical Society

60th Annual Meeting, Los Angeles, США, 27 февраля - 2 марта 2016; 27 Международная конференция “Математика. Компьютер. Образование” г. Дубна, Государственный университет “Дубна”, Россия, 27 января - 1 февраля 2020; Russian International Conference on Cryo-Electron Microscopy 2019, Москва, Россия, 2-5 июня 2019; Russian International Conference on Cryo-Electron Microscopy 2019, Москва, Россия, 2-5 июня 2019; “Ломоносовские чтения - 2019” Секция “Биология”, Москва, МГУ, Россия, 15-25 апреля 2019; XXVI международная конференция “Математика. Компьютер. Образование”, Пущино, Россия, 28 января - 2 февраля 2019 и др. Ряд тезисов опубликованы в научных журналах (см. позиции 38-54 списка публикаций [A1—A54]).

**Личный вклад.** Основные идеи, методология и результаты исследований, изложенные в диссертации, получены автором лично. Вклад автора в совместных экспериментально-теоретических работах заключается в применении разработанных методов, моделировании и интерпретации экспериментальных результатов, полученных коллегами.

**Публикации.** Основные результаты по теме диссертации изложены в 35 статьях в рецензируемых научных изданиях, индексируемых в базах данных Web of Science, Scopus, RSCI [A1—A35] (дополнительно опубликованы еще две статьи [A36; A37]). Зарегистрированы 1 патент [A55] и 1 программа для ЭВМ [A56].

**Объем и структура работы.** Диссертация состоит из введения, пяти глав и заключения. Первая глава посвящена обзору методов исследований, примененных и разработанных в ходе работы, обсуждается место методов молекулярного и интегративного моделирования в исследовательском процессе. Вторая глава посвящена применению методов суперкомпьютерной молекулярной динамики для изучения динамики нуклеосом. Третья глава посвящена разработке методов интегративного моделирования комплексов белков с ДНК на основе данных расщепления ДНК гидроксильными радикалами. Четвертая глава посвящена построению моделей комплексов нуклеосом с белками хроматина. Пятая глава посвящена моделированию амилоидоподобных фибрилл. Полный объем диссертации составляет 488 страниц, включая 106 рисунков и 9 таблиц. Список литературы содержит 677 наименований.

## Публикации автора по теме диссертации<sup>2</sup>

- A1. Histone Octamer Structure Is Altered Early in ISW2 ATP-Dependent Nucleosome Remodeling / A. Hada, S. K. Hota, J. Luo, Y.-c. Lin, S. Kale, A. K. Shaytan, S. K. Bhardwaj, J. Persinger, J. Ranish, A. R. Panchenko, B. Bartholomew // *Cell Reports*. — 2019. — July 2. — Vol. 28, no. 1. — 282–294.e6. — DOI: [10.1016/j.celrep.2019.05.106](https://doi.org/10.1016/j.celrep.2019.05.106). — **IF WoS 7.7** - (2,3/0,3).
- A2. The Effect of Oncomutations and Posttranslational Modifications of Histone H1 on Chromatosome Structure and Stability / M. V. Bass, G. A. Armeev, K. V. Shaitan, A. K. Shaytan // *Moscow University Biological Sciences Bulletin*. — 2019. — July. — Vol. 74, no. 3. — P. 121–126. — DOI: [10.3103/S0096392519030015](https://doi.org/10.3103/S0096392519030015). — **IF RINC 0.76** - (0,7/0,3).
- A3. Linking Chromatin Composition and Structural Dynamics at the Nucleosome Level / G. A. Armeev, A. K. Gribkova, I. Pospelova, G. A. Komarova, A. K. Shaytan // *Current Opinion in Structural Biology*. — 2019. — June. — Vol. 56. — P. 46–55. — DOI: [10.1016/j.sbi.2018.11.006](https://doi.org/10.1016/j.sbi.2018.11.006). — **IF WoS 6.908** - (1,2/0,6).
- A4. Structural Interpretation of DNA–Protein Hydroxyl-Radical Footprinting Experiments with High Resolution Using HYDROID / A. K. Shaytan, H. Xiao, G. A. Armeev, D. A. Gaykalova, G. A. Komarova, C. Wu, V. M. Studitsky, D. Landsman, A. R. Panchenko // *Nature Protocols*. — 2018. — Nov. — Vol. 13, no. 11. — P. 2535–2556. — DOI: [10.1038/s41596-018-0048-z](https://doi.org/10.1038/s41596-018-0048-z). — **IF WoS 11.334** - (2,5/2,2).
- A5. Joint Effect of Histone H1 Amino Acid Sequence and DNA Nucleotide Sequence on the Structure of Chromatosomes: Analysis by Molecular Modeling Methods / T. K. Gorkovets, G. A. Armeev, K. V. Shaitan, A. K. Shaytan // *Moscow University Biological Sciences Bulletin*. — 2018. — Apr. — Vol. 73, no. 2. — P. 82–87. — DOI: [10.3103/S0096392518020025](https://doi.org/10.3103/S0096392518020025). — **IF RINC 0.76** - (0,7/0,3).

---

<sup>2</sup>В скобках приведен объем публикации в печатных листах и вклад автора в печатных листах.



- A6. Molecular Basis of CENP-C Association with the CENP-A Nucleosome at Yeast Centromeres / H. Xiao, F. Wang, J. Wisniewski, A. K. Shaytan, R. Ghirlando, P. C. FitzGerald, Y. Huang, D. Wei, S. Li, D. Landsman, A. R. Panchenko, C. Wu // *Genes & Development*. — 2017. — Oct. 1. — Vol. 31, no. 19. — P. 1958–1972. — DOI: [10.1101/gad.304782.117](https://doi.org/10.1101/gad.304782.117). — **IF WoS 9.527** - (1,7/0,3).
- A7. Hydroxyl-Radical Footprinting Combined with Molecular Modeling Identifies Unique Features of DNA Conformation and Nucleosome Positioning / A. K. Shaytan, H. Xiao, G. A. Armeev, C. Wu, D. Landsman, A. R. Panchenko // *Nucleic Acids Research*. — 2017. — Sept. 19. — Vol. 45, no. 16. — P. 9229–9243. — DOI: [10.1093/nar/gkx616](https://doi.org/10.1093/nar/gkx616). — **IF WoS 11.501** - (1,7/1,5).
- A8. *Gribkova, A. K.* Investigation of Histone-DNA Binding Energy as a Function of DNA Unwrapping from Nucleosome Using Molecular Modeling / A. K. Gribkova, G. A. Armeev, A. K. Shaytan // *Moscow University Biological Sciences Bulletin*. — 2017. — July. — Vol. 72, no. 3. — P. 142–145. — DOI: [10.3103/S009639251703004X](https://doi.org/10.3103/S009639251703004X). — **IF RINC 0.76** - (0,5/0,2).
- A9. MS\_HistoneDB, a Manually Curated Resource for Proteomic Analysis of Human and Mouse Histones / S. El Kennani, A. Adrait, A. K. Shaytan, S. Khochbin, C. Bruley, A. R. Panchenko, D. Landsman, D. Pflieger, J. Govin // *Epigenetics & Chromatin*. — 2017. — Dec. — Vol. 10, no. 1. — DOI: [10.1186/s13072-016-0109-x](https://doi.org/10.1186/s13072-016-0109-x). — **IF WoS 5.333** - (2,1/0,5).
- A10. Dual Active Site in the Endolytic Transglycosylase Gp144 of Bacteriophage phiKZ / O. V. Chertkov, G. A. Armeev, I. V. Uporov, S. A. Legotsky, N. N. Sykilinda, A. K. Shaytan, N. L. Klyachko, K. A. Miroshnikov // *Acta Naturae*. — 2017. — Vol. 9, no. 1. — P. 7. — DOI: [10.32607/20758251-2017-9-1-81-87](https://doi.org/10.32607/20758251-2017-9-1-81-87). — **IF WoS 1.62** - (0,8/0,1).
- A11. Modeling of the Structure of Protein–DNA Complexes Using the Data from FRET and Footprinting Experiments / G. A. Armeev, T. K. Gorkovets, D. A. Efimova, K. V. Shaitan, A. K. Shaytan // *Moscow University Biological Sciences Bulletin*. — 2016. — Jan. — Vol. 71, no. 1. — P. 29–33. — DOI: [10.3103/S0096392516010016](https://doi.org/10.3103/S0096392516010016). — **IF RINC 0.76** - (0,6/0,2).

- A12. *Armeev, G. A.* Nucleosome Structure Relaxation during DNA Unwrapping: Molecular Dynamics Simulation Study / G. A. Armeev, K. V. Shaitan, A. K. Shaytan // *Moscow University Biological Sciences Bulletin*. — 2016. — July. — Vol. 71, no. 3. — P. 141–144. — DOI: [10.3103/S0096392516030020](https://doi.org/10.3103/S0096392516030020). — **IF RINC 0.76** - (0,5/0,2).
- A13. Genomic Profiling of Multiple Sequentially Acquired Tumor Metastatic Sites from an “Exceptional Responder” Lung Adenocarcinoma Patient Reveals Extensive Genomic Heterogeneity and Novel Somatic Variants Driving Treatment Response / R. Biswas, S. Gao, C. M. Cultraro, T. K. Maity, A. Venugopalan, Z. Abdullaev, A. K. Shaytan, C. A. Carter, A. Thomas, A. Rajan, Y. Song, S. Pitts, K. Chen, S. Bass, J. Boland, K.-I. Hanada, J. Chen, P. S. Meltzer, A. R. Panchenko, J. C. Yang, S. Pack, G. Giaccone, D. S. Schrupp, J. Khan, U. Guha // *Molecular Case Studies*. — 2016. — Nov. — Vol. 2, no. 6. — a001263. — DOI: [10.1101/mcs.a001263](https://doi.org/10.1101/mcs.a001263). — **IF WoS 1.750** - (3,1/0,5).
- A14. HistoneDB 2.0: A Histone Database with Variants—an Integrated Resource to Explore Histones and Their Variants / E. J. Draizen, A. K. Shaytan, L. Mariño-Ramírez, P. B. Talbert, D. Landsman, A. R. Panchenko // *Database*. — 2016. — Vol. 2016. — baw014. — DOI: [10.1093/database/baw014](https://doi.org/10.1093/database/baw014). — **IF WoS 2.593** - (1,2/0,6).
- A15. Structure and Functions of Linker Histones / A. V. Lyubitelev, D. V. Nikitin, A. K. Shaytan, V. M. Studitsky, M. P. Kirpichnikov // *Biochemistry (Moscow)*. — 2016. — Mar. — Vol. 81, no. 3. — P. 213–223. — DOI: [10.1134/S0006297916030032](https://doi.org/10.1134/S0006297916030032). — **IF WoS 1.978** - (1,3/0,1).
- A16. *Shaitan, K. V.* The Dynamics of Irreversible Evaporation of a Water–Protein Droplet and the Problem of Structural and Dynamic Experiments with Single Molecules / K. V. Shaitan, G. A. Armeev, A. K. Shaytan // *Biophysics*. — 2016. — Mar. — Vol. 61, no. 2. — P. 177–184. — DOI: [10.1134/S0006350916020172](https://doi.org/10.1134/S0006350916020172). — **IF SJR 0.226** - (0,9/0,1).
- A17. Coupling between Histone Conformations and DNA Geometry in Nucleosomes on a Microsecond Timescale: Atomistic Insights into Nucleosome Functions / A. K. Shaytan, G. A. Armeev, A. Goncarencu, V. B. Zhurkin, D. Landsman, A. R. Panchenko // *Journal of Molecular Biology*. — 2016. —

- Jan. — Vol. 428, no. 1. — P. 221–237. — DOI: [10.1016/j.jmb.2015.12.004](https://doi.org/10.1016/j.jmb.2015.12.004). — **IF WoS 5.04** - (2,0/1,8).
- A18. Trajectories of Microsecond Molecular Dynamics Simulations of Nucleosomes and Nucleosome Core Particles / A. K. Shaytan, G. A. Armeev, A. Goncarencu, V. B. Zhurkin, D. Landsman, A. R. Panchenko // *Data in Brief*. — 2016. — June. — Vol. 7. — P. 1678–1681. — DOI: [10.1016/j.dib.2016.04.073](https://doi.org/10.1016/j.dib.2016.04.073). — **IF WoS 0.97** - (0,5/0,5).
- A19. Large-Scale ATP-Independent Nucleosome Unfolding by a Histone Chaperone / M. E. Valieva, G. A. Armeev, K. S. Kudryashova, N. S. Gerasimova, A. K. Shaytan, O. I. Kulaeva, L. L. McCullough, T. Formosa, P. G. Georgiev, M. P. Kirpichnikov, V. M. Studitsky, A. V. Feofanov // *Nature Structural & Molecular Biology*. — 2016. — Dec. — Vol. 23, no. 12. — P. 1111–1116. — DOI: [10.1038/nsmb.3321](https://doi.org/10.1038/nsmb.3321). — **IF WoS 11.98** - (0,9/0,1).
- A20. *Armeev, G. A.* Conformational Flexibility of Nucleosomes: A Molecular Dynamics Study / G. A. Armeev, K. V. Shaitan, A. K. Shaytan // *Moscow University Biological Sciences Bulletin*. — 2015. — July. — Vol. 70, no. 3. — P. 147–151. — DOI: [10.3103/S0096392515030025](https://doi.org/10.3103/S0096392515030025). — **IF RINC 0.76** - (0,6/0,3).
- A21. *Armeev, G. A.* Molecular Dynamics Study of the Ionic Environment and Electrical Characteristics of Nucleosomes / G. A. Armeev, K. V. Shaitan, A. K. Shaitan // *Moscow University Biological Sciences Bulletin*. — 2015. — Oct. — Vol. 70, no. 4. — P. 173–176. — DOI: [10.3103/S0096392515040033](https://doi.org/10.3103/S0096392515040033). — **IF RINC 0.76** - (0,5/0,2).
- A22. Direct Prediction of Residual Dipolar Couplings of Small Molecules in a Stretched Gel by Stochastic Molecular Dynamics Simulations: Direct Prediction of Residual Dipolar Couplings by Stochastic MD Simulations / A. O. Frank, J. C. Freudenberger, A. K. Shaytan, H. Kessler, B. Luy // *Magnetic Resonance in Chemistry*. — 2015. — Mar. — Vol. 53, no. 3. — P. 213–217. — DOI: [10.1002/mrc.4181](https://doi.org/10.1002/mrc.4181). — **IF WoS 1.731** - (0,6/0,1).
- A23. Structural Analysis of Nucleosomal Barrier to Transcription / D. A. Gaykalova, O. I. Kulaeva, O. Volokh, A. K. Shaytan, F.-K. Hsieh, M. P. Kirpichnikov, O. S. Sokolova, V. M. Studitsky // *Proceedings of the National Academy of Sciences*. — 2015. — Oct. 27. — Vol. 112, no. 43. —

- E5787–E5795. — DOI: [10.1073/pnas.1508371112](https://doi.org/10.1073/pnas.1508371112). — **IF WoS 9.412** - (1,0/0,3).
- A24. Structural Perspectives on the Evolutionary Expansion of Unique Protein-Protein Binding Sites / A. Goncarenco, A. K. Shaytan, B. A. Shoemaker, A. R. Panchenko // *Biophysical Journal*. — 2015. — Sept. — Vol. 109, no. 6. — P. 1295–1306. — DOI: [10.1016/j.bpj.2015.06.056](https://doi.org/10.1016/j.bpj.2015.06.056). — **IF WoS 3.665** - (1,4/0,4).
- A25. *Shaytan, A. K.* Nucleosome Adaptability Conferred by Sequence and Structural Variations in Histone H2A-H2B Dimers / A. K. Shaytan, D. Landsman, A. R. Panchenko // *Current Opinion in Structural Biology*. — 2015. — June. — Vol. 32. — P. 48–57. — DOI: [10.1016/j.sbi.2015.02.004](https://doi.org/10.1016/j.sbi.2015.02.004). — **IF WoS 6.908** - (1,2/1,0).
- A26. Comparative Computational Study of Interaction of C60-Fullerene and Tris-Malonyl-C60-Fullerene Isomers with Lipid Bilayer: Relation to Their Antioxidant Effect / M. E. Bozdaganyan, P. S. Orekhov, A. K. Shaytan, K. V. Shaitan // *PLoS ONE* / ed. by C. M. Soares. — 2014. — July 14. — Vol. 9, no. 7. — e102487. — DOI: [10.1371/journal.pone.0102487](https://doi.org/10.1371/journal.pone.0102487). — **IF WoS 2.740** - (0,9/0,1).
- A27. Analysis of the Mechanism of Nucleosome Survival during Transcription / H.-W. Chang, O. I. Kulaeva, A. K. Shaytan, M. Kibanov, K. Kuznedelov, K. V. Severinov, M. P. Kirpichnikov, D. J. Clark, V. M. Studitsky // *Nucleic Acids Research*. — 2014. — Feb. — Vol. 42, no. 3. — P. 1619–1627. — DOI: [10.1093/nar/gkt1120](https://doi.org/10.1093/nar/gkt1120). — **IF WoS 11.501** - (1,0/0,2).
- A28. Voltage-Gated Ion Channel Modulation by Lipids: Insights from Molecular Dynamics Simulations / M. A. Kasimova, M. Tarek, A. K. Shaytan, K. V. Shaitan, L. Delemotte // *Biochimica et Biophysica Acta (BBA) - Biomembranes*. — 2014. — May. — Vol. 1838, no. 5. — P. 1322–1331. — DOI: [10.1016/j.bbamem.2014.01.024](https://doi.org/10.1016/j.bbamem.2014.01.024). — **IF WoS 3.4** - (1,2/0,1).
- A29. *Nishi, H.* Physicochemical Mechanisms of Protein Regulation by Phosphorylation / H. Nishi, A. Shaytan, A. R. Panchenko // *Frontiers in Genetics*. — 2014. — Aug. 7. — Vol. 5. — DOI: [10.3389/fgene.2014.00270](https://doi.org/10.3389/fgene.2014.00270). — **IF WoS 3.789** - (1,2/0,4).

- A30. Genome Packaging in EL and Lin68, Two Giant phiKZ-like Bacteriophages of *P. Aeruginosa* / O. Sokolova, O. Shaburova, E. Pechnikova, A. Shaytan, S. Krylov, N. Kiselev, V. Krylov // *Virology*. — 2014. — Nov. — Vol. 468—470. — P. 472—478. — DOI: [10.1016/j.virol.2014.09.002](https://doi.org/10.1016/j.virol.2014.09.002). — **IF WoS 2.464** - (0,8/0,1).
- A31. Peptide Nanofibrils Boost Retroviral Gene Transfer and Provide a Rapid Means for Concentrating Viruses / M. Yolamanova, C. Meier, A. K. Shaytan, V. Vas, C. W. Bertoncini, F. Arnold, O. Zirafi, S. M. Usmani, J. A. Müller, D. Sauter, C. Goffinet, D. Palesch, P. Walther, N. R. Roan, H. Geiger, O. Lunov, T. Simmet, J. Bohne, H. Schrezenmeier, K. Schwarz, L. Ständer, W.-G. Forssmann, X. Salvatella, P. G. Khalatur, A. R. Khokhlov, T. P. J. Knowles, T. Weil, F. Kirchhoff, J. Münch // *Nature Nanotechnology*. — 2013. — Feb. — Vol. 8, no. 2. — P. 130—136. — DOI: [10.1038/nnano.2012.248](https://doi.org/10.1038/nnano.2012.248). — **IF WoS 31.538** - (0,8/0,2).
- A32. Influence of Interionic Interactions on Functional State and Blocker Binding of Voltage-Gated Potassium Channels / K. V. Shaitan, O. S. Sokolova, A. K. Shaitan, M. A. Kasimova, V. N. Novoseletskii, M. P. Kirpichnikov // *Moscow University Biological Sciences Bulletin*. — 2013. — Mar. — Vol. 68, no. 1. — P. 8—14. — DOI: [10.3103/S0096392513010057](https://doi.org/10.3103/S0096392513010057). — **IF RINC 0.76** - (0,8/0,1).
- A33. *Orekhov, P. S.* Calculation of Spectral Shifts of the Mutants of Bacteriorhodopsin by QM/MM Methods / P. S. Orekhov, A. K. Shaytan, K. V. Shaitan // *Biophysics*. — 2012. — Mar. — Vol. 57, no. 2. — P. 144—152. — DOI: [10.1134/S0006350912020170](https://doi.org/10.1134/S0006350912020170). — **IF SJR 0.226** - (1,0/0,2).
- A34. Self-Assembling Nanofibers from Thiophene–Peptide Diblock Oligomers: A Combined Experimental and Computer Simulations Study / A. K. Shaytan, E.-K. Schillinger, P. G. Khalatur, E. Mena-Osteritz, J. Hentschel, H. G. Börner, P. Bäuerle, A. R. Khokhlov // *ACS Nano*. — 2011. — Sept. 27. — Vol. 5, no. 9. — P. 6894—6909. — DOI: [10.1021/nn2011943](https://doi.org/10.1021/nn2011943). — **IF WoS 13.7** - (1,8/1,0).
- A35. Self-Organizing Bioinspired Oligothiophene–Oligopeptide Hybrids / A. K. Shaytan, E.-K. Schillinger, E. Mena-Osteritz, S. Schmid, P. G. Khalatur, P. Bäuerle, A. R. Khokhlov // *Beilstein Journal of Nanotechnology*. —

2011. — Sept. 5. — Vol. 2. — P. 525–544. — DOI: [10.3762/bjnano.2.57](https://doi.org/10.3762/bjnano.2.57). — **IF WoS 2.44** - (2,3/2,0).
- A36. *Грешнова, А. А.* Синтетическая Биология: Конструирование Живого / А. А. Грешнова, Г. С. Глухов, А. К. Шайтан // Химия и Жизнь - XXI век. — 2019. — № 9. — С. 32–38. — (1,0/0,5).
- A37. Моделирование Нуклеосом Методами Атомистического Суперкомпьютерного Моделирования / Г. А. Армеев, В. Н. Блинов, А. А. Севенюк, В. М. Студитский, А. К. Шайтан // Journal of Qafqaz University - Physics. — 2013. — Т. 1, № 2. — С. 3–11. — (1,0/0,5).
- A38. Integrative Modeling of Nucleosomes and Supranucleosomal Structures / G. A. Armeev, A. R. Panchenko, A. V. Feofanov, A. K. Shaytan // Biophysical Journal. — 2020. — Т. 118, № 3. — DOI: [10.1016/j.bpj.2019.11.527](https://doi.org/10.1016/j.bpj.2019.11.527). — **IF WoS 3.665** - (0,03/0,01).
- A39. Analyzing Nucleosome Plasticity via Atomistic MD Simulations / A. S. Kniazeva, G. A. Armeev, I. Pospelova, A. K. Shaytan // Biophysical Journal. — 2020. — Т. 118, № 3. — DOI: [10.1016/j.bpj.2019.11.3375](https://doi.org/10.1016/j.bpj.2019.11.3375). — **IF WoS 3.665** - (0,03/0,02).
- A40. Analyzing Nucleosome Plasticity Using Molecular Dynamics Simulations / G. A. Armeev, I. Pospelova, A. Knyazeva, G. A. Komarova, A. K. Shaytan // FEBS Open Bio. — 2019. — Т. 9, sup1. — С. 19. — DOI: [10.1002/2211-5463.12674](https://doi.org/10.1002/2211-5463.12674). — **IF WoS 2.231** - (0,06/0,05).
- A41. *Gribkova, A. K.* Construction and Analysis of an Interactome between Nucleosomes and Chromatin Proteins / A. K. Gribkova, A. K. Shaytan // FEBS Open Bio. — 2019. — Т. 9, sup1. — С. 393. — DOI: [10.1002/2211-5463.12675](https://doi.org/10.1002/2211-5463.12675). — **IF WoS 2.231** - (0,06/0,02).
- A42. *Armeev, G. A.* Python Toolkit for DNA Geometry Analysis and Modeling / G. A. Armeev, I. Sukhanova, A. K. Shaytan // FEBS Open Bio. — 2019. — Т. 9, sup1. — С. 150. — DOI: [10.1002/2211-5463.12675](https://doi.org/10.1002/2211-5463.12675). — **IF WoS 2.231** - (0,06/0,01).

- A43. Mutual Influence of Linker Histone H1 Amino Acid Sequence and DNA Nucleotide Sequence on the Linker Histone H1 Position in Chromatosome / T. K. Gorkovets, G. A. Armeev, K. V. Shaitan, A. K. Shaytan // FEBS open bio. — 2018. — Vol. 8, Supplement S1. — P. 346–346. — DOI: [10.1002/2211-5463.12453](https://doi.org/10.1002/2211-5463.12453). — **IF WoS 2.231** - (0,06/0,02).
- A44. Microsecond Molecular Dynamics Simulations of Nucleosomes: Implications for Nucleosome Function / A. K. Shaytan, G. A. Armeev, A. Goncarencu, V. B. Zhurkin, D. Landsman, A. R. Panchenko // Biochemistry and Cell Biology. — 2017. — Apr. — Vol. 95. — P. 179–184. — DOI: [10.1139/bcb-2017-0043](https://doi.org/10.1139/bcb-2017-0043). — **IF WoS 2.46** - (0,03/0,03).
- A45. Nucleosome Dynamics at Microsecond Timescale: DNA-Protein Interactions, Water-Mediated Interactions and Nucleosome Formation / A. K. Shaytan, G. A. Armeev, A. Goncarencu, V. B. Zhurkin, D. Landsman, A. R. Panchenko // Biophysical Journal. — 2016. — Feb. — Vol. 110, no. 3. — 405a. — DOI: [10.1016/j.bpj.2015.11.2189](https://doi.org/10.1016/j.bpj.2015.11.2189). — **IF WoS 3.665** - (0,03/0,03).
- A46. Combined Influence of Linker DNA and Histone Tails on Nucleosome Dynamics as Revealed by Microsecond Molecular Dynamics Simulations / A. K. Shaytan, G. A. Armeev, A. Goncarencu, V. B. Zhurkin, D. Landsman, A. R. Panchenko // Journal of Biomolecular Structure and Dynamics. — 2015. — May 18. — Vol. 33, sup1. — P. 3–3. — DOI: [10.1080/07391102.2015.1032630](https://doi.org/10.1080/07391102.2015.1032630). — **IF WoS 3.220** - (0,12/0,12).
- A47. Polymorphism of Histone Tail Interactions in Nucleosome / A. K. Shaytan, G. A. Armeev, V. B. Zhurkin, D. Landsman, A. R. Panchenko // Biophysical Journal. — 2015. — Jan. — Vol. 108, no. 2. — 72a. — DOI: [10.1016/j.bpj.2014.11.431](https://doi.org/10.1016/j.bpj.2014.11.431). — **IF WoS 3.665** - (0,03/0,03).
- A48. Investigation of Ion Permeation through the Cx26 Hemichannel / M. Kasimova, A. Shaytan, K. Shaitan, M. Tarek // Biophysical Journal. — 2014. — Jan. — Vol. 106, no. 2. — 556a. — DOI: [10.1016/j.bpj.2013.11.3093](https://doi.org/10.1016/j.bpj.2013.11.3093). — **IF WoS 3.665** - (0,03/0,003).
- A49. Molecular Insights into Electroporation and Electrotransfer through Model Cell Membranes / M. Kasimova, L. Delemotte, A. Polak, D. Miklavcic, M. Casciola, F. Apollonio, M. Breton, L. Mir, A. Shaytan, K. Shaitan,

- M. Tarek // *Biophysical Journal*. — 2014. — Jan. — Vol. 106, no. 2. — 291a. — DOI: [10.1016/j.bpj.2013.11.1701](https://doi.org/10.1016/j.bpj.2013.11.1701). — **IF WoS 3.665** - (0,03/0,003).
- A50. Pausing as a Mechanism of Nucleosome Recovery / H.-W. Chang, O. I. Kulaeva, A. Shaytan, M. Kibanov, K. Severinov, D. J. Clark, V. M. Studitsky // *Epigenetics & Chromatin*. — 2013. — Mar. — Vol. 6, S1. — DOI: [10.1186/1756-8935-6-S1-P14](https://doi.org/10.1186/1756-8935-6-S1-P14). — **IF WoS 5.333** - (0,1/0,01).
- A51. *Armeev, G. A.* Abstract OR-19: Integrative Approach on Nucleosome Complexes Modeling / G. A. Armeev, A. R. Panchenko, A. K. Shaytan // *International Journal of Biomedicine*. — 2019. — T. 9, s1. — С. 13—14. — **IF WoS 2.231** - (0,1/0,03).
- A52. *Bass, M. V.* Abstract P-20: Linker Histone H1: The Interplay between Chromatosome Stability, Oncomutations and Post-Translational Modifications / M. V. Bass, G. A. Armeev, A. K. Shaytan // *International Journal of Biomedicine*. — 2019. — T. 9, S1. — С. 25. — **IF WoS 2.231** - (0,1/0,03).
- A53. *Shaytan, A. K.* Water Mediated Interactions in Nucleosomes / A. K. Shaytan, A. R. Panchenko // *Abstracts of Papers of the American Chemical Society*. — 2014. — Aug. 10. — Vol. 248. — **IF WoS 2.231** - (0,1/0,05).
- A54. Structural Analysis of the Key Intermediate Formed during Transcription through a Nucleosome / H.-W. Chang, A. K. Shaytan, F.-K. Hsieh, O. I. Kulaeva, M. P. Kirpichnikov, V. M. Studitsky // *Trends in Cell & Molecular Biology*. — 2013. — Vol. 8. — P. 13—23. — (2,0/0,5).
- A55. *Заявка 2580006 Рос. федерация, МПК G 06 F 19/100.* Способ скрининга потенциальных противоопухолевых препаратов ингибиторов ФАСТ [Текст] / В. М. Студитский, О. И. Студитская, А. К. Шайтан (Российская Федерация). — № 2013132806/10 ; заявл. 16.07.2013 ; опубл. 27.01.2015, Бюл. № 3 ; приоритет 16.07.2013 (Рос. Федерация). — 10 с.
- A56. *Свидетельство о гос. регистрации программы для ЭВМ.* Программный комплекс реконструкции пространственной структуры белков и комплексов на основе карт электронной плотности низкого разрешения [Текст] / Д. Л. Шуров, А. К. Шайтан, Г. А. Армеев, Д. А. Турченков, В. Н. Блинов, М. П. Кирпичников, К. В. Шайтан. — № 2013614397 ; заявл. 13.05.2013 ; опубл. 17.07.2013, 2013614397 (Рос. Федерация).



## Глава 1. Методы молекулярного и интегративного моделирования в структурной биологии<sup>1</sup>

Данная глава является вводной и обзорной, в ней обсуждается место методов молекулярного компьютерного моделирования в исследовании структуры и динамики биомакромолекулярных комплексов. Вводится понятие интегративного моделирования. Обсуждаются вопросы постановки задач моделирования при наличии экспериментальных данных разной степени детальности и информационного содержания. Более подробно обсуждаются методы активно применяемые автором, в частности, методы атомистической молекулярной динамики, методы расчета свободной энергии, методы огрубленного моделирования, в частности методы моделирования ДНК в динуклеотидном приближении. Кратко обсуждается ряд экспериментальных методов, данные которых использовались в ходе работ излагаемых в последующих главах диссертации.

Содержание данной главы основано на обобщении результатов по моделированию и изучению различных биологических систем, отраженного в ряде публикаций [A1—A35; A38; A42; A51]. Материалы главы также включают материалы диссертации на соискание степени Dr. rer. nat. в Университете г. Ульм, Германия [5], кандидатской диссертации Г.А. Армеева [6], руководителем которой являлся автор. Описание части методов приведено согласно кандидатской диссертации автора [7].

### 1.1 Введение

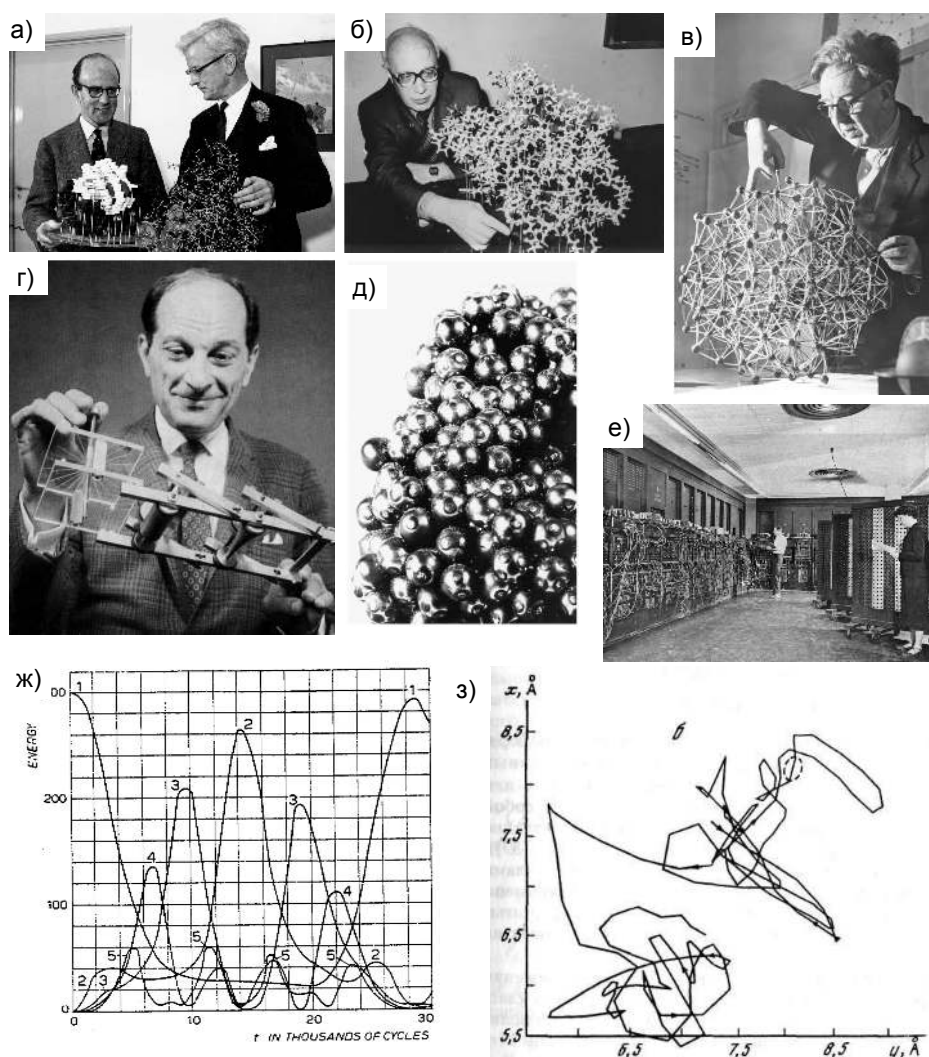
Прежде, чем приступить к более детальному обсуждению, уделим внимание обсуждению и определению ключевых понятий, вынесенных в название данной главы, а именно, “структурная биология”, “молекулярное моделирование” и “интегративное моделирование”, в том числе в историческом контексте.

---

<sup>1</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [A1—A54].

Журнал Nature дает следующее определение структурной биологии: “Структурная биология – это изучение молекулярных структур и динамики биологических макромолекул, в особенности белков и нуклеиновых кислот, а также изучение того, как изменения в их структуре влияют на функцию” [8]. Данное определение, тем не менее, не дает ответа на вопрос о том, что мы понимаем под структурой биологической макромолекулы. В биологии активно используется подразделение уровней структурной организации биомолекул на первичный (химический) уровень, и последующие уровни (вторичный, третичный, четвертичный), которые описывают укладку в пространстве биомолекул. Первые технологии секвенирования — чтения последовательности мономеров в биополимерах появились в 1950-1970-ых годах. К 1951 году при участии Фредерика Сенгера была разработана технология определения первичной последовательности белков [9], к 1964 Робертом Холли определена первичная последовательность молекулы транспортной РНК [10], к 1973 году Аланом Максамом и Вальтером Гилбертом была определена последовательность ДНК *lac* оператора [11], а к 1977 году ими [12], а также Ф. Сенгером [13] были развиты рутинные методы секвенирования ДНК, которые используются до сих пор. Среди работ отечественных ученых важный вклад внесла группа академика Александра Александровича Баева – к 1967 году была впервые определена первичная структура валиновой транспортной РНК пекарских дрожжей [14]. Несмотря на успехи определения первичной структуры биополимеров, именно, пространственная структурная и динамическая организация биомолекул является ключевой для понимания механизмов их функционирования (что описывается термином “mechanistic insight”, часто встречающимся в англоязычной литературе). В силу исторических и технических причин структура биомолекул обычно воспринимается как некоторая трехмерная модель, отображающая укладку биомолекул (белков, нуклеиновых кислот, липидов и т.д.) в пространстве. В некотором роде такой подход продолжает парадигму принятую в химии для описания структуры химических соединений в виде взаимного расположения атомов и связей между ними. На Рис. 1.1а приведены примеры белковых структур в руках основоположников белковой кристаллографии Макса Перутца и Джона Кендрю, впервые разрешивших структуру миоглобина в 1958 году [15]. На Рис. 1.1б модели белков в 1980-ых годах в руках основоположника отечественной белковой кристаллографии академика Бориса Константиновича Вайнштейна [16].

В компьютерную эпоху модели биомакромолекул принято представлять в виде файлов (например, в формате PDB) с описанием пространственных координат отдельных атомов. Следует однако отметить, что такое представление моделей позволяет описать лишь некоторые статичные конфигурации атомов. Насколько такой подход позволяет действительно представить и описать происходящие процессы в микромире? В реальном мире молекулы не являются некоторыми статическими структурами, а находятся в постоянном динамическом изменении. Для описания поведения газов, жидкостей, мягкой материи (soft matter), разупорядоченных и слабоупорядоченных систем в физике были разработаны формализмы термодинамики и статистической физики (см. Рис. 1.2). В этих подходах, приходится отказаться от описания молекулярных процессов в виде статических структур в пользу некоторых средних характеристик, ансамблей состояний, вероятностного описания различных микросостояний согласно статистике Гиббса-Больцмана. В микромире наличие, например, как у глобулярных белков более менее детерминированных пространственных структур является скорее счастливым исключением. В то же время при попытке описания многих биологических систем парадигма представления структуры в виде даже ограниченного набора статических структур испытывает значительные трудности. С одной стороны здесь можно упомянуть неструктурированные белки (англ. intrinsically disordered proteins), у которых отсутствует какая-то одна статическая структура. С другой стороны подобные проблемы характерны при попытке описания многих комплексных биологических образований, например, структуры хроматина или даже структуры нуклеосомы. В этой связи высказывается точка зрения, что понятие структуры биомолекулы следует трактовать более широко, в идеальном случае как ансамбль возможных конформаций данной биомолекулы с указанием статистических весов различных конформаций, а также их зависимости от внешних условий [17]. Такое понимание структуры биомакромолекулярных комплексов наиболее близко связано с пониманием их функционирования и, по-видимому, будет являться основой для познания функционирования живых систем. Развивая данную концепцию, стоит упомянуть о том, что большинство биологических объектов на молекулярном уровне являются неравновесными системами, в которых происходят процессы превращения энергии. Для описания таких систем не достаточно подходов только равновесной статистической физики, а требуется развитие подходов описания, так называемой, активной материи (active matter) [18].



**Рисунок 1.1** — Иллюстрации к истории структурной биологии и моделирования. а) М. Перуц (слева) и Д. Кендрю (справа) со структурными моделями миоглобина. б) Б.К. Вайнштейн со структурой белка [16], в) Дж. Д. Бернал с моделью укладки атомов в жидкости [19], г) С. Улам с тележкой Монте-Карло Энрико Ферми [20], д) результаты физического моделирования жидкости в виде стальных шаров в работах Д.Д. Бернала [19], е) первый программируемый компьютер общего назначения ENIAC [21], ж) изменение полной энергии со временем для первых пяти мод вибрирующей струны в задаче Ферми-Пасты-Улама-Цингу, первая работа по исследованию статистических свойств механических систем методом молекулярной динамики, 1955 г. [22], (з) Траектории движения атомов в жидкости, рассчитанные методом МД из кандидатской диссертации А.Г. Гривцова, начало 1970-ых годов [23].

Сам по себе термин *молекулярное моделирование* (*molecular modeling*) является достаточно широким понятием, значение которого постепенно расширяется и видоизменяется с развитием различных новых методов. В одних

случаях упор может делаться непосредственно на задачу создания модели (например, пространственной модели белка), в других случаях на анализ и исследование свойств молекул на основе физических моделей поведения и взаимодействия атомов. К первому типу задач можно отнести традиционные задачи реконструкция структуры биомолекул по экспериментальным данным. Данный процесс всегда включает в себя шаг создания модели, которая удовлетворяет этим данным. На заре структурной биологии реконструкция биомолекул по результатам кристаллографических экспериментов долгое время сводилась к построению пространственных моделей из проволоки, дерева, других материалов (Рис. 1.1аб). Позднее данный процесс был значительно автоматизирован с применением компьютерных технологий, разработаны соответствующие программные пакеты по автоматизации построения моделей и их визуализации (например, CCP4 [24], O, Phenix [25], Coot [26], PyMol [27], Chimera [28]). Построение моделей молекул в виде пространственного расположения атомов, удовлетворяющего набору экспериментальную данных, - одно из возможных значений термина компьютерное молекулярное моделирование. В то же время со времен возникновения молекулярно-кинетической теории (основы которой были в том числе заложены М.В. Ломоносовым [29]) ученые стремились к описанию и предсказанию свойств молекулярных систем на основе учета взаимодействий между отдельными атомами, описываемых законами физики. В до-компьютерную эпоху такие предсказания можно было делать либо на основе теоретических расчетов, либо на основе физических механических моделей. Так, например, в 1950-60-ых годах для молекулярного моделирования жидкостей Джон Десмонд Бернал активно использовал механические модели в виде металлических и пластиковых шариков [30](Рис. 1.1вд)<sup>2</sup>. С развитием компьютерных технологий появились возможности численного моделирования поведения молекулярных систем на основе их физических моделей. Следует отметить, что в английском языке для описания такого рода деятельности активно используется термин *molecular simulations* (см. например [33]), который на русский язык обычно переводится также “молекулярное моделирование”, однако слово “simulations” имеет важные отличия в своей смысловой нагрузке. Под этим термином понимается не столько создание моделей, сколько изу-

<sup>2</sup>Любопытным также является факт, что тот же Дж. Д. Бернал является одним из основателей области биомолекулярной кристаллографии, под его руководством работали классики структурной биологии Макс Перутц, Дороти Ходжкин, Розалинд Франклин [31], а также в честь него назван пакет по молекулярному моделированию DESMOND [32].

чение и анализ теоретических моделей с помощью вычислительных методов, так называемые “вычислительные эксперименты”. В области инженерного моделирования в русском языке для обозначения термина “simulation modeling” используется термин *имитационное моделирование*, однако он не вошел в обиход в среде исследователей, занимающихся молекулярным моделированием. Подход к изучению различных систем на основе проведения вычислительных экспериментов на данный момент многими воспринимается как одна из парадигм научного познания наравне с экспериментальным и теоретическими подходами [34] и получил название *in silico* подхода. В случае изучения свойств вещества данный подход предполагает наличие определенной физической модели исследуемого объекта, которая описывает взаимодействия между его составными частями, и использует различные вычислительные алгоритмы, чтобы исследовать свойства данной модели вещества. Если для описание физической модели вещества используются различные приближения, основанные на законах квантовой механики, то такие подходы зачатую называют *ab initio* подходами (от лат. “из первых принципов” - подразумевая, что для моделирования необходимо знать лишь базовые физические константы). Для моделирования многих задач и веществ такой формализм является слишком вычислительно затратным, поэтому исторически область “molecular simulations” развивалась в первую очередь на основе формализма классической механики (частицы вещества представлялись в виде материальных точек, взаимодействующих по законам классической механики).

Зарождение области вычислительных экспериментов в молекулярном моделировании происходило в середине-конце 1940-ых годов во время работы над созданием ядерного и термоядерного оружия в США и связано в первую очередь с именами Станислава Улама, Джона фон Неймана, Николаса Метрополиса, а также с вводом в строй в 1945 году первого электронного цифрового вычислителя общего назначения ENIAC (Рис. 1.1e). С. Улам и Д. фон Нейман, работая над проблемой транспорта нейтронов в различных типах ядерных систем, предложили в 1947 году моделировать диффузию нейтронов путем многократного численного расчета траекторий движения отдельных нейтронов с учетом случайности их взаимодействия с окружающим веществом. Таким образом, решение задачи о вычислении средних характеристик данного процесса сводилось к статистической выборке (statistical sampling), а не к полному численному решению уравнений, которые описывали этот процесс. Благода-

ря наличие компьютера ENIAC стало возможным написать компьютерную программу для соответствующих расчетов, которая использовала в том числе генератор псевдослучайных чисел для моделирования случайных событий и соответствующих статистических выборок [35]. Подходы основанные на статистическом моделировании с использованием случайных величин были названы ими методами Монте-Карло, по предложению Николаса Метрополиса, а первая статья по этому методу опубликована в 1949 году [36]. Любопытным является факт, что физик Энрико Ферми еще до формулировки метода Монте Карло Уламом и фон Нейманом использовал метод статистических выборок для своих задач, проводя расчеты вручную в предрассветные часы, когда страдал бессонницей [37]. В 1947 году, когда постановка расчетов на компьютере ENIAC была временно невозможной из-за его перевозки, Ферми придумал аналоговый компьютер, который можно было использовать для моделирования транспорта нейтронов. Данный компьютер представлял собой тележку с карандашом, которую нужно было передвигать по бумажному чертежу и периодически изменять ее направление на основе случайных чисел, таким образом моделируя передвижения нейтронов и их столкновения с атомами среды. Данное устройство получило название тележки Монте-Карло или FERMIAC (см Рис. 1.1г).

Следующей важной вехой в развитии методов Монте Карло в молекулярном моделировании и статистической механике явилась разработка Метрополисом, супругами Розенблют и Теллер в 1953 году метода Монте Карло на основе марковских цепей (Markov Chain Monte Carlo, MCMC) для изучения уравнений состояния вещества состоящего из индивидуальных взаимодействующих молекул [38]. В этой работе авторы провели исследование двумерного газа из жестких сфер на компьютере MANIAC. Инновацией данного метода стало вычисление статистической выборки состояний согласно распределению Больцмана в многомерном пространстве через построение марковской цепи состояний. Случайные переходы между состояниями принимались или отклонялись согласно критерию Метрополиса-Хастингса таким образом, что лимитирующее распределение состояний в данной марковской цепи соответствовало Больцмановскому распределению состояний по энергиям. На идеях данной работы основано большинство подходов по моделированию молекулярных систем методами Монте Карло и в наши дни [39].

1950-ые годы также ознаменовались первыми работам в области применения метода молекулярной динамики - численного решения уравнений движения

классической механики. Впервые Ферми, Паста, Улам и Цингу исследовали колебания струны с нелинейной упругостью и показали наличие необычной периодичности возникающей в нелинейных системах [22]. Идея моделирования динамики молекулярных систем была сформулирована Олдером и Вайнрайтом в 1957 году – с помощью компьютера UNIVAC исследовались фазовые переходы в системе твердых сфер [40]. К первым попыткам применить методы молекулярной динамики к реалистичным системам можно отнести работу Гибсона и др. 1960 года по моделированию радиационного поражения меди [41] и работу Анисура Рамана по моделированию жидкого аргона 1964 года [42]. В нашей стране первые шаги по развитию методов молекулярной динамики связаны с деятельностью А.А. Гривцова, Э.Э. Шноля, Н.К. Балабаева из Института прикладной математики (ИПМ, Москва) начиная с 1967 года [23]. С тех пор методы молекулярной динамики стали активно применяться для моделирования газов, жидкостей [43], а позднее стали активно применяться и для моделирования биомолекулярных систем. Развитие применения методов молекулярного моделирования к биомолекулярным системам связаны с деятельностью групп Мартина Карплуса, Арье Варшеля, Майкла Левитта, которые получили в 2013 году нобелевскую премию за “разработку мультимасштабных моделей сложных химических систем”, а также групп Питера Коллмана, Хермана Берендсена, Клауса Шультена, Дэвида Кейса, Эрика Линдаля, Алекса МакКерела и др. Работы группа М. Карплуса легли в основу набора силовых полей CHARMM, а также одноименной программы для моделирования [44]. На основе подходов, силовых полей и программного кода CHARMM были развиты многие родственные программы, например, программа для параллельных расчетов молекулярной динамики NAMD (в группе К. Шультена) [45], программа для моделирования на основе данных ЯМР XPLOR [2], программа для моделирования по гомологии MODELLER [46]. Работы группы Питера Коллмана положили основу программному пакету AMBER и одноименной группе силовых полей [47]. Работы группы Хермана Берендсена в Голландии заложили основы пакета по моделированию GROMACS [48] и силового поля GROMOS [49]. В нашей стране развитие программных продуктов в области молекулярной динамики связано в том числе с разработкой программы PUMA и ее версий под руководством Н.К. Балабаева [50].

В первые десятилетия XXI века область молекулярного моделирования продолжала активно развиваться и совершенствоваться, в том числе и в мето-



дологическом плане. Стоит отметить адаптацию программного обеспечения для использования массивно-параллельных архитектур графических процессоров, что позволило уверенно проводить расчеты МД в микросекундном диапазоне не только на суперкомпьютерах, но и отдельных рабочих станциях [51]. Усилиями нескольких групп создавались специализированные компьютерные платформы для расчета молекулярной динамики, в частности архитектуры MD GRAPE в Японии [52] и компьютеры ANTON компании D.E. Shaw Research [53]. Последние позволяют проводить вычисления в миллисекундном диапазоне времен моделируемых систем.

Серьезный прогресс был достигнут и в совершенствовании методов повышения эффективности статистических выборок (enhanced sampling methods) в ходе расчетов молекулярной динамики. К таким методам можно отнести методы метадинамики (metadynamics) [54], адаптивной смещающей силы (adaptive biasing force) [55], адаптивно смещаемой молекулярной динамики (adaptively biased molecular dynamics) [56], динамики с обменом репликами (replica exchange molecular dynamic/parallel tempering) [57], управляемой молекулярной динамики (steered molecular dynamics) [58]. Многие из этих методов получили активное распространение благодаря созданию кросс-платформенного дополнения к пакетам молекулярной динамики PLUMED [59]. Были в том числе развиты теоретические основы, связывающие характеристики расчета неравновесных молекулярных процессов с равновесными параметрами моделируемых системы. К таковым можно отнести равенство Джарзинского, связывающее работу в ходе неравновесного процесса над системой с разницей свободной энергии между конечным и начальным состояниями [60] и флуктуационную теорему Крукса (Crooks fluctuation theorem)[61]. Любопытно, что открытие и формулировка данных достаточно лаконичных и простых соотношений неравновесной статистической физики произошла только на рубеже XX и XXI веков, хотя явные предпосылки для их формулировки существовали со времен Больцмана. Отдельно стоит отметить развитие и успехи различных эмпирических подходов к моделированию структуры и дизайну белков. В отличие от методов, исходящих в первую очередь из физических представлений о структуре, подвижности и взаимодействии атомов в биомолекулах (physics-based methods), такие подходы активно используют различные эмпирические скоринговые функции для оценки вероятности той или иной конформации белка, которые основываются на статистическом анализе большого количества уже

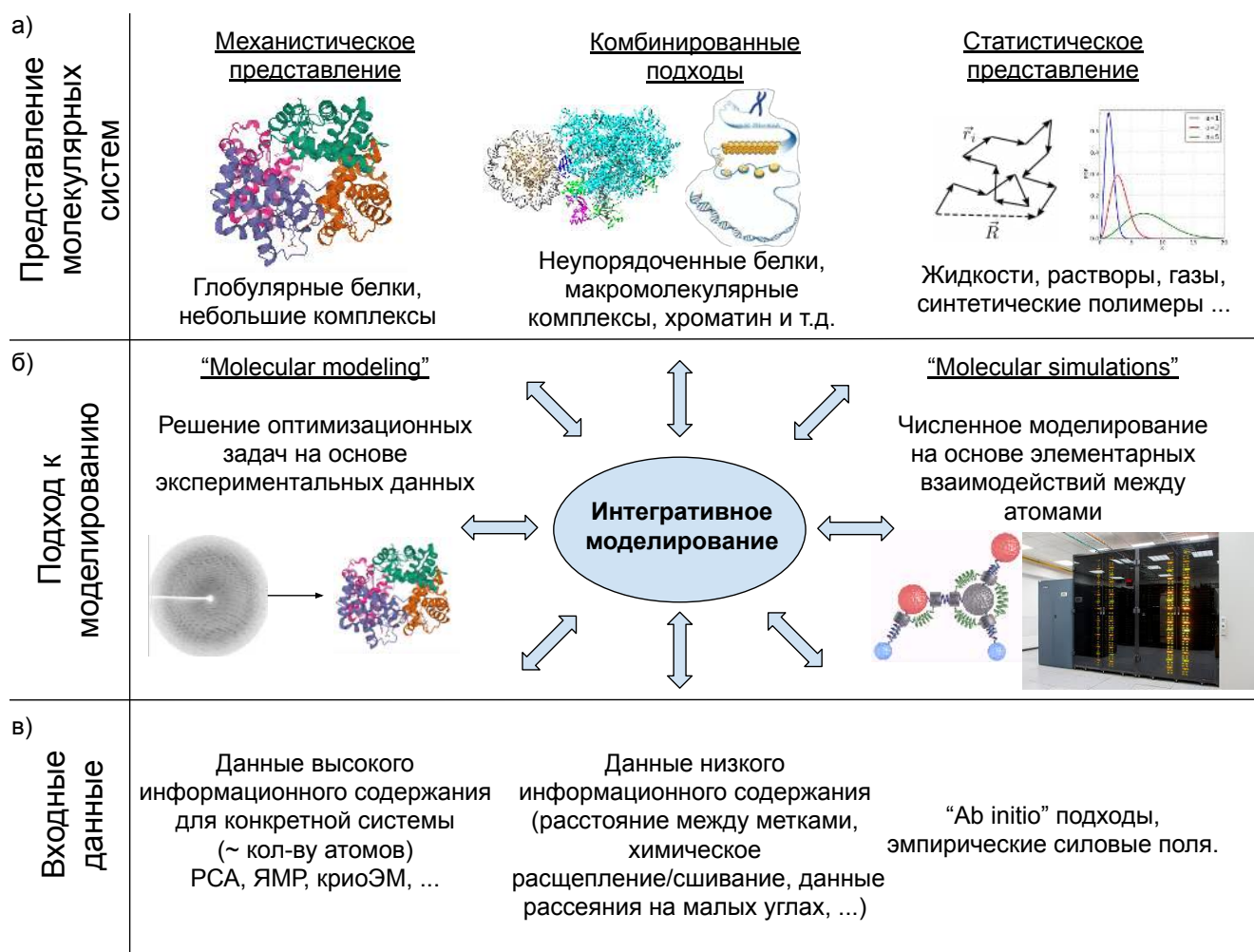
известных пространственных структур белков, а также эволюционном анализе похожести различных аминокислотных мотивов в белках. Перебор конформационного пространства полипептидной цепи также зачастую ограничивают набором конформаций фрагментов присутствующих в уже известных структурах белков. К таким методам можно отнести подходы основанные на пакете ROSETTA [62], AWSEM-MD [63] и др.

Выше мы обсудили в историческом контексте различные взгляды на задачи и методы молекулярного моделирования. В одном случае речь идет о задачах поиска оптимальной пространственной модели молекулы на основе большого количества экспериментальных данных (например, реконструкции структуры белка по набору рефлексов картины дифракции его кристаллов), в другом случае упор делается на имитационное моделирование (molecular simulations) систем на основе физических моделей взаимодействий между атомами, которые иногда называют также *ab initio* подходами (напр. *ab initio protein folding*<sup>3</sup> и т.д.). Разные аспекты данных подходов просуммированы на рисунке 1.2, они отличаются по наличию, качеству и информационному содержанию используемых экспериментальных данных (Рис. 1.2в), по используемым математическим подходам и физическим теориям (Рис. 1.2б) и по самому представлению моделей (Рис. 1.2а) - в одном случае речь обычно идет о некоторой статической структуре или наборе структур, в другом о статистических ансамблях различных конформаций.

В то же время в структурной биологии существует большое количество задач, которые не решаются в рамках одного из двух вышеописанных подходов. С одной стороны для многих систем, экспериментальные данные могут присутствовать лишь в ограниченном количестве и зачастую иметь опосредованное отношение к деталям молекулярной структуры (напр. данные спектроскопических экспериментов, где оценивается взаимодействие введенных в структуру меток, данные по химической доступности и реакционной способности различных групп, данные малоуглового рассеяния и т.д.). С другой стороны, решение структурных задач напрямую *in silico* методами физического моделирования на основе взаимодействия частиц невозможно для большинства систем в силу чрезвычайной вычислительной сложности и зачастую качества имеющихся моделей взаимодействия атомов в биомакромолекулах. Отдельную проблему

<sup>3</sup>Не стоит путать с пониманием термина *ab initio* в квантовой химии, где под этим понимается моделирование на основе решения уравнения Шредингера.

представляют задачи описания структуры биомолекул, у которых нет явной статической структуры, а она представлена ансамблем конформаций. К примерам таких систем можно отнести внутренне разупорядоченные белки (англ. *intrinsically disordered proteins*), а также большинство биомолекулярных комплексов (вплоть до структуры хроматина на масштабах клеточного ядра) – по мере роста размеров комплексов, они все реже формируют какую-то одну статическую конформацию, а их функционирование завязано сложную внутри- и межмолекулярную динамику. Для описания таких молекул динамический формализм используемый в имитационном молекулярном моделировании (*molecular simulations*) подходит значительно лучше, однако зачастую решение таких задач из первых принципов опять сводится к вышеописанным затруднениям.



**Рисунок 1.2** — Диаграмма, суммирующая различные подходы к описанию, моделированию и изучению структуры и динамики биомолекул.

Таким образом, актуальным является разработка гибридных методов и подходов по моделированию биомолекулярных комплексов, которые

сочетают возможности построения структурных моделей как на основе экспериментальных данных (в том числе данных с низким информационным содержанием), так и на основе моделирования атом-атомных взаимодействий, а также учитывают структурные динамические особенности многих биомакромолекулярных систем. Такие подходы в последнее время в литературе получили название *интегративного моделирования* (от англ. integrative modeling, см. напр. [1]). Одной из задач интегративного моделирования является в том числе интеграция различного рода экспериментальных данных при построении моделей биомакромолекулярных систем. На Рис. 1.3 представлены различные типы экспериментальных данных, которые могут быть использованы в интегративном моделировании. К ним могут относиться структуры биомакромолекул или отдельных их частей полученные традиционными методами структурной биологии с высоким разрешением, а также набор данных низкого информационного содержания (когда данных недостаточно для реконструкции структуры традиционными методами), полученных различными биофизическими, биохимическими и другими методами. К последним могут относиться данные полученные методами электронной микроскопии низкого разрешения, данные малоуглового рассеяния рентгеновских лучей или нейтронов, данные о взаимодействии введенных в структуру меток, получаемые методами ЯМР, ЭПР, спектроскопии Ферстеровского резонансного переноса энергии (FRET spectroscopy), данные по экспонированности в растворитель различных групп биомакромолекул, получаемые на основе экспериментов по дейтериевому обмену, данные экспериментов по химическому сшиванию или разрушения биомакромолекул, данные получаемые с применением методов геномики (например, методов ChIP-seq, Hi-C и др.). Актуальной является и задача обратного рода – имея некоторую структурно-динамическую модель биомакромолекулярной системы, рассчитать параметры ее отклика в различных типах экспериментов (см. Рис. 1.3). Обсуждаемые подходы также важны при дизайне и конструировании новых биомакромолекулярных комплексов с полезными функциями.

Данная диссертационная работа посвящена развитию вышеописанных подходов интегративного моделирования. В работе представлены различные элементы интегративного подхода. В главе 2 рассмотрены работы по использованию суперкомпьютерных расчетов методом атомистической молекулярной динамики для изучения динамики биомакромолекулярных комплексов на осно-

ве уже известных пространственных структур, в главе 3 рассмотрены работы по интегративному моделированию нуклеосом на основе данных расщепления ДНК гидроксильными радикалами (hydroxyl-radical footprinting), в главе 4 рассмотрены работы по созданию моделей комплексов нуклеосом с белками хроматина на основе разнообразных экспериментальных данных, в главе 5 рассмотрены работы по интегративному моделированию амилоидоподобных фибрилл.

Последующие разделы данной главы будут посвящены изложению ряда методических основ методов молекулярного моделирования, которые использовались автором в ходе решения научных задач, излагаемых в данной диссертации, а также изложению основ некоторых экспериментальных методов, данные которых использовались автором в работах по интегративному моделированию биомолекулярных комплексов.

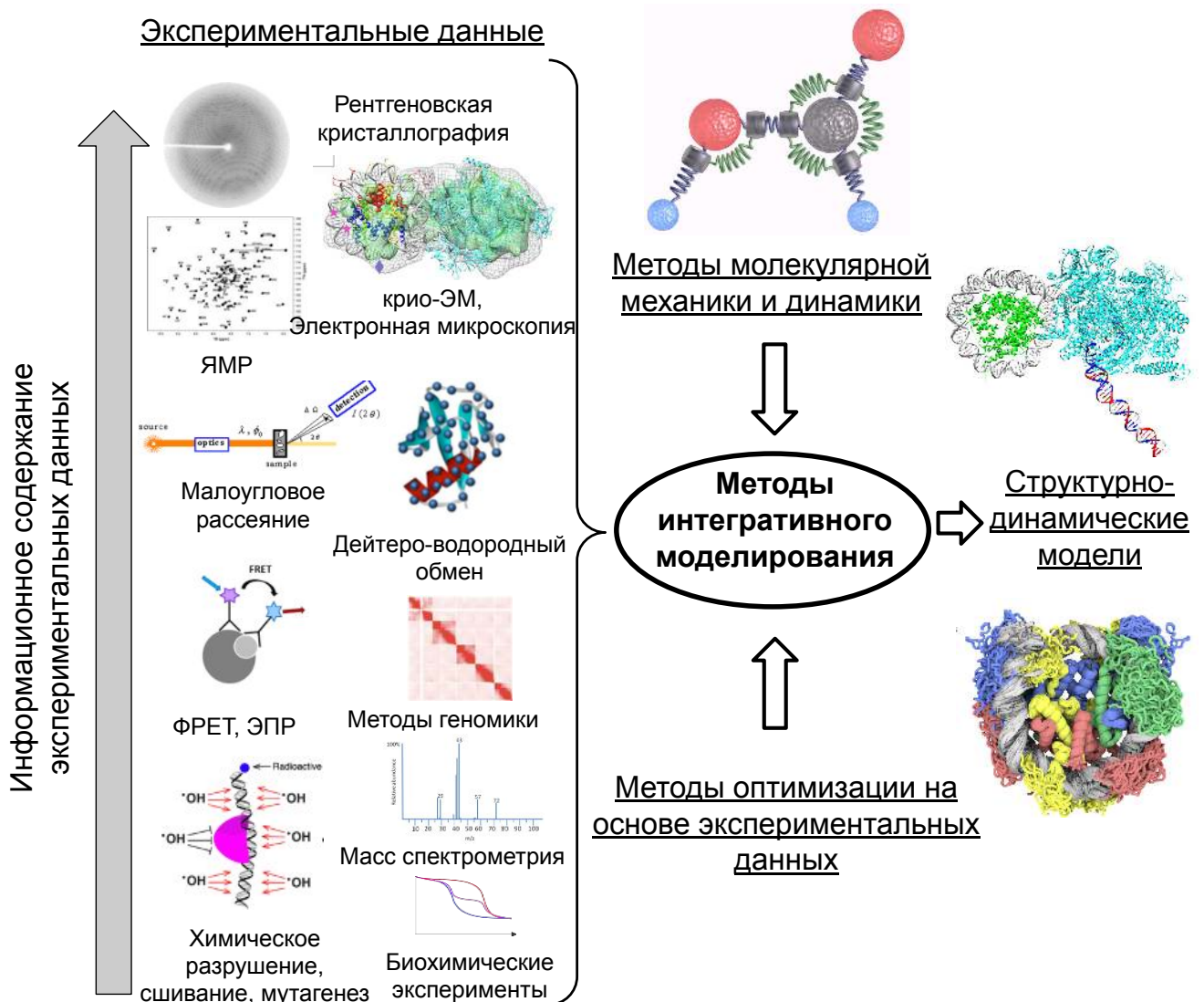


Рисунок 1.3 — Идея методов интегративного моделирования.

## 1.2 Методы атомистического суперкомпьютерного моделирования

*Начало раздела 1.2 и раздел 1.2.1 изложены согласно кандидатской диссертации автора [7] «*

О методах классического молекулярного моделирования написано множество подробных книг и монографий (см. например [33; 43]), поэтому ниже мы лишь кратко обсудим основные понятия и подходы.

В основе методов классических атомистических методов молекулярного моделирования лежит представление о молекулярной структуре как о наборе классических частиц (материальных точек или в некоторых случаях твёрдых тел) взаимодействующих по законам классической механики. Молекулярно-механическая модель обычно предполагает следующие приближения: (i) электронные степени свободы не учитываются, а учитываются лишь атомные степени свободы, (ii) молекулы полагаются находящимися в основном энергетическом состоянии, (iii) автоматически предполагается приближение Борна-Оппенгеймера. На практике при задании конкретных типов взаимодействий между атомами в виде молекулярного гамильтониана предпринимается ещё ряд серьёзных приближений.

В случае моделирования биомолекулярных систем в приближении классической механики методические основы можно грубо разделить на три больших области: 1) задание механистической модели молекулярной системы на основе набора параметров, называемого силовым полем, 2) задание метода изучения динамики системы или метода изучения конформационного пространства данной системы, 3) различные технические и алгоритмические детали, связанные с эффективным и быстрым расчетом взаимодействий между атомами, а также оптимизацией расчетов на различных вычислительных системах.

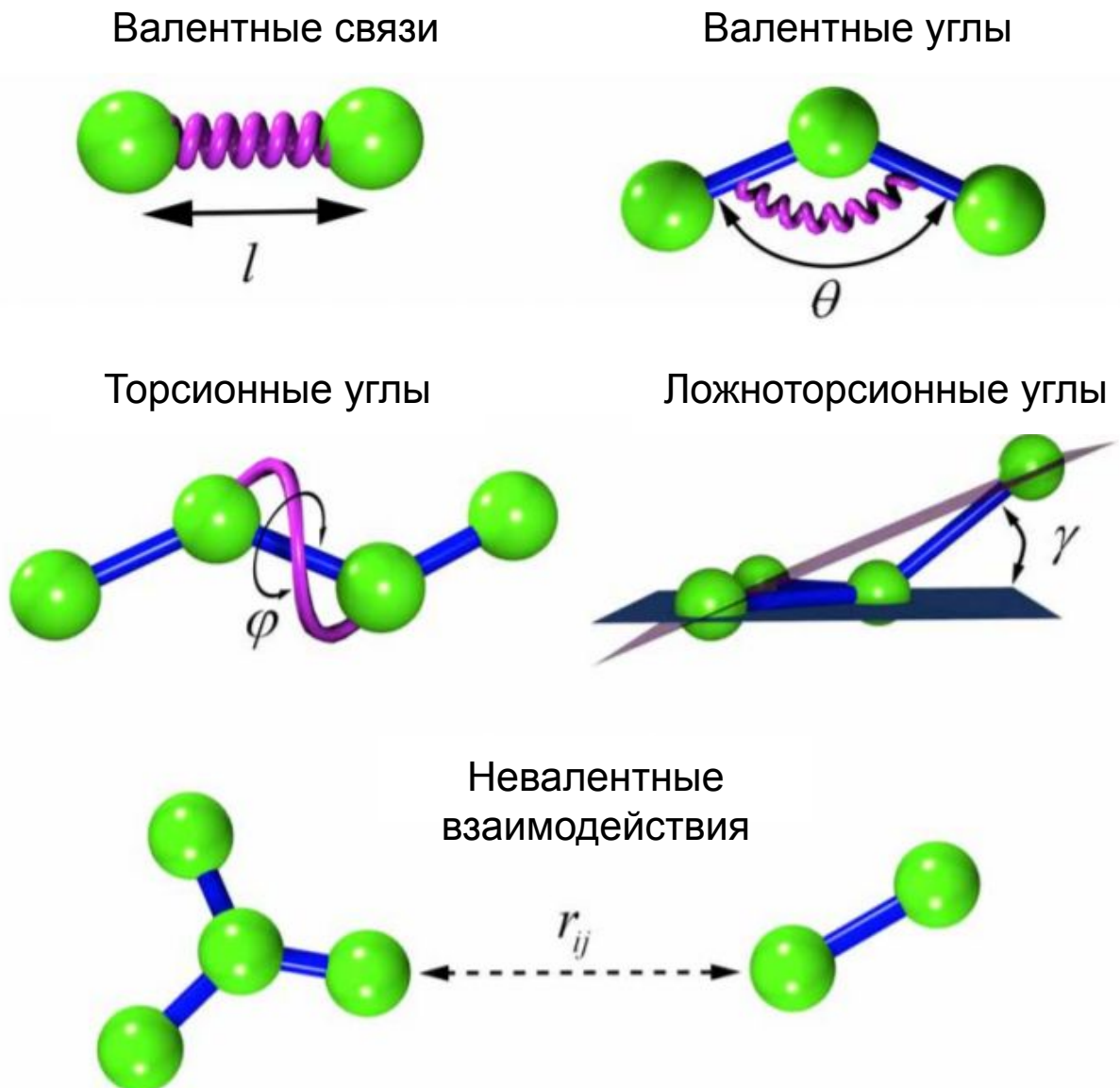
### 1.2.1 Силовые поля

В молекулярной механике молекулярная система описывается в терминах классической механики набором точечных частиц (обычно представляющих атомы или группы атомов) и их взаимодействий, заданных функцией по-

тенциальной энергии. Форма и структура функции потенциальной энергии называется силовым полем. На сегодняшний день разработано большое количество различных типов силовых полей, от обычных силовых полей до силовых полей, специфичных для определенного класса соединений. К наиболее популярным относятся группы силовых полей OPLS-AA [64], AMBER [65], CHARMM [66], GROMOS [67], CVFF [68], PCFF [69], MM3 [70] и другие. Эти силовые поля могут быть параметризованы относительно экспериментальных данных (в частности, калориметрических, спектроскопических), а также для воспроизведения поверхностей потенциальной энергии, вычисленных с помощью методов квантовой химии. Сложность функциональной формы силового поля может варьироваться, однако почти все молекулярные силовые поля содержат шесть общих основных термов для описания валентных и невалентных взаимодействий. Среди валентных взаимодействий выделяют термы описывающие растяжения ковалентных связей, деформации валентных углов, торсионных углов и ложноторсионные углы. Невалентные взаимодействия обычно описываются комбинацией парных ван-дер-Ваальсовых взаимодействий и кулоновских взаимодействий, основанных на точечных частичных зарядах атомов. Рисунок 1.4 и уравнение 1.1 дополнительно иллюстрируют эти термы. Многие из обычных силовых полей, используемых для моделирования биомолекул, включают в себя только упомянутые выше основные энергетические термины. Более сложные силовые поля (например, MM3, PCFF) могут также включать ангармонические члены и перекрестные члены для валентных взаимодействий, которые отражают связь между внутренними координатами. Например, при уменьшении валентного угла между атомами обнаруживается, что валентные связи удлиняются, чтобы уменьшить взаимодействие между атомами. Было обнаружено, что перекрестные члены играют важную роль в силовых полях, предназначенных для предсказания колебательных спектров [71]. Было высказано предположение, что наличие перекрестных членов (вместе с некоторыми другими особенностями) может обеспечить общий способ классификации силовых полей [72]. Согласно этой логике, силовое поле класса I ограничено гармоническими членами (например, для растяжения связей и валентного угла) и не имеет перекрестных членов. Силовое поле класса II будет иметь ангармонические члены и явные перекрестные члены для учета взаимоотношений между различными геометрическими параметрами. Еще одной характеристикой силового поля класса II считается возможность его

использования без модификации одновременно для моделирования свойств изолированных малых молекул, конденсированных фаз и макромолекулярных систем [71].

$$U(\{\vec{r}_i\}) = \sum_{bonds} \frac{1}{2} k_b (l - l_0)^2 + \sum_{angles} \frac{1}{2} k_\theta (\theta - \theta_0)^2 + \sum_{torsions} \frac{1}{2} V_n [1 + \cos(n\varphi - \varphi_0)] + \sum_{impropers} \frac{1}{2} k_\gamma (\gamma - \gamma_0)^2 + \sum_{j+1}^{N-1} \sum_{i=j+1}^N \left\{ 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right\} f_{ij} \quad (1.1)$$



**Рисунок 1.4** — Графическое изображение энергетических термов в силовых полях молекулярной механики. Атомы изображены зелеными сферами. Рисунок соответствует членам, представленным в уравнении 1.1



## Силовое поле OPLS-AA

В качестве примера устройства типичного силового поля, применяемого для моделирования белков и других биологических молекул, рассмотрим поле OPLS-AA. Силовое поле OPLS-AA [73; 74] является полноатомной версией поля OPLS (Optimized Potentials for Liquid Simulations), то есть атомы водорода рассматриваются явным образом. Параметры поля OPLS, в частности, подгонялись под экспериментальные свойства жидкостей, таких как плотность и теплота испарения. Общая структура энергии молекулярной системы представлена в уравнении 1.2.

$$\begin{aligned}
 U(r^N) &= U_{bond} + U_{angle} + U_{dih} + U_{nb} & (1.2) \\
 U_{bond} &= \sum_{bonds} K_r (r - r_{eq})^2 \\
 U_{angle} &= \sum_{angles} k_\theta (\theta - \theta_{eq})^2 \\
 U_{dih} &= \frac{V_1}{2} [1 + \cos(\varphi)] + \frac{V_2}{2} [1 - \cos(2\varphi)] + \frac{V_3}{2} [1 + \cos(3\varphi)] + \frac{V_4}{2} [1 - \cos(4\varphi)] \\
 U_{nb}^{ab} &= \sum_i^{on\ a} \sum_j^{on\ b} \left\{ 4\varepsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{r_{ij}} \right\} f_{ij}
 \end{aligned}$$

межмолекулярные взаимодействия  $U_{nb}$  учитываются только для атомов, которые удалены на три и более связи, для 1-4 взаимодействий коэффициент скейлинга  $f_{ij} = 0.5$ , во всех остальных случаях  $f_{ij} = 1$ .

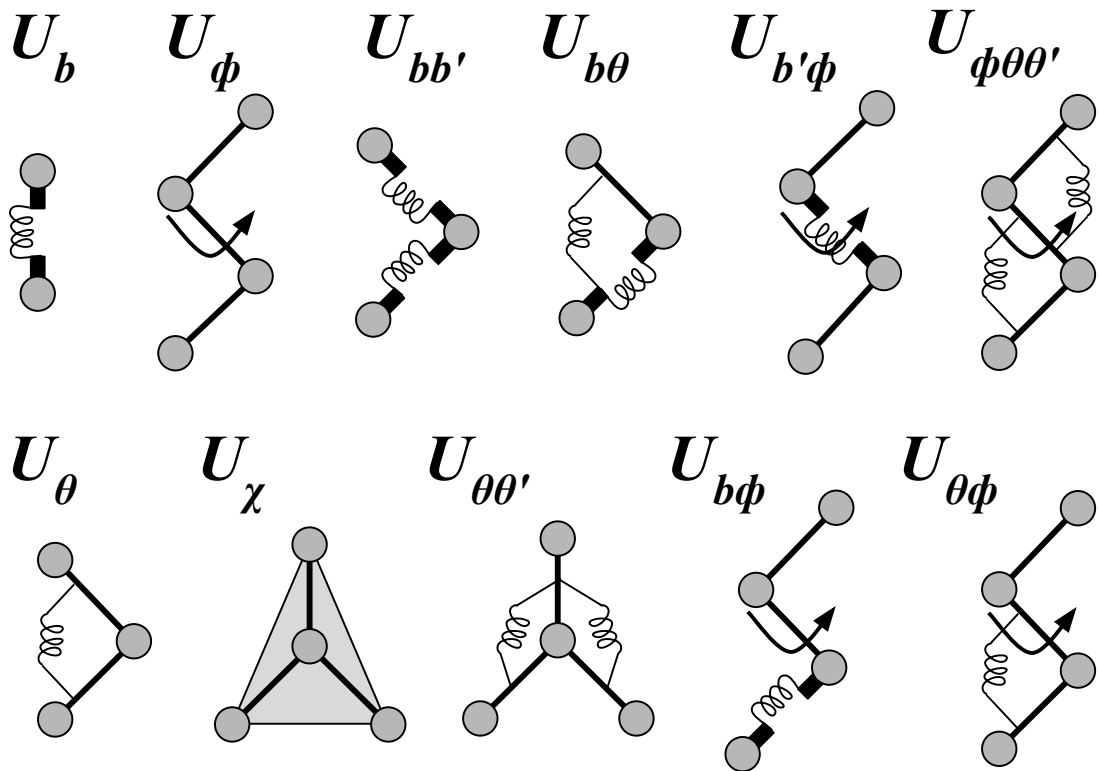
## Силовое поле PCFF

Силовое поле PCFF (Polymer Consistent Force Field) [75] основано на силовом поле CFF91 и относится к семейству силовых полей CFF (consistent force-field). Это поле создано для моделирования широкого класса органических полимеров, утверждается, что силовые поля второго поколения (к которым относится PCFF) более точно воспроизводят экспериментальные результаты, чем силовые поля первого поколения. Платой за это является более сложное

устройство поля и наличие перекрёстных членов в функции энергии. Общий вид энергии молекулярной системы в приближении поля PCFF приведён в формулах 1.4. На Рис. 1.5 приведено графическое пояснение к внутренним степеням свободы, являющимися параметрами для каждого члена потенциальной энергии.

$$\begin{aligned}
U(r^N) &= U_b + U_\theta + U_\varphi + U_\chi + U_{bb'} + U_{\theta\theta'} + U_{b\theta} + U_{b\varphi} + U_{b'\varphi} + U_{\theta\varphi} + U_{\varphi\theta\theta'} + U_{nb} \\
U_b &= \sum_b K(b - b_0)^2 \\
U_\theta &= \sum_\theta H(\theta - \theta_0)^2 \\
U_\varphi &= \sum_\varphi V(1 + \cos(\varphi - \varphi_0)) \\
U_\chi &= \sum_\chi K_\chi \chi^2 \\
U_{bb'} &= \sum_{bb'} F_{bb'}(b - b_0)(b' - b'_0) \\
U_{\theta\theta'} &= \sum_{\theta\theta'} F_{\theta\theta'}(\theta - \theta_0)(\theta' - \theta'_0) \\
U_{b\theta} &= \sum_{b\theta} F_{b\theta}(b - b_0)(\theta - \theta_0) \\
U_{b\varphi} &= \sum_{b\varphi} (b - b_0)[V_1 \cos \varphi + V_2 \cos 2\varphi + V_3 \cos 3\varphi] \\
U_{b'\varphi} &= \sum_{b'\varphi} (b' - b'_0)[V_1 \cos \varphi + V_2 \cos 2\varphi + V_3 \cos 3\varphi] \\
U_{\theta\varphi} &= \sum_{\theta\varphi} (\theta - \theta_0)[V_1 \cos \varphi + V_2 \cos 2\varphi + V_3 \cos 3\varphi] \\
U_{\varphi\theta\theta'} &= \sum_{\varphi\theta\theta'} K_{\varphi\theta\theta'} \cos \varphi (\theta - \theta_0)(\theta' - \theta'_0) \\
U_{nb} &= \sum_{i>j} \left\{ \left[ \frac{A_{ij}}{r_{ij}^9} - \frac{B_{ij}}{r_{ij}^6} \right] + \frac{q_i q_j}{r_{ij}} \right\} f_{ij}
\end{aligned} \tag{1.3}$$

»



**Рисунок 1.5** — Графическое пояснение к членам внутримолекулярной энергии уравнения 1.4.

### 1.2.2 Методы молекулярной динамики

Определив модель молекулярной механики, можно применять различные методы для ее анализа; они включают в себя минимизацию энергии, моделирование молекулярной динамики, моделирование по методу Монте-Карло, а также различные более сложные методы, основанные на этих методах. Ключевым методом данной работы (а также, вероятно, наиболее часто используемым для изучения сложных молекулярных систем) является метод моделирования молекулярной динамики (МД) и его различных вариаций. Основная идея метода МД моделирования очень проста и основана на численном решении классических уравнений движения атомов в молекулярной системе на основе второго закона Ньютона:

$$\frac{d^2 \vec{r}_i}{dt^2} = \vec{a} = \frac{\vec{F}}{m} = -\frac{1}{m_i} \nabla_i U(\{\vec{r}_i\}) \quad (1.4)$$

Численное решение обычно получается с помощью одной из схем интегрирования, такой как алгоритм Верле (velocity Verlet algorithm) [76]:

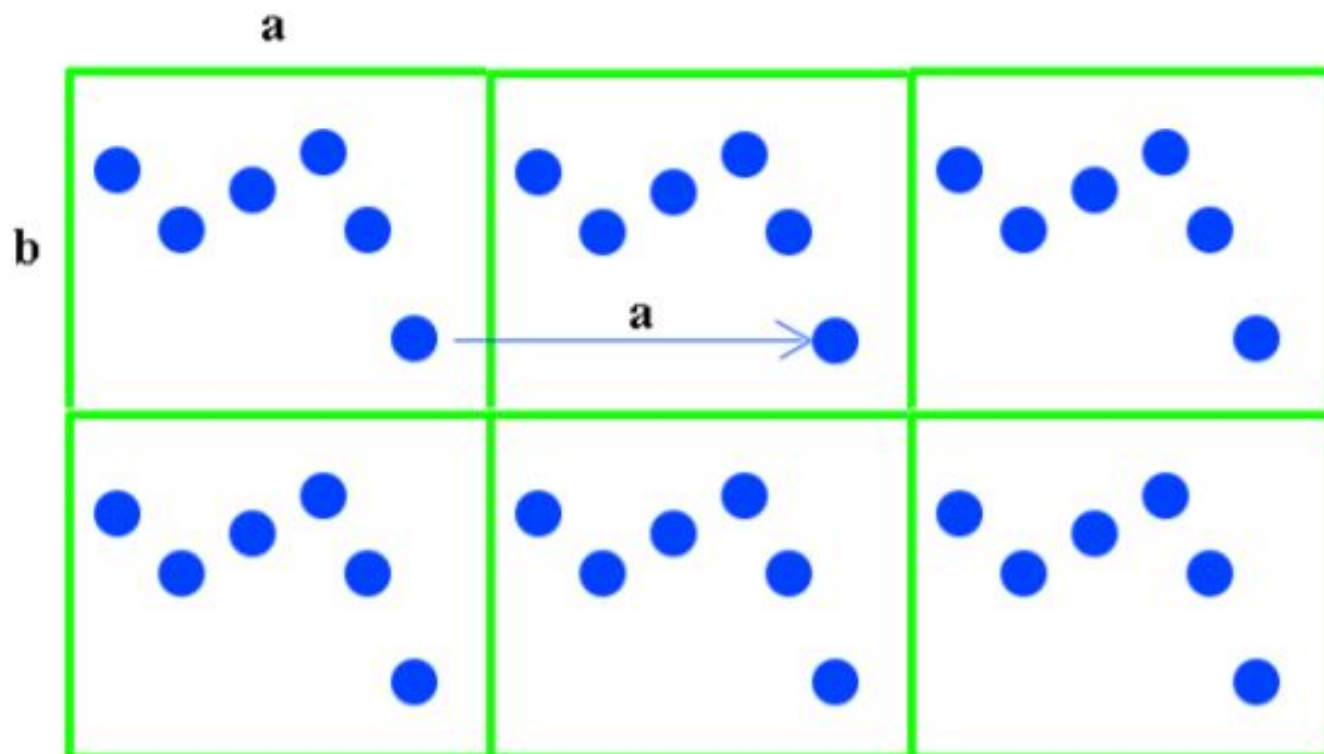
$$\begin{aligned}\vec{r}(t + \Delta t) &= \vec{r}(t) + \vec{v}(t)\Delta t + \frac{1}{2}\vec{a}(t)\Delta t^2 \\ \vec{v}(t + \Delta t) &= \vec{v}(t) + \frac{\vec{a}(t) + \vec{a}(t + \Delta t)}{2}\Delta t\end{aligned}\quad (1.5)$$

Хотя основная идея, лежащая в основе МД-моделирования, на первый взгляд кажется довольно простой и понятной, ее обоснованность и интерпретация связаны с фундаментальными проблемами теоретической физики и математики на стыке механики, статистической физики и теории хаоса [77]. Одним из таких вопросов является связь между обратимыми во времени уравнениями Ньютона, используемыми для моделирования эволюции системы во времени, и эмпирическими необратимыми законами термодинамики, приводящими систему к состоянию, в котором энтропия максимальна. Другой набор вопросов и приближений, который имеет большое значение для практической реализации МД моделирования в реальных молекулярных системах, будет кратко обсужден ниже и включает использование периодических граничных условий, реализацию статистических ансамблей, динамику Ланжевена и диссипативных частиц, вычисления несвязанных взаимодействий, параллельная реализация МД алгоритмов.

### 1.2.3 Периодические граничные условия

Периодические граничные условия (ПГУ) - это метод, используемый для уменьшения влияния граничных эффектов на молекулярную систему во время моделирования и аппроксимации свойств объемных систем (таких как газы или жидкости) с помощью модели, состоящей из конечного набора молекул. Типичный случай применения периодических граничных условий - моделирование макромолекул в явном растворителе. В периодических граничных условиях ячейка моделирования окружена своими периодическими копиями во всех трех пространственных направлениях (см. Рисунок 1.6), таким образом, эффективно образуя бесконечную кристаллическую структуру. Каждый атом, характеризуемый своими действительными координатами, также имеет координаты своих периодических образов. Распространенной формой учета частиц в периодических граничных условиях является соглашение о ближайшем образе, при

котором каждый отдельный атом в системе взаимодействует с ближайшим к себе образом каждой частицы в системе.



**Рисунок 1.6** — Иллюстрация периодических граничных условий, используемых в МД моделировании. Кружками обозначены атомы и их периодические образы в соседних периодических ячейках.

ПГУ также используются в сочетании с методами учета электростатических взаимодействий на больших расстояния, такими как суммирование по Эвальду [78]. Поэтому полный электростатический заряд системы должен быть равен нулю, чтобы избежать проблем с расхождением значений полной электростатической энергии. Однако даже если система нейтральна, общий дипольный момент ячейки может приводить к некоторым искусственным энергетическим эффектам, похожим на пьезоэлектрические эффекты в полярных кристаллах. Чтобы предотвратить артефакты связанные с ПГУ, размер ячейки моделирования должен быть достаточно большим, чтобы избежать таких нефизических эффектов. В том числе если размер ячейки слишком мал, макромолекула может взаимодействовать со своим собственным изображением в соседней ячейке, что эквивалентно взаимодействию молекулы с собой. Это может привести к нефизической динамике большинства макромолекул.

### 1.2.4 Статистические ансамбли в МД моделировании

Настоящие молекулярные системы никогда не изолированы от окружающей среды. Взаимодействие молекулярной системы с окружающей средой в плане обмена энергией, объемом и частицами с точки зрения статистической физики может быть описано в терминах статистических ансамблей [79]. Ансамбль микросостояний молекулярной системы, связанной с внешним термостатом с постоянной температурой, называется каноническим ансамблем (или NVT-ансамблем), аналогично система под поршнем, связанная с термостатом, дает изобаро-изотермический ансамбль (или NPT -ансамбль).

Поскольку базовые алгоритмы МД моделирования обеспечивают решение уравнений движения для изолированной системы (микрочанонический ансамбль), обычно вводятся специальные алгоритмы, которые имитируют связь системы с внешним резервуаром энергии или объема [80]. Цель таких алгоритмов - поддерживать правильные характеристики средней температуры (объема), в то же время обеспечивая правильную величину колебаний, которые зависят от внутренних свойств системы, таких как теплоемкость и сжимаемость. Среди алгоритмов термостатирования следующие алгоритмы нашли применение в современных кодах моделирования: термостат Берендсена [81], термостат Нозе-Гувера [82], цепи Нозе-Гувера [83], термостат Андерсена [84], термостат релаксации скоростей [85].

Термостат Берендсена, будучи самым простым, полагается на масштабирование скоростей с коэффициентом  $\lambda$  на каждом временном шаге:

$$\lambda = \sqrt{\left[1 + \frac{\Delta t}{\tau_t} \left\{ \frac{T_0}{T} - 1 \right\}\right]} \quad (1.6)$$

Влияние этого алгоритма на температуру заключается в том, что фактическая температура корректируется до эталонной температуры  $T_0$  экспоненциальным образом:

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau} \quad (1.7)$$

Однако термостат Берендсена имеет серьезные недостатки: во-первых, он не допускает колебаний температуры и, таким образом, не реализует правильный ансамбль, и, во-вторых, было показано, что процедура перемасштабирования скорости вызывает “эффект летающего куба” при моделировании системы в вакууме, когда энергия связанная с поступательным движением и вращением всей

системы непрерывно растут, в то время как энергия высокочастотных основных мод истощается [86; 87]. Формализм Нозе-Гувера - это подход, основанный на расширенном лагранжиане, содержащем дополнительные искусственные координаты и скорости, который приводит к детерминированной молекулярной динамике при постоянной температуре. Расширенный гамильтониан, первоначально предложенный Нозе [88], можно записать следующим образом:

$$H_{Nose} = \sum_{i=1}^N \frac{\tilde{\mathbf{p}}_i^2}{2m_i \tilde{s}^2} + U(\tilde{\mathbf{r}}^N) + \frac{\tilde{\mathbf{p}}_s^2}{2Q} + gkT_0 \ln(\tilde{s}) \quad (1.8)$$

где  $\mathbf{p}_s$  и  $s$  - координаты расширенной переменной,  $Q$  - ее фиктивная масса,  $k$  - постоянная Больцмана,  $T_0$ -эталонная температура,  $g$  - параметр. Этот гамильтониан порождает микроканонический ансамбль из  $6N + 2$  степеней свободы. Переменные, присутствующие в этом гамильтониане, называются виртуальными переменными. Тем временем определяется набор реальных переменных, которые соотносятся с виртуальными переменными следующим образом:

$$\mathbf{r} = \tilde{\mathbf{r}} \quad (1.9)$$

$$\mathbf{p} = \tilde{\mathbf{p}}/s \quad (1.10)$$

$$s = \tilde{s} \quad (1.11)$$

$$dt = d\tilde{t}/s \quad (1.12)$$

и псевдогамильтониан может быть тогда задан как

$$H(\mathbf{p}, \mathbf{r}) = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} + U(\mathbf{r}^N) \quad (1.13)$$

Можно показать, что микроканонический ансамбль, сгенерированный путем решения уравнений движения для расширенной системы в виртуальных переменных, будет соответствовать канонической выборке реальных переменных согласно соответствующему псевдогамильтониану (1.13), если параметр  $g = 3N$  и другие условия выполнены (см. ниже). Тепло передается в реальную систему и обратно от расширенной переменной колебательным образом, что приводит к почти периодическим колебаниям температуры. Позже Нозе и Гувер показали, что соответствующие уравнения движения могут быть сформулированы в терминах реальных системных переменных, что дает уравнения движения

Нозе-Гувера:

$$\begin{aligned}\ddot{\mathbf{r}}_i &= \frac{\mathbf{F}_i}{m_i} - \frac{\xi \dot{\mathbf{r}}_i}{m_i} \\ \dot{\xi} &= \frac{kg}{Q}(T(t) - T_0)\end{aligned}\quad (1.14)$$

Детальный теоретический анализ негамильтоновой динамики, генерируемой уравнениями Нозе-Гувера, показывает, что алгоритм генерирует правильное распределение только при наличии единственного интеграла движения, сохраняющегося во время динамики, обычно полной энергии расширенной системы, и никаких других интегралов движения [80]. Исключением является сохранения полного импульса системы, если центр масс остается неподвижным.

Данное утверждение также подразумевает, что динамика является эргодической, то есть усреднение по траектории эквивалентно усреднению по фазовому пространству. Однако было показано, что для небольших или жестких систем динамика неэргодична, и правильные распределения не генерируются. Чтобы решить эту проблему, Martyna et al. [83] предложили схему, в которой термостат Нозе-Гувера соединен с другим термостатом или цепочкой термостатов.

В подходе Андерсена к изотермическому моделированию система связана с термостатом с помощью стохастических импульсных сил, которые время от времени действуют на случайно выбранные частицы. В процессе столкновения частицы приобретают новые скорости согласно распределению Максвелла-Больцмана, соответствующие заданной температуре. Термостат Андерсена обеспечивает правильный статистический ансамбль, но делает динамику прерывистой из-за случайных столкновений.

Термостат решкалирования скорости (Velocity rescale), предложенный Bussi et al. [85] представляет собой комбинацию термостата Берендсена со стохастическим членом, который обеспечивает правильное распределение кинетической энергии.

Аналогично алгоритмам термостатирования существуют алгоритмы баростатирования, которые позволяют реализовать правильное целевое давление и NPT ансамбль. Схема баростатирования Берендсена [81] позволяет быстро отрелаксировать систему до целевого давления и поддерживать его на протяжении всего моделирования. Алгоритм Берендсена изменяет масштаб координат и векторов ячейки моделирования на каждом шаге с помощью матрицы  $\mu$  при



заданном эталонном давлении  $P_0$  следующим образом:

$$\hat{\mu}_{ij} = \delta_{ij} - \frac{\Delta t}{3\tau_p} \beta_{ij} \{P_{0ij} - P_{ij}(t)\} \quad (1.15)$$

Где  $\beta$  - изотермическая сжимаемость. Это приводит к экспоненциальной релаксации давления в соответствии с законом:

$$\frac{dP}{dt} = \frac{P_0 - P(t)}{\tau_p} \quad (1.16)$$

Как и в случае с термостатом, баростат Берендсена не полностью реализует NPT ансамбль и подавляет правильные флуктуации давления и объема, однако, если эти флуктуации не являются термодинамически значимыми, данный подход является простым и удобным методом поддержания постоянного давления в системе.

Другой схемой поддержания давления, аналогичной термостату Нозе-Гувера, которая правильно реализует NPT ансамбль, является схема Парринелло-Рамана [89].

### 1.2.5 Ланжевеновская и диссипативная динамика частиц

Ланжевеновская (или стохастическая) динамика и диссипативная динамика частиц (ДДЧ) (dissipative particle dynamics) - это методы, которые позволяют имитировать влияние растворителя на динамику молекулярной системы посредством введения диссипативных и стохастических членов в уравнения движения. Взаимодействие диссипативных и стохастических сил не только изменяет динамику, но также может гарантировать правильные характеристики NVT ансамбля. В пределе небольшого воздействия (малая вязкость растворителя) эти методы также являются удобными алгоритмами для поддержания NVT ансамбля при моделировании методом МД.

Стохастическая или скоростная динамика Ланжевена добавляет трение и шум к уравнениям движения Ньютона следующим образом:

$$m_i \ddot{\mathbf{r}} = \mathbf{F}_i - m_i \xi_i \dot{\mathbf{r}} + \hat{\mathbf{r}}_i \quad (1.17)$$

где  $\xi_i$  - постоянная трения, а  $\hat{r}_i(t)$  - это шумовой процесс с корреляционной функцией

$$\langle \hat{r}_i^\alpha(t) \hat{r}_i^\beta(t+s) \rangle = 2m_i \xi_i kT \delta(s) \delta_{ij} \delta_{\alpha\beta} \quad (1.18)$$

т.е. он нескоррелирован для разных частиц, разных компонентов и разного времени. Формальный вывод уравнения Ланжевена посредством редукции быстрых степеней свободы и их аппроксимации двумя последними членами уравнения (1.17) основан на технике проекционного оператора, введенной Мори и Цванцигом [90]. В остальных членах, отвечающих за трение и стохастический процесс, имеют интуитивно понятный физический смысл: первая представляет собой диссипативную силу, вызываемую трением с растворителем, вторая - своего рода движущую силу, которая вызывается “ударами” окружающих частиц растворителя.

Случайное трение и стохастические силы не сохраняют ни импульс, ни энергию системы. Под действием силы трения энергия отводится из системы, в то время как стохастический процесс может закачивать энергию в систему, поэтому динамика и выборка больше не соответствуют микроканоническому ансамблю, в то время как энергия системы флуктуирует, что напоминает канонический ансамбль NVT. Кроме того, можно показать, что для реализации распределения Больцмана в стохастической динамике трение и стохастические силы должны быть связаны через уравнение (1.18).

Хотя метод Ланжевена привлекателен для жидкостных систем, где структура растворителя не важна, или как алгоритм для строгой реализации ансамбля NVT, он не дает правильной гидродинамики, поскольку случайные силы и силы трения не сохраняют импульс системы. Чтобы преодолеть эту проблему, но при этом сохранить положительные черты стохастической динамики, метод диссипативной динамики частиц (ДДЧ) (dissipative particle dynamics) был первоначально разработан Хугербрюгге и Кольманом для мезоскопического моделирования простых и сложных жидкостей [91]. В ДДЧ сила, действующая на каждую частицу, состоит из трех членов [92]:

$$\mathbf{F}_i = \sum_{j \neq i} (\mathbf{F}_{ij}^C + \mathbf{F}_{ij}^D + \mathbf{F}_{ij}^R) \quad (1.19)$$

консервативная сила, диссипативная сила и случайная сила, соответственно. В то время, как консервативная сила является обычной парной силой, диссипа-

тивная и случайная силы определяются следующим образом:

$$\mathbf{F}_{ij}^D = -\gamma\omega_D(r_{ij})(\mathbf{n}_{ij} \cdot \mathbf{v}_{ij})\mathbf{n}_{ij} \quad (1.20)$$

$$\mathbf{F}_{ij}^R = \sigma\omega_R(r_{ij})\zeta_{ij}\mathbf{n}_{ij} \quad (1.21)$$

где  $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ ,  $\mathbf{n}_{ij} = \mathbf{r}_{ij}/|\mathbf{r}_{ij}|$ ,  $\gamma$  и  $\sigma$  - скалярные параметры, представляющие силу трения и случайную силу.  $\omega_D$  и  $\omega_R$  - весовые коэффициенты, которые варьируются от 0 до 1 и исчезают выше радиуса отсечки  $R_c$ .  $\zeta_{ij}$  - представляет собой гауссовский белый шум со следующими свойствами:

$$\langle \zeta_{ij}(t) \rangle = 0, \zeta_{ij} = \zeta_{ji} \quad (1.22)$$

$$\langle \zeta_{ij}(t)\zeta_{i'j'}(t') \rangle = (\delta_{ii'}\delta_{jj'} + \delta_{ij'}\delta_{i'j})\delta(t - t') \quad (1.23)$$

Замечательной особенностью уравнений ДДЧ в отличие от динамики Ланжевена является то, что диссипативные и случайные силы зависят только от относительного положения и скорости частицы и действуют по векторам, соединяющим пары частиц. Это в сочетании с симметрией индексов обеспечивает выполнение третьего закона Ньютона и сохранение количества движения. Далее можно показать, что уравнения ДДЧ описывают системы в ансамбле NVT при  $\sigma^2 = 2\gamma kT$ . Выбор одной весовой функции может быть произвольным, однако можно показать, что обе функции должны быть связаны как  $\omega_D(r) = (\omega_R(r))^2$ . Типичный выбор для этой функции:

$$\omega(r) = \left\{ 2\left(1 - \frac{r}{R_c}\right), r \leq R_c; 0, r > R_c \right\} \quad (1.24)$$

ДДЧ в основном применяется как метод мезоскопического моделирования, где частицы ДДЧ представляют собой целые молекулы или области жидкости, а не отдельные атомы. Однако трение и случайные члены могут также использоваться в атомистическом моделировании для моделирования влияния растворителя и реализации ансамбля NVT. Преимуществами термостата ДДЧ перед термостатом Ланжевена является сохранение импульса и малый радиус взаимодействий ДДЧ, оба фактора способствуют крупномасштабным конформационным переходам в молекулярной системе, если они должны произойти.

### 1.2.6 Расчет невалентных взаимодействий

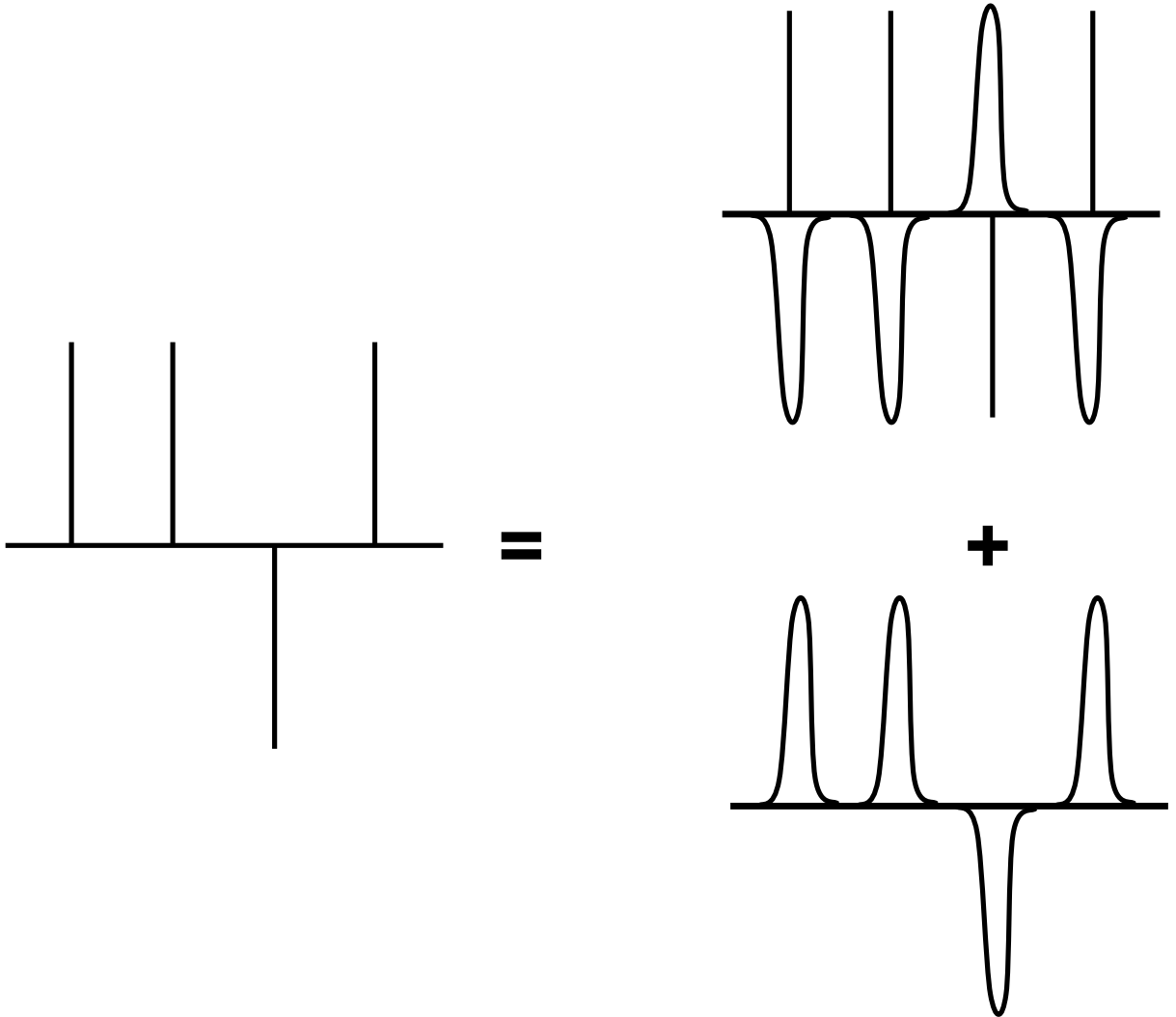
Вычисление невалентных взаимодействий и особенно кулоновских взаимодействий - одна из задач МД моделирования. Важной отличительной чертой электростатических взаимодействий является их медленный затухающий характер ( $\sim 1/r$ ), что, строго говоря, делает некорректным обрезку электростатического взаимодействия на некотором расстоянии во время вычисления сил и энергии при моделировании. В трехмерной объемной системе энергия взаимодействия точечной частицы с окружающей средой, характеризуемой плотностью заряда  $\rho(r)$ , была бы выражена интегралом:

$$E = \iiint \frac{\rho(\vec{r})}{|\vec{r} - \vec{r}_0|} d^3\vec{r} \quad (1.25)$$

который может сходиться только условно. Таким образом, необходимо соблюдать особую осторожность при расчете электростатических взаимодействий в моделировании МД. Было показано, что при простой обрезке кулоновских взаимодействий могут возникать серьезные артефакты, особенно при моделировании систем зарядов, начиная от шума и нестабильности при моделировании и заканчивая неблагоприятным воздействием на молекулярный порядок [78]. Если изучаемая система не является периодической (т.е. не используются периодические граничные условия), единственной альтернативой полному расчету электростатических взаимодействий (что для больших систем является недопустимо дорогостоящим, поскольку масштабируется как  $\sim N^2$ ) является их обрезка на некотором расстоянии  $R_c$ . Однако необходимо обосновать отсутствие серьезных побочных эффектов в таком случае. Наиболее распространенные ошибки могут возникать из-за: (i) отсутствия непрерывности силы, особенно при образовании нескомпенсированных зарядов на длине радиуса обрезания, приводящего к алгоритмическому шуму и тепловыделению, (ii) для заряженных систем и систем с внутренней периодичностью энергия и поведение системы могут зависеть от радиуса обрезания причудливым образом.

В системах, которые используют периодические граничные условия, может использоваться альтернативная более надежная стратегия суммирования электростатических взаимодействий, она основана на методе суммирования Эвальда и родственных алгоритмах PME (particle mesh Ewald) и PPPM (particle-particle particle-mesh Ewald), основанные на интерполяции функций по

значениям в узлах сеток [80]. Идея суммирования по Эвальду состоит в том, чтобы заменить бесконечную медленно сходящуюся сумму электростатического взаимодействия данного набора частиц и всех их периодических образов двумя быстро сходящимися суммами, одна в прямом, а другая - в обратном пространстве. Рассмотрим представление набора точечных зарядов, дополненных гауссовыми функциями экранирования и их обратными функциями на рисунке 1.7.



**Рисунок 1.7** — Схематическое изображение разложения заряда, используемого в суммировании по Эвальду.

Гауссовские экранирующие функции имеют вид:

$$\rho_{Gauss}(r) = -q_i(\alpha/\pi)^{3/2}exp(-\alpha r^2) \quad (1.26)$$

Можно показать, что в этом случае полный электростатический потенциал может быть выражен как сумма потенциалов обеих подсистем, представленных

на рисунке 1.7, за вычетом члена энергии самовзаимодействия. Электростатический потенциал экранированных точечных зарядов быстро спадает и может быть суммирован в прямом пространстве, в то время как потенциал компенсирующей подсистемы (плавно меняющиеся периодические гауссианы) можно легко вычислить в обратном пространстве Фурье на основе решения уравнения Пуассона. Результирующая полная энергия системы может быть записана следующим образом:

$$U = \frac{1}{2} \sum_{i \neq j}^N \frac{q_i q_j \operatorname{erfc}(\sqrt{\alpha} r_{ij})}{r_{ij}} + \frac{1}{2V} \sum_{\mathbf{k} \neq 0} \frac{4\pi}{k^2} |\rho(\mathbf{k})|^2 \exp(-k^2/4\alpha) - (\alpha/\pi)^{1/2} \sum_{i=1}^N q_i^2 \quad (1.27)$$

где  $\rho(\mathbf{k}) \equiv \sum_{i=1}^N q_i \exp(i\mathbf{k} \cdot \mathbf{r}_i)$ ,  $V$  - объем периодической ячейки.

Однако сумма Эвальда требует больших вычислительных ресурсов для больших систем с порядком масштабирования  $O(N^{3/2})$  и остается дорогостоящей по сравнению с традиционными схемами обрезки взаимодействий. Чтобы решить эту проблему, были разработаны подходы на основе интерполяционных сеток, все они пытаются ускорить решение уравнения Пуассона в периодических граничных условиях, используя преимущества быстрого преобразования Фурье (БПФ) для вычисления дискретных преобразований Фурье [78]. Эти методы могут масштабироваться как  $O(N \log(N))$ . Методы суммирования, подобные суммированию по Эвальду, могут вычислять электростатическую энергию в периодической системе с заданной точностью. Их использование привело к значительному повышению стабильности при моделировании биомолекул. Однако, некоторые фундаментальные проблемы моделирования все еще существуют и могут привести к артефактам во время моделирования [78]. Смит и Петтитт [93; 94] изучали динамические артефакты, вызванные суммированием по Эвальду. Они сосредоточились на вопросе о величине энергетических барьеров при свободном вращении диполей в системах с периодичностью (наличием дальнего порядка). Представьте себе единственный идеальный диполь в начале элементарной ячейки, который затем периодически воспроизводится в соседних ячейках. Потенциальная энергия системы будет зависеть от ориентации диполя, что не подходит для исследования растворов биомолекул. Однако Смит и Петтитт показали, что вращательные барьеры незначительны для дипольных молекул в растворителях с высокой диэлектрической проницаемостью при комнатной температуре с типичными размерами ячейки моделирования. Для

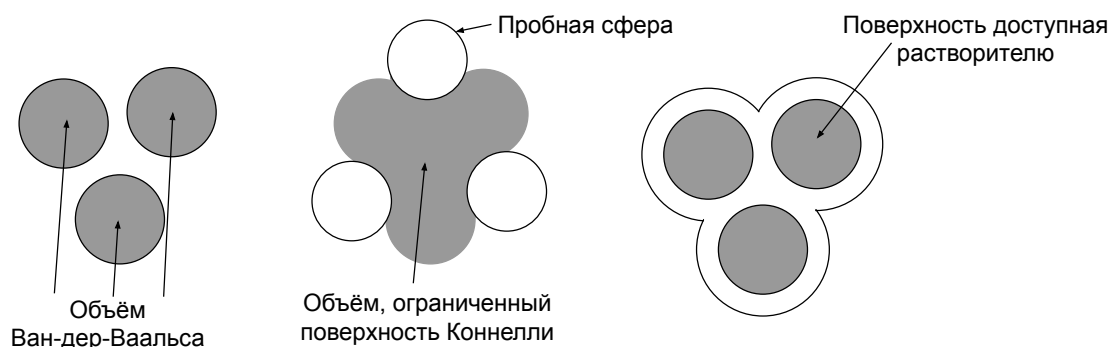
моделирования в растворителях с низкой диэлектрической проницаемостью или в вакууме отсутствие артефактов не может быть гарантировано.

### 1.3 Методы расчета поверхностей в молекулярном моделировании

*Раздел 1.3 изложен согласно кандидатской диссертации автора [7]* « Поверхность доступная растворителю (ПДР) (solvent accessible surface area (SASA)) – это поверхность биомолекулы, которая доступна растворителю. Обычно ПДР выражают в квадратных ангстремах. ПДР была впервые описана Ли и Ричардсом[95] и иногда называется молекулярной поверхностью Ли-Ричардса, она определяется как поверхность которую описывает центр пробной сферы при “обкатывании” сферы вокруг молекулы. Размер сферы обычно берётся порядка размера молекулы воды. В этом плане ПДР можно воспринимать, как расширенную поверхность Ван-дер-Ваальса.

Кроме ПДР широкое распространение получила поверхность Коннолли, которая представляет собой как бы полость, которая формируется в растворителе при помещении в него молекулы. Коннолли разработал также полуаналитический метод расчёта такой поверхности [96].

На Рис. 1.8 изображены поверхности доступные растворителю и молекулярные поверхности Коннолли для модельной системы.



**Рисунок 1.8** — Пояснения к определению поверхности доступной растворителю и поверхности Коннолли.

## 1.4 Методы моделирования ДНК в динуклеотидном приближении

Описание конформации ДНК в виде набора геометрических параметров, определяющих взаимное расположение соседних нуклеотидных пар относительно друг друга, может использоваться как для исследования структуры ДНК, так и для создания физических моделей гибкости ДНК. Последние могут быть встроены в алгоритмы интегративного моделирования ДНК и ДНК-белковых комплексов. В данном разделе излагаются представления о параметрах геометрии ДНК, эмпирических силовых полях для описания энергии изгиба ДНК в динуклеотидном приближении, а также программные продукты и методы моделирования, включая разработки научной группы автора.

### 1.4.1 Описание геометрии ДНК

Планарность азотистых оснований нуклеотидов и относительная планарность комплементарных пар оснований в двухцепочечной спирали ДНК позволяет применить для описания их взаимного расположения математический формализм, в котором основания или пары оснований представлены в виде плоских элементов с двумя уникальными гранями, отличимыми друг от друга из-за асимметрии оснований (см. Рис. 1.9). Разработке данных подходов способствовали работы Р. Дикерсона [97], В.Б. Журкина, В.И.Иванова, В.П.Лысова [98], В. Олсон, К. Лу [99], М. Эль Хассана, К. Калладайна [100] и др. Ниже сформулируем принятые на сегодняшний день подходы и алгоритмы.

Каждое азотистое основание (см. Рис. 1.9б) можно представить в виде твердого тела (при изображении обычно используются плоские прямоугольные параллелепипеды) и связанной с этим твердым телом системы отсчета. В 1999 на семинаре в Цукубе были приняты (а позднее утверждены совместной комиссией по биохимической номенклатуре Международного союза биохимии и молекулярной биологии и Международного союза теоретической и прикладной химии (IUPAC)) стандарты эталонной геометрии (расположения атомов) в идеализированных азотистых основаниях нуклеиновых кислот и их идеализированных Уотсон-Криковских парах [101]. Данные эталонной геометрии заданы



в виде координат атомов относительно стандартной системы отсчета введенной для Уотсон-Криковских пар. При анализе структуры ДНК проводится совмещение эталонного основания или пары оснований с реальными по методу минимизации среднеквадратичного отклонения атомов (RMSD), таким образом каждому реальному основанию можно сопоставить, связанную с ним систему координат (см. Рис. 1.9в). В стандартной системе отсчета ось Z перпендикулярна плоскости основания и в структуре В-ДНК совпадает с направлением 5'-3' цепи ДНК для левого основания на Рис. 1.9в, ось Y параллельна линии соединяющей C1' атомы нуклеотидов, ось X направлена в сторону большой бороздки ДНК и совпадает с осью псевдосимметрии идеальной Уотсон-Криковской пары. Таким образом, в силу наличия оси псевдосимметрии эталонные координаты необходимо задать для оснований, когда они находятся в левом положении на Рис. 1.9в, координаты для их положения с правой стороны могут быть получены путем преобразования симметрии.

Определив положения систем отсчета, связанных с основаниями или парами оснований, можно ставить задачу о вычислении геометрических параметров взаимного расположения твердых тел, характеризуемыми этими системами отсчета. Из геометрии известно, что для описания взаимного расположения двух твердых тел относительно друг друга, необходимо шесть параметров, три из которых описывают поступательные степени свободы, а три вращательные. Номенклатура таких параметров была утверждена в 1988 году и носит название "Кембриджского соглашения" (Cambridge Accord) [97]. Поступательные степени свободы для взаимного расположения оснований относительно друг друга носят названия Shear, Stretch, Stagger для смещения по осям X, Y, Z, соответственно (Рис. 1.9д). В случае определения этих параметров для Уотсон-Криковской пары M-N, по соглашению, систему координат основания N предварительно вращают вокруг оси X на 180 градусов. Вследствие этого, если рассматривать ту же самую пару как пару N-M, угловой параметр Buckle, обсуждаемый далее, поменяет знак на противоположный. Также при рассмотрении Уотсон-Криковской пары в обратном направлении (N-M вместо M-N) поменяет свой знак параметр Shear, но не параметры Stretch и Stagger. В случае рассмотрения Хугстиновских пар приняты другие правила, которые рассматривать здесь не будем.

Аналогично можно ввести параметры взаимного расположения пар оснований вдоль цепи ДНК, они носят название Shift, Slide, Rise для смещений вдоль осей X, Y, Z, соответственно. При рассмотрении двухцепочечной В-формы

ДНК для вычисления данных параметров нам необходимо выбрать направление одной из цепей ДНК в качестве референсного, которое будет задавать направление последовательного рассмотрения пар оснований для вычисления параметров их взаимного расположения. При вычислении параметров в обратном направлении (в качестве референсной цепи задается комплиментарная цепь) знак параметра Shift поменяется на противоположный, а параметры Slide и Rise не изменят знака.

Вычисление параметров вращения оснований и пар оснований относительно друг друга представляет собой более сложную задачу, поскольку в отличие от преобразований смещения в трехмерном пространстве группа преобразований вращения является некоммутативной, то есть зависит от последовательности применения операций вращения. Так в механике для описания вращения твердых тел зачастую используется углы Эйлера или матрицы поворота вокруг осей системы координат. Однако преобразования описываемые данными углами необходимо выполнять в определенной последовательности, конечный результат сложным образом зависит от всех трех углов. В случае описания взаимного расположения оснований ДНК таким образом углы Эйлера будут иметь трудно интерпретируемый физический смысл. Отдельной проблемой является то, что при изменении направления рассмотрения преобразований (например, M-N вместо N-M или рассмотрении динуклеотидных переходов вдоль по одной и другой цепи ДНК) численные значения параметров вращения будут отличаться. Данная проблема была очевидна с конца 1970-ых годов (в 1981 году была установлена первая кристаллическая структура ДНК - додекамер Дрю-Дикерсона[102]), обсуждалась на встрече 1988 года в Кембридже (“Cambridge Accord”). Для разрешения данной проблемы был сформулирован ряд подходов, которые для определения и вычисления параметров вращения оснований ДНК вводят понятие дополнительной системы отсчета, находящейся по середине между рассматриваемыми системами отсчета (см. напр. работу 1979 года В.Б. Журкина и др. [98]). Используя дополнительную систему отсчета можно сформулировать определение параметров вращения, которые будут независимы от направления рассмотрения преобразования. Наиболее используемый подход был предложен Эль Хассаном и Калладайном и описан в статьях [100; 103]. Рассмотрим этот подход на примере расчета динуклеотидных параметров вращения Twist, Roll, Tilt (Рис. 1.9e), расширение данного подхода на параметры

вращения оснований относительно друг друга Buckle, Propeller, Opening может быть произведено по аналогии.

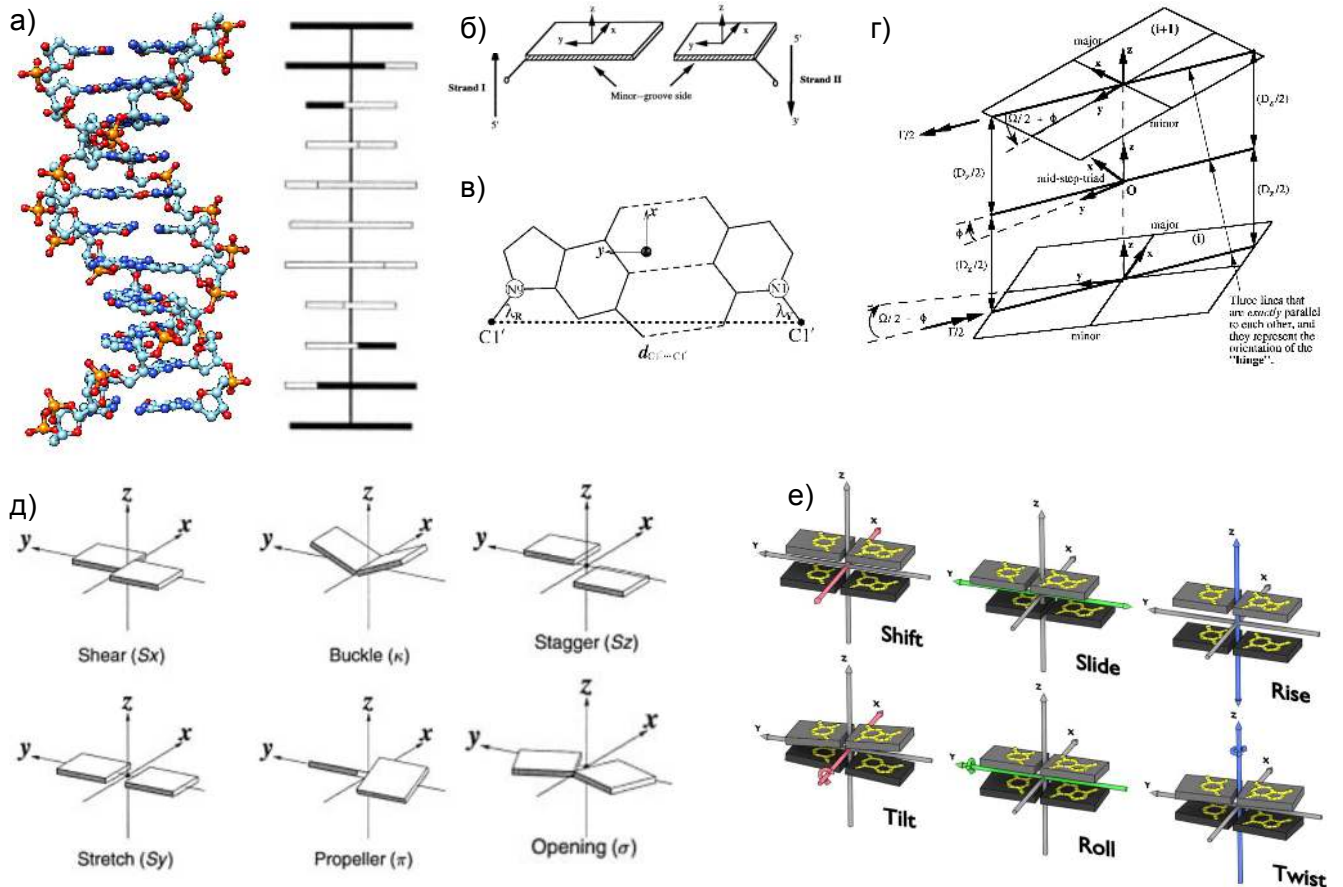
Для вычисления углов Twist, Roll и Tilt вводится система отсчета (“mid-step reference frame”), находящаяся по середине между системами отсчета, связанными с последовательными парами оснований (см. Рис. 1.9г). Для преодоления проблемы некоммутативности операций вращения используется следующий прием. Вращение систем отчета друг относительно друга может быть представлено как вращение вокруг оси Z (Twist) на угол  $\Omega$ , а затем отклонение оси Z на некоторый угол  $\Gamma$  вокруг какой-либо оси перпендикулярной оси Z (называемой остью RollTilt). В случае работы относительно промежуточной системы отсчета (“mid-step reference frame”) процедура разлагается на поворот верхнего основания на угол  $\Omega/2$  и угол  $\Gamma/2$ , а нижнего на угол  $-\Omega/2$  и угол  $-\Gamma/2$ , относительно оси RollTilt, положение которой задается углом  $\varphi$  относительно оси Y промежуточной системы отсчета. Углы Roll ( $\rho$ ) и Tilt ( $\tau$ ) определяются векторным разложением по формулам:

$$\rho = \Gamma \cos(\varphi) \quad (1.28)$$

$$\tau = \Gamma \sin(\varphi) \quad (1.29)$$

Позитивный Roll в A/B и W-формах ДНК соответствует открытию малой бороздки, знак параметра Tilt зависит от направления рассмотрения ДНК. Более подробное определение методов расчета параметров в современной реализации программы 3DNA может быть найдено в [103].

Можно показать, что соответствующие процедуры позволяют определить схему, которая с точностью до знака дает одинаковые значения параметров при рассмотрении ДНК как в прямом, так и в обратном направлениях. Важной особенностью данного подхода является то, что структура ДНК может быть однозначно восстановлена по наборам рассчитанных параметров. Таким образом, можно проводить конвертацию атомистических структур в пространство параметров оснований ДНК и обратно, а также генерировать структуры с измененной последовательностью, но идентичной геометрией.



**Рисунок 1.9** — Описание геометрии ДНК в виде геометрических параметров. а) Атомистическое представление В-формы ДНК и представление в виде диаграммы по принципам Калладайна и Дрю [104]; б) представление пары оснований в виде параллелепипедов и связанные с ними системы координат [100], в) система координат и ее расположение относительно пары оснований [101], г) иллюстрация введения промежуточной системы отсчета между парами оснований для расчета параметров вращения [100], д) параметры взаимного расположения оснований в паре оснований [99], е) параметры взаимного расположения пар оснований относительно друг друга вдоль ДНК.

### 1.4.2 Эмпирические константы жесткости изгиба ДНК

Возможность задания общей конформации ДНК в виде набора 12 параметров, определяющих взаимное расположение азотистых оснований, позволяет поставить вопрос о возможности задания некоторого силового поля, способного описывать энергию различных конформаций двухцепочечной ДНК как функ-

цию этих параметров. Для двухцепочечной ДНК любой последовательности, содержащей  $N$  пар оснований,  $6(N-1)$  параметров будут описывать взаимное расположение пар оснований относительно друг друга. Формально с точки зрения статистической физики, поскольку каждой конкретной конформации ДНК однозначно можно сопоставить численные значения этих параметров и параметры являются независимыми, то набор этих параметров можно рассматривать как набор обобщенных координат (координат реакции), в пространстве которых может быть задана функция свободной энергии.

В общем случае данная функция будет иметь сложную структуру. Для того, чтобы такой формализм мог быть использован на практике, необходимо вводить ряд приближений относительно структуры данной функции. Наиболее используемым является, так называемое, динуклеотидное аддитивное приближение. Предполагается, что функция свободной энергии аддитивно складывается из компонент, описывающих вклады отдельных динуклеотидных шагов:

$$F_{tot}(\{\theta_j^1\}, \{\theta_j^2\}, \dots, \{\theta_j^{N-1}\}) = \sum_{i=1}^{N-1} F_k(\{\theta_j^k\}) \quad (1.30)$$

где  $k$  - номер динуклеотидного шага,  $\{\theta_j^k\}$  - набор 6 переменных Roll, Tilt, Twist, Shift, Slide, Stagger для динуклеотидного шага  $k$ .

Следующим упрощением является введение гармонического приближения для представления свободной энергии каждого динуклеотидного шага в виде разложения функции до второго порядка вблизи точки равновесия.

$$F_k(\{\theta_j^k\}) = F_0 + \frac{1}{2} \sum_{i=1}^6 \sum_{j=1}^6 f_{ij} \Delta\theta_i^k \Delta\theta_j^k \Delta\theta_i^k = (\theta_i^k - \hat{\theta}_i^k) \quad (1.31)$$

Наконец, последним важным приближением является предположение о том, что коэффициенты  $f_{ij}$  и равновесные значения  $\hat{\theta}_i^k$  зависят только от типа конкретного динуклеотидного шага и не зависят от типа соседних нуклеотидов. Такое приближение называется приближением ближайших соседей (nearest neighbor). Следует отметить, что это лишь приближение, и в некоторых особых случаях учета ближайших соседей не хватает для адекватного описания - например, когда важны коллективные эффекты вдоль по цепи, связанные с бифуркациями водородных связей, образованием цепочек воды в малой бороздке и т.д. Например, такие эффекты известны при рассмотрении протяженных участков

аденина или тимина (А-траки), которые обладают высокой изгибной жесткостью.

Оценка констант  $f_{ij}$  для каждого типа динуклеотидных шагов (всего их 10: CG, CA, TA, AG, GG, AA, GA, AT, AC, GC; в силу симметрии ДНК, напр. динуклеотидные шаги AA и TT обозначают один и тот же шаг, но по разным цепям) является нетривиальной задачей. На сегодняшний день существует два подхода - анализ кристаллических структур или расчеты методами молекулярного моделирования на основе атомистических силовых полей.

Первый подход был в значительной степени реализован в работе [105]. Авторами был проанализирован набор из 92 комплексов белок-ДНК и построены распределения динуклеотидных параметров в этих структурах (см. Рис. 1.10). Предполагается, что изучая ансамбли распределений динуклеотидных параметров в комплексах ДНК-белок, можно получить информацию о подвижности и гибкости ДНК самой по себе. Аргументация данного предположения авторами сводится к следующему. Различные белки при связывании оказывают различные воздействия на конформацию молекулы ДНК, таким образом, при достаточно большом усреднении и рассмотрении конформации ДНК вблизи равновесия в статистике будет проявляться лишь естественная склонность ДНК к деформациям. Более строгим образом такое предположение обосновать сложно. Однако из анализа структур белков, известно, что распределение различных параметров полипептидной цепи в глобулярных белках обладает квази-больцмановской статистикой по отношению к энергетическим функциям, описывающим их локальные взаимодействия. Например, встречаемость различных ротамеров боковых цепей аминокислот согласуется с больцмановским распределением, описываемым функцией потенциальной энергии вращения вокруг соответствующей связи. Аналогичные наблюдения имеются для экспонированности боковых цепей аминокислот в воду и энергии их гидратации [106]. Теоретическое объяснение этого явления было дано Финкельштейном и др. [107]. Описывая белок как гетерополимер согласно модели случайных энергий (Random Energy Model), можно показать, что количество аминокислотных последовательностей, которое стабилизируют тот или иной элемент белкового фолда, экспоненциально убывает с увеличением энергии этого элемента. Таким образом, квази-Больцмановская статистика является следствием устройства энергетического спектра нативных белков, у которых должна быть одна выраженная нативная конформация, энергетический уровень которой является

самым низким по энергии и отделен энергетической щелью. Следуя этой логике, можно предположить, что количество различных белков, которые могут изгибать ДНК с той или иной силой, также будет экспоненциально зависеть от энергии изгиба ДНК.

Step	$N$	Twist, deg		Tilt, deg		Roll, deg		Shift, Å		Slide, Å		Rise, Å	
CG	88	36.1	(5.5)	0	(4.2)	5.4	(5.2)	0	(0.87)	0.41	(0.56)	3.39	(0.27)
CA	110	37.3	(6.5)	0.5	(3.7)	4.7	(5.1)	0.09	(0.55)	0.53	(0.89)	3.33	(0.26)
TA	134	37.8	(5.5)	0	(2.7)	3.3	(6.6)	0	(0.52)	0.05	(0.71)	3.42	(0.24)
AG	106	31.9	(4.5)	-1.7	(3.3)	4.5	(3.4)	0.09	(0.69)	-0.25	(0.41)	3.34	(0.23)
GG	97	32.9	(5.2)	-0.1	(3.7)	3.6	(4.5)	0.05	(0.76)	-0.22	(0.64)	3.42	(0.24)
AA	129	35.1	(3.9)	-1.4	(3.3)	0.7	(5.4)	-0.03	(0.57)	-0.08	(0.45)	3.27	(0.22)
GA	117	36.3	(4.4)	-1.5	(3.8)	1.9	(5.3)	-0.28	(0.46)	0.09	(0.70)	3.37	(0.26)
AT	140	29.3	(4.5)	0	(2.5)	1.1	(4.9)	0	(0.57)	-0.59	(0.31)	3.31	(0.21)
AC	137	31.5	(4.2)	-0.1	(3.1)	0.7	(3.9)	0.13	(0.59)	-0.58	(0.41)	3.36	(0.23)
GC	86	33.6	(4.7)	0	(3.9)	0.3	(4.6)	0	(0.61)	-0.38	(0.56)	3.40	(0.24)

**Рисунок 1.10** — Матрица средних значений и отклонений динуклеотидных параметров из статьи [105].

Вычислив параметры отклонения ДНК в зависимости от типа динуклеотидного шага, а также дисперсии этих параметров, можно ставить задачу об оценке констант жесткости изгиба ДНК по этим различными параметрам  $f_{ij}$  методом обратного гармонического анализа. В случае одномерной пружины, энергия которой описывается гармоническим законом  $E = \frac{1}{2}fx^2$ , распределение координаты согласно закону Больцмана будет выглядеть как  $P(x) \sim \exp -\frac{fx^2}{2kT}$ . Последнее является нормальным распределением с параметром дисперсии  $\langle x^2 \rangle = kT/f$ . Таким образом, мы установили связь дисперсии величины с константой жесткости. В общем случае, когда энергия описывается функцией вида 1.31, распределение состояний описывается многомерным нормальным распределением вида:

$$P_{\mathbf{X}}(x_1, \dots, x_k) = \frac{\exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right)}{\sqrt{(2\pi)^k |\boldsymbol{\Sigma}|}} \quad (1.32)$$

где  $|\boldsymbol{\Sigma}| = \det \boldsymbol{\Sigma}$ , а  $\boldsymbol{\Sigma}$  - является матрицей ковариаций. Таким образом, несложно увидеть, что

$$\langle \Delta\theta_i \Delta\theta_j \rangle = kT[\mathbf{F}^{-1}]_{ij} \quad (1.33)$$

где  $\mathbf{F}$  матрица коэффициентов  $f_{ij}$ ,  $k$  - коэффициент Больцмана,  $T$  - конформационная температура, которую можно определить, например, сравнив получаемую гибкость с известными данными о персистентной длине. Константы жесткости доступны по адресу <https://web.archive.org/web/20151025043044/http://rutchem.rutgers.edu/~olson/pdna.html>.

Альтернативным подходом является вычисление характеристик динуклеотидных параметров путем молекулярного моделирования на основе атомистических силовых полей. Наиболее актуальной и современной здесь является работа [108]. Авторы использовали наработки АВС консорциума, участники которого промоделировали методом атомистической молекулярной динамики все возможные последовательности тетрануклеотидов (всего их 136) [109]. Из данных траекторий можно получить данные о флуктуациях различных параметров. Было показано, что распределения динуклеотидных параметров в некоторых случаях являются мультимодальными и их сложно аппроксимировать в виде одного многомерного гауссова распределения. Поэтому в данной статье некоторые распределения параметров описываются линейной комбинацией многомерных гауссовых распределений. Кроме того авторы делают попытку сделать коэффициенты  $f_{ij}$  зависимыми от соседних с динуклеотидным шагом пар оснований - эффективно учитывая параметры изгибной жесткости ДНК на тетрануклеотидном уровне.

Суммируя вышесказанное, отметим достоинства и недостатки различных подходов. Гармоническое приближение при описании гибкости ДНК безусловно имеет свои ограничения, особенно в случаях, когда белок сильно деформирует ДНК. Например, в нуклеосомах динуклеотиды пиримидин-пурин (YR) обычно выгибаются в сторону малой бороздки (параметр Roll отрицательный), в то же время из Рис. 1.10 видно, что в среднем в структурах динуклеотиды имеют положительный Roll, причем, динуклеотид ТА (YR) больше склонен к положительному Roll, чем АТ (RY). Однако предполагается, что при сильных изгибах (изломах, kinks) поверхность энергии устроена более сложным образом, и при комбинации с положительными значениями Slide - как раз АТ выгоднее гнуться в сторону малой бороздки, чем ТА [110]. В этом плане мультимодальное приближение изгибной жесткости ДНК может являться более точным. Однако, для мультимодального приближения требуется большое количество качественных данных о динамике ДНК, которые, казалось бы, можно взять из молекулярно-динамических расчетов. Однако, здесь есть свой подводный камень - атомистические силовые поля, существующие на сегодняшний день, не совсем корректно описывают зависимость изгибной жесткости ДНК от последовательности в сравнении со структурными данными. Так, например, по данным вычислений молекулярной динамики максимальным значением параметра Roll



Roll обладает динуклеотид TG [111], а по данным анализа кристаллических структур CG.

### 1.4.3 Программные продукты

Для анализа геометрии ДНК был создан достаточно большое набор программных продуктов и методов, к ним можно отнести программы CompDNA[112], SCHNAaP [103], Curves+ [113], 3DNA [114]. На данный момент наиболее используемой и поддерживаемой является программа 3DNA, которая позволяет не только анализировать структуру, но и реконструировать атомистическую структуру по набору параметров. Для анализа ширины и глубины бороздок ДНК может использовать программа Curves+. Моделирование конформации ДНК в динуклеотидном приближении сводится к реализации либо алгоритмов минимизации конформации ДНК в пространстве параметров, либо к динамическим расчетам с помощью метода Монте-Карло. Многие такие программы разрабатывались для решения конкретных научных задач внутри научных групп, например, DNAMiniCarlo (N.B. Ulyanov, A.A. Gorin, V.B. Zhurkin), аналогичные подходы описаны в ряде статей [115; 116], недавно появилась программа MC-enNN и веб-сервер к ней [108]. В нашей научной группе проводится разработка программы с открытым программным кодом PyNAMod <http://github.com/intbio/PyNAMod>, способной моделировать ДНК, включая ДНК-белковые комплексы в динуклеотидном приближении методами минимизации геометрии ДНК, а также методом цепей Монте-Карло.

### 1.4.4 Алгоритмы интегративного моделирования

Введение дополнительных ограничений при моделировании ДНК в динуклеотидном приближении, а также учет стерических ограничений путем конвертации структуры из пространства динуклеотидных параметров в атомистическую структуру, позволяют сформулировать подходы для интегративного

моделирования ДНК и комплексов-ДНК белок на основе различных экспериментальных данных.

В реализованных нами подходах при минимизации конформации ДНК в динуклеотидном приближении в энергию системы вводились дополнительные члены, соответствующее отклонению набора расстояний от целевых значений, измеренных при помощи экспериментальных методов, например, spFRET. Алгоритм действия программы изображен на Рис. 1.11. На вход программа получает начальные структуры в формате PDB. Измерения методом spFRET дают информацию о расстояниях между парами меток. Алгоритм последовательно оптимизирует координаты нуклеотидных пар таким образом, чтобы привести их геометрию в соответствие с результатами spFRET. В ходе работы алгоритма учитываются ограничения, связанные с жесткостью ДНК.

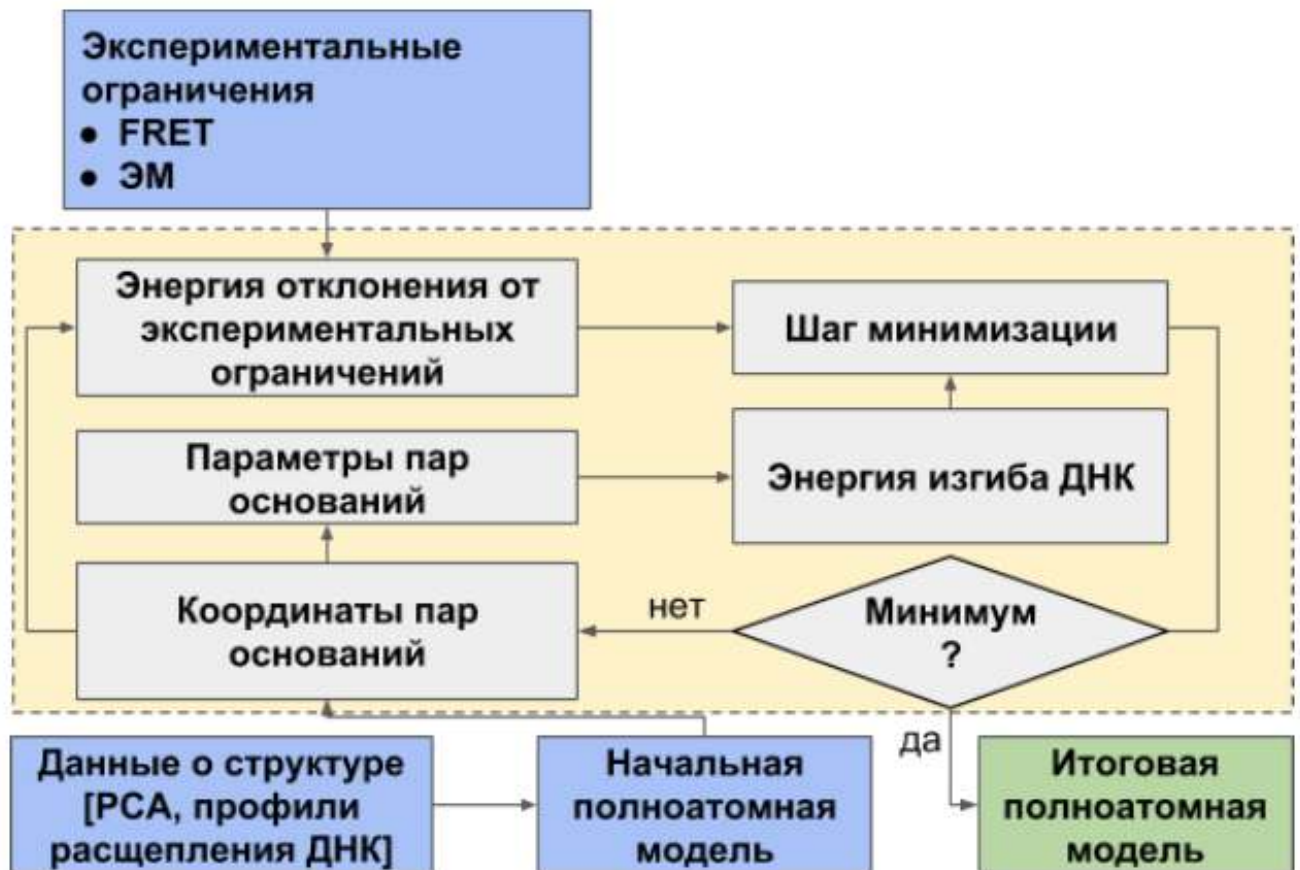


Рисунок 1.11 — Алгоритм интегративного моделирования на основе моделей ДНК в динуклеотидном приближении.

## 1.5 Методы моделирования супрануклеосомной структуры хроматина

В последние годы благодаря совершенствованию экспериментальных технологий наметилась долгожданная конвергенция методов структурной биологии и методов геномики в изучении организации хроматина на молекулярном уровне. Благодаря успехам крио-электронной микроскопии, мы получаем все больше информации о структуре не только нуклеосом, но и супрануклеосомных структур, а благодаря развитию методов геномики и 3-D геномики все более высокого разрешения, стало возможным определять контакты между локусами ДНК с суб-нуклеосомным разрешением (в частности, методы Micro-C, Micro-C-XL и др.) и определять положение и состав нуклеосом вдоль генома (напр. методы MNase-seq, MNase-ChIP-seq). Возникает необходимость в интерпретации экспериментальных данных и построении физических молекулярных моделей укладки хроматина на супрануклеосомном уровне с учетом реальных геометрических и топологических параметров молекул белков и ДНК, а также динамического характера этих взаимодействий.

Так как изучение организации хроматина на супрануклеосомном уровне довольно давно представляет собой фундаментальную задачу для молекулярной и клеточной биологии (примем за дату начала год открытия нуклеосомы и структуры в виде “бусин на нити” - 1974 [117]), было разработано множество подходов к молекулярному моделированию укладки нуклеофиламента. Большую часть прошедшего времени сомнений в концепции 30-нанометровой фибриллы не возникало (идею 30-нанометровой фибриллы предложили в 1980 году, серьезные аргументы против появились в 2000-ых, 2010-ых см., например, [118; 119]). Поэтому подавляющее число работ по молекулярному моделированию хроматина на супрануклеосомном уровне являются по сути работами по моделированию 30-нанометровой фибриллы, что, не умаляя их методологической значимости, значительно снижает ценность результатов и приводит к необходимости проведения работ по молекулярному моделированию хроматина на супрануклеосомном уровне с учетом современных знаний о нем. При этом изучение предыдущей литературы по данной теме все-таки необходимо, так как ее авторы достигли значительных результатов в разработке математического и концептуального аппарата для моделирования хроматина на этом

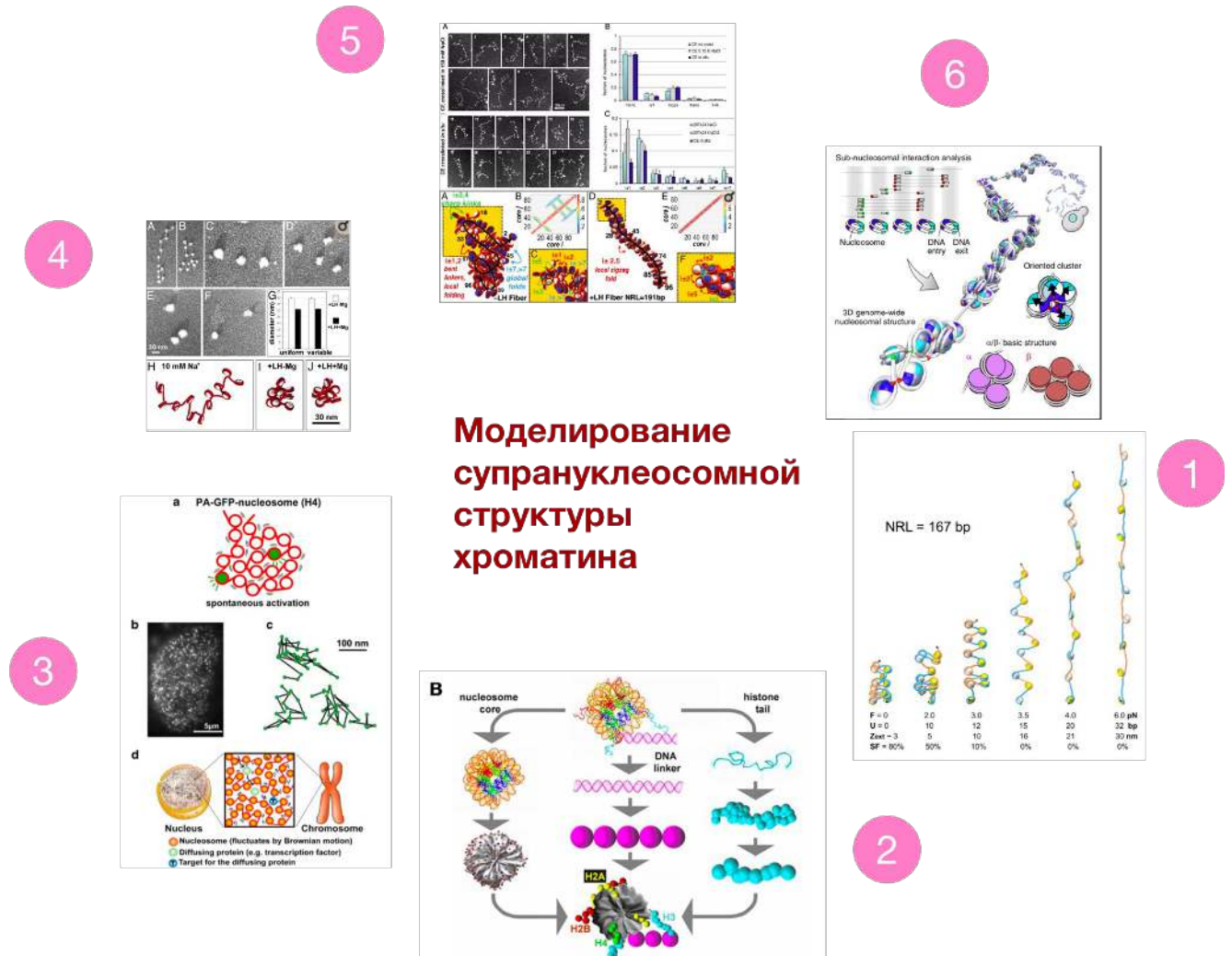
уровне вообще. Например, в работе 2015 года Norouzi, D. & Zhurkin, V. B. Topological Polymorphism of the Two-Start Chromatin Fiber [115] авторы с помощью своей модели изучили все возможности организации двухстартовой хроматиновой фибриллы (есть два гипотетических типа организации 30-нанометровой хроматиновой фибриллы: соленоидная - одностартовая фибрилла - и зигзагом - двухстартовая фибрилла - причем тип фибриллы коррелирует с длиной лнкерного участка), т.е. смоделировали все возможные ее конформационные варианты. Фибриллы моделировали симметричными и регулярными, а для расчетов вводилось 4 параметра суперспирализации: кручение, подъем, наклон нуклеосом и диаметр. Энергия фибриллы задавалась как сумма четырех слагаемых: эластической энергии линкерной ДНК, стерического отталкивания, электростатики (рассчитываемой с помощью кулоновского потенциала), феноменологического взаимодействия между двумя “сложенными в стопку” нуклеосомами. При оптимизации энергии фибриллы с учетом параметров суперспирализации исследователи впервые обнаружили два типа топологического перехода в ней: путем резкого на 360 градусов поворота в кручении линкера и путем скрещивания линкеров. Даже при моделировании структуры, отличной от 30-нанометровой фибриллы, эти принципы могут быть релевантными, так как позволяют рассчитать такие общие вещи, как энергию комплекса ДНК-нуклеосома, его оптимальную геометрию в зависимости от длины линкеров и внешних факторов. Эти же авторы в более поздней работе того же года продолжали исследовать топологический полиморфизм хроматиновой фибриллы, пытаясь установить связь между длиной линкеров и спецификой сверхспирализации ДНК (связанной с уровнем транскрипции). Интересно, что для нахождения оптимальной конформации линкерной ДНК использовался мезоскопический подход: ДНК моделируется динуклеотидными “шагами”, его траектория описывается с помощью шести параметров: Tilt, Roll, Twist, Shift, Slide, Rise. Мезоскопический подход был предложен в статье Olson et al., 1993 [120], данные шестимерные координаты - в работе Dickerson et al., 1989 [97]. Оба этих концептуальных инструмента широко используются в моделировании хроматина на низких уровнях компактизации (10-нанометровой фибриллы и следующем) и никак не зависят от концепции 30-нанометровой фибриллы, что делает их полезными и для моделирования альтернативных этой фибрилле структур. Крайне эффективным для молекулярного моделирования супрануклеосомного уровня компактизации хроматина являются методы, основанные

на принципах Монте-Карло. Познакомиться с ними можно, например, в работе — Norouzi, D. & Zhurkin, V. B. Dynamics of Chromatin Fibers: Comparison of Monte Carlo Simulations with Force Spectroscopy [121]. Шаги симуляции следующие: случайный выбор одной пары оснований в подвижном линкере и изменение его шести параметров, обновление позиций нуклеосом в нуклеосомном массиве, вычисление разницы в энергии между старым и новым состоянием и выполнение теста Метрополиса. Варьируя величину внешних сил, степень развернутости (unwrapping) нуклеосом, длину нуклеосомных повторов (NRL - nucleosome repeat length) - среднее расстояние между центрами соседних нуклеосом и значение энергии стэкинга (stacking energy), исследователи провели серию симуляций, каждая из которых включала около 100 миллионов описанных выше шагов. Подобный подход, особенно его вариант - моделирование путем имитации отжига - позволяет довольно эффективно находить глобальный энергетический минимум в заданной системе.

Также часто применяется для решения задач моделирования хроматина на разных уровнях организации метод огрубленного моделирования (coarse-grained). Суть метода состоит в представлении молекул и молекулярных структур (например, нуклеосом) простыми телами и поверхностями: сферами, цилиндрами, плоскостями. Такое приближение значительно упрощает вычисления и позволяет проводить однозначный анализ за приемлемое время. Однако при этом исходные данные подвергаются соответствующей деформации - результат моделирования соответствует им не полностью, он также зависит от выбранной стратегии огрубления. Хроматиновые фибриллы и их динамику обычно моделируют как цепочки из шаров или как червеобразные структуры [122–124]. Несмотря на популярность метода огрубленного моделирования в области изучения структуры хроматина, он является далеко не оптимальным для решения этих задач: исходные данные не используются полностью, искажаются, что может быть критичным, когда речь заходит об изучении укладки 10-нанометровой фибриллы, так как образуемая ею структура зависит от не учитываемых в огрубленных методах параметров - например, длины линкера [115; 125]. Для адекватного моделирования структуры хроматина требуется основывать его на данных высокого разрешения (например, современных протоколов Hi-C), причем они не должны искажаться так сильно, как в методе огрубленного моделирования. Таким образом необходимо разработать новый,

более чувствительный к данным и параметрам наподобие длины линкеров подход.

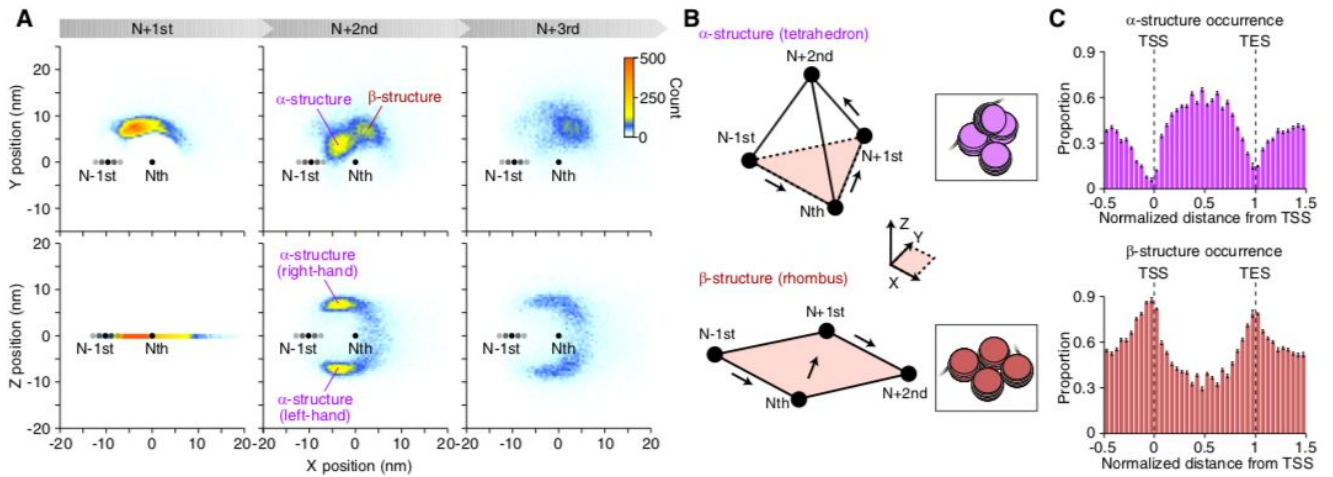
Важно отметить, что большинство описываемых выше подходов основаны на “закрытом” программном обеспечении. Авторы либо не предоставляют доступа к программному обеспечению, либо оно является слабо используемым, поскольку к нему нет описания, либо возникают проблемы с переносимостью.



**Рисунок 1.12** — Подходы к моделированию супрануклеосомной структуры, в том числе с учетом экспериментальных данных. (1) - Молекулярная динамика с использованием принципов Монте-Карло [121]; (2) - Огрубленное моделирование [126]; (3) - (6) : интегративные методы. (3) - Моделирование на основе данных метода Single nucleosome imaging - отображения одиночных нуклеосом [127]; (4), (5) - Моделирование на основе данных метода EMANIC - метода обнаружения нуклеосом с помощью электронной микроскопии [128], [129]); (6) - Моделирование супрануклеосомной структуры на основе данных модифицированного Micro-C [130]

Важным подходом является интеграции различных экспериментальных данных при построении моделей - данный подход называется интегративным моделированием. В применении к супрануклеосомной структуре он на наш взгляд развит весьма слабо. На данный момент по данным литературы известна лишь одна попытка интеграции данных Micro-C с методами моделирования. Работа опубликована в 2019 году [130]. В ней приведены довольно убедительные доказательства существования определенных супрануклеосомных структур в геноме пекарских дрожжей, альтернативных отвергнутой хроматиновой фибрилле, - хроматиновых альфа-тетраэдров и бета-ромбов - то есть этой работой был сделан большой шаг на пути к определению структуры хроматина на до сих пор не исследованных уровнях компактизации генома. Молекулярное моделирование проводилось на основании теоретической модели нуклеосомы - как комплекса из гистонового октамера и ассоциированной с ним ДНК — и данных Micro-C, позволяющих различать точки “входа” и “выхода” ДНК в/из нуклеосом. Такой комбинированный метод был назван Hi-Co. Само моделирование проводилось с помощью подхода имитации отжига в методе молекулярной динамики. Имитация отжига была необходима для оптимизации позиций и ориентаций отдельных нуклеосом, остальные структурные факторы, такие как, например, линкерная ДНК (ее длина), были неявно включены в потенциалы и по сути не рассматривались детально. Несмотря на важное значение, подходы моделирования в работе Ohno et al. не вполне реалистичны, так как, например, отсутствуют в явном виде линкерная ДНК, а следовательно не учитывается влияние топологии и кручения ДНК важные для понимания многих функциональных процессов и состояний в хроматине. Также в подходе авторов не предполагался учет нуклеосом различного типа и возможного откручивания ДНК от нуклеосомы с изменением углов входа/выхода ДНК. Кроме того работа Ohno et al. была посвящена изучению генома дрожжей — это значит, что непосредственной значимостью для медицины она не обладает. Требуется провести аналогичное исследование генома высших, многоклеточных эукариот и человека.

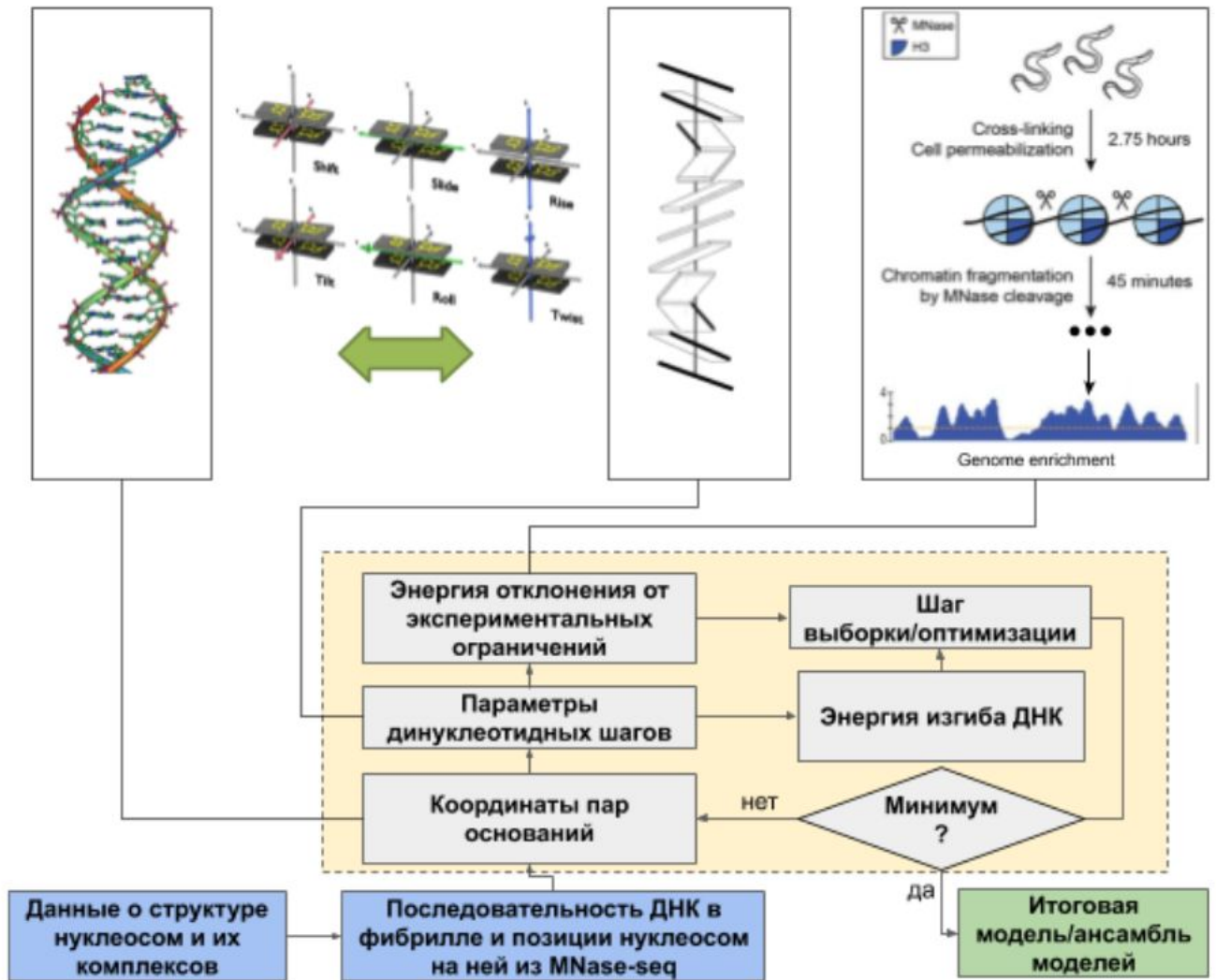
Для моделирования супрануклеосомной укладки фибрилл алгоритмы, предложенные нами в пункте 1.4.4, могут быть адаптированы путем включения дополнительных данных из экспериментов Micro-C, а также данных о позициях нуклеосом из экспериментов MNase-seq, ATAC-seq (см. Рис. 1.14). Расчет обобщенных переменных из атомистической структуры и восстановление



**Рисунок 1.13** — Визуализация результатов анализа взаимного расположения нуклеосом в пространстве и выявленные в результате него мотивы супрануклеосомной структуры ДНК из статьи Ohno et. al, 2019

координат из набора переменных, описывающих геометрию ДНК в динуклеотидном приближении, производится при помощи программы 3DNA [99], либо при помощи разрабатываемой программной библиотеки PyNAMod. Учет положения и геометрии белковых компонент можно производить в том же приближении: для этого необходимо определять их ориентацию относительно ключевых динуклеотидов. Например, при определении гистонового ядра нуклеосомы, отсчитывать его положение относительно диадного динуклеотида. Дополнительно в физическую модель нуклеосомной фибриллы вводятся парные потенциалы для учета исключенного объема и электростатических взаимодействий между соседними нуклеосомами, что необходимо для учета изменяющихся заряд посттрансляционных модификаций гистоновых хвостов. При моделировании цепочек нуклеосом, вводятся ограничения на расстояния, описывающие межнуклеосомные контакты. Возможно также введение дополнительного потенциала, позволяющего нуклеосомам смещаться относительно их изначальных положений (что обосновано точностью определения их положения в эксперименте). В рамках описываемой физической модели нуклеосомной фибриллы можно создавать модели организации хроматина путем применения численных методов минимизации значений функции, например метода градиентного спуска, однако, для обнаружения глобального минимума, а также создания множественной выборки из конфигурационного пространства более применимы стохастические подходы, такие как метод Монте-Карло с алгоритмом Метрополиса.





**Рисунок 1.14** — Алгоритмы интегративного моделирования нуклеосомных фибрилл. Слева сверху - схематическое описание ДНК в приближении динуклеотидных шагов. Справа сверху - сокращенное схематическое представление методов Micro-C/MNase-seq/MNase-ChIP-seq. Снизу - Описание подхода интеграции экспериментальных данных в физическую модель нуклеосомных фибрилл.

## 1.6 Основы некоторых экспериментальных методов

### 1.6.1 Методы футпринтинга ДНК

Одним из подходов к изучению структуры ДНК и ДНК-белковых комплексов являются методы физического, химического или ферментативного футпринтинга (от англ. footprint - след). Идея методов состоит в обработке

ДНК определенными агентами, которые вносят изменения в ее химическую структуру в зависимости от ее конформации, состава или взаимодействия с другими макромолекулами, с последующим анализом продуктов реакции для установления мест химических реакций вдоль ДНК.

К физическим методам относят облучение ДНК ультрафиолетом [131] или рентгеном, что приводит к разрыву ДНК в областях не защищенных белком. С помощью секвенирующего геля электрофореза можно определить длины продуктов реакций и следовательно участки взаимодействия с белком. Аналогичный подход может быть осуществлен с использованием генерируемых в реакционной смеси гидроксил радикалов химическими методами. Методам гидроксильного футпринтинга посвящена глава 3 данной диссертационной работы. Наиболее популярным методом ферментативного футпринтинга является использование фермента ДНКазы I, обладающего выраженной эндонуклеазной активностью. К плюсам такого подхода относится простота осуществления эксперимента, к минусам большой размер фермента (около 4 нм), что уменьшает детальность футпринтинга, и наличие некоторой специфичности по последовательности (гидролизует ДНК преимущественно около пиримидинов). Еще одним интересным методом химического футпринтинга является футпринтинг с использованием перманганата калия. Перманганат калия связывается с тиминами в одноцепочечной ДНК, окисляет их, что приводит в конечном итоге к разрыву в ДНК. Таким образом можно исследовать локализацию пузырей транскрипции в геноме, нестандартных ДНК структур [132]. Футпринтинг с использованием диметилсульфата (ДМС) используется также для изучения ДНК-белковых взаимодействий. ДМС индуцирует метилирование остатков гуанина. Два образца обрабатывают пиперидином, чтобы вызвать химическое расщепление остатков гуанина, модифицированного ДМС, с последующим расщеплением рестрикционными ферментами. После маркировки образцы запускают параллельно на геле для визуализации. Отсутствующие полосы в образце, связанном с белком, соответствуют остаткам гуанина, защищенным от модификации посредством взаимодействия [133].

## 1.6.2 FRET спектроскопия и микроскопия

### Ферстеровский резонансный перенос энергии

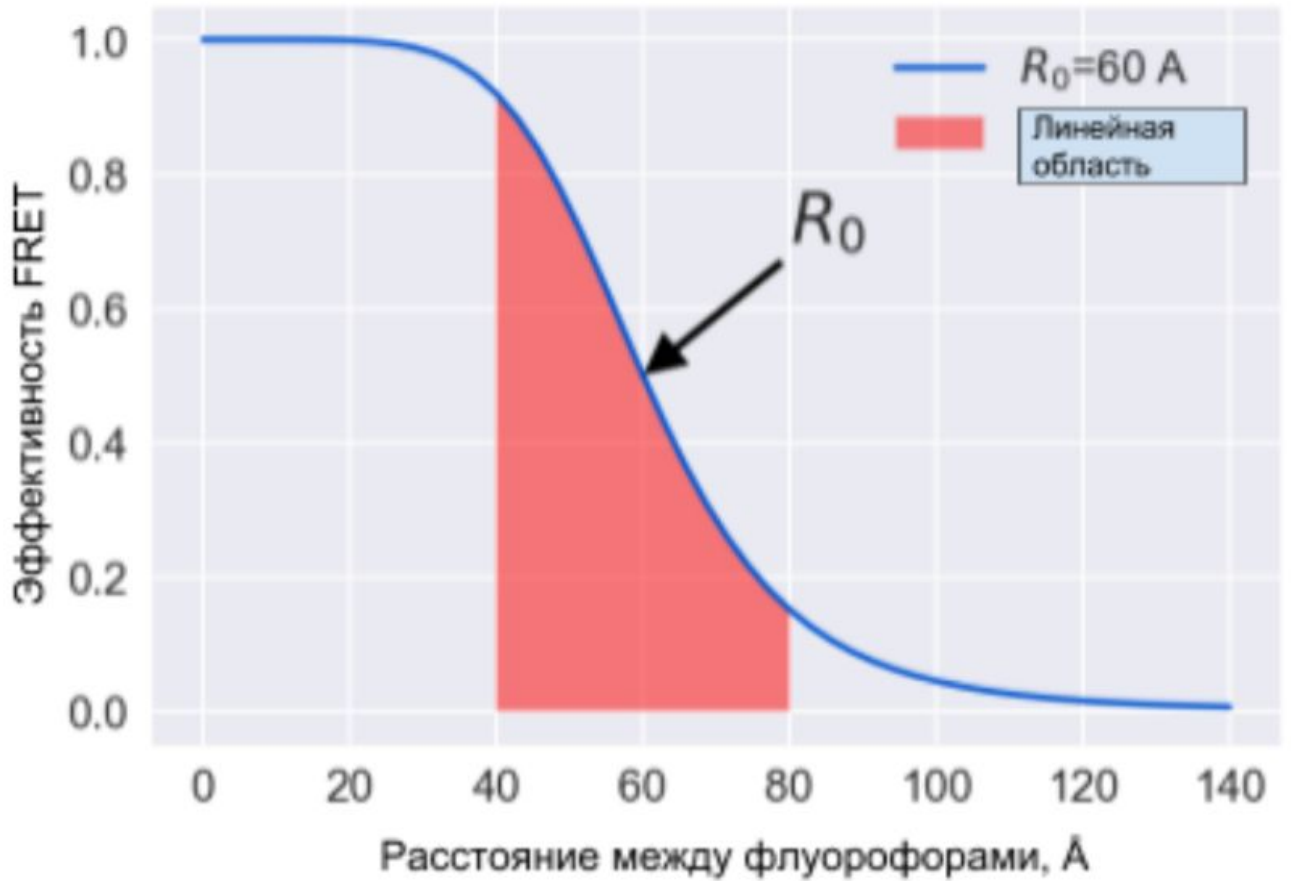
Ферстеровский резонансный перенос энергии (Forster resonance energy transfer, FRET) это эффект безызлучательного переноса энергии с одного красителя-флуорофора на другой. Пару (или более) флуорофоров подбирают таким образом, чтобы спектр поглощения акцептора перекрывался со спектром флуоресценции донора. Флуорофоры могут быть пришиты к аминокислотам или азотистым основаниям ДНК с помощью различных линкерных молекул. По определению эффективность FRET - это отношение вероятности переноса энергии с донора на акцептор по отношению ко всем другим переходам в расчете на событие возбуждение донора:

$$E_{FRET} = \frac{k_{ET}}{k_f + k_{ET} + \sum k_i} \quad (1.34)$$

где  $k_{ET}$  - скорость переноса энергии с донора на акцептор,  $k_f$  - скорость процесса флуоресценции донора,  $k_i$  - скорости других неизлучательных процессов релаксации донора, за исключением переноса энергии на акцептор. Эффективность переноса энергии существенно зависит от расстояния между метками и может быть описана следующей формулой:

$$E_{FRET} = \frac{1}{1 + (R/R_0)^6} \quad (1.35)$$

где  $R$  - расстояние между метками,  $R_0$  - радиус Ферстера, расстояние между метками, отвечающее 50% эффективности переноса энергии между флуорофорами (см. Рис. 1.15). Согласно данной формуле, измеряя эффективность FRET возможно вычислить расстояние между участками меченой макромолекулы. Для каждой пары меток радиус Ферстера индивидуален, он зависит от ряда факторов, в том числе от квантового выхода флуоресценции донора, ориентации красителя на поверхности молекулы (ориентационный фактор), коэффициента преломления среды и т.д.



**Рисунок 1.15** — Зависимость эффективности переноса энергии от расстояния. Стрелкой показана величина Ферстеровского радиуса. Красным показана область изменений, близкая к линейной.

Эффективность переноса энергии может быть выражена через отношение времен жизни флуоресценции донора в отсутствии и присутствии акцептора:

$$E_{FRET} = 1 - \tau'_D / \tau_D \quad (1.36)$$

где  $\tau'_D$  - время жизни флуоресценции донора в присутствии акцептора,  $\tau_D$  время жизни донора в отсутствии акцептора.

Альтернативным является выражение эффективности переноса через величины интенсивности флуоресценции донора и акцептора:

$$E_{FRET} = \frac{F_A}{F_A + \gamma F_D} \quad (1.37)$$

где  $F_A = I_A - B_A - \alpha_{DA}(I_D - B_D) - D_{ex}$ ,  $F_D = I_D - B_D - \alpha_{AD}(I_A - B_A)$  и в последних формулах  $I_A$  и  $I_D$  - регистрируемые интенсивности сигнала акцептора и донора, соответственно;  $B_A$  и  $B_D$  - уровни фонового сигнала в канале акцептора и донора, соответственно;  $\alpha_{DA}$  и  $\alpha_{AD}$  - перекрестные коэффициенты (crosstalk coefficients), характеризующие долю флуоресценции донора, регистрируемую

в канале акцептора ( $\alpha_{DA}$ ), и долю излучения акцептора, регистрируемую в канале донора ( $\alpha_{AD}$ ). Поправочные коэффициенты помогают учитывать особенности оптической схемы конфокального микроскопа или спектрофлуориметра, для разных пар флуоресцентных меток они отличаются.  $D_{ex}$  - это величина сигнала прямого возбуждения акцептора (возникающая в следствие возможного поглощения фотонов акцептором на длине волны поглощения донора).  $\gamma$  - так называемый фактор детекции, он позволяет учитывать различия в эффективностях детекции флуоресценции ( $\eta$ ) донора и акцептора и различия в их квантовых выходах флуоресценции ( $\Phi$ ) [134]. Фактор детекции может быть представлен в виде:

$$\gamma = \gamma_{instrument} \gamma_{dye} = \frac{\eta_A \Phi_A}{\eta_D \Phi_D} \quad (1.38)$$

где  $\Phi$  - квантовый выход,  $\eta$  - эффективность детекции. Эффективность детекции флуоресценции зависит от спектральной чувствительности детекторов (лавинных фотодиодов или фотоэлектронных умножителей) и спектральных характеристик оптических элементов, через которые проходит сигнал от образца. Измерение фактора детекции в ряде случаев может быть затруднено и его принимают за единицу. В таком случае измеряемая по формуле (1.37) величина является не истинным коэффициентом переноса, такую величину называют “коэффициентом близости” (PR - proximity ratio). Коэффициент близости нельзя непосредственно связать с расстоянием между флуоресцентными метками, однако, используя этот коэффициент можно делать выводы о гетерогенности образца, и о качественных эффектах сближения или удаления частиц.

### Детекция сигнала от одиночных молекул

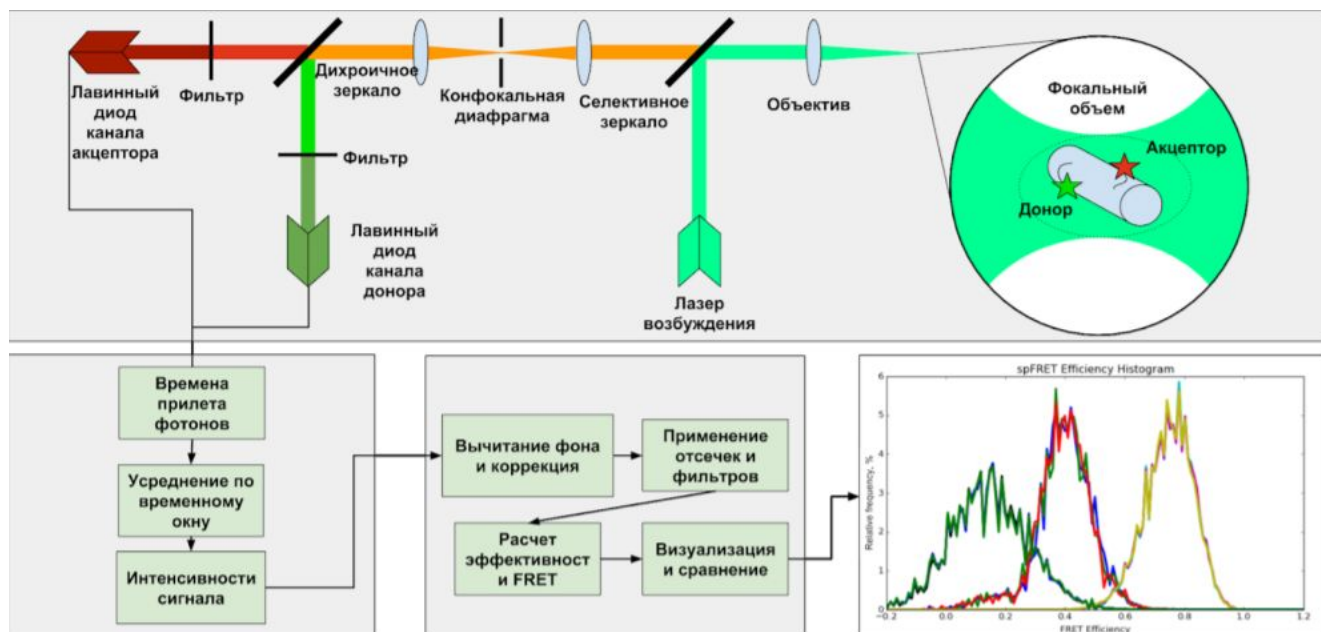
Эффективность FRET может быть измерена в объеме сразу от многих частиц при помощи прибора спектрофлуориметра. В этом случае измеренный сигнал соответствует некоторому среднему сигналу от ансамбля молекул [135]. Такой усредненный сигнал может не соответствовать ни одному конкретному состоянию системы на молекулярном уровне, а являться суперпозицией нескольких состояний разных частиц. Чтобы добиться связи сигнала FRET с

конкретными состояниями молекулярной системы, необходимо, как минимум, измерять эффективности FRET у одиночных частиц, такой метод называется spFRET (single particle FRET). С помощью метода spFRET возможно оценивать как средние расстояния [136], так и временную динамику изменения расстояний между метками [137]. Вопрос заключается в том, каким образом снимать сигнал от одной частицы. Один вариант заключается в иммобилизации частиц на поверхности. Иммобилизовав частицы, можно использовать подходы, основанные на эффекте полного внутреннего отражения, которые позволяют проводить измерения в тонком слое частиц у поверхности [138]. Преимущество такого метода в возможности наблюдения эволюции сигнала сразу для многих частиц в поле зрения микроскопа. Недостатком является возможные взаимодействия частиц с подложкой, которые могут повлиять на их структуру и динамику [139]. Альтернативный подход основан на измерении флуоресценции от единичных молекул, свободно диффундирующих сквозь фокальный объем конфокального флуоресцентного микроскопа (см. Рис. 1.16). Для реализации этого метода необходимо, чтобы исследуемые частицы находились в достаточно низкой концентрации, чтобы одновременно в фокальном объеме находилось не более одной частицы. Метод требует наличия конфокального микроскопа, чувствительных лавинных фотодиодов (APD или SAPD). По мере проплывания частиц можно собирать необходимую статистику, однако поскольку проплывание обычно занимает миллисекунды, присутствуют ограничения на возможности слежения за динамикой процессов внутри частицы.

## Подходы к измерению эффективности FRET

Величина  $E_{FRET}$ , как следует из вышеприведенных формул, может быть рассчитана двумя способами, либо через время флуоресценции донора [140], либо через измерение интенсивностей свечения донора и акцептора. Последний подход называется ратиометрическим подходом.

Для того, чтобы измерить времена флуоресценции необходимо существенно более сложное оборудование. Порядок времен флуоресценции составляет порядка наносекунд, поэтому требуются лазерные источники и синхронизированные детекторы, способные работать в пикосекундном диапазоне. Такой



**Рисунок 1.16** — Схема экспериментов spFRET. Один лазер возбуждает флуоресценцию. Флуоресценция из конфокального объема собирается на лавинных фотодиодах. Времена прилета фотонов анализируются программным обеспечением микроскопа, либо специализированным ПО.

подход реализован в методе TCSPC (Time Correlated Single Photon Counting). Он же используется в популярном методе микроскопии супер разрешения FLIM (fluorescence-lifetime imaging microscopy). Важным преимуществом данного метода является отсутствие необходимости учитывать эффекты, связанные с мощностью осветителя, различными параметрами оптической системы и системы детекции, влияющими на корректировочные факторы. Наличие временного разрешения пикосекундного диапазона, также позволяет эффективно проводить отсеивание шумовых всплесков в сигнале.

Ратиометрический подход может быть реализован с использованием более простого оборудования, например, конфокального микроскопа с лазером непрерывного излучения. Однако, этот подход чувствителен к шумовым эффектам, связанным с рассеянием света, колебаниям интенсивности лазера. Для соотношения измеренных интенсивностей с реальными фотофизическими процессами необходимо предварительно измерить фактор детекции  $\gamma$ . Отдельной проблемой является то, что данный фактор может зависеть от конформационных изменений в молекулярной системе, так как квантовые выходы чувствительны к окружению флуорофоров.

## Подходы к измерению интенсивности флуоресценции spFRET от одиночных молекул

Интенсивности флуоресценции в радиометрическом подходе, в зависимости от приборных возможностей, регистрируют либо в виде временного ряда, описывающего прилет отдельных фотонов (single photon counting), либо в виде некоторого сигнала, описывающего зависимость интенсивности от времени. В последнем случае прибор будет проводить усреднение (бинирование) сигнала с некоторым временным интервалом. В случае измерения spFRET в растворе, данный временной интервал необходимо выбирать с учетом кинетики проплывания частиц через фокальный объем микроскопа. Для оценки времени проплывания частицы через фокальный объем можно использовать диффузионное соотношение:

$$\tau \approx \langle x^2 \rangle / 2D \quad (1.39)$$

где  $\tau$  - время проплывания,  $D$  - коэффициент диффузии,  $x$  - координата частицы вдоль одной из осей системы координат.

Для обработки непосредственно времен прилета фотонов также разработаны различные алгоритмы, позволяющие выделять внутри временного ряда события, связанные с проплыванием частиц и оценивать фоновый сигнал [141]. Оценку фонового сигнала можно проводить непосредственно анализируя правый край гистограммы распределения временных задержек между прилетом фотонов. Поскольку вероятность прилета фотонов за некоторый интервал времени описывается распределением Пуассона

$$P[N(t) = k](t) = \frac{e^{-\lambda t} (\lambda t)^k}{k!} \quad (1.40)$$

можно показать, что задержки распределены согласно экспоненциальному распределению:

$$P(t) = \lambda e^{-\lambda t} \quad (1.41)$$

где  $\lambda$  - средняя частота детекции фотонов. На практике распределение будет состоять из суммы распределения фонового сигнала и сигнала от проплывающих частиц.



## Измерение фактора детекции

При ратиометрическом измерении FRET вычисление истинного  $E_{FRET}$  требует определения корректирующего фактора  $\gamma$ . Данный фактор позволяет перемасштабировать отношение регистрируемых интенсивностей сигнала в каналах донора и акцептора, чтобы учесть разницу в квантовых выходах красок и эффективности детекции фотонов от разных красок прибором. Также фактор детекции может изменяться целенаправленно варьируя параметры прибора для достижения большего разрешения в областях расстояния между метками большими или меньшими радиуса Ферстера [142].

Для измерения параметра  $\gamma$  можно использовать ряд экспериментальных методик. Квантовые выходы измеряют их сравнения с квантовыми выходами известных красителей [143]. Инструментальную часть фактора,  $\gamma_{instrument}$ , можно вычислить используя формулу:

$$\gamma_{instrument} = \frac{m_A^{smF} m_D^{Abs} \Phi_D}{m_D^{smF} m_A^{Abs} \Phi_A} \quad (1.42)$$

где  $m_A^{smF}$  и  $m_D^{smF}$  - производные интенсивности флуоресценции акцептора и донора по концентрации для заданного лазера,  $m_A^{Abs}$  и  $m_D^{Abs}$  - производные поглощения акцептора и донора от концентрации, измеренные на длине волны возбуждения донора,  $\Phi_A$  и  $\Phi_D$  квантовые выходы акцептора и донора. Более подробно методика описана в работе [144].

Альтернативой измерению фактора детекции в отдельных экспериментах (в том числе с применением спектрофотометра), является его измерение непосредственно во время эксперимента. Например, метод ALEX FRET (alternating laser excitation FRET) использует чередующиеся вспышки двух лазеров, возбуждающих поочередно донор или акцептор. Такой метод позволяет сразу определить стехиометрию красок в проплывающем комплексе. Измеряя эффективность переноса для комплексов с различной стехиометрией можно рассчитать фактор детекции [145].

### 1.6.3 Методы геномного анализа в исследовании структуры хроматина

Для определения наличия физического контакта между последовательностями ДНК в трехмерной структуре генома были разработаны методы семейства -C: 3C, 4C, 5C, Hi-C [146]. Эти методы основаны на химическом “сшивании” контактирующих в хроматине участков, фрагментации ДНК с образованием “сшитых” пар, лигирования фрагментов в этих парах с образованием гибридных молекул ДНК, детекции гибридных молекул (обнаружение молекулы, состоящей из последовательности А и последовательности В, говорит о физическом контакте локусов этих последовательностей в ДНК). Существует понятие “разрешения” метода Hi-C (см. Таблицу 1). Его можно определить как минимальное расстояние (в парах оснований) между участками, на котором они отображаются на тепловую карту в разные ее точки (пиксели). Величина разрешения зависит от величины фрагментов, на которые рестриктазы разрезают ДНК (то есть от особенностей используемых рестриктаз, размеров и встречаемости в геноме их сайтов рестрикции), и от размера бина, выбираемого на стадии обработки данных Hi-C, но не полностью произвольного: для получения тепловой карты, адекватно представляющей трехмерную структуру изучаемого генома, необходимо выбирать размеры бинов из определенного диапазона, зависящего от размеров хромосом, глубины секвенирования и качества NGS-библиотеки [147]. От разрешения метода Hi-C зависит то, какие структуры в хроматине он может выявить. Оригинальный протокол метода, описанный в статье 2009 года [148], позволял достичь разрешения 4kb и выявлял, соответственно, в трехмерной структуре генома компартменты А и В (соответствующие активному и неактивному хроматину). В протоколах Hi-C, разработанных позже, достигалось, благодаря использованию других рестриктаз, большее разрешение - до одной килобазы. Это позволяло обнаружить так называемые топологически ассоциированные домены (ТАДы) [149] и взаимодействия конкретных последовательностей друг с другом - то есть, физические контакты регуляторных элементов с регулируемыми ими генами, реализованные, как полагают современные исследователи, запетливанием ДНК. Однако этих разрешающих способностей все равно не достаточно, чтобы изучить су-

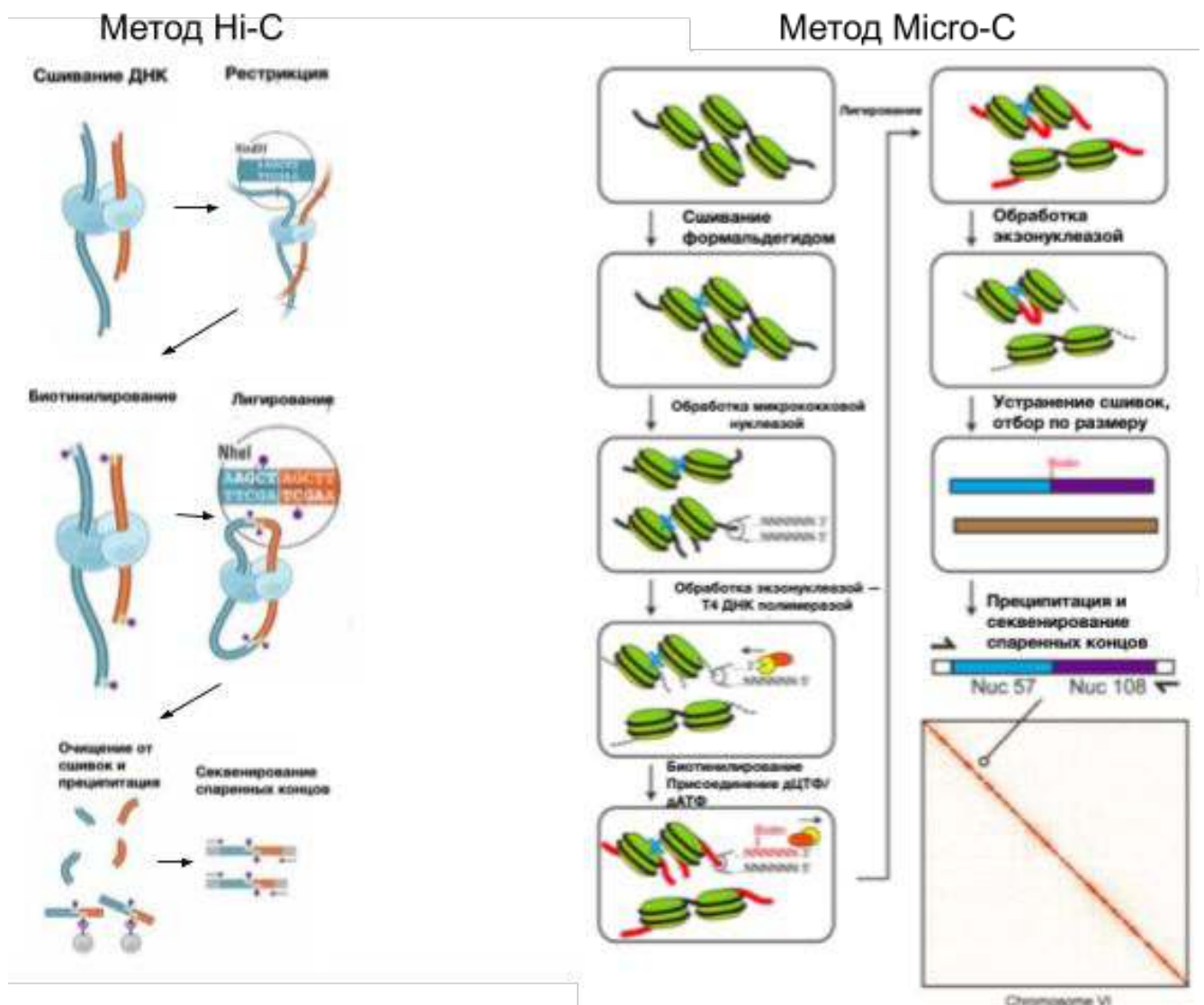
прануклеосомный уровень, на котором ранее предполагалось существование 30-нанометровой фибриллы [147].

Пролить свет на этот уровень смог метод Micro-C — модификация метода Hi-C, предполагающая использование на этапе фрагментации ДНК микрококковой нуклеазы, которая уничтожает линкерные участки, оставляя одни нуклеосомы (этот метод можно рассматривать как синтез метода позиционирования нуклеосом с помощью микрококковой нуклеазы - MNase-seq [150] - и Hi-C [151]). Это позволяет оценивать частоту взаимодействия отдельных нуклеосом друг с другом, то есть разрешение метода достигает неизученного супрануклеосомного уровня (между 200bp и 4kb), устраняя так называемый “пробел в разрешении” (resolution gap). Интересно, что метод Micro-C тоже не обнаружил 30-нанометровую фибриллу, еще раз доказав необходимость дальнейших исследований структуры хроматина супрануклеосомного уровня. В работе, в которой впервые был предложен метод Micro-C, изучался геном дрожжей - *S. cerevisiae*.

Параллельно Micro-C был разработан другой метод Hi-C высокого разрешения - in situ DNase Hi-C. С его помощью можно получить данные с разрешением около 1kb [152]. Однако этот метод, основанный на использовании фермента DNase I, неспецифически фрагментирующего ДНК на тетра-нуклеотиды, не является конкурентным для Micro-C в области изучения супрануклеосомной структуры хроматина, так как не уничтожает линкерную ДНК и, соответственно, не позволяет в отличие от Micro-C простым способом получить данные о частоте взаимодействия нуклеосом друг с другом. Сам протокол Micro-C модифицировался после его изобретения. Примером могут служить работы [153] и [130], в которых были впервые представлены модификации Micro-C: Micro-C XL и Hi-Co, соответственно. Micro-C XL - метод, позволяющий получить качественные данные в моонуклеосомном разрешении, это достигается путем снижения шума за счет изоляции нерастворимого хроматина и использования помимо формальдегида длинных сшивающих агентов (кросслинкеров) - например, DSG и EGS. Метод Hi-Co специфичен тем, что с его помощью можно определить ориентацию нуклеосом (“O” в названии - от Orientation). Это достигается с помощью особой биоинформатической обработки и использования не просто секвенирования спаренных концов, а секвенирования спаренных концов с тэгами (PET - paired-end tags [154]).

Следует отметить, что все методы -C являются статистическими и вероятностными, приблизительно оценивающими частоту взаимодействия большого

количества фрагментов ДНК. Для получения биологически значимой информации с помощью метода Hi-C вообще и, в частности Micro-C, необходим тщательный биоинформатический анализ матриц взаимодействия: фильтрация и нормализация данных, учет того — использовалась в методе клеточная популяция или единичная клетка (Single-Cell Hi-C/Micro-C [155]). Для изучения супрануклеосомной структуры хроматина с помощью Micro-C необходимо дополнительно прибегнуть к его молекулярному моделированию с учетом последних достижений в вычислительной биофизики.



**Рисунок 1.17** — Методы “захвата конформации хроматина” (chromatin conformation capture), сравнение Hi-C и Micro-C. Адаптировано из [148] и [151].

Появление методов секвенирования нового поколения дало огромное преимущество большинству методов эпигеномики, позволяющим получать данные

о местах расположения нуклеосом, в том числе с гистоновыми вариантами и с определенными пост-трансляционными метками гистонов.

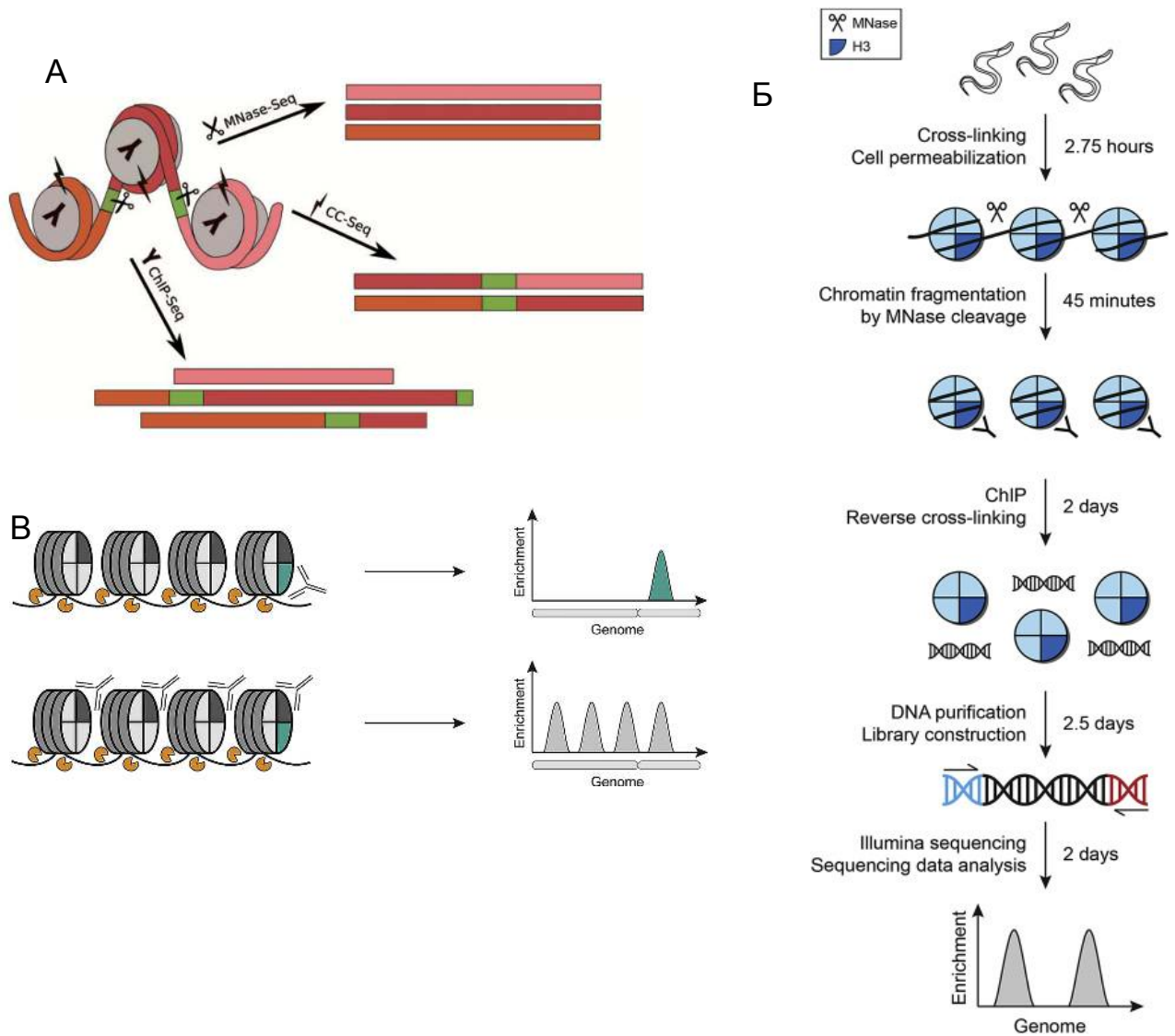
Метод MNase-seq позволяет получить данные о расположении единичных нуклеосом. В методе используется микрококковая нуклеаза, которая обладает экзо- и эндонуклеазными активностями. После работы нуклеазы, нуклеосомальная ДНК освобождается от гистонов и секвенируется. Далее прочтения картируются на референсный геном. На данный момент в публичном доступе накоплено большое количество экспериментальных наборов данных для клеток разных организмов и из разных тканей человека (110 наборов данных по данным сайта <https://generegulation.org/nucleosome-positioning-database/>). К основным проблемам использования методов Mnase-seq относится огромное количество получаемых с одного эксперимента прочтений - порядка 200-400 млн для клеточных линий человека. Для достижения оптимального разрешения позиционирования нуклеосом в человеческих клетках для сравнения здоровых и больных требуется порядка 1-4 млрд прочтений [156]. Обычно эксперименты MNase-seq выполняют на нескольких тысячах клеток, такой усредненный нуклеосомальный профиль характеризует ансамбль клеток, а не частные состояния клеток. Специализированная база данных NucMap включает 798 экспериментальных наборов данных MNase-seq из 477 образцов для 15 видов живых организмов [157]. Для дрожжей существует метод SS-seq, позволяющий определять позиционирование нуклеосом с большей точностью, чем MNase-seq [158]. Суть метода в ведении специальной мутации в гистоны, в результате которой вблизи центра нуклеосомы появляется остаток цистеина. Боковая цепь цистеина используется для реакции со специальными химическими агентами, которые разрезают ДНК в центре нуклеосомы. Далее методами секвенирования определяется положение разреза и, следовательно, центра нуклеосомы. Методы ChIP-seq осуществляют анализ ДНК-белковых взаимодействий. В открытом доступе находится большое количество наборов данных ChIP-seq гистоновых модификаций, включая метки энхансеров и промотеров, например, ChIPBase v.2.0 2466 датасетов, IHEC Data Portal - несколько тысяч образцов из разных человеческих органов, ChIP-Atlas - данные из 96000 экспериментов. Также ChIP-Atlas позволяет ответить на следующие вопросы: какие белки были связаны с определенной последовательностью ДНК, какие гены регулируются данным белком, какие белки локализованы с данным, а также позволяет предсказывать белки, связанные с данными геном-

ными локусами и генами (*in silico* ChIP). Важная модификация метода ChIP-seq - метод Mnase ChIP-seq заключается в отщеплении свободных п.н. (линкерной ДНК) микрококковой нуклеазой до этапа иммунопреципитации. В результате образуются фрагменты 20-70 п.н., позволяющие строить нуклеосомальные профили и выявлять сайты связывания транскрипционных факторов высокого разрешения. Было отмечено, что MNase-ChIP-seq также позволяет увеличить воспроизводимость результатов по сравнению с ультразвуковой фрагментацией в методе ChIP-seq [159]. В работе [160] с помощью методов ChIP-exo и MNase-seq было показано наличие субнуклеосомальных структур (гексасомы, нуклеосомы с единично представленными гистонами типов H2A, H2B, H3, H4) *in vivo* и сделаны предположения о влиянии последовательности ДНК на образование таких структур.

В работе [161] проведено сравнение обработки сырых данных экспериментов: MNase-seq, ChIP-seq и метода химического расщепления ДНК в месте диадной оси, используя разные стратегии секвенирования (одноконцевое и парноконцевое), с помощью разработанной программы NUCwave. Было показано, что обработка данных парноконцевого секвенирования MNase-seq наиболее точно определяет известные позиции нуклеосом и превосходит другие сырые данные для построения карт расположения нуклеосом.

## 1.7 Выводы главы 1

Исследование структуры многих биомакромолекулярных комплексов испытывает затруднения как при применении стандартных методов структурной биологии, так и при попытках атомистического моделирования на основе расчета атом-атомных взаимодействий. Методы интегративного моделирования, в ходе расчетов учитывающие как физические взаимодействия между атомами, так и экспериментальные данные различной природы, являются перспективным подходом для построения структурно-динамических моделей биомакромолекулярных комплексов. В главе предложены подходы по моделированию ДНК-белковых комплексов на основе моделей анизотропной гибкости ДНК с учетом разнородных экспериментальных данных.



**Рисунок 1.18** — Методы эпигеномики нуклеосомного разрешения: А) MNase-seq, ChIP-seq, CC-seq. Б-Б) MNase-ChIP-seq. Адаптировано из [161] и [159].

**Таблица 1** — Увеличение разрешения Hi-C подобных методов, благодаря использованию менее специфичных рестриктаз.

Протокол Hi-C	Рестриктазы	Сайты рестрикции	Размер сайта рестрикции
Классический (Lieberman-Aiden et al., 2009)	HindIII, NcoI	AAGCTT, CCATGG	6 пн
Sexton et al., 2012, Rao et al., 2014	Dpn-II	GATC	4 пн
COLA (Darrow et al., 2016)	CviJI	RCGY, R=A/G, Y=C/G	3 пн
Micro-C (Hsieh et al., 2016)	MNase	Отсутствует. Эндо/экзонуклеаза, почти полностью уничтожает линкерные участки ДНК	-
in situ DNase Hi-C (Ramani et al., 2016)	DNase I	Отсутствует. Режет ДНК на тетра-нуклеотиды	-



## Глава 2. Применение методов молекулярной динамики для изучения нуклеосом<sup>1</sup>

Данная глава иллюстрирует современные возможности суперкомпьютерного моделирования методом атомистической молекулярной динамики для получения и анализа структурно-динамических моделей биомакромолекулярных комплексов на основе информации об их статических структурах, получаемых методами структурной биологии. Глава посвящена расчету и анализу молекулярно-динамических моделей ключевых нуклеопротеиновых комплексов хроматина эукариот - нуклеосом. Во введении дан широкий обзор литературы по динамике нуклеосом и их роли в функционировании хроматина. В сутевой части главы приводятся результаты по рекордно долгим вычислениям траекторий молекулярной динамики в микросекундном диапазоне (в том числе свыше 10 микросекунд). Предложены оригинальные алгоритмы анализа динамического поведения нуклеосом, на основе которых проанализирован ряд физиологически важных динамических мод нуклеосом. Результаты приведенные в данной главе основаны на статьях [A1—A3; A8; A9; A12; A14; A17; A18; A20; A21; A25], тезисах [A38; A40; A41; A44—A47], а также включают новые результаты.

### 2.1 Введение в структуру и динамику нуклеосом

*Данный раздел написан по материалам собственной обзорной статьи [A3].*

Жизнь эукариотических клеток полностью управляется пространственно-временной организацией хроматина внутри ядер клеток. Хроматин не только компактизует ДНК, но и служит средой для расшифровки и интерпретации генетической информации [162]. Основной структурной единицей организации

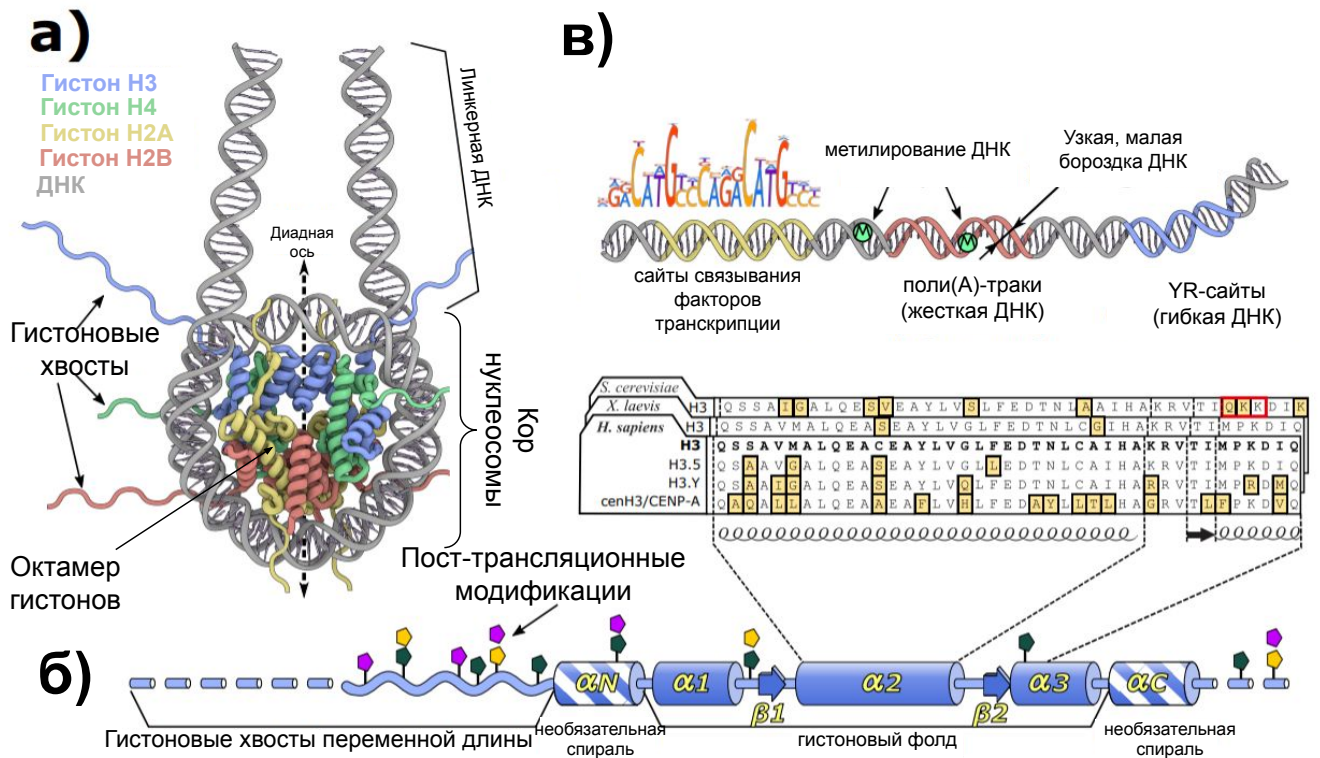
---

<sup>1</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [A1—A3; A8; A9; A12; A14; A17; A18; A20; A21; A25].

хроматина является нуклеосома - повторяющаяся единица из примерно 200 пар оснований ДНК, организованная гистоновыми белками [163; 164] (Рис. 2.1а). Центральные 145–147 пар оснований ДНК образуют коровую частицу нуклеосомы, или нуклеосомный кор (NCP, nucleosome core particle), плотно наматываясь на октамер гистонов в виде  $\sim 1,7$  витка левой суперспирали [165; 166]. Линкерные сегменты ДНК фланкируют NCP. Канонический октамер гистонов состоит из двух пар идентичных гистоновых димеров H3-H4 и H2A-H2B, образованных коровыми гистонами H3, H4, H2A и H2B, соответственно (Рис. 2.1а) [A14]. Симметрия октамера наделяет NCP осью псевдосимметрии второго порядка, которая определяет расположение центра ДНК - диады нуклеосом. Все четыре типа гистонов, вероятно, имеют общее эволюционное происхождение и имеют сходную структуру [167; 168] (Рис. 2.1б). Гистоны также подразделяются на основные “канонические” гистоны (депонируются на ДНК при репликации) и альтернативные варианты гистонов, которые синтезируются в течение клеточного цикла и могут быть тканеспецифичными [A14]. Дополнительные уровни вариабельности нуклеосом обеспечиваются репертуаром посттрансляционных модификаций гистонов (ПТМ) [169] и вариаций внутри последовательностей ДНК (Рис. 2.1в).

С годами стало ясно, что нуклеосомы претерпевают множество функционально важных структурных и динамических перестроек во время всех ключевых процессов происходящих в хроматине (транскрипция, репликация, репарация ДНК и т.д.) [170; 171]. Более того, способность нуклеосом претерпевать определенные типы конформационных переходов разумно используется многими белками хроматина, включая ремоделирующие хроматин [172–174], факторы транскрипции [175; 176], шапероны [A19], РНК-полимеразы [A23] и т.д. С момента появления первых эукариот около 2 миллиардов лет назад биологические процессы ядра приспособлялись к структуре нуклеосомы, используя тонкие детали ее динамики. Одновременно с диверсификацией репертуара нуклеосом с помощью вариантов гистонов и их ПТМ белки хроматина также диверсифицировались и эволюционировали, чтобы различать разные нуклеосомы по их структурным и динамическим характеристикам. Это тонкое взаимодействие между составом нуклеосом, структурой, функциональной динамикой и взаимодействиями с белками хроматина формирует основы многих регуляторных путей, происходящих в хроматине.

Ниже мы очертим концептуальную схему для понимания различных режимов структурной динамики нуклеосом и классифицируем основные факторы, которые влияют на эти режимы.

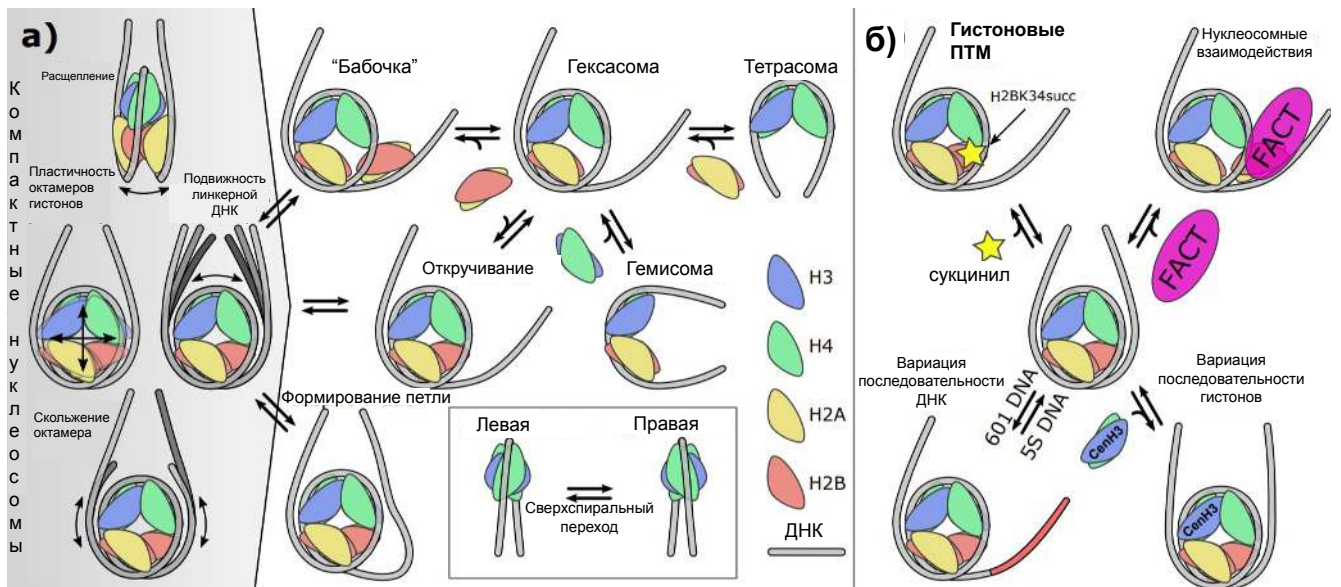


**Рисунок 2.1** — (а) Структура нуклеосомы состоит из коровой частицы нуклеосомы (NCP), фланкированной линкерными сегментами ДНК. Коровая частица образована октамером гистонов, вокруг которого обернуты 145–147 пар оснований ДНК в виде левой суперспирали. Октамер состоит из двух пар гетеродимеров гистонов (H3-H4 и H2A-H2B) и обладает осью псевдосимметрии второго порядка (диадной осью), проходящей через центральную пару оснований нуклеосомной ДНК (диаду). Неструктурированные гистоновые хвосты отходят от глобулярной части белков. Малая бороздка ДНК обращена к октамеру гистонов в 14 специально предназначенных для этого сайтах связывания (в каждом сайте связывания боковая цепь остатка аргинина вставлена в малую бороздку). (б) Схема типичной структуры гистонического белка: гистоновый фолд (спирали  $\alpha 1$ ,  $\alpha 2$  и  $\alpha 3$ , листы  $\beta 1$  и  $\beta 2$ ) с необязательными спиралями  $\alpha N$ ,  $\alpha C$  и гистоновыми хвостами переменной длины. Гистоновые ПТМ могут располагаться по всей длине белка и играть важную роль в функционировании хроматина, обеспечивая эпигенетическую разметку. Вариации гистонического сиквенса существуют как между организмами, так и внутри организма (гистоновые варианты). (в) ДНК - ключевой компонент нуклеосомы. Конкретные последовательности ДНК могут обеспечивать присутствие сайтов связывания факторов транскрипции в областях нуклеосомной ДНК, повышенную жесткость и/или изменения геометрии ДНК.

### 2.1.1 Обзор динамических мод нуклеосом

Внимание на динамическую природу нуклеосом впервые было обращено в контексте отворачивания ДНК от октамера гистонов [177] (Рис. 2.2а, в центре). Результаты экспериментов по ферментативному расщеплению ДНК [177], измерению FRET [178], крио-ЭМ [179] свидетельствуют о возможности откручивания ДНК (иногда называемом раскручиванием или дыханием ДНК). Данный процесс важен для доступа факторов транскрипции [176; 178] и РНК-полимераз [180] к последовательности ДНК. Позднее выяснилось, что откручивание ДНК может происходить вместе с раскрытием димера Н2А-Н2В (состояние “бабочка”) [181] или его полной диссоциацией (с образованием гексасомы), что часто происходит при транскрипции [182]. Частично собранные нуклеосомы (также называемые субнуклеосомными структурами), такие как тетрасома или гемисома (полунуклеосома) также эффективно способствуют отворачиванию ДНК [183]. Недавно с использованием метода ChIP-ехо было показано, что такие структуры широко распространены в динамическом хроматине в масштабе всего генома [160]. В то время как тетрасомы могут возникнуть из-за потери двух димеров Н2А-Н2В нуклеосомой, динамические пути, ведущие к гемисомам, не совсем ясны. Новые экспериментальные данные подтверждают идею о том, что сборка нуклеосом *in vivo* во время репликации ДНК происходит через депонирование тетрамеров Н3-Н4 (два димера Н3-Н4, взаимодействующие через четырехспиральную связку) с помощью фактора сборки хроматина 1 (CAF1) [184; 185]. После этого к ним присоединяются димеры Н2А-Н2В. Следовательно, маловероятно, что гемисомы образуются во время регулярной сборки нуклеосом. Расщепление нуклеосом на полунуклеосомы во время транскрипции - еще одна гипотеза (рассмотренная в [170]). Хотя есть некоторые споры, частным случаем статической структуры полунуклеосом могут быть центромерные нуклеосомы *S. cerevisiae* [186], которые также могут быть воссозданы *in vitro* в особых условиях [187]. В то время как нуклеосома в отсутствие внешнего сверхспирального стресса содержит ДНК в виде левосторонней суперспирали [188], тетрасомы или гемисомы демонстрируют повышенную тенденцию к адаптации правостороннего состояния. Вероятно, так обстоит дело с центромерными нуклеосомами пекарских дрожжей [*ibid.*] и тетрамерами гистонов архей [189].

В левой части рисунка 2.2 мы сгруппировали динамические режимы с меньшей величиной конформационных изменений (“компактная нуклеосома”), которые, тем не менее, функционально важны. Октамер гистона способен к “расщеплению” - изменению расстояния между двумя половинками нуклеосомы вдоль оси суперспирали ДНК. Экспериментально это наблюдалось на нуклеосомах [190; 191] и может быть связано со скольжением витков ДНК относительно друг друга [192]. Недавно наше понимание динамики октамера было дополнительно расширено за счет концепции пластичности октамеров на уровне отдельных димеров гистонов - точечные сшивки внутри димера Н3-Н4 ограничивает его деформируемость и ингибируют передвижение нуклеосом ремоделером SNF2h [172]. Наконец, ДНК в нуклеосоме - это компонент, который проявляет конформационную изменчивость. Октамер может скользить по ДНК, изменяя ротационное и трансляционное позиционирование ДНК [A4; A7]. Этот процесс сильно зависит от последовательности ДНК (обзор см. в [193]) и включает локальные деформации ДНК, такие как дефекты скручивания [194]. В линкерной области ДНК более гибкая, чем остальная часть ДНК [142]. Благодаря такой высокой мобильности, комплексы нуклеосом с линкерным гистонам могут образовывать ансамбль различных конфигурации (см. обзор в [195]).



**Рисунок 2.2** — (а) Обзор различных мод динамики нуклеосом. Слева (серая область) показаны режимы, сохраняющие компактность нуклеосомы. Справа показаны моды с большей амплитудой: нуклеосомная ДНК может спонтанно образовывать петли или отворачиваться от гистонов; большие состояния разворачивания ДНК способствуют образованию субнуклеосомных частиц: гексасом, тетрасом, гемисом. (б) Примеры факторов, влияющих на структурную динамику нуклеосом. Слева сверху: сукцинирование H2B по лизину 34 способствует разворачиванию ДНК [196]. Справа сверху: ассоциация с фактором транскрипции FACT изменяет динамику разворачивания нуклеосом [A19]. Слева внизу: разные последовательности ДНК имеют разную динамику разворачивания [197]. Справа внизу: вариант гистона H3 человека Cenp-A изменяет подвижность линкерной ДНК [198].

### 2.1.2 Факторы, влияющие на динамику нуклеосом: недавние примеры

Факторы, влияющие на динамику нуклеосом можно разделить на четыре группы (Рис. 2.2б): вариации последовательностей гистонов, вариации последовательности ДНК, ПТМ гистонов и различные белки, взаимодействующие с нуклеосомами.

Белки, взаимодействующие с нуклеосомами, включают в себя структурные белки хроматина (включая линкер гистона H1 [A15]), шапероны гистонов,

факторы транскрипции (включая пионерские факторы транскрипции [176]), компоненты репликационного, транскрипционного и эпигенетического аппарата. Они могут влиять на любой из обсуждаемых режимов динамики нуклеосом. Например, гистоновый шаперон FАСТ обратимо разворачивает нуклеосомную ДНК с обоих концов [A19], чтобы облегчить обмен димера гистона H2A-H2B [199]. Гистоновый шаперон Nap-1 является другим примером, он вытесняет димер H2A-H2B из нуклеосомы, и может действовать без необходимости изначального отворачивания ДНК [200]. Показано, что на раскрытие центромерных нуклеосом влияет взаимодействие с белком CENP-C [191]. Напротив, ПАРП-1 (многодоменный белок, который управляет обнаружением одно- и двухцепочечных разрывов во время репарации ДНК) значительно увеличивает расстояние между супервитками нуклеосомной ДНК в обратимой манере [201]. Линкерные участки ДНК могут притягиваться друг к другу гистоном H1, который снижает их гибкость [202].

Неструктурированные гистоновые хвосты исторически были основными целями для исследования эффектов различных РТМ гистонов на динамику нуклеосом [169] (Рис. 2.1б). По умолчанию считается, что модификации, изменяющие заряд (например, ацетилирование лизина, фосфорилирование серина или отщепление гистонового хвоста), вызывают дестабилизацию нуклеосом за счет снижения не-специфического электростатического притяжения между ДНК и гистонами [203]. Однако существуют свидетельства в пользу того, что более тонкие эффекты также могут иметь значение, такие как изменения вторичной структуры гистоновых хвостов (напр., для H4K16ac [204]). РТМ, расположенные на сайтах связывания гистонов с ДНК, могут оказывать более специфические эффекты, нарушая связывание с ДНК соответствующих мест. Например, недавно проанализированное сукцинирование H2BK34, расположенного в сайте связывания ДНК гистонами на расстоянии  $\sim 25$  п.н. от точки входа/выхода в нуклеосоме, ингибирует сборку нуклеосом и способствует разворачиванию ДНК [196]. Важность РТМ гистонов, локализованных в глобулярной части гистонов, и их функциональное значение в настоящее время также широко признаны. Например, H3K56ac облегчает разворачивание нуклеосом и играет важную роль в регуляции сборки нуклеосом [205]. Не изменяющие заряд РТМ, такие как H3R42me2, также могут способствовать дестабилизации нуклеосом [206]; последний также использовался микобактериями для изменения эпигенетического ответа хозяина [207]. РТМ на границах



раздела гистон-гистон могут как нарушать образование октамера [208], так и стабилизировать нуклеосомы (например, H4K77ac из-за устранения электростатического отталкивания внутри димера гистона [203]). Сходным образом, H1K85ac ведет к общей стабилизации хроматина и линкерной ДНК за счет увеличения связывания глобулярного домена H1 с коровыми гистонами [209].

Нет сомнений в том, что вариация последовательности ДНК может оказывать значительное функциональное воздействие, изменяя вероятность разворачивания ДНК и стабильность нуклеосом. Например, сильно взаимодействующая последовательность “601”, если ее поместить на нуклеосому, образует полярный барьер для транскрипции [180]. Это соответствует известному асимметричному разворачиванию последовательности “601” под напряжением [210] и в состоянии покоя. Последнее было недавно подтверждено малоугловым рассеянием рентгеновских лучей по сравнению с симметричным разворачиванием, наблюдаемым для 5S рибосомной последовательности [197]. Динамические эффекты также играют роль во время связывания факторов транскрипции, как первоначально было предложено J. Widom и соавторами [177]. Например, недавняя крио-ЭМ структура нуклеосомы, включающая энхансерную последовательность ALB1 (ALB1 является сайтом связывания для пионерского фактора транскрипции FoxA), аналогична структуре 601-нуклеосомы, но ДНК в области связывания FoxA является более подвижной, способствующей связыванию FoxA [211]. Несмотря на многочисленные усилия, точная взаимосвязь между последовательностью ДНК и ее влиянием на динамику нуклеосом все еще не ясна. По данным некоторых работ присутствие гибких пиримидин-пуриновых динуклеотидов (YR) в сайтах связывания ДНК является важным фактором для позиционирования ДНК и силы связывания (Рис. 2.1в) [212; 213]. Однако развитию более общих теоретических моделей в настоящее время препятствует отсутствие достаточно подробных и воспроизводимых экспериментальных наборов данных о связи последовательности ДНК с ее стабильностью на нуклеосоме [193]. Атомистическое и крупнозернистое компьютерное моделирование может восполнить этот пробел: как недавно показали несколько исследований с использованием моделирования, в зависимости от физических свойств последовательности ДНК, ее транслокация может предпочтительно происходить через винтовой механизм “распространения дефектов кручения” или механизм “повторного захвата петли” [214; 215].

Варианты гистонов обеспечивают богатый репертуар вариаций канонической гистоновой последовательности, варьирующийся от всего лишь нескольких аминокислотных различий (например, H3 vs H3.5) до довольно существенных различий, которые могут изменять несколько динамических режимов одновременно (Рис. 2.1б). Напр., CENP-A (вариант, критический для образования центромер во время митоза) имеет более короткую спираль  $\alpha N$ , чем канонический H3. Эта особенность делает линкерную ДНК в CENP-A нуклеосомах более гибкой и влияет на связывание ДНК с гистонами [198]. Кроме того, CENP-A влияет на разворачивание ДНК и способствует образованию петель [216]. Известно также, что нуклеосомы CENP-A демонстрируют расщепление половинок нуклеосомы вдоль оси суперспирали [191]. Варианты с небольшими изменениями канонической гистоновой последовательности также могут быть функционально очень важными: H3.3 важен для пластичности нейронов [217], H3.5 для сперматогенеза [218], а H3.Y изменяет регуляцию клеточного цикла [219]. В последние годы было показано, что эти небольшие вариации последовательности могут влиять на стабильность и динамику нуклеосом. Например, H3.5-специфический остаток лейцина 103 дестабилизирует нуклеосому, что важно для процесса транскрипции в тестикулярных клетках человека [218]. Как показано в [220], в H3.Y метионин 124 способствует стабильной ассоциации тетрамера H3.Y-H4 с ДНК. Другой H3.Y-специфический остаток – лизин 42 – играет роль в придании гибкости линкерной ДНК.

Более того, оказывается, что небольшие вариации между последовательностями канонических гистонов, кодируемых разными копиями генов канонических гистонов у человека, также могут функционально влиять на стабильность нуклеосом. Вариации M51L и K99R в гене гистона HIST1H2AH (изоформа H2A1H) приводят к стабилизации нуклеосом и изменяют пролиферацию клеток [221]. Точно так же теперь описаны эффекты небольших вариаций между каноническими последовательностями у разных видов. Специфичный для грибов гистоновый мотив H3 QKK (Рис. 2.1в), расположенный на оси диады нуклеосомы, вносит вклад в плохую сборку октамеров в нуклеосомах дрожжей [222].

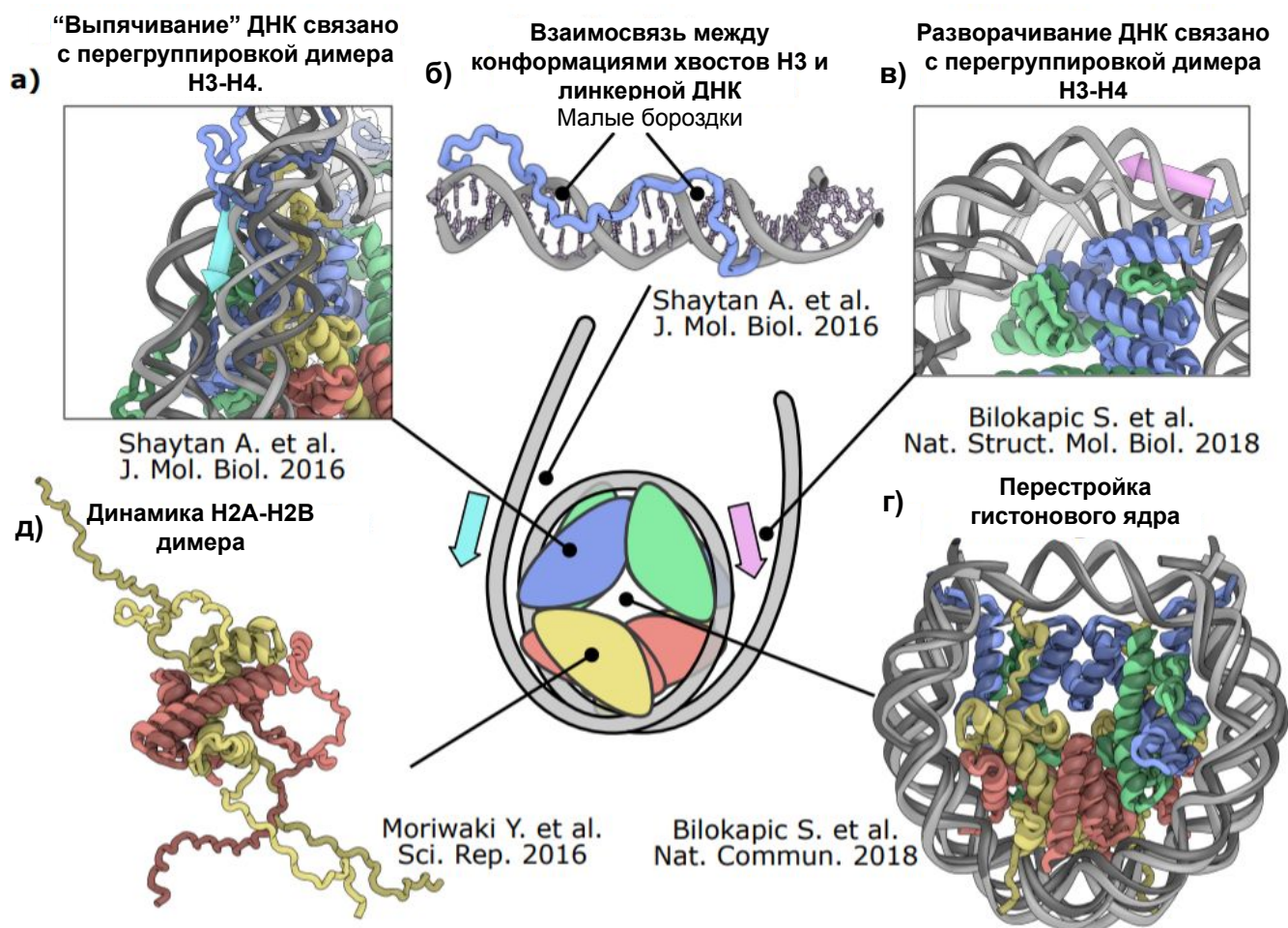
### 2.1.3 Тонкие детали динамики нуклеосом

Если распечатать димеры гистонов на 3D-принтере то их можно собрать в октамера подобно конструктору LEGO (см. 3D-печатные модели нуклеосом на <https://github.com/intbio/nucLEGO>). Этот взгляд на нуклеосому как на жесткую модульную структуру теперь уступает место альтернативному взгляду на нуклеосому как на динамическую сущность, где даже небольшие конформационные вариации функционально важны. В поддержку последней точки зрения мы обсудим несколько недавних наблюдений, представленных на рисунке 2.3. Во-первых, динамика ДНК внутри нуклеосомы выходит за рамки простого разворачивания и может демонстрировать различные искаженные конформации, включая дефекты кручения и выпячивание ДНК вблизи точки входа/выхода. Такого рода изменения недавно наблюдались как при атомистическом моделировании [A17] (Рис. 2.3 а), так и на крио-ЭМ картах [179] (Рис.2.3в). Более того, изменения в конформации ДНК связаны с тонкими изменениями конформации гистоновых белков. Например, как показано в крио-ЭМ, разворачивание 15 пар оснований ДНК с одной стороны октамера гистонов приводит к изменениям конформации гистонов вблизи точек входа / выхода с обеих сторон, а также сопровождается общим небольшим расширением октамера перпендикулярно оси симметрии нуклеосомы [179]. Более того, определенная связь между перестройками гистонового ядра и конформацией ДНК также наблюдалась для полностью завернутого состояния (Рис.2.3г), было показано, что нуклеосомы могут сокращаться на 8% вдоль оси диады и расширяться на 5% в перпендикулярном направлении [223].

Пластичность индивидуальных димеров гистонов также функционально важна. В дополнение к недавно продемонстрированной важности пластичности Н3-Н4 для ремоделирования нуклеосом [172], с помощью твердотельного ЯМР было показано, что внутренняя динамика во временных масштабах от наносекунд до миллисекунд присутствует для гистона Н4 в нуклеосомах [224]. Недавнее исследование метил-TROSY ЯМР показало, что мутантные нуклеосомы могут проявлять значительно повышенную динамику внутри гистонов Н3-Н4 [225]. Помимо Н3-Н4, пластичность Н2А-Н2В, по-видимому, также имеет решающее значение для сборки и функционирования нуклеосом. В частности, сборка нуклеосом *in vitro* со сшитыми димерами Н2А-Н2В происходит только

до стадии гексасом [179], а это означает, что для сборки требуется определенная гибкость димера. Структура ЯМР изолированного димера H2A-H2B в растворе, вероятно, отражает эти динамические режимы и демонстрирует повышенную гибкость и беспорядок, в частности, внутри дополнительных спиралей  $\alpha$ C и  $\alpha$ N по сравнению со структурой внутри нуклеосомы (Рис. 2.3e) [226]. Динамика H2A-H2B также, вероятно, используется ремоделером SWR1. Он распознает специфически канонические нуклеосомы, содержащие H2A, и заменяет H2A его вариантом H2A.Z. Способность ремоделера SWR1 распознавать и действовать на нуклеосомы H2A на не на H2A.Z недавно была связана с отличием нескольких аминокислот между H2A и H2A.Z. Причем ни одна из этих аминокислот не экспонируется на поверхности нуклеосом [227]. Поскольку рентгеновские исследования показывают очень похожую структуру нуклеосом H2A и H2A.Z [228], эти открытия подразумевают, что присутствует некая динамическая сигнатура, которая используется для распознавания нуклеосом H2A.

Наконец, тонкая динамика гибких гистоновых хвостов с их способностью взаимодействовать с ДНК и другими белками формирует еще один уровень динамической сложности. Гистоновые хвосты и другие белки хроматина несут положительно заряженные остатки (особенно аргинины [229; 230]), которые предпочитают взаимодействовать с малыми бороздками ДНК как участками отрицательного электростатического потенциала (Рис. 2.3 б). Эти взаимодействия могут, в свою очередь, модулироваться ПТМ гистонов или АТ-богатыми последовательности ДНК (что способствует образованию узких малых бороздок [231] и может распознаваться некоторыми белками в контексте нуклеосом [A6]) (Рис. 2.1 в).



**Рисунок 2.3** — Динамика ДНК в нуклеосомах тесно связана с динамикой гистонов (а, б, в, г). Цветовая схема соответствует рисунку 2.1. Голубые и фиолетовые стрелки на центральной диаграмме и панелях (а) и (в) обозначают одни и те же области входа / выхода ДНК в нуклеосому. (а) Помимо простого разворачивания ДНК, ДНК может проявлять выпуклости возле точек входа / выхода, как это наблюдается в моделировании МД и крио-ЭМ [A17; 179]. (б) Гистоновый хвост НЗ взаимодействует с линкерным сегментом ДНК. Наблюдаются преимущественные взаимодействия с малыми бороздками ДНК [A17]. (в) Раскручивание ДНК сопровождается перестройками в димере гистона НЗ-Н4 [179]. Кристаллическая структура нуклеосомы показана светлыми цветами. (г) Гистоновый октамер с полностью обернутой ДНК может принимать сжатую конформацию вдоль оси диады (темные цвета) по сравнению с кристаллической структурой по умолчанию (светлые цвета) [223]. (д) Гибкость димера гистона Н2А-Н2В, наблюдаемая с помощью ЯМР-исследований в растворе [226]. Две конформации показаны темным и светлым цветами.

## 2.2 Моделирование нуклеосом с линкерными участками ДНК

*Данный раздел написан по материалам статей [A17; A18], результаты которых основаны в том числе на предыдущих работах [A8; A12; A20; A21].*

При формировании нуклеосом октамер гистоновых белков организует около 200 пар оснований ДНК в два суперспиральных витка. Хотя статическая структура коровой частицы нуклеосомы была разрешена, детали динамических взаимодействий между гистонами и ДНК остаются не совсем понятными. Для изучения данного вопроса мы провели длительное равновесное моделирование молекулярной динамики нуклеосом, включая линкерные сегменты ДНК и полноразмерные гистоны, в явном растворителе в атомистическом приближении. Впервые мы смогли идентифицировать и охарактеризовать перестройки в нуклеосомах на микросекундном масштабе времени, изучить связь между конформацией гистоновых хвостов и геометрией ДНК. Мы обнаружили, что определенные конформации гистонового хвоста способствовали выпячиванию ДНК вблизи участков ее входа/выхода в/из нуклеосомы, что приводило к образованию дефектов кручения внутри ДНК. Мы охарактеризовали динамику гистоновых хвостов при их конденсации на нуклеосомальной и линкерной ДНК и показали, что хвосты могут принимать конформационно ограниченные позиции из-за вставки лизинов и аргининов в малые бороздки ДНК. Потенциально, эти явления влияют на доступность посттрансляционно модифицированных остатков гистонов, которые служат важными сайтами для эпигенетических меток (например, в H3K9, H3K27, H4K16). Предполагается, что взаимодействия гистоновых хвостов с коровой и линкерной ДНК модулируют процессы взаимодействия белков хроматина с гистонами, модификации хвостов и связывания эффекторных белков. В данном разделе мы обсуждаем влияние наблюдаемых явлений на функцию нуклеосом и сравниваем наши результаты с различными экспериментальными исследованиями.

### 2.2.1 Введение

Нуклеосомы - это элементарные единицы компактизации хроматина в геномах эукариот, расположенные через каждые  $200 \pm 40$  п.н. вдоль ДНК [232]. Основным источником информации об атомистической структуре нуклеосомы являются рентгеновские исследования коровых частиц нуклеосом (NCP) (см. обзор [233]), которые неизменно разрешают очень похожие структуры гистонов и ДНК независимо от вариантов последовательности гистонов, мутаций, посттрансляционные модификации, последовательности ДНК или присутствия белков, связывающих нуклеосомы, пептидов или химических агентов. Например, суперпозиция различных структур нуклеосом с использованием алгоритма VAST+ [234] показала очень небольшие вариации в среднеквадратичном отклонении (RMSD) около 1 Å для выровненных глобулярных частей гистоновых областей. Консенсусная структура NCP формируется путем наматывания  $\sim 145$ - $147$  п.н. коровой ДНК в  $\sim 1,7$  левых сверхспиральных витка вокруг октамера, состоящего из четырех типов коровых гистонов (H3, H4, H2A, H2B) [235]. Структура NCP характеризуется наличием диадной оси - оси псевдосимметрии второго порядка, сильно изогнутой структурой ДНК [236] и высоким положительным зарядом октамера гистонов [237]. Кроме того, он включает большое количество молекул воды, проникающих в структуру нуклеосомы [238], а также длинные гистоновые хвосты, выступающие из ядра нуклеосомы и обеспечивающие множество сайтов для посттрансляционных модификаций (ПТМ).

С другой стороны, решающая роль нуклеосом в функционировании хроматина (включая тонкую регуляцию экспрессии генов, репликацию ДНК, репарацию и наследование, опосредованные эпигенетическими механизмами [239; 240]) зависит от их динамической природы и конформационных переходов. Действительно, многие биофизические (FRET, эксперименты с одномолекулярными пинцетами и другие, рассмотренные в [241]) и биохимические эксперименты (анализ связывания факторов транскрипции [137], химическая сшивка [242], анализ методом ChIP-ехо [160]) предполагают, что нуклеосомы демонстрируют существенный конформационный полиморфизм как на локальном (конформация ДНК и гистоновых хвостов внутри нуклеосомы), так и на глобальном (потеря гистонов, субнуклеосомные частицы и т. д.) масштабах, которые, в свою очередь, могут быть результатом равновесных тепловых

флуктуаций или могут быть активно индуцированы внешними силами внутри клетки.

Были предложены различные функционально значимые моды динамики нуклеосом, которые включают дыхание/разворачивание/раскрытие ДНК, расщепление нуклеосом [170], а также скольжение нуклеосом [243], раскрытие или потерю димера H2A-H2B [A25]. В то же время, различные конформационные перестройки гистоновых хвостов необходимы для внутри- и межнуклеосомных взаимодействий [244], для множественных взаимодействий с ремоделерами хроматина [245; 246], гетерохроматиновыми белками [247] и со многими эффекторными белками, которые содержат домены, связывающие модификации гистонов [248]. Тем не менее, взгляду на нуклеосомы как на динамические объекты [170] все еще не хватает понимания деталей на атомистическом уровне.

Чтобы попытаться сгладить этот очевидный пробел в нашем понимании динамики и функций нуклеосом, уже были предприняты значительные усилия в области моделирования, включая моделирование молекулярной динамики (МД) на атомистическом уровне [249] и моделирование в крупнозернистом приближении [250; 251]. Однако из-за значительного размера нуклеосомы по стандартам атомистического моделирования и чрезвычайно долгих временных масштабов ее функциональной динамики в предыдущих исследованиях приходилось прибегать либо к крупнозернистому моделированию без атомистических деталей, либо к полностью атомному моделированию без явных молекул растворителя [252]. Полноатомное моделирование проводилось на временах не превышающих нескольких сотен наносекунд [253; 254]. Однако мы знаем, что изменения в атомных взаимодействиях могут иметь глубокие функциональные последствия для нуклеосом. Например, посттрансляционные модификации могут изменять внутри- и межнуклеосомные взаимодействия или обеспечивать множественные сайты связывания для так называемых эффекторных белков [255], даже однократное ацетилирование H4K16 может запускать разворачивание хроматиновых волокон [256; 257].

В работе, изложенной в данном разделе, концептуальный прогресс состоит в том, чтобы использовать наиболее надежный подход моделирования методом МД с явным растворителем для всех атомов и расширить этот подход на значительно более длительную микросекундную шкалу времени. Микросекундное моделирование, а также модели нуклеосом с реалистичными сегментами ДНК-линкера и многочисленные сравнительные модели позволили нам получить



представление о функционально значимых перестройках в нуклеосомах, включая связь между конформациями гистоновых хвостов и геометрией ДНК. Хотя многие функционально релевантные движения могут происходить в гораздо более длительных временных масштабах [241], несколько недавних исследований ЯМР показали, что гистоновые хвосты демонстрируют субмикросекундную динамику, что дает возможность сравнить наши расчетные результаты с экспериментальными данными [258–260].

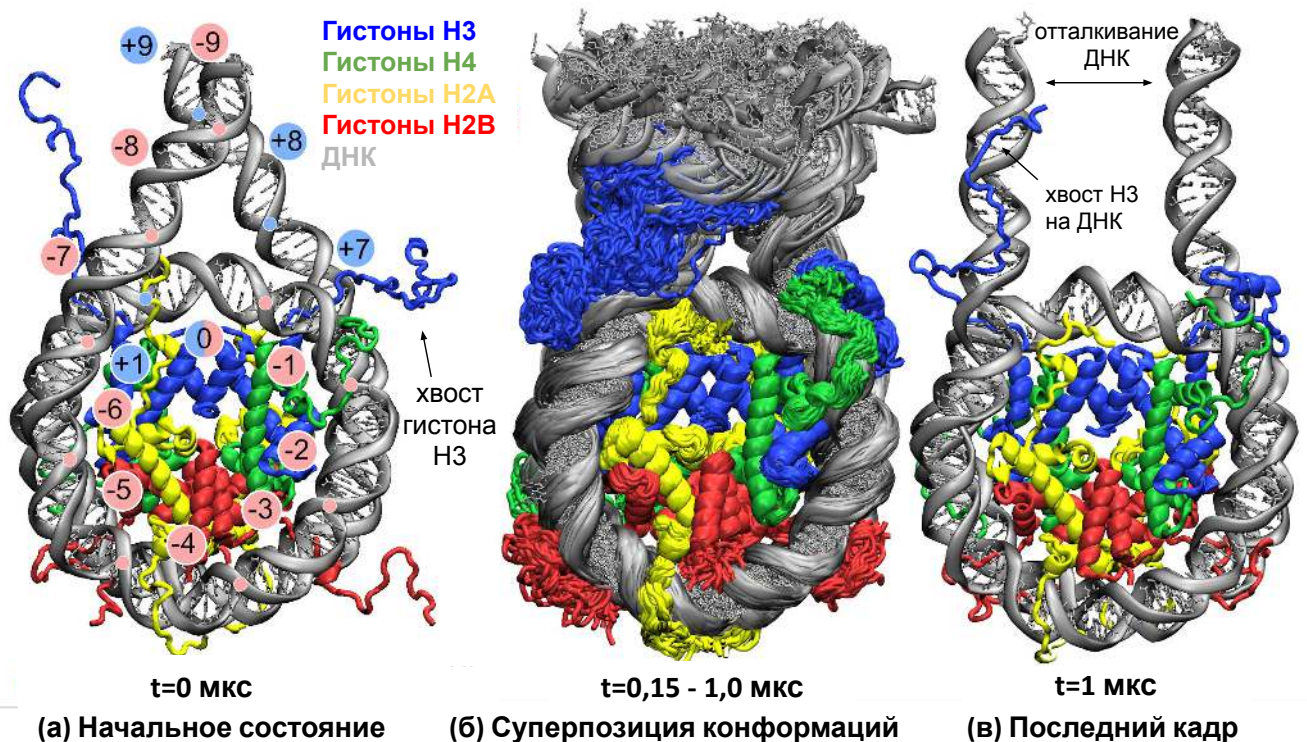
В этом исследовании мы обнаружили, что определенные конформации гистонового хвоста способствуют выпячиванию ДНК возле мест входа/выхода ДНК в нуклеосоме, образованию дефектов кручения внутри ДНК и перестройке взаимодействий гистонов с ДНК в нуклеосомном коре. Мы охарактеризовали динамику гистоновых хвостов при их конденсации на коровой и линкерной ДНК на атомистическом уровне и показали, что они могут принимать конформационно ограниченные положения, сопровождаемые вставкой определенных “закрепляющих” лизинов и аргининов в малые бороздки ДНК. Мы предполагаем, что эти процессы влияют на доступность посттрансляционно модифицированных остатков гистонов, которые служат важными эпигенетическими метками (например, H3K9, H3K27, H4K16), а взаимодействия с коровой и особенно линкерной ДНК (для хвоста H3) могут модулировать доступность различных сайтов на гистоновых хвостах для связывания эффекторными белками.

## 2.2.2 Результаты

### Обзор исследованных моделей

В нашем исследовании мы использовали несколько систем для моделирования, в том числе специально сконструированную полную нуклеосомную систему с линкерными сегментами ДНК по 20 п.н., фланкирующих коровую частицу с каждой стороны (модель FN) (Рис. 2.4а), минималистичную модель коровой частицы с усеченными гистоновыми хвостами (NCPm модель) и набор вспомогательных систем с различными условиями моделирования (подробности см. в таблице 2). Ниже мы в основном сосредотачиваемся на ре-

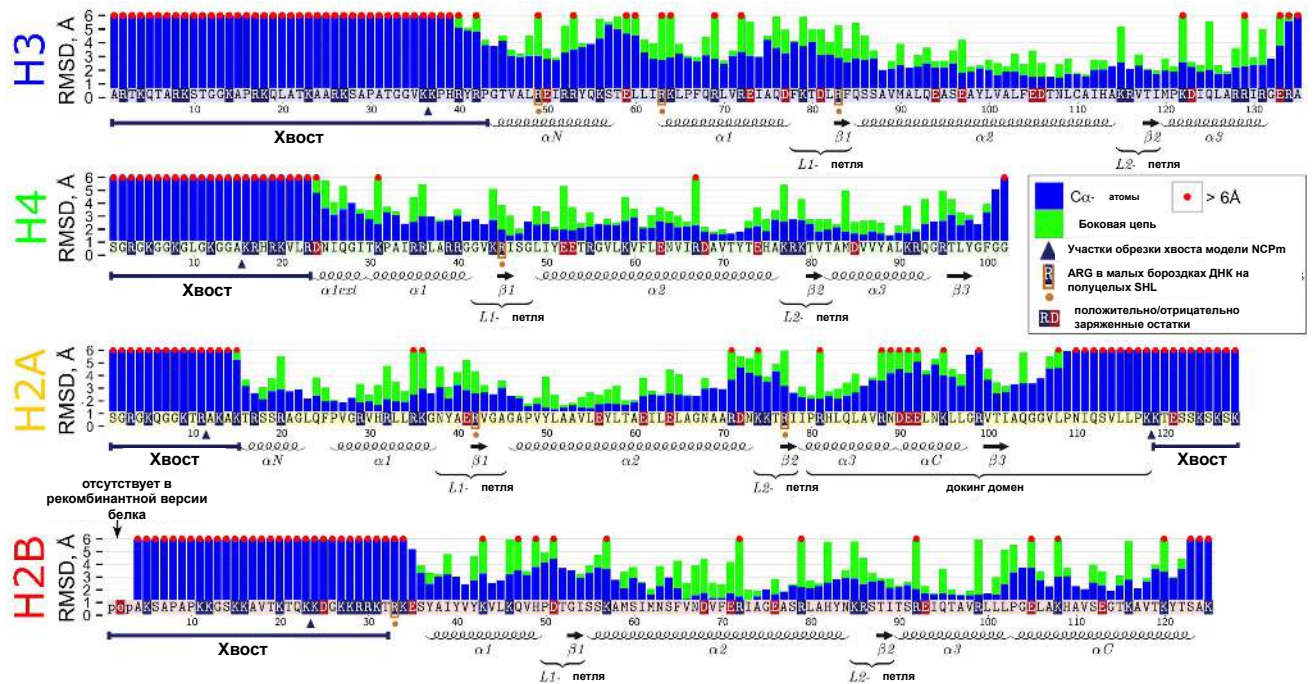
зультатах, полученных для полной нуклеосомной системы, для которой были выполнены обширные микросекундные расчеты. Мы допускаем, что динамика нуклеосом происходит в нескольких временных масштабах, а функционально значимые конформационные переходы часто превышают масштаб времени в одну микросекунду. Наше моделирование ограничивалось исследованием конформационных ансамблей вокруг определенных локальных квазиравновесных состояний в микросекундной шкале времени. Поэтому, интерпретируя результаты нашего моделирования, проверяли наши выводы, выполняя несколько вспомогательных расчетов. При обсуждении результатов ниже мы ссылаемся на ключевые структурные элементы гистоновых последовательностей и нуклеосом, как показано на рисунке 2.5 (см. Также раздел “Название элементов структуры нуклеосом” ниже).



**Рисунок 2.4** — Обзор структуры и динамики модели полной нуклеосомы (FN): (а) Исходная структура модели FN, (б) наложенные конформации из последних 75 % кадров расчетной траектории, (в) последний кадр после 1 мкс, обратите внимание на различные формы N-концевых хвостов H3 с обеих сторон.

Таблица 2 — Исследованные в моделировании системы и их параметры

Название	Описание	Количество атомов, тыс.	Эффект. концентрация соли, мМ	Размер ячейки, Å	Время моделирования, нс
FN	Система с линкерной ДНК и гистоновыми хвостами	366,8	~185	162x197x113	1000
FN1M	То же, что и не FN, но в высокой концентрации соли	366,8	~1100	161x196x112	1000
FNbb	То же, что и не FN, но в большей ячейке	3200	~160	322x358x274	1000
FNnt	То же, что и не FN, но с ацетильными и N-метил группами на концах гистонов	366,8	~185	162x197x113	1000
NCPm	Система, без линкерной ДНК и хвостов гистонов	310,3	~200	145x141x101	1000



**Рисунок 2.5** — Максимальные наблюдаемые среднеквадратичные отклонения отдельных аминокислот ( $\alpha$ -атомы - синие столбцы, атомы боковой цепи - зеленые столбцы) во время моделирования относительно их положения в исходной рентгеновской структуре. Значения, превышающие  $6\text{Å}$ , усекаются при этом значении и отмечаются красной точкой над столбцом. Аннотации структурных элементов гистонов приведены под столбиками.

## Конформационная стабильность и пластичность гистонов и ДНК

В соответствии с предыдущими экспериментальными исследованиями, наше моделирование подтвердило общую стабильность ядра нуклеосомы в отношении диссоциации компонентов нуклеосомы или крупномасштабного разворачивания ДНК (см. Обзор динамики на рисунке 2.4). В целом, RMSD центральной области гистонов оставалось в среднем в пределах  $1,5\text{Å}$  от кристаллографических положений для всех изученных моделей. Это было верно даже для минималистичной системы NCPm с усеченными гистоновыми хвостами. Таким образом, предполагается, что нуклеосомы имеют достаточный запас стабильности в отношении усечения гистоновых хвостов на микросекундных временах. Однако подробный анализ конформационной динамики, представленный ниже, предоставил новое понимание конформационной гибкости нуклеосом

и выявил конформационные перестройки внутри гистонового ядра, гистоновых хвостов, ядра ДНК и линкерных областей, которые происходят в микросекундной шкале времени.

Чтобы предоставить подробную картину динамики гистонов, мы изобразили максимальные отклонения позиций атомов, наблюдаемые в ходе динамики, от положений в кристаллической структуре (Рис. 2.5). Как можно видеть на рисунке 2.5, относительно большие конформационные изменения более чем на 6 Å наблюдались для атомов остова всех хвостов гистонов, даже некоторые части гистоновых фолдов демонстрируют высокое RMSD. Хотя, как показывает анализ кристаллических структур, докинг домен H2A (docking domain) образует тесные контакты с тетрамером  $(H3 - H4)_2$  [228], мы продемонстрировали существенные конформационные изменения для С-концевой части этого домена относительно исходной кристаллической структуры. Конформационные флуктуации остова белковой цепи часто сопровождалась переориентацией боковых цепей аминокислотных остатков. Например, изменения в  $\alpha$ С-спирали докинг-домена сопровождалась временным образованием новых солевых мостиков между H2A R88 и R99, фланкирующими эту спираль, и H3 E105, D106. Наблюдаемая пластичность докинг-домена будет влиять на его функции в качестве партнера по связыванию белков хроматина. В частности, недавно было показано, что шаперон ANP32E может быть ответственным за встраивание гистона H2A.Z в нуклеосому. ANP32E может связываться с  $\alpha$ С-спиральной областью докинг-домена H2A.Z [261], которая в нуклеосоме структурно очень похожа на каноническую  $\alpha$ С-спираль H2A. Связывание гистона с ANP32E стерически несовместимо с полной структурой нуклеосомы и может запускать разборку и вытеснение нуклеосом, когда гибкая область ANP32E вставляется в кор нуклеосомы. Последний процесс не очень хорошо изучен, однако повышенная конформационная гибкость в этой области, наблюдаемая в нашем исследовании, указывает на возможные механизмы этого процесса. Интересно, что Biswas et al. в более ранних работах по МД моделированию нуклеосом показали потенциальную связь между усечением гистоновых хвостов H3 и конформациями боковых цепей H2A R81 и R88 [262]. Визуальное исследование траекторий динамики наших систем показывает, что, хотя боковая цепь H2A R88 очень гибкая в микросекундном масштабе времени, боковая цепь H2A R81 в основном заблокирована в своей кристаллографической ориентации и иногда

перескакивает в положение, направленное в сторону от октамера гистонов. В последнем положении боковая цепь может находиться сотни наносекунд.

Кроме того, значительные смещения остова полипептидной цепи (более 3Å) и боковых цепей (более 6Å) наблюдались в центральных областях других гистонов, включая H3  $\alpha$ N-спираль, H3  $\alpha$ 1-спираль, сайт связывания L1-L2 димера H2A-H2B и H2B  $\alpha$ C-спираль) (Рис. 2.5). Как уже упоминалось, даже если остов белка не демонстрирует больших отклонений от кристаллографических положений, все же могут происходить значительные изменения конформеров боковых цепей. Фактически, все аминокислоты, кроме одной, с наблюдаемой переориентацией боковой цепи более 6Å, принадлежали заряженным остаткам. Эти переориентации часто включают разрушение и образование новых солевых мостиков внутри гистонового ядра. Изменения в модели NCPm показали аналогичные закономерности.

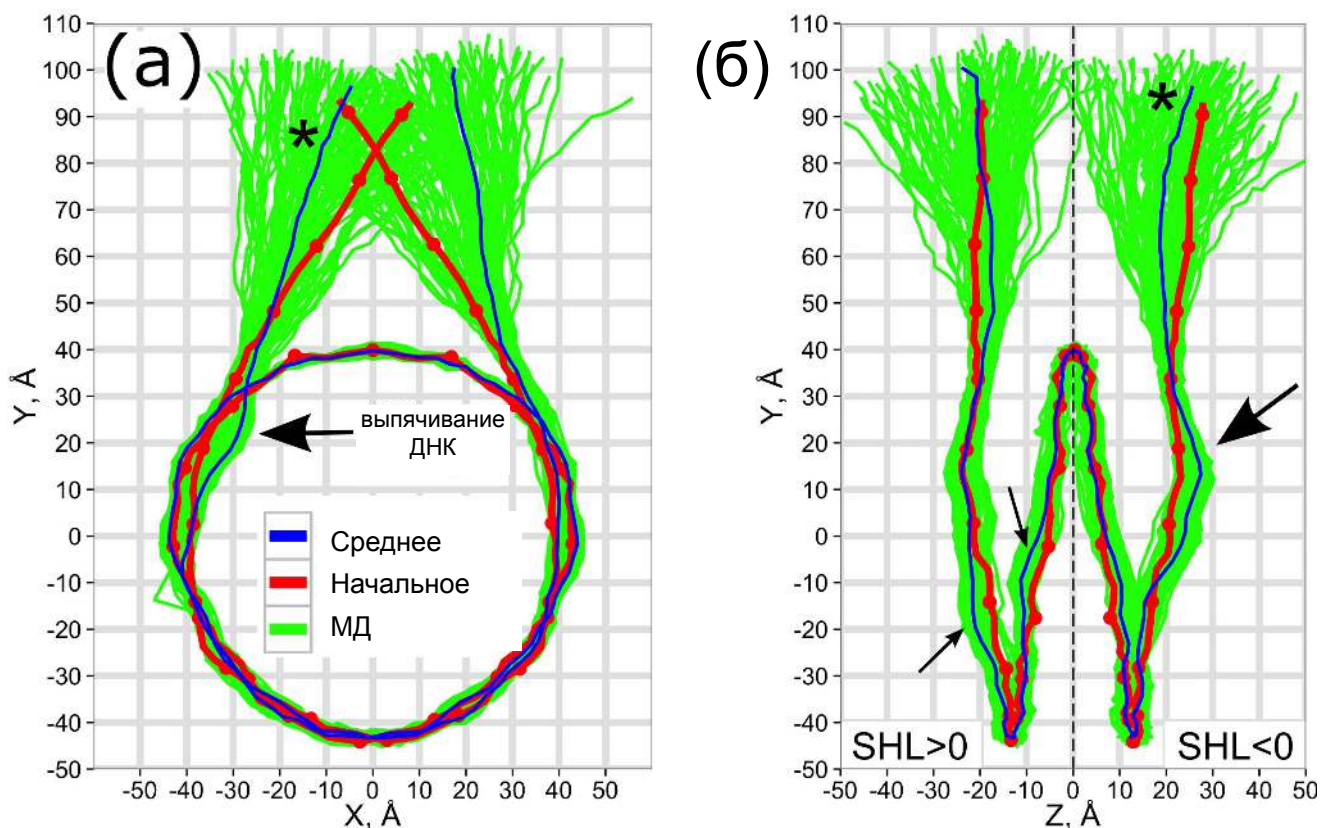
Анализ конформационной динамики коровой и линкерной ДНК показал, что ДНК оставалась в среднем в пределах 4Å (5 Å для модели NCPm) от своих кристаллографических положений, измеренных с помощью RMSD атомов N1 и N9 азотистых оснований. Кроме того, фосфатный остов ДНК характеризуется периодическим профилем средних флуктуации (RMSF): участки остова ДНК, контактирующие с гистонами, проявляли значительно меньшие колебания по сравнению с теми, которые были обращены в раствор. С другой стороны, линкерные сегменты ДНК в нашем моделировании явно демонстрировали гораздо более сильные колебания, чем коровые участки ДНК (Рис. 2.4б,в), что согласуется с известными данными [263]. Глобальная конформация ДНК была дополнительно изучена с помощью 2D-проекций многоугольников, соединяющих центры пар оснований ДНК в нуклеосомной суперспиральной системе отсчета (см. Рис. 2.6). Позиционные колебания линкерных сегментов ДНК охватывают частично перекрывающиеся конусообразные области в пространстве с угловым размахом около  $\pm 45$  градусов в обеих проекциях. Никаких существенных корреляций в позиционных колебаниях линкеров ДНК обнаружено не было. Средние положения, принятые для линкерных цепей ДНК в MD, были дальше друг от друга по сравнению с их исходными положениями, которые соответствовали прямому продолжению концов коровой ДНК (Рис. 2.6а). Моделирование той же системы, но в солевом растворе с высоким содержанием 1 М (система FN1M) ясно показало, что два линкера ДНК приближаются друг к другу и даже большую часть времени взаимодействуют своими концами.

Предполагаем, что скрининг электростатических взаимодействий значительно повлиял на угол входа/выхода ДНК. Интересно, что более ранние экспериментальные измерения расстояний между линкерами с помощью FRET показали, что расстояние монотонно уменьшается с увеличением концентрации соли [264].

Первостепенным вопросом динамики нуклеосом является амплитуда, временная шкала и механизмы отворачивания/дыхания ДНК, разворачивания или открытия поверхности октамера гистонов, что, как полагают, имеет решающее значение для доступности ДНК и связывания факторов транскрипции [137; 265]. К сожалению, на сегодняшний день нет полной картины, поскольку разные экспериментальные установки, интерпретация и терминология часто поддерживают разные взгляды на эту проблему [170; 241]. В этом исследовании мы не наблюдали крупномасштабного разворачивания или раскрытия коровой ДНК ни в системах FN, ни в системах NCPm. Такое разворачивание, если оно произойдет, обнажит достаточно частей ДНК для распознавания факторами транскрипции. Наши результаты согласуются с экспериментальными оценками временных масштабов событий разворачивания, которые лежат в субсекундном диапазоне [137; 266]. Однако, если мы рассмотрим более умеренные колебания концов коровой ДНК вместе с линкерной ДНК (часто называемые “дыханием ДНК”), такие процессы могут происходить в более коротких временных масштабах [267]. Например, недавние измерения FRET расстояний между двумя красителями, расположенными симметрично на обоих сегментах линкерной ДНК на расстоянии 5 п.н. от сайтов входа / выхода ДНК, показали две разные конформации с характерными расстояниями между двумя красителями  $\sim 4,5$  нм и  $\sim 6,5$  нм (амплитуда около 2 нм), в предположении модели с двумя состояниями [268]. В нашем моделировании расстояние между центрами пар оснований соответствующих позиций линкеров ДНК варьировалось от 4,5 нм до 6,8 нм (см. Рис. 2.6) (со средним значением 5,6 нм и RMSF 0,4 нм), показывая, что эти диапазоны расстояний находятся в пределах досягаемости быстрых субмикросекундных колебаний. Кроме того, мы наблюдали асимметрию в средних конформациях ДНК двух линкеров, а также в их отклонениях от исходной кристаллической структуры (обозначенной стрелками на рисунке 2.6). Интересно, что если мы предположим, что наблюдаемые асимметрии соответствуют двум различным стабильным конформационным состояниям ДНК в нуклеосоме, переход между этими состояниями может уже объяснить изменение положения ДНК в 0,5 нм на расстоянии 5 п.н. от входа/выхода. Однако мы признаем,

что методология для точного сравнения различных конформационных переходов, наблюдаемых при моделировании, и результатов измерений FRET все еще требует разработки и проверки: как недавно было отмечено Ленцем и др., флуктуация формы ДНК в нуклеосоме может размывать сигналы FRET от промежуточных конформационных состояний ДНК в нуклеосоме [269].

В следующих разделах мы продемонстрируем, что эти конформационные асимметрии ДНК следствие асимметричной конденсацией гистоновых хвостов. Хотя мы находим, что время, необходимое для переключения между различными конформациями гистоновых хвостов, явно превышает микросекундную шкалу времени, мы предполагаем, что динамика ДНК в микросекундной шкале времени определяется взаимодействиями между гистоновыми хвостами и ДНК, которые, в свою очередь, могут влиять на конформацию ДНК на гораздо более длительных временах.



**Рисунок 2.6** — Конформационный ансамбль ДНК в нуклеосоме, (а) - передняя и (б) - боковые проекции. Красными точками отмечены целые и полуцелые значения суперспирального положения (SHL) вдоль исходной конформации, звездочкой отмечен один линкерный сегмент ДНК на обоих рисунках. Стрелки указывают максимальные отклонения между исходной и средней конформациями в МД.



## Быстрая конденсация гистоновых хвостов на коровой и линкерной ДНК

Положительно заряженные гистоновые хвосты были тщательно изучены как экспериментально, так и с помощью методов моделирования. Известно, что межнуклеосомные взаимодействия опосредуются хвостами и играют важную роль в уплотнении хроматина. В этой работе мы охарактеризовали динамику взаимодействий между гистонами и ДНК на атомном уровне и обнаружили, что в среднем 60% всех контактов были обеспечены взаимодействиями ДНК с участками гистонового хвоста. Хвосты гистонов, которые в исходной модели выступали в объеме растворителя (в основном хвосты H3, Рис. 2.4а), быстро адсорбировались на ДНК в течение первых 20 нс моделирования. От 50 до 90% всех аминокислот в этих хвостовых областях имели прямые или опосредованные водой взаимодействия с коровой или линкерной ДНК в течение 1 мкс моделирования (Рис. 2.7). Какого-либо существенного скопления контактов к началу или концу хвостов мы не наблюдали. Предыдущие полно-атомные МД расчеты нуклеосомных частиц также указали на значительную ассоциацию гистоновых хвостов с коровой ДНК на временном масштабе 100 нс [252; 262; 270]. Однако были высказаны определенные опасения по поводу коротких временных масштабов моделирования, а также артефактов, возникающих из-за периодических граничных условий и решеточных методов суммирования электростатических взаимодействий. Чтобы решить последнюю проблему, мы выполнили моделирование на протяжении 80 нс в большой ячейке (система FNbb), размер которой должен минимизировать потенциальные неблагоприятные эффекты периодических граничных условий. Моделирование FNbb подтвердило предпочтительную ассоциацию гистоновых хвостов с ДНК и быструю конденсацию длинных N-концевых хвостов H3. Таким образом, мы снимаем ранее высказанные опасения и показываем, что конденсация гистонового хвоста не является артефактом конкретных методов, используемых в МД-моделировании.

Несколько исследований малоуглового рассеяния рентгеновских лучей (SAXS) касались вопроса о том, выступают ли гистоновые хвосты из ядра нуклеосомы в раствор или связаны с ДНК с физиологической ионной силой. Сообщалось о значительном увеличении максимального диаметра нуклеосом

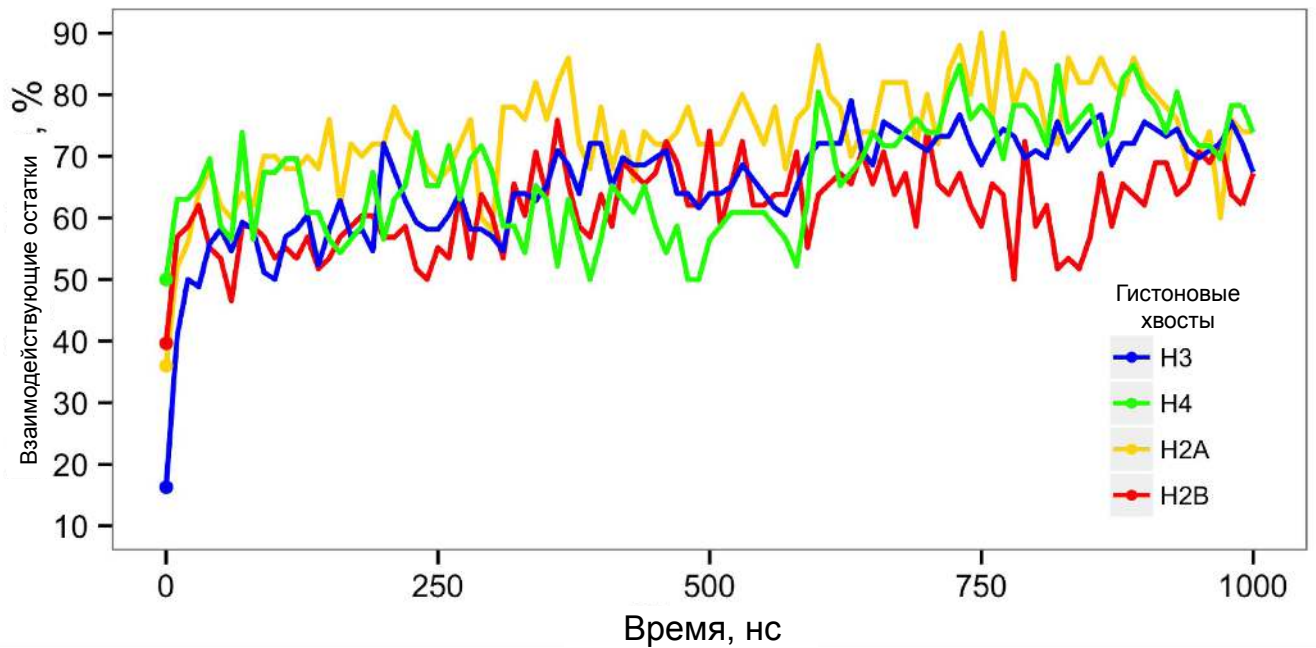
при повышении концентрации одновалентной соли с 10 мМ до физиологических концентраций и выше [271]. Этот результат был приписан процессу диссоциации гистоновых хвостов от кора нуклеосомы. Однако Ян и др. [272] указали, что сопоставимое увеличение диаметра можно объяснить конформационными изменениями нуклеосом, включая разворачивание ДНК. Наши результаты подтверждают модель, в которой гистоновые хвосты прикреплены к коровой или линкерной ДНК большую часть времени в этой временной шкале. Если по какой-то причине хвосты оторвались, характерное время их присоединения должно составить около нескольких десятков наносекунд, о чем свидетельствует их быстрая конденсация (Рис. 2.7). Наши теоретические оценки радиусов гирации предполагают, что и гистоновые хвосты, и линкерная ДНК вносят вклад в увеличение радиусов гирации.

Чтобы дополнительно выяснить влияние концентрации соли на поведение хвоста, мы выполнили моделирование при чрезвычайно высокой концентрации соли 1М (модель FN1M). Хотя известно, что октамерная форма нуклеосом нестабильна при такой концентрации соли [273], мы не наблюдали никакой разборки нуклеосом, что позволяет предположить, что это происходит во временном масштабе, намного превышающем микросекунду. В то время, как текущее силовое поле может несколько переоценивать натрий-фосфатные взаимодействия при высоких концентрациях соли [274], оно не должно препятствовать отсоединению хвостов от ДНК. Тем не менее, это позволяет нам исследовать эффекты повышенной концентрации соли на гистоновые хвосты и линкерную ДНК, в то время как ядро нуклеосомы все еще остается в компактном состоянии. Интересно, что хотя хвостам (особенно хвосту H3) потребовалось несколько больше времени (70 нс по сравнению с 20 нс при концентрации соли 150 мМ), чтобы установить контакты с ДНК из их исходной конформации, гистоновые хвосты все еще конденсируются на ДНК, несмотря на значительное усиление скрининга электростатических взаимодействий, вызванного высокой концентрацией соли. Этот факт можно рассматривать как еще одно свидетельство, подтверждающее модель тесных взаимодействий между гистоновыми хвостами и ДНК в нуклеосомах, по крайней мере, когда кор нуклеосомы находится в компактно собранном состоянии.

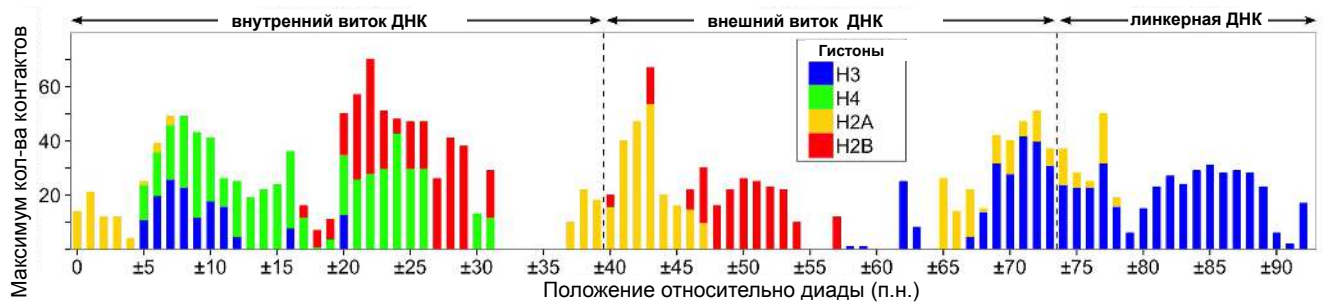
В литературе ведутся дискуссии о конденсированных и деконденсированных состояниях гистоновых хвостов в нуклеосомах, и, по нашему мнению, основные физические принципы поддерживают последнюю точку зрения. Как

обсуждалось в [275; 276], достаточно длинные олигокатионы имеют тенденцию почти полностью ассоциироваться с высоко заряженной ДНК из-за увеличения свободной энергии при высвобождении конденсированных небольших одновалентных ионов.

Комбинированный ансамбль конформаций гистоновых хвостов, отобранных в различных моделях, показал широкий диапазон участков ДНК, доступных для взаимодействия с гистоновыми хвостами (Рис. 2.8). Например, было показано, что линкерный ДНК-сегмент длиной 20 п.н. полностью доступен для взаимодействий с N-концевым хвостом Н3 и частично доступен для взаимодействий с С-концевым хвостом Н2А вблизи точки входа / выхода ДНК. В моделях FN и FNnt один N-концевой хвост Н3 оставался вытянутым и стабильно связанным с ближайшим линкерным сегментом ДНК (Рис. 2.4в и 2.9а), в то время как другой хвост образовывал  $\alpha$ -спираль возле точки входа / выхода ДНК (Н3 21-28) или находился между супервитками ДНК (Рис. 2.9д, 2.9е). Этот наблюдаемый полиморфизм конформаций хвоста Н3 перекликается с утверждениями из различных экспериментальных исследований. Связывание хвостов Н3 с линкерной ДНК недавно было подтверждено в исследовании ChIP-ехо в масштабе генома [160]. С другой стороны, исследования ЯМР и дейтериевого водородного обмена указывают на то, что в нуклеосомных массивах хвосты Н3 могут образовывать стабильные компактные структуры [258]. Возможность для Н3-хвоста конденсироваться на линкерной ДНК предполагает также его роль в частичной нейтрализации линкерной ДНК и конкурентном связывании с другими белками, взаимодействующими с линкерной ДНК (такими как гистон Н1 или определенные комплексы ремоделирования нуклеосом), что важно для понимания функционирования хроматина, учитывая очень высокие концентрации нуклеосом в ядре ( $> 200$  мг/мл) [277].



**Рисунок 2.7** — Конденсация гистоновых хвостов на ДНК. Доля остатков гистонового хвоста, которые имеют прямые или опосредованные водой взаимодействия с ДНК, представлена как функция времени моделирования. Начальные значения обозначены точками.



**Рисунок 2.8** — Доступность ДНК для взаимодействия с гистоновыми хвостами. Максимальное количество атом-атомных контактов между ДНК и гистоновыми хвостами показано для всех кадров из различных систем моделирования (FN, FNnt и FNbb).

### Хвосты гистонов “застревают” в малых бороздках ДНК и влияют на доступность эпигенетически модифицированных сайтов

Конденсация гистоновых хвостов на ДНК в ходе моделирования явно изменила конформационную динамику хвостов и самой ДНК, фактически она

напоминала двумерную диффузию на поверхности ДНК с малыми бороздками ДНК, служащими кинетическими ловушками. Несмотря на обширное моделирование, за 1 мкс мы заметили, что хвосты двух копий одного и того же гистона часто исследовали не перекрывающиеся подмножества конформационного пространства. Мы предполагаем, что переключение между различными конформационными подпространствами для таких хвостов происходило во временном масштабе более длительном, чем время нашего моделирования (Рис. 2.9). Такое поведение согласуется с экспериментальными данными по химическому сшиванию и ЯМР гистоновых хвостов [242; 259]. Интересно, что в нашем микросекундном моделировании скорость конформационной динамики гистоновых хвостов зависела от их конформации и способа связывания с ДНК. Например, амплитуда колебаний хвоста Н3 на одной стороне нуклеосомы, который адсорбируется на линкерной ДНК, была намного выше (из-за гибкости линкерной ДНК) по сравнению с другим хвостом Н3, который был компактно свернут на около входа ДНК в нуклеосому. Динамические характеристики гистоновых хвостов в мононуклеосомах и массивах нуклеосом недавно были исследованы с помощью ряда методов, включая твердотельный ЯМР и ЯМР в растворе, дейтеро-водородный обмен [258—260]. Данные исследования содержат достаточно противоречивые выводы. Принимая во внимание наши результаты, мы предполагаем, что динамика гистоновых хвостов даже в компактных массивах нуклеосом может быть организована неоднородным образом, при этом одни хвосты стабильно свернуты, в то время как другие активно исследуют конформационное пространство.

Подробный анализ взаимодействий между отдельными боковыми цепями аминокислот и основаниями ДНК указывает на то, что определенные боковые цепи хвостов гистонов могут быть глубоко вставлены в малые бороздки ДНК, взаимодействуя с основаниями ДНК и выступая в качестве якорей, стабилизирующих взаимодействия и ограничивающие конформационную динамику. Наибольшее количество контактов с основаниями ДНК образовано аргининами и лизинами, расположенными внутри гистоновых хвостов. А именно, в основной (FN) модели, следующие остатки показали высокую склонность к взаимодействию с основаниями малой бороздки ДНК: R8 и R26 гистона Н3; K16 и R17 гистона Н4; R11, K13 и K126 гистона Н2А, R29 и R30 гистона Н2В. Интересно, что для хвостов Н3 и Н4 ни одно из этих взаимодействий с основаниями ДНК не наблюдалось в исходной рентгеновской структуре. В то время, как среди

контактов белок-ДНК внутри гистонового кора преобладали взаимодействия с фосфатами (в модели FN 82% - фосфаты, 16% - сахара, 2% - основания), в области гистоновых хвостов 17% контактов - это взаимодействия с основаниями, что значительно больше, чем количество контактов между основаниями ДНК и гистонами в кристаллической структуре (8%).

Следует отметить, что многие упомянутые выше сайты потенциально могут подвергаться посттрансляционным модификациям, включая метилирование аргининов и метилирование и ацетилирование лизинов. Например, два из наиболее важных эпигенетически модифицированных остатков на гистоне H3 (H3K9 и H3K27) расположены рядом с остатками аргинина (H3R8 и H3R26), которые показали высокую тенденцию к встраиванию в малые бороздки ДНК. Следовательно, такие мотивы аргинин-лизин внутри гистоновых хвостов могут влиять на доступность определенных лизинов для эффекторных белков и их способность посттрансляционно модифицироваться.

Если мы проанализируем те сайты, которые устанавливают прямые контакты с нуклеотидами в малой бороздке ДНК в наших моделях, наиболее выраженным является H4K16, который находится в позиции ДНК SHL -1,5. Это хорошо известный сайт ацетилирования, который регулирует компактизацию хроматина и связывание ремоделера ACF [278]. Исследование мононуклеосом методом ЯМР подтвердило гибкость остатков 1-15 хвоста гистона H4, тогда как остатки 16-22 плотно взаимодействуют с нуклеосомным кором и не демонстрируют мобильности в соответствующей шкале времени ЯМР. Более того, было экспериментально подтверждено, что ацетилирование, имитирующее мутацию H4K16Q, которая, согласно нашему исследованию, вероятно, ослабит взаимодействие этого остатка с малой бороздкой ДНК из-за потери положительного заряда боковой цепи, вызывает структурное разупорядочение в этой области [259]. В соответствии с нашими наблюдениями, экспериментальное исследование показало, что H4H18 и H4K16 могут быть связаны с позициями ДНК SHL  $\pm 1.5$  [279]. Наши наблюдения о связывании малой бороздки с H4K16 также согласуются с более ранними выводами Erler et al. в REMD-моделировании нуклеосом [252]. Совсем недавно Collepardo-Guevara et al. в многомасштабном моделировании показали, что ацетилирование лизина увеличивает содержание вторичных структур в гистоновом хвосте и снижает доступность хвоста для решающих межнуклеосомных взаимодействий, компактизирующих хроматиновые волокна [280], что согласуется с более ранним моделированием конформацион-

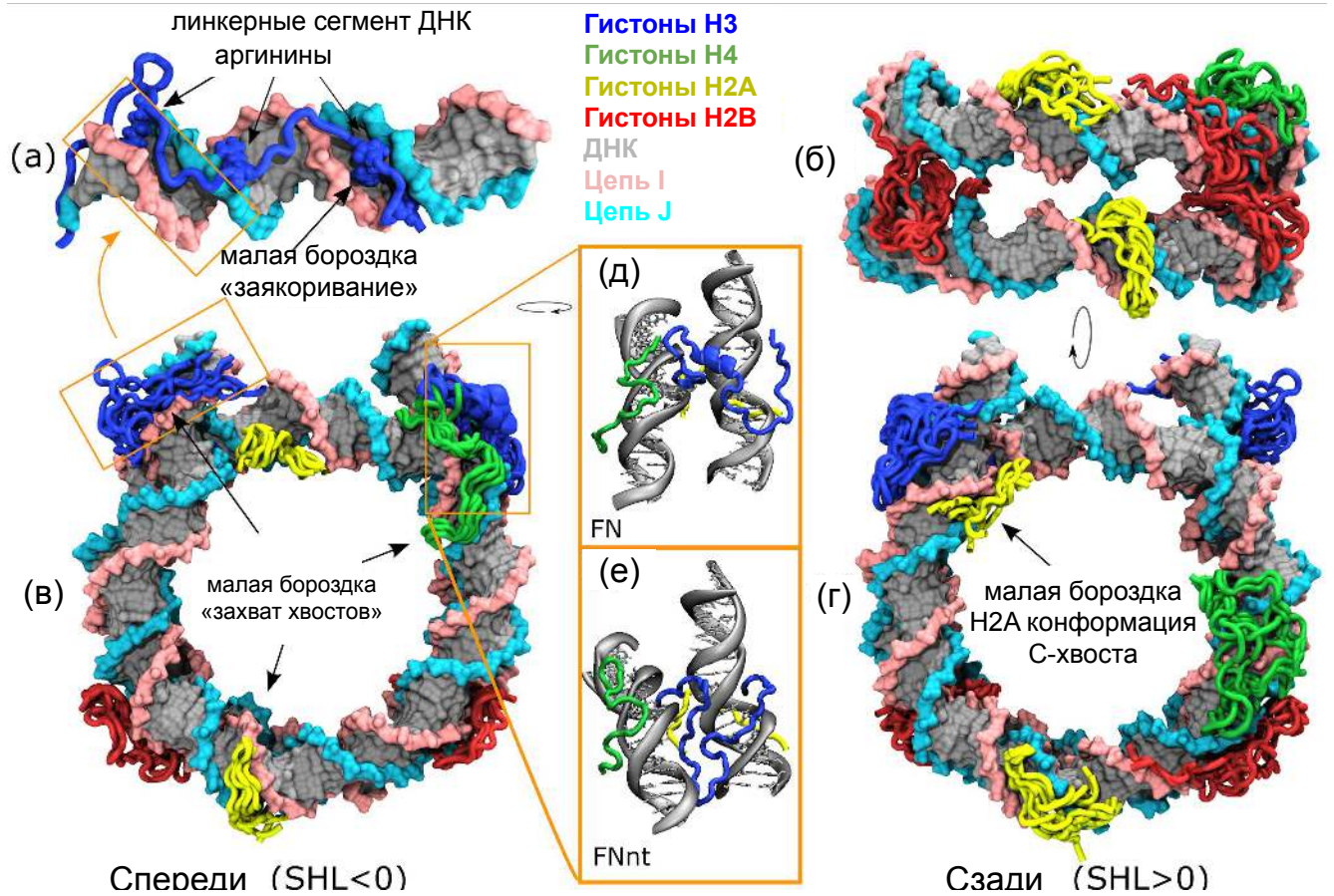
ного ансамбля фрагментов гистоновых хвостов без ДНК в растворе [204; 281]. Эти результаты также согласуются с более ранней работой по моделированию взаимодействий ДНК с пептидами из H4-хвоста [282; 283], в которой авторы предположили, что заряженные группы аргининов и лизинов играют главную роль в притяжении ДНК-ДНК через хвост гистона, образуя мосты между фосфатными группами и взаимодействуя с электроотрицательными участками в малой бороздке.

Наш сравнительный анализ взаимодействий гистонов с ДНК в различных областях нуклеосомы предполагает, что большинство взаимодействий в области глобулярных частей гистонов обеспечивается преимущественно остатками аргинина, в то время как взаимодействия между хвостами гистонов и ДНК почти в равной степени обеспечиваются остатками аргинина и лизина (Рис. 2.10). Это не относится к исходной конформации, происходящей из кристаллической структуры NCP, где взаимодействия белок-ДНК через остатки лизина в значительной степени недостаточно представлены. Хотя следует отметить, что в кристаллической структуре NCP дополнительные контакты гистон-ДНК могут формироваться с соседними нуклеосомами в кристаллической решетке, которые мы не включаем в наш анализ. Такая тенденция согласуется со склонностью аргининов находится в малых бороздках ДНК, как это наблюдалось в кристаллических структурах многих комплексов белок-ДНК [284]. Как видно из рисунка 2.10, глицин также преобладает в H3, H2A и особенно в N-концевых хвостах H4, и его карбонильная группа основной цепи может образовывать неспецифические контакты с фосфатами основной цепи ДНК. Важно отметить, что, хотя прямые контакты белков с основаниями ДНК (прямое считывание) обычно считаются почти отсутствующими в нуклеосомах и не влияют на позиционирование нуклеосом, это может быть слишком упрощенной картиной. Например, было показано, что гистоновые хвосты могут вносить вклад в зависимость от последовательности позиционирование нуклеосом [285]. Однако этот эффект не воспроизводился на последовательностях с высоким средством. Обширные контакты гистоновых хвостов с основаниями ДНК, наблюдаемые в нашем исследовании, могут помочь понять и интерпретировать эти экспериментальные результаты.

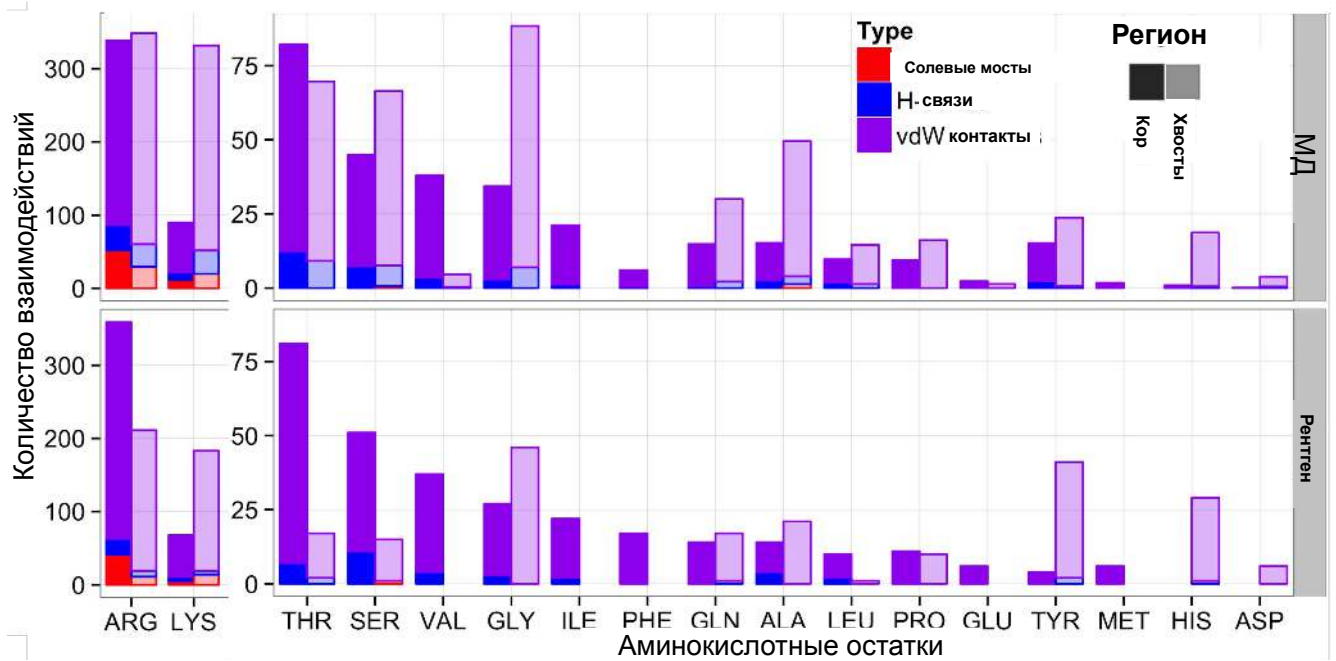
## Выпячивание ДНК возле точек входа/выхода модулируется конформациями гистонового хвоста

Теперь обратим внимание на анализ связи между конформацией гистоновых хвостов и геометрией нуклеосомной ДНК. В ходе моделирования проявились некоторые изменения в конформации ДНК (Рис. 2.6аб). Самая большая замеченная перестройка – выпячивание ДНК рядом с ее точкой входа в нуклеосому в районе  $SHL \pm 6,5$ . Данную перестройку можно было наблюдать только во временных масштабах более 100 нс, и она была обусловлена различными конформациями гистоновых хвостов (Рисунок 2.9). Взаимодействия в этой области ДНК (концевые 10 п.н. коровой ДНК) в основном представлены взаимодействиями с гистоновыми хвостами, а не с глобулярными доменами, что позволяет предположить, что гистоновые хвосты могут оказывать сильное влияние на конформацию ДНК, приводя к потенциальной стабилизации или дестабилизации всей нуклеосомы. Действительно, в моделировании FN-модели терминальная коровая область ДНК вокруг  $SHL +6,5$  образовывала намного больше контактов с гистоновыми хвостами, чем ее симметричный аналог в  $SHL -6,5$  (184 контакта против 64), что вызывало ее стабилизацию и предотвращало выпячивание ДНК. Эта стабилизация, по-видимому, была вызвана обширными взаимодействиями с N-концевым хвостом H3, который формирует компактную вторичную структуру около места входа / выхода ДНК, а также вставкой H2A K126 в малую бороздку ДНК (Рис. 2.9). Таким образом С-концевой хвост H2A также может участвовать в стабилизации ДНК. Фактически, недавно было экспериментально обнаружено, что частичная делеция С-концевой области H2A приводит к разворачиванию последних 10 п.н. коровой ДНК [286]. В соответствии с этим фактом, колебания концевых частей коровой ДНК были менее ограниченными в нашем моделировании для системы NCPm, которая имела усеченные гистоновые хвосты (Рис. 2.6аб).

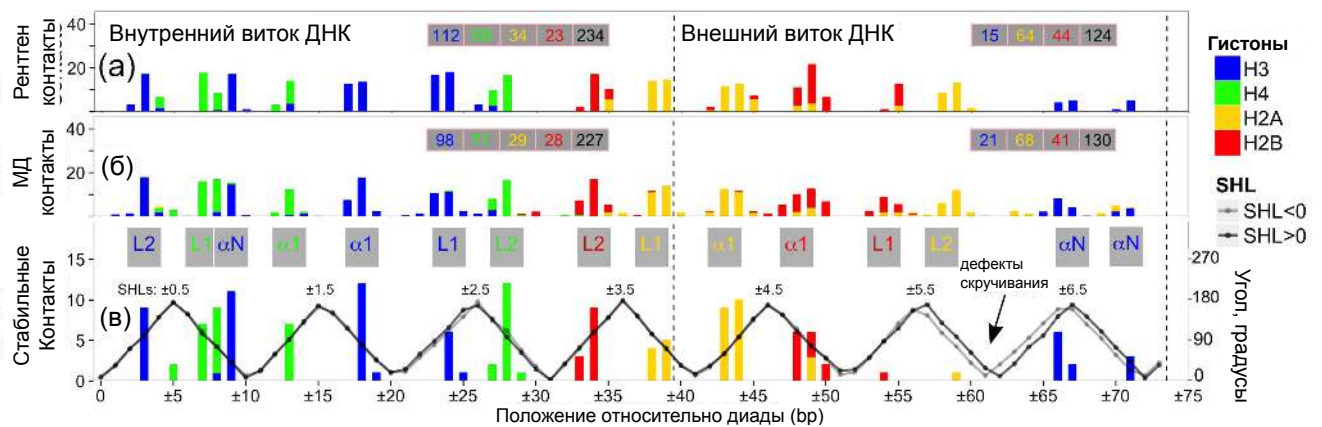




**Рисунок 2.9** — Типичные паттерны взаимодействия гистоновых хвостов с ДНК, наблюдаемые при МД моделировании. (а) Типичная конформация N-концевого хвоста H3, взаимодействующего с линкерной ДНК. (б) - снизу, (в) - спереди и (г) - виды нуклеосомы с множественными наложенными конформациями гистонического хвоста, наблюдаемыми в течение 1 мкс МД моделирования (изображены каждые 100 нс с отброшенными первыми кадрами 250 нс). (д) формирование  $\alpha$ -спирали, образованной остатками 21-28 одного из хвостов H3 (цепь A), (е) альтернативное стабильное положение хвоста H3, “состыкованное” между супервитками ДНК.



**Рисунок 2.10** — Количественная оценка взаимодействий белок-ДНК по типу, участку гистона и участвующим аминокислотам.



**Рисунок 2.11** — Взаимодействия белок-ДНК в области нуклеосомного кора. (а) Профиль контактов для рентгеновской структуры, усредненный по двум половинам нуклеосомы. (б) Среднее количество контактов по всем кадрам и по симметричным половинам нуклеосомы. Общее количество контактов для соответствующих участков ДНК и гистонов показано в розовых рамках для графиков (а) и (б). (в) Количество стабильных контактов белок-ДНК. В каждой позиции ДНК указывается количество уникальных стабильных контактов атом-атом, наблюдаемых на любой половине нуклеосомы (т.е. связанные с симметрией стабильные контакты, наблюдаемые в двух половинах нуклеосомы, подсчитывались только один раз, асимметричные контакты вносились индивидуально). Структурные элементы сайтов связывания ДНК аннотированы в верхней части этого графика разными цветами, соответствующими разным типам гистонов. Черные и серые кривые показывают периодичность вращения ДНК в структуре нуклеосом, усредненной по траектории.

### Сайты связывания ДНК в гистоновом коре демонстрируют возможность перестройки, сопровождающейся образованием дефектов кручения ДНК

В наших предыдущих разделах мы говорили о взаимодействиях между гистоновыми хвостами и ДНК; в этом разделе мы обсудим взаимодействия гистонов с ДНК, вовлекающие глобулярные домены гистонов, и их связь с перестройками геометрии ДНК. Из анализа кристаллической структуры NCP известно, что ДНК имеет семь пар симметричных сайтов связывания, где малая бороздка ДНК обращена к октамеру гистонов. В каждом сайте связывания

ключевой аргинин вставлен в малую бороздку в большинстве случаев, за исключением канонических аргининов R49 из обеих копий H3, которые расположены слишком далеко, чтобы их можно было вставить в малую бороздку в районе  $SHL \pm 6.5$ . В то же время это место имеет небольшое количество контактов с гистоновым ядром (Рис. 2.11а,б). Наш анализ динамики взаимодействий гистон-ДНК показал, что этот канонический H3R49 действительно может быть вставлен в соответствующую малую бороздку из-за выпячивания ДНК в районе  $SHL \pm 6.5$ , описанной ранее. Этому выпячиванию, по-видимому, способствовало образование дополнительных и перестройка существующих контактов между  $\alpha N$ -спиралью H3 и ДНК. Чтобы понять последствия этих изменений для общей геометрии ДНК, мы построили график эффективного вращения пар оснований ДНК относительно положения суперспиральной оси (Рис. 2.11в). Мы наблюдали сдвиг в общем скручивании ДНК на одну пару оснований вокруг  $SHL -6/-6,5$ , а в случае модели NCPm такие сдвиги наблюдались в области от  $SHL -5,5$  до конца коровой ДНК. Подобные сдвиги в геометрии скручивания ДНК ранее были названы “твист-дефектами” [194]. Некоторые экспериментальные исследования показали, что нуклеосома существует в растворе как смесь различных состояний твист-дефектов, и только некоторые из них могут быть увидены в кристаллических структурах путем изменения основной последовательности ДНК, ее длины или добавления интеркалирующих агентов [194]. В нашем исследовании мы впервые смогли наблюдать образование таких дефектов *in silico* и описать атомистические детали, лежащие в основе связи между геометрией ДНК и взаимодействиями гистон-ДНК.

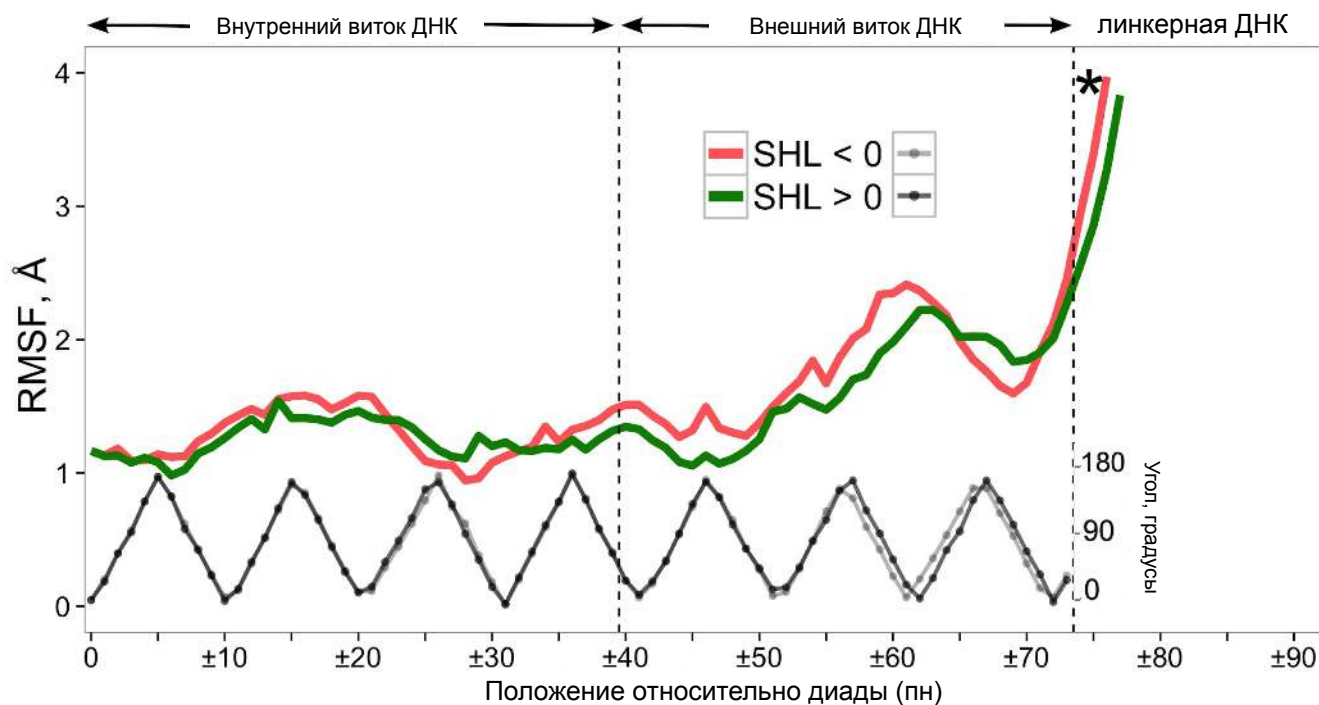
Чтобы выяснить закономерности взаимодействий в различных сайтах связывания гистонов с ДНК, мы вычислили количество так называемых стабильных контактов (см. Методы). Как видно на рисунках 2.11а-в, количество контактов между гистоновым кором и внутренним супервитком ДНК было почти вдвое больше, чем количество контактов с внешним витком ДНК. Это согласуется с более легким разворачиванием внешнего витка ДНК в экспериментах с единичными молекулами [287; 288]. Более того, рисунок 2.11 показывает четкую иерархию между различными каноническими сайтами связывания. Сайты связывания вокруг  $SHL \pm 0,5$  и  $SHL \pm 4,5$  имеют наибольшее количество динамически стабильных контактов в соответствии с картированием с высоким разрешением взаимодействий белок-ДНК в нуклеосомах, полученным путем механического растягивания ДНК. Последние эксперименты показали,

что наиболее сильные взаимодействия гистон-ДНК происходят вокруг диады и в районе расположения  $\pm 40$  пар оснований от диады [287].

Первый сайт связывания гистонов с ДНК вокруг SHL  $\pm 0,5$  образован петлями L1/L2 димера H3-H4, с дополнительным взаимодействием, исходящим от N-конца  $\alpha$ N-спирали H3, что делает этот сайт очень особенным с точки зрения его уникальной специфической организации. Различия между паттернами связывания гистонов с ДНК наблюдались на картах нуклеосом по всему геному, и недавно с помощью геномного анализа позиционирования нуклеосом был обнаружен необычный сигнал относительно содержания A|T/G|C в нуклеосомной ДНК в девяти парах оснований от центра нуклеосомы [158]. Предполагалось, что это может быть связано с внедрением H3R40 глубоко в малую бороздку ДНК, либо с образованием специфичных для последовательности контактов с основаниями [289]. Однако мы не наблюдали таких контактов в микросекундном масштабе времени, и H3R40 взаимодействовал с фосфатами ДНК, а не с основаниями.

Второй высокостабильный сайт связывания вокруг SHL  $\pm 4.5$  образован двумя  $\alpha 1$ -спиралями димера H2A-H2B, тогда как сайты связывания ДНК с наименьшим числом стабильных взаимодействий локализованы на SHL  $\pm 6.5$ , SHL  $\pm 5.5$  и SHL  $\pm 1.5$ . Карта контактов гистонов с ДНК согласуется с профилем флуктуаций ДНК (Рис. 2.12), а флуктуации ДНК показывают видимое увеличение при SHL  $\pm 1,5$ . Это согласуется с повышенной конформационной гибкостью ДНК в этих сайтах при связывании малых молекул или белков показанной в недавних экспериментах. Эти эксперименты показали, что SHL  $\pm 1,5$  является местом связывания ионов тяжелых металлов и интеркалирующих противоопухолевых агентов [166]. Отсутствие стабильных взаимодействий вокруг SHL  $\pm 5.5/\pm 6.5$  также соответствует увеличению флуктуаций ДНК. Колебания на самом конце коровой ДНК несколько меньше, что объясняется их стабилизацией через N- и C-концевые хвосты H3 (см. Предыдущие разделы). Малая стабильность и высокая гибкость сайтов связывания SHL  $\pm 5.5$  и SHL  $\pm 6.5$  может указывать на то, что эти сайты вносят лишь небольшие энергетические барьеры при разворачивании ДНК из нуклеосомы. Биологически значимый процесс, который основан на таком постепенном разрыве контактов ДНК-гистонов, включает транскрипцию РНК-полимеразами, которая может происходить без вытеснения нуклеосом [290]. Нуклеосомный барьер для транскрипции, исследованный с помощью паттернов транскрипционного паузирования, показал, что

первый барьер встречается полимеразой, когда передний край полимеразы входит в область примерно в 40 п.н. от диады [291]. Данный факт согласуется с представлением о том, что первые два сайта связывания почти не представляют барьера для транскрипции.



**Рисунок 2.12** — Флуктуации ДНК в нуклеосоме. RMSF центров пар оснований для двух симметричных половин нуклеосом (красная и зеленая кривые); черные и серые кривые показывают периодичность вращения суперспирали ДНК. Точки данных, превышающие  $\text{RMSF } 4\text{\AA}$ , не показаны. Звездочкой обозначена та же цепь, что и на рисунке 2.6.

### 2.2.3 Обсуждение

В данном разделе мы представили молекулярно-динамическое исследование полной модели нуклеосомы с полноразмерными гистоновыми хвостами и линкерными участками ДНК. Мы использовали полноатомное моделирование в явном растворителе и распространили этот подход на микросекундную шкалу времени. Микросекундная шкала времени и многочисленные сравнительные симуляции позволили нам получить представление о функционально значимых перестройках в конформациях нуклеосом, включая связь между конформациями гистоновых хвостов и ДНК.

Механистические наблюдения, полученные в нашем исследовании, предполагают, что конформационные перестройки внутри коровой ДНК зависят от паттернов взаимодействий ДНК с гистонами. В частности, мы наблюдали, что N-концевой хвост H3 и C-концевой хвост H2A устанавливают много контактов с коровой ДНК и могут стабилизировать геометрию ее концевой области, подавляя образование локализованных дефектов кручения внутри ДНК. С другой стороны, конформация хвостов, которые выступают из нуклеосомы, способствуют образованию состояний с дефектами кручения, которые могут быть важными промежуточными звеньями в скольжении и ремоделировании нуклеосом [243]. Мы предлагаем наличие механизма, с помощью которого наблюдаемое чрезмерное скручивание ДНК в терминальной области ДНК может представлять собой начальное состояние дефекта кручения ДНК, которое может быть стабилизировано последующими перестройками взаимодействий ДНК в центральной области гистонов. Последующее скольжение нуклеосом может происходить за счет распространения этого дефекта дальше по кору нуклеосомы. Интересно отметить, что несколько исследований по химическому сшиванию моонуклеосом *in vitro* ранее показали, что гистоновые хвосты занимают различные предпочтительные конформации в NCP, нуклеосомах с линкерами и нуклеосомами с гистонами H1 [292; 293]. Данный факт позволяет предположить, что гистоновые хвосты очень чувствительны к наличию/отсутствию линкерной ДНК, гистона H1 и изменений в нуклеосомном окружении.

В соответствии с результатами экспериментов по химическому сшиванию ДНК с гистонами [293; 294] мы обнаружили, что гистоновые хвосты легко адсорбируются на линкерную (в основном хвосты H3) и коровую ДНК (в основном хвосты H4, H2B и H2A), и большинство контактов гистонов с ДНК представлено контактами с гистоновыми хвостами. О возможности взаимодействия N-концевых хвостов H3 с линкерной ДНК сообщалось ранее [160; 293]. Фактически, эффективность химического сшивания в NCP оказалась в 3–4 раза ниже для H2A, H2B и H4 и в 10–12 раз ниже для H3 по сравнению с полными нуклеосомами с линкерной ДНК [294]. Более того, наше детальное изучение поведения гистоновых хвостов показало, что их динамика в микросекундном масштабе характеризуется ограниченной диффузией на поверхности ДНК и кинетическим захватом хвостовых областей в малых бороздках ДНК. Наблюдаемое стабильное прикрепление гистоновых хвостов к коровой и линкерной ДНК предполагает глубокие последствия для взаимодействий нуклеосом

с другими белками хроматина. В самом деле, когда эти белки связываются с гистоновыми хвостами или ДНК, сначала они должны проконкурировать со взаимодействиями ДНК-гистоновый хвост и вытеснить соответствующего партнера. Например, Pilotto et al. недавно экспериментально продемонстрировали элегантную практическую иллюстрацию этой концепции на примере комплекса LSD1-CoREST, который действует как деметилаза H3K4 [295]. Было высказано предположение, что этот комплекс связывается с нуклеосомой через свою субъединицу CoREST, которая вытесняет хвост H3 с ДНК. Это смещение критически необходимо для того, чтобы сделать метилированный H3K4 доступным для взаимодействий с его субъединицей LSD1. Можно предположить, что такие механизмы могут быть использованы для организации перекрестного взаимодействия между множественными сайтами связывания на гистоновых хвостах, включая сайты ПТМ и их соответствующих партнеров по связыванию [255; 296].

Атомистические детали связывания гистонового хвоста с ДНК, выявленные в ходе нашего исследования, дополнительно показали, что этому процессу способствует вставка закрепляющих боковых цепей аргинина и лизина в малые бороздки ДНК. Это, в свою очередь, повлияло на те остатки, которые служат эпигенетически модифицированными сайтами в хвостах гистонов, или на остатки, расположенные рядом с ними (например, H3K9, H3K27, H4K16 и H3R8, H3R26 и т.д.). Присутствие множества сайтов связывания на одном хвосте гистонов и окклюзия боковой цепи малыми бороздками может вызывать кооперативные эффекты. А именно, связывание одного партнера с его сайтом связывания на ДНК или сайтом гистонового хвоста может вызвать смещение гистонового хвоста от ДНК и, таким образом, облегчать связывание другого партнера. Подобные эффекты могут быть вызваны посттрансляционными модификациями, такими как метилирование аргинина и особенно ацетилирование лизина, которые не способствуют встраиванию этих остатков в малую бороздку и, таким образом, могут способствовать конформационной гибкости и доступности связывания гистоновых хвостов. Наблюдаемые эффекты помогают наметить дальнейший план экспериментальных исследований, которые должны прояснить сложное динамическое взаимодействие между посттрансляционными модификациями гистонов и доступностью гистоновых хвостов для связывания эффекторных белков.



## 2.2.4 Материалы и методы

### Построение начальной модели

В качестве отправной точки мы использовали рентгеновскую кристаллическую структуру высокого разрешения ( $1,9\text{\AA}$ ) коровой частицы нуклеосомы, образованную рекомбинантными вариантами канонических коровых гистонов *X. laevis* модифицированной  $\alpha$ -сателлитной ДНК человека (PDB ID 1KX5 [238]). Для создания полной нуклеосомной модели с линкерными сегментами ДНК с помощью программы NAB был сконструирован прямой дуплекс В-ДНК  $(AGTC)_5$  длиной 20 п.н. [297]. Используемая линкерная последовательность ДНК сбалансирована по количеству гибких и жестких динуклеотидов [105]. Она была прикреплена к основной ДНК на обоих концах NCP. Один из хвостов гистона H3 был слегка повернут, чтобы избежать стерических столкновений с линкерной ДНК (угол  $\Psi$  Lys36 был установлен на  $-35$  градусов) (Рис. 2.4а). Минималистичная модель NCP (NCPm) была получена из той же кристаллической структуры путем усечения гистоновых хвостов в сайтах, указанных на рисунке 2.5 треугольниками. Все модели были явно сольватированы в прямоугольной ячейке с минимальным расстоянием между растворенным веществом и границами ячейки в  $20\text{\AA}$ , за исключением системы FNbb (таблица 1), для которой использовался порог в  $100\text{\AA}$ . В системы добавляли ионы натрия для нейтрализации, а затем добавляли дополнительные ионы натрия и хлора в концентрации  $150\text{ mM}$  по отношению к объему воды, за исключением системы FN1M, где использовалась концентрация  $1\text{ M}$ . Точная соответствующая объемная концентрация ионов была оценена непосредственно из данных моделирования. Кристаллографические молекулы воды остались в системе, а все кристаллографические ионы были удалены. Состояния протонирования аминокислот определяли на основании значений  $pK$  их растворов при нейтральном pH, остатки гистидина считали нейтральными и протонировали на  $\epsilon$ -азотах. Модель FNnt была идентична модели FN, но имела нейтрально заряженные белковые концы для изучения устойчивости МД-моделирования к небольшим возмущениям. Структуры исходных моделей представлены на рис. 2.4а. На выбор NaCl в качестве соли для нашего моделирования повлияло его обычное

использование в экспериментах с нуклеосомами *in vitro*, а также в протоколах сборки нуклеосом [298].

## Протоколы моделирования и выбор силового поля

Силовое поле CHARMM36 использовалось для ДНК и белка [299; 300], параметры TIP3P для молекул воды и скорректированные параметры Луо и Ру для ионов [301]. Выбор силового поля всегда является деликатным вопросом из-за известных и неизвестных приближений, используемых в силовом поле, а также постоянными усилиями по улучшению силового поля, предпринимаемых разработчиками. В случае нуклеосом проблема осложняется необходимостью сочетать точные модели белков и ДНК, а также обеспечивать реалистичное моделирование взаимодействий с растворителем. Ниже мы кратко обсудим несколько недавних современных исследований о поведении различных силовых полей при моделировании нуклеосом.

Точность силового поля белка, по-видимому, особенно важна для моделирования поведения изначально неупорядоченных гистоновых хвостов. Систематические исследования биомолекулярных силовых полей [302] показали, что определенные силовые поля (например, CHARMM27, AMBER ff03) сверхстабилизируют спиральную конформацию пептидов, и поэтому эта проблема была решена в пересмотре силового поля для белков CHARMM36 [299]. Кроме того, несколько недавних исследований поставили под сомнение применимость силовых полей белков для специфического моделирования гистоновых хвостов [252; 280; 281]. Например, Ertler et al. показали, что разные версии силовых полей AMBER могут приводить к разному динамическому поведению гистоновых хвостов [252]. Collepardo-Guevara et al. сравнили результаты моделирования с использованием современных силовых полей, включая CHARMM36, с данными ЯМР по динамике гистонового хвоста и пришли к выводу, что все силовые поля дают почти идентичные результаты [280].

Силовое поле ДНК необходимо для правильного воспроизведения конформационных переходов ДНК при моделировании нуклеосом. Параметризация силовых полей нуклеиновых кислот оказывается более сложной задачей, чем аналогичная задача для белков. В настоящее время известно, что силовые по-

ля как CHARMM, так и AMBER воспроизводят стабильную динамику ДНК в микросекундном масштабе времени и за его пределами [303]. Однако точное воспроизведение экспериментальных результатов по зависимой деформируемости ДНК от последовательности с использованием доступных силовых полей все еще остается нерешенной проблемой [304].

Наши модельные системы были подготовлены с помощью программы VMD [305], а моделирование МД было выполнено с помощью пакета NAMD 2.9 [306]. Для моделирования при постоянной температуре использовался подход динамики Ланжевена с шагом интегрирования 2 фс, параметром демпфирования  $0,5 \text{ }^{-1}$  и  $T = 310\text{K}$ . Поддержание давления в 1 атм было реализовано методом Ланжевена. Моделирование проводилось с жесткими ковалентными связями, и ван-дер-ваальсовы взаимодействия постепенно выключались на расстоянии от 10 до  $12\text{\AA}$ . В электростатических расчетах использовался метод PME с шагом сетки  $1\text{\AA}$ , кубической интерполяцией, отсечкой в реальном пространстве  $12\text{\AA}$  и параметром толерантности  $10^{-6}$ . Использовались периодические граничные условия. Для устранения диффузии нуклеосом к атомам C- $\alpha$  гистоновых фолдов H3 (номера остатков 64-78, 86-114, 121-131) применялись небольшие позиционные ограничения в  $0,003 \text{ ккал/моль/\AA}^2$ . Чтобы избежать расплетания пар оснований на концах ДНК в моделировании NSPm, использовали ограничивающий потенциал в виде искусственной стенки, чтобы сохранить расстояние между центрами масс оснований в концевых парах оснований в пределах 120% от начального.

Все системы были подвергнуты минимизации энергии и первоначальному уравниванию. Затем проводилось моделирование до времени моделирования 1  $\mu\text{s}$ . Кадры траектории сохранялись каждые 100 пс. Мы запускали расчеты параллельно на высокопроизводительных компьютерных кластерах/суперкомпьютерах, используя эффективное распараллеливание, доступное в NAMD. Скорость моделирования варьировалась в зависимости от моделируемой системы, количества процессоров и архитектуры машины. Для справки: система модели FN моделировалась параллельно на 384 ядрах ЦП в течение 120 дней со скоростью  $\sim 8 \text{ нс/день}$ .

## Анализ траекторий

Анализ и визуализация траектории выполнялись с использованием набора библиотек и скриптов собственной разработки, написанных на TCL, Python и R, которые использовали возможности VMD [305] и 3DNA [114] для общего и специфического анализа структуры ДНК. Для структурного анализа нуклеосом отдельные кадры траектории были наложены на исходную кристаллическую структуру, путем минимизации значения среднеквадратичного отклонения (RMSD) между положениями атомов C- $\alpha$  спиралей гистоновых фолдов. Анализ конформации ДНК проводили по отношению к нуклеосомной суперспиральной системе отсчета, определяемой ее диадной осью и суперспиральной осью. Периодичность вращения ДНК в нуклеосоме определяли путем расчета угла между вектором пары оснований (соединяющим атомы N1 и N9 соседних оснований в паре оснований) и суперспиральной осью нуклеосомы. Максимумы и минимумы значений периодичности вращения соответствовали целочисленным и полуцелым значениям SHL.

Подробный анализ взаимодействий гистонов с ДНК проводился для каждого кадра траектории (первые кадры 250 нс не учитывались как начальный период конформационного уравнивания) путем анализа положений соответствующих атомов. Контакты (SC) между двумя атомами гистона и ДНК определялись как контакты между тяжелыми атомами на расстоянии менее 3,9Å. Контакты были дополнительно классифицированы следующим образом: солевые мостики (SB) включают два заряженных неводородных атома на расстоянии менее 3,9Å; водородные связи (HB) были определены как связи между донорными (D) и акцепторными (A) атомами с водородом между ними (DH-A), где расстояние между D и A было меньше 3,5Å, а DH-A угол составлял больше 150 градусов; контакты Ван-дер-Ваальса (vdW) были определены как контакты между атомами, которые не были связаны водородными связями и не образовывали солевой мостик; и опосредованные водой взаимодействия (WM) определялись как взаимодействия между неводородными атомами ДНК или гистонами, которые образуют водородную связь с одной и той же молекулой воды. Мы также ввели понятие стабильных контактов между ДНК и белком, чтобы описать контакты, которые сохранялись во время моделирования. А именно, они были определены как отдельные пары атомных контактов между ДНК и

молекулами белка, которые присутствовали более чем в 80% кадров траектории после начального периода конформационного уравнивания.

## Соглашения об описании структуры нуклеосом

Позиции пар оснований ДНК были пронумерованы относительно центральной пары оснований (называемой диадой), ее положение принималось равным нулю. Вращательная ориентация двойной спирали ДНК преимущественно описывается параметром сверхспирального расположения (SHL) (Рис. 2.4а), который мы расширяем, чтобы включить не только коровую ДНК (SHL от 0 до  $\pm 7$ ), но и линкерную ДНК (SHL до  $\pm 9$ ). Исходные 147 пар оснований ДНК NCP называются “коровой ДНК”, а области, где коровая ДНК соединяется с линкерной ДНК, называются сайтами входа/выхода ДНК. Мы различаем две части в каждом гистоне: область хвоста (как указано на рисунке 2.5) и остальную часть, называемую глобулярной частью гистона. Ключевыми элементами глобулярных областей являются спирали гистоновых фолдов  $\alpha 1$ ,  $\alpha 2$  и  $\alpha 3$  [307] и петли L1, L2, как показано на рисунке 2.5.

## Благодарности

Исследования данного раздела были поддержаны программами внутренних исследований Национальной медицинской библиотеки и Национального института рака, Национальных институтов здоровья; и Российским научным фондом [грант № 14-24-00031] (разработка алгоритмов визуализации нуклеосом). Автор был поддержан программой сотрудничества между США и Россией в области биомедицинских наук. В этом исследовании использовались высокопроизводительные вычислительные возможности кластера Biowulf Linux в Национальном институте здравоохранения, Бетесда, Мэриленд <http://biowulf.nih.gov>. Данное исследование было частично поддержано Суперкомпьютерным центром МГУ им. М.В. Ломоносова и вычислительными ресурсами

суперкомпьютера Hexagon Cray XE6m-200, в университете Бергена и норвежском метацентре высокопроизводительных вычислений (NOTUR).

## 2.3 Моделирование коровых частиц нуклеосом на масштабах свыше 10 микросекунд

*Данный раздел описывает работы, выполненные в 2017-2020 годах, являющиеся продолжением работ предыдущего раздела. Работы основаны на использовании более современных компьютерных архитектур с использованием графических процессоров и учитывают серьезный объем новых экспериментальных данных, опубликованных в 2017-2020 годах. Интерактивные материалы раздела доступны по ссылке [http://intbio.github.io/nucl\\_MD\\_15mus/](http://intbio.github.io/nucl_MD_15mus/).*

### 2.3.1 Введение

Многие химические и физические сигналы регулируют обработку генетической информации в живых клетках. У эукариот эта регуляция в основном происходит на уровне нуклеосом - последовательных отрезков ДНК, обернутых вокруг октамеров гистоновых белков [117; 163]. Четыре типа гистонов (H3, H4, H2A, H2B) образуют два типа димеров (H3-H4, H2A-H2B). Четыре димера образуют октамер с двойной осью симметрии (диадной осью) [308]. Как известно из кристаллографических исследований, примерно 147 пар оснований ДНК резко изгибаются вокруг октамера в  $\sim 1,7$  витка левой суперспирали, формируя коровую частицу нуклеосомы (NCP) [165] (Рис. 2.13а).

При формировании нуклеосом жесткая макромолекула ДНК с персистентной длиной около 50 нм должна резко изгибаться до радиуса кривизны всего в 5 нм. Это возможно благодаря электростатическому взаимодействию отрицательно заряженной ДНК с положительно заряженными гистонами, дополненному индивидуализированным паттерном взаимодействия белок-ДНК на атомном уровне в 14 сайтах связывания ДНК (по 7 с каждой стороны NCP). Эти

сайты связывания расположены в положениях, где малая бороздка ДНК обращена к поверхности октамера и обычно присутствует боковую цепь остатка аргинина, которая выступает в пространство малой бороздки. Общая энергия взаимодействий гистонов с ДНК в нуклеосомах оценивается как не менее 23 ккал/моль [309]. Стабильность нуклеосом может варьироваться на 2-4 ккал/моль в зависимости от гибкости последовательности ДНК [310], которая, как было показано, зависит от содержания GC, наличия гибких динуклеотидных ступеней YR, участков поли (dA:dT), эпигенетических модификации оснований ДНК, таких как 5-метилцитозин и т. д. [210; 311–313]. Еще одним важным фактором являются вариации последовательности гистонов. Несмотря на то, что гистоновые белки одни из самых консервативных белков в эволюционном смысле, они подвержены множеству функционально значимых вариаций, включая посттрансляционные модификации (ПТМ) [314] и изменение последовательности за счет включения вариантов, изоформ и мутаций гистонов [A14; 315; 316].

Представление о нуклеосомах как о средстве простой компактизации ДНК теперь устарело. Различные режимы АТФ-зависимой и АТФ-независимой динамики нуклеосом, модулируемые гистоновым составом и последовательностью ДНК, обеспечивают богатый динамический ландшафт для регуляции генома [A3]. Например, скольжение нуклеосом с помощью АТФ-зависимых ремоделеров хроматина [317] или их реконфигурация с помощью РНК-полимераз [A23; 318] в значительной степени вовлечены в транскрипцию и ее регуляцию. Пассивная динамика нуклеосом также участвует во многих процессах. Новаторские эксперименты Джона Видома показали, что временное отворачивание ДНК от октамера гистонов может быть использовано факторами транскрипции для доступа к их сайтам связывания [137]. Недавние исследования методами FRET [319; 320], ЯМР [172; 225], крио-ЭМ [223; 321] и масс-спектрометрии [A1; 172; 322] свидетельствуют о том, что существуют более тонкие режимы конформационной динамики нуклеосом, которые воспринимаются и используются белками хроматина. Например, кручение ДНК внутри NCP обеспечивает путь для АТФ-зависимого ремоделирования нуклеосом [321; 323]. Внутренняя динамика (пластичность) гистонового ядра также участвует в этом процессе (хотя на этот счет и ведутся дискуссии) [172; 324; 325]. Сообщалось, что подавление этой динамики путем введения сайт-специфичных дисульфидных сшивок в отдельные димеры Н3-Н4 ингибирует передвижение нуклеосом ремоделером SNF2h, увеличивает вытеснение октамера комплексом RSC [172], блокирует

термически индуцированную диффузию нуклеосом вдоль ДНК [223] и даже нарушает компактизацию хроматина белком HP1 [322]. Эти результаты свидетельствуют о том, что октамер гистонов может принимать конформации, альтернативные тем, которые наблюдаются в рентгеновских структурах. Действительно, используя передовые методы крио-ЭМ, Bilokapic et al. недавно наблюдали некоторые альтернативные конформации нуклеосом с развернутой ДНК, наклоненными гистоновыми  $\alpha$ -спиралями и общей сжатой формой NCP [179; 223]. Еще большие внутренние флуктуации в структуре октамера гистонов предполагают эксперименты по химической сшивке лизинов [A1] и эксперименты FRET с высоким разрешением [319].

Эти и другие исследования подчеркнули существование нового уровня динамической сложности нуклеосом, связанного с атомистическими деталями структуры. Однако структурная интерпретация этих результатов все еще остается не вполне ясной. *In silico* подходы, такие как моделирование методом молекулярной динамики (МД), являются мощными инструментами, которые могут дополнять экспериментальные методы и помочь механистически интерпретировать экспериментальные результаты с высоким уровнем детализации. До сих пор моделирование с помощью МД применялось для изучения распределения ионов вокруг нуклеосом и их гидратации [253], деталей разворачивания ДНК, скольжения и разборки нуклеосом [183; 214; 326–329], динамики гистоновых хвостов [A17; 252; 330; 331], эффектов посттрансляционных модификаций гистонов [203; 332; 333] и их вариантов [334], эффектов влияния последовательности ДНК [335] на динамику нуклеосом, взаимодействия между олигонуклеосомами [280] и гистоном H1 [336; 337] и т. д. Из-за вычислительной сложности такого рода расчетов часто прибегают к упрощениям (например, удаляют гистоновые хвосты, используют неявные модели растворителя или крупно-зернистое представление молекулярной системы), либо ограничивают временные рамки моделирования. В настоящее время наиболее известные МД исследования нуклеосом в полноатомном приближении были ограничены несколькими микросекундами времени моделирования [328; 330] даже с применением специализированных суперкомпьютеров, таких как ANTON. Однако предполагается, что важные функциональные переходы в структуре нуклеосом происходят в масштабе времени от микросекунды до миллисекунды [319]. Новые режимы функциональной динамики в этих временных масштабах



все еще ждут подробных структурных исследований с помощью МД-моделирования.

В работе данного раздела, мотивированные различными новыми экспериментальными доказательствами ранее неизвестных способов конформационной динамики и пластичности нуклеосом, мы систематически исследовали равновесную динамику нуклеосом с помощью полноатомного МД-моделирования на масштабе времени 10 микросекунд и более. Насколько нам известно, это первый случай, когда такие длинные динамические траектории анализируются для максимально реалистичных атомистических моделей NCP с полноразмерными гистоновыми хвостами при физиологической температуре и ионной силе. С помощью специально разработанных алгоритмов анализа траекторий (проекция координат в системе отсчета нуклеосом, расчет относительного скручивания ДНК, анализ стабильных контактов и т. д.) мы выявили новые режимы динамики и пластичности нуклеосом. Мы наблюдаем реконфигурацию и разворачивание концов ДНК в нуклеосомах, опосредованных гистоновыми хвостами H3 и H2A, образование дефектов кручения в ДНК и сдвиг ориентационного положения для части коровой ДНК, структурную пластичность ядра гистона в соответствии с недавними экспериментальными наблюдениями. В данном разделе также обсуждается значение наших результатов для понимания ремоделирования нуклеосом и организации структуры хроматина более высокого порядка.

### 2.3.2 Материалы и методы

#### Молекулярно-динамические расчеты

Моделирование проводилось с использованием пакета GROMACS 2018.1 с поддержкой графических ускорителей [338] с силовым полем Amber ff14SB [339], дополненным поправками parmbsc1 [340] для ДНК и поправками ионных параметров CUFIX [341]. Начальные координаты были получены из соответствующих структурных файлов в базе данных PDB (см. таблицу 3). Для моделирования системы NCP<sub>147</sub> конформация хвостов гистонов для одной по-

ловины NCP (цепи E, F, G, H) была скорректирована, чтобы соответствовать конформации симметричных цепей на другой половине NCP (цепи A, B, C, D). Для моделирования с усеченными гистоновыми хвостами хвосты усекались согласно позициям, изображенным на рисунке 2.13д. Сайты усечения были выбраны таким образом, чтобы удалить гибкие части хвостов гистонов, при этом оставляя некоторые части близкие к глобулярному ядру, которые устанавливают стабильные контакты с ДНК или гистонами (например, N-хвосты H3 и H2B, выступающие между супервитками ДНК или H2A C-хвост, взаимодействующий с димером H3-H4). Для моделирования с фиксированными гистоновыми фолдами, C  $\alpha$ -атомы спиралей  $\alpha 1$ ,  $\alpha 2$  и  $\alpha 3$  во всех гистонах удерживались в их исходных положениях с использованием гармонического потенциала с силовой постоянной 1000 кДж/моль/нм<sup>2</sup>. Все системы были помещены в ячейку для моделирования в виде усеченного октаэдра с периодическими граничными условиями, установленными на расстоянии не менее 2 нм от атомов NCP. Следующим шагом в ячейку добавлялись молекулы воды модели TIP3P [342] для сольватации NCP, затем были добавлены ионы Na и Cl, чтобы нейтрализовать заряд и довести ионную силу до 150 mM (Рис. 2.13б). Сконструированные системы были минимизированы по энергии и уравновешены в несколько этапов с постепенным снятием ограничений с тяжелых атомов в течение 1 нс. Температуру поддерживали на уровне 300 К с использованием схемы масштабирования скорости [85], а давление на уровне 1 бар с помощью баростата Парринелло-Рамана [89]. Моделирование проводилось параллельно на суперкомпьютере “Ломоносов-2” [343] с использованием 8 вычислительных узлов, каждый из которых имеет 14 ядер ЦП и один графический процессор NVidia Tesla K40. Моделирование происходило со средней скоростью 60 нс в день.

### Анализ траекторий

Скрипты для анализа были написаны на Python 3 с интеграцией функций GROMACS (предварительная обработка траектории), MDAnalysis (манипуляции с координатами, трехмерное выравнивание) [344], VMD (визуализация) [305] и 3DNA (определение центров пар оснований ДНК, расчет параметров шага пар оснований и пар оснований) [99].

Для анализа геометрии нуклеосом важной концепцией является нуклеосомальная система координат, состоящая из вектора сверхспиральной оси ( $OZ$ , вычисленный из пути ДНК [A17]) и вектора диадной оси псевдосимметрии ( $OY$ , определяемый как перпендикуляр, соединяющий  $OZ$  с центр центральной пары оснований ДНК). Вместе с вектором  $OX$ , определенным как векторное произведение  $OY$  и  $OZ$ , эти три вектора формируют систему координат нуклеосомы (СКН) (Рис. 2.13а). СКН была определена для исходной рентгеновской структуры 1КХ5, все другие структуры и конформации были выровнены с ней путем минимизации среднеквадратичного отклонения (RMSD) между  $C\alpha$ -атомами  $\alpha$ -спиралей гистонового фолда ( $\alpha1$ ,  $\alpha2$ ,  $\alpha3$ ). Используя СКН проекции координат атомов на разные плоскости системы координат, можно визуализировать и вычислить параметр относительного кручения ДНК (локальное кручение) [A7]. Последняя характеристика представляет вращение ДНК относительно поверхности октамера гистонов и полезна для отслеживания вращательного позиционирования ДНК и дефектов кручения ДНК.

Для анализа контактов ДНК-гистоны контакты атом-атом рассчитывали как пары тяжелых атомов (не водородов) с расстоянием менее 4 Å. Также с помощью MDAnalysis были рассчитаны водородные связи и водные мостики.

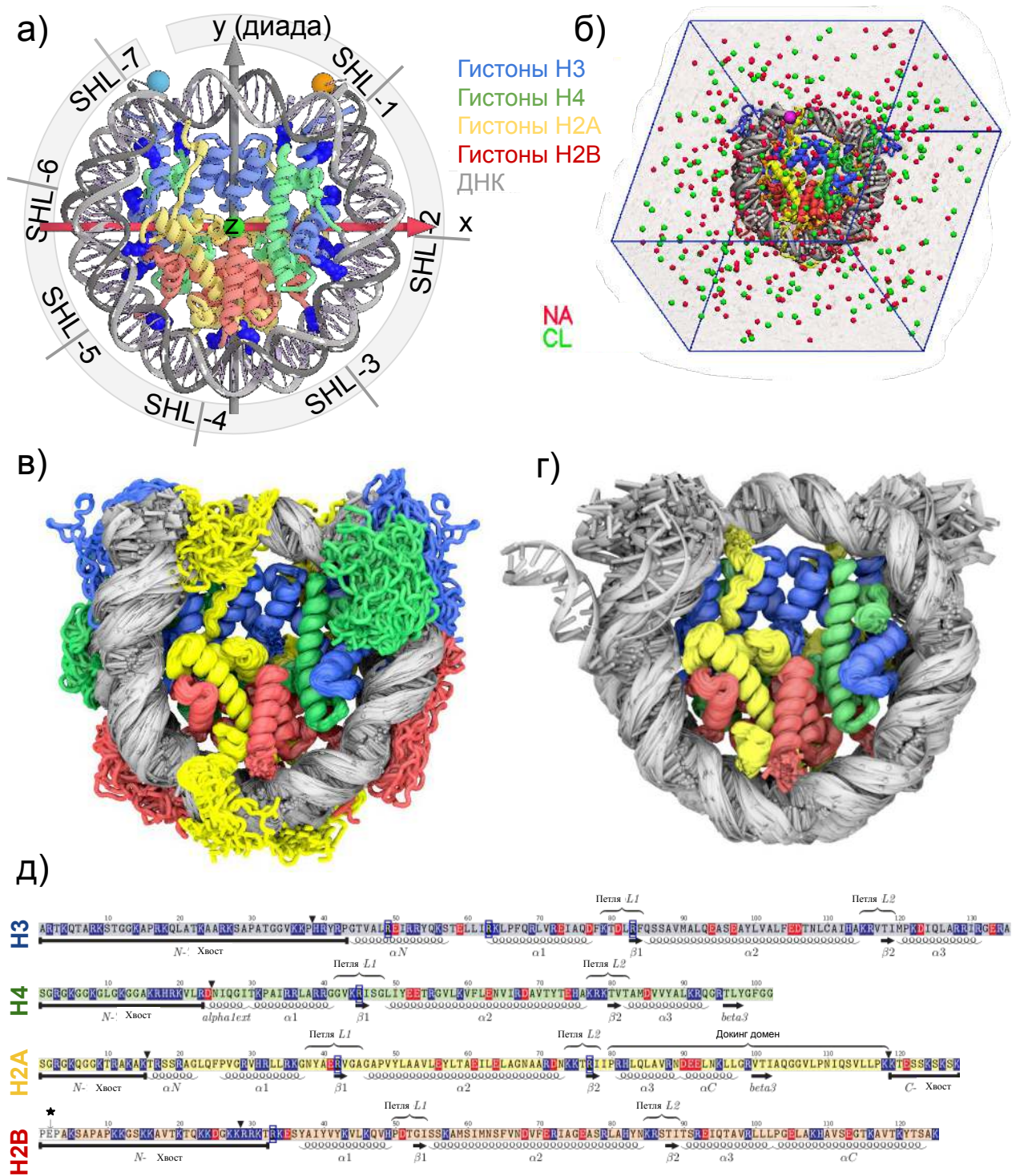
### 2.3.3 Результаты и обсуждение

#### Обзор поведения моделируемых систем

Нами было осуществлено моделирование ряда систем коровых частиц нуклеосом (NCP) на основе различных PDB структур. Список систем приведен в таблице 3. Данные системы включают системы как с полноразмерными гистонами, так и системы с обрезанными хвостами в положениях согласно рисунку 2.13д. Особенностью моделей на основе различных PDB структур является различная длина ДНК, которая укладывается на гистоновых октамер: 145, 146 либо 147 пар оснований. Данные структуры обладают концами ДНК, которые находятся в одинаковых положениях, однако в нуклеосомном коре ДНК испытывает растяжение, либо сжатие на 1-2 нуклеотида в положениях SHL

$\pm 5$  или SHL  $\pm 2$ . Такие области растяжения или сжатия называют дефектами кручения (twist defects). Также нами была промоделирована система с фиксированными C $\alpha$ -атомами  $\alpha 2$ -спиралей гистонов для изучения влияния подвижности гистонов на подвижность ДНК. На рисунках 2.13а,б представлены основные обозначения элементов нуклеосомного кора, а также вид системы в ячейке для моделирования с растворителем и ионами.

Общий вид динамики некоторых систем отражен на рисунке 2.13в,г. Интерактивные траектории доступны по адресу [http://intbio.github.io/nucl\\_MD\\_15mus/](http://intbio.github.io/nucl_MD_15mus/). К важным особенностям динамики систем без хвостов на временах порядка 10-15 микросекунд можно отнести: 1) наблюдение откручивания ДНК от гистонового ядра (такого рода эффекты не наблюдаются на временах в несколько микросекунд), 2) релаксацию дефектов кручения ДНК в районе SHL  $\pm 5$ . К важным особенностям динамики систем с гистоновыми хвостами на временах порядка 10-15 микросекунд можно отнести: 1) наблюдение откручивания ДНК от гистонового ядра (в меньших масштабах, чем в системах без гистоновых хвостов), 2) переключение хвостов между различными положениями на ДНК, 3) влияние конформации хвостов на моды откручивания ДНК. Ниже мы подробнее рассмотрим некоторые полученные результаты.



**Рисунок 2.13** — Системы нуклеосом моделируемые на временах более 10 мкс. а) Ключевые оси нуклеосомной системы отсчета. б) Вид моделируемой системы в ячейке с растворителем и ионами. в) Наложение кадров системы  $NCP_{147}$  каждые 100 нс. г) Наложение кадров  $NCP_{145}^{nt}$  каждые 100 нс. д) Последовательность коровых гистонов и особенности их вторичной структуры ( $\alpha$ -спирали,  $\beta$ -листы, петли, гибкие хвосты гистонов).  $\blacktriangledown$  отмечает позиции обрезания хвостов в модели с усеченными хвостами; аргинины, боковые цепи которых вставлены малые бороздки ДНК отмечены синими рамками; синим отмечены - положительно заряженные остатки, красным - отрицательно заряженные остатки. \* отмечает PER-конец H2B, который отсутствует в рекомбинантном варианте белка.

## Механизмы отворачивания и приворачивания ДНК

Для анализа отворачивания ДНК от гистонового октамера нами был разработан подход определения количества отвернутых пар нуклеотидов ДНК для каждого кадра траектории. Для каждой пары нуклеотидов вычислялось ее ближайшее расстояние до какой-либо пары нуклеотидов кристаллической структуры. Количество отвернутых пар нуклеотидов с каждого конца определялось как максимальная длина сегмента ДНК от конца ДНК в котором расстояние для всех нуклеотидных пар от любых кристаллических положений было более 7 Å. На основе данной меры были построены графики зависимости величины отворота ДНК от времени (Рис. 2.14, Рис. 2.15а,б и интерактивные материалы [http://intbio.github.io/nucl\\_MD\\_15mus/](http://intbio.github.io/nucl_MD_15mus/)).

Для систем без гистоновых хвостов ДНК сохраняла исходное привернутое положение в течение начального времени моделирования. Для разных расчетов и сторон нуклеосомы оно колебалось от 0.8 мкс до 4, 6 и более 7 мкс (см. Рис. 2.14). Для системы с гистоновыми хвостами отворачивание началось на 4-ой микросекунде расчетов. Величина отворота ДНК составляла порядка 12-13 пар нуклеотидов, а в случае систем без хвостов иногда достигала 28 пар нуклеотидов. Рассмотрим в начале механизмы отворота ДНК в системе без хвостов. Анализ стабильных контактов гистонов и нуклеосом показывает, что конец ДНК в нуклеосомном коре формирует плотные контакты и удерживается взаимодействиями с аминокислотными остатками Н3 гистона, в особенности Н3У41, Н3R42, Н3Т45, которые находятся в области, которую мы назовем областью “защелки” концов ДНК 2.19. Плотные взаимодействия в этой области удерживают конец ДНК в NCP. При отвороте ДНК данные взаимодействия нарушаются и ДНК переходит во флуктуирующее состояние, в котором величина отворота быстро меняется в наносекундном диапазоне (см. Рис. 2.14, состояние 2). В некоторых случаях плотные контакты застежки с ДНК восстанавливаются и система переходит в кристаллоподобное состояние (например в системе NCP<sub>145</sub><sup>nt</sup> на третьей наносекунде моделирования). Однако может произойти реконфигурация области защелки в гистоне Н3 в состояние, не способное к плотному связыванию ДНК. Это приводит к тому, что флуктуационное состояние становится более долгоживущим и сопровождается также флуктуациями области Н3 хвоста (см. Рис. 2.14, состояние 2b). Это флукту-

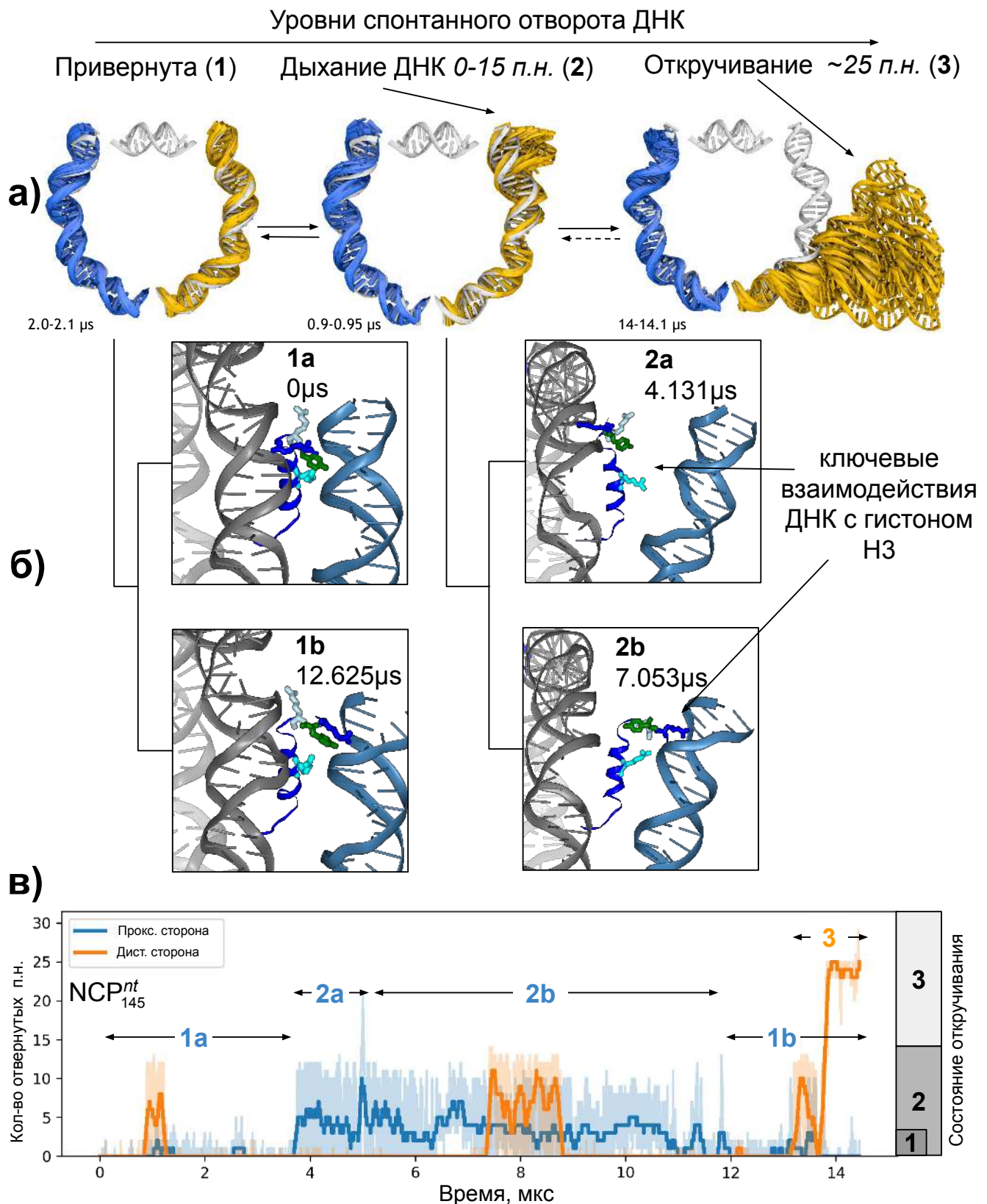
ационное состояние может разрешаться в кристаллоподобное состояние ДНК, когда остатки НЗ смогут сформировать плотные контакты с флуктуирующим концом ДНК. Более детальный визуальный анализ взаимодействия гистона НЗ с концом ДНК показал, что стабильное кристаллоподобное состояние ДНК возникает, когда ряд остатков входят в малую бороздку ДНК. В кристаллической структуре это НЗR49, НЗY41, НЗH39. Интересно, что в нашем моделировании в ряде систем ( $NCP_{145}^{nt}$ ,  $NCP_{145}^{nt}$ ) кристаллоподобное состояние ДНК также реализовывалось в состоянии, когда вместо НЗH39 в малую бороздку заходил НЗK40, который в кристалле взаимодействует с малой бороздкой другого супервитка (см. Рис. 2.14, состояние **1b**). Дальнейшему откручиванию ДНК препятствовал стабильный сайт связывания с гистонов Н2А в районе позиции -59, с которой контактируют остатки Н2AK75, Н2AT76, Н2AR77 из L2 петли. На 14-ой микросекунде в системе  $NCP_{145}^{nt}$  наблюдался дальнейший отворот до уровне 26-28 отвернутых пар оснований с диссоциацией контактов данного сайта (см. Рис. 2.14, состояние **3**). Данный факт требует дальнейшего исследования. Масштаб флуктуаций ДНК в системе без хвостов представлен на рисунках 2.13в,г. Важной особенностью является то, что ДНК откручивается не только в плоскости нуклеосомного диска, но и имеет компонент смещения вдоль суперспиральной оси нуклеосомы. Более подробно взаимоотношение смещения конца ДНК вдоль оси X и Z представлено на рисунке 2.13ж. Существенной корреляции между смещениями не наблюдается, однако очевидно, что при больших параметрах отворота наблюдается в основном состояния с некоторым компонентом смещения вдоль оси суперспирали в направлении противоположное от внутреннего супервитка ДНК. Данный факт можно объяснить стерическим и электростатическим отталкиванием витков ДНК.

В отличие от систем без гистоновых хвостов, система с хвостами ведет себя несколько по другому (Рис. 2.13б,д,е). Мы также наблюдали откручивание до 12-14 пар оснований с конца ДНК, однако существуют существенные отличия. Во-первых, величина отклонения открученных пар нуклеотидов от их кристаллического пути была меньше из-за того, что гистоновые хвосты оказывали удерживающее воздействие на конформацию ДНК и ее возможность отклоняться в различных направлениях. Во-вторых, флуктуации ДНК между состояниями с различной степенью откручивания были менее выражены по амплитуде, опять же из-за воздействия гистоновых хвостов, главным образом N-хвоста гистона НЗ и С-конца гистона Н2А. Наличие гистоновых хвостов

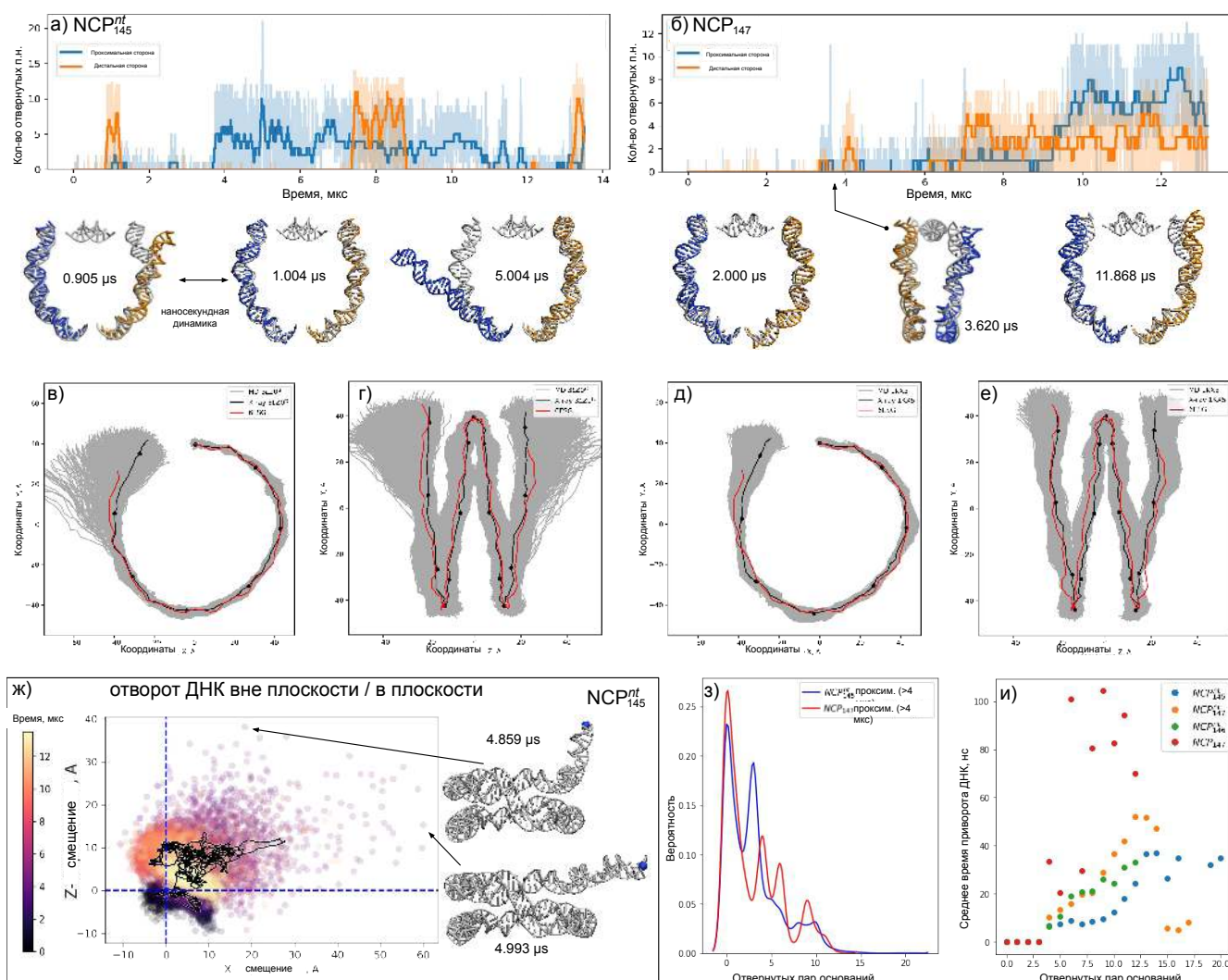
приводит к расщеплению динамического состояния отворота-приворота 12 концевых нуклеотидов, наблюдаемого в системе без хвостов, на ряд подсостояний определяемых конформацией гистоновых хвостов. Рисунок 2.13 и иллюстрирует среднее время приворота ДНК из состояний с заданной величиной отворота. Из рисунка видно, что наличие гистоновых хвостов серьезным образом влияет на увеличение времени приворота ДНК. Для величины отворота в 10 пн время возрастает с 20-40 нс до  $\sim 100$  нс. Гистоновые хвосты также существенным образом влияют на механизм отворачивания ДНК. Мы наблюдали, что область “защелки” НЗ гистона также, как и в случае системы без хвостов, теряет контакты с сегментом ДНК вблизи входа-выхода, однако, потери этих контактов способствует внедрение N-конца гистона НЗ между супервитками ДНК. Хвост НЗ с одной стороны образует дополнительные контакты с ДНК, но с другой стороны способствует замещению стабильных контактов с областью НЗ-“защелки”. Этот же эффект приводит к тому, что в отличие от системы без гистоновых хвостов, хвост НЗ может препятствовать быстрому привороту ДНК из отвернутого положения в повернутое.

Более подробно динамика взаимодействия ДНК с НЗ-“защелкой” и взаимодействие хвостов гистонов с концами нуклеосомной ДНК представлена на рисунке 2.16.

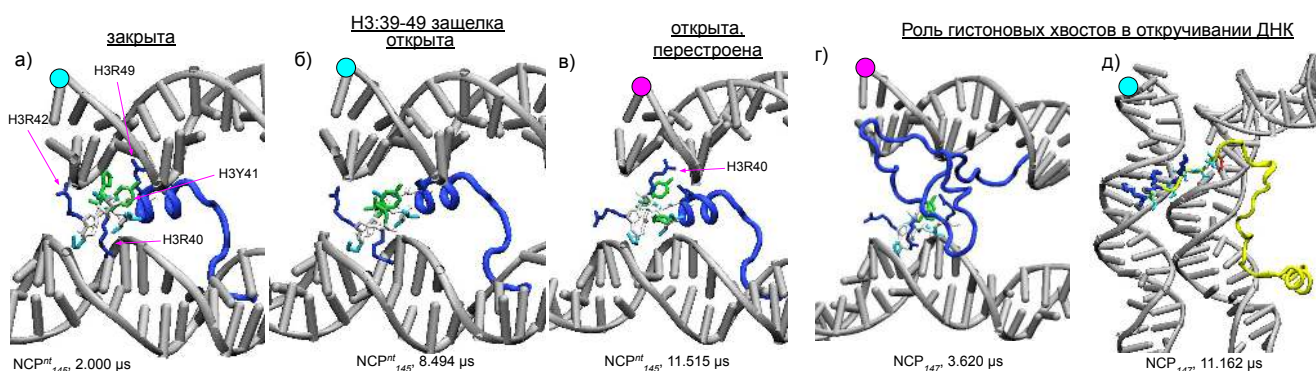




**Рисунок 2.14** — Динамика отворота ДНК в нуклеосоме с укороченными гистоновыми хвостами (наблюдаемая в системе  $NCP_{145}^{nt}$ ). Количество пар нуклеотидов отвернутых от октамера гистонов как функция времени моделирования представлена снизу. Критерий отворота ДНК – смещение центра пары оснований более, чем на 7 Å от положений оснований в исходной структуре. Наблюдается три состояния 1-3 (изображены сверху). Для состояний 1 и 2 наблюдаются подсостояния, связанные с реориентацией Н3  $\alpha$ N спирали и области Н3-“застежки” (средний ряд изображений).

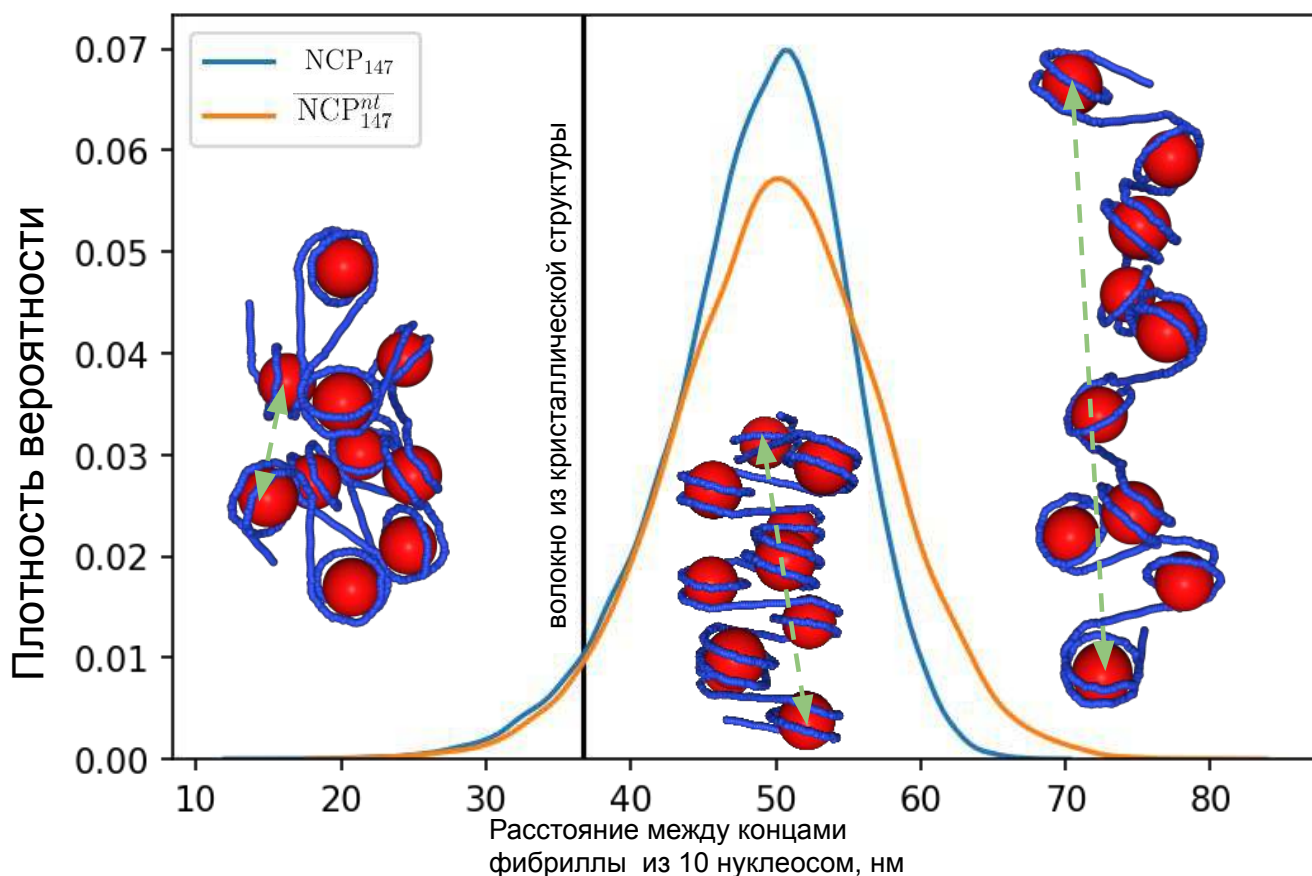


**Рисунок 2.15** — Сравнение характеристик отворота ДНК в нуклеосомах с и без гистоновых хвостов. а)-б) отворот ДНК с течением времени для NCP<sup>nt</sup><sub>145</sub> и NCP<sub>147</sub>, определенных как смещение центра пары оснований более, чем на 7 Å от положений оснований в исходной структуре. в-е) Проекции центров пар оснований ДНК на различные плоскости СКН. ж) Смещение проксимального конца ДНК в ходе МД. з) Распределений вероятности отворота ДНК. и) Среднее время приворота ДНК из состояния с определенным отворотом.



**Рисунок 2.16** — Динамика “защелки” концов ДНК фрагментом гистона НЗ

Каковым может быть функциональное значение отворачивания ДНК концов от гистонового кора? С одной стороны предполагается, что отворачивание ДНК необходимо для связывания различных факторов транскрипции [137]. В нашем моделировании для системы с гистоновыми хвостами ДНК не достигала такой степени отворота. Однако плотное взаимодействие хвостов с ДНК скорее всего распространится и на более существенные степени отворота ДНК. Таким образом, мы предполагаем, что важным фактором для определения доступности ДНК для транскрипционных факторов будет не только отворот ДНК, но и взаимодействия ДНК с гистоновыми хвостами. Другим возможным эффектом отворачивания ДНК является влияние изменение конформации нуклеосомальной ДНК на возможную укладку нуклеосом на супрануклеосомном уровне. Для оценки данных эффектов мы провели моделирование, в котором генерировали фибриллы из 10 нуклеосом, соединяя прямыми отрезками линкерной ДНК длиной 17 п.н. случайно выбранные кадры из траекторий. Таким образом, можно было получить ансамбль конформаций нуклеосомных фибрилл и оценить распределение длин между концами фибрилл (Рис. 2.17). По сравнению с фибриллами построенными на основе кристаллической структуры, фибриллы построенные с учетом флуктуаций концов коровой ДНК демонстрировали в среднем расстояние между концами на 10-15 нм больше. Ширина самого распределения составляла около 20 Å. Таким образом, можно сделать вывод о том, что динамика концов ДНК, во-первых, “разрыхляет” супрануклеосомную структуру хроматина, а, во-вторых, обеспечивает достаточную конформационную гибкость фибриллы в пределах тепловых колебаний ( $\sim kT$ ).



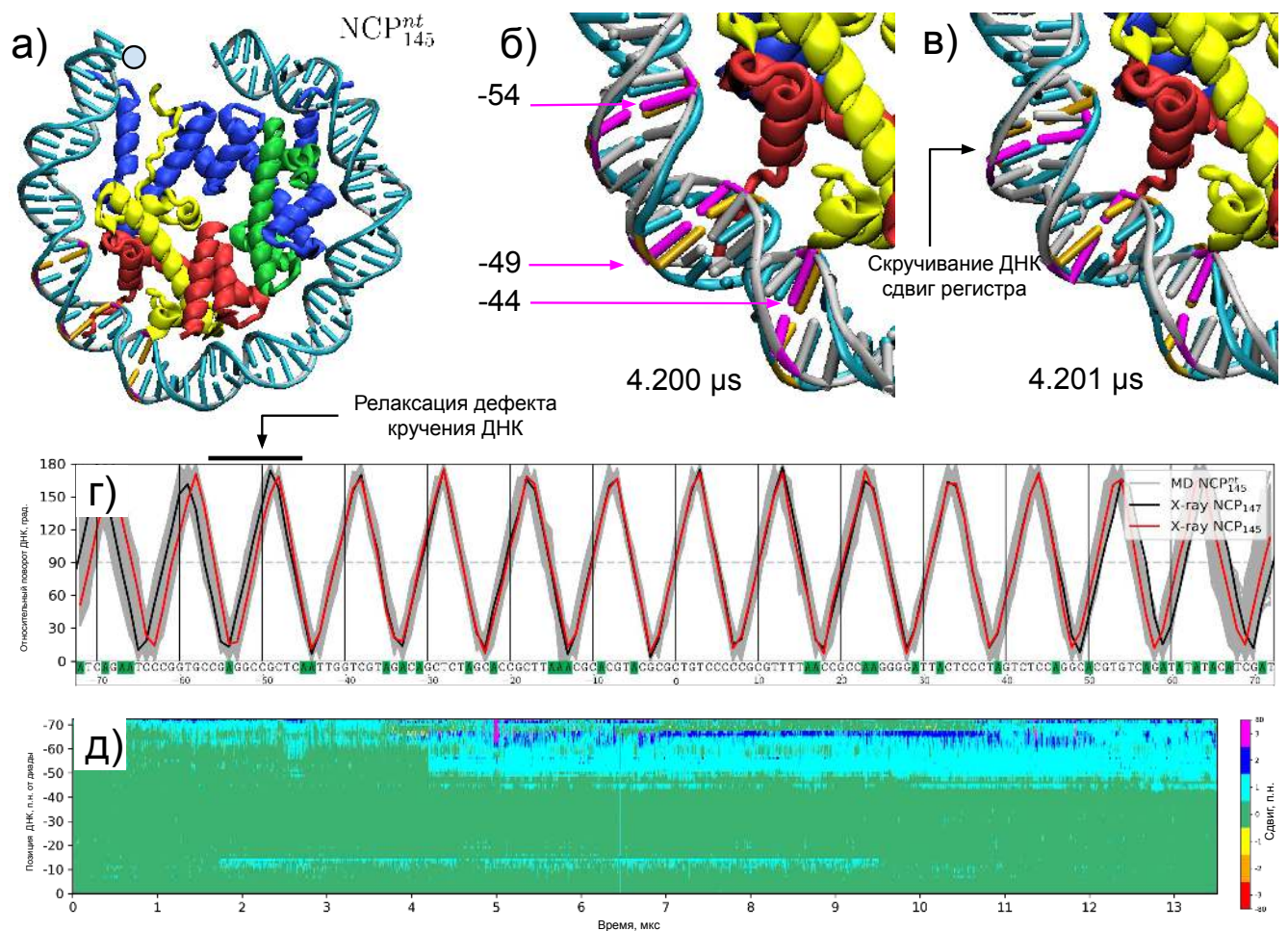
**Рисунок 2.17** — Влияние динамики ДНК на структуру хроматиновых фибрилл

### Динамика кручения ДНК в нуклеосоме

Интересным фактом в ходе моделирования явилось наблюдение релаксации дефекта кручения ДНК внутри нуклеосомного кора в системе  $NCP_{147}^{nt}$ . Дефект кручения ДНК в этой системе, основанной на 601-ой последовательности ДНК, располагается в районе  $SHL \pm 5$ . По сравнению с нуклеосомами на основе альфа-сателлитной ДНК, ДНК в районе этого положения растянута на 1 нуклеотид, при этом совершает то же самое количество оборотов (см. рисунок 2.18г). В ходе моделирования мы наблюдали резкий перескок регистра ДНК из кристаллоподобного состояния (Рис. 2.18б) в состояние, где нуклеотиды начиная с нуклеотида номер -54 сместились в сторону диады нуклеосомы. Рисунок 2.18д иллюстрирует то, что данный сдвиг регистра ДНК далее произошел для всего конца ДНК, начиная приблизительно с нуклеотида -50. Таким

образом около 20 нуклеотидов с проксимального конца ДНК испытали вращательное движение внутри нуклеосомного кора. Любопытным является изучение деталей данного механизма. Рассмотрение динамики отворачивание ДНК свидетельствует о том, что релаксация дефекта кручения произошла в области, которая не была затронута отворотом ДНК (Рис. 2.15а). В то же время нельзя исключить, что отворот концов ДНК способствовал релаксации дефекта и его распространению вдоль ДНК. С точки зрения стальных контактов в системе  $NCP_{145}^{nt}$  в районе SHL -6 в течение первой микросекунды моделирования наблюдался стабильный контакт нижней цепи ДНК в положении -58 с H2AT76 и верхней цепи ДНК в положении -54 с H2BS56. При анализе 10 микросекунды моделирования на предмет стабильных контактов между нуклеотидами и аминокислотами оказалось, что первый контакт остался неизменным, а второй контакт сместился на нуклеотид за номером -55. Таким образом, релаксация данного дефекта кручения была в некотором роде асимметрична. В большей степени наблюдалось сдвиг регистра контактов для верхней цепи ДНК. Интересным фактом является то, что как недавно было показано в структурных работах по изучению механизмов работ ремоделеров нуклеосом семейства ISWI и SWI/SNF, именно, верхняя цепь ДНК изменяет свое положение на нуклеосомном коре в первую очередь при связывании АТФ субъединицы ремоделера [321].

Еще одним интересным фактом сопровождающим релаксацию дефекта кручения ДНК в районе SHL -5 является деформация в этот момент спирали  $\alpha 2$  H2A гистона с отгибом ее С-конца в сторону центра нуклеосомного кора. С-конец данной спирали связан с сайтом связывания ДНК в районе SHL -5.5 и, вероятно, участвует в ослаблении контактов ДНК с гистоновым кором, которое помогает релаксации дефекта кручения.

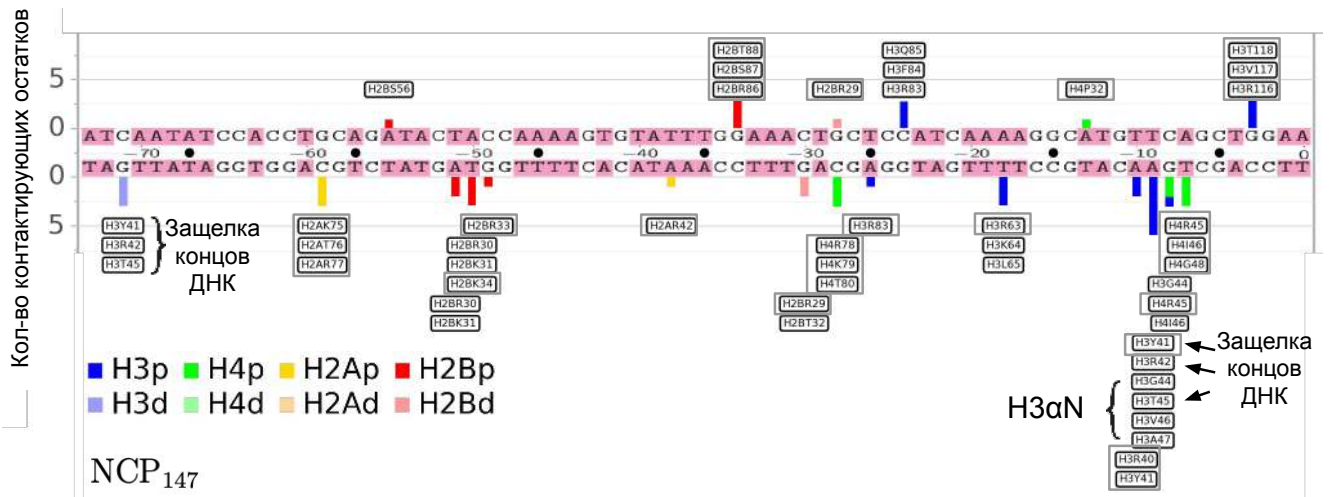


**Рисунок 2.18** — Динамика кручения ДНК в нуклеосоме. а) Структура проксимального супервитка нуклеосомы в начальный момент. Кристаллическая форма ДНК изображена голубым цветом, серым цветом ДНК в ходе МД моделирования. Оранжевым цветом отмечены пары нуклеотидов, которые участвуют изменяют положение на нуклеосоме. б) Кадр траектории до прокручивания ДНК. в) Кадр траектории после прокручивания ДНК. г) График параметра относительного кручения (relative twist) ДНК на нуклеосоме. д) Иллюстрация сдвига в положении нуклеотидов для верхней цепи ДНК относительно положения в кристаллической структуре.

### Динамика контактов ДНК-белок в нуклеосоме

Наличие траекторий молекулярной динамики позволяет достаточно точно описать взаимодействия гистонов с ДНК на уровне контактов отдельных атомов, аминокислот, нуклеотидов. Причем в отличие от анализа кристалли-

ческих структуру имеется возможность оценить устойчивость и динамическую подвижность тех или иных контактов. Нами был проведен анализ стабильных контактов между аминокислотными остатками гистонов и нуклеотидами ДНК. Для изначального анализа была выбрана система NCP<sub>147</sub> в течение первой микросекунды моделирования, когда отворота ДНК не наблюдалась. Также поскольку система является псевдосимметричной мы учитывали только те контакты, которые присутствовали с обеих сторон нуклеосомы. Результат анализа приведена на рисунке 2.19. Данный график иллюстрирует все основные особенности связывания ДНК с гистонами. Обратим внимание на две особенности, которые по нашему мнению имеют функциональное значение. Количество контактов, которые образует верхняя цепь ДНК значительно меньше, контактов которые образует нижняя цепь ДНК. Предполагаем, что это как раз может иметь значение при передвижении нуклеосом вдоль ДНК АТФ-зависимыми ремоделерами, которые как было показано недавно в первую очередь вытягивают и деформируют верхнюю цепь ДНК. Второй интересный момент связан с наличием области Н3 гистона, которую мы называем “защелкой”, которая взаимодействует как с концом коровой ДНК, так и с областью вблизи диады - около положения -9. Отметим, что вблизи положения -9 в геномных исследованиях по позиционированию нуклеосом наблюдается необычный сигнал - там чаще встречаются нуклеотиды А/Т, хотя по общему правилу нуклеосомы предпочитают в области изгиба ДНК в сторону большой бороздки нуклеотиды G/C [289]. Данный факт может отчасти объясняться плотным взаимодействием Н3R40 внутри малой бороздки ДНК с электроотрицательными группами А/Т оснований.



**Рисунок 2.19** — Стабильные контакты ДНК-гистоны в нуклеосоме. Отмечены контакты между нуклеотидами и аминокислотными остатками, присутствующие в 90% кадров в системе NCP<sub>147</sub> в течение первой микросекунды моделирования с обеих симметричных сторон нуклеосомы. В рамках отмечены контакты, которые сохраняются с обеих сторон нуклеосомы в течение всей МД траектории.

## Пластичность гистонового октамера

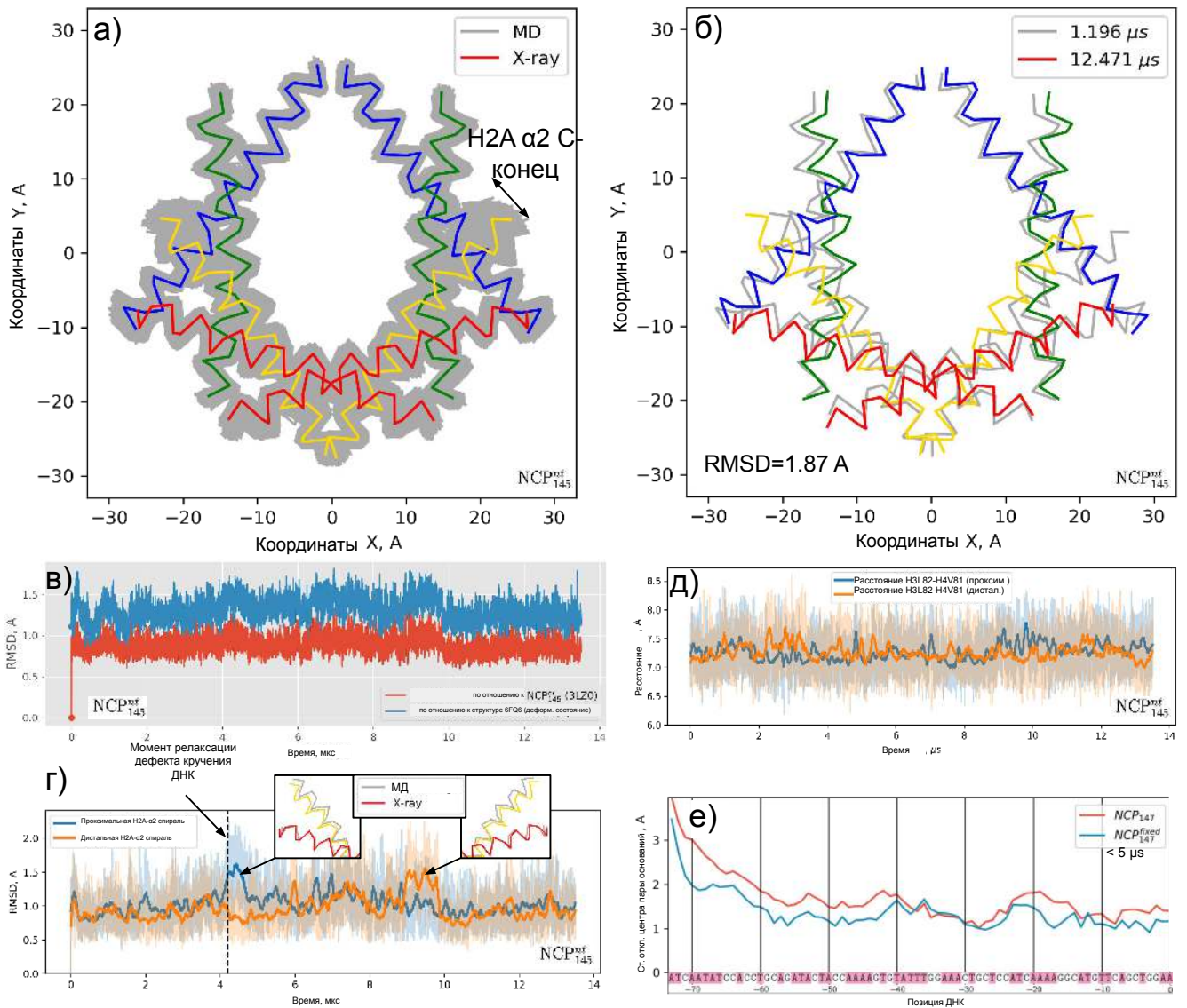
Как было описано во введении весьма актуальной темой научных исследований на данный момент является изучение пластичности октамера гистонов. Для исследования данного вопроса нами были построены проекции координат атомов  $\alpha 2$ -спиралей гистонов на плоскость нуклеосомного диска (Рис. 2.20). Видно, что спирали испытывают флуктуации в пределах от нескольких до 5-6Å. Особенно выделяются флуктуации концов С-концов спиралей гистона H2A. Недавно методами крио-ЭМ был разрешен ряд деформированных структур нуклеосомного кора, в частности структура нуклеосомы сжатая вдоль диадной оси на несколько процентов. RMSD кадров МД траектории относительно данной структуры (Рис. 2.20в) показывает, что данная структура зачастую отличается от кадров траектории не более, чем начальная структура самой траектории. Таким образом, можно сделать вывод, что наблюдаемые в эксперименте деформированные состояния нуклеосом находятся, по крайней мере, в рамках



отличий, которые мы видим между кадрами траектории на масштабе мультимикросекундного моделирования.

В ряде работ было показано, что дисульфидные сшивки внутри димеров гистонов приводят к изменению функционирования нуклеосом. В частности H3L82C-H4V81C уменьшает термическую диффузию нуклеосом по ДНК и влияет на работу ремоделера SNF2h [172; 179; 223]. Расстояние между C- $\alpha$  атомами данные остатков изображена на рисунке 2.20г. Оно колеблется в пределах от 6,1-8,5Å. Длина дисульфидной связи составляет около 2,05Å, расстояние между C- $\alpha$ -атомом и атомом серы в стандартной конформации цистеина составляет 2,81Å и не меняется при вращении торсионных углов. Таким образом максимальное расстояния между C- $\alpha$ -атомами у остатков цистеина, связанных дисульфидной связью, составляет 7,67Å. Следуя данным расчетам, видим, что дисульфидная сшивка будет препятствовать некоторым конформациям, наблюдаемым в динамике.

Для того, чтобы дополнительно изучить влияние пластичности октамера на динамику ДНК, мы провели расчеты, в которых  $\alpha$ 2-спирали гистонов были зафиксированы. Анализ флуктуационной динамики ДНК показал (см. рисунок 2.20е), что у системы с фиксированными  $\alpha$ -спиралями подвижность ДНК была снижена на всем ее протяжении. Данные эффект требует дальнейшего изучения, однако можно предположить, что уменьшение подвижности ДНК отрицательно сказывается на возможности ремоделирования и скольжения нуклеосом.

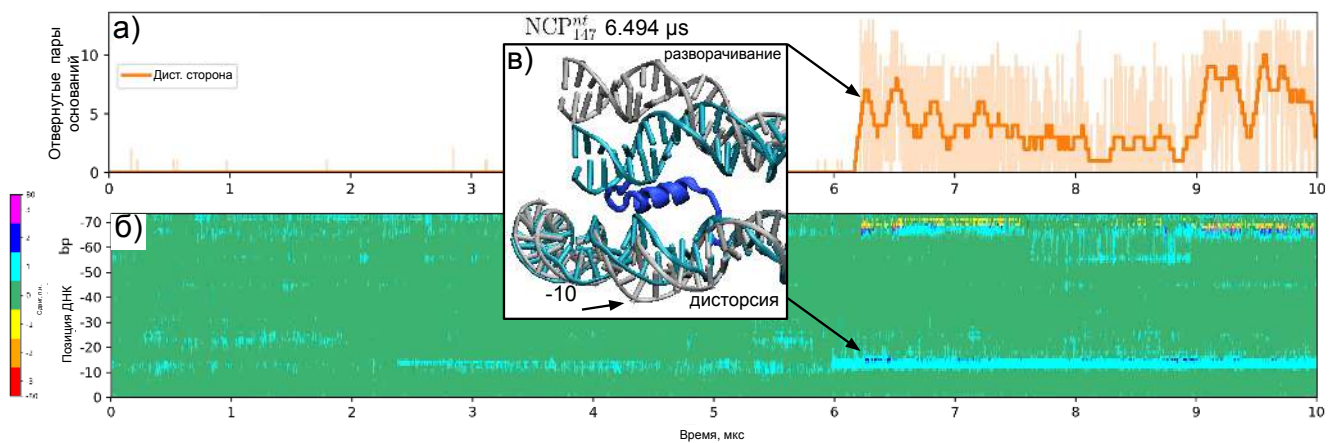


**Рисунок 2.20** — Пластичность гистонного октамера. а) Наложение проекций координат  $S\alpha$ -атомов  $\alpha 2$ -спиралей гистонов в плоскости нуклеосомального диска из кадров МД траектории. б) Аналогичные проекции координат  $S\alpha$ -атомов для двух кадров с максимально различным RMSD. в) RMSD  $\alpha 2$ -спиралей гистонов в ходе МД относительно исходной структуры и структуры 6FQ6, полученной крио-ЭМ, с деформированной конформацией ядра нуклеосомы. г) Эволюция RMSD спирали  $\alpha 2$  гистона H2A. д) Расстояние между остатками, которые подвергаются дисульфидным сшивкам в ряде экспериментов. е) Сравнение флуктуаций пар оснований ДНК в системах с закрепленными и незакрепленными  $\alpha 2$ -спиралями гистонов.

## Аллостерическая связь откручивания ДНК и конформации ДНК около диады

Область “застежки” гистона H3 взаимодействует одновременно с двумя супервитками ДНК. Это наблюдение известно достаточно давно. Например, Horn et al. в своих экспериментах наблюдали, что SIN мутация гистона H4R25C взаимодействующая в положении SHL  $\pm 0.5$  приводит к ухудшению возможности компактизации нуклеосомальных фибрилл. Они предположили, что одним из возможных механизмов является дестабилизация концов нуклеосомальной ДНК в виду коммуникации области SHL  $\pm 0.5$  с областями SHL  $\pm 6.5$  посредством H3  $\alpha$ N-спирали. Однако, Flauss et al. предположили, что этот эффект может объясняться увеличением мобильности нуклеосом и их смещению относительно исходных позиций [345]. Тем не менее, нам было интересно рассмотреть в нашем моделировании возможные эффекты коммуникации областей SHL  $\pm 0.5$  и SHL  $\pm 6.5$ . На рисунке 2.21 приведен сравнительный анализ откручивания ДНК и деформации ДНК в различных положениях от времени для системы NCP<sub>147</sub>. Любопытным является факт, что вслед за откручиванием ДНК, зачистую область “защелки” гистона H3 теряет свои контакты не только с концом ДНК, но и с областью нижнего супервитка. Такой случай изображен на панели 2.21в. Потеря контактов с нижним супервитком ДНК в свою очередь ведет к деформации ДНК в этой области и к ее усиленным флуктуациям. Можно предположить, что наличие флуктуация в нижнем супервитке будет способствовать прохождению через нуклеосому дефектов скручивания ДНК.

В результате анализа приведенного в данном разделе нами выдвигается предположение о наличие динамической (аллостерической) связи между откручиванием ДНК и скольжению нуклеосом вдоль ДНК по механизму продвижения дефектов кручения. Откручивание концов коровой ДНК, во-первых, помогает образованию и прохождению дефектов скручивания вблизи концов ДНК, а, во-вторых, способствует дестабилизации ДНК в районе диады, что также, вероятно, помогает прохождению дефектов скручивания далее через нуклеосомную ДНК.



**Рисунок 2.21** — Аллостерическая связь откручивания ДНК и конформации ДНК около диады. а) График откручивание ДНК от времени в ходе МД моделирования. б) График смещения нуклеотидов относительно кристаллической структуры.

### 2.3.4 Благодарности

Работа данного раздела поддержана грантами Российского научного фонда № 18-74-10006 (атомистическое моделирование нуклеосом и разработка и анализ траекторий нуклеосом), грантами Российского фонда фундаментальных исследований № 20-34-70039 (моделирование структуры супрануклеосомальных фибрилл) и № 19-34-51053 (метода анализа взаимодействий белков и нуклеиновых кислот). Исследования проводились на оборудовании коллективных исследовательских установок вычислительных ресурсов высокопроизводительных вычислений МГУ им. М.В. Ломоносова.

## 2.4 Выводы главы 2

В данной главе на примере моделирования нуклеосом продемонстрированы возможности современных технологий суперкомпьютерного моделирования динамики биомолекулярных систем в атомистическом приближении на микросекундных временных масштабах. Проанализирована конформационная

подвижность как коровых частиц нуклеосом (октамер гистонов + 147 пар нуклеотидов ДНК), так и нуклеосом с линкерной ДНК (октамер гистонов + 187 пар нуклеотидов ДНК). Изучена динамика хвостов и внутренняя подвижность глобулярных доменов гистонов, связь динамики белка и ДНК. Впервые в равновесной молекулярной динамике наблюдалась частичная диссоциация и реассоциация (отворот) ДНК от белкового октамера, изменение ориентационного положения ДНК, выпячивание ДНК вблизи входа-выхода в нуклеосому. На основании моделирования могут быть сформулированы следующие научные выводы и гипотезы:

- Электростатическое отталкивание двух сегментов линкерной ДНК является одним из факторов, которые влияют на конформацию ДНК и углы входа-выхода ДНК из ядра нуклеосомы.
- В процессе спонтанного ступенчатого откручивания/прикручивания ДНК в нуклеосоме последовательно теряются стабильные взаимодействия ДНК с гистонами. Конформация ДНК подвержена значительным флуктуациям в наносекундном диапазоне. В переходе нуклеосом из кристаллоподобного состояния в состояние с отвернутой ДНК ключевую роль играют взаимодействия остатков гистона Н3 между двумя супервитками ДНК.
- Формирование дефектов кручения ДНК в нуклеосоме и конформационные перестройки внутри глобулярных частей гистонового октамера, происходящие на микросекундных временах, указывают на возможность скольжения нуклеосом вдоль ДНК по винтовому механизму и транслокацию ДНК вдоль октамера.
- Диссоциации концов нуклеосомальной ДНК от гистонового октамера наблюдается на временах порядка 10 мкс.

**Таблица 3** — Список моделируемых систем

Имя	PDB ID <sup>a</sup>	Описание	Время, мкс
$NCP_{147}$	1KX5	147 п.н. псевдосимметричной $\alpha$ -сателлитной ДНК, гистоны полной длины в симметричном стартовом состоянии	13.0
$NCP_{147}^{fixed}$	1KX5	Такая же как $NCP_{147}$ , но $C\alpha$ -атомы гистоновых фолдов ограничены в движении	5.0
$NCP_{147}^{nt}$	1KX5	Такая же как $NCP_{147}$ , но гистоновых хвосты обрезаны <sup>b</sup>	10.0
$NCP_{146}^{nt}$	1AOI	146 п.н. $\alpha$ -сателлитная ДНК, обрезанные гистоновые хвосты <sup>b</sup>	7.0
$NCP_{145}^{nt}$	3LZ0	145 п.н. 601-ой последовательности ДНК, обрезанные гистоновые хвосты <sup>b</sup>	13.5

<sup>a</sup> Идентификатор структуры, использованной для создания системы, из базы данных PDB.

<sup>b</sup> Сайты обрезки хвостов определены на рис. 2.13д.

### Глава 3. Интегративное моделирование с применением данных экспериментов по футпринтингу ДНК<sup>1</sup>

В данной главе обсуждаются подходы по моделированию ДНК-белковых комплексов, основанные на использовании экспериментальных данных по футпринтингу ДНК гидроксильными радикалами. Излагаются основы метода футпринтинга ДНК. Описываются разработанные подходы по численному анализу экспериментальных данных и моделированию результатов футпринтинга на основе молекулярных моделей (реализованные в программном пакете HYDROID), которые позволяют использовать данные футпринтинга при построении молекулярных моделей. На основании разработанных подходов изучена конформация ДНК в нуклеосомах, разработан метод, позволяющий с точностью до одного нуклеотида определить позицию последовательности ДНК на нуклеосомном коре. Применение разработанных подходов интегративного моделирования продемонстрировано на примере создания структурной модели центримерной нуклеосомы пекарских дрожжей. Материалы данной главы основаны на следующих статьях [A4; A7; A11] и тезисах [A38; A51].

#### 3.1 Методы количественной обработки и структурной интерпретации экспериментов по расщеплению ДНК гидроксил радикалами с помощью программы HYDROID

*Содержание данной главы следует статье [A4].*

Метод гидроксильного футпринтинга (hydroxyl-radical footprinting, HRF) является мощным методом исследования структур комплексов нуклеиновых кислот и белков в растворе с разрешением до одного нуклеотида. Чтобы использовать весь количественный потенциал HRF, мы описываем протокол для количественной оценки данных HRF и интеграции этих данных с атомистическими структурными моделями. Наш протокол, HYDroxyl-Radical

---

<sup>1</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [A4; A7; A11].

Footprinting Interpretation for DNA (HYDROID) позволяет - анализировать данные HRF, полученные с помощью гель-электрофореза, - оценивать теоретические профили расщепления ДНК, полученные из молекулярных моделей, - визуализировать и сравнивать экспериментальные и теоретические профили расщепления. HYDROID полагается на новые алгоритмы аппроксимации данных и надежные подходы к обработке молекулярных моделей. HYDROID реализован на Python и обладает как возможность вызова из скриптов, так графического интерфейса. Этапы протокола HYDROID включают извлечение профилей дорожек из изображений геля, количественную оценку частоты расщепления ДНК по каждому нуклеотиду, теоретическую оценку частоты расщепления ДНК из атомистических структурных моделей с последующим сравнением экспериментальных и теоретических результатов. Примеры сценариев для каждого шага анализа и интерпретации данных HRF представлены для нескольких нуклеосомных систем; их можно легко адаптировать для анализа пользовательских данных. В качестве входного сигнала для HYDROID требуются изображения электрофоретических гелей продуктов реакции расщепления ДНК гидроксильными радикалами в полиакриламидном геле. Также может использоваться молекулярная модель комплекса ДНК-белок. Протокол HYDROID можно использовать для количественной оценки HRF на сегментах ДНК до 100 нуклеотидов на изображение геля. Кроме того протокол можно применять для анализа комплексов РНК-белок и свободных молекул РНК или ДНК в растворе. По сравнению с другими методами, HYDROID уникален своей способностью одновременно интегрировать данные HRF с анализом атомистических структурных моделей. HYDROID находится в свободном доступе по адресу <https://github.com/ncbi/HYDROID>. Выполнение полного протокола занимает примерно ~ 3 часа. Пользователи должны быть знакомы с интерфейсом командной строки, языком сценариев Python и форматами файлов PDB. Также доступен графический пользовательский интерфейс с базовыми функциями (HYDROID\_GUI).

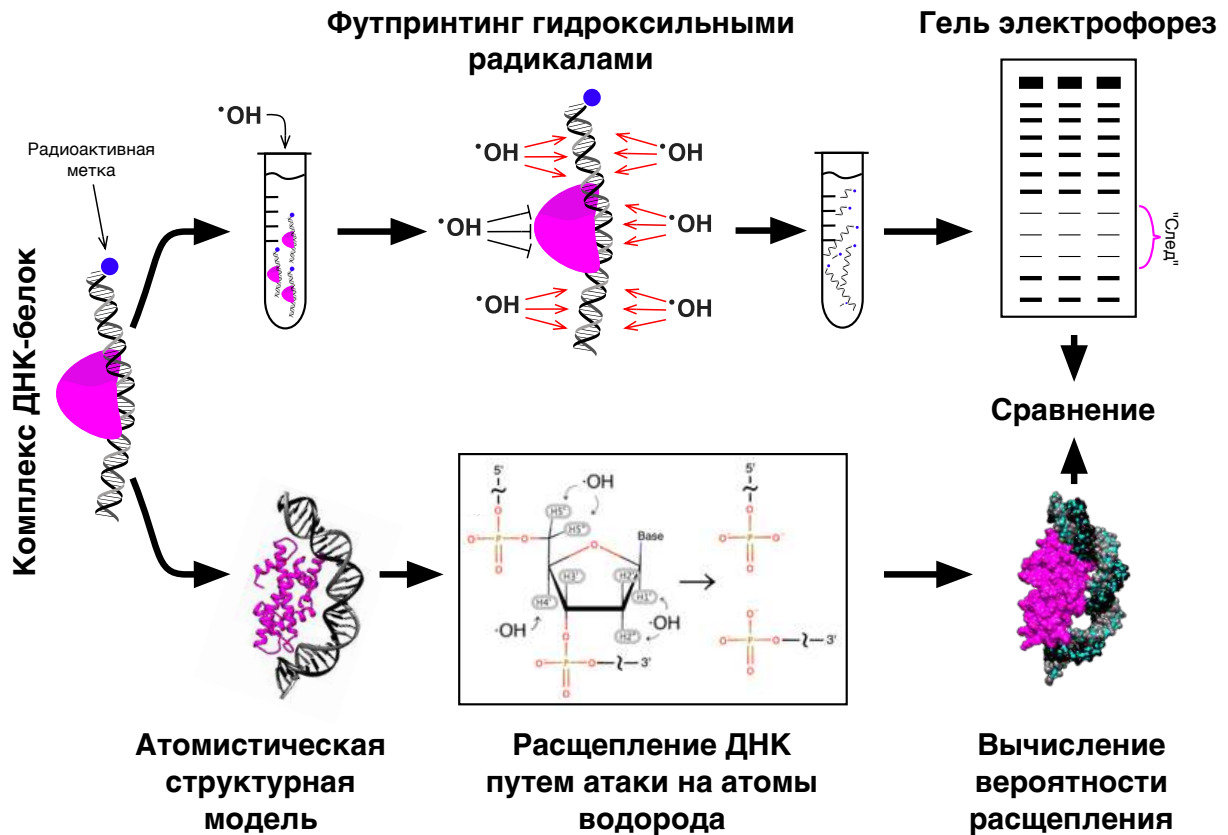


### 3.1.1 Введение

Структурная характеристика комплексов нуклеиновая кислота-белок с высокой точностью необходима для нашего понимания функционирования генома, включая регуляцию экспрессии генов, транскрипцию, репликацию и репарацию ДНК. Хотя рентгеновская кристаллография остается золотым стандартом для определения структурных характеристик на атомистическом уровне, эта методика является сложной и часто не дает результатов из-за проблем с кристаллизацией. Методы ДНК-футпринтинга традиционно использовались как грубые, но универсальные и простые способы исследования взаимодействий ДНК-белок *in vitro* [346] и *in vivo* [347; 348]. В этих методах используются различные химические или ферментные агенты (например, ДНКаза I и гидроксильные радикалы) для расщепления ДНК в позициях, подверженных воздействию растворителя. Расположение этих положений может быть впоследствии идентифицировано с помощью гель-электрофореза при условии, что нити ДНК помечены радиоактивными или флуоресцентными метками на одном конце. Точность ферментативного расщепления ДНК ограничена стерическими эффектами и эффектами, специфичными для последовательности. С другой стороны, HRF является мощным методом, так как расщепление гидроксильными радикалами не имеет эффектов специфичных для последовательности, и обеспечивают высокое разрешение (вплоть до одного нуклеотида) из-за их малых размеров (см. [349] для подробного описания метода HRF).

Основная интерпретация данных HRF часто выполняется качественно путем изучения соответствующих изображений электрофореза в полиакриламидном геле (ПААГЭ) ДНК или профилей интенсивности полос геля, полученных из этих изображений. Однако такой подход обычно не приводит к прямой трехмерной структурной интерпретации данных. Чтобы раскрыть весь потенциал этих высококачественных экспериментальных данных, необходимо выполнить две задачи: 1) количественную оценку фактических частот расщепления ДНК для каждого положения нуклеотида из данных ПААГЭ и 2) оценку частот расщепления ДНК из соответствующих атомистических молекулярных моделей. Вместе эти два компонента анализа позволяют интерпретировать данные HRF на более продвинутом уровне и позволяют включать данные в конвейеры молекулярного моделирования (рис. 3.1) [A11]. Используя

такой подход, атомистические структурные модели высокого разрешения комплексов ДНК-белок могут быть проверены и/или уточнены с использованием экспериментальных данных HRF.



**Рисунок 3.1 — Футпринтинг комплексов ДНК-белок гидроксильными радикалами и структурная интерпретация экспериментальных данных.** На схематической диаграмме изложены основные идеи использования HRF для выяснения структуры комплексов ДНК-белок. Вверху: комплексы ДНК-белок восстанавливаются на радиоактивно меченой ДНК и подвергаются расщеплению гидроксильными радикалами; гидроксильные радикалы расщепляют нити ДНК на участках, не защищенных белком, в режиме кинетики одиночных расщеплений (по одному разрезу на нить); Репертуар расщепленных фрагментов ДНК анализируется с помощью денатурирующего гель-электрофореза, и становятся очевидными положения в последовательности ДНК, защищенных белком. Внизу: атомистические структурные модели могут использоваться для прогнозирования профилей футпринтинга; частота расщепления зависит от доступности атомов водорода дезоксирибозы; сравнение предсказанных и экспериментальных результатов может быть использовано для уточнения или проверки структурных моделей.

### 3.1.2 Разработка протокола

Теоретическая основа для количественной оценки экспериментальных данных футпринтинга гидроксильными радикалами была предложена ранее [350—353]. Однако, столкнувшись с практической необходимостью структурной интерпретации данных HRF для комплексов ДНК-белок в наших недавних исследованиях [A6; A7], мы поняли, что не существует простого протокола и программного обеспечения как для количественной оценки данных HRF, так и для анализа атомистической структуры. Ранее описанные подходы [351; 352], которые использовались для количественной оценки результатов ПААГЭ или расчетов доступной для растворителя площади поверхности для анализа трехмерной структуры, утратили совместимость с современными компьютерными платформами или не имеют воспроизводимых примеров данных для их повторной реализации и проверки. Чтобы удовлетворить эту потребность, мы разработали свободно доступное программное обеспечение и здесь предоставляем пошаговые инструкции по его использованию.

Мы представляем протокол и программный пакет HYDROID (HYDroxyl-Radical fOotprinting Interpretation for DNA). HYDROID обеспечивает количественный анализ данных HRF с разрешением в один нуклеотид, оценку частот расщепления ДНК по кристаллическим структурам или атомистическим молекулярным моделям, сравнение и интеграцию этих подходов. HYDROID написан на Python и использует бесплатные кроссплатформенные компоненты ImageJ для анализа изображений [354] и библиотеку FreeSASA для расчета площади поверхности [355]. Кросс-платформенный и надежный характер структуры HYDROID поддерживается этими хорошо обслуживаемыми, регулярно обновляемыми и тщательно тестируемыми базовыми технологиями. Протокол HYDROID состоит из двух библиотек, HYDROIDexp (анализ и визуализация экспериментальных данных) и HYDROIDpred (анализ структуры) вместе с набором примеров и пошаговыми инструкциями. HYDROID количественно оценивает экспериментальные профили дорожек ПААГЭ с помощью аппроксимации данных с использованием различных функций и ограничений при проведении аппроксимации. Пакет оценивает теоретические профили частоты (интенсивности) расщепления ДНК на основе атомистических трехмерных структур путем вычисления доступной для растворителя площади

поверхности атомов водорода дезоксирибозы. Чтобы ознакомиться с протоколом HYDROID, мы создали видеоурок по адресу <https://ncbi.github.io/HYDROID/docs/video.html>. Протокол успешно применялся в двух исследованиях [A6; A7].

### 3.1.3 Применение метода

HYDROID - это надежный, но гибкий инструмент для анализа и интерпретации экспериментов по футпринтингу ДНК. Он одновременно сочетает в себе функциональность для количественной оценки данных ПААГЭ продуктов футпринтинга (HYDROIDexp) и оценки профилей расщепления ДНК в ходе гидроксильного футпринтинга из экспериментальных структур или молекулярных моделей (HYDROIDpred).

Ранее мы применили подход HYDROID для изучения центромерных нуклеосом дрожжей и их комплексов с белком CENP-C [A6; A7]. Высококачественный анализ экспериментальных данных HRF с помощью HYDROID позволил нам точно определить положение ДНК на октамере гистонов, построить атомистическую модель центромерной нуклеосомы, воссозданной на определенной последовательности ДНК, и идентифицировать сайты взаимодействий белка CENP-C с нуклеосомой.

В этом разделе подход HYDROID был протестирован и подтвержден с использованием нескольких наборов экспериментальных данных по комплексам ДНК-белок. Мы проанализировали влияние различных параметров на производительность HYDROID и дали рекомендации относительно того, какие параметры следует использовать при анализе комплексов ДНК-белок. В разделе “Экспериментальный дизайн” представлены дополнительные сведения, соответствующие примеры и рекомендации.

Основные функции модулей HYDROID можно легко адаптировать и использовать для более широкого спектра приложений. В частности, модуль HYDROIDexp сам по себе может применяться для анализа данных гель-электрофореза нуклеиновых кислот любого типа, обеспечивая точное количественное определение интенсивности полос. Алгоритмы аппроксимации полос, разрабо-

танные в HYDROIDexp, в целом могут быть адаптированы для анализа данных капиллярного электрофореза.

Модуль HYDROIDpred в его нынешнем виде способен анализировать доступность атомов водорода рибозы/дезоксирибозы с использованием подхода вычисления поверхности доступной растворителю H-SASA для свободной ДНК или РНК или в комплексе с белками. Мы не проводили обширных испытаний подхода H-SASA для оценки профилей частоты расщепления в случаях, отличных от комплексов ДНК-белок. Однако другие исследования ранее показали, что оценка профилей H-SASA для определенных атомов водорода в нативной РНК или ДНК может быть аналогичным образом использована для оценки паттернов расщепления HRF [353; 356]. В таких случаях мы рекомендуем ознакомиться с указанными выше ссылками и указать в скрипте HYDROIDpred, какие атомы водорода следует учитывать для анализа. Учитывая важность метода HRF для анализа третичной структуры РНК и РНК-белковых комплексов [357–359], мы ожидаем, что части пакета HYDROID будут полезны для этой цели и потенциально могут быть включены в конвейеры предсказания структуры РНК [360; 361].

### 3.1.4 Сравнение с другими методами

Уникальной характеристикой HYDROID является интеграция количественной оценки экспериментальных данных HRF и анализа атомистической трехмерной структуры в одном пакете. Преимущества функциональности, предоставляемой отдельными компонентами HYDROID (HYDROIDexp и HYDROIDpred), обсуждаются ниже.

Некоторые автономные программы были разработаны на протяжении многих лет для решения задач количественной оценки гелей в экспериментах HRF, такие как GelExplorer [351] или SAFA [352]. Однако одним препятствием при использовании этих программ в настоящее время является то, что они не обновлялись в течение некоторого времени и не совместимы с современными компьютерными системами. Дополнительные преимущества библиотеки HYDROIDexp включают возможность выбора гауссовой или лоренцевой формы полос и множество алгоритмов аппроксимации с ограничениями на вариации

параметров. Библиотека HYDROIDexp также включает инструменты, совместимые с Python, для построения профилей вдоль последовательности ДНК. Это бесплатное программное обеспечение, которое можно повторно использовать в других проектах.

Что касается теоретической оценки профилей расщепления ДНК из трехмерных структур, HYDROIDpred в настоящее время является единственным доступным решением. Он основан на вычислении доступности атомов водорода через библиотеку FreeSASA, которая является быстрой и бесплатной в использовании. HYDROIDpred расширяет ее, предоставляя для расчетов тщательно подобранные наборы атомных радиусов. Ранее сообщалось об альтернативном подходе (подход RADACK), который пытается учесть контролируруемую диффузией природу реакции атаки гидроксильных радикалов и общую геометрию молекулы с помощью стохастического моделирования [362]. Однако в настоящее время нет общедоступного программного обеспечения, реализующего подход RADACK.

### 3.1.5 Преимущества и ограничения

Основным преимуществом HYDROID является его конструкция 2-в-1, предоставляющая пользователю возможность одновременно количественно оценивать экспериментальные данные HRF и сравнивать их с атомистическими структурными моделями комплекса белок-нуклеиновая кислота, если они доступны. Оба компонента HYDROID, реализующие эти функции, имеют свои преимущества. В частности, преимущества HYDROIDexp включают его кроссплатформенный характер (работает в Linux, Windows, MacOS) и одновременную реализацию ряда алгоритмов аппроксимации профиля HRF (формы полос Гаусса или Лоренца в сочетании с несколькими алгоритмами аппроксимации с ограничениями). Алгоритмы аппроксимации с ограничениями хорошо работают при деконволюции сигнала от частично перекрывающихся полос геля. Преимущества HYDROIDpred включают его бесплатную природу без каких-либо патентованных или лицензионных программных компонентов, а также тщательно отобранные и проверенные наборы параметров (радиус зонда, размеры атомов), которые можно использовать для получения значимых оце-

нок профиля расщепления ДНК на основе молекулярных моделей. Алгоритм HYDROIDpred основан на оценке доступной для растворителя области атомов водорода дезоксирибозы. Этот подход не может напрямую объяснить кинетику диффузии гидроксильных радикалов (см. Раздел “Схема эксперимента” для дальнейшего обсуждения). Следовательно, следует проявлять осторожность при сравнении точных форм экспериментальных и теоретических профилей расщепления.

Для подхода HYDROIDpred на вход должны быть поданы PDB-структуры с атомами водорода. Предлагаемый алгоритм добавления атомов водорода (шаг 23) в настоящее время работает только в Linux или Mac.

### 3.1.6 Обзор протокола HYDROID

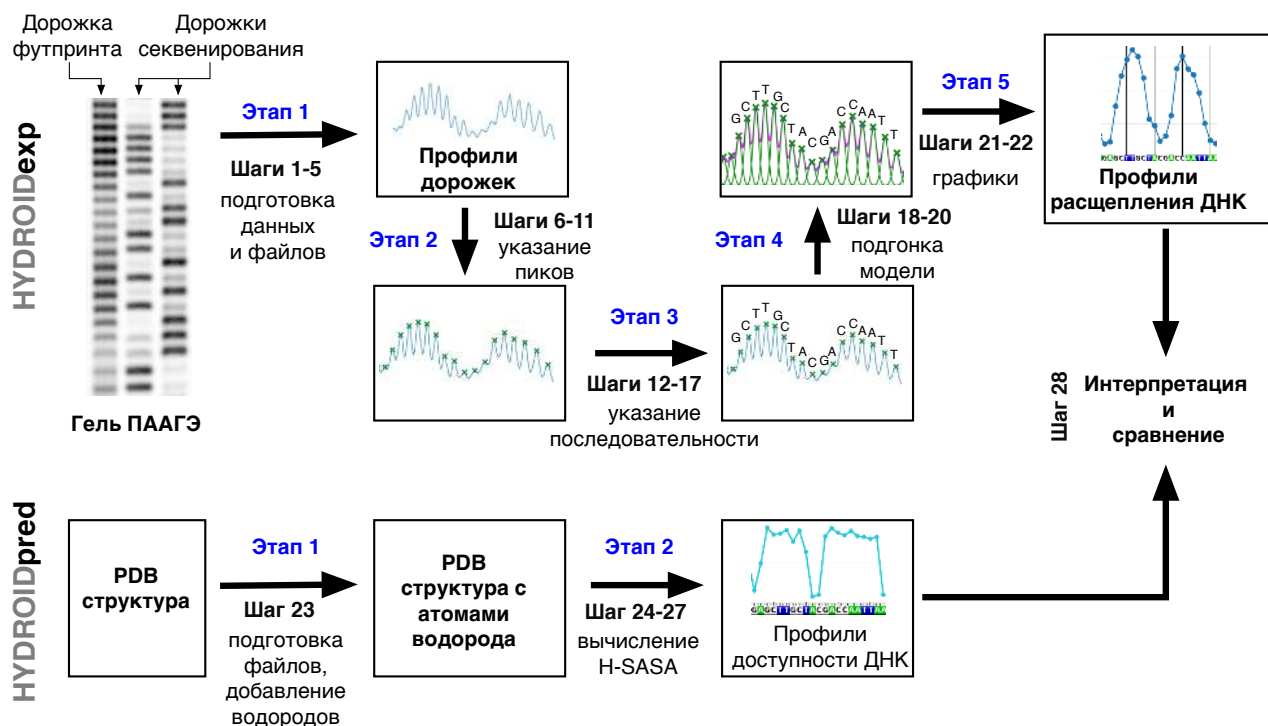
HYDROID состоит из протокола и программного пакета, написанного на языке Python, который вместе с другими компонентами (ImageJ и FreeSASA) представляет собой надежную основу для анализа и интерпретации экспериментальных данных футпринтинга гидроксильными радикалами ДНК-белковых комплексов. Протокол также может применяться для анализа комплексов РНК-белок, а также свободных молекул РНК/ДНК в растворе. Пакет программного обеспечения HYDROID можно легко установить в операционных системах Linux, Mac или Windows. Он состоит из двух дополнительных конвейеров (и соответствующих программных библиотек) HYDROIDexp и HYDROIDpred. Конвейер HYDROIDexp используется для анализа изображений геля ПААГЭ для получения значений частоты расщепления ДНК для каждой позиции на исследуемой последовательности ДНК (рис. 3.2, вверху). Конвейер HYDROIDpred используется для выполнения теоретических расчетов частоты расщепления ДНК на основе трехмерной атомистической структуры или молекулярной модели комплекса ДНК-белок (рис. 3.2, внизу). Расчеты частоты расщепления выполняются путем расчета доступной для растворителя площади поверхности атомов водорода дезоксирибозы (которые атакуются гидроксильными радикалами и вызывают расщепление ДНК) для каждого нуклеотида (профили H-SASA). Профили H-SASA позволяют оценить ожидаемую частоту расщепления ДНК. Сравнение и интегрирование экспериментальных и тео-

ретических профилей можно в дальнейшем использовать для интерпретации экспериментальных данных и/или уточнения молекулярной модели. Основные программные функции, реализованные в пакете HYDROID, перечислены в таблице 4. Они предназначены для запуска из сценария Python, а некоторые функции имеют интерактивный графический интерфейс пользователя. В пакете HYDROID полнофункциональные примеры данных HRF и анализа трехмерной структуры представлены в виде тщательно аннотированных скриптов Python. Эти примеры можно легко загрузить, и они служат в качестве изменяемых шаблонов для анализа пользовательских данных (см. ПРОЦЕДУРА И ОЖИДАЕМЫЕ РЕЗУЛЬТАТЫ). Ниже мы описываем методологию, лежащую в основе ключевых этапов структуры HYDROID. Видеоиллюстрации каждого этапа доступны по адресу <https://ncbi.github.io/HYDROID/docs/video.html>.

### 3.1.7 Конвейер HYDROIDexp

Конвейер HYDROIDexp требует в качестве входных данных изображение геля ПААГЭ с дорожками для продуктов HRF одной из нитей ДНК и соседних дорожек с продуктами реакций секвенирования ДНК для той же цепи. Конвейер дает значения относительной частоты расщепления ДНК (также называемые значениями интенсивности расщепления ДНК) вдоль последовательности ДНК. Этот конвейер можно применять параллельно к нескольким образцам, которые были обработаны на одном и том же геле. Конвейер состоит из пяти этапов. См. Диаграмму рабочего процесса на рисунке 3.2.





**Рисунок 3.2** — Конвейеры данных в HYDROID. Схематическая диаграмма для анализа экспериментальных данных по футпринтингу гидроксильными радикалами (вверху, HYDROIDexp). Теоретическая оценка профилей расщепления ДНК по молекулярным структурам (внизу, HYDROIDpred). Шаги и этапы процедуры аннотированы на схеме. Конвейер HYDROIDexp состоит из пяти этапов (отмечены синим цветом): этап 1 - профили дорожек в виде файлов данных извлекаются из изображения ПЛАГЭ, этап 2 - положения полос/пиков назначаются полуавтоматически на профилях дорожек, этап 3 - соответствие между полосами геля и последовательностью ДНК устанавливается, этап 4 - математическая модель подбирается для описания формы профиля дорожки с помощью функций Гаусса или Лоренца, этап 5 - интенсивности отдельных полос (частоты расщепления ДНК) извлекаются из модели. Конвейер HYDROIDpred состоит из двух этапов: этап 1 - загружается файл PDB интересующей системы и добавление атомов водорода к атомистической структуре при необходимости, этап 2 - профили доступности ДНК рассчитываются путем оценки доступной для растворителя области атомов водорода дезоксирибозы (H-SASA). Наконец, можно сравнить наборы теоретических и экспериментальных данных.

*Этап 1: Извлечение профилей дорожек геля.* Экспериментальные профили дорожек извлекаются из изображения геля. Экспериментальный профиль интенсивности полосы геля (профиль полосы) в дальнейшем определяется как

массив значений интенсивности на изображении геля вдоль указанной полосы гель-электрофореза (в направлении миграции ДНК по гелю). Функции HYDROID считывают данные экспериментального профиля гелевой дорожки, сохраненные в виде столбцов чисел в текстовых файлах. Для получения “цифровых” экспериментальных профилей дорожек из изображений гелей ПААГЭ мы предлагаем использовать программное обеспечение ImageJ [354] с плагином Bio-Formats [363] везде, где это необходимо для открытия файлов определенных форматов изображений гелей. Чтобы HYDROID мог считывать данные из сгенерированного файла, необходимо указать отдельный текстовый файл конфигурации, в котором указываются имена дорожек и соответствующие имена столбцов в сгенерированном файле.

*Этап 2: Назначение отдельных полос/пиков на геле.* Положение отдельных полос геля идентифицируется на профиле полосы геля с помощью полуавтоматического интерактивного алгоритма. Для деконволюции профиля полосы в набор вкладов от отдельных полос, сначала нужно знать количество и приблизительное расположение полос на профиле дорожки. Каждая полоса обычно соответствует локальному пику на профиле дорожки. Полуавтоматический интерактивный алгоритм, реализованный в функции *assign\_peaks\_interactive* в HYDROID, открывает графическое окно и позволяет пользователю изменять параметры процедуры поиска пиков до тех пор, пока все необходимые пики не будут правильно идентифицированы по оценке пользователя. Алгоритм поиска пиков основан на библиотеке PeakUtils Python. Алгоритм регулируется двумя параметрами: относительной известностью пиков (параметр пикового порога - будут обнаружены только пики с амплитудой выше порога) и минимальным расстоянием между пиками. Алгоритм поиска пиков разделяет профиль полосы на несколько сегментов и позволяет пользователю выбирать минимальное расстояние между пиками для самого левого и самого правого сегментов (параметры *min\_dist\_left* и *min\_dist\_right*); значения в промежуточных сегментах затем выводятся с использованием линейной интерполяции. Если некоторые полосы перекрываются до такой степени, что пики не могут быть идентифицированы на профиле полосы или интенсивность полосы очень мала, в HYDROID реализована опция интерполяции пиков, чтобы попытаться вывести их положения из известных положений и расстояний между соседними группами, которые более заметны. Наконец, положения пиков можно указать вручную в файле конфигурации.

*Этап 3: Привязка пиков HRF к последовательности ДНК.* Каждую полосу на дорожке геля HRF (которая эквивалентна пику на профиле дорожки HRF) можно отнести к определенному положению в последовательности ДНК путем сравнения дорожек геля HRF с дорожками с продуктами реакций секвенирования ДНК для тех же ДНК. В HYDROID функция *call\_peaks\_interactive* позволяет интерактивно строить профиль дорожки HRF вместе с профилями реакций секвенирования ДНК и последовательностью ДНК. Визуально сравнивая их с последовательностью ДНК, пользователь может указать местоположение любого одиночного пика на профиле и его соответствующее положение на последовательности ДНК. Это позволяет установить соответствие между всеми полосами и положениями в последовательности ДНК.

*Этап 4: Аппроксимация профилей интенсивности дорожек.* Математические модели, описывающие распределение интенсивности каждой полосы с помощью функций Гаусса или Лоренца, подгоняются к данным профилю полосы с помощью функции *fit\_peaks*. Эта процедура позволяет вам провести деконволюцию необработанного экспериментального профиля HRF на вклады отдельных полос геля, тем самым должным образом принимая во внимание такие эффекты, как частичное перекрытие сигналов полос и неравномерная ширина полосы вдоль полосы геля. Известно, что эта процедура подвержена от переобучению, если на алгоритм подгонки не накладываются дополнительные ограничения [364]. В HYDROID реализован ряд различных вариантов подгонки с ограничениями, перечисленных в Таблице 5. См. Раздел “Проблемы при экспериментальной количественной оценке профилей HRF” для обсуждения дополнительных деталей.

Чтобы оценить качество подгонки экспериментального профиля полосы с помощью модели, HYDROID предоставляет несколько критериев оценки: RMSD - среднеквадратичное отклонение между экспериментальными и прогнозируемыми значениями модели, относительное RMSD - вариация последнего, где отклонение в каждой точке выражается как часть экспериментального значения, и коэффициент корреляции Пирсона между экспериментальными и прогнозируемыми значениями.

Значения частоты расщепления ДНК (общие площади под пиками на профиле дорожек) оценивают методом нелинейной аппроксимации методом наименьших квадратов. Для каждого пика площадь рассчитывается под сегментом кривой, который включает этот пик. Сегмент определяется серединами

между положением данного пика (положением его максимума) и положением пиков его соседей слева и справа. Относительные различия в значениях прокси-площади, определенных для экспериментального профиля дорожки, по сравнению с профилем подобранной модели, используются как оценки неопределенности и называются “относительными ошибками площади пика”. Максимальная относительная ошибка площади пика среди всех пиков, а также среднеквадратическое значение (“RMSD относительной площади пика”) вычисляются в HYDROID.

*Этап 5: Расчет частоты расщепления ДНК.* Частота расщепления ДНК для каждой позиции на ДНК вычисляется как площадь под гауссианой/лоренцевой кривой, описывающей соответствующую полосу. Значения профиля частоты расщепления ДНК отображаются вдоль последовательности ДНК.

### 3.1.8 Конвейер HYDROIDpred

Теоретический расчет профилей расщепления ДНК с помощью HYDROID основан на анализе трехмерных атомистических структур комплексов белок-ДНК и оценке доступности атомов водорода дезоксирибозы для гидроксильных радикалов. Трехмерные координаты можно получить с помощью рентгеновской кристаллографии или молекулярного моделирования (моделирование по гомологии, интегративное моделирование, докинг и т. д.). Доступная для растворителя площадь поверхности атомов водорода дезоксирибозы рассчитывается для каждого нуклеотида в молекуле ДНК структурного комплекса, что приводит к двум профилям H-SASA, по одному для каждой цепи (в случае, если протокол применяется к РНК, это приведет к одному профилю H-SASA). Он дает оценочные значения относительной частоты расщепления ДНК. Конвейер состоит из трех основных этапов.

*Этап 1: Приготовление атомистической модели комплекса ДНК-белок с известными положениями атомов водорода.* Важным шагом является определение положений атомов водорода, если они отсутствуют в исходной структуре. Небольшие различия в положениях атомов водорода (например, разные параметры длины для связей C-H) могут повлиять на величину оцененных профилей (см. Подробное обсуждение в разделе “Проблемы оценки теорети-

ческих частотных профилей расщепления ДНК из атомистических структур”). Программа REDUCE [365] из пакета AmberTools 17 [47] с расстояниями связей Х-Н, полученными из положений ядер атомов, может использоваться при добавлении атомов водорода в рентгеновские структуры.

*Этап 2: Оценка теоретической частоты расщепления ДНК.* Теоретические профили частоты расщепления ДНК оцениваются путем расчета доступной для растворителя площади поверхности атомов водорода дезоксирибозы (профили H-SASA) с помощью функции `get_DNA_H_SASA`, которая использует библиотеку FreeSASA [355]. Алгоритм Ли-Ричардса используется [366] в FreeSASA с точностью до 200 срезов на атом по умолчанию. Для расчета SASA могут применяться другие радиусы зонда, по умолчанию 1,4Å. HYDROID реализует три набора атомных радиусов для расчетов H-SASA: а) радиусы FreeSASA по умолчанию, б) значения радиусов из силового поля CHARMM36 [299] и в) значения радиусов из силового поля AMBER ff10 (parm10) [47]. Последние два получены из параметров `rmin` ван-дер-Ваальса (расстояние/радиус, при котором энергия Леннарда-Джонса имеет минимум) соответствующих силовых полей. Затем выполняется визуализация рассчитанных профилей (см. Визуализация и сравнение данных ниже).

### 3.1.9 Визуализация и сравнение данных

Для визуализации и сравнения профиля частоты расщепления ДНК HYDROID включает модуль построения графиков (`plot_prof_on_seq`) на основе библиотеки Matplotlib [367]. Этот модуль может одновременно строить несколько профилей вместе с последовательностью ДНК и позволяет использовать несколько методов нормализации (для приведения профилей к одному масштабу). Реализованы некоторые тривиальные методы нормализации, такие как деление каждого профиля на его максимальное значение (“`every_method`”) или деление обоих профилей на их максимальное общее значение (“`together_method`”). Кроме того, “метод подгонки” выполняет линейную регрессию без свободного члена между значениями профиля и изменяет их масштаб максимизируя взаимное максимальное соответствие. Использование линейной регрессии без свободного члена соответствует физическому требова-

нию, согласно которому позиции, недоступные для расщепления, должны иметь нулевые значения на обоих профилях.

Кроме того, мы реализовали метод моделирования профилей полос геля и изображений геля из профилей частоты расщепления (функция *simulate\_gel*). Эта функция полезна для изучения ожидаемой формы профилей гелевых полос (и, следовательно, разрешения профиля в конкретной интересующей области ДНК) и планирования экспериментов. Модель подвижности ДНК Огстона используется для моделирования подвижности ДНК в геле [368].

<b>Библиотека HYDROIDexp</b>	
<code>assign_peaks_interactive</code>	Предоставляет интерактивный интерфейс для полуавтоматического алгоритма, используемого для определения положения пиков на профилях гелевых дорожек.
<code>call_peaks_interactive</code>	Предоставляет интерактивный интерфейс для назначения индивидуальных пиков относительно положения последовательности ДНК.
<code>fit_peaks</code>	Основная функция, которая подгоняет математическую модель профиля гелевой дорожки и оценивает интенсивность расщепления ДНК.
<code>plot_prof_on_seq</code>	Строит любой профиль данных по последовательности ДНК и позволяет нормализовать и подгонять два или более профилей.
<code>simulate_gel</code>	Имитирует форму профиля полосы геля и выводит изображение полосы геля на основе профиля H-SASA.
<b>Библиотека HYDROIDpred</b>	
<code>get_DNA_H_SASA</code>	Оценивает теоретические профили расщепления ДНК (профили H-SASA) по структуре ДНК-белка.

**Таблица 4** — Основные функции HYDROID и их описание.

Тип ограничения (аббревиатура)	Описание
dSIGMA $\geq$ 0	Ширина пика $\sigma$ не должна уменьшаться с увеличением $D$ , положения пика (то есть от начала геля до конца). Это значительно повышает стабильность решения и предотвращает переобучение.
SIGMA $<$ 2*dD	Ширина пика $\sigma$ не может превышать удвоенное расстояние между данным пиком/полосой и последующим пиком ( $dD$ ). Подразумевает автоматически “dSIGMA $\geq$ 0”.
SIGMA=k*D+b	Ширина пика $\sigma$ линейно связана с положением пика ( $D$ ). Это эффективно устанавливает линейную зависимость между шириной полосы и ее подвижностью.
log(SIGMA)=P2(log(M))	Логарифм ширины пика $\sigma$ пропорционален логарифму числа нуклеотидов в ДНК, $M$ , (или ее молекулярной массе) через полином второй степени. Оптимальные коэффициенты полинома определяются при подгонке: $\sigma_i = e^{a*(\log M_i)^2 + b*\log M_i + c}$
SAFA	Ограничения, реализованные в программе SAFA [352], ширина пика связана с положением соседних пиков через $\sigma_i = \sigma_0 + k * \frac{D_{i+1} - D_i}{2}$

**Таблица 5** — Список ограничений при подгонке моделей, доступные в HYDROID.

### 3.1.10 Планирование экспериментов

#### Метод HRF и интерпретация данных

В методе HRF используются гидроксильные радикалы, образованные в результате реакции Фентона или облучения, для расщепления цепей ДНК в доступных для растворителя участках [349]. Химический процесс, лежащий в основе, следующий: гидроксильный радикал атакует атомы водорода дезоксирибозы, что приводит к отщеплению водорода и последующему расщеплению основной цепи (рис. 3.1). Основными продуктами этого химического расщепления являются две цепи, оканчивающиеся 3' и 5' фосфатами, которые примыкают к атакованному нуклеотиду в исходной цепи ДНК. Анализ продуктов расщепления ДНК обычно проводят с использованием денатурирующего электрофореза в полиакриламидном геле (ПААГЭ) предварительно радиоактивно/флуоресцентно меченой цепи (нитей) ДНК. Обычно рекомендуется радиоактивное мечение 5'-конца ДНК [349]. В пределах кинетики единичного попадания (один разрез на цепь ДНК) доступность каждого нуклеотида для атаки гидроксильных радикалов должна быть пропорциональна интенсивности соответствующей полосы геля. Отнесение каждой полосы к определенному положению в последовательности ДНК является важным этапом и выполняется путем получения продуктов различных реакций секвенирования ДНК (таких, как реакции Максама-Гилберта) на соседних полосах геля (рис. 3.2).

Во время экспериментов по гидроксильному футпринтингу нуклеотид подвергается атаке гидроксильных радикалов, а затем разрушается, давая соответствующий фрагмент ДНК, который на один нуклеотид короче, чем предполагаемый фрагмент ДНК, оканчивающийся фактически атакованным нуклеотидом. Положение в последовательности ДНК, указанное пользователем на этом этапе, должно соответствовать нуклеотиду, который фактически подвергается атаке гидроксильных радикалов (и разрушается). Этот факт вместе с природой используемых реакций секвенирования ДНК следует учитывать при отнесении пика профиля HRF к положению в последовательности ДНК. Реакции секвенирования Максама-Гилберта удобны в этом отношении, поскольку



они дают продукты, которые также на один нуклеотид короче, чем предполагаемые фрагменты ДНК, оканчивающиеся атакованным нуклеотидом [12].

Количественная оценка фактических частот расщепления ДНК требует измерения интенсивности отдельных полос из профилей интенсивности полос геля. Необходимо учитывать частичное перекрытие полос и изменение их ширины [369]. Для определения общей интенсивности каждой полосы ранее была предложена математическая модель, которая использовалась для моделирования профиля интенсивности полос геля [351]. При таком подходе интенсивность сигнала каждой полосы описывалась колоколообразной функцией, а результирующий профиль моделировался как сумма колоколообразных функций. К настоящему времени было предложено несколько функций моделирования, включая функции Гаусса, Лоренца [351], упрощенную интегрированную функцию Вейбулла [370] и более сложные функции [371]. Хотя функция Гаусса является моделью по умолчанию для описания расширения полосы из-за свободной диффузии молекул, следует также учитывать некоторые другие факторы (например, электрическое поле, градиенты температуры, неоднородности геля и т. д.). В частности, методы радиографической детекции сигнала с геля могут изменять исходный сигнал, делая в некоторых случаях более подходящей модель функции Лоренца [372]. По нашему опыту, данные, полученные для гелей с использованием современных установок, хорошо соответствуют функциям Гаусса. Более того, переобучение (*overfitting*) - обычная проблема для этих типов задач, и ее можно решить, наложив дополнительные ограничения на решение [364].

Качество изображения геля, которое используется для извлечения профилей полос геля, является важной характеристикой, которая может повлиять на качество количественной оценки результатов HRF. Для получения хороших результатов ожидается, что в интересующей области будет отчетливо видна лестница из электрофоретических полос, представляющих продукты расщепления ДНК, отличающийся длиной в один нуклеотид. Однако протокол HYDROID также может быть успешным при деконволюции профилей дорожек в регионах, где некоторые полосы плохо видны или имеют значительное перекрытие. В этом случае пользователь должен иметь возможность указать количество ожидаемых диапазонов и их приблизительное расположение. С этой целью можно использовать места полос на других дорожках геля.

Вычислительное структурное моделирование комплексов ДНК-белок можно направлять и проверять с помощью методов футпринтинга ДНК с высоким

разрешением. Существует острая необходимость в программном решении для выполнения этих задач и объединения экспериментальных данных с трехмерными структурными моделями. Ранее Balasubramanian et al. провели детальное исследование интенсивности разрыва цепи ДНК при атаке на различные атомы водорода дезоксирибозы [350]. Они обнаружили, что атомы H5' и H5'' обладают самой высокой реакционной способностью, за ними следуют H4', H3', H2' + H2'' и H1'. Более того, было обнаружено, что реакционная способность атомов водорода хорошо коррелирует с их доступной для растворителя площадью поверхности (SASA), которую можно оценить с помощью рентгеновских структур или предсказать с помощью структурных моделей. Это открытие является основой для теоретической оценки частот расщепления ДНК по атомистическим структурам в HYDROID. Однако, как мы обсудим далее, надежная оценка SASA для атомов водорода в комплексе ДНК-белок требует тщательного выбора параметров, таких как атомные радиусы, радиус сферы зонда и детали алгоритмов размещения атомов водорода.

### Количественная обработка экспериментального профиля HRF

Количественная оценка экспериментальных профилей дорожек HRF выполняется в HYDROID путем деконволюции формы экспериментального профиля как суммы гауссовых (или лоренцевых) функций с центрами в положениях полос. Подгонка модели без ограничений обычно приводит к физически неправильным решениям (хотя и с лучшими характеристиками согласия), часто видны явные нарушения в положениях и ширине колоколообразных кривых, соответствующих отдельным полосам (верхняя панель рис. 3.3В, а также результирующая частота расщепления ДНК (рис. 3.4). Чтобы решить эту проблему, в HYDROID реализовано несколько ограничений в алгоритмах аппроксимации (Таблица 4). Наиболее гибким является алгоритм “dSIGMA > = 0”, который не позволяет ширине функций Гаусса уменьшаться при переходе от начала геля к его концу. Это основано на физическом предположении, что расширение полосы (более высокие коэффициенты дисперсии) увеличивается вместе с подвижностью ДНК в полосе. Это ограничение делает подобранное решение качественным в большинстве случаев, налагая только минимальные

ограничения (рисунок 3.3). Однако всегда рекомендуется визуальный осмотр полученного решения. Среди методов аппроксимации с ограничениями алгоритм “dSIGMA $\geq 0$ ” обычно обеспечивает наилучшее соответствие, если судить по характеристикам качества аппроксимации. Другие алгоритмы ограничений также обеспечивают разумные решения со значениями RMSD относительной площади пика менее 2% и почти идентичными результирующими профилями частоты расщепления ДНК.

Анализ деконволюции профиля гелевых дорожек в HYDROID может быть выполнен с использованием функций Гаусса или Лоренца для аппроксимации распределения интенсивности полос. Некоторые предыдущие исследования показали, что Лоренциан может лучше соответствовать данным электрофореза ДНК [351]. Для наборов экспериментальных данных, используемых в этом исследовании, функция Гаусса давала лучший результат, чем функция Лоренца, хотя различия в качестве аппроксимации не превышали 2%, оцененные с помощью относительного среднеквадратичного отклонения. Однако профили интенсивности расщепления ДНК имели заметную разницу: при использовании лоренциана максимальные интенсивности расщепления имели тенденцию быть выше, а минимальные интенсивности расщепления были ниже, чем те, которые получены с использованием функции Гаусса. Этот эффект может быть объяснен медленным затуханием хвостов лоренцевского распределения, что заставляет пики с более высокой интенсивностью вносить больший вклад в амплитуду сигнала в соседних полосах с низкой интенсивностью.

В профилях дорожек ПААГЭ геля иногда могут отсутствовать легко идентифицируемые максимумы для определенных полос. Это может произойти из-за шума, низкой интенсивности некоторых полос или перекрытия полос. Пока пользователь может определить количество и приблизительное положение этих полос, алгоритмы деконволюции HYDROID могут надежно их количественно оценить. Рисунок 3.3 иллюстрирует это для левой части профиля геля, где некоторые полосы перекрываются. Количественная оценка профилей, полученных электрофорезом одних и тех же продуктов реакции на двух разных гелях, дала почти идентичные результаты (неопубликованные наблюдения).

## Оценка теоретических частотных профилей расщепления ДНК из атомистических структур

Следуя идеям, приведенным в [350] HYDROID оценивает теоретические профили частоты расщепления ДНК гидроксильными радикалами путем вычисления доступной для растворителя площади поверхности (SASA) атомов водорода дезоксирибозы (называемой профилями H-SASA). В общем, тот же метод можно применять для расчета доступной для растворителя площади поверхности атомов водорода рибозы РНК. Чтобы понять влияние различных параметров на форму профилей H-SASA, мы взяли структуру нуклеосомы в качестве примера и рассчитали профили H-SASA с различными наборами параметров: радиус зонда, наборы атомных радиусов, методы добавления атомов водорода, вклады от разных атомов водорода. Интерактивные графики профилей H-SASA, рассчитанных с различными параметрами, доступны по адресу [https://ncbi.github.io/HYDROID/examples/example1/results/nucl\\_H-SASA.html](https://ncbi.github.io/HYDROID/examples/example1/results/nucl_H-SASA.html). Зависимость профилей H-SASA от радиуса зонда довольно проста: зонды меньшего размера могут различать более мелкие детали окклюзии ДНК белком, но делают профиль менее гладким и более зависимым от локальной геометрии взаимодействий ДНК-белок. Радиус зонда 1,4 Å является обычным выбором, но расчет профилей H-SASA с различными размерами зонда может быть полезным для характеристики взаимодействий ДНК-белок в различных масштабах.

Профили H-SASA довольно существенно зависят от выбора наборов атомных радиусов (FreeSASA vs CHARMM36 vs AMBER10) и от алгоритма, используемого для добавления атомов водорода (их положения в рентгеновских структурах обычно не разрешаются). Методологически последовательный подход требует генерирования положений атомов водорода с использованием параметров топологии из того же силового поля молекулярной механики, которое используется для оценки атомных радиусов. Как показано ранее, профили H-SASA чувствительны к небольшим колебаниям в геометрии остова ДНК как из-за неточностей рентгеновской структуры, так и из-за деталей конформации ДНК [A7]. Чтобы решить эту проблему, можно использовать моделирование молекулярной динамики (МД) для релаксации рентгеновских структур и анализа их динамики. Основной вклад в H-SASA вносят значения площади, доступной

для растворителя от атомов Н5', Н5'' и Н4' (рис. 3.5А). Если Н-SASA рассчитывается только для атомов Н5', Н5'' и Н4', соответствие между профилями, рассчитанными на основе МД и исходных рентгеновских структур, хорошее (рис. 3.5Б). В частности, положения минимумов Н-SASA (нуклеотиды ДНК, защищенные белком) подтверждаются между профилями (включая профили МД Н-SASA, рассчитанные со всеми рассматриваемыми атомами водорода дезоксирибозы). Следовательно, при использовании рентгеновских структур без МД-релаксации рекомендуется расчет профилей Н-SASA на основе только атомов Н5', Н5'' и Н4'.

Различия в форме экспериментальных и теоретических профилей расщепления ДНК можно объяснить двумя фактами. Во-первых, поскольку эксперименты с HRF проводятся в растворе, динамический характер конформации ДНК и взаимодействий между ДНК и белком могут привести к усреднению экспериментального профиля. Такие эффекты невозможно учесть теоретически, если не рассматривать и не анализировать динамический ансамбль структур. Во-вторых, из-за контролируемой диффузией природы реакции атаки гидроксильных радикалов, общая геометрия молекулы, включая ширину бороздок ДНК, также может модулировать вероятность атаки [353; 362; 373]. Профили Н-SASA могут лишь частично улавливать такие эффекты. Следовательно, при сравнении теоретических и экспериментальных профилей акцент следует делать на сравнении положений нуклеотидов, которые максимально защищены от атаки гидроксильными радикалами. Эффекты контролируемые диффузией отсутствуют для позиций ДНК, полностью защищенных белком.

### 3.1.11 Примеры данных

Протокол HYDROID использует два набора экспериментальных данных (названные “Пример 1” и “Пример 2”, доступные в репозитории GitHub), чтобы проиллюстрировать работу HYDROID. Пример 1 содержит данные HRF для центромерных нуклеосом *S. cerevisiae*, восстановленных на последовательности ДНК 601TA с 3'-концами, меченными радиоактивно, а пример 2 содержит данные HRF нуклеосом *G. gallus*, восстановленных на последовательности ДНК

601 с 5'-концами, меченными радиоактивно. Детали эксперимента для первого и второго наборов можно найти в ссылках [A6; A7] и [374] соответственно.

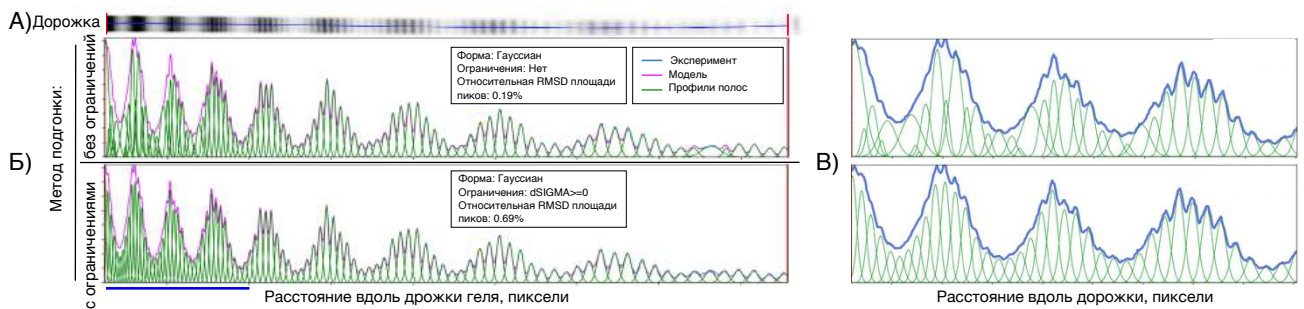
Оба набора экспериментальных данных снабжены набором скриптов Python, которые образуют конвейер для анализа данных с использованием HYDROID и ожидаемых результирующих файлов.

Пример 1 также предоставляет файлы с трехмерными координатами - молекулярную модель центромерной нуклеосомы дрожжей с последовательностью ДНК 601TA [375] (которая недоступна в PDB), которая была построена посредством моделирования по гомологии (с использованием Modeller [46] и 3DNA [114]). Последовательности гистонов были получены из HistoneDB2.0 [A14] и из соответствующей рентгеновской структуры нуклеосомы *X. laevis* (PDB 3LZ0), как описано ранее [A7]. Программа REDUCE [365] из пакета AmberTools 17 [47] использовалась для генерации положений атомов водорода с расстояниями связей X-H, полученных из положений ядер и электронных облаков. Кроме того, был использован подход молекулярной динамики. Используя команду VMD psfgen с силовым полем CHARMM36 [299], гидратированные структуры были сгенерированы с последующим моделированием минимизации, релаксации и молекулярной динамики, как описано в наших более ранних исследованиях [A7].

### **Опыт, необходимый для реализации протокола.**

Пользователи должны быть знакомы с интерфейсом командной строки в Linux, Mac или Windows. Базовые знания языка сценариев Python были бы полезны для изменения простых сценариев Python с помощью текстового редактора. Для HYDROIDpred полезна часть понимания протокола формата PDB-файла.

Оболочка графического пользовательского интерфейса (GUI) доступна для основных функций HYDROIDexp как дополнительный модуль по адресу [https://github.com/intbio/HYDROID\\_GUI](https://github.com/intbio/HYDROID_GUI) для тех пользователей, которые предпочли бы использовать GUI-интерфейс.



**Рисунок 3.3** — Количественная оценка изображения геля ПААГЭ путем подбора математической модели. А) ПААГЭ изображение продуктов расщепления гидроксильными радикалами для одной цепи ДНК нуклеосомы. Направление миграции ДНК на изображении слева направо. Для получения этого изображения центромерные нуклеосомы *S. cerevisiae* воссоздали на хорошо позиционирующей последовательности ДНК 601TA с ДНК, радиоактивно меченной на 3'-конце. Расщепление ДНК, экстракция ДНК, ПААГЭ и получение радиоактивного сигнала выполнялись, как описано в [A7]. Б) Извлеченные профили дорожек аппроксимируются с помощью модели, которая представляет каждую полосу геля как функцию Гаусса, показаны результаты аппроксимации без и с ограничениями на параметры модели. Площадь под каждым гауссом представляет интенсивность полосы и, следовательно, частоту расщепления ДНК. В) Увеличенная версия графиков для левой части геля, выделенная синей линией на панели В. Неравномерности ширины и положения гауссианы для метода аппроксимации без ограничений хорошо видны. Экспериментальные и модельные кривые невозможно различить, потому что они почти перекрываются. Ось Y на графиках представляет значения интенсивности профиля.

### 3.1.12 Материалы

#### Оборудование и данные

- Компьютер, работающий в среде Linux, Mac или Windows.
- ПААГЭ изображения продуктов реакции NRF и продуктов реакции секвенирования ДНК (обычно получаемые с помощью фосфоимаджера, если фрагменты ДНК

радиоактивно помечены). Файлы изображений должны иметь достаточно хорошее разрешение, чтобы в интересующей области расстояние между полосами было больше 10 пикселей. Формат файла изображения должен быть одним из тех, которые может читать ImageJ как таковой или через плагин BioFormats. • (Необязательно) 3D-атомистические структуры или молекулярные модели исследуемого комплекса ДНК-белок.

### Скачивание примеров

Пакет HYDROID поставляется с набором примеров (шаблонов протоколов), доступных по ссылке <https://github.com/ncbi/HYDROID/tree/master/examples>.

#### 3.1.13 Описание протокола

**ВАЖНО** Перед выполнением процедуры убедитесь, что HYDROID и сопутствующие программные инструменты установлены.

#### **Количественная оценка изображений HRF ПААГЭ (конвейер и библиотека HYDROIDexp)**

1. *Извлечение профилей полос из файлов изображений ПААГЭ и подготовка файлов данных для HYDROID.* Чтобы извлечь профиль полосы из изображения используйте программное обеспечение ImageJ. Сохраните файл изображения с линиями, обозначающими диапазоны извлеченных данных, для дальнейшего использования.

2. Загрузите образцы проектов анализа HRF, поставляемые с HYDROID. Если ДНК в экспериментах с HRF была помечена на 3'-конце, используйте пример 1; если ДНК в экспериментах с HRF была помечена на 5'-конце, используйте



пример 2. Папка примера проекта содержит сценарии Python для каждого этапа конвейера и папку данных с несколькими файлами данных, которые будут использоваться на следующих этапах.

3. Замените файл DATA/LANE\_PROFILES.XLS файлом данных профилей полос, извлеченным из изображения ПЛАГЭ на шаге 1.

4. Измените файл DATA/LANE\_CONFIG.CSV, чтобы указать, какие столбцы из файла DATA/LANE\_PROFILES.XLS следует анализировать HYDROID, и укажите понятные человеку идентификаторы. На этом этапе для всех остальных параметров должно быть установлено значение NaN (используйте верхнюю строку исходного файла в качестве шаблона).

5. Укажите исследуемую последовательность ДНК в файле DATA/DNA\_SEQ.FASTA.

6. *Определение положений пиков (местоположений электрофоретических полос) на профилях дорожек геля с помощью полуавтоматического интерактивного алгоритма.* Откройте сценарий EXP\_S2\_ASSIGN\_PEAKS.PY в текстовом редакторе и измените переменную lane\_names (в строках 37–38), чтобы включить в нее имена полос из файла DATA/LANE\_CONFIG.CSV, которые следует проанализировать.

**ВАЖНО:** Идентификация положений пиков в конечном итоге требуется только для дорожек с данными HRF, но ее можно использовать для изучения профилей дорожек секвенирования ДНК с помощью той же процедуры.

7. Запустите сценарий с помощью следующей команды в терминале: `> PYTHON EXP_S2_ASSIGN_PEAKS.PY` Для каждой дорожки, указанной в Шаге 6, откроется интерактивное окно, показывающее график профиля дорожки и позволяющее настроить параметры алгоритма идентификации пиков, а также ряд дополнительных параметров.

**ВАЖНЫЙ ШАГ:** Интерактивные окна открываются одно за другим. Пожалуйста, закройте каждое окно, чтобы перейти к следующему.

8. Цели следующих шагов - указать диапазон данных для анализа и выделить каждый пик (полосу) в этом диапазоне звездочкой. Используйте интерактивные ползунки, чтобы указать диапазон значений профиля HRF, который будет анализироваться в дальнейшем (параметры LEFTLIM и RIGHTLIM).

9. Настройте ПИКОВЫЙ ПОРОГ (регулирует чувствительность алгоритма к величине пиков), MIN\_DIST\_LEFT (регулирует минимально допустимое расстояние между пиками в левой части диапазона данных), MIN\_DIST\_RIGHT

(регулирует минимально допустимое расстояние между пиками справа от диапазона данных), BASELINE (вычитает линейную базовую линию из данных перед попыткой идентифицировать пики) и SEGMENTS (устанавливает количество сегментов для разделения диапазона данных для интерполяции минимально необходимого расстояния между пиками) до тех пор, пока каждый пик (диапазон) не будет выделен только одной звездочкой. **ВАЖНЫЙ ШАГ:** В пределах анализируемого диапазона данных каждое местоположение полосы должно быть выделено только одной звездочкой. В противном случае последующий анализ будет некорректным. **КРИТИЧЕСКИЙ ШАГ:** Иногда интенсивность полос очень мала, данные зашумлены или полосы подходят так близко друг к другу, что на профиле полосы не видно локальных максимумов, представляющих полосу. Алгоритм деконволюции может работать хорошо в таких случаях, если ожидаемое положение полосы отмечено звездочкой. Последнее может быть достигнуто путем включения переключателя INTERPOLATE в интерактивном окне. Алгоритм попытается угадать положения полос, которые не представлены определенным пиком на профиле HRF. Кроме того, положение полос можно указать вручную в файле DATA/LANE\_CONFIG.CSV. Для этого их местоположение можно ввести в столбец ADDPEAKS файла.

10. (Необязательно) Параметр ALIGNPOS позволяет указать положение любого пика, который будет использоваться на следующем этапе для выравнивания профиля HRF с профилями секвенирования ДНК для задания последовательности ДНК. Обычно это может быть заметный пик на конце геля, который однозначно идентифицируется как полоса, представляющая продукт ДНК одинаковой длины на дорожках HRF и секвенирования ДНК.

11. Нажмите кнопку Сохранить. Данные будут записаны в файл DATA/LANE\_CONFIG.CSV. Повторите процедуру для других профилей полос.

12. *Задание последовательности ДНК на профиле дорожки HRF.* Откройте сценарий EXP\_S3\_CALL\_PEAKS.PY в текстовом редакторе и измените список LANE\_SETS, включив в него имена дорожек HRF (FOOTPRINTING\_PROFILE) вместе с именами соответствующих профилей секвенирования ДНК (HELPER\_PROFILES) из файла DATA/LANE\_CONFIG.CSV.

**КРИТИЧНЫЙ ШАГ:** Обязательно укажите правильный конец маркировки ДНК в переменной МЕТКИ (начиная со строки 35). Убедитесь, что последовательность ДНК, предоставленная для каждого профиля дорожки HRF (переменная SEQ), соответствует конкретной последовательности ДНК це-

пи ДНК, которая использовалась в эксперименте (записана в обозначениях 5' → 3').

13. Запустите сценарий с помощью следующей команды: `> PYTHON EXP_S3_CALL_PEAKE.PY` Для каждого профиля дорожки HRF, указанного на шаге 12, откроется интерактивное окно, где он будет нанесен на график вместе с соответствующими профилями секвенирования ДНК. Также будет показана последовательность исследуемой цепи ДНК.

14. Цель следующих шагов 14-17 состоит в том, чтобы вручную сопоставить любой пик на профиле HRF с соответствующим положением в последовательности ДНК, просматривая профили секвенирования ДНК, нанесенные на тот же график. При необходимости используйте ползунки ALIGNPOS, чтобы выровнять профили секвенирования ДНК друг с другом и профилем HRF.

15. Используйте ползунок ПОСЛЕДОВАТЕЛЬНОСТИ, чтобы выбрать позицию в последовательности ДНК (выбранная позиция будет выделена заглавной буквой).

16. Используйте ползунок SEQPEAK, чтобы навести стрелку на пик профиля дорожки HRF, соответствующий положению в последовательности ДНК, идентифицированной на предыдущем шаге.

**КРИТИЧЕСКИЙ ШАГ:** Пользователь должен понимать природу выбранной реакции секвенирования ДНК и ее продуктов, чтобы правильно выполнить этот шаг (см. Дизайн эксперимента).

17. Нажмите кнопку Сохранить. Данные будут записаны в файл `DATA/LANE_CONFIG.CSV`. Повторите процедуру для других профилей дорожек HRF.

18. *Подбор модели к профилю дорожки HRF для количественной оценки частот расщепления ДНК.* Откройте скрипт `EXP_S4_FIT_MODEL.PY` в текстовом редакторе и измените список `LANE_DATA`, включив в него имена дорожек HRF, которые будут подвергаться количественной оценке. Параметры функции `FIT_PEAKE` также можно изменить для настройки результатов: для параметра `PEAKTYPE` можно задать значение Гаусса или Лоренца в зависимости от желаемой модели полосы, `FIT_CONSTRAINT` может принимать одно из значений, указанных в Таблице 5, `MAXFEV` управляет максимальным количеством шагов оптимизации, установите меньшее количество, чтобы получить быстрые предварительные результаты.

19. Запустите сценарий с помощью следующей команды: `> PYTHON EXP_S4_FIT_MODEL.PY`

20. Скрипт выведет интерактивные графические окна с графиками результатов подгонки (Рисунок 3.3Б). Результаты будут автоматически сохранены в CSV-файл, указанный в скрипте (`RESULTS/LANE_NAME_FITTED_INTENSITIES.CSV`). Скрипт будет запускаться последовательно для всех профилей дорожек. **ВАЖНЫЙ ШАГ.** Пользователь должен изучить полученные графики (при необходимости увеличить масштаб, используя кнопку масштабирования) и убедиться, что в результатах подгонки нет отклонений, подобных тем, которые показаны на верхних панелях рисунков 3.3Б, 3.3В.

21. *Построение графика частоты расщепления вдоль последовательности ДНК.* HYDROID предоставляет настраиваемую функцию `PLOT_PROF_ON_SEQ`, которую можно использовать для построения графиков профилей данных поверх последовательности ДНК (рисунок 3.4). Откройте сценарий `EXP_S5_PLOT_CL_FREQ.PY` в текстовом редакторе и измените список `LANE_DATA` (рядом со строкой 28), включив в него имена файлов, содержащих результаты количественной оценки из шага 20. Также можно указать заголовок графика и имя выходного файла (строки 40-43).

22. Запустите сценарий с помощью следующей команды:  
`> PYTHON EXP_S5_PLOT_CL_FREQ.PY` Откроются интерактивные окна с графиками, и графики будут сохранены в `RESULTS/LANE_NAME_FITTED_INTENSITIES.PNG`.

### **Прогнозирование интенсивности расщепления на основе атомистических трехмерных структур (конвейер и библиотека HYDROIDpred).**

23. *Подготовка файла PDB с трехмерной структурой комплекса ДНК-белок.* Файл PDB, представляющий вашу экспериментальную систему, может быть доступен в базе данных PDB ([www.pdb.org](http://www.pdb.org)) или быть продуктом молекулярного моделирования. Поскольку в структурах PDB, полученных с помощью рентгеновских лучей, часто отсутствуют атомы водорода, вы можете

использовать следующую команду для добавления атомов водорода: `> REDUCE -NUCLEAR FILE.PDB> FILE_H.PDB` где `file.pdb` - это имя исходного файла PDB.

Полученный PDB-файл должен быть сохранен в подкаталоге `DATA/STRUCTURE`.

24. *Запустите расчеты SASA, чтобы оценить профили частоты расщепления ДНК.* Откройте скрипт `PRED_S2_CALC_H-SASA.PY` в текстовом редакторе и измените список `PROF_DATA`, включив в него имя (имена) файла(ов) структуры PDB, идентификатор PDB цепи, которая представляет рассматриваемую нить ДНК, остатки. Параметр должен указывать диапазон номеров остатков из PDB-файла, который будет анализироваться, параметр `SEQ` должен указывать последовательность ДНК исследуемой цепи. Параметр `HCONTRIB` представляет собой список из семи чисел, который определяет вес вклада каждого отдельного атома водорода дезоксирибозы в следующем порядке [H1' H2' H2'' H3' H4'H5' H5'']. Для рентгеновских структур с низким и средним разрешением (выше 2,5-3Å) предлагается учитывать только вклады атомов H4', H5' и H5''.

25. Запустите сценарий с помощью следующей команды: `> PYTHON PRED_S2_CALC_H-SASA.PY` Предполагаемые профили расщепления ДНК будут сохранены в файлах `CSV RESULTS/PROF_NAME_H-SASA.CSV`, где `PROF_NAME` - имя, указанное ранее в списке `PROF_DATA`.

26. *Построение расчетных частот расщепления ДНК вдоль последовательности ДНК* Откройте сценарий `PRED_S3_PLOT_H-SASA.PY` в текстовом редакторе и измените список `LANE_DATA`, включив в него имя файла CSV с предполагаемым профилем, последовательность ДНК исследуемой цепи и имя выходного файла.

27. Запустите сценарий с помощью следующей команды: `> PYTHON PRED_S3_PLOT_H-SASA.PY` Откроются интерактивные окна с графиками, дополнительно графики будут сохранены в виде файлов PNG.

## Сравнение данных, полученных из конвейеров HYDROIDexp и HYDROIDprep

28. Данные из файлов CSV, полученные на этапах 20 (экспериментальные профили) и 27 (теоретические профили), можно открыть, нанести на график и сравнить с помощью любого программного обеспечения для построения графиков, предпочитаемого пользователем (например, MS Excel). CSV-файл, созданный конвейером HYDROIDexp (см., Например, RESULTS/SCCSE4\_601TA\_BS\_FITTED\_INTENSITIES.CSV в примере 1), имеет столбцы с именами *Site* и *Intensity*, соответствующие положению последовательности ДНК и количественной оценке интенсивности/частоты расщепления ДНК, соответственно. CSV-файл, созданный конвейером HYDROIDpred (см., Например, RESULTS/SCCSE4\_601TA\_BS\_H-SASA.CSV в примере 1), имеет столбцы с именами *Site* и *H-SASA*, соответствующие положению последовательности ДНК и теоретически оцененной вероятности расщепления ДНК, соответственно. Для тех, кто обладает продвинутыми навыками в языке Python, мы предоставляем примеры для обработки и построения результатов примера 1 в файле COM\_PLOT\_EX\_VS\_PRED.PY.

### Временные затраты

Время, необходимое для протокола, зависит от количества экспериментальных образцов для анализа и знакомства пользователя с инструментами командной строки и языком сценариев Python. Оценки для изображения ПА-АГЭ геля с двумя дорожками HRF и четырьмя дорожками секвенирования ДНК на основе Примера 1 представлены ниже:

Шаг 1. Обработка файла изображения с помощью Fiji ImageJ: 30 мин.

Шаги 2-11, Определение положения пиков (расположение электрофоретических полос): 30 мин.

Шаги 12-17, определение последовательности ДНК на профиле дорожки HRF: 30 мин.

Шаги 18-20, Подгонка модели к профилю дорожки HRF: 30 мин.

Шаги 21-22, извлечение и нанесение на график частот расщепления вдоль последовательности ДНК: 10 мин.

Шаг 23. Подготовка файла PDB с трехмерной структурой комплекса ДНК-белок: 15 мин.

Шаги 24-25, Настройка и выполнение расчетов SASA для оценки профилей частоты расщепления ДНК: 20 мин.

Шаги 26-27, нанесение на график расчетных частот расщепления ДНК вдоль последовательности ДНК: 5 мин.

### 3.1.14 Ожидаемые результаты

#### Пример из практики: ДНК-белковые взаимодействия в нуклеосомах

В качестве примера применения HYDROID к системам ДНК-белок мы проанализировали и интерпретировали эксперименты HRF для нескольких нуклеосом, поскольку они представляют собой одну из самых сложных систем для изучения и демонстрируют периодические взаимодействия ДНК-белок. Были получены два независимых набора экспериментальных данных из разных лабораторий: а) HRF нуклеосом, восстановленных на 601 хорошо позиционирующей последовательности [376] с использованием гистонов из эритроцитов *G. gallus*, б) HRF нуклеосом, восстановленных на аналогичной хорошо позиционирующей последовательности 601TA [375] с использованием рекомбинантных гистонов центромерных нуклеосом *S. cerevisiae* [A7]. В двух независимых наборах данных также использовались разные методы маркировки ДНК: в первом случае был помечен 5'-конец, а во втором - 3'-конец.

Результаты анализа этих экспериментов, а также сравнения с теоретическими профилями H-SASA представлены на рисунке 3.6. Известно, что ДНК в нуклеосоме хорошо расположена в случае последовательности 601/601TA и намотана вокруг октамера гистона в сверхспиральной форме (рис. 3.6А) интенсивности расщепления ДНК, как было ранее показано, очень похожи для верхней и нижней цепей из-за псевдосимметрии нуклеосомы [A7]. Профили расщепления ДНК отражают периодичность нуклеосомной ДНК 10-11 п.н.,

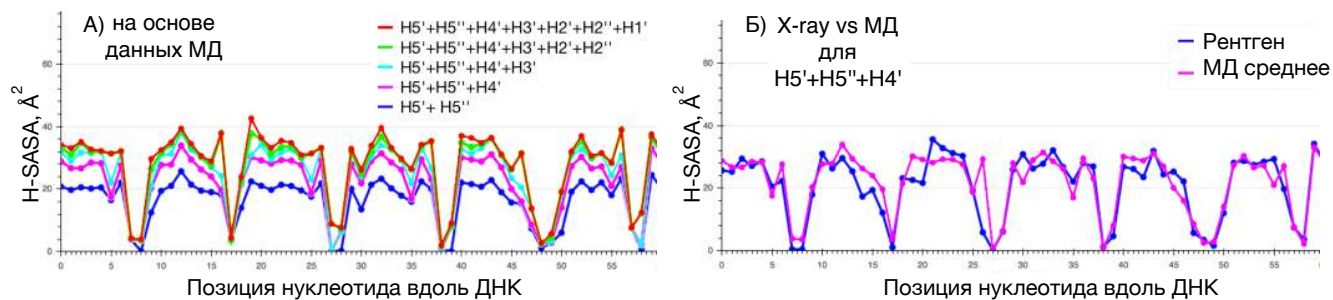
когда она вращается вокруг поверхности октамера гистонов (рис. 3.6Б). Два экспериментальных профиля, полученные с использованием разной маркировки концов, хорошо совпадают, но отличаются в некоторых аспектах. Это может отражать воспроизводимость, которую можно ожидать от анализа HRF, выполненного с использованием различных протоколов. Ключевой характеристикой является положение минимумов и максимумов, в то время как точная величина расщепления ДНК может зависеть от экспериментальных условий, наличия частично собранных нуклеосомных состояний [183], динамики ДНК и белков [193]. Положения минимумов и максимумов совпадают с точностью до 1 нуклеотида между профилями (Рисунок 3.6Б). Для ДНК, меченной 3'-концом, в большинстве случаев минимумы смещены на одну пару оснований в сторону 5'-конца ДНК по сравнению с экспериментом с ДНК, меченной по 5'-концу. Возможное объяснение этой тонкой разницы заключается в том, что во время расщепления гидроксильных радикалов помимо основных продуктов с концевыми 5'- и 3'-фосфатами образуются другие второстепенные продукты. В частности, на 5'-конце может быть образована цепь, оканчивающаяся 5'-альдегидной группой [350]. Она на один нуклеотид длиннее цепи с 5'-фосфатным концом, в ней отсутствует отрицательный заряд фосфатной группы и его подвижность в геле на 2–3 нуклеотида ниже, чем у обычного продукта [377]. На 3'-конце сайта расщепления второстепенным продуктом является 3'-фосфогликолят, который, как известно, менее распространен и мигрирует близко к обычному продукту [350]. Таким образом, обычно рекомендуется мечение 5'-конца ДНК.

Пример сравнения теоретических профилей расщепления ДНК, извлеченных из структурной модели, и экспериментальных данных HRF показан на рисунке 3.6В. Положения минимумов экспериментального и теоретического профилей совпадают с точностью до 1 нуклеотида. Форма теоретического профиля менее плавная, чем у экспериментального, что отражает ограничения теоретической модели. Более гладкая форма экспериментальных профилей, вероятно, связана с динамикой структур в растворе, а также с кинетическими эффектами, связанными с диффузией гидроксильных радикалов. Интерактивные веб-графики, сравнивающие профили H-SASA, полученные с разными параметрами, доступны в репозитории GitHub по ссылке [https://ncbi.github.io/HYDR0ID/examples/example1/results/nucl\\_H-SASAvsEXP\\_BS.html](https://ncbi.github.io/HYDR0ID/examples/example1/results/nucl_H-SASAvsEXP_BS.html).

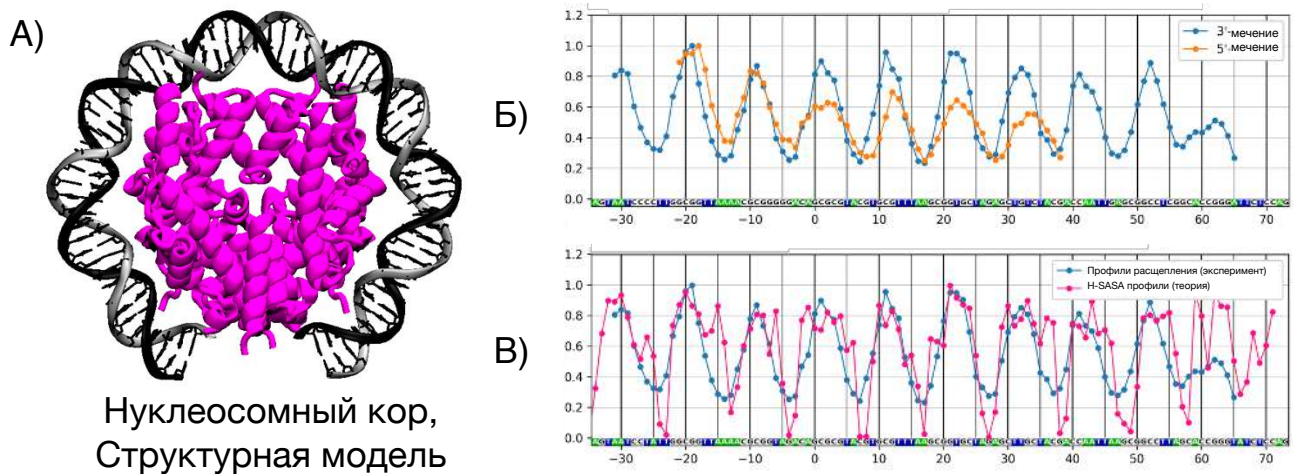




**Рисунок 3.4** — Профили частоты расщепления ДНК, извлеченные из экспериментов по гидроксильному футпринтингу. В соответствии с нашей методологией количественной оценки данных HRF кривая профиля полосы на рисунке 3.3В была построена с помощью комбинации функций Гаусса. Области этих индивидуальных функций Гаусса представляют собой интенсивности отдельных полос и, следовательно, значения частоты расщепления ДНК. Эти количественные значения частоты расщепления ДНК представлены здесь как для аппроксимации с и без органичней на параметры модели. Соответствие между последовательностью ДНК и профилем HRF было идентифицировано с использованием продуктов реакций секвенирования Максама-Гилберта. Видны отклонения в значениях, полученные в результате аппроксимации без ограничений на параметры модели.



**Рисунок 3.5** — Профили H-SASA для ДНК в нуклеосомах. А) Профили H-SASA (доступная для растворителя площадь поверхности атомов водорода дезоксирибозы) рассчитываются с использованием вкладов различных атомов водорода дезоксирибозы для каждого нуклеотида вдоль последовательности ДНК, начиная с центра нуклеосомы. Профили были рассчитаны путем усреднения более 50 снимков из МД моделирования с интервалом в 1 нс. Б) Сравнение профилей H-SASA (на основе атомов H5', H5'' и H4'), рассчитанных на основе рентгеновской структуры (с атомами водорода, добавленными через программу REDUCE) и траектории МД. Если доступны только атомы H5', H5'' и H4', профили не сильно различаются, и для надежной оценки профиля можно использовать рентгеновские структуры.



**Рисунок 3.6 — Экспериментальные и теоретические профили частоты расщепления ДНК гидроксильными радикалами в нуклеосомах.**

А) Структура ДНК-белковых взаимодействий в нуклеосоме (из модели по гомологии нуклеосомы *S. cerevisiae*, воссозданной на основе последовательности ДНК 601 на основе структуры PDB ID 3LZ0), показаны центральные 70 п.н. ДНК, гистоновые хвосты не показаны. ДНК демонстрирует периодический паттерн взаимодействий с октамером гистонов каждые 10-11 п.н. Б) Сравнение двух независимо полученных экспериментальных профилей для нуклеосом, восстановленных на последовательности ДНК 601/601TA, полученным в результате работы HYDROID. Профиль, ДНК меченной по 5'-концом, соответствует нуклеосомам, восстановленным на 601 позиционирующей последовательности с использованием гистонов из эритроцитов *G. gallus* (эти данные доступны в HYDROID как Пример 2). 3'-меченный профиль соответствует нуклеосомам, воссозданным на аналогичной позиционирующей последовательности 601TA с использованием рекомбинантных гистонов центромерных нуклеосом *S. cerevisiae* (эти данные доступны в HYDROID как Пример 1). Сравнение экспериментальных и N-SASA профилей для центромерной нуклеосомы *S. cerevisiae*. Экспериментальный профиль такой же, как и профиль, у ДНК меченной по 5'-концу из панели В. Профиль N-SASA был получен с помощью HYDROIDpred с использованием модели по гомологии центромерной нуклеосомы *S. cerevisiae* на основе структуры PDB 3LZ0 (также доступной в примере 1).

### 3.1.15 Благодарности

Эта работа была поддержана программами внутренних исследований Национальной медицинской библиотеки и Национального института рака, Национальных институтов здоровья; Российским научным фондом [грант № 14-24-00031, договор № 14-24-00031-р] (разработка алгоритмов моделирования гелей и пакета HYDROID\_GUI); Медицинским институтом Говарда Хьюза Исследовательский кампус Джанелии и Университетом Джонса Хопкинса; Российско-американской программой сотрудничества в рамках программы приглашенных стипендиатов Национального института здоровья; грантами NIH R01GM119398 и R21CA220151, R21DE025398 и P50DE019032. В этой работе использовались вычислительные ресурсы кластера NIH HPC Biowulf (<http://hpc.nih.gov>) и вычислительные ресурсы центра высокопроизводительных вычислений МГУ им. М.В. Ломоносова.

Мы благодарим Т. Туллиуса за ценные предложения, которые помогли сформулировать первоначальную идею этого исследования, Л. Заславского за обсуждения оптимальных алгоритмов аппроксимации данных и С. Миттернахта за изменения в программе FreeSASA.

## 3.2 Создание моделей нуклеосом по данным гидроксильного футпринтинга

*Содержание данной главы следует статье [A7].*

Нуклеосомы представляют собой наиболее распространенные комплексы белок-ДНК у эукариот, которые обеспечивают уплотнение геномной ДНК и участвуют в регуляции транскрипции, репликации и репарации ДНК. Детали позиционирования ДНК на нуклеосоме и конформации ДНК могут являться важными регуляторными сигналами. Метод футпринтинга комплексов белок-ДНК гидроксильными радикалами – это химический метод, который позволяет исследовать организацию нуклеосом в растворе с высокой точностью, недостижимой другими методами. В этой работе мы предлагаем метод интегративного моделирования для построения атомистических моделей нуклеосом с высоким

разрешением на основе экспериментов по гидроксильному футпринтингу. Наш метод точно определяет положение ДНК на нуклеосоме путем на основе данных футпринтинга для обеих цепей ДНК, используя свойства псевдосимметрии нуклеосомы. Мы провели эксперименты по гидроксильному футпринтингу с высоким разрешением для центромерных нуклеосом *Saccharomyces cerevisiae*, структура которых является предметом дискуссий. Мы охарактеризовали эти структуры, используя наш подход интегративного моделирования. Наша модель обеспечивает основу для дальнейшего понимания взаимодействия между белком Cse4p и последовательностями аденинов (А-трактами), важными для функции центромеры в пекарских дрожжах.

### 3.2.1 Введение

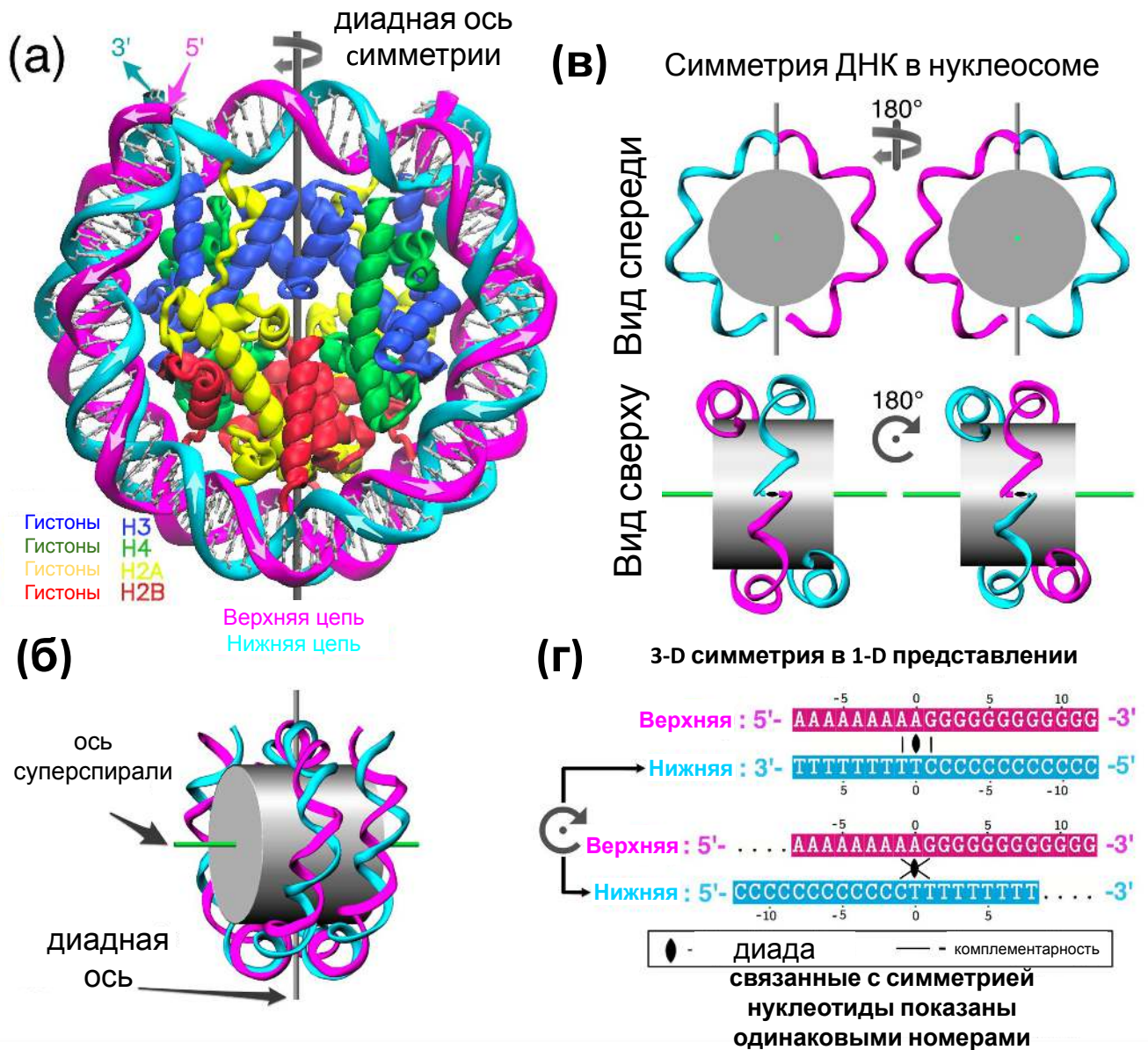
Нуклеосомы - это элементарные строительные блоки хроматина, включающие сегмент ДНК, связанный с октамером гистоновых белков (Н3, Н4, Н2А, Н2В - по две копии каждого) [164]. Кор нуклеосомы (называемый, кором нуклеосомы или NCP) состоит из 145–147 п.н. ДНК, обернутых в виде  $\sim 1,7$  левых сверхспиральных витка вокруг октамера [165] (рис. 3.7). Хотя общая структура нуклеосомы хорошо охарактеризована [166], каждая конкретная нуклеосома может значительно отличаться от своих аналогов за счет вариаций в последовательности ДНК, включения альтернативных вариантов гистонов [A14; A25], их посттрансляционных модификаций [378] и взаимодействий с другими белками [379; 380]. Эти изменения и вариации часто влияют на доступность и конформацию ДНК, которые, в свою очередь, модулируют основные процессы в хроматине, такие как транскрипция, репликация, репарация ДНК и т. д. [A23; 381]. Смещение положения ДНК в нуклеосоме только на одну пару оснований приводит к изменениям в ее ротационном позиционировании примерно на 36 градусов, что может быть достаточно, чтобы повлиять на связывание многих белков с нуклеосомами, считывающими последовательность ДНК (например, пионерские факторы транскрипции [382; 383]). Следовательно, детальная структурная характеристика нуклеосом имеет большое значение. Атомные структуры отдельных нуклеосом получались с помощью рентгеновской кристаллографии [166]. Однако этот метод требует очень много времени и

во многих случаях сильно ограничен ограничениями по кристаллизации, составу и конформации нуклеосомы, потенциально приводя к конформациям, отличным от тех, которые могут быть в растворе. Поэтому очень востребованы точные и универсальные методы характеристики структур нуклеосом. В лаборатории широко используются определенные биохимические методы для быстрой характеристики доступности и расположения ДНК на нуклеосомах в растворе. В этих методах используется визуализация “следа” белков на ДНК - разрезание ДНК ферментами или химическими веществами. Эти методы позволяют картировать взаимодействия белок-ДНК и охарактеризовать конформацию ДНК вдоль последовательности ДНК. Метод гидроксил-радикального футпринтинга (HRF) признан методом, обеспечивающим наивысшее разрешение благодаря небольшому размеру и нейтральной природе гидроксильных радикалов, приводящих к независимому от типу оснований расщеплению ДНК [349]. Считается, что разрыв цепи ДНК во время HRF происходит в основном за счет отщепления атомов водорода дезоксирибозы (рис.3.8) [384]. HRF обычно выполняется в несколько этапов: (а) каждая цепь ДНК метится на одном конце радиоактивным (или флуоресцентным) зондом; (б) комплекс белок-ДНК (или свободную ДНК) обрабатывают гидроксильными радикалами, которые расщепляют ДНК в открытых участках в режиме кинетики одиночного разрезания (по одному разрезу на цепь); (в) комплекс белок-ДНК можно снова очистить; (г) ДНК затем очищают от белков, денатурируют и подвергают денатурирующему электрофорезу в полиакриламидном геле высокого разрешения (ПААГЭ); (д) интенсивность каждой полосы (т. е. положения основания) на геле должна быть пропорциональна частоте расщепления ДНК во время реакции расщепления на конкретном нуклеотидном сайте [349]. Отображение каждой полосы в конкретное положение вдоль последовательности ДНК может быть легко выполнено путем электрофореза продуктов реакций секвенирования ДНК (например, реакций Максама-Гилберта) на соседних дорожках. Применение HRF к нуклеосомам было впервые предложено Hayes et al. задолго до появления первых рентгеновских структур нуклеосом с атомным разрешением [385–388]. Эти авторы смогли оценить спиральную периодичность ДНК в нуклеосомах, и их работа с тех пор послужила эталоном для многих других исследований, в которых для характеристики нуклеосом использовался метод футпринтинга гидроксильными радикалами. В последних исследованиях использовались нуклеосомы с альтернативными последовательностями ДНК [374; 389; 390], связями между

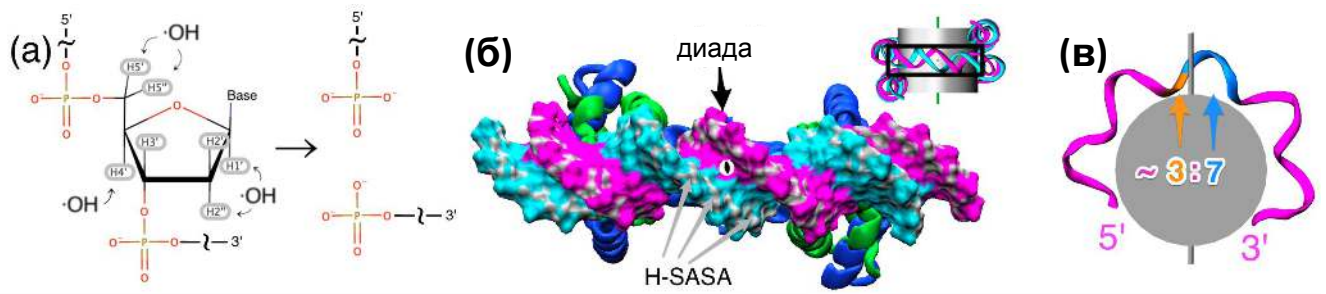
нитьями ДНК, введенными химиотерапевтическими агентами [390], нуклеосомы, взаимодействующие с ремоделерами [391] или другими белками [392]. Хотя метод HRF в нуклеосомах, несомненно, помог охарактеризовать расположение нуклеосом, геометрию ДНК и сайты взаимодействия с нуклеосомсвязывающими белками, существует дополнительный потенциал этого метода, который можно задействовать. Этот потенциал заключается в количественной оценке данных HRF и объединении этих данных с методами интегративного молекулярного моделирования для получения моделей нуклеосом на атомном уровне. Молекулярное моделирование нуклеосом является сложной задачей, поскольку оно требует не только моделирования октамера гистонов, но также включает в себя поиск правильного ротационного и трансляционного позиционирования ДНК по отношению к октамеру. Предыдущие исследования показали, что определенные нуклеосомы в дрожжах хорошо позиционированы, особенно центромерные [393], и поэтому скольжение нуклеосомной ДНК только на одну пару оснований может привести к значительным изменениям в ее экспозиции. Частоты расщепления ДНК, извлеченные из экспериментов HRF в растворе, обычно имеют разрешение в один нуклеотид и могут предоставить данные об относительной доступности ДНК для расщепления, на которое влияют взаимодействия с белками и / или конформация ДНК (например, узкая малая бороздка). Экспериментальные данные HRF могут предоставить удобный источник данных для проверки и уточнения нуклеосомных моделей. Основа для связи экспериментальных данных футпринтинга с молекулярным моделированием была сформулирована Balasubramanian et al. которые предположили, что частоты разрыва цепи ДНК гидроксильными радикалами были пропорциональны доступной для растворителя площади поверхности атомов водорода дезоксирибозы, рассчитанной по известным структурам [350]. Это облегчило интерпретацию профилей HRF с точки зрения количественных показателей, ширины больших и малых бороздок ДНК и других геометрических характеристик ДНК [353; 394]. Насколько нам известно, этот подход еще не применялся систематически к нуклеосомам. В качестве альтернативы Бегусова и соавт. предложили сложный гибридный подход к моделированию, основанный на моделировании методом Монте-Карло диффузии гидроксильных радикалов в сочетании с учетом параметров реакционной способности, полученными из подгонки моделей к экспериментальным данным [362; 395]. Ограниченный прогресс в использовании данных HRF для построения надежных моделей нуклеосом

частично связан с отсутствием исследований, которые пытаются напрямую связать данные футпринтинга нуклеосом с рентгеновскими структурами высокого разрешения и динамикой нуклеосом. Другой причиной является недостаток методологий для количественной оценки экспериментальных профилей HRF нуклеосом и понимания потенциальных источников систематических ошибок и вариаций в экспериментальных данных. В текущем исследовании мы пытаемся решить вышеупомянутые проблемы с целью предоставить надежный способ построения моделей нуклеосом на основе данных HRF и понять потенциальные различия между структурами нуклеосом, в кристаллических структурах и в растворах. Первая часть нашей работы посвящена теоретической оценке профилей HRF из трехмерных структур и их рационализации с точки зрения конформации ДНК и взаимодействий белок-ДНК. В частности, этот анализ позволил установить связь между псевдосимметрией нуклеосомы и сходством профилей HRF комплементарных нитей ДНК. Во второй части нашей работы мы выполнили HRF с высоким разрешением в растворе для двух разных систем на основе центромерных нуклеосом *S. cerevisiae*, собранных *in vitro* и содержащих вариант гистона Cse4 / CENP-A с различными последовательностями ДНК: центромерной последовательности ДНК из хромосомы III (CEN3) и хорошо позиционирующей последовательностью 601TA. Основываясь на экспериментальных данных и данных моделирования, мы предлагаем новый простой интегративный метод определения положения ДНК в нуклеосомах (как ротационных, так и трансляционных) с разрешением в одну пару оснований. В нашем методе используются экспериментальные данные HRF для обеих цепей ДНК и псевдосимметрия нуклеосом. Наконец, мы применили наш метод для определения позиционирования ДНК в нуклеосомах (позиционирование нуклеосом) и построили структурную модель центромерной нуклеосомы *S. cerevisiae*, собранной *in vitro*, ДНК которой необычайно богата А-трактами, которые, как известно, имеют решающее значение для функционирования центромер. Мы обсуждаем значение нашей структурной модели для функции центромеры.





**Рисунок 3.7** — Структура нуклеосом и псевдосимметрия верхней и нижней цепей ДНК. (а) Представление структуры нуклеосомного ядра (PDB 1KX5 ([238])). (б) Схематическое изображение нуклеосомы, показывающее расположение оси псевдосимметрии второго порядка и сверхспиральной оси. (в) Иллюстрация псевдосимметричного отношения между верхней и нижней цепями ДНК в нуклеосоме (показаны только части цепей ДНК). (г) Изображение пространственной симметрии ДНК на плоских диаграммах последовательности: обычное представление (вверху) и сонаправленное (внизу). В последнем представлении две нити выровнены путем размещения диадных нуклеотидов верхней и нижней цепей ДНК один под другим; следовательно, все связанные симметрией пары нуклеотидов верхней и нижней цепей ДНК также выровнены (один под другим). Комплементарность оснований между выбранными нуклеотидами показана на обоих изображениях.



**Рисунок 3.8** — Детали гидроксил-радикальных взаимодействий с нуклеосомной ДНК. (а) Химическая реакция гидроксильных радикалов с основной цепью ДНК: атомы водорода дезоксирибозы (выделены серым) атакуются, что приводит к разрушению остатка дезоксирибозы и расщеплению ДНК. Показаны только основные продукты расщепления. (б) Сегмент нуклеосомы, рассматриваемый сверху (см. Вставку) в представлении площади поверхности, доступной для растворителя (SASA). Участки SASA, принадлежащие атомам водорода дезоксирибозы (H-SASA), окрашены в серый цвет. (в) Асимметрия контактов гистон-ДНК относительно положения оси диады приводит к ее проявлению на профилях расщепления ДНК. Диада делит сегмент цепи ДНК между двумя соседними сайтами связывания на неравные части с соотношением  $\sim 3:7$  при подсчете в направлении 5'-3'.

### 3.2.2 Материалы и методы

#### Экспериментальные процедуры

*Приготовление фрагментов ДНК.* ДНК 601-ТА была амплифицирована полимеразной цепной реакцией (ПЦР) с асимметричным сайтом рестрикции *Ava*I, CTCGGG, на обоих концах и клонирована в модифицированный вектор pUC19, который содержит сконструированный асимметричный сайт *Ava*I. Однако в исследовании ПЦР-амплификация высоко АТ-богатой (90%) центромерной ДНК пекарских дрожжей была проблематичной. Таким образом, фрагмент ДНК SEN3 был химически синтезирован с асимметричным сайтом *Ava*I и клонирован в вектор pUC57 (Genscript USA Incorporated, Пискатауэй, Нью-Джерси, США). Асимметричный сайт *Ava*I использовали для лигирования

фрагментов ДНК в тандемные массивы. Такие тандемные массивы фрагментов ДНК 601-ТА и центромеры затем клонировали в модифицированном векторе рUC19 с сконструированным асимметричным сайтом *AvaI*. Для крупномасштабного получения фрагментов ДНК плазмиды, которые содержат тандемные массивы, трансформировали в клетки DH5- $\alpha$  или XL1-Blue для стабильного размножения длинных тандемных массивов фрагментов ДНК. Очищенные плазмиды переваривали рестрикционным ферментом *AvaI*, и фрагменты очищали электрофорезом в агарозном геле и осаждали этанолом. Нуклеотидный состав четверехнуклеотидного выступа асимметричного сайта *AvaI* различается между верхней цепью (TS) и нижней цепью (BS), и, таким образом, может использоваться для мечения фрагментов ДНК, специфичных для цепи на 3'-конце с использованием радиоактивно или флуоресцентно меченных нуклеотидов в реакциях с использованием полимеразы Кленова. Поскольку ПЦР проблематична для получения АТ-богатого фрагмента центромеры, 5'-мечение любого конца потребовало бы химического повторного синтеза ДНК центромеры с дополнительными фланкирующими сайтами рестрикции и создания тандемных массивов для получения крупномасштабной ДНК; по этой причине маркировка 5' не проводилась.

*Подготовка и восстановление нуклеосом.* Экспрессия и очистка коровых гистонов *S. cerevisiae* проводились, как описано ранее [298; 379; 396]. Октамеры гистонов были восстановлены с использованием известных протоколов [298]. Вкратце, эквимольные количества очищенных рекомбинантных гистонов (H2A, H2B, Cse4 и H4) растворяли в денатурирующем буфере (7М гуанидин-НСl, 20 мМ трис-Сl, рН 7,5, 10 мМ дитиотреитол (ДДТ)) в концентрации 2 мг/мл. Смеси диализовали против четырех замен по 2 литра каждого буфера для рефолдинга (10 мМ Трис-Сl, рН 7,5, 1 мМ ЭДТА, 5 мМ бета-меркаптоэтанол, 0,1 мМ фенилметилсульфонилфторид (ФМСФ)), содержащий 2 М NaCl в течение 2 дней при 4°C. Затем смесь центрифугировали при 15 000 об / мин в микроцентрифуге Tomu MX-300 для удаления любого нерастворимого материала. Растворимые октамеры очищали фракционированием по размеру на колонке для гель-фильтрации Superdex 200. Фрагмент ДНК 601ТА 157 п.н. [375] и фрагмент ДНК CEN3 136 п.н. получали, как описано ранее [379]. Их полноразмерные последовательности представлены на оси последовательностей на всех соответствующих рисунках. Для воссоздания нуклеосом очищенные октамеры гистонового кора и ДНК смешивали в 50 мкл высокосолевого буфера (2М

NaCl, 10 mM Tris-Cl, pH 7,5, 1 mM EDTA, 0,02% NP-40, 5 mM бета-меркаптоэтанола) с добавлением бычьего сывороточного альбумина 400 мкг / мл. Смесь переносили в диализную установку SlideA-Lyzer MINI (Thermo Scientific). Блок для диализа помещали в контейнер с 600 мл высокосолевого буфера и диализовали в течение 30-60 минут с последующим диализом в солевом градиенте, в течение которого низкосолевым буфер (100 mM NaCl, 10 mM трис-Cl, pH 7,5, 1 mM ЭДТА, 0,02% NP-40, 2 mM бета-меркаптоэтанол) закачивали в контейнер со скоростью 3,5 мл / мин в течение 16 часов. Затем диализный блок переносили в буфер с низким содержанием соли и диализовали в течение 60 мин. Диализ проводили при комнатной температуре, а образцы дополнительно обрабатывали при 65°C в течение ночи.

*Гидроксильный футпринтинг.* Нуклеосомы собирали на основе фрагментов ДНК, меченных по 3'-концам с помощью <sup>33</sup>P-dТТР для верхней цепи и <sup>33</sup>P-dАТР для нижней цепи, путем заполнения асимметричного сайта AvaI с использованием полимеразы Кленова. HRF нуклеосом выполняли на нуклеосомах с помощью реагента железа (II)-EDTA, как описано ранее [391]. Продукты реакции разделяли на 1,3% нативном агарозном геле, полосы, содержащие свободные нуклеосомы, визуализировали с помощью окрашивания SYBR Green и вырезали из геля. ДНК выделяли из срезов геля и разделяли на 8% геле для секвенирования ДНК (National Diagnostics, каталожный номер EC-833). Маркерами подвижности ДНК были реакции секвенирования G + A и C + T тех же <sup>33</sup>P-меченных фрагментов ДНК, которые выполнялись, как описано в Molecular Cloning, CSHL. Гели запускали на 40-сантиметровой стеклянной пластине при 1500-1600 В в течение 70 минут при температуре геля около 55°C, затем переносили на фильтровальную бумагу DEAE (бумага для ионообменной хроматографии Whatman Grade DE81 от GE Life Sciences) и сушили под вакуумом. Радиоактивные сигналы регистрировались с помощью PhosphorImager (Fuji Photo Film Co., Ltd) и сканера Typhoon (Typhoon 9410, Amersham Biosciences).

## Вычислительный анализ

*Количественная оценка изображений гелей в экспериментах по футпринтингу гидроксильными радикалами.* Профили интенсивности 1D дорожек

для каждой дорожки в геле были извлечены с помощью ImageJ v. 1.51f [354]. Линия или ломанная линия была проведена вручную через центры всех полос в дорожке. Ширина линии была установлена примерно на половину ширины дорожки, а профиль сигнала вдоль полосы был извлечен и сохранен в отдельный файл для дальнейшей обработки. Эта стандартная процедура автоматически запускает выпрямление изображения для каждой полосы и усреднение данных по ширине полосы, как это реализовано в ImageJ. Профили интенсивности полос, соответствующие экспериментам с гидроксильными радикалами, были дополнительно проанализированы с помощью нашего недавно разработанного пакета HYDROID (доступного по адресу <https://github.com/ncbi/HYDROID>) (публикация в стадии подготовки), чтобы получить значения интенсивности расщепления ДНК. в каждой позиции в последовательности ДНК. Хотя наша методология основана на идеях, предложенных в более ранних работах [351; 352; 364], она имеет несколько новых аспектов. Во-первых, начальные местоположения пиков интенсивности, соответствующих полосам геля, были полуавтоматически назначены и сопоставлены с положениями вдоль последовательности ДНК путем сравнения их с полосами, полученными в результате реакций Максама-Гилберта [349]. Диапазон данных на каждом профиле HRF был установлен на область, в которой можно было идентифицировать непрерывный набор отдельных положений полос (либо непосредственно по наличию пика, либо иным образом однозначно выведенный из положения соседних пиков или соответствующих положений полос на других дорожках геля). Затем аналитическая функция была использована для аппроксимации экспериментального профиля HRF. Он представляет собой линейную комбинацию функций Гаусса, каждая из которых предназначена для описания формы определенной полосы на экспериментальном профиле. Алгоритм аппроксимации на основе метода наименьших квадратов Левенберга-Марквардта был использован для решения данной задачи. Параметры ширины функций Гаусса были ограничены, чтобы обеспечить монотонное уменьшение ширины с увеличением молекулярной массы расщепленных фрагментов ДНК, соответствующих полосам геля. Было показано, что такая регуляризация увеличивает точность процедуры аппроксимации и выполняется функцией, реализованной в HYDROID. Ранее функции Лоренца вместо функций Гаусса использовались в качестве эмпирического приближения для уменьшения искажений сигнала из-за авто-радиографического метода детекции сигнала в геле ([351] и приложение к

нему). В нашем случае моделирование с помощью функций Лоренца дало несколько худшие результаты, измеренные по среднеквадратическому отклонению между аппроксимированными и экспериментальными кривыми, что оправдывает использование функций Гаусса для нашего анализа. В результате частоты расщепления ДНК для каждой позиции в последовательности ДНК были получены из значений коэффициентов (полученных интегрированием по интенсивностям для каждой полосы) перед функциями Гаусса, описывающими соответствующую полосу. Следует отметить, что профили интенсивности расщепления ДНК обычно различаются по абсолютной величине из-за их зависимости от многих экспериментальных условий (загрузка образца, время воздействия и т. д.). Чтобы обеспечить адекватное сравнение двух экспериментальных профилей, они должны быть в одном масштабе. С этой целью мы выполнили линейную регрессию, описав один профиль HRF как линейную функцию (без свободного члена) другого профиля. Значения последних значений профиля интенсивности затем были масштабированы с помощью полученного коэффициента линейной регрессии и нормализованы от нуля до единицы путем деления на максимальное значение интенсивности двух профилей. Для сравнения экспериментальных и теоретических (H-SASA, определение см. ниже) профилей использовалась простая нормализация каждого профиля по его максимальному значению.

*Молекулярное моделирование и моделирование молекулярной динамики нуклеосом.* Моделирование молекулярной динамики (МД) различных нуклеосом было выполнено в соответствии с нашим протоколом, описанным в работе [A17] если не указано иное. Первоначальные модели были основаны на рентгеновских структурах, полученных из банка данных PDB, или на моделях созданных по гомологии. Первоначально структуры были ориентированы в системе координат нуклеосомы, как определено ранее [A17], так что ось Z соответствовала суперспиральной оси ДНК (рис. 3.7Б). Моделирование проводилось в течение 80 нс, и моментальные снимки сохранялись каждые 1 нс, первые 30 нс МД моделирования отбрасывались как период релаксации. Атомы оснований N1 и N9 в ДНК были привязаны к их исходным положениям с помощью гармонических ограничений с константой жесткости в 6 ккал / моль / Å<sup>2</sup>. Кроме того, мы сравнили МД-траектории с микросекундной траекторией нуклеосомы с полными гистоновыми хвостами и без ограничений на ДНК опубликованной ранее [A18]. Структуры были визуализированы и проанализированы с исполь-

зованием VMD [305]. Первоначальная структура центромерной нуклеосомы *S. cerevisiae* содержащей белок Cse4p с последовательностью ДНК 601ТА (601ТА-NUC) была построена на основе структуры нуклеосомы *X. laevis* с 601-ой ДНК (PDB ID: 3LZ0) с использованием Modeller [46] и 3DNA. Последовательности гистонов были получены из базы данных HistoneDB 2.0 [A14]. Модель центромерной нуклеосомы дрожжей на последовательности ДНК CEN3 (CEN3-NUC) была построена с использованием тех же методов на основе структуры нуклеосомы *X. laevis* с  $\alpha$ -сателлитной ДНК (PDB ID: 1KX5). Поскольку нет данных о позиционировании ДНК для этой нуклеосомы, мы определили его, используя наш метод определения диады, примененный к экспериментальным данным HRF для CEN3-NUC (см. ниже). В обеих моделях ДНК моделировали с использованием центрального фрагмента длиной 120 п.н. (60 п.н. от диады в каждом направлении) - фрагмента, который, как известно, однозначно организован центромерными нуклеосомами [397].

*Теоретическая оценка частотных профилей расщепления ДНК по атомным структурам.* Как предложено в [350], профили частоты расщепления нуклеосомной ДНК были теоретически оценены как сумма площадей, доступных для растворителя (SASA) для всех атомов водорода дезоксирибозы данного нуклеотида, в дальнейшем называемые профилями H-SASA. Расчеты SASA были выполнены с использованием программы NACCESS [398] с радиусом зонда 1,4 Å, шириной среза 0,005 Å и радиусами атомов, установленными в параметрах `rmin` силового поля CHARMM36 [299]. Для систем с усеченными гистоновыми хвостами хвосты были усечены в следующих N-концевых (H3G44, H4D24, H2A16T и H2BR33) и C-концевых (H2AK118) положениях (нумерация остатков дана в системе отсчета канонических гистонов *X. laevis*). По умолчанию профили H-SASA были усреднены по 50 снимкам МД (с интервалом каждые 1 нс) полностью гидратированных структур (по причинам, описанным в разделе “результаты”). Для моделирования МД атомы водорода добавлялись в ходе подготовки системы к моделированию, а в случаях прямого анализа рентгеновских структур положения атомов водорода генерировались с помощью REDUCEv.3.14 из AmberTools13 [365]. Графики профилей вместе с последовательностями были созданы с помощью ggplot [399] и TexShade [400]. Усечение гистоновых хвостов и релаксация структуры и усреднение, обеспечиваемые моделированием МД, сыграли важную роль в получении профилей H-SASA, которые отражали периодичность ДНК в нуклеосомах (см. Раздел “Результаты”).

ты”). Вклад контактов ДНК-белок в профили H-SASA оценивали путем расчета профилей H-SASA для нуклеосомной ДНК с удаленными гистонами. Можно видеть, что модель H-SASA чувствительна к сужению малой бороздки ДНК в определенных местах ДНК-гистоновых взаимодействий, где ДНК изгибается в сторону малой бороздки. Однако контакты гистон-ДНК вносят основной вклад в минимумы H-SASA на этих сайтах.

*Расчет параметров геометрии ДНК.* Параметры, которые представляют пространственную ориентацию, конформацию и периодичность нуклеосомной ДНК в виде одномерных профилей, являются важными инструментами нашего анализа. В качестве основного параметра, отражающего ориентацию пар оснований ДНК в нуклеосомной суперспирали, мы использовали величину относительного кручения нуклеосомной ДНК (relative twist, rTw). Два типа параметров кручения для характеристики суперспиральной ДНК в нуклеосоме были определены ранее: внутреннее кручение (iTW) и rTw (или локальное кручение) [401; 402]. Предполагается, что последний измеряется в экспериментах по HRF или ферментативному расщеплению, потому что он подчеркивает геометрически эквивалентные положения вдоль суперспирали ДНК. iTW, с другой стороны, выделяет только собственное вращение пар оснований относительно друг друга вдоль ДНК. Однако известно, что значения rTw и iTW связаны через сверхспиральный шаг нуклеосомной ДНК [402]. Значения iTW для каждого шага пары оснований были рассчитаны с использованием 3DNA [114]; суммирование этого значения по последовательности ДНК дало совокупное значение iTW (ciTw). Чтобы охарактеризовать rTw, мы вводим здесь новую величину - угол ориентации пары оснований (ВРОА), предназначенный для отражения локальной ориентации нуклеотидов относительно октамера гистонов. ВРОА был рассчитан из атомистических структур нуклеосом следующим образом: (i) суперспиральная ось нуклеосомы (рис. 3.7Б) была определена, как описано в работе [A17]; (ii) для каждого нуклеотида был определен вектор пары оснований (BPV) как вектор, указывающий от гликозидного атома азота текущего нуклеотида к соответствующему атому комплементарного нуклеотида (таким образом соединяя атомы N1 и N9); (iii) для каждого нуклеотида его соответствующий центр пары оснований (ВРС) определяли как среднюю точку BPV; (iv) ВРОА затем определяли как угол между BPV и перпендикуляром, направленным от суперспиральной оси нуклеосомы к ВРС (радиальный вектор в цилиндрической системе координат). Значения ВРОА служили мерой rTw в



диапазоне от 0 до 180°. По аналогии с  $ciTw$ , кумулятивная  $rTw$  ( $crTw$ ) была реализована путем суммирования вращения  $BPV$  вдоль последовательности и добавления 180° каждый раз, когда значение  $BPOA$  начинало новый интервал. Относительную (локальную) периодичность нуклеосомной ДНК в каждом конкретном положении рассчитывали на основе разницы значений  $crTw$  между сайтами на 5 п.н. выше и ниже выбранного положения. Альтернативный способ расчета  $rTw$  с использованием  $BPOA$ , определяемого как угол между  $BPV$  и плоскостью, перпендикулярной супспиральной оси, дал почти идентичные результаты, что подтвердило наш выбор  $BPOA$  в качестве меры  $rTw$ . Разница между  $crTw$  и  $ciTw$  наблюдалась в соответствии с теоретическими ожиданиями. Мы использовали параметр  $rTw$  для характеристики пути ДНК в атомистических структурах  $NSP$  и для расчета локальной периодичности вдоль последовательности ДНК. Это позволило нам подробно проанализировать вращательное позиционирование ДНК в нуклеосомах и идентифицировать геометрически эквивалентные положения относительно поверхности октамера. Минимумы кривой  $rTw$  соответствуют положениям, в которых соответствующие пары оснований ориентированы почти перпендикулярно сверхспиральной оси, а нуклеотид верхней цепи ДНК взаимодействует с октамером главным образом через его атомы основной цепи. Локальная периодичность ДНК для анализируемых структур (рисунки 3.9А и 3.10) колеблется между 10 и 11 п.н. на оборот, а ДНК чрезмерно скручена вокруг положений  $\pm 20$  и  $\pm 50$  ( $SHL \pm 2$  и  $\pm 5$ ), что находится в согласии с результатами статьи [166].

### 3.2.3 Результаты

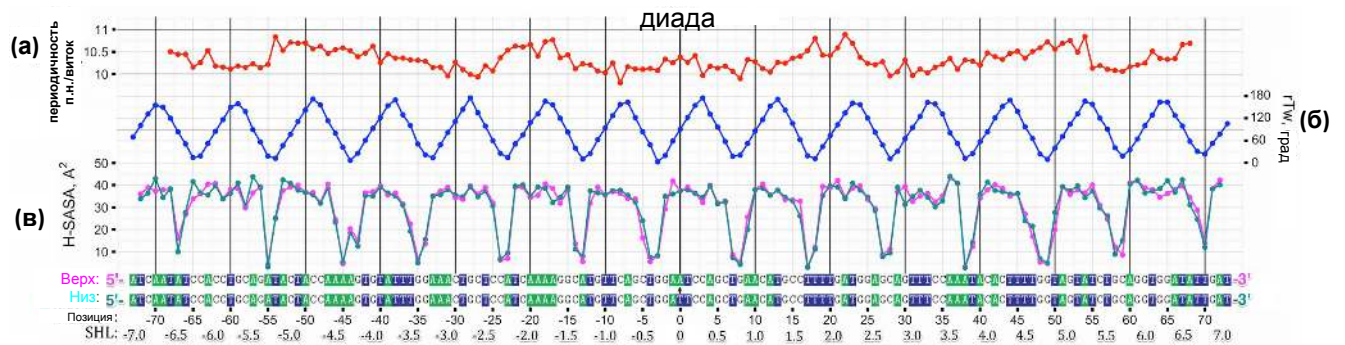
#### Анализ теоретических частотных профилей расщепления ДНК для разных нуклеосом

Расчет теоретических профилей частоты расщепления  $HRF$  основан на предположении, что он определяется профилями  $HSASA$  (см. Раздел “Материалы и методы”) [349]. Это, конечно, только приближение к сложным процессам диффузии и взаимодействия, происходящим во время реакции гидроксильно-

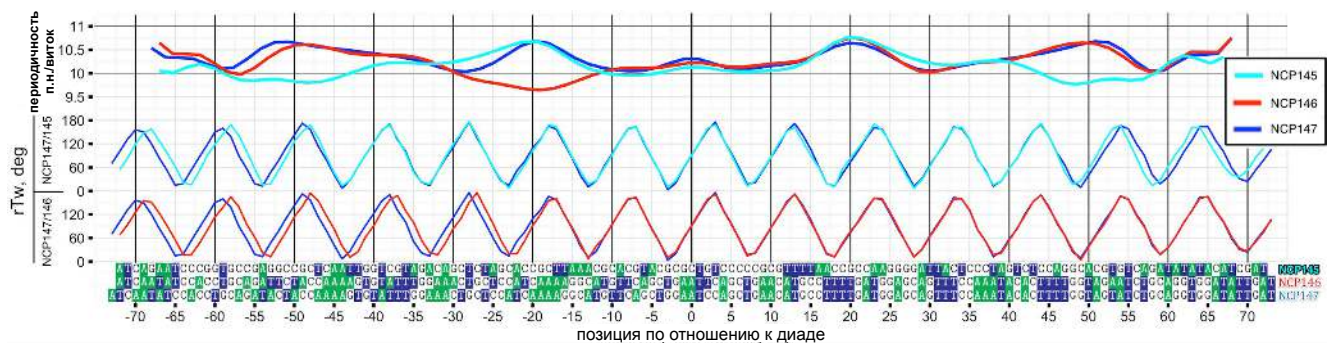
го футпринтинга. Однако есть определенные соображения в пользу такого подхода. Это позволяет надежно оценить теоретические профили с разрешением до одиночных нуклеотидов на основе атомных нуклеосомных структур или точных структурных моделей и сделать важные выводы об ожидаемой форме и сходстве экспериментальных профилей. Мы утверждаем, что симметрия нуклеосом может быть использована, чтобы обеспечить способ проверки и повышения точности теоретических профилей, облегчить интерпретацию экспериментальных профилей HRF и позволить идентифицировать диаду нуклеосом из экспериментальных данных HRF. Действительно, нуклеосомы обладают осью псевдосимметрии второго порядка (рис. 3.7Б-Г), которая связывает верхнюю и нижнюю цепи нуклеосомной ДНК таким образом, что если нуклеосома повернута на  $180^\circ$  вокруг оси диады, то верхняя цепь ДНК почти точно совпадет с положением нижней цепи и наоборот. В частности, 5'-конец верхней цепи будет совмещен с 5'-концом нижней цепи. Сначала мы исследовали структуру с самым высоким разрешением (PDB ID: 1KX5 [238]), где симметрия октамера гистонов дополняется симметрией квазипалиндромной последовательности ДНК: центральная пара оснований на диаде подразделяет нуклеосому и двухцепочечную ДНК на две идентичные половины. В случае известных структур точное положение диады можно непосредственно рассматривать как единственную пару оснований, расположенную на оси псевдосимметрии [165]. На рис. 3.9 показана четкая картина периодичности профилей H-SASA, рассчитанных на основе этой структуры для обеих цепей ДНК. Эта форма и периодичность очень похожи для обеих цепей ДНК, учитывая симметрию системы. Ось диады делит сегмент между двумя минимумами на графике H-SASA на неравные части (примерно в соотношении 3:7) с расположением диады ближе к 5' концу цепи ДНК. Это является следствием геометрии нуклеосомы и проиллюстрировано на рисунках 3.8Б и В. На рисунке 3.9В профиль H-SASA имеет отчетливые резкие минимумы в каждой позиции, где ДНК ориентирована перпендикулярно поверхности октамера, и довольно похожие на плато области высоких значений подверженности потенциальной атаке гидроксильных радикалов между этими позициями. Расстояние между минимумами на профиле H-SASA соответствует периодичности 10–11 п.н., но положения минимумов H-SASA не всегда точно соответствуют положениям минимумов  $\gamma Tw$  - они обычно отклоняются на 1 или 2 позиции в любом направлении. Последнее подчеркивает тот факт, что не только кручение ДНК отвечает за точную форму профилей H-SASA, но

и различия во взаимодействиях белок-ДНК в различных сайтах связывания ДНК на нуклеосомах влияют на локальные положения минимумов H-SASA. Чтобы получить представление об этих вариациях, мы рассчитали индивидуальный вклад аминокислот в защиту ДНК от гидроксильных радикалов, рассчитанный как изменение общего H-SASA после удаления одного определенного аминокислотного остатка. Хорошим примером локального изменения формы H-SASA является повышенная защита в профилях в положении -43, проявляющаяся в виде локального минимума на рисунке 3.9В, который вызван взаимодействиями ДНК с  $\alpha$ N-спиралью, уникальной для H2A. Несмотря на наличие локальных выбросов, общие профили H-SASA двух цепей ДНК демонстрируют хорошую суперпозицию. Как упоминалось ранее, это результат почти идеальной симметрии данной нуклеосомы как на уровне белка, так и на уровне ДНК. Эксперименты HRF проводятся в растворе на ансамбле нуклеосом, которые в совокупности представляют конформационное пространство, доступное отдельным нуклеосомам в течение длительных периодов времени, поэтому уместно предположить, что две идентичные копии гистонов и хвостов гистонов будут исследовать неразличимые ансамбли конформаций. Следовательно, введенное выше правило симметрии должно также соблюдаться. Гибкие гистоновые хвосты, которые часто встречаются в адсорбированных конформациях в структурах рентгеновских нуклеосом (если они разрешены) могут вносить асимметрию в профиль H-SASA и могут обеспечить значительную дополнительную защиту ДНК от расщепления. Мы обнаружили, что наилучшее совпадение профилей H-SASA, соответствующих двум нитям ДНК были получены, когда гибкие гистоновые хвосты были опущены и ансамблевое усреднение и релаксация остова ДНК была выполнена в ходе МД моделирования (Рисунок 3.9В). По этой причине мы реализовали протокол с использованием систем с усеченными гистоновыми хвостами и усреднение по ансамблю МД, который мы использовали на протяжении всего исследования. Далее мы проанализировали влияние вариаций последовательности ДНК на геометрию ДНК и теоретические H-SASA профили. Рисунок 3.10 суммирует различия в геометрии ДНК и периодичности для трей доступных рентгеновских структуры нуклеосомного кора с разной ДНК последовательностью:  $\alpha$ -сателлитная ДНК длиной 146 п.н. (“NCP146”) [165], модифицированная  $\alpha$ -сателлитная ДНК длиной 147 п.н. (“NCP147”) [238] и 601 высокоаффинная последовательность Видома длиной 145 п.н. (“NCP145”) [376; 403]. Во всех этих структурах одна пара нуклеотидов

расположена строго на диаде и ДНК проходит один и тот же сверхспиральный путь, несмотря на различия в их длине. Это достигается за счет растяжения/сжатия ДНК на участках, расположенных на расстоянии  $\pm 20$  или  $\pm 50$  п.н. от диады. Это иллюстрируется относительной периодичностью параметра  $rTw$ , предложенного в этом исследовании (рис. 3.10, верхний график). В то время как молекулы ДНК из NCP147 и NCP145 принимают очень похожие конформации с обеих сторон нуклеосомы, ДНК из NCP146 растянута на 1 п.н. с одной стороны. Практически идентичная конформация ДНК по обе стороны диады в NCP147 и NCP145 приводит к очень хорошему наложению между профилями H-SASA для обеих нитей ДНК (коэффициент корреляции Пирсона = 0,97, Рисунок 3.9). Что касается NCP146, несмотря на небольшую асимметрию двух половин нуклеосомной ДНК, сходство профиля H-SASA между двумя цепями поддерживается в пределах  $\pm 20$  п.н. вблизи диады. Однако небольшая асимметрия NCP146 не будет проявляться на профилях HRF для ансамбля нуклеосом в растворе. На самом деле, NCP146, имеющая чисто палиндромную последовательность ДНК, не будет иметь предпочтений относительно того, с какой стороны нуклеосомной ДНК растягиваться или расширяться, потому что две половинки ДНК полностью идентичны.



**Рисунок 3.9** — Структура ДНК и профили H-SASA в рентгеновской структуре самого высокого разрешения NCP (NCP147). (А) Локальная периодичность вращения ДНК в нуклеосоме, рассчитанная по относительному повороту ( $rTw$ ). (Б)  $rTw$  нуклеосомной ДНК вдоль последовательности, нанесенный в направлении 5'-3' верхней цепи (TS). (В) Профили H-SASA для верхней цепи (пурпурный) и нижней цепи (голубой), оба профиля и последовательности под ними даны в направлении 5'-3'. Профили H-SASA рассчитаны по МД траектории NCP147 без гистоновых хвостов (см. Раздел “Материалы и методы”).



**Рисунок 3.10** — Различия в структуре ДНК между разными рентгеновскими структурами нуклеосомы. Вверху: то же, что на рис. 3.9А, но сглаженный В-сплайном (с 30 степенями свободы). Внизу: попарное сравнение профилей  $rTw$ .

### Количественная оценка экспериментальных профилей HRF и их сравнение с теоретическими профилями H-SASA

Мы провели ряд экспериментов по гидроксильному футпринтингу для собранных *in vitro* октамерных нуклеосом на основе центромерных гистонов *S. cerevisiae* и двух различных последовательностей ДНК: высокоаффинной последовательности 601ТА и центромерной ДНК последовательность третьей хромосомы пекарских дрожжей CEN3 (601ТА-NUC и CEN3-NUC). Кроме того, мы проанализировали свободную CEN3 ДНК (CEN3-free). Согласно нашей вычислительной процедуре, каждому положению в последовательности ДНК может быть присвоено значение относительной частоты расщепления гидроксильными радикалами путем количественной оценки экспериментально полученного изображения геля ПААГЭ для продуктов реакции HRF (Рисунок 3.11). В качестве доказательства концепции на рисунке 3.11В изображен экспериментальный профиль HRF 601ТА-NUC по сравнению с теоретическим профилем H-SASA, полученным из модели высокого качества с известной позицией диады созданной на основе рентгеноструктурных исследований (раздел “Материалы и методы”). Общая периодичность экспериментальных и теоретические профилей хорошо согласуются между собой. Заметное уменьшение амплитуды сигнала к 3' концу ДНК, вероятно, связано с пониженной склонностью коротких фрагментов ДНК к осаждению на стадии очистки. Экс-

периментальные профили для верхней и нижней цепи ДНК при наложении в 5'-3' направлениях (рис. 3.11Г) очень похожи по форме, отражая псевдосимметрию нуклеосомы, как показано в предыдущем разделе.

Расположение минимумов в сайтах ДНК-гистоновых взаимодействий на H-SASA и экспериментальных профилях совпадают с точностью до  $\pm 1$  п.н., в большинстве случаев экспериментальные минимумы смещены к 5'-концу цепи по отношению к минимумам H-SASA. Возможное объяснение состоит в том, что во время расщепления гидроксильным радикалом 3'- меченой цепи ДНК в дополнение к основному продукту с 5'-фосфатным концом, как известно, присутствует второстепенный альтернативный продукт - цепь, оканчивающаяся 5'-альдегидной группой [350]. 5'-альдегидный продукт, который образуется в результате отрыва 5'-атома водорода, на 1 нуклеотид длиннее, чем цепь с концевым 5'-фосфатным концом, упомянутая выше, и она не имеет отрицательного заряда фосфатной группы и, таким образом, имеет подвижность в геле на 2-3 полосы медленнее по сравнению с соответствующим продуктами реакций Максама - Гилберта [377]. Таким образом, в фактически измеренном профиле частоты расщепления ДНК ожидается, что некоторая часть сигнала будет сдвинута на 2-3 пика в сторону 5'-конца. Это может привести к небольшому смещению минимумов на измеренном профиле относительно истинного профиля (аппроксимированного профилем H-SASA). Упомянутый эффект должен быть значительно снижен, если нить ДНК радиоактивно помечена на 5'-конце [350], но создание 5-меченых фрагментов центромерной ДНК с высоким содержанием АТ является технически сложной задачей (раздел "Материалы и методы").

Есть и другие важные различия между теоретически оцененными профилями H-SASA и экспериментальными. Последние демонстрируют гораздо более плавное изменение значений частоты расщепления вдоль последовательности, в то время как профили H-SASA имеют резкие минимумы шириной 1-2 п.н., окруженные областями с высокими значениями частоты расщепления (Рисунок 3.11В). Кроме того, хотя в определенных положениях значения H-SASA близки к нулю, что указывает на высокую защиту ДНК от расщепления, экспериментальные данные подтверждают минимальную частоту расщепления по меньшей мере в размере 25% относительно максимальных значений. В дополнение к этому, экспериментальные профили HRF имеют отчетливые локальные максимумы, которые соответствуют ДНК, обращенной от октамера, в то время

как профили H-SASA имеют форму плато в соответствующих областях. Возможные причины этих расхождений указаны в разделе “Обсуждение” ниже.

Мы наблюдали, что общий синусоидальный профиль футпринтинга ДНК в нуклеосомах имеет выбросы в определенных положениях пар оснований, их величина варьируется от эксперимента к эксперименту. Эти выбросы лучше всего проиллюстрированы на наборах данных нуклеосом CEN3-NUC и выделены стрелками на рисунке 3.12. Интересно, что расположение этих выбросов очень хорошо соответствует положениям выбросов на свободной ДНК. HRF свободной ДНК определяется в основном ее геометрией, которая зависит от последовательности. Необычная природа последовательности CEN3, которая очень богата A-трактами (определяемыми как четыре или более последовательных пары оснований A-T без шага TrA [404]), проявляется в виде пиков в местах, где эти A-тракты нарушены гуанинами или цитозинами. Расположение основных пиков, наблюдаемых экспериментально для свободной ДНК CEN3, подтверждается прогнозами веб-сервера ORCHID, который предсказывает профили HRF путем усреднения существующих измерений HRF для всех возможных свободных тринуклеотидов [371]. Тем не менее, не следует ожидать идеального соответствия между этими двумя методами, поскольку A-тракты характеризуются нелокальными разветвленными водородными связями, которые могут приводить к неаддитивным далекодействующим, а не локальным эффектам, проявляемым в тринуклеотидах [405]. Появление аналогичных выбросов на HRF профилях нуклеосомальной ДНК, вероятно, связано с взаимодействием между внутренней геометрией и динамикой свободной ДНК и геометрией, наложенной на ДНК путем связывания с гистонами. Наше моделирование теоретических профилей расщепления с использованием МД-моделирования предполагает, что значения H-SASA довольно чувствительны к небольшим колебаниям двугранных углов основной цепи ДНК, которые могут быть вызваны только небольшими смещениями положений пары оснований. Это дает один потенциальный ключ к пониманию того, как небольшие изменения в структуре и динамике нуклеосом из-за вариации последовательности ДНК могут проявляться в профилях HRF.

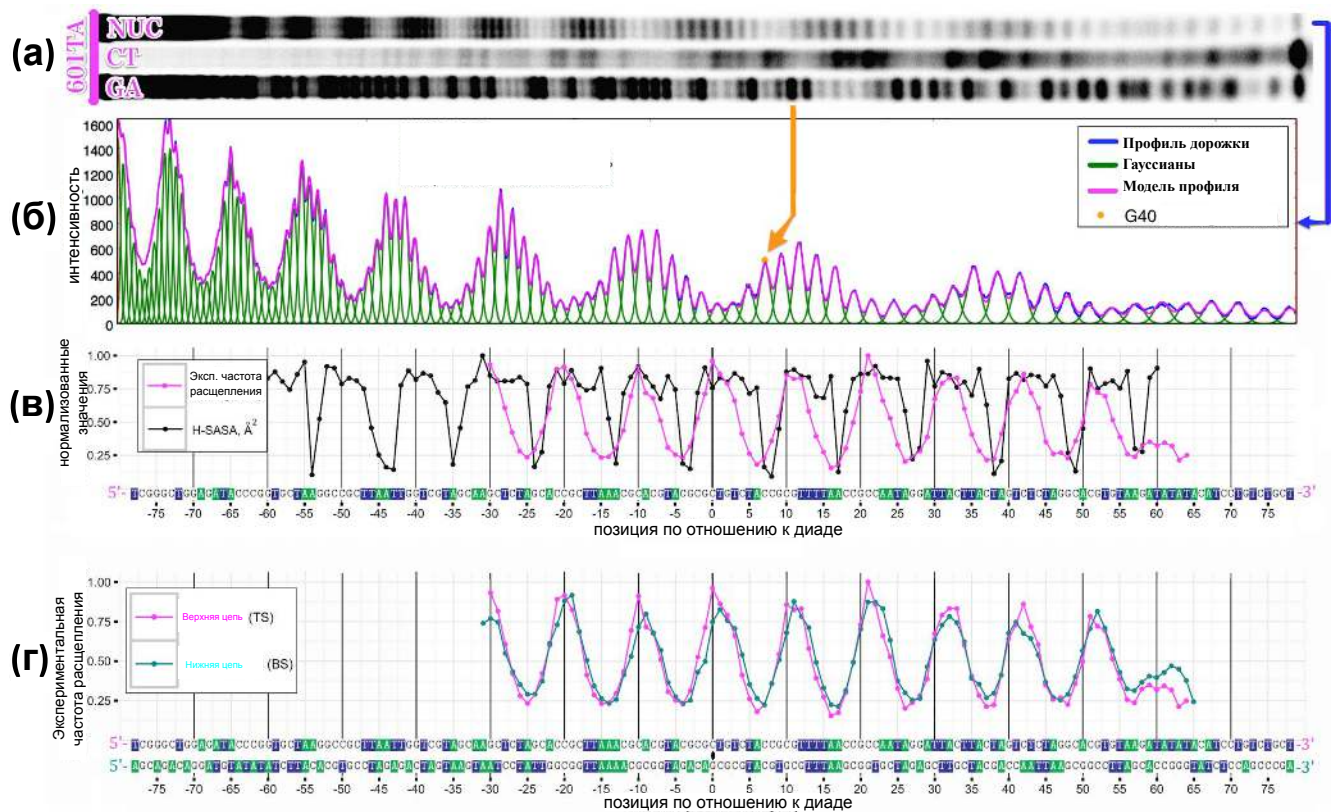
## Новый метод определения положения нуклеосомной ДНК на основе данных гидроксильного футпринтинга

Анализ как экспериментальных профилей, так и профилей H-SASA, описанных в предыдущих разделах, выявил несколько ключевых взаимосвязей между формой профилей HRF и положением ДНК на нуклеосоме. Описанный ниже метод идентификации диады основан на этих наблюдениях. Ось псевдосимметрии второго порядка нуклеосомы (рисунок 3.7) связывает верхнюю и нижнюю цепи нуклеосомной ДНК так, что если нуклеосома повернута на  $180^\circ$  вокруг диадной оси, то верхняя цепь ДНК накладывается на симметричное положение нижней цепи и наоборот (Рисунок 3.7В). Это наблюдение подразумевает, что профили HRF для верхней и нижней цепей ДНК должны совпадать (или быть очень похожими в зависимости от степени псевдосимметрии, как обсуждалось ранее), если оба профиля сравнивать в том же направлении (в нашем случае от 5' к 3'). На основе H-SASA и экспериментальных профилей HRF, а также анализе геометрии структуры нуклеосомы, можно увидеть, что положение диады расположено асимметрично между двумя локальными минимумами на соответствующем профиле любой из нитей ДНК. Диада делит отрезок между минимумами примерно в соотношении 3:7 (далее именуемое “правилом соотношения 3:7”) (Рисунок 3.7В). Важным аспектом является выбор правильных систем отсчета для сравнения профилей верхней и нижней цепей. Это особенно актуально, поскольку в профилях HRF обычно отсутствуют данные около концов нуклеосомной ДНК из-за ограниченного разрешения ПААГЭ, что делает невозможным идентифицировать нуклеосомную границу на каждом профиле и использовать ее в качестве маркера для их выравнивания. Однако, если известно положение диады в ДНК, также известны положения соответствующих нуклеотидов диады на каждом профиле. В этом случае, из-за геометрических причин и соображений комплементарности оснований цепи, правильный выбор систем отсчета может быть достигнут путем размещения положений нуклеотидов диады каждого профиля в 0, как показано на рисунке 3.7Г. Это проиллюстрировано на примере системы 601TA-NUC (рис. 3.11Г), где известное положение диады из структуры (см. Раздел “Материалы и методы”) используется для построения экспериментальных профилей HRF для верхней и нижней цепей ДНК одной и той же системы в их соответствующих системах

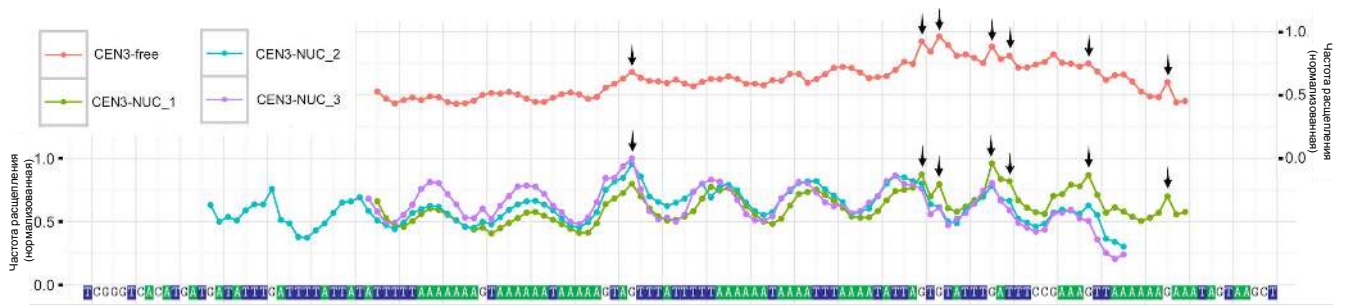


отсчета. Два профиля выглядят идеально выровненными (коэффициент корреляции Пирсона = 0,94), что подтверждает правильный выбор положения диады. Если теперь представить себе ситуацию, когда положение диады неизвестно, но есть эксперименты по гидроксильному футпринтингу для обеих цепей ДНК, описанный выше подход можно легко обобщить для определения положения диады и, следовательно, общего позиционирования ДНК (также называемого позиционированием нуклеосом). Поскольку каждое значение интенсивности профиля HRF соответствует определенному основанию на цепи ДНК, выравнивание профиля определяет выравнивание последовательностей двух цепей, нанесенных на график в направлении от 5' к 3'. Соответствующее положение диады на выравнивании последовательностей между верхней и нижней цепями ДНК затем может быть идентифицировано как единственное положение, где выровненные основания образуют пару оснований в структуре двухцепочечной нуклеосомной ДНК. На практике производится выборка различных совмещений профилей HRF для верхней и нижней цепей ДНК, и качество их соответствия оценивается путем вычисления коэффициента корреляции между двумя выровненными профилями HRF (только перекрывающиеся части профилей используются для расчета коэффициента). Такой подход вдохновлен теорией обработки сигналов, в которой сходство между двумя процессами оценивается путем вычисления функции взаимной корреляции. Положение диады должно соответствовать максимуму (хотя бы локальному максимуму) коэффициента корреляции. Корреляция между двумя профилями HRF как функция предполагаемого положения диады для данных 601TA-NUC показана на рисунке 3.13А. На этом рисунке есть несколько максимумов коэффициента корреляции, разнесенных периодически на 5 п.н., и в некоторых случаях может быть неоднозначность относительно того, какой максимум соответствует истинному положению диады. Применение “правила 3:7” полезно, поскольку половину этих максимумов корреляции можно отбросить как стереохимически “запрещенные” диадные позиции. Таким образом можно идентифицировать “настоящую” позицию диады (рис. 3.13). Упомянутая периодичность и потенциальная неоднозначность являются следствием (i) квазипериодической природы конформации ДНК в нуклеосоме и (ii) ограниченного диапазона положений в последовательности ДНК, для которых измеряли профиль расщепления ДНК. Однако на практике приблизительное положение диады (например, найденное путем определения местоположения границ нуклеосом посредством расщеп-

ления МНКазой) может использоваться для разрешения неоднозначности. В качестве альтернативы можно использовать более длинные профили HRF, которые пересекают границы нуклеосом. Точность нашего метода явно зависит от качества и отношения сигнал / шум данных. Например, для нуклеосом 601ТА-NUC из данных, полученных в этом исследовании, положение диады может быть воспроизведено с точностью до 0,5 п.н. (рис. 3.13А). Более высокая точность разрешения по сравнению с одной парой оснований связана с тем фактом, что диада потенциально может быть расположена не только в определенном положении пары оснований, но также между двумя соседними парами оснований, хотя последняя возможность не была замечена в рентгеновских структурах до сих пор. Однако стоит отметить, что положение диады в нуклеосомах в растворе является внутренне статистической характеристикой, и для палиндромных последовательностей ожидается, что среднее положение будет находиться между двумя парами оснований (как обсуждалось ранее для NCP146). Важно отметить, что наш подход к идентификации диады устойчив по отношению к систематическим экспериментальным ошибкам, например, возникающим из-за минорных продуктов расщепления ДНК, как описано в предыдущем разделе. Эти смещения обычно одинаковы для верхней и нижней цепей, и сходство конформации ДНК в нуклеосоме из-за псевдосимметрии все же должно приводить к подобию экспериментально измеренных профилей расщепления.



**Рисунок 3.11** — Количественная оценка экспериментов по гидроксильному футпринтингу и сравнение с теоретическими профилями, полученными на основе атомной структуры (для нуклеосомы 601TA-NUC). (А) Исходное изображение геля ПААГЭ для сегментов ДНК, полученных с помощью гидроксильного футпринтинга нуклеосом, восстановленных на последовательности 601TA (NUC), и соответствующих продуктов реакции Максама-Гилберта (СТ, GA). Данные показаны только для верхней цепи. (Б) соответствующий необработанный профиль интенсивности дорожек HRF, извлеченный из изображения геля, и его деконволюция в индивидуальные интенсивности полос путем подбора функций Гаусса; приводятся среднеквадратическое отклонение (RMSD) и относительное RMSD между исходным профилем и подобранной моделью; также указано RMSD для площади пика, площади пиков рассчитываются как площади под кривой между соседними минимумами. (В) суперпозиция количественных экспериментальных частот расщепления ДНК и профилей H-SASA, рассчитанных на основе атомистической структуры; оба профиля нормализованы до своих максимальных значений. (Г) Суперпозиция экспериментальных частот расщепления ДНК для верхней и нижней цепей ДНК в нуклеосоме 601TA-NUC

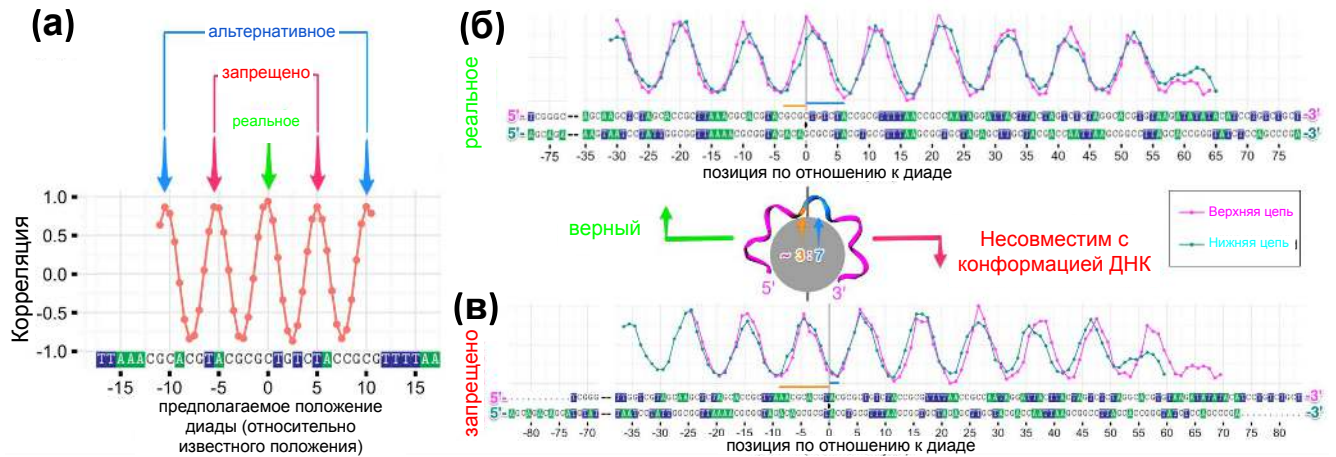


**Рисунок 3.12** — Профили частоты расщепления ДНК для нуклеосом CEN3-NUC и ДНК свободной CEN3 ДНК. Данные нескольких экспериментов для верхней цепи объединены, стрелки указывают на незначительные всплески, видимые как на свободных, так и на нуклеосомных профилях ДНК, приписываемые зависимым от последовательности вкладом.

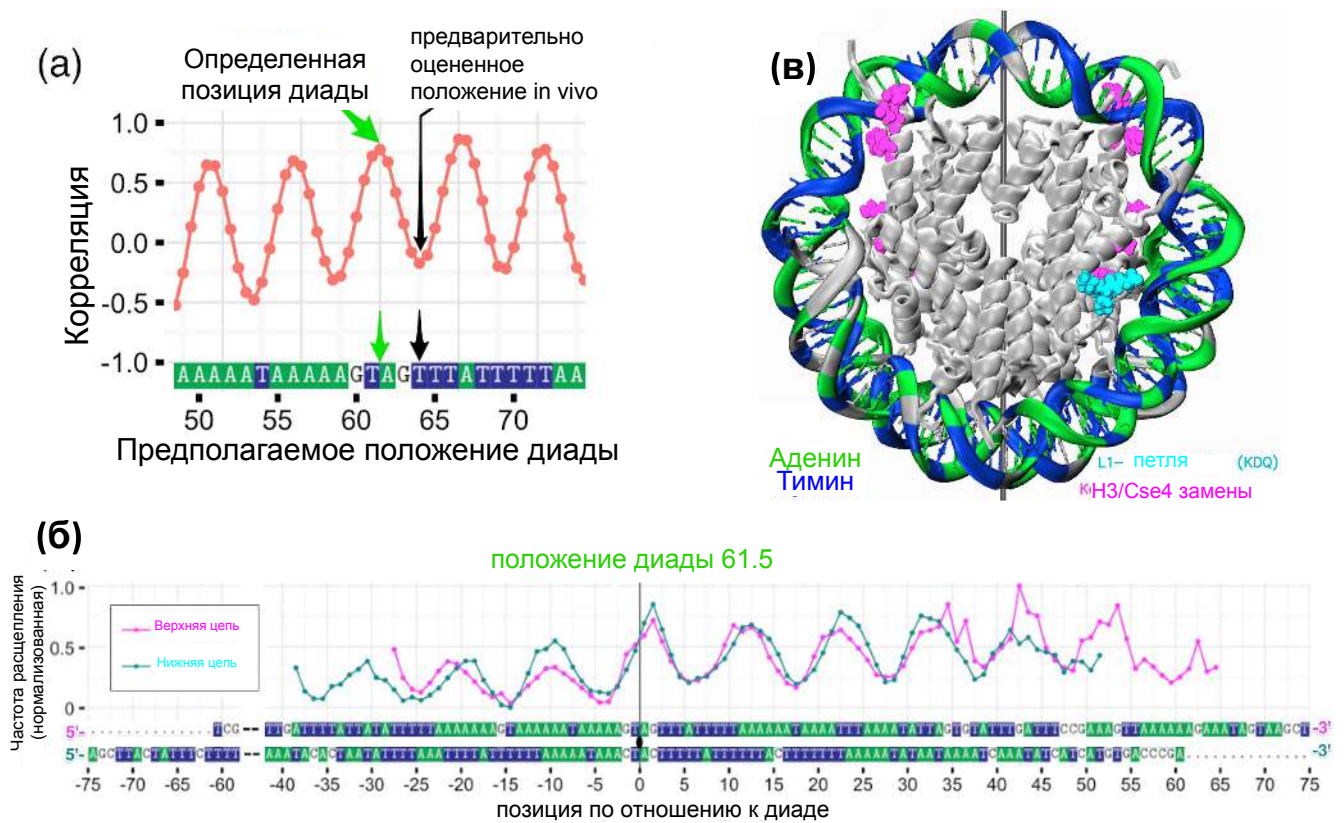
### Моделирование нуклеосомы дрожжей CEN3 помощью интегративного подхода

Мы использовали метод, описанный в предыдущем разделе, чтобы определить положение диады в восстановленной *in vitro* центромерной нуклеосоме *S. cerevisiae* из хромосомы III (CEN3-NUC) и построить ее структурную модель. С этой целью были проанализированы профили HRF для обеих нитей ДНК, полученные в результате трех независимых экспериментов (рисунок 3.12), в соответствии с предложенным нами методом. Мы проиллюстрируем применение нашего подхода к этому набору данных ниже. Были взяты образцы и протестированы различные предполагаемые положения диады нуклеосомы, как описано ранее. Специфическая интерес представляла область вблизи положения 64, которое, как предполагалось в более ранних экспериментах MNase-seq, была центром области устойчивости к расщеплению MNказой [393]. Корреляционный анализ (рис. 3.14А и Б) выявил несколько возможных позиций диад с высокими уровнями корреляции между выровненными профилями (позиции 61,5, 67 и 72). Позиция 67 была отклонена как нарушающая соотношение “правило 3:7”, что становится очевидным после просмотра выравнивания профилей HRF. Были еще две возможные позиции 61,5 и 72 с одинаковым соотношением коэффициента  $R = 0,78$ . Однако позиция 61,5 обеспечила лучшее соответствие между профилями верхней и нижней цепи ДНК в регионе вокруг диады (и,

таким образом, ее можно считать лучшим кандидатом для расположения диады). Положение 61.5 расположено намного ближе к нашей исходной области интереса, основанной на оценке положения диады *in vivo* (положение 64). Все доступные рентгеновские структуры нуклеосом указывают на положение диады на конкретном нуклеотиде, а не между ними. В то же время, как обсуждалось ранее, среднее расположение диады в полуцелом положении возможно для ансамбля нуклеосом в растворе, особенно в случае палиндромных последовательностей. Анализ коэффициента корреляции указал на позицию 61 как более предпочтительную (рис. 3.14А), и ее в дальнейшем использовали для построения точной модели нуклеосомы CEN3-NUC (рис. 3.14В). Знание точного местоположения диады позволяет впервые увидеть пространственные отношения между ключевыми элементами ДНК и белковыми элементами, которые определяют функцию центромерных нуклеосом у дрожжей. Как А-участки ДНК, так и ключевые остатки гистона Cse4p, как было показано ранее, являются критическими для функции центромеры у дрожжей, однако их коллективное взаимодействие остается не вполне понятным [406; 407].



**Рисунок 3.13** — Объяснение алгоритма идентификации диады применительно к нуклеосоме 601TA-NUC с известным положением диады. (А) Коэффициент корреляции Пирсона между экспериментальными профилями HRF верхней (TS) и нижней (BS) цепей ДНК как функция предполагаемого положения диады, используемого для выравнивания этих профилей. Реальное положение диады соответствует одному из локальных максимумов. Только половина локальных максимумов совместима со стереохимически допустимым решением (“правило 3: 7”). (Б) Наложение профилей TS и BS, выровненных с использованием реальной (известной) позиции диады. (В) Наложение профилей TS и BS, выровненных с использованием стереохимически запрещенного предполагаемого положения диады.



**Рисунок 3.14** — Применение алгоритма идентификации диады к нуклеосоме SEN3-NUC с неизвестным положением диады. (А) коэффициент корреляции Пирсона между экспериментальными профилями HRF верхней (TS) и нижней (BS) цепей ДНК как функция предполагаемого положения диады; показано положение диады, согласующееся с данными HRF (зеленая стрелка) и ранее оцененное *in vivo* положение по устойчивости к расщеплению микрококовой нуклеазой (черная стрелка) [393]. (Б) Наложение профилей TS и BS, выровненных с использованием идентифицированной позиции диады. (В) Модель нуклеосомы SEN3-NUC, построенная с использованием положения ДНК, идентифицированного в экспериментах с HRF (диада установлена в положение 61). Нити ДНК окрашены в соответствии с их последовательностью, выделяя АТ-участки: А — зеленый, Т — синий, G или С — серый.

### 3.2.4 Обсуждение

В этой работе мы провели эксперименты по гидроксильному футпринтингу на двух нуклеосомных системах и разработали вычислительную основу для анализа данных HRF, полученных в результате экспериментов с нуклеосома-

ми в растворе; интерпретации данных HRF путем сравнения с теоретическими профилями H-SASA; определения точного положения ДНК по данным HRF. Используя расширенный анализ структур нуклеосом, в сочетании с количественной оценкой экспериментальных данных HRF, мы смогли предложить интерпретацию экспериментов HRF на концептуально новом уровне. Это, в свою очередь, привело к развитию простого метода определения положения диады нуклеосом с разрешением в одну пару оснований. Мы использовали этот метод для определения положения диады в реконструированной *in vitro* центромерной нуклеосоме *S. cerevisiae* и для построения ее структурной модели. Алгоритм идентификации диады, предложенный в этом исследовании, основан на соотношении симметрии между двумя цепями нуклеосомной ДНК. Мы показали, что это требование симметрии прямо выражается в сходстве профилей HRF верхней и нижней цепей ДНК. Использование этого критерия и экспериментальных данных для двух цепей ДНК повышает точность и надежность метода идентификации диады по сравнению с ситуацией, когда профиль HRF для каждой цепи анализируется отдельно. Последний подход до сих пор использовался в нескольких предыдущих исследованиях, которые в основном прибегали к сообщению приблизительного местоположения диады на ДНК [374; 385; 389; 408; 409]. Идентификация положения диады с разрешением в одну пару оснований ранее достигалась сайт-направленным расщеплением гидроксильными радикалами или сайт-специфическим фотохимическим сшиванием [410; 411]. Однако эти методы требуют химического включения определенных зондов в октамер гистонов в симметричных положениях, и диада затем определяется как средняя точка между реакционными сайтами. Метод, описанный в данной работе, обеспечивает разрешение в одну пару оснований с использованием обычных экспериментальных данных HRF без необходимости дорогостоящих модификаций гистонов. В качестве доказательства концепции мы оценили расположение ДНК в *in vitro* нуклеосоме CEN3-NUC, которое оказалось на расстоянии 3 п.н. от ранее оцененного положения *in vivo* центра области, устойчивой к расщеплению микрококовой нуклеазе [393]. Точно определенное положение диады позволило нам построить модель этой нуклеосомы и выявить пространственную взаимосвязь между ориентацией последовательности ДНК и ключевыми белковыми мотивами белка Cse4p (Рисунок 3.14). Наша модель может служить основой для будущих исследований, направленных на понимание кооперативного взаимодействия и взаимодействия между белком



Cse4p и A-траками последовательностей CEN центромер хромосом пекарских дрожжей. В настоящее время известно, что последовательности CEN не консервативны между разными хромосомами у дрожжей, но наличие в их составе A-трактов имеет решающее значение для выживания [406]. Однако молекулярный механизм(ы) вовлечения A-трактов в функцию центромеры дрожжей остается неясным. Одним из аспектов может быть особая геометрия ДНК, присущая A-трактам, которые известны своей конформационной жесткостью и узкими малыми бороздками как в свободной ДНК [405], так и в нуклеосоме [412]. В отличие от обычного гистона H3, вариант Cse4p не имеет остатка аргинина (замещенного серином R63S153), который в случае H3 вставлен в малую бороздку ДНК около положения  $\pm 15$ . Известно, что это положение примыкает к уникальному гистоновому мотиву, который формирует чрезвычайно узкую малую бороздку через гидрофобный “сахарный зажим” [413]. С другой стороны, известно, что мотивы TnAn (в отличие от AnTn и в отличие от A-трактов) обнаруживают расширенные малые бороздки на динуклеотидах TrA [404]. Последовательности CEN с высоким содержанием AT обычно имеют несколько A-трактов, которые разделены динуклеотидами TrA, и, таким образом, последовательности CEN также содержат несколько TnAn-подобных мотивов. Эти мотивы могут способствовать функционально важному расширению малой бороздки на соответствующих участках. Интересно, что согласно нашей структурной модели центромерной нуклеосомы, последовательность CEN3 является квазипалиндромной рядом с определенным положением диады 61,5. Последовательность AAAGTAGTTT, обнаруженная на диаде, может быть преобразована в идеальный палиндром с заменой только 1 нуклеотида. Кроме того, 4 нуклеотида в диаде (GTAG) фланкированы сегментами ДНК, состоящими исключительно из А или Т, богатых А-участками. Эти структурные данные могут дополнительно обеспечить понимание структуры центромерных нуклеосом и функционального связывания с ключевыми кинетохорными белками, такими как дрожжевой Mif2 (человеческий CENP-C), который предпочтительно связывается с AT-богатыми ДНК центромерных нуклеосом дрожжей (Xiao et al., Неопубликованные данные).

Помимо определения положения диады, эксперименты HRF могут быть использованы для дальнейшего углубления нашего понимания конформации ДНК в нуклеосомах. Возможность сделать это, однако, зависит от нашей способности сопоставить форму профилей HRF со структурными деталя-

ми нуклеосомы. Понимание расхождений между теоретически полученными профилями H-SASA и экспериментальными профилями HRF для известных структур является важным шагом в этом направлении. Наш сравнительный анализ профилей H-SASA и экспериментальных профилей показал, что, хотя их общая периодичность схожа, экспериментальные профили намного более гладкие и могут иметь разную величину и / или точное расположение максимумов и минимумов. Разница между теоретическими и экспериментальными профилями также была замечена в предыдущем исследовании комплекса ДНК-ТВР: значения H-SASA для некоторых нуклеотидов были близки к нулю, в то время как определенное количество расщеплений все же было обнаружено экспериментально [394]. Значительная вероятность расщепления ДНК на сайтах связывания ДНК может быть результатом нелинейной зависимости частоты расщепления от доступности растворителю. Фактически, предыдущие исследования показали обнаруживаемые вероятности расщепления за счет отрыва атомов водорода с исчезающе малой SASA [350]. Тем не менее, существенный вклад, вероятно, исходит от нуклеосомной динамики и конформационной неоднородности [183; 414]. Нуклеосомная ДНК в растворе принимает ряд состояний с твист-дефектом ДНК [194]. В течение HRF экспериментов ансамбль нуклеосом в растворе с различными конформациями будет доступен для гидроксилрадикальной атаки. Дополнительные конформационные вариации также могут быть результатом различий в позиционировании ДНК [415].

Одним из важных аспектов является появление идентифицируемых локальных максимумов на экспериментальном профиле на расстоянии 10–11 п.н. - по одному на виток нуклеосомной ДНК, тогда как в профилях H-SASA 7–8 нуклеотидов ДНК витка почти одинаково доступны. Возможное объяснение может исходить из несовершенства модели H-SASA, служащей приближением сложной физики и химии разрыва ДНК. В частности, как подчеркивают Бегусова и др. контролируемая диффузией природа реакции атаки гидроксильных радикалов и быстрый распад радикалов из-за взаимодействия с другими атомами должны сделать вероятность разрыва зависимой не только от локальной доступности атомов водорода, но и от общей формы молекулы, которая будет управлять потоком радикалов, которые могут достичь заданного участка ДНК из раствора [362; 395]. Более того, необходимо учитывать распределение электронной плотности целевой молекулы, чтобы правильно описать реакцию атаки гидроксильных радикалов. Различные многоступенчатые реакционные

пути могут также вносить вклад в химическую сложность, например, сообщалось о механизмах расщепления ДНК посредством вторичных гистоновых радикалов [416].

С момента пионерских исследований лабораторий Вольфа и Туллиуса в начале 1990-х годов по периодичности нуклеосомной ДНК в растворе на основе экспериментов HRF [385—387; 417] стали доступны рентгеновские структуры нуклеосом с высоким разрешением [165]. Периодичность ДНК в нуклеосомах является важным параметром из-за ее влияния на сверхспирализацию ДНК [418]. Наше исследование, путем одновременного анализа профилей rTw, H-SASA и экспериментальных профилей HRF, позволяет изучить теоретические основы использования экспериментов HRF для определения периодичности ДНК. Результаты нашего анализа в настоящее время требуют осторожности при использовании положений максимумов или минимумов на экспериментальных профилях частоты расщепления ДНК для измерения периодичности нуклеосомной ДНК с высокой точностью (лучше, чем 1 п.н. / оборот). Как мы показали, минимумы профилей на профилях rTw и H-SASA не всегда совпадают, а положения минимумов могут отличаться на 1-2 п.н. из-за вариаций в локальных взаимодействиях гистон-ДНК. С другой стороны, максимумы на экспериментальных профилях частоты расщепления ДНК могут зависеть от локальной динамики ДНК, природы допустимых дефектов кручения, других динамических эффектов, а также некоторых нерегулярностей расщепления, специфичных для последовательности (локальные пики, как показано на рисунке 3.14). Кроме того, второстепенные, медленно мигрирующие продукты расщепления ДНК могут исказить форму профилей. Дальнейшие исследования могут выявить вклад этих эффектов в мельчайшие детали формы профилей HRF.

В совокупности наше исследование обеспечивает новый надежный подход к анализу и интерпретации данных экспериментов по расщеплению ДНК нуклеосом гидроксильными радикалами, определению положения ДНК с точностью до одной пары оснований и построению высокоточных молекулярных моделей неизвестных нуклеосом. Такие модели позволяют обнаружить уникальное пространственное расположение между гистонами и особенностями последовательностей ДНК и могут формировать рациональную основу для структурного понимания взаимодействий между нуклеосомами и другими белками хроматина. Учитывая аналогию между HRF и окислительным повреждением ДНК в

клетках, результаты этого исследования могут быть в дальнейшем использованы для обоснования влияния нуклеосом на повреждение ДНК с высоким разрешением.

### 3.2.5 Благодарности

Мы благодарим Т. Туллиуса за ценные предложения, которые помогли сформулировать идею этого исследования, и Л. Заславского за обсуждения оптимальных алгоритмов подбора данных. Исследования были поддержаны программой очных исследований Национальной медицинской библиотеки и Национального института рака, Национальных институтов здоровья, США; Российским научным фондом [14–24–00031] (договор № 14-24-00031-р, разработка алгоритмов визуализации нуклеосом); Исследовательским кампусом Джанелия Медицинского института Говарда Хьюза; Университетом Джона Хопкинса; программой сотрудничества США-России в области биомедицинских наук.

### 3.3 Выводы главы 3

Создано программное обеспечение для оцифровки результатов гель-электрофореза продуктов расщепления ДНК с помощью гидроксильных радикалов. Данный подход позволяет количественно описывать интенсивность расщепления ДНК по каждому конкретному нуклеотиду. Разработаны подходы для оценки вероятности расщепления ДНК в заданных позициях на основе анализа доступности атомов водорода дезоксирибозы в молекулярных моделях, что позволяет использовать данные футпринтинга для задач интегративного моделирования. На основе данных подходов создан метод, позволяющий по данным футпринтинга ДНК определять с точностью до одного нуклеотида положение последовательности ДНК на нуклеосоме. Метод использован при создании модели центромерной нуклеосомы *S. cerevisiae*.

Данные гидроксильного футпринтинга ДНК могут быть эффективно использованы для построения структурных моделей ДНК-белковых комплексов. Для ДНК-белковых комплексов, обладающих осью симметрии второго порядка, сравнение профилей расщепления ДНК для двух цепей позволяет дополнительно повысить точность определения положения ДНК относительно белка.

## Глава 4. Моделирование комплексов нуклеосом с белками хроматина<sup>1</sup>

В данной главе обсуждается применение подходов интегративного моделирования для построения и анализа различных ДНК белковых комплексов, в том числе с использованием разнородных экспериментальных данных низкого информационного содержания (футпринтинг ДНК, электронная микроскопия, флюоресцентная микроскопия с использованием эффекта Ферстеровского резонансного переноса энергии). Главу предваряет введение с описанием современных представлений о структуре хроматина и некоторых подходах к его моделированию. Материалы главы основаны на следующих статьях [A2; A5; A6; A19; A23; A54] и материалах конференций[A43].

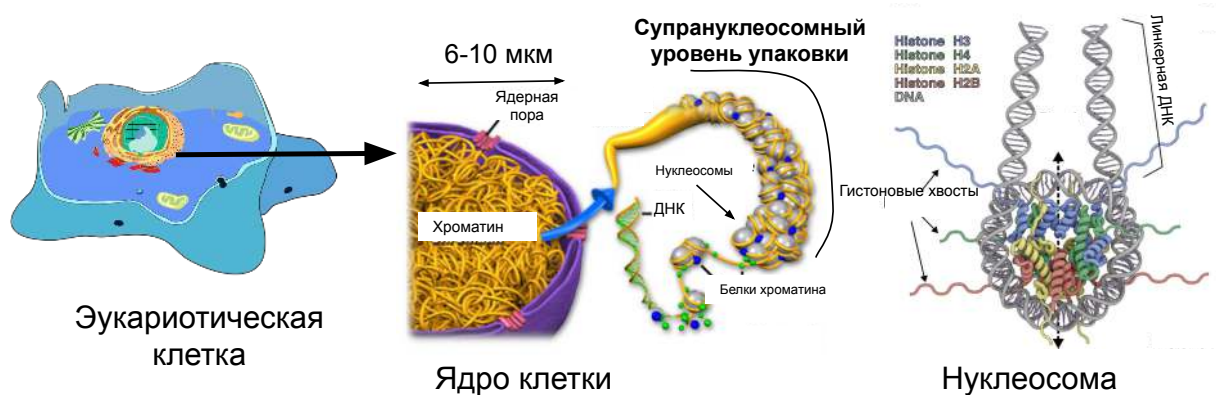
### 4.1 Введение

Ключевой задачей биологии 21 века является понимание функционирования хроматина в эукариотических клетках на молекулярном уровне (см. Рисунок 4.1.). Именно в хроматине происходят ключевые процессы интерпретации генетической программы клеток, определяющие их развитие и реакцию на изменение внешних условий. Комплексное решение этой задачи, понимание взаимодействия различных генетических и эпигенетических элементов в геноме (энхансеров, промотеров, инсуляторов, локусов, несущих эпигенетические модификации и т.д.) откроет дорогу к пониманию и управлению сложными регуляторными генетическими сетями, на основе которых функционируют клетки человека, животных, растений. Это в свою очередь позволит усовершенствовать методы диагностики и лечения многих заболеваний, осуществлять рациональный подход к генетическому редактированию и проектированию организмов важных с биотехнологической точки зрения.

---

<sup>1</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [A2; A5; A6; A11; A19; A23].

В частности известно, что многие болезни человека связаны с теми или иными нарушениями работы хроматина. Изменения в энхансер-зависимой экспрессии генов могут привести к развитию многих видов рака, а также сердечных и аутоиммунных заболеваний у людей [419]. Нарушение структуры хроматина и его гетерогенность, наблюдаемая при онкологических заболеваниях, позволяет раковым клеткам вырабатывать резистентность к химиотерапии, а воздействия на структуру хроматина активно рассматриваются как один из способов терапии онкозаболеваний [420]. Также например, петлеобразование в хроматине рассматривается как лекарственная мишень для лечения бета-талласемий и серповидноклеточной анемии [421].



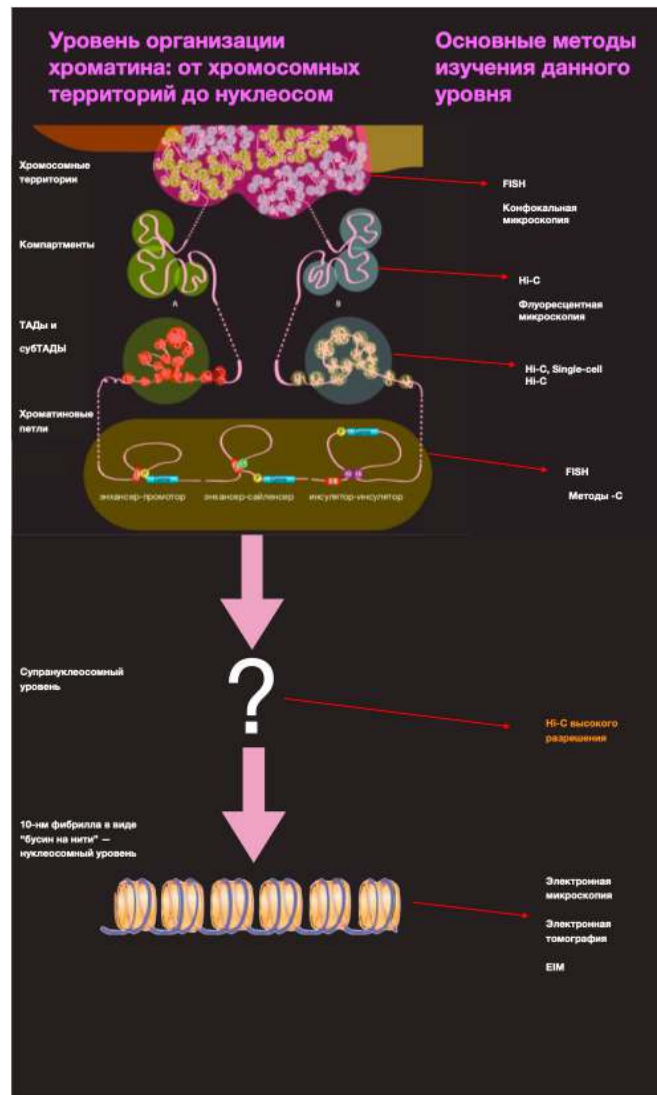
**Рисунок 4.1** — От клеток к нуклеосомам. Хроматин находится в ядре эукариотических клеток и включает в себя ДНК, белки и РНК. Элементарная единица упаковки ДНК - нуклеосома, комплекс из 8 гистонов (H3, H4, H2A, H2B) и около 200 пар оснований ДНК. Супрануклеосомальный уровень - уровень укладки нуклеом (от единиц до тысяч нуклеосом, от 200 до 200 000 пар оснований ДНК).



**Рисунок 4.2** — Хромосомные территории. Деконденсированные хромосомы визуализируются (флуоресцентной микроскопией) как хромосомные территории, которые таким образом являются высшим уровнем организации генома.

Адаптировано из [146].

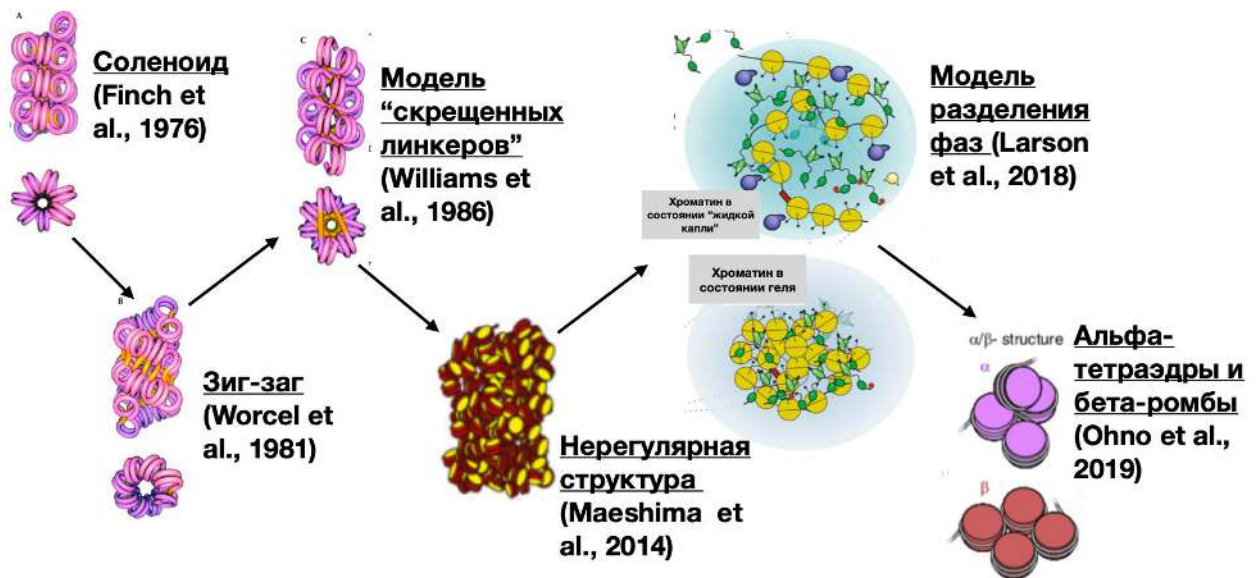




**Рисунок 4.3** — Структура хроматина и проблема ее понимания на супрануклеосомном уровне. Представлены следующие за хромосомными территориями уровни организации хроматина и основные методы их обнаружения. Множественные эксперименты с использованием Hi-C говорят о существовании иерархически организованных ТАДов (Топологически ассоциированных доменов) - участков генома, локусы внутри которых чаще физически контактируют друг с другом, чем с локусами извне. Однако о физической природе ТАДов ведутся споры. Их существование пытаются связать с хорошо изученной способностью хроматина к образованию петель и наличием петель, с помощью которых регуляторные элементы генома (сайленсеры, энхансеры) взаимодействуют с регулируемым ими промотрами генов. Также организацию хроматина на этом уровне пытаются объяснить с позиции коллоидной химии, представляя ключевым для регуляции функциональности хроматина и формирования гетерохроматина переходы между растворимым состоянием и гелеобразным состоянием (через фазу “жидкой капли”) - модель “разделения фаз” [422].

У человека длина всей ДНК, содержащейся в одной клетке, равна примерно 2 метрам (0,33 нм длина одного основания, всего около 3.1 млрд. пар оснований). При этом диаметр ядра равен примерно 6 микрометрам. То есть в ядре ДНК сложена в 3 000 000 раз. При этом она остается функциональной - гены способны к экспрессии. Однако, эта функциональность избирательна: в разных типах клеток для транскрипции доступны разные гены, их активность определяет уникальность клеточного типа. Экспрессия контролируется разными способами, но одним из наиболее важных является специфическая компактизация ДНК в хроматине. Уже с 1907 года известно [423], что хроматин бывает активным (эухроматин) и неактивным (гетерохроматин), причем последний включает в себя некоторые участки ДНК во всех клетках (конститутивный гетерохроматин), а некоторые - в зависимости от типа и состояния клетки (факультативный гетерохроматин). Эти типы, называемые также активным и неактивным компартментами хроматина, являются уровнем компактизации, следующим за наивысшим - хромосомным (в интерфазе представленном в виде хромосомных территорий, что было открыто с помощью конфокальной микроскопии и FISH - метода визуализации индивидуальных хромосом на основе флуоресцентной гибридизации и так же хорошо обнаруживающимся с помощью обычной оптической микроскопии. В 1974 году был открыт [164] низший уровень организации хроматина - нуклеосомный: на нем ДНК наматывается на гистоновые белки, образуя структуру в виде “бусин на нити” — 10-нанометровую фибриллу. Этот уровень компактизации ДНК изучен хорошо, в 1997 году методами рентгеновской кристаллографии была определена структура нуклеосомы в атомарном разрешении [165], сейчас накоплено значительное количество данных о биофизике и биохимии нуклеосом и об их цепочках, которые иногда называют 10-нанометровыми фибриллами [424]. При этом консенсуса об остальных уровнях организации хроматина в научном сообществе нет [127; 146]. Точное представление об уровне, следующем за уровнем компартментов, отсутствует. Как 10-нанометровая фибрилла компактизуется, чтобы образовать хромосому, неизвестно. Долгое время считалось, что она образует 30-нанометровую фибриллу, которая имеет вид соленоида (одностартовая спираль) или зигзага (двустартовая спираль) в зависимости от длины линкерных участков ДНК [425]. Были получены некоторые экспериментальные данные, указывающие на существование данной структуры [426; 427]. Однако на сегодняшний день многие исследователи считают ее артефакт-

ной, не существующей *in vivo*. Методы электронной томографии подтверждают существование 10-нанометровой фибриллы, но не 30-нанометровой [118; 119]. Таким образом, организация хроматина на супрануклеосомном уровне остается неясной, поэтому ее изучение является одним из главных вызовов современной молекулярной биологии.



**Рисунок 4.4** — О проблеме понимания организации и функционирования хроматина на супрануклеосомном уровне. История представлений об организации супрануклеосомного уровня в хроматине. Использованы иллюстрации из: [428]; [127]; [422]; [130].

Изучение трехмерной структуры генома имеет очень важное значение для исследования скоординированной работы генов и ее нарушений - например, при онкологических и наследственных заболеваниях, потому что в геноме эукариот существуют регуляторные элементы - последовательности, стимулирующие - тогда они называются энхансерами - или ослабляющие - их название сайленсеры - экспрессию определенных генов, посредством физического контакта с их промоторами [419; 429]. Важно, что между регуляторным элементом и регулируемым им геном может быть большое (более мегабазы) геномное расстояние (расстояние в первичной, линейной структуре ДНК). Физический контакт происходит путем образования петли из ДНК [430]. Для определения наличия физического контакта между последовательностями ДНК в трехмерной структуре генома были разработаны методы семейства -C:3C,4C,5C,Hi-C [146]. Эти методы основаны на химическом “сшивании” контактирующих в хроматине участков,

фрагментации ДНК с образованием “сшитых” пар, лигирование фрагментов в этих парах с образованием гибридных молекул ДНК, детекция гибридных молекул (обнаружение молекулы, состоящей из последовательности А и последовательности В, говорит о физическом контакте этих последовательностей в ДНК). Метод 3С позволял оценить взаимодействие только двух участков генома друг с другом, поэтому его называли “один против одного”. Аналогично: 4С — “один со многими”, 5С — “многие со многими”. Наиболее совершенным является метод Hi-C, модифицированный стадией биотинилирования гибридных молекул для обогащения ими раствора путем стрептавидинового осаждения и использованием высокоэффективного секвенирования спаренных концов [148]. Hi-C - это “все против всех”. Этот метод позволяет составить матрицу частот взаимодействий между всеми участками генома заданной длины. Метод Hi-C позволил обнаружить следующий за компартментами уровень компактизации хроматина - Топологически ассоциированные домены (ТАДы) - регионы ДНК, в пределах которых взаимодействие участков друг с другом происходит чаще, чем с участками из других регионов[149].

Хотя ТАДы проявляются как определенный паттерн на матрице взаимодействий, их физическая природа неясна [147]. Существуют две основные концепции относительно нее. Первая - DNA loop extrusion (“выпетливания ДНК”) предполагает возникновение на участке ДНК, ограниченном сайтами связывания белка CTCF, петли, расширяющейся со временем до этих сайтов [431]. В разных клетках петля в пределах сайтов связывания CTCF находится на разных стадиях расширения - соответственно в разных клетках рядом оказываются разные участки ДНК, что на матрице Hi-C отображается как ТАД. Слабость этой гипотезы в том, что она считает ТАДы популяционным феноменом, в то время как интрахромосомные ТАДы (в отличие от интерхромосомных) и хроматиновые глобулы, считающиеся соответствующими ТАДам, обнаруживаются и в отдельных клетках [155]. Другая гипотеза утверждает, что ТАДы - это особые хроматиновые домены, возникающие в клетках из-за физического взаимодействия между нуклеосомами в неактивном хроматине [432]. Эта гипотеза тоже имеет недостаток - она не объясняет расположения сайтов связывания CTCF на границах ТАДов. Таким образом очевидно, что необходимо продолжать исследования структуры хроматина и на уровне ТАДов. Так как ТАДы отчасти “фрактальны” - содержат внутри себя вложенные друг в друга ТАДы

меньших размеров - изучение хроматина даже на супрануклеосомном уровне может помочь понять структуру всей иерархии ТАДов [433; 434].

Для исследований регуляторной функции хроматина необходимо понимать его супрануклеосомальную организацию. Так, обеспечение коммуникации между энхансером и промотером, которые могут располагаться на расстояниях до нескольких сотен килобаз, обеспечивается за счет образования петель. Энхансеры - элементы, которые обеспечивают транскрипционную активацию подавляющего большинства генов человека и в значительной степени определяют состояние хроматина вблизи промоторных областей. Эпигенетическими метками активных энхансеров являются, в частности, H3K27ac, H3K4me1, H3K27me3, и гистоновые варианты H3.3 и H2A.Z. К меткам промоторов относятся: H3K4me3 и H3K27ac [419].

Исследования энхансеров затрудняются их различиями в последовательности ДНК. Большинство промоторов ассоциировано с 1 энхансером, в то время как 25% промоторов - с двумя или больше в разные моменты времени. Известно более миллиона энхансеров, а одновременно активны порядка тысячи энхансеров.

Изменения в энхансер-зависимой экспрессии генов могут привести к развитию многих видов рака, а также сердечных и аутоиммунных заболеваний у людей. К изменению активности энхансеров в онкологических заболеваниях приводит изменение копийности энхансеров, структурные перестройки генома, изменяющие таргетный ген энхансера и точечные мутации или индели, которые влияют на связывание с транскрипционными факторами и образуют новые энхансеры или наоборот, разрушают инсуляторы [435].

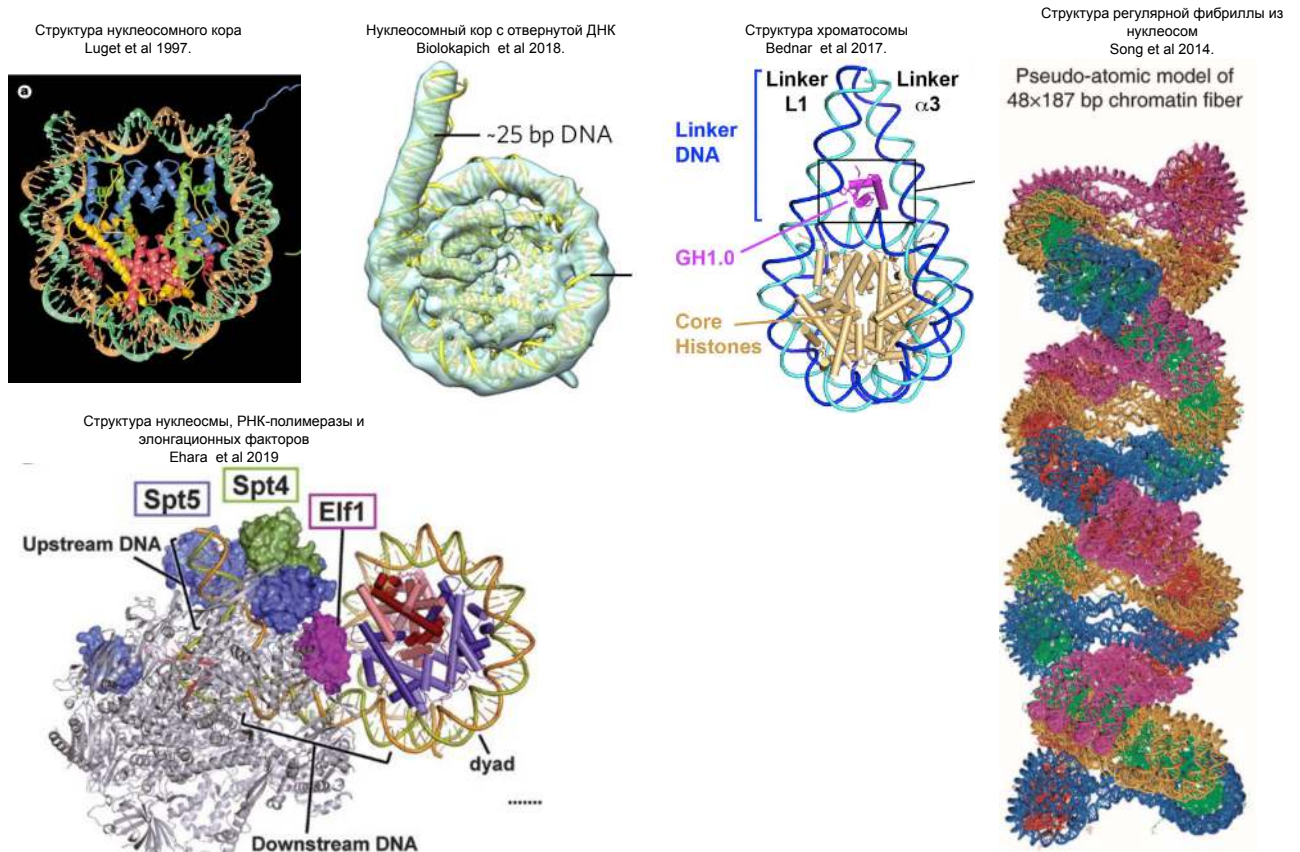
Актуальной проблемой в области лечения онкологии, в части в химиотерапии, является развитие резистентности к лекарственным агентам. Так, в работе [420] было показано, что использование агентов, влияющих на внутриядерные изменения плотности упаковки хроматина, приводит к уменьшению транскрипционной гетерогенности и, соответственно, к предотвращению возникновения химиорезистентности.

#### 4.1.1 О прогрессе методов структурной биологии в изучении супрануклеосомной организации

Кристаллическая структура коровой частицы нуклеосомы (гистоны H3, H4, H2A, H2B и 145-147 пар оснований ДНК) с атомистическим разрешением была получена в 1997 году [165] (Рисунок 4.5). Позднее стало очевидно, что данная структура является лишь некоторым компактным вариантом, который удовлетворяет условиям кристаллической упаковки, а сама нуклеосома достаточно динамична. Несмотря на биохимические данные о динамичности нуклеосом, получения различных вариантов ее структуры в динамике стало возможным лишь недавно благодаря успехам криоэлектронной микроскопии [223]. Также благодаря криоэлектронной микроскопии в последние годы получены разнообразные структуры нуклеосом с белками хроматина, в частности структура нуклеосомы при прохождении РНК-полимеразы [436], структуры нуклеосом с различными ремоделлерами (белковыми комплексами перемещающими нуклеосомы или заменяющими гистоны) [321; 324; 437].

Важной вехой явилось определение структуры хроматосомы - комплекса нуклеосомы с линкерным гистоном H1 [202].

Применение криоЭМ к определению структуры нуклеосомальных фибрилл также проводилось. Были получены структуры нуклеосомных фибрилл с и без гистона H1 [438]. Однако, данные фибриллы, состоящие из 12 нуклеосом, находящихся на равном расстоянии друг от друга, были приготовлены искусственно. Поэтому значение данной структуры для понимания биологических процессов ставится под сомнение.



**Рисунок 4.5** — Различные структуры нуклеосом, их комплексов и фибрилл из нуклеосом, полученные в последнее время методами структурной биологии.

## 4.2 Комплексы нуклеосомы и РНК-полимеразы

*Текст данного раздела основан на статьях [A23; A54].*

### 4.2.1 Введение

Транскрипция РНК-полимеразы II (Pol II) индуцирует обширное ремоделирование хроматина, сопровождающееся ковалентными модификациями гистонов и их обменом (см. напр. [439–441]). В то же время, гистоны вытесняются только из высокотранскрибируемых генов [442–444]; таким образом, Pol II обычно встречается с нуклеосомами во время транскрипции каждого участка ДНК длиной  $\sim 200$  пар оснований. Остающиеся нуклеосомы на транс-

крибируемых генах образуют два типа барьеров для транскрипции Pol II [440; 445]. Во-первых, каждая нуклеосома представляет собой барьер, на котором Pol II приостанавливается после транскрипции 40–50 п.н. от границы нуклеосомы проксимальной к промотор. Во-вторых, намного выше барьер образован первой (+1) транскрибированной нуклеосомой, по крайней мере в дрозофилы [445]; в этом случае Pol II останавливается на нуклеосомной границе. Остановившаяся Pol II обнаружена на многих генах: от 15% до более чем 50% генов у мух и человека соответственно [446–448]. В во многих случаях остановившаяся Pol II была картирована вместе с хорошо расположенной +1 нуклеосомой. Предполагается, что эта +1 нуклеосома может через паузирование влиять на общий уровень экспрессии генов [445; 449].

Данные *in vitro* экспериментов установили, что даже одна нуклеосома может образовывать высокий барьер первого типа для транскрипции Pol II [180; 287; 450–453] и что факторы транскрипции, шапероны гистонов, модификации гистонов и т.д. необходимы Pol II для преодоления этого барьера [180; 291; 452; 454–461]. Сильный барьер образуется после транскрипции 40–50 п.н. за границей нуклеосомы [область +(40–50)], является нуклеосомно-специфичным, специфичным для Pol II и был описан для всех проанализированных организмов, от дрожжей до человека [180; 451; 452]. Таким образом +40–50 нуклеосомный барьер является “универсальным” признаком транскрипции через хроматин с помощью Pol II и может использоваться для регуляции экспрессии на уровне элонгации транскрипта [180]. На любой последовательности ДНК, барьер формируется в различных положениях внутри региона +(40–50) [180].

#### 4.2.2 Создание моделей комплекса в положении активного центра +49 (нулевая петля)

Наши предыдущие исследования выявили элонгационный комплекс +49 (ЭК+49, 49 п.н. от промотор-проксимальной границы нуклеосомы), содержащий нулевую петлю, в качестве ключевого интермедиата, который помогает нуклеосомам выживать во время транскрипции [453]. В данной работе мы смоделировали комплекс ЭК+49 путем докирования структуры высокого разрешения элонгационного комплекса дрожжевой полимеразы II на нуклеосому



с 601 хорошо позиционирующей последовательностью ДНК [403; 462] (Рисунок 4.6).

ЭК+49 имеет следующие свойства (Рисунок 4.6): (i) Основная часть молекулы Pol II обращена в раствор и не имеет стерических конфликтов с гистонами нуклеосомного кора, (ii) Изгиб ДНК на  $90^\circ$ , присутствующий в ЭК, обращен к октамеру и позволяет формировать нулевую петлю, (iii) контакты ДНК-гистоны на участке ДНК длиной  $\sim 20$  п.н. за ЭК стабилизируют нулевую петлю, (iv) вытеснение 50 п.н. с конца нуклеосомы дистального к промотору снижает размер области ДНК, взаимодействующей с гистонами перед ферментом с  $\sim 100$  до  $\leq 50$  п.н. Это, вероятно, способствует дальнейшему разворачиванию ДНК с октамера впереди от Pol II и транскрипции через нуклеосому [453]. (v) Отрицательно заряженная область на поверхности Pol II (предполагаемая последовательность, стабилизирующая нулевую петлю, область ZLS - zero-loop-stabilizing sequence) находится в непосредственной близости к положительно заряженной области октамера гистонов и, таким образом, может стабилизировать нулевую-петлю и способствовать выживанию нуклеосом во время транскрипции.

### 4.2.3 Выявление отрицательно заряженных регионов на предполагаемых гистон-взаимодействующих поверхностях Pol II

Используя модель ЭК+49 (рис. 4.6), были идентифицированы предполагаемая взаимодействующая с Pol II поверхность на октамере гистонов и три взаимодействующие с гистонами поверхности Pol II (рис. 4.7A). Все взаимодействующие с гистонами поверхности заряжены отрицательно и локализуется на самой крупной субъединице дрожжевой Pol II (yRpb1, регионы 1–3, рис. 4.7A). Для дальнейшей оценки возможной роли предполагаемых областей ZLS на Pol II во время транскрипции через хроматин было оценено их присутствие и сохранение в различных структурно связанных эукариотических и прокариотических РНК-полимеразах (Таблица 6). РНКП *Escherichia coli*, которая использует механизм транскрипции типа Pol II через хроматин [463] не была включена в это исследование, потому что структуры с высоким разрешением этого фермента пока не получена. Последовательности областей 2 и 3 ZLS является

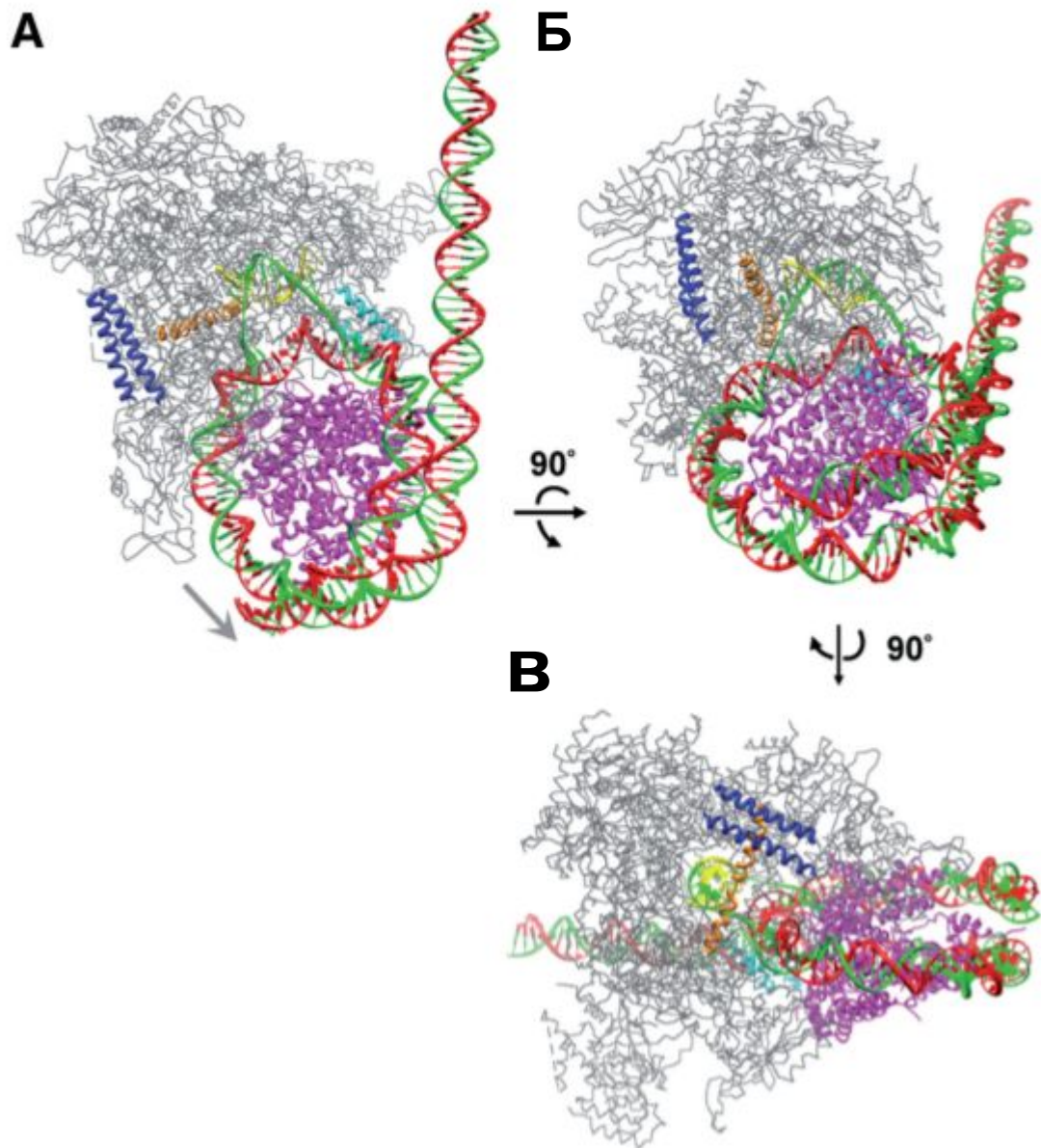
консервативными с более, чем 50% идентичностью, у мультисубъединичных прокариотических РНКП *T. thermophilus* и *T. aquaticus* (структурно сходных с Pol II [464; 465]), а как область 1 консервативна только в эукариотических РНКП. Гомологичных последовательностей в односубъединичный РНКП бактериофага T7 не было обнаружено.

Предполагается, чтобы чистый отрицательный заряд последовательности ZLS может быть важен для стабилизации нулевой петли и эффективного выживания нуклеосомы в ходе транскрипции [453]. В соответствии с этим предположением области ZLS отрицательно заряжены в дрожжевой Pol II (Рисунок 4.7А), которая не смещает нуклеосомы во время транскрипция [182; 463]. Области ZLS Pol II человека также отрицательно заряжены, и поэтому ожидается, что Pol II человека также сохраняет нуклеосомную структуры во время транскрипции. Это предложение согласуется с наблюдаемыми сходными паттернами сильной нуклеосомной паузы, которые характерны для Pol II человека и дрожжей [180].

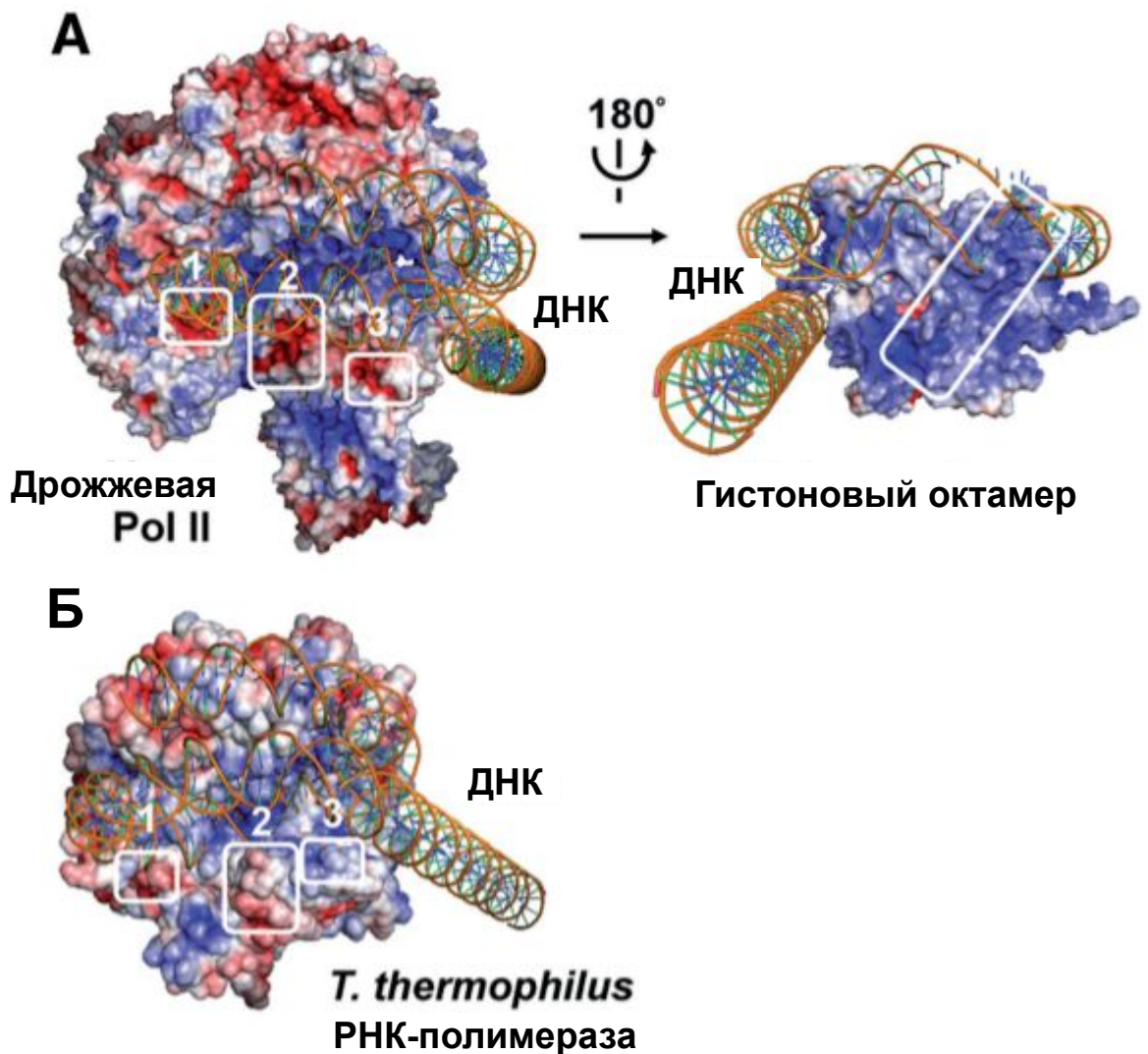
Напротив, отрицательные заряды областей ZLS не сохранились в других проанализированных прокариотических РНКП (из *T. thermophilus* и *T. aquaticus*) (Рисунок 4.7Б и Таблица 6).

Хотя предполагаемая взаимодействующая с гистонами поверхность РНКП *T. thermophilus* имеет некоторые регионы с локальным отрицательным зарядом (рис. 4.7Б), в данном регионе нет суммарного положительного заряда. Данные показывают, что чистый отрицательный заряд внутри взаимодействующего с гистонами поверхности этих РНКП положительно коррелирует с вероятностью выживания нуклеосом при транскрипции. Кроме того, расстояние между областями ZLS и поверхность октамера гистонов в среднем на  $\sim 5\text{\AA}$  больше для РНКП *T. thermophilus*, чем у дрожжевой Pol II. Более низкий полный отрицательный заряд областей ZLS и больше расстояние между РНКП и октамером гистонов, вероятно, уменьшает силу взаимодействия между РНКП и гистонами. Соответственно, РНКП, которые имеют более низкий полный отрицательный заряд области ZLS, как ожидается, формируют менее стабильный комплекс нулевой петли в положении ЭК +49. Поскольку стабильная нулевая петля, вероятно, необходима для выживания нуклеосом, ожидается, что эффективность выживаемости будет ниже во время транскрипции РНКП *T. thermophilus* и *T. aquaticus* по сравнению с Pol II. Поскольку стабильная нулевая петля также приводит нуклеосомной паузе в области +45 [466], уменьшение

паузирования в положении +45 также ожидается во время транскрипции этими РНКП. Чтобы проверить эти предсказания модели, эксперименты по транскрипции на идентичных ДНК и нуклеосомных матрицах были проведены с помощью РНКП *T. thermophilus* и *T. aquaticus*.



**Рисунок 4.6** — Модель ЭК+49 (нулевая петля) с дрожжевой Pol II. (А) Структуры нуклеосомы и комплекса элонгации Pol II дрожжей с активным сайтом в позиции +49 [PDB 3LZ0 и 1Y1W соответственно [403; 462]] были объединены с использованием подхода докинга. Чтобы позволить формирование небольшой внутринуклеосомной петли ДНК, содержащей транскрибирующую Pol II, промотор-дистальная область нуклеосомной ДНК длиной 50 п.н. была отмотана от октамера [453]. Октамер нуклеосомы изображен пурпурным. Матричная цепь ДНК, нематричная цепь ДНК и цепи РНК изображены зеленым, красным и желтым, соответственно. Мостовая спираль (bridge), зажим (clamp), С-терминальная двойная спираль (coiled coil) и остальная часть молекулы Pol II показаны оранжевым, голубым, синим и серым, соответственно. Серая стрелка указывает направление транскрипции. (Б) Структура была повернута на  $\sim 90^\circ$  вокруг горизонтальной оси. (В) Затем структура была повернута на  $90^\circ$  вокруг вертикальной оси.



**Рисунок 4.7** — Отрицательно заряженная поверхность Pol II может стабилизировать интермедиат нулевой петли посредством взаимодействия с гистоновым октамером. (А) Предполагаемые взаимодействующие с гистонами молекулярные поверхности дрожжевой Pol II. Слева: Три идентифицированные поверхности на большой субъединице Rpb1 дрожжевой Pol II внутри ЭК+49 (смоделирована как на Рисунке 4.6), взаимодействующие с октамером гистонов (не показан), показаны белыми квадратами (области 1–3). Справа: контактирующая поверхность на октамере гистонов (повернут на 180° вокруг вертикальной оси). Самый темный синий и самый темный красный цвет обозначает электростатический потенциал 5 кТ/е и -5 кТл/е, соответственно. ДНК показана желтым. (Б) Молекулярные поверхности РНКП *T. thermophilus* (PDB ID 2O5I) в гипотетической модели комплекса ЭК+49 окрашены по величине электростатического потенциала (октамер гистонов не показан). Области 1–3 гомологичны контактирующим с гистоном поверхностям Pol II (Figure 4.7A).

Фермент	Виды	Заряд ZLS области <sup>a</sup>	Нулевая петля при +49	Высота +45 барьера	Выживание нуклеосом
Pol II (RPB1)	<i>H. sapiens</i>	-8	ND	+++++ <sup>b</sup>	ND
Pol II (Rpb1)	<i>S. cerevisiae</i>	-15	+ <sup>c</sup>	+++++ <sup>b</sup>	+++++ <sup>d</sup>
РНКП ( $\beta'$ )	<i>T. thermophilus</i>	+1	ND	+ <sup>e</sup>	ND
РНКП ( $\beta'$ )	<i>T. aquaticus</i>	+1	- <sup>e</sup>	+ <sup>e</sup>	-

<sup>a</sup> Полные заряд гомологичной области 1434–1450 ZLS субъединицы Rpb1 дрожжевой Pol II.

<sup>b</sup> Ссылка [180].

<sup>c</sup> Ссылка [453].

<sup>d</sup> Ссылка [463].

<sup>t</sup> Эта работа.

**Таблица 6** — Корреляция между присутствием отрицательно заряженных областей на обращенной к октамеру поверхности РНКП, высотой нуклеосомного барьера и выживаемостью нуклеосом во время транскрипции

#### 4.2.4 Более высокий отрицательный заряд области ZLS коррелирует с более сильным барьером +45 и более эффективной выживаемостью нуклеосом при транскрипции

Данные о чистом заряде региона ZLS, силе барьера +45 нуклеосомного паузирования и судьбы нуклеосом во время транскрипции различными РНКП просуммирована в Таблице 6. Наличие более высокого отрицательного заряда область ZLS положительно коррелирует с формированием ключевого промежуточного продукта, который сильно способствует выживанию нуклеосом (нулевая петля в положении +49), с более сильным паузированием в области +45 и более эффективным выживанием нуклеосом во время транскрипции.

В целом данные свидетельствуют о том, что присутствие и полный заряд области ZLS частично определяет структуру ключевого интермедиата, сформированного в позиции +49, а в свою очередь, судьбу нуклеосом при транскрипции. Когда ZLS область присутствует и отрицательно заряжена (например, у Pol II человека или дрожжей), нуклеосомное паузирование в положении +45 является сильным, нулевая петля образуется в положении +49 и нуклеосомы эффективно выживают во время транскрипции. Когда регион ZLS присутствует, но не содержит отрицательного заряда (например, в РНКП *T. thermophilus* и *T. aquaticus*), нуклеосомная пауза в +45 положении слабая, нулевая петля не образуется, и нуклеосомы теряются во время транскрипции.

### Дискуссия и сравнение с экспериментами

В нашей работе роль отрицательно заряженной области на поверхности Pol II (области ZLS) во время транскрипции через хроматин было оценено с использованием компьютерного моделирования и экспериментального анализа механизма транскрипция через хроматин различными РНКП. Вычислительное моделирование показало, что в комплексе элонгации, определяющем выживаемость нуклеосом во время транскрипции (ЭК+49), три области ZLS находятся в непосредственной близости к положительно заряженной области октамера гистонов и таким образом могут стабилизировать нулевую-петлю и облегчать нуклеосомную выживаемость при транскрипции (рисунки 4.6 и 4.7). Анализ транскрипции РНКП *T. thermophilus* и *T. aquaticus*, у которых есть незаряженные области ZLS, предполагает, что в этом случае нуклеосомная пауза +45 слабая, нулевая-петля не образуется, и нуклеосомы теряются во время транскрипции. Напротив, транскрипция Pol II, имеющая отрицательно заряженные области ZLS, приводит к сильной нуклеосомной паузе в +45 положении, образовании нулевой-петли и выживаемости нуклеосом [180; 182; 453]. Сила +45 паузы частично определяется отрицательным зарядом в области ZLS Pol II. Таким образом свойства (в частности, заряд) области ZLS диктуют все ключевые параметры механизма транскрипции через нуклеосому.

Как может наличие и заряд ZLS региона влиять на эти множественные аспекты транскрипции через хроматин? Ранее мы предположили, чтобы регион

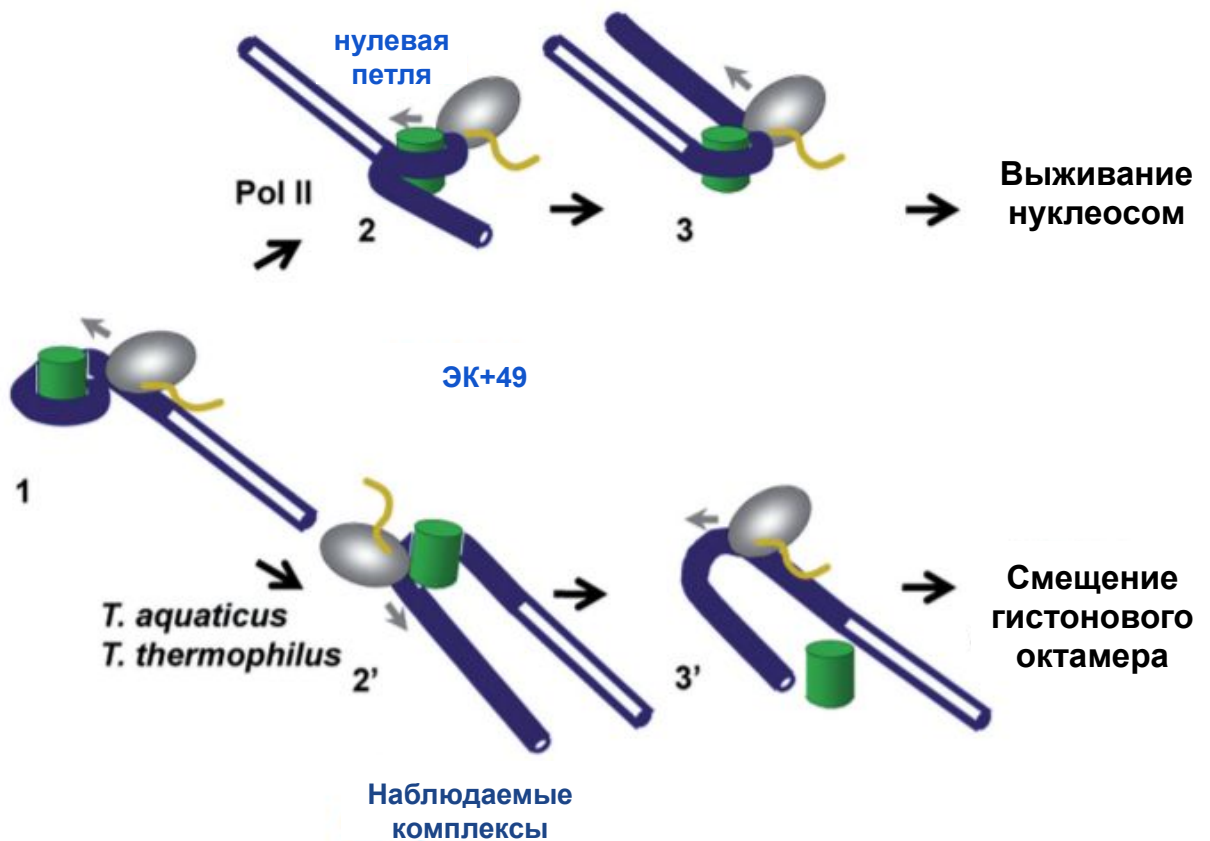
ZLS может влиять на структуру/стабильность ключевого интермедиата, содержащего нулевую-петлю (ЭК+49), образующегося во время транскрипции через нуклеосому [453]. Это промежуточное состояние, в свою очередь, как ожидалось будет играть связывающую роль между паузированием в +45 положении и эффективным выживанием нуклеосом во время транскрипции [466]. В соответствии с этим предположением наши текущие данные свидетельствуют в пользу того, что РНКП у которых отсутствует заряд в области ZLS или сама область не сталкиваются с сильным нуклеосомным барьером, не образуют нулевые-петли и не могут поддерживать эффективное выживание нуклеосом.

Основываясь на наших данных, мы предлагаем следующую модель, объясняющую наблюдаемую связь между +45 паузой, образование нулевой петли и судьбой нуклеосом. По мере того как разные РНКП входят в нуклеосому и подходят к области +45 (комплекс 1), они образуют нулевую-петлю с разной эффективностью. В частности, дрожжевой Pol II образует нулевую петлю (комплекс 2) с высокой эффективностью [453], отчасти потому, что ее ZLS-область заряжена отрицательно и стабилизирует нулевую петлю за счет взаимодействия с положительно заряженной поверхностью октамера гистонов. Формирование нулевой петли вызывает медленное раскручивание ДНК перед ферментом (комплекс 3) - то есть сопровождается сильной нуклеосомной паузой. ДНК позади РНКП повторно оборачивается вокруг октамера гистонов и нуклеосома восстанавливается в исходном положении [453]. Напротив, во время транскрипции РНКП *T. thermophilus* и *T. aquaticus* с незаряженными ZLS областями нуклеовая петля (если образуется) нестабильна (комплекс 2'). Таким образом, транскрипция сопровождается лишь незначительной нуклеосомной паузой и смещением гистонового октамера при дальнейшей транскрипции через нуклеосому [467]. Таким образом, структуры ключевых промежуточных продуктов образующиеся во время транскрипции через +45 область нуклеосомной ДНК, вероятно, обуславливают множество аспектов транскрипции через хроматин, включая нуклеосомную паузу и судьбу гистонов при транскрипции. Присутствие и заряд региона ZLS вряд ли будет единственным фактором, влияющим на результат транскрипции через хроматин. Таким образом, расстояние между областями ZLS и октамером гистонов в ЭК+49, вероятно, важно для стабильности нулевой петли.

Отрицательно заряженная область 2 ZLS дрожжевого Pol II локализована в свитч-2 домене и, вероятно, важна для разделения нитей ДНК при транскрип-



ции [462]. Она также локализована в кислотном домене, который влияет на активацию транскрипции, активность Pol II дрожжей и важен для нормального роста клеток [468]. Наши исследования предполагают дополнительную важную функцию области 2 ZLS - определение скорости прохождения РНКП через нуклеосому и судьбу нуклеосом при транскрипции в хроматине. Эта функция особенно важна для эукариотических Pol II: в данном случае консервативный отрицательно заряженный участок ZLS, скорее всего, способствует выживанию нуклеосом и поддержанию гистоновых меток во время транскрипции.



**Рисунок 4.8** — Предлагаемые механизмы транскрипции через нуклеосому разными РНКП. Мы предлагаем, чтобы общий отрицательный заряд на поверхность РНКП, которая обращена к октамеру гистонов, частично обуславливает структуру важного промежуточного комплекса (ЭК+49). Более подробное описание см. в тексте.

## Методы

### Моделирование Pol II в позиции +49 в нуклеосоме и анализ структурных особенностей моделируемых комплексов

Модель была построена с использованием программа UCSF Chimera (<http://www.cgl.ucsf.edu/chimera/>) [28] за счет объединения структуры коровой частицы нуклеосомы на основе 601 последовательности [PDB ID 3LZ0 [403]] со структурой элонгационного комплекса дрожжей Pol II [PDB ID 1Y1W [462]]. Веб-сервер 3D-DART использовался для восстановления конформации ДНК [469]. Остальные подходы следовали описанию в опубликованной статье [453]. ЭК+49 *T. thermophilus* был построен путем наложения кристаллической структуры РНКП [PDB ID: 2O5I [464]] на модель Pol II ЭК+ 49. Молекулярные электростатические поверхности белков рассчитывали с помощью APBS [470] и отображали с помощью PyMOL (<http://www.pymol.org>). Расстояния между отрицательно заряженной поверхностью Pol II и октамером гистонов, а также вовлеченные аминокислотные остатки также были идентифицированы в PyMOL. Все белковые последовательности РНКП в этой статье были взяты из базы данные RefSeq NCBI (NP\_000928.1, NP\_010141.1, NP\_418415.1, YP\_145078.1 и CAB65466.3) (<http://www.ncbi.nlm.nih.gov/guide/>). Последовательности различных РНКП были выровнены с помощью NCBI Blast (<http://blast.ncbi.nlm.nih.gov>).

#### 4.2.5 Создание моделей комплекса полимеразы и нуклеосомы в положении активного центра +42

### Экспериментальные данные

Эксперименты проводились с использованием мононуклеосомных матриц, которые повторяют многие важные аспекты транскрипции PolIII в хроматине

*in vivo* [182; 457; 466]. Механизм образования нуклеосомного барьера при транскрипции изучен с использованием 601 и 603 мононуклеосом, которые содержат ДНК последовательность с полярным барьером для транскрипции (PBS) и образуют очень сильный нуклеосомный барьер в одна из двух транскрипционных ориентаций (непермиссивная ориентация) [180; 453; 471]. Матрицы ДНК были разработаны так, чтобы Pol II останавливалась в положениях -41 и -5 (на 41 или 5 п.н. выше промоторно-проксимальной границы нуклеосомы) при наличие различных частичных комбинаций нуклеотид трифосфатов. Pol II спонтанно останавливается во время транскрипции непермиссивной нуклеосомы 601 при 150 мМ KCl на +46 и +48 позиции. Остановка обратима, так как Pol II может транскрибировать всю матрицу после дестабилизации гистонового октамера в 1М KCl.

Состояние Pol II после встречи с высоким нуклеосомный барьером изучалось с помощью 601-ой ДНК матрицы, которая представляет собой самый сильный нуклеосомный барьер для транскрипции Pol II [180]. Pol II был спонтанно останавливалась в позиции +46 и +48 на 601-нуклеосомной матрице.

Для определения положения активного центра Pol II комплексы +46 и +48 инкубировали в присутствии факторов элонгации транскрипции IIS (TFIIS). TFIIS - это фактор транскрипции, который сильно облегчает расщепление РНК в активном центре Pol II после ее отката (backtracking) [472]. Инкубация комплексов элонгации с TFIIS приводили к образованию гомогенной РНК длиной 42 нуклеотида, что указывает на то, что оба комплекса откатываются на 4 или 6 п.н. до положения +42. Нуклеосомные комплексы +42, +46 и +48 остаются полностью функциональными и могут реактивироваться после удаления нуклеосом в присутствии 1 М KCl. Таким образом, когда Pol II сталкивается с сильным нуклеосомным барьером, она останавливается и возвращается назад на 4–6 п.н. Возврат может привести к образованию щели в комплексе между Pol II и нуклеосомой, если нуклеосомная ДНК не закручивается обратно на поверхность октамера.

Вращательная ориентация Pol II в позиции +42 после отката несовместима с формированием нулевой петли [453]. Поэтому мы ожидали, что нуклеосомная ДНК будет отмотана от октамера перед комплексом +42.

Доступность нуклеосомной ДНК в комплексе +42 была анализировали с помощью анализа чувствительности к рестриктазам. Нуклеосома сильно защищает ДНК от переваривания рестрикционными ферментами [473]. В ЭК+42

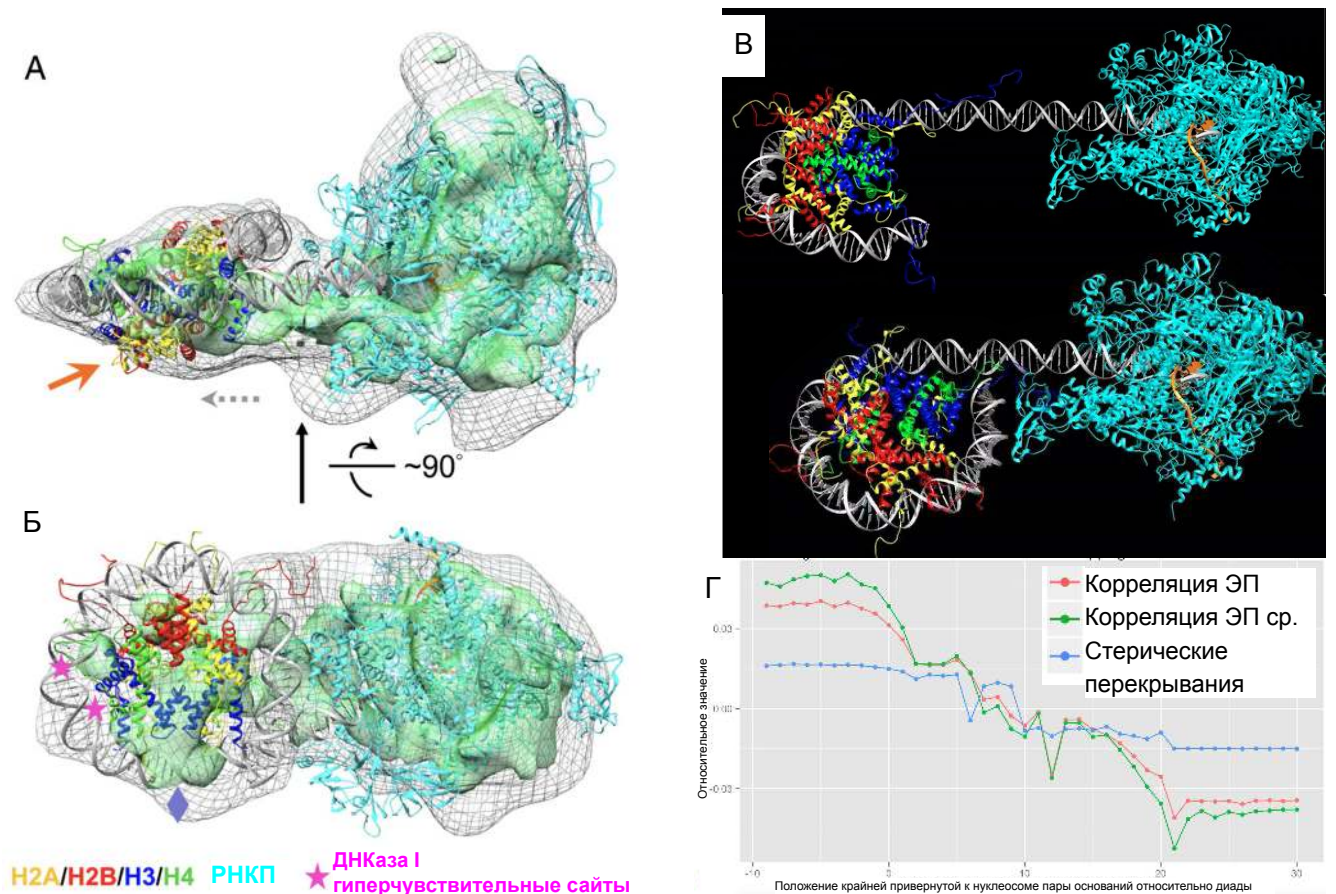
только сайт AfIII расположен выше активного центра (но не Сайты RsaI или MfeI, расположенные ниже по течению) чувствителен к расщеплению. Таким образом, ДНК в комплексе +42 отворачивается от октамер выше этой позиции, но тесно связана с гистонами ниже активного центра фермента.

### Моделирование структуры комплекса Pol II с нуклеосомами и ее анализ

Структура комплекса +42 РНКП с нуклеосомой была определена с помощью электронной микроскопии в режиме одиночного анализ частиц. ЭК+42 имеет суммарную молекулярную массу около 545 кДа, что затрудняет их наблюдение в криоЭМ. Поэтому контраст изображения был увеличен с помощью негативного контрастирования уранилацетатом. Всего 8 500 частиц комплекса ЭК+42 были проанализированы. Финальная структура содержит более крупную (диаметром около 20 нм) и меньшую (около 14 нм в диаметре) плотность. Детальные структурные особенности комплекса не могут быть разрешены при полученном разрешении в 22Å.

Поэтому нами были применены алгоритмы интегративного моделирования для восстановления структуры комплекса. Для моделирования ЭК+42, рентгеновская структура элонгационного комплекса РНКП *Thermus thermophilus* [464] была совмещена с рентгеновской структурой нуклеосомы [238] за счет удлинения нижележащего дуплекса ДНК с сегментом прямой В-ДНК и соединения его с различными соответствующими положениями вдоль нуклеосомной ДНК, соблюдая правильную геометрию ДНК и химических связей. Далее РНКП *T. thermophilus* была заменена структурой РНКП *E. coli* [474] посредством суперпозиции структур поскольку структуры двух ферментов очень похожи. Моделирование проводили с помощью UCSF Chimera [28] и программы Nucleic Acid Builder (NAB) [475]. Полученный ансамбль моделей автоматически вписывался в экспериментально полученную карту электронной плотности с помощью утилиты Chimera EM (атомистическая модель была преобразована в расчетную карту электронной плотности с разрешением 30Å и проводился обширный поиск с 40 различными пробными стартовыми пространственными положениями). Лучшая модель была выбрана на основе

максимальной кросскорреляции электронных плотностей. В полученной модели активный центр РНКП находится в позиции +42; нижестоящий дуплекс ДНК входит в нуклеосому в положении +71 (за 3 п.н. до положения оси диады). Полученная модель представлена на рисунке 4.9. Наконец, структура дрожжевого элонгационного комплекса Pol II (PDB ID 1Y1W) [462] была наложена на структуру РНКП *E. coli* для получения Pol II +42 элонгационного комплекса.



**Рисунок 4.9** — Структура комплекса элонгации остановленного в положении +42 нуклеосомы. (А) Модель ЭК +42, построенная путем подбора оптимальной структуры комплекса, вписывающейся в электронную плотность, полученную экспериментально. Изоповерхность трехмерной реконструкции электронной плотности показана с более высоким или низким порогом серой сеткой или зеленым цветом, соответственно. Серая пунктирная и оранжевая стрелки указывают направление транскрипции и промотор-проксимальный димер Н2А / Н2В, экспонированный в раствор, соответственно. (Б) Структура была повернута на  $90^\circ$  вокруг горизонтальной оси. Положение участков ДНК, гиперчувствительных к ДНКазе I, показано звездочками. Нуклеосомная диада обозначена фиолетовым ромбом. (В) Различные варианты модели, отличающиеся отворотом ДНК. (Г) График корреляции плотности модели и экспериментальной электронной плотности.

Спейсер ДНК в полученной модели, соединяющий нуклеосому и РНКП, в значительной степени скрыт белками; таким образом структура согласуется с данными футпринтинга ДНКазы I, не обнаруживающими доступности спейсерной ДНК. Плотная структура комплекса согласуется с предположением, что у остановленного комплекса откат полимеразы сопровождается прикручиванием

нуклеосомной ДНК к поверхности октамера гистонов. Докинг структуры элонгационного комплекса дрожжей Pol II в полученную модель свидетельствует о том, что подобный комплекс может быть образован эукариотический Pol II.

Самая яркая и удивительная особенность комплекса +42 - это экспонирование поверхности промотор-проксимального димера гистонов H2A/H2B, взаимодействующего с ДНК. Действительно, модель +42 комплекса (рис. 4.9) предполагает, что проксимальный участок H2A / H2B димера не стабилизируется взаимодействием с ДНК или РНКП, но остается связанным с октамером. Кроме того, наши предыдущие данные показали, что переход от комплекса +42 к +49 сопровождается полным приворотом нуклеосомной ДНК обратно на поверхность димера [453], предполагая, что димер не покидает ДНК до того, как Pol II продвинется более чем на 49 п.н. в нуклеосому. Стабильность димера во время транскрипции примечательна, учитывая, что октамер, свободный от ДНК, полностью теряет оба димера H2A / H2B менее, чем за 1 с [476]. Таким образом наличие нуклеосомной ДНК, которая остается частично накрученной вокруг тетрамера H3/H4 и промотор-дистального димера H2A/H2B приводит к снижению скорости диссоциации промотор-проксимального димера. Данные свидетельствуют о том, что связывание промотор - проксимального димера H2A/H2B с тетрамером H3/H4 в +42 комплексе может быть аллостерически стабилизировано оставшимися ДНК-гистоновыми взаимодействиями в нуклеосоме. Поскольку смещение димера представляет собой первый шаг к вытеснению всего октамера гистонов [477], этот механизм, вероятно, важен для выживания нуклеосом во время различных процессов, включая ремоделирование нуклеосом.

## Обсуждение

Наши данные идентифицируют переход из положения +48 в положение +49 как ключевой этап транскрипции через нуклеосому, где Pol II делает выбор между остановкой и дальнейшей транскрипцией. В минимальной системе, содержащей только Pol II и нуклеосому, этот выбор продиктован прежде всего последовательностью нуклеосомной ДНК, определяющей сродство ДНК-гистоновых взаимодействий и вероятность возврата Pol II. Особенно,

Последовательность ДНК в области +(99–102) имеет решающее значение для выхода из положения +48.

После преодоления нуклеосомного барьера в дальнейшем Pol II транскрипция обычно сопровождается выживанием нуклеосом благодаря образованию небольшой внутринуклеосомной петли ДНК на поверхности октамера гистонов. Высокая эффективность выживания нуклеосом оставалось загадкой, потому что транскрипция различными РНК-полимеразами сопровождается разворачиванием протяженного участка ДНК (до 80 п.н.) от октамера [A54; 453; 478], который, как ожидается, вызовет немедленную потерю димера H2A/H2B [476]. Наблюдаемая стабильность +42 комплекса, содержащего димер H2A / H2B, который экспонируется в растворитель свидетельствует о том, что связывание димера H2A/H2B с тетрамером H3/H4 может быть стабилизировано аллостерически и дает объяснение замечательной стабильности структуры нуклеосом во время различных процессов, включая транскрипцию [453], репликацию [479] и АТФ-зависимое ремоделирование хроматина [243]. Хотя один димер конститутивно вытесняется из нуклеосом при транскрипции *in vitro* [182], это вытеснение, вероятно, происходит после транскрипции через позицию +49, поскольку ЭК+49 содержит оба димера H2A/H2B [480].

Описанная здесь паузирование в положениях +42/48 наблюдается у *Drosophila* [445] и дрожжей [480]. В дрожжах первичный нуклеосомный барьер во время удлинения транскрипта Pol II встречается, когда активный центр фермента находится в области +(40–55) п.н. в нуклеосоме. С аналогичным барьером сталкивается Pol II у дрозофилы во время транскрипции через нуклеосомы, локализованные на расстоянии более ~400 п.н. от сайтов старта транскрипции [445]. Однако существуют дополнительные нуклеосомные барьеры в позициях -7 и +20 п.н. у *Drosophila*, характерные для нуклеосом, локализованные сразу после сайта старта транскрипции [445].

Недавние исследования показывают, что плотность комплексов Pol II вдоль транскрибированных генов, вероятно, определяет важные параметры транскрипции через хроматин - высоты нуклеосомных барьеров [A27; 291; 457], степень смещения гистонов и их обмен [444; 457; 481; 482]. В то же время множественные факторы, включая модификации гистонов, варианты гистонов, шапероны гистонов и ремоделеры хроматина, дополнительно изменяют динамику гистонов на транскрибируемых генах [439–441; 483]. Некоторые из эти

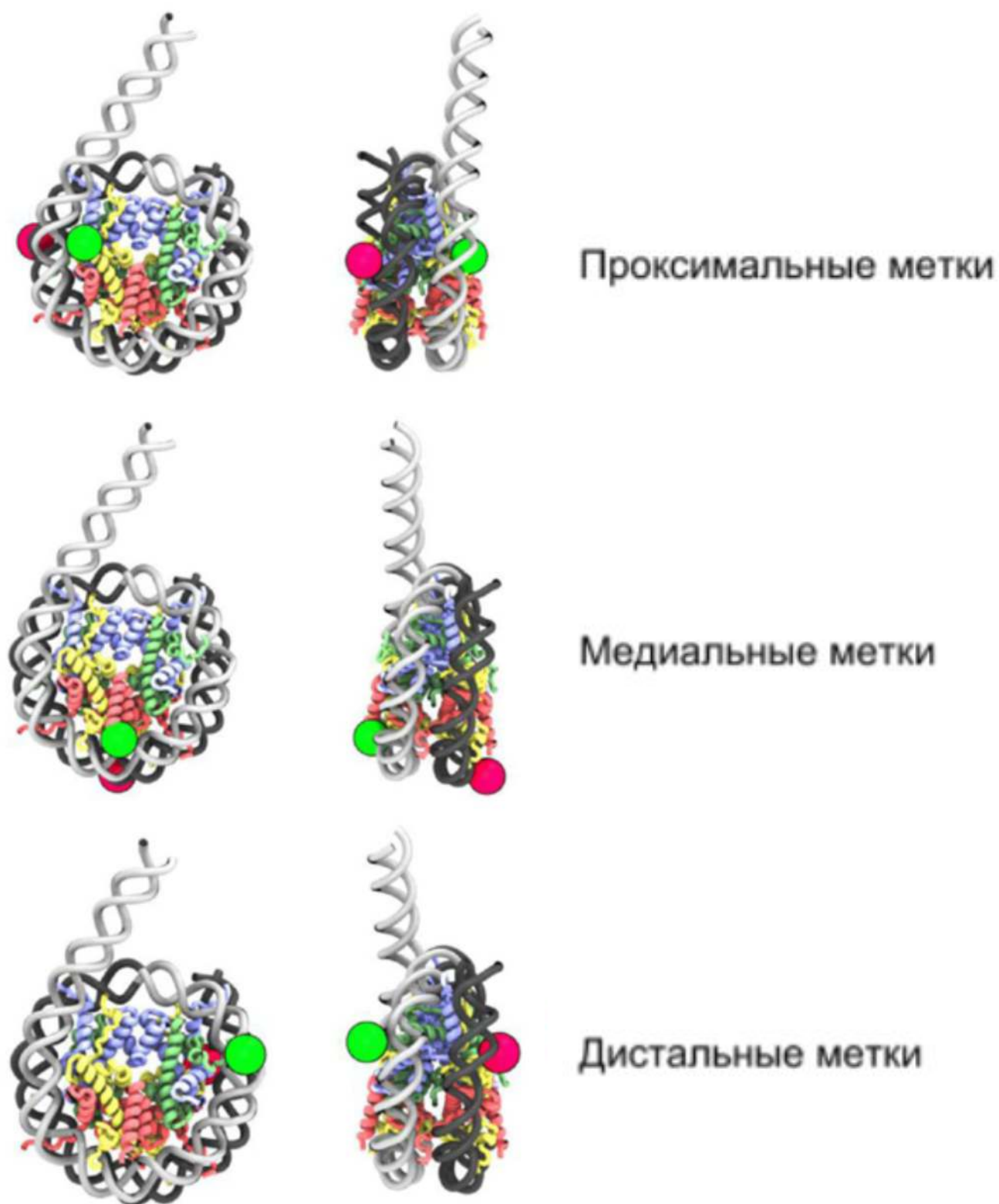


факторов взаимодействуют с опухолевыми супрессорами [484] и представляют важные мишени для разработки противоопухолевых препаратов [485].

### 4.3 Конформационные перестройки нуклеосом при взаимодействии с комплексом FACT

*Текст данного раздела основан на статье [A19].*

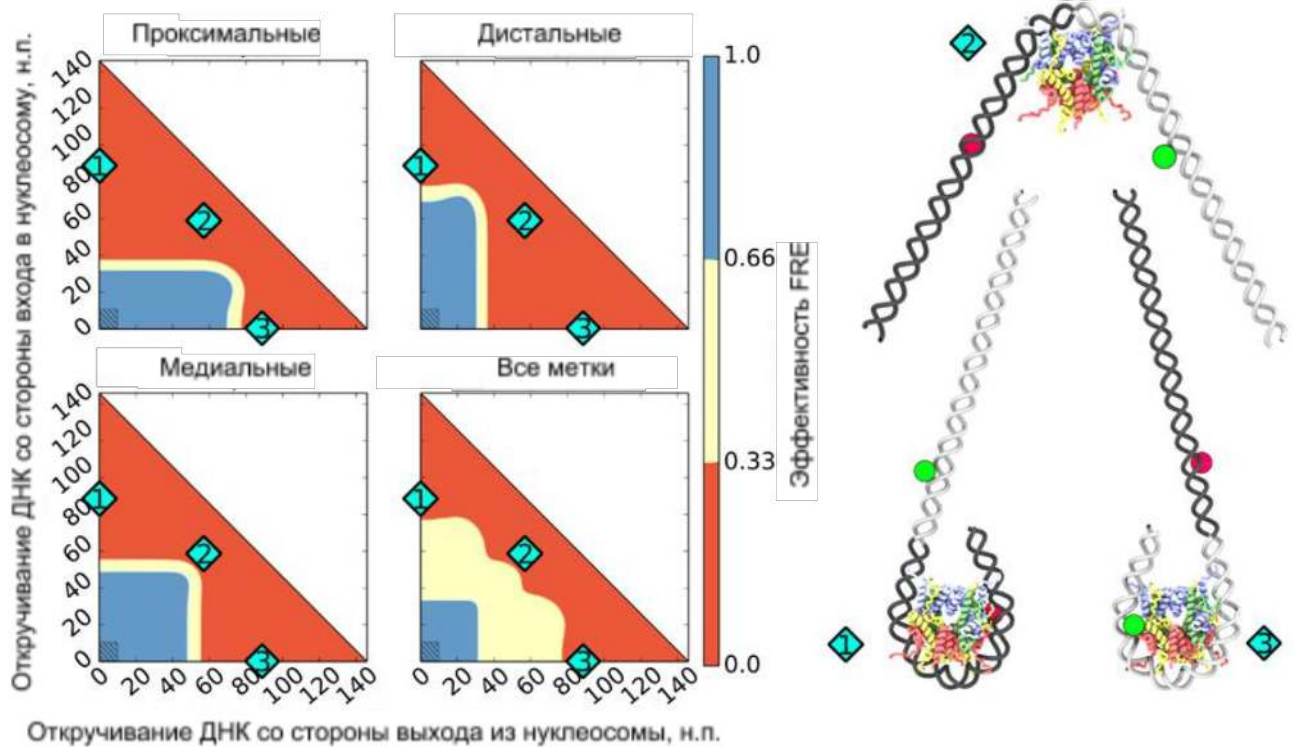
Динамика нуклеосом может активно регулироваться взаимодействиями с различными белками. Нашими коллегами был произведен ряд экспериментов методом spFRET для изучения конформационных перестроек в нуклеосомах при его взаимодействии с комплексом шаперонов FACT. В ходе эксперимента был введен ряд пар меток (Cu3-Cu5) в разные положения нуклеосомной ДНК: проксимальные метки - вблизи входа в нуклеосому, медиальные - по центру нуклеосомы и дистальные - вблизи выхода из нуклеосомы (Рис. 4.10). По изменению сигнала spFRET было обнаружено, что связывание FACT вызывает существенные, но обратимые изменения в сигнале FRET. С использованием этих данных методами интегративного моделирования нами были построены модели возможных конформационных перестроек в нуклеосомах.



**Рисунок 4.10** — Расположение флуоресцентных меток на нуклеосоме в spFRET экспериментах. Зеленым показаны положения донора (Cy3), красным - акцептора (Cy5).

На основе измерений, в том числе с использованием измерения фактора детекции (см. главу 1.2), нами был сконструирован и отобран ряд возможных

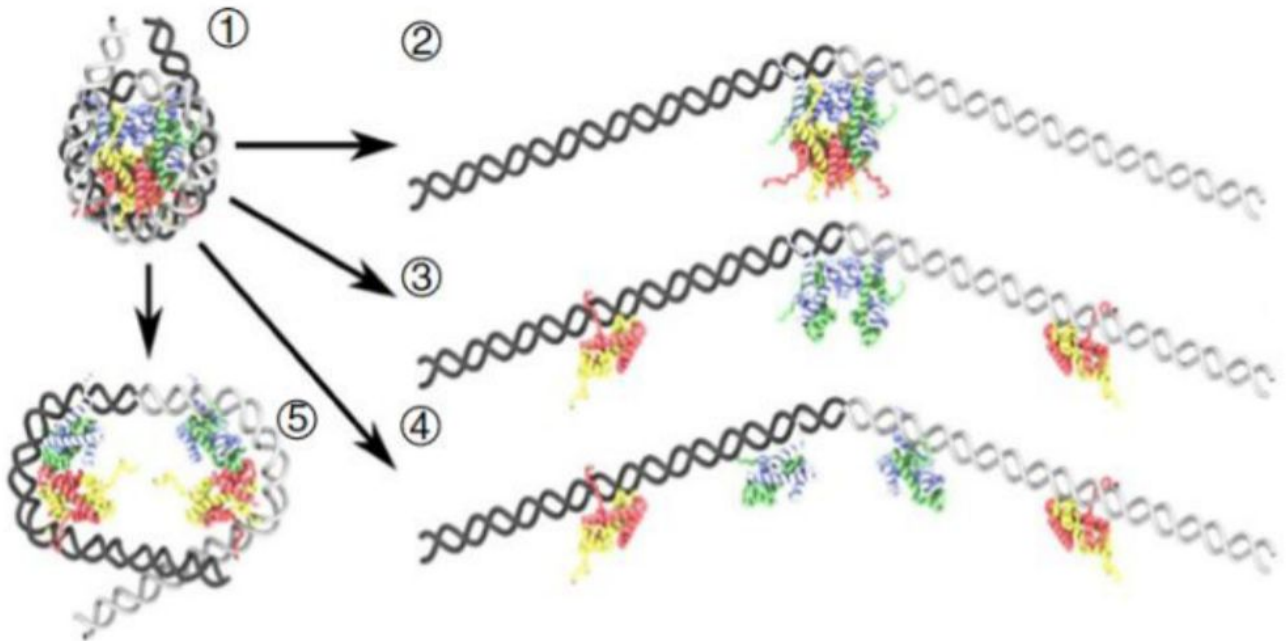
моделей, которые согласуются с экспериментальными данными. Установлено, что для значительного изменения эффективности переноса, наблюдаемого в эксперименте, величина откручивания ДНК от нуклеосомы должна составлять около 100 н.п. ДНК, при этом около 40 н.п. должно откручиваться с одной стороны и около 60 н.п. с другой стороны (рис. 4.11).



**Рисунок 4.11** — Анализ зависимости эффективности FRET от степени разворачивания нуклеосом. Слева: тепловые карты ожидаемых эффективностей переноса энергии на трех парах флуорофоров. Красные и голубые области соответствуют степеням откручивания ДНК с низкой и высокой эффективностью FRET соответственно. Затемненные участки тепловых карт (левый нижний угол) отвечают нативным состояниям нуклеосом (состояниям близким к наблюдаемым в кристаллических структурах). Справа: Модели конформаций нуклеосом, соответствующие отмеченным номерам на картах.

По этим данным были построены молекулярные модели, соответствующие таким степеням отворота ДНК (Рис. 4.12): модель откручивания ДНК с сохранением гистонового октамера (Рис. 4.12), модель откручивания ДНК с образованием структуры “бабочка” (Рис. 4.123), модель с образованием двух хемисом (Рис. 4.125), а также модель полного разрушения октамера гистонов

(Рис. 4.12). Все эти модели в равной степени соответствуют наблюдаемым изменениям в эффективности переноса энергии в присутствии гистонового шаперона ФАСТ.



**Рисунок 4.12** — Модели разворачивания нуклеосом, вызванного взаимодействием с шапероном ФАСТ. Интактные нуклеосомы (1) могут разворачиваться от интактного октамера (2), ДНК может отворачиваться с разрушением (3) или без (4) разрушения интерфейса взаимодействия гистонов Н3–Н3 или путем открытия интерфейса Н3–Н3 без откручивания ДНК (5) с формированием двух хемисом.

Потеря гистонов в процессе откручивания ДНК при низких концентрациях нуклеосом должна быть необратимой, так как в ходе экспериментов в раствор дополнительно добавляли конкурирующую за гистоны ДНК, не меченную флуорофорами. Тот факт, что наблюдаемый эффект был обратим, свидетельствует о том, что разворачивание октамера должно происходить без потери гистонов Н2А–Н2В.

#### 4.4 Модели взаимодействия нуклеосом с белком СENP-С

*Текст данного раздела основан на статье [A6].*

#### 4.4.1 Введение

Кинетохоры представляют собой большие субклеточные структуры из ~75 полипептидов, которые собираются на центромерах хромосом, чтобы сделать возможным точное разделение реплицированных дочерних хромосом во время деления клеток [486; 487]. Открытие трех центромер-специфичных аутоантигенов человека, обозначенных CENP-A (гистоновый вариант CenH3), CENP-B (специфичный для последовательности фактор связывания ДНК спираль-виток-спираль) и CENP-C (компонент внутреннего кинетохора), проложили путь к пониманию центромерного хроматина на молекулярном уровне и его роли в сборке кинетохор [488]. На фундаментальном уровне нуклеосомы замена канонического гистона H3 на CENP-A в специфичных для центромеры нуклеосомах создает платформу для рекрутирования CENP-C [489; 490]. CENP-C колокализуется с CENP-A на центромерах многоклеточных животных и служит структурной связью между нуклеосомами CENP-A и белками внешней кинетохоры, тем самым соединяя центромеры с микротрубочками для сегрегации хромосом [487; 491—494]. CENP-C является ключевым компонентом мультисубъединичного комплекса CCAN (конститутивной сети, ассоциированной с центромерой), образующего внутреннюю часть кинетохоры [495]. Следовательно, молекулярные основы взаимного узнавания между CENP-C и нуклеосомой CENP-A является центральным фактором для понимания того, как внутренняя кинетохора взаимодействует с центромерным хроматином.

Центральная часть полипептида CENP-C содержит три консервативных домена, важных для узнавания центромеры: высококонсервативный “сигнатурный мотив CENP-C”, который контактирует с гидрофобным С-концом CENP-A [489; 496], центральная ДНК-связывающая область, которая содержит CENP-C-подобный мотив, и домен гомодимеризации [497—502]. У почкующихся дрожжей, гомолог CENP-C Mif2 [503] не обладает специфической для позвоночных ДНК-связывающей областью, но он содержит один мотив CENP-C и один “АТ-крюк” [497; 504; 505]. Ранние исследования показали, что АТ-крюк в Mif2 способствует узнаванию центромер и сегрегации хромосом *in vivo* [497; 506—508]. Более того, N-концевой домен в Mif2 ассоциируется с двумя кинето-

хорными белками (Ame1-Okp1), чтобы облегчить сборку внешних кинетохор [509].

#### 4.4.2 Построение модель взаимодействия на основе данных футпринтинга ДНК

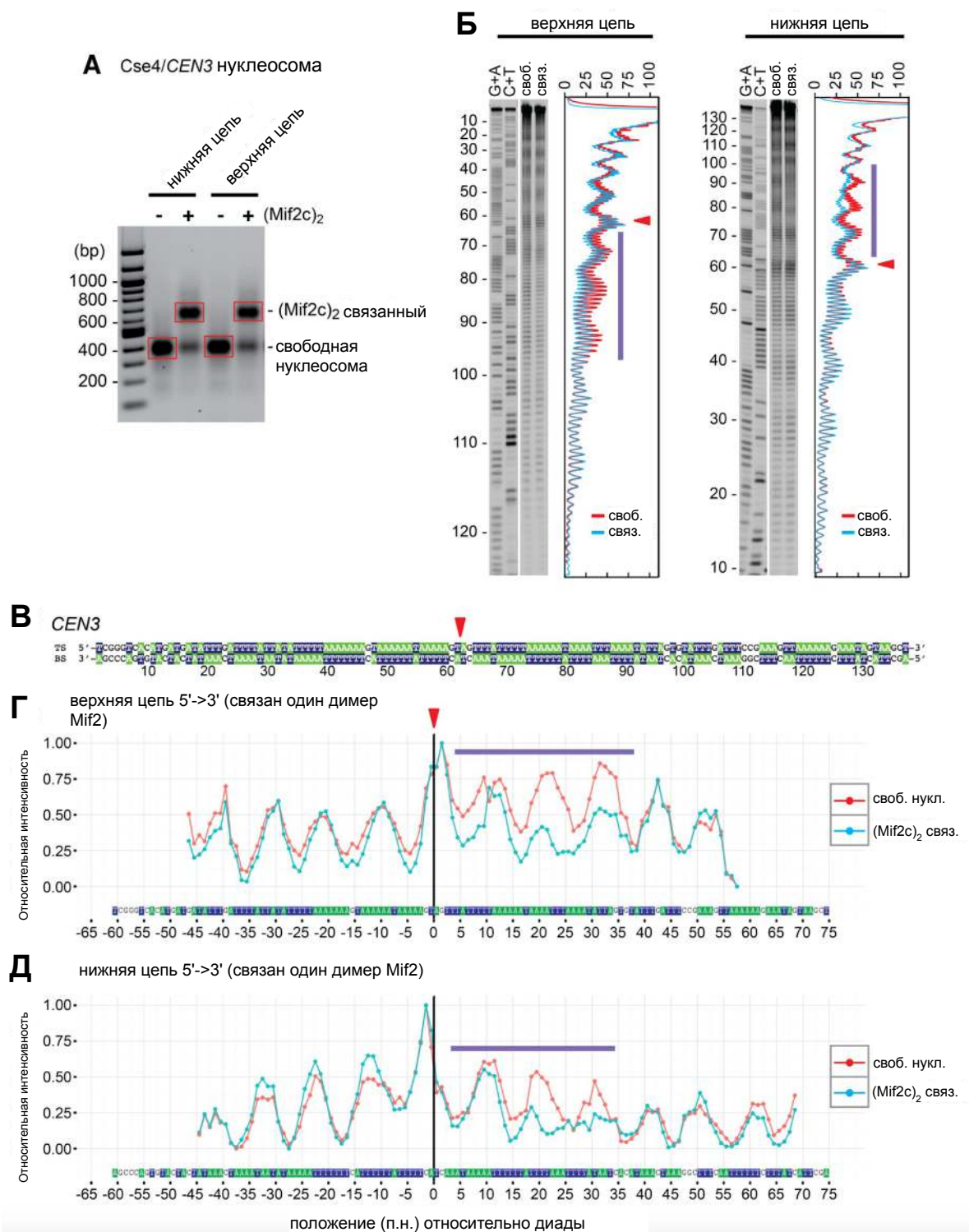
Для понимания взаимодействия Mif2 с центромерными нуклеосомами *in vitro* была проведена серия экспериментов по футпринтингу комплексов гидроксильными радикалами. Для обработки и интерпретации данных была применена программа HYDROID описанная в главе 3, а также построенная нами и описанная в той же главе модель центромерной нуклеосомы.

ДНК центромеры дрожжей длиной 125 п.н. состоит из трех смежных генетических элементов: CDEI, CDEII и CDEIII [510]. Мы использовали метод футпринтинга ДНКазой I для определения сайта связывания димера Mif2с на нуклеосоме Cse4/CEN3 – участка длиной ~30 п.н. в пределах АТ-богатого элемента CDEII (~85 п.н.) на одной стороне от диады нуклеосомы в направлении CDEIII. Для дальнейшего картирования точных контактов Mif2-ДНК мы выполнили футпринтинг димера Mif2с на нуклеосоме Cse4/CEN3 методом гидроксильного футпринтинга. Поскольку гидроксильный радикал представляет собой небольшую молекулу, от его атаки защищены только места тесных контактов. Как показано на рисунке 4.13, димер Mif2с защищает ~30 п.н. ДНК от расщепления гидроксильным радикалом, что в высокой степени согласуется с защитой от ДНКазы I. Учитывая, что минимальный сайт для связывания Mif2 лишённого димеризационного домена составляет 16-18 п.н., защита в 30 п.н. при связывании димера Mif2с предполагает, что оба ДНК-связывающих домена из димерного Mif2 взаимодействует с одной стороной нуклеосомы.

Это поднимает вопрос, почему димер Mif2 должен связываться только с одной стороной нуклеосомы, когда обе стороны состоят примерно из эквивалентного процента АТ. Мы исследовали эту проблему и обнаружили второй, более медленный по подвижности в геле комплекс при двукратном или трехкратном увеличении концентрации димера Mif2с в реакции связывания. Анализ гидроксильным футпринтингом показывает, что обе стороны нуклеосомной диады этого комплекса теперь защищены от расщепления гидроксильными радика-

лами (данные не показаны). Таким образом, одна нуклеосома Cse4 / CEN3 способна связываться с двумя димерами Mif2, занимающими каждую сторону нуклеосомы. Связывание с АТ-богатой ДНК является доминирующим фактором, но правила предпочтительного связывания со специфической АТ-богатой последовательностью остаются неясными. Мы предполагаем, что точное расположение пары оснований АТ может придавать тонкие различия в гибкости и / или конформации ДНК, делая одну АТ-богатую сторону от диады более привлекательной, чем другую.

Чтобы определить, является ли сайт между диадой нуклеосомы и CDEIII предпочтительным для взаимодействий Mif2 на других центромерных нуклеосомах, мы выполнили эксперименты по футпринтингу нуклеосом, реконструированных на CEN10. Интересно, что сайт связывания димера Mif2с был картирован в АТ-богатых регионах на противоположной стороне от нуклеосомной диады для CEN10, по направлению к CDEI. Таким образом, Mif2 взаимодействует с нуклеосомой Cse4 в соседнем с диадой сайте внутри CDEII, но его ориентация связывания относительно оси диады, по-видимому, специфична для каждой конкретной центромеры. Это означает, что определенные паттерны последовательности лежащей в основе АТ-богатой ДНК предпочтительнее для Mif2. Чтобы проверить, отражают ли области повышенной защиты, наблюдаемые в нуклеосомах при связывании димера Mif2с, какие-либо свойства ДНК самой по себе, мы выполнили реакции футпринтинга ДНКазы I с голой ДНК CEN3. Мы не наблюдали никаких специфических следов связывания для димера Mif2с на голой ДНК CEN3. Это указывает на то, что сайт-специфическая локализация Mif2 на нуклеосоме требует вовлечения как ДНК, так и гистоновых сигналов Cse4.



**Рисунок 4.13** — Анализ комплексов димеров Mif2 с центромерными нуклеосомами Cse4/CEN3 методами гидроксильного футпринтинга. (А) Анализ комплексов с помощью агарозного геля электрофореза после расщепления ДНК гидроксильными радикалами. (Б) Изображения продуктов реакции после ПА-АГЭ. Красные линии - нуклеосомы, голубые - нуклеосомы с димером Mif2c. (В) Положение дианды на ДНК для Cse4/CEN3 нуклеосомы (см. главу 3). (Г, Д) Профили интенсивности расщепления гидроксильными радикалами цепей ДНК в нуклеосомах с и без Mif2. Фиолетовые линии - дополнительная защита от расщепления.



### 4.4.3 Обсуждение

Понимание работы центромер многоклеточных животных представляют собой некоторый парадокс, потому что их последовательности сильно расходятся даже между родственными видами, несмотря на функциональную законсервированность [511]. Неуловимая природа универсального паттерна последовательности повторяющихся ДНК активных центромер многоклеточных животных привела к широкому признанию того, что идентичность центромер управляется эпигенетическими механизмами; т.е. за счет сборки и свойств унаследованного гистона CENP-A, а не лежащей в основе последовательности ДНК центромеры [494; 512; 513]. Заметным исключением является строго генетически определенная позиция центромер почкующихся дрожжей, где нуклеосомный гистон Cse4/CENP-A деградирует при переходе G1 – S каждого клеточного цикла, что исключает эпигенетическую передачу [487; 514]. Вместо этого центромерный хроматин восстанавливается *de novo* с помощью инструкций от трех специфических элементов последовательности (120 п.н.) центромеры дрожжей и их множественных родственных факторов [515]. Интересно, что ранее существовавший Mif2 на центромере почкующихся дрожжей передается дочерним клеткам; значение этой передачи еще предстоит изучить.

В этой работе мы выделили отличительные особенности АТ-богатой центромеры дрожжей ДНК. Варианты как ДНК, так и гистонов вносят вклад в тысячекратное усиление связывания Mif2 с центромерными нуклеосомами по сравнению с перицентромерным. Это огромное предпочтение, вероятно, лежит в основе сайт-специфичности нуклеации кинетохор, хотя другие механизмы могут вносить свой вклад; напр. взаимодействия с участием N-концевого хвоста Cse4 [516; 517] и взаимодействия между N концом Mif2 и кинетохорными компонентами CENP-U/Ame1-CENP-Q/Okp1 [509]. Дальнейший вклад в селективность нуклеосом CENP-A по сравнению с нуклеосомами H3 может также происходить от компонентов CCAN [495].

Мы подтвердили важность Cse4-гидрофобного С-конца для взаимодействий Mif2, что согласуется с предыдущими исследованиями [489; 496; 518]. Однако, хотя вклад гистоновых фолдов CENP-A ранее был исключен [518], мы обнаружили явную важность петли L1, несущей три дополнительных заряженных и полярных остатка (KDQ), и спирали  $\alpha$ -2 Cse4. Однако, как мы

показали весьма важной является и вклад последовательности ДНК в связывание Mif2. Какие особенности АТ-богатой СДЕП последовательности длиной 85-п.н. способствуют связыванию Mif2? Область защиты ДНК от расщепления гидроксильными радикалами длиной  $\sim 30$  п.н., возникающая в результате связывания димера Mif2, покрывает большую часть одной АТ-богатой стороны нуклеосомы. СДЕП для всех 16 центромер дрожжей содержат несколько коротких участков An·Tn или Tn·An трактов плюс динуклеотидные шаги ApT или TpA между участками. Длина и расположение трактов различаются между центромерами [406]. Тракты An·Tn и Tn·An обладают узкой малой бороздкой, большим пропеллерным разворот оснований (propeller twist), а также бифуркационными водородными связями, которые улучшают стэкинг оснований и жесткость спирали [519; 520]. С другой стороны, шаги динуклеотидов TpA (но не шаги ApT) обнаруживают расширенные малые бороздки на стыках трактов [404]. Одна из этих или обе геометрические особенности АТ-богатой ДНК могут изменять обычные спиральные параметры нуклеосомной ДНК, намотанной на гистоновое ядро (Bishop et al. 2011). Например, кристаллическая структура одного длинного тракта A16·T16 на канонической нуклеосоме выявила локальное искажение спирали нуклеосомной ДНК [412].

Учитывая корреляцию между предпочтением связывания Mif2 и АТ-богатым составом ДНК, мы предполагаем, что конформации АТ-богатого СДЕП распознаются АТ-крюк и кластерами аминокислот аргинина и лизина (RK) Mif2. Mif2 имеет один классический АТ-крюк (GRPRGRPK) на С-конце своего ДНК/гистон связывающего домена. Структурные исследования показали, как отдельные остатки АТ-крюков взаимодействуют с узкими малыми бороздками родственных AAAT и AATT элементов [504; 505; 521]. Аргинины и, в меньшей степени, лизины кластеров RK широко используются для во взаимодействии белок-ДНК не только гистонового ядра нуклеосомы [165], но также и факторов транскрипции, специфичных для последовательности животных, таких как как UBX, OCT1, бактериальный репрессор MogR [284; 522–524]. Мы предполагаем, что кластеры RK, используют этот способ взаимодействия с ДНК как часть своей стратегии по специфическому связыванию. Мы предполагаем, что наши результаты закладывают важные основы для дальнейших исследований и уточнению структурно-динамической модели взаимодействия Mif2 с центромерными нуклеосомами.

## 4.5 Анализ комплексов нуклеосомы и гистона H1

*Текст данного раздела основан на статьях [A5; A11] .*

### 4.5.1 Совместное влияние аминокислотной последовательности гистона H1 и нуклеотидной последовательности ДНК на структуру хроматосомы: анализ методами молекулярного моделирования

*Текст ниже приведен согласно статье [A5] «*

Хроматосома, состоящая из нуклеосомного ядра, линкерных участков ДНК и линкерного гистона H1 (ЛГ), является важным структурным элементом хроматина. Существует два экспериментально подтвержденных типа связывания ЛГ с нуклеосомой и линкерной ДНК, которые различаются своей геометрией - связывание “на диаде” и “вне диады”. Показано, что на тип связывания гистона и конформацию хроматосомы влияет аминокислотная последовательность ЛГ. Однако, при связывании ЛГ изменяется в том числе и геометрия линкерной ДНК. Взаимовлияние этих факторов и молекулярные основы, определяющие тип связывания ЛГ и нуклеосомы, остаются неясными. В данной работе мы применили методы молекулярного моделирования, включая моделирование по гомологии, анализ атом-атомных взаимодействий и расчет энергии деформации ДНК для изучения совместного влияния аминокислотной последовательности ЛГ и нуклеотидной последовательности ДНК на конфигурацию хроматосомы. Были проанализированы известные кристаллические и ЯМР структуры хроматосомы на предмет атом-атомных взаимодействий ЛГ и ДНК, а также энергии деформации ДНК в этих структурах для различных последовательностей ДНК. Для различных вариантов ЛГ анализ проводился с использованием методов моделирования по гомологии. Были обнаружены зависящие от последовательности различия в энергии изгиба линкерной ДНК для двух различных конформаций хроматосомы, а также предложены нуклеотидные последовательности, предпочтительные для этих структур. В результате анализа было показано, что нуклеотидная последовательность ДНК наряду с аминокислотной последовательностью ЛГ оказывает влияние на тип свя-

зывания с нуклеосомой. Сформулированы гипотезы для экспериментальной проверки, согласно которым тип связывания ЛГ может меняться при изменении нуклеотидной последовательности ДНК.

Структурными единицами хроматина являются нуклеосомы - ДНК-белковые комплексы, содержащие гистоны. Ядро нуклеосомы образовано октамером гистонов H2A, H2B, H3 и H4, вокруг которого располагается двойная спираль ДНК длиной около 146 пар оснований [165]. Нуклеосома обладает осью псевдосимметрии второго порядка, называемой также диадной осью (рисунок 4.14А). Эта ось проходит через центр нуклеосомальной ДНК, называемый диадой. Следующим уровнем компактизации является хроматосома [525], образованная при связывании линкерного гистона H1 (ЛГ) и нуклеосомы, включающей линкерные участки ДНК, выходящие за пределы ядра нуклеосомы (рисунок 4.14А). Таким образом, в хроматосому входит октамер гистонов H2A, H2B, H3 и H4, нуклеосомальная ДНК, ЛГ и линкерные участки ДНК. ЛГ представлены во многих эукариотических организмах несколькими вариантами, различающимися длиной аминокислотной последовательности и ее составом [A15]. В различных клетках и тканях могут экспрессироваться различные варианты ЛГ. Как правило, у наиболее просто устроенных организмов встречается всего один вариант ЛГ, тогда как, например, у человека известно 11 вариантов [A9], некоторые из которых экспрессируются только в половых клетках. Также стоит отметить, что присутствие того или иного варианта ЛГ зависит не только от типа клетки, но и от стадии клеточного цикла, в которой она находится. Большая часть ЛГ содержит около 200 аминокислотных остатков. ЛГ состоят из трех доменов: короткий и неупорядоченный N-конец, за которым идет глобулярный домен, образованный 70-80 аминокислотными остатками и имеющий консервативную третичную структуру, и длинный C-концевой домен из примерно 100 аминокислотных остатков. C-конец является неорганизованным и содержит много остатков лизина. Согласно экспериментальным данным [392] глобулярный домен ЛГ связывается с нуклеосомой наравне с полноразмерным ЛГ. Ввиду большой степени неупорядоченности концевых доменов отсутствуют кристаллические структуры полноразмерного ЛГ, но существуют кристаллические структуры глобулярного домена [526]. В данный момент областью для обсуждений является конфигурация связывания ЛГ с нуклеосомой. Ранее существовали различные модели связывания ЛГ с нуклеосомой. Некоторые из них даже предполагали встраивание ЛГ между витками нук-

леосомальной ДНК и октамером гистонов [527]. Подобные модели не нашли впоследствии экспериментального подтверждения [392; 528] в отличие от моделей, в которых предполагается связывание ЛГ в области диады нуклеосомы и линкерных участков ДНК. Имеющиеся на текущий момент экспериментальные данные [202; 529–531] предполагают два возможных расположения ЛГ в хроматосоме, условно называемые как «на диаде» (НД, “on-dyad”) и «вне диады» (ВД, “off-dyad”). На данный момент получена кристаллическая структура хроматосомы с глобулярным доменом ЛГ Н5 *G. gallus* (Н5 является историческим названием ЛГ Н1 у *G. gallus*) в конфигурации НД, а также модель структуры хроматосомы с глобулярным доменом ЛГ Н1 *D. melanogaster* в конфигурации ВД [530]. Важно отметить, что структура хроматосомы в конфигурации ВД была построена с помощью методов молекулярного докинга на основании структуры тетрануклеосомы (pdb-код 1zbb) на основании данных ЯМР. Известно, что ЛГ содержит ряд аминокислотных остатков, играющих важную роль в формировании контактов с нуклеосомой [532]. Эти остатки выделены в два сайта связывания: первый - Н25, R47, K69, K73, R74, K85, и второй - R42, R94, K97. Глобулярный домен гистона Н5 *G. gallus* продемонстрировал связывания типа НД с нуклеосомой, тогда как глобулярный домен Н1 *D. melanogaster* продемонстрировал связывание типа ВД. Анализ их последовательностей показал отличия, вероятно играющие роль в предпочтительности того или иного типа связывания. Так, глобулярный домен Н5 содержит положительно заряженные аминокислотные остатки (R47, K55, R74, K97) в позициях соответствующих нейтрально заряженным аминокислотным остаткам в глобулярном домене Н1 (L68, T76, S96, A119), и нейтрально заряженные аминокислотные остатки (Q51, V80, V87) в позициях соответствующим положительно заряженным остатками Н1 (K72, K102, K109) [530; 531]. Несмотря на то, что некоторые из ключевых остатков не образуют прямых контактов с нуклеосомой, при их замене в глобулярном домене Н5 на аминокислотные остатки, характерные для Н1, ЛГ показал изменение типа связывания с НД на ВД [531]. Другим фактором, который предположительно может оказывать влияние на конфигурацию хроматосомы, является нуклеотидная последовательность ДНК, как линкерной, так и входящей в состав нуклеосомы. По данным исследования позиционирования нуклеосом выдвигались предположения, что ЛГ предпочтительно связывается с АТ-богатыми участками ДНК [533]. Несмотря на то, что в полученной кристаллической структуре хроматосомы ЛГ показал

связывания НД, по данным, полученным методами молекулярного моделирования [263; 534], глобулярный домен ЛГ Н5 может также демонстрировать тип связывания ВД. Одним из объяснений может служить тот факт, что нуклеотидные последовательности ДНК, используемые при моделировании, отличались от нуклеотидной последовательности, для которой была разрешена кристаллическая структура НД конформации комплекса нуклеосомы с глобулярным доменом Н5 [530]. В описанных выше работах рассматривается связывание ЛГ с одиночной нуклеосомой *in vitro*, тогда как в клетке представлены структуры более высокого уровня компактизации - 30 нм фибриллы, представляющей собой цепочку из нуклеосом и ЛГ, связанных с линкерной ДНК. Согласно последним данным крио-электронной микроскопии в структуре 30 нм фибриллы, образованной 12 нуклеосомами, ЛГ демонстрирует связывание ВД, что может быть обусловлено геометрией линкерной ДНК в структуре фибриллы [202; 438]. Таким образом нельзя исключать влияние ДНК, вызванное “узнаванием” ЛГ специфической формы ДНК при связывании [522]. Исходя из вышесказанного, весьма вероятно, что на тип связывания ЛГ с нуклеосомой и линкерной ДНК влияет не только аминокислотная последовательность самого гистона, но и нуклеотидная последовательность ДНК, а также структура хроматиновой фибриллы. Известно также, что при образовании комплекса белок-ДНК на силу и тип связывания оказывают влияние как прямые взаимодействия белка с парами оснований и сахаро-фосфатным остовом, так и геометрия ДНК за счет энергии ее деформации [522]. Однако, несмотря на наличие структур хроматосомы в различных конфигурациях совместное влияние этих факторов ранее не изучалось. В данной работе методами молекулярного моделирования, а именно анализа атом-атомных взаимодействий, моделирования по гомологии и оценки энергии деформации ДНК, было проведено изучение совместного влияния аминокислотной последовательности ЛГ и нуклеотидной последовательности ДНК на тип конфигурации хроматосомы. Проведенный анализ позволил предсказать наиболее предпочтительные для того или иного типа связывания нуклеотидные последовательности линкерной ДНК и сформулировать гипотезы для экспериментальной проверки.

## Материалы и методы

*Анализ аминокислотных последовательностей различных гистоновых вариантов* В работе были использованы аминокислотные последовательности ЛГ, размещенные в базе данных HistoneDB [A14]. Выравнивание доступных аминокислотных последовательностей проводилось программой Muscle [535]. Для дальнейшего анализа последовательностей использовался biopython. *Моделирование по гомологии* Моделирование по гомологии осуществлялось в программе MODELLER [536; 537]. В качестве белков-шаблонов использовались кристаллическая структура хроматосомы в конфигурации НД с pdb-кодом 4qlc и модель хроматосомы в конфигурации ВД, полученная на основании данных ЯМР[530]. Для каждого исследуемого гистона с помощью моделирования по гомологии было построено по 10 моделей, из которых на основании DOPEscore были отобраны наилучшие структуры для анализа контактов. Анализ контактов проводился с помощью программного пакета Chimera [28]. *Расчет энергии деформации ДНК* Расчет зависимости деформационной энергии линкерных участков ДНК от ее последовательности производили в пространстве обобщенных переменных Tilt, Roll, Twist, Shift, Slide, Rise; переход от атомных координат к обобщенным осуществляли при помощи программы 3DNA [99]. Вычисление деформационной энергии участков ДНК производили в соответствии с формулой (1.31). Для расчетов применяли набор упругих коэффициентов и средних значений для нуклеотидных пар описанный в работе [105]. Деформационную энергию определяли для всех возможных последовательностей ДНК для четырех участков ДНК (по два на каждую модель связывания ЛГ) длиной по 12 н.п. каждый. Участок ДНК, рядом с которым располагается линкерный гистон в модели ВД был назван входом в последовательность, позиционирующую нуклеосомы (ППН), противоположный участок был назван выходом из ППН. Для каждого варианта последовательности была рассчитана разница в деформационной энергии ДНК между моделями НД и ВД. Полученное распределение разниц энергий визуализировали при помощи графического представления консервативности нуклеотидов - логотипа последовательности (sequence logo) для 5

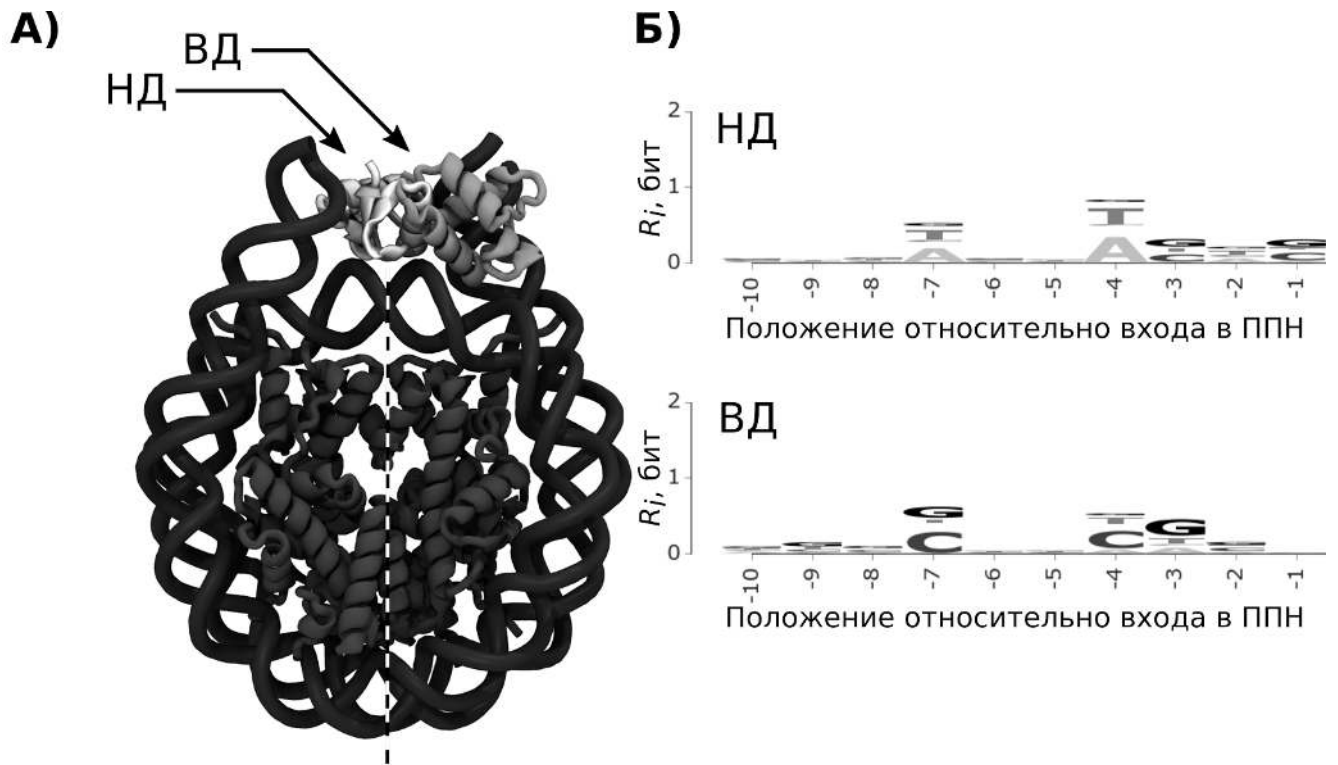
## Результаты и обсуждение

Настоящее исследование состояло из нескольких взаимодополняющих компонент. На основе анализа экспериментальных данных были выделены ключевые остатки ЛГ, способствующие тому или иному типу связывания, на основании которых была построена классификационная модель, которая была применена к ЛГ человека: H1.1, H1.2, H1.3, H1.4, H1.5, H1.0 (H1°), TS H1.6 (H1T), TS H1.7 (H1T2, HANP1), OO H1.8 (H1oo), TS H1.9 (HILS1), H1.10, H1.11, а также к ЛГ *X. laevis*. Для экспериментально определенных структур хроматосомы и структур с различными вариантами ЛГ, построенных с помощью моделирования по гомологии были проанализированы контакты между ЛГ и ДНК. Был проведен оригинальный анализ конформации ДНК в различных структурах хроматосомы и анализ пространства последовательностей линкерных ДНК с точки зрения энергии их изгиба и контактов с ЛГ.

*Определение ключевых остатков ЛГ, влияющих на тип связывания* Аминокислотная последовательность ЛГ имеет ряд ключевых позиций, мутации в которых приводят к изменению типа связывания [531]. Таким образом, на основании аминокислотных остатков, расположенных в этих позициях, для ЛГ может быть определен предполагаемый тип связывания, что позволяет предложить классификацию ЛГ по типу их связывания с нуклеосомой на основании аминокислотных последовательностей. В экспериментально определенной структуре хроматосомы в конфигурации НД и ВД присутствуют контакты непосредственно с парами оснований и линкерной ДНК, и нуклеосомальной ДНК, данные взаимодействия представлены как водородными связями, так и Ван-дер-Ваальсовыми взаимодействиями. Некоторые из аминокислотных остатков, расположенных в ключевых позициях (например, аминокислотные остатки, соответствующие K55, Q51, V87 ЛГ H5) не образуют прямых контактов с ДНК в известных структурах. Однако, несмотря на отсутствие контактов с ДНК, данные остатки предположительно могут взаимодействовать с ДНК посредством электростатического потенциала, поскольку в гистонах H1 и H5 они имеют различные заряды. *Классификация ЛГ на основании наличия ключевых остатков* На основании предложенной выше классификации, было показано, что для 3 ЛГ человека предполагаемый тип связывания был определен как НД, для 7 гистонов человека тип связывания



не может быть определен однозначно, поскольку они демонстрируют сходство в ключевых позициях и с гистоном Н5 *G. gallus* (тип связывания НД), и с гистоном Н1 *D. melanogaster* (тип связывания ВД), для 1 гистона человека предполагаемый тип связывания может быть определен однозначно как ВД. Для ЛГ *X. laevis* тип связывания был определен как НД.



**Рисунок 4.14** — Структура хроматосомы и оптимальные последовательности линкерной ДНК. А) Внешний вид структуры хроматосомы. Белок и ДНК показаны в виде отображения вторичной структуры. Линкерные гистоны выделены оттенками серого. Диадная ось показана пунктирной линией. Б) Визуализация частоты встречаемости нуклеотидов для последовательностей, склонных к формированию структуры НД (сверху) и ВД (снизу). Визуализация выполнена в виде логотипов последовательностей.

*Энергия деформации ДНК в различных конфигурациях хроматосомы*  
 Для определения зависимости типа связывания ЛГ от нуклеотидной последовательности линкерной ДНК была рассчитана разница деформационных энергий линкерных участков ДНК между моделями НД и ВД для всех возможных вариантов последовательностей. Также была произведена Z-оценка положения последовательности из оригинальных моделей в распределении рассчитанных деформационных энергий ( $Z=0,75$ ), исходя из которой оригинальная последовательность, использованная для экспериментального получения структур,

склонна к формированию структуры НД. Как видно из рисунка 4.14Б, большинство позиций в последовательностях нуклеотидов линкерных участков ДНК не значимы для определения геометрии ДНК в рамках моделей НД и ВД, кроме нуклеотидов в позициях -7, -4 и -3. В последовательностях, предпочтительных для модели НД, в этих позициях располагаются А/Т, А/Т, G/C, в то время как для модели ВД более предпочтительны C/G, C/T, G/T. Богатость линкерной ДНК тимидинами также была показана в работе [533]. Также по разнице в деформационных энергиях были найдены предпочтительные (оптимальные) последовательности линкерной ДНК (для которых энергия изгиба максимально способствует той или иной конформации) для моделей НД (CCGTCCCGTCC-ППН-ACGCCGGCGG) и ВД (GACGCCCGAC-ППН-GTGATGCTGC).

*Анализ совместного влияния аминокислотной последовательности ЛГ и нуклеотидной последовательности ДНК* На основании предложенной ранее классификации для различных вариантов ЛГ с помощью моделирования по гомологии были построены структурные модели хроматосом в конфигурациях НД и ВД. В полученных моделях был проведен анализ контактов между ЛГ и ДНК. И в тех, и в других моделях присутствуют контакты как между аминокислотными остатками ЛГ и сахаро-фосфатным остовом ДНК, так и непосредственно с азотистыми основаниями, что говорит в пользу предположений о том, что нуклеотидная последовательность ДНК может оказывать влияние на тип связывания ЛГ и нуклеосомы за счет эффекта прямого считывания последовательности ДНК белком (direct readout). Также количество этих контактов изменяется в зависимости от варианта ЛГ, для которого построена модель. Взаимодействия между ЛГ и азотистыми основаниями ДНК могут быть, как водородными связями, так и ван-дер-Ваальсовыми взаимодействиями. Также для ЛГ Н1 *D. melanogaster* и Н5 *G. gallus* с помощью моделирования по гомологии были построены модели хроматосом в конфигурациях, противоположных экспериментальным: для ЛГ Н1 использовалась конфигурация НД, для ЛГ Н5 - конфигурация ВД. Данные модели показали уменьшение количества контактов между ЛГ и ДНК, что согласуется с определенными экспериментально типами связывания. На основании расчета энергии деформации ДНК были построены модели хроматосомы в конфигурации НД и ВД с предпочтительными (оптимальными) для них последовательностями, определенными по разнице в энергии деформации ДНК, для которых также были проанализированы контакты между ЛГ и ДНК. Анализ показал, что при замене исходной нуклеотидной

последовательности в экспериментальных структурах ВН и НД на оптимальные последовательности, способствующие исходному типу связывания, определенные в нашей работе, количество контактов ЛГ с парами оснований либо не меняется (ВД 3->3), либо увеличивается (НД 2->3). В то же время, перекрестное сравнение эффектов оптимальных последовательностей ДНК в НД и ВД структурах показало, что структуры с оптимальными последовательностями ДНК, соответствующими своему типу связывания имеют такое же или большее число контактов между ЛГ и парами оснований, чем структуры, в которых используется последовательность ДНК, соответствующая альтернативному типу связывания. Таким образом, оптимальные последовательности линкерной ДНК для различных конформаций хроматосомы, предложенные выше, могут достигать своей избирательности как за счет механизмов непрямого считывания последовательности ДНК белком (indirect readout), так и прямого считывания (direct readout) - взаимодействие белка с парами оснований. Исходя из вышесказанного, можно сделать предположение, что нуклеотидная последовательность, а также геометрия и изгибная жесткость ДНК являются важным фактором определяющим тип конфигурации хроматосомы. Это предположение на данный момент согласуется косвенным образом с экспериментальными данными [202; 438]. Вероятно, аминокислотная последовательность ЛГ не является первостепенным фактором определяющим тип связывания ЛГ с нуклеосомой. Подводя итог, можно сформулировать гипотезу, согласно которой один и тот же ЛГ в зависимости от нуклеотидной последовательности ДНК, а также от геометрии линкерной ДНК, может показывать различные типы связывания. Такие пары последовательностей ДНК предложены в данной работе. В дальнейшем эта гипотеза может быть проверена путем оценки расстояний между нуклеотидами, методом измерения эффективности Фёрстеровского переноса энергии (spFRET), с использованием различных нуклеотидных последовательностей для каждого ЛГ.

### Благодарности

Авторы выражают благодарность Y. Bai за предоставленную модель хроматосомы. Работа выполнена с использованием оборудования Центра кол-

лективного пользования сверхвысокопроизводительными вычислительными ресурсами МГУ имени М.В. Ломоносова [343]. Работа выполнена при финансовой поддержке Российского научного фонда (проект №14-24-00031, соглашение №14-24-00031-п). »

#### 4.5.2 Моделирование структуры ДНК в хроматосоме по данным FRET

*Текст ниже приведен согласно статье [A11] «*

В работе данного подраздела рассматриваются методы построения трехмерных моделей ДНК в комплексе с белками на основании компьютерного моделирования и непрямых методов изучения конформации макромолекул. Предложен алгоритм поиска конформации комплексов белков с ДНК на основе данных ферстеровского резонансного переноса энергии (FRET) и информации о локальной гибкости ДНК. Алгоритм апробирован на примере построения гипотетической модели ДНК в нуклеосоме при связывании гистона H1.

Структура ДНК-белковых комплексов влияет на течение множества процессов, таких как транскрипция, репликация и репарация. В исследованиях структуры додекамеров ДНК было показано, что ДНК значительно отличается от идеальной структуры двунитевой спирали. Необходимость существования более высоких уровней организации структуры ДНК очевидна, так как геном эукариотического организма не может разместиться в относительно компактном ядре в полностью развернутом виде. В 1974 году Р. Корнбергом [117] были открыты объекты - структурные единицы компактизации хроматина, которые позже были названы нуклеосомами. Структура нуклеосом долгое время оставалась неясной, и лишь в 1997 году методом рентгеноструктурного анализа (РСА) была определена первая структура нуклеосомы с почти атомарным разрешением [165]. Нуклеосома представляет собой октамер белков гистонов, который несет на себе 145-147 нуклеотидных пар. ДНК закручена вокруг октамера, образуя 1,65 витка левозакрученной суперспирали. Белковое ядро нуклеосомы образует цилиндр диаметром 65 Å и высотой 60 Å. На уровне единичных нуклеосом происходит тонкая регуляция работы генетического аппарата клетки. Одним из ярких примеров регуляции экспрессии генов

является так называемый нуклеосомальный барьер [538]: в зависимости от типа и структуры нуклеосомы, РНК полимеразы II может иметь разную вероятность прохождения сквозь нуклеосому. Помимо этого, нуклеосомы взаимодействуют со множеством разнообразных транскрипционных факторов и структурных белков. Существует большое количество гистоновых вариантов, последовательность и структура которых отличается от канонических, что определяет их специфические свойства [539]. Например, в областях центромер хромосом ДНК находится на специфических нуклеосомах, содержащих модифицированный гистон CenH3 (CENP-A). Нуклеосомы являются ключевыми и фундаментальными единицами упаковки ДНК, их взаимодействия, а также пространственная конфигурация хроматина определяют процессы реализации и передачи генетической информации. Изучение нуклеосом было сильно облегчено с появлением информации об их атомарном устройстве, и на данный момент в банке данных PDB насчитывается более 100 структур различных нуклеосом. Однако изучение более высоких уровней организации хроматина обычными методами структурной биологии затруднено. На данный момент, одной из крупнейших структур, полученной методом РСА, является комплекс из четырех нуклеосом [539]. Хроматиновые фибриллы трудно изучать при помощи РСА, так как структуры такого размера крайне сложно кристаллизовать, к тому же места расположения нуклеосом не фиксированы строго [243], а значит хромосомы в кристалле будут значительно отличаться друг от друга. Информацию о структуре можно получать косвенно биофизическими и биохимическими методами, например, методом ферстеровского резонансного переноса энергии (FRET) и методом фут-принтинга ДНК. Метод FRET основан на явлении безизлучательного переноса энергии с одного красителя на другой. Красители подбирают таким образом, чтобы донор флуоресцировал в области спектра поглощения акцептора, и метят ими аминокислоты или азотистые основания ДНК. Вероятность переноса энергии значительно зависит от расстояния между флуоресцентными метками, таким образом, измеряя эффективность FRET можно оценить дистанцию между участками меченой макромолекулы. В работе [191] таким методом были показаны тонкие различия (5 Å) между канонической и центромерной нуклеосомой. Измерения FRET можно производить в растворе, при этом усредняется сигнал от всех, находящихся в нем молекул, но измерения такого рода не позволяют судить о конформационных переходах внутри исследуемых молекул. Помимо этого, измерения можно производить в режиме единичных молекул,

собирая флуоресценцию лишь с малого конфокального объема в растворе. Такие измерения позволяют различать конформационные переходы в структуре молекул, определять заселенность разных конформаций а так же оценивать времена таких переходов. Таким образом, FRET позволяет получить набор расстояний между парами оснований в меченой макромолекуле. Комбинируя эту информацию с механическими свойствами ДНК можно судить о трехмерной структуре изучаемых комплексов.

Моделирование ДНК и нуклеосом можно производить в полноатомном разрешении [A46] и рассчитывать энергию конформации ДНК при помощи традиционных силовых полей (AMBER, CHARMM и т.д.). Однако расчет энергии и поиск оптимальной конфигурации для таких моделей отличается высокой ресурсоемкостью. Для повышения скорости расчетов применяют огрубленные модели. В одной из таких моделей каждая нуклеотидная пара описывается шестью переменными (три трансляционных координаты и три вращательных), как и каждый шаг, между нуклеотидными парами [97]. Таким образом для цепи из  $N$  нуклеотидных пар конфигурация ДНК описывается набором из  $12N - 6$  обобщенных переменных. Такой подход, в отличие от других огрубленных моделей позволяет однозначно воссоздавать атомарную структуру молекулярного комплекса. Дополнительно в данную модель можно ввести пространственные ограничения, полученные из экспериментов FRET, позволяя получить конфигурацию исследуемого объекта в атомарном разрешении.

## Материалы и методы

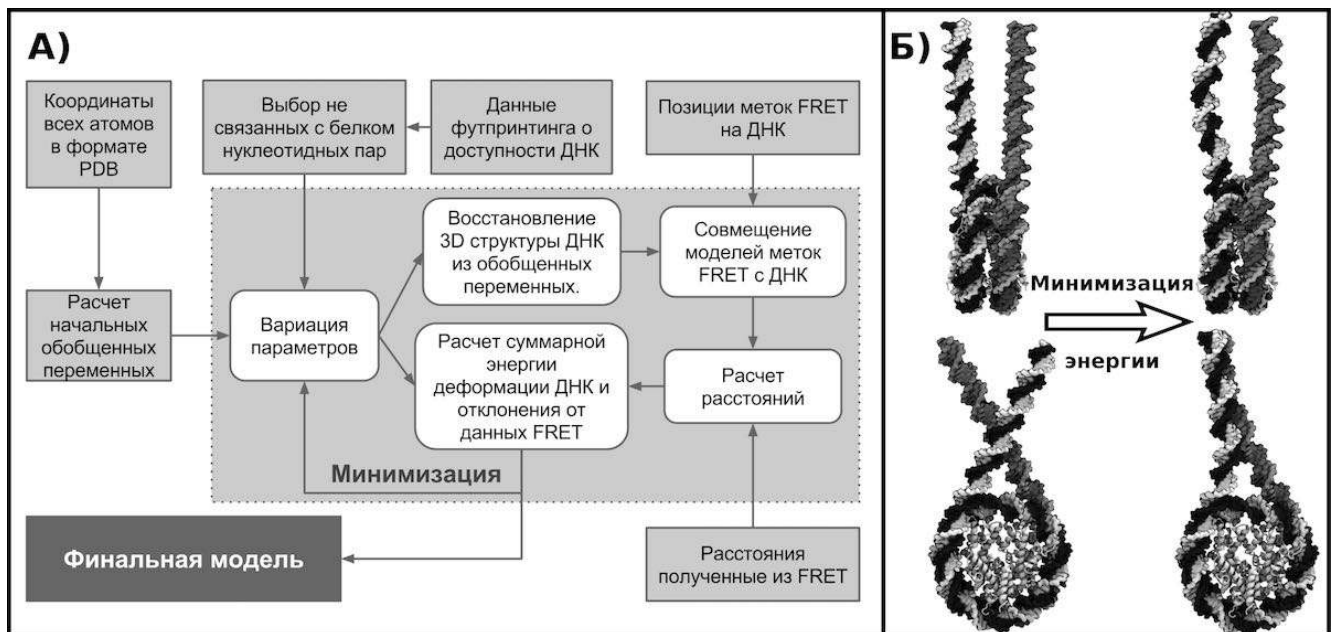
*Интерпретация данных FRET одиночных молекул* Расчет расстояний по данным FRET производился в соответствии с формулой (1.35). Ферстеровский радиус полагался равным  $56\text{Å}$  для пары меток Cy3/Cy5. Расстояние, рассчитанное для пары красителей, не соответствует расстоянию между нуклеотидными парами, так как метки закреплены на ДНК при помощи линкерного участка из 10-15 атомов углерода. Для учета смещения были созданы молекулярные модели меток. При помощи метода молекулярной динамики был построен набор их конформеров и определены координаты наиболее вероятного местонахождения меток. Моделирование производилось в программе Gromacs [338] без учета

электрических взаимодействий при температуре 500 К. Полученные координаты центров меток в дальнейшем использовались для расчета расстояния между нуклеотидными парами при оптимизации геометрии ДНК.

*Расчет энергии деформации ДНК* Энергия деформации рассчитывалась из обобщенных переменных, описывающих конфигурацию ДНК в соответствии с формулой (1.31). Коэффициенты жесткости ДНК для отклонения обобщенных переменных для каждой нуклеотидной пары взяты из работы [105]. Расчет обобщенных переменных из атомистической структуры и восстановление координат из набора переменных производился при помощи программы 3DNA [99]. Минимизация энергии производилась по методу сопряженных градиентов. Все расчеты производились на персональном компьютере в программной среде python с модулем scipy.

## Результаты и обсуждение

Энергию изгиба ДНК можно рассчитывать при помощи силовых полей, которые применяют для моделирования методом молекулярной динамики, однако ресурсоемкость таких расчетов и проблемы с явным учетом электростатических взаимодействий ограничивают применение такого подхода. Другим способом расчета энергии деформации является применение набора эмпирических силовых коэффициентов, которые были получены путем анализа геометрии ДНК из множества различных кристаллических структур [105]. Расчет конформации в обобщенных координатах можно производить быстрее, так как число переменных в модели значительно меньше. Так как минимизация происходит в пространстве обобщенных координат конформации ДНК, а расчет расстояний в трехмерном пространстве, на каждом шаге минимизации требуется производить расчет атомарной модели ДНК. Общий алгоритм поиска конфигурации ДНК, отвечающий критериям жесткости нуклеиновой кислоты, расстояниям, полученным из FRET, а также ограничениям, полученным из анализа профилей футпринтинга показан на рисунке 4.15.



**Рисунок 4.15** — Создание моделей ДНК в хроматосоме по данным FRET. А) Схема алгоритма поиска конформации ДНК с учетом FRET и других данных. Пунктиром выделен блок минимизации энергии при поиске структур. Б) Визуализация результата поиска конформации хроматосомы. Гистоны отображены в виде вторичной структуры, ДНК показана в поверхности доступной растворителю.

К недостаткам такого подхода можно отнести отсутствие учета взаимодействий между участками ДНК, что может привести к возникновению стерических перекрываний внутри молекулы. Однако при корректном выборе участков ДНК для минимизации, появление таких структур маловероятно. В качестве объекта для испытания метода была использована нуклеосома с гистоном Н1. Этот комплекс на момент проведения работы отсутствовал в банке трехмерных структур PDB, но были работы, где его структура исследуется при помощи ядерного магнитного резонанса [529]. Гистон Н1, также называемый линкерным, образует комплекс с нуклеосомами, способствуя более высокому уровню компактизации ДНК. Оказывая такое влияние на размещение/упорядочение ДНК в клетке, гистон Н1 играет большую роль в регуляции экспрессии генов. Структурные особенности взаимодействия линкерного гистона и нуклеосомы до сих пор неизвестны и их изучение вызывает большой интерес. За основу для построения модели была выбрана структура 1KX5 [238] из банка данных PDB к которой были добавлены прямые линкеры в соответствии с последовательностью, что применялась в эксперименте FRET. После проведения минимизации геометрии ДНК с ограничениями расстояний из данных



FRET была получена асимметричная структура с параллельным расположением линкерных участков ДНК. Весьма схожие модели были получены в работе [392], где структуру хроматосом исследовали методом гидроксильного футпринтинга и электронной микроскопии. В дальнейшем полученная структура может быть применена в качестве мишени для макромолекулярного докинга гистона H1. Понимание молекулярных основ функционирования хроматина встречается с проблемами ограниченности экспериментальных методик в определении параметров взаимодействий ДНК с гистонами, транскрипционными факторами и структурными белками на молекулярном уровне. Данная проблема является следствием того, что экспериментальные методы вроде РСА, ЯМР и электронной микроскопии в состоянии определить структуры лишь небольших комплексов, которые обладают компактной и упорядоченной структурой. В то же время все больший интерес привлекают к себе структуры, не отличающиеся строгой упорядоченностью и их описание хроматина должно вестись не в рамках одиночных структур, а в рамках набора конформационных ансамблей. Разработан метод, позволяющий создавать модели геометрии ДНК в комплексах с белками на основании интегрирования экспериментальных данных и данных о локальной жесткости ДНК при помощи молекулярного моделирования. Данный метод позволяет получать атомистические модели, позволяя изучать особенности комплексов как внутри, так и с другими. На примере хроматосомы показана возможность получения геометрии линкерной ДНК, что важно для понимания устройства хроматина. *Благодарности* Авторы работы выражают благодарность проф. В.М. Студитскому, проф. А.В.Феофанову и сотрудникам их лабораторий за предоставленные данные FRET. Данная работа выполнена при поддержке гранта Российского научного фонда (грант N14-24-00031).

»

#### 4.6 Выводы главы 4

Автором продемонстрировано, что методы интегративного моделирования способны решать важный класс задач при изучении супрануклеосомной структуры хроматина, исследования которой затруднительны классическими

методами структурной биологии, а именно создавать модели комплексов нуклеосом с белками хроматина на основе использования экспериментальных данных электронной микроскопии низкого разрешения, данных FRET-микроскопии, данных футпринтинга ДНК. Получены структурные и динамические модели различных комплексов, возникающих при взаимодействии нуклеосом с РНК-полимеразами, гистоном H1, комплексом FACT.

## Глава 5. Интегративное моделирование амилоидоподобных фибрилл<sup>1</sup>

В данной главе излагаются подходы, разработанные автором для интегративного моделирования структуры амилоидоподобных фибрилл. Особенностью подходов является использование экспериментальных данных как на этапе конструирования возможных межмолекулярных укладок пептидов в фибриллах, так и на этапе оценки крупномасштабной морфологии фибрилл. Излагаются работы по изучению фибрилл из олигомеров проводящего полимера тиофена и пептидов, а также конструированию положительно заряженных амилоидоподобных фибрилл, ускоряющих вирусную трансдукцию. Материалы данной главы включают данные из диссертации автора на соискание степени Dr. Rer. Nat. [5] и следующих статей [A31; A34; A35].

### 5.1 Амилоиды и материалы на основе амилоидов

В настоящем разделе мы рассмотрим феномен амилоидных фибрилл, обсудим их структуру, свойства самосборки и экспериментальные методы, используемые для их исследования. Также дадим обзор достижений в разработке новых материалов, основанных на ковалентном присоединении синтетических фрагментов к образующим  $\beta$ -лист (амилоидогенным) пептидам. Упор будет сделан на экспериментальных результатах для РЕО-функционализированных алкилированных диблок олигомеров кватротиофен-пептид, которые являются основным объектом следующего раздела.

---

<sup>1</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [A31; A34; A35].

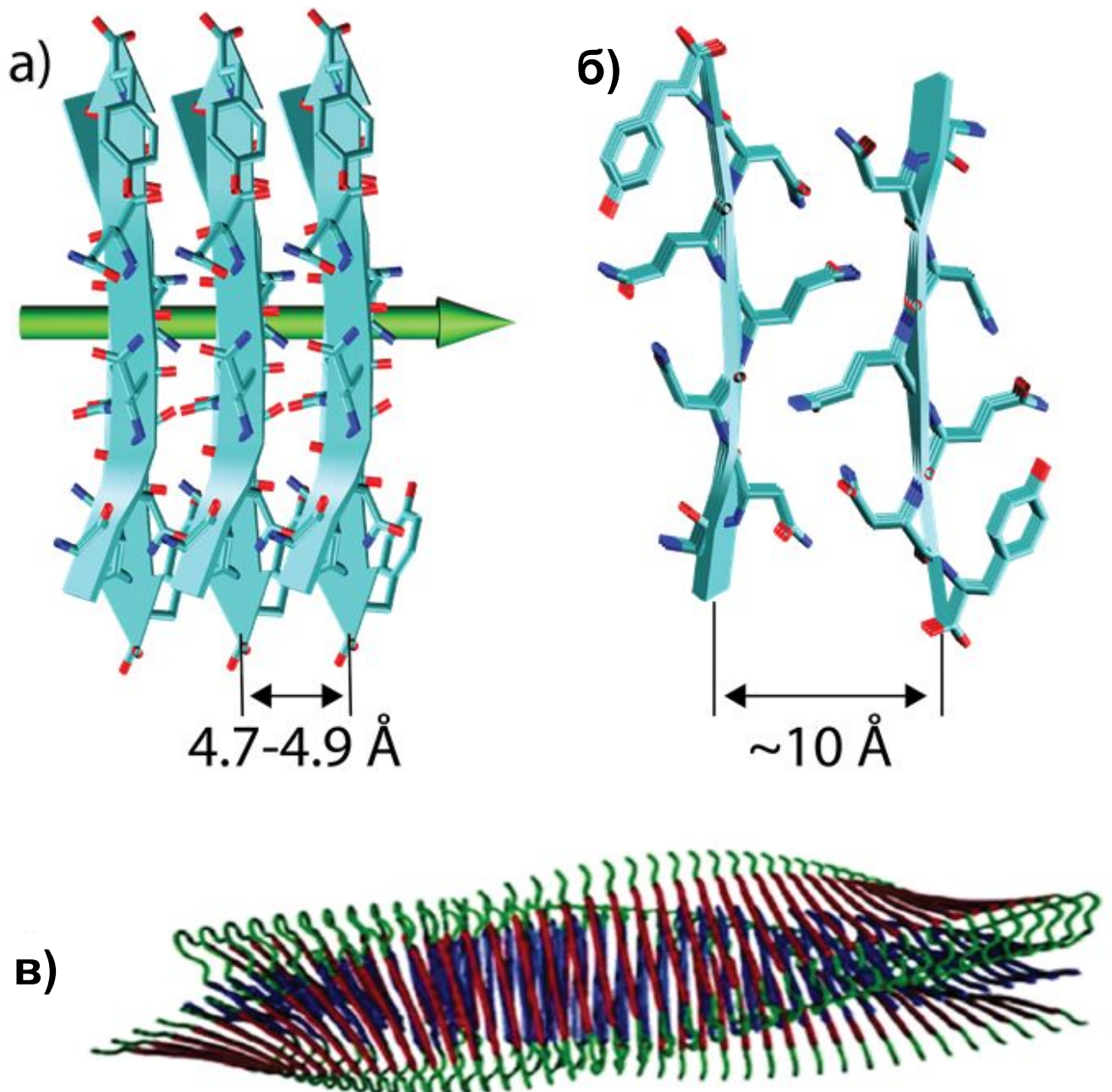
### 5.1.1 Строение природных амилоидных фибрилл и их полиморфизм

Термин “амилоидные фибриллы” в его современном понимании относится к удлиненным белковым агрегатам, характеризующимся своей длиной и относительно прямой морфологией, характерными картинами дифракции [540], специфическими свойствами связывания некоторых красителей и жесткой структурой ядра [541]. На сегодняшний день идентифицировано более 20 белков плазмы, которые могут образовывать патологические отложения амилоидного характера в организме человека, что связано со многими разрушительными заболеваниями, такими как болезнь Альцгеймера, болезнь Паркинсона, диабет II типа, болезнь Хантингтона, болезнь Крейтцфельдта-Якоба, прионные расстройства и многими другими [542–544]. Известно, что патологическое образование амилоидных фибрилл включает структурную перестройку нативного состояния белков в конформацию фибрилл, богатую  $\beta$ -слоями [545]. Недавно выяснилось, что практически все белки потенциально могут собираться в амилоидные фибриллы [542; 546] и были получены многие модельные последовательности *de novo*, которые агрегируют в фибриллы при определенных условиях [547].

Амилоидные фибриллы имеют диаметр примерно 7-10 нм (по данным электронной микроскопии (ЭМ) или атомно-силовой микроскопии (АСМ) [542–544]) и обычно состоят из 2-6 протофиламентов. Внутренняя структура всех амилоидных фибрилл имеет один и тот же структурный мотив, перекрестный  $\beta$ -мотив, который выявляется характерной картиной дифракции рентгеновских лучей [540] с меридиональными рефлексами соответствующими 4,7–4,9 Å (так называемое расстояние между основными цепями) и более диффузными экваториальными рефлексами соответствующими приблизительно 10Å (наблюдались вариации от 8,8 до 14,6Å [548]) (так называемый интервал боковых цепей), эти рефлекссы свидетельствуют о высоком порядке и периодичности внутренней структуры фибрилл вдоль и перпендикулярно направлению фибрилл. Понимание молекулярных деталей самосборки пептидов в фибриллярные агрегаты было сложной задачей из-за большого размера, низкой растворимости и некристаллической и гетерогенной природы фибрилл. В течение последнего десятилетия значительный прогресс в нашем понимании принципов образования фибрилл был достигнут благодаря многочисленным экспериментальным и

теоретическим исследованиям и, что важно, разрешению расположения пептидов на атомистическом уровне с помощью рентгеновской кристаллографии и твердотельного ЯМР. Для получения дополнительной информации можно рекомендовать ряд недавних обзоров по структуре и образованию амилоидных (-подобных) фибрилл [541; 549; 550]. Понимание структуры амилоидных фибрилл значительно улучшилось с помощью рентгеноструктурных исследований микрокристаллов амилоид-образующих сегментов амилоидогенных белков [551; 552]. Эти исследования показывают, что перекрестные  $\beta$ -мотивы в амилоидных фибриллах, образованные этими амилоидогенными сегментами, состоят из пары  $\beta$ -листов (рис. 5.1), состоящих из выровненных пептидных фрагментов. Два таких  $\beta$ -листа дополняют друг друга, образуя пару листовых структур, в которых боковые цепи, выступающие из двух листов, интеркалируют, образуя сухую “стерическую молнию” (рис. 5.1б), которая считается важным стабилизирующим взаимодействием.

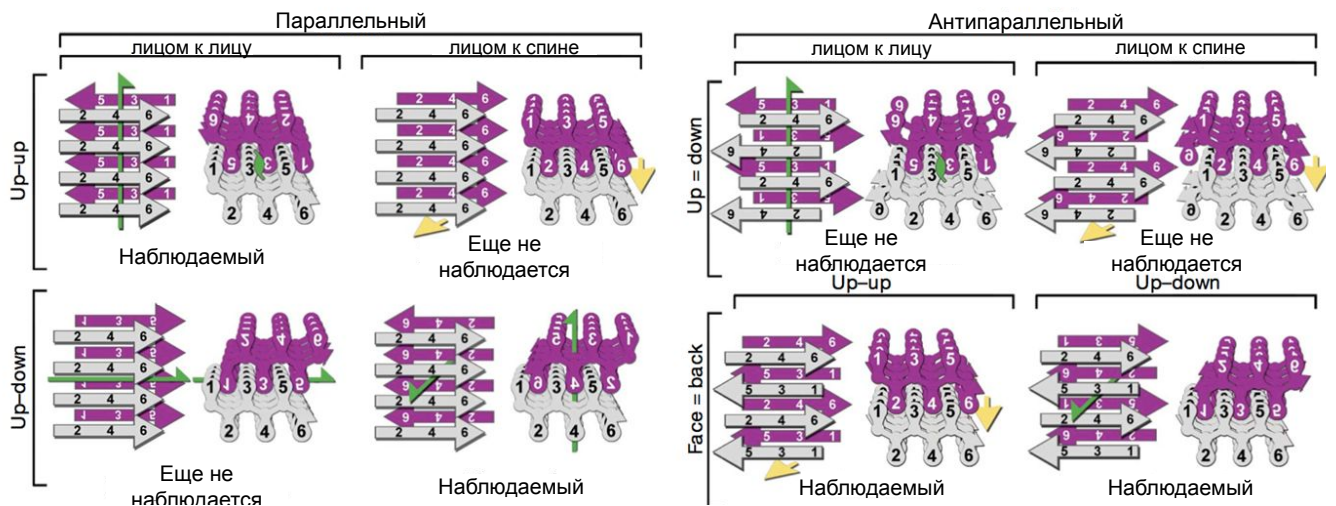
Исследования методами твердотельного ЯМР в сочетании с измерениями по изображениям электронной микроскопии позволили предложить прямые атомистические модели расположения молекул в скрученных протофиламентах амилоида- $\beta_{1-40}$  [553; 554] (рис. 5.1в). Две  $\beta$ -цепи каждой молекулы амилоид- $\beta_{1-40}$  соединены через изгибную область, содержащую остатки 25–29, и являются частями двух разных совпадающих параллельных  $\beta$ -листов, взаимодействующих через свои боковые цепи в одном протофиламенте. Одиночный протофиламент амилоид- $\beta_{1-40}$ , по-видимому, включает два перекрестных  $\beta$ -мотива, то есть четыре  $\beta$ -листа с расстоянием между листами  $\sim 1$  нм. Протофиламент проявляет левостороннюю спиральную закрутку, что характерно для  $\beta$ -листов.



**Рисунок 5.1** — Модели расположения  $\beta$ -тяжей в амилоидных протофибриллах посредством поперечной  $\beta$ -структуры. Вид сбоку (а) и вид сверху (б) перекрестной  $\beta$ -структуры ядра фибриллы на примере микрокристаллов пептидных фрагментов дрожжевого белка Sup35 [551], показано расстояние между основной и боковой цепями. Зеленая стрелка показывает направление фибрилл. На панели (б) видна стерическая застежка-молния между боковыми цепями соседних листов. (в) Ленточная диаграмма протофиламента амилоид- $\beta_{1-40}$ , если смотреть параллельно оси фибриллы. Структурная модель основана на данных твердотельного ЯМР в сочетании с ограничениями из данных электронной микроскопии. Каждая молекула  $A\beta$  вносит две  $\beta$ -цепи в параллельные  $\beta$ -листы.

Основано на материалах статьи [553]

Вышеупомянутые примеры, однако, далеко не единственные возможные молекулярные структуры, согласующиеся с перекрестным  $\beta$ -мотивом. Внутренняя структура последних может существенно различаться в разных фибриллах. Рентгеновские структуры 13 микрокристаллов ряда амилоидобразующих сегментов амилоидогенных белков, решенные Sawaya et al. [552] пролили дополнительный свет на то, как перекрестный  $\beta$  мотив может проявлять вариации. В зависимости от ориентации  $\beta$ -нитей (параллельных или антипараллельных) внутри листов, от ориентации  $\beta$ -листов (параллельных или антипараллельных) относительно друг друга или от упаковки листов (лицом к лицу или лицом к спине) 8 различных основных классов расположения “стерических застежек-молний” могут в принципе согласовываться с перекрестный  $\beta$  мотивом (см. рис. 5.2), 5 из которых уже наблюдались в решенных атомистических структурах микрокристаллов. Твердотельный ЯМР амилоидных фибрилл, образованных различными пептидами, гомологичными участкам пептида Альцгеймера, также показал, что  $\beta$ -лист может располагаться по-разному - параллельно или антипараллельно внутри протофиламентов в зависимости от свойств полипептида-предшественника [555–558], хотя, как описано ранее, исследования фибрилл, образованных из полноразмерного амилоида- $\beta$ , показали, что пептид сворачивается в  $\beta$ -изгибную структуру, которая затем связывается с другими молекулами, образуя параллельную  $\beta$ -структуру [553; 559].

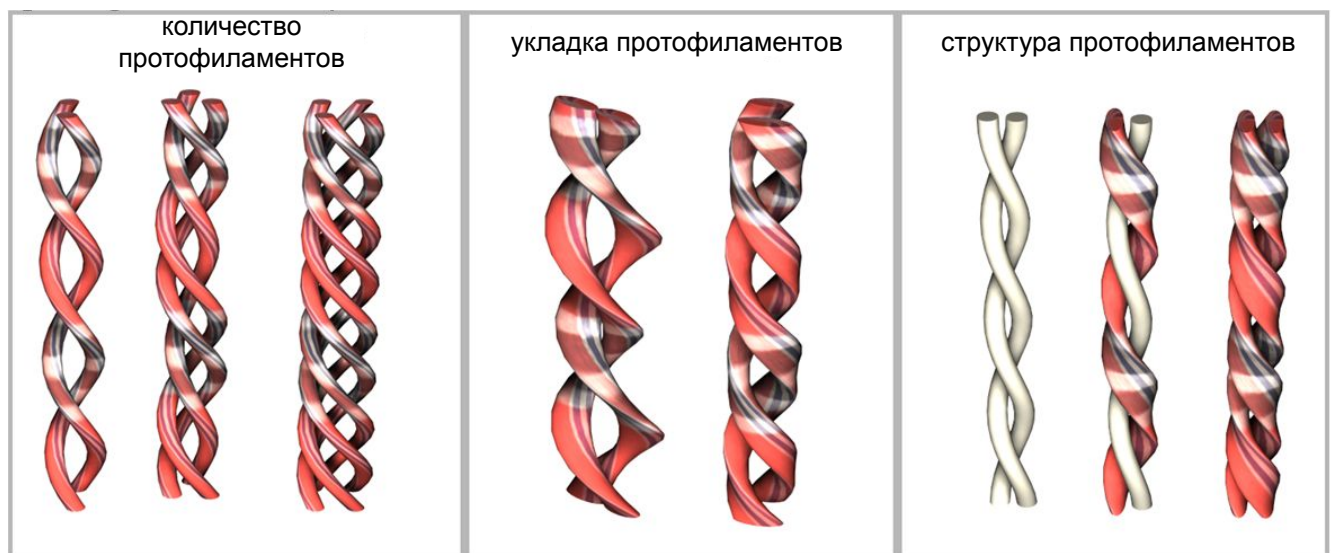


**Рисунок 5.2** — Возможные варианты расположения  $\beta$ -листов в перекрестном  $\beta$ -мотиве. Два идентичных листа можно классифицировать по: ориентации их граней (либо “лицом к лицу”, либо “лицом к спине”), ориентации их нитей (при этом оба листа имеют одинаковый край нити “вверх”, или один “вверх”, а другой “вниз”), а также от того, параллельны ли нити в листах или антипараллельны. Оба вида сбоку (слева) и виды сверху (справа) показывают, какой из шести остатков сегмента образует застежку-молнию, а какой - смотрит наружу. Зеленые стрелки показывают винтовые оси симметрии второго порядка, а желтые стрелки показывают трансляционную симметрию. Под каждым классом перечислены белковые сегменты, принадлежащие этому классу. По материалам Sawaya et al. [552].

Хотя за последние десятилетия был достигнут значительный прогресс, мы все еще далеки от того, чтобы сказать, что полная картина самосборки фибрилл теперь понятна как на молекулярном, так и на нано- и микромасштабах. Оказывается, амилоидные (-подобные) фибриллы имеют тенденцию быть высокополиморфными структурами [550] со сложным взаимодействием на молекулярном уровне, которые могут приводить к широкому спектру различных агрегатов с различной морфологией. Структурный и морфологический полиморфизм амилоидных фибрилл простирается намного дальше, чем просто зависимые от последовательности вариации структуры перекрестного  $\beta$ -мотива - одни и те же полипептидные последовательности, как известно, образуют фибриллы различной морфологии в зависимости от таких факторов, как значение pH, температуры, наличие перемешивания, солей и других соразтворенных веществ [558; 560–564]. Даже в одних и тех же условиях и в одном образце могут существовать значительные вариации морфологии фибрилл. Различное



расположение протофиламентов может приводить к различным морфологиям амилоидных фибрилл. Эти фибриллы могут различаться по нескольким структурным свойствам, таким как толщина поперечного сечения фибриллы, шаг спирали, способ слипания и переплетения амилоидных протофиламентов с образованием трехмерной структуры фибрилл. Fändrich et al. суммировали доступные данные и определили три возможности того, как амилоидные фибриллы могут различаться по структуре (см. рисунок 5.3) [550]. Во-первых, фибриллы могут состоять из разного количества протофиламентов. Во-вторых, фибриллы могут различаться относительной ориентацией протофиламентов. В-третьих, фибриллы могут различаться по своей субструктуре протофиламентов, а именно, конформации лежащих в основе пептидов в перекрестных  $\beta$ -мотивах, природе и регистре  $\beta$ -листов, в количестве остатков в  $\beta$ -цепях, а также в расстоянии между  $\beta$ -листами.



**Рисунок 5.3** — Структурные типы полиморфизма амилоидных фибрилл. Схематическое изображение различных морфологий амилоидных фибрилл, которые различаются количеством, относительной ориентацией или структурой лежащих в основе протофиламентов.

### 5.1.2 Экспериментальные методы, используемые для изучения амилоидоподобных волокон

В настоящее время для изучения амилоидных (-подобных) фибриллярных агрегатов применяется множество различных экспериментальных подходов, включая рентгеновскую и электронную дифракцию, оптическую и электронную микроскопию, АСМ, твердотельный ЯМР, рассеяние нейтронов, ИК, УФ-видимую и КД спектроскопию, связывание красителей в сочетании со сложной подготовкой и модификацией фибрилл и агрегирующих молекул.

В то время как химическая структура агрегирующих соединений почти всегда известна, а морфология фибрилл в субмикронном масштабе решается с помощью электронной или атомно-силовой микроскопии, структура фибрилл в атомистическом и наноразмерном масштабах, включая упаковку отдельных молекул, очень часто остается недоступной для возможностей экспериментальных измерений. В конкретных случаях становится возможным получить представление о внутренней структуре фибрилл, когда могут быть получены соответствующие микрокристаллы (с помощью дифракции рентгеновских лучей) или когда доступны достаточные данные ЯМР. Однако эти данные, хотя и бесценно полезны, до сих пор ограничиваются лишь несколькими небольшими пептидами при определенных условиях. Для большинства соединений и, более того, для полимер-биоконъюгатов имеющиеся экспериментальные данные о межмолекулярных взаимодействиях в большинстве случаев ограничиваются спектроскопическим анализом (ИК, УФ-видимая, КД-спектроскопия) и характерными дифракционными картинками (рентгеновские лучи, электроны) и, следовательно, точное структурное расположение в наномасштабе и его связь с морфологией остаются плохо понятными.

### 5.1.3 Самосборка и образование фибрилл

Большое внимание уделяется раскрытию механизма образования амилоидных фибрилл, поскольку понимание механизмов, лежащих в основе полимеризации растворимого мономерного пептида в зрелые нерастворимые

фибриллы, может пролить свет на возможные терапевтические подходы к остановке, обращению или предотвращению образования фибрилл. Процесс образования амилоидных фибрилл, по-видимому, начинается с частично структурированных конформеров белков [565; 566]. Частичное сворачивание/разворачивание белков, по-видимому, способствует определенным межмолекулярным взаимодействиям, таким как гидрофобные и электростатические взаимодействия, которые необходимы для стимулирования полимеризации белковых молекул в амилоидные фибриллы. Это термодинамически неблагоприятное состояние затем переходит в стабильную амилоидогенную форму.

Для некоторых белков предлагается модель полимеризации зависящей от нуклеации при описания образования фибрилл, аналогичная описанию процесса кристаллизации [567]. Гетерогенное ядро (“затравка”) или пептидная мицелла формируется выше критической “пороговой” концентрации, и фибриллы зарождаются внутри них, удлиняясь за счет необратимого связывания мономеров со своими свободными концами [568; 569]. Рост фибрилл может быть представлен схематически в виде кривой экспоненциального роста с отставанием, где фаза значительно укорачивается в присутствии затравок.

Однако становится все более очевидным, что для многих белков такие модели не подходят для извлечения информации о размере ядра из зависимости кинетики образования амилоидных фибрилл от концентрации белка [570]. В случае образования амилоидных фибрилл многими белками наблюдается быстрое образование сферических олигомеров и/или протофибрилл, и во многих случаях зрелые фибриллы появляются при длительной инкубации [571–577]. Этот механизм агрегации получил название “сборка с помощью олигомерных промежуточных продуктов” [571; 578; 579]. В этом механизме, по-видимому, образование префибриллярных агрегатов не ограничивается неблагоприятным событием нуклеации [576; 577; 580] и может рассматриваться как изодесмическая полимеризация [580].

Начальная фаза образования фибрилл многими белками характеризуется накоплением сферических олигомеров и протофибрилл. Электронная микроскопия и эксперименты с атомно-силовой микроскопией показывают, что самые ранние префибриллярные агрегаты представляют собой сферические олигомеры [571–573; 576–582], которые впоследствии, по-видимому, сливаются с образованием гранулированных удлиненных червеобразных амилоидных

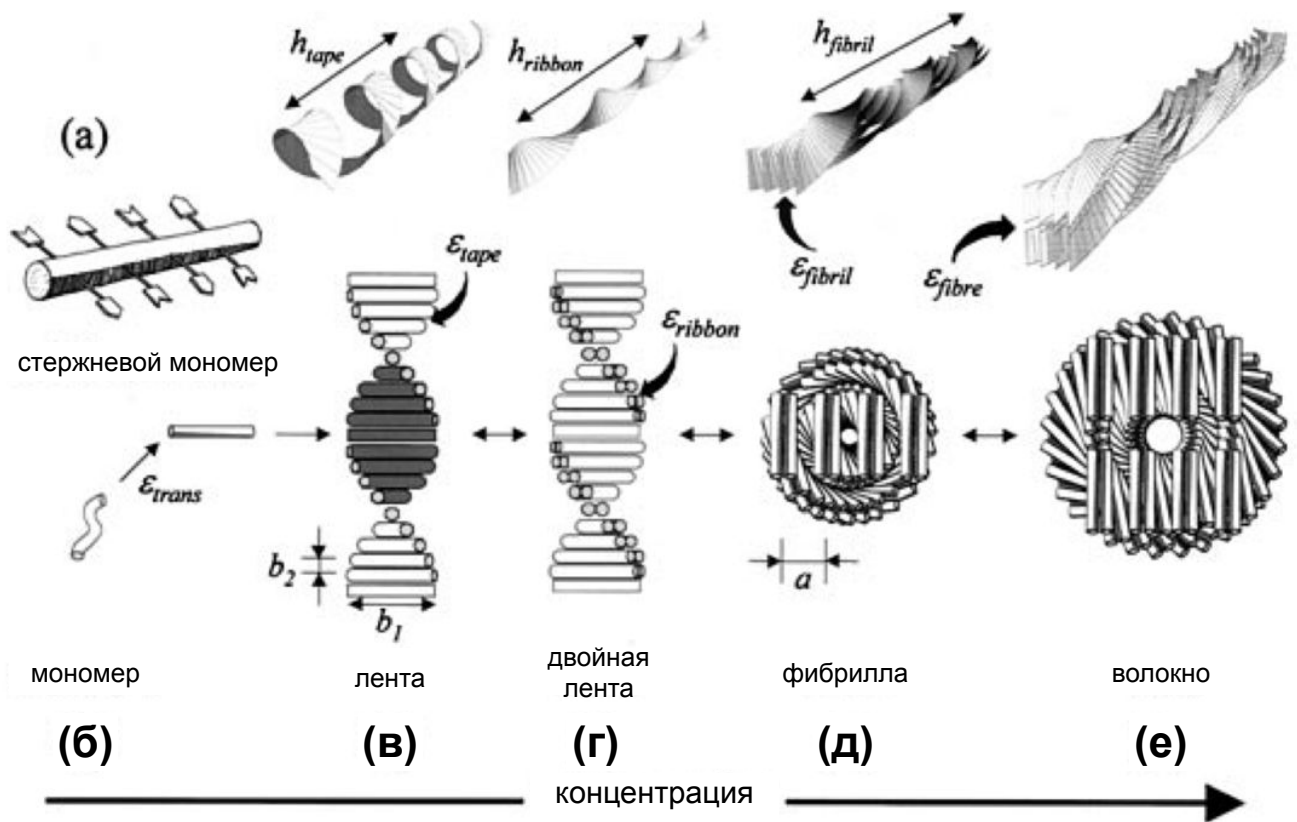
протофибрилл. Удлиненные протофибриллы иногда могут округляться с образованием кольцевых, кольцевидных протофибрилл [583–585]. Недавно было показано, что кольцевые протофибриллы белка амилоид- $\beta$  структурно отличаются от сферических олигомеров; они обнаруживают эпитоп, который отсутствует в сферических олигомерах и в фибриллах белка [585]. Альтернативно, образование фибрилл может идти по “пути смещения” без образования фибрилл, но вместо этого включает превращение промежуточных продуктов в аморфные отложения [549].

Важно определить, когда происходит конформационная конверсия  $\beta$ -листа во время образования амилоидных фибрилл. В реакциях образования амилоидных фибрилл, демонстрирующих характерные особенности механизма подобного кристаллизации, и в большинстве примеров сборки через олигомерные промежуточные соединения, рост агрегатов и приобретение  $\beta$ -листовой структуры, по-видимому, связаны [576–579; 586; 587]. Похоже, что связывающие звенья (мономеры или олигомеры) сначала присоединяются к концам растущих агрегатов, а затем претерпевают конформационные изменения  $\beta$ -слоя. Недавно для двух белков было замечено, что образование амилоидных фибрилл включает конформационно преобразованные олигомерные промежуточные соединения, т.е. конформационные изменения  $\beta$ -складок происходят в олигомерных промежуточных соединениях до того, как они присоединяются к концам растущих агрегатов [574].

Процесс образования амилоидных волокон является относительно медленным, олигомеризация белка амилоида- $\beta$  *in vitro* занимает до 72 часов, прежде чем отдельные фибриллы могут быть разделены [574].

Отличительным признаком самосборки амилоида является наличие иерархии различных взаимодействий, которые управляют фибриллярной структурой и морфологией на разных уровнях организации: например,  $\beta$ -листы  $\rightarrow$  “стерические застежки-молнии”  $\rightarrow$  протофиламенты  $\rightarrow$  амилоидные волокна. Хотя взаимодействие между различными межмолекулярными взаимодействиями и его влияние на морфологию довольно сложно и не до конца изучено, Aggeli et al. предложили хорошую статистико-механическую модель, подкрепленную экспериментальными данными, которая очерчивает основные уровни иерархии и доминирующие межмолекулярные взаимодействия [588]. Модель описывает самосборку хиральных стержнеобразных единиц, таких как пептиды, образующие  $\beta$ -листы, в спиральные ленты, которые с увеличением концентрации

объединяются в скрученные ленты (двойные ленты), фибриллы (скрученные стопки лент) и волокна (переплетенные фибриллы) (см. рисунок 5.4). Показано, что конечная ширина и спиральность фибриллы являются результатом конкуренции между приростом свободной энергии от притяжения между лентами и проигрышем в энергии из-за упругого искажения внутренне скрученных лент при включении в растущую фибриллу. Волокна стабилизируются аналогично. Поведение двух рационально сконструированных пептидов из 11 остатков аминокислот, P11-I и P11-II, иллюстрирует предложенную схему. P11-I и P11-II предназначены для принятия конформации  $\beta$ -нити и для самосборки в одном измерении с образованием антипараллельных  $\beta$ -листовых лент, двойных лент, фибрилл и волокон в четко определенных растворителем условиях.



**Рисунок 5.4** — Модель иерархической самосборки хиральных стержневых единиц. Локальные расположения (в–е) и соответствующие глобальные равновесные конформации для иерархических самособирающихся структур, сформированных в растворах хиральных молекул (а), которые имеют дополнительные донорные и акцепторные группы, показанные стрелками, через которые они взаимодействуют и выравниваются, образуя ленты (в). Черная и белая поверхности стержня (а) отображены на сторонах спиральной ленты (в), которая выбрана так, чтобы закручиваться к черной стороне (в). Все внешние стороны двойной ленты (г), фибриллы (д) и волокна (е) белые. Одна из фибрилл в волокне (е) для ясности нарисована более темным оттенком. (д и е) Вид спереди краев фибрилл и волокон соответственно. Адаптировано из Aggeli et al. [588].

#### 5.1.4 Функциональные амилоиды и синтетические амилоидные агрегаты

Однако амилоидные фибриллы не всегда вредны. Сейчас все чаще становится очевидным, что живые организмы, от прокариот до людей, ис-

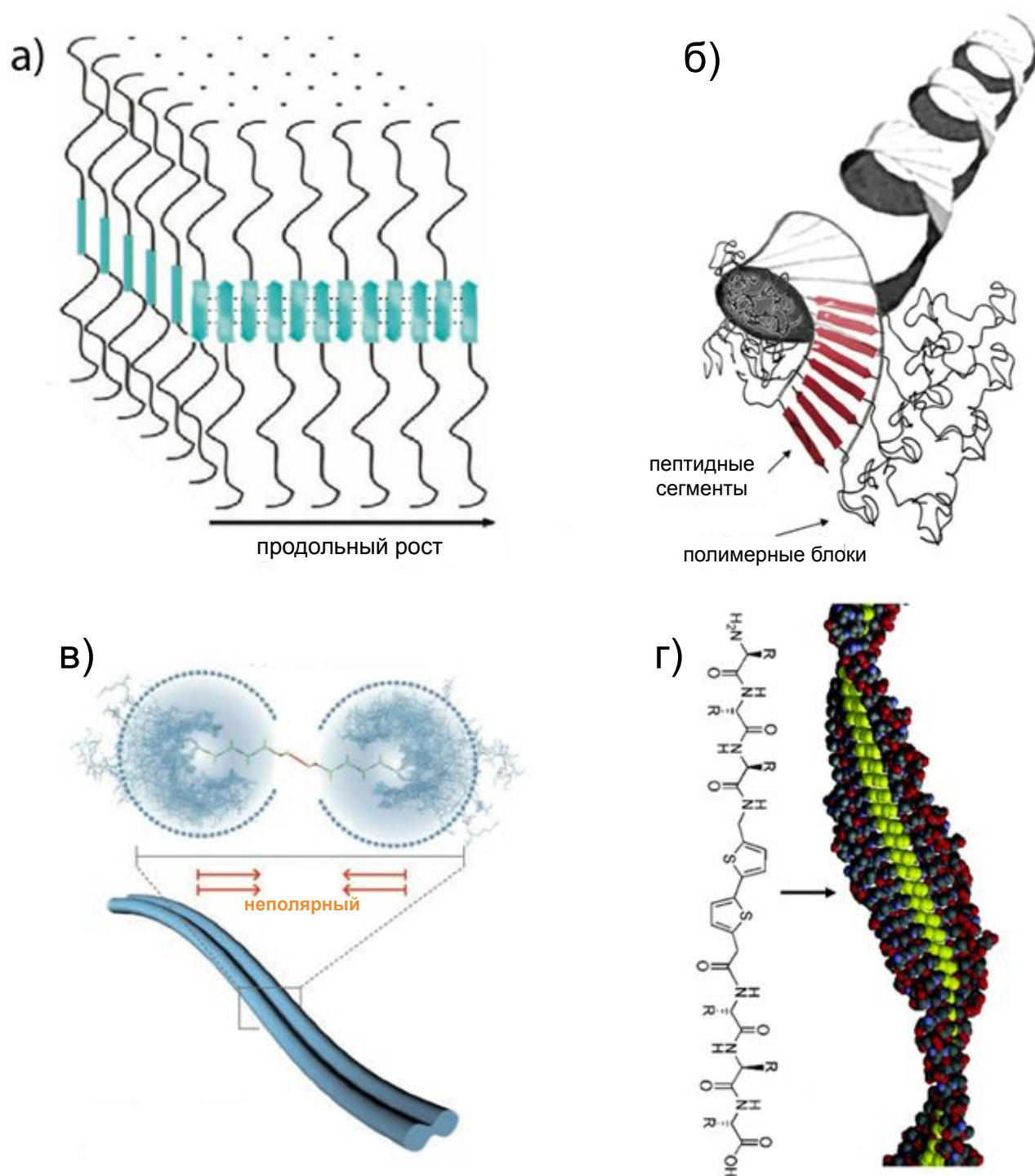
пользуют амилоидные фибриллы, образованные их эндогенными белками, для выполнения нормальных физиологических функций [542; 589]. Многие грибы производят амфипатические белки, называемые гидрофобинами [590]. Они обладают способностью собираться в фибриллы, богатые  $\beta$ -слоями, на границах раздела воздух-вода, и считается, что они играют защитную роль в грибковых структурах, таких как споры и плодовые тела. Предполагается, что прионы дрожжей играют функциональную роль, формируя цитоплазматические фибриллярные сборки, которые могут быть связаны с методом наследственной передачи информации [591]. Амилоидные и амилоидоподобные фибриллярные агрегаты, образованные природными белками или синтетическими пептидами и полимерными биоконъюгатами, привлекают большое внимание как из-за их возможного применения в качестве строительных блоков в нано- и биотехнологиях. Амилоидные фибриллы все чаще исследуются на предмет их потенциальной роли в формировании нанотрубчатых (и не только) каркасов для бионанотехнологий [592–594]. Сами фибриллы могут быть очень прочными, имеющими такую же прочность на растяжение, как сталь [595], свойство, которое они разделяют со своим структурным родственником, шелком. Пептидные волокна можно функционализировать путем присоединения белков [596] или использовать в качестве матрицы для связывания с металлами [592; 597].

Болдуин и др. [596] собрали гибридный белок, состоящий из функционального цитохрома b562 с амилоидогенной последовательностью SH3. Структуры имели амилоидоподобное ядро, при этом обладали функциональными свойствами глобулярного цитохрома. Другой пример - нанопроволоки, которые в одной из работ были изготовлены путем сборки белков, таких как N-терминальная область дрожжевого приона Sup35. Конъюгированные частицы коллоидного золота были связаны вдоль волокон с использованием открытых остатков цистеина Sup35, в результате чего были получены проволоочки диаметром около 100 нм [592]. В другой работе очень короткий пептид, состоящий из двух остатков фенилаланина, собирается с образованием амилоидоподобных нанотрубок, которые можно функционализировать с помощью ионов серебра в центре нанотрубки. В результате были получены нанопроволоки диаметром около 20 нм [597].

Другой подход, привлекающий все большее внимание, - это синтетическое конъюгация пептидов с другими молекулярными соединениями. Конъюгаты синтетических и природных макромолекул представляют большой интерес в

настоящее время из-за их многообещающих биомедицинских, микроэлектронных и других передовых технологических приложений [598–603]. Ковалентное присоединение синтетических полимеров к амилоидогенным пептидным последовательностям приводит к новому классу блок-сополимеров, которые могут унаследовать типичные свойства своих составляющих, например улучшенные рабочие характеристики, проводимость, биосовместимость и высокую склонность к самоорганизации. Обзоры [598–601] обобщают последние достижения в области химии биоконъюгации полимеров, из которых биоконъюгирование с амилоидогенными пептидами является одним из наиболее часто используемых. На результирующее взаимодействие межмолекулярных взаимодействий влияют как синтетические, так и пептидные части, что приводит к еще большему структурному полиморфизму, чем наблюдается в природных амилоидных волокнах, имея в виду, что синтетическая химия обеспечивает большую вариативность в структуре строительных блоков, включая разветвленные молекулярные топологии [604]. Ряд предполагаемых самоорганизующихся морфологий, которые могут быть адаптированы различными полимерными биоконъюгатами, изображены на рисунке 5.5. Более того, взаимодействие между различными взаимодействиями может также указывать на зависимость супрамолекулярной организации от внешних условий (таких как температура, качество растворителя, рН и др.), что перспективно для технологических приложений [605].



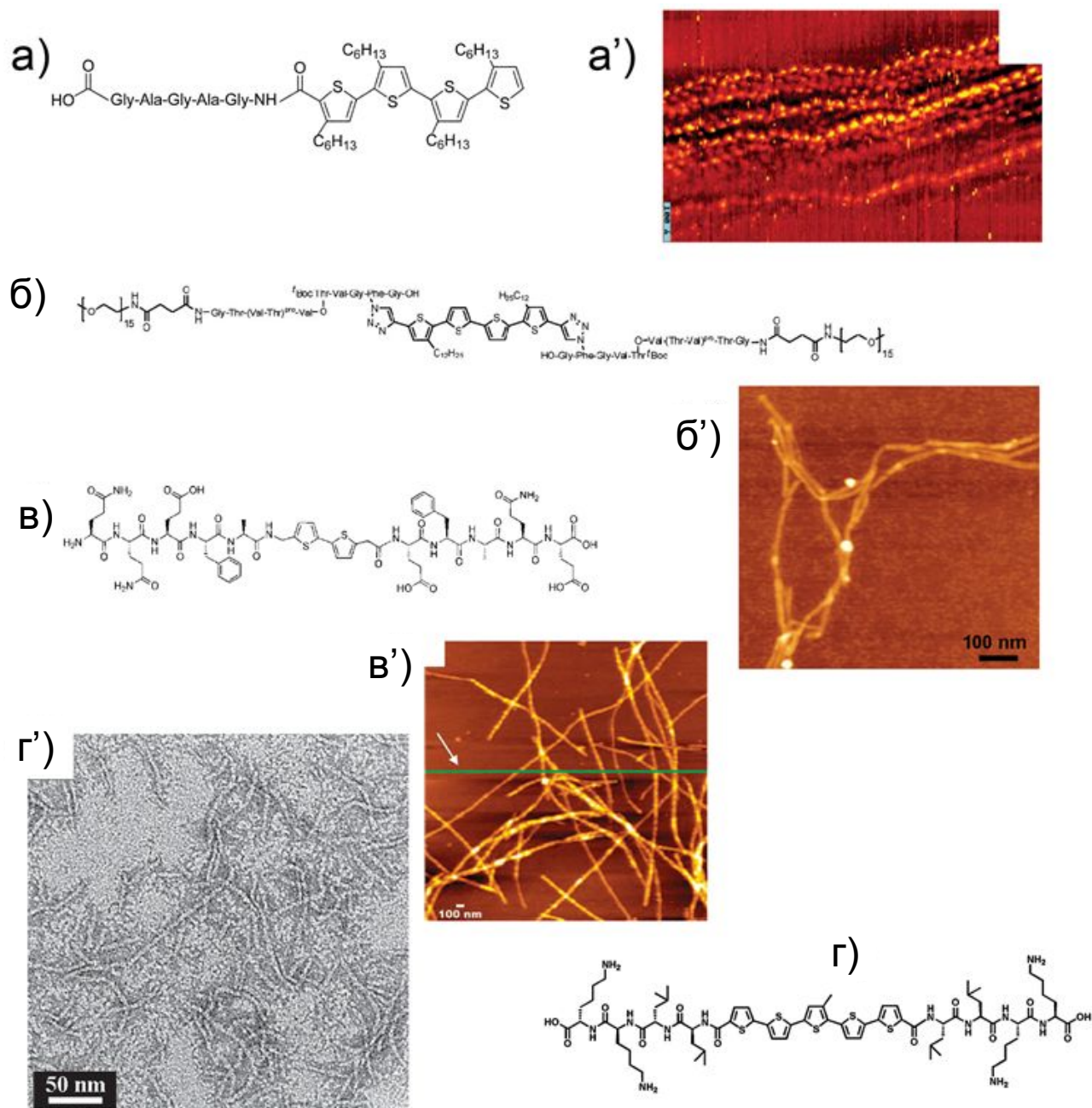


**Рисунок 5.5** — Примеры различных морфологической организации фибриллярных агрегатов, предположительно образуемых полимерными био-конъюгатами: (а) широкие плоские ленты [606], (б) спиральная надстройка [607], (в) ленты с поперечным сечением в форме гантели [608], (г) спиральные ленты с полимером посередине [609]. По материалам [606–609].

## Конъюгаты тиофен-пептид

Как подчеркивалось ранее, ярким примером конъюгатов пептид-полимер являются конъюгаты пептидов, образующих  $\beta$ -лист, и электропроводящих полимеров - тиофенов, которые, как было показано, образуют фибриллы как в воде [609], так и в органических средах [610]. Примеры синтетических “молекулярных химер”, в которых одна часть молекулы состоит из олиготиофена, а другая - из аминокислотной последовательности, образующей  $\beta$ -лист, были представлены несколькими группами. Klock et al. [611] синтезировали молекулу типа TP, в которой кватро(3-гексилтиофен) (T) функционализирован пентапептидной последовательностью Gly-Ala-Gly-Ala-Gly (P), которая, как известно, образует  $\beta$ -листовые домены в шелке тутового шелкопряда (см. рис. 5.6а, а'). Эти авторы обнаружили образование длинных линейных цепей шириной от 3,5 до 4,0 нм и высотой 3Å на поверхности из молекул TP и тенденцию пептидных блоков к формированию  $\beta$ -листов посредством множественных водородных связей. Однако точная архитектура наблюдаемой сверхструктуры еще не известна, и детальная схема водородных связей еще не может быть нарисована. Schillinger et al. [610] провели исследование симметрично замещенного P-T-P-алкилированного гибрида кватротиофен-пептида, который включает последовательность (Thr - Val)<sub>3</sub>, которая, как известно, имеет высокую склонность к принятию  $\beta$ -листовой формы (см. рис. 5.6б, б'). АСМ и ПЭМ снова указали на образование волокнистых структур длиной до 1-2 мкм (АСМ), высотой  $2,4 \pm 0,4$  нм (АСМ) и шириной  $8 \pm 2$  нм (ПЭМ) и  $11 \pm 2$  нм (АСМ). Diegelmann et al. [609] сообщили об успешном синтезе битиофенов, присоединенных к последовательностям Ala-Phe-Glu-Gln-Gln и Glu-Phe-Ala-Gln-Glu (гибрид типа P-T-P) (см. рис. 5.6в, в'). Атомно-силовая микроскопия (АСМ) образцов гелей, нанесенных на слюду, выявила одномерные наноструктуры высотой от 2 до 6 нм. Еще два примера гибридов олиготиофеновых пептидов были представлены Ступпом и сотрудниками. Первоначально был представлен гибридный тертиофен, который включает одну глутаминовую кислоту и два блока аланина [612]. В дальнейшем этот подход был расширен до двух производных центрального пентатиофена, симметрично замещенных двумя последовательностями Lys-Lys-Leu-Leu [613] (см. рис. 5.6г, г'). Обнаруженные наноструктуры имели ширину  $4,94 \pm 0,76$  нм.

Во всех исследованиях с помощью атомно-силовой микроскопии было визуализировано, что при определенных условиях эти соединения могут самоорганизовываться в длинные (в диапазоне микрометров) морфологически похожие фибриллярные агрегаты с шириной фибриллы, обычно варьирующейся от 4 до 50 нм и высотой от 0,3 до 6 нм. Однако взаимосвязь между молекулярным составом, молекулярной конформацией, межмолекулярными силами и трехмерной супрамолекулярной организацией этих материалов мало изучена. Следовательно, необходимы дополнительные экспериментальные и теоретические исследования, чтобы вывести структурную модель, описывающую расположение этих гибридов на основе олиготиофена-олигопептида.



**Рисунок 5.6** — Различные конъюгаты тиофенов и пептидов, образующих  $\beta$ -лист, и соответствующие изображения наблюдаемых фибриллярных агрегатов с помощью АСМ или ПЭМ. По материалам [609–611; 613].

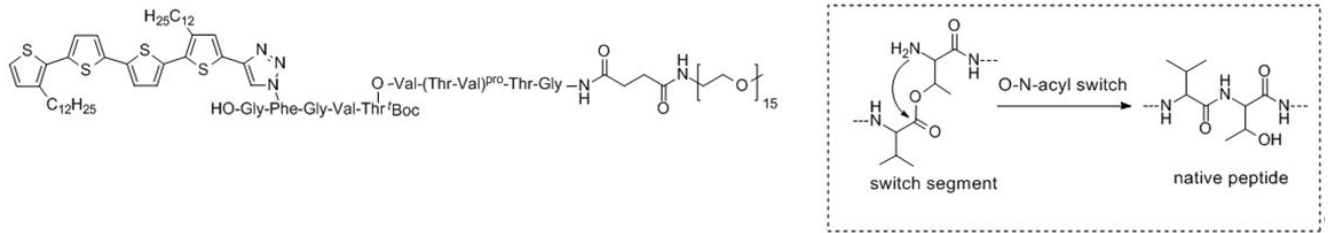
### Диблок-олигомер кватротиофена и пептида

В настоящем разделе мы опишем экспериментальные данные, доступные для самосборки диблочного олигомера кватротиофен- $\beta$ -листовой пептида (изоб-

раженного на рисунке 5.9). Представленные здесь данные были получены от группы профессора Бойерле, университет г. Ульма.

**Гибридный дизайн.** Гибридное соединение Т-Р (рис. 5.9) было разработано путем комбинации симметрично дидодецил-замещенного кватротиофена, функционализированного с одной стороны конъюгатом ПЭГ- $\beta$ -пептид. Последовательность пептида включает три повтора Val и Thr. Ранее было показано, что используемый кватротиофен самособирается в очень регулярные сверхструктуры на границе раздела жидкость-твердое тело (АСМ) и в объеме (дифракция рентгеновских лучей) [614–616]. Известно, что повторы  $(Val - Thr)_x$  образуют стабильные  $\beta$ -листы в водных и органических средах [547; 606; 607].

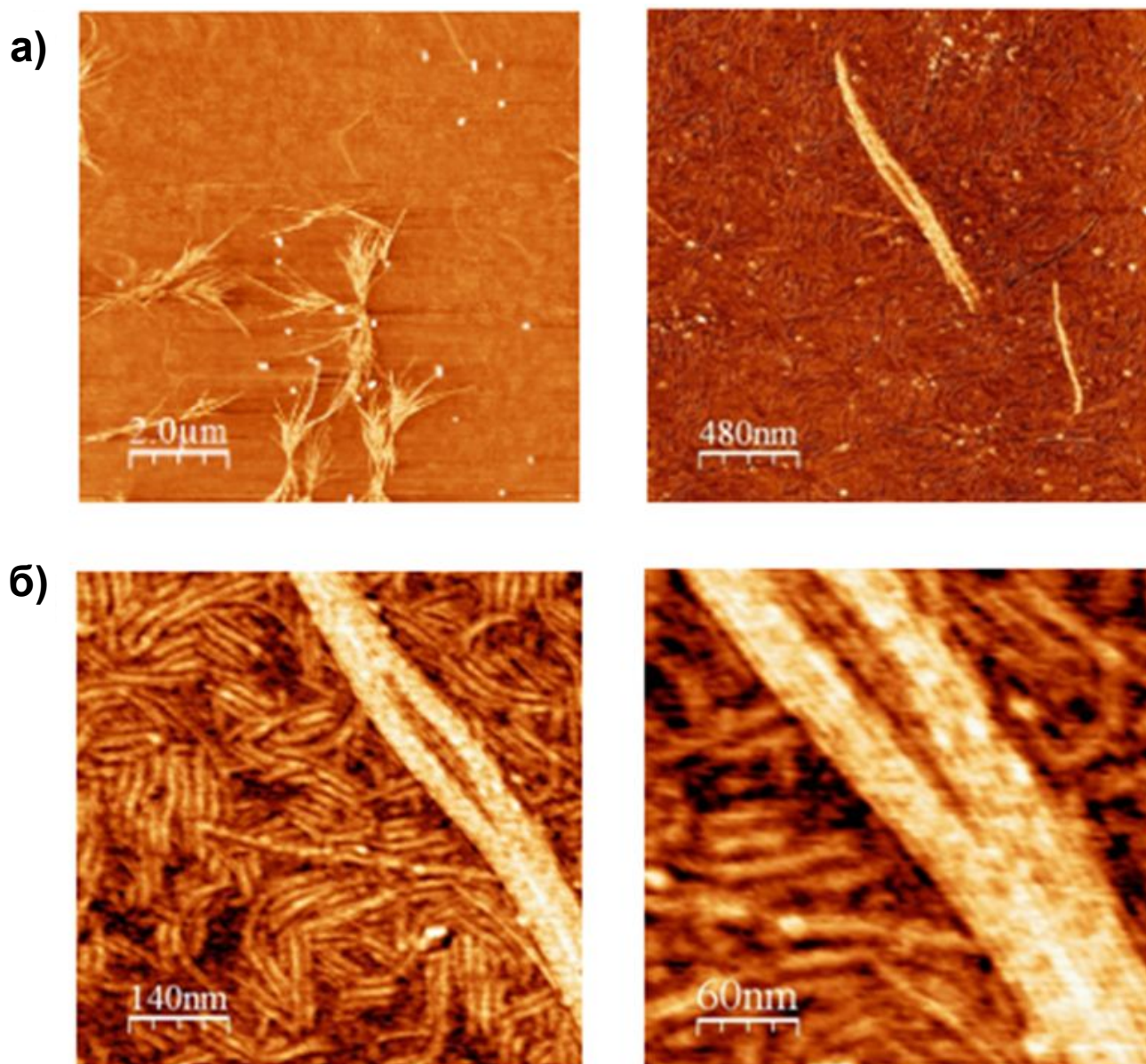
Такие последовательности, образующие  $\beta$ -лист, известны как “сложные последовательности” [604; 617], потому что они имеют тенденцию к агрегированию во время синтеза и обработки. Были реализованы две стратегии, чтобы обойти такую нежелательную агрегацию и в дополнение к получению контроля над процессом самосборки. Во-первых, в пептид была включена так называемая псевдопролиновая единица, которая после кислотного снятия защиты превращается в повтор Val-Thr [618]. Во-вторых, использовали фрагмент переключающегося сложного эфира, который вызывает временный дефект в природной амидной цепи за счет введения связи  $\beta$ -сложного эфира между  $Val^5$  и  $Thr^6$  [619–621]. Дефект переключателя сохраняется в кислых условиях, но нативный амидный остов может быть восстановлен путем увеличения pH, тем самым вызывая перегруппировку в звене переключателя сложного эфира (перенос O- $\rightarrow$  N-ацила) (см. Рисунок 5.7). Для лучшей гибкости полученного гибрида между пара-азидофенилаланином и последовательностью  $\beta$ -листа был введен глицин, потенциально разрушающий вторичную структуру. В конце концов, гибрид был снабжен ПЭГ-цепью на N-конце, чтобы повысить растворимость и предотвратить потенциальные латеральные взаимодействия фибриллярных агрегатов.



**Рисунок 5.7** — Синтезированное соединение Т-Р в “защищенной” форме. На вставке показан переключатель О-*N*-ацила, который срабатывает при повышении рН.

Чтобы способствовать превращению нарушенных пептидных сегментов гибрида (рис. 5.7) в нативные  $\beta$ -цепи, был выбран контролируемый подход путем добавления 0,001 М гидроксида натрия, растворенного в метаноле, к ранее приготовленному раствору в дихлорметане (ДХМ) с помощью шприцевого насоса до тех пор, пока не будет достигнуто соотношение растворителей 1:1. Наличие полностью протяженной последовательности  $\beta$ -листов было доказано с помощью ИК-спектроскопии. В ИК-Фурье спектре отсутствие полос для интактного фрагмента сложного эфира переключателя ( $\nu = 1785^{-1}$  и  $\nu = 1743^{-1}$ ) сопровождается сдвигом полосы амида I в область, специфичную для  $\beta$ -листа ( $\nu = 1636^{-1}$ ), что указывает на присутствие полностью протяженной пептидной структуры, которая участвует в формировании вторичной структуры  $\beta$ -слоя. Из-за довольно широкой формы полосы амида I для гибрида во вторичной структуре  $\beta$ -слоя было, к сожалению, невозможно однозначно определить, антипараллельная (специфическая полоса при  $\nu = 1690^{-1}$ ) или параллельная (отсутствие эта полоса) пептидных цепей присутствует в  $\beta$ -листе.

**Самосборка.** Для несимметричного гибрида ПЭГ- $\beta$ -пептид-олиготиофен исследовали самосборку на поверхности и в растворе. Принимая во внимание различную полярность гибрида (доминирующие гидрофобные взаимодействия со стороны олиготиофена, водородные связи пептидного сегмента) и сильную тенденцию к агрегации  $\beta$ -листового пептида в его нативной форме, для самосборки была выбрана стратегия на основе изменения свойств растворителя, использующая шприцевой насос, для одновременной перестройки пептидного остова и самосборки гибрида [610].



**Рисунок 5.8** — АСМ изображения высоты и фазы нативного ПЭГ-пептида-кватротиофена, нанесенного центрифугированием на слюду из смеси ДХМ/0,001 М NaOH в MeOH, раствор 1:1, приготовленный с помощью шприцевого насоса; а) Высотные изображения АСМ: слева: пучки волокон, полученные через 7 дней, справа: одиночное волокно; б) Фазовые изображения АСМ, слева и справа: увеличенное изображение одного волокна, показывающее состав большего волокна из мелких волокон. Предоставлено Евой-Катрин Шиллингер, группой профессора П. Бойерле.

С помощью АСМ волокна были обнаружены в образце с центрифугированием на подложке из мусковитовой слюды (рис. 5.8). Изображения были получены в тэппинг режиме после стабилизации адсорбата в течение нескольких дней. Можно было визуализировать скопление волокон, рисунок которых не

соответствует сети, а больше напоминает кластеры (рис. 5.8а, слева). Вокруг этих больших участков концентрированного материала можно найти одиночные “более крупные” волокна (рис. 5.8а, слева и справа). Можно показать, что эти более крупные волокна представляют собой пучки нескольких более тонких волокон разной длины, выровненных параллельно (рис. 5.8б, оба изображения), что объясняет нерегулярный вид более крупных отдельных волокон, которые мы поэтому называем пучками. Кроме того, короткие фибриллярные структуры, окружающие пучки, регулярно откладываются на поверхности слюды (рис. 5.8б). Эти короткие фибриллярные структуры можно рассматривать как самые маленькие одиночные структуры, которые затем иерархически самоорганизуются в более крупные структуры (пучки или даже кластеры).

Размер пучков колеблется от 1 до 2 мкм в длину, от  $1 \text{ нм} \pm 0,2 \text{ нм}$  в высоты и от 15 до 48 нм в ширину, хотя также могут наблюдаться спорадические ширины примерно до 80 нм. Самая маленькая самособирающаяся единица на слюдяной подложке (вышеупомянутые короткие фибриллярные структуры, которые предположительно являются однослойными  $\beta$ -листами) имеет длину от 100 до 500 нм и повторяющуюся ширину  $7 \text{ нм} \pm 2 \text{ нм}$  и высоту  $0,5 \pm 0,2 \text{ нм}$ . Несоответствие размеров пучков, особенно ширины, считается результатом предложенной иерархической самосборки, которая ведет от одиночных  $\beta$ -листов (наименьшие наблюдаемые фибриллярные структуры) к двойным лентам, то есть двухслойным  $\beta$ -листам (отдельные волокна в пучке), которые, в свою очередь, взаимодействуют по бокам, образуя пучки.

В просвечивающей электронной микроскопии (ПЭМ) наблюдались структуры, очень похожие на структуры, исследованные с помощью АСМ.

Поскольку фибриллярные объекты можно было наблюдать с помощью АСМ на слюде после центрифугирования, а также в ПЭМ после нанесения методом drop casting на покрытые углеродом медные сетки, можно предположить, что эти нано- и микроструктуры уже присутствуют в растворе, и преобладающие эффекты подложки на самосборке маловероятны [610]. Как упоминалось ранее, также кватротиофеновая часть в гибриде, представленном здесь, способна самоорганизовываться за счет  $\pi - \pi$ -стэкинга и ван-дер-ваальсовых взаимодействий [614–616]. Следовательно, потенциальная роль  $\pi - \pi$  взаимодействий в конечной самоорганизующейся супраструктуре была исследована методами УФ-видимой, флуоресцентной и КД-спектроскопии.



Однако флуоресцентная спектроскопия, как очень чувствительный метод, не выявила какой-либо агрегации, которая непосредственно влияет на  $\pi$ -систему. С помощью спектроскопии КД не удалось получить никаких указаний на хиральное экситонное взаимодействие сопряженных  $\pi$ -систем кватротиофенов. Однако очень слабый неотмеченный сигнал в спектре КД при энергиях ниже максимума поглощения сопряженной  $\pi$ -системы указывал на существование хромофорных агрегатов в хиральном окружении.

## 5.2 Мультимасштабное интегративное моделирование фибрилл из тиофен-пептидных олигомеров

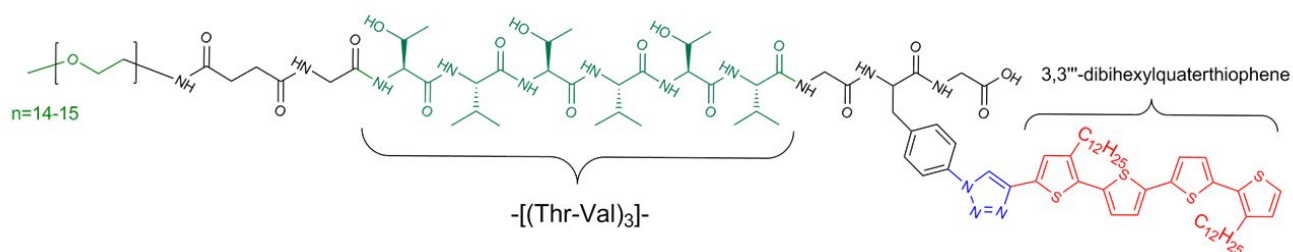
В настоящем разделе мы представляем комплексное моделирование фибриллярных агрегатов из диблочных олигомеров кватротиофена и  $\beta$ -листового пептида, основанное на различных предполагаемых межмолекулярных структурах, инспирированных структурами родственных амилоидных фибрилл. Мы выполняем мультимасштабное исследование, которое состоит из [622] предсказания различных межмолекулярных структур, [623] моделирования коротких фибрилл в явном растворителе, [624] моделирования длинных фибрилл в вакууме, [625] моделирования фибрилл на графитовой подложке. Мы описываем комбинированный интегративный экспериментальный/теоретический подход, направленный на определение структуры и свойств амилоидоподобных фибриллярных агрегатов.

### 5.2.1 Формулировка задачи

Было обнаружено, что симметрично дидодецилзамещенный кватротиофен, функционализированный с одной стороны конъюгатом ПЭГ- $\beta$ -пептид (см. Рисунок 5.9), самособирается в фибриллярные агрегаты в растворе, которые были визуализированы с помощью ПЭМ и АСМ (см. рис. 5.8). Внутренняя структура агрегатов не разрешена экспериментальными методами, хотя спектральные данные предполагают, что структура  $\beta$ -листов формируется в

растворе. В этой работе мы стремимся использовать модели молекулярной механики и моделирование молекулярной динамики для изучения возможных моделей организации гибридных молекул тиофен-пептид (Т-Р) и установления связей между межмолекулярным расположением и надмолекулярной морфологией фибриллярных агрегатов.

Поскольку использование одного подхода или методологии было бы недостаточно для решения этой проблемы, мы проводим комплексное многомасштабное исследование этой системы, которое сочетает в себе методы моделирования с доступными экспериментальными данными и рациональным пониманием иерархии взаимодействий при самосборке. Наше исследование состоит из следующих этапов: (i) построение модели молекулярной механики ТР-соединения и предложение различных межмолекулярных структур, согласующихся с экспериментальными данными и общими принципами образования амилоидных волокон, (ii) оптимизация структуры и оценка этих молекулярных расположений проводится с помощью МД-моделирования одномерных периодических кристаллов, (iii) из выбранных периодических расположений строятся короткие фибриллы ( $\sim 10$  нм в длину), которые моделируются в явном органическом растворителе, (iv) более длинные фибриллы ( $\sim 40$  нм в длину) моделируются в вакууме и изучается их морфология, (v) выполняется специальное моделирование фибрилл, адсорбированных на графите, для изучения эффектов адсорбции и получения прямого соответствия с данными АСМ.

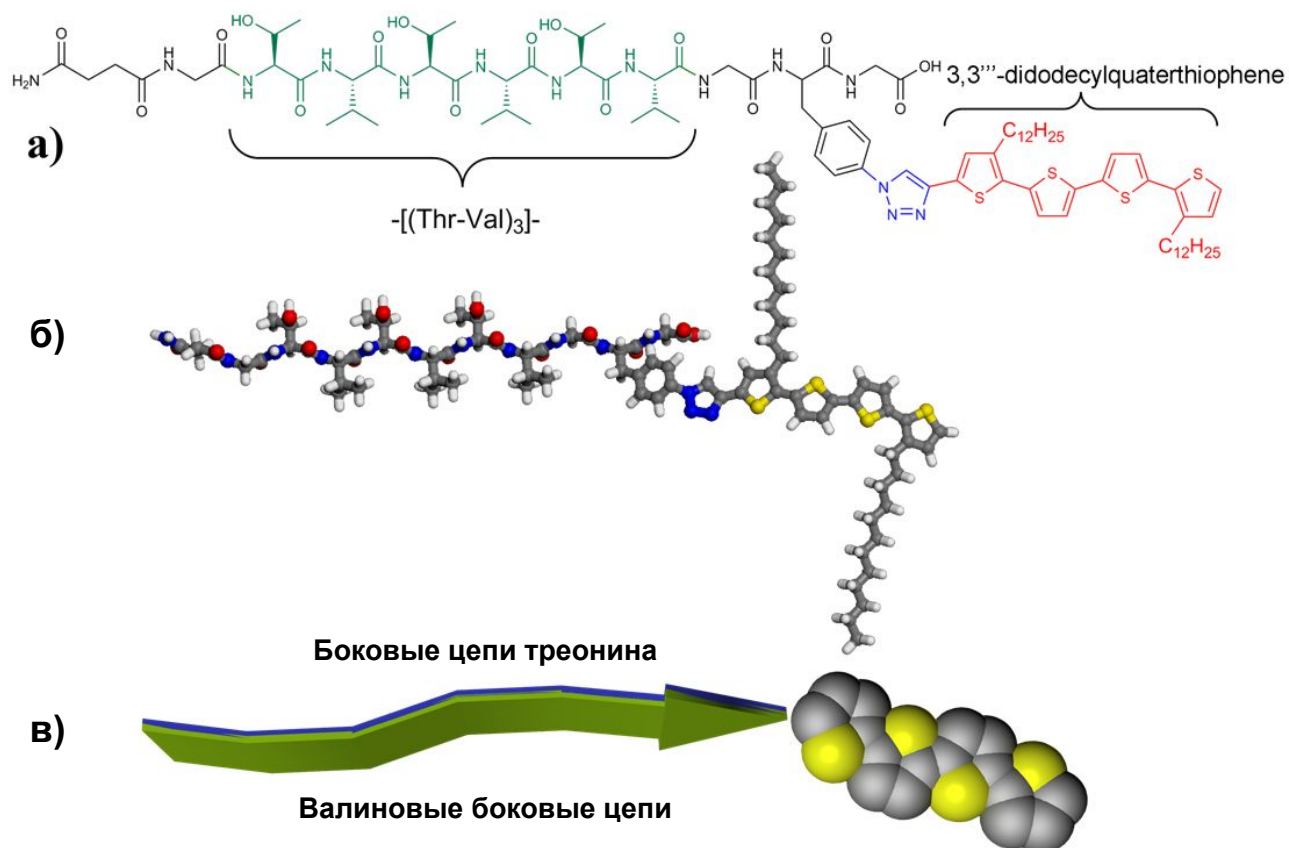


**Рисунок 5.9** — Исследуемое соединение тиофен-пептид.

### 5.2.2 Молекулярно-механическая модель

В качестве отправной точки для теоретического рассмотрения экспериментально наблюдаемых нановолокон была выбрана молекулярная структура гибрида олиготиофен-пептид, как показано на рисунке 5.10, и была построена

соответствующая полноатомная модель молекулярной механики. Модельное соединение отличается от экспериментального отсутствием гибкого ПЭГ-хвоста, поскольку ПЭГ-хвост не обладает способностью к определенным межмолекулярным взаимодействиям и первоначально был добавлен в основном по причинам растворимости. Следовательно, на нашем уровне методологии моделирования цепи ПЭГ вряд ли будут играть важную роль, определяющую структуру.



**Рисунок 5.10** — Различные представления молекулы Т-Р, используемые для компьютерного моделирования самособирающихся фибриллярных агрегатов. а) структурная формула, б) представление всех атомов с конформацией основной цепи, соответствующей конформации  $\beta$ -нити, в) схематическое представление  $\beta$ -цепи и кватертиофенового фрагмента.

Полноатомная молекулярная модель исследуемого соединения была основана на силовом поле РСFF класса II, параметризованного для воспроизведения *ab initio* энергетических поверхностей [69; 626; 627].

### 5.2.3 Построение и анализ периодических укладок

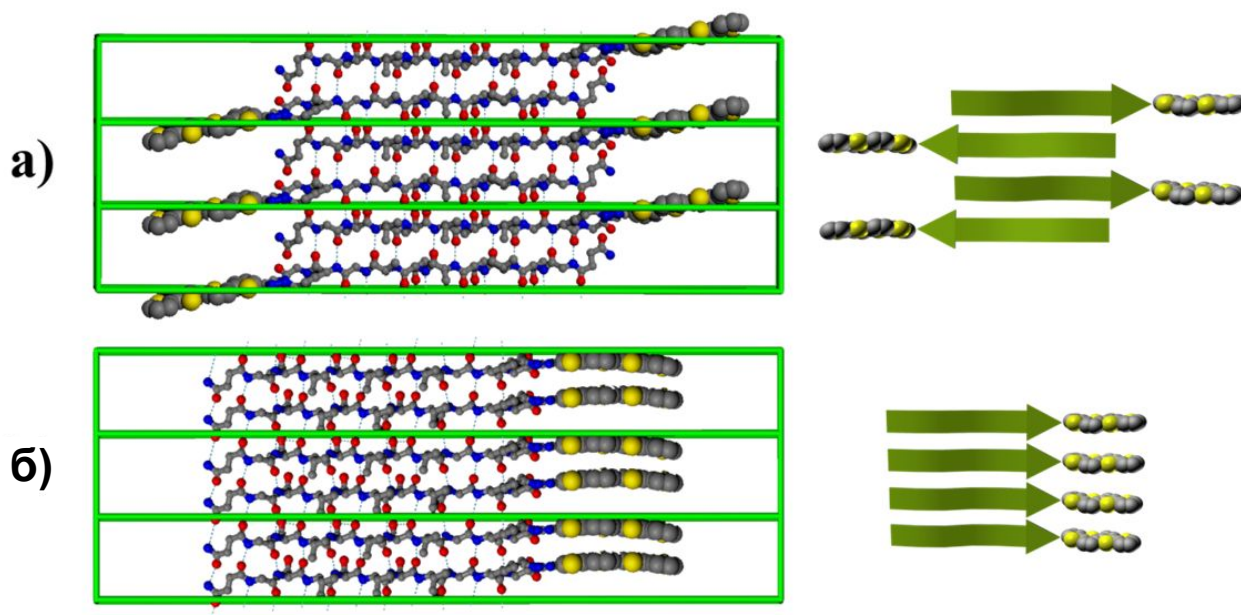
В принципе, обе части гибридной молекулы (конъюгированная часть и пептидная часть) способны к сильным межмолекулярным взаимодействиям, что приводит к образованию сильно анизотропных структур, таких как нановолокна, наностержни и т. д. Поскольку образование волокон в наномасштабе (особенно амилоидоподобных волокон) из  $\beta$ -листных пептидов происходит по известному механизму самосборки [628], подобные пептид-пептидные взаимодействия могут играть определяющую структурную роль для образования супраструктуры нашего гибридного соединения.

Большинство экспериментальных данных, доступных в литературе (включая разрешенные с помощью рентгеновских лучей и ЯМР-спектроскопии структуры амилоидоподобных фибрилл), предполагают, что в биологических амилоидных волокнах основным мотивом структурной агрегации пептидов является соединение длинных  $\beta$ -нитей в ленты с помощью латерально взаимодействующих одиночных пептидных цепей. Эти ленты могут, в свою очередь, накладываться друг на друга лицом к лицу, образуя фибриллы (филаменты) с так называемой четвертичной структурой с перекрестными  $\beta$ -ядром, которая считается основной характерной чертой амилоидных фибрилл. Стопки лент могут в дальнейшем взаимодействовать друг с другом и образовывать более толстые нитевидные агрегаты [588]. Однако, если взаимодействие между лентами затруднено, ленты могут скручиваться в спиральные структуры, как это было предложено Aggeli et al. (см. рис. 5.4) на основе их теоретических моделей, подтвержденных экспериментальными наблюдениями [588], а также обнаруженных в конъюгатах ПЭГ-пептид или поли(бутилакрилат)-пептид или поли(бутадиен)-пептид [607; 629]. Вопрос о предпочтении параллельного или антипараллельного расположения  $\beta$ -листов в таких фибриллярных структурах является спорным. Хотя было обнаружено, что многие экспериментально исследованные амилоидоподобные фибриллы имеют параллельное расположение  $\beta$ -слоев, различные амилоидные фибриллы, которые могут быть удивительно похожими по супрамолекулярной морфологии, могут существенно различаться в молекулярных деталях, включая параллельное или антипараллельное расположение  $\beta$ -слоев [630].

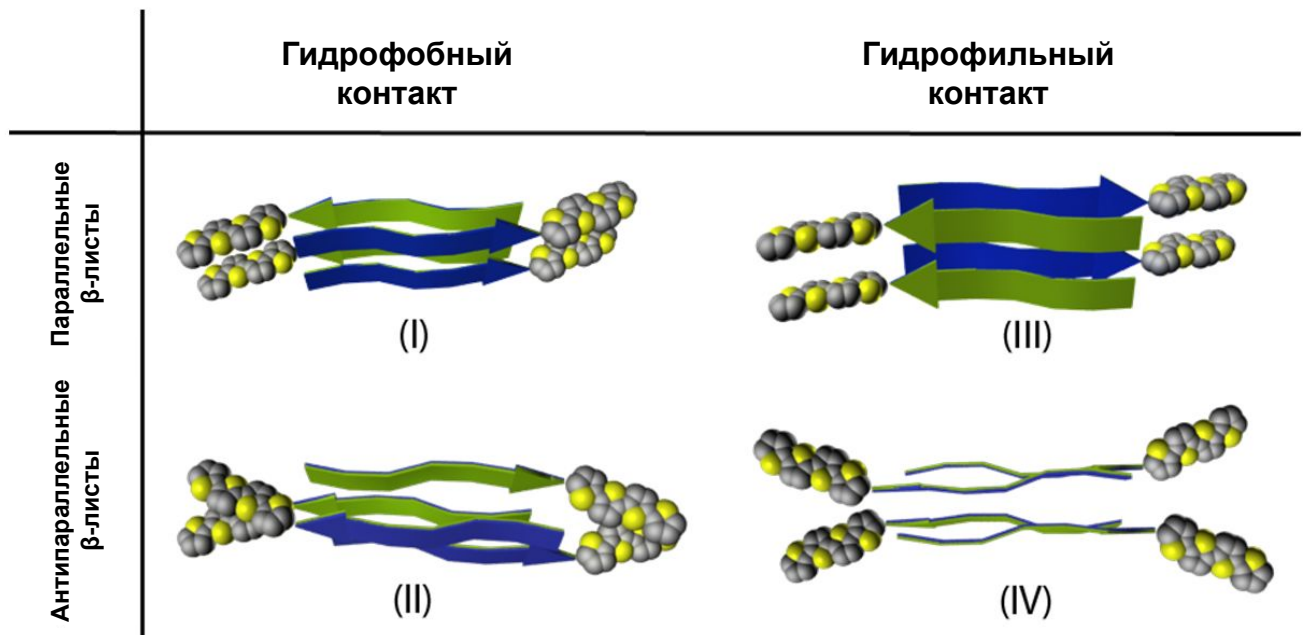
Исходя из вышеупомянутых идей, связанных со структурной организацией пептидных молекул в нитевидные агрегаты, мы предлагаем несколько периодических расположений для нашего гибридного соединения, исходя из предположения, что пептидная часть соединения организована аналогичным образом. Два предложенных основных периодических механизма включают параллельную или антипараллельную организацию пептидных фрагментов в  $\beta$ -листах. Еще четыре компоновки были получены из этих однослойных компоновок путем наложения  $\beta$ -листов лицом к лицу аналогичным структуре некоторых амилоидных фибрилл. Две основные однослойные периодические структуры, основанные на параллельных и антипараллельных  $\beta$ -листах, были построены с использованием комбинации молекулярного выравнивания и моделирования молекулярной динамики и показаны на рисунке 5.11. Полученные однослойные периодические структуры затем использовались в качестве блоков для образования различных двухслойных конструкций. Мы сконструировали эти двухслойные структуры из уравновешенных однослойных димеров, стремясь имитировать расположение пептидной части, наблюдаемое в амилоидоподобных фибриллярных структурах, разрешенных с помощью дифракции рентгеновских лучей или ЯМР-спектроскопии [551; 559] с последующей обширной релаксацией. Релаксация методом МД позволяет молекулам регулировать свое периодическое расположение до наилучшего локального минимума свободной энергии.

Двухслойные структуры были созданы следующим образом: однослойная структура была воспроизведена, повернута на  $180^\circ$  вокруг оси ленты и затем отрегулирована в латеральной плоскости так, чтобы пептидные сегменты двух однослойных фибрилл были приблизительно равны в регистре, образуя центр гибридной фибриллы.  $\beta$ -листы внутри такой фибриллы могут взаимодействовать либо с боковыми цепями валина, либо с боковыми цепями треонина каждого листа, обращенными друг к другу. Любое перекрытие молекулярных фрагментов, которое имело место (например, внутри алкильных цепей), разрешалось небольшими корректировками задействованных торсионных углов, которые в любом случае могли бы получить их равновесные значения во время цикла релаксации. Обсуждаемая процедура привела к четырем различным двухслойным периодическим системам, представленным на рисунке 5.12. В принципе для параллельного расположения-нитей можно подумать о двух других двухленточных конформациях с пониженной симметрией, когда  $\beta$ -цепи в

обеих лентах идут параллельно друг другу (см. рисунок 5.2). Однако такие структуры считались маловероятными, поскольку (i) такое структурное расположение  $\beta$ -нитей не наблюдалось экспериментально до настоящего времени для биологических амилоидных волокон, (ii) такому расположению не будет способствовать взаимодействие дипольных моментов соседних лент, и (iii) только одно такое расположение допустимо стерически из-за громоздкости тиофеновых фрагментов гибридной молекулы.



**Рисунок 5.11** — Построенные однослойные периодические ячейки укладки Т-Р соединений. Построенные периодические кристаллические ячейки для а) антипараллельного и б) параллельного расположения пептидных цепей в однослойных фибриллах. Алкильные цепи опущены для ясности; пунктирными линиями показаны водородные связи, ответственные за образование  $\beta$ -листа. На вставках справа показано принципиальное расположение пептидных и тиофеновых фрагментов.



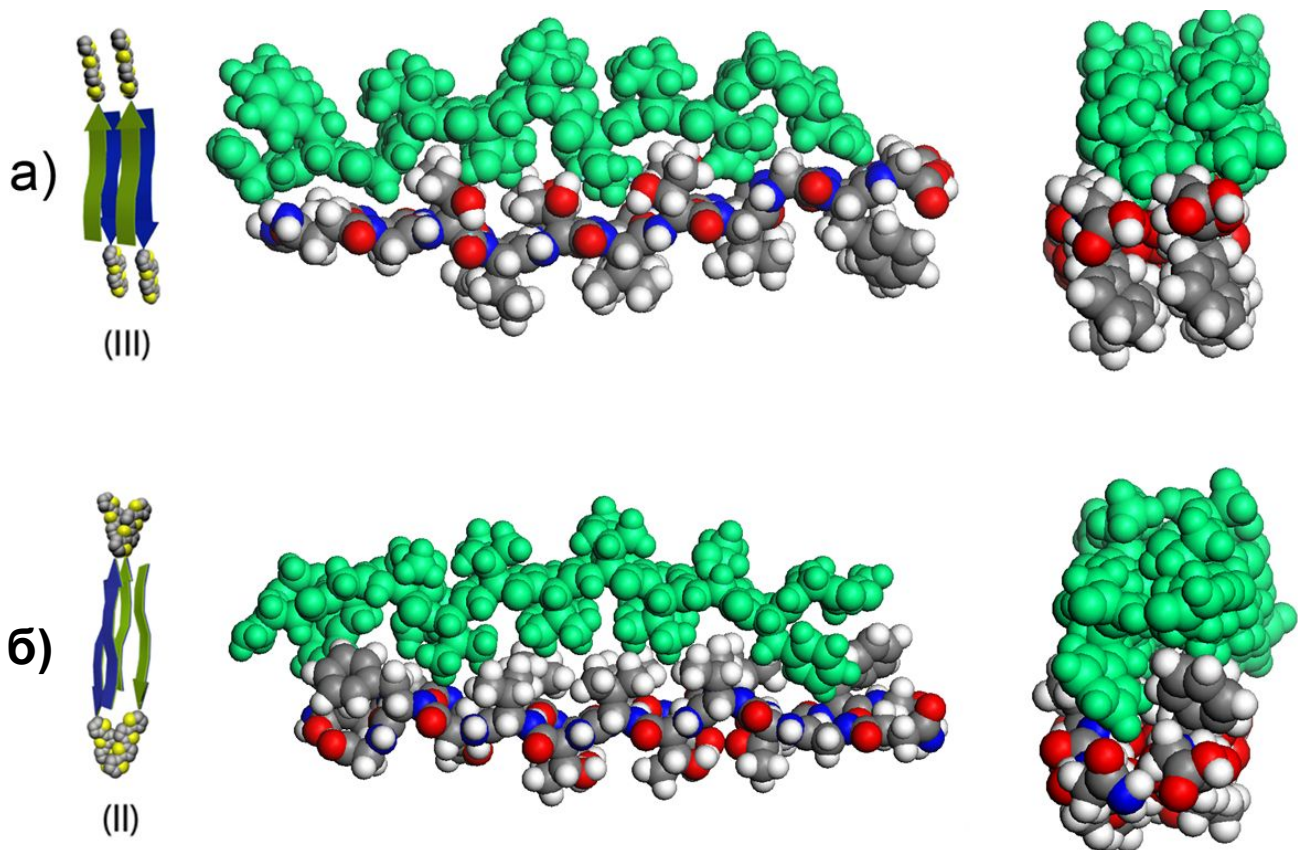
**Рисунок 5.12** — Построенные двухслойные периодические ячейки укладки Т-Р соединений. Схематические изображения сконструированных двухслойных периодических структур из гибридной молекулы, классифицированные по ориентации  $\beta$ -листов (параллельная или антипараллельная) и типу межслоевого контакта  $\beta$ -листов (гидрофобный: валин-валин, гидрофильный: треонин-треонин).

Поскольку двойные слои представляют собой значительно более сложные конформационные сборки, особенно в отношении организации поверхностей, скрытых между двумя  $\beta$ -слоями, и взаимодействия боковых цепей на этих поверхностях, для этих периодических структур был проведен цикл релаксации молекулярной динамики в 10 нс. Во время этого прогона имела место определенная перестройка конформаций боковой цепи между  $\beta$ -листами.

Относительные энтальпии образования для всех структур во время эксперимента отслеживались: самая низкая энтальпия образования наблюдалась для структуры (III) (т.е. наиболее энергетически выгодной структуры). По отношению к структуре (III) структуры (I), (II) и (IV) были менее энергетически выгодными и имели более высокие значения энтальпий образования на 11, 6 и 12 ккал/моль на молекулу соответственно (статистическая ошибка: 1-2 ккал/моль). Хотя этот набор данных был получен при моделировании в вакууме и влияние растворителя не учитывалось, он дает ценные количественные данные для понимания иерархии взаимодействий в таких системах и согласуется с гипотезой о контактах между гидрофильными слоями и антипараллельном

расположении листов – факторов, которые приводят к увеличению энтальпии образования. Однако в нашем случае из-за особой геометрии молекул картина агрегации, основанная на антипараллельных  $\beta$ -слоях и гидрофильном межслоевом контакте (см. Рисунок 5.12, система (IV)), приводит к потере плотной упаковки между тиофеновыми фрагментами и, следовательно, становится энергетически невыгодной. Благодаря специфичности присоединения тиофеновой составляющей к пептиду более плотная упаковка может быть достигнута только за счет конформационных изменений пептидного остова, что, однако, приводит к менее эффективным взаимодействиям между пептидными цепями  $\beta$ -листа.

Известно, что стабильность двухслойной структуры в амилоидных фибриллах часто приписывается “стерической застежке-молнии”, которая образуется на границе раздела  $\beta$ -листов. Межслойное расположение боковых цепей для систем (II) и (III) показано на рисунке 5.13. Можно видеть, что боковые цепи противоположных  $\beta$ -листов находятся в тесном стерическом контакте друг с другом.



**Рисунок 5.13** — Вид двух двухслойных структур, изображающих интеркаляцию боковых цепей аминокислот, образующую стерическую застежку-молнию.

Кадры делаются после выполнения уравнивания.



## Методы и детали

Построение однослойных периодических упаковок из одиночных молекул осуществлялось с использованием последующих шагов минимизации и релаксации следующим образом.

Сначала были скорректированы конформации исходных молекул: считалось, что пептидный блок имеет форму-цепи, образующей идеальный параллельный или идеальный антипараллельный  $\beta$ -лист, который определяется исключительно значениями двугранных  $\varphi$  и  $\psi$  углов пептидного остова, которые, как известно, равны  $\varphi = -119^\circ, \psi = -113^\circ$  для пептидных цепей параллельного  $\beta$ -листа и  $\varphi = -139^\circ, \psi = -135^\circ$  для  $\beta$ -цепей пептида в антипараллельном  $\beta$ -листе [631]. Считается, что тиофеновый блок, включающий боковую цепь 4-азидофенилаланина и алкильные цепи, находится в плоской протяженной конформации, которая соответствует локальному минимуму энергии, как показано на рисунке 5.11б. Чтобы получить правильное периодическое расположение, которое можно было бы использовать в качестве элементарной структуры для построения длинных фибрилл, была получена соответствующая периодическая элементарная ячейка, состоящая из двух (для однослойных фибрилл) или четырех (для двухслойных фибрилл) молекул. Период, предполагаемый элементарной ячейкой вдоль оси фибриллы, был установлен равным  $4,8\text{\AA}$  на  $\beta$ -цепь, что соответствует периоду, обычно наблюдаемому в амилоидных фибриллах ( $9,6\text{\AA}$  для двух нитей) [556]. На первом этапе две молекулы были организованы в две однослойные периодические структуры с параллельным и антипараллельным расположением  $\beta$ -цепей, пептидные сегменты в обоих случаях были выровнены так, чтобы максимизировать количество межцепочечных водородных связей (см. Рисунок 5.11). Затем была проведена минимизация энергии для таких периодических однослойных систем (которые можно рассматривать как бесконечно длинные кристаллические ленты) с последующим прогоном релаксационного МД-моделирования длительностью 1 нс.

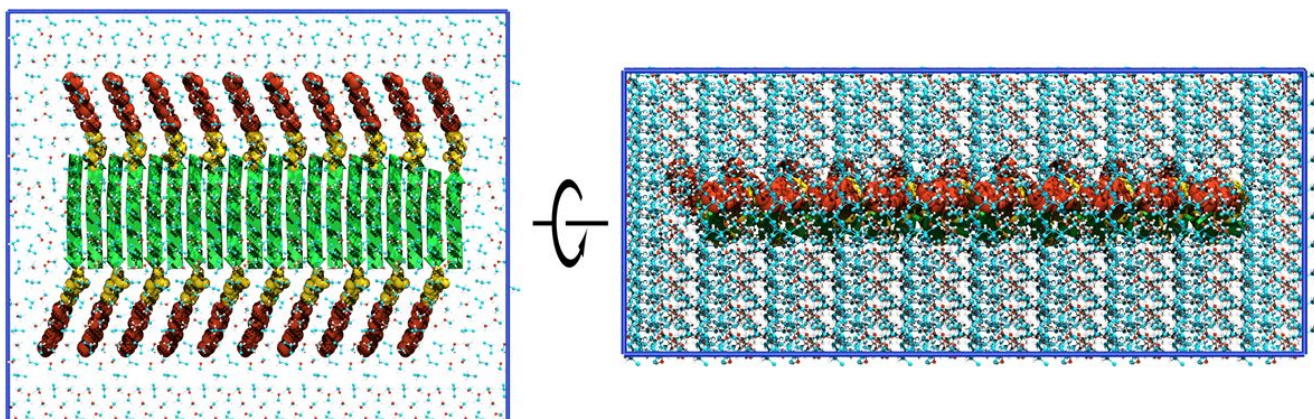
Все моделирование проводилось с использованием пакета моделирования LAMMPS [632]. Периодические системы сначала подвергались минимизации энергии с использованием версии Полака-Рибьера алгоритма сопряженного градиента с теми же параметрами силового поля (описанными далее), которые использовались во время моделирования МД. В МД-моделировании

периодических систем использовался радиус отсечения для ван-дер-ваальсовых взаимодействий 1,0 нм, кулоновские взаимодействия были учтены с использованием метода PPPM [633] с отсечкой 1,0 нм в реальном пространстве и параметром точности  $10^{-4}$ . Шаг интегрирования составлял 1 фс. Ансамбль NVT с  $T = 300K$  поддерживался добавлением трения и стохастических членов к уравнениям движения согласно уравнению Ланжевена, использовалась обратная постоянная трения 1 пс.

#### 5.2.4 Моделирование фибриллярных агрегатов в растворителе

Как полученные однослойные периодические структуры, так и две из двухслойных структур были использованы для создания коротких фибрилл путем репликации периодической ячейки вдоль оси филамента. Среди двухслойных структур компоновки (II) и (III) были выбраны для дальнейшего изучения, поскольку они (i) представляют собой компоновки с самыми низкими энтальпиями образования среди других и (ii) представляют два различных основных типа расположения *beta*-нитей - параллельную или антипараллельную упаковку, что должно позволить провести интересный сравнительный анализ влияния моделей параллельного или антипараллельного расположения на морфологию и эволюцию волокна. Для выбранных систем были сконструированы плоские прямые фибриллы длиной 20  $\beta$ -нитей (приблизительно 40 нм), фибриллы помещались в ячейку с растворителем размером  $5 \times 12 \times 10$  нм<sup>3</sup>. Сам растворитель представлял собой смесь 50% на 50% метанола и дихлорметана (по объему), соответствующую той, которая использовалась в эксперименте. Общее количество атомов в системе колебалось в районе 40 тысяч. Эти системы были подвергнуты моделированию МД в течение 10 нс при  $T = 300K$ , чтобы исследовать их стабильность и проанализировать конформационную релаксацию и поведение.

Визуальное представление одной из исследуемых систем представлено на рисунке 5.14.



**Рисунок 5.14** — Первоначальный снимок сольватированной короткой однослойной фибриллы на основе антипараллельного расположения  $\beta$ -листов. Фибрилла визуализируется в комбинированном представлении: пептидные цепи изображены зелеными стрелками, кватротиофен - в изображении атомов, красным, а пара-азидофенилаланиновый линкер - желтым. Алкильные цепи не изображены. Молекулы растворителя (метанол и дихлорметан) показаны точками.

## Результаты

### Конформационная эволюция

Во время моделирования все фибриллы оказались стабильными молекулярными агрегатами, и никакого отрыва молекул от агрегата не наблюдалось, хотя начальная конформация коротких фибрилл изменилась, чтобы принять наиболее энергетически выгодную конформацию. Кадры исходной геометрии фибрилл, а также их снимки после 10 нс времени моделирования представлены на рисунке 5.15. Представленные четыре типа фибрилл будут далее обозначаться их аббревиатурами: SL-AP (однослойная фибрилла, основанная на антипараллельном расположении  $\beta$ -нитей, изображенная на рисунке 5.15а), AL-PAR (однослойная фибрилла, основанная на параллельном расположении  $\beta$ -нитей, изображенная на рисунке 5.15б), DL-AP (двухслойная фибрилла, основанная на антипараллельном расположении  $\beta$ -нитей, тип (II) в обозначениях рис. 5.12, изображенная на рис. 5.15в), DL-PAR (двухслойная фибрилла, основанная на параллельном расположении  $\beta$ -нитей, тип (III) в обозначениях рис. 5.12, изображенная на рис. 5.15г).

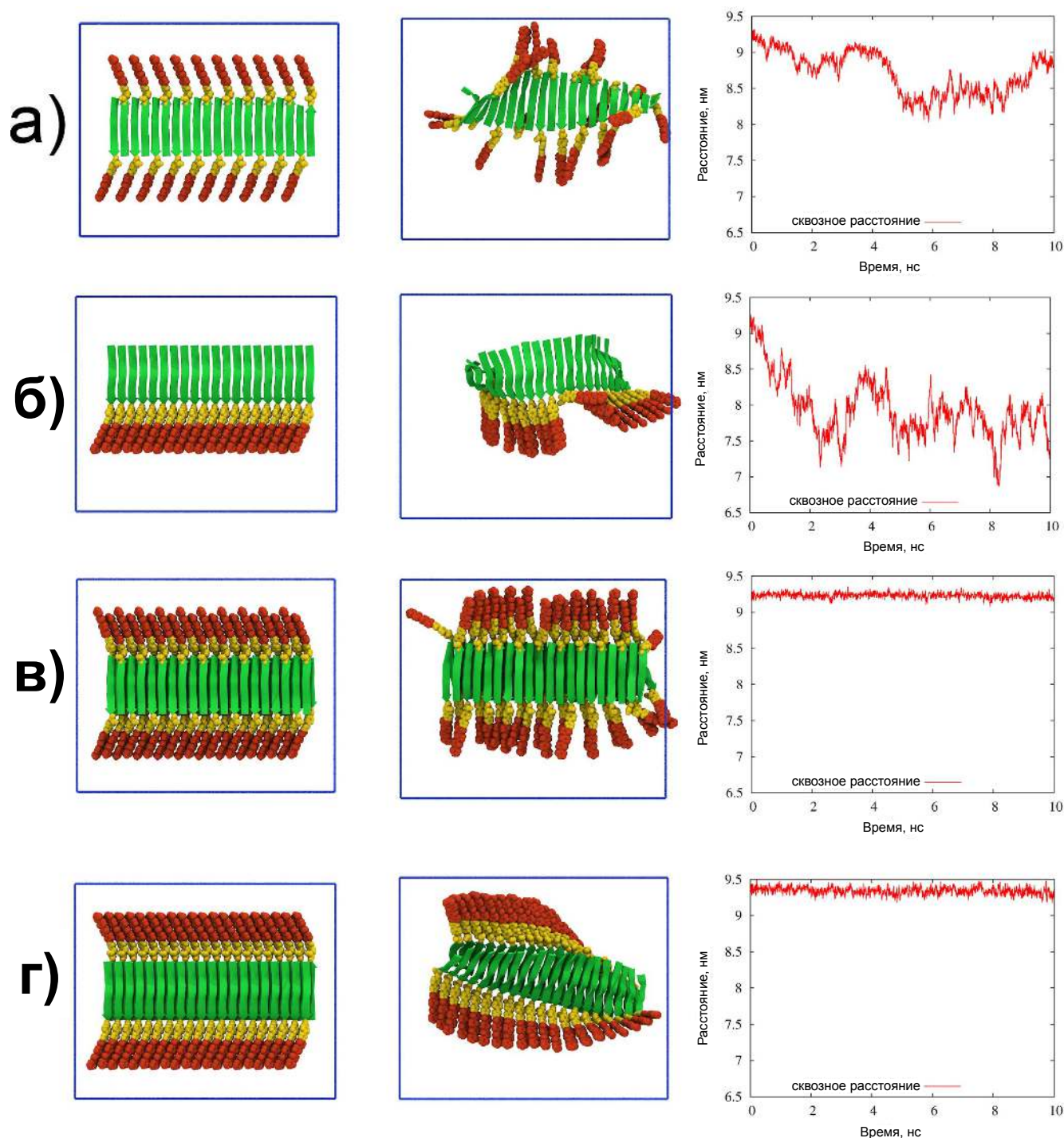
Однослойные фибриллы имеют тенденцию быть более конформационно гибкими, чем двухслойные фибриллы, как видно из графиков расстояния от конца до конца на рисунке 5.15, где расстояние между центрами масс молекул, находящихся на двух концах фибрилл, изображена как функция времени.

Фибрилла SL-AP после моделирования продемонстрировала левостороннее осевое скручивание пептидного остова с некоторыми намеками на скручивание фибриллы таким образом, чтобы гидрофильная сторона ленты (с открытыми остатками треонина) находилась на вогнутой стороне и гидрофобный (с открытыми боковыми цепями валина) на выпуклой стороне ленты (см. рис. 5.15а).

Для фибриллы SL-PAR также присутствует осевое левостороннее скручивание, которое сочетается со скручиванием фибриллы, которое имеет тенденцию принимать менее прямую конформацию, что также видно на графике расстояния от конца до конца (см. рис. 5.15б). Фибрилла SL-PAR также менее устойчива на концах, видно, что строгий узор водородных связей  $\beta$ -листа нарушен на одном конце. Топология завитка такая же, как и у фибриллы SL-AP (гидрофильная сторона ленты находится на вогнутой стороне).

Двухслойные фибриллы имеют тенденцию быть более жесткими и прямыми. В то время как фибрилла DL-AP почти плоская, фибрилла DL-PAR образует выраженный левосторонний осевой поворот на  $90^\circ$ , что соответствует примерно  $4,5^\circ$  на  $\beta$ -нить.

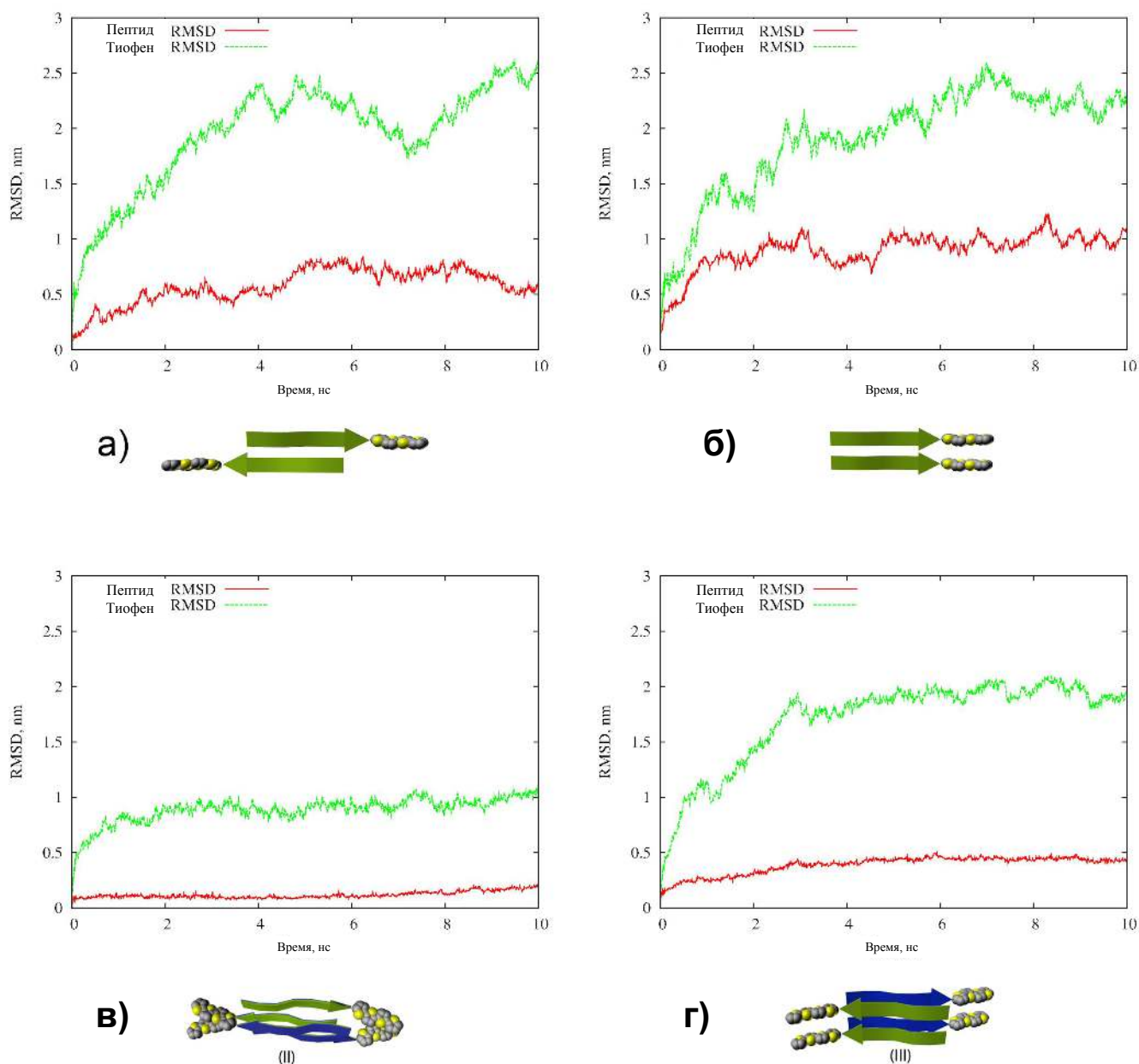
Целесообразно сравнить окончательные конфигурации фибрилл с моделями, предложенными Aggeli et al. (см. рисунок 5.4). В обозначениях рисунка 5.4 наблюдаемые фибриллы можно разделить на следующие типы: SL-AP и SL-AP - тип (в) - ленты, DL-AP и DL-PAR - тип (г) - двойные ленты.



**Рисунок 5.15** — Кадр из моделирования коротких фибрилл в растворителе после 10 нс моделирования. Фибриллы на основе одиночных  $\beta$ -листов: а) анти-параллельный-лист (SL-AP), б) параллельный  $\beta$ -лист (SL-PAR). Фибриллы на основе двухслойных  $\beta$ -листов: в) структура типа (II) на основе антипараллельного  $\beta$ -листа (DL-AP), г) структура типа (III) на основе параллельного  $\beta$ -листа (DL-PAR). Алкильные хвосты не изображены, тиофеновая часть изображена красными сферами, пара-азидофенилаланиновый линкер изображен желтым цветом, пептидные части изображены зелеными стрелками. В третьем столбце приведены графики расстояния между центрами масс крайних молекул.

Конформационная эволюция может быть дополнительно выявлена, глядя на графики среднеквадратичного отклонения (RMSD) пептидных и тиофеновых фрагментов в фибриллах на рисунке 5.16. RMSD рассчитывали относительно исходной конформации фибриллы по формуле  $RMSD = \min_{rot,trans} \sqrt{\frac{\sum_{i=1}^N |\vec{r}_i(t) - \vec{r}_i(0)|^2}{N}}$ , где N - число атомов,  $\vec{r}_i(t)$ - радиус-векторы атомов, предварительно вычисляется минимум функции по поворотам и трансляциям структуры.

Ясно видно, что флуктуации RMSD для пептидных фрагментов намного меньше для двухслойных структур, чем для однослойных. Начальный дрейф RMSD в течение первых 3-4 нс моделирования связан с конформационной перестройкой пептидных и тиофеновых фрагментов. Например, для фибриллы DL-PAR графики RMSD (рис. 5.16г) четко иллюстрируют скручивание фибриллы.



**Рисунок 5.16** — Среднеквадратичное отклонение структуры фибрилл относительно исходной структуры в зависимости от времени моделирования.

### Водородные связи

Далее мы проанализируем роль водородных связей в стабильности фибрилл. В таблице 7 и на рисунке 5.17 представлены данные по водородным связям для различных типов фибрилл. Известно, что водородные связи являются основным взаимодействием при формировании  $\beta$ -листа. Каждый остаток в  $\beta$ -листе обычно образует 2 водородные связи с остатками в следующей  $\beta$ -нити. Таким образом, максимальное количество водородных связей, которое может образовывать последовательность  $[Thr - Val]_3$ , которая, как считается, является основным звеном, стоящим за образованием  $\beta$ -листа, равно 12, или, если

рассчитывать на одну молекулу, - 6 (поскольку каждая Н- связь принадлежит 2 молекулам). Однако другие группы в гибридной молекуле, такие как боковые цепи треонина (ОН-группа) и другие линкерные аминокислоты (глицин, фенилаланин) и концевые группы, также могут образовывать Н-связи в нашей молекуле. Поскольку точное определение водородной связи может влиять на абсолютные значения водородных связей, мы лучше сосредоточимся на сравнительном анализе водородных связей в различных типах фибрилл.

Известный факт, что антипараллельные  $\beta$ -листы имеют тенденцию к образованию более прямых и прочных водородных связей [634], подтверждает вычисленное количество межфибрильных водородных связей для различных типов фибрилл. В то время как антипараллельные фибриллы имеют около 3,3-3,5 водородных связей на гибридную молекулу, параллельные - 2,1-2,5. Из графиков временной эволюции (рис. 5.17) видно, что для того, чтобы произошла конформационная релаксация структуры водородных связей, требуется приблизительно 2 нс.

Если сравнить фибриллы SL-AP и DL-AP, мы увидим, что для DL-AP наблюдается увеличение межфибрильных водородных связей (3,3 против 3,5), которое компенсируется уменьшением водородных связей фибрилла-растворитель (2,5 против 2,2). Это можно рассматривать как следствие большей регулярности упаковки пептидов в двухслойных структурах. Для фибрилл на основе параллельных  $\beta$ -слоев тенденция аналогична, но гораздо большее уменьшение водородных связей фибрилла-растворитель можно отнести к другому факту: все боковые цепи треонина, содержащие гидроксильную группу, указывают на одну и ту же сторону  $\beta$ -листа, в фибрилле SL-PAR эти ОН-группы, вероятно, образуют водородные связи с молекулами метанола. Но в фибрилле DL-PAR эти группы ограничены слоями двух соседних  $\beta$ -листов и имеют меньше возможностей для образования водородных связей.

Последнее наблюдение, по сути, важно для понимания энергетических проблем образования двухслойных фибрилл. Однослойный  $\beta$ -лист имеет в нашем случае гидрофобные и гидрофильные стороны, и, образуя двухслойные фибриллы, они могут склеиваться либо с гидрофобными, либо с гидрофильными сторонами, как показано на рисунке 5.12. Гидрофобные или гидрофильные взаимодействия фактически опосредуются растворителем и общее количество водородных связей, которые система имеет в той или иной конформации, является грубой мерой энергетической выгодности гидрофобной или гидрофильной



укладки  $\beta$ -листов. В обсуждаемом случае фибриллы DL-PAR кажется, что гидрофильная укладка  $\beta$ -листов, хотя и энергетически выгодна сама по себе, не является благоприятной в присутствии растворителя, поскольку гидроксильные группы остатков треонина могут образовывать больше водородных связей при наличии растворителя.

По общему количеству водородных связей фибрилла DL-AP превосходит все другие фибриллы и, таким образом, вероятно, энергетически более выгодна, чем другие типы фибрилл.

Тип фибриллы	Межфибрильные водородные связи на молекулу	Н-связей фибрилла-растворитель на молекулу	Всего Н-связей одна молекула составляет
SL-AP	3.3	2.5	9.1
SL-PAR	2.1	2.9	7.1
DL-AP	3.5	2.2	9.2
DL-PAR	2.5	1.4	6.4

**Таблица 7** — Среднее количество водородных связей на гибридную молекулу, усредненное за последние 2 нс моделирования. Последний столбец представляет собой общее среднее количество водородных связей, которые одна гибридная молекула создает с растворителем или другими гибридными молекулами (обратите внимание, что межфибрилярные водородные связи здесь рассчитываются дважды).

Графики зависимости количества водородных связей от времени (Рисунок 5.17) показывают величину флуктуаций водородных связей, а также изменение их количества по мере того, как фибриллы релаксируют из исходного состояния.

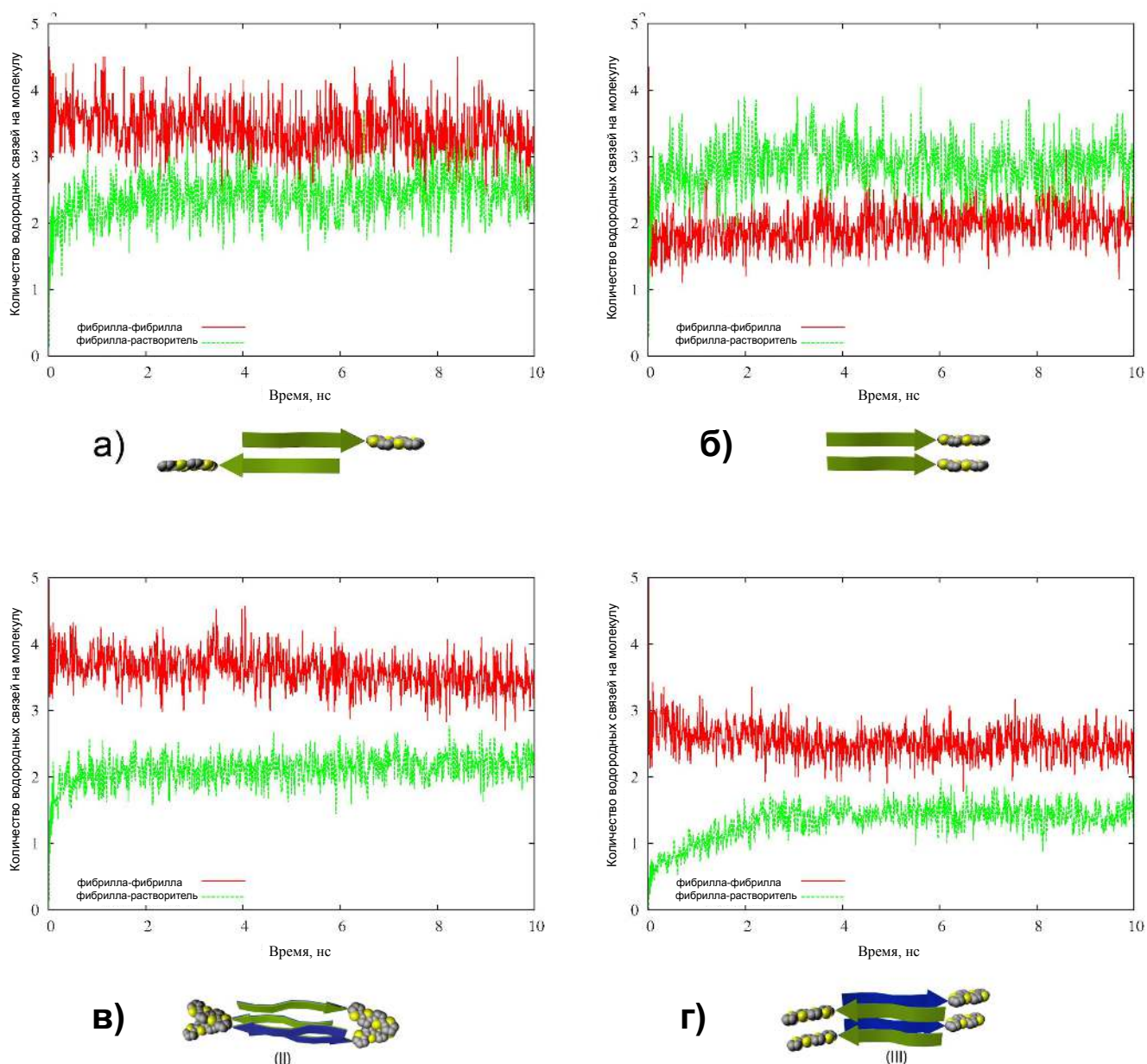


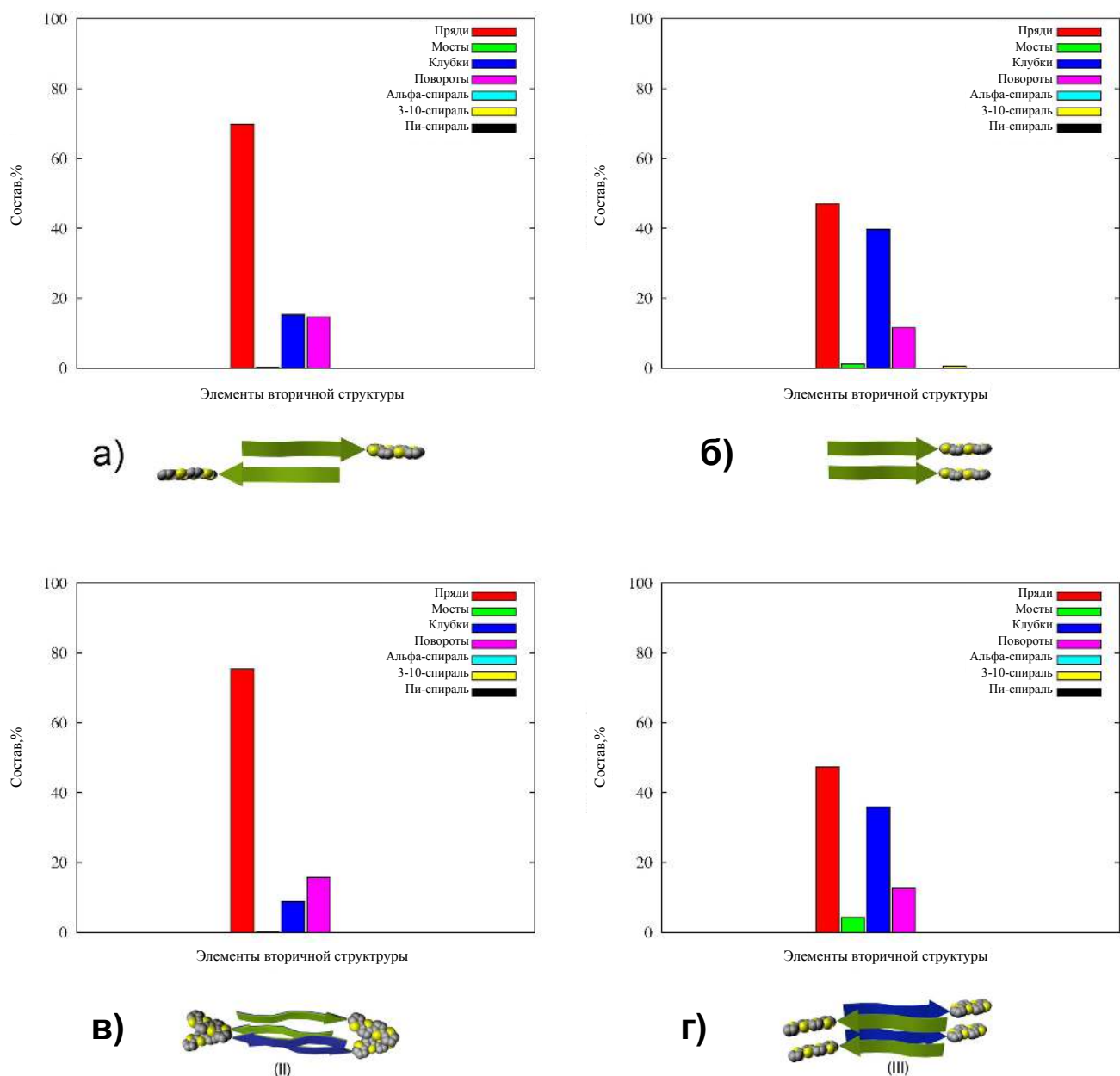
Рисунок 5.17 — Число водородных связей на одну гибридную молекулу как функция времени моделирования.

### Упаковка пептидов

Для дальнейшего анализа конформации пептидных фрагментов в фибриллах был проведен анализ вторичной структуры пептидного остова, представленного на рисунке 5.18, с использованием алгоритма STRIDE [635]. Выявлено высокое преобладание  $\beta$ -листовой структуры фибрилл на основе антипараллельного расположения (70% для SL-AP и 75% для DL-AP). Последовательность из пептидного хвоста гибридной молекулы, которая была проанализирована с помощью STRIDE, состоит из 11 аминокислот: GLY-[THR-VAL] 3-GLY-PHE-GLY и специально включенного  $\beta$ -листа, образующего

$[THR - VAL]_3$  последовательность составляет только 55% от общей последовательности. Таким образом, можно видеть, что для антипараллельного расположения склонность к  $\beta$ -слою выходит далеко за пределы последовательности  $[THR - VAL]_3$  и включает также соседние аминокислотные остатки.

Для параллельных расположений склонность формирования  $\beta$ -структуры составляла 45% (+ 1%  $\beta$ -мостов) для SL-PAR и 47% (+ 4%  $\beta$ -мостов) для фибрилл DL-PAR. Снова видно, что двухслойные укладки немного более упорядочены и имеют немного большую часть структуры в виде  $\beta$ -листов. Часть пептидной последовательности, которая организована в  $\beta$ -лист или  $\beta$ -мостики для параллельного расположения, значительно меньше, чем для антипараллельного, однако это соответствует общей менее точной структуре параллельных  $\beta$ -листов с точки зрения регулярности водородных связей. В любом случае для фибриллы DL-PAR совокупная склонность структуры к  $\beta$ -листам составляет 51%, что очень близко к доле последовательности  $[THR - VAL]_3$  в общей последовательности, и эту последовательность можно рассматривать как идеальный  $\beta$ -лист.



**Рисунок 5.18** — Гистограммы элементов вторичной структуры в различных фибриллах, смоделированных в растворителе. Данные усредняются за последние 2 нс моделирования. Фибриллы обозначены их пиктограммами под графиками: а) фибрилла SL-AP, б) фибрилла SL-PAR, в) фибрилла DL-AP, г) фибрилла DL-PAR.

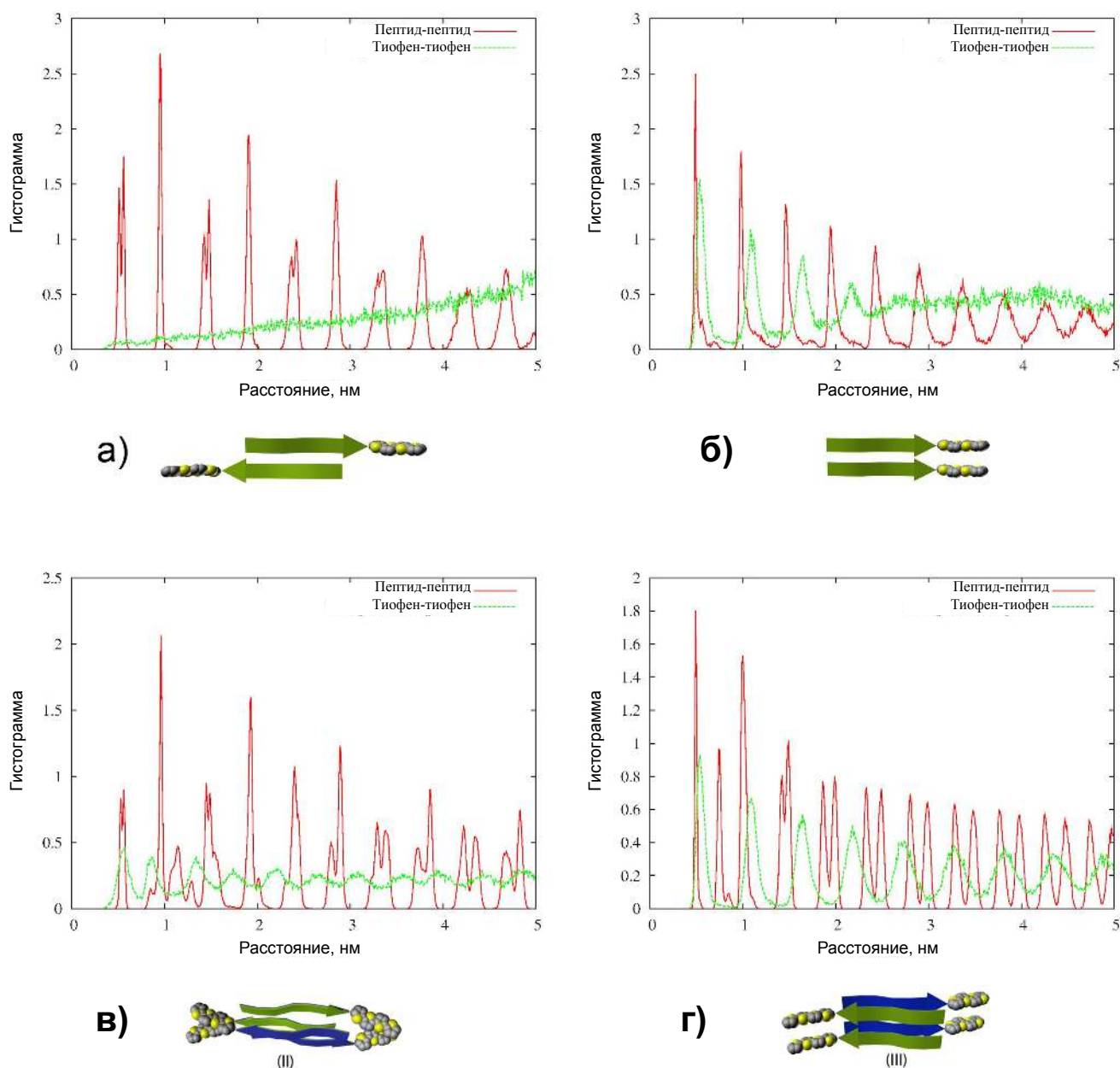
Анализ расположения пептида и тиофена также проводился с помощью гистограмм линейного распределения расстояний между центрами масс пептидных и тиофеновых фрагментов (см. рис. 5.19). Эти графики ясно указывают на то, что пептидные фрагменты надежно организованы с хорошим дальним порядком. Сравнение графиков снова показывает, что укладки с двумя слоями имеют лучший дальний порядок, чем укладки с одним слоем. Координаты

второго основного пика красных линий (распределения пептид-пептид) представляют периодичность структуры и читаются следующим образом: SL-AP - 0,95 нм, SL-PAR - 0,98 нм, DL-AP - 0,96 нм, DL-PAR - 1.0 нм, что при делении на 2 дает расстояние между нитями. Мы видим, что антипараллельное расположение имеет несколько более узкую упаковку  $\beta$ -нитей 4,8Å между пептидами по сравнению с 4,9-5,0Å при параллельном расположении.

Как видно из графиков, периодическое расположение пептидных фрагментов передается тиофеновым фрагментам в случае фибрилл SL-PAR, DL-PAR и DL-AP. Наилучший дальний порядок блоков кватротиофена достигается для фибриллы DL-PAR, тогда как в случае SL-AP упорядочение почти не наблюдается.

Пики гистограмм распределения тиофен-тиофена повторяют пики гистограмм пептид-пептид, но не в точности. Для параллельных расположений первый пик для тиофеновых фрагментов находится при 0,54 нм как для SL-PAR, так и для DL-PAR. Этот сдвиг межтиофенового расстояния по сравнению с межпептидным расстоянием можно объяснить только осевым скручиванием фибриллы. Хотя фибрилла SL-PAR является более динамичной, и ее поведение при осевом скручивании нельзя сразу увидеть при визуальном осмотре (из-за скручивания фибриллы), соответствие ее функции распределения функции DL-PAR четко указывает на то, что скручивание - это внутреннее свойство уже однослойной фибриллы.

Для фибриллы DL-AP первый и второй пики распределения тиофен-тиофен соответствуют 0,57 и 0,86 нм, что можно объяснить латеральным сдвигом тиофеновых фрагментов по отношению друг к другу, поскольку этот тип структуры менее регулярен и прост, чем тот, который основан на параллельном расположении.



**Рисунок 5.19** — Линейное распределение пептидных и тиофеновых блоков в фибриллах. Все расстояния были рассчитаны между центрами масс блоков  $[THR - VAL]_3$  или блоков кватротиофена за последние 2 нс траектории моделирования. Распределения нормированы на  $(10 - x)$ , где  $x$  - расстояние.

### *Упаковка тиофена*

Теперь мы можем приступить к анализу упорядочения тиофенов в наших фибриллах - вопрос, который имеет первостепенный интерес с точки зрения возможного применения фибрилл в области органической электроники.

Сначала приступим к качественному описанию. Увеличенные снимки упорядочения блоков кватротиофена представлены на рисунке 5.20.

*Качественное описание.*

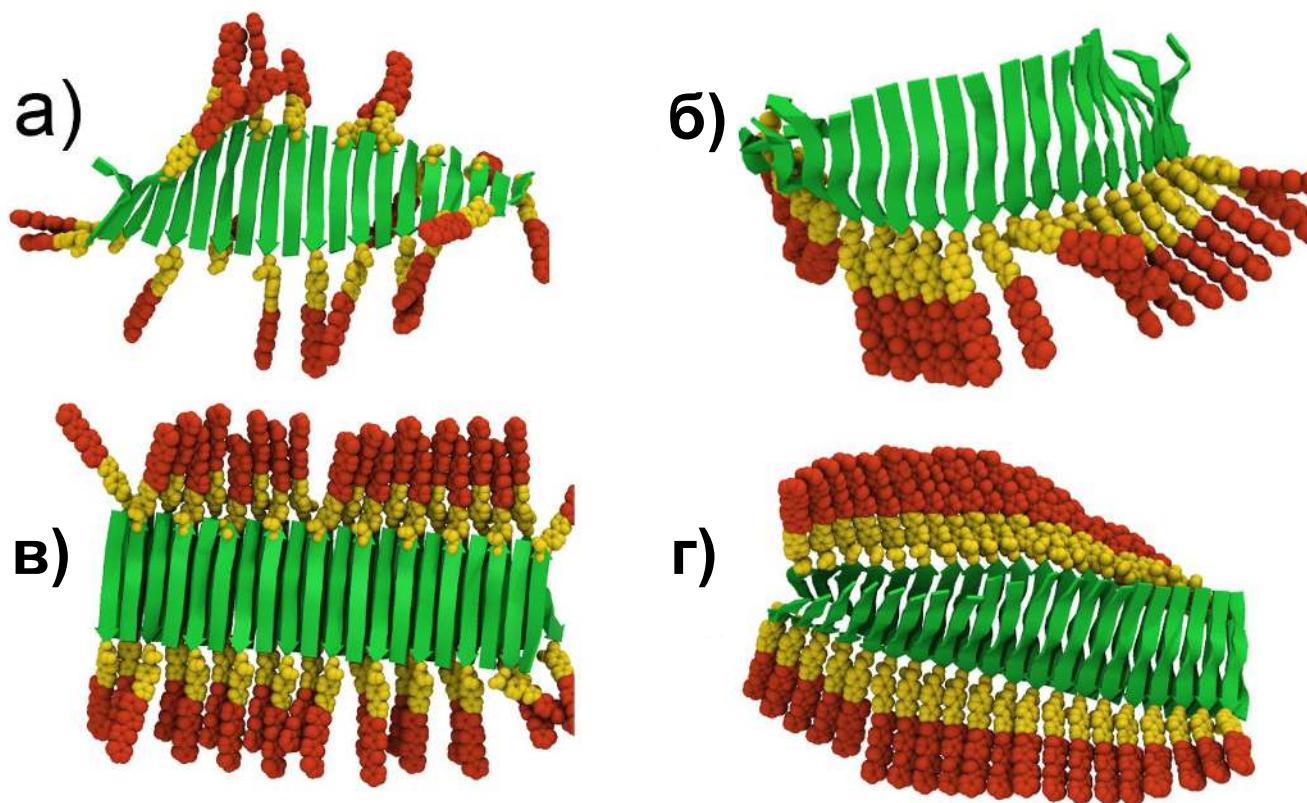
В фибрилле SL-AP (как видно во время моделирования и как видно из снимка на рис. 5.20а) блоки кватротиофена очень конформационно подвижны. Поскольку тиофены сами по себе являются жесткими, их конформационная подвижность частично обусловлена подвижностью боковой цепи пара-азидофенилаланинового линкера, через которую соединен тиофеновый блок, и частично (ii) вращением двугранного угла между фенилаланином и глицином, образуя своего рода глициновый шарнир. Последний тип конформационной подвижности может приводить к значительно различающимся ориентациям тиофеновой составляющей, как видно на рисунке 5.20а, некоторые тиофеновые блоки могут даже указывать в противоположном направлении по сравнению с другими блоками. Эта чрезвычайная гибкость также обеспечивается тем фактом, что в этом типе фибрилл тиофены расположены на расстоянии двух  $\beta$ -нитей ( $\sim 1$  нм) друг от друга, что фактически не оставляет возможности для взаимодействия и переноса электронов.

Структура тиофена в фибриллах SL-PAR (рис. 5.20б) более устойчива, поскольку тиофены здесь расположены ближе в пределах одинарного  $\beta$ -цепочечного расстояния ( $\sim 0,5$  нм) и имеют возможность стерически взаимодействовать. Это стерическое взаимодействие, однако, недостаточно сильное, чтобы удерживать вместе все тиофеновые фрагменты, они поворачиваются, чтобы сформировать области плотной упаковки, разделенные промежутками, и эта структура все еще динамична - промежутки имеют тенденцию закрываться и появляться снова. Появлению зазоров также способствует несоответствие расстояний упаковки: для тиофенов идеальное расстояние для укладки  $\pi - \pi$  составляет около  $3,3-3,5 \text{ \AA}$  [636; 637], в то время как расстояние между  $\beta$ -нитями составляет около  $5 \text{ \AA}$ , поэтому тиофены должны приспособливаться к расстоянию, обеспечиваемому  $\beta$ -листом, либо за счет образования промежутков, либо за счет наклона.

Двухслойные тиофены образуют более упорядоченную структуру; однако фибрилла DL-AP все еще не имеет степени упорядочения, необходимой для хорошего переноса электронов. Как видно на рисунке 5.20в, из-за несоответствия расстояния тиофены слипаются, но имеют промежутки между непрерывными стопками тиофенов (верхний ряд), или из-за гибкости линкерных аминокислот тиофены могут изменять свою ориентацию, вращаться так, чтобы их плоскости были выровнены вдоль фибрилл (см. нижний ряд), так что не образуется непрерывная тиофеновая стопка, способная переносить электроны вдоль фибриллы.

Кроме того, в этой фибрилле тиофены не идеально расположены латерально для стэкинг взаимодействия.

Фибрилла DL-PAR, по-видимому, содержит тиофены в лучшем порядке среди изученных, тиофены расположены в однородном порядке с некоторым изгибом и наклоном вдоль оси фибриллы (см. Рисунок 5.20г). Тиофены находятся в динамическом порядке - могут образовываться щели между соседними тиофенами в стопке, но они “заживают” за десятки пикосекунд.



**Рисунок 5.20** — Снимки расположения блоков кватротиофена в моделируемых фибриллах различных типов: а) фибрилла SL-AP, б) фибрилла SL-PAR, в) фибрилла DL-AP, г) фибрилла DL-PAR. Алкильные цепи не изображены для ясности представления.

#### *Количественное описание*

Качественно описав упорядочение тиофена в различных типах фибрилл, мы теперь количественно опишем положение соседних тиофеновых фрагментов с точки зрения их порядка, расстояния и ориентации. Для этого мы определяем три основных вектора, связанных с каждым блоком кватротиофена, вектор  $T1-T4$ , вектор нормали к поверхности и третий дополнительный вектор (см. Рисунок 5.21а). Вектор  $T1-T4$  - это вектор, соединяющий центры масс первого



и последнего тиофенового кольца в кватротиофене, вектор нормали к поверхности можно рассматривать как вектор, представляющий среднюю плоскость фрагмента кватротиофена, вычисленную как среднее из векторов нормалей каждого тиофенового кольца. Дополнительный вектор перпендикулярен двум векторам. Эти три вектора позже будут обозначаться как  $x$ ,  $y$  и  $z$ , как показано на рисунке 5.21а.

С точки зрения этих векторов, их упорядочения и положения соседних тиофеновых фрагментов по отношению к этим векторам было проанализировано упорядочение блоков кватротиофена, данные представлены в Таблице 8 и рисунке 5.21.

Тип фибриллы	Наиболее вероятное расстояние, Å	Параметр порядка T1-T4	Параметр порядка относительно нормали к поверхности	Среднее смещение $x$ , Å	Среднее смещение $y$ , Å	Среднее смещение $z$ , Å
SL-PAR	5.4	0.97	0.94	1.4	3.7	3.46
DL-AP	5.7	0.82	0.96	2.9	2.4	3.47
DL-PAR	5.4	0.98	0.95	2.0	3.3	3.55

**Таблица 8** — Параметры порядка соседних блоков кватротиофена в фибриллах, рассчитанные в течение последних 2 нс моделирования.

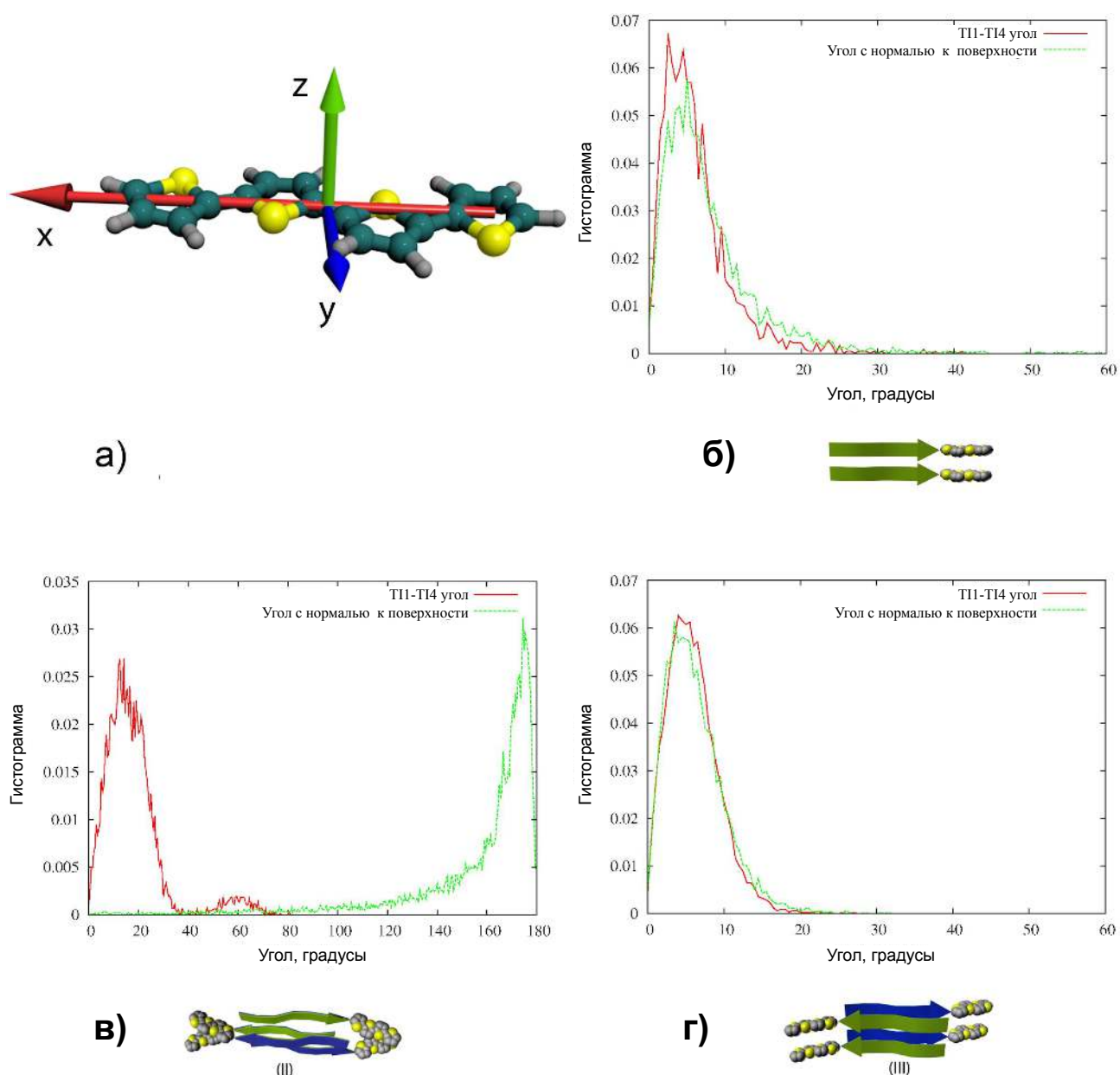
Таблица 8 включает информацию о наиболее вероятном расстоянии между центрами масс соседних блоков кватротиофена, параметрах порядка (рассчитанных как  $\langle P_2(\cos \vartheta) \rangle$ , где  $P_2$  - второй полином Лежандра) для двух векторов и проекциях среднего расстояния между центрами масс тиофена на главные векторы (называемое средним смещением). Параметры среднего смещения позволяют охарактеризовать, как следующий кватротиофен в стопке расположен относительно предыдущего.

Из таблицы можно увидеть, что укладки, основанные на параллельных  $\beta$ -листах, имеют более высокие параметры упорядочения тиофенов по сравнению с таковыми для фибриллы DL-AP для длинной оси тиофена (вектор T1-T4) (0,98 против 0,82). Однако параметр порядка для вектора нормали к поверхности, который описывает планарность соседних тиофенов, находится в

сопоставимом масштабе (фактически даже лучше) для фибриллы DL-AP, чем для фибриллы DL-PAR.

Гистограммы на рисунке 5.21 дают дополнительную информацию об этом порядке тиофеновых фрагментов в фибриллах. Для SL-PAR и DL-AP угловые распределения для обоих векторов представляют собой узкие гауссовы пики в диапазоне от  $0^\circ$  до  $10^\circ$ , тогда как для DL-AP ситуация иная. Для вектора T1-T4 максимум распределения составляет около  $15^\circ$  и охватывает более широкую область от  $0^\circ$  до  $35^\circ$ . Это указывает на то, что в направлении тиофенов все еще существует некоторый порядок, но они, как правило, не коллинеарны, а скорее имеют некоторый угол между собой. Для вектора нормали к поверхности ситуация немного иная: из-за структуры расположения смежные блоки кватротиофена имеют тенденцию быть антипараллельными по отношению к нормали к поверхности. Хотя распределение вектора нормали к поверхности в DL-AP имеет очень выраженный пик около  $170^\circ$ , что указывает на высокую компланарность тиофенов в этом типе фибрилл.

Параметры смещения в таблице позволяют понять типичную конфигурацию тиофеновых фрагментов. Из последнего столбца видно, что среднее расстояние между плоскостями тиофенов составляет около  $3,45\text{--}3,55\text{\AA}$ , что является верхним пределом идеального расстояния  $\pi\text{--}\pi$ -стэкинга в  $3,3\text{--}3,5\text{\AA}$  [636; 637]. Остальные значения дают ключ к пониманию кадров на рисунке 5.20.



**Рисунок 5.21** — Анализ упорядочения тиофенов в фибриллах. а) Графическое представление основных векторов в блоке кватротиофена, используемое для дальнейшего анализа упорядочения: красным - вектор, соединяющий центры масс первого и последнего тиофенового кольца (вектор Т1-Т4), зеленым - вектор, перпендикулярный среднему Плоскость поверхности тиофенового блока, синий - вектор, перпендикулярный двум другим векторам. б), в), г) гистограммы угловых распределений между Т1-Т4 и векторами нормали к поверхности соседних блоков кватротиофена для фибрилл SL-PAR, DL-AP и DL-PAR соответственно.

## Методы и детали

### *Методы моделирования*

Моделирование фибрилл в растворителе проводилось в пакете моделирования LAMMPS [632]. Моделирование проводилось со следующими параметрами: взаимодействие Ленарда-Джонса, радиус обрезания  $10\text{\AA}$ , электростатические взаимодействия обрабатывались с помощью алгоритма PPPM [633] с реальным пространственным отсечением  $10\text{\AA}$  и значением точности  $10^{-4}$ , динамика Ланжевена с  $T = 300\text{K}$  и параметром демпфирования  $0,5^{-1}$ , был применен алгоритм поддержания изотропного давления Берендсена с эталонным давлением 1 атм, параметром времени демпфирования 2 пс и объемным модулем сжатия системы в 20000 атм, что находится в диапазоне модуля объемной упругости типичных жидкостей. Шаг интегрирования для первых 100 пс моделирования составлял 0,1 фс, чтобы обеспечить плавную релаксацию для системы, в то время как для остальной части моделирования он составлял 1 фс.

Моделирование проводилось параллельно с использованием техники декомпозиции доменов на 512 процессорах одновременно.

### *Методы анализа*

Визуализация была выполнена с помощью программы VMD [305] и трассировщика лучей Tachyon [638].

Водородные связи определяли с использованием модуля VMD. Считалось, что водородная связь образуется между атомом с присоединенным к нему водородом (донор, D) и другим атомом (акцептор, A) при условии, что расстояние DA меньше, чем расстояние отсечки  $3,0\text{\AA}$  и угол DHA меньше предельного угла 20 градусов.

Для определения вторичной структуры использовался алгоритм STRIDE [635].

Для анализа тиофена были рассмотрены тиофены с расстояниями между ними менее  $6\text{\AA}$ . Основные векторы были рассчитаны следующим образом: вектор T1-T4 (x-вектор) - это вектор, соединяющий центры масс первого и последнего тиофенового кольца в кватротиофеновом блоке. В качестве предварительного шага для вычисления y-вектора были вычислены векторы, связывающие центр масс атомов углерода с атомом серы для каждого тиофенового кольца. Комбинация этих четырех векторов с чередующимися знаками

определила  $u$ -вектор. Затем вычисляли вектор нормали к поверхности ( $z$ -вектор) как векторное произведение  $x$ -вектора и  $u$ -вектора. Анализ проводился на основе последних 2 нс моделирования.

### 5.2.5 Крупномасштабное моделирование морфологии агрегатов в вакууме

Наша следующая цель в многомасштабном моделировании - изучить крупномасштабную конформационную морфологию фибрилл на основе выбранных молекулярных расположений в масштабе длины, сравнимом с теми, которые могут быть зафиксированы измерениями АСМ и ПЭМ (обычно десятки нанометров), таким образом, чтобы стало возможным некоторое сравнение между экспериментальными результатами и результатами моделирования.

К сожалению, моделирование длинных фибрилл в явном растворителе сейчас кажется недостижимой целью, поскольку (i) размер системы требует больших вычислительных ресурсов, (ii) время эволюции, необходимое для того, чтобы большая система приняла свою отрелаксированную конформацию, составляет намного дольше, и растворитель предотвращает крупномасштабные конформационные переходы за разумное время моделирования.

В этой работе мы решили опустить явные молекулы растворителя, но все же частично имитировать растворитель, используя подход диссипативной динамики частиц.

Как и в предыдущем разделе, две однослойных периодических укладки и две двухслойных укладки использовались для создания длинных фибрилл путем репликации периодической ячейки вдоль оси филамента. Для выбранных систем были сконструированы плоские прямые фибриллы с длиной 80  $\beta$ -нитей (приблизительно 40 нм) и подвергнуты моделированию методом МД до 10 нс с использованием термостата на основе метода диссипативной динамики частиц со специальным протоколом релаксации и нагрева (см. Подраздел “Методы”), использованного в течение первых 2,5 нс моделирования. Длина фибриллярных агрегатов была выбрана достаточно большой, чтобы избежать влияния возможных “краевых” эффектов на общую конформацию агрегатов (которая для амилоидных фибрилл родственного происхождения может быть оценена

в 10 нм [639]) с одной стороны, и вычислительная сложность задачи была разумной с другой стороны. Длина фибриллярных агрегатов, используемых в современных работах по полноатомному моделированию амилоидных фибрилл, что, как считается, позволяет проследить их геометрические характеристики, составляет около 20-30 нм [640; 641].

Исходные структуры фибрилл и их эволюцию во время моделирования можно увидеть на рисунках 5.22 и 5.23. Большие конформационные изменения, которые происходят в течение короткого времени моделирования, были одной из методологических проблем настоящей работы, другой проблемой был правильный учет дальних взаимодействий, поскольку оказалось, что далекодействующие взаимодействия (такие как электростатические взаимодействия) ответственны за крупномасштабное конформационное поведение фибрилл, и их усечение на малых расстояниях может изменить крупномасштабную конформацию фибрилл. Эти вопросы дополнительно разъясняются в следующем разделе о методах.

Чтобы изучить влияние тиофеновой и пептидной составляющих на крупномасштабную морфологию фибрилл, мы дополнительно провели моделирование фибрилл, состоящих только из пептидной части гибридных молекул (см. рисунок 5.24).

## Результаты

### *Качественное описание*

Значительное увеличение длины агрегатов фибрилл по сравнению с предыдущим исследованием позволило выявить морфологию агрегатов и их поведение в более крупном нанометровом масштабе в режиме, когда длина фибрилл намного больше их ширины.

Конформационная эволюция сконструированных агрегатов во время моделирования позволяет как изучить форму и морфологию фибрилл, так и проследить влияние межмолекулярного расположения на конформацию агрегатов на наномасштабе. Сравнительный анализ различных фибрилл позволяет обсудить роль различных межмолекулярных взаимодействий в стабильности и поведении агрегатов.

Исчерпывающий список моментальных снимков, описывающих конформационную эволюцию различных типов фибрилл, представлен на рисунках 5.22 и 5.23. Хотя все системы продемонстрировали свою стабильность в ходе моделирования, то есть все молекулы сохранили свое относительное положение в агрегатах относительно их соседей, наблюдались конформационные перестройки как на молекулярном, так и на надмолекулярном уровне. Во всех случаях организация  $\beta$ -листов доминировала в структуре и оставалась основным каркасом для организации фибрилл. Протокола релаксации в 2,5 нс в сочетании с дальнейшим моделированием в течение 1-7 нс было достаточно, чтобы понять основные характеристики морфологии фибрилл и ее эволюцию (изгиб, кручение, закручивание и т. д.).

Как видно из рисунка 5.22, однослойные фибриллы способны к более выраженной конформационной перестройке, чем двухслойные фибриллы, поскольку двухслойная организация фибрилл делает их жесткими.

Моделирование длинных однослойных фибрилл (рис. 5.22) показывает различные типы конформационной морфологии скручивания, которые могут принимать фибриллы. Фибрилла SL-AP (рис. 5.22а) изменила свою конформацию таким образом, что лист начал завиваться, образуя тип плотно намотанной ленты с левосторонней спиралью с шагом спирали около 10 нм и углом спирали около  $30^\circ$ . Фибрилла SL-PAR (рис. 5.22б) образовывала другой тип ленты с левосторонней спиралью, ее шаг спирали составляет около 7 нм, а угол спирали близок к  $90^\circ$ , образуя таким образом полую кольцеобразную структуру.

По геометрическим соображениям левая осевая скрученная лента ( $\beta$ -лист) скручивается, образуя левую спираль из этой ленты, однако эта спираль может иметь два разных типа, если поверхности ленты ( $\beta$ -лист) разные (в нашем случае гидрофобные или гидрофильные). Легко показать, что эти два типа отличаются поверхностью ленты, которая обращена на внутреннюю или внешнюю сторону спирали. Эффекты растворителя должны играть значительную роль в скручивании спирали, и это должно зависеть от баланса между взаимодействиями боковых цепей на каждой стороне ленты друг с другом, а также с растворителем. В ранее проведенном моделировании коротких фибрилл в растворителе указывает на тип скручивания, когда боковые цепи треонина (гидрофильная сторона ленты) находятся на внутренней стороне спирали, однако эти фибриллы были недостаточно длинными, чтобы выявить этот эффект в более широком масштабе. Моделирование длинных фибрилл в ваку-

уме действительно подтвердило это наблюдение: фибриллы SL-AP и SL-PAR закручены в спираль таким образом, что боковые цепи треонина находились внутри спирали. Предпочтение фибрилл скручиваться таким образом можно объяснить тем фактом, что во время скручивания боковые цепи треонина сближаются и образуют дополнительные водородные, что является энергетически выгодным. Для моделирования чистых пептидных фибрилл (см. рисунок 5.24) наблюдали такое же поведение, таким образом, этот тип скручивания является характерной особенностью пептидного остова фибриллы и на него не влияют взаимодействия тиофена или алкильной цепи.

Однако следует отметить, что моделирование в вакууме играет вспомогательную роль в выявлении геометрии изгибающихся фибрилл, а не как средство определения точного параметра их геометрии, поскольку отсутствие растворителя не только перенормирует взаимодействия, но и делает конформационный энергетический ландшафт довольно неровным, начальная планарная конформация фибриллы энергетически очень неблагоприятна и нестабильна. Что касается типа завивки (гидрофильная или гидрофобная сердцевина спирали), моделирование с растворителем следует рассматривать как более надежное свидетельство типа завивки. Так, в нашей более ранней работе [A35] было, например, замечено, что при несколько иных условиях фибрилла SL-AP в вакууме эволюционировала в спираль с гидрофобным ядром, в то время как чистая пептидная фибрилла все еще завивалась в спираль с гидрофильным ядром.

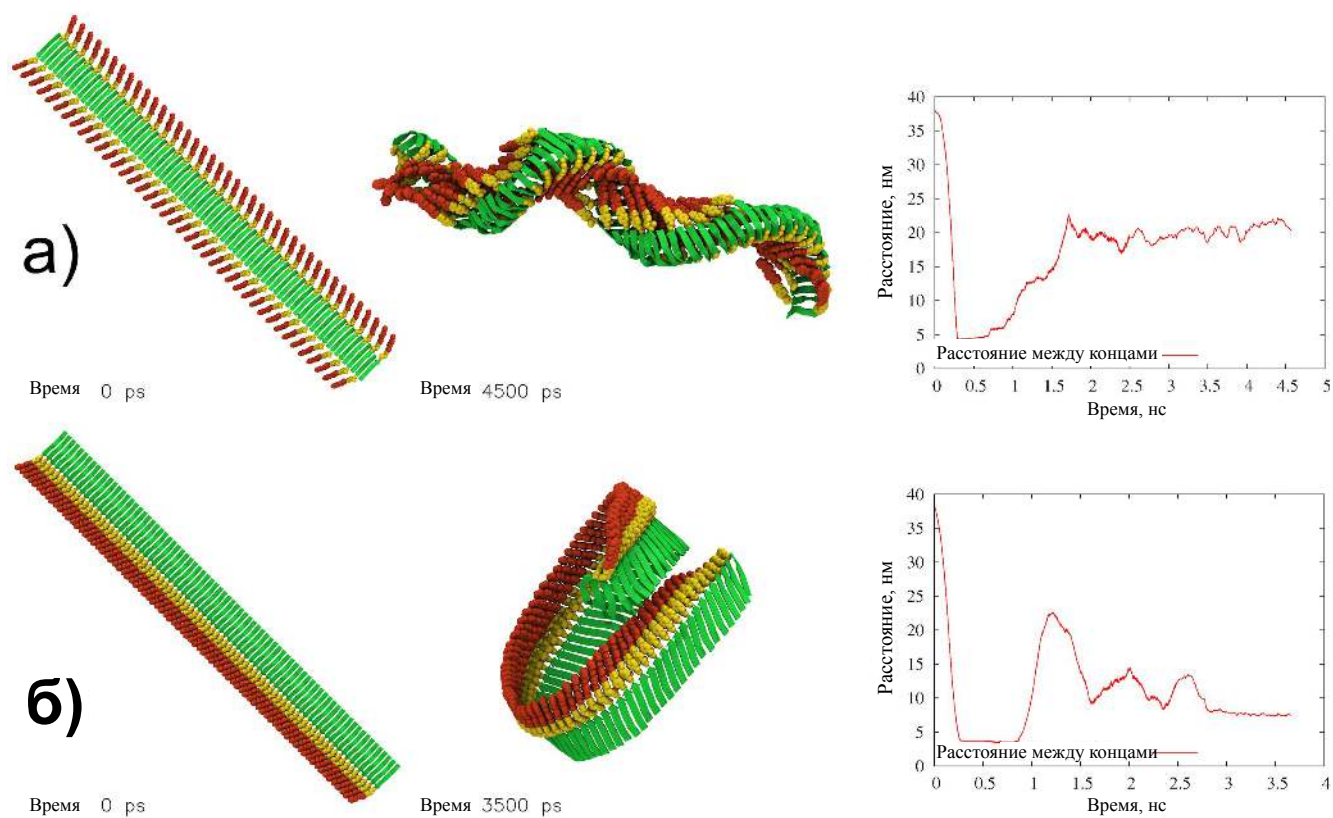
В отличие от однослойных фибрилл, двухслойные фибриллы дают гораздо более жесткие и линейные ленты из фибрилл. Фибрилла DL-AP (рис. 5.23а) превратилась в слегка скрученную влево ленту, которая изгибается примерно на  $30^\circ$  на 80  $\beta$ -нитей. В течение начального периода релаксации (см. График расстояний от конца до конца на Рисунке 5.23а) фибрилла испытывает высокие конформационные изгибные флуктуации с изменением направления изгиба, поскольку избыточная кинетическая энергия отводится от фибриллы, она принимает стабильную изогнутую форму с некоторыми колебаниями вокруг него. Общее осевое скручивание можно оценить примерно в  $1^\circ$  на  $\beta$ -нить.

Фибрилла DL-PAR (рис. 5.23а) проявляет несколько иной тип поведения; она приняла значительно более закрученную в осевом направлении конформацию с приблизительно  $2,8^\circ$  на  $\beta$ -нити, в то время как почти не наблюдалось предпочтительного изгиба фибриллы. Сравнивая поведение фибрилл DL-AP и DL-PAR, можно дополнительно предположить, что повышенная скорость скру-

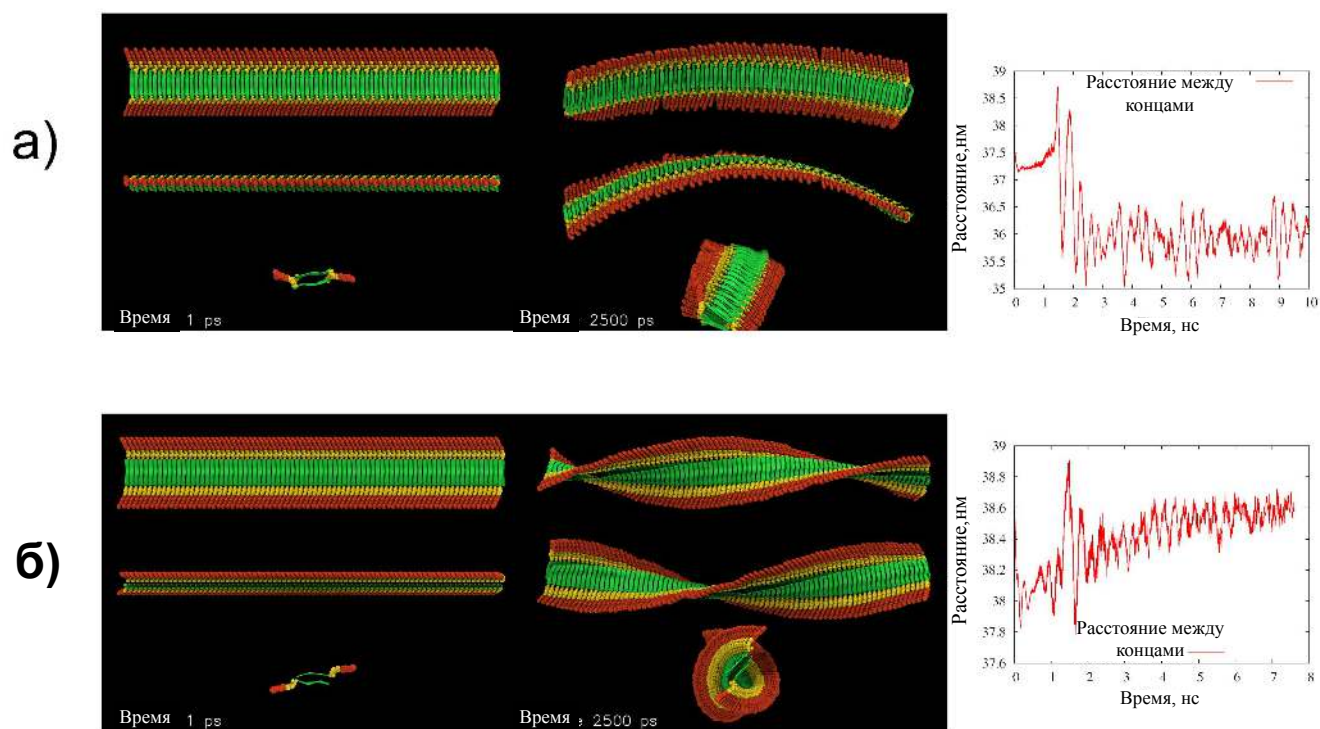


чивания приводит к увеличению жесткости фибриллы на изгиб. Более того, из соображений симметрии весьма вероятно, что наблюдаемому изгибу фибриллы DL-AP частично способствуют эффекты конечного размера, поскольку длины фибриллы недостаточно, чтобы проявить полное влияние аксиального скручивания (когда фибрилла закручивается не менее чем на  $180^\circ$ ), в то время как скручивание делает фибриллу осесимметричной и, таким образом, предпочтительное направление изгиба может быть потеряно. Следовательно, вероятно, что в еще большем масштабе морфологические различия DL-AP и DL-PAR фибрилл будут редуцированы только до различного периода осевого скручивания и персистентной длины. Наши текущие данные моделирования предсказывают приблизительное значение расстояния спиральной закрутки (когда плоскость фибриллярной ленты закручена на  $360^\circ$ ) для DL-AP - 170 нм, для DL-PAR около 60 нм.

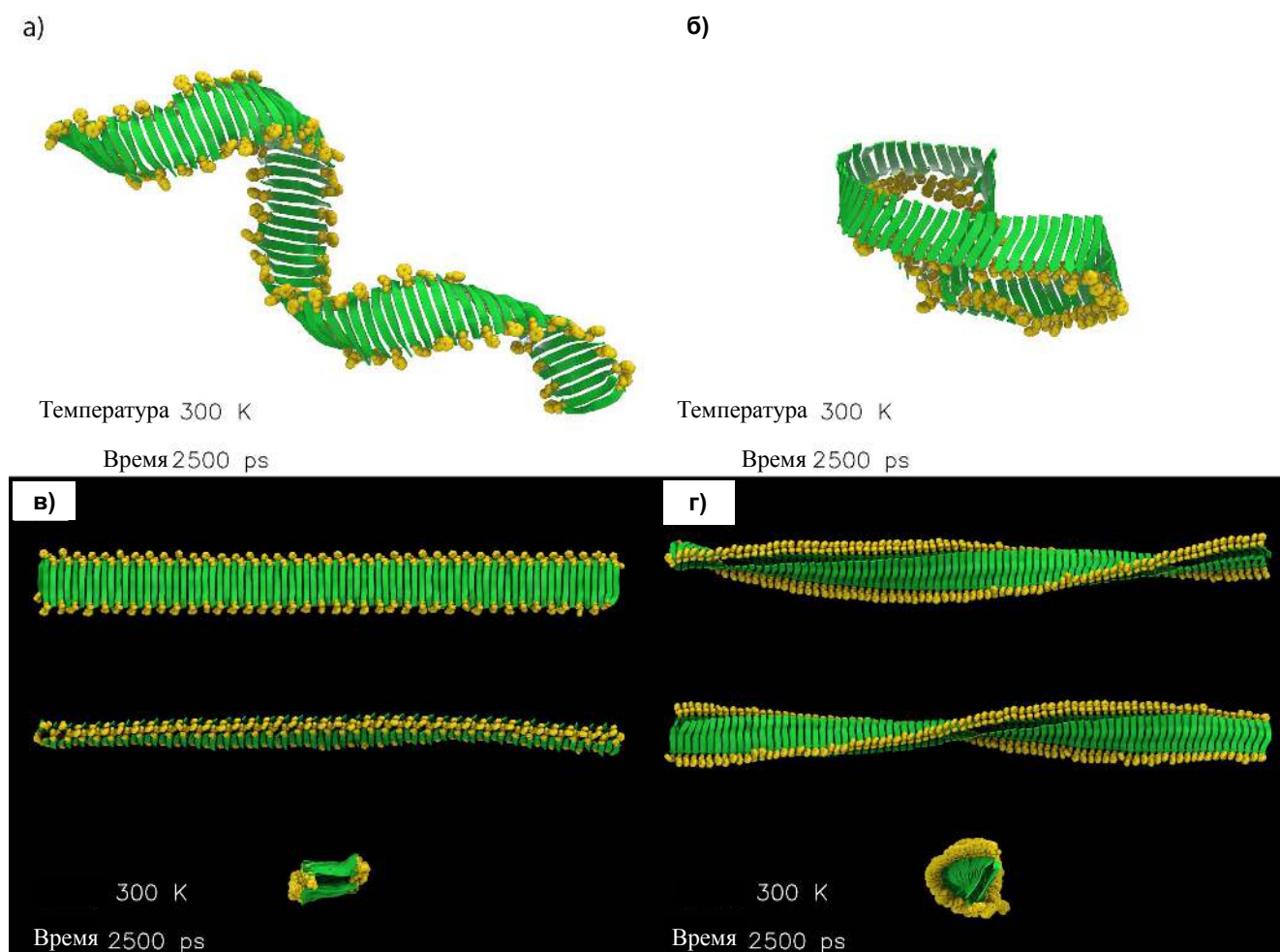
Стоит также сравнить моделирование тиофен-пептидных фибрилл и чистых пептидных фибрилл (см. рисунок 5.24). Из снимков видно, что основная конформационная структура одинакова, т.е. пептидный остов играет доминирующую роль в определении общей конформации. Однако видны определенные различия, которые можно объяснить влиянием блоков алкилированного квадратиофена: (i) фибрилла SL-AP становится более компактной в присутствии синтетических блоков, (ii) для фибриллы SL-PAR присутствие синтетических блоков препятствует взаимодействию конца фибриллы и делает структуру менее узловатой, (iii) для фибриллы чистого пептида DL-AP не наблюдалось предпочтительного изгиба, что означает, что синтетические блоки могут вызывать дополнительное изгибание фибриллы для насыщения их стерических взаимодействий.



**Рисунок 5.22** — Первоначальные снимки и эволюция фибрилл во время моделирования однослойных структур тиофен-пептид, а также график эволюции расстояния от конца до конца: а) фибрилла типа SL-AP, б) фибрилла типа SL-PAR. Алкильные хвосты не изображены, тиофеновая часть изображена красными сферами, пара-азидофенилаланиновый линкер изображен желтым цветом, пептидные части изображены зелеными стрелками.



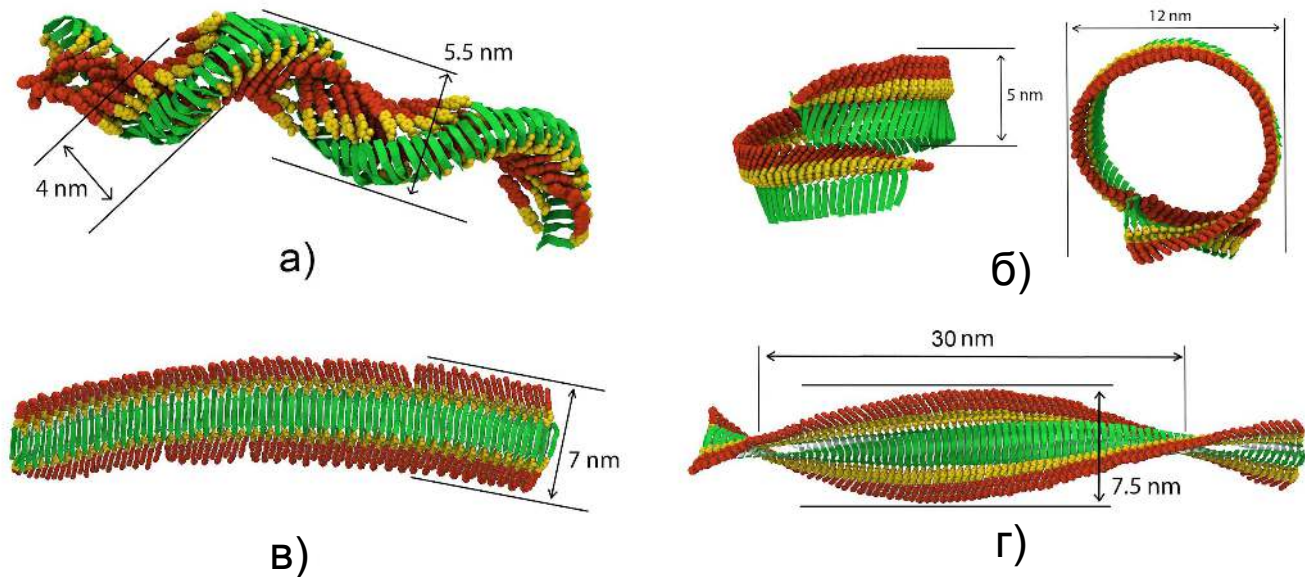
**Рисунок 5.23** — Первоначальные снимки и эволюция фибрилл во время моделирования двухслойных структур тиофен-пептид, а также график эволюции расстояния от конца до конца: а) фибрилла типа DL-AP, б) фибрилла типа DL-PAR. Одна и та же фибрилла изображена в трех проекциях. Алкильные хвосты не изображены, тиофеновая часть изображена красными сферами, пара-азидофенилаланиновый линкер изображен желтым цветом, пептидные части изображены зелеными стрелками.



**Рисунок 5.24** — Снимки окончательной конформации чистых пептидных фибрилл (без кватротиофеновых фрагментов). Пептидный остов изображен зелеными стрелками, боковая цепь фенилаланина, которая ранее использовалась для соединения тиофена, изображена желтым. Пептидные фибриллы соответствуют следующим гибридным фибриллам: SL-AP (а), SL-PAR (б), DL-AP (в), DL-PAR (г).

### *Количественное описание*

Характерные геометрические параметры окончательной конформации фибрилл представлены на рисунке 5.25. Значения винтовой закрутки и шага см. в предыдущем подразделе.



**Рисунок 5.25** — Окончательные снимки и эволюция крупномасштабных фибрилл и их размеры: а) фибрилла типа SL-AP, б) фибрилла типа SL-PAR, в) фибрилла типа DL-AP, г) фибрилла типа DL-PAR. Алкильные хвосты не изображены, тиофеновая часть изображена красными сферами, параазидофенилаланинский линкер изображен желтым цветом, пептидные части изображены зелеными стрелками.

Постоянная длина полимерной цепи является их ключевой механической характеристикой, которая определяется как длина, на которой корреляции в направлении теряются.

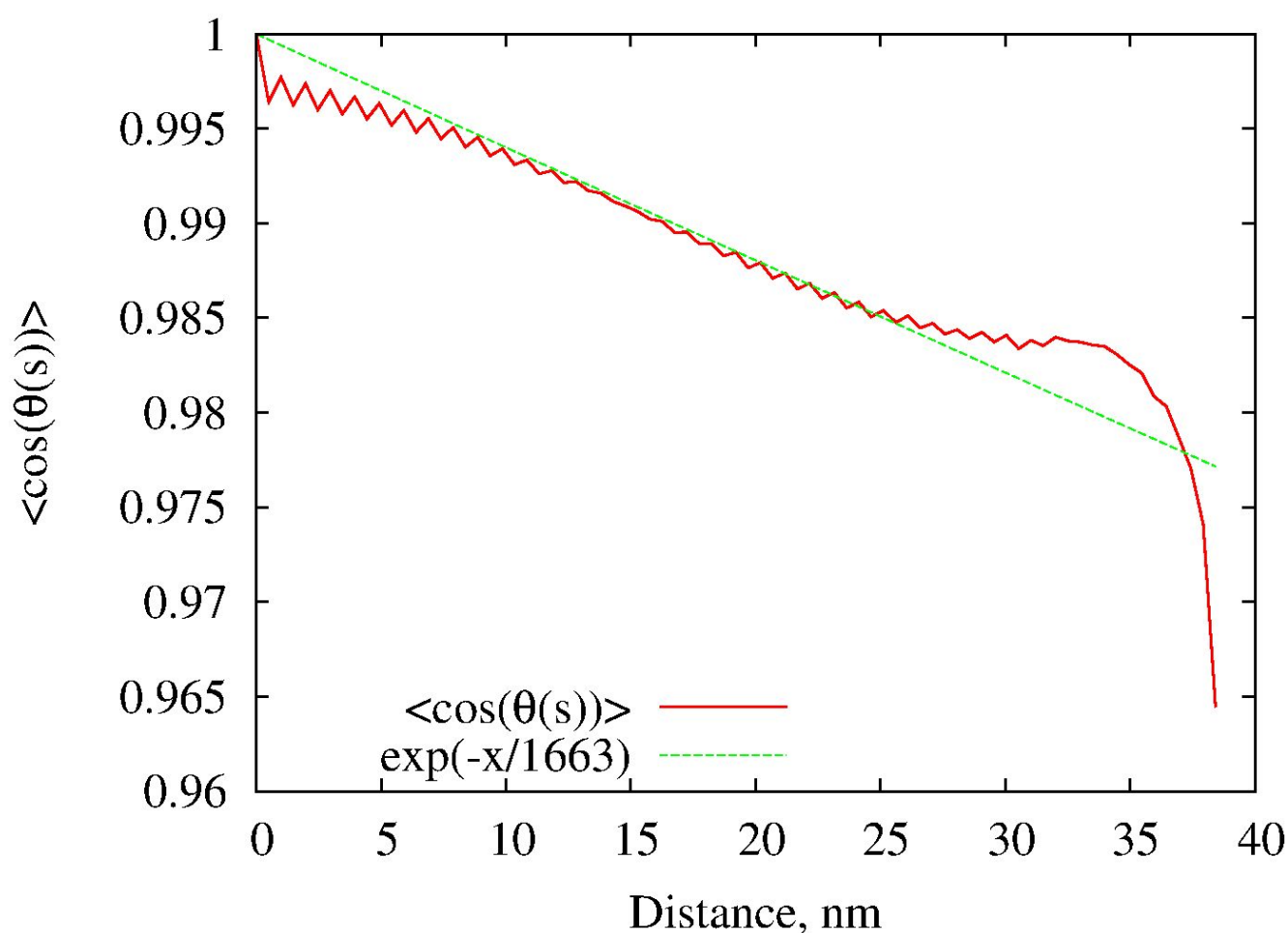
$$\langle \cos \theta(l) \rangle = \exp(-l/P) \quad (5.1)$$

В формуле (5.1)  $P$  - постоянная длина,  $l$  - контурная длина полимера,  $\theta$  - касательный угол между точками в полимере, разделенными контурной длиной  $l$ .

Концепция персистентной длины может также применяться для описания конформационного поведения и жесткости фибрилл. Однако это не всегда может быть сделано простым способом в нашем моделировании, поскольку некоторые из фибрилл (например, однослойные фибриллы) имеют довольно сложное конформационное поведение и потребуют очень обширной конформационной выборки. Для фибриллы DL-AP, которая демонстрирует постоянное поведение при изгибе, также необходимо расширить понятие персистентной длины.

В этой работе мы ограничились расчетом персистентной длины фибриллы DL-PAR, поскольку это можно было сделать прямым способом, пока фибрилла,

скорее всего, была прямой, и для ее описания можно применить персистентную модель гибкости полимера. Средний косинус угла между касательными как функция контурной длины показан на рисунке 5.26. Из графика видно, что средний косинус монотонно уменьшается с более резким уменьшением, когда расстояние сопоставимо с длиной моделируемой фибриллы. Последнее отклонение от экспоненциального спада объясняется концевыми эффектами. Подгонка зависимости показателем степени дает персистентную в длину 1,7 мкм. Это подтверждают экспериментальные данные, в которых персистентная длина была определена порядка микрометров. Для более подробного обсуждения см. раздел “Обсуждение” ниже.



**Рисунок 5.26** — Зависимость среднего косинуса между касательными от длины контура фибриллы для типа фибриллы DL-PAR, усредненного за последние 4 нс моделирования (красная линия) и аппроксимированного показателем степени (зеленая линия). Конформация фибриллы была аппроксимирована кривой, проходящей через центры масс пептидных фрагментов.

## Методы

Моделирование длинных фибриллярных агрегатов представляет собой серьезную проблему для моделирования молекулярной динамики. Основные проблемы связаны с (i) необходимостью достаточно точного учета дальнедействующих взаимодействий и (ii) необходимостью постепенной и осторожной конформационной релаксации исходной структуры в разумные сроки. Протокол конформационной релаксации и параметры моделирования, разработанные для текущего моделирования фибрилл, описаны ниже. Для моделирования использовался пакет программ LAMMPS [632].

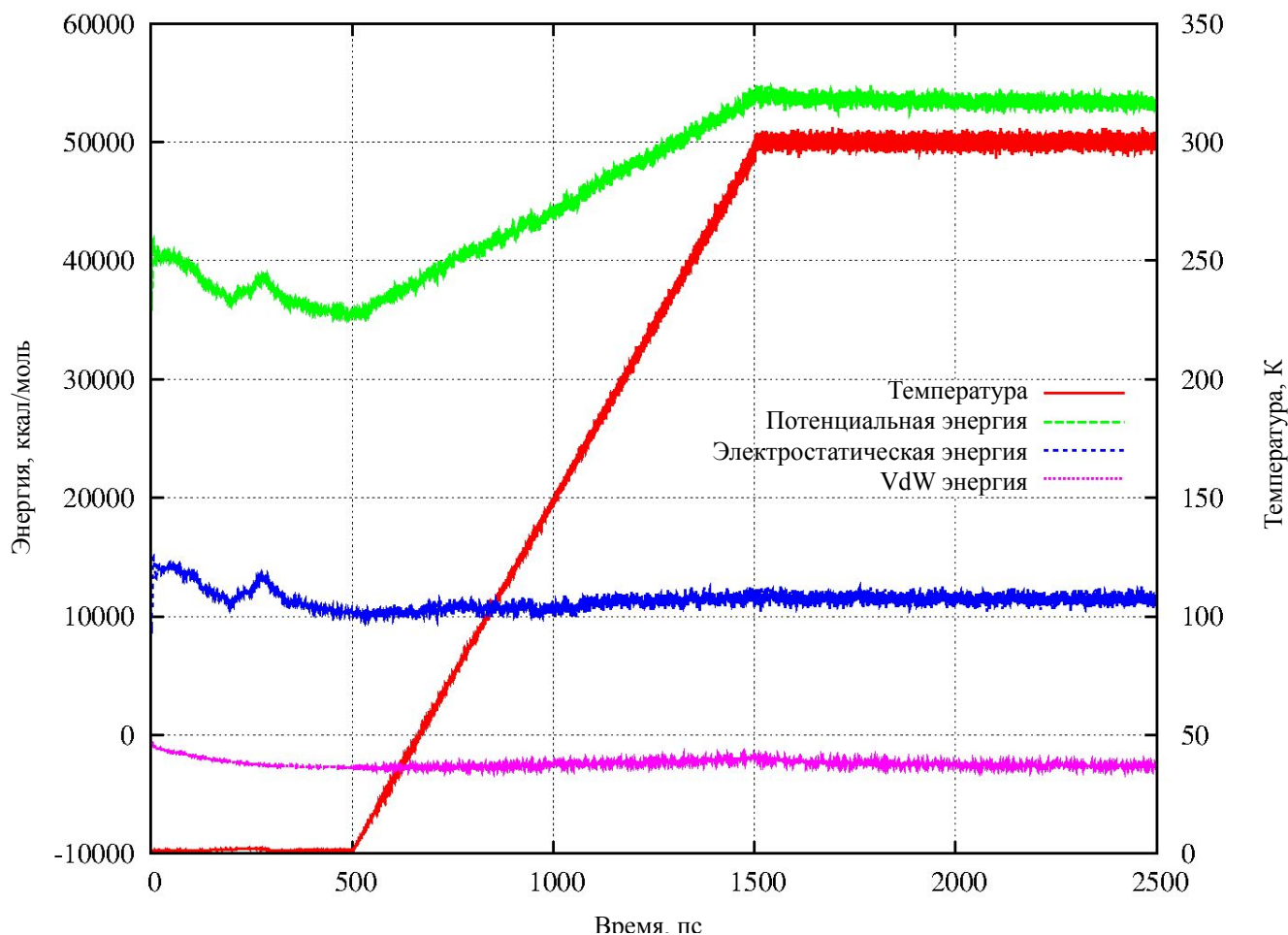
**Шаг 0:** Минимизация энергии молекулярной системы. Радиус обрезания для невалентных ван-дер-ваальсовых и кулоновских взаимодействий был установлен на  $400\text{\AA}$  (фактически без ограничений). Минимизация энергии с использованием версии Полака-Рибьера алгоритма сопряженного градиента проводилась до тех пор, пока максимальная сила в системе не стала меньше  $100$  ккал/моль/А.

**Шаг 1:** Релаксация с использованием моделирования методом МД при  $T = 1\text{K}$  с термостатированием на основе метода диссипативной динамики частиц (ДДЧ) в течение  $0,5$  нс. Радиус обрезания для невалентных ван-дер-ваальсовых взаимодействий был установлен на  $20\text{\AA}$ , а для кулоновского взаимодействия - на  $50\text{\AA}$ . Радиус обрезания ДДЧ-взаимодействий составлял  $20\text{\AA}$ . Параметр трения в термостате ДДЧ был установлен на  $1$   $^{-1}$ . Шаг интегрирования составлял  $1$  фс.

**Шаг 2:** Релаксация с использованием МД моделирования с постепенным нагревом от  $1\text{K}$  до  $300\text{K}$  в течение  $1$  нс. Все параметры были аналогичны предыдущему шагу, за исключением изменений температуры, вызванных термостатом ДДЧ.

**Шаг 3:** Релаксация с использованием МД моделирования при  $T = 300\text{K}$  в течение  $1$  нс. Все параметры были аналогичны, за исключением термостата ДДЧ. Параметр трения был установлен на  $0,2$   $^{-1}$ , что обеспечивало меньшее трение, чем на предыдущих этапах.

**Шаг 4:** Производственный МД прогон до  $10$  нс. Все параметры аналогичны предыдущему шагу.



**Рисунок 5.27** — Иллюстрация протокола моделирования MD для моделирования длинной фибриллы SL-AP. На графике показано изменение параметров энергии и температуры в течение начальных 2,5 нс моделирования.

Этапы релаксации 1-3 дополнительно проиллюстрированы на Рисунке 5.27, на котором показаны изменения температуры и энергии в системе во время работы. Ось времени начинается в тот же момент, что и на графиках-вставках рисунков 5.22 и 5.23.

### 5.2.6 Адсорбция агрегатов на поверхности

В эксперименте фибриллы обычно визуализируются с помощью атомно-силовой микроскопии (АСМ), образец наносится методом литья (drop casting) или центрифугирования из раствора на подложку. Чтобы выявить возможное влияние субстрата на морфологию фибрилл, а также понять геометрические



параметры адсорбированных фибрилл, мы провели серию расчетов по имитации адсорбции.

Длинные двухслойные фибриллы (DL-AP и DL-PAR) длиной  $80 \beta$  ( $\sim 40$  нм) мы поместили вблизи поверхности графита в их релаксированных конформациях и моделировали их адсорбцию.

Используя алгоритм виртуального АСМ (см. “Методы”), были визуализированы профили высот адсорбированных фибрилл.

## Результаты

Фибриллы были размещены в непосредственной близости от явно смоделированной поверхности графита. Исходная конформация фибрилл была принята в их наиболее благоприятном состоянии после моделирования и релаксации, описанные в предыдущем разделе. Эволюция морфологии фибрилл около подложки проиллюстрирована снимками на рис. 5.28 и рис. 5.29. Адсорбция фибрилл была быстрым процессом, и после 1,5 нс моделирования фибриллы достигли своего устойчивого адсорбированного состояния.

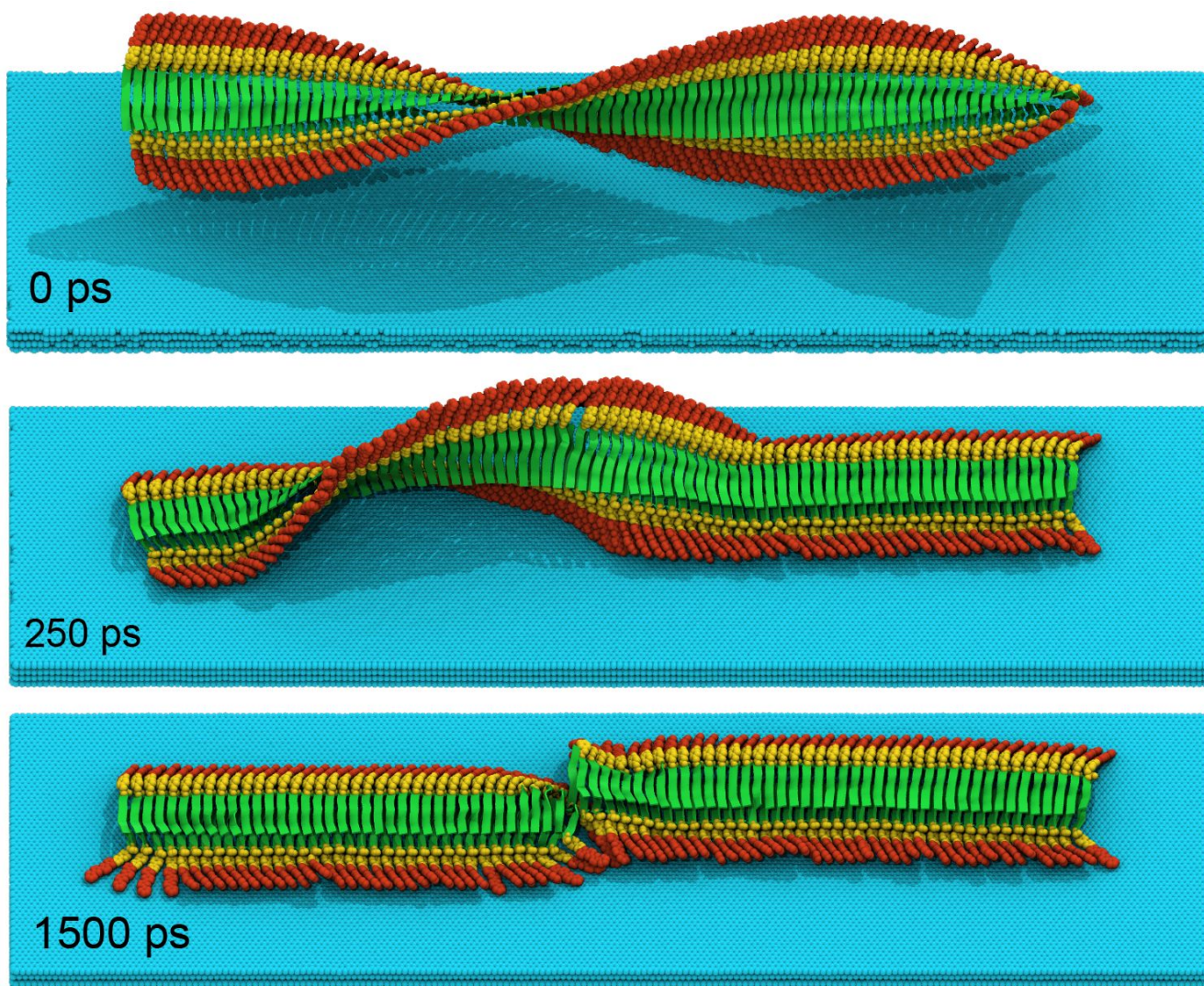
Как видно на рисунке 5.28, фибрилла DL-PAR, которая изначально имела скрученную конформацию, раскручивалась в процессе адсорбции и приобрела плоскую конформацию. Во время этого процесса сначала адсорбировались два конца фибриллы, а затем фибрилла разрывалась посередине, чтобы произошло полное раскручивание. Это замечательное поведение показывает, что: (i) поверхностные силы могут быть достаточно сильными, чтобы раскрутить фибриллу, (ii) фибрилла в целом является стабильным агрегатом, способным противостоять поверхностным силам, поскольку внутренняя структура фибриллы была сохранена. Фибрилла DL-AP (см. рис. 5.29) в процессе адсорбции также приобрела плоскую конформацию, но разрывов не наблюдалось, поскольку они не были необходимы с точки зрения структуры.

Для дальнейшего анализа структуры фибриллы в адсорбированном состоянии мы представляем моментальные снимки окончательных конформаций с разных точек зрения в более подробном представлении (также с явно показанными алкильными цепями) на рисунках 5.30 и 5.31.

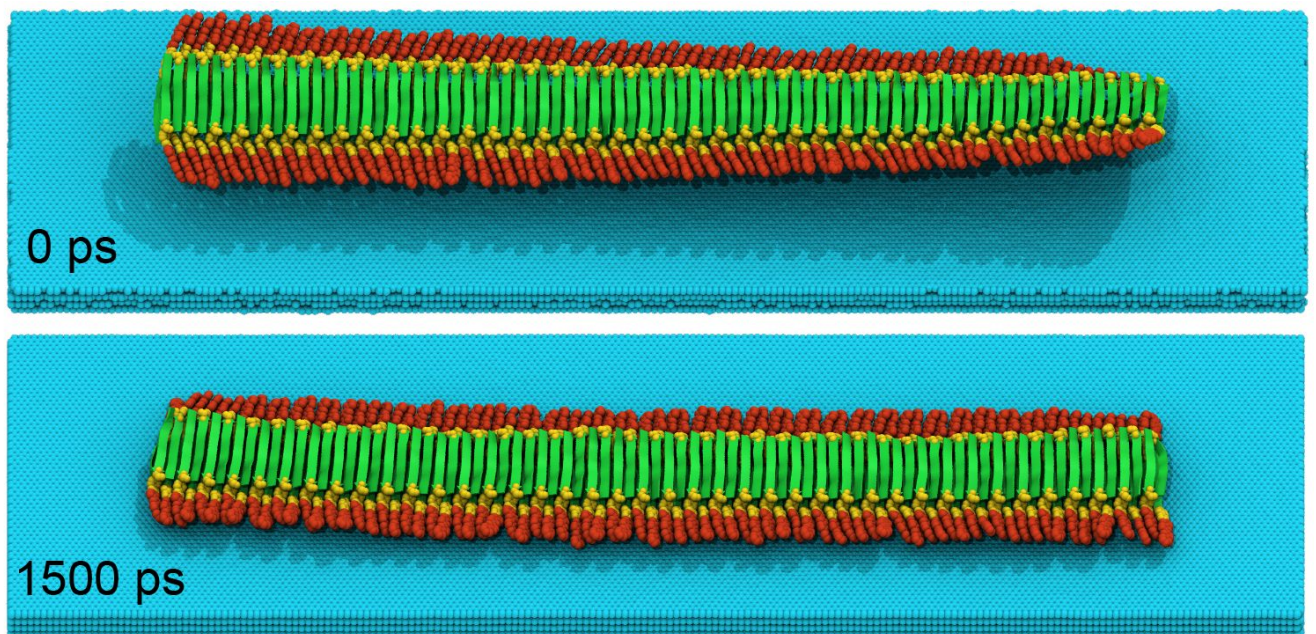
На снимках, визуализированных перпендикулярно оси фибрилл (рисунок 5.30б и рисунок 5.30б), видно, что геометрия поперечного сечения двух типов фибрилл различается, и конфигурация алкильных цепей играет здесь важную роль - они стерически мешают адсорбции кватротиофеновых фрагментов на графите. Для фибриллы DL-PAR сегменты кватротиофенов, а также алкильные цепи расположены согласованно с обеих сторон фибриллы - это приводит к более упорядоченной структуре и более плотной упаковке и адсорбции синтетических фрагментов на графите. Для фибриллы DL-AP из-за ее геометрии алкильные цепи образуют более неупорядоченную структуру и даже образуют промежуточный молекулярный слой между графитом и тиофенами. Различия в упаковке алкильных цепей приводят к тому, что поперечное сечение фибриллы DL-PAR толще в середине (пептидная часть), чем на краях (тиофен-алкильная часть), в то время как поперечное сечение фибриллы DL-AP имеет более прямоугольную форму.

Чтобы улучшить понимание геометрии фибрилл, мы рассчитали виртуальные изображения АСМ, которые соответствуют адсорбированным структурам (см. Рисунок 5.30в и Рисунок 5.31в). Эти смоделированные изображения АСМ соответствуют радиусу острия кантилевера в 1 нм. Соответствующие профили высоты поперечного сечения представлены на рисунке 5.32. Четко видно, что эти профили имеют немного другую форму: профиль для фибриллы DL-PAR (рисунок 5.32а) имеет более горбовидную форму, в то время как профиль для фибриллы DL-AP (Рисунок 5.32б) имеет более прямоугольную форму.

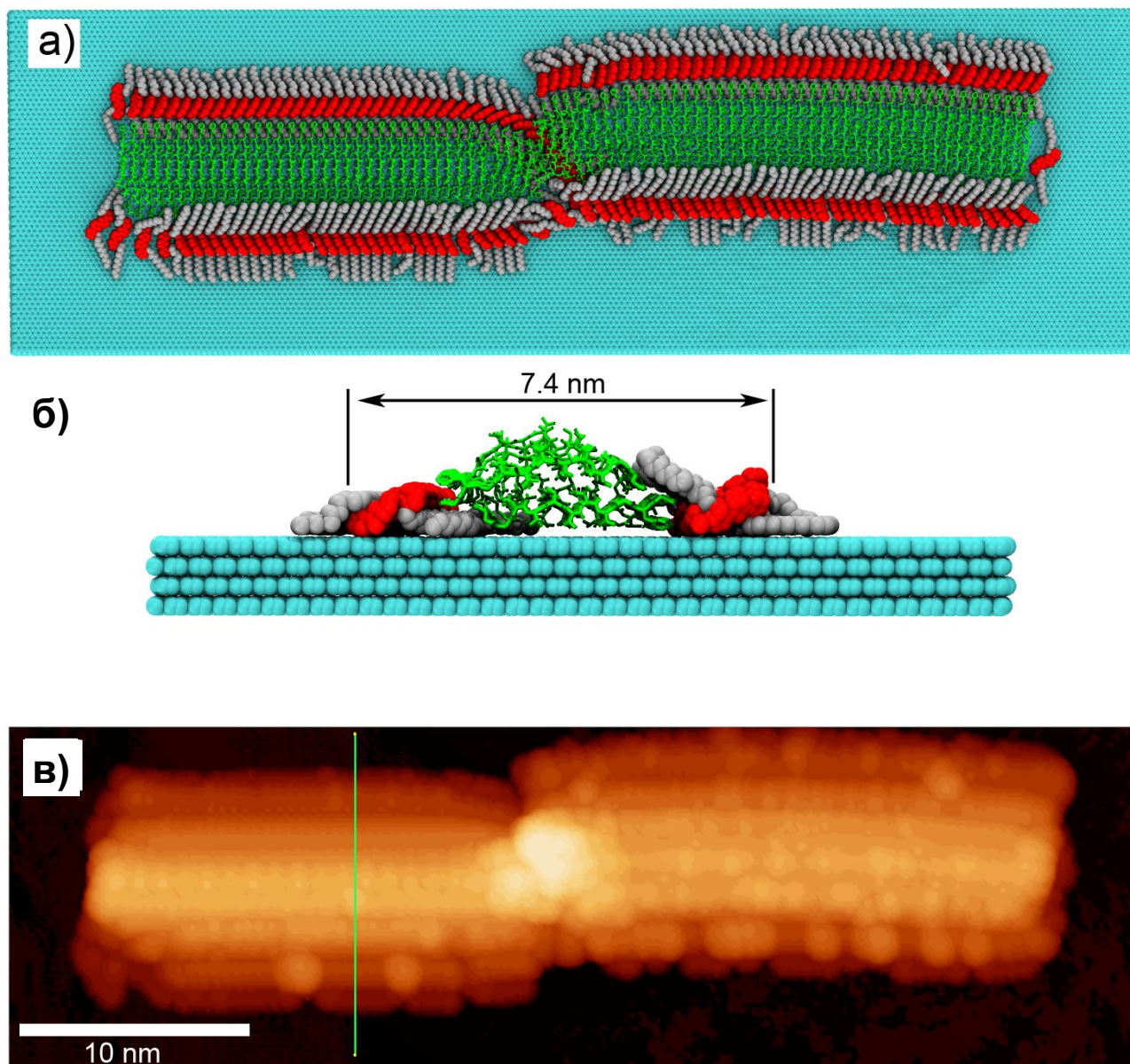
Виртуальные изображения АСМ также позволяют решать проблему оценки ширины и высоты фибрилл как *in silico*, так и экспериментально. Из рисунков 5.30б и 5.31б видно, что ширина фибриллы составляет около 7-8 нм, хотя алкильные цепи могут простираться немного дальше. Виртуальные изображения АСМ дают более подробную информацию об этой проблеме и возможность для более прямого сравнения с экспериментальными данными, так как одно только значение ширины не слишком информативно и подвержено неточностям, связанным с конечным радиусом острия кантилевера в АСМ, а также существуют проблемы с определением ширины объектов с гладкими границами. Используя наши виртуальные АСМ изображения, предполагаемые значения ширины для фибриллы DL-PAR составляют около 8-10 нм, а для фибриллы DL-AP - около 8 нм.



**Рисунок 5.28** — Снимки моделирования адсорбции фибриллы DL-PAR на поверхности графита, представляющие эволюцию процесса адсорбции. Поверхность графита (состоящая из 4-х слоев) окрашена в голубой цвет. Изображение фибриллы соответствует тому, что использовалось ранее. В нижнем левом углу каждого снимка указано соответствующее время моделирования.



**Рисунок 5.29** — Снимки моделирования адсорбции фибриллы DL-AP на поверхности графита, представляющие эволюцию процесса адсорбции. Поверхность графита (состоящая из 4-х слоев) окрашена в голубой цвет. Изображение фибриллы соответствует тому, что использовалось ранее. В нижнем левом углу каждого снимка указано соответствующее время моделирования.

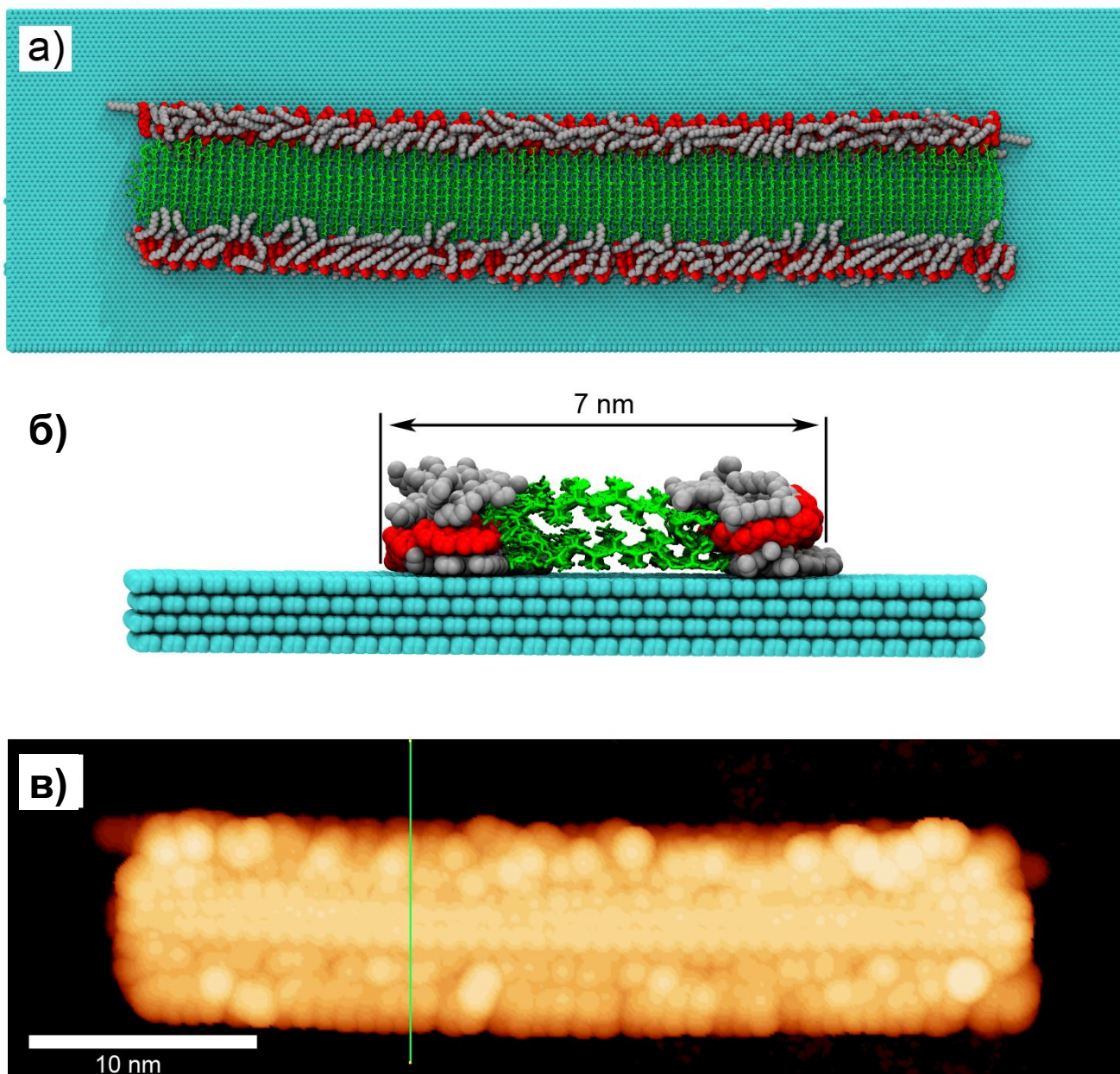


**Рисунок 5.30** — (а), (б) - снимки конечного адсорбированного состояния фибриллы DL-PAR на графите. Графит изображен голубым, кватротиофены - красным, алкильные цепи - серым, пептидные части - зелеными трубками. (в) представлена соответствующая модель АСМ-изображения, зеленая линия соответствует поперечному сечению, представленному на рисунке 5.32а.

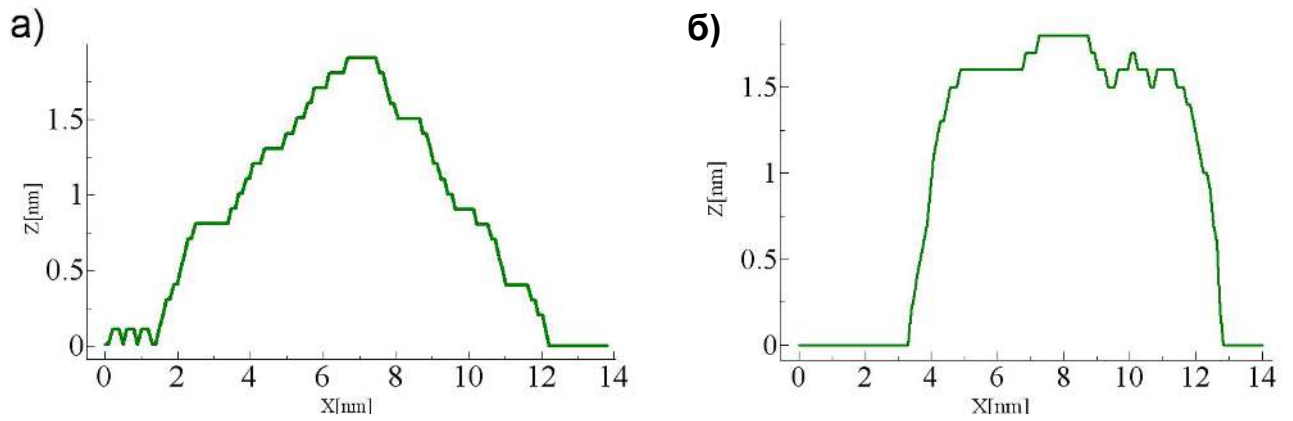
Высота фибрилл - еще один важный параметр, который можно более точно измерить в эксперименте и в меньшей степени зависит от радиуса острия кантилевера. Из наших виртуальных профилей АСМ максимальная высота фибрилл составляет 1,9 нм для фибриллы DL-PAR и 1,7–1,8Å для фибриллы DL-AP.

Дополнительные виртуальные АСМ-измерения однослойных фибрилл были выполнены путем простого отщепления верхнего молекулярного слоя. Максимальная высота однослойной фибриллы оценивается примерно в 1 нм.

Однако следует отметить, что реальный профиль фибриллы может зависеть от многих факторов (радиуса кончика кантилевера, жесткости и амплитуды кантилевера и т. д.), поэтому при сравнении с экспериментальными данными следует проводить измерения также некоторых эталонных структур.



**Рисунок 5.31** — (а), (б) - снимки конечного адсорбированного состояния фибриллы DL-PAR на графите. Графит изображен голубым, кватротиофены - красным, алкильные цепи - серым, пептидные части - зелеными трубками. (в) представлена соответствующая модель АФМ-изображения, зеленая линия соответствует поперечному сечению, представленному на рисунке 5.32б.



**Рисунок 5.32** — Высотные профили поперечного сечения фибрилл, адсорбированных на графите, для фибрилл DL-PAR (а) и DL-AP (б). Соответствующее поперечное сечение показано на Рис. 5.30в и Рис. 5.31в зелеными линиями.

## Методы

### *Построение системы и моделирование*

Графитовая подложка была смоделирована 4 слоями графита, построенными как суперячейка из элементарной ячейки графита путем репликации ее вдоль векторов элементарной ячейки. Общие размеры подложки составляли 50 нм на 15 нм. Все моделирование проводилось с использованием метода моделирования молекулярной динамики с программным обеспечением LAMMPS и силовым полем PCFF, как описано ранее.

Моделирование проводилось с периодическими граничными условиями и NVT-ансамблем. Постоянный объем, а также постоянные размеры слоев графита позволяли поддерживать подложку в вытянутой плоской форме. Неваляные взаимодействия обрабатывались следующим образом: взаимодействие Леннард-Джонса ограничивалось  $10\text{\AA}$ , электростатическое взаимодействие обрабатывалось с помощью алгоритма PPPM [633] с отсечкой в реальном пространстве  $10\text{\AA}$  и значением точности  $10^{-4}$ . Термостат ДДЧ использовался с референсной температурой 300 К, радиусом отсечки  $10\text{\AA}$ , а параметр трения был установлен на  $0,2^{-1}$ .

### *Виртуальный АСМ анализ*

Процедура виртуальных АСМ-измерений выполнялась с помощью специально разработанного программного обеспечения. В нашем алгоритме сфера

зонда диаметром 1 нм приближалась к поверхности подложки в каждой точке горизонтальной сетки с шагом 0,1 нм. Затем определялось наивысшее положение зондирующей сферы, когда сфера не перекрывалась ни с одной из сфер ван-дер-Ваальса атомов в системе. Вертикальное положение сферы определялось методом деления отрезка пополам с точностью до 0,1 нм. Полученные данные были дополнительно визуализированы с помощью программного обеспечения для визуализации WSxM [642].

### 5.2.7 Обсуждение

#### **Комбинированная вычислительная/экспериментальная интегративная методология для предсказания структуры/свойств амилоидоподобных фибрилл**

В настоящей работе мы продемонстрировали методологию мультимасштабного компьютерного моделирования для понимания структуры, морфологии и свойств конкретных фибриллярных агрегатов с ограниченными экспериментальными знаниями о межмолекулярном структурном устройстве. Методология была разработана на основе идей, сформулированных выше, и состоит из 3 этапов: (i) сначала предсказываются пробные межмолекулярные структуры, соответствующие имеющимся экспериментальным данным, (ii) фибриллы, основанные на этих структурах, затем анализируются с использованием методов мультимасштабного моделирования и (iii) их крупномасштабные свойства дополнительно сравниваются с доступными экспериментальными данными, что дает механизм обратной связи для понимания правильности первоначальных предположений. Эта методология представляет интерес сама по себе, поскольку ее можно применять не только к гибридам тиофен-пептид, но и к широкому классу амилоидоподобных волокон. В этом обсуждении мы хотели бы более подробно остановиться на особенностях методологии как комбинированного экспериментально-теоретического подхода, предназначенного для решения проблем исследования амилоидоподобных фибрилл, и описать пути его улучшения.



Предсказание возможных начальных межмолекулярных структур - первый ключевой шаг в нашей методологии. В идеальном случае желательна пространственная структура микрокристаллов, образованных исследуемыми молекулами, решаемая с помощью рентгеновской кристаллографии или данные о пространственном расположении молекул в фибриллах, обнаруженные методом ЯМР. Однако эти наборы данных могут быть получены только в некоторых редких случаях, поэтому возможно следовать альтернативной стратегии - построение возможных укладок пептидов, согласующихся с экспериментальными данными, а затем использование методов моделирования для их оптимизации к наиболее энергетически выгодному состоянию. Данные, полученные с помощью различных экспериментальных методов, могут оказаться здесь большим подспорьем и облегчить понимание и построение схемы испытаний. Экспериментальные методы, полезные на этом этапе, сведены в Таблицу 9, Часть А. Опираясь на такие экспериментальные методы, как окрашивание красителем, ИК- и КД-спектроскопия или рентгеновская дифракция и электронная дифракция, можно понять склонность образования пептидами  $\beta$ -листных фибрилл и такие детали их организации, как (i) параллельная или антипараллельная организация  $\beta$ -листных, (ii) расстояние между стопками  $\beta$ -листных, если они образуют стопки, (iii) упаковка боковых цепей и т. д. Эти детали могут быть проанализированы с помощью ряда экспериментальных методов, на практике эксперименты могут быть сложными, и в некоторых случаях единственное, что мы знаем о межмолекулярном расположении, - это данные об существовании  $\beta$ -листных в структуре фибрилл.

При построении пробных укладок на основе одиночных  $\beta$ -листных можно дополнительно руководствоваться основными энергетическими принципами, такими как: (i) количество водородных связей на молекулу, образованную основной цепью и боковыми цепями, (ii) возможные ионные взаимодействия между определенными заряженными остатками или концами пептида, (iii) возможные стэкинг-взаимодействия между ароматическими остатками, (iv) электростатическое отталкивание или притяжение заряженных остатков. Оптимизация предлагаемых схем испытаний должна быть проведена с использованием методов молекулярной механики и молекулярной динамики. Используя эти методы, можно дополнительно сравнить энтальпию образования различных укладок, что дает дополнительный критерий предпочтения одного устройства над другим. Если на основании экспериментальных данных предлагается укладка

$\beta$ -листов, то поэтапно должны быть предложены различные пробные схемы укладки на основе однослойных схем. Здесь дополнительные энергетические соображения могут направлять процесс конструирования, например, такие как (i) ионные, стэкинговые взаимодействия, водородные связи, стерические взаимодействия на границе раздела листов, (ii) влияние растворителя на укладку, включая гидрофобные взаимодействия. Здесь снова можно использовать компьютерное моделирование, чтобы оптимизировать конструкции и дать оценку предпочтения одного устройства над другим. Некоторые пробные расчеты фибрилл в растворителе и анализ взаимодействия фибрилл в растворителе (например, количество водородных связей, образованных между растворителем и фибрилой) могут дать понимание энергетической выгодности различных вариантов укладки. Вклад гидрофобных взаимодействий можно также оценить путем расчета гидрофобной поверхности, доступной для растворителя, в различных экспериментальных схемах.

Однако предлагаемая в настоящее время методология по-прежнему в значительной степени зависит от подхода проб и ошибок и участия человека в процессе. В настоящее время основной проблемой является процедура оптимизации, которая носит локальный характер, поскольку она не может увести конструкцию далеко от первоначальной пробной укладки, созданной вручную. Эта проблема становится более очевидной по мере того, как структуры становятся более сложными, то есть укладкой нескольких  $\beta$ -листов или увеличением сложности молекул, наблюдаемых в конъюгатах пептид-полимер. Возможным решением описанной проблемы может быть разработка специальных эффективных алгоритмов оптимизации, которые будут пытаться найти глобальный энергетический минимум молекулярного расположения, одновременно принимая во внимание экспериментальные ограничения на структурную организацию. Эти методы могут быть основаны на методах параллельного темпирования (позволяющих молекулам легко преодолевать барьеры свободной энергии) [643] или параллельной оптимизации, начиная с различных случайных молекулярных конформаций, как это используется в пакете GULP [644]. Идеальным была бы своего рода комбинация подходов, когда автоматический прогноз кристаллической упаковки можно было бы адаптировать к прогнозированию одномерных амилоидных фибрилл и в то же время учитывать ограничения, исходящие из экспериментальных данных

Экспериментальный подход	Полезные данные
<b>Часть А</b> -методы выявления межмолекулярной структуры	
КД-спектроскопия	Обнаруживает наличие $\beta$ -листов
ИК-спектроскопия	Обнаруживает наличие $\beta$ -листов, возможно, может различать параллельные и антипараллельные $\beta$ -листы
УФ-видимая спектроскопия	При окрашивании конго красным возможно обнаружение амилоидов
Флуоресцентная спектроскопия	При окрашивании Тиофлавином Т или Конго красным возможно обнаружение амилоидов
Рентгеновская дифракция на порошке или волокнах	Шаг-листов, укладка $\beta$ -листов, упаковка боковой цепи, параллельное или антипараллельное расположение
Окрашивание красителем (Конго Красный, Тиофлавин Т)	Обнаружение наличия амилоидов
Электронная дифракция	Шаг $\beta$ -листов, укладка $\beta$ -листов
Модификация химической структуры, зависимость самосборки от различных параметров (растворитель, температура)	Может выявить ключевые меж- и внутримолекулярные взаимодействия, а также влияние растворителя на самосборку - преобладание полярных или гидрофобных взаимодействий.
<b>Часть В</b> - методы визуализации свойств волокна	
Атомно-силовая микроскопия	Размеры фибрилл, параметры скручивания (возможное влияние поверхности)
Просвечивающая электронная микроскопия	Размеры фибрилл, параметры скручивания (возможное влияние субстрата)
Криофиксирующая электронная микроскопия	Размеры фибрилл, параметры скручивания
Рассеяние нейтронов и рентгеновских лучей	Размеры фибрилл, структурные факторы

**Таблица 9** — Экспериментальные методы, используемые для изучения амилоидных фибрилл, а также ценные данные, которые могут быть получены и использованы в обсуждаемой методологии. Для дальнейшего обсуждения мето-

Следующим шагом в нашей многомасштабной методологии является моделирование фибрилл на основе предсказанных межмолекулярных структур в объеме растворителя. Основная цель такого моделирования состоит в том, чтобы (i) проверить стабильность фибрилл в среде растворителя, (ii) оптимизировать и изучить межмолекулярную структуру молекул в фибриллярном состоянии, (iii) выявить и изучить супрамолекулярную морфологию фибрилл в различных средах (например, на подложке).

Опыт, собранный в этой работе, показывает, что наиболее важные проблемы при моделировании фибрилл проистекают из (i) их большого пространственного масштаба, который делает моделирование требовательными к вычислительным ресурсам, (ii) важности дальнедействующих взаимодействий для супрамолекулярной морфологии, правильное рассмотрение которых может значительно увеличить необходимость в вычислительных ресурсах, (iii) крупномасштабные конформационные переходы, которые могут происходить во время релаксации фибрилл, требующие особого внимания, (iv) для моделирования без явного растворителя, большинство алгоритмов термостатирования не подходят. В попытке сбалансировать эти проблемы мы предложили двухэтапный подход к моделированию, который включает (i) моделирование фибрилл среднего размера (длиной около 20  $\beta$ -нитей) в явном растворителе и (ii) моделирование одиночных фибрилл (около 80  $\beta$ -нитей) без явного растворителя. Моделирование в растворителе позволяет оценить стабильность и основные конформационные свойства фибрилл, а также описать взаимодействия растворитель-фибрилла и дополнительно оптимизировать структуру фибрилл в присутствии растворителя. Однако поведение коротких фибрилл в этих симуляциях нельзя с уверенностью экстраполировать на более крупные масштабы, поскольку (i) конформационное поведение может быть искажено краевыми эффектами и эффектами конечного размера, и (ii) некоторые конформационные черты могут быть даже не видны в этих условиях, например небольшие значения осевого скручивания, которое для многих амилоидных волокон может составлять всего 1 – 2° на  $\beta$ -слой [646], или эффекты более крупного масштаба, такие как образование суперспирали и т. д. Следовательно, необходимы некоторые крупномасштабные расчеты, однако с ними имеется несколько очевидных проблем: (i) моделирование в явном растворителе непрактично, поскольку необходимы очень большие системы для моделирования, и требуется гораздо большее время моделирования, чтобы наблюдать большие конфор-

мационные перестройки, которые могут иметь место, потому что они будут затруднены временными масштабами перегруппировки растворителя, (ii) схемы расчета для невалентных взаимодействий (ван-дер-ваальсово-кулоновское) следует применять с осторожностью, поскольку в одномерных периодических системах заряды скоррелированы в пространстве на больших расстояниях, и дальнедействующие взаимодействия играют важную роль в формировании крупномасштабной морфологии (например, скручивания), (iii) протоколы релаксации (включая алгоритмы термостатирования) для изначально планарных фибрилл должны быть тщательно выбраны, поскольку резкая релаксация напряжений в фибрилле может привести к непредсказуемым результатам. Чтобы частично решить эти проблемы, мы ввели метод моделирования фибрилл с помощью термостата ДДЧ и специальный протокол релаксации, который оказался простым и ценным методом. Термостат ДДЧ позволяет эффективно имитировать кинетические эффекты растворителя и управлять конформационной релаксацией фибриллярной структуры предсказуемым и плавным образом за короткий период времени.

Точность наших моделей в вакууме не лишена некоторых ограничений. Во-первых, в нашем моделировании не учитывались энергетические эффекты растворителя, они могут перенормировать электростатические взаимодействия между молекулами фибриллы из-за диэлектрической проницаемости растворителя, а также вызвать некоторые взаимодействия, опосредованные растворителем. Однако мы не думаем, что включение растворителя изменило бы качественное поведение фибрилл при моделировании радикально (хотя диэлектрическая проницаемость смеси метанола и дихлорметана составляет около 20), поскольку конформации фибрилл, наблюдаемые при моделировании в растворителе и вакууме в нашем исследовании были согласованы. Параметры осевого скручивания, измеренные при моделировании в вакууме и растворителе, хотя и различались, например, для фибриллы DL-PAR измеренное закручивание в растворителе составило около  $4,5$  на  $\beta$ -нить, тогда как при моделировании в вакууме около  $2,8^\circ$  на  $\beta$ -нить. Однако это различие также может быть связано с эффектами конечного размера, присутствующими в небольших фибриллах. В любом случае для моделирования в растворителях с более высокой диэлектрической проницаемостью и в воде может быть рекомендовано использование неявных моделей растворителей там, где это возможно.

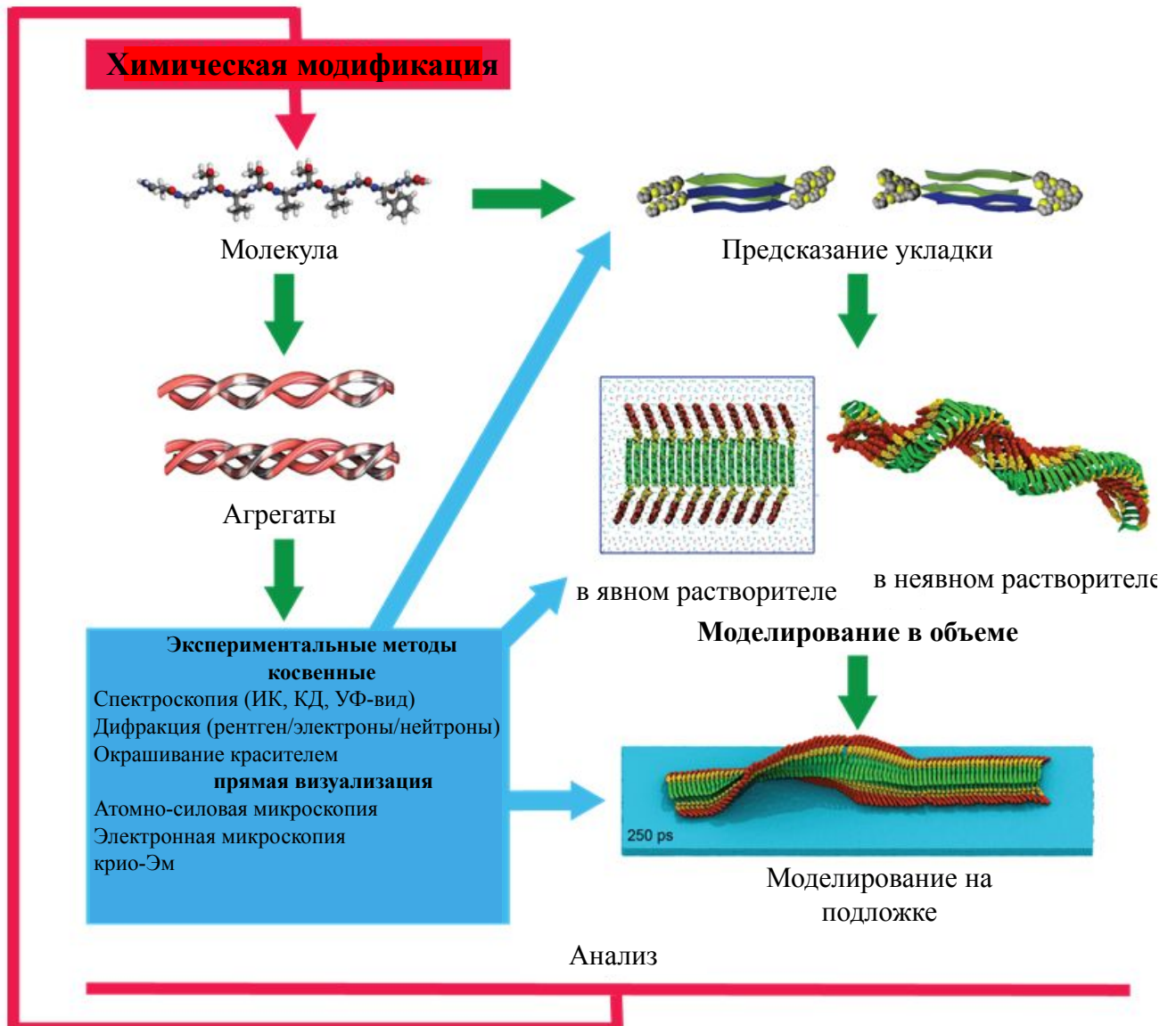
Во-вторых, крупномасштабная конформационная эволюция фибрилл может иметь несколько путей и зависеть от условий исходной конформации, а также от кинетических свойств неясной среды (термостат ДДЧ), особенно для конформационно гибких фибрилл, таких как однослойные. Например, SL-AP, начиная свою эволюцию из идеально прямой конформации, сворачивается в левую суперспираль, однако, если из-за некоторых начальных условий начальные колебания приведут фибриллу к правой суперспирали, она может продолжить этот кинетический путь. Учитывая это, следует проявлять осторожность при создании начальных структур, которые не имеют предпочтения ни к одному из крупномасштабных морфологических состояний. Также рекомендуется начать несколько расчетов с разными случайными начальными числами для термостата ДДЧ или генерации начальной скорости.

Дополнительный метод моделирования, который мы ввели, - это моделирование фибрилл на графитовой подложке после их релаксации в вакууме. Этот метод был предназначен для имитации стандартной экспериментальной процедуры, когда фибриллы, уже образовавшиеся в растворе, наносятся методом центрифугирования на подложку и там анализируется их морфология. Соответствие фибриллярной структуры в растворе и на подложке в настоящее время является предметом дискуссий (см., например, [647]). Однако считается, что структура фибрилл из конъюгатов полимер-пептид, полученных с помощью АСМ-изображений, должна соответствовать структуре фибрилл в растворе, и кручение фибрилл часто можно увидеть на АСМ-изображениях [609; 610]. Здесь мы показали, что это может быть не всегда. *Скрученные в растворе фибриллы могут разорваться и раскрутиться под действием поверхностных сил.*

Соответствующие свойства фибрилл, полученные при моделировании растворителя, вакуума или подложки, такие как ширина, высота, шаг спирали и угол спирали, можно сравнить со свойствами, полученными экспериментальными методами, приведенными в таблице 9, часть Б. Это сравнение обеспечивает механизм обратной связи, для оценки правильности моделирования и прогнозируемой структуры.

В заключение мы хотели бы подчеркнуть, что одно компьютерное моделирование все еще не может предоставить окончательный инструмент, который может предсказать структуру, свойства и морфологию амилоидоподобных фибрилл, исходя только из химической структуры молекул, однако сочетание экспериментальных методов с методами моделирования в синергетической

методологии исследования могут стать таким инструментом и даже в большей степени инструментом для рационального дизайна фибрилл с заданными свойствами для нужд био- и нанотехнологий. Идея такой методологии проиллюстрирована на схеме рисунке 5.33.



**Рисунок 5.33** — Схематическое изображение комбинированной экспериментальной/вычислительной интегративной методологии для изучения амилоидоподобных фибрилл.

## Соотношения структура-морфология-свойства в фибриллах, образованных тиофен-пептидными олигомерами

Многоуровневое моделирование тиофен-пептидных агрегатов выдвинуло на первый план основные принципы организации этих фибриллярных агрегатов, включая детальное взаимодействие между межмолекулярной организацией кватротиофеновых и пептидных блоков и его влияние на структуру, морфологию и свойства фибрилл. Ниже мы обсуждаем наблюдаемые результаты.

### *Паттерны самосборки гибридных фибрилл и иерархия взаимодействий*

Хотя образование  $\beta$ -листов пептидными частями наших молекул изначально предполагалось в нашей методологии, дальнейшее моделирование доказало, что фибриллы стабильны как в вакууме, так и в растворе. Эволюция фибрилл и сравнительный анализ организации пептидной и тиофеновой частей фибрилл ясно показали, что в структуре и морфологии фибрилл доминируют взаимодействия пептид-пептид, а не взаимодействия тиофен-тиофен. Поскольку самосборка  $\beta$ -нитей в  $\beta$ -листы по-прежнему является основной движущей силой самоорганизации наших гибридных молекул, мы утверждаем, что иерархический механизм самосборки последних может быть описан аналогично предложенному Аггели и др. для чистых пептидных фибрилл [588]. В подходе статьи [588] вводится иерархия взаимодействий, а также последовательная иерархия агрегатных структур с увеличением концентрации мономера (см. Рисунок 5.4). Предполагается, что следующее неравенство

$$\epsilon_{tape} \gg kT \gg \epsilon_{ribbon} \gg \epsilon_{fibril} \gg \epsilon_{fiber} \quad (5.2)$$

выполняется для соответствующих энергий, в расчете на один пептид внутри соответствующих структур по сравнению с пептидом внутри структуры предыдущего уровня.

В обозначениях, введенных Aggeli et al. в нашем моделировании исследовались только агрегаты типа лент и двойных лент. Из результатов моделирования можно предположить, что присоединение синтетических частей к пептидным частям вряд ли влияет на основные принципы организации на этих уровнях. Однако эффекты вполне вероятно будут проявляться на более высоких уровнях организации, когда филаменты должны взаимодействовать друг с другом.



В то же время на этих более низких уровнях самоорганизации введение синтетической части может изменить баланс между различными способами самосборки, например параллельная и антипараллельная организация нитей, укладка лент с разных сторон. Как было показано в нашем моделировании с растворителем, взаимодействие растворителя-фибриллы также может играть важную роль в выборе предпочтительного режима самосборки.

Следует иметь в виду, что уровень иерархии, достигаемый самосборкой мономеров, зависит от концентрации мономеров, поэтому при более низких концентрациях в эксперименте мы можем видеть только отдельные ленты, в то время как на высоких концентрациях – волокнах с высокой сложностью.

Электростатические и дипольные взаимодействия, по-видимому, также важны для определения паттерна самосборки фибрилл. Даже если концы пептида не заряжены, каждая  $\beta$ -цепь имеет небольшой дипольный момент, направленный от С-конца (частично отрицательный) к N-концу (частично положительный). Таким образом, энергетически выгодно, когда дипольные моменты соседних  $\beta$ -нитей антипараллельны, это может быть достигнуто либо путем использования антипараллельной организации  $\beta$ -листов, либо путем наложения вместе параллельных  $\beta$ -листов таким образом, чтобы дипольные моменты соседних листов компенсировали друг друга.

*$\pi$  –  $\pi$ -стэкинг по сравнению с водородными связями*

Что касается приложений, наиболее интересным аспектом представленной здесь работы является расположение конъюгированных тиофеновых фрагментов и взаимодействие между  $\pi$  –  $\pi$ -укладкой и образованием  $\beta$ -листа. Расстояние между пептидными цепями, обусловленное структурой  $\beta$ -листа, соответствует 4,7-5Å, в то время как расстояние между  $\pi$ -сопряженными плоскостями олиготиофенов для взаимодействий  $\pi$  –  $\pi$ -стэкинга оценивается методами квантовой химии и в идеале составляет примерно 3,3-3,5Å [636; 637]. Следовательно, эти взаимодействия могут демонстрировать конкурирующее поведение с точки зрения периодичности, подразумеваемой во время молекулярной агрегации, и это несоответствие периодичности должно быть как-то разрешено. Анализ гистограмм распределения расстояний между центрами масс пептидов, а также между кватротиофенами (рис. 5.19) ясно выявил преобладание периодичности, подразумеваемой для агрегатов за счет расположения пептидов  $\beta$ -листах для всех типов моделируемых фибрилл.

Сила и происхождение взаимодействий  $\pi - \pi$ -стэкинга в плоских ненасыщенных органических молекулах до сих пор остаются предметом споров. В [648] было показано, что для относительно небольших ароматических систем (менее 10-15 атомов углерода) стэкинг-взаимодействия могут быть описаны чисто дисперсионными взаимодействиями (силы ван-дер-Ваальса) и только для более крупных систем появляется определенный дополнительный энергетический вклад на малых расстояниях стэкинга, вступает в игру эффект электронной корреляции. В наших симуляциях мы были в пределах чисто ван-дер-ваальсова описания стэкинга, поскольку не могли быть включены эффекты электронной корреляции. Однако, поскольку кватротиофены являются относительно небольшими сопряженными молекулами, мы надеемся, что это хорошее приближение к реальности.

Чтобы понять качественное поведение кватротиофена только в стопке, мы провели моделирование минимизации энергии собранной стопки кватротиофенов с использованием того же силового поля PCFF, что и в моделировании фибрилл. Полученное межплоскостное расстояние тиофенов составило  $3,55\text{\AA}$ , что не противоречит расчетам квантовой химии. Кватертиофены, однако, также были смещены латерально по отношению к соседней молекуле на  $1,4\text{\AA}$  в направлении длинной оси (направление  $x$ ) и на  $1,8\text{\AA}$  в перпендикулярном направлении (направление  $y$ ).

Эти значения можно сравнить со значениями, определенными для расположения тиофена в фибрилле. Видно, что среднее расстояние между плоскостями тиофена (среднее смещение  $z$ ) все еще составляет около  $3,5\text{\AA}$ , что соответствует типичному расстоянию стэкинга  $\pi - \pi$ . Как этого произошло? В случае параллельного расположения-цепей (см. рис. 5.20б,г) тиофеновые фрагменты наклоняются синхронно, таким образом одновременно сохраняя высокую степень упорядоченности и выполняя ограничения периодичности, накладываемые  $\beta$ -листами. Однако это происходит за счет дополнительного бокового смещения по сравнению со значениями идеальной стопки. В случае антипараллельного расположения предлагаемая двухслойная фибрилла (рис. 5.20в) в принципе имеет такую же линейную плотность кватротиофенов вдоль фибриллы, как и в случае фибрилл с параллельным расположением, поскольку тиофеновые фрагменты обеих лент соприкасаются для образования “стерической молнии” вдоль каждой стороны двойной ленты (олиготиофеновые фрагменты с обеих сторон находятся в чередующемся контакте). Однако

при антипараллельном расположении части тиофена имеют тенденцию наклоняться друг к другу, образуя динамические кластеры с промежутками между отдельными кластерами; такое расположение уменьшает различия между оптимальными интертиофеновыми и межпептидными расстояниями. С другой стороны, из-за этой тенденции к формированию динамических кластеров степень дальнего порядка в расположении олиготиофена значительно ниже для фибрилл с антипараллельным расположением, чем для фибрилл с параллельными  $\beta$ -нитями.

Другой интересный пример влияния пептидной организации на упорядочение тиофенов - это влияние скручивания фибрилл на расстояние между кватротиофенами. Из рисунка 5.19г видно, что среднее расстояние между тиофенами в скрученной фибрилле даже больше, чем периодичность  $\beta$ -листов по чисто геометрическим причинам. Поскольку края скрученной ленты нужно растягивать больше, чем центральную часть.

#### *Кручение и изгиб фибрилл*

Осевое скручивание - это внутреннее свойство  $\beta$ -листных структур, возникающее в основном из-за хиральности аминокислот. Предполагается, что скручивание связано с чередующимися колебаниями двугранных углов, чтобы предотвратить расхождение отдельных  $\beta$ -нитей в более крупном листе [649], однако в литературе также предлагались другие вклады в угол закручивания, включая энтропию, связанную со степенями свободы остова [650], внеплоскостная деформация пептидных групп [651], внутрицепочечные [652], третичные взаимодействия [653; 654], тенденция к минимизации площади поверхности системы [655]. Однако точное понимание зависимости твиста от молекулярной организации все еще остается не вполне понятным. В пионерских исследованиях геометрии  $\beta$ -листов было предсказано большее скручивание для антипараллельных листов [656], однако современные данные предполагают, что проблема намного сложнее, например, граница раздела между двумя листами может значительно влиять на степень скручивания. и даже обратить тенденцию вспять [657]. Также утверждается, что экспериментально наблюдаемое скручивание в фибриллах может возникать не только из-за изменения симметрии упаковки, но также потенциально из-за кинетического захвата в состояниях, которые имеют твист, отличный от глобального минимума [657]. Периодическое кручение в классических амилоидных волокнах обычно наблюдается на масштабах в сотни нанометров [641].

В целом левая закрутка амилоидных фибрилл, построенных на основе левозакрученных *beta*-нитей, соответствует представлениям о часто наблюдающемся хиральном дуализме и чередовании хиральности при иерархическом структурообразовании биологических систем [658].

Как показало молекулярное моделирование, фибриллы, образованные нашими гибридными молекулами тиофен-пептид, также могут проявлять скрученное поведение. В случае фибриллы DL-PAR наблюдалось четкий период кручение в 60 нм (плоскость ленты поворачивается примерно на  $360^\circ$ ), в то время как для фибриллы DL-AP можно было бы предположить период кручения около 170 нм, хотя это сопровождается изгибом ленты. Эти значения находятся в диапазоне значений, обычно наблюдаемых в амилоидных фибриллах. Происхождение этого скручивания может быть дополнительно выяснено путем сравнения этих симуляций с отдельным исследованием симуляции, в котором рассматривалась только пептидная часть фибриллы. При сравнении окончательных структур фибрилл, представленных на рис. 5.23 и рис. 5.24, становится очевидным, что: (i) степень скручивания фибриллы DL-PAR не изменяется алкил-тиофеновыми фрагментами, (ii) пептидная фибрилла в случае DL-AP не скручена и не изогнута, в то время как добавление алкил-тиофеновых фрагментов вызывает изгиб фибриллы, который сопровождается осевым скручиванием, но гораздо меньшим, чем в случае фибриллы DL-PAR. Делать некоторые общие выводы из анализа этих двух случаев было бы преждевременным, но мы можем указать на некоторые возможные тенденции в изменении морфологии фибрилл из-за добавления синтетической части. Мы утверждаем, что если бы фибрилла имела четко выраженное скручивание, уже происходящее из взаимодействия ее пептидных фрагментов (с соответствующей упругой энергией, связанной со скручиванием), добавление алкил-тиофеновых фрагментов к химической структуре вряд ли изменило бы это поведение скручивания серьезным образом. Мы должны снова учитывать, что скручивание фибриллы не способствует плотной упаковке тиофеновых фрагментов, связанных с пептидами, однако ван-дер-ваальсово взаимодействие, ответственное за упаковку тиофена, не велико и не перевешивает потерю изгибной энергии пептидной части фибриллы, которая может быть вызвана раскручиванием фибриллы.

С другой стороны, если пептидная часть фибриллы не проявляет значительного предпочтения к скручиванию, факторы, связанные с упаковкой тиофенов, становятся более важными для определения общей формы фибрил-

лы. Поскольку даже небольшие взаимодействия могут возмущать структуру и вызывать ее скручивание-искривление, но в больших пространственных масштабах.

Еще один момент, который следует упомянуть, - это различие между однослойным и двухслойным расположением - кажется, что осевое скручивание происходит из тонкого баланса различных взаимодействий, поэтому данные об осевом скручивании однослойных фибрилл не могут быть экстраполированы на двухслойные. Это хорошо видно при сравнении их поведения в наших расчетах.

#### *Персистентная длина фибрилл*

В то время как для однослойных фибрилл трудно определить персистентную длину из компьютерного моделирования из-за высокой конформационной лабильности последних и, следовательно, сложного механизма гибкости, для двухслойных фибрилл могут быть применены идеи простого механизма персистентной гибкости. Значение 1,7 мкм было получено при моделировании фибриллы DL-PAR. К сожалению, точное измерение постоянной длины было невозможно в эксперименте, однако по изображениям фибрилл, полученным методом АСМ, можно предложить оценку в несколько микрометров, которая хорошо согласуется. В специальных исследованиях родственных им биологических амилоидных фибрилл были получены точные численные значения персистентной длины, которые составляют от 3 до 12 мкм в зависимости от типа фибриллы, что еще раз подтверждает наш вывод [659].

#### *Влияние адсорбции субстрата на фибриллы*

Одним из наиболее распространенных методов визуализации фибриллярных структур является атомно-силовая микроскопия, которая требует осаждения фибрилл на подложку, обычно мусковитовую слюду или высокоориентированный пиролитический графит. Следовательно, влияние взаимодействий субстрат-фибрилла на морфологию агрегатов важно для интерпретации экспериментальных результатов. Из результатов нашего моделирования мы можем предположить, что для двухслойных фибрилл силы поверхностной адсорбции достаточно велики, чтобы изменить скрученную морфологию фибрилл на плоскую и даже разорвать фибриллы в процессе адсорбции. Это ставит вопрос о том, можно ли экстраполировать данные о морфологии фибрилл, полученные методами АСМ или ПЭМ в сухом состоянии, на состояние фибрилл в растворе.

В Börner et al. [647] сравнили размеры и морфологию фибрилл из конъюгатов поли (этиленоксид) -пептид, полученных с помощью АСМ и ПЭМ в сухом состоянии, с данными, полученными с помощью просвечивающей электронной микроскопии с криофиксацией и малоуглового рассеяния нейтронов. Авторы предполагают, что АСМ и методы ПЭМ достаточно надежны, чтобы судить о морфологии фибрилл в растворе.

Принимая во внимание результаты нашего моделирования, мы считаем, что следует осторожно относиться к данному утверждению. По крайней мере, для определенных типов фибрилл взаимодействие с субстратом может вызывать изменения в его морфологии, такие как раскручивание, и поэтому следует проявлять осторожность при интерпретации результатов экспериментов.

По крайней мере, в фибрилле DL-PAR упругой энергии скручивания/раскручивания фибрилл было недостаточно, чтобы конкурировать с силами адсорбции. Однако можно предположить, что пока адсорбция будет происходить в растворе, силы адсорбции будут ниже, и скручивание может сохраниться. Было бы интересно исследовать это также с помощью АСМ в присутствии растворителя.

Но даже в сухом режиме АСМ амилоидные фибриллы во многих случаях демонстрируют характерное скрученное поведение на подложке. Можно предположить, что по мере дальнейшей агрегации лент в волокна и фибриллы (см. рис. 5.4) их скручивание становится более устойчивым к силам адсорбции.

#### *Проводимость гибридных фибрилл*

Проводимость возможных фибриллярных структур - один из ключевых интересных вопросов для технологического применения агрегатов тиофен-пептид.

С классической точки зрения два свойства нановолокон определенно благоприятствуют электропроводности: (i) плотная упаковка и расположение тиофеновых фрагментов и (ii) образование “перколяционных” кластеров вдоль проволоки, т.е. непрерывность тиофеновой стопки и наличие проводящего пути.

Сравнивая упорядочение тиофена в соответствии с этими принципами (см. рис. 5.20) (особенно в соответствии с последним критерием), расположение фибрилл DL-PAR можно считать лучшим среди других изученных, поскольку тиофеновые фрагменты в основном упорядочены и образуют почти непрерывную стопку, хотя и с некоторыми флуктуирующими дефектами (динамический беспорядок). В других типах укладок, как однослойных, так и двухслойных,

можно увидеть долгоживущие промежутки между кластерами тиофенов, которые определенно будут ухудшать проводящие свойства проводов, поскольку для того, чтобы произошла скачкообразное перемещение заряда, тиофены должны быть расположены близко друг к другу или, по крайней мере, находиться в динамическом взаимодействии. Например, ясно видно, что для фибриллы DL-AP в растворе тиофены наклоняются вокруг своей длинной оси, и весь путь проводимости разбивается на кластеры из 2-4 блоков кватротиофенов.

Упорядочение тиофенов при абсорбции волокон на подложке также может быть интересным для технологических приложений. Из нашего моделирования можно сделать два вывода: (i) адсорбция делает структуры и стопки тиофена более упорядоченными с меньшими динамическими колебаниями, (ii) но в то же время отсутствие динамических колебаний препятствует закрытию возможных образовавшихся промежутков в стопках тиофена при адсорбции.

Поскольку упорядочение тиофенов является ключевым вопросом для борьбы с динамическим и статическим беспорядком в системе, необходимым для электропроводности, дальнейшая агрегация лент в более крупные волокна (см. рисунок 5.4) может улучшить ситуацию.

## Структура наблюдаемых фибрилл

В этом разделе мы попытаемся предположить возможную структуру фибрилл, наблюдаемую в эксперименте. К сожалению, экспериментальные данные довольно ограничены, поэтому мы не сможем сделать окончательный вывод, однако мы попытаемся предложить правдоподобную модель, которая не будет противоречить ни одному из экспериментальных результатов или моделированию.

Как указывалось ранее и подтверждается данными для систем из пептидов и полимер-биоконъюгатов [629; 647], молекулярная агрегация в нашей системе, вероятно, регулируется рядом взаимодействий различной природы и силы, ведущих к общей иерархической модели самосборки для  $\beta$ -нитей пептидов: сначала молекулы собираются в  $\beta$ -листовые ленты, которые затем агрегируют латерально с образованием двойных лент (как в перекрестной

$\beta$ -структуре амилоидоподобных фибрилл), которые, в свою очередь, образуют более крупные структуры за счет различных типов агрегации.

Также кажется разумной гипотеза, что совокупности различных организационных уровней (одиночные ленты, двойные ленты, связки двойных лент) могут одновременно присутствовать в растворе в динамическом равновесии. Аргументом в пользу этой точки зрения может быть наблюдение различных морфологических типов фибрилл в АСМ (см. рисунок 5.8).

Судя по характерной высоте типично наблюдаемых структур (см. рисунок 5.8) и сопоставлению с расчетными данными, можно предположить, что самые короткие фибриллярные структуры, видимые на фоне АСМ-изображения, соответствуют одинарным  $\beta$ -листовым лентам (с высотой  $0,5 \text{ нм} \pm 0,2 \text{ нм}$ ), тогда как более крупные фибриллы (с высотой  $1 \text{ нм} \pm 0,2 \text{ нм}$ ) представляют собой пучки двойных лент (см. рисунок 5.34). Разница между экспериментальной и виртуальной высотой АСМ, определенной при моделировании (см. рис. 5.32), может быть связана с тем фактом, что фибрилла в эксперименте кажется внедренной в слой остаточного материала.

АСМ измерения наших фибриллярных агрегатов не выявили каких-либо периодических колебаний высоты вдоль фибрилл, также в ПЭМ, не было обнаружено никаких признаков спиральной суперструктуры волокон, экстраполяция этих данных на свойства раствора фибрилл может быть некорректной, согласно тому, что мы обсуждали в предыдущем разделе. Однако, если выбирать между организацией DL-AP и DL-PAR, можно выдвинуть аргумент в пользу варианта без скручивания: энергетически более выгодно создавать плоские пучки фибрилл (как наблюдается в АСМ) из изначально раскрученных структур, в противном случае - будет энергетических проигрыш в результате раскручивания фибрилл из их равновесного состояния.

Среди теоретических моделей, основанных на одинарных  $\beta$ -листовых лентах, обе модели продемонстрировали высокую конформационную гибкость и малую персистентную длину (по сравнению с моделями с двойной лентой), что согласуется с их поведением, наблюдаемым в АСМ (по сравнению с более крупными волокнами). Предполагается, что для небольших одиночных лент быстрая адсорбция компонентов на подложке (слоде в нашем случае) во время подготовки образца может заставить их принять плоскую конформацию. Ширина самых маленьких самоорганизующихся фибрилл, измеренная в АСМ, составляет  $7 \text{ нм} \pm 2 \text{ нм}$ , что также хорошо соответствует ширине фибрилл этого



типа, находящихся в плоской конформации. Тот факт, что такие более мелкие фибриллы не видны в ПЭМ, может снова указывать на предположение, что присутствие таких фибрилл в АСМ связано с кинетическими причинами во время подготовки образца (отдельные ленты “захватываются” на подложке за счет быстрой адсорбции) и более низкой концентрации мономеров.

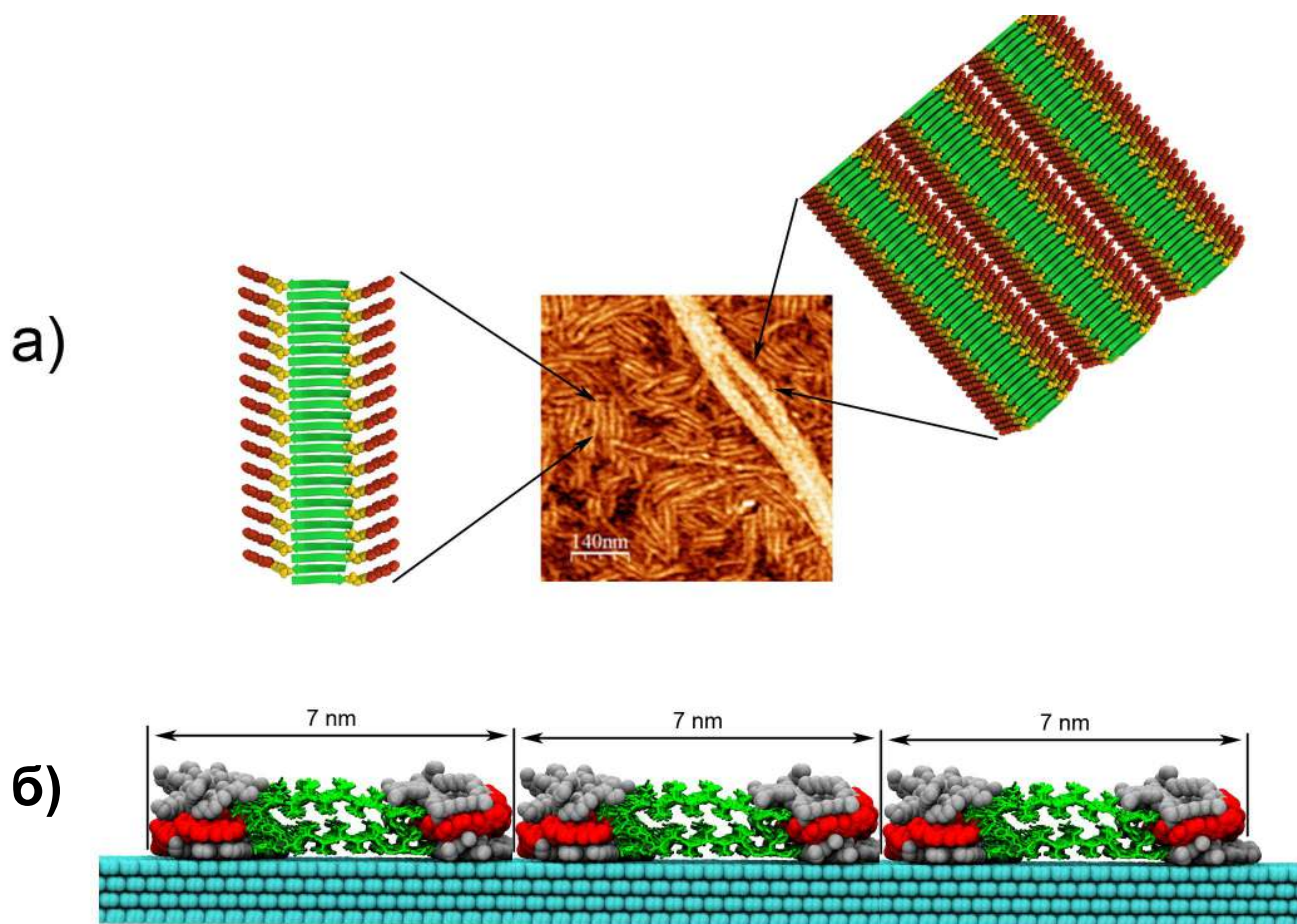
В полученных спектрах КД и флуоресценции не было признаков экситонной связи  $\pi$ -систем, которая должна происходить при  $\pi - \pi$ -стэкинге в хиральной среде на расстояниях, близких к идеальным для  $\pi - \pi$ -стэкинга  $\pi$ -сопряженных структур. Среди теоретически изученных моделей фибрилла типа SL-AP наилучшим образом подтверждает эти экспериментальные данные, поскольку имеет наименьшее количество тиофеновых колец, находящихся в тесном контакте друг с другом, по сравнению с другими моделями. Более того, в предлагаемом типе компоновки ранее не учитываемые нами гибкие ПЭГ-хвосты молекул, которые расположены в соседних  $\beta$ -листах, могут дополнительно мешать упорядоченному расположению тиофеновых фрагментов в соседнем  $\beta$ -листе и, таким образом, препятствовать прямому взаимодействию  $\pi$ -систем. Можно даже предположить, что цепи ПЭГ взаимодействуют с энергетическими уровнями  $\pi$ -систем [610]. С другой стороны, взаимодействие  $\pi$ -систем тиофеновых фрагментов на расстояниях, больших, чем те, которые характерны для идеального  $\pi - \pi$ -стэкинга, будет соответствовать предполагаемым из эксперимента взаимодействиям на больших расстояниях. Такие “дальнодействующие” взаимодействия могут осуществляться в молекулярных ансамблях, подобных агрегатам J-типа, с расположением гибридных молекул на основе двойных слоев  $\beta$ -листов.

Для агрегирования лент SL-AP в двухслойные ленты модель DL-AP (где боковые цепи валина скрыты между листами), подтверждается моделированием ее энергетической выгоды с точки зрения количества образующихся водородных связей.

Суммируя вышеупомянутые аргументы, можно предложить следующую возможную модель структуры агрегатов, наблюдаемую на изображениях АСМ (см. рисунок 5.34а): гибридные молекулы объединяются в растворе в ленты на основе антипараллельных  $\beta$ -листов, эти ленты затем объединяются с образованием двухслойных структур за счет взаимодействия их соответствующих гидрофобных поверхностей. Предполагается, что эти изображенные ленты в дальнейшем собираются в пучки за счет латеральных взаимодействий (см. рис.

5.34а,б). Таким образом формируется периодическая картина на изображениях АСМ с периодичностью приблизительно 6-7 нм и меньшими колебаниями по высоте.

Мы также должны сделать примечание о ПЭГ хвостах, которые мы не учли в наших теоретических расчетах ради простоты и предполагая, что они не играют основную структурную определяющую роль для отдельных лент и двойных лент. Однако, когда мы говорим об агрегатах более высокого порядка, таких как пучки фибрилл, влияние ПЭГ-хвостов на организацию может быть важным. В модели, представленной на рисунке 5.34, мы предполагаем, что хвосты ПЭГ достаточно гибкие, чтобы адаптироваться к структуре изображенного расположения.



**Рисунок 5.34** — Предложенная схематическая модель вероятного расположения гибридных молекул в наблюдаемых нанофибриллах и пучке, состоящем из более мелких фибрилл, образованных молекулярными гибридами тиофен-пептид. (а) Изображение АСМ и предполагаемая структура более мелких фибрилл и пучков двухслойных фибрилл.  $\beta$ -нити обозначены зелеными стрелками; кватротиофены изображены сферами ван-дер-Ваальса красного цвета. Алкильные цепи и ПЭГ не показаны. (б) Другой вид модели наблюдаемых пучков. Модельный субстрат изображен голубым цветом, кватротиофены - красным, алкильные цепи - серым, пептидные части - зелеными трубками, азидогруппа, связывающая тиофен с пептидом, и ПЭГ-хвосты не изображены.

### Дальнейший дизайн и экспериментальные стратегии

На основании результатов моделирования и анализа экспериментальных данных можно посоветовать некоторые идеи для дальнейших эксперимен-

тальных исследований, а также возможную настройку химической структуры молекул.

Дальнейшие экспериментальные исследования полученных агрегатов с помощью более совершенных методов, несомненно, будут полезны для более точного определения моделей их межмолекулярной агрегации. Могут быть применены следующие методы: (i) дифракция рентгеновских лучей на фибриллах - которая поможет определить периодичность между  $\beta$ -нитями и между стопками  $\beta$ -листов, (ii) другие методы подготовки проб АСМ (возможно, с помощью центрифугирования из более разбавленного раствора), чтобы визуализировать действительно отдельные волокна на подложке без остаточного материала, (iii) окрашивание красителем Конго красным и тиофлавином В может дать интересную информацию о структуре и кинетике самосборки фибрилл, (iv) крио-ПЭМ может быть также методом выбора для визуализации фибрилл в “квази-растворе”, (v) АСМ с измерением проводимости может пролить свет на проводящие свойства фибрилл, (vi) АСМ в растворе, если возможно, может предоставить информацию о прочности и влиянии субстрата на морфологию фибрилл.

Что касается улучшения гибридных молекул тиофен-пептид в отношении потенциального использования в качестве проводящих/полупроводниковых нанопроводов, можно предложить следующие идеи. Из измерений АСМ можно увидеть, что присутствуют как более мелкие, так и большие фибриллы (пучки), это указывает на тот факт, что одна фибриллярная конформация не имеет значительных энергетических предпочтений по сравнению с другой конформацией, но разные типы агрегатов находятся в динамическом равновесии. Другой факт - полученные фибриллы кажутся не очень стабильными и не очень упорядоченными на больших расстояниях. Следовательно, можно предположить, что могут быть внесены некоторые модификации в структуру гибридных молекул, которые будут усиливать фибриллярную конформацию агрегатов и усиливать паттерны фибриллярной самосборки. Это может быть сделано, например, путем введения зарядов на двух концах  $\beta$ -нитей, которые будут обеспечивать самосборку из-за электростатических взаимодействий, в то же время это обеспечит антипараллельную схему самосборки. В качестве альтернативы можно исследовать различные пептидные последовательности и топологии. В том числе Schillinger et al. [610] уже было показано, что трехблочный сополимер, в котором кватротиофен присоединен к двум олигопептидам на каждом конце, образует хорошо упорядоченные фибриллы.

Ассоциация молекул через параллельные  $\beta$ -листы, вероятно, будет производить более упорядоченную организацию тиофена, чем антипараллельное расположение, поэтому условия для такой сборки должны быть исследованы.

Формирование фибриллярных агрегатов высокого порядка посредством наложения и переплетения лент также может быть перспективной стратегией, поскольку добавленное упорядочение структуры снова будет способствовать улучшению электронных и проводящих свойств.

### 5.2.8 Выводы раздела

В этом разделе мы представили комплексное моделирование фибриллярных агрегатов из диблочных олигомеров кватротиофен-пептид на основе различных предполагаемых межмолекулярных структур, вдохновленных структурами родственных амилоидных фибрилл. Мы разработали мультимасштабную методологию, основанную на предсказании периодического расположения, моделировании фибрилл в явном и неявном растворителе, а также на подложке.

Выявлены и обсуждены взаимосвязи межмолекулярной организации агрегатов, супрамолекулярной морфологии фибрилл, а также их свойств.

Результаты моделирования сравнивались с имеющимися экспериментальными данными, и была предложена вероятная межмолекулярная структура наблюдаемых фибрилл. Были предложены дальнейшие экспериментальные и дизайнерские стратегии для создания проводящих нанопроволок из тиофен-пептидных олигомеров.

## 5.3 Моделирование пептидных фибрилл EF-C, взаимодействующих с ретровирусами

*Результаты данного раздела следуют публикации [A31].* В данном разделе приведены результаты интегративного моделирования амилоидоподобных положительно заряженных пептидных фибрилл, которые нашли свое приме-

нения в качестве средств ускоряющих вирусную трансдукцию генетического материала в клетки. Используют методы и подходы интегративного моделирования разработанные и описанные, в том числе, в предыдущем разделе.

### 5.3.1 Описание экспериментальных результатов

*Описание следует публикации [A31].*

#### **Пептидные нанофибриллы EF-C ускоряют ретровирусную трансдукцию генов и могут быть использованы для быстрой концентрации вирусов**

Неэффективный перенос генов и низкие концентрации вирионов это основные ограничения ретровирусной трансдукции [660]. Мы и другие ранее показали, что пептиды, полученные из человеческой спермы образуют амиллоидные фибриллы, которые ускоряют доставку генов ретровирусным путем, стимулируя прикрепления вирионов к клеткам-мишеням [661–667]. Однако применение этих природных фибриллообразующих пептидов ограничено умеренной эффективностью, высокой стоимостью синтез пептида и вариабельностью размера фибрилл и кинетики их образования. Здесь мы сообщаем о разработке нанофибрилл, которые самоорганизуются в водном растворе из пептида длиной в 12 аминокислотных остатков, которые называем ускоряющим фактором С (enhancing factor C, EF-C). Эти искусственные нанофибриллы значительно более эффективно усиливают процесс ретровирусного переноса генов, чем фибриллы полученные из спермы или другие усилители трансдукции. Более того, нанофибриллы EF-C позволяют концентрировать ретровирусные переносчиками путем обычного низкоскоростного центрифугирования и являются безопасными и эффективными по оценкам в исследованиях *ex vivo*. Наши результаты показывают, что фибриллы EF-C представляют собой очень универсальный, удобный и широко применимый наноматериал, который имеет

потенциал значительно облегчить процесс ретровирусного переноса генов в фундаментальных исследованиях и клиническом применении.

Открытие принципов, лежащих в основе конструкции нанофибрилл для ретровирусной трансдукции произошло из исследований, направленных на более точное определение взаимодействия вируса иммунодефицита человека типа 1 (ВИЧ-1) с его клеточными рецепторами. Во время этого исследования мы проанализировали влияние на вирусную инфекционность различных пептидов, полученных из гликопротеина gp120 ВИЧ-1. Мы нашли, что один пептид, EF-A (соответствует аминокислотам 413–431 gp120) приводит к увеличению инфицирования до 34 раз (рис. 5.35а,б). Для определения минимальной пептидной последовательности необходимой для этого эффекта, мы исследовали варианты различной длины (от В до Н) EF-A и пептид EFcon, полученный из консенсусной последовательности gp120 (рис. 5.35а). Обработка клеток всеми пептидами, содержащими центральную область EF-A (остатки 419–427) способствовала инфицированию ВИЧ-1 (рис. 5.35в), не вызывая цитотоксических эффектов. На основе этого скрининга была идентифицирована небольшая амфифильная 12-аминокислотная последовательность пептида EF-C с последовательностью QSKIKQIINMWQ (1,533 Да; pI=9,31, заряд =+2). EF-C стимулировал вирусную инфекцию с половинной максимальной активностью ( $EC_{50}$ )  $\sim$  11мг/мл, которая примерно в четыре раза сильнее, чем у встречающихся в природе фибрилл SEVI, полученных из спермы, (42 мг/мл; Рис. 5.35в,г). Предыдущие исследования показали, что образование фибрилл является необходимым условием для того, чтобы пептиды, полученные из спермы, могли способствовать ретровирусной инфекции [661; 666; 667]. Мы обнаружили, что все активные формы gp120 фрагменты (EF-A,-C,-D,-F и con) продемонстрировали амилоид-специфическое увеличение связывание Конго красного (рис. 5.35д) и тиофлавина Т (ThT, рис. 5.35е) связывание. Кроме того, интенсивность флуоресценции значительно коррелировала со способностью пептидов gp120 увеличивать вирусное инфицирование. Таким образом, образование фибрилл имеет решающее значение для способности фрагментов, полученных из gp120, стимулировать ретровирусную доставку генов. Примечательно, что EF-C всегда самособирается в нанофибриллы сразу после разбавления его раствора в диметилсульфоксиде (ДМСО) в различных буферах. Это отличается от сборки фибрилл SEVI, поскольку они обычно образуются только после нескольких часов или дней перемешивания синтезированного пептида [661–663; 665]. На-

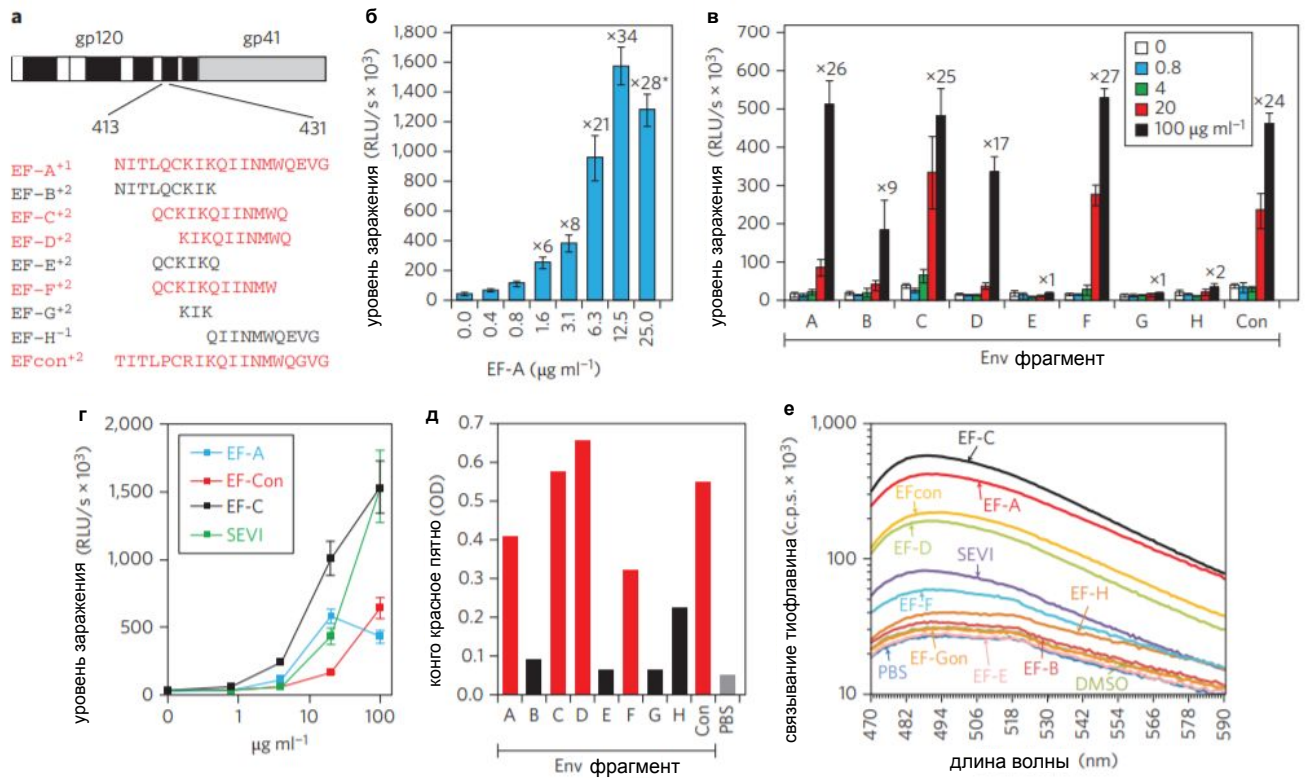
нофибриллы EF-C могут храниться не менее двух недель при комнатной температуре или  $4^{\circ}\text{C}$  без потери активности по усилению трансдукции.

Чтобы охарактеризовать нанофибриллы, мы проанализировали структуру и морфологию EF-C, наиболее эффективного усилителя трансдукции. С помощью атомно-силовой микроскопии (АСМ) мы обнаружили фибриллярные агрегаты диаметром  $3,4 \pm 1,1$  нм, длиной от  $\sim 100$  до 400 нм (рис. 5.36а, 5.37), а кручение с полушагам кручения в  $26 \pm 2$  нм (рис. 5.36б). Статистический анализ изображений более 50 отдельных нанофибрилл выявил, что персистентная длина находилась в диапазоне 0,5–2,9 мкм, а модуль Юнга (E) в диапазон 0,2–9,2 ГПа. Эти значения соответствуют тем, что наблюдаются для других амилоидных систем, и это означает, что фибриллы EF-C относительно жесткие по сравнению с другими биофиламентами, такими как актин (1-2 ГПа) или органические макромолекулы такие, как полимерные мицеллы ( $\sim 0,4$  ГПа) [668; 669]. Дальнейший биофизический анализ с использованием кругового дихроизма (КД) и инфракрасной спектроскопии с преобразованием Фурье (ИКФС) указали на присутствие преимущественно элементов вторичной структуры  $\beta$ -слоя и двух пиков в полосе амида-I, характерных для антипараллельного расположения  $\beta$ -листов [670] (рис. 5.36в). Порошковая дифракция рентгеновских лучей подтвердила типичный амилоидный паттерн дифракции с рефlekсами при  $4,8\text{\AA}$  и  $12\text{\AA}$ . Для определения молекулярной структуры фибрилл EF-C экспериментально полученные структурные параметры были объединены с молекулярным моделированием [A34] (рис. 5.36г – е). Полученная модель фибрилл предполагает, что боковые цепи лизина образуют гидрофильную поверхность с высокой плотностью катионных зарядов при физиологическом pH (рис. 5.36е). Измерения дзета-потенциала согласуются с этим анализом и показали что фибриллы EF-C имеют чистый потенциал  $+17,7 \pm 1,7$  мВ, аналогично с SEVI ( $+17,8 \pm 1,1$  мВ).

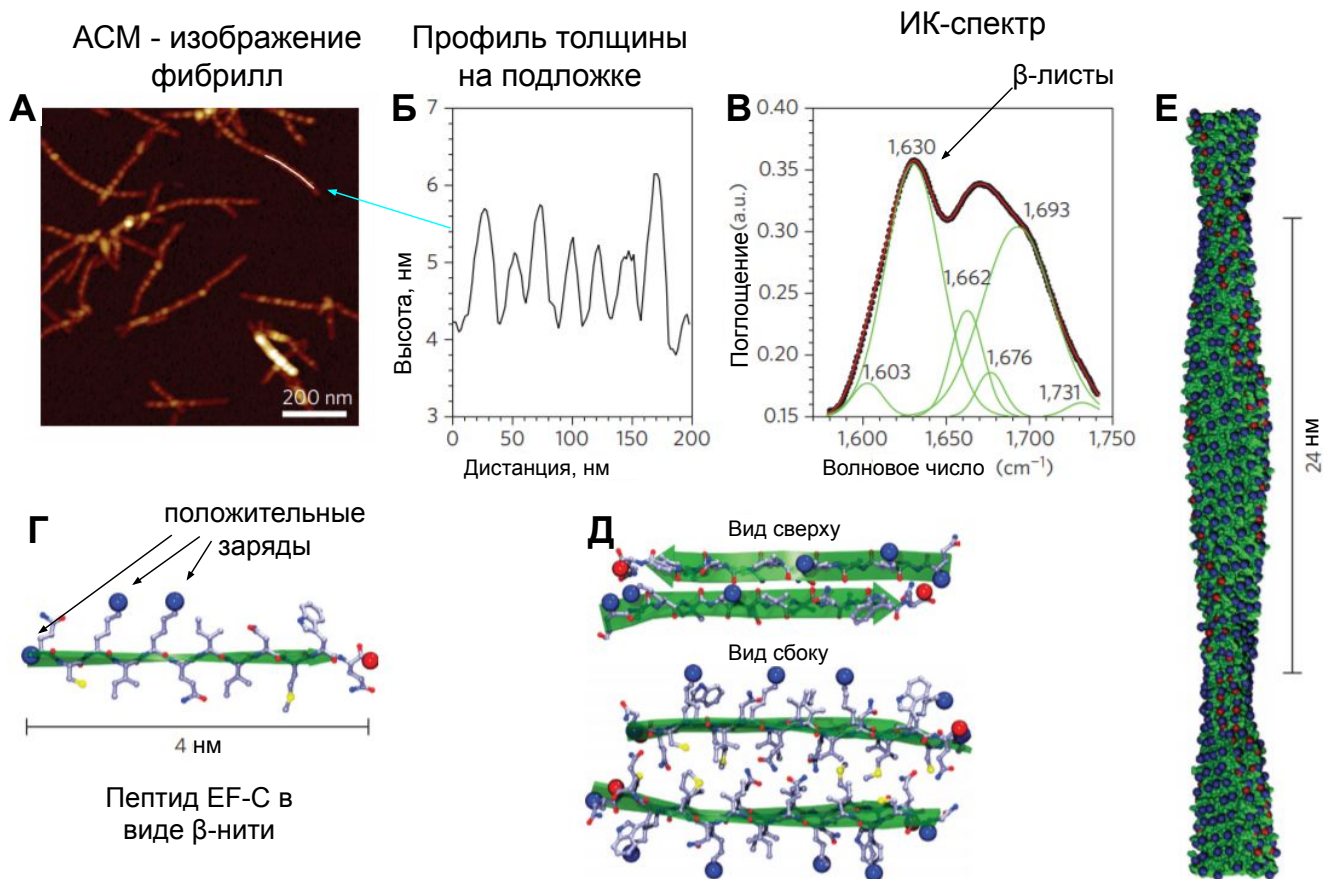
Чтобы визуализировать взаимодействие фибрилл EF-C с вирионами, нанофибриллы, меченные родамином (которые усиливают вирусную инфекцию так же эффективно, как немеченные фибриллы) были смешаны с ретровирусными частицами, меченные желтым флуоресцентным белком (YFP). Поразительно, что все вирионы связались с фибриллами EF-C, за секунды. Это эффективное образование макроскопических комплексов из вирионов и фибрилл EF-C побудили нас проверить, могут ли фибриллы использоваться для осаждения и концентрирования вирионов в настольной центрифуге. Мы обнаружили, что



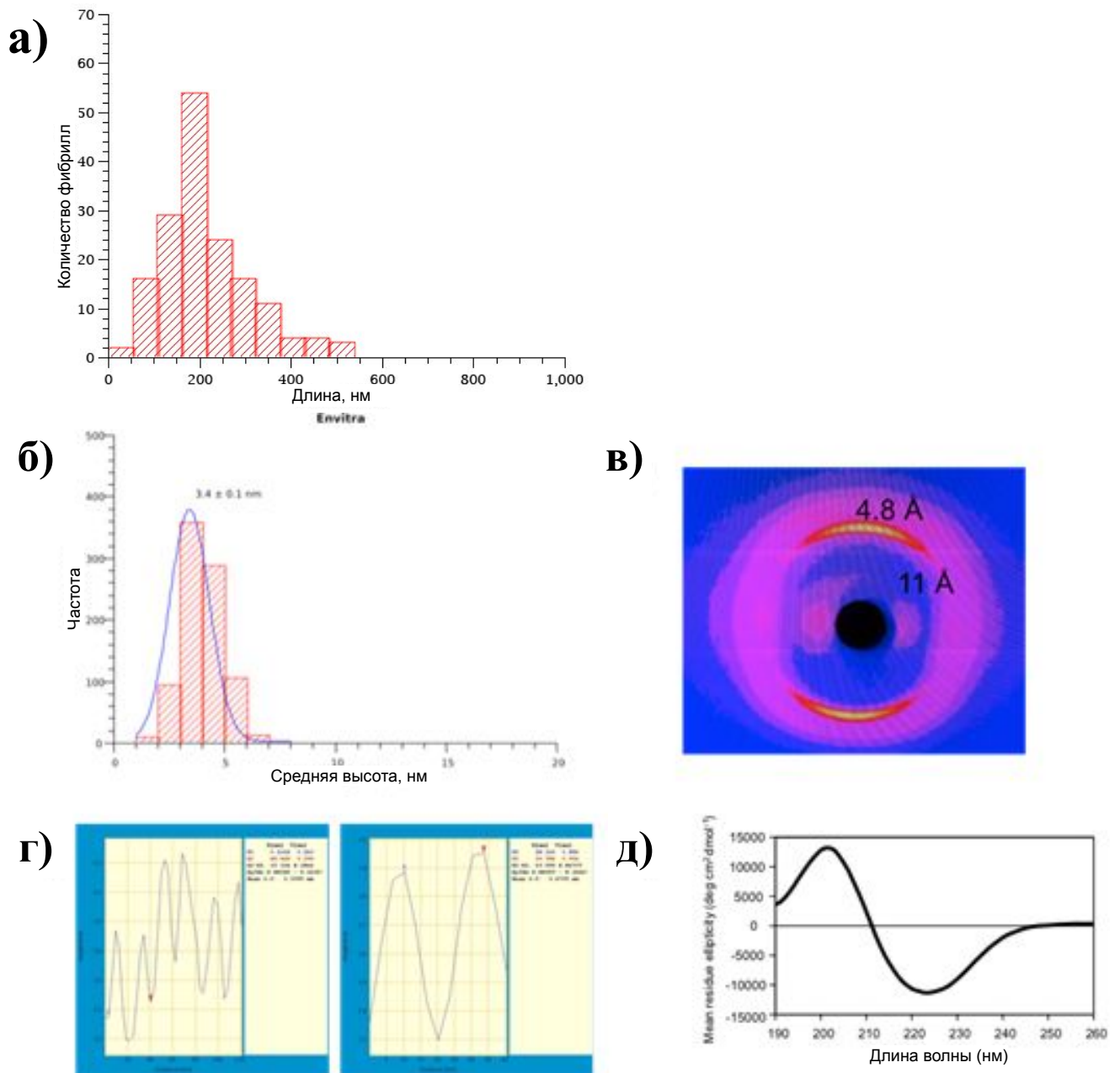
центрифугирование ВИЧ-1 или ретровирусных векторов с фибриллами EF-C в течение 5 мин при 10 000 g осаждали практически все вирионы. Удаление супернатанта (содержащего использованную среду) с последующим ресуспендированием вирусного осадка в желаемом объеме свежей среды, позволили сконцентрировать вирионы и достиг высоких скоростей трансдукции). Примечательно, что этот подход на основе нанофибрилл значительно быстрее и удобнее, чем длительное (60–90 мин) ультрацентрифугирование (50 000 g при 4°C), которое обычно используется для концентрирования ретровирусных частиц [671].



**Рисунок 5.35** — Пептиды, полученные из гликопротеина gp120 ВИЧ-1, усиливают вирусную инфекцию. **а**, Обзор проанализированных фрагментов gp120. Активные пептиды показаны красным. Указана локализация в предшественнике белке Env ВИЧ-1, а числа соответствуют положениям аминокислот в последовательности gp120 ВИЧ-1 HxB2. Цифры в верхнем индексе указывают на чистый заряд пептидов. **б**, пептид EF-A усиливает инфекцию ВИЧ-1. Вирус обрабатывали пептидом и смесь была использована для заражения клеток TZM-bl. Цифры над столбиками указывают на n-кратное усиление инфекции по сравнению с контролем без EF-A. RLU/с, относительные световые единицы в секунду; звездочка, чрезмерное заражение. **в**, активность синтетических фрагментов gp120 по стимулированию инфицирования ВИЧ-1, описанная в **а,г**, EF-C усиливает ВИЧ-1 заражение более эффективно, чем SEVI. **д,е**, окрашивание Конго красным **д** и флуоресценция ThT **е** тестируемых пептидов. OD - оптическая плотность; с.р.с., отсчеты в секунду.



**Рисунок 5.36** — Структурная характеристика и молекулярное моделирование фибрилл EF-C. **а** - АСМ-изображение фибрилл EF-C. **б**, профиль по указанной линии в **а**. Автокорреляция графика профиля дает средний полушаг фибрилл в  $26 \pm 2$  нм. **в**, ИКФС-спектр с вписыванием гауссианов для определения вторичной структуры. Пики на  $1630^{-1}$  и  $1693^{-1}$  указывают на антипараллельное расположение  $\beta$ -листов. **г** - Молекулярная модель пептида EF-C. **д**, вид сверху и сбоку элементарной единицы фибриллы, состоящей из четырех  $\beta$ -нитей, собранных в стопку из двух антипараллельных  $\beta$ -листов. **е**, Уточненная молекулярная модель фибриллы с шагом спирали 28 нм. Атомы С, серый; N, синий; O, красный; S, желтый.



**Рисунок 5.37** — Дополнительная структурная характеристика фибрилл EF-C. а и б) Гистограммы распределения длин и высот АСМ . в) Рентгеновская дифрактограмма, показывающая рефлексы, характерные для амилоидных фибрилл, при 4,8 и 11Å. г) График профиля АСМ вдоль фибриллы (слева) и соответствующая автокорреляция, показывающая половину шага спирали. д) Круговой дихроизм фибрилл EF-C в фосфатном буфере.

Для дальнейшего выяснения механизма, лежащего в основе опосредованного фибриллами Для повышения инфекционности мы провели конфокальный микроскопический анализ комплексов вирион/фибриллы в присутствии клеток. В результаты показывают, что фибриллы EF-C значительно увеличивают количество вирусных частиц, прикрепленных к поверхности клетки. Совместная

локализация с клетками указывает на стабильное взаимодействие комплексов вирион/фибриллы с клеточной мембраной. Это взаимодействие может возникнуть из-за нейтрализации отталкивания между отрицательно заряженными вирусными и клеточными мембранами в сочетании с привлекательными взаимодействиями избыточных положительных зарядов на фибриллах с поверхностью клетки. Это в конечном итоге приводит к усиленному слиянию вирусной и клеточной мембран.

В настоящее время наиболее эффективный и распространенный метод ускорения переноса ретровирусного гена основан на производном фибронектина - RetroNectin [672—674]. RetroNectin необходимо предварительно нанести на пластины для совместной локализации вирионов и клеток, что в конечном итоге приводит к увеличению поглощения вируса клетками. Чтобы проверить, могут ли нанофибриллы работать аналогично, слайды покрывали EF-C фибриллами и инкубировали с ретровирусными частицами. Конфокальная микроскопия показала, что фибриллы были иммобилизованы на поверхности и эффективно захватывали вирионы в зависимости от времени и дозы. Как и ожидалось, иммобилизованные фибриллы значительно увеличивают лентивирусную трансдукцию и этот эффект практически не зависел от используемых пластин. Мы сравнили эффекты покрытий RetroNectin и EF-C и обнаружили, что оба усиливают ретровирусный перенос с аналогичным ускорением. Комбинация покрытия RetroNectin и фибриллами EF-C в растворе приводили к еще более высоким уровням переноса генов. Как и следовало ожидать, обработка вирионов или клеток EF-C в растворе увеличивало доставку генов ретровирусами, в то время как RetroNectin такого эффекта не имел.

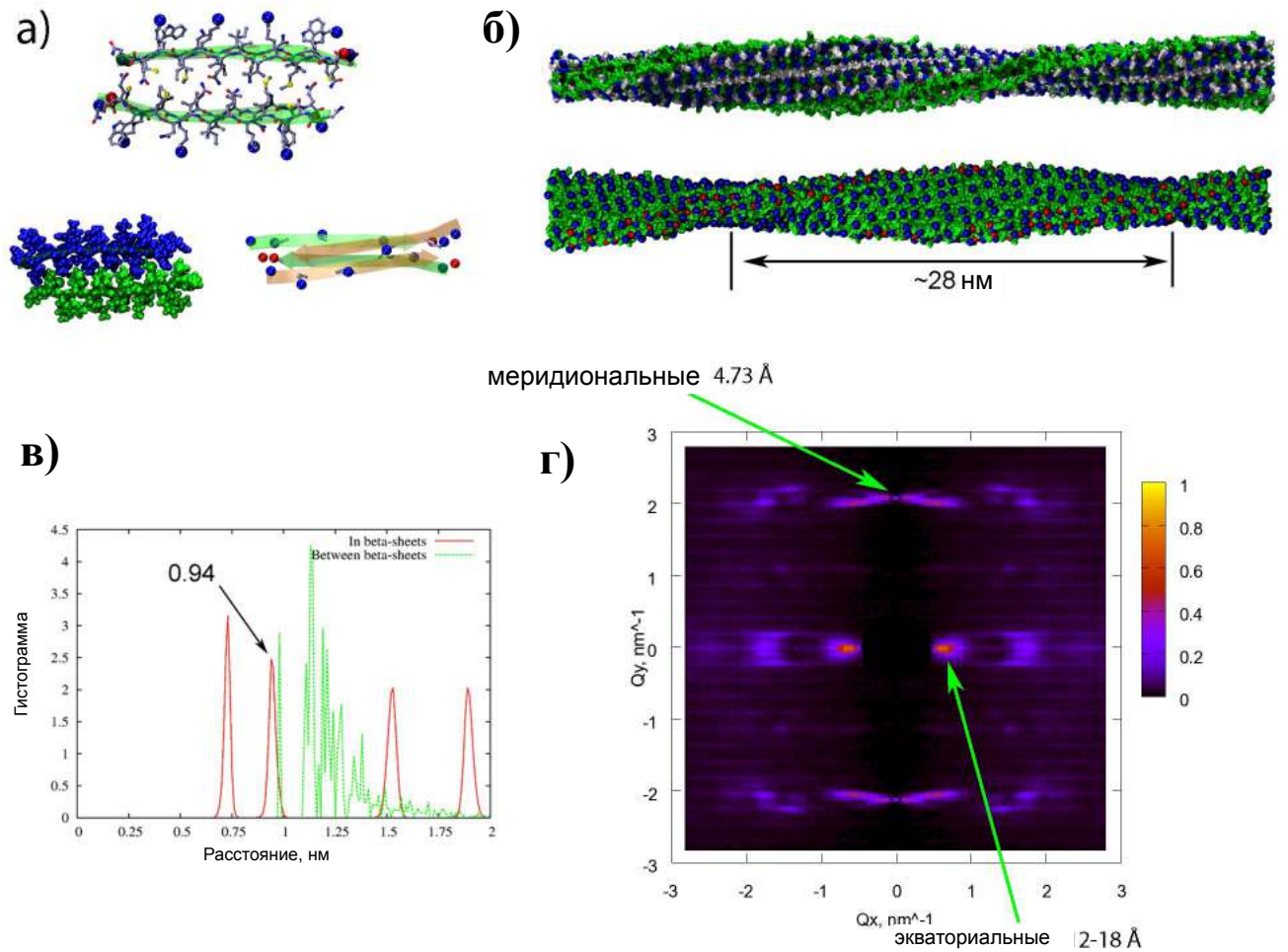
### 5.3.2 Детали моделирования

### 5.3.3 Моделирование молекулярной динамики

Чтобы построить молекулярную модель амилоидной фибриллы, был использован мультимасштабный комбинированный экспериментально-теоретический интегративный подход, описанный в других работах [A34; A35]. На

основе полностью атомной молекулярно-механической модели пептида EF-C, экспериментальных данных, основных физических принципов, а также понимания принципов образования амилоидных фибрилл была предложена базовая модель межмолекулярного расположения пептидов в фибрилле. Полностью атомная молекулярно-механическая модель этого межмолекулярного расположения была построена следующим образом: (i) пептид был взят в вытянутой конформации с соответствующими двугранными углами, (ii) два пептида были расположены в виде димера двух антипараллельных бета-цепей, так чтобы максимизировать количество водородных связей между нитями, (iii) два димера были расположены в тетрамере путем наложения “незаряженных” поверхностей бета-листов, и тетрамер был помещен в коробку для моделирования с периодом  $10\text{\AA}$  вдоль направления фибриллы, чтобы сформировать периодическую фибриллярную структуру, (iv) затем была использована модель молекулярной механики, основанная на силовом поле AMBER99SB-ILDN [675], чтобы выполнить минимизацию энергии в фиксированных периодических граничных условиях. На следующем этапе была создана фибриллярная структура длиной 50 бета-нитей (примерно 25 нм) путем воспроизведения периодической ячейки межмолекулярного расположения в одном направлении. Фибрилла сольватирована молекулами воды (модель TIP3P), добавлено 100 ионов хлора для нейтрализации системы. Общее количество атомов в системе составило около 296 тысяч с размерами ящика периодического моделирования  $11 \times 9 \times 30$  нм. Система была подвергнута минимизации энергии с последующим моделированием релаксации с помощью МД с упругими ограничениями на положения атомов фибриллы, чтобы позволить положениям молекул ионов и воды отрелаксировать. После этого было проведено моделирование равновесной МД в течение 20 нс при давлении и температуре 1 атм и 300 К соответственно с использованием алгоритмов перемасштабирования скорости и барастата Парринелло-Рамана. Электростатические взаимодействия на больших расстояниях обрабатывались с помощью алгоритма PME, для короткодействующих взаимодействий использовалось отсечение 1 нм. Моделирование проводилось параллельно с использованием пакета программ GROMACS 4.5 [676]. В ходе моделирования фибриллярная структура приняла свою благоприятную скрученную конформацию и была стабильной как на межмолекулярном, так и на морфологическом уровне. Структура фибриллы была дополнительно визуализирована и проанализирована с использованием VMD [305] и собственных

скриптов. Смоделированная дифракционная картина фибрилл была построена с использованием программы SFALL из пакета CCP4 [24] и специально написанного программного обеспечения для выполнения цилиндрического усреднения структурного фактора. Дополнительно на рис. 5.38 представлена структура фибрилл и результаты анализа.



**Рисунок 5.38** — Дополнительные данные о модели фибриллы из пептидов EF-C. а) Молекулярные модели димеров пептидов и элементарной единицы фибриллы, обрaмленной четырьмя  $\beta$ -нитями, собранными в стопку из двух антипараллельных  $\beta$ -листов. б) уточненная и проанализированная модель фибриллы с использованием обширного моделирования молекулярной динамики, фибриллы представлены с использованием их доступной для растворителя поверхности: на верхнем изображении остатки окрашены в соответствии с их гидрофобностью, на нижнем - изображены положительно и отрицательно заряженные участки синим и красным, соответственно. в) распределение расстояний между центрами масс пептидов в фибрилле, видны четкие пики, показывающие одномерную периодичность фибриллы. г) рассчитанная 2D картина дифракции фибрилл, представленная как функция компонентов вектора рассеяния вдоль и перпендикулярно оси фибрилл, можно видеть четкое указание меридиональных и экваториальных рефлексов, согласующихся с экспериментальными данными (см. рис. 5.37в).



## 5.4 Выводы главы 5

Разработаны подходы интегративного моделирования амилоидоподобных фибрилл, позволяющие реконструировать укладку пептидов и морфологию фибрилл на основе сочетания экспериментальных данных ИК- и КД-спектроскопии, рентгеновской дифракции, электронной и атомно-силовой микроскопии.

## Заключение

### Обсуждение результатов

Построение и анализ структурных моделей биомакромолекулярных комплексов имеет ключевое значение в изучении структурной организации биологических систем. В реализации этого подхода принципиальную роль играют методы компьютерного молекулярного моделирования. В данной работе рассмотрены возможности методов молекулярного моделирования в построении и анализе больших ДНК-белковых комплексов и амилодоподобных фибрилл. В результате были разработаны новые интегративные методы моделирования для больших комплексов, для изучения свойств которых недостаточно стандартных.

Одним из наиболее широко распространенных и развитых методов молекулярного моделирования является метод молекулярной динамики (МД). Физические модели молекул в методе МД основаны на рассмотрении взаимодействий между атомами в приближении классической механики Ньютона. Согласно этому для биомолекул по некоторым правилам задается эмпирическая функция зависимости потенциальной энергии системы от координат атомов (так называемое силовое поле). Теоретически, при идеальной точности задания силового поля и неограниченных вычислительных возможностях метод МД может работать в режиме “вычислительного микроскопа”, который воспроизводит реальные процессы структурообразования и взаимодействия биологических молекул. Однако на практике мы имеем дело лишь с некоторым приближенным описанием взаимодействий, построенным по аддитивному принципу. Кроме того мы ограничены во временах моделирования, что делает возможным исследование молекулярных систем лишь в ограниченной области конформационного пространства вблизи их стартовых конформаций. Современное развитие методов МД обычно представляет собой повторяющийся двухстадийный процесс, в котором вслед за совершенствованием вычислительных возможностей следует совершенствование силовых полей. Характерным примером здесь является совершенствование моделей ДНК. Так, силовые поля класса AMBER были изначально параметризованы для воспроизведения экспериментальной геомет-

рии ДНК в наносекундном диапазоне, однако, в середине 2000-ых, оказалось, что при временах моделирования свыше 10 нс изменения торсионных углов остова ДНК приводят к неправильной геометрии ДНК, и поэтому силовое поле было усовершенствовано. Были изменены параметры силового поля, отвечающие за энергетические термы вращения торсионных углов в ДНК, что привело к сохранению правильной В-формы ДНК на доступных временах моделирования. В середине 2010-ых ситуация повторилась, когда стандартные времена моделирования увеличились до микросекундного масштаба, и силовое поле было вновь усовершенствовано. Аналогичные примеры можно привести и для силовых полей белков, воды и ионов. Кроме правильного воспроизведения равновесных параметров молекулярных систем, отдельным вопросом является качество воспроизведения их динамики. Здесь с точки зрения силового поля важным является воспроизведение правильного энергетического баланса (заселенности) между различными конформациями. Мы показали [677], что на больших временах моделирования начинают проявляться энтропийные эффекты, которые в случае белков приводят к излишней стабилизации структурных моделей. На практике при исследовании динамических процессов в биомакромолекулярных комплексах методами МД моделирования, необходимо оценить с одной стороны, в каких временных диапазонах МД модель адекватно описывает систему, а с другой стороны, в каких временных диапазонах возможно компьютерное воспроизведение функционально интересных конформационных перестроек. Результаты нашего моделирования на примере нуклеосом показывают, что современные методы МД позволяют создавать структурно-динамические модели ДНК-белковых комплексов без как-либо видимых артефактов по крайней мере на временах в десятки микросекунд. Свободная ДНК сохраняет конформацию, свойственную В-форме ДНК, ионное окружение является стабильным, глобулярные домены белка являются стабильными, а неупорядоченные домены сохраняют повышенную конформационную подвижность. В тоже время проведенный нами анализ динамики показал, что на временном диапазоне в пределах одной микросекунды наблюдаются такие процессы, как флуктуации формы свободных участков ДНК, конденсация неупорядоченных белковых доменов по поверхности ДНК, появления локальных конформационных возмущений в участках ДНК, связанных с белком. На временных масштабах порядка десятков микросекунд можно уже изучать диссоциацию/реассоциацию ДНК от белка в отдельных сайтах связывания, частичную диффузию неупорядоченных

доменов вдоль поверхности ДНК. В то же время, более крупномасштабные перестройки ДНК-белковых комплексов пока находятся за пределами возможностей МД. Так, вероятно, процессы комплексообразования, например, связывания транскрипционных факторов и поиск ими сайтов на последовательности ДНК или процессы образования гетероструктур ДНК-РНК (например, при связывании комплексов системы CRISPR-Cas9) находятся за пределами возможностей МД. Аналогичная ситуация возникает и при моделировании самосборки амилоидоподобных фибрилл. Несмотря на высокую точность задания силовых полей белков по сравнению с силовыми полями нуклеиновых кислот, ситуация здесь усложняется повышенными требованиями к точности определения энергии для различных конформаций пептидов. Известно, что амилоидоподобные фибриллы обладают конформационным полиморфизмом и в экспериментальных работах небольшое изменение условий (состав растворителя, температура) и, соответственно, энергетического баланса может приводить к формированию фибрилл различного типа.

Именно для решения такого рода задач, мы обратились к развитию и разработке методов интегративного моделирования. Базовая идея разработанных методов интегративного моделирования состоит в объединении информации заложенной в силовых полях с информацией о взаимном расположении элементов структуры, которую мы получаем из дополнительных экспериментальных методов. Общая логика построения интегративных подходов может быть сформулирована следующим образом. При моделировании структурных элементов и параметров системы, на структуру и значения которых может сильно повлиять неточность силовых полей, необходимо дополнительно привлекать экспериментальные данные. Вместе с тем, моделирование конформационно устойчивых структурных элементов можно производить на основании данных силовых полей. Примером служат полученные результаты моделирования комплексов нуклеосом с РНК полимеразой или моделирования геометрии линкерной ДНК при связывании гистона H1. Предполагалось, что свободная ДНК в целом должна находиться в В-конформации двойной спирали, согласно параметрам, заложенным в силовом поле, описывающем структуру ДНК. В то же время более тонкие параметры, такие как небольшой изгиб ДНК, уровень ее отворота от нуклеосомы определяются ограничениями, накладываемыми согласно дополнительными экспериментальными данными. В целом аналогичной логике мы следовали при интегративном моделировании амилоидоподобных фибрилл.

Были выделены параметры, которые сложно воспроизвести только с помощью моделирования, а именно, параметры укладки пептидов внутри фибриллы, и параметры, которые воспроизводятся моделированием с хорошей точностью, а именно, связь морфологии фибриллы с межмолекулярной укладкой пептидов. Таким образом, экспериментальные данные использовались частично при выборе укладок, а также для отбора фибрилл на уровне их крупномасштабной морфологии.

Таким образом общий принцип разработки интегративных подходов различных биомакромолекулярных комплексов, который можно рекомендовать, состоит в следующем. В первую очередь с учетом известных экспериментальных и литературных данных проводится анализ различных уровней структурной организации моделируемой системы, выделяются уровни организации, моделирование которых эффективно может быть проведено на основе имеющихся силовых полей, а также уровни организации, при моделировании которых необходимо привлекать дополнительные экспериментальные данные. На следующем этапе необходимо использовать различные интегративные методы моделирования, которые позволяют одновременно учитывать как различные экспериментальные данные, так и данные основанные на представлениях о внутримолекулярной структуре и межмолекулярных взаимодействиях.

В интегративном моделировании весьма полезным походом является использования соображений структурной симметрии. В некоторых случаях они позволяют кардинальным образом упростить задачу, повысить точность моделирования и уменьшить влияние возможных систематических ошибок в экспериментальных данных. Именно этот подход был продемонстрирован при использовании данных футпринтинга ДНК в построении моделей нуклеосом. На качество профиля футпринтинга вдоль нити ДНК могут оказывать влияние такие факторы, как наличие минорных альтернативных продуктов химического расщепления, локальная динамика ДНК в процессе реакции, шумовой фон. Однако, в силу наличия оси симметрии второго порядка в нуклеосоме такие эффекты будут симметричны для двух цепей ДНК и не влияют на алгоритм поиска положения в ДНК, относительно которого профили симметричны. Аналогичные соображения можно рекомендовать использовать при интегративном моделировании других ДНК белковых систем, обладающих симметрией. К таким системам относятся многие транскрипционные факторы, формирующие гомодимеры.

## Заключение

В настоящей работе автором развиты новые системные подходы для построения структурно-динамических моделей сложных биомакромолекулярных комплексов. Разработанные интегративные методы молекулярного моделирования основаны на сочетании (интеграции) физических моделей взаимодействия молекул с информацией получаемой из различных источников экспериментальных данных. В ходе исследования развиты подходы, (i) позволяющие проводить анализ динамики атомистических структур в микросекундном временном диапазоне, (ii) учитывать различные экспериментальные данные при построении моделей (например, данные экспериментов по футпринтингу ДНК, данные электронной микроскопии, данные FRET-микроскопии, данные рентгеновской дифракции на фибриллах, атомно-силовой микроскопии и т.д.), (iii) сочетать атомистическое и огрубленное представление молекулярных систем (в частности при моделировании ДНК-содержащих комплексов).

Разработанные автором подходы и методы рекомендуется применять для установления структуры и динамической подвижности биомакромолекулярных комплексов, размер и свойства которых затрудняют применение стандартных методов структурной биологии. Практическое применение данных подходов продемонстрировано в диссертации на примере ряда биомакромолекулярных систем, в том числе нуклеосом, комплексов нуклеосом с белками хроматина, амилоидоподобных фибрилл.

## **Выводы диссертационной работы**

1. Разработан интегративный подход и программные решения для мультимасштабного моделирования комплексов ДНК и белков, в котором используются различные данные экспериментов по ДНК футпринтингу, электронной микроскопии, FRET-микроскопии и комбинированное представление ДНК в атомистическом и в динуклеотидном приближении.
2. В микросекундном временном диапазоне определены параметры крупномасштабных функциональных конформационных изменений структуры нуклеосом: вычислен ансамбль конформаций линкерной

- ДНК, установлено влияние электростатического отталкивания на геометрию сегментов линкерной ДНК, определено характерное время диссоциации концов нуклеосомальной ДНК от гистонового октамера (10 мкс), установлены флуктуационные структурные механизмы образования дефектов кручения ДНК, переключения конформаций гистоновых хвостов.
3. Обработка экспериментальных данных по расщеплению ДНК гидроксильными радикалами (футпринтинга ДНК) позволила вычислить вероятность расщепления ДНК для каждого нуклеотида. Показано, что профили расщепления ДНК гидроксильными радикалами в нуклеосоме мало зависят от последовательности ДНК, и определяются в основном позиционированием ДНК на нуклеосоме. Предложен алгоритм точного определения положения ДНК в нуклеосоме по данным футпринтинга для двух цепей ДНК на основании положения оси псевдосимметрии.
  4. С помощью методов интегративного моделирования были установлены структуры и параметры взаимодействий комплексов нуклеосом с РНК полимеразой, белком СENP-C, гистоном H1, белками комплекса ФАСТ. Установлено, что в положении +49 после входа в нуклеосому РНК полимеразы II может формировать компактный комплекс с нуклеосомой, в котором контакты гистонов с ДНК сохраняются по обе стороны активного центра. В модели центральной нуклеосомы дрожжей определено положение ДНК и установлено, что белок СENP-C взаимодействует с нуклеосомой в районе 20 нуклеотидов от центра симметрии нуклеосомы. Предложены модели конформации линкерных сегментов ДНК при связывании гистона H1. Установлены амплитуды конформационной подвижности ДНК в нуклеосомах при связывании с комплексом ФАСТ.
  5. Разработанные подходы построения моделей амилоидоподобных фибрилл позволили изучить связь взаимного расположения пептидов в фибриллярных структурах с крупномасштабной морфологией амилоидоподобных фибрилл и таким образом установить структурную организацию филаментов на основе диблок олигомеров кватертиофена и пептида  $(Thr - Val)_3$ , а также на основе фрагмента белка gp120. Установлено, что рассмотренные амилоидоподобные фибриллы могут

образовываться на основе двух взаимодействующих бета-листов, которые формируют либо плоскую ленту, либо левозакрученную ленту с периодом 24-30 нм.

## Перспективы

Начиная со второй половины двадцатого века и до наших дней, благодаря развитию методов цитологии, молекулярной биологии, структурной биологии достигнут существенный прогресс в понимании устройства живых систем. Мы достаточно хорошо понимаем работу систем на молекулярном уровне - уровне генетического кода ДНК, уровне работы отдельных ферментов. Также, достаточно подробно охарактеризована структура и морфология многих клеток, механизмы различных внутриклеточных процессов и структурных преобразований. Однако, промежуточный супрамолекулярный уровень устройства живых организмов в диапазоне от 10 нм и до пределов разрешения оптической микроскопии остается слабо понятным, как с точки зрения структуры, так и с точки зрения механизмов функционирования. Причины данных затруднений двояки. С одной стороны биомакромолекулярные комплексы часто обладают динамической структурой, что снижает вероятность их кристаллизации для изучения структуры с помощью методов рентгеновской кристаллографии, усложняет возможности реконструкции методами криоэлектронной микроскопии. В том числе большой размер комплексов препятствует их эффективному изучению методами биомолекулярного ЯМР. С другой стороны на больших пространственных масштабах, в том числе из-за подвижности структуры биомакромолекулярных комплексов, множества различных взаимодействий между ними, становится сложным проследить связь между структурой и механизмами работы биомакромолекул. Иными словами, если для небольших белков изучение их структуры позволяет понять механизмы их работы, то для больших биомакромолекулярных комплексов, даже имея набор их структур в различных состояниях, сложнее выяснить и понять, какими силами и внешними факторами модулируется переход между различными состояниями и каким образом данный комплекс функционирует в контексте более сложной биологической системы. Перспективы понимания функционирования



живых систем на супрамолекулярном уровне на наш взгляд связаны с современным развитием нескольких областей: 1) методов структурной биологии, способных генерировать и обрабатывать большие массивы экспериментальных данных для получения информации о конформационных структурных ансамблях биомакромолекул и их комплексов, 2) методов интегративного моделирования, позволяющих создавать структурно-динамические, кинетические и статистические модели функционирования биомакромолекул на базе разнородной экспериментальной информации (в том числе геномных данных) с учетом физических взаимодействий между атомами и молекулами, 3) методов структурно-ориентированной системной биологии, позволяющих интегрировать структурные данные в кинетические модели описывающие работу живых систем, 4) методов машинного обучения, позволяющих создавать предсказательные модели на основе больших объемов многомерных данных. В ряде вышеозвученных направлений в последнее десятилетие наблюдается существенный прогресс, что закладывает основы для дальнейшего активного развития связанных областей биологии. В частности, в области структурной биологии развиваются методы криоэлектронной микроскопии, применения синхротронного и лазерного рентгеновского излучения (FEL - free electron lasers) для изучения динамики биомолекул. В области методов машинного обучения значительный прогресс связан с развитием как математических алгоритмов, так и с совершенствованием вычислительных возможностей благодаря совершенствованию массивно-параллельных графических процессоров и способов их программирования. Развитие вычислительных технологий также способствует постоянному увеличению возможностей молекулярного моделирования. В области системной и инженерной биологии активно развиваются подходы, связанные с математическим описанием биологических систем по аналогии с различными инженерными системами.

## Благодарности

Работы, описанные в диссертации были поддержаны рядом российских и международных грантов и стипендий, в том числе, грантами РФФИ 18–74–0006, 14-24-00031, 19-74-30003, РФФИ 20-34-70039, 12-04-31942, стипендией сотрудничества России-США в области биомедицинских наук, стипендией Национальной медицинской библиотеки США, Немецким научно-исследовательским обществом, проект SFB 569 A11. Работа выполнена с использованием оборудования Центра коллективного пользования сверхвысокопроизводительными вычислительными ресурсами МГУ имени М.В. Ломоносова, кластера Biowulf Национальных институтов здоровья (США), кластера Aldan университета г. Ульм (Германия).

Автор выражает благодарность своим научным руководителям и консультантам, под руководством которых автору посчастливилось работать, М.П.Кирпичникову, А.Р.Хохлову, А.Р. Панченко, Д. Ландсману, П.Г. Халатуру, В.А. Иванову, всем соавторам своих научных работ и коллегам за плодотворное сотрудничество, в особенности, Г.А. Армееву, В.Б. Журкину, В.М. Студитскому, Х.-В. Чанг, Д.А. Гайкаловой, К.Ву, Х. Жао, А. Гончаренко, И. Драйзену, Е.-К. Шиллингер, О.С. Соколовой, А.В. Феофанову, Н.В. Малюченко, Е. Бондаренко, М. Валиевой, А. Любителеву и многим другим, коллективам кафедры физики полимеров и кристаллов физического факультета МГУ, кафедры биоинженерии биологического факультета МГУ, Национального Центра Биотехнологической Информации Национальных Институтов Здоровья за продуктивную рабочую атмосферу и обсуждение работы.

Автор выражает благодарность своей семье за поддержку, без которой написание этой работы не было бы возможным, и А.Д. Шайтану за помощь с версткой текста.

## Список сокращений и условных обозначений

- АСМ** атомно-силовая микроскопия
- ДДЧ** диссипативная динамика частиц
- ДМС** диметилсульфат
- криоЭМ** криоэлектронная микроскопия
- МД** молекулярная динамика
- ПААГЭ** полиакриламидный гель электрофорез
- ПГУ** периодические граничные условия
- ПТМ** пост-трансляционные модификации (белков)
- ПО** программное обеспечение
- ПЭГ** полиэтиленгликоль
- п.н.** пара нуклеотидов
- РНКП** РНК-полимераза
- РСА** рентгено-структурный анализ
- СКН** система координат нуклеосомы
- ТАД** топологически ассоциированный домен
- ЯМР-спектроскопия** спектроскопия ядерного магнитного резонанса
- AL-PAR** однослойная фибрилла, основанная на параллельном расположении  $\beta$ -нитей
- DL-AP** двухслойная фибрилла, основанная на антипараллельном расположении  $\beta$ -нитей
- DL-PAR** двухслойная фибрилла, основанная на параллельном расположении  $\beta$ -нитей
- DPD** dissipative particle dynamics, диссипативная динамика частиц
- FRET** Forster resonance energy transfer, Ферстеровский перенос энергии
- HRF** hydroxyl-radical footprinting, метод футпринтинга ДНК на основе гидроксильных радикалов
- H-SASA** hydrogen atoms solvent accessible surface area, поверхность, доступная растворителю атомов водорода дезоксирибозы в нуклеотиде
- NCP** ядро нуклеосомы, нуклеосомный кор (nucleosome core particle)
- Pol II** РНК-полимераза II

- REMD** replica exchange molecular dynamics
- RMSD** root-mean-squared-deviation, среднеквадратичное отклонение
- SASA** solvent accessible surface area, поверхность, доступная растворителю
- SHL** super helix location, позиция на суперспирали нуклеосомной ДНК
- SL-AP** однослойная фибрилла, основанная на антипараллельном расположении  $\beta$ -нитей
- spFRET** single particle Forster resonance energy transfer, Ферстеровский перенос энергии в экспериментах на единичных частицах
- ZLS** zero-loop-stabilizing sequence, последовательность стабилизирующая нулевую петлю

## Список литературы

1. *Braitbard, M.* Integrative Structure Modeling: Overview and Assessment / M. Braitbard, D. Schneidman-Duhovny, N. Kalisman // Annual Review of Biochemistry. — 2019. — Т. 88, № 1. — С. 113—135. — DOI: [10.1146/annurev-biochem-013118-111429](https://doi.org/10.1146/annurev-biochem-013118-111429).
2. The Xplor-NIH NMR Molecular Structure Determination Package / C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. M. Clore // Journal of Magnetic Resonance (San Diego, Calif.: 1997). — 2003. — Jan. — Vol. 160, no. 1. — P. 65—73. — DOI: [10.1016/s1090-7807\(02\)00014-9](https://doi.org/10.1016/s1090-7807(02)00014-9).
3. xMDFF: Molecular Dynamics Flexible Fitting of Low-Resolution X-Ray Structures / R. McGreevy, A. Singharoy, Q. Li, J. Zhang, D. Xu, E. Perozo, K. Schulten // Acta Crystallographica Section D: Biological Crystallography. — 2014. — Sept. 1. — Vol. 70, issue 9, no. 9. — P. 2344—2355. — DOI: [10.1107/S1399004714013856](https://doi.org/10.1107/S1399004714013856).
4. Putting the Pieces Together: Integrative Modeling Platform Software for Structure Determination of Macromolecular Assemblies / D. Russel, K. Lasker, B. Webb, J. Velázquez-Muriel, E. Tjioe, D. Schneidman-Duhovny, B. Peterson, A. Sali // PLoS biology. — 2012. — Jan. — Vol. 10, no. 1. — e1001244. — DOI: [10.1371/journal.pbio.1001244](https://doi.org/10.1371/journal.pbio.1001244).
5. *Shaytan, A. K.* Computer Simulations of Self-Assembling Nanofibers from Thiophene-Peptide Oligomers : Dissertation / Shaytan A. K. — Universität Ulm, 2012. — DOI: <http://dx.doi.org/10.18725/OPARU-1936>.
6. *Армеев, Г.А.* Определение Структуры Нуклеосом и Их Комплексов с Белками Хроматина Методами Молекулярного Моделирования : диссертация на соискание степени кандидата физико-математических наук / Армеев, Г.А. — МГУ имени М.В. Ломоносова, 2018.
7. *Шайтан, А.К.* Компьютерное Моделирование и Статистический Анализ Самоорганизующихся Молекулярных Систем На Основе Пептидов : диссертация на соискание степени кандидата физико-математических наук / Шайтан, А.К. — МГУ имени М.В. Ломоносова, 2010.

8. Structural Biology | Nature / <https://www.nature.com/subjects/structural-biology>  
URL: <https://www.nature.com/subjects/structural-biology> (дата обр. 22.06.2019).
9. *Sanger, F.* The Amino-Acid Sequence in the Phenylalanyl Chain of Insulin. 1. The Identification of Lower Peptides from Partial Hydrolysates / F. Sanger, H. Tuppy // *Biochemical Journal*. — 1951. — Sept. 1. — Vol. 49, no. 4. — P. 463—481. — DOI: [10.1042/bj0490463](https://doi.org/10.1042/bj0490463).
10. Structure of a Ribonucleic Acid / R. W. Holley, J. Apgar, G. A. Everett, J. T. Madison, M. Marquisee, S. H. Merrill, J. R. Penswick, A. Zamir // *Science*. — 1965. — Mar. 19. — Vol. 147, no. 3664. — P. 1462—1465. — DOI: [10.1126/science.147.3664.1462](https://doi.org/10.1126/science.147.3664.1462).
11. *Gilbert, W.* The Nucleotide Sequence of the Lac Operator / W. Gilbert, A. Maxam // *Proceedings of the National Academy of Sciences*. — 1973. — Dec. 1. — Vol. 70, no. 12. — P. 3581—3584. — DOI: [10.1073/pnas.70.12.3581](https://doi.org/10.1073/pnas.70.12.3581).
12. *Maxam, A. M.* A New Method for Sequencing DNA / A. M. Maxam, W. Gilbert // *Proceedings of the National Academy of Sciences*. — 1977. — Feb. 1. — Vol. 74, no. 2. — P. 560—564. — DOI: [10.1073/pnas.74.2.560](https://doi.org/10.1073/pnas.74.2.560).
13. *Sanger, F.* DNA Sequencing with Chain-Terminating Inhibitors / F. Sanger, S. Nicklen, A. R. Coulson // *Proceedings of the National Academy of Sciences*. — 1977. — Dec. 1. — Vol. 74, no. 12. — P. 5463—5467. — DOI: [10.1073/pnas.74.12.5463](https://doi.org/10.1073/pnas.74.12.5463).
14. Первичная Структура Валиновой Транспортной РНК 1 Пекарских Дрожжей / А. Баев, Т. Венкстерн, А. Мирзабеков, А. Крутилина, Л. Ли, В. Аксельрод // *Молекулярная биология*. — 1967. — Т. 1, № 5. — С. 745—766.
15. A Three-Dimensional Model of the Myoglobin Molecule Obtained by x-Ray Analysis. / K. Jc, B. G, D. Hm, P. Rg, W. H, P. Dc // *Nature*. — 1958. — Mar. 1. — Vol. 181, no. 4610. — P. 662—666. — DOI: [10.1038/181662a0](https://doi.org/10.1038/181662a0).
16. Кристаллография и жизнь / под ред. М. V. Kovalchuk. — Moskva : Fizmatlit, 2012. — 376 с. — ISBN 978-5-9221-1403-5.
17. *Mannige, R. V.* Dynamic New World: Refining Our View of Protein Structure, Function and Evolution / R. V. Mannige // *Proteomes*. — 2014. — 7 марта. — Т. 2, № 1. — С. 128—153. — DOI: [10.3390/proteomes2010128](https://doi.org/10.3390/proteomes2010128).

18. *Needleman, D.* Active Matter at the Interface between Materials Science and Cell Biology / D. Needleman, Z. Dogic // Nature Reviews Materials. — 2017. — July 20. — Vol. 2, issue 9, no. 9. — P. 1–14. — DOI: [10.1038/natrevmats.2017.48](https://doi.org/10.1038/natrevmats.2017.48).
19. *Finney, J. L.* Bernal's Road to Random Packing and the Structure of Liquids / J. L. Finney // Philosophical Magazine. — 2013. — 5 нояб. — Т. 93, № 31–33. — С. 3940–3969. — DOI: [10.1080/14786435.2013.770179](https://doi.org/10.1080/14786435.2013.770179).
20. *Gregg, C.* MNCP Software Quality: Then and Now / C. Gregg. — URL: <https://permalink.lanl.gov/object/tr?what=info:lanl-repo/lareport/LA-UR-00-2532>.
21. English: ENIAC (Electronic Numerical Integrator And Computer) in Philadelphia, Pennsylvania. Glen Beck (Background) and Betty Snyder (Foreground) Program the ENIAC in Building 328 at the Ballistic Research Laboratory (BRL). [https:// Commons.Wikimedia.Org / Wiki/ File:Eniac.Jpg](https://commons.wikimedia.org/wiki/File:Eniac.jpg). — [1947].
22. Studies of the Nonlinear Problems / E. Fermi, P. Pasta, S. Ulam, M. Tsingou ; Los Alamos Scientific Lab., N. Mex. — 05/01/1955. — LA—1940. — DOI: [10.2172/4376203](https://doi.org/10.2172/4376203).
23. *Гривцов, А.Г.* Методика Численных Экспериментов и Динамика Микрорегетерогенных Систем / Гривцов, А.Г. // Метод Молекулярной Динамики в Физической Химии. — Москва : Наука, 1996. — С. 16. — ISBN 5-02-001907-0.
24. Overview of the CCP4 Suite and Current Developments / M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson // Acta Crystallographica Section D: Biological Crystallography. — 2011. — Apr. 1. — Vol. 67, issue 4, no. 4. — P. 235–242. — DOI: [10.1107/S0907444910045749](https://doi.org/10.1107/S0907444910045749).
25. Macromolecular Structure Determination Using X-Rays, Neutrons and Electrons: Recent Developments in Phenix / D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczy, V. B. Chen, T. I. Croll, B. Hintze, L.-W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon,

- M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams, P. D. Adams // *Acta Crystallographica Section D: Structural Biology*. — 2019. — Oct. 1. — Vol. 75, issue 10, no. 10. — P. 861–877. — DOI: [10.1107/S2059798319011471](https://doi.org/10.1107/S2059798319011471).
26. *Emsley, P.* Coot: Model-Building Tools for Molecular Graphics / P. Emsley, K. Cowtan // *Acta Crystallographica Section D: Biological Crystallography*. — 2004. — Dec. 1. — Vol. 60, issue 12, ARRAY(0xa66f734). — P. 2126–2132. — DOI: [10.1107/S0907444904019158](https://doi.org/10.1107/S0907444904019158).
27. *Schrodinger.* The PyMOL Molecular Graphics System, Version 1.8 / Schrodinger. — 2015.
28. UCSF Chimera—a Visualization System for Exploratory Research and Analysis / E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin // *Journal of Computational Chemistry*. — 2004. — Oct. — Vol. 25, no. 13. — P. 1605–1612. — DOI: [10.1002/jcc.20084](https://doi.org/10.1002/jcc.20084).
29. *Ломоносов, М.* Полное Собрание Сочинений: Ломоносов Михаил Васильевич — Алфавитный Каталог — Электронная Библиотека Руниверс / М. Ломоносов. — 1950–1983. — URL: <https://runivers.ru/lib/book6882/> (дата обр. 22.09.2020).
30. *Bernal, J.* The Bakerian Lecture, 1962 The Structure of Liquids / J. Bernal // *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*. — 1964. — 28 июля. — Т. 280, № 1382. — С. 299–322. — DOI: [10.1098/rspa.1964.0147](https://doi.org/10.1098/rspa.1964.0147).
31. *Breathnach, C. S.* Desmond Bernal and His Role in the Biological Exploitation of X-Ray Crystallography / C. S. Breathnach // *Journal of Medical Biography*. — 1995. — Nov. — Vol. 3, no. 4. — P. 197–200. — DOI: [10.1177/096777209500300403](https://doi.org/10.1177/096777209500300403).
32. *Kevin, J.* Schrödinger Desmond High-Performance Molecular Dynamics Simulations / J. Kevin ; Schrödinger. — 2020. — URL: <https://www.schrodinger.com/desmond>.
33. *Frenkel, D.* Understanding molecular simulation / D. Frenkel, B. Smit. — Academic Press, 2002.



34. *Hey, T.* The Fourth Paradigm: Data-Intensive Scientific Discovery / T. Hey, S. Tansley, K. Tolle. — Microsoft Research, 10.2009. — ISBN 978-0-9825442-0-4.
35. *Eckhardt, R.* And the MONTE CARLO METHOD / R. Eckhardt // Los Alamos Science Special Issue. — 1987. — P. 131—141.
36. *Metropolis, N.* The Monte Carlo Method / N. Metropolis, S. Ulam // Journal of the American Statistical Association. — 1949. — Sept. — Vol. 44, no. 247. — P. 335—341. — DOI: [10.1080/01621459.1949.10483310](https://doi.org/10.1080/01621459.1949.10483310).
37. *Metropolis, N.* The Beginnig of the Monte Carlo Method / N. Metropolis // Los Alamos Science Special Issue. — 1987. — С. 125—130.
38. Equation of State Calculations by Fast Computing Machines / N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller // The Journal of Chemical Physics. — 1953. — 1 июня. — Т. 21, № 6. — С. 1087—1092. — DOI: [10.1063/1.1699114](https://doi.org/10.1063/1.1699114).
39. *Иванов, В.* Методы Компьютерного Моделирования Для Исследования Полимеров и Биополимеров / В. Иванов. — 2009. — URL: <http://www.chem.msu.su/rus/books/polymer/welcome.html> (дата обр. 18.09.2020).
40. *Alder, B. J.* Phase Transition for a Hard Sphere System / B. J. Alder, T. E. Wainwright // The Journal of Chemical Physics. — 1957. — 1 нояб. — Т. 27, № 5. — С. 1208—1209. — DOI: [10.1063/1.1743957](https://doi.org/10.1063/1.1743957).
41. Dynamics of Radiation Damage / J. B. Gibson, A. N. Goland, M. Milgram, G. H. Vineyard // Physical Review. — 1960. — 15 нояб. — Т. 120, № 4. — С. 1229—1253. — DOI: [10.1103/PhysRev.120.1229](https://doi.org/10.1103/PhysRev.120.1229).
42. *Rahman, A.* Correlations in the Motion of Atoms in Liquid Argon / A. Rahman // Physical Review. — 1964. — 19 окт. — Т. 136, 2A. — A405—A411. — DOI: [10.1103/PhysRev.136.A405](https://doi.org/10.1103/PhysRev.136.A405).
43. *Allen, M. P.* Computer Simulation of Liquids / M. P. Allen, D. J. Tildesley. — New York, NY, USA : Clarendon Press, 1989. — ISBN 0-19-855645-4.
44. *Karplus, M.* Molecular Dynamics of Biological Macromolecules: A Brief History and Perspective / M. Karplus // Biopolymers. — 2003. — Mar. — Vol. 68, no. 3. — P. 350—358. — DOI: [10.1002/bip.10266](https://doi.org/10.1002/bip.10266).

45. Scalable Molecular Dynamics on CPU and GPU Architectures with NAMD / J. C. Phillips, D. J. Hardy, J. D. C. Maia, J. E. Stone, J. V. Ribeiro, R. C. Bernardi, R. Buch, G. Fiorin, J. Hénin, W. Jiang, R. McGreevy, M. C. R. Melo, B. K. Radak, R. D. Skeel, A. Singharoy, Y. Wang, B. Roux, A. Aksimentiev, Z. Luthey-Schulten, L. V. Kalé, K. Schulten, C. Chipot, E. Tajkhorshid // The Journal of Chemical Physics. — 2020. — 28 июля. — Т. 153, № 4. — С. 044130. — DOI: [10.1063/5.0014475](https://doi.org/10.1063/5.0014475).
46. *Sali, A.* Comparative Protein Modelling by Satisfaction of Spatial Restraints / A. Sali, T. L. Blundell // J Mol Biol. — 1993. — Dec. 5. — Vol. 234, no. 3. — P. 779—815. — DOI: [10.1006/jmbi.1993.1626](https://doi.org/10.1006/jmbi.1993.1626).
47. AMBER 17 / D. A. Case, D. S. Cerutti, T. E. C. III, T. Darden, R. Duke, T. J. Giese, H. Gohlke, A. Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mermelstein, K. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. Roe, A. Roitberg, C. Sagui, C. Simmerling, W. Botello-Smith, J. Swails, R. Walker, J. Wang, R. Wolf, X. Wu, L. Xiao, D. York, P. Kollman. — University of California, San Francisco.
48. GROMACS 2020 Manual / Lindahl, Abraham, Hess, V. D. Spoel. — 2020. — 1 янв. — DOI: [10.5281/ZENODO.3562512](https://doi.org/10.5281/ZENODO.3562512).
49. An Improved Nucleic Acid Parameter Set for the GROMOS Force Field / T. A. Soares, P. H. Hünenberger, M. A. Kastenholtz, V. Kräutler, T. Lenz, R. D. Lins, C. Oostenbrink, W. F. van Gunsteren // Journal of Computational Chemistry. — 2005. — May. — Vol. 26, no. 7. — P. 725—737. — DOI: [10.1002/jcc.20193](https://doi.org/10.1002/jcc.20193).
50. *Likhachev, I.* Parallelism of different levels in the program of molecular dynamics simulation PUMA-CUDA / I. Likhachev, N. Balabaev // (Mathematical Biology and Bioinformatics). — Pushchino, Moscow Region, Russia, 07.11.2018. — e44. — DOI: [10.17537/icmbb18.53](https://doi.org/10.17537/icmbb18.53).
51. NVIDIA GPU and Amber 16 Configurations / NVIDIA. — URL: <https://www.nvidia.com/es-la/data-center/gpu-accelerated-applications/amber/> (дата обр. 30.09.2020).
52. MDGRAPE-4: A Special-Purpose Computer System for Molecular Dynamics Simulations / I. Ohmura, G. Morimoto, Y. Ohno, A. Hasegawa, M. Taiji //

Philosophical transactions. Series A, Mathematical, physical, and engineering sciences. — 2014. — 6 аБГ. — Т. 372, № 2021. — DOI: [10.1098/rsta.2013.0387](https://doi.org/10.1098/rsta.2013.0387).

53. Anton 2: Raising the Bar for Performance and Programmability in a Special-Purpose Molecular Dynamics Supercomputer / D. E. Shaw, J. P. Grossman, J. A. Bank, B. Batson, J. A. Butts, J. C. Chao, M. M. Deneroff, R. O. Dror, A. Even, C. H. Fenton, A. Forte, J. Gagliardo, G. Gill, B. Greskamp, C. R. Ho, D. J. Ierardi, L. Iserovich, J. S. Kuskin, R. H. Larson, T. Layman, L. Lee, A. K. Lerer, C. Li, D. Killebrew, K. M. Mackenzie, S. Y. Mok, M. A. Moraes, R. Mueller, L. J. Nociolo, J. L. Peticolas, T. Quan, D. Ramot, J. K. Salmon, D. P. Scarpazza, U. B. Schafer, N. Siddique, C. W. Snyder, J. Spengler, P. T. P. Tang, M. Theobald, H. Toma, B. Towles, B. Vitale, S. C. Wang, C. Young // SC '14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC '14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis). — 11.2014. — С. 41–53. — DOI: [10.1109/SC.2014.9](https://doi.org/10.1109/SC.2014.9).
54. *Laio, A.* Escaping Free-Energy Minima / A. Laio, M. Parrinello // Proceedings of the National Academy of Sciences. — 2002. — Oct. 1. — Vol. 99, no. 20. — P. 12562–12566. — DOI: [10.1073/pnas.202427399](https://doi.org/10.1073/pnas.202427399).
55. Smoothed Biasing Forces Yield Unbiased Free Energies with the Extended-System Adaptive Biasing Force Method / A. Lesage, T. Lelièvre, G. Stoltz, J. Hénin // The Journal of Physical Chemistry B. — 2017. — 20 апр. — Т. 121, № 15. — С. 3676–3685. — DOI: [10.1021/acs.jpccb.6b10055](https://doi.org/10.1021/acs.jpccb.6b10055).
56. *Marchi, M.* Adiabatic Bias Molecular Dynamics: A Method to Navigate the Conformational Space of Complex Molecular Systems / M. Marchi, P. Ballone // The Journal of Chemical Physics. — 1999. — 11 февр. — Т. 110, № 8. — С. 3697–3702. — DOI: [10.1063/1.478259](https://doi.org/10.1063/1.478259).
57. *Sugita, Y.* Replica-Exchange Molecular Dynamics Method for Protein Folding / Y. Sugita, Y. Okamoto // Chemical Physics Letters. — 1999. — Nov. 26. — Vol. 314, no. 1. — P. 141–151. — DOI: [10.1016/S0009-2614\(99\)01123-9](https://doi.org/10.1016/S0009-2614(99)01123-9).
58. Неравновесная Молекулярная Динамика Наноструктур, Включая Биологические / К. В. Шайтан, Е. В. Турлей, Д. Н. Голик, К. Б. Терешкина,

О. В. Левцова, И. В. Федик, А. К. Шайтан, М. П. Кирпичников // Химическая физика. — 2006. — Т. 25, № 9. — С. 31—48. — IF WoS 0.58.

59. Promoting Transparency and Reproducibility in Enhanced Molecular Simulations / M. Bonomi, G. Bussi, C. Camilloni, G. A. Tribello, P. Banáš, A. Barducci, M. Bernetti, P. G. Bolhuis, S. Bottaro, D. Branduardi, R. Capelli, P. Carloni, M. Ceriotti, A. Cesari, H. Chen, W. Chen, F. Colizzi, S. De, M. De La Pierre, D. Donadio, V. Drobot, B. Ensing, A. L. Ferguson, M. Filizola, J. S. Fraser, H. Fu, P. Gasparotto, F. L. Gervasio, F. Giberti, A. Gil-Ley, T. Giorgino, G. T. Heller, G. M. Hocky, M. Iannuzzi, M. Invernizzi, K. E. Jelfs, A. Jussupow, E. Kirilin, A. Laio, V. Limongelli, K. Lindorff-Larsen, T. Löhner, F. Marinelli, L. Martin-Samos, M. Masetti, R. Meyer, A. Michaelides, C. Molteni, T. Morishita, M. Nava, C. Paissoni, E. Papaleo, M. Parrinello, J. Pfendtner, P. Piaggi, G. Piccini, A. Pietropaolo, F. Pietrucci, S. Pipolo, D. Provasi, D. Quigley, P. Raiteri, S. Raniolo, J. Rydzewski, M. Salvalaglio, G. C. Sosso, V. Spiwok, J. Šponer, D. W. H. Swenson, P. Tiwary, O. Valsson, M. Vendruscolo, G. A. Voth, A. White, The PLUMED consortium // Nature Methods. — 2019. — Aug. — Vol. 16, issue 8, no. 8. — P. 670—673. — DOI: [10.1038/s41592-019-0506-8](https://doi.org/10.1038/s41592-019-0506-8).
60. *Jarzynski, C.* Nonequilibrium Equality for Free Energy Differences / C. Jarzynski // Physical Review Letters. — 1997. — 7 апр. — Т. 78, № 14. — С. 2690—2693. — DOI: [10.1103/PhysRevLett.78.2690](https://doi.org/10.1103/PhysRevLett.78.2690).
61. *Crooks, G. E.* Entropy Production Fluctuation Theorem and the Nonequilibrium Work Relation for Free Energy Differences / G. E. Crooks // Physical Review E. — 1999. — 1 сент. — Т. 60, № 3. — С. 2721—2726. — DOI: [10.1103/PhysRevE.60.2721](https://doi.org/10.1103/PhysRevE.60.2721).
62. Macromolecular Modeling and Design in Rosetta: Recent Methods and Frameworks / J. K. Leman [et al.] // Nature Methods. — 2020. — July. — Vol. 17, issue 7, no. 7. — P. 665—680. — DOI: [10.1038/s41592-020-0848-2](https://doi.org/10.1038/s41592-020-0848-2).
63. *Wu, H.* AWSEM-IDP: A Coarse-Grained Force Field for Intrinsically Disordered Proteins / H. Wu, P. G. Wolynes, G. A. Papoian // The Journal of Physical Chemistry. B. — 2018. — Dec. 13. — Vol. 122, no. 49. — P. 11115—11125. — DOI: [10.1021/acs.jpcc.8b05791](https://doi.org/10.1021/acs.jpcc.8b05791).

64. *Jorgensen, W. L.* Development of the OPLS-AA Force Field for Organic and Biomolecular Systems / W. L. Jorgensen, J. Tirado-Rives // Journal of the Chemical Society, Abstracts. — 1998. — T. 216. — U696.
65. A 2nd Generation Force-Field for the Simulation of Proteins, Nucleic-Acids, and Organic-Molecules / W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman // Journal of the American Chemical Society. — 1995. — T. 117, № 19. — C. 5179—5197.
66. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins / a. D. MacKerell, D. Bashford, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus // The Journal of Physical Chemistry B. — 1998. — T. 102, № 18. — C. 3586—3616. — DOI: [10.1021/jp973084f](https://doi.org/10.1021/jp973084f).
67. *Schuler, L. D.* An Improved GROMOS96 Force Field for Aliphatic Hydrocarbons in the Condensed Phase / L. D. Schuler, X. Daura, W. F. Van Gunsteren // Journal of Computational Chemistry. — 2001. — T. 22, № 11. — C. 1205—1218. — DOI: [10.1002/jcc.1078](https://doi.org/10.1002/jcc.1078).
68. Structure and Energetics of Ligand Binding to Proteins: Escherichia Coli Dihydrofolate Reductase-Trimethoprim, a Drug-Receptor System / P. Dauber-Osguthorpe, V. a Roberts, D. J. Osguthorpe, J. Wolff, M. Genest, a. T. Hagler // Proteins. — 1988. — T. 4, № 1. — C. 31—47. — DOI: [10.1002/prot.340040106](https://doi.org/10.1002/prot.340040106).
69. *Sun, H.* Force Field for Computation of Conformational Energies, Structures, and Vibrational Frequencies of Aromatic Polyesters / H. Sun // Journal of Computational Chemistry. — 1994. — T. 15, № 7. — C. 752—768.
70. *Allinger, N. L.* Molecular Mechanics. The MM3 Force Field for Hydrocarbons. / N. L. Allinger, Y. H. Yuh, J. H. Lii // Journal of the American Chemical Society. — 1989. — T. 111, № 23. — C. 8551—8551. — DOI: [10.1021/ja00205a001](https://doi.org/10.1021/ja00205a001).

71. *Leach, A. R.* Molecular Modelling: Principles and Applications (2nd Edition) / A. R. Leach. — Prentice Hall, 2001.
72. *Hwang, M. J.* Derivation of Class II Force Fields. 2. Derivation and Characterization of a Class II Force Field, CFF93, for the Alkyl Functional Group and Alkane Molecules / M. J. Hwang, T. P. Stockfisch, a. T. Hagler // Journal of the American Chemical Society. — 1994. — T. 116, № 6. — C. 2515—2525. — DOI: [10.1021/ja00085a036](https://doi.org/10.1021/ja00085a036).
73. *Jorgensen, W. L.* The OPLS Force Field for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin. / W. L. Jorgensen, J. Tirado-Rives // J. Am. Chem. Soc. — 1988. — T. 110. — C. 1657—1666.
74. *Jorgensen, W. L.* Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids / W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives // J. Am. Chem. Soc. — 1996. — T. 118, № 45. — C. 11225—11236. — DOI: [10.1021/ja9621760](https://doi.org/10.1021/ja9621760).
75. *Sun, H.* Force field for computation of conformational energies, structures, and vibrational frequencies of aromatic polyesters / H. Sun // J. Comp. Chem. — Biosym Technologies, Inc., 9685 Scranton Road, San Diego, California 92121, 1994. — T. 15, № 7. — C. 752—768. — DOI: [10.1002/jcc.540150708](https://doi.org/10.1002/jcc.540150708).
76. *Swope, W. C.* A Computer Simulation Method for the Calculation of Equilibrium Constants for the Formation of Physical Clusters of Molecules: Application to Small Water Clusters / W. C. Swope, H. C. Anderson // The Journal of Chemical Physics. — 1982. — T. 76, № 1. — C. 637—637. — DOI: [10.1063/1.442716](https://doi.org/10.1063/1.442716).
77. *Hoover, W. G.* Time Reversibility, Computer Simulation, and Chaos. T. 13 / W. G. Hoover. — World Scientific Publishing, 2001. — (Advanced Series in Nonlinear Dynamics).
78. *Sagui, C.* Molecular Dynamics Simulations of Biomolecules: Long-Range Electrostatic Effects / C. Sagui, T. a Darden // Annual Review of Biophysics and Biomolecular Structure. — 1999. — T. 28. — C. 155—79. — DOI: [10.1146/annurev.biophys.28.1.155](https://doi.org/10.1146/annurev.biophys.28.1.155).
79. *Landau, L. D.* Statistical Physics. T. 5 / L. D. Landau, E. M. Lifshitz. — 3-е изд. — Butterworth-Heinemann, 2000. — (Course of Theoretical Physics).

80. *Frenkel, D.* Understanding Molecular Simulation: From Algorithms to Applications / D. Frenkel, B. Smith. — London : Academic Press, 2002.
81. 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. — T. 81, № 8. — C. 3684—3690.
82. *Hoover, W. G.* Canonical Dynamics: Equilibrium Phase-Space Distributions / W. G. Hoover // Physical Review A. — 1985. — T. 31, № 3. — C. 1695.
83. *Martyna, G.* Nose-Hoover Chains: The Canonical Ensemble via Continuous Dynamics / G. Martyna, M. Klein, M. Tuckerman // The Journal of Chemical Physics. — 1992. — T. 97, № 4. — C. 2635—2643.
84. *Andersen, H.* Molecular Dynamics Simulations at Constant Pressure and/or Temperature / H. Andersen // The Journal of Chemical Physics. — 1980. — T. 72, № 4. — C. 2384—2393.
85. *Bussi, G.* Canonical Sampling through Velocity Rescaling / G. Bussi, D. Donadio, M. Parrinello // The Journal of Chemical Physics. — 2007. — Jan. 7. — Vol. 126, no. 1. — P. 014101. — DOI: [10.1063/1.2408420](https://doi.org/10.1063/1.2408420).
86. *Harvey, S.* The Flying Ice Cube: Velocity Rescaling in Molecular Dynamics Leads to Violation of Energy Equipartition / S. Harvey, R. Tan, T. Cheatham // Journal of Computational Chemistry. — 1998. — T. 19, № 7. — C. 726—740.
87. *Golo, V. L.* [Dynamic attractor for the Berendsen thermostat an the slow dynamics of biomacromolecules] / V. L. Golo, K. V. Shaĭtan // Biofizika. — 2002. — T. 47, № 4. — C. 611—617.
88. *Nosé, S.* A Molecular Dynamics Method for Simulations in the Canonical Ensemble / S. Nosé // Molecular Physics: An International Journal at the Interface Between Chemistry and Physics. — 1984. — T. 52, № 2. — C. 255—268.
89. *Parrinello, M.* Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method / M. Parrinello, A. Rahman // Journal of Applied Physics. — 1981. — Dec. — Vol. 52, no. 12. — P. 7182—7190. — DOI: [10.1063/1.328693](https://doi.org/10.1063/1.328693).

90. *Zwanzig, R.* Nonequilibrium Statistical Mechanics / R. Zwanzig. — Illustrated Edition. — Oxford ; New York : Oxford University Press, 04/19/2001. — 240 p. — ISBN 978-0-19-514018-7.
91. *Hoogerbrugge, P. J.* Simulating Microscopic Hydrodynamic Phenomena with Dissipative Particle Dynamics / P. J. Hoogerbrugge, J. Koelman // EPL (Europhysics Letters). — 1992. — T. 19, № 3. — C. 155–160.
92. *Moeendarbary, E.* Dissipative Particle Dynamics: Introduction, Methodology and Complex Fluid Applications — A Review / E. Moeendarbary, T. Y. Ng, M. Zangeneh // International Journal of Applied Mechanics. — 2009. — T. 1, № 4. — C. 737–763.
93. *Smith, P.* Ewald Artifacts in Liquid State Molecular Dynamics Simulations / P. Smith, M. Pettitt // The Journal of Chemical Physics. — 1996. — T. 105, № 10. — C. 4289–4293.
94. *Smith, P. E.* On the Presence of Rotational Ewald Artifacts in the Equilibrium and Dynamical Properties of a Zwitterionic Tetrapeptide in Aqueous Solution / P. E. Smith, H. D. Blatt, B. M. Pettitt // The Journal of Physical Chemistry B. — 1997. — T. 101, № 19. — C. 3886–3890. — DOI: [10.1021/jp9637643](https://doi.org/10.1021/jp9637643).
95. *Lee, B.* The interpretation of protein structures: estimation of static accessibility. / B. Lee, F. M. Richards // J. Mol. Biol. — 1971. — T. 55, № 3. — C. 379–400.
96. *Connolly, M. L.* Analytical molecular surface calculation / M. L. Connolly // J. App. Cryst. — 1983. — T. 16, № 5. — C. 548–558. — DOI: [10.1107/S0021889883010985](https://doi.org/10.1107/S0021889883010985).
97. *Dickerson, R. E.* Definitions and Nomenclature of Nucleic Acid Structure Components / R. E. Dickerson // Nucleic Acids Research. — 1989. — Mar. 11. — Vol. 17, no. 5. — P. 1797–1803. — DOI: [10.1093/nar/17.5.1797](https://doi.org/10.1093/nar/17.5.1797).
98. *Zhurkin, V. B.* Anisotropic Flexibility of DNA and the Nucleosomal Structure. / V. B. Zhurkin, Y. P. Lysov, V. I. Ivanov // Nucleic Acids Research. — 1979. — Март. — T. 6, № 3. — C. 1081–1096.



99. *Lu, X.-J.* 3DNA: A Software Package for the Analysis, Rebuilding and Visualization of Three-Dimensional Nucleic Acid Structures / X.-J. Lu, W. K. Olson // *Nucleic Acids Research*. — 2003. — Sept. 1. — Vol. 31, no. 17. — P. 5108–5121. — DOI: [10.1093/nar/gkg680](https://doi.org/10.1093/nar/gkg680).
100. *El Hassan, M. A.* The Assessment of the Geometry of Dinucleotide Steps in Double-Helical DNA; a New Local Calculation Scheme / M. A. el Hassan, C. R. Calladine // *Journal of Molecular Biology*. — 1995. — Sept. 1. — Vol. 251, no. 5. — P. 648–664. — DOI: [10.1006/jmbi.1995.0462](https://doi.org/10.1006/jmbi.1995.0462).
101. A Standard Reference Frame for the Description of Nucleic Acid Base-Pair Geometry / W. K. Olson, M. Bansal, S. K. Burley, R. E. Dickerson, M. Gerstein, S. C. Harvey, U. Heinemann, X. J. Lu, S. Neidle, Z. Shakked, H. Sklenar, M. Suzuki, C. S. Tung, E. Westhof, C. Wolberger, H. M. Berman // *Journal of Molecular Biology*. — 2001. — Oct. 12. — Vol. 313, no. 1. — P. 229–237. — DOI: [10.1006/jmbi.2001.4987](https://doi.org/10.1006/jmbi.2001.4987).
102. Structure of a B-DNA Dodecamer: Conformation and Dynamics / H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura, R. E. Dickerson // *Proceedings of the National Academy of Sciences*. — 1981. — Apr. 1. — Vol. 78, no. 4. — P. 2179–2183. — DOI: [10.1073/pnas.78.4.2179](https://doi.org/10.1073/pnas.78.4.2179).
103. *Lu, X. J.* Structure and Conformation of Helical Nucleic Acids: Analysis Program (SCHNAaP) / X. J. Lu, M. A. El Hassan, C. A. Hunter // *Journal of Molecular Biology*. — 1997. — Oct. 31. — Vol. 273, no. 3. — P. 668–680. — DOI: [10.1006/jmbi.1997.1346](https://doi.org/10.1006/jmbi.1997.1346).
104. *Understanding DNA: The Molecule and How It Works* / C. R. Calladine, H. Drew, B. Luisi, A. Travers. — 3Rd Edition. — Academic Press, 03/13/2004. — 352 p.
105. DNA Sequence-Dependent Deformability Deduced from Protein-DNA Crystal Complexes / W. K. Olson, A. A. Gorin, X. J. Lu, L. M. Hock, V. B. Zhurkin // *Proceedings of the National Academy of Sciences of the United States of America*. — 1998. — Sept. 15. — Vol. 95, no. 19. — P. 11163–11168. — DOI: [10.1073/pnas.95.19.11163](https://doi.org/10.1073/pnas.95.19.11163).

106. *Shaytan, A. K.* Solvent Accessible Surface Area of Amino Acid Residues in Globular Proteins: Correlation of Apparent Transfer Free Energies with Experimental Hydrophobicity Scales / A. K. Shaytan, K. V. Shaitan, A. R. Khokhlov // *Biomacromolecules*. — 2009. — May 11. — Vol. 10, no. 5. — P. 1224–1237. — DOI: [10.1021/bm8015169](https://doi.org/10.1021/bm8015169). — IF WoS 5.667.
107. *Finkelstein, A. V.* Why Do Protein Architectures Have Boltzmann-like Statistics? / A. V. Finkelstein, A. Y. Badretdinov, A. M. Gutin // *Proteins: Structure, Function, and Bioinformatics*. — 1995. — Vol. 23, no. 2. — P. 142–150. — DOI: [10.1002/prot.340230204](https://doi.org/10.1002/prot.340230204).
108. A Multi-Modal Coarse Grained Model of DNA Flexibility Mappable to the Atomistic Level / J. Walther, P. D. Dans, A. Balaceanu, A. Hospital, G. Bayarri, M. Orozco // *Nucleic Acids Research*. — 2020. — Mar. 18. — Vol. 48, no. 5. — e29–e29. — DOI: [10.1093/nar/gkaa015](https://doi.org/10.1093/nar/gkaa015).
109. The Static and Dynamic Structural Heterogeneities of B-DNA: Extending Calladine–Dickerson Rules / P. D. Dans, A. Balaceanu, M. Pasi, A. S. Patelli, D. Petkevičiūtė, J. Walther, A. Hospital, G. Bayarri, R. Lavery, J. H. Maddocks, M. Orozco // *Nucleic Acids Research*. — 2019. — Dec. 2. — Vol. 47, no. 21. — P. 11090–11102. — DOI: [10.1093/nar/gkz905](https://doi.org/10.1093/nar/gkz905).
110. *Wang, D.* Sequence-Dependent Kink-and-Slide Deformations of Nucleosomal DNA Facilitated by Histone Arginines Bound in the Minor Groove / D. Wang, N. B. Ulyanov, V. B. Zhurkin // *Journal of Biomolecular Structure & Dynamics*. — 2010. — June. — Vol. 27, no. 6. — P. 843–859. — DOI: [10.1080/07391102.2010.10508586](https://doi.org/10.1080/07391102.2010.10508586).
111. muABC: A Systematic Microsecond Molecular Dynamics Study of Tetranucleotide Sequence Effects in B-DNA / M. Pasi, J. H. Maddocks, D. Beveridge, T. C. Bishop, D. A. Case, T. Cheatham, P. D. Dans, B. Jayaram, F. Lankas, C. Laughton, J. Mitchell, R. Osman, M. Orozco, A. Pérez, D. Petkevičiūtė, N. Spackova, J. Sponer, K. Zakrzewska, R. Lavery // *Nucleic Acids Research*. — 2014. — 29 окт. — T. 42, № 19. — C. 12272–12283. — DOI: [10.1093/nar/gku855](https://doi.org/10.1093/nar/gku855).

112. *Gorin, A. A.* B-DNA Twisting Correlates with Base-Pair Morphology / A. A. Gorin, V. B. Zhurkin, W. K. Olson // *Journal of Molecular Biology*. — 1995. — Mar. 17. — Vol. 247, no. 1. — P. 34–48. — DOI: [10.1006/jmbi.1994.0120](https://doi.org/10.1006/jmbi.1994.0120).
113. *Conformational Analysis of Nucleic Acids Revisited: Curves+* / R. Lavery, M. Moakher, J. H. Maddocks, D. Petkeviciute, K. Zakrzewska // *Nucleic Acids Research*. — 2009. — Sept. — Vol. 37, no. 17. — P. 5917–5929. — DOI: [10.1093/nar/gkp608](https://doi.org/10.1093/nar/gkp608).
114. *Lu, X. J.* 3DNA: A Versatile, Integrated Software System for the Analysis, Rebuilding and Visualization of Three-Dimensional Nucleic-Acid Structures / X. J. Lu, W. K. Olson // *Nat Protoc*. — 2008. — T. 3, № 7. — C. 1213–27. — DOI: [10.1038/nprot.2008.104](https://doi.org/10.1038/nprot.2008.104).
115. *Norouzi, D.* Topological Polymorphism of the Two-Start Chromatin Fiber / D. Norouzi, V. B. Zhurkin // *Biophysical Journal*. — 2015. — May 19. — Vol. 108, no. 10. — P. 2591–2600. — DOI: [10.1016/j.bpj.2015.04.015](https://doi.org/10.1016/j.bpj.2015.04.015).
116. *Internucleosomal Interactions Mediated by Histone Tails Allow Distant Communication in Chromatin* / O. I. Kulaeva, G. Zheng, Y. S. Polikanov, A. V. Colasanti, N. Clauvelin, S. Mukhopadhyay, A. M. Sengupta, V. M. Studitsky, W. K. Olson // *The Journal of Biological Chemistry*. — 2012. — 8 июня. — Т. 287, № 24. — С. 20248–20257. — DOI: [10.1074/jbc.M111.333104](https://doi.org/10.1074/jbc.M111.333104).
117. *Kornberg, R. D.* Chromatin Structure; Oligomers of the Histones / R. D. Kornberg, J. O. Thomas // *Science (New York, N.Y.)* — 1974. — May 24. — Vol. 184, no. 4139. — P. 865–868. — DOI: [10.1126/science.184.4139.865](https://doi.org/10.1126/science.184.4139.865).
118. *Chromosomes without a 30-Nm Chromatin Fiber* / Y. Joti, T. Hikima, Y. Nishino, F. Kamada, S. Hihara, H. Takata, T. Ishikawa, K. Maeshima // *Nucleus*. — 2012. — 1 сент. — Т. 3, № 5. — С. 404–410. — DOI: [10.4161/nucl.21222](https://doi.org/10.4161/nucl.21222).
119. *Razin, S. V.* Chromatin without the 30-Nm Fiber / S. V. Razin, A. A. Gavrilov // *Epigenetics*. — 2014. — 6 мая. — Т. 9, № 5. — С. 653–657. — DOI: [10.4161/epi.28297](https://doi.org/10.4161/epi.28297).

120. Influence of Fluctuations on DNA Curvature. A Comparison of Flexible and Static Wedge Models of Intrinsically Bent DNA / W. K. Olson, N. L. Marky, R. L. Jernigan, V. B. Zhurkin // Journal of Molecular Biology. — 1993. — July 20. — Vol. 232, no. 2. — P. 530–554. — DOI: [10.1006/jmbi.1993.1409](https://doi.org/10.1006/jmbi.1993.1409).
121. *Norouzi, D.* Dynamics of Chromatin Fibers: Comparison of Monte Carlo Simulations with Force Spectroscopy / D. Norouzi, V. B. Zhurkin // Biophysical Journal. — 2018. — Nov. 6. — Vol. 115, no. 9. — P. 1644–1655. — DOI: [10.1016/j.bpj.2018.06.032](https://doi.org/10.1016/j.bpj.2018.06.032).
122. *Korolev, N.* A Systematic Analysis of Nucleosome Core Particle and Nucleosome-Nucleosome Stacking Structure / N. Korolev, A. P. Lyubartsev, L. Nordenskiöld // Scientific Reports. — 2018. — Jan. 24. — Vol. 8, no. 1. — P. 1543. — DOI: [10.1038/s41598-018-19875-0](https://doi.org/10.1038/s41598-018-19875-0).
123. *Savel'ev, A.* Molecular Renormalization Group Coarse-Graining of Polymer Chains: Application to Double-Stranded DNA / A. Savel'ev, G. A. Papoian // Biophysical Journal. — 2009. — 20 мая. — Т. 96, № 10. — С. 4044–4052. — DOI: [10.1016/j.bpj.2009.02.067](https://doi.org/10.1016/j.bpj.2009.02.067).
124. *Ozer, G.* The Chromatin Fiber: Multiscale Problems and Approaches / G. Ozer, A. Luque, T. Schlick // Current opinion in structural biology. — 2015. — Apr. — Т. 31. — С. 124–139. — DOI: [10.1016/j.sbi.2015.04.002](https://doi.org/10.1016/j.sbi.2015.04.002).
125. DNA Topology in Chromatin Is Defined by Nucleosome Spacing / T. Nikitina, D. Norouzi, S. A. Grigoryev, V. B. Zhurkin // Science Advances. — 2017. — Oct. — Vol. 3, no. 10. — e1700957. — DOI: [10.1126/sciadv.1700957](https://doi.org/10.1126/sciadv.1700957).
126. *Schlick, T.* Monte Carlo, Harmonic Approximation, and Coarse-Graining Approaches for Enhanced Sampling of Biomolecular Structure / T. Schlick // F1000 Biology Reports. — 2009. — 29 июня. — Т. 1. — DOI: [10.3410/B1-48](https://doi.org/10.3410/B1-48).
127. Chromatin as Dynamic 10-Nm Fibers / K. Maeshima, R. Imai, S. Tamura, T. Nozaki // Chromosoma. — 2014. — Т. 123, № 3. — С. 225–237. — DOI: [10.1007/s00412-014-0460-2](https://doi.org/10.1007/s00412-014-0460-2).
128. Evidence for Heteromorphic Chromatin Fibers from Analysis of Nucleosome Interactions / S. A. Grigoryev, G. Arya, S. Correll, C. L. Woodcock, T. Schlick // Proceedings of the National Academy of Sciences of the United

- States of America. — 2009. — 11 авг. — Т. 106, № 32. — С. 13317—13322. — DOI: [10.1073/pnas.0903280106](https://doi.org/10.1073/pnas.0903280106).
129. Hierarchical Looping of Zigzag Nucleosome Chains in Metaphase Chromosomes / S. A. Grigoryev, G. Bascom, J. M. Buckwalter, M. B. Schubert, C. L. Woodcock, T. Schlick // Proceedings of the National Academy of Sciences of the United States of America. — 2016. — 2 февр. — Т. 113, № 5. — С. 1238—1243. — DOI: [10.1073/pnas.1518280113](https://doi.org/10.1073/pnas.1518280113).
130. Sub-Nucleosomal Genome Structure Reveals Distinct Nucleosome Folding Motifs / M. Ohno, T. Ando, D. G. Priest, V. Kumar, Y. Yoshida, Y. Taniguchi // Cell. — 2019. — 24 янв. — Т. 176, № 3. — 520—534.e25. — DOI: [10.1016/j.cell.2018.12.014](https://doi.org/10.1016/j.cell.2018.12.014).
131. "Ultraviolet Footprinting" Accurately Maps Sequence-Specific Contacts and DNA Kinking in the EcoRI Endonuclease-DNA Complex. / M. M. Becker, D. Lesser, M. Kurpiewski, A. Baranger, L. Jen-Jacobson // Proceedings of the National Academy of Sciences of the United States of America. — 1988. — Сент. — Т. 85, № 17. — С. 6247—6251.
132. Permanganate/S1 Nuclease Footprinting Reveals Non-B DNA Structures with Regulatory Potential across a Mammalian Genome / F. Kouzine, D. Wojtowicz, L. Baranello, A. Yamane, S. Nelson, W. Resch, K.-R. Kieffer-Kwon, C. J. Benham, R. Casellas, T. M. Przytycka, D. Levens // Cell Systems. — 2017. — Mar. 22. — Vol. 4, no. 3. — 344—356.e7. — DOI: [10.1016/j.cels.2017.01.013](https://doi.org/10.1016/j.cels.2017.01.013).
133. *Hornstra, I. K.* In Vivo Footprinting and Genomic Sequencing by Ligation-Mediated PCR / I. K. Hornstra, T. P. Yang // Analytical Biochemistry. — 1993. — Sept. — Vol. 213, no. 2. — P. 179—193. — DOI: [10.1006/abio.1993.1407](https://doi.org/10.1006/abio.1993.1407).
134. Ratiometric measurement and identification of single diffusing molecules / M. Dahan, A. A. Deniz, T. Ha, D. S. Chemla, P. G. Schultz, S. Weiss // Chemical Physics. — 1999. — 1 авг. — Т. 247, № 1. — С. 85—106. — DOI: [10.1016/S0301-0104\(99\)00132-9](https://doi.org/10.1016/S0301-0104(99)00132-9).

135. Simulation vs. Reality: A Comparison of In Silico Distance Predictions with DEER and FRET Measurements / D. Klose, J. P. Klare, D. Grohmann, C. W. M. Kay, F. Werner, H.-J. Steinhoff // PLOS ONE. — 2012. — June 25. — Vol. 7, no. 6. — e39492. — DOI: [10.1371/journal.pone.0039492](https://doi.org/10.1371/journal.pone.0039492).
136. Vesicle Encapsulation Studies Reveal That Single Molecule Ribozyme Heterogeneities Are Intrinsic / B. Okumus, T. J. Wilson, D. M. J. Lilley, T. Ha // Biophysical Journal. — 2004. — Oct. 1. — Vol. 87, no. 4. — P. 2798–2806. — DOI: [10.1529/biophysj.104.045971](https://doi.org/10.1529/biophysj.104.045971).
137. Rapid Spontaneous Accessibility of Nucleosomal DNA / G. Li, M. Levitus, C. Bustamante, J. Widom // Nature Structural & Molecular Biology. — 2005. — Jan. — Vol. 12, no. 1. — P. 46–53. — DOI: [10.1038/nsmb869](https://doi.org/10.1038/nsmb869).
138. *Roy, R.* A Practical Guide to Single-Molecule FRET / R. Roy, S. Hohng, T. Ha // Nature Methods. — 2008. — June. — Vol. 5, no. 6. — P. 507–516. — DOI: [10.1038/nmeth.1208](https://doi.org/10.1038/nmeth.1208).
139. Hydrophobic Fluorescent Probes Introduce Artifacts into Single Molecule Tracking Experiments Due to Non-Specific Binding / L. C. Zanetti-Domingues, C. J. Tynan, D. J. Rolfe, D. T. Clarke, M. Martin-Fernandez // PLOS ONE. — 2013. — Sept. 16. — Vol. 8, no. 9. — e74200. — DOI: [10.1371/journal.pone.0074200](https://doi.org/10.1371/journal.pone.0074200).
140. Accurate Single-Molecule FRET Studies Using Multiparameter Fluorescence Detection / E. Sisamakias, A. Valeri, S. Kalinin, P. J. Rothwell, C. A. M. Seidel // Methods in Enzymology. — 2010. — Vol. 475. — P. 455–514. — DOI: [10.1016/S0076-6879\(10\)75018-7](https://doi.org/10.1016/S0076-6879(10)75018-7).
141. FRETbursts: An Open Source Toolkit for Analysis of Freely-Diffusing Single-Molecule FRET / A. Ingargiola, E. Lerner, S. Chung, S. Weiss, X. Michalet // PLoS ONE. — 2016. — 17 abr. — T. 11, № 8. — DOI: [10.1371/journal.pone.0160716](https://doi.org/10.1371/journal.pone.0160716).
142. Structural Variability of Nucleosomes Detected by Single-Pair Förster Resonance Energy Transfer: Histone Acetylation, Sequence Variation, and Salt Effects / A. Gansen, K. Tóth, N. Schwarz, J. Langowski // The Journal of Physical Chemistry B. — 2009. — T. 113, № 9. — C. 2604–2613. — DOI: [10.1021/jp7114737](https://doi.org/10.1021/jp7114737).

143. *Williams, A. T. R.* Relative Fluorescence Quantum Yields Using a Computer-Controlled Luminescence Spectrometer / A. T. R. Williams, S. A. Winfield, J. N. Miller // *Analyst*. — 1983. — Jan. 1. — Vol. 108, no. 1290. — P. 1067–1071. — DOI: [10.1039/AN9830801067](https://doi.org/10.1039/AN9830801067).
144. Interplay of Alpha-Synuclein Binding and Conformational Switching Probed by Single-Molecule Fluorescence / A. C. M. Ferreon, Y. Gambin, E. A. Lemke, A. A. Deniz // *Proceedings of the National Academy of Sciences*. — 2009. — Apr. 7. — Vol. 106, no. 14. — P. 5645–5650. — DOI: [10.1073/pnas.0809232106](https://doi.org/10.1073/pnas.0809232106).
145. Accurate FRET Measurements within Single Diffusing Biomolecules Using Alternating-Laser Excitation / N. K. Lee, A. N. Kapanidis, Y. Wang, X. Michalet, J. Mukhopadhyay, R. H. Ebright, S. Weiss // *Biophysical Journal*. — 2005. — Apr. 1. — Vol. 88, no. 4. — P. 2939–2953. — DOI: [10.1529/biophysj.104.054114](https://doi.org/10.1529/biophysj.104.054114).
146. An Overview of Genome Organization and How We Got There: From FISH to Hi-C / J. Fraser, I. Williamson, W. A. Bickmore, J. Dostie // *Microbiology and molecular biology reviews: MMBR*. — 2015. — Sept. — Vol. 79, no. 3. — P. 347–372. — DOI: [10.1128/MMBR.00006-15](https://doi.org/10.1128/MMBR.00006-15).
147. *Pal, K.* Hi-C Analysis: From Data Generation to Integration / K. Pal, M. Forcato, F. Ferrari // *Biophysical Reviews*. — 2019. — Feb. 1. — Vol. 11, no. 1. — P. 67–78. — DOI: [10.1007/s12551-018-0489-1](https://doi.org/10.1007/s12551-018-0489-1).
148. Comprehensive Mapping of Long Range Interactions Reveals Folding Principles of the Human Genome / E. Lieberman-Aiden, N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, J. Dekker // *Science (New York, N.Y.)* — 2009. — 9 окт. — Т. 326, № 5950. — C. 289–293. — DOI: [10.1126/science.1181369](https://doi.org/10.1126/science.1181369).
149. Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions / J. R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, B. Ren // *Nature*. — 2012. — May. — Vol. 485, no. 7398. — P. 376–380. — DOI: [10.1038/nature11082](https://doi.org/10.1038/nature11082).

150. *Cui, K.* Genome-Wide Approaches to Determining Nucleosome Occupancy in Metazoans Using MNase-Seq / K. Cui, K. Zhao // *Methods in Molecular Biology* (Clifton, N.J.) — 2012. — Vol. 833. — P. 413–419. — DOI: [10.1007/978-1-61779-477-3\\_24](https://doi.org/10.1007/978-1-61779-477-3_24).
151. Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C / T.-H. S. Hsieh, A. Weiner, B. Lajoie, J. Dekker, N. Friedman, O. J. Rando // *Cell*. — 2015. — July 2. — Vol. 162, no. 1. — P. 108–119. — DOI: [10.1016/j.cell.2015.05.048](https://doi.org/10.1016/j.cell.2015.05.048).
152. Mapping 3D Genome Architecture through in Situ DNase Hi-C / V. Ramanani, D. A. Cusanovich, R. J. Hause, W. Ma, R. Qiu, X. Deng, C. A. Blau, C. M. Disteche, W. S. Noble, J. Shendure, Z. Duan // *Nature Protocols*. — 2016. — Nov. — Vol. 11, no. 11. — P. 2104–2121. — DOI: [10.1038/nprot.2016.126](https://doi.org/10.1038/nprot.2016.126).
153. Micro-C XL: Assaying Chromosome Conformation from the Nucleosome to the Entire Genome / T.-H. S. Hsieh, G. Fudenberg, A. Goloborodko, O. J. Rando // *Nature Methods*. — 2016. — Dec. — Vol. 13, no. 12. — P. 1009–1011. — DOI: [10.1038/nmeth.4025](https://doi.org/10.1038/nmeth.4025).
154. Next-Generation DNA Sequencing of Paired-End Tags (PET) for Transcriptome and Genome Analyses / M. J. Fullwood, C.-L. Wei, E. T. Liu, Y. Ruan // *Genome Research*. — 2009. — Apr. — Vol. 19, no. 4. — P. 521–532. — DOI: [10.1101/gr.074906.107](https://doi.org/10.1101/gr.074906.107).
155. Single-Cell Hi-C Reveals Cell-to-Cell Variability in Chromosome Structure / T. Nagano, Y. Lubling, T. J. Stevens, S. Schoenfelder, E. Yaffe, W. Dean, E. D. Laue, A. Tanay, P. Fraser // *Nature*. — 2013. — Oct. 3. — Vol. 502, no. 7469. — P. 59–64. — DOI: [10.1038/nature12593](https://doi.org/10.1038/nature12593).
156. *Teif, V. B.* Nucleosome Positioning: Resources and Tools Online / V. B. Teif // *Briefings in Bioinformatics*. — 2016. — Sept. — Vol. 17, no. 5. — P. 745–757. — DOI: [10.1093/bib/bbv086](https://doi.org/10.1093/bib/bbv086).
157. NucMap: A Database of Genome-Wide Nucleosome Positioning Map across Species / Y. Zhao, J. Wang, F. Liang, Y. Liu, Q. Wang, H. Zhang, M. Jiang, Z. Zhang, W. Zhao, Y. Bao, Z. Zhang, J. Wu, Y. W. Asmann, R. Li, J. Xiao // *Nucleic Acids Research*. — 2019. — Jan. 8. — Vol. 47, no. D1. — P. D163–D169. — DOI: [10.1093/nar/gky980](https://doi.org/10.1093/nar/gky980).



158. A Map of Nucleosome Positions in Yeast at Base-Pair Resolution / K. Brogaard, L. Xi, J.-P. Wang, J. Widom // *Nature*. — 2012. — June 28. — Vol. 486, no. 7404. — P. 496–501. — DOI: [10.1038/nature11142](https://doi.org/10.1038/nature11142).
159. *Wedel, C.* Genome-Wide Analysis of Chromatin Structures in *Trypanosoma Brucei* Using High-Resolution MNase-ChIP-Seq / C. Wedel, T. N. Siegel // *Experimental Parasitology*. — 2017. — 1 сент. — Т. 180. — С. 2–12. — (Proceedings of the 27th Meeting of the German Society for Parasitology 2016). — DOI: [10.1016/j.exppara.2017.03.003](https://doi.org/10.1016/j.exppara.2017.03.003).
160. Subnucleosomal Structures and Nucleosome Asymmetry across a Genome / H. S. Rhee, A. R. Bataille, L. Zhang, B. F. Pugh // *Cell*. — 2014. — Dec. 4. — Vol. 159, no. 6. — P. 1377–1388. — DOI: [10.1016/j.cell.2014.10.054](https://doi.org/10.1016/j.cell.2014.10.054).
161. *Quintales, L.* Comparative Analysis of Methods for Genome-Wide Nucleosome Cartography / L. Quintales, E. Vázquez, F. Antequera // *Briefings in Bioinformatics*. — 2015. — July. — Vol. 16, no. 4. — P. 576–587. — DOI: [10.1093/bib/bbu037](https://doi.org/10.1093/bib/bbu037).
162. *Van Holde, K. E.* Chromatin / K. E. Van Holde. — Springer-Verlag, 1989. — (Springer Series in Molecular Biology). — ISBN 978-0-387-96694-6.
163. *Olins, A. L.* Spheroid Chromatin Units (v Bodies) / A. L. Olins, D. E. Olins // *Science (New York, N.Y.)* — 1974. — Jan. 25. — Vol. 183, no. 4122. — P. 330–332. — DOI: [10.1126/science.183.4122.330](https://doi.org/10.1126/science.183.4122.330).
164. *Kornberg, R. D.* Chromatin Structure: A Repeating Unit of Histones and DNA / R. D. Kornberg // *Science (New York, N.Y.)* — 1974. — May 24. — Vol. 184, no. 4139. — P. 868–871. — DOI: [10.1126/science.184.4139.868](https://doi.org/10.1126/science.184.4139.868).
165. Crystal Structure of the Nucleosome Core Particle at 2.8 Å Resolution / K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent, T. J. Richmond // *Nature*. — 1997. — Sept. 18. — Vol. 389, no. 6648. — P. 251–260. — DOI: [10.1038/38444](https://doi.org/10.1038/38444).
166. *Tan, S.* Nucleosome Structural Studies / S. Tan, C. A. Davey // *Curr Opin Struct Biol*. — 2011. — Февр. — Т. 21, № 1. — С. 128–36. — DOI: [10.1016/j.sbi.2010.11.006](https://doi.org/10.1016/j.sbi.2010.11.006).
167. *Malik, H. S.* Phylogenomics of the Nucleosome / H. S. Malik, S. Henikoff // *Nature Structural Biology*. — 2003. — Т. 10, № 11. — С. 882–891. — DOI: [10.1038/nsb996](https://doi.org/10.1038/nsb996).

168. *Bhattacharyya, S.* Archaeal DNA on the Histone Merry-Go-Round / S. Bhattacharyya, F. Mattioli, K. Luger // The FEBS Journal. — 2018. — T. 285, № 17. — C. 3168–3174. — DOI: [10.1111/febs.14495](https://doi.org/10.1111/febs.14495).
169. *Bowman, G. D.* Post-Translational Modifications of Histones That Influence Nucleosome Dynamics / G. D. Bowman, M. G. Poirier // Chemical Reviews. — 2015. — Mar. 25. — Vol. 115, no. 6. — P. 2274–2295. — DOI: [10.1021/cr500350x](https://doi.org/10.1021/cr500350x).
170. The Nucleosome Family: Dynamic and Growing / J. Zlatanova, T. C. Bishop, J. M. Victor, V. Jackson, K. van Holde // Structure. — 2009. — 13 февр. — T. 17, № 2. — C. 160–71. — DOI: [10.1016/j.str.2008.12.016](https://doi.org/10.1016/j.str.2008.12.016).
171. Asymmetric Unwrapping of Nucleosomal DNA Propagates Asymmetric Opening and Dissociation of the Histone Core / Y. Chen, J. M. Tokuda, T. Topping, S. P. Meisburger, S. A. Pabit, L. M. Gloss, L. Pollack // Proceedings of the National Academy of Sciences of the United States of America. — 2017. — Jan. 10. — Vol. 114, no. 2. — P. 334–339. — DOI: [10.1073/pnas.1611118114](https://doi.org/10.1073/pnas.1611118114).
172. *Sinha, K. K.* Distortion of Histone Octamer Core Promotes Nucleosome Mobilization by a Chromatin Remodeler / K. K. Sinha, J. D. Gross, G. J. Narlikar // Science. — 2017. — Jan. 20. — Vol. 355, no. 6322. — eaaa3761. — DOI: [10.1126/science.aaa3761](https://doi.org/10.1126/science.aaa3761).
173. Mechanism of Chromatin Remodelling Revealed by the Snf2-Nucleosome Structure / X. Liu, M. Li, X. Xia, X. Li, Z. Chen // Nature. — 2017. — Apr. — Vol. 544, no. 7651. — P. 440–445. — DOI: [10.1038/nature22036](https://doi.org/10.1038/nature22036).
174. H2A Histone-Fold and DNA Elements in Nucleosome Activate SWR1-Mediated H2A.Z Replacement in Budding Yeast / A. Ranjan, F. Wang, G. Mizuguchi, D. Wei, Y. Huang, C. Wu // eLife. — 2015. — T. 4, JUNE 2015. — C. 1–11. — DOI: [10.7554/eLife.06845.001](https://doi.org/10.7554/eLife.06845.001).
175. P53 Binding To Nucleosomes Within the P21 Promoter in Vivo Leads To Nucleosome Loss and Transcriptional Activation / O. Laptenko, R. Beckerman, E. Freulich, C. Prives // Proceedings of the National Academy of Sciences. — 2011. — T. 108, № 26. — C. 10385–10390. — DOI: [10.1073/pnas.1105680108](https://doi.org/10.1073/pnas.1105680108).

176. *Zaret, K. S.* Pioneer Transcription Factors, Chromatin Dynamics, and Cell Fate Control / K. S. Zaret, S. E. Mango // *Current Opinion in Genetics and Development*. — 2016. — T. 37. — C. 76–81. — DOI: [10.1016/j.gde.2015.12.003](https://doi.org/10.1016/j.gde.2015.12.003).
177. *Polach, K. J.* Mechanism of Protein Access to Specific DNA Sequences in Chromatin: A Dynamic Equilibrium Model for Gene Regulation / K. J. Polach, J. Widom // *Journal of Molecular Biology*. — 1995. — 24 нояб. — Т. 254, № 2. — С. 130–149. — DOI: [10.1006/jmbi.1995.0606](https://doi.org/10.1006/jmbi.1995.0606).
178. *Li, G.* Nucleosomes Facilitate Their Own Invasion / G. Li, J. Widom // *Nature Structural and Molecular Biology*. — 2004. — Т. 11, № 8. — С. 763–769. — DOI: [10.1038/nsmb801](https://doi.org/10.1038/nsmb801).
179. *Bilokapic, S.* Histone Octamer Rearranges to Adapt to DNA Unwrapping / S. Bilokapic, M. Strauss, M. Halic // *Nature Structural & Molecular Biology*. — 2018. — Jan. — Vol. 25, no. 1. — P. 101–108. — DOI: [10.1038/s41594-017-0005-5](https://doi.org/10.1038/s41594-017-0005-5).
180. Nucleosomes Can Form a Polar Barrier to Transcript Elongation by RNA Polymerase II / V. A. Bondarenko, L. M. Steele, A. Újvári, D. A. Gaykalova, O. I. Kulaeva, Y. S. Polikanov, D. S. Luse, V. M. Studitsky // *Molecular Cell*. — 2006. — Т. 24, № 3. — С. 469–479. — DOI: [10.1016/j.molcel.2006.09.009](https://doi.org/10.1016/j.molcel.2006.09.009).
181. Nucleosome Accessibility Governed by the Dimer/Tetramer Interface / V. Böhm, A. R. Hieb, A. J. Andrews, A. Gansen, A. Rocker, K. Tóth, K. Luger, J. Langowski // *Nucleic Acids Research*. — 2011. — App. — Т. 39, № 8. — С. 3093–3102. — DOI: [10.1093/nar/gkq1279](https://doi.org/10.1093/nar/gkq1279).
182. Nucleosome Remodeling Induced by RNA Polymerase II: Loss of the H2A/H2B Dimer during Transcription / M. L. Kireeva, W. Walter, V. Tchernajenko, V. Bondarenko, M. Kashlev, V. M. Studitsky // *Molecular Cell*. — 2002. — Т. 9, № 3. — С. 541–552. — DOI: [10.1016/S1097-2765\(02\)00472-0](https://doi.org/10.1016/S1097-2765(02)00472-0).
183. Partially Assembled Nucleosome Structures at Atomic Detail / G. N. Rychkov, A. V. Ilatovskiy, I. B. Nazarov, A. V. Shvetsov, D. V. Lebedev, A. Y. Konev, V. V. Isaev-Ivanov, A. V. Onufriev // *Biophysical*

- Journal. — 2017. — Feb. 7. — Vol. 112, no. 3. — P. 460–472. — DOI: [10.1016/j.bpj.2016.10.041](https://doi.org/10.1016/j.bpj.2016.10.041).
184. Insights into the Molecular Architecture and Histone H3-H4 Deposition Mechanism of Yeast Chromatin Assembly Factor 1 / P. V. Sauer, J. Timm, D. Liu, D. Sitbon, E. Boeri-Erba, C. Velours, N. Mücke, J. Langowski, F. Ochsenbein, G. Almouzni, D. Panne // *eLife*. — 2017. — T. 6. — DOI: [10.7554/eLife.23474](https://doi.org/10.7554/eLife.23474).
185. DNA-Mediated Association of Two Histone-Bound Complexes of Yeast Chromatin Assembly Factor-1 (CAF-1) Drives Tetrasome Assembly in the Wake of DNA Replication / F. Mattioli, Y. Gu, T. Yadav, J. L. Balsbaugh, M. R. Harris, E. S. Findlay, Y. Liu, C. A. Radebaugh, L. A. Stargell, N. G. Ahn, I. Whitehouse, K. Luger // *eLife*. — 2017. — T. 6. — C. 1–23. — DOI: [10.7554/eLife.22799](https://doi.org/10.7554/eLife.22799).
186. Remarkable Evolutionary Plasticity of Centromeric Chromatin / S. Henikoff, J. Thakur, S. Kasinathan, P. B. Talbert // *Cold Spring Harbor Symposia on Quantitative Biology*. — 2017. — T. LXXXII. — C. 033605–033605. — DOI: [10.1101/sqb.2017.82.033605](https://doi.org/10.1101/sqb.2017.82.033605).
187. *Furuyama, T.* Reconstitution of Hemisomes on Budding Yeast Centromeric DNA / T. Furuyama, C. A. Codomo, S. Henikoff // *Nucleic Acids Research*. — 2013. — T. 41, № 11. — C. 5769–5783. — DOI: [10.1093/nar/gkt314](https://doi.org/10.1093/nar/gkt314).
188. Structural Plasticity of Single Chromatin Fibers Revealed by Torsional Manipulation / A. Bancaud, N. Conde E Silva, M. Barbi, G. Wagner, J. F. Allemand, J. Mozziconacci, C. Lavelle, V. Croquette, J. M. Victor, A. Prunell, J. L. Viovy // *Nature Structural and Molecular Biology*. — 2006. — T. 13, № 5. — C. 444–450. — DOI: [10.1038/nsmb1087](https://doi.org/10.1038/nsmb1087).
189. Archaeal Histone Tetramerization Determines DNA Affinity and the Direction of DNA Supercoiling / F. Marc, K. Sandman, R. Lurz, J. N. Reeve // *Journal of Biological Chemistry*. — 2002. — T. 277, № 34. — C. 30879–30886. — DOI: [10.1074/jbc.M203674200](https://doi.org/10.1074/jbc.M203674200).
190. *Ngo, T. T. M.* Nucleosomes Undergo Slow Spontaneous Gaping / T. T. M. Ngo, T. Ha // *Nucleic Acids Research*. — 2015. — Apr. — T. 43, № 8. — C. 3964–3971. — DOI: [10.1093/nar/gkv276](https://doi.org/10.1093/nar/gkv276).

191. CENP-C Reshapes and Stabilizes CENP-A Nucleosomes at the Centromere / S. J. Falk, L. Y. Guo, N. Sekulic, E. M. Smoak, T. Mani, G. A. Logsdon, K. Gupta, L. E. Jansen, G. D. Van Duyne, S. A. Vinogradov, M. A. Lampson, B. E. Black // *Science*. — 2015. — T. 348, № 6235. — C. 699—703. — DOI: [10.1126/science.1259308](https://doi.org/10.1126/science.1259308).
192. CENP-C Directs a Structural Transition of CENP-A Nucleosomes Mainly through Sliding of DNA Gyres / S. J. Falk, J. Lee, N. Sekulic, M. A. Sennett, T. H. Lee, B. E. Black // *Nature Structural and Molecular Biology*. — 2016. — T. 23, № 3. — C. 204—208. — DOI: [10.1038/nsmb.3175](https://doi.org/10.1038/nsmb.3175).
193. *Eslami-Mossallam, B.* Nucleosome Dynamics: Sequence Matters / B. Eslami-Mossallam, H. Schiessel, J. van Noort // *Advances in Colloid and Interface Science*. — 2016. — T. 232. — C. 101—113. — DOI: [10.1016/j.cis.2016.01.007](https://doi.org/10.1016/j.cis.2016.01.007).
194. Nucleosomes in Solution Exist as a Mixture of Twist-Defect States / R. S. Edayathumangalam, P. Weyermann, P. B. Dervan, J. M. Gottesfeld, K. Luger // *J Mol Biol*. — 2005. — Jan. 7. — Vol. 345, no. 1. — P. 103—14. — DOI: [10.1016/j.jmb.2004.10.012](https://doi.org/10.1016/j.jmb.2004.10.012).
195. *Öztürk, M. A.* Toward an Ensemble View of Chromosome Structure: A Paradigm Shift from One to Many / M. A. Öztürk, V. Cojocaru, R. C. Wade // *Structure*. — 2018. — DOI: [10.1016/j.str.2018.05.009](https://doi.org/10.1016/j.str.2018.05.009).
196. Site-Specific Installation of Succinyl Lysine Analog into Histones Reveals the Effect of H2BK34 Succinylation on Nucleosome Dynamics / Y. Jing, Z. Liu, G. Tian, X. Bao, T. Ishibashi, X. D. Li // *Cell Chemical Biology*. — 2018. — T. 25, № 2. — 166—174.e7. — DOI: [10.1016/j.chembiol.2017.11.005](https://doi.org/10.1016/j.chembiol.2017.11.005).
197. Local DNA Sequence Controls Asymmetry of DNA Unwrapping from Nucleosome Core Particles / A. W. Mauney, J. M. Tokuda, L. M. Gloss, O. Gonzalez, L. Pollack // *Biophysical Journal*. — 2018. — Sept. 4. — Vol. 115, no. 5. — P. 773—781. — DOI: [10.1016/j.bpj.2018.07.009](https://doi.org/10.1016/j.bpj.2018.07.009).
198. The Flexible Ends of CENP-A Nucleosome Are Required for Mitotic Fidelity / Y. Roulland, K. Ouararhni, M. Naidenov, L. Ramos, M. Shuaib, S. H. Syed, I. N. Lone, R. Boopathi, E. Fontaine, G. Papai, H. Tachiwana, T. Gautier, D. Skoufias, K. Padmanabhan, J. Bednar, H. Kurumizaka, P. Schultz, D. Angelov, A. Hamiche, S. Dimitrov // *Molecular Cell*. — 2016. — T. 63, № 4. — C. 674—685. — DOI: [10.1016/j.molcel.2016.06.023](https://doi.org/10.1016/j.molcel.2016.06.023).

199. The Histone Chaperone FACT Modulates Nucleosome Structure by Tethering Its Components / T. Wang, Y. Liu, G. Edwards, D. Krzizike, H. Scherman, K. Luger // Life Science Alliance. — 2018. — Т. 1, № 4. — e201800107—e201800107. — DOI: [10.26508/lsa.201800107](https://doi.org/10.26508/lsa.201800107).
200. *Lee, J.* Single-Molecule Investigations on Histone H2A-H2B Dynamics in the Nucleosome / J. Lee, T. H. Lee // Biochemistry. — 2017. — Т. 56, № 7. — С. 977—985. — DOI: [10.1021/acs.biochem.6b01252](https://doi.org/10.1021/acs.biochem.6b01252).
201. Unfolding of Core Nucleosomes by PARP-1 Revealed by spFRET Microscopy / D. Sultanov, N. Gerasimova, K. Kudryashova, N. Maluchenko, E. Kotova, M.-F. Langelier, J. Pascal, M. Kirpichnikov, A. Feofanov, V. Studitsky // AIMS Genetics. — 2017. — Т. 4, № 1. — С. 21—31. — DOI: [10.3934/genet.2017.1.21](https://doi.org/10.3934/genet.2017.1.21).
202. Structure and Dynamics of a 197 Bp Nucleosome in Complex with Linker Histone H1 / J. Bednar, I. Garcia-Saez, R. Boopathi, A. R. Cutter, G. Papai, A. Reymer, S. H. Syed, I. N. Lone, O. Tonchev, C. Crucifix, H. Menoni, C. Papin, D. A. Skoufias, H. Kurumizaka, R. Lavery, A. Hamiche, J. J. Hayes, P. Schultz, D. Angelov, C. Petosa, S. Dimitrov // Molecular Cell. — 2017. — 4 мая. — Т. 66, № 3. — 384—397.e8. — DOI: [10.1016/j.molcel.2017.04.012](https://doi.org/10.1016/j.molcel.2017.04.012).
203. Modulation of Nucleosomal DNA Accessibility via Charge-Altering Post-Translational Modifications in Histone Core / A. T. Fenley, R. Anandakrishnan, Y. H. Kidane, A. V. Onufriev // Epigenetics & Chromatin. — 2018. — Mar. 16. — Vol. 11, no. 1. — P. 11. — DOI: [10.1186/s13072-018-0181-5](https://doi.org/10.1186/s13072-018-0181-5).
204. *Potoyan, D. A.* Regulation of the H4 Tail Binding and Folding Landscapes via Lys-16 Acetylation / D. A. Potoyan, G. A. Papoian // Proc Natl Acad Sci U S A. — 2012. — 30 окт. — Т. 109, № 44. — С. 17857—62. — DOI: [10.1073/pnas.1201805109](https://doi.org/10.1073/pnas.1201805109).
205. Multisite Substrate Recognition in Asf1-Dependent Acetylation of Histone H3 K56 by Rtt109 / L. Zhang, A. Serra-Cardona, H. Zhou, M. Wang, N. Yang, Z. Zhang, R.-M. M. Xu // Cell. — 2018. — АВГ. — Т. 174, № 4. — 818—830.e11. — DOI: [10.1016/j.cell.2018.07.005](https://doi.org/10.1016/j.cell.2018.07.005).

206. H3R42me2a Is a Histone Modification with Positive Transcriptional Effects / F. Casadio, X. Lu, S. B. Pollock, G. LeRoy, B. A. Garcia, T. W. Muir, R. G. Roeder, C. D. Allis // *Proceedings of the National Academy of Sciences of the United States of America*. — 2013. — 10 сент. — Т. 110, № 37. — С. 14894–14899. — DOI: [10.1073/pnas.1312925110](https://doi.org/10.1073/pnas.1312925110).
207. Mycobacteria Modulate Host Epigenetic Machinery by Rv1988 Methylation of a Non-Tail Arginine of Histone H3 / I. Yaseen, P. Kaur, V. K. Nandicoori, S. Khosla // *Nature Communications*. — 2015. — Т. 6. — С. 1–13. — DOI: [10.1038/ncomms9922](https://doi.org/10.1038/ncomms9922).
208. Histone H4 Lysine 91 Acetylation: A Core Domain Modification associated with Chromatin Assembly / J. Ye, X. Ai, E. E. Eugeni, L. Zhang, L. R. Carpenter, M. A. Jelinek, M. A. Freitas, M. R. Parthun // *Molecular Cell*. — 2005. — Т. 18, № 1. — С. 123–130. — DOI: [10.1016/j.molcel.2005.02.031](https://doi.org/10.1016/j.molcel.2005.02.031).
209. Histone H1 Acetylation at Lysine 85 Regulates Chromatin Condensation and Genome Stability upon DNA Damage / Y. Li, Z. Li, L. Dong, M. Tang, P. Zhang, C. Zhang, Z. Cao, Q. Zhu, Y. Chen, H. Wang, T. Wang, D. Lv, L. Wang, Y. Zhao, Y. Yang, H. Wang, H. Zhang, R. G. Roeder, W.-G. Zhu // *Nucleic Acids Research*. — 2018. — July. — С. 1–15. — DOI: [10.1093/nar/gky568](https://doi.org/10.1093/nar/gky568).
210. Asymmetric Unwrapping of Nucleosomes under Tension Directed by DNA Local Flexibility / T. T. M. Ngo, Q. Zhang, R. Zhou, J. G. Yodh, T. Ha // *Cell*. — 2015. — Mar. 12. — Vol. 160, no. 6. — P. 1135–1144. — DOI: [10.1016/j.cell.2015.02.001](https://doi.org/10.1016/j.cell.2015.02.001).
211. Cryo-EM Structure of the Nucleosome Containing the ALB1 enhancer DNA Sequence. / Y. Takizawa, H. Tanaka, S. Machida, M. Koyama, K. Maehara, Y. Ohkawa, P. A. Wade, M. Wolf, H. Kurumizaka // *Open biology*. — 2018. — Т. 8, № 3. — С. 170255–170255. — DOI: [10.1098/rsob.170255](https://doi.org/10.1098/rsob.170255).
212. *Cui, F.* Structure-Based Analysis of DNA Sequence Patterns Guiding Nucleosome Positioning in Vitro / F. Cui, V. B. Zhurkin // *Journal of Biomolecular Structure & Dynamics*. — England, 2010. — June. — Vol. 27, no. 6. — P. 821–841. — DOI: [10.1080/073911010010524947](https://doi.org/10.1080/073911010010524947).

213. *Segal, E.* From DNA Sequence to Transcriptional Behaviour: A Quantitative Approach / E. Segal, J. Widom // Nature Reviews Genetics. — 2009. — T. 10, № 7. — C. 443–456. — DOI: [10.1038/nrg2591](https://doi.org/10.1038/nrg2591).
214. *Lequieu, J.* In Silico Evidence for Sequence-Dependent Nucleosome Sliding / J. Lequieu, D. C. Schwartz, J. J. de Pablo // Proceedings of the National Academy of Sciences of the United States of America. — 2017. — Oct. 31. — Vol. 114, no. 44. — E9197–E9205. — DOI: [10.1073/pnas.1705685114](https://doi.org/10.1073/pnas.1705685114).
215. Sequence-Dependent Nucleosome Sliding in Rotation-Coupled and Uncoupled Modes Revealed by Molecular Simulations / T. Niina, G. B. Brandani, C. Tan, S. Takada // PLOS Computational Biology. — 2017. — Dec. 1. — Vol. 13, no. 12. — e1005880. — DOI: [10.1371/journal.pcbi.1005880](https://doi.org/10.1371/journal.pcbi.1005880).
216. Nanoscale Dynamics of Centromere Nucleosomes and the Critical Roles of CENP-A / M. P. Stumme-Diers, S. Banerjee, M. Hashemi, Z. Sun, Y. L. Lyubchenko // Nucleic Acids Research. — 2018. — T. 46, № 1. — C. 94–103. — DOI: [10.1093/nar/gkx933](https://doi.org/10.1093/nar/gkx933).
217. Critical Role of Histone Turnover in Neuronal Transcription and Plasticity / I. Maze, W. Wenderski, K. M. Noh, R. C. Bagot, N. Tzavaras, I. Purushothaman, S. J. Elsässer, Y. Guo, C. Ionete, Y. L. Hurd, C. A. Tamminga, T. Halene, L. Farrelly, A. A. Soshnev, D. Wen, S. Rafii, M. R. Birtwistle, S. Akbarian, B. A. Buchholz, R. D. Blitzler, E. J. Nestler, Z. F. Yuan, B. A. Garcia, L. Shen, H. Molina, C. D. Allis // Neuron. — 2015. — T. 87, № 1. — C. 77–94. — DOI: [10.1016/j.neuron.2015.06.014](https://doi.org/10.1016/j.neuron.2015.06.014).
218. Histone H3.5 Forms an Unstable Nucleosome and Accumulates around Transcription Start Sites in Human Testis / T. Urahama, A. Harada, K. Maehara, N. Horikoshi, K. Sato, Y. Sato, K. Shiraishi, N. Sugino, A. Osakabe, H. Tachiwana, W. Kagawa, H. Kimura, Y. Ohkawa, H. Kurumizaka // Epigenetics and Chromatin. — 2016. — T. 9, № 1. — C. 1–16. — DOI: [10.1186/s13072-016-0051-y](https://doi.org/10.1186/s13072-016-0051-y).
219. Identification and Characterization of Two Novel Primate-Specific Histone H3 Variants, H3.X and H3.Y / S. M. Wiedemann, S. N. Mildner, C. B. Nisch, L. Israel, A. Maiser, S. Matheisl, T. Straub, R. Merkl, H. Leonhardt, E. Kremmer, L. Schermelleh, S. B. Hake // Journal of Cell Biology. — 2010. — T. 190, № 5. — C. 777–791. — DOI: [10.1083/jcb.201002043](https://doi.org/10.1083/jcb.201002043).



220. Identification of the Amino Acid Residues Responsible for Stable Nucleosome Formation by Histone H3.Y / T. Kujirai, N. Horikoshi, Y. Xie, H. Taguchi, H. Kurumizaka // *Nucleus*. — 2017. — T. 8, № 3. — C. 1–10. — DOI: [10.1080/19491034.2016.1277303](https://doi.org/10.1080/19491034.2016.1277303).
221. Histone Isoform H2A1H Promotes Attainment of Distinct Physiological States by Altering Chromatin Dynamics / S. Bhattacharya, D. Reddy, V. Jani, N. Gadewal, S. Shah, R. Reddy, K. Bose, U. Sonavane, R. Joshi, S. Gupta // *Epigenetics and Chromatin*. — 2017. — T. 10, № 1. — C. 1–19. — DOI: [10.1186/s13072-017-0155-z](https://doi.org/10.1186/s13072-017-0155-z).
222. Unique Yeast Histone Sequences Influence Octamer and Nucleosome Stability / A. Leung, M. Cheema, R. González-Romero, J. M. Eirin-Lopez, J. Ausió, C. J. Nelson // *FEBS Letters*. — 2016. — T. 590. — C. 2629–2638. — DOI: [10.1002/1873-3468.12266](https://doi.org/10.1002/1873-3468.12266).
223. *Bilokapic, S.* Structural Rearrangements of the Histone Octamer Translocate DNA / S. Bilokapic, M. Strauss, M. Halic // *Nature Communications*. — 2018. — Dec. — Vol. 9, no. 1. — DOI: [10.1038/s41467-018-03677-z](https://doi.org/10.1038/s41467-018-03677-z).
224. Structure and Dynamics in the Nucleosome Revealed by Solid-State NMR / X. Shi, C. Prasanna, T. Nagashima, T. Yamazaki, K. Pervushin, L. Nordenskiöld // *Angewandte Chemie - International Edition*. — 2018. — T. 57, № 31. — C. 9734–9738. — DOI: [10.1002/anie.201804707](https://doi.org/10.1002/anie.201804707).
225. Investigating the Dynamics of Destabilized Nucleosomes Using Methyl-TROSY NMR / J. L. Kitevski-LeBlanc, T. Yuwen, P. N. Dyer, J. Rudolph, K. Luger, L. E. Kay // *Journal of the American Chemical Society*. — 2018. — Apr. 11. — Vol. 140, no. 14. — P. 4774–4777. — DOI: [10.1021/jacs.8b00931](https://doi.org/10.1021/jacs.8b00931).
226. Solution Structure of the Isolated Histone H2A-H2B Heterodimer / Y. Moriwaki, T. Yamane, H. Ohtomo, M. Ikeguchi, J. I. Kurita, M. Sato, A. Nagadoi, H. Shimojo, Y. Nishimura // *Scientific Reports*. — 2016. — T. 6, February. — C. 1–11. — DOI: [10.1038/srep24999](https://doi.org/10.1038/srep24999).
227. H2A Histone-Fold and DNA Elements in Nucleosome Activate SWR1-Mediated H2A.Z Replacement in Budding Yeast / A. Ranjan, F. Wang, G. Mizuguchi, D. Wei, Y. Huang, C. Wu // *eLife*. — 2015. — June 27. — Vol. 4. — e06845. — DOI: [10.7554/eLife.06845](https://doi.org/10.7554/eLife.06845).

228. Crystal Structure of a Nucleosome Core Particle Containing the Variant Histone H2A.Z / R. K. Suto, M. J. Clarkson, D. J. Tremethick, K. Luger // *Nature Structural Biology*. — 2000. — Dec. — Vol. 7, no. 12. — P. 1121—1124. — DOI: [10.1038/81971](https://doi.org/10.1038/81971).
229. Electrostatic Interactions between Arginines and the Minor Groove in the Nucleosome / S. M. West, R. Rohs, R. S. Mann, B. Honig // *Journal of Biomolecular Structure and Dynamics*. — 2010. — T. 27, № 6. — C. 861—866. — DOI: [10.1080/07391102.2010.10508587](https://doi.org/10.1080/07391102.2010.10508587).
230. The Energetics of Specific Binding of AT-Hooks from HMGA1 to Target DNA / A. I. Dragan, J. R. Liggins, C. Crane-Robinson, P. L. Privalov // *Journal of Molecular Biology*. — 2003. — Mar. 21. — Vol. 327, no. 2. — P. 393—411.
231. DNA Shape Dominates Sequence Affinity in Nucleosome Formation / G. S. Freeman, J. P. Lequieu, D. M. Hinckley, J. K. Whitmer, J. J. De Pablo // *Physical Review Letters*. — 2014. — T. 113, № 16. — DOI: [10.1103/PhysRevLett.113.168101](https://doi.org/10.1103/PhysRevLett.113.168101).
232. *McGhee, J. D.* Nucleosome Structure / J. D. McGhee, G. Felsenfeld // *Annu Rev Biochem*. — 1980. — Vol. 49. — P. 1115—56. — DOI: [10.1146/annurev.bi.49.070180.005343](https://doi.org/10.1146/annurev.bi.49.070180.005343).
233. *Dechassa, M. L.* Nucleosomes as Control Elements for Accessing the Genome / M. L. Dechassa, K. Luger // *Genome Organization and Function in the Cell Nucleus*. — Wiley-VCH Verlag GmbH & Co. KGaA, 2011. — C. 55—87. — ISBN 978-3-527-63999-1. — DOI: [10.1002/9783527639991.ch3](https://doi.org/10.1002/9783527639991.ch3).
234. MMDB and VAST+: Tracking Structural Similarities between Macromolecular Complexes / T. Madej, C. J. Lanczycki, D. Zhang, P. A. Thiessen, R. C. Geer, A. Marchler-Bauer, S. H. Bryant // *Nucleic Acids Res*. — 2014. — Jan. — Vol. 42, Database issue. — P. D297—303. — DOI: [10.1093/nar/gkt1208](https://doi.org/10.1093/nar/gkt1208).
235. The Histone Database: An Integrated Resource for Histones and Histone Fold-Containing Proteins / L. Mariño-Ramírez, K. M. Levine, M. Morales, S. Zhang, R. T. Moreland, A. D. Baxevanis, D. Landsman // *Database: The Journal of Biological Databases and Curation*. — 2011. — 22 окт. — T. 2011. — DOI: [10.1093/database/bar048](https://doi.org/10.1093/database/bar048).

236. A Novel Roll-and-Slide Mechanism of DNA Folding in Chromatin: Implications for Nucleosome Positioning / M. Y. Tolstorukov, A. V. Colasanti, D. M. McCandlish, W. K. Olson, V. B. Zhurkin // *Journal of Molecular Biology*. — 2007. — Aug. 17. — Vol. 371, no. 3. — P. 725–738. — DOI: [10.1016/j.jmb.2007.05.048](https://doi.org/10.1016/j.jmb.2007.05.048).
237. *Fenley, A. T.* Charge State of the Globular Histone Core Controls Stability of the Nucleosome / A. T. Fenley, D. A. Adams, A. V. Onufriev // *Biophys J*. — 2010. — 8 сент. — Т. 99, № 5. — С. 1577–85. — DOI: [10.1016/j.bpj.2010.06.046](https://doi.org/10.1016/j.bpj.2010.06.046).
238. Solvent Mediated Interactions in the Structure of the Nucleosome Core Particle at 1.9 Å Resolution / C. A. Davey, D. F. Sargent, K. Luger, A. W. Maeder, T. J. Richmond // *Journal of Molecular Biology*. — 2002. — June 21. — Vol. 319, no. 5. — P. 1097–1113. — DOI: [10.1016/S0022-2836\(02\)00386-8](https://doi.org/10.1016/S0022-2836(02)00386-8).
239. *Henikoff, S.* Nucleosome Destabilization in the Epigenetic Regulation of Gene Expression / S. Henikoff // *Nat Rev Genet*. — 2008. — ЯНВ. — Т. 9, № 1. — С. 15–26. — DOI: [10.1038/nrg2206](https://doi.org/10.1038/nrg2206).
240. *Petty, E.* Balancing Chromatin Remodeling and Histone Modifications in Transcription / E. Petty, L. Pillus // *Trends Genet*. — 2013. — Ноябрь. — Т. 29, № 11. — С. 621–9. — DOI: [10.1016/j.tig.2013.06.006](https://doi.org/10.1016/j.tig.2013.06.006).
241. *Choy, J. S.* Structural Dynamics of Nucleosomes at Single-Molecule Resolution / J. S. Choy, T. H. Lee // *Trends Biochem Sci*. — 2012. — Oct. — Vol. 37, no. 10. — P. 425–35. — DOI: [10.1016/j.tibs.2012.06.006](https://doi.org/10.1016/j.tibs.2012.06.006).
242. *Lee, K. M.* The N-Terminal Tail of Histone H2A Binds to Two Distinct Sites within the Nucleosome Core / K. M. Lee, J. J. Hayes // *Proc Natl Acad Sci U S A*. — 1997. — Aug. 19. — Vol. 94, no. 17. — P. 8959–64.
243. *Mueller-Planitz, F.* Nucleosome Sliding Mechanisms: New Twists in a Looped History / F. Mueller-Planitz, H. Klinker, P. B. Becker // *Nature Structural & Molecular Biology*. — 2013. — Sept. — Vol. 20, no. 9. — P. 1026–1032. — DOI: [10.1038/nsmb.2648](https://doi.org/10.1038/nsmb.2648).
244. *Pepenella, S.* Intra- and Inter-Nucleosome Interactions of the Core Histone Tail Domains in Higher-Order Chromatin Structure / S. Pepenella, K. J. Murphy, J. J. Hayes // *Chromosoma*. — 2014. — Mar. — Vol. 123, no. 1/2. — P. 3–13. — DOI: [10.1007/s00412-013-0435-8](https://doi.org/10.1007/s00412-013-0435-8).

245. Histone H4 Tail Mediates Allosteric Regulation of Nucleosome Remodelling by Linker DNA / W. L. Hwang, S. Deindl, B. T. Harada, X. Zhuang // Nature. — 2014. — June 29. — DOI: [10.1038/nature13380](https://doi.org/10.1038/nature13380).
246. The Histone H4 Tail Regulates the Conformation of the ATP-Binding Pocket in the SNF2h Chromatin Remodeling Enzyme / L. R. Racki, N. Naber, E. Pate, J. D. Leonard, R. Cooke, G. J. Narlikar // J Mol Biol. — 2014. — May 15. — Vol. 426, no. 10. — P. 2034–44. — DOI: [10.1016/j.jmb.2014.02.021](https://doi.org/10.1016/j.jmb.2014.02.021).
247. Heterochromatin Protein Sir3 Induces Contacts between the Amino Terminus of Histone H4 and Nucleosomal DNA / F. Wang, G. Li, M. Altaf, C. Lu, M. A. Currie, A. Johnson, D. Moazed // Proc Natl Acad Sci U S A. — 2013. — May 21. — Vol. 110, no. 21. — P. 8495–500. — DOI: [10.1073/pnas.1300126110](https://doi.org/10.1073/pnas.1300126110).
248. *Rando, O. J.* Combinatorial Complexity in Chromatin Structure and Function: Revisiting the Histone Code / O. J. Rando // Curr Opin Genet Dev. — 2012. — Apr. — T. 22, № 2. — C. 148–55. — DOI: [10.1016/j.gde.2012.02.013](https://doi.org/10.1016/j.gde.2012.02.013).
249. *Biswas, M.* Atomistic Simulations of Nucleosomes / M. Biswas, J. Langowski, T. C. Bishop // Wiley Interdisciplinary Reviews: Computational Molecular Science. — 2013. — T. 3, № 4. — C. 378–392. — DOI: [10.1002/wcms.1139](https://doi.org/10.1002/wcms.1139).
250. *Arya, G.* Role of Histone Tails in Chromatin Folding Revealed by a Mesoscopic Oligonucleosome Model / G. Arya, T. Schlick // Proc Natl Acad Sci U S A. — 2006. — 31 окт. — T. 103, № 44. — C. 16236–41. — DOI: [10.1073/pnas.0604817103](https://doi.org/10.1073/pnas.0604817103).
251. *Collepardo-Guevara, R.* Chromatin Fiber Polymorphism Triggered by Variations of DNA Linker Lengths / R. Collepardo-Guevara, T. Schlick // Proc Natl Acad Sci U S A. — 2014. — June 3. — Vol. 111, no. 22. — P. 8061–6. — DOI: [10.1073/pnas.1315872111](https://doi.org/10.1073/pnas.1315872111).
252. The Role of Histone Tails in the Nucleosome: A Computational Study / J. Erler, R. Zhang, L. Petridis, X. Cheng, J. C. Smith, J. Langowski // Biophysical Journal. — 2014. — Dec. 16. — Vol. 107, no. 12. — P. 2911–2922. — DOI: [10.1016/j.bpj.2014.10.065](https://doi.org/10.1016/j.bpj.2014.10.065).

253. *Materese, C. K.* Counterion Atmosphere and Hydration Patterns near a Nucleosome Core Particle / C. K. Materese, A. Savelyev, G. A. Papoian // Journal of the American Chemical Society. — 2009. — Oct. 21. — Vol. 131, no. 41. — P. 15005–15013. — DOI: [10.1021/ja905376q](https://doi.org/10.1021/ja905376q).
254. *Ruscio, J. Z.* A Computational Study of Nucleosomal DNA Flexibility / J. Z. Ruscio, A. Onufriev // Biophys J. — 2006. — 1 дек. — Т. 91, № 11. — С. 4121–32. — DOI: [10.1529/biophysj.106.082099](https://doi.org/10.1529/biophysj.106.082099).
255. Multivalent Engagement of Chromatin Modifications by Linked Binding Modules / A. J. Ruthenburg, H. Li, D. J. Patel, C. D. Allis // Nature Reviews. Molecular Cell Biology. — 2007. — Dec. — Vol. 8, no. 12. — P. 983–994. — DOI: [10.1038/nrm2298](https://doi.org/10.1038/nrm2298).
256. Chromatin Fiber Folding: Requirement for the Histone H4 N-Terminal Tail / B. Dorigo, T. Schalch, K. Bystricky, T. J. Richmond // J Mol Biol. — 2003. — Mar. 14. — Vol. 327, no. 1. — P. 85–96.
257. Topological Diversity of Chromatin Fibers: Interplay between Nucleosome Repeat Length, DNA Linking Number and the Level of Transcription / D. Norouzi, A. Katebi, F. Cui, V. B. Zhurkin // AIMS biophysics. — 2015. — Vol. 2, no. 4. — P. 613–629. — DOI: [10.3934/biophy.2015.4.613](https://doi.org/10.3934/biophy.2015.4.613).
258. Characterization of the N-Terminal Tail Domain of Histone H3 in Condensed Nucleosome Arrays by Hydrogen Exchange and NMR / H. Kato, J. Gruschus, R. Ghirlando, N. Tjandra, Y. Bai // J Am Chem Soc. — 2009. — Oct. 28. — Vol. 131, no. 42. — P. 15104–5. — DOI: [10.1021/ja9070078](https://doi.org/10.1021/ja9070078).
259. Histone H4 K16Q Mutation, an Acetylation Mimic, Causes Structural Disorder of Its N-Terminal Basic Patch in the Nucleosome / B. R. Zhou, H. Feng, R. Ghirlando, H. Kato, J. Gruschus, Y. Bai // J Mol Biol. — 2012. — Aug. 3. — Vol. 421, no. 1. — P. 30–7. — DOI: [10.1016/j.jmb.2012.04.032](https://doi.org/10.1016/j.jmb.2012.04.032).
260. Histone H3 and H4 N-Terminal Tails in Nucleosome Arrays at Cellular Concentrations Probed by Magic Angle Spinning NMR Spectroscopy / M. Gao, P. S. Nadaud, M. W. Bernier, J. A. North, P. C. Hammel, M. G. Poirier, C. P. Jaroniec // J Am Chem Soc. — 2013. — Oct. 16. — Vol. 135, no. 41. — P. 15278–81. — DOI: [10.1021/ja407526s](https://doi.org/10.1021/ja407526s).

261. ANP32E Is a Histone Chaperone That Removes H2A.Z from Chromatin / A. Obri, K. Ouararhni, C. Papin, M. L. Diebold, K. Padmanabhan, M. Marek, I. Stoll, L. Roy, P. T. Reilly, T. W. Mak, S. Dimitrov, C. Romier, A. Hamiche // *Nature*. — 2014. — Jan. 30. — Vol. 505, no. 7485. — P. 648–53. — DOI: [10.1038/nature12922](https://doi.org/10.1038/nature12922).
262. Role of Histone Tails in Structural Stability of the Nucleosome / M. Biswas, K. Voltz, J. C. Smith, J. Langowski // *PLoS Comput Biol*. — 2011. — Дек. — Т. 7, № 12. — e1002279. — DOI: [10.1371/journal.pcbi.1002279](https://doi.org/10.1371/journal.pcbi.1002279).
263. *Pachov, G. V.* On the Structure and Dynamics of the Complex of the Nucleosome and the Linker Histone / G. V. Pachov, R. R. Gabdoulline, R. C. Wade // *Nucleic Acids Res*. — 2011. — July. — Vol. 39, no. 12. — P. 5255–63. — DOI: [10.1093/nar/gkr101](https://doi.org/10.1093/nar/gkr101).
264. *Toth, K.* Chromatin Compaction at the Mononucleosome Level / K. Toth, N. Brun, J. Langowski // *Biochemistry*. — 2006. — Feb. 14. — Vol. 45, no. 6. — P. 1591–8. — DOI: [10.1021/bi052110u](https://doi.org/10.1021/bi052110u).
265. *Mirny, L. A.* Nucleosome-Mediated Cooperativity between Transcription Factors / L. A. Mirny // *Proc Natl Acad Sci U S A*. — 2010. — Dec. 28. — Vol. 107, no. 52. — P. 22534–9. — DOI: [10.1073/pnas.0913805107](https://doi.org/10.1073/pnas.0913805107).
266. *Tomschik, M.* Nucleosome Dynamics as Studied by Single-Pair Fluorescence Resonance Energy Transfer: A Reevaluation / M. Tomschik, K. van Holde, J. Zlatanova // *J Fluoresc*. — 2009. — Jan. — Vol. 19, no. 1. — P. 53–62. — DOI: [10.1007/s10895-008-0379-1](https://doi.org/10.1007/s10895-008-0379-1).
267. Nucleosome Disassembly Intermediates Characterized by Single-Molecule FRET / A. Gansen, A. Valeri, F. Hauger, S. Felekyan, S. Kalinin, K. Toth, J. Langowski, C. A. Seidel // *Proc Natl Acad Sci U S A*. — 2009. — Sept. 8. — Vol. 106, no. 36. — P. 15308–13. — DOI: [10.1073/pnas.0903005106](https://doi.org/10.1073/pnas.0903005106).
268. Clipping of Flexible Tails of Histones H3 and H4 Affects the Structure and Dynamics of the Nucleosome / N. P. Nurse, I. Jimenez-Useche, I. T. Smith, C. Yuan // *Biophys J*. — 2013. — 5 марта. — Т. 104, № 5. — С. 1081–8. — DOI: [10.1016/j.bpj.2013.01.019](https://doi.org/10.1016/j.bpj.2013.01.019).

269. The Influence of DNA Shape Fluctuations on Fluorescence Resonance Energy Transfer Efficiency Measurements in Nucleosomes / L. Lenz, M. Hoenderdos, P. Prinsen, H. Schiessel // *J Phys Condens Matter*. — 2015. — Feb. 18. — Vol. 27, no. 6. — P. 064104. — DOI: [10.1088/0953-8984/27/6/064104](https://doi.org/10.1088/0953-8984/27/6/064104).
270. *Roccatano, D.* Structural Flexibility of the Nucleosome Core Particle at Atomic Resolution Studied by Molecular Dynamics Simulation / D. Roccatano, A. Barthel, M. Zacharias // *Biopolymers*. — 2007. — 5 apr. — T. 85, № 5/6. — C. 407—21. — DOI: [10.1002/bip.20690](https://doi.org/10.1002/bip.20690).
271. Salt-Induced Conformation and Interaction Changes of Nucleosome Core Particles / S. Mangenot, A. Leforestier, P. Vachette, D. Durand, F. Livolant // *Biophys J*. — 2002. — Jan. — Vol. 82, 1 Pt 1. — P. 345—56. — DOI: [10.1016/s0006-3495\(02\)75399-x](https://doi.org/10.1016/s0006-3495(02)75399-x).
272. Biophysical Analysis and Small-Angle X-Ray Scattering-Derived Structures of MeCP2-Nucleosome Complexes / C. Yang, M. J. van der Woerd, U. M. Muthurajan, J. C. Hansen, K. Luger // *Nucleic Acids Res*. — 2011. — May. — Vol. 39, no. 10. — P. 4122—35. — DOI: [10.1093/nar/gkr005](https://doi.org/10.1093/nar/gkr005).
273. Reconstitution of Chromatin: Assembly of the Nucleosome / F. X. Wilhelm, M. L. Wilhelm, M. Erard, M. P. Duane // *Nucleic Acids Res*. — 1978. — Feb. — Vol. 5, no. 2. — P. 505—21.
274. *Yoo, J.* Improved Parametrization of Li<sup>+</sup>, Na<sup>+</sup>, K<sup>+</sup>, and Mg<sup>2+</sup> Ions for All-Atom Molecular Dynamics Simulations of Nucleic Acid Systems / J. Yoo, A. Aksimentiev // *The Journal of Physical Chemistry Letters*. — 2012. — 5 янв. — Т. 3, № 1. — С. 45—50. — DOI: [10.1021/jz201501a](https://doi.org/10.1021/jz201501a).
275. *Iwaki, T.* How Are Small Ions Involved in the Compaction of DNA Molecules? / T. Iwaki, T. Saito, K. Yoshikawa // *Colloids and Surfaces B: Biointerfaces*. — 2007. — Т. 56, № 1/2. — С. 126—133. — DOI: [10.1016/j.colsurfb.2006.10.040](https://doi.org/10.1016/j.colsurfb.2006.10.040).
276. *Korolev, N.* Physicochemical Analysis of Electrostatic Foundation for DNA-Protein Interactions in Chromatin Transformations / N. Korolev, O. V. Vorontsova, L. Nordenskiöld // *Prog Biophys Mol Biol*. — 2007. — Sept. — Vol. 95, no. 1—3. — P. 23—49. — DOI: [10.1016/j.pbiomolbio.2006.11.003](https://doi.org/10.1016/j.pbiomolbio.2006.11.003).

277. *Dehghani, H.* Organization of Chromatin in the Interphase Mammalian Cell / H. Dehghani, G. Dellaire, D. P. Bazett-Jones // *Micron*. — 2005. — T. 36, № 2. — С. 95—108. — DOI: [10.1016/j.micron.2004.10.003](https://doi.org/10.1016/j.micron.2004.10.003).
278. Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions / M. Shogren-Knaak, H. Ishii, J.-M. Sun, M. J. Pazin, J. R. Davie, C. L. Peterson // *Science*. — 2006. — Feb. 10. — Vol. 311, no. 5762. — P. 844—847. — DOI: [10.1126/science.1124000](https://doi.org/10.1126/science.1124000).
279. *Weng, L.* Probing Interactions between Lysine Residues in Histone Tails and Nucleosomal DNA via Product and Kinetic Analysis / L. Weng, C. Zhou, M. M. Greenberg // *ACS Chem Biol*. — 2014. — Dec. 5. — DOI: [10.1021/cb500737y](https://doi.org/10.1021/cb500737y).
280. Chromatin Unfolding by Epigenetic Modifications Explained by Dramatic Impairment of Internucleosome Interactions: A Multiscale Computational Study / R. Collepardo-Guevara, G. Portella, M. Vendruscolo, D. Frenkel, T. Schlick, M. Orozco // *Journal of the American Chemical Society*. — 2015. — Aug. 19. — Vol. 137, no. 32. — P. 10205—10215. — DOI: [10.1021/jacs.5b04086](https://doi.org/10.1021/jacs.5b04086).
281. *Potoyan, D. A.* Energy Landscape Analyses of Disordered Histone Tails Reveal Special Organization of Their Conformational Dynamics / D. A. Potoyan, G. A. Papoian // *J Am Chem Soc*. — 2011. — 18 мая. — T. 133, № 19. — С. 7405—15. — DOI: [10.1021/ja1111964](https://doi.org/10.1021/ja1111964).
282. *Korolev, N.* H4 Histone Tail Mediated DNA-DNA Interaction and Effects on DNA Structure, Flexibility, and Counterion Binding: A Molecular Dynamics Study / N. Korolev, L. Nordenskiöld // *Biopolymers*. — 2007. — 5 авг. — T. 86, № 5/6. — С. 409—23. — DOI: [10.1002/bip.20749](https://doi.org/10.1002/bip.20749).
283. Molecular Dynamics Simulations Demonstrate the Regulation of DNA-DNA Attraction by H4 Histone Tail Acetylations and Mutations / N. Korolev, H. Yu, A. P. Lyubartsev, L. Nordenskiöld // *Biopolymers*. — 2014. — Oct. — Vol. 101, no. 10. — P. 1051—64. — DOI: [10.1002/bip.22499](https://doi.org/10.1002/bip.22499).
284. The Role of DNA Shape in Protein-DNA Recognition / R. Rohs, S. M. West, A. Sosinsky, P. Liu, R. S. Mann, B. Honig // *Nature*. — 2009. — 29 окт. — T. 461, № 7268. — С. 1248—53. — DOI: [10.1038/nature08473](https://doi.org/10.1038/nature08473).



285. *Yang, Z.* The Core Histone Tail Domains Contribute to Sequence-Dependent Nucleosome Positioning / Z. Yang, C. Zheng, J. J. Hayes // *J Biol Chem.* — 2007. — 16 марта. — Т. 282, № 11. — С. 7930—8. — DOI: [10.1074/jbc.M610584200](https://doi.org/10.1074/jbc.M610584200).
286. The Docking Domain of Histone H2A Is Required for H1 Binding and RSC-Mediated Nucleosome Remodeling / M. S. Shukla, S. H. Syed, D. Goutte-Gattat, J. L. Richard, F. Montel, A. Hamiche, A. Travers, C. Faivre-Moskalenko, J. Bednar, J. J. Hayes, D. Angelov, S. Dimitrov // *Nucleic Acids Res.* — 2011. — Apr. — Vol. 39, no. 7. — P. 2559—70. — DOI: [10.1093/nar/gkq1174](https://doi.org/10.1093/nar/gkq1174).
287. High Resolution Dynamic Mapping of Histone-DNA Interactions in a Nucleosome / M. A. Hall, A. Shundrovsky, L. Bai, R. M. Fulbright, J. T. Lis, M. D. Wang // *Nature structural & molecular biology.* — 2009. — Февр. — Т. 16, № 2. — С. 124—129. — DOI: [10.1038/nsmb.1526](https://doi.org/10.1038/nsmb.1526).
288. Mechanical Disruption of Individual Nucleosomes Reveals a Reversible Multistage Release of DNA / B. D. Brower-Toland, C. L. Smith, R. C. Yeh, J. T. Lis, C. L. Peterson, M. D. Wang // *Proceedings of the National Academy of Sciences of the United States of America.* — 2002. — Feb. 19. — Vol. 99, no. 4. — P. 1960—1965. — DOI: [10.1073/pnas.022638399](https://doi.org/10.1073/pnas.022638399).
289. *Davey, C. A.* Does the Nucleosome Break Its Own Rules? / C. A. Davey // *Curr Opin Struct Biol.* — 2013. — Apr. — Vol. 23, no. 2. — P. 311—3. — DOI: [10.1016/j.sbi.2013.01.011](https://doi.org/10.1016/j.sbi.2013.01.011).
290. Mechanism of Transcription through the Nucleosome by Eukaryotic RNA Polymerase / V. M. Studitsky, G. A. Kassavetis, E. P. Geiduschek, G. Felsenfeld // *Science.* — 1997. — Dec. 12. — Vol. 278, no. 5345. — P. 1960—3.
291. Synergistic Action of RNA Polymerases in Overcoming the Nucleosomal Barrier / J. Jin, L. Bai, D. S. Johnson, R. M. Fulbright, M. L. Kireeva, M. Kashlev, M. D. Wang // *Nat Struct Mol Biol.* — 2010. — ИЮНЬ. — Т. 17, № 6. — С. 745—52. — DOI: [10.1038/nsmb.1798](https://doi.org/10.1038/nsmb.1798).
292. *Lee, K. M.* Linker DNA and H1-Dependent Reorganization of Histone-DNA Interactions within the Nucleosome / K. M. Lee, J. J. Hayes // *Biochemistry.* — 1998. — June 16. — Vol. 37, no. 24. — P. 8622—8. — DOI: [10.1021/bi980499y](https://doi.org/10.1021/bi980499y).

293. Preferential Interaction of the Core Histone Tail Domains with Linker DNA / D. Angelov, J. M. Vitolo, V. Mutskov, S. Dimitrov, J. J. Hayes // Proc Natl Acad Sci U S A. — 2001. — 5 ИЮНЯ. — Т. 98, № 12. — С. 6599–604. — DOI: [10.1073/pnas.121171498](https://doi.org/10.1073/pnas.121171498).
294. Laser-Induced Crosslinking of Histones to DNA in Chromatin and Core Particles: Implications in Studying Histone-DNA Interactions / V. Stefanovsky, S. I. Dimitrov, V. R. Russanova, D. Angelov, I. G. Pashev // Nucleic Acids Res. — 1989. — Dec. 11. — Vol. 17, no. 23. — P. 10069–81.
295. Interplay among Nucleosomal DNA, Histone Tails, and Corepressor CoREST Underlies LSD1-Mediated H3 Demethylation / S. Pilotto, V. Speranzini, M. Tortorici, D. Durand, A. Fish, S. Valente, F. Forneris, A. Mai, T. K. Sixma, P. Vachette, A. Mattevi // Proc Natl Acad Sci U S A. — 2015. — Mar. 3. — Vol. 112, no. 9. — P. 2752–7. — DOI: [10.1073/pnas.1419468112](https://doi.org/10.1073/pnas.1419468112).
296. *Nishi, H.* Crosstalk between Signaling Pathways Provided by Single and Multiple Protein Phosphorylation Sites / H. Nishi, E. Demir, A. R. Panchenko // J Mol Biol. — 2015. — Jan. 30. — Vol. 427, no. 2. — P. 511–20. — DOI: [10.1016/j.jmb.2014.11.001](https://doi.org/10.1016/j.jmb.2014.11.001).
297. *Macke Thomas, J.* Modeling Unusual Nucleic Acid Structures / J. Macke Thomas, A. Case David // Molecular Modeling of Nucleic Acids : В 0 Т. Т. 682. — American Chemical Society, 1997. — С. 379–393. — (ACS Symposium Series ; 682). — ISBN 0-8412-3541-4. — DOI: [10.1021/bk-1998-0682.ch024](https://doi.org/10.1021/bk-1998-0682.ch024).
298. Reconstitution of Nucleosome Core Particles from Recombinant Histones and DNA / P. N. Dyer, R. S. Edayathumangalam, C. L. White, Y. Bao, S. Chakravarthy, U. M. Muthurajan, K. Luger // Methods Enzymol. — 2004. — Vol. 375. — P. 23–44.
299. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone Phi, Psi and Side-Chain Chi(1) and Chi(2) Dihedral Angles / R. B. Best, X. Zhu, J. Shim, P. E. Lopes, J. Mittal, M. Feig, A. D. Mackerell Jr. // J Chem Theory Comput. — 2012. — 11 СЕНТ. — Т. 8, № 9. — С. 3257–3273. — DOI: [10.1021/ct300400x](https://doi.org/10.1021/ct300400x).

300. Optimization of the CHARMM Additive Force Field for DNA: Improved Treatment of the BI/BII Conformational Equilibrium / K. Hart, N. Foloppe, C. M. Baker, E. J. Denning, L. Nilsson, A. D. MacKerell // Journal of Chemical Theory and Computation. — 2012. — Jan. 10. — Vol. 8, no. 1. — P. 348–362. — DOI: [10.1021/ct200723y](https://doi.org/10.1021/ct200723y).
301. *Luo, Y.* Simulation of Osmotic Pressure in Concentrated Aqueous Salt Solutions / Y. Luo, B. Roux // The Journal of Physical Chemistry Letters. — 2010. — 7 янв. — Т. 1, № 1. — С. 183–189. — DOI: [10.1021/jz900079w](https://doi.org/10.1021/jz900079w).
302. 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, no. 2. — e32131. — DOI: [10.1371/journal.pone.0032131](https://doi.org/10.1371/journal.pone.0032131).
303. *Galindo-Murillo, R.* Convergence and Reproducibility in Molecular Dynamics Simulations of the DNA Duplex d(GCACGAACGAACGAACGC) / R. Galindo-Murillo, D. R. Roe, T. E. Cheatham 3rd // Biochim Biophys Acta. — 2015. — May. — Vol. 1850, no. 5. — P. 1041–58. — DOI: [10.1016/j.bbagen.2014.09.007](https://doi.org/10.1016/j.bbagen.2014.09.007).
304. *Perez, A.* Frontiers in Molecular Dynamics Simulations of DNA / A. Perez, F. J. Luque, M. Orozco // Acc Chem Res. — 2012. — 21 февр. — Т. 45, № 2. — С. 196–205. — DOI: [10.1021/ar2001217](https://doi.org/10.1021/ar2001217).
305. *Humphrey, W.* VMD: Visual Molecular Dynamics / W. Humphrey, A. Dalke, K. Schulten // Journal of Molecular Graphics. — 1996. — Feb. — Vol. 14, no. 1. — P. 33–38, 27–28. — DOI: [10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).
306. Scalable Molecular Dynamics with NAMD / J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, K. Schulten // J Comput Chem. — 2005. — Дек. — Т. 26, № 16. — С. 1781–802. — DOI: [10.1002/jcc.20289](https://doi.org/10.1002/jcc.20289).
307. A Variety of DNA-Binding and Multimeric Proteins Contain the Histone Fold Motif / A. D. Baxevanis, G. Arents, E. N. Moudrianakis, D. Landsman // Nucleic Acids Res. — 1995. — 25 июля. — Т. 23, № 14. — С. 2685–91.

308. Crystallographic Structure of the Octameric Histone Core of the Nucleosome at a Resolution of 3.3 Å / R. W. Burlingame, W. E. Love, B. C. Wang, R. Hamlin, H. X. Nguyen, E. N. Moudrianakis // *Science* (New York, N.Y.) — 1985. — May 3. — Vol. 228, no. 4699. — P. 546–553. — DOI: [10.1126/science.3983639](https://doi.org/10.1126/science.3983639).
309. *Onufriev, A. V.* The Nucleosome: From Structure to Function through Physics / A. V. Onufriev, H. Schiessel // *Current Opinion in Structural Biology*. — 2019. — Jan. 30. — Vol. 56. — P. 119–130. — DOI: [10.1016/j.sbi.2018.11.003](https://doi.org/10.1016/j.sbi.2018.11.003).
310. Sequence Motifs and Free Energies of Selected Natural and Non-Natural Nucleosome Positioning DNA Sequences / A. Thåström, P. T. Lowary, H. R. Widlund, H. Cao, M. Kubista, J. Widom // *Journal of Molecular Biology*. — 1999. — Apr. 30. — Vol. 288, no. 2. — P. 213–229. — DOI: [10.1006/jmbi.1999.2686](https://doi.org/10.1006/jmbi.1999.2686).
311. The Mechanics behind DNA Sequence-Dependent Properties of the Nucleosome / E. Y. D. Chua, D. Vasudevan, G. E. Davey, B. Wu, C. A. Davey // *Nucleic Acids Research*. — 2012. — July. — Vol. 40, no. 13. — P. 6338–6352. — DOI: [10.1093/nar/gks261](https://doi.org/10.1093/nar/gks261).
312. *Zhurkin, V. B.* Sequence-Dependent Bending of DNA and Phasing of Nucleosomes / V. B. Zhurkin // *Journal of Biomolecular Structure & Dynamics*. — 1985. — Feb. — Vol. 2, no. 4. — P. 785–804. — DOI: [10.1080/07391102.1985.10506324](https://doi.org/10.1080/07391102.1985.10506324).
313. A Genomic Code for Nucleosome Positioning / E. Segal, Y. Fondufe-Mittendorf, L. Chen, A. Thåström, Y. Field, I. K. Moore, J.-P. Z. Wang, J. Widom // *Nature*. — 2006. — Aug. 17. — Vol. 442, no. 7104. — P. 772–778. — DOI: [10.1038/nature04979](https://doi.org/10.1038/nature04979).
314. *Zhao, Y.* Comprehensive Catalog of Currently Documented Histone Modifications / Y. Zhao, B. A. Garcia // *Cold Spring Harbor Perspectives in Biology*. — 2015. — Sept. 1. — Vol. 7, no. 9. — a025064. — DOI: [10.1101/cshperspect.a025064](https://doi.org/10.1101/cshperspect.a025064).

315. Replication-Dependent Histone Isoforms: A New Source of Complexity in Chromatin Structure and Function / R. Singh, E. Bassett, A. Chakravarti, M. R. Parthun // *Nucleic Acids Research*. — 2018. — Sept. 28. — Vol. 46, no. 17. — P. 8665–8678. — DOI: [10.1093/nar/gky768](https://doi.org/10.1093/nar/gky768).
316. The Expanding Landscape of 'oncohistone' Mutations in Human Cancers / B. A. Nacev, L. Feng, J. D. Bagert, A. E. Lemiesz, J. Gao, A. A. Soshnev, R. Kundra, N. Schultz, T. W. Muir, C. D. Allis // *Nature*. — 2019. — Mar. — Vol. 567, no. 7749. — P. 473–478. — DOI: [10.1038/s41586-019-1038-1](https://doi.org/10.1038/s41586-019-1038-1).
317. *Paul, S.* Regulation of ATP-Dependent Chromatin Remodelers: Accelerators/Brakes, Anchors and Sensors / S. Paul, B. Bartholomew // *Biochemical Society Transactions*. — 2018. — Dec. 17. — Vol. 46, no. 6. — P. 1423–1430. — DOI: [10.1042/BST20180043](https://doi.org/10.1042/BST20180043).
318. *Kujirai, T.* Transcription through the Nucleosome / T. Kujirai, H. Kuruizaka // *Current Opinion in Structural Biology*. — 2020. — Apr. — Vol. 61. — P. 42–49. — DOI: [10.1016/j.sbi.2019.10.007](https://doi.org/10.1016/j.sbi.2019.10.007).
319. High Precision FRET Studies Reveal Reversible Transitions in Nucleosomes between Microseconds and Minutes / A. Gansen, S. Felekyan, R. Kühnemuth, K. Lehmann, K. Tóth, C. A. M. Seidel, J. Langowski // *Nature Communications*. — 2018. — Nov. 6. — Vol. 9, no. 1. — P. 1–13. — DOI: [10.1038/s41467-018-06758-1](https://doi.org/10.1038/s41467-018-06758-1).
320. Direct Observation of Coordinated DNA Movements on the Nucleosome during Chromatin Remodelling / A. Sabantsev, R. F. Levendosky, X. Zhuang, G. D. Bowman, S. Deindl // *Nature Communications*. — 2019. — Apr. 12. — Vol. 10, no. 1. — P. 1720. — DOI: [10.1038/s41467-019-09657-1](https://doi.org/10.1038/s41467-019-09657-1).
321. Mechanism of DNA Translocation Underlying Chromatin Remodelling by Snf2 / M. Li, X. Xia, Y. Tian, Q. Jia, X. Liu, Y. Lu, M. Li, X. Li, Z. Chen // *Nature*. — 2019. — Mar. — Vol. 567, no. 7748. — P. 409–413. — DOI: [10.1038/s41586-019-1029-2](https://doi.org/10.1038/s41586-019-1029-2).
322. HP1 Reshapes Nucleosome Core to Promote Phase Separation of Heterochromatin / S. Sanulli, M. J. Trnka, V. Dharmarajan, R. W. Tibble, B. D. Pascal, A. L. Burlingame, P. R. Griffin, J. D. Gross, G. J. Narlikar // *Nature*. — 2019. — Nov. — Vol. 575, no. 7782. — P. 390–394. — DOI: [10.1038/s41586-019-1669-2](https://doi.org/10.1038/s41586-019-1669-2).

323. *Bowman, G. D.* Remodeling the Genome with DNA Twists / G. D. Bowman, S. Deindl // *Science* (New York, N.Y.) — 2019. — Oct. 4. — Vol. 366, no. 6461. — P. 35–36. — DOI: [10.1126/science.aay4317](https://doi.org/10.1126/science.aay4317).
324. Structures of the ISWI–Nucleosome Complex Reveal a Conserved Mechanism of Chromatin Remodeling / L. Yan, H. Wu, X. Li, N. Gao, Z. Chen // *Nature Structural & Molecular Biology*. — 2019. — Apr. — Vol. 26, no. 4. — P. 258. — DOI: [10.1038/s41594-019-0199-9](https://doi.org/10.1038/s41594-019-0199-9).
325. Cryo-EM Structures of Remodeler-Nucleosome Intermediates Suggest Allosteric Control through the Nucleosome / J. P. Armache, N. Gamarra, S. L. Johnson, J. D. Leonard, S. Wu, G. J. Narlikar, Y. Cheng // *eLife*. — 2019. — June 18. — Vol. 8. — DOI: [10.7554/eLife.46057](https://doi.org/10.7554/eLife.46057).
326. Dissecting DNA-Histone Interactions in the Nucleosome by Molecular Dynamics Simulations of DNA Unwrapping / R. Ettig, N. Kepper, R. Stehr, G. Wedemann, K. Rippe // *Biophysical Journal*. — 2011. — Oct. 19. — Vol. 101, no. 8. — P. 1999–2008. — DOI: [10.1016/j.bpj.2011.07.057](https://doi.org/10.1016/j.bpj.2011.07.057).
327. Exploring the Free Energy Landscape of Nucleosomes / B. Zhang, W. Zheng, G. A. Papoian, P. G. Wolynes // *Journal of the American Chemical Society*. — 2016. — July 6. — Vol. 138, no. 26. — P. 8126–8133. — DOI: [10.1021/jacs.6b02893](https://doi.org/10.1021/jacs.6b02893).
328. *Winogradoff, D.* Molecular Mechanism of Spontaneous Nucleosome Unraveling / D. Winogradoff, A. Aksimentiev // *Journal of Molecular Biology*. — 2019. — Jan. 18. — Vol. 431, no. 2. — P. 323–335. — DOI: [10.1016/j.jmb.2018.11.013](https://doi.org/10.1016/j.jmb.2018.11.013).
329. DNA Sliding in Nucleosomes via Twist Defect Propagation Revealed by Molecular Simulations / G. B. Brandani, T. Niina, C. Tan, S. Takada // *Nucleic Acids Research*. — 2018. — Apr. 6. — Vol. 46, no. 6. — P. 2788–2801. — DOI: [10.1093/nar/gky158](https://doi.org/10.1093/nar/gky158).
330. *Chakraborty, K.* Molecular Mechanism for the Role of the H2A and H2B Histone Tails in Nucleosome Repositioning / K. Chakraborty, M. Kang, S. M. Loverde // *The Journal of Physical Chemistry. B*. — 2018. — Dec. 20. — Vol. 122, no. 50. — P. 11827–11840. — DOI: [10.1021/acs.jpcc.8b07881](https://doi.org/10.1021/acs.jpcc.8b07881).

331. The Conformation of the Histone H3 Tail Inhibits Association of the BPTF PHD Finger with the Nucleosome / E. A. Morrison, S. Bowerman, K. L. Sylvers, J. Wereszczynski, C. A. Musselman // *eLife*. — 2018. — Apr. 12. — Vol. 7. — DOI: [10.7554/eLife.31481](https://doi.org/10.7554/eLife.31481).
332. *Li, Z.* Investigating the Influence of Arginine Dimethylation on Nucleosome Dynamics Using All-Atom Simulations and Kinetic Analysis / Z. Li, H. Kono // *The Journal of Physical Chemistry. B*. — 2018. — Oct. 25. — Vol. 122, no. 42. — P. 9625–9634. — DOI: [10.1021/acs.jpccb.8b05067](https://doi.org/10.1021/acs.jpccb.8b05067).
333. Structural Dynamics of Nucleosome Mediated by Acetylations at H3K56 and H3K115,122 / M. Rajagopalan, S. Balasubramanian, I. Ioshikhes, A. Ramaswamy // *European biophysics journal: EBJ*. — 2017. — July. — Vol. 46, no. 5. — P. 471–484. — DOI: [10.1007/s00249-016-1191-5](https://doi.org/10.1007/s00249-016-1191-5).
334. *Bowerman, S.* Unique Dynamics in Asymmetric macroH2A-H2A Hybrid Nucleosomes Result in Increased Complex Stability / S. Bowerman, R. J. Hickok, J. Wereszczynski // *The Journal of Physical Chemistry. B*. — 2019. — Jan. 17. — Vol. 123, no. 2. — P. 419–427. — DOI: [10.1021/acs.jpccb.8b10668](https://doi.org/10.1021/acs.jpccb.8b10668).
335. *Sun, R.* TMB Library of Nucleosome Simulations / R. Sun, Z. Li, T. C. Bishop // *Journal of Chemical Information and Modeling*. — 2019. — Oct. 28. — Vol. 59, no. 10. — P. 4289–4299. — DOI: [10.1021/acs.jcim.9b00252](https://doi.org/10.1021/acs.jcim.9b00252).
336. Chromatosome Structure and Dynamics from Molecular Simulations / M. A. Öztürk, M. De, V. Cojocaru, R. C. Wade // *Annual Review of Physical Chemistry*. — 2020. — Apr. 20. — Vol. 71. — P. 101–119. — DOI: [10.1146/annurev-physchem-071119-040043](https://doi.org/10.1146/annurev-physchem-071119-040043).
337. *Perišić, O.* Sensitive Effect of Linker Histone Binding Mode and Subtype on Chromatin Condensation / O. Perišić, S. Portillo-Ledesma, T. Schlick // *Nucleic Acids Research*. — 2019. — June 4. — Vol. 47, no. 10. — P. 4948–4957. — DOI: [10.1093/nar/gkz234](https://doi.org/10.1093/nar/gkz234).
338. 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. — 1 сент. — Т. 1/2. — С. 19–25. — DOI: [10.1016/j.softx.2015.06.001](https://doi.org/10.1016/j.softx.2015.06.001).

339. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB / J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling // *Journal of chemical theory and computation*. — 2015. — 11 авг. — Т. 11, № 8. — С. 3696–3713. — DOI: [10.1021/acs.jctc.5b00255](https://doi.org/10.1021/acs.jctc.5b00255).
340. Parmbsc1: A Refined Force Field for DNA Simulations / I. Ivani, P. D. Dans, A. Noy, A. Pérez, I. Faustino, A. Hospital, J. Walther, P. Andrio, R. Goñi, A. Balaceanu, G. Portella, F. Battistini, J. L. Gelpí, C. González, M. Vendruscolo, C. A. Laughton, S. A. Harris, D. A. Case, M. Orozco // *Nature Methods*. — 2016. — Jan. — Vol. 13, no. 1. — P. 55–58. — DOI: [10.1038/nmeth.3658](https://doi.org/10.1038/nmeth.3658).
341. Yoo, J. New Tricks for Old Dogs: Improving the Accuracy of Biomolecular Force Fields by Pair-Specific Corrections to Non-Bonded Interactions / J. Yoo, A. Aksimentiev // *Physical Chemistry Chemical Physics*. — 2018. — Mar. 28. — Vol. 20, no. 13. — P. 8432–8449. — DOI: [10.1039/C7CP08185E](https://doi.org/10.1039/C7CP08185E).
342. Comparison of Simple Potential Functions for Simulating Liquid Water / W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein // *The Journal of Chemical Physics*. — 1983. — July 15. — Vol. 79, no. 2. — P. 926–935. — DOI: [10.1063/1.445869](https://doi.org/10.1063/1.445869).
343. Supercomputer Lomonosov-2: Large Scale, Deep Monitoring and Fine Analytics for the User Community / V. V. Voevodin, A. S. Antonov, D. A. Nikitenko, P. A. Shvets, S. I. Sobolev, I. Y. Sidorov, K. S. Stefanov, V. V. Voevodin, S. A. Zhumatiy // *Supercomputing Frontiers and Innovations*. — 2019. — June 26. — Vol. 6, issue 2, no. 2. — P. 4–11. — DOI: [10.14529/jsfi190201](https://doi.org/10.14529/jsfi190201).
344. 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, no. 10. — P. 2319–2327. — DOI: [10.1002/jcc.21787](https://doi.org/10.1002/jcc.21787).
345. Sin Mutations Alter Inherent Nucleosome Mobility / A. Flaus, C. Rencurel, H. Ferreira, N. Wiechens, T. Owen-Hughes // *The EMBO Journal*. — 2004. — 28 янв. — Т. 23, № 2. — С. 343–353. — DOI: [10.1038/sj.emboj.7600047](https://doi.org/10.1038/sj.emboj.7600047).



346. *Tullius, T. D.* Physical Studies of Protein-DNA Complexes by Footprinting / T. D. Tullius // *Annu Rev Biophys Biophys Chem.* — 1989. — Vol. 18. — P. 213–37. — DOI: [10.1146/annurev.bb.18.060189.001241](https://doi.org/10.1146/annurev.bb.18.060189.001241).
347. *Adilakshmi, T.* Hydroxyl Radical Footprinting in Vivo: Mapping Macromolecular Structures with Synchrotron Radiation / T. Adilakshmi, R. A. Lease, S. A. Woodson // *Nucleic Acids Res.* — 2006. — Май. — Т. 34, № 8. — e64. — DOI: [10.1093/nar/gkl291](https://doi.org/10.1093/nar/gkl291).
348. *Song, L.* DNase-Seq: A High-Resolution Technique for Mapping Active Gene Regulatory Elements across the Genome from Mammalian Cells / L. Song, G. E. Crawford // *Cold Spring Harb Protoc.* — 2010. — 1559–6095 (Electronic). — pdb.prot5384. — DOI: [10.1101/pdb.prot5384](https://doi.org/10.1101/pdb.prot5384).
349. *Jain, S. S.* Footprinting Protein-DNA Complexes Using the Hydroxyl Radical / S. S. Jain, T. D. Tullius // *Nat Protoc.* — 2008. — Vol. 3, no. 6. — P. 1092–1100.
350. *Balasubramanian, B.* DNA Strand Breaking by the Hydroxyl Radical Is Governed by the Accessible Surface Areas of the Hydrogen Atoms of the DNA Backbone / B. Balasubramanian, W. K. Pogozelski, T. D. Tullius // *Proceedings of the National Academy of Sciences.* — 1998. — Aug. 18. — Vol. 95, no. 17. — P. 9738–9743. — DOI: [10.1073/pnas.95.17.9738](https://doi.org/10.1073/pnas.95.17.9738).
351. Quantitative Analysis of Electrophoresis Data: Novel Curve Fitting Methodology and Its Application to the Determination of a Protein-DNA Binding Constant / S. E. Shadle, D. F. Allen, H. Guo, W. K. Pogozelski, J. S. Bashkin, T. D. Tullius // *Nucleic Acids Res.* — 1997. — Feb. 15. — Vol. 25, no. 4. — P. 850–60.
352. SAFA: Semi-Automated Footprinting Analysis Software for High-Throughput Quantification of Nucleic Acid Footprinting Experiments / R. Das, A. Laederach, S. M. Pearlman, D. Herschlag, R. B. Altman // *RNA.* — 2005. — Mar. — Vol. 11, no. 3. — P. 344–54. — DOI: [10.1261/rna.7214405](https://doi.org/10.1261/rna.7214405).
353. A Map of Minor Groove Shape and Electrostatic Potential from Hydroxyl Radical Cleavage Patterns of DNA / E. P. Bishop, R. Rohs, S. C. Parker, S. M. West, P. Liu, R. S. Mann, B. Honig, T. D. Tullius // *ACS Chem Biol.* — 2011. — Dec. 16. — Vol. 6, no. 12. — P. 1314–20. — DOI: [10.1021/cb200155t](https://doi.org/10.1021/cb200155t).

354. *Schneider, C. A.* NIH Image to ImageJ: 25 Years of Image Analysis / C. A. Schneider, W. S. Rasband, K. W. Eliceiri // Nat Methods. — 2012. — July. — Vol. 9, no. 7. — P. 671–5.
355. *Mitternacht, S.* FreeSASA: An Open Source C Library for Solvent Accessible Surface Area Calculations / S. Mitternacht // F1000Research. — 2016. — T. 5. — C. 189.
356. Chemical Probing of RNA with the Hydroxyl Radical at Single-Atom Resolution / S. Ingle, R. N. Azad, S. S. Jain, T. D. Tullius // Nucleic Acids Res. — 2014. — 10 нояб. — Т. 42, № 20. — С. 12758–67. — DOI: [10.1093/nar/gku934](https://doi.org/10.1093/nar/gku934).
357. *Costa, M.* Probing RNA Folding by Hydroxyl Radical Footprinting / M. Costa, D. Monachello // Methods Mol Biol. — 2014. — Т. 1086. — С. 119–42. — DOI: [10.1007/978-1-62703-667-2\\_7](https://doi.org/10.1007/978-1-62703-667-2_7).
358. *Hampel, K. J.* Time-Resolved Hydroxyl-Radical Footprinting of RNA Using Fe(II)-EDTA / K. J. Hampel, J. M. Burke // Methods. — 2001. — Март. — Т. 23, № 3. — С. 233–9. — DOI: [10.1006/meth.2000.1134](https://doi.org/10.1006/meth.2000.1134).
359. *Nilsen, T. W.* Mapping RNA-Protein Interactions Using Hydroxyl-Radical Footprinting / T. W. Nilsen // Cold Spring Harb Protoc. — 2014. — 1 дек. — Т. 2014, № 12. — С. 1333–6. — DOI: [10.1101/pdb.prot080952](https://doi.org/10.1101/pdb.prot080952).
360. Three-Dimensional RNA Structure Refinement by Hydroxyl Radical Probing / F. Ding, C. A. Lavender, K. M. Weeks, N. V. Dokholyan // Nat Methods. — 2012. — 15 апр. — Т. 9, № 6. — С. 603–8. — DOI: [10.1038/nmeth.1976](https://doi.org/10.1038/nmeth.1976).
361. *Tian, S.* RNA Structure through Multidimensional Chemical Mapping / S. Tian, R. Das // Q Rev Biophys. — 2016. — ЯНВ. — Т. 49. — e7. — DOI: [10.1017/S0033583516000020](https://doi.org/10.1017/S0033583516000020).
362. RADACK, a Stochastic Simulation of Hydroxyl Radical Attack to DNA / M. Begusova, M. Spothem-Maurizot, D. Sy, V. Michalik, M. Charlier // J Biomol Struct Dyn. — 2001. — АВГ. — Т. 19, № 1. — С. 141–58. — DOI: [10.1080/07391102.2001.10506727](https://doi.org/10.1080/07391102.2001.10506727).

363. Metadata Matters: Access to Image Data in the Real World / M. Linkert, C. T. Rueden, C. Allan, J. M. Burel, W. Moore, A. Patterson, B. Loranger, J. Moore, C. Neves, D. Macdonald, A. Tarkowska, C. Sticco, E. Hill, M. Rossner, K. W. Eliceiri, J. R. Swedlow // *J Cell Biol.* — 2010. — May 31. — Vol. 189, no. 5. — P. 777–82. — DOI: [10.1083/jcb.201004104](https://doi.org/10.1083/jcb.201004104).
364. *Takamoto, K.* Semi-Automated, Single-Band Peak-Fitting Analysis of Hydroxyl Radical Nucleic Acid Footprint Autoradiograms for the Quantitative Analysis of Transitions / K. Takamoto, M. R. Chance, M. Brenowitz // *Nucleic Acids Research.* — 2004. — T. 32, № 15. — e119. — DOI: [10.1093/nar/gnh117](https://doi.org/10.1093/nar/gnh117).
365. Asparagine and Glutamine: Using Hydrogen Atom Contacts in the Choice of Side-Chain Amide Orientation / J. M. Word, S. C. Lovell, J. S. Richardson, D. C. Richardson // *J Mol Biol.* — 1999. — Jan. 29. — Vol. 285, no. 4. — P. 1735–47. — DOI: [10.1006/jmbi.1998.2401](https://doi.org/10.1006/jmbi.1998.2401).
366. *Lee, B.* The Interpretation of Protein Structures: Estimation of Static Accessibility / B. Lee, F. M. Richards // *J Mol Biol.* — 1971. — 14 февр. — T. 55, № 3. — C. 379–400.
367. *Hunter, J. D.* Matplotlib: A 2D Graphics Environment / J. D. Hunter // *Computing in Science Engineering.* — 2007. — Май. — T. 9, № 3. — C. 90–95. — DOI: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55).
368. *Slater, G. W.* DNA Separation Mechanisms During Electrophoresis | Springer Nature Experiments / G. W. Slater. — 2001. — URL: <https://experiments.springernature.com/articles/10.1385/1-59259-055-1:27> (visited on 10/29/2020).
369. Mobility, Diffusion and Dispersion of Single-Stranded DNA in Sequencing Gels / S. N. Brahmamandra, D. T. Burke, C. H. Mastrangelo, M. A. Burns // *Electrophoresis.* — 2001. — Apr. — Vol. 22, no. 6. — P. 1046–62. — DOI: [10.1002/1522-2683\(200104\)22:6<1046::aid-elps1046>3.0.co;2-e](https://doi.org/10.1002/1522-2683(200104)22:6<1046::aid-elps1046>3.0.co;2-e).
370. *Maramis, C. F.* Improved Modeling of Lane Intensity Profiles on Gel Electrophoresis Images / C. F. Maramis, A. N. Delopoulos // XII Mediterranean Conference on Medical and Biological Engineering and Computing 2010: May 27 – 30, 2010 Chalkidiki, Greece / под ред. P. D. Bamidis, N. Pallikarakis. — Berlin, Heidelberg : Springer Berlin

- Heidelberg, 2010. — C. 671—674. — ISBN 978-3-642-13039-7. — DOI: [10.1007/978-3-642-13039-7\\_169](https://doi.org/10.1007/978-3-642-13039-7_169).
371. *Greenbaum, J. A.* Construction of a Genome-Scale Structural Map at Single-Nucleotide Resolution / J. A. Greenbaum, B. Pang, T. D. Tullius // *Genome Res.* — 2007. — June. — Vol. 17, no. 6. — P. 947—53. — DOI: [10.1101/gr.6073107](https://doi.org/10.1101/gr.6073107).
372. *Berg, J.* APPENDIX: Lorentzian Lineshapes Are Intrinsic to Autoradiographic Detection / J. Berg // *Nucleic Acids Res.* — 1997. — Feb. 15. — Vol. 25, no. 4. — P. 850—60.
373. *Spotheim-Maurizot, M.* Radiation Damage to DNA in DNA-Protein Complexes / M. Spotheim-Maurizot, M. Davidkova // *Mutat Res.* — 2011. — June 3. — Vol. 711, no. 1/2. — P. 41—8. — DOI: [10.1016/j.mrfmmm.2011.02.003](https://doi.org/10.1016/j.mrfmmm.2011.02.003).
374. Using DNA Mechanics to Predict in Vitro Nucleosome Positions and Formation Energies / A. V. Morozov, K. Fortney, D. A. Gaykalova, V. M. Studitsky, J. Widom, E. D. Siggia // *Nucleic Acids Research.* — 2009. — Aug. — Vol. 37, no. 14. — P. 4707—4722. — DOI: [10.1093/nar/gkp475](https://doi.org/10.1093/nar/gkp475).
375. *Cloutier, T. E.* DNA Twisting Flexibility and the Formation of Sharply Looped Protein-DNA Complexes / T. E. Cloutier, J. Widom // *Proc Natl Acad Sci U S A.* — 2005. — Mar. — Vol. 102, no. 10. — P. 3645—50. — DOI: [10.1073/pnas.0409059102](https://doi.org/10.1073/pnas.0409059102).
376. *Lowary, P. T.* New DNA Sequence Rules for High Affinity Binding to Histone Octamer and Sequence-Directed Nucleosome Positioning / P. T. Lowary, J. Widom // *Journal of Molecular Biology.* — 1998. — Feb. 13. — Vol. 276, no. 1. — P. 19—42. — DOI: [10.1006/jmbi.1997.1494](https://doi.org/10.1006/jmbi.1997.1494).
377. *Kappen, L. S.* Deoxyribonucleic Acid Damage by Neocarzinostatin Chromophore: Strand Breaks Generated by Selective Oxidation of C-5' of Deoxyribose / L. S. Kappen, I. H. Goldberg // *Biochemistry.* — 1983. — 11 окт. — Т. 22, № 21. — С. 4872—8.
378. *Gardner, K. E.* Operating on Chromatin, a Colorful Language Where Context Matters / K. E. Gardner, C. D. Allis, B. D. Strahl // *J Mol Biol.* — 2011. — May 27. — Vol. 409, no. 1. — P. 36—46. — DOI: [10.1016/j.jmb.2011.01.040](https://doi.org/10.1016/j.jmb.2011.01.040).

379. Nonhistone Scm3 Binds to AT-Rich DNA to Organize Atypical Centromeric Nucleosome of Budding Yeast / H. Xiao, G. Mizuguchi, J. Wisniewski, Y. Huang, D. Wei, C. Wu // *Mol Cell*. — 2011. — Aug. 5. — Vol. 43, no. 3. — P. 369–80. — DOI: [10.1016/j.molcel.2011.07.009](https://doi.org/10.1016/j.molcel.2011.07.009).
380. *Becker, P. B.* Nucleosome Remodeling and Epigenetics / P. B. Becker, J. L. Workman // *Cold Spring Harb Perspect Biol*. — 2013. — Sept. — Vol. 5, no. 9. — a017905. — DOI: [10.1101/cshperspect.a017905](https://doi.org/10.1101/cshperspect.a017905).
381. *Luger, K.* New Insights into Nucleosome and Chromatin Structure: An Ordered State or a Disordered Affair? / K. Luger, M. L. Dechassa, D. J. Tremethick // *Nature Reviews Molecular Cell Biology*. — 2012. — July. — Vol. 13, no. 7. — P. 436–447. — DOI: [10.1038/nrm3382](https://doi.org/10.1038/nrm3382).
382. *Iwafuchi-Doi, M.* Pioneer Transcription Factors in Cell Reprogramming / M. Iwafuchi-Doi, K. S. Zaret // *Genes Dev*. — 2014. — Dec. 15. — Vol. 28, no. 24. — P. 2679–2692. — DOI: [10.1101/gad.253443.114](https://doi.org/10.1101/gad.253443.114).
383. *Cui, F.* Rotational Positioning of Nucleosomes Facilitates Selective Binding of P53 to Response Elements Associated with Cell Cycle Arrest / F. Cui, V. B. Zhurkin // *Nucleic Acids Res*. — 2014. — ЯНВ. — Т. 42, № 2. — С. 836–47. — DOI: [10.1093/nar/gkt943](https://doi.org/10.1093/nar/gkt943).
384. *Pogozelski, W. K.* Oxidative Strand Scission of Nucleic Acids: Routes Initiated by Hydrogen Abstraction from the Sugar Moiety / W. K. Pogozelski, T. D. Tullius // *Chem Rev*. — 1998. — May 7. — Vol. 98, no. 3. — P. 1089–1108.
385. *Hayes, J. J.* The Structure of DNA in a Nucleosome / J. J. Hayes, T. D. Tullius, A. P. Wolffe // *Proc Natl Acad Sci U S A*. — 1990. — Oct. — Vol. 87, no. 19. — P. 7405–9.
386. The Histone Core Exerts a Dominant Constraint on the Structure of DNA in a Nucleosome / J. J. Hayes, J. Bashkin, T. D. Tullius, A. P. Wolffe // *Biochemistry*. — 1991. — Aug. 27. — Vol. 30, no. 34. — P. 8434–40.
387. *Hayes, J. J.* Histone Contributions to the Structure of DNA in the Nucleosome / J. J. Hayes, D. J. Clark, A. P. Wolffe // *Proc Natl Acad Sci U S A*. — 1991. — Aug. 1. — Vol. 88, no. 15. — P. 6829–33.

388. *Churchill, M. E.* Detection of Drug Binding to DNA by Hydroxyl Radical Footprinting. Relationship of Distamycin Binding Sites to DNA Structure and Positioned Nucleosomes on 5S RNA Genes of *Xenopus* / M. E. Churchill, J. J. Hayes, T. D. Tullius // *Biochemistry*. — 1990. — June 26. — Vol. 29, no. 25. — P. 6043–50.
389. Nucleosome Structural Features and Intrinsic Properties of the TATAAACGCC Repeat Sequence / H. R. Widlund, P. N. Kuduvalli, M. Bengtsson, H. Cao, T. D. Tullius, M. Kubista // *J Biol Chem*. — 1999. — Nov. 5. — Vol. 274, no. 45. — P. 31847–52.
390. *Bjorklund, C. C.* Attenuation of DNA Charge Transport by Compaction into a Nucleosome Core Particle / C. C. Bjorklund, W. B. Davis // *Nucleic Acids Res*. — 2006. — Vol. 34, no. 6. — P. 1836–46. — DOI: [10.1093/nar/gkl030](https://doi.org/10.1093/nar/gkl030).
391. *Schwanbeck, R.* Spatial Contacts and Nucleosome Step Movements Induced by the NURF Chromatin Remodeling Complex / R. Schwanbeck, H. Xiao, C. Wu // *J Biol Chem*. — 2004. — Sept. 17. — Vol. 279, no. 38. — P. 39933–41. — DOI: [10.1074/jbc.M406060200](https://doi.org/10.1074/jbc.M406060200).
392. Single-Base Resolution Mapping of H1-Nucleosome Interactions and 3D Organization of the Nucleosome / S. H. Syed, D. Goutte-Gattat, N. Becker, S. Meyer, M. S. Shukla, J. J. Hayes, R. Everaers, D. Angelov, J. Bednar, S. Dimitrov // *Proceedings of the National Academy of Sciences of the United States of America*. — 2010. — May 25. — Vol. 107, no. 21. — P. 9620–9625. — DOI: [10.1073/pnas.1000309107](https://doi.org/10.1073/pnas.1000309107).
393. *Cole, H. A.* The Centromeric Nucleosome of Budding Yeast Is Perfectly Positioned and Covers the Entire Centromere / H. A. Cole, B. H. Howard, D. J. Clark // *Proceedings of the National Academy of Sciences*. — 2011. — Aug. 2. — Vol. 108, no. 31. — P. 12687–12692. — DOI: [10.1073/pnas.1104978108](https://doi.org/10.1073/pnas.1104978108).
394. A Detailed Interpretation of OH Radical Footprints in a TBP-DNA Complex Reveals the Role of Dynamics in the Mechanism of Sequence-Specific Binding / N. Pastor, H. Weinstein, E. Jamison, M. Brenowitz // *J Mol Biol*. — 2000. — Nov. 17. — Vol. 304, no. 1. — P. 55–68. — DOI: [10.1006/jmbi.2000.4173](https://doi.org/10.1006/jmbi.2000.4173).

395. Radiolysis of Nucleosome Core DNA: A Modelling Approach / M. Begusova, D. Sy, M. Charlier, M. Spothem-Maurizot // *Int J Radiat Biol.* — 2000. — АБГ. — Т. 76, № 8. — С. 1063—73.
396. Nonhistone Scm3 and Histones CenH3-H4 Assemble the Core of Centromere-Specific Nucleosomes / G. Mizuguchi, H. Xiao, J. Wisniewski, M. M. Smith, C. Wu // *Cell.* — 2007. — June 15. — Vol. 129, no. 6. — P. 1153—64. — DOI: [10.1016/j.cell.2007.04.026](https://doi.org/10.1016/j.cell.2007.04.026).
397. Crystal Structure of the Human Centromeric Nucleosome Containing CENP-A / H. Tachiwana, W. Kagawa, T. Shiga, A. Osakabe, Y. Miya, K. Saito, Y. Hayashi-Takanaka, T. Oda, M. Sato, S. Y. Park, H. Kimura, H. Kurumizaka // *Nature.* — 2011. — 11 АБГ. — Т. 476, № 7359. — С. 232—5. — DOI: [10.1038/nature10258](https://doi.org/10.1038/nature10258).
398. *Hubbard, S. J.* NACCESS / S. J. Hubbard, J. M. Thornton. — Department of Biochemistry and Molecular Biology, University College: London, 1993.
399. *Wickham, H.* Ggplot2: Elegant Graphics for Data Analysis / H. Wickham. — New York : Springer-Verlag, 2009. — ISBN 978-0-387-98140-6.
400. *Beitz, E.* TEXshade: Shading and Labeling of Multiple Sequence Alignments Using LATEX2 Epsilon / E. Beitz // *Bioinformatics.* — 2000. — Feb. — Vol. 16, no. 2. — P. 135—9.
401. *Richmond, T. J.* The Structure of DNA in the Nucleosome Core / T. J. Richmond, C. A. Davey // *Nature.* — 2003. — 8 мая. — Т. 423, № 6936. — С. 145—50. — DOI: [10.1038/nature01595](https://doi.org/10.1038/nature01595).
402. *Travers, A.* DNA Structure / A. Travers // *DNA-Protein Interactions* / под ред. A. Travers. — Dordrecht : Springer Netherlands, 1993. — С. 1—27. — ISBN 978-94-011-1480-6. — DOI: [10.1007/978-94-011-1480-6\\_2](https://doi.org/10.1007/978-94-011-1480-6_2).
403. *Vasudevan, D.* Crystal Structures of Nucleosome Core Particles Containing the ‘601’ Strong Positioning Sequence / D. Vasudevan, E. Y. D. Chua, C. A. Davey // *Journal of Molecular Biology.* — 2010. — 15 окт. — Т. 403, № 1. — С. 1—10. — DOI: [10.1016/j.jmb.2010.08.039](https://doi.org/10.1016/j.jmb.2010.08.039).
404. DNA A-Tract Bending in Three Dimensions: Solving the dA4T4 vs. dT4A4 Conundrum / R. Stefl, H. Wu, S. Ravindranathan, V. Sklenar, J. Feigon // *Proc Natl Acad Sci U S A.* — 2004. — Feb. 3. — Vol. 101, no. 5. — P. 1177—82. — DOI: [10.1073/pnas.0308143100](https://doi.org/10.1073/pnas.0308143100).

405. *Haran, T. E.* The Unique Structure of A-Tracts and Intrinsic DNA Bending / T. E. Haran, U. Mohanty // *Q Rev Biophys.* — 2009. — Feb. — Vol. 42, no. 1. — P. 41–81. — DOI: [10.1017/s0033583509004752](https://doi.org/10.1017/s0033583509004752).
406. *Baker, R. E.* Genetic and Genomic Analysis of the AT-Rich Centromere DNA Element II of *Saccharomyces Cerevisiae* / R. E. Baker, K. Rogers // *Genetics.* — 2005. — Dec. — Vol. 171, no. 4. — P. 1463–75. — DOI: [10.1534/genetics.105.046458](https://doi.org/10.1534/genetics.105.046458).
407. *Malik, H. S.* Major Evolutionary Transitions in Centromere Complexity / H. S. Malik, S. Henikoff // *Cell.* — 2009. — Sept. 18. — Vol. 138, no. 6. — P. 1067–82. — DOI: [10.1016/j.cell.2009.08.036](https://doi.org/10.1016/j.cell.2009.08.036).
408. *Ober, M.* A 1,2-d(GpG) Cisplatin Intrastrand Cross-Link Influences the Rotational and Translational Setting of DNA in Nucleosomes / M. Ober, S. J. Lippard // *J Am Chem Soc.* — 2008. — Mar. 5. — Vol. 130, no. 9. — P. 2851–61. — DOI: [10.1021/ja710220x](https://doi.org/10.1021/ja710220x).
409. Rapid Deamination of Cyclobutane Pyrimidine Dimer Photoproducts at TCG Sites in a Translationally and Rotationally Positioned Nucleosome in Vivo / V. J. Cannistraro, S. Pondugula, Q. Song, J. S. Taylor // *J Biol Chem.* — 2015. — Oct. 30. — Vol. 290, no. 44. — P. 26597–609. — DOI: [10.1074/jbc.M115.673301](https://doi.org/10.1074/jbc.M115.673301).
410. Mapping Nucleosome Position at Single Base-Pair Resolution by Using Site-Directed Hydroxyl Radicals / A. Flaus, K. Luger, S. Tan, T. J. Richmond // *Proc Natl Acad Sci U S A.* — 1996. — Feb. 20. — Vol. 93, no. 4. — P. 1370–5.
411. High-Resolution Mapping of Changes in Histone-DNA Contacts of Nucleosomes Remodeled by ISW2 / S. R. Kassabov, N. M. Henry, M. Zofall, T. Tsukiyama, B. Bartholomew // *Mol Cell Biol.* — 2002. — Nov. — Vol. 22, no. 21. — P. 7524–34.
412. *Bao, Y.* Nucleosome Core Particles Containing a Poly(dA.dT) Sequence Element Exhibit a Locally Distorted DNA Structure / Y. Bao, C. L. White, K. Luger // *J Mol Biol.* — 2006. — Aug. 25. — Vol. 361, no. 4. — P. 617–24. — DOI: [10.1016/j.jmb.2006.06.051](https://doi.org/10.1016/j.jmb.2006.06.051).



413. Structural Insight into the Sequence Dependence of Nucleosome Positioning / B. Wu, K. Mohideen, D. Vasudevan, C. A. Davey // *Structure* (London, England: 1993). — 2010. — Mar. 14. — Vol. 18, no. 4. — P. 528–536. — DOI: [10.1016/j.str.2010.01.015](https://doi.org/10.1016/j.str.2010.01.015).
414. Shearing of the CENP-A Dimerization Interface Mediates Plasticity in the Octameric Centromeric Nucleosome / D. Winogradoff, H. Zhao, Y. Dalal, G. A. Papoian // *Sci Rep.* — 2015. — Nov. 25. — Vol. 5. — P. 17038. — DOI: [10.1038/srep17038](https://doi.org/10.1038/srep17038).
415. *Widom, J.* Role of DNA Sequence in Nucleosome Stability and Dynamics / J. Widom // *Q Rev Biophys.* — 2001. — Aug. — Vol. 34, no. 3. — P. 269–324.
416. *Zhou, C.* DNA Damage by Histone Radicals in Nucleosome Core Particles / C. Zhou, M. M. Greenberg // *J Am Chem Soc.* — 2014. — 7 мая. — Т. 136, № 18. — С. 6562–5. — DOI: [10.1021/ja501285s](https://doi.org/10.1021/ja501285s).
417. Structure of DNA in a Nucleosome Core at High Salt Concentration and at High Temperature / J. Bashkin, J. J. Hayes, T. D. Tullius, A. P. Wolffe // *Biochemistry.* — 1993. — Mar. 2. — Vol. 32, no. 8. — P. 1895–8.
418. *Klug, A.* The Helical Periodicity of DNA on the Nucleosome / A. Klug, L. C. Lutter // *Nucleic Acids Res.* — 1981. — Sept. 11. — Vol. 9, no. 17. — P. 4267–83.
419. Towards Quantitative Analysis of Gene Regulation by Enhancers / E. V. Nizovtseva, S. Todolli, W. K. Olson, V. M. Studitsky // *Epigenomics.* — 2017. — Sept. — Vol. 9, no. 9. — P. 1219–1231. — DOI: [10.2217/epi-2017-0061](https://doi.org/10.2217/epi-2017-0061).
420. Macrogenomic Engineering via Modulation of the Scaling of Chromatin Packing Density / L. M. Almassalha, G. M. Bauer, W. Wu, L. Cherkezyan, D. Zhang, A. Kendra, S. Gladstein, J. E. Chandler, D. VanDerway, B.-L. L. Seagle, A. Ugolkov, D. D. Billadeau, T. V. O’Halloran, A. P. Mazar, H. K. Roy, I. Szleifer, S. Shahabi, V. Backman // *Nature Biomedical Engineering.* — 2017. — Nov. — Vol. 1, no. 11. — P. 902–913. — DOI: [10.1038/s41551-017-0153-2](https://doi.org/10.1038/s41551-017-0153-2).

421. *Krivega, I.* Chromatin Looping as a Target for Altering Erythroid Gene Expression / I. Krivega, A. Dean // *Annals of the New York Academy of Sciences*. — 2016. — Март. — Т. 1368, № 1. — С. 31–39. — DOI: [10.1111/nyas.13012](https://doi.org/10.1111/nyas.13012).
422. *Larson, A. G.* The Role of Phase Separation in Heterochromatin Formation, Function, and Regulation / A. G. Larson, G. J. Narlikar // *Biochemistry*. — 2018. — May 1. — Vol. 57, no. 17. — P. 2540–2548. — DOI: [10.1021/acs.biochem.8b00401](https://doi.org/10.1021/acs.biochem.8b00401).
423. *S., G.* Zur Kenntnis Der Heterochromosomen / G. S. // *Arch mikrosk. Anat.* — 1907. P. 69. —
424. The 10-Nm Chromatin Fiber and Its Relationship to Interphase Chromosome Organization / J. C. Hansen, M. Connolly, C. J. McDonald, A. Pan, A. Pryamkova, K. Ray, E. Seidel, S. Tamura, R. Rogge, K. Maeshima // *Biochemical Society Transactions*. — 2018. — Feb. 19. — Vol. 46, no. 1. — P. 67–76. — DOI: [10.1042/BST20170101](https://doi.org/10.1042/BST20170101).
425. *Perišić, O.* Modeling Studies of Chromatin Fiber Structure as a Function of DNA Linker Length / O. Perišić, R. Collepardo-Guevara, T. Schlick // *Journal of molecular biology*. — 2010. — 12 нояб. — Т. 403, № 5. — С. 777–802. — DOI: [10.1016/j.jmb.2010.07.057](https://doi.org/10.1016/j.jmb.2010.07.057).
426. *Tremethick, D. J.* Higher-Order Structures of Chromatin: The Elusive 30 Nm Fiber / D. J. Tremethick // *Cell*. — 2007. — Feb. 23. — Vol. 128, no. 4. — P. 651–654. — DOI: [10.1016/j.cell.2007.02.008](https://doi.org/10.1016/j.cell.2007.02.008).
427. Chromatin Conformation in Living Cells: Support for a Zig-Zag Model of the 30 Nm Chromatin Fiber / B. Rydberg, W. R. Holley, I. S. Mian, A. Chatterjee // *Journal of Molecular Biology*. — 1998. — Nov. 20. — Vol. 284, no. 1. — P. 71–84. — DOI: [10.1006/jmbi.1998.2150](https://doi.org/10.1006/jmbi.1998.2150).
428. *Wu, C.* A Variable Topology for the 30-Nm Chromatin Fibre / C. Wu, A. Bassett, A. Travers // *EMBO reports*. — 2007. — Dec. — Vol. 8, no. 12. — P. 1129–1134. — DOI: [10.1038/sj.embor.7401115](https://doi.org/10.1038/sj.embor.7401115).
429. *Cavalli, G.* Functional Implications of Genome Topology / G. Cavalli, T. Misteli // *Nature Structural & Molecular Biology*. — 2013. — Mar. — Vol. 20, no. 3. — P. 290–299. — DOI: [10.1038/nsmb.2474](https://doi.org/10.1038/nsmb.2474).

430. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping / S. S. P. Rao, M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, E. L. Aiden // *Cell*. — 2014. — Dec. 18. — Vol. 159, no. 7. — P. 1665—1680. — DOI: [10.1016/j.cell.2014.11.021](https://doi.org/10.1016/j.cell.2014.11.021).
431. Chromatin Extrusion Explains Key Features of Loop and Domain Formation in Wild-Type and Engineered Genomes / A. L. Sanborn, S. S. P. Rao, S.-C. Huang, N. C. Durand, M. H. Huntley, A. I. Jewett, I. D. Bochkov, D. Chinnappan, A. Cutkosky, J. Li, K. P. Geeting, A. Gnirke, A. Melnikov, D. McKenna, E. K. Stamenova, E. S. Lander, E. L. Aiden // *Proceedings of the National Academy of Sciences of the United States of America*. — 2015. — Nov. 24. — Vol. 112, no. 47. — E6456—6465. — DOI: [10.1073/pnas.1518552112](https://doi.org/10.1073/pnas.1518552112).
432. Active Chromatin and Transcription Play a Key Role in Chromosome Partitioning into Topologically Associating Domains / S. V. Ulianov, E. E. Khrameeva, A. A. Gavrilov, I. M. Flyamer, P. Kos, E. A. Mikhaleva, A. A. Penin, M. D. Logacheva, M. V. Imakaev, A. Chertovich, M. S. Gelfand, Y. Y. Shevelyov, S. V. Razin // *Genome Research*. — 2016. — Jan. — Vol. 26, no. 1. — P. 70—84. — DOI: [10.1101/gr.196006.115](https://doi.org/10.1101/gr.196006.115).
433. Architectural Protein Subclasses Shape 3-D Organization of Genomes during Lineage Commitment / J. E. Phillips-Cremins, M. E. G. Sauria, A. Sanyal, T. I. Gerasimova, B. R. Lajoie, J. S. K. Bell, C.-T. Ong, T. A. Hookway, C. Guo, Y. Sun, M. J. Bland, W. Wagstaff, S. Dalton, T. C. McDevitt, R. Sen, J. Dekker, J. Taylor, V. G. Corces // *Cell*. — 2013. — 6 июня. — T. 153, № 6. — С. 1281—1295. — DOI: [10.1016/j.cell.2013.04.053](https://doi.org/10.1016/j.cell.2013.04.053).
434. Detecting Hierarchical Genome Folding with Network Modularity / H. K. Norton, D. J. Emerson, H. Huang, J. Kim, K. R. Titus, S. Gu, D. S. Bassett, J. E. Phillips-Cremins // *Nature methods*. — 2018. — Февр. — T. 15, № 2. — С. 119—122. — DOI: [10.1038/nmeth.4560](https://doi.org/10.1038/nmeth.4560).
435. *Sur, I.* The Role of Enhancers in Cancer / I. Sur, J. Taipale // *Nature Reviews Cancer*. — 2016. — Aug. — Vol. 16, no. 8. — P. 483—493. — DOI: [10.1038/nrc.2016.62](https://doi.org/10.1038/nrc.2016.62).

436. Structural Insight into Nucleosome Transcription by RNA Polymerase II with Elongation Factors / H. Ehara, T. Kujirai, Y. Fujino, M. Shirouzu, H. Kurumizaka, S.-i. Sekine // *Science*. — 2019. — Feb. 15. — Vol. 363, no. 6428. — P. 744–747. — DOI: [10.1126/science.aav8912](https://doi.org/10.1126/science.aav8912).
437. Nucleosome–Chd1 Structure and Implications for Chromatin Remodelling / L. Farnung, S. M. Vos, C. Wigge, P. Cramer // *Nature*. — 2017. — Oct. — Vol. 550, no. 7677. — P. 539–542. — DOI: [10.1038/nature24046](https://doi.org/10.1038/nature24046).
438. Cryo-EM Study of the Chromatin Fiber Reveals a Double Helix Twisted by Tetranucleosomal Units / F. Song, P. Chen, D. Sun, M. Wang, L. Dong, D. Liang, R.-M. Xu, P. Zhu, G. Li // *Science*. — 2014. — Apr. 25. — Vol. 344, no. 6182. — P. 376–380. — DOI: [10.1126/science.1251413](https://doi.org/10.1126/science.1251413).
439. *Smolle, M.* reSETting Chromatin during Transcription Elongation / M. Smolle, J. L. Workman, S. Venkatesh // *Epigenetics*. — 2013. — 1 янв. — Т. 8, № 1. — С. 10–15. — DOI: [10.4161/epi.23333](https://doi.org/10.4161/epi.23333).
440. *Das, C.* Histone Exchange and Histone Modifications during Transcription and Aging / C. Das, J. K. Tyler // *Biochimica Et Biophysica Acta*. — 2013. — Vol. 1819, no. 3/4. — P. 332–342. — DOI: [10.1016/j.bbagr.2011.08.001](https://doi.org/10.1016/j.bbagr.2011.08.001).
441. *Zentner, G. E.* Regulation of Nucleosome Dynamics by Histone Modifications / G. E. Zentner, S. Henikoff // *Nature Structural and Molecular Biology*. — 2013. — Т. 20, № 3. — С. 259–266. — DOI: [10.1038/nsmb.2470](https://doi.org/10.1038/nsmb.2470).
442. *Kristjuhan, A.* Evidence for Distinct Mechanisms Facilitating Transcript Elongation through Chromatin in Vivo / A. Kristjuhan, J. Q. Svejstrup // *The EMBO Journal*. — 2004. — 27 окт. — Т. 23, № 21. — С. 4243–4252. — DOI: [10.1038/sj.emboj.7600433](https://doi.org/10.1038/sj.emboj.7600433).
443. *Schwabish, M. A.* Evidence for Eviction and Rapid Deposition of Histones upon Transcriptional Elongation by RNA Polymerase II / M. A. Schwabish, K. Struhl // *Molecular and Cellular Biology*. — 2004. — Dec. — Vol. 24, no. 23. — P. 10111–10117. — DOI: [10.1128/MCB.24.23.10111-10117.2004](https://doi.org/10.1128/MCB.24.23.10111-10117.2004).

444. Evidence for Nucleosome Depletion at Active Regulatory Regions Genome-Wide / C.-K. Lee, Y. Shibata, B. Rao, B. D. Strahl, J. D. Lieb // *Nature Genetics*. — 2004. — Aug. — Vol. 36, no. 8. — P. 900–905. — DOI: [10.1038/ng1400](https://doi.org/10.1038/ng1400).
445. *Weber, C. M.* Nucleosomes Are Context-Specific, H2A.Z-Modulated Barriers to RNA Polymerase / C. M. Weber, S. Ramachandran, S. Henikoff // *Molecular Cell*. — 2014. — Mar. — Vol. 53, no. 5. — P. 819–830. — DOI: [10.1016/j.molcel.2014.02.014](https://doi.org/10.1016/j.molcel.2014.02.014).
446. A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells / M. G. Guenther, S. S. Levine, L. A. Boyer, R. Jaenisch, R. A. Young // *Cell*. — 2007. — July 13. — Vol. 130, no. 1. — P. 77–88. — DOI: [10.1016/j.cell.2007.05.042](https://doi.org/10.1016/j.cell.2007.05.042).
447. RNA Polymerase Is Poised for Activation across the Genome / G. W. Muse, D. A. Gilchrist, S. Nechaev, R. Shah, J. S. Parker, S. F. Grissom, J. Zeitlinger, K. Adelman // *Nature Genetics*. — 2007. — Dec. — Vol. 39, no. 12. — P. 1507–1511. — DOI: [10.1038/ng.2007.21](https://doi.org/10.1038/ng.2007.21).
448. RNA Polymerase Stalling at Developmental Control Genes in the *Drosophila Melanogaster* Embryo / J. Zeitlinger, A. Stark, M. Kellis, J.-W. Hong, S. Nechaev, K. Adelman, M. Levine, R. A. Young // *Nature Genetics*. — 2007. — Dec. — Vol. 39, no. 12. — P. 1512–1516. — DOI: [10.1038/ng.2007.26](https://doi.org/10.1038/ng.2007.26).
449. Nucleosome Organization in the *Drosophila* Genome / T. N. Mavrich, C. Jiang, I. P. Ioshikhes, X. Li, B. J. Venters, S. J. Zanton, L. P. Tomsho, J. Qi, R. L. Glaser, S. C. Schuster, D. S. Gilmour, I. Albert, B. F. Pugh // *Nature*. — 2008. — May 15. — Vol. 453, no. 7193. — P. 358–362. — DOI: [10.1038/nature06929](https://doi.org/10.1038/nature06929).
450. *Brown, S. A.* Activator-Dependent Regulation of Transcriptional Pausing on Nucleosomal Templates. / S. A. Brown, A. N. Imbalzano, R. E. Kingston // *Genes & Development*. — 1996. — June 15. — Vol. 10, no. 12. — P. 1479–1490. — DOI: [10.1101/gad.10.12.1479](https://doi.org/10.1101/gad.10.12.1479).

451. *Izban, M. G.* Transcription on Nucleosomal Templates by RNA Polymerase II in Vitro: Inhibition of Elongation with Enhancement of Sequence-Specific Pausing / M. G. Izban, D. S. Luse // *Genes & Development*. — 1991. — Apr. — Vol. 5, no. 4. — P. 683–696. — DOI: [10.1101/gad.5.4.683](https://doi.org/10.1101/gad.5.4.683).
452. Nature of the Nucleosomal Barrier to RNA Polymerase II / M. L. Kireeva, B. Hancock, G. H. Cremona, W. Walter, V. M. Studitsky, M. Kashlev // *Molecular Cell*. — 2005. — Apr. 1. — Vol. 18, no. 1. — P. 97–108. — DOI: [10.1016/j.molcel.2005.02.027](https://doi.org/10.1016/j.molcel.2005.02.027).
453. Mechanism of Chromatin Remodeling and Recovery during Passage of RNA Polymerase II / O. I. Kulaeva, D. A. Gaykalova, N. A. Pestov, V. V. Golovastov, D. G. Vassylyev, I. Artsimovitch, V. M. Studitsky // *Nature Structural & Molecular Biology*. — 2009. — Dec. — Vol. 16, no. 12. — P. 1272–1278. — DOI: [10.1038/nsmb.1689](https://doi.org/10.1038/nsmb.1689).
454. *Brown, S. A.* Disruption of Downstream Chromatin Directed by a Transcriptional Activator / S. A. Brown, R. E. Kingston // *Genes & Development*. — 1997. — 1 дек. — Т. 11, № 23. — С. 3116–3121.
455. *Izban, M. G.* Factor-Stimulated RNA Polymerase II Transcribes at Physiological Elongation Rates on Naked DNA but Very Poorly on Chromatin Templates / M. G. Izban, D. S. Luse // *The Journal of Biological Chemistry*. — 1992. — July 5. — Vol. 267, no. 19. — P. 13647–13655.
456. Histone Chaperone FACT Action during Transcription through Chromatin by RNA Polymerase II / F.-K. Hsieh, O. I. Kulaeva, S. S. Patel, P. N. Dyer, K. Luger, D. Reinberg, V. M. Studitsky // *Proceedings of the National Academy of Sciences of the United States of America*. — 2013. — May 7. — Vol. 110, no. 19. — P. 7654–7659. — DOI: [10.1073/pnas.1222198110](https://doi.org/10.1073/pnas.1222198110).
457. *Kulaeva, O. I.* RNA Polymerase Complexes Cooperate to Relieve the Nucleosomal Barrier and Evict Histones / O. I. Kulaeva, F.-K. Hsieh, V. M. Studitsky // *Proceedings of the National Academy of Sciences of the United States of America*. — 2010. — June 22. — Vol. 107, no. 25. — P. 11325–11330. — DOI: [10.1073/pnas.1001148107](https://doi.org/10.1073/pnas.1001148107).

458. *Kim, J.* The Human PAF1 Complex Acts in Chromatin Transcription Elongation Both Independently and Cooperatively with SII/TFIIS / J. Kim, M. Guermah, R. G. Roeder // *Cell*. — 2010. — Feb. 19. — Vol. 140, no. 4. — P. 491—503. — DOI: [10.1016/j.cell.2009.12.050](https://doi.org/10.1016/j.cell.2009.12.050).
459. Synergistic Functions of SII and P300 in Productive Activator-Dependent Transcription of Chromatin Templates / M. Guermah, V. B. Palhan, A. J. Tackett, B. T. Chait, R. G. Roeder // *Cell*. — 2006. — Apr. 21. — Vol. 125, no. 2. — P. 275—286. — DOI: [10.1016/j.cell.2006.01.055](https://doi.org/10.1016/j.cell.2006.01.055).
460. *Carey, M.* The RSC Complex Exploits Histone Acetylation to Abrogate the Nucleosomal Barrier to RNA Polymerase II Elongation / M. Carey, B. Li, J. L. Workman // *Molecular cell*. — 2006. — 3 нояб. — Т. 24, № 3. — С. 481—487. — DOI: [10.1016/j.molcel.2006.09.012](https://doi.org/10.1016/j.molcel.2006.09.012).
461. Nucleosomal Elements That Control the Topography of the Barrier to Transcription / L. Bintu, T. Ishibashi, M. Dangkulwanich, Y.-Y. Wu, L. Lubkowska, M. Kashlev, C. Bustamante // *Cell*. — 2012. — Nov. 9. — Vol. 151, no. 4. — P. 738—749. — DOI: [10.1016/j.cell.2012.10.009](https://doi.org/10.1016/j.cell.2012.10.009).
462. *Kettenberger, H.* Complete RNA Polymerase II Elongation Complex Structure and Its Interactions with NTP and TFIIS / H. Kettenberger, K.-J. Armache, P. Cramer // *Molecular Cell*. — 2004. — Dec. 22. — Vol. 16, no. 6. — P. 955—965. — DOI: [10.1016/j.molcel.2004.11.040](https://doi.org/10.1016/j.molcel.2004.11.040).
463. Bacterial Polymerase and Yeast Polymerase II Use Similar Mechanisms for Transcription through Nucleosomes / W. Walter, M. L. Kireeva, V. M. Studitsky, M. Kashlev // *The Journal of Biological Chemistry*. — 2003. — Sept. 19. — Vol. 278, no. 38. — P. 36148—36156. — DOI: [10.1074/jbc.M305647200](https://doi.org/10.1074/jbc.M305647200).
464. Structural Basis for Transcription Elongation by Bacterial RNA Polymerase / D. G. Vassylyev, M. N. Vassylyeva, A. Perederina, T. H. Tahirov, I. Artsimovitch // *Nature*. — 2007. — July 12. — Vol. 448, no. 7150. — P. 157—162. — DOI: [10.1038/nature05932](https://doi.org/10.1038/nature05932).
465. Crystal Structure of *Thermus Aquaticus* Core RNA Polymerase at 3.3 Å Resolution / G. Zhang, E. A. Campbell, L. Minakhin, C. Richter, K. Severinov, S. A. Darst // *Cell*. — 1999. — Sept. 17. — Vol. 98, no. 6. — P. 811—824. — DOI: [10.1016/s0092-8674\(00\)81515-9](https://doi.org/10.1016/s0092-8674(00)81515-9).

466. Histone Sin Mutations Promote Nucleosome Traversal and Histone Displacement by RNA Polymerase II / F.-K. Hsieh, M. Fisher, A. Újvári, V. M. Studitsky, D. S. Luse // *EMBO Reports*. — 2010. — Сент. — Т. 11, № 9. — С. 705—710. — DOI: [10.1038/embor.2010.113](https://doi.org/10.1038/embor.2010.113).
467. *Walz, A.* Sequence of the PR Promoter of Phage Lambda / A. Walz, V. Pirrotta // *Nature*. — 1975. — Mar. 13. — Vol. 254, no. 5496. — P. 118—121. — DOI: [10.1038/254118a0](https://doi.org/10.1038/254118a0).
468. *Xiao, H.* A Highly Conserved Domain of RNA Polymerase II Shares a Functional Element with Acidic Activation Domains of Upstream Transcription Factors. / H. Xiao, J. D. Friesen, J. T. Lis // *Molecular and Cellular Biology*. — 1994. — Ноябрь. — Т. 14, № 11. — С. 7507—7516.
469. *Van Dijk, M.* 3D-DART: A DNA Structure Modelling Server / M. van Dijk, A. M. J. J. Bonvin // *Nucleic Acids Research*. — 2009. — 1 июля. — Т. 37, Web Server issue. — W235—W239. — DOI: [10.1093/nar/gkp287](https://doi.org/10.1093/nar/gkp287).
470. Electrostatics of Nanosystems: Application to Microtubules and the Ribosome / N. A. Baker, D. Sept, S. Joseph, M. J. Holst, J. A. McCammon // *Proceedings of the National Academy of Sciences of the United States of America*. — 2001. — Aug. 28. — Vol. 98, no. 18. — P. 10037—10041. — DOI: [10.1073/pnas.181342398](https://doi.org/10.1073/pnas.181342398).
471. A Polar Barrier to Transcription Can Be Circumvented by Remodeler-Induced Nucleosome Translocation / D. A. Gaykalova, V. Nagarajavel, V. A. Bondarenko, B. Bartholomew, D. J. Clark, V. M. Studitsky // *Nucleic Acids Research*. — 2011. — Май. — Т. 39, № 9. — С. 3520—3528. — DOI: [10.1093/nar/gkq1273](https://doi.org/10.1093/nar/gkq1273).
472. *Fish, R. N.* Promoting Elongation with Transcript Cleavage Stimulatory Factors / R. N. Fish, C. M. Kane // *Biochimica Et Biophysica Acta*. — 2002. — Sept. 13. — Vol. 1577, no. 2. — P. 287—307. — DOI: [10.1016/s0167-4781\(02\)00459-1](https://doi.org/10.1016/s0167-4781(02)00459-1).
473. *Polach, K. J.* Restriction Enzymes as Probes of Nucleosome Stability and Dynamics / K. J. Polach, J. Widom // *Methods in Enzymology*. — 1999. — Vol. 304. — P. 278—298. — DOI: [10.1016/s0076-6879\(99\)04017-3](https://doi.org/10.1016/s0076-6879(99)04017-3).



474. *Zuo, Y.* The Mechanism of E. Coli RNA Polymerase Regulation by ppGpp Is Suggested by the Structure of Their Complex / Y. Zuo, Y. Wang, T. A. Steitz // *Molecular Cell*. — 2013. — May 9. — Vol. 50, no. 3. — P. 430–436. — DOI: [10.1016/j.molcel.2013.03.020](https://doi.org/10.1016/j.molcel.2013.03.020).
475. *Bomble, Y. J.* Multiscale Modeling of Nucleic Acids: Insights into DNA Flexibility / Y. J. Bomble, D. A. Case // *Biopolymers*. — 2008. — Vol. 89, no. 9. — P. 722–731. — DOI: [10.1002/bip.21000](https://doi.org/10.1002/bip.21000).
476. *Feng, H. P.* Lifetime of the Histone Octamer Studied by Continuous-Flow Quasielastic Light Scattering: Test of a Model for Nucleosome Transcription / H. P. Feng, D. S. Scherl, J. Widom // *Biochemistry*. — 1993. — Aug. 3. — Vol. 32, no. 30. — P. 7824–7831. — DOI: [10.1021/bi00081a030](https://doi.org/10.1021/bi00081a030).
477. The Elongation Rate of RNA Polymerase Determines the Fate of Transcribed Nucleosomes / L. Bintu, M. Kopaczynska, C. Hodges, L. Lubkowska, M. Kashlev, C. Bustamante // *Nature structural & molecular biology*. — 2011. — 13 нояб. — Т. 18, № 12. — С. 1394–1399. — DOI: [10.1038/nsmb.2164](https://doi.org/10.1038/nsmb.2164).
478. The Nature of the Nucleosomal Barrier to Transcription: Direct Observation of Paused Intermediates by Electron Cryomicroscopy / J. Bednar, V. M. Studitsky, S. A. Grigoryev, G. Felsenfeld, C. L. Woodcock // *Molecular Cell*. — 1999. — Sept. — Vol. 4, no. 3. — P. 377–386. — DOI: [10.1016/s1097-2765\(00\)80339-1](https://doi.org/10.1016/s1097-2765(00)80339-1).
479. *Randall, S. K.* The Fate of Parental Nucleosomes during SV40 DNA Replication / S. K. Randall, T. J. Kelly // *The Journal of Biological Chemistry*. — 1992. — July 15. — Vol. 267, no. 20. — P. 14259–14265.
480. *Churchman, L. S.* Nascent Transcript Sequencing Visualizes Transcription at Nucleotide Resolution / L. S. Churchman, J. S. Weissman // *Nature*. — 2011. — Jan. 20. — Vol. 469, no. 7330. — P. 368–373. — DOI: [10.1038/nature09652](https://doi.org/10.1038/nature09652).
481. Dynamics of Replication-Independent Histone Turnover in Budding Yeast / M. F. Dion, T. Kaplan, M. Kim, S. Buratowski, N. Friedman, O. J. Rando // *Science (New York, N.Y.)* — 2007. — Mar. 9. — Vol. 315, no. 5817. — P. 1405–1408. — DOI: [10.1126/science.1134053](https://doi.org/10.1126/science.1134053).

482. Genome-Wide Replication-Independent Histone H3 Exchange Occurs Predominantly at Promoters and Implicates H3 K56 Acetylation and Asf1 / A. Rufiange, P.-E. Jacques, W. Bhat, F. Robert, A. Nourani // *Molecular Cell*. — 2007. — Aug. 3. — Vol. 27, no. 3. — P. 393–405. — DOI: [10.1016/j.molcel.2007.07.011](https://doi.org/10.1016/j.molcel.2007.07.011).
483. *Weber, C. M.* Histone Variants: Dynamic Punctuation in Transcription / C. M. Weber, S. Henikoff // *Genes & Development*. — 2014. — 1 ая. — Т. 28, № 7. — С. 672–682. — DOI: [10.1101/gad.238873.114](https://doi.org/10.1101/gad.238873.114).
484. ZMYND11 Links Histone H3.3K36me3 to Transcription Elongation and Tumour Suppression / H. Wen, Y. Li, Y. Xi, S. Jiang, S. Stratton, D. Peng, K. Tanaka, Y. Ren, Z. Xia, J. Wu, B. Li, M. C. Barton, W. Li, H. Li, X. Shi // *Nature*. — 2014. — Apr. 10. — Vol. 508, no. 7495. — P. 263–268. — DOI: [10.1038/nature13045](https://doi.org/10.1038/nature13045).
485. Facilitates Chromatin Transcription Complex Is an Accelerator of Tumor Transformation and Potential Marker and Target of Aggressive Cancers with Low Survival Rates / H. Garcia, J. C. Miecznikowski, A. Safina, M. Commane, A. Ruusulehto, S. Kilpinen, R. W. Leach, K. Attwood, Y. Li, S. Degan, A. R. Omilian, O. Guryanova, O. Papantonopoulou, J. Wang, M. Buck, S. Liu, C. Morrison, K. V. Gurova // *Cell reports*. — 2013. — 11 июля. — Т. 4, № 1. — С. 159–173. — DOI: [10.1016/j.celrep.2013.06.013](https://doi.org/10.1016/j.celrep.2013.06.013).
486. *Fukagawa, T.* Centromere DNA, Proteins and Kinetochore Assembly in Vertebrate Cells / T. Fukagawa // *Chromosome Res.* — 2004. — Vol. 12, no. 6. — P. 557–67. — DOI: [10.1023/B:CHRO.0000036590.96208.83](https://doi.org/10.1023/B:CHRO.0000036590.96208.83).
487. *Biggins, S.* The Composition, Functions, and Regulation of the Budding Yeast Kinetochore / S. Biggins // *Genetics*. — 2013. — Aug. — Vol. 194, no. 4. — P. 817–46. — DOI: [10.1534/genetics.112.145276](https://doi.org/10.1534/genetics.112.145276).
488. *Fukagawa, T.* The Centromere: Chromatin Foundation for the Kinetochore Machinery / T. Fukagawa, W. C. Earnshaw // *Dev Cell*. — 2014. — Sept. 8. — Vol. 30, no. 5. — P. 496–508. — DOI: [10.1016/j.devcel.2014.08.016](https://doi.org/10.1016/j.devcel.2014.08.016).

489. *Carroll, C. W.* Dual Recognition of CENP-A Nucleosomes Is Required for Centromere Assembly / C. W. Carroll, K. J. Milks, A. F. Straight // *J Cell Biol.* — 2010. — June 28. — Vol. 189, no. 7. — P. 1143–55. — DOI: [10.1083/jcb.201001013](https://doi.org/10.1083/jcb.201001013).
490. Induced Ectopic Kinetochore Assembly Bypasses the Requirement for CENP-A Nucleosomes / K. E. Gascoigne, K. Takeuchi, A. Suzuki, T. Hori, T. Fukagawa, I. M. Cheeseman // *Cell.* — 2011. — Apr. 29. — Vol. 145, no. 3. — P. 410–22. — DOI: [10.1016/j.cell.2011.03.031](https://doi.org/10.1016/j.cell.2011.03.031).
491. Autoantibody to Centromere (Kinetochore) in Scleroderma Sera / Y. Moroi, C. Peebles, M. J. Fritzler, J. Steigerwald, E. M. Tan // *Proc Natl Acad Sci U S A.* — 1980. — Mar. — Vol. 77, no. 3. — P. 1627–31.
492. *Earnshaw, W. C.* Identification of a Family of Human Centromere Proteins Using Autoimmune Sera from Patients with Scleroderma / W. C. Earnshaw, N. Rothfield // *Chromosoma.* — 1985. — Vol. 91, no. 3/4. — P. 313–21.
493. CENP-C, an Autoantigen in Scleroderma, Is a Component of the Human Inner Kinetochore Plate / H. Saitoh, J. Tomkiel, C. A. Cooke, H. Ratrie 3rd, M. Maurer, N. F. Rothfield, W. C. Earnshaw // *Cell.* — 1992. — July 10. — Vol. 70, no. 1. — P. 115–25.
494. *Sullivan, B. A.* Determining Centromere Identity: Cyclical Stories and Forking Paths / B. A. Sullivan, M. D. Blower, G. H. Karpen // *Nat Rev Genet.* — 2001. — Aug. — Vol. 2, no. 8. — P. 584–96. — DOI: [10.1038/35084512](https://doi.org/10.1038/35084512).
495. Insights from Biochemical Reconstitution into the Architecture of Human Kinetochores / J. R. Weir, A. C. Faesen, K. Klare, A. Petrovic, F. Basilico, J. Fischbock, S. Pentakota, J. Keller, M. E. Pesenti, D. Pan, D. Vogt, S. Wohlgemuth, F. Herzog, A. Musacchio // *Nature.* — 2016. — 31 авг. — Т. 537, № 7619. — С. 249–253. — DOI: [10.1038/nature19333](https://doi.org/10.1038/nature19333).
496. A Conserved Mechanism for Centromeric Nucleosome Recognition by Centromere Protein CENP-C / H. Kato, J. Jiang, B.-R. Zhou, M. Rozendaal, H. Feng, R. Ghirlando, T. S. Xiao, A. F. Straight, Y. Bai // *Science (New York, N.Y.)* — 2013. — May 31. — Vol. 340, no. 6136. — P. 1110–1113. — DOI: [10.1126/science.1235532](https://doi.org/10.1126/science.1235532).

497. *Brown, M. T.* Sequence Similarities between the Yeast Chromosome Segregation Protein Mif2 and the Mammalian Centromere Protein CENP-C / M. T. Brown // *Gene*. — 1995. — July 4. — Vol. 160, no. 1. — P. 111–6.
498. Identification of Overlapping DNA-Binding and Centromere-Targeting Domains in the Human Kinetochore Protein CENP-C / C. H. Yang, J. Tomkiel, H. Saitoh, D. H. Johnson, W. C. Earnshaw // *Mol Cell Biol*. — 1996. — July. — Vol. 16, no. 7. — P. 3576–86.
499. Characterization of Internal DNA-Binding and C-Terminal Dimerization Domains of Human Centromere/Kinetochore Autoantigen CENP-C in Vitro: Role of DNA-Binding and Self-Associating Activities in Kinetochore Organization / K. Sugimoto, K. Kuriyama, A. Shibata, M. Himeno // *Chromosome Res*. — 1997. — Apr. — Vol. 5, no. 2. — P. 132–41.
500. CENP-C Binds the Alpha-Satellite DNA in Vivo at Specific Centromere Domains / V. Politi, G. Perini, S. Trazzi, A. Pliss, I. Raska, W. C. Earnshaw, G. Della Valle // *J Cell Sci*. — 2002. — June 1. — Vol. 115, Pt 11. — P. 2317–27.
501. *Milks, K. J.* Dissection of CENP-C-Directed Centromere and Kinetochore Assembly / K. J. Milks, B. Moree, A. F. Straight // *Mol Biol Cell*. — 2009. — Oct. — Vol. 20, no. 19. — P. 4246–55. — DOI: [10.1091/mbc.E09-05-0378](https://doi.org/10.1091/mbc.E09-05-0378).
502. The C-Terminal Domain of CENP-C Displays Multiple and Critical Functions for Mammalian Centromere Formation / S. Trazzi, G. Perini, R. Bernardoni, M. Zoli, J. C. Reese, A. Musacchio, G. Della Valle // *PLoS One*. — 2009. — Vol. 4, no. 6. — e5832. — DOI: [10.1371/journal.pone.0005832](https://doi.org/10.1371/journal.pone.0005832).
503. Isolation of Two Genes That Affect Mitotic Chromosome Transmission in *S. Cerevisiae* / D. Meeks-Wagner, J. S. Wood, B. Garvik, L. H. Hartwell // *Cell*. — 1986. — Jan. 17. — Vol. 44, no. 1. — P. 53–63.
504. The Solution Structure of an HMG-I(Y)-DNA Complex Defines a New Architectural Minor Groove Binding Motif / J. R. Huth, C. A. Bewley, M. S. Nissen, J. N. Evans, R. Reeves, A. M. Gronenborn, G. M. Clore // *Nat Struct Biol*. — 1997. — Aug. — Vol. 4, no. 8. — P. 657–65.

505. *Reeves, R.* Structure and Function of the HMGI(Y) Family of Architectural Transcription Factors / R. Reeves // Environ Health Perspect. — 2000. — Oct. — Vol. 108 Suppl 5. — P. 803–9.
506. *Lanini, L.* Domains Required for CENP-C Assembly at the Kinetochores / L. Lanini, F. McKeon // Mol Biol Cell. — 1995. — Aug. — Vol. 6, no. 8. — P. 1049–59.
507. *Meluh, P. B.* Evidence That the MIF2 Gene of *Saccharomyces Cerevisiae* Encodes a Centromere Protein with Homology to the Mammalian Centromere Protein CENP-C / P. B. Meluh, D. Koshland // Mol Biol Cell. — 1995. — July. — Vol. 6, no. 7. — P. 793–807.
508. Structural and Functional Dissection of Mif2p, a Conserved DNA-Binding Kinetochores Protein / R. L. Cohen, C. W. Espelin, P. De Wulf, P. K. Sorger, S. C. Harrison, K. T. Simons // Mol Biol Cell. — 2008. — Oct. — Vol. 19, no. 10. — P. 4480–91. — DOI: [10.1091/mbc.E08-03-0297](https://doi.org/10.1091/mbc.E08-03-0297).
509. A Cooperative Mechanism Drives Budding Yeast Kinetochores Assembly Downstream of CENP-A / P. Hornung, P. Troc, F. Malvezzi, M. Maier, Z. Demianova, T. Zimniak, G. Litos, F. Lampert, A. Schleiffer, M. Brunner, K. Mechtler, F. Herzog, T. C. Marlovits, S. Westermann // J Cell Biol. — 2014. — Aug. 18. — Vol. 206, no. 4. — P. 509–24. — DOI: [10.1083/jcb.201403081](https://doi.org/10.1083/jcb.201403081).
510. *Hegemann, J. H.* The Centromere of Budding Yeast / J. H. Hegemann, U. N. Fleig // Bioessays. — 1993. — Июль. — Т. 15, № 7. — С. 451–60. — DOI: [10.1002/bies.950150704](https://doi.org/10.1002/bies.950150704).
511. *Henikoff, S.* The Centromere Paradox: Stable Inheritance with Rapidly Evolving DNA / S. Henikoff, K. Ahmad, H. S. Malik // Science. — 2001. — 10 авг. — Т. 293, № 5532. — С. 1098–102. — DOI: [10.1126/science.1062939](https://doi.org/10.1126/science.1062939).
512. *Allshire, R. C.* Epigenetic Regulation of Centromeric Chromatin: Old Dogs, New Tricks? / R. C. Allshire, G. H. Karpen // Nat Rev Genet. — 2008. — Dec. — Vol. 9, no. 12. — P. 923–37. — DOI: [10.1038/nrg2466](https://doi.org/10.1038/nrg2466).
513. *Black, B. E.* Epigenetic Centromere Propagation and the Nature of CENP-a Nucleosomes / B. E. Black, D. W. Cleveland // Cell. — 2011. — 18 февр. — Т. 144, № 4. — С. 471–9. — DOI: [10.1016/j.cell.2011.02.002](https://doi.org/10.1016/j.cell.2011.02.002).

514. Imaging the Fate of Histone Cse4 Reveals de Novo Replacement in S Phase and Subsequent Stable Residence at Centromeres / J. Wisniewski, B. Hajj, J. Chen, G. Mizuguchi, H. Xiao, D. Wei, M. Dahan, C. Wu // *Elife*. — 2014. — Vol. 3. — e02203. — DOI: [10.7554/eLife.02203](https://doi.org/10.7554/eLife.02203).
515. *Westermann, S.* Structures and Functions of Yeast Kinetochores / S. Westermann, D. G. Drubin, G. Barnes // *Annu Rev Biochem*. — 2007. — Vol. 76. — P. 563–91. — DOI: [10.1146/annurev.biochem.76.052705.160607](https://doi.org/10.1146/annurev.biochem.76.052705.160607).
516. The N Terminus of the Centromere H3-like Protein Cse4p Performs an Essential Function Distinct from That of the Histone Fold Domain / Y. Chen, R. E. Baker, K. C. Keith, K. Harris, S. Stoler, M. Fitzgerald-Hayes // *Mol Cell Biol*. — 2000. — Sept. — Vol. 20, no. 18. — P. 7037–48.
517. Methylation of CenH3 Arginine 37 Regulates Kinetochores Integrity and Chromosome Segregation / A. Samel, A. Cuomo, T. Bonaldi, A. E. Ehrenhofer-Murray // *Proc Natl Acad Sci U S A*. — 2012. — June 5. — Vol. 109, no. 23. — P. 9029–34. — DOI: [10.1073/pnas.1120968109](https://doi.org/10.1073/pnas.1120968109).
518. In Vitro Centromere and Kinetochores Assembly on Defined Chromatin Templates / A. Guse, C. W. Carroll, B. Moree, C. J. Fuller, A. F. Straight // *Nature*. — 2011. — Sept. 15. — Vol. 477, no. 7364. — P. 354–8. — DOI: [10.1038/nature10379](https://doi.org/10.1038/nature10379).
519. A Bifurcated Hydrogen-Bonded Conformation in the d(A.T) Base Pairs of the DNA Dodecamer d(CGCAAATTTGCG) and Its Complex with Distamycin / M. Coll, C. A. Frederick, A. H. Wang, A. Rich // *Proc Natl Acad Sci U S A*. — 1987. — Dec. — Vol. 84, no. 23. — P. 8385–9.
520. The Structure of an Oligo(dA).Oligo(dT) Tract and Its Biological Implications / H. C. Nelson, J. T. Finch, B. F. Luisi, A. Klug // *Nature*. — 1987. — Nov. 19. — Vol. 330, no. 6145. — P. 221–6. — DOI: [10.1038/330221a0](https://doi.org/10.1038/330221a0).
521. *Reeves, R.* The A.T-DNA-Binding Domain of Mammalian High Mobility Group I Chromosomal Proteins. A Novel Peptide Motif for Recognizing DNA Structure / R. Reeves, M. S. Nissen // *J Biol Chem*. — 1990. — May 25. — Vol. 265, no. 15. — P. 8573–82.

522. Origins of Specificity in Protein-DNA Recognition / R. Rohs, X. Jin, S. M. West, R. Joshi, B. Honig, R. S. Mann // *Annu Rev Biochem.* — 2010. — Vol. 79. — P. 233–69. — DOI: [10.1146/annurev-biochem-060408-091030](https://doi.org/10.1146/annurev-biochem-060408-091030).
523. *Shen, A.* Recognition of AT-Rich DNA Binding Sites by the MogR Repressor / A. Shen, D. E. Higgins, D. Panne // *Structure* (London, England : 1993). — 2009. — 13 мая. — Т. 17, № 5. — С. 769–777. — DOI: [10.1016/j.str.2009.02.018](https://doi.org/10.1016/j.str.2009.02.018).
524. Functional Interplay between the RK Motif and Linker Segment Dictates Oct4–DNA Recognition / X. Kong, J. Liu, L. Li, L. Yue, L. Zhang, H. Jiang, X. Xie, C. Luo // *Nucleic Acids Research.* — 2015. — 19 мая. — Т. 43, № 9. — С. 4381–4392. — DOI: [10.1093/nar/gkv323](https://doi.org/10.1093/nar/gkv323).
525. *Simpson, R. T.* Structure of the Chromatosome, a Chromatin Particle Containing 160 Base Pairs of DNA and All the Histones / R. T. Simpson // *Biochemistry.* — 1978. — 12 дек. — Т. 17, № 25. — С. 5524–5531. — DOI: [10.1021/bi00618a030](https://doi.org/10.1021/bi00618a030).
526. Crystal Structure of Globular Domain of Histone H5 and Its Implications for Nucleosome Binding / V. Ramakrishnan, J. T. Finch, V. Graziano, P. L. Lee, R. M. Sweet // *Nature.* — 1993. — Mar. — Vol. 362, no. 6417. — P. 219. — DOI: [10.1038/362219a0](https://doi.org/10.1038/362219a0).
527. An Asymmetric Model for the Nucleosome: A Binding Site for Linker Histones inside the DNA Gyres / D. Pruss, B. Bartholomew, J. Persinger, J. Hayes, G. Arents, E. N. Moudrianakis, A. P. Wolffe // *Science* (New York, N.Y.) — 1996. — Oct. 25. — Vol. 274, no. 5287. — P. 614–617. — DOI: [10.1126/science.274.5287.614](https://doi.org/10.1126/science.274.5287.614).
528. Position and Orientation of the Globular Domain of Linker Histone H5 on the Nucleosome / Y. B. Zhou, S. E. Gerchman, V. Ramakrishnan, A. Travers, S. Muyldermans // *Nature.* — 1998. — Sept. 24. — Vol. 395, no. 6700. — P. 402–405. — DOI: [10.1038/26521](https://doi.org/10.1038/26521).
529. Structural Insights into the Histone H1-Nucleosome Complex / B.-R. Zhou, H. Feng, H. Kato, L. Dai, Y. Yang, Y. Zhou, Y. Bai // *Proceedings of the National Academy of Sciences.* — 2013. — Nov. 26. — Vol. 110, no. 48. — P. 19390–19395. — DOI: [10.1073/pnas.1314905110](https://doi.org/10.1073/pnas.1314905110).

530. Structural Mechanisms of Nucleosome Recognition by Linker Histones / B.-R. Zhou, J. Jiang, H. Feng, R. Ghirlando, T. S. Xiao, Y. Bai // *Molecular Cell*. — 2015. — 20 авр. — Т. 59, № 4. — С. 628—638. — DOI: [10.1016/j.molcel.2015.06.025](https://doi.org/10.1016/j.molcel.2015.06.025).
531. A Small Number of Residues Can Determine If Linker Histones Are Bound On or Off Dyad in the Chromatosome / B.-R. Zhou, H. Feng, R. Ghirlando, S. Li, C. D. Schwieters, Y. Bai // *Journal of Molecular Biology*. — 2016. — 9 окт. — Т. 428, № 20. — С. 3948—3959. — DOI: [10.1016/j.jmb.2016.08.016](https://doi.org/10.1016/j.jmb.2016.08.016).
532. *Brown, D. T.* Mapping the Interaction Surface of Linker Histone H1(0) with the Nucleosome of Native Chromatin in Vivo / D. T. Brown, T. Izard, T. Misteli // *Nature Structural & Molecular Biology*. — 2006. — Mar. — Vol. 13, no. 3. — P. 250—255. — DOI: [10.1038/nsmb1050](https://doi.org/10.1038/nsmb1050).
533. *Cui, F.* Distinctive Sequence Patterns in Metazoan and Yeast Nucleosomes: Implications for Linker Histone Binding to AT-Rich and Methylated DNA / F. Cui, V. B. Zhurkin // *Nucleic Acids Research*. — 2009. — May. — Vol. 37, no. 9. — P. 2818—2829. — DOI: [10.1093/nar/gkp113](https://doi.org/10.1093/nar/gkp113).
534. Conformational Selection and Dynamic Adaptation upon Linker Histone Binding to the Nucleosome / M. A. Öztürk, G. V. Pachov, R. C. Wade, V. Cojocaru // *Nucleic Acids Research*. — 2016. — Aug. 19. — Vol. 44, no. 14. — P. 6599—6613. — DOI: [10.1093/nar/gkw514](https://doi.org/10.1093/nar/gkw514).
535. *Edgar, R. C.* MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput / R. C. Edgar // *Nucleic Acids Research*. — 2004. — Vol. 32, no. 5. — P. 1792—1797. — DOI: [10.1093/nar/gkh340](https://doi.org/10.1093/nar/gkh340).
536. *Webb, B.* Comparative Protein Structure Modeling Using MODELLER / B. Webb, A. Sali // *Current Protocols in Bioinformatics*. — 2016. — June 20. — Vol. 54. — P. 5.6.1—5.6.37. — DOI: [10.1002/cpbi.3](https://doi.org/10.1002/cpbi.3).
537. Comparative Protein Structure Modeling of Genes and Genomes / M. A. Martí-Renom, A. C. Stuart, A. Fiser, R. Sánchez, F. Melo, A. Sali // *Annual Review of Biophysics and Biomolecular Structure*. — 2000. — Vol. 29. — P. 291—325. — DOI: [10.1146/annurev.biophys.29.1.291](https://doi.org/10.1146/annurev.biophys.29.1.291).
538. *Studitsky, V. M.* Overcoming a Nucleosomal Barrier to Transcription / V. M. Studitsky, D. J. Clark, G. Felsenfeld // *Cell*. — 1995. — Oct. 6. — Vol. 83, no. 1. — P. 19—27. — DOI: [10.1016/0092-8674\(95\)90230-9](https://doi.org/10.1016/0092-8674(95)90230-9).



539. X-Ray Structure of a Tetranucleosome and Its Implications for the Chromatin Fibre / T. Schalch, S. Duda, D. F. Sargent, T. J. Richmond // *Nature*. — 2005. — July 7. — Vol. 436, no. 7047. — P. 138–141. — DOI: [10.1038/nature03686](https://doi.org/10.1038/nature03686).
540. Common Core Structure of Amyloid Fibrils by Synchrotron X-Ray Diffraction / M. Sunde, L. C. Serpell, M. Bartlam, P. E. Fraser, M. B. Pepys, C. C. Blake // *Journal of Molecular Biology*. — 1997. — Oct. 31. — Vol. 273, no. 3. — P. 729–739. — DOI: [10.1006/jmbi.1997.1348](https://doi.org/10.1006/jmbi.1997.1348).
541. *Kumar, S.* Mechanisms of Amyloid Fibril Formation by Proteins / S. Kumar, J. B. Udgaonkar // *CURRENT SCIENCE*. — 2010. — Vol. 98, no. 5. — P. 18.
542. *Chiti, F.* Protein Misfolding, Functional Amyloid, and Human Disease / F. Chiti, C. M. Dobson // *Annual Review of Biochemistry*. — 2006. — Vol. 75. — P. 333–366. — DOI: [10.1146/annurev.biochem.75.101304.123901](https://doi.org/10.1146/annurev.biochem.75.101304.123901).
543. *Makin, O. S.* Structures for Amyloid Fibrils / O. S. Makin, L. C. Serpell // *The FEBS journal*. — 2005. — Dec. — Vol. 272, no. 23. — P. 5950–5961. — DOI: [10.1111/j.1742-4658.2005.05025.x](https://doi.org/10.1111/j.1742-4658.2005.05025.x).
544. *Rochet, J. C.* Amyloid Fibrillogenesis: Themes and Variations / J. C. Rochet, P. T. Lansbury // *Current Opinion in Structural Biology*. — 2000. — Feb. — Vol. 10, no. 1. — P. 60–68. — DOI: [10.1016/s0959-440x\(99\)00049-4](https://doi.org/10.1016/s0959-440x(99)00049-4).
545. *Sipe, J.* The Beta-Pleated Sheet Conformation and Protein Folding: A Brief History / J. Sipe // *Amyloid Proteins*. — 2008. — C. 48–61.
546. Designing Conditions for in Vitro Formation of Amyloid Protofilaments and Fibrils / F. Chiti, P. Webster, N. Taddei, A. Clark, M. Stefani, G. Ramponi, C. M. Dobson // *Proceedings of the National Academy of Sciences of the United States of America*. — 1999. — 30 марта. — T. 96, № 7. — C. 3590–3594.
547. Water-Soluble Beta-Sheet Models Which Self-Assemble into Fibrillar Structures / K. Janek, J. Behlke, J. Zipper, H. Fabian, Y. Georgalis, M. Beyermann, M. Bienert, E. Krause // *Biochemistry*. — 1999. — T. 38, № 26. — C. 8246–8252.

548. *Fändrich, M.* The Behaviour of Polyamino Acids Reveals an Inverse Side Chain Effect in Amyloid Structure Formation / M. Fändrich, C. Dobson // The EMBO Journal. — 2002. — T. 21, № 21. — C. 5682—5690.
549. *Rambaran, R. N.* Amyloid Fibrils: Abnormal Protein Assembly / R. N. Rambaran, L. C. Serpell // Prion. — 2008. — Vol. 2, no. 3. — P. 112—117. — DOI: [10.4161/pri.2.3.7488](https://doi.org/10.4161/pri.2.3.7488).
550. *Fändrich, M.* Structural Polymorphism of Alzheimer Abeta and Other Amyloid Fibrils / M. Fändrich, J. Meinhardt, N. Grigorieff // Prion. — 2009. — T. 3, № 2. — C. 89—93.
551. Structure of the Cross-[Beta] Spine of Amyloid-like Fibrils / R. Nelson, M. Sawaya, M. Balbirnie, A. Madsen, C. Riek, R. Grothe, D. Eisenberg // Nature. — 2005. — T. 435, № 7043. — C. 773—778.
552. Atomic Structures of Amyloid Cross-Beta Spines Reveal Varied Steric Zippers / M. R. Sawaya, S. Sambashivan, R. Nelson, M. I. Ivanova, S. A. Sievers, M. I. Apostol, M. J. Thompson, M. Balbirnie, J. J. W. Wiltzius, H. T. McFarlane, A. Ø. Madsen, C. Riek, D. Eisenberg // Nature. — 2007. — May 24. — Vol. 447, issue 7143, no. 7143. — P. 453—457. — DOI: [10.1038/nature05695](https://doi.org/10.1038/nature05695).
553. *Petkova, A. T.* Experimental Constraints on Quaternary Structure in Alzheimer's Beta-Amyloid Fibrils / A. T. Petkova, W.-M. Yau, R. Tycko // Biochemistry. — 2006. — Jan. 17. — Vol. 45, no. 2. — P. 498—512. — DOI: [10.1021/bi051952q](https://doi.org/10.1021/bi051952q).
554. *Tycko, R.* Molecular Structure of Amyloid Fibrils: Insights from Solid-State NMR / R. Tycko // Quarterly Reviews of Biophysics. — 2006. — Feb. — Vol. 39, no. 1. — P. 1—55. — DOI: [10.1017/S0033583506004173](https://doi.org/10.1017/S0033583506004173).
555. Increasing the Amphiphilicity of an Amyloidogenic Peptide Changes the Beta-Sheet Structure in the Fibrils from Antiparallel to Parallel / D. J. Gordon, J. J. Balbach, R. Tycko, S. C. Meredith // Biophysical Journal. — 2004. — ЯНВ. — T. 86, № 1. — C. 428—434.
556. Supramolecular Structure in Full-Length Alzheimer's Beta-Amyloid Fibrils: Evidence for a Parallel Beta-Sheet Organization from Solid-State Nuclear Magnetic Resonance / J. Balbach, A. Petkova, N. Oyler, O. Antzutkin, D. Gordon, S. Meredith, R. Tycko // Biophys. J. — 2002. — T. 83, № 2. — C. 1205—1216.

557. Solid State NMR Reveals a pH-Dependent Antiparallel Beta-Sheet Registry in Fibrils Formed by a Beta-Amyloid Peptide / A. T. Petkova, G. Buntkowsky, F. Dyda, R. D. Leapman, W.-M. Yau, R. Tycko // *Journal of Molecular Biology*. — 2004. — Jan. 2. — Vol. 335, no. 1. — P. 247–260. — DOI: [10.1016/j.jmb.2003.10.044](https://doi.org/10.1016/j.jmb.2003.10.044).
558. Self-Propagating, Molecular-Level Polymorphism in Alzheimer's Beta-Amyloid Fibrils / A. T. Petkova, R. D. Leapman, Z. Guo, W.-M. Yau, M. P. Mattson, R. Tycko // *Science (New York, N.Y.)* — 2005. — Jan. 14. — Vol. 307, no. 5707. — P. 262–265. — DOI: [10.1126/science.1105850](https://doi.org/10.1126/science.1105850).
559. 3D Structure of Alzheimer's Amyloid-Beta(1-42) Fibrils / T. Lührs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Döbeli, D. Schubert, R. Riek // *Proceedings of the National Academy of Sciences*. — 2005. — T. 102, № 48. — C. 17342–17347.
560. The Changing Face of Glucagon Fibrillation: Structural Polymorphism and Conformational Imprinting / J. S. Pedersen, D. Dikov, J. L. Flink, H. A. Hjuler, G. Christiansen, D. E. Otzen // *Journal of Molecular Biology*. — 2006. — Jan. 20. — Vol. 355, no. 3. — P. 501–523. — DOI: [10.1016/j.jmb.2005.09.100](https://doi.org/10.1016/j.jmb.2005.09.100).
561. The Structural Basis of Yeast Prion Strain Variants / B. H. Toyama, M. J. S. Kelly, J. D. Gross, J. S. Weissman // *Nature*. — 2007. — Sept. 13. — Vol. 449, no. 7159. — P. 233–237. — DOI: [10.1038/nature06108](https://doi.org/10.1038/nature06108).
562. *Makarava, N.* The Same Primary Structure of the Prion Protein Yields Two Distinct Self-Propagating States / N. Makarava, I. V. Baskakov // *The Journal of Biological Chemistry*. — 2008. — 6 июня. — Т. 283, № 23. — C. 15988–15996. — DOI: [10.1074/jbc.M800562200](https://doi.org/10.1074/jbc.M800562200).
563. Polymorphism in an Amyloid-Like Fibril-Forming Model Peptide / R. Verel, I. T. Tomka, C. Bertozzi, R. Cadalbert, R. A. Kammerer, M. O. Steinmetz, B. H. Meier // *Angewandte Chemie International Edition*. — 2008. — Т. 47, № 31. — C. 5842–5845. — DOI: [10.1002/anie.200800021](https://doi.org/10.1002/anie.200800021).
564. Effect of Different Salt Ions on the Propensity of Aggregation and on the Structure of Alzheimer's Abeta(1-40) Amyloid Fibrils / K. Klement, K. Wieligmann, J. Meinhardt, P. Hortschansky, W. Richter, M. Fändrich // *Journal of Molecular Biology*. — 2007. — Т. 373, № 5. — C. 1321–1333.

565. *Uversky, V. N.* Conformational Constraints for Amyloid Fibrillation: The Importance of Being Unfolded / V. N. Uversky, A. L. Fink // *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*. — 2004. — May. — Vol. 1698, no. 2. — P. 131–153. — DOI: [10.1016/j.bbapap.2003.12.008](https://doi.org/10.1016/j.bbapap.2003.12.008).
566. *Kelly, J. W.* The Alternative Conformations of Amyloidogenic Proteins and Their Multi-Step Assembly Pathways / J. W. Kelly // *Current Opinion in Structural Biology*. — 1998. — Feb. 1. — Vol. 8, no. 1. — P. 101–106. — DOI: [10.1016/S0959-440X\(98\)80016-X](https://doi.org/10.1016/S0959-440X(98)80016-X).
567. *Come, J. H.* A Kinetic Model for Amyloid Formation in the Prion Diseases: Importance of Seeding / J. H. Come, P. E. Fraser, P. T. Lansbury // *Proceedings of the National Academy of Sciences of the United States of America*. — 1993. — July 1. — Vol. 90, no. 13. — P. 5959–5963. — DOI: [10.1073/pnas.90.13.5959](https://doi.org/10.1073/pnas.90.13.5959).
568. On the Nucleation and Growth of Amyloid Beta-Protein Fibrils: Detection of Nuclei and Quantitation of Rate Constants / A. Lomakin, D. S. Chung, G. B. Benedek, D. A. Kirschner, D. B. Teplow // *Proceedings of the National Academy of Sciences of the United States of America*. — 1996. — Feb. 6. — Vol. 93, no. 3. — P. 1125–1129. — DOI: [10.1073/pnas.93.3.1125](https://doi.org/10.1073/pnas.93.3.1125).
569. Amyloid Beta-Protein Fibrillogenesis. Structure and Biological Activity of Protofibrillar Intermediates / D. M. Walsh, D. M. Hartley, Y. Kusumoto, Y. Fezoui, M. M. Condron, A. Lomakin, G. B. Benedek, D. J. Selkoe, D. B. Teplow // *The Journal of Biological Chemistry*. — 1999. — Sept. 3. — Vol. 274, no. 36. — P. 25945–25952. — DOI: [10.1074/jbc.274.36.25945](https://doi.org/10.1074/jbc.274.36.25945).
570. *Xue, W.-F.* Systematic Analysis of Nucleation-Dependent Polymerization Reveals New Insights into the Mechanism of Amyloid Self-Assembly / W.-F. Xue, S. W. Homans, S. E. Radford // *Proceedings of the National Academy of Sciences of the United States of America*. — 2008. — July 1. — Vol. 105, no. 26. — P. 8926–8931. — DOI: [10.1073/pnas.0711664105](https://doi.org/10.1073/pnas.0711664105).
571. *S, K.* Conformational Conversion May Precede or Follow Aggregate Elongation on Alternative Pathways of Amyloid Protofibril Formation. / K. S, U. Jb // *Journal of Molecular Biology*. — 2008. — Nov. 25. — Vol. 385, no. 4. — P. 1266–1276. — DOI: [10.1016/j.jmb.2008.11.033](https://doi.org/10.1016/j.jmb.2008.11.033).

572. Characterization of the Formation of Amyloid Protofibrils from Barstar by Mapping Residue-Specific Fluorescence Dynamics / S. Mukhopadhyay, P. K. Nayak, J. B. Udgaonkar, G. Krishnamoorthy // *Journal of Molecular Biology*. — 2006. — May 12. — Vol. 358, no. 4. — P. 935–942. — DOI: [10.1016/j.jmb.2006.02.006](https://doi.org/10.1016/j.jmb.2006.02.006).
573. Secondary Structure of Alpha-Synuclein Oligomers: Characterization by Raman and Atomic Force Microscopy / M. M. Apetri, N. C. Maiti, M. G. Zagorski, P. R. Carey, V. E. Anderson // *Journal of Molecular Biology*. — 2006. — Jan. 6. — Vol. 355, no. 1. — P. 63–71. — DOI: [10.1016/j.jmb.2005.10.071](https://doi.org/10.1016/j.jmb.2005.10.071).
574. Evidence of Fibril-like Beta-Sheet Structures in a Neurotoxic Amyloid Intermediate of Alzheimer's Beta-Amyloid / S. Chimon, M. Shaibat, C. Jones, D. Calero, B. Aizezi, Y. Ishii // *Nature structural & molecular biology*. — 2007. — T. 14, № 12. — C. 1157–1164.
575. Multiple Assembly Pathways Underlie Amyloid-Beta Fibril Polymorphisms / C. Goldsbury, P. Frey, V. Olivieri, U. Aebi, S. A. Müller // *Journal of Molecular Biology*. — 2005. — Sept. 16. — Vol. 352, no. 2. — P. 282–298. — DOI: [10.1016/j.jmb.2005.07.029](https://doi.org/10.1016/j.jmb.2005.07.029).
576. *Jain, S.* Evidence for Stepwise Formation of Amyloid Fibrils by the Mouse Prion Protein / S. Jain, J. B. Udgaonkar // *Journal of Molecular Biology*. — 2008. — Oct. 24. — Vol. 382, no. 5. — P. 1228–1241. — DOI: [10.1016/j.jmb.2008.07.052](https://doi.org/10.1016/j.jmb.2008.07.052).
577. *Kumar, S.* Mechanism of Formation of Amyloid Protofibrils of Barstar from Soluble Oligomers: Evidence for Multiple Steps and Lateral Association Coupled to Conformational Conversion / S. Kumar, S. K. Mohanty, J. B. Udgaonkar // *Journal of Molecular Biology*. — 2007. — Apr. 6. — Vol. 367, no. 4. — P. 1186–1204. — DOI: [10.1016/j.jmb.2007.01.039](https://doi.org/10.1016/j.jmb.2007.01.039).
578. Nucleated Conformational Conversion and the Replication of Conformational Information by a Prion Determinant / T. R. Serio, A. G. Cashikar, A. S. Kowal, G. J. Sawicki, J. J. Moslehi, L. Serpell, M. F. Arnsdorf, S. L. Lindquist // *Science (New York, N.Y.)* — 2000. — Aug. 25. — Vol. 289, no. 5483. — P. 1317–1321. — DOI: [10.1126/science.289.5483.1317](https://doi.org/10.1126/science.289.5483.1317).

579. Assembly of Amyloid Protofibrils via Critical Oligomers—a Novel Pathway of Amyloid Formation / A. J. Modler, K. Gast, G. Lutsch, G. Damaschun // *Journal of Molecular Biology*. — 2003. — Jan. 3. — Vol. 325, no. 1. — P. 135–148. — DOI: [10.1016/s0022-2836\(02\)01175-0](https://doi.org/10.1016/s0022-2836(02)01175-0).
580. Transthyretin Aggregation under Partially Denaturing Conditions Is a Downhill Polymerization / A. R. Hurshman, J. T. White, E. T. Powers, J. W. Kelly // *Biochemistry*. — 2004. — June 15. — Vol. 43, no. 23. — P. 7365–7381. — DOI: [10.1021/bi049621l](https://doi.org/10.1021/bi049621l).
581. Protofibril Formation of Amyloid Beta-Protein at Low pH via a Non-Cooperative Elongation Mechanism / R. Carrotta, M. Manno, D. Bulone, V. Martorana, P. L. San Biagio // *The Journal of Biological Chemistry*. — 2005. — Aug. 26. — Vol. 280, no. 34. — P. 30001–30008. — DOI: [10.1074/jbc.M500052200](https://doi.org/10.1074/jbc.M500052200).
582. Hierarchical Assembly of Beta2-Microglobulin Amyloid In Vitro Revealed by Atomic Force Microscopy / N. Kad, S. Myers, D. Smith, D. Alastair Smith, S. Radford, N. Thomson // *Journal of Molecular Biology*. — 2003. — T. 330, № 4. — C. 785–797.
583. Alpha-Synuclein, Especially the Parkinson's Disease-Associated Mutants, Forms Pore-like Annular and Tubular Protofibrils / H. A. Lashuel, B. M. Petre, J. Wall, M. Simon, R. J. Nowak, T. Walz, P. T. Lansbury // *Journal of Molecular Biology*. — 2002. — Oct. 4. — Vol. 322, no. 5. — P. 1089–1102. — DOI: [10.1016/s0022-2836\(02\)00735-0](https://doi.org/10.1016/s0022-2836(02)00735-0).
584. Neurodegenerative Disease: Amyloid Pores from Pathogenic Mutations / H. A. Lashuel, D. Hartley, B. M. Petre, T. Walz, P. T. Lansbury // *Nature*. — 2002. — July 18. — Vol. 418, no. 6895. — P. 291. — DOI: [10.1038/418291a](https://doi.org/10.1038/418291a).
585. Annular Protofibrils Are a Structurally and Functionally Distinct Type of Amyloid Oligomer / R. Kaye, A. Pensalfini, L. Margol, Y. Sokolov, F. Sarsoza, E. Head, J. Hall, C. Glabe // *The Journal of Biological Chemistry*. — 2009. — Feb. 13. — Vol. 284, no. 7. — P. 4230–4237. — DOI: [10.1074/jbc.M808591200](https://doi.org/10.1074/jbc.M808591200).
586. Alzheimer's Disease Amyloid Propagation by a Template-Dependent Dock-Lock Mechanism / W. P. Esler, E. R. Stimson, J. M. Jennings, H. V. Vinters,

- J. R. Ghilardi, J. P. Lee, P. W. Mantyh, J. E. Maggio // *Biochemistry*. — 2000. — 1 мая. — Т. 39, № 21. — С. 6288—6295. — DOI: [10.1021/bi992933h](https://doi.org/10.1021/bi992933h).
587. *Scheibel, T.* The Elongation of Yeast Prion Fibers Involves Separable Steps of Association and Conversion / T. Scheibel, J. Bloom, S. L. Lindquist // *Proceedings of the National Academy of Sciences of the United States of America*. — 2004. — Feb. 24. — Vol. 101, no. 8. — P. 2287—2292. — DOI: [10.1073/pnas.0308754101](https://doi.org/10.1073/pnas.0308754101).
588. Hierarchical Self-Assembly of Chiral Rod-like Molecules as a Model for Peptide  $\alpha$ -Sheet Tapes, Ribbons, Fibrils, and Fibers / A. Aggeli, I. A. Nyrkova, M. Bell, R. Harding, L. Carrick, T. C. B. McLeish, A. N. Semenov, N. Boden // *Proceedings of the National Academy of Sciences*. — 2001. — Т. 98, № 21. — С. 11857—11862.
589. Functional Amyloid Formation within Mammalian Tissue / D. Fowler, A. Koulov, C. Alory-Jost, M. Marks, W. Balch, J. Kelly // *PLoS biology*. — 2006. — Т. 4, № 1. — e6.
590. Structural Analysis of Hydrophobins / M. Sunde, A. H. Y. Kwan, M. D. Templeton, R. E. Beever, J. P. Mackay // *Micron* (Oxford, England: 1993). — 2008. — Oct. — Vol. 39, no. 7. — P. 773—784. — DOI: [10.1016/j.micron.2007.08.003](https://doi.org/10.1016/j.micron.2007.08.003).
591. *Uptain, S. M.* Prions as Protein-Based Genetic Elements / S. M. Uptain, S. Lindquist // *Annual Review of Microbiology*. — 2002. — Vol. 56. — P. 703—741. — DOI: [10.1146/annurev.micro.56.013002.100603](https://doi.org/10.1146/annurev.micro.56.013002.100603).
592. Conducting Nanowires Built by Controlled Self-Assembly of Amyloid Fibers and Selective Metal Deposition / T. Scheibel, R. Parthasarathy, G. Sawicki, X.-M. Lin, H. Jaeger, S. L. Lindquist // *Proceedings of the National Academy of Sciences*. — 2003. — Apr. 15. — Vol. 100, no. 8. — P. 4527—4532. — DOI: [10.1073/pnas.0431081100](https://doi.org/10.1073/pnas.0431081100).
593. *Hamada, D.* Engineering Amyloidogenicity towards the Development of Nanofibrillar Materials. - Abstract - Europe PMC / D. Hamada, I. Yanagihara, K. Tsumoto. — 2004. — URL: <https://europepmc.org/article/med/14757044> (дата обр. 24.10.2020).

594. Electrochemical Devices Made from Conducting Nanowire Networks Self-Assembled from Amyloid Fibrils and Alkoxysulfonate PEDOT / M. Hamedi, A. Herland, R. H. Karlsson, O. Inganäs // Nano Letters. — 2008. — June. — Vol. 8, no. 6. — P. 1736–1740. — DOI: [10.1021/nl0808233](https://doi.org/10.1021/nl0808233).
595. Characterization of the Nanoscale Properties of Individual Amyloid Fibrils / J. Smith, T. Knowles, C. Dobson, C. MacPhee, M. Welland // Proceedings of the National Academy of Sciences. — 2006. — T. 103, № 43. — C. 15806–15811.
596. Cytochrome Display on Amyloid Fibrils / A. J. Baldwin, R. Bader, J. Christodoulou, C. E. MacPhee, C. M. Dobson, P. D. Barker // Journal of the American Chemical Society. — 2006. — Feb. 22. — Vol. 128, no. 7. — P. 2162–2163. — DOI: [10.1021/ja0565673](https://doi.org/10.1021/ja0565673).
597. *Reches, M.* Casting Metal Nanowires within Discrete Self-Assembled Peptide Nanotubes / M. Reches, E. Gazit // Science (New York, N.Y.) — 2003. — Apr. 25. — Vol. 300, no. 5619. — P. 625–627. — DOI: [10.1126/science.1082387](https://doi.org/10.1126/science.1082387).
598. *Börner, H.* Strategies Exploiting Functions and Self-Assembly Properties of Bioconjugates for Polymer and Materials Sciences / H. Börner // Progress in Polymer Science. — 2009. — T. 34, № 9. — C. 811–851.
599. *Lutz, J.* Modern Trends in Polymer Bioconjugates Design / J. Lutz, H. Börner // Progress in Polymer Science. — 2008. — T. 33, № 1. — C. 1–39.
600. *Börner, H.* Bioinspired Functional Block Copolymers / H. Börner, H. Schlaad // Soft Matter. — 2007. — T. 3, № 4. — C. 394–408.
601. *Vandermeulen, G.* Peptide/Protein Hybrid Materials: Enhanced Control of Structure and Improved Performance through Conjugation of Biological and Synthetic Polymers / G. Vandermeulen, H.-A. Klok // Macromolecular Bioscience. — 2004. — T. 4, № 4. — C. 383–398.
602. *Hardy, J. D.* Silk-Inspired Polymers and Proteins / J. D. Hardy, T. R. Scheibel // Biochemical Society Transactions. — 2009. — T. 37. — C. 677–681.



603. *Schlaad, H.* Block Copolymers with Amino Acid Sequences: Molecular Chimeras of Polypeptides and Synthetic Polymers / H. Schlaad, M. Antonietti // *The European Physical Journal E: Soft Matter and Biological Physics*. — 2003. — T. 10, № 1. — C. 17–23.
604. Rational Design of Oligopeptide Organizers for the Formation of Poly(Ethylene Oxide) Nanofibers / D. Eckhardt, M. Groenewolt, E. Krause, H. Börner // *Chemical Communications*. — 2005. — № 22. — C. 2814–2816.
605. *Ikkala, O.* Hierarchical Self-Assembly in Polymeric Complexes: Towards Functional Materials / O. Ikkala, G. Brinke // *Chemical Communications*. — 2004. — № 19. — C. 2131–2137.
606. *Hentschel, J.* Switch-Peptides to Trigger the Peptide Guided Assembly of Poly(Ethylene Oxide)-Peptide Conjugates into Tape Structures / J. Hentschel, E. Krause, H. Börner // *Journal of the American Chemical Society*. — 2006. — T. 128, № 24. — C. 7722–7723.
607. *Hentschel, J.* Peptide-Directed Microstructure Formation of Polymers in Organic Media / J. Hentschel, H. Börner // *Journal of the American Chemical Society*. — 2006. — T. 128, № 43. — C. 14142–14149.
608. *Frauenrath, H.* A General Concept for the Preparation of Hierarchically Structured Pi-Conjugated Polymers / H. Frauenrath, E. Jahnke // *Chemistry - A European Journal*. — 2008. — T. 14, № 10. — C. 2942–2955.
609. *Diegelmann, S.* One-Dimensional Optoelectronic Nanostructures Derived from the Aqueous Self-Assembly of Pi-Conjugated Oligopeptides / S. Diegelmann, J. Gorham, J. Tovar // *Journal of the American Chemical Society*. — 2008. — T. 130, № 42. — C. 13840–13841.
610. Oligothiophene Versus Beta-Sheet Peptide: Synthesis and Self-Assembly of an Organic Semiconductor-Peptide Hybrid / E.-K. Schillinger, E. Mena-Osteritz, J. Hentschel, H. Börner, P. Bäuerle // *Advanced Materials*. — 2009. — T. 21, № 16. — C. 1562–1567.
611. Synthesis of a Silk-Inspired Peptide-Oligothiophene Conjugate / H.-A. Klok, A. Rosler, G. Gotz, E. Mena-Osteritz, P. Bäuerle // *Organic & Biomolecular Chemistry*. — 2004. — T. 2, № 24. — C. 3541–3544.

612. Self-Assembly of Amphiphiles with Terthiophene and Tripeptide Segments into Helical Nanostructures / W. Tsai, L. Li, H. Cui, H. Jiang, S. Stupp // *Tetrahedron*. — 2008. — T. 64, № 36. — C. 8504—8514.
613. *Stone, D.* Self-Assembling Quinquethiophene-Oligopeptide Hydrogelators / D. Stone, L. Hsu, S. Stupp // *Soft Matter*. — 2009. — T. 5, № 10. — C. 1990—1993.
614. Coincidence of the Molecular Organization of B-Substituted Oligothiophenes in Two-Dimensional Layers and Three-Dimensional Crystals / R. Azumi, G. Götz, T. Debaerdemaeker, P. Bäuerle // *Chem. Eur. J.* — 2000. — T. 6, № 4. — C. 735—744.
615. Oligothiophenes - Yet Longer? Synthesis, Characterization, and Scanning Tunneling Microscopy Images of Homologous, Isomerically Pure Oligo(Alkylthiophene)s / P. Bäuerle, T. Fischer, B. Bidlingmeier, J. P. Rabe, A. Stabel // *Angewandte Chemie International Edition*. — 1995. — T. 34, № 3. — C. 303—307.
616. *Mena-Osteritz, E.* Superstructures of Self-Organizing Thiophenes / E. Mena-Osteritz // *Advanced Materials*. — 2002. — T. 14, № 8. — C. 609—616.
617. *Hentschel, J.* Peptide-Guided Organization of Peptide-Polymer Conjugates: Expanding the Approach from Oligo- to Polymers / J. Hentschel, M. ten Cate, H. Borner // *Macromolecules*. — 2007. — T. 40, № 26. — C. 9224—9232.
618. Pseudo-Prolines as a Solubilizing, Structure-Disrupting Protection Technique in Peptide Synthesis / T. Wöhr, F. Wahl, A. Nefzi, B. Rohwedder, T. Sato, X. Sun, M. Mutter // *Journal of the American Chemical Society*. — 1996. — T. 118, № 39. — C. 9218—9227.
619. Novel and Efficient Synthesis of Difficult Sequence-Containing Peptides through O-N Intramolecular Acyl Migration Reaction of O-Acyl Isopeptides / Y. Sohma, M. Sasaki, Y. Hayashi, T. Kimura, Y. Kiso // *Chemical Communications*. — 2004. — № 1. — C. 124—125.
620. Switch Peptides In Statu Nascendi: Induction of Conformational Transitions Relevant to Degenerative Diseases / M. Mutter, A. Chandravarkar, C. Boyat, J. Lopez, S. Dos Santos, B. Mandal, R. Mimna, K. Murat, L. Patiny,

- L. Saucède, G. Tuchscherer // *Angewandte Chemie International Edition*. — 2004. — T. 43, № 32. — C. 4172—4178.
621. Synthesis of Difficult Peptide Sequences: Application of a Depsipeptide Technique to the Jung-Redemann 10- and 26-Mers and the Amyloid Peptide A[ $\beta$ ](1-42) / L. A. Carpino, E. Krause, C. D. Sferdean, M. Schümann, H. Fabian, M. Bienert, M. Beyermann // *Tetrahedron Letters*. — 2004. — T. 45, № 40. — C. 7519—7523.
622. *Schopf, G.* Polythiophenes - Electrically Conductive Polymers / G. Schopf, G. Kobmehl. — Berlin : Springer, 1997. — (Advances in Polymer Science ; 129).
623. Light-Emitting Diodes Based on Polythiophene: Influence of the Metal Workfunction on Rectification Properties / F. Garten, J. Vrijmoeth, A. Schlatmann, R. Gill, T. Klapwijk, G. Hadziioannou // *Synthetic Metals*. — 1996. — T. 76, № 1—3. — C. 85—89.
624. *Choi, M.* Polymers for Flexible Displays: From Material Selection to Device Applications / M. Choi, Y. Kim, C. Ha // *Progress in Polymer Science*. — 2008. — T. 33, № 6. — C. 581—630.
625. *Cai, W.* Polymer Solar Cells: Recent Development and Possible Routes for Improvement in the Performance / W. Cai, X. Gong, Y. Cao // *Solar Energy Materials and Solar Cells*. — 2009.
626. Derivation of Class II Force Fields. I. Methodology and Quantum Force Field for the Alkyl Functional Group and Alkane Molecules / J. R. Maple, M. J. Hwang, T. P. Stockfish, U. Dinur, M. Waldman, C. S. Ewig, A. T. Hagler // *Journal of Computational Chemistry*. — 1994. — T. 15, № 2. — C. 162—182.
627. Derivation of Class II Force Fields. 4. van Der Waals Parameters of Alkali Metal Cations and Halide Anions / Z. Peng, C. Ewig, M.-J. Hwang, M. Waldman, A. Hagler // *Journal of Physical Chemistry A*. — 1997. — T. 101, № 39. — C. 7243—7252.
628. *Ecroyd, H.* Crystallin Proteins and Amyloid Fibrils / H. Ecroyd, J. Carver // *Cellular and Molecular Life Sciences*. — 2009. — T. 66, № 1. — C. 62—81.

629. Molecular Level Control over Hierarchical Structure Formation and Polymerization of Oligopeptide-Polymer Conjugates / E. Jahnke, N. Severin, P. Kreutzkamp, J. P. Rabe, H. Frauenrath // *Advanced Materials*. — 2008. — T. 20, № 3. — C. 409—414.
630. *Margittai, M.* Fibrils with Parallel In-Register Structure Constitute a Major Class of Amyloid Fibrils: Molecular Insights from Electron Paramagnetic Resonance Spectroscopy / M. Margittai, R. Langen // *Quarterly Reviews of Biophysics*. — 2008. — T. 41, № 3/4. — C. 265—297.
631. *Nelson, D.* *Lehninger Principles of Biochemistry*. T. 5 / D. Nelson, M. Cox. — 01.01.2000. — DOI: [10.1007/978-3-662-08289-8](https://doi.org/10.1007/978-3-662-08289-8).
632. *Plimpton, S.* Fast Parallel Algorithms for Short-Range Molecular Dynamics / S. Plimpton // *Journal of Computational Physics*. — 1995. — T. 117, № 1. — C. 1—19.
633. *Hockney, R. W.* *Computer Simulation Using Particles* / R. W. Hockney, J. W. Eastwood. — Taylor & Francis, 1989. — ISBN 0-85274-392-0.
634. *Lehninger, A.* *Lehninger Principles of Biochemistry* / A. Lehninger, D. Nelson, M. Cox. — W. H. Freeman, 2008.
635. *Frishman, D.* Knowledge-Based Protein Secondary Structure Assignment / D. Frishman, P. Argos // *Proteins: Structure, Function, and Genetics*. — 1995. — T. 23, № 4. — C. 566—579.
636. *Rodriguez-Ropero, F.* Ab Initio Calculations on Pi-Stacked Thiophene Dimer, Trimer, and Tetramer: Structure, Interaction Energy, Cooperative Effects, and Intermolecular Electronic Parameters / F. Rodriguez-Ropero, J. Casanovas, C. Aleman // *Journal of Computational Chemistry*. — 2008. — T. 29, № 1. — C. 69—78.
637. *Tsuzuki, S.* Model Chemistry Calculations of Thiophene Dimer Interactions: Origin of PI-Stacking / S. Tsuzuki, K. Honda, R. Azumi // *Journal of the American Chemical Society*. — 2002. — T. 124, № 41. — C. 12200—12209.
638. *Stone, J.* *An Efficient Library for Parallel Ray Tracing and Animation* / Stone J. — University of Missouri-Rolla, 1998.
639. *Xu, Z.* Alzheimer's Abeta(1-40) Amyloid Fibrils Feature Size-Dependent Mechanical Properties / Z. Xu, R. Paparcone, M. Buehler // *Biophysical journal*. — 2010. — T. 98, № 10. — C. 2053—2062.

640. *Paparcone, R.* Atomistic Simulation of Nanomechanical Properties of Alzheimer's Abeta(1-40) Amyloid Fibrils under Compressive and Tensile Loading / R. Paparcone, S. Keten, M. Buehler // *Journal of biomechanics*. — 2010. — T. 43, № 6. — C. 1196—1201.
641. *Paparcone, R.* Microscale Structural Model of Alzheimer Abeta(1-40) Amyloid Fibril / R. Paparcone, M. J. Buehler // *Applied Physics Letters*. — 2009. — T. 94. — C. 243904.
642. WSXM: A Software for Scanning Probe Microscopy and a Tool for Nanotechnology / I. Horcas, R. Fernandez, G. Rodriguez, J. Colchero, G. Herrero, A. M. Baro // *Review of Scientific Instruments*. — 2007. — T. 78, № 1. — C. 013705.
643. *Deem, M.* Parallel Tempering: Theory, Applications, and New Perspectives / M. Deem // *Phys. Chem. Chem. Phys.* — 2005. — T. 7, № 23. — C. 3910—3916.
644. *Gale, J.* The General Utility Lattice Program "GULP" / J. Gale, A. Rohl // *Molecular Simulation*. — 2003. — T. 29, № 5. — C. 291—341.
645. *Amyloid Proteins: Methods and Protocols*. Т. 299 / под ред. E. M. Sigurdsson. — Humana Press, 2005. — (Methods in Molecular Biology).
646. Molecular Structural Basis for Polymorphism in Alzheimer's Beta-Amyloid Fibrils / A. Paravastu, R. Leapman, W.-M. Yau, R. Tycko // *Proceedings of the National Academy of Sciences*. — 2008. — T. 105, № 47. — C. 18349—18354.
647. Organization of Self-Assembled Peptide-Polymer Nanofibers in Solution / H. Börner, B. Smarsly, J. Hentschel, A. Rank, R. Schubert, Y. Geng, D. Discher, T. Hellweg, A. Brandt // *Macromolecules*. — 2008. — T. 41, № 4. — C. 1430—1437.
648. *Grimme, S.* Do Special Noncovalent Pi-Pi Stacking Interactions Really Exist? / S. Grimme // *Angewandte Chemie International Edition*. — 2008. — T. 47, № 18. — C. 3430—3434.
649. *He, M.* Mathematics of Bioinformatics: Theory, Methods and Applications / M. He, S. Petoukhov. — Hoboken, New Jersey : John Wiley & Sons, 2011.

650. *Chothia, C.* Conformation of Twisted B-Pleated Sheets in Proteins / C. Chothia // *Journal of Molecular Biology*. — 1973. — Т. 75, № 2. — С. 295—302.
651. *Salemme, F. R.* Structural Properties of Protein Beta-Sheets / F. R. Salemme // *Progress in Biophysics and Molecular Biology*. — 1983. — Т. 42. — С. 95—133.
652. *Chou, K.* Role of Interchain Interactions in the Stabilization of the Right-Handed Twist of Beta-Sheets / K. Chou, G. Nemethy, H. Scheraga // *Journal of Molecular Biology*. — 1983. — Т. 168, № 2. — С. 389—407.
653. *Yang, A.* Free Energy Determinants of Secondary Structure Formation: II. Antiparallel Beta-Sheets / A. Yang // *Journal of Molecular Biology*. — 1995. — Т. 252, № 3. — С. 366—376.
654. *Wang, L.* Molecular Simulations of Beta-Sheet Twisting / L. Wang // *Journal of Molecular Biology*. — 1996. — Т. 262, № 2. — С. 283—293.
655. *Koh, E.* Minimal Surface as a Model of Beta-Sheets / E. Koh, T. Kim // *Proteins*. — 2005. — Т. 61, № 3. — С. 559—569.
656. *Chou, K. C.* Origin of the Right-Handed Twist of Beta-Sheets of Poly(LVal) Chains / K. C. Chou, H. A. Scheraga // *Proceedings of the National Academy of Sciences*. — 1982. — Т. 79, № 22. — С. 7047—7051.
657. Prediction of Twist of Amyloid Fibrils Using Molecular Dynamics / ed. by J. Berryman, S. Radford, S. Harris. — Jülich : NIC, 2008. — 430 p. — (NIC Series ; 40). — ISBN 978-3-9810843-6-8.
658. *Твердислов, В.А.* Хиральный Дуализм Как Системный Фактор Иерархического Структурообразования в Молекулярно-Биологической Эволюции / Твердислов, В.А., Малышко, Е.В. // *Сложность. Разум. Постнеклассика*. — 2016. — Т. 1, № 1. — С. 78—83.
659. Spatial Persistence of Angular Correlations in Amyloid Fibrils / T. Knowles, J. Smith, A. Craig, C. Dobson, M. Welland // *Physical Review Letters*. — 2006. — Т. 96, № 23. — С. 238301.
660. *Mátrai, J.* Recent Advances in Lentiviral Vector Development and Applications / J. Mátrai, M. K. L. Chuah, T. VandenDriessche // *Molecular Therapy: The Journal of the American Society of Gene Therapy*. — 2010. — Mar. — Vol. 18, no. 3. — P. 477—490. — DOI: [10.1038/mt.2009.319](https://doi.org/10.1038/mt.2009.319).

661. Semen-Derived Amyloid Fibrils Drastically Enhance HIV Infection / J. Münch, E. Rücker, L. Ständker, K. Adermann, C. Goffinet, M. Schindler, S. Wildum, R. Chinnadurai, D. Rajan, A. Specht, G. Giménez-Gallego, P. Sánchez, D. Fowler, A. Koulov, J. Kelly, W. Mothes, J.-C. Grivel, L. Margolis, O. Keppler, W.-G. Forssmann, F. Kirchhoff // *Cell*. — 2007. — Т. 131, № 6. — С. 1059—1071.
662. The Cationic Properties of SEVI Underlie Its Ability To Enhance Human Immunodeficiency Virus Infection / N. R. Roan, J. Münch, N. Arhel, W. Mothes, J. Neidleman, A. Kobayashi, K. Smith-McCune, F. Kirchhoff, W. C. Greene // *Journal of Virology*. — 2009. — Jan. 1. — Vol. 83, no. 1. — P. 73—80. — DOI: [10.1128/JVI.01366-08](https://doi.org/10.1128/JVI.01366-08).
663. Semen-Mediated Enhancement of HIV Infection Is Donor-Dependent and Correlates with the Levels of SEVI / K.-A. Kim, M. Yolamanova, O. Zirafi, N. R. Roan, L. Staendker, W.-G. Forssmann, A. Burgener, N. Dejucq-Rainsford, B. H. Hahn, G. M. Shaw, W. C. Greene, F. Kirchhoff, J. Münch // *Retrovirology*. — 2010. — 23 июня. — Т. 7, № 1. — С. 55. — DOI: [10.1186/1742-4690-7-55](https://doi.org/10.1186/1742-4690-7-55).
664. The Influence of Semen-Derived Enhancer of Virus Infection on the Efficiency of Retroviral Gene Transfer / M. Wurm, A. Schambach, D. Lindemann, H. Hanenberg, L. Ständker, W.-G. Forssmann, R. Blasczyk, P. A. Horn // *The Journal of Gene Medicine*. — 2010. — Feb. — Vol. 12, no. 2. — P. 137—146. — DOI: [10.1002/jgm.1429](https://doi.org/10.1002/jgm.1429).
665. Improved Lentiviral Gene Transfer into Human Embryonic Stem Cells Grown in Co-Culture with Murine Feeder and Stroma Cells / M. Wurm, B. Gross, M. Sgodda, L. Ständker, T. Müller, W.-G. Forssmann, P. A. Horn, R. Blasczyk, T. Cantz // *Biological Chemistry*. — 2011. — Oct. — Vol. 392, no. 10. — P. 887—895. — DOI: [10.1515/BC.2011.085](https://doi.org/10.1515/BC.2011.085).
666. Peptides Released by Physiological Cleavage of Semen Coagulum Proteins Form Amyloids That Enhance HIV Infection / N. R. Roan, J. A. Müller, H. Liu, S. Chu, F. Arnold, C. M. Stürzel, P. Walther, M. Dong, H. E. Witkowska, F. Kirchhoff, J. Münch, W. C. Greene // *Cell Host & Microbe*. — 2011. — Dec. 15. — Vol. 10, no. 6. — P. 541—550. — DOI: [10.1016/j.chom.2011.10.010](https://doi.org/10.1016/j.chom.2011.10.010).

667. Naturally Occurring Fragments from Two Distinct Regions of the Prostatic Acid Phosphatase Form Amyloidogenic Enhancers of HIV Infection / F. Arnold, J. Schnell, O. Zirafi, C. Stürzel, C. Meier, T. Weil, L. Ständker, W.-G. Forssmann, N. R. Roan, W. C. Greene, F. Kirchhoff, J. Münch // *Journal of Virology*. — 2012. — ЯНВ. — Т. 86, № 2. — С. 1244—1249. — DOI: [10.1128/JVI.06121-11](https://doi.org/10.1128/JVI.06121-11).
668. *Knowles, T. P. J.* Nanomechanics of Functional and Pathological Amyloid Materials / T. P. J. Knowles, M. J. Buehler // *Nature Nanotechnology*. — 2011. — July 31. — Vol. 6, no. 8. — P. 469—479. — DOI: [10.1038/nnano.2011.102](https://doi.org/10.1038/nnano.2011.102).
669. *Dalhaimer, P.* Single Molecule Visualization of Stable, Stiffness-Tunable, Flow-Conforming Worm Micelles / P. Dalhaimer, F. S. Bates, D. E. Discher // *Macromolecules*. — 2003. — 1 сент. — Т. 36, № 18. — С. 6873—6877. — DOI: [10.1021/ma034120d](https://doi.org/10.1021/ma034120d).
670. Antiparallel Beta-Sheet: A Signature Structure of the Oligomeric Amyloid Beta-Peptide / E. Cerf, R. Sarroukh, S. Tamamizu-Kato, L. Breydo, S. Derclaye, Y. F. Dufrene, V. Narayanaswami, E. Goormaghtigh, J.-M. Ruyschaert, V. Raussens // *The Biochemical Journal*. — 2009. — July 15. — Vol. 421, no. 3. — P. 415—423. — DOI: [10.1042/BJ20090379](https://doi.org/10.1042/BJ20090379).
671. *Tiscornia, G.* Production and Purification of Lentiviral Vectors / G. Tiscornia, O. Singer, I. M. Verma // *Nature Protocols*. — 2006. — Vol. 1, no. 1. — P. 241—245. — DOI: [10.1038/nprot.2006.37](https://doi.org/10.1038/nprot.2006.37).
672. Colocalization of Retrovirus and Target Cells on Specific Fibronectin Fragments Increases Genetic Transduction of Mammalian Cells / H. Hanenberg, X. L. Xiao, D. Dilloo, K. Hashino, I. Kato, D. A. Williams // *Nature Medicine*. — 1996. — Aug. — Vol. 2, no. 8. — P. 876—882. — DOI: [10.1038/nm0896-876](https://doi.org/10.1038/nm0896-876).
673. Optimization of Fibronectin-Assisted Retroviral Gene Transfer into Human CD34+ Hematopoietic Cells / H. Hanenberg, K. Hashino, H. Konishi, R. A. Hock, I. Kato, D. A. Williams // *Human Gene Therapy*. — 1997. — Dec. 10. — Vol. 8, no. 18. — P. 2193—2206. — DOI: [10.1089/hum.1997.8.18-2193](https://doi.org/10.1089/hum.1997.8.18-2193).



674. High-Efficiency Gene Transfer into Normal and Adenosine Deaminase-Deficient T Lymphocytes Is Mediated by Transduction on Recombinant Fibronectin Fragments / K. E. Pollok, H. Hanenberg, T. W. Noblitt, W. L. Schroeder, I. Kato, D. Emanuel, D. A. Williams // *Journal of Virology*. — 1998. — June. — Vol. 72, no. 6. — P. 4882–4892. — DOI: [10.1128/JVI.72.6.4882-4892.1998](https://doi.org/10.1128/JVI.72.6.4882-4892.1998).
675. 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*. — 2010. — June. — Vol. 78, no. 8. — P. 1950–1958. — DOI: [10.1002/prot.22711](https://doi.org/10.1002/prot.22711).
676. GROMACS: Fast, Flexible, and Free / D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. C. Berendsen // *Journal of Computational Chemistry*. — 2005. — Dec. — Vol. 26, no. 16. — P. 1701–1718. — DOI: [10.1002/jcc.20291](https://doi.org/10.1002/jcc.20291).
677. Free Energy Profiles of Amino Acid Side Chain Analogs near Water-Vapor Interface Obtained via MD Simulations / A. K. Shaytan, V. A. Ivanov, K. V. Shaitan, A. R. Khokhlov // *Journal of Computational Chemistry*. — 2010. — Jan. 15. — Vol. 31, no. 1. — P. 204–216. — DOI: [10.1002/jcc.21267](https://doi.org/10.1002/jcc.21267). — IF WoS 3.28.

## Список рисунков

1.1	Иллюстрации к истории структурной биологии и моделирования. . .	28
1.2	Диаграмма, суммирующая различные подходы к описанию, моделированию и изучению структуры и динамики биомакромолекул.	35
1.3	Идея методов интегративного моделирования. . . . .	37
1.4	Энергетические термы силового поля . . . . .	40
1.5	Графическое пояснение к членам внутримолекулярной энергии уравнения 1.4. . . . .	43
1.6	Иллюстрация периодических граничных условий . . . . .	45
1.7	Схематическое изображение разложения заряда, используемого в суммировании по Эвальду. . . . .	53
1.8	Пояснения к определению поверхности доступной растворителю и поверхности Коннолли. . . . .	55
1.9	Описание геометрии ДНК в виде геометрических параметров . . . .	60
1.10	Матрица средних значений и отклонений динуклеотидных параметров из статьи [105]. . . . .	63
1.11	Алгоритм интегративного моделирования на основе моделей ДНК в динуклеотидном приближении. . . . .	66
1.12	Подходы к моделированию супрануклеосомной структуры . . . . .	70
1.13	Визуализация анализа данные Micro-C, Ohno et al., 2019 . . . . .	72
1.14	Алгоритмы интегративного моделирования нуклеосомных фибрилл .	73
1.15	Зависимость эффективности переноса энергии от расстояния между флуорофорами. . . . .	76
1.16	Схема экспериментов spFRET . . . . .	79
1.17	Методы “захвата конформации хроматина”. . . . .	84
1.18	Методы эпигеномики нуклеосомного разрешения . . . . .	87
2.1	Структура и вариабельность нуклеосомы . . . . .	92
2.2	Моды динамики нуклеосом . . . . .	95
2.3	Тонкие детали динамики нуклеосом . . . . .	101
2.4	Обзор структуры и динамики модели полной нуклеосомы (FN) . . .	106
2.5	Максимальные наблюдаемые среднеквадратичные отклонения отдельных аминокислот гистонов в МД моделировании . . . . .	108
2.6	Конформационный ансамбль ДНК в нуклеосоме . . . . .	112
2.7	Конденсация гистоновых хвостов на ДНК . . . . .	116

2.8	Доступность ДНК для взаимодействия с гистоновыми хвостами . . .	116
2.9	Типичные паттерны взаимодействия гистоновых хвостов с ДНК . . .	121
2.10	Взаимодействия белок-ДНК в нуклеосоме . . . . .	122
2.11	Взаимодействия белок-ДНК в области нуклеосомного кора . . . . .	123
2.12	Флуктуации ДНК в нуклеосоме . . . . .	126
2.13	Системы нуклеосом моделируемые на временах более 10 мкс . . . . .	141
2.14	Динамика отворота ДНК в нуклеосоме с укороченными гистоновыми хвостами . . . . .	145
2.15	Сравнение характеристик отворота ДНК в нуклеосомах с и без гистоновых хвостов . . . . .	146
2.16	Динамика “защелки” концов ДНК фрагментом гистона H3 . . . . .	146
2.17	Влияние динамики ДНК на структуру хроматиновых фибрилл . . .	148
2.18	Динамика кручения ДНК в нуклеосоме . . . . .	150
2.19	Стабильные контакты ДНК-гистоны в нуклеосоме . . . . .	152
2.20	Пластичность гистонового октамера . . . . .	154
2.21	Аллостерическая связь откручивания ДНК и конформации ДНК около диады . . . . .	156
3.1	Футпринтинг комплексов ДНК-белок гидроксильными радикалами и структурная интерпретация экспериментальных данных. . . . .	162
3.2	Конвейеры данных в HYDROID . . . . .	169
3.3	Количественная оценка изображения геля ПААГЭ путем подбора математической модели. . . . .	183
3.4	Профили частоты расщепления ДНК, извлеченные из экспериментов по гидроксильному футпринтингу. . . . .	193
3.5	Профили H-SASA для ДНК в нуклеосомах . . . . .	194
3.6	Экспериментальные и теоретические профили частоты расщепления ДНК гидроксильными радикалами в нуклеосомах . . .	195
3.7	Структура нуклеосом и псевдосимметрия верхней и нижней цепей ДНК. . . . .	201
3.8	Детали гидроксил-радикальных взаимодействий с нуклеосомной ДНК. . . . .	202
3.9	Структура ДНК и профили H-SASA в рентгеновской структуре самого высокого разрешения NCP (NCP147) . . . . .	212

3.10	Различия в структуре ДНК между разными рентгеновскими структурами нуклеосомы. . . . .	213
3.11	Количественная оценка экспериментов по гидроксильному футпринтингу и сравнение с теоретическими профилями, полученными на основе атомной структуры (для нуклеосомы 601TA-NUC) . . . . .	219
3.12	Профили частоты расщепления ДНК для нуклеосом SEN3-NUC и ДНК свободной SEN3 ДНК . . . . .	220
3.13	Объяснение алгоритма идентификации диады применительно к нуклеосоме 601TA-NUC с известным положением диады . . . . .	222
3.14	Применение алгоритма идентификации диады к нуклеосоме SEN3-NUC с неизвестным положением диады. . . . .	223
4.1	От клеток к нуклеосомам. . . . .	231
4.2	Хромосомные территории. . . . .	232
4.3	Структура хроматина и проблема ее понимания на супрануклеосомном уровне. . . . .	233
4.4	О проблеме понимания организации и функционирования хроматина на супрануклеосомном уровне . . . . .	235
4.5	Различные структуры нуклеосом, их комплексов и фибрилл из нуклеосом, полученные в последнее время методами структурной биологии. . . . .	239
4.6	Модель ЭК+49 (нулевая петля) с дрожжевой Pol II . . . . .	244
4.7	Отрицательно заряженная поверхность Pol II может стабилизировать интермедиат нулевой петли посредством взаимодействия с гистоновым октамером . . . . .	245
4.8	Предлагаемые механизмы транскрипции через нуклеосому разными РНКП. . . . .	249
4.9	Структура комплекса элонгации остановленного в положении +42 нуклеосомы. . . . .	254
4.10	Расположение флуоресцентных меток на нуклеосоме в spFRET экспериментах . . . . .	258
4.11	Анализ зависимости эффективности FRET от степени разворачивания нуклеосом . . . . .	259

4.12	Модели разворачивания нуклеосом, вызванного взаимодействием с шапероном FАСТ . . . . .	260
4.13	Анализ комплексов димеров Mif2 с центромерными нуклеосомами Cse4/CEN3 методами гидроксильного футпринтинга . . . . .	264
4.14	Структура хроматосомы и оптимальные последовательности линкерной ДНК. . . . .	273
4.15	Создание моделей ДНК в хроматосоме по данным FRET. . . . .	280
5.1	Модели расположения $\beta$ -тяжей в амилоидных протофибриллах . . .	286
5.2	Возможные варианты расположения $\beta$ -листов в перекрестном $\beta$ -мотиве . . . . .	288
5.3	Структурные типы полиморфизма амилоидных фибрилл . . . . .	289
5.4	Модель иерархической самосборки хиральных стержневых единиц .	294
5.5	Примеры различных морфологической организации фибриллярных агрегатов, предположительно образуемых полимерными биоконъюгатами . . . . .	297
5.6	Различные конъюгаты тиофенов и пептидов, образующих $\beta$ -лист . .	300
5.7	Синтезированное соединение Т-Р в “защищенной” форме . . . . .	302
5.8	АСМ изображения высоты и фазы нативного ПЭГ-пептида-кватротиофена . . . . .	303
5.9	Исследуемое соединение тиофен-пептид . . . . .	306
5.10	Различные представления молекулы Т-Р, используемые для компьютерного моделирования самособирающихся фибриллярных агрегатов . . . . .	307
5.11	Построенные однослойные периодические ячейки укладки Т-Р соединений . . . . .	310
5.12	Построенные двухслойные периодические кристаллические ячейки укладки Т-Р соединений . . . . .	311
5.13	Вид двух двухслойных структур, изображающих интеркаляцию боковых цепей аминокислот, образующих стерическую застежку-молнию . . . . .	312
5.14	Первоначальный снимок сольватированной короткой однослойной фибриллы на основе антипараллельного расположения $\beta$ -листов . .	315
5.15	Кадр из моделирования коротких фибрилл в растворителе после 10 нс моделирования . . . . .	317

5.16	Среднеквадратичное отклонение структуры фибрилл относительно исходной структуры в зависимости от времени моделирования . . . . .	319
5.17	Число водородных связей на одну гибридную молекулу как функция времени моделирования . . . . .	322
5.18	Гистограммы элементов вторичной структуры в различных фибриллах, смоделированных в растворителе . . . . .	324
5.19	Линейное распределение пептидных и тиофеновых блоков в фибриллах . . . . .	326
5.20	Снимки расположения блоков кватротиофена в моделируемых фибриллах различных типов: а) фибрилла SL-AP, б) фибрилла SL-PAR, в) фибрилла DL-AP, г) фибрилла DL-PAR. Алкильные цепи не изображены для ясности представления. . . . .	328
5.21	Анализ упорядочения тиофенов в фибриллах . . . . .	331
5.22	Первоначальные снимки и эволюция фибрилл во время моделирования однослойных структур тиофен-пептид . . . . .	338
5.23	Первоначальные снимки и эволюция фибрилл во время моделирования двухслойных структур тиофен-пептид . . . . .	339
5.24	Снимки окончательной конформации чистых пептидных фибрилл (без кватротиофеновых фрагментов) . . . . .	340
5.25	Окончательные снимки и эволюция крупномасштабных фибрилл и их размеры . . . . .	341
5.26	Зависимость среднего косинуса между касательными от длины контура фибриллы для типа фибриллы DL-PAR . . . . .	342
5.27	Иллюстрация протокола моделирования МД для моделирования длинной фибриллы SL-AP . . . . .	344
5.28	Снимки моделирования адсорбции фибриллы DL-PAR на поверхности графита, представляющие эволюцию процесса адсорбции	347
5.29	Снимки моделирования адсорбции фибриллы DL-AP на поверхности графита, представляющие эволюцию процесса адсорбции	348
5.30	Снимки конечного адсорбированного состояния фибриллы DL-PAR на графите . . . . .	349
5.31	Снимки конечного адсорбированного состояния фибриллы DL-PAR на графите . . . . .	350
5.32	Высотные профили поперечного сечения фибрилл, адсорбированных на графите . . . . .	351

5.33	Схематическое изображение комбинированной экспериментальной/вычислительной интегративной методологии для изучения амилоидоподобных фибрилл . . . . .	359
5.34	Предложенная схематическая модель вероятного расположения гибридных молекул в наблюдаемых нанофибриллах . . . . .	371
5.35	Пептиды, полученные из гликопротеина gp120 ВИЧ-1, усиливают вирусную инфекцию . . . . .	378
5.36	Структурная характеристика и молекулярное моделирование фибрилл EF-C . . . . .	379
5.37	Дополнительная структурная характеристика фибрилл EF-C . . . . .	380
5.38	Дополнительные данные о модели фибриллы из пептидов EF-C . . . . .	384

**Список таблиц**

1	Увеличение разрешения Hi-C подобных методов, благодаря использованию менее специфичных рестриктаз. . . . .	88
2	Исследованные в моделировании системы и их параметры . . . . .	107
3	Список моделируемых систем . . . . .	158
4	Основные функции HYDROID и их описание. . . . .	174
5	Список ограничений при подгонке моделей, доступные в HYDROID. . . . .	175
6	Корреляция между присутствием отрицательно заряженных областей на обращенной к октамеру поверхности РНКП, высотой нуклеосомного барьера и выживаемостью нуклеосом во время транскрипции . . . . .	246
7	Среднее количество водородных связей на гибридную молекулу, усредненное за последние 2 нс моделирования . . . . .	321
8	Параметры порядка соседних блоков кватротиофена в фибриллах, рассчитанные в течение последних 2 нс моделирования. . . . .	329
9	Экспериментальные методы, используемые для изучения амилоидных фибрилл, а также ценные данные, которые могут быть получены и использованы в обсуждаемой методологии. Для дальнейшего обсуждения методов см. также [645]. . . . .	355